数学 IB 講義録

2014 年度 夏学期

金井 雅彦 著

目次

第1章

数列の極限

A　実数の必要性

A.1　フィボナッチ数列と黄金比

皆さんすでに数列の極限についてはある程度の知識を持っていることと思います．そこで，復習として，まずは問題をひとつ解いていただきましょう．

問題 1.1. ――――――――――――――――――――――――――――――――

$a_0 = a_1 = 1$，および

$$a_{n+1} = a_n + a_{n-1} \qquad (n = 1, 2, \cdots) \tag{1.1}$$

により帰納的に定義される[1] 数列 $\{a_n\}$ に対し，極限 $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ を計算せよ．

――――――――――――――――――――――――――――――――――――

ここでちょっと脱線して，この問題に現れる数列 $\{a_n\}$ 自体についてお話しをしたいと思います．この数列の最初のいくつかの項を求めてみましょう．まず，最初のふたつの項はともに 1 です．その次の項は，$1 + 1 = 2$ です．そして，その次の項は，$1 + 2 = 3$, さらにその次の項は $2 + 3 = 5$ といった仕方で，この数列の項を順次求めていくことができます．先だつふたつの項の和が，その次の項の値になっているわけです．

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55, \cdots$$

と言った具合にこの数列は続きます．とくにこの数列 $\{a_n\}$ 自体は発散します．

[1]恐らく，高校までは初項の番号は 1 とするのが普通だったのではないでしょうか．しかし，初項の番号をゼロとする「流儀」もあります．このノートでは，番号を ∞ に発散させたときの極限を問題にしていますので，どちらの流儀を採用しようが，実質的な差異は一切生じません．そこで，これらふたつの流儀を混在させることにします．

この数列には**フィボナッチ数列** という名前がついています．また，この数列の項として現れる自然数を，**フィボナッチ数**と言います．フィボナッチ数は，自然界にしばしば現れる神秘的な数として，古くから知られています．[2]

寄り道はこのくらいにして，問題 1.1 について考えることにしましょう．ただ，問題の解答を始める前に，少し一緒にその解法について考えてみることにします．少々唐突かも知れませんが，漸化式 (1.1) を変形して，

$$a_{n+2} - \lambda a_{n+1} = \mu(a_{n+1} - \lambda a_n) \qquad (n = 0, 1, \cdots) \tag{1.2}$$

の形にできたとします．ただし，λ, μ は適当な定数とします．いま，新たな数列 $\{b_n\}$ を，

$$b_n = a_{n+1} - \lambda a_n \qquad (n = 0, 1, \cdots)$$

で定義すれば，(1.2) は，$b_{n+1} = \mu b_n$ ということに他なりません．すなわち，$\{b_n\}$ は公比 μ の等比数列です．

さて，ここでさらに $|\mu| < 1$ であると仮定しましょう．すると，$b_n = \mu^n b_0 \to 0$ $(n \to \infty)$ であることが従います．また，$a_n \to \infty$ $(n \to \infty)$ でした．したがって，

$$\frac{a_{n+1}}{a_n} - \lambda = \frac{b_n}{a_n} \longrightarrow 0,$$

すなわち，

$$\frac{a_{n+1}}{a_n} \longrightarrow \lambda$$

を得ます．

いま行った計算においては，次のふたつを仮定していました：まず第一に，式 (1.1) が (1.2) の形に変形可能であること，また第二に，(1.2) に現れる定数 μ が $|\mu| < 1$ を満たすようにとれること．これらふたつの仮定が実際に満たされることを確かめましょう．そのために，まず (1.2) は

$$a_{n+2} = (\lambda + \mu)a_{n+1} - \lambda\mu a_n$$

と変形されることに注意してください．この式と式 (1.1) の係数を比較することにより，

$$\lambda + \mu = 1, \qquad \lambda\mu = -1 \tag{1.3}$$

[2] フィボナッチ数列，あるいはフィボナッチ数に興味がある人は，例えば以下の書籍をご覧下さい：

- 中村滋著『フィボナッチ数の小宇宙 – フィボナッチ数，リュ数，黄金分割』，日本評論社，2002.
- R. A. ダンラップ著『黄金比とフィボナッチ数』日本評論社，2003.

を得ます．逆に，もし λ, μ が (1.3) を満たすならば，式 (1.1) は式 (1.2) と同値，すなわち，一方から他方に変形することが可能です．さらに，定数 λ, μ を決定するには，2 次方程式に対する解と係数の関係を思い出せば十分です．実際 x に対する 2 次方程式

$$x^2 - x - 1 = 0$$

のふたつの解 $(1 \pm \sqrt{5})/2$ の一方を λ と他方を μ とすればよいことになります．ここまでは，ふたつの実数 $(1 \pm \sqrt{5})/2$ のどちらを μ としても構いませんでした．しかし，μ にはもうひとつ条件 $|\mu| < 1$ が課されています．いま，$-1 < (1-\sqrt{5})/2 < 0 < 1 < (1+\sqrt{5})/2$ ですから，$\lambda = (1+\sqrt{5})/2$, $\mu = (1-\sqrt{5})/2$ とすればよいことが分かります．そして，このとき，

$$\frac{a_{n+1}}{a_n} \longrightarrow \lambda = \frac{1+\sqrt{5}}{2}$$

であることが結論されます．

以上が問題 1.1 の解法です．さて，これを答案の形にまとめ上げてみましょう．[3]

問題 1.1 の解答． λ, μ を 2 次方程式 $x^2 - x - 1 = 0$ のふたつの解とすると，$\lambda + \mu = 1$, $\lambda\mu = -1$ が成り立つ．したがって (1.1) は，

$$a_{n+1} - \lambda a_n = \mu(a_n - \lambda a_{n-1})$$

と変形できる．これは，$b_n = a_{n+1} - \lambda a_n$ で定義される数列 $\{b_n\}$ が，公比 μ の等比数列であることに他ならない．ここで，λ, μ を，$\lambda = (1+\sqrt{5})/2$, $\mu = (1-\sqrt{5})/2$ と取れば，とくに $|\mu| < 1$ であるから，$b_n \to 0$ $(n \to \infty)$ が成り立つ．一方，明らかに $n \to \infty$ のとき $a_n \to \infty$ である．したがって，$n \to \infty$ のとき，

$$\frac{a_{n+1}}{a_n} - \lambda = \frac{b_n}{a_n} \longrightarrow 0, \quad \text{すなわち} \quad \frac{a_{n+1}}{a_n} \longrightarrow \lambda = \frac{1+\sqrt{5}}{2}$$

が従う． □

1.1 の答に現れる極限値 $(1+\sqrt{5})/2$ は，しばしば**黄金比**と呼ばれます．[4] すでに皆さんご存知かもしれませんが，この黄金比は無理数であることが知られています．この黄金比こそが私達人類が最初に「遭遇」した無理数とのことです． ここ

[3] 上で述べた「考え方」と，下にまとめた「答案」の間には，大きな差異があるように感じられるのではないかと想像します．解答を書くに先だって考えたことと，答案における最終的な表現の仕方には，このような違いがあることが，数学においてはむしろ普通です．このような差異は，定理の証明と，その証明を見出すための考察との間にも存在します．教科書に書かれた証明や，あるいは演習問題の解答を読む際には，その背後にあるはずの着想を推し量る必要があります．そして，この点が多くの初学者を惑わす点でもあります．

[4] フィボナッチ数列に関する文献として挙げた 2 冊の本には，黄金比に関する記述もあります．黄金比に関しより詳しく知りたい人は，そちらをご覧下さい．

では黄金比の無理性を，とくにユークリッドの互除法と呼ばれる方法に基づき，証明したいと思います．

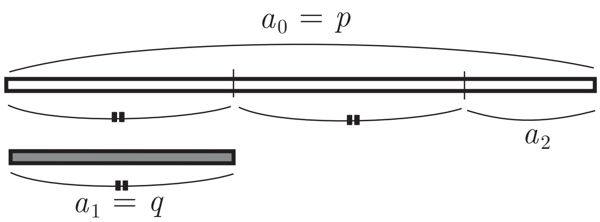
A.2　ユークリッドの互除法

いま，実数 p, q が与えられたとします．ただし，$q > 0$ とします．このとき，

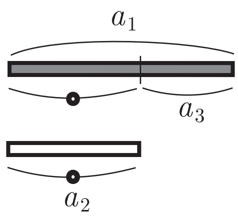
$$p = dq + r, \qquad 0 \leq r < q$$

を満たす整数 d および実数 r が一意的に存在します．これらはそれぞれ p を q で割ったときの商，および余りと考えられます．とくに r を，それが p, q の関数であることを強調するために，記号 $R(p, q)$ で表すことにします．

さて，次に，ふたつの正の実数 p, q が与えられたとしましょう．すると，$a_0 = p,\quad a_1 = q$，および $a_{n+1} = R(a_{n-1}, a_n)$ により，数列 $\{a_n\}$ が帰納的に定義出来ます．ただし，ある番号 N に対し $a_N = 0$ となった場合には，それより大きな番号 $n > N$ に対しては a_n は定義されていないものとします．この数列 $\{a_n\}$ を生成するアルゴリズムが**ユークリッドの互除法**と呼ばれるものです．そして，もしある番号 N に対し $a_N = 0$ となり，したがってそれより大きな番号 n に対しては a_n が定義されない場合には，p，q **に対する互除法は有限回で終了する**，と言うことにします．



(a) 最初のステップ



(b) 次のステップ

図 1.1: ユークリッドの互除法

ここで，この互除法をもう少し直感的な仕方で説明してみましょう．いま，白色と黒色のテープを２本用意します．ただし，白色のテープは長さ p，また黒色のテープの長さは q であるとします．話を簡単にするために，$p > q$，すなわち，白色のテープの方が長いと仮定しましょう．さてこの白色のテープから，黒色のテープと同じ長さを持つものを出来るだけたくさん切り取り，それらを捨ててしまいます．すると，残った白色のテープはもはや黒色のテープより短くなっているはずです．その残った白色のテープの長さが a_2 に他なりません．さて今度は黒色のテープと白色のテープの役割を交換して，黒色のテープから白色のテープと同じ長さを持つテープを出来るだけたくさん切り取り，捨ててしまいます．そのときに，残っ

た黒色のテープの長さが a_3 になります（図 1.1）．この作業を繰り返すことが，互除法に他なりません．なぜ「互除」法という言い方がなされるかも，これで理解いただけたのではないかと思います．

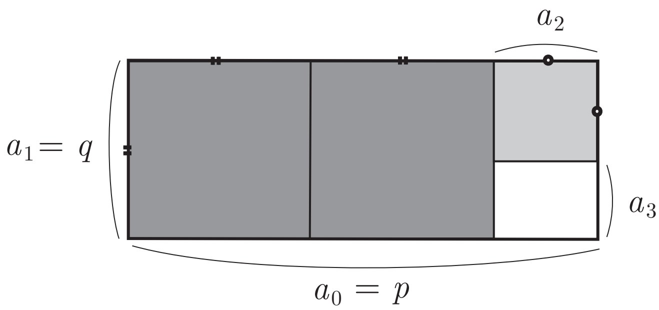


図 1.2: 長方形を用いたユークリッドの互除法の図示

この互除法の解釈の仕方は，また次のようにもう少し幾何学的に言い換えることが出来るかもしれません．上述の p, q に対し，隣り合う 2 辺の長さがそれぞれ $p = a_0, q = a_1$ に等しい長方形 A_1 を考えます（図 1.2）．このとき，この長方形の中に，1 辺の長さが q の正方形を，図のように出来るだけたくさん描き，それらをもとの長方形から取り除きます．するともともとの長方形より小さな長方形が新たに得られますので，それを A_2 と呼ぶことにします．その 4 つの辺のうち 2 つの長さは $a_1 = q$，そして残り 2 つの辺の長さは a_2 になります．いま長方形 A_1 に対して行ったのと同じ操作を今度は長方形 A_2 に対して行うことにより，さらに小さな長方形 A_3 を得ます．そしてその辺の長さとして a_3 が実現されます．もちろんこの後も同じ操作を可能な限り繰り返すことは言うまでもありません．そして，互除法が有限回で終了するということは，ある番号 N に対し，長方形 A_{N-1} がその短い方の辺と同じ長さの辺を持つ正方形いくつかにより過不足なく覆い尽くされる，ということと同値になります．

次の定理に述べるように，ユークリッドの互除法を用いて有理性の判定を行うことが出来ます．

定理 1.2.
正の実数 p, q に対し，その比 p/q が有理数であるための必要十分条件は，p, q に対する互除法が有限回で終了することである．

この定理の証明は簡単です．一緒に証明してみることにしましょう．まず，十分性を示します．そのために，p, q に対する互除法が有限回で終了すると仮定しま

しょう．このとき，番号 N を，$a_N \neq 0, a_{N+1} = 0$ を満たすよう取ることが出来ます．すると，数列 $\{a_n\}$ の定義によれば，

$$\begin{aligned} p = a_0 &= d_1 a_1 + a_2, \\ q = a_1 &= d_2 a_2 + a_3, \\ a_2 &= d_3 a_3 + a_4, \\ &\vdots \\ a_{N-3} &= d_{N-2} a_{N-2} + a_{N-1}, \\ a_{N-2} &= d_{N-1} a_{N-1} + a_N, \\ a_{N-1} &= d_N a_N \end{aligned}$$

が成り立ちます．ただし，$d_1, \cdots, d_N$ は正の整数です．まず，1 番最後の式によれば，a_{N-1} は a_N の整数倍です．次いで，1 番最後の式をそのひとつ前の式に代入すれば，a_{N-2} も a_N の整数倍であることが判ります．したがって，下から 3 番目の式とにより，a_{N-3} もまた a_N の整数倍であることが見て取れます．そして，この議論を繰り返すことにより，$q = a_0, p = a_1, a_2, \cdots, a_N$ すべてが a_N の整数倍であることが結論されます．とくに，$p = P \cdot a_N$, $q = Q \cdot a_N$ なる整数 P, Q を取ることが出来るから，比 $p/q = P/Q$ は有理数でなければなりません．

　次に逆を証明しましょう．そのためにいま比 p/q が有理数 P/Q（ただし P, Q は互いに素な正の整数）に等しいと仮定します．このときに，p, q に対する互除法が有限回で終了することを示せばよいわけです．ここで，互除法を前述のような仕方で長方形を用い幾何学的解釈するにあたって，一番最初に取った長方形を，それと相似な長方形に置き換えても互除法が有限回で終了するか否かに変わりはない，と言う事実に着目します．これより，勝手な正の実数 λ に対し，$\lambda p, \lambda q$ に対する互除法が有限回で終了するとき，およびそのときに限り，p, q に対する互除法が有限回で終了することが従います．この事実はまた，互除法の定義に基づき直接証明することも可能です．いずれにしても，この事実により，実数 p, q の代わりに，整数 P, Q から出発する互除法が有限回で終了することを示せばよいことになりました．これを背理法で証明しましょう．そこで，問題の互除法が有限回では終了しない，すなわち，a_n がすべての $n = 0, 1, \cdots$ に対し定義されたとしましょう．このとき数列 $\{a_n\}$ が

$$a_n > 0 \qquad (n = 0, 1, \cdots)$$

および

$$a_1 > a_2 > \cdots \tag{1.4}$$

を満たしてなければならないことが，互除法の定義から直ちに導かれます．ちなみに，第 2 の主張において番号 $n = 0$ を除外したのは，$P < Q$ の場合を配慮してのことです．さらにもうひとつ重要なことは，いま正の整数 P, Q から出発して互除

法を行っているために，数列 $\{a_n\}$ のすべての項もまた正の整数でなければならないことです．しかし，a_1 より小さい正の整数は有限個しかありませんから，(1.4) は明らかに矛盾しています．したがって，互除法が有限回で終了しなければならないことが結論されます．以上で定理 1.2 の証明が終了しました．

さて，ここで本当に簡単な問題を出します．

問題 1.3. ――――――――――――――――――――――――

いま，隣り合う 2 辺の長さが 1，および λ（ただし $\lambda > 1$）であるような長方形が与えられたとする．この長方形の内部に，この長方形の長さ 1 の辺を 1 辺として持つような正方形を取る (図 1.3)．その正方形を長方形から取り除いて得られる小さな長方形が，もとの大きな長方形と相似であるとき，もとの長方形の長い方の辺の長さ λ を求めよ．

――――――――――――――――――――――――――――――

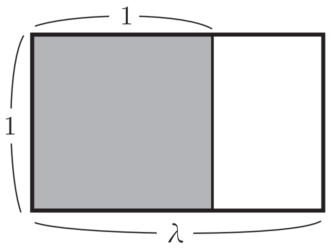


図 1.3:

大小ふたつの長方形に対する相似条件より，

$$1 : \lambda = \lambda - 1 : 1, \quad \text{すなわち} \quad \lambda^2 - \lambda - 1 = 0$$

を得ます．ところが，この 2 次方程式は問題 1.1 の解答に現れたものに他なりません．そして，そのふたつの解のうち $\lambda > 1$ を満たすのが黄金比 $(1+\sqrt{5})/2$ でした．

ここまで述べれば，黄金比が無理数であることが容易に理解されます．実際，黄金比が上述の問題の解として得られること，および互除法の長方形を用いた幾何学的解釈により，1，および λ から出発する互除法が有限回では決して終了しないことが直ちに判ります．そして，定理 1.2 から黄金比が無理数であることが結論されます．

A.3　数の体系

ここで数の体系について手短に復習しておきたいと思います．[5] ご存知の通り，数の概念は，

自然数，整数，有理数，実数，複素数

といった仕方で，順次拡張されていきます．[6] そもそもなぜ，このように数の概念を拡張していく必要があったのでしょうか．皆さんが小学校に入学する前後に最初に習ったのが，$1, 2, \cdots$ という数，すなわち自然数です．[7] そしてそこで自然数の加法・乗法を学習したと思います．大事なことは，自然数全体はこれら２種類の演算について閉じている，すなわちふたつの自然数を任意[8] にとったとき，それらの和，および積は再び自然数になる，という事実です．しかしながら，例えば $1-3$ を考えれば判るように，自然数全体は減法に関しては閉じていません．減法を自由に行うことは自然数の世界では許されないわけです．この不自由を解消するために，新たに負の数 $-1, -2, \cdots$，およびゼロが導入されました．そしてその結果得られる数の全体 $0, \pm 1, \pm 2, \cdots$ — そのおのおのを整数と呼ぶことは言うまでもないことですが — を考えると，今度は３つの演算，加法・減法・乗法に関し閉じた体系となります．しかしまだ除法を自由に行うことは出来ません．この問題を克服するためには有理数，すなわち整数を分子・分母とする分数を考える必要があったわけです．有理数まで話を進めると，今度はいわゆる四則演算，すなわち，加減乗除，４つの演算すべてに関し閉じた体系が得られます．有理数全体は代数的観点からすれば満足のいく対象に思われます．

---

[5]より詳しくは，

- H.-D. エビングハウス他著『数（上)』シュプリンガー・フェアラーク東京，1991

をご覧下さい．また，古い本ですが，

- 高木貞治著『数の概念（改版)』岩波書店，1970

もお勧めです．ただし，もはや書店で入手することは困難ではないかと思います．図書館などをあたってみて下さい．

[6]通常，自然数全部を集めて得られる集合，整数全部を集めて得られる集合，有理数全部を集めて得られる集合，実数全部を集めて得られる集合，複素数全部を集めて得られる集合を，それぞれ，$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ と書きます．これらの各々はそれに続くものの部分集合になっています：

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

[7]ゼロを自然数に含めることもあります．どちらが良いのかに関しては，必ずしも定見はないようです．

[8]「**任意**」という言葉も数学では非常によく使われます.「ふたつの自然数を任意にとったとき」とは,「どのようにふたつの自然数をとろうとも」といった意味です.「任意の」という表現と「すべての」という表現は，ニュアンスは異なるものの，形式的には同じ意味だと考えても間違いないと思います．

それではなぜ，さらに数の体系を実数まで拡張する，言い換えれば，無理数を導入する必要があるのでしょうか．その答を問題 1.1 に見い出すことが出来ます．いま問題 1.1 数列 $\{a_n\}$ に対し $r_n = a_{n+1}/a_n$ とおけば，その解答で示したことは，$\lim_{n\to\infty} r_n = (1+\sqrt{5})/2$ ということに他なりません．ここで極限値として得られた黄金比 $(1+\sqrt{5})/2$ は既に見た通り無理数です．ところが一方，数列の各項 r_n は（$\{a_n\}$ の各項が自然数であることより），有理数です．したがってこの問題は，すべての項が有理数であるような数列，すなわち有理数列の極限が必ずしも有理数でないことを示す例となっているわけです．このような例はこれ以外にも，文字通り無数に存在します．もし無理数を考えることを放棄するならば，こういった数列の極限を論ずることはもはや不可能になります．ところが一方，微分・積分といった微積分学の基本概念は何らかの極限として定義されます．したがって有理数の世界の中では，微積分の理論を展開することは不可能です．無理数導入の必要性はここにあったわけです．換言すれば，有理数全体をさらに極限操作に関し閉じた体系になるよう拡張したもの，それが実数の体系です．なお，蛇足になりますが，実数から複素数への拡張は，任意の代数方程式 $a_n z^n + \cdots + a_1 z + a_0 = 0$ の解の存在が後者においては保証される，という点にあります．

ふたたび，p, q を正の実数，$\{a_n\}$ を p, q に対し互除法を適用して得られる（有限ないし無限の長さを持つ）数列とします．

$$
\begin{aligned}
p = a_0 &= d_1 a_1 + a_2, \\
q = a_1 &= d_2 a_2 + a_3, \\
a_2 &= d_3 a_3 + a_4, \\
a_3 &= d_4 a_4 + a_5, \\
&\vdots
\end{aligned}
$$

が成り立ちます．したがって，

$$
\begin{aligned}
\frac{p}{q} = \frac{a_0}{a_1} &= d_1 + \frac{a_2}{a_1} \\
&= d_1 + \cfrac{1}{d_2 + \cfrac{a_3}{a_2}} \\
&= d_1 + \cfrac{1}{d_2 + \cfrac{1}{d_3 + \cfrac{a_3}{a_2}}} \\
&\vdots
\end{aligned}
$$

$$= d_1 + \cfrac{1}{d_2 + \cfrac{1}{d_3 + \cfrac{1}{d_4 + \cfrac{1}{\ddots}}}}$$

を得ます．最後の式を，実数 p/q の**連分数展開**と呼びます．たとえば，黄金比の連分数展開は，

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}$$

となります．

B　有界単調数列の収束

B.1　極限計算に関するある教訓

ところで皆さんの中には，問題 1.1 に対して次のような解答を考えた人がいるかもしれません．

問題 1.1 に対する「別解」． $r_n = a_{n+1}/a_n$ と置けば，漸化式 (1.1) の両辺を a_n で割ることにより，数列 $\{r_n\}$ に対する漸化式

$$r_n = 1 + \frac{1}{r_{n-1}} \tag{1.5}$$

を得る．したがって，$\rho = \lim_{n\to\infty} r_n$ と置けば，この漸化式において $n \to \infty$ とすることにより

$$\rho = 1 + \frac{1}{\rho}$$

を得る．これを解けば $\rho = (1 \pm \sqrt{5})/2$. 一方，明らかに $a_n > 0 \quad (n = 0, 1, \cdots)$ であるから, $r_n > 0$ $(n = 1, 2, \cdots)$，したがって $\rho \geq 0$ でなければならないことが判る．よって，$\lim_{n\to\infty} r_n = \rho = (1 + \sqrt{5})/2$ が結論される． □

この第 2 の解答の方が，第 1 のものと比べてエレガントに見えると感じる人も多いと思います．しかし残念ながらこの第 2 の解答には，極めて重大な欠陥があります．もちろんそこで得られた最終的な結果，すなわち極限の値自体は，最初の解答で得られたものと一致しているわけで，なんら問題はありません．問題なのは，漸化式 (1.5) から $\rho = (1 \pm \sqrt{5})/2$ を結論する議論の進め方にあります．

何が誤りかをはっきりさせるため，次のような例を考えましょう．$\{s_n\}$ を，

$$s_0 = 1 \qquad \text{および} \qquad s_{n+1} = 2s_n + 1 \quad (n = 0, 1 \cdots) \tag{1.6}$$

で定義される数列とします．このとき，問題 1.1 に対する「別解」と同様に $\sigma = \lim_{n\to\infty} s_n$ とおけば，(1.6) において $n \to \infty$ とすることにより $\sigma = 2\sigma + 1$，すなわち $\sigma = -1$ を得ます．しかしこれは，明らかに間違いです．なぜなら，(1.6) によれば, 数列 $\{s_n\}$ は発散しなければならないからです．

この例が示すとおり，問題 1.1 の「別解」で用いた論法は，われわれを誤った答に導く可能性があります．そして，そのような論法を用いているという意味で，問題 1.1 の「別解」は誤った解答であると言わざるを得ません．それではなぜ第 2 の例でわれわれは正しくない答 $\sigma = -1$ を得たのでしょうか．その理由は，問題の数列の極限が（有限の値を持つ通常の数として）存在すると初めから暗黙のうちに仮定している点にあります．例えば上述の例では，$\sigma = \lim_{n\to\infty} s_n$ とおいて σ を通常の数のように取り扱い，とくにそれに対する方程式 $\sigma = 2\sigma + 1$ から $\sigma = -1$ であると結論する部分がこれに相当します．そこで，**教訓 — 極限値を求めるにあたって，はじめからその極限値が存在すると証明なしに仮定してはならない．**

B.2　有界単調数列とその収束

それではどのような状況において，数列の極限は存在するのでしょうか．もし適当な仮定の下で，問題の数列の極限の存在を保証するような定理があれば，上述のような問題を解くにあたって，極めて有効なはずです．このような定理のひとつとして，このあと述べる有界単調数列の収束定理があります．

ただ定理を述べるに先だって，いくつか定義を与える必要があります．

定義 1.4. ――――――――――――――――――――――――――――――

$\{a_n\}$ を数列とする．

(1) もし

$$a_0 \leq a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots$$

が成り立つならば，$\{a_n\}$ は**広義単調増加**，あるいは単に**単調増加**であると言われる．また，逆に

$$a_0 \geq a_1 \geq \cdots \geq a_n \geq a_{n+1} \geq \cdots$$

が成り立つならば，$\{a_n\}$ は**広義単調減少**，あるいは単に**単調減少**であると言われる．[9]

[9]通常，ここで述べた広義単調性とは別に，狭義単調性が次のようにして定義されます．数列 $\{a_n\}$

(2) 適当[10] な定数 K に対し

$$a_n \leq K \qquad (n = 0, 1, \cdots)$$

が成り立つとき，数列 $\{a_n\}$ は**上に有界**であると言われる．また，適当な定数 K に対し

$$a_n \geq K \qquad (n = 0, 1, \cdots)$$

が成り立つならば，数列 $\{a_n\}$ は**下に有界**であると言われる．

例えば，$a_n = n$ で定義される数列 $\{a_n\}$ は単調増加ですが，上に有界ではありません．一方，$\{b_n\}$ を $b_n = -2^{-n}$ で定義される数列とすると，これは単調増加，かつ上に有界となります．

定理 1.5 (有界単調数列の収束)．

上に有界な単調増加数列は必ず収束する．また，下に有界な単調減少数列も必ず収束する．

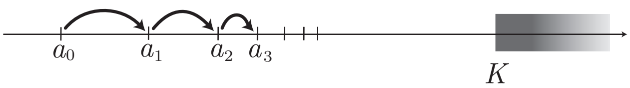


図 1.4: 上に有界な単調増加数列

たとえば，$\{a_n\}$ を上に有界な単調増加数列としましょう．その各項 a_n を実数直線上にプロットするとすと，単調増加性から，a_n に対応する点は，番号 n が増えるにつれて右に移動する，あるいはより厳密な言い方をすると，決して左方向に

に対し，もし

$$a_0 < a_1 < \cdots < a_n < a_{n+1} < \cdots$$

が成り立つならば，$\{a_n\}$ は**狭義単調増加**であると，また

$$a_0 > a_1 > \cdots > a_n > a_{n+1} > \cdots$$

が成り立つならば，$\{a_n\}$ は**狭義単調減少**であると言われます．広義の単調性と狭義のそれとの差異は，前者においてはその定義に等号付きの不等号が用いられておりその結果隣り合う 2 項が同じ値をとることが許されているのに対し，後者においては数列の値が本当に増え続けるか，あるいは減り続けなければならないことにあります．この本では広義の単調性しか利用しない予定ですので，それを単に単調性と称することにします．

[10]「**適当に**」という表現は，日常会話では「いいかげんに」といった意味で使われることがしばしばです．例えば,「適当なことを言うんじゃありません！」なんていう表現を思い出して下さい．しかし,「適当に」という表現を,「(目的などに）ぴったりあった」,「適切な」という意味で用いることもあります.「適当な人材に恵まれる」などといった用法がこれにあたります．数学においては，常に後者の意味でこの言葉を用います．

引き返すことがないことになります．一方，上からの有界性によれば，実数直線上に定数 K に対応する点が存在し，数列 $\{a_n\}$ の各項に対応する点はそれより右側には存在しないことが保証されます．真っ直ぐな道を，決して後戻りはしないという頑固なうさぎが，ぴょこぴょこはねながらやってくるのですが，その前方には非常に高い塀があって，そのうさぎはそれを決して飛び越えることが出来ないといった，うさぎにとっては少々気の毒な状況に例えることが出来るかも知れません．時刻を無限大に発散させたときに，そのうさぎの位置は，塀の手前のある点に収束する，というのが，有界単調数列の収束定理が主張している内容です．こう考えると，定理の主張の正当性がある程度直感的に把握できるかも知れません．また，このように視覚的に解釈すれば，たとえば単調増加数列の収束性を保証するため必要なのは，下への有界性ではなく，上への有界性であることも直ちに理解されると思います．

B.3　簡単な応用例

有界単調数列の収束定理の証明は後回しにして，むしろ，定理の使い方，あるいは定理の有効性を示したいと思います．まず簡単な例題として，

$$\lim_{n\to\infty} \lambda^n = 0$$

（ただし λ は $0 < \lambda < 1$ なる定数）が成り立つことを，有界単調数列の収束定理を用いて証明してみましょう（ちなみに，この事実に類するものが，既に問題 1.1 の解答の中でも利用されています）．[11] 以下，$a_n = \lambda^n$ $(n = 0, 1, \cdots)$ と置きます．するとただちに

$$a_n = \lambda \cdot a_{n-1} \tag{2.1}$$

であることが判ります．ここで $0 < \lambda < 1$，かつ $a_n > 0$ $(n = 0, 1, \cdots)$ であることから $\{a_n\}$ が単調減少であることが判ります．一方，先ほど述べた通り，$a_n > 0$，従って特に $\{a_n\}$ は下に有界です．よって有界単調数列の収束定理より，数列 $\{a_n\}$ に対し極限 $\alpha = \lim_{n\to\infty} a_n$ が存在することが判ります．そこで，(2.1) において $n \to \infty$ とすると $\alpha = \lambda\alpha$，すなわち $(1-\lambda)\alpha = 0$ を得ます．ここで $\lambda \neq 1$ であることを思い起こせば $\alpha = 0$ が結論されます．

類似の，しかし少しだけ進んだ問題を考えましょう．

[11]有界単調数列の収束定理 1.5 を適用する際には，まず与えられた数列が，単調増加なのか，それとも単調減少なのかを見極める必要があります．その後に，示すべきは上への有界性なのか，それとも下への有界性なのかを考えればいいのです．

問題 1.6. ――――――――――――――――――――――――――――――――――

以下を証明せよ：

$$\frac{2^n}{n!} \longrightarrow 0 \qquad (n \to \infty)$$

――――――――――――――――――――――――――――――――――――――

解答. $a_n = 2^n/n!$ $(n \geq 1)$ とおけば，

$$a_{n+1} = \frac{2}{n+1} a_n \tag{1.7}$$

が成り立つ．$n \geq 1$ のもとで $2/(n+1) \leq 1$ であるから，数列 $\{a_n\}$ は $n = 1$ より先で単調減少である．また，数列 $\{a_n\}$ は明らかに下に有界である．したがって，数列 $\{a_n\}$ は収束する．[12] いま，その極限を α としよう．すると，漸化式 (1.7) において $n \to \infty$ とすることにより，$\alpha = 0 \cdot \alpha = 0$ を得る． □

上の問題に関し，少し補足説明をしましょう．ふたつの数列

$$b_n = 2^n, \qquad c_n = n! \qquad (n = 1, 2, \cdots)$$

は，ともに $n \to \infty$ において $+\infty$ に発散する数列です．しかし，それらの発散の速さを比較したとき，数列 $\{c_n\}$ の方が，数列 $\{b_n\}$ より早く発散するということを，問題 1.6 は意味します．

B.4　ネピア数

さらに進んだ応用として次のような問題を考えましょう．

問題 1.7. ――――――――――――――――――――――――――――――――――

数列 $\{s_n\}$ を，

$$s_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}, \qquad n = 1, 2, \cdots$$

により定義する．

[12]通常，$0! = 1$ と約束します．したがって，

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 2, \quad a_3 = \frac{4}{3}, \quad \cdots$$

といった具合に，この数列は続きます．とくに，$n = 0$ から $n = 1$ にかけて，この数列は増加します．しかし，それから先では単調減少です．最初に単調減少でない部分があっても，ある番号から先ずっと単調減少であり，さらに下に有界であるならば，その数列は収束します．なぜなら，$n \to \infty$ における極限を問題にする限り，数列の最初の有限個の項がどう振る舞おうが全く問題ないからです．数列の最初の有限個の項を取り去ったり，あるいは値を置き換えても，数列の収束や極限の値には一切影響がありません．

(1)　この数列が収束することを証明せよ.

(2)　極限 $e = \lim_{n\to\infty} s_n$ の値を，誤差 10^{-4} 以内で求めよ.

以下でこの問題に対する解説を行います．ただ，それを読み始める前に各自解答を試みることを是非ともお勧めしたいと思います.

まずは設問 (1)．数列 $\{s_n\}$ が単調増加であることは明らかです．したがって，あとはそれが上に有界であることを示せば，有界単調数列の収束定理により数列 $\{s_n\}$ が収束することが結論されます．では，どうやって数列 $\{s_n\}$ が上に有界であることを示したらよいでしょうか．ところで，この数列 $\{s_n\}$ は，$a_n = 1/n!\ (n = 0, 1, \cdots)$ で定義される数列 $\{a_n\}$ の最初の項 a_0 から $(n+1)$ 番目の項 a_n までの和と考えることができます．すなわち，$s_n = a_0 + a_1 + \cdots + a_n$ です．したがって，数列 $\{s_n\}$ の極限は（もしそれが存在することが判ったときには），数列 $\{a_n\}$ の総和 $\displaystyle\sum_{n=0}^{\infty} a_n$ と見なせます．仮にこの数列 $\{a_n\}$ に対しても，等比数列に対するような和の公式があれば，それを用いることにより総和 $\displaystyle\sum_{n=0}^{\infty} a_n$ が有限の値を持つこと，したがってまた元の数列 $\{s_n\}$ が上から有界であることが結論されるかも知れません．しかし，残念ながらそのような公式は数列 $a_n = 1/n!$ に対しては存在しません．そこで少しだけ考え方を変えてみましょう．いま別の数列 $\{b_n\}$ で，$a_n \leq b_n$ が（十分大きな）任意の番号 n に対し成立し，しかもこの新たな数列 $\{b_n\}$ に対しては和の公式が存在し，その結果総和 $\displaystyle\sum_{n=0}^{\infty} b_n$ が有限であることが判ったとしたらどうでしょう．実は，数列 $\{s_n\}$ が上に有界であることを示すためにはこれで十分なのです．実際，もし $a_n \leq b_n$ がすべての n に対し成り立つとすると，

$$s_n = a_0 + a_1 + \cdots + a_n \leq b_0 + b_1 + \cdots + b_n \leq \sum_{n=0}^{\infty} b_n < \infty$$

となります．もし $a_n \leq b_n$ が最初の有限個の番号 n に対しては成立していないとしても，ある番号から先では必ず成り立っているならば，この論法は簡単な改良を行うだけで適用可能です．それでは，ここで期待されているような性質を有する数列 $\{b_n\}$ として，どの様なものをとってきたらよいのでしょうか．答えは $b_n = 1/2^n$ です．実際，任意の $n \geq 4$ に対し，$2^n \leq n!$ ですから，確かに $a_n \leq b_n$ が成り立ち，またこの数列 $\{b_n\}$ は公比 $1/2$ の等比数列ですから，$\displaystyle\sum_{n=0}^{\infty} b_n = 2 < \infty$ です.

上述の議論を解答として書き直してみましょう.

問題 1.7 (1) の解答． $2^n \leq n!\ (n \geq 4)$ であるから

$$\begin{aligned} s_n &= 1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!} \\ &\leq 1+1+\frac{1}{2}+\frac{1}{6}+\frac{1}{2^4}+\frac{1}{2^5}+\cdots+\frac{1}{2^n} \\ &\leq \frac{8}{3}+\sum_{n=4}^{\infty}\frac{1}{2^n}=\frac{8}{3}+\frac{1}{8} \end{aligned}$$

(ただし $n \geq 4$) したがって $\{s_n\}$ は上に有界である．一方，$\{s_n\}$ が単調増加であることは明らかである．よって，有界単調数列に対する収束定理より数列 $\{s_n\}$ は確かに収束することが分かる． □

さあ，設問 (2) に進みましょう．記号 e を用いていることから気づいた人も多いと思いますが，問題の極限 $e=\lim_{n\to\infty} s_n$ は通常自然対数の底として用いられる定数です．また，イギリスの数学者の名前をとって**ネピア数**と呼ばれることもしばしばあります．[13] そして後に章末問題 1.18 において示すように，e は無理数であることが知られています．したがって，それを小数表示しようとすると，循環しない無限小数となります．たとえば数学辞典[14] には，その最初の数十桁が記載されています．しかし，ときには，人には頼らずに自分で近似値を求めなければならないこともあるかも知れません．そのときに備えるための練習問題が，設問 (2) です．

さて，定数 e は無限和

$$e=1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots$$

と考えられます．したがってそれを途中で打ち切り，有限和にした

$$s_n=1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}$$

は e の近似値と，また残りの部分

$$\epsilon_n=\frac{1}{(n+1)!}+\frac{1}{(n+2)!}+\cdots$$

はその誤差と考えれれます．そもそも e は数列 $\{s_n\}$ の極限として定義されたわけですから，番号 n が大きくなればなるほど s_n による e の近似の精度が向上するのは言うまでもありません．それでは，求められている精度，すなわち $\epsilon_n \leq 10^{-4}$ が満たされるようにするには果たして n としてどんな数をとればよいのでしょうか．

数学的な根拠なしに，例えば「$n=10^{10}$ とすれば十分だろう」と答える人が，もしかすると皆さんの中にもいるかもしれません．でも，本当はそれでは不十分だっ

[13]ただし，この定数を「発見」したのは彼ではないとのことです．
[14]日本数学会編『岩波数学辞典（第４版）』岩波書店，2007.

たらどうしましょう．やはり要求されている精度が達成されている確実な保証，すなわち証明が欲しくなります．またこの問題に関して言えば，例えば $n=10^{10}$ とすれば確かに求められている精度を十二分に満たしていますが，一方計算量を節約するために可能な限り小さな n をとりたい，というのも極めて自然な要求です．この問題は近似値計算としては簡単なものなので，計算量の多寡があまり実感できないかもしれません．しかし，より高度な近似値計算を実行するにあたっては，計算量の問題はコンピュータを利用する際にもしばしば極めて切実な問題となるそうです．

余談はこれくらいにしましょう．結局，問題は

$$\epsilon_n \leq 10^{-4}$$

を満たす n であって「可能な限り小さいもの」を求めよ，ということになります．そしてこれをとくにあたって設問 (1) で用いた手法が再び利用できます．実際，

$$\begin{aligned}
\epsilon_n &= \frac{1}{(n+1)!}+\frac{1}{(n+2)!}+\cdots \\
&= \frac{1}{(n+1)!}\left\{1+\frac{1}{n+2}+\frac{1}{(n+2)(n+3)}+\cdots\right\} \\
&< \frac{1}{(n+1)!}\left\{1+\frac{1}{n+2}+\frac{1}{(n+2)^2}+\cdots\right\} \\
&= \frac{1}{(n+1)!}\cdot\frac{1}{1-\frac{1}{n+2}} \\
&= \frac{n+2}{(n+1)(n+1)!}.
\end{aligned}$$

ここで $(n+2)/(n+1)(n+1)!$ の値を $n=1,2,\cdots$ に対し順次計算していくと，$n=7$ のときその値が初めて 10^{-4} 以下になっていることが分かります．[15] したがって，求めるべき近似値は，

$$s_7 = 1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{7!}$$

であることが結論されます．

C　収束定理の証明

話題を有界単調数列の収束定理の証明に転ずることにします．その準備として，まず次のような補題[16] を考えましょう．

[15] しかし，これは必ずしも，「$\epsilon_n \leq 10^{-4}$ となる最小の n は $n=7$ である」ということではありません．

[16] 「**定理**」，「**命題**」，「**補題**」，「**系**」などは，いずれも数学的な正しい主張を意味します．そして，そのいずれもが，それに対する証明を必要とします．証明なしには，その正しさは認められません．

補題 1.8.
$\{a_n\}$ をそのすべての項が整数であるような数列とする．もし $\{a_n\}$ が上に有界かつ単調増加であるか，あるいは下に有界かつ単調減少であるならば，ある番号 N から先，数列 $\{a_n\}$ の各項は一定の値をとり続ける：すなわち，$a_N = a_{N+1} = a_{N+2} = \cdots$ が成り立つ．

この補題において注意すべきことは，そこで扱う数列は整数列，つまりそのすべての項が整数であるような数列に限られている点です．さてここで，第B節に現れた頑固うさぎのことを思い出して下さい．この補題1.8の設定は，頑固うさぎの例を借りるならば次のように例えることが出来るでしょう．やはり，以前と同様，真っ直ぐな道を決して後戻りをしない頑固なうさぎが飛び跳ねながらやってきます．一方，その前方にそのうさぎには飛び越えることの出来ない塀があるというのも，以前と同様です．前回との相違点は，その道には敷石が一列に（しかも等間隔に）置かれていて，うさぎはその上を飛び跳ねてくると言う点です．もちろんこれは補題1.8における数列が，整数列であることの比喩です．するとある時刻以降，その可哀想なうさぎはあるひとつの敷石の上で飛び跳ね続けるしかない，というのが補題1.8の主張です．このように考えると，補題1.8で主張されている内容が正しいことが，苦もなく理解されるでしょう．すなわち，この補題1.8は証明を必要としないほど自明[17] な主張なわけです．そしてこれは，数列の取り得る値が整数という（実数全体の中に）離散的に分布する数に限られていることに起因します．ちなみに，この補題1.8で述べられていることと実質的には同等な事柄が，すでに定理1.2の証明の後半部分でも利用されています．

さて，定理1.5の証明を始めましょう．$\{a_n\}$ が上に有界な単調増加数列であるとしたとき，これが必ず収束することを，以下において証明します．$\{a_n\}$ が下に有界な単調減少数列の場合にも本質的には同様な仕方で，あるいは $\{a_n\}$ から $b_n = -a_n$ により定義される数列 $\{b_n\}$ が，上に有界かつ単調増加であることを利用して容易に示されるので，ここでは省略することにします．まず数列 $\{a_n\}$ の各項を整数部分と小数部分とに分解します．すなわち，各 $n = 0, 1, \cdots$ に対し，

$$a_n = M_n + r_n$$

形式的には，これら4つの単語の意味するものは一致します．それでは，なぜわざわざ言い方を変えるのでしょうか.「命題」は「定理」ほど重要ではない主張を意味します．一方,「補題」は別の「定理」や「命題」を証明するのに際して必要となる主張です．また,「系」は，別の「定理」や「命題」から（比較的容易に）導出されるものを意味します．しかし，いずれにせよ，これらの使い分けは，多分に感覚的なものであり，それが身につくには若干の慣れを必要とします．

[17]「**自明**」とは,「当たり前」，あるいは「当然成り立つ」と言った意味です．数学ではよく使われる言葉です．

（ただし M_n は整数，$0 \leq r_n < 1$）とするわけです．さらに各 r_n を，

$$r_n = 0, d_{n1} d_{n2} \cdots$$

（ただし $d_{n1}, d_{n2}, \cdots$ は 0 以上 9 以下の整数）のよう，10 進小数として表記しましょう．もちろん，

$$a_n = M_n + \frac{d_{n1}}{10} + \frac{d_{n2}}{100} + \frac{d_{n3}}{1000} + \cdots$$

が成り立ちます．[18] もとの数列 $\{a_n\}$ が上から有界，かつ単調増加であることより，その整数部分を取ることにより得られる数列 $\{M_n\}$ も上から有界，かつ単調増加であることが直ちに判ります．しかもこれは整数列であるから，補題 1.8 で述べたように，ある番号 N_0 から先，M_n は同一の値 M を取らなければならないことが保証されます．

次いで，小数点以下第 1 位の数字を集めて得られる数列 $\{d_{n1}\}$ に注目しましょう．これも明らかに上に有界な整数列です．しかも, もともとの数列 $\{a_n\}$ が単調増加であること，またその各項 a_n の整数部分 M_n が，番号 N_0 から先変化しないことより，$\{d_{n1}\}$ は番号 N_0 から先では単調に増加することが判ります．したがって，再び補題 1.8 によれば，ある番号 N_1 （ただし $N_1 \geq N_0$) から先，d_{n1} は同一の値 d_1 をとり続けねばなりません. 全く同様な議論を，小数点以下第 2 位以降に対しても行うことにより $d_2, d_3, \cdots$ が定義されます.

以上の議論は，自動車に付いている走行距離計のようなデジタル・メーターを思い浮かべると，容易に理解できるかも知れません．そのメーターには，まず整数部分を表す部分があり，さらにその右側に，小数点以下の各位を表す部分が一列にしかも無限に並んでいます（実際の走行距離計は有限桁しか持ちませんが，その点には目をつぶって下さい）．さて，そのメーターに表示されている数値が刻々変化している状況を想定しましょう．まず，第一にそこに表れる数値は，決してある数よりは大きくならないと仮定します．第二に，そのメーターは，自動車の走行距離計と同様，逆回転はしない，したがって，その数値が時間の経過とともに減少することはあり得ない，と仮定します．すると，どうでしょう．ある時間が経過すると，まず整数部分をあらわす桁が停止し，その後は決して値を変えなくなるはずです．そして，さらに時間が経過するにつれ，小数部分を表す各桁が，左から順に停止していくはずです．

[18] 厳密なことを言うと，ここで少々注意が必要です．実数の小数表記はひと通りとは限らないからです．実際，例えば，無限小数 $0.0999\cdots$ は 0.1 に等しいと考えられます．すなわち，小数 0.1 はもうひとつの小数表示 $0.0999\cdots$ を有するわけです．このような場合，無限小数表示をするのではなく，有限小数表示をすることにしておくと約束しておきましょう．

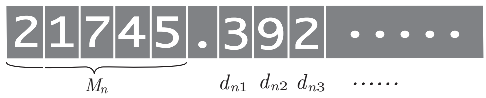


図 1.5: 走行距離計

さて，先程求めた $d_1, d_2, \cdots$ を用い，まず実数 r を 10 進表記を用い

$$r = 0.d_1d_2\cdots$$

で，次いで実数 a を

$$a = M + r$$

で定義しましょう．すると，$a_n \to a$ であることを結論するのは容易です．以上で，有界単調数列の収束定理の証明は終わりです．[19]

D　章末問題

まずは，問題 1.1 の極めて簡単な類題から始めましょう．

問題 1.9. ――――――――――――――――――――――――――――――――

数列 $\{a_n\}$ を，

$$a_1 = 1, \qquad a_2 = 2,$$
$$a_{n+2} = 2a_{n+1} + a_n \quad (n = 1, 2, \cdots) \tag{*}$$

により定義する．

(1) 漸化式 $(*)$ は，$a_{n+2} - \lambda a_{n+1} = \mu(a_{n+1} - \lambda a_n)$（ただし，$\lambda, \mu$ は定数）の形に変形できることを示せ．また，λ, μ の値を求めよ．

(2) $n \to \infty$ としたときの a_{n+1}/a_n の極限を求めよ．

――――――――――――――――――――――――――――――――――――

解答. (1) 等式 $a_{n+2} - \lambda a_{n+1} = \mu(a_{n+1} - \lambda a_n)$ は，

$$a_{n+2} = (\lambda + \mu)a_{n+1} - \lambda\mu a_n$$

と変形できる．これと漸化式 $(*)$ の係数を比較することにより，

$$\lambda + \mu = 2, \qquad \lambda\mu = -1$$

[19] 数列の収束概念の厳密な定式化として，ϵ-N 論法と称されるものがあります．上に述べた有界単調数列の収束定理の証明を，ϵ-N 論法による数列の収束の定義に基づいたものに書き直すことは，別に難しいことではありません．しかし，この講義では，数列の収束は ϵ-N 論法を用いずあくまで直感的に扱うに止めることにします．

を得る．したがって，λ, μ を，x に対する 2 次方程式

$$x^2 - 2x - 1 = 0$$

の解 $1 \pm \sqrt{2}$ とすれば，確かに $(*)$ は求められた形の式に変形可能である．

(2) $b_n = a_{n+1} - \lambda a_n$ とおく．(1) によれば，$\{b_n\}$ は公比 μ の等比数列である．ここで，とくに

$$\lambda = 1 + \sqrt{2}, \qquad \mu = 1 - \sqrt{2}$$

とおくと，$|\mu| < 1$ であるから，

$$b_n = a_{n+1} - \lambda a_n = (a_1 - \lambda a_0)\mu^n \longrightarrow 0 \qquad (n \to \infty)$$

が従う．一方，明らかに $a_n \to \infty$ $(n \to \infty)$ であるから，$n \to \infty$ のとき，

$$\frac{a_{n+1}}{a_n} - \lambda = \frac{b_n}{a_n} \longrightarrow 0, \quad \text{すなわち，} \quad \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lambda = 1 + \sqrt{2}$$

を得る．

（答） $\displaystyle \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1 + \sqrt{2}$

□

次の問題は，フィボナッチ数列そのものに関するものです．皆さんご存じの通り，フィボナッチ数列の一般項を具体的に表示することが可能です．

問題 1.10.

フィボナッチ数列の一般項を求めよ．

解答. $\{a_n\}$ をフィボナッチ数列とする．一方，λ, μ を 2 次方程式 $x^2 - x - 1 = 0$ の 2 解 $(1 \pm \sqrt{5})/2$ とする．このとき，問題 1.1 の解答で見たとおり，

$$a_{n+1} - \lambda a_n = \mu(a_n - \lambda a_{n-1})$$

が成り立つ．すなわち，$b_n = a_{n+1} - \lambda a_n$ $(n = 0, 1, \cdots)$ で定義される数列 $\{b_n\}$ は公比 μ，初項 $b_0 = a_1 - \lambda a_0 = 1 - \lambda$ の等比数列である．よって，

$$b_n = a_{n+1} - \lambda a_n = (1 - \lambda)\mu^n$$

を得る．以上の議論は λ と μ を入れ替えても成り立つので，

$$a_{n+1} - \mu a_n \quad = (1 - \mu)\lambda^n$$

も成立する．これら 2 式に対し，(第 2 式) $\times\lambda -$ (第 1 式) $\times\mu$ を計算することにより a_n を消去すれば，$\lambda+\mu=1$ であることより，

$$(\lambda-\mu)a_{n+1}=(1-\mu)\lambda^{n+1}-(1-\lambda)\mu^{n+1}=\lambda^{n+2}-\mu^{n+2}$$

を得る．最後に，$\lambda,\mu=(1\pm\sqrt{5})/2$ であったことを思い起こせば，以下が結論される：

$$\textbf{（答）}\qquad a_n=\frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right\}$$

□

フィボナッチ数列の各項は当然整数ですが，ここで得た一般項の表示には黄金比やその共役といった無理数が現れます．ちょっと不思議だとは思いませんか？

ところで，いま紹介したフィボナッチ数列の一般項の求め方は，行列やベクトルと言った線形代数の言葉を使って述べることができます．それを問題として出題しましょう．

問題 1.11. ――――――――――――――――――――――――――――

以下，$\{a_n\}$ をフィボナッチ数列とする．また，

$$\lambda=\frac{1+\sqrt{5}}{2},\qquad \mu=\frac{1-\sqrt{5}}{2}$$

とおく．

(1) 各 $n=0,1,\cdots$ に対し，2 次元列ベクトル $\mathbf{a}_n$ を，

$$\mathbf{a}_n=\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

により定義する．このとき，

$$\mathbf{a}_n=X\mathbf{a}_{n-1}\qquad (n=1,2,\cdots)$$

が成り立つことを示せ．ただし，X は以下で与えられる 2×2 行列である．

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

(2) さらに，

$$P=\begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix},\qquad Y=\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

としたとき，

$$P^{-1}XP=Y$$

が成り立つことを確かめよ．

(3) (1), (2) において示したことを用い，フィボナッチ数列の一般項 a_n を決定せよ．

解答．(1), (2) 略．

(3) (1) より，

$$\mathbf{a}_n = X\mathbf{a}_{n-1} = X^2\mathbf{a}_{n-2} = \cdots = X^n\mathbf{a}_0$$

が従う．一方，(2) によれば，$X = PYP^{-1}$ であるから

$$X^n = \underbrace{(PYP^{-1})(PYP^{-1})\cdots(PYP^{-1})}_{n} = PY^nP^{-1}$$

が成り立つ．また明らかに

$$Y^n = \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix}$$

である．以上をあわせると，

$$\begin{aligned}
\mathbf{a}_n = X^n\mathbf{a}_0 &= PY^nP^{-1}\mathbf{a}_0 \\
&= \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix}\cdot\frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -\mu \\ -1 & \lambda \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda^{n+1}-\mu^{n+1} & -\lambda^{n+1}\mu+\lambda\mu^{n+1} \\ \lambda^n-\mu^n & -\lambda^n\mu+\lambda\mu^n \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda^{n+1}(1-\mu)+(\lambda-1)\mu^{n+1} \\ \lambda^n(1-\mu)+(\lambda-1)\mu^n \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda^{n+2}-\mu^{n+2} \\ \lambda^{n+1}-\mu^{n+1} \end{pmatrix}
\end{aligned}$$

を得る．これより以下が従う．

（答） $$a_n = \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right\}$$

□

実は，フィボナッチ数列の一般項の求め方の背後には，正方行列の対角化や標準形といった線形代数の一般論が隠れています．しかし，それについては，線形代数の講義に譲ることにしましょう．

さて，次の問題に進みましょう．問題 1.1 に対する「不完全」な別解を，§B.1 で述べました．それを完全な解答にしようというのが，次の問題の意図です．

問題 1.12.

$\{a_n\}$ をフィボナッチ数列とする．このとき，$r_n = a_{n+1}/a_n$ $(n = 0, 1, \cdots)$ により定義される数列 $\{r_n\}$ に対し，

$$r_0 < r_2 < r_4 < \cdots ,$$
$$r_1 > r_3 > r_5 > \cdots$$

が成り立つことを示せ．また，このことを用い，数列 $\{r_n\}$ が収束することを示せ．

解答． まず，

$$r_{n+2} - r_n = \frac{a_{n+3}a_n - a_{n+2}a_{n+1}}{a_{n+2}a_n}$$

であることに注意する．右辺に現れる分数の分子を b_n とおく：

$$b_n = a_{n+3}a_n - a_{n+2}a_{n+1}.$$

このとき，

$$\begin{aligned} b_{n+1} &= a_{n+4}a_{n+1} - a_{n+3}a_{n+2} \\ &= (a_{n+3} + a_{n+2})a_{n+1} - a_{n+3}a_{n+2} \\ &= -a_{n+3}(a_{n+2} - a_{n+1}) + a_{n+2}a_{n+1} \\ &= -a_{n+3}a_n + a_{n+2}a_{n+1} \\ &= -b_n \end{aligned}$$

が成り立つ．さらに，$b_0 = a_3a_0 - a_2a_1 = 3 \cdot 1 - 2 \cdot 1 = 1$ であるから，結局

$$b_n = (-1)^n$$

であることが分かる．そしてこれより，

$$r_{2m+2} - r_{2m} > 0, \qquad r_{2m+3} - r_{2m+1} < 0 \qquad (m = 0, 1, \cdots) \tag{1.8}$$

が従う．

以上で，数列 $\{r_{2m}\}, \{r_{2m+1}\}$ がそれぞれ単調増加，単調減少であることが示された．次にこれらの数列がそれぞれ上に有界，下に有界であることを示したい．そのために，まず

$$r_{n+1} - r_n = \frac{a_{n+2}a_n - a_{n+1}^2}{a_{n+1}a_n}$$

であることに着目する．さて，右辺の分子を c_n とおこう：

$$c_n = a_{n+2}a_n - a_{n+1}^2.$$

すると，

$$\begin{aligned} c_{n+1} &= a_{n+3}a_{n+1} - a_{n+2}^2 = (a_{n+2} + a_{n+1})a_{n+1} - a_{n+2}^2 \\ &= -a_{n+2}(a_{n+2} - a_{n+1}) + a_{n+1}^2 = -a_{n+2}a_n + a_{n+1}^2 = -c_n \end{aligned}$$

が成り立つ．これと，$c_0 = a_2a_0 - a_1^2 = 2 \cdot 1 - 1^2 = 1$ であることより，

$$c_n = (-1)^n$$

を得る．以上から，とくに

$$r_{2m} < r_{2m+1} \qquad \text{かつ} \qquad r_{2m+1} > r_{2m+2} \qquad (m = 0, 1, \cdots) \tag{1.9}$$

が成り立つことが分かる．ここで，最初に示したとおり数列 $\{r_{2m+1}\}$ が単調減少であることを思い出せば，

$$r_{2m} < r_{2m+1} < r_{2m-1} < \cdots < r_1 \qquad (m = 0, 1, \cdots)$$

であることが分かる．よって，数列 $\{r_{2m}\}$ は上に有界である．数列 $\{r_{2m+1}\}$ が下に有界であることも同様にして示される．

数列 $\{r_{2m}\}, \{r_{2m+1}\}$ はそれぞれ上に有界な単調増加数列，下に有界な単調減少数列であるから，それらはともに収束する．これらふたつ数列の極限の値が一致することを示せば，元の数列 $\{r_n\}$ が収束することが証明される．ところが，前の段落で見たこと，および $a_n \to \infty$ $(n \to \infty)$ であることとより，

$$0 < r_{2m+1} - r_{2m} = \frac{c_{2m}}{a_{2m+1}a_{2m}} = \frac{1}{a_{2m+1}a_{2m}} \longrightarrow 0$$

が従う．よって，ふたつの数列 $\{r_{2m}\}, \{r_{2m+1}\}$ は確かに共通の極限に収束する．□

この問題の解答において重要な役割を果たすのが，不等式 (1.8), (1.9) です．数列 $\{r_n\}$ 自身は単調ではありません（実際，(1.9) によれば，数列 $\{r_n\}$ は増減を交互に繰り返します）．したがって，数列 $\{r_n\}$ に対し有界単調数列の収束定理を直接適用することはできません．しかし，番号が偶数の項だけを取り出すと単調増加に，また奇数の項だけを取り出すと単調減少になるということが，(1.8) より分かります．実際，この問題の数列 $\{r_n\}$ の最初の数項を計算し，折れ線グラフとしてプロットしたのが，図 1.6 です．不等式 (1.9) は一種のヒントとして問題文の中で

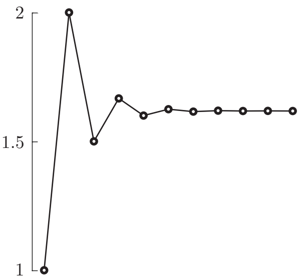


図 1.6:

与えられています．しかし，数列 $\{r_n\}$ の最初のいくつかの項を試しに計算してみれば，(1.8) や (1.9) が成り立つことは容易に予想できるはずです．[20] その後解答を完成させるにはさしたる手間は必要ありません．

次のふたつの問題も有界単調数列に対する収束定理を適用して解くことが可能です．

問題 1.13. ――――――――――――――――――――――――――――――

$a_0 = 1$, および

$$a_{n+1} = 1 + \frac{1}{1+a_n} \qquad (n = 0, 1, \cdots)$$

により定義される数列 $\{a_n\}$ に対し，その極限が $\sqrt{2}$ であることを証明せよ．

―――――――――――――――――――――――――――――――――――

解答． まず，

$$a_0 < a_2 < \cdots < \sqrt{2} < \cdots < a_3 < a_1 \tag{1.10}$$

が成り立つことを示そう．与えられた漸化式から

$$a_{n+1}^2 - 2 = -\frac{a_n^2 - 2}{(1+a_n)^2}$$

であることが分かる．とくに $a_n^2 - 2$ は交互に符号を変える．しかも，$a_0^2 - 2 < 0$ であるから，$a_{2m} < \sqrt{2} < a_{2m+1}$ である．一方，与えられた漸化式から

$$a_{n+2} - a_n = -\frac{2(a_n^2 - 2)}{3 + 2a_n}$$

であることが分かる．右辺の分母はつねに正であるから，結局 $\{a_{2m}\}$ は単調増加であり，$\{a_{2m+1}\}$ は単調減少であることが従う．以上で，(1.10) が示された．

[20] 与えられた数列に対し，試しにその最初の数項を計算してみる，ということはとても大事なことです．そういった試みは，数学における一種の「実験」と考えられます．皆さんは是非，自らの手で実験してみて下さい．そこから多くのことが読み取れるはずです．

とくに，有界単調数列の収束性定理によれば，$\{a_{2m}\}$, $\{a_{2m+1}\}$ はともに収束する．それらの極限がともに $\sqrt{2}$ であることを確かめよう．与えられた漸化式から，

$$a_{n+2} = \frac{4+3a_n}{3+2a_n} \tag{1.11}$$

を得る．とくに，$n = 2m \to \infty$ とすることにより，$\alpha = \lim_{m\to\infty} a_{2m}$ に対し，$\alpha = (4+3\alpha)/(3+2\alpha)$, すなわち，$\alpha = \pm\sqrt{2}$ であることが従う．ところが，$a_n > 0$ $(n = 0, 1, \cdots)$ より $\alpha \geq 0$ でなければならない．よって，数列 $\{a_{2m}\}$ の極限が $\sqrt{2}$ であることが分かる．(1.11) において，$n = 2m+1 \to \infty$ とすることにより，$\{a_{2m+1}\}$ の極限も $\sqrt{2}$ であることが結論される． □

この問題の数列の各項を順次計算すると，

$$1, \quad 1+\frac{1}{2}, \quad 1+\frac{1}{2+\frac{1}{2}}, \quad 1+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}, \cdots$$

と成ります．すなわち，$\sqrt{2}$ は

$$\sqrt{2} = 1 + \frac{1}{2+\frac{1}{2+\frac{1}{2+\cdots}}}$$

という連分数表示を持つことが分かります．ということは，また $\sqrt{2}$ と 1 に対する互除法が無限に続くことですから，$\sqrt{2}$ が無理数であることが，このことからも分かります．

問題 1.14. ―――――――――――――――――――――――――――――

正の実数 x に対し正の実数を対応させる関数 $f(x)$ を，

$$f(x) = \frac{3x+2}{x+1}$$

により定義する．さらに，正の定数 α が任意に与えられたとする．このとき，

$$a_0 = \alpha, \qquad a_{n+1} = f(a_n) \quad (n = 0, 1\cdots)$$

で定義された数列 $\{a_n\}$ は必ず収束し，しかもその極限は定数 α の取り方に依らないことを証明せよ．

―――――――――――――――――――――――――――――――――――

解答. まず，数列 $\{a_n\}$ が収束することを仮定し，その極限を求める．いま，数列 $\{a_n\}$ の極限を λ と書くことにする．すると，漸化式 $a_{n+1} = f(a_n)$ において $n \to \infty$ とすることにより，$\lambda = f(\lambda)$ を得る．これより，極限 λ は，x に対する 2 次方程式

$$x^2 - 2x - 2 = 0$$

の解でなければならないことが分かる．さらに，$a_n > 0$ より $\lambda \geq 0$ であるから，

$$\lambda = 1 + \sqrt{3}$$

であることが従う．とくに，極限 λ の値が初項 a_0 の値 α に依らないことも，以上の議論から従う．

次に，数列 $\{a_n\}$ が必ず収束することを証明したい．以下，

$$q(x) = x^2 - 2x - 2$$

とする（λ は 2 次方程式 $q(x) = 0$ の正の解であったことを思い出そう）．このとき，

$$q(f(x)) = \frac{q(x)}{(x+1)^2} \qquad (x > 0)$$

であることより以下が従う：

$$\begin{cases} q(x) > 0 & \Longrightarrow \quad q(f(x)) > 0, \\ q(x) < 0 & \Longrightarrow \quad q(f(x)) < 0. \end{cases} \tag{1.12}$$

一方，

$$x - f(x) = \frac{q(x)}{x+1}$$

であることも明かである．ここで，$x > 0$ であり，とくに上式右辺に現れる分数の分母は常に正であることに注意すれば，

$$\begin{cases} q(x) > 0 & \Longrightarrow \quad f(x) < x, \\ q(x) < 0 & \Longrightarrow \quad f(x) > x \end{cases} \tag{1.13}$$

であることが分かる．

さらに，$\lambda > 0$ が 2 次方程式 $q(x) = 0$ のふたつの解のひとつであり，もうひとつの解 $1 - \sqrt{3}$ は負であることに注意すれば，正の実数 x に対し以下が成り立つことは明かである：

$$\begin{cases} x > \lambda & \Longleftrightarrow \quad q(x) > 0, \\ x < \lambda & \Longleftrightarrow \quad q(x) < 0. \end{cases} \tag{1.14}$$

以上 (1.12), (1.13), (1.14) から直ちに，数列 $\{a_n\}$ に対し以下が成り立つことが従う：

$$\begin{cases} a_n > \lambda & \Longrightarrow \quad a_{n+1} < a_n \quad \text{かつ} \quad a_{n+1} > \lambda, \\ a_n < \lambda & \Longrightarrow \quad a_{n+1} > a_n \quad \text{かつ} \quad a_{n+1} < \lambda. \end{cases}$$

そしてこのことから，$\alpha > \lambda$ であるときには数列 $\{a_n\}$ は単調減少かつ下に有界であることが，また $\alpha < \lambda$ であるときには単調増加かつ上に有界であることが分か

る．一方，$\alpha = \lambda$ であるときには，$a_n = \lambda\ (n = 0, 1, \cdots)$ であり，数列 $\{a_n\}$ が収束することは自明である．以上で，いかなる場合にも数列 $\{a_n\}$ は収束ししかもその極限の値は初項 $\alpha > 0$ の値に依らないことが証明された． $\square$

実は，この問題とほとんど同じやり方で，フィボナッチ数列の隣り合う 2 項の比が黄金比に収束することを示すことが可能です．是非，皆さん試してみてください．ところで，この問題 1.14 に現れる関数は 1 **次分数変換**と呼ばれるものです．1 次分数変換は，関数論や双曲幾何などといった数学の諸分野に登場する，興味深い対象です．皆さんも，近い将来どこかでそれにまた出会うことがあるのではないかと思います．

黄金比が無理数であることの幾何学的な（しかし本文中で述べたのとは別の）証明を与えることを目的としたのが，次の問題です．

問題 1.15. ――――――――――――――――――――――――――――――

(1) 1 辺の長さが 1 の正 5 角形の対角線の長さは，黄金比 $(1+\sqrt{5})/2$ に等しいことを示せ（図 1.7 (a)）．

(2) 1 辺の長さが 1 の正 5 角形 $\mathrm{A_0B_0B_1C_1C_0}$ を考える（図 1.7 (b)）．その 2 本の対角線 $\mathrm{B_0C_1}$, $\mathrm{C_0B_1}$ の交点を $\mathrm{A_1}$ とする．さらに，辺 $\mathrm{B_0B_1}$ の延長上に点 $\mathrm{B_2}$ を，辺 $\mathrm{C_0C_1}$ の延長上に点 $\mathrm{C_2}$ を，$\mathrm{A_1B_1B_2C_2C_1}$ が正 5 角形になるようとる．この操作を繰り返すことにより，点 $\mathrm{A_2}$, $\mathrm{B_3}$, $\mathrm{C_3}$, $\mathrm{A_3}$, $\mathrm{B_4}$, $\mathrm{C_4}$, $\cdots$ を定義し，$a_n = \mathrm{B}_n\mathrm{C}_n$ $(n = 0, 1, \cdots)$ とおく．このとき，

$$a_n = a_{n+1} + a_{n+2} \qquad (n = 0, 1, \cdots)$$

が成り立つことを示せ．

(3) 以上のことから，黄金比が無理数であることを（互除法を使うことなく）証明せよ．

――――――――――――――――――――――――――――――――――――

解答． (1) 設問 (2) と同じく，$\mathrm{A_0B_0B_1C_1C_0}$ を 1 辺の長さが 1 の正 5 角形とする．その対角線の長さを λ で表そう．そして，やはり設問 (2) と同様，2 本の対角線 $\mathrm{B_0C_1}$, $\mathrm{C_0B_1}$ の交点を $\mathrm{A_1}$ とする．容易に分かるように，$\angle\mathrm{B_0B_1A_1} = \angle\mathrm{B_0A_1B_1} = 2\pi/5$ であるから，3 角形 $\mathrm{B_0B_1A_1}$ は $\mathrm{B_0B_1} = \mathrm{B_0A_1}$ なる 2 等辺 3 角形である．よって，$\mathrm{B_0A_1} = 1$, $\mathrm{C_1A_1} = 1 - \lambda$ が従う．一方，ふたつの 2 等辺 3 角形 $\mathrm{A_1C_0B_0}$, $\mathrm{A_1B_1C_1}$ が相似であることより，$\mathrm{B_0C_0 : B_1C_1 = B_0A_1 : C_1A_1}$ を得る．以上より，$\lambda : 1 = 1 : \lambda - 1$，すなわち，$\lambda^2 - \lambda - 1 = 0$ が従う．よって，λ が黄金比 $(1+\sqrt{5})/2$ に等しいことが結論される．

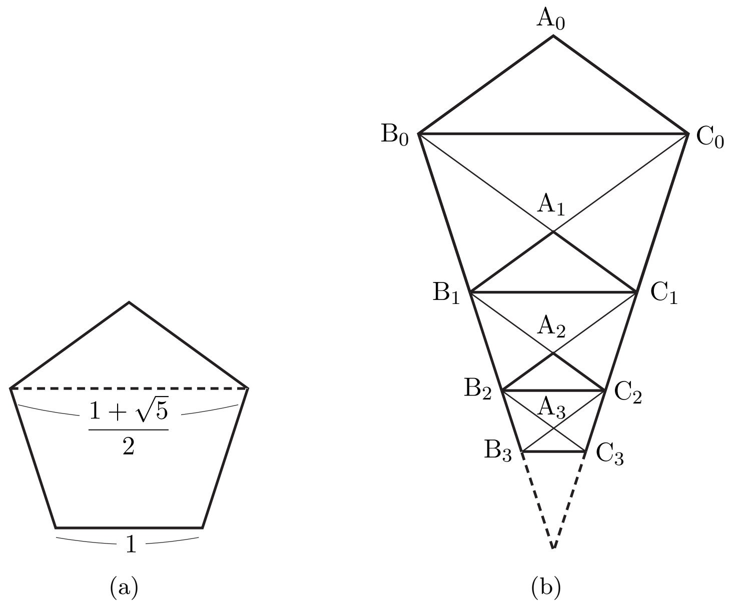


図 1.7:

(2) 3 角形 $\mathrm{B}_n\mathrm{B}_{n+1}\mathrm{A}_{n+1}$ が，$\mathrm{B}_n\mathrm{B}_{n+1} = \mathrm{B}_n\mathrm{A}_{n+1}$ なる 2 等辺 3 角形であることから，

$$\begin{aligned} a_n &= \mathrm{B}_n\mathrm{C}_n = \mathrm{B}_n\mathrm{C}_{n+1} = \mathrm{B}_n\mathrm{A}_{n+1} + \mathrm{A}_{n+1}\mathrm{C}_{n+1} \\ &= \mathrm{B}_n\mathrm{B}_{n+1} + \mathrm{B}_{n+2}\mathrm{C}_{n+2} = \mathrm{B}_{n+1}\mathrm{C}_{n+1} + \mathrm{B}_{n+2}\mathrm{C}_{n+2} = a_{n+1} + a_{n+2} \end{aligned}$$

が成り立つことが分かる．

(3) まず，$a_0 = \lambda$, $a_1 = 1$ であることに注意しよう．また，数列 $\{a_n\}$ に対し，$a_0 > a_1 > \cdots$ かつ $a_n > 0\ (n = 1, 2, \cdots)$ であることが，図より容易に分かる．以下，背理法で議論を進めて行く．そこで，黄金比 λ が有理数であると仮定する．すると，互いに素なふたつの正の整数 p, q を用い，$\lambda = p/q$ と書ける．すると，(2) において示した等式より，数列 $\{a_n\}$ の各項 a_n が分子が整数，分母が q の分数の形に書けることが分かる．その分子を p_n としよう．数列 $\{a_n\}$ は狭義単調減少，かつその項はすべて正であるから，正の整数を項とする数列 $\{p_n\}$ も狭義単調数列であることになる．しかし，これはあり得ないことである．[21] すなわち，矛盾が生じた．以上から，黄金比 λ が無理数であることが結論された． □

ところで，「ピタゴラスの定理」などでその名を知られたピタゴラスは，紀元前のギリシャで活躍した数学者・哲学者です．彼はピタゴラス学派と呼ばれる集団の

[21] あるいは，補題 1.8 を用いてもよい．

長でもありました．そのピタゴラス学派は，正 5 角形の 5 本の対角線が形づくる星形を彼らの紋章としていました．ところで，一方，今日では想像しにくいことではありますが，彼らにとって無理数とは決してその存在を認めることの出来ない対象であったとのことです．宗教的とも言える彼ら固有の哲学や自然観がその背景にあってのことです．しかし，皮肉なことに，彼らの紋章の中にも，黄金比という無理数が含まれていたのでした．

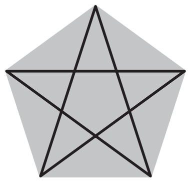


図 1.8:

問題 1.16. ――――――――――――――――――――――――――――――

$k > 0$, $\lambda > 1$ を定数とする．このとき，以下の数列の $n \to \infty$ における発散の速さを比較せよ：

$$n^k, \qquad \lambda^n.$$

――――――――――――――――――――――――――――――――――――

略解. $n^k/\lambda^n \to 0$ $(n \to \infty)$ であるから，数列 n^k より早く数列 λ^n は発散する． □

ネピア数を $\left(1+\dfrac{1}{n}\right)^n$ の $n \to \infty$ での極限として定義することがあります（皆さんは，むしろこの定義の方になじみがあるのではないかと思います）．§B.4 で与えたネピア数の定義と，このもうひとつの定義が一致することを確かめるのが，次の問題です．

問題 1.17. ――――――――――――――――――――――――――――――

数列

$$\left(1+\frac{1}{n}\right)^n$$

は，(§B.4 で定義した）ネピア数 e に収束することを示せ．

――――――――――――――――――――――――――――――――――――

解答. $n = 1, 2, \cdots$ に対し，

$$a_n = \left(1+\frac{1}{n}\right)^n, \qquad b_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}$$

とおく．$b_n \to e$ $(n \to \infty)$ であった．したがって，$b_n - a_n \to 0$ $(n \to \infty)$ を示せばよい．

まず，2 項定理[22] より

$$a_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k}$$

を得る．この等式の右辺の各項に対しては，以下が成り立つ：

$$\begin{aligned}\binom{n}{k}\frac{1}{n^k} &= \frac{n!}{k!(n-k)!}\frac{1}{n^k} \\ &= \frac{1}{k!}\cdot\frac{n-k+1}{n}\frac{n-k+2}{n}\cdots\frac{n}{n} \\ &= \frac{1}{k!}\cdot\left(1-\frac{k-1}{n}\right)\left(1-\frac{k-2}{n}\right)\cdots\left(1-\frac{0}{n}\right) \leq \frac{1}{k!}.\end{aligned}$$

以上より，

$$a_n \leq b_n \qquad (n = 1, 2, \cdots)$$

が従う．

一方，

$$\begin{aligned}b_n - a_n &= \sum_{k=2}^{n} \frac{1}{k!}\left\{1 - \left(1-\frac{k-1}{n}\right)\left(1-\frac{k-2}{n}\right)\cdots\left(1-\frac{0}{n}\right)\right\} \\ &\leq \sum_{k=2}^{n} \frac{1}{k!}\left\{1 - \left(1-\frac{k-1}{n}\right)^k\right\}\end{aligned}$$

である．ここで，簡単のため，$x = 1 - \dfrac{k-1}{n}$ とおくと，

$$\begin{aligned}1 - \left(1-\frac{k-1}{n}\right)^k &\leq 1 - x^k = (1-x)(1+x+\cdots+x^{k-1}) \\ &= \frac{k-1}{n}(1+x+\cdots+x^{k-1}) < \frac{k(k-1)}{n}\end{aligned}$$

が成り立つ．ただし，最後の不等式は，$x < 1$ より従う．以上から，

$$b_n - a_n \leq \sum_{k=2}^{n} \frac{1}{k!}\frac{k(k-1)}{n} = \frac{1}{n}\sum_{j=0}^{n-2}\frac{1}{j!} \leq \frac{e}{n} \longrightarrow 0 \qquad (n \to \infty)$$

を得る．以上で，$b_n - a_n \to 0$ の証明が終わった．　□

[22]公式

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n-1}x^{n-1} + x^n$$

を 2 **項定理**と呼ぶ．

次の問題もネピア数に関するものです．このノートの中では，この問題は難問の部類に入るかも知れません．

問題 1.18. ――――――――――――――――――――――――――――

ネピア数 e が無理数であることを証明せよ．

――――――――――――――――――――――――――――――――――

解答. 背理法で証明する．ネピア数 e が有理数であるとしよう．すると $e = p/q$（ただし p，q は互いに素な整数，しかも $q > 0$）と書くことができる．このとき，e の定義より，

$$\begin{aligned} \frac{p}{q} &= 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots \\ &= \left\{1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{q!}\right\} + \frac{1}{q!}\left\{\frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots\right\}. \end{aligned}$$

この両辺に $q!$ を乗じることにより，

$$p(q-1)! = M + r,$$

ただし，

$$\begin{aligned} M &= q! + (2 \cdot \cdots \cdot q) + (3 \cdot \cdots \cdot q) + \cdots + q + 1, \\ r &= \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots, \end{aligned}$$

を得る．ここで $p(q-1)!$ および M は明らかに整数であるから，$r = p(q-1)! - M$ も整数でなければならない．ところが一方，

$$\begin{aligned} 0 < r &= \frac{1}{q+1}\left\{1 + \frac{1}{q+2} + \frac{1}{(q+2)(q+3)} + \cdots\right\} \\ &< \frac{1}{q+1}\left\{1 + \frac{1}{q+2} + \frac{1}{(q+2)^2} + \cdots\right\} \\ &= \frac{q+2}{(q+1)^2} < 1. \end{aligned}$$

したがって r は整数ではあり得ないこととなり矛盾が生じた．以上で e が無理数であることが証明された．　□

問題 1.19. ――――――――――――――――――――――――――――

数列 $\{a_n\}$ を以下で定義する：

$$a_n = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \qquad (n = 1, 2, \cdots).$$

(1) 以下の不等式を示せ（**ヒント**：図 1.9 を参考にしながら，定積分を利用せよ）：

$$\log(n+1) < a_n < 1 + \log n \qquad (n \geq 2).$$

(2) $n \to \infty$ としたとき，数列 $\{a_n\}$ は発散することを示せ．
(3) $n \to \infty$ としたとき，

$$a_n - \log n = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} - \log n$$

は収束することを証明せよ．

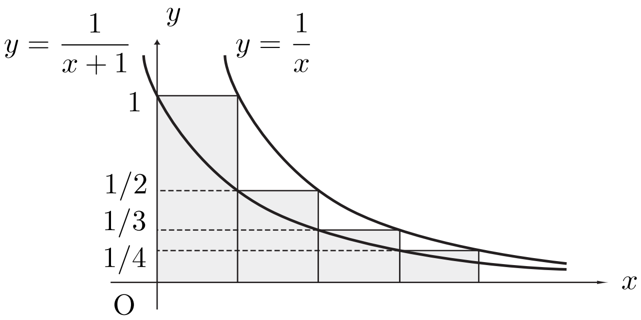


図 1.9:

解答. (1) 関数 $y = f(x)$ $(x \geq 0)$ を，$f(x) = \dfrac{1}{[x]+1}$ により定義する．ただし，ここで実数 x に対し，$[x]$ は x を越えない最大の整数を表す．

$$a_n = \int_0^n f(x)\,dx$$

であることに注意しよう（とくに，図 1.9 における「棒グラフ」の面積が a_n である）．一方，

$$\frac{1}{x+1} \leq f(x) \quad (x \geq 0) \qquad \text{かつ} \qquad f(x) < \frac{1}{x} \quad (x \geq 1)$$

が成り立つことより，

$$\log(n+1) = \int_0^n \frac{dx}{x+1} < \int_0^n f(x)\,dx < 1 + \int_1^n \frac{dx}{x} = 1 + \log n,$$

すなわち，問題の不等式を得る．

(2) (1) において得た不等式 $a_n > \log(n+1)$ において，$n \to \infty$ とすることにより，ただちに，$a_n \to \infty$ $(n \to \infty)$ を得る．

(3) (1) において行った考察より，

$$\int_1^n \left\{\frac{1}{x} - f(x)\right\} dx = \log n - (a_n - 1)$$

が n に関し単調増加であることが分かる（実際，この量は，関数 $y = 1/x$ のグラフ，関数 $y = f(x)$ のグラフ，およびふたつの直線 $x = 1$, $x = n$ に挟まれた部分の面積である）．したがって，$a_n - \log n$ は単調減少である．一方，(1) より，

$$a_n - \log n > \log(n+1) - \log n \geq 0 \qquad (n \geq 1)$$

であるから，$a_n - \log n$ は下から有界である．よって $n \to \infty$ としたとき，$a_n - \log n$ が収束することが結論される． □

解説. (2) で証明した事実はこの後もしばしば必要となるはずです．是非，覚えておいてください．なお，$s > 1$ としたとき，

$$b_n = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} \qquad (n = 1, 2, \cdots)$$

により定義される数列は必ず収束します．あるいは言い換えれば，無限和

$$\frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

は有限の値に収束します．どうしてそうなるか，各自考えてみて下さい．

ところで，(3) に現れる極限 $\lim_{n\to\infty} (a_n - \log n)$ として得られる数は**オイラー定数**と呼ばれます．この数が有理数かあるいは無理数かも分かっていないとのことです．

第 2 章

連続関数

A　中間値の定理とその応用

A.1　関数・グラフ・中間値の定理

第 1 章で主役を勤めたのは数列でした．第 2 章では関数がそれにとって代わります．関数の例として，

$$y = x^3 - 2x + 1 \tag{2.1}$$

を考えましょう．変数 x の値を決めれば，変数 y の値が確定します．このようなとき，変数 y は変数 x の**関数**である，といった言い方をします．また，変数 x は**独立変数**と呼ばれます．その値を他から独立して自由に決めることが出来るからです．一方，変数 y の値は変数 x の値に従属して決まることから，**従属変数**と呼ばれます．この章では，独立変数のとる値は必ず実数であると仮定します．ただし，独立変数は必ずしもすべての実数をその値としてとり得るわけではありません．例えば，$y = \log x$ を考えて見てください．この場合には，$x > 0$ という制約が課されます．この章では，独立変数の動く範囲は，実直線 $\mathbb{R}$ 全体，[1] ないし

$$\begin{aligned}
(a,b) &= \{x \in \mathbb{R} : a < x < b\}, & (a,b] &= \{x \in \mathbb{R} : a < x \leq b\},\\
[a,b) &= \{x \in \mathbb{R} : a \leq x < b\}, & [a,b] &= \{x \in \mathbb{R} : a \leq x \leq b\},\\
(a,\infty) &= \{x \in \mathbb{R} : a < x\}, & (-\infty,b) &= \{x \in \mathbb{R} : x < b\},\\
[a,\infty) &= \{x \in \mathbb{R} : a \leq x\}, & (-\infty,b] &= \{x \in \mathbb{R} : x \leq b\}
\end{aligned}$$

といった**区間**であると仮定します．[2] 独立変数の動く範囲を，その関数の**定義域**と呼びます．一方，この章においては，従属変数のとる値は必ず実数であるとします．

[1] 記号 $\mathbb{R}$ は，実数全部を集めて得られる集合を表します．したがって，幾何学的には，それは実直線を表すと考えられます．

[2] これらの区間のうち，(a,b), $(-\infty,b)$, (a,∞) を**開**区間と，また，$[a,b]$, $(-\infty,b]$, $[a,\infty)$ を**閉**区間と呼びます．一方，(a,b), $(a,b]$, $[a,b)$, $[a,b]$ を**有界**区間と呼びます．

再度 (2.1) を例にとりましょう．$y = x^3 - 2x + 1$ という表示，あるいはその右辺 $x^3 - 2x + 1$ を**関数**と呼んだりすることもあります．変数 x を独立変数とする一般の関数について議論したいときには，それらを $f(x)$, $g(x)$ などと表記することがあります．関数 $f(x)$ の定義域が I であるとき，関数 $f(x)$ は I 上で**定義された**関数である，といいます．そして，$f(x)$ $(x \in I)$ と書きます．また，I が例えば $[a, b]$ の形の区間のときには，$f(x)$ $(x \in I)$ と書く代わりに，$f(x)$ $(a \leq x \leq b)$ と書くこともしばしばです．

この他に関数の例としては，例えば，指数関数 e^x や三角関数 $\sin x, \cos x, \tan x$ などが考えられます．また，ふたつの関数 $f(x), g(x)$ に対し，それらの和 $f(x)+g(x)$, 差 $f(x) - g(x)$，積 $f(x)g(x)$ が定義されます．また，分母がゼロにならない範囲では，商 $f(x)/g(x)$ を考えることも可能です．例えば，$\sin x/e^x$ は，実直線全体で定義された関数です．

さて，区間 I 上で定義された関数 $f(x)$ が与えられたとしましょう．各 $x \in I$ に対し，実数 $f(x)$ が定まります．そしてそれに応じて，(x, y)-平面内に点 $(x, f(x))$ が定まります．次いで x を区間 I 内すべてに渡って動かしましょう．すると，点 $(x, f(x))$ も平面内を動き回ります．そして，その軌跡として得られるのが，関数 $f(x)$ の**グラフ**です．もしそのグラフが切れ目のない，いわゆる連続な曲線になるならば，関数 $f(x)$ は**連続**であると言うことにしましょう．[3]

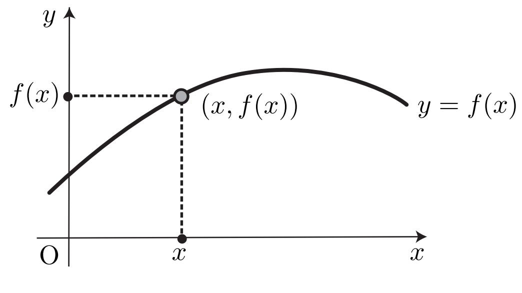
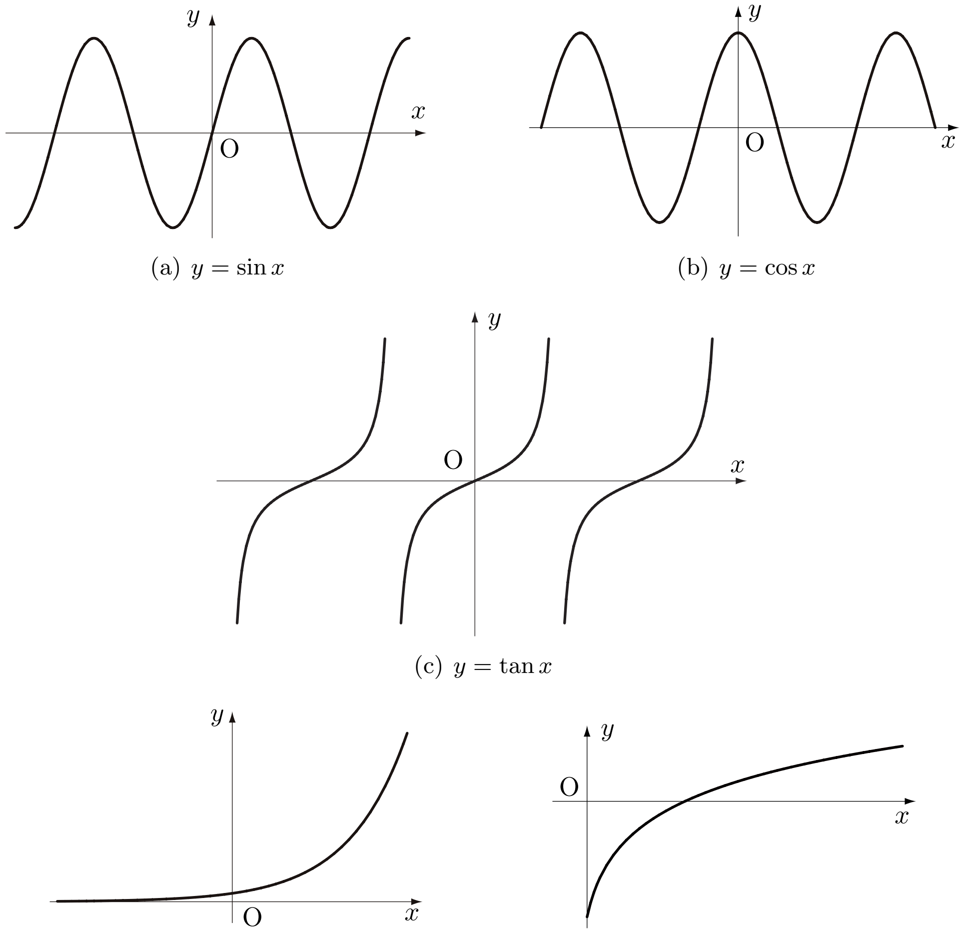


図 2.1: 関数のグラフ

図 2.2 に，典型的な関数のグラフを例として挙げておきます．ただし，関数 $y = \tan x$ の定義域は区間ではなく，集合 $\mathbb{R} \setminus \{(n + \frac{1}{2})\pi : n \in \mathbb{Z}\}$ [4] です．そこに挙げた関数のいずれもが連続です．連続関数のその他の例としては，多項式で定義される関数が挙げられます．また，区間 I 上で定義された関数 $f(x), g(x)$ がともに連続であるならば，それらの和 $f(x) + g(x)$, 差 $f(x) - g(x)$, 積 $f(x)g(x)$ も連続関数です．さらに，分母がゼロにならなければ，商 $f(x)/g(x)$ も連続です．一方，不連

[3] ここで，述べた連続関数の定義は実は仮のものです．後に本当の定義を与えます．
[4] ふたつの集合 X, Y に対し，$X \setminus Y$ は**差集合** $\{x \in X : x \notin Y\}$ を表します．



(a) $y = \sin x$　(b) $y = \cos x$

(c) $y = \tan x$

(d) $y = \exp x$　(e) $y = \log x$

図 2.2: グラフの例

続な関数としては，例えば，

$$f(x) = \begin{cases} x+1 & (x < 0) \\ x-1 & (x \geq 0) \end{cases} \tag{2.2}$$

が挙げられます．実際，そのグラフは $x = 0$ において不連続です（図 2.3）．

以降，$f(x)$ は連続であるとします．さらに関数 $f(x)$ の定義域 I が有界閉区間 $[a, b]$ であり，しかも両端点 a, b における関数 $f(x)$ の値が，不等式

$$f(a) \cdot f(b) < 0 \tag{2.3}$$

を満たすと仮定しましょう．このとき，$f(a)$ と $f(b)$ は異なる符号を持たなければなりません．もう少し詳しく述べると，$f(a) < 0$ かつ $f(b) > 0$ であるか，あるいは $f(a) > 0$ かつ $f(b) < 0$ であるかのどちらかです．ここで関数 $y = f(x)$

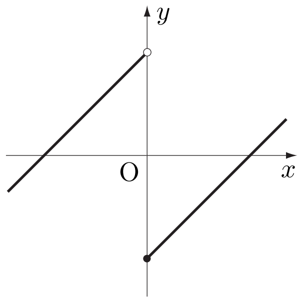
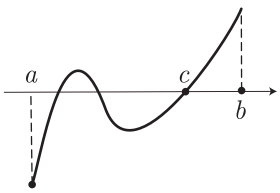
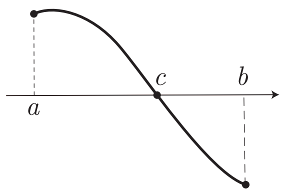


図 2.3: 不連続関数の例

$(a \leq x \leq b)$ のグラフを左側から描き始めると，第 1 の場合には，x-軸の下側から始まり x-軸の上側で終わり，一方，第 2 の場合には逆に x-軸の上側から始まり x-軸の下側で終わります．ところで，関数 $f(x)$ は連続であると仮定していたわけですから，そのグラフは切れ目のない連続曲線です．したがっていずれの場合にも，関数 $f(x)$ のグラフは少なくとも 1 点で x-軸と交わるはずです．そしてその交点の x-座標を c（ただし $a < c < b$）とすれば，$f(c) = 0$ が成り立つのは言うまでもありません．また (2.3) の代わりに $f(a) \cdot f(b) = 0$ が成り立つときには，$f(a)$, $f(b)$ の少なくとも一方はゼロでなければなりません．したがってこの場合には，a, b のうち適当な方をとり c とすれば，やはり $f(c) = 0$ が成り立ちます．



(a) $f(a) < 0$ かつ $f(b) > 0$ の場合



(b) $f(a) > 0$ かつ $f(b) < 0$ の場合

図 2.4: 中間値の定理

以上の議論をまとめたのが，次に述べる中間値の定理です．

定理 2.1 (中間値の定理).
関数 $f(x)$ $(a \leq x \leq b)$ が連続，かつ

$$f(a) \cdot f(b) \leq 0$$

を満たすならば，

$$f(c) = 0$$

なる c（ただし，$a \leq c \leq b$）が少なくともひとつ存在する．

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

言い換えると，この定理の仮定のもと，方程式 $f(x) = 0$（ただし $a \leq x \leq b$）が，少なくとも解をひとつ持つことが保証されたわけです．

A.2　問題ふたつ

皆さんに中間値の定理の面白さを知ってもらうために，それを使って解くことができるパズルふうの問題をふたつほど出すことにします．ただし，それらに対し厳密な解答を与えるには，本書の範囲を超えた知識が若干必要になるかも知れないことを，あらかじめお断りしておきます．ともかく，解答を読み始める前に，是非とも皆さん各自で考えてみることを強くお勧めします．

問題 2.2. ————————————————————————————

平面内に有界な領域がふたつ与えられたとき，その平面内に直線を 1 本適当に引くことにより，それら領域のおのおのを面積の等しい 2 つの部分に分けられることを示せ．

————————————————————————————————

ここでは，平面内の**領域**とは，平面内の「図形」であると考えて下さい．[5] 例えば，円板（すなわち，円の内部）や多角形（の内部）は領域です．ただし，領域は「つながって」いなければならないとします．例えば，(重なりのない）ふたつの円板をあわせたものは，領域とは考えません．とくに,「広がりが有限」の領域を，**有界**領域と呼びます．円板など先に挙げた領域の例はすべて有界です．一方，有界でない領域の例としては，第 1 象限 $\{(x, y) : x, y > 0\}$ などが挙げられます．

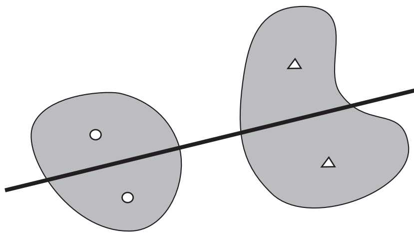


図 2.5: ケーキの問題

この問題は，次の様に例えることが出来るかも知れません．いま，テーブルの上にケーキが 2 個のっているとします．例えば，一方はチョコレートケーキ，もう一

[5] より厳密な定義もありますが，この問題を考える限り，ここで述べた直観的な定義で十分なはずです．

方はチーズケーキということにしましょう．それをふたりの人が仲良く分けあって食べることになりました．このふたりは公平を好む質の人たちで，ふたつのケーキのおのおのを同じ分量だけ食べたいそうです．そこで，それらふたつのケーキを，ナイフを使って等分することにしました．しかも，この人たちは随分と無精らしく，ケーキをひとつひとつ等分するのがおっくうとのこと．でも，大丈夫，もしながーいナイフがあれば，願ったとおり，一度だけそれを使うことにより，ふたつのケーキを同時に等分することが可能であるということを，上の問題は述べています．

問題 2.3. ——————————————————————————————

Ａ駅からＢ駅へ至るまっすぐな線路がある．そしてその線路の上を走る１台の電車の床の上に，下端がちょうつがいで床に固定された棒がついていて，それが電車の加速・減速に応じて前後にのみに揺れ動くようになっているとしよう．さて，その電車がある時刻にＡ駅を出発し，一定時間後Ｂ駅に到着するとする．ただしその間電車は加速・減速を何度でも繰り返すことが出来るし，さらには途中で停止したり，あるいはバックしたりする可能性もあるものとする．このとき，もし出発時に棒を適当な角度に立てておけば，到着時に棒がまっすぐ立っているように出来ることを示せ．ただし電車の運行はあらかじめ決まっていて，それに応じて出発時の棒の角度を選べるものとする．また仮に棒が途中で倒れた場合には，床の表面でバウンドしたりしないでそのまま倒れ続けているものとする．

————————————————————————————————————

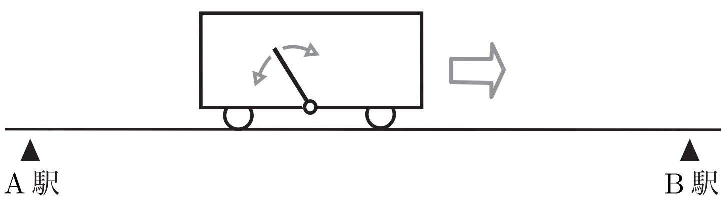


図 2.6: 電車の問題

A.3 ケーキの問題解説

さて，問題 2.2 の解説を始めることにしましょう．ところで，この問題を初めて聞いた人のなかには，一体何から考え始めたらいいのか，全く見当がつかないと感じた人も少なくないと思います．そこで，そういった人達へのヒントを．一般に，問題が難しすぎて何をやったらいいの分からないような場合には，その問題自体のかわりに， それを単純化した問題を手始めに考えることが，そもそもの問題の解決の大きな助けになることがよくあります．実は，問題 2.2 はこの考え方が有効な

典型例です．では，どうしたら問題 2.2 を単純化できるでしょうか．そもそも問題は，平面内に与えられたふたつの領域をいっぺんに等分せよ，というものだったわけです．しかし，もし仮に領域がひとつだけ与えられていたら話は随分と簡単になるのではないでしょうか．「平面内に与えられたひとつの有界領域を，その平面内に直線を 1 本引くことにより，面積の等しい 2 つの部分に分割せよ」という，より易しい問題を，問題 2.2 を解くための準備として，最初に考えようというわけです．もちろん，これが出来なければ最初の問題 2.2 が解けるわけがありません．それでは，仮にこの単純化された問題，すなわち，平面内に与えられたひとつの有界領域を 1 本の直線により等分する問題が肯定的に解けたとして，それをどうやって本来の問題 2.2 の解答に活かすことが出来るでしょうか．平面内に有界領域が 2 つ与えられた場合には，上述の単純化された問題に対する解答を，ふたつの領域のうちの一方 — これを領域 Ω_1 と呼ぶことにしましょう — に適用することにより，その領域 Ω_1 を等分する直線が少なくとも 1 本存在することが分かるはずです．したがって問題 2.2 の解答を完成させるためには，領域 Ω_1 を等分する直線のなかに，第 2 の領域 — これを Ω_2 と呼ぶことにします — をも等分するものがあることが言えればよいわけです．そしてもちろんそのためには，第 1 の領域 Ω_1 を等分する直線が沢山あったほうが有利であることは言うまでもないことです．以上のように考えると，問題 2.2 を解くための準備として，次の問題を考えたらよいのではないか，という結論にたどり着きます．

問題 2.4. ————————————————————————

平面内に有界領域がひとつ与えられたとき，それを面積の等しい 2 つの部分に分ける直線はどれくらいたくさんあるかを，記述せよ．

————————————————————————————

まず問題 2.4 に対する答えを述べることにします．[6] いま，Ω で平面内の任意の有界領域を表すことにしましょう．一方 ℓ を，その平面内の原点 O を通る任意の直線とします．このとき，問題 2.4 に対するひとつの解答は次のようなものです．

> 直線 ℓ と平行で，しかも領域 Ω を面積の等しいふたつの部分に分割する直線が必ず存在する．しかも各 ℓ に対し，このような直線はただ 1 本しか存在しない．

言い換えれば，領域 Ω を等分する直線が，原点を通過する直線と「同じだけ沢山」存在する，ということになるでしょうか．

さて，いま述べたことを証明するために，平面内に直交座標系 (u, v) をとります．ただし，u 軸が直線 ℓ と平行になるようにして下さい．すると，直線 ℓ と同

[6] こういったある意味曖昧な問題に対する答えは，複数あるのが普通です．実際，問題 2.4 に対する答えもいろいろあると思います．以下で述べるものは，単にそのうちのひとつだと思って下さい．

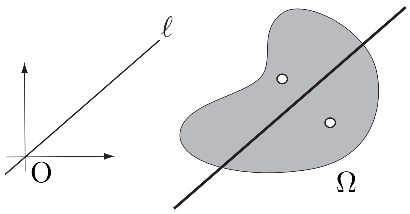


図 2.7: ひとつだけのケーキ

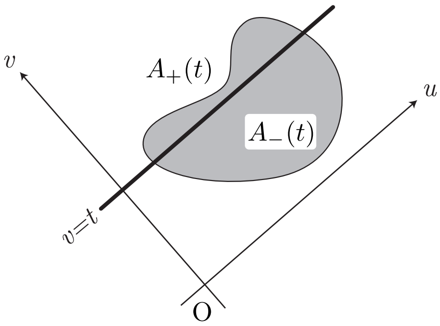


図 2.8: $A_-(t)$ と $A_+(t)$

じ向きを持つ任意の直線は $v = t$ （ただし，t は定数）の形に表されます．さて，ここで直線 $v = t$ の v 切片 t の値を変化させたときに，どんなことが起こるかを見てみることにしましょう．まず最初に，t の値が非常に小さいときには，領域 Ω 全体が直線 $v = t$ の上側[7]にあることが（領域 Ω の有界性より）判ると思います．ところが徐々に t の値を増加させていくと，そのうち直線 $v = t$ と領域 Ω とが共通部分をもつようになります．そしてさらに t の値を増加させていくと，今度は逆に領域 Ω 全体が直線 $v = t$ の下側にある，といった状態になります．いま一般に，直線 $v = t$ 上の下側にある部分の面積を $A_-(t)$，上側にある部分の面積を $A_+(t)$ としましょう（図 2.8）．ただし，領域 Ω 全体が直線 $v = t$ の上側にある場合には，$A_-(t) = 0$ と，また逆に領域 Ω 全体が直線 $v = t$ の下側にある場合には，$A_+(t) = 0$ と約束することにします．すると，ふたつの面積の差 $A_+(t) - A_-(t)$ は変数 t の関数と思えるので，それを記号 $f(t)$ で表すことにします．そのとき，関数 $f(t)$ が広義単調減少関数であること，すなわち，$t_1 < t_2$ ならば $f(t_1) \geq f(t_2)$ となることがただちに判ります．しかも，2 本の直線 $v = t_1$, $v = t_2$ の少なくとも一方が領域 Ω と共通部分を持つならば，$t_1 < t_2$ から $f(t_1) > f(t_2)$ が従います．また，t が十分小さいときには $f(t) > 0$，一方 t が非常に大きいときには $f(t) < 0$ が成り立つことも明らかだと思います．そして最後に，関数 $f(t)$ が t の連続関数

[7]ただし，ここで「上側」とは，あくまで (u, v)-座標系に関してのものです．

であることも，少なくとも直観的には困難なく見てとれるのではないでしょうか．ここまで来れば，後は簡単です．実際，関数 $f(t)$ に対し，中間値の定理を適用すれば，$f(t_0)=0$ なる実数 t_0 が存在することがただちに結論されます．さらに関数 $f(t)$ の単調性より，このような t_0 の一意性もただちに判ります．そしてこの t_0 に対応する直線 $v=t_0$ が，領域 Ω を面積の等しいふたつの部分に分割することは，関数 $f(t)$ の定義からすぐに見てとれると思います．以上が，問題 2.4 に対する解答です．

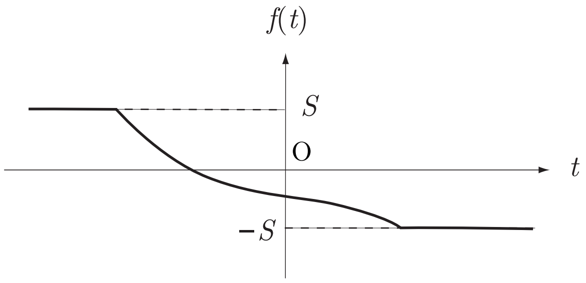


図 2.9: 関数 $f(t)$ のグラフ（ただし，Ω の面積を S とした）

ようやく問題 2.2 に対する解答を述べる準備が出来ました．いま，Ω_1，Ω_2 を平面内の任意の有界領域としましょう．一方，その平面内の原点から発する単位ベクトル[8] $\mathbf{u}$ を任意に取ります．この単位ベクトル $\mathbf{u}$ は，あらかじめ平面内に与えられた xy 座標系の x 軸正の部分とそれ自身がなす角 θ（ただし $0\leq\theta<2\pi$）により決定されることに注意しましょう．さて，ここで問題 2.4 の解答を，とくに第 1 の領域 Ω_1 に対し適用します．すると，ベクトル $\mathbf{u}$ に平行でしかも領域 Ω_1 を面積の等しいふたつの部分に分割する直線がただひとつ存在することが分かります．その直線は，結局単位ベクトル $\mathbf{u}$ と x 軸正の方向のなす角 θ により一意的に確定します．そこで，今後はその直線を ℓ_θ と記すことにします．直線 ℓ_θ は第 1 の領域 Ω_1 を等分しますが，第 2 の領域 Ω_2 を等分するとは限りません．そこで，次のような θ の関数 $g(\theta)$ を導入します．まず，直線 ℓ_θ が領域 Ω_2 をふたつの部分に分ける場合を考えます．直線 ℓ_θ によって分割された Ω_2 のふたつの部分のうち，直線 ℓ_θ 上をベクトル $\mathbf{u}$ と同じ向きに進んだとき，その右側に見える部分をとり，その面積を $A_R(\theta)$ と，またその左側に見える部分の面積を $A_L(\theta)$ とします．一方，直線 ℓ_θ 上をベクトル $\mathbf{u}$ の向きに進んだとき，領域 Ω_2 全体がその右側にのみにある場合には $A_L(\theta)=0$ と，また逆に，左側にのみにある場合には，$A_R(\theta)=0$ と定義す

[8] **単位**ベクトルとは長さが 1 のベクトルを意味します．

るのはきわめて自然です．そして，最後に

$$g(\theta) = A_R(\theta) - A_L(\theta)$$

とおくことにより，関数 $g(\theta)$ を定義します．

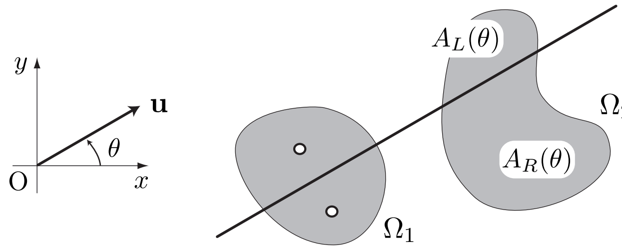


図 2.10: $A_R(\theta)$ と $A_L(\theta)$

たったいま導入した関数 $g(\theta)$ $(0 \leq \theta < 2\pi)$ の最も重要な性質として

$$g(\theta + \pi) = -g(\theta) \qquad (0 \leq \theta < \pi) \tag{2.4}$$

が成り立つことがあります．どうしてこの (2.4) が成り立つのかを，説明しましょう．ふたつの直線 ℓ_θ と $\ell_{\theta+\pi}$ を考えると，定義によれば，これらはともに領域 Ω_1 を等分し，しかも，互いに平行でなければなりません．さてここで，問題 2.4 に対する解答を思い起しましょう．与えられた直線に平行で，しかも与えられたひとつの領域を等分する直線の存在だけでなく，その一意性もそこで示したはずです．とくに，その一意性を適用することにより，２つの直線 $\ell_\theta, \ell_{\theta+\pi}$ が一致することが分かります．ただし，ベクトル $\mathbf{u}$ から定まる直線 ℓ_θ と直線 $\ell_{\theta+\pi}$ の向きは反対なので，直線 ℓ_θ から見て右側にあるものは直線 $\ell_{\theta+\pi}$ から見ると左側に，また逆に ℓ_θ から見て左側にあるものは $\ell_{\theta+\pi}$ から見ると右側に位置することになります．これより，$A_R(\theta + \pi) = A_L(\theta)$, $A_L(\theta + \pi) = A_R(\theta)$ $(0 \leq \theta < \pi)$ が成り立つことが分かります．そして，これより直ちに，関数 $g(\theta)$ が性質 (2.4) を有することが結論されるわけです．

性質 (2.4) から特に $g(0) \cdot g(\pi) \leq 0$ が成り立つことが見て取れます．さらに関数 $g(\theta)$ のもうひとつの重要な性質として— ただし，その証明は容易ではありませんが — それが連続であることが挙げられます．したがって，関数 $g(\theta)$ に対し中間値の定理を適用すれば，$g(\theta_0) = 0$ なる角度 θ_0（ただし，$0 \leq \theta_0 < \pi$）の存在が従います．この角 θ_0 に対応する直線 ℓ_{θ_0} が，第 1 の領域 Ω_1 のみならず第 2 の領域 Ω_2 をも等分します．これで，問題 2.2 に対する解答が（関数 $f(t)$，および $g(\theta)$ の連続性の証明を除いて）終了したわけです．

A.4　電車の問題解説

問題2.3の解説に進むことにしましょう．はたして，本当にこの問題のなかで述べられているようなことは可能なのでしょうか．極めて不思議なこと，難しいことのように感じられるのではないでしょうか．ところで，この問題で述べられている棒の運動自体はニュートンの運動方程式という，一種の常微分方程式により記述されます．けれども，この問題を考えるにあたっては，運動方程式を具体的に解こうとすることは得策ではありません．いえ，そもそも電車の運行自体が具体的に与えられていないわけですから，運動方程式を具体的に解けるはずがないのです．それでは，電車の運行が具体的に与えられていた場合に限ったらどうでしょうか．たしかにその場合には，運動方程式を解き棒の運動を決定することで問題に答えられるかもしれません．しかし実際それを実行することは，一般には決して簡単ではないはずです．ところが，これから見るように，中間値の定理を用いると，なんの技術的困難もなく，しかもどんな（とくに具体的に与えられているとは限らない）電車の運行に関しても，この問題を解くことが出来ます．

ともかく中間値の定理をこの問題に適用するためには，なんらかの関数を問題から「抽出」する必要があります．そのためにもう一度問題を見直してみましょう．出発時における棒の角度を選ぶ，これだけがこの問題のなかでわれわれが自由に出来ることです．そこで，出発時の棒の角度 θ （ただし $-\pi/2 \leq \theta \leq \pi/2$) を図2.11のように計り（ただし，その符号を，前向きに傾いているときにはマイナス，後ろ向きに傾いているときはプラス，と約束します），それをいま捜している関数の独立変数と考えることは，自然のよう思えます．一方，われわれがコントロールしなければならないのは，到着時の棒の角度です．途中のことは問題にはなりません．しかも,「電車の運行はあらかじめ決まっている」，という約束からすると，到着時の棒の角度は出発時の棒の角度 θ のみで決定される，すなわち，変数 θ の関数であると考えられます．そこで到着時の棒の角度を出発時と同様な仕方で計り，それを $h(\theta)$ と書くことにしましょう．

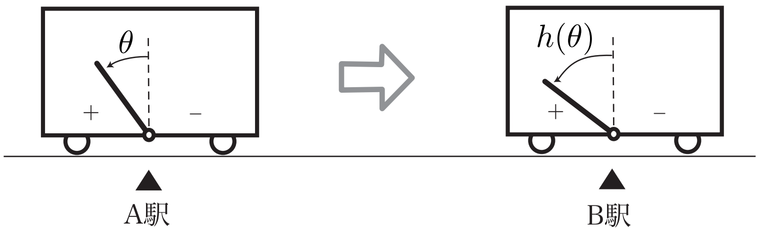


図 2.11: 棒の傾き具合の計り方

さて，ここでもうひとつの取り決め，すなわち，「棒はいったん倒れたらそれ以降ずっと倒れたままである」という約束を思い出してください．したがって，とくに出発時に棒を倒しておけば，到着時にも棒は同じ向きに倒れている，ということになります．これを関数 $h(\theta)$ を用いて表現すれば，

$$h\Big(-\frac{\pi}{2}\Big)=-\frac{\pi}{2} \qquad \text{かつ} \qquad h\Big(\frac{\pi}{2}\Big)=\frac{\pi}{2}$$

ということに他なりません．一方，先程述べた通り，棒の運動はニュートンの運動方程式という，一種の常微分方程式で記述されます．そして，「常微分方程式の解は，初期条件に連続的に依存する」という，常微分方程式に関する基本的な定理のひとつを適用することにより，特に関数 $h(\theta)$ が連続であることが結論されます．あとは単に中間値の定理を関数 $h(\theta)$ に適用するだけです．すると， $h(\theta_0)=0$ となる角度 θ_0 が存在することが分かります．ところが，これは，出発時に棒を角度 θ_0 で立たせておけば，到着時には棒が鉛直である，ということにほかなりません．以上が問題 2.2 に対する答えです．

B　連続性の定義再考

B.1　連続性の新たな定義

§A で述べた定義によれば，区間 I 上で定義された関数が連続であるとは，そのグラフが連続曲線となることでした．しかし，実はこの定義，いくつか本質的な難点を持ちます．この定義によれば，関数の連続性は，そのグラフとして得られる曲線の連続性に帰結されます．しかし，そもそも，曲線，およびその連続性とは一体何なのでしょう．これに対する厳密な解答無くしては，関数の連続性に関する厳密な議論は不可能です．これが，§A で述べた連続性の定義の第 1 の問題点です．

第 2 の問題点は，例えば前節で取り扱ったような，中間値の定理の応用問題の中に現れます．問題 2.2，2.3 に対する解答を思い出しましょう．その中には，f, g, h 合計 3 つの関数が現れました．そして，それらの連続性が解答において本質的な役割を果たします．これらの関数の連続性，直観的には正しいであろうと感じられるのではないでしょうか．でも，本当にそれらの関数は連続なのでしょうか．一旦疑いを抱くと，厳密な証明を与えない限り心が晴れることはありません．しかし，どうやってこれらの関数の連続性を証明すればよいのでしょうか．§A で与えた定義に基づくならば，まずこれらの関数のグラフを描き，そのグラフが連続曲線になっていることを確認すればよいはずです．ところが，前節の問題の解答で利用した 3 つの関数は，具体的な式で与えられているようなものではありません．したがって，そのグラフを具体的に描くことは出来ません．これらの関数の連続性をグ

ラフを介して判定することは困難なはずです．これらの関数以外にも，抽象的な仕方でしか定義できない関数を取り扱う必要がしばしばあります．その際も，事情は全く同様です．これが第 2 の問題点です．

第 3 の問題点は，高次元化が困難であることです．より進んだ数学においては，区間上で定義された関数のみならず，ふたつのユークリッド空間の間で定義された写像 $\mathbb{R}^n \to \mathbb{R}^m$ や，さらには「無限次元」空間の間で定義された写像の連続性を議論する必要があります．ところが，このような関数・写像のグラフを描くことは，通常の意味では不可能です． 例えば，n-次元ユークリッド空間 $\mathbb{R}^n$ から m-次元ユークリッド空間 $\mathbb{R}^m$ への写像を考えましょう．このとき，$(m+n)$-次元ユークリッド空間 $\mathbb{R}^{m+n}$ の中の部分集合として，その写像のグラフを形式的に定義することは可能です．しかし，$m+n>3$ の場合には，そのグラフを実際に描くことは不可能ですし，それどころか，そのグラフを頭の中に「思い浮かべる」ことさえ極めて困難なはずです．

関数の連続性をそのグラフの連続性に還元することにより定義するという方法は，とても直観的で，分かりやすいものであるように思えるかも知れません．しかし，実はそれが種々の問題を内包することを見てきました．そこで，われわれはこの問題の多い定義を放棄して，全く別の定義を採用することにします．

定義 2.5.

区間 I 上で定義された関数 $f(x)$ を考える．点 $a \in I$ に対し以下の条件が満たされるとき，関数 $f(x)$ は $x=a$ において**連続**であると言われる：$\{a_n\}$ を，すべての番号 n に対し $a_n \in I$，かつ a に収束するような任意の数列としたとき，必ず，

$$\lim_{n\to\infty} f(a_n) = f(a) \tag{2.5}$$

が成り立つ．

またすべての $a \in I$ において $f(x)$ が連続であるとき，$f(x)$ は I **上連続**，あるいは単に**連続**であると言われる．

条件 (2.5) において $a = \lim_{n\to\infty} a_n$ でしたから，これを右辺に代入すれば，

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n)$$

を得ます．言い換えれば，関数 $f(x)$ は極限操作 $\lim_{n\to\infty}$ と「可換」である，ということになります．そして，これが定義区間 I 内の任意の収束数列に対し成り立つということが，連続性の意味です．

最も簡単な「反例」として，まず (2.2) で定義された関数 $f(x)$ を考えましょう．この関数が，$x=0$ において連続でないことを，新たな定義に基づき確かめることにします．そのために，数列 $\{a_n\}$ を $a_n=-1/n$ $(n=1,2,\cdots)$ により定義しましょう．このとき，$a_n \to 0$ $(n\to\infty)$ ですが，

$$f(a_n)=-1/n+1 \longrightarrow 1 \neq -1 = f(0)$$

となります．これで，確かに関数 (2.2) が $x=0$ において不連続であることが，新たな定義にのっとって示されました．[9]

一方，連続な関数の例としては，多項式で定義された関数や，図 2.2 に挙げた関数が挙げられます．これらの関数は定義 2.5 の意味でも連続です．また，連続なふたつの関数に対し，その和・差・積・商（ただし，分母がゼロにならない範囲で）が連続であるということも，以前と同様です．[10]

この定義，関数の連続性をそのグラフの連続性により定義するという流儀と比べると，随分と形式的です．直観的な理解を拒んでいるように感じられるかも知れません．しかし，この形式性こそが連続性の厳密な取り扱いを初めて可能にします．また後に見るように，幾何学的直観に依存していないことから，その高次元化も容易です．[11]

定義 2.5 を関数の連続性の定義として採用したからには，それがわれわれの直観と合致することを確認しなければなりません．そのためには，この新たな定義に基

[9]これで，連続性の新たな定義の正当性が，少しだけ確かめられたことになります．しかし，確信を得るには，もっと多くの場面でこの定義が私たちの直観と一致することを確かめねばなりません．そのような場面の典型として，以下で述べる中間値の定理や最大値・最小値の定理，あるいはそれらの証明が挙げられます．

[10]本当は，この章の最初で扱ったふたつの問題 2.2, 2.3 の解答に現れた 3 つの関数 $f(t)$, $g(\theta)$, $h(\theta)$ の連続性も，この新たな定義に基づき検証せねばなりません．最初の関数 $f(t)$ の連続性は，それほど難しいものではないはずです．しかし，第 2 の関数 $g(\theta)$ の連続性の証明は，かなり込み入ったものになります．要点は，直線 ℓ_θ の一意性です．ちなみに，この例に限らず，一意性から連続性が従うということがよくあります．前にも触れたとおり，第 3 の関数 $h(\theta)$ の連続性は，常微分方程式に対する基本的な定理からの帰結で，その証明を自力で行うことはとても大変なことです．いずれにせよ，これら 3 つの関数の連続性の証明は，本著の範囲を超えるものであり，ここでは省略せざるを得ません．

[11]ただし，いろいろな現象を予期したりあるいは理解したりするに際しては，グラフを介して連続性を理解するという考え方は，（少なくもと区間上で定義された関数に対しては）極めて重要であることに変わりありません．現に，関数のグラフを利用すれば，中間値の定理が成り立つことは（その証明は別として），たちどころに理解されたわけでした．

これに限らず，数学においては直観的理解が極めて重要な役割を果たします．もちろん，数学における厳密性の志向は，他の学問分野と比べても顕著です．とくに，数学的主張，あるいは仮説は，それが証明されて初めて正しいもの，すなわち定理として認知されます．しかし，私たちが新たな定理を獲得するに先だち，その定理が成り立つことを期待し，そして，直観に基づきそれが正しいであろうことを判断します．通常，こういった過程を経て定理は生まれてきます．直観力こそ数学の原動力である，と言っても過言ではないでしょう．

しかし，一方，数学の対象は極めて抽象的かつ一般的で，その中にはときに一種病理的なものが含まれています．そのようなものに対してまで私たちの直観を安易に適用すると，ただちに誤謬を犯すことになります．その危険を回避するために，厳密かつ論理的な思考が必要となります．

づき，連続関数の種々の性質を調べるという作業が必要となります．そのような経験を積み重ねることで，私たちの定義が徐々に正当化されていきます．その手始めとして，次節で中間値の定理の証明を考えることにします．

B.2　中間値の定理の証明

定義 2.5 を関数の連続性の定義として採用したいまとなっては，もはや中間値の定理は自明な主張ではありません．定義 2.5 に基づいた中間値の定理の証明を与える必要があります．

定理 2.1（中間値の定理）の証明． $f(a) \cdot f(b) = 0$ が成り立つときには，$f(a) = 0$ または $f(b) = 0$ でなければならない．したがって，この場合には $c = a$ あるいは $c = b$ とおけば定理が成り立つことは明らかである．そこで以下では $f(a) \cdot f(b) < 0$ の場合を議論する．このとき，さらに，$f(a) < 0$ かつ $f(b) > 0$ が成り立つか，あるいは $f(a) > 0$ かつ $f(b) < 0$ が成り立つかのふたつの場合が考えられる．しかし，第 2 の場合には，第 1 の場合の議論を少々手直しすることで，あるいはは $-f$ を改めて f と取り直し第 1 の場合に帰着することにより証明可能である．そこで，以下では第 1 の場合，すなわち

$$f(a) < 0 \qquad \text{かつ} \qquad f(b) > 0$$

が満たされている場合のみを取り扱うことにする．

最初に，ふたつの数列 $\{a_n\}, \{b_n\}$ を以下の条件を満たすように定義する：

$$a \leq a_n \leq b_n \leq b; \tag{2.6}$$

$$b_n - a_n = \frac{b-a}{2^n}; \tag{2.7}$$

$$f(a_n) < 0, \quad f(b_n) \geq 0; \tag{2.8}$$

$$\{a_n\} \text{ は単調増加,} \quad \{b_n\} \text{ は単調減少}; \tag{2.9}$$

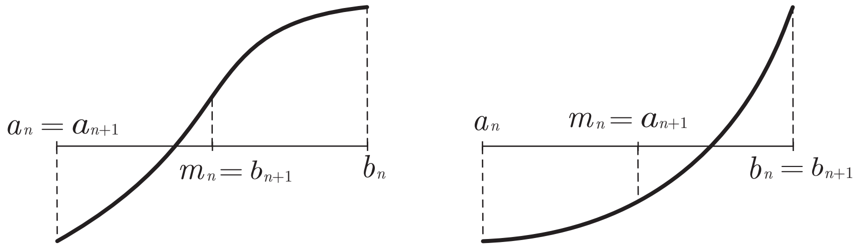
実際，これらの数列は以下のように帰納的に定義される．まず，

$$a_0 = a, \qquad b_0 = b$$

で初項 a_0，b_0 を定義する．次にすでに a_n，b_n が (2.6) – (2.9) を満たすように定義されているとき，次の項 a_{n+1}，b_{n+1} を，

$$\begin{aligned} f(m_n) \geq 0 &\quad \text{のときには} \quad a_{n+1} = a_n, \qquad b_{n+1} = m_n, \\ f(m_n) < 0 &\quad \text{のときには} \quad a_{n+1} = m_n, \qquad b_{n+1} = b_n \end{aligned}$$

で定義しよう．ただし，$m_n = (a_n + b_n)/2$ は 2 点 a_n, b_n の中点である．このようにして定義されたこれらふたつの数列 $\{a_n\}$，$\{b_n\}$ が実際に条件 (2.6) – (2.9) を満たしていることを確かめることは極めて容易である．



(a) $f(m_n) \geq 0$ の場合　(b) $f(m_n) < 0$ の場合

図 2.12: 数列 $\{a_n\}$, $\{b_n\}$ の定義

(2.6) および (2.9) によれば，$\{a_n\}$，$\{b_n\}$ は，おのおの上に有界な単調増加数列および下に有界な単調減少数列である．したがって，有界単調数列の収束性定理により，これらふたつの数列がともに収束することが保証される．しかも (2.7) によれば $b_n - a_n \to 0$（$n \to \infty$）であるから，これらふたつの数列の極限は一致しなければならない．そこで，$\{a_n\}$，$\{b_n\}$ の共通の極限を c で表わすことにする：

$$c = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$

(2.6) よりもちろん $a \leq c \leq b$ である．

この c に対し $f(c) = 0$ が成り立つことを以下で示したい．まず，$c = \lim_{n\to\infty} a_n$，関数 $f(x)$ の連続性，および $f(a_n) < 0$ より，

$$f(c) = f(\lim_{n\to\infty} a_n) = \lim_{n\to\infty} f(a_n) \leq 0$$

を得る．一方，$c = \lim_{n\to\infty} b_n$，関数 $f(x)$ の連続性，および $f(b_n) \geq 0$ より，

$$f(c) = f(\lim_{n\to\infty} b_n) = \lim_{n\to\infty} f(b_n) \geq 0$$

である．これらふたつの不等式から直ちに $f(c) = 0$ が結論される．　□

B.3　区間縮小法

中間値の定理の証明で用いられた論法は，しばしば辞書の引き方に例えられます．それを説明するために，ある単語の意味を英和辞典を使って調べたいという状況を考えます．例えば，その単語は，そう，“intermediate” だとしましょう．[12] 辞書

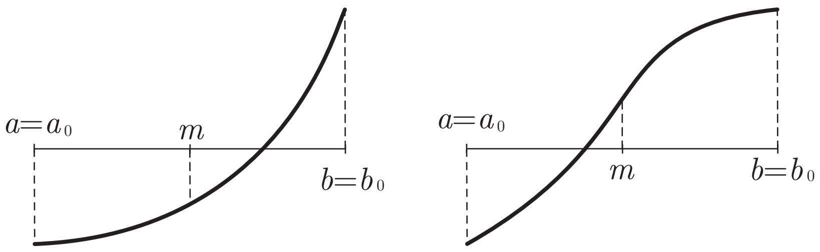
[12]「中間値の定理」は，英語では，“The intermediate value theorem” と呼ばれます．

の一番最初から順次調べていくのがひとつのやり方です．英和辞典には単語がアルファベット順に配列されているわけですから，その一番最初のページから順番に 1 ページずつ調べていけば，使っている辞書に “intermediate” という項目がある限り，やがてはそれにたどり着くはずです．しかしながら，これは一般的には効率のいい方法とは言えません．[13] これより賢明な辞書の引き方として次のような方法があります．いま，先ほどと同様 “intermediate” という単語の意味を調べたいとします．そのために利用する英和辞典のページ数を仮に $2^{10} = 1024$ としましょう．また，その辞書には “intermediate” という項目が確かに含まれているとします．とりあえずその辞書のどのページに “intermediate” という項目が記載されているかが判れば十分としましょう．それを知るためにまず辞書の（ほぼ）真ん中第 512 ページを開いてみます．もし，そのページにはアルファベット k から始まる単語のみが記載されているとしたら，探している単語 “intermediate” はその辞書の前半分，すなわち第 1 ページ以降第 512 ページ以前にある，ということが判ったことになります．次に，辞書の前半分の真ん中，第 256 ページを開きましょう．もしそのページには h から始まる単語が並んでいるようだったら，単語 “intermediate” は第 256 ページから 512 ページの間にあることがわかります．そして，この真ん中，真ん中を開くという操作を繰り返し，探している単語が載っている範囲を狭めていくことで，やがでは探している単語の記載されているページに到達するはずです．

いま述べた辞書の引き方と，中間値の定理の証明を比較してみましょう．辞書を使って意味を調べたい単語，すなわち上の例においては “intermediate” という単語が，中間値の定理における $f(c) = 0$ なる c の対応物であると考えるならば，全 1024 ページからなる辞書を，中間値の定理における関数 $f(x)$ の定義域 $[a, b]$ と見なすことができるでしょう．そして，辞書の中央第 512 ページを開いてそこに載っている単語と “intermediate” という単語のアルファベット順を比較するという作業は，中間値の定理の証明では，区間 $[a, b]$ の中点 $(a + b)/2 = (a_0 + b_0)/2$ における関数 $f(x)$ の符号による場合分けに相当します．上の例において，第 512 ページに掲載されている単語が k から始まるようなら単語 “intermiditate” は第 1 ページから第 512 ページまでのいずれかに現れるであろう様に，中間値の定理の証明においては（$f(a) < 0$，$f(b) > 0$ を仮定していたので）もし $f((a + b)/2) \geq 0$ ならば $a_1 = a_0 = a$，$b_1 = (a_0 + b_0)/2 = (a + b)/2$ と，また逆に $f((a + b)/2) < 0$ ならば $a_1 = (a_0 + b_0)/2 = (a + b)/2$，$b_1 = b_0 = b$ とおけば $f(c) = 0$ なる c が区間 $[a_1, b_1]$ 内に少なくともひとつ存在するであろうと「推測」されるわけです．あるいは言い換えると，最初に与えられた区間 $[a, b]$ をその中点 $m = (a + b)/2$ において分割す

[13] そもそも,「効率の良い」方法とは一体どんなものなのでしょう．各自，その点についても考えて見て下さい．

ることにより得られるふたつの区間 $[a,m]$，$[m,b]$ のうちの一方を，f の m における符号に応じてうまくとりあらたに $[a_1,b_1]$ とおくと，$f(a_1)<0$，$f(b_1)\geq 0$ とすることができます．(中間値の定理を認めるならば）この区間 $[a_1,b_1]$ の中に方程式 $f(x)=0$ の解が少なくともひとつあるはずです．さらにこの区間 $[a_1,b_1]$ をふたつの区間 $[a_1,m_1]$，$[m_1,b_1]$（ただし $m_1=(a_1+b_1)/2$）に等分しそのうちの一方をうまくとり $[a_2,b_2]$，その区間の左端 a_2 では f は負，右端 b_2 では f は非負となります．したがって，区間 $[a_2,b_2]$ は方程式 $f(x)=0$ の解を少なくともひとつ含むはずです．この議論を繰り返すと，$f(x)=0$ の解を含んでいる区間の幅がどんどん小さくなり，したがって極限をとることにより解を捉えることができるであろう，というのが中間値の定理の証明で用いられた論法です．そしてこの論法は数学では通常**区間縮小法**と呼ばれます．



(a)　(b)

図 2.13: 区間縮小法

B.4　近似解の計算

区間縮小法は，ただ単に中間値の定理の証明に用いられるだけではありません．方程式の近似解を求める最も簡単な方法としても知られています．それに関連した例題を出しましょう．

問題 2.6. ――――――――――――――――――――――――――

方程式

$$x^5+3x^2+2x-3=0 \tag{2.10}$$

の $0\leq x\leq 1$ における近似解を，誤差 2^{-3} 以内で求めよ．[14]

[14]方程式 (2.10) は 5 次方程式です．もしこれが 2 次方程式だったら，皆さんご承知の通り，解の公式を適用することによりその解を厳密に求めることができます．また，実は 3 次方程式，および 4 次方程式に対しても，2 次方程式に対する解の公式に相当するものが知られているので（ただしその形は 2 次方程式の解の公式と比べるとはるかに複雑ですが)，それを利用することにより 3 次ないし 4 次の方程式の厳密解を求めることもできます．ところがこれに反し，5 次以上の方程式に対しては，驚くべきことに,「解の公式」が存在しないことが知られています．もはや，解の公式を利用して解を求めることが原理的に不可能なわけです．そのために近似解を求めることが，なおさら重要になります．

解答. 簡単のため，(2.10) の左辺を $f(x)$ とおく：

$$f(x) = x^5 + 3x^2 + 2x - 3.$$

このとき，

$$f(0) = -3 < 0, \qquad f(1) = 3 > 0$$

であるから，中間値の定理により，方程式 (2.10) が区間 $[0,1]$ 内に確かに解を有することが保証される．次いで，中点 $1/2$ における関数 $f(x)$ の値を計算すると，

$$f\left(\frac{1}{2}\right) = -\frac{39}{2^5} < 0$$

である．これより，区間 $[1/2,1]$ に解が存在することが分かる．さらにこの区間の中点 $3/4$ における関数 $f(x)$ の値を調べる．

$$f\left(\frac{3}{4}\right) = \frac{435}{2^{10}} > 0$$

であるから，区間 $[1/2,3/4]$ の中に解が少なくともひとつ存在することが分かる．いま，この区間の中点 $5/8$ とこの区間内の任意の数との差はたかだか $1/8 = 2^{-3}$ であることに注意する．これより，とくに $5/8$ と $[1/2,3/4]$ 内にある方程式 (2.10) の解との距離も，2^{-3} 以下であることが従う．よって，$5/8$ を誤差 2^{-3} 以内の近似解としてとることができる． □

C　最大値・最小値の定理

例えば,「利潤は最大，損失は最小」，誰もがこれを望むはずです．何かしらの量を最大あるいは最小にせよという問題が，このような実利的な場面に限らず，数学・物理学・化学等，自然科学の諸分野にも見出されます．この「最大・最小問題」に対する最も基本的な事実を学ぶのがこの節の目的です．まず何はともあれ，最大値，最小値とは何であるか，その定義を述べることにしましょう．

定義 2.7.

$f(x)$ を区間 I 上で定義された関数，また $a \in I$ とする．もし

すべての $x \in I$ に対し　$f(x) \leq f(a)$　が成り立つ

ならば，関数 $f(x)$ は $x = a$ において**最大**であると言われる．また逆に

すべての $x \in I$ に対し　$f(x) \geq f(a)$　が成り立つ

が成り立つならば，関数 $f(x)$ は $x = a$ において**最小**であると言われる．さらに，関数 $f(x)$ が a において最大，あるいは最小のとき，値 $f(a)$ をそれぞれ関数 $f(x)$ の**最大値**，**最小値**と呼ぶ．

さて，定義を理解するためにいくつか例を挙げることにしましょう．まず，関数 $f(x) = -4x(x-1)$（ただし $x \in [0,1]$）を考えます．この関数のグラフを描くと，図 2.14 (a) のようになります．そのグラフで一番「標高」の高いところで関数 $f(x)$ は最大，また一番低いところで最小です．すなわち，関数 $f(x)$ は，$x = 1/2$ において最大で，最大値は $f(1/2) = 1$，また $x = 0, 1$ において最小で，最小値は $f(0) = f(1) = 0$ です．次に，関数 $g(x) = e^x$（ただし $x \in (-\infty, \infty)$）を考えましょう（図 2.14 (b)）．関数 $g(x)$ は x が大きくなるにつれて，いくらでも大きい値を取りますから最大値を持ちません．それでは，最小値はどうでしょうか．皆さんの中にはゼロが最小値である，と答えた人がいるのではないかと思います．確かに，関数 $g(x) = e^x$ は，負の値を取りませんし，また一方，ゼロにいくらでも近い正の実数をその値として持ちますから，最小値がゼロであるという答えは一見正しいようにも思えます．しかし，ちょっと待って下さい．最小値の定義によれば，ゼロが関数 $g(x)$ の最小値であるためには，$g(x) = 0$ となるような x が $g(x)$ の定義域 $(-\infty, \infty)$ の中に存在しなければなりません．ところが，$g(x) = e^x$ は決してゼロを値としてを取り得ません．したがって，ゼロは $g(x)$ の最小値ではないのです．この関数 $g(x)$ は最小値を持たない，というのが正しい答えです．それでは，

$$h(x) = \begin{cases} 1 + x & -1 \le x < 0 \\ 0 & x = 0 \\ -1 + x & 0 < x \le 1 \end{cases}$$

で定義される関数 $h(x)$（ただし，$x \in [-1, 1]$）はどうでしょう（図 2.14 (c)）．この関数に対しても，一見 1, -1 がそれぞれ最大値，最小値のように見えるかも知れません．しかし，関数 $h(x)$ はそれらを値として持ちませんので，これらは最大値・最小値ではありません．関数 $h(x)$ に対しては，最大値・最小値いずれも存在しない，というのが正しい答えです．

これらの例が示すように，最大値・最小値はいつでも存在するわけではありません．そこで，数列の極限の場合と同様，最大値・最小値の存在を保証するような定理が欲しくなります．そして，その種の定理の典型的なものとして，次のようなものがあります．

定理 2.8 (最大値・最小値の定理).
有界閉区間 $I = [a, b]$ 上で定義された連続関数 $f(x)$ は，必ず最大値・最小値を有

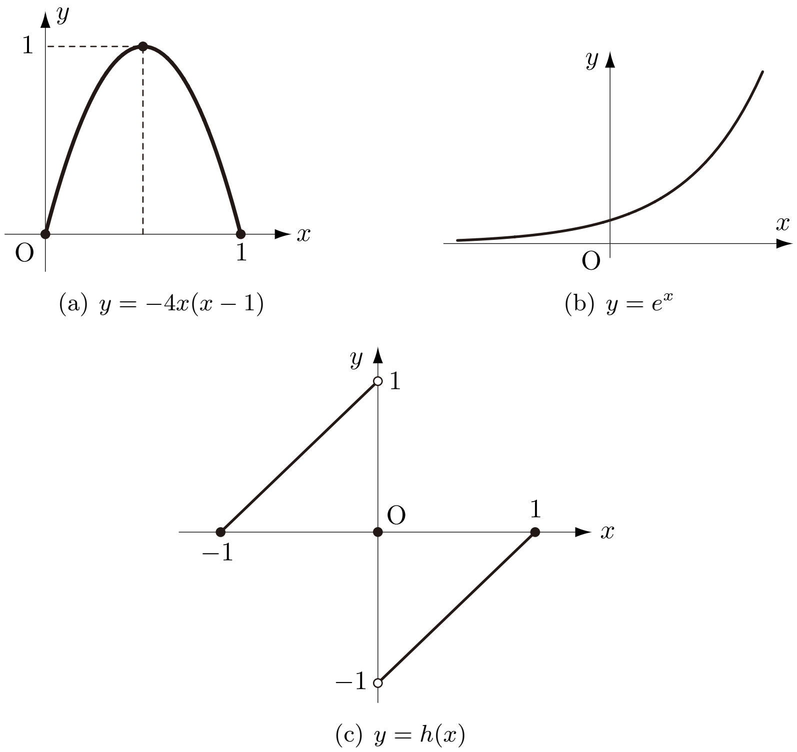


図 2.14: 関数の最大値・最小値

する．

関数 $f(x)$ が連続であること，およびその定義域が有界閉区間あること，このふたつが仮定されていることに注意して下さい．実際，先程の例によれば，これらふたつの仮定のうちどちらか一方をはずしてしまうと，もはや定理は成り立ちません．

この最大値・最小値の定理の証明も，区間縮小法を用いて実行することが可能です．

定理 2.8（最大値・最小値の定理）の証明. 関数 $f(x)$ が最大値を持つことを証明する（最小値の存在も同様な仕方で示すことが可能なので，ここではその証明は省略する）.

最初に，有界閉区間 $I_n=[a_n,b_n]$ (ただし，$n=0,1,\cdots$) を，以下のように帰納法により定義する．まず，$n=0$ のときには，$I_0=[a,b]$，すなわち，$a_0=a, b_0=b$ とする．次に，区間 $I_n=[a_n,b_n]$ が定義できたと仮定して，区間 $I_{n+1}=[a_{n+1},b_{n+1}]$ を以下の仕方で定義する．ただし，$m_n=(a_n+b_n)/2$ とする．

(i) 適当に $x \in [a_n, m_n]$ を取ったとき，すべての $x' \in (m_n, b_n]$ に対し $f(x) > f(x')$ が成り立つ場合：その場合には，

$$a_{n+1} = a_n, \qquad b_{n+1} = m_n$$

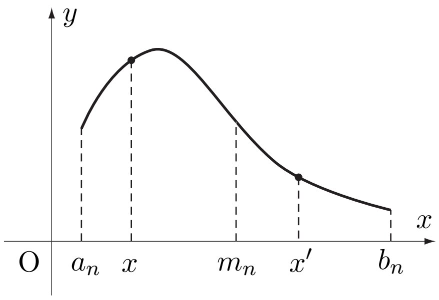
とする．

(ii) そうでない場合，すなわち，各 $x \in [a_n, m_n]$ に対し，$f(x) \leq f(x')$ なる $x' \in (m_n, b_n]$ が存在する場合：このときには，

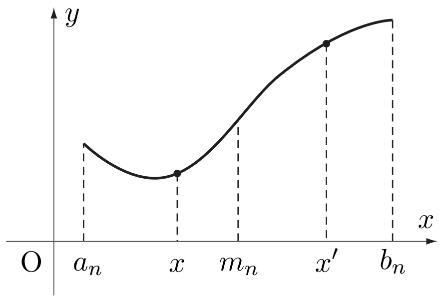
$$a_{n+1} = m_n, \qquad b_{n+1} = b_n$$

とする．

このとき，中間値の定理 2.1 の証明でおいてもそうであったように，ふたつの数列 $\{a_n\}$, $\{b_n\}$ が同一の極限 $c \in [a, b]$ に収束することが，有界単調数列の収束性定理により導かれる．



(a) (i) の場合



(b) (ii) の場合

図 2.15: 区間 $I_n = [a_n, b_n]$ の定義

そこで，

$$c = \lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n \tag{2.11}$$

とおく．すると，

$$M = f(c)$$

が関数 $f(x)$ の最大値になる．ただし，そのことを示すためには，ふたつほど準備が必要である．それらを述べよう．

(a) 各 $x' \in I \setminus I_n$ に対して，$f(x') \leq f(x)$ なる $x \in I_n$ を見い出すことが出来る．

(b) $\epsilon > 0$ を任意の定数とする．このとき，番号 n を適当にとれば，

$$\text{すべての } x \in I_n \text{ に対し，} f(x) \leq M + \epsilon$$

が成り立つ．

第 1 の主張 (a) は，区間 $I_n = [a_n, b_n]$ の定義からの容易な帰結であるので，ここではその証明は省略する．[15] 一方，第 2 の主張 (b) は，背理法により次のような仕方で証明される．いま，仮に (b) が正しくないとしよう．すると，$\epsilon > 0$ を適当にとれば，各番号 n に対し，区間 $I_n = [a_n, b_n]$ 内の点 x_n で，$f(x_n) > M + \epsilon$ を満たすものが存在しなければならない．ところが，$x_n \in [a_n, b_n]$，すなわち，$a_n \leq x_n \leq b_n$ であるから，(2.11) から，$x_n \to c$ $(n \to \infty)$ であることが従う．さらに関数 $f(x)$ の連続性によれば，$\lim_{n\to\infty} f(x_n) = f(c) = M$ でなけばならない．ところが，先ほど述べたとおり，$f(x_n) > M + \epsilon$ であったから，$M = f(c) = \lim_{n\to\infty} f(x_n) \geq M + \epsilon > M$ となり，矛盾に達した．これで，(b) が正しいことが確認された．

さて，定数 M が関数 $f(x)$ の最大値であることを証明しよう．これもまた背理法による．仮に，$f(\gamma) > M$ なる $\gamma \in I$ が存在したとする．ここで，とくに $\epsilon = (f(\gamma) - M)/2 > 0$ に対し主張 (b) を適用しよう．すると，適当に番号 n をとれば，すべての $x \in I_n$ に対し，$f(x) \leq M + \epsilon < f(\gamma)$ が成り立たねばならない．従って，とくに γ は I_n の要素ではあり得ないことが判る．ところが，これは主張 (a) に矛盾している．これで，M が関数 $f(x)$ の最大値であることが証明できた．　□

最大値・最小値の定理により，有界閉区間上で定義された連続関数に対しその最大値・最小値の存在が保証されます．しかしながら，どうやって最大値・最小値を求めることが出来るのか，その方法について最大値・最小値の定理は何の情報をも与えていません．また，中間値の定理の場合はその区間縮小法による証明が $f(x) = 0$ の近似解を見いだす方法を与えていましたが，最大値・最小値の定理の証明は最大値・最小値の近似値を求めるための手がかりを与えてはくれません．[16] 最大値・最小値を決定する方法については，次の章で議論する予定です．

D　章末問題

まずは，問題 2.2 の類題をふたつほど出題します．

問題 2.9. ————————————————————

平面内に有界な領域がひとつ与えられたとする．このとき，その平面内の直交する 2 直線で，とくにその領域を面積の等しい 4 つの部分に分けるものが存在することを示せ．

————————————————————————————

[15] もちろん証明は，n に関する帰納法によります．

[16] 証明に背理法が利用されているのが，その理由です．これは，背理法を用いた存在証明に共通する難点です．

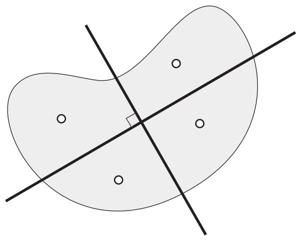
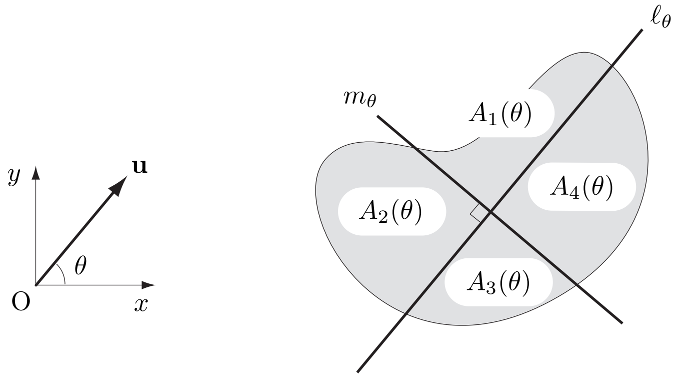


図 2.16: ケーキの 4 等分問題

解答. Ω を xy 平面内の任意の有界領域とする．原点 O を始点とし x-軸正の方向となす角が θ $(0 \leq \theta < 2\pi)$ であるような単位ベクトルを $\mathbf{u}$ とする．このとき，$\mathbf{u}$ と平行でしかも Ω を面積の等しいふたつの部分に分ける直線がただひとつ存在する（問題 2.4）．その直線を ℓ_θ と呼ぼう．一方，$\mathbf{u}$ と垂直で Ω を面積の等しいふたつの部分に分割する直線がやはり一意的に存在する．それを m_θ としよう．これら



ふたつの直線は Ω を合計 4 つの部分に分ける．それらの面積を図のように $A_1(\theta)$, $A_2(\theta)$, $A_3(\theta)$, $A_4(\theta)$ とする（厳密に言えば，直線 ℓ_θ をベクトル $\mathbf{u}$ の向きに進んだとき左手側にあるふたつの部分のうち，「奥側」にある部分の面積を $A_1(\theta)$ とし，あとは反時計回りに $A_2(\theta)$, $A_3(\theta)$, $A_4(\theta)$ と名前を付ける）．このとき，

$$A_1(\theta) + A_2(\theta) = A_3(\theta) + A_4(\theta), \qquad A_1(\theta) + A_4(\theta) = A_2(\theta) + A_3(\theta)$$

であった．したがって

$$A_1(\theta) = A_3(\theta), \qquad A_2(\theta) = A_4(\theta)$$

が常に成り立つ．しかし，$A_1(\theta) = A_2(\theta)$ が成り立つとは限らない（実際，これが成り立つとき，直交する 2 直線 ℓ_θ, m_θ は Ω を面積の等しい 4 つの部分に分割する訳である）．そこで，

$$f(\theta) = A_2(\theta) - A_1(\theta)$$

とおけば，$f(\theta)$ は $[0, 2\pi)$ 上で定義された連続関数である．しかも，

$$A_1\left(\theta + \frac{\pi}{2}\right) = A_2(\theta), \quad A_2\left(\theta + \frac{\pi}{2}\right) = A_3(\theta) = A_1(\theta)$$

であるから，

$$f\left(\theta + \frac{\pi}{2}\right) = -f(\theta)$$

が成り立つ．したがって，とくに $f(\pi/2) = -f(0)$ が従う．よって中間値の定理により，$f(\theta) = 0$ なる角 θ $(0 \leq \theta \leq \pi/2)$ が存在することが保証される．そしてこの角 θ に対応する直線 ℓ_θ, m_θ をとれば，それらが直交ししかも領域 Ω を面積の等しい 4 つの部分に分割することが直ちに分かる．□

問題 2.10. ————————————————————————

平面内に滑らかな境界を持つ有界凸領域が与えられているとする．このとき，この領域に 4 点で外接する正方形が存在することを示せ．

————————————————————————————

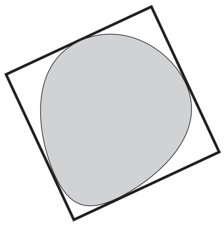


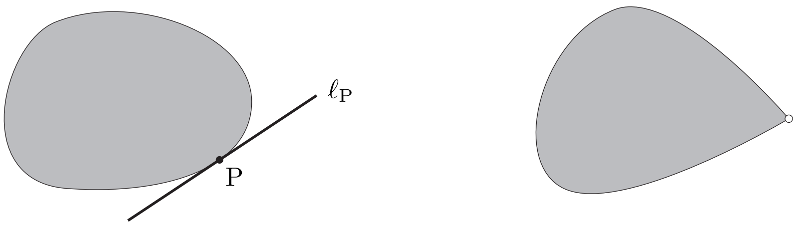
図 2.17: 有界凸領域に外接する正方形

この問題の中に現れるいくつかの用語の意味を説明しておきましょう．Ω を有界領域とします．その境界の各点 P において，境界の接線 ℓ_{P} が存在し，しかも P が Ω の境界上を動くときそれに応じて接線 ℓ_{P} の傾きが連続的に変化するならば，領域 Ω は**滑らかな**境界を有すると言われます．さらに，境界の各点 P における接線 ℓ_{P} の片側のみに Ω があるとき，領域 Ω は**凸**であると言われます．

問題 2.10 の解答. Ω を滑らかな境界を有する有界凸領域とする．このとき，次が成り立つ：

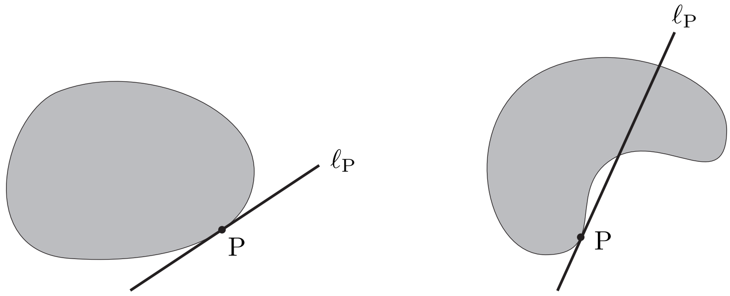
> 任意の単位ベクトル $\mathbf{u}$ に対し，Ω の境界 $\partial\Omega$ の接線でとくに $\mathbf{u}$ に平行なものがちょうど 2 本存在する．しかも，Ω はこのふたつの接線の間にある．

単位ベクトル $\mathbf{u}$ を任意にとる．それと x 軸正の方向のなす角を θ $(0 \leq \theta < 2\pi)$ としよう．このとき，Ω の境界 $\partial\Omega$ の接線であってベクトル $\mathbf{u}$ と平行なものがちょ



(a) 滑らかな境界を有する領域　　(b) 境界が滑らかでない領域

図 2.18: 境界の滑らかさ



(a) 凸領域　　(b) 非凸領域

図 2.19: 領域の凸性

うど 2 本，$\mathbf{u}$ と垂直な接線がちょうど 2 本存在する．これら 4 本の接線に囲まれる長方形は，当然 Ω に 4 点で外接する．$\mathbf{u}$ と平行な辺の長さを $a(\theta)$，垂直な辺の長さを $b(\theta)$ としよう．さらに θ の関数 $f(\theta)$ を

$$f(\theta) = a(\theta) - b(\theta)$$

で定義する．これは連続関数である．しかも，

$$a\Big(\theta + \frac{\pi}{2}\Big) = b(\theta), \qquad b\Big(\theta + \frac{\pi}{2}\Big) = a(\theta), \qquad f\Big(\theta + \frac{\pi}{2}\Big) = -f(\theta)$$

が成り立つ．これよりとくに $f(\pi/2) = -f(0)$ が従う．よって中間値の定理より，$f(\theta) = 0$，すなわち $a(\theta) = b(\theta)$ を満たす角 θ $(0 \leq \theta \leq \pi/2)$ の存在が保証される．この θ に対応する Ω の外接長方形がとくに正方形であることは明らかである．□

さて，次の問題は，関数の連続性の定義に関するものです．

問題 2.11. ————————————————————————————
関数 $f(x)$ $(a < x < b)$ が以下の条件を満たすとする：任意の実数 c, d（ただし，$a < c < d < b$）に対し，

$$f(x) \leq \frac{f(d) - f(c)}{d - c}(x - c) + f(c) \qquad (c \leq x \leq d) \tag{2.12}$$

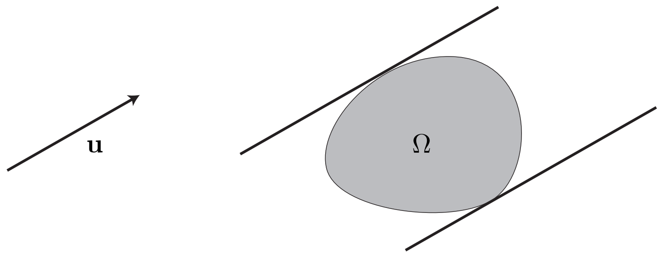


図 2.20: 与えられたベクトルに平行な 2 本の接線

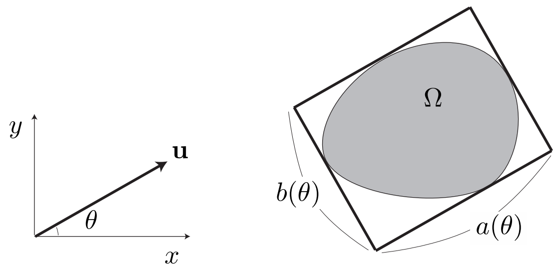


図 2.21: 領域に外接する長方形

が成り立つ．このとき，関数 $f(x)$ は連続であることを証明せよ．

区間 $[c,d]$ においては，関数 $y=f(x)$ のグラフは，その両端点 $(c,f(c)),(d,f(d))$ を結ぶ線分の下側にあることを，条件 (2.12) は意味しています（図 2.22）．ちなみに，この問題にあるような条件を満たす関数は**下に凸**であると言われます．

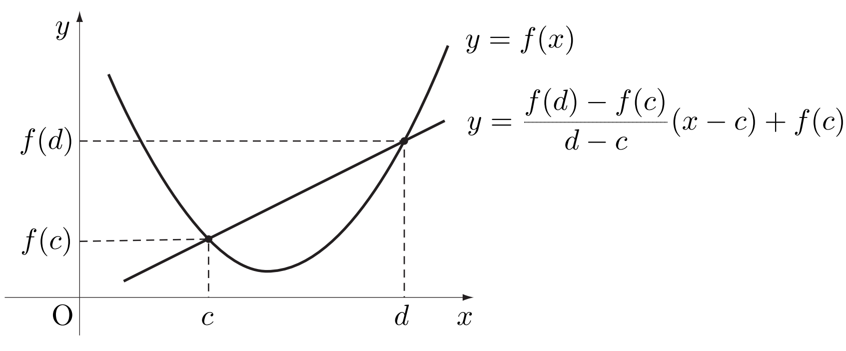


図 2.22: 凸関数

問題 2.11 の解答． 点 $c\in(a,b)$ を任意にとる．その点における関数 $f(x)$ の連続性を証明しよう．そのために，数列 $\{c_n\}$ を，$a<c_n<b$ $(n=0,1,\cdots)$ かつ $c_n\to c$ $(n\to\infty)$ となるようとる．そのとき，$f(c_n)\to f(c)$ であることを示せばよい．

[場合 1]　まず，数列 $\{c_n\}$ のすべての項が $c_n \leq c$ を満たす場合を考える．いま，$a < a_0 < c_n < c < b_0 < b$ $(n = 0, 1, \cdots)$ を満たすように定数 a_0, b_0 をとる．このとき，点 $(c_n, f(c_n))$ は 2 点 $(a_0, f(a_0))$, $(c, f(c))$ を結ぶ線分の下側になければならない（図 2.23）．すなわち，

$$f(c_n) \leq \frac{f(c) - f(a_0)}{c - a_0}(c_n - a_0) + f(a_0)$$

である．これより，

$$f(c_n) \leq \frac{c - c_n}{c - a_0} f(a_0) + \frac{c_n - a_0}{c - a_0} f(c)$$

を得る．一方，点 $(c, f(c))$ は 2 点 $(c_n, f(c_n))$, $(b_0, f(b_0))$ を結ぶ線分の下側にあるから，

$$f(c) \leq \frac{b_0 - c}{b_0 - c_n} f(c_n) + \frac{c - c_n}{b_0 - c_n} f(b_0)$$

でなければならない（図 2.23）．これらふたつの不等式より，

$$\frac{b_0 - c_n}{b_0 - c} f(c) - \frac{c - c_n}{b_0 - c} f(b_0) \leq f(c_n) \leq \frac{c - c_n}{c - a_0} f(a_0) + \frac{c_n - a_0}{c - a_0} f(c)$$

を得る．ここで，$n \to \infty$ とすると，上記不等式の最左辺，最右辺ともに $f(c)$ に収束する．したがって，$f(c_n) \to f(c)$ $(n \to \infty)$ が従う．

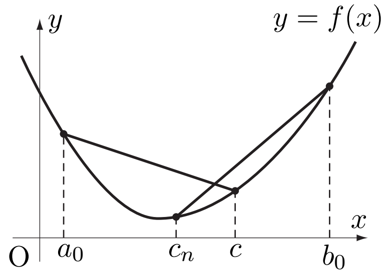


図 2.23: 凸関数の連続性

[場合 2]　数列 $\{c_n\}$ のすべての項が $c_n \geq c$ を満たす場合も同様にして，$f(c_n) \to f(c)$ $(n \to \infty)$ を示すことが出来る．

[一般の場合]　最後に，数列 $\{c_n\}$ が単に $a < c_n < b$, $c_n \to c$ $(n \to \infty)$ のみを満たす場合を考える．数列 $\{c_n\}$ に対し，$c_n > c$ なる項 c_n が有限個しかない場合には，ある番号から先では必ず $c_n \leq c$ がなりたつ．したがって，[場合 1] で見たとおり，$f(c_n) \to f(c)$ $(n \to \infty)$ が成り立つ．また，$c_n \leq c$ なる項が有限個しかない場合も同様である．したがって，$c_n > c$ なる項も，$c_n \leq c$ なる項も無限個存在する場合を考えればよいことになる．そのとき，数列 $\{c_n\}$ の項であって $c_n \leq c$

なるものを抜き出して得られる新たな数列を，$c_0^-, c_1^-, \cdots$ としよう．一方，$c_n > c$ なる項を抜き出して得られる数列を $c_0^+, c_1^+, \cdots$ とする．これらふたつの数列は明らかに c に収束する．したがって，[場合 1]，[場合 2] で見たとおり，これら新たな数列に対しては，$f(c_n^-) \to f(c), f(c_n^+) \to f(c)$ $(n \to \infty)$ が成り立つ．そしてこのことよりただちに元の数列 $\{c_n\}$ に対しても $f(c_n) \to f(c)$ $(n \to \infty)$ が成り立つことが従う．□

次の問題は連続性を利用した証明問題です．

問題 2.12. ————————————————————————

連続関数 $f(x)$ $(x \in \mathbb{R})$ に対し，

$$f(x+y) = f(x) + f(y) \qquad (x, y \in \mathbb{R}) \tag{2.13}$$

が成り立つとき，関数 $f(x)$ は，

$$f(x) = ax \qquad (x \in \mathbb{R}) \tag{2.14}$$

の形に限ることを証明せよ．ただし，a は定数とする．

————————————————————————————————

解答. (1) 定数 a を，

$$a = f(1)$$

により定める．このとき，(2.14) が成り立つことを以下で証明する．

まず，x が自然数 $1, 2, \cdots$ に等しいとき (2.14) が成り立つことを，帰納法により証明する．定数 a の取り方から，$x = 1$ に対し (2.14) が成り立つことは明らかである．次に，x の値が自然数 n に等しいとき (2.14) が成り立つと仮定する．すると，(2.13) より，

$$f(n+1) = f(n) + f(1) = an + a = a(n+1),$$

すなわち，(2.14) は，$x = n+1$ に対しても成立することが分かる．以上で，x が自然数の場合における (2.14) の証明が終わった．

次に，x がゼロ以下の整数である場合を考えよう．$0 = 0 + 0$ であるから，(2.13) より

$$f(0) = f(0+0) = f(0) + f(0) = 2f(0)$$

が従う．そして，これより直ちに

$$f(0) = 0$$

を得る．さて，$-n$ $(n > 0)$ を負の整数としよう．すると，

$$0 = f(0) = f(n + (-n)) = f(n) + f(-n)$$

であるから，

$$f(-n) = -f(n) = -an$$

が従う．これで，x がゼロ以下の整数である場合にも (2.14) が成り立つことが分かった．結局，任意の整数 x に対し (2.14) が成り立つことが示された訳である．

さらに，x が有理数である場合にも (2.14) が成り立つことを証明したい．そのために，任意の自然数 $m = 1, 2, \cdots$ に対し，$f(mx) = f(x+\cdots+x) = f(x)+\cdots+f(x) = mf(x)$ すなわち，

$$f(mx) = mf(x) \tag{2.15}$$

が成り立つことに注意しよう．また，この等式は非正の整数 m に対しても成り立つことが容易に分かる．

さて，p, q（ただし，$q > 0$）を互いに素な整数としよう．すると，(2.15) より，

$$f\left(\frac{p}{q}\right) = pf\left(\frac{1}{q}\right)$$

を得る．一方，再び (2.15) によれば，

$$qf\left(\frac{1}{q}\right) = f(1) = a, \quad \text{すなわち，} \quad f\left(\frac{1}{q}\right) = a\frac{1}{q}$$

である．以上をあわせると，

$$f\left(\frac{p}{q}\right) = a\frac{p}{q},$$

すなわち，(2.14) が任意の有理数 $x = p/q$ に対し成立することが分かる．

最後に，任意の実数 x に対し (2.14) 成り立つことを証明したい．いま，x が無理数である場合を考えれば十分である．このとき，各項が有理数であるような数列 $\{x_n\}$ で，$x_n \to x$ $(n \to \infty)$ であるようなものをとることができる．[17] このとき，関数 $f(x)$ の連続性によれば，

$$f(x_n) \to f(x) \qquad (n \to \infty)$$

である．一方，x_n が有理数であることから，

$$f(x_n) = ax_n \longrightarrow ax \qquad (n \to \infty)$$

が従う．以上から，実数 x に対しても (2.14) が成り立つことが分かった．　□

[17]実際，x を 10 進小数表記し，それを小数点以下第 n 桁で打ち切ることにより得られる有限小数を x_n とすればよい．

上の問題と同じやり方で，以下を示すことが可能です．

事実 2.13.

任意の実数 x に対し正の実数を対応させる連続関数 $f(x)$ に対し

$$f(x+y)=f(x)f(y) \qquad (x,y\in\mathbb{R}) \tag{2.16}$$

が成り立つならば，関数 $f(x)$ は，

$$f(x)=a^x \qquad (x\in\mathbb{R}) \tag{2.17}$$

の形に限る．ただし，a は正の定数とする．

ところで，ここで指数関数 a^x（ただし，$a>0$）がどう定義されたかを思い出してみたいと思います．まず，非負の整数 n に対しては，

$$a^0=1, \qquad a^{n+1}=a\cdot a^n \quad (n=0,1,\cdots)$$

というふうに帰納法を用いて定義します．あるいは，とくに n が正の整数のときには，

$$a^n=\underbrace{a\cdot\cdots\cdot a}_{n\text{ 個}}$$

としてもよいでしょう．さらに，負の整数 $-n$ $(n>0)$ に対しては，

$$a^{-n}=\frac{1}{a^n}$$

と定義します．次に，有理数 p/q（ただし，p, q は互いに素な整数；また q は正であるとする）に対し $a^{p/q}$ を定義したいと思います．そのために，まず，$x^q=a$ なる正の実数 x がただひとつ存在する（このことは，関数 x^q $(x>0)$ の狭義単調増加性および中間値の定理からの帰結です）ことに着目します．そのような x を $a^{1/q}$ であると定義します：

$$(a^{1/q})^q=1, \qquad a^{1/q}>0.$$

そして，

$$a^{p/q}=(a^{1/q})^p$$

とおくことにより，a の有理数乗 $a^{p/q}$ を定義します．

以上で有理数 x に対しては a^x が全く代数的な仕方で定義されることを見ました．それでは，x が無理数の場合はどうでしょう．実は，次のようなことが言えます．[18]

[18]証明に関しては，以下の本を参考にして下さい：

・小林昭七著『微分積分読本 – 1変数 –』裳華房，2000.

事実 2.14.

各項が有理数であるような数列 $\{x_n\}$ に対し，もしそれが収束するならば，数列 $\{a^{x_n}\}$ も収束する．また，ふたつの有理数列 $\{x_n\}, \{x'_n\}$ がある共通の極限に収束するならば，$a^{x_n}, a^{x'_n}$ も共通の極限に収束する．

いま，x を任意の実数としたとき，それに収束するような有理数列 $\{x_n\}$ をとることができます．そのとき，事実2.14 により，a^{x_n} がある実数に収束することが分かります．しかもその極限が有理数列 $\{x_n\}$ のとり方に依らないことも，事実2.14 により保証されます．そこで，その極限を a^x と定義します：

$$a^x = \lim_{n\to\infty} a^{x_n}, \qquad \text{ただし，}\quad x = \lim_{n\to\infty} x_n, \quad x_n \in \mathbb{Q}\,(n = 0, 1, \cdots).$$

以上が指数関数 a^x の厳密な定義です．[19]

いずれにせよ，厳密な仕方で関数を定義するのは，結構やっかいな仕事であることが理解して頂けたのではないかと思います．[20]

最後に重要な注意をひとつ．問題 2.12 で証明されたことも，事実 2.13 において主張されていることも，ともに関数 $f(x)$ の連続性を仮定してはじめて成り立つものです．言い換えれば，連続性を仮定しなければ，反例が見つかります．実際，例えば問題 2.12 において関数 $f(x)$ の連続性を仮定しないとしましょう．その場合にも，$a = f(1)$ の値を使って有理数 $x \in \mathbb{Q}$ に対する関数 $f(x)$ の値が $f(x) = ax$ と表されることが示せます．しかし，例えば無理数 $x = (1+\sqrt{5})/2$ における関数 $f(x)$ の値は，$a = f(1)$ の値によらず自由に定めることが可能です．連続性の仮定なしに，関数 $f(x)$ は $a = f(1)$ の値によって決定されることはありません．こういったところにも，連続な関数とそうでない関数との間に重大な差異が見出されます．

本文で区間縮小法の応用をいくつか紹介しました．それ以外にも応用があります．ここではそんな応用のうち，とくに重要なものを演習問題として出題したいと思います．まずは準備として，以下のような定義が必要となります．数列 $\{a_n\}$ が与えられたとしましょう．その項 $a_0, a_1, \cdots$ の中から無限個の項を取り出し，順

[19] あるいは，これとは別の定義として，巾級数を利用する方法があります．そのためには，まず e^x を

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

により定義します（もちろん，これを定義として採用するためには，この定義式の右辺がどんな実数 x に対しても収束することを言わねばなりません）．次に，自然対数 $\log x$ を e^x の逆関数として定義し，そして最後に，任意の $a > 0$ に対し，a^x を $a^x = e^{x \log a}$ により定義する，といった具合です．

[20] このノートでは，これ以上，関数の厳密な定義について深入りはしないことにします．

番を変えずに並べ上げたものを，元の数列の**部分列**と呼びます．例えば，数列

$$1,\ 2,\ 3,\ 4,\ \cdots$$

に対し，以下のいずれもが部分列になります：

$$2,\ 4,\ 6,\ \cdots\ ;$$
$$1,\ 2,\ 4,\ 8,\ \cdots\ ;$$
$$2,\ 3,\ 5,\ 7,\ 11,\ \cdots\ .$$

厳密な定義は以下のようになります．数列 $\{a_n\}$ および狭義単調増加な非負の整数列

$$0 \le k_0 < k_1 < k_2 < \cdots$$

が与えられたとき，新たな数列 $\{b_n\}$ を

$$b_n = a_{k_n} \qquad (n = 0, 1, \cdots)$$

により定義することができます．この新たな数列を元の数列 $\{a_n\}$ の**部分列**と呼びます．

問題 2.15. ――――――――――――――――――――――――――――

有界数列は収束部分列を有することを証明せよ．

――――――――――――――――――――――――――――――――――

解答． 数列 $\{c_n\}$ が $a \le c_n \le b$ $(n = 0, 1, \cdots)$ を満たすとしよう．各 $k = 0, 1, \cdots$ に対し実数 a_k, b_k（ただし，$a_k \le b_k$）および数列 $\{c_{k,n}\}$（ただし，$a_k \le c_{k,n} \le b_k$）を以下のように帰納的に定義する．$a_0 = a$, $b_0 = b$, $c_{0,n} = c_n$ $(n = 0, 1, \cdots)$ とする．いま，a_k, b_k, $\{c_{k,n}\}$ が定義されたとしよう．このとき，a_{k+1}, b_{k+1}, $\{c_{k+1,n}\}$ を次の規則に従い定義する．$m_k = (a_k + b_k)/2$ とおく．また，

$$A_k = \{n : c_{k,n} \in [a_k, m_k]\}, \qquad B_k = \{n : c_{k,n} \in (m_k, b_k]\}$$

とする．[21] これらふたつの集合のうち少なくとも一方は無限集合であることに注意しよう．

(i) A_k が無限集合の場合．この場合には，$a_{k+1} = a_k$, $b_{k+1} = m_k$ とする．また，A_k の要素を小さい順に並べ

$$\ell_0 < \ell_1 < \cdots$$

[21] 集合 A_k, B_k は番号 n を集めて得られる集合であって，数列 $\{c_{k,n}\}$ の項を集めて得られる集合ではありません．とくにその要素は非負の**整数**です．

としたとき，$c_{k+1,n} = c_{k,\ell_n}$ $(n = 0, 1, \cdots)$ とする．

(ii) A_k が有限集合の場合．この場合には，$a_{k+1} = m_k$, $b_{k+1} = b_k$ とおく．またこの場合には，B_k は無限集合である．その要素を小さい方から順に並べて

$$\ell_0 < \ell_1 < \cdots$$

とし，$c_{k+1,n} = c_{k,\ell_n}$ $(n = 0, 1, \cdots)$ とおく．

各 $k = 0, 1, \cdots$ に対しこのように定義した実数 a_k, b_k, および数列 $\{c_{k,n}\}$ は以下の条件を満たす．

(i) $a_0 \le a_1 \le \cdots\cdots \le b_1 \le b_0$;

(ii) $b_k - a_k = (b-a)/2^k$;

(iii) 数列 $\{c_{k+1,n}\}$ は数列 $\{c_{k,n}\}$ の部分列である；

(iv) $a_k \le c_{k,n} \le b_k$ $(n = 0, 1, \cdots)$.

数列 $\{a_k\}$ は単調増加かつ上に有界，一方数列 $\{b_k\}$ は単調減少かつ下に有界であるから，これらの数列はともに収束する．しかも，(ii) によればそれらの極限は一致する．そこで，$c = \lim a_k = \lim b_k$ とする．いま，数列 $\{d_n\}$ を，$d_n = c_{n,n}$ $(n = 0, 1, \cdots)$ により定義する． この数列 $\{d_n\}$ はもとの数列 $\{c_n\}$ の部分列である．しかも (iv) より 数列 $\{d_n\}$ も c に収束する． □

問題 2.16. ——————————————————————————

有界数列が収束部分列を有することを用い，最大値・最小値の定理を証明せよ．

——————————————————————————————

解答. 有界閉区間 $[a, b]$ において定義された実数値連続関数 f が与えられたとしよう．この f が最大値を有することを証明する．そのためにまず以下のような定数 M が存在することを証明したい：

$$\text{すべての } x \in [a, b] \text{ に対し } f(x) \le M \text{ が成り立つ.} \qquad (*)$$

証明は背理法による．このような定数 M が存在しないと仮定しよう．すると，各 $n = 1, 2, \cdots$ に対し，$f(c_n) > n$ なる $c_n \in [a, b]$ が存在する．数列 $\{c_n\}$ は収束部分列を有する．それを $\{c_{k_n}\}$ と，またその極限を c とすれば，f の連続性より，数列 $\{f(c_{k_n})\}$ は実数 $f(c)$ に収束する．ところが一方 c_n の取り方によれば，$\{f(c_{k_n})\}$ は $n \to \infty$ のとき発散する．これは矛盾である．以上で，(*) を満たす定数 M が存在することが分かった．

条件 (*) を満たす定数 M としてとくに整数をとることもできる．いま，(*) を満たす整数のうち，最小のものを M_0 とする．[22] 最小性より，$f(x) > M_0 - 1$ を満たす $x \in [a, b]$ が存在することもわかる．さて，ふたつの数列 $\{\alpha_n\}$, $\{\beta_n\}$ を以下のように定義しよう．まず，$\alpha_0 = M_0 - 1$, $\beta_0 = M_0$ とする．次に，α_n, β_n がすでに定義されたとして，α_{n+1}, β_{n+1} を以下の仕方で定めよう．ただし，$\gamma_n = (\alpha_n + \beta_n)/2$ とする．

(a) $f(x) > \gamma_n$ なる $x \in [a, b]$ が存在する場合．このときには，$\alpha_{n+1} = \gamma_n$, $\beta_{n+1} = \beta_n$ とする．

(b) すべての $x \in [a, b]$ に対し $f(x) \le \gamma_n$ である場合．このときには $\alpha_{n+1} = \alpha_n$, $\beta_{n+1} = \gamma_n$ とおく．

このようにして定義したふたつの数列 $\{\alpha_n\}$, $\{\beta_n\}$ が以下を満たすことは明らかである．

(i) $\alpha_0 \le \alpha_1 \le \cdots\cdots \le \beta_1 \le \beta_0$;

(ii) $\beta_n - \alpha_n = 2^{-n}(\beta - \alpha)$;

(iii) 各 $n = 0, 1, \cdots$ に対し，$f(x_n) > \alpha_n$ なる $x_n \in [a, b]$ が存在する；

(iv) すべての $n = 0, 1, \cdots$ およびすべての $x \in [a, b]$ に対し $f(x) \le \beta_n$ が成り立つ．

条件 (i), (ii) から，これらふたつの数列 $\{\alpha_n\}$, $\{\beta_n\}$ は共通の極限を有することがただちに分かる．その共通の極限を γ とする．すると条件 (iv) から

$$f(x) \le \lim_{n\to\infty} \beta_n = \gamma$$

がすべての $x \in [a, b]$ に対し成り立つことが分かる．一方，条件 (iii) より得られた数列 $\{x_n\}$ は収束部分列 $\{x_{k_n}\}$ を持つ．その極限を $c \in [a, b]$ とする．すると，条件 (iii) より

$$f(c) = \lim_{n\to\infty} f(x_{k_n}) \ge \lim_{n\to\infty} \alpha_{k_n} = \gamma$$

が従う．以上で $\gamma = f(c)$ が最大値であることが結論される．

関数 f が最小値を有することも同様に証明できる． □

[22] 整数に限定しなければ，最小のものが存在するかどうかは分かりません．

第 3 章

1 変数関数の微分

A　微分係数の定義と計算方法

A.1　高校数学の復習

何はさておき，1 変数関数の微分の定義を復習しましょう．この章でも，関数の値はつねに実数であると仮定します．

定義 3.1. ————————————————————————————

$f(x)$ を開区間 I 上で定義された関数とする．もし，$a \in I$ に対し，極限

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{3.1}$$

が存在するならば，関数 $f(a)$ は $x = a$ において**微分可能**であると言われる．また，その極限値 (3.1) を，関数 $f(x)$ の点 $x = a$ における**微分係数**と呼び，記号

$$f'(a) \qquad \text{あるいは} \qquad \frac{df}{dx}(a)$$

で表す．さらに，すべての点 $x \in I$ において $f(x)$ が微分可能であるとき，単に関数 $f(x)$ は（区間 I 上）**微分可能**であるという言い方がなされる．また，各点 $x \in I$ に対し，その点における微分係数 $f'(x) = \dfrac{df}{dx}(x)$ を対応させることにより得られる関数を，関数 $f(x)$ の**導関数**と呼ぶ．

————————————————————————————————

実際に導関数を計算する場合によく使われる公式をまとめておきましょう．いずれも，高校時代に習ったことのあるものだと思います．

公式 3.2. ————————————————————————————

(1)　$(x^a)' = ax^{a-1}$	(2)　$(e^x)' = e^x$
(3)　$(\log\lvert x\rvert)' = \dfrac{1}{x}$	(4)　$(\sin x)' = \cos x$
(5)　$(\cos x)' = -\sin x$	(6)　$(\tan x)' = \dfrac{1}{\cos^2 x}$

公式 3.3.
関数 $f(x)$, $g(x)$ に対し，以下が成り立つ．

$$\begin{aligned}
&(1) \qquad \big(f(x)+g(x)\big)' = f'(x)+g'(x) \\
&(2) \qquad \big(f(x)g(x)\big)' = f'(x)g(x)+f(x)g'(x) \\
&(3) \qquad \left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2} \\
&(4) \qquad \big(g(f(x))\big)' = g'(f(x))f'(x)
\end{aligned}$$

与えられた関数 $f(x)$ に対し，$y = f(x)$ とおき，関数 $f(x)$ の x における微分係数，あるいは関数 $f(x)$ の導関数を，

$$\frac{dy}{dx}(x) \qquad \text{あるいは単に，} \qquad \frac{dy}{dx}$$

と表記することがあります．いま，公式 3.3 (4) におけるように，ふたつの関数 $f(x)$, $g(x)$ が与えられたとしましょう．このとき，$y = f(x)$, $z = g(y)$ とおけば，公式 3.3 (4) は，

$$\frac{dz}{dx}(x) = \frac{dz}{dy}(y)\cdot\frac{dy}{dx}(x), \quad \text{あるいは，} \quad \frac{dz}{dx} = \frac{dz}{dy}\cdot\frac{dy}{dx}$$

と書くことが可能です．とくに 2 番目の式を見てください．あたかも分数のかけ算における約分のように右辺の dy が「消去」できる，と解釈することが出来ます．このおかげで，この公式を容易に記憶することが可能です．

これらの公式の使い方を復習するために，次の問題を解いてみてください．

問題 3.4.
次の関数の導関数を計算せよ．

(1) $x^5 - 2x^3 + 7x^2 - x + 9$　　(2) $\dfrac{x}{x^2+1}$

(3) $e^x - \log x \quad (x > 0)$　　(4) $\sin x - \cos x + \tan x$

(5) $\sin x^2$

解答. (1) $5x^4 - 6x^2 + 14x - 1$.　(2) $\dfrac{(x^2+1) - x \cdot 2x}{(x^2+1)^2} = \dfrac{1-x^2}{(1+x^2)^2}$.
(3) $e^x - \dfrac{1}{x}$.　(4) $\cos x + \sin x + \dfrac{1}{\cos^2 x}$.　(5) $(\cos x^2) \cdot 2x = 2x \cos x^2$.　□

まず，微分可能性と，前章で議論した連続性との間にどのような関係があるのかを確かめておきましょう．

命題 3.5.
開区間 I 上で定義された関数 $f(x)$ が，点 $a \in I$ において微分可能であれば，$f(x)$ は点 a において連続である．とくに $f(x)$ が I 上いたるところで微分可能であれば，$f(x)$ は連続関数である．

証明. $h \neq 0$ に対し，$(f(a+h) - f(a))/h$ を h の関数と見なし，$\delta(h)$ と書くことにする．定義によれば $\delta(h)$ は，$h \to 0$ のとき，(有限の値）$f'(a)$ に収束する．したがって，$f(a+h) - f(a) = h\delta(h) \to 0$ $(h \to 0)$ が従う．さて，関数の連続性の定義 2.5 によれば，$f(x)$ が $x = a$ において連続であるとは，以下が成り立つことであった：$\{a_n\}$ を，$a_n \in I$ $(n = 1, 2, \cdots)$，かつ $a_n \to a$ なる任意の数列としたとき，$f(a_n) \to f(a)$．さて，上記のような数列 $\{a_n\}$ が任意に与えられたとき，$h_n = a_n - a$ とおけば，$h_n \to 0$ $(n \to \infty)$ である．そしてこれより，実際に $f(a_n) - f(a) = f(a + h_n) - f(a) \to 0$ $(n \to \infty)$ が成り立つことが分かる．以上で，関数 $f(x)$ が点 $x = a$ において連続であることが示された．命題の後半の主張は自明である．　□

A.2　逆関数とその微分

まず，指数関数と対数関数の間にどのような関係があったか，復習しましょう．そのために，$y = e^x$ とおきます．すると，$x = \log y$ が成り立ちます．あるいは，$f(x) = e^x$, $g(y) = \log y$ $(y > 0)$ とするならば，

$$g(f(x)) = x, \qquad f(g(y)) = y \quad (y > 0)$$

が成り立ちます．指数関数と対数関数の間にあるこの関係を一般化することにより，逆関数の概念に至ります．

ただし，逆関数の定義を与えるに先だち，若干の準備が必要です．まずは，次の定義から．

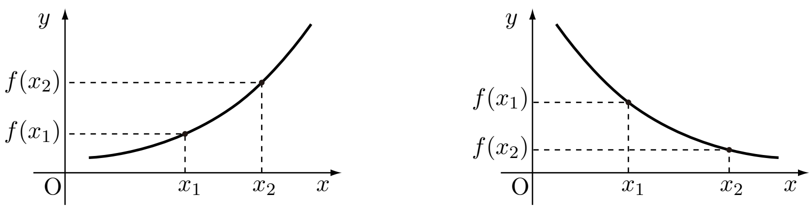
定義 3.6.
区間 I 上で定義された関数 $f(x)$ を考える．以下の条件が満たされるとき，関数

$f(x)$ は（**狭義**）**単調増加**であると言われる：

$$x_1, x_2 \in I \text{ に対し，} x_1 < x_2 \text{ ならば } f(x_1) < f(x_2).$$

また，以下が成り立つとき，関数 $f(x)$ は（**狭義**）**単調減少**であると言われる：

$$x_1, x_2 \in I \text{ に対し，} x_1 < x_2 \text{ ならば } f(x_1) > f(x_2).$$



(a) 単調増加関数　(b) 単調減少関数

図 3.1: 関数の単調性

グラフが右肩上がりならば単調増加，右肩下がりならば単調減少です．例えば，関数 $f(x) = x^2$ $(x \geq 0)$ は単調増加，関数 $f(x) = x^2$ $(x \leq 0)$ は単調減少，といった具合です．

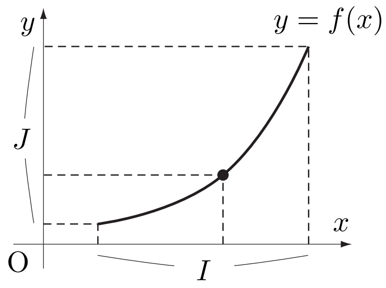


図 3.2: 単調関数とその逆関数

さて，逆関数の定義に進みましょう．区間 I 上で定義された関数 $f(x)$ に対し，それが単調増加，あるいは単調減少であるとします．さらに，

$$J = \{f(x) : x \in I\}$$

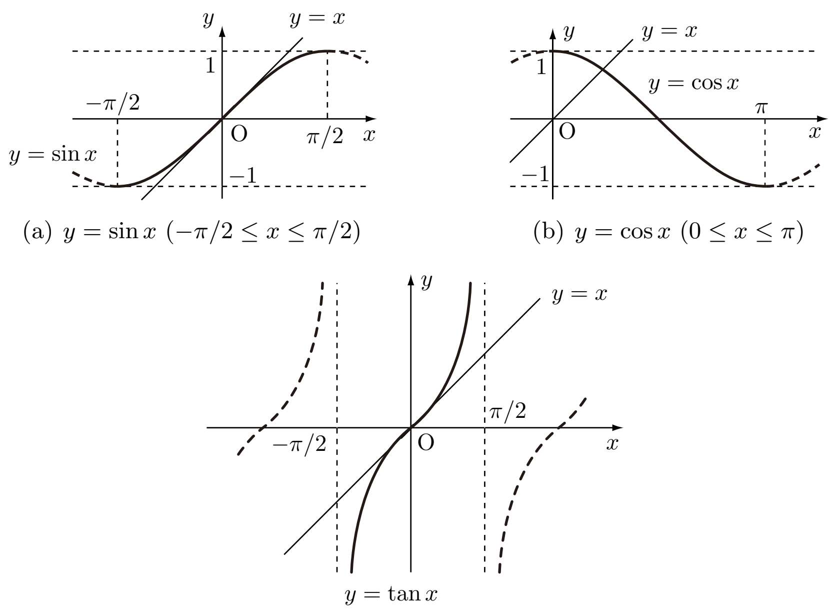
としましょう．このとき，各 $y \in J$ に対し，$y = f(x)$ なる $x \in I$ がちょうどひとつ存在します．そこで，各 $y \in J$ に対しこのような $x \in I$ を対応させる関数を考

えることが出来ます．その関数を元の関数 $f(x)$ の**逆関数**と言います．また，関数 $y = f(x)$ の逆関数を，しばしば，

$$x = f^{-1}(y)$$

と書きます．

いまとくに例として，$I = \mathbb{R}$, $f(x) = e^x$ の場合を考えましょう．確かに，指数関数 $f(x) = e^x$ は（狭義）単調増加関数です．また，この場合には $J = (0, \infty)$ です．そして，指数関数 $f(x) = e^x$ の逆関数として得られるのが，対数関数 $g(y) = \log y$ $(y \in J)$ に他なりません．対数関数 $g(y) = \log y$ の逆関数が指数関数 $f(x) = e^x$ であることにも注意してください．



(c) $y = \tan x$ $(-\pi/2 < x < \pi/2)$

図 3.3: 三角関数

逆関数の例をさらにいくつか挙げましょう．まず，正弦関数 $y = \sin x$ $(x \in \mathbb{R})$ を考えます．残念ながら，この関数は単調増加でも単調減少でもありません．しかし，その定義域を区間 $[-\pi/2, \pi/2]$ に制限すれば，単調増加になります．そして，その逆関数を

$$x = \operatorname{Arcsin} y \quad \text{ないし} \quad x = \operatorname{Sin}^{-1} y \qquad (y \in [-1, 1])$$

と書きます．同様に，余弦関数 $y = \cos x$ の区間 $[0, \pi]$ への制限として得られる単調減少関数の逆関数を

$$x = \mathrm{Arccos}\, y \quad \text{ないし} \quad x = \mathrm{Cos}^{-1}\, y \qquad (y \in [-1, 1])$$

と，また関数 $y = \tan x$ の区間 $(-\pi/2, \pi/2)$ への制限として得られる単調増加関数の逆関数を

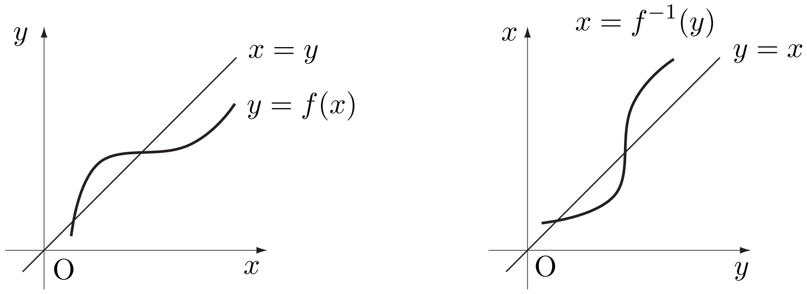
$$x = \mathrm{Arctan}\, y \quad \text{ないし} \quad x = \mathrm{Tan}^{-1}\, y \qquad (y \in \mathbb{R})$$

と書きます．また，これら3つの関数 $x = \mathrm{Arcsin}\, y$, $x = \mathrm{Arccos}\, y$, $x = \mathrm{Arctan}\, y$ を**逆三角関数**と総称することがあります．

問題 3.7.

3 つの逆三角関数のグラフを描け．

区間 I において定義された関数 $y = f(x)$ が，単調増加あるいは単調減少であるとします．そのとき，その関数の逆関数 $x = f^{-1}(y)$ のグラフは，元の関数 $y = f(x)$ のグラフを，直線 $y = x$ に関し対称に折り返すことに得られます．ただし，逆関数 $x = f^{-1}(y)$ においては，y が独立変数，x が従属変数ですから，そのグラフを描くにあたっては，横軸は y 軸，縦軸は x 軸となっていることに注意してください．

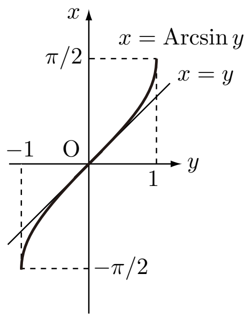


(a) 元の関数のグラフ　(b) 逆関数のグラフ

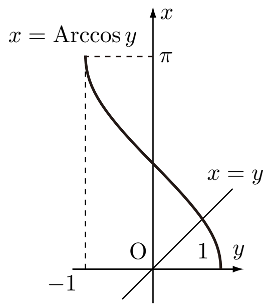
図 3.4: 逆関数のグラフ

とくに，問題 3.7 の解答として，図 3.5 を得ます．

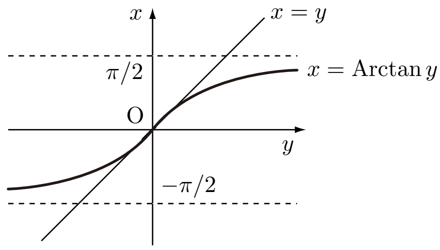
開区間上で定義された関数 $f(x)$ に対し，そのグラフ $y = f(x)$ を考えましょう．もし $f(x)$ が $x = a$ において微分可能であるならば，そのグラフの点 $(a, f(a))$ における接線が存在し，しかもその傾きがまさに微分係数 $f'(a)$ に等しいことは，皆さんすでにご存じのことと思います．これは，微分係数の定義に立ち返れば，納得のいくことだと思います．



(a) $x = \mathrm{Arcsin}\, y$



(b) $x = \mathrm{Arccos}\, y$



(c) $x = \mathrm{Arctan}\, y$

図 3.5: 逆三角関数のグラフ（問題 3.7 の解答）

いま，関数 $y = f(x)$ は単調増加ないし単調減少であると仮定します．すでに述べたとおり，逆関数 $x = f^{-1}(y)$ のグラフは，元の関数 $y = f(x)$ のグラフを直線 $y = x$ に関し対称に折り返したものになっています．すると，この操作により，元の関数 $y = f(x)$ の点 $(a, f(a))$ における接線（これを ℓ と書くことにしましょう）は，逆関数 $x = f^{-1}(y)$ の点 $(b, f^{-1}(b))$ （ただし，$b = f(a)$）における接線（これを m と書くことにします）に移ります．一方，ℓ の傾きを γ とすれば，m の傾きは $1/\gamma$（ただし，$\gamma \neq 0$ と仮定）になります．最後に，接線の傾きが微分係数に等しいことを思い出せば，次の命題が成り立つことが分かります．

命題 3.8.

単調増加ないし単調減少な関数 $y = f(x)$ が $x = a$ において微分可能であり，しかも $f'(a) \neq 0$ であるとする．このとき，逆関数 $x = f^{-1}(y)$ は $y = b$（ただし，$b = f(a)$） において微分可能であり，しかも

$$(f^{-1})'(b) = \frac{1}{f'(a)} \tag{3.2}$$

が成り立つ．

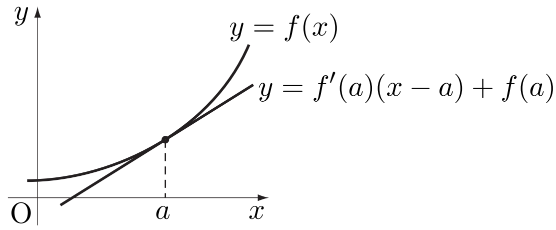


図 3.6: 微分係数と接線の傾き

この命題の中の公式 3.2 を，

$$\frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}}$$

と表すことがあります．この公式を覚えるにあたっては，この書き方の方が好都合でしょう．

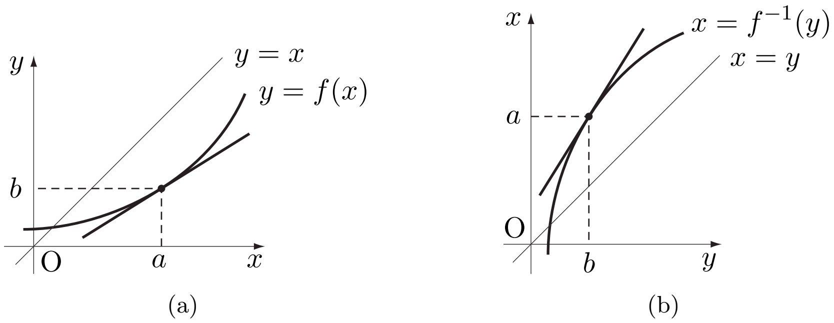


図 3.7: 逆関数の微分係数

B 微分法の応用とその理論的背景

B.1 1変数関数に対する極値問題

再び高校数学の復習として，問題をひとつ出します．

問題 3.9. ――――――――――――――――――――――――――――

次の関数の極値を求めよ：

$$f(x) = \frac{2x^3 + 3x^2 - 1}{x^2 + 1} \qquad (x \in \mathbb{R}).$$

――――――――――――――――――――――――――――――――

解答. まず，関数 $f(x)$ の導関数を求めよう．

$$f'(x) = \frac{(6x^2+6x)(x^2+1)-(2x^3+3x^2-1)\cdot 2x}{(x^2+1)^2} = \frac{2x(x+1)(x^2-x+4)}{(x^2+1)^2}$$

である．ここで，$f'(x)=0$ を x に関する方程式とみなしそれを解けば，$x=0,-1$ を得る．さらに 2 次関数 x^2-x+4 は常に正の値をとることに注意すれば，以下のような増減表を得る．

x	$\cdots$	-1	$\cdots$	0	$\cdots$
$f(x)$	$\nearrow$	極大	$\searrow$	極小	$\nearrow$
$f'(x)$	$+$	0	$-$	0	$+$

よって，

（答）　関数 $f(x)$ は $x=-1$ において極大，極大値は $f(-1)=0$.
また，$x=0$ において極小，極小値は $f(0)=-1$.

□

ちなみに，さらに $x\to\pm\infty$ のとき $f(x)\to\pm\infty$（複合同順）であることに注意すれば，関数 $y=f(x)$ のグラフの概形は図 3.8 のようであることが分かります．

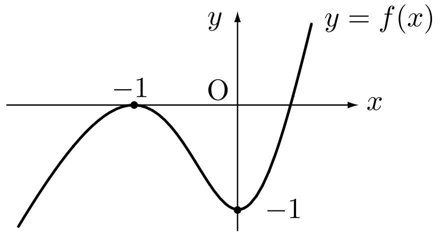


図 3.8:

ここで念のため，極値の定義の復習をしておきましょう．

定義 3.10.

区間 I 上で定義された関数 $f(x)$ を考える．また，a を区間 I 内の点とする．以下の条件が満たされるとき，関数 $f(x)$ は $x=a$ において**極大**であると言われる：

正の実数 $\epsilon>0$ を十分小さくとったとき，関数 $f(x)$ の区間 $I\cap(a-\epsilon,a+\epsilon)$ への制限は，点 $x=a$ において最大である．

関数 $f(x)$ が $x=a$ において極大であるとき，その点における $f(x)$ の値 $f(a)$ を**極大値**と呼ぶ．一方，以下の条件が満たされるとき，関数 $f(x)$ は $x=a$ において**極小**であると言われる：

正の実数 $\epsilon > 0$ を十分小さくとったとき，関数 $f(x)$ の区間 $I \cap (a-\epsilon, a+\epsilon)$ への制限は，点 $x = a$ において最小である．

関数 $f(x)$ が $x = a$ において極小であるとき，その点における $f(x)$ の値を**極小値**と呼ぶ．

問題 3.9 の答えとして得た極大値，極小値が，本当にこの定義の条件を満たしていることを，まずは確認して下さい．また，極大ではあっても必ずしも最大ではないこと，また極小ではあっても必ずしも最小ではないことも，この例から分かります．ちなみに，逆はいつでも成立します．すなわち，関数 $f(x)$ に対し，それが $x = a$ で最大であれば，その点において極大です．また最小であれば極小です．極大というのは，十分小さな範囲でみれば最大，と言うことです．富士山は世界で一番高い山ではないけれど，日本国内に制限すれば一番です．「井の中の蛙」という表現も，これに類するものかも知れません．

問題 3.9 の解答で述べた仕方で 1 変数関数に対する極値問題を解くことは，皆さんもう十分に身についていることと思います．では，その理論的背景に関する理解も十分でしょうか．1 変数関数の極値問題に対する解法の背景には，以下に述べるふたつの定理（定理 3.11，および定理 3.12）があります．

定理 3.11.

開区間 I において定義された微分可能関数 $f(x)$ を考える．もしこの関数が $x = a$ において極大あるいは極小であるならば，

$$f'(a) = 0$$

が成り立つ．

この定理の証明を少しだけ後まわしにして，極値問題の解法について説明をしたいと思います．微分可能関数 $f(x)$ に対し，$f'(x) = 0$ となる点 $x = a$ を，関数 $f(x)$ の**臨界点**と呼びます．すると，定理 3.11 は，

$$f(x) \text{ が } x = a \text{ において極大ないし極小} \quad \Longrightarrow \quad x = a \text{ は } f(x) \text{ の臨界点}$$

と述べることが出来ます．ここで注意すべきは，一般には逆が成り立たないことです．実際，例えば，関数 $f(x) = x^3$ は $x = 0$ を臨界点として持ちますが，$x = 0$ において極大でも極小でもありません．しかし，いずれにせよ，関数 $f(x)$ の臨界点をすべて求めておけば，その中に関数 $f(x)$ が極値をとる点がもれなく含まれていることを，定理 3.11 は保証します．

関数 $f(x)$ に対しその臨界点をすべて求めることが出来たとしましょう．すると，次にすべきは，各臨界点において，関数 $f(x)$ が実際に極値をとるかどうか，また極値をとる場合には，極大値をとるのか極小値をとるのかを判定する必要があります．その際に利用されるのが次の定理です．

定理 3.12.
区間 I において定義された微分可能関数 $f(x)$ を考える．I 上のすべての点 x において $f'(x) > 0$ であるならば，関数 $f(x)$ は I 上狭義単調増加である．また，I 上のいたる点 x において $f'(x) < 0$ であるならば，$f(x)$ は I 上狭義単調減少である．

この定理の証明は，次の節で行うことにします．

さて，極値問題に対する解法の解説を続けましょう．定理 3.12 から直ちに次の系が従います．

系 3.13.
開区間 I において定義された微分可能関数 $f(x)$ に対し，$x = a$ がその臨界点だとする．正の数 ϵ を適当にとったとき，

$$\begin{array}{lll} a - \epsilon < x < a & \text{において} & f'(x) > 0, \quad \text{かつ} \\ a < x < a + \epsilon & \text{において} & f'(x) < 0 \end{array}$$

であるならば，$f(x)$ は $x = a$ において極大である．また，逆に，適当にとった正の数 ϵ に対し，

$$\begin{array}{lll} a - \epsilon < x < a & \text{において} & f'(x) < 0, \quad \text{かつ} \\ a < x < a + \epsilon & \text{において} & f'(x) > 0 \end{array}$$

であるならば，$f(x)$ は $x = a$ において極小である．

この系を用いて，臨界点において極値をとるかどうか，また極値をとる場合には，極大となるのかそれとも極小となるのかが，判定されます．[1] ちなみに，後に§B.4 において，別の判定方法を紹介する予定です．

すでに，前の章でも述べたとおり，何らかの関数の最大値・最小値を求めたり，あるいはその関数が最大値・最小値を取る点を決定したりすることは，応用上極めて重要な問題です．もし問題とする関数が連続であり，しかもその定義域が有界閉

[1] この系によって判定が出来ない場合もあります．どのような場合判定不能となるのか，各自考えてみて下さい．

集合であるならば，最大値・最小値の定理 により，その関数が最大値・最小値を有することが保証されます．しかし，最大値・最小値の定理，あるいは前章で与えたその証明は，最大値・最小値の求め方や，あるいは関数が最大値・最小値をとる点の求め方を教えてはくれません．しかし，扱っている関数が微分可能な場合には，興味の対象を最大値・最小値だけでなく極大値・極小値まで広げると，上で述べた仕方により極値を決定することが出来ます．[2]

さて，以上で見たとおり，問題 3.9 の解法は，上のふたつの定理 3.11, 3.12 によって正当化されます．また，後に見るように，2番目の定理は 1 番目の定理から導かれます．その意味で，定理 3.11 は 1 変数関数に対する極値問題においてもっとも基本的なものであると言うことができます．しかし，それだけでなく，実はこの章でこの後扱うすべての定理が，実は最終的にはこの定理 3.11 から導出されます． 定理 3.11 はそのくらい重要な定理なのです．しかし，一方，その証明は案外簡単です．その証明をまず済ませてしまいましょう．

定理 3.11 の証明. 関数 $f(x)$ が $x = a$ において極大の場合を考える（極小である場合も，同じ方法で証明できるので，その場合の証明は省略する）．まず，

$$f'(x) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

であったことを思い起こそう．いま，$|h|$ が十分小さいとき，$a + h \in I$ であり，しかも

$$f(a+h) \leq f(a)$$

が成り立つ．したがって，さらに $h > 0$ である場合には，$\big(f(a+h) - f(a)\big)/h \leq 0$ である．これより，

$$\lim_{h\to +0} \frac{f(a+h) - f(a)}{h} \leq 0$$

を得る．一方，$|h|$ が十分小さく，しかも $h < 0$ である場合には，$\big(f(a+h) - f(a)\big)/h \geq 0$ であり，したがって，

$$\lim_{h\to -0} \frac{f(a+h) - f(a)}{h} \geq 0$$

である．以上から，

$$f'(x) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = 0$$

でなければならないことが従う．　□

[2]ただし，求めた極大値・極小値が，最大値・最小値になるかどうかを決定するにあたっては，別の議論が必要です．

B.2　平均値の定理

極値問題の解法の正当化に必要であったもうひとつの定理，定理 3.12 の証明には，平均値の定理と呼ばれる定理が必要になります．[3] この平均値の定理も，1 変数関数の微分法に関わるいろいろな定理の証明に顔を出す，重要な定理です．

定理 3.14 (平均値の定理). 関数 $f(x)$ を閉区間 $[a,b]$ （ただし，$a<b$）において定義された連続関数とする．ただし，$a<b$ とする．さらに，$f(x)$ は開区間 (a,b) 上いたるところで微分可能であるとする．このとき，

$$f'(c)=\frac{f(b)-f(a)}{b-a} \tag{3.3}$$

なる点 c が区間 (a,b) 内に存在する．

等式 (3.3) の右辺は，2 点 $(a,f(a))$, $(b,f(b))$ を通る直線の傾きに一致します．その直線を適当に平行移動することにより，関数 $f(x)$ のグラフに接するようにすることが出来る，ということを平均値の定理は述べているわけです．

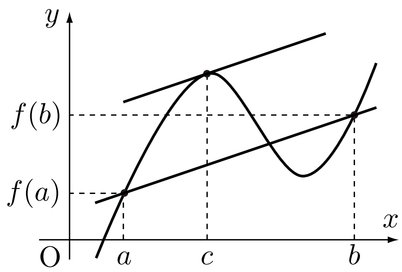


図 3.9: 平均値の定理

まずは，この平均値の定理を仮定して定理 3.12 の証明を終わらせてしまいましょう．

定理 3.12 の証明. 第一の場合，すなわち，区間 I のいたる点 x において $f'(x)>0$ である場合を考える（残るもうひとつの場合も同じ仕方で扱うことが可能であるので，その場合の証明は省略する）．証明は背理法による．すなわち，区間 I 内に，$a<b$ かつ $f(a)\geq f(b)$ なる点 a, b が存在したと仮定しよう．このとき，もちろん

$$\frac{f(b)-f(a)}{b-a}$$

[3] その他に積分を利用した証明も可能です．

の値はゼロまたは負である．一方，平均値の定理によれば，微分係数 $f'(c)$ の値がこれと一致するような点 c が区間 I 内に存在せねばならない．すると，$f'(c) \leq 0$ となり，仮定 $f'(x) > 0$ $(a < x < b)$ に反することになる．以上で定理 3.12 の証明が完了した． □

平均値の定理において，とくに $f(a) = f(b)$ と仮定したものをロルの定理と呼びます．

定理 3.15 (ロルの定理). 関数 $f(x)$ を有界閉区間 $[a, b]$ で定義された連続関数とする．ただし，$a < b$ とする．また，さらに開区間 (a, b) のいたる点において，関数 $f(x)$ は微分可能であるとする．もし，$f(a) = f(b)$ であるならば，$f'(c) = 0$ なる点 c が区間 (a, b) 内に存在する．

実は，この平均値の定理の特殊形であるロルの定理から，平均値の定理を簡単に導くことができます．実際，関数 $f(x)$ を平均値の定理のようにとってきましょう．そのとき，新たな関数 $g(x)$ を，

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

で定義すれば，この関数はロルの定理における仮定をすべて満たします．したがって，

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0, \qquad \text{すなわち} \qquad f'(c) = \frac{f(b) - f(a)}{b - a}$$

なる $c \in (a, b)$ の存在が，ロルの定理により保証されます．以上で，ロルの定理から平均値の定理が導かれることが示されました．

したがって，平均値の定理の証明を完成させるには，ロルの定理を証明すればいいことになります．さて，ロルの定理 3.15 を証明しましょう．

定理 3.15 の証明. 関数 $f(x)$ をロルの定理にあるようにとる．とくに関数 $f(x)$ は有界閉区間 $[a, b]$ 上で連続である．したがって，最大値・最小値の定理（定理 2.8）により，$f(x)$ は最大値，最小値をとる．最大値，最小値がともに $f(a) = f(b)$ の値と一致する場合には，$f(x)$ は区間 $[a, b]$ 上一定で，したがってとくに区間 $[a, b]$ 上のいたる点 x において $f'(x) = 0$ が成り立つ．すなわち，この場合ロルの定理は自明である．そこで，最大値，最小値のいずれか少なくとも一方 – それを $f(c)$ $(c \in (a, b))$ とする – が，$f(a) = f(b)$ とは異なる値をとると仮定しよう．すると，このときにも定理 3.11 より $f'(c) = 0$ であることが分かる．以上で，ロルの定理の証明が完了した． □

ご覧の通り，ロルの定理，および平均値の定理の証明は意外と簡単です．そしてその証明において本質的な役割を果たしたのが，定理 3.11 および連続関数に対する最大値・最小値の定理 2.8 でした．

B.3　ロピタルの定理

微分法の最初の応用として，1 変数関数に対する極値問題の解法を，§B.1 において復習しました．一方，§B.2 においては，その解法の理論的な背景のひとつとして，平均値の定理について学びました．さて，次の話題はロピタルの定理です．皆さん恐らくすでに高校などで勉強したおぼえがあるのではないかと思います．これも，微分法の応用のひとつです．その証明には前出の平均値の定理が用いられます．

定理 3.16 (ロピタルの定理).

関数 $f(x)$, $g(x)$ を原点 0 を含む開区間で定義された $f(0)=g(0)=0$ なる微分可能関数とする．さらに，導関数 $f'(x)$, $g'(x)$ は $x=0$ において連続，かつ $g'(0)\neq 0$ であると仮定する．このとき，

$$\lim_{x\to 0}\frac{f(x)}{g(x)}=\frac{f'(0)}{g'(0)}$$

が成り立つ．

先ほど述べたとおり，このロピタルの定理も，平均値の定理から導くことが出来ます．以下で実際にそれを見てみましょう．いま，極限

$$\lim_{x\to 0}\frac{f(x)}{g(x)}$$

を求めよ，という問題を考えます．ただし，$f(x)$, $g(x)$ は原点 $x=0$ を含む開区間で定義された関数とします．さらに，関数 $f(x)$, $g(x)$ が $x=0$ において連続であると仮定します．しかも $g(0)\neq 0$ であるならば，

$$\lim_{x\to 0}\frac{f(x)}{g(x)}=\frac{f(0)}{g(0)}$$

であることがただちに分かります．また，$g(0)=0$ ではあるが $f(0)\neq 0$ である場合には，$f(x)/g(x)$ は発散します．

それでは，

$$f(0)=g(0)=0$$

の場合はどうでしょう．この場合には，極限 $\lim\limits_{x\to 0} f(x)/g(x)$ はいわゆる不定形になります．しかしそれにも関わらず，微分法を用いて極限を計算できることがありま

す．いまさらに，関数 $f(x)$, $g(x)$ はともに微分可能であると仮定します．すると，平均値の定理により，$|x|$ が十分小さい x に対し，以下の等式を満たす $\xi, \bar{\xi}$ が原点 0 と点 x の間に存在することが分かります（$f(0)=g(0)=0$ と仮定していたことを思い起こしてください）：

$$f(x)=f'(\xi)x, \qquad g(x)=g'(\bar{\xi})x.$$

したがって，

$$\frac{f(x)}{g(x)}=\frac{f'(\xi)}{g'(\bar{\xi})}$$

が成り立ちます．ここで，$x \to 0$ とすると，$\xi, \bar{\xi} \to 0$ となることに注意しましょう．さて，ここでさらに導関数 $f'(x)$, $g'(x)$ が $x=0$ において連続，かつ $g'(0) \neq 0$ であると仮定します．すると，$\lim_{\xi \to 0} f'(\xi)=f'(0)$, $\lim_{\bar{\xi} \to 0} g'(\bar{\xi})=g'(0)$,

$$\lim_{x \to 0}\frac{f(x)}{g(x)}=\frac{f'(0)}{g'(0)}$$

を得ます．以上の議論をまとめたのがロピタルの定理です．

ロピタルの定理を利用しての極限計算の具体例を，演習問題として出題します．

問題 3.17.

次の極限を計算せよ：

$$\lim_{x \to 0}\frac{\sin x}{e^{2x}-\cos x}.$$

解答. $f(x)=\sin x$, $g(x)=e^{2x}-\cos x$ とおく．$f(0)=g(0)=0$ である．また，これらふたつの関数ははともに無限回連続的に微分可能な関数であり，とくに，$f'(x)=\cos x$, $g'(x)=2e^{2x}+\sin x$ である．したがって，

$$\lim_{x \to 0}\frac{f(x)}{g(x)}=\frac{f'(0)}{g'(0)}=\frac{1}{2}$$

を得る． □

B.4　2 階導関数と凸性

いま，$f(x)$ を開区間 I で定義された関数とします．さらに，$f(x)$ は微分可能であるとしましょう．すると，その導関数 $f'(x)$ が再び区間 I 上の関数として定義されます．区間 I 上の関数として $f'(x)$ が微分可能であるとき，その導関数を $f''(x)$ と書き，$f(x)$ の **2 階導関数**と呼びます：

$$f''(x)=\bigl(f'(x)\bigr)'.$$

例えば,

$$(x^3)'' = 6x, \qquad (\sin x)'' = -\sin x$$

といった具合です．また，$y = f(x)$ とおいたとき，2 階導関数を

$$\frac{d^2y}{dx^2}(x), \quad \text{あるいは，単に} \quad \frac{d^2y}{dx^2}$$

と書くこともあります.

2 階導関数の符号はその関数のグラフの凸性に関係しています．まずは，凸性の定義から始めましょう（問題 2.11，およびそのあとに続く注釈を参考のこと）.

定義 3.18.

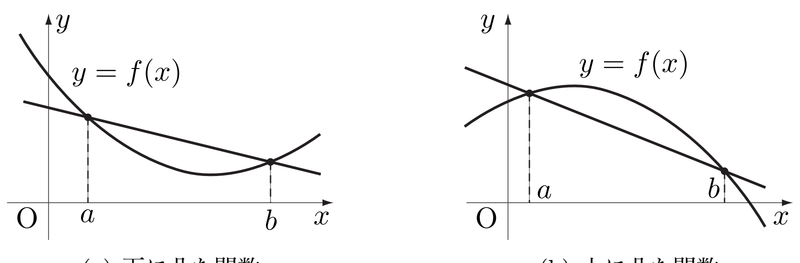
区間 I で定義された関数 $f(x)$ を考える．任意の実数 a, b（ただし，$a < b$）に対し，

$$f(x) \leq \frac{f(b)-f(a)}{b-a}(x-a) + f(a) \qquad (a \leq x \leq b) \tag{3.4}$$

が成り立つとき，関数 $f(x)$ は**下に凸**であると言われる．逆に，任意の実数 a, b（ただし，$a < b$）に対し，

$$f(x) \geq \frac{f(b)-f(a)}{b-a}(x-a) + f(a) \qquad (a \leq x \leq b) \tag{3.5}$$

が成り立つとき，関数 $f(x)$ は**上に凸**であると言われる.



(a) 下に凸な関数　(b) 上に凸な関数

図 3.10: 関数の凸性

条件 (3.4) は，次と同値です：

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b) \qquad (0 \leq t \leq 1). \tag{3.6}$$

同様に，条件 (3.5) は

$$f((1-t)a + tb) \geq (1-t)f(a) + tf(b) \qquad (0 \leq t \leq 1). \tag{3.7}$$

と同値です.

微分可能関数に関しては，以下の仕方で凸性を判定することも可能です.

命題 3.19.
いま，$f(x)$ を開区間 I で定義された微分可能関数とする．このとき，$f(x)$ が下に凸であるための必要十分条件は，以下が成り立つことである：区間 I 内の任意の点 a に対し，

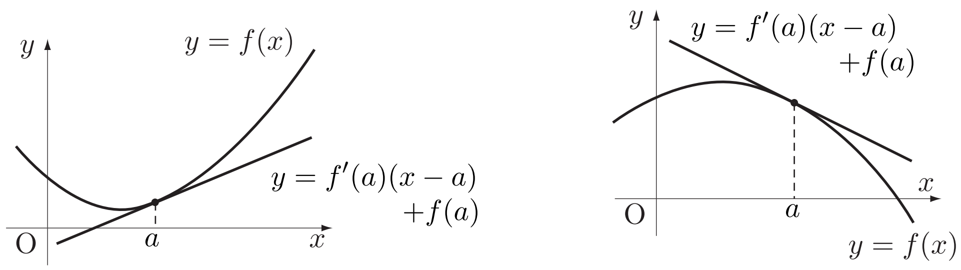
$$f(x) \geq f'(a)(x-a) + f(a) \qquad (x \in I). \tag{3.8}$$

また逆に，関数 $f(x)$ が上に凸であるための必要十分条件は，区間 I 内のすべての点 a に対し，

$$f(x) \leq f'(a)(x-a) + f(a) \qquad (x \in I) \tag{3.9}$$

が成り立つことである．

関数 $y = f(x)$ のグラフはその点 $(a, f(a))$ における接線 $y = f'(a)(x-a) + f(a)$ の上側にある，ということを式 (3.8) は意味しています．また，逆に (3.9) は，関数 $y = f(x)$ のグラフはその各点における接線の下側に位置するということを意味しています．



(a) 下に凸な関数　(b) 上に凸な関数

図 3.11: 微分可能関数に対する凸性の判定

命題 3.19 の証明. 命題の第 1 の主張，すなわち，微分可能関数 $f(x)$ に対し，それが下に凸であることと，すべての a に対し (3.8) が成り立つことが同値であることを証明する（第 2 の主張の証明は省略する）．

その準備として，区間 I 上で定義された連続関数 $f(x)$ に対し，それが下に凸であるならば，以下の主張が成り立つことを示す：任意の $a, b \in I$（ただし，$a < b$）に対し，

$$f(x) \geq \frac{f(b) - f(a)}{b - a}(x-a) + f(a) \qquad (x \in I \setminus [a, b]). \tag{3.10}$$

[4] [5]

さて，連続関数 $f(x)$ が下に凸であると仮定しよう．さらに，$a, b \in I$ が $a < b$ を満たすよう任意に与えられたとする．このとき，(3.10) が $x = c \in I \setminus [a, b]$ に対し成り立つことを示したい．以下，$c < a$ の場合を考える．2 点 $(c, f(c)), (b, f(b))$ を通る直線 ℓ を考える（図 3.12）．関数 $f(x)$ は下に凸であると仮定されていたので，関数 $y = f(x)$ のグラフは区間 (c, b) 上においてはその直線 ℓ の下側になければならない．とくに，関数 $f(x)$ のグラフ上の点 $(a, f(a))$ は直線 ℓ の下側になければならない．すなわち，

$$f(a) \leq \frac{f(b) - f(c)}{b - c}(a - c) + f(c)$$

が成り立つ．このことより，

$$f(c) \geq \frac{f(b) - f(a)}{b - a}(c - a) + f(a)$$

が従う．これは，不等式 (3.10) において $x = c$ とおいたものにほかならない．残る

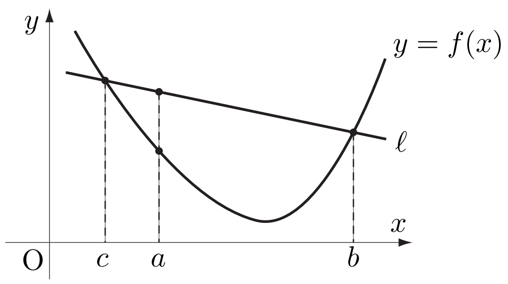


図 3.12: (3.10) の証明

もうひとつの場合，$c > b$ の場合も証明は同様である．以上で，連続関数 $f(x)$ に対し，それが下に凸であるならば，条件 (3.10) がすべての $a, b \in I$ $(a < b)$ に対し成り立つことが示された．

次に，命題 3.19 の前半分の証明を行う．以下，$f(x)$ を開区間 I で定義された微分可能関数であるとする．関数 $f(x)$ が下に凸であると仮定する．さらに，$a \in I$ が任意に与えられたとしよう．このとき，(3.8) が成り立つことを示したい．そのために，$b \in I$ を $b \neq a$ となるようとる．すると，関数 $y = f(x)$ のグラフ上の 2 点 $(a, f(a)), (b, f(b))$ を通る直線 ℓ が一意的に決まる（図 3.13）．いま，関数 $f(x)$ は下に凸であると仮定しているから，(3.10) が成り立つ．すなわち，独立変数 x が

[4]独立変数 x が区間 $[a, b]$ の外側を動くときの関数 $y = f(x)$ のグラフは，そのグラフ上の 2 点 $(a, f(b))$ を通る直線の上側にある，というのがこの最後の条件の幾何学的な意味です．

[5]実は，逆も成り立ちます．すなわち，任意の $a, b \in I$ $(a < b)$ に対し (3.10) が成り立つことが，連続関数 $f(x)$ が下に凸であるための必要十分条件になっています．

区間 $[a,b]$ の外側を動くときには，関数 $y=f(x)$ のグラフは直線 ℓ の上側にある．一方，$b\to a$ としたとき，直線 ℓ は，関数 $y=f(x)$ のグラフの点 $(a,f(a))$ における接線に収束する．したがって，関数のグラフ全体がその接線の上側にあることが分かる．

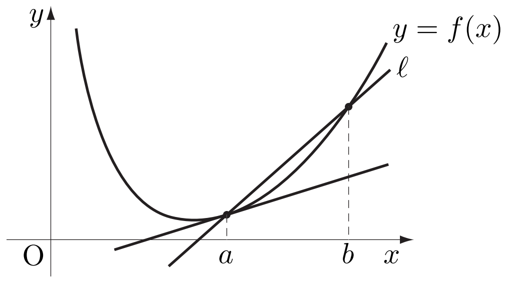


図 3.13: (3.8) の証明

最後に逆を示そう．すなわち，任意の $a\in I$ に対し (3.8) が成り立つと仮定する．その仮定のもと，$f(x)$ が下に凸であることを言えばよい．そのために，$a,b\in I$ を $a<b$ を満たすようとる．さらに，

$$g(x)=f(x)-\left\{\frac{f(b)-f(a)}{b-a}(x-a)+f(a)\right\}$$

により関数 $g(x)$ $(a\le x\le b)$ を定義する．関数 $g(x)$ の値がゼロ以下であることを示せばよい．ここで，$g(a)=g(b)=0$ であることに注意しよう．いま関数 $g(x)$ は $x=c$ において最大値をとるとしよう．もし $c=a$ または b の場合にはただちに $g(x)\le 0$ であることが分かる．そこで，$c\ne a,b$ の場合を考える．[6] すると，$0=g'(c)=f'(c)-\bigl(f(b)-f(a)\bigr)/(b-a)$, すなわち

$$f'(c)=\frac{f(b)-f(a)}{b-a} \tag{3.11}$$

を得る．したがって，

$$\begin{aligned}g(c)&=f(c)-\left\{\frac{f(b)-f(a)}{b-a}(c-a)+f(a)\right\}\\&=f(c)-\bigl\{f'(c)(c-a)+f(a)\bigr\}=\bigl\{f'(c)(a-c)+f(c)\bigr\}-f(a)\end{aligned}$$

が成り立つ．一方，(3.8) において，a を c に置き換え，その上で，$x=a$ とおけば，

$$f(a)\ge f'(c)(a-c)+f(c)$$

を得る．以上から，$g(c)\le 0$ が従う．ここで $g(c)$ は関数 $g(x)$ $(a\le x\le b)$ の最大値であったことを思い起こせば，$g(x)\le 0$ $(a\le x\le b)$ が結論される． □

[6] 実際にはこうはならないのですが，まだそのことは証明されていません．

関数の凸性は 2 階導関数とも関係します．次に定理の形でそれを示しましょう．

定理 3.20.
以下，$f(x)$ を開区間 I において定義された 2 回微分可能な関数とする．もしすべての点 x において $f''(x) > 0$ であるならば，$f(x)$ は下に凸である．また，逆にすべての点 x において $f''(x) < 0$ であるならば，$f(x)$ は上に凸である．

定理の証明は少しだけ後回しにして，1 変数関数に対する極値問題への応用について触れたいと思います．上述の定理より，ただちに以下が導かれます．

系 3.21.
いま，$f(x)$ を開区間 I において定義された 2 回連続的に微分可能な関数とする．さらに，$a \in I$ が関数 $f(x)$ の臨界点であるとする．

(i) さらにもし $f''(a) > 0$ であるならば，関数 $f(x)$ は $x = a$ において極小である．

(ii) またもし $f''(a) < 0$ であるならば，$f(x)$ は $x = a$ において極大である．

関数 $f(x)$ が $x = a$ を臨界点として持つならば，その関数のグラフ $y = f(x)$ の点 $(a, f(a))$ における接線は水平です．さらに，(i) にあるように，$f''(a) > 0$ であるとしましょう．正の実数 ϵ を十分小さくとれば，区間 $(a - \epsilon, a + \epsilon)$ において $f''(x) > 0$ が成り立つことが，2 階導関数 $f''(x)$ の連続性より保証されます．したがって，定理 3.20 によれば，関数 $f(x)$ は区間 $(a - \epsilon, a + \epsilon)$ において下に凸です．ここで，命題 3.19 を思い出しましょう．すると，区間 $(a - \epsilon, a + \epsilon)$ 上では，関数 $y = f(x)$ のグラフはその点 $(a, f(a))$ における接線の上側にあることが分かります．そして，これは関数 $f(x)$ が $x = a$ において極小であることに他なりません．これが，系 3.21 (i) が成り立つことの説明です．もちろん，(ii) も全く同じ仕方で説明可能です．

定理 3.20 の証明. 定理の第 1 の主張 (i) を証明する（(ii) の証明も同様なので，省略する）．そのために，以下の主張を証明する．

区間 I 上で定義された 2 回微分可能関数 $g(x)$ と $a \in I$ が与えられたとする．

$$g(a) = g'(a) = 0, \qquad g''(x) > 0 \quad (x \in I) \tag{3.12}$$

が成り立つならば，

$$g(x) \geq 0 \quad (x \in I)$$

である．

この主張を証明するために，まず関数 $g(x)$ 自体に平均値の定理を適用する．すると，仮定 $g(a) = 0$ より以下が従う：各 $x \in I$ に対し，適当に θ（ただし，$0 < \theta < 1$）をとれば，

$$g(x) = g'(a + \theta(x-a))(x-a)$$

が成り立つ．さらに，導関数 $g'(x)$ に対し再び平均値の定理を適用する．その際に，$g'(a) = 0$ であったことを思い起こそう．すると，各 $x \in I$ に対し，

$$g'(x) = g''(a + \bar{\theta}(x-a))(x-a)$$

なる定数 $\bar{\theta}$（ただし，$0 < \bar{\theta} < 1$）が存在することが従う．とくに，後者を x の代わりに $a + \theta(x-a)$ に対し適用することが可能である．そして，その結果，

$$g'(a + \theta(x-a)) = g''(a + \bar{\theta}\theta(x-a))\theta(x-a)$$

および

$$g(x) = g''(a + \bar{\theta}\theta(x-a))\theta(x-a)^2$$

を得る．仮定によれば右辺はいつもゼロ以上である．以上をあわせて，$g(x) \geq 0$ $(x \in I)$ を得る．

次に，定理の中の主張 (i) の証明を行おう．開区間 I において定義された 2 回微分可能な関数 $f(x)$ が，$f''(x) > 0$ $(x \in I)$ を満たすとする．この関数 $f(x)$ が下に凸であることを，命題 3.19 を利用して示したい．そのために，実数 $a \in I$ を任意にとる．このとき，

$$g(x) = f(x) - \left\{ f'(a)(x-a) + f(a) \right\} \qquad (a \leq x \leq b)$$

により定義された関数 $g(x)$ は，条件 (3.12) を満たす．したがって，$g(x) \geq 0$, すなわち，(3.8) が成り立つ．よって，関数 $f(x)$ が下に凸であることが結論される．□

B.5　テイラーの定理

多項式（で定義される関数）に対しては，その導関数も不定積分も極めて容易に計算可能です．しかも，導関数や不定積分自体，再び多項式になります．その意味で，多項式は最も扱いの容易な関数であるということが出来るでしょう．もちろん，多項式以外にも関数は存在します．例えば，指数関数や三角関数がそんな関数の例となります．そこで，そういった関数を，多項式で近似することを考えてみましょう．

まずは，その準備として，2 階導関数の定義を一般化して，n 階導関数を定義します．いま，$f(x)$ を開区間で定義された関数とします，そのとき，その n **階導関数**

$$f^{(n)}(x) \quad \text{あるいは} \quad \frac{d^n y}{dx^n}$$

を，帰納的に

$$f^{(n)}(x) = (f^{(n-1)}(x))' \qquad \text{あるいは} \qquad \frac{d^n y}{dx^n} = \frac{d}{dx}\left(\frac{d^{n-1}y}{dx^{n-1}}\right)$$

により定義します．ただし，

$$f^{(0)}(x) = f(x), \qquad \text{あるいは} \qquad \frac{d^0 y}{dx^0} = y$$

と約束します．関数 $f(x)$ に対し，n 階までの導関数 $f'(x), f''(x), \cdots, f^{(n)}(x)$ が存在するとき，$f(x)$ は n **回微分可能**であると言われます．また，さらに n 階導関数 $f^{(n)}(x)$ が連続であるとき，$f(x)$ は n **回連続的に微分可能**であると言われます．[7]

いま，関数 $f(x)$ は何回でも必要なだけ微分できる関数であるとします．その関数 $f(x)$ が，

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + g(x) x^{n+1} \tag{3.13}$$

と書けたとしましょう．ただし，$a_0, a_1, \cdots, a_n$ は定数，$g(x)$ は関数とします．右辺の最後の項を除いたもの，すなわち，$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ は x の n 次多項式です．この多項式を関数 $f(x)$ の近似と見なそう，というのが上の仮定の意図するところです．すると，右辺の最後の項は誤差と見なされるべきものです．それが，$g(x)x^{n+1}$ という形をしている，というのもまた仮定の一部であると考えてください．いま，さらに，関数 $g(x)$ もまた何回でも微分可能であると仮定します．この仮定のもと，関数 $f(x)$ を近似する多項式 $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ の係数を決定しましょう．

[7] このときには，低階の導関数 $f'(x), \cdots, f^{(n-1)}(x)$ は自動的に連続になることに注意してください．

まず，(3.13) において $x=0$ とおくことにより，ただちに，

$$f(0)=a_0$$

を得ます．次に (3.13) の両辺を x で微分してみましょう．すると，

$$f'(x)=a_1+2a_2x+3a_3x^2+\cdots+na_nx^{n-1}+g'(x)x^{n+1}+(n+1)g(x)x^n \tag{3.14}$$

が従います．この式で $x=0$ とすれば，

$$f'(0)=a_1$$

を得ます．さらに，(3.14) をもう 1 回微分すれば，

$$\begin{aligned}f''(x)&=2a_2+3\cdot 2a_3x+4\cdot 3a_4x^2+\cdots+n\cdot(n-1)a_nx^{n-2}\\&\quad+g''(x)x^{n+1}+2(n+1)g'(x)x^n+(n+1)\cdot ng(x)x^{n-1},\end{aligned}$$

そしてこの式において $x=0$ とおけば，

$$f''(0)=2a_2$$

となります．以上の議論をさらに繰り返すことにより，

$$f^{(k)}(0)=k!a_k \qquad (k=0,1,\cdots,n)$$

を得ます．これで，関数 $f(x)$ を近似する多項式 $a_0+a_1x+a_2x^2+\cdots+a_nx^n$ の係数 $a_0,a_1,a_2,\cdots,a_n$ の係数がすべて決定されたことになります．

次に述べるマクローリンの定理は，$n+1$ 回微分可能な関数は，必ず (3.13) のように多項式近似でき，しかも誤差項 $g(x)x^{n+1}$ の形も大きく制約されることを保証します．

定理 3.22 (マクローリンの定理).

関数 $f(x)$ を開区間 I 上で定義された $n+1$ 回微分可能な関数とする．ただし，区間 I は原点 0 を含むとする．このとき，各 $x\in I$ に対し，以下の等式を満たすような定数 θ（ただし，$0<\theta<1$）が存在する：

$$\begin{aligned}f(x)&=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2\\&\quad+\cdots\cdots+\frac{f^{(n)}(0)}{n!}x^n+\frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}\end{aligned} \tag{3.15}$$

証明. $x=a$ に対し，(3.15) を証明する．そのために，新たな関数 $g(x)$ を以下で定義する（ただし，A は定数とする）：

$$\begin{aligned}g(x) = f(a) - \Big\{ f(x) &+ \frac{f'(x)}{1!}(a-x) + \frac{f''(x)}{2!}(a-x)^2 \\ &+ \cdots\cdots + \frac{f^{(n)}(x)}{n!}(a-x)^n + A(a-x)^{n+1} \Big\}.\end{aligned}$$

明らかに $g(a)=0$ が成り立つ．さらに $g(0)=0$ が成り立つように定数 A をとることができる．以下，A をそのようにとったとしよう．すると，$g(0)=g(a)=0$ であるから，ロルの定理によれば，

$$g'(b) = 0$$

なる点 b が 0 および a の間に存在する．ところが，一方，$g(x)$ の定義から

$$\begin{aligned}g'(b) &= -\left\{ f'(b) + \frac{f''(b)}{1!}(a-b) - \frac{f'(b)}{1!} \right.\\ &\qquad + \frac{f'''(x)}{2!}(a-b)^2 - \frac{f''(b)}{1!}(a-b) + \cdots\cdots \\ &\qquad\quad \left. + \frac{f^{(n+1)}(b)}{n!}(a-b)^n - \frac{f^{(n)}(b)}{(n-1)!}(a-b)^{n-1} - (n+1)A(a-b)^n \right\} \\ &= -\left\{ \frac{f^{(n+1)}(b)}{n!}(a-b)^n - (n+1)A(a-b)^n \right\}\end{aligned}$$

である．この値がゼロであることから，

$$A = \frac{f^{(n+1)}(b)}{(n+1)!}$$

であることが直ちに従う．これと $g(0)=0$ より，

$$f(a) = f(0) + \frac{f'(0)}{1!}a + \frac{f''(0)}{2!}a^2 + \cdots + \frac{f^{(n)}(0)}{n!}a^n + \frac{f^{(n+1)}(b)}{(n+1)!}a^{n+1},$$

を得る． □

さあ，さらに話を進めましょう．マクローリンの定理からただちに次を得ます．

定理 3.23 (テイラーの定理).
関数 $f(x)$ を開区間 I 上で定義された $n+1$ 回微分可能な関数とする．ただし，区間 I は点 a を含むとする．このとき，各 $x \in I$ に対し，以下の等式を満たすような定数 θ（ただし，$0<\theta<1$）が存在する：

$$\begin{aligned}f(x) = f(a) &+ \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots\cdots \\ &+ \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(a+\theta(x-a))}{(n+1)!}(x-a)^{n+1}. \qquad (3.16)\end{aligned}$$

テイラーの定理はマクローリンの定理からただちに導出されます．実際，$t = x - a$ により新たな独立変数 t を導入し，関数 $f(x)$ を t の関数と見なしたものを $g(t)$ をし，その $g(t)$ に対しマクローリンの定理を適用すれば良いわけです．

とくに $n = 1$ の場合，上述のテイラーの定理は平均値の定理に他なりません．すなわち，テイラーの定理は平均値の定理の一般化と位置付けられます．しかし一方，すでに見たとおり，テイラーの定理の原型であるマクローリンの定理の証明においては，平均値の定理の特殊形であるところのロルの定理が本質的な仕方で利用されます．その意味では，マクローリンの定理やテイラーの定理を，平均値の定理の応用と考えることも可能です．

さて，マクローリンの定理の応用についてお話ししましょう．まずは，関数の値の近似計算です．

問題 3.24. ――――――――――――――――――――――――――――

正弦関数 $f(x) = \sin x$ に対し，マクローリンの定理を適用せよ．またその結果を用いて，$\sin 0.1$ の近似値を，誤差 10^{-10} 未満で計算せよ．

―――――――――――――――――――――――――――――――――――

解答. 関数 $f(x) = \sin x$ に対し，

$$f^{(2m)}(x) = (-1)^m \sin x, \qquad f^{(2m+1)}(x) = (-1)^m \cos x$$

であるから，とくに

$$f^{(2m)}(0) = 0, \qquad f^{(2m+1)}(0) = (-1)^m$$

である．よって，$f(x) = \sin x$ に対しマクローリンの定理を適用すれば，

$$\begin{aligned}\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\cdots \\ + (-1)^m \frac{x^{2m+1}}{(2m+1)!} + (-1)^{m+1} \frac{x^{2m+3}}{(2m+3)!} \cdot \cos \theta x\end{aligned}$$

を得る．[8] ただし，$0 < \theta < 1$ である．この等式の右辺から最後の項を取り除いたものを $\sin x$ の近似値と見なしたとき，その誤差は，$|x|^{2m+3}/(2m+3)!$ 以下であることになる．さて，とくに $x = 0.1$ としよう．そのとき，

$$\frac{0.1^{2m+3}}{(2m+3)!} < 10^{-10}$$

であるような最小の m は $m = 2$ であることが簡単な計算により分かる．

$$\frac{0.1}{1!} - \frac{0.1^3}{3!} + \frac{0.1^5}{5!} = 0.09983341666\cdots$$

[8] ここでは，$n = 2m + 2$ としてマクローリンの定理を使っています．

であるから，

$$0.0998334167$$

が，$\sin 0.1$ の誤差 10^{-10} 未満の近似値であることが従う． □

B.6　高階のロピタルの定理

より高階の微分を含むマクローリンの定理を用いることにより，「高階のロピタルの定理」が得られます．それを問題としましょう．

問題 3.25 (高階ロピタルの定理). ――――――――――――――――
関数 $f(x)$, $g(x)$ を原点 0 を含む開区間で定義された $n+1$ 回連続的に微分可能な関数とする．さらに，

$$f(0)=g(0)=f'(0)=g'(0)=\cdots=f^{(n)}(0)=g^{(n)}(0)=0, \quad g^{(n+1)}(0)\neq 0,$$

かつ，$f^{(n+1)}(x)$, $g^{(n+1)}(x)$ は $x=0$ において連続であると仮定する．このとき，

$$\lim_{x\to 0}\frac{f(x)}{g(x)}=\frac{f^{(n+1)}(0)}{g^{(n+1)}(0)} \tag{3.17}$$

が成り立つことを証明せよ．

――――――――――――――――――――――――――――――

解答. マクローリンの定理，および仮定より，

$$f(x)=\frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}, \qquad g(x)=\frac{g^{(n+1)}(\bar{\theta}x)}{(n+1)!}x^{n+1},$$

を得る（ただし，$0<\theta<1, 0<\bar{\theta}<1$）．よって，

$$\frac{f(x)}{g(x)}=\frac{f^{(n+1)}(\theta x)}{g^{(n+1)}(\bar{\theta}x)}$$

が成り立つ．この等式において，$x\to 0$ とすると，$0<\theta<1, 0<\bar{\theta}<1$ であること，および $f^{(n+1)}(x)$, $g^{(n+1)}(x)$ の連続性より，(3.17) が従う． □

高階ロピタルの定理を利用した極限計算の例を問題として出すことにします．

問題 3.26. ――――――――――――――――――――――
次の極限を計算せよ：

$$\lim_{x\to 0}\frac{x-\sin x}{x^3}.$$

――――――――――――――――――――――――――――――

解答. 問題に現れる分数の分子，分母を，それぞれ $f(x)$, $g(x)$ とすると，

$$
\begin{aligned}
&f'(x) = 1 - \cos x, \qquad &&f''(x) = \sin x, \qquad &&f'''(x) = \cos x,\\
&g'(x) = 3x^2, &&g''(x) = 6x, &&g'''(x) = 6
\end{aligned}
$$

である．したがって，とくに $f(0) = f'(0) = f''(0) = g(0) = g'(0) = g''(0) = 0$ であるが，$g'''(0) = 6 \neq 0$ であるから，

$$\lim_{x\to 0} \frac{x - \sin x}{x^3} = \frac{f'''(0)}{g'''(0)} = \frac{1}{6}$$

を得る． □

B.7　テイラー展開

「無限次の多項式」

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

を**巾級数**と呼びます． 適当な仮定の下，関数が巾級数表示可能であることを示すのが，以下の定理です．[9]

定理 3.27.
いま，$f(x)$ を開区間 $(a-R, a+R)$ (ただし，$R > 0$) で定義された無限回微分可能な関数とする．さらに，任意の $n = 0, 1, \cdots$, および $x \in (a-R, a+R)$ に対し，

$$|f^{(n)}(x)| \leq M$$

が成り立つと仮定する．ただし，ここで M は n や x に依らない定数であるとする．このとき，各 $x \in (a-R, a+R)$ に対し，

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 \tag{3.18}$$

$$+ \cdots\cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots\cdots \tag{3.19}$$

が成り立つ．

証明. $x \in (a-R, a+R)$ をとり，固定する．さらに，数列 $\{y_n\}$ を，

$$y_n = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots\cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

[9]関数 $f(x)$ $(x \in I)$ が巾級数 $a_0 + a_1x + a_2x^2 + \cdots$ に等しいとは，各 $x \in I$ に対し，$y_n = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ $(n = 0, \cdots, n)$ とおいたとき，$y_n \to f(x)$ $(n \to \infty)$ が成り立つことです．

により定義する．すると，テイラーの定理によれば，θ_n（ただし，$0<\theta_n<1$）を適当にとれば，

$$f(x)-y_n=\frac{f^{(n+1)}(a+\theta_n(x-a))}{(n+1)!}(x-a)^{n+1}$$

が成り立つ．これより，$n\to\infty$ のとき，

$$|f(x)-y_n|\le M\frac{R^{n+1}}{(n+1)!}\longrightarrow 0$$

であることが従う．[10] □

この定理に現れる巾級数 (3.18) は，関数 $f(x)$ の**テイラー展開**と呼ばれます．また，とくに $a=0$ のときには，**マクローリン展開**と呼ばれることがあります．

問題 3.28. ――――――――――――――――――――――――――――

関数 e^x, $\sin x$, $\cos x$ が定理 3.27 の仮定を満たすことを確かめよ．また，これらの関数のマクローリン展開を求めよ．

――――――――――――――――――――――――――――――――――

解答． まず，指数関数 $y=e^x$ について考えよう．この関数に対しては，$d^ny/dx^n=e^x$ $(n=0,1,\cdots)$ である．したがって，$R>0$ を任意にとったとき，

$$\left|\frac{d^ny}{dx^n}\right|\le e^R \qquad (n=0,1,\cdots,|x|<R)$$

が成り立つ．よって，マクローリン展開が可能である．実際，指数関数 e^x のマクローリン展開は，

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots\cdots$$

で与えられる．

次に，正弦関数 $y=\sin x$ について考えよう．この場合，

$$\frac{d^{2m}y}{dx^{2m}}=(-1)^m\sin x,\qquad \frac{d^{2m+1}y}{dx^{2m+1}}=(-1)^m\cos x,\qquad (m=0,1,\cdots)$$

であるから，任意の $n=0,1,\cdots$, $x\in\mathbb{R}$ に対し，$|(d^ny/dx^n)(x)|\le 1$ が成り立つ．よって，マクローリン展開が可能であり，実際，

$$\sin x=\frac{x}{1!}-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots\cdots$$

である．

[10] 最後の極限計算に関しては，問題 1.6 を参考のこと．

最後に余弦関数 $y = \cos x$ を考える．この場合，

$$\frac{d^{2m}y}{dx^{2m}} = (-1)^m \cos x, \qquad \frac{d^{2m+1}y}{dx^{2m+1}} = (-1)^{m+1} \sin x, \qquad (m = 0, 1, \cdots)$$

であり，とくにこれより，マクローリン展開可能なことが直ちに分かる．また，マクローリン展開は以下で与えられる：

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\cdots.$$

□

C　章末問題

問題 3.29.

微分係数の定義に基づき，公式 3.3 を証明せよ．

解答は省略する．

問題 3.30.

逆三角関数 $\mathrm{Arcsin}\, y$, $\mathrm{Arccos}\, y$, $\mathrm{Arctan}\, y$ の導関数を計算せよ．

解答. まず，$x = \mathrm{Arcsin}\, y$ の導関数を計算しよう．そのために，$y = \sin x$ $(-\pi/2 \le x \le \pi/2)$ とおく．その逆関数が，$x = \mathrm{Arcsin}\, y$ であった．$dy/dx = \cos x$ であるから，$dx/dy = 1/\cos x$．この右辺を，変数 y を使って表したい．そこで，$-\pi/2 \le x \le \pi/2$ において $\cos x \ge 0$ であることに注意しよう．これより，$\cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2}$ であることが従う．[11] 以上から，$dx/dy = 1/\sqrt{1 - y^2}$，すなわち，

$$(\mathrm{Arcsin}\, y)' = \frac{1}{\sqrt{1 - y^2}}$$

が従う．

次に，$x = \mathrm{Arccos}\, y$ の導関数を計算しよう．これは，$y = \cos x$ $(0 \le x \le \pi)$ の逆関数であった．そして，$dy/dx = -\sin x$ より，$dx/dy = -1/\sin x$ を得る．ここで，$0 \le x \le \pi$ において $\sin x \ge 0$ であるから，$\sin x = \sqrt{1 - \cos^2 x} = \sqrt{1 - y^2}$ を得る．以上から，以下が結論される：

$$(\mathrm{Arccos}\, y)' = -\frac{1}{\sqrt{1 - y^2}}.$$

[11] 一般には，$\cos x = \pm\sqrt{1 - \sin^2 x}$ です．ところが，いま変数 x の動く範囲に制限が加えられていて，それが理由で $\cos x = -\sqrt{1 - \sin^2 x}$ が排除されます．

最後に，$x = \mathrm{Arctan}\, y$ の導関数の計算を行う．そのために，$y = \tan x$ $(-\pi/2 < x < \pi/2)$ とおく．問題の関数 $x = \mathrm{Arctan}\, y$ はこの関数の逆関数であった．そして，$dy/dx = 1/\cos^2 x$ より，$dx/dy = \cos^2 x = 1/(\tan^2 x + 1) = 1/(y^2+1)$, すなわち，

$$(\mathrm{Arctan}\, y)' = \frac{1}{y^2+1}$$

を得る． □

問題 3.31. ————————————————————————

(1) 開区間 I 上で定義された微分可能関数 $f(x)$ が以下の条件を満たすと仮定する：

$$f(x) \in I \qquad (x \in I);$$

ある $\lambda \in I$ に対し $f(\lambda) = \lambda$ が成り立つ;

$$|f'(x)| < K \qquad (x \in I).$$

ただし，K は $K < 1$ なる正の定数であるとする．いま，任意に $a \in I$ が与えられたとする．このとき，

$$x_0 = a, \qquad x_{n+1} = f(x_n) \quad (n = 0, 1, \cdots)$$

で定義される数列 $\{x_n\}$ に対し，

$$x_n \to \lambda \qquad (n \to \infty)$$

が常に成り立つことを示せ．

(2) フィボナッチ数列 $\{a_n\}$ に対し，その隣り合う 2 項の比 a_{n+1}/a_n が黄金比 $(1+\sqrt{5})/2$ に収束することを，(1) を用いて証明せよ．

————————————————————————————

解答. (1) 数列 $\{x_n\}$ を問題のようにとる．すると，

$$x_{n+1} - \lambda = f(x_n) - f(\lambda) = f'(y_n)(x_n - \lambda)$$

が平均値の定理より従う．ただし，ここで y_n は 2 点 λ, x_n の間にある点である．これより，

$$|x_{n+1} - \lambda| = |f'(y_n)||x_n - \lambda| \le K|x_n - \lambda|$$

であるから，

$$|x_n - \lambda| \le K^n|a - \lambda| \longrightarrow 0,$$

すなわち，$x_n \to \lambda$ $(n \to \infty)$ を得る．

(2) 簡単のため，$x_n = a_{n+1}/a_n$ とおく．すると，数列 $\{x_n\}$ は漸化式

$$x_{n+1} = 1 + \frac{1}{x_n}$$

を満たす．そこで，

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

とおく（ただし，この関数 $f(x)$ の定義域は後ほど定めることにする）．すると，数列 $\{x_n\}$ に対する漸化式は，

$$x_{n+1} = f(x_n)$$

と書くことが出来る．黄金比を λ で表すことにすれば，

$$f(\lambda) = \lambda$$

が成り立つ．また，$f'(x) = -1/x^2$ である．そこで，定数 b を，

$$1 < b < 3/2$$

を満たすようとり，関数 $f(x)$ の定義域を区間 $I = (b, 2)$ に制限しよう．すると，

$$f(x) \in I \qquad (x \in I) \qquad \text{および} \qquad |f'(x)| < \frac{1}{b^2} < 1$$

が成り立つ．したがって，(1) より $x_n \to \lambda$ $(n \to \infty)$ が従う． □

問題 3.32.

関数 $f(x)$ $(a \le x \le b)$ について，以下の問に答えよ．

(1) 関数 $f(x)$ が，

$$f(a) < 0, \qquad f(b) > 0, \qquad f'(x) > 0 \quad (a \le x \le b) \tag{3.20}$$

を満たすとする．このとき，

$$f(\gamma) = 0 \quad (a < \gamma < b)$$

を満たす γ がただひとつ存在することを示せ．

(2) 関数 $y = f(x)$ のグラフ上の点 $(c, f(c))$（ただし，$a \le c \le b$）における接線 ℓ の方程式を求めよ．また，その接線 ℓ と x-軸の交点の x-座標 $\sigma(c)$ を求めよ．

(3) 以降，さらに

$$f''(x) > 0 \quad (a \le x \le b) \tag{3.21}$$

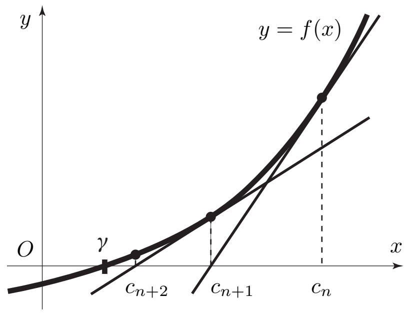


図 3.14: ニュートン法

を仮定する．(2) において，$\gamma < c \leq b$ であるならば，

$$f(\sigma(c)) > 0 \quad \text{かつ} \quad \gamma < \sigma(c) < c$$

が成り立つことを示せ．

(4) 数列 $\{c_n\}$ を，

$$c_0 = b, \qquad c_{n+1} = \sigma(c_n) \quad (n = 0, 1, \cdots)$$

により定義する．このとき，数列 $\{c_n\}$ が収束することを証明せよ．また，

$$\lim_{n\to\infty} c_n = \gamma$$

を示せ．

(5) とくに，関数

$$f(x) = x^2 - 2 \quad (1 \leq x \leq 2)$$

を考える．この関数が，条件 (3.20), (3.21) を満たしていることを確かめよ．また，c_1, c_2, c_3 を求めよ．ただし，$c_0 = 2$ である．

(6) c_3 を $\sqrt{2}$ の近似値と見なしたとき，その誤差を評価せよ．[12]

解答． (1) 関数 $f(x)$ は微分可能であるから，とくに連続である．さらに，$f(a)\cdot f(b) < 0$ であるから，中間値の定理により，$f(\gamma) = 0$ なる γ $(a < \gamma < b)$ の存在が従う．一方，$f'(x) > 0$ であるから，関数 $f(x)$ は（狭義）単調増加であり，したがって $f(\gamma) = 0$ なる γ は一意的である．

(2) 点 $(c, f(c))$ における接線 ℓ の傾きは $f'(c)$ であるから，ℓ の方程式は，

$$y - f(c) = f'(c)(x - c)$$

[12] すなわち，$|c_3 - \sqrt{2}| \leq K$ なる定数 K を見出せ．もちろん，K の値は出来るだけ小さい方が望ましい．

である．この方程式に $(x, y) = (\sigma(c), 0)$ を代入することにより，

$$\sigma(c) = c - \frac{f(c)}{f'(c)}$$

を得る．

(3) 条件 (3.21) によれば，関数 $y = f(x)$ は下に凸である．したがって，関数 $y = f(x)$ のグラフは，その点 $(c, f(c))$ における接線 ℓ の上側にある（図 3.15 を参照のこと）．このことより，$f(\sigma(c)) > 0$ であることが従う．一方，

$$\begin{cases} a \leq x < \gamma & \text{において} \quad f(x) < 0, \\ \gamma < x \leq b & \text{において} \quad f(x) > 0 \end{cases}$$

である．したがって，$f(\sigma(c)) > 0$ から $\sigma(c) > \gamma$ が従う．さらに，$c > \gamma$ より $f(c) > 0$ である．よって (3.20) とより，

$$\sigma(c) - c = -\frac{f(c)}{f'(c)} < 0,$$

すなわち，$\sigma(c) < c$ を得る．

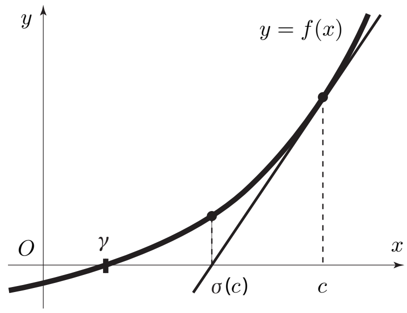


図 3.15:

(4) (3) によれば，$c_{n+1} = \sigma(c_n) < c_n$，かつ $c_n > \gamma$ である．すなわち，$\{c_n\}$ は下に有界な単調減少数列である．したがって，$\{c_n\}$ は収束する．その極限値を $\bar{\gamma}$ とすると，漸化式

$$c_{n+1} = \sigma(c_n) = c_n - \frac{f(c_n)}{f'(c_n)}$$

において，$n \to \infty$ とすることにより，

$$\bar{\gamma} = \bar{\gamma} - \frac{f(\bar{\gamma})}{f'(\bar{\gamma})},$$

すなわち，$f(\bar{\gamma}) = 0$ を得る．これより直ちに $\bar{\gamma} = \gamma$ が従う．

(5) 関数 $f(x) = x^2 - 2$ $(1 \leq x \leq 2)$ に対し，

$$f(1) = -1 < 0, \qquad f(2) = 2 > 0, \qquad f'(x) = 2x \geq 2 > 0 \quad (1 \leq x \leq 2),$$
$$f''(x) = 2 > 0$$

であるから，確かに (3.20), (3.21) が成り立つ．また，

$$\sigma(c) = c - \frac{c^2 - 2}{2c} = \frac{c}{2} + \frac{1}{c}$$

であるから，$c_0 = 2$ としたとき，

$$c_1 = \frac{3}{2}, \qquad c_2 = \frac{17}{12}, \qquad c_3 = \frac{577}{408}$$

である．

(6) $c_3^2 - (\sqrt{2})^2 = (c_3 + \sqrt{2})(c_3 - \sqrt{2}) > 2\sqrt{2}(c_3 - \sqrt{2}) > 2(c_3 - \sqrt{2})$ であるから，

$$0 < c_3 - \sqrt{2} < \frac{c_3^2 - (\sqrt{2})^2}{2} = \frac{1}{332928}$$

を得る．すなわち，誤差は

$$\frac{1}{332928}$$

未満である． □

この問題にある仕方で，方程式の近似解を求めることが出来ます．近似解を求めるこの方法を**ニュートン法**と呼びます．この問題にあるような設定の下では，区間縮小法によっても近似解が得られますが，ニュートン法の方が，解への収束が「速い」という違いがあることが知られています．その意味で，ニュートン法の方が優れた解法です．

問題 3.33. ────────────────────────

関数

$$f(x) = \log(x+1) \qquad (x > -1)$$

を考える．

(1) 関数 $f(x)$ の n 次導関数 $f^{(n)}(x)$ を求めよ．

(2) マクローリンの定理を用い，関数 $f(x)$ を近似する 4 次関数を求めよ．また，そのときの誤差を評価せよ．

(3) (2) において求めた 4 次関数において，とくに $x = 0.1$ とおくことにより，$\log 1.1$ の近似値を求めたとき，その誤差の大きさを評価せよ．

────────────────────────────

解答. (1) 与えられた関数を逐次微分して，

$$f'(x) = \frac{1}{x+1}, \quad f''(x) = -\frac{1}{(x+1)^2}, \quad f'''(x) = \frac{2}{(x+1)^3}, \quad \cdots,$$
$$f^{(n)}(x) = (-1)^{(n-1)}\frac{(n-1)!}{(x+1)^n}$$

を得る．

(2) マクローリンの定理によれば，

$$\log(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{1}{5}\cdot\frac{x^5}{(\theta x+1)^5}$$

が成り立つ（ただし，$0 < \theta < 1$ ）．よって，

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$$

を，関数 $\log(x+1)$ を近似する 4 次関数と見なすことが出来る．そして，その誤差は，

$$\frac{1}{5}\cdot\frac{x^5}{(\theta x+1)^5}$$

である．

(3) (2) で求めた誤差において，とくに $x = 0.1$ とおくと，

$$0 < \frac{1}{5}\cdot\frac{(0.1)^5}{(0.1\theta+1)^5} < \frac{1}{5}\cdot(0.1)^5 = 2\times 10^{-6}$$

を得る．すなわち，誤差は

$$2\times 10^{-6}$$

未満である． □

問題 3.34. ――――――――――――――――――――――――――――――

次の極限が存在するような定数 A の値を決定せよ．また，そのときの極限値を求めよ．

$$\lim_{x\to 0}\frac{e^x - \sin x - 1 + Ax^2}{x^3}.$$

――――――――――――――――――――――――――――――――――――

解答. まず，指数関数と正弦関数のテイラー展開を思い起こそう：

$$\begin{array}{rcccccccccl} e^x &=& 1 &+& \dfrac{x}{1!} &+& \dfrac{x^2}{2!} &+& \dfrac{x^3}{3!} &+& \cdots, \\ \sin x &=& &+& \dfrac{x}{1!} && &-& \dfrac{x^3}{3!} &+& \cdots. \end{array}$$

したがって，

$$e^x - \sin x - 1 + Ax^2 = \left(\frac{1}{2} + A\right)x^2 + \frac{x^3}{3} + \cdots$$

である．よって，問題の極限が存在するためには，$A = -1/2$ でなければならない．また，そのとき，極限の値は，$1/3$ である．　□

第 4 章

多変数関数の微分

A　多変数極値問題

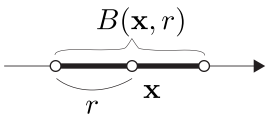
A.1　偏微分と極値問題

1 変数関数に対しその極大値・極小値を決定する方法を多変数関数に対するそれに拡張すること，それがこの章の目的です．多変数関数の場合には，その定義域の「自由度が大きい」がゆえに，極値問題の解法は，1 変数関数の場合と比べ，難しくなります．しかし，基本的な考え方は大きく変わりませんので，心配の必要はありません．

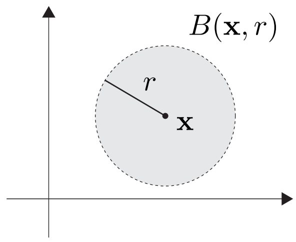
まずは多変数関数の極値を厳密に定式化する必要があります．そのための簡単な準備から話を始めることにしましょう．実数を n 個並べたもの $(x_1, \cdots, x_n)$ 全部からなる集合を，n **次元ユークリッド空間**と呼び，記号 $\mathbb{R}^n$ で表します．また，その要素を**点**と呼びます．とくに，$n = 1, 2, 3$ のときは，$\mathbb{R}^n$ は幾何学的に，それぞれ，直線，平面，空間と見なされることは皆さんご存じの通りです．次元 n が 3 より大きいときにも，$\mathbb{R}^n$ を「空間」と呼びます．しかし，低次元の場合のようにそれを（通常の意味で）視覚的にとらえることは不可能です．（それにも関わらず，数学では，高次元のユークリッド空間やその部分集合をある種の「図形」として幾何学的に取り扱いますが）．n 次元ユークリッド空間 $\mathbb{R}^n$ の点 $\mathbf{x}$，および正の実数 r に対し，

$$B(\mathbf{x}; r) = \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{x}| < r\}$$

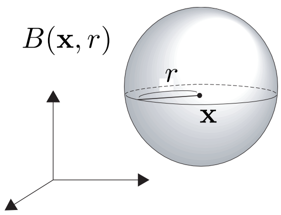
とおきましょう．集合 $B(\mathbf{x}; r)$ は，$n = 1$ のときには点 $\mathbf{x}$ を中点とする長さ $2r$ の開区間，$n = 2$ のときには平面内の点 $\mathbf{x}$ を中心とする半径 r の円板（の内部），そして $n = 3$ の場合には（3 次元）空間内の点 $\mathbf{x}$ を中心とする半径 r の球体（の内部）となります（図 4.1）．一般次元においても，通常 $B(\mathbf{x}; r)$ を，点 $\mathbf{x} \in \mathbb{R}^n$ を中心とする半径 r の n **次元開円板**，ないし n **次元開球体**と呼びます．



(a) $n = 1$ の場合



(b) $n = 2$ の場合



(c) $n = 3$ の場合

図 4.1: $\mathbb{R}^n$ $(n = 1, 2, 3)$ の開球体 $B(\mathbf{x}; r)$

定義 4.1. ―――――――――――――――――――――――――――――――

X を n 次元ユークリッド空間 $\mathbb{R}^n$ の空でない部分集合と，また $f : X \to \mathbb{R}$ を X において定義された関数とする．さらに $\mathbf{a} \in X$ とする．もし

$$\mathbf{x} \in X \cap B(\mathbf{a}; r) \qquad \Longrightarrow \qquad f(\mathbf{a}) \leq f(\mathbf{x})$$

を満たすような $r > 0$ が存在するならば，関数 f は点 $\mathbf{a}$ において**極小**であると言われる．またそのとき，値 $f(\mathbf{a})$ を f の**極小値**と呼ぶ．一方，もし十分小さい $r > 0$ に対し，

$$\mathbf{x} \in X \cap B(\mathbf{a}; r) \qquad \Longrightarrow \qquad f(\mathbf{a}) \geq f(\mathbf{x})$$

が成り立つならば，f は $\mathbf{a}$ において**極大**であると言い，またそのとき $f(\mathbf{a})$ を f の**極大値**と呼ぶ．

―――――――――――――――――――――――――――――――――――

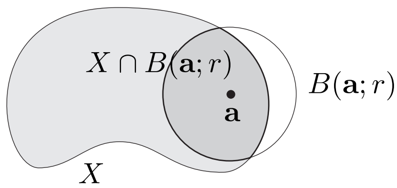


図 4.2:

言い換えるならば，十分小さな $r > 0$ をとったとき，関数 f の $X \cap B(\mathbf{a} : r)$ への制限が $\mathbf{a}$ において最小であるとき，f は $\mathbf{a}$ において極小であると，また $f \cap B(\mathbf{a}; r)$ への制限が $\mathbf{a}$ において最大であるとき，f は $\mathbf{a}$ において極大であると言われます．もちろん，関数 f がある点で最小ならば，その点において f は極小であるし，また反対に最大であれば，そこで極大です．しかし，1 変数関数の場合とまったく同様，一般には逆は成立しません．富士山は日本で一番高い山だけれど，世界で一番高い山ではないのと同じです．あるいは，極小値を「井の中の蛙」と例えることもできるかも知れません．

次に，この章で学ぶ予定の多変数極値問題に対する解法が適用できる関数のクラスを特定しなければなりません．そのためにまず，多変数関数の連続性を定義する必要があります．ところで，1 変数関数の場合には，数列の収束を利用してその連続性が定義されました．類似の仕方で多変数関数に対してもその連続性を定義しようと思うと，数列の「高次元化」が必要となります．そもそも，数列とは，非負，または正の自然数のおのおのに，実数，すなわち集合 $\mathbb{R}$ の要素を対応させる規則です．したがって，その高次元版は，非負ないし正の自然数 ℓ に対し，n 次元ユークリッド空間 $\mathbb{R}^n$ の点 $\mathbf{x}_\ell$ を対応させる規則となります．かような $\{\mathbf{x}_\ell\}$ を，$\mathbb{R}^n$ の**点列**と呼びます．また，$\mathbb{R}^n$ の部分集合 X が存在して，すべての番号 ℓ に対し $\mathbf{x}_\ell \in X$ が成り立つとき，$\{\mathbf{x}_\ell\}$ を X の点列と呼ぶこともあります．さて，点列の収束を定義しましょう．

定義 4.2. ――――――――――――――――――――――――――――――――

n 次元ユークリッド空間内の点列 $\{\mathbf{x}_\ell\}$，および点 $\mathbf{x} \in \mathbb{R}^n$ に対し，

$$|\mathbf{x}_\ell - \mathbf{x}| \to 0 \qquad (\ell \to \infty) \tag{4.1}$$

が成り立つとき，点列 $\{\mathbf{x}_\ell\}$ は点 $\mathbf{x}$ に**収束**すると言う．またそのとき，$\mathbf{x}_\ell \to \mathbf{x}$ $(\ell \to \infty)$，あるいは，$\mathbf{x} = \lim_{\ell\to\infty} \mathbf{x}_\ell$ と記す．

――――――――――――――――――――――――――――――――――――

この定義の条件 (4.1) に現れる $|\mathbf{x}_\ell - \mathbf{x}|$ は実数の列，つまり数列ですから，その収束は既知のものであることに注意しましょう．また，点列 $\{\mathbf{x}_\ell\}$ の収束は，各点 $\mathbf{x}_\ell$ の成分を用いても判定可能です．実際，$\mathbb{R}^n$ の点列 $\{\mathbf{x}_\ell\}$ および $\mathbf{x} \in \mathbb{R}^n$ に 対し $\mathbf{x}_\ell \to \mathbf{x}$ であるための必要十分条件は，$\mathbf{x}_\ell = (x_1^{(\ell)} \cdots x_n^{(\ell)})$，$\mathbf{x} = (x_1 \cdots x_n)$ とおいたとき，各 $i = 1, \cdots, n$ に対し $x_i^{(\ell)} \to x_i$ $(\ell \to \infty)$ が成り立つことです．

次に多変数関数の連続性を定義しましょう．

定義 4.3. ――――――――――――――――――――――――――――――――

$\mathbb{R}^n$ の部分集合 X 上の関数 f に対し，もしそれが以下に述べる条件を満たすなら

ば，f は**連続**であると言われる：

> $\mathbf{x}$ を X の任意の点，$\{\mathbf{x}_\ell\}$ を点 $\mathbf{x}$ に収束するような X の任意の点列としたとき，$f(\mathbf{x}_\ell) \longrightarrow f(\mathbf{x})$ $(\ell \to \infty)$ が必ず成り立つ．

この定義も，1 変数関数の場合と同様，形式的には $\lim f(\mathbf{x}_\ell) = f(\lim \mathbf{x}_\ell)$，すなわち，$f$ と極限操作が可換であることを意味しています．多変数関数の連続性については，とりあえず定義を与えるにとどめ，これ以上深入りしないことにします．

さて，ここで再び，1 変数関数のことを思い起こして下さい．通常，その定義域は何らかの区間であると仮定します．ところで，区間には開区間，閉区間等，いくつかの種類がありました．しかし，少なくとも 1 変数関数に対しその微分を議論する上では，開区間が都合が良かったわけです．次の目的は，開区間の「高次元版」を定義することです．もちろん，その第一候補は開円板 $B(\mathbf{x}, r)$ でしょうが，それより広いクラスを設定しておいた方が，後々便利なことが多いので，ここでもそうすることにします．

定義 4.4. ———

X を n 次元ユークリッド空間 $\mathbb{R}^n$ の部分集合とする．もし，各 $\mathbf{x} \in X$ に対し，$B(\mathbf{x}; r) \subset X$ なる $r > 0$ が存在するならば，X は $\mathbb{R}^n$ の**開部分集合**，あるいは単に**開集合**であると言われる．

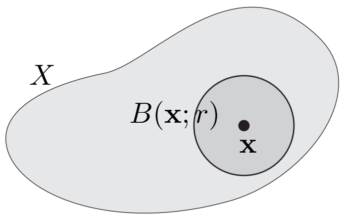


図 4.3: 開集合

各点 $\mathbf{x} \in X$ に対し，それを中心とししかも十分小さな半径 $r > 0$ を持つ開球体 $B(\mathbf{x}; r)$ が X の中に含まれるならば，X は開集合であると言われる．

例えば，平面，すなわち 2 次元ユークリッド空間の部分集合 $X = \{(x, y) \in \mathbb{R}^2 : |x| < 1, |y| < 1\}$ は開集合です．[1] 実際，各 $(x, y) \in X$ に対し，$(1+x)/2$, $(1-x)/2$, $(1+y)/2$, $(1-y)/2$ のうち一番小さいものを r とすれば，$r > 0$ であり，また

[1] ここでは，変数 x_1, x_2 の代わりに，変数 x, y を用いています．

$B((x,y);r) \subset X$ も成り立ちます．一方，$Y = \{(x,y) \in \mathbb{R}^2 : |x| \leq 1, |y| \leq 1\}$ は開ではありません．実際，$(x,y) \in Y$ を集合 Y の境界 $\{(x,-1) \in \mathbb{R}^2 : |x| \leq 1\} \cup \{(x,1) \in \mathbb{R}^2 : |x| \leq 1\} \cup \{(-1,y) \in \mathbb{R}^2 : |y| \leq 1\} \cup \{(1,y) : |y| \leq 1\}$ の点とすれば，いかなる $r > 0$ に対しても $B((x,y);r) \not\subset Y$ です．

ようやく多変数関数に対する微分を議論する準備ができました．以降この節を通じて，$f : X \to \mathbb{R}$ を $\mathbb{R}^n$ の開集合 X 上で定義された関数とします．また，$\mathbf{x}$ を X の点，また $\mathbf{v}$ を n 次元行ベクトルとします．もし，実数 t に対し $|t|$ が十分小さければ，$\mathbf{x}+t\mathbf{v}$ もまた X の点となることが，X が開集合であるという仮定から従います．したがって，$f(\mathbf{x}+t\mathbf{v})$ を，実数 t（ただし，$|t|$ は十分小）の関数と見なすことが可能です．そもそも，これが，f の定義域 X は開集合であると仮定した理由です．この t の関数のとくに $t=0$ における微分係数が存在するとき，関数 f は点 $\mathbf{x}$ において $\mathbf{v}$ 方向に**方向微分可能**である，と言います．また，その微分係数を，f の点 $\mathbf{x}$ における $\mathbf{v}$ 方向への**方向微分係数**，あるいは単に**方向微分**と呼び，記号 $(\nabla_{\mathbf{v}} f)(\mathbf{x})$ で表します：

$$(\nabla_{\mathbf{v}} f)(\mathbf{x}) = \frac{d}{dt}\Big|_{t=0} f(\mathbf{x}+t\mathbf{v}).$$

結局のところ，多変数関数を $\mathbb{R}^n$ 内の直線（あるいはより厳密に言えば，直線と定義域 X との交わり）に制限することにより 1 変数関数とみなし，それに対して 1 変数関数の微分を適用することにより得られるのが，方向微分に他なりません．

とくに，$\mathbf{v} = \mathbf{e}_i$（ただし，$\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ は $\mathbb{R}^n$ の標準基底）としたときの $\mathbf{v}$ 方向への方向微分 $(\nabla_{\mathbf{v}} f)(\mathbf{x})$ を，関数 f の点 $\mathbf{x}$ における**偏微分係数**と呼び，記号 $(\nabla_i f)(\mathbf{x})$，あるいは $\dfrac{\partial f}{\partial x_i}(\mathbf{x})$ により表します：

$$(\nabla_i f)(\mathbf{x}) = \frac{\partial f}{\partial x_i}(\mathbf{x}) = (\nabla_{\mathbf{e}_i} f)(\mathbf{x}).$$

さらに，もしすべての点 $\mathbf{x} \in X$ および $i = 1, \cdots, n$ に対して，偏微分係数 $(\nabla_i f)(\mathbf{x})$ が存在するならば，f は**偏微分可能**であると言われます．もし f が偏微分可能であるならば，各 $i = 1, \cdots, n$ に対し，偏微分係数 $(\nabla_i f)(\mathbf{x}) = \partial f/\partial x_i$ を点 $\mathbf{x} \in X$ の関数と見なすことが可能です．そして，それを f の**偏導関数**と呼びます．偏導関数の計算は極めて容易です．実際，関数 $f(\mathbf{x})$ (ただし，$\mathbf{x} = {}^t(x_1 \cdots x_n)$) の偏導関数 $\nabla_i f = \partial f/\partial x_i$ を求めるためには，変数 $x_1, \cdots, x_n$ のうち x_i を除いたものを定数と思い，変数 x_i に対し 1 変数関数の微分を適用すればよいわけです．たとえば，

$$f(x,y) = \sin x \, \cos y$$

に対しては，

$$\frac{\partial f}{\partial x} = \cos x \, \cos y, \qquad \frac{\partial f}{\partial y} = -\sin x \, \sin y$$

である，といった具合です．偏導関数の計算に関する簡単な類題をひとつ出しましょう．

問題 4.5. ―――――――――――――――――――――――――――――
関数

$$f(x,y) = x^4 + y^4 - 10x^2 + 16xy - 10y^2 \tag{4.2}$$

に対し，その偏導関数を求めよ．

―――――――――――――――――――――――――――――――――

解答. $$\frac{\partial f}{\partial x} = 4(x^3 - 5x + 4y), \qquad \frac{\partial f}{\partial y} = 4(y^3 + 4x - 5y)$$ □

偏微分は，方向微分の特殊な場合として定義されました．しかし，偏微分に若干の仮定を課せば，偏微分から任意の方向の方向微分を求めることが可能です．それを述べるために，定義をひとつだけ与えることにしましょう．

定義 4.6. ―――――――――――――――――――――――――――――
関数 f が偏微分可能であり，しかも各偏導関数 $\nabla_i f$ $(i = 1, \cdots, n)$ が X 上連続であるならば，f は C^1 **級**であると言われる．

―――――――――――――――――――――――――――――――――

命題 4.7. ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
関数 f が C^1 級であるならば，任意の点 $\mathbf{x} \in X$ および任意のベクトル $\mathbf{v} = (v_1, \cdots, v_n) \in \mathbb{R}^n$ に対し，方向微分 $(\nabla_{\mathbf{v}} f)(\mathbf{x})$ が存在し，しかも

$$(\nabla_{\mathbf{v}} f)(\mathbf{x}) = \sum v_i (\nabla_i f)(\mathbf{x}) \tag{4.3}$$

が成立する．

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

証明. しばらくの間 $|\mathbf{v}|$ は十分小さいと仮定する．このとき，X の点 $\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_n$ を，

$$\mathbf{x}_0 = \mathbf{x}, \qquad \mathbf{x}_i = \mathbf{x}_{i-1} + v_i \mathbf{e}_i \quad (i = 1, \cdots, n)$$

により帰納的に定義する（図 4.4 を参照のこと）．$\mathbf{x}_n = \mathbf{x} + \mathbf{v}$ であることに注意しよう．

$$f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) = \sum_{i=1}^{n} \{f(\mathbf{x}_i) - f(\mathbf{x}_{i-1})\}$$

が成り立つことは明らかである．一方，各 $i = 1, \cdots, n$ に対し，点 $\mathbf{x}_{i-1}$ と点 $\mathbf{x}_i$ を結ぶ線分 $\mathbf{x}_{i-1} + s(\mathbf{x}_i - \mathbf{x}_{i-1})$ $(s \in [0,1])$ をとり，その上に関数 f を制限する．

そのとき得られる変数 s のみの関数に対し，平均値の定理を適用すれば，

$$f(\mathbf{x}_i) - f(\mathbf{x}_{i-1}) = \frac{d}{ds}\Big|_{s=s_i} f(\mathbf{x}_{i-1} + s(\mathbf{x}_i - \mathbf{x}_{i-1}))$$

を満たすような $s_i \in [0,1]$ が存在することが判る．ところが，$\mathbf{x}_i - \mathbf{x}_{i-1} = v_i\mathbf{e}_i$ であったから，上式の右辺に現れる微分係数は，偏微分を使って，$v_i \cdot (\nabla_i f)(\mathbf{x}_{i-1} + s_i(\mathbf{x}_i - \mathbf{x}_{i-1}))$ と書ける．以上をまとめれば，結局，

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) = \sum_{i=1}^{n} v_i(\nabla_i f)(\mathbf{x}_{i-1} + s_i(\mathbf{x}_i - \mathbf{x}_{i-1}))$$

（ただし，$s_i \in [0,1]$）を得る．

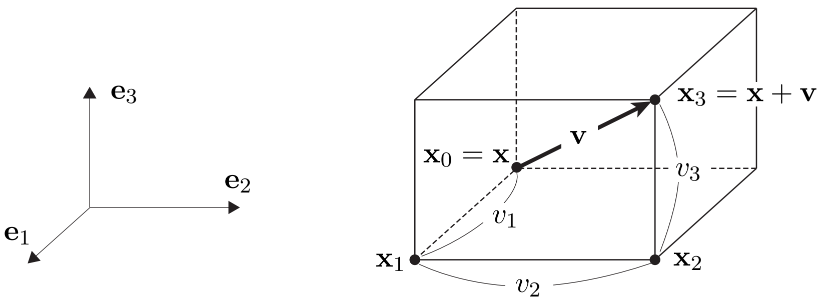


図 4.4: $n = 3$ の場合

いま得た結果において，$\mathbf{v} = (v_1 \cdots v_n)$ を $t\mathbf{v} = (tv_1 \cdots tv_n)$（$t \neq 0$ は実数）で置き換える．ただし，今度は $\mathbf{v}$ は任意にとり，その代わりに $|t|$ が十分小さいと仮定する．すると，

$$\frac{1}{t}\{f(\mathbf{x}+t\mathbf{v}) - f(\mathbf{x})\} = \sum_{i=1}^{n} v_i(\nabla_i f)(\mathbf{x}_{i-1} + s_i(\mathbf{x}_i - \mathbf{x}_{i-1}))$$

が従う．この式において $t \to 0$ とすると，左辺は方向微分 $(\nabla_{\mathbf{v}} f)(\mathbf{x})$ の定義そのものになる．一方，$t \to 0$ のとき $\mathbf{x}_i \to \mathbf{x}$ $(i = 0, 1, \cdots, n)$ であるから，偏導関数 $\nabla_i f$ の連続性により，右辺は $t \to 0$ のとき $\sum v_i(\nabla_i f)(\mathbf{x})$ に収束する．これで，方向微分 $(\nabla_{\mathbf{v}} f)(\mathbf{x})$ が存在し，しかも (4.3) を満たすことが証明された．[2]　□

関数 f が C^1 級であるとき，点 $\mathbf{x} \in X$ における偏微分係数 $(\nabla_i f)(\mathbf{x})$ を第 i 成分とするような n 次元列ベクトルを $(\nabla f)(\mathbf{x})$ と書き，f の点 $\mathbf{x}$ における**勾配ベク**

[2] この証明でも見たとおり，1 変数関数に対する平均値の定理は，多変数関数の微分に関連した結果を示す上でも極めて有効です．実際，この後も偏微分に関係した定理・命題を証明するにあたって，平均値の定理が主要な役割を果たすことが何度もあるはずです．

トルと呼びます：

$$(\nabla f)(\mathbf{x}) = \begin{pmatrix} (\nabla_1 f)(\mathbf{x}) \\ \vdots \\ (\nabla_n f)(\mathbf{x}) \end{pmatrix}.$$

すると，(4.3) の右辺は，行ベクトル $\mathbf{v}$ と列ベクトル $(\nabla f)(\mathbf{x})$ の積 $\mathbf{v}(\nabla f)(\mathbf{x})$ と一致します．[3] したがって，方程式 (4.3) 自体も

$$(\nabla_{\mathbf{v}} f)(\mathbf{x}) = \mathbf{v}(\nabla f)(\mathbf{x}) \tag{4.4}$$

と書き換えられます．

いよいよ極値問題に関する議論を始めることが可能になりました．最初に示すべきは，次の命題です．

命題 4.8.

関数 f が C^1 級，かつ点 $\mathbf{x} \in X$ において極大，あるいは極小であるならば，$(\nabla_i f)(\mathbf{x}) = \cdots = (\nabla_n f)(\mathbf{x}) = 0$，すなわち，

$$(\nabla f)(\mathbf{x}) = \mathbf{0}$$

が成り立つ．

証明． 番号 $i = 1, \cdots, n$ を任意にとる．実数 t（ただし，$|t|$ は十分小さいとする）に対し，実数 $g(t) = f(\mathbf{x} + t\mathbf{e}_i)$ を対応させる関数を考える：とくに $g(0) = f(\mathbf{x})$ であることに注意しよう．すると，f が点 $\mathbf{x}$ において極大，あるいは極小であるという仮定により，関数 $g(t)$ もまた $t = 0$ において極大，ないし極小でなければならない．したがって，$g'(0) = 0$ を得る．ところが，偏微分係数の定義によれば，この等式の左辺は $(\nabla_i f)(\mathbf{x})$ にほかならない．したがって，$(\nabla_i f)(\mathbf{x}) = 0$ が結論される． □

等式 $(\nabla f)(\mathbf{x}) = \mathbf{0}$ が成り立つとき，$\mathbf{x}$ は f の**臨界点**と呼ばれます．あるいは，

$$(\nabla_1 f)(\mathbf{x}) = \quad \cdots \quad = (\nabla_n f)(\mathbf{x}) = 0 \tag{4.5}$$

を $\mathbf{x}$ に対する連立方程式と見なしたとき，その解が臨界点にほかなりません．

[3] 厳密に言うならば，n 次元行ベクトル $\mathbf{v}$ と n 次元列ベクトル $(\nabla f)(\mathbf{x})$ の積は，$(1, 1)$ 型の行列，すなわち実数ひとつを括弧で「くくった」ものになります．ここでは，それを，その成分である実数と同一視しています．

問題 4.9. ――――――――――――――――――――――――――――
関数 (4.2) の臨界点をすべて求めよ．

―――――――――――――――――――――――――――――――――

解答. 問題 4.5 に対する解答により，連立方程式

$$\begin{cases} x^3 - 5x + 4y &= 0, \\ y^3 + 4x - 5y &= 0 \end{cases} \tag{4.6}$$

を解くこと，すなわち，その解をすべて求めることに問題は帰着される．まず，これらふたつの方程式の和，および差を計算すると，それぞれ $(x^3+y^3)-x-y=0$, $(x^3-y^3)-9x+9y=0$ を得る．左辺を因数分解すると，

$$\begin{cases} (x+y)(x^2-xy+y^2-1) &= 0, \\ (x-y)(x^2+xy+y^2-9) &= 0 \end{cases} \tag{4.7}$$

となる．まず，(4.7) の第 1 式から，$x+y=0$ または $x^2-xy+y^2-1=0$ でなければならないことが，また第 2 式から $x-y=0$ または $x^2+xy+y^2-9=0$ でなければならないことがわかる．したがって，以下に挙げる合計 4 つの場合が考えられる:

$$(a)\quad \begin{cases} x+y &= 0, \\ x-y &= 0 \end{cases} \qquad (b)\quad \begin{cases} x+y &= 0, \\ x^2+xy+y^2 &= 9 \end{cases}$$

$$(c)\quad \begin{cases} x^2-xy+y^2 &= 1, \\ x-y &= 0 \end{cases} \qquad (d)\quad \begin{cases} x^2-xy+y^2 &= 1, \\ x^2+xy+y^2 &= 9 \end{cases}$$

もちろん，(a) の場合には $(x,y)=(0,0)$ が解である．また，(b), (c) からは, それぞれ $(x,y)=(\pm 3,\mp 3)$, $(x,y)=(\pm 1,\pm 1)$（ただし，複号同順）を得る．最後に (d) の場合であるが，この場合には，まずそれら 2 式の和，および差をとることにより，$x^2+y^2=5$, $xy=4$ を得る．よって，$(x-y)^2=x^2-2xy+y^2=-3<0$ が従い，この場合には（実数）解は存在しないことがわかる．

（答）　$(x,y)=(0,0)$,　$(\pm 1,\pm 1)$,　$(\pm 3,\mp 3)$　（ただし，複号同順）

□

ところで，命題 4.8 によれば，

$$f \text{ が点 } \mathbf{x} \text{ において極値をとる} \quad \Longrightarrow \quad \mathbf{x} \text{ は } f \text{ の臨界点である} \tag{4.8}$$

が成り立ちます．ただし，**極値**とは，極大値・極小値の総称であることは言うまでもないと思います．これによれば，臨界点，すなわち，連立方程式 (4.5) の解をす

べて求めれば，その中に f が極大，あるいは極小になる点はもれなく含まれていることになります．したがって，たとえば先程の問題で扱った関数 (4.2) が極値をとる点をすべて求めたいときには，実際問題 4.9 で行ったように，その関数の臨界点をすべて求めることが，最初に行うべき仕事です．これを極値問題に対する解法の（第 1 段 ）と呼びましょう．しかし，すべての臨界点において，関数が極値をとるとは限りません．これは，(4.8) の逆は一般には成立しないことに起因します．そこで，次に問題になるのは，極値問題の解法の（第 1 段）で求めた臨界点のおのおのにおいて，問題の関数が極値をとるかどうか，また仮に極値をとる場合にはそれが極大値なのかそれとも極小値なのかを，判定しなければなりません．この作業を，極値問題に対する解法の（第 2 段）と呼ぶことにしましょう．

ここで 1 変数関数に対する極値問題はどうやって解かれたのかを思い出して下さい．$g = g(x)$ を，ある開区間上で定義された連続的に微分可能な関数としましょう．その極値を決定するためのひとつの方法は，次のようなものでした．まず g の導関数 $g'(x)$ を計算します．そして，とくに $g'(x) = 0$ となる x をすべて求めます．以上が，1 変数関数に対する極値問題の解法の（第 1 段）でした．次に,（第 2 段）に進みます．第 1 段で求めた x 以外の点においては g' の値はゼロではありませんから，その符号，すなわち正・負が決まります．そして，それをもとにして，関数 g の増減表，あるいはグラフを書くことにより，先程求めた $g'(x) = 0$ なる点 x のおのおのにおいて，g が極大，あるいは極小であるかを決定したわけです．さて，この方法はそのまま多変数関数に一般化できるでしょうか．1 変数関数 $g(x)$ に対する微分係数 $g'(x)$ は，多変数関数 f においては勾配ベクトル $(\nabla f)(\mathbf{x})$ に対応するはずです．しかしながら，1 変数関数の場合にはその微分係数 $g'(x)$ は実数であり，したがってその符号が議論できたのに反し，多変数関数に対する勾配ベクトル $(\nabla f)(\mathbf{x})$ は n 次元ベクトルですから,（極値問題に対し有効な仕方で）その符号を定義することは出来そうもありません．あるいは別の言い方をすれば，先に復習した 1 変数関数に対する極値問題の解法のとくに（第 2 段）においては，関数 g の単調性を g' の符号から判定し，それを本質的な仕方で利用していましたが，そもそも，関数の単調性は，その定義域が「1 次元的」であって初めて意味をなします．このような理由から，先程要約した，1 変数関数に対する極値問題の解法をそのまま多変数関数に拡張するのは極めて困難であろうことが予測されます．

しかし，1 変数関数の場合には，極値問題に対しもうひとつ別の解法がありました．関数の 2 階微分を利用する方法です．それを述べるために，g をある（開）区間上で定義された C^2 級関数とします．そのとき，g が極小・極大となる点は，次のようにしても決定することが可能でした．

（第 1 段）　1 階導関数 $g'(x)$ を計算し，とくに $g'(x) = 0$ となる点をすべて求めよ．

（第 2 段）　（第 1 段）で求めた各 x において，2 階微分係数 $g''(x)$ を計算せよ．そして，

$$\begin{cases} g''(x) > 0 & \Longrightarrow & g \text{ は } x \text{ において極小} \\ g''(x) < 0 & \Longrightarrow & g \text{ は } x \text{ において極大} \end{cases}$$

により，g が x において極小・極大であるかを判断せよ．

（第 1 段）は以前の解法におけるそれと全く変わりありません．違いは（第 2 段）にあります．ところで，そこで主張されていることの正当性は，2 階導関数の符号と関数の凹凸との間にある関係 — すなわち，2 階導微分が正ならば下に凸，また逆に負ならば上に凸であること — により保証されます．またテイラーの定理をその代わりに用いることも可能です．ちなみに,（第 2 段）において $g''(x) = 0$ となる場合には，g が 臨界点 x に おいて極値をとるか，また極値をとる場合には，極大値をとるのか極小値をとるのか，上記の方法により決定することは出来ません．その場合には，g の点 x におけるより高次の微分係数を計算し，テイラーの定理を適用してやる必要があります．

ところで，1 変数関数極値問題に対するこれらふたつの方法を比べたとき，第 2 の解法の短所は C^2 級を仮定する必要がある点です．しかし,「自然」な関数は C^2 級であることが多いので，これが問題になることはあまり多くないはずです．また，2 階微分に対しても，臨界点においてのみその符号を決定すれば十分で，その意味では，計算量がそう増えるわけでもないことに注意しましょう．

多変数関数に対しその極値を決定する問題も，基本的にはたったいま述べた 1 変数関数に対する方法を拡張することにより解決されます．ただ，そのためには多変数関数に対しても 2 階微分を考える必要があります．さて，いま f が C^1 級であるとしましょう．すると，各 $j = 1, \cdots, n$ に対し，f の偏導関数 $\nabla_j f = \partial f/\partial x_j$ が再び X 上の関数として定義されます．もし，$\nabla_j f = \partial f/\partial x_j$ の 変数 x_i による偏微分係数 $(\nabla_i(\nabla_j f))(\mathbf{x})$ が点 $\mathbf{x}$ において存在するとき，それを f の $\mathbf{x}$ における **2 階偏微分係数**と言い，記号 $(\nabla_{ij}^2 f)(\mathbf{x})$，ないし $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})$ により表します．すなわち，

$$(\nabla_{ij}^2 f))(\mathbf{x}) = (\nabla_i(\nabla_j f))(\mathbf{x}), \quad \text{あるいは} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\mathbf{x}) \tag{4.9}$$

とするわけです．ただし，同じ変数で 2 度偏微分することにより得られる $\dfrac{\partial^2 f}{\partial x_i \partial x_i}(\mathbf{x})$

は $\dfrac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$ と略記されます．また，X の各点 $\mathbf{x}$ に おいて 2 階偏微分係数 $(\nabla_{ij}^2 f)(\mathbf{x}) = (\partial^2 f/\partial x_i \partial x_j)(\mathbf{x})$ が存在するとき，各点 $\mathbf{x}$ に対しこの 2 階偏微分係数を対応させる関数 $\nabla_{ij}^2 f = \partial^2 f/\partial x_i \partial x_j$ を，関数 f の 2 **階偏導関数**と呼びます．

ここで，とくに $\nabla_{ij}^2 f = \partial^2 f/\partial x_i \partial x_j$ と $\nabla_{ji}^2 f = \partial^2 f/\partial x_j \partial x_i$ とを比較してみましょう．第 1 のものは，関数 f を最初変数 x_j で次いで変数 x_i で偏微分して得られたのに対し，第 2 のものは最初 x_i で偏微分し次に x_j で偏微分することにより得られました．注意すべきは，x_i, x_j，これらふたつの変数で偏微分する，その順序が逆になっていることです．それが理由で，$\nabla_{ij}^2 f = \partial^2 f/\partial x_i \partial x_j, \nabla_{ji}^2 f = \partial^2 f/\partial x_j \partial x_j$，両方が存在するとしても，必ずしもそれらが一致するとは限りません．しかし，これらふたつの偏微分が一致しないような関数は，言ってみるならば，かなり不自然な，あるいは病理的な関数です．そこで，そのような病理的な関数を排除するための条件を導入します．

定義 4.10.

関数 f が，まず第一に C^1 級であり，さらに，すべての 2 階偏導関数 $\nabla_{ij}^2 f = \partial^2 f/\partial x_i \partial x_j$ $(i, j = 1, \cdots, n)$ が存在してしかも X 上連続であるならば，f は C^2 **級**である，と言われる．

このとき，次が成り立ちます．

命題 4.11.

もし関数 f が C^2 級であるならば，すべての $i, j = 1, \cdots, n$ に対し，

$$\nabla_{ij}^2 f = \nabla_{ji}^2 f, \qquad \text{すなわち} \qquad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

が成り立つ．

証明． ふたつの変数 x_i, x_j による偏微分の可換性を示すにあたっては，それ以外の変数は定数と考えたままでよいから，結局問題は 2 変数関数の場合に帰着される．そこで，$f(x, y)$ を $\mathbb{R}^2$ のある開集合上で定義された C^2 級関数とする．示すべきは，

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

である．まず，偏微分の定義から，

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x \partial y}(a,b) \\
= \quad & \lim_{h\to 0}\frac{1}{h}\left\{\frac{\partial f}{\partial y}(a+h,b)-\frac{\partial f}{\partial y}(a,b)\right\} \\
= \quad & \lim_{h\to 0}\frac{1}{h}\left\{\lim_{k\to 0}\frac{1}{k}\bigl(f(a+h,b+k)-f(a+h,b)\bigr)\right. \\
& \qquad\qquad\qquad\qquad \left. -\lim_{k\to 0}\frac{1}{k}\bigl(f(a,b+k)-f(a,b)\bigr)\right\} \\
= \quad & \lim_{h\to 0}\lim_{k\to 0}\frac{\Delta}{hk}
\end{aligned}$$

を得る．ただし，

$$\Delta = f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b)$$

である．ここで，$g(x)=f(x,b+k)-f(x,b)$ とおけば，明らかに $\Delta = g(a+h)-g(a)$ が成り立つ．一方，

$$\frac{dg}{dx}(x)=\frac{\partial f}{\partial x}(x,b+k)-\frac{\partial f}{\partial x}(x,b)$$

であるから，$g(x)$ は微分可能である．したがって，$g(x)$ に平均値の定理を適用することができる．そして，その結果，

$$\begin{aligned}
\Delta = g(a+h)-g(a) &= h\tfrac{dg}{dx}(a+\alpha h) \\
&= h\left\{\tfrac{\partial f}{\partial x}(a+\alpha h,b+k)-\tfrac{\partial f}{\partial x}(a+\alpha h,b)\right\}
\end{aligned}$$

（ただし，α は $0\le\alpha\le 1$）を得る．さらに，最後の等号の右辺において，変数 y に対し平均値の定理を適用すれば，

$$\Delta = hk\,\frac{\partial^2 f}{\partial y\partial x}(a+\alpha h,b+\beta k)$$

（ただし，$0\le\beta\le 1$）を得る．しかも，仮定から $\partial^2 f/\partial y\partial x$ は連続であるから，

$$\frac{\partial^2 f}{\partial x\partial y}(a,b)=\lim_{h\to 0}\lim_{k\to 0}\frac{\Delta}{hk}=\lim_{h\to 0}\lim_{k\to 0}\frac{\partial^2 f}{\partial y\partial x}(a+\alpha h,b+\beta k)=\frac{\partial^2 f}{\partial y\partial x}(a,b)$$

が結論される． □

C^2 級関数という，言ってみれば「自然」な関数に対しては，2 階偏微分係数は偏微分する順序によらず決まることが，この命題により保証されます．したがって，

f が C^2 級であるならば，X の各点 $\mathbf{x}$ において，対称な n 次正方行列

$$(\nabla^2 f)(\mathbf{x}) = \begin{pmatrix} (\nabla^2_{11} f)(\mathbf{x}) & \cdots & (\nabla^2_{1n} f)(\mathbf{x}) \\ \vdots & & \vdots \\ (\nabla^2_{n1} f)(\mathbf{x}) & \cdots & (\nabla^2_{nn} f)(\mathbf{x}) \end{pmatrix} \tag{4.10}$$

が定まります．ただし，n 次正方行列 A が**対称**であるとは，その (i,j)-成分を a_{ij} と書いたとき，$a_{ij} = a_{ji}$ が すべての $i, j = 1, \cdots, n$ に対し成り立つことを意味します．[4] たとえば，

$$\begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix}$$

は対称行列です．この例のように，行列の（左上から右下へ走る）対角線に関し対称に反転させる，すなわち，転置をとっても変わらない行列を対称行列と呼ぶわけです．なお,「行列が対称である」といったときには，その行列が正方行列であることも同時に意味することにします．とくに，C^2 級関数 f の 2 階偏微分係数から作られた対称行列 (4.10) を，関数 f の点 $\mathbf{x}$ における**ヘッセ行列**と呼びます．

問題 4.12. ————————————————————————

関数 (4.2) のヘッセ行列を計算せよ．

————————————————————————————————

解答.
$$(\nabla^2 f)(x, y) = \begin{pmatrix} 4(3x^2 - 5) & 16 \\ 16 & 4(3y^2 - 5) \end{pmatrix}$$
□

もちろん，多変数関数のヘッセ行列は，1 変数関数に対する 2 階導関数に対応するものと考えられます．そこで，1 変数極値問題に対する 2 階導関数を用いた解法を，多変数関数に適用できるよう拡張するためには，ヘッセ行列，あるいは，より一般に対称行列に対し，実数の符号に相当するような概念を定義してやる必要があろうことが予想されます．

定義 4.13. ————————————————————————

A を n 次対称行列とする．もし，ゼロベクトルとは異なる任意の n 次元行ベクトル $\mathbf{v}$ に対し，

$$\mathbf{v} A\, {}^t\mathbf{v} > 0^5$$

[4] この講義録では，対称行列はつねにその成分が実数であると仮定します．

が成り立つならば，A は**正定値**であると言われる．逆にゼロでないすべての n 次元行ベクトル $\mathbf{v}$ に対し，

$$\mathbf{v}A\,{}^t\mathbf{v} < 0$$

が成り立つならば，A は**負定値**であると言われる．また，A が正定値，ないし負定値であるとき，A は**定値**であると言われる．

正方行列に対し，もしそれが正定値，あるいは負定値ならば，必ず正則，すなわち，逆行列を持ちます．実際，もし仮に A が正則でない正方行列とすると，$\mathrm{Ker}\,A \neq \{\mathbf{0}\}$ [6] ですから，$A\,{}^t\mathbf{v} = \mathbf{0}$ となるような n 次元行ベクトル $\mathbf{v} \neq \mathbf{0}$ が存在します．そして，そのとき，もちろん $\mathbf{v}A\,{}^t\mathbf{v} = 0$ が成り立ちますから，A は正定値でも，負定値でもありません．ところで，実数に対してそれがゼロでないという条件は，正方行列の正則性に対応している，と考えられないこともありません．実際，$(1,1)$-行列 $A = (a)$ を考えると，それが正則である必要十分条件が，$a \neq 0$ であることは自明です．実数の場合には，もしそれがゼロでなければ，それは正あるいは負でなければならないことは言うまでもありません．それに反し，一般の次数の正方行列に対しては，正則であるからといって，正定値，あるいは負定値であるとは限りません．それを示すために，とくに行列 $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ を例に取りましょう．このとき，たとえば，$\mathbf{v} = (1\ 0)$ とすると，$\mathbf{v}A\,{}^t\mathbf{v} = 1$，一方，$\mathbf{v} = (0\ 1)$ に対しては $\mathbf{v}A\,{}^t\mathbf{v} = -1$ となります．ベクトル $\mathbf{v}$ の 取り方によって $\mathbf{v}A\,{}^t\mathbf{v}$ の符号が変化するので，A は定値ではありません．しかし，$\det A = -1 \neq 0$ から判るとおり，A は正則です．より一般に，次の事実が成り立ちます．

命題 4.14.
行列

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \text{\huge 0} & \\ & \text{\huge 0} & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{4.11}$$

に対し次が成り立つ：

---

[5] ここで，行ベクトル $\mathbf{v} = (v_1, \cdots, v_n)$ を $(1,n)$-行列と見なしたときのその転置行列として得られる列ベクトル $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ を，${}^t\mathbf{v}$ は表します．一方，n 次元行ベクトル $\mathbf{v}$, n 次正方行列 A, および n 次列ベクトル ${}^t\mathbf{v}$ をこの順で乗じることにより得られる $\mathbf{v}A\,{}^t\mathbf{v}$ は，厳密に言えば，$(1,1)$ 行列であり，したがってその符号の正負を問うことはできません．しかし，その $(1,1)$ 行列の成分は実数であり，その正負が考えられます．ここでは，記号を少々濫用して，その $(1,1)$ 行列の成分として得られる実数を，$\mathbf{v}A\,{}^t\mathbf{v}$ と表記しています．

[6] $\mathrm{Ker}\,A$ は行列 A が誘導する線形写像の核 (kernel) $\{\mathbf{v} \in \mathbb{R}^n : A\,{}^t\mathbf{v} = 0\}$ を表す記号です．

(1)　A が正則であるための必要十分条件は，$\lambda_1, \lambda_2, \cdots, \lambda_n \neq 0$ が成り立つことである；

(2)　A が正定値であるための必要十分条件は，$\lambda_1, \lambda_2, \cdots, \lambda_n > 0$ が成り立つことである；

(3)　A が負定値であるための必要十分条件は，$\lambda_1, \lambda_2, \cdots, \lambda_n < 0$ が成り立つことである．

ちなみに，(4.11) のような正方行列，すなわち，その (i,j)-成分を a_{ij} としたとき $a_{ij}=0$ が $i \neq j$ なるすべての番号 i,j に対し成立するような正方行列を，**対角行列**と呼びます．結局，命題 4.14 は，対角行列の正則性・定値性が，極めて容易に調べられることを主張しています．

命題 4.14 の証明． (1) これは，正方行列に対しそれが正則であるための必要十分条件はその行列の行列式がゼロでないこと，および $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$ からただちに帰結される．

(2) まず必要性を示すために，A が正定値であると仮定する．すると，とくに $\mathbf{e}_i A\, {}^t\mathbf{e}_i > 0$ が従う．ただし，$\mathbf{e}_i$ は第 i 成分が 1 で他の成分が全部ゼロである n 次元行ベクトル $(0, \cdots, \overset{i}{1}, \cdots, 0)$ を表す．ところが，A は (4.11) の形をした対角行列であるから，$A\, {}^t\mathbf{e}_i = \lambda_i\, {}^t\mathbf{e}_i$，よってまた $\mathbf{e}_i A\, {}^t\mathbf{e}_i = \lambda_i$ である．これで $\lambda_i > 0$ $(i = 1, \cdots, n)$ を得た．次に逆を証明しよう．A の形からただちに，$\mathbf{v} A\, {}^t\mathbf{v} = \sum_{i=1}^{n} \lambda_i v_i^2$ が任意の n 次元行ベクトル $\mathbf{v} = (v_1 \ \cdots \ v_n)$ に対し成り立つことが判る．よって，$\lambda_1, \cdots, \lambda_n > 0$ を仮定すれば，$\mathbf{v} A\, {}^t\mathbf{v} > 0$ (ただし，$\mathbf{v} \neq \mathbf{0}$ とする)，すなわち，A が正定値であることが従う．

(3) これは，(2) とまったく同じ仕方で証明できる．□

多変数関数の各臨界点において極大・極小の判定をするのに際して，最も重要な役割を果たす定理を述べる準備がようやく整いました．

定理 4.15．

関数 f が C^2 級であると仮定する．このとき，f の各臨界点 $\mathbf{x} \in X$ において次が成り立つ：

(1)　ヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が正定値ならば，f は $\mathbf{x}$ において極小である；

(2)　ヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が負定値ならば，f は $\mathbf{x}$ において極大である；

(3)　ヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が正則ではあるが，正定値でも負定値でもないならば，f は $\mathbf{x}$ において極大，極小のいずれでもない．

この定理において，ヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が正則でないときには，残念ながらこの方法によっては極大・極小の判定が出来ません．その点に関しては，1 変数関数の場合と事情は変わりません．なお，この定理の証明はこの章の最後の節まで保留することにします．

A.2　２次対称行列の定値性の判定

前節の最後で述べた定理 4.15 を適用することにより多変数関数に対する極値問題を解くにあたっては，臨界点のおのおのにおいて，ヘッセ行列という一種の対称行列が，正定値あるいは負定値か，またはそのどちらでもないかを判定せねばなりません．そこでこの節では，とくに 2 次対称行列に話を限って，その定値性を判定する方法を勉強し，それを用いて 2 変数関数の極値問題を実際に解いてみたいと思います．なお，一般の次数の対称行列の定値性については，次節で議論する予定です．

とくに，2 次正方行列に対しては，その定値性の判定はいたって初等的な問題ですので，皆さん自身に考えてもらうことにします．

問題 4.16.

2 次対称行列 $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ は，いつ正定値，あるいは負定値になるか．

解答. $\mathbf{v} = (v\ w)$ とすると，

$$\mathbf{v}A\,{}^t\mathbf{v} = v(av + bw) + w(bv + cw) = av^2 + 2bvw + cw^2.$$

まず $a = 0$ の場合を考える．このときには，$\mathbf{v}A\,{}^t\mathbf{v} = w(2bv + cw)$，しかも b, c がいかなる値を取ろうとも，$2bv + cw = 0$ を満たす $(v\ w) \neq (0, 0)$ が存在することは明らかである．したがって，$a = 0$ の場合には A は定値ではあり得ない．

つぎに $a \neq 0$ と仮定しよう．すると，

$$\mathbf{v}A\,{}^t\mathbf{v} = a\left\{\left(v + \frac{b}{a}w\right)^2 + \frac{ac - b^2}{a^2}w^2\right\}$$

を得る．ここで，右辺に現れる $ac - b^2$ は，A の行列式 $\det A$ にほかならない．さらに，もし $\det A = ac - b^2 < 0$ が成り立つならば，上式右辺の $\{\cdots\}$ の部分は，

$(v\ w)$ のとりかたに応じて符号を変えることに注意しよう．これより，この場合にも A が定値でないことが従う．一方，$\det A = 0$ のときには，A は正則ではないから，また定値でもないことはすでに述べた通りである．残るは $\det A > 0$ の場合であるが，このときには，A が正定値であるための必要十分条件は $a > 0$ であること，また，逆に A が負定値であるための必要十分条件は $a < 0$ であることは，ともに明らかである．以上から，結局次の結論が導かれる．

（答） $\begin{cases} \det A > 0 \text{ かつ } a > 0 \text{ のときおよびそのときに限り，} A \text{ は正定値,} \\ \det A > 0 \text{ かつ } a < 0 \text{ のときおよびそのときに限り，} A \text{ は負定値.} \end{cases}$

□

いま皆さんが解いた問題，それと，命題 4.8，および定理 4.15 とから，f が 2 変数関数の場合には，それが極値をとる点を次のようにして決定できることが結論されます．

[2 変数関数に対する極値問題の解法] $f = f(x, y)$ を，平面 $\mathbb{R}^2$ の開集合 X 上で定義された C^2 級関数とする．

（第 1 段） まず，1 階偏導関数 $\partial f/\partial x, \partial f/\partial y$ を計算し，ついで連立方程式

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0$$

を (x, y) について解くことにより，f の臨界点 (x, y) をすべて求める．

（第2段） 上で求めた f の臨界点 (x, y) のおのおのにおいて，f のヘッセ行列

$$(\nabla^2 f)(x, y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(x, y) & \dfrac{\partial^2 f}{\partial x \partial y}(x, y) \\[2ex] \dfrac{\partial^2 f}{\partial y \partial x}(x, y) & \dfrac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix},$$

およびその行列式 $\det(\nabla^2 f)(x, y)$ を計算する．さらに，その点における極小・極大を，次の規則に基づき判定する：

$$\begin{cases} \det(\nabla^2 f)(x, y) > 0, \\ \quad \text{かつ} \quad \dfrac{\partial^2 f}{\partial x^2}(x, y) > 0 & \Longrightarrow \quad f \text{ は点 } (x, y) \text{ において極小；} \\ \det(\nabla^2 f)(x, y) > 0, \\ \quad \text{かつ} \quad \dfrac{\partial^2 f}{\partial x^2}(x, y) < 0 & \Longrightarrow \quad f \text{ は } (x, y) \text{ において極大；} \\ \det(\nabla^2 f)(x, y) < 0 & \Longrightarrow \quad f \text{ は } (x, y) \text{ において極大,} \\ & \qquad \text{極小のいずれでもない.} \end{cases} \tag{4.12}$$

ただし，$\det(\nabla^2 f)(x,y)=0$ の場合には，ここで述べた方法は適用できないことに注意されたい．

この 2 変数関数に対する極値問題の解法が正しく理解できたことを確認するために，再び問題を解いてもらうことにします．

問題 4.17. ————————————————

関数 (4.2) が極小・極大となる点をすべて決定せよ．

————————————————

問題の関数に対しては，そのすべての臨界点をすでに問題 4.9 において求めてあります．それによると，$(0,0)$, $(\pm 1,\pm 1)$, $(\pm 3,\mp 3)$（ただし，複号同順）が f の臨界点でした．あとはその各点に対し，(4.12) を適用するだけです．ここで，問題 4.12 に対する解答を思い出して下さい：

$$(\nabla^2 f)(x,y)=\begin{pmatrix} 4(3x^2-5) & 16 \\ 16 & 4(3y^2-5) \end{pmatrix}$$

でした．これより，とくに

$$\det(\nabla^2 f)(x,y)=16\left\{(3x^2-5)(3y^2-5)-4^2\right\}$$

を得ます．まず，f の臨界点 $(0,0)$ を考えてみましょう．$\det(\nabla^2 f)(0,0)=16(5^2-4^2)>0$，かつ $(\partial^2 f/\partial x^2)(0,0)=4\cdot(-5)<0$ ですから，(4.12) により，f は点 $(0,0)$ において極大であることがわかります．一方，臨界点 $(\pm 1,\pm 1)$ においては，$\det(\nabla^2 f)(\pm 1,\pm 1)=16(2^2-4^2)<0$ ですから，(4.12) によれば，f はこれらの点においては極値をとらないことが分かります．そして，最後に臨界点 $(\pm 3,\mp 3)$ ですが，これらの点においては $\det(\nabla^2 f)(\pm 3,\mp 3)=16(22^2-4^2)>0$，かつ $(\partial^2 f/\partial x^2)(\pm 3,\mp 3)=4(27-5)>0$ ですから，f はこれらの点において極小となることが，再び (4.12) より従います．以上をまとめると，

f は点 $(0,0)$ で極大，点 $(\pm 3,\mp 3)$（複号同順）で極小；

これが問題 4.17 に対する解答です．

念のため，同種の問題をもう 1 題だけ出しておきます．

問題 4.18. ————————————————

次の関数が極小・極大となる点を決定せよ：

$$f(x,y)=x^4+y^4-4xy+1. \tag{4.13}$$

————————————————

解答．（第 1 段）まず，関数 f の臨界点を決定する．そのために，f の 1 階偏導関数を計算すると，

$$\frac{\partial f}{\partial x}(x,y) = 4(x^3 - y), \qquad \frac{\partial f}{\partial y}(x,y) = 4(y^3 - x)$$

となる．したがって，$(\partial f/\partial x)(x,y) = (\partial f/\partial y)(x,y) = 0$ の解として，

$$(x,y) = (0,0),\ (\pm 1, \pm 1) \qquad \text{（複号同順）}$$

を得る．これらが f の臨界点である．

（第 2 段）まず，関数 f のヘッセ行列を計算しよう：

$$(\nabla^2 f)(x,y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

である．よって，その行列式は

$$\det(\nabla^2 f)(x,y) = 16(3^2x^2y^2 - 1)$$

で与えられる．とくに，臨界点 $(0,0)$ においては，$\det(\nabla^2 f)(0,0) = -16 < 0$ であるから，f は $(0,0)$ において極値をとらない．一方，臨界点 $(\pm 1, \pm 1)$ においては，$\det(\nabla^2 f)(\pm 1, \pm 1) = 16(3^2 - 1) > 0$，かつ $(\partial^2 f/\partial x^2)(\pm 1, \pm 1) = 12 > 0$ であるから，f はこれらふたつの臨界点 $(\pm 1, \pm 1)$ において極小であることが分かる．

（答）　　点 $(\pm 1, \pm 1)$（複号同順）において極小．□

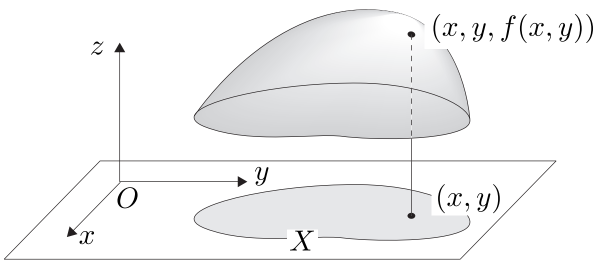


図 4.5: 2 変数関数のグラフ
f を $\mathbb{R}^2$ の開集合 X 上で定義された関数とする．点 (x,y) を X 内にとると，それに応じて点 $(x,y,f(x,y))$ が (x,y,z) 空間内に得られる．点 (x,y) が X 全体を動き回ったとき，点 $(x,y,f(x,y))$ の軌跡として得られるのが関数 f のグラフである．

ここで，2 変数関数に対する極値問題，およびその解法の幾何学的背景について，少々考えてみたいと思います．いま，$f = f(x,y)$ を，平面 $\mathbb{R}^2$ の開集合 X

で定義された C^2 級関数としましょう．各 $(x, y) \in X$ に対し，(x, y, z) 空間の点 $(x, y, f(x, y))$ をとります．これは，(x, y) 平面内の点 (x, y) のちょうど真上，または真下に位置します．ところで，$\mathbb{R}^2$ の部分集合 X は，(x, y) 平面内のある「図形」と捉えることが出来ます．その図形の中を点 (x, y) が勝手に動き回れば，それに連れて (x, y, z) 空間内の点 $(x, y, f(x, y))$ もその空間内を移動します．その軌跡を関数 $z = f(x, y)$ の**グラフ**と呼びます．たとえば，図 A.2 は，問題 4.18 で取り扱った関数 (4.13) のグラフです．また関数 (4.2) のグラフが図 4.7 に与えてあります．両者の差異がどこにあるのか，注意して観察して下さい．

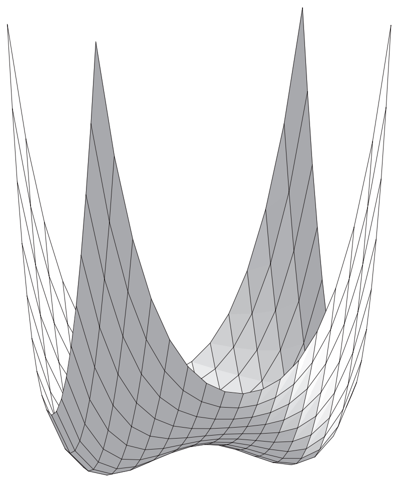


図 4.6: 関数 $z = x^4 + y^4 - 4xy + 1$ のグラフ

いま，$\mathbf{x} = (x, y) \in X$ および $\mathbf{v} = (v, w) \in \mathbb{R}^2$ を勝手に取ります．すると，(x, y) 平面内に，実数 t によりパラメーター付けられた直線 $\mathbf{x} + t\mathbf{v} = (x + tv, y + tw)$ を考えることが出来ます．この直線を含み，しかも z 軸に平行な平面が (x, y, z) 空間内に一意的に決まります．一般に，その平面と関数 $z = f(x, y)$ のグラフの交わりは，その平面内の曲線になります．そして，その曲線の点 $(t, z) = (0, f(x, y))$ における接線の傾きが，方向微分係数 $(\nabla_{\mathbf{v}} f)(x, y)$ に他ならないことが，方向微分の定義からただちに従います．このようにして，方向微分，とくに偏微分の意味が幾何学的に捉えられます．また，とくに点 (x, y) が $f(x, y)$ の臨界点であるときには，その関数のグラフ上の点 $(x, y, f(x, y))$ における接平面が水平，すなわち，(x, y) 平面に平行であることにも注意しておきましょう．

それでは，2 階偏微分係数はどうでしょう．そのために，補題をひとつ用意します．なお，以下に述べる補題は 2 変数関数に限らず，任意の多変数関数に対しても成り立つことに注意して下さい．

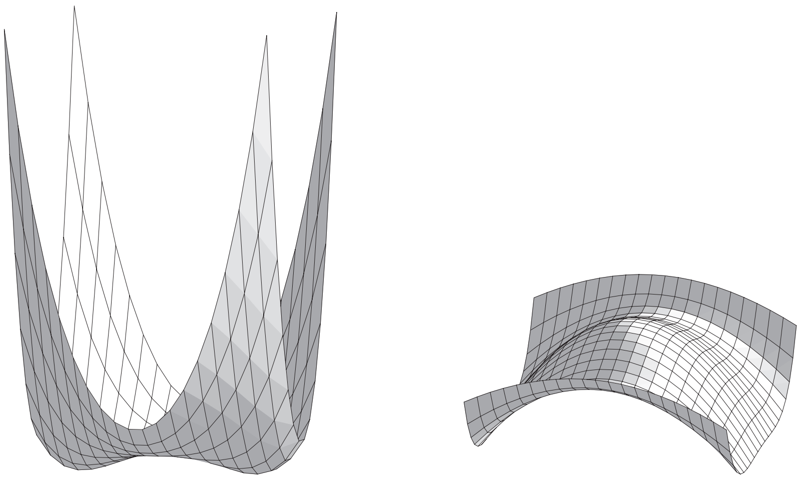


図 4.7: 関数 $z = x^4 + y^4 - 10x^2 + 16xy - 10y^2$ のグラフ
左側の図は，この関数のグラフをとくに $|x+y| \le 3.5$, $|x-y| \le 4.5$ の範囲で描いたもの．また，右はその中央部分の拡大図．真ん中に小さなこぶ状の突起があることに注意されたい．

補題 4.19.
$f : X \to \mathbb{R}$ を，n 次元ユークリッド空間 $\mathbb{R}^n$ の開集合 X の上で定義された C^2 級関数と，また $\mathbf{x}$ を X の任意の点，$\mathbf{v}$ を任意の n 次元行ベクトルとする．このとき，実数 t（ただし，$|t|$ は十分小）の関数 $f(\mathbf{x}+t\mathbf{v})$ に対し，その $t=0$ における 2 階微分係数は，$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v}$ に一致する：

$$\left.\frac{d^2}{dt^2}\right|_{t=0} f(\mathbf{x}+t\mathbf{v}) = \mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v}. \tag{4.14}$$

ただし，ここで $(\nabla^2 f)(\mathbf{x})$ は，f の点 $\mathbf{x}$ におけるヘッセ行列である．

この補題の関数 f に対し，そのヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ $(\mathbf{x} \in X)$ は，その (i,j) 成分を $(\nabla^2_{ij} f)(\mathbf{x})$ とするような n 次正方行列です．行ベクトル $\mathbf{v}$ を $\mathbf{v} = (v_1, \cdots, v_n)$ と成分表示したとき，(4.14) の右辺は，

$$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} = \sum_{i,j=1}^{n} v_i v_j (\nabla^2_{ij} f)(\mathbf{x})$$

と書くことができます．

補題 4.19 の証明. 命題 4.7 において

$$\frac{d}{dt} f(\mathbf{x}+t\mathbf{v}) = \sum_{i=1}^{n} v_i (\nabla_i f)(\mathbf{x}+t\mathbf{v}) \qquad (\mathbf{v} = (v_1 \ \cdots \ v_n))$$

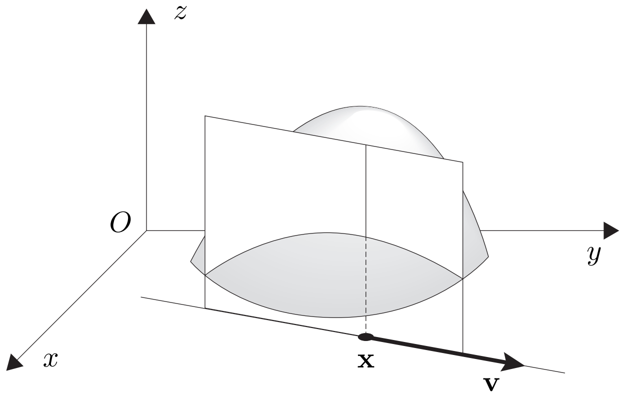


図 4.8:

を証明したのを思い起こそう．関数 f の代わりにその偏導関数 $\nabla_j f$ に対しこれを適用すれば，2 階偏微分係数の定義 $(\nabla^2_{ij} f)(\mathbf{x}) = (\nabla_i(\nabla_j f))(\mathbf{x})$ とから，

$$\frac{d}{dt}(\nabla_j f)(\mathbf{x}+t\mathbf{v}) = \sum_{i=1}^{n} v_i(\nabla^2_{ij} f)(\mathbf{x}+t\mathbf{v})$$

を得る．そして，これらふたつの等式から，

$$\frac{d^2}{dt^2}\Big|_{t=0} f(\mathbf{x}+t\mathbf{v}) = \frac{d}{dt}\Big|_{t=0} \sum_{j=1}^{n} v_j(\nabla_j f)(\mathbf{x}+t\mathbf{v}) = \sum_{i,j=1}^{n} v_i v_j(\nabla^2_{ij} f)(\mathbf{x})$$

が従う．この最後の等式の右辺は $\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v}$ に等しい．したがって，補題が従う． □

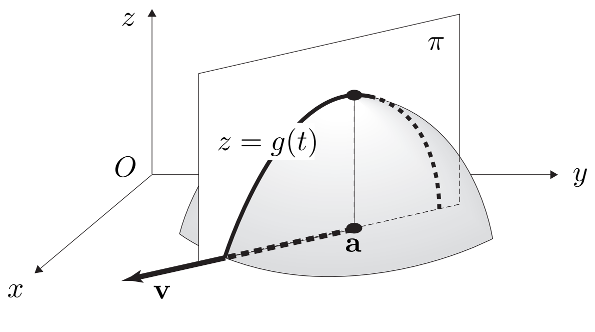


図 4.9:

さて，再び 2 変数関数 $f(x,y)$ の場合に帰ります．さらに，点 $\mathbf{x} = (x,y)$ を f の定義域 X 内の点と，また $\mathbf{v} = (v,w)$ を 2次元ベクトルとします．このとき，(x,y) 平面内の直線 $\mathbf{x}+t\mathbf{v} = (x+tv, y+tw)$ に着目し，その上に関数 f を制限するこ

とにより，実数 t（ただし，$|t|$ は十分小）のみの関数 $g(t) = f(x+tv, y+tw)$ が定義されるのも，先程と同様です．このとき，命題 4.7，およびたったいま証明した補題 4.19 によれば，

$$g'(0) = \mathbf{v}\,(\nabla f)(\mathbf{x}), \qquad g''(0) = \mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \tag{4.15}$$

が成り立ちます．さて，$\mathbf{x} = (x, y)$ がとくに関数 f の臨界点であると仮定しましょう．すると，定義により $(\nabla f)(x, y) = \mathbf{0}$ ですから，結局 $g'(0) = 0$ が成り立ちます．したがって，もし $g''(0) > 0$，すなわち，$g(t)$ が $t = 0$ の近辺で下に凸ならば，関数 $z = g(t)$ のグラフを (t, z) 平面内に描くと，それは関数 $z = t^2$ のグラフのように $t = 0$ において極小であることが判ります（図 4.9）．このグラフ $z = g(t)$ は，もとの 2 変数関数 $z = f(x, y)$ のグラフ，および直線 $(x+tv, y+tw)$ を含み z 軸に平行な平面 π との交わりとして得られたことを思い出しておいて下さい．さて，今度は (x, y) が f の臨界点であるという仮定に加えて，さらに，関数 f の点 (x, y) におけるヘッセ行列 $(\nabla^2 f)(x, y)$ が正定値であると仮定してみましょう．すると，ゼロでないどんな 2 次元行ベクトル $\mathbf{v}$ をとっても，それから得られる 1 変数関数 $g(t)$ が $g''(0) > 0$ を満たすことが，(4.15) より従います．これは，言い換えれば，先程の平面 π を，$\mathbf{v}$ を変化させることにより回転させても，その平面と関数 $z = f(x, y)$ の交わりとして得られる曲線 $z = g(t)$ は，$t = 0$ において極小であることを意味しています．したがって，関数 $z = f(x, y)$ のグラフは，点 $(x, y, f(x, y))$ の近辺では，お椀のような形をしていて，その一番底の部分が点 $(x, y, f(x, y))$ になっているはずです．そして，このことからとくに $f(x, y)$ が点 (x, y) において極小であることが納得されると思います．実際，問題 4.18 で扱った関数のグラフをもう一度見ると，その関数のふたつの極小点 $(\pm1, \pm1)$ の近辺では，その関数のグラフが，いま述べたように，お椀状になっていることが確認できるはずです．

逆に，もし点 (x, y) において f のヘッセ行列が負定値ならば，今度は関数 $z = f(x, y)$ のグラフは，点 $(x, y, f(x, y))$ の近辺で，お椀を伏せたような形をしていて，しかもそのてっぺんが点 $(x, y, f(x, y))$ になっています．

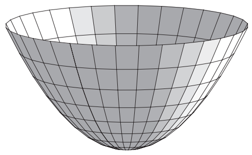
それでは，f の臨界点 (x, y) におけるヘッセ行列が，正則ではあるけれど定値ではないとき，一体どのようなことが起こるのでしょうか．そのためには，とくに 2 次関数

$$f_\pm(x, y) = \frac{1}{2}(x^2 - y^2)$$

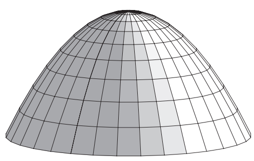
を例に取ると分かりやすいと思います．この関数に対しては，

$$(\nabla f_\pm)(x, y) = \begin{pmatrix} x \\ -y \end{pmatrix}, \qquad (\nabla^2 f_\pm)(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

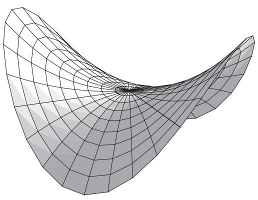
が成り立ちますから，原点 $(0,0)$ は $f_{\pm}(x,y)$ の唯一の臨界点で，しかもその点においてヘッセ行列は，正則ですが定値ではありません．図 4.10 (iii) に，そのグラフを描きました．先程の平面 π は，この場合 z 軸を含む平面になります．そして，その平面を z 軸を含んだまま回転させると，ある場合には関数 $z = f_{\pm}(x,y)$ のその平面による切り口として得られる曲線が下に凸になり，またある場合には上に凸になることが，図から容易に理解できると思います．そして，$f_{\pm}(0,0)=0$，かつ $f_{\pm}(x,y)$ は原点 $(x,y)=(0,0)$ の近辺で符号を変えますから，原点においては極小でも極大でもありません．問題 4.18 において，原点 $(0,0)$ が臨界点ではあったけれど，そこで極小値も，極大値もとらなかったのも，これと同じ理由によるものです．実際，そのグラフは，原点の近くで関数 $f_{\pm}(x,y)=(x^2-y^2)/2$ のグラフと似た形をしています．その形状が，馬の鞍に似ていることから，このような臨界点，すなわち，2 変数関数の臨界点であって，そこにおいてヘッセ行列が正則ではあるけれど定値ではないようなものを，**鞍点**と呼ぶことがあります．[7]



(a) $z=\frac{1}{2}(x^2+y^2)$ のグラフ



(b) $z=-\frac{1}{2}(x^2+y^2)$ のグラフ



(c) $z=\frac{1}{2}(x^2-y^2)$ のグラフ

図 4.10:

ちなみに，ヘッセ行列が正則かつ定値であるような 2 変数関数の最も単純な例も，やはり，2 次関数により与えられます．実際，

$$f_+(x,y)=\frac{1}{2}(x^2+y^2), \qquad f_-(x,y)=-\frac{1}{2}(x^2+y^2)$$

とおくと，

$$(\nabla f_+)(x,y)=\begin{pmatrix} x \\ y \end{pmatrix}, \qquad (\nabla^2 f_+)(x,y)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$(\nabla f_-)(x,y)=\begin{pmatrix} -x \\ -y \end{pmatrix}, \qquad (\nabla^2 f_-)(x,y)=\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

[7]「馬の鞍」と言われてもピンとこない場合には，ポテトチップを思い浮かべて下さい．

ですから，原点はこれらふたつの関数のおのおのの臨界点になります．そして，$f_+(x,y)$ の原点におけるヘッセ行列は正定値，また $f_-(x,y)$ に対するそれは負定値であることが分かります．また，図 4.10 (i), (ii) にこれらふたつの関数のグラフを描いておきました．以上の 3 つの 2 次関数 f_+, f_-, $f_\pm$ は，2 変数関数に対する極値問題の解法のとくに（第 2 段）における場合分けを忘れてしまった場合，それを思い出すのにも有効であることを，最後に指摘しておきましょう．

A.3　一般次数の対称行列の定値性の判定

さて次に，次数が 3 以上の対称行列に対しても，その定値性を判定する方法を考えなければなりません．その出発点となるのが，次に述べる定理です．

定理 4.20.

A が n 次対称行列であるとき，

$$R^{-1}AR = {}^tRAR = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \text{\huge 0} & \\ & \text{\huge 0} & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{4.16}$$

なる n 次直交行列 R が必ず存在する．

　ただし，n 次実正方行列 R が**直交行列**であるとは，

$$R\,{}^tR = {}^tRR = I \qquad (I \text{ は } n \text{ 次単位行列})$$

が成り立つことを意味します．とくに，直交行列 R に対し，

$$R^{-1} = {}^tR$$

が成り立つことに注意しましょう．この定理の証明も後回しにすることにして，とりあえず話を進めることにします．ところで，この定理で主張されていることを一言で述べるならば，対称行列は直交行列を用いて**対角化**することが可能である，と表現できます．対角化することの利点のひとつは，次の系の証明においても見るとおり，定値性等の性質を有しているか否かの判定が極めて容易になることです．

系 4.21.

n 次対称行列 A に対し，(4.22) がなりたつような直交行列 R をとる．このとき，

(1)　A が正定値であるための必要十分条件は，$\lambda_1, \cdots, \lambda_n > 0$ が成り立つことである．

(2)　また，A が負定値であるための必要十分条件は，$\lambda_1, \cdots, \lambda_n < 0$ が成り立つことである．

証明. 最初に証明の準備として，任意の n 次対称行列 A，および n 次直交行列 R に対し，

$$\begin{cases} A \text{ が正定値} & \Longleftrightarrow \quad R^{-1}AR \text{ が正定値}, \\ A \text{ が負定値} & \Longleftrightarrow \quad R^{-1}AR \text{ が負定値} \end{cases} \tag{4.17}$$

が成り立つことを示す．まず，A が正定値であると仮定する．このとき，定義によれば，任意の n 次元行ベクトル $\mathbf{v} \neq \mathbf{0}$ に対し，$\mathbf{v}\,A\,{}^t\mathbf{v} > 0$ が成り立つ．ここで，$\mathbf{v} \neq \mathbf{0}$ は任意である．一方，直交行列は常に正則である．[8] したがって，$\mathbf{v} \neq \mathbf{0}$ である限り $\mathbf{v}R^{-1} = \mathbf{v}\,{}^tR = {}^t(R\,{}^t\mathbf{v}) \neq \mathbf{0}$ でなければならない．よって，先の不等式に現れる $\mathbf{v}$ を，$\mathbf{v}R^{-1}$（ただし，$\mathbf{v} \neq \mathbf{0}$）で置き換えることが可能である．そしてその結果，$(\mathbf{v}R^{-1})A\,{}^t(\mathbf{v}R^{-1}) > 0$ を任意の $\mathbf{v} \neq \mathbf{0}$ に対し得る．ところが，$(\mathbf{v}R^{-1})A\,{}^t(\mathbf{v}R^{-1}) = \mathbf{v}R^{-1}AR\,{}^t\mathbf{v}$ である．よって，任意の $\mathbf{v} \neq \mathbf{0}$ に対し $\mathbf{v}\,R^{-1}AR\,{}^t\mathbf{v} > 0$ が示されたことになり，$R^{-1}AR$ が正定値であることが結論される．

逆を示すには，$B = R^{-1}AR$ とおいたとき，$A = RBR^{-1} = Q^{-1}BQ$（ただし，$Q = R^{-1}$）が成り立ち，しかも $Q = R^{-1} = {}^tR$ が再び直交行列であることに気付けばよい．実際，たったいま証明したことから，ただちに，$B = R^{-1}AR$ が正定値であるならば $Q^{-1}BQ = A$ が正定値であることが従う．また，第 2 の主張，すなわち A が負定値であるための必要十分条件が，$R^{-1}AR$ が負定値であることも，全く同様な仕方で証明できる．

以上で (4.17) が証明された．それと，すでに示した命題 4.14 を組み合わせば，ただちに系の主張を得る．　□

この系 4.21 によれば，対称行列 A が正定値，あるいは負定値であるかどうかを判定するためには，それを適当な直交行列 R を用いて (4.22) のように対角化したとき，その結果得られる対角行列の対角成分がすべて正であるか，あるいはすべて負であるかを調べればよいことになります．そこで，(4.22) に現れる対角行列の対角成分 $\lambda_1, \cdots, \lambda_n$ が，どのように特徴づけられるかを考えてみましょう．

[8] この事実は，次のようにして示すことができます．R を直交行列とします．定義によれば，${}^tRR = I$ が成り立たねばなりません．この等式の両辺の行列式をとれば，$\det({}^tRR) = 1$ を得ます．ところが，$\det({}^tRR) = \det {}^tR \cdot \det R = (\det R)^2$ ですから，$\det R = \pm 1$，とくに $\det R \neq 0$ であることが分かります．よって，R が正則であることが結論されます．

ただそれに先立って，列ベクトルのみならず行列に対しても，その和とスカラー積を定義してやる必要があります．いま，$A = (a_{ij})$, $B = (b_{ij})$ をともに (m, n)-行列としましょう．そのときそれらの**和**が，$A + B = (a_{ij} + b_{ij})$，すなわち，

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

で定義されます．また，(m, n)-行列 $A = (a_{ij})$ と実数 λ に対し，その**スカラー積**が $\lambda A = (\lambda a_{ij})$，すなわち，

$$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

により定義されます．これらの定義はともに極めて自然で分かりやすいものですので，これ以上何の説明もいらないでしょう．

さて，話を元に戻しましょう．$\lambda_1, \cdots, \lambda_n$ を (4.22) に現れる対角行列の対角成分とします．さらに，番号 $i = 1, \cdots, n$ をとり，$\xi = \lambda_i$ と置きます．すると，$(\xi I - R^{-1}AR)\,{}^t\mathbf{e}_i = \xi\,{}^t\mathbf{e}_i - \lambda_i\,{}^t\mathbf{e}_i = \mathbf{0}$ が成り立ちます．ただし，$\mathbf{e}_i$ は第 i 成分のみが 1 で残りの成分がすべて 0 であるような n 次元行ベクトル $(0, \cdots, \overset{i}{1}, \cdots, 0)$ を表します．よって，行列 $\xi I - R^{-1}AR$ は正則でないことになります．したがって，$\det(\xi I - R^{-1}AR) = 0$ を得ます．一方，$R^{-1}(\xi I - A)R = \xi I - R^{-1}AR$ ですから，$\det(\xi I - R^{-1}AR) = \det(R^{-1}(\xi I - A)R) = \det R^{-1} \cdot \det(\xi I - A) \cdot \det R = \det(\xi I - A)$ を得ます．以上から，$\det(\xi I - A) = 0$，すなわち $\xi I - A$ が非正則であることが従います．よって，$A\,{}^t\mathbf{v} = \xi\,{}^t\mathbf{v}$ なる n 次元行ベクトル $\mathbf{v}$（ただし，$\mathbf{v} \neq \mathbf{0}$）が存在しなければなりません．そこで，この観察に基づき，次のような定義を与えることにします．

定義 4.22. ――――――――――――――――――――――――

A を n 次正方行列，また ξ を定数とする．もし，

$$A\,{}^t\mathbf{v} = \xi\,{}^t\mathbf{v}$$

なる n 次元行ベクトル $\mathbf{v} \neq \mathbf{0}$ が存在するならば，ξ は A の**固有値**であると言われる．

――――――――――――――――――――――――――――――

先程調べたことは，結局次のように述べることができます：(4.22) **に現れる対角成分** $\lambda_1, \cdots, \lambda_n$ **のおのおのは，行列** A **の固有値でなければならない．**

さて，それでは，与えられた正方行列に対し，その固有値，ないしその符号を決定するにはどうしたらよいでしょうか．実は，そのためには，任意の n 次正方行列 A に対し次が成り立つことに気付けば十分です：

$$\begin{aligned}
&\quad \xi \text{ が } A \text{ の固有値である} \\
\Longleftrightarrow &\quad \text{ある } \mathbf{v} \neq \mathbf{0} \text{ に対し，} A\,{}^t\mathbf{v} = \xi\,{}^t\mathbf{v} \text{ が成り立つ} \qquad \text{（固有値の定義より）} \\
\Longleftrightarrow &\quad \det(\xi I - A) = 0
\end{aligned}$$

この最後の条件に現れる行列式 $\det(\xi I - A)$ は，行列式の定義によれば，ξ を変数とする n 次多項式と考えられますから，結局固有値は，ξ に対する n 次代数方程式 $\det(\xi I - A) = 0$ の解として特徴付けられることが分かります．変数 ξ に対するこの多項式 $\det(\xi I - A)$ を正方行列 A の**固有多項式**と，また，ξ に対する代数方程式 $\det(\xi I - A) = 0$ を A の**固有方程式**と呼びます．さて，ここでちょっとした演習問題を解いてみて下さい．

問題 4.23. ————————————————————————

次の正方行列の固有値を求めよ．

(1) $\begin{pmatrix} 7 & 4 & 4 \\ 4 & 1 & -8 \\ 4 & -8 & 1 \end{pmatrix}$　　(2) $\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$

————————————————————————————

解答. (1) 問題の行列の固有多項式は，

$$\begin{aligned}
&\det\begin{pmatrix} \xi-7 & -4 & -4 \\ -4 & \xi-1 & 8 \\ -4 & 8 & \xi-1 \end{pmatrix} \\
=&\det\begin{pmatrix} \xi-7 & -4 & -4 \\ -4 & \xi-1 & 8 \\ & -\xi+9 & \xi-9 \end{pmatrix} && \text{（第 3 行）} \leftarrow \text{（第 3 行）} - \text{（第 2 行）} \\
=&(\xi-9)\cdot\det\begin{pmatrix} \xi-7 & -4 & -4 \\ -4 & \xi-1 & 8 \\ & -1 & 1 \end{pmatrix} && \text{（第 3 行）から } \xi-9 \text{ をくくり出す} \\
=&(\xi-9)\cdot\det\begin{pmatrix} \xi-7 & -8 & -4 \\ -4 & \xi+7 & 8 \\ & & 1 \end{pmatrix} && \text{（第 2 列）} \leftarrow \text{（第 2 列）} + \text{（第 3 列）} \\
=&(\xi-9)\{(\xi-7)(\xi+7)-(-8)\cdot(-4)\} && \text{（第 3 行）に関し余因子展開} \\
=&(\xi+9)(\xi-9)^2
\end{aligned}$$

である．従って，固有値は ± 9 である．

(2) 問題の行列に対し，

$$\det\begin{pmatrix}\xi-1 & 2\\ -1 & \xi-1\end{pmatrix}=(\xi-1)^2+2$$

がその固有多項式である．この多項式はどんな実数 ξ に対しても正の値をとるから，問題の行列は（実数の範囲で）固有値を持たないことが判る．

（答）　(1)　± 9　　(2) (実) 固有値を持たない．　□

いま解いた問題からも判るとおり，固有値は必ずしも（実数の範囲では）存在するとは限りません．ところが，定理 4.20 によれば，対称行列に関しては事情が異なります．実際，定理 4.20 といままでの議論とから，ただちに次の結果が得られます．

系 4.24.
n 次対称行列は，n 個の実固有値を有する．

与えられた n 次対称行列に対しその固有値を求めることは，その固有多項式と呼ばれる n 次代数方程式を解くことに他ならないことは，既に述べた通りです．そしていま解答を与えた問題のように，行列の次数 n が小さいときには，これは容易に実行可能です．ところが，次数 n が大きくなると，事情は全く異なります．というのも，5 次以上の代数方程式に対しては，たとえば 2 次方程式に対する解の公式に相当するような，解を求めるための公式が存在しないことが知られているからです．5 次以上の対称行列に対してその固有値の厳密な値を知ることは，このような原理的な困難が伴います．しかし，もう一度考えてみると，私達がここで本当に必要なのは固有値の値そのものではなく，その符号です．そして以下の補題に見る通り，固有値の符号を調べるだけなら，話は簡単に済みます．

補題 4.25.
変数 ξ に対する n 次代数方程式

$$\xi^n+c_{n-1}\xi^{n-1}+\quad\cdots\cdots\quad+c_1\xi+c_0=0 \tag{4.18}$$

が，n 個の実数解 $\lambda_1,\cdots,\lambda_n$ を持つとする．このとき，次が成り立つ：

(1)　$\lambda_1,\cdots,\lambda_n<0$ が成り立つための必要十分条件は，各 $i=0,1,\cdots,n-1$ に対し $c_i>0$ が成り立つことである；

(2)　$\lambda_1,\cdots,\lambda_n>0$ が成り立つための必要十分条件は，各 $i=0,1,\cdots,n-1$ に対し，$(-1)^{n+i}c_i>0$ が成り立つことである．

証明. 方程式 (4.18) の左辺を $f(\xi)$ とおく．すると，方程式 $(-1)^n f(-\xi) = 0$ に対し (1) を適用することにより (2) が得られる．したがって，(1) のみを証明すれば十分である．

まず，十分性を証明しよう．そのために，$c_0, c_1, \cdots, c_{n-1} > 0$ と仮定する．すると，任意の $\xi \geq 0$ に対し $f(\xi) > 0$ が成り立つから，方程式 $f(\xi) = 0$ の解 $\lambda_1, \cdots, \lambda_n$ はすべて負でなければならない．

次に必要性を証明するために，$\lambda_1, \cdots, \lambda_n < 0$ と仮定しよう．n 次方程式 $f(\xi) = 0$ が実数 $\lambda_1, \cdots, \lambda_n$ を解として持つという仮定から，$f(\xi) = (\xi - \lambda_1) \cdots (\xi - \lambda_n)$ を得る．この右辺を展開することにより，

$$c_i = \sum_{1 \leq k_1 < \cdots < k_{n-i} \leq n} (-\lambda_{k_1}) \cdots (-\lambda_{k_{n-i}})$$

が従う．$\lambda_1, \cdots, \lambda_n < 0$ であったから，$c_i > 0$ が結論される．□

いま証明した補題 4.25 をとくに対称行列の固有多項式に適用することにより，次の結果を得ます．

系 4.26.
n 次正方行列 A に対し，その固有多項式 $\det(\xi I - A)$ を変数 ξ に関し展開したときの i 次の係数を c_i とする：

$$\det(\xi I - A) = \xi^n + c_{n-1}\xi^{n-1} + \quad \cdots\cdots \quad + c_1 \xi + c_0.$$

このとき，

(1)　A が負定値であるための必要十分条件は，すべての $i = 0, 1, \cdots, n-1$ に対し，$c_i > 0$ が成り立つことである．また，

(2)　A が正定値であるための必要十分条件は，すべての $i = 0, 1, \cdots, n-1$ に対し，$(-1)^{n+i} c_i > 0$ が成り立つことである．

これでようやく一般の多変数関数に対する極値問題の解法が完成です．実際，系 4.26 と定理 4.15 を組み合わせれば，次の仕方で極値問題が解けることが結論されます．

[**多変数関数に対する極値問題の解法**]　$f : X \to \mathbb{R}$ を，n 次元ユークリッド空間の開集合 X 上で定義された C^2 級関数とする．

（第 1 段） 関数 f の臨界点をすべて求める．もちろんそのためには，まず f の勾配ベクトル $(\nabla f)(\mathbf{x})$ を各点 $\mathbf{x} \in X$ において計算し，ついで，$(\nabla f)(\mathbf{x}) = \mathbf{0}$ を $\mathbf{x}$ に対する方程式と思って解けばよい．

（第 2 段） 関数 f の各臨界点 $\mathbf{x} \in X$ において，f のヘッセ行列 $(\nabla^2 f)(\mathbf{x})$，およびその行列式 $\det((\nabla^2 f)(\mathbf{x}))$ を計算する．とくに，$\det((\nabla^2 f)(\mathbf{x})) \neq 0$ であることを確認せよ（もしそうでない場合には，残念ながら以下に述べる方法は適用できない）．さらに，ヘッセ行列の固有多項式 $\det(\xi I - (\nabla^2 f)(\mathbf{x}))$ を変数 ξ に関し展開し，

$$\det(\xi I - (\nabla^2 f)(\mathbf{x})) = \xi^n + c_{n-1}\xi^{n-1} + \quad \cdots \quad + c_1\xi + c_0$$

と置く．そして最後に，以下の規則に従って，f が臨界点 $\mathbf{x}$ において極大，極小，あるいはそのいずれでもないかを判定する：

(1) すべての $i = 0, 1, \cdots, n-1$ に対し $c_i > 0$ であるとき，f は $\mathbf{x}$ において極大である；

(2) すべての $i = 0, 1, \cdots, n-1$ に対し $(-1)^{n+i}c_i > 0$ であるとき，f は $\mathbf{x}$ において極小である；

(3) それ以外の場合，f は $\mathbf{x}$ において極値をとらない．

ちなみに，（第 2 段）において，もしある臨界点 $\mathbf{x}$ におけるヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が簡単な形をしていて，その定値性を系 4.26 を用いずに直接判定出来る場合には，もちろんそれでもかまいません．

この節の締めくくりとして，上述の解法を利用して解ける問題をひとつ出します．

問題 4.27.

次の関数の極値を決定せよ：

$$f(x, y, z) = xyz + \frac{1}{2}(x^2 + y^2 + z^2).$$

解答. （第 1 段） 与えられた関数 f の臨界点を決定する．そのために，まず各点 (x, y, z) における勾配ベクトルを計算すると，

$$(\nabla f)(x, y, z) = \begin{pmatrix} x + yz \\ y + zx \\ z + xy \end{pmatrix},$$

よって，連立方程式

$$\begin{cases} x + yz &= 0 \\ y + zx &= 0 \\ z + xy &= 0 \end{cases} \tag{4.19}$$

を解けばよい．(4.19) の（第 1 式）と（第 2 式）の和，および差より，それぞれ，$(x + y)(1 + z) = 0$, $(x - y)(1 - z) = 0$ を得る．これらふたつの等式のうち，第 1 のものから，$x + y = 0$ または $z + 1 = 0$ であることが，また第 2 のものから，$x - y = 0$ または $z - 1 = 0$ であることが従う．もちろん，$z + 1 = 0$ と $z - 1 = 0$ とは両立しないから，全部で以下に述べる 3 つの場合が考えられる：

(a)　$x + y = x - y = 0$,

(b)　$x + y = z - 1 = 0$,

(c)　$x - y = z + 1 = 0$.

このうち，(a) から $(x, y, z) = (0, 0, 0)$ が従うことは明らかである．また，(b) が成り立つ場合には，$z = 1$ と (4.19) の（第 3 式）から，$xy = -1$ が従う．これと，$x + y = 0$ とから，$(x, y, z) = (\pm 1, \mp 1, 1)$（複号同順）を得る．最後に (c) が成り立つ場合であるが，このときも $z = -1$ と (4.19) の（第 3 式）からまず，$xy = 1$ を，ついで $x - y = 0$ とから $(x, y, z) = (\pm 1, \pm 1, -1)$（複号同順）が従う．

以上で得た解をもう一度まとめて述べると，

$$(x, y, z) = (0, 0, 0), \quad (-1, 1, 1), \quad (1, -1, 1), \quad (1, 1, -1), \quad (-1, -1, -1)$$

となる．これらが，関数 f の臨界点のすべてである．

（第 2 段）最初に，関数 f のヘッセ行列を計算すると，

$$(\nabla^2 f)(x, y, z) = \begin{pmatrix} 1 & z & y \\ z & 1 & x \\ y & x & 1 \end{pmatrix}.$$

よって，その固有多項式は，

$$\begin{aligned} &\det \begin{pmatrix} \xi - 1 & -z & -y \\ -z & \xi - 1 & -x \\ -y & -x & \xi - 1 \end{pmatrix} \\ =&(\xi - 1)^3 - 2xyz - (x^2 + y^2 + z^2)(\xi - 1) \\ =&\xi^3 - 3\xi^2 + (3 - x^2 - y^2 - z^2)\xi \\ &\qquad\qquad + (x^2 + y^2 + z^2 - 2xyz - 1) \end{aligned} \tag{4.20}$$

で与えられる．また，この計算結果において $\xi = 0$ とおくことにより，点 (x, y, z) におけるヘッセ行列の行列式の値が $-(x^2 + y^2 + z^2 - 2xyz - 1)$ であることも判る．

まず，f の臨界点 $(0,0,0)$ においては，ヘッセ行列の固有多項式の奇数次の係数は正，偶数次の係数は負であるから，f のヘッセ行列は点 $(0,0,0)$ において正定値である．したがって，関数 f は点 $(0,0,0)$ において極小であることが結論される．

一方，f の $(0,0,0)$ 以外の臨界点 $(-1,1,1)$, $(1,-1,1)$, $(1,1,-1)$, $(-1,-1,-1)$ においては，ヘッセ行列の行列式は値 -4 をとるから，ヘッセ行列は正則である．ところが，これらの点においては，(4.20) の第 2 の等号の右辺における 1 次の係数がゼロになる．よって，ヘッセ行列はこれらの臨界点では定値ではない．したがって，関数 f がこれらの点で極値をとらないことが結論される．

（答）　f は点 $(0,0,0)$ において極小 □

A.4　コンパクト集合

前々節で 2 変数関数に対する，また前節では一般の多変数関数に対する極値問題の解法を勉強しました．とくにそこで一緒に解いた例題を通じて，その使用方法自体は十分理解して頂けたのではないかと思います．しかし，なぜそこでやったような仕方で極値問題が解けるのか，その解法の正当化はまだ終わっていません．定理 4.15，および定理 4.20 の証明がまだ済んでいないからです．もう一度極値問題に対する解法の導出を吟味してみれば判ることと思いますが，これらふたつの定理は，少なくとも理論的観点から言えば，極値問題に対する解法の核をなすものです．それが故に，それら定理を証明するにはそれなりの労力が必要であろうことも納得して頂けるのではないかと思います．実際，この節全部がそれらふたつの定理を証明するための準備に当てられます．なお，この節で取り扱う内容は，極値問題に限らず数学の種々の分野で重要な役割を果たすことを一言注意しておきましょう．

さて，$\{\mathbf{x}_\ell\}$ を，n 次元ユークリッド空間 $\mathbb{R}^n$ の点列としましょう．一方，k_ℓ $(\ell = 0,1,\cdots)$ を，$k_0 < k_1 < \cdots$ を満たす非負の整数とします．あるいは，言い換えるならば，$\{k_\ell\}$ は非負の整数からなる狭義単調増加数列です．このとき，$\mathbf{y}_\ell = \mathbf{x}_{k_\ell}$ $(\ell = 0,1,\cdots)$ とおくことにより，新たな点列 $\{\mathbf{y}_\ell\}$ が得られます．このようにして得られる点列 $\{\mathbf{y}_\ell\}$ を，もとの点列 $\{\mathbf{x}_\ell\}$ の**部分列**と呼びます．

定義 4.28.

X を $\mathbb{R}^n$ の部分集合とする．X の任意の点列に対し，その部分列であって，しかも X の点に収束するものが存在するとき，X は**点列コンパクト**，あるいは単に**コンパクト**であると言われる．

このコンパクト性という新しい概念を導入した最大の理由は，次の定理にあります．

定理 4.29 (最大値・最小値の定理).
$\mathbb{R}^n$ のコンパクト集合上で定義された連続関数は，必ず最大値・最小値を有する．

この定理は，1次元の場合（問題 2.16）と同様に証明できます．そこで証明は章末の問題として各自にお任せしたいと思います．定理によれば，コンパクト集合は，その上のいかなる連続関数も最大値・最小値を有するという，私たちにとっては極めて好都合な性質を持っています．しかし，どのような集合がコンパクト集合なのでしょう．それに答えなければ，せっかく証明した定理も無用の長物に過ぎません．次に話題とすべきは，コンパクト性の判定法です．しかし，その前に，コンパクト集合の定義自身を理解し直すために，簡単な反例を挙げさせて下さい．まず，実直線 $\mathbb{R}$ 内の区間 $[0,\infty)$，これはコンパクトではありません．それを見るためには，たとえば $x_\ell = \ell$ により定義される数列 $\{x_\ell\}$ を考えれば十分です．この数列は，もちろん $x_\ell \in [0,\infty)$ $(\ell = 0,1,\cdots)$ を満たします．一方，数列 $\{x_\ell\}$ は $\ell \to \infty$ のとき $+\infty$ に発散します．したがって，その任意の部分列もまた $\ell \to \infty$ としたとき発散します．このように，$\{x_\ell\}$ は収束部分列を持ちませんから，区間 $[0,\infty)$ はコンパクトでないことになります．また，区間 $(0,1]$ もコンパクトではありません．なぜならば，たとえば，その中の数列として，$x_\ell = 1/\ell$ なる数列 $\{x_\ell\}$ をとると，この数列自身，したがってまたその任意の部分列も，0 に収束しますが，0 は区間 $(0,1]$ に入っていないからです．

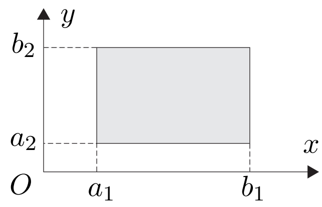
これらが，コンパクトでない集合の初等的な例です．それでは，コンパクト集合の例としては，どのようなものがあるでしょう．問題 2.15 を思い出して下さい．そこで示したことを，再度定理として述べましょう．

定理 4.30.
有界閉区間 $[a,b]$ はコンパクトである．

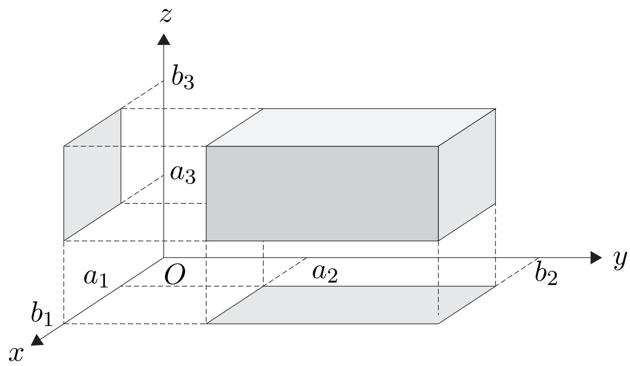
定理 4.30 により，有界閉区間がコンパクト集合であることが判りました．次に，この結果を高次元化したいと思います．そのために，ひとつ定義を与えましょう．いま，a_i, b_i $(i = 1,\cdots,n)$ を，$a_i \leq b_i$ を満たす実数としたとき，n 次元ユークリッド空間の部分集合

$$\{\mathbf{x} = (x_1 \cdots x_n) \in \mathbb{R}^n : a_1 \leq x_1 \leq b_1, \cdots, a_n \leq x_n \leq b_n\}$$

を考えることが出来ます．ちなみに，これは $n = 1$ のときには有界閉区間そのものです．また，$n = 2$ のときには平面内の長方形，また $n = 3$ のときには（3次元）空間内の直方体になっています（図 4.11）．そこで，一般の n に対しても，こ



(a) $[a_1, b_1] \times [a_2, b_2]$



(b) $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$

図 4.11: n 次元直方体 $(n = 2, 3)$

の集合を n **次元閉直方体**，あるいは単に，n **次元直方体**と呼ぶことにします．そして，それを表す記号としては，$[a_1, b_1] \times \cdots \times [a_n, b_n]$ を用いることにします．

系 4.31.
n 次元直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n]$ はコンパクトである．

証明. 次元 n に関する帰納法により証明する．まず，$n = 1$ の場合であるが，これは前出の定理 4.30 に他ならない．そこで，任意の n 次元直方体に対してはそのコンパクト性が証明できたと仮定して，$n+1$ 次元直方体に対してもコンパクト性を示すことが出来れば，定理の証明は完了する．$\{\mathbf{x}_\ell\}$ を $n+1$ 次元直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n] \times [a_{n+1}, b_{n+1}]$ の任意の点列とする．各項 $\mathbf{x}_\ell$ を，その第 1 成分から n 成分までが作る n 次元ベクトル $\mathbf{y}_\ell$ と第 $n+1$ 成分 z_ℓ とに分解する．すなわち，$\mathbf{x}_\ell = (\mathbf{y}_\ell, z_\ell)$ $(\mathbf{y}_\ell \in \mathbb{R}^n,\ z_\ell \in \mathbb{R})$ とおくわけである．これにより，n 次元点列 $\{\mathbf{y}_\ell\}$，および数列 $\{z_\ell\}$ が得られる．すると，とくに $\{\mathbf{y}_\ell\}$ は，n 次元直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n]$ の点列であるから，帰納法の仮定より，その直方体のある点 $\mathbf{y}$ に収束する部分列 $\{\mathbf{y}_{k_\ell}\}$ が存在する．このとき，最初に与えられた $n+1$ 次元点列 $\{\mathbf{x}_\ell\}$ の部分列 $\{\mathbf{x}_{k_\ell}\}$ も考えられることに注意しよう．同様に，もとの点列 $\{\mathbf{x}_\ell\}$ の各項の第 $n+1$ 成分から作られた数列 $\{z_\ell\}$ に対し，$\{z_{k_\ell}\}$ はその部分列になっ

ている．しかも，その各項は有界閉区間 $[a_{n+1}, b_{n+1}]$ の点である．したがって，定理 4.30 によれば，数列 $\{z_{k_\ell}\}$ の部分列であって，区間 $[a_{n+1}, b_{n+1}]$ の点に収束するようなものが存在する．そのような部分列を $\{z_{j_\ell}\}$, その極限を $z \in [a_{n+1}, b_{n+1}]$ としよう．また，$\{\mathbf{y}_{k_\ell}\}$ の部分列 $\{\mathbf{y}_{j_\ell}\}$ が考えられるが，それも再び点 $\mathbf{y}$ に収束する．よって，もとの数列 $\{\mathbf{x}_\ell\}$ は点 $\mathbf{x} = (\mathbf{y}, z)$ に収束する部分列 $\{\mathbf{x}_{j_\ell}\}$ を有することが分かった．以上で，直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n] \times [a_{n+1}, b_{n+1}]$ がコンパクトであることが結論される． □

n 次元ユークリッド空間 $\mathbb{R}^n$ の点列 $\{\mathbf{x}_\ell\}$ に対し，もし

$$\mathbf{x}_\ell \in [a_1, b_1] \times \cdots \times [a_n, b_n] \qquad (\ell = 0, 1, \cdots)$$

なる直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n]$ を見いだすことが出来るとき，点列 $\{\mathbf{x}_\ell\}$ は**有界**であると言うことにすると，いま証明した系 4.31 からただちに次の結論を得ます．

系 4.32.

ユークリッド空間 $\mathbb{R}^n$ の任意の有界点列は，必ず収束部分列を有する．

実は，n 次元ユークリッド空間のコンパクト部分集合は，別の言葉を用いて完全に特徴付けることが可能です．ただし，そのためには新たな概念をふたつほど導入する必要があります．

定義 4.33.

$\mathbb{R}^n$ の部分集合 X に対し，もし X を含むような n 次元直方体が存在する，すなわち，

$$X \subset [a_1, b_1] \times \cdots \times [a_n, b_n]$$

なる a_i, b_i $(i = 1, \cdots, n)$ が存在するならば，X は**有界**であると言われる．

少々乱暴な言い方をするならば，有界集合とはユークリッド空間 $\mathbb{R}^n$ の中での「広がり具合」が有限のものを言います．たとえば，$\mathbb{R}^n$ の開円板 $B(\mathbf{x}; r) = \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{x}| < r\}$，および閉円板 $\overline{B}(\mathbf{x}; r) = \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{x}| \leq r\}$ がともに有界であることが容易に判ります．一方，たとえば $\mathbb{R}^2$ の部分集合 $\{(x, y) \in \mathbb{R}^2 : |\mathbf{y}| \leq 1\}$ は有界でないことも明らかだと思います．

定義 4.34.

n 次元ユークリッド空間 $\mathbb{R}^n$ の部分集合 X が，以下に述べる条件を満たすとき，X は $\mathbb{R}^n$ の**閉集合**であると言う：

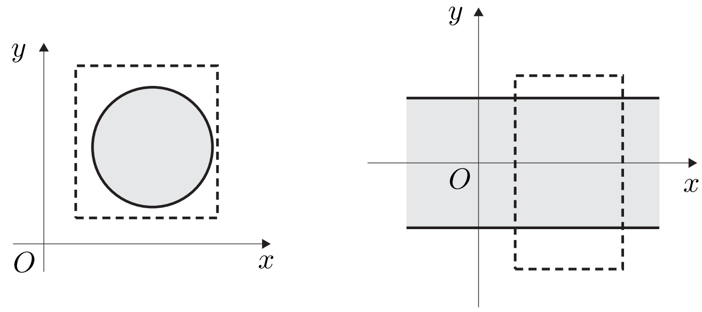


図 4.12:
平面内の円板はそれが開であるか閉であるかによらず有界である．ところが一方，領域 $\{(x,y) \in \mathbb{R}^2 : |y| \leq 1\}$ は有界ではない．

$\mathbb{R}^n$ の点 $\mathbf{x}$ に対し，もしその点に収束するような X の点列が存在するならば，$\mathbf{x} \in X$ が成り立つ．

例をいくつか挙げましょう．$\mathbb{R}$ の閉区間はいつでも閉集合です．しかしそれ以外のタイプの区間，たとえば半開区間 $[0,1)$ などは閉集合ではありません．また $\mathbb{R}^n$ 内の閉円板 $\{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{x}| \leq r\}$ （ただし $\mathbf{x} \in \mathbb{R}^n$）は閉集合ですが，開円板はそうではありません．これらの事実の証明は容易だと思いますので皆さんにお任せします．

定理 4.35.
ユークリッド空間 $\mathbb{R}^n$ の部分集合に対し，それがコンパクトであるための必要十分条件は，それが有界かつ閉であることである．

この後すぐに，実際に経験することと思いますが，コンパクト性に比べて，有界性および閉であることの判定の方が容易です．この定理の最大の利点はそこにあります．

定理 4.35 の証明. X を $\mathbb{R}^n$ の任意の部分集合とする．最初に十分性を示そう．そこで，X が有界かつ閉であると仮定する．一方，$\{\mathbf{x}_\ell\}$ を X の任意の点列とする．この点列が，X の点に収束するような部分列を持つことを示せばよいわけである．まず，X の有界性から，$X \subset [a_1, b_1] \times \cdots \times [a_n, b_n]$ なる n 次元直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n]$ が存在する．したがって，X の点列 $\{\mathbf{x}_\ell\}$ は，同時にこの直方体の点列でもある．ところが，系 4.31 によれば，直方体 $[a_1, b_1] \times \cdots \times [a_n, b_n]$

はコンパクトである．したがって，その点列 $\{\mathbf{x}_\ell\}$ が収束部分列 $\{\mathbf{y}_\ell\}$ を持つことが保証される．その極限を $\mathbf{y}$ としよう．あとは，$\mathbf{y} \in X$ を言えばよいだけであるが，これは X が閉であるという仮定，および $\{\mathbf{y}_\ell\}$ が X の 点列であるという事実からただちに従う．これで X がコンパクトであることが示された．

十分性は，対偶，すなわち，X が有界でない場合にも，また閉でない場合にも，X はコンパクトではあり得ないことを示すことにより証明される．まず，X が有界でない場合を考えよう．すると，いかなる a_i, b_i $(i = 1, \cdots, n)$ に対しても，$\mathbf{x} \notin [a_1, b_1] \times \cdots \times [a_n, b_n]$ なる $\mathbf{x} \in X$ が存在する．とくに，$a_1 = \cdots = a_n = -\ell$, $b_1 = \cdots = b_n = \ell$（ただし，$\ell = 0, 1, \cdots$）としたとき，この条件を満たす $\mathbf{x}$ をひとつとりそれを $\mathbf{x}_\ell$ とすることにより，X の点列 $\{\mathbf{x}_\ell\}$ を得る．もちろん，この点列，およびその任意の部分列が収束しないことは明らかである．したがって，この場合には，X がコンパクトでないことが証明された．

もし X が閉でないならば，X はまたコンパクトでもないことを最後に示したい．そこで，X が閉でないと仮定しよう．すると，X の点列 $\{\mathbf{x}_\ell\}$，およびユークリッド空間 $\mathbb{R}^n$ の点 $\mathbf{x}$ で，$\mathbf{x}_\ell \to \mathbf{x}$ $(\ell \to \infty)$ であるが，$\mathbf{x} \notin X$ であるようなものが存在せねばならない．このとき，$\{\mathbf{x}_\ell\}$ のいかなる部分列も再び点 $\mathbf{x}$ に収束するが，その点 $\mathbf{x}$ は X の点ではないので，これで X がコンパクトではないことが結論される．[9]

以上で，定理 4.35 の証明が終わった．□

A.5　定理 4.15 および定理 4.20 の証明

この節では，お約束通り，表題にあるふたつの定理の証明を与えます．まずは定理 4.15 の証明から始めましょう．この定理により，与えられた C^2 級関数が臨界点で極大，あるいは極小になるかが，その関数のヘッセ行列の定値性と関連づけられたわけです．少々くどいようではありますが，もう一度その定理をここで述べることにします．

定理 4.15

f を n 次元ユークリッド空間 $\mathbb{R}^n$ の開集合 X 上で定義された C^2 級関数と，また $\mathbf{x} \in X$ をその臨界点とする．このとき，次のおのおのが成り立つ：

(1)　もし $(\nabla^2 f)(\mathbf{x})$ が正定値ならば，f は $\mathbf{x}$ において極小である；

(2)　もし $(\nabla^2 f)(\mathbf{x})$ が負定値ならば，f は $\mathbf{x}$ において極大である；

[9] この定理における十分性，およびその証明は，それぞれ，区間 $[0, \infty)$，$(0, 1]$ が非コンパクトである，という事実，およびその証明の直接的な一般化になっていることに気付かれたでしょうか．

(3)　$(\nabla^2 f)(\mathbf{x})$ が正則ではあるが定値でないときには，f は $\mathbf{x}$ において極小でも極大でもない．

ここで，定理の主張 (1), (2), (3) のおのおのにおいて，その逆は必ずしも成立しないことに注意して下さい．実際，たとえば関数 $f(x,y)=x^4+y^4$ を考えると，その関数は原点 $(0,0)$ において極小ですが，原点におけるそのヘッセ行列の成分はすべてゼロですから，もちろん定値ではありません．この例が示すように，極小あるいは極大であるからと言って，ヘッセ行列が定値であるとは限らないわけです．しかし，定値性の条件を若干弱めることにより，この言明は正しいものになります．それを述べるために，弱い意味での定値性を定義することにしましょう．

定義 4.36.

A を n 次対称行列とする．もし，任意の n 次元行ベクトル $\mathbf{v}$ に対し，

$$\mathbf{v}A\,{}^t\mathbf{v} \geq 0$$

が成り立つならば，A は**半正定値**であると，また逆に，

$$\mathbf{v}A\,{}^t\mathbf{v} \leq 0$$

が成り立つならば，A は**半負定値**であると言われる．

いま述べた半正定値性・半負定値性の定義と，正定値性・負定値性に対するそれとを，ここで簡単に比較しましょう．たとえば，n 次対称行列 A に対し，その正定値性は，$\mathbf{v}A\,{}^t\mathbf{v}>0$ が，任意の n 次元行ベクトル $\mathbf{v}\neq\mathbf{0}$ に対し成り立つこととして定義されました．$\mathbf{v}\neq\mathbf{0}$ を仮定するのは，もちろん，ゼロベクトルに対しては，$\mathbf{0}A\,{}^t\mathbf{0}=0$ が成り立つからです．結局，不等号として等号付きのものを用いるか，それとも等号無しのものを使うか，その点が半正定値性と正定値性の定義の差異です．とくに，正定値であるならば必ず半正定値でなければなりません．また，同様に，もし負定値であるならば半負定値であるはずです．しかし，逆は一般には成立しません．そのことは，たとえばこれから述べる半定値性の判定方法からも，容易に見て取れます．

命題 4.37.

A を n 次対称行列，また R を，

$${}^tRAR = R^{-1}AR = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \text{\Large 0} & \\ & \text{\Large 0} & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

なる n 次直交行列とする．このとき，

(1)　A が半正定値であるための必要十分条件は，$\lambda_1, \lambda_2, \cdots, \lambda_n \geq 0$ が成り立つことである．

(2)　一方，A が半負定値であるための必要十分条件は，$\lambda_1, \lambda_2, \cdots, \lambda_n \leq 0$ が成り立つことである．

この命題は，系 4.21 と全く同じ仕方で示すことが可能ですので，その証明は省略したいと思います．ともかく，この命題と系 4.21 を見比べると，ただちに次の結論を得ます．

系 4.38.
対称行列は，正則かつ半正定値であるとき，およびそのときに限り正定値である．また同様に，対称行列が負定値であるのは，それが正則かつ半負定値であるとき，およびそのときに限る．

定理 4.15 のふたつの主張 (1), (2) において逆が成立しないことは，すでに述べたとおりです．しかし，次の命題に述べるように，(1), (2) の逆に類することは成立します．

命題 4.39.
f を $\mathbb{R}^n$ の開集合 X 上で定義された C^2 級関数とする．

(1)　もし f が点 $\mathbf{x} \in X$ において極小であるならば，f のその点におけるヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ は半正定値である．

(2)　また，もし f が点 $\mathbf{x}$ において極大であるならば，f の $\mathbf{x}$ におけるヘッセ行列は半負定値である．

証明. (1), (2) は基本的に同じ仕方で証明できるので，ここでは (1) のみを証明する．そこで，f が点 $\mathbf{x} \in X$ において極小であると仮定しよう．一方，$\mathbf{v}$ を任意の n 次元行ベクトルとする．このとき，実数 t（ただし，$|t|$ は十分小）の関数 $g(t)$ を，$g(t) = f(\mathbf{x} + t\mathbf{v})$ により定義できる．この関数 $g(t)$ に対し，テイラーの定理を適用すると，

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\alpha t)t^2$$

（ただし，$0 \leq \alpha \leq 1$）を得る．ところで，$g(t)$ は $t = 0$ において極小であるから，$g'(0) = 0$，および $g(t) \geq g(0)$ が成り立つ．一方，補題 4.19 によれば，$g''(t) = \mathbf{v}\,(\nabla^2 f)(\mathbf{x} + t\mathbf{v})\,{}^t\mathbf{v}$ である．そして以上から，$g''(\alpha t) = \mathbf{v}\,(\nabla^2 f)(\mathbf{x} + \alpha t\mathbf{v})\,{}^t\mathbf{v} \geq 0$ が従う．ここで，$t \to 0$ とすると，2 階偏導関数の連続性から，$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \geq 0$ を得る．$\mathbf{v} \in \mathbb{R}^n$ は任意であったから，これによりヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ が半正定値であることが結論される． □

この命題から，定理 4.15 のとくに主張 (3) がただちに従います．実際にそれを示すために，定理 4.15 の関数 f のヘッセ行列が，f の臨界点 $\mathbf{x}$ において正則ではあるが，定値ではないと仮定しましょう．もし，仮に f が $\mathbf{x}$ において極小であるならば，命題 4.39 により，ヘッセ行列 $(\nabla^2 f)(\mathbf{x})$ は半正定値でなければなりません．しかも，仮定によれば $(\nabla^2 f)(\mathbf{x})$ は正則でした．したがって，系 4.38 から，$(\nabla^2 f)(\mathbf{x})$ が正定値であることが従いますが，これはもうひとつの仮定，すなわち，$(\nabla^2 f)(\mathbf{x})$ が点 $\mathbf{x}$ において定値でないという仮定に反します．よって，f は $\mathbf{x}$ において極小でないことが示されました．f が $\mathbf{x}$ において極大でないことも全く同じ仕方で示すことができます．以上が，定理 4.15 の主張 (3) の証明です．

まずは，定理 4.15 のふたつの主張 (1), (2) の証明を行います．その主張の要点となるのは，次の補題です．

補題 4.40.
f を $\mathbb{R}^n$ の開集合 X 上の C^2 級関数と，また $\mathbf{x} \in X$ をその臨界点とする．

(1)　f が $\mathbf{x}$ において極小でないならば，

$$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \leq 0$$

なる n 次元行ベクトル $\mathbf{v}$（ただし，$|\mathbf{v}| = 1$）が存在する．

(2)　また，もし f が $\mathbf{x}$ において極大でないならば，

$$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \geq 0$$

なる n 次元行ベクトル $\mathbf{v}$（ただし，$|\mathbf{v}| = 1$）が存在する．

まず，この補題を仮定して定理 4.15 のふたつの主張 (1), (2) を証明してしまいましょう．そこで，定理に現れる関数 f に対し，その臨界点 $\mathbf{x}$ におけるヘッセ行列が正定値であると仮定しましょう．この場合，証明すべきは，f が $\mathbf{x}$ において極小になることです．もし仮に f が $\mathbf{x}$ において極小でないとすると，補題 4.40 に

より，$\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \leq 0$，かつ $|\mathbf{v}| = 1$ を満たす $\mathbf{v} \in \mathbb{R}^n$ が存在することになり，ヘッセ行列の正定値性に矛盾します．したがって，f が $\mathbf{x}$ で極小でなければならないことが確認できた訳です．これが (1) の証明です．そして，もちろん，主張 (2) も全く同じ方法で証明可能です．

補題 4.40 の証明を済ませてしまいましょう．

補題 4.40 の証明. 関数 f が，その臨界点 $\mathbf{x} \in X$ において極小でないと仮定する．すると，任意の $r > 0$ に対し，$\mathbf{x}$ を中心とする半径 r の開球体 $B(\mathbf{x}, r)$ の点 $\mathbf{y}$ であって，$f(\mathbf{y}) < f(\mathbf{x})$ を満たすものが存在しなければならない．もちろん $\mathbf{y} \neq \mathbf{x}$ であるから，$|\mathbf{v}| = 1$ なる n 次元列ベクトル $\mathbf{v}$ を，$\mathbf{v} = (\mathbf{y} - \mathbf{x})/|\mathbf{y} - \mathbf{x}|$ で定義できる．さらに，実数 t（ただし，$|t|$ は十分小）の関数 $g(t)$ を，$g(t) = f(\mathbf{x} + t\mathbf{v})$ で定義する．この関数 $g(t)$ に対し，マクローリンの定理 3.22 を適用すると，

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\theta t)t^2$$

を得る．ただし，$0 \leq \theta \leq 1$ である．とくに，$t = |\mathbf{y} - \mathbf{x}|$ とおけば，点 $\mathbf{y}$ の選び方から $g(t) = f(\mathbf{y}) < f(\mathbf{x}) = g(0)$ が従う．さらに，点 $\mathbf{x}$ の臨界点であること，および公式 (4.4), (4.14) から，$g'(0) = 0$, $g''(\theta t) = \mathbf{v}\,(\nabla^2 f)(\mathbf{x} + \theta t\mathbf{v})\,{}^t\mathbf{v}$ が従う．よって，$\mathbf{v}\,(\nabla^2 f)(\mathbf{x} + \theta t\mathbf{v})\,{}^t\mathbf{v} < 0$ でなければならない．そして，ここで，$\mathbf{x} + \theta t\mathbf{v}$ を $\mathbf{z}$ とすれば，以下の主張を得る：任意の $r > 0$ に対し，$\mathbf{z} \in B(\mathbf{x}, r)$，および $|\mathbf{v}| = 1$ なる $\mathbf{v} \in \mathbb{R}^n$ を適当にとると，

$$\mathbf{v}\,(\nabla^2 f)(\mathbf{z})\,{}^t\mathbf{v} < 0$$

が成り立つ．

とくに，$r = 1/\ell$ $(\ell = 1, 2, \cdots)$ とすると，開球体 $B(\mathbf{x}, 1/\ell)$ の点 $\mathbf{z}_\ell$ からなる点列 $\{\mathbf{z}_\ell\}$，および $\mathbb{R}^n$ の部分集合 $Y = \{\mathbf{v} \in \mathbb{R}^n : |\mathbf{v}| = 1\}$ の点列 $\{\mathbf{v}_\ell\}$ で，条件

$$\mathbf{v}_\ell\,(\nabla^2 f)(\mathbf{z}_\ell)\,{}^t\mathbf{v}_\ell < 0 \tag{4.21}$$

を満たすものが存在することが従う．点列 $\{\mathbf{z}_\ell\}$ が点 $\mathbf{x}$ に収束することは明らかである．一方，$\mathbb{R}^n$ の部分集合 $Y = \{\mathbf{v} \in \mathbb{R}^n : |v| = 1\}$ は，有界かつ閉である．定理 4.35 により，Y がコンパクトであることが保証される．とくにその点列 $\{\mathbf{v}_\ell\}$ は Y の点 $\mathbf{v}$ に収束するような部分列 $\{\mathbf{v}_{k_\ell}\}$ を持つ．それに応じて，点列 $\{\mathbf{z}_\ell\}$ の部分列 $\{\mathbf{z}_{k_\ell}\}$ が得られるが，それがもとの点列 $\{\mathbf{z}_\ell\}$ と同じく点 $\mathbf{x}$ に収束することは明らかである．そこで，不等式 (4.21) において，$\mathbf{z}_\ell$ を $\mathbf{z}_{k_\ell}$ で，また $\mathbf{v}_\ell$ を $\mathbf{v}_{k_\ell}$ で置き換え，さらに $\ell \to \infty$ とすると，2 階偏導関数の連続性から $\mathbf{v}\,(\nabla^2 f)(\mathbf{x})\,{}^t\mathbf{v} \leq 0$ が従う．これで補題 4.40 の主張 (1) が証明された．主張 (2) の証明も同様であるので，ここでは省略することにする．　□

最後に残ったのが，定理 4.20 の証明です．もう一度定理を思い出しましょう．

定理 4.20

A が n 次対称行列であるとき，

$$R^{-1}AR = {}^tRAR = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \text{\huge 0} & \\ & \text{\huge 0} & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{4.22}$$

なる n 次直交行列 R が必ず存在する．

この定理の証明には，直交行列以外にも線形代数からの概念がいくつか必要になります．それを手短にまとめましょう．[10] n 次元ユークリッド空間 $\mathbb{R}^n$ の点を n 次元行ベクトルと見なします．ふたつのベクトル $\mathbf{x} = (x_1, \cdots, x_n), \mathbf{y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$ に対し，それらの**内積**が

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$$

により定義されます．[11] とくに，$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ですから，

$$|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

を考えることができます．これをベクトル $\mathbf{x}$ の**長さ**，あるいは**ノルム**と呼びます．[12] 一方，$\langle \mathbf{x}, \mathbf{y} \rangle = 0$ が成り立つとき，ふたつのベクトル $\mathbf{x}$, $\mathbf{y}$ は**垂直**であると言われます．[13] そして，n 個のベクトル $\mathbf{v}_1, \cdots, \mathbf{v}_n \in \mathbb{R}^n$ に対し，

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}, \qquad i, j = 1, \cdots, n$$

が成り立つとき，$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ を $\mathbb{R}^n$ の**正規直交基底**と呼びます．[14] ただし，δ_{ij} は**クロネッカーのデルタ記号**と呼ばれるもので，$i = j$ のときには値 1 を，それ以外の時には値ゼロをとります．したがって，$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ が正規直交基底であるとは，各ベクトル $\mathbf{v}_i$ は長さが 1 であり，$\mathbf{v}_i$, $\mathbf{v}_j$ $(i \neq j)$ は垂直であることです．いま，$\mathbf{v}_i \in \mathbb{R}^n$ $(i = 1, \cdots, n)$ に対し，そのおのおのを第 i 行ベクトルに持つ n 次正方行列を考えることができます．それを記号

$$\begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}$$

[10] これらのことがらは，冬学期に「数学 II」で学習するはずです．

[11] 内積を $\mathbf{x} \cdot \mathbf{y}$ と書く流儀もあります．

[12] 長さ $|\mathbf{x}|$ を $\|\mathbf{x}\|$ と書くこともあります．ちなみに，$n = 2$, ないし 3 のときには，これが本当にベクトルの長さを与えることは，皆さんご存知でしょう．

[13] このことも，2 次元および 3 次元においてはご存知のことと思います．

[14] 通常はまず「基底」を定義し，次いで「正規直交」基底を定義するのですが，ここでは時間を節約するためにいっきに「正規直交基底」を定義しています．

で表すことにします．容易に分かる通り，この正方行列が直交行列であるための必要十分条件は，$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ が正規直交基底であることです．そして，このことを用いると，定理 4.20 を以下のように書き直すことができます．

定理 4.41.
n 次対称行列 A に対し，

$$\mathbf{v}_i A\,{}^t\mathbf{v}_j = 0 \qquad (\text{ただし，} i, j = 1, \cdots, n,\ i \neq j) \tag{4.23}$$

となるような正規直交基底 $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ が存在する．

証明. 行列の次数 n に関する帰納法による．すると，まず考えるべきは $n = 1$ の場合であるが，このときにはなにも証明すべきことはない．そこで，いかなる n 次対称行列 A に対しても $\mathbb{R}^n$ の (4.23) を満たすような正規直交基底 $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ が存在すると仮定して，同様なことを $n+1$ 次対称行列に対して証明する．いま，$n+1$ 次対称行列を任意にとってきて B とする．一方，$n+1$ 次元ユークリッド空間 $\mathbb{R}^{n+1}$ の部分集合として，とくに

$$S^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| = 1\}$$

を考える．[15] これは $\mathbb{R}^{n+1}$ の有界閉集合である．[16] したがって，この集合 S^n がコンパクト集合であることが定理 4.35 から従う．一方，S^n 上の関数 $f : S^n \to \mathbb{R}$ を，

$$f(\mathbf{x}) = \mathbf{x}\,B\,{}^t\mathbf{x}, \qquad \mathbf{x} \in S^n$$

により定義する．これが連続であることも明らかである．しかも，その定義域 S^n はすでに述べたようにコンパクトであるから，最大値・最小値の定理 4.29 により，関数 f が最大値・最小値を有することが保証される．とくに，関数 f が点 $\mathbf{a} \in S^n$ において最大であるとしよう．このとき，

$$\begin{aligned} &\langle \mathbf{v}, \mathbf{a} \rangle = 0 \text{ を満たす任意の } \mathbf{v} \in \mathbb{R}^{n+1} \text{ に対し，} \\ &\qquad \mathbf{v}\,B\,{}^t\mathbf{a} = \mathbf{a}\,B\,{}^t\mathbf{v} = 0 \text{ が成り立つ} \end{aligned} \tag{4.24}$$

ことを示したい．

仮定 $\langle \mathbf{v}, \mathbf{a} \rangle = 0$ から，$\mathbf{a}$ と $\mathbf{v}$ は 1 次独立，あるいは $\mathbf{v} = \mathbf{0}$ でなければならないことが従う．いずれにしても，$\mathbf{a} + t\mathbf{v}$ $(t \in \mathbb{R})$ はゼロベクトルにはなり得ないから，

[15] この集合 S^n は，$n = 1$ のとき平面内の原点を中心とする半径 1 の円周，また $n = 2$ のとき 3 次元ユークリッド空間内の原点を中心とした半径 1 の球面となります．

[16] なぜそうなのかは，各自で考えてみて下さい．

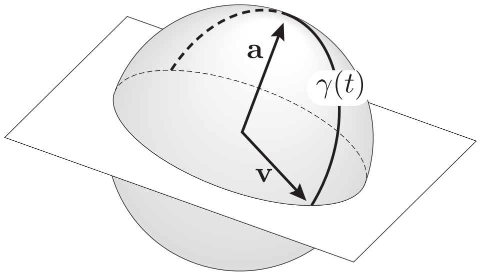


図 4.13:

この $\mathbf{v}$ に応じて S^n 内の曲線 $\gamma(t)$ を，

$$\gamma(t) = \frac{\mathbf{a} + t\mathbf{v}}{|\mathbf{a} + t\mathbf{v}|} \qquad (t \in \mathbb{R})$$

により定義できる．さらに，実数 t の関数 $g(t)$ を，

$$g(t) = f(\gamma(t)) = \frac{(\mathbf{a} + t\mathbf{v})B\,{}^t(\mathbf{a} + t\mathbf{v})}{|\mathbf{a} + t\mathbf{v}|^2}$$

により定義する．この関数が微分可能であることも明らかであろう．しかも，関数 f が点 $\mathbf{a} = g(0)$ で最大であることから，関数 $g(t)$ は $t = 0$ で最大であることが従う．よって，$g'(0) = 0$ が成り立たなければならない．ところが，

$$g'(t) = \frac{\mathbf{v}B\,{}^t(\mathbf{a} + t\mathbf{v}) + (\mathbf{a} + t\mathbf{v})B\,{}^t\mathbf{v}}{|\mathbf{a} + t\mathbf{v}|^2} - 2\,\frac{\langle \mathbf{v}, \mathbf{a} + t\mathbf{v}\rangle \cdot (\mathbf{a} + t\mathbf{v})B\,{}^t(\mathbf{a} + t\mathbf{v})}{|\mathbf{a} + t\mathbf{v}|^4}$$

である．さらに，仮定 $|\mathbf{a}| = 1$，および $\langle \mathbf{v}, \mathbf{a}\rangle = 0$ とから，結局，$0 = g'(0) = \mathbf{v}B\,{}^t\mathbf{a} + \mathbf{a}B\,{}^t\mathbf{v}$ が従う．一方，B は対称，すなわち，${}^tB = B$ を満たすから，$\mathbf{v}B\,{}^t\mathbf{a} = \mathbf{a}B\,{}^t\mathbf{v}$ が成り立つ．以上で (4.24) の証明が完了した．

次に，$n+1$ 次対称行列 B に対しても

$$\mathbf{v}_i B\,{}^t\mathbf{v}_j = 0 \quad （ただし，i, j = 1, \cdots, n+1,\ i \neq j） \tag{4.25}$$

を満たす $\mathbb{R}^{n+1}$ の正規直交基底 $\{\mathbf{v}_1, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$ が存在することを証明しよう．そのために，まず，$\mathbb{R}^{n+1}$ の正規直交基底 $\{\mathbf{v}_1, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$ で，とくに $\mathbf{v}_{n+1} = \mathbf{a}$ を満たすものをとる．[17] すると，(4.24) より，

$$(\mathbf{v}_i B\,{}^t\mathbf{v}_j) = \begin{pmatrix} * & \cdots & * & 0 \\ \vdots & & \vdots & \vdots \\ * & \cdots & * & 0 \\ 0 & \cdots & 0 & * \end{pmatrix} \tag{4.26}$$

[17] このような正規直交基底が存在することは，直感的には受け入れやすいことと思います．しかし，厳密にはやはり証明をせねばならないことです．証明には**グラム=シュミットの直交化法**と呼ばれる方法が用いられます．詳しいことは，線形代数の講義や教科書に任せることにしましょう．

となることが分かる．右辺の左上に現れる n 次正方行列 — それを A としよう — は対称行列であるから，帰納法の仮定を適用することができる．すると，$\{\mathbf{v}_1, \cdots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$ を $\mathbb{R}^{n+1}$ の正規直交基底としたまま，$\mathbf{v}_1, \cdots, \mathbf{v}_n$ を適当にとりかえることにより，(4.26) の右辺の行列の左上の部分 A を対角行列にすることができること，すなわち (4.25) が従う．以上で，定理の証明が終了した．□

B　章末問題

問題 4.42. ————————————————————————

次の関数の 2 階偏微分係数 $f_{xy}(0,0)$ および $f_{yx}(0,0)$ を計算し，$f_{xy}(0,0) \neq f_{yx}(0,0)$ であることを確かめよ：

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

————————————————————————————————

解答. $f(x,y) + f(y,x) = 0$ であるから，$f_{xy}(0,0) + f_{yx}(0,0) = 0$ である．以下において $f_{xy}(0,0)$ を計算する．まず，

$$f_{xy}(0,0) = \lim_{y \to 0} \frac{f_x(0,y) - f_x(0,0)}{y}$$

であることに注意しよう．さらに，$y \neq 0$ のとき

$$f_x(0,y) = -y, \qquad f_x(0,0) = \lim_{x \to 0} \frac{f(x,0) - f(0,0)}{x} = 0$$

である．これより，$f_{xy}(0,0) = -1$，$f_{yx}(0,0) = 1$ が従う．とくに，$f_{xy}(0,0) \neq f_{yx}(0,0)$ である．□

問題 4.43. ————————————————————————

次の関数の極値を求めよ．

(1)　$f(x,y) = xy(x^2+y^2-1)$　　　　(2)　$f(x,y) = xy \cdot e^{-\frac{1}{2}(x^2+y^2)}$

————————————————————————————————

解答. (1)（第 1 段）

$$f_x = 3x^2y + y^3 - y, \qquad f_y = x^3 + 3xy^2 - x$$

であるから，臨界点を求めるためには連立方程式

$$\begin{cases} 3x^2y + y^3 - y &= 0, \\ x^3 + 3xy^2 - x &= 0 \end{cases}$$

を解けばよい．これら2式の両辺の和・差をとることにより，

$$\begin{cases} (x+y)(x^2+2xy+y^2-1) &= 0, \\ (x-y)(x^2-2xy+y^2-1) &= 0 \end{cases}$$

を得る．(i) $x+y=x-y=0$ のときには，$(x,y)=(0,0)$ が従う．(ii) $x+y=x^2-2xy+y^2-1=0$ のときには，$(x,y)=(\pm 1/2, \mp 1/2)$（複号同順）を，また (iii) $x-y=x^2+2xy+y^2-1=0$ のときには，$(x,y)=(\pm 1/2, \pm 1/2)$（複号同順）を得る．(iv) $x^2+2xy+y^2-1=x^2-2xy+y^2-1=0$ のときには，$(x,y)=(\pm 1, 0), (0, \pm 1)$ が従う．以上をまとめて，

$$(0,0), \quad (\pm 1, 0), \quad (0, \pm 1), \quad (\pm 1/2, \pm 1/2)\ (\text{複号任意}),$$

これらが，関数 $f(x,y)$ の臨界点である．

（第2段）

$$\nabla^2 f = \begin{pmatrix} 6xy & 3x^2+3y^2-1 \\ 3x^2+3y^2-1 & 6xy \end{pmatrix}$$

である．臨界点 $(0,0), (0,\pm 1), (\pm 1, 0)$ においては，$\det \nabla^2 f < 0$ であるから，これらは鞍点であり，f はこれらの点においては極値をとらない．一方，点 $(\pm 1/2, \pm 1/2)$（複号任意）においては，$\det \nabla^2 f > 0$ である．しかも，点 $(\pm 1/2, \pm 1/2)$（複号同順）においては $f_{xx} > 0$ であるから f はこれらの点において極大である．また，点 $(\pm 1/2, \mp 1/2)$（複号同順）においては $f_{xx} < 0$ であるから f はこれらの点において極大である．

（答） 点 $(\pm 1/2, \pm 1/2)$（複号同順）において極小で極小値は $-1/8$，
点 $(\pm 1/2, \mp 1/2)$（複号同順）において極大で極大値は $1/8$．

(2)（第1段）まず，1階導関数を計算すると，

$$f_x = y(1-x^2)e^{-\frac{1}{2}(x^2+y^2)}, \qquad f_y = x(1-y^2)e^{-\frac{1}{2}(x^2+y^2)}.$$

したがって，関数 $f(x,y)$ の臨界点は，

$$(0,0), \qquad (\pm 1, \pm 1)\ (\text{複号任意})$$

である．

（第 2 段）関数 $f(x,y)$ のヘッセ行列は以下で与えられる：

$$\nabla^2 f = \begin{pmatrix} -xy(3-x^2)e^{-\frac{1}{2}(x^2+y^2)} & (1-x^2)(1-y^2)e^{-\frac{1}{2}(x^2+y^2)} \\ \\ (1-x^2)(1-y^2)e^{-\frac{1}{2}(x^2+y^2)} & -xy(3-y^2)e^{-\frac{1}{2}(x^2+y^2)} \end{pmatrix}.$$

とくに原点 $(0,0)$ においては，$\nabla^2 f = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ である．この対称行列は正則ではあるが定値ではない．よって，関数 $f(x,y)$ は原点においては定値をとらない．一方，点 $(\pm1,\pm1)$（複号同順）においては，$\nabla^2 f = \begin{pmatrix} -2/e & 0 \\ 0 & -2/e \end{pmatrix}$ である．これは負定値であり，関数 $f(x,y)$ はこれらの点において極大値 $1/e$ をとる．そして，点 $(\pm1,\mp1)$（複号同順）においては $\nabla^2 f = \begin{pmatrix} 2/e & 0 \\ 0 & 2/e \end{pmatrix}$ である．この行列は正定値であり，関数 $f(x,y)$ はこれらの点において極小値 $-1/e$ をとる．

（答） 点 $(\pm1,\pm1)$（複号同順）において極大で，極小値は $1/e$,
点 $(\pm1,\mp1)$（複号同順）において極小で，極小値は $-1/e$.

□

問題 4.44. ————————————————————————————

次の関数の極値を決定せよ．

$$f(x,y,z) = xy + yz + zx - xyz.$$

————————————————————————————————————

解答. （第 1 段）まず臨界点を求めたい．そのために，関数 $f(x,y,z)$ の偏導関数を計算すると，

$$f_x = y + z - yz, \qquad f_y = x + z - xz, \qquad f_z = x + y - xy$$

を得る．したがって，臨界点を決定するためには連立方程式

$$\left.\begin{array}{rcl} y + z - yz &=& 0, \\ x \quad + z - xz &=& 0, \\ x + y \quad - xy &=& 0 \end{array}\right\} \tag{*}$$

を解けばよい．この連立方程式において，(第 2 式) − (第 1 式), および (第 3 式) − (第 2 式) を計算すると，

$$(x-y)(1-z) = 0 \quad \text{かつ} \quad (y-z)(1-x) = 0$$

を得る．したがって，以下の 4 つの場合が考えられる．

(i) $x-y=y-z=0$ の場合．このときには，$x=y=z$ である．とくに $x=y$ と (*) の第3式から，$x=0,2$ を得る．そして，$x=0$ の場合には $(x,y,z)=(0,0,0)$ が，$x=2$ のときには $x=y=z=2$ が従う．

(ii) $x-y=1-x=0$ の場合．このときには，$x=y=1$ である．ところが，すでに (i) で見たとおり，$x=y$ が成り立つためには，$x=0,2$ でなければならない．したがって，この場合には解が存在しないことが分かる．

(iii) $y-z=1-z=0$ の場合．この場合にも解は存在しない．

(iv) $1-z=1-x=0$ の場合．このときにも解は存在しない．

以上から，関数 $f(x,y,z)$ の臨界点は

$$(0,0,0), \qquad (2,2,2)$$

であることが分かる．

（第2段）関数 $f(x,y,z)$ のヘッセ行列を計算すると，

$$\nabla^2 f = \begin{pmatrix} 0 & 1-z & 1-y \\ 1-z & 0 & 1-x \\ 1-y & 1-x & 0 \end{pmatrix}.$$

原点 $(0,0,0)$ における $\nabla^2 f$ の固有多項式は

$$\det\begin{pmatrix} \xi & -1 & -1 \\ -1 & \xi & -1 \\ -1 & -1 & \xi \end{pmatrix} = \xi^3 - 3\xi - 2$$

である．一方，臨界点 $(2,2,2)$ におけるヘッセ行列の固有多項式は，

$$\det\begin{pmatrix} \xi & 1 & 1 \\ 1 & \xi & 1 \\ 1 & 1 & \xi \end{pmatrix} = \xi^3 - 3\xi + 2$$

である．したがって，これらの臨界点におけるヘッセ行列 $\nabla^2 f$ は正則ではあるが定値ではない．よって，関数 $f(x,y,z)$ はこれらの臨界点において極値を持たないことが従う．

（答）関数 $f(x,y,z)$ は極値を持たない．

□

問題 4.45. ————————————————————

次の関数が原点において極値をとるかどうか判定せよ．ただし，(1) においては a, b は $a^2+b^2 \neq 1$ なる定数，また (2) においては a, b, c, p, q, r は $ac-b^2 \neq 0$, $pr-q^2 \neq 0$ なる定数とする．

(1)　$f(x,y,z)=\dfrac{1}{2}(x^2+y^2+z^2)+axy+byz+3xyz.$

(2)　$f(x,y,z,w)=ax^2+2bxy+cy^2+pz^2+2qzw+rw^2.$

解答. (1) 関数 $f(x,y,z)$ の勾配ベクトルは原点において値ゼロをとる．したがって，原点は確かに臨界点である．原点における関数 $f(x,y,z)$ のヘッセ行列は

$$\begin{pmatrix} 1 & a & 0 \\ a & 1 & b \\ 0 & b & 1 \end{pmatrix}$$

であり，その固有多項式は以下で与えられる．ただし，$\sigma=a^2+b^2$ とした．

$$\begin{aligned}\det\begin{pmatrix} \xi-1 & -a & 0 \\ -a & \xi-1 & -b \\ 0 & -b & \xi-1 \end{pmatrix} &= (\xi-1)^3-(a^2+b^2)(\xi-1) \\ &= \xi^3-3\xi^2+(3-\sigma)\xi+(\sigma-1).\end{aligned}$$

仮定により，定数項はゼロにはならない．よって，ヘッセ行列は正則である．また，$0\le\sigma<1$ のとき，ヘッセ行列は正定値であるが，$\sigma>1$ のときには定値ではない．以上をまとめると，

（答） $a^2+b^2<1$ のときには極小値をとる．それ以外の場合には極値をとらない．

□

問題 4.46. ―――――――――――――――――――――――――――――――

関数

$$f(x,y)=x^4+y^4-4xy+1 \tag{4.27}$$

を考える．[18] 以下，r を正の定数とする．

(1) この関数の円周 $C_r=\{(x,y)\in\mathbb{R}^2 : x^2+y^2=r^2\}$ における最大値・最小値を求めよ．

(2) この関数の円板 $D_r=\{(x,y)\in\mathbb{R}^2 : x^2+y^2\le r^2\}$ における最大値・最小値を決定せよ．

---

[18] この関数の極値はすでに決定済みである．実際，臨界点は，$(x,y)=(0,0)$, $(\pm1,\pm1)$（複号同順）である．原点 $(0,0)$ は鞍点であり，そこでは極値をとらない．一方，臨界点 $(\pm1,\pm1)$（複号同順）において極小，極小値は -1 であった．

解答. (1) $x = r\cos\theta$, $y = r\sin\theta$ を $f(x, y)$ に代入することにより，

$$f(\theta) = f(r\cos\theta, r\sin\theta) = -\frac{1}{2}r^4\sin^2 2\theta - 2r^2\sin 2\theta + r^4 + 1$$

を得る．この関数の導関数は

$$f'(\theta) = -2r^2\cos 2\theta(r^2\sin 2\theta + 2)$$

である．したがって，$r < \sqrt{2}$ のとき，関数 $f(\theta)$ の臨界点は，

$$\frac{\pi}{4},\quad \frac{3\pi}{4},\quad \frac{5\pi}{4},\quad \frac{7\pi}{4}$$

である．一方，$r > \sqrt{2}$ のときの臨界点は，小さい方から順に

$$\frac{\pi}{4},\quad \frac{\pi}{2} + \alpha_0,\quad \frac{3\pi}{4},\quad \pi - \alpha_0,\quad \frac{5\pi}{4},\quad \frac{3\pi}{2} + \alpha_0,\quad \frac{7\pi}{4},\quad 2\pi - \alpha_0$$

（ただし，$\sin 2\alpha_0 = -2/r^2$, $0 < \alpha_0 < \pi/4$）である．残る $r = \sqrt{2}$ の場合には，$r > \sqrt{2}$ の場合に新たに生じた臨界点たちが $r < \sqrt{2}$ の場合のもともとあった臨界点 $3\pi/4$, $7\pi/4$ の一方と一致することが分かる．したがって，その場合の臨界点は $r < \sqrt{2}$ の場合のそれらと一致する．

次に関数 $f(\theta)$ の 2 階導関数を計算しよう：

$$f''(\theta) = 8r^4\sin^2 2\theta + 8r^2\sin 2\theta - 4r^4.$$

$r < \sqrt{2}$ のときには，

$$f''(\pi/4) = f''(5\pi/4) = 4r^2(r^2 + 2) > 0,$$
$$f''(3\pi/4) = f''(7\pi/4) = 4r^2(r^2 - 2) < 0$$

であるから，関数 $f(\theta)$ は $\theta = \pi/4$, $5\pi/4$ において極小，$\theta = 3\pi/4$, $7\pi/4$ において極大であることが分かる．次に，$r = \sqrt{2}$ の場合を考えよう．この場合にも $\theta = \pi/4$, $5\pi/4$ における 2 階微分係数 $f''(\theta)$ は正であるから，これらの点において関数 $f(\theta)$ は極小である．ところが，残るふたつの臨界点 $\theta = 3\pi/4$, $7\pi/4$ における 2 階微分係数はゼロとなる．したがって，それを見ただけで極大・極小性の判定は出来ない．しかし，$f(\theta) = -2(\sin 2\theta + 1)^2 + 7$, および $\theta = 3\pi/4$, $7\pi/4$ において $\sin 2\theta = -1$ であることに注意すれば，それらの点において関数 $f(\theta)$ は最大であることが分かる．最後に $r > \sqrt{2}$ の場合である．このときには，臨界点 $\pi/4$, $3\pi/4$, $5\pi/4$, $7\pi/4$ における 2 階微分係数の値は $4r^2(r^2 + 2)$ ないし $4r^2(r^2 - 2)$ であるからいずれにしても正，残り 4 つの臨界点における値は $4(4 - r^4) < 0$ である．よって，前者に

おいては極小，後者においては極大である．以上をまとめると次のようになる：

$$
\begin{aligned}
r \le \sqrt{2}\ \text{のとき，}\quad & \theta = \frac{\pi}{4}, \frac{5\pi}{4}\ \text{において極小，極小値は}\ \frac{1}{2}r^4 - 2r^2 + 1, \\
& \theta = \frac{3\pi}{4}, \frac{7\pi}{4}\ \text{において極大，極大値は}\ \frac{1}{2}r^4 + 2r^2 + 1; \\
r > \sqrt{2}\ \text{のとき，}\quad & \theta = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\ \text{において極小であり，} \\
& \qquad f\left(\frac{\pi}{4}\right) = f\left(\frac{5\pi}{4}\right) = \frac{1}{2}r^4 - 2r^2 + 1, \\
& \qquad f\left(\frac{3\pi}{4}\right) = f\left(\frac{7\pi}{4}\right) = \frac{1}{2}r^4 + 2r^2 + 1; \\
& \theta = \frac{\pi}{2} + \alpha_0, \pi - \alpha_0, \frac{3\pi}{2} + \alpha_0, 2\pi - \alpha_0\ \text{において極大，} \\
& \qquad\qquad \text{極大値は}\ r^4 + 3.
\end{aligned}
$$

最大値・最小値を取り出して，さらに変数 θ ではなくもとの (x, y) 座標を用いて記述すれば，以下のようになる．

（答）　$r \le \sqrt{2}$ のとき $(x, y) = (\pm r/\sqrt{2}, \mp r/\sqrt{2})$（複号同順）において最大で最大値は $\frac{1}{2}r^4 + 2r^2 + 1$,
$r > \sqrt{2}$ のとき $(x, y) = (r\cos\beta_0, r\sin\beta_0)$
（ただし，$\sin 2\beta_0 = -2/r^2$）において最大で最大値は $r^4 + 3$.
一方，r の値に依らず $(x, y) = (\pm r/\sqrt{2}, \pm r/\sqrt{2})$（複号同順）において最小，最小値は $\frac{1}{2}r^4 - 2r^2 + 1$.

(2)

すでに見たとおり，もとの関数 $f(x, y)$ は極大値を持たない．したがって，関数 $f(x, y)$ の円板 $x^2 + y^2 \le r^2$ への制限が最大になるのは常にその境界上の点においてである．その点は関数 $f(x, y)$ の円周への制限が最大値をとる点でもある．したがって，(1) によれば，関数 $f(x, y)$ の円板 $x^2 + y^2 \le r^2$ への制限が最大になるのは，$r < \sqrt{2}$ のときには $(x, y) = (\pm r/\sqrt{2}, \mp r/\sqrt{2})$（複号同順）においてでありその点における関数 $f(x, y)$ の値は $\frac{1}{2}r^4 + 2r^2 + 1$，また $r \ge 2$ のときには点 $(x, y) = (r\cos\beta_0, r\sin\beta_0)$（ただし，$\sin 2\beta_0 = -2/r^2$）でありその点における値は $r^4 + 3$ である．次に最小値について考えよう．$r < \sqrt{2}$ であるときには，もとの関数 $f(x, y)$ が極小値をとる点 $(\pm 1, \pm 1)$ を円板 $x^2 + y^2 \le r^2$ は含まないから，やはり境界上で最小になる．したがって，再び (1) より，関数 $f(x, y)$ の円板 $x^2 + y^2 \le r^2$ への制限が最小になるのは，境界上の点 $(x, y) = (\pm r/\sqrt{2}, \pm r/\sqrt{2})$（複号同順）においてであり，その点における関数 $f(x, y)$ の値は $\frac{1}{2}r^4 - 2r^2 + 1$ であることが従う．一方，$r \ge \sqrt{2}$ のときにはもとの関数 $f(x, y)$ が極小値をとる点 $(\pm 1, \pm 1)$ は

円板 $x^2+y^2 \leq r^2$ に含まれる．境界に制限したときの最小値 $\frac{1}{2}r^4-2r^2+1$ と点 $(\pm1,\pm1)$ における値 -1 を比較すると，$\frac{1}{2}r^4-2r^2+1 \geq -1$．したがって，関数 $f(x,y)$ の円板 $x^2+y^2 \leq r^2$ への制限が最小になるのは，点 $(\pm1,\pm1)$（複号同順）であり，その点における値は -1 である．　□

設問 (1) の解答に登場した関数 $f(\theta)$ のグラフを図 4.14 にあげておきました．

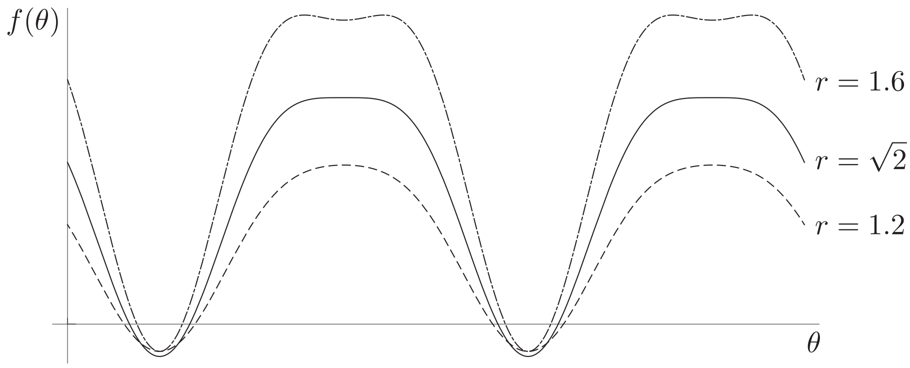


図 4.14:

問題 4.47.

以下，A を 2 次対称行列とする．

(1) A がふたつの実固有値 λ, μ を有することを示せ．

(2) もし $\lambda \neq \mu$ ならば，λ の固有ベクトルと μ のそれとは垂直であることを示せ．

(3) ${}^tRAR = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ なる 2 次直交行列 R が存在することを示せ．

(4) 3 次対称行列に対しても，同様なことを議論せよ．

問題 4.48.

n 次元ユークリッド空間 $\mathbb{R}^n$ からそれ自身の中への写像 f が，任意の $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ に対し

$$|f(\mathbf{y})-f(\mathbf{x})| = |\mathbf{y}-\mathbf{x}|$$

を満たすとする．このとき，f を

$$f(\mathbf{x}) = R\mathbf{x}+\mathbf{a}$$

（ただし，R は n 次直交行列，また $\mathbf{a} \in \mathbb{R}^n$）と書けることを証明せよ．

問題 4.49. ――――――――――――――――――――――――――――

定理 4.29 を証明せよ．

――――――――――――――――――――――――――――

解答. 一般に，任意の関数 f に対し，その最小値は，$g(\mathbf{x}) = -f(\mathbf{x})$ で定義される関数 g の最大値に他ならない．したがって，最大値の存在のみを示せば十分である．そこで，以下では，$\mathbb{R}^n$ のコンパクト集合 X 上で定義された連続関数 $f : X \to \mathbb{R}$ に対し，その最大値の存在を証明する．

（第 1 段） まず，準備として，次の主張を証明する：適当に定数 b をとれば，

$$\text{すべての } \mathbf{x} \in X \text{ に対し，} \qquad f(\mathbf{x}) \leq b \tag{4.28}$$

が成り立つ．証明は背理法による．そのために，(4.28) を満たすような定数 b が存在しない，すなわち，いかなる定数 b に対しても，$f(\mathbf{x}) > b$ なる点 $\mathbf{x} \in X$ が存在する，と仮定する．すると，とくに各 $\ell = 0, 1, \cdots$ に対し，$f(\mathbf{x}_\ell) > \ell$ なる点 $\mathbf{x}_\ell \in X$ が存在しなければならない．こうして，X の点列 $\{\mathbf{x}_\ell\}$ を得る．ところで，仮定によれば，X はコンパクトであったから，この点列 $\{\mathbf{x}_\ell\}$ は収束部分列を持たねばならない．$\{\mathbf{y}_\ell\}$ を，$\{\mathbf{x}_\ell\}$ の部分列であって，しかも $\mathbf{y} \in X$ に収束するものとする．すると，関数 f の連続性から，$f(\mathbf{y}_\ell) \to f(\mathbf{y})$ $(\ell \to \infty)$ を得る．とくに，$\ell \to \infty$ のとき，$f(\mathbf{y}_\ell)$ は（有限の値に）収束することに注意されたい．ところが，一方，点列 $\{\mathbf{x}_\ell\}$ の取り方によれば，$f(\mathbf{x}_\ell) \to \infty$ $(\ell \to \infty)$，したがってその部分列 $\{\mathbf{y}_\ell\}$ に対しても，$f(\mathbf{y}_\ell)$ は $\ell \to \infty$ のとき発散しなければならない．すなわち，矛盾が生じた．よって，(4.28) を満たす定数 b の存在が保証される．

（第 2 段）区間縮小法をとくに関数 f の値域に対し適用することにより，f が最大値を有することを証明する．そのために，X の点列 $\{\mathbf{x}_\ell\}$，およびふたつの数列 $\{a_\ell\}, \{b_\ell\}$ を，以下のように帰納的に定義する．まず，X から勝手な点をとってきて，$\mathbf{x}_0$ とする．また，a_0 を，$a_0 = f(\mathbf{x}_0)$ により定義する．一方，b_0 を，(4.28) を満たすような定数 b とする．もちろん，$a_0 \leq b_0$ が成り立つ．つぎに，$\mathbf{x}_\ell, a_\ell, b_\ell$ がすでに定義できたと仮定しよう．このとき，$\mathbf{x}_{\ell+1}, a_{\ell+1}, b_{\ell+1}$ を，以下に述べる規則に従い定義する：

(i)　もし

$$f(\mathbf{x}) \geq \frac{a_\ell + b_\ell}{2}$$

を満たす点 $\mathbf{x} \in X$ が存在するならば，その $\mathbf{x}$ を $\mathbf{x}_{\ell+1}$ と，また，

$$a_{\ell+1} = \frac{a_\ell + b_\ell}{2}, \qquad b_{\ell+1} = b_\ell$$

とする；

(ii)　それ以外の場合には,

$$\mathbf{x}_{\ell+1} = \mathbf{x}_\ell, \qquad a_{\ell+1} = a_\ell, \qquad b_{\ell+1} = \frac{a_\ell + b_\ell}{2}$$

とする.

このようにして定義された点列 $\{\mathbf{x}_\ell\}$，および数列 $\{a_\ell\}$, $\{b_\ell\}$ が，以下に挙げる性質を持っていることを確かめるのは容易である：

(1)　数列 $\{a_\ell\}$ は，上に有界，かつ単調増加である；

(2)　数列 $\{b_\ell\}$ は，下に有界，かつ単調減少である；

(3)　$b_\ell - a_\ell = 2^{-\ell}(b_0 - a_0)$;

(4)　$a_\ell \leq f(\mathbf{x}_\ell) \qquad (\ell = 0, 1, \cdots)$;

(5)　いかなる $\mathbf{x} \in X$，および $\ell = 0, 1, \cdots$ に対しても，$f(\mathbf{x}) \leq b_\ell$ が成り立つ.

まず，性質 (1)，(2)，および有界単調数列の収束性定理（定理 1.5）により，数列 $\{a_\ell\}$, $\{b_\ell\}$ がともに収束することが保証される．さらに，性質 (3) から，これらふたつの数列の極限は一致しなければならないことが従う．その共通の極限を c としよう：

$$\lim_{\ell\to\infty} a_\ell = \lim_{\ell\to\infty} b_\ell = c.$$

すると，性質 (5) から，

(6)　$f(\mathbf{x}) \leq c$ が，すべての $\mathbf{x} \in X$ に対して成り立つ

ことが判る．あとは $f(\mathbf{y}) = c$ なる $\mathbf{y} \in X$ の存在さえ示すことができれば，c が f の最大値であることになり，証明は完了する．さてここで，X の点列 $\{\mathbf{x}_\ell\}$ に注目する．X のコンパクト性より，それは収束部分列を有する．そこで，$\{\mathbf{y}_\ell\}$ を，$\{\mathbf{x}_\ell\}$ の部分列であって，しかもある点 $\mathbf{y} \in X$ に収束するようなものとする．このとき，f の連続性から，$f(\mathbf{y}_\ell) \to f(\mathbf{y})$ $(\ell \to \infty)$ が成り立たねばならない．ところが，性質 (4)，および $a_\ell \to c$ $(\ell \to \infty)$ より，$f(\mathbf{y}) \geq c$ が従う．これと (6) とから，$f(\mathbf{y}) = c$，すなわち，関数 f が c を値としてとることが判った．以上で定理 4.29 の証明が終わった． □