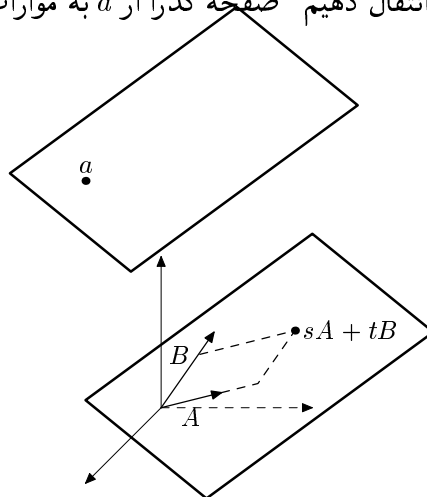


برای دریافت نمونه سؤالات، جزوات و کتب بیشتر به ما پیوندید در

https://t.me/iust_scientific

زیرفضاهای مستوی (۱)

در جلسه قبل مفهوم نقطه و خط راست را که ابتدایی‌ترین مفاهیم هندسه‌اند در \mathbb{R}^n در نظر گرفتیم. گام بعدی ما تعریف یک "صفحه راست" (یا به طور خلاصه "صفحه") در \mathbb{R}^n است ($n \geq 2$). یادآوری می‌کنیم که "خط گذرا از a به موازات A "، $A \neq \mathbf{0}$ ، این گونه تعریف شد که نخست مضارب حقیقی A را در نظر گرفتیم، که یک "خط گذرا از $\mathbf{0}$ در راستای A " تشکیل می‌دهند، و سپس با انتقال این خط با مقدار ثابت a ، "خط گذرا از a به موازات A " به دست آمد. برای تعریف صفحه به طریق مشابه عمل می‌کنیم. فرض کنید A, B دو تایی ناهمراستا باشند یعنی هیچ‌یک مضرب حقیقی دیگری نباشد. نخست مجموعه نقاط به شکل $sA + tB$ را در نظر می‌گیریم. تصور شهودی ما از این مجموعه، صفحه گذرا از مبدأ و شامل نقاط A و B است. توجه کنید که به ازای $s = t = 0$ ، مبدأ به دست می‌آید، به ازای $s = 1, t = 0$ ، نقطه A و به ازای $s = 0, t = 1$ ، نقطه B . حال اگر نقطه‌ای a در نظر بگیریم و همه نقاط این مجموعه را به اندازه a انتقال دهیم "صفحه گذرا از a به موازات A و B " حاصل می‌شود.



بدین ترتیب تعریف می‌کنیم

$$\langle a; A, B \rangle = \{a + sA + tB \mid s, t \in \mathbb{R}\} \quad (1)$$

این مجموعه را صفحه گذرا از a به موازات A و B می‌نامیم.

مثال ۱ (صفحات مختصاتی). در $\{1, \dots, n\}$ به تعداد $\binom{n}{2} = \frac{n(n-1)}{2}$ زوج اندیس $\{i, j\}$ ، $i \neq j$ وجود دارد. به ازای هر مجموعه $\{i, j\}$

$$\{se_i + te_j \mid s, t \in \mathbb{R}\}$$

یک صفحه مختصاتی در \mathbb{R}^n است.

اگر E یک صفحه در \mathbb{R}^n باشد، $E = \langle a; A, B \rangle$ ، E° یا انتقال یافته E به مبدأ را به صورت

$$E^\circ = \langle \underline{0}; A, B \rangle$$

تعریف می‌کنیم. گاهی به جای $\langle \underline{0}; A, B \rangle$ می‌نویسیم $\langle A, B \rangle$. اگر E_1 و E_2 هر یک خط یا صفحه باشد، E_1 و E_2 را موازی می‌نامیم و می‌نویسیم $E_1 \parallel E_2$ مشروط بر این که اشتراک E_1 و E_2 تهی باشد و به علاوه:

$$E_2 \subset E_1 \quad \text{یا} \quad E_1 \subset E_2$$

مثال ۲. اگر E یک صفحه در \mathbb{R}^n باشد و $\underline{0} \notin E$ نشان می‌دهیم $E \parallel E^\circ$. باید نشان دهیم E و E° نقطه مشترک ندارند. فرض کنید $E = \langle a; A, B \rangle$ ، پس $E^\circ = \langle \underline{0}; A, B \rangle$. اگر نقطه مشترکی وجود داشته باشد، مثلاً P ، آنگاه $P = a + s_1A + t_1B$ برای s_1 و t_1 مناسب، و نیز $P = s_2A + t_2B$ برای s_2 و t_2 مناسب، پس $a + s_1A + t_1B = s_2A + t_2B$ یا

$$a = (s_2 - s_1)A + (t_2 - t_1)B \quad (2)$$

در نتیجه می توان نوشت

$$\begin{aligned}\underline{o} &= (s_2 - s_1)A + (t_2 - t_1)B + (s_1 - s_2)A + (t_1 - t_2)B \\ &= a + (s_1 - s_2)A + (t_1 - t_2)B\end{aligned}$$

یعنی $\underline{o} \in E$ که خلاف فرض است.

مثال ۳. اگر $E = \langle a; A, B \rangle$ یک صفحه باشد، $b \notin E$ و $L = \langle b; B \rangle$ ، آنگاه خط L با صفحه E موازی است زیرا که اولاً $E^\circ = \langle \underline{o}; A, B \rangle \subset \langle \underline{o}; B \rangle = L^\circ$ ، و ثانیاً نشان می دهیم L و E نقطه مشترک ندارند. فرض کنید P یک نقطه مشترک باشد، پس $P = b + tB$ و نیز $P = a + s_1A + t_1B$ ،

بنابراین

$$b + tB = a + s_1A + t_1B$$

$$b = a + s_1A + (t_1 - t)B$$

یعنی $b \in E$ که خلاف فرض است.

مثال ۴. در \mathbb{R}^3 ، اگر دو صفحه نقطه مشترک نداشته باشند، لزوماً موازی اند. این وضعیت در \mathbb{R}^n ، برای $n \geq 4$ ، لزوماً حاکم نیست، یعنی دو صفحه فاقد نقطه مشترک ممکن است موازی نباشند! مثال زیر را در \mathbb{R}^4 در نظر بگیرید:

$$E_1 = \langle \underline{o}; e_1, e_2 \rangle \quad , \quad E_2 = \langle e_4; e_1, e_3 \rangle$$

E_1 متشکل از نقاط به شکل $(s, t, 0, 0)$ است و E_2 متشکل از نقاط به صورت $(s', 0, t', 1)$. بدیهی است که E_1 و E_2 نقطه مشترک ندارند زیرا که مؤلفه چهارم هر عنصر E_1 صفر است و مؤلفه چهارم هر عنصر E_2 برابر ۱ می باشد. حال اگر E_1 و E_2 را به مبدأ منتقل کنیم، داریم

$$E_1^\circ = E_1 = \{(s, t, 0, 0) \mid s, t \in \mathbb{R}\} \quad , \quad E_2^\circ = \{(s', 0, t', 0) \mid s', t' \in \mathbb{R}\}$$

که نه E_1° زیرمجموعه E_2° است و نه بالعکس، پس E_1 و E_2 در تعریف توازی صدق نمی کنند. دو صفحه در \mathbb{R}^n ، $n \geq 4$ ، را که نه نقطه مشترک داشته باشند و نه موازی باشند، دو صفحه متنافر می نامیم.

در اینجا می‌توان گزاره‌هایی مشابه گزاره‌های (۱-۴) و (۱-۵) جلسه قبل ثابت کرد به این مضمون که:

(۱) اگر $E = \langle a; A, B \rangle$ ، C و D دو عنصر ناهمراستای E° باشند و $b \in E$ ، آنگاه $E = \langle b; C, D \rangle$.

(۲) اگر P, Q, R سه نقطه متمایز و ناهمراستای مشترک بین دو صفحه E_1 و E_2 باشند، آنگاه $E_1 = E_2$.

چون به زودی احکام کلی‌تری ثابت خواهیم کرد در حال حاضر اثبات این دو حکم را به دانشجویان واگذار می‌کنیم و به معرفی مفاهیم جدیدی فرای خط و صفحه می‌پردازیم. تصور ذهنی ما از خط و صفحه، اشکال مسطح "یک بعدی" و "دو بعدی" است. هنوز تعریفی دقیق از کلمه "بعد" ارائه نکرده‌ایم و لیکن می‌دانیم که نقاط خط با یک پارامتر t در مجموعه $\{a + tA \mid t \in \mathbb{R}\}$ مدرج می‌شوند و نقاط یک صفحه با دو پارامتر s, t در مجموعه $\{a + sA + tB \mid s, t \in \mathbb{R}\}$. به طور شهودی می‌توان تعداد پارامترهای حقیقی مستقل لازم برای تمیز دادن نقاط یک مجموعه را "بعد" آن مجموعه تلقی کرد. این مفهوم را به طور دقیق‌تر ضمن معرفی "زیرفضاهای مستوی k -بعدی در \mathbb{R}^n "، $k \leq n$ ، که تعمیم خط و صفحه به عنوان "اشیاء مسطح k -بعدی" خواهند بود، بررسی خواهیم کرد.

فرض کنید می‌خواهیم در \mathbb{R}^n ، $n \geq 3$ ، تعریفی برای اشیاء سه‌بعدی مسطح مشابه خط و صفحه دو بعدی تعریف کنیم. طبیعی است که نخست مجموع‌های به شکل $rA + sB + tC$ ، r, s, t اعداد حقیقی را در نظر بگیریم که در آن A, B و C سه تایی هستند. سپس اگر $a \in \mathbb{R}^n$ داده شده باشد، مجموعه انتقال یافته، یعنی $\{a + rA + sB + tC \mid r, s, t \in \mathbb{R}\}$ را "یک زیرفضای سه‌بعدی" در \mathbb{R}^n تلقی کنیم. نکته قابل توجه این که در مورد خط راست فرض کردیم $A \neq \underline{0}$ که در غیر این صورت مجموعه $\{a + tA \mid t \in \mathbb{R}\}$ به یک تک نقطه $\{a\}$ مبدل می‌شود. همین‌طور در مورد صفحه، فرض ناهمراستا بودن A و B را اعمال کردیم که در غیر این صورت مجموعه $\{a + sA + tB \mid s, t \in \mathbb{R}\}$ به یک خط راست یا حتی یک نقطه (وقتی $A = B = \underline{0}$) مبدل می‌شود. حال می‌خواهیم شرطی مناسب برای سه تایی (یا حتی به طور کلی k ، تایی) بیابیم که تعمیم طبیعی صفر نبودن A در $\{A\}$ و ناهمراستا بودن A و B در $\{A, B\}$ باشد.

(۲-۱) تعریف. فرض کنید A_1, \dots, A_k عناصر \mathbb{R}^n باشند. مجموعه $\{A_1, \dots, A_k\}$ را مستقل خطی می‌نامیم در صورتی که هیچ مجموع مضارب حقیقی A_1, \dots, A_k ، یعنی $t_1 A_1 + \dots + t_k A_k$ ، صفر نشود مگر این که همه ضرایب t_1, \dots, t_k صفر باشند. اگر مجموعه $\{A_1, \dots, A_k\}$ مستقل خطی نباشند، آن را وابسته خطی می‌نامیم. هر مجموع به شکل $t_1 A_1 + \dots + t_k A_k$ یک ترکیب خطی A_1, \dots, A_k خوانده می‌شود.

مثال ۱. برای $k = 1$ ، مجموعه $\{A\}$ وابسته خطی خواهد بود به شرطی که $tA = \underline{0}$ بتواند برقرار باشد بدون این که $t = 0$. این تنها وقتی میسر است که $A = \underline{0}$. پس شرط اعمال شده برای تعریف خط راست $\langle a; A \rangle$ روی A (یعنی $A \neq \underline{0}$) معادل این است که $\{A\}$ مستقل خطی است.

مثال ۲. برای $k = 2$ ، وابستگی خطی مجموعه دو عنصری $\{A, B\}$ معادل این است که $t_1 A + t_2 B = \underline{0}$ بدون این که هر دو t_1 و t_2 صفر شوند. مثلاً اگر $t_1 \neq 0$ ، آنگاه $A = -\frac{t_2}{t_1} B$ ، یعنی A مضربی حقیقی از B خواهد شد. بدین ترتیب در تعریف صفحه $\langle a; A, B \rangle$ شرط همراستا نبودن A و B را می‌توان بدین صورت بیان کرد که $\{A, B\}$ مستقل خطی است.

مثال ۳. اگر در مجموعه $\{A_1, \dots, A_k\}$ ، یک یا بیشتر از A_i ها n -تایی صفر، $\underline{0}$ ، باشد، مجموعه، وابسته خطی می‌شود. مثلاً فرض کنید $A_j = \underline{0}$. ضرایب t_i را این طور در نظر بگیرید:

$$t_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

در این صورت $t_1 A_1 + \dots + t_k A_k = \underline{0}$ بدون این که همه A_i ها صفر باشند.

(۲-۲) گزاره. شرطی لازم و کافی برای وابستگی خطی $\{A_1, \dots, A_k\}$ بدین شرح است:

(الف) در حالت $k = 1$ ، $\{A_1\}$ وابسته خطی است اگر و تنها اگر $A_1 = \underline{0}$.

(ب) در حالت $k > 1$ ، $\{A_1, \dots, A_k\}$ وابسته خطی است اگر و تنها اگر بتوان یکی از A_i ها را به صورت ترکیبی خطی از سایر A_i ها نوشت.

اثبات. حالت $k = 1$ در مثال 1 بالا بررسی شد، فرض کنید $k \geq 2$. اگر $\{A_1, \dots, A_k\}$ وابسته خطی باشد، t_1, \dots, t_k حقیقی وجود دارند، نه همه صفر، که $t_1 A_1 + \dots + t_k A_k = \underline{0}$ مثلاً $t_j \neq 0$. در این صورت

$$A_j = -\frac{t_1}{t_j} A_1 - \dots - \frac{t_{j-1}}{t_j} A_{j-1} - \frac{t_{j+1}}{t_j} A_{j+1} - \dots - \frac{t_n}{t_j} A_n$$

یعنی A_j ترکیبی خطی از سایر A_i ها است.

بالعکس اگر داشته باشیم:

$$A_j = c_1 A_1 + \dots + c_{j-1} A_{j-1} + c_{j+1} A_{j+1} + \dots + c_k A_k$$

با انتقال A_j به طرف دیگر رابطه داریم

$$c_1 A_1 + \dots + c_{j-1} A_{j-1} - A_j + c_{j+1} A_{j+1} + \dots + c_k A_k = \underline{0}$$

□ و این در حالی است که دست کم یکی از ضرایب (ضریب A_j) صفر نیست.

با این مقدمه، زیرفضاهای با بعد بالاتر از 1 و 2 در \mathbb{R}^n به صورت زیر تعریف می شود. فرض کنید $\{A_1, \dots, A_k\}$ یک مجموعه مستقل خطی از عناصر \mathbb{R}^n باشد و $a \in \mathbb{R}^n$ یک نقطه داده شده، در این صورت مجموعه زیر یک زیرفضای مستوی k -بعدی ("زیرفضای مستوی k -بعدی گذرا از a به موازات A_1, \dots, A_k ") خوانده می شود:

$$\langle a; A_1, \dots, A_k \rangle = \{a + t_1 A_1 + \dots + t_k A_k \mid t_1, \dots, t_k \in \mathbb{R}\} \quad (3)$$

اگر مجموعه تعریف شده در (3) را به E نمایش دهیم، مقصود از E° (انتقال یافته E به $\underline{0}$)، مجموعه $\langle \underline{0}; A_1, \dots, A_k \rangle = \langle \underline{0}, A_1, \dots, A_k \rangle$ است. به طور کلی یک زیرفضای مستوی k -بعدی که شامل $\underline{0}$ باشد یک زیرفضای خطی خوانده می شود. طبق قرارداد، تنها زیرفضای صفر بعدی تک عنصری $\{\underline{0}\}$ است، و بدین ترتیب زیرفضاهای مستوی صفر بعدی تک عنصری های $\{a\}$ هستند، $a \in \mathbb{R}^n$ ، که

از انتقال \underline{e} به a به دست می آیند. بدین ترتیب جدول زیر را داریم:

زیرفضای مستوی صفر بعدی \longleftrightarrow نقطه

زیرفضای مستوی یک بعدی \longleftrightarrow خط

زیرفضای مستوی دو بعدی \longleftrightarrow صفحه

:

تعریف زیرفضای k -بعدی جای تأمل دارد. سؤال‌های متعددی می‌توان این مقطع مطرح کرد. آیا عدد k به طور منحصر به فرد تعیین می‌شود، یعنی اگر $\{A_1, \dots, A_k\}$ و $\{B_1, \dots, B_l\}$ دو مجموعه، هر یک مستقل خطی، باشند، آیا می‌توان رابطه‌ای به شکل $\langle a; A_1, \dots, A_k \rangle = \langle b; A_1, \dots, A_k \rangle$ ، برای a و b مناسب، برقرار ساخت بدون این که $k = l$ ؟ به چه اعتباری زیرمجموعه‌های $\langle a; A_1, \dots, A_k \rangle$ را "مسطح" تجسم می‌کنیم؟ همان طور که دو خط راست دارای دو نقطه متمایز مشترک بر هم منطبق می‌شوند، آیا دو زیرفضای مستوی k -بعدی با $(k+1)$ نقطه مشترک، لزوماً بر هم منطبق‌اند؟ در باقیمانده این جلسه و جلسه بعد به جواب‌های قاطعی در مورد این گونه سؤال‌ها خواهیم رسید.

(۲-۳) گزاره. فرض کنید $\{A_1, \dots, A_k\}$ مستقل خطی ست و $a \in \mathbb{R}^n$ ، در این صورت نمایش عناصر $\langle a; A_1, \dots, A_k \rangle$ به شکل $a + t_1 A_1 + \dots + t_k A_k$ منحصر به فرد است.

اثبات. توجه کنید که هر عنصر $\langle a; A_1, \dots, A_k \rangle$ طبق تعریف به شکل $a + t_1 A_1 + \dots + t_k A_k$ برای t_1, \dots, t_k مناسب است. می‌خواهیم ثابت کنیم t_1, \dots, t_k به طور یگانه تعیین می‌شوند. فرض کنید:

$$a + t_1 A_1 + \dots + t_k A_k = a + t'_1 A_1 + \dots + t'_k A_k$$

از این رابطه نتیجه می‌شود که $(t_1 - t'_1)A_1 + \dots + (t_k - t'_k)A_k = \underline{e}$. از آنجا که $\{A_1, \dots, A_k\}$ مستقل خطی است نتیجه می‌شود که همه ضرایب صفر هستند، پس $t'_i = t_i$ برای $i = 1, \dots, k$ و یگانگی ضرایب به اثبات می‌رسد. \square

گزاره فوق نشان می‌دهد هر زیرفضای مستوی k -بعدی در تناظر یک به یک با مجموعه k -تایی‌های مرتب (t_1, \dots, t_k) ، یعنی مجموعه \mathbb{R}^k ، قرار می‌گیرد، همان طور که خط راست با \mathbb{R} و

صفحه با \mathbb{R}^2 مدرج می‌شود.

نکته زیر نیز در مورد مفهوم استقلال خطی مکرراً مورد استفاده قرار خواهد گرفت:

(۲-۴) گزاره. اگر مجموعه $\{A_1, \dots, A_k\}$ مستقل خطی باشد، هر زیرمجموعه ناتهی آن نیز مستقل خطی است. اگر $\{A_1, \dots, A_k\}$ وابسته خطی باشد، هر مجموعه شامل آن نیز وابسته خطی است.

اثبات. اگر زیرمجموعه‌ای از $\{A_1, \dots, A_k\}$ وابسته خطی شود، یک ترکیب خطی از اعضای آن زیرمجموعه برابر صفر می‌شود بدون آن که همه ضرایب صفر باشند. اگر سایر A_i ها را با ضریب صفر به این رابطه بیافزاییم یک رابطه وابسته خطی برای $\{A_1, \dots, A_k\}$ به دست می‌آید. همین طور اگر یک رابطه وابستگی خطی برای $\{A_1, \dots, A_k\}$ موجود باشد، همین رابطه، وابستگی خطی برای هر مجموعه شامل $\{A_1, \dots, A_k\}$ را نشان می‌دهد. \square

زیرفضاهای مستوی (۲)

در پایان جلسه قبل به تعدادی سؤال بنیادی در مورد تعریف زیرفضای مستوی k -بعدی اشاره کردیم که قرار شد به پاسخگویی کامل آنها اقدام کنیم. قضیه زیر و روش اثبات آن کلید این بحث است.

(۱-۳) قضیه تبادلی. فرض کنید $\{A_1, \dots, A_k\}$ و $\{B_1, \dots, B_l\}$ دو زیرمجموعه مستقل خطی

از عناصر \mathbb{R}^n باشند به طوری که $\langle B_1, \dots, B_l \rangle = \langle A_1, \dots, A_k \rangle$ در این صورت $l = k$.

اثبات. نشان می‌دهیم $l \neq k$ منجر به تناقض می‌شود. مثلاً فرض کنید $k \leq l$. چون

$$B_1 \in \langle B_1, \dots, B_l \rangle = \langle A_1, \dots, A_k \rangle$$
 می‌توان نوشت:

$$B_1 = t_1 A_1 + \dots + t_k A_k \quad (1)$$

دست کم یکی از ضرایب t_1, \dots, t_k باید ناصفر باشد چه در غیر این صورت $B_1 = \mathbf{0}$ و مجموعه $\{B_1, \dots, B_l\}$ نمی‌تواند مستقل خطی باشد. مثلاً فرض می‌کنیم $t_1 \neq 0$ (در صورت لزوم اندیس‌ها را تعویض کنید به طوری که ضریب با اندیس ۱ ناصفر باشد)، پس با تبادلی جای B_1 و $t_1 A_1$ در دو طرف (۱) داریم:

$$-t_1 A_1 = -B_1 + t_2 A_2 + \dots + t_k A_k \quad (2)$$

$$A_1 = \frac{1}{t_1} B_1 - \frac{t_2}{t_1} A_2 - \dots - \frac{t_k}{t_1} A_k$$

از (۲) نتیجه می‌شود که هر ترکیب خطی A_1, \dots, A_k ، یعنی هر عنصر $\langle A_1, \dots, A_k \rangle$ ، ترکیبی خطی از $\{B_1, A_2, \dots, A_k\}$ است. حال عنصر $B_2 \in \langle B_1, \dots, B_l \rangle = \langle A_1, \dots, A_k \rangle$ را در نظر می‌گیریم.

چون ثابت کرده‌ایم هر عنصر $\langle A_1, \dots, A_k \rangle$ ترکیبی خطی از $\{B_1, A_2, \dots, A_k\}$ است، می‌توان نوشت:

$$B_2 = s_1 B_1 + s_2 A_2 + \dots + s_k A_k \quad (3)$$

در اینجا همه ضرایب s_2, \dots, s_k نمی‌توانند صفر باشند زیرا در این صورت $B_2 = s_1 B_1$ و مجموعه $\{B_1, \dots, B_l\}$ وابسته خطی خواهد شد که خلاف فرض قضیه است. پس دست کم یکی از s_2, \dots, s_k صفر نیست که مثلاً (با تعویض نام اندیس‌ها در صورت لزوم) فرض می‌کنیم $s_2 \neq 0$. پس مجدداً با تبادل مکان B_2 و $s_2 A_2$ داریم:

$$-s_2 A_2 = -s_1 B_1 - B_2 + s_3 A_3 + \dots + s_k A_k \quad (4)$$

$$A_2 = -\frac{s_1}{s_2} B_1 - \frac{1}{s_2} B_2 - \frac{s_3}{s_2} A_3 - \dots - \frac{s_k}{s_2} A_k$$

این رابطه نشان می‌دهد که هر ترکیب خطی B_1, A_2, \dots, A_k (در نتیجه هر ترکیب خطی $\langle A_1, \dots, A_k \rangle$) ترکیبی خطی از $\{B_1, B_2, A_3, \dots, A_k\}$ است. به همین روش ادامه داده، تک تک A_i ها را با B_i ها مبادله می‌کنیم. از آنجا که $k \leq l$ ، پس از k مرحله به این نتیجه می‌رسیم که هر ترکیب خطی A_1, \dots, A_k یک ترکیب خطی B_1, \dots, B_k است. از آنجا که $\langle B_1, \dots, B_l \rangle = \langle A_1, \dots, A_k \rangle$ ، اگر l اکیداً بزرگتر از k باشد، B_{k+1} باید ترکیبی خطی از B_1, \dots, B_k باشد که این خلاف استقلال خطی $\{B_1, \dots, B_l\}$ است. نتیجه این که $l = k$ و حکم به اثبات می‌رسد. \square

قضیه تبادل نشان می‌دهد که اگر یک زیرفضای خطی E° متشکل از ترکیب‌های خطی یک مجموعه مستقل خطی $\{A_1, \dots, A_k\}$ باشد، هر مجموعه مستقل خطی دیگر که ترکیب‌های خطی عناصر آن همین مجموعه را ایجاد کند باید دارای دقیقاً k عضو باشد. بدین ترتیب عدد k ، که از این پس بعد زیرفضای خطی $\langle A_1, \dots, A_k \rangle$ خوانده خواهد شد عددی است که به طور ذاتی بر این زیرفضای خطی تحمیل می‌شود. هر زیرمجموعه مستقل خطی k -عضوی $\{A_1, \dots, A_k\}$ از عناصر E° را یک پایه برای E می‌نامیم. اگر زیرفضای مستوی $E = \langle a; A_1, \dots, A_k \rangle$ از انتقال زیرفضای

خطی $\langle A_1, \dots, A_k \rangle$ به دست آمده باشد، بعد E را نیز k می‌گیریم. اکنون می‌توانیم گزاره‌ای مشابه (۴-۱) در حالت کلی ارائه کنیم:

(۲-۳) گزاره. اگر $\{A_1, \dots, A_k\}$ مستقل خطی باشد $E = \langle a; A_1, \dots, A_k \rangle$ ، $\{B_1, \dots, B_k\}$ یک مجموعه مستقل خطی از عناصر $\langle A_1, \dots, A_k \rangle$ $E^\circ = \langle A_1, \dots, A_k \rangle$ باشد، و $b \in E$ ، آنگاه $E = \langle b; B_1, \dots, B_k \rangle$.

اثبات. چون $b \in E$ داریم

$$b = a + t_1 A_1 + \dots + t_k A_k \quad (5)$$

برای اعداد حقیقی مناسب t_1, \dots, t_k . از طرفی دیگر دیدیم که هر ترکیب خطی $\{A_1, \dots, A_k\}$ یک ترکیب خطی $\{B_1, \dots, B_k\}$ است، و بالعکس. این امر را به صورت زیر می‌نویسیم:

$$\begin{aligned} x &= a + s_1 A_1 + \dots + s_k A_k \\ &= b + (s_1 - t_1) A_1 + \dots + (s_k - t_k) A_k \quad (\text{طبق } 5) \end{aligned} \quad (6)$$

و طبق (۶) می‌توان $(s_1 - t_1) A_1 + \dots + (s_k - t_k) A_k$ را به صورت ترکیبی خطی از B_1, \dots, B_k نوشت، پس $x \in \langle b; B_1, \dots, B_k \rangle$. بالعکس اگر $y \in \langle b; B_1, \dots, B_k \rangle$ آنگاه:

$$y = b + r'_1 B_1 + \dots + r'_k B_k \quad \text{برای اعداد حقیقی مناسب } r'_1, \dots, r'_k$$

ترکیب خطی $r'_1 B_1 + \dots + r'_k B_k$ را می‌توان طبق (۶) به صورت $r_1 A_1 + \dots + r_k A_k$ نوشت، پس به کمک (۵):

$$y = a + (t_1 + r_1) A_1 + \dots + (t_k + r_k) A_k$$

□ یعنی $y \in \langle a; A_1, \dots, A_k \rangle$ و حکم به اثبات می‌رسد.

گزاره زیر نیز مشابه (۱-۵) است:

(۳-۳) گزاره. اگر زیرفضاهای مستوی k -بعدی E_1, E_2 در \mathbb{R}^n دارای $(k+1)$ نقطهٔ مشترک

باشند که در یک زیرفضای با بعد کوچکتر از k در \mathbb{R}^k قرار نمی‌گیرند، آنگاه $E_1 = E_2$.

اثبات. فرض کنید P_0, P_1, \dots, P_k نقاط مشترک E_1, E_2 باشند. اگر $\{A_1, \dots, A_k\}$ یک پایه

برای E_1° باشد، عناصر E_1 را می‌توان به صورت

$$P_0 + t_1 A_1 + \dots + t_k A_k$$

نمایش داد. حال چون $P_1, \dots, P_k \in E$ ، هر یک به صورت فوق نوشته می‌شود،

پس $P_1 - P_0, \dots, P_k - P_0$ هر یک ترکیبی خطی از A_1, \dots, A_k است. می‌نویسیم

$P'_1 = P_1 - P_0, \dots, P'_k = P_k - P_0$. حال ادعا می‌کنیم $\{P'_1, \dots, P'_k\}$ مستقل خطی است. نشان

می‌دهیم فرض وابستگی خطی $\{P'_1, \dots, P'_k\}$ منجر به تناقض با این فرض گزاره می‌شود که

P_0, P_1, \dots, P_k در یک زیرفضای با بعد کوچکتر از k جای نمی‌گیرند. اگر $k = 1$ که در حالت گزاره

(۵-۱) هستیم که قبلاً ثابت شده است، پس فرض می‌کنیم $k > 1$. حال فرض کنید $\{P'_1, \dots, P'_k\}$

وابسته خطی است. دست کم یکی از P'_1, \dots, P'_k ناصفر است زیرا اگر همهٔ اینها صفر باشند داریم

$P_0 = P_1 = \dots = P_k$ پس $\{P_0, P_1, \dots, P_k\}$ در یک زیرفضای مستوی صفر بعدی \mathbb{R}^n قرار می‌گیرد

که خلاف فرض است. بدین ترتیب $\{P'_1, \dots, P'_k\}$ دارای حداقل یک زیرمجموعهٔ مستقل خطی است

(یک تک عنصری ناصفر). بزرگترین زیرمجموعه (از نظر تعداد اعضا) از $\{P'_1, \dots, P'_k\}$ را در نظر

می‌گیریم که مستقل خطی است، تعداد اعضای آن را l می‌نامیم. چنین زیرمجموعهٔ $\{P'_1, \dots, P'_k\}$

ممکن است منحصر به فرد نباشد ولیکن به هر حال $l < k$ زیرا $\{P'_1, \dots, P'_k\}$ وابسته خطی فرض شده

است. با تعویض نام اندیس‌ها در صورت لزوم، فرض می‌کنیم $\{P'_1, \dots, P'_l\}$ این زیرمجموعهٔ مستقل

خطی (دارای بیشترین تعداد ممکن عضو) باشد. زیرفضای خطی ایجاد شده توسط P'_1, \dots, P'_l را F°

می‌نامیم، $F^\circ = \langle P'_1, \dots, P'_l \rangle$. داریم $P_{l+1}, \dots, P_k \in F^\circ$ زیرا که افزودن هر یک به $\{P'_1, \dots, P'_l\}$

وابستگی خطی ایجاد می‌کند. F° یک زیرفضای خطی l -بعدی است، $l \leq k$ ، و اگر آن را با P_0

انتقال دهیم، یک زیرمجموعه مستوی l بعدی F به دست می آید:

$$F = \langle P_0; P'_1, \dots, P'_l \rangle$$

که $P_0 = P_0 + P'_1, \dots, P_l = P_0 + P'_l$ همه عضو آن هستند. چون $l < k$ ، این خلاف فرض گزاره است. نتیجه این که وابستگی خطی $\{P'_1, \dots, P'_k\}$ منجر به تناقض شد و این مجموعه باید مستقل خطی باشد. بنابراین $\{P'_1, \dots, P'_k\}$ یک پایه برای E_1° است و طبق گزاره (۳-۲) داریم

$$E_1 = \langle P_0; P'_1, \dots, P'_k \rangle$$

چون $P_0, P_1, \dots, P_k \in E_2$ عین همین استدلال برای E_2 نیز کار می کند و نتیجه می شود که:

$$E_2 = \langle P_0; P'_1, \dots, P'_k \rangle$$

و $E_1 = E_2$ نتیجه می شود. □

مفهوم توازی را که برای خط و صفحه تعریف کردیم می توانیم به زیرفضاهای مستوی از هر بعد تعمیم دهیم. فرض کنید E_1, E_2 دو زیرفضای مستوی \mathbb{R}^n باشند. می گوییم E_1, E_2 موازی هستند، و می نویسیم $E_1 \parallel E_2$ ، در صورتی که اشتراک E_1, E_2 تهی باشد و برای انتقال یافته های آنها به مبدأ، یعنی E_1°, E_2° ، داشته باشیم $E_1^\circ \subset E_2^\circ$ یا $E_2^\circ \subset E_1^\circ$.

مثال. وضعیت نسبی خط راست $l: \frac{x_1-2}{1} = \frac{x_2}{2} = \frac{x_3}{0} = \frac{x_4-1}{1} = \frac{x_5+2}{-1}$ را با زیرفضای سه بعدی $E = \langle e_3; e_1, e_4, e_5 \rangle$ از \mathbb{R}^5 بررسی کنید. عناصر E به شکل $(x_1, 0, 1, x_4, x_5)$ ، و برای نقاط l مؤلفه سوم صفر است، پس $l \cap E$ تهی است. داریم $l^\circ: \frac{x_1}{1} = \frac{x_2}{2} = \frac{x_3}{0} = \frac{x_4}{1} = \frac{x_5}{-1}$ و $E^\circ = \langle e_1, e_4, e_5 \rangle$ یعنی مؤلفه های دوم و سوم عناصر E° صفر هستند. حال نقطه $(1, 2, 0, 1, -1)$ روی l° قرار دارد ولی در E° نیست، پس $l^\circ \not\subset E^\circ$. از طرفی دیگر $E^\circ \not\subset l^\circ$ زیرا که بعد E° از بعد l° بزرگتر است. پس E و l موازی نیستند. E و l را مثل گذشته متنافر می نامیم.

ضرب داخلی و هندسه اقلیدسی در \mathbb{R}^n (۱)

تاکنون مفاهیم ابتدایی نقطه، خط، صفحه و ترازوی را به \mathbb{R}^n تعمیم داده‌ایم. در هندسه مقدماتی طول و زاویه نیز نقش مهمی ایفا می‌کنند. هدف بعدی ما معرفی این مفاهیم در \mathbb{R}^n و بحث پیرامون موضوع‌های هندسی وابسته به آنهاست. برای این کار رابطه طول، زاویه و ضرب داخلی بردارها در \mathbb{R}^2 و \mathbb{R}^3 را یادآوری می‌کنیم و خاطر نشان می‌سازیم که هر دو مفهوم "طول" و "زاویه" را می‌توان برحسب ضرب داخلی بیان کرد. از این رو در \mathbb{R}^n ضرب داخلی را مبنای بحث قرار می‌دهیم.

اگر u, v دو بردار در \mathbb{R}^3 باشند اغلب حاصل ضرب داخلی آنها، $u \cdot v$ ، به صورت زیر تعریف

می‌شود:

$$u \cdot v = |u||v| \cos \alpha \quad (1)$$

که $\alpha \in [0, \pi]$ زاویه بین u و v است و $|u|$ طول بردار را نمایش می‌دهد. با قرار دادن $u = v$ ، از (۱) نتیجه می‌شود که:

$$|u| = \sqrt{u \cdot u} \quad (2)$$

یعنی طول بردار را می‌توان برحسب ضرب داخلی بیان کرد. حال اگر $u \neq 0$ و $v \neq 0$ ، به طوری که زاویه بین آنها، $\alpha \in [0, \pi]$ ، معنی داشته باشد، از (۱) می‌توان نوشت:

$$\cos \alpha = \frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}} \quad (3)$$

تابع کسینوس روی بازه $[0, \pi]$ یک به یک است و همه مقادیر ممکن برای کسینوس، یعنی اعداد بازه $[-1, 1]$ را (یک بار) اتخاذ می کند، پس تابع معکوس $\cos^{-1} : [-1, 1] \rightarrow [0, \pi]$ وجود دارد و می توان نوشت:

$$\alpha = \cos^{-1} \left(\frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}} \right) \quad (4)$$

بدین ترتیب مفهوم زاویه بین دو بردار نیز به کمک (4) از ضرب داخلی به دست می آید. حال اگر بتوان حاصل ضرب داخلی دو بردار را مستقل از "طول" و "زاویه" تعریف کرد، می توان با به کار گرفتن (2) و (4) به تعریف مفاهیم طول و زاویه رسید. در واقع در هندسه تحلیلی دو بعدی و سه بعدی عبارتی جبری برای حاصل ضرب داخلی به دست می آید (به کمک قضیه مثلثاتی "قاعده کسینوس") که این خواست را برآورده می کند. در \mathbb{R}^2 اگر $\mathbf{u} = (u_1, u_2)$ و $\mathbf{v} = (v_1, v_2)$ ثابت می شود که:

$$u \cdot v = u_1 v_1 + u_2 v_2 \quad (5)$$

و نیز در \mathbb{R}^3 ، برای $\mathbf{u} = (u_1, u_2, u_3)$ و $\mathbf{v} = (v_1, v_2, v_3)$:

$$u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3 \quad (6)$$

عبارت های جبری (5) و (6) تعریف زیر را در \mathbb{R}^n القا می کنند:

(4-1) تعریف. برای $\mathbf{u} = (u_1, \dots, u_n)$ و $\mathbf{v} = (v_1, \dots, v_n)$ حاصل ضرب داخلی، $\mathbf{u} \cdot \mathbf{v}$ ، به

صورت زیر تعریف می شود:

$$u \cdot v = u_1 v_1 + \dots + u_n v_n$$

خواص زیر همه به سادگی از تعریف فوق نتیجه می شوند:

(4-2) خواص ابتدایی حاصل ضرب داخلی

$$(۱-۲-۴) \quad \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \text{ برای هر}$$

$$(۲-۲-۴) \quad (\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + (\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = (\mathbf{u} \cdot \mathbf{v}) + (\mathbf{u} \cdot \mathbf{w}), \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n \text{ برای هر}$$

$$(\mathbf{u} \cdot \mathbf{w}) + (\mathbf{v} \cdot \mathbf{w})$$

$$(۳-۲-۴) \quad \mathbf{u} \cdot (r\mathbf{v}) = r(\mathbf{u} \cdot \mathbf{v}) \text{ و } (r\mathbf{u}) \cdot \mathbf{v} = r(\mathbf{u} \cdot \mathbf{v}), r \in \mathbb{R} \text{ و هر } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \text{ برای هر}$$

$$(۴-۲-۴) \quad \mathbf{u} \cdot \mathbf{u} \geq 0 \text{ و } \mathbf{u} \cdot \mathbf{u} = 0 \text{ اگر و تنها اگر } \mathbf{u} = \mathbf{0}$$

با توجه به (۴-۲-۴) و با الهام از (۲)، طول $\mathbf{u} \in \mathbb{R}^n$ را به صورت زیر تعریف می‌کنیم:

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}} \quad (۷)$$

تنها n -تایی دارای طول صفر، $\mathbf{0}$ است. برای $|\mathbf{u}|$ ، علاوه بر طول، اصطلاح‌های نرم و قدر مطلق نیز به کار می‌رود. بعضی $|\mathbf{u}|$ را به $\|\mathbf{u}\|$ نمایش می‌دهند.

اکنون می‌خواهیم مفهوم زاویه بین \mathbf{u} و \mathbf{v} را نیز همانند (۴) تعریف کنیم. برای این کار باید اطمینان

حاصل کرد که با تعریف ارائه شده در \mathbb{R}^n ، مقدار عبارت $\frac{\mathbf{u} \cdot \mathbf{v}}{\sqrt{\mathbf{u} \cdot \mathbf{u}} \sqrt{\mathbf{v} \cdot \mathbf{v}}}$ همواره در بازه $[-1, 1]$ قرار دارد.

$$(۳-۴) \quad \text{نامساوی کوشی-شوارتس. برای هر } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}|$$

به علاوه شرطی لازم و کافی برای برقراری تساوی این است که \mathbf{u} و \mathbf{v} همراستا باشند.

اثبات. نخست حالتی را در نظر بگیرید که \mathbf{v} و \mathbf{u} همراستا هستند. اگر یکی از \mathbf{u} و \mathbf{v} صفر باشد که دو

طرف نامساوی بالا صفر می‌شود، در غیر این صورت $\mathbf{v} = r\mathbf{u}$ برای عدد حقیقی مناسب r . در این

صورت هر دو طرف نامساوی به $|r| |\mathbf{u}|^2$ تبدیل می‌شود و تساوی برقرار است. حال فرض کنید \mathbf{u} و

\mathbf{v} همراستا نباشند، بالاخص $\mathbf{u} \neq \mathbf{0}$ و $\mathbf{v} \neq \mathbf{0}$. n -تایی $\mathbf{v} = x\mathbf{u} + \mathbf{w}$ ، $x \in \mathbb{R}$ را در نظر بگیرید. داریم

$xu + v \neq 0$ زیرا در غیر این صورت $v = -xu$ و همراستایی ایجاد می شود. پس:

$$(xu + v) \cdot (xu + v) > 0 \quad (\text{طبق } 4-2-4)$$

$$(xu) \cdot (xu) + (xu) \cdot v + v \cdot (xu) + (v \cdot v) > 0 \quad (\text{با استفاده مکرر از } 2-2-4)$$

$$(u \cdot u)x^2 + 2(u \cdot v)x + (v \cdot v) \quad (\text{با استفاده مکرر از } 3-2-4 \text{ و } 1-2-4)$$

این نامساوی درجه دوم نسبت به x برای هر عدد حقیقی x برقرار است، پس مبین عبارت منفی است:

$$(u \cdot v)^2 - (u \cdot u)(v \cdot v) < 0$$

یا

$$|u \cdot v|^2 < |u|^2 |v|^2$$

□ که نامساوی مورد نظر است.

حال با توجه به نامساوی ثبت شده داریم

$$-1 \leq \frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}} = \frac{u \cdot v}{|u||v|} \leq 1$$

ویگانه مقدار واقع در $[0, \pi]$ که کسینوس آن برابر $\frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}}$ است زاویه بین u و v می نامیم:

$$\angle(u, v) = \cos^{-1} \frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}} \quad (8)$$

قضیه زیر که از ابتدایی ترین قضایای هندسه است، از نامساوی کوشی-شوارتس نتیجه می شود:

$$(4-4) \text{ نامساوی مثلث برای هر } u, v \in \mathbb{R}^n$$

$$|u + v| \leq |u| + |v|$$

شرط لازم و کافی برای برقراری تساوی این است که یکی از u و v مضربی نامنفی از دیگری باشد.

اثبات. کافی است نامساوی برای مجذور دو طرف ثابت شود، یعنی:

$$|u + v|^2 \leq |u|^2 + |v|^2 + 2|u||v|$$

که معادل است با

$$(u + v) \cdot (u + v) \leq (u \cdot u) + (v \cdot v) + 2|u||v|$$

و پس از ساده کردن:

$$u \cdot v \leq |u||v|$$

$$|u||v| \cos \angle(u, v) \leq |u||v| \quad (\text{طبق نتیجه نامساوی کوشی-شوارتس})$$

اگر u و v صفر باشد (که در این صورت یکی مضرب صفر از دیگری است) تساوی برقرار می شود، و اگر هر دو ناصفر باشند، رابطه بالا معادل است با:

$$\cos \angle(u, v) \leq 1$$

که همواره برقرار است. تساوی در صورتی حاصل می شود که $\cos \angle(u, v) = 1$ ، یعنی u و v همراستا و هم جهت باشند. □

برای دو عنصر u و v از \mathbb{R}^n می نویسیم $u \perp v$ ، و می گوئیم u بر v عمود است در صورتی که $u \cdot v = 0$. قضیه فیثاغورس که پایه هندسه اقلیدسی است در \mathbb{R}^n با تعاریف طول و زاویه ذکر شده برقرار است:

(۴-۵) قضیه فیثاغورس. هرگاه برای $u, v \in \mathbb{R}^n$ داشته باشیم $u \perp v$ ، آنگاه:

$$|u + v|^2 = |u|^2 + |v|^2$$

اثبات. عبارت بالا را می توان به صورت

$$(u + v) \cdot (u + v) = (u \cdot u) + (v \cdot v)$$

نوشت که با توجه به $u \cdot v = 0$ برقرار است. □

توجه کنید که با همین استدلال (یا جایگزینی $-v$ به جای v)، رابطه $|u - v|^2 = |u|^2 + |v|^2$ نیز تحت فرض $u \cdot v = 0$ برقرار است. در \mathbb{R}^n نیز، مانند \mathbb{R}^2 ، قضیه فیثاغورس به شکل $|u - v|^2 = |u|^2 + |v|^2$ حالت خاص قاعده کسینوس است که می توان به شکل زیر نوشت:

$$|u - v|^2 = |u|^2 + |v|^2 - 2|u||v| \cos \angle(u, v) \quad v \neq 0, u \neq 0 \quad (9)$$

این نیز از محاسبه ای مشابه محاسبه بالا با استفاده از (8) به دست می آید:

$$\begin{aligned} |u - v|^2 &= (u - v) \cdot (u - v) \\ &= |u|^2 + |v|^2 - 2(u \cdot v) \\ &= |u|^2 + |v|^2 - 2|u||v| \cos \angle(u, v) \quad (\text{طبق (8)}) \end{aligned}$$

به طور کلی، هرگاه $x = (x_1, \dots, x_n)$ و $y = (y_1, \dots, y_n)$ دو عنصر \mathbb{R}^n باشند، فاصله x از y ، که گاهی به $d(x, y)$ نمایش داده می شود، برابر $|x - y|$ تعریف می شود. این تعریف با آنچه از بردارها در \mathbb{R}^2 و \mathbb{R}^3 می دانیم سازگار است. قضیه های هندسی اخیر را می توانیم با توجه به این تعریف به صورت های آشناتری نیز ارائه کنیم:

(4-6) نامساوی مثلث. برای x, y, z در \mathbb{R}^n :

$$|y - z| \leq |y - x| + |x - z|$$

(4-7) قضیه فیثاغورس. هرگاه برای x, y, z در \mathbb{R}^n داشته باشیم $(x - y) \perp (x - z)$ ، آنگاه:

$$|y - z|^2 = |y - x|^2 + |z - x|^2$$

(4-8) قاعده کسینوس. برای هر x, y, z در \mathbb{R}^n داریم:

$$|y - z|^2 = |y - x|^2 + |z - x|^2 - 2|y - x||z - x| \cos \angle(y - x, z - x)$$

هریک از این روابط با جایگزینی در رابطه متناظر به دست می آید. در (۶-۴) بنویسید $u = y - x$ و $v = x - z$ ، آنگاه $u + v = y - z$. در مورد (۷-۴) و (۸-۴)، می نویسیم $u = y - x$ و $v = z - x$ ، آنگاه $u - v = y - z$.

(۹-۴) تصویر قائم روی یک راستا. یکی از موارد استفاده مهم ضرب داخلی در \mathbb{R}^2 و \mathbb{R}^3 محاسبه تصویر قائم یک بردار روی راستایی است که یک بردار ناصفر دیگر پدید می آورد (شکل ۱).

شکل ۱

اگر $u \neq v$ در \mathbb{R}^3 باشند، u' تصویر قائم u بر راستای v ، برداری است که طول آن برابر $|u| \cos \angle(u, v)$ است (در حالتی که $u = 0$ ، این طول برابر صفر در نظر گرفته می شود هرچند که زاویه بین u, v تعریف نشده است)، u' مضربی از v است، $u' = r u$ ، و علامت r با علامت کسینوس زاویه بین u و v یکی است. بدین ترتیب اگر $\frac{v}{|v|}$ بردار واحد در جهت v باشد، می توان نوشت:

$$u' = |u| \cos \angle(u, v) \frac{v}{|v|} \quad (10)$$

یا معادلاً:

$$u' = \frac{u \cdot v}{|v|^2} v \quad (11)$$

با توجه به این که عبارتهای سمت راست (۱۱) همه در \mathbb{R}^n معنی دارند، می توانیم تصویر قائم u بر $v \neq 0$ در \mathbb{R}^n را نیز به صورت (۱۱) تعریف کنیم. در حالت خاصی که v یک بردار واحد باشد:

$$|v| = 1 \quad u' = (u \cdot v)v \quad (12)$$

مثال (پایه متداول \mathbb{R}^n). e_1, \dots, e_n را مانند مثال صفحه ۸ جلسه قبل در نظر بگیرید، یعنی

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$$

هر $\mathbf{x} = (x_1, \dots, x_n)$ را می توان به صورت:

$$x = x_1 e_1 + \dots + x_n e_n$$

نوشت. در واقع $x_i e_i$ تصویر قائم x بر راستای e_i (محور i ام) است:

$$(x \cdot e_i) e_i = x_i e_i$$

پس می توان نوشت:

$$x = \sum_{i=1}^n (x \cdot e_i) e_i \quad (۱۳)$$

مجموعه $\{e_1, \dots, e_n\}$ را پایه متداول \mathbb{R}^n می نامند.

ضرب داخلی و هندسه اقلیدسی در \mathbb{R}^n (۲)

در پایان جلسه گذشته مشاهده کردیم که اگر e_1, \dots, e_n پایه متداول \mathbb{R}^n باشد، هر عنصر x از \mathbb{R}^n را می توان به صورت

$$x = \sum_{i=1}^n (x \cdot e_i) e_i \quad (1)$$

نمایش داد. حال به یک تعمیم این مطلب اشاره می کنیم. به طور کلی، اگر $\{b_1, \dots, b_k\}$ یک مجموعه در \mathbb{R}^n باشد (در حالت خاص، پایه ای برای زیرفضای خطی E از \mathbb{R}^n)، $\{b_1, \dots, b_k\}$ را متعامد می نامیم در صورتی که برای $b_i \perp b_j$ برای $i \neq j$. مجموعه $\{b_1, \dots, b_k\}$ را راست هنجار می نامیم در صورتی که متعامد باشد و به علاوه $|b_i| = 1$ برای $i = 1, \dots, k$. به عنوان نمونه، $\{e_1, \dots, e_n\}$ یک پایه راست هنجار برای \mathbb{R}^n است و هرگاه اعداد حقیقی ناصفر c_1, \dots, c_n داده شده باشند، $\{c_1 e_1, \dots, c_n e_n\}$ یک پایه متعامد برای \mathbb{R}^n است. حال فرض کنید $\{b_1, \dots, b_k\}$ یک پایه راست هنجار برای زیرفضای خطی E باشد و $x \in E$. داریم $x = \sum_{i=1}^k t_i b_i$. اگر ضرب داخلی دو طرف را با b_j ، j ثابت، محاسبه کنیم، حاصل می شود $x \cdot b_j = \sum_{i=1}^k t_i (b_i \cdot b_j)$. از آنجا که پایه راست هنجار است، نتیجه می شود که $b_i \cdot b_j = 0$ اگر $i \neq j$ و $b_j \cdot b_j = 1$ پس $x \cdot b_j = t_j$. بنابراین عین فرمول (۱) در اینجا نیز برقرار است:

$$x = \sum_{i=1}^k (x \cdot b_i) b_i \quad (2)$$

بدین ترتیب محاسبه ضرایب نمایش نسبت به یک پایه راست هنجار بسیار ساده است. به زودی یک روش عمومی برای ساختن پایه‌های راست هنجار برای زیرفضاهای خطی ارائه خواهیم کرد، ولی مقدماً گزاره زیر را مطرح می‌کنیم:

(۱-۱۰) گزاره. اگر $\{B_1, \dots, B_k\}$ یک مجموعه متعامد متشکل از عناصر ناصفر \mathbb{R}^n باشد، آنگاه $\{B_1, \dots, B_k\}$ مستقل خطی است.

اثبات. فرض کنید $c_1 B_1 + \dots + c_k B_k = \underline{0}$ باید ثابت کنیم همه c_i ها صفر هستند. برای j ثابت و دلخواه، حاصل ضرب داخلی دو طرف رابطه را با B_j محاسبه می‌کنیم:

$$c_1(B_1 \cdot B_j) + \dots + (B_k \cdot B_j) = 0$$

ولی $B_i \cdot B_j = 0$ مگر وقتی که $i = j$ که در این صورت $B_j \cdot B_j = |B_j|^2$ ناصفر است زیرا که همه B_1, \dots, B_k ناصفر فرض شده‌اند. پس از رابطه $c_j |B_j|^2 = 0$ نتیجه می‌گیریم که $c_j = 0$. چون j در بین $1, \dots, k$ دلخواه بود، حکم به اثبات می‌رسد. \square

اکنون نشان می‌دهیم چگونه می‌توان به کمک هر پایه داده شده برای زیرفضای خطی E ، یک پایه راست هنجار برای E ساخت:

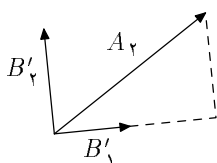
(۱-۲) روش گرام-اشمیت. فرض کنید E یک زیرفضای خطی k -بعدی از \mathbb{R}^n است که متشکل از ترکیب‌های خطی $t_1 A_1 + \dots + t_k A_k$ از مجموعه مستقل خطی $\{A_1, \dots, A_k\}$ می‌باشد. می‌خواهیم یک پایه راست هنجار $\{B_1, \dots, B_k\}$ برای E بسازیم. برای این کار نخست یک پایه متعامد $\{B'_1, \dots, B'_k\}$ برای E به دست می‌آوریم، سپس با قرار دادن $B_i = \frac{1}{|B'_i|} B'_i$ ، یک پایه راست هنجار حاصل می‌شود.

برای ساختن $\{B'_1, \dots, B'_k\}$ ، گام به گام به صورت زیر عمل می‌کنیم:

$$(1) \text{ قرار می دهیم } B'_1 = A_1.$$

(2) برای ساختن B'_2 ، تصویر قائم A_2 بر $B'_1 = A_1$ را از A_2 کم می کنیم:

$$B'_2 = A_2 - \frac{A_2 \cdot B'_1}{B'_1 \cdot B'_1} B'_1 \quad (3)$$



توجه کنید که $B'_2 \perp B'_1$ زیرا حاصل ضرب داخلی طرف راست با B'_1 صفر می شود. به علاوه $B'_2 \neq 0$ زیرا که اگر $B'_2 = 0$ ، آنگاه با توجه به این که $A_1 = B'_1$ ، یک رابطه وابستگی خطی در $A_2 - \frac{A_2 \cdot B'_1}{B'_1 \cdot B'_1} A_1$ پدید می آید که خلاف فرض استقلال خطی $\{A_1, A_2\}$ و $\{B'_1, B'_2\}$ است. نتیجه استقلال خطی هر زیرمجموعه آن، است. نتیجه این که بنابر گزاره ۱-۱، $\langle B'_1, B'_2 \rangle = \langle A_1, A_2 \rangle$ زیرا که $B'_1 = A_1$ و طبق (3) مستقل خطی است. ضمناً توجه کنید که $\langle B'_1, B'_2 \rangle = \langle A_1, A_2 \rangle$ زیرا که $B'_1 = A_1$ و طبق (3) می توان نقش A_2 و B'_2 را مبادله کرد.

(3) به طور استقرایی می توان به این روش ادامه داد. اگر $\{B'_1, \dots, B'_j\}$ به دست آمده باشد، $j < k$ ، که عناصر آن ناصفر و دو به دو بر هم عمودند به طوری که $\langle B'_1, \dots, B'_j \rangle = \langle A_1, \dots, A_j \rangle$ را به روش زیر می سازیم:

$$B'_{j+1} = A_{j+1} - \frac{A_{j+1} \cdot B'_1}{B'_1 \cdot B'_1} B'_1 - \dots - \frac{A_{j+1} \cdot B'_j}{B'_j \cdot B'_j} B'_j \quad (4)$$

در واقع در (4) تصویر قائم A_{j+1} بر تک تک B'_1, \dots, B'_j را از A_{j+1} کم کرده ایم. حاصل باید بر B'_1, \dots, B'_j عمود باشد که این موضوع با محاسبه حاصل ضرب داخلی دو طرف راست (4) با $B'_i, i = 1, \dots, j$ ، مشاهده می شود. همچنین توجه کنید که $B'_{j+1} \neq 0$ زیرا که طبق فرض

استقراء $\langle A_1, \dots, A_j \rangle = \langle B'_1, \dots, B'_j \rangle$ و $\{A_1, \dots, A_j, A_{j+1}\}$ مستقل خطی است. و بالاخره مجموعه متعامد $\{B'_1, \dots, B'_{j+1}\}$ که همه عناصرش ناصفرند، طبق $1-10$ مستقل خطی است و مجدداً با استدلال مشابه گام ۲ داریم $\langle B'_1, \dots, B'_{j+1} \rangle = \langle A_1, \dots, A_{j+1} \rangle$.

بدین ترتیب با ادامه روش، در گام k به مجموعه $\{B'_1, \dots, B'_k\}$ دست می‌یابیم که از عناصر ناصفر دو به دو برهم عمود در E تشکیل شده است. چون تعداد عناصر برابر بعد E است و مجموعه طبق $1-10$ مستقل خطی است، به یک پایه متعامد برای E دست یافته‌ایم. \square

مثال. تحقیق می‌کنیم که مجموعه $\{A_1, A_2, A_3\}$ در \mathbb{R}^4 که در آن $A_1 = (0, 0, -1, 1)$ ، $A_2 = (1, 0, 2, 0)$ و $A_3 = (1, 1, 3, 0)$ یک مجموعه مستقل خطی است. فرض کنید پس $c_1 A_1 + c_2 A_2 + c_3 A_3 = \underline{0}$

$$(c_2 + c_3, c_3 - c_1 + 2c_2 + 3c_3, c_1) = (0, 0, 0, 0)$$

نتیجه این که $c_3 = 0$ و $c_1 = 0$ که از اینها نتیجه می‌شود $c_2 = 0$ پس $\{A_1, A_2, A_3\}$ مستقل خطی است. $E = \langle A_1, A_2, A_3 \rangle$ را زیرفضای متشکل از ترکیب‌های خطی A_1, A_2, A_3 می‌گیریم، توجه کنید که $\{A_1, A_2, A_3\}$ متعامد نیست، مثلاً $A_1 \cdot A_2 = -2 \neq 0$. روش گرام-اشمیت را به کار می‌گیریم:

$$B'_1 = A_1 = (0, 0, -1, 1)$$

$$B_2 = (1, 0, 2, 0) - \frac{(1, 0, 2, 0) \cdot (0, 0, -1, 1)}{(0, 0, -1, 1) \cdot (0, 0, -1, 1)} (0, 0, -1, 1) = (1, 0, 2, 0) + (0, 0, -1, 1)$$

یا $B'_2 = (1, 0, 1, 1)$. بالاخره:

$$B'_3 = (1, 1, 3, 0) - \frac{(1, 1, 3, 0) \cdot (0, 0, -1, 1)}{(0, 0, -1, 1) \cdot (0, 0, -1, 1)} (0, 0, -1, 1) - \frac{(1, 1, 3, 0) \cdot (1, 0, 1, 1)}{(1, 0, 1, 1) \cdot (1, 0, 1, 1)} (1, 0, 1, 1)$$

$$= (1, 1, 3, 0) + \frac{3}{4} (0, 0, -1, 1) - \frac{4}{5} (1, 0, 1, 1)$$

$$= \left(-\frac{1}{5}, 1, \frac{1}{5}, \frac{1}{5}\right)$$

پس مجموعه $\{(0, 0, -1, 1), (1, 0, 1, 1), (-\frac{1}{\sqrt{3}}, 1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$ یک پایه متعامد برای E است، که می توان این مطلب را مستقیماً نیز تحقیق کرد. برای یافتن یک پایه راست هنجار، هر یک از B'_i ها را در معکوس طول آن ضرب می کنیم:

$$B_1 = (0, 0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$$

$$B_2 = (\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$$

$$B_3 = (-\frac{2}{\sqrt{42}}, \frac{7}{\sqrt{42}}, \frac{1}{\sqrt{42}}, \frac{1}{\sqrt{42}})$$

روش گرام-اشمیت در تکمیل یک مجموعه متعامد (یا راست هنجار) به یک پایه کامل نیز به کار گرفته می شود:

(۳-۱۰) تکمیل مجموعه متعامد به پایه. فرض کنید $\{B_1, \dots, B_k\}$ یک مجموعه راست هنجار از عناصر E باشد که $k < n$. نشان می دهیم چگونه می توان با استفاده از روش ۱۰-۲، عناصر B_{k+1}, \dots, B_n از \mathbb{R}^n یافت، به طوری که $\{B_1, \dots, B_n\}$ یک پایه راست هنجار برای تمام \mathbb{R}^n باشد. چون $k < n$ ، عنصری A_{k+1} از \mathbb{R}^n یافت می شود که در $\langle A_1, \dots, A_k \rangle$ نیست. به روش گرام-اشمیت با کم کردن تصویر قائم A_{k+1} بر B_1, \dots, B_k ، عنصری B'_{k+1} به دست می آوریم که $\{B_1, \dots, B_k, B'_{k+1}\}$ متعامد و مستقل خطی است:

$$B'_{k+1} = A_{k+1} - (A_{k+1} \cdot B_1)B_1 - \dots - (A_{k+1} \cdot B_k)B_k$$

سپس با تعریف $B_{k+1} = \frac{1}{|B'_{k+1}|} B'_{k+1}$ ، یک مجموعه $\{B_1, \dots, B_k, B_{k+1}\}$ حاصل می شود که عناصر آن طول واحد دارند و دو به دو بر هم عمودند (یعنی یک مجموعه راست هنجار). اگر $n = k + 1$ که چون این مجموعه مستقل خطی به تعدادی برابر بعد \mathbb{R}^n عضو دارد، خود یک پایه راست هنجار برای \mathbb{R}^n می شود، $\mathbb{R}^n = \langle B_1, \dots, B_{k+1} \rangle$. اگر $k + 1 < n$ ، عنصری A_{k+2} در \mathbb{R}^n یافت می شود که در $\langle B_1, \dots, B_{k+1} \rangle$ نیست. به روش بالا، از A_{k+2} ، عنصری B_{k+2} می سازیم که

$\{B_1, \dots, B_{k+2}\}$ راست هنجار است. اگر این عمل را $(n - k)$ بار انجام دهیم به یک مجموعه راست هنجار n عضوی $\{B_1, \dots, B_n\}$ می‌رسیم که بنابراین یک پایه برای \mathbb{R}^n خواهد بود.

(۱۰-۴) کاربرد در معادله زیرفضاها. زیرفضای مستوی E از \mathbb{R}^n را یک ابرصفحه در \mathbb{R}^n می‌نامیم در صورتی که بعد E برابر $(n - 1)$ باشد. بدین ترتیب ابرصفحه‌های \mathbb{R} ، نقاط \mathbb{R} هستند، ابرصفحه‌های \mathbb{R}^2 ، خطوط راست در \mathbb{R}^2 ، و ابرصفحه‌های \mathbb{R}^3 ، صفحات واقع در \mathbb{R}^3 می‌باشند. حال یک ابرصفحه E از \mathbb{R}^n در نظر بگیرید. اگر E° انتقال یافته E به \circ باشد، می‌توان برای E° یک پایه متعامد $\{B_1, \dots, B_{n-1}\}$ طبق روش گرام-اشمیت در نظر گرفت. با استفاده از ۱۰-۳، عنصری C ، $C \neq \circ$ در \mathbb{R}^n وجود دارد که بر B_1, \dots, B_{n-1} عمود است و در نتیجه $\{B_1, \dots, B_{n-1}, C\}$ یک پایه متعامد برای \mathbb{R}^n می‌باشد. توجه کنید که شرطی لازم و کافی برای این که عنصر x از \mathbb{R}^n در E° باشد این است که:

$$C \cdot x = \circ \quad (5)$$

زیرا اگر بنویسیم $x = \sum_{i=1}^{n-1} c_i B_i + c_n C$ عناصر E° دقیقاً آن x هایی هستند که $c_n = \circ$ پس اگر $x \in E$ ، آنگاه $x = \sum_{i=1}^{n-1} c_i (B_i \cdot B_n) = \circ$ بالعکس اگر $C \cdot x = \circ$ داریم

$$\circ = C \cdot x = \sum_{i=1}^{n-1} c_i (C \cdot B_i) + c_n (C \cdot C) = c_n$$

حال فرض کنید ابرصفحه E به شکل $E = \langle p; B_1, \dots, B_{n-1} \rangle$ توصیف شده است که $p = (p_1, \dots, p_n)$ نقطه‌ای در E است. نقطه $x = (x_1, \dots, x_n)$ در E است اگر و تنها اگر $x - p \in E^\circ$ پس شرطی لازم و کافی برای این است که

$$(x - p) \cdot C = \circ \quad (6)$$

اگر n تایی C را به (C_1, \dots, C_n) نمایش دهیم؛ (۶) معادل است با

$$C_1(x_1 - p_1) + \dots + C_n(x_n - p_n) = 0 \quad (7)$$

این رابطه را معادله ابرصفحه گذرا از p عمود بر C می نامند. نمایش متداول خط در صفحه و صفحه در فضای سه بعدی حالت های خاص (۷) هستند.

(۷) را می توان به زیرفضاهای مستوی ابعاد غیر از $(n - 1)$ تعمیم داد. اگر E یک زیرفضای مستوی k -بعدی به شکل $E = \langle p; B_1, \dots, B_k \rangle$ باشد که در آن $\{B_1, \dots, B_k\}$ یک پایه متعامد برای انتقال یافته E به مبدأ، E° است؛ طبق روش $(3-10)$ ، $\{B_1, \dots, B_k\}$ را به پایه ای متعامد $\{B_1, \dots, B_k, B_{k+1}, \dots, B_n\}$ برای \mathbb{R}^n تکمیل می کنیم. با استدلالی مشابه آنچه برای ابرصفحه ها آمد؛ شرطی لازم و کافی برای این که نقطه x در E باشد این خواهد بود که $x - p$ بر B_{k+1}, \dots, B_n عمود باشد. پس مجموعه نقاطی است که در $(n - k)$ رابطه زیر صدق می کنند:

$$B_{k+1} \cdot (x - p) = 0, \dots, B_n \cdot (x - p) = 0 \quad (8)$$

مثال. می خواهیم در \mathbb{R}^4 فاصله نقطه $(1, -1, 0, 2)$ را از خط راست $\frac{x_1}{1} = \frac{x_2-1}{-1} = \frac{x_3+1}{-1} = \frac{x_4}{3}$ محاسبه کنیم. مقصود از فاصله نقطه از خط؛ کوتاهترین فاصله ممکن از نقطه داده شده به نقاط روی خط است. توجه کنید که نقطه داده شده، $(1, -1, 0, 2)$ ؛ روی خط داده شده قرار ندارد (چرا؟)؛ پس خط و نقطه روی یک صفحه مستوی منحصر به فرد قرار می گیرند و می توان به روال هندسه عادی فرض کرد کوتاهترین فاصله با رسم عمود از نقطه بر خط به دست می آید (اثباتی مستقیم از این مطلب در تمرین زیر آمده است). با فرض این مطلب برای یافتن پای عمود از $(1, -1, 0, 2)$ به خط داده شده؛ معادله ابرصفحه گذرا از $(1, -1, 0, 2)$ عمود بر $(2, 1, -1, 3)$ را می نویسیم و اشتراک آن را با خط پیدا می کنیم. نقطه تقاطع نزدیکترین نقطه خط به $(1, -1, 0, 2)$ است. طبق (۷)؛ معادله

ابرفه مورد نظر هست:

$$2(x_1 - 1) + (x_2 + 1) - x_3 + 3(x_4 - 2) = 0$$

حال با جایگزینی از $\frac{x_2}{3} = \frac{x_3 - 1}{-1} = \frac{x_4 + 1}{-1} = \frac{x_4}{-3}$ در رابطه بالا برحسب x_1 داریم:

$$2(x_1 - 1) + \left(\frac{x_1}{3} + 2\right) + \left(\frac{x_1}{3} + 1\right) + 3\left(\frac{2}{3}x_1 - 2\right) = 0$$

$$x_1 = \frac{2}{3}$$

پس $x_2 = \frac{4}{3}, x_3 = -\frac{4}{3}, x_4 = 1$ و نقطه $(\frac{2}{3}, \frac{4}{3}, -\frac{4}{3}, 1)$ نزدیکترین نقطه خط به $(1, -1, 0, 2)$ می باشد که فاصله اش برابر مقدار زیر است:

$$\sqrt{\left(1 - \frac{2}{3}\right)^2 + \left(-1 - \frac{4}{3}\right)^2 + \left(\frac{4}{3}\right)^2 + (2 - 1)^2} = \frac{5}{\sqrt{3}}$$

تمرین. فرض کنید در \mathbb{R}^n خط گذرا از a به موازات A است، $a = (a_1, \dots, a_n)$ ، $A = (A_1, \dots, A_n)$ و $a \neq 0$ یک نقطه در \mathbb{R}^n است. نشان دهید به ازای نقطه منحصر به فردی q روی l نزدیکترین فاصله به p حاصل می شود و در واقع $q = a + \frac{(p-a) \cdot A}{A \cdot A} A$ (یعنی از نقطه a روی l باید به اندازه تصویر قائم $(p - a)$ روی A جدا کرد تا به نقطه q رسید).

راهنمایی. معادله پارامتری خط l ، $x_i = a_i + tA_i$ را در نظر بگیرید و مجذور فاصله p از نقطه $a + tA$ را برحسب t بنویسید. عبارت به دست آمده، یک عبارت درجه ۲ برحسب t خواهد بود که با توجه به نقطه مینیموم سهمی می توان t مطلوب را از آن محاسبه کرد.

نگاشت‌های خطی (۱)

وقتی $m \times n$ عدد حقیقی a_{ij} ، $i = 1, \dots, m$ ، $j = 1, \dots, n$ را در یک قالب m در n ، با m ردیف و n ستون به شکل زیر بنویسیم، یک ماتریس $m \times n$ (حقیقی) تعریف می‌شود:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & & a_{mn} \end{bmatrix}$$

اعداد مندرج در ماتریس را درایه‌های ماتریس می‌نامند. مقصود از a_{ij} درایه‌ای است که در ردیف i ام و ستون j ام قرار دارد. گاهی اوقات A را به صورت $[a_{ij}]$ ، یا به طور کاملتر $[a_{ij}]_{i=1, \dots, m, j=1, \dots, n}$ می‌نویسیم.

برای ماتریس‌های هم اندازه می‌توان عمل جمع، و عمل ضرب در اعداد حقیقی، همانند عملیات مشابه برای n تایی‌ها (بردارها) را تعریف کرد. اگر $A = [a_{ij}]$ و $B = [b_{ij}]$ دو ماتریس $m \times n$ باشند، مجموع دو ماتریس، $A + B = [c_{ij}]$ با فرمول

$$c_{ij} = a_{ij} + b_{ij} \quad (1)$$

تعریف می‌شود. تحقیق صحت خواص زیر از تعریف سراسر است:

(۱-۱) خواص ابتدایی جمع ماتریس‌ها

(۱-۱-۱) تعویض‌پذیری (جابجایی). اگر A و B دو ماتریس $m \times n$ باشند، داریم

$$.A + B = B + A$$

(۱۱-۱-۲) شرکت پذیری. اگر A, B و C ماتریس های $m \times n$ باشند، داریم $A + (B + C) = (A + B) + C$.

(۱۱-۱-۳) عنصر بی اثر. ماتریس صفر $m \times n$ که همه درایه های آن صفر است و به O نمایش می دهیم (یگانه ماتریس $m \times n$) دارای این ویژگی است که برای هر ماتریس $m \times n$ ، A داریم $A + O = O + A = A$.

(۱۱-۱-۴) عنصر قرینه. برای $A = [a_{ij}]$ ، ماتریس $-A = [b_{ij}]$ که با $b_{ij} = -a_{ij}$ تعریف می شود (یگانه ماتریس) دارای این ویژگی است که $A + (-A) = (-A) + A = O$.

حال اگر $A = [a_{ij}]$ یک ماتریس $m \times n$ باشد و $r \in \mathbb{R}$ ، ماتریس $rA = [b_{ij}]$ با فرمول $b_{ij} = ra_{ij}$ تعریف می شود، یعنی همه درایه های A در عدد r ضرب می شوند. تحقیق خواص زیر نیز سراسر است:

(۱۱-۲) خواص ابتدایی ضرب اعداد در ماتریس ها

$$(۱۱-۲-۱) 1 \cdot A = A \text{ برای هر ماتریس } A.$$

$$(۱۱-۲-۲) (rs)A = r(sA) \text{ برای هر ماتریس } A \text{ و هر زوج عدد حقیقی } r, s.$$

$$(۱۱-۲-۳) (r+s)A = rA + sA \text{ برای هر ماتریس } A \text{ و هر زوج عدد حقیقی } r, s.$$

$$(۱۱-۲-۴) r(A+B) = rA + rB \text{ برای هر عدد حقیقی } r \text{ و هر دو ماتریس هم اندازه } A \text{ و } B.$$

علاوه بر عملیات بالا، مفهوم ضرب ماتریس ها نیز مابین ماتریس های اندازه های مناسب که توضیح

داده خواهد شد، تعریف می‌شود. اگر $A = [a_{ij}]$ یک ماتریس $m \times n$ و $B = [b_{ij}]$ یک ماتریس $n \times p$ باشد، ماتریس حاصل ضرب، $AB = [c_{ij}]$ ، به صورت زیر تعریف می‌شود:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} \quad (2)$$

توجه کنید که لازمه تعریف حاصل ضرب این است که تعداد ستون‌های A برابر تعداد ردیف‌های B باشد، بنابراین عبارت (۲) معنی دارد و در واقع می‌توان c_{ij} را حاصل ضرب داخلی ردیف i ام ماتریس A به عنوان n تایی (a_{i1}, \dots, a_{in}) با ستون j ام ماتریس B ، به عنوان n تایی (b_{j1}, \dots, b_{nj}) تلقی کرد. انگیزه این تعریف را در جلسه آینده بررسی خواهیم کرد.

(۱۱-۳) خواص ابتدایی ضرب ماتریس‌ها

(۱۱-۳-۱) شرکت‌پذیری. A ، B و C ماتریس‌های دارای اندازه به ترتیب $m \times n$ ، $n \times p$ و $p \times q$ هستند. در این صورت:

$$(AB)C = A(BC)$$

(۱۱-۳-۲) قانون پخش. داریم

$$(A+B)C = (AC) + (BC) \quad , \quad A(B+C) = (AB) + (AC)$$

مشروط بر این که اندازه ماتریس‌های فوق برای عملیات ذکر شده مناسب باشد.

(۱۱-۳-۳) ماتریس واحد $n \times n$ ، $I_n = [\delta_{ij}]$ بدین صورت تعریف می‌شود که:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

برای هر ماتریس $A, m \times n$ و هر ماتریس $B, n \times p$ داریم:

$$A \cdot I_n = A \quad , \quad I_n \cdot B = B$$

اثبات (۱۱-۳-۱) را ارائه می‌کنیم، دوتای دیگر ساده‌ترند و تحقیق آنها به خواننده واگذار

می‌شود. برای (۱۱-۳-۱) بنویسید $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$, $AB = [d_{ij}]$, $BC = [e_{ij}]$

و $(AB)C = [x_{ij}]$ و $A(BC) = [y_{ij}]$. طبق تعریف حاصل ضرب ماتریس‌ها عمل می‌کنیم:

$$\begin{aligned} x_{ij} &= \sum_{k=1}^p d_{ik} c_{kj} \\ &= \sum_{k=1}^p \left(\sum_{l=1}^n a_{il} b_{lk} \right) c_{kj} \\ &= \sum_{k=1}^p \sum_{l=1}^n a_{il} b_{lk} c_{kj} \\ &= \sum_{l=1}^n a_{il} \left(\sum_{k=1}^p b_{lk} c_{kj} \right) \\ &= \sum_{l=1}^n a_{il} e_{lj} \\ &= y_{ij} \end{aligned}$$

و حکم به اثبات می‌رسد.

در اینجا قابل ذکر است که اگر AB تعریف شده باشد، لزومی ندارد BA نیز تعریف شدنی باشد

زیرا که تعریف AB مستلزم این است که تعداد ستون‌های A برابر تعداد ردیف‌های B باشد، و این

حکمی در مورد برابری تعداد ستون‌های B با تعداد ردیف‌های A نمی‌کند. حتی اگر AB و BA هر دو

تعریف شدنی باشند، لزومی ندارد هر دو یک اندازه باشند، مثلاً اگر A یک ماتریس $1 \times n$ و B یک

ماتریس $n \times 1$ باشند، $n > 1$

$$A = [a_1 \dots a_n] \quad , \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

AB یک ماتریس 1×1 ، با تک درایه $\sum_{i=1}^n a_i b_i$ می‌شود، در حالی که BA یک ماتریس $n \times n$ است که درایه ردیف i ام و ستون j ام آن عبارت است از $b_j a_i$. حتی اگر A و B هر دو $n \times n$ ، با $n \geq 2$ باشند، لزومی بر تساوی AB و BA نیست، مثلاً:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad , \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

البته در مورد ماتریس I_n و هر ماتریس $n \times n$ ، A ، از $3-3-11$ نتیجه می‌شود که $AI_n = I_n A$ ($A =$).

توجه کنید که برای یک n -تایی مرتب از اعداد، مثلاً $x = (x_1, \dots, x_n)$ اکنون سه نمایش ممکن داریم، یکی به صورت بالا که معمولاً تجسم برداری دارد، دیگر به صورت یک ماتریس $1 \times n$ ، $[x_1 \dots x_n]$ و بالاخره به صورت یک ماتریس $n \times 1$ ، $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. اغلب لازم است که بین این سه نمایش تمایز قائل شویم تا ابهامی در عملیات حاصل نشود. برای این منظور، هرگاه n تایی $x = (x_1, \dots, x_n)$ به صورت یک ماتریس $1 \times n$ نوشته شود، آن را به $\langle x |$ نمایش می‌دهیم، و هرگاه به صورت یک ماتریس $n \times 1$ نمایش داده شود، آن را به $|x\rangle$ نمایش می‌دهیم.

با این مقدمه در مورد ماتریس‌ها، به موضوع اصلی درس که "نگاشت‌های خطی" است می‌پردازیم. کلمات "تابع"، "نگاشت" و "تبدیل" در اینجا به صورت کاملاً مترادف به کار خواهند رفت هر چند که به لحاظ تاریخی گاهی تصاویر ذهنی متفاوتی از این کلمات القاء می‌شود. اگر S و T دو مجموعه باشند، مقصود از یک تابع (\equiv نگاشت، \equiv تبدیل) از S به T ، قانونی f است که به هر عنصر s از S ، عنصر مشخصی $f(s)$ از T نسبت می‌دهد. می‌نویسیم $f: S \rightarrow T$. در گذشته کلمه "تابع" معمولاً وقتی به کار می‌رفته است که T برابر \mathbb{R} یا \mathbb{C} باشد، "تبدیل" در حالت $S = T$ ، یعنی تابع‌های از یک مجموعه

به خود آن، و "نگاشت" برای تابع‌های $\mathbb{R}^m \rightarrow \mathbb{R}^n$ ، بزرگتر از ۱، به کار می‌رفته است. ما چنین تمایزهایی قابل نمی‌شویم. برای تابع $f: S \rightarrow T$ ، مجموعه S را دامنه (= قلمرو)، و مجموعه T را بُرد می‌نامیم. مجموعه $\{f(s) \mid s \in S\}$ که یک زیرمجموعه T است، تصویر f ، یا تصویر S تحت f خوانده می‌شود. در واقع برای هر زیرمجموعه E از S ، تصویر E تحت f ، که به $f(E)$ نمایش داده می‌شود به صورت:

$$f(E) = \{f(s) \mid s \in E\}$$

تعریف می‌شود. تابع f را یک به یک می‌نامیم در صورتی که $f(s_1) = f(s_2)$ همواره نتیجه دهد $s_1 = s_2$ ، و f پوشا می‌نامیم اگر $f(S) = T$. برای تابع‌های یک به یک و پوشا، $f: S \rightarrow T$ ، تابعی به نام وارون f (یا معکوس f) به صورت زیر تعریف می‌شود که به $f^{-1}: T \rightarrow S$ نمایش می‌دهیم. هرگاه $t \in T$ ، چون f پوشا است، عنصری s از S وجود دارد که $f(s) = t$. به علاوه چون f یک به یک فرض شده است، چنین عنصر s به طور منحصر به فرد تعیین می‌شود، پس f^{-1} به صورت قانونی بی‌ابهام $f^{-1}(t) = s$ تعریف شدنی است و داریم:

$$f \circ f^{-1} = \mathbb{1}_T \quad , \quad f^{-1} \circ f = \mathbb{1}_S$$

که مقصود از $\mathbb{1}_X$ تابع همانی مجموعه X است، $\mathbb{1}_X(x) = x$ برای هر $x \in X$.

نماد $f^{-1}(t)$ حتی وقتی f^{-1} تعریف شدنی نباشد نیز به کار می‌رود. مقصود از $f^{-1}(t)$ مجموعه نقاط S است که تحت f به t فرستاده (= نگاشته) می‌شوند، یعنی:

$$f^{-1}(t) = \{s \in S \mid f(s) = t\}$$

$f^{-1}(t)$ را گاهی مجموعه تراز منسوب به t می‌خوانیم. به طور کلی، اگر Y زیرمجموعه‌ای از T باشد، زیرمجموعه $f^{-1}(Y)$ (تصویر وارون Y) یک زیرمجموعه S است که بدین صورت تعریف می‌شود:

$$f^{-1}(Y) = \{s \in S \mid f(s) \in Y\}$$

موضوع اصلی درس ما بررسی تابع های $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (یا به طور کلی تر، تابع های $f: S \rightarrow \mathbb{R}^m$ ، $S \subset \mathbb{R}^n$) است. بدین ترتیب چنین تابع f به هر n تایی مرتب از اعداد حقیقی در دامنه، یک m تایی مرتب از اعداد حقیقی نسبت می دهد:

$$f(x_1, \dots, x_n) = (y_1, \dots, y_m) \quad (3)$$

هر y_i وابسته به (x_1, \dots, x_n) است، پس در واقع ارائه f همانند ارائه m تابع f_1, \dots, f_m ، هر یک از \mathbb{R}^n به \mathbb{R} است، $f_i(x_1, \dots, x_n) = y_i$ ، $i = 1, \dots, m$. (3) را به صورت زیر نیز می نویسیم:

$$\begin{cases} f_1(x_1, \dots, x_n) = y_1 \\ \dots \\ f_m(x_1, \dots, x_n) = y_m \end{cases} \quad (4)$$

گاهی می نویسیم $f = (f_1, \dots, f_m)$ و هر f_i را یک مؤلفه f می نامیم.

تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ در (3) یا (4) را خطی می نامیم در صورتی که هر $f_i(x_1, \dots, x_n)$ یک

عبارت همگن درجه اول نسبت به x_1, \dots, x_n باشد، یعنی:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \dots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{cases} \quad (5)$$

مقصود از "همگن" این است که همه جملات سمت راست از یک درجه اند، در اینجا همه از درجه 1، بالاخص جمله ثابت وجود ندارد. اگر هر عبارت سمت راست را فقط درجه 1 در نظر بگیریم و محدودیت همگن را حذف کنیم، یک عدد ثابت نیز ممکن است به هر جمله افزوده شود. این گونه تابع ها را مستوی می نامیم:

$$\begin{cases} y_1 = a_{10} + a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ y_m = a_{m0} + a_{m1}x_1 + \dots + a_{mn}x_n \end{cases} \quad (6)$$

مقدار تابع مستوی (۶) از انتقال مقدار تابع (۵) با $-m$ تایی ثابت (a_1, \dots, a_m) به دست می آید و نیازی به بررسی جداگانه تابع های مستوی نیست.

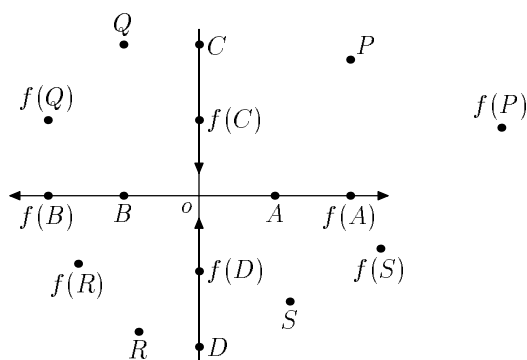
(۱۱-۴) چند مثال

(۱۱-۴-۱) هر تابع خطی $f: \mathbb{R} \rightarrow \mathbb{R}$ به شکل $f(x) = mx$ است که در آن $m \in \mathbb{R}$ یک عدد حقیقی داده شده است. تابع های مستوی $g: \mathbb{R} \rightarrow \mathbb{R}$ به شکل $g(x) = mx + b$ ، m و b اعداد حقیقی داده شده، می باشند.

(۱۱-۴-۲) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ را به صورت زیر تعریف می کنیم

$$f(x_1, x_2) = (2x_1, \frac{1}{4}x_2)$$

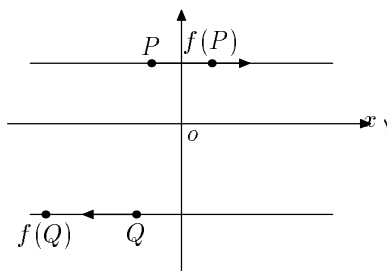
هر دو عبارت $2x_1$ و $\frac{1}{4}x_2$ درجه ۱ همگن نسبت به (x_1, x_2) هستند، پس f خطی است. عملکرد این تابع بدین صورت است که فاصله هر نقطه را از محور x_2 به دو برابر افزایش می دهد و فاصله آن از محور x_1 را نصف می کند. محور x_1 با ضریب تجانس ۲ به خود آن نگاشته می شود (انبساط با ضریب ۲)، و محور x_2 با ضریب تجانس $\frac{1}{4}$ روی خود منقبض می شود. (شکل ۱)



(۱۱-۴-۲) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ را به صورت زیر تعریف می کنیم:

$$f(x_1, x_2) = (x_1 + x_2, x_2)$$

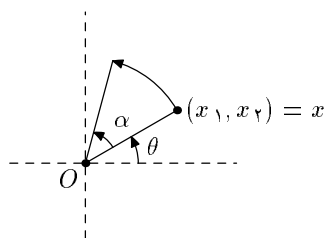
f خطی است. توجه کنید که هر خط راست افقی $x_2 = c$ به خود آن نگاشته می شود ولی هر نقطه روی این خط به اندازه مقدار c انتقال افقی می یابد. بدین ترتیب خطوط راست افقی بالای محور x_1 به طرف راست و خطوط افقی پایین محور x_1 به سمت چپ روی خود می لغزند. نقاط محور x_1 سر جای خود ثابت می مانند.



(۱۱-۴-۳) (دوران حول \underline{e} در \mathbb{R}^2) دوران حول \underline{e} در \mathbb{R}^2 با زاویه α را در نظر بگیرید. اگر

$x = (x_1, x_2)$ نقطه ای $\underline{e} \neq$ در صفحه باشد، نمایش آن به صورت قطبی در نظر بگیرید:

$$x_2 = |x| \sin \theta, \quad x_1 = |x| \cos \theta$$



اگر دوران زاویه α حول \underline{e} را به $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ نمایش دهیم، نمایش قطبی $f(x_1, x_2)$ می شود:

$$\begin{aligned} f(x_1, x_2) &= (|x| \cos(\theta + \alpha), |x| \sin(\theta + \alpha)) \\ &= (|x| \cos \theta \cos \alpha - |x| \sin \theta \sin \alpha, |x| \sin \theta \cos \alpha + |x| \cos \theta \sin \alpha) \\ &= ((\cos \alpha)x_1 - (\sin \alpha)x_2, (\sin \alpha)x_1 + (\cos \alpha)x_2) \end{aligned}$$

هر یک از دو مؤلفه f یک تابع درجه یک همگن نسبت به (x_1, x_2) است، پس f خطی است.

(۱۱-۴-۴) (تصویر قائم روی یک زیرفضای مختصاتی) فرض کنید $n > m$ و $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ را

به صورت زیر تعریف کنید

$$f(x_1, \dots, x_n) = (x_1, \dots, x_m)$$

این تابع خطی است و عملکرد آن این است که نقطه (x_1, \dots, x_n) را به (x_1, \dots, x_m) ، متشکل از m

مؤلفه اول آن، می‌نگارد. مثلاً برای $n = 3$ و $m = 2$ ، $f(x, y, z) = (x, y)$ تصویر قائم (x, y, z) روی صفحه xy است.

(۱۱-۴-۵) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ را به صورت زیر تعریف کنید:

$$f(x_1, x_2) = (x_1, x_1 - x_2, 2x_1 + x_2, -x_2)$$

چون هر مؤلفه f یک عبارت درجه یک همگن نسبت به (x_1, x_2) است، این تابع خطی است.

(۱۱-۴-۶) $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ به صورت زیر تعریف می‌شود:

$$f(x_1, x_2, x_3) = (x_1 - x_2 + x_3, x_1 x_2)$$

این تابع خطی نسبت زیرا مؤلفه دوم f ، یعنی $x_1 x_2$ از درجه یک نیست.

اکنون ارتباط میان دو بحث این جلسه را بیان می‌کنیم. توجه کنید که می‌توان (۵) را به صورت زیر

نوشت:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (7)$$

یا به طور موجز:

$$|y\rangle = A|x\rangle \quad (8)$$

یعنی محاسبه مقدار یک تابع خطی بدین صورت حاصل می شود که ضرایب a_{ij} در (۵) را در یک ماتریس $m \times n$ قرار می دهیم و حاصل ضرب این ماتریس را با ستون $|x\rangle$ محاسبه می کنیم. بدین ترتیب یک تناظر یک به یک میان تابع های خطی $\mathbb{R}^n \rightarrow \mathbb{R}^m$ و ماتریس های $m \times n$ ایجاد می شود.

برای مثال های خطی بالا، ماتریس های مربوط به ترتیب از بالا به پایین عبارتند از:

$$[m]$$

$$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

$$\begin{bmatrix} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}$$

با نگاه کردن به ماتریس متناظر به یک تابع خطی می توان عمل تابع خطی بر اعضای پایه متداول، یعنی e_1, \dots, e_n را فوراً دریافت. توجه کنید که اگر ماتریس $A = [a_{ij}]$ از سمت چپ در ستون $|e_j\rangle$ ضرب شود، ستون j -ام ماتریس $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ به دست می آید. بدین ترتیب، برای نوشتن ماتریس مربوط به یک تابع خطی f ، کافی است $f(e_1), \dots, f(e_n)$ را محاسبه کنیم و نتایج حاصل را به ترتیب در ستون های اول تا n ام درج کنیم.

نگاشت‌های خطی (۲)

مثال‌هایی که از تابع‌های خطی در جلسه گذشته ارائه کردیم از تنوع قابل ملاحظه‌ای برخوردار بود و ممکن است این احساس را ایجاد کرده باشد که وجه تشابه چندانی میان تابع‌های خطی موجود نیست. در چند جلسه آینده، برعکس، خواهیم دید که تابع‌های خطی از بعضی وجوه شباهت‌های فوق‌العاده‌ای به هم دارند. بررسی این خواص دست‌آوردهای مهمی برای ما خواهد داشت. یکی از این دستاوردها امکان بحث کامل در مورد وجود و نظام جواب‌های یک دستگاه m معادله n مجهولی درجه اول است که به آن خواهیم پرداخت. دستاوردی دیگر زمینه‌سازی برای بررسی توابع غیرخطی در باقیمانده درس است که در آنجا، آنچه در چند جلسه آینده در مورد تابع‌های خطی خواهیم دید رهنمودی مهم خواهد بود. در این جلسه زمینه را برای مطالعه تابع‌های خطی فراهم می‌کنیم. کار عمده ما بازبینی بعضی تعاریف و مفاهیمی خواهد بود که در جلسات قبل مطرح شده‌اند و اکنون آنها را به صورتی قابل استفاده در می‌آوریم.

(۱-۱۲) گزاره. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ خطی است در صورتی که واجد دو شرط زیر باشد:

الف) برای هر x و x' در \mathbb{R}^n ، $f(x + x') = f(x) + f(x')$.

ب) برای هر x در \mathbb{R}^n و هر عدد حقیقی r ، $f(rx) = rf(x)$.

اثبات. اگر $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ خطی باشد، ماتریس متناظر با آن، A ، را در نظر می‌گیریم. عملکرد f روی عناصر x از \mathbb{R}^n به صورت:

$$|f(x)\rangle = A|x\rangle \quad (1)$$

است. بنابراین:

$$\begin{aligned} |f(x+x')\rangle &= A|x+x'\rangle \\ &= A|x\rangle + A|x'\rangle \quad \text{قانون پخشی ضرب ماتریسی} \\ &= |f(x)\rangle + |f(x')\rangle \end{aligned}$$

پس (الف) برقرار است. همین طور:

$$\begin{aligned} |f(rx)\rangle &= A|rx\rangle \\ &= r(A|x\rangle) \end{aligned}$$

و (ب) برقرار است. بالعکس فرض کنید شرطهای (الف) و (ب) برای تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ برقرارند. باید نشان دهیم ماتریسی $A, m \times n$ وجود دارد که برای هر $x \in \mathbb{R}^n$ داریم $|f(x)\rangle = A|x\rangle$. از بحث پایان جلسه قبل به یاد می‌آوریم که ستون‌های ماتریس یک تابع خطی به ترتیب از مؤلفه‌های $f(e_1)$ تا $f(e_n)$ تشکیل می‌شوند، بنابراین نامزد واضحی برای ماتریس مورد نظر A وجود دارد که باید در مورد آن ادعای $|f(x)\rangle = A|x\rangle$ را ثابت کرد. ماتریس A را این گونه در نظر می‌گیریم که ستون j -ام آن، برای $j = 1, \dots, n$ برابر $|f(e_j)\rangle$ باشد. برای $x = (x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n$ داریم:

$$\begin{aligned} |f(x)\rangle &= |f(x_1 e_1 + \dots + x_n e_n)\rangle \\ &= |x_1 f(e_1) + \dots + x_n f(e_n)\rangle \quad \text{طبق (الف) و (ب)} \\ &= |x_1 f(e_1)\rangle + \dots + |x_n f(e_n)\rangle \\ &= x_1 |f(e_1)\rangle + \dots + x_n |f(e_n)\rangle \\ &= x_1 (A|e_1\rangle) + \dots + x_n (A|e_n\rangle) \\ &= A|x_1 e_1\rangle + \dots + A|x_n e_n\rangle \\ &= A|x_1 e_1 + \dots + x_n e_n\rangle \quad \text{قانون پخشی ضرب ماتریسی} \\ &= A|x\rangle \end{aligned}$$

□ و حکم به اثبات می‌رسد.
 بدین ترتیب از این پس برای استفاده از خطی بودن یک تابع به کارگیری دو خاصیت (الف) و (ب) بالا کافی است و نیازی به نوشتن صریح ماتریس مربوط نیست.
 در جلسه قبل ضرب ماتریسی را تعریف کردیم. این تعریف در نظر اول دور از ذهن و فاقد انگیزه مشخص به نظر می‌رسد، ولی گزاره زیر نشان می‌دهد که حاصل ضرب دو ماتریس در واقع بیانگر ترکیب تابع‌های خطی متناظر است.

(۱۲-۲) گزاره. تابع‌های خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ با ماتریس A و $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ با ماتریس B داده شده‌اند. در این صورت ماتریس $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^p$ برابر BA است.

اثبات. مقدماً توجه کنید که $g \circ f$ در واقع یک تابع خطی است که این مطلب را می‌توان با استفاده از گزاره ۱۲-۱ ملاحظه کرد:

$$\begin{aligned} (g \circ f)(x + x') &= g(f(x + x')) \\ &= g(f(x) + f(x')) && \text{چون } f \text{ خطی است} \\ &= g(f(x)) + g(f(x')) && \text{چون } g \text{ خطی است} \\ &= (g \circ f)(x) + (g \circ f)(x') \end{aligned}$$

به همین روش

$$\begin{aligned} (g \circ f)(rx) &= g(f(rx)) \\ &= g(rf(x)) && \text{چون } f \text{ خطی است} \\ &= rg(f(x)) && \text{چون } g \text{ خطی است} \\ &= r(g \circ f)(x) \end{aligned}$$

حال به اثبات ادعای گزاره می پردازیم. ماتریس $g \circ f$ را به C نمایش می دهیم. داریم:

$$\begin{aligned} C|x\rangle &= |(g \circ f)(x)\rangle \\ &= |g(f(x))\rangle \\ &= B|f(x)\rangle \\ &= B|(A|x)\rangle \\ &= BA|x\rangle \end{aligned}$$

بنابر ویژگی شرکت پذیری ضرب ماتریسی

پس $C|x\rangle = BA|x\rangle$. اگر حاصل ضرب دو ماتریس $p \times n$ در هر ستون n تایی برابر باشد، دو ماتریس برابرند زیرا که با گرفتن $x = e_j, j = 1, \dots, n$ ، کلیه ستون های ماتریس ها به دست می آیند. □
 بالاخره ضابطه ساده ای برای تحقیق کردن این مطلب که آیا یک زیرمجموعه E از \mathbb{R}^n یک زیرفضای خطی است ارائه می کنیم.

(۱۲-۳) گزاره. زیرمجموعه ناتهی E از \mathbb{R}^n یک زیرفضای خطی \mathbb{R}^n است اگر و تنها اگر دو شرط زیر برقرار باشند:

(الف) هرگاه x و x' در E باشند، آنگاه $x + x'$ نیز در E است.

(ب) هرگاه x در E باشد و $r \in \mathbb{R}$ آنگاه rx نیز در E است.

اثبات. نخست نشان می دهیم هر زیرفضای خطی E از \mathbb{R}^n از دو ویژگی فوق برخوردار است. زیرفضای خطی E را به شکل $\langle A_1, \dots, A_k \rangle$ ، یعنی مجموعه ترکیب های خطی $\{A_1, \dots, A_k\}$ ، که یک مجموعه مستقل خطی عناصر E است، در نظر می گیریم. پس x و x' به شکل

$$x = t_1 A_1 + \dots + t_k A_k \quad \text{و} \quad x' = t'_1 A_1 + \dots + t'_k A_k$$

بنابراین داریم:

$$x + x' = (t_1 + t'_1)A_1 + \dots + (t_k + t'_k)A_k$$

و $x, x' \in E$ همین طور:

$$r(t_1 A_1 + \dots + t_k A_k) = (rt_1)A_1 + \dots + (rt_k)A_k$$

و $rx \in E$

بالعکس فرض کنید زیرمجموعه ناتهی E از \mathbb{R}^n واجد شرطهای (الف) و (ب) باشد، نشان می‌دهیم E یک زیرفضای خطی است، یعنی برابر مجموعه ترکیب‌های خطی یک مجموعه مستقل خطی می‌باشد. قطعاً داریم $\underline{0} \in E$ زیرا که E ناتهی است یعنی عنصری x در E هست، پس طبق (ب) $\underline{0} = \underline{0} \cdot x$ در E قرار دارد.

اگر E عضو دیگری نداشته باشد، یعنی $E = \{\underline{0}\}$ که طبق قرارداد E یگانه زیرفضای خطی صفر بعدی \mathbb{R}^n است. اگر عضو دیگری $\underline{0} \neq A_1$ در E باشد، طبق (ب)، همه مضارب A_1 ، یعنی rA_1 ها، $r \in \mathbb{R}$ در E هستند. حال اگر عضو دیگری در E نباشد، یعنی $E = \langle A_1 \rangle$ ، E یک زیرفضای خطی یک بعدی است. در غیراین صورت عنصری A_2 در E وجود دارد که مضرب A_1 نیست، و در نتیجه $\{A_1, A_2\}$ مستقل خطی است. طبق (الف) و (ب)، همه ترکیب‌های خطی A_1 و A_2 در E قرار دارند، یعنی $\langle A_1, A_2 \rangle \subset E$.

اگر $\langle A_1, A_2 \rangle$ همه E باشد که حکم ثابت شده است، و گرنه، عنصری A_3 در E وجود دارد که $\{A_1, A_2, A_3\}$ مستقل خطی است. استدلال بالا را مجدداً به کار می‌گیریم. همه ترکیب‌های خطی A_1, A_2, A_3 طبق (الف) و (ب)، باید در E باشند. در صورتی که $\langle A_1, A_2, A_3 \rangle$ همه E نباشد، عنصری A_4 انتخاب می‌کنیم و غیره. این استدلال باید در حداکثر n گام به نتیجه رسد زیرا در \mathbb{R}^n یک مجموعه مستقل خطی نمی‌تواند بیش از n عضو داشته باشد. بنابراین اگر E فقط از $\underline{0}$ تشکیل نشده باشد، در گامی $k, 1 \leq k \leq n$ ، به این نتیجه می‌رسیم که $\langle A_1, \dots, A_k \rangle = E$ و حکم به اثبات می‌رسد. □

توجه کنید که این گزاره نوعی تایید برقرار بودن ایده شهودی "مسطح بودن" زیرفضاهای خطی است. اگر نقطه‌ای x در E باشد خط گذرا از $\underline{0}$ و x به تمامی در E واقع است، و اگر x و x' در E باشند، صفحه گذرا از $\underline{0}$ ، x و x' در E واقع است. طبق قسمت دوم گزاره، اگر این دو شرط برای E برقرار باشند می‌توان حکم کرد که E متشکل از همه ترکیب‌های خطی k عضو \mathbb{R}^n است.

نگاشت‌های خطی (۳)

با مقدمات جلسه قبل اکنون آماده هستیم بررسی عمومی توابع خطی

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (1)$$

را آغاز کنیم. آنچه در زیر خواهد آمد در رابطه با تابع خطی (۱) خواهد بود.

(۱-۱۳) تحت اثر f ، $\underline{o} \in \mathbb{R}^n$ به $\underline{o} \in \mathbb{R}^m$ نگاشته می‌شود.

اثبات. می‌نویسیم $\underline{o} = \circ \cdot x$ که x عنصری دلخواه از \mathbb{R}^n است. داریم:

$$f(\underline{o}) = f(\circ \cdot x) = \circ \cdot f(x) = \underline{o}$$

(۲-۱۳) تحت اثر f ، تصویر هر زیرفضای خطی (به ترتیب هر زیرفضای مستوی) یک زیرفضای

خطی (به ترتیب یک زیرفضای مستوی) است.

اثبات. نخست حالت یک زیرفضای خطی E را در نظر بگیرید، باید ثابت کنیم

$f(E) = \{f(x) \mid x \in E\}$ یک زیرفضای خطی \mathbb{R}^m است. با استفاده از ضابطه (۱-۱۲) جلسه قبل

عمل می‌کنیم. اگر y و y' دو عنصر $f(E)$ باشند، داریم $y = f(x)$ و $y' = f(x')$ برای x و x' مناسب در

E پس

$$y + y = f(x) + f(x') = f(x + x')$$

چون $x, x' \in E$ طبق ضابطه (۱۲-۳) جلسه قبل، $x + x' \in E$ پس $y + y' \in f(E)$ به همین ترتیب

اگر $r \in \mathbb{R}$

$$\begin{aligned} ry &= rf(x) \\ &= f(rx) \end{aligned} \quad \text{طبق (۱۲-۱)}$$

و $rx \in E$ طبق ۱۲-۳، پس $ry \in f(E)$ بدین ترتیب ثابت کرده ایم که تصویر یک زیرفضای خطی،

خود یک زیرفضای خطی است. حال هر زیرفضای مستوی را می توان به صورت زیر نوشت:

$$a + E = \{a + x \mid x \in E\}$$

که E یک زیرفضای خطی و a یک عنصر داده شده از E است. بنابراین هر عنصر تصویر به شکل $f(a + x) = f(a) + f(x)$ می باشد. چون $f(x) \in f(E)$ و در بالا دیدیم که $f(E)$ یک زیرفضای خطی است، تصویر زیرفضای مستوی بالا برابر انتقال یافته $f(E)$ توسط عنصر ثابت $f(a)$ است، پس مستوی است. □

توجه کنید که در استدلال بالا، در صورتی که $f(a) \in f(E)$ ، تصویر زیرفضای مستوی یک زیرفضای خطی خواهد بود که تناقض نیست زیرا که زیرفضاهای خطی خود نوعی زیرفضای مستوی هستند. ضمناً اگر دو زیرفضای موازی و همبعد $a + E$ و $b + E$ در \mathbb{R}^n را در نظر بگیریم؛ طبق آنچه در بالا مشاهده شد، تصویر آنها به ترتیب $f(a) + f(E)$ و $f(b) + f(E)$ خواهد بود، یعنی دو زیرفضای موازی (یا منطبق) از \mathbb{R}^m :

(۱۳-۳) اگر E_1 و E_2 دو زیرفضای مستوی همبعد و موازی از \mathbb{R}^n باشند، $f(E_1)$ و $f(E_2)$ همبعد و

موازی (یا منطبق) خواهند بود. □

(۱۸-۴) مثال. تابع خطی $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ را که با ماتریس مشخص $\begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & -1 \\ 5 & 1 & -2 \end{bmatrix}$ می شود در نظر می گیریم. تصویر هر نقطه (x, y, z) تحت f را به (X, Y, Z) نمایش می دهیم:

$$(X, Y, Z) = f(x, y, z). \text{ داریم:}$$

$$X = -x + y, \quad Y = 2x + y - z, \quad Z = 5x + y - 2z \quad (2)$$

توجه کنید که $Z = 2Y - X$ پس همه نقاط \mathbb{R}^3 به یک صفحه گذرا از $z=0$ (زیرفضای خطی دوبعدی) یعنی $x_1 - 2x_2 + x_3 = 0$ نگاشته می شوند. حال اثر f را بر چند صفحه خاص در نظر می گیریم. نخست صفحه مختصاتی xy متشکل از نقاط $(x, y, 0)$ را در نظر بگیرید. طبق (۲) داریم:

$$f(x, y, 0) = (-x + y, 2x + y, 5x + y)$$

ادعا می کنیم هر نقطه صفحه $x_1 - 2x_2 + x_3 = 0$ در تصویر صفحه مختصاتی xy قرار دارد. اگر $(x_1, x_2, -x_1 + 2x_2)$ یک چنین نقطه ای باشد، باید نشان دهیم دستگاه زیر دارای جواب بر حسب x و y است:

$$\begin{cases} -x + y = x_1 \\ 2x + y = x_2 \\ 5x + y = -x_1 + 2x_2 \end{cases}$$

دو معادله اول جواب منحصر به فرد $x = \frac{-x_1 + x_2}{3}, y = \frac{2x_1 + x_2}{3}$ را دارند که با جایگزینی ملاحظه می کنیم در معادله سوم نیز صدق می کند. پس در واقع صفحه مختصاتی xy به طور یک به یک بر صفحه تصویر همه \mathbb{R}^3 نگاشته می شود. محاسبه مشابه با صفحات مختصاتی yz و zx نشان خواهد داد که هر یک از این صفحات نیز به طور یک به یک بر $f(\mathbb{R}^3)$ نگاشته می شود. حال تصویر صفحه $x_1 + 2x_2 - x_3 = 0$ تحت f بررسی می کنیم. نقاط این صفحه به شکل $(x, y, x + 2y)$ هستند. طبق (۲) داریم:

$$f(x, y, x + 2y) = (-x + y, x - y, 3x - 3y)$$

ملاحظه می کنیم که هر نقطه تصویر مضربی از سه تایی $(-1, 1, 3)$ است، یعنی تصویر، خط راست

$$\frac{x}{-1} = \frac{y}{1} = \frac{z}{3}$$

می باشد. بدین ترتیب تصویر صفحاتی که تاکنون بررسی کردیم، بعضی یک صفحه (دوبعدی) و بعضی یک خط راست (یک بعدی) می شوند. سوآلی که مطرح می شود این است که چه قانونی بر تعیین بعد تصویر حاکم است؟ بررسی های بعدی این موضوع را روشن خواهد ساخت.

برای تابع خطی (۱)، هسته f را به صورت زیر تعریف می کنیم و به $\ker(f)$ نمایش می دهیم:

$$\ker(f) = \{x \in \mathbb{R}^n \mid f(x) = \underline{0}\} \quad (3)$$

(۱۳-۴) $\ker(f)$ یک زیرفضای خطی \mathbb{R}^n است.

اثبات. اگر x و x' در $\ker(f)$ باشند، باید نشان دهیم $x + x'$ در $\ker(f)$ است، و اگر به علاوه $r \in \mathbb{R}$ ، باید نشان دهیم rx نیز در $\ker(f)$ است:

$$f(x + x') = f(x) + f(x') \quad \text{طبق ۱-۱۲}$$

$$= \underline{0} + \underline{0} = \underline{0}$$

$$f(rx) = rf(x) \quad \text{طبق ۱-۱۲}$$

$$= r\underline{0} = \underline{0}$$

و حکم به اثبات می رسد. \square

حال طبق (۱۳-۳)، اگر L از انتقال هسته f با یک عضو ثابت \mathbb{R}^n به دست آید، تصویر L تحت f موازی و همبعد تصویر هسته تحت f است، ولی هسته به تک نقطه $\{0\}$ نگاشته می شود، بنابراین تصویر هر L یک تک نقطه خواهد بود:

(۱۳-۵) اگر K هسته f باشد و $a \in \mathbb{R}^n$ ، تصویر $a + K$ تحت f تک نقطه $\{f(a)\}$ است. بالعکس، اگر b نقطه‌ای در \mathbb{R}^m باشد، $f^{-1}(b)$ (= مجموعه تراز منسوب به b = مجموعه نقاطی که تحت f به b نگاشته می‌شوند) یا تهی است یا یک زیرفضای مستوی همبند و موازی K است.

اثبات. قسمت اول حکم را در بالا توضیح دادیم و برای قسمت دوم فرض کنید $f^{-1}(b)$ تهی نباشد، آنگاه عنصری a در \mathbb{R}^n وجود دارد که $f(a) = b$. نشان می‌دهیم:

$$f^{-1}(a) = a + K$$

نخست برای هر عنصر $a + K$ ، یعنی عنصر به شکل $a + x$ که $x \in K$ داریم:

$$\begin{aligned} f(a + x) &= f(a) + f(x) \\ &= f(a) \\ &= b \end{aligned}$$

بالعکس اگر $a' \in f^{-1}(b)$ می‌نویسیم:

$$a' = a + (a' - a) \quad (۴)$$

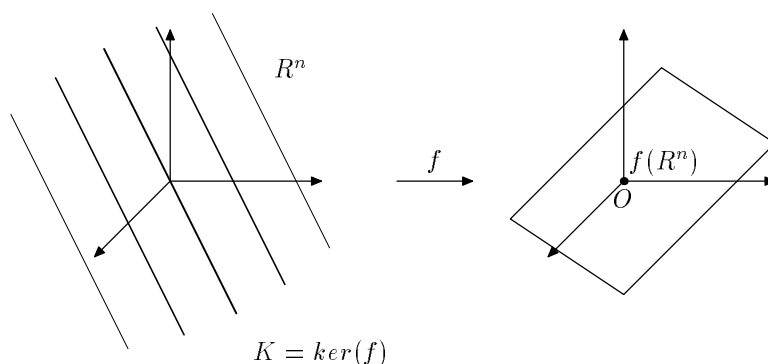
$a' - a$ در هسته است زیرا که

$$\begin{aligned} f(a' - a) &= f(a') - f(a) \\ &= b - b \\ &= \underline{0} \end{aligned}$$

بنابراین (۴) نشان می‌دهد که هر عضو $f^{-1}(b)$ مجموع a با یک عنصر هسته است و اثبات حکم کامل می‌شود. \square

نتیجه‌های اخیر تصویری ساده و به لحاظی کامل از عملکرد یک تابع خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ارائه می‌کند. طبق (۱۳-۲)، کل \mathbb{R}^n به یک زیرفضای خطی \mathbb{R}^m نگاشته می‌شود، هسته به نقطه $\underline{0}$ و

زیرفضاهای موازی و همبند هسته هر یک به تمامی به یک نقطه تصویر فرستاده می‌شوند. در واقع از هر نقطه \mathbb{R}^n یک زیرفضای منحصر به فرد همبند و موازی K می‌گذرد که همه نقاط آن به یک



شکل ۱

نقطه در $f(\mathbb{R}^n)$ نگاشته می‌شوند. می‌توان این توصیف عملکرد تابع خطی را این گونه تجسم کرد که در اثر عملکرد تابع خطی هر یک از زیرفضاهای مستوی موازی و همبند هسته به یک نقطه در \mathbb{R}^m منقبض می‌شود و حاصل یک زیرفضای خطی \mathbb{R}^m است. به نظر می‌آید که در اثر انقباض این زیرفضاهای مستوی به تک نقطه‌ها، باید تصویر \mathbb{R}^n تحت f ، بعدی برابر $(n - k)$ داشته باشد که k بعد هسته است. در واقع حکم کلی زیر برقرار است:

(۱۳-۶) فرض کنید E یک زیرفضای خطی \mathbb{R}^n باشد، در این صورت:

$$\dim E = \dim f(E) + \dim(E \cap \ker(f)) \quad (۵)$$

(در اینجا \dim مخفف dimension به معنای بُعد است)

اثبات. مقدمتاً توجه کنید که اشتراک هر دو زیرفضای خطی از \mathbb{R}^n یک زیرفضای خطی است، بنابراین $E \cap \ker(f)$ یک زیرفضای خطی است و بُعد آن معنی دارد. برای مشاهده این امر فرض کنید E_1 و E_2 دو زیرفضای خطی \mathbb{R}^n باشند. اگر x و x' هر دو در $E_1 \cap E_2$ باشند، چون هر دو در E_1

و هر دو در E_2 هستند، داریم $x + x' \in E_1$ و $x + x' \in E_2$ پس $x + x' \in E_1 \cap E_2$. همچنین اگر $x \in E \cap E'$ و $x \in \mathbb{R}$ داریم $rx \in E_1$ و $rx \in E_2$ پس $rx \in E_1 \cap E_2$.

حال به اثبات حکم می پردازیم. فرض کنید $\dim(E \cap \ker(f)) = l$. در این صورت مجموعه مستقل خطی $\{A_1, \dots, A_l\}$ از عناصر $E \cap \ker(f)$ وجود دارد که پایه برای این زیرفضاست، یعنی $E \cap \ker(f) = \langle A_1, \dots, A_l \rangle$. اگر $E \cap \ker(f) = E$ آنگاه لزوماً $E \subset \ker(f)$ ، همه نقاط E به صفر نگاشته می شوند، $\dim f(E) = 0$ ، و فرمول (5) برقرار است. اگر $E \cap \ker(f)$ همه E نباشد، عنصری A_{l+1} در E وجود دارد که ترکیبی خطی از A_1, \dots, A_l نیست. مجموعه $\{A_1, \dots, A_{l+1}\}$ مستقل خطی است (دقیقاً چرا؟). اگر $\langle A_1, \dots, A_{l+1} \rangle = E$ داریم $\dim E = l + 1$ و $f(E)$ از مضارب $f(A_{l+1})$ تشکیل شده است زیرا که

$$\begin{aligned} f(t_1 A_1 + \dots + t_l A_l + t_{l+1} A_{l+1}) &= \sum_{i=1}^l t_i f(A_i) + t_{l+1} f(A_{l+1}) \\ &= t_{l+1} f(A_{l+1}) \end{aligned}$$

از طرفی دیگر $f(A_{l+1}) \neq 0$ چون $A_{l+1} \notin \ker(f)$. بنابراین $\dim f(E) = 1$ و مجدداً فرمول (5) برقرار است. حال اگر $\langle A_1, \dots, A_{l+1} \rangle$ همه E نباشد، عنصری A_{l+2} از E به مجموعه $\{A_1, \dots, A_{l+1}\}$ می افزاییم که ترکیبی خطی از A_1, \dots, A_{l+1} نباشد و استدلال بالا را تکرار می کنیم. به طور کلی، اگر $\dim E = k$ ، در $k - l$ گام به مجموعه مستقل خطی k عضوی $\{A_1, \dots, A_l, A_{l+1}, \dots, A_k\}$ می رسیم که لزوماً یک پایه برای E است. حال بعد $f(E)$ را محاسبه می کنیم. برای هر عضو x از E داریم:

$$\begin{aligned} x &= t_1 A_1 + \dots + t_l A_l + t_{l+1} A_{l+1} + \dots + t_k A_k \\ f(x) &= t_1 f(A_1) + \dots + t_l f(A_l) + t_{l+1} f(A_{l+1}) + \dots + t_k f(A_k) \quad \text{بنابر ۱-۱۲} \\ &= t_{l+1} f(A_{l+1}) + \dots + t_k f(A_k) \quad (\text{چون } A_i \in \ker(f) \text{ برای } i = 1, \dots, l) \end{aligned}$$

بنابراین هر عضو تصویر E یک ترکیب خطی از $f(A_{l+1}), \dots, f(A_k)$ است. اگر ثابت کنیم مجموعه $\{f(A_{l+1}), \dots, f(A_k)\}$ مستقل خطی است، نتیجه می شود که $\dim(f(E)) = k - l$ و رابطه (5) به

اثبات می‌رسد. برای اثبات این واقعیت فرض کنید مجموعه فوق مستقل خطی نباشد، یعنی اعداد حقیقی c_{l+1}, \dots, c_k وجود داشته باشند که:

$$c_{l+1}f(A_{l+1}) + \dots + c_k f(A_k) = \underline{0}$$

$$f(c_{l+1}A_{l+1} + \dots + c_k A_k) = \underline{0}$$

این نشان می‌دهد که $c_{l+1}A_{l+1} + \dots + c_k A_k$ عضوی از هسته f است و در ضمن چون $A_{l+1}, \dots, A_k \in E$ ، $c_{l+1}A_{l+1} + \dots + c_k A_k \in E \cap \ker f$ است. بنابراین اعداد حقیقی c_1, \dots, c_l وجود دارند که:

$$c_{l+1}A_{l+1} + \dots + c_k A_k = c_1 A_1 + \dots + c_l A_l$$

$$c_1 A_1 + \dots + c_l A_l - c_{l+1}A_{l+1} - \dots - c_k A_k = \underline{0}$$

با توجه به استقلال خطی $\{A_1, \dots, A_k\}$ همه ضرایب باید صفر باشند، بالاخص $c_{l+1} = \dots = c_k = 0$. و اثبات استقلال خطی $\{f(A_{l+1}), \dots, f(A_k)\}$ و حکم (۵) به اتمام می‌رسد. □

با توجه به ۱۳-۶، اکنون می‌توانیم با نگاهی مجدد به مثال ۱۸-۴، این مطلب که تصاویر صفحات گوناگون تحت تابع خطی آن مثال، بعضی دوبعدی و بعضی یک بعدی بودند به طور کامل تجزیه و تحلیل کنیم. برای این کار نخست هسته f در مثال ۱۸-۴ را پیدا می‌کنیم. باید داشته باشیم:

$$A|x\rangle = |\underline{0}\rangle$$

یا $(-x_1 + x_2, 2x_1 + x_2 - x_3, 5x_1 + x_2 - 2x_3) = (0, 0, 0)$. بنابراین $x_2 = x_1$ و $x_3 = 3x_1$. نتیجه این که هسته f عبارت است از خط راست:

$$\frac{x_1}{1} = \frac{x_2}{1} = \frac{x_3}{3} \quad (6)$$

حال اگر E یک زیرفضای خطی دوبعدی \mathbb{R}^3 باشد، بعد تصویر E طبق ۱۸-۶ برابر $l-2$ است که l بعد اشتراک E با هسته است. دو امکان وجود دارد، یا اشتراک E و هسته فقط $\{\underline{0}\}$ است که در این

صورت $l = 0$ و تصویر E دوبعدی است، و یا اشتراک E و هسته یک بعدی است، یعنی هسته به تمامی در E قرار دارد، که در این صورت $l = 1$ و تصویر یک بعدی است. در مثال ۱۸-۴، هر یک از صفحات مختصاتی با هسته (یعنی خط راست (۶)) فقط در اشتراک دارند، بنابراین تصویر هر یک، دوبعدی است. ولی صفحه $0 = x_1 + 2x_2 - x_3$ شامل خط راست (۶) است، یعنی برای آن $l = 1$ ، و در نتیجه تصویر آن تحت f ، یک بعدی می باشد.

نگاشت‌های خطی (۴)

آنچه در جلسه قبل دیدیم تصویر به نسبت جامعی از عملکرد تابع‌های خطی روی زیرفضاهای مستوی می‌دهد. گزاره‌های ذکر شده اکثراً به طور مستقیم به زیرفضاهای خطی مربوط می‌شدند ولی نتیجه‌گیری درست در مورد زیرفضاهای مستوی دشوار نیست. مثلاً آخرین گزاره:

$$\dim(E) = \dim(f(E)) + \dim(E \cap \ker f) \quad (1)$$

برای زیرفضاهای خطی E برقرار است. اگر E' یک زیرفضای مستوی همبند و موازی E باشد، یعنی $E' = p + E$ ، دیگر $E' \cap \ker f$ مطرح نیست، بلکه می‌دانیم که $\dim f(E') = \dim f(E)$ (تصویر زیرفضاهای مستوی موازی همبند، موازی و همبند است)، پس

$$\dim(E') = \dim f(E') + \dim(E \cap \ker f) \quad (2)$$

یعنی در اینجا بعد اشتراک انتقال یافته E' به $\underline{0}$ با هسته، تعیین کننده افت بعد E' در اثر عملکرد f است. با توجه به این تذکر، به مثال جلسه قبل باز می‌گردیم.

مثال. تابع خطی $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ را با ماتریس $\begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & -1 \\ -5 & 1 & -2 \end{bmatrix}$ را بررسی می‌کنیم. هسته این تابع خط راست $x_1 = x_2 = x_3$ است. دیدیم که زیرفضاهای خطی دوبعدی شامل هسته به یک خط راست و زیرفضاهای خطی دوبعدی که فقط در $\underline{0}$ با هسته اشتراک دارند به تمامی صفحه $x_1 - 2x_2 + x_3 = 0$ نگاشته می‌شوند. حال اگر F یک صفحه دلخواه در \mathbb{R}^3 باشد (زیرفضای مستوی

دو بعدی)، برای تعیین بعد تصویر آن باید انتقال یافته آن به مبدأ، F° را در نظر بگیریم. اگر F° شامل هسته باشد، یعنی F موازی هسته باشد، تصویر F یک خط راست می شود. در غیر این صورت، F° با هسته فقط در \underline{e} مشترک است و تصویر آن دو بعدی می شود. اکنون وضعیت تصویر خطوط راست \mathbb{R}^3 را تحت اثر f بررسی می کنیم. اگر خط راست L به مبدأ منتقل شود تا خط راست L° به دست آید، L° یا برابر هسته است، که در این صورت L یک خط موازی هسته است و یا L° با هسته فقط در \underline{e} مشترک است. در حالت اول تصویر L یک نقطه می شود و در حالت دوم تصویر آن یک خط راست است.

اکنون به ذکر چند نکته تکمیلی در مورد بحث جلسه گذشته می پردازیم. پس تابع خطی

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

مورد نظر است.

(۱۴-۱) شرطی لازم و کافی برای یک به یک بودن f این است که $\ker(f) = \{0\}$. در این حالت لزوماً $n \leq m$.

اثبات. اگر عنصری $x \neq \underline{e}$ در هسته باشد، چون $f(x) = f(\underline{e}) = \underline{e}$ ، f یک به یک نیست. حال فرض کنید $\ker f = \{\underline{e}\}$ و $f(x_1) = f(x_2)$ چون f خطی است نتیجه می گیریم که $f(x_1 - x_2) = \underline{e}$ ، یعنی $x_1 - x_2 \in \ker(f)$ پس طبق فرض $x_1 = x_2$ و یک به یک بودن به اثبات می رسد. در این حالت $\dim(f(\mathbb{R}^n)) = n$ ، و چون $f(\mathbb{R}^n) \subset \mathbb{R}^m$ ، نتیجه می گیریم که $n \leq m$. □

در فرمول (۱)، وقتی $E = \mathbb{R}^n$ داریم:

$$n = \dim(f(\mathbb{R}^n)) + \dim(\ker(f))$$

پوشا بودن f معادل این است که $f(\mathbb{R}^n) = \mathbb{R}^m$ ، یعنی $n = m + \dim(\ker(f))$. بنابراین:

(۱۴-۲) شرطی لازم و کافی برای پوشا بودن f این است که $\dim(\ker(f)) = n - m$. در این حالت

□ لزوماً $n \geq m$.

برای تابع‌های $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ، یعنی در حالت $m = n$ ، از (۱۴-۱) و (۱۴-۲) نتیجه می‌شود که:

□ (۱۴-۳) تابع خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک به یک است اگر و تنها اگر پوشا باشد.

می‌دانیم که شرطی لازم و کافی برای وجود تابع معکوس (وارون) یک به یک و پوشا بودن تابع است. بدین ترتیب وقتی $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک به یک (معادلاً پوشا) باشد، تابعی $f^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ وجود دارد که

$$f^{-1}(f(x)) = x \quad , \quad f(f^{-1}(x)) = x$$

ادعا می‌کنیم که f^{-1} نیز خطی است. فرض کنید $y_1, y_2 \in \mathbb{R}^n$. چون f پوشا است، x_1 و x_2 وجود دارند که $y_1 = f(x_1)$ و $y_2 = f(x_2)$. بنابراین

$$\begin{aligned} f^{-1}(y_1 + y_2) &= f^{-1}(f(x_1) + f(x_2)) \\ &= f^{-1}(f(x_1 + x_2)) \quad \text{چون } f \text{ خطی است} \\ &= x_1 + x_2 \\ &= f^{-1}(y_1) + f^{-1}(y_2) \end{aligned}$$

به همین ترتیب اگر $y = f(x)$ و r یک عدد حقیقی باشد،

$$\begin{aligned} f^{-1}(ry) &= f^{-1}(rf(x)) \\ &= f^{-1}(f(rx)) \quad \text{چون } f \text{ خطی است} \\ &= rx \\ &= rf^{-1}(y) \end{aligned}$$

چون f^{-1} خطی است، به آن ماتریسی B ، $n \times n$ ، منسوب می‌شود. اگر A ماتریس تابع خطی f باشد، رابطه B را با A بررسی می‌کنیم. داریم $f \circ f^{-1} = \mathbb{1}_{\mathbb{R}^n}$ و $f^{-1} \circ f = \mathbb{1}_{\mathbb{R}^n}$ تابع همانی

\mathbb{R}^n است. از آنجا که ماتریس I_n برابر $1_{\mathbb{R}^n}$ است و ترکیب تابع های خطی متناظر با حاصل ضرب ماتریس های مربوط است، داریم:

$$BA = I_n \quad , \quad AB = I_n$$

چنین ماتریس B وارون یا معکوس ماتریس A خوانده می شود و معمولاً آن را به A^{-1} نمایش می دهیم.

مثال زیر نشان می دهد چگونه می توان بعضی مسائل هندسی را به بحث در مورد جواب های یک دستگاه معادله درجه یک تبدیل کرد و سپس از مطالب فوق در حل مسأله یاری جست.

مثال. فرض کنید E_1 و E_2 دو صفحه (زیرفضای مستوی دوبعدی) در \mathbb{R}^4 هستند. نشان می دهیم:
 (الف) اگر $E_1 \cap E_2$ یک تک نقطه باشد و E'_1 یک صفحه موازی E_1 و E'_2 یک صفحه موازی E_2 باشد، آنگاه $E'_1 \cap E'_2$ نیز یک تک نقطه است.

(ب) اگر $E_1 \cap E_2$ یک خط راست باشد، صفحه ای E'_1 موازی E_1 و صفحه ای E'_2 موازی E_2 وجود دارند که $E'_1 \cap E'_2$ تهی است.

فرض کنید E_1 صفحه گذرا از p به موازات A_1 و A_2 باشد، $\{A_1, A_2\}$ مستقل خطی، و E_2 صفحه گذرا از q به موازات B_1 و B_2 و $\{B_1, B_2\}$ مستقل خطی. این که E_1 و E_2 دقیقاً یک نقطه مشترک دارند بدین صورت قابل بیان است که چهارتایی منحصر به فردی (s_1, s_2, t_1, t_2) از اعداد حقیقی وجود دارد که:

$$p + s_1 A_1 + s_2 A_2 = q + t_1 B_1 + t_2 B_2 \quad (3)$$

حال p, q, A_1, A_2, B_1, B_2 همه عناصر \mathbb{R}^4 ، یعنی چهارتایی هستند. می نویسیم
 $A_2 = (a_{12}, a_{22}, a_{32}, a_{42})$ ، $A_1 = (a_{11}, a_{21}, a_{31}, a_{41})$ ، $q = (q_1, \dots, q_4)$ ، $p = (p_1, \dots, p_4)$
 $B_2 = (b_{12}, b_{22}, b_{32}, b_{42})$ ، $B_1 = (b_{11}, b_{21}, b_{31}, b_{41})$ پس اگر مؤلفه های اول تا چهارم دو طرف

(۳) را برابر قرار دهیم حاصل می شود:

$$\begin{cases} p_1 + s_1 a_{11} + s_2 a_{12} = q_1 + t_1 b_{11} + t_2 b_{12} \\ p_2 + s_1 a_{21} + s_2 a_{22} = q_2 + t_1 b_{21} + t_2 b_{22} \\ p_3 + s_1 a_{31} + s_2 a_{32} = q_3 + t_1 b_{31} + t_2 b_{32} \\ p_4 + s_1 a_{41} + s_2 a_{42} = q_4 + t_1 b_{41} + t_2 b_{42} \end{cases} \quad (4)$$

با انتقال جملات دارای s_1, s_2, t_1, t_2 به سمت چپ و سایر جملات به سمت راست، دستگاه زیر حاصل می شود:

$$\begin{bmatrix} a_{11} & a_{12} & -b_{11} & -b_{12} \\ a_{21} & a_{22} & -b_{21} & -b_{22} \\ a_{31} & a_{32} & -b_{31} & -b_{32} \\ a_{41} & a_{42} & -b_{41} & -b_{42} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \\ q_4 - p_4 \end{bmatrix} \quad (5)$$

ماتریس 4×4 سمت چپ یک تابع خطی $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ تعریف می کند. اینکه E_1 و E_2 فقط یک نقطه اشتراک داشته باشند بدین صورت قابل بیان است که یک و تنها یک نقطه \mathbb{R}^4 به نقطه $(q-p)$ نگاشته می شود، مجموعه تراز، $f^{-1}(q-p)$ تک عضوی است. می دانیم که همه مجموعه های تراز ناتهی از انتقال هسته به دست می آیند و با هم موازی هستند. پس در این حالت هسته برابر $\{0\}$ است و در نتیجه f یک به یک و پوشا است. حال اگر به جای E_1 و E_2 ، دو زیرفضای موازی جایگزین کنیم، فقط طرف راست (۵) عوض می شود، ولی از آنجا که f یک به یک و پوشا است، مجدداً جواب منحصر به فردی (s_1, s_2, t_1, t_2) برای دستگاه به دست می آید، یعنی اشتراک دو زیرفضای موازی نیز یک تک نقطه است.

اکنون حالتی را در نظر بگیرید $E_1 \cap E_2$ یک خط راست است. به زبان تابع های خطی، این مطلب را می توان به این صورت بیان کرد که مجموعه تراز $f^{-1}(q-p)$ یک بعدی است. نتیجه می شود که هسته تابع خطی $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ یک بعدی است. بنابراین f یک به یک نیست و طبق (۱۴-۳) پوشا نیست. بنابراین طرف راست (۵) را می توان طوری اختیار کرد که مجموعه تراز منسوب به آن تهی باشد. مثلاً با ثابت نگاه داشتن p و تغییر مناسب q ، می توان به چنین نقطه ای رسید. نتیجه این که با

انتقال موازی مناسب E_1 و E_2 می توان به دو صفحه غیر متقاطع رسید.

مثال بالا به خوبی نشان می دهد که بسیاری سؤال های مربوط به دستگاه های معادلات خطی را اکنون می توان به کمک اطلاعاتی که در مورد تابع های خطی کسب کرده ایم جواب داد. دستگاه کلی m معادله n مجهولی درجه یک را به صورت زیر در نظر می گیریم:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases} \quad (6)$$

یا به طور خلاصه:

$$A|x\rangle = |b\rangle \quad (7)$$

که در آن $A = [a_{ij}]$. دستگاهی که طرف چپ آن مانند بالا و طرف راست از صفر تشکیل شده باشد، دستگاه همگن (یا متجانس) وابسته به (6) یا (7) می نامند:

$$A|x\rangle = |0\rangle \quad (8)$$

جدول زیر رابطه بین جبر حل دستگاه های معادلات بالا و هندسه تابع های خطی را خلاصه می کند:

هندسه	جبر
یافتن مجموعه تراز منسوب b	حل (6) برای b داده شده
ناتهی بودن مجموعه تراز منسوب به b	وجود جواب به ازای b داده شده
پوشا بودن تابع خطی f	وجود جواب به ازای هر b
هسته $\{0\}$	یگانگی جواب در صورت وجود
یافتن هسته	حل دستگاه همگن

چند نمونه از احکام جبری را که اکنون می توان با توسل به هندسه تابع های خطی جواب داد در زیر می آوریم. مثال های دیگری در تمرین ها آمده اند.

(14-4) اگر تعداد معادلات بیش از تعداد مجهول ها باشد، می توان b را طوری اختیار کرد که دستگاه جواب نداشته باشد.

اثبات. این حالت $m > n$ است که طبق ۱۴-۲ تابع خطی مربوط پوشا نیست، یعنی $b \in \mathbb{R}^m$ وجود دارد که $f^{-1}(b)$ تهی است.

(۱۴-۵) اگر $(\bar{x}_1, \dots, \bar{x}_n) = \bar{x}$ یک جواب (۶) باشد، همه جواب‌های (۶) عبارتند از $x + \bar{x}$ که در آن x یک جواب دستگاه همگن متناظر است.

اثبات. داریم $f(\bar{x}) = b$. اگر x یک جواب دستگاه همگن باشد، یعنی x در هسته f است، پس $f(x + \bar{x}) = f(x) + f(\bar{x})$. بالعکس فرض کنید \bar{x} جواب دیگری از (۶) باشد، آنگاه

$$f(\bar{x} - \bar{x}) = f(\bar{x}) - f(\bar{x}) = b - b = \underline{0}$$

پس $\bar{x} - \bar{x}$ در هسته است و می‌توان نوشت $\bar{x} = (\bar{x} - \bar{x}) + \bar{x}$.

حالت مهم $m = n$ (تعداد مجهول = تعداد معادله) را در نظر بگیرید. دیدیم که در این حالت شرایط زیر معادل هستند: f یک به یک، f پوشا، f وارون پذیر، هسته $f = \{0\}$. در این حالت، ماتریس مربوط، A ، نیز وارون پذیر است و با ضرب کردن دو طرف (۷) در A^{-1} از سمت چپ، جواب (منحصر به فرد) دستگاه به دست می‌آید:

$$|x\rangle = A^{-1}|b\rangle \quad (9)$$

حجم و دترمینان (۱)

در بین مفاهیم هندسی که تاکنون در \mathbb{R}^n بررسی کرده‌ایم اثری از مساحت، حجم و تعمیم آنها به نوعی مفهوم برای سنجش "محتوای n -بعدی" نبوده است. برای اقدام به این کار لازم است وجه مشترکی میان طول، مساحت و حجم بیابیم که قابلیت تعمیم به هر بعد را داشته باشد. در اینجا نخست مفهوم ضرب برداری در \mathbb{R}^3 را یادآوری می‌کنیم.

برای $u, v \in \mathbb{R}^3$ ، $u = (u_1, u_2, u_3)$ و $v = (v_1, v_2, v_3)$ حاصل ضرب برداری، $u \times v \in \mathbb{R}^3$ را

به صورت زیر تعریف می‌کنیم:

$$u \times v = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1) \quad (1)$$

(۱-۱۰) ویژگی‌های ابتدایی ضرب خارجی

(۱-۱-۱۰) (ضد تقارن) برای هر $u, v \in \mathbb{R}^3$ داریم

$$u \times v = -v \times u \quad (2)$$

بالاخص

$$u \times u = \underline{0} \quad (3)$$

اثبات. در فرمول (۱) اگر نقش u, v تعویض شود، هر مؤلفه در (۱) ضرب می شود. اگر در (۲) قرار

□ دهیم $u = v$ ، نتیجه می شود که $u \times u = -(u \times u)$ ، پس لزوماً $u \times u = 0$.

(۱۰-۱-۲) (قوانین پخشی) برای هر $u, v, w \in \mathbb{R}^3$ داریم:

$$(u + v) \times w = (u \times w) + (v \times w) \quad , \quad u \times (v + w) = (u \times v) + (u \times w) \quad (4)$$

اثبات. این دو قانون با محاسبه سرراست از (۱) حاصل می شوند. محاسبه به خواننده واگذار

□ می شود.

(۱۰-۱-۳) برای $u, v \in \mathbb{R}^3$ و $r \in \mathbb{R}$ داریم:

$$(ru) \times v = r(u \times v) \quad , \quad u \times (rv) = r(u \times v) \quad (5)$$

□ اثبات. اگر در (۱) یکی از u یا v در r ضرب شود، هر مؤلفه طرف راست در r ضرب می شود.

(۱۰-۱-۴) برای $u, v \in \mathbb{R}^3$ داریم:

$$v \cdot (u \times v) = 0 \quad , \quad u \cdot (u \times v) = 0 \quad (6)$$

اثبات. این نیز یک محاسبه سرراست است:

$$\begin{aligned}
 u \cdot (u \times v) &= (u_1, u_2, u_3) \cdot (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1) \\
 &= u_1(u_2 v_3 - u_3 v_2) + u_2(u_3 v_1 - u_1 v_3) + u_3(u_1 v_2 - u_2 v_1) \\
 &= 0
 \end{aligned}$$

□ و به همین ترتیب اتحاد دیگر ثابت می شود.

(۱۰-۱-۵) (رابطه با ضرب داخلی) برای هر $u, v \in \mathbb{R}^3$ داریم:

$$|u \times v|^2 + |u \cdot v|^2 = |u|^2 |v|^2 \quad (7)$$

اثبات. توجه کنید که مقصود از $|u \times v|$ طول سه تایی $u \times v$ است و مقصود از $|u \cdot v|$ قدرمطلق عدد $u \cdot v$. اتحاد (۶) به صورت زیر بسط داده می شود:

$$\begin{aligned} & (u_2 v_3 - u_3 v_2)^2 + (u_3 v_1 - u_1 v_3)^2 + (u_1 v_2 - u_2 v_1)^2 + (u_1 v_1 + u_2 v_2 + u_3 v_3)^2 \\ &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) \end{aligned} \quad (8)$$

□ صحت این اتحاد جبری را می توان با محاسبه تحقیق کرد.

اگر $\{u, v\}$ وابسته خطی باشد، مثلاً $v = ru$ ، برای عدد حقیقی r ، آنگاه از (۵) و (۲) نتیجه می شود $u \times v = 0$. البته این مطلب از (۷) نیز نتیجه می شود زیرا که اگر $\{u, v\}$ وابسته خطی باشد، آنگاه $|u \cdot v| = |u||v|$. حال فرض کنید $\{u, v\}$ مستقل خطی باشد، بالاخص $u \neq 0$ و $v \neq 0$. در این صورت u, v هر یک، یک نیم خط تعریف می کند (مضارب مثبت این بردار). α را زاویه بین دو نیم خط می گیریم. از آنجا که $u \cdot v = |u||v| \cos \alpha$ ، از (۷) نتیجه می شود که:

$$|u \times v| = |u||v| |\sin \alpha| \quad (9)$$

عبارت طرف راست در واقع مساحت متوازی الاضلاعی است که توسط u, v ایجاد می شود. به طور دقیقتر، مقصود از متوازی الاضلاع ایجاد شده توسط u, v مجموعه زیر است:

$$P(u, v) = \{t_1 u_1 + t_2 v \mid 0 \leq t_1 \leq 1, 0 \leq t_2 \leq 1\}$$

شکل ۱

توجه کنید که ارتفاع وارد بر "قاعده" u ، دارای طول $|v| \sin \alpha$ است، پس $|u||v| \sin \alpha$ فی الواقع مساحت $P(u, v)$ است.

بدین ترتیب تاکنون به این نتایج در مورد $u \times v$ رسیده‌ایم: اگر $\{u, v\}$ وابسته خطی باشد، $u \times v$ برابر $\underline{0}$ است و اگر $\{u, v\}$ مستقل خطی، آنگاه $u \times v$ بر صفحه u, v عمود است (بنابر (۶)) و طول آن برابر $|u||v| \sin \alpha$ می‌باشد. در واقع حالت وابستگی خطی را می‌توان حالت خاص حکم دوم دانست زیرا که در آن صورت $\sin \alpha = 0$ و چون $u \times v = \underline{0}$ ، این حاصل ضرب بر هر صفحه شامل u, v عمود است (به مفهوم صفر شدن ضرب داخلی). حال در حالت استقلال خطی، دقیقاً دوسه‌تایی با طول $|u||v| \sin \alpha$ وجود دارند که بر صفحه u, v عمودند. برای مشخص کردن این که کدام یک برابر $u \times v$ است از مفهوم "دترمینان" کمک می‌گیریم.

اگر $u = (u_1, u_2)$ و $v = (v_1, v_2)$ دو عنصر \mathbb{R}^2 باشند، مؤلفه‌های u, v را به ترتیب در ستون‌های

اول و دوم یک ماتریس 2×2 به صورت زیر قرار می‌دهیم

$$M = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}$$

دترمینان ماتریس M بنابر تعریف برابر $u_1 v_2 - u_2 v_1$ است. معنی هندسی این عبارت را بررسی می‌کنیم. اگر (v_1, v_2) را به اندازه $\frac{\pi}{4}$ در جهت عقربه ساعت دوران دهیم بردار v' به صورت زیر حاصل می‌شود:

$$v' = (v_2, -v_1)$$

شکل ۲

بنابراین داریم:

$$\begin{aligned} u_1 v_2 - u_2 v_1 &= u \cdot v' \\ &= |u||v'| \cos \angle(u, v') \\ &= |u||v| \cos(\alpha - \frac{\pi}{4}) \\ &= |u||v| \sin \alpha \end{aligned}$$

قدرمطلق عبارت سمت راست اندازه مساحت $P(u, v)$ و علامت آن به علامت $\sin \alpha$ وابسته است. اگر $\{u, v\}$ وابسته خطی باشد، $\sin \alpha = 0$ و $\det M$ صفر می شود. ولی چنانچه $\{u, v\}$ مستقل خطی باشد، اگر زاویه α (از نیم خط تعیین شده توسط u به نیم خط تعیین شده توسط v) بین 0 و π باشد، $\sin \alpha > 0$ و $\det M$ مثبت است. ولی اگر زاویه α بین π و 2π باشد، $\sin \alpha < 0$ و $\det M$ منفی است.

شکل ۳

یک زوج مرتب برداری (u, v) در صفحه مانند (الف) در شکل (۳) را "راستگرد" و یک زوج مرتب برداری مانند (ب) در شکل (۳) را "چپگرد" می نامند. تمایز این دو وضعیت را می توان به صورت زیر نیز بیان کرد. چنانچه با گردش کوچکتر از یک نیم صفحه در جهت مثلثاتی از بردار اول به بردار دوم برسیم، زوج مرتب راستگرد خوانده می شود ولی چنانچه با گردش کوچکتر از یک نیم صفحه در جهت عقربه ساعت از بردار اول به بردار دوم برسیم، زوج مرتب چپگرد محسوب می شود.

با توجه به بحث بالا، می توان زوج مرتب (u, v) را راستگرد (به ترتیب چپگرد) تعریف کرد اگر $\det M > 0$ (به ترتیب $\det M < 0$) بنابراین $\det M$ مشخص کننده دو اطلاع زیر است:

(الف) قدرمطلق دترمینان M برابر مساحت متوازی الاضلاع $P(u, v)$ است.

(ب) $\det M = 0$ اگر و تنها اگر $\{u, v\}$ وابسته خطی باشد، $\det M > 0$ (به ترتیب $\det M < 0$) اگر و تنها اگر زوج مرتب (u, v) راستگرد (به ترتیب چپگرد) باشد.

این ملاحظات را می توان به \mathbb{R}^3 تعمیم داد. فرض کنید $u = (u_1, u_2, u_3)$ ، $v = (v_1, v_2, v_3)$ و

$w = (w_1, w_2, w_3)$ عناصر \mathbb{R}^3 باشند. مقصود از متوازی السطوح ایجاد شده توسط u, v و w مجموعه زیر است:

$$P(u, v, w) = \{t_1 u + t_2 v + t_3 w \mid 0 \leq t_i \leq 1, i = 1, 2, 3\} \quad (10)$$

اگر سه تایی $\{u, v, w\}$ وابسته خطی باشد، سه بردار در یک صفحه قرار می گیرند و حجم این متوازی السطوح صفر است. در غیر این صورت، برای به دست آوردن حجم باید مساحت یک قاعده (مثلاً متوازی الاضلاع $P(u, v)$) را در ارتفاع وارد بر این قاعده ضرب کنیم. ارتفاع وارد قاعده $P(u, v)$ برابر طول تصویر قائم w بر امتداد عمود بر صفحه u, v است. امتداد عمود بر صفحه توسط $u \times v$ تعیین می شود، پس

$$\text{طول ارتفاع} = \frac{|w \cdot (u \times v)|}{|u \times v|^2} |u \times v|$$

و در نتیجه چون مساحت قاعده برابر $|u \times v|$ است:

$$\begin{aligned} \text{حجم } P(u, v, w) &= |w \cdot (u \times v)| \\ &= |w_1(u_2 v_3 - u_3 v_2) + w_2(u_3 v_1 - u_1 v_3) + w_3(u_1 v_2 - u_2 v_1)| \end{aligned} \quad (11)$$

حال مانند حالت دو بعدی، مؤلفه های سه بردار u, v و w را به ترتیب در ستون های یک ماتریس 3×3 قرار می دهیم:

$$M = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

طبق تعریف، دترمینان ماتریس 3×3 بالا برابر است با

$$\det M = u_1 v_2 w_3 - u_1 v_3 w_2 + v_1 w_2 u_3 - v_1 w_3 u_2 + w_1 u_2 v_3 - w_1 u_3 v_2 \quad (12)$$

که می توان آن را به صورت های زیر نیز نوشت:

$$\det M = u_1(v_2 w_3 - v_3 w_2) + u_2(v_3 w_1 - v_1 w_3) + u_3(v_1 w_2 - v_2 w_1) \quad (13)$$

$$\det M = v_1(w_2u_3 - w_3u_2) + v_2(w_3u_1 - w_1u_3) + v_3(w_1u_2 - w_2u_1) \quad (14)$$

$$\det M = w_1(u_2v_3 - u_3v_2) + w_2(u_3v_1 - u_1v_3) + w_3(u_1v_2 - u_2v_1) \quad (15)$$

از مقایسه (۱۵) و (۱۱) می بینیم که قدرمطلق دترمینان M همان حجم متوازی السطوح $P(u, v, w)$ است. در مورد علامت، سه تایی مرتب (u, v, w) در \mathbb{R}^3 را راستگرد (به ترتیب چپگرد) می نامیم در صورتی که $\det M > 0$ (به ترتیب $\det M < 0$). توجه کنید که $\det M = 0$ اگر و تنها اگر سه تایی $\{u, v, w\}$ وابسته خطی باشد زیرا که اگر u, v و w در یک صفحه باشند، حاصل ضرب داخلی w با $u \times v$ که بر این صفحه عمود است صفر می شود. به این ترتیب مجدداً $\det M$ مشخص کننده دو اطلاع زیر است:

الف) قدرمطلق دترمینان M برابر حجم متوازی السطوح $P(u, v, w)$ است.

ب) $\det M = 0$ اگر و تنها اگر $\{u, v, w\}$ وابسته خطی باشد، $\det M > 0$ (به ترتیب $\det M < 0$) اگر و تنها اگر سه تایی مرتب (u, v, w) راستگرد (به ترتیب چپگرد) باشد.

ضمناً با توجه به (۱۳) و (۱۴) نیز می توان نوشت:

$$u \cdot (v \times w) = v \cdot (w \times u) = w \cdot (u \times v) = -w \cdot (v \times u) = -v \cdot (u \times w) = -u \cdot (w \times v) \quad (16)$$

قدرمطلق هر یک از این عبارات حجم متوازی السطوح $P(u, v, w)$ به صورت حاصل ضرب یک قاعده در ارتفاع وارد بر آن است.

در اینجا راهی برای تعریف "حجم n -بعدی"، دست کم برای تعمیم متوازی الاضلاع و

متوازی السطوح به نظر می رسد. اگر $u_1 = (u_{11}, u_{21}, \dots, u_{n1})$ ، \dots ، $u_n = (u_{1n}, \dots, u_{nn})$

عنصر \mathbb{R}^n باشند، متوازی السطوح n -بعدی ایجاد شده توسط u_1, \dots, u_n را به صورت زیر تعریف می کنیم:

$$P(u_1, \dots, u_n) = \{t_1 u_1 + \dots + t_n u_n \mid 0 \leq t_i \leq 1, i = 1, \dots, n\} \quad (17)$$

با قرار دادن مؤلفه های u_1 تا u_n در یک ماتریس $n \times n$ داریم:

$$M = \begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix}$$

اگر مفهوم دترمینان یک ماتریس $n \times n$ در دست باشد، قاعدتاً باید n تایی مرتب (u_1, \dots, u_n) را “راستگرد” (به ترتیب “چپگرد”) بنامیم اگر $\det M > 0$ (به ترتیب $\det M < 0$)، و $\det M = 0$ باید معادل وابستگی خطی $\{u_1, \dots, u_n\}$ باشد. همچنین $|\det M|$ را باید “حجم n -بعدی $P(u_1, \dots, u_n)$ ” تعریف کرد. تعمیم دترمینان به ماتریس های $n \times n$ و پیگیری این مفاهیم در جلسه آینده انجام خواهد شد.

به سؤال اولیه ای که بحث بالا را پیش آورد باز می گردیم. اگر $\{u, v\}$ در \mathbb{R}^3 مستقل خطی باشد، می خواستیم بدانیم $u \times v$ کدامیک از دو بردار عمود بر صفحه u, v به طول $|u||v| \sin \alpha$ است. اکنون ادعا می کنیم:

(۱۰-۲) اگر $\{u, v\}$ یک مجموعه مستقل خطی در \mathbb{R}^3 باشد، سه تایی مرتب $(u, v, u \times v)$ راستگرد است.

اثبات. باید نشان دهیم دترمینان ماتریس زیر مثبت است:

$$M = \begin{bmatrix} u_1 & v_1 & u_2 v_3 - u_3 v_2 \\ u_2 & v_2 & u_3 v_1 - u_1 v_3 \\ u_3 & v_3 & u_1 v_2 - u_2 v_1 \end{bmatrix}$$

که طبق (۱۵) می شود:

$$\det M = (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2$$

- اگر $\{u, v\}$ مستقل خطی باشد، دست کم یکی از سه عبارت بالا صفر نیست و $\det M > 0$.
بدین ترتیب $u \times v$ از نظر هندسی به طور منحصر به فرد مشخص می شود.

حجم و دترمینان (۲)

هدف ما در این بخش تعمیم مفاهیم دترمینان و حجم به \mathbb{R}^n است. نخست تعریف هندسی و تعریف

جبری دترمینان ماتریس‌های 2×2 و 3×3 را یادآوری می‌کنیم. فرض کنید

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{یا} \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

ستون‌های ماتریس 2×2 را به ترتیب از چپ به راست به A^1 و A^2 ، و ستون‌های ماتریس 3×3 را به

A^1 ، A^2 و A^3 نمایش می‌دهیم. در \mathbb{R}^2 ، متوازی‌الاضلاع $P(A)$ را به صورت

$$P(A) = \{t_1 A^1 + t_2 A^2 \mid 0 \leq t_1 \leq 1, 0 \leq t_2 \leq 1\}$$

و در \mathbb{R}^3 ، متوازی‌الاضلاع $P(A)$ را به صورت

$$P(A) = \{t_1 A^1 + t_2 A^2 + t_3 A^3 \mid 0 \leq t_1 \leq 1, 0 \leq t_2 \leq 1, 0 \leq t_3 \leq 1\}$$

در نظر می‌گیریم.

(۱-۱) تعریف هندسی دترمینان

در حالت 2×2 : مساحت $|\det A| = P(A)$

$$\det A \text{ علامت} = \begin{cases} + & \text{اگر زوج مرتب } (A^1, A^2) \text{ راستگرد باشد} \\ - & \text{اگر زوج مرتب } (A^1, A^2) \text{ چپگرد باشد} \\ 0 & \text{اگر } \{A^1, A^2\} \text{ وابسته خطی باشد} \end{cases}$$

در حالت 3×3 : حجم $|\det A| = P(A)$

$$\det A \text{ علامت} = \begin{cases} + & \text{اگر سه تایی مرتب } (A^1, A^2, A^3) \text{ راستگرد باشد} \\ - & \text{اگر زوج مرتب } (A^1, A^2, A^3) \text{ چپگرد باشد} \\ \circ & \text{اگر } \{A^1, A^2, A^3\} \text{ وابسته خطی باشد} \end{cases}$$

(۱۱-۲) تعریف جبری دترمینان

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (1)$$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \quad (2)$$

برای تعمیم این مفاهیم به ماتریس‌های $n \times n$ و \mathbb{R}^n ، طبعاً ما ذهنیتی پیشینی نسبت به حجم n -بعدی و "راستگرد بودن یک n -تایی بردارها در \mathbb{R}^n " نداریم. بنابراین راهبرد ما این خواهد بود که تعریف جبری را به گونه‌ای تعمیم دهیم که قدرمطلق آن خواصی مشابه مساحت و حجم در \mathbb{R}^n را داشته باشد و علامت آن، n -تایی‌های برداری مستقل خطی را به دو دسته "راستگرد" و "چپگرد" تقسیم کند به نحوی که قرابت مورد نظر با دوگونگی متناظر در \mathbb{R}^2 و \mathbb{R}^3 برقرار شود. به این منظور سه ویژگی دترمینان در حالت 2×2 و 3×3 را در زیر مورد نظر قرار می‌دهیم و خواهیم دید که به طور کلی برای ماتریس‌های $n \times n$ یک و تنها یک روش نسبت دادن یک عدد به یک ماتریس وجود دارد که این سه ویژگی را داراست. این عدد را "دترمینان ماتریس" نام خواهیم گذارد.

پایه متداول \mathbb{R}^n ، یعنی (e_1, \dots, e_n) را در نظر بگیرید. مقصود از مکعب n -بعدی واحد، مجموعه

زیر است:

$$K^n = \{t_1 e_1 + \dots + t_n e_n \mid 0 \leq t_1 \leq 1, \dots, 0 \leq t_n \leq 1\} \quad (3)$$

برای $n = 1$ داریم $K^1 = [0, 1]$ ، برای $n = 2$ $K^2 = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$ و برای

$$K^3 = \{(x, y, z) \mid 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}, n = 3 \quad (\text{شکل ۱})$$

به طور کلی، برای n عضو A^1, \dots, A^n در \mathbb{R}^n ، متوازی السطوح n -بعدی ایجاد شده توسط

A^1, \dots, A^n را به صورت زیر تعریف می‌کنیم:

$$P(A^1, \dots, A^n) = \{t_1 A^1 + \dots + t_n A^n \mid 0 \leq t_i \leq 1, i = 1, \dots, n\} \quad (4)$$

اگر n -تایی‌های A^1 تا A^n را به ترتیب به عنوان ستون‌های یک ماتریس A در نظر می‌گیریم،

$P(A^1, \dots, A^n)$ را به طور خلاصه به $P(A)$ نمایش می‌دهیم.

(۱۱-۳) سه ویژگی اساسی دترمینان (برای ماتریس‌های 2×2 و 3×3)

(۱۱-۳-۱) دترمینان نسبت به هر ستون خطی است اگر سایر ستون‌های ثابت نگاه داشته شوند.

مقصود از "خطی بودن" برقراری دو شرط ۱-۷ است، یعنی اگر همه درایه‌های یک ستون در

عددی r ضرب شوند، دترمینان در r ضرب می‌شود. اگر یک ستون (به عنوان یک n -تایی عددی)

برابر مجموع دو ستون باشد، دترمینان برابر مجموع دترمینان‌های ماتریس‌هایی خواهد شد که از تفکیک

دو ستون به دست می‌آیند. مثلاً:

$$\det \begin{bmatrix} a_{11} + a'_{11} & a_{12} & a_{13} \\ a_{21} + a'_{21} & a_{22} & a_{23} \\ a_{31} + a'_{31} & a_{32} & a_{33} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} a'_{11} & a_{12} & a_{13} \\ a'_{21} & a_{22} & a_{23} \\ a'_{31} & a_{32} & a_{33} \end{bmatrix}$$

این دو شرط به سادگی از تعریف جبری (۱) و (۲) به دست می‌آیند. نکته این است که در هر جمله

سمت راست (۱) و (۲) فقط یک درایه از هر ستون ظاهر می‌شود و با درجه یک. وقتی $r > 0$ ، تعبیر

هندسی شرط اول این است که با ضرب کردن طول یک ضلع متوازی الاضلاع (به ترتیب یک ضلع متوازی السطوح) در r و ثابت نگاه داشتن اضلاع دیگر، مساحت متوازی الاضلاع (به ترتیب حجم متوازی السطوح) در r ضرب می شود. وقتی $r < 0$ ، جهت گردش (راستگردی یا چپگردی) اضلاع معکوس می شود، بنابراین علامت دترمینان عوض می شود (شکل ۲).

تعبیر هندسی شرط دوم در حالت $n = 2$ در شکل ۳ نمایش داده شده است. اگر یک ضلع متوازی الاضلاع، مثلاً ضلع u ، برابر مجموع $u' + u''$ باشد به طوری که (u, v) ، (u', v) و (u'', v) هر سه راستگرد یا هر سه چپگرد باشند، مساحت $P(u, v)$ برابر مجموع مساحت های $P(u', v)$ و $P(u'', v)$ می شود. وقتی جهت گردش سه دوتایی یکسان نباشد، باید علامت دترمینان را نیز منظور کرد. تعبیر مشابهی در حالت سه بعدی برقرار است و این حکم را می توان از اینکه مساحت برابر $|u \times v|$ و حجم برابر $|(u \times v) \cdot w|$ است نتیجه گرفت.

اکنون ویژگی اساسی دوم دترمینان را بیان می کنیم:

(۱۱-۳-۲) دترمینان نسبت به ستون ها "ضدمتقارن" است، یعنی هرگاه جای دو ستون تعویض شود، دترمینان در (-1) ضرب می شود.

این مطلب را می توان به طور جبری از تعریف های جبری (۱) و (۲) مشاهده کرد. تعبیر هندسی این است که تعویض ترتیب دو ضلع جهت گردش را معکوس می کند. توجه کنید که اگر k بار تعویض ستونی صورت گیرد، هر بار دترمینان در (-1) ضرب می شود، بنابراین پس از k تعویض، دترمینان در $(-1)^k$ ضرب خواهد شد.

بالاخره، ویژگی اساسی سوم دترمینان به شرح زیر است:

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1, \quad \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1 \quad (11-3-3)$$

در حالت 2×2 ، این بدین معنی است که مساحت مربع واحد برابر یک است و دوتایی مرتب

(e_1, e_2) راستگرد می باشد. مشابهاً در حالت 3×3 ، حجم مکعب واحد برابر یک است و سه تایی مرتب (e_1, e_2, e_3) راستگرد است.

خواهیم دید که این سه ویژگی دترمینان را به طور منحصر به فرد مشخص می کنند. لازم است نخست مفهوم "ضدتقارن" را کمی بیشتر بررسی کنیم. به طور کلی جابجا کردن ترتیب n شیء، مانند n ستون یک ماتریس $n \times n$ ، یک "جایگشت" خوانده می شود. اگر n شیء را با شماره گذاری به مجموعه $\{1, 2, \dots, n\}$ نمایش دهیم، یک جایگشت $\{1, 2, \dots, n\}$ در واقع یک تابع یک به یک و پوشا از $\{1, \dots, n\}$ به خود آن است.

$$\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$$

مقصود از $\sigma(i)$ جای جدید i است. تابع همانی جایگشتی است که هر یک از 1 تا n را سر جای خود نگاه می دارد. یک نوع جایگشت ساده، "تبادل" است. مقصود از یک تبادل یا ترانهش جایگشتی است که $(n-2)$ عنصر را تثبیت می کند و جای دو عنصر باقیمانده را تعویض می کند. مثلاً

$$1 \rightarrow 1, \quad 2 \rightarrow 3, \quad 3 \rightarrow 2, \quad 4 \rightarrow 4$$

یک تبادل $\{1, 2, 3, 4\}$ است.

$(4-1)$ قضیه. هر جایگشت $\{1, \dots, n\}$ را می توان به صورت ترکیبی متناهی از تبادلهای نمایش داد. روش نمایش یکتا نیست ولی تعداد تبادلهای لازم همواره زوج یا همواره فرد است.

برهان. حکم اول را می توان با استقراء ثابت کرد. برای $n = 1$ ، حکم واضح است. فرض کنید حکم تا n ثابت شده است، آن را برای $(n+1)$ ثابت می کنیم. پس فرض کنید σ یک جایگشت $\{1, \dots, n, n+1\}$ است. اگر σ همانی باشد، می توان نوشت $\sigma = \tau \circ \tau$ که τ یک تبادل دلخواه است. اگر σ همانی نباشد، σ دست کم یک i را جابجا می کند، مثلاً $\sigma(i) = j \neq i$. حال تبادل ρ را در نظر بگیرید که بدین صورت تعریف می شود: $\rho(i) = j$ ، $\rho(j) = i$ و $\rho(k) = k$ اگر $k \neq i, j$ ترکیب $\rho \circ \sigma$

جایگشتی است که i را تثبیت می کند زیرا که $(\rho \circ \sigma)(i) = \rho(j) = i$. بدین ترتیب می توان $\rho \circ \sigma$ را یک جایگشت n شیء $\{1, \dots, i-1, i+1, \dots, n+1\}$ تلقی کرد. طبق فرض استقراء، ρ باید ترکیب تعدادی تبادله τ_i باشد (که i را ثابت نگاه می دارند)، $\rho \circ \sigma = \tau_k \circ \dots \circ \tau_1$. حال اگر روی دو طرف، تبادله ρ را اثر دهیم، چون همانی $\rho \circ \rho = \text{id}$ ، نتیجه می شود که $\sigma = \rho \circ \tau_k \circ \dots \circ \tau_1$ یعنی σ ترکیبی از تبادله ها است.

برای حکم دوم نخست به هر جایگشت σ از $\{1, \dots, n\}$ عدد $\varepsilon(\sigma)$ را به صورت زیر نسبت

می دهیم:

$$\varepsilon(\sigma) = \frac{1}{(1)!(2)!\dots(n-1)!} \prod_{i < j} (\sigma(j) - \sigma(i)) \quad (5)$$

نشان می دهیم که $\varepsilon(\sigma) = \pm 1$. توجه کنید که به ازای هر $k, l \in \{1, \dots, n\}$ ، $k \neq l$ جمله $(k-l)$ یا

$(l-k)$ یک و فقط یک بار در حاصل ضرب سمت راست ظاهر می شود، بنابراین می توان نوشت:

$$\begin{aligned} \prod_{i < j} (\sigma(i) - \sigma(j)) &= \pm \prod_{k < l} (l - k) \\ &= \pm (\prod_{1 < l} (l - 1)) (\prod_{2 < l} (l - 2)) \dots (\prod_{n-1 < l} (l - (n-1))) \end{aligned}$$

و طرف راست برابر $(1)!\dots(n-1)!$ است. بنابراین $\varepsilon(\sigma) = \pm 1$. حال اثر ترکیب یک تبادله τ را با

σ بر $\varepsilon(\sigma)$ بررسی می کنیم. نشان می دهیم $\varepsilon(\tau \circ \sigma) = -\varepsilon(\sigma)$. فرض کنید تبادله τ دو عدد متمایز μ و

ν را جابجا می کند و سایر اعداد را ثابت نگاه می دارد، مثلاً $\mu < \nu$. در حاصل ضرب طرف راست (5)

جملات $(\sigma(j) - \sigma(i))$ که در آنها $\sigma(i) \neq \mu, \nu$ و $\sigma(j) \neq \mu, \nu$ تغییری نمی کنند، داریم:

$$(\tau \circ \sigma)(j) - (\tau \circ \sigma)(i) = \sigma(j) - \sigma(i)$$

یکی از دو جمله $(\nu - \mu)$ یا $(\mu - \nu)$ یک و فقط یک بار به صورت $(\sigma(i) - \sigma(j))$ ظاهر می شود که با

اثر دادن τ در (-1) ضرب می شود. سایر جملات $(\sigma(j) - \sigma(i))$ را که در آنها یکی از جمله ها برابر μ

یا ν است. می توان دوتا، دوتا، به صورت زیر در نظر گرفت:

$$\pm (\sigma(k) - \nu)(\sigma(k) - \mu) \quad , \quad \sigma(k) \neq \mu, \nu \quad (6)$$

و (۶) تحت اثر τ تغییر علامت نمی دهد:

$$\begin{aligned}
 ((\tau \circ \sigma)(k) - \tau(\nu))((\tau \circ \sigma)(k) - \tau(\mu)) &= (\sigma(k) - \mu)(\sigma(k) - \nu) \\
 &= (\sigma(k) - \nu)(\sigma(k) - \mu)
 \end{aligned}$$

بنابراین ماحصل اثر دادن τ این است که $\varepsilon(\sigma)$ در (-1) ضرب می شود:

$$\varepsilon(\tau \circ \sigma) - \varepsilon(\sigma) \quad , \quad \tau : \text{تبادل} \quad (۷)$$

توجه کنید که اگر σ همانی باشد، $\varepsilon(\sigma) = 1$ زیرا به ازای $i < j$ ، $\sigma(j) - \sigma(i) = j - i > 0$. بنابراین برای هر تبادل τ داریم $\varepsilon(\tau) = -1$. حال اگر σ یک جایگشت دلخواه باشد، طبق قسمت اول قضیه، σ را به صورت ترکیبی از تبدل ها می نویسیم، $\sigma = \tau_k \circ \dots \circ \tau_1$ با استفاده مکرر از (۷) داریم:

$$\varepsilon(\sigma) = (-1)^k \quad (۸)$$

که k تعداد تبدل هاست. چون به هر σ یک $\varepsilon(\sigma)$ مشخص نسبت داده شده است، (۸) نشان می دهد تعداد تبدل های لازم برای نمایش σ باید همواره زوج یا همواره فرد باشد. □

به عنوان دستاوردی از قضیه بالا، اگر σ_1 و σ_2 دو جایگشت باشند، داریم:

$$\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1) \cdot \varepsilon(\sigma_2) \quad (۹)$$

یک جایگشت را زوج (به ترتیب فرد) می نامیم در صورتی که $\varepsilon(\sigma) = +1$ (به ترتیب $\varepsilon(\sigma) = -1$). هر جایگشت زوج ترکیب تعدادی زوج تبادل است و هر جایگشت فرد ترکیب تعدادی فرد تبادل.

به کمک قضیه ۱۱-۴ اکنون می توانیم نسبت به تعمیم مفهوم دترمینان به ماتریس های $n \times n$ اقدام کنیم.

(۱۱-۵) گزاره. فرض کنید به هر ماتریس $n \times n$ ، A ، عددی $D(A)$ نسبت داده ایم که واجد دو شرط

زیر است:

(الف) $D(A)$ نسبت به هر ستون خطی است اگر سایر ستون ها ثابت نگاه داشته شوند.

ب) $D(A)$ نسبت به ستون‌ها ضدمتقارن است، یعنی جابجایی دو ستون آن را در (-1) ضرب می‌کند.

در این صورت مقدار $D(I_n)$ ، D را به طور منحصر به فرد مشخص می‌کند. برهان. ستون‌های ماتریس $A = [a_{ij}]$ را به A^1, \dots, A^n نمایش می‌دهیم. اگر ستون j ام، A_j را به عنوان یک n تایی، یعنی عضوی از \mathbb{R}^n در نظر بگیریم، داریم:

$$A^j = a_{1j}e_1 + \dots + a_{nj}e_n$$

با توجه به (الف) داریم

$$D(A) = \sum_{i_1=1}^n \dots \sum_{i_n=1}^n a_{i_1,1} a_{i_2,2} \dots a_{i_n,n} D[e_{i_1} | e_{i_2} | \dots | e_{i_n}] \quad (10)$$

در ماتریس بالا چنانچه برای $\mu \neq \nu$ داشته باشیم $e_{i_\mu} = e_{i_\nu}$ ، مقدار $D[e_{i_1} | \dots | e_{i_n}]$ صفر می‌شود زیرا که جابجایی e_{i_μ} و e_{i_ν} از یک طرف ماتریس سمت راست بالا را عوض نمی‌کند و از طرفی دیگر طبق (ب) مقدار D باید در (-1) ضرب شود. بنابراین $D[e_{i_1} | \dots | e_{i_n}]$ صفر است مگر اینکه (i_1, i_2, \dots, i_n) جایگشتی σ از $(1, 2, \dots, n)$ باشد. بنابراین می‌توان (10) را به صورت زیر نوشت:

$$D(A) = \sum_{\sigma: \text{جایگشت}} a_{\sigma(1),1} \dots a_{\sigma(n),n} D[e_{\sigma(1)} | \dots | e_{\sigma(n)}]$$

ولی از خاصیت ضد تقارن (ب) می‌توانیم نتیجه بگیریم که:

$$\begin{aligned} D[e_{\sigma(1)} | \dots | e_{\sigma(n)}] &= \varepsilon(\sigma) D[e_1 | \dots | e_n] \\ &= \varepsilon(\sigma) D(I_n) \end{aligned}$$

پس

$$D(A) = \sum_{\sigma: \text{جایگشت}} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} D(I_n) \quad (11)$$

□

و گزاره به اثبات می‌رسد.

همچنان که در بررسی سه ویژگی اصلی دترمینان 2×2 و 3×3 دیدیم، ویژگی‌های اول و دوم، که مشابه (الف) و (ب) هستند، ویژگی‌های ابتدایی مساحت و حجم متوازی‌الاضلاع و متوازی‌السطوح با منظور کردن علامت هستند ویژگی سوم، یعنی نسبت دادن عددی به مکعب واحد، به منزله ارائه واحد یا مقیاس برای مساحت یا حجم است.

توجه کنید که ستون‌های I_n در واقع بردارهای تعریف کننده مکعب واحد هستند. با تعیین $D(I_n)$ مقدار $D(A)$ برای هر ماتریس A مشخص می‌شود. چنانچه $D(I_n)$ را برابر ۱ بگیریم تابع D حاصل، طبق تعریف، $\det(A)$ (دترمینان A)، خوانده می‌شود، پس:

$$\det A = \sum_{\sigma: \text{جایگشت}} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \quad (12)$$

(۱۱-۶) نتیجه. تابع \det که به هر ماتریس $n \times n$ مقدار (۱۲) را نسبت می‌دهد یگانه تابع از ماتریس‌های $n \times n$ به \mathbb{R} است که سه ویژگی زیر را دارد: (الف) نسبت به هر ستون خطی است وقتی سایر ستون‌ها ثابت نگاه داشته شوند. (ب) نسبت به ستون‌ها ضدمتقارن است، و (ج) $\det(I_n) = 1$. □
تعریف. برای $A^1, \dots, A^n \in \mathbb{R}^n$ ، حجم n -بعدی متوازی‌السطوح $P(A)$ برابر $|\det A|$ است.

تعبیر زیر برحسب نگاشت‌های خطی نیز مورد استفاده قرار خواهد گرفت. ماتریس A که از کنار هم قراردادن n تایی‌های A^1, \dots, A^n به دست می‌آید در واقع نماینده یک نگاشت خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ است که $f(e_j) = A^j$ برای هر $j = 1, \dots, n$. $\det f$ را همان $\det A$ تعریف می‌کنیم. نگاشت خطی f مکعب واحد را به متوازی‌السطوح $P(A)$ می‌نگارد. پس $|\det f|$ در واقع حجم n -بعدی تصویر مکعب واحد تحت f است.

(۱۱-۸) گزاره. برای تابع‌های خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ و $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ داریم:

$$\det(g \circ f) = (\det g)(\det f) \quad (13)$$

یا اگر ماتریس f را به A و ماتریس g را به B نمایش دهیم:

$$\det(BA) = (\det B)(\det A) \quad (۱۴)$$

برهان. برای اثبات، g (معادلاً B) را تثبیت می‌کنیم و دو تابع زیر از مجموعه ماتریس‌های $n \times n$ به \mathbb{R}

را در نظر می‌گیریم:

$$D_1, D_2 : n \times n \text{ مجموعه ماتریس‌های } \rightarrow \mathbb{R}$$

$$D_1(A) = (\det B)(\det A)$$

$$D_2(A) = \det(BA)$$

اگر نشان دهیم D_1 و D_2 هر دو واجد شرط‌های (الف) و (ب) گزاره (۱۱-۵) هستند، از این گزاره

نتیجه می‌شود که مقدار D_1 و D_2 با دانستن مقدار آنها به ازای $A = I_n$ تعیین می‌شود. ولی داریم:

$$D_1(I_n) = (\det B)(\det I_n) = \det B$$

$$D_2(I_n) = \det(BI_n) = \det B$$

پس اگر ثابت شود که شرط‌های (الف) و (ب) برای D_1 و D_2 برقرارند، خواهیم داشت

$D_1(A) = D_2(A)$ برای هر ماتریس A و حکم گزاره به اثبات می‌رسد. برای $D_1(A)$ که مضرب ثابتی

از $\det A$ است برقرار بودن (الف) و (ب) واضح است. موضوع را در مورد D_2 تحقیق می‌کنیم. فرض

کنید A^j ، ستون j -ام A ، به صورت زیر است:

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix} = \begin{bmatrix} a'_{1j} \\ \vdots \\ a'_{nj} \end{bmatrix} + \begin{bmatrix} a''_{1j} \\ \vdots \\ a''_{nj} \end{bmatrix}$$

در این صورت ستون j -ام ماتریس BA می‌شود:

$$\begin{bmatrix} \sum_{\nu=1}^n b_{1\nu} a_{\nu j} \\ \vdots \\ \sum_{\nu=1}^n b_{n\nu} a_{\nu j} \end{bmatrix} = \begin{bmatrix} \sum_{\nu=1}^n b_{1\nu} a'_{\nu j} \\ \vdots \\ \sum_{\nu=1}^n b_{n\nu} a'_{\nu j} \end{bmatrix} + \begin{bmatrix} \sum_{\nu=1}^n b_{1\nu} a''_{\nu j} \\ \vdots \\ \sum_{\nu=1}^n b_{n\nu} a''_{\nu j} \end{bmatrix}$$

چون ویژگی (الف) در مورد \det برقرار است داریم:

$$\det(BA) = \det(BA') + \det(BA'')$$

که در آن A' و A'' ماتریس‌های هستند که از تفکیک ستون j -ام ماتریس A به دست آمده‌اند. پس

$$D_r(A) = D_r(A') + D_r(A'')$$

همین‌طور اگر ستون j -ام A در عدد r ضرب شود، ستون j -ام BA نیز در r ضرب خواهد شد و خطی بودن نسبت به ستون j -ام ثابت می‌شود. در مورد ویژگی (ب)، اگر A' ماتریس باشد که از تعویض ستون‌های i و j ماتریس A به دست آید، BA' نیز از تعویض ستون‌های i و j ماتریس BA حاصل می‌شود زیرا که به طور کلی در حاصل ضرب دو ماتریس MN ، ستون j ام از ضرب کردن ردیف‌های M در ستون j -ام N به دست می‌آید. پس چون \det واجد شرط (ب) است و این شرط برای D_r نیز نتیجه می‌شود و گزاره به اثبات می‌رسد. \square

گزاره بالا نتایج مهمی در پی دارد از جمله:

(۱۱-۹) نتیجه. شرطی لازم و کافی برای وارون‌پذیری تابع خطی $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ این است که $\det f \neq 0$. مشابهاً شرطی لازم و کافی برای وارون‌پذیری ماتریس $n \times n$ ، A ، این است که $\det A \neq 0$.

برهان. اگر A وارون‌پذیر باشد، ماتریسی A^{-1} وجود دارد که $AA^{-1} = I$ ، پس طبق (۱-۸)، $(\det A)(\det A^{-1}) = 1$ ، در نتیجه $\det A \neq 0$. بالعکس اگر A (یا معادلاً تابع خطی متناظر، f) وارون‌پذیر نباشد، می‌دانیم ستون‌های A وابسته خطی خواهند شد، پس می‌توان یک ستون را به صورت ترکیب خطی ستون‌های دیگر نوشت. پس با استفاده از بسط دادن به کمک ویژگی (الف)، به مجموع $(n-1)$ دترمینان ماتریس‌هایی می‌رسیم که یک ستون آنها تکرار شده است. دترمینان ماتریسی که دو ستون برابر داشته باشد صفر است زیرا که با تعویض دو ستون از یک سو ماتریس عوض نمی‌شود و از سویی طبق (ب) مقدار دترمینان باید در (-1) ضرب شود. \square

یک دستاورد برهان بالا را جداگانه ثبت می‌کنیم:

(۱۱-۱۰) نتیجه. مجموعه $\{A^1, \dots, A^n\}$ از عناصر \mathbb{R}^n مستقل خطی است اگر و تنها اگر

□ $\det A \neq 0$.

اکنون می‌توانیم مفهوم "راستگرد" و "چپگرد" را در \mathbb{R}^n تعریف کنیم. فرض کنید $\{A^1, \dots, A^n\}$ مستقل خطی است. n تایی مرتب (A^1, \dots, A^n) را راستگرد (به ترتیب چپگرد) می‌نامیم چنانچه $\det A > 0$ (به ترتیب $\det A < 0$). بدین ترتیب عیناً مانند حالت‌های دوبعدی و سه‌بعدی $\det A$ حاوی دو اطلاع زیر است:

(i) $|\det A|$ برابر حجم n -بعدی متوازی‌السطوح $P(A)$ است.

(ii) علامت $\det A$ نشان‌دهنده راستگردی یا چپگردی n تایی مرتب ستون‌های (A^1, \dots, A^n) است.

خواص تابع‌های پیوسته (۲)

فرض کنید $f: S \rightarrow \mathbb{R}$ یک تابع باشد. نقطه x_0 از S را یک نقطه بیشینه یا ماکسیمم برای تابع f می‌نامیم در صورتی که $f(x_0) \geq f(x)$ برای هر x در S . به همین ترتیب نقطه کمینه یا می‌نیمم به عنوان نقطه‌ای x_0 که در آن $f(x_0) \leq f(x)$ برای هر x در S ، تعریف می‌شود. در حالت اول $f(x_0)$ را بیشینه یا ماکسیمم تابع f در S ، و در حالت دوم، $f(x_0)$ را کمینه یا می‌نیمم تابع f در S می‌نامند.

به طور کلی، به دلایل مختلف، تابع $f: S \rightarrow \mathbb{R}$ ممکن است فاقد ماکسیمم یا می‌نیمم باشد. در شکل ۱ سه تابع نمایش داده شده‌اند. در (الف) تابع پیوسته نیست، به‌ازای مقادیر صحیح می‌نیمم تابع اتخاذ می‌شود ولی تابع ماکسیمم ندارد. در واقع تابع به دلخواه به کوچکترین کران بالایی مقادیر خود نزدیک می‌شود ولی به‌ازای هیچ نقطه دامنه برابر کوچکترین کران بالایی، یعنی $+$ ، نمی‌شود. در (ب)، تابع پیوسته و صعودی است ولی از آنجا که دامنه تابع یک بازه باز است، تابع در هیچ نقطه دامنه به کوچکترین کران بالایی خود یا بزرگترین کران پایینی نمی‌رسد. در (ج) نیز تابع پیوسته و صعودی است ولی در نزدیک شدن به دو انتهای بازه تابع بی‌کران می‌شود. برای تابعی که مجموعه مقادیرش کران بالایی داشته باشد، نقطه ماکسیمم نقطه‌ای در دامنه است که تابع این مقدار را بگیرد، و به همین ترتیب، برای تابعی که مجموعه مقادیرش کران پایینی داشته باشد، نقطه می‌نیمم نقطه‌ای از دامنه است که مقدار تابع در آن برابر بزرگترین کران پایینی باشد. قضیه زیر نشان می‌دهد که برای یک تابع پیوسته تعریف شده روی $[a, b]$ ، a, b عدد حقیقی، همواره نقطه ماکسیمم و نقطه می‌نیمم وجود دارد، بالاخص چنین تابعی لزوماً کراندار است.

(۱-۱۲) قضیه. هر تابع پیوسته $f: [a, b] \rightarrow \mathbb{R}$ دارای نقطه ماکسیمم و می‌نیمم در $[a, b]$ است.

برهان این قضیه را نیز، مانند برهان قضیه مقدار بینی در بخش پیش، نخست برای $[a, b] = [0, 1]$ ارائه می‌کنیم. حالت کلی به روشی کاملاً مانند اثبات قضیه مقدار بینی از همین حالت خاص نتیجه خواهد شد که این نتیجه‌گیری را به خواننده واگذار می‌کنیم.

برهان. فرض می‌کنیم $[a, b] = [0, 1]$ ، $a = 0$ ، $b = 1$. نقاط بازه $[0, 1]$ را در مبنای ۲ می‌نویسیم. بدین ترتیب هر عضو $[0, 1]$ نمایشی به شکل

$$c = 0/c_1c_2c_3\dots$$

دارد که در آن c_i ها رقم ۰ یا ۱ هستند. نقطه‌ای با نمایش بالا جستجو می‌کنیم که نقطه ماکسیم تابع f باشد. مانند اثبات قضیه مقدار بینی ارقام c_1, c_2, c_3, \dots را به ترتیب می‌سازیم ولی روش کار در اینجا از پیچیدگی بیشتری برخوردار است. بازه $[0, 1]$ را به صورت اجتماع دو زیربازه $[0, \frac{1}{4}]$ و $[\frac{1}{4}, 1]$ در نظر می‌گیریم. ادعا می‌کنیم دست کم یکی از دو بازه $[0, \frac{1}{4}]$ و $[\frac{1}{4}, 1]$ ویژگی زیر را دارد: (*) هیچ نقطه t از این بازه وجود ندارد که به ازای هر t در بازه دیگر داشته باشیم:

$$f(t_0) > f(t)$$

این ادعا که دست کم یکی از دو بازه $[0, \frac{1}{4}]$ و $[\frac{1}{4}, 1]$ ویژگی (*) را دارد بدین طریق توجیه می‌شود: اگر یکی از این دو بازه ویژگی ذکر شده را نداشته باشد، آنگاه عنصر t_0 از این بازه هست که $f(t_0)$ اکیداً بزرگتر از $f(t)$ برای هر t در بازه دیگر است. بنابراین هیچ عنصری از بازه دیگر وجود ندارد که مقدار f در آن اکیداً بزرگتر از مقدار f در هر نقطه این بازه (بالاخص $f(t_0)$) باشد، یعنی بازه دیگر ویژگی (*) را داراست.

اگر فقط یکی از دو بازه ویژگی (*) را داشته باشد، بازه دیگر را I_1 می‌نامیم، و اگر هر دو بازه این ویژگی را داشته، یکی از آنها را به دلخواه I_1 می‌نامیم. رقم c_1 را به صورت زیر تعریف می‌کنیم:

$$c_1 = \begin{cases} 0 & \text{اگر } I_1 = [0, \frac{1}{4}] \\ 1 & \text{اگر } I_1 = [\frac{1}{4}, 1] \end{cases}$$

از این پس جستجو برای نقطهٔ ماکسیمم را به بازهٔ I_1 محدود می‌کنیم. I_1 را به صورت اجتماع دو زیربازهٔ چپ و راست، هر یک به طول $\frac{1}{4}$ می‌نویسیم. مجدداً به همان استدلالی که در بالا آمد ادعا می‌کنیم دست‌کم یکی از این دو زیربازهٔ I_1 باید واجد شرط (*) نسبت به دیگری باشد. مانند قبل اگر فقط یک زیربازه شرط (*) را احراز کند، دیگری را I_2 می‌نامیم، و اگر هر دو واجد شرط (*) باشند، یکی را به دلخواه I_2 می‌نامیم. تعریف می‌کنیم:

$$c_2 = \begin{cases} 0 & \text{اگر } I_2 \text{ زیربازهٔ چپ باشد} \\ 1 & \text{اگر } I_2 \text{ زیربازهٔ راست باشد} \end{cases}$$

و از این جستجو برای نقطهٔ ماکسیمم را به I_2 محدود می‌کنیم. این فرایند را با نیمه کردن I_2 ، ادعای اینکه (*) باید برای دست‌کم یک نیمهٔ آن برقرار باشد، انتخاب I_3 و تعیین c_3 برابر صفر یا یک بسته به این که نیمهٔ چپ I_3 باشد یا نیمهٔ راست، ادامه می‌دهیم. با ادامهٔ روش به ترتیب رقم‌های c_n ساخته می‌شوند. ثابت می‌کنیم نقطهٔ $c = 0.c_1c_2c_3\dots$ که بدین طریق به دست می‌آید یک نقطهٔ ماکسیمم است. استدلال به طریق برهان خلف است. فرض می‌کنیم c یک نقطهٔ ماکسیمم نباشد و به تناقض می‌رسیم. اگر $f(c)$ ماکسیمم نباشد، نقطه‌ای d در $[0, 1]$ وجود دارد، $d \neq c$ ، که:

$$f(d) > f(c) \quad (1)$$

توجه کنید که هر نقطهٔ غیر از c در یک مرحلهٔ فرایند نصف کردن بالا باید از c جدا شده باشد زیرا فاصله c تا d هرچه قدر کوچک باشد، عددی n وجود دارد $\frac{1}{2^n}$ (یعنی طول بازهٔ I_n)، که c همواره در آن است) از فاصله c تا d کوچکتر است و بازهٔ شامل c نمی‌تواند شامل d نیز باشد. فرض کنید n_1 اولین مرحله‌ای است که d در زیربازهٔ کنار گذاشته شده واقع شده است. در این صورت طبق (*) نقطه‌ای d' وجود دارد، در بازهٔ I_{n_1} ، که:

$$f(d) \leq f(d') \quad (2)$$

از (1) و (2) می‌بینیم که $f(d') > f(c)$. حال در مورد d' نیز استدلالی مشابه d انجام می‌دهیم. داریم $d' \neq c$ چون $f(d') \neq f(c)$ ، پس n_2 را اولین مرحله‌ای می‌گیریم که d' از c جدا شده است، یعنی d' در

I_{n_2} قرار ندارد. پس طبق (*) عنصری d'' در I_{n_2} وجود دارد که:

$$f(d') \leq f(d'') \quad (3)$$

بدین ترتیب تاکنون داریم $f(d'') \geq f(d') \geq f(d) \geq f(c)$ با ادامه این استدلال، دنباله‌ای $d^{(k)}$ از نقاط $[0, 1]$ پیدا می‌کنیم که $d^{(k)} \in I_{n_k}$ و $n_3 > n_2 > n_1 > \dots$. چون طول I_j برابر $\frac{1}{j}$ است و دنباله $(n_i)_i$ اکیداً صعودی است، دنباله I_{n_k} به سوی نقطه c منقبض می‌شود و داریم:

$$d^{(k)} \rightarrow c \text{ وقتی } k \rightarrow +\infty$$

نشان می‌دهیم این در تناقض با $f(c) < f(d)$ است. اگر $f(d) - f(c) > \epsilon$ را به e نمایش دهیم، نتیجه می‌شود $\delta > 0$ وجود دارد که برای هر x با $|x - c| < \delta$ داریم $|f(x) - f(c)| < e$ ، بالاخص

$$f(x) - f(c) < e$$

$$f(x) < f(c) + e \quad (4)$$

چون $d^{(k)} \rightarrow c$ برای k بزرگ داریم $|d^{(k)} - c| < \delta$ ، پس

$$f(d^{(k)}) < f(c) + e = f(d)$$

که در تناقض با $f(d'') \geq f(d') \geq f(d) > f(c)$ است. این تناقض نشان می‌دهد فرض $f(d) > f(c)$ نادرست است و $f(c)$ باید ماکسیمم تابع f در $[0, 1]$ باشد.

استدلال مربوط به می‌نیمم کاملاً مشابه است و به خواننده واگذار می‌شود. \square

(۱۲-۲) یادداشت. نکته مهمی در مورد اثبات بالا باید ذکر شود. در اثبات اینکه $f(c)$ ماکسیمم است، تنها استفاده ما از پیوستگی f ، استفاده از نامساوی (۴) بود. این در واقع "نصف پیوستگی" است زیرا که طبق پیوستگی f در c ، برای $\epsilon > 0$ ، $\delta > 0$ وجود دارد که $|x - a| < \delta$ نتیجه می‌دهد

$$|f(x) - f(c)| < \epsilon. \text{ نامساوی اخیر شامل دو حکم است:}$$

$$f(x) - f(c) < \epsilon$$

$$-e < f(x) - f(c)$$

در اثبات قبل فقط از نامساوی اول استفاده شد. اگر اثبات وجود می‌نیمم را به طور مشابه دنبال کنیم خواهیم دید که در آن اثبات فقط از نامساوی دوم بالا استفاده می‌شود. تابع‌هایی که فقط یکی از دو نامساوی بالا برایشان برقرار باشد تابع‌های "نیم پیوسته" خوانده می‌شوند. به طور خاص، تابع $f: S \rightarrow \mathbb{R}$ را نیم پیوسته از بالا در نقطه c از دامنه می‌نامیم در صورتی که برای هر $\epsilon > 0$ ، $\delta > 0$ وجود داشته باشد که هرگاه $x \in S$ و $|x - a| < \delta$ ، آنگاه

$$f(x) < f(c) + \epsilon$$

به همین ترتیب f نیم پیوسته از پایین در نقطه c خوانده می‌شود اگر برای $\epsilon > 0$ ، $\delta > 0$ وجود داشته باشد که هرگاه $x \in S$ و $|x - a| < \delta$ ، آنگاه:

$$f(x) > f(c) - \epsilon$$

f را نیم پیوسته از بالا در S (به ترتیب نیم پیوسته از پایین در S) می‌نامیم در صورتی که f در همه نقاط S نیم پیوسته از بالا (به ترتیب نیم پیوسته از پایین) باشد.

بدین ترتیب می‌توان برای بهره‌برداری بیشتر از قضیه ۱۲-۱، قضیه را به دو قسمت تجزیه کرد:

(۱۲-۳) قضیه. تابع $f: [a, b] \rightarrow \mathbb{R}$ داده شده است. (الف) اگر f نیم پیوسته از بالا باشد، f دارای ماکسیمم در $[a, b]$ است، (ب) اگر f نیم پیوسته از پایین باشد، f دارای می‌نیمم در $[a, b]$ است. □
 به زودی کاربرد مهمی از این صورت جامعتر قضیه خواهیم دید، ولی نخست مثال‌هایی ارائه می‌کنیم.

مثال ۱. تابع $f: [0, 1] \rightarrow \mathbb{R}$ را به صورت $f(x) = x - [x]$ در نظر می‌گیریم که مقصود از $[x]$ جزء صحیح x است. در واقع $f(x) = \begin{cases} x & 0 \leq x < 1 \\ 0 & x = 1 \end{cases}$. این تابع در همه نقاط به استثنای $x = 1$ پیوسته است و در $x = 1$ نیم پیوسته از پایین است و نیم پیوسته از بالا نیست. فرض کنید $\epsilon > 0$ داده شده است. قطعاً برای هر x داریم:

$$f(x) > f(0) - \epsilon = -\epsilon$$

زیرا که f در $[0, 1]$ نامنفی است. از طرفی دیگر، اگر $e = \frac{1}{4}$ را در نظر می‌گیریم، $\delta > 0$ هرچه باشد برای $x \in [0, 1]$ که $|x - 1| < \delta$ نمی‌توان حکم کرد که

$$f(x) < f(0) + \frac{1}{4} = \frac{1}{4}$$

زیرا که اگر x از طرف چپ به 1 نزدیک باشد، $f(x) = x$ نزدیک 1 است. ضمناً توجه کنید که این تابع در $[0, 1]$ دارای می‌نیم است ولی دارای ماکسیم نیست.

مثال 2. تابعی که به صورت $f: [0, 1] \rightarrow \mathbb{R}$ ، $f(x) = [x] - x$ تعریف می‌شود در نقطه 1 نیم‌پیوسته از بالا است ولی نیم‌پیوسته از پایین نیست. در واقع داریم $f(x) = \begin{cases} -x & 0 \leq x < 1 \\ 0 & x = 1 \end{cases}$ برای $e > 0$ داده شده، x هر نقطه‌ای که در دامنه باشد، داریم $f(x) < f(0) + e = e$ زیرا که مقادیر f نامثبت‌اند، پس نیم‌پیوستگی از بالا برقرار است. بالعکس برای $e = \frac{1}{4}$ ، $\delta > 0$ هرچه باشد، $|x - 1| < \delta$ ، $x \in [0, 1]$ دلالت بر $f(x) > f(1) - \frac{1}{4} = -\frac{1}{4}$ نمی‌کند زیرا که برای x نزدیک 1 از سمت چپ مقدار f نزدیک 1- است.

اکنون کاربرد مهمی از نیم‌پیوستگی را در رابطه با بحث "پیوستگی یکنواخت" که در پایان بخش 9 آمد ارائه می‌کنیم. فرض کنید $f: [a, b] \rightarrow \mathbb{R}$ یک تابع پیوسته باشد. از گزاره 9-4 به یاد می‌آوریم که هرگاه $e > 0$ داده شده باشد، برای هر نقطه $t \in [a, b]$ عددی $\delta > 0$ وجود دارد که برای هر دو نقطه t_1 و t_2 از $[a, b]$ که در $[t - \delta, t + \delta]$ قرار گیرند داریم $|f(t_1) - f(t_2)| < e$. برای هر $t \in [a, b]$ $\delta(t)$ را برابر کوچکترین کران بالایی $\delta > 0$ هایی می‌گیریم که در ویژگی بالا صدق می‌کنند ولی سقف $b - a$ را برای $\delta(t)$ منظور می‌کنیم. توجه کنید که به هر حال اگر δ از $b - a$ بزرگتر شود، آنگاه $[t - \delta, t + \delta]$ از بازه $[a, b]$ بزرگتر می‌شود و بزرگتر کردن آن نقاط جدیدی از $[a, b]$ به دست نمی‌دهد. بدین ترتیب اگر $\delta < \delta(t)$ ، آنگاه برای هر دو نقطه t_1 و t_2 که در $[t - \delta, t + \delta]$ باشند، داریم $|f(t_1) - f(t_2)| < e$.

(12-4) لم. برای $e > 0$ ثابت، تابع $\delta(t)$ که برای $t \in [a, b]$ تعریف شده است، نیم‌پیوسته از پایین است.

برهان. باید ثابت کنیم برای هر $\epsilon > 0$ داده شده، $\delta' > 0$ وجود دارد که هرگاه $t' \in]t - \delta', t + \delta'[$ آنگاه $\delta(t') > \delta(t) - \epsilon$ یعنی باید ثابت کنیم که برای این δ' ، اگر $|t' - t| < \delta$ و t_1 و t_2 در فاصله $\delta(t) - \epsilon$ از t' قرار گیرند، آنگاه $|f(t_1) - f(t_2)| < \epsilon$. برای مطلب اخیر، کافی است نشان دهیم که اگر t_1 و t_2 در این بازه باشند، آنگاه $]t - \delta(t), t + \delta(t)[$ قرار می‌دهیم $\epsilon < \delta' < \epsilon'$. بنابراین اگر t_1, t_2 در $]t' - \delta', t' + \delta'[$ باشند، داریم:

$$i = 1, 2 : |t_i - t| \leq |t_i - t'| + |t' - t| < \delta' + \delta(t) - \epsilon' < \delta(t)$$

و حکم به اثبات می‌رسد. □

به کمک این لم اکنون می‌توان به سادگی دید که هر تابع پیوسته $f : [a, b] \rightarrow \mathbb{R}$ به طور یکنواخت پیوسته است، یعنی هرگاه $\epsilon > 0$ داده شده باشد، آنگاه $\delta > 0$ وجود دارد که برای هر دو نقطه t_1 و t_2 از $[a, b]$ با $|t_1 - t_2| < \delta$ داریم $|f(t_1) - f(t_2)| < \epsilon$. به این منظور، برای $\epsilon > 0$ داده شده، برای هر $t \in [a, b]$ را طبق بحث فوق در نظر می‌گیریم δ تابعی است با مقادیر اکیداً مثبت که طبق لم نیم پیوسته از پایین است. بنابراین طبق ۱۲-۳، ب، $\delta(t)$ در $[a, b]$ می‌نیمم خود را اتخاذ می‌کند. مقدار این می‌نیمم را به δ نمایش می‌دهیم، پس $\delta \leq \delta(t)$ برای هر t . در نتیجه اگر $|t_1 - t_2| < \delta$ ، آنگاه t_2 در فاصله کوچکتر یا مساوی $\delta(t_1)$ از t_1 قرار دارد، بنابراین $|f(t_2) - f(t_1)| < \epsilon$ و حکم به اثبات می‌رسد:

□ (۱۲-۵) قضیه. هر تابع پیوسته $f : [a, b] \rightarrow \mathbb{R}$ به طور یکنواخت پیوسته است.

(κ : حرف یونانی "کاپا"، معمولاً برای انحنای به کار گرفته می‌شود). تعریف دقیق θ را باید (ضریب زاویه خط مماس) \arctan گرفت. دو شکل در اینجا ظاهر می‌شود:

شکل ۲

(۱) ضریب زاویه مربوط به خط مماس است که با یک بار مشتق‌گیری از خم γ به دست می‌آید، بنابراین مشتق‌گیری مجدد، $\frac{d\theta}{ds}$ ، مستلزم این است که خم اولیه دوبار مشتق‌پذیر باشد.
(۲) وقتی خط مماس در وضعیت قائم قرار گیرد، ضریب زاویه تعریف نشده است (یا ∞ است) هر چند که به نظر می‌رسد θ قابل تعریف شدن باشد.

در مورد (۱)، از این پس برای بحث در مورد انحنای فرض می‌کنیم که γ دوبار مشتق‌پذیر باشد. در مورد (۲)، اشکال تعریف θ برطرف شدنی است و در تعریف معادلی که بعداً از انحنای ارائه خواهد شد، به آن اشاره خواهیم کرد. فعلاً با همین تعریف (۱)، وضعیت خط راست و دایره را بررسی می‌کنیم:
مثال ۱. برای خط راست در صفحه، θ ، ثابت است، پس $\frac{d\theta}{ds} = 0$ ، بین انحنای خط راست، همچنان که انتظار داریم، صفر است.

شکل ۳

مثال ۲. دایره‌ای به شعاع R در نظر بگیرید که در جهت مثلثاتی پرمایش شده است (شکل ۴).

شکل ۴

داریم

$$\theta = \frac{\pi}{2} + \alpha$$

$$\alpha = \frac{s}{R} \quad (\text{برحسب رادیان})$$

پس $\kappa = \frac{d\theta}{ds} = \frac{1}{R}$. بدین ترتیب همان گونه که انتظار می‌رفت، انحنای متناسب با معکوس شعاع دایره است.

در مثال بالا، اگر دایره در جهت عقربه‌ساعت پرمایش شود، خواهیم داشت $\theta = \alpha - \frac{\pi}{2}$ و $\alpha = -\frac{s}{R}$ (را همواره در جهت مثلثاتی مثبت، و s را در جهت حرکت مثبت فرض می‌کنیم)، پس

بعضی کتابها از قدرمطلق $\frac{d\theta}{ds}$ به عنوان انحنای صحبت می کنند، که در آن صورت علامت منفی ظاهر نمی شود، ولی این روش باعث از دست رفتن پاره‌ای اطلاعات مهم می شود. در تعریف ما، صعود θ در جهت حرکت به منزله انحنای مثبت و نزول آن به معنای انحنای منفی است.

شکل ۵

به زودی به مفهوم فیزیکی این علامت اشاره خواهیم کرد.

با تعریف (۱) می توان به محاسبه انحنای انواع خمها اقدام کرد، ولی در اینجا ترجیح می دهیم تعریف معادلی از انحنای ارائه کنیم که قابلیت تعمیم به خمهای واقع در ابعاد بالاتر را نیز دارد. خم هموار دوبار مشتق پذیر $\gamma: I \rightarrow \mathbb{R}^2$ را در نظر بگیرید. مماس واحد در جهت حرکت به ازای هر $t \in I$ قابل تعریف کردن است:

$$\vec{T} = \frac{1}{|\gamma'(t)|} \gamma'(t) \quad (\gamma'(t) \neq 0 \text{ هموار بودن موجب می شود که } \gamma'(t) \neq 0) \quad (2)$$

ضمناً بنابر (۵) بخش قبل، داریم:

$$\vec{T} = \frac{d\vec{r}}{ds} \quad (3)$$

که s مقدار تابع طول در جهت حرکت است. چون \vec{T} یک بردار واحد است، می توان مؤلفه‌های افقی و قائم آن را به ترتیب به صورت کسینوس و سینوس زاویه θ نوشت:

$$\vec{T} = (\cos \theta, \sin \theta) \quad (4)$$

اگر γ دوبار مشتق پذیر باشد، می توان ثابت کرد (ولی اثبات دقیق از حوصله این درس خارج است) که θ را می توان در طول خم به صورت تابعی مشتق پذیر از t (یا s) در نظر گرفت. نکته قابل توجه در مورد چنین تابع θ این که معمولاً نمی توان برای یک خم پیچیده، مقدار θ را به بازه خاصی به طول 2π محدود کرد. به شکل ۶، الف، توجه کنید:

ب

الف

شکل ۶

در نقطه A می‌توانیم θ را برابر 0 یا هر مضرب دیگر 2π انتخاب کنیم، ولی پس از این انتخاب، وقتی θ به طور پیوسته در طول خم تغییر کند، مقدار آن در نقاط دیگر با انتخاب آن در A تحمیل می‌شود. مثلاً اگر θ را در A برابر 0 بگیریم، مقدار θ در B برابر $\frac{\pi}{4}$ ، در C برابر π ، در D برابر $\frac{3\pi}{4}$ و در E برابر 2π خواهد شد. با این که توابع مثلثاتی سینوس و کسینوس برای 0 و 2π یک مقدار دارند، ولیکن اگر θ به طور پیوسته تغییر کند، مقدار θ در A و در E برابر نخواهد بود، بلکه به اندازه 2π افزایش خواهد داشت که یک گردش کامل مماس واحد حول دایره مثلثاتی را نشان می‌دهد. در شکل ۶ ب، مماس‌های واحد در طول خم را به مبداء مختصات منتقل کرده‌ایم؛ A', B', C', D' و E' منناظر با A, B, C, D و E هستند. وقتی متحرک در طول خم از A به E حرکت کند، مماس واحد روی دایره واحد یک دور کامل دایره مثلثاتی را در جهت مثبت می‌پیماید و بدین ترتیب از A تا E مقدار θ به اندازه 2π افزایش می‌یابد. به (۳) و (۴) باز می‌گردیم. \vec{T} همواره طول ثابت واحد را دارد بنابراین آهنگ تغییر آن نسبت به s باید صرفاً خمیدن γ را مشخص کند، یعنی با انحنا رابطه داشته باشد. در واقع:

$$\begin{aligned}\frac{d\vec{T}}{ds} &= \left(-\sin\theta \frac{d\theta}{ds}, \cos\theta \frac{d\theta}{ds}\right) \\ &= \kappa(-\sin\theta, \cos\theta)\end{aligned}$$

با توجه به این که $\cos(\theta + \frac{\pi}{2}) = -\sin\theta$ و $\sin(\theta + \frac{\pi}{2}) = \cos\theta$ بردار $(-\sin\theta, \cos\theta)$ عمود بر \vec{T} و چرخش یافته نسبت به آن به اندازه $\frac{\pi}{2}$ در جهت مثلثاتی است. این بردار را به \vec{N} نمایش می‌دهیم و قائم واحد γ می‌نامیم. توجه کنید که به ازای هر t ، دو قائم واحد قرینه برای γ وجود دارد که \vec{N} آن قائم است که از گردش \vec{T} به اندازه $\frac{\pi}{2}$ در جهت مثلثاتی به دست می‌آید. پس داریم:

$$\frac{d\vec{T}}{ds} = \kappa\vec{N} \quad (5)$$

زوج (\vec{T}, \vec{N}) و امتدادهای تعیین شده توسط آنها را به صورت یک دستگاه مختصات متحرک چسبیده به تصویر γ تجسم می‌کنیم (\vec{T}, \vec{N}) معمولاً کنج متحرک γ خوانده می‌شود. (۵) را می‌توان این گونه تعبیر کرد که κ آهنگ تغییر این دستگاه مختصات متحرک وابسته به خم است. ضمناً اگر از

$$\vec{N} = (-\sin\theta, \cos\theta) \text{ نسبت به } s \text{ مشتق بگیریم، حاصل می‌شود:}$$

$$\frac{d\vec{N}}{ds} = -\kappa\vec{T} \quad (6)$$

قبلاً اشاره کردیم که $\kappa = \frac{d\theta}{ds}$ متضمن دوبار مشتق گیری از γ است. از این رو رابطه κ را با بردار شتاب

$\vec{a} = \frac{d^2\vec{r}}{dt^2}$ بررسی می کنیم:

$$\begin{aligned}\vec{a} &= \frac{d}{dt}\left(\frac{d\vec{r}}{dt}\right) \\ &= \frac{d}{dt}\left(\frac{ds}{dt} \frac{d\vec{r}}{ds}\right) = \frac{d}{dt}\left(\frac{ds}{dt} \vec{T}\right) \\ &= \frac{d^2s}{dt^2} \vec{T} + \left(\frac{ds}{dt}\right) \left(\frac{ds}{dt} \cdot \frac{d\vec{T}}{ds}\right)\end{aligned}$$

پس

$$\vec{a} = \left(\frac{d^2s}{dt^2}\right) \vec{T} + \kappa \left(\frac{ds}{dt}\right)^2 \vec{N} \quad (7)$$

فرمول (7) را می توان تجزیه بردار شتاب در راستای دستگاه مختصات متحرک وابسته به γ تعبیر کرد. توجه کنید که برای $\kappa > 0$ ، مؤلفه \vec{a} در جهت \vec{N} مثبت و برای $\kappa < 0$ ، مؤلفه \vec{a} در جهت \vec{N} منفی است. اگر مسیر حرکت را مسیر یک حرکت فیزیکی فرض کنیم، \vec{a} و نیرو همجهت هستند و \vec{a} باید همواره به سوی تقعر منحنی اشاره کند. این نکته با تعبیری که قبلاً از علامت κ ارائه کردیم سازگار است (شکل 7).

شکل 7

حال سؤال اساسی زیر را مطرح می کنیم: انحنا تا چه حد معرف شکل هندسی (تصویر) یک خم است؟ قضیه زیر پاسخی قاطع به این سؤال می دهد. در واقع برای خم های هموار، کمیت انحنا شکل منحنی را به طور کامل تعیین می کند، یعنی دو منحنی که در نقاط متناظر انحناهای برابر داشته باشند با یک حرکت اقلیدسی در صفحه (انتقال + دوران) برهم قابل انطباقند.

(۱۳-۲) قضیه اساسی خم ها فرض کنید $\alpha, \beta : I \rightarrow \mathbb{R}^2$ دو خم هموار پرمایش شده بر حسب طول باشند. انحناهای α و β را به ترتیب به κ_1 و κ_2 نمایش می دهیم. اگر به ازای هر $s \in I$ ، داشته باشیم $\kappa_1(s) = \kappa_2(s)$ ، آنگاه می توان با یک حرکت اقلیدسی (انتقال + دوران) تصویر دو خم را برهم منطبق کرد.

اثبات یک نقطه مرجع s_0 در I اختیار می‌کنیم. نخست انتقالی در صفحه انجام می‌دهیم که $\alpha(s_0) = \beta(s_0)$ و $\beta(s_0)$ را برهم منطبق کند. "انتقال" به معنی تابعی $f(x, y) = (x, y) + (a, b)$ است که در آن برداری ثابت (a, b) بردار انتقال است. بدین ترتیب اگر $\alpha(s) = (\alpha_1(s), \alpha_2(s))$ ، $f(\alpha(s)) = (\alpha_1(s) + a, \alpha_2(s) + b)$ ، و چون a و b ثابت هستند، \vec{T} ، \vec{N} و κ (که از مشتق‌گیری از α نسبت به s به دست می‌آیند) تغییر نمی‌کنند. بدین ترتیب می‌توان از اول فرض کرد که $\alpha(s_0) = \beta(s_0)$. حال کنج‌های متحرک وابسته به α و β را به ترتیب به (\vec{T}_1, \vec{N}_1) و (\vec{T}_2, \vec{N}_2) نمایش می‌دهیم. از آنجا که (\vec{T}_2, \vec{N}_2) و (\vec{T}_1, \vec{N}_1) هر دو راستگرد هستند، می‌توان با یک دوران به مرکز $\alpha(s_0) = \beta(s_0)$ فرض کرد $\vec{T}_2(s_0) = \vec{T}_1(s_0)$ و $\vec{N}_2(s_0) = \vec{N}_1(s_0)$. توجه کنید که دوران، زاویه θ را به مقداری ثابت تغییر می‌دهد، پس تابع‌های انحنا تغییر نخواهد کرد. بدین ترتیب می‌توان از آغاز فرض کرد که $\vec{T}_2(s_0) = \vec{T}_1(s_0)$ و $\vec{N}_2(s_0) = \vec{N}_1(s_0)$. حال تابع $\varphi: I \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$\varphi(s) = \vec{T}_1(s) \cdot \vec{T}_2(s) + \vec{N}_1(s) \cdot \vec{N}_2(s)$$

مشتق‌گیری نسبت به s با استفاده از فرمول (۷) جلسه قبل و فرمول‌های (۵) و (۶) این جلسه نتیجه می‌دهد:

$$\varphi'(s) = \kappa_1(s) \vec{N}_1(s) \cdot \vec{T}_2(s) + \kappa_2(s) \vec{T}_1(s) \cdot \vec{N}_2(s) - \kappa_1(s) \vec{T}_1(s) \cdot \vec{N}_2(s) - \kappa_2(s) \vec{N}_1(s) \cdot \vec{T}_2(s)$$

از فرض قضیه، $\kappa_1(s) = \kappa_2(s)$ داریم $\varphi'(s) = 0$ برای هر s ، پس φ ثابت است. بنابراین:

$$\varphi(s) = \varphi(s_0) = \vec{T}_1(s_0) \cdot \vec{T}_2(s_0) + \vec{N}_1(s_0) \cdot \vec{N}_2(s_0)$$

و چون به ازای s_0 ، دو کنج برهم منطبقند، داریم $\varphi(s) = 2$ به ازای هر s . ولی بردارهای \vec{T}_i و \vec{N}_i بردارهای واحدند، بنابراین $\varphi(s)$ تنها در صورتی می‌تواند همواره برابر ۲ باشد که برای هر s داشته باشیم $\vec{T}_1(s) = \vec{T}_2(s)$ و $\vec{N}_1(s) = \vec{N}_2(s)$. رابطه اول نتیجه می‌دهد که $\frac{d}{ds}(\alpha(s) - \beta(s)) = 0$ پس $\alpha(s) - \beta(s)$ ثابت است. ولی به ازای $s = s_0$ ، با انتقال فرض کردیم $\alpha(s_0) = \beta(s_0)$ ، پس $\alpha(s) - \beta(s) = 0$ برای هر s ، یعنی با انتقال و دوران انجام شده، دو خم کاملاً برهم منطبق شده‌اند. ■

خم‌های هموار در \mathbb{R}^n

اکنون به بررسی خم‌های هموار $\gamma : I \rightarrow \mathbb{R}^3$ می‌پردازیم. روشن است که برای یک خم در \mathbb{R}^3 امکان خمیدن و تابیدن بیشتری وجود دارد تا برای خم مسطح. اگر مانند سابق یک نیم‌خط را به عنوان جهت مرجع تعریف کنیم، هر مقدار θ به جای این که یک نیم‌خط منحصر بفرد تعریف کند، یک مخروط در فضا در تعیین خواهد کرد که هر مولد این مخروط زاویه θ با محور مخروط می‌سازد (شکل ۱، الف). همچنین، هرچند که مماس واحد \vec{T} در جهت حرکت

(الف)

(ب)

شکل ۱

برای خم هموار فضایی معنی دارد (طبق معمول $\gamma'(t)$ یا $\frac{d\vec{r}}{ds}$)، ولی برای بردار قائم واحد بی‌نهایت نامزد وجود دارد که در یک صفحه عمود بر تصویر خم قرار می‌گیرند. آیا می‌توان \vec{N} مشخصی را از این میان انتخاب کرد؟

چون $\vec{T}(s) \cdot \vec{T}(s) = 1$ برای هر s ، با مشتق‌گیری نسبت به s و استفاده از فرمول (۷) بخش (۱۲)

نتیجه می‌شود که:

$$\frac{d\vec{T}}{ds} \cdot \vec{T} = 0 \quad \text{یا} \quad \frac{d\vec{T}}{ds} \perp \vec{T} \quad (1)$$

دو حالت می‌توان در نظر گرفت:

حالت ۱: $\frac{d\vec{T}}{ds} = \underline{0}$ در این حالت $\kappa(s)$ را همانند (۵) بخش (۱۳) برابر صفر تعریف می‌کنیم.

حالت ۲: $\frac{d\vec{T}}{ds} \neq \underline{0}$ در این حالت $\frac{d\vec{T}}{ds}$ جهتی مشخص عمود بر \vec{T} مشخص می‌کند. تعریف می‌کنیم:

$$\kappa(s) = \left| \frac{d\vec{T}}{ds} \right|, \quad \vec{N} = \frac{1}{\kappa(s)} \frac{d\vec{T}}{ds} \quad (2)$$

بدین ترتیب برای یک خم هموار $\mathbb{R}^3 \rightarrow I : \gamma$ ، همواره $\kappa(s) \geq 0$. در هر نقطه که $\kappa(s) \neq 0$ ، قائم واحدی \vec{N} از میان قائم‌های واحد ممکن مشخص کرده‌ایم. از این پس فقط به نقاطی که در آن $\kappa(s) \neq 0$ توجه می‌کنیم (حالت ۲ بالا). برای تکمیل کنج متحرک، یا دستگاه مختصات متحرک وابسته به γ در این حالت، بردار واحد سومی \vec{B} را به صورت زیر در نظر می‌گیریم:

$$\vec{B} = \vec{T} \times \vec{N} \quad (3)$$

چون \vec{B} بر \vec{T} عمود است، \vec{B} نیز در صفحه عمود بر خم در نقطه $\gamma(s)$ قرار دارد. گاهی \vec{N} را قائم واحد اول (یا قائم واحد اصلی)، و \vec{B} را قائم واحد دوم می‌نامند. به تبعیت از حالت دوبعدی، آهنگ تغییرات $(\vec{T}, \vec{N}, \vec{B})$ نسبت به s را بررسی می‌کنیم، با این امید که کمیت‌هایی که به دست می‌آیند گویای شکل کامل تصویر خم باشند. شکل کلی باید چنین باشد:

$$\begin{cases} \frac{d\vec{T}}{ds} = a_{11}\vec{T} + a_{12}\vec{N} + a_{13}\vec{B} \\ \frac{d\vec{N}}{ds} = a_{21}\vec{T} + a_{22}\vec{N} + a_{23}\vec{B} \\ \frac{d\vec{B}}{ds} = a_{31}\vec{T} + a_{32}\vec{N} + a_{33}\vec{B} \end{cases} \quad (4)$$

که در آن a_{ij} ها تابع‌هایی نسبت به s هستند. تاکنون طبق تعریف داریم $a_{11} = a_{13} = 0$ و $a_{12}(s) = \kappa(s)$. به بررسی شش ضریب ردیف‌های دوم و سوم می‌پردازیم. اگر از $\vec{N}(s) \cdot \vec{N}(s) \equiv 1$ و $\vec{B}(s) \cdot \vec{B}(s) \equiv 1$ نسبت به s مشتق بگیریم نتیجه می‌شود که $\frac{d\vec{N}}{ds} \cdot \vec{N} = 0$ و $\frac{d\vec{B}}{ds} \cdot \vec{B} = 0$. نتیجه این که $a_{22} = 0$ و $a_{33} = 0$ (ضرب داخلی سطر دوم را با \vec{N} و ضرب داخلی سطر سوم را با \vec{B} در نظر بگیرد). همچنین از $\vec{T}(s) \cdot \vec{B}(s) \equiv 0$ ، با مشتق‌گیری نسبت به s نتیجه می‌شود که:

$$\frac{d\vec{T}}{ds} \cdot \vec{B} + \vec{T} \cdot \frac{d\vec{B}}{ds} = 0$$

$$a_{12} + a_{31} = 0$$

ولی $a_{13} = 0$ ، پس $a_{31} = 0$. بدین ترتیب، علاوه بر دو ضریب سطر اول، ضرایب a_{22} ، a_{31} و a_{33} نیز صفر هستند. حال با مشتق‌گیری از $\vec{T} \cdot \vec{N} \equiv 0$ و $\vec{N} \cdot \vec{B} \equiv 0$ به روشی مشابه بالا نتیجه می‌گیریم

که $a_{21} + a_{12} = 0$ و $a_{22} + a_{23} = 0$. چون $a_{12} = \kappa$ (طبق تعریف)، داریم $a_{21} = -\kappa$. ضریب تابعی a_{23} را به τ نمایش می‌دهیم و تاب می‌نامیم. پس:

$$\begin{cases} \frac{d\vec{T}}{ds} = & +\kappa\vec{N} \\ \frac{d\vec{N}}{ds} = & -\kappa\vec{T} & +\tau\vec{B} \\ \frac{d\vec{B}}{ds} = & & -\tau\vec{N} \end{cases} \quad (5)$$

لازم به تأکید است تعریف \vec{N} ، \vec{B} و τ و برقراری سطرهای دوم و سطرهای سوم (5) وابسته به فرض $\kappa \neq 0$ است. یک خم هموار در \mathbb{R}^3 که در سراسر آن $\kappa \neq 0$ ، خم فرنه خوانده می‌شود. کنج متحرک $(\vec{T}, \vec{N}, \vec{B})$ یک کنج فرنه و روابط (5) معادلات فرنه نامگذاری شده‌اند.

شهود ما از κ تا حدی مشابه حالت دو بعدی تأمین می‌شود، κ ضریب \vec{N} در رابطه آهنک تغییر \vec{T} نسبت به s می‌باشد. τ چه اهمیت هندسی دارد که "تاب" نام گرفته است؟ در زیر می‌بینیم که τ شاخص "تمایل" یک خم فضایی به خروج از صفحه مسطح است.

(4-4) گزاره برای خم فرنه $\gamma: I \rightarrow \mathbb{R}^3$ ، $\tau \equiv 0$ اگر و تنها اگر تصویر γ به تمامی در یک صفحه قرار داشته باشد.

اثبات فرض می‌کنیم γ برحسب طول پرمایش شده است (که می‌دانیم همواره می‌توان خم را با حفظ جهت نسبت به طول پرمایش کرد). نقطه مرجع $\gamma(s_0)$ در نظر بگیرید و تعریف کنید:

$$f: I \rightarrow \mathbb{R}, \quad f(s) = (\gamma(s) - \gamma(s_0)) \cdot \vec{B}(s)$$

داریم:

$$\begin{aligned} f'(s) &= \gamma'(s) \cdot \vec{B}(s) + (\gamma(s) - \gamma(s_0)) \cdot \frac{d\vec{B}}{ds} \\ &= \vec{T}(s) \cdot \vec{B}(s) + (-\tau)(\gamma(s) - \gamma(s_0)) \cdot \vec{N} \\ &= -\tau(\gamma(s) - \gamma(s_0)) \cdot \vec{N} \end{aligned}$$

اگر $\tau(s) \equiv 0$ ، نتیجه می‌شود که $f'(s) \equiv 0$ ، پس $f(s)$ ثابت است و

$$f(s) = f(s_0) = 0$$

ضمناً $\tau(s) \equiv 0$ از سطر سوم معادلات فرنه نتیجه می دهد که $B(s)$ نیز ثابت و همواره برابر $B(s_0)$ است. پس بردار واصل از $\gamma(s_0)$ به هر نقطه دیگر تصویر خم، $\gamma(s)$ ، هموار در صفحه گذرا از $\gamma(s_0)$ و عمود بر $B(s_0)$ قرار دارد.

بالعکس اگر $\gamma(s)$ همواره در یک صفحه قرار داشته باشد، $T(s) = \gamma'(s)$ خمی در همان صفحه ترسیم می کند، پس $\vec{N}(s) = \frac{1}{\kappa} \frac{d\vec{T}}{ds}$ نیز در همان صفحه است. نتیجه این که \vec{B} باید ثابت باشد،
 $\tau \equiv 0$ و $\frac{d\vec{B}}{ds} \equiv 0$.

(۱۴-۲) صفحه بوسان فرض کنید $\gamma : I \rightarrow \mathbb{R}^3$ یک خم فرنه باشد. در این صورت در هر نقطه بردار \vec{N} تعریف شده است. زوج متعامد (\vec{T}, \vec{N}) یک صفحه در \mathbb{R}^3 ایجاد می کند. این صفحه را صفحه بوسان γ در نقطه مربوط می نامند. به تعابیر مختلف می توان نشان داد این صفحه "نزدیکترین" صفحه به تصویر خم است. در واقع دیدیم که شرط لازم و کافی برای $\tau \equiv 0$ این است که تصویر γ به تمامی در صفحه بوسان بماند. هر صفحه شامل امتداد $\vec{T}(t)$ را می توان صفحه ای مماس بر تصویر γ در نقطه $\gamma(t)$ تلقی کرد. از بین این صفحات، صفحه بوسان به مفهومی که در آینده خواهیم دید "مماس ترین" این صفحه هاست. تعبیر دیگری از صفحه بوسان به صورت زیر به دست می آید. $t_0 \in I$ را تثبیت کنید و t_1 و t_2 دو مقدار دیگر در I بگیرید. سه نقطه $\gamma(t_0)$ ، $\gamma(t_1)$ و $\gamma(t_2)$ معمولاً یک صفحه پدید می آورند. وقتی $t_1 \rightarrow t_0$ و $t_2 \rightarrow t_0$ می توان نشان داد که این صفحه به صفحه بوسان میل می کند. تعبیر فیزیکی زیر نیز از صفحه بوسان قابل توجه است. بردار شتاب \vec{a} را می توان همانگونه که در حالت دو بعدی دیدیم در راستاهای تعیین شده توسط \vec{T} و \vec{N} تجزیه کرد:

$$\begin{aligned} \vec{a} &= \frac{d}{dt} \left(\frac{d\vec{r}}{dt} \right) = \frac{d}{dt} \left(\frac{ds}{dt} \vec{T} \right) \\ &= \left(\frac{d^2s}{dt^2} \right) \vec{T} + \left(\frac{ds}{dt} \right) \left(\frac{d\vec{T}}{ds} \frac{ds}{dt} \right) \end{aligned}$$

پس

$$\vec{a} = \left(\frac{d^2s}{dt^2} \right) \vec{T} + \kappa \left(\frac{ds}{dt} \right)^2 \vec{N} \quad (6)$$

یعنی \vec{a} همواره در صفحه بوسان قرار می گیرد. بدین ترتیب برای هر حرکت فیزیکی ممکن روی تصویر خم γ بردار شتاب همواره در صفحه بوسان می ماند.

(۱۴-۳) قضیه اساسی خم ها به شیوه ای کاملاً مشابه اثبات قضیه اساسی خم ها در \mathbb{R}^2 می توان نشان داد که هرگاه $\alpha, \beta : I \rightarrow \mathbb{R}^3$ دو خم فرنه پرمایش شده برحسب طول باشند با انحنا و تاب به ترتیب (κ_1, τ_1) و (κ_2, τ_2) به طوری که به ازای هر $s \in I$ داشته باشیم $\kappa_1(s) = \kappa_2(s)$ و $\tau_1(s) = \tau_2(s)$ آنگاه یک حرکت اقلیدسی در \mathbb{R}^3 وجود دارد که تصویر γ_1 و γ_2 را بر هم منطبق می کند. مقصود از یک حرکت اقلیدسی در اینجا نیز ترکیب انتقال و دوران است. اما مقصود از دوران در \mathbb{R}^3 چیست؟ دوران همواره به مفهوم چرخش \mathbb{R}^3 با یک زاویه ثابت حول یک محور ثابت است. بدین ترتیب در دوران یک خط راست (محور دوران) نقطه به نقطه ساکن می ماند و زاویه ای φ وجود دارد که هر صفحه عمود بر این محور روی خود به اندازه زاویه φ حول نقطه تقاطع با محور گردش می کند. در اثبات قضیه اساسی پس از آن که یک نقطه تصویر خم β را به نقطه متناظر روی خم α منتقل می کنیم، باید کنج فرنه β در آن نقطه را نیز بر کنج فرنه α در همان نقطه منطبق کنیم. در واقع قضیه ای در هندسه سه بعدی هست که طبق آن می توان هر سه تایی متعامد واحد راستگرد ساطع از یک نقطه را به هر سه تایی متعامد راستگرد دیگر در آن نقطه با یک دوران به صورت ذکر شده منطبق کرد. نکته دیگری که در اینجا شایان توجه است این است که در اثر چین دورانی مقادیر κ و τ تغییر نمی کنند. این موضوع نیاز به اثبات دارد. با فرض این نکات بقیه اثبات عیناً مانند حالت دوبعدی است با این تفاوت که به جای تابع $f(s) = \vec{T}_1(s) \cdot \vec{T}_2(s) + \vec{N}_1(s) + \vec{N}_2(s)$ در ۱۴-۱، از تابع $f(s) = \vec{T}_1(s) \cdot \vec{T}_2(s) + \vec{N}_1(s) + \vec{N}_2(s) + \vec{B}_1(s) \cdot \vec{B}_2(s)$ استفاده می کنیم. بدین ترتیب برای خم های فرنه در \mathbb{R}^3 ، κ و τ شکل تصویر خم را به طور کامل توصیف می کنند.

(۱۴-۴) به عنوان یک مثال، انحنا و خم مارپیچ دوار را که در بخش ۱۲ معرفی کردیم محاسبه می کنیم. یادآوری می کنیم که مارپیچ دوار $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$ به صورت $\gamma(t) = (a \cos t, a \sin t, bt)$ تعریف می شود که در آن $a > 0$ و b اعداد حقیقی داده شده اند. داریم $\gamma'(t) = (-a \sin t, a \cos t, b)$ پس $\frac{ds}{dt} = |\gamma'(t)| = \sqrt{a^2 + b^2}$ بدین ترتیب در این مثال پارامتر طول، $s = \sqrt{a^2 + b^2} t$ به سادگی

به دست می آید و می توان خم را برحسب s پرمایش کرد:

$$\tilde{\gamma}(s) = \left(a \cos \frac{s}{\sqrt{a^2 + b^2}}, a \sin \frac{s}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} s \right)$$

$$\vec{T}(s) = \frac{d\tilde{\gamma}}{ds} = \frac{1}{\sqrt{a^2 + b^2}} \left(-a \sin \frac{s}{\sqrt{a^2 + b^2}}, a \cos \frac{s}{\sqrt{a^2 + b^2}}, b \right)$$

$$\frac{d\vec{T}}{ds} = \frac{1}{a^2 + b^2} \left(-a \cos \frac{s}{\sqrt{a^2 + b^2}}, -a \sin \frac{s}{\sqrt{a^2 + b^2}}, 0 \right)$$

پس

$$\kappa = \frac{a}{a^2 + b^2}, \quad \vec{N} = \left(-\cos \frac{s}{\sqrt{a^2 + b^2}}, -\sin \frac{s}{\sqrt{a^2 + b^2}}, 0 \right)$$

بنابراین وقتی $a > 0$ ، یک خم فرنه به دست می آید. حال به محاسبه \vec{B} و τ می پردازیم:

$$\vec{B} = \vec{T} \times \vec{N} = \frac{1}{\sqrt{a^2 + b^2}} \left(b \sin \frac{s}{\sqrt{a^2 + b^2}}, -b \cos \frac{s}{\sqrt{a^2 + b^2}}, a \right)$$

برای محاسبه τ ، از $\frac{d\vec{B}}{ds} = -\tau \vec{N}$ استفاده می کنیم:

$$\frac{d\vec{B}}{ds} = \frac{b}{a^2 + b^2} \left(\cos \frac{s}{\sqrt{a^2 + b^2}}, \sin \frac{s}{\sqrt{a^2 + b^2}}, 0 \right)$$

پس

$$\tau = \frac{b}{a^2 + b^2}$$

توجه کنید که τ با b هم علامت است. وقتی $\tau > 0$ ، جهت γ به طرف بالا، و وقتی $\tau < 0$ جهت γ به سوی پایین است. برای $b = 0$ ، تصویر خم γ در صفحه xy می ماند و $\tau = 0$. به طور کلی، برای خم های فرنه، می توان ثابت کرد که $\tau > 0$ به مفهوم تابیدن خم خارج از صفحه بوسان در طرف \vec{B} است و $\tau < 0$ تابیدن در طرف مقابل \vec{B} را نشان می دهد.

(۱۴-۵) محاسبه κ و τ در حالت کلی در مثال بالا محاسبه پارامتر طول برحسب پارامتری که خم بر حسب آن ارائه شده بود آسان بود و توانستیم از آغاز خم را برحسب s پرمایش کنیم. این یک وضعیت استثنایی است. بسیاری اوقات نمی توان عبارتهای ساده و قابل استفاده ای برای بازپرمایش خم برحسب طول به دست آورد. در حالت کلی باید با استفاده از قاعده زنجیره ای و مشتق گیری نسبت به

پارامتر داده شده به نتایج مورد نظر رسید. بخشی از این کار در فرمول (۶) انجام شده است. اگر ضرب خارجی دو طرف (۱) را با بردار سرعت $\gamma'(t)$ محاسبه کنیم، نتیجه می‌شود:

$$\gamma'(t) \times \gamma''(t) = \left(\frac{d^2s}{dt^2}\right) \gamma'(t) \times \vec{T} + \kappa \left(\frac{ds}{dt}\right)^2 \gamma'(t) \times \vec{N}$$

$\gamma'(t)$ و \vec{T} همراستا هستند، پس جمله اول طرف راست صفر است. از طرفی دیگر

$$\gamma'(t) \times \vec{N} = \left(\frac{ds}{dt}\right) \vec{T} \times \vec{N} = \left(\frac{ds}{dt}\right) \vec{B}$$

پس

$$\gamma'(t) \times \gamma''(t) = \kappa \left(\frac{ds}{dt}\right)^2 \vec{B} \quad (۷)$$

و بالاخره با در نظر گرفتن طول بردار در دو طرف داریم

$$\kappa = \frac{|\gamma'(t) \times \gamma''(t)|}{|\gamma'(t)|^2} \quad (۸)$$

این فرمول انحنای برحسب پارامتر دلخواه t است.

برای به دست آوردن فرمولی برای τ ، از (۱) نسبت به t مشتق می‌گیریم که در آن فقط محاسبه ضرب \vec{B} بعداً مورد نیاز قرار خواهد گرفت:

$$\gamma'''(t) = \left(\frac{d^2s}{dt^2}\right) \vec{T} + \left(\frac{ds}{dt}\right) \vec{N} + \kappa \left(\frac{ds}{dt}\right)^2 \tau \vec{B}$$

از محاسبه ضرب داخلی دو طرف با \vec{B} نتیجه می‌شود

$$\vec{B} \cdot \gamma'''(t) = \kappa \left(\frac{ds}{dt}\right)^2 \tau$$

و با جایگزینی از (۸) داریم:

$$\tau = \frac{(\gamma'(t) \times \gamma''(t)) \cdot \gamma'''(t)}{|\gamma'(t) \times \gamma''(t)|^2} \quad (۹)$$

ضمناً این فرمول وابستگی τ به مشتق سوم γ را نشان می‌دهد. این موضوع نباید تعجب آور باشد زیرا که تعریف \vec{N} ، مشتق دوم γ را می‌طلبید و τ در معادلات فرنه از مشتق‌گیری \vec{N} به دست می‌آید

نتایج اولیه مشتق پذیری

در جلسه قبل مفهوم مشتق و مشتق پذیری مورد بررسی قرار گرفت. در این جلسه نخست با بیان و اثبات یک گزاره در مورد آمیختن جبری توابع مشتق پذیر، چند دسته تابع مشتق پذیر معرفی می کنیم، سپس به ذکر پاره ای خواص ابتدایی مشتق می پردازیم.

(۱-۱۵) گزاره. فرض کنید تابع های $f, g : S \rightarrow \mathbb{R}$ در نقطه درونی x_0 از دامنه مشتق پذیرند. در این صورت:

الف) $(f + g)'(x_0) = f'(x_0) + g'(x_0)$ در x_0 مشتق پذیر است و

ب) (قانون لایب نیتس) $f \cdot g$ در x_0 مشتق پذیر است و $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$

ج) قرار دهید $S' = \{x \in S \mid g(x) \neq 0\}$. فرض کنید x_0 یک نقطه درونی S' است. در این صورت تابع $\frac{f}{g} : S' \rightarrow \mathbb{R}$ در x_0 مشتق پذیر است و

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$$

برهان. (الف) $\frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0}$ و حکم از این که حد مجموع برابر مجموع حدهاست نتیجه می شود.

(ب)

$$\begin{aligned} \frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} &= \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x)}{x - x_0} + \frac{f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \left(\frac{f(x) - f(x_0)}{x - x_0} \right) g(x) + f(x_0) \left(\frac{g(x) - g(x_0)}{x - x_0} \right) \end{aligned}$$

توجه کنید که طبق گزاره ۱۴-۸، چون g در x_0 مشتق پذیر است، g در x_0 پیوسته نیز هست، پس $g(x) \rightarrow g(x_0)$ وقتی $x \rightarrow x_0$. حکم از اینکه حد مجموع و حاصل ضرب برابر مجموع و حاصل ضرب حد است نتیجه می شود.

(ج)

$$\begin{aligned} \frac{\left(\frac{f}{g}\right)(x) - \left(\frac{f}{g}\right)(x_0)}{x - x_0} &= \frac{f(x)g(x_0) - f(x_0)g(x)}{(x - x_0)g(x)g(x_0)} \\ &= \left[\left(\frac{f(x) - f(x_0)}{x - x_0} \right) g(x_0) - f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \right] \frac{1}{g(x)g(x_0)} \end{aligned}$$

□ در اینجا نیز حکم از پیوستگی g در x_0 و قوانین حد مجموع و حاصل ضرب نتیجه می شود.

(۱۵-۲) تابع های گویا. نخست یک تابع چند جمله ای

$$p(x) = a_0 + a_1x + \dots + a_mx^m \quad (1)$$

را در نظر بگیرید. در مثال های ۱۴-۵-۱ و ۱۴-۵-۲ دیدیم که هر تک جمله ای a_kx^k ، مشتق پذیر است و فرمولی برای مشتق آن پیدا کردیم. پس با توجه به ۱۵-۱ (الف)، این چند جمله ای به ازای هر x مشتق پذیر است و در واقع

$$p'(x) = a_1 + 2a_2x + \dots + ma_mx^{m-1} \quad (2)$$

حال فرض کنید $q(x) = b_0 + b_1x + \dots + b_nx^n$ یک چند جمله ای دیگر باشد. از ۱۵-۱ (ج) نتیجه می شود که تابع گویای تعریف شده به صورت $f(x) = \frac{p(x)}{q(x)}$ روی دامنه $S' = \{x \in \mathbb{R} \mid q(x) \neq 0\}$ در همه نقاط دامنه خود مشتق پذیر است و می توان مشتق آن را به کمک ۱۵-۱ محاسبه کرد. توجه کنید که هر نقطه x_0 از S' یک نقطه درونی S' است زیرا بنابر پیوستگی q ، اگر $q(x_0) \neq 0$ ، آنگاه برای همه x های نزدیک x_0 نیز داریم $q(x) \neq 0$.

(۱۵-۳) تابع های مثلثاتی. در گزاره^{۱۰} ۱-۴ از بخش ۱۰ دیدیم که تابع های مثلثاتی \sin ، \cos ، \tan ، \cot ، \sec و \csc هر یک در دامنه^{۱۱} تعریف خود پیوسته هستند. اکنون نشان می دهیم این تابع ها در دامنه^{۱۲} تعریف خود مشتق پذیر نیز هستند و فرمول هایی برای مشتق آنها به دست می آوریم. با توجه به گزاره ۱۵-۱ کافی است مشتق پذیری سینوس و کسینوس ثابت شود زیرا چهار تابع دیگر به صورت خارج قسمت این تابع ها تعریف می شوند. بدین ترتیب نخست تابع سینوس را در نظر می گیریم:

$$\begin{aligned} \frac{\sin(x+h) - \sin x}{h} &= \frac{(\sin x)(\cos h) + (\cos x)(\sin h) - \sin x}{h} \\ &= (\sin x) \frac{\cos h - 1}{h} + (\cos x) \frac{\sin h}{h} \end{aligned}$$

در ۱۳-۷-۲ دو حد اساسی مثلثاتی $\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1$ و $\lim_{h \rightarrow 0} \frac{1 - \cos h}{h} = 0$ را ثابت کردیم. بنابراین

$$\lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} = \cos x$$

به بیان دیگر، تابع \sin به ازای هر x مشتق پذیر است و

$$\sin' x = \cos x \quad (3)$$

به همین ترتیب ثابت می شود که تابع \cos به ازای هر x مشتق پذیر است و

$$\cos' x = -\sin x \quad (4)$$

حال با استفاده از ۱۵-۱ (ج) ثابت می شود که توابع $\tan = \frac{\sin}{\cos}$ ، $\cot = \frac{\cos}{\sin}$ ، $\sec = \frac{1}{\cos}$ و $\csc = \frac{1}{\sin}$ به ازای هر x در دامنه^{۱۳} تعریف (یعنی به ازای x هایی که مخارج عبارت تعریف کننده صفر نشود) مشتق پذیرند و فرمول های زیر به سادگی نتیجه می شوند:

$$\tan' x = \sec^2 x = 1 + \tan^2 x \quad (5)$$

$$\cot' x = -\csc^2 x = -(1 + \cot^2 x) \quad (6)$$

$$\sec' x = (\tan x)(\sec x) \quad (7)$$

$$\csc' x = -(\cot x)(\csc x) \quad (۸)$$

در باقیمانده این بخش به بررسی دسته‌ای از خواص مشتق می‌پردازیم که به علامت مشتق و اندازه آن بستگی دارند.

(۱۵-۴) علامت مشتق در یک نقطه

فرض کنید $f: S \rightarrow \mathbb{R}$ در نقطه درونی x_0 از S مشتق پذیر باشد. سه حالت $f'(x_0) > 0$ ، $f'(x_0) < 0$ و $f'(x_0) = 0$ وجود دارد که هر یک را بررسی می‌کنیم.

نخست فرض کنید $f'(x_0) > 0$. اگر عددی e طوری اختیار کنیم که $0 < e < f'(x_0)$ ، طبق

تعریف حد، $0 < \delta$ وجود دارد که برای $0 < |x - x_0| < \delta$ داریم:

$$-e < \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) < e$$

بالاخص

$$\frac{f(x) - f(x_0)}{x - x_0} > f'(x_0) - e > 0 \quad (۹)$$

بنابراین صورت و مخرج کسر سمت چپ باید هم علامت باشند و می‌توان حکم کرد که:

(۱۵-۴-۱) اگر $f'(x_0) > 0$ ، آنگاه $0 < \delta$ وجود دارد که اگر $x_0 < x < x_0 + \delta$ ، آنگاه

$$f(x) > f(x_0) \text{، و اگر } x_0 - \delta < x < x_0 \text{، آنگاه } f(x) < f(x_0).$$

به بیان دیگر اگر $f'(x_0) > 0$ ، آنگاه برای x های نزدیک و بزرگتر از x_0 ، مقدار $f(x)$ بزرگتر از $f(x_0)$ است و برای x های نزدیک و کوچکتر از x_0 ، $f(x)$ کوچکتر از $f(x_0)$ می‌باشد. نکته قابل تذکر این است که این حکم فقط مقدار $f(x_0)$ را با مقدار $f(x)$ ، نزدیک x_0 ، مقایسه می‌کند و دال بر صعودی بودن تابع f در یک بازه کوچک حول x_0 نیست. شکل ۱ وضعیتی را نشان می‌دهد که حکم ۱-۴-۱۵ برقرار است ($f'(0) > 0$) ولیکن f در هیچ بازه حول 0 صعودی نیست.

در این شکل نمودار تابع در نزدیکی \circ بی نهایت "دندانه" یا شاخه صعودی- نزولی دارد که دامنه آنها به تدریج کوچک می شود ولی هر قدر هم که به \circ نزدیک شویم هنوز شاخه های نزولی و صعودی در دو طرف \circ وجود دارند. در آینده عبارت صریحی برای تعریف چنین تابعی ارائه خواهیم کرد.

حالت $\circ < f'(x_0)$ مشابه است. در اینجا اگر e را طوری بگیریم که $\circ < e < -f'(x_0)$ ، آنگاه $\circ > \delta$ وجود دارد که برای $\delta < |x - x_0| < \circ$ داریم:

$$-e < \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) < e$$

بلاخص

$$\frac{f(x) - f(x_0)}{x - x_0} < f'(x_0) + e < \circ \quad (10)$$

بنابراین صورت و مخرج کسر سمت چپ باید علامت مختلف داشته باشند که از آن نتیجه می شود:

(۱۵-۴-۲) اگر $\circ < f'(x_0)$ ، آنگاه $\circ > \delta$ وجود دارد که اگر $x_0 < x < x_0 + \delta$ ، آنگاه

$$f(x) < f(x_0) \text{، و اگر } x_0 - \delta < x < x_0 \text{، آنگاه } f(x) > f(x_0)$$

در اینجا اگر x نزدیک و بزرگتر از x_0 باشد، داریم $f(x) < f(x_0)$ ، و اگر x نزدیک و کوچکتر از x_0

$$\text{باشد، } f(x) > f(x_0)$$

از ۱۵-۴-۱ و ۱۵-۴-۲ نتیجه جالب توجهی حاصل می شود. برای یک تابع $f: S \rightarrow \mathbb{R}$ ،

نقطه درونی x_0 از S را یک نقطه بیشینه موضعی (به ترتیب نقطه کمینه موضعی) می نامیم اگر

$\circ > \delta$ وجود داشته باشد که برای هر $x \in S$ که $|x - x_0| < \delta$ داشته باشیم $f(x) \leq f(x_0)$ (به ترتیب

$f(x) \geq f(x_0)$). حال فرض کنید تابع f در نقطه بیشینه یا کمینه موضعی x_0 مشتق پذیر است. در

این صورت هیچ یک از دو وضعیت $\circ > f'(x_0)$ و $\circ < f'(x_0)$ در x_0 ممکن نیست زیرا که بنابر

۱۵-۴-۱ و ۱۵-۴-۲، مقدار تابع باید در یک طرف x_0 بزرگتر از $f(x_0)$ و در طرف دیگر کوچکتر

از $f(x_0)$ باشد. بدین ترتیب لاجرم:

(۱۵-۴-۳) اگر تابع $f: S \rightarrow \mathbb{R}$ در نقطه درونی x_0 از S بیشینه یا کمینه موضعی داشته باشد و f

در x_0 مشتق پذیر باشد، آنگاه $\circ = f'(x_0)$.

بدین ترتیب در نقاط بیشینه و کمینه موضعی درونی که تابع دارای خط مماس باشد، این خط مماس باید لزوماً افقی باشد (شکل ۲).

لازم به ذکر است که افقی شدن خط مماس لزوماً دال بر وجود بیشینه یا کمینه موضعی نیست. در شکل ۲، در نقاط x_1 و x_3 کمینه موضعی موجود است، در x_2 بیشینه موضعی، ولی در x_4 که خط مماس افقی است، نه بیشینه موضعی ظاهر شده است و نه کمینه موضعی. مثال‌های صریح زیر در تأیید این مطلب هستند.

مثال ۱. تابع $f: \mathbb{R} \rightarrow \mathbb{R}$ به صورت $f(x) = x^4 - x^3$ تعریف شده است. این تابع به ازای هر x مشتق پذیر است. نقاطی را که در آن مشتق صفر می‌شود بررسی می‌کنیم. داریم:

$$f'(x) = 4x^3 - 3x^2 = x^2(4x - 3)$$

پس مشتق در دو نقطه $x = 0$ و $x = \frac{3}{4}$ صفر می‌شود. با توجه به علامت یابی $x^4 - x^3 = x^3(x - 1)$ ملاحظه می‌شود که مقدار f برای $x < 0$ مثبت و برای $0 < x < 1$ منفی است، پس $x = 0$ نمی‌تواند بیشینه یا کمینه موضعی باشد. چون تابع پیوسته $f(x) = x^4 - x^3$ روی $[0, 1]$ باید دارای کمینه باشد و مقدار تابع در $[0, 1]$ منفی است، مقدار کمینه باید منفی باشد. بنابراین این نقطه کمینه باید یک نقطه درونی $[0, 1]$ باشد زیرا که $f(0) = f(1) = 0$. از طرفی دیگر مشتق در کمینه درونی باید صفر باشد، پس نقطه $x = \frac{3}{4}$ لزوماً یک کمینه است. نمودار f در شکل ۳ (الف) نمایش داده شده است.

مثال ۲. تابع مشتق پذیر $f(x) = x - \sin x$ را در نظر می‌گیریم. داریم $f'(x) = 1 - \cos x$ که در مضارب صحیح 2π صفر می‌شود. هیچ یک از این نقاط بیشینه یا کمینه موضعی برای تابع نیست (شکل ۳ (ب)).

در گام بعدی به بررسی مثبت یا منفی بودن علامت مشتق در سراسر یک بازه می‌پردازیم. حربه مناسب برای این کار "قضیه میانگین" است که کاربردهای بسیار دیگری نیز خواهد داشت. نخست حالت خاصی از این قضیه را که به قضیه رُل معروف است بیان و ثابت می‌کنیم.

(۱۵-۵) قضیه رُل. فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ تابعی پیوسته است که در همه نقاط درونی $[a, b]$ مشتق پذیر می باشد و $f(a) = f(b) = 0$. در این صورت نقطه‌ای c وجود دارد، $a < c < b$ ، که $f'(c) = 0$.

برهان. اگر f در سراسر $[a, b]$ صفر باشد که مشتق آن در هر نقطه صفر است و نقطه c مورد نظر وجود دارد. حال فرض کنید نقطه‌ای x در $[a, b]$ وجود دارد که $f(x) \neq 0$ ، مثلاً فرض کنید $f(x) > 0$. تابع پیوسته f روی $[a, b]$ دارای بیشینه است و از آنجا که f در حداقل یک نقطه مثبت است، مقدار این بیشینه باید مثبت باشد. از طرفی دیگر $f(a) = f(b) = 0$ ، پس نقطه بیشینه باید یک نقطه درونی بازه باشد، مثلاً c که $a < c < b$. حال طبق ۱۵-۴-۳ داریم $f'(c) = 0$. به همین ترتیب اگر $f(x) < 0$ ، با استفاده از کمینه، نقطه مورد نظر را پیدا می کنیم. □

یک تعبیر قضیه بالا این است که نقطه‌ای c بین a و b وجود دارد که مماس بر نمودار به ازای c موازی خط واصل بین دو انتهای نمودار است. قضیه زیر را می توان دقیقاً این گونه تعبیر کرد.

(۱۵-۶) قضیه میانگین. فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ تابعی پیوسته است که در همه نقاط درونی $[a, b]$ مشتق پذیر می باشد. در این صورت نقطه‌ای c وجود دارد، $a < c < b$ ، که

$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad (11)$$

برهان. خط راست واصل بین $(a, f(a))$ و $(b, f(b))$ معادله زیر را دارد:

$$y = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

اگر مقدار سمت راست را از $f(x)$ کم کنیم در وضعیت قضیه رُل قرار می گیریم. به طور دقیق، تعریف کنید

$$g(x) = f(x) - \left[f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right]$$

تابع g در $[a, b]$ پیوسته و در $[a, b]$ مشتق پذیر است زیرا که مجموع دو تابع با این ویژگی هاست. از طرفی دیگر:

$$g(a) = 0, \quad g(b) = 0$$

پس طبق قضیه رل نقطه‌ای c وجود دارد $a < c < b$ که $g'(c) = 0$ ، یعنی:

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0$$

□ که حکم قضیه است.

به یک تعبیر هندسی این قضیه اشاره کردیم. اگر متغیر x را زمان و $y = f(x)$ را مکان یک ذره متحرک در زمان x فرض کنیم، $\frac{f(b)-f(a)}{b-a}$ سرعت متوسط ذره در بازه زمانی $[a, b]$ است. طبق این قضیه، زمانی c بین زمان شروع و زمان پایان حرکت وجود دارد که سرعت ذره در آن زمان برابر سرعت متوسط در طول مسیر است. در واقع مهم‌ترین کاربردهای ۱۵-۶ به صورت نامساوی برای تخمین نمود یک تابع خواهد بود که بعداً به آن خواهیم پرداخت ولی فعلاً به چند کاربرد در تکمیل بررسی علامت مشتق می پردازیم.

(۱۵-۷) علامت مشتق در یک بازه

فرض کنید I یک بازه باشد و $f : I \rightarrow \mathbb{R}$ یک تابع پیوسته که در نقاط درونی بازه مشتق پذیر است.

(۱۵-۷-۱) اگر $f'(x) = 0$ برای هر نقطه درونی x از I ، آنگاه f در سراسر I ثابت است.

برهان. کافی است نشان دهیم برای هر دو نقطه متمایز a و b از I داریم $f(a) = f(b)$. مثلاً فرض کنید $a < b$. طبق قضیه میانگین نقطه‌ای c بین a و b وجود دارد که $f'(c) = \frac{f(b)-f(a)}{b-a}$ و حکم به اثبات می رسد.

(۱۵-۷-۲) نتیجه. فرض کنید $f, g : I \rightarrow \mathbb{R}$ دو تابع پیوسته باشند که در نقاط درونی I مشتق پذیر

□ بوده و مشتق برابر دارند. در این صورت $f - g$ یک ثابت است.

(۱۵-۷-۳) اگر $f'(x) > 0$ برای هر نقطه درونی x از I ، آنگاه f در I صعودی است، یعنی برای هر a و b در I که $a < b$ ، داریم $f(a) < f(b)$.

برهان. طبق قضیه میانگین $f'(c) = \frac{f(b)-f(a)}{b-a} > 0$ ، پس $b-a$ و $f(b)-f(a)$ هم علامت هستند. □

(۱۵-۷-۴) اگر $f'(x) < 0$ برای هر نقطه درونی x در I ، آنگاه f در I نزولی است، یعنی برای هر a و b در I که $a < b$ ، داریم $f(a) > f(b)$. □

(۱۵-۸) تخمین نمو تابع. فرض کنید $f: [a, b] \rightarrow \mathbb{R}$ پیوسته و f در نقاط درونی $[a, b]$ مشتق پذیر است. اگر $M \geq 0$ وجود داشته باشد که $|f'(x)| \leq M$ برای هر x در $[a, b]$ ، آنگاه از (۱۱) نتیجه می شود که:

$$|f(b) - f(a)| \leq M|b - a| \quad (12)$$

از آنجا که $f'(x)$ آهنگ تغییر کمیت $y = f(x)$ در نقطه x محسوب می شود. نامساوی (۱۲) بیانگر این واقعیت است که نمو y در بازه $[a, b]$ از حاصل ضرب طول بازه در حداکثر آهنگ نمو بیشتر نیست. گاهی نمو x یعنی $b - a$ را به Δx و نمو y ، یعنی $f(b) - f(a)$ را به Δy نمایش می دهند. پس با این نمادگذاری:

$$|\Delta y| \leq M|\Delta x| \quad (13)$$

مثال ۱. نشان دهید برای هر α و β داریم

$$|\sin \alpha - \sin \beta| \leq |\alpha - \beta| \quad (14)$$

از آنجا که برای $f(x) = \sin x$ ، داریم $f'(x) = \cos x$ و $|\cos x| \leq 1$ ، این حکم از (۱۲) نتیجه می شود.

مثال ۲. نشان دهید برای هر a و b مثبت داریم:

$$\left| \frac{1}{1+a} - \frac{1}{1+b} \right| \leq |a - b| \quad (15)$$

تابع $f(x) = \frac{1}{1+x}$ را روی $]-1, +\infty[$ در نظر می‌گیریم. در این بازه تابع تعریف شده، مشتق پذیر است،
و داریم:

$$f'(x) = \frac{-1}{(1+x)^2}$$

برای $x > 0$ ، مخرج از ۱ بزرگتر است؛ پس $|f'(x)| < 1$ و (۱۵) حاصل می‌شود.

مجموعه‌های تراز (۱)

به دنبال معرفی کلی مجموعه‌های تراز، اکنون مثال‌هایی مطرح می‌کنیم.

مثال ۱. فرض کنید $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ یک نگاشت مستوی باشد. A را می‌توان به صورت زیر نوشت:

$$A(x) = L(x) + B \quad x \in \mathbb{R}^n \quad \text{برای هر} \quad (1)$$

که در آن L یک نگاشت خطی از \mathbb{R}^n به \mathbb{R}^m است و B یک عنصر ثابت \mathbb{R}^m . مجموعه‌های تراز A و L یکی هستند، در واقع

$$A^{-1}(c + B) = L^{-1}(c)$$

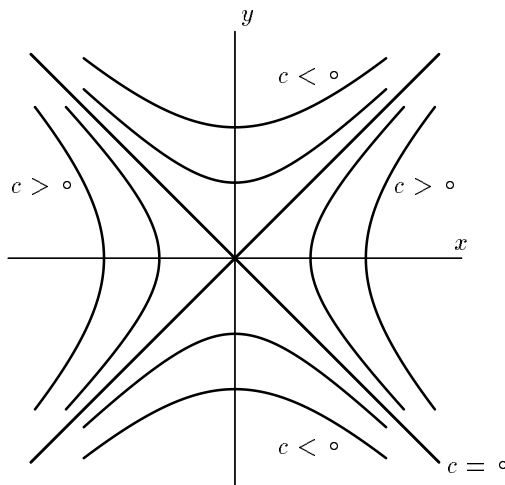
یعنی مجموعه تراز منسوب به c برای L همان مجموعه تراز منسوب به $c + B$ برای A است. برای تابع خطی L می‌دانیم که مجموعه‌های تراز یا تهی هستند و یا زیرفضاهای خطی موازی و همبند هستند L می‌باشند.

مثال ۲. $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ را به صورت $f(x, y, z) = x^2 + y^2 + z^2$ تعریف می‌کنیم. مجموعه‌های تراز f را برای مقادیر مختلف $c \in \mathbb{R}$ بررسی می‌کنیم. برای $c < 0$ ، مجموعه تراز منسوب تهی است؛ برای $c = 0$ ، تک نقطه‌ای $\{(0, 0, 0)\}$ حاصل می‌شود؛ و برای $c > 0$ ، کره شعاع \sqrt{c} ، یعنی مجموعه نقاط (x, y, z) بقیه به شرط $x^2 + y^2 + z^2 = c$ به دست می‌آید.

مثال ۳. تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را در نظر می‌گیریم که به صورت $f(x, y) = x^2 - y^2$ تعریف شده است. نمودار این تابع را در مثال ۴ بخش قبل بررسی کردیم، اکنون به بررسی مجموعه‌های تراز می‌پردازیم. برای $c = 0$ ، مجموعه

$$\{(x, y) \mid x^2 - y^2 = 0\} = \{(x, y) \mid x - y = 0 \text{ یا } x + y = 0\}$$

از دو خط راست $x + y = 0$ و $x - y = 0$ تشکیل شده است. برای $c > 0$ ، هذلولوی دو شاخه از $x^2 - y^2 = c$ به دست می‌آید که رئوس شاخه‌ها در نقاط $(\pm\sqrt{c}, 0)$ قرار دارند و شاخه‌ها به طرف راست و چپ باز می‌شوند. برای $c < 0$ ، نقش x و y تعویض می‌شود. این بار دو رأس هذلولوی در $(0, \pm\sqrt{c})$ قرار دارند و هذلولوی به طرف بالا و پایین باز می‌شود (شکل ۱).



شکل ۱

ارتباط این شکل با نمودار $z = x^2 - y^2$ به طریق زیر حاصل می‌شود. به ازای $z = c$ با تقاطع صفحه $z = c$ با نمودار متشکل از نقاط به شکل (x, y, c) است. اگر این مقطع را بر صفحه xy تصویر کنیم دقیقاً مجموعه تراز منسوب به c حاصل می‌شود.

مطلب فوق برای هر تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ معتبر است. اگر نمودار تابع را با صفحه $z = c$ قطع کنیم دقیقاً آن بخش نمودار به دست می‌آید که بالا سر مجموعه تراز منسوب به c قرار دارد. بدین ترتیب اگر

نمودار تابع در دست باشد با قطع کردن آن به وسیله صفحات $z = c$ و تصویر کردن روی صفحه xy ، مجموعه‌های تراز f حاصل می‌شوند. بالعکس چنانچه مجموعه‌های تراز رسم شده باشند، با انتقال هر مجموعه تراز $f^{-1}(c)$ به ارتفاع $z = c$ ، نمودار تابع بازسازی می‌شود.

مثال ۴. فرض کنید $a \in \mathbb{R}^n$ داده شده باشد که $a \neq \underline{0}$. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$f(x) = a \cdot x \quad (2)$$

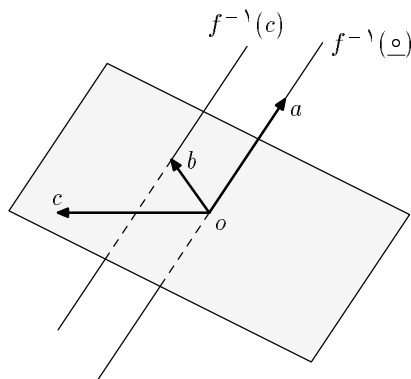
چون $a \cdot x = a_1 x_1 + \dots + a_n x_n$ ، این در واقع یک تابع خطی است و در مثال ۱ می‌گنجد ولی بررسی مستقیم آن نیز قابل توجه است. می‌دانیم که مجموعه‌های x هایی که در $a \cdot x = c$ صدق می‌کنند یک ابرصفحه در \mathbb{R}^n است و این ابرصفحه‌ها همه موازی $a \cdot x = 0$ (هسته f) مجموعه‌های x هایی که $x \perp a$ می‌باشند.

مثال ۵. عضوی ثابت $a \neq \underline{0}$ در \mathbb{R}^3 در نظر می‌گیریم و $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ را به صورت زیر تعریف می‌کنیم:

$$f(x) = a \times x \quad (3)$$

که مقصود از $a \times x$ ضرب برداری a و x است. تابع (۳) نیز خطی است (هر مؤلفه عبارت درجه ۱ همگن نسبت به x_1, x_2, x_3 است) و در چارچوب مثال ۱ قرار می‌گیرد. این مثال را نیز مستقیماً بررسی می‌کنیم. از آنجا که $a \times x$ بر a عمود است، اگر $c \in \mathbb{R}^3$ در صفحه گذرا از $\underline{0}$ و عمود بر a قرار نداشته باشد، مجموعه تراز $f^{-1}(c)$ تهی است. برای $c = \underline{0}$ ، مجموعه‌های x هایی که در $a \times x = \underline{0}$ صدق می‌کنند مضارب بردار a هستند، یعنی نقاط روی خطی که از امتداد a پدید می‌آید. این در واقع هسته تابع خطی f است. پس سایر مجموعه‌های تراز ناتهی خطوط راست موازی a می‌باشند. در واقع

اشتراک $f^{-1}(c)$ با صفحه عمود بر a باید یگانه بردار b در این صفحه باشد که (a, b, c) یک سه تایی راستگرد است و مساحت مستطیل تعیین شده توسط a و b برابر طول c می باشد.



شکل ۲

مثال ۶. تابع $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ را به صورت

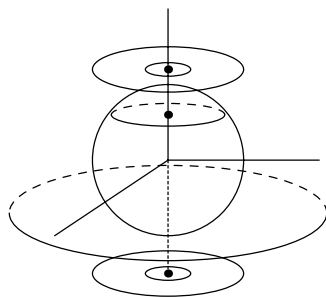
$$f(x, y, z) = (z, x^2 + y^2 + z^2)$$

تعریف می کنیم. به ازای $c = (c_1, c_2) \in \mathbb{R}^2$ ، مجموعه تراز $f^{-1}(c)$ عبارت است از:

$$\{(x, y, z) \mid z = c_1\} \cap \{(x, y, z) \mid x^2 + y^2 + z^2 = c_2\}$$

برای $c_2 < 0$ ، مجموعه سمت راست و در نتیجه $f^{-1}(c)$ تهی است. برای $c_2 = 0$ ، مجموعه سمت راست از تک نقطه ای $\{(0, 0, 0)\}$ تشکیل شده است؛ پس $f^{-1}(c_1, 0)$ تهی است مگر وقتی که $c_1 = 0$ و $f^{-1}(0, 0)$ نیز تک نقطه ای $\{(0, 0, 0)\}$ است. حال فرض کنید $c_2 > 0$. مجموعه سمت راست؛ یعنی $f^{-1}(c_2)$ کره شعاع $\sqrt{c_2}$ به مرکز 0 است. برای این که مجموعه سمت چپ؛ یعنی $f^{-1}(c_1)$ با $f^{-1}(c_2)$ اشتراک ناتهی داشته باشد؛ شرط لازم و کافی این است که $-\sqrt{c_2} \leq c_1 \leq \sqrt{c_2}$. وقتی $c_1 = \pm\sqrt{c_2}$ ، صفحه $z = \pm\sqrt{c_2}$ بر کره مماس می شود و یک تک

نقطه‌ای به عنوان مجموعه تراز حاصل می‌شود $(0, 0, \pm\sqrt{c_2})$ بسته به این که صفحه $z = \sqrt{c_2}$ یا صفحه $z = -\sqrt{c_2}$ در نظر گرفته شود). برای $-\sqrt{c_2} < c_1 < \sqrt{c_2}$ ، دایره تقاطع کره با صفحه $z = c_1$ است. بدین ترتیب مجموعه‌های تراز ناتهی تابع f عبارتند از تک نقطه‌ای‌های روی محور z و دایره‌های به مرکز روی محور z که در صفحات افقی $z = c$ قرار دارند.



شکل ۳

در اینجا مسأله زیر را مطرح می‌کنیم. در مورد تابع‌های مستوی دیدیم که مجموعه‌های تراز شکل‌های ساده و یکنواختی دارند. به عنوان گام بعدی، آیا می‌توان حکمی کلی در مورد مجموعه‌های تراز یک تابع درجه ۲ عنوان کرد؟ تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$f(x_1, \dots, x_n) = a_0 + a_1x_1 + \dots + a_nx_n + \sum_{i,j=1}^n a_{ij}x_ix_j \quad (4)$$

که در اینجا a_i ها و a_{ij} ها اعداد حقیقی داده شده‌اند. طرف راست (۴) کلی‌ترین عبارت درجه دوم برحسب x_1, \dots, x_n است. سؤال این است که مجموعه‌های تراز تابعی به شکل f در (۴) چگونه اشکالی هستند. در آینده خواهیم دید که با "دوران" مناسب در فضای \mathbb{R}^n می‌توان فقط صورت‌هایی از (۴) را در نظر گرفت که در آن جملات مخلوط درجه دوم، یعنی x_ix_j با $i \neq j$ ، وجود نداشته باشند، یعنی کافی است تابع‌های درجه دوم به شکل زیر را در نظر بگیریم:

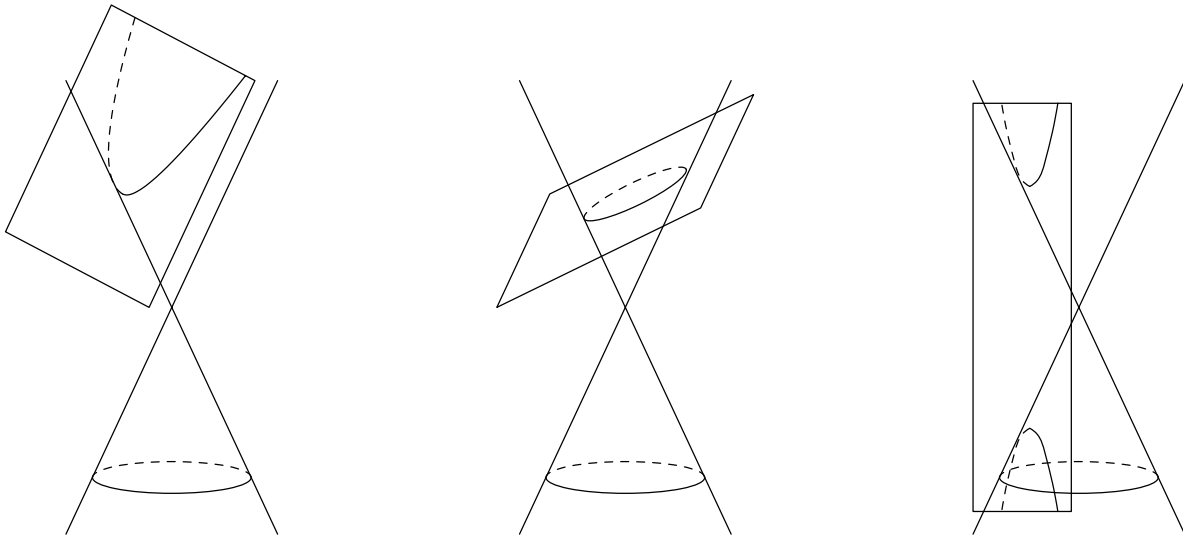
$$f(x_1, \dots, x_n) = a_0 + a_1x_1 + \dots + a_nx_n + a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{nn}x_n^2 \quad (5)$$

در این بخش و بخش بعد دو حالت $n = 2$ و $n = 3$ را توصیف می‌کنیم.

در حالت $n = 2$ اگر به جای x_1 و x_2 نمادهای معمولتر x و y را به کار گیریم، داریم

$$f(x, y) = a_0 + a_1x + a_2y + a_{11}x^2 + a_{22}y^2 \quad (6)$$

مجموعه‌های تراز چنین تابعی به مقاطع مخروطی معروفند به این دلیل که از نظر هندسی این منحنی‌ها را می‌توان به صورت مقطع یک صفحه و یک مخروط به دست آورد (شکل ۴). در واقع مقاطع مخروطی از دوران باستان شناخته شده بودند و کتاب مخروطات آپولونیوس که بیش از ۲۰۰ سال قبل از میلاد نوشته شده رسالهٔ جامعی در بررسی این اشکال است. سه شکل "سه‌می"، "بیضی" و "هذلولی" که در شکل نمایش داده شده‌اند، سه نوع عمدهٔ مقاطع مخروطی هستند.



شکل ۴

حال برای f در (۶) می‌خواهیم مجموعهٔ (x, y) هایی را که در $f(x, y) = 0$ صدق می‌کنند بررسی کنیم. در (۶) می‌توانیم قرار دهیم $a_0 = 0$ زیرا با جذب a_0 در c مجموعه‌های تراز $f(x, y) = c$ همهٔ حالات ممکن را در بر می‌گیرد. چند حالت در نظر می‌گیریم:

الف) $a_{11} = a_{22} = 0$. در این حالت در وضعیت تابع‌های مستوی قرار می‌گیریم که قبلاً بررسی شده‌اند.

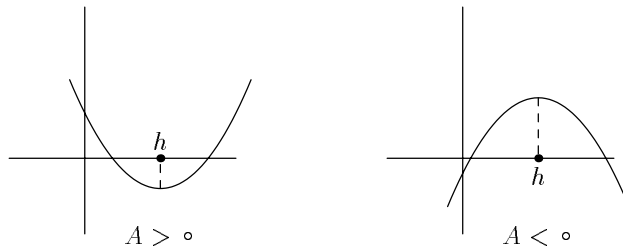
ب) یکی از a_{11} و a_{22} صفر و دیگری ناصفر. مثلاً فرض کنید $a_{11} \neq 0$ و $a_{22} = 0$. در این صورت با تکمیل مجذور، f به شکل زیر در می آید:

$$f(x, y) = A_0 + a_{11}(x - h)^2 + a_{22}y \quad (7)$$

اگر $a_{22} \neq 0$ ، با قرار دادن $f(x, y) = 0$ و تقسیم بر a_{22} ، عبارتی به شکل:

$$y = A(x - h)^2 + B, \quad A = \frac{a_{11}}{a_{22}} \neq 0$$

به دست می آید. که سهمی نام دارد. بسته به این که $A > 0$ یا $A < 0$ یکی از دو منحنی نمایش داده شده در شکل ۵ حاصل می شود.



شکل ۵

به عنوان مقطع مخروطی، سهمی از تقاطع یک صفحه موازی یکی از خطوط مولد مخروط به دست می آید. اگر در (۷) داشته باشیم $a_{22} = 0$ ، آنگاه $f(x, y) = 0$ ، بسته به علامت A_0 و a_{11} ، دو خط راست قائم، یک خط راست یا تهی است.

حال فرض کنید a_{11} و a_{22} هر دو ناصفر باشند. در این صورت با تکمیل مجذورها در (۶) به عبارت زیر می رسیم:

$$f(x, y) = a_{11}(x - h)^2 + a_{22}(y - k)^2 + A_0 \quad (8)$$

در اینجا دو حالت در نظر می گیریم:

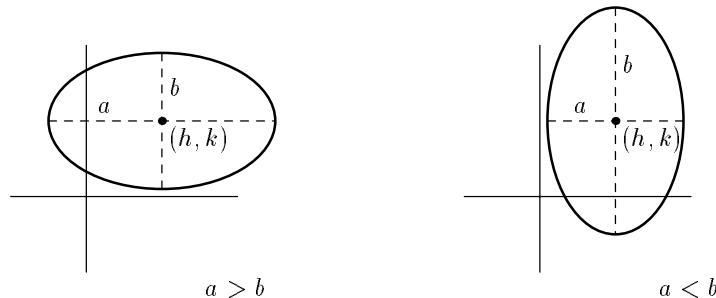
ج) $a_{۲۲}$ و $a_{۱۱}$ همعلامت باشند. در این صورت اگر A_0 نیز همعلامت با $a_{۱۱}$ و $a_{۲۲}$ باشد، $f(x, y)$ همواره مثبت یا همواره منفی است؛ یعنی در $f(x, y) = 0$ هیچ نقطه‌ای صدق نمی‌کند. اگر $A_0 = 0$ ، آنگاه تنها نقطه (h, k) در $f(x, y) = 0$ صدق می‌کند. بالاخره اگر علامت A_0 از علامت $a_{۱۱}$ و $a_{۲۲}$ متفاوت باشد، در $f(x, y) = 0$ ، A_0 را به طرف دیگر معادله برده بر $-A_0$ تقسیم می‌کنیم:

$$\frac{(x-h)^2}{-A_0/a_{۱۱}} + \frac{(y-k)^2}{-A_0/a_{۲۲}} = 1$$

از آنجا که علامت A_0 با علامت $a_{۱۱}$ و $a_{۲۲}$ متفاوت بود، هر یک از دو مخرج بالا مثبت است و می‌توان هر یک را به صورت یک مجذور نوشت، پس داریم:

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \quad (9)$$

این معادله یک بیضی به مرکز نقطه (h, k) است (شکل ۶).



شکل ۶

د) $a_{۲۲}$ و $a_{۱۱}$ مختلف‌العلامه باشند. در این صورت نیز اگر $A_0 \neq 0$ ، با محاسبه‌ای عیناً مانند فوق

به یکی از دو عبارت زیر می‌رسیم (بسته به این که A_0 با کدامیک از $a_{۱۱}$ و $a_{۲۲}$ همعلامت باشد):

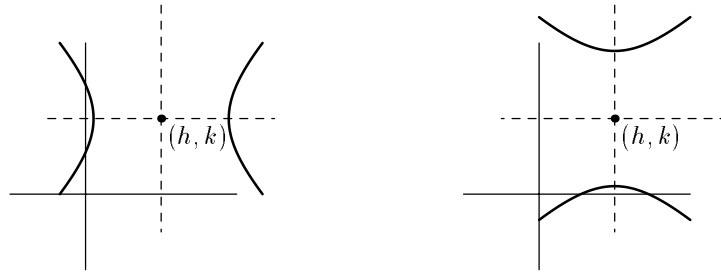
$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \quad (10)$$

یا

$$-\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \quad (11)$$

هر یک از این دو منحنی یک هذلولی نام دارد. هر هذلولی از دو شاخه قرینه تشکیل شده است (شکل

(۷)



شکل ۷

بالاخره در این حالت اگر $A_0 = 0$ ، آنگاه $f(x, y) = 0$ را می توان به صورت زیر نوشت

$$a^2(x-h)^2 - b^2(y-k)^2 = 0 \quad (12)$$

که در آن $a^2 = |a_{11}|$ و $b^2 = |a_{22}|$. این نمایش جبری دو خط راست متقاطع با ضریب زاویه های $\pm \frac{a}{b}$ است که از نقطه (h, k) می گذرند. ضمناً هر یک از شاخه های هذلولی های (۱۰) و (۱۱) بر دو خط راست (۱۲) مجانبند.

مجموعه‌های تراز (۲)

به بررسی مجموعه‌های تراز تابع‌های درجه دو ادامه می‌دهیم. این بار تابع‌های درجه دو $\mathbb{R} \rightarrow \mathbb{R}^3$ را در نظر می‌گیریم. همان طور که در جلسه قبل اشاره کردیم، می‌توان با دوران محورهای مختصات جملات مخلوط را حذف و f را به صورت زیر نمایش داد:

$$f(x, y, z) = a_0 + a_1x + a_2y + a_3z + a_{11}x^2 + a_{22}y^2 + a_{33}z^2 \quad (1)$$

مجموعه‌های تراز چنین تابعی رویه‌های درجه دوم خوانده می‌شوند. در زیر تعدادی مثال بررسی می‌کنیم.

مثال ۱. (بیضی وار) رویه

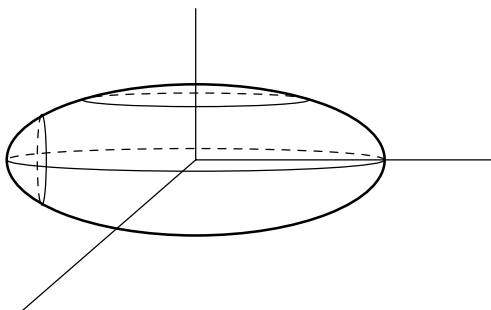
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad a, b, c > 0 \quad (2)$$

بیضی وار نام دارد. برای دستیابی به شکل این رویه در \mathbb{R}^3 ، تقاطع آن را با صفحات موازی صفحات مختصاتی بررسی می‌کنیم. مثلاً اشتراک (۲) را با صفحه $z = k$ در نظر بگیرید:

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{k^2}{c^2} \\ z = k \end{cases} \quad (3)$$

اشتراک به ازای $|k| > c$ تهی است، به ازای $k = \pm c$ ، دو نقطه $(0, 0, \pm c)$ به دست می‌آیند، و برای $|k| < c$ ، با تقسیم دو طرف رابطه اول (۳) بر $1 - \frac{k^2}{c^2}$ یک بیضی در صفحه $z = k$ به دست می‌آوریم.

عیناً وضعیت مشابه برای اشتراک‌های $x = k$ و $y = k$ وجود دارد. رویه (۲) در شکل ۱ نمایش داده شده است.

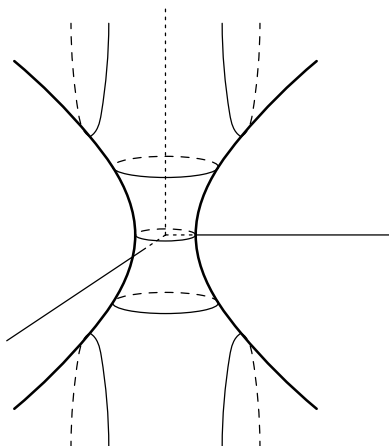


شکل ۱

مثال ۲. (هدلولی وار یکپارچه) رویه زیر را در نظر می‌گیریم که هدلولی وار یکپارچه نام دارد:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad a, b, c > 0 \quad (4)$$

مجدداً روش بررسی شکل این رویه، یافتن تقاطع رویه با صفحات موازی صفحات مختصاتی است. برای $z = k$ بیضی $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{k^2}{c^2}$ در صفحه $z = k$ به دست می‌آید. کوچکترین بیضی به‌ازای $k = 0$ در صفحه xy حاصل می‌شود و به طور کلی طول نیم‌قطرها برابر $a\sqrt{1 + \frac{k^2}{c^2}}$ و $b\sqrt{1 + \frac{k^2}{c^2}}$ است. برای تقاطع $x = k$ و $y = k$ به جای بیضی، به ترتیب هدلولی $\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 - \frac{k^2}{a^2}$ در صفحه $x = k$ و هدلولی $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{k^2}{b^2}$ را به دست می‌آوریم. در شکل (۲) رویه (۴) نمایش داده شده است.



شکل ۲

مثال ۳. (هذلولی وار دوپارچه) این نام به مجموعه (x, y, z) ها که در رابطه‌ای به صورت زیر صدق می‌کنند اطلاق می‌شود:

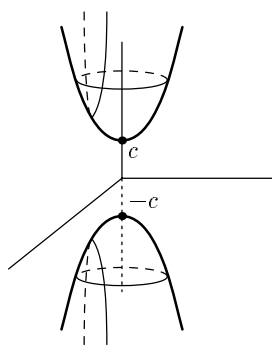
$$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad a, b, c > 0 \quad (5)$$

در اینجا تقاطع $x = k$ و $y = k$ همانند مثال قبل هذلولی هستند. برای $z = k$ داریم:

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{k^2}{c^2} - 1 \\ z = k \end{cases} \quad (6)$$

مشاهده می‌کنیم که به‌ازای $|k| > c$ یک بیضی در صفحه $z = k$ به‌دست می‌آید، برای $k = \pm c$ دو تک نقطه‌ای $(0, 0, \pm c)$ به ترتیب در صفحات $z = c$ و $z = -c$ حاصل می‌شوند و برای $|k| < c$

اشتراک تهی است. هذلولی وار دو پارچه در شکل ۳ نمایش داد شده است.

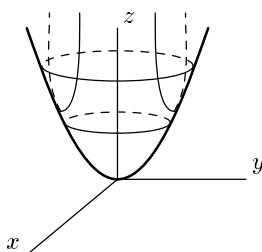


شکل ۳

مثال ۴. (سهمی وار بیضوی) اصطلاح سهمی وار بیضوی به مجموعه های (x, y, z) که در معادلاتی چون $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ یا $z = -(\frac{x^2}{a^2} + \frac{y^2}{b^2})$ صدق می کنند اطلاق می شود. مثلاً مجموعه (x, y, z) هایی را در نظر بگیرید که در رابطه زیر صدق می کنند:

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2} \quad a, b > 0 \quad (7)$$

مقطع این مجموعه با $z = k$ به ترتیب یک بیضی، تک نقطه ای $(0, 0, 0)$ و تهی است بسته به این که $k > 0$ یا $k = 0$ یا $k < 0$. هر مقطع $x = k$ و $y = k$ نیز یک سهمی است. بدین ترتیب شکل ۴ حاصل می شود.

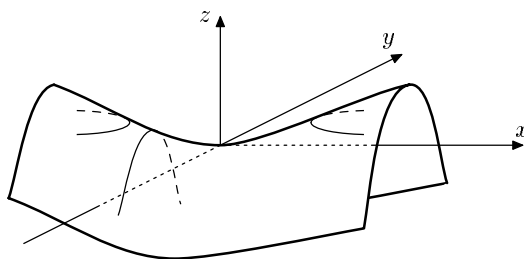


شکل ۴

مثال ۵. (سهمی وار هذلولوی) اگر علامت یکی از دو عبارت سمت راست (۷) را به منفی مبدل کنیم یک سهمی وار هذلولوی به دست می آید:

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2} \quad a, b > 0 \quad (8)$$

مقطع $z = 0$ یک جفت خط راست متقاطع $bx \pm ay = 0$ است و به ازای $z = k \neq 0$ یک هذلولوی به دست می آید که امتداد باز شدن آن برای $k > 0$ و $k < 0$ به اندازه $\frac{\pi}{4}$ چرخش دارد. تقاطع $x = k$ و $y = k$ سهمی هستند (شکل ۵)



شکل ۵

مثال ۶. (مخروط) مجموعه نقاط (x, y, z) که در رابطه زیر صدق می کنند در نظر بگیرید:

$$\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2} \quad a, b, c > 0 \quad (9)$$

تقاطع $z = k$ به ازای $k \neq 0$ بیضی و مقطع با $z = 0$ تک نقطه ای $(0, 0, 0)$ است. برای $x = k$ و $y = k$ مقطع یک هذلولی است. مقطع صفحه قائم $Ax + By = 0$ را با این شکل به دست می آوریم. دست کم یکی از A و B ناصفر است، مثلاً $A \neq 0$. با جایگزینی در (۹) داریم:

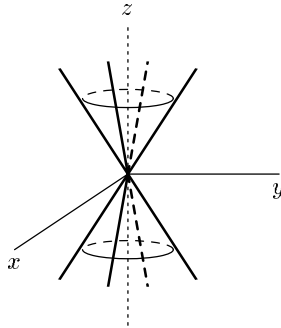
$$\frac{z^2}{c^2} = \frac{B^2}{A^2 a^2} y^2 + \frac{y^2}{b^2}$$

اگر $\frac{1}{b^2} + \frac{B^2}{A^2 a^2}$ را به $\frac{1}{p^2}$ نمایش دهیم، داریم:

$$\frac{z^2}{c^2} - \frac{y^2}{p^2} = 0$$

$$\left(\frac{z}{c} - \frac{y}{p}\right)\left(\frac{z}{c} + \frac{y}{p}\right) = 0$$

پس تصویر مقطع روی صفحه yz دو خط متقاطع است. با توجه به این که مقطع روی صفحه $Ax + By = 0$ قرار دارد، مقطع خود از دو خط راست تشکیل شده است. این مخروط در شکل ۶ نمایش داده شده است.



شکل ۶

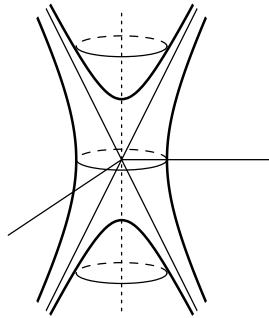
مخروط را می توان حالت گذر از هذلولی وار یکپارچه به هذلولی وار دوپارچه تلقی کرد. تابع

$f: \mathbb{R}^3 \rightarrow \mathbb{R}$ را به صورت زیر در نظر بگیرید:

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} \quad (10)$$

که در آن $a, b, c > 0$ داده شده اند. مجموعه های تراز f را بررسی می کنیم. به ازای $k > 0$ ، $f^{-1}(k)$ یک هذلولی وار یکپارچه است. با میل دادن $k > 0$ به سمت صفر، بیضی مرکزی این هذلولی وار تدریجاً کوچکتر می شود تا به ازای $k = 0$ این بیضی تبدیل به یک نقطه می شود و مخروط $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ پدید می آید. وقتی k از صفر گذر کرده و منفی می شود، یعنی $k < 0$ ، رویه به یک هذلولی وار دوپارچه مبدل می شود. می توان موضوع را اینگونه

تلقی کرد که با گذر از صفر به مقادیر منفی، دو شاخهٔ مخروط از هم جدا می‌شوند (شکل ۷)



شکل ۷

قابل ذکر است که همهٔ هذلولی‌وارهایی که مجموعهٔ تراز این تابع هستند نسبت به مخروط فوق
مجانبنند.

شش مثال بالا بیشتر حالت‌های غیر استثنایی مجموعه‌های تراز یک تابع درجهٔ دوم سه متغیری
را در بر می‌گیرد (در واقع مخروط را می‌توان یک حالت استثنایی تلقی کرد). یک مورد اساسی
دیگر موجود است و آن وقتی است که یکی از سه متغیر به درجه دو و دوتای دیگر با درجه اول ظاهر
می‌شوند. برای بررسی این مورد نخست وضعیت زیر را در نظر بگیرید.

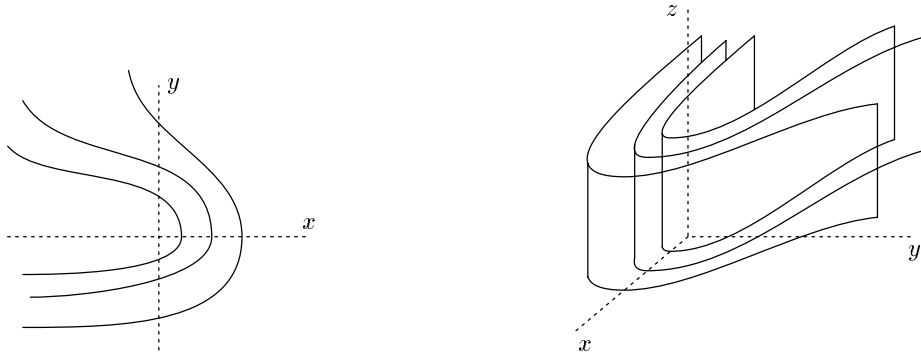
مثال ۷. (استوانه) فرض کنید $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ به صورت زیر تعریف شده است:

$$f(x, y, z) = g(x, y) \quad (11)$$

که در آن $g: \mathbb{R}^2 \rightarrow \mathbb{R}$. مجموعه‌های تراز f را بررسی می‌کنیم.

$$\begin{aligned} f^{-1}(k) &= \{(x, y, z) \mid f(x, y, z) = k\} \\ &= \{(x, y, z) \mid g(x, y) = k, z \in \mathbb{R}\} \end{aligned}$$

بدین ترتیب مجموعه تراز $f^{-1}(k)$ بدین طریق به دست می آید که مجموعه تراز $g^{-1}(k)$ را در صفحه xy در نظر می گیریم و آن را به موازات محور z در فضای سه بعدی امتداد می دهیم. شکل به دست آمده یک استوانه نام دارد. برای تابع های هر تعداد متغیر، هرگاه یک یا چند متغیر در عبارت تعریف کننده تابع ظاهر نشوند یک استوانه به مفهوم عام به دست می آید.



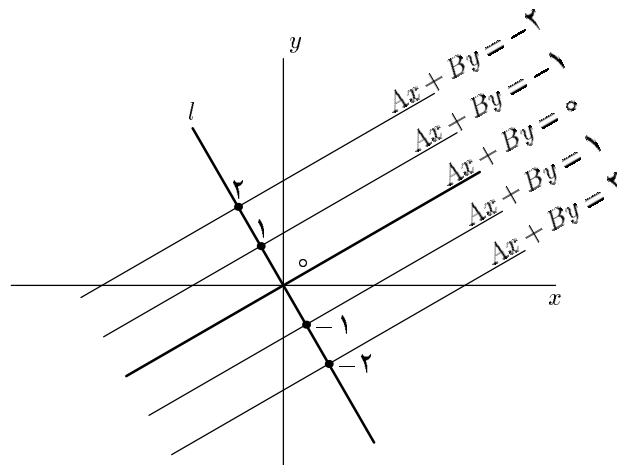
شکل ۸

مثال ۸. حال مجموعه (x, y, z) هایی را در نظر بگیرید که در رابطه

$$Ax + By + Cz^2 = 0 \quad (12)$$

صدق می کنند. در اینجا A, B و C اعداد حقیقی داده شده اند. قرار می دهیم $w = Ax + By$. خطوط راست $Ax + By = k$ ، $k \in \mathbb{R}$ داده شده، در صفحه xy موازی هستند. فرض کنید l خط گذرا از o در صفحه xy عمود بر این خطوط باشد. خط l را محور w می نامیم و هر نقطه آن را با مقدار ثابت w از

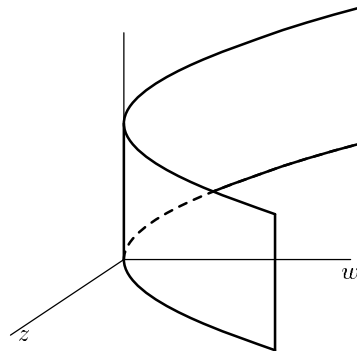
خط راست $Ax + By = w$ که در آن نقطه خط l را قطع می کند مدرج می کنیم (شکل ۹)



شکل ۹

حال در صفحه zw منحنی $w + cz^2 = 0$ را رسم می کنیم که یک سهمی است. اگر این سهمی را در جهت عمود بر صفحه zw امتداد دهیم استوانه ای به دست می آید که مکان هندسی (۱۲) است

(شکل ۱۰)



شکل ۱۰

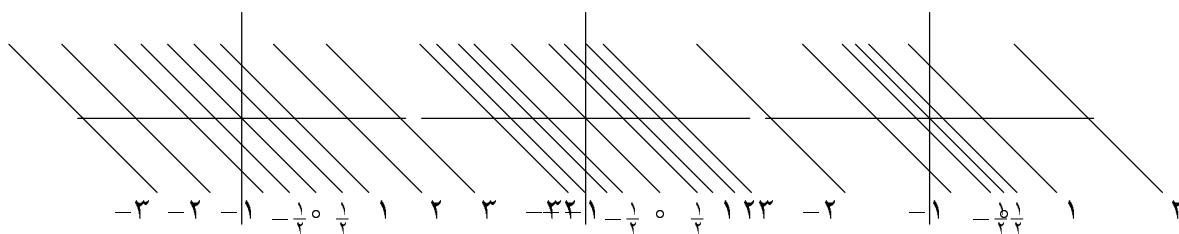
در پایان سؤال زیر را مطرح می کنیم. دیدیم که تابع های درجه ۱ (مستوی، خطی) دارای مجموعه های تراز مستوی، موازی و همباعد هستند، و تابع های درجه ۲ نیز از شکل های ویژه ای برخوردارند. تا چه حد می توان ماهیت یک تابع را از شکل مجموعه های تراز آن تعیین کرد؟ به عنوان

ساده‌ترین حالت، فرض کنید تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ طوری باشد که مجموعه‌های تراز آن خطوط راست موازی باشند. آیا f لزوماً خطی یا مستوی است؟ جواب این سؤال منفی است:

مثال ۹. سه تابع زیر از \mathbb{R}^2 به \mathbb{R} را در نظر بگیرید:

$$h(x, y) = \sqrt{x+y} \quad , \quad g(x, y) = (x+y)^3 \quad , \quad f(x, y) = x+y$$

طرف راست هر یک از این عبارتها را برابر مقداری ثابت قرار دهیم نتیجه می‌شود که $x+y$ ثابت است، و هر تابع هر مقدار ثابت را روی دقیقاً یک خط راست موازی $x+y = 0$ می‌گیرد. در شکل (۱۱) مجموعه‌های تراز این سه تابع به‌ازای مقدار مختلف نمایش داده شده‌اند. مجموعه‌های تراز به‌ازای $c = 0, \pm 1$ برای هر سه تابع یکی هستند. در مورد تابع خطی f به‌ازای مقادیر متوالی و



شکل ۱۱

هم فاصله c ، مثلاً $2, 1, 0, -1, -2$ ، مجموعه‌های تراز مربوط از هم هم فاصله‌اند. برای تابع g ، برای $|c| > 1$ ، مجموعه‌های تراز برای مقادیر هم فاصله c کندتر از مورد f از هم دور می‌شوند، و برعکس برای h ، برای $|c| > 1$ ، مجموعه‌های تراز برای مقادیر هم فاصله c فواصل فزاینده‌ای دارند. عکس این وضعیت برای مقادیر $|c| < 1$ رخ می‌دهد. به طور کلی نزدیکی و تمرکز مجموعه‌های تراز برای مقادیر هم فاصله دلیل بر رشد سریع تابع و بالعکس دوری نسبی آنها نشان رشد کند تابع است. مثلاً در این مثال، برای $|c| > 1$ ، تابع با مقدار $(x+y)^3$ سریعتر از $x+y$ و $x+y$ سریعتر از $(x+y)^{\frac{1}{2}}$ رشد می‌کند. این روند برای مقادیر $|c| < 1$ وقتی از 1 به 0 نزدیک می‌شویم معکوس می‌شود.

فرض کنید $\phi : \mathbb{R} \rightarrow \mathbb{R}$ یک تابع یک به یک باشد و $f : \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع دلخواه. در این صورت مجموعه‌های تراز $g = \phi \circ f : \mathbb{R}^n \rightarrow \mathbb{R}$ همان مجموعه‌های تراز f هستند منتها به‌ازای مقادیر متفاوت، که تحت ϕ جابجا می‌شوند. در مثال بالا، برای g داریم $\phi(t) = t^3$ و برای h $\phi(t) = t^{\frac{1}{3}}$. بالعکس اگر تابع $f : \mathbb{R}^n \rightarrow \mathbb{R}$ داده شده باشد و $g : \mathbb{R}^n \rightarrow \mathbb{R}$ دارای همان مجموعه‌های تراز f با مقادیر جابجا شده باشد، می‌توان نوشت $g = \phi \circ f$ که $\phi : S \rightarrow \mathbb{R}$ ، S مجموعه مقادیر ممکن f ، و ϕ جابجا کننده مقادیر f است.

پیوستگی و حد (۱)

مفهومی‌های پیوستگی و حد در وضعیت چندمتغیری شباهت کامل به وضعیت یک متغیری دارند. همان‌انگیزه‌ها و همان روشها در اینجا نیز حکمفرماست. فرض کنید S زیرمجموعه‌ای از \mathbb{R}^n است و $f : S \rightarrow \mathbb{R}$ یک تابع. نقاط S را به $x = (x_1, \dots, x_n)$ و مقدار $f(x)$ را به y نمایش می‌دهیم. در یک مسأله تجربی می‌توان x_1, \dots, x_n را عواملی دانست که با آزمایش اندازه‌گیری می‌شوند و $f(x_1, \dots, x_n)$ یک عبارت یا فرمول است که از آن مقدار y محاسبه می‌شود. نقطه‌ای $a = (a_1, \dots, a_n)$ در دامنه f در نظر بگیرید، یعنی $a \in S$. برای محاسبه مقدار $f(a)$ ، خطایی $e > 0$ را به عنوان "خطای قابل تحمل" منظور می‌کنیم. می‌خواهیم بدانیم که آیا می‌توان حدود خطایی $e'_1 > 0, \dots, e'_n > 0$ برای به ترتیب x_1, \dots, x_n پیدا کرد به طوری که اگر برای x در دامنه f ، $|x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n$ ، آنگاه خطای حاصل در محاسبه، یعنی $|f(x) - f(a)|$ کوچکتر از e باشد؟ اگر بتوان برای هر $e > 0$ ، مقادیر متنظاری $e'_1 > 0, \dots, e'_n > 0$ با ویژگی بالا پیدا کرد، می‌گوییم "محاسبه f در a پایدار است". اصطلاح معمول‌تر این است که تابع f را در نقطه a پیوسته بنامیم.

مثال ۱. (تابع‌های مستوی) تابع مستوی $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را در نظر می‌گیریم:

$$f(x_1, \dots, x_n) = c_0 + c_1 x_1 + \dots + c_n x_n$$

که در آن c_0, c_1, \dots, c_n اعداد حقیقی داده شده‌اند. نشان می‌دهیم f در همه نقاط \mathbb{R}^n پیوسته است. نقطه‌ای $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ در نظر بگیرید و خطای دلخواهی $e > 0$ منظور کنید. می‌خواهیم $e'_1 > 0, \dots, e'_n > 0$ را طوری تعیین کنیم که:

$$(|x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n) \implies |f(x) - f(a)| < e$$

معمولاً یافتن e'_1, \dots, e'_n به این طریق حاصل می‌شود که سعی می‌کنیم عبارت $|f(x) - f(a)|$ یا کمیتی بزرگتر از آن را، طوری بازنویسی کنیم که اثر هر یک از عوامل $|x_1 - a_1|, \dots, |x_n - a_n|$ در آن ظاهر شود. در اینجا:

$$\begin{aligned}
 |f(x) - f(a)| &= |(c_1 x_1 + \dots + c_n x_n) - (c_1 a_1 + \dots + c_n a_n)| \\
 &= |c_1(x_1 - a_1) + \dots + c_n(x_n - a_n)| \\
 &\leq |c_1||x_1 - a_1| + \dots + |c_n||x_n - a_n|
 \end{aligned}$$

حال اگر بتوانیم هر $|c_j||x_j - a_j|$ را کوچکتر از $\frac{e}{n}$ کنیم، نتیجه مطلوب در مورد خطای $f(x)$ به دست می‌آید. چون c_j داده شده و معلوم است اگر e'_j برابر (یا کوچکتر از) $\frac{e}{|c_j|n}$ اختیار شود مقصود حاصل می‌شود. تنها مشکل احتمالی این است که ممکن است c_j صفر باشد. برای احتراز از این مشکل $e_j > 0$ را کوچکتر یا مساوی $\frac{e}{(|c_j|+1)n}$ می‌گیریم. در این صورت:

$$\begin{aligned}
 (|x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n) \implies |f(x) - f(a)| &\leq |c_1||x_1 - a_1| + \dots + |c_n||x_n - a_n| \\
 &< |c_1| \frac{e}{(|c_1|+1)n} + \dots + |c_n| \frac{e}{(|c_n|+1)n} \\
 &< \frac{e}{n} + \dots + \frac{e}{n} = e
 \end{aligned}$$

و حکم به اثبات می‌رسد.

توجه کنید که تابع‌های مستوی شامل مثال‌های مهم زیر می‌شوند:

(الف) تابع‌های ثابت. در اینجا $c_1 = \dots = c_n = 0$.

(ب) تابع‌های افکنش. مقصود از "تابع افکنش روی مؤلفه j ام" تابع $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}$ که به صورت

$\pi_j(x_1, \dots, x_n) = x_j$ تعریف می‌شود. در اینجا $c_j = 1$ و سایر c_j ها صفر هستند.

ج) تابع جمع. تابع $s: \mathbb{R}^n \rightarrow \mathbb{R}$ که به صورت زیر تعریف می‌شود تابع جمع می‌نامیم:

$$s(x_1, \dots, x_n) = x_1 + \dots + x_n$$

در اینجا $c_0 = 0$ و $c_1 = \dots = c_n = 1$.

مثال ۲. (تابع ضرب) $p: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$p(x, y) = x \cdot y$$

ادعا می‌کنیم p در هر نقطه (a, b) از \mathbb{R}^2 پیوسته است. برای خطای داده شده $\epsilon > 0$ ، باید $\epsilon' > 0$ و $\epsilon' > 0$ را طوری تعیین کنیم که:

$$(|x - a| < \epsilon', |y - b| < \epsilon') \implies |xy - ab| < \epsilon$$

در اینجا نیز $|xy - ab|$ را بازنویسی می‌کنیم:

$$\begin{aligned} |xy - ab| &= |xy - xb + xb - ab| \\ &\leq |x||y - b| + |b||x - a| \end{aligned}$$

مجدداً می‌خواهیم با کوچکتر کردن هر یک از دو عامل جمع در سمت راست از $\frac{\epsilon}{2}$ ، به نتیجه مطلوب برسیم. تنها نکته جدید در اینجا این است که ضریب $|x|$ در عبارت اول خود متغیر است. برای رفع این مشکل، چون قرار است x به a نزدیک باشد، مقدمتاً x را به بازه $[a - 1, a + 1]$ محدود می‌کنیم، پس با این قید:

$$|x| < |a| + 1$$

و در نتیجه

$$|xy - ab| \leq (|a| + 1)|y - b| + |b||x - a|$$

حال با گرفتن $0 < e' \leq \min\{1, \frac{e}{\sqrt{(|b|+1)}}\}$ و $0 < e' \leq \frac{e}{\sqrt{(|a|+1)}}$ ، هم قید مقدماتی تأمین می شود و هم نامساوی مورد نظر در مورد $|xy - ab|$ به دست می آید.

مثال ۳. (تابع خارج قسمت) U را مجموعه زیر از \mathbb{R}^2 می گیریم:

$$U = \{(x, y) \mid y \neq 0\}$$

یعنی U مکمل محور x در صفحه xy است. تابع $q: U \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم:

$$q(x, y) = \frac{x}{y}$$

ثابت می کنیم q در هر نقطه دامنه خود، (a, b) ، $b \neq 0$ ، پیوسته است. برای $e > 0$ داده شده، $e' > 0$ و $e' > 0$ جستجو می کنیم که

$$((x, y) \in U, |x - a| < e', |y - b| < e') \implies \left| \frac{x}{y} - \frac{a}{b} \right| < e$$

داریم

$$\begin{aligned} \left| \frac{x}{y} - \frac{a}{b} \right| &= \frac{|bx - ay|}{|y||b|} \\ &= \frac{|bx - ba + ba - ay|}{|y||b|} \\ &\leq \frac{|x - a|}{|y|} + \frac{|a||y - b|}{|y||b|} \end{aligned}$$

می خواهیم با انتخاب مناسب $e' > 0$ و $e' > 0$ هر یک از دو عامل طرف راست را از $\frac{e}{2}$ کوچکتر کنیم. با این که $y \neq 0$ ، از آنجا که فعلاً محدودیتی برای $y \neq 0$ قائل نشده ایم، مخرج هر یک از دو عبارت می تواند به دلخواه کوچک شود و از این رو در نگاه اول محدود کردن اندازه کسرها مشکل به نظر می رسد. ولی توجه کنید که $b \neq 0$ مقداری داده شده است و قرار است y نهایتاً نزدیک b اختیار شود. مقدمتاً y را از نصف فاصله b به 0 بزرگتر می گیریم، یعنی $|y - b| < \frac{b}{2}$ ، یا $|y| > \frac{|b|}{2}$. این موجب می شود که $\frac{1}{|y|} < \frac{2}{|b|}$ ، پس با این قید داریم:

$$\left| \frac{x}{y} - \frac{a}{b} \right| \leq \frac{2}{|b|} |x - a| + \frac{2|a|}{|b|^2} |y - b|$$

حال ضرایب $|x - a|$ و $|y - b|$ مقادیر ثابتی هستند. $e'_1 > 0$ و $e'_2 > 0$ را به طریق زیر اختیار می‌کنیم:

$$0 < e'_1 \leq \frac{|b|}{4} e, \quad 0 < e'_2 \leq \min\left\{\frac{|b|^2}{4(|a| + 1)} e, \frac{|b|}{2}\right\}$$

به این ترتیب اگر $|x - a| < e'_1$ و $|y - b| < e'_2$ ، هم قید مقدماتی رعایت می‌شود و هم $|\frac{x}{y} - \frac{a}{b}| < e$ حاصل می‌گردد.

تا اینجا فقط تابع‌هایی را در نظر گرفته‌ایم که برد آنها \mathbb{R} است. حال تابعی $f: S \rightarrow \mathbb{R}^m$ در نظر بگیریم که در آن $S \subset \mathbb{R}^n$. داریم $f(x) = (f_1(x), \dots, f_m(x))$ که در اینجا هر یک از f_i ها تابعی از S به \mathbb{R} است. از آنجا که محاسبه تابع f به منزله محاسبه m تابع f_1, \dots, f_m است، طبیعی است که پیوستگی f در نقطه a را به صورت زیر تعریف کنیم: f را در نقطه $a \in S$ پیوسته می‌نامیم در صورتی که هر یک از f_i ها در نقطه a پیوسته باشد. بدین ترتیب پیوستگی یک تابع با مقدار در \mathbb{R}^m نکته تازه‌ای ندارد، بلکه باید پیوستگی m تابع مؤلفه را بررسی کرد. با این حال می‌توان صورت معادلی از این تعریف ارائه کرد که شباهت ظاهری بیشتری به تعریف پیوستگی تابع‌های $\mathbb{R} \rightarrow S$ دارد و تعبیر هندسی بعضاً سودمندی می‌پذیرد. برای این کار نخست نامساوی‌های ساده و اساسی زیر را مطرح می‌کنیم.

(۱-۱۸) فرض کنید $z = (z_1, \dots, z_p)$ و $w = (w_1, \dots, w_p)$ که در آن z_i ها و w_i ها اعداد حقیقی هستند. در این صورت

$$|z_i - w_i| \leq |z - w| \quad \text{برای هر } i = 1, \dots, p \quad (1)$$

$$|z - w| \leq \sqrt{p} \max\{|z_1 - w_1|, \dots, |z_p - w_p|\} \quad (2)$$

توجه کنید که $|z - w|$ طول (نرم) p تایی $z - w$ است و هر $|z_i - w_i|$ قدرمطلق عدد $z_i - w_i$ می‌باشد. نامساوی (۱) از تعریف $|z - w|^2 = \sum_{i=1}^p |z_i - w_i|^2$ نتیجه می‌شود. برای (۲) نیز از همین تعریف

داریم:

$$|z - w|^2 \leq p \max\{|z_1 - w_1|^2, \dots, |z_p - w_p|^2\}$$

که با جذرگیری نتیجه مورد نظر را می دهد.

نامساوی (۱) نتیجه می دهد که اگر دو p تایی z و w در \mathbb{R}^p نزدیکتر از e باشند، تفاضل مؤلفه های متناظر آنها نیز کوچکتر از e است. بالعکس از (۲) می بینیم که هرگاه قدرمطلق تفاضل همه مؤلفه های متناظر z و w کوچکتر از e باشند، آنگاه فاصله z از w کوچکتر از $\sqrt{p}e$ است. با توجه به این که در \mathbb{R}^p یک عدد ثابت است، این دو نامساوی نشان می دهند که برای نزدیک کردن z و w لازم و کافی است که همه مؤلفه های متناظر z و w به هم نزدیک شوند. با این مقدمه، تعریف پیوستگی تابع های $f: S \rightarrow \mathbb{R}^m$ را به صورت زیر بازنویسی می کنیم:

(۱۸-۲) گزاره. S زیرمجموعه ای از \mathbb{R}^n است، $a \in S$ و $f: S \rightarrow \mathbb{R}^m$ یک تابع. در این صورت شرطی لازم و کافی برای پیوستگی f در a این است که برای هر $e > 0$ وجود داشته باشد $e' > 0$ که

$$(x \in S, |x - a| < e') \implies |f(x) - f(a)| < e$$

اثبات. نخست فرض می کنیم هر f_i در نقطه a پیوسته است و $e > 0$ داده شده است. بنابراین فرض پیوستگی، $e'_{i_1} > 0, \dots, e'_{i_m} > 0$ وجود دارند که:

$$(x \in S, |x_1 - a_1| < e'_{i_1}, \dots, |x_m - a_m| < e'_{i_m}) \implies |f_i(x) - f_i(a)| < \frac{e}{\sqrt{m}}$$

بنابراین تحت این شرایط از (۲) نتیجه می گیریم که:

$$|f(x) - f(a)| < e$$

از سویی دیگر اگر قرار دهیم $e' = \min\{e'_{ij} \mid i = 1, \dots, m; j = 1, \dots, n\}$ ، هرگاه $|x - a| < e'$ ، می‌توان طبق (۱) نتیجه گرفت که $|x_j - a_j| < e'$ برای هر j ، و در نتیجه $|x_j - a_j| < e'_{ij}$ ، بنابراین:

$$(x \in S, |x - a| < e') \implies |f(x) - f(a)| < e$$

بالعکس فرض کنید شرط بالا را بتوان به‌ازای هر $e > 0$ تأمین کرد، نشان می‌دهیم هر f_i در نقطه a پیوسته است. فرض کنید $e > 0$ داده شده باشد، می‌خواهیم $e'_1 > 0, \dots, e'_n > 0$ را طوری اختیار کنیم که:

$$(x \in S, |x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n) \implies |f_i(x) - f_i(a)| < e$$

طبق فرض، $e' > 0$ وجود دارد که $|x - a| < e'$ و $x \in S$ نتیجه می‌دهند $|f(x) - f(a)| < e$ ، و در نتیجه طبق (۱)، $|f_i(x) - f_i(a)| < e$ برای هر i . حال تعریف می‌کنیم:

$$e'_1 = \dots = e'_n = \frac{e'}{\sqrt{n}}$$

بنابراین طبق (۲):

$$\begin{aligned} (x \in S, |x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n) &\implies (x \in S, |x - a| < e') \\ &\implies |f(x) - f(a)| < e \\ &\implies |f_i(x) - f_i(a)| < e \end{aligned}$$

□ و حکم به اثبات می‌رسد.

حکم گزاره ۱۸-۲ را می‌توان به صورت هندسی زیرتجسم کرد. مجموعه y های \mathbb{R}^m که در $|y - f(a)| < e$ صدق می‌کنند دقیقاً نقاط گوی باز شعاع e حول $f(a)$ هستند. همین طور مجموعه نقاط $x \in S$ که در $|x - a| < e'$ صدق می‌کنند، دقیقاً آن نقاط گوی باز شعاع e' حول a هستند که در مجموعه S قرار دارند. بنابراین پیوستگی در نقطه a بدین معنی است که به‌ازای هر گوی باز B به شعاع مثبت حول $f(a)$ ، گوی بازی B' به شعاع مثبت حول a وجود دارد که $f(B' \cap S) \subset B$ (شکل ۱). اگر

به تعریف اولیه مؤلفه‌ای پیوستگی در نقطه a بازگردیم، برای هر f_i و هر e_i باید $e'_{i1} > 0, \dots, e'_{in} > 0$ وجود داشته باشند که $|x_j - a_j| < e'_{ij}$ ، $j = 1, \dots, n$ و $x \in S$ نتیجه دهد $|f_i(x) - f_i(a)| < e_i$. قرار می‌دهیم $e'_j = \min\{e'_{1j}, \dots, e'_{nj}\}$ پس

$$(x \in S, |x_1 - a_1| < e'_1, \dots, |x_n - a_n| < e'_n) \implies |f_i(x) - f_i(a)| < e_i, i = 1, \dots, m)$$

تجسم هندسی این مطلب این است که برای هر مستطیل باز R به مرکز $f(a)$ ، مستطیل بازی R' به مرکز a وجود دارد که $f(R' \cap S) \subset R$ (شکل ۲). استدلال بالا و استدلال گزاره ۱۸-۲ را می‌توان در این نکته خلاصه کرد که درون هر گوی باز به مرکز یک نقطه p ، مستطیل بازی به همان مرکز وجود دارد، و بالعکس درون هر مستطیل باز به مرکز p ، یک گوی باز به مرکز p وجود دارد.

پیوستگی و حد (۲)

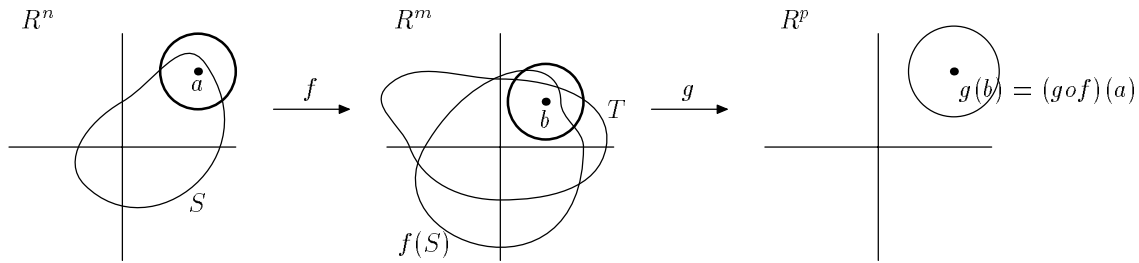
تابع $f : S \rightarrow \mathbb{R}^m$ را پیوسته می‌نامیم در صورتی که f در همه نقاط دامنه خود، یعنی S ، پیوسته باشد. مثال عمده‌ای که از تابع‌های پیوسته در جلسه گذشته دیدیم تابع‌های مستوی $\mathbb{R}^n \rightarrow \mathbb{R}^m$ بودند. در واقع ثابت کردیم هر تابع مستوی $\mathbb{R}^n \rightarrow \mathbb{R}$ پیوسته است ولی از آنجا که شرط لازم و کافی برای پیوستگی یک تابع به \mathbb{R}^m پیوستگی همه مؤلفه‌های f است و هر مؤلفه تابع مستوی، مستوی است، حکم در مورد تابع‌های مستوی $\mathbb{R}^n \rightarrow \mathbb{R}^m$ نتیجه می‌شود. با مثال‌هایی که در مورد تابع‌های ضرب و خارج قسمت ثابت کردیم، اکنون می‌توان گردایه بزرگی از تابع‌های پیوسته فراهم آورد. به این منظور نخست حکم کلی و مهم زیر را ثابت می‌کنیم:

(۱-۱۹) گزاره. فرض کنید $f : S \rightarrow \mathbb{R}^m$ ، $S \subset \mathbb{R}^n$ و $g : T \rightarrow \mathbb{R}^p$ ، $T \subset \mathbb{R}^m$ دو تابع باشند، $a \in S$ ، $b = f(a) \in T$ در f پیوسته باشد و g در b پیوسته باشد. در این صورت تابع $g \circ f$ در نقطه a پیوسته است.

اثبات. نخست دامنه $g \circ f$ را در نظر بگیرید:

$$(g \circ f) \text{ دامنه} = \{x \in S \mid f(x) \in T\}$$

که زیرمجموعه‌ای از S است. نقطه a در این دامنه است زیرا که فرض کرده‌ایم $b = f(a)$ عضو T است. اثبات گزاره با توجه به گزاره $18-2$ بخش قبل سراسر است. هرگاه گوی باز شعاع $e > 0$



حول $(g \circ f)(a)$ داده شده باشد، باید $e' > 0$ پیدا کنیم که هرگاه $x \in S$ و x در گوی باز شعاع e' حول a باشد، آنگاه $(g \circ f)(x)$ در گوی باز شعاع e حول $(g \circ f)(a)$ قرار گیرد. نخست از پیوستگی g در b استفاده می‌کنیم. طبق این فرض گوی شعاع $e'' > 0$ حول $b = f(a)$ وجود دارد که:

$$(y \in T, |y - b| < e'') \implies |g(y) - g(b)| < e \quad (1)$$

حال برای گوی شعاع $e'' > 0$ حول b ، بنابر پیوستگی f در a ، گوی $e' > 0$ حول a وجود دارد که:

$$(x \in S, |x - a| < e') \implies |f(x) - b| < e'' \quad (2)$$

بنابراین اگر x به علاوه در دامنه $g \circ f$ باشد، اگر $f(x)$ در (2) را y بنامیم، حکم از (1) نتیجه می‌شود. □
 با این گزاره و مثال‌های بخش قبل اکنون می‌توانیم پیوستگی "تابع‌های گویا" را بررسی کنیم.
 نخست مقصود از یک تک جمله‌ای $\mathbb{R}^n \rightarrow \mathbb{R}$ تابعی به شکل زیر است:

$$f(x_1, \dots, x_n) = cx_1^{\alpha_1} \dots x_n^{\alpha_n} \quad (3)$$

که در اینجا $\alpha_1, \dots, \alpha_n$ اعداد صحیح نامنفی هستند و $c \in \mathbb{R}$. به عنوان یک حالت ساده تابع $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ را در نظر بگیرید که به صورت $f(x, y, z) = xyz$ تعریف شده است. نشان می‌دهیم

پیوسته است برای این کار؛ f را به صورت ترکیب $f = p \circ g$ می نویسیم که در آن:

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad g(x, y, z) = (xy, z)$$

و

$$p : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad p(u, v) = uv$$

برای اثبات پیوستگی f ، کافی است بنا بر گزاره ۱۹-۱ ملاحظه کنیم که g و p پیوسته اند. p تابع حاصل ضرب است که در بخش گذشته پیوستگی آن ثابت شد. برای تابع g ، باید تحقیق کنیم که هر مؤلفه پیوسته است؛ مؤلفه دوم h تابع افکنش $z \rightarrow (x, y, z)$ است که پیوسته می باشد. مؤلفه اول خود ترکیب افکنش و حاصل ضرب است:

$$(x, y, z) \rightarrow (x, y) \rightarrow xy$$

افکنش روی هر تعداد مؤلفه پیوسته است زیرا هر مؤلفه افکنش روی یک محور است. حال برای تک جمله ای کلی (۳) نیز با استفاده مکرر از این استدلال، یا استقراء، دیده می شود که (۳) یک تابع پیوسته از \mathbb{R}^n به \mathbb{R} تعریف می کند.

مجموع تعداد متناهی تک جمله ای، یک چندجمله ای می شود:

$$f(x_1, \dots, x_n) = c_1 x_1^{\alpha_1} \dots x_n^{\alpha_n} + \dots + c_k x_1^{\beta_1} \dots x_n^{\beta_n} \quad (4)$$

f حاصل ترکیب تک جمله ای ها با تکرار عمل جمع است که هر دو نوع تابع پیوسته اند، پس f پیوسته است. به عنوان مثال برای دو جمله ای $f(x_1, \dots, x_n) = c_1 x_1^{\alpha_1} \dots x_n^{\alpha_n} + c_2 x_1^{\beta_1} \dots x_n^{\beta_n}$ ، f را به صورت ترکیب زیر می نویسیم:

$$(x_1, \dots, x_n) \xrightarrow{g} (c_1 x_1^{\alpha_1} \dots x_n^{\alpha_n}, c_2 x_1^{\beta_1} \dots x_n^{\beta_n}) \xrightarrow{s} c_1 x_1^{\alpha_1} \dots x_n^{\alpha_n} + c_2 x_1^{\beta_1} \dots x_n^{\beta_n}$$

تابع جمع، s ، پیوسته است (مثال بخش قبل) و هر یک از دو مؤلفه تابع g تک جمله‌ای، پس پیوسته‌اند، بنابراین $f = s \circ g$ پیوسته می‌باشد. حالت کلی با استقراء یا استفاده مکرر از تابع جمع حاصل می‌شود. بدین ترتیب چند جمله‌ای‌ها تابع‌های پیوسته تعریف می‌کنند.

یک گام فرای چند جمله‌ای‌ها، "تابع‌های گویا" هستند. اگر $g(x)$ و $h(x)$ دو چند جمله‌ای بر حسب $x = (x_1, \dots, x_n)$ باشند، تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌شود یک تابع گویا می‌نامیم:

$$f(x_1, \dots, x_n) = \frac{g(x_1, \dots, x_n)}{h(x_1, \dots, x_n)} \quad (5)$$

دامنه f از (x_1, \dots, x_n) ‌هایی در \mathbb{R}^n تشکیل شده که $h(x_1, \dots, x_n) \neq 0$. تابع f را می‌توان به صورت ترکیب دو تابع p و q به صورت زیر نوشت:

$$f = q \circ p, \quad x \xrightarrow{p} (g(x), h(x)) \xrightarrow{q} \frac{g(x)}{h(x)} \quad (6)$$

تابع p پیوسته است زیرا که هر مؤلفه آن یک چند جمله‌ای است، و تابع q نیز، که تابع خارج قسمت است، در بخش قبل نشان داده شده که در نقاطی که مؤلفه دوم صفر نباشد پیوسته است. بنابراین f پیوسته می‌باشد.

اکنون به بررسی مفهوم "حد" می‌پردازیم که قرابت‌هایی با مفهوم پیوستگی دارد. اگر $\rho > 0$ داده شده باشد، مقصود از گوی باز محذوف به شعاع ρ و مرکز $a \in \mathbb{R}^n$ ، مجموعه زیر است:

$$\{x \in \mathbb{R}^n \mid 0 < |x - a| < \rho\}$$

به بیان دیگر، با حذف مرکز از گوی باز شعاع ρ ، گوی محذوف شعاع ρ به مرکز a به دست می‌آید. حال زیرمجموعه S از \mathbb{R}^n را در نظر بگیرید. نقطه $a \in \mathbb{R}^n$ را یک نقطه حدی S می‌نامیم در صورتی هر گوی محذوف مثبت به مرکز a حاوی نقطه‌ای از S باشد. توجه کنید که این تعریف شرطی بر

تعلق نقطه a به مجموعه S نمی‌گذارد؛ a ممکن است نقطه‌ای از S باشد یا نباشد. در واقع می‌توان از تعریف نتیجه گرفت که در هر گوی محذوف به مرکز نقطه حدی، بی‌نهایت نقطه از مجموعه S وجود دارد. زیرا که هرگاه نقطه‌ای x از S در یک گوی محذوف به مرکز a باشد، داریم $|x - a| > 0$ و حال طبق تعریف باید نقطه‌ای y از S در گوی محذوف شعاع $|x - a|$ حول a باشد، یعنی $|y - a| < |x - a|$ پس $y \neq x$ و به همین ترتیب با کوچک کردن شعاع کره محذوف همواره نقطه جدیدی از S در گوی محذوف اولیه حول a پیدا می‌شود.

مثال ۱. برای مستطیل نیمه باز $a < x < b$ ، $c \leq y \leq d$ در \mathbb{R}^2 ، که در آن $a < b$ و $c < d$ ، مجموعه نقاط حدی این مجموعه عبارت است از مستطیل بسته $a \leq x \leq b$ ، $c \leq y \leq d$.

مثال ۲. برای حلقه $0 < |x| < 1$ در \mathbb{R}^n ، یعنی مجموعه $\{x \in \mathbb{R}^n \mid 0 < |x| < 1\}$ ، مجموعه نقاط حدی برابر است با گوی بسته $\{x \in \mathbb{R}^n \mid |x| \leq 1\}$.

اکنون می‌توانیم مفهوم "حد" یک تابع را مطرح کنیم. فرض کنید S زیرمجموعه‌ای از \mathbb{R}^n باشد، a یک نقطه حدی S ، $f : S \rightarrow \mathbb{R}^m$ یک تابع، و $L \in \mathbb{R}^m$ می‌نویسیم:

$$\lim_{x \rightarrow a} f(x) = L \quad (7)$$

و می‌خوانیم حد تابع f وقتی x به a میل کند برابر L است در صورتی که برای هر $\epsilon > 0$ ، وجود داشته باشد $\delta > 0$ که:

$$(x \in S, 0 < |x - a| < \delta) \implies |f(x) - L| < \epsilon \quad (8)$$

به بیان دیگر، برای هر گوی با شعاع مثبت ϵ حول L ، گوی محذوفی به شعاع مثبت δ حول a وجود داشته باشد که اشتراک آن با S به تمامی به داخل گوی شعاع ϵ حول L نگاهشته شود.

این تعریف با تعریف پیوستگی f در نقطه a در نکات زیر متمایز است:

(الف) نقطه a ممکن است یک نقطه دامنه تعریف f نباشد و بالعکس یک نقطه پیوستگی f ممکن است نقطه حدی نباشد، یعنی حد تابع وقتی x به a میل کند معنی نداشته باشد.

(ب) در تعریف حد، حتی اگر f در a تعریف شده باشد، شرطی بر مقدار $f(a)$ قایل نشده ایم.

حکم زیر بلافاصله از تعریف نتیجه می شود:

(۱۹-۲) گزاره. برای تابع $f: S \rightarrow \mathbb{R}^m$ ، اگر a در دامنه تعریف f ، یعنی S ، باشد، و نقطه حدی S نیز باشد، آنگاه شرطی لازم و کافی برای پیوستگی f در a این است که حد f وقتی x به a میل می کند وجود داشته و برابر $f(a)$ باشد. \square

توجه کنید که حد، در صورت وجود یگانه است. فرض کنید داشته باشیم:

$$\lim_{x \rightarrow a} f(x) = L' \quad , \quad \lim_{x \rightarrow a} f(x) = L$$

اگر $L \neq L'$ ، $|L - L'|$ اکیداً مثبت است. عدد $\epsilon > 0$ را طوری می گیریم که $\frac{1}{4}\epsilon < |L - L'|$. در این صورت $\epsilon' > 0$ و $\epsilon'' > 0$ وجود دارند که:

$$(x \in S, 0 < |x - a| < \epsilon') \implies |f(x) - L| < \epsilon$$

$$(x \in S, 0 < |x - a| < \epsilon'') \implies |f(x) - L'| < \epsilon$$

بدین ترتیب اگر $\bar{\epsilon} = \min\{\epsilon', \epsilon''\}$ ، همه نقاط S که در گوی محذوف شعاع $\bar{\epsilon}$ حول a قرار دارند هم به داخل گوی شعاع ϵ حول L و هم به داخل گوی شعاع ϵ حول L' نگاشته می شوند. این غیرممکن است زیرا L و L' از یکدیگر فاصله بیش از دو برابر ϵ دارند.

مثال. S را زیرمجموعه $\{(x, y) \mid (x, y) \neq (0, 0)\}$ از \mathbb{R}^2 بگیرید. نقطه $(0, 0)$ عضو این مجموعه نیست ولی نقطه حدی آن است. تابع $f: S \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم:

$$f(x, y) = \frac{x^m y^n}{x^2 + y^2}$$

که در آن $m \geq 0$ و $n \geq 0$ اعداد صحیح داده شده اند. می خواهیم $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ را بررسی کنیم. در مختصات قطبی عبارت تابع به صورت زیر در می آید:

$$f(x, y) = r^{m+n-2} \cos^m \theta \sin^n \theta \quad , \quad (0, 0) \text{ از } (x, y) \text{ فاصله} = r \quad (9)$$

حالت اول. $(m + n > 2)$ بنویسید $p = m + n - 2$. $0 < p$ داریم.

$$|f(x, y)| \leq r^p$$

زیرا که $|\cos \theta|$ و $|\sin \theta|$ هر دو کوچکتر یا مساوی واحد هستند. حال اگر $e > 0$ داده شده باشد، قرار می دهیم $0 < e' \leq \sqrt[p]{e}$. اگر (x, y) در گوی باز محذوف شعاع e' حول $(0, 0)$ باشد، داریم:

$$|(x, y) - (0, 0)| < \sqrt[p]{e} \implies r^p < e$$

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0 \quad \text{بنابراین}$$

حالت دوم. $(m + n \leq 2)$. در اینجا نیز می نویسیم $p = m + n - 2$ ولی $p \leq 0$ داریم:

$$f(x, y) = r^p \cos^m \theta \sin^n \theta$$

اگر $p = 0$ ، داریم $f(x, y) = \cos^m \theta \sin^n \theta$ و f روی هر نیم خط منتهی به $(0, 0)$ مقداری ثابت دارد. بنابراین هرگاه (x, y) از روی یک نیم خط به $(0, 0)$ میل کند مقدار (ثابت) تابع به $\cos^m \theta \sin^n \theta$

میل خواهد کرد. با توجه به این که مقدار به θ بستگی دارد و حد، در صورت وجود، یگانه است،

$\lim_{(x,y) \rightarrow (0,0)} f(x,y)$ وجود ندارد. برای p منفی، قرار می دهیم $p = -q$ که $q > 0$ و داریم:

$$|f(x,y)| = \frac{1}{r^q} |\cos^m \theta| |\sin^n \theta|$$

مثلاً برای $\theta = \frac{\pi}{4}$ ، $|f(x,y)| = \frac{2^{-\frac{m+n}{4}}}{r^q}$. وقتی فاصله (x,y) به $(0,0)$ به صفر میل کند، r^q به صفر

میل می کند و $|f(x,y)|$ به طور نامحدود بزرگ می شود، یعنی به هیچ عدد L میل نمی کند.

شباهت های تعریف حد و پیوستگی نتایج مشابهی را ایجاد می کند. احکام زیر همه اثبات هایی

کاملاً مشابه احکام نظیر در مورد پیوستگی دارند و اثبات آنها به خواننده واگذار می شود.

(۳-۱۹) $S \subset \mathbb{R}^n$ ، a یک نقطه حدی برای S ، $f : S \rightarrow \mathbb{R}^m$ ، $f = (f_1, \dots, f_m)$ و

$L = (L_1, \dots, L_m)$ عضوی از \mathbb{R}^m است. در این صورت $\lim_{x \rightarrow a} f(x) = L$ اگر و تنها اگر

$$\lim_{x \rightarrow a} f_i(x) = L_i \quad \text{برای } i = 1, \dots, m$$

(۴-۱۹) $S \subset \mathbb{R}^n$ ، a یک نقطه حدی برای S ، $f : S \rightarrow \mathbb{R}$ ، $g : S \rightarrow \mathbb{R}$ ، $\lim_{x \rightarrow a} f(x) = L$ و

$\lim_{x \rightarrow a} g(x) = L'$. در این صورت تابع های $f+g : S \rightarrow \mathbb{R}$ که به صورت $(f+g)(x) = f(x)+g(x)$

و $f \cdot g : S \rightarrow \mathbb{R}$ که به صورت $(f \cdot g)(x) = f(x)g(x)$ تعریف می شوند دارای حد هستند وقتی x به

a میل کند و در واقع:

$$\lim_{x \rightarrow a} ((f+g)(x)) = L + L' \quad (10)$$

$$\lim_{x \rightarrow a} ((f \cdot g)(x)) = LL' \quad (11)$$

□

در مورد خارج قسمت در وضعیت بالا، تابع خارج قسمت را به صورت

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

تعریف می‌کنیم. دامنه $\frac{f}{g}$ از آن نقاط $x \in S$ تشکیل شده است که در آن $g(x) \neq 0$. فرض کنید $L' \neq 0$. اگر e را برابر مثلاً $\frac{1}{4}|L'|$ در نظر بگیریم، $e' > 0$ وجود دارد که:

$$(s \in S, |x - a| < e') \implies |g(x) - L'| < e = \frac{1}{4}|L'|$$

از این مطلب نتیجه می‌شود که برای $x \in S$ که در گوی محذوف شعاع e' حول a باشند، لزوماً $g(x) \neq 0$. بنابراین دامنه $\frac{f}{g}$ شامل اشتراک S با یک گوی محذوف حول a می‌شود و a یک نقطهٔ حدی برای دامنه $\frac{f}{g}$ است. با این مقدمه، می‌توان مانند حالت پیوستگی ثابت کرد که:

$$\lim_{x \rightarrow a} \left(\frac{f}{g}\right)(x) = \frac{L}{L'} \quad (12)$$

مشتق و تقریب خطی (۱)

در بررسی توابع حقیقی یک متغیری دیدیم که دو انگیزه متفاوت، یکی صورتبندی مفهوم آهنگ تغییرات، و دیگری یافتن مماس و تقریب خطی برای تابعهای غیرخطی، منجر به پیدایش مفهوم مشتق شدند. در اینجا قصد داریم همین بحثها را برای تابعهای چندمتغیری مطرح کنیم. خواهیم دید که دو رهیافت ذکر شده در اینجا منجر به تعاریف متمایز ولی مرتبط می شوند.

در آنچه خواهد آمد، S زیرمجموعه‌ای از \mathbb{R}^n است، $f: S \rightarrow \mathbb{R}^m$ یک تابع و a یک نقطه درونی S است. اگر بخواهیم برای آهنگ تغییر f در نقطه a عبارتی مشابه عبارت مربوط به مشتق تابع یک متغیری در نظر بگیریم، باید مفهوم $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$ را تعمیم دهیم. با توجه به این که $a \in S$ یک n تایی است، "تغییر کوچک" در a ، یعنی h ، نیز باید عنصری از \mathbb{R}^n در نظر گرفته شود (a ممکن است در هر جهتی از نقطه a در \mathbb{R}^n تغییر کند). اشکال واضحی که در اینجا پدید می آید این است که مخرج کسر بالا به یک n -تایی تبدیل می شود ولی مفهوم "تقسیم بر یک بردار" تعریف نشده است. اگر چنین عمل تقسیمی تعریف شدنی بود، خارج قسمت تقسیم یک m -تایی بر یک n -تایی چگونه موجودی می بایست باشد؟ در بخش بعد خواهیم دید که جواب این سؤال به اعتباری یک ماتریس $m \times n$ است. توجه کنید که اگر m -تایی‌ها و n -تایی‌ها را به صورت ستون بنویسیم، حاصل ضرب یک ماتریس $m \times n$ در یک ستون n تایی یک ستون m تایی است، بنابراین دور از ذهن نیست که "نسبت" یک m تایی به یک n تایی از "جنس" ماتریس $m \times n$ باشد. ولی در ابتدا هدف محدودتری را مد نظر قرار

خواهیم داد. نمودار یک تابع با مقدار حقیقی، مثلاً تابعی $\mathbb{R} \rightarrow \mathbb{R}^2$ را در نظر بگیرید. با توجه به خمیدگی‌های ممکن برای این نمودار در جهت‌های مختلف باید انتظار داشت که به ازای هر جهت ساطع از نقطه a ، بتوان عددی نسبت داد که آهنگ تغییر f در آن جهت را بیان کند.

”جهت‌های“ مختلف در \mathbb{R}^n به وسیله n تایی‌های به طول واحد مشخص می‌شوند. از آنجا که برای هر n تایی ناصفر v می‌توان نوشت $v = |v| \frac{v}{|v|}$ ؛ با قرار دادن $u_v = \frac{v}{|v|}$ یک n تایی به طول واحد به دست می‌آید و داریم $v = |v|u_v$ ، یعنی هر n تایی v با یک طول $|v|$ و یک ’جهت‘ u_v مشخص می‌شود.

اکنون برای $f: S \rightarrow \mathbb{R}^m$ ، $a \in S$ ، $S \subset \mathbb{R}^n$ یک نقطه درونی S ، و n تایی واحد u در \mathbb{R}^n ، مشتق سویی یا مشتق جهندار تابع f در نقطه a در جهت u را به صورت زیر مطرح می‌کنیم:

$$D_u f(a) = \lim_{h \rightarrow 0} \frac{f(a + hu) - f(a)}{h} \quad (1)$$

که در اینجا $h \in \mathbb{R}$. حد فوق ممکن است وجود داشته باشد یا نباشد. کمی توضیح در مورد این تعریف لازم است. در اینجا نیز مانند حالت یک متغیری نسبت تغییر تابع به تغییر متغیر را در نظر گرفته حد آن را وقتی تغییر متغیر به صفر میل کند مورد نظر قرار داده‌ایم. ولی توجه کنید که بررسی تغییرات متغیر به امتداد خط راست گذرا از a به موازات u محدود شده است. برای تابع‌های $f: S \rightarrow \mathbb{R}$ ، یعنی برای $m = 1$ ، می‌توان یک تعبیر هندسی مشابه حالت یک متغیری ارائه کرد. مثلاً در شکل ۱ تابعی $\mathbb{R}^2 \rightarrow \mathbb{R}$ در نظر گرفته شده

است، $z = f(x, y)$. خط گذرا از a به موازات u را l می‌نامیم. فقط شامل مقادیر f روی این خط است. حال صفحه π را در نظر بگیرید که شامل خط l و عمود بر صفحه xy است. این صفحه نمودار را در یک منحنی قطع می‌کند که نمودار تحدید دامنه f به خط l است. در شکل (ب) صفحه π و بخشی از نمودار را که بالای سر آن قرار دارد رسم کرده‌ایم. اگر برای l جهت مثبت را جهت u و برای امتداد عمود بر آن جهت مثبت را موازی جهت مثبت محور z منظور کنیم می‌توان در صفحه

π ، عیناً مانند صفحه دکارتی xy ، شیب خطوط راست مماس بر نمودار تابع‌های یک متغیری را مطرح نمود. در وضعیت موجود، $D_u f(a)$ دقیقاً ضریب زاویه خط مماس بر منحنی تقاطع π با نمودار f در صفحه π است.

ذکر این نکته لازم است که برای $f : S \rightarrow \mathbb{R}^m$ ، $S \subset \mathbb{R}^n$ ، کسر $\frac{f(a+hu)-f(a)}{h}$ عنصری از \mathbb{R}^m است، پس حد آن نیز، در صورت وجود، عنصری از \mathbb{R}^m خواهد بود. ضمناً چون حد یک m تایی مؤلفه به مؤلفه منظور می‌شود، اگر $f = (f_1, \dots, f_m)$ ، آنگاه:

$$D_u f(a) = (D_u f_1(a), \dots, D_u f_m(a)) \quad (2)$$

مثال ۱. $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت $f(x, y) = x^2 - y^2$ تعریف می‌کنیم. در نقطه $a = (1, 0)$ ، مشتق جهتدار f را در جهت‌های گوناگون بررسی می‌کنیم. هر عنصر با طول یک از \mathbb{R}^2 را می‌توان به صورت $u = (\cos \theta, \sin \theta)$ نوشت که در آن θ مختصه قطبی در \mathbb{R}^2 است. بنابراین داریم:

$$\begin{aligned} \frac{f(a+hu) - f(a)}{h} &= \frac{f(1+h\cos\theta, h\sin\theta) - f(1, 0)}{h} \quad h \neq 0 \\ &= \frac{(1+h\cos\theta)^2 - h^2\sin^2\theta - 1}{h} \\ &= 2\cos\theta + h(\cos^2\theta - \sin^2\theta) \end{aligned}$$

بنابراین

$$D_u f(1, 0) = 2\cos\theta$$

چند نمونه را در نظر می‌گیریم و تعبیر هندسی مشتق سویی را روی نمودار (شکل ۲) بررسی می‌کنیم. نخست برای $\theta = 0$ خط موازی محور x گذرا از $(1, 0)$ را در نظر می‌گیریم. مقطع صفحه قائمی از این خط می‌گذرد و بر صفحه xy عمود است یک سهمی است ($z = x^2$) که در شکل ۲ (الف) نمایش

داده شده است ضریب زاویه مماس برابر $2 \cos \theta = 2$ می باشد. برای $\theta = \frac{\pi}{4}$ ، اگر دامنه نمودار f را به خط گذرا از $(1, 0)$ با ضریب زاویه $\frac{\pi}{4}$ محدود کنیم حاصل می شود:

$$\begin{aligned} z &= f\left((1, 0) + h\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right) \\ &= \left(1 + \frac{h}{\sqrt{2}}\right)^2 - \left(\frac{h}{\sqrt{2}}\right)^2 \\ &= 1 + \sqrt{2}h \end{aligned}$$

چون عبارت سمت راست نسبت به h درجه یک است مقطع با نمودار یک خط راست می باشد، شکل ۲ (ب) که ضریب زاویه آن $\sqrt{2} \cos \frac{\pi}{4} = \sqrt{2}$ است. بالاخره با محدود کردن دامنه نمودار f به خط قائم گذرا از $(1, 0)$ ، داریم:

$$\begin{aligned} z &= f\left((1, 0) + h(0, 1)\right) \\ &= f(1, h) \\ &= 1 - h^2 \end{aligned}$$

که یک سهمی به دست می دهد، شکل ۲ (ج)، و ضریب زاویه مماس $2 \cos \frac{\pi}{4} = 0$ است. بدین ترتیب ملاحظه می کنیم که مشتق سوپی در جهت های مختلف مقادیر مختلف دارد. قابل ذکر است که اگر جهت معکوس شود (یا به اندازه π به θ افزوده شود) مقدار مشتق جهتدار در (-1) ضرب می شود. دلیل این است که با جایگزینی $-u$ به جای u ، جهت محور افقی که جهت افزایش h است معکوس می شود. اگر یک تابع مشتق پذیر یک متغیری در یک جهت افزایش یابد، در جهت معکوس به همان آهنگ کاهش می یابد.

در یک مورد خاص، نماد ویژه ای برای مشتق جهتدار وجود دارد. فرض کنید u یکی از اعضای پایه متعارف \mathbb{R}^n باشد، $u = e_j = (0, \dots, 0, 1, 0, \dots, 0)$ که در مکان j -ام n تایی است. در این صورت به جای $D_{e_i} f(a)$ برای تابع $y = f(x)$ معمولاً یکی از نمادهای زیر به کار می رود:

$$\frac{\partial y}{\partial x_j}(a), \frac{\partial f}{\partial x_j}(a), f_{x_j}(a), D_j f(a), D_{x_j} f(a)$$

و $D_{e_j} f(a)$ مشتق پاره‌ای (یا مشتق نسبی، یا مشتق جزئی) تابع f نسبت به x_j در نقطه a خوانده می‌شود. به عبارت تعریف کننده مشتق پاره‌ای توجه کنید:

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h} \quad (3)$$

ملاحظه کنید که این در واقع عمل مشتق‌گیری نسبت به متغیر x_j است وقتی سایر متغیرها یعنی مقدار $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$ برابر ثابت‌های به ترتیب $a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n$ نگاه داشته شود. بنابراین محاسبه $\frac{\partial f}{\partial x_j}(a)$ بسیار ساده است، کافی است متغیرهای غیر از x_j را ثابت فرض کرده از قوانین مشتق‌گیری یک متغیری برای محاسبه مشتق نسبت به متغیر x_j استفاده کنیم.

مثال ۱. تابع $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ به صورت $f(x, y, z) = x^2 y - z^3$ تعریف شده است. مشتق‌های پاره‌ای f را نسبت به x, y, z در نقطه دلخواه (x, y, z) محاسبه کنید. داریم:

$$\frac{\partial f}{\partial x} = 2xy \quad , \quad \frac{\partial f}{\partial y} = x^2 \quad , \quad \frac{\partial f}{\partial z} = -3z^2$$

مثال ۲. تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ به صورت زیر تعریف شده است:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

فرمولی برای $\frac{\partial f}{\partial y}$ در نقطه (x, y) پیدا کنید.

دو حالت $(x, y) = (0, 0)$ و $(x, y) \neq (0, 0)$ را جداگانه بررسی می‌کنیم. نخست برای $(0, 0)$ ، $\frac{\partial f}{\partial y}(0, 0)$ توجه کنید که برای $x = 0$ ثابت، داریم $f(0, y) = 0$ برای هر x ، پس $\frac{\partial f}{\partial x}(0, 0) = 0$ در واقع $\frac{\partial f}{\partial x}(0, y) = 0$ برای هر y . حال فرض کنید $(x, y) \neq (0, 0)$. چون برای حد‌گیری فقط مقادیر دو متغیر نزدیک به این (x, y) مطرح هستند، می‌توانیم همواره فرض کنیم که فقط عبارت $\frac{x^2 y}{x^2 + y^2}$ مورد

استفاده است و از این عبارت نسبت به y مشتق بگیریم:

$$\begin{aligned}\frac{\partial f}{\partial y} &= \frac{x^2(x^2 + y^2) - 2y(x^2y)}{(x^2 + y^2)^2} \\ &= \frac{x^4 - x^2y^2}{(x^2 + y^2)^2}\end{aligned}$$

مثال ۳. طبق (۲)، اگر برد تابع، \mathbb{R}^m باشد، مشتق‌های پاره‌ای تابع مؤلفه به مؤلفه محاسبه می‌شوند. به

عنوان مثال فرض کنید $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ به صورت زیر داده شده است:

$$f(x, y, z) = (xyz^2, 2y - z)$$

در این صورت داریم:

$$\frac{\partial f}{\partial x} = (yz^2, 0)$$

$$\frac{\partial f}{\partial y} = (xz^2, 2)$$

$$\frac{\partial f}{\partial z} = (2xyz, -1)$$

یادداشت. در آنچه گذشت عبارت (۱) را به عنوان مشتق سویی، فقط برای n تایی‌های واحد u مطرح

کردیم، ولی می‌توان این عبارت را برای هر n تایی v نوشت و حاصل، مشتق f نسبت به v در نقطه a

خوانده می‌شود. نخست اگر $v \neq 0$ ، می‌نویسیم $v = |v|u_v$ که در اینجا $u_v = \frac{v}{|v|}$ بردار واحد در جهت v

است. داریم:

$$\begin{aligned}D_v f(a) &= \lim_{h \rightarrow 0} \frac{f(a + hv) - f(a)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(a + (h|v|)u_v) - f(a)}{h} \\ &= \lim_{h \rightarrow 0} |v| \frac{f(a + (h|v|)u_v) - f(a)}{h|v|}\end{aligned} \quad (4)$$

اگر $|v|$ را به k نمایش دهیم، چون $|v|$ ثابت و ناصفر است، $h \rightarrow 0$ معادل $k \rightarrow 0$ است، پس:

$$\begin{aligned}
 D_v f(a) &= |v| \lim_{k \rightarrow 0} \frac{f(a + ku_v) - f(a)}{k} \\
 &= |v| D_{u_v} f(a)
 \end{aligned}$$

برای $v = 0$ ، تعریف می‌کنیم $D_v f(a) = 0$ ، که با این رابطه سازگار است و به هر حال صورت کسر طرف راست (۴) همواره صفر است که بنابراین حد آن نیز صفر خواهد شد.

پس مشتق f نسبت به v در نقطه a ، $|v|$ برابر مشتق سویی f در جهت u_v در نقطه a است. مطلب را می‌توان اینگونه تعبیر کرد: اگر متحرکی از نقطه a با سرعت ثابت $|v|$ در جهت u_v در دامنه حرکت کند، $D_v f(a)$ آهنگ تغییر لحظه‌ای f در نقطه a در جهت حرکت است، که متناسب با مقدار $|v|$ می‌باشد.

مشتق و تقریب خطی (۲)

در بخش گذشته مفهوم مشتق به عنوان آهنگ تغییر لحظه‌ای را تعمیم دادیم و به مفهوم مشتق سویی دست یافتیم. اکنون می‌خواهیم تعبیر دیگر مشتق، به عنوان خط مماس یا تقریب خطی، را برای تابع‌های $f : S \rightarrow \mathbb{R}^m$ پیاده کنیم. طبیعی است که به جای "خط" باید شیء مسطح با بعد مناسب مطرح شود. در واقع کوشش می‌کنیم تعریف مماس در حالات یک متغیری را عیناً تعمیم دهیم.

فرض کنید $f : S \rightarrow \mathbb{R}^m$ و $g : T \rightarrow \mathbb{R}^m$ دو تابع باشند که به ترتیب روی دامنه‌های S و T ، زیر مجموعه‌هایی از \mathbb{R}^n ، تعریف شده‌اند، و a یک نقطهٔ درونی هر دو مجموعهٔ S و T است. در این صورت می‌گوییم f در نقطهٔ a بر g مماس است در صورتی که:

$$f(a) = g(a) \quad (۱)$$

$$\lim_{x \rightarrow a} \frac{|f(x) - g(x)|}{|x - a|} = 0 \quad (۲)$$

برای $f : S \rightarrow \mathbb{R}^m$ داده شده، f را مشتق‌پذیر در نقطه a می‌نامیم در صورتی که تابعی مستوی $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ وجود داشته باشد که f در نقطهٔ a بر A مماس باشد، یعنی:

$$f(a) = A(a) \quad (۳)$$

$$\lim_{x \rightarrow a} \frac{|f(x) - A(x)|}{|x - a|} = 0 \quad (4)$$

در این صورت A ، تقریب خطی f در نقطه a خوانده می شود (اصطلاح "تقریب مستوی" دقیق تر است ولی "تقریب خطی" بسیار رایج تر). این تعریف را کمی می شکافیم. هر تابع مستوی A به شکل زیر است

$$A(x) = L(x) + B \quad (5)$$

که در آن $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ یک تابع خطی است و B یک m تایی ثابت در \mathbb{R}^m . از (3) داریم:

$$B = f(a) - L(a) \quad (6)$$

پس $A(x) = L(x) + f(a) - L(a)$ ، یا:

$$A(x) = f(a) + L(x - a) \quad (7)$$

زیرا که L خطی است. با جایگزینی در (4) داریم:

$$\lim_{x \rightarrow a} \frac{|f(x) - [f(a) + L(x - a)]|}{|x - a|} = 0 \quad (8)$$

اگر $x - a$ را به \vec{h} نمایش دهیم، (8) را می توان به صورت معادل زیر نیز نوشت:

$$\lim_{\vec{h} \rightarrow \underline{0}} \frac{|f(a + \vec{h}) - f(a) - L(\vec{h})|}{|\vec{h}|} = 0 \quad (9)$$

البته تابع خطی $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ را معمولاً به صورت یک ماتریس $m \times n$ ، مانند M ، می نویسیم.

بنابراین تقریب خطی، A ، شکل زیر را دارد:

$$|A(x)\rangle = M|x\rangle + |B\rangle \quad (10)$$

که شباهت آن با $A(x) = mx + b$ واضح است. تابع خطی L (با معادلاً ماتریس M ، $m \times n$) را، در صورت مشتق‌پذیری f در a ، مشتق (یا دیفرانسیل) f در نقطه a می‌نامند و به $Df(a)$ نمایش می‌دهند.

(۲۱-۱) محاسبه مشتق. اکنون نشان می‌دهیم که ماتریس مربوط به $Df(a)$ را می‌توان به سادگی محاسبه کرد. اگر M ماتریس تابع خطی $Df(a)$ باشد، می‌دانیم که ستون j -ام آن از مؤلفه‌های $(Df(a))(e_j)$ تشکیل شده است. اگر در کسر (۹) به جای L ، $Df(a)$ و به جای h ، he_j را جایگزین کنیم (که $h \in \mathbb{R}$)، نتیجه می‌شود:

$$\frac{f(a + he_j) - f(a) - (Df(a))(he_j)}{|he_j|} = \frac{|f(a + he_j) - f(a) - h(Df(a))(e_j)|}{|h|}$$

$$= \left| \frac{f(a + he_j) - f(a)}{h} - (Df(a))(e_j) \right|$$

بنابراین اگر f در نقطه a مشتق‌پذیر باشد، داریم:

$$\frac{\partial f}{\partial x_j}(a) = (Df(a))(e_j) \quad (11)$$

بدین ترتیب اگر f در نقطه a مشتق‌پذیر باشد، ستون j -ام ماتریس $Df(a)$ ، از مشتق‌های پاره‌ای مؤلفه‌های f تشکیل شده، پس $Df(a)$ نمایش ماتریسی زیر را دارد:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix} \quad (12)$$

گاهی اصطلاح ماتریس ژاکوبی نیز برای این ماتریس به کار می‌رود.

مثال ۱. برای تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ، $f(x, y) = x^2 - y^2$ ، با فرض مشتق‌پذیری f در نقطه $(2, 1)$ ، معادله صفحه مماس بر نمودار را در نقطه $(2, 1, 3)$ پیدا کنید. در اینجا $Df(2, 1)$ به وسیله ماتریس 1×2 زیر داده می‌شود

$$\left[\frac{\partial f}{\partial x}(2, 1) \quad \frac{\partial f}{\partial y}(2, 1) \right] = [4 \quad -2]$$

بنابراین طبق (۷):

$$A(x, y) = 3 + 4(x - 2) - 2(y - 1)$$

که اگر بنویسیم $z = A(x, y)$ ، معادلهٔ صفحهٔ مماس در فضای x, y, z عبارت است از:

$$4x - 2y - z - 3 = 0$$

مثال ۲. فرض کنید S زیرمجموعه‌ای از \mathbb{R}^n ، $a = (a_1, \dots, a_n)$ یک نقطهٔ درونی S و $f: S \rightarrow \mathbb{R}$ مشتق پذیر باشد. معادلهٔ ابرصفحهٔ مماس بر نمودار f را در نقطهٔ $(a_1, \dots, a_n, f(a))$ محاسبه کنید. نقاط دامنه را به $x = (x_1, \dots, x_n)$ و مقدار تابع را به $x_{n+1} = f(x_1, \dots, x_n)$ نمایش می‌دهیم. $Df(a)$ نمایش ماتریسی زیر را به عنوان یک ماتریس $1 \times n$ دارد:

$$\left[\frac{\partial f}{\partial x_1}(a) \cdots \frac{\partial f}{\partial x_n}(a) \right]$$

بنابراین طبق (۷)، معادلهٔ ابرصفحهٔ مماس، یا عبارت تقریب خطی، عبارت است از:

$$x_{n+1} = f(a) + \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n) \quad (13)$$

شبهات این عبارت به معادلهٔ خط مماس برای تابع‌های $\mathbb{R} \rightarrow \mathbb{R}$ ، یعنی $y = f(a) + f'(a)(x - a)$ مشهود است.

در اینجا می‌توانیم رابطهٔ مشتق پذیری، مشتق، و مشتق‌های سوپی را تشریح کنیم. فرض کنید $f: S \rightarrow \mathbb{R}^m$ در نقطهٔ درونی a از S مشتق پذیر باشد. با محاسبه‌ای مشابه آنچه به یافتن درایه‌های ماتریس ژاکوبی منجر شد، رابطهٔ $Df(a)$ و مشتق سوپی را به دست می‌آوریم. فرض کنید v یک عضو ناصفر \mathbb{R}^n باشد. اگر در کسر (۹)، به جای \vec{h} مقدار $h\vec{v}$ را جایگزین کنیم، $h \in \mathbb{R}$ ، و به جای L بنویسیم

$Df(a)$ ، داریم:

$$\begin{aligned} \frac{|f(a+hv) - f(a) - Df(a)(hv)|}{|hv|} &= \frac{1}{|v|} \frac{|f(a+hv) - f(a) - h(Df(a))(v)|}{|h|} \\ &= \frac{1}{|v|} \left| \frac{f(a+hv) - f(a)}{h} - (Df(a))(v) \right| \end{aligned}$$

طبق فرض مشتق‌پذیری f در a ، عبارت بالا دارای حد صفر است، ولی حد عبارت اول داخل $| \quad |$ در فوق همان مشتق نسبت به v ، یعنی $D_v f(a)$ ، است. ضمناً برای $v = 0$ نیز تساوی $D_v f(a) = Df(a)(v)$ برقرار است زیرا دوطرف رابطه صفر هستند. بنابراین:

(۲۱-۲) گزاره. اگر f در a مشتق‌پذیر باشد برای هر $v \in \mathbb{R}^n$ داریم

$$D_v f(a) = (Df(a))(v) \quad (14)$$

■

بالاخص این نتیجه مهم نشان می‌دهد که اگر تابع در a مشتق‌پذیر باشد، میان مشتق‌های جهت‌دار در جهت‌های گوناگون رابطه‌ای وجود دارد زیرا که همه آنها با دانستن مشتق‌های پاره‌ای (درایه‌های $Df(a)$) و دانستن جهت، یعنی u ، قابل محاسبه‌اند. از نظر هندسی نیز باید همین را انتظار داشت زیرا که در صورت مشتق‌پذیری، خطوط مماس در جهات مختلف همه در یک شیء مسطح، یعنی نمودار تابع مستوی مماس باشد، قرار می‌گیرند. که همه خطوط مماس در یک زیرفضای مستوی قرار می‌گیرند برای تابعهایی که مشتق‌پذیر نباشند ممکن است برقرار نباشد. مثال زیر نمونه ساده‌ای از چنین وضعیتی است.

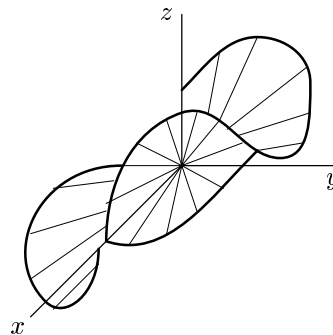
مثال. تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

نشان می دهیم f در $(0, 0)$ مشتق پذیر نیست. عبارت $\frac{x^2y}{x^2+y^2}$ در مختصات قطبی تبدیل می شود به $r \cos^2 \theta \sin \theta$. مقدار این عبارت در مبدأ مختصات، یعنی $r = 0$ ، صفر است، پس به طور کلی

$$f(x, y) = r \cos^2 \theta \sin \theta$$

اگر دامنه f را به یک خط راست گذرا از مبدأ محدود کنیم، نمودار یک خط راست است زیرا که برای نیم خط تعریف شده توسط θ ثابت داریم $z = (\cos^2 \theta \sin \theta)r$ ، یعنی z تابعی خطی نسبت به r (فاصله از 0) است، و برای نیم خط مقابل (یعنی به ازای $\theta + \pi$)، $\sin(\theta + \pi) = -\sin \theta$ ، پس در واقع $z = (\cos^2 \theta \sin \theta)r$ یک خط راست بالا سر خط $\{\theta, \theta + \pi\}$ تعریف می کند. به ازای $\theta = 0, \pi$ و نیز به ازای $\theta = \frac{\pi}{4}, \frac{3\pi}{4}$ ، داریم $f(x, y) = 0$ ، یعنی محورهای x و y جزء نمودار تابع f هستند. بنابراین اگر f در $(0, 0)$ مشتق پذیر می بود، می بایست دو محور x و y جزء صفحه مماس باشند و در نتیجه صفحه مماس می بایست همان صفحه xy باشد. ولی به ازای مقادیر θ غیر از مقادیر مربوط به دو محور، داریم $(\cos^2 \theta)(\sin \theta) \neq 0$ و خط راست $z = (\cos^2 \theta \sin \theta)r$ با صفحه xy در 0 برخورد غیر مماس دارد. این نکته نشان می دهد که تنها نامزد ممکن صفحه مماس برای نمودار f در $(0, 0)$ در واقع بر همه نمودار مماس نیست، یعنی تابع f در $(0, 0)$ مشتق پذیر نیست (شکل ۱)!



شکل ۱

ضمناً توجه کنید که در $(0, 0)$ ، f دارای مشتق سویی در همه جهتهاست زیرا با قرار دادن

داریم: $u = (\cos \theta, \sin \theta)$

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(\underline{0} + hu) - f(\underline{0})}{h} &= \lim_{h \rightarrow 0} \frac{h^2 \cos^2 \theta \sin \theta}{h^2} \\ &= \cos^2 \theta \sin \theta \end{aligned}$$

اکنون مشتق ناپذیری f در $(0, 0)$ را به طور دقیق تر با توجه به تعاریف و نتایج به دست آمده توضیح می دهیم. فرض کنید f در $(0, 0)$ مشتق پذیر باشد. در این صورت مساتریس $Df(a)$ هست $[\frac{\partial f}{\partial x}(0, 0) \quad \frac{\partial f}{\partial y}(0, 0)]$. با توجه به این که f روی هر دو محور مقدار ثابت صفر را داراست، نتیجه می شود که $\frac{\partial f}{\partial x}(0, 0) = 0$ و $\frac{\partial f}{\partial y}(0, 0) = 0$. بنابراین با توجه به این که $f(0, 0) = 0$ ، صفحه مماس در $(0, 0)$ باید $z = 0$ باشد. حال اگر تابع در $(0, 0)$ مشتق پذیر می بود، طبق (۹) با قراردادن $a = (0, 0)$ و $\vec{h} = (x, y)$ باید می داشتیم:

$$\lim_{(x,y) \rightarrow (0,0)} \frac{|\frac{x^2 y}{x^2 + y^2}|}{\sqrt{x^2 + y^2}} = 0$$

یا در مختصات قطبی:

$$\lim_{(x,y) \rightarrow (0,0)} (\cos^2 \theta |\sin \theta|) = 0$$

چون $\cos^2 \theta |\sin \theta|$ روی هر خط گذرا از مبدأ مقدار ثابتی دارد و این مقدار ثابت غیر از روی دو محور صفر نیست، حد بالا نمی تواند برقرار باشد.

این نکته را باید تأکید کرد که علیرغم وجود $\frac{\partial f}{\partial x}(0, 0)$ و $\frac{\partial f}{\partial y}(0, 0)$ در واقع همه مشتقهای سویی در $(0, 0)$ ، تابع در $(0, 0)$ مشتق پذیر نیست. مطالب این جلسه نشان دادند اگر تابع در نقطه a مشتق پذیر باشد، مشتقهای پاره‌ای، و در واقع مشتقهای سویی، همه در نقطه a وجود دارند و رابطه کلی (۱۴) برقرار است. ولی وجود مشتقهای پاره‌ای (و حتی وجود همه مشتقهای سویی)، بر مشتق پذیری دلالت نمی کند. در مثال فوق خطوط مماس در جهات مختلف در یک زیرفضای مسطح واحد قرار نگرفتند، یعنی یک تقریب خطی در برگیرنده همه جهات وجود نداشت.

در مثالی که در زیر می آید، با اینکه خطوط مماس در همه جهتها در یک صفحه قرار می گیرند، این صفحه بر نمودار مماس نخواهد بود! در واقع در مثال زیر تابع داده شده در نقطه مزبور حتی پیوسته نیست در حالیکه مشتقهای سوئی تابع در آن نقطه در همه جهات موجود است.

مثال ۴. خم C در صفحه را که در مختصات قطبی به صورت زیر تعریف شده است در نظر بگیرید:

$$r = 2\pi - \theta, \quad 0 \leq \theta < 2\pi$$

تصویر C در شکل ۲ نمایش داده شده است. توجه کنید که C هر خط راست گذرا از \underline{e} را

شکل ۲

در دقیقاً دو نقطه قطع می کند. حال تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم:

$$f(x, y) = \begin{cases} 0 & \text{اگر } (x, y) \text{ روی } C \text{ نباشد} \\ 1 & \text{اگر } (x, y) \text{ روی } C \text{ باشد} \end{cases}$$

تحدید این تابع به هر خط گذرا از مبدأ در همه جا به استثنای دو نقطه غیر \underline{e} صفر است، پس مشتق سوئی تابع در \underline{e} در همه جهتها وجود دارد و برابر صفر است. بدین ترتیب خطوط مماس در \underline{e} در همه جهتها افقی هستند و روی صفحه xy قرار می گیرند. ولی این تابع در $(0, 0)$ پیوسته نیست زیرا اگر به $(0, 0)$ در طول C نزدیک شویم مقدار f به ۱ میل می کند و اگر از بیرون C نزدیک شویم مقدار f به ۰ میل می کند. این تابع در $(0, 0)$ مشتق پذیر نیست زیرا طبق گزاره زیر اگر f در نقطه ای مشتق پذیر باشد، در آن نقطه پیوسته نیز هست.

(۲۱-۳) گزاره. اگر f در a مشتق پذیر باشد، f در a پیوسته است.

برهان. طبق تعریف مشتق پذیری داریم

$$\lim_{\vec{h} \rightarrow 0} \frac{|f(a + \vec{h}) - f(a) - Df(a)(\vec{h})|}{|\vec{h}|} = 0$$

پس حد صورت زیر صفر است

$$\lim_{\vec{h} \rightarrow 0} |f(a + \vec{h}) - f(a) - Df(a)(\vec{h})| = 0$$

چون $Df(a)$ یک تابع خطی است، $Df(a)$ پیوسته است و مقدار تابع خطی در $\vec{h} = 0$ برابر صفر

است، پس داریم:

$$\lim_{\vec{h} \rightarrow 0} |f(a + \vec{h}) - f(a)| = 0$$

■ که این در واقع بیان پیوستگی f در نقطه a است.

مشتق و تقریب خطی (۳)

مثالهای بخش پیشین نشان داد که تحقیق مشتق‌پذیری در یک تابع یک نقطه از دامنه ممکن است موضوع ظریفی باشد. خوشبختانه ضابطه‌ای کافی برای مشتق‌پذیری وجود دارد که در بسیاری موارد کارساز و قابل استفاده است. قضیه زیر بیان این ضابطه است:

(۲۴-۱) قضیه. فرض کنید $f : S \rightarrow \mathbb{R}^m$ ، $S \subset \mathbb{R}^n$ طوری باشد که تابع‌های $\frac{\partial f}{\partial x_i} : S \rightarrow \mathbb{R}^m$ همه در یک گوی باز حول نقطه درونی a از S وجود داشته و در نقطه a پیوسته باشند. در این صورت f در نقطه a مشتق‌پذیر است.

بدین ترتیب برای تحقیق مشتق‌پذیری f کافی است پس از اطمینان از این که مشتق‌های پاره‌ای در نزدیکی a وجود دارند، پیوستگی هر یک را در آن نقطه تحقیق کنیم. اثبات زیر در واقع نشان می‌دهد که پیوستگی $(n-1)$ مشتق پاره‌ای در نقطه a و وجود مشتق پاره‌ای باقیمانده در آن نقطه برای مشتق‌پذیری کافی است.

برهان. قضیه را نخست در حالت دو متغیری ثابت می‌کنیم که ایده اصلی را نشان می‌دهد، سپس روش تعمیم آن به n متغیر را توضیح می‌دهیم. بدین ترتیب فرض کنید $S \subset \mathbb{R}^2$ ، (a, b) یک نقطه درونی S است، $\frac{\partial f}{\partial x}(a, b)$ وجود دارد، و $\frac{\partial f}{\partial y}$ در یک گوی باز حول (a, b) وجود دارد و در نقطه (a, b) پیوسته است. برای مشتق‌پذیری f در (a, b) باید ثابت کنیم:

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(a+h, b+k) - [f(a, b) + h \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b)]}{\sqrt{h^2 + k^2}} = 0 \quad (1)$$

چون (a, b) یک نقطه درونی S است، در بالا و در باقیمانده بحث می‌توان فرض کرد که $|h| > 0$ و $|k| > 0$ آنقدر کوچک هستند که مستطیل $a - |h| \leq x \leq a + |h|$ ، $b - |k| \leq y \leq b + |k|$ در دامنه تعریف S قرار دارد. صورت کسر (۱) را با افزودن و کم کردن عبارت $f(a + h, b)$ به صورت زیر بازنویسی می‌کنیم:

$$[f(a + h, b + k) - f(a + h, b) - k \frac{\partial f}{\partial y}(a, b)] + [f(a + h, b) - f(a, b) - h \frac{\partial f}{\partial x}(a, b)]$$

بنابراین کافی است دو حد زیر را تحقیق کنیم

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|f(a + h, b) - f(a, b) - h \frac{\partial f}{\partial x}(a, b)|}{\sqrt{h^2 + k^2}} = 0 \quad (2)$$

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|f(a + h, b + k) - f(a + h, b) - k \frac{\partial f}{\partial y}(a, b)|}{\sqrt{h^2 + k^2}} = 0 \quad (3)$$

در مورد (۲) داریم

$$\begin{aligned} \left| \frac{f(a+h, b) - f(a, b) - h \frac{\partial f}{\partial x}(a, b)}{\sqrt{h^2 + k^2}} \right| &\leq \left| \frac{f(a+h, b) - f(a, b) - h \frac{\partial f}{\partial x}(a, b)}{h} \right| \\ &= \left| \frac{f(a+h, b) - f(a, b)}{h} - \frac{\partial f}{\partial x}(a, b) \right| \end{aligned}$$

و اینکه حد عبارت سمت راست صفر است به معنای وجود مشتق $\frac{\partial f}{\partial x}$ در (a, b) است. برای (۳)، از آنجا که $\frac{\partial f}{\partial y}$ در یک گوی حول (a, b) تعریف شده است، با تثبیت h کوچک، $f(a + h, y)$ برای y در نزدیکی b تابعی مشتق‌پذیر نسبت به متغیر دوم است و با استفاده از قضیه میانگین یک متغیری می‌توان نوشت:

$$f(a + h, b + k) - f(a + h, b) = k \frac{\partial f}{\partial y}(a + h, b')$$

که در آن b' بین b و $b + k$ است. پس صورت کسر (۳) را می‌توان به صورت زیر نوشت

$$k \left[\frac{\partial f}{\partial y}(a + h, b') - \frac{\partial f}{\partial y}(a, b) \right]$$

حال چون $1 \leq \left| \frac{k}{\sqrt{h^2+k^2}} \right|$ ، قدرمطلق کسر (۳) کوچکتر یا مساوی $\left| \frac{\partial f}{\partial y}(a+h, b') - \frac{\partial f}{\partial y}(a, b) \right|$ است. از آنجاکه b' بین b و $b+k$ است، وقتی $(h, k) \rightarrow (0, 0)$ ، نقطه $(a+h, b')$ به (a, b) میل می کند و چون $\frac{\partial f}{\partial y}$ در (a, b) پیوسته فرض شده است، این عبارت به صفر میل می کند و اثبات حکم کامل می شود. برای اثبات حکم در حالت n متغیری، روش بالا را به صورت زیر تعمیم می دهیم. فرض کنید نقطه درونی $a = (a_1, \dots, a_n)$ از مجموعه S داده شده است و نقطه $(a_1 + h_1, \dots, a_n + h_n)$ در دامنه f طوری است که مستطیل $a_1 - |h_1| \leq x_1 \leq a_1 + |h_1|, \dots, a_n - |h_n| \leq x_n \leq a_n + |h_n|$ به تمامی در دامنه تعریف f قرار می گیرد. حال با تفریق و افزودن $(n-1)$ مقدار $f(a_1 + h_1, \dots, a_{n-1} + h_{n-1}, a_n), f(a_1 + h_1, \dots, a_{n-2} + h_{n-2}, a_{n-1}, a_n), \dots, f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a)$ عبارت $f(a_1 + h_1, a_2, \dots, a_n)$ را به صورت مجموع n عبارت زیر می نویسیم:

$$[f(a_1 + h_1, \dots, a_n + h_n) - f(a_1 + h_1, \dots, a_{n-1} + h_{n-1}, a_n) - h_n \frac{\partial f}{\partial x_n}(a)] + \dots + [f(a_1 + h_1, a_2, \dots, a_n) - f(a_1, \dots, a_{n-1}, a_n) - h_1 \frac{\partial f}{\partial x_1}(a)]$$

و نشان می دهیم نسبت تقسیم قدرمطلق هر گروه بر $\sqrt{h_1^2 + \dots + h_n^2}$ به صفر میل می کند. برای گروه آخر، این امر از وجود $\frac{\partial f}{\partial x_1}(a)$ نتیجه می شود و برای سایر گروه ها مثل استدلال بالا از قضیه میانگین و پیوستگی $\frac{\partial f}{\partial x_i}$ در نقطه a استفاده می کنیم. استدلال عین بالا است. \square

مثال. به مثالی که در هر یک از دو بخش قبل مورد بحث قرار گرفت باز می گردیم، یعنی تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ با ضابطه:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

دیدیم که این تابع در $(0, 0)$ مشتق پذیر نیست. و در عین حال $\frac{\partial f}{\partial x}(0, 0)$ و $\frac{\partial f}{\partial y}(0, 0)$ وجود دارند. بنابراین طبق قضیه هیچ یک از $\frac{\partial f}{\partial y} \frac{\partial f}{\partial x}$ نمی توانند در $(0, 0)$ پیوسته باشند. ضمناً برای $\frac{\partial f}{\partial y}$ در دو

بخش قبل عبارت زیر را به دست آوریم:

$$\frac{\partial f}{\partial y}(x, y) = \begin{cases} \frac{x^2 - x^2 y^2}{(x^2 + y^2)^{\frac{3}{2}}} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

در مختصات قطبی خارج از $(0, 0)$ داریم $\frac{\partial f}{\partial y} = \cos^2 \theta (\cos^2 \theta - \sin^2 \theta)$ روی محور x ، $\theta = 0, \pi$ و $\frac{\partial f}{\partial x}(x, 0) = 1$ برقرار است. بنابراین حد $\frac{\partial f}{\partial y}$ وقتی $(x, y) \rightarrow (0, 0)$ برابر 0 نیست، و $\frac{\partial f}{\partial x}$ در $(0, 0)$ پیوسته نیست. مشابهاً با محاسبه $\frac{\partial f}{\partial x}$ در نقاط خارج از مبدأ می توان مشاهده کرد که $\frac{\partial f}{\partial x}$ نیز در $(0, 0)$ پیوسته نیست.

اکنون به جنبه های کاربردی تر تقریب خطی می پردازیم. همان طور که دیدیم در بین تابع های درجه یک که نمودارشان از یک نقطه خاص نمودار تابع داده شده می گذرد، تقریب خطی یگانه تابعی است که تفاضل مقادارش با مقدار متناظر تابع داده شده سریعتر از فاصله متغیر از نقطه مورد بحث به صفر میل می کند. این ویژگی باید روش تقریب مناسبی را پیش پای ما بگذرد زیرا که از یک سو تقریب خوبی مطرح است و از سویی دیگر محاسبه تابع های درجه یک بسیار ساده است.

مثال ۱. به کمک روش تقریب خطی، مقداری تقریبی برای $\sqrt{x^2 - y}$ پیدا کنید وقتی $x = 2/9$ و $y = 1/1$. ملاحظه می کنیم که به ازای $a = 3$ و $b = 1$ داریم $\sqrt{a^2 - b} = 2$ و نقطه $(x, y) = (2/9, 1/1)$ را می توان برای بعضی مقاصد نزدیک نقطه (a, b) فرض کرد. تابع $f(x, y) = \sqrt{x^2 - y}$ را در نظر می گیریم. مشتق های پاره ای این تابع نسبت به x و y وجود دارند و پیوسته اند وقتی زیر $\sqrt{\quad}$ صفر شود، یعنی وقتی $y \neq x^2$. نقطه $(3, 1)$ روی $y = x^2$ نیست، پس در یک گوی باز حول آن (که شامل $(2/9, 1/1)$ نیز می شود) مشتق های پاره ای وجود دارند. در واقع

$$\frac{\partial f}{\partial x} = \frac{2x}{3}(x^2 - y)^{-\frac{3}{2}}, \quad \frac{\partial f}{\partial y} = -\frac{1}{3}(x^2 - y)^{-\frac{3}{2}}$$

تقریب خطی f را در نقطه $(a, b) = (3, 1)$ می‌نویسیم:

$$A(x, y) = 2 + \frac{\partial f}{\partial x}(3, 1)(x - 3) + \frac{\partial f}{\partial y}(3, 1)(y - 1)$$

$$\begin{aligned} A(2/9, 1/1) &= 2 + \frac{1}{9}(-0/1) + \left(-\frac{1}{12}\right)(0/1) \\ &= 2 - \frac{1}{90} - \frac{1}{120} \end{aligned}$$

پس طبق تقریب خطی:

$$\sqrt[3]{(2/9)^2 - (1/1)} \simeq 2 - \frac{1}{90} - \frac{1}{120} \simeq 1/941667$$

در اینجا روش دیگری برای استفاده از تقریب خطی به نظر می‌رسد. داریم

$$\sqrt[3]{(2/9)^2 - (1/1)} = \sqrt[3]{7/31}$$

مطرح کرد. تابع $f(t) = \sqrt[3]{t}$ در نظر بگیرید. داریم $f(8) = 2$ و می‌خواهیم مقداری تقریبی برای

$$f(7/31) = f(8 - 0/69)$$

$$f(7/31) \simeq f(8) - (0/69)f'(8)$$

$$\text{حال } f'(t) = \frac{1}{3}t^{-2/3} \text{ پس } f'(8) = \frac{1}{12}$$

$$f(7/31) \simeq 2 - \frac{69}{1200} = 1/9425$$

کدامیک از دو تقریب به واقعیت نزدیکتر است؟ تابع دو متغیری $f(x, y) = \sqrt[3]{x^2 - y}$ را در

نظر بگیرید. ترکیب‌های مختلفی از x و y منجر به $\sqrt[3]{7/31}$ می‌شوند. در واقع با قرار دادن

$$x^2 - y = 7/31$$

یک سهمی در صفحه xy به دست می‌آید که جزیی از آن "نزدیک" نقطه $(3, 1)$

شکل ۱

است. بخشی از نمودار $f(x, y) = \sqrt[3]{x^2 - y}$ را که بالای سر این منحنی قرار دارد در نظر بگیرید

که یک منحنی روی نمودار می‌شود. نقاط مختلف این منحنی از تقریب خطی در نقطه $(3, 1)$ فواصل

گوناگون دارند، یعنی تقریب خطی در $(1, 3)$ جوابی یکسان برای همه این نقاط نمی دهد، که امری طبیعی است. برای هر چنین (x, y) ، دو مقدار $(x - 3)$ و $(y - 1)$ بر تقریب خطی به دست آمده اثر دارند. در تقریب خطی یک متغیری، همه این نقاط به صورت یک نقطه روی محور t دیده می شوند و تقریبی که به دست می آید نسبت به پستی بلندی های نمودار $f(x, y) = \sqrt{x^2 - y}$ حساس نیست. جمع بندی بحث بالا این است که اگر مسأله اولیه به طور طبیعی یک مسأله دو متغیری باشد، بهتر است تقریب خطی دو متغیری در نظر گرفته شود که تمایز میان نقاط مختلفی که منجر به یک مسأله یک متغیری می شوند نادیده گرفته نشود. بالعکس اگر یک مسأله یک متغیری را به روش های گوناگون به مسأله ای دو متغیری مبدل کنیم، جواب های گوناگونی حاصل خواهند شد که هر یک ممکن است وابسته به فرض های اضافی و بعضاً غیرطبیعی باشد. در مثال ۱ که مسأله به صورت دو متغیری مطرح بود، روش طبیعی، استفاده از تقریب خطی دو متغیری است.

مثال ۲. $S \subset \mathbb{R}^n$ را به صورت زیر در نظر می گیریم:

$$S = \{(x_1, \dots, x_n) \mid x_i > 0, i = 1, \dots, n\}$$

و تابع $f: S \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم:

$$f(x_1, \dots, x_n) = cx_1^{\alpha_1} \dots x_n^{\alpha_n} \quad (4)$$

که در آن $c, \alpha_1, \dots, \alpha_n$ اعداد حقیقی مثبت داده شده اند. این گونه تابع در بعضی مسائل اقتصادی مطرح می شود که در آن x_i ها مقادیر عوامل مختلف تولید هستند ولی منحصر به این نوع کاربرد نیستند. سؤالی که مطرح می کنیم این است که اگر عوامل x_1, \dots, x_n به ترتیب با دقت نسبی r_1, \dots, r_n درصد معلوم باشند، نتیجه محاسبه، یعنی $f(x_1, \dots, x_n)$ با چه دقت نسبی تعیین می شود؟ اگر a_1, \dots, a_n "مقادیر واقعی" متغیرهای x_1, \dots, x_n باشند، خطاهای نسبی به ترتیب

هستند $\frac{x_1 - a_1}{a_1}, \dots, \frac{x_n - a_n}{a_n}$ را به $x_i - a_i$ نیز نمایش می دهیم. طبق فرمول تقریب خطی داریم:

$$f(x_1, \dots, x_n) \simeq f(a_1, \dots, a_n) + \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(a) \Delta x_i \quad (5)$$

با مشتق گیری از (۱) داریم:

$$\frac{\partial f}{\partial x_i} = c \alpha_i x^{\alpha_1} \dots x_{i-1}^{\alpha_{i-1}} x_i^{\alpha_i - 1} x_{i+1}^{\alpha_{i+1}} \dots x_n^{\alpha_n} = \frac{\alpha_i}{x_i} f(x)$$

بنابراین با جایگزینی در (۲):

$$\Delta f = f(x) - f(a) \simeq f(x) \sum_{i=1}^n \alpha_i \Delta x_i$$

در نتیجه خطای نسبی f ، یعنی $\frac{\Delta f}{f}$ ، از تقریب خطی به صورت زیر است:

$$\frac{\Delta f}{f} \simeq \sum_{i=1}^n \alpha_i \frac{\Delta x_i}{x_i}$$

حال اگر خطای نسبی x_i برابر r_i درصد باشد، نتیجه می شود که خطای نسبی در محاسبه f حدوداً

$\sum_{i=1}^n \alpha_i r_i$ درصد خواهد بود.

میدان گرادیان

فرض کنید تابع $f : S \rightarrow \mathbb{R}$ داده شده است، $S \subset \mathbb{R}^n$ ، که در سراسر S مشتق پذیر است. می دانیم که برای هر $a \in S$ و هر $v \in \mathbb{R}^n$:

$$D_v f(a) = (Df(a))(v) \quad (1)$$

بازنویسی جالب توجهی از (۱) وجود دارد که منشاء بحث این بخش است. چون برد تابع f فضای یک بعدی \mathbb{R} است، ماتریس $Df(a)$ یک ماتریس $1 \times n$ می باشد و وقتی $v = (v_1, \dots, v_n)$ به صورت یک ستون نوشته شود سمت راست (۱) به صورت زیر در می آید:

$$\frac{\partial f}{\partial x_1}(a)v_1 + \dots + \frac{\partial f}{\partial x_n}(a)v_n \quad (2)$$

که حاصل ضرب داخلی دو بردار n تایی v و $(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a))$ است. بردار متشکل از مشتقات پاره ای f را گرادیان f در نقطه a می نامیم و با $\nabla f(a)$ یا $grad f(a)$ نمایش می دهیم:

$$\nabla f(a) = (\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)) \quad (3)$$

تفاوت $\nabla f(a)$ و $Df(a)$ در این است که $\nabla f(a)$ یک n تایی عضو \mathbb{R}^n تلقی می شود و $Df(a)$ یک ماتریس $1 \times n$ بدین ترتیب وقتی تابع f در نقطه a مشتق پذیر باشد، داریم:

$$D_v f(a) = \nabla f(a) \cdot v \quad (4)$$

∇f در واقع به هر نقطه مشتق پذیری تابع f یک بردار n تایی $\nabla f(a)$ نسبت می دهد. معمول است که این تابع را به صورت یک "میدان" تجسم می کنیم به این معنی که بردار $\nabla f(a)$ را روئیده از مبدأ نقطه a در نظر می گیریم.

مثال ۱. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت $f(x) = x^2 - y$ در نظر بگیرید. داریم $\nabla f(x, y) = (2x, -1)$. برای رسم میدان گرادیان، به مبدأ هر نقطه (x, y) بردار $(2x, -1)$ را رسم می کنیم (شکل ۱).

مثال ۲. $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت $f(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2$ در نظر بگیرید. داریم $\nabla f(x) = (2x_1, \dots, 2x_n)$ پس در هر نقطه، برداری به دو برابر بردار حامل از مبدأ به آن و در همان جهت نصب می شود (شکل ۲ برای $n = 2$).

مثال ۳. در مکانیک نیوتنی و دستگاه های پایسته، میدان نیرو در هر نقطه برابر گرادیان $(-U)$ است که U انرژی پتانسیل دستگاه است. مثلاً برای گرانش مرکزی در فضای سه بعدی که جسم سنگینی در مبدأ مختصات منشاء گرانش است، انرژی پتانسیل برابر است با $U(x, y, z) = -\frac{k}{r}$ که $k > 0$ ثابت است و $r = \sqrt{x^2 + y^2 + z^2}$ فاصله از مبدأ می باشد. در اینجا میدان نیرو \vec{F} برابر است با $\vec{F}(x, y, z) = -\frac{k}{r^3}(x, y, z)$.

اکنون نشان خواهیم داد که چگونه میدان گرادیان همراه با مجموعه های تراز تصویری بسیار گویا از رفتار یک تابع حقیقی چند متغیری ارائه می کند. از (۴) نتیجه می گیریم که:

$$D_v f(a) = |\nabla f(a)| |v| \cos \angle(\nabla f(a), v) \quad (5)$$

وقتی v یک بردار به طول واحد باشد و در نتیجه $D_v f(a)$ مشتق سویی در جهت v ، داریم:

$$D_v f(a) = |\nabla f(a)| \cos \angle(\nabla f(a), v) \quad (6)$$

از آنجا که $|\cos \angle| \leq 1$ ، نتیجه می گیریم که:

$$|D_v f(a)| \leq |\nabla f(a)| \quad (7)$$

یعنی قدرمطلق مشتق سویی در یک نقطه حداکثر برابر طول بردار گرادیان در آن نقطه است. در واقع اگر در نقطه a ، $\nabla f(a) \neq 0$ ، برای بردار واحد در جهت $\nabla f(a)$ ، یعنی $v = \frac{\nabla f(a)}{|\nabla f(a)|}$ ، $\cos \angle(\nabla f(a), v) = 1$ متقابلاً برای بردار واحد v در جهت مخالف $\nabla f(a)$ ، کسینوس زاویه، (-1) می شود، پس:

(۲۳-۱) گزاره. اگر در نقطه a ، $\nabla f(a) \neq 0$ ، جهت حداکثر افزایش تابع، جهت گرادیان، و جهت حداکثر کاهش تابع، جهت معکوس گرادیان است. □

مثال ۴. تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ به صورت $f(x, y) = x^2 - y^2$ را در نظر می گیریم که در جلسات قبل نیز با آن آشنایی پیدا کرده ایم. می خواهیم جهت حداکثر افزایش و کاهش تابع را در نقطه $(2, 0)$ پیدا کنیم. داریم $\nabla f(x, y) = (2x, -2y)$ ، پس $\nabla f(2, 0) = (4, 0)$. بردار واحد در جهت گرادیان بردار $(1, 0)$ است، که جهت حداکثر افزایش تابع می باشد. مشتق سویی در جهت $(1, 0)$ یعنی $\frac{\partial f}{\partial x}(2, 0)$ که برابر ۴ است. همچنین جهت حداکثر کاهش تابع جهت مقابل $(1, 0)$ ، یعنی جهت $(-1, 0)$ است که مشتق سویی نسبت به آن برابر (-4) می شود. روی نمودار تابع (شکل ۳) مشاهده می شود که جهت حداکثر افزایش تابع متناظر با صعود روی "یال" زین است و جهت حداکثر کاهش جهت مقابل آن. می توان (۷) را به شکل زیر نوشت:

$$-|\nabla f(a)| \leq D_v f(a) \leq |\nabla f(a)| \quad (8)$$

بین جهت حداکثر افزایش تابع و جهت حداکثر کاهش تابع، انتظار داریم جهتی موجود باشد که تابع کمترین تغییر را در آن جهت دارد، یعنی $D_v f(a) = 0$. در واقع وقتی $v \perp \nabla f(a)$ ، کسینوس زاویه بین

آنها صفر است و در نتیجه $D_v f(a) = 0$. نکته جالب این است که جهت‌های عمود بر گرادیان، در واقع جهت‌های مماس بر مجموعه تراز گذرا از نقطه a هستند:

(۲۳-۲) گزاره. اگر در نقطه a داشته باشیم $\nabla f(a) \neq 0$ ، آنگاه ابرصفحه عمود بر $\nabla f(a)$ ابرصفحه مماس بر مجموعه تراز تابع f است که از نقطه a می‌گذرد.

به طور خلاصه گفته می‌شود که گرادیان همواره بر مجموعه‌های تراز عمود است. البته وقتی گرادیان صفر شود، عمود بودن صرفاً به معنای صفر شدن حاصل ضرب داخلی گرادیان با بردارهای مماس بر مجموعه تراز (در صورت وجود) است. اثبات دقیق این گزاره را به بعد موکول می‌کنیم ولی کمی توضیح و استدلال شهودی در مورد آن بجاست. سؤال اولی که به ذهن می‌رسد این است که آیا مجموعه‌های تراز مجموعه‌هایی "هموار" هستند به این معنی که در هر نقطه دارای یک "ابرفضحه مماس" هستند؟ بعدها نشان داده خواهد شد که اگر در نقطه a داشته باشیم $\nabla f(a) \neq 0$ ، آنگاه دست کم بخش کوچکی از مجموعه تراز گذرا از a که در مجاورت نقطه a قرار دارد "هموار" است، ولی اگر $\nabla f(a) = 0$ ، آنگاه در نقطه a ممکن است این "هموار بودن" نقض شود یا بعد فضای مماس کوچکتر از $(n-1)$ شود (به مثال‌های بعد مراجعه کنید). نقاطی را که در آن $\nabla f(a) = 0$ نقاط بحرانی تابع f و نقاطی را که در آن $\nabla f(a) \neq 0$ ، نقاط عادی یا نقاط منظم می‌نامیم. با فرض این مقدمه، می‌توان دلایل شهودی زیر را برای عمود بودن گرادیان بر مجموعه تراز ارائه کرد:

استدلال شهودی ۱. تابع f روی مجموعه تراز ثابت است، بنابراین "مشتق در جهت مجموعه تراز"، یعنی مشتق نسبت به هر جهت مماس بر مجموعه تراز باید صفر باشد.

استدلال شهودی ۲. فرض کنید u یک بردار مماس بر مجموعه تراز در نقطه a باشد و به کسر

$$\frac{f(a + hu) - f(a)}{h} \quad (9)$$

توجه کنید که حد آن وقتی h به صفر میل کند برای $D_u f(a)$ است. از نقطه $a + hu$ عمودی بر مجموعه تراز رسم کنید. با توجه به تعریفی که از "مماس" ارائه کردیم انتظار داریم طول پاره خط عمود سریعتر از h به صفر میل کند وقتی $h \rightarrow 0$. ولی پای عمود روی مجموعه تراز است یعنی مقدار f در آن برابر مقدار f در a است، بنابراین نامعقول نیست که با اثر دادن تابع مشتق پذیر f ، انتظار داشته باشیم تفاضل $f(a+h) - f(a)$ نیز سریعتر از h به صفر میل کند، یعنی حد کسر (۹) صفر باشد.

با قبول این گزاره، تصویر زیر در مورد رفتار تابع مشتق پذیر $f: S \rightarrow \mathbb{R}$ ، $S \subset \mathbb{R}^n$ ، حاصل می شود:

گرادیان f همواره بر مجموعه های تراز f عمود است و به جهت حداکثر افزایش f اشاره می کند.

مثال ۵. تابع $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ را به صورت $f(x, y, z) = x^2 + y^2 - z^2$ تعریف می کنیم. داریم:

$$\nabla f(x, y, z) = (2x, 2y, -2z)$$

به ازای $c > 0$ ، مجموعه تراز منسوب به c ، یعنی $x^2 + y^2 - z^2 = c$ یک هذلولی وار یکپارچه است، به ازای $c = 0$ یک مخروط به دست می آید و به ازای $c < 0$ ، مجموعه تراز یک هذلولی وار دو پارچه است (شکل ۵). جهت گرادیان یعنی بردار $(2x, 2y, -2z)$ ، خارج از محور z ، هموار دور از محور z ، به طرف پایین وقتی $z > 0$ و به طرف بالا وقتی $z < 0$ می باشد.

روی محور z ، جهت گرادیان به طرف مبدأ است. تنها نقطه بحرانی $(0, 0, 0)$ است که در آن مجموعه تراز مربوط (مخروط) هموار نیست.

مثال ۶. تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت $f(x, y) = (x^2 - 1)^2 + y^2$ تعریف می کنیم. داریم:

$$\nabla f(x, y) = (4x(x^2 - 1), 2y)$$

سه نقطه بحرانی $(\pm 1, 0)$ و $(0, 0)$ وجود دارند. نخست مجموعه های تراز شامل این نقاط بحرانی را در نظر می گیریم. به ازای $(\pm 1, 0)$ داریم $f(\pm 1, 0) = 0$ و در واقع با قرار دادن $(x^2 - 1)^2 + y^2 = 0$

فقط این دو نقطه به دست می آیند. پس مجموعه تراز منسوب به صفر از دو نقطه $(1, 0)$ و $(-1, 0)$ تشکیل شده است. می بینیم که این مجموعه تراز حالتی استثنایی دارد، یعنی به جای حالت معمول، که مجموعه تراز یک منحنی است، در اینجا از دو نقطه مجزا تشکیل شده است. به ازای نقطه بحرانی دیگر، یعنی $(0, 0)$ ، $f(0, 0) = 1$. مجموعه تراز $1 = y^2 + (x^2 - 1)^2$ را می توان با نقطه یابی، استفاده از مختصات قطبی، یا ترسیم کامپیوتری به دست آورد که یک منحنی پروانه شکل است (شکل ۶). توجه کنید که این مجموعه تراز در نقطه بحرانی $(0, 0)$ هموار نیست، یعنی دو شاخه منحنی در آن نقطه متقاطع می شوند. برای $c < 0$ مجموعه تراز $c = y^2 + (x^2 - 1)^2$ تهی است. برای $0 < c < 1$ ، مجموعه تراز از دو منحنی بسته کوچک در داخل دو شاخه پروانه و حول دو نقطه بحرانی $(\pm 1, 0)$ تشکیل شده است، و بالاخره برای $c > 1$ ، مجموعه تراز یک منحنی است که از بیرون پروانه را احاطه می کند. میدان گرادیان روی مجموعه های تراز ترسیم شده است.

نمودار این تابع را می توان به کمک مجموعه های تراز بازسازی کرد به این ترتیب که مجموعه تراز منسوب به c را به ارتفاع $z = c$ منتقل می کنیم. صفحه $z = 0$ بر نمودار در دو نقطه $(\pm 1, 0, 0)$ مماس است. برای $c > 0$ کوچک، $0 < c < 1$ ، صفحه $z = c$ نمودار را در دو منحنی قطع می کند، به ازای $z = 1$ شکل پروانه ای پدید می آید و برای $z > 1$ مقطع یک منحنی بسته است (شکل ۷).

مثال ۷. تابعهای $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$ در مثال ۹، بخش ۱۷ را در نظر می گیریم. داریم

$$\nabla f(x, y) = (1, 1)$$

$$\nabla g(x, y) = 3(x + y)^2(1, 1)$$

$$\nabla h(x, y) = \frac{1}{3}(x + y)^{-\frac{2}{3}}(1, 1)$$

گرادیان هر سه تابع بر مجموعه های تراز که خطوط راست (ثابت $x + y$) هستند عمودند. توجه کنید که در مورد f میدان گرادیان ثابت است که شدت رشد یکنواخت تابع f را نشان می دهد. در مورد g ، با دور شدن از 0 ، طول بردار گرادیان ∇g سریعتر از ∇f رشد می کند که نشانگر افزایش سریعتر $(x + y)^3$ در

مقایسه با $x + y$ است. ضمناً مجموعه‌های تراز g نیز نسبت به هم نزدیکتر می‌شوند که حاکی از همین امر است. بالعکس در مورد $\nabla h, h$ با دور شدن از $\underline{0}$ کوچکتر می‌شود و مجموعه‌های تراز نیز از هم فاصله می‌گیرند که نمایانگر رشد کندتر h نسبت به تابع خطی $f(x, y) = x + y$ است.

قاعده زنجیره‌ای (۲)

در این جلسه مثال‌هایی از نحوه استفاده از قاعده زنجیری ارائه خواهیم کرد.

مثال ۱. فرض کنید دمای ناحیه‌ای در \mathbb{R}^3 در همسایگی نقطه $(0, 0, 0)$ از فرمول $T = \frac{100}{\sqrt{1+x^2+y^2+z^2}}$ تبعیت می‌کند. مکان متحرکی مجهز به یک دماسنج در زمان t به وسیله $(x, y, z) = (1+t, t^2, t-t^3)$ داده شده است. آهنگ تغییر دمای نمایش داده شده توسط دماسنج را در لحظه $t = 1$ به دست آورید. در اینجا T به عنوان تابعی از (x, y, z) ارائه شده است و (x, y, z) به عنوان تابعی از زمان، t :

$$t \rightarrow (x, y, z) \rightarrow T$$

بنابراین قاعده زنجیری به صورت زیر در می‌آید:

$$\left[\frac{dT}{dt} \right] = \left[\frac{\partial T}{\partial x} \quad \frac{\partial T}{\partial y} \quad \frac{\partial T}{\partial z} \right] \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{bmatrix}$$

$$\frac{dT}{dt} = \frac{\partial T}{\partial x} \frac{dx}{dt} + \frac{\partial T}{\partial y} \frac{dy}{dt} + \frac{\partial T}{\partial z} \frac{dz}{dt}$$

در $t = 1$ داریم $(x, y, z) = (2, 1, 0)$:

$$\begin{aligned} \left[\frac{\partial T}{\partial x} \quad \frac{\partial T}{\partial y} \quad \frac{\partial T}{\partial z} \right]_{(2,1,0)} &= \left[\frac{-200 \cdot x}{(1+x^2+y^2+z^2)^2}, \frac{-200 \cdot y}{(1+x^2+y^2+z^2)^2}, \frac{-200 \cdot z}{(1+x^2+y^2+z^2)^2} \right]_{(2,1,0)} \\ &= \left[-\frac{400}{36} \quad -\frac{200}{36} \quad 0 \right] \\ &= \left[-\frac{100}{9} \quad -\frac{50}{9} \quad 0 \right] \end{aligned}$$

از طرفی دیگر

$$\left. \left(\frac{dx}{dt} \quad \frac{dy}{dt} \quad \frac{dz}{dt} \right) \right|_{t=1} = \left. (1 \quad 2t \quad 1 - 3t^2) \right|_{t=1}$$

$$= (1 \quad 2 \quad -2)$$

بنابراین

$$\frac{dT}{dt} = -\frac{1}{9} - \frac{1}{9} + 0$$

$$\simeq -22/22$$

مثال ۲. تابعی $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ به صورت $z = x^2 - y^2$ داده شده است. اگر (r, θ) مختصات قطبی در صفحه xy باشند، $\frac{\partial z}{\partial \theta}$ و $\frac{\partial z}{\partial r}$ را محاسبه کنید.

در اینجا باید $z = x^2 - y^2$ را به عنوان تابعی از (r, θ) در نظر بگیریم و سپس $\frac{\partial z}{\partial \theta}$ و $\frac{\partial z}{\partial r}$ را محاسبه کنیم. در این مثال ساده می توان به راحتی z را بر حسب (r, θ) نوشت و محاسبه مشتق های پاره ای را انجام داد ولی اگر عبارت تعریف کننده تابع پیچیده تر باشد اغلب استفاده از قاعده زنجیری کوتاه تر است. برای روشن شدن روش از قاعده زنجیری استفاده می کنیم. داریم:

$$(r, \theta) \rightarrow (x, y) \rightarrow z$$

بنابراین قاعده زنجیری به صورت زیر در می آید:

$$\begin{bmatrix} \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}$$

هر یک از دو ماتریس طرف راست را محاسبه می کنیم:

$$\begin{bmatrix} \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{bmatrix} = [2x \quad -2y]$$

با توجه به $y = r \sin \theta$ ، $x = r \cos \theta$ داریم:

$$\begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & r \cos \theta \\ \sin \theta & -r \sin \theta \end{bmatrix}$$

بنابراین

$$\begin{aligned}
 \frac{\partial z}{\partial r} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial r} \\
 &= (\sqrt{r} \cos \theta)(\cos \theta) + (-\sqrt{r} \sin \theta)(\sin \theta) \\
 &= \sqrt{r} \cos 2\theta \\
 \frac{\partial z}{\partial \theta} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial \theta} \\
 &= (\sqrt{r} \cos \theta)(r \cos \theta) + (-\sqrt{r} \sin \theta)(-r \sin \theta) \\
 &= \sqrt{r} r^2
 \end{aligned}$$

مثال ۳. در بحث مربوط به گرادیان اشاره کردیم به این مطلب که اگر $f: S \rightarrow \mathbb{R}$ ، $S \subset \mathbb{R}^n$ ، تابعی مشتق پذیر باشد و $\nabla f(a) \neq \underline{0}$ ، آنگاه $\nabla f(a)$ بر مجموعه تراز تابع f که از نقطه a می گذرد در آن نقطه عمود است. عمود بودن بردار $\nabla f(a)$ بر مجموعه تراز را بدین معنی گرفتیم که $\nabla f(a)$ بر هر بردار مماس بر مجموعه تراز در نقطه a عمود است و دلایلی شهودی برای این واقعیت ارائه کردیم. اکنون اثبات دقیق تری از این مطلب به کمک قاعده زنجیری ارائه می کنیم. فرض کنید v برداری مماس بر مجموعه تراز مورد نظر در نقطه a باشد. در این صورت یک منحنی پارامتری مشتق پذیر $\alpha: I \rightarrow \mathbb{R}^n$ یک بازه در \mathbb{R} ، وجود دارد که تصویر آن به تمامی روی مجموعه تراز قرار دارد، در زمانی $t_0 \in I$ از نقطه a می گذرد، $\alpha(t_0) = a$ ، و بردار سرعت آن برابر v است، $\alpha'(t_0) = v$ (این مطلب نیاز به اثبات دارد که در جلسات بعد ارائه خواهد شد).

طبق تعریف مجموعه تراز، f روی مجموعه تراز گذرا از a مقداری ثابت دارد، پس $f(\alpha(t)) = c$ که $c \in \mathbb{R}$ مقداری ثابت است، پس $\frac{d}{dt}(f(\alpha(t))) = 0$ ، و طبق قاعده زنجیره ای:

$$(D(f \circ \alpha))(t_0) = Df(\alpha(t_0)) \circ D\alpha(t_0)$$

یا با نوشتن $\alpha(t) = (x_1, \dots, x_n)$

$$\frac{d}{dt}((f \circ \alpha)(t))|_{t=t_0} = \left[\frac{\partial f}{\partial x_1}(a) \quad \dots \quad \frac{\partial f}{\partial x_n}(a) \right] \begin{bmatrix} \frac{dx_1}{dt}(t_0) \\ \vdots \\ \frac{dx_n}{dt}(t_0) \end{bmatrix} = \nabla f(a) \cdot v$$

پس $\nabla f(a) \cdot v = 0$ و حکم به اثبات می‌رسد.

مثال ۴ (اصل بقای انرژی برای دستگاه‌های پایسته). یک دستگاه مکانیکی نیوتنی در نظر می‌گیریم با مختصات مکانی (x_1, \dots, x_n) ، مختصات سرعتی (v_1, \dots, v_n) ، و جرم‌های m_1, \dots, m_n . مثلاً می‌توان k ذره در \mathbb{R}^3 را در نظر گرفت که تحت اثر متقابل گرانشی در حرکتند. در این صورت $n = 3k$ ، مختصات مکانی ذره اول، (x_1, x_2, x_3) ، سرعت ذره اول، (v_1, v_2, v_3) ، و جرم ذره اول $m_1 = m_2 = m_3$ است.

برای ذره دوم، (x_4, x_5, x_6) ، (v_4, v_5, v_6) ، و $m_4 = m_5 = m_6$ در نظر گرفته می‌شوند و غیره. این دستگاه را پایسته می‌نامیم در صورتی که تابعی مشتق‌پذیر U از (x_1, \dots, x_n) وجود داشته باشد (موسوم به پتانسیل) که:

$$\vec{F} = -\nabla U$$

که در اینجا \vec{F} بردار نیرو است و در فرمول نیوتن $\vec{F} = m\vec{a}$ صدق می‌کند. مقصود از $\vec{a} = (a_1, \dots, a_n)$ بردار شتاب است و $m\vec{a}$ یک خلاصه نویسی برای $(m_1 a_1, \dots, m_n a_n)$ می‌باشد. انرژی جنبشی دستگاه را به صورت $\frac{1}{2} \sum_{i=1}^n m_i v_i^2$ تعریف می‌کنیم و $E = U + K$ را انرژی دستگاه می‌نامیم. E تابعی از $(x_1, \dots, x_n, v_1, \dots, v_n)$ است. در سیر تحول دستگاه، این $(2n)$ تایی به عنوان تابعی از زمان، t ، تغییر می‌کنند. می‌خواهیم ثابت کنیم

$$\frac{dE}{dt} = 0$$

یعنی E در طی تحول دستگاه ثابت می ماند. در واقع باید ثابت کنیم که مشتق ترکیب دو تابع زیر صفر است:

$$t \longrightarrow (x_1, \dots, x_n, v_1, \dots, v_n) \longrightarrow E$$

طبق قاعده زنجیره ای داریم:

$$\begin{aligned} \frac{dE}{dt} &= \sum_{i=1}^n \frac{\partial E}{\partial x_i} \frac{dx_i}{dt} + \sum_{i=1}^n \frac{\partial E}{\partial v_i} \frac{dv_i}{dt} \\ &= \sum_{i=1}^n \left(-\frac{\partial U}{\partial x_i}\right) v_i + \sum_{i=1}^n (m_i v_i) (a_i) \\ &= \sum_{i=1}^n (-m_i a_i) v_i + \sum_{i=1}^n (m_i v_i) (a_i) \\ &= 0 \end{aligned}$$

مثال ۵ (فرمول اویلر). S را زیرمجموعه \mathbb{R}^n متشکل از (x_1, \dots, x_n) هایی می گیریم که همه x_i ها مثبت هستند. تابع $f: S \rightarrow \mathbb{R}$ را همگون از درجه d می نامیم در صورتی که به ازای هر $r > 0$ داشته باشیم:

$$f(rx_1, \dots, rx_n) = r^d f(x_1, \dots, x_n) \quad (1)$$

رابطه زیر به فرمول اویلر معروف است:

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = df(x_1, \dots, x_n) \quad (2)$$

برای اثبات (۱)، از دو طرف (۱) نسبت به r مشتق می گیریم. طبق قاعده زنجیره ای داریم:

$$\begin{aligned} \frac{\partial}{\partial r} [f(rx_1, \dots, rx_n)] &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} (rx_1, \dots, rx_n) \cdot \frac{\partial (rx_i)}{\partial r} \\ &= \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} (rx_1, \dots, rx_n) \\ \frac{\partial}{\partial r} [r^d f(x_1, \dots, x_n)] &= dr^{d-1} f(x_1, \dots, x_n) \end{aligned}$$

با برابر قرار دادن دو طرف و جایگزینی $r = 1$ فرمول اوپلر به دست می آید.

مشتق‌های پاره‌ای مرتبه بالا و چند جمله‌ای تیلور

فرض کنید $a, S \subset \mathbb{R}^n$ یک نقطه درونی S است و برای $f : S \rightarrow \mathbb{R}$ ، $y = f(x)$ ، مشتق پاره‌ای $\frac{\partial f}{\partial x_i}$ در همه نقاط یک گوی باز حول a تعریف شده است. در این صورت $\frac{\partial f}{\partial x_i}$ خود یک تابع با مقدار حقیقی است که نقطه a یک نقطه درونی دامنه آن است. بدین ترتیب می‌توان برای تابع $\frac{\partial f}{\partial x_i}$ وجود مشتق پاره‌ای نسبت به متغیر x_j را در نقطه a مطرح کرد. $(\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i}))(a)$ را در صورت وجود به نمادهای زیر نمایش می‌دهیم

$$\frac{\partial^2 y}{\partial x_j \partial x_i}(a) \quad , \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \quad , \quad D_{ij}f(a) \quad , \quad f_{x_i x_j}(a) \quad (1)$$

و آن را یک مشتق پاره‌ای مرتبه دوم f در نقطه a می‌نامیم. برای تابع n متغیری $y = f(x_1, \dots, x_n)$ می‌توان n^2 مشتق پاره‌ای مرتبه دوم در نظر گرفت زیرا که در (۱)، i, j هر یک عدد صحیحی بین ۱ و n هستند. معمولاً به جای $\frac{\partial^2 y}{\partial x_j \partial x_i}$ می‌نویسیم $\frac{\partial^2 y}{\partial x_i \partial x_j}$. به همین ترتیب، اگر مشتق پاره‌ای مرتبه دوم $\frac{\partial^2 y}{\partial x_j \partial x_i}$ در یک گوی باز حول a تعریف شده باشد، می‌توان وجود مشتق پاره‌ای مرتبه سوم $\frac{\partial^3 y}{\partial x_k \partial x_j \partial x_i} = \frac{\partial}{\partial x_k}(\frac{\partial^2 y}{\partial x_j \partial x_i})$ را بررسی نمود، و همین طور برای مشتق‌های پاره‌ای مرتبه‌های بالاتر.

نکته مهم در مورد مشتق‌های پاره‌ای مرتبه دو به بالا این است که تحت شرایط سهل الوصول، ترتیب مشتقگیری نسبت به متغیرهای مختلف اثری بر نتیجه نهایی ندارد و تنها عامل اثرگذار تعداد دفعاتی است که مشتقگیری نسبت به متغیر خاص صورت گرفته است. قضیه زیر نمونه‌ای از این نوع احکام است:

(۲۶-۱) قضیه. فرض کنید S زیرمجموعه‌ای از \mathbb{R}^n باشد، a یک نقطه درونی S و $f : S \rightarrow \mathbb{R}$ یک

تابع به طوری که $\frac{\partial^2 y}{\partial x_j \partial x_i}$ و $\frac{\partial^2 f}{\partial x_i \partial x_j}$ هر دو در یک گوی باز حول a تعریف شده و در نقطه a پیوسته

باشند. در این صورت:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \quad (2)$$

بدیهی است که می توان با به کارگیری مکرر این قضیه، حکم را به مشتق های پاره ای مرتبه های بالاتر تعمیم داد.

اثبات قضیه. برای سهولت در نوشتن، حالت $n = 2$ را در نظر می گیریم. اثبات حالت کلی عیناً به همین صورت است، کافی است استدلال را به صفحه (x_i, x_j) محدود کنیم. بدین ترتیب نقاط S را به (x, y) و نقطه درونی مورد نظر را به (a, b) نمایش می دهیم. اگر $|h|$ و $|k|$ به اندازه کافی کوچک باشند، مستطیل $b \leq y \leq b+k, a \leq x \leq a+h$ (در صورتی که هر یک از h, k منفی باشد، جهت نامساوی مربوط را معکوس کنید) در درون دامنه تعریف $\frac{\partial^2 f}{\partial x_i \partial x_j}$ و $\frac{\partial^2 f}{\partial x_j \partial x_i}$ قرار می گیرد. از سه تابع کمکی u, Q و v به شرح زیر استفاده می کنیم:

$$Q(h, k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) \quad (3)$$

Q تفاضل مجموع مقدار f در دو رأس مقابل مستطیل و دو رأس مقابل دیگر است.

$$u(x) = f(x, b+k) - f(x, b) \quad (\text{برای هر } k \text{ ثابت}) \quad (4)$$

$$v(y) = f(a+h, y) - f(a, y) \quad (\text{برای هر } h \text{ ثابت}) \quad (5)$$

بدین ترتیب u و v ، به ترتیب، تفاضل f در دو انتهای پاره خط های قائم و افقی را نمایش می دهند. داریم

$$Q(h, k) = u(a+h) - u(a) \quad (6)$$

و نیز

$$Q(h, k) = v(b+k) - v(b) \quad (7)$$

طبق قضیه مقدار میانگین یک متغیری، برای u داریم:

$$u(a+h) - u(a) = h \cdot u'(\xi) \quad (8)$$

که در آن ξ بین a و $a+h$ است. ضمناً مشتق u طبق (۴) از مشتق‌گیری f نسبت به x به دست می‌آید، پس از (۶) و (۸):

$$Q(h, k) = h \left[\frac{\partial f}{\partial x}(\xi, b+k) - \frac{\partial f}{\partial x}(\xi, b) \right] \quad (9)$$

حال اگر تابع $\frac{\partial f}{\partial x}$ را به عنوان تابعی از متغیر دوم آن در نظر بگیریم، طبق قضیه میانگین یک متغیری داریم:

$$\frac{\partial f}{\partial x}(\xi, b+k) - \frac{\partial f}{\partial x}(\xi, b) = k \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \quad (10)$$

که در اینجا η بین b و $b+k$ است. پس

$$Q(h, k) = kh \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \quad (11)$$

حال به همین ترتیب، اگر به جای u ، از v استفاده کنیم، نقطه η' بین b و $b+k$ ، نقطه ξ' بین a و $a+h$ به دست می‌آیند که:

$$Q(h, k) = kh \frac{\partial^2 f}{\partial x \partial y}(\xi', \eta') \quad (12)$$

از مقایسه (۱۱) و (۱۲) نتیجه می‌شود که

$$\frac{\partial^2 f}{\partial x \partial y}(\xi', \eta') = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \quad (13)$$

توجه کنید که (ξ, η) و (ξ', η') دو نقطه در مستطیل شکل ۱ هستند، پس وقتی $(h, k) \rightarrow (0, 0)$ هر دو نقطه به (a, b) میل می‌کنند. از پیوستگی $\frac{\partial^2 f}{\partial y \partial x}$ و $\frac{\partial^2 f}{\partial x \partial y}$ در (a, b) نتیجه می‌شود که:

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

چنان که حکم بود.

مثال ۱. برای $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ، $f(x, y) = xe^y - y^2x^3$ ، مشتق‌های پاره‌ای مرتبه دوم را محاسبه کنید.

$$\frac{\partial f}{\partial x} = e^y - 3x^2y^2, \quad \frac{\partial f}{\partial y} = xe^y - 2yx^3$$

$$\frac{\partial^2 f}{\partial x^2} = -6xy^2, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = e^y - 6x^2y, \quad \frac{\partial^2 f}{\partial y^2} = xe^y - 2x^3$$

مثال ۲. کلیه مشتق‌های پاره‌ای از همه مراتب $f(x, y) = ye^x$ را محاسبه کنید.

از عبارات بالا مشخص است که همه مشتق‌های پاره‌ای وجود دارند و پیوسته خواهند بود، بنابراین ترتیب مشتق‌گیری اهمیتی ندارد. اگر از عبارت فوق یک بار نسبت به y مشتق‌گیری شود، متغیر y حذف می‌شود، پس می‌توان نوشت

$$\frac{\partial^{k+l} f}{\partial x^k \partial y^l} = 0 \quad \text{اگر } l \geq 2$$

همچنین نتایج زیر نیز به دست می‌آیند:

$$\frac{\partial^k f}{\partial x^k} = ye^x, \quad \frac{\partial^{k+1} f}{\partial x^k \partial y} = e^x$$

یکی از کاربردهای مهم مشتق‌های مرتبه بالا ارائه تقریب‌های مرتبه بالا برای تابع‌ها به صورت چندجمله‌ای تیلور است. در همین رابطه آزمون مشتق دوم برای تعیین نوع نقاط بحرانی (یعنی نقاطی که در آن مشتق صفر است) و تعیین جهت تحدب نمودار یک تابع را نیز باید نام برد. مشابه همین ملاحظات را می‌توان تابع‌های چندمتغیری نیز مطرح نمود. برای این منظور نخست حالت یک متغیری چندجمله‌ای‌های تیلور و خواص آنها را یادآوری می‌کنیم. فرض کنید I یک بازه در \mathbb{R} ، a یک نقطه درونی I ، و $f: I \rightarrow \mathbb{R}$ تابعی باشد که در نقطه a مشتق‌پذیر است. تقریب خطی f در نقطه a ، چندجمله‌ای درجه ۱ زیر است:

$$f(a) + f'(a) \cdot (x - a) \quad (14)$$

این چندجمله‌ای را به $T_1(x)$ نمایش می‌دهیم. هر یک از دو ویژگی زیر $T_1(x)$ را از میان چندجمله‌ای‌های درجه یک به طور منحصر به فرد متمایز می‌کنند:

الف) $T_1(x)$ یگانه چندجمله‌ای درجه یک است که $T_1(a) = f(a)$ و $T'_1(a) = f'(a)$.

ب) $T_1(x)$ یگانه چندجمله‌ای درجه یک است که $T_1(a) = f(a)$ و $\lim_{x \rightarrow a} \frac{f(x) - T_1(x)}{|x-a|} = 0$.

قبلاً دیده‌ایم که (ب) را می‌توان به این صورت تعبیر کرد که $T_1(x)$ "نزدیک‌ترین تقریب درجه یک" به f حول $x = a$ است، و چندجمله‌ای‌های تیلور این وضعیت را به تقریب‌های درجه بالاتر تعمیم می‌دهند. اگر f در نقطه a دارای مشتق تا مرتبه p (با شمول درجه p) باشد، چندجمله‌ای تیلور درجه p تابع f حول a به صورت زیر تعریف می‌شود:

$$T_p(x) = f(a) + f'(a) \cdot (x - a) + \dots + \frac{1}{p!} f^{(p)}(a) \cdot (x - a)^p$$

به طور کلی می‌توان انتظار داشت که با بالا بردن درجه، تقریب‌های بهتری از تابع f به دست می‌آید زیرا که برای $q < p$ ، می‌توان هر چندجمله‌ای درجه q را یک چندجمله‌ای درجه p (با ضرایب صفر پس از درجه q) فرض کرد و بدین ترتیب "بهترین تقریب درجه q " نمی‌تواند اکیداً بهتر از "بهترین تقریب درجه p " باشد. در واقع مشابه خواص (الف) و (ب) برای $T_p(x)$ نیز برقرار است:

الف) $T_p(x)$ یگانه چندجمله‌ای درجه p است که $T_p(a) = f(a)$ ، $T'_p(a) = f'(a)$ ، ...
 $T_p^{(p)}(a) = f^{(p)}(a)$

ب) $T_p(x)$ یگانه چندجمله‌ای درجه p است که $T_p(a) = f(a)$ و $\lim_{x \rightarrow a} \frac{f(x) - T_p(x)}{|x-a|^p} = 0$.

بدین ترتیب مشتقات f تا مرتبه p در نقطه a با مشتقات متناظر f برابرند و این امر، طبق (ب)، تقریبی آن چنان نزدیک از f حول a می‌دهد که تفاضل $f(x) - T_p(x)$ سریعتر از $|x-a|^p$ به صفر میل می‌کند وقتی $x \rightarrow a$. در مورد دقت این تقریب، راه‌های گوناگونی برای تخمین خطا وجود دارد. یکی از معمولترین روش‌ها، استفاده از "باقیمانده لاگرانژ" است. طبق این روش، اگر f در بازه بین a و x دارای مشتق مرتبه $(p+1)$ باشد، آنگاه نقطه‌ای ξ بین a و x وجود دارد که:

$$f(x) - T_p(x) = \frac{1}{(p+1)!} f^{(p+1)}(\xi) \cdot (x-a)^{p+1} \quad (15)$$

بدین ترتیب، مثلاً دقت تقریب خطی به قدرمطلق مشتق دوم تابع وابسته است. چنانچه مشتق مرتبه دوم دارای قدرمطلق کوچک باشد، میزان تحدب یا تقعر تابع (به انحناء به انحراف از خط راست) کوچک است و نمودار تابع نزدیک به تقریب خطی می ماند، ولی اگر قدرمطلق مشتق دوم بزرگ باشد، نمودار تابع به سرعت از تقریب خطی دور می شود.

هدف ما در باقیمانده این بخش این است که ملاحظات بالا را به تابع های چند متغیری تعمیم دهیم. فرض کنید S زیرمجموعه ای از \mathbb{R}^n ، a یک نقطه درونی S و $f: S \rightarrow \mathbb{R}$ یک تابع باشد. چون a یک نقطه درونی S است، اگر نقطه x در S به اندازه کافی نزدیک به a باشد، پاره خط واصل بین a و x به تمامی در S قرار می گیرد. این پاره خط را به صورت $a + t(x - a)$ ، $0 \leq t \leq 1$ ، پرمایش می کنیم. گاهی به جای $x - a$ می نویسیم h پس $h = (h_1, \dots, h_n) = (x_1 - a_1, \dots, x_n - a_n)$. مقدار تابع f روی پاره خطی بالا را به $\varphi(t)$ نمایش می دهیم، پس:

$$\begin{cases} \varphi(t) = f(a + th) \\ \varphi(0) = f(a), \varphi(1) = f(x) \end{cases} \quad (16)$$

اگر f در همه نقاط پاره خط $a + th$ مشتق پذیر باشد، طبق قاعده زنجیره ای داریم:

$$\begin{aligned} \varphi'(t) &= Df(a + th) \cdot h \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + th) \cdot (x_i - a_i) \end{aligned} \quad (17)$$

پس $\varphi'(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot (x_i - a_i)$ و چند جمله ای تیلور φ به ازای $t = 1$ برابر می شود با:

$$\varphi(0) + \varphi'(0) \cdot (1 - 0) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot (x_i - a_i) \quad (18)$$

عبارت سمت راست همان چند جمله ای درجه یک نسبت به x_1, \dots, x_n است که به عنوان تقریب درجه یک تابع f در نقطه a می شناسیم. به همین ترتیب، فرض می کنیم هر یک از مشتقات پاره ای $\frac{\partial f}{\partial x_i}$ خود تابعی مشتق پذیر در طول پاره خط $a + th$ است (مثلاً با این فرض که مشتقات پاره ای مرتبه دوم f

پیوسته‌اند) و با مشتق‌گیری مجدد از (۵) داریم:

$$\begin{aligned}\varphi''(t) &= \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i} (a+th) \cdot h_j \right) (x_i - a_i) \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i} (a+th) \cdot (x_j - a_j) (x_i - a_i)\end{aligned}\quad (19)$$

پس چندجمله‌ای تیلور درجه دوم φ به‌ازای $t = 1$ برابر می‌شود با:

$$\begin{aligned}\varphi(1) &= f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} (a) \cdot (x_i - a_i) + \frac{1}{2!} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} (a) \cdot (x_i - a_i) (x_j - a_j)\end{aligned}\quad (20)$$

توجه کنید که این عبارت یک چندجمله‌ای درجه دوم نسبت به x_1, \dots, x_n است. طرف راست (۷) را چندجمله‌ای تیلور درجه دوم تابع f حول a می‌نامیم. به‌طور کلی، اگر f دارای مشتقات پاره‌ای پیوسته از مرتبه p -ام باشد، با مشتق‌گیری مکرر می‌توانیم چندجمله‌ای تیلور درجه p تابع یک متغیری φ را تشکیل دهیم. مقدار این چندجمله‌ای به‌ازای $t = 1$ ، چندجمله‌ای تیلور درجه p تابع f حول a خوانده می‌شود. جزء ثابت این چندجمله‌ای، $f(a)$ ؛ جزء درجه یک (نسبت به $x_1 - a_1, \dots, x_n - a_n$) برابر $\sum_{i=1}^n \frac{\partial f}{\partial x_i} (a) (x_i - a_i)$ ؛ جزء درجه دو برابر $\frac{1}{2!} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} (a) \cdot (x_i - a_i) (x_j - a_j)$ و ... تعریف می‌شوند. توجه کنید که جزء درجه k از مجموع جملات به شکل زیر تشکیل شده است:

$$\left(\frac{1}{k!} \right) \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} (a) \cdot (x_{i_1} - a_{i_1}) \dots (x_{i_k} - a_{i_k}) \quad (21)$$

(تکرار اندیس در بین i_j ها ممکن است) که در آن هر i_j همه مقادیر ممکن بین ۱ تا n را اتخاذ می‌کند. اگر چندجمله‌ای تیلور درجه p تابع f را به $T_p(x_1, \dots, x_n)$ نمایش دهیم، از (۱) نتیجه می‌شود که برای $k = p$:

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} (a) = \frac{\partial^k T_p}{\partial x_{i_1} \dots \partial x_{i_k}} (a) \quad (22)$$

(چرا؟)، یعنی ویژگی مشابه (الف) حالت یک متغیری همچنان برقرار است.

قبل از ادامه بحث، برای مانوس شدن بیشتر با شکل چندجمله‌ای تیلور، چندجمله‌ای تیلوریک تابع دو متغیری را به طور صریح محاسبه می‌کنیم. به جای (x_1, x_2) می‌نویسیم (x, y) و به جای (a_1, a_2) می‌نویسیم (a, b) . اجزاء زیر به دست می‌آیند:

$$\begin{aligned} f(a, b) & \text{ جزء درجه صفر:} \\ \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b) & \text{ جزء درجه ۱:} \\ \frac{1}{2!} \left[\frac{\partial^2 f}{\partial x^2}(a, b) \cdot (x - a)^2 + \frac{\partial^2 f}{\partial y \partial x}(a, b) \cdot (y - b)(x - a) + \frac{\partial^2 f}{\partial x \partial y}(a, b) \cdot (x - a)(y - b) + \frac{\partial^2 f}{\partial y^2}(a, b) \cdot (y - b)^2 \right] & \text{ جزء درجه ۲:} \end{aligned}$$

اگر مشتقات پاره‌ای مرتبه دوم f پیوسته باشند، نتیجه می‌شود که $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$ ، پس جزء درجه دوم به شکل زیر خلاصه می‌شود:

$$\frac{1}{2!} \left[\frac{\partial^2 f}{\partial x^2}(a, b) \cdot (x - a)^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(a, b) \cdot (x - a)(y - b) + \frac{\partial^2 f}{\partial y^2}(a, b) \cdot (y - b)^2 \right]$$

به همین ترتیب اگر مشتقات پاره‌ای مرتبه سوم f پیوسته باشند، به جای ۸ نوع جمله ممکن درجه ۳، فقط چهارنوع به صورت زیر ظاهر می‌شوند:

$$\frac{1}{3!} \left[\frac{\partial^3 f}{\partial x^3}(a, b) \cdot (x - a)^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y}(a, b) \cdot (x - a)^2 (y - b) + 3 \frac{\partial^3 f}{\partial x \partial y^2}(a, b) \cdot (x - a)(y - b)^2 + \frac{\partial^3 f}{\partial y^3}(a, b) \cdot (y - b)^3 \right]$$

ضرایب اجزاء درجه ۲ و درجه ۳ که به ترتیب $(1, 2, 1)$ و $(1, 3, 3, 1)$ هستند دقیقاً ضرایب بسط دو جمله‌ای $(A + B)^2$ و $(A + B)^3$ می‌باشند. در واقع با مشتق‌گیری مکرر و استقرای به سادگی دیده می‌شود که جزء درجه k نیز به همین شکل عمومی است، یعنی:

$$\frac{1}{k!} \left[\sum_{i=0}^k \binom{k}{i} \frac{\partial^k f}{\partial x^i \partial y^{k-i}}(a, b) \cdot (x - a)^i (y - b)^{k-i} \right] \quad (23)$$

که در اینجا $\binom{k}{i} = \frac{k!}{i!(k-i)!}$. برای به خاطر سپردن (10) بیان نمادین زیر می‌تواند مفید باشد:

$$\frac{1}{k!} \left[(x - a) \frac{\partial}{\partial x} + (y - b) \frac{\partial}{\partial y} \right]^k \cdot f(a, b) \quad (24)$$

در حالت n متغیری نیز، ضرایب برابر ضرایب بسط توان k یک جمله‌ای، یعنی ضرایب

$(A_1 + \dots + A_n)^k$ می شوند و می توان جزء درجه k را به شکل نمادین زیر نمایش داد:

$$\frac{1}{k!} [(x_1 - a_1) \frac{\partial}{\partial x_1} + \dots + (x_n - a_n) \frac{\partial}{\partial x_n}]^k \cdot f(a) \quad (25)$$

گاهی $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ را به "بردار نمادین" ∂ و $(x_1 - a_1, \dots, x_n - a_n)$ را به Δ نمایش داده و عبارت داخل کروشه را به عنوان ضرب داخلی نمادین این دو "بردار" در نظر می گیرند که در این صورت (۱۲) به شکل زیر در می آید:

$$\frac{1}{k!} (\Delta \cdot \partial)^k f(a) \quad (26)$$

بدین ترتیب چندجمله ای تیلور درجه p تابع f را می توان به شکل زیر نمایش داد:

$$f(a) + \frac{1}{1!} (\Delta \cdot \partial) f(a) + \frac{1}{2!} (\Delta \cdot \partial)^2 f(a) + \dots + \frac{1}{p!} (\Delta \cdot \partial)^p f(a) \quad (27)$$

شکل فوق مستقل از تعداد متغیرهاست. قبل از پرداختن به خواص و کاربردهای چندجمله ای تیلور، یکی دو مثال از محاسبه چندجمله ای تیلور را مطرح می کنیم.

مثال ۱ چندجمله ای تیلور درجه ۲ تابع $f(x, y) = \sqrt[3]{x^2 - y}$ را حول $(3, 1)$ محاسبه می کنیم. تابع $\sqrt[3]{\quad}$ جز در صفر از همه مراتب مشتق دارد و عبارت $x^2 - y$ نسبت به (x, y) از هر مرتبه مشتق پذیر است، پس در نقطه $(3, 1)$ که زیر $\sqrt[3]{\quad}$ صفر نمی شود می توان چندجمله ای تیلور از هر مرتبه را در نظر گرفت. در اینجا:

$$\frac{\partial f}{\partial x} = \frac{2x}{3} (x^2 - y)^{-\frac{2}{3}} \quad , \quad \frac{\partial f}{\partial y} = -\frac{1}{3} (x^2 - y)^{-\frac{2}{3}}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{2}{3} (x^2 - y)^{-\frac{2}{3}} - \frac{4x}{9} (x^2 - y)^{-\frac{5}{3}} \quad , \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{2x}{9} (x^2 - y)^{-\frac{5}{3}} \quad , \quad \frac{\partial^2 f}{\partial y^2} = -\frac{2}{9} (x^2 - y)^{-\frac{5}{3}}$$

پس چندجمله‌ای تیلور درجه ۲ عبارت است از:

$$\begin{aligned} f(3, 1) + \frac{\partial f}{\partial x}(3, 1) \cdot (x - 3) + \frac{\partial f}{\partial y}(3, 1) \cdot (y - 1) + \frac{1}{2!} \left[\frac{\partial^2 f}{\partial x^2}(3, 1)(x - 3)^2 \right. \\ \left. + 2 \frac{\partial^2 f}{\partial x \partial y}(3, 1)(x - 3)(y - 1) + \frac{\partial^2 f}{\partial y^2}(3, 1)(y - 1)^2 \right] \\ = 2 + \frac{1}{4}(x - 3) - \frac{1}{12}(y - 1) + \frac{1}{2!} \left[\left(-\frac{1}{12}\right)(x - 3)^2 + \frac{1}{12}(x - 3)(y - 1) - \frac{1}{144}(y - 1)^2 \right] \\ = 2 + \frac{1}{4}(x - 3) - \frac{1}{12}(y - 1) - \frac{1}{48}(x - 3)^2 + \frac{1}{48}(x - 3)(y - 1) - \frac{1}{288}(y - 1)^2 \end{aligned}$$

مثال ۲ چندجمله‌ای تیلور تابع $f(x, y, z) = x^2 y^3 - xz$ از همه درجات را حول $(1, 1, 0)$ محاسبه می‌کنیم. نتیجه در هر حالت باید یک چندجمله‌ای برحسب $(x - 1)$ ، $(y - 1)$ و $(z - 0) = z$ باشد. در عبارت داده شده برای f ، به جای x ، $(x - 1) + 1$ ، و به جای y ، $(y - 1) + 1$ قرار می‌دهیم:

$$f(x, y, z) = ((x - 1) + 1)^2 ((y - 1) + 1)^3 - ((x - 1) + 1)z$$

و جملات را به ترتیب توان برحسب $(x - 1)$ ، $(y - 1)$ و z مرتب می‌کنیم:

۱	جزء ثابت:
$2(x - 1) + 2(y - 1) - z$	جزء درجه ۱:
$(x - 1)^2 + 3(y - 1)^2 + 6(x - 1)(y - 1) - (x - 1)z$	جزء درجه ۲:
$(y - 1)^3 + 3(x - 1)^2(y - 1) + 6(x - 1)(y - 1)^2$	جزء درجه ۳:
$3(x - 1)^2(y - 1)^2 + 2(x - 1)(y - 1)^3$	جزء درجه ۴:
$(x - 1)^2(y - 1)^3$	جزء درجه ۵:

برای به دست آوردن چندجمله‌ای تیلور درجه k حول $(1, 1, 0)$ اجزاء تا درجه k را با هم جمع می‌کنیم. بدین ترتیب چندجمله‌ای‌های تیلور درجه ۵ به بالا همه برابرند. توجه این روش یگانگی چندجمله‌ای تیلور است. در واقع طبق (۹) ضریب جمله $(x_{i_1} - a_{i_1}) \cdots (x_{i_k} - a_{i_k})$ برابر است با

$$\frac{1}{k!} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a)$$

حال اگر از

$$f(x, y, z) = 1 + [2(x - 1) + 2(y - 1) - z] + \cdots + [(x - 1)(y - 1)]^2$$

از هر مرتبه مشتق گرفته و $(x, y, z) = (1, 1, 0)$ را جایگزین می‌کنیم ضرایب بالا به دست می‌آیند.
اکنون به ذکر پاره‌ای از خواص چندجمله‌ای تیلور می‌پردازیم. این خواص به سادگی از خواص مشابه در حالت یک متغیری به یاری تابع کمکی $\varphi(t) = f(a + th)$ که در آغاز این جلسه ذکر شد به دست می‌آیند. ویژگی مشابه (الف) در (۹) آمد. برای (ب) توجه کنید که:

$$f(a_1 + th_1, \dots, a_n + th_n) - T_p(a_1 + th_1, \dots, a_n + th_n) = \varphi(t) - [\varphi(0) + \varphi'(0)t + \dots + \frac{\varphi^{(p)}(0)}{p!}t^p]$$

عبارت طرف راست وقتی بر $|t|^p$ تقسیم شود، بنابر ویژگی چندجمله‌ای تیلور یک متغیری به صفر میل می‌کند وقتی $t \rightarrow 0$ ، از طرفی دیگر با قرار دادن $x = a + th$ می‌بینیم که $x \rightarrow a$ معادل $t \rightarrow 0$ است. بالاخره نوعی تخمین خطا به شکل باقیمانده لاگرانژ را بررسی می‌کنیم. برای $t = 1$ در عبارت بالا داریم

$$f(x) - T_p(x) = \varphi(1) - [\varphi(0) + \varphi'(0) \cdot 1 + \dots + \frac{\varphi^{(p)}(0)}{p!} \cdot 1^p]$$

فرض می‌کنیم f دارای مشتق‌های پاره‌ای مرتبه $(p + 1)$ پیوسته باشد که در نتیجه $\varphi(t) = f(a + th)$ ، $(p + 1)$ بار مشتق پذیر خواهد شد و با استفاده از باقیمانده لاگرانژ در حالت یک متغیری عددی $t = 0$ و 1 وجود دارد که:

$$\varphi(1) - [\varphi(0) + \varphi'(0) \cdot 1 + \dots + \frac{\varphi^{(p)}(0)}{p!} \cdot 1^p] = \frac{\varphi^{(p+1)}(t_0)}{(p+1)!} \cdot 1^{p+1}$$

اگر نقطه $a + t_0 h$ را ξ بنامیم، عبارت طرف راست مشابه جزء درجه $(p + 1)$ چندجمله‌ای تیلور به شکل زیر محاسبه می‌شود:

$$\frac{1}{(p+1)!} (\Delta \cdot \partial)^{p+1} \cdot f(\xi) \quad (28)$$

عبارت فوق را باقیمانده لاگرانژ در حالت n متغیری می‌نامند. مقدار آن خطای تقریب به وسیله چندجمله‌ای تیلور درجه p را نمایش می‌دهد و همان طور که انتظار می‌رود به مشتق‌های پاره‌ای مرتبه $(p + 1)$ (در یک نقطه نامشخص روی پاره‌خط واصل از a به x) وابسته است. برای به کار گرفتن (۲۸) باید کرانی برای مشتق‌های پاره‌ای مرتبه $(p + 1)$ در دست داشت.

مثال ۳ مقداری تقریبی برای $\sqrt{(3/1)^2 - (0/9)}$ از روش تقریب خطی به دست آورید و خطا را تخمین بزنید. در اینجا تابع $f(x, y) = \sqrt{x^2 - y}$ که قبلاً بررسی شد به ذهن می رسد. چون $f(3, 1) = 2$ و $f(3/1, 0/9)$ نزدیک $(3, 1)$ است، از تقریب خطی f حول $(3, 1)$ استفاده می کنیم:

$$f(3 + 0/1, 1 - 0/1) \sim 2 + \frac{1}{4}(0/1) - \frac{1}{14}(-0/1) = 2 + \frac{1}{4} + \frac{1}{14} = 2/0.583$$

با استفاده از باقیمانده لاگرانژ خطای این تقریب را تخمین می زنیم. برای نقطه ای (ξ, η) بین $(3, 1)$ و $(3/1, 0/9)$ خطا برابر است با:

$$\frac{1}{2!} \left[\frac{\partial^2 f}{\partial x^2}(\xi, \eta) \cdot (0/1)^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\xi, \eta) \cdot (0/1)(-0/1) + \frac{\partial^2 f}{\partial y^2}(\xi, \eta) \cdot (-0/1)^2 \right]$$

طبق محاسبات مشتق دوم در مثال ۱:

$$\left| \frac{\partial^2 f}{\partial x^2}(\xi, \eta) \right| \leq \left(\frac{2}{3} \right) \frac{1}{\sqrt{(\xi^2 - \eta)^2}} + \left(\frac{8}{9} \right) \frac{\xi^2}{\sqrt{(\xi^2 - \eta)^5}}$$

ماکزیمم عبارت طرف راست وقتی حاصل می شود که مخرج کوچکترین مقدار ممکن را داشته باشد و در پاره خط واصل بین $(3, 1)$ و $(3/1, 0/9)$ این وقتی حاصل می شود که $\xi = 3$ و $\eta = 1$ پس

$$\left| \frac{\partial^2 f}{\partial x^2}(\xi, \eta) \right| \leq \frac{1}{6} + \left(\frac{8}{9} \right) \frac{10}{32} = \frac{4}{9}$$

(توجه کنید که برای ξ^2 کران بالایی $(3/1)^2 > 10$ در نظر گرفته شده است.) به همین ترتیب:

$$\left| \frac{\partial^2 f}{\partial x \partial y}(\xi, \eta) \right| \leq \left(\frac{4}{9} \right) \frac{3/1}{32} = \frac{1}{8} \times \frac{3/1}{9} < \frac{3/2}{8 \times 9} = \frac{4}{90}$$

$$\left| \frac{\partial^2 f}{\partial y^2}(\xi, \eta) \right| \leq \left(\frac{2}{9} \right) \frac{1}{32} = \frac{1}{144}$$

پس

$$|خطا| \leq \frac{1}{200} \left(\frac{4}{9} + \frac{8}{90} + \frac{1}{144} \right) = \frac{1162}{200 \times 1440} < \frac{1}{200} = 5 \times 10^{-3}$$

پس تقریب به دست آمده حداقل تا دو رقم اعشار درست است. محاسبه با ماشین حساب $(\frac{0}{9}) - \sqrt{(\frac{3}{1})^2 - (\frac{0}{9})}$ را $2/0574978$ می دهد که با این نتایج سازگار است.

در اینجا باید ذکر کرد که روش های دیگری به جز باقیمانده لاگرانژ برای تخمین خطا وجود دارد. از جمله، روش "باقیمانده انتگرال" که در آن به جای استفاده از مشتق های مرتبه $(p + 1)$ در یک نقطه نامشخص، از نوعی میانگین موزون مشتق های مرتبه $(p + 1)$ در سراسر پاره خط واصل میان a و x استفاده می شود.

در پایان مختصراً به سری تیلور اشاره می کنیم. اگر تابع f در نقطه a (یک نقطه درونی زیرمجموعه S از \mathbb{R}^n) دارای مشتق های پاره ای از همه مراتب باشد، می توان سری تیلور آن حول a را به صورت زیر تشکیل داد:

$$f(a) + \sum_{i=0}^{\infty} \frac{1}{i!} ((\Delta \cdot \partial)^i f)(a)$$

موضوع مهم در اینجا، مانند حالت یک متغیری، وجود عددی $\rho > 0$ است که در گوی شعاع ρ حول a ، یعنی به ازای x های دامنه که $|x - a| < \rho$ ، سری بالا به مقدار تابع f میل کند. چنین $\rho > 0$ ممکن است وجود داشته باشد یا نباشد. مانند حالت یک متغیری، اگر بتوانیم نشان دهیم که باقیمانده لاگرانژ (یا هر صورت دیگر باقیمانده)، به ازای $|x - a| < \rho$ ، به صفر میل می کند وقتی $p \rightarrow +\infty$ ، آنگاه همگرایی سری تیلور ثابت می شود.

نقاط بحرانی

طبق معمول S را زیرمجموعه‌ای از \mathbb{R}^n ، a را یک نقطه درونی S ، و $f : S \rightarrow \mathbb{R}$ را یک تابع دارای مشتق‌های پاره‌ای مرتبه اول پیوسته می‌گیریم. در این بخش به بررسی نقاط درونی a که در آن $\nabla f(a) = \underline{0}$ می‌پردازیم. این نقاط را نقاط بحرانی تابع f می‌نامیم. با ذکر تعدادی مثال بحث را آغاز می‌کنیم.

مثال ۱ (نقطه می‌نیمم منزوی) f را به صورت $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ، $f(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2$ در نظر می‌گیریم. $\nabla f(a) = (2x_1, \dots, 2x_n)$ و $(0, \dots, 0)$ تنها نقطه بحرانی f است. تابع f در این نقطه صفر است و در سایر نقاط \mathbb{R}^n مقدار مثبت دارد، پس $(0, \dots, 0)$ مینیمم تابع f است. نمودار تابع f در حالت $n = 2$ همراه با مجموعه‌های تراز در شکل ۱ نمایش داده شده است.

مثال ۲ (نقطه ماکسیمم منزوی) اگر $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت $f(x_1, \dots, x_n) = -\sum_{i=1}^n x_i^2$ تعریف کنیم، $(0, \dots, 0)$ تنها نقطه بحرانی و ماکزیمم تابع خواهد بود (شکل ۲).

در واقع هر نقطه ماکسیمم یا مینیمم موضعی که نقطه درونی دامنه تابع بوده و تابع در آن نقطه مشتق پذیر باشد لزوماً بحرانی است. دلیل ساده این موضوع این است که اگر دامنه f را به خط راست گذرا از a به موازات محور x_i محدود کنیم f روی این دامنه نیز ماکسیمم یا مینیمم موضعی در نقطه a خواهد داشت، پس $\frac{\partial f}{\partial x_i}(a) = 0$. در واقع استدلال نشان می‌دهد که اگر تحدید f به خطوط گذرا به موازات محورها ماکسیمم موضعی یا مینیمم موضعی باشد، نقطه a بحرانی است.

مثال ۳ (نقطه زینی ساده) $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت $f(x_1, \dots, x_n) = \sum_{i=1}^k x_i^2 - \sum_{i=k+1}^n x_i^2$ تعریف می‌کنیم. مجدداً $\underline{0} = (0, \dots, 0)$ تنها نقطه بحرانی است. اگر تابع f به محورهای x_1, \dots, x_k محدود شود، $\underline{0}$ یک نقطه مینیمم روی این محورها خواهد بود ولی تحدید f به هر یک از

محورهای x_{k+1}, \dots, x_n در نقطه \underline{e} یک ماکسیمم دارد. ساده‌ترین نقطه زینی در حالت $n = 2$ برای $f(x, y) = x^2 - y^2$ در $(0, 0)$ مشاهده می‌شود. در شکل ۳ نمودار این تابع و مجموعه‌های تراز آن نمایش داده شده‌اند.

مثال ۴ $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ به صورت $f(x, y) = x^2$ تعریف شده است. $\nabla f(x, y) = (2x, 0)$ که نشان می‌دهد f در سراسر محور y بحرانی است. مقدار تابع در همه این نقاط صفر است و در سایر نقاط مثبت. مجموعه تراز $f(x, y) = c$ به ازای $c > 0$ از دو خط راست $x = \pm\sqrt{c}$ تشکیل شده است که برای $c = 0$ دو خط بر هم منطبق می‌شوند. (شکل ۴)

توجه کنید که برای یک تابع دو متغیری $f(x, y)$ ، صفحه مماس بر نمودار تابع در یک نقطه بحرانی افقی است. در مورد ماکزیمم و مینیمم منزوی این صفحه در یک طرف نمودار تابع قرار می‌گیرد، برای نقاط زینی بخشی از نمودار در یک طرف و بخشی دیگر در طرف دیگر قرار می‌گیرد. در مثال ۴ بالا صفحه مماس و نمودار در یک خط راست اشتراک دارند.

مثال ۵ $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را به صورت $f(x, y) = x^2 + y^3$ در نظر می‌گیریم. مجدداً تنها نقطه بحرانی، نقطه \underline{e} است. در اینجا روی محور x ، f یک مینیمم در \underline{e} دارد و روی محور y ، یک نقطه عطف در همان نقطه. نمودار تابع در امتداد سه نیمخط از محورهای مختصات با دور شدن از \underline{e} صعودی است و در امتداد نیمخط چهارم با دور شدن از \underline{e} نزولی (شکل ۵).

نمودار تابع‌های مثال‌های ۱-۵ و مجموعه‌های تراز آنها

مثال ۶ (زین میمون) تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ را در نظر بگیرید که به صورت $f(x, y) = x^3 - 3xy^2$ تعریف شده است. داریم $\nabla f(x, y) = (3x^2 - 3y^2, -6xy)$ و $(0, 0)$ تنها نقطه بحرانی است. توجه کنید که عبارت f به صورت زیر تجزیه می‌شود:

$$f(x, y) = x(x - \sqrt{3}y)(x + \sqrt{3}y)$$

پس مجموعه تراز $f(x, y) = 0$ از اجتماع سه خط متقاطع $x = \sqrt{3}y$ ، $x = -\sqrt{3}y$ و $x = 0$ تشکیل شده است. این سه خط صفحه xy را به ۶ ناحیه تقسیم می‌کنند (شکل ۶) که علامت f روی هر یک

نمایش داده شده است. برای $c > 0$ کوچک یک مجموعه تراز نمونه در شکل ۶ رسم شده است که یک خم دارای سه بخش (شاخه) است. برای $c < 0$ نیز شکل مشابهی در سه بخش دیگر صفحه مشاهده می شود. نمودار این تابع روی سه نیمخط $\frac{2\pi}{3}, \frac{4\pi}{3}, \frac{5\pi}{3}$ با $\theta = 0$ دور شدن از $\underline{0}$ صعودی و روی سه نیمخط $\frac{\pi}{3}, \pi, \frac{5\pi}{3}$ با دور شدن از $\underline{0}$ نزولی است.

شکل ۶

مثال های بالا نشان می دهند که تنوع رفتار تابع ها حول یک نقطه بحرانی حتی در حالت دو بعدی به مراتب گسترده تر از حالت یک بعدی است. هدف بعدی ما ارائه نوعی "آزمون مشتق دوم" است که می تواند مانند آزمون مشتق دوم یک متغیری در تشخیص ماکسیمم و مینیمم موضعی یا نقاط زینی ساده می تواند مؤثر واقع شود. فرض کنید برای تابع $f: S \rightarrow \mathbb{R}$ زیرمجموعه \mathbb{R}^n ، نقطه درونی a از S یک نقطه بحرانی است و f در یک گوی باز حول a دارای مشتق های پاره ای مرتبه دوم پیوسته است. چون مشتق های پاره ای مرتبه اول در نقطه a صفر هستند، قضیه تیلور با باقیمانده لاگرانژ به صورت زیر در می آید:

$$f(x) - f(a) = \frac{1}{2!} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi) \cdot (x_i - a_i)(x_j - a_j) \quad (1)$$

که در اینجا ξ نقطه ای روی پاره خط واصل از a به x است. $\frac{\partial^2 f}{\partial x_i \partial x_j}(\xi)$ را A_{ij} و $x_i - a_i$ را X_i می نامیم و عبارت زیر را در نظر می گیریم:

$$X = (X_1, \dots, X_n) \quad , \quad Q(X) = \sum_{i,j=1}^n A_{ij} X_i X_j \quad (2)$$

اگر عبارت درجه دوم همگون $Q(X)$ طوری باشد که برای $X \neq \underline{0}$ و $|X|$ کوچک، $Q(X) > 0$ ، از (۱) نتیجه می شود که f در نقطه a یک مینیمم موضعی اکید دارد، و نیز اگر برای $|X|$ کوچک $X \neq \underline{0}$ ، $Q(X)$ منفی باشد، نتیجه می شود که a یک نقطه ماکسیمم موضعی اکید برای f است. هرگاه برای مقادیر کوچک و ناصفر $|X|$ ، $Q(X)$ هم مقادیر مثبت و هم مقادیر منفی بپذیرد، به همین ترتیب می بینیم که a نه یک نقطه مینیمم موضعی و نه یک نقطه ماکسیمم موضعی است. بدین ترتیب چنانچه

علامت یابی $Q(X)$ میسر باشد، می توان در مورد ماهیت نقطه بحرانی a حکم کرد. برای $n = 2$ یک چنین علامت یابی بر پایه جبر مقدماتی در دسترس است. در اینجا:

$$Q(X) = A_{11}X_1^2 + 2A_{12}X_1X_2 + A_{22}X_2^2 \quad (3)$$

قرار می دهیم $\Delta = A_{22}^2 - 4A_{11}A_{12}$. داریم:

• اگر $\Delta < 0$ ، $Q(X)$ هم علامت A_{11} (و نیز A_{22}) است مگر وقتی $(X_1, X_2) = (0, 0)$.

• اگر $\Delta > 0$ ، $Q(X)$ تغییر علامت می دهد.

حالت دوم را دقیقتر بررسی می کنیم. اگر $A_{11} = A_{22} = 0$ ، که آنگاه $A_{12} \neq 0$ ، و علامت $Q(X) = 2A_{12}X_1X_2$ با علامت X_1 و X_2 تغییر می کند. بدین ترتیب دو خط راست $x_1 = a_1$ و $x_2 = a_2$ که از نقطه بحرانی $a = (a_1, a_2)$ می گذرند صفحه را به چهاربخش تقسیم می کنند که علامت Q در آنها متناوباً مثبت و منفی است. بنابراین در این حالت a یک نقطه زینی است. حال فرض کنید دست کم یکی از A_{11} و A_{22} صفر نیست، مثلاً $A_{11} \neq 0$. اگر α و β ریشه های معادله درجه دوم $A_{11}t^2 + 2A_{12}t + A_{22} = 0$ باشند، می توان نوشت:

$$Q(X) = A_{11}(X_1 - \alpha X_2)(X_1 - \beta X_2)$$

توجه کنید که $\alpha \neq \beta$ زیرا $\Delta > 0$. بنابراین وقتی نسبت $\frac{X_1}{X_2}$ از ریشه ها عبور می کند تغییر علامت حاصل می شود. معادلاً $Q = A_{11}(x_1 - \alpha x_2 - a_1 + \alpha a_2)(x_1 - \beta x_2 - a_1 + \beta a_2)$ ، دو خط راست متمایز $x_1 - \alpha x_2 - a_1 + \alpha a_2 = 0$ و $x_1 - \beta x_2 - a_1 + \beta a_2 = 0$ که از نقطه بحرانی $a = (a_1, a_2)$ می گذرند صفحه را به چهاربخش تقسیم می کنند که Q در بخش های مجاور علامت مختلف دارد. بنابراین a یک نقطه زینی است. بحث بالا را در زیر خلاصه می کنیم:

آزمون مشتق دوم برای تابع های دو متغیری S زیرمجموعه \mathbb{R}^2 است، a یک نقطه درونی S ، $f : S \rightarrow \mathbb{R}$ تابعی با مشتق های پاره ای مرتبه دوم پیوسته در یک گوی باز حول a که a یک نقطه بحرانی برای آن است. در این صورت:

(الف) اگر $f_{xx} > 0$ و $f_{xy}^2 - f_{xx}f_{yy} < 0$ در نقطه a ، یک نقطه مینیم موضعی اکید است.

(ب) اگر $f_{xx} < 0$ و $f_{xy}^2 - f_{xx}f_{yy} < 0$ در نقطه a ، یک نقطه ماکسیم موضعی اکید است.

(ج) اگر $f_{xy}^2 - f_{xx}f_{yy} > 0$ ، یک نقطه زینی ساده است.

ذکر چند نکته ضروری است. اول اینکه از آنجا که مشتقات پاره‌ای مرتبه دوم پیوسته فرض شده‌اند، هر علامت (مثبت یا منفی) که f_{xx} و $f_{xy}^2 - f_{xx}f_{yy}$ در a داشته باشند، همین علامت را در یک گوی باز حول a حفظ می‌کنند، پس علامت باقیمانده لاگرانژ نیز به همین صورت است. دوم اینکه در (الف) و (ب)، از منفی بودن $f_{xy}^2 - f_{xx}f_{yy}$ نتیجه می‌شود که f_{xx} و f_{yy} هم علامت هستند، پس می‌توان به جای علامت f_{xx} از علامت f_{yy} در نقطه a کمک گرفت. بالاخره اینکه وقتی $f_{xy}^2 - f_{xx}f_{yy}$ در نقطه a صفر شود، نمی‌توان حکمی در مورد نوع نقطه a ابراز کرد. مثال‌های ۴، ۵ و ۶ سه نمونه متفاوت از این نوع بودند. برای تابع‌های $f(x, y) = x^4 + y^4$ و $f(x, y) = -(x^4 + y^4)$ و $f(x, y) = x^4 - y^4$ عبارت $f_{xy}^2 - f_{xx}f_{yy}$ در نقطه بحرانی $(0, 0)$ صفر است، در حالی که این نقطه برای سه تابع فوق به ترتیب مینیم، ماکسیم و زینی می‌باشد.

اکنون تعمیم آزمون مشتق دوم به حالت n متغیری را در نظر می‌گیریم. همان طور که قبلاً اشاره کردیم اگر a یک نقطه بحرانی برای تابع n متغیری f باشد، تعیین نوع نقطه بحرانی به بررسی علامت عبارت $Q(X) = \sum_{i,j=1}^n A_{ij} X_i X_j$ منجر می‌شود (به بحث پیرامون (۲) مراجعه کنید). در این حالت کلی نیز می‌توان با استفاده از قضیه‌ای در جبرخطی در مورد علامت $Q(X)$ بحث کرد. ماتریس هسیان تابع f در نقطه بحرانی a را به صورت زیر تعریف می‌کنیم:

$$Hf(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix}$$

از آنجا که مشتق‌های پاره‌ای مرتبه دوم را پیوسته فرض کرده‌ایم این ماتریس متقارن است. برای هر ماتریس $A = [a_{ij}]$ ، $n \times n$ مقادیر ویژه A ریشه‌های معادله درجه n زیر برحسب λ هستند:

$$\det(A - \lambda I) = 0 \quad (4)$$

که در آن I ماتریس واحد $n \times n$ ، یعنی $\begin{bmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$ است. در جبرخطی ثابت می شود که اگر ماتریس A متقارن باشد (مانند ماتریس هسیان)، همه ریشه های (۴) حقیقی هستند (تکرار ریشه ها ممکن است).

آزمون مشتق دوم برای تابع های n متغیری S زیرمجموعه \mathbb{R}^n است، a یک نقطه درونی S ، $f: S \rightarrow \mathbb{R}$ تابعی با مشتق های پاره ای مرتبه دوم پیوسته در یک گوی باز حول a که a یک نقطه بحرانی برای آن است. مقادیر ویژه ماتریس هسیان را با منظور کردن تعدد ریشه ها به $\lambda_1, \dots, \lambda_n$ نمایش می دهیم. در این صورت:

(الف) اگر همه λ_i ها مثبت باشند، a یک نقطه مینیم موضعی اکید است.

(ب) اگر همه λ_i ها منفی باشند، a یک نقطه ماکسیم موضعی اکید است.

(ج) اگر همه λ_i ها ناصفر باشند، k تا مثبت و $(n - k)$ تا منفی، آنگاه a یک نقطه زینی ساده است و n امتداد متمایز در \mathbb{R}^n وجود دارد که f روی k امتداد از آنها در نقطه a مینیم موضعی دارد و روی $(n - k)$ امتداد دیگر ماکسیم موضعی.

در حالتی که حتی یک λ_i صفر شود، رفتارهای دیگری نیز می توان انتظار داشت. اثبات آزمون خارج از برنامه ما است ولی توجه کنید که در حالت $n = 2$ ، همان آزمون مشتق مرتبه دوم توابع دو متغیری به دست می آید. در اینجا ماتریس هسیان عبارت است از:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

و مقادیر ویژه از حل معادله زیر به دست می آیند:

$$\det \begin{bmatrix} f_{xx} - \lambda & f_{xy} \\ f_{xy} & f_{yy} - \lambda \end{bmatrix} = \lambda^2 - (f_{xx} + f_{yy})\lambda + f_{xx}f_{yy} - f_{xy}^2 = 0 \quad (5)$$

شرط های (الف) و (ب) بالا وقتی برقرار می شوند که $\lambda_1 \lambda_2 > 0$ ، یعنی $f_{xy}^2 - f_{xx}f_{yy} < 0$ که این عیناً شرط داشتن ماکسیم یا مینیم موضعی است. مجموع دو ریشه معادله (۵) برابر می شود با $f_{xx} + f_{yy}$.

پس در حالت (الف) و (ب) مثبت بودن (به ترتیب منفی بودن) هر دو مقدار ویژه معادل است با این که $f_{xx} + f_{yy} > 0$ (به ترتیب $f_{xx} + f_{yy} < 0$). ولی در این حالت f_{xx} و f_{yy} همعلامتند، پس همان نتیجه آزمون مشتق توابع دو متغیری به دست می آید.

مثال نقاط بحرانی تابع $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ که به صورت $f(x, y, z) = x^2 + 12yz + (y - z)^3$ تعریف شده است را پیدا کنید و نوع آنها را شناسایی کنید.

حل داریم $(2x, 12z + 3(y - z)^2, 12y - 3(y - z)^2) = \nabla f(x, y, z)$ ، و دو نقطه بحرانی $P = (0, 0, 0)$ و $Q = (0, 1, -1)$ به دست می آیند. ماتریس هسیان در هر مورد به صورت زیر محاسبه می شود:

$$H(P) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 12 \\ 0 & 12 & 0 \end{bmatrix}, \quad H(Q) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 12 \end{bmatrix}$$

در مورد Q ، مقادیر ویژه عبارتند از $2, 12, 12$ که هر سه مثبت هستند، پس $(0, 1, -1)$ یک نقطه مینیمم موضعی اکید است. برای P مقادیر ویژه را بررسی می کنیم:

$$\det \begin{bmatrix} 2 - \lambda & 0 & 0 \\ 0 & -\lambda & 12 \\ 0 & 12 & -\lambda \end{bmatrix} = 0$$

$$(2 - \lambda)(\lambda^2 - 144) = 0$$

مقادیر ویژه عبارتند از $\lambda = 2, \pm 12$ ، پس $(0, 0, 0)$ یک نقطه زینی ساده است.

توابع ضمنی

در حساب دیفرانسیل یک متغیری به مطلب زیر برخورد کنید. رابطه‌ای $F(x, y) = 0$ داده شده است. اگر فرض کنیم بتوان از این رابطه y را به عنوان تابعی مشتق پذیر از x در نظر گرفت، $y = f(x)$ ، می‌خواهیم مشتق این تابع، یعنی $f'(x) = \frac{dy}{dx}$ ، را محاسبه کنیم. اینکه بتوان y را به عنوان تابعی از x از $F(x, y) = 0$ استخراج کرد نتیجه می‌دهد که $F(x, f(x))$ به ازای همه مقادیر x صفر است. بنابراین مشتق عبارت $F(x, f(x))$ نسبت به x همواره صفر خواهد بود. با استفاده از قاعده زنجیره‌ای دو متغیری داریم:

$$\frac{d}{dx}(F(x, f(x))) = \frac{\partial F}{\partial x} \frac{dx}{dx} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0 \quad (1)$$

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0$$

حال اگر در نقطه (x, y) داشته باشیم $\frac{\partial F}{\partial y}(x, y) \neq 0$ ، می‌توان نوشت:

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \quad (2)$$

روش فوق گاهی زیرعنوان "مشتق‌گیری ضمنی" مطرح می‌شود. این روش اغلب بر حل رابطه $F(x, y) = 0$ برای y بر حسب x و سپس مشتق گرفتن، ارجحیت دارد. دلایل عمده برتری ادعا شده به این شرح است:

الف) ممکن است استخراج صریح y بر حسب x ناممکن یا دشوار باشد ولی مشتق را بتوان از (2) به سادگی محاسبه کرد.

ب) اگر در نقطه‌ای (x, y) داشته باشیم $\frac{\partial F}{\partial y} = 0$ و $\frac{\partial F}{\partial x} \neq 0$ ، این امر می‌تواند دلیل بر این باشد که مماس بر خم تعریف شده به وسیله $F(x, y) = 0$ قائم باشد. در این صورت باید بتوان x را به صورت تابعی مشتق پذیر از y حول (x, y) فوق نوشت.

ج) مهمترین دلیل این است که رابطه $F(x, y) = 0$ عموماً y را به عنوان تابعی از x تعریف نمی‌کند. مجموعه نقاط (x, y) که در $F(x, y) = 0$ صدق می‌کنند ممکن است مثلاً خمی مانند شکل ۱ باشد.

شکل ۱

در اینجا از شکل چنین بر می‌آید که در نقاط P, Q و S مماس بر خم قائم است و در نقطه R ، نه y تابعی مشتق‌پذیر از x است و نه x تابعی مشتق‌پذیر از y . اگر نقاط P, Q, R و S را حذف کنیم، ظاهراً می‌توان پنج تابع به شکل $y = f(x)$ از رابطه فوق استخراج کرد. حتی اگر بتوان عبارت این تابع‌ها را صریحاً به دست آورد، مشتق‌گیری جداگانه از آنها دشوارتر از این است که از $F(x, y) = 0$ یک بار به طور ضمنی مشتق‌گیری شود. مشتق‌گیری ضمنی این ویژگی را نیز دارد که نقاط P, Q و S را که در آنها مماس قائم وجود دارد بدون تبعیض نمایان می‌کند. حتی در مورد عبارت ساده $x^2 + y^2 = R^2$ ، دایره شعاع R حول $(0, 0)$ ، از مشتق‌گیری ضمنی $x + y \frac{dy}{dx} = 0$ یا $x \frac{dx}{dy} + y = 0$ می‌توان ضریب زاویه خط مماس در هر نقطه دایره را به سادگی محاسبه کرد، و این روش طبیعی‌تر از استخراج دو تابع $y = \pm \sqrt{R^2 - x^2}$ و کارکردن با آنهاست که به هر حال دو نقطه $(\pm R, 0)$ را پوشش نمی‌دهند.

فرمول مشابه (۲) را می‌توان برای مشتق‌گیری ضمنی از یک عبارت با تعداد متغیر بیشتر نیز به دست آورد. فرض کنید $F(x_1, \dots, x_n, y) = 0$. اگر فرض کنیم از این عبارت، y به صورت تابعی مشتق‌پذیر از x_1, \dots, x_n استخراج شدنی است، $y = f(x_1, \dots, x_n)$ ، آنگاه $F(x_1, \dots, x_n, f(x_1, \dots, x_n))$ به‌ازای همه مقادیر (x_1, \dots, x_n) صفر است و اگر نسبت به x_j مشتق بگیریم حاصل صفر خواهد شد.

$$\frac{\partial}{\partial x_j} (F(x_1, \dots, x_n, f(x_1, \dots, x_n))) = 0$$

$$\sum_{i=1}^n \frac{\partial F}{\partial x_i} \frac{\partial x_i}{\partial x_j} + \frac{\partial F}{\partial y} \frac{\partial y}{\partial x_j} = 0$$

$$\frac{\partial F}{\partial x_j} + \frac{\partial F}{\partial y} \frac{\partial y}{\partial x_j} = 0$$

بنابراین وقتی $\frac{\partial F}{\partial y} \neq 0$ داریم:

$$\frac{\partial y}{\partial x_j} = - \frac{\frac{\partial F}{\partial x_j}}{\frac{\partial F}{\partial y}} \quad (3)$$

که مشابه (۲) است. توجه کنید که در (۲) و (۳)، علاوه بر این فرض که $\frac{\partial F}{\partial y} \neq 0$ ، به طور ضمنی این نیز فرض شده است که واقعاً می‌توان y را به صورت تابعی مشتق‌پذیر از x ، یا (x_1, \dots, x_n) ، نوشت. موضوع اصلی بخش حاضر بررسی شرایط صحت این فرض و بحث پیرامون نتایج آن است. در واقع به جای یک رابطه $F(x_1, \dots, x_n, y) = 0$ ، حالت کلی یک مجموعه رابطه را در نظر می‌گیریم.

فرض کنید m رابطه زیر میان $(m+n)$ متغیر $x_1, \dots, x_n, y_1, \dots, y_m$ داده شده‌اند:

$$\begin{cases} F_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \\ \vdots \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \end{cases} \quad (4)$$

که در آن هر F_i دارای مشتق‌های پاره‌ای پیوسته نسبت به کلیه متغیرهای $x_1, \dots, x_n, y_1, \dots, y_m$ است. با توجه به تعداد روابط، m ، انتظار می‌رود که معمولاً بتوان m متغیر y_1, \dots, y_m را برحسب متغیرهای باقیمانده، x_1, \dots, x_n ، به صورت تابع‌هایی $y_1 = f_1(x_1, \dots, x_n)$ ، \dots ، $y_m = f_m(x_1, \dots, x_n)$ استخراج کرد. البته این انتظار بعضی اوقات برآورده نمی‌شود. مثلاً در شکل ۱ دیدیم که پنج تابع مختلف $y = f(x)$ از رابطه داده شده استخراج می‌شود، یا از رابطه $x^2 + y^2 + 1 = 0$ هیچ تابعی نمی‌توان استخراج کرد چون هیچ نقطه (x, y) در این رابطه صدق نمی‌کند. به همین ترتیب اگر دستگاه دو معادله سه مجهولی زیر را در نظر بگیریم:

$$\begin{cases} x^2 + y^2 + z^2 - 1 = 0 \\ x + y + z - 2 = 0 \end{cases}$$

این دستگاه جواب ندارد (چرا؟). موضوعی که در اینجا بررسی خواهیم کرد به شرح زیر است. فرض می‌کنیم (۴) دارای دست کم یک جواب باشد، یعنی نقطه‌ای $(x_1, \dots, x_n) = x = a = (a_1, \dots, a_n)$ و $(y_1, \dots, y_m) = y = b = (b_1, \dots, b_m)$ وجود داشته باشند که در (۴) صدق کند. در این صورت "قضیه تابع ضمنی" که در زیر خواهد آمد، شرطی کافی ارائه می‌کند برای این که بخشی (احتمالاً

کوچک) از مجموعه جواب (۴) که حول (a, b) قرار دارد به شکل نمودار تابعی مشتق پذیر $y = f(x)$ باشد. چون نمودار یک تابع مشتق پذیر را معمولاً "هموار" تعبیر می‌کنیم، به این مفهوم، و تحت شرط کافی که ارائه خواهد شد، مجموعه جواب در نزدیکی (a, b) "هموار" تلقی می‌شود.

(۲۸-۱) قضیه تابع ضمنی فرض می‌کنیم F_i ها در (۴) دارای مشتق‌های پاره‌ای پیوسته نسبت به متغیرهای x و y هستند، نقطه $(x, y) = (a, b)$ در (۴) صدق می‌کند و:

$$\det\left[\frac{\partial F_i}{\partial y_i}(a, b)\right] \neq 0 \quad (5)$$

در این صورت $r_1 > 0$ و $r_2 > 0$ وجود دارند و تابعی مشتق پذیر f از گوی باز شعاع r_1 حول a به گوی باز شعاع r_2 حول b ، به طوری که نمودار تابع f دقیقاً برابر مجموعه نقاط (x, y) واجد شرط‌های $|x| < r_1$ و $|y| < r_2$ است که در (۴) صدق می‌کنند.

در حالت $m = n = 1$ ، شرط (۵) به این معنی است که $\frac{\partial F}{\partial y}(a, b) \neq 0$ و شکل ۲ حکم قضیه را نمایش می‌دهد. توجه کنید که فقط بخش کوچکی از مجموعه جواب که حول (a, b) قرار دارد به صورت نمودار یک تابع مشتق پذیر $y = f(x)$ ظاهر می‌شود.

شکل ۲

اثبات قضیه تابع ضمنی در این درس ارائه نخواهد شد ولی از صورت آن استفاده فراوان خواهد شد. در زیر به بحث پیرامون بعضی کاربردها و حواشی قضیه خواهیم پرداخت.

(۲۸-۲-۱) شرط (۵) یک شرط کافی بسیار قابل استفاده است. بدون آن نیز گاهی حکم قضیه برقرار می‌شود. مثلاً اگر رابطه $y^3 - x^4 = 0$ را در نظر بگیریم، برای $F(x, y) = y^3 - x^4$ داریم $\frac{\partial F}{\partial y}(0, 0) = 0$ ولی مجموعه جواب برابر نمودار تابع مشتق پذیر $y = x^{\frac{3}{4}}$ است. با این حال برای درک اهمیت (۵) لازم است با بعضی مشکلاتی که ممکن است در صورت عدم برقراری آن رخ دهد آشنا شویم. برای $m = n = 1$ سه مثال در نظر می‌گیریم. در ابعاد بالا تنوع مشکلات گسترده‌تر است. برای هر یک از سه تابع $F, G, H : \mathbb{R}^2 \rightarrow \mathbb{R}$ که به صورت

$$F(x, y) = x^2 + y^2, \quad G(x, y) = x + y^2, \quad H(x, y) = x + y^3$$

تعریف شده‌اند، مجموعهٔ جواب، به ترتیب $F(x, y) = 0$ ، $G(x, y) = 0$ و $H(x, y) = 0$ در شکل‌های ۳الف، ۳ب و ۳ج نمایش داده شده‌اند. در هر مورد نقطه $(a, b) = (0, 0)$ در رابطه صدق می‌کند. در (الف) $\frac{\partial F}{\partial y}(0, 0) = 0$ و مجموعهٔ جواب $F(x, y) = 0$ از یک تک نقطه‌ای، $(0, 0)$ ، تشکیل شده است. طبعاً این تک نقطه نمودار تابعی نیست که روی یک گوی باز حول $a = 0$ تعریف شده باشد. در اینجا انتظار متعارف ما که مجموعهٔ جواب یک خم (یک بعدی) باشد برآورده نمی‌شود. در (ب)، $\frac{\partial G}{\partial y}(0, 0) = 0$. مجموعهٔ جواب نمی‌تواند نمودار یک تابع تعریف شده در مجموعهٔ بازی حول $a = 0$ باشد. در (ج)، $\frac{\partial H}{\partial y}(0, 0) = 0$. مجموعهٔ جواب در این مورد نمودار یک تابع از \mathbb{R} به \mathbb{R} است، ولیکن این تابع در $a = 0$ مشتق‌پذیر نیست (ماس قائم).

(۲۸-۲-۲) هموار بودن مجموعهٔ تراز. فرض کنید S زیرمجموعه‌ای از \mathbb{R} باشد، $F : S \rightarrow \mathbb{R}$ تابعی مشتق‌پذیر با مشتق‌های پاره‌ای پیوسته و a یک نقطهٔ درونی S . قبلاً اشاره کرده‌ایم که اگر $\nabla F(a) \neq 0$ ، آنگاه مجموعهٔ تراز گذرا از a به مفهومی ”هموار“ است. این مطلب را می‌توانیم اکنون به طور دقیق توضیح دهیم. $\nabla F(a) = (\frac{\partial F}{\partial x_1}(a), \dots, \frac{\partial F}{\partial x_n}(a)) \neq 0$ بدین معنی است که دست کم یکی از مشتق‌های پاره‌ای F در نقطهٔ a صفر نیست. مثلاً فرض کنید $\frac{\partial F}{\partial x_n}(a) \neq 0$. اگر $F(a) = c$ ، یعنی a روی مجموعهٔ تراز منسوب به c قرار داشته باشد، تابع $G : S \rightarrow \mathbb{R}$ را به صورت $G(x) = F(x) - c$ تعریف می‌کنیم. مجموعه تراز فوق‌الذکر به مجموعهٔ منسوب به 0 برای G تبدیل می‌شود. برای این مجموعهٔ تراز داریم $G(x_1, \dots, x_n) = 0$ ، نقطه a در این رابطه صدق می‌کند و پس طبق قضیهٔ تابع ضمنی، تابعی مشتق‌پذیر f تعریف شده از یک گوی باز حول (a_1, \dots, a_{n-1}) به بازهٔ بازی حول a_n در \mathbb{R} وجود دارد که نمودار آن برابر بخشی از مجموعهٔ تراز است که حول a قرار دارد. به این مفهوم مجموعه تراز حول a هموار است.

برای بهره‌گیری کاملتر از قضیهٔ تابع ضمنی، لازم است همانند فرمول‌های (۲) و (۳)، فرمولی در حالت کلی برای مشتق تابع مشتق‌پذیر f که وجود آن به وسیلهٔ قضیهٔ تابع ضمنی تضمین می‌شود به دست آوریم. برای $y = f(x)$ ، ماتریس مشتق، $[\frac{\partial y_i}{\partial x_j}(a)]$ یک ماتریس $m \times n$ است، ماتریس مشتق‌های پاره‌ای F_i ها نسبت به (x_1, \dots, x_n) ، $[\frac{\partial F_i}{\partial x_j}(a, b)]$ نیز یک ماتریس $m \times n$ و ماتریس

مشتق‌های پاره‌ای f_i ها نسبت به (y_1, \dots, y_m) یک ماتریس مربعی $m \times m$ می‌باشد. ضمناً شرط $\det\left[\frac{\partial F_i}{\partial x_j}(a, b)\right] \neq 0$ تضمین می‌کند که وارون ماتریس $\left[\frac{\partial F_i}{\partial y_j}(a, b)\right]$ وجود دارد.

(۲۸-۳) متمم قضیه تابع ضمنی در شرایط قضیه تابع ضمنی، اگر (x, y) به اندازه کافی به (a, b) نزدیک باشد، داریم:

$$\left[\frac{\partial y_i}{\partial x_j}(x)\right] = -\left[\frac{\partial F_i}{\partial y_j}(x, y)\right]^{-1} \left[\frac{\partial F_i}{\partial x_j}(x, y)\right] \quad (6)$$

توجه کنید که چون مشتق‌های پاره‌ای پیوسته فرض شده‌اند و دترمینان تابعی پیوسته از درایه‌هاست (برحسب ضرب و جمع تعریف شده است)، صفر نبودن دترمینان $\left[\frac{\partial F_i}{\partial y_j}(x, y)\right]$ به‌ازای $(x, y) = (a, b)$ موجب می‌شود که این دترمینان در نزدیکی (a, b) نیز ناصفر باشد، پس وارون ماتریس $\left[\frac{\partial F_i}{\partial y_j}(x, y)\right]$ برای (x, y) به اندازه کافی نزدیک به (a, b) وجود خواهد داشت. توجه کنید که در حالت $m = 1$ ، (۶) به (۳) تبدیل می‌شود و در حالت خاص تر $m = 1$ و $n = 1$ ، این همان رابطه (۲) است.

اثبات (۶) با استفاده از قاعده زنجیره‌ای سرراست است. برای x در دامنه تابع f ذکر شده در قضیه

تابع ضمنی، $F(x, f(x))$ همواره ثابت صفر است، پس

$$\frac{\partial}{\partial x_j}(F(x, f(x))) = 0$$

$$\begin{bmatrix} \frac{\partial}{\partial x_j}(F_1(x, f(x))) \\ \vdots \\ \frac{\partial}{\partial x_j}(F_m(x, f(x))) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

مشتق‌گیری طرف چپ را به کمک قاعده زنجیره‌ای انجام می‌دهیم:

$$\begin{cases} \sum_{k=1}^n \frac{\partial F_1}{\partial x_k} \frac{\partial x_k}{\partial x_j} + \sum_{l=1}^m \frac{\partial F_1}{\partial y_l} \frac{\partial y_l}{\partial x_j} = 0 \\ \vdots \\ \sum_{k=1}^n \frac{\partial F_m}{\partial x_k} \frac{\partial x_k}{\partial x_j} + \sum_{l=1}^m \frac{\partial F_m}{\partial y_l} \frac{\partial y_l}{\partial x_j} = 0 \end{cases}$$

یا

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_j} \\ \vdots \\ \frac{\partial F_m}{\partial x_j} \end{bmatrix} + \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{\partial y_1}{\partial x_j} \\ \vdots \\ \frac{\partial y_m}{\partial x_j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

با انتقال ستون سمت چپ به طرف راست و ضرب کردن در وارون ماتریس $m \times m$ داریم

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_j} \\ \vdots \\ \frac{\partial y_m}{\partial x_j} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_j} \\ \vdots \\ \frac{\partial F_m}{\partial x_j} \end{bmatrix}$$

حال اگر برای $j = 1, \dots, m$ این ستون‌ها را کنار هم قرار دهیم، فرمول (۶) نتیجه می‌شود.

مثال ۱. روابط زیر داده شده‌اند:

$$\begin{cases} x^3 y + 2y^2 z^2 - xz^3 = 0 \\ e^{y+z} - x = 0 \end{cases}$$

نقطه $(x, y, z) = (1, 0, 0)$ در این روابط صدق می‌کند. انتظار عادی این است که هر یک از دو معادله فوق یک رویه (سطح خمیده) در \mathbb{R}^3 تعریف کند و اشتراک این دو رویه یک خم باشد. می‌دانیم $(1, 0, 0)$ در این اشتراک قرار دارد. آیا می‌توان برای بخشی از این اشتراک که در نزدیکی $(1, 0, 0)$ قرار دارد، دو متغیر را به صورت تابعی مشتق‌پذیر از متغیر سوم نوشت؟ در صورت جواب مثبت، می‌خواهیم مشتق دو متغیر را نسبت به متغیر سوم در نقطه $(1, 0, 0)$ محاسبه کنیم. این مسأله ظاهراً در چارچوب قضیه تابع ضمنی با $m = 2$ و $n = 1$ قرار می‌گیرد ولی مشخص نیست کدام دو از متغیرهای x, y, z می‌توانند نقش متغیرهای y_i را در قضیه ایفاء کنند. به منظور یافتن متغیرهای مناسب، ماتریس مشتق‌های عبارت‌های بالا را نسبت به متغیرهای x, y, z تشکیل می‌دهیم:

$$\begin{bmatrix} 3x^2 y - z^3 & x^3 + 4yz^2 & 4y^2 z - 3xz^2 \\ -1 & e^{y+z} & e^{y+z} \end{bmatrix}_{(1,0,0)} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

(ستون‌های اول تا سوم، به ترتیب، مشتق نسبت به x, y, z را نمایش می‌دهند). برای اینکه از قضیه تابع ضمنی استفاده کنیم باید دو متغیر را طوری انتخاب کنیم که ماتریس 2×2 متشکل از ستون‌های مربوط به آن دو متغیر دارای دترمینان ناصفر باشد. اگر $x^3 y + 2y^2 z^2 - xz^3$ را به $F_1(x, y, z)$ و

داریم $e^{y+z} - x$ را به $F_2(x, y, z)$ نمایش دهیم، داریم

$$\det \frac{\partial(F_1, F_2)}{\partial(x, y)} = \det \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} = 1$$

$$\det \frac{\partial(F_1, F_2)}{\partial(x, z)} = \det \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} = 0$$

$$\det \frac{\partial(F_1, F_2)}{\partial(y, z)} = \det \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = 1$$

بدین ترتیب، طبق قضیه تابع ضمنی، می‌توان حول $(1, 0, 0)$ ، (x, y) را به عنوان تابعی مشتق‌پذیر از z و نیز (y, z) را به عنوان تابعی مشتق‌پذیر از x در نظر گرفت، ولی قضیه تابع ضمنی تضمین نمی‌کند که بتوان (x, z) را به صورت تابعی مشتق‌پذیر از y نمایش داد (خواهیم دید که در واقع نمی‌توان حول $(1, 0, 0)$ ، (x, z) را تابعی مشتق‌پذیر از y پنداشت). برای محاسبه مشتق، فرض کنید (x, y) را تابعی مشتق‌پذیر از z در نظر گرفته‌ایم. طبق متمام قضیه تابع ضمنی داریم:

$$\begin{bmatrix} \frac{dx}{dz} \\ \frac{dy}{dz} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial z} \end{bmatrix}$$

و در نقطه $(1, 0, 0)$:

$$\begin{bmatrix} \frac{dx}{dz} \\ \frac{dy}{dz} \end{bmatrix} = - \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = - \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\frac{dx}{dz} = 1, \quad \frac{dy}{dz} = 0$$

محاسبه مشتق‌های فوق را می‌توان بدون استفاده از متمام قضیه تابع ضمنی انجام داد به این شرط که متغیرهای مناسب را اختیار کرده باشیم. به عنوان مثال، به فرض اینکه بدانیم می‌توان (x, y) را تابعی مشتق‌پذیر از z در نظر گرفت (حول $(1, 0, 0)$ ، $(x, y) = f(z)$ ، روابط $F_1(f(z), z) = 0$ و $F_2(f(z), z) = 0$ برای z در نزدیکی $z = 0$ برقرار است، پس مشتق عبارتهای طرف چپ نسبت به z

صفر خواهد بود. با استفاده از قاعده زنجیره‌ای داریم:

$$\begin{cases} \frac{\partial F_1}{\partial x} \frac{dx}{dz} + \frac{\partial F_1}{\partial y} \frac{dy}{dz} + \frac{\partial F_1}{\partial z} \frac{dz}{dz} = 0 \\ \frac{\partial F_2}{\partial x} \frac{dx}{dz} + \frac{\partial F_2}{\partial y} \frac{dy}{dz} + \frac{\partial F_2}{\partial z} \frac{dz}{dz} = 0 \end{cases}$$

$$\begin{cases} (3x^2y - z^3) \frac{dx}{dz} + (x^2 + 4yz^2) \frac{dy}{dz} + (4y^2z - 3xz^2) = 0 \\ -\frac{dx}{dz} + e^{y+z} \frac{dy}{dz} + e^{y+z} = 0 \end{cases}$$

در نقطه $(1, 0, 0)$ داریم

$$\begin{cases} \frac{dy}{dz} = 0 \\ -\frac{dx}{dz} + \frac{dy}{dz} + 1 = 0 \end{cases}$$

که مجدداً نتیجه می‌دهد $\frac{dy}{dz} = 0$ و $\frac{dx}{dz} = 1$

اگر فرض می‌کردیم (x, z) را می‌توان به عنوان تابعی مشتق‌پذیر از y حول $(1, 0, 0)$ در نظر گرفت

و محاسبه مشابهی انجام می‌دادیم حاصل می‌شد:

$$\begin{cases} \frac{\partial F_1}{\partial x} \frac{dx}{dy} + \frac{\partial F_1}{\partial y} \frac{dy}{dy} + \frac{\partial F_1}{\partial z} \frac{dz}{dy} = 0 \\ \frac{\partial F_2}{\partial x} \frac{dx}{dy} + \frac{\partial F_2}{\partial y} \frac{dy}{dy} + \frac{\partial F_2}{\partial z} \frac{dz}{dy} = 0 \end{cases}$$

و در نقطه $(1, 0, 0)$:

$$\begin{cases} 1 = 0 \\ -\frac{dx}{dy} + 1 + \frac{dz}{dy} = 0 \end{cases}$$

این نتیجه متناقض گویای این واقعیت است که نمی‌توان (x, z) را تابعی مشتق‌پذیر از y حول $(1, 0, 0)$

انگاشت. در واقع با توجه به محاسبه پیشین که نتیجه داد $\frac{dy}{dz} = 0$ ، اگر مجموعه جواب حول $(1, 0, 0)$

را روی صفحه (y, z) تصویر کنیم، خمی مشاهده خواهد شد که مماس بر آن محور z مماس است.

بدین ترتیب اگر (x, z) به صورت تابعی از y نوشته شود، مشتق z نسبت به y تعریف شدنی نیست.

مثال ۲. در ترمودینامیک از رابطه معروف $PV - kT = 0$ (P : فشار، V : حجم، T : دمای کلوین)

رابطه زیر نوشته می شود:

$$\left(\frac{\partial P}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial T}{\partial P}\right)_V = -1 \quad (7)$$

مقصود از $\left(\frac{\partial P}{\partial V}\right)_T$ این است که از رابطه $P, PV - kT = 0$ به عنوان تابعی از (V, T) در نظر گرفته شده و سپس مشتق آن نسبت به V با ثابت نگاهداشتن T محاسبه شده است، به همین ترتیب در مورد دو عبارت دیگر. صحت (7) به رابطه خاص $PV - kT = 0$ وابسته نیست، بلکه هرگاه $F(x, y, z) = 0$ داده شده باشد که در آن مشتق های پاره ای F نسبت به سه متغیر وجود داشته و پیوسته باشند، در هر نقطه (x, y, z) که سه مشتق پاره ای $\frac{\partial F}{\partial x}$ ، $\frac{\partial F}{\partial y}$ و $\frac{\partial F}{\partial z}$ ناصفر باشند، می توان نوشت:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \quad (8)$$

این رابطه را نیز می توان به n متغیر تعمیم داد. فرض کنید $F(x_1, \dots, x_n) = 0$ مشتق های پاره ای F نسبت به x_1, \dots, x_n همه وجود داشته و پیوسته باشند، و به ازای مقداری از (x_1, \dots, x_n) داشته باشیم $\frac{\partial F}{\partial x_j} \neq 0$ برای هر $j = 1, \dots, n$. در این صورت طبق قضیه تابع ضمنی، می توان حول این نقطه، هر x_j را تابع مشتق پذیر از سایر متغیرها فرض کرد و طبق (3) داریم:

$$\frac{\partial x_j}{\partial x_{j+1}} = -\frac{\frac{\partial F}{\partial x_{j+1}}}{\frac{\partial F}{\partial x_j}} \quad (9)$$

(برای $j = n$ ، به جای $j + 1$ در دو مخرج، بنویسید 1). با ضرب کردن n رابطه (9) به ازای $j = 1, \dots, n$ حاصل می شود:

$$\left(\frac{\partial x_1}{\partial x_2}\right) \left(\frac{\partial x_2}{\partial x_3}\right) \dots \left(\frac{\partial x_n}{\partial x_1}\right) = (-1)^n \quad (10)$$

(برای $n = 2$ ، رابطه $\frac{dy}{dx} \frac{dx}{dy} = 1$ به دست می آید.) البته در اینجا نیز مقصود از $\left(\frac{\partial x_1}{\partial x_2}\right)$ ، به طور مبسوط تر، $\left(\frac{\partial x_1}{\partial x_2}\right)_{x_3, \dots, x_n}$ است، یعنی x_1 به عنوان تابعی مشتق پذیر از (x_2, \dots, x_n) در نظر گرفته شده و سپس مشتق گیری نسبت به x_2 با ثابت نگاهداشتن x_3, \dots, x_n منظور شده است.

قضیه تابع وارون، تعویض مختصات

فرض کنید روابط زیر داده شده‌اند که در آن هر f_i تابعی دارای مشتق‌های پاره‌ای پیوسته است:

$$\begin{cases} y_1 = f_1(x_1, \dots, x_n) \\ \vdots \\ y_n = f_n(x_1, \dots, x_n) \end{cases} \quad (1)$$

در اینجا n کمیت y_1, \dots, y_n برحسب کمیت‌های x_1, \dots, x_n داده شده‌اند. سؤالی طبیعی که تکرراً در ریاضیات و کاربردهای آن مطرح می‌شود این است که در چه شرایطی می‌توان از کمیت‌های y_1, \dots, y_n به جای x_1, \dots, x_n به عنوان متغیرهای مسأله خاصی که برحسب x_1, \dots, x_n بیان شده است استفاده کرد. در واقع می‌توان این سؤال را به عنوان "تعویض متغیر" یا "تعویض مختصات" مطرح کرد. از سویی دیگر می‌توان (۱) را بیان یک تابع $f = (f_1, \dots, f_n)$ از دامنه مشترک f_i ها در \mathbb{R}^n به \mathbb{R}^n فرض کرد. تابع وارون یا تابع معکوس f ، تابعی g با این ویژگی است که هرگاه $y = f(x)$ آنگاه $x = g(y)$ و بالعکس (شکل ۱). بدین ترتیب f و g تناظری یک به یک میان دو زیرمجموعه \mathbb{R}^n برقرار می‌کنند و می‌توان نوشت:

$$\begin{cases} x_1 = g_1(y_1, \dots, y_n) \\ \vdots \\ x_n = g_n(y_1, \dots, y_n) \end{cases} \quad (2)$$

شکل ۱

اگر g_1, \dots, g_n تابع‌های دارای مشتق‌های پاره‌ای پیوسته باشند، جایگزینی y_1, \dots, y_n به جای x_1, \dots, x_n را توسط (۲) تعویض مختصات مجاز از x به y می‌نامیم. بدین ترتیب اگر

$Z = \varphi(x_1, \dots, x_n)$ یک تابع مشتق پذیر از x_1, \dots, x_n باشد، می توان نوشت $Z = \varphi(g(x_1, \dots, x_n))$

و بنا بر قاعده زنجیره ای، $\varphi \circ g$ کمیت Z را به عنوان تابعی مشتق پذیر از y_1, \dots, y_n بیان می کند.

در مورد تابع های یک متغیری، $n = 1$ دیده ایم که شرطی لازم و کافی برای وجود وارون برای تابعی $f: I \rightarrow \mathbb{R}$ یک بازه باز در \mathbb{R} ، این است که f یکنوا (صعودی یا نزولی) باشد. در این وضعیت، اگر f مشتق پذیر با مشتق همواره مثبت (یا همواره منفی) باشد، تابع وارون، g ، نیز مشتق پذیر خواهد شد و از رابطه $g(f(x)) = x$ بنا بر قاعده زنجیره ای نتیجه می گیریم که:

$$g'(x) = \frac{1}{f'(g(x))} \quad \text{اگر } y = f(x) \quad (3)$$

البته شرط صعودی یا نزولی بودن یک تابع در سراسر دامنه به ندرت حاصل می شود. از این رو معمولاً به مفهوم "وارون موضعی" متوسل می شویم که به صورت زیر است. دامنه تابع داده شده f را به بازه ای کوچکتر محدود می کنیم به طوری که در این بازه f صعودی یا نزولی باشد. در این صورت برای تحدید تابع به این زیربازه وارونی وجود خواهد داشت که یک وارون موضعی f خوانده می شود. در این صورت تحدید f به یک زیربازه باز از دامنه اولیه، این بازه را به طور یک به یک به یک بازه باز می نگارد.

مثالی باز تابع $f: \mathbb{R} \rightarrow \mathbb{R}$ ، $f(x) = \sin x$ است. چنانچه دامنه f را بازه $[-\frac{\pi}{2}, \frac{\pi}{2}]$ محدود کنیم، f در این بازه صعودی است و وارون آن معمولاً به \sin^{-1} یا Arcsin نمایش داده می شود. بدین ترتیب $\sin^{-1}(y)$ برای $-1 \leq y \leq 1$ ، یگانه x در بازه $[-\frac{\pi}{2}, \frac{\pi}{2}]$ است که $\sin x = y$. اگر به جای بازه بسته $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ، از بازه باز $]-\frac{\pi}{2}, \frac{\pi}{2}[$ استفاده کنیم، مشتق f در این بازه اکیداً مثبت است، پس \sin^{-1} در تصویر این بازه، یعنی $] -1, 1 [$ مشتق پذیر است و

$$\begin{aligned} (\sin^{-1})'(y) &= \frac{1}{\sin' x} = \frac{1}{\cos x} \\ &= \frac{1}{\sqrt{1-y^2}} \end{aligned}$$

(فقط جذر مثبت $\cos^2 x = (1 - \sin^2 x)$ در نظر گرفته شده است زیرا که $\cos x$ برای $-\frac{\pi}{2} < x < \frac{\pi}{2}$ مثبت است).

هدف ما اکنون این است که موضوع "وارون موضعی" را به توابع چند متغیری تعمیم دهیم. قضیه زیر تعمیم مورد نظر است:

(۲۹-۱) قضیه تابع وارون دستگاه (۱) را در نظر بگیرید و فرض کنید هر f_i دارای مشتق‌های پاره‌ای مرتبه اول پیوسته است. فرض کنید نقطه $a = (a_1, \dots, a_n)$ یک نقطه درونی دامنه $f = (f_1, \dots, f_n)$ باشد که در آن $\det(Df(a)) \neq 0$. در این صورت زیرمجموعه‌ی بازی U از دامنه f حول a وجود دارد که $V = f(U)$ یک مجموعه‌ی باز است و تحدید f به U تناظری یک به یک میان U و V ایجاد می‌کند. به علاوه اگر $g: U \rightarrow V$ وارون موضعی f باشد، دارای مشتق‌های پاره‌ای مرتبه اول پیوسته است. این قضیه را می‌توان به صورت حالت خاصی از قضیه تابع ضمنی به صورت زیر نتیجه گرفت، ولی شایان ذکر است که معمولاً نخست این قضیه به اثبات می‌رسد و سپس می‌توان قضیه تابع ضمنی را از این قضیه نتیجه گرفت. برای اثبات، (۱) را به صورت زیر می‌نویسیم:

$$\begin{cases} -y_1 + f_1(x_1, \dots, x_n) = 0 \\ \vdots \\ -y_n + f_n(x_1, \dots, x_n) = 0 \end{cases} \quad (4)$$

و قرار می‌دهیم $b = (b_1, \dots, b_n) = f(a)$. نقطه $(y, x) = (b, a)$ در (۴) صدق می‌کند. اگر به جای $-y_i + f_i(x_1, \dots, x_n)$ بنویسیم $F_i(y_1, \dots, y_n, x_1, \dots, x_n)$ ، نتیجه می‌شود که دترمینان ماتریس $[\frac{\partial F_i}{\partial x_j}(b, a)]$ ناصفر است. پس بنا بر قضیه تابع ضمنی می‌توان برای (y, x) های نزدیک (b, a) که در (۴) صدق می‌کنند، x را به صورت تابعی از y استخراج کرد، یا $x = g(y)$ ، یا

$$\begin{cases} x_1 = g_1(y_1, \dots, y_n) \\ \vdots \\ x_n = g_n(y_1, \dots, y_n) \end{cases} \quad (5)$$

که در آن g_i ها تابع‌های دارای مشتق‌های پاره‌ای مرتبه اول پیوسته هستند. بدین ترتیب یک وارون موضعی برای f حول a به دست می‌آید.

مضافاً توجه کنید که با مشتق‌گیری از ترکیب $g(f(x)) = x$ و $f(g(y)) = y$ روابطی مشابه (۳)

به دست می آیند:

$$Dg(y) \cdot Df(x) = I \quad , \quad Df(x) \cdot Dg(y) = I \quad (6)$$

که در اینجا ماتریس واحد $n \times n$ یعنی $\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ است. پس ماتریس مشتق‌های پاره‌ای y_i ها نسبت به x_j ها، وارون ماتریس مشتق‌های پاره‌ای x_i ها نسبت به y_j هاست:

$$\left[\frac{\partial x_i}{\partial y_j}(y) \right] = \left[\frac{\partial y_i}{\partial x_j}(x) \right]^{-1} \quad (7)$$

در زیر چند نوع تعویض مختصات معروف را در این چارچوب بررسی می‌کنیم.

مثال ۱ (مختصات قطبی (x, y)) می‌نویسیم

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \quad (8)$$

داریم

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}$$

پس $\det\left(\frac{\partial(x, y)}{\partial(r, \theta)}\right) = r$. طبق قضیه تابع وارون اگر (r_0, θ_0) طوری باشد که $r_0 \neq 0$ ، آنگاه (۸) یک تعویض مختصات مجاز از یک مجموعه باز U حول (r_0, θ_0) به مجموعه باز V حول $(x_0, y_0) = (r_0 \cos \theta_0, r_0 \sin \theta_0)$ ایجاد می‌کند ولی می‌توان نوشت:

$$(x, y) \in V \text{ و } (r, \theta) \in U \text{ برای } \begin{cases} r = g_1(x, y) \\ \theta = g_2(x, y) \end{cases} \quad (9)$$

در مورد بزرگی اندازه U و V توجه کنید که نباید $r = 0$ را شامل شود زیرا که به ازای $r = 0$ یک به یک بودن تابع تعریف شده در (۸) نقض می‌شود. همچنین U نباید شامل دو نقطه (r, θ_1) و (r, θ_2) باشد که $\theta_2 - \theta_1 = 2\pi$ است زیرا که مجدداً یک به یک بودن (۸) نقض خواهد شد.

شکل ۲

مثال ۲ (مختصات استوانه‌ای در \mathbb{R}^3) این در واقع جایگزینی (x, y) در مختصات دکارتی فضای

سه‌بعدی (x, y, z) با مختصات قطبی است، یعنی

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = z \end{cases} \quad (10)$$

سه‌تایی (r, θ, z) را مختصات استوانه‌ای نقطه (x, y, z) می‌نامیم. داریم

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \begin{bmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

و $\det \frac{\partial(x, y, z)}{\partial(r, \theta, z)} = r$. در اینجا ملاحظاتی مشابه مثال ۱ برقرار است.

مثال ۳ (مختصات کروی) برای نقطه (x, y, z) در \mathbb{R}^3 ، ρ را برابر فاصله نقطه از z ، θ را به مفهوم

مختصات قطبی، و φ را زاویه از نیمه بالای محور z به نیمخط واصل از z به (x, y, z) می‌گیریم.

شکل ۳

بدین ترتیب:

$$\begin{cases} x = \rho \sin \varphi \cos \theta \\ y = \rho \sin \varphi \sin \theta \\ z = \rho \cos \varphi \end{cases} \quad (11)$$

و داریم:

$$\frac{\partial(x, y, z)}{\partial(\rho, \varphi, \theta)} = \begin{bmatrix} \sin \varphi \cos \theta & \rho \cos \varphi \cos \theta & -\rho \sin \varphi \sin \theta \\ \sin \varphi \sin \theta & \rho \cos \varphi \sin \theta & \rho \sin \varphi \cos \theta \\ \cos \varphi & -\rho \sin \varphi & 0 \end{bmatrix}$$

و محاسبه دترمینان این ماتریس نشان می‌دهد که $\det = \rho^2 \sin \varphi$. این دترمینان دقیقاً در نقاط خارج

از محور z ناصفر است، پس حول این نقاط می‌توان یک تعویض مختصات موضعی مجاز به مختصات

کروی انجام داد. در واقع اگر محدودیت‌های $0 \leq \varphi \leq \pi$ ، $0 \leq \theta \leq 2\pi$ و $\rho > 0$ را اعمال کنیم،

دستگاه (۱۱) یک تابع یک به یک از (ρ, φ, θ) به (x, y, z) تعریف می‌کند. اگر مشتق‌پذیری تابع

وارون مطرح نباشد می‌توان از این تعویض مختصات خارج از z استفاده کرد.

بهینه‌سازی (۱)

مسائل بهینه‌سازی چند متغیری بدین صورت ارائه می‌شوند که تابعی $f : S \rightarrow \mathbb{R}$ ، S زیرمجموعه‌ای از \mathbb{R}^n ، داده شده است، و هدف یافتن ماکسیمم یا مینیمم مطلق f در مجموعه S است. روش بخش ۲۷ تا حدودی جوابگوی این مسأله است به این مفهوم که اگر ماکسیمم یا مینیمم مورد نظر در یک نقطه درونی مجموعه S ظاهر شود و اگر f در این نقطه مشتق‌پذیر باشد، این نقطه یک نقطه بحرانی تابع خواهد بود. روشن است که برای حل کامل مسأله، علاوه بر در نظر گرفتن نقاط بحرانی، باید نقاط نوع زیر نیز در نظر گرفته شوند تا با مقایسه مقدار f در این نقاط و نقاط بحرانی جواب نهایی مسأله معلوم گردد. این نقاط عبارتند از:

(۱) نقاط مرزی ناحیه تعریف تابع.

(۲) نقاط تکین، یعنی نقاطی که در آن f مشتق‌پذیر نیست.

اگر S یک مجموعه بسته (شامل کلیه نقاط مرزی) و کراندار باشد، می‌دانیم هر تابع پیوسته $f : S \rightarrow \mathbb{R}$ روی S ماکسیمم و مینیمم مطلق اتخاذ می‌کند، بنابراین اگر S بسته و کراندار و $f : S \rightarrow \mathbb{R}$ پیوسته باشد، مقایسه نتایج به دست آمده از نقاط بحرانی، نقاط مرزی و نقاط تکین قطعاً جواب مورد نظر را به دست خواهد داد. در حالتی که مجموعه S بی‌کران باشد، یا بعضی نقاط مرزی خود را شامل نشود، امکان دارد که ماکسیمم یا مینیمم در S موجود نباشد. برای این‌گونه مجموعه‌ها باید رفتار تابع $f(x)$ را وقتی x به یک نقطه مرزی خارج S نزدیک می‌شود و نیز وقتی $|x| \rightarrow \infty$ (در داخل مجموعه S ، وقتی S بی‌کران است) جداگانه بررسی کرد. امکان دارد در این موارد اصلاً ماکسیمم یا مینیمم وجود نداشته باشد. در زیر به ذکر دو مثال می‌پردازیم.

مثال ۱ ماکسیمم و مینیمم تابع $f(x, y) = x^2 + y^2 + x$ را در ناحیه $x^2 + y^2 \leq 1$ به دست آورید.

حل ناحیه داده شده گوی بسته شعاع ۱ حول مبدأ در \mathbb{R}^2 است که بسته و کراندار می باشد. تابع داده شده به ازای هر (x, y) مشتق پذیر است، پس نقطه تکین وجود ندارد. کافی است نقاط بحرانی و نقاط مرزی را بررسی و مقایسه کنیم. برای نقاط بحرانی:

$$\nabla f(x, y) = (2x + 1, 2y) = (0, 0)$$

نتیجه می دهد $(-\frac{1}{2}, 0)$ تنها نقطه بحرانی است، و $f(-\frac{1}{2}, 0) = -\frac{1}{4}$. نقاط مرزی دقیقاً آن (x, y) ها هستند که $x^2 + y^2 = 1$ ، پس در یک نقطه مرزی $f(x, y) = 1 + x$ و از آنجا که برای نقاط مرزی $-1 \leq x \leq 1$ ، نتیجه می گیریم که ماکسیمم f روی مرز در نقطه $(1, 0)$ و مینیمم آن در نقطه $(-1, 0)$ به دست می آید. داریم $f(1, 0) = 2$ و $f(-1, 0) = 0$. از مقایسه سه مقدار به دست آمده نتیجه می گیریم که ماکسیمم تابع برابر ۲ است و در نقطه مرزی $(1, 0)$ ظاهر می شود، در حالی که مینیمم آن، $-\frac{1}{4}$ ، در نقطه بحرانی درونی $(-\frac{1}{2}, 0)$ حاصل می شود.

مثال ۲ ماکسیمم و مینیمم تابع $f(x, y) = \frac{|xy|}{x^2 + y^2 + 1}$ را در \mathbb{R}^2 به دست آورید.

حل وجود قدر مطلق نشان می دهد تابع ممکن است وقتی $xy = 0$ مشتق پذیر نباشد. در هر حال به ازای $xy = 0$ ، یعنی روی اجتماع محوره های x, y ، مقدار تابع صفر است که قطعاً مینیمم تابع می باشد زیرا که به ازای $xy \neq 0$ مقدار تابع مثبت است. در خارج اجتماع محورها، کافی است مقدار تابع به ازای $x > 0, y > 0$ بررسی شود زیرا که به سبب قدر مطلق، تقارن نسبت به دو محور دیگر وجود دارد. در ربع اول داریم

$$\nabla f(x, y) = \left(\frac{y(y^2 + 1 - x^2)}{(x^2 + y^2 + 1)^2}, \frac{x(x^2 + 1 - y^2)}{(x^2 + y^2 + 1)^2} \right)$$

برای $x > 0$ و $y > 0$ ، $\nabla f(x, y) = 0$ جواب ندارد. زیرا که خم های $x^2 + 1 = y^2$ و $y^2 + 1 = x^2$ نقطه مشترک ندارند، پس نقطه بحرانی وجود ندارد. از سویی دیگر برای $x \neq 0$ و $y \neq 0$:

$$\frac{|xy|}{x^2 + y^2 + 1} = \frac{\frac{|xy|}{x^2 + y^2}}{1 + \frac{1}{x^2 + y^2}} < \frac{1}{2}$$

و در امتداد خط $y = x$ مقدار تابع به $\frac{1}{4}$ میل می کند وقتی $x \rightarrow +\infty$. نتیجه این که تابع ماکسیمم ندارد.

در مثال اول، مجموعه مرزی، دایره $x^2 + y^2 = 1$ بود و تابع مورد نظر $f(x, y) = x^2 + y^2 + x$ روی این مجموعه به شکل ساده $f(x, y) = 1 + x$ خلاصه می شد. در حالت کلی، یافتن ماکسیمم یا مینیمم روی مرز یک مجموعه می تواند دشوار باشد. مجموعه های مرزی که در عمل ظاهر می شوند معمولاً از یک یا چند قطعه هموار تشکیل شده اند. "روش لاگرانژ" که در زیر می آید حربه مؤثری برای یافتن نقاط ماکسیمم یا مینیمم موضعی تحدید یک تابع به یک مجموعه هموار (مثلاً مجموعه تراز غیر بحرانی) است. به طور خاص وضعیت زیر را در نظر بگیرید. فرض کنید S یک زیرمجموعه باز \mathbb{R}^n است، $g : U \rightarrow \mathbb{R}$ تابعی با مشتق های پاره ای مرتبه اول پیوسته، و M یک مجموعه تراز تابع g ، مثلاً

$$M = \{x \in U \mid g(x) = c\}$$

در مثال های مورد نظر ما گاهی M مرز یک ناحیه است (مثلاً مرز ناحیه ای که در آن $g(x) \geq c$ یا ناحیه ای که در آن $g(x) \leq c$). فرض کنید تابع $f : U \rightarrow \mathbb{R}$ نیز داده شده است که مشتق های پاره ای مرتبه اول پیوسته دارد. می خواهیم نقاط ماکسیمم و مینیمم موضعی تحدید f به M را شناسایی کنیم.

(۱۶-۱) قضیه لاگرانژ فرض کنید a یک نقطه ماکسیمم یا مینیمم موضعی تحدید f به M باشد و a یک نقطه بحرانی برای g نباشد، در این صورت $\nabla f(a)$ در نقطه a بر M عمود است: بدین ترتیب، چون a برای g بحرانی نیست، یعنی $\nabla g(a) \neq 0$ ، عددی حقیقی λ وجود دارد که:

$$\nabla f(a) = \lambda \nabla g(a)$$

توجه کنید که $\nabla g(a) \neq 0$ نشان می دهد که مجموعه تراز M حول a هموار است، پس عمود بودن $\nabla f(a)$ بر آن معنی دارد، این در واقع به معنی عمود بودن بر کلیه بردارهای مماس بر M در نقطه a است. فرض کنید v یک چنین بردار باشد. خمی هموار $\gamma : I \rightarrow \mathbb{R}^n$ در نظر می گیریم که تصویر آن به تمامی روی M قرار گیرد، $\gamma(t) \in M$ برای هر $t \in I$ و $\gamma(t_0) = a$ و $\gamma'(t_0) = v$ ، چون a یک نقطه مینیمم موضعی (یا ماکسیمم موضعی) برای تحدید f به M است $f(\gamma(t))$ به ازای $t = t_0$ دارای

شکل ۱

مینیم موضعی (یا ماکسیم موضعی) است، پس $\frac{d}{dt}((f \circ \gamma)(t))|_{t=t_0} = 0$. طبق قاعده زنجیره‌ای:

$$0 = \frac{d}{dt}((f \circ \gamma)(t))|_{t=t_0} = \nabla f(\gamma(t_0)) \cdot \gamma'(t_0) = \nabla f(a) \cdot v$$

یعنی $\nabla f(a) \perp v$ و قضیه لاگرانژ به اثبات می‌رسد.

قبل از ارائه مثال و تکمیل کردن صورت قضیه لاگرانژ که بعداً خواهد آمد، دو بحث شهودی در تأیید قضیه لاگرانژ بیان می‌کنیم:

(۱۶-۲) بحث شهودی ۱ محدود ساختن دامنه f به M و جستجو برای نقاط ماکسیم و مینیم آن را می‌توان به شکل یک آزمایش ذهنی بدین صورت مطرح کرد. موجوداتی را تصور می‌کنیم که چسبیده به M زندگی می‌کنند و از جهان خارج (یعنی \mathbb{R}^n) هیچ‌گونه ادراکی ندارند. برای این موجودات تابع f فقط روی M تعریف شده است. چون نقطه a یک نقطه غیربحرانی g است، M حول a از هموار بودن کافی برخوردار است و این موجودات ناحیه حول a را به شکل \mathbb{R}^{n-1} تجسم می‌کنند. اگر a یک نقطه ماکسیم یا مینیم موضعی برای تحدید f به M باشد، برای این موجودات a یک نقطه بحرانی f (محدود به جهان آنان) است و بنابراین باید "گرادیان f " (به مفهوم جهان M) از نظر این موجودات صفر شود. ولی "گرادیان f " از دیدگاه این موجودات با "گرادیان f " به مفهوم ما چه رابطه‌ای دارد؟ طبیعی به نظر می‌رسد که فرض کنیم گرادیان f برای موجودات محدود به M ، "سایه" یا تصویر قائم گرادیان f عادی روی M است. در این صورت صفر شدن گرادیان برای این موجودات معادل با عمود بودن آن بر M از نظر ما است.

(۱۶-۳) بحث شهودی ۲ یک مورد خاص را در نظر می‌گیریم که وضعیت کلی را به طرزی بارز نشان می‌دهد. این مثال را می‌توان در هر بعد ارائه کرد ولی برای سهولت تجسم آن را در \mathbb{R}^2 یا \mathbb{R}^3 تصور کنید. شکل ۲ مربوط \mathbb{R}^2 است. فرض کنید هدف یافتن مینیمم (مجذور) فاصله نقطه $p = (p_1, \dots, p_n)$ به مجموعه تراز هموار $M = g^c$ است که p خارج آن قرار دارد. برای این کار بالنی کروی به مرکز p تجسم می‌کنیم که تدریجاً شعاع آن افزایش می‌یابد. اولین برخورد سطح این بالن

با M در نقطه حداقل فاصله M از p صورت می‌گیرد. در این نقطه سطح بالن و M بر هم مماس می‌شوند پس قائم‌های این دو در یک راستا قرار می‌گیرند. جهت قائم بر M در نقطه تماس a به وسیله $\nabla g(a)$ مشخص می‌شود. تابعی که یافتن مینیمم آن مطرح است، مجذور فاصله از p ، یعنی $f(x) = \sum_{i=1}^n (x_i - p_i)^2$ می‌باشد که گرادیان آن $\nabla f(x) = 2(x_1 - p_1, \dots, x_n - p_n)$ در امتداد شعاع حامل از p به نقطه x بر سطح بالن عمود است (در واقع سطح بالن در لحظات مختلف مجموعه‌های تراز گوناگون f است). بدین ترتیب همراستا بودن $\nabla f(a)$ و $\nabla g(a)$ توجیه می‌شود.

شکل ۲

اکنون مثال‌هایی از نحوه به کار گرفتن قضیه لاگرانژ در حل مسایل بهینه‌سازی ارائه می‌کنیم. در هر مسأله نقاط a را که در رابطه $\nabla f(a) = \lambda \nabla g(a)$ صدق می‌کنند پیدا کرده و با مقایسه مقدار f در این نقاط به جواب مسأله می‌رسیم. توجه کنید که اگر مسأله در \mathbb{R}^n باشد، رابطه $\nabla f(a) = \lambda \nabla g(a)$ در واقع n معادله $(n+1)$ - مجهولی به دست می‌دهد (λ نیز نامعلوم است) ولی رابطه $g(x) = c$ (مجموعه تراز) که نقطه مورد نظر روی آن قرار دارد) نیز برقرار است که با انضمام آن می‌توان امید داشت از $(n+1)$ معادله $(n+1)$ مجهولی به جواب برسیم.

مثال ۱ نزدیکترین و دورترین نقاط بیضی وار $1 = \frac{x^2}{4} + y^2 + \frac{z^2}{3}$ را از نقطه $(-1, 0, 0)$ پیدا کنید. مجذور فاصله، یعنی $f(x, y, z) = (x+1)^2 + y^2 + z^2$ را در نظر می‌گیریم. هدف یافتن مینیمم و ماکسیمم این تابع برای (x, y, z) محدود به سطح بیضی وار است. اگر قرار دهیم $g(x, y, z) = \frac{x^2}{4} + y^2 + \frac{z^2}{3}$ ، هدف یافتن ماکسیمم و مینیمم تحدید f به مجموعه تراز $g(x, y, z) = 1$ است. داریم $\nabla g(x, y, z) = (\frac{x}{2}, 2y, \frac{2}{3}z)$ پس در هیچ نقطه مجموعه تراز $g(x, y, z) = 1$ صفر نمی‌شود و روش لاگرانژ باید جواب دهد. رابطه $\nabla f = \lambda \nabla g$ به صورت زیر در می‌آید:

$$\begin{cases} 2(x+1) = \lambda(\frac{x}{2}) \\ 2y = \lambda(2y) \\ 2z = \lambda(\frac{2}{3}z) \end{cases}$$

رابطه دوم نشان می‌دهد $y = 0$ یا $\lambda = 1$ که این دو حالت را جداگانه بررسی می‌کنیم.

حالت $y = 0$ از معادله سوم نتیجه می شود که $z = 0$ یا $\lambda = 2$. اگر $z = 0$ ، علاوه بر $y = 0$ ، داریم $x = \pm 2$ و دو نقطه $(\pm 2, 0, 0)$ برای بررسی مطرح می شوند. اگر $\lambda = 2$ ، معادله اول می دهد $2x + 2 = x$ یا $x = -2$ که در نتیجه $y = z = 0$ و همان نقطه $(-2, 0, 0)$ به دست می آید.

حالت $\lambda = 1$ در این صورت معادله سوم نتیجه می دهد $z = 0$ و معادله اول $2x + 2 = \frac{x}{3}$ منجر می شود به $x = -\frac{4}{3}$. چون $z = 0$ داریم $y = \pm \frac{\sqrt{5}}{3}$. در نتیجه دو نقطه $(-\frac{4}{3}, \pm \frac{\sqrt{5}}{3}, 0)$ مطرح می گردند.

در مجموع به چهار نقطه $(\pm 2, 0, 0)$ و $(-\frac{4}{3}, \pm \frac{\sqrt{5}}{3}, 0)$ دست یافته ایم که با مقایسه مقدار f در این چهار نقطه باید جواب مورد نظر به دست آید:

$$f(2, 0, 0) = 9 \quad , \quad f(-2, 0, 0) = 1 \quad , \quad f(-\frac{4}{3}, \pm \frac{\sqrt{5}}{3}, 0) = \frac{2}{3}$$

نتیجه این که $(2, 0, 0)$ دورترین نقطه بیضی وار، و $(-\frac{4}{3}, \pm \frac{\sqrt{5}}{3}, 0)$ نزدیکترین نقطه های بیضی وار به نقطه $(-1, 0, 0)$ هستند.

مثال ۲ فرض کنید $K, A, B, C, D, \alpha, \beta, \gamma$ داده های ثابت مثبت هستند. می خواهیم ماکسیمم تابع $f(x, y, z) = Kx^\alpha y^\beta z^\gamma$ را مشروط به این که $Ax + By + Cz = D$ ، $x \geq 0$ ، $y \geq 0$ و $z \geq 0$ بررسی کنیم. این نوع بهینه سازی در بعضی مسایل اقتصادی مطرح می شود. در این تعبیر x ، y و z عوامل تولید هستند (مثلاً نیروی کار، ماده اولیه و سرمایه گذاری در ماشین آلات) و A ، B و C هزینه واحد این عوامل. D مقدار ثابتی است که مجموع هزینه گذاری را نمایش می دهد. f تابع تولید برحسب عوامل x ، y ، z است. هدف، انتخاب نسبت مناسب از x ، y و z با قید $Ax + By + Cz = D$ است به نحوی که میزان تولید حداکثر ممکن باشد. از نظر هندسی باید ماکسیمم تابع $f(x, y, z) = Kx^\alpha y^\beta z^\gamma$ روی مثلث واقع بر صفحه $Ax + By + Cz = D$ محصور در ثمن اول ($x \geq 0$ ، $y \geq 0$ ، $z \geq 0$) بررسی شود. توجه کنید که روی اضلاع مثلث، یکی از سه متغیر x ، y یا z صفر می شود و در نتیجه مقدار f صفر می شود، بنابراین جواب باید یک نقطه درونی مثلث باشد.

شکل ۳

ضمناً با نوشتن $g(x, y, z) = Ax + By + Cz$ می‌بینیم که $\nabla g(x, y, z) = (A, B, C)$ همه‌جا ناصفر است و روش لاگرانژ باید جواب مورد نظر را به دست دهد. رابطه $\nabla f = \lambda \nabla g$ به صورت مؤلفه به مؤلفه هست:

$$\begin{cases} \alpha K x^{\alpha-1} y^{\beta} z^{\gamma} = \lambda A \\ \beta K x^{\alpha} y^{\beta-1} z^{\gamma} = \lambda B \\ \gamma K x^{\alpha} y^{\beta} z^{\gamma-1} = \lambda C \end{cases}$$

چون برای جواب مطلوب $x \neq 0$ ، $y \neq 0$ و $z \neq 0$ می‌توانیم سه رابطه را به ترتیب در x ، y ، و z ضرب کرده نتیجه بگیریم:

$$\begin{cases} \alpha f(x, y, z) = \lambda Ax \\ \beta f(x, y, z) = \lambda By \\ \gamma f(x, y, z) = \lambda Cz \end{cases}$$

بنابراین

$$\frac{A}{\alpha} x = \frac{B}{\beta} y = \frac{C}{\gamma} z$$

که با جایگزین در $Ax + By + Cz = D$ به جواب منحصر به فردی می‌رسد. همان طور که بحث شد این جواب یکتا باید ماکسیمم باشد. رابطه بالا نشان می‌دهد سهم مناسب x ، y و z با ماکسیمم کردن تابع تولید نسبت $(x : y : z) = (\frac{\alpha}{A} : \frac{\beta}{B} : \frac{\gamma}{C})$ است.

بالاخره مثال ساده‌ی زیر نشان می‌دهد که رعایت بررسی جداگانه نقاط احتمالاً ناهموار، یعنی نقاطی که در آن ∇g صفر می‌شود کاملاً ضروری است.

مثال ۳ نزدیکترین نقطه $x^3 - y^2 = 0$ به نقطه $(-1, 0)$ را پیدا کنید.

مجدور فاصله $(-1, 0)$ از نقطه (x, y) هست $f(x, y) = (x+1)^2 + y^2$. برای $g(x, y) = y^2 - x^3$ ،

رابطه $\nabla f = \lambda \nabla g$ به صورت زیر در می‌آید:

$$\begin{cases} 2(x+1) = \lambda(-3x^2) \\ 2y = \lambda(2y) \end{cases}$$

جواب بدیهی مسأله، نقطه $(0, 0)$ است که در رابطه اول بالا صدق نمی‌کند. توجه کنید که $\nabla g(0, 0) = \underline{0}$ و خم $y^2 - x^3 = 0$ در $(0, 0)$ هموار نیست.

شکل ۴

بهینه‌سازی (۲)

در این جلسه نخست به تشریح نقش فرض $\nabla g(a) \neq 0$ در قضیه لاگرانژ می‌پردازیم. همان طور که مثال آخر جلسه قبل نشان داد وقتی $\nabla g(a) = 0$ ، ممکن است مجموعه تراز تابع g که از a می‌گذرد در نقطه a هموار نباشد و در این صورت مماس وقائم در آن نقطه فاقد معنی می‌شوند. یاد آوری می‌کنیم که "هموار بودن" مجموعه تراز $M = g^c$ در نقطه a را بدین معنی گرفتیم که بتوانیم بخشی از مجموعه M حول a را به صورت نمودار یک تابع مشتق‌پذیر $(n-1)$ متغیری بنویسیم. حال $\nabla g(a) = (\frac{\partial g}{\partial x_1}(a), \dots, \frac{\partial g}{\partial x_n}(a)) \neq 0$ و شرط $\nabla g(a) \neq 0$ معادل این است که دست کم یکی از $\frac{\partial g}{\partial x_i}(a)$ ها صفر نباشد. مثلاً فرض کنید $\frac{\partial g}{\partial x_n}(a) \neq 0$. در این صورت بنا بر قضیه تابع ضمنی نقاط (x_1, \dots, x_n) از مجموعه تراز $M = g^c$ را که نزدیک a هستند می‌توان به صورت نمودار تابعی مشتق‌پذیر $x_n = \varphi(x_1, \dots, x_{n-1})$ نمایش داد و تحدید تابع f به M را می‌توان به صورت تابعی از x_1, \dots, x_{n-1} نوشت:

$$f(x_1, \dots, x_n) = f(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$$

که عبارت سمت راست را به $h(x_1, \dots, x_{n-1})$ نمایش می‌دهیم. حال اگر a یک نقطه ماکسیمم یا مینیمم موضعی تحدید f به M باشد، $(a_1, \dots, a_{n-1}) = a'$ یک نقطه بحرانی برای h است، و در واقع در هر نقطه بحرانی a' از h داریم $\nabla h(a') = 0$ ، یا $\frac{\partial h}{\partial x_i}(a') = 0$ برای هر i . طبق قاعده زنجیره‌ای داریم:

$$\frac{\partial h}{\partial x_i} = \sum_{j=1}^{n-1} \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial x_i} + \frac{\partial f}{\partial x_n} \frac{\partial \varphi}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n} \frac{\partial \varphi}{\partial x_i}$$

پس در نقطه بحرانی a :

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n} \frac{\partial \varphi}{\partial x_i} = 0, \quad i = 1, \dots, n-1 \quad (1)$$

از طرفی دیگر چون مجموعه M در واقع مجموعه تراز منسوب به $\underline{0}$ تابع $\varphi(x_1, \dots, x_{n-1}) - x_n$ است، جهت عمود بر آن توسط گرادیان این تابع، یعنی بردار زیر داده می شود:

$$\left(\frac{\partial \varphi}{\partial x_1}, \dots, \frac{\partial \varphi}{\partial x_{n-1}}, -1 \right) \quad (2)$$

از مقایسه روابط (1) و (2) نتیجه می شود که:

$$\left(\frac{\partial \varphi}{\partial x_1} : \dots : \frac{\partial \varphi}{\partial x_{n-1}} : -1 \right) = \left(\frac{\partial f}{\partial x_1} : \dots : \frac{\partial f}{\partial x_{n-1}} : \frac{\partial f}{\partial x_n} \right)$$

پس در نقطه a ، $\nabla f(a)$ عمود بر M است. استدلال بالا در واقع اثبات دیگری از قضیه لاگرانژ است که در آن به صراحت نشان داده شد که هرگاه مجموعه تراز M در a هموار باشد و a یک نقطه بحرانی برای تحدید f به M ، آنگاه $\nabla f(a)$ بر مجموعه تراز M عمود است.

هدف بعدی ما تعمیم قضیه لاگرانژ به حالتی است که به جای $g(x_1, \dots, x_n) = c$ ، تعدادی رابطه از این نوع ملحوظ شده باشد. به طور کلی، به خصوص در کاربردها، اصطلاحات زیر معمول است. تابع f که ماکسیمم یا مینیمم تحدید آن به زیرمجموعه خاصی مطرح است تابع عینی، و شرط $g(x_1, \dots, x_n) = c$ (تعریف کننده مجموعه تراز $M = g^c$) را یک قید می نامند. تا اینجا فقط مسائل تک قیدی را مطرح کرده ایم و اکنون قصد داریم مسائل چند قیدی را نیز بررسی کنیم.

وضعیت زیر را در نظر می گیریم. تابع $f : U \rightarrow \mathbb{R}$ داده شده است که تابعی مشتق پذیر روی زیرمجموعه باز U از \mathbb{R}^n است. به علاوه تابع های $g_1, \dots, g_k : U \rightarrow \mathbb{R}$ نیز داده شده اند که دارای مشتق های پاره ای مرتبه اول پیوسته می باشند. می خواهیم ماکسیمم یا مینیمم تحدید تابع f به اشتراک مجموعه های تراز $M_1 : g_1(x_1, \dots, x_n) = c_1, \dots, M_k : g_k(x_1, \dots, x_n) = c_k$ را بیابیم. به بیان دیگر، مقصود یافتن ماکسیمم یا مینیمم f تحت قیود $g_1(x_1, \dots, x_n) = c_1, \dots, g_k(x_1, \dots, x_n) = c_k$ است. در حالت $k = 1$ دیدیم که اگر در نقطه مطلوب a ، $\nabla g(a) \neq \underline{0}$ ، آنگاه مجموعه تراز $M : g(x) = c$ حول نقطه a هموار است، و $\nabla f(a)$ عمود بر این مجموعه عمود خواهد بود. حال اگر اشتراک مجموعه های تراز حول a هموار باشد، آنگاه خواهیم دید که به دلیلی کاملاً مشابه حالت تک قیدی، $\nabla f(a)$ بر این اشتراک عمود است. بدین ترتیب دو سؤال زیر مطرح می شود:

(1) تحت چه شرایطی اشتراک مجموعه های تراز حول a هموار است؟

(۲) به شرط هموار بودن اشتراک مجموعه‌های تراز حول a ، چگونه می‌توان عمود بودن $\nabla f(a)$ بر این اشتراک را به صورت یک رابطه جبری قابل استفاده بیان کرد؟

نخست سؤال (۱) را بررسی می‌کنیم. اشتراک مجموعه‌های تراز، مجموعه (x_1, \dots, x_n) هایی است که در روابط زیر صدق می‌کنند:

$$\begin{cases} g_1(x_1, \dots, x_n) - c_1 = 0 \\ \vdots \\ g_k(x_1, \dots, x_n) - c_k = 0 \end{cases}$$

فرض کنید نقطه a در این روابط صدق کند. اگر بتوان از این روابط، حول a ، k متغیر را به صورت تابع‌های مشتق‌پذیر از $(n - k)$ متغیر باقیمانده نوشت، آنگاه بخشی از اشتراک $M_1 \cap \dots \cap M_k$ که حول a قرار دارد به صورت نمودار یک تابع مشتق‌پذیر $(n - k)$ متغیری ظاهر می‌شود و بدین معنی "هموار" است. طبق قضیه تابع ضمنی اگر یک زیرماتریس $k \times k$ از ماتریس $k \times n$ زیر، متشکل از k ستون، دارای دترمینان ناصفر باشد، آنگاه k متغیر مربوط به این k ستون را می‌توان به صورت تابع‌های مشتق‌پذیر نسبت به $(n - k)$ متغیر باقیمانده نوشت:

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1}(a) & \dots & \frac{\partial g_k}{\partial x_n}(a) \end{bmatrix} \quad (3)$$

توجه کنید که ردیف i -ام این ماتریس از مؤلفه‌های $\nabla g_i(a)$ تشکیل شده است. در حالت $k = 1$ ، که قبلاً بررسی شد، وجود یک ماتریس 1×1 با دترمینان مخالف صفر در واقع معادل این است که $\nabla g(a)$ صفر نباشد (یعنی حداقل یک مؤلفه ناصفر داشته باشد). برای $k > 1$ توجه کنید اگر یکی از $\nabla g_i(a)$ ها را بتوان به صورت "ترکیبی خطی" از سایر $\nabla g_i(a)$ ها نوشت، یعنی مثلاً:

$$\nabla g_k(a) = c_1 \nabla g_1(a) + \dots + c_{k-1} \nabla g_{k-1}(a)$$

که در آن c_1, \dots, c_{k-1} عدد حقیقی هستند، آنگاه برای هر زیرماتریس $k \times k$ از ماتریس بالا نیز، یک ردیف ترکیبی خطی از سایر ردیف‌ها می‌شود و در نتیجه دترمینان صفر خواهد شد. طبق قضایای جبر خطی، عکس موضوع فوق نیز صحیح است، یعنی شرطی لازم و کافی برای این که ماتریس (۳) دارای

زیر ماتریس $k \times k$ با دترمینان ناصفر باشد این است که نتوان هیچ‌یک از $\nabla g_i(a)$ ها را به صورت ترکیبی خطی از سایر $\nabla g_i(a)$ ها نوشت.

بدین ترتیب جایگزین مناسب برای شرط $\nabla g(a) \neq \underline{0}$ در حالتی که k قید موجود باشند این خواهد بود که $\{\nabla g_1(a), \dots, \nabla g_k(a)\}$ مستقل خطی باشد. این شرط تضمین می‌کند که بخشی از $M = M_1 \cap \dots \cap M_k$ که در نزدیکی a قرار دارد نمودار یک تابع $(n - k)$ متغیری مشتق پذیر است و به این مفهوم از "هموار بودن" برخوردار است.

حال به سؤال دوم می‌پردازیم. فرض کنید اشتراک $M = M_1 \cap \dots \cap M_k$ حول a هموار است و بردار w بر این اشتراک عمود می‌باشد. با توجه به ملاحظات بالا که M حول a نمودار یک تابع مشتق پذیر $(n - k)$ متغیری است و با $(n - k)$ پارامتر به طور مشتق پذیر توصیف می‌شود، می‌توان آن را یک شیء هموار $(n - k)$ بعدی تصور کرد که در \mathbb{R}^n قرار گرفته است. بنابراین جهت عمود بر M در نقطه a ، k بعدی می‌شود. توجه کنید که k بردار مستقل خطی $\nabla g_1(a), \dots, \nabla g_k(a)$ همه در راستای عمود بر M قرار دارند زیرا که $\nabla g_i(a)$ بر M_i عمود است و در نتیجه هر زیرمجموعه M از M_i نیز عمود می‌باشد. پس ترکیب‌های خطی $c_1 \nabla g_1(a) + \dots + c_k \nabla g_k(a)$ یک فضای k بعدی پدید می‌آورند که همان فضای k بعدی عمود بر M است. بدین ترتیب باید بتوان w را به صورت ترکیبی خطی از $\nabla g_1(a), \dots, \nabla g_k(a)$ نوشت. با توجه به این توضیحات، اکنون آماده‌ایم تعمیم قضیه لاگرانژ را ارائه کنیم:

(۱۷-۱) قضیه لاگرانژ (صورت کامل) U زیرمجموعه‌ای باز از \mathbb{R}^n است، $f : U \rightarrow \mathbb{R}$ تابعی مشتق پذیر، و $g_1, \dots, g_k : U \rightarrow \mathbb{R}$ تابع‌هایی با مشتق‌های پاره‌ای مرتبه اول پیوسته. فرض کنید برای اعداد حقیقی داده شده c_1, \dots, c_k ، مجموعه‌های M_1, \dots, M_k ، به ترتیب، مجموعه‌های تراز g_1, \dots, g_k منسوب به c_1, \dots, c_k باشند، یعنی $M_i = g_i^{c_i}$ ، و $M = M_1 \cap \dots \cap M_k$. اگر نقطه a در M یک نقطه ماکسیمم یا مینیمم موضعی تحدید تابع f به M باشد و اگر $\nabla g_1(a), \dots, \nabla g_k(a)$ مستقل خطی باشند، آنگاه اعداد حقیقی $\lambda_1, \dots, \lambda_k$ وجود دارند که:

$$\nabla f(a) = \lambda_1 \nabla g_1(a) + \dots + \lambda_k \nabla g_k(a) \quad (۴)$$

با توجه به مقدمات ذکر شده بیشتر اثبات این قضیه قبلاً آماده شده است، فقط باید ثابت کنیم که اگر a یک نقطهٔ ماکسیمم یا مینیمم موضعی برای تحدید f به M باشد، آنگاه $\nabla f(a)$ بر M عمود است. ولی اثبات این مطلب عیناً مانند حالت $k = 1$ است. اگر $\gamma: I \rightarrow \mathbb{R}^n$ خمی مشتق‌پذیر باشد که تصویر آن به تمامی در M قرار می‌گیرد و $\gamma(t_0) = a$ ، باید ثابت کنیم $\nabla f(a)$ بر $\gamma'(t_0)$ عمود است. برای این کار توجه می‌کنیم که $f \circ \gamma$ در t_0 دارای ماکسیمم یا مینیمم موضعی است، پس $(f \circ \gamma)'(t_0) = 0$ ، و طبق قاعدهٔ زنجیره‌ای $\nabla f(a) \cdot \gamma'(t_0) = 0$ و حکم به اثبات می‌رسد.

در زیر به ذکر دو مثال می‌پردازیم. توجه کنید که (۴) در واقع n معادله با $(n+k)$ مجهول $x_1, \dots, x_n, \lambda_1, \dots, \lambda_k$ به دست می‌دهد، و با کمک k قید داده شده مجموعاً $(n+k)$ معادله با $(n+k)$ مجهول در دست است.

مثال ۱ برای اشتراک استوانهٔ $x^2 + y^2 = 1$ و صفحهٔ $x + 2y + 3z = 1$ نقطه‌هایی که بزرگترین و کوچکترین z را دارند پیدا کنید.

حل تابع عینسی در اینجا $f(x, y, z) = z$ است و قیدها عبارتند از $x^2 + y^2 - 1 = 0$ و $x + 2y + 3z - 1 = 0$. می‌نویسیم $g_1(x, y, z) = x^2 + y^2 - 1$ و $g_2(x, y, z) = x + 2y + 3z - 1$. داریم $\nabla g_1(x, y, z) = (2x, 2y, 0)$ و $\nabla g_2(x, y, z) = (1, 2, 3)$. ∇g_2 هیچ‌گاه مضربی از ∇g_1 نیست زیرا که مؤلفه سوم ∇g_2 برابر ۳ است و مؤلفه سوم ∇g_1 برابر ۰. تنها به‌ازای $x = y = 0$ ∇g_1 مضرب (صفر) ∇g_2 می‌شود، ولی هیچ نقطهٔ به شکل $(0, 0, z)$ روی استوانهٔ داده شده قرار ندارد، پس شرط استقلال خطی در همهٔ نقاط اشتراک برقرار است. بدین ترتیب می‌توانیم از رابطهٔ $\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$ استفاده کنیم:

$$\begin{cases} 0 = \lambda_1(2x) + \lambda_2(1) \\ 0 = \lambda_1(2y) + \lambda_2(2) \\ 1 = \lambda_1(0) + \lambda_2(3) \end{cases}$$

از معادله سوم نتیجه می‌شود که $\lambda_2 = \frac{1}{3}$ و با جایگزین در معادله اول معلوم می‌شود که $\lambda_1 \neq 0$ ، پس $x = \frac{-1}{6\lambda_1}$ و $y = \frac{-1}{3\lambda_1}$. با جایگزین در $x^2 + y^2 = 1$ داریم $\lambda_1 = \pm \frac{\sqrt{5}}{6}$ ، پس با استفاده

از $x + 2y + 3z = 1$ ، دو نقطه $(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, \frac{1-\sqrt{5}}{3})$ و $(\frac{-1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}, \frac{1+\sqrt{5}}{3})$ به دست می آیند، که اولی بزرگترین z و دومی کوچکترین z ممکن را دارد.

مثال ۲ فرض کنید که $g_1(x, y, z) = 0$ و $g_2(x, y, z) = 0$ دو رویه هموار در \mathbb{R}^3 هستند که اشتراک تهی دارند، و نقطه P روی رویه اول و نقطه Q روی رویه دوم طوری هستند که پاره خط PQ کوتاهترین طول ممکن میان پاره خطهای واصل بین دو رویه را دارد. نشان دهید پاره خط PQ بر هر دو رویه عمود است.

حل هموار بودن دو رویه را بدین معنی می گیریم که تابع های g_1 و g_2 دارای مشتق های پاره ای مرتبه اول پیوسته اند، و ∇g_1 در سراسر $g_1(x, y, z) = 0$ و ∇g_2 در سراسر $g_2(x, y, z) = 0$ ناصفر هستند. اگر یک نقطه رویه اول را به (x, y, z) و یک نقطه رویه دوم را به (x', y', z') نمایش دهیم، تابعی که مینیمم آن مطرح است تابع f متغیری $f(x, y, z, x', y', z') = (x - x')^2 + (y - y')^2 + (z - z')^2$ می باشد. تابع های g_1 و g_2 را به صورت زیر بازنویسی می کنیم:

$$G_1(x, y, z, x', y', z') = g_1(x, y, z) \quad , \quad G_2(x, y, z, x', y', z') = g_2(x', y', z')$$

بدین ترتیب قیدهای مسأله $G_1(x, y, z, x', y', z') = 0$ و $G_2(x, y, z, x', y', z') = 0$ می باشند. توجه

کنید که

$$\nabla G_1(x, y, z, x', y', z') = \left(\frac{\partial g_1}{\partial x}, \frac{\partial g_1}{\partial y}, \frac{\partial g_1}{\partial z}, 0, 0, 0 \right)$$

$$\nabla G_2(x, y, z, x', y', z') = \left(0, 0, 0, \frac{\partial g_2}{\partial x'}, \frac{\partial g_2}{\partial y'}, \frac{\partial g_2}{\partial z'} \right)$$

که مستقل خطی هستند زیرا که $\nabla g_1 \neq 0$ و $\nabla g_2 \neq 0$ در سراسر رویه های داده شده فرض شده

است. رابطه لاگرانژ $\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$ به صورت زیر در می آید:

$$\left\{ \begin{array}{l} 2(x - x') = \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2(0) \\ 2(y - y') = \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2(0) \\ 2(z - z') = \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2(0) \\ -2(x - x') = \lambda_1(0) + \lambda_2 \frac{\partial g_2}{\partial x'} \\ -2(y - y') = \lambda_1(0) + \lambda_2 \frac{\partial g_2}{\partial y'} \\ -2(z - z') = \lambda_1(0) + \lambda_2 \frac{\partial g_2}{\partial z'} \end{array} \right.$$

اگر $P = (x, y, z)$ و $Q = (x', y', z')$ جواب مسأله باشند، که باید در این روابط صدق کنند، از سه رابطه اول داریم $2\vec{QP} = \lambda_1 \nabla g_1(P)$ ، پس \vec{PQ} بر $g_1(x, y, z) = 0$ در نقطه P عمود است. همین طور از سه رابطه دوم نتیجه می شود که $-2\vec{PQ} = \lambda_2 \nabla g_2(Q)$ ، یعنی \vec{PQ} در نقطه Q بر $g_2(x', y', z') = 0$ عمود است، و نتیجه مطلوب به دست می آید.

دو توضیح در مورد مثال بالا بجا است. یکی اینکه حکم مشابه برای اشکال مسطح، مثلاً دو خط متنافر در \mathbb{R}^3 ، را معمولاً در چارچوب هندسه اقلیدسی ثابت می کنیم، ولی حتی بیان مطلب برای اشیاء خمیده نیاز به مفهوم "شیء هموار" دارد که در اینجا به تعریفی از آن رسیده ایم. نکته دیگر این که برای حل این مسأله سه بعدی (دو رویه در \mathbb{R}^3) عملاً مطلب را به یک مسأله در \mathbb{R}^1 تبدیل کردیم و به سهولت به جواب رسیدیم. توجه کنید که $G_1(x, y, z, x', y', z') = 0$ و $G_2(x, y, z, x', y', z') = 0$ هر یک یک "شیء هموار ۵- بعدی" در \mathbb{R}^1 تعریف می کند که اشتراک آنها به سبب استقلال خطی ∇G_1 و ∇G_2 یک "شیء هموار ۴- بعدی" در \mathbb{R}^1 می شود. آنگاه مینیمم تحدید تابع f را به این شیء چهار بعدی مد نظر قرار دادیم.

انتگرال چندگانه

یکی از دو رکن اصلی حساب دیفرانسیل و انتگرال، فرایند انتگرال گیری است. یک تعبیر متداول از انتگرال $\int_a^b f$ وقتی f تابعی نامنفی و تعریف شده روی $[a, b]$ باشد، اندازه مساحت زیر نمودار تا محور x بین a و b است. در اینجا می خواهیم مفهوم انتگرال تابع های تعریف شده روی ناحیه های واقع در \mathbb{R}^n را بررسی کنیم. یک تفاوت فوری که به ذهن می رسد تنوع شکل ناحیه ها در حالت دو بعدی و ابعاد بالاتر است. در حالت یک بعدی دامنه تعریف تابع معمولاً یک بازه یا اجتماع چند بازه در نظر گرفته می شود و این در تقریباً همه موارد کفاف نیاز می کند. در حالت دو بعدی انواع ناحیه های محصور به پاره خط یا قطعه خم را می توان تصور کرد که به طور طبیعی در مسائل گوناگون ظاهر می شوند. مثلاً در شکل ۱، ناحیه داده شده ممکن است یک صفحه نازک با چگالی متغیر $\rho(x, y)$ باشد. برای محاسبه جرم این صفحه باید از $\rho(x, y)$ در سراسر ناحیه "انتگرال گیری" شود. برای مطرح ساختن مفهوم انتگرال در بعد ۲ به بالا، نخست مفهوم انتگرال را روی "مستطیل" که تعمیم طبیعی بازه است مطرح می کنیم، سپس به ناحیه های پیچیده تر می پردازیم.

شکل ۱

فرض کنید $2n$ عدد حقیقی $a_1, \dots, a_n, b_1, \dots, b_n$ داده شده اند که $a_1 \leq b_1, \dots, a_n \leq b_n$. مجموعه زیر یک مستطیل بسته در \mathbb{R}^n خوانده می شود:

$$R = \{(x_1, \dots, x_n) \mid a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\}$$

اصطلاح "بسته" ارجاع به علامت \leq است. وقتی به جای \leq ، نماد $<$ به کار برده شود، یک "مستطیل باز" تعریف می شود و در صورت اختلاط $<$ و \leq ، انواع مستطیل ها که شامل بخشی از مرز خود می شوند به دست می آیند. مقصود از یک افراز R ، تجزیه R به زیرمستطیل هایی از همین نوع به شکل دقیق زیر

است. برای هر مؤلفه، مثلاً مؤلفه i -ام، بازه $[a_i, b_i]$ را به صورت زیر افراز می‌کنیم:

$$a_i = a_{i_0} \leq a_{i_1} \leq \dots \leq a_{i_{p_i}} = b_i$$

p_i تعداد زیربازه‌هایی است که $[a_i, b_i]$ به آن تجزیه شده است. اگر این افراز برای هر مؤلفه انجام گیرد، تجزیه‌ای که برای R حاصل می‌شود مرکب از زیرمستطیل‌های به شکل زیر است:

$$a_{1,j_1} \leq x_i \leq a_{1,j_1+1}, \dots, a_{n,j_n} \leq x_n \leq a_{n,j_n+1}$$

در شکل ۲، یک نمونه از زیرمستطیل‌های تجزیه هاشور زده شده است.

شکل ۲

حجم n -بعدی مستطیل R را برابر حاصل ضرب طول اضلاع، یعنی $(b_1 - a_1) \times \dots \times (b_n - a_n)$ تعریف می‌کنیم. بدین ترتیب حجم R برابر مجموع حجم‌های زیرمستطیل‌های پدید آمده در افراز است. در حالت $n = 1$ ، حجم 1 -بعدی همان طول، در حالت $n = 2$ ، حجم دو بعدی برابر مساحت، و برای $n = 3$ ، همان مفهوم حجم معمولی است.

حال فرض کنید تابعی کراندار $f: R \rightarrow \mathbb{R}$ داده شده است، می‌خواهیم مفهومی $\int_R f$ تعریف کنیم که موارد استفاده‌ای مشابه انتگرال روی یک بازه $[a, b]$ را داشته باشد. بدین ترتیب در حالت دو بعدی، $n = 2$ ، اگر $f(x) \geq 0$ برای هر x در R ، آنگاه $\int_R f$ باید بیانگر "حجم" زیر نمودار $f: R \rightarrow \mathbb{R}$ باشد. همچنین برای $n = 2$ و $n = 3$ ، اگر f در هر نقطه R ، چگالی منسوب به آن نقطه را بیان کند، $\int_R f$ باید جرم کل R باشد. به این منظور به روشی مشابه حالت یک بعدی عمل می‌کنیم. فرض کنید افرازی از R داده شده باشد و زیرمستطیل‌های افراز را R_1, \dots, R_p نامگذاری کنید. در هر نقطه‌ای Q_i به دلخواه انتخاب کنید و مجموع زیر را در نظر بگیرید:

$$\sum_{i=1}^p \text{vol}(R_i) \cdot f(Q_i) \quad (1)$$

که در اینجا مقصود از $\text{vol}(R_i)$ حجم (n -بعدی) R_i است. مجموع‌های به شکل (۱) را مجموع ریمان می‌نامند. در حالت دو بعدی مجموع فوق تقریب معقولی برای حجم زیر نمودار f است چه

یعنی حجم یک مکعب مستطیل. مجموع حجم های این مکعب مستطیل ها تقریبی است برای حجم زیر نمودار. به تقلید از حالت یک بعدی می خواهیم "حد" مجموع های ریمان را، وقتی "ضخامت افراز" به صفر میل می کند، حجم زیر نمودار یا انتگرال تابع f بنامیم. ضخامت افراز را می توان مثلاً به شکل $\max_{i,j}(a_{i,j+1} - a_{i,j})$ تعریف کرد. تعریف دقیق انتگرال به صورت زیر است:

(۳۲-۱) تعریف تابع کراندار f روی مستطیل بسته R تعریف شده است. اگر عددی I وجود داشته باشد با این ویژگی که به ازای هر $\varepsilon > 0$ ، $\delta > 0$ وجود داشته باشد که هرگاه ضخامت یک افراز R از δ کوچکتر باشد، آنگاه برای مجموع (۱) داشته باشیم:

$$|I - \text{مجموع ریمان}| < \varepsilon$$

در آن صورت f را انتگرال پذیر روی R خوانده و I را، که معمولاً به $\int_R f$ نمایش می دهیم، انتگرال f روی R می نامیم.

خواص زیر به روش هایی کاملاً مشابه حالت یک متغیری قابل اثبات هستند.

(۳۲-۲) خواص ابتدایی انتگرال روی مستطیل

(۳۲-۲-۱) (خطی بودن) اگر f, g روی R انتگرال پذیر باشند و c یک عدد حقیقی باشد، آنگاه $f + g$ و cf روی R انتگرال پذیرند و

$$\int_R (f + g) = \int_R f + \int_R g \quad (۲)$$

$$\int_R (cf) = c \int_R f \quad (۳)$$

(۳۲-۲-۲) اگر f روی R انتگرال پذیر باشد و $f(x) \geq 0$ برای هر x در R ، آنگاه:

$$\int_R f \geq 0$$

(۳-۲-۳۲) اگر f تابع ثابت $f(x) = k$ روی R باشد، f انتگرال پذیر است و:

$$\int_R f = k \cdot \text{vol}(R) \quad (۴)$$

بدین ترتیب اگر از تابع ثابت با مقدار ۱ روی R انتگرال بگیریم عدد حاصل برابر حجم n -بعدی R است.

(۴-۲-۳۲) اگر مستطیل R برابر اجتماع دو زیرمستطیل R' و R'' باشد که فقط در مرز اشتراک دارند و f روی هر یک از R' و R'' انتگرال پذیر باشد، آنگاه f روی R انتگرال پذیر است و

$$\int_R f = \int_{R'} f + \int_{R''} f \quad (۵)$$

شکل ۳

(۳-۳۲) مثال مهم: تابع پیوسته. فرض کنید $f: R \rightarrow \mathbb{R}$ یک تابع پیوسته باشد. می دانیم هر تابع پیوسته روی مجموعه بسته و کراندار، کراندار است (در واقع حتی، ماکسیمم و مینیمم اتخاذ می کند). طبق قضیه‌ای که اثبات آن در اینجا نخواهد آمد، هر تابع پیوسته روی یک مجموعه بسته و کراندار \mathbb{R}^n ، در واقع به طور یکنواخت پیوسته است، بدین معنی که هرگاه $\varepsilon > 0$ داده شده باشد، $\delta > 0$ وجود دارد که برای هر دو عضو x, x' از دامنه تابع که $|x - x'| < \delta$ داریم $|f(x) - f(x')| < \varepsilon$. تفاوت این امر با پیوستگی عادی در سراسر دامنه این است که در حالت عادی یک δ ی واحد ممکن است در سراسر دامنه برای ε داده شده کار نکنند، بلکه به ازای هر x دامنه، باید $\delta > 0$ مناسب آن نقطه یافت شود که هرگاه $|x - x'| < \delta$ ، آنگاه $|f(x) - f(x')| < \varepsilon$. به کمک مفهوم پیوستگی یکنواخت می توان نشان داد که هر تابع پیوسته روی مستطیل بسته انتگرال پذیر است. در واقع انتگرال پذیری محدود به تابع‌های پیوسته نیست. در حالت یک متغیری دیده‌ایم که تعدادی متناهی "جهش" در مقدار تابع اثری بر انتگرال پذیری ندارد. قضیه‌ای قاطع در این زمینه وجود دارد که طبق آن شرطی لازم و کافی برای انتگرال پذیری تابع کراندار f روی یک مستطیل n -بعدی این است که نقاط

ناپیوستگی f مجموعه‌اً "حجم n -بعدی ناچیز" داشته باشد. تعریف دقیق چنین است: زیرمجموعه S از \mathbb{R}^n را دارای اندازه صفر n -بعدی می‌نامیم در صورتی که برای هر $\varepsilon > 0$ ، دنباله‌ای R_1, R_2, \dots از مستطیل‌های n -بعدی وجود داشته باشد که S را پوشانند، یعنی $S \subset \bigcup_{i=1}^{\infty} R_i$ و $\sum_{i=1}^{\infty} \text{vol}(R_i) < \varepsilon$.

مثال ۱ هر مجموعه متناهی در \mathbb{R}^n اندازه n -بعدی صفر دارد. قرار دهید $S = \{p_1, \dots, p_k\}$ به مرکز p_i مکعب n -بعدی (یعنی مستطیل n -بعدی با اضلاع برابر) به ضلع کوچکتر از $\sqrt[n]{\frac{\varepsilon}{k}}$ در نظر بگیرید و آن را R_i بنامید. داریم $S \subset \bigcup_{i=1}^k R_i$ و $\text{vol}(R_i) < \frac{\varepsilon}{k}$ پس $\sum \text{vol}(R_i) < \varepsilon$.

مثال ۲ در واقع مثال بالا را می‌توان به هر مجموعه شمارا $S = \{p_1, p_2, \dots\}$ تعمیم داد. حول p_i مکعبی R_i به ضلع کوچکتر از $\sqrt[n]{\frac{\varepsilon}{2^i}}$ در نظر بگیرید. داریم $\sum_{i=1}^{\infty} \text{vol}(R_i) < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon$.

مثال ۳ استدلال مشابه بالا نشان می‌دهد اگر S_1, S_2, \dots دنباله‌ای از مجموعه‌های با اندازه n -بعدی صفر باشد، اجتماع آنها نیز اندازه صفر دارد. در واقع می‌توان S_k را با دنباله‌ای از مستطیل‌ها به مجموع کوچکتر از $\frac{\varepsilon}{k}$ پوشاند. مجموعه همه این مستطیل‌ها یک پوشش شمارا برای اجتماع S_k ها می‌دهد و مجموع حجم مستطیل از ε کوچکتر است.

مثال ۴ فرض کنید R یک مستطیل بسته در \mathbb{R}^{n-1} با حجم $(n-1)$ -بعدی ناصفر V است و $f: \mathbb{R} \rightarrow \mathbb{R}$ یک تابع پیوسته. نمودار f یک زیرمجموعه \mathbb{R}^n است. نشان می‌دهیم نمودار f ، به عنوان یک زیرمجموعه \mathbb{R}^n ، دارای اندازه n -بعدی صفر است. برای $\frac{\varepsilon}{V}$ ، طبق پیوستگی یکنواخت f روی R ، $\delta > 0$ وجود دارد که هرگاه $|x - x'| < \delta$ ، آنگاه $|f(x) - f(x')| < \frac{\varepsilon}{V}$. افزای از R در نظر می‌گیریم با ضخامت آنقدر کوچک که هر دو نقطه در یک زیرمستطیل افراز، فاصله کوچکتر از δ داشته باشند. حال برای هر زیرمستطیل $(n-1)$ بعدی R_i از این افراز، مستطیلی n بعدی S_i به ترتیب زیر می‌سازیم. فرض کنید M_i و m_i به ترتیب ماکسیمم و مینیمم تابع پیوسته f روی R_i باشند. S_i را متشکل از نقاط $(x_1, \dots, x_{n-1}, x_n)$ می‌گیریم که (x_1, \dots, x_{n-1}) در R_i باشد و $m_i \leq x_n \leq M_i$. داریم

$$\text{vol}(S_i) = \text{vol}(R_i) \times (M_i - m_i)$$

که در طرف چپ $vol(S_i)$ حجم n بعدی است و در طرف راست، $vol(R_i)$ ، حجم $(n - 1)$ بعدی. بنابراین انتخاب $\delta > 0$ ، داریم $M_i - m_i < \frac{\varepsilon}{V}$ ، پس $vol(S_i) < \frac{\varepsilon}{V} vol(R_i)$. حال توجه کنید که S_i ها نمودار f را می پوشانند و

$$\sum_i vol(S_i) = \sum_i (vol(R_i) \times (M_i - m_i)) < \frac{\varepsilon}{V} \sum_i vol(R_i) = \varepsilon$$

مثال فوق وسیله سودمندی برای محاسبه انتگرال روی ناحیه های غیرمستطیل شکل خواهد شد. اکنون به بررسی انتگرال روی مجموعه های غیرمستطیلی می پردازیم. فرض کنید D یک مجموعه کراندار در \mathbb{R}^n باشد و $f: D \rightarrow \mathbb{R}$ یک تابع کراندار. چون D کراندار است، مستطیلی بسته R وجود دارد که $D \subset R$. تابع $\tilde{f}: R \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم:

$$\tilde{f}(x) = \begin{cases} f(x) & x \in D \\ 0 & x \notin D \end{cases}$$

شکل ۴

حال اگر \tilde{f} روی R انتگرال پذیر باشد، چون \tilde{f} خارج D صفر است، طبیعی است که $\int_D f$ را برابر $\int_R \tilde{f}$ تعریف کنیم. یک سوال اساسی که بلافاصله مطرح می شود این است که حتی اگر f روی D پیوسته نباشد، جهش مقادیر \tilde{f} به صفر خارج از D می تواند ناپیوستگی در سراسر مرز D پدید آورد. بدین ترتیب اگر ناحیه D دارای این ویژگی باشد که مرز D اندازه n بعدی صفر داشته باشد، آنگاه \tilde{f} روی R انتگرال پذیر خواهد شد مشروط بر این که مجموعه نقاط ناپیوستگی f در D نیز یک مجموعه با اندازه n بعدی صفر باشد (بالاخص اگر f روی D پیوسته باشد). ناحیه های کراندار D در \mathbb{R}^n را که مرز دارای اندازه n بعدی صفر داشته باشند ناحیه های مجاز خواهیم خواند. هر تابع پیوسته و کراندار روی یک چنین ناحیه مجاز انتگرال پذیر می شود (اگر D به علاوه بسته باشد، یعنی شامل مرز خود باشد، تابع پیوسته f روی D خود به خود کراندار می شود). اگر D یک ناحیه مجاز در \mathbb{R}^n باشد، حجم n بعدی D را به صورت زیر تعریف می کنیم:

$$vol(D) = \int_D 1$$

که در طرف راست، مقصود از ۱ تابع ثابت با مقدار ۱ روی D است.
 اکنون می‌توانیم خواصی مشابه آنچه در (۲-۳۲) آمد برای انتگرال روی ناحیه‌های مجاز بیان کنیم. این خواص عموماً از (۲-۳۲) نتیجه می‌شوند:

(۴-۳۲) خواص ابتدایی انتگرال روی ناحیه‌های مجاز

(۳۲-۴-۱) (خطی بودن) اگر f, g روی ناحیه مجاز D انتگرال پذیر باشند و c یک عدد حقیقی باشد، آنگاه $f + g$ و cf روی D انتگرال پذیرند و:

$$\int_D (f + g) = \int_D f + \int_D g \quad (6)$$

$$\int_D (cf) = c \int_D f \quad (7)$$

(۳۲-۴-۲) اگر f روی ناحیه مجاز D انتگرال پذیر باشد و $f(x) \geq 0$ برای هر x در D ، آنگاه:

$$\int_D f \geq 0 \quad (8)$$

از (۳۲-۴-۲) و (۱-۴-۳۲) نتیجه می‌شود که اگر g, h روی ناحیه مجاز D انتگرال پذیر باشند و $g(x) \geq h(x)$ برای هر x در D ، آنگاه

$$\int_D g \geq \int_D h \quad (9)$$

(۳۲-۴-۳) اگر f تابع ثابت $f(x) = k$ روی ناحیه مجاز D باشد، آنگاه f روی D انتگرال پذیر است

و

$$\int_D f = k \cdot \text{vol}(D) \quad (10)$$

(۳۲-۴-۴) اگر ناحیه مجاز D اجتماع دو ناحیه مجاز D' و D'' باشد که فقط در مرز اشتراک دارند و f روی هر یک از D' و D'' انتگرال پذیر باشد، آنگاه f روی D انتگرال پذیر است و

$$\int_D f = \int_{D'} f + \int_{D''} f \quad (11)$$

زیرمجموعه U از \mathbb{R}^n را همبند مسیری می نامیم در صورتی که برای هر دو نقطه A, B از U ، خمی پیوسته $\gamma : [a, b] \rightarrow U$ وجود داشته باشد که $\gamma(a) = A$ و $\gamma(b) = B$.

(۳۲-۴-۵) قضیه میانگین انتگرال فرض کنید D یک ناحیه بسته و همبند مسیری در \mathbb{R}^n باشد و $f : D \rightarrow \mathbb{R}$ یک تابع پیوسته. در این صورت نقطه ای P در D وجود دارد که:

$$\int_D f = f(P) \cdot \text{vol}(D) \quad (12)$$

اثباتی از این حکم ارائه می کنیم. چون f پیوسته است روی مجموعه بسته و کراندار D در نقطه ای A مینیمم و در نقطه ای B ماکسیمم خود را اتخاذ می کند، $f(A) = m$ و $f(B) = M$. برای هر x در D داریم $m \leq f(x) \leq M$ پس بنابر (۹) و (۱۰):

$$m \cdot \text{vol}(D) \leq \int_D f \leq M \cdot \text{vol}(D)$$

یا

$$m \leq \frac{\int_D f}{\text{vol}(D)} \leq M \quad (13)$$

چون D همبند مسیری فرض شده است، خمی پیوسته $\gamma : [a, b] \rightarrow D$ وجود دارد که $\gamma(a) = A$ و $\gamma(b) = B$. پس $(f \circ \gamma)(a) = m$ و $(f \circ \gamma)(b) = M$. حال $f \circ \gamma$ تابعی پیوسته از $[a, b]$ به \mathbb{R} است که در نقطه a مقدار m و در نقطه b مقدار M را می گیرد. بنابر قضیه مقدار بینی باید نقطه ای c در $[a, b]$ وجود داشته باشد که $(f \circ \gamma)(c) = \frac{\int_D f}{\text{vol}(D)}$. نقطه $P = \gamma(c)$ در حکم قضیه صدق می کند.

محاسبه انتگرال چندگانه

پس از بحث در مورد کلیات انتگرال که در بخش قبل گذشت، اکنون به تشریح روش محاسبه انتگرال چندگانه می‌پردازیم. مقدمتاً ذکر یکی دو نکته در مورد نمادها لازم است. انتگرال تابع f روی مجموعه D از \mathbb{R}^n را تاکنون به $\int_D f$ نمایش داده‌ایم. برای تأکید n بعدی بودن ناحیه D ، بعضی نماد $\int \cdots \int_D f$ (با n بار f) را به کار می‌برند. همچنین اگر نقاط D به $x = (x_1, \dots, x_n)$ نمایش داده شده باشند، نماد $\int \cdots \int f(x_1, \dots, x_n) dV$ یا $\int \cdots \int f(x_1, \dots, x_n) dV_x$ نیز به کار گرفته می‌شود. گاهی از dV به عنوان "عنصر حجم n -بعدی" یاد می‌شود. در حالت دو بعدی، $n = 2$ ، به جای dV معمولاً از dA ("عنصر سطح") استفاده می‌شود.

محاسبه انتگرال را با انتگرال دوگانه روی مستطیل آغاز می‌کنیم. قضیه زیر راهگشای حالت کلی است.

(۳۳-۱) قضیه فرض کنید مستطیل بسته $D = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$ و تابع انتگرال پذیر $f : D \rightarrow \mathbb{R}$ داده شده است. در این صورت:

$$\int \int_D f(x, y) dA = \int_a^b \left[\int_c^d f(x, y) dy \right] dx \quad (1)$$

$$= \int_c^d \left[\int_a^b f(x, y) dx \right] dy \quad (2)$$

نخست برای (۱) (معادلاً (۲)) یک تعبیر هندسی ارائه می‌کنیم. حالت $f(x, y) \geq 0$ را در نظر بگیرد. در انتگرال داخلی (۱)، برای x ثابت یک انتگرال گیری نسبت به y انجام شده است. این به منزله محاسبه مساحت مقطع ثابت $x =$ در زیر نمودار f است. بنابراین (۱) بیانگر این مطلب است که برای محاسبه حجم زیر نمودار، برای هر x ثابت مساحت مقطع مربوط زیر نمودار را محاسبه می‌کنیم که تابعی از x می‌شود سپس با محاسبه انتگرال این تابع مساحت برای $a \leq x \leq b$ ، حجم مورد نظر را

به دست می آوریم. در واقع حجم زیر نمودار از انباشتن لایه های بینهایت نازک به ازای x های ثابت گوناگون به دست می آید.

تعبیر دیگر (۱) را می توان بر اساس مجموع های ریمان ارائه کرد. فرض کنید افراز $a = x_0 < x_1 < \dots < x_p = b$ برای $[a, b]$ و افراز $c = y_0 < y_1 < \dots < y_q = d$ برای $[c, d]$ در نظر گرفته شوند و در زیر مستطیل $R_{ij} : x_{i-1} \leq x \leq x_i$ و $y_{j-1} \leq y \leq y_j$ نقطه دلخواهی z_{ij} در نظر بگیرید. مجموع ریمان مربوط عبارت است از

$$\sum_{i=1}^p \sum_{j=1}^q f(z_{ij}) \cdot (y_j - y_{j-1})(x_i - x_{i-1}) = \sum_{i=1}^p \left(\sum_{j=1}^q f(z_{ij})(y_j - y_{j-1}) \right) (x_i - x_{i-1}) \quad (3)$$

در پراکنش داخلی، برای i ثابت، نسبت به j جمع بندی شده است. برای گذر به حد، باید $(x_i - x_{i-1})$ ها و $(y_j - y_{j-1})$ ها توأمأً به صفر میل کنند. فرمول (۱) بیانگر این است که تحت شرایط انتگرال گیری می توان نخست برای x ثابت گذر به حد را با کوچک کردن $(y_j - y_{j-1})$ انجام داد و سپس حاصل را که وابسته به x است مجدداً با سوق دادن $(x_i - x_{i-1})$ به صفر حدگیری کرد. با دقیق تر کردن این ایده می توان اثبات کاملی از (۱) ارائه کرد ولی این کار در اینجا انجام نخواهد شد.

مثال ۱ انتگرال تابع $f(x) = x^2 y + e^{-2y}$ را روی مستطیل $0 \leq y \leq 1$ ، $1 \leq x \leq 2$ محاسبه کنید. این محاسبه را از هر دو روش (۱) و (۲) انجام می دهیم. اگر R مستطیل فوق باشد:

$$\begin{aligned} \iint_R (x^2 y + e^{-2y}) dA &= \int_1^2 \left[\int_0^1 (x^2 y + e^{-2y}) dy \right] dx \quad \text{از فرمول (۱)} \\ &= \int_1^2 \left[\frac{1}{2} x^2 y^2 - \frac{1}{2} e^{-2y} \right]_0^1 dx \\ &= \int_1^2 \left(\frac{1}{2} x^2 - \frac{1}{2} e^{-2} + \frac{1}{2} \right) dx \\ &= \left[\frac{1}{6} x^3 + \frac{1}{2} (1 - e^{-2}) x \right]_1^2 \\ &= \frac{8}{3} + (1 - e^{-2}) - \frac{1}{6} - \frac{1}{2} (1 - e^{-2}) = \frac{5}{3} - \frac{1}{2} e^{-2} \end{aligned}$$

$$\begin{aligned}
 \iint_R (x^2 y + e^{-2y}) dA &= \int_0^1 \left[\int_1^2 (x^2 y + e^{-2y}) dx \right] dy \quad (2) \text{ از فرمول} \\
 &= \int_0^1 \left[\frac{1}{3} x^3 y - e^{-2y} x \right]_1^2 dy \\
 &= \int_0^1 \left(\frac{7}{3} y + e^{-2y} \right) dy \\
 &= \left(\frac{7}{6} y^2 - \frac{1}{2} e^{-2y} \right) \Big|_0^1 = \frac{7}{6} - \frac{1}{2} e^{-2} + \frac{1}{2} = \frac{5}{3} - \frac{1}{2} e^{-2}
 \end{aligned}$$

قضیه (۱-۲۳) تعمیمی واضح به حالت مستطیل n -بعدی دارد. به طور کلی می توان یک انتگرال روی مستطیل n بعدی را به $n!$ نوع مختلف به عنوان انتگرال مکرر بیان کرد.

مثال ۲ انتگرال تابع $f(x, y, z) = x + e^{-y+z}$ را روی مستطیل R با مشخصات $1 \leq x \leq 2$ ، $0 \leq y \leq 1$ ، $0 \leq z \leq 2$ ، محاسبه کنید. این انتگرال را به یکی از شش ترتیب ممکن محاسبه می کنیم.

$$\begin{aligned}
 \iiint_R (x + e^{-y+z}) dV &= \int_0^2 \left[\int_0^1 \left[\int_1^2 (x + e^{-y+z}) dx \right] dy \right] dz \\
 &= \int_0^2 \left[\int_0^1 \left[\left(\frac{1}{2} x^2 + e^{-y+z} x \right) \Big|_1^2 \right] dy \right] dz \\
 &= \int_0^2 \left[\int_0^1 \left(\frac{3}{2} + e^{z-y} \right) dy \right] dz \\
 &= \int_0^2 \left(\frac{3}{2} y - e^{z-y} \right) \Big|_0^1 dz \\
 &= \int_0^2 \left(\frac{3}{2} - e^{z-1} + e^z \right) dz \\
 &= \left(\frac{3}{2} z - e^{z-1} + e^z \right) \Big|_0^2 = 3 - e + e^2 + e^{-1} - 1 \\
 &= 2 - e + e^2 + e^{-1}
 \end{aligned}$$

برای محاسبه انتگرال روی ناحیه های غیر مستطیل شکل، به روش زیر عمل می کنیم. اگر ناحیه مجاز D و تابع انتگرال پذیر $f: D \rightarrow \mathbb{R}$ داده شده باشد، مانند جلسه قبل، D را در یک مستطیل R جای داده، تابع f را به تابعی \tilde{f} توسعه می دهیم که در R تعریف شده ولی خارج D صفر است، به عنوان مثال مورد زیر را در نظر بگیرید. فرض کنید تابع های $\alpha, \beta: [a, b] \rightarrow \mathbb{R}$ داده شده اند که $\alpha(x) \leq \beta(x)$

برای هر x در $[a, b]$. ناحیه D به شکل زیر را یک ناحیه y -محدب می‌نامیم:

$$D : \begin{cases} a \leq x \leq b \\ \alpha(x) \leq y \leq \beta(x) \end{cases}$$

حال ناحیه y -محدب D را در مستطیل R با تعریف $a \leq x \leq b, c \leq y \leq d$ که $c \leq \alpha(x) \leq \beta(x) \leq d$ قرار می‌دهیم و داریم:

$$\iint_D f(x, y) dA = \int \int_R \tilde{f}(x, y) dA = \int_a^b \left[\int_c^d \tilde{f}(x, y) dy \right] dx$$

برای x ثابت، $\tilde{f}(x, y)$ به‌ازای $y < \alpha(x)$ و $y > \beta(x)$ صفر است، پس

$$\int_c^d \tilde{f}(x, y) dy = \int_{\alpha(x)}^{\beta(x)} f(x, y) dy$$

و در نتیجه:

$$\iint_D f(x, y) dA = \int_a^b \left[\int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right] dx \quad (4)$$

ضمناً توجه کنید که اگر α و β پیوسته باشند، همچنان که در درس قبل دیدیم نمودار هر یک، یک مجموعه دارای اندازه صفر دو بعدی خواهد بود و ناحیه D یک ناحیه مجاز انتگرال‌گیری است.

به همین ترتیب یک ناحیه x -محدب ناحیه‌ای با توصیف زیر است:

$$D : \begin{cases} c \leq y \leq d \\ \gamma(y) \leq x \leq \delta(y) \end{cases}$$

و اگر تابع‌های γ و δ پیوسته باشند، D یک ناحیه مجاز انتگرال‌گیری خواهد بود.

برای این نوع ناحیه به ترتیب قبل خواهیم داشت:

$$\iint_D f(x, y) dA = \int_c^d \left[\int_{\gamma(y)}^{\delta(y)} f(x, y) dx \right] dy \quad (5)$$

مثال ۳ ربع گوی زیر را در نظر بگیرید:

$$D : \begin{cases} 0 \leq x \leq 1 \\ 1 \leq y \leq 1 + \sqrt{1 - x^2} \end{cases}$$

می‌خواهیم $\int \int_D x dA$ را محاسبه کنیم. طبق فرمول (۴) داریم:

$$\begin{aligned} \int \int_D x dA &= \int_0^1 \left[\int_1^{1+\sqrt{1-x^2}} x dy \right] dx \\ &= \int_0^1 (xy) \Big|_1^{1+\sqrt{1-x^2}} dx \\ &= \int_0^1 x \sqrt{1-x^2} dx \\ &= \left(-\frac{1}{3} \right) \left(\frac{2}{3} \right) (1-x^2)^{\frac{3}{2}} \Big|_0^1 = \frac{1}{3} \end{aligned}$$

ناحیه D ، x محدب نیز هست و می‌توان آن را به شکل زیر تعریف کرد:

$$\begin{cases} 1 \leq y \leq 2 \\ 0 \leq x \leq \sqrt{2y-y^2} \end{cases}$$

بنابراین،

$$\begin{aligned} \int \int_D x dA &= \int_1^2 \left[\int_0^{\sqrt{2y-y^2}} x dx \right] dy \\ &= \int_1^2 \left(\frac{1}{2} x^2 \right) \Big|_0^{\sqrt{2y-y^2}} dy \\ &= \frac{1}{2} \int_1^2 (2y-y^2) dy \\ &= \left(\frac{1}{2} \right) \left(y^2 - \frac{1}{3} y^3 \right) \Big|_1^2 = \frac{1}{3} \end{aligned}$$

گاهی اوقات محاسبه یک انتگرال به یکی از دو روش (۴) یا (۵) مقدور نیست یا دشوار است ولی به روش دیگر امکان‌پذیر است. مثال زیر نمونه‌ای از این وضعیت است.

مثال ۴ می‌خواهیم انتگرال تابع $f(x, y) = e^{-y^2}$ را روی ناحیه مثلثی زیر محاسبه کنیم:

$$T : \begin{cases} 0 \leq x \leq 1 \\ x \leq y \leq 1 \end{cases}$$

اگر این ناحیه را به شکل y -محدب تلقی کنیم، داریم:

$$\int \int_T e^{-y^2} dA = \int_0^1 \left[\int_x^1 e^{-y^2} dy \right] dx$$

در اینجا به این مشکل برمی‌خوریم که تابع اولیه e^{-y^2} را نمی‌شناسیم و ادامه محاسبه به این طریق عملی نیست. ولی اگر T را به صورت x - محدب بنویسیم، نتیجه حاصل خواهد شد:

$$T : \begin{cases} 0 \leq y \leq 1 \\ 0 \leq x \leq y \end{cases}$$

$$\begin{aligned} \iint_T e^{y^2} dA &= \int_0^1 \left[\int_0^y e^{y^2} dx \right] dy \\ &= \int_0^1 (e^{-y^2} x) \Big|_0^y dy \\ &= \int_0^1 e^{-y^2} y dy \\ &= \left(-\frac{1}{2}\right) e^{-y^2} \Big|_0^1 = \left(-\frac{1}{2}\right)(e^{-1} - 1) \end{aligned}$$

روشی که برای محاسبه انتگرال برای ناحیه‌های دو بعدی غیر مستطیل شکل ذکر شد به حالت n بعدی نیز تعمیم پذیر است و اگر بتوانیم ناحیه D در \mathbb{R}^n را به صورت زیر توصیف کنیم:

$$\begin{cases} \alpha_n(x_1, \dots, x_{n-1}) \leq x_n \leq \beta_n(x_1, \dots, x_{n-1}) \\ \vdots \\ \alpha_2(x_1) \leq x_2 \leq \beta_2(x_1) \\ \alpha_1 \leq x_1 \leq \beta_1 \end{cases}$$

آنگاه $\int \dots \int_D f$ را می‌توان به صورت زیر محاسبه کرد:

$$\int \dots \int_D f = \int_{\alpha_1}^{\beta_1} \left[\int_{\alpha_2(x_1)}^{\beta_2(x_1)} \left[\dots \left[\int_{\alpha_n(x_1, \dots, x_{n-1})}^{\beta_n(x_1, \dots, x_{n-1})} f(x_1, \dots, x_n) dx_n \right] dx_{n-1} \dots \right] dx_1 \quad (7)$$

البته هر ترتیب دیگری از متغیرهای x_1, \dots, x_n نیز ممکن است به کار رود. توجه کنید که در هر مرحله انتگرال گیری در فرمول (7) حداقل یکی از متغیرها از بین می‌رود به طوری که در مرحله آخر فقط یک انتگرال گیری یک متغیری با حدود عددی باقی می‌ماند.

مثال 5 حجم ناحیه محصوره به $z \geq x^2 + y^2$ و داخل استوانه $z = 2 - x^2$ را به کمک انتگرال سه‌گانه محاسبه کنید.

ناحیه مورد نظر را می توان به صورت زیر توصیف کرد:

$$D : \begin{cases} -\sqrt{z-x^2} \leq y \leq \sqrt{z-x^2} \\ 2 - \sqrt{1-x^2} \leq z \leq 2 + \sqrt{1-x^2} \\ -1 \leq x \leq 1 \end{cases}$$

بنابراین

$$\iiint_D 1 dV = \int_{-1}^1 \left[\int_{2-\sqrt{1-x^2}}^{2+\sqrt{1-x^2}} \left[\int_{-\sqrt{z-x^2}}^{\sqrt{z-x^2}} dy \right] dz \right] dx$$

و با سه بار انتگرال گیری حجم مورد نظر به دست می آید.

مثال ۶ حجم ناحیه محصور بین دو استوانه $x^2 + y^2 = a^2$ و $x^2 + z^2 = a^2$ را محاسبه کنید ($a > 0$).

توجه کنید که این جسم روی گوی $x^2 + y^2 \leq a^2$ بنا شده است، جدار بالایی آن روی نیمه بالایی سطح استوانه $x^2 + z^2 = a^2$ و جدار پایینی آن روی نیمه پایینی سطح همین استوانه قرار دارد. ناحیه سه بعدی را می توان به صورت زیر توصیف کرد:

$$D : \begin{cases} -\sqrt{a^2-x^2} \leq z \leq \sqrt{a^2-x^2} \\ -\sqrt{a^2-x^2} \leq y \leq \sqrt{a^2-x^2} \\ -a \leq x \leq a \end{cases}$$

پس

$$\iiint_D 1 dV = \int_{-a}^a \left[\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \left[\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dy \right] dz \right] dx$$

مثال ۷ حجم چهار بعدی گوی $x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq R^2$ در \mathbb{R}^4 را محاسبه کنید.

این ناحیه را می توان به شکل زیر توصیف کرد:

$$D : \begin{cases} -\sqrt{R^2 - (x_1^2 + x_2^2 + x_3^2)} \leq x_4 \leq \sqrt{R^2 - (x_1^2 + x_2^2 + x_3^2)} \\ -\sqrt{R^2 - (x_1^2 + x_2^2)} \leq x_3 \leq \sqrt{R^2 - (x_1^2 + x_2^2)} \\ -\sqrt{R^2 - x_1^2} \leq x_2 \leq \sqrt{R^2 - x_1^2} \\ -R \leq x_1 \leq R \end{cases}$$

می‌توان به روش ذکر شده به نتیجه رسید ولی با توجه به این که فرمول حجم گوی سه بعدی را می‌دانیم، محاسبه را می‌توان در یک گام خلاصه کرد. توجه کنید که برای x_1 ثابت، مقطع این گوی، یک گوی سه بعدی به شعاع $\sqrt{R^2 - x_1^2}$ است که حجم سه بعدی آن $\frac{4}{3}\pi(R^2 - x_1^2)^{\frac{3}{2}}$ می‌باشد این در واقع نتیجه سه انتگرال گیری اول روی ناحیه بالا است. بنابراین

$$\text{حجم گوی چهار بعدی} = \int_{-R}^R \frac{4}{3}\pi(R^2 - x_1^2)^{\frac{3}{2}} dx_1$$

با قرار دادن $x_1 = R \sin \theta$ ، انتگرال بالا تبدیل می‌شود به

$$\begin{aligned} \left(\frac{4}{3}\pi\right) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (R^2 \cos^2 \theta)(R \cos \theta) d\theta &= \left(\frac{4}{3}\pi R^3\right) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3 \theta d\theta \\ &= \left(\frac{4}{3}\pi R^3\right) \left(\frac{2\pi}{\pi}\right) = \frac{8}{3}\pi R^3 \end{aligned}$$

بالاخره این نکته در مورد نماد انتگرال مکرر قابل ذکر است که معمولاً گروه‌ها نوشته نمی‌شوند و مقصود از نماد $\int \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n$ این است که انتگرال گیری از داخل به خارج به ترتیب نسبت به x_1 ، بعد نسبت به x_2 ، ... و نهایتاً نسبت به x_n انجام می‌شود.

تعویض متغیر در انتگرال چندگانه

در حالت یک متغیری مشاهده کرده‌ایم که کمتر ممکن است یک انتگرال مستقیماً از دستوره‌های متداول جداول تابع اولیه محاسبه شود، بلکه معمولاً نوعی تعویض متغیر و استفاده هر چند ضمنی از قاعده زنجیره‌ای برای محاسبه انتگرال لازم است. در مورد توابع چند متغیری نیاز به یک روش مؤثر تعویض متغیر از اهمیت کمتری برخوردار نیست و نقشی تعیین کننده هم در عمل و هم در تئوری دارد. قبل از پرداختن به فرمول تعویض متغیر انتگرال چند متغیری لازم است حالت یک متغیری را مورد بازبینی قرار دهیم. فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ یک تابع انتگرال پذیر، مثلاً تابعی پیوسته، باشد، $\varphi : [\alpha, \beta] \rightarrow \mathbb{R}$ یک تابع مشتق پذیر با مشتق پیوسته به طوری که φ بازه $[\alpha, \beta]$ را بر بازه $[a, b]$ می‌نگارد به این صورت که $\varphi(\alpha) = a$ و $\varphi(\beta) = b$ یا $\varphi(\alpha) = b$ و $\varphi(\beta) = a$. اگر نقاط $[a, b]$ را به t و نقاط $[a, b]$ را به τ نمایش دهیم، داریم

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(t) dt = \int_{\alpha}^{\beta} f(\varphi(\tau)) \varphi'(\tau) d\tau \quad (1)$$

این فرمول تعویض متغیر تابع‌های یک متغیری است و به سادگی از قاعده زنجیره‌ای نتیجه می‌شود. می‌خواهیم این فرمول را از دیدگاه هندسی تعبیر انتگرال به عنوان مساحت بررسی کنیم و به خصوص به نقش $\varphi'(\tau)$ آگاهی عمیقتری کسب کنیم. به این منظور f را مثبت فرض کنید و موقتاً فرض کنید φ' تغییر علامت نمی‌دهد، یعنی همواره $\varphi'(\tau) > 0$ یا همواره $\varphi'(\tau) < 0$ برای هر τ در $[\alpha, \beta]$. بدین ترتیب φ صعودی یا نزولی و در هر حال یک به یک است. وقتی $\varphi' > 0$ داریم $\varphi(\alpha) = a$ و $\varphi(\beta) = b$ و در حالتی که $\varphi' < 0$ داریم $\varphi(\alpha) = b$ و $\varphi(\beta) = a$. در شکل ۲ نمودارهای f و $f \circ \varphi$ را

شکل ۱

در حالت $\varphi' > 0$ نمایش داده‌ایم. توجه کنید که به طور کلی طول بازه‌های $[\alpha, \beta]$ و $[a, b]$ لازم نیست برابر باشد ولی چون φ یک به یک فرض شده است تناظری یک به یک میان نقاط $[\alpha, \beta]$ و $[a, b]$ وجود دارد و مقادیر f و $f \circ \varphi$ در نقاط متناظر برابر است: $f(t) = (f \circ \varphi)(\tau)$ وقتی $t = \varphi(\tau)$. به طور کلی نمودار f از نمودار $f \circ \varphi$ با کشش وقتی $\varphi' > 1$ و فشردن وقتی $0 < \varphi' < 1$ به دست می‌آید.

شکل ۲

از آنجا که تناظری یک به یک بین مقادیر دو تابع f و $f \circ \varphi$ وجود دارد ولی توزیع این مقادیر با حفظ ترتیب به سبب انبساط ($\varphi' > 1$) یا انقباض ($\varphi' < 1$) پایه دستخوش جابجایی شده است، نمی‌توان انتظار داشت که دو مساحت $\int_a^b f(t) dt$ و $\int_\alpha^\beta f(\varphi(\tau)) d\tau$ برابر باشند. در اینجا نقش هندسی ضریب $\varphi'(\tau)$ که باید در داخل انتگرال دوم ضرب شود نمایان می‌شود. وقتی $\varphi'(\tau) > 1$ بازه‌های کوچک حول τ تحت φ بزرگتر می‌شوند، بنابراین افزایش مساحت روی بازه متناظر صورت می‌گیرد، و بالعکس وقتی $0 < \varphi'(\tau) < 1$ بازه‌های کوچک حول τ را منقبض می‌کند و کاهش مساحت زیر منحنی رخ می‌دهد. حال با ضرب کردن $f(\varphi(\tau))$ در $\varphi'(\tau)$ عملاً نوعی "تصحیح تغییر طول پایه" با تعدیل نمودار صورت می‌گیرد به طوری که $\int_a^b f(t) dt$ و $\int_\alpha^\beta f(\varphi(\tau)) \varphi'(\tau) d\tau$ دو مساحت برابر را نشان می‌دهند. شکل ۳ بیانگر این مطلب است.

شکل ۳

در حالت $\varphi' < 0$ وضع مشابه این است با این تفاوت که دو انتهای بازه تعویض می‌شوند. در این حالت منفی بودن φ' و این که در \int_α^β حد بالای انتگرال کوچکتر از حد پایین است (شکل ۱، سمت راست) توأمأ منجر به نتیجه مثبت برای حاصل انتگرال‌گیری می‌شوند. یک نکته آخر در مورد فرمول تعویض متغیر این که در (۱) لازم نیست φ تابعی یک به یک باشد. نکته این است که وقتی φ' تغییر علامت می‌دهد، همان طور که در بالا ذکر شد، تعویض حدود بالا و پایین انتگرال خود به خود نتیجه نهایی را تصحیح می‌کند. در شکل ۴ انتگرال‌گیری $\int_a^b 1 \cdot dt$ تحت تعویض متغیر غیر یک به یک نمایش داده شده است.

شکل ۴

با این مقدمات آماده‌ایم تعویض متغیر در انتگرال چندگانه را توصیف کنیم:

(۳۴-۱) قضیه (تعویض متغیرانتگرال چندگانه) W' و W ناحیه‌های همبند مسیری و مجاز انتگرال گیری در \mathbb{R}^n هستند، φ یک تابع دارای مشتق‌های پاره‌ای پیوسته از یک زیرمجموعه \mathbb{R}^n به \mathbb{R}^n که دامنه آن شامل W' است و W' را به طور یک به یک بر W می‌نگارد، $\varphi(W') = W$ و $\det D\varphi(x) \neq 0$ برای هر x در W' . در این صورت برای هر تابع انتگرال پذیر $f: W \rightarrow \mathbb{R}$ داریم:

$$\int_W f = \int_{W'} (f \circ \varphi) \cdot |\det D\varphi| \quad (2)$$

تحت شرایط قضیه $|\det D\varphi| (f \circ \varphi)$ یک تابع انتگرال پذیر روی W' خواهد شد، و طبق قضیه تابع وارون، وارون تحدید f به W وجود دارد، $\varphi^{-1}: W \rightarrow W'$ و φ^{-1} دارای مشتقات پاره‌ای پیوسته است.

فرمول (۲) را می‌توان به صورت زیر نیز نوشت. اگر نقاط W را به $y = (y_1, \dots, y_n)$ و نقاط W' را به $x = (x_1, \dots, x_n)$ نمایش دهیم، $y = \varphi(x)$ ، آنگاه

$$\int \dots \int_W f(y_1, \dots, y_n) dV_y = \int \dots \int_{W'} f(\varphi(x_1, \dots, x_n)) \left| \det \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \right| dV_x \quad (3)$$

مقصود از dV_y و dV_x به ترتیب $dx_1 \dots dx_n$ و $dy_1 \dots dy_n$ ، هر یک به ترتیب مناسب برای انتگرال گیری مکرر است.

قبل از توضیح بیشتر در مورد صحت این فرمول، که در پایان جلسه خواهد آمد، نکته‌ای در مورد مقایسه این فرمول با حالت یک متغیری ذکر می‌کنیم و چند مثال می‌آوریم.

(۳۴-۲) مقایسه با حالت یک متغیری در نگاه اول وجود علامت قدرمطلق در اینجا متفاوت از حالت یک متغیری به نظر می‌رسد. در حالت یک متغیری نیازی به $||$ نیست زیرا که در صورت منفی بودن φ' ، حدود انتگرال گیری خود به خود وارونه می‌شوند یعنی حد بالایی کوچکتر از حد پایینی خواهد شد ولی در حالت $n \geq 2$ ، روش انتگرال گیری روی مستطیل‌ها چنان بود که همواره حد پایینی انتگرال

کوچکتر از حد بالایی است. از این رو علامت قدر مطلق الزامی می‌شود. در حالت یک متغیری نیز می‌توان فرمول (۱) را به صورت

$$\int_a^b f(t)dt = \int_\alpha^\beta f(\varphi(\tau))|\varphi'(\tau)|d\tau$$

نوشت مشروط بر این که به جای $\varphi(\alpha) = a$ و $\varphi(\beta) = b$ داشته باشیم $\{\varphi(\alpha), \varphi(\beta)\} = \{a, b\}$ و $a \leq b$ و $\alpha \leq \beta$. شرط یک به یک بودن φ نیز در اینجا الزامی می‌شود زیرا به علت فقدان مکانیسمی مشابه وارونه کردن حد بالایی و پایینی انتگرال، حذف خود به خود انتگرال گیری مثبت و منفی زائد صورت نمی‌گیرد.

مثال ۱ حجم درون بیضی‌وار، یعنی ناحیه $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1$ را محاسبه کنید که در آن a, b, c اعداد حقیقی مثبت داده شده‌اند.

حل با یک تعویض متغیر ساده می‌توان بیضی‌وار را به کره تبدیل کرد که حجم درون آن را می‌دانیم. قرار دهید $X = \frac{x}{a}$ و $Y = \frac{y}{b}$ و $Z = \frac{z}{c}$ ، و:

$$(x, y, z) = \varphi(X, Y, Z) = (aX, bY, cZ)$$

$$\det \frac{\partial(x, y, z)}{\partial(X, Y, Z)} = \det \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} = abc$$

$$\iiint_{\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1} 1 \cdot dV_{x,y,z} = \iiint_{X^2 + Y^2 + Z^2 \leq 1} (abc) dV_{X,Y,Z} = (abc) \frac{4}{3} \pi$$

مثال ۲ حجم جسم محصور درون کره $x^2 + y^2 + z^2 \leq 1$ و درون استوانه $(x - \frac{1}{2})^2 + y^2 \leq \frac{1}{4}$ را محاسبه کنید.

حل

شکل ۵

از مختصات استوانه‌ای (r, θ, z) استفاده می‌کنیم. یادآوری می‌کنیم که $\det \frac{\partial(x,y,z)}{\partial(r,\theta,z)} = r$. معادله دایره $(x - \frac{1}{2})^2 + y^2 = \frac{1}{4}$ در مختصات قطبی می‌شود $r = \cos \theta$ ، بنابراین جسم سه‌بعدی مورد نظر در مختصات استوانه‌ای توصیف زیر را دارد:

$$\begin{cases} -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \\ 0 \leq r \leq \cos \theta \\ -\sqrt{1-r^2} \leq z \leq \sqrt{1-r^2} \end{cases}$$

بنابراین

$$\begin{aligned} \text{حجم مورد نظر} &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{\cos \theta} \int_{-\sqrt{1-r^2}}^{\sqrt{1-r^2}} r dz dr d\theta \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{\cos \theta} (2r) \sqrt{1-r^2} dr d\theta \\ &= - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{2}{3} [(1 - \cos^2 \theta)^{\frac{3}{2}} - 1] d\theta \\ &= - \frac{2}{3} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (|\sin \theta|^3 - 1) d\theta = - \frac{4}{3} \int_0^{\frac{\pi}{2}} (\sin^3 \theta - 1) d\theta \\ &= \left(-\frac{4}{3}\right) \int_0^{\frac{\pi}{2}} [\sin \theta - \sin \theta \cos^2 \theta - 1] d\theta \\ &= \left(-\frac{4}{3}\right) [(0 + 1) + \frac{1}{3}(0 - 1) - \frac{\pi}{2}] \\ &= \frac{2\pi}{3} - \frac{8}{9} \end{aligned}$$

مثال ۳ حجم قیفی که از بالا به سطح کره $x^2 + y^2 + z^2 = R^2$ و از پایین به مخروط $\varphi = \alpha$:

مختصه کروی، α : ثابت در $[0, \pi]$ محصور است محاسبه کنید

شکل ۶

حل می‌دانیم $\det \frac{\partial(x,y,z)}{\partial(r,\varphi,\theta)} = \rho^2 \sin \varphi$. ناحیه مورد نظر در مختصات کروی توصیف زیر را دارد:

$$\begin{cases} 0 \leq \theta \leq 2\pi \\ 0 \leq \varphi \leq \alpha \\ 0 \leq \rho \leq R \end{cases}$$

بنابراین

$$\begin{aligned} \text{حجم مورد نظر} &= \int_0^{2\pi} \int_0^\alpha \int_0^R \rho^2 \sin \varphi d\rho d\varphi d\theta \\ &= (2\pi) \left(\int_0^\alpha \sin \varphi d\varphi \right) \left(\int_0^R \rho^2 d\rho \right) \\ &= \frac{2\pi}{3} R^3 (-\cos \alpha + 1) \end{aligned}$$

توجه کنید برای $\alpha = \pi$ ، حجم گوی سه‌بعدی شعاع R ، یعنی $\frac{4}{3}\pi R^3$ ، حاصل می‌شود.

مثال ۴ مساحت متوازی الاضلاع محصور به چهار خط دو به دو موازی زیر را محاسبه کنید:

$$\begin{cases} A_1 x + B_1 y = C_1 \\ A_1 x + B_1 y = C'_1 \end{cases} \quad \begin{cases} A_2 x + B_2 y = C_2 \\ A_2 x + B_2 y = C'_2 \end{cases}$$

حل این مسأله به روش عادی انتگرال‌گیری یک متغیری برای مساحت در تئوری سرراست است ولیکن با توجه به وضعیت‌های گوناگونی که ممکن است چهار رأس قرارگیرند انتگرال مربوط را باید به چند انتگرال تقسیم کرد و محاسبات در هر صورت جذاب و طبیعی نیستند. با فرمول تعویض متغیر انتگرال دوگانه می‌توانیم متوازی الاضلاع را به صورت یک مستطیل در آوریم و با یک محاسبه ساده جواب مورد نظر در حالت کلی مستقل از مکان نسبی چهار رأس به دست می‌آید. تعویض متغیر طبیعی زیر را در نظر بگیرید:

$$\begin{cases} u = A_1 x + B_1 y \\ v = A_2 x + B_2 y \end{cases}$$

پس $\frac{\partial(u,v)}{\partial(x,y)} = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix}$ که دترمینان آن $A_1 B_2 - A_2 B_1$ است. با توجه به این که دترمینان

معکوس ماتریس برابر معکوس دترمینان آن ماتریس است، داریم

$$\left| \det \frac{\partial(x, y)}{\partial(u, v)} \right| = \frac{1}{|A_1 B_2 - A_2 B_1|}$$

بدین ترتیب

$$\begin{aligned} (\text{مساحت متوازی الاضلاع}) &= \iint_{\text{متوازی الاضلاع}} 1 \cdot dV_{x,y} \\ &= \left| \int_{C_1}^{C'_1} \int_{C_2}^{C'_2} \frac{1}{|A_1 B_2 - A_2 B_1|} du dv \right| \end{aligned}$$

(قدرمطلق بیرونی بدین سبب است که معلوم نیست C_i و C'_i ، $i = 1, 2$ ، کدام بزرگترند).

$$= \frac{|C'_1 - C_1| |C'_2 - C_2|}{|A_1 B_2 - A_2 B_1|}$$

شکل ۷

(۳-۳۴) توضیح در مورد اثبات قضیه (۱-۳۴) ارائه اثبات کامل این قضیه از حد بحث ما کمی خارج است ولی می‌توان به اصول اساسی آن با توجه به بازبینی فرمول (۱) که در آغاز جلسه مطرح شد پی برد. حالتی را در نظر بگیرید که W' (در قضیه) یک مستطیل n -بعدی است (در هر صورت، طبق تعریف، همه انتگرال گیری‌های ما به انتگرال گیری روی مستطیل تحویل شد). تحت اثر φ ، W' به ناحیه W نگاشته می‌شود که حجم n -بعدی آن می‌تواند متفاوت از حجم W باشد، یا حتی اگر برابر باشد در بعضی قسمت‌ها گسترده‌تر شده و در بعضی نواحی دیگر فشرده‌تر شده باشد (شکل ۸ حالت دو بعدی را نشان می‌دهد)، چگونه می‌توان به یک ضریب "تصحیح تغییر حجم پایه"، مانند $\varphi'(\tau)$ در حالت یک بعدی، دست یافت؟ مستطیل W' را با ترسیم ابرصفحه‌هایی به موازات ابرصفحه‌های مختصاتی به قطعات کوچکتر افراز می‌کنیم و تصویر این

شکل ۸

برای هر زیرمستطیل کوچک نقطه‌ای p در آن اختیار می‌کنیم. اگر این زیرمستطیل به اندازه کافی کوچک باشد حجم تصویر آن حدوداً برابر حجم تصویر آن تحت تقریب خطی یعنی $D\varphi(p)$ است. تحت اثر تابع‌های خطی حجم n -بعدی متوازی‌السطوح در $|\det D\varphi(p)|$ ضرب می‌شود، بنابراین $|\det D\varphi(p)|$ به طور تقریب ضریب "تصحیح تغییر حجم پایه" مناسب است. بدین ترتیب می‌توان انتظار داشت که $|\det D\varphi(p)| \cdot \Sigma(f \circ \varphi)(p)$ (مجموع نسبت به زیرمستطیل‌های افراز W') تقریب مناسبی برای $\int_W f$ باشد. با میل دادن ضخامت افراز به صفر، می‌توان نشان داد که مجموع فوق به انتگرال $\int_{W'} (f \circ \varphi) |\det D\varphi|$ میل می‌کند.

(۳۴-۴) انتگرال‌های ناسره در حالت یک بعدی علی‌الاصول دو نوع "انتگرال ناسره" وجود دارد که همه انتگرال‌های ناسره مجموعی از آنها هستند. یک مورد حالتی است که تابع کراندار می‌ماند ولی بازه انتگرال‌گیری از یک سو بی‌کران می‌شود. این انتگرال‌ها به $\int_a^\infty f$ یا $\int_{-\infty}^a f$ نمایش داده می‌شوند. مثلاً $\int_a^\infty f = \lim_{b \rightarrow \infty} (\int_a^b f)$ نوع دیگر انتگرال ناسره به وضعیتی اطلاق می‌شود که بازه انتگرال‌گیری کراندار است ولی وقتی متغیر به یکی از دو نقطه انتهایی بازه میل می‌کند تابع بی‌کران می‌شود. مثلاً اگر $f(x) \rightarrow \infty$ وقتی $x \rightarrow b^-$ ، $\int_a^b f$ در صورت وجود، برابر $\int_a^{b-\epsilon} f$ است. در حالت $n \geq 2$ شیوه بی‌کران شدن ناحیه انتگرال‌گیری بی‌نهایت تنوع دارد، و نیز به جای فقط دو نقطه مرزی، مرز یک ناحیه کراندار ممکن است دارای بی‌نهایت نقطه باشد که تابع ممکن است در نزدیکی هر یک بی‌کران شود. بدین ترتیب موضوع انتگرال ناسره برای توابع چند متغیری می‌تواند بسیار پیچیده‌تر شود. در اینجا توجه خود را به ناحیه‌های نوع زیر محدود می‌کنیم:

(۱) یک ناحیه مجاز انتگرال‌گیری که از آن تعدادی متناهی نقطه حذف شده باشد.

(۲) \mathbb{R}^n یا $\mathbb{R}^n - \{p_1, \dots, p_k\}$ ، یعنی \mathbb{R}^n که از آن تعدادی متناهی نقطه حذف شده باشد.

(۳) بخشی X از \mathbb{R}^n که محصور به تعدادی متناهی ابرصفحه در \mathbb{R}^n باشد، یا $X - \{p_1, \dots, p_k\}$

برای نوع (۳) می‌توان مثلاً ناحیه $\{(x_1, \dots, x_n) \mid x_i \geq 0, i = 1, \dots, n\}$ را مثال زد. علاوه بر این محدودیت، فقط تابع‌های غیرمنفی را در نظر می‌گیریم (مشابه‌ها می‌توان توابع غیرمثبت را در نظر

گرفت). قضیه زیر را بدون اثبات مورد استفاده قرار می دهیم.

(۳۴-۵) قضیه فرض کنید D یک ناحیه در \mathbb{R}^n از نوع بالا باشد و $(D_n)_{n=1}^{\infty}$ یک دنباله از زیرمجموعه های بسته، کراندار و مجاز انتگرال گیری از \mathbb{R}^n به طوری که:

$$D_1 \subset D_2 \subset D_3 \subset \dots, \quad \bigcup_{n=1}^{\infty} D_n = D$$

اگر $f: D \rightarrow \mathbb{R}$ یک تابع پیوسته و نامنفی باشد و $\lim_{n \rightarrow \infty} \int_{D_n} f$ وجود داشته باشد، آنگاه برای هر دنباله $(D'_n)_{n=1}^{\infty}$ دیگر با شرایط ذکر شده برای $(D_n)_{n=1}^{\infty}$ وجود دارد و برابر $\lim_{n \rightarrow \infty} \int_{D_n} f$ است.

این حد مشترک را انتگرال ناسره f روی D می نامیم و در صورت وجود حد می گوئیم $\int_D f$ همگراست. لازم به ذکر است که اگر f همه جا غیرمنفی (یا همه جا غیرمثبت) نباشد، شکل نواحی D_n می تواند بر نتیجه نهایی، یعنی حد $\int_{D_n} f$ اثر بگذارد. این پدیده در حالت یک متغیری نیز دیده می شود. مثلاً برای تابع $f(x) = \frac{1}{x}$ که روی بازه محذوف $\{0\} - [1, -1]$ تعریف شده و پیوسته است، اگر D_n را برابر $[\frac{1}{n}, 1] \cup [-1, -\frac{1}{n}]$ بگیریم، داریم $\int_{D_n} f = 0$ و در نتیجه حد این انتگرال ها صفر است. ولی برای $D'_n = [-1, -\frac{1}{n}] \cup [\frac{1}{n}, 1]$ داریم $\int_{D'_n} f = \ln 2$ که حد آن نیز $\ln 2$ است.

مثال در مورد همگرایی انتگرال ناسره $\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dA$ بحث کنید.

حل D_n را گوی بسنه شعاع n حول مبدأ می گیریم، $D_n = \mathbb{R}^2$. داریم:

$$\begin{aligned} \iint_{D_n} e^{-(x^2+y^2)} dA &= \int_0^{2\pi} \int_0^n e^{-r^2} r dr d\theta \\ &= (2\pi) \left(-\frac{1}{r}\right) e^{-r^2} \Big|_0^n \\ &= \pi(1 - e^{-n^2}) \end{aligned}$$

وقتی $n \rightarrow \infty$ ، مقدار فوق به π میل می کند، پس $\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dA$ همگراست و برابر π می باشد.

کاربرد نشان می دهیم $\int_{-\infty}^{\infty} e^{-x^2} dx$ همگراست و برابر $\sqrt{\pi}$ می باشد. مقصود از $\int_{-\infty}^{\infty} e^{-x^2} dx$

$\lim_{a \rightarrow \infty} \int_{-a}^a e^{-x^2} dx$ است هر چند که به سبب مثبت بودن e^{-x^2} نتیجه می شود که لازم نیست حدود انتگرال لزوماً متقارن گرفته شوند. توجه کنید که $\int_{-a}^a e^{-x^2} dx = \int_{-a}^a e^{-y^2} dy$ پس اگر $\int_{-a}^a e^{-x^2} dx$ را I_a بنامیم، داریم $\int_{-a}^a \int_{-a}^a e^{-(x^2+y^2)} dx dy = I_a^2$ ناحیه انتگرال گیری مربع $-a \leq x, y \leq a$ است. اگر این مربع را D'_a بنامیم، داریم $D'_1 \subset D'_2 \subset D'_3 \subset \dots$ و $\bigcup_{n=1}^{\infty} D'_n = \mathbb{R}^2$ ، پس طبق قضیه و مثال قبل $\lim_{n \rightarrow \infty} I_n^2 = \pi$. نتیجه این که $\lim_{n \rightarrow \infty} I_n = \sqrt{\pi}$. این حد در نظریه احتمال زیاد به کار می رود.

رویه‌های هموار

در بخش‌های ۱۲، ۱۳ و ۱۴ به بررسی نسبتاً مسبوطی از 'خمهای هموار' پرداختیم. در اینجا می‌خواهیم به توصیف اشیاء هموار دوبعدی احتمالاً خمیده (غیر مسطح) بپردازیم. به این منظور نخست مفهوم 'رویه پارامتری' را مطرح می‌کنیم، سپس شرطی را که متضمن خاصیت 'هموار بودن' است به تعریف می‌افزاییم. برای توصیف رویه که برداشت ذهنی ما از آن یک شی دوبعدی است دو پارامتر مورد نیاز است همانگونه که برای توصیف خم (یک بعدی) یک پارامتر t مورد استفاده قرار می‌گرفت.

اگر W یک ناحیه در صفحه uv باشد، تابع

$$\vec{r} = \varphi(u, v) \quad , \quad \varphi : W \rightarrow \mathbb{R}^3$$

یک رویه پارامتری در \mathbb{R}^3 تعریف می‌کند. صفات پیوسته، مشتق‌پذیر، مشتق‌پذیر با مشتقات پاره‌ای پیوسته (اصطلاحاً C^1) وقتی به رویه پارامتری اطلاق شوند مقصود این است که تابع φ از این صفات برخوردار است. البته $\varphi(u, v)$ دارای سه مؤلفه (x, y, z) است، $x = \varphi_1(u, v)$ ، $y = \varphi_2(u, v)$ ، $z = \varphi_3(u, v)$ و صفت مربوط باید برای هر یک از تابع‌های φ_1 ، φ_2 و φ_3 برقرار باشد. شکل ۱ تجسم ما از رویه‌های پارامتری را نشان می‌دهد انتظار داریم خطوط راست (ثابت u) و (ثابت v) تحت اثر φ به خم‌هایی روی تصویر φ منتقل شوند و نوعی "دستگاه مختصات خمیده" روی $\varphi(W)$ بسازند

شکل ۱

این انتظار در حالت کلی کمی خوشبینانه است زیرا که به طور کلی تصویر φ ممکن است "دوبعدی" نشود، مثلاً φ ممکن است کل W را به یک پاره‌خط یا حتی یک نقطه بنگارد (تابع ثابت). در مورد خم‌ها، افزودن شرط پیوستگی مشتق و ناصفر ماندن بردار سرعت، این تضمین را ایجاد کرد که تصویر

خم پارامتری "یک بعدی و هموار" است. در اینجا باید شرط مشابهی را جستجو کنیم. مشتق‌های پاره‌ای φ نسبت به u و v را به ترتیب به \vec{r}_u و \vec{r}_v نمایش می‌دهیم. بدین ترتیب:

$$\vec{r}_u = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right), \quad \vec{r}_v = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right)$$

\vec{r}_u و \vec{r}_v تعبیر هندسی جالب توجهی دارند. با ثابت نگاهداشتن $v = v_0$ و مشتق‌گیری نسبت به u ، بردار سرعت خم $\varphi(u, v_0)$ را که روی $\varphi(u, v)$ قرار دارد در نظر گرفته‌ایم، و به همین ترتیب \vec{r}_v بردار سرعت خم $\varphi(u_0, v)$ است. پس \vec{r}_u و \vec{r}_v بردارهای مماس بر $\varphi(W)$ در نقطه $\varphi(u, v)$ تلقی می‌شوند. اگر \vec{r}_u و \vec{r}_v همراستا نباشند، انتظار داریم u و v دو جهت مستقل روی رویه $\varphi(W)$ را مدرج کنند و بدین ترتیب شیئی "دو بعدی" حاصل شود. با این زمینه، رویه پارامتری $\varphi: W \rightarrow \mathbb{R}^3$ را هموار می‌نامیم در صورتی که شرایط زیر برقرار باشند:

الف) φ مشتق‌پذیر با مشتق‌های پاره‌ای پیوسته باشد.

ب) به‌ازای هر (u, v) در W داشته باشیم $\vec{r}_u \times \vec{r}_v \neq \mathbf{0}$ (معادلاً \vec{r}_u و \vec{r}_v همراستا نباشند).

ضمناً توجه کنید که چون $\vec{r}_u \times \vec{r}_v$ بر \vec{r}_u و \vec{r}_v عمود است، و \vec{r}_u, \vec{r}_v دو جهت مماس مستقل بر رویه را نشان می‌دهند، بردار واحد $\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}$ بر $\varphi(W)$ عمود خواهد بود. \vec{n} را قائم واحد قراردادی رویه هموار φ می‌نامیم.

برای توجیه اصطلاح هموار، نشان می‌دهیم که تحت این شرط، اگر (u_0, v_0) نقطه‌ای از W باشد، آنگاه تحت شرط هموار بودن، گوی بازی D حول (u_0, v_0) وجود دارد که تصویر آن تحت φ ، در واقع نموداریک تابع مشتق‌پذیر دو متغیری است. اگر ماتریس ژاکوبی φ (مشتق φ) در نقطه (u_0, v_0) را در نظر بگیریم، داریم:

$$D\varphi(u_0, v_0) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix}$$

یعنی \vec{r}_u و \vec{r}_v ستون‌های این ماتریس هستند. همراستا نبودن \vec{r}_u و \vec{r}_v دلالت بر این می‌کند که تصویرهای قائم این دو بردار بردست کم یکی از سه صفحه مختصات xy ، yz یا zy همراستا نیستند.

برای صراحت فرض کنید تصویر این دو بر صفحه xy همراستا نباشند، یعنی $(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u})$ و $(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v})$ همراستا نباشند. نتیجه اینکه

$$\det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}_{(u_0, v_0)} \neq 0$$

طبق قضیه تابع وارون مجموعه باز D' حول (x_0, y_0) (تصویر قائم $\varphi(u_0, v_0)$ روی صفحه xy) وجود دارد و مجموعه باز D حول (u_0, v_0) به طوری که φ یک تناظر یک به یک میان D و D' برقرار می کند و وارون آن، $\varphi^{-1}: D' \rightarrow D$ نیز مشتق پذیر است. بدین ترتیب (u, v) در D تابعی مشتق پذیر از (x, y) در D' می شود. نتیجه این که $z = \varphi_3(u, v)$ نیز تابعی مشتق پذیر از (x, y) است،

$$z = (\varphi_3 \circ \varphi^{-1})(x, y)$$

پس نقاط $\varphi(u, v) = (x, y, z)$ برای (u, v) نزدیک (u_0, v_0) به صورت $(x, y, (\varphi_3 \circ \varphi^{-1})(x, y))$ یعنی نمودار تابع مشتق پذیر $\varphi_3 \circ \varphi^{-1}$ قابل نمایش اند. بدین مفهوم شرط های (الف) و (ب) مفهوم شهودی هموار بودن را برقرار می کنند.

استدلال بالا ضمناً نشان می دهد که برای هر $w_0 = (u_0, v_0)$ در W ، دامنه تعریف یک رویه هموار، مجموعه بازی حول w_0 وجود دارد که تحدید φ به آن یک به یک است، هر چند که رویه هموار ممکن است نهایتاً خود را قطع کند. نکته دیگر این که \vec{r}_u و \vec{r}_v ، ستون های $D\varphi(u, v)$ ، تصاویر $e_1 = (1, 0)$ و $e_2 = (0, 1)$ تحت اثر تقریب خطی هستند، بنابراین ضریب تغییر مساحت در نقطه (u, v) برابر مساحت متوازی الاضلاع تعریف شده توسط \vec{r}_u و \vec{r}_v ، یعنی $|\vec{r}_u \times \vec{r}_v|$ است (مراجعه کنید به بحث تعویض متغیر در انتگرال)، از این رو $dS = |\vec{r}_u \times \vec{r}_v| du dv$ را عنصر سطح رویه هموار φ می نامند. اگر D زیرمجموعه ای از W باشد که روی آن φ یک به یک است و D یک ناحیه مجاز انتگرال گیری در صفحه uv باشد، تعریف می کنیم:

$$\varphi(M) \text{ مساحت} = \iint_D dS = \iint_D |\vec{r}_u \times \vec{r}_v| du dv$$

مثال ۱ اگر W یک زیرمجموعهٔ باز \mathbb{R}^2 باشد و $f: W \rightarrow \mathbb{R}$ یک تابع دارای مشتق‌های پاره‌ای پیوسته، نمودار f را می‌توان به صورت یک رویهٔ هموار پارامتری کرد:

$$\varphi: W \rightarrow \mathbb{R}^3, \quad \varphi(u, v) = (u, v, f(u, v))$$

$$\text{داریم } \vec{r}_u \times \vec{r}_v = \left(-\frac{\partial f}{\partial u}, -\frac{\partial f}{\partial v}, 1\right), \vec{r}_v = \left(0, 1, \frac{\partial f}{\partial v}\right), \vec{r}_u = \left(1, 0, \frac{\partial f}{\partial u}\right)$$

$$dS = \sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2}, \quad \vec{n} = \frac{\left(-\frac{\partial f}{\partial u}, -\frac{\partial f}{\partial v}, 1\right)}{\sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2}}$$

به کمک dS بالا می‌توانیم مساحت نمودارها را محاسبه کنیم.

مثال ۱-۱ سطح جانبی جسم به دست آمده از اشتراک استوانه‌های $x^2 + y^2 \leq 1$ و $x^2 + z^2 \leq 1$ را محاسبه کنید.

برای این کار توجه کنید که به سبب تقارن، مساحت مورد نظر چهار برابر مساحت نمودار $z = \sqrt{1-x^2}$ روی گوی $x^2 + y^2 \leq 1$ است. داریم $\frac{\partial z}{\partial x} = \frac{-x}{\sqrt{1-x^2}}$ ، $\frac{\partial z}{\partial y} = 0$ پس مساحت نمودار برابر است با:

$$\begin{aligned} \int \int_{x^2+y^2 \leq 1} \sqrt{1 + \frac{x^2}{1-x^2}} dA &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\sqrt{1-x^2}} dy dx \\ &= \int_{-1}^1 2 dx = 4 \end{aligned}$$

بنابراین سطح جانبی مورد نظر برابر ۱۶ است. همین نتیجه را از روش انتگرال‌گیری روی خم نیز به دست آورده بودیم.

مثال ۲ کره شعاع R حول z در \mathbb{R}^3 را به کمک مختصات کروی پارامتری می‌کنیم. تعریف می‌کنیم:

$$\varphi(u, v) = (R \sin u \cos v, R \sin u \sin v, R \cos u)$$

$$0 \leq u \leq \pi, \quad 0 \leq v \leq 2\pi$$

در اینجا مختصه‌های کروی φ و θ به ترتیب به u و v نمایش داده شده‌اند. داریم:

$$\vec{r}_u = (R \cos u \cos v, R \cos u \sin v, -R \sin u), \quad \vec{r}_v = (-R \sin u \sin v, R \sin u \cos v, 0)$$

$$\vec{r}_u \times \vec{r}_v = (R^2 \sin^2 u \cos v, R^2 \sin^2 u \sin v, R^2 \sin u \cos v), \quad dS = R^2 \sin u du dv$$

توجه کنید که شرط هموار بودن به ازای $u = 0, \pi$ (دو قطب کره) نقض می شود. این پدیده فقط ناشی از روش پارامتری کردن است و ارتباطی با شکل کره ندارد. توجه کنید که در دو قطب خطوط مختصاتی، یعنی تصاویر ثابت $u = 0$ و ثابت $v = 0$ (مدارها و نصف النهارها) از حالت متقاطع بودن خارج می شوند زیرا تصویر هر یک از $u = 0$ و $u = \pi$ یک تک نقطه است. با این حال، برای منظورهایی انتگرال گیری، چنانچه تابعی در سراسر کره انتگرال پذیر باشد، دو خط $u = 0$ و $u = \pi$ که اندازه دو بعدی صفر در مستطیل $0 \leq u \leq \pi$ ، $0 \leq v \leq 2\pi$ ، دارند، بی اثر هستند.

مثال ۳ (چنبره دوار) برای ثابت های $a > b > 0$ ، تعریف می کنیم:

$$\vec{r} = \varphi(u, v) = ((a + b \cos u) \cos v, (a + b \cos u) \sin v, b \sin u)$$

$$0 \leq u \leq 2\pi, \quad 0 \leq v \leq 2\pi$$

این رویه از دوران دایره شعاع b به مرکز $(a, 0)$ در صفحه xz حول محور z به دست می آید. تعبیر زاویه های u و v در شکل ۲ نمایش داده شده است. دایره اولیه را مدرج می کند و v در واقع مختصه θ در مختصات استوانه ای است، داریم:

$$\begin{cases} \vec{r}_u = ((-b \sin u) \cos v, (-b \sin u) \sin v, b \cos u) \\ \vec{r}_v = (-(a + b \cos u) \sin v, (a + b \cos u) \cos v, 0) \end{cases}$$

$$\vec{r}_u \times \vec{r}_v = (-b(a + b \cos u) \cos u \cos v, -b(a + b \cos u) \cos u \sin v, -b \sin u(a + b \cos u))$$

$$|\vec{r}_u \times \vec{r}_v| = b(a + b \cos u), \quad dS = b(a + b \cos u) du dv$$

توجه کنید که شرط هموار بودن همه جا برقرار است. مساحت این چنبره برابر می شود با:

$$\int_0^\pi \int_0^{2\pi} b(a + b \cos u) du dv = (2\pi b)(2\pi a) = 4\pi^2 ab$$

شکل ۲

مثال ۴ برای $a, b, c > 0$ داده شده تعریف می کنیم:

$$\vec{r} = \varphi(u, v) = (a \cosh u \cos v, b \cosh u \sin v, c \sinh u)$$

$$0 \leq u \leq 1, \quad 0 \leq v \leq 2\pi$$

اگر سه مؤلفه $\varphi(u, v)$ را به ترتیب x, y, z بخوانیم، خواهیم داشت:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

پس تصویر جزیی از هذلولی واریکپارچه $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ است، در واقع قسمتی که برای آن $0 \leq z \leq c \sinh 1$ داریم (شکل ۳).

$$\vec{r}_u = (a \sinh u \cos v, b \sinh u \sin v, c \cosh u), \vec{r}_v = (-a \cosh u \sin v, b \cosh u \cos v, 0)$$

$$\vec{r}_u \times \vec{r}_v = (-bc \cosh^2 u \cos v, -ac \cosh^2 u \sin v, ab \sinh u \cosh u) \neq \underline{0}$$

پس شرط هموار بودن برقرار است.

شکل ۳

در مثال‌های بالا عمدتاً شرط هموار بودن نقض نمی‌شد و در یک مورد (کره) نقض این شرط تعبیر هندسی نداشت، بلکه یک ویژگی روش پارامتری کردن را منعکس می‌کرد. در زیر چند مثال از ناهنجاری‌های هندسی که ممکن است در اثر نقض شرط هموار بودن پیش آید نشان می‌دهیم:

مثال ۵ (الف) برای a, b, c داده شده تعریف کنید

$$\varphi(u, v) = (a, b, c), \quad (u, v) \in \mathbb{R}^2$$

در اینجا به وضوح $\vec{r}_u = \vec{r}_v = \underline{0}$ و شرط هموار بودن همه‌جا نقض می‌شود. تصویر φ در واقع یک تک نقطه‌ای است و یک "رویه" به مفهوم دو بعدی آن نیست.

(ب) برای تابع $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ که دارای مشتق‌های پاره‌ای پیوسته است، تعریف کنید:

$$\varphi(u, v) = (f(u, v), f(u, v), f(u, v))$$

تصویر φ بخشی از خط راست $x = y = z$ است، یعنی یک بعدی است و با شهود رویه دو بعدی

$$\vec{r}_u \times \vec{r}_v = \underline{0} \text{ و } \vec{r}_v = \left(\frac{\partial f}{\partial v}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial v} \right), \vec{r}_u = \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial u}, \frac{\partial f}{\partial u} \right)$$

(ج)

$$\vec{r} = \varphi(u, v) = (u \cos v, u \sin v, u), \quad (u, v) \in \mathbb{R}^2$$

داریم $(\vec{r}_u = (\cos v, \sin v, 1), \vec{r}_v = (-u \sin v, u \cos v, 0))$ پس

$$\vec{r}_u \times \vec{r}_v = (-u \cos v, -u \sin v, u)$$

به‌ازای $u = 0$ شرط هموار بودن نقض می‌شود. در واقع تصویر φ مخروط $x^2 + y^2 = z^2$ است و $u = 0$ متناظر با رأس مخروط می‌باشد که در آن نقطه مخروط صفحه مماس ندارد

شکل ۳

(د) به مثال ۴، بخش ۱۲، $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$ ، نگاه کنید که در آن تصویر γ نمودار $y = |x|$ را به صورت مشتق‌پذیر با مشتق پیوسته پارامتری می‌کند ولی $\gamma'(0) = 0$. حال تعریف کنید:

$$\vec{r} = \varphi(u, v) = (u, \gamma_1(v), \gamma_2(v)) \quad , \quad (u, v) \in \mathbb{R}^2$$

که در آن $\gamma(v) = (\gamma_1(v), \gamma_2(v))$ داریم:

$$\vec{r}_u = (1, 0, 0) \quad , \quad \vec{r}_v = (0, \gamma'_1(v), \gamma'_2(v)) \quad , \quad \vec{r}_u \times \vec{r}_v = (0, -\gamma'_2(v), \gamma'_1(v))$$

برای $v = 0$ ، مستقل از u ، شرط هموار بودن نقض می‌شود. $v = 0$ در شکل ۴ متناظر با خط "تا شدن" رویه است که در سراسر آن صفحه مماس بر تصویر رویه وجود ندارد. در اینجا لازم است یک نمایش معمول دیگر برای $dS = |\vec{r}_u \times \vec{r}_v| dudv$ را نیز معرفی کنیم. داریم:

$$|\vec{r}_u \times \vec{r}_v|^2 = (\vec{r}_u \times \vec{r}_v) \cdot (\vec{r}_u \times \vec{r}_v)$$

طبق اتحاد لاگرانژ

$$(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}) \quad (1)$$

می‌توان نوشت:

$$|\vec{r}_u \times \vec{r}_v|^2 = (\vec{r}_u \cdot \vec{r}_u)(\vec{r}_v \cdot \vec{r}_v) - (\vec{r}_u \cdot \vec{r}_v)^2$$

معمولاً $\vec{r}_u \cdot \vec{r}_u$ را به E ، $\vec{r}_u \cdot \vec{r}_v$ را به F و $\vec{r}_v \cdot \vec{r}_v$ را به G نمایش می‌دهند (نمادگذاری گاوس)، پس:

$$dS = \sqrt{EG - F^2} dudv \quad (2)$$

یادداشت تصویر یک رویه هموار ممکن است پرمایش‌های گوناگون به عنوان رویه هموار داشته باشد، می‌توان «بازپرمایش» را همچنان که در مورد خم‌های هموار مطرح کردیم در اینجا نیز مطرح کنیم. در ساده‌ترین حالت فرض کنید W و W' دو مجموعه باز و همبند مسیری در \mathbb{R}^2 باشند، نقاط W را به (u, v) و نقاط W' را به (u', v') نمایش دهید. فرض کنید $\tau: W' \rightarrow W$ یک نگاشت یک به یک، پوشا و دارای مشتق‌های پاره‌ای پیوسته باشد که به‌ازای هر $w' = (u', v')$ در W' داشته باشیم $\det D\tau(w') \neq 0$. اگر $\varphi: W \rightarrow \mathbb{R}^3$ یک رویه هموار تعریف کند، $\varphi \circ \tau: W' \rightarrow \mathbb{R}^3$ نیز یک رویه هموار است و همان مجموعه $\varphi(W)$ را توصیف می‌کند. چون W' همبند مسیری فرض شده است، $\det D\tau(w')$ برای همه نقاط w' از W' یا مثبت یا منفی است (در غیر این صورت اگر مسیر پیوسته‌ای بین دو نقطه با علامت مختلف در نظر بگیریم، به سبب پیوستگی $\det D\tau(w')$ ، این کمیت باید جایی در طول مسیر صفر شود که خلاف فرض است). در حالت اول τ را یک بازپرمایش جهت‌نگهدار و در حالت دوم τ را یک بازپرمایش جهت‌برگردان می‌نامند. برای بازپرمایش‌های جهت‌نگهدار، قائم واحد قراردادی تغییر نمی‌کند، ولی برای بازپرمایش‌های جهت‌برگردان، \vec{n} وارونه می‌شود. پس در واقع یک رویه هموار در \mathbb{R}^3 دو 'جهت' می‌پذیرد که با انتخاب جهت قائم واحد مشخص می‌شود.

انتگرال روی خم و رویه (۱)

تاکنون مفهوم انتگرال فقط روی مجموعه‌های "مسطح" مطرح بوده است: یک بازه در \mathbb{R} ، یک ناحیه در صفحه و به طور کلی یک ناحیه توپر در \mathbb{R}^n . نمونه اشیاء خمیده، یک خم در \mathbb{R}^n ، یک رویه در \mathbb{R}^2 ، یا یک مجموعه تراز یک تابع غیرخطی $\mathbb{R}^n \rightarrow \mathbb{R}$ است. مفهوم انتگرال روی اینگونه مجموعه‌ها کاربردهای اساسی هم در درون ریاضیات و هم در بیرون آن دارد. در باقیمانده درس مفهوم انتگرال روی خم‌ها و رویه‌ها محور اصلی بحث خواهد بود.

دو نوع انتگرال در اینجا مورد بحث قرار خواهد گرفت. نوع اول که به 'انتگرال چگالی' معروف است برای خم γ به $\int_{\gamma} f ds$ و برای رویه φ به $\int_{\varphi} f dS$ نمایش داده خواهد شد. در اینجا f تابعی است، معمولاً پیوسته، که روی تصویر خم یا رویه تعریف شده است. در حالتی که f مثبت باشد، اگر آن را به چگالی تصویر خم یا تصویر رویه تعبیر کنیم، مقدار انتگرال مربوط برابر جرم تصویر خواهد بود. مقدار این انتگرال به جهت خم یا رویه بستگی ندارد. نوع دوم انتگرال، که محور اصلی بحث‌های بعدی خواهد بود، 'انتگرال یک میدان برداری'، یا 'انتگرال یک فرم دیفرانسیل' نام دارد و با تعویض جهت یا رویه مقدار آن تغییر علامت می‌دهد.

(۳۶-۱) انتگرال چگالی روی خم یک خم هموار $\gamma : [a, b] \rightarrow \mathbb{R}^n$ در نظر بگیرید و فرض کنید f تابعی با مقدار حقیقی است که دامنه آن شامل تصویر γ می‌باشد. معمولاً با f پیوسته سروکار داریم. طبق معمول پارامتر طول خم را به s نمایش می‌دهیم. تعریف می‌کنیم:

$$\int_{\gamma} f ds = \int_a^b f(\gamma(t)) \frac{ds}{dt} dt \quad (1)$$

طرف چپ این تعریف به پارامتر خاص t بستگی ندارد ولی طرف راست ظاهراً به t وابسته است. برای این که این تعریف معنی داشته باشد، باید نحوه وابستگی سمت راست به پارامتر t را بررسی کنیم.

فرض کنید $\varphi : [\alpha, \beta] \rightarrow [a, b]$ یک بازپرمایش γ را تعریف کند، $t = \varphi(\tau)$ داریم:

$$\begin{aligned} \int_a^b f(\gamma(t)) \frac{ds}{dt} dt &= \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} (f \circ \gamma)(\varphi(\tau)) \frac{ds}{dt}(\varphi(\tau)) \frac{dt}{d\tau} d\tau \\ &= \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} (f \circ \gamma)(\varphi(\tau)) \frac{ds}{d\tau} d\tau \end{aligned}$$

اگر φ جهت‌نگهدار باشد، داریم $\varphi^{-1}(a) = \alpha$ و $\varphi^{-1}(b) = \beta$ پس عبارت سمت راست برابر $\int_{\alpha}^{\beta} f((\gamma \circ \varphi)(\tau)) \frac{ds}{d\tau} d\tau$ می‌شود. اگر φ جهت برگردان باشد، $\varphi^{-1}(a) = \beta$ و $\varphi^{-1}(b) = \alpha$ و در نتیجه عبارت سمت راست برابر می‌شود با $-\int_{\alpha}^{\beta} f((\gamma \circ \varphi)(\tau)) \frac{ds}{d\tau} d\tau$ که در اینجا s پارامتر طول خمیده اولیه یعنی γ است. اگر پارامتر طول خم معکوس را به \bar{s} نمایش دهیم، داریم $\frac{d\bar{s}}{dt} = -\frac{ds}{dt}$ ، پس مجدداً مقدار انتگرال نسبت به پارامتر جدید، یعنی $\int_{\alpha}^{\beta} f((\gamma \circ \varphi)(t)) \frac{d\bar{s}}{dt} dt$ همان مقدار سابق است. برای $\int_{\gamma} f ds$ دو تعبیر زیر را مطرح می‌کنیم.

(۳۶-۱-۱) فرض کنید $\gamma : [a, b] \rightarrow \mathbb{R}^n$ یک به یک باشد. اگر f تابعی مثبت باشد و تصویر γ را یک مفتول نازک به چگالی $f(\gamma(t))$ در نقطه متناظر با t فرض کنیم، $\int_{\gamma} f ds$ جرم کل مفتول تلقی می‌شود.

(۳۶-۱-۲) در حالت خاصی که تصویر خم یک به یک $\gamma : [a, b] \rightarrow \mathbb{R}^3$ در یک صفحه مسطح P قرار گیرد، و f مثبت باشد، می‌توان به $\int_{\gamma} f ds$ یک تعبیر مساحت به صورت زیر نسبت داد. روی تصویر خم γ و عمود بر صفحه P ، در هر نقطه $\gamma(t)$ عمودی در یک طرف صفحه به ارتفاع $f(\gamma(t))$ در نظر بگیرید. این پاره‌خط‌های عمود رویه‌ای عمود بر P می‌سازند که مساحت آن برابر $\int_{\gamma} f ds$ است (شکل ۱).

شکل ۱

مثال مساحت سطح جانبی جسمی را که از تقاطع دو استوانه $x^2 + y^2 = 1$ و $x^2 + z^2 = 1$ ساخته می‌شود محاسبه کنید.

حل قبلاً حجم داخل این جسم را محاسبه کرده‌ایم. سطح جانبی این جسم از دو قسمت متساوی تشکیل شده است که یک بخش آن روی دایره $x^2 + y^2 = 1$ ساخته شده و متشکل از پاره‌خط‌هایی

است که روی استوانه $x^2 + y^2 = 1$ قرار دارند و توسط $x^2 + z^2 = 1$ بریده شده‌اند. روی نقطه $(x, y, 0)$ از دایره $x^2 + y^2 = 1$ ، این پاره‌خط با $-\sqrt{1-x^2} \leq z \leq \sqrt{1-x^2}$ مشخص می‌شود که طول آن $2\sqrt{1-x^2}$ است. بنابراین تقارن، بخش دیگر سطح جانبی نیز وضعیت مشابه دارد (با تعویض نقش y و z). بدین ترتیب:

$$\text{سطح جانبی} = 2 \times \int_{\gamma} 2\sqrt{1-x^2} ds$$

که در اینجا γ دایره $x^2 + y^2 = 1$ را نمایش می‌دهد که یک بار در جهت مثلثاتی طی شده است. با استفاده از پرمایش $x = \cos t, y = \sin t, 0 \leq t \leq 2\pi, ds = dt$ داریم:

$$\text{سطح جانبی} = 4 \times \int_0^{2\pi} |\sin t| dt = 8 \times \int_0^{\pi} \sin t dt = 16$$

(۳۶-۱) انتگرال چگالی روی رویه فرض کنید $\varphi: W \rightarrow \mathbb{R}^3$ یک رویه هموار باشد

و $\vec{r} = \varphi(u, v)$ یک ناحیه مجاز انتگرال‌گیری در صفحه (u, v) . اگر f تابعی با مقدار حقیقی باشد که دامنه آن شامل $\varphi(W)$ است، انتگرال $\int_{\varphi} f dS$ به صورت زیر تعریف می‌شود:

$$\int_{\varphi} f dS = \int_W \int f(\varphi(u, v)) |\vec{r}_u \times \vec{r}_v| dudv \quad (2)$$

برای اینکه طرف راست (۲) وجود داشته باشد، تابع $f \circ \varphi$ باید روی ناحیه مجاز W انتگرال پذیر باشد. اگر f پیوسته باشد، ترکیب $f \circ \varphi$ نیز پیوسته است، $|\vec{r}_u \times \vec{r}_v|$ نیز پیوسته می‌باشد زیرا که مشتقات پاره‌ای φ نسبت به u و v پیوسته فرض شده‌اند، پس انتگرال طرف راست وجود دارد.

اکنون به کمک قاعده زنجیره‌ای و فرمول تعویض متغیر انتگرال دوگانه نشان می‌دهیم که طرف راست (۲) تحت بازپرمایش پا برجا می‌ماند. فرض کنید $\Psi: W' \rightarrow W$ تابعی یک‌به‌یک و C^1 از ناحیه مجاز انتگرال‌گیری W' به W است به طوری که دترمینان ماتریس ژاکوبی Ψ همه جا ناصفر است. نقطه $w' = (u', v')$ را به $w = (x, y, z)$ نمایش می‌دهیم و می‌نویسیم $\vec{r} = (x, y, z)$. از قاعده زنجیره‌ای داریم:

$$\begin{bmatrix} y_u & y_v \\ z_u & z_v \end{bmatrix} = \begin{bmatrix} y_{u'} & y_{v'} \\ z_{u'} & z_{v'} \end{bmatrix} \begin{bmatrix} u'_u & u'_v \\ v'_u & v'_v \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} z_u & z_v \\ x_u & x_v \end{bmatrix} = \begin{bmatrix} z_{u'} & z_{v'} \\ x_{u'} & x_{v'} \end{bmatrix} \begin{bmatrix} u'_u & u'_v \\ v'_u & v'_v \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} = \begin{bmatrix} x_{u'} & x_{v'} \\ y_{u'} & y_{v'} \end{bmatrix} \begin{bmatrix} u'_u & u'_v \\ v'_u & v'_v \end{bmatrix} \quad (5)$$

با توجه به اینکه دترمینان تمرینهای طرف چپ مؤلفه‌های $\vec{r}'_u \times \vec{r}'_v$ و دترمینان ماتریسهای سمت چپ طرف راست مؤلفه‌های $\vec{r}'_{u'} \times \vec{r}'_{v'}$ هستند، داریم

$$|\vec{r}'_u \times \vec{r}'_v| = |\vec{r}'_{u'} \times \vec{r}'_{v'}| \cdot \left| \det \frac{\partial(u', v')}{\partial(u, v)} \right| \quad (6)$$

با جایگزینی (6) در (2) و با استفاده از فرمول تعویض متغیر نتیجه می‌شود که:

$$\int_W \int f(\varphi(u, v)) |\vec{r}'_u \times \vec{r}'_v| du dv = \int_{W'} \int f(\varphi \circ \Psi(u', v')) |\vec{r}'_{u'} \times \vec{r}'_{v'}| du' dv'$$

پس تعریف (2) تحت بازپرمایش تغییر نمی‌کند.

تعبیر جرم و چگالی در اینجا نیز مطرح می‌شود. اگر φ یک‌به‌یک باشد و به ازای هر $(u, v) \in W$ ، $f(\varphi(u, v)) > 0$ چگالی روبه در نقطه $\varphi(u, v)$ تلقی شود، طبیعی است که مقدار $\int_{\varphi} \int f dS$ را جرم کل $\varphi(W)$ تلقی کنیم. مفهوم مرکز ثقل یک مفتول نازک (تصویر یک خم) و مرکز ثقل یک پیوسته نازک (تصویر یک روبه) را نیز اکنون می‌توان به شیوه معمول تعریف کرد. برای خم $\gamma: [0, b] \rightarrow \mathbb{R}^3$ و روبه $\varphi: W \rightarrow \mathbb{R}^3$ چگالی را به δ نمایش می‌دهیم. مختصات مرکز ثقل، $(\bar{x}, \bar{y}, \bar{z})$ به صورت زیر تعریف می‌شوند:

برای خم γ :

$$\bar{x} = \frac{\int_{\gamma} x \delta ds}{\int_{\gamma} \delta ds}, \quad \bar{y} = \frac{\int_{\gamma} y \delta ds}{\int_{\gamma} \delta ds}, \quad \bar{z} = \frac{\int_{\gamma} z \delta ds}{\int_{\gamma} \delta ds} \quad (7)$$

و برای روبه φ :

$$\bar{x} = \frac{\int_{\gamma} \int y \delta dS}{\int_{\gamma} \int \delta dS}, \quad \bar{y} = \frac{\int_{\gamma} \int x \delta dS}{\int_{\gamma} \int \delta dS}, \quad \bar{z} = \frac{\int_{\gamma} \int z \delta dS}{\int_{\gamma} \int \delta dS} \quad (8)$$

در (۸) و (۹) مخارجها جرم کل را بیان می‌کنند. در حالتی که δ ثابت باشد (چگالی یکنواخت)، δ از صورت و مخرج حذف می‌شود و مرکز هندسی تصویر خم با تصویر رویه به دست می‌آید.
مثال مختصات مرکز نیمکره $z^2 + y^2 + x^2 = R^2$ ، $z \geq 0$ را به دست آورید وقتی این نیمکره از ماده‌ای با چگالی یکنواخت ساخته شده باشد.

از تقارن کره واضح است که $\bar{x} = \bar{y} = 0$. برای محاسبه \bar{z} از (۹) استفاده می‌کنیم. پس از حذف کردن δ از صورت و مخرج کسر، مخرج عبارت \bar{z} برابر مساحت نیمکره یعنی $2\pi R^2$ خواهد شد. نیمکره را طبق مثال ۲، بخش ۳۵، پرمایش می‌کنیم که در آن $dS = R^2 \sin u \, du \, dv$. پس

$$\begin{aligned} \int_{\varphi} \int z \, dS &= \int \int R^3 \cos u \sin u \, du \, dv \\ &= (2\pi) R^3 \left(\frac{1}{2}\right) \sin^2 u \Big|_0^{\frac{\pi}{2}} \\ &= \pi R^3 \end{aligned}$$

بنابراین $\bar{z} = \frac{R}{4}$.

برای معرفی انتگرالهای روی خم و رویه از نوع دوم لازم است که نخست کلیاتی را در مورد 'میدانهای برداری' مطرح کنیم. اگر U زیرمجموعه‌ای از \mathbb{R}^n باشد و $F: U \rightarrow \mathbb{R}^n$ یک تابع، به F یک میدان برداری نیز می‌گویند. معمولاً وقتی اصطلاح میدان برداری به کار می‌رود تجسمی بدین صورت مورد نظر است که بردار $F(x)$ را به مبدأ نقطه x رسم می‌کنیم و کلمه "میدان" یک توزیع پیکان‌ها در نقاط مختلف U را تداعی می‌کند. با این تعبیر، معمولاً برای تأکید برداری بودن مقدار $F(x)$ ، به جای F از \vec{F} استفاده می‌کنیم.

شکل ۲

در \mathbb{R}^3 ، میدان‌های پایه \vec{i} ، \vec{j} و \vec{k} به صورت زیر تعریف می‌شوند:

$$\vec{i}(x, y, z) = (1, 0, 0) \quad , \quad \vec{j}(x, y, z) = (0, 1, 0) \quad , \quad \vec{k}(x, y, z) = (0, 0, 1) \quad (9)$$

بدین ترتیب میدان دلخواه $\vec{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$ را می‌توان به صورت $\vec{F} = F_1\vec{i} + F_2\vec{j} + F_3\vec{k}$ نمایش داد. برقراری صفاتی مانند پیوستگی و مشتق‌پذیری برای \vec{F} معادل برقراری این صفات برای هر سه مؤلفه است.

شکل ۳

به همین ترتیب در \mathbb{R}^n ، میدان‌های پایه $\vec{e}_1, \dots, \vec{e}_n$ به صورت زیر تعریف می‌شوند:

$$\vec{e}_j(x_1, \dots, x_n) = (0, \dots, 0, 1, 0, \dots, 0)$$

که در اینجا درایه ۱ در مکان j ام است. هر میدان $\vec{F} = (F_1, \dots, F_n)$ را می‌توان به صورت $\vec{F}(x) = \sum_{i=1}^n F_i(x) \vec{e}_i$ نوشت.

یک مفهوم اساسی در رابطه با میدان برداری، مفهوم "خم انتگرال" است. اگر F یک میدان برداری تعریف شده روی زیرمجموعه U از \mathbb{R}^n باشد، مقصود از یک خم انتگرال (یا خط میدان) برای \vec{F} ، خمی مشتق‌پذیر $\gamma: I \rightarrow U$ است که $\gamma'(t) = \vec{F}(\gamma(t))$ به‌ازای هر t در I . بدین ترتیب تصویر خم انتگرال طوری قرار می‌گیرد که بردار سرعت آن همواره منطبق بر بردار میدان داده شده است. پس خم‌های انتگرال در واقع جواب‌های دستگاه معادله دیفرانسیل زیر می‌شوند:

$$\begin{cases} \frac{dx_1}{dt} = F_1(x) \\ \vdots \\ \frac{dx_n}{dt} = F_n(x) \end{cases} \quad (10)$$

قضیه اساسی وجود و یگانگی جواب معادلات دیفرانسیل حکم می‌کند که اگر F_1, \dots, F_n دارای مشتق‌های پاره‌ای پیوسته باشند، از هر نقطه x در U ، خم انتگرال منحصر به فردی عبور می‌کند. در زیر به ذکر چند مثال در صفحه می‌پردازیم:

مثال ۱ میدان $\vec{F}(x, y) = x\vec{i} + y\vec{j}$ را در نظر بگیرید. این میدان در مبدأ صفر است و در هر نقطه دیگر به اندازه بردار حامل از مبدأ به آن نقطه، از آن نقطه ساطع می‌کند. بدین ترتیب بردارهای میدان همه به دور از مبدأ اشاره می‌کنند. نقطه $(0, 0)$ را در این صورت یک چشمه می‌نامند (شکل ۴، الف) خم‌های انتگرال این میدان نیم‌خط‌های راست هستند که می‌توان به صورت $\gamma(t) = (x_0 e^t, y_0 e^t)$ نمایش داد. توجه کنید که $\gamma(0) = (x_0, y_0)$ و $\gamma(t) \rightarrow (0, 0)$ وقتی $t \rightarrow -\infty$.

مثال ۲ برای $\vec{G}(x, y) = -x\vec{i} - y\vec{j}$ جهت میدان مثال ۱ وارونه می‌شود و میدان به سوی مبدأ اشاره می‌کند. $(0, 0)$ را در این حالت یک چاهک می‌نامند. خم‌های انتگرال این میدان برداری را می‌توان به صورت $\gamma(t) = (x_0 e^{-t}, y_0 e^{-t})$ نوشت.

مثال ۳ میدان $\vec{R}(x, y) = (-y)\vec{i} + x\vec{j}$ را در نظر بگیرید. در نقطه (x, y) بردار میدان عمود بر بردار حامل است و در واقع از دوران $\frac{\pi}{2}$ در جهت مثلثاتی حاصل می‌شود (شکل ۴، ب). خم انتگرال گذرا از نقطه $(x_0, y_0) = (r_0 \cos \theta_0, r_0 \sin \theta_0)$ را می‌توان به صورت:

$$\gamma(t) = (r_0 \cos(\theta_0 + t), r_0 \sin(\theta_0 + t))$$

نوشت که تصویر آن یک دایره به شعاع r_0 حول مبدأ است. در این حالت $(0, 0)$ را یک مرکز می‌نامند.

مثال ۴ میدان $\vec{S}(x, y) = x\vec{i} + (-y)\vec{j}$ را در نظر بگیرید. این میدان در شکل ۴، ج، نمایش داده شده است و خم انتگرال گذرا از (x_0, y_0) را می‌توان به $\gamma(t) = (x_0 e^t, y_0 e^{-t})$ نمایش داد.

شکل ۴

که یک شاخهٔ هذلولی $xy = x_0 y_0$ است. برای $x_0 = 0, y_0 \neq 0$ و $x_0 \neq 0, y_0 = 0$ ، نیم خط‌های مختصاتی به دست می‌آیند. نقطه $(0, 0)$ را در اینجا یک نقطه زینی می‌نامند.

نوع مهمی از میدان برداری که قبلاً نیز به آن برخوردیم میدان گرادیان است. اگر U زیرمجموعه‌ای باز از \mathbb{R}^n باشد و $f: U \rightarrow \mathbb{R}$ تابعی مشتق‌پذیر، گرادیان f ، ∇f ، را به صورت زیر تعریف کرده‌ایم:

$$\nabla f(x) = \frac{\partial f}{\partial x_1} \vec{e}_1 + \dots + \frac{\partial f}{\partial x_n} \vec{e}_n$$

تابع f را یک پتانسیل برای میدان ∇f می‌نامیم و از آنجا که میدان گرادیان بر مجموعه‌های تراز عمود است، خم‌های انتگرال که بر میدان مماس هستند نیز بر مجموعه‌های تراز عمود خواهند شد. میدان‌های \vec{F} ، \vec{G} و \vec{S} در مثال‌های ۱، ۲ و ۴ میدان‌های گرادیان هستند با پتانسیل‌های به ترتیب

$f(x, y) = \frac{1}{r}(x^2 + y^2)$ ، $g(x, y) = \frac{-1}{r}(x^2 + y^2)$ و $s(x, y) = \frac{1}{r}(x^2 - y^2)$ برای f و g ، دواير به مرکز مبدأ مجموعه‌های تراز هستند که بر نیم‌خط‌های مماس بر میدان عمودند. برای s ، هذلولی‌های ثابت $x^2 - y^2 = r$ مجموعه‌های تراز هستند که بر خانواده هذلولی‌های ثابت $xy = r$ عمود می‌باشند. نشان می‌دهیم میدان \vec{R} (مثال ۳) از نوع گرادیان نیست. فرض کنید تابعی $r(x, y)$ وجود داشته باشد که $\nabla r = \vec{R}$ ، یعنی $\frac{\partial r}{\partial x} = -y$ و $\frac{\partial r}{\partial y} = x$. از آنجایی که تابع‌های $(-y)$ و x خود مشتق‌پذیر با مشتق‌های پاره‌ای پیوسته‌اند، نتیجه می‌شود که $r(x, y)$ باید دارای مشتق‌های پاره‌ای مرتبه دوم پیوسته باشد، پس $\frac{\partial^2 r}{\partial x \partial y} = \frac{\partial^2 r}{\partial y \partial x}$ ، که برقرار نیست زیرا $\frac{\partial}{\partial x}(\frac{\partial r}{\partial y}) = 1$ و $\frac{\partial}{\partial y}(\frac{\partial r}{\partial x}) = -1$. این استدلال عمومیت دارد و منجر به شرط لازم زیر برای گرادیان بودن می‌شود:

(۳۶-۲) شرط لازم برای گرادیان بودن یک میدان فرض کنید میدان برداری $\vec{F} = F_1 \vec{e}_1 + \dots + F_n \vec{e}_n$ که دارای مشتق‌های پاره‌ای پیوسته مرتبه اول است، گرادیان یک تابع باشد، در این صورت:

$$i, j = 1, \dots, n \quad \text{برای هر زوج} \quad \frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$$

اگر f پتانسیل \vec{F} باشد، تساوی بالا همان $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ است. میدان‌های برداری گرادیان را میدان‌های پایسته نیز می‌نامند.

انتگرال روی خم و رویه (۲)

در بخش گذشته انتگرال چگالی روی خم و رویه را بررسی کردیم. اکنون به بررسی نوع دیگری انتگرال روی خم و رویه می پردازیم که به انتگرال یک میدان برداری یا انتگرال یک 'فرم دیفرانسیل' شهرت دارد. مقدار این انتگرال نیز تحت اثر بازپرمایش جهت نگهدار عوض نمی شود ولی تحت اثر بازپرمایش جهت برگردان تغییر علامت می دهد. بحث را از انتگرال روی خم آغاز می کنیم.

(۲۲-۱) انتگرال میدان برداری در طول خم فرض کنید $\gamma : [a, b] \rightarrow \mathbb{R}^n$ یک خم هموار است که گاهی آن راه به صورت $\vec{r} = \gamma(t)$ می نویسیم و $\vec{F} = (F_1, \dots, F_n)$ یک میدان برداری پیوسته که دامنه تعریف آن شامل تصویر خم γ در \mathbb{R}^n می باشد. انتگرال میدان F در امتداد خم γ که به نماد $\int_{\gamma} \vec{F} \cdot d\vec{r}$ نمایش داده می شود، به صورت زیر تعریف می کنیم:

$$\int_{\gamma} \vec{F} \cdot d\vec{r} = \int_{\gamma} (\vec{F} \cdot \frac{d\vec{r}}{ds}) ds \quad (1)$$

بدین ترتیب $\int_{\gamma} \vec{F} \cdot d\vec{r}$ به یک انتگرال چگالی $\int_{\gamma} f ds$ تبدیل می شود که در آن $f = \vec{F} \cdot \frac{d\vec{r}}{ds}$. نکته تمایز این است که تحت تعویض متغیر جهت برگردان، $\frac{d\vec{r}}{ds} = \vec{T}$ (مماس واحد) در (-1) ضرب می شود، پس f تغییر علامت می دهد و در نتیجه (۱) به جهت خم وابسته است. مقدار طرف راست (۱) را می توان عیناً به صورت انتگرال چگالی محاسبه کرد و با استفاده از قاعده زنجیره ای داریم:

$$\int_{\gamma} \vec{F} \cdot d\vec{r} = \int_a^b (\vec{F}(\gamma(t)) \cdot \frac{d\vec{r}}{dt}) dt \quad (2)$$

با توجه به اینکه $\frac{d\vec{r}}{ds} = \vec{T}$ مماس واحد در جهت پرمایش خم γ است، تعبیر هندسی این انتگرال نیز مشهود است. $\vec{F} \cdot \vec{T} = |\vec{F}| \cos(\vec{F}, \vec{T})$ مؤلفه \vec{F} در جهت \vec{T} است، پس در واقع $\int_{\gamma} \vec{F} \cdot d\vec{r}$ نمایانگر میزان انباشته شدن میدان \vec{F} در راستای حرکت روی γ است. در فیزیک، اگر \vec{F} یک میدان نیرو باشد

$\int_{\gamma} \vec{F} \cdot d\vec{r}$ را "کار انجام شده" در امتداد γ می نامند. قبل از ارائه مثال و کاربردهای این نوع انتگرال، نمادگذاری دیگری را که نیز معمول است معرفی می کنیم. بردار $\vec{r} = \gamma(t)$ در واقع مکان متحرک را در زمان t نشان می دهد. می نویسیم $\vec{r} = (x_1, \dots, x_n)$ و در نتیجه $\gamma'(t) = (\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt})$. بنابراین (۲) به صورت زیر در می آید:

$$\int_a^b (F_1 \frac{dx_1}{dt} + \dots + F_n \frac{dx_n}{dt}) dt \quad (3)$$

چون تعویض پارامتر جهت نگهدار اثری بر مقدار انتگرال فوق ندارد، گاهی با حذف dt ها انتگرال $\int_{\gamma} \vec{F} \cdot d\vec{r}$ را به صورت زیر نیز می نویسند:

$$\int_{\gamma} \vec{F} \cdot d\vec{r} = \int_{\gamma} F_1 dx_1 + \dots + F_n dx_n \quad (4)$$

مثال ۱ برای $\vec{R}(x, y) = (-y)\vec{i} + x\vec{j}$ و γ دایره شعاع A به مرکز $(0, 0)$ که یک بار در جهت مثلثاتی پیموده شده است، $\int_{\gamma} \vec{R} \cdot d\vec{r}$ را محاسبه کنید.

حل داریم $\vec{r} = \gamma(t) = (A \cos t, A \sin t)$ ، $0 \leq t \leq 2\pi$ پس $x = A \cos t$ و $y = A \sin t$. بدین ترتیب:

$$\begin{aligned} \int_{\gamma} \vec{R} \cdot d\vec{r} &= \int_0^{2\pi} (\vec{R} \cdot \frac{d\vec{r}}{dt}) dt \\ &= \int_0^{2\pi} [(-A \sin t)(-A \sin t) + (A \cos t)(A \cos t)] dt \\ &= 2\pi A^2 \end{aligned}$$

لازم است که مفهوم انتگرال روی خم را به خم های "قطعه قطعه هموار" توسعه دهیم. خم پیوسته $\gamma: [a, b] \rightarrow \mathbb{R}^n$ را قطعه قطعه هموار می نامیم در صورتی که افراز $a = t_0 < t_1 < \dots < t_p = b$ وجود داشته باشد که γ به دامنه $[t_{i-1}, t_i]$ برای هر $i = 1, \dots, n$ یک خم هموار تعریف کند. بدین ترتیب γ ممکن است در $\gamma(t_i)$ "گوشه" داشته باشد زیرا که مماس چپ $\vec{T}(t_i)^-$ و مماس راست $\vec{T}(t_i)^+$ ممکن است بر هم منطبق نباشند. برای خم قطعه قطعه هموار فوق، $\int_{\gamma} f ds$ را برابر مجموع $\sum_{i=1}^p \int_{\gamma_i} f ds$ تعریف می کنیم در آن γ_i تحدید γ به $[t_{i-1}, t_i]$ است.

مثال ۲ اگر \vec{R} مانند مثال قبل باشد، $\int_{\gamma} \vec{R} \cdot d\vec{r}$ را برای γ : مربع با رئوس $(\pm A, \pm A)$ یک بار پیموده شده در جهت مثلثاتی، محاسبه کنید.

حل

شکل ۵

خم های $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ چهار ضلع مربع را طبق شکل توصیف می کنند. باید $\sum_{i=1}^4 \int_{\gamma_i} \vec{R} \cdot d\vec{r}$ را محاسبه کنیم. خم ها را می توان به شکل زیر توصیف کرد

$$\begin{cases} \gamma_1(t) = (A, t), & -A \leq t \leq A \\ \gamma_2(t) = (-t, A), & -A \leq t \leq A \\ \gamma_3(t) = (-A, -t), & -A \leq t \leq A \\ \gamma_4(t) = (t, -A), & -A \leq t \leq A \end{cases}$$

پس

$$\int_{\gamma_1} \vec{R} \cdot d\vec{r} = \int_{-A}^A [(-t, A) \cdot (0, 1)] dt = \int_{-A}^A A dt = 2A^2$$

$$\int_{\gamma_2} \vec{R} \cdot d\vec{r} = \int_{-A}^A [(-A, -t) \cdot (-1, 0)] dt = \int_{-A}^A A dt = 2A^2$$

به همین ترتیب انتگرال های روی γ_3 و γ_4 نیز هر یک برابر $2A^2$ خواهند شد و $\int_{\gamma} \vec{R} \cdot d\vec{r}$ برابر $8A^2$ می شود.

توجه کنید که در این مثال و مثال قبل $\int_{\gamma} -y dx + x dy$ دو برابر مساحت محصور شده توسط γ درآمد. این یک واقعیت کلی است که بعداً به دلیل آن پی خواهیم برد.

در ادامه بحث، انتگرال های روی خم $\int_{\gamma} \vec{F} \cdot d\vec{r}$ را برای میدان های پایسته \vec{F} بررسی می کنیم.

فرض کنید $\gamma: [a, b] \rightarrow \mathbb{R}^n$ یک خم هموار باشد با نقطه آغازی $\gamma(a) = A$ و نقطه پایانی $\gamma(b) = B$.

و $\vec{F} = \nabla f$ برای تابع مشتق پذیر f . در این صورت:

$$\begin{aligned}\int_{\gamma} \vec{F} \cdot d\vec{r} &= \int_a^b \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \right) dt \\ &= \int_a^b (f \circ \gamma)'(t) dt \\ &= f(B) - f(A)\end{aligned}$$

یعنی مقدار $\int_{\gamma} \vec{F} \cdot d\vec{r}$ ، مستقل از مسیر هموار γ ، برابر اختلاف پتانسیل در دو انتهای مسیر است. این مطلب را می توان به خم های پیوسته قطعه قطعه هموار $\gamma: [a, b] \rightarrow \mathbb{R}^n$ نیز تعمیم داد زیرا که اگر $a = t_0 < t_1 < \dots < t_p = b$ افزایش دهیم، آنگاه بنا بر پیوستگی خم، $\gamma_i(t_i) = \gamma_{i+1}(t_i)$ و در نتیجه:

$$\begin{aligned}\sum_{i=1}^p \int_{\gamma_i} \vec{F} \cdot d\vec{r} &= \sum_{i=1}^p (f(\gamma_i(t_i)) - f(\gamma_i(t_{i-1}))) \\ &= f(\gamma_p(b)) - f(\gamma_1(a))\end{aligned}$$

نکته مهم این که این ویژگی مختص میدان های پایسته است.

(۲-۳۷) قضیه فرض کنید U یک مجموعه باز و همبند مسیری در \mathbb{R}^n باشد و \vec{F} یک میدان پیوسته تعریف شده در U . در این صورت \vec{F} پایسته است اگر و تنها اگر $\int_{\gamma} \vec{F} \cdot d\vec{r}$ مستقل از مسیر باشد، یعنی برای هر دو خم قطعه قطعه هموار α و β در U با نقاط آغازی و پایانی مشترک، داشته باشیم

$$\int_{\alpha} \vec{F} \cdot d\vec{r} = \int_{\beta} \vec{F} \cdot d\vec{r}$$

اثبات لزوم این شرط را در بالا دیدیم. حال فرض کنید شرط استقلال از مسیر برقرار باشد، می خواهیم وجود پتانسیل برای میدان \vec{F} را ثابت کنیم. نقطه ای P را در U به عنوان نقطه مرجع تثبیت می کنیم. تابعی f روی U تعریف می کنیم که $\nabla f = \vec{F}$.

شکل ۱

برای نقطه دلخواه Q ، $f(Q)$ را به صورت زیر تعریف می کنیم. از آنجا که U همبند مسیری و باز است، مسیری قطعه قطعه هموار از P به Q وجود دارد، و به سبب شرط استقلال از مسیر، $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای هر

مسیر قطعه قطعه هموار از P به Q یک مقدار را دارد، پس با انتخاب یک چنین مسیر γ ، می توان تعریف کرد:

$$f(Q) = \int_{\gamma} \vec{F} \cdot d\vec{r}$$

باید نشان دهیم f پتانسیل مورد نظر است، یعنی $\frac{\partial f}{\partial x_i} = F_i$ ، که F_i مؤلفه i -ام \vec{F} است. به این منظور از Q پاره خط راستی به موازات محور x_i رسم می کنیم که در U بماند و آن را به صورت زیر پارامتری می کنیم:

$$\lambda(t) = Q + te_i, \quad 0 \leq t \leq h$$

نقطه $Q_h = Q + he_i$ ، نقطه پایانی مسیری است که از P شروع می شود، نخست تصویر γ را می بینیم و سپس تصویر λ را. این مسیر را، که قطعه قطعه هموار است، به $\gamma * \lambda$ نمایش می دهیم. طبق تعریف:

$$f(Q_h) = \int_{\gamma * \lambda} \vec{F} \cdot d\vec{r} = \int_{\gamma} \vec{F} \cdot d\vec{r} + \int_{\lambda} \vec{F} \cdot d\vec{r} = f(Q) + \int_0^h F_i(\lambda(t)) dt$$

(توجه کنید که غیر از مؤلفه i -ام، سایر مؤلفه های λ ثابت هستند) بنابراین:

$$\frac{f(Q_h) - f(Q)}{h} = \frac{1}{h} \int_0^h F_i(\lambda(t)) dt$$

بنابر قضیه میانگین انتگرال، انتگرال سمت راست برابر است با $h \cdot F_i(Q^*)$ که Q^* نقطه ای روی پاره خط λ است. بدین ترتیب $\frac{f(Q_h) - f(Q)}{h} = F_i(Q^*)$. وقتی $h \rightarrow 0$ به Q_h به Q میل می کند و در نتیجه Q^* که روی پاره خط واصل قرار دارد نیز به Q میل می کند. طبق پیوستگی F_i ، سمت راست به $F_i(Q)$ میل می کند و ثابت می شود که $\frac{\partial f}{\partial x_i}(Q) = F_i(Q)$. ■

مثال ۳ میدان \vec{F} روی $\mathbb{R}^n - \{0\}$ به صورت زیر تعریف شده است:

$$\vec{F}(x_1, \dots, x_n) = \sum_{i=1}^n \frac{x_i}{x_1^2 + \dots + x_n^2} \vec{e}_i$$

اگر P و Q دو نقطه به فواصل $|P|$ و $|Q|$ از 0 باشند و γ یک خم قطعه قطعه هموار در $\mathbb{R}^n - \{0\}$ با نقطه آغازی P و نقطه پایانی Q ، ادعا می کنیم $\int_{\gamma} \vec{F} \cdot d\vec{r} = \ln \frac{|Q|}{|P|}$. بالاخص اگر P و Q هم فاصله از

باشد، داریم $\int_{\gamma} \vec{F} \cdot d\vec{r} = 0$. در اینجا می‌توان به سادگی حدس زد و تحقیق نمود که میدان \vec{F} دارای پتانسیل $f(x_1, \dots, x_n) = \frac{1}{2} \ln(x_1^2 + \dots + x_n^2)$ است و حکم فوراً نتیجه می‌شود.

مثال ۴ میدان \vec{F} در $\mathbb{R}^2 - \{0\}$ را به صورت زیر در نظر می‌گیریم:

$$\vec{F}(x, y) = \frac{-y}{x^2 + y^2} \vec{i} + \frac{x}{x^2 + y^2} \vec{j}$$

آیا این میدان پایسته است؟ توجه کنید که استقلال از مسیر برای انتگرال میدان‌های پایسته نتیجه می‌دهد که انتگرال حول یک مسیر بسته قطعه قطعه هموار باید صفر شود. اگر معلوم شود که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای یک مسیر قطعه قطعه هموار بسته صفر نیست، \vec{F} نمی‌تواند پایسته باشد. γ را دایره‌ای به شعاع R به مرکز 0 یک بار پیموده شده در جهت مثلثاتی می‌گیریم:

$$\gamma(t) = (R \cos t, R \sin t) \quad , \quad 0 \leq t \leq 2\pi$$

$$\begin{aligned} \int_{\gamma} \vec{F} \cdot d\vec{r} &= \int_0^{2\pi} \left[\frac{-R \sin t}{R^2} (-R \sin t) + \frac{R \cos t}{R^2} (R \cos t) \right] dt \\ &= \int_0^{2\pi} 1 dt = 2\pi \neq 0 \end{aligned}$$

بدین ترتیب \vec{F} پایسته نیست. ضمناً توجه کنید که نتیجه انتگرال گیری به شعاع R بستگی ندارد، این مطلب بعداً بررسی خواهد شد.

مثال را بالا نباید به همین صورت (فقدان پتانسیل) رها کرد. توجه کنید که این میدان بر بردار حامل از 0 به نقطه (x, y) عمود است، بنابراین بر نیم خط (ثابت θ) عمود می‌باشد. به نظر می‌آید که باید پتانسیلی برای این میدان وجود داشته باشد که مجموعه‌های تراز آن نیم خط‌های ثابت θ باشند.

در واقع با در نظر گرفتن $f(x, y) = \theta = \tan^{-1}\left(\frac{y}{x}\right)$ داریم:

$$\frac{\partial f}{\partial x} = \frac{-\frac{y}{x^2}}{1 + \frac{y^2}{x^2}} = \frac{-y}{x^2 + y^2}$$

$$\frac{\partial f}{\partial y} = \frac{\frac{1}{x}}{1 + \frac{y^2}{x^2}} = \frac{x}{x^2 + y^2}$$

یعنی، علی‌رغم آنچه ثابت شد، ظاهراً پتانسیلی برای این میدان موجود است. اشکالی که فوراً به ذهن می‌رسد این است که $\frac{y}{x}$ روی محور y تعریف نشده و $\tan^{-1}(\frac{y}{x})$ نمی‌تواند معنی داشته باشد. ولیکن با قرار دادن $\theta = \frac{\pi}{4}$ روی نیم خط بالای محور y و $\theta = \frac{3\pi}{4}$ روی نیم خط پایین به نظر می‌آید که مشکل حل شود چون θ چیزی جز مختصه قطبی زاویه نیست. مشکل واقعی در اینجا "چند مقداری بودن" θ است. توجه کنید که اگر نیم خط مثبت محور x را با $\theta = 0$ مشخص کنیم و یک دور کامل

شکل ۲

دور مبداء گردش کنیم در حالی که تابع θ به طور پیوسته تغییر کند، آنگاه در بازگشت به نیمه راست محور x ، مقدار θ به (2π) می‌رسد، نه صفر! جالب اینجاست که این مقدار 2π در واقع مقداری است که برای انتگرال مثال قبل به دست آورده‌ایم و این امر تصادفی نیست. در واقع به اعتباری، نوعی "پتانسیل چند مقداری" برای \vec{F} وجود دارد که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ همچنان برابر اختلاف پتانسیل بین دو سر مسیر خواهد بود. کار کردن با این پتانسیل‌ها دقت می‌طلبد و بررسی دقیق آنها از کار این درس خارج است. موضع رسمی ما در مورد این میدان همچنان فقدان پتانسیل خواهد بود. پس از مطالعه قضیه گرین در جلسه بعد، به نحو مؤثرتری از این میدان استفاده خواهیم کرد. بالاخص خواهیم دید که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ ، همچنان که با توجه به θ انتظار داریم، برابر $2\pi k$ خواهد بود که k تعداد دفعات گردش مسیر γ حول 0 در جهت مثلثاتی است.

بالاخره در مورد مثال بالا ذکر این نکته نیز لازم است که چنانچه یک نیم خط کامل ساطع از 0 را از \mathbb{R}^2 حذف کنیم، تابع θ بدون ابهام به عنوان یک تابع مشتق‌پذیر روی باقیمانده صفحه قابل تعریف است زیرا در این ناحیه امکان گردش دور 0 و بازگشت به نقطه آغازی وجود ندارد. مثلاً با حذف $\theta = \theta_0$ و تعریف کردن $\theta_0 + 2\pi < \theta < \theta_0$ برای باقیمانده صفحه، θ تابعی مشتق‌پذیر در دامنه خود خواهد شد. تحدید \vec{F} به چنین ناحیه‌ای پایسته است و $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای هر مسیر بسته کاملاً واقع شده در این ناحیه صفر است (شکل ۳).

شکل ۳

انتگرال روی خم و رویه (۳)

در ادامه بحث انتگرال روی اجسام خمیده، اکنون مفهوم انتگرال یک میدان برداری 'یا فرم دیفرانسیل' را روی رویه بررسی می‌کنیم. فرض کنید $\varphi : W \rightarrow \mathbb{R}^3$ ، $\vec{r} = \varphi(u, v)$ ، یک رویه هموار باشد و W یک ناحیه مجاز انتگرالگیری در \mathbb{R}^2 . طبق معمول نقاط W را به $W = (u, v)$ ، عنصر سطح $dS = |\varphi_u \times \varphi_v| du dv$ ، را به dS ، و قائم واحد قراردادی $\frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}$ را به \vec{n} نمایش می‌دهیم. گاهی $d\vec{S} = (\vec{r}_u \times \vec{r}_v) du dv = \vec{n} dS$ را به $d\vec{S}$ نیز نمایش می‌دهند. فرض کنید $\vec{F} = (F_1, F_2, F_3)$ یک میدان برداری پیوسته باشد که دامنه آن شامل تصویر φ ، یعنی $\varphi(W)$ ، است. در این صورت تعریف می‌کنیم:

$$\int \int_{\varphi} \vec{F} \cdot d\vec{S} = \int \int_{\varphi} (\vec{F} \cdot \vec{n}) dS \quad (1)$$

توجه کنید به ازای هر $(u, v) \in W$ ، $\vec{n}((u, v)) = \vec{n}(\varphi(u, v))$ تعریف شده است، پس طرف راست بالا یک انتگرال چگالی به مفهوم عادی می‌باشد. از طرفی دیگر چون \vec{n} معرف جهت رویه است، $\vec{F} \cdot \vec{n}$ تحت تعویض متغیر جهت برگردان تغییر علامت می‌دهد، پس (۱) نیز تحت تعویض متغیر جهت برگردان تغییر علامت می‌دهد.

تعبیر فیزیکی زیر در مورد انتگرال (۱) منشاء کاربردهای بسیاری است. \vec{F} را یک میدان نیرو فرض کنید و $\varphi(W)$ را یک لایه بسیار نزدیک در فضای سه بعدی تصور کنید که در معرض این نیرو قرار دارد. مقدار $\vec{F} \cdot \vec{n}$ برابر مؤلفه \vec{F} در جهت عمود قراردادی بر لایه است. انتگرال $\vec{F} \cdot \vec{n}$ نسبت به dS روی این لایه نمایشگر کل شار گذرا از لایه محسوب می‌شود.

مثال ۱ میدان واحد قائم $\vec{F} = \vec{k}$ در \mathbb{R}^3 را در نظر بگیرید. شار گذرا از نیمکره $x^2 + y^2 + z^2 = R^2$ ، $z \geq 0$ ، را محاسبه کنید اگر قائم واحد \vec{n} با مؤلفه سوم مثبت برای نیمکره منظور شود.

در نقطه (x, y, z) از نیمکره داریم $\vec{n} = (\frac{x}{R}, \frac{y}{R}, \frac{z}{R})$ ، باید نماد دیگر $\vec{n} = \frac{x}{R}\vec{i} + \frac{y}{R}\vec{j} + \frac{z}{R}\vec{k}$ پس $\vec{F} \cdot \vec{n} = \frac{z}{R}$ و یا به انتگرال $\iint \frac{z}{R} dS$ را روی نیمکره محاسبه کنیم. در مختصات کروی داریم:

$$z = R \cos \varphi$$

$$dS = R^2 \sin \varphi d\varphi d\theta$$

پس

$$\begin{aligned} \int_0^{2\pi} \int_0^{\frac{\pi}{2}} (\cos \varphi) (R^2 \sin \varphi) d\varphi d\theta &= (2\pi R^2) \left(\frac{1}{\varphi}\right) \sin^2 \varphi \Big|_0^{\frac{\pi}{2}} \\ &= \pi R^2 \end{aligned}$$

عبارت داخل انتگرال در طرف راست (۱) را می توان به صورتهای دیگری نیز بازنویسی کرد. از آنجا

$$\text{که داریم } \vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}$$

$$\vec{n} dS = (\vec{r}_u \times \vec{r}_v) dudv \quad (2)$$

پس

$$\iint \vec{F} \cdot d\vec{S} = \iint_W \vec{F} \cdot (\vec{r}_u \times \vec{r}_v) dudv \quad (3)$$

حال اگر بنویسیم $\vec{r} = (x, y, z)$ ، داریم

$$\vec{r}_u \times \vec{r}_v = \begin{pmatrix} \begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix}, \begin{vmatrix} z_u & z_v \\ x_u & x_v \end{vmatrix}, \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \end{pmatrix}$$

که طبق معمول مقصود از نماد y_u مشتق پاره‌ای y نسبت به u ، یعنی $\frac{\partial y}{\partial u}$ ، است و به همین ترتیب برای سایر درایه‌های دترمینان‌های بالا. پس برای $\vec{F} = (F_1, F_2, F_3)$ نتیجه می شود که:

$$\iint \vec{F} \cdot d\vec{S} = \iint_W \left[F_1 \begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix} + F_2 \begin{vmatrix} z_u & z_v \\ x_u & x_v \end{vmatrix} + F_3 \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \right] dudv \quad (4)$$

طرف راست (۴) روش محاسبه صریح این نوع انتگرال را برحسب پرمایش φ عرضه می کند. از آنجا که تحت تعویض پارامتر جهت نگهدار مقدار این انتگرال تغییر نمی کند، مطلوب است که عبارت سمت راست (۴) به صورتی مستقل از پارامترهای خاص (u, v) نوشته شود. توجه کنید که دترمینان

$$\begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix} \text{ دترمینان ماتریس مشتقهای } y \text{ و } z \text{ نسبت به } u \text{ و } v \text{ است. به جای } dudv \text{ گاهی}$$

می نویسیم $dydz$. و به همین ترتیب $dzdx$ و $dx dy$ (به ترتیب توجه کنید) را جایگزین عبارتهای متناظر کرده و (۴) را به صورت زیر نیز می نویسیم:

$$\int_{\varphi(W)} \int \vec{F} \cdot d\vec{S} = \int_{\varphi} \int F_1 dydz + F_2 dzdx + F_3 dx dy \quad (5)$$

مقصود از عبارت سمت راست این است که نسبت به هر پرمایش مجاز φ با جایگزین مناسب برای $dydz$, $dzdx$ و $dx dy$ برحسب پارامترهای ذیربط می توان انتگرال را محاسبه کرد.

مثال ۲ انتگرال مثال (۱) را به روش دیگری محاسبه می کنیم. داریم $F_1 = F_2 = 0$ و $F_3 = 1$ ، پس برای نیمکره شمالی H ، به شعاع R ، داریم:

$$\int_H \int \vec{k} \cdot d\vec{S} = \int_H \int (1) dx dy$$

حال می توان نیمکره شمالی را به صورت نمودار تابعی برحسب x و y یعنی $z = \sqrt{R^2 - x^2 - y^2}$ پرمایش کرد که ضمناً دیده ایم مؤلفه سوم $\vec{r}_x \times \vec{r}_y$ همواره مثبت است پس جهت روبه منظور شده در مثال (۱) را دارد. پس:

$$\int_H \int dx dy = \int_{x^2 + y^2 \leq R^2} \int dx dy = \pi R^2$$

در باقیمانده این بخش دو مفهوم 'دیورژانس' و 'چرخه' را معرفی می کنیم که همراه با گرادیان نقش اساسی در قضایای مهم انتگرال روی خمها و رویه ها دارند. فرض کنید U زیرمجموعه ای باز از \mathbb{R}^n است و $\vec{F} = (F_1, \dots, F_n)$ یک میدان برداری C^1 تعریف شده روی U . دیورژانس \vec{F} ، که به $div \vec{F}$ ، نمایش داده می شود، تابع زیر است:

$$div \vec{F} = \frac{\partial F_1}{\partial x_1} \dots + \frac{\partial F_n}{\partial x_n} \quad (6)$$

اگر \vec{F} تابعی از U به \mathbb{R}^n تصور کنیم، $div \vec{F}$ در واقع مجموع درایه های قطر اصلی ماتریس ژاکوبی \vec{F} است. مجموع عناصر قطری یک ماتریس مربعی، مانند دترمینان، از خواص مهمی برخوردار است. در بخشهای آینده خواهیم دید که دیورژانس \vec{F} نوعی شاخص وضعیت میدان \vec{F} از نظر انبساط یا انقباض

است. مثبت بودن $\text{div} \vec{F}$ در یک نقطه نوعی انبساط حول آن نقطه، یا دور شدن پیکانهای میدان حول آن نقطه را نشان می‌دهد، و بالعکس منفی بودن $\text{div} \vec{F}$ دالّ بر تقارب پیکانها یا انقباض است. بدین ترتیب دیورژانس به هر میدان برداری C^1 روی U یک تابع پیوسته $\text{div} \vec{F} : U \rightarrow \mathbb{R}$ نسبت می‌دهد، در حالی که گرادیان، به هر تابع C^1 ، $f : U \rightarrow \mathbb{R}$ یک میدان برداری پیوسته $\text{grad} f = \nabla f$ روی U منسوب می‌کند.

در حالت $n = 3$ ، یعنی وقتی که U زیرمجموعه بازی از \mathbb{R}^3 باشد، یک عملگر دیگر روی میدانهای برداری مجموعه گرادیان و دیورژانس را کامل می‌کند. این عملگر که چرخه نام دارد و به نمادهای curl یا rot نمایش داده می‌شود، به هر میدان برداری مشتقپذیر $\vec{F} = (F_1, F_2, F_3)$ روی U میدان دیگری به شکل زیر نسبت می‌دهد:

$$\text{curl} \vec{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \quad (7)$$

در بخشهای آینده خواهیم دید که $\text{curl} \vec{F}$ به اعتباری تقریب \vec{F} با یک میدان چرخشی است، $|\text{curl} \vec{F}|$ مقدار چرخش و راستای $\text{curl} \vec{F}$ (چنانچه $\text{curl} \vec{F}$ صفر نباشد) محور این چرخش است. دو حکم کلیدی زیر ارتباط مهمی را میان سه عملگر grad ، curl و div بیان می‌کنند.

(۱-۳۸) الف) اگر f تابعی C^2 باشد داریم:

$$\text{curl}(\text{grad} f) = \underline{\underline{0}} \quad (8)$$

ب) اگر \vec{F} یک میدان C^2 باشد، داریم

$$\text{div}(\text{curl} \vec{F}) = 0 \quad (9)$$

برهان اثبات (۸) و (۹) سراسر است ولی مشاهده محاسبات لازم حایز اهمیت است:

$$\begin{aligned} \text{curl}(\text{grad} f) &= \text{curl} \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \\ &= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}, \frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial y \partial x} - \frac{\partial^2 f}{\partial x \partial y} \right) \end{aligned}$$

از آنجا که f تابعی C^2 ، یعنی دارای مشتقهای پاره‌ای مرتبه دوم پیوسته، فرض شده است، هر سه مؤلفه بالا صفر هستند. به همین ترتیب داریم:

$$\begin{aligned} \operatorname{div}(\operatorname{curl}\vec{F}) &= \operatorname{div}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) \\ &= \frac{\partial^2 F_z}{\partial x \partial y} - \frac{\partial^2 F_y}{\partial x \partial z} + \frac{\partial^2 F_x}{\partial y \partial z} - \frac{\partial^2 F_z}{\partial y \partial x} + \frac{\partial^2 F_y}{\partial z \partial x} - \frac{\partial^2 F_x}{\partial z \partial y} \end{aligned}$$

و در اینجا نیز چون مؤلفه‌های \vec{F} دارای مشتقهای پاره‌ای پیوسته مرتبه دوم فرض شده‌اند، شش جمله بالا دوبه‌دو حذف می‌شوند و حکم نتیجه می‌شود.

□

توجه کنید که هر دو حکم بالا دقیقاً بیانگر مجاز بودن تعویض ترتیب مشتقگیری برای تابعهای C^2 هستند. می‌توان تعبیری شهودی از (۸) و (۹) ارائه داد. در مورد (۸)، میدان گرادیان میدان است که تمایل گردشی ندارد زیرا که خطوط میدان همواره در جهت بیشترین افزایش تابع f حرکت می‌کنند. بنابراین $\operatorname{curl}(\operatorname{grad}f)$ ، که میزان چرخش میدان $\operatorname{grad}f$ است، صفر می‌شود. به همین ترتیب، $\operatorname{curl}\vec{F}$ حالت چرخشی دارد، یعنی نه منبسط کننده و نه منقبض کننده است، بنابراین دیورژانس آن صفر می‌شود.

نماد زیر برای به یاد ماندن سه عملگر ذکر شده رویه مفید واقع می‌شود و به هر حال خلاصه‌نویسی

سودمندی است. اگر ∇ را به‌طور نمادین به صورت

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \quad (10)$$

بنویسیم، گرادیان f از نوشتن f در جوار مؤلفه‌های ∇ به دست می‌آید، یعنی:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \quad (11)$$

اگر $\vec{F} = (F_x, F_y, F_z)$ یک میدان باشد، ضرب داخلی نمادین ∇ و \vec{F} برابر دیورژانس می‌شود:

$$\nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \quad (12)$$

و بالاخره ضرب خارجی نمادین ∇ و \vec{F} ، میدان $\operatorname{curl}\vec{F}$ را به دست می‌دهد:

$$\nabla \times \vec{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) \quad (13)$$

عملگرهای گرادیان، چرخه و دیورژانس هریک رابطه نزدیکی با مشتقگیری دارند. علاوه بر (۸) و (۹)، اتحادهای زیر نیز گاهی مورد استفاده قرار خواهد گرفت. بعضی از این اتحادها شباهت زیادی به قواعد مشتقگیری دارند. در زیر اعداد حقیقی ثابت با حروف a و b ، تابعهای با مقدار حقیقی با حروفی مانند f و g ، و میدانهای برداری با نمادهایی چون \vec{F} و \vec{G} نمایش داده شده‌اند. تابعها و میدانها مشتقپذیر فرض شده‌اند.

$$\nabla(af + bg) = a(\nabla f) + b(\nabla g) \quad (14)$$

$$\nabla(fg) = g(\nabla f) + f(\nabla g) \quad (15)$$

$$(g \neq 0) \quad \nabla\left(\frac{f}{g}\right) = \frac{g(\nabla f) - f(\nabla g)}{g^2} \quad (16)$$

$$\nabla \cdot (a\vec{F} + b\vec{G}) = a(\nabla \cdot \vec{F}) + b(\nabla \cdot \vec{G}) \quad (17)$$

$$\nabla \cdot (f\vec{G}) = f(\nabla \cdot \vec{G}) + (\nabla f) \cdot \vec{G} \quad (18)$$

$$\nabla \cdot (\vec{F} \times \vec{G}) = \vec{G} \cdot (\nabla \times \vec{F}) - \vec{F} \cdot (\nabla \times \vec{G}) \quad (19)$$

$$\nabla \times (a\vec{F} + b\vec{G}) = a(\nabla \times \vec{F}) + b(\nabla \times \vec{G}) \quad (20)$$

$$\nabla \times (f\vec{G}) = f(\nabla \times \vec{G}) + (\nabla f) \times \vec{G} \quad (21)$$

اثبات این اتحادها سراسر است و به خواننده واگذار می‌شود.

آنالیز برداری (۱)

همان طور که قبلاً ذکر شد هدف ما اکنون بررسی انتگرال روی مجموعه هموار خمیده یا دارای مرز خمیده است. دستاورد مهم این بررسی سه قضیه به نام‌های قضیه گرین، قضیه استوکس و قضیه دیورژانس خواهد بود که هر یک به نوعی انتگرال روی یک مجموعه را به انتگرال یک بعد پایین‌تر روی مرز آن مرتبط می‌سازد. از دیدگاهی، هر سه قضیه، تعمیم‌های قضیه حساب دیفرانسیل و انتگرال هستند، و در واقع عناصر اصلی اثبات هر یک از آنها همان قضیه اساسی است. در این سیر نخست به بحث پیرامون قضیه گرین می‌پردازیم.

فرض کنید D یک ناحیه در صفحه است که محصور به یک یا چند (تعدادی متناهی) تصویر سه خم قطعه قطعه هموار می‌باشد. در شکل ۱ (الف) مرز ناحیه D تصویر، یک خم، در شکل ۱ (ب)، مرز ناحیه D متشکل از تصویر خم قطعه قطعه هموار است. ناحیه D قسمت هاشورزده شکل است. یادآوری می‌کنیم که خم پیوسته $\gamma : [a, b] \rightarrow \mathbb{R}^n$ را قطعه قطعه هموار می‌نامیم در صورتی که افراز $a = t_0 < t_1 < \dots < t_p = b$ وجود داشته باشد به طوری که تحدید γ به هر $[t_{i-1}, t_i]$ یک خم هموار باشد. در نقاط انتهایی $[t_{i-1}, t_i]$ مشتق یک طرفه منظور می‌کنیم. اجتماع تصاویر خمهای مرزی را به ∂D نمایش می‌دهیم. هر نقطه $\gamma(t_i)$ یک "گوشه" ∂D است.

شکل ۱

در نقاط غیر گوشه‌ای مرز D دو انتخاب برای مماس واحد \vec{T} وجود دارد. جهت قراردادی ∂D ، یعنی انتخاب مماس واحد، را به صورت زیر منظور می‌کنیم. در هر نقطه غیر گوشه‌ای \vec{T} به گونه‌ای انتخاب می‌شود که اگر به اندازه $\frac{\pi}{4}$ در جهت مثلثاتی گردش کند به سوی درون D قرار گیرد. بدین ترتیب جهت ∂D در شکل ۱ (الف) جهت مثلثاتی است و در شکل ۱ (ب)،

خم بیرونی در جهت مثلثاتی و خمهای داخلی در جهت عقربه ساعت جهت داده می شوند. در زیر همواره ∂D با جهت قراردادی موردنظر است.

(۳۹-۱) قضیه گرین ناحیه‌ای در صفحه است که مرز آن ∂D از تعدادی متناهی تصویر خمهای قطعه قطعه هموار تشکیل شده است و $\vec{F} = (F_1, F_2)$ یک میدان برداری C^1 که دامنه تعریف آن شامل D و ∂D است.

در این صورت داریم

$$\int_{\partial D} \vec{F} \cdot d\vec{r} = \int_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA_{x,y} \quad (1)$$

قبل از پرداختن به اثبات این قضیه، تعدادی از کاربردهای آن را بررسی می کنیم. از نظر محاسباتی قضیه گرین می تواند در بسیاری موارد کارساز باشد. در عمل گاهی طرف راست فرمول (۱) و گاهی طرف چپ از نظر محاسباتی ساده ترند و در هر صورت می توان از قضیه گرین بهره جست.

کاربرد ۱ با انتخاب مناسب مؤلفه های \vec{F} ، می توان فرمول های متعددی برای مساحت ناحیه D برحسب انتگرال روی خم ∂D پیدا کرد. مثلاً با گرفتن $\vec{F} = \frac{1}{2}(-y, x)$ ، داریم $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 1$ پس

$$D \text{ مساحت} = \frac{1}{2} \int_{\partial D} (-ydx + xdy) \quad (2)$$

کاربرد ۲ (مسأله پتانسیل در صفحه) قبلاً دیدیم که اگر $\vec{F} = F_1\vec{i} + F_2\vec{j}$ یک میدان گرادیان در ناحیه ای U از صفحه xy باشد، لزوماً داریم:

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x} \quad (3)$$

اکنون ملاحظه خواهیم کرد که اگر U واجد شرط هندسی مناسبی باشد، شرط (۳) برای این که \vec{F} به شکل گرادیان باشد (معادلاً \vec{F} پایسته باشد) کفایت می کند:

(۳۹-۲) قضیه فرض کنید مجموعه باز و همبند مسیری U در صفحه واجد شرط هندسی زیر باشد:

برای هر خم بسته ساده قطعه قطعه هموار در U ، ناحیه درون خم به تمامی در U قرار دارد. در این

صورت اگر میدان دارای مشتق‌های پاره‌ای پیوسته $\vec{F} = F_1\vec{i} + F_2\vec{j}$ که در U تعریف شده است حایز شرط (۳) باشد، \vec{F} پایسته است.

شرط ذکر شده بدین صورت بیان می‌شود که "ناحیه U فاقد حفره است". در شکل (۲ الف) یک ناحیه فاقد حفره و در شکل (۲ ب) یک ناحیه حفره دار دیده می‌شود. توجه کنید که برای ناحیه

شکل ۲

حفره‌دار درون یک خم بسته ساده حول حفره به تمامی در U قرار ندارد.

اثبات قضیه به سادگی از قضیه گرین به دست می‌آید. قبلاً دیده‌ایم که شرطی لازم و کافی برای پایسته بودن \vec{F} در U این است که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای هر خم بسته ساده در U صفر شود. حال اگر شرط هندسی ذکر شده برقرار باشد، برای هر خم بسته ساده γ در U ، ناحیه درون آن D به تمامی در U ، قلمرو \vec{F} ، قرار دارد، و می‌توان از قضیه گرین و اینکه $\frac{\partial F_1}{\partial x} - \frac{\partial F_2}{\partial y} = 0$ نتیجه گرفت که $\int_{\gamma} \vec{F} \cdot d\vec{r} = 0$ پس حکم نتیجه می‌شود. □

کاربرد ۳ فرض کنید $\gamma, \gamma_1, \dots, \gamma_k$ خمهای بسته ساده قطعه هموار و همجهت باشند که تصویرهای $\gamma_1, \dots, \gamma_k$ دوه‌دو متخارج بوده و همگی در درون تصویر γ قرار داشته باشند (شکل ۳). اگر میدان برداری \vec{F} در سراسر ناحیه بین خمهای γ_i و γ تعریف شده، C^1 باشد، و داشته باشیم $\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$ ، آنگاه:

$$\int_{\gamma} \vec{F} \cdot d\vec{r} = \sum_{i=1}^k \int_{\gamma_i} \vec{F} \cdot d\vec{r} \quad (4)$$

در شکل (۳) ناحیه بین γ_i ها و γ با نقطه‌چین مشخص شده است. برای اثبات (۴)، D را ناحیه بین γ_i ها و γ می‌گیریم. اگر ∂D (اجتماع تصویر γ_i ها و γ) را با جهت قراردادی منظور کنیم، طبق قضیه گرین داریم:

$$\left(\int_{\partial D} \vec{F} \cdot d\vec{r} \right) = 0 \quad (5)$$

ولی ∂D از اجتماع تصویر γ (با جهت مثلثاتی) و تصاویر γ_i ها (با جهت عقربه‌ ساعت) تشکیل شده است. بنابراین اگر γ_i ها و γ را همجهت بگیریم (یعنی یا γ_i ها را تغییر جهت دهیم، یا γ را)، از (۵)

رابطه (۴) حاصل می شود زیرا که تعویض جهت علامت مقدار انتگرال می شود.

بدین ترتیب هر چند شرط (۳)، به نهایبی، ممکن است دلالت بر پایسته بودن \vec{F} نکند، ولیکن مقدار انتگرال $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای هر دو خم ساده بسته ای که ناحیه بین آنها در دامنه تعریف قرار داشته باشد برابر است. نمونه مهمی از این کاربرد، برای $\vec{F} = (\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2})$ ظاهر می شود که شرط (۳) برای آن برقرار است. دیده ایم که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برای هر دایره شعاع R به مرکز \underline{e} که یک بار در جهت مثلثاتی طی شود برابر (2π) است. نتیجه این که اگر به جای γ ، هر خم بسته ساده قطعه قطعه هموار که یک بار در جهت مثلثاتی حول نقطه \underline{e} می چرخد نیز برابر 2π می باشد زیرا که میدان \vec{F} در ناحیه بین این خم در یک دایره شعاع مناسب تعریف شده است. در واقع اگر خم بسته γ ساده نباشد، یعنی خود را قطع کند، با تجزیه γ به چند قطعه بسته ساده مشاهده می شود که $\int_{\gamma} \vec{F} \cdot d\vec{r}$ برابر $(2k\pi)$ است که k تعداد دفعات چرخش γ حول \underline{e} در جهت مثلثاتی است (برای جهت عقربه ساعت، k منفی می شود). از این رو، تعریف دقیقی برای "دفعات گردش" خم بسته γ ، که در $\mathbb{R}^2 - \{\underline{e}\}$ قرار دارد، حول \underline{e} به صورت زیر ارائه می شود:

$$I_{\gamma}(\underline{e}) = \left(\frac{1}{2\pi}\right) \int_{\gamma} \left(\frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy\right) \quad (6)$$

طبیعی است که اگر به جای \underline{e} ، نقطه دیگری p در صفحه در نظر بگیریم، با انتقال مختصات می توان به طور مشابه $I_{\gamma}(p)$ را تعریف کرد. در شکل ۴ چند نمونه نمایش داده شده است.

شکل ۴

(۳-۳۹) اثبات قضیه گرین نخست قضیه گرین را برای یک مستطیل ثابت می کنیم. فرض کنید:

$$D = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$$

∂D از چهار پاره خط $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ ، به صورت مشخص شده در شکل (۵) تشکیل شده است.

شکل ۵

خم های مرزی را می توان به شکل زیر پرمایش کرد:

$$\gamma_1(t) = (t, c), a \leq t \leq b, \quad \gamma_3(t) = (a + b - t, d), a \leq t \leq b$$

$$\gamma_2(t) = (b, t), c \leq t \leq d, \quad \gamma_4(t) = (a, c + d - t), c \leq t \leq d$$

برای محاسبه $\iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA$ (طرف راست (۱))، دو انتگرال دوگانه جداگانه محاسبه می کنیم:

$$\begin{aligned} \iint_D \frac{\partial F_2}{\partial x} dA &= \int_c^d \int_a^b \frac{\partial F_2}{\partial x} dx dy \\ &= \int_c^d [F_2(b, y) - F_2(a, y)] dy \\ &= \int_c^d F_2(b, y) dy - \int_c^d F_2(a, y) dy \end{aligned}$$

از طرفی دیگر:

$$\int_{\gamma_2} \vec{F} \cdot d\vec{r} = \int_c^d (F_1 \frac{dx}{dt} + F_2 \frac{dy}{dt}) dt$$

در امتداد γ_2 ، $x = b$ ثابت است، پس $\frac{dx}{dt} = 0$ و $\frac{dy}{dt} = 1$ ، پس:

$$\int_{\gamma_2} \vec{F} \cdot d\vec{r} = \int_c^d F_2(b, t) dt = \int_c^d F_2(b, y) dy$$

و نیز مشابهاً می بینیم که $\int_{\gamma_4} \vec{F} \cdot d\vec{r} = - \int_c^d F_2(a, y) dy$ ، پس

$$\iint_D \frac{\partial F_2}{\partial x} dA = \int_{\gamma_2} \vec{F} \cdot d\vec{r} + \int_{\gamma_4} \vec{F} \cdot d\vec{r}$$

به همین ترتیب می توان دید که $-\iint_D \frac{\partial F_1}{\partial y} dA = \int_{\gamma_1} \vec{F} \cdot d\vec{r} + \int_{\gamma_3} \vec{F} \cdot d\vec{r}$ و اثبات قضیه گرین برای مستطیل کامل می شود. توجه کنید که در این اثبات عنصر اصلی، استفاده از قضیه اساسی حساب دیفرانسیل و انتگرال است.

در گام بعد، ناحیه هایی را در نظر می گیریم که مرز آنها اجتماع پاره خط هایی به موازات محورهای x و y است. این ناحیه ها را می توان با رسم خطوطی به موازات دو محور به اجتماع مستطیل هایی از نوع بالا تجزیه کرد که فقط در مرز اشتراک دارند (شکل ۶)

شکل ۶

ناحیه D به صورت اجتماعی متناهی از مستطیل‌های D_1, \dots, D_k نوشته می‌شود که برای هر یک قضیهٔ گرین ثابت شده است. در جمع زدن $\int \vec{F} \cdot d\vec{r}$ برای مرزهای این مستطیل‌ها، مشاهده می‌کنیم که پاره‌خط‌های واقع در داخل D دوبار در جهت‌های مخالف طی می‌شوند، پس اثر آنها در مجموع صفر است در حالی که پاره‌خط‌های تشکیل دهندهٔ ∂D فقط یک بار ظاهر می‌شوند و در مجموع، $\int_{\partial D} \vec{F} \cdot d\vec{r}$ را می‌سازند. بدین ترتیب قضیهٔ گرین برای این نوع نواحی ثابت می‌شود.

شکل ۷

بالاخره در حالت کلی، با رسم خطوط راستی موازی دو محور، افرازی از D به دست می‌آوریم و تقریبی از این ناحیه که به شکل گام بالا است (شکل ۷). می‌توان نشان داد (با جزئیاتی که در اینجا مطرح نخواهد شد) که باظریفتر کردن افراز، انتگرال دوگانه روی ناحیهٔ تقریب D به انتگرال دوگانه روی D میل می‌کند و نیز انتگرال روی خم مرزی به انتگرال روی مرز D میل خواهد کرد، و بدین ترتیب اثبات قضیهٔ گرین در حالت کلی به دست می‌آید.

آنالیز برداری (۲)

قضیه گرین انتگرال دوگانه عبارت $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ روی یک ناحیه مسطح در صفحه را برحسب انتگرال $\vec{F} \cdot d\vec{r}$ روی مرز ناحیه بیان می کنند. در نظر اول ممکن است عبارت $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ مصنوعی به نظر رسد. ولی اگر $\vec{F} = (F_1, F_2, F_3)$ را در نظر بگیریم، عبارت بالا مؤلفه سوم میدان $\text{curl} \vec{F}$ است. در واقع تعمیم مهمی از قضیه گرین به ناحیه های خمیده (تصویر رویه های قطعه قطعه هموار) وجود دارد که در آن به جای این عبارت از $\text{curl} \vec{F}$ استفاده می شود. این قضیه را در زیر نخست در ساده ترین حالت در نظر می گیریم.

فرض کنید $\varphi : W \rightarrow \mathbb{R}^3$ یک رویه هموار با شرایط اضافی زیر باشد:

الف) W دارای مرزی است که از یک یا چند (تعدادی متناهی) تصویر خم قطعه قطعه هموار تشکیل شده است.

ب) φ یک به یک است.

در این صورت $D = \varphi(W)$ را یک رویه ساده می نامیم. مرز D ، ∂D ، برابر $\varphi(\partial W)$ تعریف می شود.

مثال ۱ فرض کنید $W = \{(u, v) \mid 0 \leq u \leq \pi, -1 \leq v \leq 1\}$ و $\varphi : W \rightarrow \mathbb{R}^3$ به صورت زیر تعریف می شود:

$$\varphi(u, v) = (\cos u, \sin u, v)$$

به سادگی دیده می شود که φ یک به یک است. داریم

$$\begin{aligned} \varphi_u \times \varphi_v &= (-\sin u, \cos u, 0) \times (0, 0, 1) \\ &= (\cos u, \sin u, 0) \end{aligned}$$

به ازای هر (u, v) داریم $\varphi_u \times \varphi_v \neq \mathbf{0}$ ، پس φ یک رویه هموار تعریف می‌کند. $D = \varphi(W)$ روی سطح استوانه $x^2 + y^2 = 1$ دارد و نیمی از قطعه $-1 \leq z \leq 1$ ، یعنی نیمه $y \geq 0$ ، را دربرمی‌گیرد (شکل ۱). مرز D از دو پاره خط راست و دو نیمدایره تشکیل شده است.

شکل ۱

مثال ۲ فرض کنید

$$W = \{(u, v) \mid 1 \leq |u| + |v| \leq 2\}$$

W ناحیه بین دو مربع در صفحه uv است (شکل ۲)

شکل ۲

حال $\varphi: W \rightarrow \mathbb{R}^3$ را به صورت $\varphi(u, v) = (u, v, u^2 + v^2)$ تعریف می‌کنیم. $D = \varphi(W)$ قطعه ای از نمودار تابع $z = x^2 + y^2$ است و می‌دانیم که نمودار یک تابع C^1 تصویر یک رویه هموار است. φ در اینجا به وضوح یک به یک است. مرز D از دو خم تشکیل شده است که هر یک قطعه هموار و اجتماع چهار قطعه هموار است (شکل ۳).

شکل ۳

یادآوری می‌کنیم که برای رویه ساده $D = \varphi(W)$ جهت قراردادی با قائم واحد

$$\vec{n} = \frac{\varphi_u \times \varphi_v}{|\varphi_u \times \varphi_v|}$$

تعریف می‌شود. همچنان که در مورد مرز یک ناحیه دوبعدی عمل کردیم، اکنون جهت مشخصی برای ∂D منظور می‌کنیم.

در هر نقطه مرزی غیر گوشه‌ای، \vec{t} را بردار واحدی می‌گیریم که بر رویه D مماس است، بر مرز یعنی ∂D ، عمود است و به درون D اشاره می‌کند. حال مماس واحد، \vec{T} ، در یک نقطه غیر گوشه‌ای را آن مماس واحد بر ∂D در آن نقطه می‌گیریم که $(\vec{T}, \vec{v}, \vec{n})$ یک سه‌تایی مرتب راستگرد باشد. اگر

انسانی را تصور کنیم که روی ∂D ، در جهت \vec{T} حرکت می کند و دست چپ او به طرف درون D است، جهت سر این انسان متحرک باید به طرف قائم قراردادی روبه، یعنی \vec{n} باشد.

برای مثالهای ۱ و ۲، \vec{T} را محاسبه می کنیم. در مثال ۱ دیدیم که $\vec{n} = (\cos u, \sin u, 0)$ روی قطعه افقی $t = -1$ که با $v = -1$ مشخص می شود، داریم $\vec{v} = (0, 0, 1)$ پس

$$\vec{T} = \vec{v} \times \vec{n} = (-\sin u, \cos u, 0)$$

به همین ترتیب ملاحظه می کنیم که روی قطعه قائم متناظر با $u = \pi$ ، داریم $\vec{n} = (0, 0, 1)$ روی قطعه افقی متناظر با $v = 1$ ، $\vec{T} = (\sin u, -\cos u, 0)$ و بالاخره روی قطعه قائم متناظر با $u = 0$ ، $\vec{T} = (0, 0, -1)$

در مثال ۲، قائم واحد عبارت است از:

$$\begin{aligned} \frac{\varphi_u \times \varphi_v}{|\varphi_u \times \varphi_v|} &= \frac{(1, 0, 2u) \times (0, 1, 2v)}{|(1, 0, 2u) \times (0, 1, 2v)|} \\ &= \frac{(-2u, -2v, 1)}{\sqrt{4u^2 + 4v^2 + 1}} \end{aligned}$$

که قائم بر نمودار سهمی وار $z = x^2 + y^2$ است که مؤلفه سوم آن مثبت است، یعنی به طرف 'داخل' سهمی وار، یعنی ناحیه $z > x^2 + y^2$ اشاره دارد. نتیجه این که جهت ∂D طبق شکل ۱۰ می باشد. توجه کنید که اگر این جهت ها را بر صفحه xy (= صفحه uv) تصویر کنیم (ناحیه شکل ۹)، مربع داخلی در جهت عقربه ساعت و مربع خارجی در جهت مثلثاتی جهت می گیرد که سازگار با جهت مرز در قضیه گرین است.

این در واقع کلیت دارد زیرا که φ_u تصویر $(1, 0)$ تحت $D\phi(u, v)$ است، φ_v تصویر $(0, 1)$ تحت $D\phi(u, v)$ است، پس $\vec{n} = \frac{\vec{\varphi}_u \times \vec{\varphi}_v}{|\vec{\varphi}_u \times \vec{\varphi}_v|}$ مانند $(\vec{i}, \vec{j}, \vec{k})$ راستگرد است. حال چون تحت φ ، درون W به درون $D = \varphi(W)$ نگاشته می شود، جهت مرزها نیز متناظر خواهد شد.

اکنون می توانیم قضیه استوکس را برای روبه های ساده بیان کنیم. در زیر مرز روبه طبق قرارداد بالا جهت داده شده است.

(۴۰-۲) قضیه استوکس (روبه های ساده) $D = \varphi(W)$ یک روبه ساده در \mathbb{R}^3 است که φ تابعی C^2 است و مرز روبه، ∂D ، طبق قرارداد جهت داده شده است و \vec{n} قائم واحد قراردادی روی D می باشد،

اگر \vec{F} یک میدان برداری C^1 باشد که دامنه تعریف آن شامل D و مرز آن است، داریم:

$$\int_{\partial D} \vec{F} \cdot d\vec{r} = \int_D \int \text{curl} \vec{F} \cdot \vec{n} dS \quad (1)$$

برهان همان طور که انتظار می رود فرمول بالا با نوشتن (x, y, z) به صورت $\varphi(u, v)$ و استفاده از قاعده زنجیره ای به فرمول گرین در صفحه uv برای W و ∂W تبدیل می شود. طبق معمول می نویسیم:

$$(x, y, z) = \vec{r} = \varphi(u, v) = (\varphi_1(u, v), \varphi_2(u, v), \varphi_3(u, v))$$

نخست عبارت داخل انتگرال سمت چپ را بر حسب (u, v) می نویسیم:

$$\begin{aligned} \vec{F} \cdot d\vec{r} &= F_1 dx + F_2 dy + F_3 dz \\ &= F_1 \left(\frac{\partial \varphi_1}{\partial u} du + \frac{\partial \varphi_1}{\partial v} dv \right) + F_2 \left(\frac{\partial \varphi_2}{\partial u} du + \frac{\partial \varphi_2}{\partial v} dv \right) + F_3 \left(\frac{\partial \varphi_3}{\partial u} du + \frac{\partial \varphi_3}{\partial v} dv \right) \end{aligned}$$

پس

$$\vec{F} \cdot d\vec{r} = (\vec{F} \cdot \vec{r}_u) du + (\vec{F} \cdot \vec{r}_v) dv$$

و طبق قضیه گرین در صفحه داریم

$$\begin{aligned} \int_{\partial D} \vec{F} \cdot d\vec{r} &= \int_{\partial W} [(\vec{F}(\varphi(u, v)) \cdot \vec{r}_u) du + (\vec{F}(\varphi(u, v)) \cdot \vec{r}_v) dv] \\ &= \int \int_W \left[\frac{\partial}{\partial u} (\vec{F}(\varphi(u, v)) \cdot \vec{r}_v) - \frac{\partial}{\partial v} (\vec{F}(\varphi(u, v)) \cdot \vec{r}_u) \right] dA \end{aligned}$$

عبارت داخل انتگرال را بسط می دهیم:

$$\begin{aligned} & \left[\frac{\partial}{\partial u} (F_1 x_v + F_2 y_v + F_3 z_v) - \frac{\partial}{\partial v} (F_1 x_u + F_2 y_u + F_3 z_u) \right] dudv \\ &= \left[\frac{\partial (F_1 \circ \varphi)}{\partial u} x_v + (F_1 \circ \varphi) x_{vu} - \frac{\partial (F_1 \circ \varphi)}{\partial v} x_u - (F_1 \circ \varphi) x_{uv} \right. \\ & \left. + (\text{دو دسته چهارتایی جملات مشابه}) \right] dudv \end{aligned}$$

چون مشتقات پاره ای مرتبه دوم φ پیوسته فرض شده اند داریم $x_{vu} = x_{uv}$ و نصف جملات حذف

می‌شوند. با استفاده از قاعده زنجیره‌ای عبارت بالا برابر می‌شود با:

$$\begin{aligned} & \left[\left(\frac{\partial F_1}{\partial x} x_u + \frac{\partial F_1}{\partial y} y_u + \frac{\partial F_1}{\partial z} z_u \right) x_v - \left(\frac{\partial F_1}{\partial x} x_v + \frac{\partial F_1}{\partial y} y_v + \frac{\partial F_1}{\partial z} z_v \right) x_u \right. \\ & \left. + (\text{دو دسته جملات مشابه}) \right] dudv \\ & = \left[-\frac{\partial F_1}{\partial y} \cdot \frac{\partial(x, y)}{\partial(u, v)} + \frac{\partial F_1}{\partial z} \cdot \frac{\partial(z, x)}{\partial(u, v)} - \frac{\partial F_2}{\partial z} \cdot \frac{\partial(y, z)}{\partial(u, v)} + \frac{\partial F_2}{\partial x} \cdot \frac{\partial(x, y)}{\partial(u, v)} - \frac{\partial F_2}{\partial x} \cdot \frac{\partial(z, x)}{\partial(u, v)} \right. \\ & \left. + \frac{\partial F_2}{\partial y} \cdot \frac{\partial(y, z)}{\partial(u, v)} \right] dudv \\ & = \left[\left(\frac{\partial F_2}{\partial y} - \frac{\partial F_2}{\partial z} \right) \frac{\partial(y, z)}{\partial(u, v)} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_2}{\partial x} \right) \frac{\partial(z, x)}{\partial(u, v)} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \frac{\partial(x, y)}{\partial(u, v)} \right] dudv \end{aligned}$$

■ که انتگرال عبارت بالا روی W دقیقاً برابر $\int \int_D (\vec{F} \cdot \vec{n}) dS$ است و حکم به اثبات می‌رسد.

مثال ۱ M را بخشی از نمودار $z = f(x, y) = 1 - x^2 - y^2$ می‌گیریم که در بالای صفحه xy ، یعنی در $z \geq 0$ قرار دارد و \vec{n} را قائم واحد به طرف z های مثبت روی این رویه می‌گیریم. میدان \vec{F} به صورت زیر داده شده است:

$$\vec{F}(x, y, z) = (x - y)\vec{i} + (x + y^2)\vec{j} + z^2\vec{k}$$

می‌خواهیم $\int \int_D \text{curl} \vec{F} \cdot \vec{n} dS$ را محاسبه کنیم. ∂D عبارت است از دایره واحد $x^2 + y^2 = 1$ ، $z = 0$. با توجه به این که D به وسیله $(u, v, 1 - u^2 - v^2) \mapsto (u, v)$ ، $u^2 + v^2 \leq 1$ ، پرمایش می‌شود و قائم واحد عبارت است از:

$$\frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|} = \frac{(1, 0, -2u) \times (0, 1, -2v)}{|(1, 0, -2u) \times (0, 1, -2v)|} = \frac{(2u, 2v, 1)}{\sqrt{4u^2 + 4v^2 + 1}}$$

که مؤلفه z آن مثبت است و برای به کارگیری قضیه گرین برای $u^2 + v^2 \leq 1$ ، باید مرز را در جهت مثلثاتی پرمایش کرد، جهت ∂D نیز باید جهت مثلثاتی در صفحه xy اتخاذ شود تا قضیه استوکس برقرار گردد:

$$\int \int_D \text{curl} \vec{F} \cdot \vec{n} dS = \int_{\partial D} \vec{F} \cdot d\vec{r}$$

می‌توان در اینجا انتگرال سمت راست را محاسبه کرد ولی کار ساده‌تر این است که مجدداً قضیه استوکس را این بار برای رویه $x^2 + y^2 \leq 1, z = 0$ به کار گیریم که نتیجه می‌دهد:

$$\begin{aligned} \int_{\partial M} \vec{F} \cdot d\vec{r} &= \int \int_{x^2+y^2 \leq 1} (\text{curl} \vec{F} \cdot \vec{k}) dA \\ &= \int \int_{x^2+y^2 \leq 1} ((0 - 0)\vec{i} + (0 - 0)\vec{j} + (1 + 1)\vec{k}) \cdot \vec{k} dA \\ &= \int \int_{x^2+y^2 \leq 1} 2 dA = 2\pi \end{aligned}$$

برای بهره‌گیری کامل از قضیه استوکس لازم است که آن را به "رویه‌های بزرگتر"، یعنی رویه‌هایی که نتوان (یا به آسانی نتوان) آنها را به صورت رویه ساده پرمایش کرد تعمیم دهیم. زیرمجموعه D از \mathbb{R}^3 را یک رویه جهت‌پذیر می‌نامیم در صورتی $D = D_1 \cup \dots \cup D_k$ و شرایط زیر برقرار باشند:

(الف) هر D_i یک رویه ساده است.

(ب) دو D_i متمایز فقط در مرز می‌توانند اشتراک داشته باشند و $D_i \cap D_j$ را می‌توان به صورت اجتماعی متناهی از تصاویر خم‌های هموار نوشت

(ج) هر قطعه خم مرزی D_i با حداکثر یک $D_j, j \neq i$ ، مشترک است.

(د) برای D_i ها می‌توان طوری جهت اختیار کرد که جهت قراردادی مرز مشترک بین D_i و D_j متضاد باشد.

در شکل ۳ دو نمونه رویه جهت‌پذیر نمایش داده شده است. ۳ (الف) از چسباندن سه چنبره به هم به دست آمده است. از چنبره میانی دو قطعه و از هر یک چنبره‌های کناری یک قطعه بریده و در امتداد برش‌ها چنبره‌ها را به هم چسبانده‌ایم. به برقرار بودن شرط (د) توجه کنید. در شکل (ب) از استوانه افقی سه خم بسته γ_1, γ_2 و γ_3 را بریده‌ایم به طوری که D_1 دارای مرزی متشکل از پنج خم بسته $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$ است. در امتداد $\gamma_1, \gamma_2, \gamma_3$ و γ_4 سه استوانه دیگر به D_1 چسبانده‌ایم. سه خم $\gamma_1, \gamma_2, \gamma_3$ و γ_4 جهت‌های متضاد از D_i ها می‌گیرند. مرز $D = D_1 \cup D_2 \cup D_3 \cup D_4$ از پنج قطعه $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$ و γ_6 با جهت نمایش داده شده تشکیل شده است. در ۳ (الف) مرز D تهی است. این نمونه یک رویه "جهت‌پذیر بسته" است.

شکل ۳

در شکل ۴ "نوار مویبوس" نمایش داده شده است که جهت پذیر نیست. اگر دو سر یک نوار کاغذ را پس از یک دو تاب دادن در فضا به هم بچسبانیم نوار مویبوس حاصل می شود. این رویه را نمی توان به صورت یک رویه جهت پذیر نمایش داد.

شکل ۴

مثلاً اگر دو برش cd و ef را طبق شکل ۵ انجام دهیم و روی هر یک از رویه حاصل شده جهتی اختیار کنیم

شکل ۵

مشاهده می شود که بسته به جهت های اختیار شده، یکی از دو قطعه مرزی cd یا ef جهت های متوافق کسب می کنند و دیگری جهت متضاد. نمود دیگر جهت ناپذیری یک رویه در \mathbb{R}^3 این است که نمی توان برای آن یک میدان برداری قائم واحد اختیار کرد که به طور پیوسته تغییر کند. مثلاً اگر در یک نقطه دایره نقطه چین شکل ۴ یک قائم واحد بر رویه اختیار کنیم و آن را یک دو کامل در طول این خم بسته به طور پیوسته دنبال کنیم، در بازگشت به نقطه آغازی جهت آن معکوس خواهد شد.

حال فرض کنید $D = D_1 \cup \dots \cup D_k$ یک رویه جهت پذیر باشد و برای D_i ها جهت هایی طبق (د) اختیار کنید که خم های اشتراک جهت های متضاد کسب کنند. اگر فرمول استوکس را برای هر D_i بنویسیم و طرف های متناظر را با هم جمع کنیم، در طرف راست (۱) $\int \int_D \text{curl } \vec{F} \cdot \vec{n} dS$ حاصل می شود و در طرف چپ کلیه مرزهای مشترک به دلیل جهت های متضاد حذف شده و تنها انتگرال $\int \vec{F} \cdot d\vec{r}$ برای آن خم های مرزی باقی می ماند که متعلق به فقط یک قطعه هستند. این خم ها دقیقاً مرز D را تشکیل می دهند. در ضمن اگر رویه D بسته باشد، همه خم های مرزی مشترک هستند و مجموع انتگرال های سمت چپ صفر می شود. بدین ترتیب فرمول کلی استوکس به دست می آید:

(۴۰-۲) قضیه استوکس اگر D یک رویه جهت پذیر با مرز ∂D باشد و بردار قائم واحد و جهت ∂D به طور سازگار با قرارداد انتخاب شده باشند، برای هر میدان برداری با مشتق های پاره ای پیوسته \vec{F} که

دامنه آن شامل D و ∂D باشد داریم:

$$\int_{\partial D} \vec{F} \cdot d\vec{r} = \int \int_D \text{curl} \vec{F} \cdot \vec{n} dS \quad (2)$$

اگر رویه بسته نیز باشد، لزوماً:

$$\int \int_D \text{curl} \vec{F} \cdot \vec{n} dS = 0 \quad (3)$$

مثال ۲ نمودار تابعی مشتق پذیر $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ استوانه $x^2 + y^2 = 1$ را در خمی γ قطع می کند. میدان برداری $\vec{F}(x, y, z) = (-y)\vec{i} + x\vec{j}$ را در نظر بگیرید. نشان دهید که $\int_{\gamma} \vec{F} \cdot d\vec{r} = \pm 2\pi$ مستقل از تابع خاص f .

دو تابع مشتق پذیر مختلف $f, g: \mathbb{R}^2 \rightarrow \mathbb{R}$ در نظر بگیرید که اشتراک نمودارشان با $x^2 + y^2 = 1$ از یکدیگر مجزاست

شکل ۶

ناحیه بین دو اشتراک روی سطح استوانه را D می نامیم. در شکل ۶ داریم $\partial D = \beta \cup (-\alpha)$ ، که مقصود از $(-\alpha)$ بازپرمایش جهت برگردان α است. داریم $\text{curl} \vec{F} = \underline{0}$ ، پس $\int_{\beta} \vec{F} \cdot d\vec{r} - \int_{\alpha} \vec{F} \cdot d\vec{r} = 0$ ، یعنی همان طور که ادعا شده بود $\int_{\gamma} \vec{F} \cdot d\vec{r}$ مستقل از تابع خاص f است. بنابراین برای محاسبه $\int_{\gamma} \vec{F} \cdot d\vec{r}$ کافی است یک حالت خاص، مثلاً تابع ثابت $f(x, y) \equiv 0$ در نظر گرفته شود که اشتراک نمودار آن با استوانه، دایره واحد در صفحه xy است. برای این دایره می دانیم که $\int_{\gamma} -y dx + x dy = \pm 2\pi$ جهت γ مثلثاتی یا ضد مثلثاتی باشد.

(۳-۴۰) کاربرد در وجود پتانسیل

به عنوان یک کاربرد قضیه گرین در صفحه دیدیم که هرگاه $\vec{F} = F_1\vec{i} + F_2\vec{j}$ یک میدان برداری دارای مشتق های پاره ای پیوسته باشد که در مجموعه باز و همبند مسیری U در \mathbb{R}^2 تعریف شده است و شرط لازم پایستگی $\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$ برقرار باشد، \vec{F} پایسته است در صورتی که U ویژگی هندسی اضافی

زیرا داشته باشد: برای هر خم ساده بسته قطعه قطعه هموار در U ، ناحیه محصور توسط این خم به تمامی در U واقع باشد. اکنون می‌توانیم با استفاده از قضیه استوکس تعمیم مناسبی برای این مطلب در \mathbb{R}^3 ارائه کنیم. U را یک مجموعه باز و همبند مسیری در \mathbb{R}^3 می‌گیریم و فرض می‌کنیم میدان برداری \vec{F} ، که دارای مشتق‌های پاره‌ای پیوسته است، در U تعریف شده است. شرطی لازم برای پایسته بودن \vec{F} این است که:

$$\text{curl} \vec{F} = \underline{0} \quad (4)$$

این شرط برای پایسته بودن \vec{F} کافی نیست زیرا که در مورد

$$\vec{F}(x, y, z) = \frac{-y}{x^2 + y^2} \vec{i} + \frac{x}{x^2 + y^2} \vec{j}$$

که در مجموعه باز و همبند مسیری $U = \mathbb{R}^3 - \{ \text{محور } z \}$ تعریف شده است داریم $\text{curl} \vec{F} = 0$ ولیکن انتگرال $\int \vec{F} \cdot d\vec{r}$ روی دایره واحد در صفحه xy صفر نیست. ویژگی هندسی اضافی مورد نیاز در اینجا بدین شرح است: برای هر خم بسته ساده قطعه قطعه هموار γ در U ، روبه‌ای جهت‌پذیر D ، کاملاً واقع شده در U وجود داشته باشد که $\partial D = \gamma$.

در صورت احراز این شرط، $\int_D \text{curl} \vec{F} \cdot \vec{n} dS = 0$ نتیجه می‌دهد $\int_\gamma \vec{F} \cdot d\vec{r} = 0$ برای هر خم ساده بسته قطعه قطعه هموار در U ، پس \vec{F} دارای پتانسیل است.

به عنوان مثال توجه کنید که $\mathbb{R}^3 - \{(0, 0, 0)\}$ ویژگی مورد نظر را دارد زیرا که مثلاً برای هر خم بسته ساده حول z در صفحه xy می‌توان روبه‌ای در $z > 0$ در نظر گرفت که γ را به عنوان مرز دارد. مثلاً نیمکره شمالی این ویژگی را دارد.

(4-40) تعبیر curl فرض کنید D_ρ یک گوی مسطح دو بعدی به شعاع ρ و به مرکز نقطه P در \mathbb{R}^3 است. \vec{F} یک میدان برداری است با مشتق‌های پاره‌ای پیوسته، که دامنه آن شامل D_ρ و مرز آن است. طبق فرمول استوکس داریم:

$$\int_{\partial D_\rho} \vec{F} \cdot d\vec{r} = \int \int_{D_\rho} \text{curl} \vec{F} \cdot \vec{n} dS$$

با استفاده از قضیه میانگین انتگرال، انتگرال سمت راست برابر می شود با $(\text{curl } \vec{F} \cdot \vec{n})(Q)$ که $(\pi \rho^2)$ نقطه ای در D_ρ است. پس:

$$(\text{curl } \vec{F} \cdot \vec{n})(Q) = \frac{\int_{\partial D_\rho} \vec{F} \cdot d\vec{r}}{\pi \rho^2} \quad (5)$$

حال $\int_\gamma \vec{F} \cdot d\vec{r}$ گردش \vec{F} حول γ نام دارد زیرا که به مفهومی "مجموع" مؤلفه های \vec{F} در راستای مماس بر γ است. بدین ترتیب طرف راست (5) گردش متوسط \vec{F} نسبت به مساحت را حول P در صفحه عمود بر \vec{n} نمایش می دهد. وقتی $\rho \rightarrow 0$ ، بنابر پیوستگی مشتق های پاره ای، $(\text{curl } \vec{F} \cdot \vec{n})(Q)$ به $(\text{curl } \vec{F} \cdot \vec{n})(P)$ میل می کند. بدین ترتیب تعبیر زیر برای $\text{curl } \vec{F}$ حاصل می شود: $\text{curl } \vec{F}(P)$ برداری است که تصویر آن روی هر راستا برابر گردش لحظه ای \vec{F} حول آن راستا نسبت به واحد مساحت است. چون حداکثر $\text{curl } \vec{F} \cdot \vec{n}(P)$ وقتی به دست می آید که \vec{n} در راستای $\text{curl } \vec{F}$ باشد، امتداد $\text{curl } \vec{F}$ را محور چرخش لحظه ای \vec{F} در نقطه P می نامند. به علاوه $|\text{curl } \vec{F}|$ نیز معمولاً سرعت زاویه ای لحظه ای این چرخش خوانده می شود. دلیل این نامگذاری مثال ساده زیر است:

مثال میدان $\vec{F}(x, y, z) = -\omega y \vec{i} + \omega x \vec{j}$ را در نظر بگیرید. این میدان نمایش دهنده یک دوران حول محور z با سرعت زاویه ای ثابت ω است. داریم:

$$\text{curl } \vec{F} = (0, 0, 2\omega) = (2\omega) \vec{k}$$

بدین ترتیب مشاهده می شود که امتداد $\text{curl } \vec{F}$ در واقع محور چرخش است و ضریب \vec{k} دو برابر سرعت زاویه ای می باشد.

بدین ترتیب $\text{curl } \vec{F}(P)$ تمایل میدان \vec{F} به حالت چرخشی حول P را نمایش می دهد و امتداد $\text{curl } \vec{F}(P)$ محور بیشترین چرخش است. با این تعبیر می توان رابطه اساسی $\text{curl}(\text{grad } f) = \underline{0}$ را بازبینی کرد. به یاد آورید که خطوط میدان $\text{grad } f$ همواره در جهت بیشترین صعود f در حرکتند، پس خطوط میدان $\text{grad } f$ تمایلی به بازگشت به نقطه شروع (چرخش) را ندارند که با $\text{curl}(\text{grad } f) = \underline{0}$ سازگار است.

از این تعبیر باید با احتیاط و دقت استفاده کرد. به مثال زیر توجه کنید:

مثال میدان $\vec{F}(x, y, z) = x\vec{j}$ را در نظر می‌گیریم. این میدان روی صفحه yz صفر است و خارج از آن دارای خطوط میدان موازی محور y می‌باشد (شکل ۷ وضعیت میدان را در صفحه xy نشان می‌دهد).

شکل ۷

به نظر نمی‌رسد هیچ گونه تمایلی به چرخش در این میدان وجود داشته باشد زیرا که خطوط میدان خارج از صفحه yz خط راست هستند. ولیکن محاسبه ساده نشان می‌دهد که $\text{curl}\vec{F} = \vec{k}$. چگونه باید این "تناقض" را توجیه کرد؟ یک راه توجیه این است که مشاهده کنیم در دو طرف صفحه yz جهت میدان عوض می‌شود. این امر به سبب صفر شدن میدان روی صفحه yz امکان پذیر است. می‌توان تصور کرد که در اینجا نوعی چرخش در جریان است که مسیر آن تا "بی‌نهایت" در امتداد محور y ادامه دارد. راه دیگر این است که یک چرخه سبک چهارپره‌ای را تصور کنیم که تحت تأثیر "میدان نیروی" $\vec{F} = x\vec{j}$ قرار دارد. مثلاً اگر این چرخه در نیمه $x > 0$ قرار داشته باشد، نیرویی که به پره‌های راست آن وارد می‌آید بزرگتر از نیرویی است که در همان جهت به پره‌های سمت چپ وارد می‌شود، پس چرخه دور خود خواهد چرخید. بدین ترتیب شاید بهتر باشد به جای این که $\text{curl}\vec{F}$ را نمایشگر تمایل خود میدان \vec{F} به چرخش فرض کنیم، آن را نمایشگر توان میدان \vec{F} برای چرخاندن اشیاء سبکی که در معرض اثر آن قرار می‌گیرند تعبیر کنیم.

آنالیز برداری (۳)

آخرین قضیه از رشته قضایایی که انتگرال روی یک مجموعه را به انتگرال روی مرز مرتبط می‌سازد 'قضیه دیوژرانس' است که در این بخش به آن خواهیم پرداخت. نخست حالت دوبعدی این قضیه را که در واقع یک بازنویسی قضیه گرین است بررسی می‌کنیم.

ناحیه D در صفحه xy را در نظر بگیرید که در شرایط قضیه گرین (قضیه ۳۹-۱) صدق می‌کند یعنی مرز آن ∂D از اجتماعی متناهی خمهای قطعه قطعه هموار تشکیل شده است و ∂D طبق قرارداد جهت داده شده است. مماس واحد در نقاط غیرگوشه‌ای را در نظر بگیرید، $\vec{T} = \left(\frac{dx}{ds}, \frac{dy}{ds}\right)$ قائم واحد برونگرا نسبت به D را، که به \vec{n} نمایش می‌دهیم، از دوران \vec{T} به اندازه $\frac{\pi}{2}$ در جهت عقربه‌ساعت به دست می‌آید، پس $\vec{n} = \left(\frac{dy}{ds}, -\frac{dx}{ds}\right)$. حال فرض کنید $\vec{G} = (G_1, G_2)$ یک میدان برداری C^1 است که دامنه تعریف آن شامل D و ∂D است. در این صورت داریم:

$$\vec{G} \cdot \vec{n} = G_1 \frac{dy}{ds} - G_2 \frac{dx}{ds}$$

میدان برداری $\vec{F} = (F_1, F_2)$ را با روابط $F_2 = G_1$ و $F_1 = -G_2$ تعریف می‌کنیم، پس

$$\vec{G} \cdot \vec{n} = \vec{F} \cdot \vec{T} \quad (1)$$

از طرفی دیگر

$$\operatorname{div} \vec{G} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \quad (2)$$

پس با جایگزینی (۱) و (۲) در قضیه گرین برای میدان برداری \vec{F} داریم:

$$\int_{\partial D} \vec{G} \cdot \vec{n} ds = \int_D \operatorname{div} G dA_{x,y} \quad (3)$$

فرمول (۳) صورت دیورژانس قضیه گرین نام دارد. این قضیه گاهی بدین صورت بیان می شود که انتگرال دیورژانس در ناحیه D برابر شار خروجی از مرزهای ناحیه است. توجه کنید که $\vec{G} \cdot \vec{n}$ نمایانگر مؤلفه خروجی میدان \vec{G} از مرز ناحیه است. در واقع (۳) منجر به تعبیر هندسی - فیزیکی مورد نظر از مفهوم دیورژانس می شود. فرض کنید D_ρ یک گوی به شعاع ρ حول نقطه ای p در صفحه باشد. چون \vec{G} دارای مشتق های پاره ای پیوسته فرض شده است، $div \vec{G}$ پیوسته است و طبق قضیه میانگین انتگرال، $\iint_{D_\rho} div \vec{G} dA$ برابر حاصل ضرب مساحت D_ρ در $(div \vec{G})(Q)$ می باشد که Q نقطه ای مناسب در D_ρ است:

$$\iint_{D_\rho} (div \vec{G}) dA = (\pi \rho^2)(div \vec{G})(Q)$$

حال اگر ρ به صفر میل داده شود، از پیوستگی $div \vec{G}$ نتیجه می شود که $div \vec{G}(Q)$ به $div \vec{G}(P)$ میل می کند، پس:

$$(div \vec{G})(P) = \lim_{\rho \rightarrow 0} \frac{\iint_{D_\rho} \vec{G} \cdot \vec{n} ds}{\pi \rho^2} \quad (4)$$

توجه کنید که صورت کسر، شار خروجی از دایره ρ حول P است و مخرج مساحت درون دایره، بنابراین $div \vec{G}(P)$ معمولاً بدین صورت تعبیر می شود که نمایانگر میزان شار خروجی حول P در واحد مساحت است. مثبت بودن $div \vec{G}(P)$ دلالت بر این دارد که برای دایره های کوچک حول P ، $\vec{G} \cdot \vec{n}$ بیشتر مثبت است تا منفی، یعنی تمایل میدان بیشتر به دور شدن از P (انبساط حول P) است، و برعکس منفی بودن $div \vec{G}(P)$ ، نمایانگر تمایل میدان به انقباض به سوی P است. این تعبیر شهود مناسبی از کاربردهای دیورژانس در فیزیک می دهد.

(۴-۱) یک کاربرد هندسی فرض کنید D یک ناحیه بسته، کراندار و محدب در صفحه باشد که مرز آن، ∂D ، تصویر یک خم پیوسته ساده، بسته و قطعه قطعه هموار است. ناحیه محدب ناحیه ای است که برای هر دو نقطه P و Q از آن، پاره خط واصل بین P و Q به تمامی در ناحیه قرار گیرد. می خواهیم نوعی مفهوم "مرکز" و "شعاع" برای D منظور کنیم. برای هر نقطه p از D ، r_p به صورت

$$r_p = \max\{|q - p| \mid q \in \partial D\}$$

تعریف می شود، یعنی r_p حداکثر فاصله p از نقاط مرزی D است. از پیوستگی خم تعریف کننده مرز می توان نتیجه گرفت که این ماکسیمم به ازای نقطه ای از مرز اتخاذ می شود. همچنین از محدب بودن ناحیه می توان نشان داد که r_p به طور پیوسته با p تغییر می کند و بنابراین نقطه ای C در D یافت می شود که در آن r_C حداقل ممکن است:

$$r_C = \min\{r_p \mid p \in D\}$$

$r_C = r$ را شعاع D و C را یک مرکز D می نامیم. برای گوی بسته معمولی، مفهوم شعاع و مرکز همان مفهوم متداول است. حال اگر A مساحت D و L محیط ∂D باشد، ثابت می کنیم:

$$A \leq \frac{1}{4} r \cdot L \quad (5)$$

برای گوی مدور، در (5) تساوی برقرار است. برای اثبات (5)، میدان \vec{G} را به صورت زیر در نظر بگیرید:

$$\vec{G}(x, y) = (x - x_C, y - y_C)$$

که در آن $C = (x_C, y_C)$ مرکز است. داریم $\text{div} \vec{G} = 2$. از سوی دیگر، اگر (x, y) یک نقطه مرزی باشد:

$$\vec{G} \cdot \vec{n} = |\vec{G}| |\vec{n}| \cos \angle(\vec{G}, \vec{n}) \leq |\vec{G}|$$

ولی $|\vec{G}(x, y)|$ برابر فاصله (x, y) از C است، پس $|\vec{G}| \leq r$ ، یا

$$\vec{G} \cdot \vec{n} \leq r$$

بنابراین

$$2A = \int \int_D \text{div} \vec{G} = \int_{\partial D} (\vec{G} \cdot \vec{n}) ds \leq r \cdot L$$

و حکم به اثبات می رسد.

قضیه دیورژانس تعمیمی در همه ابعاد دارد. در زیر تعمیم آن به ناحیه های سه بعدی را تشریح می کنیم.

(۴۱-۲) قضیه دیورژانس M یک ناحیه سه بعدی مجاز انتگرال گیری در \mathbb{R}^3 است که به اجتماع متناهی رویه‌های قطعه قطعه هموار ∂M محصور می‌باشد و \vec{G} یک میدان برداری دارای مشتق‌های پاره‌ای پیوسته است که در سراسر M و ∂M تعریف شده است. در این صورت اگر \vec{n} قائم واحد برون‌گرای ∂M باشد، داریم:

$$\int \int \int_M \text{div} \vec{G} dV = \int \int_{\partial M} \vec{G} \cdot \vec{n} dS \quad (6)$$

ملاحظات زیر مشابه نکات متناظر در حالت دو بعدی هستند.

(۴۱-۳) تعبیر دیورژانس. M_ρ را گوی بسنه شعاع ρ حول نقطه p می‌گیریم. طبق قضیه میانگین انتگرال، نقطه‌ای Q در این گوی وجود دارد که

$$\int \int \int_M \text{div} \vec{G} dV = (M_\rho \text{ حجم}) \cdot (\text{div} \vec{G})(Q)$$

بنابراین

$$\text{div} \vec{G}(Q) = \frac{\int \int_{\partial M_\rho} \vec{G} \cdot \vec{n} dS}{M_\rho \text{ حجم}}$$

حال وقتی $\rho \rightarrow 0$ ، $\text{div} \vec{G}(Q)$ به $\vec{G}(p)$ میل می‌کند زیرا که مشتق‌های پاره‌ای G ، و در نتیجه $\text{div} \vec{G}$ پیوسته فرض شده‌اند. از این رو $\text{div} \vec{G}(p)$ به آهنگ انبساط میدان حول p نسبت به واحد حجم تعبیر می‌شود.

(۴۱-۴) همان‌گونه که برای ناحیه‌های محدب دو بعدی، "شعاع" و "مرکز" تعریف کرده‌ایم، برای ناحیه‌های محدب سه بعدی نیز مرکز و شعاع تعریف می‌شوند. در اینجا نیز با در نظر گرفتن میدان

$$\vec{G}(x, y, z) = (x - x_c, y - y_c, z - z_c) \quad \text{مرکز ناحیه} : (x_c, y_c, z_c)$$

می‌توان ثابت کرد که:

$$V \leq \frac{1}{3} r S$$

که V حجم ناحیه، S سطح جانبی و r شعاع است. برای گوی سه بعدی تساوی برقرار است.

اثبات قضیه دیورژانس اثبات قضیه دیورژانس سه بعدی نیز کاملاً مانند اثبات قضیه گرین دو بعدی است، نخست قضیه برای مستطیل‌های با اضلاع موازی محورها ثابت می‌شود، سپس برای ناحیه‌های قابل تجزیه به این گونه مستطیل‌ها، و بالاخره در حد برای ناحیه‌های مجاز ذکر شده در صورت قضیه. برای مستطیل $a_1 \leq x \leq a_2$ ، $b_1 \leq y \leq b_2$ ، $c_1 \leq z \leq c_2$ و $\vec{G} = G_1\vec{i} + G_2\vec{j} + G_3\vec{k}$ ، سه انتگرال $\int \int \int_M \frac{\partial G_1}{\partial x} dV$ ، $\int \int \int_M \frac{\partial G_2}{\partial y} dV$ و $\int \int \int_M \frac{\partial G_3}{\partial z} dV$ را به طور جداگانه محاسبه می‌کنیم. برای اولی، نخست نسبت به x انتگرال گیری می‌کنیم که بنابر قضیه اساسی حساب دیفرانسیل و انتگرال نتیجه می‌دهد:

$$\int_{c_1}^{c_2} \int_{b_1}^{b_2} [G_1(a_2, y, z) - G_1(a_1, y, z)] dy dz$$

روی وجه‌های $x = a_1$ و $x = a_2$ ، به ترتیب، قائم واحد برابر $-\vec{i}$ و \vec{i} است و در نتیجه:

$$x = a_1 \quad \text{روی} \quad -G_1(a_1, y, z) = \vec{G} \cdot \vec{n}$$

$$x = a_2 \quad \text{روی} \quad G_1(a_2, y, z) = \vec{G} \cdot \vec{n}$$

پس $\int \int \int_M \frac{\partial G_1}{\partial x} dV$ برابر $\int \int \vec{G} \cdot \vec{n} dS$ روی دو وجه $x = a_1$ و $x = a_2$ می‌شود. به همین ترتیب دو انتگرال سه‌گانه دیگر برابر انتگرال‌های $\vec{G} \cdot \vec{n}$ روی چهار وجه باقیمانده می‌شوند و حکم به اثبات می‌رسد. اکنون دو مثال در نظر می‌گیریم.

مثال ۱. ناحیه‌ای مسطح واقع بر یک صفحه در \mathbb{R}^3 است و P نقطه‌ای خارج این صفحه. از P به همه نقاط D وصل می‌کنیم و جسم حاصل را مخروط به رأس P و قاعده D می‌نامیم. نشان می‌دهیم:

$$\text{مساحت}(D) (\text{ارتفاع}) = \frac{1}{3} \text{حجم مخروط}$$

که در اینجا مقصود از ارتفاع فاصله قائم P از صفحه‌ای است که D در آن واقع است. میدان برداری

$$\vec{G}(x, y, z) = (x - x_P)\vec{i} + (y - y_P)\vec{j} + (z - z_P)\vec{k}$$

را در نظر می‌گیریم که در آن $P = (x_P, y_P, z_P)$. داریم $\text{div} \vec{G} = 3$. "سطح جانبی مخروط" اجتماع پاره‌خط‌هایی است که رأس P را به نقاط مرزی D وصل می‌کنند. روی این سطح \vec{n} بر این خطوط

واصل عمود است، پس $\vec{G} \cdot \vec{n} = 0$ و $\iint_D \vec{G} \cdot \vec{n} dS$ روی سطح جانبی صفر می‌شود. برای سطح قاعده، D داریم:

$$\vec{G} \cdot \vec{n} = |\vec{G}| |\vec{n}| \cos \angle(\vec{G}, \vec{n}) = |\vec{G}| \cos \angle(\vec{G}, \vec{n})$$

برای نقطه (x, y, z) در D ، $G(x, y, z)$ برابر بردار $(x - x_P, y - y_P, z - z_P)$ است که اگر در کسینوس زاویه بین \vec{G} و \vec{n} ضرب شود تصویر قائم \vec{G} بر امتداد عمود بر صفحه D را می‌دهد که همان ارتفاع است، پس

$$\iint_D \vec{G} \cdot \vec{n} dS = (\text{مساحت } D)(\text{ارتفاع})$$

و حکم از قضیه دیورژانس نتیجه می‌شود.

مثال ۲. میدان $\vec{G}(x, y, z) = A\vec{i} + B\vec{j} + C\vec{k}$ را در نظر می‌گیریم که در آن A, B, C اعداد حقیقی داده شده‌اند. T^+ را نیمه بالایی چنبره توصیف شده در مثال ۳، بخش ۳۵، می‌گیریم، یعنی:

$$T^+ : (x, y, z) = ((a + b \cos u) \cos v, (a + b \cos u) \sin v, a \sin u)$$

$$\vec{n}, \text{ قائم قراردادی که در اینجا درونگرا است } 0 \leq u \leq \pi, 0 \leq v \leq 2\pi$$

انتگرال روی رویه $\iint_{T^+} \vec{G} \cdot \vec{n} dS$ را به کمک قضیه دیورژانس محاسبه می‌کنیم. به صورت داده شده، T^+ هیچ ناحیه سه بعدی را محصور نمی‌کند و نمی‌توان از قضیه دیورژانس استفاده کرد. در این گونه مسایل معمولاً رویه یا رویه‌های دیگری را جستجو می‌کنیم که در مرز به رویه داده شده وصل شده و با هم یک ناحیه را محصور کنند. برای رویه دوم کاندیداهای متنوعی وجود دارد، انتخاب مناسب انتخابی است که برای آن محاسبه $\iint \vec{G} \cdot \vec{n} dS$ ساده باشد. در این مثال رویه حلقوی و مسطح:

$$R : z = 0, (a - b)^2 \leq x^2 + y^2 \leq (a + b)^2$$

در صفحه xy همراه با T^+ یک ناحیه سه بعدی M را محصور می‌کند. به علاوه چون $\text{div} \vec{G} = 0$ ، از قضیه دیورژانس داریم:

$$\iint_{T^+} \vec{G} \cdot \vec{n} dS + \iint_R \vec{G} \cdot \vec{n} dS = 0$$

بدین ترتیب برای R ، $\vec{n} = -\vec{k}$ و $\vec{G} \cdot \vec{n} = -C$ پس (مساحت R) $\int_{T^+} \vec{G} \cdot \vec{n} dS = (-C)(R)$ ، یعنی $-4\pi abc$.

(۴۱-۵) پتانسیل برداری در بررسی وجود پتانسیل برای میدانهای برداری دیدیم که اگر $\text{curl } \vec{F} = \vec{Q}$ و دامنه تعریف \vec{F} از شرط هندسی مناسبی (نداشتن حفره) برخوردار باشد، آنگاه \vec{F} دارای پتانسیل است، یعنی تابع f وجود دارد که $\vec{F} = \nabla f$.

اکنون با توجه به این اتحاد که $\text{div}(\text{curl } \vec{F}) = 0$ (به شرط C^2 بودن میدان \vec{F})، این سؤال مطرح می شود که اگر برای میدانی C^1 ، \vec{G} داشته باشیم $\text{div } \vec{G} = 0$ ، آیا میدان برداری \vec{F} وجود دارد که $\vec{G} = \text{curl } \vec{F}$ ؟

در صورت وجود \vec{F} را یک پتانسیل برداری برای \vec{F} می نامند.

در اینجا نیز با افزودن شرطی هندسی بر دامنه تعریف \vec{G} می توان به نتیجه مثبت رسید، ولی نخست مثال زیر را در نظر بگیرید که نشان می دهد بدون افزودن شرط مناسب، جواب ممکن است منفی باشد.

مثال میدان \vec{G} روی $\mathbb{R}^3 - \{(0, 0, 0)\}$ را به صورت زیر تعریف می کنیم:

$$\vec{G}(x, y, z) = \left(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3} \right) \quad (7)$$

که در آن $r = \sqrt{x^2 + y^2 + z^2}$ (فاصله از مبدا). میدان \vec{G} یک میدان C^1 در U است. محاسبه سر راست نشان می دهد که در سراسر U داریم $\text{div } \vec{G} = 0$ ، ولی نشان می دهیم \vec{G} فاقد پتانسیل است. فرض کنید میدان \vec{F} وجود دارد که $\vec{G} = \text{curl } \vec{F}$. در این صورت برای هر رویه بسته D در U باید طبق قضیه استوکس داشته باشیم:

$$\int_D \vec{G} \cdot \vec{n} dS = 0$$

D را کره واحد $x^2 + y^2 + z^2 = 1$ با قائم واحد برونگرایی $\vec{n} = (x, y, z)$ در نظر می گیریم. داریم

$$\vec{G} \cdot \vec{r} = \frac{1}{r} = 1$$

پس $\int_D \vec{G} \cdot \vec{r} dS = 4\pi \neq 0$ که نشان می دهد \vec{G} نمی تواند به شکل $\text{curl } \vec{F}$ باشد.

شرط هندسی برای هر رویه قطعه قطعه هموار بسته D در U ، ناحیه درون در U قرار می گیرد. به شباهت این شرط با شرط متناظر در مورد وجود پتانسیل اسکالر توجه کنید. در اینجا به جای خم بسته، رویه بسته (یک بعد بالاتر) را منظور کرده ایم. در واقع در اینجا طبق قضیه استوکس می توان نتیجه گرفت که انتگرال $\vec{G} \cdot \vec{r}$ روی هر رویه بسته واقع شده در U صفر خواهد شد و نکته اصلی اثبات وجود پتانسیل برداری این شرط است.

در پایان شایان ذکر است که اگر ناحیه U از شرط هندسی محدود کننده تری که در زیر خواهد آمد برخوردار باشد، به روشهای موجود می توانیم وجود پتانسیل را ثابت کنیم. ناحیه U را ستاره وار نسبت به نقطه P می نامیم در صورتی که پاره خط واصل از هر نقطه (x, y, z) ناحیه U به نقطه P به تمامی در ناحیه U قرار گیرد.

(۴۱-۶) (لم پوانکاره) فرض کنید U ناحیه ای ستاره وار نسبت به نقطه P در \mathbb{R}^3 باشد و \vec{F} و \vec{G} دو میدان برداری C^1 پیوسته در U . در این صورت:

(الف) اگر $\text{curl} \vec{F} = \vec{0}$ ، آنگاه \vec{F} پایسته است و تابع f زیر یک پتانسیل برای آن است:

$$f(x, y, z) = \int_0^1 (\vec{F} \cdot \frac{d\vec{r}}{dt}) dt$$

که $\vec{r}(t) = P + t((x, y, z) - P)$ ، $0 \leq t \leq 1$ ، پاره خط واصل از P به (x, y, z) است.

(ب) اگر $\text{div} \vec{G} = 0$ ، آنگاه \vec{G} دارای پتانسیل برداری است. میدان \vec{H} زیر یک چنین پتانسیل

برداری است:

$$\vec{H}(x, y, z) = \int_0^1 (\vec{G} \times \frac{d\vec{r}}{dt}) t dt$$

که در آن $\vec{r}(t) = P + t((x, y, z) - P)$ ، $0 \leq t \leq 1$ ، پاره خط واصل از P به (x, y, z) است. در

اینجا مقصود از انتگرال تابع برداری، برداری است که هر مؤلفه آن انتگرال مؤلفه متناظر است.

اثبات در هر دو مورد داریم

$$\frac{d\vec{r}}{dt} = (x - x_P, y - y_P, z - z_P)$$

نخست به اثبات (الف) می پردازیم. باید ثابت کنیم $\frac{\partial f}{\partial x} = F_1$ ، $\frac{\partial f}{\partial y} = F_2$ و $\frac{\partial f}{\partial z} = F_3$. یک مورد، مثلاً

$\frac{\partial f}{\partial x} = F_x$ را ثابت می‌کنیم، دو مورد دیگر مشابه است. داریم:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \left(\int_0^1 (\vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt}) dt \right)$$

چون $\vec{F} \cdot \frac{d\vec{r}}{dt}$ دارای مشتق پاره‌ای پیوسته نسبت به x است، می‌توان طبق قضیه‌ای مشتق‌گیری $\frac{\partial}{\partial x}$ را به داخل انتگرال منتقل کرد، پس

$$\begin{aligned} \frac{\partial f}{\partial x} &= \int_0^1 \left(\frac{\partial}{\partial x} (\vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt}) \right) dt \\ &= \int_0^1 \frac{\partial}{\partial x} [(x - x_P)F_x(\vec{r}(t)) + (y - y_P)F_y(\vec{r}(t)) + (z - z_P)F_z(\vec{r}(t))] dt \\ &= \int_0^1 [F_x(\vec{r}(t)) + (x - x_P)t \frac{\partial F_x}{\partial x}(\vec{r}(t)) + (y - y_P)t \frac{\partial F_y}{\partial x}(\vec{r}(t)) + (z - z_P)t \frac{\partial F_z}{\partial x}(\vec{r}(t))] dt \end{aligned}$$

حال از فرض $\text{curl } \vec{F} = 0$ نتیجه می‌شود که $\frac{\partial F_y}{\partial x} = \frac{\partial F_x}{\partial z}$ و $\frac{\partial F_z}{\partial x} = \frac{\partial F_x}{\partial y}$ پس

$$\begin{aligned} \frac{\partial f}{\partial x} &= \int_0^1 [F_x(\vec{r}(t)) + t(x - x_P) \frac{\partial F_x}{\partial x}(\vec{r}(t)) + t(y - y_P) \frac{\partial F_x}{\partial y}(\vec{r}(t)) + t(z - z_P) \frac{\partial F_x}{\partial z}(\vec{r}(t))] dt \\ &= \int_0^1 \frac{d}{dt} (tF_x(\vec{r}(t))) dt \\ &= (tF_x(\vec{r}(t))) \Big|_0^1 \\ &= F_x(x, y, z) \end{aligned}$$

و (الف) به اثبات می‌رسد.

اثبات (ب) مشابه است. اگر بنویسیم $\vec{G} = G_x \vec{i} + G_y \vec{j} + G_z \vec{k}$ و $H = H_x \vec{i} + H_y \vec{j} + H_z \vec{k}$ باید

نشان دهیم $G_x = \frac{\partial H_y}{\partial y} - \frac{\partial H_z}{\partial z}$ و مشابهاً برای دو مؤلفه دیگر. تساوی $G_x = \frac{\partial H_y}{\partial y} - \frac{\partial H_z}{\partial z}$ را نشان می‌دهیم، دو مورد دیگر مشابه است. داریم:

$$\begin{cases} H_y(x, y, z) = \int_0^1 [G_z(\vec{r}(t))(x - x_P) - G_x(\vec{r}(t))(z - z_P)] t dt \\ H_z(x, y, z) = \int_0^1 [G_x(\vec{r}(t))(y - y_P) - G_y(\vec{r}(t))(x - x_P)] t dt \end{cases}$$

$$\frac{\partial H_y}{\partial z} = \int_0^1 \left[\frac{\partial G_z}{\partial z}(\vec{r}(t)) \cdot t \cdot (x - x_P) - \frac{\partial G_x}{\partial z}(\vec{r}(t)) \cdot t \cdot (z - z_P) - G_x(\vec{r}(t)) \right] t dt$$

$$\frac{\partial H_z}{\partial y} = \int_0^1 [G_x(\vec{r}(t)) + \frac{\partial G_x}{\partial y}(\vec{r}(t)) \cdot t \cdot (y - y_P) - \frac{\partial G_y}{\partial y}(\vec{r}(t)) \cdot t \cdot (x - x_P)] t dt$$

پس با توجه به فرض $div \vec{G} = 0$ که نتیجه می دهد $\frac{\partial G_x}{\partial x} - \frac{\partial G_y}{\partial y} - \frac{\partial G_z}{\partial z} = 0$ داریم:

$$\begin{aligned}
 \frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial z} &= \int_0^1 \{ t G_x(\vec{r}(t)) + t^2 \left[\frac{\partial G_x}{\partial x}(\vec{r}(t)) \cdot (x - x_P) + \frac{\partial G_x}{\partial y} \cdot (y - y_P) + \frac{\partial G_x}{\partial z} \cdot (z - z_P) \right] \} dt \\
 &= \int_0^1 \frac{d}{dt} [t^2 G_x(\vec{r}(t))] dt \\
 &= (t^2 G_x(\vec{r}(t))) \Big|_0^1 \\
 &= G_x(x, y, z)
 \end{aligned}$$

و حکم به اثبات می رسد.

والسلام