

微積分学テキスト構成案：高校数学からの飛躍

2026 年 2 月 12 日

まえがき

「計算」から「解析」へ、そして「応用」へ

本書は、高校数学（数 III）を履修し終え、これから大学レベルの数学、あるいはデータサイエンスや工学の世界へ踏み出そうとしている読者を対象に執筆されました。

現代社会において、微積分は単なる「計算のルール」ではありません。それは、変化する世界を記述し、未来を予測し、最適解を導き出すための「言語」です。AI（人工知能）が画像を見分けるとき、その裏側では数億次元の空間で「微分（勾配降下法）」が行われています。金融市場のリスクを計算するとき、そこには「ガウス積分（正規分布）」が使われています。ロボットが滑らかに動くとき、その制御には「ベクトル解析（ヤコビアン）」が不可欠です。

しかし、多くの学習者は、高校数学の「直感的な計算」と、大学数学の「厳密な論理（ $\epsilon - \delta$ 論法など）」との間の深い溝に直面し、数学の本来の面白さに触れる前に挫折してしまうのが現実です。

本書の目的は、この溝に橋を架けることです。そのために、以下の 3 つの視点を大切にしました。

1. 直感と論理の両立: 「なぜその定義が必要なのか？」という動機（工学的な誤差制御など）を先に示し、その後に厳密な証明を行うスタイルをとりました。
2. 多次元への拡張: 現代の課題は多変数です。1 変数の微積分を丁寧に解説した後、スムーズに多変数（偏微分・重積分）へ接続し、高次元空間のイメージを養います。
3. 計算機との接続: 解析学はコンピュータシミュレーションの基礎です。巻末には Python を用いた可視化コードを掲載し、数式がプログラムとして動く様子を体験できるようにしました。

本書を読み終えたとき、無機質に見えていた数式が、実は世界の動きを鮮やかに記述する「生きた言葉」であることを実感していただけるはずです。この本が、皆様の理系ライフにおける頼もしい羅針盤となることを願っています。

2026 年 2 月 著者

目次

まえがき	i
第Ⅰ部 極限と連続性の厳密化（直感から論理へ）	1
第1章 実数の連続性と数列：無限を有限で制御する	3
1.1 実数の連続性と上限・下限	3
1.2 数列の収束 ($\epsilon - N$ 論法)	4
1.3 収束の一意性	5
1.4 第1章のまとめと演習	6
第2章 関数の極限と連続：入力と出力の誤差制御	7
2.1 関数の極限 ($\epsilon - \delta$ 論法)	7
2.2 関数の連続性	8
2.3 中間値の定理 (The Intermediate Value Theorem)	9
2.4 最大値・最小値の定理	10
2.5 一様連続性～より強い連続性～	11
2.6 第2章のまとめ	12
第Ⅱ部 1変数関数の深化（線形近似から高次近似へ）	13
第3章 微分法の基礎と平均値の定理：最強の近似ツール	15
3.1 微分可能性と線形近似	15
3.2 ロルの定理 (Rolle's Theorem)	15
3.3 平均値の定理 (Mean Value Theorem)	17
3.4 コーシーの平均値の定理とロピタルの定理	18
3.5 第3章のまとめ	19
第4章 テイラー展開と級数：関数を多項式に分解する	21
4.1 高次導関数と多項式近似	21
4.2 テイラーの定理 (Taylor's Theorem)	21
4.3 マクローリン展開	23
4.4 ランダウの記号 (Big-O, Small-o) の拡張版	25
4.5 無限級数の収束判定	26
4.6 第4章のまとめ	27

第 III 部 多変数関数の微分（高次元世界への入り口）	29
第 5 章 多変数関数と偏微分：多次元空間の傾き	31
5.1 n 次元ユークリッド空間と位相	31
5.2 多変数関数の極限と連続	31
5.3 偏微分（Partial Differentiation）	32
5.4 全微分（Total Differential）～真の微分可能性～	32
5.5 合成関数の微分法（連鎖律：Chain Rule）	34
5.6 第 5 章のまとめ	35
第 6 章 多変数微分の応用：最適解を求めて	37
6.1 多変数関数の極値判定	37
6.2 ラグランジュの未定乗数法（Lagrange Multipliers）	38
6.3 【発展】不等式制約と KKT 条件～壁に寄りかかるか、部屋の中にいるか～	40
6.4 第 6 章のまとめ	42
第 IV 部 多変数関数の積分（重積分）	43
第 7 章 重積分：体積と総和の計算	45
7.1 重積分の定義	45
7.2 累次積分とフビニの定理	45
7.3 積分領域の変更と順序交換	46
7.4 第 7 章のまとめ	47
第 8 章 変数変換とヤコビアン：歪んだ空間の測り方	49
8.1 変数変換の一般論とヤコビアン	49
8.2 極座標変換（Polar Coordinates）	50
8.3 ガウス積分（Gaussian Integral）～統計学の至宝～	51
8.4 一般の変数変換と応用	52
8.5 第 8 章のまとめ	52
第 V 部 ベクトル解析の入り口（オプション）	53
第 9 章 ベクトル解析：場の流れを読む	55
9.1 スカラー場とベクトル場	55
9.2 ナブラ演算子 ∇ と 3 つの微分	55
9.3 線積分と面積分～流れに乗る・流れを測る～	56
9.4 平面での積分定理：グリーンの定理	57
9.5 第 9 章のまとめ	59
第 10 章 積分定理の統合と展望	61
10.1 ガウスの発散定理（Gauss's Divergence Theorem）	62
10.2 ストークスの定理（Stokes' Theorem）	63
10.3 積分定理の統一：微分形式へ	63
10.4 一般化されたストークスの定理	65

10.5	おわりに：解析学のその先へ	65
Python で見る微積分 (Computational Calculus)		67
.1	テイラー展開の可視化～多項式近似の威力～	67
.2	勾配降下法シミュレーション～2変数関数の最適化～	68
.3	モンテカルロ積分～乱数で面積・体積を測る～	69
参考文献・推奨図書		71

第Ⅰ部

極限と連續性の厳密化（直感から論理へ）

第 1 章

実数の連続性と数列：無限を有限で制御する

高校数学（数 III）では、数列 $\{a_n\}$ がある値 α に収束することを $\lim_{n \rightarrow \infty} a_n = \alpha$ と書き、「 n が大きくなると a_n は α に限りなく近づく」と説明しました。しかし、「限りなく」とは具体的にどれくらいでしょうか？0.1まで？0.00001まで？

実社会、特に工学や数値計算の現場では、「誤差」を許容範囲内に収めることが至上命題です。「なんとなく近づく」ではなく、「どんなに厳しい誤差範囲 (ϵ) を突きつけられても、ある時点 (N) 以降なら必ずその範囲内に収まる」という論理で収束を定義し直します。これが有名な「 $\epsilon - N$ （イプシロン・エヌ）論法」です。

1.1 実数の連続性と上限・下限

微積分を厳密に行うには、舞台となる「実数」の性質を知る必要があります。有理数（分数で書ける数）だけでは、数直線上に「穴」が開いてしまいます（例： $\sqrt{2}$ の場所）。実数はこの穴が完全に埋まっている（連続性を持つ）体系です。

定義 1.1 (上界・下界・有界). 実数の集合 $A \subset \mathbb{R}$ について、

1. ある実数 M が存在して、任意の $x \in A$ に対して $x \leq M$ が成り立つとき、 A は上に有界であるといい、 M を A の上界 (upper bound) という。
2. ある実数 m が存在して、任意の $x \in A$ に対して $x \geq m$ が成り立つとき、 A は下に有界であるといい、 m を A の下界 (lower bound) という。
3. 上に有界かつ下に有界な集合を有界 (bounded) であるという。

定義 1.2 (上限 (Supremum)・下限 (Infimum)). 集合 A が上に有界であるとする。

1. A の上界の中で「最小のもの」を A の上限と呼び、 $\sup A$ と書く。
2. 同様に、 A の下界の中で「最大のもの」を A の下限と呼び、 $\inf A$ と書く。

コメント 1.1 (実社会のイメージ：最大値 (Max) と上限 (Sup) の違い). ある工場で、製品の耐久テストを行っているとします。強度は 0 から始まり、理論上の限界値 100 に近づくことはできますが、物理的な制約で 100 ジャストには絶対なりません。

- データセット $D = \{90, 99, 99.9, 99.99, \dots\}$
- この集合に「最大値 (Max)」は存在しません（どれだけ高い値を取っても、それより高い値があるため）。
- しかし、「上限 (Sup)」は存在し、それは 100 です。
- 「到達はできないかもしれないが、実質的な天井となる値」が上限です。最適化問題で「最小値が存在しないが、下限 (Inf) を目指してコストを下げる」場面などで頻出します。

例 1.1. 集合 $A = \left\{1 - \frac{1}{n} \mid n \in \mathbb{N}\right\} = \left\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\}$ について、 $\max A, \min A, \sup A, \inf A$ を求めよ。

【解答】

- 最小値・下限: $n = 1$ のとき $1 - 1 = 0$ 。これより小さい要素はないので、 $\min A = 0$ 。よって $\inf A = 0$ 。
- 最大値・上限: n が大きくなると値は 1 に近づくが、決して 1 にはならない ($1 - 1/n < 1$)。よって $\max A$ は存在しない。
- 一方、1 は明らかに上界である。1 より小さい任意の値 $1 - \epsilon$ (ただし $\epsilon > 0$) を考えると、十分大きな n を取れば $1 - 1/n > 1 - \epsilon$ ができる。つまり、1 より少しでも小さい値は上界になれない。よって最小の上界（上限）は $\sup A = 1$ 。
- 実数の連続性公理: 「空でない実数の部分集合が上に有界ならば、必ず上界が存在する。」(これは証明するものではなく、実数というものを規定するルールの 1 つです。有理数だけではこの性質は成り立たません。)

1.2 数列の収束 ($\epsilon - N$ 論法)

ここが最初の難関ですが、論理記号 (\forall, \exists) に慣れれば、これほど明確な定義はありません。

定義 1.3 (数列の収束). 数列 $\{a_n\}$ が α に収束する ($\lim_{n \rightarrow \infty} a_n = \alpha$) とは、以下の条件が成り立つことである。

$$\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ s.t. } \forall n \geq N, |a_n - \alpha| < \epsilon$$

【言葉での翻訳】

1. $\forall \epsilon > 0$ (挑戦状) : 「誤差を ϵ 以内にせよ」と、どんなに小さな正の数 ϵ を突きつけられても、
2. $\exists N \in \mathbb{N}$ (受けて立つ) : 「わかりました。では第 N 項目以降を見てください」と、ある番号 N を提示でき、
3. $\forall n \geq N, |a_n - \alpha| < \epsilon$ (保証) : 実際に N 項目以降の全ての a_n が、誤差 $\pm \epsilon$ の範囲に収まっている。

コメント 1.2 (実社会のイメージ：品質保証). あなたがスマートフォンの画面を作るエンジニアだとします。

- クライアント : 「画面サイズ (α) との誤差は 0.1mm 以内 ($\epsilon = 0.1$) にしてくれ」
- あなた : 「製造プロセスを $N = 100$ 回改良した以降のロットなら、全てその誤差内に収まります」
- クライアント : 「やっぱり 0.001mm 以内 ($\epsilon = 0.001$) にしてくれ」
- あなた : 「それなら、 $N = 5000$ 回改良した以降なら対応できます」

どんなに厳しい要求 (ϵ) が来ても、それに対応する改良回数 (N) を提示できるなら、あなたの技術は理想値に「収束」していると言えます。

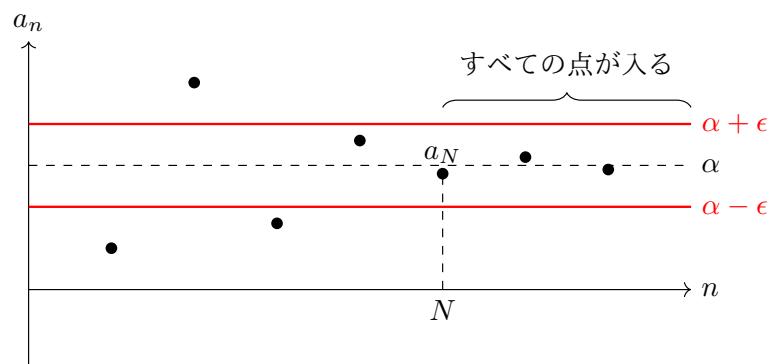


図 1.1 $\epsilon - N$ 論法のイメージ：ある時点 N 以降はすべて誤差帯（赤い線）の中に収まる。

例 1.2 (定義に基づいた証明). 数列 $a_n = \frac{1}{n}$ が 0 に収束することを、定義に従って証明せよ。

【証明】示したいことは、任意の $\epsilon > 0$ に対して、ある自然数 N が存在し、 $n \geq N \implies |\frac{1}{n} - 0| < \epsilon$ が成り立つことである。

1. 任意の $\epsilon > 0$ をとる。(これは相手が決める値なので固定して考える)
2. 不等式 $|\frac{1}{n}| < \epsilon$ を満たす n を見つけたい。 $\frac{1}{n} < \epsilon \iff n > \frac{1}{\epsilon}$
3. ここで、アルキメデスの性質(どんな実数よりも大きな自然数が存在する)により、 $N > \frac{1}{\epsilon}$ となる自然数 N を一つ選ぶことができる。(例えば、 $\epsilon = 0.01$ なら $1/\epsilon = 100$ なので、 $N = 101$ とすればよい)
4. この N を用いると、任意の $n \geq N$ に対して、 $n \geq N > \frac{1}{\epsilon} \implies \frac{1}{n} < \epsilon$ すなわち $|a_n - 0| < \epsilon$ が成り立つ。
5. よって定義により、 $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ である。(証明終)

例 1.3 (少し複雑な証明). 数列 $a_n = \frac{2n+1}{n+1}$ が 2 に収束することを証明せよ。

【証明】任意の $\epsilon > 0$ をとる。目標は、ある N を見つけて、 $n \geq N$ ならば $|a_n - 2| < \epsilon$ となるようにすることである。

まず、差の絶対値を計算して簡単にする。

$$|a_n - 2| = \left| \frac{2n+1}{n+1} - 2 \right| = \left| \frac{2n+1 - 2(n+1)}{n+1} \right| = \left| \frac{-1}{n+1} \right| = \frac{1}{n+1}$$

この値を ϵ 未満にしたい。

$$\frac{1}{n+1} < \epsilon \iff n+1 > \frac{1}{\epsilon} \iff n > \frac{1}{\epsilon} - 1$$

そこで、 $\frac{1}{\epsilon} - 1$ より大きい自然数 N を一つ選ぶ(例えはガウス記号を用いて $N = \lfloor \frac{1}{\epsilon} \rfloor + 1$ としてもよい)。すると、任意の $n \geq N$ に対して、

$$n > \frac{1}{\epsilon} - 1 \implies \frac{1}{n+1} < \epsilon$$

つまり $|a_n - 2| < \epsilon$ が成立する。よって、 $\lim_{n \rightarrow \infty} \frac{2n+1}{n+1} = 2$ である。(証明終)

1.3 収束の一意性

「収束するなら、行き先は一つしかない」という直感も、証明が必要です。

定理 1.1 (極限の一意性). 数列 $\{a_n\}$ が α に収束し、かつ β にも収束するならば、 $\alpha = \beta$ である。

Proof. 背理法で示す。 $\alpha \neq \beta$ と仮定する。 $\epsilon = \frac{|\alpha - \beta|}{2}$ と置く。 $\alpha \neq \beta$ なので $\epsilon > 0$ である。(直感的には、 α と β の距離の半分を誤差許容範囲とする、ということ)

1. $a_n \rightarrow \alpha$ なので、この ϵ に対してある N_1 が存在し、 $n \geq N_1 \implies |a_n - \alpha| < \epsilon$
2. $a_n \rightarrow \beta$ なので、同じ ϵ に対してある N_2 が存在し、 $n \geq N_2 \implies |a_n - \beta| < \epsilon$

ここで、 $N = \max(N_1, N_2)$ と置く。 $n \geq N$ なる n を一つ選ぶと、 $|a_n - \alpha| < \epsilon$ かつ $|a_n - \beta| < \epsilon$ が同時に成り立つ。

三角不等式 $|x + y| \leq |x| + |y|$ を用いると、

$$\begin{aligned} |\alpha - \beta| &= |(\alpha - a_n) + (a_n - \beta)| \leq |\alpha - a_n| + |a_n - \beta| \\ &= |a_n - \alpha| + |a_n - \beta| < \epsilon + \epsilon = 2\epsilon \end{aligned}$$

ところが、最初に $\epsilon = \frac{|\alpha - \beta|}{2}$ と定義したため、 $2\epsilon = |\alpha - \beta|$ である。よって不等式は

$$|\alpha - \beta| < |\alpha - \beta|$$

となり、矛盾する（ある数が自分自身より真に小さいことはあり得ない）。したがって仮定が誤りであり、 $\alpha = \beta$ でなければならない。□

1.4 第1章のまとめと演習

本章では「極限」を、曖昧な動的なプロセスではなく、「不等式によって確定した静的な論理関係」として再定義しました。この厳密さがあるからこそ、我々はコンピュータで無限を扱うアルゴリズム（ニュートン法や勾配降下法）の停止条件を設計できるのです。

(演習問題)

1. 集合 $B = \{x \in \mathbb{Q} \mid x^2 < 2\}$ について、 \mathbb{Q} 内での上限は存在しないが、 \mathbb{R} 内での上限は $\sqrt{2}$ であることを確認せよ。
2. $\epsilon - N$ 論法を用いて、 $\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} = 0$ を証明せよ。

第 2 章

関数の極限と連続：入力と出力の誤差制御

第 1 章では、自然数 n を大きくするという「離散的な無限」を扱いました。第 2 章では、実数 x をある値に近づけるという「連続的な無限」を扱います。

ここでは、高校数学で最も直感任せにされていた「関数が連続である（グラフがつながっている）」という概念を、論理のメスで解剖します。これは、制御工学における「安定性」や、機械学習における「汎化性能（入力の微小変化が出力を大きく変えないこと）」の数学的保証そのものです。

2.1 関数の極限 ($\epsilon - \delta$ 論法)

「 x が a に近づくとき、 $f(x)$ は b に近づく」高校数学ではこう言いましたが、これを工学的な「入力」と「出力」の関係として捉え直しましょう。

- 出力の要求精度 (ϵ) : クライアントは「目標値 b からの誤差を ϵ 未満にしてほしい」と言う。
- 入力の許容誤差 (δ) : エンジニアは「それなら、入力 x を設定値 a から δ の範囲内に収めれば、その要求を満たせます」と答える。

これが可能なとき、極限が存在すると言います。

定義 2.1 (関数の極限 ($\epsilon - \delta$ 論法)). 関数 $f(x)$ について、 $\lim_{x \rightarrow a} f(x) = b$ であるとは、以下の条件が成り立つことである。

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } 0 < |x - a| < \delta \implies |f(x) - b| < \epsilon$$

解説: $0 < |x - a|$ は「 x は a 自身ではない ($x \neq a$)」ことを意味します。極限はあくまで「近づく」状態であり、その点での値 ($f(a)$) は無関係です。

コメント 2.1 (実社会のイメージ：機械加工の公差 (Tolerance)). ある部品の直径 x が決まると、その強度 $f(x)$ が決まると言います。

- 設計部門 ($\forall \epsilon$) : 「強度の誤差は ± 0.1 (ϵ) 以内に収めてくれ」
- 製造現場 ($\exists \delta$) : 「わかりました。では旋盤の加工精度を $\pm 0.05\text{mm}$ (δ) 以内に制御します。そうすれば強度はその範囲に収まります」

どんなに厳しい強度の要求 (ϵ) が来ても、それに見合う加工精度 (δ) を提示できるなら、そのプロセスは制御可能です。

例 2.1 (1 次関数の極限). $f(x) = 3x - 1$ について、 $\lim_{x \rightarrow 2} f(x) = 5$ であることを証明せよ。

【証明】 任意の $\epsilon > 0$ をとる。我々が見つけたいのは、 $0 < |x - 2| < \delta$ ならば $|(3x - 1) - 5| < \epsilon$ となるような $\delta > 0$ である。不等式を変形する：

$$|(3x - 1) - 5| = |3x - 6| = 3|x - 2|$$

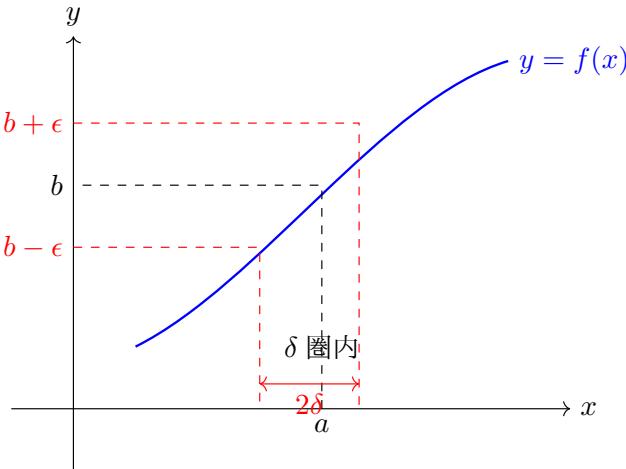


図 2.1 $\epsilon - \delta$ 論法のイメージ：出力の誤差 ϵ を決めると、入力の許容範囲 δ が決まる。

この値を ϵ 未満にしたいので、

$$3|x - 2| < \epsilon \iff |x - 2| < \frac{\epsilon}{3}$$

そこで、 $\delta = \frac{\epsilon}{3}$ と選べばよい。

実際に確認する： $\delta = \epsilon/3$ とおく。 $0 < |x - 2| < \delta$ なる任意の x に対して、

$$|f(x) - 5| = 3|x - 2| < 3\delta = 3 \cdot \frac{\epsilon}{3} = \epsilon$$

よって証明された。（証明終）

例 2.2 (2 次関数の極限 (δ の依存性)). $\lim_{x \rightarrow 2} x^2 = 4$ を証明せよ。

【証明】任意の $\epsilon > 0$ をとる。目標は $|x^2 - 4| < \epsilon$ を満たす δ を見つけることである。

$$|x^2 - 4| = |x - 2||x + 2|$$

ここで $|x - 2|$ は δ で制御できるが、 $|x + 2|$ が邪魔である。そこで、まず x の範囲を大雑把に限定する。とりあえず $\delta \leq 1$ と決めてしまう（ δ は小さければ小さいほど良いので、上限を設けても構わない）。すると $|x - 2| < 1$ より $1 < x < 3$ なので、

$$|x + 2| < 3 + 2 = 5$$

評価式はこうなる：

$$|x^2 - 4| = |x - 2||x + 2| < 5|x - 2|$$

これが ϵ 未満になればよいので、 $5|x - 2| < \epsilon \iff |x - 2| < \epsilon/5$ 。

以上より、 $\delta = \min(1, \frac{\epsilon}{5})$ と選ぶ。これを用いれば、

$$|x^2 - 4| = |x - 2||x + 2| < \frac{\epsilon}{5} \cdot 5 = \epsilon$$

となり、成立する。（証明終）

2.2 関数の連續性

極限の定義ができれば、「連續」は一瞬で定義できます。

定義 2.2 (点 a における連續). 関数 $f(x)$ が $x = a$ で連續であるとは、以下が成り立つことである。

$$\lim_{x \rightarrow a} f(x) = f(a)$$

$\epsilon - \delta$ 論法で書けば：

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } |x - a| < \delta \implies |f(x) - f(a)| < \epsilon$$

(極限と違い、 $x = a$ の場合も含めて考えるため $0 < |x - a|$ の条件が消えています)

2.3 中間値の定理 (The Intermediate Value Theorem)

直感的には「一本書きでグラフを書くとき、線をまたぐには必ず交点を通る」という定理ですが、これは「方程式の解の存在」を保証する強力なツールです。

定理 2.1 (中間値の定理). 関数 $f(x)$ が閉区間 $[a, b]$ で連続であり、 $f(a) < \gamma < f(b)$ ならば、 $f(c) = \gamma$ となる $c \in (a, b)$ が少なくとも一つ存在する。

上限の性質を利用した厳密な証明。集合 S を次のように定義します。

$$S = \{x \in [a, b] \mid f(x) < \gamma\}$$

Step 1: 上限 c の存在を示す

1. 空でないこと：仮定より $f(a) < \gamma$ なので、 $a \in S$ です。よって $S \neq \emptyset$ (空集合ではない)。
2. 上に有界であること： S は閉区間 $[a, b]$ の部分集合なので、すべての要素は b 以下です。よって S は上に有界です。

実数の連続性公理（上に有界な空でない実数の部分集合は、必ず上限を持つ）により、 S の上限（最小の上界）が存在します。これを c と置きます。

$$c = \sup S$$

Step 2: c が求める解であること ($f(c) = \gamma$) を示す背理法を用います。 $f(c) \neq \gamma$ 、すなわち $f(c) < \gamma$ または $f(c) > \gamma$ と仮定して矛盾を導きます。

Case A: もし $f(c) < \gamma$ ならば

1. f は連続なので、 c の十分近くでも値は γ より小さいはずです。厳密には、 $\epsilon = \gamma - f(c) > 0$ とおけば、連続性の定義より、ある $\delta > 0$ が存在して、 $|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$ すなわち $f(x) < f(c) + \epsilon = \gamma$ となります。
2. これは、 c より少し右側の点（例えば $c + \frac{\delta}{2}$ ）でも $f(x) < \gamma$ が成り立つことを意味します。
3. すると $c + \frac{\delta}{2} \in S$ となります。これは c が S の上界（上界の中で最小のもの）であること (c より大きい要素は S にないはず) に矛盾します。（※ただし $c = b$ の場合は $f(b) < \gamma$ となり仮定 $f(b) > \gamma$ に矛盾するので除外できます）

Case B: もし $f(c) > \gamma$ ならば

1. 同様に連続性より、 c の十分近くで値は γ より大きいはずです。ある $\delta > 0$ が存在して、 $(c - \delta, c]$ の範囲で $f(x) > \gamma$ となります。
2. これは、 c より少し左側の点（例えば $c - \frac{\delta}{2}$ ）より右には S の要素が存在しない（値が γ を超えてしまうため）ことを意味します。
3. すると $c - \frac{\delta}{2}$ も S の上界となってしまいます。これは c が最小の上界であることに矛盾します（ c より小さい上界があつてはならない）。

結論: Case A も Case B も矛盾するため、消去法により

$$f(c) = \gamma$$

でなければなりません。また、 $f(a) < \gamma < f(b)$ より $c \neq a$ かつ $c \neq b$ なので、 $c \in (a, b)$ です。 \square

コメント 2.2 (実社会のイメージ：二分法 (Bisection Method))。コンピュータで方程式 $f(x) = 0$ の解を数値的に求める際、「 $f(a) < 0$ かつ $f(b) > 0$ 」となる区間を見つけ、その中点を調べる作業を繰り返します。このアルゴリズムが「必ず解に収束する」ことを保証しているのが、この中間値の定理です。

2.4 最大値・最小値の定理

「閉区間上の連続関数は、必ず最大値と最小値を持つ」当たり前に見えますが、開区間 $(0, 1)$ での $f(x) = x$ （最大値なし、上限は 1）などを考えると、閉区間であることの重要性がわかります。

定理 2.2 (最大値・最小値の定理). 関数 $f(x)$ が有界閉区間 $[a, b]$ 上で連続ならば、 $f(x)$ はこの区間で最大値および最小値を持つ。

最大値・最小値の定理の完全証明。この証明には、以下の実数の性質（ボルツァノ・ワイエルシュトラスの定理）を用います。補題（ボルツァノ・ワイエルシュトラスの定理）：有界な数列 $\{x_n\}$ は、収束する部分列 $\{x_{n_k}\}$ を持つ。

証明は 2 段階で行います。

1. 有界性の証明： $f(x)$ の値域が無限大に発散しないことを示す。
2. 最大値の存在証明：上限に到達する点 c が存在することを示す。

Step 1: $f(x)$ の有界性の証明

背理法を用いる。 $f(x)$ が I 上で有界でないと仮定する。（ここでは「上に有界でない」と仮定する。下に有界でない場合も同様）

有界でないならば、どのような自然数 n に対しても、

$$f(x_n) > n$$

となる点 $x_n \in I$ が存在するはずである。これによって数列 $\{x_n\}$ が作れる。

数列 $\{x_n\}$ はすべて区間 $I = [a, b]$ に含まれるので、有界な数列である。よって、ボルツァノ・ワイエルシュトラスの定理より、収束する部分列 $\{x_{n_k}\}$ が存在する。その極限を c としよう。

$$\lim_{k \rightarrow \infty} x_{n_k} = c$$

I は閉区間なので、極限 c も I に含まれる ($a \leq x_{n_k} \leq b$ なら極限もその範囲内)。 f は c で連続なので、関数の極限値は関数值と一致するはずである。

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = f(c)$$

しかし、数列の作り方から $f(x_{n_k}) > n_k$ であり、 $k \rightarrow \infty$ と共に $n_k \rightarrow \infty$ なので、

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = \infty$$

となる。これは「 $f(c)$ という有限の値に収束する」ことと矛盾する。よって、仮定が誤りであり、 $f(x)$ は I 上で有界でなければならない。

Step 2: 最大値の存在証明（上限への到達性）

Step 1 より、 $f(x)$ の値域 $f(I) = \{f(x) \mid x \in I\}$ は上に有界な実数の集合である。実数の連続性公理より、必ず上限 (Supremum) が存在する。これを M とおく。

$$M = \sup\{f(x) \mid x \in I\}$$

これから、 $f(c) = M$ となる $c \in I$ が存在することを示す。

上限の定義より、 M は最小の上界なので、 M より少しでも小さい値は上界ではない。任意の自然数 n に対して、 $M - \frac{1}{n}$ は上界ではないので、これより大きい値をとる点 $y_n \in I$ が存在する。

$$M - \frac{1}{n} < f(y_n) \leq M$$

この数列 $\{y_n\}$ は I に含まれる有界数列なので、再びボルツァノ・ワイエルシュトラスの定理より、収束する部分列 $\{y_{n_k}\}$ が存在する。その極限を $c \in I$ とする（閉区間なので極限も中に入る）。

$$\lim_{k \rightarrow \infty} y_{n_k} = c$$

f の連続性より、

$$\lim_{k \rightarrow \infty} f(y_{n_k}) = f(c)$$

一方で、不等式 $M - \frac{1}{n_k} < f(y_{n_k}) \leq M$ において $k \rightarrow \infty$ とすると、挟み撃ちの原理により

$$M - 0 \leq \lim_{k \rightarrow \infty} f(y_{n_k}) \leq M$$

すなわち

$$f(c) = M$$

これにより、値域の上限 M を値としてとる点 c が区間に実在することが示された。したがって M は最大値である。（最小値についても、 $-f(x)$ の最大値を考えることで同様に証明できる） \square

コメント 2.3 (実社会のイメージ：最適化と損失関数). 機械学習において、モデルのパラメータ空間（閉領域に制限された空間）で損失関数が連続であれば、「最適なパラメータ（損失を最小にする点）」が必ず存在することが保証されます。「解が存在する」とわかっているからこそ、我々は勾配降下法などでそれを探しに行くことができるのです。

2.5 一様連続性～より強い連続性～

最後に、積分の可能性を議論するために必要な「一様連続」という概念に触れます。これは $\epsilon - \delta$ の順序に関わる深い概念です。

定義 2.3 (一様連続 (Uniform Continuity)). 区間 I 上の関数 f が一様連続であるとは、

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } \forall x, y \in I, |x - y| < \delta \implies |f(x) - f(y)| < \epsilon$$

通常の連続との違い：

- 通常の連続: δ は「場所 x 」と「誤差 ϵ 」に依存して決まる ($\delta(x, \epsilon)$)。急な坂道では δ を小さくする必要がある。
- 一様連続: δ は「誤差 ϵ 」だけで決まり、「場所 x 」に依らない ($\delta(\epsilon)$)。区間内のどこであっても、共通の δ が通用する。

例 2.3 (一様連続でない関数). $f(x) = 1/x$ を区間 $(0, 1)$ で考える。 x が 0 に近づくとグラフの傾きは無限大になる。どんなに小さな δ を固定しても、 x を十分に 0 に近づければ、幅 δ の区間での $f(x)$ の変化量は無限に大きくなり、 ϵ に収まらない。（逆に、閉区間 $[0.1, 1]$ ならば一様連続になる。これがハイネ・カントールの定理「有界閉区間上の連続関数は一様連続である」の意味です）

2.6 第2章のまとめ

- $\epsilon - \delta$ 論法: 入力誤差 δ と出力誤差 ϵ の制御関係。
- 中間値の定理: 連続変化する量は、途中を飛ばさない。
- 最大値・最小値の定理: コンパクトな領域上の連続関数は、最適解を持つ。

これらの道具を使えば、次章からの「微分（線形近似）」の議論において、「誤差のオーダー」を厳密に評価することが可能になります。

第 II 部

1 変数関数の深化（線形近似から高次近似へ）

第 3 章

微分法の基礎と平均値の定理：最強の近似ツール

第 1 章・第 2 章で、実数という「舞台」と、連続関数という「役者」の性質を固めました。第 3 章からは、いよいよ解析学の最大の武器である「線形近似 (Linear Approximation)」、すなわち微分法に入ります。

高校数学では「接線の傾き」として教わりましたが、大学数学・工学・統計学においては、微分とは「複雑な関数を、一次関数（直線）で近似すること」です。この視点がなければ、多変数関数や機械学習の最適化理論は理解できません。

3.1 微分可能性と線形近似

ある点 $x = a$ において関数 $f(x)$ が微分可能であるとは、その点の周りを超拡大すると直線に見えるということです。

定義 3.1 (微分係数). 関数 f の $x = a$ における微分係数 $f'(a)$ を次のように定義する。

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

この極限が存在するとき、 f は $x = a$ で微分可能であるという。

【視点の転換：線形近似式】この定義式を変形すると、ランダウの記号 $o(\cdot)$ （スモールオー：高次の微小項）を用いて次のように書けます。

$$f(a + h) = f(a) + f'(a)h + o(h) \quad (h \rightarrow 0)$$

ここで $o(h)$ は $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$ を満たす項 (h よりも速く 0 に収束する誤差項) です。

この式の意味: 「関数 f の値の変化は、定数項 $f(a)$ と、 h に比例する項 $f'(a)h$ （線形部分）で大体説明できる。残りの誤差 $o(h)$ は無視できるほど小さい」

コメント 3.1 (実社会のイメージ：感度解析 (Sensitivity Analysis)). ある化学プラントで、温度 T を制御して収率 $Y = f(T)$ を最大化したいとします。現在の温度 $T = 100^\circ$ での微分係数が $f'(100) = 2.5$ だったとします。これは、「温度を 1° 上げれば、収率は約 2.5% 上がる（線形近似）」ことを意味します。実際の関数 $f(T)$ がどんなに複雑な非線形曲線であっても、「局所的には比例関係」とみなして制御則を作れるのが微分の強みです。

3.2 ロルの定理 (Rolle's Theorem)

平均値の定理、そしてテイラー展開へと続く、微分積分学の論理的な土台となる定理です。

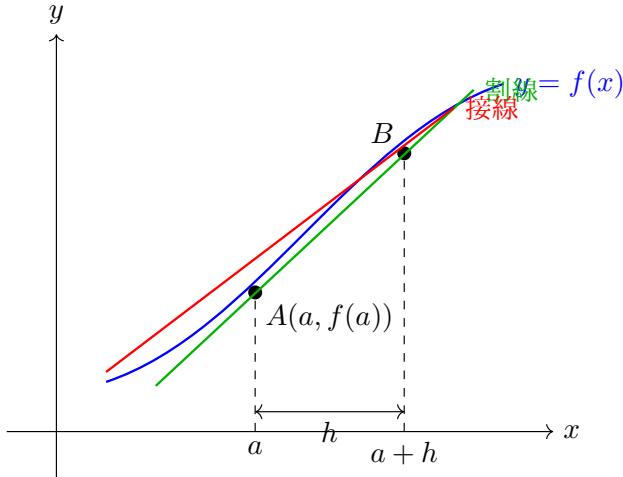
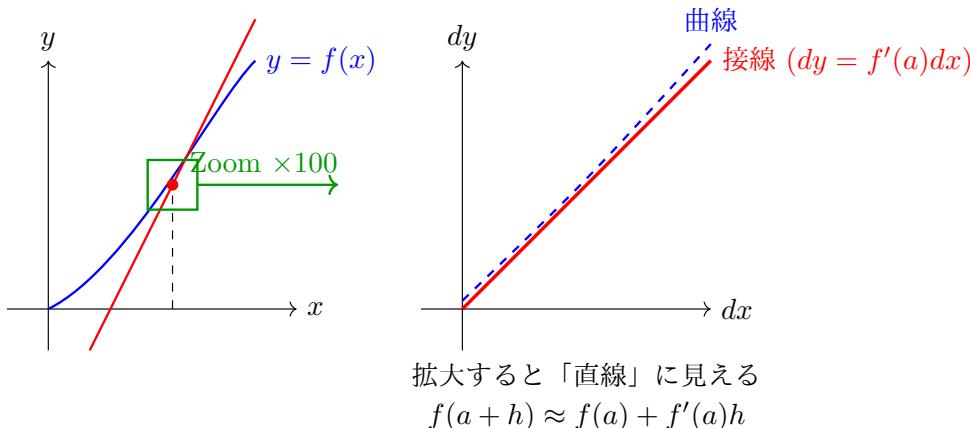
図 3.1 平均変化率（緑）と微分係数（赤）： $h \rightarrow 0$ で割線は接線に近づく。

図 3.2 微分の本質：局所的な線形化（どんな曲線も、拡大すれば直線になる）。

定理 3.1 (ロルの定理). 関数 $f(x)$ が閉区間 $[a, b]$ で連続、開区間 (a, b) で微分可能であり、 $f(a) = f(b)$ ならば、

$$f'(c) = 0$$

となる $c \in (a, b)$ が少なくとも一つ存在する。

Proof. $f(x)$ は閉区間 $[a, b]$ で連続なので、最大値・最小値の定理（定理 2.2）により、この区間内で最大値 M と最小値 m を必ず持ります。

以下の 2 つのケースに分けて考えます。

Case 1: 最大値 M と最小値 m が一致する場合 ($M = m$) このとき、区間内のすべての x に対して $f(x) = M$ (定数関数) となります。定数関数の微分は常に 0 なので、開区間 (a, b) 内のすべての点 c で $f'(c) = 0$ が成り立ちます。よって定理は成立します。

Case 2: 最大値 M と最小値 m が異なる場合 ($M \neq m$) 仮定より $f(a) = f(b)$ なので、最大値 M または最小値 m の少なくとも一方は、端点 (a や b) の値とは異なります。すなわち、 $f(c) = M$ (または $f(c) = m$) となる点 c が、区間の内部 (a, b) に存在します。ここでは、 $f(c) = M$ (最大値) をとる点 $c \in (a, b)$ が存在すると仮定して証明を進めます (最小値の場合も同様です)。

点 c で最大値をとるということは、任意の h (ただし $c + h$ が区間内にあるような小さな h) に対して

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

が成り立ちます。よって、 $f(c + h) - f(c) \leq 0$ です。

ここで、微分係数の定義式（右極限と左極限）の符号を調べます。

1. 右側からの極限 ($h > 0$): 分子は ≤ 0 、分母は $h > 0$ なので、分数は 0 以下です。

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

よって、右側微分係数は

$$\lim_{h \rightarrow +0} \frac{f(c+h) - f(c)}{h} \leq 0$$

2. 左側からの極限 ($h < 0$): 分子は ≤ 0 、分母は $h < 0$ なので、分数は 0 以上（負÷負）になります。

$$\frac{f(c+h) - f(c)}{h} \geq 0$$

よって、左側微分係数は

$$\lim_{h \rightarrow -0} \frac{f(c+h) - f(c)}{h} \geq 0$$

仮定より $f(x)$ は c で微分可能なので、右側極限と左側極限は一致し、有限な値 $f'(c)$ にならなければなりません。「0 以下」かつ「0 以上」である値は 0 しかありません。

$$f'(c) \leq 0 \quad \text{かつ} \quad f'(c) \geq 0 \implies f'(c) = 0$$

よって、 $f'(c) = 0$ となる点 c が存在することが証明されました。 \square

3.3 平均値の定理 (Mean Value Theorem)

微分の歴史において最も重要な定理の一つです。「瞬間の変化率」と「平均の変化率」をつなぎます。

定理 3.2 (平均値の定理). 関数 $f(x)$ が閉区間 $[a, b]$ で連続、開区間 (a, b) で微分可能ならば、

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

となる $c \in (a, b)$ が少なくとも一つ存在する。

Proof. ロルの定理に帰着させるための補助関数を作るテクニックを使う。関数 $g(x)$ を、「曲線 $y = f(x)$ 」と「2 点 $(a, f(a)), (b, f(b))$ を結ぶ直線」との差として定義する。直線の式は $y = f(a) + \frac{f(b)-f(a)}{b-a}(x - a)$ なので、

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

この $g(x)$ について確認すると：

1. $g(x)$ は閉区間 $[a, b]$ で連続、開区間 (a, b) で微分可能 ($f(x)$ と一次関数の差だから)。
2. $g(a) = f(a) - f(a) = 0$
3. $g(b) = f(b) - f(b) = 0$

よって $g(a) = g(b) = 0$ となり、ロルの定理の条件を満たす。

したがって、 $g'(c) = 0$ となる $c \in (a, b)$ が存在する。 $g(x)$ を微分すると：

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

$g'(c) = 0$ より、

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0 \iff f'(c) = \frac{f(b) - f(a)}{b - a}$$

\square

コメント 3.2 (実社会のイメージ：区間速度取締システム). 高速道路で、地点 A と地点 B (距離 100km) を通過する時間を計測するシステムがあります。

- A 地点通過 : 10:00
- B 地点通過 : 10:30
- 平均速度 : $100\text{km}/0.5\text{h} = 200\text{km/h}$

平均値の定理によれば、「瞬間の速度計がちょうど 200km/h を指した瞬間 (時点 c)」が、道中のどこかで少なくとも一回は必ず存在します。だから「ずっと安全運転でした」という言い訳は数学的に却下され、スピード違反となります。

3.4 コーシーの平均値の定理とロピタルの定理

平均値の定理を 2 つの関数に拡張します。これは不定形の極限 ($\frac{0}{0}$ や $\frac{\infty}{\infty}$) を計算する「ロピタルの定理」の証明に使われます。

定理 3.3 (コーシーの平均値の定理). 関数 $f(x), g(x)$ が $[a, b]$ で連続、 (a, b) で微分可能であり、 $g'(x) \neq 0$ ならば、

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

となる $c \in (a, b)$ が存在する。

定理 3.4 (ロピタルの定理). $\lim_{x \rightarrow a} f(x) = 0, \lim_{x \rightarrow a} g(x) = 0$ であり、 $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = L$ が存在するならば、

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = L$$

Proof. $f(a) = g(a) = 0$ と定義して連続関数に拡張する。コーシーの平均値の定理より、区間 $[a, x]$ の間のどこか c ($a < c < x$) において、

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

が成り立つ。ここで $x \rightarrow a$ とすると、挟み撃ちの原理のように c も a に近づく ($a < c < x$ ので)。仮定より $\lim_{c \rightarrow a} \frac{f'(c)}{g'(c)} = L$ なので、

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = L$$

□

例 3.1 (ロピタルの定理の適用). $\lim_{x \rightarrow 0} \frac{\sin x - x}{x^3}$ を求めよ。

【解答】 $x \rightarrow 0$ で分母・分子ともに 0 になる ($\frac{0}{0}$ 型)。

1. 微分を実行 : 分母 $(x^3)' = 3x^2$ 分子 $(\sin x - x)' = \cos x - 1 \lim_{x \rightarrow 0} \frac{\cos x - 1}{3x^2}$ (まだ $\frac{0}{0}$ 型)
2. もう一度微分 : 分母 $(3x^2)' = 6x$ 分子 $(\cos x - 1)' = -\sin x \lim_{x \rightarrow 0} \frac{-\sin x}{6x}$ (まだ $\frac{0}{0}$ 型だが、 $\frac{\sin x}{x} \rightarrow 1$ を知っていればここで終了。知らなければもう一度)
3. もう一度微分 : 分母 $(6x)' = 6$ 分子 $(-\sin x)' = -\cos x \lim_{x \rightarrow 0} \frac{-\cos x}{6} = \frac{-1}{6}$

よって極限値は $-\frac{1}{6}$ である。(これは次章のテイラー展開 $\sin x \approx x - \frac{x^3}{6}$ からも見通しよく理解できます)

数学史上、最も有名な定理の一つ「ロピタルの定理」ですが、実はこの定理を発見したのはロピタル (Guillaume de l'Hôpital, 1661 - 1704) ではありません。真の発見者は、彼の家庭教師であったヨハン・ベルヌーイ (Johann Bernoulli) です。

1. 貴族ロピタルと天才ベルヌーイの出会い 17世紀後半、フランスの裕福な貴族であったロピタル侯爵は、数学愛好家でもありました。彼は当時最先端の「微積分学（ライプニッツが開発したばかりの手法）」を学びたいと考え、当時最高の数学者の一人、ヨハン・ベルヌーイを家庭教師として雇います。

2. 歴史に残る「300 リーブルの契約」 1694年、ロピタルはベルヌーイに対して驚くべき契約を持ちかけます。「年間 300 リーブル（現在の価値で数百万円相当）を支払う代わりに、あなたの数学的発見を私に譲り、他言しないこと」

お金に困っていた若きベルヌーイはこの契約を受け入れました。そして彼は、不定形の極限を計算する画期的な方法（現在のロピタルの定理）を発見し、それをロピタルに手紙で送ったのです。

3. 世界初の微積分教科書 1696年、ロピタルはこの成果をまとめ、世界初の微積分学の教科書『曲線の理解のための無限小解析』を出版しました。この本は大ベストセラーとなり、その中に記された定理は「ロピタルの定理」として世界中に広りました。

4. 死後の暴露 ロピタルの生前、ベルヌーイは契約を守り沈黙していましたが、ロピタルの死後、「あれは私の発見だ！」と暴露しました。当初は誰も信じませんでしたが、20世紀になってからベルヌーイからロピタルへの手紙が発見され、その主張が真実であったことが証明されたのです。

結論：ロピタルの定理は、「ベルヌーイが発見し、ロピタルが買い取り、世界に広めた定理」です。しかし、ロピタルの功績がゼロというわけではありません。難解だった微積分を体系化し、教科書として普及させた彼の手腕があったからこそ、この定理は現代まで受け継がれてきたのです。

3.5 第3章のまとめ

- **微分係数：** 関数の局所的な「線形近似（傾き）」。
- **平均値の定理：** 局所的な変化 ($f'(c)$) と大域的な変化 ($f(b) - f(a)$) を結びつける、解析学の架け橋。
- **ロピタルの定理：** 複雑な関数の比の極限を、導関数の比に帰着させて解くツール。

第4章では、この「線形近似（1次式）」を「高次多項式近似」へとアップグレードするテイラーフェルヌーイ展開が登場します。統計学や機械学習において、なぜ2次近似（分散やヘッセ行列）が重要なのか、その数理的背景が明らかになります。

第 4 章

テイラー展開と級数：関数を多項式に分解する

第 3 章では、関数を「直線（1 次式）」で近似する微分法を学びました。第 4 章では、それをさらに推し進め、「関数を n 次多項式で近似する」という、応用数学における最強のツール「テイラー展開」を扱います。

これは、「どんなに複雑な関数も、足し算と掛け算（多項式）だけで表現できる」ことを意味し、コンピュータが $\sin x$ や e^x を計算できる原理そのものです。また、統計学においては「分散（2 次のモーメント）」や「歪度・尖度（3 次・4 次）」といった概念の基礎となります。

4.1 高次導関数と多項式近似

まず、なぜ高次導関数を考える必要があるのか、直感的に理解しましょう。

- $f(a)$: 現在地（0 次の情報）
- $f'(a)$: 速度（1 次の情報）。これだけだと「ずっと同じ速度で進む（直線）」予測しかできません。
- $f''(a)$: 加速度（2 次の情報）。これがあれば「カーブ（曲がり具合）」まで予測できます。
- $f^{(n)}(a)$: さらに詳しい未来の挙動。

情報を増やすほど、元の関数 $f(x)$ にピタリと重なる多項式を作ることができます。

4.2 テイラーの定理 (Taylor's Theorem)

平均値の定理の一般化です。「 n 次多項式で近似した時の誤差」を定量的に評価します。

定理 4.1 (テイラーの定理). 関数 $f(x)$ が閉区間 $[a, x]$ で n 回連続微分可能、開区間 (a, x) で $n+1$ 回微分可能であるとする。このとき、

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_{n+1}$$

を満たす剩余項 R_{n+1} が存在する。ラグランジュの剩余形式では、ある $c \in (a, x)$ を用いて次のように書ける：

$$R_{n+1} = \frac{f^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1}$$

証明（補助関数 $F(t)$ を使う方法）。まず、剩余項 R_{n+1} の形を特定するための定数 K を導入します。次の等式を満たすように定数 K を定義します。

$$f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + K(x - a)^{n+1}$$

(つまり、真の値 $f(x)$ と多項式近似との「誤差」を $K(x-a)^{n+1}$ という形で表現したときの係数が K です)
次に、以下の補助関数 $F(t)$ を区間 $[a, x]$ 上で定義します。

$$F(t) = f(x) - \left(f(t) + f'(t)(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n)}(t)}{n!}(x-t)^n + K(x-t)^{n+1} \right)$$

この関数 $F(t)$ の性質を調べます。

1. 端点 $t = x$ での値: $t = x$ を代入すると、 $(x-t)$ の項はすべて 0 になるので、

$$F(x) = f(x) - f(x) = 0$$

2. 端点 $t = a$ での値: $t = a$ を代入すると、

$$F(a) = f(x) - \left(\sum_{k=0}^n \frac{f^{(k)}(a)}{k!}(x-a)^k + K(x-a)^{n+1} \right)$$

これは最初に K を定義した式そのものなので、移項すれば 0 になります。

$$F(a) = 0$$

3. 微分可能性とロルの定理の適用: $F(t)$ は仮定より $[a, x]$ で連続、 (a, x) で微分可能です。 $F(a) = F(x) = 0$ なので、ロルの定理より、 $F'(c) = 0$ となる $c \in (a, x)$ が存在します。

4. $F'(t)$ の計算: $F(t)$ を t で微分します。積の微分法 $(uv)' = u'v + uv'$ を各項に適用します。定数項 $f(x)$ の微分は 0 です。カッコの中身を微分します。

- 第 1 項: $(f(t))' = f'(t)$
- 第 2 項: $(f'(t)(x-t))' = f''(t)(x-t) + f'(t)(-1)$
- 第 3 項: $(\frac{f''(t)}{2!}(x-t)^2)' = \frac{f'''(t)}{2!}(x-t)^2 + \frac{f''(t)}{2!} \cdot 2(x-t)(-1) = \frac{f'''(t)}{2!}(x-t)^2 - f''(t)(x-t)$
- 一般項 (k 番目): $\left(\frac{f^{(k)}(t)}{k!}(x-t)^k \right)' = \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1}$

これらを並べて足し合わせると、見事に「隣り合う項が打ち消し合う（望遠鏡和）」ことがわかります。

$$\begin{aligned} -\frac{d}{dt}(\dots) &= - \left[f'(t) + (f''(t)(x-t) - f'(t)) + \left(\frac{f'''(t)}{2!}(x-t)^2 - f''(t)(x-t) \right) + \dots \right] \\ &\quad \dots - \left[\frac{f^{(n+1)}(t)}{n!}(x-t)^n - \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} \right] - [K(n+1)(x-t)^n(-1)] \end{aligned}$$

残るのは、多項式部分の最後の微分の項と、 K の項だけです。

$$F'(t) = - \left(\frac{f^{(n+1)}(t)}{n!}(x-t)^n - K(n+1)(x-t)^n \right)$$

共通因数 $(x-t)^n$ でくくります。

$$F'(t) = (x-t)^n \left(K(n+1) - \frac{f^{(n+1)}(t)}{n!} \right)$$

5. 結論の導出: ロルの定理より、 $F'(c) = 0$ となる c が存在します。 $c \in (a, x)$ なので $(x-c)^n \neq 0$ です。よってカッコの中身が 0 でなければなりません。

$$K(n+1) - \frac{f^{(n+1)}(c)}{n!} = 0$$

これを K について解くと：

$$K = \frac{f^{(n+1)}(c)}{(n+1)!}$$

最初に定義した剩余項は $R_{n+1} = K(x - a)^{n+1}$ でした。これに求めた K を代入すると、

$$R_{n+1} = \frac{f^{(n+1)}(c)}{(n+1)!} (x - a)^{n+1}$$

となり、ラグランジュの剩余形式が導かれました。

□

コメント 4.1 (実社会のイメージ：物理シミュレーションの精度). ゲーム物理や流体解析では、時間の刻み幅 Δt で未来の状態を予測します。

- オイラー法（1次近似）： $x(t + \Delta t) \approx x(t) + v(t)\Delta t$ 。精度は低いですが計算は速い。
- 高次解法（ルンゲ・クッタ法など）：テイラー展開の高次の項（加速度や躍度など）相当の情報を取り込み、驚異的な精度を実現します。

剩余項 R_{n+1} は、このシミュレーションで「切り捨てられた誤差」が最大どれくらいになるかを見積もるために使われます。

4.3 マクローリン展開

$a = 0$ の場合のテイラー展開を特にマクローリン展開と呼びます。公式を暗記するのではなく、「微分して 0 を代入する」という単純作業の繰り返しで導けることを確認しましょう。

マクローリン展開の公式：

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n$$

主要な関数の導出プロセス

1. 指数関数 $f(x) = e^x$

微分しても形が変わらない、最も基本的な関数です。

- $f(x) = e^x \implies f(0) = e^0 = 1$
- $f'(x) = e^x \implies f'(0) = 1$
- $f''(x) = e^x \implies f''(0) = 1$
- $f^{(n)}(x) = e^x \implies f^{(n)}(0) = 1$

よって、すべての係数の分子は 1 です。

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

2. 正弦関数 $f(x) = \sin x$

4 回微分すると元に戻る周期性を持ちます。

- $f(0) = \sin 0 = 0$
- $f'(x) = \cos x \implies f'(0) = \cos 0 = 1$
- $f''(x) = -\sin x \implies f''(0) = 0$
- $f'''(x) = -\cos x \implies f'''(0) = -1$
- $f^{(4)}(x) = \sin x \implies f^{(4)}(0) = 0$ (ここから繰り返し)

係数の列は $0, 1, 0, -1, 0, 1, \dots$ となります。偶数次の項はすべて消え、奇数次の項だけが符号を変えながら残ります。

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

(直感：原点で $y = x$ に接し、次は x^3 で下側に曲がり… という補正を繰り返しています)

3. 余弦関数 $f(x) = \cos x$

同様に微分します。

- $f(0) = \cos 0 = 1$
- $f'(x) = -\sin x \implies f'(0) = 0$
- $f''(x) = -\cos x \implies f''(0) = -1$
- $f'''(x) = \sin x \implies f'''(0) = 0$
- $f^{(4)}(x) = \cos x \implies f^{(4)}(0) = 1$

係数の列は $1, 0, -1, 0, 1, 0, \dots$ となります。今度は奇数次の項が消え、偶数次の項だけが残ります。

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

(直感：原点で $y = 1$ (水平) から始まり、放物線 $1 - x^2/2$ で落ちていく形です)

4. 対数関数 $f(x) = \log(1+x)$

$x = 0$ で $\log x$ は定義できないため、 $\log(1+x)$ を展開します。

- $f(x) = \log(1+x) \implies f(0) = \log 1 = 0$
- $f'(x) = (1+x)^{-1} \implies f'(0) = 1$
- $f''(x) = -(1+x)^{-2} \implies f''(0) = -1$
- $f'''(x) = 2(1+x)^{-3} \implies f'''(0) = 2$
- $f^{(4)}(x) = -6(1+x)^{-4} \implies f^{(4)}(0) = -6$
- 一般項: $f^{(n)}(0) = (-1)^{n-1}(n-1)!$

これをテイラー係数 $\frac{1}{n!}$ と掛け合わせると、 $\frac{(-1)^{n-1}(n-1)!}{n!} = \frac{(-1)^{n-1}}{n}$ となり、階乗が消えます。

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

(※収束半径は $|x| < 1$ です。階乗で割らないため、収束は遅いです)

5. 幾何級数（等比級数） $f(x) = \frac{1}{1-x}$

これは微分しなくても、「等比級数の和の公式」から逆算できますが、微分でも確認できます。

- $f(x) = (1-x)^{-1} \implies f(0) = 1$
- $f'(x) = (1-x)^{-2} \implies f'(0) = 1$
- $f''(x) = 2(1-x)^{-3} \implies f''(0) = 2 = 2!$
- $f'''(x) = 6(1-x)^{-4} \implies f'''(0) = 6 = 3!$
- 一般項: $f^{(n)}(0) = n!$

係数は $\frac{n!}{n!} = 1$ となり、すべて 1 になります。

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

例 4.1 (オイラーの公式の「発見」). e^{ix} を形式的にマクローリン展開してみよ (i は虚数単位)。

【解答】 e^z の展開式に $z = ix$ を代入する。

$$e^{ix} = 1 + (ix) + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \dots$$

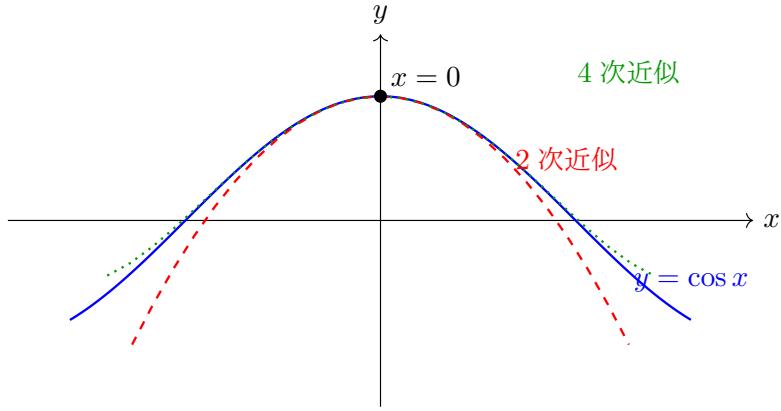


図 4.1 テイラー展開のイメージ：次数を上げるほど元の関数（青）に近づく。

$i^2 = -1, i^3 = -i, i^4 = 1$ などを整理すると：

$$e^{ix} = \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots\right) + i \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots\right)$$

カッコの中身を見ると、実部は $\cos x$ 、虚部は $\sin x$ の展開式そのものである。よって、人類至高の公式：

$$e^{ix} = \cos x + i \sin x$$

が導かれる。

コメント 4.2 (実社会のイメージ：信号処理とフーリエ変換). このオイラーの公式は、電気信号（交流電流）や音波を解析する「フーリエ変換」の核心です。「回転する複素数 e^{ix} 」こそが、波の重ね合わせを記述する最もシンプルな言葉なのです。

4.4 ランダウの記号 (Big-O, Small-o) の拡張版

近似を行う際、「どの程度の速さで変化するか（誤差の減り方、あるいは計算量の増え方）」を表す世界共通の記法です。「どこに近づくか ($x \rightarrow a$ なのか $x \rightarrow \infty$ なのか）」によって意味合いが変わる点に注意が必要です。

4.4.1 定義と意味

関数 $f(x), g(x)$ について、ある極限操作 ($x \rightarrow a$ または $x \rightarrow \infty$) を考えます。

定義 4.1 (ランダウの記号). • スモールオー (o): 「 $g(x)$ より無視できるほど小さい」

$$f(x) = o(g(x)) \iff \lim \frac{f(x)}{g(x)} = 0$$

– $x \rightarrow 0$ の例: $x^3 = o(x)$ (x が小さくなるとき、 x^3 はもっと急激に小さくなる)

– $x \rightarrow \infty$ の例: $\sqrt{x} = o(x)$ (x が大きくなるとき、 \sqrt{x} の増え方は x に比べて無視できるほど遅い)

• ビッグオー (O): 「 $g(x)$ と同程度か、それ以下である（抑えられる）」

$$f(x) = O(g(x)) \iff \limsup \left| \frac{f(x)}{g(x)} \right| < \infty$$

（比が有限の値に収まる、あるいは有界である）

4.4.2 ケーススタディ：状況による使い分け

Case A: 微小極限 ($x \rightarrow 0$) ~テイラー展開・解析学~

「主要な項（次数の低い項）」が支配し、次数の高い項はゴミ（誤差）として扱われます。

- $\sin x = x + O(x^3)$
- 意味：「 $\sin x$ はだいたい x と同じだが、誤差は x^3 程度ある」
- ここで $O(x^3)$ は「 x が 0 に近づくとき、 x^3 と同じ速さで 0 になる微小量」を指します。

Case B: 無限大極限 ($x \rightarrow \infty$) ~アルゴリズム・計算量理論~

「主要な項（次数の高い項）」が支配し、次数の低い項や定数倍は無視されます。

- $T(n) = 3n^2 + 100n + 500 = O(n^2)$
- 意味：「データ数 n が巨大になると、計算時間は n^2 に比例して増える（係数や下位の項は気にしない）」
- ここで $O(n^2)$ は「 n が増えるとき、増加のペースは n^2 を超えない」ことを指します。

表 4.1 記号の比較 ($x \rightarrow \infty$ のとき)

記号	読み方	意味 ($x \rightarrow \infty$)	例 ($x \rightarrow \infty$)
$f = o(g)$	スモールオー	f は g より圧倒的に遅い	$\log x = o(x)$
$f = O(g)$	ビッグオー	f は g 以下のオーダー	$2x + 10 = O(x)$
$f \sim g$	同値	f と g は漸近的に等しい	$x^2 + x \sim x^2$

コメント 4.3 (実社会のイメージ：ソートアルゴリズムの壁). プログラミングでデータを並べ替えるとき、計算量のオーダーが勝負を分けます。

- バブルソート: $O(n^2)$
- クイックソート: $O(n \log n)$

データ数 n が 100 程度なら差はありませんが、 $n = 100$ 万になると、 n^2 (1 兆回) は宇宙の寿命より長くなりかねませんが、 $n \log n$ (2000 万回) は一瞬で終わります。「無限大の彼方でどうなるか」を見通すのが、この記号の役割です。

4.5 無限級数の収束判定

テイラー展開が元の関数と一致するには、無限の彼方で剩余項 $R_n \rightarrow 0$ となる必要があります。数列の和（級数）が有限の値に収まるか（収束）、無限大に吹っ飛ぶか（発散）を見極める判定法です。

定理 4.2 (ダランベールの判定法 (比判定法)). 正項級数 $\sum a_n$ について、

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = r$$

とする。

1. $r < 1$ ならば、級数は収束する。
2. $r > 1$ ならば、級数は発散する。
3. $r = 1$ ならば、判定不能（別の方針が必要）。

例 4.2 (指數関数の収束半径). $e^x = \sum \frac{x^n}{n!}$ が任意の実数 x で収束することを示せ。

【証明】 $a_n = \frac{|x|^n}{n!}$ とおく (絶対収束を調べる)。比をとると :

$$\frac{a_{n+1}}{a_n} = \frac{|x|^{n+1}}{(n+1)!} \cdot \frac{n!}{|x|^n} = \frac{|x|}{n+1}$$

$n \rightarrow \infty$ のとき、分母が無限大になるため、任意の固定された x について

$$\lim_{n \rightarrow \infty} \frac{|x|}{n+1} = 0 = r$$

常に $r = 0 < 1$ なので、ダランベールの判定法より、どんな x に対してもこの級数は収束する。(証明終)

4.6 第4章のまとめ

- テイラー展開: 関数を「多項式の和」に分解する、解析学の万能ナイフ。
- マクローリン展開: 原点付近での近似。 $e^x, \sin x, \cos x$ の正体を暴く。
- ランダウの記号: 近似の精度（誤差の減り方）を議論する共通言語。
- 収束判定法: 無限和が意味を持つかどうかをチェックする検問所。

ここまでで、1変数関数の微積分（解析学の基礎）は完了です。次章から、いよいよ世界を多次元に拡張します。「多変数関数と偏微分」により、3次元空間の曲面や、高次元データの解析が可能になります。

第 III 部

多変数関数の微分（高次元世界への入り口）

第 5 章

多変数関数と偏微分：多次元空間の傾き

第 4 章までで、1 つの変数 x が変化するときに関数 $f(x)$ がどう動くかを完全に理解しました。しかし、現実世界やデータサイエンスの問題は、変数が 1 つであることなど稀です。

- **経済:** 株価は、金利、為替、企業業績など無数の要因で決まる。
- **AI:** ニューラルネットワークの損失関数は、数億個のパラメータ（重み）を持つ超多変数関数である。

第 5 章では、この多次元世界 (n 次元ユークリッド空間) を解析するための「偏微分」と「全微分」を導入します。

5.1 n 次元ユークリッド空間と位相

まず、舞台となる空間を定義します。

定義 5.1 (n 次元ユークリッド空間 \mathbb{R}^n). 実数を n 個並べた組 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 全体の集合。距離（ユークリッド距離）は次のように定義される：

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

近傍 (Neighborhood) : 点 \mathbf{a} から距離 ϵ 以内にある点の集まり。

$$U_\epsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{a}| < \epsilon\}$$

(1 次元なら区間、2 次元なら円板、3 次元なら球体です)

5.2 多変数関数の極限と連続

2 変数関数 $z = f(x, y)$ の極限 $\lim_{(x,y) \rightarrow (a,b)} f(x, y) = A$ を定義しますが、ここで 1 変数の時とは決定的に違う恐ろしい点があります。

「近づき方が無限にある」1 次元なら「右から」か「左から」しかありませんが、2 次元では「直線で」「放物線を描いて」「螺旋を描いて」など、あらゆる方向からの近づき方において、同じ値 A に収束しなければなりません。

例 5.1 (極限が存在しない例). $f(x, y) = \frac{xy}{x^2+y^2}$ について、原点 $(0, 0)$ での極限を調べよ。

【解答】

1. x 軸に沿って近づく ($y = 0, x \rightarrow 0$) : $f(x, 0) = \frac{x \cdot 0}{x^2+0} = 0 \rightarrow 0$
2. 直線 $y = x$ に沿って近づく ($y = x, x \rightarrow 0$) : $f(x, x) = \frac{x \cdot x}{x^2+x^2} = \frac{x^2}{2x^2} = \frac{1}{2} \rightarrow \frac{1}{2}$

近づく方向によって値が違う ($0 \neq 1/2$) ため、極限は存在しない。(原点付近のグラフは、方向によって高さが違う複雑な形状をしています)

5.3 偏微分 (Partial Differentiation)

多次元の変化を一気に考えるのは難しいので、「とりあえず1つの変数だけ動かして、他は固定する」という戦略をとります。

定義 5.2 (偏微分係数). 2変数関数 $f(x, y)$ において、 y を定数みなして x だけで微分したものを、 x に関する偏微分といい、 $\frac{\partial f}{\partial x}$ や f_x と書く。

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h}$$

同様に、 y に関する偏微分は：

$$\frac{\partial f}{\partial y}(a, b) = \lim_{k \rightarrow 0} \frac{f(a, b+k) - f(a, b)}{k}$$

コメント 5.1 (実社会のイメージ：A/B テストと感度分析). Web マーケティングで、売上 (z) が「広告費 (x)」と「サイトのデザイン (y)」で決まると言います。

- $\frac{\partial f}{\partial x}$: デザインを変えずに、広告費だけを増やしたときの売上の伸び率。
- $\frac{\partial f}{\partial y}$: 広告費を変えずに、デザインだけを変えたときの売上の変化。

現実にはこれらが絡み合っていますが、分析の第一歩として「他を固定 (ceteris paribus)」して考えるのが偏微分です。

例 5.2 (偏微分の計算). $f(x, y) = x^3 + 2xy^2 - y^4$ の偏導関数を求めよ。

【解答】

1. x で偏微分 (y は数字扱い) : $\frac{\partial f}{\partial x} = 3x^2 + 2(1)y^2 - 0 = 3x^2 + 2y^2$
2. y で偏微分 (x は数字扱い) : $\frac{\partial f}{\partial y} = 0 + 2x(2y) - 4y^3 = 4xy - 4y^3$

5.4 全微分 (Total Differential) ~真の微分可能性~

ここが多変数解析の最大の落とし穴であり、最も重要な概念です。「偏微分可能（各軸方向の傾きがある）だからといって、関数が滑らか（微分可能）とは限らない」のです。

グラフがカクカクしていても、たまたま軸方向だけ滑らかに見えることがあります。真に「滑らか（接平面が存在する）」であるためには、全微分可能性が必要です。

定義 5.3 (全微分可能性). 関数 $z = f(x, y)$ が点 (a, b) で全微分可能であるとは、定数 A, B が存在して、増分 Δz が次のように書けることである。

$$f(a+h, b+k) - f(a, b) = Ah + Bk + o(\sqrt{h^2 + k^2})$$

ここで $A = f_x(a, b), B = f_y(a, b)$ である。

直感的意味: 「曲面 $z = f(x, y)$ が、点 (a, b) の周りで平面 $z - f(a, b) = A(x - a) + B(y - b)$ で近似できる」ということ。この平面を接平面 (Tangent Plane) と呼びます。

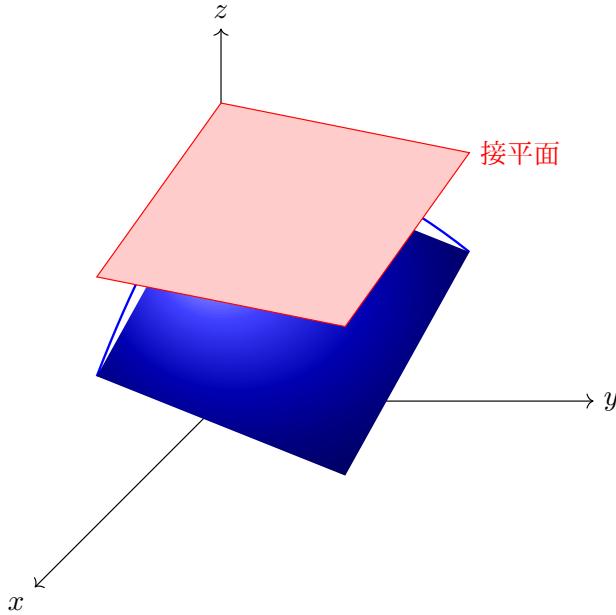


図 5.1 全微分のイメージ：曲面の局所的な挙動を「平面（接平面）」で近似する。

全微分のありがたみ～なぜ偏微分だけではダメなのか？～

偏微分の限界：偏微分 f_x, f_y は、東西南北（十字方向）の傾きしか見ていません。極端な例ですが、「十字方向だけは滑らかな坂道だけど、斜め方向には崖がある」ような関数も作れてしまいます。そんな危険な関数では、最適化（勾配降下法）で斜めに進もうとした瞬間に計算が破綻します。

全微分の恩恵：全微分可能であれば、「どの方向に進んでも、その近傍は平面（接平面）で近似できる」ことが保証されます。この「全方向への安心感」こそが全微分の価値です。しかし、全微分可能かどうかを定義通り確認するのは大変です。そこで、次の定理が「安心保証書」として活躍します。

定理 5.1 (偏微分の連続性と全微分可能性)。偏導関数 f_x, f_y が点 (a, b) の近傍で存在し、かつ (a, b) で連続ならば、 f は (a, b) で全微分可能である。

Proof. 目標：増分 $\Delta f = f(a + h, b + k) - f(a, b)$ が、以下の形に変形できることを示す。

$$\Delta f = f_x(a, b)h + f_y(a, b)k + \epsilon_1 h + \epsilon_2 k$$

(ただし $h, k \rightarrow 0$ のとき $\epsilon_1, \epsilon_2 \rightarrow 0$)

Step 1: 増分を 2 段階に分解する点 (a, b) から $(a + h, b + k)$ への移動を、1. $(a, b) \rightarrow (a + h, b)$ (x だけ動く) 2. $(a + h, b) \rightarrow (a + h, b + k)$ (y だけ動く) という 2 ステップに分けます。

$$\Delta f = [f(a + h, b) - f(a, b)] + [f(a + h, b + k) - f(a + h, b)]$$

Step 2: 平均値の定理を適用するそれぞれのカッコ内は「1 变数関数の変化」と見なせるので、平均値の定理が使えます。

- 第 1 項 (x の変化)：ある $\theta_1 \in (0, 1)$ が存在して、

$$f(a + h, b) - f(a, b) = f_x(a + \theta_1 h, b) \cdot h$$

- 第 2 項 (y の変化)：ある $\theta_2 \in (0, 1)$ が存在して、

$$f(a + h, b + k) - f(a + h, b) = f_y(a + h, b + \theta_2 k) \cdot k$$

よって、

$$\Delta f = f_x(a + \theta_1 h, b)h + f_y(a + h, b + \theta_2 k)k$$

Step 3: 連續性を使って誤差を評価するここで、仮定「 f_x, f_y は (a, b) で連続」を使います。連続なので、 $h, k \rightarrow 0$ のとき、

$$f_x(a + \theta_1 h, b) \rightarrow f_x(a, b), \quad f_y(a + h, b + \theta_2 k) \rightarrow f_y(a, b)$$

となります。その差（誤差）を ϵ_1, ϵ_2 と置きます。

$$\epsilon_1 = f_x(a + \theta_1 h, b) - f_x(a, b), \quad \epsilon_2 = f_y(a + h, b + \theta_2 k) - f_y(a, b)$$

これらは $(h, k) \rightarrow (0, 0)$ で 0 に収束します。これらを Step 2 の式に代入すると：

$$\begin{aligned} \Delta f &= (f_x(a, b) + \epsilon_1)h + (f_y(a, b) + \epsilon_2)k \\ \Delta f &= f_x(a, b)h + f_y(a, b)k + (\epsilon_1 h + \epsilon_2 k) \end{aligned}$$

結論: 残りの項 $\epsilon_1 h + \epsilon_2 k$ は、 h, k に対して高次の微小量 ($o(\sqrt{h^2 + k^2})$) です。なぜなら、極座標 $h = r \cos \theta, k = r \sin \theta$ を考えると

$$\frac{|\epsilon_1 h + \epsilon_2 k|}{\sqrt{h^2 + k^2}} \leq |\epsilon_1| |\cos \theta| + |\epsilon_2| |\sin \theta| \leq |\epsilon_1| + |\epsilon_2| \rightarrow 0$$

となるからです。よって、全微分可能であることが証明されました。□

コメント 5.2 (実社会のイメージ：勾配 (Gradient) と最適化). ベクトル解析の記法を使うと、全微分の係数 (f_x, f_y) を並べたものを勾配ベクトル (**Gradient**) と呼び、 ∇f (ナブラエフ) と書きます。

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

このベクトルは、「関数が最も急激に増加する方向」を向いています。機械学習の「勾配降下法」は、この ∇f の逆方向（最も下る方向）にパラメータを少しづつ動かすことで、損失関数の谷底（最適解）を探すアルゴリズムです。

5.5 合成関数の微分法（連鎖律：Chain Rule）

ニューラルネットワークの学習アルゴリズム（バックプロパゲーション）の数学的原理です。多変数の場合、変数の依存関係が複雑になります。

定理 5.2 (多変数の連鎖律). $z = f(x, y)$ が全微分可能で、 $x = x(t), y = y(t)$ が微分可能ならば、合成関数 $z(t) = f(x(t), y(t))$ の微分は次となる：

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

Proof. t が微小量 Δt だけ変化したときの、 x, y, z の変化を考えます。 x の増分を $\Delta x = x(t + \Delta t) - x(t)$ 、 y の増分を $\Delta y = y(t + \Delta t) - y(t)$ 、 z の増分を $\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$ と置きます。

Step 1: 全微分の定義を使う f は全微分可能なので、増分 Δz を以下のように分解できます。

$$\Delta z = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

ここで、 $(\Delta x, \Delta y) \rightarrow (0, 0)$ のとき $\epsilon_1, \epsilon_2 \rightarrow 0$ です。

Step 2: 両辺を Δt で割る求めたいのは $\frac{dz}{dt}$ なので、微小時間 Δt での比率を見ます。

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \frac{\Delta y}{\Delta t} + \epsilon_1 \frac{\Delta x}{\Delta t} + \epsilon_2 \frac{\Delta y}{\Delta t}$$

Step 3: 極限 $\Delta t \rightarrow 0$ をとることで $\Delta t \rightarrow 0$ とすると、以下のことが起こります。

1. 微分係数への収束: $\frac{\Delta z}{\Delta t} \rightarrow \frac{dz}{dt}$, $\frac{\Delta x}{\Delta t} \rightarrow \frac{dx}{dt}$, $\frac{\Delta y}{\Delta t} \rightarrow \frac{dy}{dt}$
2. 誤差項の消滅: $\Delta t \rightarrow 0$ のとき、 $\Delta x, \Delta y \rightarrow 0$ なので、 $\epsilon_1, \epsilon_2 \rightarrow 0$ となります。一方で $\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}$ は有限の値（速度）に収束します。よって、後ろの2項 ($\epsilon_1 \frac{dx}{dt}$ など) は「 $0 \times \text{有限値}$ 」となり、消滅します。

Step 4: 結論残った項を整理すると：

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

となり、連鎖律が導かれました。 \square

イメージ（影響の伝播）： t が変化すると $\rightarrow x$ と y が同時に変化し \rightarrow その両方の変化が合わさって z が変化する。「 x 経由の影響」と「 y 経由の影響」を足し合わせるのが多変数のチェインルールです。

例 5.3 (連鎖律の計算). $z = x^2y$, $x = \cos t$, $y = \sin t$ のとき、 $\frac{dz}{dt}$ を求めよ。

【解答】公式通りに計算する。

1. 各パートの準備 : $\frac{\partial z}{\partial x} = 2xy$, $\frac{\partial z}{\partial y} = x^2$, $\frac{dx}{dt} = -\sin t$, $\frac{dy}{dt} = \cos t$

2. 組み合わせる :

$$\frac{dz}{dt} = (2xy)(-\sin t) + (x^2)(\cos t)$$

t だけの式に戻す (x, y を代入) : $= 2 \cos t \sin t (-\sin t) + \cos^2 t \cos t = -2 \cos t \sin^2 t + \cos^3 t$

(検算：直接代入すると $z = \cos^2 t \sin t$ 。これを t で微分しても同じ結果になります)

5.6 第5章のまとめ

- 偏微分：多変数関数の「ある一方向」の変化を見る。
- 全微分：多変数関数を「接平面」で近似する（真の線形近似）。
- 勾配 (Gradient) : 最急登坂方向を示すベクトル。最適化の羅針盤。
- 連鎖律：複数の経路で伝わる変化の影響を足し合わせるルール。

これで「多次元の坂道を歩く」準備ができました。次章では、この坂道の頂上や谷底を見つける「多変数関数の極値問題」を扱います。ラグランジュの未定乗数法など、経済学や AI で頻出のトピックが登場します。

第5章補足：全微分の実用と応用～誤差はどう伝わるか～

全微分 $dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ は、単なる定義式ではありません。これは「入力のズレが、出力にどれだけ響くか」を見積もるための最強の実践ツールです。

近似計算への応用

全微分の式を、微小変化量 $\Delta x, \Delta y$ を用いた近似式に書き換えます。

$$f(a + \Delta x, b + \Delta y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y$$

これは「関数値を毎回真面目に計算し直さなくても、微分係数（感度）を使えば手計算で予測できる」ことを意味します。

例 5.4 (抵抗値の誤差見積もり). 抵抗 $R_1 = 100\Omega$, $R_2 = 200\Omega$ の並列回路がある。抵抗値の測定にそれぞれ 1% の誤差が生じる可能性があるとき、合成抵抗 R の誤差は何%程度になるか見積もれ。

【解答】合成抵抗の式は $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$ より、 $R(R_1, R_2) = \frac{R_1 R_2}{R_1 + R_2}$ 。まず、真値を計算する： $R = \frac{100 \cdot 200}{300} = \frac{200}{3} \approx 66.67\Omega$

次に全微分をとる（対数微分法を使うと楽）： $\ln R = \ln R_1 + \ln R_2 - \ln(R_1 + R_2)$ 両辺を全微分すると、

$$\begin{aligned}\frac{dR}{R} &= \frac{dR_1}{R_1} + \frac{dR_2}{R_2} - \frac{d(R_1 + R_2)}{R_1 + R_2} \\ \frac{dR}{R} &= \frac{dR_1}{R_1} + \frac{dR_2}{R_2} - \frac{dR_1 + dR_2}{R_1 + R_2}\end{aligned}$$

係数を整理する：

$$\begin{aligned}\frac{dR}{R} &= \left(\frac{1}{R_1} - \frac{1}{R_1 + R_2} \right) dR_1 + \left(\frac{1}{R_2} - \frac{1}{R_1 + R_2} \right) dR_2 \\ \frac{dR}{R} &= \frac{R_2}{R_1(R_1 + R_2)} dR_1 + \frac{R_1}{R_2(R_1 + R_2)} dR_2\end{aligned}$$

これを「相対誤差 (%/100)」の形 $\frac{dR}{R}$ に変形する。 $dR_1/R_1 = 0.01, dR_2/R_2 = 0.01$ （1% の誤差）を代入すると、

$$\frac{\Delta R}{R} \approx \frac{R_2}{R_1 + R_2} \left(\frac{\Delta R_1}{R_1} \right) + \frac{R_1}{R_1 + R_2} \left(\frac{\Delta R_2}{R_2} \right)$$

数値を代入：係数 $C_1 = \frac{200}{300} = \frac{2}{3}, C_2 = \frac{100}{300} = \frac{1}{3}$

$$\frac{\Delta R}{R} \approx \frac{2}{3}(0.01) + \frac{1}{3}(0.01) = 0.01$$

結論：並列回路では、各抵抗が 1% ズレても、合成抵抗のズレも 1% 程度に収まる。（全微分を使えば、再計算せずにこの特性がわかります）

完全微分形式 (Exact Differential Forms)

全微分に関連して、物理（熱力学）や微分方程式で重要な概念も補足します。

形式 $P(x, y)dx + Q(x, y)dy$ が、ある関数 $\phi(x, y)$ の全微分 $d\phi = \phi_x dx + \phi_y dy$ になっているとき、これを完全微分形式といいます。

【判定条件】 $Pdx + Qdy$ が完全微分であるための必要十分条件は（領域が単連結なら）：

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

これは第9章で登場する「保存場の条件（回転が0）」と同じものです。

コメント 5.3 (実社会のイメージ：状態量)。熱力学において、仕事 W や熱量 Q は完全微分ではありません（経路によって値が変わる）。しかし、エントロピー S や内部エネルギー U は完全微分（状態量）です。「今の状態だけで値が決まるのか、そこに至る履歴に依存するのか」を見分ける数学的判定法がこれです。

第 6 章

多変数微分の応用：最適解を求めて

第 5 章で手に入れた「偏微分」と「勾配」という武器を使って、いよいよ実利的な問題、すなわち「最適化（Optimization）」に挑みます。「利益を最大化したい」「コストを最小化したい」「誤差をゼロにしたい」——これらはすべて、多変数関数の極値問題に帰着されます。

ここでは、単純な極値問題だけでなく、統計学や機械学習で必須となる「制約条件付き最適化問題（ラグランジュの未定乗数法）」までを、幾何学的な直感を交えて解説します。

6.1 多変数関数の極値判定

1 変数のときは $f'(x) = 0$ で候補を探し、 $f''(x)$ の符号で山か谷かを判定しました。2 変数関数 $f(x, y)$ でも基本は同じですが、「鞍点（Saddle Point）」という厄介な地形が登場します。

定義 6.1 (極大・極小). 関数 $f(x, y)$ が点 (a, b) のある近傍内のすべての点 (x, y) に対して、

- $f(x, y) \leq f(a, b)$ ならば、 (a, b) で極大であるという。
- $f(x, y) \geq f(a, b)$ ならば、 (a, b) で極小であるという。

定理 6.1 (極値の必要条件). 偏微分可能な関数 $f(x, y)$ が点 (a, b) で極値をとるならば、勾配はゼロになる。

$$f_x(a, b) = 0 \quad \text{かつ} \quad f_y(a, b) = 0$$

この条件を満たす点 (a, b) を停留点（Stationary Point）と呼ぶ。

※注意：停留点だからといって極値とは限りません（1 変数の $y = x^3$ の $x = 0$ と同じです）。

定理 6.2 (極値の十分条件 (ヘッセ行列式による判定)). f が C^2 級（2 回偏微分可能で連続）とする。停留点 (a, b) において、以下の行列式（ヘッシアン）を考える。

$$H = \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = f_{xx}f_{yy} - (f_{xy})^2$$

(ただし $f_{xy} = f_{yx}$ はシュワルツの定理より成立)

1. $H > 0$ かつ $f_{xx} > 0 \implies$ 極小値（お椀の底）
2. $H > 0$ かつ $f_{xx} < 0 \implies$ 極大値（山の頂上）
3. $H < 0 \implies$ 鞍点（馬の背：ある方向には山、ある方向には谷。極値ではない）
4. $H = 0 \implies$ 判定不能（高次の項を調べる必要がある）

直感的意味: f_{xx} は x 方向の曲がり具合、 f_{yy} は y 方向の曲がり具合です。 $H > 0$ は「曲がり具合の符号が全方向で一致している（凸または凹）」ことを示し、 $H < 0$ は「方向によって曲がり方が逆（サドル）」であることを示します。

コメント 6.1 (実社会のイメージ：敵対的生成ネットワーク (GAN))。AI 画像生成技術である GAN の学習は、生成 AI (Generator) と識別 AI (Discriminator) のいたちごっこです。損失関数において、生成 AI は値を下げようとし（極小化）、識別 AI は上げようとします（極大化）。この均衡点こそが鞍点（Saddle Point）です。単純な勾配法では鞍点付近で学習が停滞してしまうため、AI 研究において鞍点の攻略は重要なテーマです。

例 6.1 (極値の判定). $f(x, y) = x^3 + y^3 - 3xy$ の極値を求めよ。

【解答】

1. 勾配 = 0 を解く: $f_x = 3x^2 - 3y = 0 \Rightarrow y = x^2$ $f_y = 3y^2 - 3x = 0 \Rightarrow x = y^2$ これらを連立させると、 $x = (x^2)^2 = x^4 \Rightarrow x(x^3 - 1) = 0$ 。実数解は $x = 0, 1$ 。対応する y は $y = 0, 1$ 。よって停留点は $(0, 0)$ と $(1, 1)$ の 2 つ。
2. ヘッシアンを計算: $f_{xx} = 6x$, $f_{yy} = 6y$, $f_{xy} = -3$ $H(x, y) = (6x)(6y) - (-3)^2 = 36xy - 9$
3. 各点の判定:
 - 点 $(0, 0)$: $H(0, 0) = -9 < 0$ よって鞍点である（極値ではない）。
 - 点 $(1, 1)$: $H(1, 1) = 36 - 9 = 27 > 0$ かつ $f_{xx}(1, 1) = 6 > 0$ （下に凸）。よって極小値をとる。極小値は $f(1, 1) = 1 + 1 - 3 = -1$ 。

6.2 ラグランジュの未定乗数法 (Lagrange Multipliers)

「予算 1 万円以内で（制約）、満足度を最大化せよ（目的）」現実の問題は常に制約付きです。これを魔法のように解く手法です。

定理 6.3 (ラグランジュの未定乗数法). 目的関数 $f(x, y)$ を、制約条件 $g(x, y) = 0$ の下で最大化・最小化したい。極値を与える点 (a, b) では、ある定数 λ (ラグランジュ乗数) が存在して、以下が成り立つ（ただし $\nabla g \neq 0$ ）。

$$\nabla f(a, b) = \lambda \nabla g(a, b) \quad \text{かつ} \quad g(a, b) = 0$$

成分で書くと：

$$\begin{cases} f_x = \lambda g_x \\ f_y = \lambda g_y \\ g(x, y) = 0 \end{cases}$$

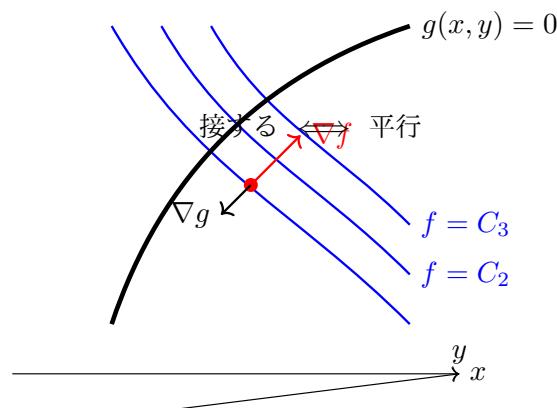


図 6.1 ラグランジュの未定乗数法：等高線（青）と制約曲線（黒）が接する点で極値をとる。このとき勾配ベクトル（矢印）は平行になる。

ラグランジュの未定乗数法：もっとわかりやすい解説

1. イメージ：「登山とフェンス」

あなたは山登りをしています。山の高さが「目的関数 $f(x, y)$ 」です。しかし、山には「立ち入り禁止のフェンス（制約条件 $g(x, y) = 0$ ）」があり、フェンスの上しか歩けません。

質問：フェンスの上を歩くとき、一番高い地点はどこでしょうか？

- 状況 A：等高線とフェンスが交差しているときフェンスが等高線を横切っているなら、フェンス沿いに少し進めば、もっと高い等高線へ移動できます。つまり、そこは頂上（最大点）ではありません。
- 状況 B：等高線とフェンスが接しているときフェンス沿いに進んでも、もう高い等高線には移れません。ここが「フェンス上の頂上」です！

数学的な結論：「2つの曲線（等高線とフェンス）が接する」ということは、その点での「法線ベクトル（勾配）」が同じ向き（または逆向き）ということです。

$$\nabla f \parallel \nabla g$$

これを式にしたのが、 $\nabla f = \lambda \nabla g$ なのです。（ λ はベクトルの長さを合わせるための調整係数です）

2. ラムダ λ の正体：「制約の値段」

λ は単なる調整係数ではありません。実は「もし制約を少し緩めたら、どれだけ得をするか」を表す重要な指標です（経済学では「シャドープライス」と呼びます）。

具体例：予算 1 万円でお菓子を買う

- 目的 f ：満足度を最大化する
- 制約 g ：合計金額 = 10000 円

計算の結果、最大満足度が得られ、そのときの $\lambda = 5$ だったとします。これは、「もし予算を 1 円増やして 10001 円にすれば、満足度が約 5 ポイント上がる」ということを意味します。

- λ が大きい → 制約がキツイ（予算を増やせば満足度が激増する）。
- λ が小さい → 制約は緩い（予算を増やしてもあまり変わらない）。
- $\lambda = 0$ → 制約は無意味（もう十分満足している、あるいは欲しいものがない）。

この解説により、 $\nabla f = \lambda \nabla g$ という式が、「ベクトルが平行である（幾何学的意味）」と「制約の強さ（経済学的意味）」の両方を表していることが深く理解できます。

例 6.2 (制約付き最大化). 制約条件 $x^2 + y^2 = 1$ (単位円上) の下で、関数 $f(x, y) = x + y$ の最大値・最小値を求めよ。

【解答】

1. ラグランジュ関数: $g(x, y) = x^2 + y^2 - 1 = 0$ 条件式: $\nabla f = (1, 1)$ $\nabla g = (2x, 2y)$ $\nabla f = \lambda \nabla g \implies$

$$\begin{cases} 1 = 2\lambda x & \dots (1) \\ 1 = 2\lambda y & \dots (2) \\ x^2 + y^2 = 1 & \dots (3) \end{cases}$$
2. 連立方程式を解く: (1)(2) より、 λ が 0 だと $1 = 0$ となり矛盾するので $\lambda \neq 0$ 。よって $x = \frac{1}{2\lambda}$, $y = \frac{1}{2\lambda}$ 。つまり $x = y$ 。 (3) に代入: $x^2 + y^2 = 1 \implies 2x^2 = 1 \implies x = \pm \frac{1}{\sqrt{2}}$ 。よって候補点は $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ と $(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ 。

3. 値を比較: $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{2}{\sqrt{2}} = \sqrt{2}$ (最大値) $f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\sqrt{2}$ (最小値)

6.3 【発展】不等式制約と KKT 条件～壁に寄りかかるか、部屋の中にいるか～

これまで「等式制約 $g(x, y) = 0$ (フェンスの上)」だけを考えましたが、現実には「不等式制約 $g(x, y) \leq 0$ (フェンスの内側ならどこでも OK)」の場合が多いです。

このとき、最適解（一番高い場所）は 2 つのパターンのどちらかになります。

6.3.1 2 つのパターン

- パターン 1: 制約が「効いていない」場合（内部解）頂上がフェンスの内側 ($g < 0$) にある場合です。このとき、フェンスは無関係なので、制約なしの極値問題と同じです。

$$\nabla f = 0, \quad \lambda = 0 \quad (\text{制約の値段は } 0 \text{ 円})$$

- パターン 2: 制約が「効いている」場合（境界解）頂上がフェンスの外にあり、行きたくても行けないため、フェンス際 ($g = 0$) で我慢する場合です。このとき、等式制約と同じ状態になります。

$$\nabla f = \lambda \nabla g, \quad \lambda > 0 \quad (\text{制約を緩めればもっと登れる})$$

6.3.2 KKT 条件 (Karush-Kuhn-Tucker 条件)

これら 2 つのパターンを、「 λ と g の積がゼロになる」という 1 つの式でまとめたのが KKT 条件です。

定理 6.4 (KKT 条件). 不等式制約 $g(\mathbf{x}) \leq 0$ の下で $f(\mathbf{x})$ を最大化する問題の必要条件：

1. 勾配条件: $\nabla f = \lambda \nabla g$ (力の釣り合い)
2. 制約条件: $g(\mathbf{x}) \leq 0$ (ルールを守る)
3. 非負条件: $\lambda \geq 0$ (制約の価値はマイナスにはならない)
4. 相補性条件: $\lambda \cdot g(\mathbf{x}) = 0$ (ここが最重要！)

相補性条件の意味：「もし壁に寄りかかっている ($g = 0$) なら、壁を押す力が発生する ($\lambda > 0$)」。「もし壁から離れている ($g < 0$) なら、壁を押す力はゼロである ($\lambda = 0$)」。つまり、「位置」か「力」のどちらかは必ずゼロになる、という条件です。

例 6.3 (不等式制約付き最適化). 制約条件 $g(x, y) = x^2 + y^2 - 1 \leq 0$ の下で、目的関数 $f(x, y) = 2x + y$ の最大値を求めよ。

【解答】KKT 条件より：1. $\nabla f = \lambda \nabla g \implies (2, 1) = \lambda(2x, 2y)$ 2. $x^2 + y^2 - 1 \leq 0$ 3. $\lambda \geq 0$ 4. $\lambda(x^2 + y^2 - 1) = 0$

相補性条件でケース分けを行う。

- **Case A:** $\lambda = 0$ (内部解) 条件 1 より $2 = 0$ となり矛盾。よってこのケースはない。
- **Case B:** $x^2 + y^2 - 1 = 0$ (境界解) 条件 1 より $x = \frac{1}{\lambda}, y = \frac{1}{2\lambda}$ 。境界式に代入して $\lambda^2 = \frac{5}{4} \implies \lambda = \frac{\sqrt{5}}{2}$ ($\lambda \geq 0$ より)。よって $x = \frac{2}{\sqrt{5}}, y = \frac{1}{\sqrt{5}}$ 。このとき最大値は $f = \sqrt{5}$ 。

6.3.3 KKT 条件の導出 (幾何学的証明)

Proof. 最大点 \mathbf{x}^* がどこにあるかで場合分けする。

1. 内部の場合 ($g(\mathbf{x}^*) < 0$): 制約なしの極値問題と同じなので $\nabla f = 0$ 。これは $\lambda = 0$ とすれば $\nabla f = \lambda \nabla g$ を満たす。
2. 境界の場合 ($g(\mathbf{x}^*) = 0$): \mathbf{x}^* から領域内（内側）へ動くと、値が下がらなければならない（最大値だから）。つまり、勾配 ∇f （登り方向）は「領域の外側」を向いている必要がある。一方、制約の勾配 ∇g も定義より「領域の外側 (g が増える方向)」を向いている。よって 2 つのベクトルは同じ向きなので、 $\nabla f = \lambda \nabla g$ （ただし $\lambda \geq 0$ ）と書ける。

いずれの場合も $\lambda g(\mathbf{x}^*) = 0$ が成り立つ。 \square

コラム：ヤコビアンとヘシアン ~多次元の「1 次近似」と「2 次近似」~

「ヘッセ行列 (Hessian)」や「ヤコビ行列 (Jacobian)」は、高校数学の言葉で言えば、それぞれ「放物線近似（2 次関数）」と「接線近似（1 次関数）」の多変数版です。

1. ヤコビアン (Jacobian) ~多次元の「接線の傾き」~

高校数学: 1 変数関数 $y = f(x)$ の接線の傾きは $f'(x)$ です。これは、「 x がちょっと動くと、 y は $f'(x)$ 倍動く ($dy \approx f'(x)dx$)」という拡大率を表します。

大学数学（ヤコビアン）: 多変数関数（座標変換など） $\mathbf{y} = \mathbf{f}(\mathbf{x})$ の場合、「 \mathbf{x} がちょっと動くと、 \mathbf{y} はどう動くか？」を表すのがヤコビ行列 J です。

$$d\mathbf{y} \approx J d\mathbf{x}$$

この行列式（ヤコビアン）は、「局所的な面積（体積）の拡大率」を表します。積分で変数変換するときに、「升目の面積がどれくらい変わったか」を補正するために必要になります。

2. ヘシアン (Hessian) ~多次元の「凹凸（曲がり具合）」~

高校数学: 関数のグラフが「下に凸（谷）」か「上に凸（山）」かを調べるには、2 階微分 $f''(x)$ の符号を見ました。

- $f''(x) > 0 \implies$ 下に凸（極小の候補）
- $f''(x) < 0 \implies$ 上に凸（極大の候補）

これは、関数を放物線 $y = ax^2$ （2 次関数）で近似していることに相当します。

大学数学（ヘシアン）: 多変数関数 $f(x, y)$ の曲がり具合を調べるのがヘッセ行列 H です。

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

この行列の性質（固有値の符号など）を見ることで、

- 全部プラス \implies お椀型（極小）
- 全部マイナス \implies 山型（極大）
- プラスとマイナスが混在 \implies 馬の鞍型（鞍点）

を判定できます。つまり、多次元の放物面近似を行っているのです。

3. なぜ「3 次以上」はあまり登場しないのか？

理由は単純で、「2 次までで用が足りることが多いから」です。

1. 最適化（極値問題）のゴール：私たちの大半が知りたいのは、「どこが頂上（最大値）か？」「どこが谷底（最小値）か？」です。

- 1階微分（ヤコビアン・勾配） $= 0 \rightarrow$ 「平らな場所」を探す。
- 2階微分（ヘシアン） \rightarrow そこが「山か谷か」を判定する。

これだけで目的（最適解の発見）は達成されます。

2. 3次の役割：3階微分は「曲がり具合の変化率（非対称性や歪み）」を表します。統計学では「歪度（Skewness）」として登場しますが、最適化や物理法則の記述（運動方程式など）においては、2次（加速度やエネルギー）までの情報で記述できる現象が圧倒的に多いため、主役になることは稀です。

結論：

- ヤコビアン：1次近似（線形・拡大率）
- ヘシアン：2次近似（放物線・凹凸）
- 3次以降：通常の用途では「微細な補正」に過ぎず、コスト対効果が薄い。

6.4 第6章のまとめ

- ヘッセ行列：2階微分の行列。多次元空間でのグラフの凹凸（山、谷、鞍）を判定する。
- 鞍点：極値のふりをした偽物。高次元最適化の敵。
- ラグランジュの未定乗数法：「勾配が平行になる」という幾何学的条件を使って、制約付き問題を解く強力な手法。

ここまでで、微分学（微小変化の解析）の主要なトピックは網羅しました。第7章からは、いよいよ「積分学（集積量の解析）」を多次元へ拡張します。面積、体積、そして確率分布の計算へつながる「重積分」の世界へ進みます。

第 IV 部

多変数関数の積分（重積分）

第 7 章

重積分：体積と総和の計算

第 6 章までで、多変数関数の「局所的な変化（微分）」を解析し、最適解を見つける手法を学びました。第 7 章からは、逆の操作である「大域的な集積（積分）」に入ります。

1 変数の積分 $\int f(x)dx$ は「曲線の下の面積」でしたが、2 変数の重積分 $\iint f(x, y)dxdy$ は「曲面の下の体積」を表します。さらに高次元になれば、それは「質量」「電荷」「確率」といった物理量や統計量の総和を計算することになります。統計学における「確率密度関数から確率を求める計算」は、まさにこの重積分そのものです。

7.1 重積分の定義

長方形領域 $D = [a, b] \times [c, d]$ 上で定義された関数 $f(x, y) \geq 0$ を考えます。この関数がつくる曲面と xy 平面に挟まれた立体の体積を定義します。

定義 7.1 (重積分 (リーマン和による定義)). 領域 D を細かい小長方形 ΔD_{ij} に分割し、その面積を ΔS_{ij} 、区間内の代表点を (ξ_{ij}, η_{ij}) とします。このとき、直方体の体積の総和 (リーマン和) の極限を重積分と定義します。

$$\iint_D f(x, y) dx dy = \lim_{n \rightarrow \infty} \sum_{i,j} f(\xi_{ij}, \eta_{ij}) \Delta S_{ij}$$

(※分割の最大径を 0 に近づけたとき、この値が収束する場合、「積分可能」といいます)

コメント 7.1 (実社会のイメージ：デジタル画像の画素集計). デジタル画像はピクセル（小領域）の集合です。各ピクセルの座標 (i, j) における輝度を $f(i, j)$ とすると、画像全体の明るさの総和は $\sum f(i, j)$ です。これを連続的な世界（アナログ写真や光の強度分布）で考えたものが重積分です。センサーが受け取る総光子数や、ある地域の総雨量を計算する際に使われます。

7.2 累次積分とフビニの定理

定義通りにリーマン和の極限を計算するのは困難です。実際には、「輪切りにして断面積を求め、それを積み重ねる」という方法をとります。これが累次積分です。

定理 7.1 (フビニの定理). $f(x, y)$ が長方形領域 $D = [a, b] \times [c, d]$ で連続ならば、重積分は以下の累次積分と等しい。

$$\iint_D f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

意味: 「まず y で積分 (x を固定して短冊の面積を求める) し、その結果を x で積分 (短冊を並べて体積にする)」しても、順序を逆にしても、結果は同じである。

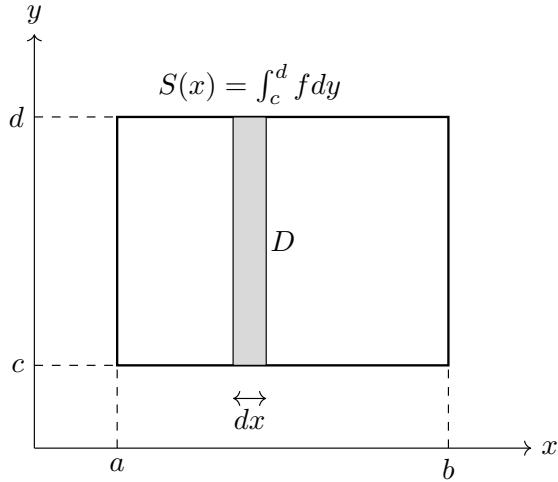


図 7.1 フビニの定理のイメージ：重積分を「断面積 \$S(x)\$ の積分」として計算する（スライス法）。

例 7.1 (累次積分の計算). 領域 \$D = [0, 1] \times [0, 2]\$ における \$f(x, y) = xy^2\$ の重積分を求めよ。

【解答】フビニの定理より、計算しやすい順序で積分する。

$$I = \int_0^1 \left(\int_0^2 xy^2 dy \right) dx$$

まず内側の積分 (\$y\$ について) :

$$\int_0^2 xy^2 dy = \left[\frac{1}{3} xy^3 \right]_{y=0}^{y=2} = \frac{1}{3} x (8) - 0 = \frac{8}{3} x$$

次に外側の積分 (\$x\$ について) :

$$I = \int_0^1 \frac{8}{3} x dx = \left[\frac{8}{3} \cdot \frac{1}{2} x^2 \right]_0^1 = \frac{4}{3} (1 - 0) = \frac{4}{3}$$

(逆に先に \$x\$ で積分しても \$\int_0^2 (\frac{1}{2} y^2) dy = [\frac{1}{6} y^3]_0^2 = \frac{8}{6} = \frac{4}{3}\$ となり一致する)

7.3 積分領域の変更と順序交換

積分領域 \$D\$ が長方形でない場合（三角形や円など）、積分範囲の上限・下限が変数になります。このとき、積分順序を交換するには領域の図示と範囲の書き換えが必要です。

例 7.2 (積分順序の交換). 次の積分の順序を交換せよ。

$$I = \int_0^1 \left(\int_x^1 e^{y^2} dy \right) dx$$

（※注：\$e^{y^2}\$ の不定積分は初等関数では表せないため、このままでは計算不可能。順序交換が必須となる有名な問題です）

【解答】

1. 領域の図示: 内側の積分範囲 : \$x \leq y \leq 1\$ 外側の積分範囲 : \$0 \leq x \leq 1\$ これらを図示すると、\$xy\$ 平面上の「3点 \$(0, 0), (1, 1), (0, 1)\$ を結ぶ三角形領域」となる。
2. 視点の切り替え: この三角形を、先に \$x\$ 方向（横）にスライスすると見る。\$y\$ を固定したとき、\$x\$ が動く範囲は \$0\$ から \$y\$ まで。\$y\$ 全体の範囲は \$0\$ から \$1\$ まで。よって新しい範囲は : \$0 \leq x \leq y, 0 \leq y \leq 1\$

3. 積分の実行:

$$I = \int_0^1 \left(\int_0^y e^{y^2} dx \right) dy$$

内側の積分 (x について): e^{y^2} は x に無関係な定数なので、 $\int_0^y e^{y^2} dx = e^{y^2} [x]_0^y = ye^{y^2}$

外側の積分 (y について): $\int_0^1 ye^{y^2} dy$ 置換積分 $t = y^2$ とおくと、 $dt = 2ydy \implies ydy = \frac{1}{2}dt$ 。範囲は $0 \rightarrow 1$ 。

$$= \int_0^1 e^t \frac{1}{2} dt = \frac{1}{2} [e^t]_0^1 = \frac{1}{2}(e - 1)$$

このように、順序交換によって「計算不能」が「計算可能」になることがあります。

7.4 第7章のまとめ

- **重積分:** 曲面の下の体積。
- **フビニの定理:** 多重積分を「1変数積分の繰り返し」に分解するツール。
- **順序交換:** 積分領域の幾何学的理解が不可欠。計算の難易度を劇的に変えるテクニック。

第8章では、直交座標 (x, y) では計算しづらい積分（円形の領域など）を、極座標 (r, θ) などに変換して解くための「変数変換とヤコビアン」を扱います。ここで有名な「ガウス積分（正規分布の正規化定数）」の証明が登場します。

第 8 章

変数変換とヤコビアン：歪んだ空間の測り方

第 7 章では、積分領域が「長方形」や「単純な形」の場合を扱いました。しかし、現実世界には円、球、あるいはもっと複雑な歪んだ空間があふれています。

直交座標 (x, y) では計算が困難な積分も、「座標系を取り替える（視点を変える）」ことで劇的に簡単になります。その際の「変換の代償（通貨レート）」として登場するのがヤコビアン（Jacobian）です。

これは統計学における確率変数の変換や、ロボット工学における運動制御の基礎となる極めて重要な概念です。

8.1 変数変換の一般論とヤコビアン

積分変数 (x, y) を、新しい変数 (u, v) に変換することを考えます（例： $x = r \cos \theta, y = r \sin \theta$ ）。このとき、微小面積 $dxdy$ は単純に $dudv$ にはなりません。座標の網目が歪んだり、拡大縮小したりするからです。この「局所的な面積の拡大率」がヤコビアンです。

定義 8.1 (ヤコビ行列とヤコビアン). 変換 $x = \varphi(u, v), y = \psi(u, v)$ が与えられたとき、以下の行列をヤコビ行列（Jacobian Matrix）という。

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

この行列の行列式 $\det J$ をヤコビアン（Jacobian Determinant）と呼ぶ（単にヤコビアンと呼ぶときは行列式を指すことが多い）。

定理 8.1 (重積分の変数変換公式). 変換が一対一であり、偏導関数が連続であるとき、

$$\iint_D f(x, y) dx dy = \iint_E f(\varphi(u, v), \psi(u, v)) |\det J| du dv$$

ここで E は uv 平面における D の対応領域である。重要：絶対値 $|\det J|$ を忘れてはいけません（面積に負の値はないため）。

コメント 8.1 (幾何学的証明の直感（微小領域の近似）). uv 平面上の微小な正方形（一辺 du, dv ）が、 xy 平面上でどのような形に移るかを考えます。微分可能であれば、微小領域は線形変換（1次近似）とみなせるため、正方形は「平行四辺形」に移ります。この平行四辺形を張る 2 つのベクトルは、 $\vec{a} \approx (\frac{\partial x}{\partial u} du, \frac{\partial y}{\partial u} du)$ $\vec{b} \approx (\frac{\partial x}{\partial v} dv, \frac{\partial y}{\partial v} dv)$ となります。平行四辺形の面積は、この 2 つのベクトルの外積（2 次元なら行列式）の絶対値で与えられます。

$$\text{Area} = |\det(\vec{a}, \vec{b})| = \left| \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right| du dv = |J| du dv$$

つまり、「 $|J|$ は局所的な面積の拡大倍率」なのです。

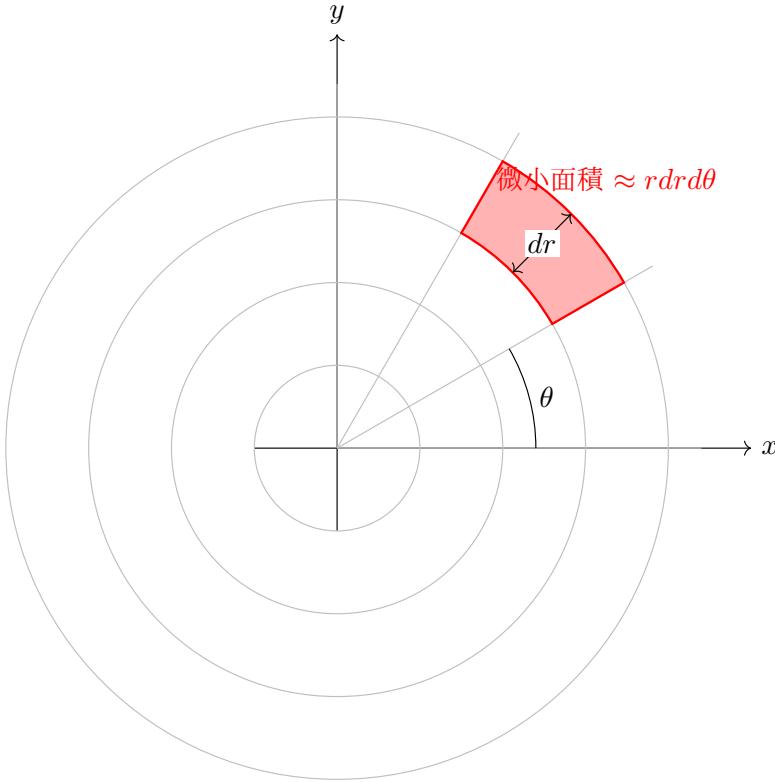


図 8.1 極座標の微小面積：中心から遠いほど (r が大)、同じ角度幅でも弧が長くなるため、面積は r 倍される（ヤコビアン r の意味）。

8.2 極座標変換 (Polar Coordinates)

最も頻繁に使われる変換です。円対称な物理現象や統計分布を扱う際に必須です。

【計算】 極座標のヤコビアン

変換式： $x = r \cos \theta$, $y = r \sin \theta$ ヤコビ行列を計算する：

$$\frac{\partial x}{\partial r} = \cos \theta, \quad \frac{\partial x}{\partial \theta} = -r \sin \theta$$

$$\frac{\partial y}{\partial r} = \sin \theta, \quad \frac{\partial y}{\partial \theta} = r \cos \theta$$

ヤコビアン（行列式）：

$$\det J = (\cos \theta)(r \cos \theta) - (-r \sin \theta)(\sin \theta) = r(\cos^2 \theta + \sin^2 \theta) = r$$

よって、変換公式は以下のようになる：

$$dx dy = r dr d\theta$$

直感： 中心から遠ざかるほど (r が大きいほど)、角度 $d\theta$ の開きによる弧の長さ $rd\theta$ は長くなります。だから微小面積は単なる $drd\theta$ ではなく、 r 倍されるのです。

例 8.1 (原点を中心とする円の面積). 半径 R の円板 $D : x^2 + y^2 \leq R^2$ の面積を重積分で求めよ。

【解答】 面積は関数 $f(x, y) = 1$ の重積分である。領域 D は極座標で見ると、 $0 \leq r \leq R$, $0 \leq \theta \leq 2\pi$ という長方形領域 E になる。

$$\begin{aligned} S &= \iint_D 1 dx dy = \iint_E 1 \cdot r dr d\theta \\ &= \int_0^{2\pi} \left(\int_0^R r dr \right) d\theta = \int_0^{2\pi} \left[\frac{1}{2} r^2 \right]_0^R d\theta \end{aligned}$$

$$= \int_0^{2\pi} \frac{1}{2} R^2 d\theta = \frac{1}{2} R^2 \cdot 2\pi = \pi R^2$$

(小学校からの公式が、解析的に証明されました)

8.3 ガウス積分 (Gaussian Integral) ~統計学の至宝~

「正規分布の全確率が 1 になる」ことを保証する、解析学で最も美しい結果の一つです。 e^{-x^2} の原始関数は初等関数では書けませんが、重積分を経由することで定積分だけが求まるという奇跡的な手法です。

定理 8.2 (ガウス積分).

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

Proof. 求める積分値を I とする。

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx$$

ここで、積分変数を y に変えても値は同じなので、 $I = \int_{-\infty}^{\infty} e^{-y^2} dy$ 。この 2 つを掛け合わせる（ここが最大のアイデア）。

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$$

(フビニの定理の逆を行い、2 重積分にまとめた)

ここで極座標変換を行う。積分領域は全平面 \mathbb{R}^2 なので、 $r : 0 \rightarrow \infty$, $\theta : 0 \rightarrow 2\pi$ 。被積分関数は $e^{-(x^2+y^2)} = e^{-r^2}$ 。微小面積は $dxdy = rdrd\theta$ 。

$$I^2 = \int_0^{2\pi} \left(\int_0^{\infty} e^{-r^2} \cdot r dr \right) d\theta$$

内側の積分 (r について) を計算する。 $t = r^2$ と置換すると、 $dt = 2rdr \Rightarrow rdr = \frac{1}{2}dt$ 。範囲は $0 \rightarrow \infty$ 。

$$\int_0^{\infty} e^{-r^2} r dr = \int_0^{\infty} e^{-t} \frac{1}{2} dt = \frac{1}{2} [-e^{-t}]_0^{\infty} = \frac{1}{2}(0 - (-1)) = \frac{1}{2}$$

(※ヤコビアン由来の r があったおかげで、積分が可能になった！)

これを外側の積分に戻す：

$$I^2 = \int_0^{2\pi} \frac{1}{2} d\theta = \frac{1}{2} \cdot 2\pi = \pi$$

$I > 0$ (被積分関数が正) なので、

$$I = \sqrt{\pi}$$

□

コメント 8.2 (実社会のイメージ：正規分布と偏差値). 統計学における正規分布の確率密度関数は以下の通りです。

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

この係数 $\frac{1}{\sqrt{2\pi\sigma^2}}$ は、全確率を 1 にするための「正規化定数」ですが、その正体はガウス積分の結果 ($\sqrt{\pi}$) から来ているのです。この積分ができなければ、偏差値も、検定も、AI のベイズ推定も成立しません。

8.4 一般の変数変換と応用

例 8.2 (楕円の面積). 楕円領域 $D : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$ の面積を求めよ。

【解答】単位円に変形するような変数変換を行う。 $x = au, y = bv$ とおく。ヤコビ行列は対角行列になる：

$$J = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \implies \det J = ab$$

変換後の領域 E は $u^2 + v^2 \leq 1$ (単位円) となる。

$$S = \iint_D dx dy = \iint_E |ab| du dv = ab \iint_E du dv$$

$\iint_E du dv$ は単位円の面積 π なので、

$$S = ab\pi = \pi ab$$

(円の面積公式 πr^2 の一般化になっています)

コメント 8.3 (実社会のイメージ：ロボットアームの特異点). ロボット工学では、関節角度空間 (θ_1, θ_2) から手先位置座標 (x, y) への変換 f を考えます。手先の速度 v と関節速度 $\dot{\theta}$ の関係は、ヤコビ行列 J を用いて $v = J\dot{\theta}$ と書かれます。もある姿勢で $\det J = 0$ になるとどうなるか？その姿勢（特異点）では、特定の方向に手先を動かすことが不可能になります（腕が伸びきった状態など）。ヤコビアンが 0 になる点は、制御不能に陥る危険な点として、工学的に非常に重要な意味を持ちます。

8.5 第8章のまとめ

- ヤコビアン：座標変換に伴う「面積の拡大率」。行列式で計算される。
- 極座標変換: $dxdy \rightarrow rdrd\theta$ 。円対称な問題のキーツール。
- ガウス積分: $\int e^{-x^2} dx = \sqrt{\pi}$ 。極座標変換を経由して初めて解ける、統計学の魂。

これで、多変数関数の微分と積分（解析学基礎）の主要な山場を越えました。もしさらに続けるならば、

- 第9章：ベクトル解析（場の解析、グリーンの定理、ストークスの定理）
- 第10章：フーリエ解析（波の解析）

などが考えられます。ひとまず、この「第8章」までの構成で、テキストの骨子は完成したと言えます。

第 V 部

ベクトル解析の入り口（オプション）

第 9 章

ベクトル解析：場の流れを読む

第 8 章までで、多変数関数の微積分、特に重積分を用いた「体積」や「重み付き総和」の計算をマスターしました。第 9 章では、これを物理的・幾何学的な「場 (Field)」の世界へ応用します。

電磁気学のマクスウェル方程式や、流体力学のナビエ・ストークス方程式を理解するための言語、それが「ベクトル解析」です。ここでは、勾配 (grad)、発散 (div)、回転 (rot) という 3 つの微分演算子と、それらを結ぶ積分定理 (グリーン、ガウス、ストークス) を学びます。

9.1 スカラー場とベクトル場

定義 9.1 (場 (Field)). 空間 \mathbb{R}^3 の各点 (x, y, z) に対して、ある量を対応させる関数。

- スカラー場 $\phi(x, y, z)$: 各点に「数値」が割り当てられる (例: 温度分布、気圧分布、電位)。
- ベクトル場 $\mathbf{V}(x, y, z)$: 各点に「ベクトル」が割り当てられる (例: 風速分布、電場、磁束密度)。

9.2 ナブラ演算子 ∇ と 3 つの微分

ベクトル微分演算子 ∇ (ナブラ) を次のように定義します。

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

これを「ベクトル」のように扱うことで、3 種類の微分操作を統一的に記述できます。

1. 勾配 (Gradient) : $\nabla\phi$

スカラー場 ϕ に作用し、ベクトル場を作ります。

$$\text{grad } \phi = \nabla\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z} \right)$$

意味: 「スカラー場が最も急激に増加する方向と、その傾斜の大きさ」を表す。(例: 山の高さ分布 ϕ に対し、 $\nabla\phi$ は頂上へ向かう登り坂の方向)

2. 発散 (Divergence) : $\nabla \cdot \mathbf{V}$

ベクトル場 $\mathbf{V} = (V_x, V_y, V_z)$ との内積を取ります。結果はスカラー場になります。

$$\text{div } \mathbf{V} = \nabla \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$$

意味: 「その点からベクトルがどれだけ湧き出しているか (湧き出し量)」を表す。

- 正なら「湧き出し源 (Source)」がある（例：噴水）。
- 負なら「吸い込み口 (Sink)」がある（例：排水溝）。
- 0なら「素通り（湧きもしなければ消えもしない）」している（例：定常流）。

3. 回転 (Rotation / Curl) : $\nabla \times \mathbf{V}$

ベクトル場 \mathbf{V} との外積を取ります。結果はベクトル場になります。

$$\text{rot } \mathbf{V} = \nabla \times \mathbf{V} = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}, \quad \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}, \quad \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right)$$

意味：「その場がどれだけ渦を巻いているか」を表す。（例：台風の周りの風速分布では、中心付近で回転成分が強くなります）

発散 (div > 0) 回転 (rot ≠ 0)

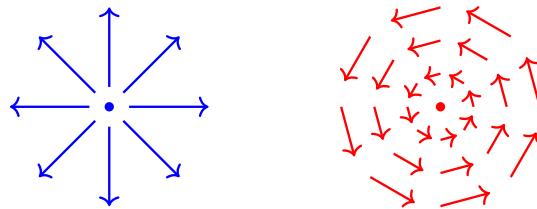


図 9.1 場のイメージ：左は湧き出し（発散）、右は渦（回転）。

9.3 線積分と面積分～流れに乗る・流れを測る～

通常の積分 $\int_a^b f(x)dx$ は「直線上」での足し算でしたが、これを曲がった空間へ拡張します。

直感でわかる！線積分と面積分～高校物理からの拡張～

「線積分」や「面積分」という言葉を聞くと難しそうに感じますが、実はこれらは高校物理で習った「仕事」や「流量」を、曲がった世界に拡張したものに過ぎません。

1. 線積分：曲がりくねった道での「仕事」

高校物理では、「仕事 W 」を次のように習いました。

$$W = F \times d \quad (\text{力} \times \text{距離})$$

ただし、これは「力が一定」で「道が真っ直ぐ」な場合だけの話です。

もし、「道が曲がりくねっていて、風向き（力）も場所ごとに変わる」としたらどう計算すればよいでしょうか？答えはシンプルです。「道を極限まで細かく区切って、一瞬だけ真っ直ぐとみなす」のです。

- 道を微小な区間 Δr に刻みます。
- その一瞬の区間での仕事 $\mathbf{F} \cdot \Delta \mathbf{r}$ を計算します（内積を使うのは、進む方向の成分だけが有効だからです）。
- それらを全部足し合わせます（積分）。

これが線積分 $\int_C \mathbf{F} \cdot d\mathbf{r}$ の正体です。「変化する環境の中で、ゴールまで進むのにどれだけエネルギーを使ったか（または助けてもらったか）」を表しています。

2. 面積分：曲がった網でくう「流量」

次に、「流量 Q 」を思い出しましょう。川の中に網を置いたとき、通過する水の量は

$$Q = v \times S \quad (\text{流速} \times \text{断面積})$$

でした。これも「流れが一定」で「網が平面」の場合の話です。

もし、「流れが複雑で、網も丸く歪んでいる（金魚すくいのポイが破れそうな状態）」としたら？やはり同じアプローチをとります。

1. 網（曲面）を微小なタイル ΔS に刻みます。
2. 各タイルを通過する水量 $\mathbf{F} \cdot \mathbf{n} dS$ を計算します（ \mathbf{n} はタイルの向き。正面から受ける水だけをカウントするため内積を使います）。
3. それらを全部足し合わせます。

これが面積分 $\iint_S \mathbf{F} \cdot \mathbf{n} dS$ の正体です。「歪んだフィルター全体で、合計どれだけの量が通過したか」を表しています。

9.3.1 線積分 (Line Integral)

「流れの中を泳ぐときの負担」

曲線 C に沿って進むとき、ベクトル場（風や川の流れ） \mathbf{F} からどれくらい「後押し（または抵抗）」を受けるかを計算します。

- 数式:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{t_0}^{t_1} (\text{力のベクトル } \mathbf{F}) \cdot (\text{進む方向 } \mathbf{v}) dt$$

（内積 \cdot を取るのは、「進行方向と一致する成分」だけが仕事をするからです。真横からの風は仕事をしません）

- 物理的イメージ（仕事）：重い荷物を運ぶとき、追い風（ \mathbf{F} と $d\mathbf{r}$ が同じ向き）ならプラスの仕事、向かい風ならマイナスの仕事になります。その合計が線積分です。

9.3.2 面積分 (Surface Integral)

「網でくう水の量」

曲面 S （網）を川の中に置いたとき、その面を通過する「流体の総量（フラックス）」を計算します。

- 数式:

$$\iint_S \mathbf{F} \cdot \mathbf{n} dS$$

ここで \mathbf{n} は面の「法線ベクトル（面の向き）」です。

- 物理的イメージ（流量）：川の流れ \mathbf{F} に対して、網を垂直に立てれば大量の水が通過します（内積最大）。網を水平に寝かせれば、水は網の目を素通りせず、通過量はゼロになります（内積ゼロ）。「面を垂直に貫く成分」だけを足し合わせたものが面積分です。

9.4 平面での積分定理：グリーンの定理

「微分の積分は、境界での値になる（微積分の基本定理）」という哲学の、2次元平面バージョンです。

9.4.1 定理の主張と直感的意味

定理 9.1 (グリーンの定理). 平面領域 D とその境界曲線 ∂D (反時計回り) において、

$$\oint_{\partial D} (Pdx + Qdy) = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy$$

- 直感的イメージ (台風の目の集合体) : お風呂にたくさんの「小さな渦巻き (ミクロな回転)」が敷き詰められていると想像してください。隣り合う渦巻き同士は、接している部分で流れが逆向きになるため、打ち消し合います ($\uparrow\downarrow$)。しかし、一番外側の境界線だけは、打ち消す相手がいません。
- 結論: 「内部の小さな渦 (回転 rot の z 成分) の合計」は、「外周を回る大きな流れ (線積分)」と等しくなります。

9.4.2 厳密な証明

領域 D が「縦線形領域 (y 軸方向に単純)」かつ「横線形領域 (x 軸方向に単純)」である場合について示せば十分です (複雑な領域はこれらに分割できるため)。ここでは、第 2 項 $-\iint_D \frac{\partial P}{\partial y} dxdy = \oint_{\partial D} Pdx$ を示します (第 1 項も同様)。

1. 領域の設定: D を $a \leq x \leq b$, $g_1(x) \leq y \leq g_2(x)$ と表す。境界 ∂D は、以下の 4 つの部分で構成される:

- 下側の曲線 $C_1 : y = g_1(x)$ ($x : a \rightarrow b$)
- 右側の縦線 ($dx = 0$ なので積分値は 0)
- 上側の曲線 $C_2 : y = g_2(x)$ ($x : b \rightarrow a$ 。向きが逆!)
- 左側の縦線 ($dx = 0$ なので積分値は 0)

2. 重積分 (左辺に相当する部分) の計算:

$$-\iint_D \frac{\partial P}{\partial y} dxdy = - \int_a^b \left[\int_{g_1(x)}^{g_2(x)} \frac{\partial P}{\partial y} dy \right] dx$$

内側の積分は微積分学の基本定理より :

$$\begin{aligned} &= - \int_a^b (P(x, g_2(x)) - P(x, g_1(x))) dx \\ &= \int_a^b P(x, g_1(x)) dx - \int_a^b P(x, g_2(x)) dx \end{aligned}$$

3. 線積分 (右辺に相当する部分) の計算: $\oint_{\partial D} Pdx$ を計算する。

- 下側 C_1 : $y = g_1(x)$ で x は $a \rightarrow b$ 。寄与は $\int_a^b P(x, g_1(x)) dx$ 。
- 上側 C_2 : $y = g_2(x)$ で x は $b \rightarrow a$ 。寄与は $\int_b^a P(x, g_2(x)) dx = - \int_a^b P(x, g_2(x)) dx$ 。

これらを合計すると :

$$\int_a^b P(x, g_1(x)) dx - \int_a^b P(x, g_2(x)) dx$$

4. 結論: 2 と 3 の結果が完全に一致した。同様に Q の項についても示せるので、定理は証明された。(証明終)

コメント 9.1 (実社会のイメージ : マクスウェル方程式). 電磁気学の基礎方程式は、これらを使って記述されます。

- $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$ (ガウスの法則：電荷があるところから電場が湧き出す)
- $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ (ファラデーの法則：磁場が変化すると、渦を巻く電場が生まれる)

ベクトル解析は、自然界の「場」の法則を記述するための最も洗練された言語なのです。

9.5 第9章のまとめ

- ∇ (ナブラ)：勾配、発散、回転を生み出す万能演算子。
- 線積分・面積分：仕事や流量を計算するための積分。
- 積分定理：「内部の導関数（源）」と「境界の値（出入り）」を結ぶ、解析学の集大成。

第 10 章

積分定理の統合と展望

第 9 章で学んだグリーンの定理を、3 次元空間へ拡張します。

コラム：微分形式とは何か？～ dx の正体～

高校数学で、積分の式に出てくる dx を「お飾り」だと思っていませんでしたか？

$$\int_a^b f(x)dx$$

実はこの dx （微分形式）こそが、解析学の真の主役です。

1. なぜ登場したのか？（動機）

これまでのベクトル解析（勾配、発散、回転）には、致命的な欠点がありました。「次元が変わると公式の形が変わる」のです。

- 2 次元の回転（rot）はスカラー量ですが、3 次元の回転はベクトル量です。
- グリーンの定理（2 次元）、ガウスの定理（3 次元）、ストークスの定理（曲面）は、バラバラの式に見えます。

「これらをたった一つの式で統一したい！」その願いから生まれたのが、微分形式です。 dx や dy に厳密な代数ルール（外積代数）を与えることで、次元に関係なく共通の計算が可能になりました。

2. 高校数学からの接続

高校で習った「置換積分」を思い出して下さい。 $x = g(t)$ と置くとき、 $dx = g'(t)dt$ と書き換えましたよね？実は、これこそが微分形式の計算ルールそのものです。

微分形式のルール（外積代数）：

- 順序を入れ替えるとマイナスになる（反対称性）：

$$dx \wedge dy = -dy \wedge dx$$

（面積には「表と裏」があるため。向きが変わると符号が変わる）

- 同じものが重なるとゼロになる：

$$dx \wedge dx = 0$$

（幅のない線の面積はゼロだから）

これを使うと、重積分の変数変換（ヤコビアン）が自動的に導かれます。例えば、 $x = r \cos \theta, y = r \sin \theta$ のとき、

$$dx = \cos \theta dr - r \sin \theta d\theta$$

$$dy = \sin \theta dr + r \cos \theta d\theta$$

これらを「掛け算（ウェッジ積 \wedge ）」すると：

$$\begin{aligned} dx \wedge dy &= (\cos \theta dr - r \sin \theta d\theta) \wedge (\sin \theta dr + r \cos \theta d\theta) \\ &= r(\cos^2 \theta + \sin^2 \theta) dr \wedge d\theta = r dr \wedge d\theta \end{aligned}$$

面倒なヤコビ行列式を計算しなくても、ただの展開計算で r が出てきました！

3. 究極の定理：ストークスの定理

微分形式を使うと、グリーンの定理もガウスの定理も、すべて以下の形になります。

$$\int_D d\omega = \int_{\partial D} \omega$$

「中身の変化（微分 d ）を足し合わせると、縁での値になる」この美しさこそが、現代数学が微分形式を採用する理由です。

10.1 ガウスの発散定理 (Gauss's Divergence Theorem)

「中身の湧き出しの合計は、外への流出量に等しい」という定理です。

10.1.1 定理の主張と直感的意味

定理 10.1 (ガウスの発散定理). 空間領域 V とその表面 S (外向き法線 \mathbf{n}) において、

$$\iiint_V (\nabla \cdot \mathbf{F}) dV = \iint_S \mathbf{F} \cdot \mathbf{n} dS$$

- 直感的イメージ（満員電車の乗客数）：ある空間（部屋や車両）を考えます。各地点で人が「湧き出したり ($\text{div} > 0$)」、「消滅したり ($\text{div} < 0$)」しています。空間全体で「合計どれだけの人が増えたか（ミクロな湧き出しの総和）」を知りたければ、「壁（表面）から出て行った人の数」を数えればよいのです。

10.1.2 厳密な証明

ベクトル場 $\mathbf{F} = (P, Q, R)$ の各成分ごとに示す。ここでは z 成分 $\iiint_V \frac{\partial R}{\partial z} dV = \iint_S R n_z dS$ について示す。

- 領域の設定: V を z 軸方向に単純な領域とする。すなわち、 D を xy 平面の射影として、 $g_1(x, y) \leq z \leq g_2(x, y)$ 。
- 体積積分の計算:

$$\iiint_V \frac{\partial R}{\partial z} dx dy dz = \iint_D \left[\int_{g_1}^{g_2} \frac{\partial R}{\partial z} dz \right] dx dy$$

基本定理より：

$$= \iint_D (R(x, y, g_2) - R(x, y, g_1)) dx dy$$

- 面積分の計算: 表面 S は、上面 S_{top} 、下面 S_{bottom} 、側面 S_{side} に分かれる。

- 側面: 法線 \mathbf{n} は z 軸と垂直なので $n_z = 0$ 。積分値は 0。
- 上面 ($z = g_2$): 法線は上向き。面素の関係 $n_z dS = dx dy$ が成り立つ。寄与は $\iint_D R(x, y, g_2) dx dy$ 。
- 下面 ($z = g_1$): 法線は下向き (外向きだから)。面素の関係 $n_z dS = -dx dy$ 。寄与は $-\iint_D R(x, y, g_1) dx dy$ 。

4. 結論: 2 と 3 の結果が一致した。 x, y 成分も同様に示せる。(証明終)

10.2 ストークスの定理 (Stokes' Theorem)

「曲面上の小さな渦の合計は、縁 (ふち) の循環に等しい」という定理です。

10.2.1 定理の主張と直感的意味

定理 10.2 (ストークスの定理). 曲面 S とその境界 ∂S において、

$$\oint_{\partial S} \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n} dS$$

- 直感的イメージ (曲面上のパッチワーク) : 曲面全体を小さな網目で区切れます。各網目で「小さな渦」が回っています。隣同士の渦は打ち消し合いますが、丘のふもと (境界線) だけは打ち消されずに残ります。

10.2.2 厳密な証明

曲面 S が $z = f(x, y)$ というグラフで表される場合で証明する。「曲面上の線積分を、 xy 平面上のグリーンの定理に帰着させる」のがポイント。

1. 左辺 (線積分) の変形: 曲線 ∂S 上の点は $(x, y, f(x, y))$ 。微小変位は $dz = f_x dx + f_y dy$ 。 $\mathbf{F} = (P, Q, R)$ の線積分は :

$$\oint_{\partial S} (P dx + Q dy + R dz) = \oint_{\partial D} [(P + R f_x) dx + (Q + R f_y) dy]$$

これは xy 平面上の領域 D の境界 ∂D での線積分である。

2. グリーンの定理を適用: $P^* = P + R f_x$, $Q^* = Q + R f_y$ とおいて、グリーンの定理 $\iint_D (\frac{\partial Q^*}{\partial x} - \frac{\partial P^*}{\partial y}) dx dy$ を計算する。偏微分の連鎖律に注意して差をとると ($f_{xy} = f_{yx}$ で項が消える) :

$$\frac{\partial Q^*}{\partial x} - \frac{\partial P^*}{\partial y} = (\text{rot } \mathbf{F})_z - (\text{rot } \mathbf{F})_y f_y - (\text{rot } \mathbf{F})_x f_x$$

3. 右辺 (面積分) の計算: 曲面 $z = f(x, y)$ の上向き単位法線ベクトル \mathbf{n} と面素 dS の積は公式より $\mathbf{n} dS = (-f_x, -f_y, 1) dx dy$ 。これと回転ベクトル $\nabla \times \mathbf{F}$ の内積をとる :

$$\iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n} dS = \iint_D [(\text{rot } \mathbf{F})_x (-f_x) + (\text{rot } \mathbf{F})_y (-f_y) + (\text{rot } \mathbf{F})_z (1)] dx dy$$

4. 結論: 2 と 3 の結果が見事に一致した。よって定理は成立する。(証明終)

10.3 積分定理の統一 : 微分形式へ

これまで学んだ定理は、すべて「多様体上の微分形式の積分」として統一されます。

10.3.1 例題1：高校数学で解く微分形式～置換積分の正体～

問題：次の微分形式の計算を行い、関係式を導け。

1. 1変数の場合（置換積分）： $x = t^2$ のとき、 dx を t と dt で表せ。また、 $f(x) = x^2$ について、 $df(x) = f'(x)dx$ を計算し、それを t で表した結果と一致することを確認せよ。
2. 2変数の場合（極座標の面積要素）： $x = r \cos \theta$, $y = r \sin \theta$ のとき、微分形式のウェッジ積（外積）のルール

$$dx \wedge dy = -dy \wedge dx, \quad dx \wedge dx = 0, \quad dy \wedge dy = 0$$

を使って、 $dx \wedge dy$ を $dr \wedge d\theta$ の形で表せ。

【解答】

1. 1変数の場合

- $x = t^2 \implies dx = 2t dt$.
- $f(x) = x^2 \implies df = 2x dx = 2(t^2)(2t dt) = 4t^3 dt$.
- 一方、 $f(t^2) = t^4$ を t で微分しても $4t^3 dt$ となり一致する。

2. 2変数の場合全微分をとる：

$$dx = \cos \theta dr - r \sin \theta d\theta, \quad dy = \sin \theta dr + r \cos \theta d\theta$$

ウェッジ積をとる（展開）：

$$\begin{aligned} dx \wedge dy &= (\cos \theta dr - r \sin \theta d\theta) \wedge (\sin \theta dr + r \cos \theta d\theta) \\ &= (\cos \theta)(\sin \theta) \underbrace{dr \wedge dr}_0 + r \cos^2 \theta (dr \wedge d\theta) \\ &\quad - r \sin^2 \theta (d\theta \wedge dr) - r^2 \sin \theta \cos \theta \underbrace{d\theta \wedge d\theta}_0 \\ &= r \cos^2 \theta (dr \wedge d\theta) + r \sin^2 \theta (dr \wedge d\theta) \quad (\because d\theta \wedge dr = -dr \wedge d\theta) \\ &= r(\cos^2 \theta + \sin^2 \theta) dr \wedge d\theta = r dr \wedge d\theta \end{aligned}$$

高校数学で暗記したヤコビアン r が、代数計算だけで自然に導かれた。

10.3.2 例題2：微分形式の積分～グリーンの定理の確認～

問題：単位円内部 $D : x^2 + y^2 \leq 1$ とその境界 ∂D について、微分形式 $\omega = x dy$ の積分を計算せよ。

【解答】

1. 境界での積分（線積分）： $x = \cos t, y = \sin t$ ($0 \leq t \leq 2\pi$) と置くと、 $dy = \cos t dt$ 。

$$\int_{\partial D} x dy = \int_0^{2\pi} \cos t \cdot \cos t dt = \int_0^{2\pi} \frac{1 + \cos 2t}{2} dt = \pi$$

2. 内部での積分（重積分）：まず外微分を計算する。 $d\omega = d(x) \wedge dy = dx \wedge dy$ 。

$$\int_D d\omega = \iint_D dx \wedge dy = \iint_{x^2+y^2 \leq 1} 1 dx dy = (\text{単位円の面積}) = \pi$$

結論： $\int_{\partial D} \omega = \int_D d\omega$ が成立している（一般化ストークスの定理の確認）。

10.4 一般化されたストークスの定理

微積分学の基本定理、グリーンの定理、ガウスの定理、ストークスの定理。これらはすべて、次の美しい数式に帰着されます。

定理 10.3 (一般化ストークスの定理). n 次元多様体 M と、その境界 ∂M 上の微分形式 ω について、

$$\int_M d\omega = \int_{\partial M} \omega$$

解釈:

- 左辺: 領域 M の内部での「変化 (微分 $d\omega$)」をすべて足し合わせたもの。
- 右辺: 境界 ∂M での「値 (ω)」を足し合わせたもの。

「中身の合計は、縁 (ふち) で決まる」。これが、高校数学で習った $\int_a^b f'(x)dx = f(b) - f(a)$ から続く、解析学を貫く大原則です。

10.5 おわりに : 解析学のその先へ

本テキストでは、以下の道のりを歩んできました。

1. **極限:** 無限を論理で飼いならす ($\epsilon - \delta$)。
2. **微分:** 複雑なものを線形 (直線・平面) で近似する。
3. **積分:** 微小な断片を積み上げて全体を知る。
4. **ベクトル解析:** 空間の構造と流れを理解する。

この先には、関数そのものをベクトルとみなす「関数解析」、曲がった空間 (リーマン幾何学) で行う「多様体論」、そして複素数の世界で微分を行う「複素解析」といった、さらに広大で美しい数学の世界が広がっています。

しかし、どのような高度な数学に進んでも、ここで学んだ「近似」「極限」「局所と大域の関係」という視点は、常に強力な羅針盤となるはずです。本テキストが、皆様の大学数学、そしてその先の科学技術への冒険の確かな第一歩となることを願っています。

Python で見る微積分 (Computational Calculus)

解析学の概念を、プログラミング言語 Python を使って視覚化・実験してみましょう。以下のコードは、Google Colab などのノートブック環境でそのまま実行可能です。（必要なライブラリ：‘numpy’, ‘matplotlib’, ‘scipy’）

.1 テイラー展開の可視化～多項式近似の威力～

第4章で学んだ「関数を多項式で近似する」様子をアニメーション的に描画します。次数 n を上げるにつれて、近似が元の関数 $\sin x$ にピタリと張り付いていく様子を確認してください。

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 import math
4
5 def plot_taylor_sin(max_order=15, x_range=(-10, 10)):
6     """sin(x)のマクローリン展開を次数ごとにプロットする"""
7     x = np.linspace(x_range[0], x_range[1], 400)
8     y_true = np.sin(x)
9
10    plt.figure(figsize=(10, 6))
11    plt.plot(x, y_true, 'k', linewidth=2, label='sin(x) (True)')
12    plt.ylim(-2, 2)
13
14    # 近似多項式の計算
15    # sin(x) = x - x^3/3! + x^5/5! - ...
16    y_approx = np.zeros_like(x)
17
18    for n in range(1, max_order + 2, 2):  # 奇数のみ
19        term = ((-1)**((n-1)//2)) * (x**n) / math.factorial(n)
20        y_approx += term
21
22        # 次数が低いもの(1, 3, 5...)を描画
23        if n in [1, 3, 5, 9, 13]:
24            idx = (n-1)//2
25            plt.plot(x, y_approx, '--', label=f'Order {n}', alpha=0.8)
26
27    plt.title(f'Taylor Expansion of sin(x) up to Order {max_order}')
28    plt.xlabel('x')
29    plt.ylabel('y')
30    plt.legend()
31    plt.grid(True)

```

```

32     plt.show()
33
34 # 実行
35 plot_taylor_sin()

```

Listing 1 Taylor Expansion Visualization

.2 勾配降下法シミュレーション～2変数関数の最適化～

第6章の「最適化」を体験します。2変数関数 $f(x, y) = x^2 + y^2 + \sin(5x) + \sin(5y)$ の複雑な地形で、ボールが転がり落ちる（極小値を探す）様子を可視化します。

```

1 def gradient_descent_demo():
2     # 対象関数（少し波打ったお椀型）
3     def f(x, y):
4         return 0.5 * (x**2 + y**2) + 2 * np.sin(x) * np.sin(y)
5
6     # 勾配（偏微分）
7     def grad_f(x, y):
8         df_dx = x + 2 * np.cos(x) * np.sin(y)
9         df_dy = y + 2 * np.sin(x) * np.cos(y)
10        return np.array([df_dx, df_dy])
11
12    # 勾配降下法の設定
13    learning_rate = 0.1
14    current_pos = np.array([2.5, 2.0]) # スタート地点
15    path = [current_pos]
16
17    # 更新ループ
18    for _ in range(30):
19        grad = grad_f(current_pos[0], current_pos[1])
20        current_pos = current_pos - learning_rate * grad
21        path.append(current_pos)
22
23    path = np.array(path)
24
25    # プロット（等高線図 + 軌跡）
26    x = np.linspace(-4, 4, 100)
27    y = np.linspace(-4, 4, 100)
28    X, Y = np.meshgrid(x, y)
29    Z = f(X, Y)
30
31    plt.figure(figsize=(8, 8))
32    plt.contourf(X, Y, Z, levels=20, cmap='coolwarm', alpha=0.6)
33    plt.colorbar(label='f(x, y)')
34    plt.plot(path[:, 0], path[:, 1], 'o-', color='black', label='Gradient Path')
35    plt.scatter(path[0, 0], path[0, 1], color='red', s=100, label='Start')
36    plt.scatter(path[-1, 0], path[-1, 1], color='green', s=100, label='End')
37    plt.title('Gradient Descent Optimization')
38    plt.legend()
39    plt.show()
40
41 # 実行

```

```
42 gradient_descent_demo()
```

Listing 2 Gradient Descent Simulation

.3 モンテカルロ積分～乱数で面積・体積を測る～

第7章の「重積分」の応用です。積分公式を使わず、ランダムに点を打って「領域内に入った割合」から面積（この場合は円周率）を求める手法です。高次元積分で威力を発揮します。

```
1 def monte_carlo_pi(n_samples=10000):
2     # [0,1]x[0,1] の正方形内にランダムに点を打つ
3     x = np.random.rand(n_samples)
4     y = np.random.rand(n_samples)
5
6     # 原点からの距離が以下の点（扇形の中）を判定1
7     inside_circle = (x**2 + y**2) <= 1.0
8     pi_estimate = 4 * np.sum(inside_circle) / n_samples
9
10    # 可視化
11    plt.figure(figsize=(6, 6))
12    plt.scatter(x[inside_circle], y[inside_circle], s=1, c='blue', alpha=0.5,
13                 label='Inside')
14    plt.scatter(x[~inside_circle], y[~inside_circle], s=1, c='red', alpha=0.5,
15                 label='Outside')
16    plt.title(f'Monte Carlo Integration of PI\nEstimate: {pi_estimate:.5f} (
17         Samples: {n_samples})')
18    plt.axis('equal')
19    plt.legend(loc='upper right')
20    plt.show()
21
22    # 実行
23    monte_carlo_pi()
```

Listing 3 Monte Carlo Integration

参考文献・推奨図書

本書で学んだ内容をさらに深めたい、あるいは応用分野へ繋げたい読者のために、名著を厳選しました。

【解析学の基礎を深める】

- 杉浦光夫『解析入門 I・II』(東京大学出版会)

日本における解析学の聖書（バイブル）。本書で扱った定理の証明が、より一般的かつ厳密な形で記述されています。「 $\epsilon - \delta$ 論法」を完全にマスターしたいならこの本です。

- W. ルディン『現代解析学』(共立出版)

世界標準の教科書。無駄を削ぎ落としたエレガントな証明が特徴で、数学的な美意識を磨くことができます。

【物理・工学への応用】

- ファインマン『ファインマン物理学 II 電磁気学』(岩波書店)

本書の第9章・第10章（ベクトル解析）が、物理の世界でどのように「実体」として使われているか、著者の圧倒的な語り口で理解できます。

- ギルバート・ストラング『ストラング：線形代数とデータサイエンス』(近代科学社)

多変数関数の解析（第5章以降）には、線形代数（行列・ベクトル）の知識が不可欠です。現代的な視点で線形代数を学び直すのに最適な一冊です。

【機械学習・統計学への接続】

- C.M. ビショップ『パターン認識と機械学習』(丸善出版)

通称「PRML」。本書で学んだ「最大値問題（第6章）」「ガウス積分（第8章）」が、最先端のAIモデルの中でどのように生きているかを目撃できます。

- 甘利俊一『情報幾何学の新展開』(サイエンス社)

統計学における確率分布の空間を「曲がった空間」として捉える理論。本書の延長線上にある、現代数学の深淵に触れることができます。