دانشگاه گیلان
۱۳۵۳_1974

آنالیز برداری

مرضیه شمس یوسفی
استادیار دانشکده علوم ریاضی

چاپ اول

اداره چاپ و انتشارات دانشگاه گیلان

بسم الله الرحمن الرحیم

آنالیز برداری

تدوین (به شیوه گردآوری)

دکتر مرضیه شمس یوسفی

استادیار دانشکده علوم ریاضی دانشگاه گیلان

انتشارات دانشگاه گیلان

۱۳۹۶

دانشگاه گیلان
۱۳۵۳-1974

شابک: ۹ ـ ۱۵۶ ـ ۱۵۳ ـ ۶۰۰ ـ ۹۷۸

سرشناسه	:	شمس یوسفی، مرضیه، ۱۳۵۹-
عنوان و نام پدیدآور	:	آنالیز برداری/ تدوین (به شیوه گردآوری) مرضیه شمس یوسفی.
مشخصات نشر	:	رشت: دانشگاه گیلان، ۱۳۹۶.
مشخصات ظاهری	:	خ، ۱۵۷ص.
شابک	:	978-600-153-156-9
وضعیت فهرست‌نویسی	:	فیپا
یادداشت	:	کتابنامه: ص. [۱۵۴].
یادداشت	:	نمایه.
موضوع	:	بردارها -- راهنمای آموزشی (عالی)
موضوع	:	Vector Analysis -- Study and teaching (Higher)
موضوع	:	بردارها -- مسائل، تمرین‌ها و غیره (عالی)
موضوع	:	Vector Analysis -- Problems, exercises, etc. (Higher)
موضوع	:	حساب تانسورها
موضوع	:	Calculus of tensors
شناسه افزوده	:	دانشگاه گیلان
رده بندی کنگره	:	۱۳۹۶ ۸آ۸ش/QA۴۳۳
رده بندی دیویی	:	۵۱۵/۶۳
شماره کتابشناسی ملی	:	۴۷۹۱۶۴۹

اداره چاپ و انتشارات دانشگاه گیلان

نام کتاب	:	آنالیز برداری
گردآورنده	:	دکتر مرضیه شمس یوسفی
ویراستار علمی	:	دکتر ساناز لامعی جوان
ویراستار ادبی	:	فرشته گلچین‌راد
نوبت چاپ	:	اول، ۱۳۹۶
ناشر	:	انتشارات دانشگاه گیلان

فهرست مطالب

پیشگفتار مؤلف

با تصمیم جدید شورای عالی برنامه‌ریزی و تغییر و تحول در سر فصل‌ها و عناوین دروس رشته ریاضی و همچنین به‌دلیل کمبود مرجع مناسب برای درس آنالیز ۲ رشته ریاضی بر آن شدم تا به‌تدوین و نگارش کتابی که پاسخگوی نیاز دانشجویان این درس و درس نظریه اندازه و کاربردها باشد، بپردازم. عمدتاً برای تدریس این دروس در دورهٔ کارشناسی مجبور به انتخاب مطالب از کتاب‌های پیشرفته‌تر دوره‌های تحصیلات تکمیلی یا برخی کتب ترجمه نشده به‌صورت پراکنده بوده‌ام. این کتاب جمع‌آوری مناسب مطالبی از کتاب‌های مختلف در باب آنالیز چندمتغیره، هندسه دیفرانسیل و نظریه اندازه است به‌نحوی‌که نیاز دانشجویان را پوشش دهد. در این مسیر کوشش شده است با ارائه نمونه مثال‌ها و تمرین‌های متنوع بر عمق و تنوع مطالب افزوده گردد.

این کتاب مشتمل بر ۶ فصل است. در فصل اول به مطالبی از جبر خطی با دیدگاه توپولوژی می‌پردازیم. از آنجا که مفاهیم مربوط به ماتریس‌ها در فصل‌های بعدی بسیار مورد نیاز است، در این فصل فضای ماتریس‌ها و توپولوژی‌های حاکم بر فضای آن‌ها را مورد مطالعه قرار می‌دهیم. در فصل دوم که بنیادی‌ترین قسمت کتاب است به‌مفهوم مشتق در

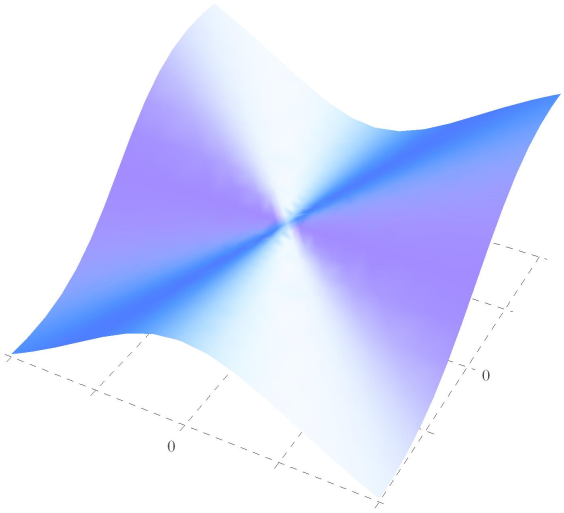
فضاهای با بعد بالاتر می‌پردازیم. فصل سوم شامل قضایای بنیادین ریاضیات چند متغیره است. نظریه انتگرال چند متغیره در فضاهای اقلیدسی در فصل چهارم مورد مطالعه قرار می‌گیرد. نظریه اندازه و فرم‌های دیفرانسیل به‌ترتیب فصل‌های پنجم و ششم را تشکیل می‌دهند.

در پایان از زحمات سرکار خانم ربابه احدزاده کارشناس ارشد ریاضی محض به دلیل همکاری صمیمانه در تایپ نسخهٔ دست‌نویس و آقای سیدکریم محمدی به دلیل ویراستاری و صفحه‌آرایی این کتاب سپاسگزاری می‌نمایم.

مرضیه شمس یوسفی
استادیار دانشکده علوم ریاضی دانشگاه گیلان

مقدمه

ریاضیات چند متغیره و نقش اساسی آن در فیزیک اعم از نیوتونی و کوانتومی ریاضی‌دانان را به جست و جو و تلاش در این شاخه ترغیب کرده است. بررسی خواص تحلیلی توابع چندمتغیره تا آنجا که به‌مباحثی چون تعیین دامنه، برد و حد و پیوستگی مربوط است از لحاظ نظری چندان دشوار نیست و ریاضیات مقدماتی توانایی پاسخ‌گویی در این زمینه را داراست. اما به محض برخورد به‌مفهوم مشتق و نیاز به تعمیم‌بخشی به‌این مفهوم اساسی، با یک مشکل جدی مواجه هستیم. شاید طبیعی‌ترین تعمیم برای مفهوم مشتق یک تابع چند متغیره حقیقی مقدار، استفاده از ابزار مشتق جزئی (مشتق پاره‌ای) آن باشد که عموماً در ریاضیات مقدماتی به آن پرداخته می‌شود. اما چنان که از ریاضیات یک متغیره می‌دانیم، وجود مشتق یک تابع در یک نقطه پیوستگی تابع در آن نقطه را نتیجه می‌دهد. اما اگر وجود مشتقات جزئی را تعمیم مشتق‌پذیری یک تابع در یک نقطه بدانیم، مثال ذیل نشان می‌دهد که این تعمیم مناسب نیست.



تابعی که مورد آزمایش قرار می‌گیرد تابع ذیل است

$$f(x,y)=\begin{cases}\dfrac{x^{۲}y}{x^{۲}+y^{۲}} & (x,y)\neq(۰,۰)\\ ۰ & (x,y)=(۰,۰).\end{cases}$$

مشتقات جزئی تابع f در مبدا صفر است. $\frac{\partial f}{\partial y}(۰,۰)=۰=\frac{\partial f}{\partial x}(۰,۰)$. بنابراین تنها احتمال برای وجود صفحه‌ی مماس بر نمودار تابع در مبدأ صفحه‌ی افقی است. اما باتوجه به این‌که در راستاهای گوناگون شیب نمودار تابع به صفر نمی‌گراید صفحه‌ی افقی نمی‌تواند مماس باشد در واقع هیچ صفحه‌ی مماس در صفر بر این نمودار تابع موجود نیست و لذا تابع در صفر مشتق‌پذیر نیست.

مثال از توابع غیر پیوسته و لذا مشتق‌ناپذیر که دارای مشتقات جزئی هستند هم موجود است. از این جهت باید به دنبال جای‌گزین مناسبی برای مفهوم مشتق بود همان‌گونه که در فصل مشتق‌پذیری این کتاب خواهید دید، دیدگاه جدیدی برای ارائه این تصمیم ارائه می‌شود که در آن دیدگاه عملگری و ماتریسی برای مشتق نقش اساسی ایفا می‌کند. در واقع

در دیدگاه ماتریسی، مشتق تابع $f : \mathbb{R} \longrightarrow \mathbb{R}$ با یکی دانستن همه عملگرهای خطی از $\mathbb{R}$ به $\mathbb{R}$ با ماتریس‌های یک در یک تعریف می‌شود. به عبارتی مشتق تابع f در نقطه‌ی a را نه به‌عنوان یک عدد حقیقی بلکه به‌عنوان یک ماتریس یک در یک و یا یک تابع (عملگر) خطی از $\mathbb{R}$ به $\mathbb{R}$ که h را به $f'(a)h$ می‌برد در نظر می‌گیریم.

فصل ۱

جبر خطی با رویکرد توپولوژی

۱.۱ عملگرهای خطی

در این بخش به مطالعه مختصری بر جبر خطی می‌پردازیم. جبر خطی از مهم‌ترین ابزارهای مطالعه ساختارهای جبری است که مجردسازی دستگاه‌های جبری و مطالعه را آسان می‌کند.

تعریف ۱.۱.۱. یک فضای برداری عبارت است از یک مجموعه مانند E که اعضای آن را بردار، یک میدان مانند $\mathbb{F}$ که اعضای آن را اسکالر، همراه با یک عمل که آن را جمع برداری می‌نامیم و عبارت است از نگاشتی مانند

$$\begin{aligned} +: \quad & E \times E \longrightarrow E \\ & (x, y) \longrightarrow x + y, \end{aligned}$$

و یک عمل که آن را ضرب اسکالر می‌نامیم و عبارت است از نگاشتی مانند

$$\cdot: \quad \mathbb{F} \times E \longrightarrow E$$

$$(c, x) \longrightarrow cx,$$

که در شرایط ذیل صادق باشند:

(الف) جمع جابجایی است؛ $x + y = y + x$ $(x, y \in E)$.
(ب) جمع شرکت‌پذیر است؛ $(x + y) + z = x + (y + z)$ $((x, y, z \in E))$.
(ج) بردار یکتا مانند ۰ در E موجود است به‌طوری‌که برای هر $x \in E$ داریم $x + ۰ = x$.
(د) برای هر بردار $x \in E$، برداری مانند $-x$ در E موجود است به‌طوری‌که $x+(-x) = ۰$.
(ه) برای هر $x \in E$، $۱x = x$.
(و) برای هر $x \in E$ و $c_۱$ و $c_۲$ متعلق به $\mathbb{F}$، $(c_۱c_۲)x = c_۱(c_۲x)$.
(ز) برای هر $c \in \mathbb{F}$ و $x, y \in E$، $c(x + y) = cx + cy$.
(ح) برای هر $x \in E$ و $c_۱$ و $c_۲$ متعلق به $\mathbb{F}$، $(c_۱ + c_۲)x = c_۱ + c_۲x$.

در این صورت فضای برداری را با $(E, \cdot, +)$ نمایش می‌دهیم و اگر جای ابهام نباشد به E اکتفا می‌کنیم.

به‌عنوان مثال اگر $\mathbb{F}$ یک میدان دلخواه و

$$E = \{(x_۱, \dots, x_n) \mid x_i \in \mathbb{F}\},$$

آن‌گاه E با اعمال طبیعی جمع و ضرب اسکالر

$$(x_۱, \dots, x_n) + (y_۱, \dots, y_n) = (x_۱ + y_۱, \dots, x_n + y_n)$$

و

$$\forall\, c \in \mathbb{F} \quad c(x_۱, \dots, x_n) = (cx_۱, \dots, cx_n),$$

یک فضای برداری است.

همچنین فضای ماتریس‌های $m \times n$ روی میدان $\mathbb{F}$ یک فضای برداری است. منظور از فضای ماتریس‌های $m \times n$ روی میدان $\mathbb{F}$ عبارت است از همهٔ ماتریس‌های $m \times n$ با درایه‌های از $\mathbb{F}$ که آن را معمولاً با $\mathbb{M}_{m\times n}(\mathbb{F})$ نمایش می‌دهیم. دراین صورت اعمال طبیعی جمع و ضرب اسکالر زیر به‌عنوان اعمال فضای برداری در نظر گرفته می‌شود:

$$[x_{ij}]_{m\times n} + [y_{ij}]_{m\times n} = [x_{ij} + y_{ij}]_{m\times n} ,$$

و

$$c[x_{ij}]_{m\times n} = [cx_{ij}]_{m\times n} .$$

تعریف ۲.۱.۱. فرض کنیم E یک فضای برداری روی میدان $\mathbb{F}$ و $S \subseteq E$ ناتهی و دلخواه باشد. گوییم بردار x از E یک ترکیب خطی از اعضای S است هرگاه بردارهایی چون $x_۱$، ... وx_n از S و اسکالرهایی چون $c_۱$، ... و c_n از $\mathbb{F}$ موجود باشند به‌طوری‌که $x = \sum_{i=۱}^{n} \alpha_i x_i$.

تعریف ۳.۱.۱. فرض کنیم $S \subset \mathbb{R}$ یک مجموعه ناتهی و دلخواه باشد. اگر E با مجموعهٔ تمام ترکیبات خطی اعضای S برابر باشد، گوییم که S مجموعهٔ E را **تولید می‌کند** و یا این‌که E **تولید** S است.

بدیهی است که هر تولید، یک فضای برداری است.

تعریف ۴.۱.۱. فرض کنیم E یک فضای برداری روی میدان $\mathbb{F}$ باشد. زیر مجموعهٔ ناتهی S از E را وابسته خطی می‌نامیم هرگاه بردارهای متمایز $x_۱$، ... و x_n از S و اسکالرهای $c_۱$، ... و c_n که همگی صفر نباشند در $\mathbb{F}$ یافت شوند، به‌طوری‌که

$$c_۱x_۱ + \cdots + c_nx_n = ۰ .$$

مجموعه‌ای که وابسته خطی نباشد، مستقل خطی نامیده می‌شود.

تعریف ۵.۱.۱. اگر فضای برداری X شامل یک مجموعهٔ مستقل خطی با r بردار باشد ولی هیچ مجموعهٔ مستقل خطی با $r+1$ بردار را شامل نباشد، گوییم که X دارای بعد r است و می‌نویسیم $\dim X = r$.

مجموعه‌ای که تنها از 0 تشکیل شده باشد، یک فضای برداری است؛ بعد آن به‌طور قراردادی صفر تعریف می‌شود.

قضیه ۶.۱.۱. فرض کنیم r عدد صحیح مثبتی است. اگر یک فضای برداری X به‌وسیلهٔ مجموعه‌ای از r بردار تولید شود، آن‌گاه $\dim X \leq r$.

برهان. اگر قضیه برقرار نباشد، یک فضای برداری مانند X وجود دارد که به‌وسیله مجموعهٔ $S_0 = \{x_1, \ldots, x_r\}$ از r بردار تولید می‌شود و در عین حال شامل مجموعهٔ مستقل خطی $Q = \{y_1, \ldots, y_{r+1}\}$ است. فرض کنیم $1 \leq i < r$ و فرض می‌کنیم مجموعه S_i را به‌قسمی داریم که X را تولید می‌کند و شامل تمام y_jهایی که $1 \leq j \leq i$ به‌علاوهٔ $r-i$ عضو از S_0، مثلاً $x_{r-i}, \ldots, x_1$ است(توجه می‌کنیم که حالت $i=0$ همان فرض خلف قضیه است(.به‌عبارت دیگر S_i با جای‌گزین کردن i عضو از S_0 با اعضایی از Q، بدون این‌که تولید آن تغییر کند حاصل می‌شود. چون S_i فضای X را تولید می‌کند پس y_{i+1} در تولید S_i قرار دارد. بنابراین اسکالرهایی مانند $b_1, \ldots, b_{r-i}$ و $a_1, \ldots, a_{i+1}$ با $a_{i+1} = 1$ وجود دارد، به‌قسمی که

$$\sum_{j=1}^{i+1} a_j y_j + \sum_{k=1}^{r-i} b_k x_k = 0\,.$$

اگر تمام b_k ها 0 باشند، استقلال Q ایجاب می‌کند که تمام a_jها نیز 0 باشند، که یک تناقض است. بنابراین $x_k \in S_i$ چنان موجود است که ترکیب خطی از سایر اعضای $T_i = S_i \cup \{y_{i+1}\}$ است. این x_k را از T_i خارج می‌کنیم و مجموعهٔ باقیمانده را S_{i+1} می‌نامیم. دراین‌صورت S_{i+1} داری تمام ویژگی‌های S_i است. با شروع از S_0 به‌این‌ترتیب

مجموعه‌های $S_1 \dots S_n$ و را می‌سازیم که آخرین آن‌ها یعنی S_r از $y_1 \dots y_r$ و تشکیل شده است و نحوه ساختن نشان می‌دهد که X را تولید می‌کند. ولی Q مستقل خطی است و بنابراین y_{r+1} در تولید S_r جای ندارد. این تناقض قضیه را ثابت می‌کند. $\square$

تعریف ۷.۱.۱. یک مجموعهٔ مستقل خطی از فضای برداری X که X را تولید کند، یک **پایهٔ** X خوانده می‌شود.

اگر $B = \{x_1, \dots, x_r\}$ پایه‌ای از X باشد، آن‌گاه چون B فضای X را تولید می‌کند، هر $x \in X$ را می‌توان به‌صورت $x = \sum_{i=1}^{r} c_i x_i$ نمایش داد. به‌علاوه این نمایش منحصربه‌فرد است زیرا اگر $x = \sum_{i=1}^{r} c_i x_i$ و $x = \sum_{i=1}^{r} c'_i x_i$، آن‌گاه استقلال خطی B ایجاب می‌کند که برای $i = 1, \dots, r$، $c_i = c'_i$. در نمایش $x = \sum_{i=1}^{r} c_i x_i$، اعداد $c_1 \dots c_r$ و را **مختصات** x نسبت به پایهٔ B می‌خوانیم.

قضیه ۸.۱.۱. فرض کنیم X یک فضای برداری است و $\dim X = n$ دراین‌صورت:

(الف) یک مجموعهٔ E از n بردار X فضای X را تولید می‌کند، اگر و فقط اگر E مستقل خطی باشد.

(ب) X یک پایه دارد و هر پایهٔ X از n بردار تشکیل شده است.

(پ) اگر $1 \le r \le n$، $\{y_1, \dots, y_r\}$ مجموعه‌ای مستقل خطی در X باشد، آن‌گاه X پایه‌ای دارد که شامل $\{y_1, \cdots, y_r\}$ است.

برهان. (الف) فرض کنیم $E = \{x_1, \dots, x_n\}$. چون $\dim X = n$، برای هر $y \in X$، مجموعه $\{x_1, \dots, x_n, y\}$ وابسته خطی است. اگر E مستقل خطی باشد پس y در تولید E جای دارد یعنی E فضای X را تولید می‌کند. برعکس اگر E وابسته خطی باشد می‌توان یکی از اعضای آن را بدون این‌که تولید آن تغییر کند، برداریم. پس بنا بر قضیه ۶.۱.۱، E نمی‌تواند X را تولید کند و (الف) ثابت می‌شود.

چون $\dim X = n$ پس مجموعه‌ای مستقل خطی از n بردار را شامل است و بنا بر (الف) یک چنین مجموعه‌ای پایه‌ای از X است. به‌علاوه اگر S یک پایه X باشد، بنا بر تعریف ۵.۱.۱ تعداد اعضای S کوچک‌تر از $n+1$ است، از طرفی بنا بر قضیه ۶.۱.۱ تعداد اعضای S بزرگ‌تر از $\dim X = n$ یا مساوی آن است. پس تعداد بردارهای S دقیقاً برابر n است و این (ب) را ثابت می‌کند.

برای اثبات (پ) فرض می‌کنیم $\{x_1, \ldots, x_n\}$ پایه‌ای برای X باشد.
مجموعه $S = \{y_1, \ldots, y_r, x_1, \ldots, x_n\}$، X را تولید می‌کند و نامستقل خطی است زیرا بیش از n عضو دارد. بنا بر بحثی که در برهان قضیه ۶.۱.۱ دیدیم، یکی از x_i ها ترکیب خطی از سایر اعضای S است. اگر این عمل را r مرتبه تکرار کنیم، به مجموعه‌ای با n عضو می‌رسیم که X را تولید می‌کند که بنا بر (الف) مستقل خطی است، پس یک پایه می‌باشد که شامل $\{y_1, \cdots, y_r\}$ است. □

۲.۱ ساختار فضای $\mathbb{R}^n$

یکی از مهم‌ترین فضاهای برداری که در این کتاب به تفصیل راجع به ساختار و هندسهٔ آن صحبت خواهد شد فضای $\mathbb{R}^n$ است. در بخش حاضر به ساختار جبری و توپولوژیکی فضای $\mathbb{R}^n$ می‌پردازیم.

مجموعهٔ $\{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \quad i = 1, \ldots, n\}$ را با $\mathbb{R}^n$ نشان می‌دهیم. با اعمال طبیعی جمع و ضرب اسکالر مؤلفه به مؤلفه، $\mathbb{R}^n$ یک فضای برداری است. به‌عبارتی برای هر $X = (x_1, \cdots, x_n)$ و $Y = (y_1, \cdots, y_n)$ متعلق به $\mathbb{R}^n$ تعریف می‌کنیم

$$X + Y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

و

$$\forall k \in \mathbb{R} \quad k(x_1, \ldots, x_n) = (kx_1, \ldots, kx_n) \ .$$

با اعمال جمع و ضرب اسکالر معرفی شده در بالا $\mathbb{R}^n$ یک فضای برداری است. به‌سادگی دیده می‌شود که بردارهای

$$e_i = (۰, \cdots, \overbrace{۱}^{\text{مکان}\, i\text{ـام}}, \ldots, ۰),$$

برای $i = ۱, \ldots, n$ یک پایه برای $\mathbb{R}^n$ تشکیل می‌دهد و لذا $\mathbb{R}^n$ یک فضای برداری با بعد n روی $\mathbb{R}$ است.

بیش از این $\mathbb{R}^n$ به ساختار دیگری نیز مجهز است که در ادامه آن را در حالت کلی تعریف و در حالت $\mathbb{R}^n$ ارائه می‌کنیم.

تعریف ۱.۲.۱. ضرب داخلی (ضرب عددی) روی فضای برداری مختلط V، نگاشتی چون

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}$$

است به‌گونه‌ای که به‌ازای هر x، y و z متعلق به V و هر λ متعلق به $\mathbb{C}$

(۱) $\langle x, y \rangle = \langle y, x \rangle$,

(۲) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,

(۳) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,

(۴) $\langle x, x \rangle \geq ۰$.

زوج $(V, \langle \cdot, \cdot \rangle)$ را فضای ضرب داخلی می‌نامیم.

تابع ضرب داخلی در $\mathbb{R}^n$ به‌صورت زیر تعریف می‌شود:

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$\langle x, y \rangle = \sum_{i=۱}^{n} x_i y_i \ .$$

که در آن $x = (x_1, \ldots, x_n)$ و $y = (y_1, \ldots, y_n)$. به‌این‌ترتیب $\mathbb{R}^n$ یک فضای ضرب داخلی است.

در ادامه ساختار توپولوژیک فضای $\mathbb{R}^n$ را مورد مطالعه قرار می‌دهیم. به‌این‌منظور ابتدا برخی تعریف‌های مهم و کلی را در این راستا ارائه می‌کنیم.

تعریف ۱.۲.۲. فرض کنیم E یک فضای برداری باشد. دراین صورت منظور از یک نرم روی E تابع

$$\|\cdot\| : E \longrightarrow [0, \infty)$$

است که در خواص ذیل صدق می‌کند

$$(1)\ \ \|x\| \geq 0\ , \quad x \in E,$$
$$(2)\ \ \|x\| = 0 \ \text{ اگر و تنها اگر } \ x = 0\ ,\ x \in E,$$
$$(3)\ \ \|\alpha x\| = |\alpha|\|x\|\ ,\ \alpha \in \mathbb{R},\ x \in E,$$
$$(4)\ \ \|x + y\| \leq \|x\| + \|y\|.$$

در این صورت $(E, \|\cdot\|)$ را یک فضای نرم‌دار می‌نامیم.

هر فضای نرم‌دار یک فضای متری است. کافی است قرار دهیم

$$\begin{aligned} d: \quad & E \times E \longrightarrow [0, \infty) \\ & (x, y) \longrightarrow \|x - y\|\ . \end{aligned}$$

بررسی این حقیقت ساده است و به خواننده واگذار می‌شود. لذا کلیه مفاهیم توپولوژیکی مانند مفاهیم بازبودن، بسته بودن، فشردگی و همگرایی در فضاهای نرم‌دار بامعنی است به نمونه‌هایی در ذیل اشاره می‌کنیم.

منظور از یک گوی باز به شعاع r و به مرکز $x \in E$ مجموعه زیر است:

$$B_r(x) = \{y \in E \ \mid \ \|x-y\| < r\}$$

و همچنین گوی بسته به شعاع r و به مرکز $x \in E$ به‌صورت

$$\overline{B_r(x)} = \{y \in E \ \mid \ \|x-y\| \leq r\}$$

تعریف می‌شود.

گوییم $\{x_n\} \subseteq E$ به نقطهٔ $x \in E$ همگراست هرگاه

$$\forall \varepsilon > 0 ,\ \exists N \in \mathbb{N} ,\ \forall n \geq N \ ,\ \|x_n - x\| < \varepsilon \, .$$

در این حالت می‌نویسیم $x_n \xrightarrow{\|\cdot\|} x$.

در واقع $x_n \xrightarrow{\|\cdot\|} x$، اگر و تنها اگر $\{\|x_n - x\|\}$ یک دنباله عددی همگرا به صفر باشد. به وضوح در این صورت $\|x_n\| \longrightarrow \|x\|$ و عکس این حکم برقرار نیست.

مثال ۳.۲.۱. فضای $\mathbb{R}^n$ به‌عنوان یکی از مهم‌ترین ساختارهای مورد بررسی این کتاب نیز مجهز به انواع نرم است که طی موارد ذیل به آن‌ها اشاره می‌کنیم:

(الف) ۲ـنرم یا نرم اقلیدسی. طبیعی‌ترین نرم روی فضای $\mathbb{R}^n$ است که با $\|\cdot\|_2$ نمایش می‌دهیم و به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} \|\cdot\|_2 : \quad & \mathbb{R}^n \longrightarrow [0, \infty) \\ (x_1, \ldots, x_n) & \longrightarrow \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} . \end{aligned}$$

نرم اقلیدسی روی فضای $\mathbb{R}^n$ ارتباط نزدیکی با ضرب داخلی این فضا دارد. در واقع داریم

$$\langle X, X \rangle = \|X\|_2^2 \, ,$$

و این ارتباط ویژگی خاصی به نرم اقلیدسی می‌دهد.

در قضیه ذیل به خواص از ۲ـ نرم اشاره می‌کنیم.

قضیه ۴.۲.۱. (نامساوی کوشی‌ـ شوارتس) فرض کنید $x = (x_1, \ldots, x_n)$ و $y = (y_1, \ldots, y_n)$ به $\mathbb{R}^n$ متعلق باشند. دراین صورت

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \|x\|_2 \cdot \|y\|_2 \ .$$

برهان. اگر x, y مضربی از هم باشند، تساوی بدیهی است. در غیر این صورت برای هر $\lambda \in \mathbb{R}$، $\lambda y - x \neq 0$ و بنابراین

$$\begin{aligned} 0 < \|\lambda x - y\|^2 &= \sum_{i=1}^{n} (\lambda y_i - x_i)^2 \\ &= \lambda^2 \sum_{i=1}^{n} (y_i)^2 - 2\lambda \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} (x_i)^2 \end{aligned}$$

بنابراین معادلهٔ درجه دو سمت راست بدون جواب است و باید دلتای منفی داشته باشد. بنابراین

$$4(\sum_{i=1}^{n} x_i y_i)^2 - 4\sum_{i=1}^{n} (x_i)^2 \sum_{i-1}^{n} (y_i)^2 < 0,$$

و این همان حکم مطلوب است. □

(ب) ۱ـنرم. نرم دیگری می‌توان روی $\mathbb{R}^n$ تعریف کرد که آن را با $\|\cdot\|_1$ نمایش می‌دهیم و به‌صورت زیر تعریف می‌کنیم.

$$\begin{aligned} &\|\cdot\|_1 : \mathbb{R}^n \longrightarrow [0, \infty) \\ &(x_1, \ldots, x_n) \longrightarrow \textstyle\sum_{i=1}^{n} |x_i|. \end{aligned}$$

(ج) **∞ـنرم**. نرمی که به‌صورت

$$\begin{aligned}\|\cdot\|_{\infty} &: \mathbb{R}^n \longrightarrow [0,\infty)\\ (x_1,...,x_n) &\longrightarrow \max\{|x_1|,\ldots,|x_n|\},\end{aligned}$$

تعریف می‌شود را بی‌نهایت‌ـ نرم یا ∞ـ نرم روی $\mathbb{R}^n$ می‌نامیم.

(د) **pـنرم**. تابع نرم با تعریف

$$\begin{aligned}\|\cdot\|_{p} &: \mathbb{R}^n \longrightarrow [0,\infty)\\ (x_1,...,x_n) &\longrightarrow \left(\textstyle\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}},\end{aligned}$$

را یک pـ نرم روی $\mathbb{R}^n$ می‌نامیم که در آن $1 < p < \infty$.

بررسی نرم بودن نگاشت‌های بالا در موارد ۲ـ نرم و ۱ـ نرم و ∞ـ نرم ساده است. برای بررسی نرم‌بودن pـ نرم می‌توان به کتاب‌های پیشرفته‌تر آنالیز حقیقی مراجعه کرد. در اغلب کتاب‌های پیشرفته‌تر $\left(\mathbb{R}^n, \|\cdot\|_p\right)$ برای $1 \leq p \leq \infty$ را با $\ell_p^n(\mathbb{R})$ نمایش می‌دهند.

به‌سادگی دیده می‌شود که رابطه‌های زیر برقرار است

$$\begin{aligned}\tfrac{1}{\sqrt{n}}\|\cdot\|_1 \leq \|\cdot\|_2 &\leq \|\cdot\|_1 \ ,\\ \|\cdot\|_{\infty} \leq \|\cdot\|_2 &\leq \sqrt{n}\|\cdot\|_{\infty} \ ,\\ \|\cdot\|_{\infty} \leq \|\cdot\|_1 &\leq n\|\cdot\|_{\infty} \ .\end{aligned}$$

روابط بالا خواص ویژه‌ای بر $\mathbb{R}^n$ تحمیل می‌کند و در ذیل بیشتر به آن می‌پردازیم.

تعریف ۵.۲.۱. فرض کنیم $(E, \|\cdot\|)$ و $(E, \|\cdot\|')$ دو فضای نرم‌دار باشند در این صورت گوییم $\|\cdot\|$ و $\|\cdot\|'$ روی E با هم معادل هستند، هرگاه

$$\exists K, M > 0 \quad , \quad K\|\cdot\|' \leq \|\cdot\| \leq M\|\cdot\|'.$$

به وضوح رابطه معادل بودن نرم‌ها این حکم را ایجاب می‌کند که کلیه خواص توپولوژیکی مانند باز بودن، بسته بودن، همگرایی و ... در دو فضای نرم‌دار با هم معادل باشند. زیرا اگر گوی واحد به مرکز صفر فضای $(E, \|\cdot\|)$ را با $B_۱(۰)$ و گوی واحد به مرکز صفر فضای $(E, \|\cdot\|')$ را با $B'_۱(۰)$ نمایش دهیم آن‌گاه،

$$\cdots \subseteq B_۱(۰) = \{x \in E \mid \|x\| \leq ۱\} \subseteq B'_{\frac{۱}{K}}(۰) = \left\{x \in E \mid \|x\|' < \frac{۱}{K}\right\}$$
$$\subseteq \left\{x \in E \mid \|x\| < \frac{M}{K}\right\} = B_{\frac{M}{K}}(۰) \subseteq \cdots$$

لذا تمامی گوی‌ها در دو فضای نرم‌دار تو در تو بوده و در نتیجه موجب می‌شود توپولوژی حاصل از $\|\cdot\|$ و $\|\cdot\|'$ بر هم منطبق باشند. لذا تمامی خواص توپولوژیکی یکسان باقی می‌ماند.

فضای برداری دیگری که می‌تواند مورد بحث قرار گیرد فضای $\mathbb{C}^n$ است که عبارت است از

$$\mathbb{C}^n = \{(z_۱, \ldots, z_n) \mid z_i \in \mathbb{C}, i = ۱, \ldots, n\}.$$

کلیه نرم‌هایی که روی $\mathbb{R}^n$ در بالا معرفی شد، $\mathbb{C}^n$ را نیز به یک فضای نرم‌دار تبدیل می‌کنند.

در اغلب کتاب‌های پیشرفته برای نمایش فضای نرم‌دار $(\mathbb{C}^n, \|\cdot\|_p)$ برای $۱ \leq p \leq \infty$ از نماد ℓ_p^n استفاده می‌شود.

به‌عنوان یک تمرین ساده می‌توان نشان داد که اگر $\{x_n\}_{n=۱}^{\infty} \subseteq \mathbb{C}^k$ یک دنباله همگرا به x در ۲ـ نرم (نرم معرفی شده بالا) باشد آن‌گاه برای هر $i = ۱, \ldots, k$ دنباله $x_n^i \longrightarrow x^i$ که در آن $x = (x^۱, \ldots, x^k)$ و $x_n = (x_n^۱, \ldots, x_n^k)$ و برعکس.

مثال ۶.۲.۱. فضای $\mathbb{M}_{n\times m}$

مجموعهٔ همهٔ ماتریس‌های $n\times m$ـ بعدی روی $\mathbb{R}$ $(\mathbb{C})$ را با $\mathbb{M}_{n\times m}(\mathbb{R})$ $(\mathbb{M}_{n\times m}(\mathbb{C}))$ نمایش می‌دهیم و اگر جای ابهام نباشد، از نوشتن $\mathbb{R}$ یا $\mathbb{C}$ صرفنظر می‌کنیم. از آنجا که

$\mathbb{M}_{n\times m}(\mathbb{R})$ از لحاظ فضای برداری با $\mathbb{R}^{nm}$ یکریخت است لذا کلیه نرم‌هایی که در بالا برای $\mathbb{R}^k$ معرفی شد را می‌توان برای فضای ماتریس‌ها نیز در نظر گرفت. هر چند این نرم مطلوب ما نیست و در ادامه با دیدی متفاوت به فضای ماتریس‌ها خواهیم نگریست و نرم مناسب‌تری جای‌گزین این نرم خواهیم کرد و در سرتاسر این کتاب با این نرم کار خواهیم کرد.

قضیه ذیل یکی از قضایای معروف آنالیز تابعی است.

قضیه ۱.۲.۷. اگر E یک فضای برداری متناهی بعد باشد، آن‌گاه کلیه نرم‌های روی E با هم معادل هستند.

برهان. بدون خلل به کلیت فرض می کنیم، فضای متناهی بعد مفروض $\mathbb{R}^n$ است. فرض کنیم $\|\cdot\|_۲$ نمایش دهنده نرم اقلیدسی و $\|\cdot\|$ نرم دلخواه دیگری باشد. نشان می دهیم این دو نرم با هم معادل هستند و این برای اثبات حکم کفایت می‌کند.

اگر $(x_۱,\ldots,x_n) = x_۱e_۱ + \cdots + x_ne_n$، که در آن e_i نمایش دهنده پایه یکه استاندارد فضای $\mathbb{R}^n$، آن‌گاه با توجه به نامساوی مثلث داریم :

$$\|x\| = \left\|\sum_{i=۱}^{n} x_ie_i\right\| \le \sum_{i=۱}^{n} |x_i|\cdot\|e_i\| \le \left(\sum_{i=۱}^{n}\|e_i\|\right)\cdot\|x\|_۲.$$

بنابراین با قرار دادن $M = \sum_{i=۱}^{n}\|e_i\|$ برای هر $x\in\mathbb{R}^n$ داریم ،

$$\|x\| \le M\|x\|_۲.$$

این رابطه نیمی از آن‌چه می‌خواستیم، ثابت می‌کند. روابط زیر برقرار است

$$|\|x\| - \|y\|| \le \|x-y\| \le M\|x-y\|_۲.$$

این روابط نشان می‌دهد که نگاشت $x \to \|x\|$ از $\mathbb{R}^n$ با نرم اقلیدسی به اعداد حقیقی پیوسته است.

حال قرار می‌دهیم $S = \{x \in \mathbb{R}^n | \|x\|_۲ = ۱\}$ که کره واحد در $\mathbb{R}^n$ نامیده می‌شود. بنابر قضیه هاینه بورل این مجموعه با نرم اقلیدسی فشرده است. بنابراین نگاشت $x \to \|x\|$ روی این مجموعه مینیمم خود را مثلا در نقطه $x_۰$ اخذ می‌کند. بنابراین برای هر $x \in S$ داریم $\|x\| \geq \|x_۰\|$. قرار می‌دهیم $K = \|x_۰\|$. از انجا که $\|x_۰\|_۲ = ۱$، داریم $K > ۰$. و برای هر $x \in \mathbb{R}^n$ داریم،

$$K\|x\|_۲ \leq \|x\|$$

و حکم ثابت می‌شود. $\square$

از قضیه بالا برمی‌آید که روی فضاهایی که در مثال‌های قبل معرفی شد همهٔ نرم‌ها با هم معادل هستند و لذا بررسی مفاهیم توپولوژیکی روی این فضاها با هر کدام از نرم‌ها میسر است و خواص به دست آمده به سایرین گسترش می‌یابد.

در سرتاسر این کتاب $\mathbb{R}^n$ و $\mathbb{C}^n$ را با نرم اقلیدسی $\|\cdot\|_۲$ در نظر می‌گیریم و برای راحتی و خلاصه‌نویسی از نوشتن اندیس ۲ صرفنظر می‌کنیم.

مشابه مطالعه کلیه رسته‌های ریاضیات، بلافاصله پس از شناخت اشیاء مورد نظر بحث در این کتاب نیازمند شناخت نگاشت‌های مناسب بین آن‌ها نیز هستیم.

تعریف ۸.۲.۱. فرض کنیم X و Y یک فضای برداری روی میدان $\mathbb{F}$ باشند. در این صورت نگاشت T از فضای برداری X به توی فضای برداری Y را یک **تبدیل خطی** خوانیم درصورتی‌که برای هر $x_۱, x_۲$ و x از X و هر اسکالر c داشته باشیم

$$T(x_۱ + x_۲) = T(x_۱) + T(x_۲)\ , \qquad T(cx) = cTx\ .$$

یک تبدیل خطی از X به‌توی خودش را یک **عملگر خطی** روی X می خوانیم.

بدیهی است که اگر T خطی باشد، آن‌گاه $T(0)=0$. به‌علاوه اگر $\{x_1, \cdots, x_n\}$ پایه‌ای از X باشد آن‌گاه هر $x \in X$ نمایشی یکتا به‌صورت $x = \sum_{i=1}^{n} c_i x_i$ دارد و باتوجه به خطی بودن T داریم $T(x) = \sum_{i=1}^{n} c_i T(x_i)$ یعنی یک تابع خطی از X به Y کاملاً به‌وسیله مقادیرش در اعضای پایه مشخص می‌شود.

عملگر خطی T روی X را معکوس‌پذیر خوانیم در صورتی‌که عملگری خطی مانند T^{-1} روی X وجود داشته باشد، به‌طوری‌که برای هر $x \in X$، $T^{-1}T(x) = x = TT^{-1}(x)$. به‌راحتی می‌توان نشان داد که T معکوس‌پذیر است، اگر و فقط اگر یک‌به‌یک و بروی باشد. ولی در مورد فضا با بعد متناهی، قضیه زیر نشان می‌دهد هر یک از دو شرط یک‌به‌یک یا بروی بودن برای معکوس‌پذیری T کافی است.

قضیه ۹.۲.۱. هر عملگر خطی روی یک فضای برداری با بعد متناهی یک‌به‌یک است، اگر و فقط اگر بروی باشد.

برهان. فرض کنیم X یک فضای با بعد متناهی و T یک عملگر خطی روی X باشد. در این صورت اگر $\{x_1, \ldots, x_n\}$ یک پایه از X باشد، آن‌گاه خطی بودن T ایجاب می‌کند که

$$Q = \{T(x_1), \cdots, T(x_n)\}$$

حوزه مقادیر T را تولید می‌کند که با $R(T)$ نشان می‌دهیم . بنا بر قضیه ۸.۱.۱ (الف) $R(T) = X$ اگر و فقط اگر Q مستقل خطی باشد. بنابراین بایستی ثابت کنیم که شرط لازم و کافی برای استقلال خطی Q آن است که T یک‌به‌یک باشد. فرض کنیم T یک‌به‌یک است و $\sum c_i T(x_i) = 0$. دراین‌صورت $T(\sum c_i x_i) = 0$ و بنابراین $\sum c_i x_i = 0$. پس $c_1 = \cdots = c_n = 0$ و نتیجه می‌گیریم Q مستقل خطی است. برعکس فرض کنیم Q مستقل خطی است. فرض کنیم برای $x = \sum_{i=1}^{n} c_i x_i$ داشته باشیم $T\left(\sum_{i=1}^{n} c_i x_i\right) = 0$.

در این صورت $\sum_{i=1}^{n} c_i T(x_i) = 0$ و در نتیجه $c_1 = \cdots = c_n = 0$ پس $T(x) = 0$ ایجاب می‌کند که $x = 0$ یعنی T یک‌به‌یک است. □

ملاحظه ۱۰.۲.۱. مرسوم است که برای نمایش عملگر خطی بر اعضا از گذاشتن پرانتز صرف‌نظر می‌کنیم. دلیل این امر به زودی روشن خواهد شد.

فرض کنیم X و Y فضاهای برداری روی میدان $\mathbb{F}$ باشند. مجموعه تمام تبدیلات خطی از فضای برداری X به توی فضای برداری Y را به $L(X,Y)$ نشان می‌دهیم و معمولاً به جای $L(X,X)$ می‌نویسیم $L(X)$. اگر T_1, T_2 به $L(X,Y)$ متعلق باشند و c_1 و c_2 دو اسکالر در $\mathbb{F}$ باشند، $c_1T_1 + c_2T_2 : X \longrightarrow Y$ را با ضابطهٔ

$$(c_1T_1 + c_2T_2)x = c_1T_1x + c_2T_2x \qquad (x \in X)$$

تعریف می‌کنیم. به‌سادگی می‌توان ملاحظه کرد که $c_1T_1 + c_2T_2$ نیز یک تبدیل خطی از X به Y است. اگر X، Y و Z فضاهای برداری باشند و $T \in L(X,Y)$ و $S \in L(Y,Z)$، حاصلضرب $ST : X \longrightarrow Z$ را با ضابطهٔ

$$(ST)x = S(Tx)\,,$$

یعنی به‌شکل ترکیب T و S، تعریف می‌کنیم. بدیهی است که $ST \in L(X,Z)$.

تعریف ۱۱.۲.۱. برای هر $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ نرم T را به‌صورت

$$\|T\| = \sup\big\{\|Tx\| \quad | \quad \|x\| \leq 1\big\},$$

تعریف می‌کنیم. با توجه به این تعریف برای هر $0 \neq x \in \mathbb{R}^n$، اگر بردار $z = \dfrac{x}{\|x\|}$ را در نظر بگیریم داریم $\|z\| = 1$ پس $\|Tz\| \leq \|T\|$. بنابراین

$$\frac{1}{\|x\|}\|Tx\| = \left\|T\frac{x}{\|x\|}\right\| \leq \|T\|\,.$$

در نتیجه برای هر $x \in \mathbb{R}^n$ داریم

$$\|Tx\| \leq \|T\|\|x\|$$

به‌علاوه اگر عددی مانند λ وجود داشته باشد که برای هر $x \in \mathbb{R}^n$، $\|Tx\| \leq \lambda\|x\|$ آن‌گاه $\|T\| \leq \lambda$.

قضیه ۱.۲.۱۲. اگر $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ آن‌گاه $\|T\| \leq \infty$ و T نگاشتی به‌طور یکنواخت پیوسته از $\mathbb{R}^n$ به توی $\mathbb{R}^m$ است.

برهان. فرض کنیم $\{e_1, \ldots, e_n\}$ پایه استاندارد در $\mathbb{R}^n$ باشد. فرض کنیم $x = \sum c_i e_i$ و $\|x\| \leq 1$. بنابراین برای هر $i = 1, \ldots, n$، $|c_i| \leq 1$. پس

$$\|Tx\| = \left\|\sum_{i=1}^{n} c_i T e_i\right\| \leq \sum_{i=1}^{n} |c_i|\|Te_i\| \leq \sum_{i=1}^{n} \|Te_i\| \,.$$

در نتیجه

$$\|T\| \leq \sum_{i=1}^{n} \|Te_i\| < \infty$$

پیوستگی یکنواخت f از رابطهٔ

$$\|Tx - Ty\| \leq \|T\|\|x - y\| \qquad x, y \in \mathbb{R}^n$$

نتیجه می‌شود. □

قضیه ۱.۲.۱۳. (الف) اگر $S, T \in L(\mathbb{R}^n, \mathbb{R}^m)$ و c یک اسکالر باشد آن‌گاه

$$\|S + T\| \leq \|S\| + \|T\| \,, \qquad \|cT\| = |c|\|T\|$$

و با $d(T,S)=\|T-S\|$ مجموعهٔ $L(\mathbb{R}^n,\mathbb{R}^m)$ یک فضای متری است.

(ب) اگر $T\in L(\mathbb{R}^n,\mathbb{R}^m)$ و $S\in L(\mathbb{R}^m,\mathbb{R}^n)$ آن‌گاه

$$\|ST\|\leq \|S\|\|T\|\ .$$

برهان. **(الف)** برای هر $x\in X$

$$\|(T+S)x\|=\|Tx+Sx\|\leq \|Tx\|+\|Sx\|\leq (\|T\|+\|S\|)\|x\|\ .$$

پس $\|T+S\|\leq\|T\|+\|S\|$. رابطه دوم (الف) نیز از نامساوی‌های

$$\|cTx\|=|c|\|Tx\|\leq |c|\|T\|\|x\|,$$

حاصل می‌شود. اثبات نامساوی در جهت عکس مشابه است. اثبات متر بودن d نیز آسان است. نامساوی مثلثی با استفاده از (الف) به دلیل زیر برقرار است:

$$\|T-U\|=\|(T-S)+(S-U)\|\leq \|T-S\|+\|S-U\|\ .$$

بالاخره (ب) از

$$\|(ST)x\|=\|S(Tx)\|\leq \|S\|\|Tx\|\leq \|S\|\|T\|\|x\|,$$

حاصل می‌شود. $\square$

حال که دیدیم $L(\mathbb{R}^n,\mathbb{R}^m)$ یک فضای متری است، کلیه مفاهیم توپولوژیکی در این مجموعه دارای معنی است.

قضیه ۱۴.۲.۱. فرض کنیم $GL(\mathbb{R}^n)$ مجموعهٔ تمام عملگرهای معکوس‌پذیر روی $\mathbb{R}^n$ باشد.

(الف) اگر $T \in GL(\mathbb{R}^n)$، $S \in L(\mathbb{R}^n)$ و $\|S-T\| \cdot \|T^{-1}\| < 1$، آن‌گاه S معکوس‌پذیر است.

(ب) $GL(\mathbb{R}^n)$ در $L(\mathbb{R}^n)$ باز است و نگاشت $T \mapsto T^{-1}$ روی $GL(\mathbb{R}^n)$ پیوسته است. (معکوس این نگاشت به‌وضوح با خود آن برابر است پس این نگاشت یک‌به‌یک نیز می‌باشد).

برهان. **(الف)** قرار می‌دهیم $\|T^{-1}\| = \frac{1}{\alpha}$ و $\|S - T\| = \beta$. در این صورت بنا بر فرض $\beta < \alpha$. برای هر $x \in \mathbb{R}^n$ داریم

$$\begin{aligned} \alpha\|x\| = \alpha\left\|T^{-1}Tx\right\| &\leq \alpha\left\|T^{-1}\right\| \cdot \|Tx\| = \|Tx\| \\ &\leq \|(T-S)x\| + \|Sx\| \leq \beta\|x\| + \|Sx\| \, . \end{aligned}$$

بنابراین

$$(\alpha - \beta)\|x\| \leq \|Sx\| \qquad (x \in \mathbb{R}^n) \tag{۱.۱}$$

چون با توجه به فرض $\alpha - \beta > 0$، این رابطه نشان می‌دهد که اگر $x \neq 0$ آن‌گاه $Sx \neq 0$. یعنی S یک‌به‌یک است و بنا بر قضیه ۹.۲.۱، $S \in GL(\mathbb{R}^n)$.

(ب) در (الف) دیدیم که اگر $T \in GL(\mathbb{R}^n)$ آن‌گاه برای هر S که $\|S - T\| < \alpha$، داریم $S \in GL(\mathbb{R}^n)$، پس $GL(\mathbb{R}^n)$ باز است. برای قسمت دوم حکم، در رابطهٔ (۱.۱)، x را با $S^{-1}y$ جای‌گزین می‌کنیم و در این صورت

$$(\alpha - \beta)\|S^{-1}y\| \leq \|SS^{-1}y\| = \|y\|$$

که در نتیجه $\|S^{-1}\| \leq \frac{1}{\alpha - \beta}$. حال بنا بر قضیه ۱۲.۲.۱ (ب)

$$\|S^{-1} - T^{-1}\| = \|(S^{-1}(T-S)T^{-1}\| \leq \|S^{-1}\|\|T-S\|\|T^{-1}\| \leq \frac{\beta}{\alpha(\alpha - \beta)} \, .$$

نامساوی بالا پیوستگی نگاشت $T \mapsto T^{-1}$ را نتیجه می‌دهد. □

فرض کنیم $\{x_۱,\dots,x_n\}$ و $\{y_۱,\dots,y_m\}$ به‌ترتیب پایه‌هایی از فضای برداری X و Y باشند. در این صورت هر $T\in L(X,Y)$ مجموعه‌ای از اعداد a_{ij} را معین می‌کند به‌طوری‌که

$$Tx_j=\sum_{i=۱}^{m}a_{ij}y_i\,. \qquad (۱\le j\le n) \tag{۲.۱}$$

این اعداد را معمولاً در یک نمایش مستطیلی با m سطر و n ستون قرار می‌دهیم و یک ماتریس m در n می‌نامیم:

$$[T]=\begin{bmatrix} a_{۱۱} & a_{۱۲} & \dots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \dots & a_{۲n} \\ \dots & \dots & \ddots & \dots \\ a_{m۱} & a_{m۲} & \dots & a_{mn} \end{bmatrix}.$$

با توجه به این‌که مؤلفه‌های بردار Tx_j نسبت به پایهٔ $\{y_۱,\dots,y_m\}$ در ستون j ام T ظاهر می‌شوند، گاهی اوقات Tx_j یک بردار ستونی T خوانده می‌شود. با این نام‌گذاری حوزهٔ مقادیر T به‌وسیلهٔ بردارهای ستونی تولید می‌شوند. فرض کنیم $x=\sum_{j=۱}^{n}c_jx_j$. از خطی بودن T و رابطهٔ (۲.۱) نتیجه می‌گردد که

$$Tx=\sum_{i=۱}^{m}\left(\sum_{j=۱}^{n}a_{ij}c_j\right)y_i\,. \tag{۳.۱}$$

بنابراین مختصات Tx برابر است با بردار $\left(\sum_{j=۱}^{n}a_{۱j}c_j,\dots,\sum_{j=۱}^{n}a_{mj}c_j\right)$.

همان‌گونه که ملاحظه شد، هر تبدیل خطی از فضای n بعدی X به فضای m بعدی Y یک ماتریس $m\times n$ به دست می‌دهد. برعکس اگر یک ماتریس $m\times n$ با درایه‌های حقیقی $\{a_{ij}\}_{i,j=۱}^{n,m}$ داشته باشیم و T توسط (۳.۱) تعریف شود، بدیهی است که T تابعی خطی از X به Y است و T برابر ماتریس داده شده است. بنابراین تناظری یک‌به‌یک بین $L(X,Y)$ و مجموعهٔ تمام ماتریس‌های $m\times n$ وجود دارد و این امر دلیل ملاحظه (۱۰.۲.۱) را روشن

می‌سازد. بایستی تأکید کرد که این ماتریس نه تنها به T بلکه به انتخاب پایه‌های X و Y نیز وابسته است و درصورتی‌که پایه‌ها را تغییر دهیم یک تبدیل خطی ممکن است ماتریس‌های مختلفی را به دست دهد. این امر در کار ما زیاد دخالت نخواهد داشت زیرا ما معمولاً با پایه ثابتی کار می‌کنیم. اگر Z فضای برداری دیگری با پایهٔ $\{z_1, \dots, z_p\}$ باشد و T به‌وسیله (۲.۱) معین شود و اگر

$$Sy_i = \sum_{k=1}^{p} b_{ki} z_k \qquad \text{و} \qquad (ST)x_j = \sum_{k=1}^{p} c_{kj} z_k \ ,$$

آن‌گاه $T \in L(X,Y)$، $S \in L(Y,Z)$، $ST \in L(X,Z)$، و چون

$$S(Tx_j) = S\sum_i a_{ij} y_j = \sum_i a_{ij} S y_j = \sum_i a_{ij} \sum_k b_{ki} z_k = \sum_k \left(\sum_i b_{ki} a_{ij} \right) z_k$$

از استقلال خطی $\{z_1, \dots, z_p\}$ نتیجه می‌گردد که

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij} \ . \qquad (1 \le k \le p \ , \quad 1 \le j \le n)$$

این رابطه نشان می‌دهد که چگونه می‌توان ماتریس $p \times n$، مانند $[ST]$ را به کمک $[T]$ و $[S]$ محاسبه کرد که همان قاعدهٔ معمول در ضرب ماتریس‌ها می‌باشد. اگر $\{x_1, \dots, x_n\}$ و $\{y_1, \dots, y_m\}$ پایه‌های استاندارد $\mathbb{R}^n$ و $\mathbb{R}^m$ باشند و T توسط (۳.۱) تعریف شده باشد، آن‌گاه از نامساوی کوشی‌ـ شوارتس داریم

$$\|Tx\|^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right)^2 \le \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}^2 \sum_{j=1}^{n} c_j^2 \right) = \sum_{i,j} a_{ij}^2 \|x\|^2 \ .$$

بنابراین

$$\|T\| \le \left[\sum_{i,j} a_{ij}^2 \right]^{\frac{1}{2}} \ .$$

اگر این رابطه را برای $S-T$ به‌جای T بنویسیم (وقتی که $T,S\in L(\mathbb{R}^n,\mathbb{R}^m)$) نتیجه می‌گردد که اگر درایه‌های ماتریس یعنی a_{ij} ها توابعی پیوسته از یک پارامتر باشند، آن‌گاه خود T نیز تابعی پیوسته از آن پارامتر است. به‌عبارت دقیق‌تر فرض کنیم A یک فضای متری و $a_{11},\dots,a_{mn}$ توابعی پیوسته روی A باشند و فرض کنیم برای هر $p\in A$، T_p تبدیل خطی از $\mathbb{R}^n$ به $\mathbb{R}^m$ باشد که ماتریس آن دارای درایه‌های $a_{ij}(p)$ است. در این صورت نگاشت $p\mapsto T_p$ از A به توی $L(\mathbb{R}^n,\mathbb{R}^m)$ پیوسته است.

۳.۱ تمرین

۱. ثابت کنید برای هر $A\in L(\mathbb{R}^n,\mathbb{R}^1)$ بردار منحصر به فردی مانند $y\in\mathbb{R}^n$ وجود دارد به‌قسمی‌که $Ax=x\cdot y$ و $\|A\|=\|y\|$.

(راهنمایی: تحت شرایط خاصی نامساوی شوارتس به تساوی تبدیل می‌گردد.)

۲. احکام مثال ۳.۲.۱ را ثابت کنید.

۳. فرض کنید توابع حقیقی مقدار f و g روی $[a,b]$ انتگرال‌پذیرند.

(الف) ثابت کنید

$$\left|\int_a^b f(x)g(x)\,dx\right| \le \left(\int_a^b f^2(x)\,dx\right)^{\frac{1}{2}}\left(\int_a^b g^2(x)\,dx\right)^{\frac{1}{2}}.$$

(راهنمایی: دو حالت $\int_a^b (f-\lambda g)^2=0$ را برای $\lambda\in\mathbb{R}$ و $\int_a^b (f-\lambda g)^2>0$ برای هر $\lambda\in\mathbb{R}$ را در نظر بگیرید.)

(ب) اگر تساوی برقرار باشد آیا لزوماً $\lambda\in\mathbb{R}$ وجود دارد که $f=\lambda g$؟ اگر f و g پیوسته باشند چطور؟

۴. تبدیل خطی $T:\mathbb{R}^n\to\mathbb{R}^n$ را حافظ نرم خوانیم، در صورتی که $\|T(x)\|=\|x\|$ و آن را حافظ ضرب داخلی گوییم در صورتی که $\langle Tx,Ty\rangle=\langle x,y\rangle$.

(الف) ثابت کنید T حافظ نرم است، اگر و فقط اگر حافظ ضرب داخلی باشد.
(ب) ثابت کنید اگر T یک چنین تبدیل خطی باشد، آن‌گاه یک‌به‌یک است و T^{-1} نیز از همان نوع است.

۵. فرض کنید $T : \mathbb{R}^n \to \mathbb{R}^m$ یک تبدیل خطی است. بدون استفاده از تعریف $\|T\|$ نشان دهید، عددی مانند M وجود دارد به‌قسمی که برای هر $h \in \mathbb{R}^n$، $\|T(h)\| \leq M\|h\|$.

فصل ۲

مشتق‌گیری

۱.۲ مشتق‌گیری از توابع چندمتغیره

در ریاضیات عمومی دیده‌ایم که مشتق تابع $f : \mathbb{R} \longrightarrow \mathbb{R}$ در نقطهٔ $x = a$ به‌صورت

$$\lim_{h\to ۰} \frac{f(x+h)-f(x)}{h} \tag{۱.۲}$$

تعریف می‌شود و این مقدار در صورت وجود با $f'(x)$ نمایش داده می‌شود. به‌دنبال جای‌گزین مناسب مفهوم مشتق در نقطهٔ a هستیم که تعمیم مناسبی از کلیه پدیده‌های حاصل از مشتق‌پذیری یک تابع در یک نقطه را به دست بدهد. لذا باید دید مناسب‌تری از مقدار $f'(x)$ داشته باشیم. به هر عدد حقیقی می‌توان به‌صورت یک تبدیل خطی روی $\mathbb{R}$ نگریست به‌عبارتی تناظر دوسویی بین همهٔ عملگرهای خطی روی $\mathbb{R}$ یا همان ماتریس‌های ۱ × ۱ و همهٔ اعداد حقیقی وجود دارد:

$$\mathbb{R} \longrightarrow \mathbb{M}_{۱\times۱}(\cong L(\mathbb{R}))$$

$$\alpha \longrightarrow [\alpha]$$

که در آن

$$[\alpha] : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longrightarrow \alpha x\ .$$

لذا $f'(x)$ نه فقط یک عدد حقیقی بلکه یک ماتریس(عملگر خطی) است. بنابراین می‌توان تعریف (۲.۱) را به‌صورت زیر بازنویسی کرد:

تابع $f : \mathbb{R} \longrightarrow \mathbb{R}$ در x دارای مشتق است هرگاه ماتریس ۱ × ۱ حقیقی مانند $[f'(x)]$ چنان موجود باشد که

$$\lim_{h\to 0}\left|\frac{f(x+h)-f(x)-[f'(x)]h}{h}\right| = 0\ .$$

حال اگر تابع $f : (a,b) \subseteq \mathbb{R} \longrightarrow \mathbb{R}^m$ تعریف شده باشد نیز می‌توان به همان ترتیب عمل کرد. در این حالت همان‌گونه که می‌دانیم $f'(x)$ به‌صورت برداری مانند $y \in \mathbb{R}^n$ (در صورت وجود) تعریف می‌شود که

$$\lim_{h\to 0}\left\|\frac{f(x+h)-f(x)}{h} - y\right\| = 0\ .$$

حال اگر $\mathbb{R}$ را با $\mathbb{R}^n$ جای‌گزین نماییم، می‌توان تعریف مشتق توابعی که از $\mathbb{R}^n$ به $\mathbb{R}^m$ تعریف می‌شوند را به‌صورت زیر بیان کنیم.

تذکر. برای تأکید تکرار می کنیم که در سرتاسر این کتاب، فضای $\mathbb{R}^n$ را با نرم اقلیدسی ($\|\cdot\|_2$) در نظر می‌گیریم.

تعریف ۲.۱.۱. فرض کنیم $E \subseteq \mathbb{R}^n$ باز است و $f : E \to \mathbb{R}^m$ و $x \in E$ دلخواه باشد. اگر تبدیل خطی $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ وجود داشته باشد به‌قسمی که

$$\lim_{h\to 0}\frac{\|f(x+h)-f(x)-Ah\|}{\|h\|} = 0. \qquad (۲.۲)$$

آن‌گاه گوییم f در x مشتق‌پذیراست و می‌نویسیم $f'(x) = A$. اگر f در هر نقطهٔ $x \in E$ مشتق‌پذیر باشد، آن‌گاه گوییم f در E مشتق‌پذیر است.

تذکر. منظور از همگرایی $h \to \circ$، همگرایی در نرم فضای $\mathbb{R}^n$ است؛ به‌این معنی که $\|h\| \to \circ$.

توجه می‌کنیم در (۲.۲)، $h \in \mathbb{R}^n$ و در صورتی که بردار h از لحاظ نرم کوچک باشد، چون E باز است، پس $x + h \in E$ و در نتیجه $f(x+h)$ تعریف می‌شود و متعلق به $\mathbb{R}^m$ است. همچنین از آنجا که $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ پس $Ah \in \mathbb{R}^m$. بنابراین

$$f(x+h) - f(x) - Ah \in \mathbb{R}^m \ .$$

پس در (۲.۲) نرم صورت همان نرم $\mathbb{R}^m$ و نرم مخرج، نرم $\mathbb{R}^n$ است. همانند توابع حقیقی مقدار (۲.۲) را می‌توان به‌صورت $f(x+h)-f(x) = Ah+r(h)$ نوشت که $\lim\limits_{h\to\circ}\frac{\|r(h)\|}{\|h\|} = \circ$. پس می‌توان گفت که مشتق تابع $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ در نقطه $x \in E$ عبارت است از تابعی خطی از $\mathbb{R}^n$ به $\mathbb{R}^m$ که مقدارش در نقطهٔ h تقریباً برابر $f(x+h) - f(h)$ است.

اولین قضیه‌ای که ثابت می‌کنیم نشان می‌دهد که مشتق تابع در یک نقطه در صورت وجود، یکتا است.

قضیه ۲.۱.۲. فرض کنیم E و f مطابق تعریف ۱.۱.۲ باشند و $x \in E$ و فرض کنیم A_1 و A_2 در (۲.۲) صدق کنند. در این صورت $A_1 = A_2$.

برهان. قرار می‌دهیم $B = A_1 - A_2$. نامساوی

$$\|Bh\| \leq \|f(x+h) - f(x) - A_1 h\| + \|f(x+h) - f(x) - A_2 h\|$$

نشان می‌دهد که

$$\lim_{h\to\circ} \frac{\|Bh\|}{\|h\|} = \circ \ .$$

پس برای بردار ثابت و ناصفر h

$$\lim_{t\to \circ}\frac{\|B(th)\|}{\|th\|}=\circ\,.$$

چون B خطی است پس در سمت چپ این رابطه می‌توان $|t|$ را از صورت و مخرج حذف کرد و بنابراین برای هر $h\in\mathbb{R}^n$، $Bh=\circ$ پس $B=\circ$. $\square$

تذکر. اگر f با علایم تعریف ۱.۱.۲ در E مشتق‌پذیر باشد، آن‌گاه برای هر $x\in E$، $f'(x)$ یک تبدیل خطی از $\mathbb{R}^n$ به $\mathbb{R}^m$ است. همچنین f' نیز یک نگاشت است؛ نگاشتی از E به $L(\mathbb{R}^n,\mathbb{R}^m)$.

قضیه ۳.۱.۲. با علایم تعریف ۱.۱.۲، اگر f در نقطهٔ $x\in E$ مشتق‌پذیر باشد، آن‌گاه در این نقطه پیوسته است.

برهان. چون f در x مشتق‌پذیر است، لذا $f(x+h)-f(x)=f'(x)h+r(h)$ که وقتی $h\longrightarrow\circ$ آن‌گاه $r(h)\longrightarrow\circ$. بنابراین $\lim\limits_{h\to\circ}(f(x+h)-f(x))=\circ$ $\square$

به‌عنوان مثالی از محاسبه مشتق یک تابع فرض کنیم $A:\mathbb{R}^n\to\mathbb{R}^m$ یک تبدیل خطی باشد. در این صورت بنا بر خطی بودن A،

$$A(x+h)-Ax=Ah\,.$$

بنابراین در (۲.۲) برای A به‌جای f، صورت کسر برابر با صفر است و در نتیجه برای تبدیل خطی A و به‌ازای برای هر $x\in\mathbb{R}^n$ داریم $A'(x)=A$؛ یعنی مشتق یک تابع خطی در هر نقطه، برابر خود آن تابع خطی است. دیدیم با تعریفی که برای مشتق توابع چندمتغیره بیان شد، همان‌گونه که انتظار می‌رفت همانند توابع یک متغیره، مشتق‌پذیری پیوستگی را ایجاب می‌کند. اکنون نشان می‌دهیم که قاعدهٔ زنجیری نیز برای توابع چند متغیره برقرار است.

قضیه ۴.۱.۲ (قاعدهٔ زنجیری)**.** فرض کنیم $E \subseteq \mathbb{R}^n$ باز و f نگاشتی از E به‌توی $\mathbb{R}^n$ باشد و g نگاشتی از یک مجموعه شامل $f(E)$ به‌توی $\mathbb{R}^k$ باشد و فرض کنیم f در $x_\circ \in E$ و g در $f(x_\circ)$ مشتق‌پذیر باشد. در این صورت نگاشت $g \circ f$ از E به‌توی $\mathbb{R}^k$ در $x_\circ$ مشتق‌پذیر است و $(g \circ f)'(x_\circ) = g'(f(x_\circ))f'(x_\circ)$ توجه می‌کنیم که در سمت راست این رابطه حاصلضرب دو تبدیل خطی را به مفهوم بخش ۲.۱ داریم.

برهان. نگاشت F از E به‌توی $\mathbb{R}^k$ را که به‌صورت $F(x) = g(f(x))$ تعریف می‌شود در نظر می‌گیریم. می‌خواهیم ثابت کنیم که

$$F'(x_\circ) = g'(f(x_\circ))f'(x_\circ). \tag{۳.۲}$$

قرار می‌دهیم $y_\circ = f(x_\circ)$، $A = f'(x_\circ)$ و $B = g'(y_\circ)$. برای هر $h \in \mathbb{R}^n$ و هر $k \in \mathbb{R}^n$ که به ازای آن‌ها $f(x_\circ + h)$ و $g(y_\circ + k)$ تعریف شده باشند، توابع $U(h)$ و $V(k)$ را به‌صورت

$$U(h) = f(x_\circ + h) - f(x_\circ) - Ah$$
$$V(k) = g(y_\circ + k) - g(y_\circ) - Bk$$

تعریف می‌کنیم. در این صورت

$$\|U(h)\| = \varepsilon(h)\|h\| \quad \text{و} \quad \|V(k)\| = \eta(k)\|k\|, \tag{۴.۲}$$

که در آن ε و η توابعی حقیقی هستند به‌طوری که وقتی $(h, k) \longrightarrow \circ$ آن‌گاه $(\varepsilon(h), \eta(k)) \longrightarrow \circ$.
فرض کنیم h داده شده باشد، قرار می‌دهیم $k = f(x_\circ + h) - f(x_\circ)$. پس

$$\|k\| = \|Ah + U(h)\| \leq \big[\|A\| + \varepsilon(h)\big]\|h\| \tag{۵.۲}$$

$$F(x_\circ + h) - F(x_\circ) - BAh = g(y_\circ + k) - g(y_\circ) - BAh$$

$$= B(k - Ah) + V(k)$$
$$= BU(h) + V(k)\ .$$

پس بنا بر (۴.۲) و (۵.۲)

$$\frac{\|F(x_\circ + h) - F(x_\circ) - BAh\|}{\|h\|} \le \|B\|\varepsilon(h) + \big[\|A\| + \varepsilon(h)\big]\eta(k)\ .$$

حال اگر $h \to \circ$ آن‌گاه $\varepsilon(h) \to \circ$. همچنین بنا بر (۵.۲)، $k \to \circ$ و در نتیجه $\eta(k) \to \circ$.
پس $F'(x_\circ) = BA$ که همان (۳.۲) است. □

نگاشت f از زیرمجموعهٔ باز $E \subseteq \mathbb{R}^n$ به‌توی $\mathbb{R}^n$ را در نظر می‌گیریم. فرض کنیم $\{e_1, \ldots, e_n\}$ و $\{u_1, \ldots, u_m\}$ به‌ترتیب پایه استاندارد در $\mathbb{R}^n$ و $\mathbb{R}^m$ باشد. مؤلفه‌های f توابع حقیقی f_1، ... و f_m می‌باشند که با

$$f(x) = \big(f_1(x), \ldots, f_m(x)\big) = \sum_{i=1}^{m} f_i(x)u_i \qquad (x \in E)$$

که با $f_i(x) = f(x) \cdot u_i$، ($1 \le i \le m$) تعریف می‌شوند. همانند توابع حقیقی مقدار، برای $x \in E$، $1 \le i \le m$ و $1 \le j \le n$ مقدار

$$\lim_{t \to \circ} \frac{f_i(x + te_j) - f_i(x)}{t}$$

را در صورت وجود، مشتق جزئی f در نقطهٔ x خوانیم و به $\frac{\partial f_i(x)}{\partial x_j}$ و یا $D_j f_i(x)$ نشان می‌دهیم. اگر u یک بردار واحد در $\mathbb{R}^n$ باشد، مشتق جهتی f در x در جهت u نیز به‌صورت $\lim_{t\to\circ}\frac{f(x+tu)-f(x)}{t}$ تعریف می‌شود (در صورت وجود حد) و آن را به $(D_u f)(x)$ نشان می‌دهیم. در مثال ذیل نشان می‌دهیم که وجود تمام مشتق‌های جزئی برای وجود مشتق یک تابع کفایت نمی‌کند ولی همان‌گونه که قضیه بعدی زیر نشان می‌دهد عکس این مطلب همواره برقرار است.

مثال ۵.۱.۲. تابع $f : \mathbb{R}^۲ \longrightarrow \mathbb{R}$ با ضابطه

$$f(x,y) = \begin{cases} \dfrac{xy}{(x^۲+y)} & x^۲ \neq -y \\ ۰ & x^۲ = y \end{cases}$$

را در نظر می‌گیریم. در این صورت اگر $e = (e_۱, e_۲)$ آن‌گاه

$$\lim_{t\to ۰} \frac{۱}{t} f(te_۱, te_۲) = \lim_{t\to ۰} \frac{۱}{t} \frac{t^۲ e_۱ e_۲}{t^۲ e_۱^۲ + te_۲} = \lim_{t\to ۰} \frac{te_۱ e_۲}{t^۲ e_۱^۲ + te_۲} = e_۱ \ .$$

بنابراین کلیه مشتقات جهتی در $(۰,۰)$ موجودند، اما f در $(۰,۰)$ پیوسته نیست زیرا برای $x^۲$ های نزدیک به $-y$ با xy های کوچک، f خیلی بزرگ خواهد بود. به‌عبارتی برای ϵ کوچک اگر قرار دهیم $y = -x^۲ + \epsilon$، در این صورت با انتخاب ϵ، مقدار xy در کنترل خواهد بود و می‌توان آن را کوچک انتخاب کرد. به‌عبارت دیگر

$$z(x) = xy = x(-x^۲ + \epsilon) = -x^۳ + \epsilon x$$

به‌ازای $x = \sqrt{۴/۳}$ به مینیمم خود می‌رسد. یعنی برای ϵ به‌قدر کافی کوچک، نقاط به‌فرم $(\sqrt{۴/۳}, -۴/۳ + \epsilon)$ نقاطی هستند که فاصلهٔ y و $-x^۲$ به اندازهٔ ϵ و xy کوچک است. بنابراین f در $(۰,۰)$ مشتق‌پذیر نیست.

قضیه ۶.۱.۲. فرض کنیم f نگاشتی از مجموعهٔ باز $E \subseteq \mathbb{R}^n$ به‌توی $\mathbb{R}^m$ باشد، و فرض کنیم f در نقطهٔ $x \in E$ مشتق‌پذیر باشد. در این صورت مشتق جزئی $(D_j f_i)(x)$ وجود دارد و

$$f'(x)e_j = \sum_{i=۱}^{m} (D_j f_i)(x) u_i \ . \qquad (۱ \leq j \leq n)$$

که در آن، u_i به‌ازای $۱ \leq i \leq m$ پایهٔ استاندارد $\mathbb{R}^m$ است.

برهان. j را بین ۱ و n در نظر می‌گیریم. چون f در x مشتق‌پذیر است داریم:

$$f(x+te_j)-f(x)=f'(x)(te_j)+r(te_j)$$

که در آن $\frac{\|r(te_j)\|}{t}\to \circ$ وقتی که $t\to\circ$. بنابراین از خطی بودن $f'(x)$ نتیجه می‌گردد که

$$\lim_{t\to\circ}\frac{f(x+te_j)-f(x)}{t}=f'(x)e_j\ .$$

حال اگر f را برحسب مؤلفه‌هایش بنویسیم خواهیم داشت

$$\lim_{t\to\circ}\sum_{i=\text{۱}}^{m}\frac{f_i(x+te_j)-f_i(x)}{t}u_i=f'(x)e_j\ .$$

پس بنا بر آن‌چه در مورد حد نگاشت‌های برداری می‌دانیم، بایستی حد هر یک از خارج‌قسمت‌های موجود در حاصل جمع بالا وجود داشته باشد. یعنی هر $(D_jf_i)(x)$ موجود است و قضیه ثابت می‌شود. □

اگر ماتریس نمایشگر $f'(x)$ نسبت به پایه‌های استاندارد را به‌گونه‌ای که در انتهای بخش ۲ دیدیم به $[f'(x)]$ نشان دهیم، آن‌گاه $f'(x)e_j$ بردار ستونی jام از ماتریس $[f'(x)]$ می‌باشد. قضیه قبل نشان می‌دهد که عدد $(D_jf_i)(x)$ درایهٔ سطر iام و ستون jام ماتریس است؛ یعنی

$$[f'(x)]=\begin{bmatrix}(D_\text{۱}f_\text{۱})(x) & \dots & (D_nf_\text{۱})(x)\\ \vdots & \dots & \vdots\\ (D_\text{۱}f_m)(x) & \dots & (D_nf_m)(x)\end{bmatrix}.$$

حال اگر $h=\sum_{j=\text{۱}}^{n}h_je_j$ برداری در $\mathbb{R}^n$ باشد، بنا بر قضیه قبل

$$f'(x)h=\sum_{i=\text{۱}}^{m}\left(\sum_{j=\text{۱}}^{n}(D_jf_i)(x)h_j\right)u_i$$

مثال ۷.۱.۲. فرض کنیم $f:\mathbb{R}^۲\longrightarrow\mathbb{R}^۳$ با ضابطهٔ $f(x,y)=(x^۲,x^۳y,x^۴y^۲)$ تابع باشد. دراین صورت داریم

$$Df(x,y)=\begin{bmatrix}\frac{\partial f_۱}{\partial x} & \frac{\partial f_۱}{\partial y}\\ \frac{\partial f_۲}{\partial x} & \frac{\partial f_۲}{\partial y}\\ \frac{\partial f_۳}{\partial x} & \frac{\partial f_۳}{\partial y}\end{bmatrix}=\begin{bmatrix}۲x & ۰\\ ۳x^۲y & x^۳\\ ۴x^۳y^۲ & ۲x^۴y\end{bmatrix}.$$

تعریف ۸.۱.۲. تابع $f:E\subseteq\mathbb{R}^n\to\mathbb{R}$ را در نظر می‌گیریم و فرض می‌کنیم مشتقات جزئی آن $D_۱f$ و ... و D_nf در نقطه x موجود باشند. تابع با مقادیر برداری ∇f را که به‌صورت

$$\nabla f(x)=\big(D_۱f(x),\dots,D_nf(x)\big)=\sum_{i=۱}^{n}(D_if)(x)e_i$$

تعریف می‌شود، گرادیان f می‌خوانیم.

قضیه ۹.۱.۲. فرض کنیم $f:E\subseteq\mathbb{R}^n\to\mathbb{R}$ در نقطهٔ $x\in E$ مشتق‌پذیر باشد. در این صورت

(الف) گرادیان $\nabla f(x)$ موجود است و $f'(x)t=\nabla f(x)\cdot t$

(ب) مشتق جهتی $D_uf(x)$ موجود است و $D_uf(x)=\nabla f(x)\cdot u$.

برهان. **(الف)** قضیه ۶.۱.۲ را برای حالت $m=۱$ می‌نویسیم، نتیجه می‌گردد که برای $۱\le j\le n$،

$$f'(x)e_j=(D_jf)(x).$$

فرض کنیم $t=(t_۱,\dots,t_n)$ عضوی از $\mathbb{R}^n$ باشد، در این صورت باتوجه به رابطهٔ بالا

$$f'(x)t=f'(x)(t_۱e_۱+\cdots+t_ne_n)=\sum_{j=۱}^{n}t_j(D_jf)(x)=\nabla f(x)\cdot t\ .$$

(ب) فرض کنیم u یک بردار واحد است. (الف) را برای u به‌جای t می‌نویسیم

$$f'(x)u = \nabla f(x) \cdot u.$$

پس کافی است نشان دهیم که $f'(x)u$ برابر است با $D_u f(x)$. چون $f'(x)$ موجود است اگر $\|h\|$ کوچک باشد، می‌توان نوشت $f(x+h) - f(x) - f'(x) = r(h)$ که $\lim_{h\to 0} \frac{r(h)}{h} = 0$.

پس برای هر $\varepsilon > 0$ می‌توان δ را چنان معین کرد که اگر $\|h\| < \delta$ آن‌گاه $\|r(h)\| < \varepsilon\|h\|$ و یا

$$\|f(x+h) - f(x) - f'(x)h\| < \varepsilon\|h\| \ .$$

برای عدد حقیقی کوچک $\lambda \neq 0$، این رابطه را به‌ازای $h = \lambda u$ می‌نویسیم

$$\|f(x+\lambda u) - f(x) - f'(x)\lambda u\| < \varepsilon\|\lambda\| \ .$$

بنابراین اگر $0 < |\lambda| < \delta$، آن‌گاه

$$\left\|\frac{f(x+\lambda u) - f(x)}{\lambda} - f'(x)u\right\| < \varepsilon \ .$$

یعنی $D_u f(x) = f'(x)u$ و (ب) ثابت می‌شود. $\square$

قبلاً در مبحث مشتق توابع با مقادیر برداری در قضیه‌ای مشابه قضیهٔ مقدار میانگین دیده‌ایم که اگر $f : [a,b] \to \mathbb{R}^m$ پیوسته باشد و در بازه باز (a,b) مشتق‌پذیر باشد، آن‌گاه $x \in (a,b)$ ی وجود دارد به‌قسمی که $f(b) - f(a) = f'(c)(b-a)$ و لذا داریم

$$|f(b) - f(a)| \leq |b-a|\,|f'(x)| \ .$$

متأسفانه چنین حکمی به‌طور کامل برای توابع چندمتغیره برقرار نیست. به‌عنوان مثال خواننده علاقه‌مند می‌تواند وجود چنین c ای را برای تابع $f : \mathbb{R} \longrightarrow \mathbb{R}^۲$ با تعریف $f(x, y) = (x^۲, x^۳)$ بررسی کند. اما مشاهده خواهیم کرد که قضیه‌ای نظیر این، برای توابع n متغیره نیز وجود دارد.

تعریف ۱۰.۱.۲. برای $a, b \in \mathbb{R}^n$ منظور از قطعهٔ خطی که a و b را به هم وصل می‌کند، منحنی

$$\begin{aligned} \gamma: \quad & [۰, ۱] \longrightarrow \mathbb{R}^n \\ & t \quad \longrightarrow tb + (۱ - t)a, \end{aligned}$$

است.

قضیه ۱۱.۱.۲ (قضیه مقدار میانگین)**.** (الف) فرض کنیم $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ روی مجموعه باز E مشتق‌پذیر باشد. برای هر $a, b \in E$ اگر قطعه خطی که a و b را به یکدیگر وصل می‌کند در E جای داشته باشد، آن‌گاه نقطه‌ای مانند c روی این قطعه خط وجود دارد که

$$f(b) - f(a) = f'(c)(b - a).$$

(ب) فرض کنیم $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ روی مجموعه باز E مشتق‌پذیر باشد. فرض کنیم $a, b \in E$ و $f = (f_۱, \ldots, f_m)$. اگر خطی که a و b را به یکدیگر وصل می‌کند در E جای داشته باشد. آن‌گاه نقاط $c_۱$ ، ... و c_m روی این قطعه خط وجود دارند به‌قسمی که،

$$f_i(b) - f_i(a) = f_i'(c_i)(b - a) . \qquad i = ۱, \ldots, m$$

برهان. (الف) تابع $h:[۰,۱]\to\mathbb{R}$ را که به‌صورت $h(t)=f\big((۱-t)a+tb\big)$ تعریف شده است، در نظر می‌گیریم. این تابع در $(۰,۱)$ مشتق‌پذیر است، پس بنا بر قضیه مقدار میانگین معمولی، $t_۰\in(۰,۱)$ ی وجود دارد که

$$h(۱)-h(۰)=h'(t_۰)(۱-۰)\ .$$

ولی $h(۱)=f(b)$ و $h(۰)=f(a)$ و بنا بر قاعده زنجیری چون مشتق $(۱-t)x+ty$ نسبت به t برابر $y-x$ است. داریم

$$h'(t_۰)=Df\big((۱-t_۰)a+t_۰b\big)(b-a)\ .$$

پس کافی است $c=(۱-t)a+t_۰b$ اختیار شود.

(ب) کافی است (الف) را برای هر مؤلفهٔ f_i به کار گیریم.

□

نتیجه ۱. فرض کنیم f نگاشتی از مجموعهٔ محدب و باز $E\subset\mathbb{R}^n$ به‌توی $\mathbb{R}^m$ باشد. اگر f روی E مشتق‌پذیر بوده و عددی مانند M وجود داشته باشد که برای هر $x\in M$، $\|Df(x)\|\le M$، آن‌گاه برای هر $a,b\in E$

$$\|f(b)-f(a)\|\le M\|b-a\|\ .$$

برهان. بدیهی است. □

نتیجه ۲. اگر علاوه بر مفروضات نتیجه ۱، برای هر $x\in E$ داشته باشیم $f'(x)=۰$، آن‌گاه f تابع ثابت است.

برهان. کافی است توجه کنیم که در این حالت در نتیجه ۱ داریم $M=۰$. □

تعریف ۱۲.۱.۲. نگاشت مشتق‌پذیر f از زیرمجموعهٔ باز $E \subseteq \mathbb{R}^n$ به‌توی $\mathbb{R}^m$ را پیوسته‌ـ مشتق‌پذیر خوانیم هرگاه f' نگاشتی پیوسته از E به‌توی $L(\mathbb{R}^n, \mathbb{R}^m)$ باشد. مجموعه تمام توابع پیوسته‌ـ مشتق‌پذیر بر E را با $C^1(E)$ نشان می‌دهیم در حقیقت $f \in C^1(E)$، اگر و فقط اگر برای هر $x \in E$ و برای هر $\varepsilon > 0$ عددی مانند $\delta > 0$ وجود داشته باشد، به‌قسمی که برای هر $y \in E$، اگر $\|x - y\| < \delta$، آن‌گاه $\|f'(y) - f'(x)\| < \varepsilon$.

لزوماً مشتق یک تابع، پیوسته نیست. به‌عنوان مثال ساده و شناخته شده می‌توان به تابع

$$f(x) = \begin{cases} x^2 \sin\left(\dfrac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

اشاره کرد. قبلاً دیدیم که برای تابع $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^m$ وجود تمامی مشتق‌های جزئی، مشتق‌پذیری را ایجاب نمی‌کند. اکنون نشان می‌دهیم که اگر همه مشتق‌های جزئی موجود و پیوسته باشند، آن‌گاه نه تنها f' موجود است بلکه پیوسته نیز می‌باشد.

قضیه ۱۳.۱.۲. فرض کنیم $E \subseteq \mathbb{R}^n$ باز و $f : E \longrightarrow \mathbb{R}$ یک تابع باشد. اگر برای هر $j = 1, ..., n$ مشتق جزئی $D_j f$ روی E موجود و پیوسته باشد، آن‌گاه $f \in C^1(E)$.

برهان. فرض کنیم x عضوی از E و $\varepsilon > 0$ داده شده باشد. چون E باز است گوی بازی به مرکز x وشعاع r وجود دارد به‌قسمی که $B_r(x) \subset E$. باتوجه به پیوستگی $D_j f$ می‌توان r را چنان اختیار کرد که برای هر $y \in B_r(x)$؛

$$\|(D_j f)(y) - (D_j f)(x)\| < \frac{\varepsilon}{n} \,. \tag{۶.۲}$$

فرض کنیم $h = (h_1, \ldots, h_n)$ و $\|h\| < r$. قرار می‌دهیم $V_0 = 0$ و برای $1 \leq k \leq n$،

$$V_k = (h_1, \ldots, h_k, 0, \ldots, 0) = h_1 e_1 + \ldots + h_k e_k \,.$$

در این صورت

$$f(x+h)-f(x)=\sum_{j=1}^{n}\big[f(x+V_j)-f(x+V_{j-1})\big]\ . \qquad (۷.۲)$$

چون برای $1\leq k\leq n$، $r>|V_k|$ و به علاوه $B_r(x)$ محدب است، قطعه خط واصل انتهای $x+V_{j-1}$ و $x+V_j$ در $B_r(x)$ جای دارد. چون $V_j=V_{j-1}+h_je_j$، از قضیه مقدار میانگین برای هر یک از مؤلفه‌های (۷.۲) داریم

$$f(x+V_j)-f(x+V_{j-1})=h_j(D_jf)(w_j),$$

که در آن

$$w_j=(x_1+h_1,\dots,x_{j-1}+h_{j-1},x_j+u_j,x_{j+1},\dots,x_n)$$

و $0\leq u_j\leq 1$. ولی بنا بر (۶.۲) اختلاف بین عبارت با $h_j(D_jf)(x)$ از $|h_j|\frac{\varepsilon}{n}$کم‌تر است. حال بنا بر (۷.۲) برای هر h، اگر $\|h\|<r$ آن‌گاه

$$\Big\|f(x+h)-f(x)-\sum h_j(D_jf)(x)\Big\|\leq\frac{1}{n}\sum_{j=1}^{n}|h_j|\varepsilon\leq\|h\|\varepsilon\ .$$

پس f در نقطهٔ x مشتق‌پذیر است و مشتق آن عبارت است از تابع خطی $h\mapsto\sum h_j(D_jf)(x)$.
برای اثبات پیوستگی f' توجه می‌کنیم که در این حالت ماتریس $[f'(x)]$ ماتریسی سطری است که تشکیل شده است از سطر

$$\big[(D_1f)(x),\dots,(D_nf)(x)\big]$$

و چون D_1f ، … و D_nf روی E پیوسته‌اند پس بنا بر آن‌چه در انتهای بخش ۲.۱ آوردیم $f\in C^1(E)$. این قضیه را به شکل زیر نیز می‌توان بیان کرد. □

قضیه ۱۴.۱.۲. فرض کنیم f نگاشتی از زیرمجموعهٔ باز $E \subseteq \mathbb{R}^n$ به‌توی $\mathbb{R}^m$ باشد. در این صورت $f \in C^1(E)$، اگر و فقط اگر برای $1 \leq i \leq m$ و $1 \leq j \leq n$، مشتق‌های جزئی $D_j f_i$ روی E موجود و پیوسته باشند.

برهان. فرض کنیم $f \in C^1(E)$. در قضیه ۶.۱.۲ دیدیم که مشتق‌های جزئی موجودند و

$$f'(x)e_j = \sum_{i=1}^{m}(D_j f_i)(x)u_i \ .$$

پس برای هر i و j و هر $x \in E$

$$(D_j f_i)(x) = [f'(x)e_j] \cdot u_i \ .$$

بنابراین

$$(D_j f_i)(y) - (D_j f_i)(x) = \{[f'(y) - f'(x)]e_j\} \cdot u_i$$

که با توجه به $\|u_i\| = |e_j| = 1$ نتیجه می‌گردد

$$\big|(D_j f_i)(y) - (D_j f_i)(x)\big| \leq \big|\big[f'(y) - f'(x)\big]e_j\big| \leq \big\|f'(y) - f'(x)\big\| \ .$$

پس $D_j f_i$ پیوسته است.

برعکس، نگاشت $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ را به‌صورت $f = (f_1, \ldots, f_m)$ می‌نویسیم. چون برای $1 \leq i \leq m$، f_i حقیقی مقدار است.

در قضیهٔ ۱۴.۱.۲ قرار می‌دهیم $Df_i(x)h = \sum (D_j f_i)(x)h_j$. در این صورت برای هر $\varepsilon > 0$ عدد حقیقی $r > 0$ چنان یافت می‌شود که اگر $\|h\| > r$، آن‌گاه

$$\|f_i(x+h) - f_i(x) - Df_i(x)h\| < \varepsilon\|h\| \ .$$

بنابراین اگر $\|h\| < r$ آن‌گاه

$$\begin{aligned}\left\| f(x+h)-f(x) - \sum_{i=1}^{m}(Df_i(x)h)u_i \right\| &= \left\| \sum [f_i(x+h) - f_i(x) - Df_i(x)h]u_i \right\| \\ &\leq \sum_{i=1}^{m} \left\| [f_i(x+h) - f_i(x) - (Df_i(x))h]u_i \right\| \\ &< m\varepsilon \|h\| .\end{aligned}$$

چون ε دلخواه است، f در نقطهٔ x مشتق‌پذیر است و مشتق آن برابر است با تابع خطی

$$h \mapsto \sum_{i=1}^{m} Df_i(x)h .$$

پیوستگی از قضیهٔ ۱۴.۱.۲ نتیجه می‌گردد. $\square$

۲.۲ تمرین

۱. ثابت کنید اگر S و T تبدیلات خطی باشند، آن‌گاه ST نیز خطی است. همچنین ثابت کنید T^{-1} نیز در صورت وجود خطی و معکوس‌پذیر است.

۲. اگر

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

نشان دهید با اینکه f در $(0,0)$ پیوسته نیست، $D_1 f$ و $D_2 f$ در هر نقطهٔ $\mathbb{R}^2$ موجودند.

۳. فرض کنید f یک تابع حقیقی باشد که در مجموعهٔ باز $E \subseteq \mathbb{R}^n$ تعریف شده و مشتقات جزئی $D_1 f$ ، ... و $D_n f$ در E کراندارند. ثابت کنید f در E پیوسته است.

۴. فرض کنید f یک تابع حقیقی است که در مجموعهٔ باز و محدب $E \subseteq \mathbb{R}^n$ تعریف شده و برای هر $x \in E$، $D_1 f(x) = 0$، نشان دهید f تنها به x_2 ، ... و x_n بستگی خواهد داشت.

۵. تابع $f:\mathbb{R}^{2}\longrightarrow\mathbb{R}$ مستقل از متغیر دوم نامیده می‌شود، هرگاه برای هر $x\in\mathbb{R}, y_{1}, y_{2}\in\mathbb{R}$ داشته باشیم $f(x,y_{1})=f(x,y_{2})$.
نشان دهید f مستقل از متغیر دوم است، اگر وتنها اگر تابع $g:\mathbb{R}\longrightarrow\mathbb{R}$ چنان موجود باشد که $f(x,y)=g(x)$. در این حالت $f'(a,b)$ برحسب g چیست؟

۶. تابع $f:\mathbb{R}^{2}\longrightarrow\mathbb{R}$ به‌صورت زیر تعریف می‌شود:

$$f(x,y)=\begin{cases}\dfrac{xy}{\sqrt{x^{2}+y^{2}}} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}.$$

نشان دهید f در $(0,0)$ مشتق‌پذیر نیست.

۷. فرض کنید $f:\mathbb{R}^{n}\longrightarrow\mathbb{R}$ تابعی با شرط $|f(x)|\leq\|x\|^{2}$ است. نشان دهید f در صفر مشتق‌پذیر است.

۸. توابع $f,g:\mathbb{R}\longrightarrow\mathbb{R}$ در نقطهٔ a از درجهٔ nـ‌ام یکسان نامیده می‌شود، هرگاه

$$\lim_{h\longrightarrow 0}\frac{f(a+h)-g(a+h)}{h^{n}}=0\ .$$

اگر $f'(a)$ ، ... و $f^{(n)}(a)$ موجود باشند. نشان دهید f و g که به‌صورت

$$g(x)=\sum_{i=0}^{n}\frac{f^{i}(a)}{i!}(x-a)^{i}$$

تعریف می‌شود از درجهٔ nـ‌ام در نقطهٔ a یکسان هستند.

۹. اگر $f:\mathbb{R}^{2}\longrightarrow\mathbb{R}$ به‌صورت $f(x,y)=x\cdot y$ تعریف شود آن‌گاه نشان دهید

$$D_{f}(a,b)(x,y)=bx+a,$$

و بنابراین $f'(a,b)=(b,a)$.

۱۰. اگر $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}$ در نقطهٔ a مشتق‌پذیر باشد نشان دهید

$$D(f+g)(a) = Df(a) + Dg(a) ,$$
$$D(f \cdot g)(a) = g(a)Df(a) + f(a)Dg(a) .$$

علاوه براین اگر $g(a) \neq ۰$ آن‌گاه

$$D\left(\frac{f}{g}\right)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{[g(a)]^۲} .$$

۱۱. تابع $f : \mathbb{R}^۳ \longrightarrow \mathbb{R}^۲$ با ضابطه $f(x,y,z) = (x^۴ y, ye^x)$ را در نظر می‌گیریم. Df را محاسبه کنید.

۱۲. فرض کنیم $l : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ یک نگاشت خطی و $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ مفروض باشد، چنان‌که $\|g(x)\| \leq M\|x\|^۲$ و فرض کنیم $f(x) = l(x) + g(x)$. نشان دهید $Df(۰) = l$.

۱۳. نشان دهید

$$f(x,y) = \begin{cases} \dfrac{(xy)^۲}{\sqrt{x^۲ + y^۲}} & (x,y) \neq (۰,۰), \\ ۰ & (x,y) = (۰,۰) \end{cases}$$

در $(۰,۰)$ مشتق‌پذیر است.

۱۴. فرض کنیم $f : [۰, +\infty) \longrightarrow \mathbb{R}$ پیوسته و $f(۰) = ۰$ و f بر $(۰, +\infty)$ مشتق‌پذیر و f' صعودی باشد. نشان دهید $g(x) = \frac{f(x)}{x}$ روی $(۰, +\infty)$ صعودی است.

۱۵. فرض کنید E_i برای $i = ۱, \ldots, k$ فضاهای اقلیدسی با ابعاد متفاوت باشند. تابع

$$f : E_۱ \times \cdots \times E_k \longrightarrow \mathbb{R}^p$$

چندخطی نامیده می‌شود، هرگاه برای هر انتخاب $x_j \in E_j$ و $j \neq i$ تابع $g : E_i \longrightarrow \mathbb{R}^p$ با تعریف

$$g(x) = f(x_۱, \ldots, x_{i-۱}, x, x_{i+۱}, \ldots, x_k),$$

یک نگاشت خطی باشد.

(الف) اگر f یک نگاشت چندخطی و $i \neq j$، نشان دهید برای $h = (h_۱, \ldots, h_k)$ با $h_i \in E_i$ داریم

$$\lim_{h \to ۰} \frac{| f(a_۱, \ldots, h_i, \ldots h_j, \ldots, a_k) |}{| h |} = ۰ \; .$$

(ب) ثابت کنید

$$Df(a_۱, \ldots, a_k)(x_۱, \ldots, x_k) = \sum_{i=۱}^{k} f(a_۱, \ldots, a_{i-۱}, x_i, a_{i+۱}, \ldots, a_k) \; .$$

(ج) ثابت کنید تابع دترمینان $(\det : \mathbb{R}^n \times \ldots \mathbb{R}^n \longrightarrow \mathbb{R})$ مشتق‌پذیر است و

$$D(\det)(a_۱, \ldots, a_n)(x_۱, \ldots, x_n) = \sum_{i=۱}^{k} \det \begin{bmatrix} a_۱ \\ \vdots \\ x_i \\ \vdots \\ a_n \end{bmatrix} \; .$$

۱۶. فرض کنید $g_۱, g_۲ : \mathbb{R}^۲ \longrightarrow \mathbb{R}$ توابع به‌طور پیوسته مشتق باشد و فرض کنیم $D_۱ g_۲ = D_۲ g_۱$. فرض کنیم

$$f(x, y) = \int_۰^x g_۱(t, ۰) \, \mathrm{d}t + \int_۰^y g_۲(x, t) \, \mathrm{d}t \; .$$

نشان دهید $D_۱ f(x, y) = g_۱(x, y)$.

۱۷. فرض کنید $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ مشتق‌پذیر و $f(0) = 0$. ثابت کنید توابع $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ چنان موجود است که

$$f(x) = \sum_i^n x^i g_i(x) \ .$$

فصل ۳

قضیه تابع معکوس- تابع ضمنی- رتبه و مطالب مربوطه

این فصل شامل مطالب و قضایای اصلی مربوط به آنالیز برداری است.

۱.۳ قضیه تابع معکوس

فرض کنیم $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ یک تابع دلخواه باشد. در این صورت دترمینان ژاکوبی f که به $Jf(x)$ نشان داده می‌شود، عبارت است از دترمینان ماتریس $Df(x)$ یعنی

$$Jf(x) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{vmatrix}.$$

از جبر خطی می‌دانیم که $Jf(x) \neq 0$، اگر و فقط اگر $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$ یک یکریختی باشد. قضیهٔ تابع معکوس به دنبال آن است که نشان دهد اگر $Df(x)$، که بهترین تقریب

خطی f است، یکریختی باشد یعنی معکوس‌پذیر باشد، آن‌گاه خود f نیز معکوس پذیر است. اگر حالت $n = ۱$، یعنی $f : \mathbb{R} \to \mathbb{R}$ را در نظر بگیریم و فرض کنیم $f'(x_۰) \neq ۰$ آن‌گاه شیب f در $x_۰$ و بنابراین در یک همسایگی $x_۰$ مخالف صفر می‌باشد یعنی f در یک همسایگی $x_۰$ معکوس‌پذیر است. پس بایستی بدنبال معکوس‌پذیری موضعی باشیم. قبل از بیان قضیه تابع معکوس، قضیه نگاشت انقباض را ثابت می‌کنیم.

تعریف ۱.۱.۳. فرض کنیم (X, d) یک فضای متری است. نگاشت $f : X \longrightarrow X$ را انقباض خوانیم، هرگاه عددی مانند $۰ < c < ۱$ وجود داشته باشد، به‌طوری که برای هر $x, y \in X$،

$$d(f(x), f(y)) \leq cd(x, y) \; . \qquad (۱.۳)$$

بدیهی است که هر تابع انقباض پیوسته است؛ کافی است برای هر $\varepsilon > ۰$، δ را برابر $\frac{\varepsilon}{c}$ اختیار کنیم.

به‌منظور پیگیری بهتر قضیه ذیل یادآوری می‌کنیم فضای متری کامل، فضایی است که همه دنباله‌های کوشی آن همگرا باشند.

قضیه ۲.۱.۳ (قضیهٔ نگاشت انقباض)**.** فرض کنیم X یک فضای متری کامل و $f : X \longrightarrow X$ تابع انقباض باشد. در این صورت نقطه منحصر به فرد $x \in X$ وجود دارد، به‌قسمی که $f(x) = x$ (x را نقطه ثابت f می‌خوانیم).

برهان. نقطه دلخواه $x_۰ \in X$ را اختیار می‌کنیم و دنبالهٔ $\{x_n\}$ را به‌صورت بازگشتی

$$x_{n+۱} = f(x_n) \qquad (n = ۰, ۱, ۲, \cdots) \; ,$$

تعریف می‌کنیم. ابتدا نشان می‌دهیم که $\{x_n\}$ یک دنباله کشی است. عدد $۰ < c < ۱$ را چنان اختیار می‌کنیم که (۱.۳) برقرار باشد. برای $n \geq ۱$ داریم

$$d(x_{n+۱}, x_n) \leq d\big(f(x_n), f(x_{n-۱})\big) \leq cd(x_n, x_{n-۱}),$$

که در نتیجه به استقراء خواهیم داشت:

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0) \qquad (n = 0, 1, 2, \cdots) .$$

بنابراین برای هر دو عدد طبیعی n و k

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{n+k-1}, x_{n+k}) \\ &\leq (c^n + c^{n+1} + \cdots + c^{n+k-1}) d(x_1, x_0) \\ &\leq \frac{c^n}{1-c} d(x_1, x_0) . \end{aligned}$$

چون $0 < c < 1$ پس $\{x_n\}$ یک دنباله کشی است. چون X کامل است پس $\{x_n\}$ همگرا به نقطه‌ای منحصر به فرد مانند $x \in X$ است. این نقطه x، نقطه ثابت نگاشت f است. زیرا از این‌که $\lim_{n\to\infty} x_n = x$ و با توجه به پیوستگی f (هر نگاشت انقباض پیوسته است) داریم

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x .$$

حال اگر x' نقطهٔ ثابت دیگری باشد به‌راحتی دیده می‌شود که $x' = x$. □

قضیه ۳.۱.۳ (قضیه تابع معکوس)**.** فرض کنیم f یک نگاشت C^1 از زیرمجموعهٔ باز $E \subset \mathbb{R}^n$ به‌توی $\mathbb{R}^n$ باشد و در نقطهٔ $a \in E$، عملگر $f'(x)$ معکوس‌پذیر باشد (و یا $Jf(x) \neq 0$). در این صورت همسایگی U از a در E و همسایگی V از $b = f(a)$ وجود دارد، به‌قسمی که $f(U) = V$ و f دارای تابع معکوس C^1 چون $f^{-1} : V \to U$ می‌باشد. به‌علاوه برای $y \in V$ و $x = f^{-1}(y)$، داریم

$$Df^{-1}(y) = [Df(x)]^{-1} .$$

برهان. به دلیل طولانی بودن اثبات، آن را به چند مرحله تقسیم می‌کنیم.

مرحله ۱: تبدیل به حالت خاص. کافی است قضیه را در حالتی ثابت کنیم که $Df(a)$ تبدیل همانی باشد. زیرا اگر فرض کنیم $\lambda = Df(a)$ در این صورت λ^{-1} وجود دارد و بنابر قاعدهٔ زنجیری

$$D(\lambda^{-1} \circ f)(a) = D(\lambda^{-1})(f(a)) \circ Df(a) = \lambda^{-1} \circ Df(a) = \mathrm{id} \ .$$

حال اگر قضیه برای $\lambda^{-1} \circ f$ برقرار باشد و g معکوس نگاشت $\lambda^{-1} \circ f$ باشد، آن‌گاه $g \circ \lambda^{-1}$ معکوس f است و قضیه برای f برقرار خواهد بود. همچنین می‌توان فرض کرد که $a = 0$ و $b = f(a) = 0$. برای توضیح این عمل فرض کنیم قضیه در حالت خاص $a = 0$ و $b = 0$ ثابت شده است. فرض می‌کنیم

$$h(x) = f(x+a) - f(a) \ .$$

در این صورت $h(0) = 0$ و $Dh(0) = Df(a)$ و بنابراین $Dh(0)$ معکوس‌پذیر است. حال اگر h در یک همسایگی کوچک از نقطه ۰ دارای معکوس باشد، آن‌گاه نگاشت f^{-1} که به‌وسیله $f^{-1}(y) = h^{-1}(y - f(a)) + a$ تعریف می‌شود، معکوس f در نقطهٔ (نزدیک به a) $x + a$ می‌باشد. پس در این مرحله دیدیم که کافی است قضیه را برای $a = 0$، $b = 0$ و $Df(0) = \mathrm{id}$ ثابت کنیم. با این مفروضات برهان را ادامه می‌دهیم.

مرحله ۲: استفاده از قضیه نقطه ثابت برای یافتن معکوس موضعی. باتوجه به مرحله اول برهان، آن‌چه می‌خواهیم در همسایگی از ۰ است به‌قسمی که برای هر y از همسایگی اول نقطهٔ منحصر به فردی مانند x از همسایگی دوم به‌قسمی موجود باشد که $f(x) = y$. برای این کار، نگاشت g_y را که به وسیلهٔ $g_y(x) = y + x - f(x)$ تعریف شده است، در نظر می‌گیریم. اگر در یک همسایگی بسته از ۰ این نگاشت انقباضی باشد، آن‌گاه دارای نقطه ثابتی مانند x است. یعنی

$$x = y + x - f(x) \ .$$

پس x نقطه منحصر به فردی از این همسایگی است که $y = f(x)$. حال همسایگی مورد نظر را می‌سازیم. نگاشت $g(x) = x - f(x)$ را در نظر می‌گیریم. در این صورت $Dg(\circ) = \circ$. چون g نگاشتی $C^{۱}$ است پس Dg پیوسته است، و بنابر پیوستگی در نقطه ∘ عدد حقیقی مانند $r > \circ$ وجود دارد که اگر $|x| < r$ آن‌گاه $\|Dg(x)\| < \frac{۱}{۲n}$ که $g = (g_{۱}, \ldots, g_n)$.
بنا بر قضیهٔ مقدار میانگین ۱۱.۱.۲ ، برای هر $x \in B_r(\circ)$ نقاط $w_{۱}$ ، ... و w_n در $B_r(x)$ وجود دارند به‌قسمی که

$$g_i(x) = g_i(x) - g_i(\circ) = Dg_i(w_i)(x - \circ) = Dg_i(w_i)(x)\ .$$

پس

$$\|g(x)\| \leq \sum_{i=۱}^{n} \|g_i(x)\| = \sum_{i=۱}^{n} \|Dg_i(w_i)(x)\| \leq \sum_{i=\circ}^{n} \|Dg_i(w_i)\|\|x\| \leq \frac{\|x\|}{۲} < \frac{r}{۲}\ .$$

این نشان می‌دهد که g گوی بسته $\overline{B_r(\circ)}$ را به‌توی $\overline{B_{\frac{r}{۲}}(\circ)}$ نقش می‌کند. زیرا اگر $|y| < \frac{r}{۲}$ و $x \in \overline{B_r(\circ)}$ آن‌گاه $\|g_y(x)\| = \|y + g(x)\| \leq \|y\| + \|g(x)\| < \frac{r}{۲} + \frac{r}{۲} = r$. فرض کنیم $x_{۱}$ و $x_{۲}$ دو نقطهٔ دلخواه $\overline{B_r(\circ)}$ باشند. در این صورت اگر قضیه مقدار میانگین را به کارگیریم خواهیم داشت

$$\left\|g_y(x_{۱}) - g_y(x_{۲})\right\| = \|g(x_{۱}) - g(x_{۲})\| \leq \frac{۱}{۲}\|x_{۱} - x_{۲}\|\ .$$

پس g_y نگاشتی انقباضی (با $c = \frac{۱}{۲}$) است. حال بنابر قضیه نگاشت انقباض، نقطه منحصر به فردی مانند $x \in \overline{B_r(\circ)}$ وجود دارد که نقطه ثابت g_y است. پس $f(x) = y$. این بیان می‌کند که f دارای معکوسی به‌صورت

$$f^{-۱} : \overline{B_{\frac{r}{۲}}(\circ)} \subset \mathbb{R}^n \to \overline{B_r(\circ)} \subset \mathbb{R}^n$$

می باشد.

مرحله ۳: $f^{-۱}$ پیوسته است. فرض کنیم $x \in B_r(\circ)$ و $h > \circ$ عدد کوچکی باشد،

به‌طوری‌که $x+h \in B_r(\circ)$. در این صورت باتوجه به تعریف g،

$$\begin{aligned}\|h\| = \|x+h-x\| &\leq \|f(x+h)-f(x)\| + \|g(x+h) = g(x)\| \\ &\leq \|f(x+h)-f(x)\| + \frac{۱}{۲}\|h\| .\end{aligned}$$

پس $\|h\| \leq ۲\|f(x+h)-f(x)\|$. بنابراین اگر $y \in B_r(\circ)$ و k عدد مثبت کوچکی باشد به‌طوری که $y+k \in B_r(\circ)$، آن‌گاه اگر قرار دهید $x = f^{-۱}(y)$ و $x+h = f^{-۱}(y+k)$، نتیجه می‌شود $\|f^{-۱}(y+k) - f^{-۱}(y)\| \leq ۲\|k\|$ که پیوستگی $f^{-۱}$ را نتیجه می‌دهد.

مرحله ۴: نشان می‌دهیم اگر r به‌قدر کافی کوچک باشد، تابع معکوس روی $B_{\frac{r}{۲}}(\circ)$ مشتق‌پذیر است. چون $Df(\circ)$ معکوس‌پذیر است و تابع $Df : U \subset \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^n)$ پیوسته است، باتوجه به قضیه ۱۳.۲.۱ **(ب)**، $GL(\mathbb{R}^n)$ باز است و در نتیجه $[Df(x)]$ در یک همسایگی ۰ وجود دارد. اگر این همسایگی شامل گوی باز $B_{\frac{r}{۲}}(\circ)$ نباشد، r را کوچکتر اختیار می‌کنیم تا این شرط برقرار شود. حال فرض کنیم $y \in B_{\frac{r}{۲}}(\circ)$ و k را چنان اختیار می‌کنیم که $y+k \in B_{\frac{r}{۲}}(\circ)$. قرار می‌دهیم $x = f^{-۱}(y)$ و $x+h = f^{-۱}(y+k)$. در این صورت

$$\begin{aligned}&\frac{\left\|f^{-۱}(y+k) - f^{-۱}(y) - [Df(x)]^{-۱}(k)\right\|}{\|k\|} \\ &= \frac{\left\|h - [Df(x)]^{-۱}(k)\right\|}{\|k\|} \\ &= \frac{\left\|[Df(x)][Df(x)]^{-۱}(h) - [Df(x)]^{-۱}(f(x+h)-f(x))\right\|}{\|k\|} \\ &\leq \left\|[Df(x)]^{-۱}\right\| \frac{\|f(x+h)-f(x)-[Df(x)](h)\|}{\|k\|} \cdot \frac{\|h\|}{\|k\|} .\end{aligned}$$

در این عبارت باتوجه به مرحلهٔ ۳، $h = f^{-۱}(y+k) - f^{-۱}(y)$ کوچکتر است از

$$۲\left\|[Df(x)]^{-۱}\right\| \frac{\|f(x+h)-f(x)-[Df(x)](h)\|}{h} .$$

ولی f^{-1} پیوسته است، پس $k \to 0$ ایجاب می‌کند $h \to 0$. بنابراین

$$\lim_{k\to 0} \frac{\left\|f^{-1}(y+k) - f^{-1}(k) - [Df(x)]^{-1}(k)\right\|}{\|k\|} = 0 \,.$$

یعنی f^{-1} در نقطه y مشتق‌پذیر است و مشتق آن برابر است با

$$[Df(x)]^{-1} = \Big[Df\big(f^{-1}(y)\big)\Big]^{-1}$$

حال در قضیه کافی است $V = B_{\frac{r}{2}}(0)$ و $U = f^{-1}(V)$ اختیار شود و برای کامل شدن برهان فقط بایستی C^1 بودن f^{-1} و پیوستگی $[Df(x)]^{-1}$ را ثابت کنیم که باتوجه به پیوستگی Df و این که تابع $A \mapsto A^{-1}$ از $GL(\mathbb{R}^n)$ به روی $GL(\mathbb{R}^n)$ پیوسته است (قضیه ۱۴.۲.۱) بدیهی است. به این ترتیب برهان قضیه کامل می‌شود. □

مثال ۴.۱.۳. معادله‌های $u(x,y) = \dfrac{x^4+y^4}{x}$ و $v(x,y) = \sin x + \cos x$ را در نظر بگیرید. در نزدیکی چه نقاط (x,y) می‌توان x و y را بر حسب u و v محاسبه کرد؟

حل. تابع $f : \mathbb{R}^2 \to \mathbb{R}^2$ به‌صورت $f(f_1, f_2)$ را در نظر می‌گیریم که

$$f_1(x) = \frac{x^4+y^4}{x} = u(x,y) \,, \quad f_2(x) = \sin x + \cos x = v(x,y) \,.$$

می‌خواهیم نقاطی را معین کنیم که در نزدیکی آنها بتوان x و y را به‌صورت تابعی از u و v نوشت و به علاوه مشتقات جزئی $\frac{\partial f}{\partial u}$ و نظایر آن را محاسبه نمائیم. طبق قضیهٔ تابع معکوس، ابتدا بایستی دترمینان ژاکوبی F را محاسبه کنیم. توجه داریم که حوزه تعریف f عبارت است از $A = \{(x,y) \in \mathbb{R}^2 |\ x \neq 0\}$. حال

$$Jf(x,y) = \begin{vmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{3x^4-y^4}{x^2} & \frac{4y^3}{x} \\ \cos x & -\sin y \end{vmatrix} = \frac{\sin y}{x^2}(y^4 - 3x^4) - \frac{4y^3}{x}\cos x \,.$$

بنابراین در نقاطی که این عبارت صفر نشود، می‌توانیم x و y را بر حسب u و v به‌دست آوریم. یعنی در نزدیکی x و y هایی که $y \neq 0$ یا $\sin y(y^{4} - 3x^{4}) \neq 4xy^{3} \cos x$، می‌توان x و y را به‌صورت توابعی از u و v بنویسیم. این شرط را نمی‌توان به‌صورت کلی حل کرد. مثلاً اگر $x_0 = \frac{\pi}{2}$ و $y_0 = \frac{\pi}{2}$ آن‌گاه می‌توانیم x و y را در نزدیکی x_0 و y_0 به‌دست آوریم. زیرا در این نقطه $Jf(x,y) \neq 0$. مشتقات جزئی $\frac{\partial x}{\partial u}, \cdots$ را می‌توان بر اساس قضیه تابع معکوس به‌وسیله ماتریس ژاکوبی به‌دست آورد. در حالت دو بعدی خواهیم داشت:

$$\frac{\partial x}{\partial u} = \frac{1}{Jf(x,y)} \frac{\partial V}{\partial y}\,, \qquad \frac{\partial x}{\partial v} = \frac{-1}{Jf(x,y)} \frac{\partial u}{\partial y}\,,$$
$$\frac{\partial y}{\partial u} = \frac{-1}{Jf(x,y)} \frac{\partial u}{\partial x}\,, \qquad \frac{\partial y}{\partial v} = \frac{1}{Jf(x,y)} \frac{\partial u}{\partial x}\,.$$

در این مثال

$$\frac{\partial x}{\partial u} = \frac{-x^{2} \sin y}{\sin y\,(y^{4} - 3x^{4}) - 4y^{2} x \cos x}\,.$$

ملاحظه شد که قضیه تابع معکوس به ما گفت که می‌توان x و y را محاسبه کرد و چگونگی محاسبهٔ مشتقات جزئی را نیز بیان نمود. گرچه حل صریح معادله ممکن است میسر نباشد.

۲.۳ قضیه تابع ضمنی

مجموعهٔ x و y هایی را در نظر می‌گیریم که به‌وسیله معادله $F(x,y) = 0$، به یکدیگر مربوط شده‌اند. در این صورت می‌گوییم این معادله (به‌صورت ضمنی) تابعی مانند $y = f(x)$ را تعریف می‌کند. به‌طور کلی اگر چنین F ای داشته باشیم نوشتن y به‌صورت تابع صریحی از x عمل ساده‌ای نیست. بنابراین یافتن شرایطی که تحت آن بتوان حکم کرد که چه تابعی وجود دارد، بدون آنکه مجبور باشیم y را بدست آوریم، از اهمیت ویژه‌ای برخوردار است.

برای آشنایی بیشتر با مسئله، تابع $F(x,y) = x^{2} + y^{2} - 1$ را در نظر می‌گیریم. به‌دنبال x و y هایی هستیم که در $F(x,y) = 0$ صدق می‌کنند، که دایره واحد است. تابع $f(x)$

جواب است اگر و فقط اگر برای هر x از حوزه تعریف f داشته باشیم، $F(x,f(x))=0$. در این صورت بدیهی است که $f(x)=\pm\sqrt{1-x^2}$ و از این رو ممکن است منحصر به فرد نباشد. علاقه‌مندیم ببینیم برای (x_0,y_0) های دلخواهی که در $F(x_0,y_0)=0$ صدق می‌کنند، آیا می‌توان f را چنان یافت که $f(x_0,f(x_0))=0$ و f در نزدیکی (x_0,y_0) منحصر به فرد و مشتق‌پذیر باشد. اگر $x_0\neq\pm1$ چنین f ای وجود دارد؛ کافی است ریشه دوم را در نظر بگیریم.

در دو نقطهٔ $x_0=\pm1$ تابع f منحصر به فرد نیست و می‌توان ریشه دوم مناسب (اگر $y_0>0$ علامت مثبت و اگر $y_0<0$ علامت منفی) را در نظر بگیریم. در این دو نقطهٔ استثنایی داریم $\frac{\partial F}{\partial y}=0$. بنابراین در حالت کلی شرطی نظیر $\frac{\partial F}{\partial y}\neq 0$ لازم داریم تا بتوانیم وجود (لااقل موضعی) تابع منحصر به فرد مشتق‌پذیر f را که در $F(x,f(x))=0$ صدق می‌کند، تضمین کنیم. در حالت کلی تابعی به‌صورت $F:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ خواهیم داشت.

معادله $F(x,y)=0$ را می‌توان به‌صورت

$$\begin{gathered}F_1(x_1,\ldots,x_n,y_1,\ldots,y_m)=0\,,\\ \vdots\\ F_m(x_1,\ldots,x_n,y_1,\ldots,y_m)=0\,,\end{gathered}$$

نوشت. می‌خواهیم m مجهول y_1 ، ... و y_m را برحسب x_1 ، ... و x_n بدست آوریم. برای سادگی در بیان قضیه، فرض کنیم A_y ماتریس

$$A_y=\begin{bmatrix}\frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m}\\ \vdots & \ddots & \vdots\\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m}\end{bmatrix}$$

محاسبه شده در نقطهٔ $(a,b)\in\mathbb{R}^n\times\mathbb{R}^m$ باشد و دترمینان آن در این نقطه را به Δ نشان

می‌دهیم، یعنی

$$\Delta = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{vmatrix},$$

محاسبه شده در نقطهٔ (a,b).

قضیه ۱.۲.۳ (قضیه تابع ضمنی)**.** فرض کنیم $E \subseteq \mathbb{R}^n \times \mathbb{R}^m$ باز، $F : E \to \mathbb{R}^m$ تابعی C^1 و $(a,b) \in E$ و $F(a,b) = 0$. همچنین فرض کنیم A_y معکوس‌پذیر باشد (که معادل است با $\Delta \neq 0$). در این صورت همسایگی U از a در $\mathbb{R}^n$، همسایگی V از $\mathbb{R}^m$ و تابع منحصر به فرد $f : U \to V$، به‌قسمی موجودند که برای هر $x \in U$، $F(x, f(x)) = 0$. به‌علاوه f، C^1 است.

برهان. تابع $G : E \to \mathbb{R}^n \times \mathbb{R}^m$ را به‌صورت $G(x,y) = (x, F(x,y))$ تعریف می‌کنیم. F بنا به فرض C^1 است و چون نگاشت همانی تا هر مرتبه‌ای مشتق‌پذیر است. پس G، نگاشتی C^1 است. ماتریس مشتقات جزئی G (ماتریس ژاکوبی) عبارت است از

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots & \\ 0 & & \cdots & 1 & 0 & \cdots & 0 \\ \frac{\partial F_1}{\partial x_1} & & \cdots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & & & & & \vdots \\ \frac{\partial F_m}{\partial x_1} & & \cdots & \frac{\partial F_m}{\partial y_n} & \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}.$$

دترمینان این ماتریس محاسبه شده در نقطه (a,b) برابر است با

$$\Delta = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{vmatrix}.$$

بنابراین باتوجه به فرض $JG(a,b) \neq 0$. پس بنا به قضیه تابع معکوس مجموعهٔ باز S شامل (a,b) و مجموعه باز $(a,0)$ وجود دارد، به‌قسمی که $G(S) = W$. G دارای معکوس $G^{-1}: S \to W$ است که C^1 می‌باشد. چون S مجموعهٔ باز شامل (a,b)است، مجموعه‌های باز $U \subseteq \mathbb{R}^n$ و $V \subseteq \mathbb{R}^m$ را چنان اختیار می‌کنیم که $a \in U$ و $b \in V$ و $U \times V \subset S$. فرض کنیم $G(U \times V) = N \subset W$. در این صورت $G: U \times V \to N$ نگاشتی C^1 است که دارای معکوس $G^{-1}: N \to U \times V$ می‌باشد که G^{-1} نیز C^1 است. پس باتوجه به تعریف G، نگاشت G^{-1} به‌صورت $G^{-1}(x,w) = (x, H(x,w))$ می‌باشد که H نگاشتی C^1 از N به V است. فرض کنیم $\Pi_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ نگاشت تصویر روی مؤلفهٔ دوم باشد یعنی $\Pi_2(x,y) = y$. در این صورت

$$F(x, H(x,w)) = \Pi_2 \circ G(x, H(x,y)) = \Pi_2 \circ G \circ G^{-1}(x,w) = w \,.$$

همچنین توجه می‌کنیم که چون G^{-1} به‌صورت $G^{-1}(x,w) = (x, H(x,w))$ می‌باشد. پس اگر $(x,w) \in N$ آن‌گاه $x \in U$. نگاشت $f: U \to V$ را به‌صورت $f(x) = H(x,0)$ تعریف می‌کنیم. در این صورت چون $F(x, H(x,w)) = w$ داریم $F(x, f(x)) = 0$. به‌علاوه f از ردهٔ C^1 است. زیرا H از رده C^1 است. همچنین بنا به قضیه تابع معکوس، $H(x,w)$ منحصر به فرد است و در نتیجه $f(x) = H(x,0)$ منحصر به فرد بوده و به این ترتیب برهان کامل می‌شود. □

نتیجه ۳. در قضیه ۱.۲.۳ از رابطهٔ $\dfrac{\partial f_j}{\partial x_i}$

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_n} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix} \tag{۲.۳}$$

به‌دست می‌آید که (-1) در بالای ماتریس به معنی معکوس ماتریس است.

برهان. با علایم قضیهٔ تابع ضمنی فرض کنیم $\phi(x) = (x, f(x))$ و فرض کنیم $A = DF(\phi(a))$، یعنی A برابر است با ماتریس زیر در نقطه $\phi(a)$

$$A = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial y_n} & \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix} .$$

باتوجه به خطی بودن A برای $k \in \mathbb{R}^n$ و $h \in \mathbb{R}^m$ داریم

$$A(h, k) = A(k, 0) + A(0, h) .$$

حال مشتق عبارت $F(\phi(x)) = F(x, f(x)) = 0$ را در a به کمک قاعدهٔ زنجیری حساب می‌کنیم. خواهیم داشت

$$AD\phi(x) = D(F \circ \phi)(a) = 0 .$$

ولی باتوجه به تعریف ϕ، $D\phi(x)(k) = \big(k, Df(x)(k)\big)$. پس

$$A(k, 0) + A(0, Df(a)(k)) = A(k, Df(a)k) = 0 .$$

این معادل است با

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix} + \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_n} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = 0$$

که همان رابطه (۲.۳) است. □

مثال ۲.۲.۳. دستگاه معادلات

$$xu + yv^2 = 0$$

$$xv^{۲} + y^{۲}u^{۶} = ۰$$

را درنظر بگیرید. آیا می‌توان u و v را به طور منحصر به فرد برحسب x و y در نزدیکی نقطهٔ $x = ۰$، $y = ۱$، $u = ۰$، $v = ۰$ حل کرد؟

حل. در اینجا با علایم قضیهٔ تابع ضمنی، سمت چپ معادلات عبارت است از تابع F و معادله اول و دوم به ترتیب برابرند با $F_۱$ و $F_۲$. به علاوه $x_۱ = x$، $x_۲ = y$، $y_۱ = u$ و $y_۲ = v$. می‌خواهیم ببینیم آیا می‌توان $u(x, y)$ و $v(x, y)$ را به‌دست آورد. بنابراین دترمینان

$$\Delta = \begin{vmatrix} \frac{\partial F_۱}{\partial u} & \frac{\partial F_۱}{\partial v} \\ \frac{\partial F_۲}{\partial u} & \frac{\partial F_۲}{\partial v} \end{vmatrix} = \begin{vmatrix} x & ۲yx \\ ۶y^{۲}u^{۵} & ۳xv^{۲} \end{vmatrix} = ۳x^{۲}v^{۲} - ۱۲y^{۳}u^{۵}v$$

را تشکیل می‌دهیم که در نقطه مفروض برابر ۵ است. پس بنابر قضیهٔ تابع ضمنی می‌توان u و v را بر حسب x و y به‌صورت منحصر به فرد به‌دست آورد.

مثال ۳.۲.۳. تابع $F : \mathbb{R}^{۳} \times \mathbb{R}^{۲} \to \mathbb{R}^{۲}$ را که به‌وسیله معادلات

$$۲e^{u} + vx - ۴y + ۳ = ۰$$
$$v\cos u - ۶u + ۲x - z = ۰$$

تعریف شده است، در نظر می‌گیریم. در صورتی که $a = (۳, ۲, ۷)$ و $b = (۰, ۱)$ آیا می‌توان u و v را به طور منحصر به فرد بر حسب x و y در نقطهٔ (a, b) به‌دست آورد؟ به‌علاوه $\frac{\partial u}{\partial x}$ را در صورت وجود محاسبه می‌کنیم.

حل. در اینجا $F = (F_۱, F_۲)$ که $F_۱$ و $F_۲$ به ترتیب برابر معادله اول و دوم می‌باشند. به‌علاوه $x_۱ = x$، $x_۲ = y$، $x_۳ = z$، $y_۱ = u$ و $y_۲ = v$. می‌خواهیم ببینیم آیا u و v

برحسب x و y در نزدیکی نقطهٔ $x_1 = ۳$، $x_2 = ۲$، $x_3 = ۷$، $u = ۰$ و $v = ۱$ قابل حل است. دترمینان

$$\Delta = \begin{vmatrix} \frac{\partial F_1}{\partial u} & \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial u} & \frac{\partial F_2}{\partial v} \end{vmatrix} = \begin{vmatrix} ۲e^u & x \\ -v\cos u - ۶ & \cos u \end{vmatrix} = \begin{vmatrix} ۲ & ۳ \\ -۶ & ۱ \end{vmatrix} = ۲۰$$

در نقطه مفروض مخالف صفر است و در نتیجه قضیه ضمنی وجود تابع $f \in C^1$ را در یک همسایگی نقطهٔ $(۳, ۲, ۷)$ ایجاب می‌کند. یعنی توابع $u(x,y,z)$ و $v(x,y,z)$، به طور منحصر به فرد موجودند. حال باتوجه به نتیجه ۳ داریم

$$Df(۳,۲,۷) = -\frac{۱}{۲۰}\begin{bmatrix} ۱ & -۳ \\ ۶ & ۲ \end{bmatrix}\begin{bmatrix} ۱ & -۴ & ۰ \\ ۲ & ۰ & ۱ \end{bmatrix} = \begin{bmatrix} \frac{۱}{۴} & \frac{۱}{۵} & \frac{-۳}{۲۰} \\ \frac{-۱}{۲} & \frac{۶}{۵} & \frac{۱}{۱۰} \end{bmatrix}.$$

به این ترتیب مشتقات جزئی u و v نسبت به x و y به دست می‌آید

$$\frac{\partial u}{\partial x} = \frac{۱}{۴}, \qquad \frac{\partial u}{\partial y} = \frac{۱}{۵}, \qquad \frac{\partial u}{\partial z} = -\frac{۳}{۲۰},$$
$$\frac{\partial v}{\partial x} = -\frac{۱}{۲}, \qquad \frac{\partial v}{\partial y} = \frac{۶}{۵}, \qquad \frac{\partial v}{\partial z} = -\frac{۱}{۱۰}.$$

۳.۳ قضیه رتبه

در این قسمت یکی دیگر از قضایای مهم آنالیز را اثبات می‌کنیم. قبل از طرح قضیه بیان چند تعریف در تسهیل امر مؤثر است. در جبر خطی رتبه یک ماتریس را به یکی از سه صورت معادل زیر تعریف می‌کنیم:

(الف) بعد زیرفضای تولید شده توسط سطرها،

(ب) بعد زیرفضای تولید شده توسط ستون‌ها،

(پ) ماکزیمم مرتبهٔ یک کهاد دترمینان مخالف صفر از آن.

تعریف ۱.۳.۳. رتبه یک تبدیل خطی عبارت از بعد تصویر آن است. رتبهٔ تبدیل خطی A را با $\mathrm{rank}(A)$ نشان می‌دهیم.

می‌توان نشان داد که رتبهٔ یک تبدیل خطی برابر است با رتبهٔ هر ماتریسی که نمایشگر آن تبدیل خطی باشد.

تعریف ۲.۳.۳. فرض کنیم $U \subseteq \mathbb{R}^n$ و $V \subseteq \mathbb{R}^m$ باز هستند. نگاشت $F : U \longrightarrow V$ را یک همان‌سانی می‌نامیم هرگاه F یک‌به‌یک، پوشا، پیوسته و با معکوس پیوسته باشد.

تعریف ۳.۳.۳. فرض کنیم $U \subseteq \mathbb{R}^n$ و $V \subseteq \mathbb{R}^m$ باز هستند. نگاشت $F : U \to V$ را یک وابرسانی خوانیم، در صورتی‌که اولاً F همان‌سانی باشد، ثانیاً F و F^{-1} هر دو C^1 باشند.

قضیه ۴.۳.۳ (قضیه رتبه)**.** فرض کنیم $A_0 \subseteq \mathbb{R}^n$ و $B_0 \subseteq \mathbb{R}^m$ مجموعه‌هایی باز بوده و همچنین $F : A_0 \longrightarrow B_0$ نگاشتی C^1 باشد. فرض کنیم رتبه DF روی A_0 برابر k باشد. اگر $a \in A_0$ و $b = F(a)$، آن‌گاه مجموعه‌های باز $A \subseteq A_0$ و $B \subseteq B_0$ که $a \in A$ و $b \in B$ و وابرسانی‌های $G : A \to U \subset \mathbb{R}^n$ و $H : B \to V \subset \mathbb{R}^m$ (U و V باز هستند) وجود دارند به‌قسمی که $H \circ F \circ G^{-1}(U) \subset V$ و

$$H \circ F \circ G^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0) \, .$$

برهان. همانند قضیه تابع معکوس اگر در حالتی که $a = 0 \in \mathbb{R}^n$ و $b = 0 \in \mathbb{R}^m$، قضیه را ثابت کنیم، حالت کلی نتیجه خواهد شد. زیرا فرض کنیم $\tilde{F}(u) = F(u + a) - b$ دراین‌صورت $\tilde{F}(0) = 0$ و اگر $\tilde{G}$ و $\tilde{H}$ی که در حکم قضیه صدق کنند، موجود باشند کافی است $G^{-1} = \tilde{G}^{-1} + a$ و $H = \tilde{H}(\cdot - b)$ اختیار شوند. به‌علاوه به کمک تغییر اندیس‌ها (در صورت لزوم) می‌توان فرض کرد که کهاد دترمینان $k \times k$ مخالف صفر در $DF(a)$ عبارت

است از

$$\begin{vmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_k} \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial u_1} & \cdots & \frac{\partial f_k}{\partial u_k} \end{vmatrix}.$$

نگاشت $G : A_0 \longrightarrow \mathbb{R}^n$ را که یک نگاشت C^1 است، به‌صورت

$$G(u_1, \ldots, u_n) = \Big(f_1(u_1, \ldots, u_n), \ldots, f_k(u_1, \ldots, u_n), u_{k+1}, \ldots, u_n\Big)$$

تعریف می‌کنیم. دراین‌صورت ماتریس مشتق های جزئی G عبارت است از

$$\Delta G = \left[\begin{array}{ccc|c} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_k} & \\ \vdots & \ddots & \vdots & * \\ \frac{\partial f_k}{\partial u_1} & \cdots & \frac{\partial f_k}{\partial u_k} & \\ \hline & 0 & & I_{n-k} \end{array}\right]$$

که I_{n-k} ماتریس همانی $(n-k) \times (n-k)$ می‌باشد. جملات قسمت پایینی چپ همگی صفرند و قسمت بالایی راست مورد توجه ما نیست. این ماتریس در $x = a$ معکوس‌پذیر است. پس بنابر قضیه تابع معکوس، زیرمجموعه باز A_1 از A_0 شامل a وجود دارد که G روی آن یک وابرسانی بر روی مجموعه باز $U_1 = G(A_1)$ است. باتوجه به تعریف G داریم $F \circ G(0) = 0$، $F \circ G^{-1}(U_1) \subseteq B_0$ و

$$F \circ G^{-1}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = \big(x_1, \ldots, x_k, \overline{f}_{k+1}(x), \ldots, \overline{f}_m(x)\big)$$

که $\overline{f}_{k+j}(x) = f_{k+j} \circ G^{-1}(x)$ زیرا G^{-1} روی U_1 یک‌به‌یک است و برای $l = 1, \ldots, m$ داریم $x_l = f_l(u)$ و $u \in G^{-1}(u_1)$ پس $f_l \circ G^{-1}(x) = f_l(u) = x_l$. تا اینجا فقط از اینکه رتبهٔ Df در نقطهٔ a (و در نتیجه در یک همسایگی آن) برابر k است استفاده کرده‌ایم

و از این حقیقت که رتبهٔ Df روی $A_\circ$ همه جا برابر k است بهره‌ای نبرده‌ایم ولی این مطلب را در مرحله بعد مورد استفاده قرار می‌دهیم. ماتریس مشتق $D(F\circ G^{-\text{۱}})$ را از روی فرمول $F\circ G^{-\text{۱}}$ حساب می‌کنیم خواهیم داشت

$$\Delta(F\circ G^{-\text{۱}})(x)=\left[\begin{array}{c|ccc} I_k & & \circ & \\ \hline & \frac{\partial\bar f_{k+\text{۱}}}{\partial x_n} & \cdots & \frac{\partial\bar f_{\text{۱}}}{\partial x_{k+\text{۱}}} \\ * & \vdots & \ddots & \vdots \\ & \frac{\partial\bar f_m}{\partial x_n} & \cdots & \frac{\partial\bar f_m}{\partial x_{k+\text{۱}}} \end{array}\right].$$

این تساوی روی $U_\text{۱}$ که به‌صورت $F\circ G^{-\text{۱}}$ تعریف شده است، برقرار است. از طرفی $DG^{-\text{۱}}$ روی $U_\text{۱}$ معکوس‌پذیر است و $G^{-\text{۱}}(u_\text{۱})=A_\text{۱}\subset A_\circ$. بنابراین روی $u_\text{۱}$

$$\operatorname{rank}\Big(D(F\circ G^{-\text{۱}})\Big)=\operatorname{rank}\Big(DF\circ DG^{-\text{۱}}\Big)=k$$

که نتیجه می‌دهد تمامی جملات قسمت پایینی راست ماتریس بالا روی $U_\text{۱}$ برابر صفر است. یعنی توابع $\bar f_{k+\text{۱}}$ ، ... و $\bar f_m$ فقط به $x_\text{۱}$ ، $\cdots$ و x_k وابسته‌اند. حال نگاشت T از همسایگی $V_\text{۱}$ از $\circ$ در $\mathbb{R}^m$ را به‌توی $B_\circ\subset\mathbb{R}^m$ به‌صورت

$$\begin{aligned}T(y_\text{۱},\ldots,y_k,y_{k+\text{۱}},\ldots,y_m)=\big(&y_\text{۱},\ldots,y_k,y_{k+\text{۱}}+\bar f_{k+\text{۱}}(y_\text{۱},\ldots,y_k),\\&\ldots,y_m+\bar f_m(y_\text{۱},\ldots,y_k)\big)\end{aligned}$$

تعریف می‌کنیم. مجموعهٔ $V_\text{۱}$ را بایستی آن‌قدر کوچک اختیار کرد که اولاً برای $y=(y_\text{۱},\ldots,y_m)\in V_\text{۱}$ توابع $\bar f_{k+j}(y_\text{۱},\ldots,y_k)$ تعریف شوند و ثانیاً $T(V_\text{۱})\subset B_\circ$. باتوجه به $F(\circ)=\circ$ واضح است که $T(\circ)=\circ$. اگر DT را محاسبه کنیم در هر نقطه $y\in V_\text{۱}$ داریم

$$DT(y)=\left[\begin{array}{c|c} I_k & \circ \\ \hline \circ & I_{m-k}\end{array}\right].$$

پس T از یک همسایگی V_0 در V_1 به‌روی مجموعهٔ بازی مانند $B \subset \mathbb{R}^m$ وابرسانی است که در آن B شامل مبدأ $\mathbb{R}^m$ بوده و $B \subset B_0$. همسایگی $U \subset U_1$ از مبدا در $\mathbb{R}^n$ را در نظر می‌گیریم به‌قسمی که $F \circ G^{-1}(u) \subset B$. فرض کنیم $A = G^{-1}(u)$ و $H = T^{-1}$.
در این صورت نگاشت‌های $U \xrightarrow{G^{-1}} A \xrightarrow{F} B \xrightarrow{H} V$ همگی C^1 می‌باشند و G^{-1} و H وابرسانی‌هایی بروی A و V هستند. به‌علاوه چون

$$F \circ G^{-1}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = \big(x_1, \ldots, x_k, \bar{f}_{k+1}(x_1, \ldots, x_k), \ldots, \bar{f}_m(x_1, \ldots, x_k)\big) ,$$

پس

$$H \circ F \circ G^{-1}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0) .$$

زیرا T نقطه $(x_1, \ldots, x_k, 0, \ldots, 0)$ را به نقطه

$$\big(x_1, \ldots, x_k, \bar{f}_{k+1}(x_1, \ldots, x_k), \ldots, \bar{f}_m(x_1, \ldots, x_k)\big)$$

می‌برد و به این ترتیب برهان قضیه کامل می‌شود. $\square$

۴.۳ تمرین

۱. فرض کنیم $A \subseteq \mathbb{R}^n$ یک مجموعه باز و $f : A \longrightarrow \mathbb{R}^n$ یک تابع به طور پیوسته مشتق‌پذیر و یک‌به‌یک باشد، به‌نحوی که برای هر $x \in A$، $\det f'(x) \neq 0$. نشان دهید $f(A) \subseteq \mathbb{R}^n$ باز و $f^{-1} : f(A) \longrightarrow A$ مشتق‌پذیر است.

۲. فرض کنیم $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ یک تابع به‌طور پیوسته مشتق‌پذیر باشد. نشان دهید f نمی‌تواند یک‌به‌یک باشد.

۳. فرض کنیم $f:\mathbb{R}\longrightarrow\mathbb{R}$ در شرط $f'(a)\neq 0$ برای هر $a\in\mathbb{R}$ صادق باشد. نشان دهید f یک‌به‌یک است.

۴. فرض کنیم $f:\mathbb{R}\times\mathbb{R}\longrightarrow\mathbb{R}$ مشتق‌پذیر باشد. برای هر $x\in\mathbb{R}$ تعریف می‌کنیم

$$g_x:\mathbb{R}\longrightarrow\mathbb{R}\,,\quad g_x(y)=f(x,y).$$

فرض کنیم برای هر x، عضو منحصر به فرد y چنان موجود است که $g'_x(y)=0$. این y را $c(x)$ می‌نامیم. در این صورت اگر $D_{2,2}f(x,y)\neq 0$ آن‌گاه برای هر (x,y) نشان دهید c مشتق‌پذیر است و

$$c'(x)=-\frac{D_{2,1}f\big(x,c(x)\big)}{D_{2,2}f\big(x,c(x)\big)}\ .$$

فصل ۴

انتگرال‌گیری چندمتغیره

۱.۴ توابع انتگرال‌پذیر

در ریاضیات عمومی انتگرال ریمان را برای توابع کران‌دار $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$ دیده‌ایم. اکنون می‌خواهیم ببینیم در حالتی که $[a,b]$ با زیرمجموعه‌ای از $\mathbb{R}^n$ جای‌گزین گردد، انتگرال به چه صورتی تعریف می‌شود.

فرض کنیم A زیرمجموعه‌ای کران‌دار از $\mathbb{R}^n$ باشد و $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ تابعی کران‌دار باشد. فرض کنیم $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ مستطیلی در $\mathbb{R}^n$ باشد که A را شامل است. تابع f را به تمام B توسیع می‌دهیم به این ترتیب که اگر $x \notin A$ قرار می‌دهیم $f(x) = 0$. یک افراز از مستطیل $[a_1, b_1] \times \cdots \times [a_n, b_n]$ عبارت است از گردایهٔ $P = \{P_1, \cdots, P_n\}$ که هر P_i افرازی از بازهٔ $[a_i, b_i]$ است. فرض کنیم $P_i = \{x^i_0, \ldots, x^i_{m_i}\}$. در این صورت افراز P از تعداد $m_1 \times m_2 \times \cdots \times m_n$ مستطیل به‌صورت

$$[x^1_{j_1}, x^1_{j_1+1}] \times \cdots \times [x^n_{j_n}, x^n_{j_n+1}] , \qquad 1 \le j_i \le m_i$$

تشکیل گردیده است. برای مستطیل $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ حجم مستطیل را به‌صورت

$$V(B) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$$

تعریف می‌کنیم. اگر S یکی از مستطیل‌های افراز باشد، قرار می‌دهیم:

$$m_S(f) = \inf\{f(x) | x \in S\} , \qquad M_S(f) = \sup\{f(x) | x \in S\} .$$

و حاصل جمع پایینی و بالایی f نسبت به افراز P عبارتند از

$$L(f, P) = \sum_{S \in P} m_S(f) V(S) , \qquad U(f, P) = \sum_{S \in P} M_S(f) V(S) .$$

از تعریف بدیهی است که برای هر افراز P

$$L(f, P) \leq U(f, P) .$$

فرض کنیم P' افراز ظریف‌تری از P باشد (یعنی هر زیر مستطیل P' در یک زیر مستطیل P جای داشته باشد). اگر مستطیل S از P متشکل از مستطیل‌های $S_1, \cdots, S_k$ باشد، اولاً

$$V(S) = V(S_1) + \cdots + V(S_k) ,$$

ثانیاً چون برای $1 \leq i \leq k$، $S_i \subset S$ پس

$$\inf\{f(x) \mid x \in S_i\} \geq \inf\{f(x) \mid x \in S\} .$$

بنابراین نتیجه می‌شود که $L(f, P) \leq L(f, P')$ و به طریق مشابه $U(f, P') \leq U(f, P)$.
از این جا می‌توان لم زیر را نتیجه گرفت.

لم ۱.۱.۴. اگر P و P' دو افراز دلخواه از مستطیل $B \subseteq \mathbb{R}^n$ باشند، آن‌گاه $L(f,P) \leq U(f,P')$.

برهان. فرض کنید P'' افرازی از B باشد که ظریف‌تر از P و P' است. در این صورت

$$L(f,P) \leq L(f,P'') \leq U(f,P'') \leq U(f,P') \ .$$

□

از این لم نتیجه می‌گردد که مجموعهٔ $\{L(f,P) |$ افرازی ازمستطیل B است $P\}$ از بالا کران‌دار و بنا بر اصل تمامیت اعداد حقیقی، دارای سوپریمم است.

همچنین مجموعهٔ $\{U(f,P) |$ افرازی ازمستطیل B است $P\}$ از پایین کران‌دار و در نتیجه دارای اینفیمم است.

تعریف ۲.۱.۴. با علایم فوق انتگرال بالایی f عبارت است از

$$\inf\left\{U(f,P) \mid P \text{ افرازی از } B\right\} ,$$

و با $\overline{\int}_A f$ نشان داده می‌شود و انتگرال پایینی f که با $\underline{\int}_A$ نشان داده می‌شود، عبارت است از

$$\sup\left\{L(f,P) \mid P \text{ افرازی از } B\right\} .$$

در صورتی که $\underline{\int}_A f = \overline{\int}_A f$ گوییم که f روی A انتگرال‌پذیر ریمان است و مقدار مشترک انتگرال پایینی و بالایی را انتگرال f روی A خوانیم و آن را با یکی از نمادهای $\int_A f$، $\int_A f(x)\,\mathrm{d}x$ و یا $\int_A f(x_1,\cdots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n$ نشان می‌دهیم.

قضیه ۳.۱.۴ (شرط ریمان)**.** فرض کنیم $A \subset \mathbb{R}^n$ کران‌دار بوده و در مستطیل S واقع باشد. تابع کران‌دار $f : A \to \mathbb{R}$ انتگرال‌پذیر است، اگر و فقط اگر برای هر $\epsilon > ۰$ افرازی مانند P از S وجود داشته باشد به‌قسمی که $U(f,P) - L(f,p) < \epsilon$.

برهان. فرض کنیم f انتگرال‌پذیر است و فرض کنیم $\epsilon > ۰$ داده شده است. چون

$$\int_A f = \inf\{U(f,p) \mid \text{افرازی از } S \text{ است } P\}\,,$$

پس افرازی مانند $P_۱$ از S وجود دارد که

$$U(f,P_۱) < \int_A f + \frac{\epsilon}{۲}\ .$$

و به‌دلیل مشابه افرازی مانند $P_۲$ از S وجود دارد که

$$\int_A f - \frac{\epsilon}{۲} < L(f,P_۲)\ .$$

اگر P افرازی ظریف‌تر از $P_۱$ و $P_۲$ باشد، باتوجه به آن‌چه قبل از لم ۱.۱.۴ بیان شد، خواهیم داشت

$$U(f,P) - L(f,p) < \epsilon\ .$$

برعکس، اگر شرط ریمان برقرار باشد $\sup\{L(f,P)\}=\inf\{U(f,P)\}$، یعنی f انتگرال‌پذیر است (بررسی این طرف قضیهٔ را به‌عنوان تمرین واگذار می‌کنیم!). □

قبل از ادامه بحث دو مثال ذکر می‌کنیم که اولی انتگرال‌پذیر است ولی دومی انتگرال پذیر نیست.

مثال ۴.۱.۴. فرض کنید $A \subseteq \mathbb{R}^n$ یک مستطیل و $f : A \to \mathbb{R}$ تابع ثابت باشد یعنی $f(x) = c$. اگر P افراز دلخواهی از A باشد آن‌گاه در هر مستطیل S_i داریم $m_{S_i}(f) = M_{S_i}(f) = c$. پس

$$L(f,P) = U(f,P) = \sum_{s_i} cV(S_i) = cV(A) \ .$$

مثال ۵.۱.۴. تابع $f : [۰,۱] \times [۰,۱] \to \mathbb{R}$ را که به‌صورت

$$f(x,y) = \begin{cases} ۰ & x \text{ گویا است} \\ ۱ & x \text{ گنگ است} \end{cases}$$

تعریف شده است، در نظر بگیرید. اگر P افرازی از $[۰,۱] \times [۰,۱]$ باشد، آن‌گاه به‌دلیل چگال بودن اعداد گویا در اعداد حقیقی، هر زیر مستطیل S_i نقاط (x,y)ی با x گویا و همچنین نقاط (x,y)ای با x گنگ را شامل است. پس $m_s = ۰$ و $M_s = ۱$. در نتیجه

$$L(f,P) = \sum_s ۰V(s) = ۰ \ ,$$
$$U(f,P) = \sum_s ۱V(s) = V\big([۰,۱] \times [۰,۱]\big) = ۱ \ .$$

پس f انتگرال‌پذیر نیست.

قضیهٔ زیر که نظیر آن را برای توابعی از $\mathbb{R}$ به $\mathbb{R}$ دیده‌ایم، بدون اثبات ارائه می‌شود. خوانندهٔ علاقه‌مند برای دیدن برهان می‌تواند به [۲] مراجعه کند.

قضیه ۶.۱.۴ (قضیهٔ داربو)**.** فرض کنیم $A \subset \mathbb{R}^n$ کران‌دار باشد و در مستطیل S جای داشته باشد. فرض کنیم $f : A \to \mathbb{R}$ کران‌دار است و آن را به S توسیع می‌دهیم. به این ترتیب که در خارج A تعریف می‌کنیم $f = ۰$. در این صورت f انتگرال‌پذیربا انتگرالی برابر I است، اگر و فقط اگر برای هر $\epsilon > ۰$ عددی مانند $\delta > ۰$ وجود داشته باشد به‌قسمی

که اگر P افرازی دلخواه به مستطیل‌های $S_۱, \cdots, S_N$ با اضلاع با طول کمتر از δ باشد و اگر $x_۱ \in S_۱, \cdots, x_n \in S_N$، آن‌گاه

$$\left|\sum_{i=۱}^{N} f(x_i)V(S_i) - I\right| < \epsilon \ .$$

مجموع $\sum_{i=۱}^{N} f(x_i)V(S_i)$ را مجموع ریمان خوانیم.

۲.۴ اندازهٔ صفر و محتوای صفر

برای توابعی که بر خط حقیقی تعریف شده باشند، معمولاً انتگرال روی یک بازه گرفته می‌شود. ولی در $\mathbb{R}^n$ انتگرال‌گیری بر روی مجموعه‌های پیچیده‌تری انجام خواهد گرفت. درواقع لازم است این مجموعه‌ها به گونه‌ای باشند که افرازی که در تعریف انتگرال‌گیری ذکر شد، به شکل «معقولی» باشد. به طوری که خواهیم دید معقول بودن به این معناست که بایستی این مجموعه‌ها دارای مرز خیلی «پیچیده» نباشند. به زودی معنا و مفهوم دقیق کلمات کیفی عرضه‌شده در سطور بالا ارائه خواهد شد.

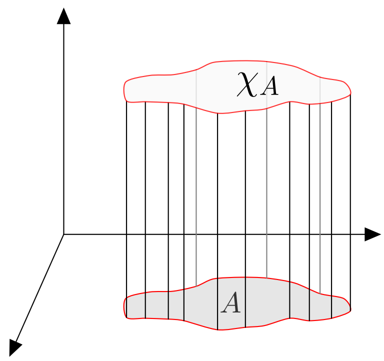
برای بیان دقیق این مطلب احتیاج به مقدماتی داریم.

تعریف ۱.۲.۴. اگر $A \subseteq \mathbb{R}^n$ آن‌گاه تابع مشخصهٔ A عبارت است از

$$\chi_A : \mathbb{R}^n \to \mathbb{R}^n; \qquad \chi_A(x) = \begin{cases} ۱ & x \in A \text{ اگر} \\ ۰ & x \neq A \text{ اگر} \end{cases}$$

گوییم که A دارای حجم است، اگر χ_A انتگرال‌پذیر باشد و در این صورت حجم A عبارت است از عدد

$$\int_A \chi_A(x)\,\mathrm{d}x = V(A).$$



از نظر شهودی این تعریف منطقی است زیرا همان‌گونه که شکل نشان می‌دهد ناحیهٔ زیر نمودار χ_A یک استوانه‌گون به ارتفاع ۱ و قاعدهٔ A است.
بدیهی است که در حالت $n = 1$ برای $A \subset \mathbb{R}$ حجم $V(A)$ برابر است با طول A و برای $A \subset \mathbb{R}^2$ مساحت A برابر است با $V(A)$.

تعریف ۲.۲.۴. گوییم خانوادهٔ $\{A_i\}_{i=1}^{\infty}$ از زیرمجموعه‌های $\mathbb{R}^n$ یک پوشش برای مجموعهٔ $A \subseteq \mathbb{R}^n$ است، هرگاه $A \subseteq \cup_{i=1}^{\infty} A_i$.

تعریف ۳.۲.۴. مجموعهٔ $A \subset \mathbb{R}^n$ دارای اندازهٔ صفر است، در صورتی که برای هر $\epsilon > 0$ پوششی از A با تعداد شمارا مستطیل مانند $S_1, S_2, \cdots$ وجود داشته باشد، به قسمی که $\sum_{i=1}^{\infty} V(S_i) < \epsilon$.

همچنین گوییم مجموعهٔ A دارای محتوای صفر است، در صورتی که برای هر $\epsilon > 0$ پوششی از A به وسیلهٔ تعداد متناهی مستطیل مانند $S_1, \cdots, S_N$ وجود داشته باشد، به طوری که $\sum_{i=1}^{n} V(S_i) < \epsilon$.

بدیهی است که اگر A دارای اندازهٔ (محتوای صفر) باشد و $B \subset A$، آن‌گاه B نیز

دارای اندازهٔ (محتوای صفر) است. بنابراین یک مجموعه با اندازهٔ صفر نمی‌تواند مستطیلی نابدیهی را شامل باشد. یعنی مستطیلی به‌صورت $[a_۱, b_۱] \times \cdots \times [a_n, b_n]$ به‌قسمی که $a_i < b_i$. به‌علاوه، هر مجموعهٔ متشکل از تعداد شمارا نقطه، دارای اندازهٔ صفر است زیرا اگر $A = \{a_۱, a_۲, a_۳, \cdots\}$ کافی است S_i را مستطیلی شامل a_i با $V(S_i) < \frac{\epsilon}{۲^i}$ اختیار کنیم در این صورت

$$\sum_{i=۱}^{\infty} V(S_i) < \sum_{i=۱}^{\infty} \frac{\epsilon}{۲^i} = \epsilon \ .$$

از این جا نتیجه می‌گردد که مجموعهٔ اعداد گویا دارای اندازهٔ صفر است.

می‌توان نشان داد که زیرمجموعهٔ $A \subset \mathbb{R}^n$ دارای اندازهٔ صفر است اگر و تنها اگر $V(A) = ۰$. بررسی این گزاره حالت خاصی از قضیهٔ ۱.۲.۴ است که در ادامه می‌آید. باتوجه به این مطلب، گاهی اوقات به جای محتوای صفر عبارت «حجم صفر» نیز به کار برده می‌شود.

مثال ۴.۲.۴. نشان می‌دهیم که $\mathbb{R}$ به‌عنوان زیرمجموعه‌ای از $\mathbb{R}^۲$ دارای اندازهٔ صفر است. ولی به‌عنوان زیرمجموعه‌ای از $\mathbb{R}$ اندازهٔ صفر ندارد.

حل. فرض کنیم $۰ < \epsilon$ داده شده است. به‌دنبال مستطیل‌های $S_۱, S_۲, \cdots$ می باشیم که محور x ها را بپوشاند و مجموع مساحت‌های آنها کمتر از ϵ باشد. فرض کنیم برای هر عدد طبیعی i،

$$S_i = [-i, i] \times \left[-\frac{\epsilon}{۳i \times ۲^{i+۱}}, \frac{\epsilon}{۳i \times ۲^{i+۱}}\right] .$$

در این صورت $V(S_i) = (۲i).\frac{۲\epsilon}{۳i \times ۲^{i+۱}} < \frac{\epsilon}{۲^i}$ و بنابراین

$$\sum_{i=۱}^{\infty} V(S_i) < \sum_{i=۱}^{\infty} \frac{\epsilon}{۲^i} = \epsilon \ .$$

این ادعای اول را ثابت می‌کند. برای اثبات ادعای دوّم کافی است توجه کنیم که اگر $\mathbb{R}$ را با بازه‌ها بپوشانیم، مجموع طول این بازه‌ها به $+\infty$ میل می‌کند.

قضیه ۵.۲.۴. اگر $A = A_۱ \cup A_۲ \cup \cdots$ و هر A_i دارای اندازهٔ صفر باشد، آن‌گاه اندازهٔ A صفر است.

برهان. فرض کنیم $\epsilon > ۰$ داده شده است. چون هر A_i دارای اندازهٔ صفر است، پس پوششی از مستطیل‌ها مانند $B_{i۱}$ ، $B_{i۲}$ ، $\cdots$ برای A_i وجود دارد که $\sum_{j=۱}^{\infty} V(B_{ij}) < \frac{\epsilon}{۲^i}$. حال گردایهٔ شمارای $\{B_{ij} \,،\, j \in \mathbb{N}\}$ پوششی برای A است و

$$\sum_{i,j=۱}^{\infty} V(B_{ij}) = \sum_{i=۱}^{\infty}\sum_{j=۱}^{\infty} V(B_{ij}) < \sum_{i=۱}^{\infty} \frac{\epsilon}{۲^i} = \epsilon$$

و این نشان می‌دهد A دارای اندازهٔ صفر است. $\square$

اکنون یکی از قضایای مهم انتگرال‌گیری را بیان می‌کنیم. انتظار داریم که یک تابع رفتاری معقول همانند توابع پیوسته داشته باشد، تا بتوانیم انتگرال‌پذیری آن را ثابت کنیم. قضیه‌ای که توسط لبگ ثابت شده است، دقیقاً رفتار «معقول» را مشخص نموده و با وارد نمودن مفهوم اندازهٔ صفر، انتگرال‌گیری را به دنیای دیگری کشانیده است. قبل از بیان قضیه، ذکر مقدماتی لازم است.

فرض کنیم $A \subset \mathbb{R}^n$ و $f : A \to \mathbb{R}$ کران‌دار باشند. میزان اشکال موجود بر سر راه پیوستگی f در $x_۰ \in A$ را می‌توان به‌صورت دقیقی اندازه گرفت. برای $\delta > ۰$ فرض کنیم

$$M(x_۰, f, \delta) = \sup\{f(x) \mid x \in A \,, |x - x_۰| < \delta\} \,,$$

$$m(x_۰, f, \delta) = \inf\{f(x) \mid x \in A \,, |x - x_۰| < \delta\} \,.$$

تعریف ۶.۲.۴. نوسان f در نقطهٔ $x_۰$ عبارت است از

$$o(f, x_۰) = \lim_{\delta \to ۰}[M(x_۰, f, \delta) - m(x_۰, f, \delta)] \,.$$

این حد همواره وجود دارد زیرا $M(x_\circ, f, \delta) - m(x_\circ, f, \delta)$ با کاهش δ نزول می‌کند.

قضیه ۷.۲.۴. تابع کران‌دار f در نقطهٔ $x_\circ$ پیوسته است، اگر و فقط اگر $o(f, x_\circ) = \circ$.

برهان. فرض کنیم f در $x_\circ$ پیوسته است و $\epsilon > \circ$ داده شده است. عدد $\delta > \circ$ را می‌توان چنان اختیار کرد که برای هر $x \in A$، اگر $|x - x_\circ| < \delta$ آن‌گاه $|f(x) - f(x_\circ)| < \epsilon$ و در نتیجه

$$M(x_\circ, f, \delta) - m(x_\circ, f, \delta) \leq ۲\epsilon.$$

چون ϵ دلخواه بود پس $o(f, x_\circ) = \circ$. عکس مطلب نیز به همین ترتیب ثابت می‌شود (اثبات به‌عهدهٔ خواننده). $\square$

لم ۸.۲.۴. فرض کنیم $A \subset \mathbb{R}^n$ بسته است. اگر $f : A \to \mathbb{R}$ تابعی کران‌دار باشد و $\epsilon > \circ$ آن‌گاه $D_\epsilon = \{x \in A |\ o(f, x) \geq \epsilon\}$ بسته است.

برهان. نشان می‌دهیم که $\mathbb{R}^n - D_\epsilon$ باز است. اگر $x \in \mathbb{R}^n - D_\epsilon$، آن‌گاه یا $x \notin A$ و یا $x \in A$ ولی $o(f, x) < \epsilon$. در حالت اوّل چون A بسته است، مستطیل بازی مانند C وجود دارد که

$$x \in C \subset \mathbb{R}^n - A \subset \mathbb{R}^n - D_\epsilon.$$

در حالت دوّم $\delta > \circ$ وجود دارد که

$$M(x, f, \delta) - m(x, f, \delta) \leq \epsilon.$$

فرض کنیم C مستطیل بازی باشد و برای هر $y \in C$ داشته باشیم $|x - y| < \delta$. در این صورت برای هر $y \in C$ عددی مانند $\delta_۱ > \circ$ وجود دارد که برای هر z اگر $|z - y| < \delta_۱$ آن‌گاه $|x - z| < \delta$. بنابراین

$$M(y, f, \delta) - m(y, f, \delta) < \epsilon$$

و در نتیجه $o(y,f) < \epsilon$ پس $C \subset \mathbb{R}^n - D_\varepsilon$. $\square$

اکنون می‌توانیم قضیهٔ زیر را که صورتی از قضیهٔ لبگ است بیان کنیم.

قضیه ۹.۲.۴ (قضیه لبگ)**.** فرض کنیم A مستطیل بسته و $f : A \longrightarrow \mathbb{R}$ یک تابع کران‌دار باشد. و فرض کنیم $D = \{x \mid$ در x پیوسته نیست $f\}$. در این صورت f انتگرال‌پذیر است، اگر و تنها اگر D یک مجموعه از اندازهٔ صفر باشد.

برهان. فرض کنیم D دارای اندازهٔ صفر است و فرض کنیم $\varepsilon > 0$ و $D_\varepsilon = \{x \mid o(f,x) \geq \varepsilon\}$. در این صورت $D_\varepsilon \subseteq D$ و بنابراین D_ε دارای اندازهٔ صفر است. چون D_ϵ طبق لم فشرده است پس دارای محتوای صفر می‌باشد. یعنی تعداد متناهی مستطیل B_1،...و B_N وجود دارد که D_ϵ را می‌پوشاند و $\sum_{i=1}^N V(B_i) < \epsilon$. فرض کنیم P افرازی از B باشد به‌قسمی که هر زیر مستطیل S از P در یکی از دو دسته زیر جای داشته باشد.

۱. دستهٔ C_1 متشکل از زیرمستطیل‌های S، به‌قسمی که i ی وجود داشته باشد که $S \subset B_i$.

۲. دستهٔ C_2 متشکل از مستطیل‌هایی که D_ε را قطع نمی‌کنند.

فرض کنیم $|f(x)| < M$ برای هر $x \in B$. در این صورت برای هر S

$$M_S(f) - m_S(f) < 2M$$

. بنابراین

$$\sum_{S \in C_1} [M_S(f) - m_S(f)] \cdot V(S) < 2M \sum_{i=1}^{N} V(B_i) < 2M\varepsilon\ .$$

حال اگر $S \in C_2$ آن‌گاه برای هر $x \in S$، $o(f,x) < \varepsilon$. پس باتوجه به لم برای هر $x \in S$، همسایگی $U(x)$ وجود دارد به قسمی که $M_{U(x)}(f) - m_{U(x)}(f) < \varepsilon$. چون S فشرده

است تعداد متناهی از $U(x)$ ها مانند $U(x_۱),\dots,U(x_n)$ مجموعه S را می‌پوشاند. فرض کنیم P' افرازی ظریف‌تر از P باشد به‌قسمی که هر زیر مستطیل S' از P' که D_ε را قطع نکند، در داخل یک $U(x_i)$ جای داشته باشد. در این صورت برای هر زیر مستطیل S' از P'

$$M_{S'}(f)-m_{S'}(f)<\varepsilon\ .$$

بنابراین برای هر $S\in C_۲$

$$\sum_{S'\subset S}[M_{S'}(f)-m_{S'}(f)]\cdot V(S')<\varepsilon V(S)$$

و داریم

$$\begin{aligned}U(f,p')-L(f,p')&=\sum_{S'\subset S\in C_۱}[M_{S'}(f)-m_{S'}(f)]\cdot V(S')\\&\quad+\sum_{s'\subset S\in C_۲}[M_{S'}(f)-m_{S'}(f)]\cdot V(S')\\&\leq ۲M\varepsilon+\sum_{S\in C_۲}\varepsilon\cdot V(S)\leq ۲M\varepsilon+\varepsilon\cdot V(B)\ .\end{aligned}$$

چون ϵ دلخواه بود پس بنا بر شرط ریمان f انتگرال‌پذیر است.
برعکس فرض کنیم f انتگرال‌پذیر است. چون

$$D=D_۱\cup D_{\frac{۱}{۲}}\cup D_{\frac{۱}{۳}}\cup\cdots\ ,$$

بنا بر قضیه کافی است نشان دهیم که اندازهٔ هر $D_{\frac{۱}{n}}$ برابر صفر است. در حقیقت نشان خواهیم داد که $D_{\frac{۱}{n}}$ دارای محتوای صفر است.

فرض کنیم $\varepsilon>۰$ و P افرازی از B باشد که

$$U(f,P)-L(f,P)<\frac{\varepsilon}{n}\ .$$

فرض کنیم C گردایه‌ای از زیر مستطیل‌های S از P باشد که $D_{\frac{1}{n}}$ را قطع می‌کنند. در این صورت C پوششی از $D_{\frac{1}{n}}$ است. حال اگر $S \in C$ آن‌گاه

$$M_S(f) - m_S(f) \geq \frac{1}{n} \,.$$

بنابراین

$$\begin{aligned}\frac{1}{n}\sum_{S\in C} V(S) &\leq \sum_{S\in C}[M_S(f) - m_S(f)]\cdot V(S)\\ &\leq \sum_{S}[M_S(f) - m_S(f)]\cdot V(S) \leq \frac{\varepsilon}{n}\end{aligned}$$

و در نتیجه $\sum_{S\in C} V(S) < \varepsilon$. □

نتیجه ۴. زیر مجموعه کران‌دار A از $\mathbb{R}^n$ دارای حجم است (یعنی χ_A انتگرال‌پذیر است)، اگر و فقط اگر مرز A دارای اندازه صفر باشد.

برهان. بنا بر قضیه ۹.۲.۴ کافی است نشان دهیم که نقاط ناپیوستگی تابع

$$\chi_A = \begin{cases} 0 & x \notin A \\ 1 & x \in A \end{cases}$$

برابر مرز A است. مرز A را با $\partial(A)$ نشان می‌دهیم. اگر $x \in \partial(A)$ آن‌گاه هر همسایگی x هم A را قطع می‌کند و هم $\mathbb{R}^n - A$ را. بنابراین نقطه‌ای مانند y در این همسایگی وجود دارد که

$$|\chi_A(x) - \chi_A(y)| = 1 \,.$$

پس χ_A در x پیوسته نیست. اگر $x \notin \partial(A)$ آن‌گاه یک همسایگی از x وجود دارد که به‌طور کامل یا در A واقع است و یا در $\mathbb{R} - A$. در هر دو صورت χ_A در این همسایگی تابع ثابت است و در نتیجه در x پیوسته است. □

نتیجه ۵. فرض کنیم $A \subset \mathbb{R}^n$ کران‌دار و دارای حجم است. هر تابع کران‌دار $f : A \longrightarrow \mathbb{R}$ که تعداد نقاط ناپیوستگی آن متناهی یا شمارا باشد (یا به طور کلی مجموعه‌ای با اندازه صفر)، انتگرال‌پذیر است.

برهان. نقاط ناپیوستگی تابع توسیع یافتهٔ g که روی A برابر است با f و در خارج A برابر است با صفر عبارتند از نقاط ناپیوستگی f به‌علاوهٔ احتمالاً نقاطی از مرز A. چون بنابر نتیجه ۴ مرز A دارای اندازهٔ صفر است پس باتوجه به قضیه ۹.۲.۴، کافی است نشان دهیم که یک مجموعهٔ شمارا دارای اندازه صفر است. ولی این مطلب با در نظر داشتن این که یک نقطه دارای اندازهٔ صفر است، نتیجه‌ای از قضیه ۵.۲.۴ است . □

این نتیجه، بسیاری از توابعی را که در عمل با آن برخورد می‌کنیم، در بر دارد. مثلاً هر تابع پیوسته بر بازهٔ $[a,b]$ انتگرال‌پذیر است زیرا $[a,b]$ دارای حجم است و مرز آن شامل دو نقطهٔ a و b است.

یادآوری می‌کنیم که تابع f را تکه‌ای پیوسته خوانیم، درصورتی‌که تعداد نقاط ناپیوستگی آن متناهی باشد. توابع تکه‌ای پیوسته کران‌دار نیز به‌همین دلیل انتگرال‌پذیرند.

توجه کنید که در قضیه ۹.۲.۴، انتگرال‌پذیری f وابسته به تابع توسیع یافتهٔ g است. مثلاً اگر A مجموعهٔ نقاط در $[\text{۰},\text{۱}]$ باشد و f روی A برابر ۱ تعریف شود، گرچه f روی A پیوسته است ولی g در هیچ نقطه‌ای پیوسته نیست و در نتیجه انتگرال‌پذیر نمی‌باشد. در نتیجهٔ ۵ لازم نیست که تابع توسیع یافته را در نظر گیریم، ولی در عوض فرض شده است که A دارای حجم است.

مثال ۱۰.۲.۴. نشان می‌دهیم که تابع

$$f(x) = \begin{cases} x & -\text{۱} \leq x \leq \text{۰} \\ \text{۳}x + \text{۸} & \text{۰} < x \leq \text{۱} \end{cases}$$

در $[-۱,۱]$ انتگرال‌پذیر است. در اینجا f فقط یک نقطهٔ ناپیوستگی در $x=۰$ دارد. پس باتوجه به نتیجهٔ ۵ انتگرال‌پذیر می‌باشد، زیرا کران‌دار است.

مثال ۱۱.۲.۴. فرض کنید

$$f(x,y)=\begin{cases} x^{۲}+\sin\frac{۱}{y} & \text{اگر } y\neq ۰\\ x^{۲} & \text{اگر } y=۰\end{cases}$$

نشان دهید که f روی $A=\{(x,y)\ \mid\ x^{۲}+y^{۲}<۱\}$ انتگرال‌پذیر است.

حل. چون مرز A دارای اندازهٔ صفر است، پس بنابر نتیجه ۴ مجموعهٔ A دارای حجم است. حال تابع f روی A کران‌دار است و نقاط ناپیوستگی آن روی خط $y=۰$ است که مجموعه‌ای با اندازهٔ صفر می‌باشد. بنابراین باتوجه به نتیجهٔ ۵ تابع f روی A انتگرال‌پذیر است.

قضیه ۱۲.۲.۴. (الف) فرض کنیم $A\subseteq\mathbb{R}^n$ کران‌دار و دارای اندازهٔ صفر و $f:A\longrightarrow\mathbb{R}$ انتگرال‌پذیر باشد، در این صورت $\int_A f(x)\,\mathrm{d}x=۰$.
(ب) اگر $f:A\longrightarrow\mathbb{R}$ انتگرال‌پذیر باشد و برای هر x، $f(x)\geq ۰$ و $\int_A f(x)\,\mathrm{d}x=۰$، آن‌گاه مجموعهٔ $\{x\in A\ \mid\ f(x)\neq ۰\}$ دارای اندازهٔ صفر است.

برهان. (الف) فرض کنیم B یک مستطیل شامل A باشد. f را به B توسیع می‌دهیم به این ترتیب که روی $B-A$ تعریف می‌کنیم $f(x)=۰$. فرض کنیم $p=\{S_۱,\dots,S_N\}$ افرازی از B باشد و فرض کنیم M یک کران بالای f روی A باشد. در این صورت

$$L(f,p)=\sum_{i=۱}^{N} m_{S_i}(f)V(S_i)\leq M\sum_{i=۱}^{N} m_{S_i}(\chi_A)V(S_i)\ .$$

فرض کنیم زیر مستطیل ناتهی S_i وجود داشته باشد که $m_{S_i}(\chi_A)\neq ۰$. این بدین معنی است که $S_i\subset A$ ولی این غیر ممکن است زیرا A دارای اندازهٔ صفر است. پس برای هر S_i ناتهی

داریم $m_{S_i}(\chi_A) = \circ$. برای S_iهای تهی نیز $V(S_i) = \circ$. بنابراین $\sum_{i=1}^{N} m_{S_i}(\chi_A)V(S_i) = \circ$ و یا $L(f,p) \leq \circ$. به‌علاوه از

$$\sup\{f(x) \ | \ x \in S_i\} = -\inf\{-f(x) \ | \ x \in S_i\}$$

نتیجه می‌گردد که

$$\begin{aligned} U(f,p) &= \sum_{S_i \in p} \sup\{f(x) \ | \ x \in S_i\} \cdot V(S_i) \\ &= -\sum_{S_i \in p} \inf\{-f(x) \ | \ x \in S_i\} \cdot V(S_i) = -L(-f,p) \ . \end{aligned}$$

باتوجه به بحث بالا، $L(-f,p) \leq \circ$. پس

$$U(f,p) = -L(-f,p) \geq \circ \ .$$

چون p دلخواه بود، نتیجه می‌گردد که

$$\overline{\int_A} f \geq \circ \geq \underline{\int_A} f \ .$$

پس چون f انتگرال‌پذیر است

$$\overline{\int_A} f = \underline{\int_A} f = \int_A f = \circ \ .$$

(ب) فرض کنیم $A_n = \{x \in A \ | \ f(x) > \frac{1}{n}\}$. اگر نشان دهیم که A_n دارای محتوای صفر است چون

$$\{x \in A \ | \ f(x) \neq \circ\} = \cup_{n=1}^{\infty} A_n$$

باتوجه به قضیه ۵.۲.۴ حکم ثابت می‌شود.

فرض کنیم $\varepsilon > 0$ داده شده است. فرض کنیم B مستطیلی شامل A باشد و f را به B توسیع می‌دهیم به این ترتیب که روی $B - A$ قرار می‌دهیم $f = 0$. فرض کنیم p افرازی از مستطیل B است، به‌قسمی که $U(f,p) < \frac{\varepsilon}{n}$. چون $\int_A f = 0$ چنین افرازی وجود دارد. اگر S_1، $\cdots$ و S_N زیر مستطیل‌هایی از p باشند که A_n را قطع می‌کنند، چون $nM_{S_i}(f) \geq 1$ پس

$$\sum_{i=1}^{N} V(S_i) \leq \sum_{i=1}^{N} nM_{S_i}(f)V(S_i) < \varepsilon \ .$$

بنابراین S_1، $\cdots$ و S_N پوششی از مستطیل‌های بسته برای A_n است، که $\sum_{i=1} V(S_i) < \varepsilon$، یعنی A_n دارای محتوای صفر است و برهان تمام می‌شود. $\square$

قضیه ذیل به خواص انتگرال می‌پردازد که قبلاً مشابه آن را در انتگرال توابع یک متغیره دیده‌ایم. برهان آن را به خواننده واگذار می‌کنیم.

قضیه ۱۳.۲.۴ (خواص انتگرال)**.** فرض کنیم $A, B \subseteq \mathbb{R}^n$ و $c \in \mathbb{R}$، $f, g : A \longrightarrow \mathbb{R}$ انتگرال‌پذیر باشد. در این صورت

۱. $f+g$ انتگرال‌پذیر است و $\int_A f+g = \int_A f + \int_A g$.

۲. cf انتگرال‌پذیر است و $\int_A cf = c\int_A f$.

۳. $|f|$ انتگرال‌پذیر است و $|\int_A f| \leq \int_A |f|$.

۴. اگر $f \leq g$ آن‌گاه $\int_A f \leq \int_A g$.

۵. اگر A دارای حجم باشد و $|f| \leq M$ آن‌گاه $|\int_A f| \leq MV(A)$.

۶. (قضیه مقدار میانگین برای انتگرال‌ها) اگر $f : A \longrightarrow \mathbb{R}$ پیوسته و A دارای حجم، فشرده و همبند باشد، آن‌گاه $x_0 \in A$ چنان موجود است که $\int_A f(x)\,\mathrm{d}x = f(x_0)V(A)$. مقدار $\frac{\int_A f}{V(A)}$ میانگین f روی A نامیده می‌شود.

۷. فرض کنیم $f : A \cup B \longrightarrow \mathbb{R}$ یک تابع باشد. اگر A و B چنان باشند که $A \cap B$ با اندازه صفر باشد و $f\mid_{A\cap B}$ و $f\mid_A$ و $f\mid_B$ انتگرال‌پذیر باشند، آن‌گاه f روی $A \cup B$ انتگرال‌پذیر است و $\int_{A\cup B} f = \int_A + \int_B f$.

۳.۴ قضیهٔ فوبینی و تعویض متغیر

دو قضیهٔ مهم وجود دارد که ما را در محاسبهٔ انتگرال‌های چندگانه یاری می‌کند. قضیهٔ نخست، انتگرال‌های چندگانه را به کمک انتگرال‌گیری مکرر قابل محاسبه می‌سازد. مثلاً از درس حسابان می‌دانیم که اگر A مربع $\{(x,y) \mid ۰ \leq x \leq ۱, ۰ \leq y \leq ۱\}$ باشد آن‌گاه

$$\begin{aligned}\int_A (x+y)x \,\mathrm{d}x\,\mathrm{d}y &= \int_{x=۰}^{۱}\left(\int_{y=۰}^{۱}(x^۲+yx)\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_۰^۱ (x^۲ + \frac{۱}{۲}x)\,\mathrm{d}x = \frac{۱}{۳}+\frac{۱}{۴} = \frac{۷}{۱۲}.\end{aligned}$$

این قضیه که به "قضیهٔ فوبینی" معروف است، به ابعاد بالاتری تعمیم می‌یابد. در این کتاب این قضیه را در حالت خاص بیان می‌کنیم و برای دیدن صورت کامل‌تر این قضیه دانشجوی علاقمند را به کتب پیشرفته‌تر آنالیز حقیقی ارجاع می‌دهیم [۱].

قضیه ۱.۳.۴ (قضیهٔ فوبینی[۱]). فرض کنیم $A \subseteq \mathbb{R}^n$ و $B \subseteq \mathbb{R}^m$ مستطیل‌های بسته و $f\colon A\times B \to \mathbb{R}$ انتگرال‌پذیر باشد. برای هر $x \in A$، تابع $f_x\colon B \subseteq \mathbb{R}^m \to \mathbb{R}$ را به‌صورت $f_x(y) = f(x,y)$ تعریف می‌کنیم. اگر به ازای هر $x \in A$، f_x انتگرال‌پذیر باشد آن‌گاه

$$\int_{A\times B} f = \int_A\left(\int_B f(x,y)\,\mathrm{d}y\right)\mathrm{d}x.$$

1. Fubini

برهان. فرض کنیم $g : A \to \mathbb{R}$ تابعی باشد که به‌صورت

$$g(x) = \int_B f(x,y)\,\mathrm{d}y$$

تعریف می شود. بایستی نشان دهیم که g روی A انتگرال‌پذیر است و

$$\int_{A\times B} f = \int_A g(x)\,\mathrm{d}x \ . \tag{۱}$$

فرض کنیم p افرازی از $A \times B$ باشد. افراز p_A از A و افراز p_B از B را چنان در نظر می‌گیریم که اگر مستطیل‌های جزئی این دو افراز را به ترتیب به S_A و S_B نشان دهیم، هر مربع S از p به‌صورت $S_A \times S_B$ باشد. بنابراین

$$\begin{aligned} L(f,p) = \sum_S m_S(f)\cdot V(S) &= \sum_{S_A,S_B} m_{S_A\times S_B}(f)\,V(S_A\times S_B) \\ &= \sum_{S_A}\left(\sum_{S_B} m_{S_A\times S_B}(f)\cdot V(S_B)\right)V(S_A). \end{aligned}$$

برای هر $x \in S_A$ داریم $m_{S_A\times S_B}(f) \le m_{S_B}(f_x)$. بنابراین

$$\sum_{S_B} m_{S_A\times S_B}(f)\cdot V(S_B) \le \sum_{S_B} m_{S_B}(f_x)\,V(S_B) \le \int_B f_x(y)\,\mathrm{d}y = g(x) \ .$$

چون این رابطه برای هر $x \in S_A$ برقرار است، پس

$$\sum_{S_B} m_{S_A\times S_B}(f)\cdot V(S_B) \le m_{S_A}(g).$$

در نتیجه

$$L(f,p) = \sum_{S_A,S_B} m_{S_A\times S_B}(f)\cdot V(S_A\times S_B) \le \sum_{S_A} m_{S_A}(g)\cdot V(S_A) \le L(g,p_A) \ .$$

از این رابطه و از بحثی مشابه برای حاصل‌جمع بالایی خواهیم داشت،

$$L(f,p) \le L(g,p_A) \le U(g,p_A) \le U(f,p) \ .$$

چون f روی $A \times B$ انتگرال‌پذیر است، از این نامساوی نتیجه می‌گردد که g انتگرال‌پذیر است و

$$\int_{A\times B} f = \int_A g(x)\,\mathrm{d}x$$

که همان رابطهٔ (۱) است. □

تذکر. برهانی مشابه نشان می‌دهد که با شرایط متناظر داریم

$$\int_{A\times B} f = \int_B \left(\int_A f(x,y)\,\mathrm{d}x \right) \mathrm{d}y.$$

تذکر. اگر تابع f در قضیهٔ بالا پیوسته باشد، آن‌گاه f_x خودبه‌خود پیوسته و در نتیجه انتگرال‌پذیر است و بنابراین

$$\int_{A\times B} f = \int_A \left(\int_B f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

مثال ۲.۳.۴. فرض کنید A هرمی باشد که رأس آن مبدأ مختصات است و قاعدهٔ آن مثلثی با رئوس $(۱,۰,۰)$، $(۰,۱,۰)$ و $(۰,۰,۱)$. انتگرال $\int_A (x+y+z)^۲\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x$ را حساب کنید.

حل. A عبارت است از مجموعهٔ

$$\{(x,y,z) \in \mathbb{R}^۳ |\ x \ge ۰\ \ ,\ y \ge ۰\ \ ,\ z \ge ۰\ \ ,\ x+y+z \le ۱\} \ .$$

بنابراین A شامل نقاط (x,y,z) می‌باشد که $x \ge ۰$، $y \ge ۰$، $x+y \le ۱$ و $۰ \le z \le ۱-(x+y)$.

حال فرض کنید

$$B = \{(x,y) |\ x \geq ۰ \ ,\ y \geq ۰ \ ,\ x+y \leq ۱\} \ .$$

در این صورت بنا بر قضیه ۱.۳.۴ و باتوجه به این که f در خارج A برابر صفر است داریم

$$\int_A (x+y+z)^۲ \,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = \int_B \left(\int_۰^{۱-(x+y)} (x+y+z)^۲ \,\mathrm{d}z \right) \mathrm{d}y\,\mathrm{d}x \ .$$

ولی B تشکیل شده است از نقاط (x,y) ای که $۰ \leq x \leq ۱$ و $۰ \leq y \leq ۱ - x$، پس

$$\begin{aligned} \int_A (x+y+z)^۲ \,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x &= \int_۰^۱ \int_۰^{۱-x} \int_۰^{۱-(x+y)} (x+y+z)^۲ \,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_۰^۱ \int_۰^{۱-x} \left[\frac{(x+y+z)^۳}{۳} \right]_۰^{۱-(x+y)} \mathrm{d}y\,\mathrm{d}x \\ &= \int_۰^۱ \int_۰^{۱-x} \left(\frac{(x+y+۱-(x+y))^۳}{۳} - \frac{(x+y+۰)^۳}{۲} \right) \mathrm{d}y\,\mathrm{d}x \\ &= \int_۰^۱ \int_۰^{۱-x} \left(\frac{۱}{۳} - \frac{(x+y)^۳}{۳} \right) \mathrm{d}y\,\mathrm{d}x \\ &= \int_۰^۱ \left(\frac{۱-x}{۳} - \frac{(x+(۱-x))^۴}{۱۲} + \frac{x^۴}{۱۲} \right) \mathrm{d}x \\ &= \int_۰^۱ \left(\frac{۱}{۳} - \frac{x}{۳} - \frac{۱}{۱۲} + \frac{x^۴}{۱۲} \right) \mathrm{d}x \\ &= \frac{۱}{۴} - \frac{۱}{۶} + \frac{۱}{۶۰} = \frac{۱}{۱۰} \ . \end{aligned}$$

قضیهٔ مهّم دوم در این فصل "قضیهٔ تعویض متغیر" است. قبل از بیان قضیهٔ تعویض متغیر مفهومی دیگر به نام افراز واحد را مطالعه می‌کنیم.

تعریف ۳.۳.۴. محمل تابع حقیقی یا مختلط f که روی $\mathbb{R}^n$ تعریف شده است، عبارت است از بستار مجموعهٔ نقاط $x \in \mathbb{R}^n$ ای که $f(x) \neq ۰$. محمل تابع f را به $\mathrm{Supp}(f)$

نشان خواهیم داد. به‌عبارت دیگر

$$\operatorname{Supp}(f) = \overline{\{x \in \mathbb{R}^n |\ f(x) \neq 0\}}\ .$$

تعریف ۴.۳.۴. یک افراز واحد C^∞ برای مجموعهٔ $A \subset \mathbb{R}^n$، عبارت است از گردایه‌ای از توابع بی‌نهایت بار مشتق‌پذیر C^∞، مانند $\{\varphi_\gamma\}$، که روی A تعریف شده‌اند و در شرایط زیر صدق کنند.

۱. برای هر $x \in A$، $\varphi_\gamma(x) \geq 0$.
۲. هر $x \in A$ یک همسایگی مانند V داشته باشد که محمل تعداد متناهی از φ_γ ها را قطع کند. یعنی به جز تعداد متناهی، بقیهٔ φ_γ ها روی V برابر ۰ باشند.
۳. برای هر $x \in A$، $\sum_\gamma \varphi_\gamma(x) = 1$. (توجه کنید که بنا بر ۲ این مجموع در یک همسایگی x، متناهی است).

اگر $\{U_\alpha\}$ پوششی باز از A باشد، افراز واحد $\{\varphi_\gamma\}$ را پیرو پوشش $\{U_\alpha\}$ خوانیم، اگر برای هر γ، U_α ای وجود داشته باشد، به‌طوری که $\operatorname{Supp}\varphi_\gamma \subset U_\alpha$.

قضیه ۵.۳.۴. فرض کنیم $A \subset \mathbb{R}^n$ و $\xi = \{U_\alpha\}$ پوششی باز از A باشد. در این صورت یک افراز واحد پیوسته برای A وجود دارد که پیرو $\{U_\alpha\}$ است.

برهان. **حالت ۱:** A فشرده است. برای هر $x \in A$، $\alpha(x)$ ای وجود دارد که $x \in U_{\alpha(x)}$. گوی‌های باز $B(x)$ و $W(x)$ را چنان اختیار می‌کنیم که

$$\overline{B(x)} \subset W(x) \subset \overline{W(x)} \subset U_{\alpha(x)}\ . \qquad (1)$$

چون A فشرده است، تعداد متناهی نقطهٔ x_1، ...و x_k وجود دارد که

$$A \subset B(x_1) \cup \cdots \cup B(x_k)\ .$$

باتوجه به (۱) و بنا بر لم اوریسون که در کتب توپولوژی عمومی به آن اشاره می‌شود، توابع پیوستهٔ $\Psi_۱$، $\ldots ,$ و $\Psi_s,$

وجود دارند به‌قسمی که $\Psi_i(x) = ۱$ روی $B(x_i)$، $\Psi_i(x) = ۰$ در خارج $W(x_i)$ و $۰ \leq \Psi_i(x) \leq ۱$ برای هر $x \in \mathbb{R}^n$. حال تعریف می‌کنیم $\varphi_۱ = \Psi_۱$ و

$$\varphi_{i+۱} = (۱ - \Psi_۱) \cdots (۱ - \Psi_i)\Psi_{i+۱} \,. \qquad (۲)$$

شرایط ۱ و ۲ از تعریف ۴.۳.۴ به وضوح برای گردآیهٔ $\{\Psi_۱, \ldots , \Psi_s\}$ برقرارند. به‌علاوه به کمک استقرار داریم

$$\mathrm{Supp}\ \varphi_i \subset \overline{W(x_i)} \subset U_{\alpha(x_i)} \,.$$

زیرا برقراری آن برای $i = ۱$ بدیهی است و اگر برای i برقرار باشد از جمع آن با (۲)، رابطهٔ بالا برای $i + ۱$ به‌جای i به دست می‌آید. پس

$$\sum_{i=۱}^{s} \varphi_i(x) = ۱ - \prod_{i=۱}^{s} \left(۱ - \Psi_i(x)\right) \,.$$

پس اگر $x \in A$ آن‌گاه $B(x_i)$ ای وجود دارد که $x \in B(x_i)$. بنابراین $\Psi_i(x) = ۱$ و در نتیجه $\sum_{i=۱}^{s} \varphi_i(x) = ۱$. پس شرط ۳ نیز برقرار است. پس $\{\varphi_۱, ..., \varphi_s\}$ افراز واحد مورد نظر است.

حالت ۲.

$$A = A_۱ \cup A_۲ \cup A_۳ \cup \cdots$$

که هر A_i فشرده است و $A_i \subset \mathrm{Int}(A_{i+۱})$، که در آن منظور از نماد $\mathrm{Int}(A_{i+۱})$، درون مجموعهٔ $A_{i+۱}$ است. برای هر i فرض کنیم

$$\xi_i = \{U \cap (\mathrm{Int}(A_{i+۱}) - A_{i-۲}) |\ U \in \xi\} \,.$$

در این صورت ξ_i پوششی باز از مجموعهٔ فشردهٔ

$$B_i = A_i - \operatorname{Int}(A_{i-۱})$$

می‌باشد و بنا بر حالت ۱، یک افراز واحد پیرو ξ_i مانند Φ_i برای B_i وجود دارد. برای هر $x \in A$ مجموع

$$\sigma(x) = \sum \varphi(x),$$

که عمل جمع روی $\varphi \in \Phi_i$ برای تمام i ها محاسبه شده است، در یک همسایگی x مجموعی متناهی است. زیرا اگر $x \in A_i$، آن‌گاه برای $\varphi \in \Phi_j$ با $j \geq i + ۲$ داریم $\varphi(x) = ۰$. برای هر φ از یک Φ_i قرار می‌دهیم

$$\varphi'(x) = \frac{\varphi(x)}{\sigma(x)}$$

در این صورت گردایهٔ تمام φ' ها افراز واحد مطلوب است.

حالت ۳. A باز است. قرار می‌دهیم

$$A_i = \{x \in A \mid |x| \leq i\,,\ \text{فاصلهٔ } x \text{ از مرز } A \text{ بزرگتر یا مساوی است با } \frac{۱}{i}\}$$

و حالت ۲ را به کار می‌بریم.

حالت ۴. A دلخواه است. فرض کنیم

$$B = \bigcup \{U \mid U \in \xi\}$$

بنابر حالت ۳ یک افراز واحد برای B وجود دارد. همین برای A نیز یک افراز واحد است. $\square$

اکنون به بیان قضیهٔ تعویض متغیر می‌پردازیم. در درس حسابان دیده‌ایم که اگر $g : [a,b] \to \mathbb{R}$ دارای مشتق پیوسته باشد و $f : \mathbb{R} \to \mathbb{R}$ پیوسته باشد، آن‌گاه

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) \cdot g'.$$

اثبات این قضیه بسیار ساده است؛ اگر قرار دهیم $F(x) = \int_{g(a)}^{x} f$، آن‌گاه بنا بر دومین قضیهٔ اساسی حسابان طرف چپ برابر است با $F\big(g(b)\big) - F\big(g(a)\big)$. به‌علاوه، بنا بر اولین قضیهٔ اساسی حسابان و قاعدهٔ زنجیری داریم، $(F \circ g)' = (f \circ g) \cdot g'$. پس طرف راست برابر است با $F \circ g(b) - F \circ g(a)$ و قضیه ثابت می‌شود. اگر g علاوه بر پیوسته‌ـ مشتق‌پذیر بودن یک‌به‌یک نیز باشد، آن‌گاه فرمول بالا را می‌توان به‌صورت

$$\int_{g(a,b)} f = \int_{(a,b)} (f \circ g) \cdot |g'|$$

نوشت (این رابطه را ثابت کنید! دو حالتی را که g صعودی یا نزولی باشد به‌صورت جداگانه بررسی کنید). تعمیم این قضیه به بعدهای بالاتر برهانی پیچیده دارد که اکنون به ذکر آن می‌پردازیم.

قضیه ۶.۳.۴ (تعویض متغیر). فرض کنیم $A \subseteq \mathbb{R}^n$ مجموعه‌ای باز و کران‌دار و $g : A \to \mathbb{R}^n$ تابعی C^1 و یک‌به‌یک باشد، به‌قسمی که در هر نقطهٔ $x \in A$، $J\big(g(x)\big) \neq 0$ (یعنی $\det g'(x) \neq 0$). اگر $f : g(A) \to \mathbb{R}$ انتگرال‌پذیر باشد، آن‌گاه

$$\int_{g(A)} f = \int_A (f \circ g) \cdot |J_g|$$

یعنی

$$\int_{g(A)} f(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n = \int_A f(g(x_1, \ldots, x_n)) \left| \frac{\delta(g_1, \ldots, g_n)}{\delta(x_1, \ldots, x_n)} \right| \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

برهان. قضیه را در چند مرحله اثبات می‌کنیم.

مرحلهٔ ۱. فرض کنیم ξ پوششی باز از A باشد به قسمی که برای هر $U \in \xi$ و هر تابع انتگرال‌پذیر f داشته باشیم،

$$\int_{g(U)} f = \int_U (f \circ g)|J_g| \,.$$

در این صورت قضیه برقرار است (فقط در این مرحله یک‌به‌یک بودن g روی A مورد استفاده قرار می‌گیرد). گردآیهٔ $g(U)$، پوششی باز از $g(A)$ است (چرا؟). فرض کنیم Φ یک افراز واحد پیرو این پوشش باز باشد. اگر داشته باشیم $\Psi = ۰$ در خارج $g(U)$، آن‌گاه g یک‌به‌یک است و در خارج U خواهیم داشت $(\varphi \cdot f) \cdot g = ۰$. بنابراین می‌توان

$$\int_{g(U)} \varphi \cdot f = \int_U ((\varphi \cdot f) \circ g)|J_g|$$

را به‌صورت

$$\int_{g(A)} \varphi \cdot f = \int_A ((\varphi \cdot f) \circ g)|Jg|,$$

نوشت. بنابراین

$$\begin{aligned}\int_{g(A)} f &= \sum_{\varphi \in \Phi} \int_{g(A)} \varphi \cdot f = \sum_{\varphi \in \Phi} \int_A ((\varphi \cdot f) \circ g)|Jg| \\ &= \sum_{\varphi \in \Phi} \int_A (\varphi \circ g) \cdot (f \circ g)|Jg| \\ &= \int_A (f \circ g)|Jg|.\end{aligned}$$

مرحلهٔ ۲. کافی است قضیه را برای حالت $f = ۱$ ثابت کنیم. اگر قضیه برای $f = ۱$ برقرار باشد آن‌گاه برای هر تابع ثابت برقرار است. حال فرض کنیم f یک تابع انتگرال‌پذیر دلخواه روی $g(A)$ باشد. فرض کنیم B یک مستطیل شامل $g(A)$ و $p = \{S_۱, \ldots, S_N\}$ افرازی از B باشد، برای هر $p \in S_i$ فرض کنیم f_{S_i} تابع ثابت $m_{S_i}(f)$ باشد. در این صورت

$$\begin{aligned}L(f,p) &= \sum_{i=۱}^{N} m_{S_i}(f) \cdot V(S_i) = \sum_{i=۱}^{N} \int_{S_i} f_{S_i} \\ &= \sum_{i=۱}^{N} \int_{g^{-۱}(s_i)} (f_{S_i} \circ g)|Jg| \leq \sum_{i=۱}^{N} \int_{g^{-۱}(S_i)} (f \circ g)|Jg|\end{aligned}$$

$$= \int_{g^{-۱}(B)} (f \circ g)|Jg| = \int_A (f \circ g)|Jg| \ .$$

پس

$$\int_{g(A)} f \leq \int_A (f \circ g)|Jg| \ .$$

با استدلالی مشابه برای $f_{S_i} = M_{s_i}(f)$ نتیجه می‌گردد که $\int_{g(A)} f \geq \int_A (f \circ g)|Jg|$ و بنابراین

$$\int_{g(A)} f = \int_A (f \circ g)|Jg| \ .$$

مرحلهٔ ۳. اگر قضیه برای $g : A \longrightarrow \mathbb{R}$ و $h : B \longrightarrow \mathbb{R}^n$ برقرار باشد که $g(A) \subset B$، آن‌گاه برای $h \circ g : A \longrightarrow \mathbb{R}^n$ برقرار است. زیرا

$$\begin{aligned} \int_{h \circ g(A)} f = \int_{h(g(A))} f &= \int_{g(A)} f \circ h|Jh| \\ = \int_A (f \circ h \circ g)(|Jh| \circ g)|Jg| &= \int_A f \circ (h \circ g)|J(h \circ g)| \ . \end{aligned}$$

مرحلهٔ ۴. اگر g یک تبدیل خطی باشد، آن‌گاه قضیه برای g برقرار است.

برای اثبات مرحلهٔ ۴ ادعا می‌کنیم که اگر $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ نگاشتی خطی و $A \subset \mathbb{R}^n$ مجموعه‌ای دارای حجم باشد (یعنی $\int_A \chi_A$ موجود است)، آن‌گاه حجم $L(A)$ برابر است با $|\det A| \cdot V(A)$. یعنی

$$\int_{L(A)} ۱ = \int_{L(A)} \chi_{L(A)} = \int_A |\det L|.$$

اگر این ادعا ثابت شود، آن‌گاه باتوجه به این که برای g، $g = Dg$، نتیجه می‌گردد

$$\int_{g(A)} ۱ = \int_A |\det g| = \int_A |Jg| \ .$$

و بنا بر مرحلهٔ ۲ اثبات خواهیم داشت

$$\int_{g(A)} = \int_A f \circ g|Jg| \; .$$

برای اثبات ادعا ابتدا حالتی را که A یک مستطیل و L یک تبدیل خطی که ماتریس آن

$$L_1 = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

یا

$$L_2 = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{vmatrix}$$

باشد در نظر می‌گیریم (این ماتریس‌ها را ماتریس‌های مقدماتی می‌خوانند، در اولی ماتریس همانی را در نظر می‌گیریم که یک جا روی قطر ۱ را با عدد ثابت c جای‌گزین کرده باشیم و دومی از ماتریس همانی با وارد کردن عدد ۱ یک جا در خارج قطر حاصل شده است.)

اگر A به‌صورت

$$A = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

باشد و c در سطر iـ‌ام واقع باشد آن‌گاه

$$L_1(A) = [a_1, b_1] \times \cdots \times [ca_i, cb_i] \times \cdots \times [a_n, b_n] \; .$$

پس حجم $L_1(A)$ برابر است با

$$V(L_1(A)) = |c|V(A) = |\det L_1| \cdot V(A) \; .$$

اگر عدد ۱ خارج از قطر در مکان (i,j)ـام قرار داشته باشد، آن‌گاه

$$L_۲(A)=\big\{(x_۱,\ldots,x_{i-۱},x_i+x_j,x_{i+۱},\ldots,x_n)\mid ۱\leq k\leq n \text{ برای } x_k\in[a_k,b_k]\big\}.$$

مجموعهٔ $L_۲(A)$ را به سه ناحیه تفکیک می‌کنیم:

$$\{(x_۱,\ldots,x_{i-۱},x_i+x_j,x_{i+۱},\ldots,x_n)\in L_۲(A)\ \mid\ a_i+a_j\leq x_i+x_j\leq a_i+b_j\},$$
$$\{(x_۱,\ldots,x_{i-۱},x_i+x_j,x_{i+۱},\ldots,x_n)\in L_۲(A)\ \mid\ a_i+b_j\leq x_i+x_j\leq a_j+b_i\},$$
$$\{(x_۱,\ldots,x_{i-۱},x_i+x_j,x_{i+۱},\ldots,x_n)\in L_۲(A)\ \mid\ a_j+b_i\leq x_i+x_j\leq b_i+b_j\}\ .$$

حجم مجموعهٔ اول بنا بر قضیهٔ فوبینی برابر است با

$$(b_۱-a_۱)\cdots(b_{i-۱}-a_{i-۱})(b_{i+۱}-a_{i+۱})\cdots(b_{j-۱}-a_{j-۱})(b_{j+۱}-a_{j+۱})$$
$$\cdots(b_n-a_n)\cdot\textstyle\int_{a_۱+a_j}^{a_۱+b_j}(\int_{a_j}^{a_j+a_i+b_j-x_i}۱\,\mathrm{d}x_j)\,\mathrm{d}x_i\ .$$

ولی

$$\int_{a_i+a_j}^{a_i+b_j}(\int_{a_j}^{a_j+a_i+b_j-x_i}۱\,\mathrm{d}x_j)\,\mathrm{d}x_i=\int_{a_i+a_j}^{a_i+b_j}(a_i+b_j-x_i)\,\mathrm{d}x_i=\frac{۱}{۲}(b_j-a_j)^۲\ .$$

پس حجم مجموعهٔ اول برابر است با

$$(b_۱-a_۱)\cdots(b_{i-۱}-a_{i-۱})(b_{i+۱}-a_{i+۱})\cdots(b_{j-۱}-a_{j-۱})(b_{j+۱}-a_{j+۱})$$
$$\cdots(b_n-a_n)\tfrac{۱}{۲}(b_j-a_j)^۲\ .$$

همین روش را می‌توان برای مجموعهٔ سوم به کار گرفت و نشان داد که حجم آن برابر حجم مجموعهٔ اول است.

مجموعهٔ دوم یک مستطیل است با حجم

$$(b_۱ - a_۱) \cdots (b_{i-۱} - a_{i-۱})(b_{i+۱} - a_{i+۱}) \cdots (b_{j-۱} - a_{j-۱})(b_{j+۱} - a_{j+۱})$$
$$\cdots (b_n - a_n)(a_j + b_i - a_i - b_j)(b_j - a_j) \ .$$

حجم $L_۲(A)$ برابر مجموع سه حجم بالا است که برابر است با

$$(b_۱ - a_۱) \cdots (b_n - a_n)$$

و همان حجم مستطیل

$$[a_۱, b_۱] \times \cdots \times [a_n, b_n]$$

می‌باشد. بنابراین باتوجه به این که $\det L_۲ = ۱$، داریم

$$V(L_۲(A)) = |\det L_۲| \cdot V(A) \ .$$

حال فرض کنیم A یک مجموعهٔ دلخواه دارای حجم است و فرض کنیم L_i یک ماتریس مقدماتی است و $\det L_i \neq ۰$ (یعنی اگر L_i ماتریس مقدماتی نوع اول باشد آن‌گاه $c \neq ۰$). فرض کنیم B یک مستطیل شامل A باشد و $p = \{S_۱, \ldots, S_N\}$ افرازی از A به‌قسمی‌که

$$U(\chi_A, p) - V(A) < \frac{\varepsilon}{۲|\det L_i|} \ , \ \ V(A) - L(\chi_A, p) < \frac{\varepsilon}{۲|\det L|} \ .$$

دراین صورت برای دو مجموعهٔ $\{S_i \ \ | \ \ S_i \subset A\}$ $V_\varepsilon = \cup$ و $W_\varepsilon = \cup\{S_i \ \ | \ \ S_i \cap A \neq \varnothing\}$ بنا بر نتیجه‌ای که برای مستطیل‌ها به دست آوردیم، داریم $\nu(L_i(V_\varepsilon)) = |\det L_i| L(\chi_A, p)$ و $\nu(L_i(W_\varepsilon)) = |\det L_i| \cup (\chi_A, p)$. بنابراین $\nu(L_i(W_\varepsilon)) - \nu(L_i(V_\varepsilon)) < \varepsilon$. در نتیجه $L_i(A)$ دارای حجم است و $\nu(L_i(A)) = |\det L_i| V(A)$.
اگر $L_i = ۰$، یعنی L_i ماتریسی از نوع اول، $c = ۰$ باشد آن‌گاه برای هر مستطیل S، $\nu(L_i(S)) = ۰$ و در نتیجه برای هر مجموعهٔ دارای حجم مانند A، $\nu(L_i(A)) = ۰$.

اکنون فرض کنیم L یک نگاشت خطی و A مجموعه‌ای دارای حجم باشد. چون هر ماتریس را می‌توان به‌صورت حاصلضرب ماتریس‌های مقدماتی نوشت، داریم

$$L = L_۱ L_۲ \dots L_k$$

که هر L_i یک ماتریس مقدماتی است. با تکرار آن‌چه در بالا ثابت شد نتیجه می‌گردد که $L(A)$ دارای حجم است و

$$\nu(L(A)) = |\det L_۱||\det L_۲|\dots|\det L_n|V(A) = |\det L|V(A).$$

به این ترتیب اثبات مرحلهٔ ۴ کامل می‌شود.

اکنون آمادگی داریم که قضیه را به کمک استقراء ثابت کنیم. آن‌چه قبل از بیان قضیه گفته شد به‌علاوهٔ مرحله‌های ۱ و ۲ حالت $n = ۱$ را ثابت می‌کند. فرض کنیم قضیه در بعد $n - ۱$ برقرار است آن را در بعد n ثابت می‌کنیم. بنا بر مرحلهٔ ۱ کافی است برای هر $a \in A$ همسایگی $U(a) \subset A$ را چنان بیابیم که قضیه برای U برقرار باشد. به‌علاوه می‌توان فرض کرد $g'(a) = I$. زیرا اگر T تبدیل خطی $Dg(a)$ باشد، آن‌گاه $(T^{-۱} \circ g)'(a) = I$ و چون بنا بر مرحلهٔ ۴ قضیه برای T برقرار است، پس باتوجه به مرحلهٔ ۳ برقراری آن برای $T^{-۱} \circ g$ برقراری آن را برای g ایجاب می‌کند.

نگاشت $h : A \longrightarrow \mathbb{R}^n$ را به‌صورت $h(x) = (g^۱(x), \dots, g^{n-۱}(x), x^n)$ تعریف می‌کنیم. دراین صورت $h'(a) = I$. پس بنا بر قضیه تابع معکوس همسایگی $U' \subset A$ از a وجود دارد که در آن h یک به یک است و $J(h) \neq ۰$. بنابراین می‌توان $k : h(U') \longrightarrow \mathbb{R}^n$ را به‌صورت $k(x) = \left(x^۱, \dots, x^{n-۱}, g^n(h^{-۱}(x)\right)$ تعریف کرد، که در نتیجه $g = k \circ h$. چون

$$(g^n \circ h^{-۱})'(h(a)) = (g^n)'(a) \cdot [h'(a)]^{-۱} = (g^n)'(a)\ ,$$

پس

$$D_n(g^n \circ h^{-۱})(h(a)) = D_n g^n(a) = ۱$$

و بنابراین $k'(h(a)) = ۱$. پس در یک مجموعهٔ باز V که $h(a) \in V \subset h(U')$، نگاشت k یک به یک است و $Jk \neq ۰$. قرار می‌دهیم $U = k^{-۱}(V)$ در این صورت داریم $g = k \circ h$ که $h : U \longrightarrow \mathbb{R}^n$، $k : V \longrightarrow \mathbb{R}^n$ و $h(U) \subset V$. پس بنا بر مرحلهٔ ۳ کافی است قضیه را برای h و k ثابت کنیم. اثبات را برای h می‌نویسیم و برای k اثبات مشابه است.

فرض کنیم $W \subset U$ مستطیلی به‌صورت $D_x[a_n, b_n]$ باشد که D مستطیلی در $\mathbb{R}^{n-۱}$ است بنا بر قضیهٔ فوبینی

$$\int_{h(W)} ۱ = \int_{[a_n,b_n]} \left(\int_{h(Dx\{x^n\})} ۱ \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} \right) \mathrm{d}x^n \, .$$

نگاشت $h_{x^n} : D \longrightarrow \mathbb{R}^{n-۱}$ را که به‌صورت

$$h_{x^n}(x^۱, \ldots, x^{n-۱}) = \left(g^۱(x^۱, \ldots, x^n), \ldots, g^{n-۱}(x^۱, \ldots, x^n) \right) .$$

تعریف شده است، در نظر می‌گیریم در این صورت هر h_{x^n} یک به یک است و

$$\det(h_{x^n})'(x^۱, \ldots, x^{n-۱}) = \det h'(x^۱, \ldots, x^n) \neq ۰ \, .$$

به‌علاوه

$$\int_{h(Dx\{x^n\})} ۱ \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} = \int_{hx^n(D)} ۱ \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} \, .$$

پس اگر قضیه را برای حالت $n - ۱$ به کار بریم خواهیم داشت

$$\begin{aligned} \int_{h(W)} ۱ &= \int_{[a_n,b_n]} \left(\int_{hx^n(D)} ۱ \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} \right) \mathrm{d}x^n \\ &= \int_{[a_n,b_n]} \left(\int_D |J(h_{x^n})| \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} \right) \mathrm{d}x^n \\ &= \int_{[a_n,b_n]} \left(\int_D |Jh| \, \mathrm{d}x^۱ \ldots \mathrm{d}x^{n-۱} \right) \mathrm{d}x^n = \int_W |Jh|. \end{aligned}$$

و به‌این ترتیب برهان کامل می‌شود. $\square$

۴.۴ تمرین

۱. با یک مثال نقض نشان دهید که قضیه تغییر متغیر در حالتی که g یک‌به‌یک نباشد، درست نیست. حتی اگر $Jg(x) \neq \circ$.
راهنمایی: قرار دهید $f = ۱$ و $g(x,y) = (e^x \cos y, e^x \sin y)$.

۲. $\int_A xy \sin(x^۲ - y^۲)\, \mathrm{d}x\, \mathrm{d}y$ را محاسبه کنید که در آن

$$A = \Big\{(x,y) \ \mid \ \circ < y < ۱ \ , x > y \ , \ x^۲ - y^۲ < ۱\Big\} \ .$$

۳. فرض کنید $f, g : A \longrightarrow \mathbb{R}$ انتگرال‌پذیر و $f \leq g$. نشان دهید

$$\int_A f \leq \int_A g \ .$$

۴. فرض کنید $f : [\circ, ۱] \times [\circ, ۱] \longrightarrow \mathbb{R}$ به‌صورت

$$f(x,y) = \begin{cases} \circ & x \in \mathbb{Q} \\ \circ & x \in \mathbb{Q} \ \ , y \notin \mathbb{Q} \\ \frac{۱}{q} & x \in \mathbb{Q} \ , \ y = \frac{p}{q} \ , \ (p,q) = ۱ \end{cases}$$

نشان دهید f انتگرال‌پذیر است و $\int_{[\circ,۱]\times[\circ,۱]} f = \circ$.

۵. نشان دهید مرز هر مجموعه دارای محتوای صفر، خود با محتوای صفر است. مثالی ارائه دهید از یک مجموعه با اندازهٔ صفر که مرز آن از اندازهٔ صفر نیست.

۶. فرض کنید $f : [a,b] \longrightarrow \mathbb{R}$ صعودی باشد. نشان دهید مجموعه نقاطی از $[a,b]$ که f در آن پیوسته نیست، از اندازهٔ صفر است.

۷. فرض کنید f و g انتگرال‌پذیر باشند. نشان دهید $f \cdot g$ نیز انتگرال‌پذیر است.

۸. اگر $f : A \longrightarrow \mathbb{R}$ تابعی نامنفی و $\int_A f = \circ$. نشان دهید $\{x \ \mid \ f(x) \neq \circ\}$ از اندازهٔ صفر است.

۹. اگر $f : [a,b] \times [a,b] \longrightarrow \mathbb{R}$ پیوسته باشد، نشان دهید

$$\int_a^b \int_a^y f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \int_x^b f(x,y)\,\mathrm{d}y\,\mathrm{d}x$$

فصل ۵

اندازه و انتگرال لبگ

۱.۵ اندازه

مفهوم اندازه که به‌منظور تعمیم مفاهیم طول و سطح و حجم توسط کانتور[۱]، لبگ[۲] و کاراتئودوری[۳] و سایرین معرفی شد، انقلابی در مفهوم "انتگرال‌گیری ریمان" پدید آورد. این کتاب مجال اولیه‌ای برای پرداختن به این موضوع در $\mathbb{R}^n$ است. کتب سطوح بالاتر آنالیز به مفهوم عام اندازه خواهند پرداخت. مفهوم اندازه به‌شکل گسترده‌ای در آنالیز تابعی، نظریه احتمال، نظریه سیستم‌های دینامیکی و سایر شاخه‌های ریاضیات مورد استفاده قرار می‌گیرد.

در این بخش اندازهٔ یک مجموعهٔ کران‌دار $A \subseteq \mathbb{R}^n$ را تعریف می‌کنیم. ابتدا ساده‌ترین

1. Cantor
2. Lebesgue
3. Carathéodory

حالت را یعنی وقتی که A به‌صورت یک مستطیل nـ بعدی باشد، در نظر می‌گیریم. فرض کنیم

$$A = I_۱ \times \cdots \times I_n$$

که در آن هر I_i $(۱ \leq i \leq n)$ یک بازهٔ بسته در $\mathbb{R}$ است. در این صورت اندازهٔ A را که با $\mu(A)$ نشان می‌دهیم به‌صورت حاصلضرب طول بازه‌های $I_۱$ و $\ldots$ و I_n تعریف می‌کنیم. $\mu(A)$ همان $V(A)$ است که در فصل قبل از آن بهره بردیم. سپس اندازه را برای مجموعه‌هایی که به‌صورت اجتماع تعداد متناهی بازه در $\mathbb{R}^n$ باشند، تعریف می‌کنیم. برای این کار برای هر $i = ۱, \ldots, n$ مجموعهٔ متناهی $p^i = \{x^i_۰, \ldots, x^i_{m_i}\}$ را در نظر می‌گیریم. فرض کنیم p^i_j ابرصفحه‌ای با معادلهٔ $x^i = x^i_j$ و p اجتماع تمام این ابرصفحه‌ها باشد. در این صورت p را یک تور خوانیم. هر تور $\mathbb{R}^n$ را به‌تعداد متناهی بازه، که بازه‌های p خوانده می‌شوند و تعداد متناهی مجموعه‌های بی‌کران تقسیم می‌کند. هر بازه از p به‌صورت $J_۱ \times \cdots \times J_n$ است که $J_i = \left[x^i_{l_i}, x^i_{l_i+۱}\right]$ و برای هر J_i، $۰ \leq l_i \leq m_i$. مجموعهٔ Y را یک مجموعه مقدماتی خوانیم، در صورتی که برابر اجتماع تعداد متناهی $I_۱ \ldots$ و I_p از بازه‌های یک تور p باشد. در این صورت اندازهٔ Y را به‌صورت

$$\mu(Y) = \mu(I_۱) + \cdots + \mu(I_p)$$

تعریف می‌کنیم. برای معتبر بودن این تعریف باید نشان دهیم که اندازهٔ Y مستقل از انتخاب تور است. فرض کنیم p' توری ظریف‌تر از p باشد یعنی $p \subset p'$. اگر p' از افزودن یک ابرصفحه به p حاصل شده باشد به‌وضوح دیده می‌شود که $\mu(Y)$ نسبت به دو تور برابر است و به‌کمک استقرا می‌توان صحت آن را در حالتی که p' هر افراز ظریف‌تر از p باشد، نتیجه گرفت. حال اگر p و p' دو تور باشند به قسمی که Y برابر باشد با اجتماع بازه‌های p و همچنین اجتماع بازه‌های p'، آن‌گاه $p \cup p'$ یک تور ظریف‌تر از p و p' است. پس $\mu(Y)$ نسبت به انتخاب p و p' ثابت باقی خواهد ماند. اگر Y و Z مجموعه‌های مقدماتی باشند،

می‌توان تور p را چنان معین کرد که Y و Z را بتوان به‌صورت اجتماعی از بازه‌های p نوشت. بنابراین $Y \cup Z$ نیز یک مجموعهٔ مقدماتی است. به‌علاوه داریم

$$\mu(Y \cup Z) \leq \mu(Y) + \mu(Z) \ .$$

همچنین تساوی در حالتی برقرار است که $Y \cap Z = \varnothing$. اکنون اندازه را برای یک مجموعهٔ کران‌دار A تعریف خواهیم کرد. برای این کار ابتدا اندازهٔ یک مجموعهٔ باز به‌وسیلهٔ مجموعه‌هایی مقدماتی که "از داخل به آن نزدیک می‌شوند"، و اندازهٔ یک مجموعهٔ فشرده به‌وسیلهٔ مجموعه‌های "مقدماتی که از خارج به آن نزدیک می‌شوند" را تعریف می‌کنیم. سپس A را از داخل به وسیله یک مجموعهٔ فشرده و از خارج به‌وسیلهٔ یک مجموعهٔ باز تقریب می‌کنیم. فرض کنیم G مجموعه‌ای باز باشد. اگر Y مجموعه‌ای مقدماتی مشمول در G باشد، در این صورت باید اندازهٔ G بزرگ‌تراز یا مساوی با $\mu(Y)$ باشد.

تعریف ۵.۱.۱. اندازهٔ مجموعهٔ باز G عبارت است از

$$\mu(G) = \sup\{\mu(Y) \mid Y \subset G \text{ و } Y \text{ یک مجموعهٔ مقدماتی است}\} \ .$$

اگر مجموعهٔ $S = \{\mu(Y) | Y \subset G\}$ از بالا کران‌دار نباشد، قرار می‌دهیم $\mu(G) = +\infty$. مثلاً $\mu(\mathbb{R}^n) = +\infty$. اگر G کران‌دار باشد، آن‌گاه G در مستطیل nـ بعدی مانند I جای دارد و $\mu(I)$ یک کران بالا برای S است. در چنین حالتی $\mu(G)$ متناهی است.

اگر H زیرمجموعهٔ بازی از G باشد و $T = \{\mu(Y) | Y \subset H\}$، آن‌گاه چون $T \subset S$ پس $\sup T \leq \sup S$ و در نتیجه $\mu(H) \leq \mu(G)$.

مثال ۵.۱.۲. فرض کنیم Z یک مجموعهٔ مقدماتی باشد و $G = \text{Int } Z$ (توجه کنید که هر مجموعهٔ مقدماتی بسته است). نشان می‌دهیم که $\mu(\text{Int } Z) = \mu(Z)$. برای هر مجموعهٔ مقدماتی Y اگر $Y \subset \text{Int } Z$ آن‌گاه $\mu(Y) \leq \mu(Z)$. از طرفی برای هر $\varepsilon > 0$ می‌توان مجموعهٔ

مقدماتی $Y \subset \mathrm{Int}\ Z$ را به‌قسمی معین کرد که $\mu(Z) < \mu(Y) + \varepsilon$. بنابراین $\mu(Z) = \sup\{\mu(Y) |\, Y \subset \mathrm{Int}\ Z\}$.

اکنون فرمولی برای مجموع اندازه دو مجموعهٔ باز ثابت می‌کنیم. به این منظور ابتدا یک لم از توپولوژی که حالتی خاص از لم عدد لُبگ است را بیان می‌کنیم.

لم ۳.۱.۵. فرض کنیم G و H دو مجموعهٔ باز و K زیرمجموعه‌ای فشرده از $G \cup H$ باشد. در این صورت عددی مانند r وجود دارد به قسمی که برای هر $x \in K$ گوی باز $B_r(x)$ یا در G واقع است و یا در H.

برهان. مجموعه‌های $\mathbb{R}^n - G$ و $\mathbb{R}^n - H$ بسته‌اند. توابع f و g را به‌صورت

$$f(x) = \mathrm{d}(x, \mathbb{R}^n - G)\ , \quad g(x) = \mathrm{d}(x, \mathbb{R}^n - H)$$

در نظر می‌گیریم که $\mathrm{d}(x, A)$ به معنی فاصلهٔ x از مجموعهٔ A می‌باشد. توابع f و g پیوسته‌اند(ثابت کنید!) و برای هر $x \in G \cup H$، $f(x) + g(x) > ۰$. چون K فشرده است پس تابع پیوستهٔ $f + g$ یک مینیمم مقدار مثبت مانند c روی K دارد. قرار می‌دهیم $r = \frac{c}{۲}$. پس برای هر $x \in K$ یا $f(x) \geq r$ و یا $g(x) > r$. □

لم ۴.۱.۵. فرض کنیم G و H دو مجموعهٔ باز با اندازهٔ متناهی باشند. در این صورت

$$\mu(G \cup H) \leq \mu(G) + \mu(H).$$

برهان. فرض کنیم W یک مجموعهٔ مقدماتی باشد که $W \subset G \cup H$ و فرض کنیم r همان عدد لم قبل برای $K = W$ باشد. مجموعهٔ مقدماتی W اجتماع بازه‌های یک تور مانند p است. با تظریف p در صورت لزوم می‌توان فرض کرد که قطر هر بازهٔ p از r کوچک‌تر

است. فرض کنیم Y و Z به‌ترتیب برابر اجتماع بازه‌هایی از p باشند که مشمول در G و H می‌باشند. باتوجه به لم ۳.۱.۵، $W \subset Y \cup Z$. پس

$$\mu(W) \leq \mu(Y \cup Z) \leq \mu(Y) + \mu(Z) \leq \mu(G) + \mu(H)$$

و چون W یک مجموعهٔ مقدماتی دلخواه بود قضیه ثابت می‌شود. □

اکنون اندازهٔ یک مجموعهٔ فشردهٔ K را به‌وسیلهٔ تقریب آن از خارج به‌وسیلهٔ مجموعه‌های مقدماتی، تعریف می‌کنیم. برای این کار مجموعه‌های مقدماتی Z را در نظر می‌گیریم که درون آن‌ها شامل K باشد.

تعریف ۵.۱.۵. اندازهٔ مجموعهٔ فشردهٔ K عبارت است از

$$\mu(K) = \inf\{\mu(Z) |\, K \subset \operatorname{Int} Z \,, \text{مقدماتی است } Z\}\,.$$

مثال ۶.۱.۵. هر مجموعهٔ مقدماتی Y فشرده است. به‌سادگی ملاحظه می‌شود که تعریف جدید و قدیم $\mu(Y)$ بر هم منطبق‌اند.

لم ۷.۱.۵. فرض کنیم K و L دو مجموعهٔ فشرده باشند، به قسمی که $K \cap L = \varnothing$. در این صورت

$$\mu(K \cup L) \geq \mu(K) + \mu(L).$$

برهان. فرض کنیم $f(x) = \mathrm{d}(x, L)$. چون $K \cap L = \varnothing$ و L بسته است، پس برای هر $x \in K$، $f(x) > ۰$. به‌علاوه K فشرده است و f پیوسته بنابراین f دارای مقدار مینیمم مثبتی مانند r است. فرض کنیم W مجموعهٔ مقدماتی دلخواهی باشد که $K \cup L \subset \operatorname{Int} W$. دیدیم که اندازهٔ یک مجموعهٔ مقدماتی از انتخاب تور مستقل است، پس می‌توان فرض کرد

که W برابر است با اجتماع بازه‌های $I_۱$، ...و I_p که قطر هر کدام کوچک‌تر از $\frac{r}{۲}$ است. فرض کنیم

$$Y = \bigcup \{I_j \mid I_j \cap K \neq \varnothing \ , \quad ۱ \leq j \leq p\} \ ,$$
$$Z = \bigcup \{I_j \mid I_j \cap L \neq \varnothing \ , \quad ۱ \leq j \leq p\} \ .$$

در این صورت $Y \cup Z \subset W$ و چون قطر هر بازه کوچک‌تر از $\frac{r}{۲}$ است، پس $Y \cap Z = \varnothing$. به‌علاوه $K \subset \operatorname{Int} Y$ و $L \subset \operatorname{Int} Z$. بنابراین

$$\mu(K) + \mu(L) \leq \mu(Y) + \mu(Z) = \mu(Y \cup Z) \leq \mu(W) \ ,$$

پس $\mu(K) + \mu(L)$ یک کران پائین برای $\{\mu(W) | \, K \cup L \subset W\}$ است و در نتیجه از اینفیمم مجموعه یعنی $\mu(K \cup L)$ بزرگ‌تر نمی‌باشد. $\square$

تعریف ۸.۱.۵. فرض کنیم A یک مجموعهٔ کران‌دار است. اندازهٔ خارجی A که به $\overline{m}(A)$ نشان داده می‌شود، عبارت است از تقریب آن از خارج به‌وسیلهٔ مجموعه‌های باز. یعنی

$$\overline{m}(A) = \inf\{\mu(G) | \, A \subset G, \text{باز است } G\}$$

و اندازهٔ داخلی A که به $\underline{m}(A)$ نشان داده می‌شود عبارت است از تقریب آن از داخل به‌وسیلهٔ مجموعه‌های فشرده. یعنی

$$\underline{m}(A) = \sup\{\mu(K) | \, K \subset A\} \ .$$

اگر K یک مجموعهٔ فشرده و G مجموعهٔ بازی شامل K باشد، برای هر $x \in K$ یک بازهٔ $I(x)$ وجود دارد که $I(x) \subset G$. باتوجه به‌فشردگی K تعداد متناهی از $\operatorname{Int} I(x)$ ها مجموعهٔ K را می‌پوشاند. پس می‌توان مجموعهٔ مقدماتی Y را چنان معین کرد که

$$K \subset \operatorname{Int} Y \subset Y \subset G \ .$$

بنابراین اگر $K \subset A$ آن‌گاه برای هر G که شامل A باشد داریم $\mu(K) \leq \mu(G)$. پس $\mu(K)$ یک کران پائین برای مجموعهٔ $\{\mu(G) | A \subset G\}$ می‌باشد و $\mu(K) \leq \overline{m}(A)$. بنابراین $\overline{m}(A)$ یک کران بالا برای مجموعهٔ $\{\mu(K) | K \subset A\}$ است، و $\underline{m}(A) \leq \overline{m}(A)$.

تعریف ۹.۱.۵. مجموعهٔ کران‌دار A را اندازه‌پذیر خوانیم، اگر اندازهٔ خارجی و داخلی آن با هم برابر باشند. اگر A اندازه‌پذیر باشد عدد

$$m(A) = \underline{m}(A) = \overline{m}(A)$$

را اندازهٔ (nـ بعدی) A خوانیم.

به‌سادگی می‌توان دید که اگر $B \subset A$ آن‌گاه

$$\underline{m}(B) \subset \underline{m}(A) \qquad \text{و} \qquad \overline{m}(B) \subset \overline{m}(A).$$

اگر H یک مجموعهٔ باز کران‌دار باشد، آن‌گاه برای هر مجموعهٔ باز G که شامل H باشد داریم $\mu(H) \leq \mu(G)$ و تساوی فقط برای $G = H$ برقرار است. بنابراین $\overline{m}(H) = \mu(H)$. به‌علاوه برای هر $\varepsilon > 0$ می‌توان مجموعهٔ مقدماتی $Y \subset H$ را چنان معین کرد که $\mu(H) - \varepsilon < \mu(Y)$. چون Y فشرده است پس $\mu(Y) \leq \underline{m}(H)$ و در نتیجه $\mu(H) - \varepsilon < \underline{m}(H)$، چون این رابطه برای هر $\varepsilon > 0$ برقرار است پس $\mu(H) \leq \underline{m}(H)$. ولی $\underline{m}(H) \leq \overline{m}(H)$ و بنابراین $\underline{m}(H) = \overline{m}(H)$؛ یعنی هر مجموعهٔ کران‌دار باز اندازه‌پذیر است و اندازهٔ آن با اندازه‌ای که قبلاً تعریف شده است، برابر است. به طریق مشابه می‌توان نشان داد (ثابت کنید) که هر مجموعهٔ فشرده اندازه‌پذیر است. مجموعه‌های بسیار دیگری که نه باز هستند و نه فشرده، اندازه‌پذیر می‌باشند. اکنون نشان می‌دهیم که اجتماع، اشتراک و تفاضل تعداد متناهی مجموعه‌های اندازه‌پذیر، اندازه‌پذیر است. ابتدا لم زیر را ثابت می‌کنیم.

لم ۱۰.۱.۵. فرض کنیم A و B مجموعه‌های کران‌دار باشند. در این صورت

$$\overline{m}(A\cup B)\leq \overline{m}(A)+\overline{m}(B)\ .$$

و اگر $A\cap B=\varnothing$ آن‌گاه

$$\underline{m}(A\cup B)\geq \underline{m}(A)+\underline{m}(B)$$

برهان. فرض کنیم $\varepsilon>۰$ داده شده است. مجموعه‌های باز G و H موجودند به‌قسمی که $A\subset G$ و $B\subset H$ و

$$m(G)<\overline{m}(A)+\frac{\varepsilon}{۲}\ ,\qquad m(H)<\overline{m}(B)+\frac{\varepsilon}{۲}.$$

مجموعهٔ $G\cup H$ باز و شامل $A\cup B$ است و بنا بر لم ۴.۱.۵

$$\begin{aligned}\overline{m}(A\cup B)&\leq m(G\cup H)\leq m(G)+m(H)\\&<\overline{m}(A)+\overline{m}(B)+\varepsilon\ .\end{aligned}$$

چون ε دلخواه بود قسمت اوّل قضیه ثابت می‌شود اثبات رابطهٔ دوم مشابه است (با استفاده از لم ۷.۱.۵). □

قضیه ۱۱.۱.۵. فرض کنیم A و B دو مجموعهٔ کران‌دار اندازه‌پذیر باشند به قسمی که $A\cap B=\varnothing$. در این صورت $A\cup B$ اندازه‌پذیر است و

$$m(A\cup B)=m(A)+m(B)\ .$$

برهان. بنا بر لم قبل

$$m(A)+m(B)\leq \underline{m}(A\cup B)\leq \overline{m}(A\cup B)\leq m(A)+m(B)\ ,$$

پس

$$\underline{m}(A \cup B) = \overline{m}(A \cup B).$$

□

از این قضیه به‌کمک استقرا نتیجه می‌گردد:

نتیجه ۶. اگر $\{A_1, \dots, A_r\}$ گردایه‌ای از مجموعه‌های مجزای (برای هر $k \neq l$، $A_k \cap A_l = \varnothing$) کران‌دار اندازه‌پذیر باشد، آن‌گاه $A_1 \cup \dots \cup A_r$ اندازه‌پذیر است و

$$m(A_1 \cup \dots \cup A_r) = \sum_{n=1}^{r} m(A_n) \ .$$

نتیجه ۷. فرض کنیم A مجموعه‌ای کران‌دار است. در این صورت A اندازه‌پذیر است، اگر و فقط اگر برای هر $\varepsilon > 0$ مجموعه‌ای فشرده مانند K و مجموعه‌ای باز مانند G وجود داشته باشد، به‌قسمی که $K \subset A \subset G$ و $m(G - K) < \varepsilon$.

برهان. ساده است و به خواننده واگذار می‌شود (راهنمایی: از خاصیت مشخصهٔ سوپریمم و اینفیمم و این‌که $m(G) = m(G - K) + m(K)$ استفاده کنید). □

قضیه ۱۲.۱.۵. اگر A و B مجموعه‌هایی کران‌دار و اندازه‌پذیر باشند، آن‌گاه $A - B$، $A \cap B$، $A \cup B$ اندازه‌پذیر است.

برهان. ابتدا اندازه‌پذیری $A - B$ را ثابت می‌کنیم. فرض کنیم $\varepsilon > 0$ داده شده است. فرض کنیم G و G' مجموعه‌هایی باز و K و K' مجموعه‌هایی بسته باشند به‌قسمی که $K \subset A \subset G$ و $K' \subset B \subset G'$ و

$$m(G - K) < \frac{\varepsilon}{2} \ , \qquad m(G' - K') < \frac{\varepsilon}{2}.$$

قرار می‌دهیم $H = G - K'$ و $L = K - G'$. در این صورت H باز و L فشرده است و

$$L \subset A - B \subset H.$$

به‌علاوه $H - L$ باز است و

$$H - L \subset (G - K) \cup (G' - K') \; .$$

پس باتوجه به لم ۴.۱.۵،

$$m(H - L) \leq m(G - K) + m(G' - K') < \varepsilon \; .$$

بنابراین $A - B$ اندازه‌پذیر است. حال

$$A \cap B = A - (A - B) \; .$$

چون A و $A - B$ هر دو اندازه‌پذیرند، بنا بر قسمت اوّل قضیه تفاضل آن‌ها یعنی $A \cap B$ اندازه‌پذیر است. در نتیجه

$$A \cup B = (A - B) \cup B \; .$$

هر دو مجموعهٔ سمت راست اندازه‌پذیرند و اشتراک آن‌ها تهی است. پس $A \cup B$ بنا بر قضیهٔ ۱۱.۱.۵ اندازه‌پذیر است. $\square$

نتیجهٔ ۶ را می‌توان برای تعداد نامتناهی مجموعه نیز تعمیم داد. فرض کنیم $\{A_n\}$ دنباله‌ای از مجموعه‌ها باشد. دنبالهٔ $\{A_n\}$ را مجزا خوانیم در صورتی که اگر $k \neq l$، آن‌گاه $A_k \neq A_l$. ابتدا لم زیر را ثابت می‌کنیم.

لم ۱۳.۱.۵. اگر G_1 و G_2 و ... دنباله‌ای از مجموعه‌های باز باشند که هر کدام دارای اندازهٔ متناهی است و فرض کنیم $G = \bigcup\limits_{n=1}^{\infty} \{G_n\}$، آن‌گاه

$$m(G) \leq \sum_{n=1}^{\infty} m(G_n) \ .$$

برهان. فرض کنیم $Y \subset G$ مجموعه‌ای مقدماتی باشد. چون Y فشرده است تعداد متناهی از G_n ها مجموعهٔ Y را می‌پوشاند. پس عددی مانند m وجود دارد که

$$Y \subset G_1 \cup \cdots \cup G_m \ .$$

بنا بر لم ۴.۱.۵ و به کمک استقرا

$$m(G_1 \cup \cdots \cup G_m) \leq \sum_{n=1}^{m} m(G_n)$$

بنابراین

$$m(Y) \leq \sum_{n=1}^{m} m(G_n) \leq \sum_{n=1}^{\infty} m(G_n)$$

و چون Y دلخواه بود لم ثابت می‌شود. □

قضیه ۱۴.۱.۵. اگر $\{A_n\}$ دنباله‌ای از مجموعه‌های مجزای اندازه‌پذیر باشد و اگر $A = \bigcup\limits_{n=1}^{\infty} \{A_n\}$ کران‌دار باشد، آن‌گاه A اندازه‌پذیر است و

$$m(A) = \sum_{n=1}^{\infty} m(A_n).$$

برهان. برای هر m داریم

$$A_1 \cup \cdots \cup A_m \subset A \ .$$

پس بنا بر نتیجهٔ ۶
$$\sum_{n=1}^{m} m(A_n) \leq \underline{m}(A)\ .$$
چون این رابطه برای هر m برقرار است، پس
$$\sum_{n=1}^{\infty} m(A_n) \leq \underline{m}(A)\ .$$
از طرف دیگر فرض کنیم $\varepsilon > 0$ داده شده است و فرض کنیم برای هر عدد طبیعی n، G_n مجموعه‌ای باز باشد، به‌قسمی که $A_n \subset G_n$ و
$$m(G_n) < m(A_n) + \frac{\varepsilon}{2^n}.$$
قرار می‌دهیم
$$G = G_1 \cup G_2 \cup \cdots\ .$$
در این صورت $A \subset G$ و در نتیجه $\overline{m}(A) \leq m(G)$ و بنا بر لم قبل
$$\overline{m}(A) < \sum_{n=1}^{\infty} m(A_n) + \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=1}^{\infty} m(A_n) + \varepsilon\ .$$
چون این رابطه برای هر $\varepsilon > 0$ برقرار است پس A اندازه‌پذیر است و
$$m(A) = \sum_{n=1}^{\infty} m(A_n).$$
□

اگر A_n ها در قضیهٔ قبل مجزا نباشند، رابطهٔ تساوی به‌صورت نامساوی درمی‌آید. برای یک مجموعهٔ بی‌کران A می‌توان مفهوم اندازه‌پذیری را چنین تعریف کرد: فرض کنیم

$B_r = \{x | |x| < 1\}$. مجموعهٔ A را اندازه‌پذیر خوانیم، اگر برای هر $r > 0$، $A \cap B_r$ اندازه‌پذیر باشد، اندازهٔ A عبارت است از

$$m(A) = \lim_{r \to \infty} m(A \cap B_r).$$

اگر تابع φ را به‌صورت $\varphi(r) = m(A \cap B_r)$ تعریف کنیم آن‌گاه φ تابعی صعودی است و بنابراین حد آن موجود است که ممکن است متناهی یا $+\infty$ باشد. اگر A کران‌دار باشد آن‌گاه r_0 ای وجود دارد که $A \subset B_{r_0}$. برای $r \geq r_0$ داریم $A = A \cap B_r$.

قضیه ۱۵.۱.۵. (الف) هر مجموعهٔ باز G با تعریف معادل بالا اندازه‌پذیر است.

(ب) اگر A اندازه‌پذیر باشد، آن‌گاه متمم A $(\mathbb{R}^k - A)$ اندازه‌پذیر است.

(پ) اگر A_1، A_2، ... اندازه‌پذیر باشند، آن‌گاه $\bigcup\limits_{n=1}^{\infty} A_n$ و $\bigcap\limits_{n=1}^{\infty} A_n$ نیز اندازه‌پذیر است.

برهان. فرض کنیم G باز است، در این صورت برای هر $r > 0$، $G \cap B_r$ باز و کران‌دار است. بنابراین برای هر $r > 0$، $G \cap B_r$ اندازه‌پذیر است و در نتیجه G اندازه‌پذیر است و (الف) ثابت می‌شود.

برای اثبات (ب) فرض کنیم A اندازه‌پذیر است. داریم که

$$(\mathbb{R}^k - A) \cap B_r = B_r - (A \cap B_r).$$

چون B_r و $A \cap B_r$ اندازه‌پذیرند بنا بر قضیهٔ ۱۴.۱.۵، $(\mathbb{R}^k - A) \cap B_r$ اندازه‌پذیر است. چون این برای هر $r > 0$ برقرار است پس $\mathbb{R}^k - A$ اندازه‌پذیر است.

برای اثبات (پ) ابتدا فرض کنیم مجموعهٔ

$$A = A_1 \cup A_2 \cup \dots$$

کران‌دار است. قرار می‌دهیم

$$C_1 = A_1 \;,\; C_2 = A_2 - A_1 \;,\; \dots \;,\; C_k = A_k - (A_1 \cup \dots \cup A_{k-1}) \;,\; \cdots .$$

هر C_k اندازه‌پذیر است. به‌علاوه C_k ها مجزا بوده و اجتماعشان برابر است با A پس بنا بر قضیهٔ ۱۱.۱.۵ A اندازه‌پذیر است. حالتی که A کران‌دار نباشد باتوجه به

$$A \cap B_r = (A_۱ \cap B_r) \cup (A_۲ \cap B_r) \cup \cdots$$

به حالتی که A کران‌دار است برمی‌گردد. در نهایت اندازه‌پذیری $\{A_n\}$ $\bigcap\limits_{n=۱}^{\infty}$ از

$$\bigcap_{n=۱}^{\infty} A_n = \mathbb{R}^k - \bigcup_{n=۱}^{\infty}\{(\mathbb{R}^k - A_n)\}$$

نتیجه می‌گردد. □

۲.۵ توابع اندازه‌پذیر

در تمامی این بخش فرض می‌کنیم، f تابعی از $\mathbb{R}^n$ به‌توی مجموعهٔ اعداد حقیقی توسیع یافته است.

تعریف ۱.۲.۵. تابع f را اندازه‌پذیر گوییم، هرگاه برای هر عدد حقیقی c مجموعهٔ $\{x|\, f(x) > c\}$ اندازه‌پذیر باشد.

قضیه ۲.۲.۵. شرایط زیر معادلند

(الف) f اندازه‌پذیر است.

(ب) برای هر عدد حقیقی c مجموعهٔ $\{x|\, f(x) \geq c\}$ اندازه‌پذیر است.

(پ) برای هر عدد حقیقی c مجموعهٔ $\{x|\, f(x) < c\}$ اندازه‌پذیر است.

(ت) برای هر عدد حقیقی c مجموعهٔ $\{x|\, f(x) \leq c\}$ اندازه‌پذیر است.

برهان. چون

$$\{x|\, f(x) \geq c\} = \bigcap_{n=۱}^{\infty} \{x|\, f(x) > c - \frac{۱}{n}\}$$

پس (الف) درستی (ب) را ایجاب می‌کند. چون (پ) متمم (ب) است پس بنا بر قضیهٔ ۱۵.۱.۵، (پ) از (ب) نتیجه می‌گردد. (ت) نیز بنا بر همین قضیه از (الف) نتیجه می‌شود. به علاوه

$$\{x|\, f(x) > c\} = \mathbb{R}^k - \{x|\, f(x) \leq c\}\ .$$

پس (ت) درستی (الف) را ایجاب می‌کند. □

بدیهی است که هر تابع پیوسته (با حوزه تعریف اندازه‌پذیر) اندازه‌پذیر است. همچنین اگر f تابعی اندازه‌پذیر روی مجموعهٔ (اندازه‌پذیر) E باشد، آن‌گاه $f|_E$ نیز اندازه‌پذیر است.

تعریف ۳.۲.۵. یک تابع اندازه‌پذیر f را ساده خوانیم، در صورتی که فقط تعداد متناهی مقدار را اختیار کند.

باتوجه به این تعریف هر تابع سادهٔ f به‌صورت

$$\sum_{i=1}^{n} c_i \chi_i$$

می‌باشد که c_i مقادیر ثابتی است که f اختیار می‌کند و χ_i تابع مشخصهٔ مجموعهٔ اندازه‌پذیر $A_i = \{x|\, f(x) = c_i\}$ است.

قضیه ۴.۲.۵. فرض کنیم k مقداری ثابت و f و g دو تابع حقیقی مقدار اندازه‌پذیر (با حوزه تعریف برابر) باشند. در این صورت kf، $f+g$ و fg اندازه‌پذیر است.

برهان. برای اثبات قضیهٔ ۲.۲.۵ (پ) را به کار می‌گیریم. چون

$$\{x|\, f(x) + k < c\} = \{x|\, f(x) < c - k\}\ ,$$

پس اگر f اندازه‌پذیر باشد آن‌گاه $f+k$ اندازه‌پذیر است. فرض کنیم $f(x)+g(x)<c$، در این صورت $f(x)<c-g(x)$. پس عددی گویا مانند r وجود دارد که $f(x)<r<c-g(x)$ و بنابراین

$$\{x|\, f(x)+g(x)<c\}=\bigcup_{r}(\{x|\, f(x)<r\}\cap\{x|\, g(x)<c-r\}).$$

چون مجموعهٔ اعداد گویا شماراست، پس مجموعهٔ بالا و در نتیجه $f+g$ اندازه‌پذیر است.

برای اثبات اندازه‌پذیری fg، چون

$$fg=\frac{۱}{۲}\left[(f+g)^{۲}-f^{۲}-g^{۲}\right],$$

کافی است نشان دهیم که اگر f اندازه‌پذیر باشد، آن‌گاه $f^{۲}$ اندازه‌پذیر است. فرض کنیم f اندازه‌پذیر است. اگر $c\geq ۰$، آن‌گاه

$$\left\{x|\, f^{۲}(x)>c\right\}=\left\{x|\, f(x)>\sqrt{c}\right\}\cup\left\{x|\, f(x)<-\sqrt{c}\right\}$$

و اگر $c<۰$، آن‌گاه برای هر x از حوزه تعریف f داریم $f^{۲}(x)>c.$ یعنی

$$\{x|\, f^{۲}(x)>c\}=D_f\ .$$

پس در هر حال $f^{۲}$ اندازه‌پذیر است. □

قضیه ۵.۲.۵. فرض کنیم $\{f_n\}$ دنباله‌ای از توابع اندازه‌پذیر (با حوزه تعریف برابر) باشند. در این صورت $\max\{f_۱,\ldots,f_n\}$، $\sup\limits_n f_n$ و $\limsup f_n$ اندازه‌پذیرند (همین‌طور $\min\{f_۱,\ldots,f_n\}$، $\inf\limits_n f_n$ و $\liminf f_n$).

برهان. فرض کنیم $h(x)=\max\{f_۱(x),\ldots,f_n(x)\}$. در این صورت

$$\{x|\, h(x)>c\}=\bigcup_{i=۱}^{n}\{x|\, f_i(x)>c\}.$$

پس اندازه‌پذیری f_iها، اندازه‌پذیری h را ایجاب می‌کند. به‌همین ترتیب اگر $g(x) = \sup_n f_n(x)$، آن‌گاه

$$\{x|\, g(x) > c\} = \bigcup_{n=۱}^{\infty} \{x|\, f_n(x) > c\}$$

و در نتیجه g اندازه‌پذیر است. بالاخره $\limsup f_n = \inf_n \sup_{k\geq n} f_c$ و بنابراین $\limsup f_n$ اندازه‌پذیر است. □

تعریف ۶.۲.۵. گوئیم یک ویژگی تقریباً همه‌جا برقرار است، در صورتی که مجموعهٔ نقاطی که در این ویژگی صدق نمی‌کنند، دارای اندازهٔ صفر باشد.

باتوجه به‌تعریف بالا گوئیم تقریباً همه‌جا $f = g$ در صورتی که f و g دارای حوزهٔ تعریف برابر باشند و $m(\{x|\, f(x) \neq g(x)\}) = ۰$. به همین ترتیب گوئیم $\{f_n\}$ تقریباً همه‌جا همگرا به g است هرگاه مجموعه‌ای با اندازهٔ صفر مانند E وجود داشته باشد که برای هر $x \notin E$ دنبالهٔ $f_n(x)$ همگرا به $g(x)$ باشد.

قضیه ۷.۲.۵. اگر f اندازه‌پذیر باشد و تقریباً همه‌جا $f = g$، آن‌گاه g اندازه‌پذیر است.

برهان. فرض کنیم $E = \{x|\, f(x) \neq g(x)\}$. داریم

$$\begin{aligned}\{x \mid g(x) > c\} &= \{x \mid g(x) > c\} \cup \{x \notin E \mid g(x) > c\} \\ &= \{x \in E \mid g(x) > c\} \cup \{x \mid f(x) > c\}\,.\end{aligned}$$

چون f اندازه‌پذیر است پس اولین مجموعه طرف راست اندازه‌پذیر است. به علاوه مجموعهٔ دیگر طرف راست نیز اندازه‌پذیر است، زیرا زیرمجموعه‌ای از E می‌باشد و $m(E) = ۰$. پس g اندازه‌پذیر است. □

۳.۵ انتگرال لبگ

ابتدا فرض کنیم f تابعی حقیقی مقدار، نامنفی و اندازه‌پذیر است. فرض کنیم I یک بازه در $\mathbb{R}^k$ باشد. مطابق معمول از نظر شهودی $\int_I f$ را که به A نشان می‌دهیم، به‌عنوان حجم مجموعهٔ

$$\{(x, y) \mid x \in I\,,\ \circ < y < f(x)\}$$

در نظر می‌گیریم. انتگرال لبگ یک تابع به‌صورت‌های مختلف تعریف می‌شود که می‌توان نشان داد با یکدیگر معادلند. ما ابتدا روش لبگ را یادآوری می‌کنیم ولی سپس روش دیگری را برمی‌گزینیم. در روش لبگ فرض می‌کنیم برای هر y، $g_f(y) = m\{x \in I \mid f(x) > y\}$. برای هر مجموعهٔ متناهی از اعداد $\{y_j\}$ مانند $\circ < y_\circ < y_۱ < \cdots < y_n < \infty$ عبارت

$$\sum_{j=۱}^{n} g_f(y_{j-۱})(y_j - y_{j-۱}),$$

تقریبی برای A می‌باشد. در این صورت $\int_I f$ را به‌صورت حد عبارت فوق در نظر می‌گیریم که در آن وقتی $n \to \infty$ آن‌گاه $\max(y_j - y_{j-۱}) \to \circ$.

روش دیگری که برای تعریف انتگرال لبگ معمول است، روش توابع ساده می‌باشد. در این روش به‌گونه ای که خواهیم دید، ابتدا همانند توابع پله‌ای انتگرال توابع ساده را تعریف می‌کنیم. سپس انتگرال تابع نامنفی اندازه‌پذیر f به صورت سوپریمم انتگرال‌های توابع سادهٔ h که $h \le f$ تعریف می‌شود. تنها اختلاف آن با انتگرال‌های ریمان (که اختلاف مهمی است) در این است که مجموعه‌هایی که توابع ساده روی آن‌ها مقدار ثابتی دارند، لازم نیست به شکل بازه باشند بلکه فقط باید اندازه‌پذیر باشند. در تعریف انتگرال به روش لبگ و روش توابع ساده می‌توان انتگرال را برای توابعی که اندازه‌پذیر باشند نیز تعریف کرد. ولی در این صورت بسیاری از ویژگی‌های انتگرال‌ها را نمی‌توان ثابت کرد. بنابراین

در تعریف فرض شده است که f اندازه‌پذیر است. فرض کنیم h یک تابع ساده است که مقادیر c_1، ...و c_n را اختیار می‌کند. دیدیم که در این صورت

$$h = \sum c_i \chi_{E_i}$$

که $E_i = \{x | h(x) = c_i\}$. این نمایش از h را نمایش متعارف h خوانیم و مشخص‌کنندهٔ آن است که E_i ها مجزایند و c_i ها متمایز و یا صفرند.

تعریف ۵.۳.۱. فرض کنیم E یک مجموعهٔ اندازه‌پذیر باشد و f تابعی ساده باشد که روی مجموعهٔ اندازه‌پذیر E_i مقدار ثابت c_i را اختیار می‌کند ($i = 1, 2, \ldots, r$) و در خارج آن صفر است. در این صورت تعریف می‌کنیم

$$\int_E f = \sum_{i=1}^{r} c_i m(E \cap E_i) \ .$$

همان‌گونه که لم زیر نشان می‌دهد، اگر E_i را به زیرمجموعه‌های (اندازه‌پذیر) کوچک‌تر و مجزا تقسیم کنیم مقدار انتگرال تغییر نمی‌کند، زیرا m خاصیت جمعی دارد.

لم ۵.۳.۲. با علائم بالا فرض کنیم $\varphi = \sum_{i=1}^{n} c_i \chi_{A_i}$ که $A_i \cap A_j = \varnothing$ برای $i \neq j$ و c_iها متمایز هستند. فرض کنیم هر A_i اندازه‌پذیر است. در این صورت $\int_E \varphi = \sum_{i=1}^{n} c_i m(A_i)$.

برهان. واضح است. $\square$

تعریف ۵.۳.۳. فرض کنیم f یک تابع اندازه‌پذیر نامنفی باشد که روی مجموعهٔ اندازه‌پذیر E تعریف شده است. در این صورت انتگرال f عبارت است از

$$\int_E f = \sup \int_E h.$$

که سوپریمم روی مجموعهٔ تمام توابع سادهٔ نامنفی h که $h \leq f$ گرفته شده است ($h \leq f$ یعنی برای هر $x \in E$، $h(x) \leq f(x)$). در صورتی که f مقادیر مثبت و منفی هر دو را اختیار کند تعریف می‌کنیم

$$\int_E f = \int_E f^+ - \int_E f^-$$

که $f^+(x) = \max\{0, f(x)\}$ و $f^-(x) = \max\{0, -f(x)\}$، البته مشروط بر این‌که یکی ازانتگرال‌های $\int_E f^+$ یا $\int_E f^-$ دارای مقدار متناهی باشد. در غیر این صورت $\int_E f$ تعریف نمی‌شود. اگر $\int_E f^+$ و $\int_E f^-$ هر دو متناهی باشند، f را روی E انتگرال‌پذیر لبگ (و اگر جای ابهام نباشد انتگرال‌پذیر) خوانیم و می‌نویسیم $f \in I(E)$ (وقتی یکی از انتگرال‌های $\int_E f^+$ و $\int_E f^-$ متناهی و دیگری نامتناهی باشد واژهٔ انتگرال‌پذیری به کار نمی‌رود. در این حالت گویند $\int_E f$ وجود دارد با مقدار نامتناهی). توجه می‌کنیم که بنابراین تعریف، انتگرال‌پذیری اندازه‌پذیری را ایجاب می‌کند.

لم ۴.۳.۵. فرض کنیم E یک مجموعهٔ با اندازهٔ متناهی و h و k دو تابع ساده باشند که در خارج E صفر شوند. در این صورت

(الف) برای هر $a, b \in \mathbb{R}$

$$\int_E (ah + bk) = a\int_E h + b\int_E k.$$

(ب) اگر تقریباً همه‌جا $h \geq k$ آن‌گاه

$$\int_E h \geq \int_E k\ .$$

برهان. **(الف)** فرض کنیم $\{A_i\}$ و $\{B_i\}$ به‌ترتیب مجموعه‌های ظاهرشده در نمایش متعارف توابع h و k باشند و فرض کنیم A_0 و B_0 مجموعه‌هایی باشند که به‌ترتیب h و k روی آن‌ها

صفر می‌شوند. در این صورت مجموعه‌های E که به‌صورت مقطع $A_i \cap B_j$ می‌باشند، گردآیه‌ای متناهی از مجموعه‌های مجزا می‌باشند و می‌توان نوشت

$$h = \sum_{k=۱}^{N} a_k \chi_{E_k} , \qquad k = \sum_{k=۱}^{N} b_k \chi_{E_k}$$

که N به‌تعداد E_k ها است. بنابراین

$$ah + bk = \sum_{k=۱}^{N} (aa_k + bb_k)\chi_{E_k} = a\sum_{k=۱}^{N} a_k \chi_{E_k} + b\sum_{c=۱}^{N} b_c \chi_{E_k} .$$

پس باتوجه به لم ۲.۳.۵، نتیجه می‌شود که

$$\int_E ah + bk = a\int_E h + b\int_E K .$$

(ب) باتوجه به تعریف به‌وضوح اگر برای تابع ساده مانند f داشته باشیم تقریباً همه‌جا $f \geq ۰$ آن‌گاه $\int_E f \geq ۰$. حال توجه می‌کنیم که چون تقریباً همه‌جا $h - k \geq ۰$ پس

$$\int_E h - \int_E k = \int_E (h-k) \geq ۰.$$

□

قضیه ۵.۳.۵. اگر E اندازه‌پذیر و f و g توابع اندازه‌پذیر نامنفی یا توابع انتگرال‌پذیر روی E باشند و تقریباً همه‌جا $f \leq g$ آن‌گاه،

$$\int_E f \leq \int_E g .$$

برهان. اگر f و g نامنفی باشند، فرض کنیم h یک تابع ساده باشد و $h \leq f$. در این صورت $h \leq g$. پس باتوجه به تعریف ۳.۳.۵ $\int_E h \leq \int_E g$ و در نتیجه

$$\int_E f = \sup_{h \leq f} \int_E h \leq \int_E g.$$

اگر f و g انتگرال‌پذیر ولی مقادیر مثبت و منفی هر دو اختیار کنند، آن‌گاه $f \leq g$ ایجاب می‌کند که $f^+ \leq g^+$ و $f^- \geq g^-$. پس بنا بر قسمت اوّل برهان $\int_E f^+ \leq \int_E g^+$ و $\int_E f^- \geq \int_E g^-$. چون f و g انتگرال‌پذیرند، پس تمام این عبارت‌ها متناهی‌اند و در نتیجه می‌توان آن‌ها را از یکدیگر کم کرد. بنابراین

$$\int_E f = \int_E f^+ - \int_E f^- \leq \int_E g^+ - \int_E g^- = \int_E g.$$

□

قضیه ۶.۳.۵. فرض کنیم E و $A \subset E$ اندازه‌پذیر باشند و f تابعی اندازه‌پذیر و نامنفی و χ_A تابع مشخصهٔ A باشد، آن‌گاه

$$\int_A f = \int_E \chi_A f.$$

برهان. داریم

$$\begin{aligned}\int_A f &= \sup\left\{ \int_A h \mid \text{ساده است } h \text{ و } A \text{ روی } 0 \leq h \leq f \right\}\\ &= \sup\left\{ \int_A h \mid \text{ساده است } h \text{ و } A \text{ روی } 0 \leq h \leq \chi_A f \right\}\\ &= \sup\left\{ \int_E h \mid h \leq \chi_A f \right\} = \int_E \chi_A f\,.\end{aligned}$$

□

قضیه ۷.۳.۵. فرض کنیم E اندازه‌پذیر و α عددی ثابت باشد. در این صورت اگر روی E، f نامنفی و انتگرال‌پذیر باشد، آن‌گاه

$$\int_E \alpha f = \alpha \int_E f\,.$$

برهان. برای توابع ساده قضیه به‌وضوح برقرار است. بنابراین باتوجه به تعریف ۳.۳.۵ و خواص سوپریمم برای هر تابع f نتیجه به‌دست می‌آید (ثابت کنید). □

قضیه ۸.۳.۵ فرض کنیم $\{A_n | n = ۱, ۲, \ldots\}$ دنباله ای صعودی (یعنی برای هر n، $A_n \subset A_{n+۱}$) از مجموعه‌های اندازه‌پذیر و h تابع نامنفی ساده‌ای روی $E = \bigcup_{n=۱}^{\infty} A_n$ باشد. در این صورت

$$\int_E h = \lim_{n\to\infty} \int_{A_n} h \,.$$

برهان. فرض کنیم مقدار h روی مجموعهٔ اندازه‌پذیر B_j $(۱ \leq j \leq r)$ برابر c_j باشد. در این صورت

$$\int_{A_n} h = \sum_{j=۱}^{r} c_j m(A_n \cap B_j),$$

که با توجه به لم ۲.۳.۵ وقتی $n \to \infty$ این مجموع به

$$\sum_{j=۱}^{r} c_j m(E \cap B_j) = \int_E h.$$

میل می‌کند. □

قضیه ۹.۳.۵ اگر f تابعی پیوسته روی I باشد، آن‌گاه برای هر دنبالهٔ $\{I_n\}$ از بازه‌ها به‌قسمی که $a \in I_n \subset I$ داریم

$$\lim_{n\to\infty} \frac{۱}{m(I_n)} \int_{I_n} f = f(a) \,.$$

برهان. بنا بر پیوستگی، برای هر $\varepsilon > ۰$ عددی مانند $\delta(\varepsilon) > ۰$ وجود دارد که

$$|x - a| < \delta(\varepsilon) \quad \Rightarrow \quad f(a) - \varepsilon < f(x) < f(a) + \varepsilon \,.$$

حال بنا بر قضیهٔ ۵.۳.۵

$$\int_{I_n}(f(a)-\varepsilon) \leq \int_{I_n} f \leq \int_{I_n}(f(x)+\varepsilon)\ .$$

پس باتوجه به قضیهٔ ۷.۳.۵

$$(f(a)-\varepsilon)m(I_n) \leq \int_{I_n} f \leq (f(a)+\varepsilon)m(I_n)$$

و چون ε دلخواه بود

$$\lim_{n\to\infty}\frac{۱}{m(I_n)}\int_{I_n} f = f(a)\ .$$

□

قضیه ۱۰.۳.۵. فرض کنیم $\{f_n\}$ دنباله‌ای صعودی از توابع اندازه‌پذیر نامنفی روی E باشد. اگر تقریباً همه‌جا روی E، $f_n(x) \to f(x)$، آن‌گاه

$$\int_E f = \lim \int_E f_n\ .$$

برهان. باتوجه به این‌که انتگرال روی یک مجموعه با اندازهٔ صفر برابر صفر است، پس بدون آن‌که به کلیت استدلال خللی وارد شود، می‌توان فرض کرد که همگرایی روی تمام مجموعه است. (فرض کنیم $g = \sup f_n$ در این‌صورت g اندازه‌پذیر است (ثابت کنید!) پس $\int g$ تعریف می‌شود.) چون $f_n \leq f_{n+۱}$، پس $\int f_n \leq \int f_{n+۱}$ و در نتیجه $\lim_{n\to\infty}\int f_n$ (در $\mathbb{R}^*$) موجود است. به‌علاوه به ازای هر n، $f_n \leq f$ زیرا $\{f_n\}$ صعودی است. بنابراین

$$\lim_{n\to\infty}\int_E f_n \leq \int_E f\ .$$

از طرف دیگر فرض کنیم h یک تابع اندازه‌پذیر است به‌قسمی که $۰ \leq h \leq f$. فرض کنیم $\varepsilon > ۰$ دلخواه و از این به بعد، ثابت است. برای هر n مجموعهٔ A_n را به‌صورت

$$A_n = \{x \in E |\ f_n(x) \geq (۱-\varepsilon)h(x), \}$$

تعریف می‌کنیم. بدیهی است که برای هر n، $A_n \subset A_{n+1}$ و $E = \bigcup_{n=1}^{\infty} A_n$. به‌علاوه

$$\int_E f_n \geq \int_{A_n} f_n \geq (1-\varepsilon)\int_{A_n} h \,.$$

بنابراین

$$\lim_{n\to\infty}\int_E f_n \geq (1-\varepsilon)\lim_{n\to\infty}\int_{A_n} h = (1-\varepsilon)\int_E h \,.$$

چون تابع سادهٔ h دلخواه بود پس $\lim_{n\to\infty}\int f_n \geq \int f$ و برهان کامل می‌شود. □

به‌کمک این قضیه می‌توان انتگرال برخی از توابع را محاسبه نمود.

مثال ۱۱.۳.۵. فرض کنید f تابعی باشد که روی بازهٔ $[0,1]$ به‌صورت

$$f(x) = \begin{cases} \dfrac{1}{\sqrt[3]{x}} & 0 < x \leq 1 \\ 0 & x = 0 \end{cases}$$

تعریف شده است. انتگرال لبگ f روی $[0,1]$ را محاسبه می‌کنیم. به‌این منظور دنبالهٔ $\{f_n\}$ از توابع را که برای هر n به‌صورت

$$f_n(x) = \begin{cases} f(x) & \text{اگر } 0 \leq f(x) \leq n \\ n & \text{اگر } f(x) > n \end{cases}$$

تعریف می‌شود، در نظر می‌گیریم. یعنی اگر $x \geq \frac{1}{n^3}$ آن‌گاه $f_n(x) = \frac{1}{\sqrt[3]{x}}$ و اگر $x < \frac{1}{n^3}$ آن‌گاه $f_n(x) = n$. پس

$$\begin{aligned}\int_0^1 f_n &= \int_0^{\frac{1}{n^3}} n\,\mathrm{d}x + \int_{\frac{1}{n^3}}^1 \frac{1}{\sqrt[3]{x^3}}\mathrm{d}x \\ &= nx\Big|_0^{\frac{1}{n^3}} + \frac{3}{2}x^{\frac{2}{3}}\Big|_{\frac{1}{n^3}}^1 = \frac{1}{n^2} + \frac{3}{2}(1-\frac{1}{n^2}) = \frac{3}{2} - \frac{1}{2n^2}\,.\end{aligned}$$

پس
$$\lim_{n\to\infty}\int_{۰}^{۱} f_n = \frac{۳}{۲}$$
و چون $\{f_n\}$ صعودی است پس
$$\int_{۰}^{۱} f = \frac{۳}{۲}.$$

۴.۵ تمرین

۱. فرض کنید E مربع واحد بسته باشد ثابت کنید

(الف) هر زیرمجموعه باز E اندازه‌پذیر است.

(ب) هر زیر مجموعه بسته E اندازه‌پذیر است.

۲. نشان دهید $\mathbb{Q} \subseteq \mathbb{R}$ اندازه‌پذیر و دارای اندازهٔ صفر است.

۳. نشان دهید تعریف ارائه شده پیش از قضیه ۱۵.۱.۵، برای مجموعه‌های کران‌دار، با تعریف اولیه معادل است.

۴. نشان دهید مجموعهٔ کانتور اندازه‌پذیر و دارای اندازهٔ صفر است.

۵. ثابت کنید هر مجموعه با اندازه مثبت در $[۰, ۱]$ شامل دو نقطه است که به اندازه یک عدد گویا فاصله دارند.

۶. نشان دهید تابع
$$f(x) = \begin{cases} ۱ & x \in \mathbb{Q} \\ ۰ & x \notin \mathbb{Q} \end{cases}$$
روی هر بازهٔ $[a, b]$ اندازه‌پذیر است. سپس نشان دهید انتگرال‌پذیر لبگ نیز می‌باشد.

۷. نشان دهید اگر f اندازه‌پذیر باشد آن‌گاه $|f|$ نیز اندازه‌پذیر است.

۸. نشان دهید حد تقریباً همه‌جا از یک دنباله از توابع اندازه‌پذیر یکتاست.

۹. (قضیه همگرایی تسلطی لبگ) فرض کنیم f_n دنباله‌ای از توابع همگرای تقریباً همه‌جا به تابع f روی مجموعه A باشد و فرض کنیم

$$|f_n(x)| \leq \varphi(x) \quad (x \in A\ ,\ n = ۱, ۲, ...)$$

که در آن f روی A انتگرال‌پذیر لبگ است. در این‌صورت f انتگرال‌پذیر است و

$$\lim_{n\to\infty} \int f_n(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x\ .$$

فصل ۶

انتگرال‌گیری روی زنجیره‌ها

هدف اصلی این فصل ارائه نمادها، تعاریف و قضایایی است که صورت تعمیم یافته قضیه "استوکس" را ارائه کند. به‌قضیه استوکس روی $\mathbb{R}^۲$، در ریاضی عمومی مقدماتی پرداخته می‌شود و رابطه بین انتگرال‌گیری روی نواحی و مرز نواحی را به هم مربوط می‌کند.

۱.۶ مقدمات جبری

فرض کنیم V یک فضای برداری (روی $\mathbb{R}$) باشد، حاصلضرب دکارتی $\overbrace{V \times \cdots \times V}^{k \text{ مرتبه}}$ را با V^k نمایش می‌دهیم.

تابع $T : V^k \longrightarrow \mathbb{R}$ را چندخطی‌می‌نامیم، هرگاه برای هر $۱ \leq i \leq k$ داشته باشیم

$$T(v_۱, \ldots, v_i + v_{i'}, \ldots, v_k) = T(v_۱, \ldots, v_i, \ldots, v_k) + T(v_۱, \ldots, v_{i'}, \ldots, v_k)$$

و

$$T(v_۱, \ldots, \alpha v_i, \ldots, v_k) = \alpha T(v_۱, \ldots, v_k)\ .$$

یک نگاشت چندخطی، یک kـ تانسور روی V نامیده می‌شود. مجموعه همهٔ kـ تانسورها که با $\mathcal{T}^k(V)$ نمایش داده می‌شود، یک فضای برداری روی $\mathbb{R}$ است. زیرا اگر $S,T \in \mathcal{T}^k(V)$ و $\alpha \in \mathbb{R}$ تعریف می‌کنیم

$$(S+T)(v_1, \ldots, v_k) = S(v_1, \ldots, v_k) + T(v_1, \ldots, v_k)$$

$$(\alpha S)(v_1, \ldots, v_k) = \alpha \cdot S(v_1, \ldots, v_k) \ .$$

همچنین عملی وجود دارد که فضاهای مختلف $\mathcal{T}^k(V)$ را به‌هم مربوط می‌کند. اگر $S \in \mathcal{T}^k(V)$ و $T \in \mathcal{T}^k(V)$. ضرب تانسوری $S \otimes T \in \mathcal{T}^{k+l}(V)$ را به‌صورت زیر تعریف می‌کنیم

$$S \otimes T(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}) = S(v_1, \ldots, v_k)T(v_{k+1}, \ldots, v_{k+l}) \ .$$

به‌وضوح ترتیب در عمل بالا نقش اساسی دارد به‌این معنا که $S \otimes T$ و $T \otimes S$ با هم برابر نیستند. بررسی خواص زیر ساده است و به‌عنوان تمرین به‌خواننده واگذار می‌شود

$$(S_1 + S_2) \otimes T = S_1 \otimes T + S_2 \otimes T,$$
$$S(T_1 + T_2) = (S \otimes T_1) + (S \otimes T_2),$$
$$(\alpha S) \otimes T = S \otimes (\alpha T) = \alpha(S \otimes T),$$
$$(S \otimes T) \otimes U = S \otimes (T \otimes U) \ .$$

$(S \otimes T) \otimes U$ و $S \otimes (T \otimes U)$ معمولاً با $S \otimes T \otimes U$ نمایش داده می‌شود. حاصلضرب‌های مرتبهٔ بالاتر $T_1 \otimes \cdots \otimes T_r$ به‌طریق مشابه تعریف می‌شوند.

خواننده تاکنون دریافته است که $\mathcal{T}^1(V)$ دقیقاً V^* ، فضای دوگان است. عمل $\otimes$ ما را قادر می‌سازد که فضاهای $\mathcal{T}^k(V)$ را بر حسب $\mathcal{T}^1(V)$ بیان کنیم.

قضیه ۱.۱.۶. فرض کنیم $\{v_1,\dots,v_k\}$ پایه فضای برداری V و $\{\phi_1,\dots,\phi_n\}$ پایه دوگان باشد و $\phi_i(v_i)=\delta_{ij}$. در این‌صورت مجموعهٔ همهٔ ضرب‌های تانسوری kـگانه

$$\phi_{i_1}\otimes\dots\otimes\phi_{i_k}\qquad 1\leq i_1,\dots,i_k\leq n$$

پایه $\mathcal{T}^k(v)$ است و لذا بُعد این فضا n^k است.

برهان. توجه می‌کنیم که

$$\begin{aligned}\phi_{i_1}\otimes\dots\otimes\phi_{i_k}\Big(v_{j_1},\dots,v_{j_k}\Big)&=\delta_{i_1,j_1},\dots,\delta_{i_k,j_k}\\&=\begin{cases}1 & \text{اگر } i_1=j_1,\dots,i_k=j_k\\ 0 & \text{در غیر این‌صورت}\end{cases}\end{aligned}$$

برای بردارهای $w_1,\dots,w_k$ با نمایش $w_i=\sum_{j=1}^n a_{ij}v_j$ و $T\in\mathcal{T}^k(V)$ آنگاه،

$$\begin{aligned}T(w_1,\dots,w_k)&=\sum_{j_1,j_k=1}^{n}a_{1,j_1}\cdot\ldots\cdot a_{k,j_k}T\Big(v_{j_1},\dots,v_{j_k}\Big)\\&=\sum_{i_1,\dots,i_k=1}^{n}T\Big(v_{i_1},\dots,v_{i_k}\Big)\phi_{i_1}\otimes\cdots\otimes\phi_{i_k}(w_1,\dots,w_k)\end{aligned}$$

بنابراین

$$T=\sum_{i_1,\dots,i_k=1}^{n}T\Big(v_{i_1},\dots,v_{i_k}\Big)\phi_{i_1}\otimes\cdots\otimes\phi_{i_k}$$

و در نتیجه $\phi_{i_1}\otimes\cdots\otimes\phi_{i_k}$ فضای $\mathcal{T}^k(V)$ را تولید می‌کند.

حال فرض می‌کنیم اعداد $a_{i_1,\dots,i_k}$ چنان باشند که

$$\sum_{i_1,\dots,i_k=1}^{n}a_{i_1,\dots,i_k}\cdot\phi_{i_1}\otimes\cdots\otimes\phi_{i_k}=0\,.$$

حال دو طرف را بر $\Big(v_{j_1},\dots,v_{j_k}\Big)$ اثر می‌دهیم. لذا $a_{i_1,\dots,i_k}=0$ و بنابراین $\phi_{i_1}\otimes\cdots\otimes\phi_{i_k}$ ها مستقل خطی هستند. □

یک ساختار آشنای فضاهای دوگان روی فضای تانسورها هم وجود دارد. اگر $f: V \longrightarrow W$ یک نگاشت خطی باشد، نگاشت خطی $f^*: \mathcal{T}^k(W) \longrightarrow \mathcal{T}^k(V)$ به‌صورت

$$f^*T(v_1, \dots, v_k) = T\big(f(v_1), \dots, f(v_k)\big),$$

تعریف می‌شود که در آن $T \in \mathcal{T}^k(W)$ و $v_1, \dots, v_k \in V$. به‌راحتی دیده می‌شود که $f^*(T \otimes S) = f^*(T) \otimes f^*(S)$.

خواننده از پیش با برخی از تانسورها مانند اعضای V^* آشنایی دارد. اولین مثال ضرب داخلی $\langle \cdot, \cdot \rangle \in \mathcal{T}^2(\mathbb{R}^n)$ است.

در اینجا به‌منظور تعمیم بخشی به‌مفهوم ضرب داخلی، تعریف می‌کنیم:
T روی فضای برداری V یک ضرب داخلی است، هرگاه یک ۲ـ تانسور متقارن باشد؛ به این معنا که برای هر $v, w \in V$، $T(v,w) = T(w,v)$. همچنین معین مثبت باشد یعنی $T(v,v) > 0$ برای $v \neq 0$. قضیه ذیل نشان می‌دهد این تعمیم خیلی هم کلی نیست.

قضیه ۲.۱.۶. اگر T یک ضرب داخلی روی V باشد، آن‌گاه پایه‌ای چون $\{v_1, \dots, v_n\}$ برای V چنان موجود است که $T\big(v_i, v_j\big) = \delta_{ij}$ (چنین پایه‌ای نسبت به T متعامد خوانده می‌شود). در نتیجه یکریختی $f: \mathbb{R}^n \longrightarrow V$ چنان موجود است که برای هر $x, y \in \mathbb{R}^n$، $T(f(x), f(y)) = \langle x, y \rangle$. به‌عبارتی $f^* \cdot T = \langle \cdot, \cdot \rangle$.

برهان. فرض کنیم $\{w_1, \dots, w_n\}$ یک پایه برای V باشد. تعریف می‌کنیم

$$w'_1 = w_1\ , \quad w'_2 = w_2 - \frac{T\big(w'_1, w_2\big)}{T\big(w'_1, w'_1\big)} \cdot w'_1\ ,$$
$$w'_3 = w_3 - \frac{T\big(w'_1, w_3\big)}{T\big(w'_1, w'_1\big)} \cdot w'_1 - \frac{T\big(w'_2, w_3\big)}{T\big(w'_2, w'_2\big)} \cdot w'_2\ \ ,\ \ \dots .$$

در این‌صورت $T(w'_i, w'_j) = 0$ برای $i \neq j$ و برای $w'_i \neq 0$، $T(w'_i, w'_i) > 0$. حال قرار می‌دهیم

$$v_i = \frac{w'_i}{\sqrt{T(w'_i, w'_i)}}$$

و تعریف کنیم $f(e_i) = v_i$. □

باوجود اهمیت آن، ضرب داخلی نسبت به nـ تانسور دترمینان نقش کمتری ایفا می‌کند. در تلاش برای تعمیم مناسب مفهوم دترمینان یادآوری می‌کنیم که جابجایی دو سطر یک ماتریس علامت دترمینان را تغییر می‌دهد. این اتفاق الهام بخش تعمیم زیر است.

k ـ تانسور $w \in \mathcal{T}^k(V)$ متناوب خوانده می‌شود هرگاه

$$w(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -w(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)\ , \quad v_1, \ldots, v_k \in V$$

(در این تساوی v_j و v_i جابجا شده‌اند و مابقی بردارها ثابت مانده‌اند.)

مجموعه همهٔ kـ تانسورهای متناوب یک زیر فضای $\mathcal{T}^k(V)$ است که با $\Lambda^k(V)$ نمایش داده می‌شود.

نمایش kـ تانسورهای متناوب راحت نیست. هر چند که یک راه عمومی برای نمایش همهٔ آن‌ها موجود است. علامت یک جایگشت σ که با $\operatorname{Sgn}\sigma$ نمایش داده می‌شود ۱+ است، اگر σ زوج و ۱ـ است اگر فرد باشد. یادآوری می‌کنیم یک جایگشت زوج نامیده می‌شود، هرگاه تجزیه آن به تعداد زوج جایگشت‌های دوری امکان‌پذیر باشد. برای $T \in \mathcal{T}^k(V)$، $\operatorname{Alt}(T)$ با تساوی زیر تعریف می‌شود:

$$\operatorname{Alt}(T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{Sgn}\sigma \cdot T\Big(v_{\sigma(1)}, \ldots, v_{\sigma(k)}\Big),$$

که در آن S_k همهٔ جایگشت‌های اعداد ۱ تا k است.

قضیه ۳.۱.۶. **(الف)** اگر $T \in \mathcal{T}^k(V)$، آن‌گاه $\mathrm{Alt}(T) \in \Lambda^k(V)$.

(ب) اگر $w \in \Lambda^k(V)$، آن‌گاه $\mathrm{Alt}(w) = w$.

(ج) اگر $T \in \mathcal{T}^k(V)$، آن‌گاه $\mathrm{Alt}(\mathrm{Alt}(T)) = \mathrm{Alt}(T)$.

برهان. **(الف)** فرض کنیم (i,j) جایگشت روی $۱, \dots, k$ باشد که i, j را جابجا می‌کند و مابقی را ثابت نگه می‌دارد. اگر $\sigma \in S_k$ باشد قرار می‌دهیم $\sigma' := \sigma \cdot (i,j)$. ، آن‌گاه

$$\begin{aligned}
&\mathrm{Alt}(T)(v_۱, \dots, v_i, \dots, v_j, \dots, v_k) \\
&= \frac{۱}{k!} \sum_{\sigma \in S_k} \mathrm{Sgn}\,\sigma T\Big(v_{\sigma(۱)}, \dots, v_{\sigma(j)}, \dots, v_{\sigma(i)}, \dots, v_{\sigma(k)}\Big) \\
&= \frac{۱}{k!} \sum_{\sigma \in S_k} \mathrm{Sgn}\,\sigma T\Big(v_{\sigma'(۱)}, \dots, v_{\sigma'(i)}, \dots, v_{\sigma'(j)}, \dots, v_{\sigma'(k)}\Big) \\
&= \frac{۱}{k!} \sum_{\sigma \in S_k} -\mathrm{Sgn}\,\sigma' T\Big(v_{\sigma'(۱)}, \dots, v_{\sigma'(i)}, \dots, v_{\sigma'(j)}, \dots, v_{\sigma'(k)}\Big) \\
&= \frac{۱}{k!} \sum_{\sigma' \in S_k} -\mathrm{Sgn}\,\sigma' T\Big(v_{\sigma'(۱)}, \dots, v_{\sigma'(k)}\Big) \\
&= -\mathrm{Alt}(T)(v_۱, \dots, v_k)\ .
\end{aligned}$$

(ب) اگر $w \in \Lambda^k(V)$ و $\sigma = (i,j)$، آن‌گاه $w\big(v_{\sigma(۱)}, \dots, v_{\sigma(k)}\big) = \mathrm{Sgn}\,\sigma \cdot w(v_۱, \dots, v_k)$. از آن‌جایی که هر σ حاصلضرب جایگشت‌هایی به‌صورت (i,j) است، این رابطه برای هر جایگشت σ درست است. بنابراین

$$\begin{aligned}
\mathrm{Alt}(w)(v_۱, \dots, v_k) &= \frac{۱}{k!} \sum_{\sigma \in S_k} \mathrm{Sgn}\,\sigma \cdot w\big(v_{\sigma(۱)}, \dots, v_{\sigma(k)}\big) \\
&= \frac{۱}{k!} \sum_{\sigma \in S_k} \mathrm{Sgn}\,\sigma \cdot \mathrm{Sgn}\,\sigma w\big(v_۱, \dots, v_k\big) \\
&= w(v_۱, \dots, v_k)\ .
\end{aligned}$$

(ج) از قسمت (الف) و (ب) حاصل می‌شود. □

برای تعیین بعد فضای $\Lambda^k(V)$ به یک قضیه مشابه ۱.۱.۶ علاقه‌مند هستیم. البته اگر $w \in \Lambda^k(V)$ و $\eta \in \Lambda^l(V)$ آن‌گاه $w \otimes \eta$ لزوماً به $\Lambda^{k+l}(V)$ متعلق نیست. بنابراین ضرب جدیدی در این فضا تعریف می‌کنیم که "ضرب وج" نامیده می‌شود و به‌صورت زیر تعریف می‌شود

$$w \wedge \eta = \frac{(k+l)!}{k!l!} \mathrm{Alt}(w \otimes \eta) \ .$$

دراین‌صورت $w \wedge \eta \in \Lambda^{k+l}(V)$ (ضریب ظاهر شده به زودی توجیح خواهد شد). بررسی خواص ساده زیر به‌خواننده واگذار می‌شود:

$$(w_۱ + w_۲) \wedge \eta = w_۱ \wedge \eta + w_۲ \wedge \eta,$$
$$w \wedge (\eta_۱ + \eta_۲) = w \wedge \eta_۱ + w \wedge \eta_۲,$$
$$aw \wedge \eta = w \wedge a\eta = a(w \wedge \eta),$$
$$w \wedge \eta = (-۱)^{kl} \eta \wedge w,$$
$$f^*(w \wedge \eta) = f^*(w) \wedge f^*(\eta).$$

تساوی $w \wedge (\eta \wedge \theta) = (w \wedge \eta) \wedge \theta$ برقرار است اما نیاز به محاسبات بیشتری دارد که در قضیه ذیل دیده می‌شود.

قضیه ۴.۱.۶. **(۱)** اگر $S \in \mathcal{T}^k(V)$ و $T \in \mathcal{T}^l(V)$ و $\mathrm{Alt}(S) = ۰$، آن‌گاه

$$\mathrm{Alt}(S \otimes T) = \mathrm{Alt}(T \otimes S) = ۰ \ .$$

(۲) $\mathrm{Alt}(\mathrm{Alt}(w \otimes \eta) \otimes \theta) = \mathrm{Alt}(w \otimes \eta \otimes \theta) = \mathrm{Alt}(w \otimes \mathrm{Alt}(\eta \otimes \theta))$

(۳) اگر $w \in \Lambda^k(V)$ و $\eta \in \Lambda^l(V)$ و $\theta \in \Lambda^m(V)$، آن‌گاه

$$(w \wedge \eta) \wedge \theta = w \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k!\ l!\ m!} \mathrm{Alt}(w \otimes \eta \otimes \theta) \ .$$

به‌صورت طبیعی $(w \wedge \eta) \wedge \theta = w \wedge (\eta \wedge \theta)$ را با $w \wedge \eta \wedge \theta$ نمایش می‌دهیم و ضرب مراتب بالاتر $w_۱ \wedge \cdots \wedge w_r$ نیز به‌شکل مشابه تعریف می‌شود. اگر $\{v_۱, \ldots, v_n\}$ پایه فضای برداری V و $\{\phi_۱, \ldots, \phi_n\}$ پایه فضای دوگان باشند، پایه فضای $\Lambda^k(V)$ در قضیه ذیل به سادگی ساخته می‌شود.

قضیه ۶.۱.۵. مجموعه همهٔ $\phi_{i_۱} \wedge \cdots \wedge \phi_{i_k}$ برای $۱ \leq i_۱ < \cdots < i_k \leq n$، پایه فضای $\Lambda^k(V)$ است و لذا بُعد این فضا $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ است.

برهان. اگر $w \in \Lambda^k(V)$، آن‌گاه $w \in \mathcal{T}^k(V)$. لذا

$$w = \sum_{۱ \leq i_۱, \ldots, i_k \leq n} a_{i_۱, \ldots, i_k} \phi_{i_۱} \otimes \cdots \otimes \phi_{i_k} \ ,$$

بنابراین

$$w = \mathrm{Alt}(w) = \sum_{i_۱, \ldots, i_k} a_{i_۱, \ldots i_k} \mathrm{Alt}\big(\phi_{i_۱} \otimes \cdots \otimes \phi_{i_k}\big) \ .$$

اما $\mathrm{Alt}\big(\phi_{i_۱} \otimes \cdots \otimes \phi_{i_k}\big)$ مضربی از $\phi_{i_۱} \wedge \cdots \wedge \phi_{i_k}$ است. $\square$

اگر V دارای بعد n باشد، از قضیهٔ بالاتر نتیجه می‌شود که $\Lambda^n(V)$ دارای بعد ۱ است. یعنی هر nـ تانسور متناوب مضرب یک nـ تانسور ناصفر است. از آن‌جا که دترمینان مثالی از چنین تابعی است لذا آن‌چه در قضیه ذیل می‌آید، عجیب نیست.

قضیه ۶.۱.۶. فرض کنیم $\{v_۱, \ldots, v_n\}$ پایه‌ای برای فضای V باشند و $w \in \Lambda^n(V)$. اگر $w_i = \sum_{j=۱}^n a_{ij} v_j$، nـ بردار در V باشند، آن‌گاه

$$w(w_۱, \ldots, w_n) = \det(a_{ij}) \cdot w(v_۱, \ldots, v_n) \ .$$

برهان. تعریف می‌کنیم

$$\eta\big((a_{11},...,a_{1n}),\dots,(a_{n1},\dots,a_{nn})\big) = w\left(\sum_{j=1}^{n} a_{1j}v_j,\dots,\sum_{j=1}^{n} a_{nj}v_j\right).$$

در این صورت $\eta \in \mathcal{T}^n(\mathbb{R}^n)$ و به وضوح $\eta \in \Lambda^n(\mathbb{R}^n)$، لذا η مضربی از دترمینان است. یعنی

$$\lambda \in \mathbb{R} \quad \eta = \lambda \cdot \det$$

که در آن $\lambda = \eta(e_1,\dots,e_n) = w(v_1,\dots,v_n)$. □

قضیه بالا نشان می‌دهد که عنصر ناصفر $w \in \Lambda^n(V)$ ، پایه‌های V را به دو گروه مجزا تقسیم می‌کند؛ گروهی که $w(v_1,\dots,v_n) > 0$ و همچنین گروهی که $w(v_1,\dots,v_n) < 0$.

اگر $\{v_1,\dots,v_n\}$ و $\{w_1,\dots,w_n\}$ دو پایه برای V باشند و $A=(a_{ij})$ به‌وسیلهٔ $w_i = \sum a_{ij}v_j$ تعریف شده باشد، آن‌گاه $\{v_1,\dots,v_n\}$ و $\{w_1,\dots,w_n\}$ در یک گروه هستند اگروتنهااگر $\det A > 0$. هرکدام از این دو گروه یک جهت روی v نامیده می‌شود. جهتی که پایهٔ $\{v_1,\dots,v_n\}$ به آن متعلق است با $[v_1,\dots,v_n]$ نمایش داده می‌شود و دیگری با $-[v_1,\dots,v_n]$. در $\mathbb{R}^n$ جهت معمول با $[e_1,\dots,e_n]$ تعریف می‌شود.

حال فرض می‌کنیم ضرب داخلی T برای V داده شده است. اگر $\{v_1,\dots,v_n\}$ و $\{w_1,\dots,w_n\}$ دو پایه متعامد(نسبت به T) باشند و ماتریس $A=(a_{ij})$ از تساوی $w_i = \sum_{j=1}^{n} a_{ij}v_j$ حاصل شده باشد، آن‌گاه

$$\delta_{ij} = T(w_i,w_j) = \sum_{k,l=1}^{n} a_{ik}a_{jl}T(v_k,v_l) = \sum_{k=1}^{n} a_{ik}a_{jk}.$$

به‌عبارتی اگر A^T نمایشگر ترانهاده ماتریس A باشد، آن‌گاه $AA^T = I$ و در نتیجه $\det A = \pm 1$.

بنا بر قضیه ۶.۱.۶، اگر $w \in \Lambda^n(V)$ و $w(v_n,\dots,v_1) = \mp 1$، آن‌گاه $w(v_n,\dots,v_1) =$

۱∓. در نتیجه اگر جهت μ برای V داده شده باشد، یک $w \in \Lambda^n(V)$ یکتا چنان موجود است که $w(v_n, \ldots, v_1) = 1$ در حالی که $\{v_1, \ldots, v_n\}$ یک پایهٔ متعامد یکه است که $\mu = [v_n, \ldots, v_1]$. این nـ تانسور یکتای w عنصر حجمV نامیده می‌شود که به‌وسیلهٔ ضرب داخلی T و جهت μ تعیین می‌شود. توجه می‌کنیم که $\det$ عنصر حجم $\mathbb{R}^n$ تعیین شده به وسیلهٔ ضرب داخلی معمولی و جهت معمولی روی $\mathbb{R}^n$ است و $|\det(v_n, \ldots, v_1)|$ حجم تولید شده توسط خط‌های گذرنده از ۰ تا هر کدام از $v_n, \ldots, v_1$ ها است.

در ادامه به یک ساختار می‌پردازیم که به $\mathbb{R}^n$ محدود می‌شود. اگر v_1 ، ...و v_{n-1} به $\mathbb{R}^n$ متعلق باشند و ϕ به وسیلهٔ

$$\phi(w) = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ w \end{pmatrix}$$

تعریف شود آن‌گاه $\phi \in \Lambda^1(\mathbb{R}^n)$ و بنابراین $z \in \mathbb{R}^n$ چنان موجود است که

$$\phi(w) = < w, z > \ .$$

z را با $v_1 \times \cdots \times v_{n-1}$ نمایش داده و آن را ضرب صلیبی v_1 ، ...و v_n می‌نامیم. روابط زیر برقرارند:

$$\begin{gathered} v_{\sigma(1)} \times \cdots \times v_{\sigma(n-1)} = \operatorname{Sgn}(\sigma) \cdot v_1 \times \cdots \times v_{n-1}, \\ v_1 \times \cdots \times a v_i \times \cdots \times v_{n-1} = a \cdot (v_1 \times \cdots \times v_{n-1}), \\ v_1 \times \cdots \times (v_i + vi') \times \cdots \times v_{n-1} = \\ v_1 \times \cdots \times v_i \times \cdots \times v_{n-1} + v_1 \times \cdots \times v_i' \times \cdots \times v_{n-1}. \end{gathered}$$

در واقع در ریاضیات معمول نیست که حاصلضربی داشته باشیم که به بیش از دو عنصر وابسته باشد. بنابراین اگر خود را به $\mathbb{R}^3$ محدود کنیم به مفهوم ملموس‌تری می‌رسیم. در

واقع برای $v, w \in \mathbb{R}^3$ داریم $v \times w \in \mathbb{R}^3$. به‌همین دلیل ضرب صلیبی را تنها در $\mathbb{R}^3$ تعریف می‌کنند.

۲.۶ میدان‌ها و صُوَر

برای $p \in \mathbb{R}^n$ مجموعهٔ $\{(p, v) \mid v \in \mathbb{R}^n\}$ با $\mathbb{R}^n_p$ نمایش داده می‌شود و فضای مماس $\mathbb{R}^n$ در p خوانده می‌شود. این مجموعه با اعمال زیر یک فضای برداری است:

$$(p, v) + (p, w) = (p, v + w) \, ,$$
$$a \cdot (p, w) = (p, aw) \, .$$

بردار $v \in \mathbb{R}^n$ اغلب به‌وسیلهٔ یک فلش از ۰ تا v تصویر می‌شود. بردار $(p, v) \in \mathbb{R}^n$ نیز می‌تواند به‌وسیلهٔ فلشی در همان جهت با همان طول اما با نقطهٔ آغازی p تصویر شود. این فلش از p به نقطهٔ $p + v$ می‌رود و لذا $p + v$ نمایشگر انتهای (p, v) است. معمولاً (p, v) را با v_p (بردار v در p) نمایش می‌دهیم.

فضای برداری $\mathbb{R}^n_p$ از خیلی جهات مشابه $\mathbb{R}^n$ است. در واقع ضرب داخلی معمولی $\langle \cdot, \cdot \rangle_p$ برای $\mathbb{R}^n_p$ به‌وسیلهٔ $\langle v_p, w_p \rangle_p = \langle v, w \rangle$ تعریف می‌شود و "جهت معمول" $\mathbb{R}^n_p$ نیز $[(e_1)_p, \ldots, (e_n)_p]$ است.

هر عملی که روی یک فضای برداری محتمل است، روی $\mathbb{R}^n_p$ نیز تعریف می‌شود. یکی از ساده‌ترین اعمال روی یک فضای برداری انتخاب یک بردار از آن است. اگر چنین انتخابی در یک $\mathbb{R}^n_p$ انجام شود یک "میدان برداری" حاصل خواهد شد. برای دقت بیشتر، یک میدان برداری تابعی است چون F به طوری که برای هر $p \in \mathbb{R}^n$ ، $F(p) \in \mathbb{R}^n_p$. برای هر p اعدادی چون $F^1(p), \ldots, F^n(p)$ چنان موجودند که

$$F : \mathbb{R}^n \longrightarrow \cup_{p \in \mathbb{R}^n} \mathbb{R}^n_p \, , \qquad F(p) = F^1(p)(e_1)_p + \cdots + F^n(p)(e_n)_p \, .$$

بنابراین n تابع مؤلفه‌ای $F^i : \mathbb{R}^n \longrightarrow \mathbb{R}$ حاصل خواهد شد. میدان برداری F پیوسته، مشتق‌پذیر و ... نامیده می‌شود هرگاه F^iها چنین باشند. اگر F و G دو میدان برداری باشند و f یک تابع باشد، تعریف می‌کنیم

$$(F+G)(p) = F(p) + G(p) \; ,$$
$$\langle F, G\rangle(p) = \langle F(p), G(p)\rangle \; ,$$
$$(f \cdot F)(p) = f(p) \cdot F(p) \; .$$

اگر $F_۱, \ldots, F_{n-۱}$ میدان‌های برداری روی $\mathbb{R}^n$ باشند، در این صورت به‌طور مشابه می‌توانیم تعریف کنیم

$$(F_۱ \times \cdots \times F_{n-۱})(p) = F_۱(p) \times \cdots \times F_{n-۱}(p) \; .$$

برخی تعاریف دیگر استاندارد و مفید است. دیورژانس به‌صورت زیر تعریف می‌شود:

$$\mathrm{div} F = \sum_{i=۱}^{n} D_i F^i \; .$$

اگر نماد زیر را بپذیریم

$$\nabla = \sum_{i=۱}^{n} D_i \cdot e_i$$

با کمک نمادهای معرفی شده در بالا می‌توان نوشت

$$\mathrm{div} F = \langle \nabla, F\rangle \, .$$

برای $n = ۳$ می‌نویسیم

$$(\nabla \times F)(p) = \big(D_۲ F^۳ - D_۳ F^۲\big)(e_۱)_p + \big(D_۳ F^۱ - D_۱ F^۳\big)(e_۲)_p$$

$$+ \big(D_۱F^۲ - D_۲F^۱\big)(e_۳)_p\,.$$

میدان برداری $\nabla \times F$، کرل F نامیده می‌شود. مفاهیم دیورژانس و کرل از فیزیک وارد شده‌اند. مطالعات مشابهی را می‌توان برای تابعی چون w داشت که $w(p) \in \Lambda^k(\mathbb{R}^n_p)$. چنین تابعی یک kـ صورت دیفرانسیلی (kـ فرم دیفرانسیلی) روی $\mathbb{R}^n$

$$\big(w : \mathbb{R}^n \longrightarrow \cup_{p\in\mathbb{R}^n}\Lambda^k(\mathbb{R}^n_p)\big)$$

و یا برای سادگی یک kـ صورت (kـ فرم) خوانده می‌شود. اگر $\{\phi_۱(p), \ldots, \phi_n(p)\}$ پایهٔ دوگان برای $\{(e_۱)_p, \ldots, (e_n)_p\}$ باشند، آن‌گاه

$$w(p) = \sum_{i_۱<\cdots<i_k} w_{i_۱,\ldots,i_k}(p) \cdot \big[\phi_{i_۱}(p) \wedge \cdots \wedge \phi_{i_k}(p)\big]\,.$$

یک صورت دیفرانسیلی پیوسته، مشتق‌پذیر و ... نامیده می‌شود، هرگاه توابع $w_{i_۱,\ldots,i_k}$ این چنین باشند.

اغلب فرض می‌کنیم صور و میدان‌های برداری مشتق‌پذیر هستند و برای جلوگیری از رخداد برخی مشکلات فنی احتمالی در دفعات مشتق‌پذیری، اغلب مشتق‌پذیری با "C^∞" بودن تعبیر می‌شود.

$w + \eta$ ، $f \cdot w$ و $w \wedge \eta$ به‌طور مشابه تعریف می‌شوند تابع $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ یک ۰ـ صورت (۰ـ فرم) تعریف می‌شود و لذا $f \cdot w$ معمولاً با $f \wedge w$ نمایش داده می‌شود.

اگر $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ مشتق‌پذیر باشد، آن‌گاه $Df(p) \in \Lambda^۱(\mathbb{R}^n)$. با کمی دقت یک ۱ـ فرم به نام $\mathrm{d}f$ می‌سازیم

$$\big(df\big)(p) \in \Lambda^۱(\mathbb{R}^n_p)\,.$$

$$\big(df\big)(p) : \mathbb{R}^n(p) \longrightarrow \mathbb{R}$$

$$(v_p) \longrightarrow (Df)(p)(v)$$

حال به‌طور خاص ۱ـ فرم‌های $\mathrm{d}\pi^i$ را در نظر می‌گیریم. فرض کنیم π_i توابع تصویر

$$\begin{aligned} \pi_i : \quad \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longmapsto x_i \end{aligned}$$

باشند و

$$d\pi^i : \mathbb{R}^n \longrightarrow \cup_{p \in \mathbb{R}^n} \Lambda^1(\mathbb{R}^n_p)$$

$$\begin{aligned} d\pi^i(p) : \quad \mathbb{R}^n_p &\longrightarrow \mathbb{R} \\ (v_p) &\longmapsto D\pi^i(p)(v) \end{aligned}$$

$$\begin{aligned} D\pi^i : \quad \mathbb{R}^n &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \\ (x_1, \dots, x_n) &\longmapsto e_i[x_1, \dots, x_n] \ . \end{aligned}$$

برای راحتی π^i را با x^i نمایش می‌دهیم.

$$\mathrm{d}x^i(p)(v_p) = \mathrm{d}\pi^i(p)(v_p) = D\pi^i(p)(v) = v^i \ .$$

بنابراین می‌بینیم که $\{\mathrm{d}x^1(p), \dots, \mathrm{d}x^n(p)\}$ پایه دوگان $\{(e_1)_p, \dots, (e_n)_p\}$ هستند.

$$\phi_i\big((e_j)_p\big) = \delta_{ij} \ \ , \ \ \big(\mathrm{d}\pi^i\big)(p)\big((e_j)_p\big) = \delta_{ij} \ .$$

بنابراین هر kـ فرم w به‌صورت

$$w = \sum_{i_1 < \cdots < i_k} w_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

نوشته می‌شود.

به‌طور ویژه $\mathrm{d}f$ را مورد بررسی قرار می‌دهیم.

قضیه ۱.۲.۶. اگر $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ مشتق‌پذیر باشد، آن‌گاه

$$\mathrm{d}f = D_1 f \cdot \mathrm{d}x^1 + \cdots + D_n f \cdot \mathrm{d}x^n.$$

با استفاده از نمادهای کلاسیک داریم

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\,\mathrm{d}x^1 + \cdots + \frac{\partial f}{\partial x_n}\,\mathrm{d}x^n\ .$$

برهان.

$$\begin{aligned}\mathrm{d}f(p)(v_p) = (Df)(p)(v) &= \sum_{i=1}^{n} v^i \cdot D_i f(p)\\ &= \sum_{i=1}^{n} \big(\mathrm{d}x^i\big)(p)(v_p) \cdot D_i f(p)\ .\end{aligned}$$

□

اگر تابع مشتق‌پذیر $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ را در نظر بگیریم، نگاشت خطی $Df(p) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ موجود است و نگاشت خطی $f_* : \mathbb{R}^n_p \longrightarrow \mathbb{R}^m_{f(p)}$ با ضابطه $f_*(v_p) = (Df(p)(v))_{f(p)}$ تعریف می‌شود.

این نگاشت تبدیل خطی $f^* : \Lambda^k(\mathbb{R}^m_{f(p)}) \longrightarrow \Lambda^k(\mathbb{R}^n_p)$ القا خواهد کرد لذا می‌توان kـ فرم f^*w روی $\mathbb{R}^n$ را تعریف کرد $(f^*w)(p) = f^*(w(f(p)))$.

یک ساختار مهم فرم‌های دیفرانسیل، تعمیم مفهوم عملگر d است که یک ۰ـ فرم را به یک ۱ـ فرم تبدیل می‌کند.

اگر $w = \sum_{i_1<\cdots<i_k} w_{i_1\ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ یک kـ فرم باشد، آن‌گاه $(k+1)$ـ فرم $\mathrm{d}w$ به‌صورت

$$\mathrm{d}w = \sum_{i_1<\cdots<i_k} \mathrm{d}w_{i_1\ldots i_k} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

$$= \sum_{i_1<\cdots<i_k} \sum_{\alpha=1}^{n} D_\alpha(w_{i_1,\dots,i_k}) \cdot \mathrm{d}x^\alpha \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

تعریف می‌شود.

قضیه ۲.۲.۶. اگر w یک kـ فرم و η یک lـ فرم باشد، آن‌گاه احکام زیر برقرارند:

(۱) $\mathrm{d}(w+\eta) = \mathrm{d}w + \mathrm{d}\eta$.

(۲) $\mathrm{d}(w \wedge \eta) = \mathrm{d}w \wedge \eta + (-1)^k w \wedge \mathrm{d}\eta$.

(۳) $\mathrm{d}(\mathrm{d}w) = 0$ در واقع می‌توان نوشت $\mathrm{d}^2 = 0$.

(۴) اگر w یک kـ فرم دیفرانسیل روی $\mathbb{R}^m$ و $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ مشتق‌پذیر باشد، آن‌گاه $f^*(\mathrm{d}w) = \mathrm{d}(f^*w)$.

فرم w بسته نامیده می‌شود هرگاه $\mathrm{d}w = 0$ و فرم دقیق نامیده می‌شود هرگاه فرم η چنان موجود باشد که $w = \mathrm{d}\eta$.

قضیه بالا نشان می‌دهد که هر فرم دقیق، بسته است و سؤال طبیعی این است که عکس آن درست است یا نه؟

اگر w، ۱ـ فرم $P\,\mathrm{d}x + Q\,\mathrm{d}y$ روی $\mathbb{R}^2$ باشد، آن‌گاه

$$\begin{aligned}\mathrm{d}w &= (D_1 p\,\mathrm{d}x + D_2 p\,\mathrm{d}y) \wedge \mathrm{d}x + (D_1 Q\,\mathrm{d}x + D_2 Q\,\mathrm{d}y) \wedge \mathrm{d}y \\ &= (D_1 Q - D_2 p)\,\mathrm{d}x \wedge \mathrm{d}y\,.\end{aligned}$$

بنابراین اگر $\mathrm{d}w = 0$، آن‌گاه $D_1 Q = D_2 P$. تمرین ۱۶ فصل ۲ نشان می‌دهد که ۰ـ فرم f چنان موجود است که $w = \mathrm{d}f = D_1 f\,\mathrm{d}x + D_2 f\,\mathrm{d}y$.

اگر w روی یک زیر مجموعه دلخواه $\mathbb{R}^2$ تعریف شده باشد، ممکن است چنین f ای موجود نباشد. مثال زیر یک مثال کلاسیک از این گونه فرم‌ها است.

$$w = \frac{-y}{x^2+y^2}\,\mathrm{d}x + \frac{x}{x^2+y^2}\,\mathrm{d}y,$$

که روی $\mathbb{R}^2 - \{0\}$ تعریف شده است. این فرم عموماً با $\mathrm{d}\theta$ نمایش داده می‌شود.
فرض کنیم $w = \sum_{i=1}^{n} w_i \, \mathrm{d}x^i$ یک ۱ـ فرم روی $\mathbb{R}^n$ و همچنین w برابر با $\mathrm{d}f = \sum_{i=1}^{n} D_i f \cdot \mathrm{d}x^i$ باشد. به‌وضوح می‌توانیم فرض کنیم $f(0) = 0$. بنا بر تمرین ۱۷ فصل ۲ داریم

$$f(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f(tx) \, \mathrm{d}t = \int_0^1 \sum_{i=1}^{n} D_i f(tx) \cdot x^i \, \mathrm{d}t = \int_0^1 \sum_{i=1}^{n} w_i(tx) \cdot x^i \, \mathrm{d}t \, .$$

با توضیحات بالا بر حسب این که برای w داده شده f چگونه به‌دست می‌آید، به‌نظر مناسب می‌رسد که نماد زیر را معرفی کنیم

$$Iw(x) = \int_0^1 \sum_{i=1}^{n} w_i(tx) x^i \, \mathrm{d}t \, .$$

باید توجه کنیم Iw بر حسب w، زمانی معنی‌دار است که زیرمجموعهٔ باز $A \subseteq \mathbb{R}^n$ چنان باشد که برای $x \in A$ پاره‌خط واصل 0 به x در A واقع باشد. چنین مجموعهٔ بازی را ستاره‌ـگون نسبت به صفر می‌نامیم. محاسبات بالا نشان می‌دهد که روی یک مجموعهٔ باز ستاره‌ـگون داریم $w = \mathrm{d}(Iw)$ البته با این شرط اضافه که $\mathrm{d}w = 0$.

محاسبات بالا را می‌توان در قضیه ذیل گسترش داد:

قضیه ۳.۲.۶ (لم پوانکاره)**.** اگر $A \subseteq \mathbb{R}^n$ باز و نسبت به صفر ستاره‌ـگون باشد، آن‌گاه هر فرم بسته روی A دقیق است.

برهان. تابع I را از lـفرم‌ها به $(l-1)$ـفرم‌ها (برای هر l) چنان تعریف می‌کنیم که $I(0) = 0$ و برای هر فرم w، $w = I(\mathrm{d}w) + \mathrm{d}(Iw)$. این نشان می‌دهد که برای فرم دقیق w، $w = \mathrm{d}(Iw)$. فرض کنیم

$$w = \sum_{i_1 < \cdots < i_l} w_{i_1,\ldots,i_l} \, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \, .$$

چون A ، ستاره‌ـگون است می‌توانیم Iw را تعریف کنیم:

$$Iw(x)=\sum_{i_1<\cdots<i_l}\sum_{\alpha=1}^{l}(-1)^{\alpha}\left(\int_0^1 t^{l-1}w_{i_1,\ldots,i_l}(tx)\,\mathrm{d}t\right)x^{i_\alpha}\ \ldots$$
$$\mathrm{d}x^{i_1}\wedge\cdots\wedge\widehat{\mathrm{d}x^{i_\alpha}}\wedge\cdots\wedge\mathrm{d}x^{i_\alpha}.$$

که در آن منظور از نماد $\wedge$ روی $\mathrm{d}x^{i_\alpha}$ به این معنی است که $\mathrm{d}x^{i_\alpha}$ حذف شده است.

حال بنا بر مسأله داریم

$$\mathrm{d}(Iw)=l\cdot\sum_{i_1<\cdots<i_l}\left(\int_0^1 t^{l-1}w_{i_1,\ldots,i_l}(tx)\,\mathrm{d}t\right)\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_l}$$
$$+\sum_{i_1<\cdots<i_l}\sum_{\alpha=1}^{l}\sum_{j=1}^{n}(-1)^{\alpha-1}\left(\int_0^1 t^{l}D_j(w_{i_1,\ldots,i_l})(tx)\,\mathrm{d}t\right)x^{i_\alpha}\ \ldots$$
$$\mathrm{d}x^{j}\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\widehat{\mathrm{d}x^{i_\alpha}}\wedge\cdots\wedge\mathrm{d}x^{i_l}.$$

همچنین داریم

$$\mathrm{d}w=\sum_{i_1<\cdots<i_l}\sum_{j=1}^{n}D_j(w_{i_1,\ldots,i_l})\cdot\mathrm{d}x^{j}\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_l}$$

حال I را روی $(l+1)$ـ فرم $\mathrm{d}w$ اعمال می‌کنیم. خواهیم داشت

$$I(x)=\sum_{I_1<\cdots<i_l}\sum_{j=1}^{n}\left(\int_0^1 t^{l}D_j(w_{i_1,\ldots,i_l})(tx)\,\mathrm{d}t\right)x^{j}\,\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_l}$$
$$-\sum_{i_1<\cdots<i_l}\sum_{j=1}^{n}\sum_{\alpha=1}^{l}(-1)^{\alpha-1}\left(\int_0^1 t^{l}D_j(w_{i_1,\ldots,i_l}(tx)\,\mathrm{d}t\right)x^{i_\alpha}\ \ldots$$
$$\mathrm{d}x^{j}\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_\alpha}\wedge\cdots\wedge\mathrm{d}x^{i_l}.$$

با جمع کردن، جمع‌های سه‌گانه حذف می‌شوند و خواهیم داشت

$$
\begin{aligned}
\mathrm{d}(Iw) + I(\mathrm{d}w) & \sum_{i_1<\cdots<i_l} l \cdot \left(\int_0^1 t^{l-1} w_{i_1,\ldots,i_l}(tx)\,\mathrm{d}t \right) \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \\
& + \sum_{i_1<\cdots<i_l} \sum_{j=1}^{n} \left(\int_0^1 t^l x^j D_j(w_{i_1,\ldots,i_l})(tx)\,\mathrm{d}t \right) \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \\
& = \sum_{i_1<\cdots<i_l} \left(\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \left[t^l w_{i_1,\ldots,i_l}(tx) \right] \mathrm{d}t \right) \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \\
& = \sum_{i_1<\cdots<i_l} w_{i_1,\ldots,i_l}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \\
& = w \,.
\end{aligned}
$$

$\square$

۳.۶ مقدمات هندسی

منظور از یک nـ مکعب منفرد در $A \subseteq \mathbb{R}^n$، تابع پیوسته $c : [0,1]^n \longrightarrow A$ است که در آن $[0,1]^n := \overbrace{[0,1] \times \cdots \times [0,1]}^{n \text{ مرتبه}}$ از این پس $[0,1]^0$ یا $\mathbb{R}^0$ را با $\{0\}$ نشان می‌دهیم. در این‌صورت یک ۰ـ مکعب منفرد تابع $f : \{0\} \longrightarrow A$ و یا یک نقطه در A است. یک ۱ـ مکعب منفرد اغلب یک منحنی نامیده می‌شود. یک مثال خاص و البته ساده از nـ مکعب در $\mathbb{R}^n$، nـ مکعب استاندارد $I^n : [0,1]^n \longrightarrow \mathbb{R}^n$ با ضابطه $I^n(x) = x$، $x \in [0,1]^n$ است.

اغلب، جمع‌های صوری nـ مکعب‌های منفرد در A با مضارب صحیح را در نظر می‌گیریم که نمایشی مثل $2c_1 + 3c_2 - 4c_3$ دارند که در آن c_1, c_2, c_3، nـ مکعب‌های منفرد در A هستند. چنین جمع‌های متناهی از nـ مکعب‌های منفرد یک nـ زنجیر نامیده

می‌شود. به‌طور خاص یک nـ مکعب c را می‌توان یک nـ زنجیر $1 \cdot c$ در نظر گرفت.

برای هر nـ زنجیر c در A یک $(n-1)$ـ زنجیر در A را تعریف خواهیم کرد و آن را مرز c می‌نامیم و با ∂c نمایش می‌دهیم. به‌عنوان مثال مرز I^2 باید جمع ۱ـ مکعب‌های منفرد تعریف شود که جهت عکس عقربه‌های ساعت منظم شده‌اند. تعریف دقیق ∂I^n نیازمند برخی مقدمات است. برای هر $1 \leq i \leq n$ دو $(n-1)$ـ مکعب منفرد $I^n_{(i,0)}$ و $I^n_{(i,1)}$ مطابق آن‌چه در ذیل می‌آید تعریف می‌شود. برای $x \in [0,1]^{n-1}$

$$I^n_{(i,0)}(x) = I^n(x^1, \dots, x^{i-1}, 0, x^i, \dots, x^{n-1}) = (x^1, \dots, x^{i-1}, 0, x^i, \dots, x^{n-1})$$
$$I^n_{(i,1)}(x) = I^n(x^1, \dots, x^{i-1}, 1, x^i, \dots, x^{n-1}) = (x^1, \dots, x^{i-1}, 1, x^i, \dots, x^{n-1})$$

$I^n_{(i,0)}$ را وجه $(i,0)$ـ ام و $I^n_{(i,1)}$ را وجه $(i,1)$ـ ام آن می‌نامیم. حال تعریف می‌کنیم

$$\partial I^n = \sum_{i=1}^{n} \sum_{\alpha=(0,1)} (-1)^{i+\alpha} I^n_{(i,\alpha)} \,.$$

برای یک nـ مکعب منفرد دلخواه $c : [0,1]^n \longrightarrow A$، ابتدا وجه (i,α)ـ ام را به‌صورت زیر تعریف می‌کنیم

$$c_{(i,\alpha)} = c_\circ\Big(I^n_{(i,\alpha)}\Big),$$

و سپس تعریف می‌کنیم

$$\partial c = \sum_{i=1}^{n} \sum_{\alpha=0,1} (-1)^{i+\alpha} c_{(i,\alpha)}.$$

در نهایت مرز یک nـ زنجیر $\sum \alpha_i c_i$ را به‌وسیلهٔ

$$\partial\Big(\sum \alpha_i c_i\Big) = \sum \alpha_i \partial c_i,$$

تعریف می‌کنیم.

اولین خاصیت استاندارد مرز را در قضیه ذیل ارائه می‌کنیم:

قضیه ۱.۳.۶. اگر c یک nـ زنجیر در A باشد، آن‌گاه $\partial(\partial c) = ۰$ به‌اختصار $\partial^۲ = ۰$.

برهان. فرض کنیم $i \leq j$، $\left(I^n_{(i,\alpha)}\right)_{(j,\beta)}$ را در نظر می‌گیریم. اگر $x \in [۰,۱]^{n-۲}$ آن‌گاه با یادآوری تعریف وجه (j,β)ـ ام یک nـ مکعب منفرد، داریم

$$\begin{aligned}\left(I^n_{(i,\alpha)}\right)_{(j,\beta)}(x) &= I^n_{(i,\alpha)}\left(I^{n-۱}_{(j,\beta)}(x)\right)\\ &= I^n_{(i,\alpha)}\left(x^۱,\ldots,x^{j-۱},\beta,x^j,\ldots,x^{n-۲}\right)\\ &= I^n\left(x^۱,\ldots,x^{i-۱},\alpha,x^i,\ldots,x^{j-۱},\beta,x^j,\ldots,x^{n-۲}\right).\end{aligned}$$

به‌طور مشابه

$$\begin{aligned}\left(I^n_{(j+۱,\beta)}\right)_{(i,\alpha)} &= I^n_{(j+۱,\beta)}\left(I^{n-۱}_{(i,\alpha)}(x)\right)\\ &= I^n_{j+۱,\beta)}\left((x^۱,\ldots,x^{i-۱},\alpha,x^i,\ldots,x^{n-۲}\right)\\ &= I^n\left(x^۱,\ldots,x^{i-۱},\alpha,x^i,\ldots,x^{j-۱},\beta,x^j,\ldots,x^{n-۲}\right).\end{aligned}$$

بنابراین برای هر $i \leq j$، $\left(I^n_{(i,\alpha)}\right)_{(j,\beta)} = \left(I^n_{(j+۱,\beta)}\right)_{(i,\alpha)}$. حال به‌سادگی برای هر nـ مکعب منفرد c داریم $\left(c_{(i,\alpha)}\right)_{(j,\beta)} = \left(c_{(j+۱,\beta)}\right)_{(i,\alpha)}$، $i \leq j$. از این رو

$$\begin{aligned}\partial(\partial c) &= \partial\left(\sum_{i=۱}^{n}\sum_{\alpha=۰,۱}(-۱)^{i+\alpha}c_{(i,\alpha)}\right)\\ &= \sum_{i=۱}^{n}\sum_{\alpha=۰,۱}\sum_{j=۱}^{n}\sum_{\beta=۰,۱}(-۱)^{i+\alpha+j+\beta}\left(c_{(i,\alpha)}\right)_{(j,\beta)}.\end{aligned}$$

که در این جمع $\left(c_{(i,\alpha)}\right)_{(j,\beta)}$ و $\left(c_{(j+۱,\beta)}\right)_{(i,\alpha)}$ با علائم مخالف هم ظاهر می‌شوند و بنابراین همهٔ جملات حذف می‌شوند و $\partial(\partial c) = ۰$. در نتیجه قضیه برای هر nـ مکعب و لذا برای هر nـ زنجیر درست است. □

۴.۶ قضیه اساسی حسابان

برقراری $\mathrm{d}^2 = 0$ و $\partial^2 = 0$ فقط تشابه رسم‌الخطی نیست. بلکه به‌دلیل ارتباطی است که بین زنجیرها و فرم‌ها موجود است. این ارتباط به‌وسیلهٔ انتگرال‌گیری از فرم‌ها روی زنجیرها بیشتر آشکار می‌شود. در ادامه تنها به nـ مکعب‌های مشتق‌پذیر خواهیم پرداخت.

فرض کنیم w یک kـ فرم روی $[0,1]^k$ باشد در این‌صورت $w = f\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k$ که در آن f تابع یکتایی است.

تعریف می‌کنیم

$$\int_{[0,1]^k} w = \int_{[0,1]^k} f.$$

در این حالت می‌توان نوشت

$$\int_{[0,1]^k} f\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k = \int_{[0,1]^k} f(x^1,\ldots,x^k)\,\mathrm{d}x^1 \ldots \mathrm{d}x^k.$$

اگر w یک kـ فرم روی A و c یک kـ منفرد روی A باشد، تعریف می‌کنیم

$$\int_c w = \int_{[0,1]^k} c^* w\ .$$

دقت می‌کنیم در این حالت

$$\begin{aligned}\int_{I^k} f\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k &= \int_{[0,1]^k} \left(I^k\right)^* \Big(f\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k\Big)\\ &= \int_{[0,1]^k} f(x_1,\ldots,x_k)\,\mathrm{d}x^1 \ldots \mathrm{d}x^k\ .\end{aligned}$$

در حالت خاصی یک 0ـ فرم w یک تابع است، اگر $c : \{0\} \longrightarrow A$ یک 0ـ مکعب منفرد در A باشد. تعریف می‌کنیم

$$\int_c w = w(c(0))\ .$$

انتگرال w روی یک kـ زنجیر $c = \sum \alpha_i c_i$ نیز به‌وسیلهٔ

$$\int_c w = \sum \alpha_i \int_{c_i} w$$

تعریف می‌شود.

انتگرال یک ۱ـ فرم روی یک ۱ـ زنجیر اغلب انتگرال خطی تعریف می‌شود. اگر $P\,\mathrm{d}x + Q\,\mathrm{d}y$ یک ۱ـ فرم روی $\mathbb{R}^2$ و $c : [0,1] \longrightarrow \mathbb{R}^2$ یک ۱ـ مکعب (که در این حالت به وضوح یک منحنی است) باشد آن‌گاه می‌توان ثابت کرد

$$\begin{aligned}\int_c P\,\mathrm{d}x + Q\,\mathrm{d}y = &\lim_{n\to\infty} \sum_{i=1}^{n} \left[c^1(t_i) - c^1(t_{i-1})\right] \cdot P(c^i(t)) \\ &+ \lim_{n\to\infty} \sum_{i=1}^{n} \left[c^2(t_i) - c^2(t_{i-1})\right] \cdot Q(c(t^i))\,.\end{aligned}$$

که در آن $t_0 < \cdots < t_n$ افرازی برای $[0,1]$ هستند. انتخاب t^i در $[t_{i-1}, t_i]$ دلخواه است و حد روی همهٔ افرازهای ممکن گرفته می‌شود که نرم افراز به سمت صفر میل کند.

ارتباط بین فرم‌ها، زنجیرها، d و ∂ به وسیله "قضیه استوکس" مشخص می‌شود که برخی مواقع به قضیه اساسی (بنیادی) حسابان در بعد بالا از آن یاد می‌شود.

قضیه ۱.۴.۶ (استوکس)**.** اگر w یک $(k-1)$ـ فرم روی یک مجموعهٔ باز $A \subseteq \mathbb{R}^n$ و c یک kـ زنجیر در A باشد، آن‌گاه

$$\int_c \mathrm{d}w = \int_{\partial c} w\ .$$

برهان. ابتدا فرض می‌کنیم $c = I^k$ و w یک $(k-1)$ـ فرم روی $[0,1]^k$ باشد. دراین‌صورت w حاصل جمع $(k-1)$ـ فرم‌های از نوع

$$f\,\mathrm{d}x^1 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^k$$

است و کافی است قضیه را برای چنین فرم‌هایی ثابت کنیم
دقت می‌کنیم

$$\int_{[0,1]^{k-1}} I^{k*}_{(j,\alpha)}(f\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k)=\begin{cases}0 & \text{اگر } i\neq j\\ \int_{[0,1]^k} f(x^1\cdots,\alpha,\cdots,x^k)\,\mathrm{d}x^1\ldots\mathrm{d}x^k & \text{اگر } i=j.\end{cases}$$

بنابراین

$$\begin{aligned}&\int f\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k\\&=\sum_{j=1}^{k}\sum_{\alpha=0,1}(-1)^{j+\alpha}\int_{(j,\alpha)^*} I^k_{(i,\alpha)^*}(f\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k)\\&=(-1)^{i+1}\int_{[0,1]^k} f\big(x^1,\cdots,1,\cdots x^k\big)\,\mathrm{d}x^1\ldots\mathrm{d}x^k+(-1)^i\int_{[0,1]^k} f(x^1,\cdots,0,\cdots x^k)\,\mathrm{d}x^1\ldots\mathrm{d}x^k\end{aligned}$$

از سوی دیگر

$$\begin{aligned}\int \mathrm{d}(f\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k)&=\int_{[0,1]^k} D_i f\,\mathrm{d}x^i\wedge\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k\\&=(-1)^{i-1}\int_{[0,1]^k} D_i f\,.\end{aligned}$$

با استفاده از قضیه فوبینی و قضیه بنیادی حسابان (در بعد یک) داریم:

$$\begin{aligned}&\int \mathrm{d}(f\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^i}\wedge\cdots\wedge\mathrm{d}x^k)\\&=(-1)^{i-1}\int_0^1\ldots\left(\int_0^1 D_i f(x^1,\ldots,x^k)\,\mathrm{d}x^i\right)\mathrm{d}x^1\ldots\widehat{\mathrm{d}x^i}\ldots\mathrm{d}x^k\\&=(-1)^{i-1}\int_0^1\ldots\int_0^1\left[f(x^1,\ldots,1,\ldots x^k)-f(x^1,\ldots,0,\ldots,x^k)\right]\mathrm{d}x^1\ldots\widehat{\mathrm{d}x^i}\ldots\mathrm{d}x^k\end{aligned}$$

$$= (-1)^{i-1}\int_{[0,1]^k} f(x^1,\ldots,1,\ldots,x^k)\,\mathrm{d}x^1\ldots\mathrm{d}x^k$$
$$+ (-1)^i \int_{[0,1]^k} f(x^1,\ldots,0,\ldots,x^k)\,\mathrm{d}x^1\ldots\mathrm{d}x^k\,.$$

بنابراین

$$\int_{I^k}\mathrm{d}w = \int_{\partial I^k} w.$$

حال اگر c یک kـ مکعب منفرد دلخواه باشد، بنا بر تعریف خواهیم داشت

$$\int_{\partial c} w = \int_{\partial I^k} c^* w\,.$$

از این رو

$$\int_c \mathrm{d}w = \int_{I^k} c^*(\mathrm{d}w) = \int_{I^k}\mathrm{d}(c^*w) = \int_{\partial I^k} c^*w = \int_{\partial c} w\,.$$

در نهایت اگر c یک kـ زنجیر $\sum \alpha_i c_i$ باشد، خواهیم داشت

$$\int_c \mathrm{d}w = \sum\alpha_i\int_{c_i}\mathrm{d}w = \sum\alpha_i\int_{\partial c_i} w = \int_{\partial c} w.$$

□

۵.۶ تمرین

۱. فرض کنید $\{e_1\ldots e_n\}$ ، پایهٔ استاندارد $\mathbb{R}^n$ باشد و فرض کنید $\{\varphi_1\ldots\varphi_n\}$ پایهٔ دوگان باشد. نشان دهید

$$\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_k}(e_{i_1},\ldots,e_{i_k}) = 1\,.$$

۲. فرض کنید V یک فضای برداری nـ بعدی و $f : V\longrightarrow V$ یک نگاشت خطی باشد. نشان دهید ضابطه $f^* : \Lambda^n(V)\longrightarrow\Lambda^n(V)$ عبارت است از $f^*(w) = \det f\cdot w$.

۳. فرض کنیم $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ و $g : \mathbb{R}^m \longrightarrow \mathbb{R}^p$ دو تابع باشند. نشان دهید

$$(g \circ f)^* = f^* \circ g^* \ .$$

۴. اگر F یک فضای برداری روی $\mathbb{R}^۳$ باشد فرم‌های زیر را تعریف می‌کنیم

$$w_F^۱ = F^۱ \, \mathrm{d}x + F^۲ \, \mathrm{d}y + F^۳ \, \mathrm{d}z \ ,$$
$$w_F^۲ = F^۱ \, \mathrm{d}y \wedge \mathrm{d}z + F^۲ \, \mathrm{d}z \wedge \mathrm{d}x + F^۳ \, \mathrm{d}x \wedge \mathrm{d}y \ .$$

ثابت کنید:

(الف) $f = w_{\mathrm{grad} f}^۱$، $\mathrm{d}(w_F^۱) = w_{\mathrm{curl} F}^۲$، $\mathrm{d}(w_F^۲) = (\mathrm{div} F) \, \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$

(ب) $\mathrm{curl}(\mathrm{grad}) f = ۰$، $\mathrm{div}(\mathrm{curl}) F = ۰$

۵. (استقلال از پارامتر) فرض کنید c یک kـ مکعب منفرد و $p : [۰, ۱]^k \longrightarrow [۰, ۱]^k$ یک تابع یک به یک باشد، به نحوی که $p([۰, ۱])^k) = [۰, ۱]^k$ و برای هر $x \in [۰, ۱]^k$، $\det p'(x) \geq ۰$. اگر w یک kـ فرم باشد. نشان دهید

$$\int_c w = \int_{c o p} w \ .$$

کتاب‌نامه

[1] Aliprantis, C. D. and O. Burkinshaw (1998). *Principles of real analysis*. Gulf Professional Publishing.

[2] Hoffman, M. J. and J. E. Marsden (1993). *Elementary Classical Analysis*. WH Freeman and Co., New York.

[3] Kolmogorov, A. N. and S. V. Fomin (2012). *Introductory real analysis*. Courier Corporation.

[4] Rudin, W. (1976). *Principles of mathematical analysis*. McGraw Hill Kogakusha Limited.

[5] Rudin, W. (1991). Functional analysis. *International series in pure and applied mathematics*.

[6] Spivak, M. (1971). *Calculus On Manifolds: A Modern Approach To Classical Theorems Of Advanced Calculus*. Westview Press.

نمایه

م

ن

و

ه

Vector Analysis

Marzieh Shams Yousefi
Assistant Professor, Faculty of Mathematical Sciences

University of Guilan Press

ISBN: 978-600-153-156-9