初等数学ノート

渕野 昌 (Sakaé Fuchino)

主な更新日: 16.01.04(月 02:04(JST)) Feb.20,2003 Mar.1,2003 Dec.2,2004 Jun.17,2005 Jul.5,2005 Oct.22,2005 Nov.13,2005 Dec.26,2005 May 02,2006 Mar.22,2007 May 28,2007 Jul.14,2008 Aug.3,2008 Dec.1,2008 Mar.22,2011 Jun.18,2011 Feb.17,2014 Jun.01,2014 Jun.15,2015

2017 年 05 月 18 日 (10:19)

“初等数学ノート” というのは題としてあまり適当でないかもしれません．要するに，数学の「研究ノート」のようなものとして書くほどの内容では全然ない，初等的な（しかも多くの場合よく知られている）内容の数学関連のノートを整理したものです．リクリエーション数学，教養の微積や線型代数などに関連したメモ（微積や線型代数を教えているうちに気のついた問題や，このように言えば分りやすいのではないか，というような工夫，いつか後で教科書を書くときに使えそうな別証明や記述方法，etc.）また，大学の数学科の講義で話せる内容などを集めたものです．読者は特に想定していません．全体的には自分のための忘備録という性格が強いし，内容や予備知識のレベルにもばらつきがあります．使用言語も日本語になったり英語になったりドイツ語になったりしてばらばらです．しかし，高校数学くらいの予備知識で読めるものも少なくないはずと思います．分野ごとの章分けになっています（下の目次を参照）．このテキストは常に work in progress 状態です．Download して読んだ方はぜひコメントをください[1)]．

目次

[1)] このテキストの最新版は，

`http://fuchino.ddo.jp/notes/math-notes-elementary.pdf`

としてダウンロードできます．2011 年 03 月 22 日 (火 21:45(JST)) に，本テキストに含まれていたトポロジーの節を，

`http://fuchino.ddo.jp/notes/math-notes-top.pdf`

として分離しました．トポロジーに関することを，そのうち書こうと思っている集合論的トポロジーの教科書というかモノグラフのようなものの準備をかねて順不同に書き出しているうちに，だんだん「初等的」でなくなってきてしまったからですが，このテキストもできるかぎり self-contained に書く努力はしていますので興味のある方はこちらも覗いてみてください．

1 解析学

analysis

1.1 三角関数の (基礎)^2 知識

trigonometry

この節は中部大学工学部で 2007 年春学期に開講している微分積分学 I の補足授業として行なった三角関数の導入の講義を基にしています．私が中部大学で担当しているこの講義を含め，色々な事項の説明を「高校で習っているから」という理由で省略してしまうことが，難しくなりつつある，というのが今日（2007 年現在）の日本の大学での理学基礎教育のかなり一般的な現状のようです．

いずれにしても，手をこまねいて教えることを放棄するわけにはいかないので，「高校で習っている」はずのことや，「中学校で習っている」はずのことでも大学の講義としての品格を失なわないようなやり方で，ていねいに（再）導入する必要があると思われます．

以下は，そのような再導入の試みの 1 つです．といっても特に特別な導入の仕方を目指しているわけではありません．逆に，目標は，できるだけ標準的な導入に近いやり方でしかも，ごまかしのない，self-contained な，中学校の知識だけを前提にしても理解にさしさわりのないような記述を試みるということです．

* * *

角度は日常生活では $45°, 90°, 180°$ など度数で表現されることが多いですが，この度数は，ぐるっと一回りした角度を $360°$ として，これを等分して角度を表すシステムです．ところが，この 360 という数はよく考えてみると何の必然性もなさそうに思えます．数学では，通常この度数法の代わりに弧度法 (radiant [′reɪdɪənt]) が用いられます．これは $360°$ に相当する角度を（単位なしの）2π で表す（一回りの角度との比率が α の角度は $2\pi\alpha$ で表す）ものです．2π は（π の定義から）半径が 1 の円の円周の長さなので，弧度法の角度は，次のように考えることができます:

(1.1) $x \in \mathbb{R}$ の弧度法で表している角度は，xy-平面の原点を中心に半径が 1 の円を描いたとき，x 軸の正の部分を原点を中心にこの円上の円弧の長さが x になるように左回りに回転させたとき，この直線ともとの x 軸の正の部分とのなす角とする．

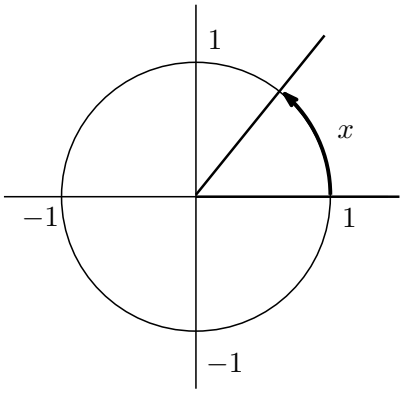


図 1

弧度法での角度を上のように定義すると，x が 0 と 2π の間の数でなくても，またマイナスの数でも，x の表している角度を考えることができることになります．ちなみにマイナスの数が弧度法で表す角度は，右回りの回転をマイナスの円弧の長さと考えることで，この文脈で自然に解釈できます．

弧度法での円周一回りに相当する角度は 2π で，これは 0 の表す角度と同じものです．より一般的には次が成立つことがわかります:

(1.2) $x, y \in \mathbb{R}$ を弧度法で考えるとき，x と y が同じ角（度）を表しているのは，ある整数 m があって $x = y + 2\pi m$ となるときである．

また数の足し算で弧度法での角度の足し算が自然に導入できていますが，これは次のようにして見ることができます: 今 $x, x' \in \mathbb{R}$ が弧度法で同じ角度をあらわし，$y, y' \in \mathbb{R}$ も弧度法で同じ角度をあらわしているとする．このとき (1.2) から，$x = x' + 2\pi m, y = y' + 2\pi n$ となる整数 m, n がとれる．$x + y = x' + y' + 2\pi(m + n)$ となるから，ふたたび (1.2) により，$x + y$ と $x' + y'$ は弧度法で同じ角度をあらわす数になっている．

度数法での $a°$ が弧度法で x のとき，$a : 360 = x : 2\pi$ ですから，$a°$ は弧度法では，$\dfrac{2a\pi}{360}$ になることに注意します．たとえば，$90°$ は，$\dfrac{2 \times 90\pi}{360} = \dfrac{\pi}{2}$ です．

さて，三角関数の代表的なものは $\sin x$, $\cos x$, $\tan x$ の 3 つです．まずこれらを $[0, \frac{\pi}{2})$ 上の関数（つまり $[0, \frac{\pi}{2})$ を定義域とする関数）として定義してみましょう．このために，$x \in [0, \frac{\pi}{2})$ に対し，$\angle \mathrm{ABC} = x$, $\angle \mathrm{BCA} = \frac{\pi}{2}$ となる直角三角形 $\triangle \mathrm{ABC}$ を考えます．ただし $x = 0$ のときには，高さが 0 の A と C が一致した “つぶれた三角形” を考えることにします．三角形の内角の和は π (180°) なので，条件 $x \in [0, \frac{\pi}{2})$ は，このような三角形が考えられることと同値であることに注意してください[2)]．

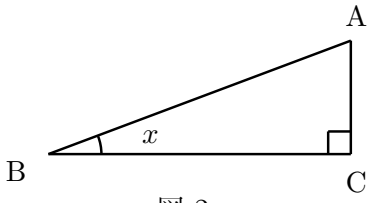


図 2

ここで，$\sin x$, $\cos x$, $\tan x$ を，

$$\sin x = \frac{\mathrm{AC}}{\mathrm{AB}}, \quad \cos x = \frac{\mathrm{BC}}{\mathrm{AB}}, \quad \tan x = \frac{\sin x}{\cos x} = \frac{\mathrm{AC}}{\mathrm{BC}}$$

と定義します．

直角三角形に関しては，三平方の定理としても知られている次のピタゴラスの定理が成立ちます．

定理 1.1（ピタゴラスの定理）AB, BC, CA をそれぞれ斜辺，底辺，高さとする直角三角形で，$\mathrm{AB}^2 = \mathrm{BC}^2 + \mathrm{CA}^2$ が成り立つ． pythagoras

この定理は測量や作図，設計の基礎でもあるので，人類の歴史の中で大規模な測地や建設が行われるようになった時期には少なくとも経験則としては知られていたはずですが，実際に，古バビロニアの遺跡[3)]から出土した刻板で，$a^2 + b^2 = c^2$ の整数解を並べたものが残っていることから，この時期にはすでにピタゴラスの定理が知られていて，その測量や作図，設計での高度な応用がなされていたことがわかります．エジプト 3 大ピラミッドの造営されたのは，紀元前 2550 年と古バビロニアよりさらにずっと前の時期なので，ピタゴラスの定理は，もっと古い時代にもすでに知られていたのではないかと想像されますが，エジプトでは，3 : 4 : 5 の長さの比率の縄で直角を作ることを専門とする職人がいたというこ

[2)] ここで直角三角形に関する用語を復習しておきましょう．直角三角形が図 2 のように描かれているとき，AB を $\triangle$ABC の斜辺と呼ぶのでした．また BC は $\triangle$ABC の底辺 AC は $\triangle$ABC の高さと呼びます．ただし，“高さ” と言ったときには，文脈によって辺 AC のことを指すことも，その長さのことを指すこともあります．これに対して，斜辺や底辺の長さはそれぞれ “斜辺の長さ”, “底辺の長さ” と言って “斜辺”, “底辺” とは言葉として区別することにします．これは単に日本語の語呂の都合で他には特に意味はありません．また，たとえば，AB で，三角形 $\triangle$ABC の辺 AB のことも，この辺の長さのこともあらわすことにします．

[3)] 紀元前 1900 年から紀元前 1200 年くらいの時期に栄えた古代国家で，首都バビロンは現在のイラクの，バグダットの南方約 90km くらいのところにあった古代都市です．

とで，ピタゴラスの定理（とその逆）はこの比率に特化された形で使われていたようです．また古代インドでは 15 : 36 : 39 という比率が使われていた（$15^2 + 36^2 = 1521 = 39^2$），ということです（[2]）．

ピタゴラスの定理の現存している証明はこれよりずっと新しいものです．紀元前 3 世紀ごろに書かれたユークリッドの「原論」に載っている証明は，現存している証明の最初のものの 1 つで,「ピタゴラスの定理」という名称もここで使われています．

なお漢朝 (紀元前 202 年～紀元 220 年) の「周髀算経」には 3 : 4 : 5 の比率に特化した（ピタゴラスの定理の逆の）証明が記載されているということです．

ピタゴラスの定理の証明は 300 以上の異なるものが知られていますが，以下の証明は，ユークリッドによる証明とほぼ同一のものです．

ピタゴラスの定理の証明

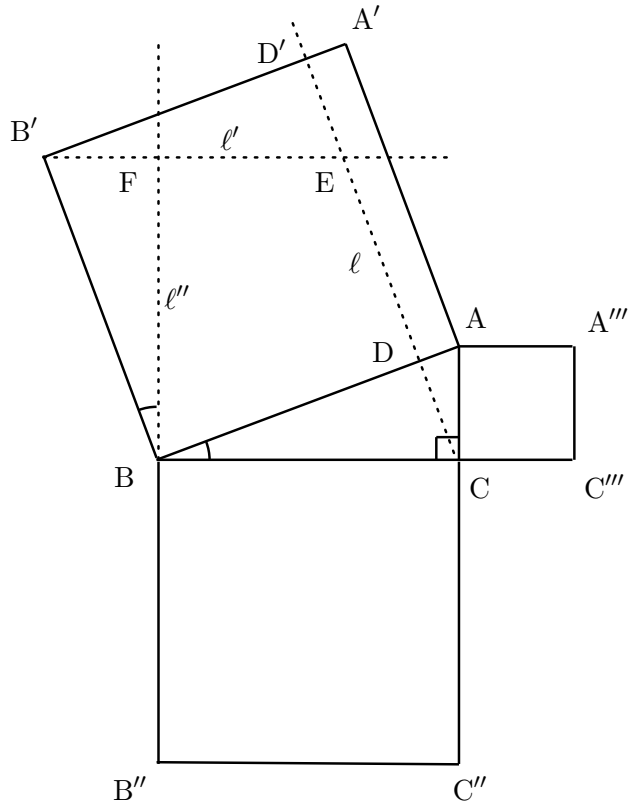


図 3

AB^2, BC^2, CA^2 は，それぞれ直角三角形の斜辺，底辺，高さを一辺とする正方形の面積だから，方程式 $AB^2 = BC^2 + CA^2$ は，この直角三角形の斜辺を一辺する正方形の面積が，底辺を一辺とする正方形の面積と高さを一辺とする正方形の面積を足したものと等しくなることを主張していることに注意する．

そこで，任意の直角三角形 △ABC をとり，AB^2, BC^2, CA^2 のそれぞれを一辺とする正方形をこの三角形の外側に付け加え，それらを □ABB′A′, □BCC″B″, □$CAA'''C'''$ とする（図 3）．

C を通る BB′ と平行な直線を ℓ として，ℓ と AB の交点を D とし，A′B′ との交点を D′ とする．また B' を通る BC と平行な直線を ℓ' として ℓ と ℓ' の交点を E とする．このとき □BB′D′D の面積は □BB′EC の面積と等しくなる．B を通り BC に垂直な直線を ℓ'' として ℓ'' と ℓ' の交点を F とすると，□BB′EC は，BC を底辺として高さが BF の平行四辺形と見ることができるが， △B′BF は △ABC と合同だから， BC = BF である．したがって □BB′EC の面積は BC^2 である．

以上をまとめると，

(1.3) □BB′D′D の面積 $= \mathrm{BC}^2$ trig-0

となることがわかったが，同様に議論すると，

(1.4) □AA′D′D の面積 $= \mathrm{CA}^2$ trig-1

もいえる．ところが

(1.5) AB^2 = □ABB′A′ の面積 = □BB′D′D の面積 + □AA′D′D の面積 trig-2

だから，(1.3), (1.4), (1.5) から，$\mathrm{AB}^2 = \mathrm{BC}^2 + \mathrm{CA}^2$ が示せた． ❑ (定理 1.1)

次のようなピタゴラスの定理の「逆」も成立します．普通，単に「ピタゴラスの定理」と言う場合にも，この逆の命題も含めて言っているいることが多いようです．

pythagoras-0

定理 1.2 a, b, c を $a^2 + b^2 = c^2$ を満たす正の実数とするとき，AB = c, BC = a, CA = b となるような三角形は，$\angle \mathrm{BCA} = \dfrac{1}{2}\pi$ となる直角三角形である．

証明． a, b, c を $a^2 + b^2 = c^2$ を満たす正の実数として，AB = c, BC = a, CA = b となるような三角形について $\angle \mathrm{BCA} = \dfrac{1}{2}\pi$ が成り立たなかったとして矛盾を示す．

例えば $\angle \mathrm{BCA} < \dfrac{1}{2}\pi$ としてみる．$\angle \mathrm{BCA} > \dfrac{1}{2}\pi$ の場合の証明も同様である．このときには，三角形 △ABC の 頂点 A から BC に下した垂線の足を D とすると，D は B と C の間の点となる（図 4）．

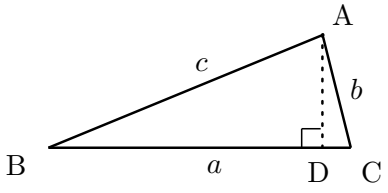


図 4

定理 1.1 を △ACD に適用すると，$\mathrm{DA} = \sqrt{b^2 - \mathrm{CD}^2}$ だから，定理 1.1 を △ABD に適用して，$(a - \mathrm{CD})^2 + (\sqrt{b^2 - \mathrm{CD}^2})^2 = c^2$ となるが，これに $c^2 = a^2 + b^2$ を代入すると，$-2a\mathrm{CD} = 0$ となるから CD = 0 となってしまい．仮定に矛盾する． ❑ (定理 1.2)

[この節はまだ書きかけです]

References

[1] 足立恒雄，フェルマーの大定理，ちくま学芸文庫 (2006).

[2] http://de.wikipedia.org/wiki/Satz_des_Pythagoras

1.2 三角関数の加法定理の導出

additiveth

$R_\tau : \mathbb{R}^2 \to \mathbb{R}^2$ をベクトル $\mathbf{x} \in \mathbb{R}^2$ を原点を中心に左回りに角度 τ だけ回転させて得られるベクトル $R_\tau(\mathbf{x})$ に対応させる写像とする．このとき明らかに，R_τ は線型写像となる．

$R_\tau(\begin{bmatrix}1\\0\end{bmatrix}) = \begin{bmatrix}\cos\tau\\ \sin\tau\end{bmatrix}$, $R_\tau(\begin{bmatrix}0\\1\end{bmatrix}) = \begin{bmatrix}-\sin\tau\\ \cos\tau\end{bmatrix}$ だから，$\begin{bmatrix}\cos\tau & -\sin\tau\\ \sin\tau & \cos\tau\end{bmatrix}$ が R_τ の表現行列になる．

今，$\mathbf{x} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ を考えると，R_τ の定義から，

$$R_\tau(\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}) = \begin{bmatrix}\cos(\theta+\tau)\\ \sin(\theta+\tau)\end{bmatrix} \tag{1.6}$$

additivth-1

となる．一方，上 R_τ の表現行列を用いると，

$$R_\tau(\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}) = \begin{bmatrix}\cos\tau & -\sin\tau\\ \sin\tau & \cos\tau\end{bmatrix}\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} = \begin{bmatrix}\cos\theta\cos\tau - \sin\theta\sin\tau\\ \sin\theta\cos\tau + \cos\theta\sin\tau\end{bmatrix} \tag{1.7}$$

additivth-2

となるから

$$\begin{bmatrix}\cos(\theta+\tau)\\ \sin(\theta+\tau)\end{bmatrix} = \begin{bmatrix}\cos\theta\cos\tau - \sin\theta\sin\tau\\ \sin\theta\cos\tau + \cos\theta\sin\tau\end{bmatrix} \tag{1.8}$$

additivth-3

である．したがって，(1.8) の成分の比較から，

$$\begin{aligned}\cos(\theta+\tau) &= \cos\theta\cos\tau - \sin\theta\sin\tau;\\ \sin(\theta+\tau) &= \sin\theta\cos\tau + \cos\theta\sin\tau\end{aligned} \tag{1.9}$$

がわかる．

この導入法は，学部の微積の講義での三角関数の再導入の教材としては (線型代数と微積の相互関連の一例を見せるという意味でも) 好ましいものと言えるが，数学の基礎の厳密な導入としては適切ではない．解析的な命題の証明を幾何の直観に頼っているからである．幾何の直観を排除した導入としては，三角関数を微分方程式の解として導入する方法 (この場合微分方程式の解の一意性の議論をまずやらなくてはならないという欠点あり)， 三角関数をマクロラン展開により導入する (関数列の収束についての議論をまずしておく必要あり) などが考えられる．

このような導入をして，幾何学的解釈を排除して三角関数の加法定理の証明を行った場合には，上の幾何学的議論は，三角関数の加法定理の幾何学的解釈での追証として捉えられることになる．

上のような思考は大半の日本人が無駄なことと捉えるかもしれない．しかし，私自身は高校生のときに，ここで述べたようなことが解答となっているような疑問のために数学の学習がブロックされて先に進めなくなってしまった苦い思い出がある．

1.3 合成関数の微分法 etc.

chain-rule

2015 年前期の微分積分学 I で使っている教科書 (吹田信之，新保経彦 著，理工系の微分積分学，以下 [吹田，新保] として引用する) では，第 1 章で ε-δ 論法による現代的な極限計算を導入しているのに，第 2 章以降では，コーシーの教科書にでも出ているような古色蒼然とした議論を (直観的な説明としてではなく)“(証)” と称して載せている，という大変気分の悪い書き方になっています．例えば以下の定理 1.4 や補題 1.5 は，この教科書の定理 1 やその系として与えられているものですが，この証明も [吹田，新保] では，“$\Delta x \to 0$ のとき …” という感じのものになっています．講義では，これを ε-δ 論法での厳密な証明に翻訳して示しました．

厳密な証明の欠点として，証明の背後にあるアイデアが見えにくい，ということがあると思いますが，逆に，きっちりと証明を追えば，必ず理解できる，ことが長所もになります．前近代的な証明 (もどき) では，理解に必要となる直観が会得できれば，アイデアが鮮明に見えてくるということは言えるかもしれませんが，この「直観の会得」は誰でもできることではなくなってしまうように思えます．

analysis-a-0

補題 1.3 (1) $\lim_{x\to a} f(x) = b$ で，$\lim_{x\to b} g(x) = c$ なら，$\lim_{x\to a} g(f(x)) = c$ である．

(2) $f(x)$ が a で連続で，$g(x)$ が $f(a)$ で連続なら，$g(f(x))$ は a で連続である．

証明. (1): $\lim_{x\to b} g(x) = c$ により，$\varepsilon > 0$ を任意にとるとき，$\delta_1 > 0$ で，すべての $x \in \mathrm{dom}(g)$ で[4]，$|x-b| < \delta_1$ となるものに対し $|g(x)-c| < \epsilon$ となるようなものがとれる．$\lim_{x\to a} f(x) = b$ により，この δ_1 に対し，$\delta > 0$ で，すべての $x \in \mathrm{dom}(f)$ で $|x-a| < \delta$ となるものに対し，$|f(x)-b| < \delta_1$ となるものがとれる．このとき，すべての $x \in \mathrm{dom}(f)$ で $|x-a| < \delta$ となるものに対し，$|g(f(x))-c| < \varepsilon$ である．

(2): 仮定から，$\lim_{x\to a} f(x) = f(a)$ で，$\lim_{x\to f(a)} g(f(a))$ だから，(1) により，$\lim_{x\to a} g(f(x)) = g(f(a))$ である．したがって，$g(f(x))$ は a で連続である． ❑ (補題 1.3)

次の定理は [吹田，新保] の定理 1 の系として与えられている．

analysis-a

定理 1.4 $f(x)$ が $f(x)$ の定義域に含まれる x_0 で微分可能なら，$f(x)$ は x_0 で連続である．

証明. $f(x)$ が x_0 で微分可能なら，$A = \lim_{h\to 0} \dfrac{f(x_0+h)-f(x_0)}{h}$ が存在する[5]．

したがって，ある $\delta_0 > 0$ で，すべての $|h| < \delta_0$ となる $h \in \mathbb{R}$ に対し，

$$\left|\frac{f(x_0+h)-f(x_0)}{h} - A\right| \le 1 \tag{1.10}$$

difconti-0

[4] $\mathrm{dom}(f)$ で f の定義域を表す．

[5] “A” という記号は，[吹田，新保] からの流用だが，これは，ドイツ語の,,Ableitung“ (微分係数) から来た記号のように思える．他の記号でも，f の定義域を $D(f)$ (‘D’ für Definitionsbereich — このノートでは f の定義域は $\mathrm{dom}(f)$ と表している) とするなど，ドイツ語由来と思える記号の使い方が多いことから，[吹田，新保] は，ドイツ語で書かれた古い教科書が種本になっているのではないか，と推測される．

となるようなものが存在する．

ここで，任意の $\varepsilon > 0$ に対して，

(1.11) $\delta = \min\{\delta_0, \dfrac{\varepsilon}{1+|A|}\}$ difconti-1

とすると，

(1.12) すべての $x \in D(f)$ に対し，$|x - x_0| < \delta$ なら，$|f(x) - f(x_0)| < \varepsilon$ となる． difconti-2

ε は任意だったので，このことから，$\lim_{x \to x_0} f(x) = f(x_0)$ が導かれる．つまり $f(x)$ は x_0 で連続である．

上の (1.11) は，次のようにして確かめられる: $h = x - x_0$ とおく．$x = x_0 + h$ である．

(1.13) $|x - x_0| < \delta$ つまり $|h| < \delta$ とすると， difconti-2-0

δ の定義 (1.11) から，$|h| < \delta_0$ だから，(1.10) から，

(1.14) $|f(x_0 + h) - f(x_0) - hA| \leq |h|$ difconti-3

である，したがって[6]，

(1.15) difconti-4

$$\begin{aligned} |(x_0 + h) - f(x_0)| &\leq |h| + |hA| \\ &= |h|(1 + |A|) \\ &\leq \frac{\varepsilon}{1+|A|} \cdot (1 + |A|) \qquad ; (1.13) \text{ と } (1.11) \text{ により} \\ &\qquad\qquad |h| < \delta \leq \frac{\varepsilon}{1+|A|} \\ &\leq \varepsilon. \end{aligned}$$

❑ (定理 1.4)

次の補題は，[吹田，新保] の p.36 での定理 1 に対応するものである． varepsilon

補題 1.5 $\mathbb{I} \subseteq \mathbb{R}$ を区間として，$f : \mathbb{I} \to \mathbb{R}$ を微分可能な関数とする．
$x \in \mathbb{I}$ と $d \in \mathbb{I} - x = \{a - x : a \in \mathbb{I}\}$ に対し，$\varepsilon_f(x, d)$ を

(1.16) $\varepsilon_f(x, d) = f(x + d) - f(x) - f'(x)d$ a-0

で定義する．つまり，

(1.17) $f(x + d) = f(x) + f'(x)d + \varepsilon_f(x, d)$ a-0-a

である．このとき，$\varepsilon_f(x, d)$ は (x を固定したとき d の関数として) 連続で，すべての $x \in \mathbb{I}$ に対し，$\lim_{d \to 0} \dfrac{\varepsilon_f(x, d)}{d} = 0$ が成り立つ．特に，$\lim_{d \to 0} \varepsilon_f(x, d) = 0$ である．

---

[6] ここでは，$|a - b| \leq c$ なら $|a| \leq c + |b|$ となることを使っている．

証明. $\varepsilon_f(x,d)$ の連続性は，(1.16) から明らかである．$x \in \mathbb{I}$ を固定する．$d \neq 0$ として (1.16) の両辺を d で割ると，

$$\frac{\varepsilon_f(x,d)}{d} = \frac{f(x+d)-f(x)}{d} \ - \ f'(x)$$

となる，したがって，

$$\begin{aligned}\lim_{d\to 0}\frac{\varepsilon_f(x,d)}{d} &= \lim_{d\to 0}\left(\frac{f(x+d)-f(x)}{d} \ - \ f'(x)\right) = \lim_{d\to 0}\left(\frac{f(x+d)-f(x)}{d}\right) \ - \ f'(x)\\ &= f'(x) - f'(x) = 0\end{aligned}$$

となる． ❑ (補題 1.5)

analysis-0

定理 1.6 （合成関数の微分法） ある区間 $\mathbb{I} \subseteq \mathbb{R}$ 上の関数 $h : \mathbb{I} \to \mathbb{R}$ が，微分可能な関数 f と g の合成 $h(x) = g(f(x))$ $(x \in \mathbb{I})$ として与えられているとする．このとき h も微分可能で，すべての $x \in \mathbb{I}$ に対し，

$$h'(x) = g'(f(x)) \cdot f'(x)$$

が成り立つ．

証明. 各 $x \in \mathbb{I}$ に対し，$\lim_{d\to 0}\dfrac{h(x+d)-h(x)}{d}$ が存在して，これが $g'(f(x)) \cdot f'(x)$ と等しくなることを示せばよい．$x \in \mathbb{I}$ を固定する．このとき，(1.17) により，

(1.18)
$$\begin{aligned}&\lim_{d\to 0}\frac{h(x+d)-h(x)}{d}\\ &= \lim_{d\to 0}\frac{g(f(x+d)) \ - \ g(f(x)))}{d}\\ &= \lim_{d\to 0}\frac{g(f(x)+f'(x)d+\varepsilon_f(x,d)) \ - \ g(f(x)))}{d}\\ &= \lim_{d\to 0}\frac{g(f(x)) \ + \ g'(f(x))(f'(x)d+\varepsilon_f(x,d)) \ + \ \varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d)) \ - \ g(f(x)))}{d}\\ &= \lim_{d\to 0}\frac{g'(f(x))(f'(x)d+\varepsilon_f(x,d))+\varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d))}{d}\\ &= g'(f(x))f'(x) \ + \ g'(f(x))\lim_{d\to 0}\frac{\varepsilon_f(x,d)}{d} \ + \ \lim_{d\to 0}\frac{\varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d))}{d}\end{aligned}$$

a-0-a-0

補題 1.5 により，$\lim_{d\to 0}\dfrac{\varepsilon_f(x,d)}{d} = 0$ かつ

(1.19)
$$\begin{aligned}&\lim_{d\to 0}\frac{\varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d))}{d}\\ &= \lim_{d\to 0}\frac{\varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d))}{f'(x)d+\varepsilon_f(x,d)} \cdot \frac{f'(x)d+\varepsilon_f(x,d)}{d}\\ &= \lim_{d\to 0}\frac{\varepsilon_g(f(x),f'(x)d+\varepsilon_f(x,d))}{f'(x)d+\varepsilon_f(x,d)} \cdot \lim_{d\to 0}\frac{f'(x)d+\varepsilon_f(x,d)}{d}\end{aligned}$$

$$= \lim_{d\to 0} \frac{\varepsilon_g(f(x), f'(x)d + \varepsilon_f(x,d))}{f'(x)d + \varepsilon_f(x,d)} \cdot \left(f'(x) + \lim_{d\to 0} \frac{\varepsilon_f(x,d)}{d} \right)$$
$$= 0$$

となるから，(1.18) は $g'(f(x))f'(x)$ と等しくなることがわかる． ❑ (定理 1.6)

上の定理の証明をよく見てみると，定理 1.6 は，実は，同じ証明により，各点ごとの微分可能性に関する次のような主張[7] に改良できることが分る:

analysis-0-0

定理 1.7 （合成関数の微分法 — local version） ある区間 $\mathbb{I} \subseteq \mathbb{R}$ 上の関数 $h : \mathbb{I} \to \mathbb{R}$ が関数 f と g の合成 $h(x) = g(f(x))$ $(x \in \mathbb{I})$ として与えられているとする．ある $x_0 \in \mathbb{I}$ に対し，$f(x)$ が x_0 で微分可能で，$g(x)$ が $f(x_0)$ で微分可能なら，h は x_0 で微分可能で，

(1.20) $\quad h'(x_0) = g'(f(x_0)) \cdot f'(x_0)$

が成り立つ． ❑

逆関数の微分法は，一見したところでは，合成関数の微分法の系として，次のようにして得られるように思える:

analysis-0-1

系 1.8 区間 $\mathbb{I}$ 上の微分可能な関数 f について，$f'(x)$ が $\mathbb{I}$ で常に 0 と異る値をとるものとする．f の逆関数 f^{-1} が存在するときには，f^{-1} も微分可能で，

(1.21) $\quad (f^{-1})'(x) = \dfrac{1}{f'(f^{-1}(x))}$ a-0-a-1

が成り立つ．

証明． x を f^{-1} の定義域の要素とするとき，合成関数の微分法から，

(1.22) $\quad 1 = (x)' = (f \circ f^{-1})'(x) = f'(f^{-1}(x))(f^{-1})'(x)$ i-a

だから，両辺を $f'(f^{-1}(x))$ で割れば求める式が得られる． ❑ (系 1.8)

しかし，この証明では $f^{-1}(x)$ が微分可能であることが前提として議論しているので，この「証明」で言えているのは，もし逆関数が微分可能なら，微分係数は，(1.21) のようなものにならなくてはならない，ということのみであり，その意味で完全な証明とは言えない．

以下で，系 1.8 の主張を更に一般化したものに，厳密な証明を与える．このために，まず，次の準備をする:

inverse-0

補題 1.9 $f : \mathbb{I} \to \mathbb{R}$ を単調関するとする．このとき，すべての $\mathbb{I}$ の内点 x_0 に対し，左極限 $y^+ = \lim_{x\to x_0+0} f(x)$ と右極限 $y^- = \lim_{x\to x_0-0} f(x)$ が存在して，

(1.23) $\quad \min\{y^+, y^-\} \le f(x_0) \le \max\{y^+, y^-\}$ i-0

[7] [吹田，新保] では対応する定理は p.38 にこの形で与えられている．

が成り立つ．したがって，f が x_0 で連続となるのは，$y^+ = y^-$ となる丁度そのときである．

証明． 簡単のために f が単調増加の場合を考える．f が単調現象の場合にも同様に示せる．このときには $f(x_0)$ は $\{f(x) : x < x_0\}$ の上界で，$\{f(x) : x_0 < x\}$ の下界だから，$y^+ = \inf\{f(x) : x_0 < x\}$ と $y^- = \sup\{f(x) : x < x_0\}$ が存在して，$y^- \leq f(x_0) \leq y^+$ となる．このとき，$y^+ = \lim_{x \to x_0 + 0} f(x)$, $y^- = \lim_{x \to x_0 - 0} f(x)$ となることが示せる (演習).

❑ (補題 1.9)

inverse-1

補題 1.10 $\mathbb{I}$ を区間とする．
(1) : $f : \mathbb{I} \to \mathbb{R}$ を 1 対 1 で連続な関数とするとき，f は真に単調[8] で，f の像[9] $f(\mathbb{I})$ は区間になる．

(2) : $f : \mathbb{I} \to \mathbb{R}$ を 1 対 1 で連続な関数とするとき，区間 $\mathbb{J} = f(\mathbb{I})$ 上の f の逆関数 f^{-1} も連続になる．

証明． (1) : : $\mathbb{I} = [a, a] = \{a\}$ のときには主張は自明である．そうでなければ，$x_0, x_1 \in \mathbb{I}$ を $x_0 < x_1$ となるようにとる．このとき，f は 1 対 1 だから，

$$f(x_0) < f(x_1) \tag{1.24}$$ i-1

または，

$$f(x_1) < f(x_0) \tag{1.25}$$ i-2

のどちらかが成り立つ．ここでは (1.24) が成り立つと仮定して議論を進める．(1.25) が成り立つ場合も同様にして示せる．

$x_0', x_1' \in \mathbb{I}$ で $x_0' < x_1'$ となるものを任意にとる．このとき，$f(x_0') < f(x_1')$ となることが示せればよい．

関数 $u : [0,1] \to [\min\{x_0, x_0'\}, \max\{x_0, x_0'\}]$, $v : [0,1] \to [\min\{x_1, x_1'\}, \max\{x_1, x_1'\}]$ を，

$$u(t) = tx_0' + (1-t)x_0, \tag{1.26}$$ i-3

$$v(t) = tx_1' + (1-t)x_1 \tag{1.27}$$ i-4

で定義する．このとき，すべての $t \in [0,1]$ に対して，

$$v(t) - u(t) = t(x_1' - x_0') + (1-t)(x_1 - x_0) > 0 \tag{1.28}$$ i-5

だから，$u(t) \neq v(t)$ である．したがって，f が 1 対 1 であることから，すべての $t \in [0,1]$ に対し，

$$f(u(t)) \neq f(v(t)) \tag{1.29}$$ i-6

である．ここで，$t \in [0,1]$ に対し，

$$g(t) = f(v(t)) - f(u(t)) \tag{1.30}$$ i-7

とすると，すべての $t \in [0,1]$ に対し，$g(t) \neq 0$ となるが，g は定義から連続で，$g(0) = f(x_1) - f(x_0) > 0$ だから，中間値の定理により，すべての $t \in [0,1]$ に対し，$g(t) > 0$ となることがわかる．特に，$g(1) = f(x_1') - f(x_0') > 0$ つまり， $f(x_0') < f(x_1')$ である．

f は単調で連続だから，補題 1.9 により，$f(\mathbb{I})$ はある区間の稠密な部分集合になる[10] したがって，再び中間値定理により， $f(\mathbb{I})$ は区間となることがわかる．

(2)：：$\mathbb{J} = f(\mathbb{I})$ とすると，(1)： から $\mathbb{J}$ は区間となる．区間 $\mathbb{J}$ 上の関数 f^{-1} も単調で 1 対 1 となり，補題 1.9 により，f^{-1} は連続である． ❑ (補題 1.10)

analysis-0-2

定理 1.11 f をある区間 $\mathbb{I} \subseteq \mathbb{R}$ 上で連続関数な 1 対 1 関数とする．f が $x_0 \in \mathbb{I}$ で微分可能で $f'(x_0) \neq 0$ なら，f^{-1} は $y_0 = f(x_0)$ で微分可能で，

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} \quad \left(= \frac{1}{f'(f^{-1}(y_0))} \right) \tag{1.31}$$

である．

証明. 補題 1.10 により，f は真に単調で，f^{-1} も真に単調で連続である．

$$\begin{aligned} \lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} &= \lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(f(x_0))}{f(f^{-1}(y)) - f(f^{-1}(y_0))} \\ &= \lim_{y \to y_0} \frac{1}{\frac{f(f^{-1}(y)) - f(f^{-1}(y_0))}{f^{-1}(y) - f^{-1}(f(x_0))}} \\ &= \frac{1}{\lim_{y \to y_0} \frac{f(f^{-1}(y)) - f(f^{-1}(y_0))}{f^{-1}(y) - f^{-1}(f(x_0))}} \end{aligned} \tag{1.32}$$

となる．f の x_0 での連続性から $y_0 = f(x_0) = \lim_{x \to x_0} f(x_0)$ だから，補題 1.3 により，最後の式の分母は，

$$\begin{aligned} \lim_{y \to y_0} \frac{f(f^{-1}(y)) - f(f^{-1}(y_0))}{f^{-1}(y) - f^{-1}(f(x_0))} &= \lim_{x \to x_0} \frac{f(f^{-1}(f(x))) - f(f^{-1}(y_0))}{f^{-1}(f(x)) - f^{-1}(f(x_0))} \\ &= \lim_{x \to x_0} \frac{f(x) - f(x_0))}{x - x_0} \\ &= f'(x_0) \end{aligned} \tag{1.33}$$

である．したがって，$(f^{-1})'(y_0) = \lim_{y \to y_0} \dfrac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0}$ は存在して，その値は $\dfrac{1}{f'(x_0)}$ である．

❑ (定理 1.11)

analysis-1

---

[8] [吹田，新保] の用語では狭義に単調.

[9] [吹田，新保] の用語では f の値域

[10] つまり，ある $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$, $a < b$ で，$f(\mathbb{I}) \subseteq [a,b]$ となり，任意の $a < a' < b' < b$ に対し $f(\mathbb{I}) \cap (a', b') \neq \emptyset$ となるようなものが存在する．

定理 1.12 (合成関数の微分法 その 2) $\mathbb{I}, \mathbb{J}, \mathbb{K} \subseteq \mathbb{R}$ を区間として，$f : \mathbb{J} \times \mathbb{K} \to \mathbb{R}$ を C_1-級とする．$\varphi : \mathbb{I} \to \mathbb{J}$ と $\psi : \mathbb{I} \to \mathbb{K}$ を微分可能とするとき，合成関数 $f(\varphi(t), \psi(t)) : \mathbb{I} \to \mathbb{R}$; $t \mapsto f(\varphi(t), \psi(t))$ も微分可能で，$t \in \mathbb{I}$ に対し，

$$\frac{d}{dt} f(\varphi(t), \psi(t)) = f_x(\varphi(t), \psi(t))\varphi'(t) + f_y(\varphi(t), \psi(t))\psi'(t) \tag{1.34}$$ a-1

となる[11)]．

証明. $t \in \mathbb{I}$ に対し，

$$\frac{d}{dt} f(\varphi(t), \psi(t)) = \lim_{h \to 0} \frac{f(\varphi(t+h), \psi(t+h)) - f(\varphi(t), \psi(t))}{h} \tag{1.35}$$ a-2

だから，この極限が存在して (1.34) の右辺と等しいことが示せればよい．(1.35) は，

$$\lim_{h \to 0} \frac{f(\varphi(t+h), \psi(t+h)) - f(\varphi(t), \psi(t+h))}{h} + \lim_{h \to 0} \frac{f(\varphi(t), \psi(t+h)) - f(\varphi(t), \psi(t))}{h} \tag{1.36}$$ a-3

と等しいから，

$$\Psi(t) = \lim_{h \to 0} \frac{f(\varphi(t), \psi(t+h)) - f(\varphi(t), \psi(t))}{h}, \tag{1.37}$$ a-4

$$\Phi(t) = \lim_{h \to 0} \frac{f(\varphi(t+h), \psi(t+h)) - f(\varphi(t), \psi(t+h))}{h} \tag{1.38}$$ a-5

として，次の Claim 1.12.1, 1.12.2 が示せれば十分である．

anal-0

Claim 1.12.1 $\Psi(t)$ はすべての $t \in \mathbb{I}$ に対しうまく定義できて，$\Psi(t) = f_y(\varphi(t), \psi(t)))\psi'(t)$ となる．

$\vdash$ 各 $t \in \mathbb{I}$ に対し，$\varphi(t)$ を定数と見ると，定理 1.6 から，

$$\Psi(t) = \lim_{h \to 0} \frac{f(\varphi(t), \psi(t+h)) - f(\varphi(t), \psi(t))}{h} = f_y(\varphi(t), \psi(t))\psi'(t)$$

となることがわかる． $\dashv$ (Claim 1.12.1)

anal-1

Claim 1.12.2 $\Phi(t)$ はすべての $t \in \mathbb{I}$ に対しうまく定義できて，$\Phi(t) = f_x(\varphi(t), \psi(t)))\varphi'(t)$ となる．

$\vdash$ 補題 1.5 と同様に，$a \in \mathbb{J}$, $y \in \mathbb{K}$ と $d \in \mathbb{K} - y$ に対し，

$$\varepsilon_{f(a,\cdot)}(y, d) = f(a, y+d) - f(a, y) - f_y(a, y)d \tag{1.39}$$ a-6

とすれば，定理の仮定から，$\varepsilon_{f(a,\cdot)}(y, d)$ は変数 a, y, d に関し連続な関数で，すべての $a \in \mathbb{J}$, $y \in \mathbb{K}$ に対し，

$$\lim_{d\to 0}\frac{\varepsilon_{f(a,\cdot)}(y,d)}{d}=0 \tag{1.40}$$ a-7

となる．同様に，

$$\varepsilon_{\psi}(t,h)=\psi(t+h)-\psi(t)-\psi'(t)h \tag{1.41}$$ a-8

とすれば，各 $t\in\mathbb{I}$ に対し，

$$\lim_{h\to 0}\frac{\varepsilon_{\psi}(t,h)}{h}=0 \tag{1.42}$$ a-9

である．

(1.41) により，

$$\begin{aligned}\Phi(t)&=\lim_{h\to 0}\frac{f(\varphi(t+h),\psi(t+h))-f(\varphi(t),\psi(t+h))}{h}\\&=\lim_{h\to 0}\frac{f(\varphi(t+h),\psi(t)+\psi'(t)h+\varepsilon_{\psi}(t,h))-f(\varphi(t),\psi(t)+\psi'(t)h+\varepsilon_{\psi}(t,h))}{h}\end{aligned} \tag{1.43}$$

ここで $\Delta(t,h)=\psi'(t)h+\varepsilon_{\psi}(t,h)$ と置くと，(1.39) により，

$$\begin{aligned}&=\lim_{h\to 0}\frac{f(\varphi(t+h),\psi(t))+f_y(\varphi(t+h),\psi(t))\Delta(t,h)+\varepsilon_{f(\varphi(t+h),\cdot)}(\psi(t),\Delta(t,h))}{h}\\&\quad-\lim_{h\to 0}\frac{f(\varphi(t),\psi(t))+f_y(\varphi(t),\psi(t))\Delta(t,h)+\varepsilon_{f(\varphi(t),\cdot)}(\psi(t),\Delta(t,h))}{h}\\&=\lim_{h\to 0}\frac{f(\varphi(t+h),\psi(t))-f(\varphi(t),\psi(t))}{h}\\&\quad+\lim_{h\to 0}\frac{(f_y(\varphi(t+h),\psi(t))-f_y(\varphi(t),\psi(t)))\cdot\Delta(t,h)}{h}\\&\quad+\lim_{h\to 0}\frac{\varepsilon_{f(\varphi(t+h),\cdot)}(\psi(t),\Delta(t,h))-\varepsilon_{f(\varphi(t),\cdot)}(\psi(t),\Delta(t,h))}{h}\end{aligned}$$

最後の式に表われる 3 つの項のうち，最初の項については，Claim 1.12.1 と同様の議論により，$f_x(\varphi(t),\psi(t))\varphi'(t)$ となることがわかる．したがって，2 番目の項と 3 番目の項が 0 になることが示せれば証明が完了する．

2 番目の項については，

$$\begin{aligned}&\lim_{h\to 0}\frac{(f_y(\varphi(t+h),\psi(t))-f_y(\varphi(t),\psi(t)))\cdot\Delta(t,h)}{h}\\&=\lim_{h\to 0}\Big(f_y(\varphi(t+h),\psi(t))-f_y(\varphi(t),\psi(t))\Big)\cdot\lim_{h\to 0}\frac{\psi'(t)h+\varepsilon_{\psi}(t,h)}{h}\end{aligned}$$

したがって，(1.42) により

$$=\Big(\big(\lim_{h\to 0}f_y(\varphi(t+h),\psi(t))\big)-f_y(\varphi(t),\psi(t))\Big)\cdot\psi'(t)$$

ここで仮定により f_y は連続だから，

[11] ここでの “$\frac{d}{dt}f(\varphi(t),\psi(t))$” は微分演算子 $\frac{d}{dt}$ の “t” と変数に代入された値 t が混在した書き方になっている．そういう点についてもっと神経質に書くとすると，以下の定理 1.13 での記法を用いて，$\frac{d}{du}f(\varphi(u),\psi(u))\Big|_{u=t}$ とでもするべきであろう．

$$= 0$$

である．3 番目の項は，

$$\begin{aligned}
&\lim_{h\to 0} \frac{\varepsilon_{f(\varphi(t+h),\cdot)}(\psi(t),\Delta(t,h)) - \varepsilon_{f(\varphi(t),\cdot)}(\psi(t),\Delta(t,h))}{h} \\
&= \lim_{h\to 0} \frac{\varepsilon_{f(\varphi(t+h),\cdot)}(\psi(t),\Delta(t,h)) - \varepsilon_{f(\varphi(t),\cdot)}(\psi(t),\Delta(t,h))}{\Delta(t,h)} \cdot \lim_{h\to 0} \frac{\Delta(t,h)}{h} \\
&= \left(\lim_{h\to 0} \frac{\varepsilon_{f(\varphi(t+h),\cdot)}(\psi(t),\Delta(t,h))}{\Delta(t,h)} - \lim_{h\to 0} \frac{\varepsilon_{f(\varphi(t),\cdot)}(\psi(t),\Delta(t,h))}{\Delta(t,h)} \right) \\
&\quad \cdot \lim_{h\to 0} \frac{\psi'(t)h + \varepsilon_{\psi}(t,h)}{h}
\end{aligned}$$

ここで，(1.42) により，$h \to 0$ なら $\Delta(t,h) \to 0$ となることと，(1.40), また (1.42) により，

$$= (0-0)\cdot\psi'(t) = 0$$

である． ⊣ (Claim 1.12.2)

❑ (定理 1.12)

substitution

定理 1.13 (置換積分の定理) g を連続関数として，f を微分可能な関数とする．このとき[12]，

$$\int g(f(x))f'(x)\,dx = \int g(u)\,du\Big|_{u=f(x)}$$

が成り立つ．

証明． $G(x)$ を $g(x)$ の原始関数とする[13]．このとき，$G'(x) = g(x)$ だから，合成関数の微分法により，

$$g(f(x))f'(x) = G'(f(x))f'(x) = (G(f(x)))'$$

である．特に，$G(f(x))$ は g(f(x))f'(x) の原始関数になっている．したがって，

$$\int g(f(x))f'(x)\,dx = G(f(x)) + C = \int g(u)\,du\Big|_{u=f(x)}$$

である． ❑ (定理 1.13)

1.4 合成関数の微分法の意味

kettenregel

合成関数の微分法は,「関数の合成の線型近似は線型近似の合成である」ということを意味している．しかし，不思議なことに，このことは私の見たかぎり日本語で書かれたどの教科書にも明示的に注意されていない．

---

[12] $\int g(u)\,du\Big|_{u=f(x)}$ は，u をパラメタとして持つ $\int g(u)\,du$ の u を $f(x)$ で置き換えて得られる表現である．

[13] 微分積分学の基本定理により，連続関数に対しては，常にその原始関数が存在する．

簡単のために 1 変数関数 $f:\mathbb{R}\to\mathbb{R}, g:\mathbb{R}\to\mathbb{R}$ の合成を考えてみよう．$h(x)=g(f(x))$ とする．$f(x)$ は点 $a\in\mathbb{R}$ で微分可能で，$g(x)$ は点 $f(a)$ で微分可能とする．このとき合成関数の微分法から，$h(x)$ は a で微分可能となるが，点 $a\in\mathbb{R}$ での $h(x)$ の線型近似（つまり接線をグラフとして持つ一次関数）を h_0 とよぶことにすると，

$$h_0(x)=h'(a)x+(h(a)-a\cdot h'(a))$$

である．一方 $f(x)$ の a での線型近似と $g(x)$ の $f(a)$ での線型近似は，それぞれ，

(1.44) $f_0(x)=f'(a)x+(f(a)-a\cdot f'(a)),$ a-10

(1.45) $g_0(x)=g'(f(a))x+(g(f(a))-f(a)\cdot g'(f(a)))$ a-11

となる．したがって，(1.44) と (1.45) から，

$$\begin{aligned}(1.46)\quad g_0(f_0(x)) &= g'(f(a))(f'(a)x+(f(a)-a\cdot f'(a)))+(g(f(a))-f(a)\cdot g'(f(a)))\\ &= g'(f(a))f'(a)x+(g(f(a))-a\cdot g'(f(a))f'(a))\\ &= g'(f(a))f'(a)x+(h(a)-a\cdot g'(f(a))f'(a))\end{aligned}$$

となる．ところが合成関数の微分法から，$g'(f(a))f'(a)=h'(a)$ だから，これを上に代入すると，

$$g_0(f_0(x))=h'(a)x+(h(a)-a\cdot h'(a))=h_0(x)$$

となって，最初に述べた「関数の合成の線型近似は線型近似の合成である」が確かに成り立っていることがわかる．

1.5 三角関数の微分法を最小の三角関数の基礎知識から導出する

trigono-x

$(\sin x)'$ を

(1.47) 三角関数の基本性質: $\sin^2 x+\cos^2 x=1$ trigo-0

(1.48) 加法定理: $\sin(\alpha+\beta)=\sin\alpha\cdot\cos\beta+\cos\alpha\cdot\sin\beta$ trigo-1

(1.49) $\cos\ x$ の連続性 trigo-2

(1.50) $\displaystyle\lim_{x\to 0}\frac{\sin x}{x}=1$ trigo-3

のみから導出する．三角関数の積の公式から導くよりも力ずくだが，アイデアの飛躍は少ない．

微分の定義から，

(1.51) $\displaystyle(\sin\ x)'=\lim_{h\to 0}\frac{\sin(x+h)-\sin\ x}{h}$

である．これの右辺は，

trigo-4

$$\begin{aligned}(1.52)\quad &\lim_{h\to 0}\frac{\sin(x+h)-\sin\ x}{h}\\ &=\lim_{h\to 0}\frac{\sin x\cos h+\cos x\sin h-\sin x}{h}\\ &=\lim_{h\to 0}\frac{\sin\ x\cdot(\cos h-1)}{h}+\lim_{h\to 0}\cos x\frac{\sin h}{h}\end{aligned}$$

となる．(1.52) の最右辺の第 1 項は，(1.47) により，

$$
\begin{aligned}
(1.53)\quad & \lim_{h\to 0}\frac{\sin x\cdot(\cos h-1)}{h}\\
&=\sin x\cdot\lim_{h\to 0}\frac{(\cos h-1)(\cos h+1)}{h(\cos h+1)}\\
&=\sin x\cdot\lim_{h\to 0}\frac{\cos^2 h-1}{h(\cos h+1)}\\
&=\sin x\cdot\lim_{h\to 0}\frac{-\sin^2 h}{h^2}\cdot\frac{h}{\cos h+1}\\
&=\sin x\cdot\lim_{h\to 0}\frac{-\sin^2 h}{h^2}\cdot\lim_{h\to 0}\frac{h}{\cos h+1}\\
&=\sin x\cdot\lim_{h\to 0}-\left(\frac{\sin h}{h}\right)^2\cdot\lim_{h\to 0}\frac{h}{\cos h+1}=\sin x\cdot(-1)\cdot 0=0
\end{aligned}
$$

となる．最後の行で，二乗をとる関数の連続性，(1.50) と $\cos x$ の連続性が用いられている．一方 (1.52) の最右辺の第 2 項は，やはり (1.50) から $\cos x$ となることがわかるから，全体として，

$$(\sin x)'=\cos x \tag{1.54}$$

が示せた．

1.6 対数関数の計算

log

定理 1.14 $a,\ b,\ c\in\mathbb{R}$ で $0<a,b,\ a,b\neq 1$ とするとき，

$$\log_a b=\frac{\log_c b}{\log_c a} \tag{1.55}$$ log-0

が成り立つ．

証明． (1.55) を示すには，

$$\log_a b\cdot\log_c a=\log_c b \tag{1.56}$$ log-1

が示せれば十分である．(1.56) を示すには，(1.56) の両辺をそれぞれ c の肩に載せたものが等しいこと:

$$c^{\log_a b\cdot\log_c a}=c^{\log_c b} \tag{1.57}$$ log-2

を示せば十分である．ところがこれは，

$$c^{\log_a b\cdot\log_c a}=\left(c^{\log_c a}\right)^{\log_a b}=a^{\log_a b}=b$$

また，$c^{\log_c b}=b$ により成り立つ．したがって (1.55) が示せた． ❑

1.7 Anwendungen des Mittelwertsatzes

mittelwertsatz

$f : D \to \mathbb{R}$ heißt monoton steigend auf dem Intervall $I \subseteq D$, falls gilt: $f(x_0) \leq f(x_1)$ für alle x_0, $x_1 \in I$ mit $x_0 < x_1$. f heißt streng monoton fallend anf I, falls $f(x_0) < f(x_1)$ für alle x_0, $x_1 \in I$ mit $x_0 < x_1$.

Analog heißt $f : D \to \mathbb{R}$ monoton fallend auf dem Intervall $I \subseteq D$, falls gilt: $f(x_0) \geq f(x_1)$ für alle x_0, $x_1 \in I$ mit $x_0 < x_1$. f heißt streng monoton fallend anf I, falls $f(x_0) > f(x_1)$ für alle x_0, $x_1 \in I$ mit $x_0 < x_1$.

mwz-0

Lemma 1.15 Sei f differenzierbar auf $[a, b]$.

(a) Falls es gilt $f'(x) \geq 0$ für alle $x \in [a, b]$, so ist f monoton steigend auf $[a, b]$.

(b) Falls es gilt $f'(x) \geq 0$ für alle $x \in [a, b]$, so ist f monoton fallend auf $[a, b]$.

Beweis. (a): Sonst gibt es x_0, $x_1 \in [a, b]$ mit $x_0 < x_1$ d.d. $f(x_0) > f(x_1)$. Nach dem Mittelwertsatz gibt es dann ein $c\mathbb{R}$ mit $x_0 < c < x_1$ d.d.

$$f'(c) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} < 0.$$

Dies ist ein Widerspruch, da $c \in [a, b]$. (b): Analog. ❑

$f : D \to \mathbb{R}$ heißt konkav (oder nach unten konvex) auf dem Intervall $I \subseteq D$, falls gilt: $f(x_1) \leq f(x_0) + \frac{f(x_2) - f(x_0)}{x_2 - x_0}(x_1 - s_0)$ für alle x_0, x_1, $x_2 \in I$ mit $x_0 < x_1 < x_2$.

Analog heißt $f : D \to \mathbb{R}$ (nach oben) konvex auf dem Intervall $I \subseteq D$, falls gilt: $f(x_1) \geq f(x_0) + \frac{f(x_2) - f(x_0)}{x_2 - x_0}(x_1 - s_0)$ für alle x_0, x_1, $x_2 \in I$ mit $x_0 < x_1 < x_2$.

hl-0

Lemma 1.16 (a) f ist konkav auf I gdw. gilt: $\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$ für alle x_0, x_1, $x_2 \in I$ mit $x_0 < x_1 < x_2$.

(b) f ist konvex auf I gdw. gilt: $\frac{f(x_1) - f(x_0)}{x_1 - x_0} \geq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$ für alle x_0, x_1, $x_2 \in I$ mit $x_0 < x_1 < x_2$.

Lemma 1.17 Sei f zweimal diferenzierbar auf dem Intervall I.

(a) Falls $f''(x) \geq 0$ für alle $x \in I$, so ist f konkav auf I.

(b) Falls $f''(x) \leq 0$ für alle $x \in I$, so ist f konvex auf I.

Beweis. (a): Sonst gibt es nach Lemma 1.16, (a) x_0, x_1, $x_2 \in I$ mit $\frac{f(x_1) - f(x_0)}{x_1 - x_0} > \frac{f(x_2) - f(x_1)}{x_2 - x_1}$. Nach dem Mittelwertsatz gibt es also c_0, $c_1 \in I$ mit $x_0 < c_0 < x_1 < c < x_2$ und

$$f'(c_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} > \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c_1).$$

Auf der anderen Seite, ist f' monoton steigend auf I nach der Annahme von $f''(x) = (f')'(x) \geq 0$ für alle $x \in I$ und nach Lemma 1.15. Da $c_0 < c_1$, folgt also $f'(c_0) \leq f'(c_1)$. Ein Widerspruch.

(b): Analog. ❑

1.8 コンパクト性の微積での扱い

compactness

ラグランジュの未定常数法（あるいは未定乗数法）は,「$\mathbb{R}^n$ 有界閉集合上の連続関数は最大値と最小値をとる」という有界閉集合のコンパクト性からの帰結と組にして応用されることが多い．以下に，微積の講義での予備知識のみから，この定理を証明してみる．

compact-0

補題 1.18 すべての $\mathbb{R}$ での有界点列[14] $\langle a_\ell : \ell \in \mathbb{N}\rangle$ は（有限の値に）収束する部分点列を含む．

証明. 区間 $I_n = [c_n, d_n]$ と $k_n \in \mathbb{N}$ $(n = 0, 1, 2, ...)$ を帰納的に次を満たすようにとる:

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots ; \tag{1.58}$$ comp-0

$$d_{n+1} - c_{n+1} = \frac{1}{2}(d_n - c_n); \tag{1.59}$$ comp-1

$$\{\ell \in \mathbb{N} : a_\ell \in [c_n, d_n]\} \text{ は無限集合}; \tag{1.60}$$ comp-2

$$a_{k_n} \in I_n. \tag{1.61}$$ comp-3

このような $I_n = [c_n, d_n]$ と $k_n \in \mathbb{N}$ $(n = 0, 1, 2, ...)$ がとれることは，次のようにしてわかる:

まず，$\langle a_\ell : \ell \in \mathbb{N}\rangle$ は有界だから，(1.60) を満たすような，I_0 がとれる．k_0 を $a_{k_0} \in I_0$ となるようにとる．$I_n = [c_n, d_n]$ と $k_0 < k_1 < \cdots < k_n$ がとれたとき，$b_n = c_n + \frac{1}{2}(d_n - c_n)$ とする．このとき，(1.60) により，$\{\ell \in \mathbb{N} : a_\ell \in [c_n, b_n]\}$ または，$\{\ell \in \mathbb{N} : a_\ell \in [b_n, d_n]\}$ の少なくとも片方は無限集合になる．もし前者が無限集合となるときには，$c_{n+1} = c_n$, $d_{n+1} = b_n$ とし，そうでないときには，$c_{n+1} = b_n$, $d_{n+1} = d_n$ とする．このとき，$\{\ell \in \mathbb{N} : a_\ell \in [c_{n+1}, d_{n+1}]\}$ は無限集合だから，この集合の要素 k で $k_n < k$ となるものがとれる．このような k を k_{n+1} としてとればよい．

(1.59) と (1.61) により，$\langle a_\ell : \ell \in \mathbb{N}\rangle$ の部分列 $\langle a_{k_n} : n \in \mathbb{N}\rangle$ は収束する[15]．

❑ (補題 1.18)

compact-1

補題 1.19 任意の自然数 $n \geq 1$ に対し，すべての $\mathbb{R}^n$ での有界点列[16] $\langle \bar{a}_\ell : \ell \in \mathbb{N}\rangle$ は収束する部分点列を含む．

証明. $n = 1$ のときには，補題の主張は，補題 1.18 と一致するからよい．主張が $n = m \geq 1$ に対し成り立つとして，$n = m+1$ のときにも成り立つことを示す．$\langle \bar{a}_\ell : \ell \in \mathbb{N}\rangle$ を $\mathbb{R}^{m+1}$ の有界点列とする．$\ell \in \mathbb{N}$ に対し，$\bar{a}_\ell = (a^0_\ell, ..., a^m_\ell)$ として，

$$\bar{a}'_\ell = (a^0_\ell, ..., a^{m-1}_\ell), \quad \ell \in \mathbb{N}$$

---

[14] $\mathbb{R}$ で点列 $\langle a_\ell : \ell \in \mathbb{N}\rangle$ が有界とは，ある区間 $[a, b]$ $(a, b \in \mathbb{R})$ にすべての a_ℓ, $\ell \in \mathbb{N}$ が含まれることである．

[15] 厳密に言うと，ここでは,「すべてのコーシー列は収束する」という $\mathbb{R}$ の性質を仮定している．

[16] $\mathbb{R}^n$ の点列 $\langle \bar{a}_\ell : \ell \in \mathbb{N}\rangle$ が有界とは，ある区間 $[a, b]$ $(a, b \in \mathbb{R})$ の積集合 $[a, b]^n = \{(x_0, ..., x_{n-1}) : x_0, ..., x_n \in [a, b]\}$ にすべての $\bar{a}_\ell$, $\ell \in \mathbb{N}$ が含まれることを言う．

とする．このとき，帰納法の仮定により $\langle \bar{a}'_\ell : \ell \in \mathbb{N}\rangle$ の部分列，$\langle \bar{a}'_{k_\ell} : \ell \in \mathbb{N}\rangle$ で収束するものがとれる．一方補題 1.18 により，$\mathbb{R}$ の有界点列 $\langle a^m_{k_\ell} : \ell \in \mathbb{N}\rangle$ の部分列 $\langle a^m_{k'_\ell} : \ell \in \mathbb{N}\rangle$ で収束するものが存在する．$\langle \bar{a}_\ell : \ell \in \mathbb{N}\rangle$ の部分列

$$\langle \bar{a}_{k'_\ell} : \ell \in \mathbb{N}\rangle$$

は収束する． ❑ (補題 1.19)

定理 1.20 任意の自然数 $n \geq 1$ に対し，$\mathbb{R}^n$ の有界閉集合[17] D 上の連続関数は常に最大値と最小値をとる[18]．

証明. 最大値の存在を示す．最小値の存在についても同様に示せる．まず，定理より弱い次の命題を示す:

Claim 1.20.1 有界閉集合 D 上の連続関数 $f(x)$ の値の全体 $f(D)$ は有界である．

⊢ そうでなかったとして，たとえば D 上の連続関数 $f(x)$ の値が上に有界でないとする．このとき $a_0, a_1, a_2, ... \in D$ で，

$$f(a_n) \geq n \quad (1.62)$$ comp-4

がすべての $n \in \mathbb{N}$ に対し成り立つようなものがとれる．D が有界だから，点列 $\langle a_n : n \in \mathbb{N}\rangle$ も有界である．したがって，補題 1.19 により，$k_0 < k_1 < k_2 < \cdots, (k_n \in \mathbb{N})$ を $a_{k_n}, n \in \mathbb{N}$ が収束するようなものとする． $a = \lim_{n\to\infty} a_{k_n}$ とすると，D は閉集合だから，$a \in D$ である．よって $f(x)$ の連続性から，$f(a) = \lim_{n\to\infty} f(a_{k_n})$ だが，(1.62) から $\lim_{n\to\infty} f(a_{k_n}) = \infty$ とならなくてはならず矛盾である． ⊣ (Claim 1.20.1)

Claim 1.20.1 により，$b^* = \sup\{f(a) : a \in D\}$ とすると，$b^* < \infty$ である．a_0, a_1, $a_2, ... \in D$ を，

$$f(a_n) \geq b^* - \frac{1}{n}, \ (n \in \mathbb{N}\ ) \quad (1.63)$$ comp-5

となるようにとる．Claim 1.20.1 の証明でと同じように，$k_0 < k_1 < k_1 < \cdots, (k_n \in \mathbb{N})$ を，$a_{k_n}, n \in \mathbb{N}$ が収束するようにとり，$a^* = \lim_{n\to\infty} a_{k_n}$ とする．D が閉集合であることから，$a^* \in D$ である．f の連続性から，$f(a^*) = \lim_{n\to\infty} f(a_{k_n})$ であるが，(1.63) から $f(a^*) = b^*$ となることがわかる．つまり $f(x)$ は $a^* \in D$ で最大値をとる． ❑ (定理 1.20)

1.9 微分演算子の特徴付け

diff-op

定理 1.21 $D : C^\infty(\mathbb{R}) \to \mathbb{R}$ が線型で，ある実数 $a \in \mathbb{R}$ に対し，

[17] $D \subseteq \mathbb{R}^n$ が有界とは，ある区間 $[a, b]$ $(a, b \in \mathbb{R})$ に対し，$D \subseteq [a, b]^n$ となることである．D が閉集合とは，D の要素からなる点列 $a_n, n \in \mathbb{N}$ が収束するとき，常に $\lim_{n\in\mathbb{N}} a_n \in D$ となることである．$\mathbb{R}^n$ での閉曲線（端点をすべて含む曲線）は閉集合の例の一つである．

[18] この定理で D の有界性は必要である．たとえば，$f(x) = x^2$ を $\mathbb{R}$ 全体で考えると， $\mathbb{R}$ は閉集合だが，$f(x)$ は最大値をとらない．

(1) すべての $f, g \in C^\infty(\mathbb{R})$ に対し，$D(fg) = f(a)D(g) + D(f)g(a)$;
(2) $D(x) = 1$

を満たすなら，すべての $f \in C^\infty(\mathbb{R})$ に対し，$D(f) = \dfrac{df}{dx}(a)$ となる．

証明． $a = 0$ に対して主張が証明できれば十分である．そこで D は，線型で，$a = 0$ に対する (1), と (2) を満たす，とする．

(1) から，

$$D(1) = D(1) \cdot 1 + 1 \cdot D(1)$$

となる[19]．したがって，$D(1) = 1$ となる．よって，線型性から，すべての $b \in \mathbb{R}$ に対し

$$D(b) = D(b \cdot 1) = bD(1) = 0$$

である．

$f \in C^\infty(\mathbb{R})$ とする．

$$\begin{aligned} f(x) - f(0) &= \int_0^1 \frac{df(tx)}{dt} dt \\ &= \int_0^1 x \left[\frac{df(x)}{dx}\right]_{x=tx} dt \\ &= x \int_0^1 \left[\frac{df(x)}{dx}\right]_{x=tx} dt \end{aligned}$$

したがって，

$$f(x) = f(0) + x \int_0^1 \left[\frac{df(x)}{dx}\right]_{x=tx} dt$$

となる[20]．両辺に D を施すと，D の線型性と，(1) から，

$$D(f(x)) = D(f(0)) + D(x) \cdot \left[\int_0^1 \left[\frac{f(x)}{dx}\right]_{x=tx} dt\right]_{x=0} + [x]_{x=0} \cdot D(\cdots)$$

となるが，上で見たように $D(f(0)) = 0$ だから，

$$D(f(x)) = D(x) \cdot \int_0^1 [\frac{df}{dx}]_{x=0} dt = 1 \cdot [\frac{df}{dx}]_{x=0} = \frac{df}{dx}(0)$$

となる． ❑

---

[19] ここで 1 は恒等的に 1 を返す関数のことである．

[20] ここでは，式 t に対し，その式にあらわれる x に式 u を代入して得られる式を $[t]_{x=u}$ であらわしている．

2 線型代数

linalg

以下のテキストは，神戸大学の理学系や工学系の線型代数のテキストとして使っている長谷川 [1] の補足として書かれている．記号の使い方などは，概ね [1] に準ずるものとなっている．ただし，ここでは，自然数の全体 $\mathbb{N}$ は 0 を含むものとすること[21] とし，そのことに対応して自然数で添字をつけるときには 0 から始めることにする．特に，ベクトルの一番 “上” の成分は 0-成分と呼ぶことにし，行列の一番 “上” の行は 0-行，一番 “左” の列は 0-列と呼ぶことにする[22]．

集合（集合論ではない）に関する語彙は積極的に使うことにする．特に，自然数 n はそれより小さな自然数の集合として導入されているもの，とする．つまり $n=\{0,1,...,n-1\}$ である．この記法を用いることのメリットの一つは，$\sum_{i=0}^{n-1}$ と書くかわりに，$\sum_{i\in n}$ と書け，$\sum_{i=k}^{n-1}$ と書くかわりに $\sum_{i\in n\setminus k}$ と書けることである[23]．

2.1 n-次元ベクトル空間とその部分空間

vector-sp

2.2 基底と次元

dimension

以下では，スカラー体 $\mathbb{R}$ 上の n-次元ベクトル空間 $\mathbb{R}^n$ とその部分空間が考察の対象となっているが，ここでの議論は，ほとんどすべて，任意の体[24] K 上の n-次元ベクトル空間 K^n について成り立つ．応用上は，とりあえずスカラー体 K として $\mathbb{C}$ の部分体（0 と 1 を含み加減乗除で閉じているような $\mathbb{C}$ の部分集合）を考えていれば十分であろう．

$\mathbf{a}_0,\dots,\mathbf{a}_{k-1}\in\mathbb{R}^n$ が **線型独立** とは，

(2.1) 任意の $c_0,...,c_{k-1}\in\mathbb{R}$ に対し，$\sum_{i\in k}c_i\mathbf{a}_i=\mathbf{0}$ なら，$c_0=c_1=\cdots=c_{k-1}=0$ となること，

dim-a

とする．ただし，$\mathbf{0}$ は $\mathbb{R}^n$ のゼロベクトルである．

この定義は，

(2.2) $\sum_{i\in k}c_i\mathbf{a}_i=\mathbf{0}$ となるような $c_0,...,c_{k-1}\in\mathbb{R}$ は $c_0=c_1=\cdots=c_{k-1}=0$ に限る，

dim-a-0

と表現したほうが理解しやすい日本語になるかもしれないが，線型独立性に関する数学的議論の中では，最初の書き方の方が自然に使える（たとえば以下の証明を参照）．この定義から，$\mathbf{a}_0,\dots,\mathbf{a}_{k-1}\in\mathbb{R}^n$ が線型独立なら，$\mathbf{a}_0,\dots,\mathbf{a}_{k-1}$ はすべてゼロベクトル $\mathbf{0}$ と異なり，互いにも異ることがわかる（演習）．

(2.1) の否定は，

(2.3) ある $c_0,...,c_{k-1}\in\mathbb{R}$ に対し，$\sum_{i\in k}c_i\mathbf{a}_i=\mathbf{0}$ だが，$c_0=c_1=\cdots=c_{k-1}=0$ では

dim-a-0-0

[21] つまり $\mathbb{N}=\{0,1,2,...\}$ である．

[22] このような添字の使い方は，多くのコンピュータ言語での行列（配列）の添字の扱いとも一致する．

[23] 集合 X,Y に対し，X と Y の集合差 $X\setminus Y$ は，$X\setminus Y=\{x\in X:x\notin Y\}$ と定義される．

[24] 四則演算が定義されて，それらの演算が $\mathbb{R}$ での四則演算と同様の計算則を満たすものを体 (field) という．

ない

である．あるいは，このことは，

(2.4) すべてが 0 ではないような，ある $c_0,...,c_{k-1}\in\mathbb{R}$ に対し，$\sum_{i\in k} c_i\mathbf{a}_i=\mathbf{0}$ となる dim-a-0-1

とも表現できるが，これが，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が線型独立でない，ということの意味である．

$\mathbf{a}_0,...,\mathbf{a}_{k-1}\in\mathbb{R}^n$ に対し，$c_0,...,c_{k-1}\in\mathbb{R}$ として，$\sum_{i\in k} c_i\mathbf{a}_i$ の形の表現（および，このような表現のあらわす $\mathbb{R}^n$ の元）のことを，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の **線型結合** (*linear combination*) とよぶのだった． $\mathbb{R}^n$ の元としての $\sum_{i\in k} c_i\mathbf{a}_i$ を区別する必要があるときには，これを線型結合 $\sum_{i\in k} c_i\mathbf{a}_i$ の値とよぶことにする．

$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の線型結合 $\sum_{i\in k} c_i\mathbf{a}_i$ で， $c_0=\cdots=c_{k-1}=0$ となるもの（このような線型結合の値は当然ゼロベクトルとなる）のことを，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の自明な線型結合とよぶことにすると，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が線型独立であることは，

(2.5) $\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の線型結合で，その値が $\mathbf{0}$ となるものは，自明な線型結合に限る dim-a-1

こと，と表現することもできる．また，この表現を使うと，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が線型独立でないのは，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の自明でない線型結合で，その値が $\mathbf{0}$ となるものが存在することである，と言える．

lin-L-a

補題 2.1 $\mathbf{a}_0,...,\mathbf{a}_{k-1}\in\mathbb{R}^n$ として，ある $\ell\le k$ と $0\le i_0<i_1\cdots<i_{\ell-1}<k$ に対して，$\mathbf{a}_{i_0},...,\mathbf{a}_{i_{\ell-1}}$ が線型独立でないなら， $\mathbf{a}_a,...,\mathbf{a}_{k-1}$ も線型独立でない．

証明． $\mathbf{a}_{i_0},...,\mathbf{a}_{i_{\ell-1}}$ が線型独立でないなら，すべてが 0 でないような $c_{i_0},...,c_{i_{\ell-1}}\in\mathbb{R}$ で，$\sum_{j<\ell} c_{i_j}\mathbf{a}_{i_j}=\mathbf{0}$ となるものが存在する．$i<k$ で $i=i_j$ となるような $j<\ell$ の存在しないものに対して $c_i=0$ とすると，$\sum_{i<k} c_i\mathbf{a}_i=\mathbf{0}$ となるが，$c_i, i<k$ はすべては 0 ではないから，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が線形独立でないことが示せた． ❑ (補題 2.1)

> $\mathbf{a}_0,...,\mathbf{a}_{k-1}\in\mathbb{R}^n$ が線型独立なら，任意の $\ell\le k$ と $0\le i_0<i_1\cdots<i_{\ell-1}<k$ に対して， $\mathbf{a}_{i_0},...,\mathbf{a}_{i_{\ell-1}}$ も線型独立である

という主張は上の補題と論理的に同値であることに注意する．

より一般的には，$X\subseteq\mathbb{R}^n$ が **線型独立** とは，

(2.6) 任意の $k\in\mathbb{N}$ と，任意の互いに異る $\mathbf{a}_0,\ldots,\mathbf{a}_{k-1}\in X$ に対し，$\mathbf{a}_0,\ldots,\mathbf{a}_{k-1}$ が線型独立となること X-lin-indep

とする．この定義から，空集合 $\emptyset$ は線型独立であることに注意する．一方 X が線型独立なら $\mathbf{0}\notin X$ である（演習）．この線型独立性の定義は，(2.1) による線型独立性の定義の拡張になっている:

lin-L-0

補題 2.2 $X \subseteq \mathbb{R}^n$ が空でない有限集合のとき，たとえば，$X = \{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}$ として（ただし，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ は互いに異るとする），$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ が線型独立であることと，X が線型独立であることは同値である[25)]．

証明. X が (2.6) の意味で線型独立なら，この線型独立の定義から $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ は (2.5) の意味で線型独立である．

X が (2.6) の意味で線型独立でないなら，$\ell \leq k$ と $0 \leq i_0 < i_1 < \cdots < i_{\ell-1} < k$ で，$\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_{\ell-1}}$ が ((2.5) の意味で) 線型独立でないようなものが存在する．このとき，補題 2.1 により，$\mathbf{a}_0, \mathbf{a}_1, ..., \mathbf{a}_{k-1}$ も線形独立でない． ❑ (補題 2.2)

注意. 教科書 [1] では,「線型独立」ではなく **一次独立** という用語が用いられていた（講義では両方の用語を導入して主に「線型独立」という言い方を用いていた）.「線型独立」も「一次独立」も英語では共に “linearly independent”（名詞形では linear independence）である．また，[1] では，“一次独立” は上の定義ではなく，それと同値な以下の (2.7) により導入されていた．

lin-L-1

補題 2.3 $\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ が線型独立であることは次と同値である:

(2.7) 任意の $c_0, ..., c_{k-1}, c'_0, ..., c'_{k-1} \in \mathbb{R}$ に対し，$\sum_{i\in k} c_i\mathbf{a}_i = \sum_{i\in n} c'_i\mathbf{a}_i$ なら，$c_i = c'_i$ がすべて $i \in k$ に対し成り立つ． dim-0

証明. $\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ が線型独立であると仮定する．$c_0, ..., c_{k-1}, c'_0, ..., c'_{k-1} \in \mathbb{R}$ を，

$$\sum_{i\in k} c_i\mathbf{a}_i = \sum_{i\in k} c'_i\mathbf{a}_i \tag{2.8}$$ dim-1

が成り立つようなものとすると，(2.8) の右辺を移項して整理すると，

$$\sum_{i\in k} (c_i - c'_i)\mathbf{a}_i = \mathbf{0} \tag{2.9}$$ dim-2

となる．したがって，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ の線型独立性から，すべての $i \in k$ に対し $c_i - c'_i = 0$ つまり $c_i = c'_i$ となることがわかる．$c_0, ..., c_{k-1}, c'_0, ..., c'_{k-1}$ は (2.8) の成り立つような任意の $\mathbb{R}$ の元だったから，(2.7) が成り立つことが示せた．

逆に，$\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ に対して，(2.7) が成り立つと仮定してみる．このときには，$c_0, ..., c_{k-1} \in \mathbb{R}$ に対し，$\sum_{i\in k} c_i\mathbf{a}_i = \mathbf{0}$ となるなら， (2.7) で，$c'_0 = \cdots c'_{k-1} = 0$ とした場

[25)] 二つの主張（命題） A, B が同値である，とは，二つのに現れるパラメタ（ここでの命題では $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$）をどうとっても，$A$ と B の真偽が一致する，ということである．もう少し具体的には，パラメタをどのようにとっても，“A が成りたてば B が成り立つ” と “B が成りたてば A が成り立つ” の両方が言えることである．

ここで，“B が成りたてば A が成り立つ” はこの命題の対偶命題である “A が成り立たなければ B が成り立たない” と同値であるから，A と B が同値になることを示すには，“A が成りたてば B が成り立つ” と A が成り立たなければ B が成り立たない” を示せばよいことがわかる．以下の証明では，この二つを任意の $\mathbf{a}_0$, ..., $\mathbf{a}_{k-1}$ に対して示している．

合を考えると，すべての $i \in k$ に対し，$c_i = 0$ となることがわかる[26]．したがって，$\mathbf{a}_0$, ..., $\mathbf{a}_{k-1} \in \mathbb{R}^n$ は線形独立である． ❑ (補題 2.3)

$\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ とするとき，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ で，$\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}$ から生成される $\mathbb{R}^n$ の部分空間をあらわすことにする．つまり，

$$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} = \{\sum_{i\in k} c_i\mathbf{a}_i : c_0, ..., c_{k-1} \in \mathbb{R}\} \quad (2.10)$$ dim-2-a

である．より一般的には，$X \subseteq \mathbb{R}^n$ から生成される $\mathbb{R}^n$ の部分空間 $[X]_{\mathbb{R}^n}$ を，

$$[X]_{\mathbb{R}^n} = \{\sum_{i\in k} c_i\mathbf{a}_i : k \in \mathbb{N}, c_0, ..., c_{k-1} \in \mathbb{R}, \mathbf{a}_0, ..., \mathbf{a}_{k-1} \in X\} \quad (2.11)$$ dim-2-a-0

とする．ただし，$k = 0$ のときの $\sum_{i\in k} c_i\mathbf{a}_i$ は $\mathbf{0}$ とすることにして，特殊な場合として，$[\emptyset]_{\mathbb{R}^n} = \{\mathbf{0}\}$ と考えることにする．ここに，$\mathbf{0}$ は $\mathbb{R}^n$ のゼロベクトルである．(2.10) と (2.11) は互いに抵触しないものになっていることに注意する．つまり次が成り立つ:

lin-L-1-a

補題 2.4 $X \subseteq \mathbb{R}^n$ が有限集合のとき，$X = \{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}$ とすると，(2.11) による $[X]_{\mathbb{R}^n}$ と (2.10) による $[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ は一致する．

証明. 演習. ❑ (補題 2.4)

上の記法を用いると，線型独立性は次のように特徴付けることもできる:

lin-L-1-0

補題 2.5 $\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ が線型独立となるのは，以下が成り立つことと同値である:

(2.12) すべての $i \in k$ に対し， $\mathbf{a}_i \notin [\{\mathbf{a}_0, ..., \mathbf{a}_{i-1}\}]_{\mathbb{R}^n}$ が成り立つ． dim-2-0

証明. $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ が線型独立でなければ，$c_0, ..., c_{k-1} \in \mathbb{R}$ で，$[c_i] \neq \mathbf{0}$ だが[27]，$\sum_{i\in k} c_i\mathbf{a}_i = \mathbf{0}$ となるものが存在する．このような組 $c_0, ..., c_{k-1}$ のうちの 1 つについて，$i_0 \in k$ を $c_i \neq 0$ となるような $i \in k$ のうち最大のものとすると，$\mathbf{a}_{i_0} = -\dfrac{1}{c_{i_0}} \sum_{i\in i_0} c_i\mathbf{a}_i$ となるから，$\mathbf{a}_{i_0} \in [\{\mathbf{a}_0, ..., \mathbf{a}_{i_0-1}\}]_{\mathbb{R}^n}$ である．

逆に，ある $i_0 \in k$ に対し，$\mathbf{a}_{i_0} \in [\{\mathbf{a}_0, ..., \mathbf{a}_{i_0-1}\}]_{\mathbb{R}^n}$ だったとすると，$c_0, ..., c_{i_0-1} \in \mathbb{R}$ で，$\mathbf{a}_{i_0} = \sum_{i\in i_0} c_i\mathbf{a}_i$ となるものが存在する．$c_{i_0} = -1$ として，$i \in k \setminus i_0 + 1$ に対しては，$c_i = 0$ とすれば，$c_{i_0} \neq 0$ により $[c_i] \neq \mathbf{0}$ だが，$\sum_{i\in k} c_i\mathbf{a}_i = \mathbf{0}$ となる．したがって，$\mathbf{a}_0$, ..., $\mathbf{a}_{k-1}$ は線型独立ではない． ❑ (補題 2.5)

lin-L-1-1

補題 2.6 $\mathbf{a}_0, , ..., \mathbf{a}_{k-1}, \mathbf{a}'_0, ..., \mathbf{a}'_{k'-1}, \mathbf{b}_0, ..., \mathbf{b}_{\ell-1} \in \mathbb{R}^n$ とする．このとき，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} = [\{\mathbf{a}'_0, ..., \mathbf{a}'_{k'-1}\}]_{\mathbb{R}^n}$ なら，

---

[26] つまり，このときには，$\sum_{i\in k} c_i\mathbf{a}_i = \mathbf{0} = \sum_{i\in k} c'_i\mathbf{a}_i$ だから，(2.7) により，$c_0 = c'_0 = 0$,..., $c_{k-1} = c'_{k-1} = 0$ である．

[27] $[c_i]$ で 各 $i \in k$ に対し，c_i を i-成分とするベクトルを表わしている．したがって，$[c_i] \neq \mathbf{0}$ とは,「$c_i \neq 0$ となるような $i \in k$ が少なくとも 1 つは存在する」ということである．

$$[\{\mathbf{a}_0,...,\mathbf{a}_{k-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}=[\{\mathbf{a}'_0,...,\mathbf{a}'_{k'-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}$$

である．

証明． $\mathbf{c}\in[\{\mathbf{a}_0,...,\mathbf{a}_{k-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}$ とすると，$c_0,...,c_{k-1},\,d_0,...,d_{\ell-1}\in\mathbb{R}$ で，$\mathbf{c}=\sum_{i<k}c_i\mathbf{a}_i+\sum_{j<\ell}d_j\mathbf{b}_j$ となるものがとれる．$\sum_{i<k}c_i\mathbf{a}_i\in[\{\mathbf{a}_0,...,\mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ で，仮定により $[\{\mathbf{a}_0,...,\mathbf{a}_{k-1}\}]_{\mathbb{R}^n}=[\{\mathbf{a}'_0,...,\mathbf{a}'_{k-1}\}]_{\mathbb{R}^n}$ だから，$c'_0,...,c'_{k'-1}\in\mathbb{R}$ で，$\sum_{i<k}c_i\mathbf{a}_i=\sum_{i<k'}c'_i\mathbf{a}'_i$ となるものがとれる．したがって，$\mathbf{c}=\sum_{i<k'}c'_i\mathbf{a}'_i+\sum_{j<\ell}d_j\mathbf{b}_j$ で，$\mathbf{c}\in[\{\mathbf{a}'_0,...,\mathbf{a}'_{k'-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}$ である．

$\mathbf{c}\in[\{\mathbf{a}'_0,...,\mathbf{a}'_{k'-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}$ なら $\mathbf{c}\in[\{\mathbf{a}_0,...,\mathbf{a}_{k-1},\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}$ となることも同様に示せる． ❑ (補題 2.6)

$W\subseteq\mathbb{R}^n$ を $\mathbb{R}^n$ の部分空間とするとき，$\mathbf{a}_0,...,\mathbf{a}_{k-1}\in W$ が W **の基底** (*basis*) **である**とは，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ は線形独立で，$W=[\{\mathbf{a}_0,...,\mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ となること，とする．同様に $X\subseteq W$ が W **の基底** (*basis*) であるとは，X が線型独立で，$W=[X]_{\mathbb{R}^n}$ となること，とする．

補題 2.2 と補題 2.4 により，$X=\{\mathbf{a}_0,...,\mathbf{a}_{k-1}\}$ のとき，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が W の基底であることと X が W の基底であることは同値である．

補題 2.3 により，$\mathbf{a}_0,...,\mathbf{a}_{k-1}$ が W の基底であることは，

(2.13) W の各要素が $\mathbf{a}_0,...,\mathbf{a}_{k-1}$ の線型結合として一意に表わせる dim-2-1

ことと同値である．

lin-E-0

例 1 $\mathbb{R}^n$ の **基本ベクトル** $\mathbf{e}^n_i$ $(i\in n)$ を

$$\mathbf{e}^n_i=\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}\leftarrow\ i\text{ 番目}$$

と定義する．このとき，$\mathbf{e}^n_0,...,\mathbf{e}^n_{n-1}$ は $\mathbb{R}^n$ の基底である（演習）．

lin-L-2

補題 2.7 W を $\mathbb{R}^n$ の部分空間として $\mathbf{a}_0,...,\mathbf{a}_{k-1}\in W$ を線型独立とする．$\mathbf{b}_0,...,\mathbf{b}_{\ell-1}$ を W の基底とするとき，$0\le i_k<i_{k+1}<\cdots<i_{m-1}<\ell$ をうまく選んで，$\{\mathbf{a}_0,...,\mathbf{a}_{k-1},\mathbf{b}_{i_k},...,\mathbf{b}_{i_{m-1}}\}$ が W の基底になるようにできる．

証明． $\mathbf{b}_0,...,\mathbf{b}_{\ell-1}$ は W の基底だから，

(2.14) $[\{\mathbf{b}_0,...,\mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n}=W$ dim-2-2

となることにまず注意しておく．

$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ がすでに W の基底となっているときには $m = 0$ とすればよい．

そうでないときには，$\mathbf{b}_0, ..., \mathbf{b}_{\ell-1}$ の中に $[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ に含まれないものがあるから[28)]，

(2.15) $\mathbf{b}_i \notin [\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ となるような $i \in \ell$ のうち最小のものを i_k とする． dim-3

$i_k, ..., i_{j-1}$ がすでに選ばれたとき，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{j-1}}\}]_{\mathbb{R}^n} = W$ なら，$m = j$ として構成を終える．そうでなければ，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{j-1}}\}]_{\mathbb{R}^n} \subsetneqq W$ だから，(2.14) により（脚注 28) と同様の議論で），$\mathbf{b}_0, ..., \mathbf{b}_{\ell-1}$ のうちで $[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{j-1}}\}]_{\mathbb{R}^n}$ に含まれないものがあることが示せる．そこで，

(2.16) $\mathbf{b}_i \notin [\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{j-1}}\}]_{\mathbb{R}^n}$ となるような $i \in \ell$ のうち最小のものを i_j とする． dim-4

この構成を続けると，ある $j < \ell$ ステップ目で，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{j-1}}\}]_{\mathbb{R}^n} = W$ となって，$m = j$ として，この帰納的構成が停止することがわかる．

ここで構成した，$i_k, ..., i_{m-1}$ が求めるようなものであることを示す．

dim-claim-0

Claim 2.7.1 $0 \le i_k < i_{k+1} < \cdots < i_{m-1}$ である．

⊢ $m \le k+1$ のときには，不等式は自明に成り立つ $m > k+1$ として，この不等式が成り立たないとすると，$k \le k_0 < k_1 < m$ で，$i_{k_1} < i_{k_0}$ となるものがとれる． i_{k_1} は (2.16) により選ばれているので，特に $\mathbf{b}_{i_{k-1}} \notin [\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{k_0-1}}\}]_{\mathbb{R}^n}$ となるが，これは，(2.16) または (2.15) での i_{k_0} の最小性に矛盾である． ⊣ (Claim 2.7.1)

dim-claim-1

Claim 2.7.2 $\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{m-1}}$ は線型独立である．

⊢ $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ が独立であることと，補題 2.5, (2.15) と (2.16) により，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{m-1}}$ は (2.12) を満たす．したがって，再び 補題 2.5 により，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{b}_{i_k}, ..., \mathbf{b}_{i_{m-1}}$ は線型独立である． ⊣ (Claim 2.7.2)

❑ (補題 2.7)

lin-C-0

系 2.8 $\mathbb{R}^n$ の任意の線型独立な要素 $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ は $\mathbb{R}^n$ の（有限個の要素からなる）基底に拡張できる．

証明. $\mathbf{e}^n_0, ..., \mathbf{e}^n_{n-1}$ は $\mathbb{R}^n$ の基底だから， $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ と $\mathbf{e}^n_0, ..., \mathbf{e}^n_{n-1}$ に対して 補題 2.7 を適用すればよい． ❑ (系 2.8)

lin-L-3

補題 2.9 $\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ とするとき，$0 \le i_0 < i_1 < \cdots < i_{m-1} < k$ で，

(2.17) $[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} = [\{\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_{m-1}}\}]_{\mathbb{R}^n}$, かつ， dim-5

(2.18) $\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_{m-1}}$ は線型独立 dim-6

となるようなものが存在する．

証明． $i_0 = 0$ とする．$i_0, ..., i_j$ がすでに求まったとき，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} = [\{\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_j}\}]_{\mathbb{R}^n}$ なら，$m = j+1$ として構成を終える．そうでなければ，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} \supsetneqq [\{\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_j}\}]_{\mathbb{R}^n}$ だから，$\mathbf{a}_i \notin [\{\mathbf{a}_{i_0}, \mathbf{a}_{i_1}, ..., \mathbf{a}_{i_j}\}]_{\mathbb{R}^n}$ となるような $i \in k$ が存在するから，そのようなもののうち最小のものを i_{j+1} とする．この構成は高々 k ステップで終了し，そのときには，構成の定義から (2.17) が成立している．

Claim 2.7.1, Claim 2.7.2 と同様の議論で，ここで構成した $i_0, ..., i_{m-1}$ が求めるようなものになっていることが示せる． ❑ (補題 2.9)

次の 補題 2.10 は，[1], p.140 の補題に対応するものである．[1] での，この補題の証明により示すことができるが，[1] での議論の道筋では，$\mathbb{R}^n$ の部分空間が無限集合になっているような基底を持つ，という（実際にはありえないことが後で証明されるところの）可能性が処理しきれていない．この問題を選択公理を用いずに処理するためには[29]，以下でのような議論の筋道が必要になってくる．

well-def-of-dim

補題 2.10 n を任意の自然数とする．ある自然数 k に対し，$\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^n$ が $\mathbb{R}^n$ の基底になっているなら，$k = n$ が成り立つ．

証明． [1] の p.140 の補題の証明を参照． ❑ (補題 2.10)

lin-C-1

系 2.11 $W \subseteq \mathbb{R}^n$ を $\mathbb{R}^n$ の部分空間として，$X \subseteq W$ を線型独立とするとき，$|X| \leq n$ が成り立つ[30]．

証明． 要素を n 個以上持つ W の基底 X があったとして矛盾を導く．$\mathbf{a}_0, ..., \mathbf{a}_n \in X$ を互いに異る $n+1$ 個の X の要素とする．$\mathbf{a}_0, ..., \mathbf{a}_n$ は線型独立だから，系 2.8 により， $\mathbf{a}_0, ..., \mathbf{a}_n$ は W の（有限個の要素からなる）基底 B に拡張できるが，B の要素の数は $n+1$ 大きいものになるから，補題 2.10 に矛盾である． ❑ (系 2.11)

正の自然数 n に対して，$\mathbb{R}^n$ には標準的な基底 $e^n_0, ..., e^n_{n-1}$ が存在するのだった．しかし，$\mathbb{R}^n$ のすべての部分空間が基底を持つかどうかは，よく考えてみると，それほど自明なことではない．選択公理を仮定すると，すべての線型空間は基底を持つことが証明でき

---

[28] もし $\mathbf{b}_0, ..., \mathbf{b}_{\ell-1}$ がすべて $[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n}$ に含まれているとすると，(2.14) により，$W = [\{\mathbf{b}_0, ..., \mathbf{b}_{\ell-1}\}]_{\mathbb{R}^n} \subseteq [\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} \subseteq \mathbb{R}^n$ だから，$[\{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}]_{\mathbb{R}^n} = \mathbb{R}^n$ が成り立つ．したがって，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ はすでに W の基底になっていることになり，仮定に矛盾である．

[29] 一般の線型空間で基底や次元を議論するためには選択公理が必要である．さらに具体的に言うと,「すべての線型空間に基底が存在する」という主張は，集合論の他の公理上，選択公理と同値になることが知られている．ここでは，ある自然数 n に対する，$\mathbb{R}^n$ の部分空間という特別な形の線型空間に議論を制限しているため，選択公理なしで基底や次元の議論ができているのだが，そのことを明らかにするためには，実際に選択公理を用いずに議論を展開してみせる必要がある．

[30] 集合 X に対し，$|X|$ で X の要素の個数をあらわす．

るため，これを仮定して議論していれば，基底の存在で悩む必要はないのだが，$\mathbb{R}^n$ の部分空間に関しては，選択公理の仮定なしで基底の存在が保証できる:

lin-C-2

系 2.12 $W \subseteq \mathbb{R}^n$ を $\mathbb{R}^n$ の部分空間として，$\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in W$ が線型独立のとき，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ を拡張して W の基底を作ることができる．特に，W は基底を持つ．

証明. $\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ を W の線型独立な要素とする．このとき，$\mathbf{a}_k, ..., \mathbf{a}_{m-1}$ を，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{a}_k, ..., \mathbf{a}_{m-1}$ が W の基底になるように，次のように帰納的に構成する: $\mathbf{a}_k, ..., \mathbf{a}_{\ell-1}$ が，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{a}_k, ..., \mathbf{a}_{\ell-1}$ が線型独立になるように，既に選ばれたとき，$[\{\mathbf{a}_0, ..., \mathbf{a}_{\ell-1}\}]_{\mathbb{R}^n} = W$ なら，$m = \ell$ として構成を終える．そうでなければ，

$$\mathbf{a}_\ell \in W \setminus [\{\mathbf{a}_0, ..., \mathbf{a}_{\ell-1}\}]_{\mathbb{R}^n} \tag{2.19}$$

dim-8

となるように $\mathbf{a}_\ell$ をとる．仮定から，$\mathbf{a}_0, ..., \mathbf{a}_{\ell-1}$ は線型独立だから，(2.19) と補題 2.5 により，$\mathbf{a}_0, ..., \mathbf{a}_{\ell-1}, \mathbf{a}_\ell$ も線型独立である．系 2.11 により，この構成は n ステップ内の，あるステップ $m-1$ で停止する．このときには，構成の仕方から $[\mathbf{a}_0, ..., \mathbf{a}_{m-1}]_{\mathbb{R}^n} = W$ が成立しており，したがって，$\{\mathbf{a}_0, ..., \mathbf{a}_{m-1}\}$ は W の基底である． ❑ (系 2.12)

lin-C-3

系 2.13 $W \subseteq \mathbb{R}^n$ を $\mathbb{R}^n$ の部分空間として，X と X' をそれぞれ W の基底とする．このとき X も X' も W の有限部分集合で，$|X| = |X'|$ が成り立つ．

証明. X と X' がともに W の有限部分集合になることは，系 2.11 によりよい．$X = \{\mathbf{a}_0, ..., \mathbf{a}_{k-1}\}$, $X' = \{\mathbf{a}'_0, ..., \mathbf{a}'_{k'-1}\}$ で，$k < k'$ だったとして矛盾を導く．

系 2.8 により，$\mathbf{a}_0, ..., \mathbf{a}_{k-1}$ を拡張して $\mathbb{R}^n$ の基底 $\tilde{X} = \{\mathbf{a}_0, ..., \mathbf{a}_{k-1}, \mathbf{a}_k, ..., \mathbf{a}_{\ell-1}\}$ が得られる．ここで，$\tilde{X}' = \{\mathbf{a}'_0, ..., \mathbf{a}'_{k'-1}, \mathbf{a}_k, ..., \mathbf{a}_{\ell-1}\}$ とすると，補題 2.6 と 補題 2.5 により，$\tilde{X}'$ は線型独立になる．したがって，$\tilde{X}'$ も $\mathbb{R}^n$ の基底になるが，$|\tilde{X}'| > |\tilde{X}|$ だから，これは 補題 2.10 に矛盾である． ❑ (系 2.13)

n をある自然数として，W を $\mathbb{R}^n$ の部分空間とする．このとき，系 2.12 により，W の基底 X が存在する．系 2.11 により，X は有限集合で，系 2.13 により，W のすべての基底の要素の数は $|X|$ と一致する．そこで $|X|$ を W の **次元** (*dimension*) とよび，$\dim(W)$ とあらわすことにする．例 1 により $\dim(\mathbb{R}^n) = n$ である．

次の定理では $\mathbb{R}^n$ の部分空間が有限次元であることが本質的である．実際，無限次元の線型空間に対しては，次の (a) も (b) も，一般には成り立たない．

定理 2.14 ある自然数 n に対し W と W' を $\mathbb{R}^n$ の部分空間とする．

(a) $W \subsetneqq W' \subseteq \mathbb{R}^n$ なら，$\dim(W) < \dim(W')$ が成り立つ．

(b) $X \subseteq W$ が線型独立で $|X| = \dim(W)$ なら，X は W の基底である．

証明. (a): X を W の基底とする．系 2.12 により X は W' の基底 X' に拡張できるが，$W \subsetneqq W'$ から X は W' の基底でないから，$X \subsetneqq X'$ である．X と X' は有限集合だから，このことから $\dim(W) = |X| < |X'| = \dim(W')$ がわかる．

(b): $X \subseteq W$ が線型独立なら，系 2.12 により X は W の基底 X' に拡張できるが，$|X'| = \dim(W)$ で，$\dim(W)$ は有限だから，$X = X'$ である．つまり X はすでに W の基底となっている． ❑ (定理 2.14)

2.3 線型写像の行列表現

matrix

以下では m, n は自然数とする．ただし m （または n）が 0 のときは，$\mathbb{R}^m$ （または $\mathbb{R}^n$）は $\{0\}$ のこととする．区別をする必要があるときには，$\mathbb{R}^m$ のゼロベクトルを $\mathbf{0}_{\mathbb{R}^m}$, $\mathbb{R}^n$ のゼロベクトルを $\mathbf{0}_{\mathbb{R}^n}$ であらわすことにする．

写像 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ が **線型写像** である，とは，すべての $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ と $a \in \mathbb{R}$ に対し，

(2.20) $\varphi(\mathbf{a}+\mathbf{b}) = \varphi(\mathbf{a}) + \varphi(\mathbf{b})$; lin-map-0

(2.21) $\varphi(a\mathbf{a}) = a\varphi(\mathbf{a})$ lin-map-1

が成り立つことである．

ここで，等式 (2.20) の左辺の $+$ はベクトル空間 $\mathbb{R}^m$ の加法であるのに対し，この式の右辺の $+$ はベクトル空間 $\mathbb{R}^n$ での加法であることに注意する．同様に等式 (2.21) の左辺の “a 倍” はベクトル空間 $\mathbb{R}^n$ でのものであるのに対し，この式の右辺での “a 倍” はベクトル空間 $\mathbb{R}^n$ での演算である．

このことを注意してもう一度線型写像の定義を見てみると，

> φ が $\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像である，ということは，φ が $\mathbb{R}^m$ の加法と定数倍を保存しつつ $\mathbb{R}^m$ の要素を $\mathbb{R}^n$ に移すような写像になっていることである

と解釈することができることがわかる．特に φ が 1 対 1 写像のときには，φ は $\mathbb{R}^m$ の（加法と定数倍に関する構造を保つような） $\mathbb{R}^n$ への埋め込みになっている，と考えることができる．

lin-a-0

例 2 (1) すべての $\mathbf{a} \in \mathbb{R}^m$ に対し，$\mathbf{0}_{\mathbb{R}^n}$ を対応させる写像は，$\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像である．

(2) $m \geq n > 0$ として，$\mathbf{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix} \in \mathbb{R}^m$ に，$\mathbf{a}' = \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{R}^n$ を対応させる写像は，$\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像である．

(3) $m < n$ のとき，$\mathbf{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix} \in \mathbb{R}^m$ に，$\mathbf{a}' = \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ を対応させる写像は，$\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像である．

(4) $n, m > 0$ として，A を $n \times m$-行列とするとき[31)]，$\varphi_A : \mathbb{R}^m \to \mathbb{R}^n$ を，$\mathbf{a} \in \mathbb{R}^m$ に

[31)] “$m \times n$” ではなく “$n \times m$” となっていることに注意する（誤殖ではない!!）．

対して $\varphi_A(\mathbf{a}) = A\mathbf{a}$ （行列 A とベクトル $\mathbf{a}$ の積）となるものとして定義すると，φ_A は，$\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像となる．

上の 例 2, (4) は本質的である: 実は，すべての線型写像 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ が，ある $n \times m$-行列 A により φ_A として一意に表わせることを後で示す（補題 2.18 を参照）．

lin-a

補題 2.15 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ が線型写像で，$k \in \mathbb{N}$, $\mathbf{a}_0, ..., \mathbf{a}_{k-1} \in \mathbb{R}^m$, $c_0, ..., c_{k-1} \in \mathbb{R}$ のとき，

$$\varphi(\sum_{i\in k} c_i\mathbf{a}_i) = \sum_{i\in k} c_i\varphi(\mathbf{a}_i)$$

が常に成り立つ．

証明． k に関する帰納法で示せる（演習）． ❑ (補題 2.15)

写像 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ が線型写像のとき，φ の **核** (*kernel*) と **像** (*image*) をそれぞれ以下で定義する:

(2.22) $Ker(\varphi) = \{\mathbf{a} \in \mathbb{R}^m : \varphi(\mathbf{a}) = \mathbf{0}_{\mathbb{R}^n}\}$; lin-map-2

(2.23) $Im(\varphi) = \{\mathbf{b} \in \mathbb{R}^n : \mathbf{b} = \varphi(\mathbf{a}) \text{ となる } \mathbf{a} \in \mathbb{R}^m \text{ が存在する }\}$. lin-map-3

核も像も $\mathbb{R}^m$ から $\mathbb{R}^n$ への（必ずしも線型写像ではない）任意の写像に対しても同様に定義できるが，φ が線型写像のときには，これらは特別な意味を持つ対象となる[32]:

lin-0

補題 2.16 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ を線型写像とする．このとき以下が成り立つ:

(1) $\varphi(\mathbf{0}_{\mathbb{R}^m}) = \mathbf{0}_{\mathbb{R}^n}$ である．特に，$\mathbf{0}_{\mathbb{R}^m} \in Ker(\varphi)$, $\mathbf{0}_{\mathbb{R}^n} \in Im(\varphi)$ が成り立つ．

(2) φ が 1 対 1 写像となるのは， $Ker(\varphi) = \{\mathbf{0}_{\mathbb{R}^m}\}$ となるちょうどそのときである．

(3) $Ker(\varphi)$ は $\mathbb{R}^m$ の部分空間である．

(4) $Im(\varphi)$ は $\mathbb{R}^n$ の部分空間である．

(5) (次元定理) $\dim(Ker(\varphi)) + \dim(Im(\varphi)) = m$ である．特に $\dim(Im(\varphi)) \leq n$ だから， $\dim(Ker(\varphi)) \geq m - n$ が成り立つ．

証明． (1): $\mathbf{0}_{\mathbb{R}^m} = 0\mathbf{0}_{\mathbb{R}^m}$ （右辺は $\mathbf{0}_{\mathbb{R}^m}$ のゼロ倍）に注意すると，(2.21) により，

$$\varphi(\mathbf{0}_{\mathbb{R}^m}) = \varphi(0\mathbf{0}_{\mathbb{R}^m}) = 0\varphi(\mathbf{0}_{\mathbb{R}^m}) = \mathbf{0}_{\mathbb{R}^n}$$

である．

(2): φ が 1 対 1 写像なら，$Ker(\varphi)$ はただ 1 つの要素からなるから，(1) により，$Ker(\varphi) = \{\mathbf{0}_{\mathbb{R}^m}\}$ でなくてはならないことがわかる．

逆に $Ker(\varphi) = \{\mathbf{0}_{\mathbb{R}^m}\}$ だと仮定してみる．このとき，任意の $\mathbf{a}$, $\mathbf{a}' \in \mathbb{R}^m$ に対し，$\varphi(\mathbf{a}) = \varphi(\mathbf{a}')$ とすると，(2.20) と (2.21) により，

[32] 集合論で通常使われる記号を用いると，$Ker(\varphi)$ と $Im(\varphi)$ はそれぞれ，$Ker(\varphi) = \varphi^{-1}[\{\mathbf{0}_{\mathbb{R}^n}\}]$, $Im(\varphi) = \varphi[\mathbb{R}^m]$ と表わすこともできる．ただし，ここで用いられている “φ^{-1}” は φ の逆対応（二項関係としての φ の逆を考えたもの）を表わしていて，必ずしも φ の逆写像ではない（φ が逆写像を持たないときにもこの記号を使う）ことに注意する．

$$\varphi(\mathbf{a}-\mathbf{a}')=\varphi(\mathbf{a})-\varphi(\mathbf{a}')=\mathbf{0}_{\mathbb{R}^m}$$

となるから，仮定により，$\mathbf{a}-\mathbf{a}'=\mathbf{0}_{\mathbb{R}^m}$, したがって， $\mathbf{a}=\mathbf{a}'$ である．よって，φ は 1 対 1 である．

(3): 任意の $\mathbf{a}_0, \mathbf{a}_1 \in Ker(\varphi)$ と $c_0, c_1 \in \mathbb{R}$ に対し，$\sum_{i\in 2} c_i\mathbf{a}_i \in Ker(\varphi)$ となることを示せばよいが，これは，

$$\varphi(\sum_{i\in 2} c_i\mathbf{a}_i)=\sum_{i\in 2} c_i\varphi(\mathbf{a}_i)=\sum_{i\in 2} c_i\mathbf{0}_{\mathbb{R}^n}=\mathbf{0}_{\mathbb{R}^n}$$

により成り立つ．

(4): 任意の $\mathbf{b}_0, \mathbf{b}_1 \in Im(\varphi)$ と，$c_0, c_1 \in \mathbb{R}$ に対し，$\sum_{i\in 2} c_i\mathbf{b}_i \in Im(\varphi)$ となることを示せばよい．

$\mathbf{b}_0, \mathbf{b}_1 \in Im(\varphi)$ により，$\mathbf{a}_0, \mathbf{a}_1 \in \mathbb{R}^m$ で，$\mathbf{b}_0=\varphi(\mathbf{a}_0)$, $\mathbf{b}_1=\varphi(\mathbf{a}_1)$ となるものがとれる．このとき，

$$\sum_{i\in 2} c_i\mathbf{b}_i=\sum_{i\in 2} c_i\varphi(\mathbf{a}_i)=\varphi(\sum_{i\in 2} c_i\mathbf{a}_i)$$

だから，$\sum_{i\in 2} c_i\mathbf{b}_i \in Im(\varphi)$ である．

(5): $\dim(Ker(\varphi))=d_1$ として，$\mathbf{a}_0,...,\mathbf{a}_{d_1-1}$ を $Ker(\varphi)$ の基底とする．このとき，系 2.8 と 補題 2.10 により，$\mathbf{a}_0,...,\mathbf{a}_{d_1-1}$ の拡張 $\mathbf{a}_0,...,\mathbf{a}_{d_1-1}, \mathbf{a}_{d_1},...,\mathbf{a}_m$ で $\mathbb{R}^m$ の基底になっているようなものがとれる．$d_2=m-d_1$ として，$\mathbf{b}_0=\varphi(\mathbf{a}_{d_1})$, $\mathbf{b}_1=\varphi(\mathbf{a}_{d_1+1})$,..., $\mathbf{b}_{d_2-1}=\varphi(\mathbf{a}_{m-1})$ とする．$\mathbf{b}_0,...,\mathbf{b}_{d_2-1}$ が $Im(\varphi)$ の基底になっていることが示せれば十分である．

まず，$\mathbf{b}_0,...,\mathbf{b}_{d_2-1}$ の線型独立性を示す．もし $\mathbf{b}_0,...,\mathbf{b}_{d_2-1}$ が線型独立でなかったとすると，すべては 0 でない $c_0,...,c_{d_2-1} \in \mathbb{R}$ で，$\sum_{i<d_2} c_i\mathbf{b}_i=\mathbf{0}$ となるようなものが存在する．補題 2.15 により，$\sum_{i<d_2} c_i\mathbf{b}_i=\sum_{i<d_2} c_i\varphi(\mathbf{a}_{d_1+i})=\varphi(\sum_{i<d_2} c_i\mathbf{a}_{d_1+i})$ だから，$\sum_{i<d_2} c_i\mathbf{a}_{d_1+i} \in Ker(\varphi)$ となる．したがって，$\mathbf{a}_0,...,\mathbf{a}_{d_1-1}$ の選び方から，$c'_0,...,c'_{d_1-1} \in \mathbb{R}$ で，$\sum_{i<d_2} c_i\mathbf{a}_{d_1+i}=\sum_{j<d_1} c'_j\mathbf{a}_j$ となるものがある．したがって，$\sum_{j<d_1} c'_j\mathbf{a}_j-\sum_{i<d_2} c_i\mathbf{a}_{d_1+i}=\mathbf{0}$ となるが，$c_i, i<d_2$ は全部は 0 でないのだったから，これは，$\mathbf{a}_i, i<m$ の独立性に矛盾である．

あとは，$\mathbf{b}_0,...,\mathbf{b}_{d_2-1}$ が $Im(\varphi)$ を張ることを示せばよい．このためには任意の $\mathbf{b} \in Im(\varphi)$ が $\mathbf{b}_0,...,\mathbf{b}_{d_2-1}$ の線型結合としてあらわされることを示せばよい．$\mathbf{b} \in Im(\varphi)$ から，$\mathbf{a} \in \mathbb{R}^m$ で $\varphi(\mathbf{a})=\mathbf{b}$ となるものがとれる．$\mathbf{a}_0,...,\mathbf{a}_{m-1}$ は $\mathbb{R}^m$ の基底だったから，c_0, ..., $c_{m-1} \in \mathbb{R}$ で $\mathbf{a}=\sum_{i<n} c_i\mathbf{a}_i$ となるものがとれるが，このとき φ の線型性から，

$$\mathbf{b}=\varphi(\mathbf{a})=\varphi(\sum_{i<n} c_i\mathbf{a}_i)=\sum_{i<n} c_i\varphi(\mathbf{a}_i)$$

である．ここで，$\varphi(\mathbf{a}_i)=\mathbf{0}, i<d_1$ で，$\varphi(\mathbf{a}_{d_1+j})=\mathbf{b}_j, j<d_2$ だから， $\mathbf{b}=\sum_{j<d_2} c_{d_1+j}\mathbf{b}_j$ である． ❑ (補題 2.16)

lin-1

補題 2.17 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ と $\psi : \mathbb{R}^m \to \mathbb{R}^n$ を線型写像とする．$\mathbf{a}_0, ..., \mathbf{a}_{m-1}$ を $\mathbb{R}^m$ の基底とするとき[33]，

(2.24) $\varphi(\mathbf{a}_0) = \psi(\mathbf{a}_0)$,..., $\varphi(\mathbf{a}_{m-1}) = \psi(\mathbf{a}_{m-1})$ lin-map-4

なら，φ と ψ は等しい[34]．

証明． $\mathbf{a}$ を $\mathbb{R}^m$ の任意の要素とするとき，$\varphi(\mathbf{a}) = \psi(\mathbf{a})$ が成り立つことを示せばよい．これは次のようにして示せる: $\mathbf{a}_0, ..., \mathbf{a}_{m-1}$ は $\mathbb{R}^m$ の基底だったから，$c_0, ..., c_{m-1} \in \mathbb{R}$ で $\mathbf{a} = \sum_{i \in m} c_i \mathbf{a}_i$ となるものがとれる．このとき，φ と ψ の線型性 (補題 2.15) と，(2.24) の仮定から，

$$\varphi(\mathbf{a}) = \varphi(\sum_{i \in m} c_i \mathbf{a}_i) = \sum_{i \in m} c_i \varphi(\mathbf{a}_i) = \sum_{i \in m} c_i \psi(\mathbf{a}_i) = \psi(\sum_{i \in m} c_i \mathbf{a}_i) = \psi(\mathbf{a})$$

が成り立つ． ❑ (補題 2.17)

任意の線型写像 $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ に対し，$\mathbf{a}_j = \varphi(\mathbf{e}^m_j)$ $(j < m)$ として，$M_\varphi = [\mathbf{a}_0, ..., \mathbf{a}_{m-1}]$ とする．M_φ は $n \times m$-行列である．

$n \times m$-行列 M に対し，線型写像 $\varphi_M : \mathbb{R}^m \to \mathbb{R}^n$ を $\varphi_M(\mathbf{a}) = M\mathbf{a}$ で定義するのだった（$M\mathbf{a}$ は $n \times m$-行列 M と m-次元ベクトル $\mathbf{a}$ の積である）．

lin-2

補題 2.18 (1) $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ を任意の線型写像とするとき，$\varphi_{M_\varphi} = \varphi$ である．特に，M_φ は，$\varphi_A = \varphi$ となるような唯一の $n \times m$ 行列 A である．

(2) A を任意の $n \times m$-行列とするとき，$M_{\varphi_A} = A$ である．特に，φ_A は，$M_\varphi = A$ となるような唯一の $\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像である．

証明． (1): $j < m$ に対し，$\varphi_{M_\varphi}(\mathbf{e}^m_j) = M_\varphi \mathbf{e}^m_j = [\mathbf{a}_0, ..., \mathbf{a}_{m-1}]\mathbf{e}^m_j = \mathbf{a}_j = \varphi(\mathbf{e}^m_j)$ である．したがって，補題 2.17 により，$\varphi_{M_\varphi} = \varphi$ が示せた．

M_φ の一意性の証明には，二つの異る $n \times m$-行列 A, B に対して $\varphi_A \neq \varphi_B$ が常に成り立つこをが示せれば十分である．$A = [\mathbf{a}_0, ..., \mathbf{a}_{m-1}]$, $B = [\mathbf{b}_0, ..., \mathbf{b}_{m-1}]$ として，$A \neq B$ なら，$j^* < m$ で，$\mathbf{a}_{j^*} \neq \mathbf{b}_{j^*}$ となるものが存在するが，このとき，

$$\varphi_A(\mathbf{e}^m_{j^*}) = A\mathbf{e}^m_{j^*} = \mathbf{a}_{j^*} \neq \mathbf{b}_{j^*} = B\mathbf{e}^m_{j^*} = \varphi_B(\mathbf{e}^m_{j^*})$$

となるから，$\varphi_A \neq \varphi_B$ である．

(2): $A = [\mathbf{a}_0, ..., \mathbf{a}_{m-1}]$ として，$M_{\varphi_A} = [\mathbf{a}'_0, ..., \mathbf{a}'_{m-1}]$ とすると，M_{φ_A} の定義から，$j < m$ に対し $\mathbf{a}'_j = \varphi_A(\mathbf{e}^m_j) = A\mathbf{e}^m_j = \mathbf{a}_j$ である．したがって，$A = M_{\varphi_A}$ となることがわかる．

---

[33] 補題 2.10 により，$\mathbb{R}^m$ の任意の基底は m 個のベクトルからなる．

[34] 2 つの写像 f, g が等しいとは，f と g の定義域が等しく，定義域のすべての要素 x について $f(x) = g(x)$ が成り立つことである（このことは写像の概念の（厳密な）定義から導ける）．

φ_A の一意性の証明ためには，φ, ψ を異る $\mathbb{R}^m$ から $\mathbb{R}^n$ への線型写像とするとき，$\varphi \neq \psi$ なら，$M_\varphi \neq M_\psi$ となることが示せればよい．$\varphi \neq \psi$ とすると，補題 2.17 により，$j^* < m$ で，$\varphi(\mathbf{e}^m_{j^*}) \neq \psi(\mathbf{e}^m_{j^*})$ となるものがとれる．このとき，

$$M_\varphi = [\varphi(\mathbf{e}^m_0), ..., \varphi(\mathbf{e}^m_{j^*}), ..., \varphi(\mathbf{e}^m_{m-1})] \neq [\psi\mathbf{e}^m_0), ..., \psi(\mathbf{e}^m_{j^*}), ..., \psi(\mathbf{e}^m_{m-1})] = M_\psi$$

である． ❑ (補題 2.18)

写像 $\varphi : X \to Y$, $\psi : Y \to Z$ の合成を $\psi \circ \varphi$ であらわすことにする．$\psi \circ \varphi : A \to Z$ で，$x \in X$ に対し，$(\psi \circ \varphi)(x) = \psi(\varphi(x))$ である．

lin-2-0

補題 2.19 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ と $\psi : \mathbb{R}^m \to \mathbb{R}^\ell$ を線型写像とするとき，$\psi \circ \varphi$ も線型写像である．

証明． 演習． ❑ (補題 2.19)

対応 $A \mapsto \varphi_A$ は，行列の積と線型写像の合成を対応させるものになっている．

lin-2-1

補題 2.20 A を $\ell \times m$-行列，B を $m \times n$-行列とするとき，$\varphi_{AB}, \varphi_A \circ \varphi_B$ はともに $\mathbb{R}^n$ から $\mathbb{R}^\ell$ への線型写像だが，さらに $\varphi_{AB} = \varphi_A \circ \varphi_B$ が成り立つ．

証明． 任意の $\mathbf{a} \in \mathbb{R}^n$ に対し，$\varphi_{AB}(\mathbf{a}) = (AB)\mathbf{a} = A(B\mathbf{a}) = A\varphi(\mathbf{a}) = \varphi_A(\varphi_B(\mathbf{a})) = \varphi_A \circ \varphi_B(\mathbf{a})$ である． ❑ (補題 2.20)

2.4 連立方程式の解の全体の構造

sytem-of-eqs

前節までに導入した用語や結果を連立一次方程式の解の全体の構造に関して適用すると，次が分る:

eq-system

定理 2.21 A を $n \times m$-行列として，連立方程式 $A\mathbf{x} = \mathbf{c}$ を考える，ここに $\mathbf{x}$ は m 個の変数 x_0,..., x_{m-1} を成分とするベクトルで $\mathbf{c} \in \mathbb{R}^n$ である．このとき:

(1) $A\mathbf{x} = \mathbf{c}$ に対応する斉次方程式 $A\mathbf{x} = \mathbf{0}$ の解の全体 L_0 は $Ker(\varphi_A)$ と一致する．したがって，補題 2.16, (3) により，L_0 は $\mathbb{R}^m$ の部分空間で，補題 2.16, (5) により，$\dim(L_0) \geq m - n$ である．

(2) $A\mathbf{x} = \mathbf{c}$ が解を持つのは，$\mathbf{c} \in Im(\varphi_A)$ となるちょうどそのときである．このときには，$\mathbf{a}$ をこの連立方程式の解の一つとすれば，$A\mathbf{x} = \mathbf{c}$ の解の全体を L とすると

$$L = L_0 + \mathbf{a} = \{\mathbf{b} + \mathbf{a} : \mathbf{b} \in L_0\} \quad (2.25)$$

lin-3

である．特に $A\mathbf{x} = \mathbf{c}$ が解を持てば，解の一般解は，m 個以下 $m - n$ 個以上のパラメタを持つものとなる． ❑

証明． (2.25) を示せば，残りは前に述べたことから明らかである．

$\mathbf{d} \in L_0 + \mathbf{a}$ とすれば，$\mathbf{b} \in L_0$ で $\mathbf{d} = \mathbf{b} + \mathbf{a}$ となるものがあるが，$A\mathbf{b} = \mathbf{0}$ だから，$A\mathbf{d} = A(\mathbf{b} + \mathbf{a}) = A\mathbf{b} + A\mathbf{a} = \vec{0} + A\mathbf{a} = A\mathbf{a} = \mathbf{c}$ となり $\mathbf{d}$ も $A\mathbf{x} = \mathbf{c}$ の解である，つまり $\mathbf{d} \in L$ であることがわかる．

逆に $\mathbf{d} \in L$ とすると，$\mathbf{b} = \mathbf{d} - \mathbf{a}$ として $\mathbf{d} = \mathbf{b} + \mathbf{a}$ だが，このとき $A\mathbf{b} = A(\mathbf{d} - \mathbf{a}) = A\mathbf{d} - A\mathbf{a} = \mathbf{c} - \mathbf{c} = \mathbf{0}$ となるから，$\mathbf{b}$ は $A\mathbf{x} = \mathbf{0}$ の解で，つまり $\mathbf{b} \in L_0$ となり，$\mathbf{d} \in L_0 + \mathbf{a}$ がわかる． ❑ (定理 2.21)

References

[1] 長谷川 浩司，線型代数，日本評論社 (2004).

3 確率と統計

statistics

3.1 付値の和の期待値

expected-value

次の演習問題とその解答を見てください:

coins

演習問題 3.1 10 円硬貨が 7 枚，50 円硬貨が 5 枚，100 円硬貨が 8 枚ある．これらの中から無作為に[35] 3 枚取り出すとき，それらの金額の和の期待値を求めよ。

解答: 上のような貨幣の全体から硬貨を無作為に 1 枚取り出すときの金額の期待値は，

$$\frac{10\times 7+50\times 5+100\times 8}{7+5+8}=56$$

だから，この貨幣の全体から硬貨を無作為に 3 枚取り出したときの金額の期待値は，$56\times 3=168$ である． ❑

上の議論は本当に正しいのでしょうか？ 特に脳天気に 3 倍しているところなど，本当にこれでいいのか不安です．しかし，この計算で得られる答え自身は正しいのです．そこで，以下で，上の計算がなぜ正しい答を導くのかを見てみようと思います．

まず，上の論法をほぼそのまま正当化することを試みてみましょう．そのために，次のような一般的な定理を用意しておきます（証明は，たとえば，[小寺平治:ゼロから学ぶ統計解析] の p.65 を参照）:

定理 3.2 $X_0,\ldots,X_{n-1}$ を確率変数とする．$a_0,\ldots,a_{n-1}$ を実数とするとき，

$$E(a_0X_0+\cdots a_{n-1}X_{n-1})=a_0E(X_0)+\cdots+a_{n-1}E(X_{n-1}) \tag{3.1}$$

p-0

が成り立つ．ただし，$E(\cdots)$ で確率変数 $\cdots$ の期待値 (expected value) をあらわす．

ここで,「確率変数とは測定（あるいは観測）をすると実数値が返ってくるようなもの」のことだと思ってください．たとえば上の例では,「10 円硬貨 7 枚，50 円硬貨 5 枚，100 円硬貨 8 枚から中から無作為に三枚とり出したときの一枚目の硬貨の金額」を確率変数ととらえることができます．一方 $a_0X_0+\cdots+a_{n-1}X_{n-1}$ は X_0 の返した値の a_0 倍，$\ldots$， X_{n-1} の返した値の a_{n-1} 倍を全部足して得られる値を返す確率変数のことです．

簡単のために，$a_i,\ i\in I$ というとびとびの値をとる確率変数 X を考えることにして，X が値 a_i を返す確率が p_i とすると，X の期待値は，$\sum_{i\in I}a_ip_i$ で定義されます．

[35] ここでは「無作為」とは “randomly” の訳語として使っています．無作為という単語は難しすぎると判断されたためか高校の教科書には出ていないようです．ようするに「でたらめに」ということですが，どうもこの「でたらめ」という言葉も高校の教科書では使われていないようです．多分「でたらめ」のネガティヴな響きを嫌った結果だのだろうと思うのですが，科学では，日常語のしがらみをたちきって思いきった言葉の使い方をする，ということがぜひとも必要なのです．日常語の「でたらめ」には，作為的にでたらめをする，という意味も含まれているので，誤解をまねく可能性ももちろんあるわけですが，それにもかかわらず，このようなビビッドな日本語を何かのガイドラインのようなもので教科書から締め出してしまうのはやはり問題のような気がします．

上の定理で注意したいのは，$X_0,...,X_{n-1}$ の間の関係如何によらず (3.1) が成り立つということです．

さて，定理 3.2 を用いて上の問題を再考してみましょう．X_0, X_1, X_2 をそれぞれ「10 円硬貨 7 枚，50 円硬貨 5 枚，100 円硬貨 8 枚から中から無作為に三枚とり出したときの一枚目の硬貨の金額」,「10 円硬貨 7 枚，50 円硬貨 5 枚，100 円硬貨 8 枚から中から無作為に三枚とり出したときの二枚目の硬貨の金額」,「10 円硬貨 7 枚，50 円硬貨 5 枚，100 円硬貨 8 枚から中から無作為に三枚とり出したときの三枚目の硬貨の金額」を返す確率変数とします．このとき，$E(X_0+X_1+X_2)$ が求める期待値となります．定理 3.2 により $E(X_0+X_1+X_2)=E(X_0)+E(X_1)+E(X_2)$ です．ところが，各 $X_i, i=0,1,2$ の期待値 $E(X_i)$ はそれぞれ 56 だから，$E(X_0+X_1+X_2)=56\times 3=168$ となる．

さて，これで一件落着のように見えますが，ごく厳密には，上で「各 $X_i, i=0,1,2$ の期待値 $E(X_i)$ はそれぞれ 56 だから」と言ったところで，まだ問題が残っています.「10 円硬貨 7 枚，50 円硬貨 5 枚，100 円硬貨 8 枚から中から無作為に一枚とり出したときの一枚目の硬貨の金額」を与える確率変数だったら期待値が 56 になることは明らかですが，三枚とったうちの一枚の期待値がこれと同じ値になる，というのは，直観的には正しそうに思えても，それ以上の保証がないようにも思えるからです．この点を厳密な議論でうめることも，もちろんできますが，ここでは，新しくアプローチしなおして，演習問題 3.1 の一般化となっている，次の命題を直接証明することを試みてみることにします:

定理 3.3 ある対象 $o_0,...,o_{n-1}$ を考える．各 o_k $(0\le k<n)$ には値 a_k が付されているとする．$o_0,...,o_{n-1}$ から m 個（ただし $m\le n$）を無作為に取り出すとき，それらの値の和の期待値は，$\dfrac{m}{n}\cdot\sum_{k=0}^{n-1}a_k$ である．

演習問題 3.1 は，上の定理で，たとえば，$n=7+5+8=20$, $m=3$, $o_0,...,o_6$ は 10 円硬貨，$o_7,...,o_{11}$ は 50 円硬貨，$o_{12},...,o_{19}$ は 100 円硬貨として，$a_0=\cdots a_6=10$, $a_7=\cdots=a_{11}=50$, $a_{12}=\cdots a_{19}=100$ としたときの定理による期待値:

$$\frac{3}{20}\sum_{k=0}^{19}a_k=\frac{3}{20}\Big(\underbrace{10+\cdots+10}_{7\text{個}}+\underbrace{50+\cdots+50}_{5\text{個}}+\underbrace{100+\cdots+100}_{8\text{個}}\Big)$$

と一致します．

この定理の証明は次のようにして行うことができます．まず，

$$I=\{Y : Y\subseteq\{0,1,...,n-1\},\ |Y|=m\}$$

とします．ただし $|Y|$ で集合 Y の要素の個数をあらわします．つまり，I は $\{0,1,...,n-1\}$ から m 個の要素を集めてできる集合のすべてを集めてできる集合となっています．$|I|={}_nC_m$ に注意します．このとき，$o_0,...,o_{n-1}$ から m 個選ぶ選び方は

$$\{o_k : k\in Y\}, Y\in I$$

と列挙でき，これらはすべて同じ確率で選ばれるとすると，それらの一つが選ばれる確率は $\dfrac{1}{{}_nC_m}$ となることがわかります．$\{o_k : k \in Y\}$ が選ばれたときの値の和は，$\displaystyle\sum_{k\in Y} a_k$ とあらわせるので，問題となっている期待値は，

$$(3.2)\qquad \sum_{Y\in I}\left(\frac{1}{{}_nC_m}\left(\sum_{k\in Y} a_k\right)\right) = \frac{1}{{}_nC_m}\left(\sum_{Y\in I}\sum_{k\in Y} a_k\right) \qquad \text{p-1}$$

となることがわかります．この式の右辺をよく見ると，$\displaystyle\sum_{Y\in I}\sum_{k\in Y} a_k$ のところは，各 a_k をそれぞれ ${}_{n-1}C_{m-1}$ 個足しあわせたの和になっていることがわかります．つまり

$$\begin{aligned}(3.2) &= \frac{1}{{}_nC_m}\cdot {}_{n-1}C_{m-1}\cdot\sum_{k=0}^{n-1} a_k = \frac{m!(n-m)!}{n!}\cdot\frac{(n-1)!}{(n-m)!(m-1)!}\cdot\sum_{k=0}^{n-1} a_k\\ &= \frac{m}{n}\cdot\sum_{k=0}^{n-1} a_k\end{aligned}$$

となることがわかり，定理が証明されました． ❑ (定理 3.3)

上の定理からわかることの一つは，o_0,...,o_{n-1} から m 個順にとり出していったときのそれぞれの値の和の期待値も，m 回とりだして戻すことを繰り返したときのそれぞれの値の和の期待値も同じになる，ということですが，直観的には，これはなんとなく不思議な気がします[36]．

3.2 ポアソン分布

poison

ある $m > 0$ に対し，確率変数 X が，

$$(3.3)\qquad P(X=k) = \frac{m^k}{k!}e^{-m},\ \ k = 0,1,2,\ldots \qquad \text{poisson-0}$$

を満たすとき，X はポアソン分布 (Poisson[37] distribution) $Po(m)$ に従う，という．ポアソン分布 $Po(m)$ に従う確率変数は，

(3.4)　単位時間あたりの平均発生回数が m の事象で， poisson-1

(∗) それぞれの時刻での，この事象の発生（または非発生）が他の時刻での，この事象の発生（または非発生）と独立である

ようなものについて，その事象が（ある測定での）単位時間内に起る回数を返す

---

[36] ここで書いたような考察は,「うるさい」と感じる人も多いのではないかと思います．確かに，数学的能力の十分にある人は，最初に解答としてあげた説明を読めば，自動的に，その後に書いたことに対応する数学的内容を頭の中で補間して，次に進むことができるはずです．また，それのできない人はいずれにしても数学は分らないのだから説明しても無駄だ，という議論も成立するのかもしれません．

[37] Siméon-Denis Poisson (1781–1840) (発音は “プワソン” の方が原音に近い) フランスの物理学者，数学者．音響学，剛体の弾性，熱伝導，電気伝導などに関する仕事がある．

ものになっていると考えられる．このことを念頭に置くと，ポアソン分布の次のようなタイプの応用が考えられる:

例 3 ある国で 1000 年の間に歴史上の大地震が 73 回起っている．大地震の起る回数がポアソン分布に従うと仮定して計算すると，この国での一年間の大地震の平均回数は，$\dfrac{73}{1000} = 0.073$ と考えられるから，この国である年に大地震が起る回数を与える確率変数を X として，2005 年にここで大地震が起る確率は，

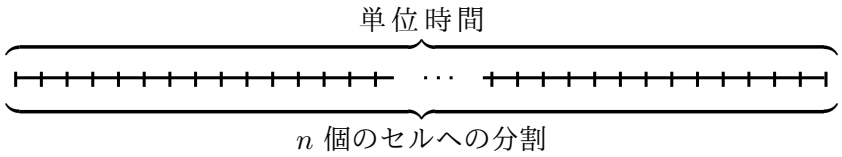
$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{0.073^0 \cdot e^{-0.073}}{0!} = 0.07039\cdots$$

となる．

上の応用例では，地震の発生の状況が時間変化しないことと，発生が (3.4) での $(*)$ を満たしていると仮定してよいかどうかが，ポアソン分布をこの問題に応用するのが妥当かどうかを評価する際のポイントになるであろう．地震の研究がご専門の中部大学理学教室の工藤健先生によると，地震の発生回数の分布がポアソン分布によく合致する地域もある，ということである．

以下で，(3.4) の分布が，なぜ (3.3) の式で与えられると考えられるかを考察する．

今，単位時間を n 個の微小時間のセル (細胞) に等分することを考える:



n は十分に大きくとってあり，一つ一つの微小時間のセル内で，ここで考えている事象が 2 回以上起る確率は無視できるものとする．このとき，各セル内で，この事象が発生する確率は，$p = \dfrac{m}{n}$ と考えられる．(3.4), $(*)$ により，それぞれのセル内で事象が起るかどうかは，他のセル内で事象が起るかどうかとは独立と考えられるから，事象が起った微小時間のセルの個数 $\approx$ 単位時間内に事象の起った回数は二項分布 $Bin(n, p)$ に従うと考えられる．したがって，このような事象がちょうど k 個のセルで起きる確率は ${}_nC_k \cdot p^k(1-p)^{n-k}$ で与えられる．この値が，(3.4) のような分布を持つ確率変数 X の $P(X = k)$ の値の近似となっていると考えられるが，この近似は n を大きくするほど精度が上るので，

$$P(X = k) = \lim_{n\to\infty} \left({}_nC_k \cdot p^k(1-p)^{n-k}\right)$$

とすればよいことがわかる．ところが，この右辺は $\dfrac{m^k}{k!}e^{-m}$ に一致する:

命題 3.4 $m > 0$ として，$k \in \mathbb{N}$ とする．このとき，

$$\lim_{n\to\infty} \left({}_nC_k \cdot \left(\frac{m}{n}\right)^k \cdot \left(1 - \frac{m}{n}\right)^{n-k}\right) = \frac{m^k}{k!}e^{-m}$$

が成り立つ．

証明.

$$ {}_nC_k = \frac{n!}{k!\,(n-k)!} $$

だった．このことと，数列の積の極限は各々の数列の極限の積と一致することから，

$$ \begin{aligned} &\lim_{n\to\infty}\left({}_nC_k\cdot\left(\frac{m}{n}\right)^k\cdot\left(1-\frac{m}{n}\right)^{n-k}\right)\\ &=\lim_{n\to\infty}\left(\frac{n!}{k!\,(n-k)!}\cdot\left(\frac{m}{n}\right)^k\cdot\left(1-\frac{m}{n}\right)^{n-k}\right)\\ &=\lim_{n\to\infty}\left(\frac{m^k}{k!}\cdot\frac{n!}{(n-k)!\,n^k}\cdot\left(1-\frac{m}{n}\right)^{n}\cdot\left(1-\frac{m}{n}\right)^{-k}\right)\\ &=\frac{m^k}{k!}\cdot\underbrace{\lim_{n\to\infty}\frac{n!}{(n-k)!\,n^k}}_{=1}\cdot\underbrace{\lim_{n\to\infty}\left(1-\frac{m}{n}\right)^{n}}_{=e^{-m}}\cdot\underbrace{\lim_{n\to\infty}\left(1-\frac{m}{n}\right)^{-k}}_{=1}\\ &=\frac{m^k}{k!}\cdot e^{-m} \end{aligned} $$

となる．上式の最後の等号のところでは，以下のような部分計算 (1), (2), (3) が用いられている:

(1):

$$ \begin{aligned} &\lim_{n\to\infty}\frac{n!}{(n-k)!\,n^k}\\ &=\lim_{n\to\infty}\frac{n\cdot(n-1)\cdot\,,\ldots\,\cdot(n-k+1)\cdot(n-k)\cdot(n-k-1)\cdot\,,\ldots\,\cdot2\cdot1}{\underbrace{n\cdot\,,\ldots\,\cdot n}_{k\ \text{コ}}\cdot(n-k)\cdot(n-k-1)\cdot\,,\ldots\,\cdot2\cdot1}\\ &=\lim_{n\to\infty}\frac{n\cdot(n-1)\cdot\,,\ldots\,\cdot(n-k+1)}{\underbrace{n\cdot\,,\ldots\,\cdot n}_{k\ \text{コ}}}=1 \end{aligned} $$

(2): e が，

$$ e=\lim_{h\to0}(1+h)^{\frac{1}{h}} $$

と表わせることと，任意の α に対し，関数 $x\mapsto x^\alpha$ が連続であること（ここでは $\alpha=-m$ を考える）を用いる．

今，$h=-\dfrac{m}{n}$ とおくと，$n\to\infty$ なら $h\to0$ だから，

$$ \lim_{n\to\infty}\left(1-\frac{m}{n}\right)^n=\lim_{h\to0}(1+h)^{-\frac{m}{h}}=\lim_{h\to0}\left((1+h)^{\frac{1}{h}}\right)^{-m}=\left(\lim_{h\to0}(1+h)^{\frac{1}{h}}\right)^{-m}=e^{-m} $$

である．

(3): (2) と同様に，関数 $x\mapsto x^\alpha$ の連続性から，

$$ \lim_{n\to\infty}\left(1-\frac{m}{n}\right)^{-k}=\left(\lim_{n\to\infty}\left(1-\frac{m}{n}\right)\right)^{-k}=1^{-k}=1 $$

である． ❑ (命題 3.4)

3.3 Chebyshev の定理と大数の法則

Chebyshev

X を確率変数として，$\varphi(x)$ を X の確率密度関数とする．つまり，可測集合 $D \subseteq \mathbb{R}$ に対し，X の返す値が D に入っている確率 $P(X \in D)$ は，

$$P(X \in D) = \int_D \varphi(x)dx \tag{3.5}$$

random-2

となる．

$E(X)$ で X の期待値を，$V(X)$ で X の分散をあらわす．

$$E(X) = \int_{-\infty}^{\infty} x\varphi(x)dx \tag{3.6}$$

random-0

$$V(X) = \int_{-\infty}^{\infty} \left(x^2 - E(X)\right)^2 \varphi(x)dx \tag{3.7}$$

random-1

である．

$f(x)$ を（$\mathbb{R}$ から $\mathbb{R}$ への）可測関数するとき，$f(X)$ で X と共時的な確率変数で，試行したとき X の返した値が x のとき $f(x)$ を返すようなものをあらわすことにする．

これは，理論的な枠組の中では，$f(X)$ を，すべての可測な $T \subseteq \mathbb{R}$ に対して

$$P(Y \in T) = P(X \in f^{-1}{}''T) \tag{3.8}$$

random-3

となるような（確率密度関数を持つ），確率変数 Y とする，ということである．このとき，

$$E(f(X)) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx \tag{3.9}$$

random-4

となる．

Che-0

補題 3.5 X を確率変数とするとき，

$$V(X) = E(X^2) - (E(X))^2$$

が成り立つ．

証明. $\varphi(x)$ を X の確率密度関数とすると，

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 \varphi(x)dx \\
&= \int_{-\infty}^{\infty} \left(x^2 - 2E(X)x + (E(X))^2\right) \varphi(x)dx \\
&= \int_{-\infty}^{\infty} x^2\varphi(x)dx - 2E(X) \underbrace{\int_{-\infty}^{\infty} x\varphi(x)dx}_{=E(X)} + (E(x))^2 \underbrace{\int_{-\infty}^{\infty} \varphi(x)dx}_{=1} \\
&= E(X^2) - (E(X))^2
\end{aligned}$$

□ (補題 3.5)

定理 3.6 (P. L. Chebyshev) X を確率変数とするとき，任意の $c>0$ に対し， cheby-0

$$P(|X-E(X)|\geq c)\leq V(X)/c^2$$

が成り立つ．余事象に翻訳すると，

$$P(|X-E(X)|\leq c)\geq 1-V(X)/c^2$$

である．

証明. $\varphi(x)$ を X の確率密度関数とする．このとき，

$$\begin{aligned}(3.10)\quad V(X)&=\int_{-\infty}^{\infty}(x-E(X))^2\varphi(x)dx \qquad \text{chebyshev-0}\\ &=\int_{E(X)-c}^{E(X)+c}(x-E(X))^2\varphi(x)dx\\ &\quad+\int_{-\infty}^{E(X)-c}(x-E(X))^2\varphi(x)dx+\int_{E(X)+c}^{\infty}(x-E(X))^2\varphi(x)dx\\ &\geq\int_{-\infty}^{E(X)-c}(x-E(X))^2\varphi(x)dx+\int_{E(X)+c}^{\infty}(x-E(X))^2\varphi(x)dx\\ &\geq\int_{-\infty}^{E(X)-c}c^2\varphi(x)dx+\int_{E(X)+c}^{\infty}c^2\varphi(x)dx\\ &=c^2P(|X-E(X)|\geq c)\end{aligned}$$

となるから，この両辺を c^2 で割ると求める不等式が得られる． ❑ (定理 3.6)

random-5

補題 3.7 X_1 と X_2 を共時的な確率変数として，$c_1, c_2\in\mathbb{R}$ を定数とする．このとき
(1) $E(c_1X_1+c_2X_2)=c_1E(X_1)+c_2E(X_2)$ が成り立つ．
(2) X_1 と X_2 が独立なら，$V(c_1X_1+c_2X_2)=(c_1)^2V(X_1)+(c_2)^2V(X_2)$ が成り立つ．

証明. $\psi(x_1,x_2)$ を X_1 と X_2 の同時確率密度関数として，$\varphi_1(x_1)$, $\varphi_2(x_2)$ をそれぞれ X_1 と X_2 の周辺確率密度関数とする．

$$\varphi_1(x_1)=\int_{-\infty}^{\infty}\psi(x_1,x_2)dx_2,\qquad \varphi_2(x_2)=\int_{-\infty}^{\infty}\psi(x_1,x_2)dx_1$$

である．このとき，

(1)

$$\begin{aligned}E(c_1X_1+c_2X_2)&=\iint_{\mathbb{R}^2}(c_1x_1+c_2x_2)\psi(x_1,x_2)dx_1dx_2\\ &=c_1\iint_{\mathbb{R}^2}x_1\psi(x_1,x_2)dx_1dx_2+c_2\iint_{\mathbb{R}^2}x_2\psi(x_1,x_2)dx_1dx_2\\ &=c_1\int_{-\infty}^{\infty}x_1\varphi_1(x_1)dx_1+c_2\int_{-\infty}^{\infty}x_2\varphi_2(x_2)x_2\\ &=c_1E(X_1)+c_2E(X_2).\end{aligned}$$

(2) 仮定から $\psi(x_1,x_2)=\varphi_1(x_1)\varphi_2(x_2)$ である．したがって，(1) を使うと，

$$\begin{aligned}
V(c_1X_1+c_2X_2) &= \iint_{\mathbb{R}^2} \left(c_1x_1+c_2x_2-E(c_1X_1+c_2X_2)\right)^2 \psi(x_1,x_2)dx_1dx_2 \\
&= \iint_{\mathbb{R}^2} \left(c_1x_1+c_2x_2-(c_1E(X_1)+c_2E(X_2))\right)\right)^2 \varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&= \iint_{\mathbb{R}^2} (c_1x_1)^2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 + \iint_{\mathbb{R}^2} (c_2x_2)^2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad + \iint_{\mathbb{R}^2} (c_1E(X_1))^2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 + \iint_{\mathbb{R}^2} (c_2E(X_2))^2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad + \iint_{\mathbb{R}^2} 2c_1c_2x_1x_2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad + \iint_{\mathbb{R}^2} 2c_1c_2E(X_1)E(X_2)\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad - \iint_{\mathbb{R}^2} 2c_1c_2x_1E(X_2)\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad - \iint_{\mathbb{R}^2} 2c_1c_2E(X_1)x_2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad - \iint_{\mathbb{R}^2} 2(c_1)^2E(X_1)x_1\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&\quad - \iint_{\mathbb{R}^2} 2(c_2)^2E(X_2)x_2\varphi_1(x_1)\varphi(x_2)dx_1dx_2 \\
&= (c_1)^2E((X_1)^2)+(c_2)^2E((X_2)^2) \\
&\quad +(c_1)^2(E(X_1))^2+(c_2)^2(E(X_2))^2 \\
&\quad +2c_1c_2E(X_1)E(X_2) \\
&\quad +2c_1c_2E(X_1)E(X_2) \\
&\quad -2c_1c_2E(X_1)E(X_2) \\
&\quad -2c_1c_2E(X_1)E(X_2) \\
&\quad -2(c_1)^2(E(X_1))^2 \\
&\quad -2(c_2)^2(E(X_2))^2 \\
&= \left((c_1)^2E((X_1)^2)-(c_1)^2(E(X_1))^2\right) + \left((c_2)^2E((X_2)^2)-(c_2)^2(E(X_2))^2\right) \\
&= (c_1)^2V(X_1)+(c_2)^2V(X_2).
\end{aligned}$$

❑ (補題 3.7)

定理 3.8 (大数の法則) X_1, X_2, $X_3\ldots$ を，独立な，同じ確率分布を持つ確率変数で $\mu =$ cheby-1 $E(X_1)=E(X_2)=E(X_3)=\cdots$, $\sigma^2=V(X_1)=V(X_2)=V(X_3)=\cdots$ とする．

$$\bar{X}_{(n)}=\frac{1}{n}\sum_{k=1}^{n}X_k$$

とするとき，任意の $\varepsilon>0$ に対し，

$$\lim_{n\to\infty} P(|\bar{X}_{(n)}-\mu|\le\varepsilon)=1$$

が成り立つ．

証明. 補題 3.7, (1) により，$E(\bar{X}_{(n)}) = \mu$ である．一方，補題 3.7, (2) により，$V(\bar{X}_{(n)}) = \dfrac{1}{n}\sigma^2$ である．

したがって Chebyshev の定理（定理 3.6）により，

$$P(|\bar{X}_{(n)} - \mu| \leq \varepsilon) = P(|\bar{X}_{(n)} - E(\bar{X}_{(n)})| \leq \varepsilon) \geq 1 - V(\bar{X}_{(n)})/\varepsilon^2 = 1 - \frac{1}{n}\sigma^2/\varepsilon^2 \xrightarrow{n\to\infty} 1$$

である． ❑ (定理 3.8)

3.4 正規分布と Kurtosis

kurtosis

正規分布の密度関数 (probability density function, pdf) の正規化のための積分計算:

$$\begin{aligned}\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{a}}dx\right)^2 &= \lim_{c\to\infty}\int_{-c}^{c}\int_{-c}^{c} e^{-\frac{x^2+y^2}{a}}dx\,dy\\ &= \lim_{s\to\infty}\int_0^{2\pi}\int_0^{s} e^{-\frac{r^2}{a}}r\,dr\,d\theta\\ &= \lim_{s\to\infty}\int_0^{2\pi}\left[-\frac{a}{2}e^{-\frac{r^2}{a}}\right]_{r=0}^{s} d\theta\\ &= \pi a\end{aligned}$$

したがって,

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{a}}dx = \sqrt{\pi a}$$

である.

x を密度関数 $\dfrac{1}{\sqrt{\pi a}}\,e^{-\frac{x^2}{a}}$ を持つ確率変数とする. このとき, 密度関数の対称性から $E(x)=0$ だから,

$$\begin{aligned}Var(x) &= E((x-E(x))^2) = E(x^2)\\ &= \frac{1}{\sqrt{\pi a}}\int_{-\infty}^{\infty} x^2\cdot e^{-\frac{x^2}{a}}\,dx\\ &= \frac{1}{\sqrt{\pi a}}\int_{-\infty}^{\infty} x\cdot\left(-\frac{a}{2}\cdot e^{-\frac{x^2}{a}}\right)'dx\\ &= \frac{1}{\sqrt{\pi a}}\left(\left[-\frac{a}{2}\cdot xe^{-\frac{x^2}{a}}\right]_{-\infty}^{\infty} + \frac{a}{2}\int_{-\infty}^{\infty} e^{-\frac{x^2}{a}}\,dx\right) \qquad \text{(部分積分の定理による)}\\ &= \frac{1}{\sqrt{\pi a}}\left(0+\frac{a}{2}\sqrt{\pi a}\right) = \frac{a}{2}\end{aligned}$$

平均値が 0 分散が 1 となるような正規分布を標準正規分布とよぶが, 標準正規分布は, 上の計算から, a に 2 を代入した

$$\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}$$

を確率分布関数として持つことがわかる.

x を標準正規分布に従う確率変数とするとき,

$$\begin{aligned}E(x^4) &= \frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty} x^4e^{-\frac{x^2}{2}}dx\\ &= \frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty} -x^3\left(e^{-\frac{x^2}{2}}\right)'dx\\ &= \frac{1}{\sqrt{2\pi}}\left(\left[-x^3e^{-\frac{x^2}{2}}\right]_{\infty}^{\infty} - \int_{\infty}^{\infty} -3x^2e^{-\frac{x^2}{2}}\,dx\right)\\ &= 3\cdot\left(\frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty} x^2e^{-\frac{x^2}{2}}\,dx\right)\\ &= 3\end{aligned}$$

となる．より一般には，x を正規分布に従う確率変数とするとき，$E(x^4) = 3(E(x^2))^2$ となることがわかる．ここで，確率変数 y に対しその kurtosis を

$$kurt(y) = E(y^4) - 3(E(y^2))^2$$

と定義すると，これは y が正規分布のときに $kurt(y) = 0$ となるような y の不変量となる．

4 初等数論

number-theory

4.1 n^9

n29

以下の文章は，今のところ，ある程度の数学の素養を読者に仮定したものとなっているが，後で拡張してもっと初心者向けのテキストにするつもりである:

先日，職員食堂でコンピュータ・リテラシーの講義を受けもっている先生から

(4.1) どんな数でも，9 乗した結果の 1 の桁と，もとの数の 1 の桁は同じになる n-to-9-0

というのが昔から気になっていたのだが，これは何でなのか，という質問を受けた．もちろん，(4.1) の証明には，0 から 9 までの数について，このことが成り立つことを計算で示せば十分である．試しにやってみると:

```
1^9 = 1
2^9 = 512
3^9 = 19683
4^9 = 262144
5^9 = 1953125
6^9 = 10077696
7^9 = 40353607
8^9 = 134217728
9^9 = 387420489
```

となるから，確かによい．しかし，この計算は何でそうなるのかということを何も説明していないように思える．

そこで，この場合は，10 進法の数表示だったが，一般の n 進法ではどうかを考えてみることにする．つまり，任意の n について，

(4.2) どんな数 $k \in \mathbb{N}$ でも k の n 進法表示の 1 桁目の数字と k^{n-1} の n 進法表示の 1 桁目の数字が同じになる n-to-9-1

という主張を考えてみる．この主張が，すべての n に対し成り立つわけでないことは，n が素数の場合について考えてみれば分かる．このときには，フェルマーの小定理が答を与えてくれる．$\equiv_n$ で n を法とする整数の同値関係を表すことにする．つまり $k, \ell \in \mathbb{Z}$ のとき，$k \equiv_n \ell \ \Leftrightarrow \ k - \ell \in n\mathbb{Z}$ である．

定理 4.1 (フェルマーの小定理からの帰結) p を素数とするとき，任意の $k \in \mathbb{N} \setminus \{0\}$ に対し，$k^{p-1} \equiv_p 1$ である．特に，任意の $k \in \mathbb{N}$ に対し $k^p \equiv_p k$ となる． ❑ fermat

したがって p が素数の場合には任意の数の p 進数表示の 1 桁目は，その数の（$p-1$ 乗ではなく）p 乗の p 進数表示の第 1 桁と等しくなる．

定理 4.1 を使うと，次の補題が示せる:

2p

補題 4.2 $n = 2p$ で p は素数とする．このとき，任意の $k \in \mathbb{Z}$ に対し $k^{n-1} \equiv_n k$ となる．

証明. $k=0$ のときは，等式は明らかである．$k\in\mathbb{Z}\setminus\{0\}$ とするとき，$k^{n-1}-k=k^{2p-1}-k=k(k^{2(p-1)}-1)=k\left(\left(k^{p-1}\right)^2-1\right)=k(k^{p-1}-1)(k^{p-1}+1)$ となるが，定理 4.1 により，$k^{p-1}-1$ は p の倍数となり，k か $k^{p-1}+1$ の，少なくとも 1 つは 2 の倍数である．したがって，$k^{n-1}-k$ は n の倍数であることがわかる．つまり $k^{n-1}-k\equiv_n 0$ である．

❑ (補題 4.2)

系 4.3 n が素数の 2 倍のとき，任意の正の数 k の n 進数表示の第 1 桁目の数字は，k^{n-1} の n 進数表示の第 1 桁目の数字と等しくなる．❑

$10=5\times 2$ だから，10 進法表記は系 4.3 の適用範囲に入っていることに注意する．したがって，補題 4.2 が (4.1) の背後にある数学的事実を述べたものとなっている，と考えることができる．

次の補題も同様に証明できる:

補題 4.4 $n=2^m$ $(m\in\mathbb{N})$ とするとき，任意の $k\in\mathbb{Z}$ に対し，

$$k^{n-1}\equiv_n k \tag{4.3}$$ n-to-9-2

となる．

証明. m に関する帰納法で示す．$m=1$ のときは明らかである．$n=2^m$ に対し (4.3) が成り立つことが示せたとして，$n=2^{m+1}$ に対しても，(4.3) が成り立つことを示す．$n=2^{m+1}$ とする．$k=0$ のときは (4.3) が成り立つことは明らかだから，$k\neq 0$ とする．

$$\begin{aligned}k^{n-1}-k&=k^{2^{m+1}-1}-k=k^{2(2^m-1)+1}-k=k(k^{2(2^m-1)}-1)\\&=k\left(\left(k^{2^m-1}\right)^2-1\right)=k(k^{2^m-1}-1)(k^{2^m-1}+1)\end{aligned}$$

となるが，$k^{2^m-1}-1$ は帰納法の仮定から 2^m の倍数で，k か $k^{2^m-1}+1$ の，少なくとも 1 つは 2 の倍数である．したがって，$k^{n-1}-k$ は $n=2^{m+1}$ の倍数であることがわかる．つまり $k^{n-1}\equiv_n k$ である．❑ (補題 4.4)

系 4.5 n が 2 の冪乗のとき，任意の正の数 k の n 進数表示の第 1 桁目の数字は，k^{n-1} の n 進数表示の第 1 桁目の数字と等しくなる．❑

これらの結果から，2 進数，4 進数, 8 進数，10 進数，16 進数といった我々が通常に用いる数表示では，常に (4.1) と同様の性質が成り立つことが結論される．

以上を書いた後で，補題 4.2 はさらに以下のように一般化できることを [1] で知った．

補題 4.6 n を平方因子を持たないような正の自然数とする．つまり，n は異なる素数の（1 乗の）積の形に表せるようなものとする．このとき，正の自然数 t が $t\equiv_{\varphi(n)}1$ を満たせば．すべての整数 a に対し，$a^t\equiv_n a$ が成り立つ．ただし，$\varphi(n)$ でオイラーの関数を表す．❑

$\varphi(10) = 4$ だから，$5 \equiv_{\varphi(10)} 1$ である．したがって，補題 4.6 により，十進数表示については，数の第 1 桁の数字と，その数の 5 乗の第 1 桁の数字も常に等しくなることがわかる．

上の議論は，この節の初めに述べた，計算による (4.1) の証明に比べて，(4.1) の，より本質的な説明を与えているとは言えないだろうか.「一般論は分りにくい」というのは一般的な偏見のような気がするが，上の議論は，一般論を行うことによって，本質がより深く見えてくることの 1 つの好例になっていると思う．

References

[1] 楫 元，公開鍵暗号を解読せよ! — 君もスパイになれる? —, 数学通信，10(2), (2005), 5–34.

4.2 a^b

a2b

補題 4.7 無理数 a, b で a^b が有理数になるようなものが存在する．

証明. $\sqrt{2}^{\sqrt{2}}$ が有理数なら，$a=\sqrt{2}, b=\sqrt{2}$ とすればよい．そうでないなら，$a=\sqrt{2}^{\sqrt{2}}$, $b=\sqrt{2}$ とすれば，

$$a^b=\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}}=(\sqrt{2})^{(\sqrt{2}\cdot\sqrt{2})}=(\sqrt{2})^2=2$$

となる[38)]． ❑ (補題 4.7)

$$\frac{1}{n}=\frac{1}{n+1}+\frac{1}{n\,(n+1)}$$

と

$$\frac{m}{n}=\underbrace{\frac{1}{n}+\cdots+\frac{1}{n}}_{m\text{ times}}$$

を組み合せると，すべての有理数は，それぞれ異なる分母を持つ単位分数 (unit fraction — 分子が 1 の分数) の和に書けることがわかる．しかも，1 つの数に対し，このような表現は無限に存在する（これは Fibonacci (Leonardo, (1180?-1250?)) によるものらしい [Paul Hoffman, The man who loved only numbers] からの孫引き）．

以下の議論も Fibonacci による: $\dfrac{m}{n}$ が与えられたとき，$m_0=m$, $n_0=n$ として，$m_k\neq 0$ のときには，

$$\ell_k=\min\{\ell\,:\,\frac{m_k}{n_k}\geq\frac{1}{\ell}\}$$

として，

$$n_{k+1}=n_k\ell_k,$$
$$m_{k+1}=m_k\ell_k-n_k$$

とする．

$\dfrac{m_{k+1}}{n_{k+1}}=\dfrac{m_k}{n_k}-\dfrac{1}{\ell_k}$ である．

演習問題 4.8 このプロセスは，有限回のステップの後，ある k^* で $m_{k^*}=0$ となって停止する．

[ヒント ℓ_k の選び方から，$\dfrac{1}{\ell_k}\leq\dfrac{m_k}{n_k}<\dfrac{1}{\ell_k-1}$ である．$\dfrac{m_k}{n_k}-\dfrac{1}{\ell_k}=\dfrac{m_k\ell_k-n_k}{n_k\ell_k}$ とすると，$\dfrac{1}{\ell_k}\leq\dfrac{m_k}{n_k}$ から，$0<m_k\ell_k-n_k$, また $\dfrac{m_k}{n_k}<\dfrac{1}{\ell_k-1}$ から，$m_k(\ell_k-1)<n_k$ したがって $m_k\ell_k-n_k<m_k$ である．]

このとき，

$$\frac{m}{n}=\frac{1}{\ell_0}+\cdots+\frac{1}{\ell_{k^*-1}}$$

となる．

[38)] 実は $\sqrt{2}^{\sqrt{2}}$ が無理数となることは知られているらしい．しかし，この証明の面白いところは，この事実を知らなくても証明ができてしまうところであろう．

4.3 Bertrand's Postulate

bertrand
bertrand-T

Theorem 4.9 (Bertrand's Postulate, Chebyshev's Theorem) *For any natural number $n > 1$ there is always at least one prime number p such that $n < p < 2n$.*

Corollary 4.10 *The sequence of prime numbers (starting with 1) is complete. That is, any natural number can be represented as the sum of pairwise distinct prime numbers (anc 1).*

Proof. Let p_n denote nth prime number where we set $p_0 = 1$. We prove

$(*)_n$ for any $m < p_n$, m can be represented as a sum of finite elements of p_k, $k < n$ where each p_k appears at most once in the sum

holds for all $n \in \omega$ by induction on n. For $n = 0$ and $n - 1$ this is easy to check.

Assume that $(*)_n$ holds and we show that $(*)_{n+1}$ holds. By Theorem 4.9 we have $p_n < p_{n+1} < 2p_n$. For any $\ell \leq p_{n+1}$, if $\ell \leq p_n$ then ℓ can be represented as a sum of some of p_k, $k < n$. If $\ell \geq p_n$ then $\ell - p_n < p_n$ and hence, by the induction hypothesis, $\ell - p_n$ can be represented as a sum S of some of p_k, $k < n$ Thus $\ell = S + p_n$. ❑ (Corollary 4.10)

5 ブール代数

boolean-algebras

補題 5.1 X と Y を Boolean spaces として，$f : X \to Y$ を continuous な surjection とする．このとき，$\tilde{f} : Clopen(Y) \to Clopen(Z); U \mapsto f^{-1}(U)$ は embedding となるが，次の同値が成り立つ：

f は open mapping $\Leftrightarrow$ $\tilde{f}$ は relatively complete embedding.

（この同値は Koppelberg の “Projective Algebras” では sheaf representation を介して証明してあるが以下で直接証明を与える．）

証明． $\Rightarrow$: $U \subseteq X$ を clopen とする．f は closed mapping だから，$f''U$ は Y の clopen subset である．つまり，$f''U \in Clopen(Y)$ $U \subseteq f^{-1}(f''U)$（つまり $Clopen(X) \models U \leq \tilde{f}(f''U)$）だが，$\tilde{f}(f''U)$ は U の $\tilde{f}''Clopen(Y)$ への upper projection になっている：$V \in Clopen(Y)$ で $U \subseteq f^{-1}(V)$ なら，$f''U \subseteq f''f^{-1}(V) = V$ だから，$f^{-1}(f''U) \subseteq f^{-1}(V)$ となるからである．

$\Rightarrow$: f が open mapping でないなら，clopen な $O \subseteq X$ で $f''O \notin \mathcal{O}_Y$ となるようなものがとれる．f は closed mapping だから，$f''O$ は closed であることに注意する．O は $\tilde{f}''Clopen(Y)$ への upper projection を持たないことを示す：$U \in Clopen(Y)$ を $f^{-1}(U) \supseteq O$ となるものとする．このとき，$U \supseteq f''O$ だが，$f''O$ は open ではないので，$U \setminus f''O$ は空でない open set となる．したがって（X は Boolean だから）空でない clopen set $V \subseteq U \setminus f''O$ がとれる．このとき，$U' = U \setminus V$ とすると，$U' \subsetneqq U$ で $U' \supseteq f''O$ となる．したがって $\tilde{f}(U)$ は O の upper projection ではない． ❑ (補題 5.1)

6 初等幾何?

geometry
geometry-0

補題 6.1 平面は $< 2^{\aleph_0}$ 個の直線では覆えない．

証明． L を濃度が $< 2^{\aleph_0}$ の平面上の直線の集合とする．各 $\ell \in L$ に対し，$w(\ell)$ を ℓ の傾き (角度) とすると，$r \in [0, \pi) \setminus \{w(\ell) : \ell \in L\}$ がとれる．ℓ^* を，傾き r を持つ任意の直線とすると，ℓ^* は各 $\ell \in L$ とちょうど 1 点でしか交じわらないから，ℓ^* 上に L のどの直線とも交じわらない点が存在する． ❑ (補題 6.1)

命題 6.2 (S. Mazurkiewicz 1914, see [1]) どの直線ともちょうど 2 点で交じわるような平面上の集合が存在する．

証明． 平面上の直線の全体を $\langle \ell_\alpha : \alpha < 2^{\omega_0} \rangle$ と enumerate する．点の集合の continuous な increasig sequence P_α, $\alpha < 2^{\omega_0}$ を以下が成り立つように構成する．

(6.1) P_α の任意の 3 点は同一直線上にない (更に $P_{\alpha+1} \setminus P_\alpha$ の点は P_α のどの 2 点を結んだ直線上にもない); geom-0

(6.2) 任意の $\beta < \alpha$ に対し，$|\ell_\beta \cap P_\alpha| = 2$. geom-1

この構成が可能なことは 補題 6.1 からわかる: $\alpha+1$ 番目のステップでは，L_α を P_α の 2 点を結んでできる直線の全体とする．ℓ_α が L_α に含まれているなら， $P_{\alpha+1}=P_\alpha$ とする．そうでなければ，(6.1) により，$k=2-|\,\ell_\alpha\cap P_\alpha\,|>0$ だから，k (= 1 or 2) 個の点を $\ell_\alpha\setminus(\bigcup L_\alpha)$ からとって，それらを P_α に加えたものを $P_{\alpha+1}$ とすればよい．

$P=\bigcup_{\alpha<2^{\omega 0}} P_\alpha$ とすれば，これが求めるようなものとなる． ❑ (命題 6.2)

References

[1] Ben Chad, Robin Knight and Rolf Suabedissen, Set-theoretic constructions of two-point sets, in Fundamenta Mathematicae, 203 (2009), 179-189.

7 グラフ理論

graph-th

定理 7.1 (平面上) 多角形の頂点を結ぶことで得られる平面グラフの頂点は，辺で結ばれた 2 点が異る色を割り当てられるように 3 色で塗り分けることができる．

証明. ある多角形 P の対角線による分割から得られたグラフを G とする．必要なら対角線を足して P は G で三角形に分割されているとしてよい，このとき，

Claim 7.1.1 G の頂点で対角線で結ばれていないようなものが少なくとも 2 つは存在する．

⊢ P の頂点の数 n に関する帰納法で証明する．$n=3$ のときには主張は自明に成り立つ．

すべての $n<k$ に対し主張が成り立つとして $n=k$ に対しても主張が成り立つことを示す．P を k 個の頂点を持つ多角形として，G を P の頂点を結んで得られた三角形分割のグラフとする．このとき，G の対角線の一つを d とすると，d の両側のグラフを考えることで，G は d を一辺とする 2 つの多角形の分割のグラフ G', G'' に分割される．帰納法の仮定から，G', G'' はともに少なくとも二つの対角線で結ばれていないような頂点を持つ．たとえば G' でのそれらの 2 つの頂点が d の両端だとするると G' は三角形でなくてはならない．したがって，この場合にも，G' 頂点で d の両端以外の点で対角線で結ばれていない点が存在する． G'' についても同様である．したがって，G も対角線で結ばれていない頂点が少なくとも 2 つは存在することがわかる． ⊣ (Claim 7.1.1)

定理を証明する．P の頂点の数 n に関する帰納法で証明する．$n=3$ のときには自明である．

今，すべての $n<k$ に対して定理が成り立つとして $n=k$ のときにも定理が成り立つことを示す．P を k 個の頂点を持つ多角形として，G を P の頂点を結んで得られた三角形分割のグラフとする．補題により，G の頂点で対角線で結ばれていないものが存在するが，その一つを v として G から，v と v につながっている二つの辺 e_0, e_1 を取り除いてできるグラフ G' を考える．G' に帰納法の仮定を適用すると G' の頂点は 3 色で塗り分けられるが，この色分けで e_0 と e_1 の端点で v でない方のものの G' の色分けでの色を c_0,

c_1 として 3 色のうち c_0, c_1 以外のものを v に割り当てることで拡張した色分けは G の頂点の 3 色塗り分けになっている． ❑ (定理 7.1)

8 雑

misc

8.1 2 次方程式

quadratic

複素数体 $\mathbb{C}$ で考える．一般の代数閉体で考えても同様だが，この場合については $\sqrt{a}$ は，$b^2 = a$ となるような b のうちの一つというような定義に変更する必要がある．

$a, b, c \in K$ として $a \neq 0$ とする．このとき，2 次方程式

$$ax^2 + bx + c = 0 \tag{8.1}$$ quad-1

を解く．$a \neq 0$ だから方程式の両辺を a で割ることができて

$$ax^2 + bx + c = 0 \ \Leftrightarrow \ x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \tag{8.2}$$ quad-2

である．ここで (8.1) の二つの解を α, β とすれば，

$$x^2 + \frac{b}{a}x + \frac{c}{a} = (x - \alpha)(x - \beta) \tag{8.3}$$ quad-3

だから，係数の比較から

$$\alpha + \beta = -\frac{b}{a}, \tag{8.4}$$ quad-4

$$\alpha\beta = \frac{c}{a} \tag{8.5}$$ quad-5

である．したがって，

$$(\alpha - \beta)^2 = (\alpha + \beta)^2 - 4\alpha\beta = \frac{b^2}{a^2} - 4\frac{c}{a} = \frac{b^2 - 4ac}{a^2} \tag{8.6}$$ quad-6

となり，このことから

$$\alpha - \beta = \pm\frac{\sqrt{b^2 - 4ac}}{a} \tag{8.7}$$ quad-7

である．したがって，

$$\alpha = \frac{1}{2}((\alpha + \beta) + (\alpha - \beta)) = \frac{1}{2}(-\frac{b}{a} \pm \frac{\sqrt{b^2 - 4ac}}{a}) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{8.8}$$ quad-8

である．β についても同様だから，

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{8.9}$$ quad-9

が方程式 (8.1) の 2 つの解であることがわかる．