線形代数の構造とダイナミクス
〜 現代数学の基礎から高次元データ解析へ 〜

Antigravity

2026 年 2 月 9 日

目次

第 0 章

序章：高校数学からの接続と拡張 〜「計算」から「構造」の理解へ〜

本書は，あなたが高校で学んだ「ベクトル」と「行列」の知識を基礎としてスタートします．しかし，大学（およびデータ科学）での使い方は，高校までとは少し視点が異なります．

本格的な定義に入る前に，**「何が同じで，何が変わるのか」**を整理しておきましょう．

0.1 「矢印」から「データのリスト」へ ($\mathbb{R}^n$ の導入)

【高校数学】

高校では，ベクトルを主に**「向きと大きさを持つ矢印」**として扱い，2 次元（平面）や 3 次元（空間）の図形問題を解くために使いました．

$$\vec{a} = (2, 3), \quad \vec{b} = (1, -1, 4)$$

【線形代数での拡張】

データ解析の現場では，変数は 2 つや 3 つではありません．顧客データなら「年齢・年収・購入額・来店回数…」と 10 個，100 個の数字が並びます．これらもベクトルとして扱いたいのです．

そこで，矢印のイメージを一旦脇に置き，**「数を n 個並べたリスト（数ベクトル）」**として定義し直します．これを $\mathbb{R}^n$（エヌ次元実ベクトル空間）と呼びます．

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

- **記号の変化**: 高校の $\vec{a}$ から，太字の $\boldsymbol{x}$ や $\boldsymbol{a}$ へと表記が変わりますが，実体は同じです．
- **次元の拡張**: 図には描けませんが，計算ルール（足し算・スカラー倍）は 2 次元のときと全く同じです．

0.2 「連立方程式」から「写像」へ

【高校数学】

連立方程式を解くとき，行列を行基本変形（ガウスの消去法）して「解 (x, y) を求めること」がゴールでした．

$$\begin{cases} 2x + y = 5 \\ x - y = 1 \end{cases} \xrightarrow{\text{行列で計算}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

【線形代数での拡張】

大学では，「解けるかどうか」や「解がどんな形をしているか」という**構造**に注目します．行列 A を，ベクトル $\boldsymbol{x}$ を別のベクトル $\boldsymbol{y}$ に変換する**「関数（写像）」**だと見なします．

$$\boldsymbol{y} = A\boldsymbol{x}$$

これは，関数 $y = f(x)$ の多次元版です．

- 高校：連立方程式を解く ＝ 逆算して x を求める．
- 大学：行列 A の性質（潰れているか？ 回転しているか？）を調べる．

0.3 「重ね合わせ」という最強のルール

高校数学で，ベクトル $\vec{a}, \vec{b}$ を使って $s\vec{a} + t\vec{b}$ という形を何度も書いたはずです．これこそが線形代数の魂であり，**「線形結合（一次結合）」**と呼ばれます．

線形代数とは，究極的には以下の式が成り立つ世界のことです．

$$f(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1 f(\boldsymbol{x}_1) + c_2 f(\boldsymbol{x}_2)$$

- **入力の足し算は，出力の足し算になる．**
- **入力の定数倍は，出力の定数倍になる．**

この「単純なルール（線形性）」が成り立つからこそ，複雑なデータを単純な要素（基底）に分解して解析できるのです．

この章の意義

このように，

1. **表記:** $\vec{a}$ から $\boldsymbol{x} \in \mathbb{R}^n$ へ
2. **目的:** 「解を求める」から「変換の性質を知る」へ
3. **核心:** 「図形」から「線形性（重ね合わせ）」へ

と視点をずらすことで，第 1 章からの「線形空間の公理」が，「新しい難しいルール」ではなく，**「高校で扱っていたベクトルの性質を，どんな次元でも使えるように整理したもの」**としてスムーズに入ってくるはずです．

第 I 部

線形空間の構造論

第 1 章

線形空間と基底 (Vector Spaces and Basis)

【導入】矢印だけが「ベクトル」ではない

高校までの数学では,「ベクトル」と言えば「向きと大きさを持つ矢印(幾何ベクトル)」や「数のペア(数ベクトル)」のことでした.しかし,大学の線形代数では,**「ベクトル空間(線形空間)」**という,より抽象的な定義からスタートします.なぜ,わざわざ面倒な定義をするのでしょうか?

その理由は,**「一見まったく別物に見えるデータたちを,同じ道具で扱いたいから」**です.

例えば,以下の 3 つを見てください.

1. **物理の矢印**: 力の合成($F_1 + F_2$),力の 2 倍($2F_1$)ができる.
2. **音声波形(関数)**: 波の重ね合わせ($f(t) + g(t)$),ボリュームアップ($2f(t)$)ができる.
3. **多項式**: 式の足し算($(x^2 + 1) + (x - 1)$),定数倍($3(x^2 + 1)$)ができる.

これらは,「見た目」は全く違いますが,**「足し算ができる」「定数倍ができる」という構造(ルール)は完全に同じ**です.

数学者たちはこう考えました.

> 「中身が矢印か関数かはどうでもいい.『足し算』と『定数倍』のルール(公理)さえ満たしていれば,すべて『ベクトル』と呼んでしまおう.そうすれば,矢印のために作った便利な定理(次元定理や固有値など)が,そのまま関数や画像データにも使い回せるはずだ」

つまり,線形空間を定義するメリットは以下の 2 点に集約されます.

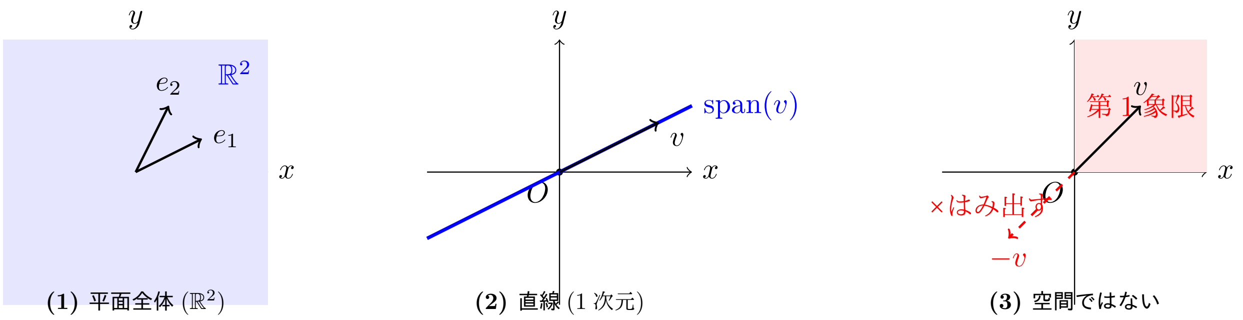
1. **統一(Unification)**: 画像も,音声も,確率分布も,すべて「ベクトル」という同じ土俵に乗せることで,統一的に扱えるようになります.
2. **道具の再利用(Reusability)**: 「ベクトルの一次独立性」や「基底」といった概念を一度証明してしまえば,それを「多項式の補間問題」や「信号の周波数解析」に,証明なしでそのまま適用できます.

プログラミングに例えるなら,線形空間とは**「Interface(インターフェース)」**の定義です.`VectorSpace` というインターフェース(8 つの公理)さえ実装していれば,中身が `Arrow` クラスだろうが `Wave` クラスだろうが `Polynomial` クラスだろうが,同じ `linear_analysis()` ライブラリに放り込んで解析できるのです.

これから学ぶ抽象的な定義は,**「世界中のあらゆる現象を『重ね合わせ(線形結合)』で説明するための,汎用的なアダプターを作る作業」**だと思ってください.

1.1　線形空間の可視化：到達可能な領域

ベクトル空間をイメージするときは，矢印そのものではなく，**「そのベクトルたちを使って，足し算とスカラー倍でどこまで行けるか（到達可能な点の集合）」**を想像するのがコツです．これを**生成する部分空間 (Span)** と呼びます．



(1) 平面全体 ($\mathbb{R}^2$)
2 本の独立なベクトルで
張られる空間

(2) 直線 (1 次元)
1 本のベクトルで
張られる空間

(3) 空間ではない
逆ベクトル ($-v$) が
集合から出るため

図 1.1: ベクトル空間のイメージ：原点を通る「無限に広がる」直線や平面．(3) のように「途切れる」ものは空間ではない．

「線形空間である」とは，**「足し算とスカラー倍をいくら繰り返しても，その世界から決してはみ出さない（閉じている）」**ことを意味します．

1.2　線形空間の定義

高校数学では，ベクトルを「向きと大きさを持つ矢印」として習いました．しかし，データサイエンスや高度な工学においては，**画像，音声波形，確率分布，多項式**など，一見「矢印」に見えないものもすべて「ベクトル」として扱います．本章では，これらを統一的に扱うための舞台である「線形空間（ベクトル空間）」を定義し，その構造を解明します．

「足し算」と「スカラー倍（定数倍）」ができる集合を，線形空間と呼びます．

定義 1.1 (線形空間 (Vector Space))**.** 集合 V と体 K（実数 $\mathbb{R}$ や複素数 $\mathbb{C}$）に対し，2 つの演算（加法 $+ : V \times V \to V$，スカラー倍 $\cdot : K \times V \to V$）が定義され，以下の公理を満たすとき，V を K 上の線形空間と呼ぶ．

1. **加法の群構造:**
   - (交換律) $\forall u, v \in V, u + v = v + u$
   - (結合律) $\forall u, v, w \in V, (u + v) + w = u + (v + w)$
   - (ゼロ元) $\exists \mathbf{0} \in V$ s.t. $\forall u \in V, u + \mathbf{0} = u$
   - (逆元) $\forall u \in V, \exists (-u) \in V$ s.t. $u + (-u) = \mathbf{0}$
2. **スカラー倍の作用:**
   - (分配律 1) $\forall a \in K, \forall u, v \in V, a(u + v) = au + av$
   - (分配律 2) $\forall a, b \in K, \forall u \in V, (a + b)u = au + bu$
   - (結合律) $\forall a, b \in K, \forall u \in V, (ab)u = a(bu)$
   - (単位元) $1 \cdot u = u$

例題 1.1 (多項式空間)**.** n 次以下の多項式全体 $P_n(\mathbb{R}) = \{a_n x^n + \cdots + a_0 \mid a_i \in \mathbb{R}\}$ を考えます．

- 多項式同士は足せます（$(x^2+1)+(x-1)=x^2+x$）.
- 定数倍できます（$3(x^2+1)=3x^2+3$）.

よって，**多項式はベクトルの一種**です．機械学習で曲線をフィッティングする際，我々は「多項式空間」というベクトル空間の中で最適な点を探していることになります.

例題 1.2 (RGB カラーモデル)**.** 画面上の色は，赤 (R)・緑 (G)・青 (B) という「基底」の線形結合 $c_1\vec{r}+c_2\vec{g}+c_3\vec{b}$ で表されます．この 3 つのベクトルが張る空間が「色空間」であり，次元は 3 です.

例題 1.3 (音声信号)**.** 音声データ（波形関数 $f(t)$）を考えます.

- 2 つの音を重ねる（$f(t)+g(t)$）ことはベクトルの和です.
- ボリュームを上げる（$2f(t)$）ことはスカラー倍です.
- 無音（0）はゼロベクトルです.

つまり，**「音声信号全体の集合」は無限次元の線形空間**となります．この視点を持つことで，ベクトルの理論を信号処理にそのまま応用（フーリエ変換など）できるようになります.

1.3　線形独立と基底

空間の「広さ」や「自由度」を定義するために，基底（Basis）を導入します.

定義 1.2 (線形独立 (Linearly Independent))**.** V のベクトル $v_1,\dots,v_k$ が線形独立であるとは，

$$c_1v_1+\cdots+c_kv_k=\mathbf{0} \quad \Longrightarrow \quad c_1=\cdots=c_k=0$$

が成り立つことをいう.（自明な組み合わせ以外でゼロを作れない＝どのベクトルも他のベクトルの合成で作れない，という意味）

定義 1.3 (基底 (Basis) と次元 (Dimension))**.** ベクトル群 $\{e_1,\dots,e_n\}$ が以下の 2 条件を満たすとき，V の基底と呼ぶ.

1. **線形独立性**: $\{e_1,\dots,e_n\}$ は線形独立である.
2. **生成性**: 任意の $v\in V$ は $v=\sum_{i=1}^n c_ie_i$ と表せる.

このとき，基底の本数 n は（基底の選び方によらず）一定となる．これを V の次元と呼び，$\dim V=n$ と書く.

定理 1.1 (基底の延長定理)**.** V を有限次元線形空間とする．V の線形独立なベクトル集合 $\{v_1,\dots,v_k\}$ は，適当なベクトル $v_{k+1},\dots,v_n$ を追加することで，V の基底 $\{v_1,\dots,v_n\}$ に延長できる.

Proof.
1. 現在の集合 $S_k=\{v_1,\dots,v_k\}$ が V を生成する（つまり $\mathrm{span}(S_k)=V$）ならば，定義より S_k は基底である．証明終了.
2. 生成しない場合，$\mathrm{span}(S_k)$ に含まれないベクトル v_{k+1} が存在する.
3. 新しい集合 $S_{k+1}=\{v_1,\dots,v_k,v_{k+1}\}$ を考える．これが線形独立であることを背理法で示す.
   - $\sum_{i=1}^{k+1}c_iv_i=\mathbf{0}$ とし，少なくとも 1 つの係数が非ゼロと仮定する.
   - もし $c_{k+1}\neq 0$ なら，$v_{k+1}=-\frac{1}{c_{k+1}}\sum_{i=1}^{k}c_iv_i$ となり，$v_{k+1}\in\mathrm{span}(S_k)$ となって矛盾する（選び方に反する）.
   - よって $c_{k+1}=0$ でなければならない.
   - すると $\sum_{i=1}^{k}c_iv_i=\mathbf{0}$ となるが，S_k は線形独立なので $c_1=\cdots=c_k=0$.

- ゆえに全ての係数が 0 となり，S_{k+1} は線形独立である．

4. この操作を繰り返す．V は有限次元であるため，このプロセスは有限回（n 回）で必ず終了し，そのとき V 全体を生成する．よって基底が得られる．

□

例題 1.4 (行列もベクトルである)**.** 2×2 の行列全体 $M_{2,2}(\mathbb{R})$ を考えます．行列 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ に対し，

- 足し算 $A + B$ ができます．
- スカラー倍 $2A$ ができます．
- ゼロ行列 O がゼロ元として機能します．

よって，**行列も「ベクトル」の一種**です．この空間の基底は $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ の 4 つなので，**次元は 4** です．

例題 1.5 (【反例】第 1 象限の世界)**.** 平面上のベクトルで，成分がすべて正のものだけの集合 $V = \{(x, y) \mid x > 0, y > 0\}$ を考えます．

- 足し算は大丈夫です（正＋正＝正）．
- しかし，スカラー倍（-1 倍）をすると，$(-x, -y)$ となり，集合からはみ出してしまいます．

よって，これは**線形空間ではありません**．「逆ベクトルが存在しない世界」は線形空間と呼べないのです．

第 2 章

線形写像と次元定理 (Linear Mappings & Rank-Nullity)

空間と空間をつなぐ「関数」について考えます．ここでの主役は，行列そのものではなく，行列が引き起こす「変換」です．

2.1 線形写像・像・核

定義 2.1 (線形写像 (Linear Mapping))**.** 2 つの線形空間 V, W に対し，写像 $f: V \to W$ が以下を満たすとき，線形写像と呼ぶ．

$$f(ax+by) = af(x) + bf(y) \quad (\forall x, y \in V, \forall a, b \in K)$$

(「足してから変換」と「変換してから足す」が等しい＝重ね合わせの理が成り立つ)

定義 2.2 (像 (Image) と 核 (Kernel))**.**

- **像**: $\mathrm{Im}\, f = \{f(x) \in W \mid x \in V\}$
  f によって写される値域．情報の「到達範囲」．
- **核**: $\mathrm{Ker}\, f = \{x \in V \mid f(x) = \mathbf{0}\}$
  f によってゼロに潰される領域．情報の「損失部分」．

例題 2.1 (微分作用素)**.** $V = P_3(\mathbb{R})$（3 次以下の多項式空間）とする．写像 $D: V \to V$ を $D(p(x)) = \frac{d}{dx}p(x)$ で定義する．

- $D(ax^2+bx+c) = 2ax+b$．これは線形写像である．
- **核**: 微分して 0 になるのは定数関数．よって $\mathrm{Ker}\, D = \{c \mid c \in \mathbb{R}\}$（1 次元）．
- **像**: 3 次式を微分すると 2 次式になる．よって $\mathrm{Im}\, D = P_2(\mathbb{R})$（3 次元）．
- 元の次元 $\dim V = 4$（基底は $1, x, x^2, x^3$）．
- 確認：$\dim(\mathrm{Im}) + \dim(\mathrm{Ker}) = 3 + 1 = 4$．次元が保存されている．

例題 2.2 (回転行列)**.** 平面ベクトルを原点中心に θ 回転させる変換 f は線形写像です．行列で書くと $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

- 像 $\mathrm{Im}\, f$：平面全体（回転しても次元は潰れない）．
- 核 $\mathrm{Ker}\, f$：$\{\mathbf{0}\}$ のみ（ゼロ以外でゼロに移るものはない）．

例題 2.3 (射影行列（影を落とす）)**.** 3 次元空間から xy 平面への射影 $f(x, y, z) = (x, y, 0)$．行列で書くと

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

- Im f : xy 平面（2 次元）.
- Ker f : z 軸上のベクトル $(0, 0, c)$（1 次元）.
- 次元定理の確認：$3(\text{全体}) = 1(\text{核}) + 2(\text{像})$.

【コラム】射影行列 ～情報を捨てる潔い行列～

射影行列 P は，$P^2 = P$（べき等性）を満たす行列と定義されます．

- **直感**: 「一度影を落としたら，その影にもう一度光を当てても，影の形は変わらない」
- **意味**: これは数学的に**「不可逆な操作（情報を捨てること）」**を定義しています．最小二乗法や量子力学の観測（波動関数の収縮）は，すべてこの射影行列の仕業です．

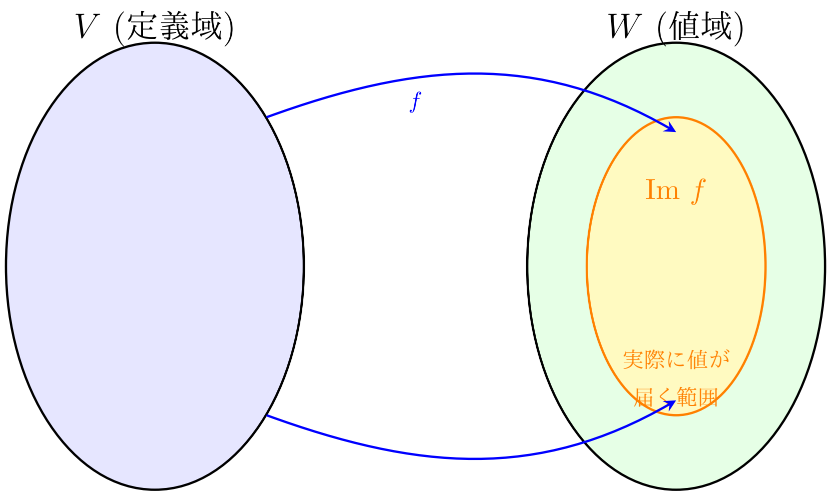


図 2.1: **像 (Im f)** のイメージ：定義域全体が写される「影」の範囲.

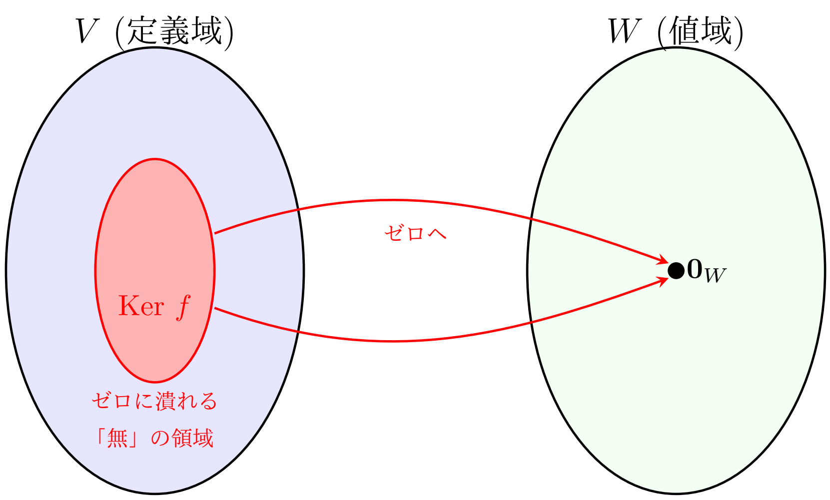


図 2.2: **核 (Ker f)** のイメージ：ゼロに潰れて消えてしまう領域.

2.2　次元定理（Rank-Nullity Theorem）

線形代数の「基本定理」です．この定理は，連立方程式の解の自由度から，データ圧縮の限界までを支配します．

定理 2.1 (次元定理)**.** 有限次元線形空間 V から W への線形写像 f について，

$$\dim V = \dim(\mathit{Ker}\ f) + \dim(\mathit{Im}\ f)$$

Proof.

1. **核の基底をとる**: $k = \dim(\mathrm{Ker}\ f)$ とする．Ker f の基底を $\{u_1, \ldots, u_k\}$ とおく．
2. **基底を延長する**: 定理 1.4（基底の延長定理）より，これに r 個のベクトル $\{v_1, \ldots, v_r\}$ を加え，V 全体の基底 $\{u_1, \ldots, u_k, v_1, \ldots, v_r\}$ を作る．ここで $\dim V = k + r$ である．我々の目標は，$\dim(\mathrm{Im}\ f) = r$ を示すことである．
3. **像の基底候補を作る**: $S = \{f(v_1), \ldots, f(v_r)\}$ という r 個のベクトルの集合を考える．これが Im f の基底になればよい．
4. **生成性の確認**: 任意の $y \in \mathrm{Im}\ f$ をとる．定義より $y = f(x)$ となる $x \in V$ が存在する．x を V の基底で展開する：
$$x = c_1u_1 + \cdots + c_ku_k + d_1v_1 + \cdots + d_rv_r$$
これに f を作用させる：
$$y = f(x) = \sum c_i \underbrace{f(u_i)}_{\mathbf{0}} + \sum d_j f(v_j) = \sum_{j=1}^{r} d_j f(v_j)$$
（$u_i \in \mathrm{Ker}\ f$ なので $f(u_i) = \mathbf{0}$ となる点が重要）よって，y は $f(v_j)$ 達の線形結合で書けるので，S は Im f を生成する．
5. **線形独立性の確認**:
$$a_1f(v_1) + \cdots + a_rf(v_r) = \mathbf{0}$$
という関係式があるとする．線形性より $f(\sum a_jv_j) = \mathbf{0}$．これはベクトル $v^* = \sum a_jv_j$ が Ker f に属することを意味する．Ker f の元は $\{u_i\}$ の線形結合で書けるので，
$$\sum_{j=1}^{r} a_jv_j = \sum_{i=1}^{k} b_iu_i$$
移項して，
$$\sum_{j=1}^{r} a_jv_j - \sum_{i=1}^{k} b_iu_i = \mathbf{0}$$
ここで $\{u_1, \ldots, u_k, v_1, \ldots, v_r\}$ は V の基底（線形独立）であった．したがって，全ての係数は 0 でなければならない．特に $a_1 = \cdots = a_r = 0$．よって S は線形独立である．
6. **結論**: S は Im f の基底であり，その数は r 個である．$\dim V = k + r = \dim(\mathrm{Ker}\ f) + \dim(\mathrm{Im}\ f)$.

□

【コラム】次元定理の図解 ～情報はどこへ消えた？～

数式 $\dim V = \dim(\mathrm{Ker}\ f) + \dim(\mathrm{Im}\ f)$ は，情報の「エネルギー保存則」のようなものです．これを「圧縮機（コンプレッサー）」に例えてみましょう．

- V **(素材):** 圧縮前のデータ．例えば「1000 個の変数」があるとします ($\dim = 1000$)．
- f **(圧縮):** データを「重要な 3 個の特徴」だけに要約する処理です．
- **Im** f **(成果物):** 出力された 3 個の特徴データ．ランクは 3 です．
- **Ker** f **(廃棄物):** 捨てられた 997 個分の情報．これが「核（カーネル）」の次元です．

次元定理は，**「入力した 1000 個の情報は，"成果物 (3)"と"廃棄物 (997)"にきれいに分かれる．行方不明になる情報はない」**ということを保証しているのです．

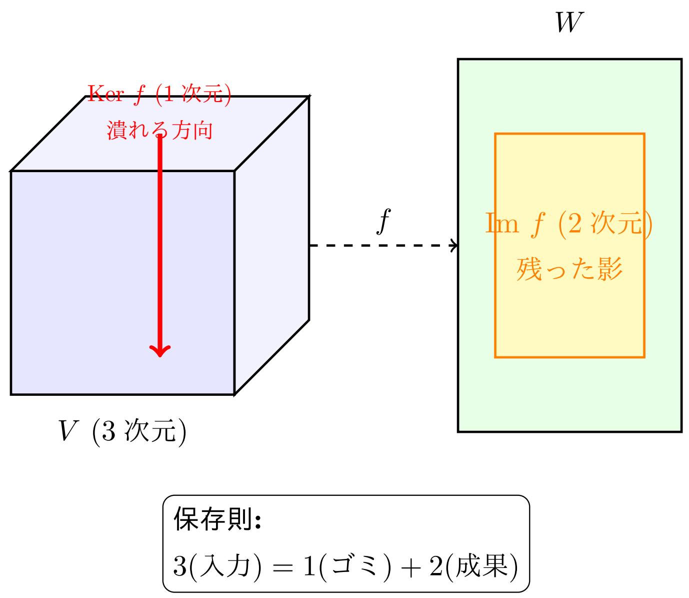


図 2.3: 次元定理のイメージ：入力空間が「潰れる方向」と「残る影」に分解される．

このイメージを持つことで，Ker を単なる「ゼロになる集合」ではなく，**「システムの不可逆性（情報を捨てること）の尺度」**として捉えられるようになります．

2.3　ランク（階数）の計算と意味

「ランク」は抽象的な概念ではなく，**ガウスの消去法（行基本変形）**を使って機械的に計算できる数値です．

定義 2.3 (ランクの 3 つの顔)**.** 行列 A のランク（rank A）は，以下の 3 つの意味を持ち，これらはすべて一致します．

1. **像の次元**: $\dim(\mathrm{Im}\ A)$（出力情報の広がり）
2. **線形独立な列の数**: 情報の重複がない列ベクトルの最大本数
3. **ピボットの数**: 行基本変形後の「階段の段数」（計算上の定義）

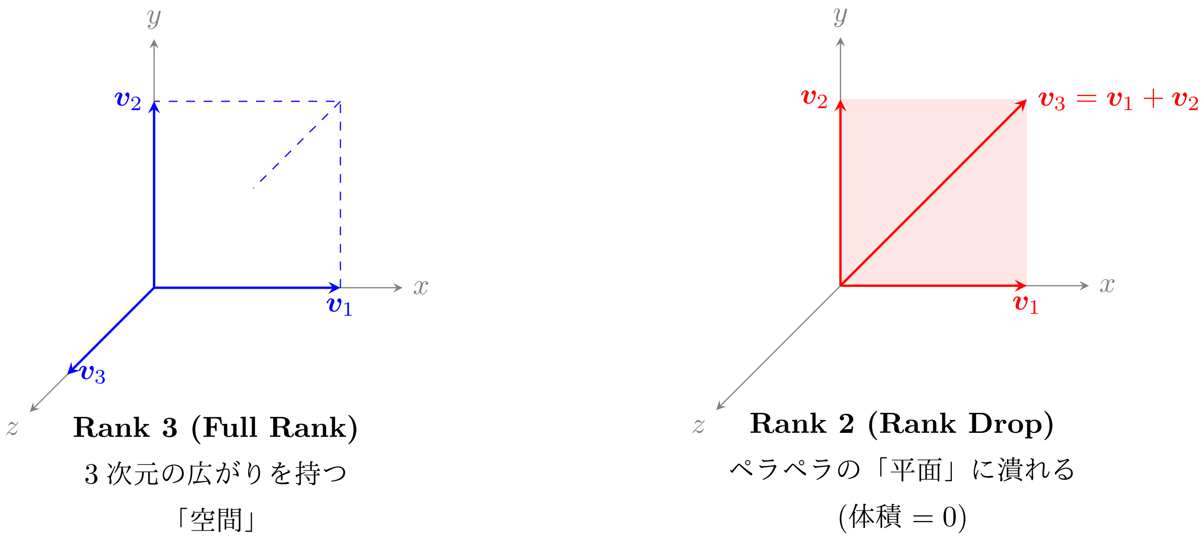


図 2.4: ランクと「情報の次元」．左は 3 本のベクトルが独立して 3 次元空間を張っている．右は 3 本目が他の 2 本のコピー（線形結合）であるため，実質的に 2 次元の情報しかなく，空間が潰れている（ランク落ち）．

例題 2.4 (ランクの実践計算)．**問題**: 次の行列 A のランクを求めよ．

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 2 \end{pmatrix}$$

解答手順: 行基本変形（掃き出し法）を行います．

1. **1 列目を消去**: 2 行目から 1 行目の 2 倍を引き，3 行目から 1 行目を引く．
$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \end{pmatrix}$$
2. **行の入れ替え**: 0 だけの行を一番下に送る．
$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$
3. **段数を数える**: 非ゼロの成分（ピボット）を持つ行は **2** つ残りました．よって，**rank** $A = 2$ です．

意味の解釈:

- 元の行列には 3 本の列ベクトルがありましたが，ランクが 2 ということは，**「実質的に 2 次元の情報しか持っていない」**（3 本目は他の 2 本の合成で作れる）ことを意味します．
- 次元定理より，$\dim(\mathrm{Ker}\,A) = 3 - 2 = 1$．入力のうち 1 次元分の情報は潰れて消えます．

2.4　部分空間と基底の構成法

「基底」は空間の「座標軸」です．計算しやすい基底を選ぶことが，問題を解く鍵になります．

例題 2.5 (平面の基底)．**問題**: ベクトル空間 $V = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$（平面）の基底を一組求めよ．

解答手順:

1. パラメータ表示する: $z = -x - y$ なので，任意のベクトルは

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ -x-y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

と書けます．

2. 基底を取り出す: パラメータ x, y についているベクトルを取り出します．

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

これらは明らかに平行ではない（線形独立）ので，V の基底となります．よって $\dim V = 2$ です．

2.5　線形写像の行列表現（補足）

(※前節で解説した「ランク」は，以下の行列表現を用いることで具体的に計算できます)

定義 2.4 (表現行列)**.** n 次元空間 V の基底 $\{e_1, \ldots, e_n\}$ と，m 次元空間 W の基底 $\{u_1, \ldots, u_m\}$ を固定する．線形写像 $f : V \to W$ が，基底の行き先として以下のように決まるとする．

$$f(e_j) = a_{1j}u_1 + a_{2j}u_2 + \cdots + a_{mj}u_m$$

この係数 a_{ij} を並べた $m \times n$ 行列 $A = (a_{ij})$ を，f の**表現行列**と呼ぶ．

- **写像** f: 「入力データを加工する機械（ブラックボックス）」
- **表現行列** A: 「その機械の設計図（数値データ）」

例題 2.6 (微分の行列表現)**.** $V = P_2(\mathbb{R})$（2 次以下の多項式空間，基底は $\{1, x, x^2\}$）とする．微分写像 $D(f) = f'$ を考える．

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

これにより，「関数の微分」という操作が「行列の掛け算」に翻訳されました．

例題 2.7 (基底を変える：同じ写像でも行列が変わる)**.** $f : \mathbb{R}^2 \to \mathbb{R}^2$ を $f(x, y) = (x + y,\ y)$ とする．

1. 標準基底: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
2. 基底 $B = \{(1, 0), (1, 1)\}$: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ （偶々同じ）

一般に**基底が変わると行列も変わる**（相似変換）．

例題 2.8 ($\mathbb{R}^3 \to \mathbb{R}^2$：標準基底・混合)**. 問題**: 線形写像 $f : \mathbb{R}^3 \to \mathbb{R}^2$ を

$$f(x, y, z) = (x + z,\ y - z)$$

で定める．標準基底に関する表現行列を求めよ．

解答: $\mathbb{R}^3$ の標準基底 $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$，$\mathbb{R}^2$ の標準基底 $u_1 = (1, 0), u_2 = (0, 1)$ を用いる．

$$f(e_1) = (1, 0), \quad f(e_2) = (0, 1), \quad f(e_3) = (1, -1).$$

よって列に並べて

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

（確認）任意の (x, y, z) に対し $A(x, y, z)^{\mathsf{T}} = (x + z,\ y - z)^{\mathsf{T}}$.

例題 2.9 ($\mathbb{R}^2 \to \mathbb{R}^2$：幾何 $\to$ 行列)**.** **問題:** $f : \mathbb{R}^2 \to \mathbb{R}^2$ を「x **軸方向は 2 倍**，y **軸方向は 3 倍**」する伸縮変換とする．標準基底に関する表現行列を求めよ．

解答: 標準基底 $e_1 = (1, 0), e_2 = (0, 1)$ について

$$f(e_1) = (2, 0), \quad f(e_2) = (0, 3).$$

したがって

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

（直感）固有ベクトルがそのまま座標軸（対角行列）になる典型例.

例題 2.10 (多項式空間：シフト作用素 $p(x) \mapsto p(x + 1)$)**.** **問題:** $V = P_2(\mathbb{R})$ とし，線形写像 $T : V \to V$ を

$$(Tp)(x) = p(x + 1)$$

で定める．基底 $(1, x, x^2)$ に関する表現行列を求めよ．

解答: 基底 $e_1 = 1, e_2 = x, e_3 = x^2$ に対し

$$T(1) = 1, \quad T(x) = x + 1, \quad T(x^2) = (x + 1)^2 = x^2 + 2x + 1.$$

各々を $(1, x, x^2)$ で座標化すると

- $T(1) = 1 \Rightarrow (1, 0, 0)$
- $T(x) = 1 \cdot 1 + 1 \cdot x + 0 \cdot x^2 \Rightarrow (1, 1, 0)$
- $T(x^2) = 1 \cdot 1 + 2 \cdot x + 1 \cdot x^2 \Rightarrow (1, 2, 1)$

よって

$$[T] = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

（直感）「x を $x + 1$ にずらす」操作が，係数ベクトルに対して上三角の“混ざり”を起こす.

例題 2.11 (行列空間 $M_{2,2} \to \mathbb{R}^2$：入力が行列)**.** **問題:** $V = M_{2,2}(\mathbb{R})$ に対し

$$f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (a + d,\ b - c)$$

で $f : V \to \mathbb{R}^2$ を定める．V の基底を

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

とし，$\mathbb{R}^2$ は標準基底とする．表現行列を求めよ．

解答: 各基底の像を計算する：

$$f(E_{11}) = (1, 0), \quad f(E_{12}) = (0, 1), \quad f(E_{21}) = (0, -1), \quad f(E_{22}) = (1, 0).$$

よって列に並べて

$$[f] = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

410: （意味）第 1 成分は $\mathrm{tr}(A) = a + d$（対角の和），第 2 成分は「上三角－下三角」の差を抜き出している.
411:

【コラム】なぜ掃き出し法（行基本変形）をやっていいのか？

ランクを求める計算（掃き出し法）は，実は**「基本行列」**と呼ばれる特殊な行列を左から次々と掛けていく操作そのものです．その正体は，**「正則行列（可逆な変換）を左から掛けている」**だけです．正則行列を掛ける操作は情報を壊さない（可逆）ため，本質的な情報量（ランク）は保存されます．

- **(I) 行の c 倍**: 情報の「拡大・縮小」.
- **(II) 行の入れ替え**: 情報の「並べ替え」.
- **(III) 行の加算**: 情報の「視点変更（ある行を基準に他の行を見る）」.

これら行列 E を左から掛けて $E_k \dots E_1 A =$ (階段行列) にしているのが掃き出し法の正体です．これを理解すると，第 10 章の行列分解（LU 分解や SVD）がスムーズに理解できます．

基本行列の 3 種 〜行列操作の「原子」〜

基本行列は，単位行列 I に「たった 1 回だけ行基本変形」を施したものです．これらは「行列操作の原子（アトム）」とも言えます．

1. 行の入れ替え（置換行列）

$$E_{\text{swap}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

単位行列の 1 行目と 2 行目を入れ替えたもの．これを左から掛けると，相手の 1 行目と 2 行目が入れ替わります．

```
[ 行１ ]      [ 行２ ]
[ 行２ ]  ->  [ 行１ ]
```

（幾何学的意味）座標軸の入れ替え（鏡映変換の一種）.

2. ある行を c 倍（スケーリング行列）

$$E_{\text{scale}} = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix} \quad (c \neq 0)$$

単位行列の 2 行目を c 倍したもの．これを左から掛けると，相手の 2 行目だけが c 倍されます．

```
[ 行１ ]      [ 行１   ]
[ 行２ ]  ->  [ c*行２ ]
```

（幾何学的意味）軸方向の伸縮．$c = 0$ は不可逆（情報を潰す）になるため禁止です．

3. 行の加算（せん断行列）

$$E_{\text{add}} = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

単位行列の 2 行目に，1 行目の c 倍を足したもの．これを左から掛けると，相手の 2 行目に 1 行目の c 倍が加算されます．

```
[ 行 1 ]            [ 行 1        ]
[ 行 2 ]    ->      [ 行 2 + c*行 1]
```

（幾何学的意味）せん断（Shear）．長方形を「平行四辺形」に歪める操作です．これがガウス消去法の主役です．

2.6　行基本変形の戦略的ドリル（5 選）

行基本変形は，闇雲に計算するのではなく，「ゴール（階段行列）」を見据えて，「情報を整理・圧縮」していくプロセスです．ここでは，典型的な 5 つのパターンを通して，その「手筋」を習得します．

例題 1：基本のキ（上から順に消す）

問題: 次の行列を行基本変形し，ランクを求めよ．

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 4 \end{pmatrix}$$

解答と戦略: 「左上の 1」をピボット（基準）にして，その下の数字をすべて 0 にします（掃除）．

1. **1** 列目の掃除: 2 行目 $-$ 1 行目 $\times 2$，3 行目 $-$ 1 行目 $\times 1$

$$\to \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

2. **2** 列目の掃除: 次は「真ん中の 1」をピボットにします．3 行目 $-$ 2 行目 $\times 1$

$$\to \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

3. 仕上げ: 非ゼロ行は 3 つ．答え: $\operatorname{rank} A = 3$（フルランク）．

例題 2：行の入れ替え（ピボットが 0 の時）

問題: 左上が 0 の場合，どうすればよいか？

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix}$$

解答と戦略: 「 0 で割る」ことはできないため，計算しやすい行を一番上に持ってきます．

1. 行の交換: 1 行目 $\leftrightarrow$ 2 行目

$$\to \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 2 & 4 & 2 \end{pmatrix}$$

2. **1** 列目の掃除: 3 行目 $-$ 1 行目 $\times 2$

$$\to \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

ポイント: 3 行目が一気に消えました．これは「3 行目が 1 行目の定数倍（情報の重複）」だったことを意味します．

3. 答え: rank $A = 2$.

例題 3：ランク落ち（情報が消える瞬間）

問題: 一見複雑に見えるが，実は情報量が少ない行列．

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

解答と戦略:

1. **1** 列目の掃除: 2 行目 $-$ 1 行目，3 行目 $-$ 1 行目 $\times 2$

$$\to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

2. 発見: 2 行目と 3 行目が全く同じ形になりました．3 行目 $-$ 2 行目 をすると，3 行目は完全に 0 になります．

$$\to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

3. 答え: rank $A = 2$. 直感: 3 行目 $=$ 1 行目 $+2$ 行目 という関係があったため，3 行目の情報は「冗長（無駄）」として消去されました．

例題 4：横長行列（変数が多すぎる場合）

問題: 式が 2 つ，変数が 4 つの場合．

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 3 & 6 & -3 & 5 \end{pmatrix}$$

解答と戦略:

1. **1** 列目の掃除: 2 行目 $-$ 1 行目 $\times 3$

$$\to \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

2. 観察: 2 行目の左側（x_2, x_3 の係数）も一緒に消えてしまいました．これは，x_2, x_3 という変数が x_1 と強い従属関係（または無関係）にあることを示唆します．
3. 答え: rank $A = 2$. 自由度の計算: 変数 4 $-$ ランク 2 $= 2$. 解空間は 2 次元の広がりを持ちます．

例題 5：パラメータを含む行列（分岐条件）

問題: 定数 a によってランクが変わるケース．

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & a \\ 1 & 4 & a^2 \end{pmatrix}$$

解答と戦略: 最後まで変形してから，a の条件を考えます．

1. **1 列目の掃除**: 2 行目 − 1 行目，3 行目 − 1 行目
$$\to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & a-1 \\ 0 & 3 & a^2-1 \end{pmatrix}$$
2. **2 列目の掃除**: 3 行目 − 2 行目 ×3．$(a^2-1)-3(a-1)=(a-1)(a+1)-3(a-1)=(a-1)(a-2)$ なので，
$$\to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & a-1 \\ 0 & 0 & (a-1)(a-2) \end{pmatrix}$$
3. **分岐**: 階段の段数が 3 段になる（ランク 3）には，右下が 0 であってはいけません．
   - **Case 1**: $a \neq 1$ かつ $a \neq 2$ のとき，rank $A=3$.
   - **Case 2**: $a=1$ または $a=2$ のとき，3 行目がすべて 0 になるので rank $A=2$.

   応用: これがファンデルモンド行列式の最小版であり，多項式補間の可解条件につながります．

【ランク計算の極意】列基本変形との違いと混合技

行列のランクを計算する際，「行基本変形」だけでなく「列基本変形」も存在します．これらはどう違い，どう使い分けるべきなのでしょうか？

1. 行変形と列変形の「意味」の違い

- **行基本変形**（左から掛ける）：**「出力空間（値域）」の座標変換**．方程式の「式」を足し引きして解きやすくする操作．連立方程式を解くときは**これしか使えません**（x の値が変わらないため）．
- **列基本変形**（右から掛ける）：**「入力空間（定義域）」の座標変換**．変数 x を別の変数 x' に置き換える操作．方程式でこれを使うと，解の意味（x の中身）が変わってしまうため，方程式を解くときには使えません．

2. ランク計算における「混合技」

ランク（情報の次元）だけを知りたい場合，**行変形と列変形を混ぜて使っても OK** です．なぜなら，ランクは「座標変換（正則行列を掛ける操作）」に対して不変だからです．
$$\mathrm{rank}(A) = \mathrm{rank}(PAQ)$$

例題 2.12 (混合変形によるランク計算)**.**
$$A = \begin{pmatrix} 1 & 100 & 1000 \\ 2 & 201 & 2005 \\ 1 & 102 & 1006 \end{pmatrix}$$
これを行変形だけでやろうとすると，桁が大きくて大変です．列変形を混ぜると「瞬殺」できます．

1. **列変形で「巨大な数」を消す**: 2 列目 − 1 列目 ×100，3 列目 − 1 列目 ×1000
$$\to \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 5 \\ 1 & 2 & 6 \end{pmatrix}$$

   （右側のゴミを一掃しました）

2. 行変形で仕上げ: 2 行目 − 1 行目 ×2，3 行目 − 1 行目

$$\to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 2 & 6 \end{pmatrix}$$

3 行目 − 2 行目 ×2

$$\to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & -4 \end{pmatrix}$$

3. 答え: 非ゼロ成分が 3 つ残ったので，rank $A = 3$.

結論:

- 方程式を解くとき: 「行変形」のみ（厳禁・列変形）.
- ランクを知りたいとき: 「行」と「列」の合わせ技が最強．特に「1 列目をピボットにして，右側の列を掃除する」と計算が劇的に楽になります.

第 II 部

スペクトル理論と対角化

第 3 章

行列式 (Determinant) —体積の幾何学

固有値 λ を求める際，我々は特性方程式 $\det(A - \lambda I) = 0$ を解きます．なぜここで「行列式」が登場するのか？ それは行列式が**「体積拡大率」**を表す指標だからです．

【コラム】行列式とランクの歴史：人間の知恵

数学の歴史を紐解くと，「行列式」と「ランク」は，現代の学ぶ順序（行列 → 行列式 → ランク）とは逆の順序や，意外な動機から生まれています．

1. 行列式 (Determinant)：連立方程式の「判別式」

「連立方程式を一発で解くための『魔法の数字』が欲しい」

実は，**行列式は行列よりも先に発見されました**．17 世紀後半，関孝和（日本）やライプニッツ（ドイツ）は，連立方程式の解の有無を一発で判定する方法を探していました．例えば，$ax + by = 0, cx + dy = 0$ が 0 以外の解を持つ条件は $ad - bc = 0$ です．彼らは，変数がふえても係数を特定のルールで計算した値がゼロになれば解が見つかることに気づき，この値を **Determinant（決定子＝行列式）**と名付けました．当初は「体積」の意味はなく，純粋な**「解の有無判定ツール」**でした．

2. ランク (Rank)：情報の「純度」

「見た目は 3 本だけど，実質何本なの？」

19 世紀後半，行列自体が研究対象になると，「式が大量にあっても，実は重複しているだけかもしれない」という問題が浮上しました．例えば，$x + y = 2$ と $2x + 2y = 4$ は実質的に同じ情報です．数学者たちは，**「行列式がゼロにならない最大のサイズ」**を調べることで，**「実質的に独立な情報の数」**を数えられることに気づきました．これを軍隊の階級になぞらえて **Rank（階数）**と呼びました．

まとめ

現代の私たちがデータ解析で使うランクや行列式も，結局は**「ビッグデータの中から『真に意味のある情報』だけを抜き出したい」**という，先人と同じ動機に基づいているのです．

3.1　定義と幾何学的意味

定義 3.1 (行列式 (Determinant))**.** n 次正方行列 $A = (a_{ij})$ に対し，行列式 $\det(A)$（または $|A|$）を以下で定義する．

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

ここで S_n は n 個の数字の置換全体，$\mathrm{sgn}(\sigma)$ は置換の符号（偶置換なら $+1$，奇置換なら-1）である．

例題 3.1 (ヤコビアンと変数変換)**.** 多重積分における変数変換 $\iint f(x,y)dxdy = \iint f(u,v)|J|dudv$ の $|J|$ はヤコビ行列の行列式です．

- $|A| = 2$: 空間の体積を 2 倍に引き伸ばす変換．
- $|A| = 0$: 空間を「ぺちゃんこ」にする変換（体積 0）．
  - 3 次元の物体を 2 次元平面に押しつぶすような変換です．一度潰れたら，元の立体には戻せません．これが「**行列式が 0** $\iff$ **逆行列が存在しない（不可逆）**」の直観的理由です．

3.2　計算テクニックと幾何学的イメージ

定義通りに計算するのは大変なので，実戦的な計算の手法と，その背後にある「世界観」を紹介します．

3.2.1　サラスの法則 (Sarrus' Rule) ：円筒状の世界

3×3 行列式の計算公式（サラスの法則）は単なる記憶術ではありません．視覚的イメージを「**円筒状の世界（トーラス）**」として捉えると，計算ミスが激減します．

1. **図解：世界を拡張する**
   3 次の行列式を計算するとき，「**最初の 2 列**」を右側にコピーして世界を拡張します．

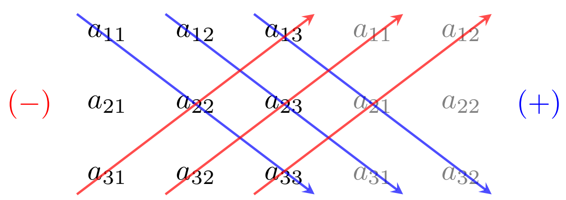


図 3.1: サラスの法則の視覚化．右側に 2 列分拡張することで，すべての対角線（ダウンヒル・アップヒル）が見えるようになる．

2. **計算の矢印：ダウンヒルとアップヒル**
   この拡張された世界に，2 種類の矢印を引きます．

- **ダウンヒル（↘）：正の項 (+)** 左上から右下へ降りる，3 本の「順風満帆なライン」．
  1. $a_{11}a_{22}a_{33}$
  2. $a_{12}a_{23}a_{31}$ （拡張部へ突入）
  3. $a_{13}a_{21}a_{32}$ （拡張部へ突入）
- **アップヒル（↙）：負の項 (-)** 右上から左下へ降りる，3 本の「逆風のライン」．
  1. $a_{13}a_{22}a_{31}$

2. $a_{11}a_{23}a_{32}$（拡張部へ突入）
3. $a_{12}a_{21}a_{33}$（拡張部へ突入）

3. 直感的イメージ：円筒（トーラス）

この「拡張」は，実は行列の左端と右端がつながっている**「円筒状の世界」**を表しています．RPG のワールドマップのように，右端に行くと左端から出てくる感覚です．サラスの法則とは，この「ループする世界」における 3 本のメインストリーム（+）と 3 本の逆流（-）の総和なのです．

注意：4 次元の罠

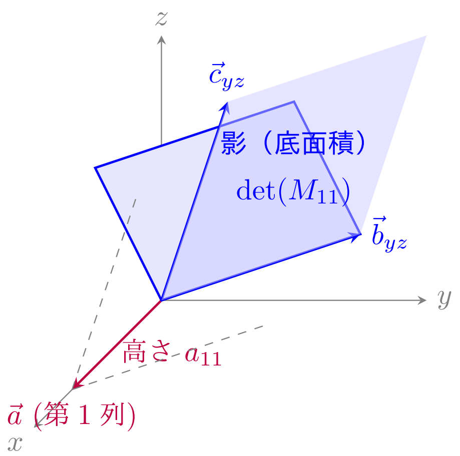
この「斜めに掛ける」視覚的テクニックは，**3 次までしか通用しません**．4 次以上の行列では，対角線の数（$4+4=8$）と項の数（$4!=24$）が合わないため，この図解は崩壊します．「4 次のサラス」は存在しないので注意しましょう．

3.2.2　ラプラス展開（余因子展開）：影による体積計算

ラプラス展開の幾何学的解釈は，**「体積（n 次元）」を「底面積（$n-1$ 次元）× 高さ」に分解して計算する方法**と言い換えると，驚くほど直感的に理解できます．

体積＝底面積×高さ

3×3 行列式（体積）を例に考えます．ベクトル $\vec{a}=(a_{11},a_{21},a_{31})^T$, $\vec{b},\vec{c}$ が張る平行六面体の体積 $\det(\vec{a},\vec{b},\vec{c})$ を，**第 1 列 $\vec{a}$ に沿って展開**します．



第 1 項の幾何学的意味:
$a_{11}\times\det(M_{11})$
$=(\vec{a}$の x 成分$)\times(yz$ 平面への影の面積$)$

図 3.2: ラプラス展開の第 1 項 $a_{11}\det(M_{11})$ の幾何学的意味．ベクトル $\vec{a}$ の「高さ（x 成分）」と，残りのベクトルが作る「影（yz 平面の面積）」の積として体積の一部を計算している．

$$\det A = \underbrace{a_{11}\det(M_{11})}_{\text{x 方向の寄与}} - \underbrace{a_{21}\det(M_{21})}_{\text{y 方向の寄与}} + \underbrace{a_{31}\det(M_{31})}_{\text{z 方向の寄与}}$$

各項の幾何学的意味

- $a_{11}\det(M_{11})$:
  - a_{11} **(高さ)**: ベクトル $\vec{a}$ の x 成分．
  - $\det(M_{11})$ **(底面積)**: 残りのベクトル $\vec{b},\vec{c}$ を yz 平面に射影したときにできる「影（平行四辺形）」の面積．

つまり，「x 方向の高さ × yz 平面の影の面積」です．

- $a_{21}\det(M_{21})$:
  - a_{21} **(高さ)**: ベクトル $\vec{a}$ の y 成分．
  - $\det(M_{21})$ **(底面積)**: $\vec{b}, \vec{c}$ を xz 平面に射影した影の面積．

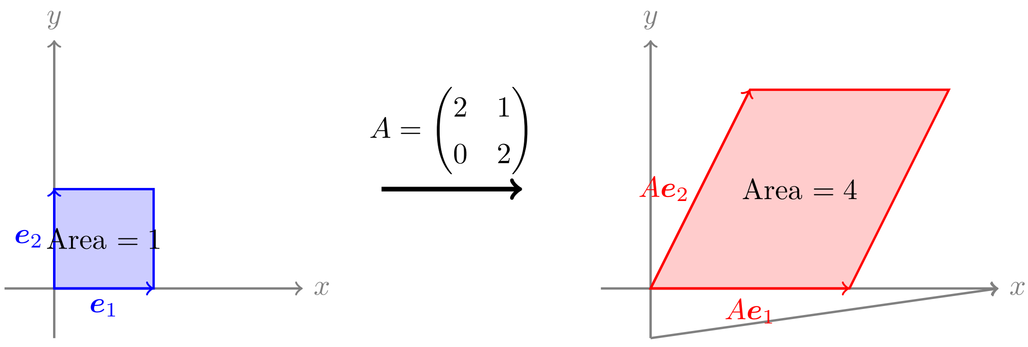
  （符号がマイナスなのは，座標系の向き（右手系・左手系）の関係です）

結論：体積の輪切り計算

ラプラス展開とは，歪んだ平行六面体の体積を，真正面から測るのではなく，**「x 軸方向の影 × x 成分」＋「y 軸方向の影 × y 成分」＋「z 軸方向の影 × z 成分」**というふうに，**投影図の面積（余因子）と奥行き（成分）の積和として計算する手法**なのです．

3.3 直感的理解：行列式の正体＝「体積拡大率」

行列式（Determinant）の幾何学的な正体は，**「線形変換によって空間の体積が何倍になるか」**を表す拡大率です．



Input Space
(単位正方形)

Output Space
(平行四辺形)

図 3.3: 行列式の意味：1×1 の正方形が，面積 4 の平行四辺形に変形された．この「面積倍率 4」こそが $\det(A) = 2 \times 2 - 1 \times 0 = 4$ の正体．

- $\det(A) = 2$: 面積が 2 倍になる変換．
- $\det(A) = 1$: 形は変わる（回転やせん断）けれど，面積は変わらない変換．（例：回転行列，直交行列）
- $\det(A) = 0$: 面積が 0 になる（ペチャンコに潰れる）変換．
  - 平行四辺形が「線」や「点」に潰れてしまった状態です．これでは逆再生（逆行列）ができません．

この「体積の変化率」というイメージを持つだけで，「積の行列式は行列式の積（$\det(AB) = \det(A)\det(B)$）」という性質が，「倍率 A のあとに倍率 B をかけたら，全体の倍率は $A \times B$ になる」という当たり前のことに思えてきます．

3.4 行列式の交代性：鏡の世界と裏返し

行列式の「交代性」は，単なる計算ルール（行を入れ替えるとマイナスがつく）以上に，**「向き（配向）の反転」**という幾何学的な意味を持っています．

定義：交代性 (Alternating Property)

行列の行（または列）を入れ替えると，行列式の符号が反転する．

$$\det(\dots, \boldsymbol{a}, \dots, \boldsymbol{b}, \dots) = -\det(\dots, \boldsymbol{b}, \dots, \boldsymbol{a}, \dots)$$

幾何学的イメージ：裏返し（Flip）

この性質は，図形を**「鏡に映す（裏返す）」**操作に対応しています．

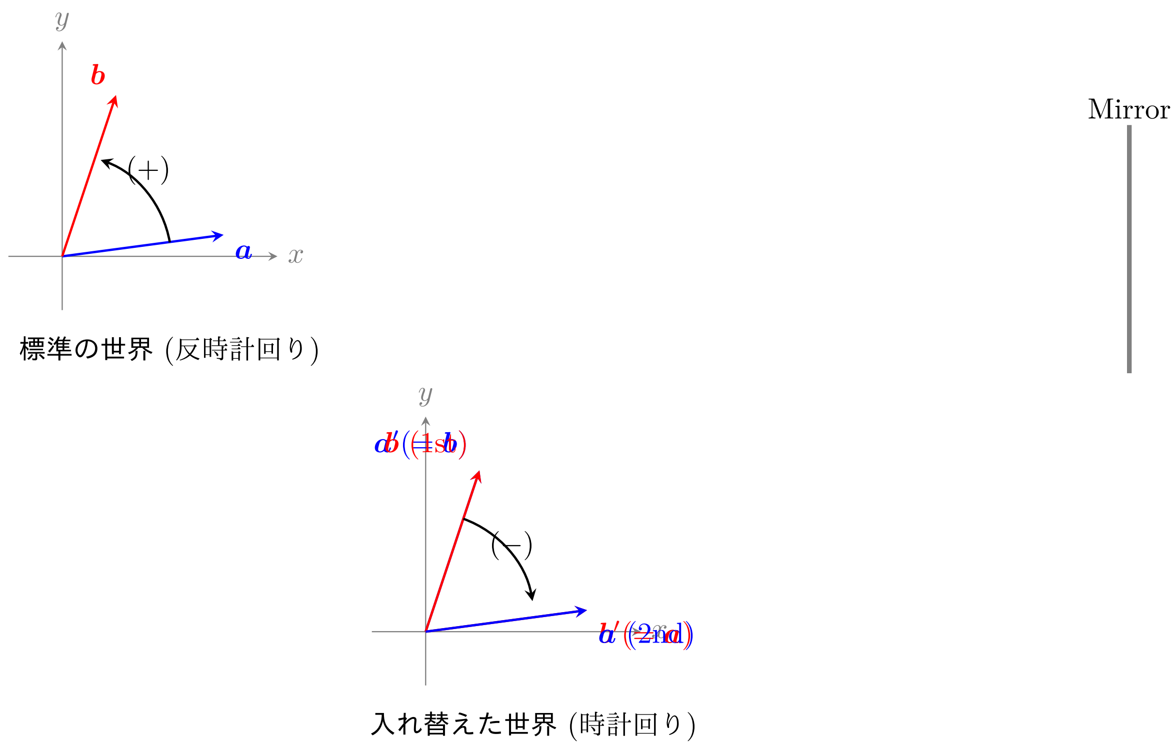


図 3.4: 行列式の交代性の幾何学的意味．ベクトルの順序を入れ替える（$\boldsymbol{a}, \boldsymbol{b} \to \boldsymbol{b}, \boldsymbol{a}$）ことは，回転の向きを逆転させることと同じであり，面積の符号が反転する．

1. **2 次元の場合（平行四辺形）**：
   - ベクトル $\boldsymbol{a}$ から $\boldsymbol{b}$ への回転が「反時計回り」だと面積はプラス．
   - $\boldsymbol{a}$ と $\boldsymbol{b}$ を入れ替えると，回転は「時計回り」になり，面積はマイナスになります．
   - これは，紙をひっくり返して裏側から見た状態と同じです．
2. **3 次元の場合（右手系と左手系）**：
   - 親指 ($\boldsymbol{a}$)，人差し指 ($\boldsymbol{b}$)，中指 ($\boldsymbol{c}$) の順序を入れ替える（例えば親指と人差し指を交換する）と，それは「右手」から「左手」に変わります．
   - 鏡に映った自分（左手系）の世界では，体積の符号が反転するのです．

同じ行があるとゼロになる理由

交代性を使うと，$\det(\boldsymbol{a}, \boldsymbol{a}) = 0$ が一瞬で証明できます．

$$\det(\boldsymbol{a}, \boldsymbol{a}) = -\det(\boldsymbol{a}, \boldsymbol{a})$$

ある数 x が $x = -x$ を満たすなら，その数は 0 しかありません．

直感的な意味: 2 つのベクトルが重なってしまったら，平行四辺形は「線」になり，体積は「0（ペチャンコ）」になります．交代性は，**「空間が潰れているかどうか」**を判定するセンサーの役割も果たしているのです．

3.5　行列式の乗法性

「変換を合成したときの倍率は，各変換の倍率の積になる」という直観を定理にします．

定理 3.1 (行列式の乗法性)**.** n 次正方行列 A, B について，以下が成り立つ．

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

Proof. 行列式を定義から直接計算して証明するのは計算量が爆発するため，**基本変形（ガウスの消去法）**の原理を用います．

1. **基本行列の性質**: 任意の行列は，基本行列 E_i（行の入れ替え，定数倍，行の加算を行う行列）の積と，階段行列 R の積に分解できる：$A = E_1E_2 \dots E_kR$．各基本行列について，$\det(EB) = \det(E)\det(B)$ が成り立つことを個別に確認する．
   - **行の入れ替え**: 行列式の符号が変わる（$\det(E) = -1$）．積も符号が変わる．成立．
   - **行の定数倍 (c 倍)**: 行列式は c 倍になる（$\det(E) = c$）．積も c 倍になる．成立．
   - **行の加算 (i 行 $+c \times j$ 行)**: 行列式は変わらない（$\det(E) = 1$）．積も変わらない．成立．
2. **可逆・非可逆による場合分け**:
   - **Case 1: A が可逆（正則）の場合** A は行基本変形で単位行列 I にできる．つまり $R = I$ であり，$A = E_1 \dots E_k$ と書ける．
     $$\det(AB) = \det(E_1 \dots E_kB)$$
     基本行列の性質を繰り返し適用して，
     $$= \det(E_1)\dots\det(E_k)\det(B)$$
     一方で $B = I$ とすれば $\det(A) = \det(E_1)\dots\det(E_k)$ なので，
     $$\det(AB) = \det(A)\det(B)$$
     が成り立つ．
   - **Case 2: A が非可逆（特異）の場合**このとき $\mathrm{rank}(A) < n$ である．積のランクの性質より $\mathrm{rank}(AB) \leq \mathrm{rank}(A) < n$．よって AB も非可逆である．非可逆行列の行列式は（体積が潰れるので）0 である．
     $$左辺 = \det(AB) = 0$$
     $$右辺 = \det(A)\det(B) = 0 \cdot \det(B) = 0$$
     よって両辺は等しい．
3. 以上より証明された．

□

第 4 章

固有値と対角化 (Eigenvalues)

行列を「座標変換」によって単純な形（対角行列）にし，システムの挙動を解析します．

【直観と起源】回転の中心にある「動かない軸」

固有ベクトル (Eigenvector) の発想の原点は，**「回転する世界の中で，止まって見える軸を探したい」**という，物理学（特に力学）の切実な要求にあります．

1. 発想の原点：オイラーの独楽（コマ）の研究

18 世紀，レオンハルト・オイラーは「剛体の回転運動」を研究していました．地球やコマのように，物体がどんなに複雑にグルグル回っていても，ある瞬間において**「回転の軸（Axis of Rotation)」**だけは動いていません．

> 「どんな回転運動も，ある軸を中心とした回転として記述できるはずだ．つまり，変換しても**向きが変わらない特別なベクトル**（回転軸）が必ず存在するのではないか？」

これが固有ベクトルの原点です．

- **Eigen**（ドイツ語）＝「固有の」「独自の」「自分自身の」
- **Eigenvector** ＝ その変換が持っている「独自の（向きが変わらない）ベクトル」

2. 直感的イメージ：変換の「骨組み」

行列 A を「空間をゴム膜のように引き伸ばしたり回転させたりする装置」だと考えます．

- **普通のベクトル:** 変換されると，向きも長さも変わってしまいます（例：青い矢印 $(1,0) \to (3,1)$ にグニャッと曲がる）．
- **固有ベクトル:** 変換されても，**向きが変わりません**．ただ長さが伸び縮みするだけです（例：赤い矢印 $(1,1) \to (4,4)$ に真っ直ぐ伸びる）．

この「向きが変わらない線（固有空間)」こそが，この変換の**「骨組み（主軸)」**なのです．

図 4.1: 固有ベクトルの直感的イメージ．普通のベクトル（左）は変換で行き先が変わってしまうが，固有ベクトル（右）は一直線上に留まり，単に定数倍されるだけである．

3. なぜこれを思いついた？

「複雑な動き（行列 A）を，単純な伸び縮み（スカラー倍 λ）の集まりに分解したい」という欲求から生まれました．

「行列 A を掛ける」という複雑な計算も，固有ベクトルの上では単なる「λ 倍」という掛け算になります．これが，**対角化（計算を簡単にする魔法）**の正体です．

4.1　固有値問題の定式化

行列 A の構造を知るために，以下の条件を満たす数 λ とベクトル $x \neq \mathbf{0}$ を探します．

$$Ax = \lambda x$$

これを変形すると $(A - \lambda I)x = \mathbf{0}$ となります．$x \neq \mathbf{0}$（非自明な解）が存在するためには，行列 $A - \lambda I$ が「情報を潰す行列（核を持つ行列）」でなければなりません．つまり，**行列式が 0 になる**必要があります．

$$\Phi_A(\lambda) = \det(\lambda I - A) = 0 \quad (\text{特性方程式})$$

例題 4.1 (共振現象)**.** ある物体を外部から揺らすとき，その周波数が固有値 λ（の虚部）と一致すると，振幅が無限大に発散する「共振」が起きます．$\det(\lambda I - A) \approx 0$ となることは，入力に対する応答（逆行列 $(\lambda I - A)^{-1}$）が無限大に近づくことを数学的に意味しています．

4.2　固有値と固有ベクトルの線形独立性

定理 4.1 (固有ベクトルの独立性)**.** 行列 A の相異なる固有値 $\lambda_1, \ldots, \lambda_m$ に対応する固有ベクトル $v_1, \ldots, v_m$ は，線形独立である．

Proof. 数学的帰納法を用いる．

1. $m = 1$ のとき，$v_1 \neq \mathbf{0}$ より明らか．
2. $m - 1$ 個まで線形独立と仮定する．
3. m 個の間に線形従属の関係があると仮定する．すなわち，全てが 0 ではない係数 c_i を用いて以下のように書けるとする．

$$c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = \mathbf{0} \quad \ldots (*1)$$

4. この式の両辺に左から行列 A を掛ける．$Av_i = \lambda_i v_i$ より，
$$c_1\lambda_1 v_1 + c_2\lambda_2 v_2 + \cdots + c_m\lambda_m v_m = \mathbf{0} \quad \ldots(*2)$$
5. また，式 (*1) 全体を λ_m 倍する．
$$c_1\lambda_m v_1 + c_2\lambda_m v_2 + \cdots + c_m\lambda_m v_m = \mathbf{0} \quad \ldots(*3)$$
6. (*2) から (*3) を引く．第 m 項が消去される．
$$c_1(\lambda_1 - \lambda_m)v_1 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = \mathbf{0}$$
7. 帰納法の仮定により，$v_1, \ldots, v_{m-1}$ は線形独立なので，各係数は 0 でなければならない．
$$c_i(\lambda_i - \lambda_m) = 0 \quad (i = 1, \ldots, m-1)$$
8. 固有値は相異なる（$\lambda_i \neq \lambda_m$）ので，$c_1 = c_2 = \cdots = c_{m-1} = 0$ となる．
9. これを (*1) に戻すと $c_m v_m = \mathbf{0}$ となるが，$v_m \neq \mathbf{0}$ なので $c_m = 0$.
10. 全ての c_i が 0 となり，線形独立性が示された．

□

例題 4.2 (主成分分析 (PCA))**.** データ群に対して，分散（情報量）が最大になる軸を探す．異なる固有値を持つ軸同士は，情報の重複がない（独立している）ため，効率的な座標系となる．

例題 4.3 (振動モード)**.** ビルや橋梁の揺れを解析する場合，行列 A は構造物の「剛性」を表します．

- 固有値 λ は「固有振動数（の 2 乗)」に対応します．
- 固有ベクトル v は「揺れの形（モード)」に対応します．

定理は，**「異なる周波数の揺れ方は，互いに独立しており，混ざり合わない」**ことを数学的に保証しています．これにより，複雑な揺れを単純な波の重ね合わせとして分解（モード分解）できます．

【実践ドリル】固有空間の基底計算（5 選）

固有空間 $W_\lambda = \mathrm{Ker}(A - \lambda I)$ の基底を求めることは，連立方程式を解いて「解のパラメータ」をベクトルの形で取り出すことと同じです．ここでは，典型的な 5 つのパターンを通して，その「手筋」と「幾何学的意味」を習得します．

パターン 1：1 次元の固有空間（基本）

問題: $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ の固有値 $\lambda = 3$ に対する固有空間の基底を求めよ．

解答:

1. **連立方程式を作る:** $(A - 3I)\boldsymbol{x} = \mathbf{0}$ を解く．
$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
2. **簡約化:** $-2x + 2y = 0 \iff x = y$.
3. **基底の抽出:** $y = t$ とおくと，$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ t \end{pmatrix} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}$．よって基底は $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

（幾何学的意味：直線 $y = x$ 上のベクトルは，変換で 3 倍に伸びる）

パターン 2：2 次元の固有空間（平面全体が固有ベクトル）

問題: $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ の固有値 $\lambda = 2$ に対する固有空間の基底を求めよ．

解答:

1. **連立方程式**: $(A - 2I)\boldsymbol{x} = \mathbf{0}$.
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
2. **条件**: $3z = 0 \iff z = 0$．x, y は自由（条件なし）．
3. **基底の抽出**: $x = s, y = t$ とおくと，
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} s \\ t \\ 0 \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
よって基底は $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

（幾何学的意味：xy 平面上のすべてのベクトルが，変換で 2 倍される）

パターン 3：重解だが 1 次元（対角化不可能）

問題: $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ の固有値 $\lambda = 2$（重解）に対する固有空間の基底を求めよ．

解答:

1. **連立方程式**:
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
2. **条件**: $y = 0$．x は自由．
3. **基底の抽出**: $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \end{pmatrix}$．よって基底は $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ だけ．

（解説：固有値は重解（代数的重複度 2）だが，固有空間の次元（幾何学的重複度）は 1 しかない．これがジョルダン標準形が必要になるサイン）

パターン 4：虚数の固有空間（回転行列）

問題: $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ の固有値 $\lambda = i$ に対する固有空間の基底を求めよ．

解答:

1. **連立方程式**: $(A - iI)\boldsymbol{x} = \mathbf{0}$.
$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
2. **簡約化**: 1 行目を i 倍すると $1 \cdot x - i \cdot y = 0$．2 行目と同じ式になる．よって $x - iy = 0 \iff x = iy$.

3. **基底の抽出**: $y=1$ とおくと $x=i$．よって基底は $\begin{pmatrix} i \\ 1 \end{pmatrix}$.

（解説：複素ベクトル空間では，回転も「対角化」できる）

パターン 5：変数が混ざるケース（対称行列）

問題: $A=\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ の固有値 $\lambda=-1$ に対する固有空間の基底を求めよ．

解答:

1. **連立方程式**: $(A-(-1)I)\boldsymbol{x}=\mathbf{0}$.
$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
2. **条件**: $x+y+z=0$ のみ．自由度は 2.
3. **基底の抽出**: $y=s, z=t$ とおくと $x=-s-t$.
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} -s-t \\ s \\ t \end{pmatrix}=s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}+t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$
よって基底は $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

（解説：対称行列なので，もう一つの固有値 $\lambda=2$ の固有ベクトル $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ と直交していることを確認せよ）

【コラム】トラブルシューティング 〜固有ベクトルが見つからないときは？〜

固有値 λ を求めた後，連立方程式 $(A-\lambda I)\boldsymbol{x}=\mathbf{0}$ を解いたら $\boldsymbol{x}=\mathbf{0}$ （ゼロベクトル）しか残らなかった．．．．そんなときは，以下の 3 つのポイントを疑ってください．

1. 固有値の計算ミス（最重要）

固有ベクトルが見つからない理由の 9 割はこれです．固有値 λ が間違っていると，行列 $A-\lambda I$ は正則（逆行列を持つ）になってしまい，方程式の解は $\boldsymbol{x}=\mathbf{0}$ だけになります．

- **確認法**: 求めた λ を特性方程式 $\det(A-\lambda I)$ に代入して，本当に 0 になるか検算しましょう．
  - $\det=0$ にならない → **固有値の計算ミス**です．
  - $\det=0$ になる → 次のステップへ．

2. 「移項」の符号ミス

連立方程式を立てる際，成分の引き算を間違えていませんか？

- 正しい式:

$$(A - \lambda I) = \begin{pmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{pmatrix}$$

- よくあるミス: 対角成分以外から引いている，引くのを忘れている，符号を逆（$\lambda - a_{11}$）にしている（方程式としては合っていますが，計算ミスの元です）．

3. 「勝手に変数を消している」

方程式を解く際，**「不定（解が無数にある）」**状態に不安を感じて，勝手に変数を 0 にしてしまうケースです．

- 例: $2x + 2y = 0$ という式が出た．
  - 間違い: 「$x = 0, y = 0$ も成り立つから，これでいいや」 $\to$ これは自明解です．
  - 正解: 「$x = -y$ だから，$y = t$ とおいて自由に動かそう」 $\to$ $\begin{pmatrix} -1 \\ 1 \end{pmatrix} t$ が基底になります．
- 鉄則: 式が足りない（ランク落ちしている）ときは，**足りない分の変数をパラメータ（s, t など）に置き換えること**．決して 0 にしてはいけません．

チェックリスト（心の救急箱）

症状	原因	対処法
解が $\boldsymbol{x} = \mathbf{0}$ だけになる	固有値 λ が違う	特性方程式 $\det(A - \lambda I) = 0$ を再計算
式が全部消えて $0 = 0$ になった	正常です	残った変数をパラメータ t と置け
固有ベクトルが分数だらけ	計算は合っているかも	整数倍してきれいにせよ（定数倍は自由）
重解なのに基底が 1 つしかない	正常（対角化不能）	ジョルダン標準形の出番です（第 7 章へ）

4.3　対角化の十分条件

行列 A が対角化できるとは，「固有ベクトルたちを基底として選べば，変換が単純な定数倍に見える」ということです．

定理 4.2 (対角化の判定条件)**.** n 次正方行列 A が n 個の線形独立な固有ベクトル $v_1, \ldots, v_n$ を持つならば，A は対角化可能である．すなわち，正則行列 $P = (v_1 \ldots v_n)$ を用いて以下のように変形できる．

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

標準基底で見ると...
グニャッと歪んでいる
（非対角成分がある）

固有基底で見ると...
軸に沿って伸縮するだけ
（対角行列になる！）

図 4.2: 対角化の幾何学的イメージ．固有ベクトルを座標軸（基底）に選ぶことで，複雑な行列変換を「グリッド（座標軸）に沿った素直な伸縮」として捉え直すことができる．

Proof. 1. **行列の積を作る**: 固有ベクトルの定義 $Av_i = \lambda_i v_i$ を n 本並べて行列形式で書く．

$$A(v_1 \dots v_n) = (\lambda_1 v_1 \dots \lambda_n v_n)$$

ここで $P = (v_1 \dots v_n)$ と置くと，左辺は AP である．右辺は，各列が λ_i 倍されているので，行列の積として以下のように分解できる．

$$(\lambda_1 v_1 \dots \lambda_n v_n) = (v_1 \dots v_n) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = PD$$

よって，$AP = PD$ が成り立つ．

2. **正則性の利用**: 仮定より $v_1, \dots, v_n$ は線形独立なので，これらを並べた行列 P は可逆（正則）である．つまり P^{-1} が存在する．
3. 左から P^{-1} を掛ければ，

$$P^{-1}AP = D$$

となり，対角行列が得られる．

□

例題 4.4 (フィボナッチ数列の一般項)**.** 漸化式 $a_{n+2} = a_{n+1} + a_n$ を行列で解く．$\begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$ と書ける．$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ とすると，状態は $x_n = A^n x_0$ で求まる．A の固有値を計算すると $\lambda = \frac{1\pm\sqrt{5}}{2}$（黄金比）．$A$ を対角化して $A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}$ とすれば，$A^n = P \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} P^{-1}$ となり，行列の累乗が単純なスカラーの累乗になります．これがフィボナッチ数列の一般項（ビネの公式）の正体です．

例題 4.5 (回転の固有値)**.** 90 度回転行列 $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ の固有値を求めよ．

- 特性方程式：$\det(A - \lambda I) = \lambda^2 + 1 = 0$.
- 解：$\lambda = i, -i$（虚数）.
- 意味：実数の世界では「回転して向きが変わらないベクトル」は存在しません．しかし複素数の世界まで拡張すると固有値が存在します．

例題 4.6 (三角行列の固有値)**.** 上三角行列 $A = \begin{pmatrix} 2 & 5 & 1 \\ 0 & 3 & 8 \\ 0 & 0 & 1 \end{pmatrix}$ の固有値を求めよ．

- 特性方程式：$(2-\lambda)(3-\lambda)(1-\lambda) = 0$.
- 解：$\lambda = 2, 3, 1$.
- ポイント：三角行列の固有値は，**対角成分そのもの**になります．計算不要で一瞬で求まります．

4.4　対角化 vs 三角化の幾何学的イメージ

「対角化」と「三角化（ジョルダン標準形への布石）」の違いを視覚的に理解しましょう．

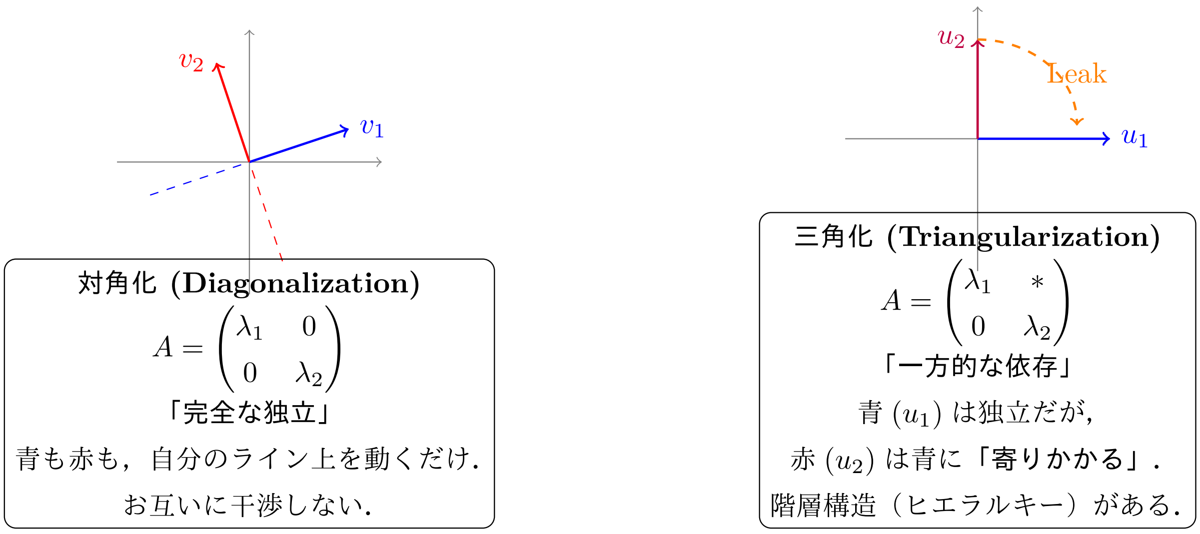


図 4.3: 基底の取り方による行列の見え方の違い．対角行列は「分離」を，三角行列は「従属」を表す．

4.5　行列式と固有値の幾何学的関係

行列式のイメージ（体積拡大率）と，固有値（各軸方向の伸縮率）の関係は，図にすると**「長方形の面積公式」**のようにシンプルになります．

コンセプト：行列式＝すべての固有値の積

$$\det A = \lambda_1 \times \lambda_2 \times \cdots \times \lambda_n$$

この数式は，幾何学的には**「体積変化率（行列式）は，各軸方向の伸縮率（固有値）の掛け算で決まる」**ということを意味しています．

図解：長方形の変形

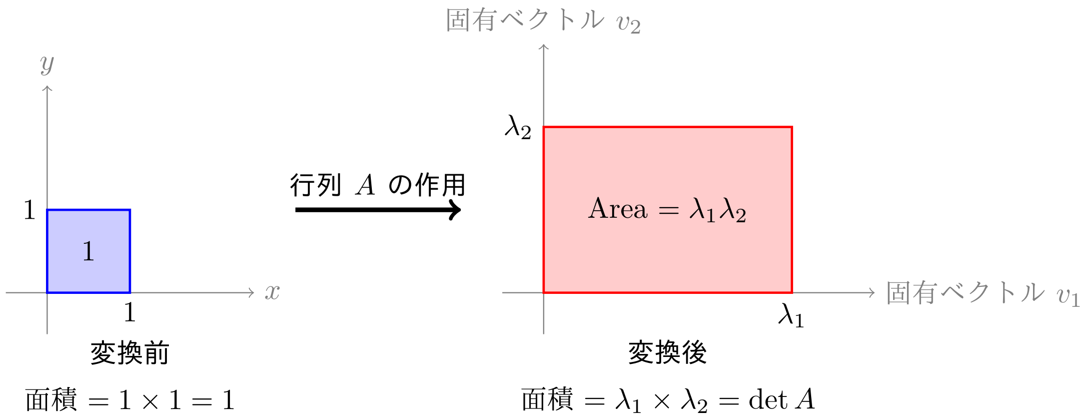


図 4.4: 各固有ベクトル方向に λ_i 倍されるため，全体の面積（体積）はそれらの積倍になる．

幾何学的解説

1. **固有ベクトルの軸**: 対角化できる行列の場合，固有ベクトル v_1, v_2 の方向は，互いに独立な「新しい座標軸」とみなせます．
2. **各軸ごとの伸縮**: 行列 A の作用は，固有ベクトルの方向に見ると単純です．
   - v_1 方向に λ_1 倍伸ばす．
   - v_2 方向に λ_2 倍伸ばす．
3. **全体の面積変化**: 縦に λ_1 倍，横に λ_2 倍引き伸ばされたゴム膜の面積は，当然 $\lambda_1 \times \lambda_2$ 倍になります．これが**行列式** $\det A$ の正体です．

- **$\det A = 0$ のとき**: 少なくとも 1 つの固有値 λ が 0 です．「ある一方向の厚みがゼロになる（ペチャンコになる）」ため，体積もゼロになります．
- **$\det A < 0$ のとき**: 固有値の積がマイナスです（例：$\lambda_1 = 2, \lambda_2 = -1$）．「片方が伸びて，もう片方が裏返る（反転）」ため，図形全体が裏返しになります．

【コラム】エルミート行列 ～実直で裏切らない行列たち～

- **定義**: $A^* = A$（複素共役転置が自分自身と等しい）．実数なら対称行列 $A^\top = A$.
- **性質**:
  - 固有値はすべて実数である．
  - 異なる固有値に対応する固有ベクトルは直交する．

なぜ物理学（量子力学）ではエルミート行列ばかり使うのか？量子力学では，世界を複素数（波動関数）で記述します．しかし，私たちが実験で観測する「エネルギー」や「運動量」といった物理量は，必ず**実数**として観測されます（複素数のエネルギーなんてありません）．

複素数の世界に住んでいるのに，結果だけは必ず実数で返してくれる．この「実直さ」を持つ行列こそがエルミート行列です．これを**「観測可能量（オブザーバブル）」**と呼びます．

数学的には，エルミート行列による変換は，空間を「回転」させず，ある直交する軸に沿って「伸縮」させるだけの**純粋な変形**（スペクトル分解）を意味します．これが最も素直で扱いやすい行列のクラスです．

第 III 部

ジョルダン標準形とシステムのダイナミクス

第 5 章

ケーリー・ハミルトンの定理と最小多項式

ジョルダン標準形の証明には，「行列を多項式に代入してゼロにする」という操作が不可欠です．その基礎となる定理を証明します．

5.1 ケーリー・ハミルトンの定理 (Cayley-Hamilton Theorem)

定理 5.1 (ケーリー・ハミルトンの定理)**.** n 次正方行列 A の特性多項式を $\Phi_A(t) = \det(tI - A) = t^n + c_{n-1}t^{n-1} + \cdots + c_0$ とする．この多項式に行列 A 自身を代入すると，ゼロ行列になる．

$$\Phi_A(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_0 I = O$$

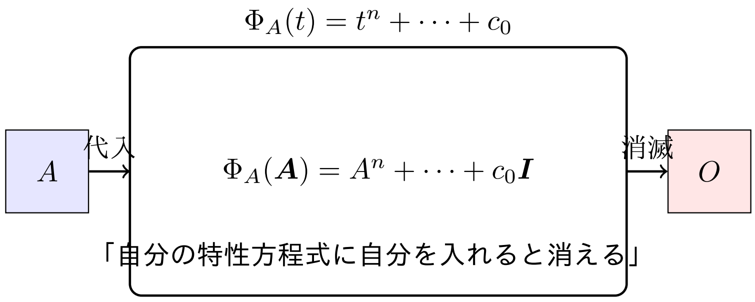


図 5.1: ケーリー・ハミルトンの定理のイメージ．行列 A は，自身の特性多項式 Φ_A という「型」にぴったりハマって消滅する．

Proof.「$\Phi_A(t) = \det(tI - A)$ なので，t に A を代入すれば $\det(A - A) = \det(O) = 0$」という説明は誤りです（スカラと行列を混同している）．厳密には以下のように証明します．

1. **余因子行列の導入**: 行列 $B(t) = tI - A$ を考える．$B(t)$ の成分は t の 1 次式である．クラメルの公式の変形より，余因子行列 $\mathrm{adj}(B(t))$ を用いて以下が成り立つ．

$$B(t) \cdot \mathrm{adj}(B(t)) = \det(B(t))I = \Phi_A(t)I$$

2. **多項式展開**: $\mathrm{adj}(B(t))$ の各成分は，$B(t)$ の $(n-1)$ 次小行列式なので，t についての $(n-1)$ 次多項式である．よって，行列係数 B_k を用いて以下のように展開できる．

$$\mathrm{adj}(tI - A) = t^{n-1}B_{n-1} + t^{n-2}B_{n-2} + \cdots + B_0$$

3. **恒等式の比較**: 元の式 $(tI - A)(t^{n-1}B_{n-1} + \cdots + B_0) = (t^n + c_{n-1}t^{n-1} + \cdots + c_0)I$ の両辺の t^k の係数を比較する．
   - t^n の係数: $B_{n-1} = I$

- t^{n-1} の係数: $B_{n-2} - AB_{n-1} = c_{n-1}I$
- $\vdots$
- t^1 の係数: $B_0 - AB_1 = c_1 I$
- t^0 の係数: $-AB_0 = c_0 I$

4. **望遠鏡和（Telescoping Sum）**：上から順に，$A^n, A^{n-1}, \ldots, I$ を左から掛けて足し合わせる．
   - $A^n B_{n-1} = A^n$
   - $A^{n-1}(B_{n-2} - AB_{n-1}) = c_{n-1}A^{n-1}$
   - $\ldots$
   - $A(B_0 - AB_1) = c_1 A$
   - $-AB_0 = c_0 I$

   これらを全て足すと，左辺は隣り合う項が打ち消し合い（$A^k B_{k-1}$ と $-A^k B_{k-1}$），最終的にゼロ行列 O になる．右辺は $\Phi_A(A)$ そのものになる．よって $\Phi_A(A) = O$ が示された．

□

例題 5.1 (自分の過去に従うシステム)**.** この定理は，**「現在の状態 A^n は，過去の状態 $A^{n-1}, \ldots, I$ の線形結合で書ける」**ことを意味します．これは時系列解析において，自己回帰モデル（AR モデル）が成立する数学的根拠であり，行列の 100 乗などを計算する際に計算量を劇的に減らすアルゴリズムの基礎となります．

【コラム】ハミルトンの発見 ～逆行列を「足し算」で求める～

ケーリー・ハミルトンの定理 $\Phi_A(A) = O$ は，単なる恒等式ではありません．これは**「行列自身の情報（スペクトル）を使えば，逆行列計算という『割り算』を，行列の『積と和』だけで実行できる」**ことを示唆しています．

逆行列の公式

定数項 c_0 が非ゼロ（つまり $\det A \neq 0$）のとき，定理の式を変形します．

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = O$$

$c_0 I$ 以外を移項して，A で括り出します：

$$A(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I) = -c_0 I$$

両辺を $-c_0$ で割り，A^{-1} を掛け合わせる（あるいは単に $A \times (\ldots) = I$ の形と見る）と，以下の公式が得られます．

$$A^{-1} = -\frac{1}{c_0}(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I)$$

意味

通常，逆行列 A^{-1} を求めるには「掃き出し法」などの割り算的な手続きが必要です．しかしこの公式は，**「自分の冪乗（$A, A^2, \ldots$）を足し合わせるだけで，自分を打ち消す存在（逆行列）を作れる」**と言っています．これは，行列が「自己完結した世界（閉じた代数系）」を持っていることの現れです．

第 6 章

広義固有空間 (Generalized Eigenspace)

対角化できない行列に対処するため，固有ベクトルの概念を拡張します．

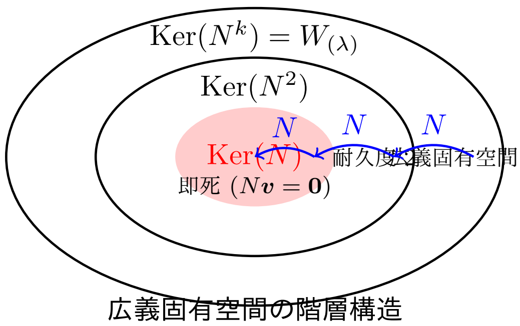
6.1　広義固有ベクトルの定義

対角化できない行列の典型例（ジョルダン細胞）を考えます．$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ この行列に対し，固有ベクトル $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ は 1 つしかありません．しかし，$(J - \lambda I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ は，2 乗するとゼロ行列になります（冪零行列）．

定義 6.1 (広義固有ベクトル)**.** 行列 A と固有値 λ に対し，ある自然数 k が存在して

$$(A - \lambda I)^k v = \mathbf{0}$$

となるとき，v を広義固有ベクトルと呼ぶ．これら全体の集合を**広義固有空間** $W_{(\lambda)}$ と呼ぶ．



広義固有空間の階層構造
「攻撃 $N = A - \lambda I$」を受けるたびに
内側の核へ落ちていく

図 6.1: 広義固有空間の仕組み．核が入れ子構造（ネスト）になっており，外側のベクトルも演算 N を繰り返すことで中心の核（真の固有空間）に吸い込まれていく．

【直観】広義固有空間＝「ゾンビ感染」の階層構造

方程式 $(A - \lambda I)^k v = 0$ だけ見てもイメージが湧きにくい「広義固有空間」．これを**「ゾンビ感染」**や**「耐久度付きの壁」**として理解しましょう．

1. 固有ベクトル＝「即死するベクトル」

行列 A の固有値 λ に対応する行列を $N = A - \lambda I$ とします．通常の固有ベクトル v は，この特殊的攻撃 N を**たった 1 回**受けただけでゼロ（無）になります．

$$Nv = \mathbf{0}$$

いわば，**「N という攻撃に耐性がない（即死する）弱いベクトル」**です．これが「真の固有空間」です．

2. 広義固有ベクトル＝「耐久力のあるベクトル」

しかし，世の中にはもっと「しぶとい」ベクトルがいます．

- N を 1 回掛けてもゼロにならない（$Nv \neq \mathbf{0}$）．
- でも，2 回掛けるとゼロになる（$N^2 v = \mathbf{0}$）．
- あるいは，100 回掛けてやっとゼロになる．

このように，**「何度か攻撃すれば最終的に消滅するベクトルたち」**をすべて集めたのが**広義固有空間**です．

3. 階層構造（タマネギの皮）

広義固有空間は，耐久度ごとに層になっています．

- **レベル 1 (真の固有空間)**: $Nv = \mathbf{0}$ （1 発で消える）
- **レベル 2**: $N^2 v = \mathbf{0}$ （2 発以内で消える．レベル 1 も含む）
- **レベル 3**: $N^3 v = \mathbf{0}$ （3 発以内で消える）
- ...
- **最大レベル (広義固有空間)**: $N^k v = \mathbf{0}$ （いつかは消える）

対角化できない行列とは，**「レベル 1（真の固有ベクトル）だけでは空間全体を埋め尽くせない行列」**のことです．だから，レベル 2，レベル 3... の「しぶといベクトル」も動員して，なんとか空間全体をカバーしようとする．これが広義固有空間分解の正体です．

4. バケツリレー（ジョルダン細胞の正体）

レベル 3 のベクトル v_3 に N を掛けるとどうなるでしょうか？

$$v_3(\text{耐久 } 3) \xrightarrow{N} v_2(\text{耐久 } 2) \xrightarrow{N} v_1(\text{耐久 } 1) \xrightarrow{N} \mathbf{0}(\text{消滅})$$

攻撃を受けるたびに「耐久度」が減っていき，最後は真の固有ベクトルになって消滅します．この**「死へのカウントダウン（連鎖）」**こそが，ジョルダン細胞の構造そのものなのです．

6.2　広義固有空間分解定理

定理 6.1 (空間の完全分解)**.** 複素ベクトル空間 V は，行列 A の相異なる固有値 $\lambda_1, \dots, \lambda_m$ に対応する広義固有空間の直和に分解される．

$$V = W_{(\lambda_1)} \oplus W_{(\lambda_2)} \oplus \cdots \oplus W_{(\lambda_m)}$$

しかも，各 $W_{(\lambda_i)}$ は A に関して不変部分空間（A を掛けてもその空間からはみ出さない）である．

【直感的イメージ】証明の心：万能フィルターによる「ケーキの切り分け」

広義固有空間分解の証明は，数式だけだと非常に抽象的ですが，その本質は**「混ぜこぜになったジュース（ベクトル）から，成分ごとのフィルターを使って，100% 果汁（広義固有空間）を分離抽出する」**プロセスです．

1. **目標：空間全体をきれいに分割したい**全空間 V を，それぞれの固有値に対応する「広義固有空間 W_{λ_i}」に切り分けたい．
$$V = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_k}$$
2. **道具：多項式フィルター**証明の鍵となるのは，行列 A を使って作れる**「特定の成分だけを通すフィルター（射影作用素）」**です．
固有多項式 $\Phi(t) = (t - \lambda_1)^{n_1} \dots (t - \lambda_k)^{n_k}$ を考えます．ここから，λ_i 以外の成分をすべて消してしまう多項式 $f_i(A)$ を作ります．
   - **フィルター P_1 の作り方:** $\Phi(t)$ から $(t - \lambda_1)^{n_1}$ だけを取り除いたものを使います．これを行列 A に代入した P_1 は，$\lambda_2, \lambda_3, \dots$ の成分をすべて「0（消滅）」にします．結果として，λ_1 **の成分（広義固有空間）だけが生き残ります．**
3. **「1 の分解」：すべてを足すと元通り**ベズーの等式（多項式の互除法）を使うと，これらのフィルターをうまい具合に係数調整して足し合わせることで，**「恒等変換 I（何もしない変換）」**を作ることができます．
$$I = Q_1(A) + Q_2(A) + \cdots + Q_k(A)$$
任意のベクトル $\boldsymbol{v}$ にこれを掛けると：
$$\boldsymbol{v} = I\boldsymbol{v} = \underbrace{Q_1(A)\boldsymbol{v}}_{\in W_{\lambda_1}} + \underbrace{Q_2(A)\boldsymbol{v}}_{\in W_{\lambda_2}} + \cdots + \underbrace{Q_k(A)\boldsymbol{v}}_{\in W_{\lambda_k}}$$
4. **結論：成分抽出**この式は，**「どんなベクトル $\boldsymbol{v}$ も，それぞれの広義固有空間の成分に『ろ過』して取り出せる」**ことを意味しています．これが $V = \bigoplus W_{\lambda_i}$ の証明の正体です．

Proof. 証明の概略は以下の通りです．

1. **核の安定化:** 写像 $N = A - \lambda I$ を考える．核の列 $\mathrm{Ker}(N) \subset \mathrm{Ker}(N^2) \subset \dots$ は単調増加するが，次元有限なのでいつか停止する．その最大空間が $W_{(\lambda)}$ である．
2. **互いに素な多項式による分解:** 特性多項式を $\Phi_A(t) = (t - \lambda_1)^{n_1} \dots (t - \lambda_m)^{n_m}$ と因数分解する．多項式 $f_i(t) = \Phi_A(t)/(t - \lambda_i)^{n_i}$ を考える（λ_i の成分だけ取り除いたもの）．これら $f_1, \dots, f_m$ は互いに素（公約数を持たない）なので，ベズーの等式より多項式 $g_1, \dots, g_m$ が存在して $\sum g_i(t) f_i(t) = 1$ となる．
3. **射影作用素の構成:** $P_i = g_i(A) f_i(A)$ と置くと，$\sum P_i = I$（単位行列の分解）となる．ケーリー・ハミルトンの定理を利用して $P_i P_j = O (i \neq j)$ および $P_i^2 = P_i$ を示せる．つまり P_i は空間を $W_{(\lambda_i)}$ へ射影する作用素となり，空間全体が直和分解されることが証明される．

□

第 7 章

ジョルダン標準形 (Jordan Canonical Form)

広義固有空間分解とは？

「真の固有ベクトル」だけではスカスカで支えきれない空間を，「広義固有ベクトル（準・固有ベクトル）」まで動員して，**空間全体を余すところなく自分の領土として分割する**定理のことです．これにより，複雑な行列も「ブロックごとの独立した世界」に切り分けて考えることができるようになります．

【直感的イメージ】不完全な柱を補強して部屋を仕切る

(a) 失敗：対角化できない場合
(真の固有空間だけでは隙間だらけ)

(b) 成功：広義固有空間分解
(空間全体を余すことなく分割)

図 7.1: 広義固有空間分解のイメージ．「線（1 次元）」の固有空間を「面（多次元）」の広義固有空間に拡張することで，空間全体 V を隙間なく埋め尽くすことができる．

【直感的イメージ】広義固有空間＝「植物の根」

ジョルダン標準形の核心は，**「対角化できない（＝壊れている）部分を修理して，空間全体を再構築する」**ことにあります．

1. 固有空間の役割：「空間の骨組み」

対角化できる行列（健全な行列）において，固有空間は**「空間を支える独立した柱」**です．空間全体 V は，それぞれの固有値に対応する固有空間 E_λ の直和（きれいな分割）で表現でき，どの方向もいずれかの固有空間に属するか，その組み合わせで表せます．

2. ジョルダン標準形での役割：「足りない柱の補強」

対角化できない行列（病的な行列）では，真の固有空間 E_λ だけでは次元が足りず，空間全体を支えきれません（柱が折れている状態）．そこで，**「広義固有空間 W_λ」**が役割を果たします．

- **役割**: 真の固有空間（核）の周囲に，「準・固有空間（雲のような広がり）」を形成し，不足している次元を埋め合わせる．
- **構造**: $V = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_m}$

ジョルダン標準形の理論は，**「真の固有ベクトルだけではスカスカな空間を，広義固有ベクトルという充填剤で埋めて，再び完全な直和分解を作る」**試みなのです．

3. 図解イメージ：「植物の根」

この役割を直感的に理解するために，**「植物の根」**のメタファーを使います．

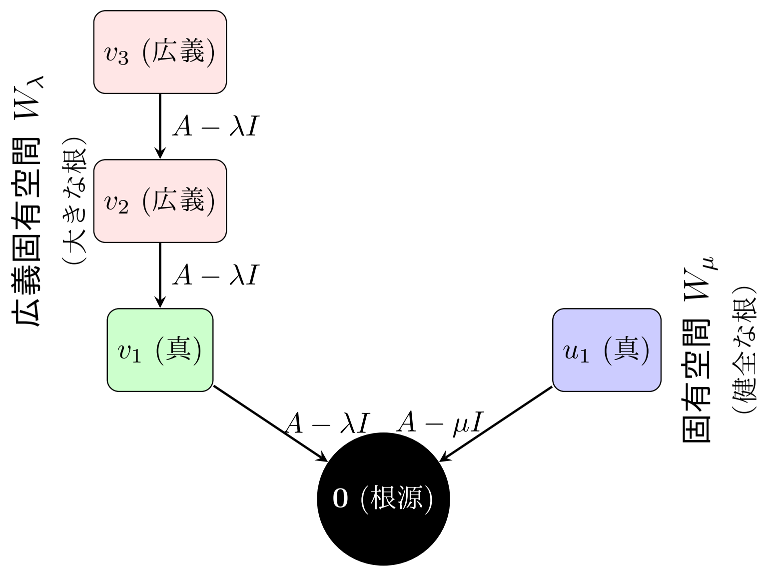


図 7.2: ジョルダン鎖の「植物の根」イメージ：真の固有ベクトル v_1 を起点に，広義固有ベクトル v_2, v_3 が茎のように伸びて空間を支えている．

- **真の固有ベクトル (v_1)**: 根っこの先端．これ以上変換してもゼロ（土）に行き着くだけ．対角化できる場合は，この「先端」しか存在しません．
- **広義固有ベクトル (v_2, v_3)**: 根の茎や幹．v_3 は v_2 に支えられ，v_2 は v_1 に支えられています．この**「支え合い（連鎖）」**全体が 1 つのブロック（ジョルダン細胞）を形成しています．

【直感的イメージ】バケツリレーとしてのジョルダン標準形

前述の通り，ジョルダン標準形の内部構造（細胞の中身）は**「バケツリレー」**として理解できます．

1. 対角化できる世界（理想郷）

対角化できる行列 A では，空間全体が「独立した固有ベクトルたち」によって張られます．これらは互いに干渉せず，それぞれの軸上でただ「伸び縮み（λ 倍）」するだけです．

- **イメージ: 「全員がソリスト」**．各楽器（基底）が勝手に自分のパート（固有値）を演奏している状態です．

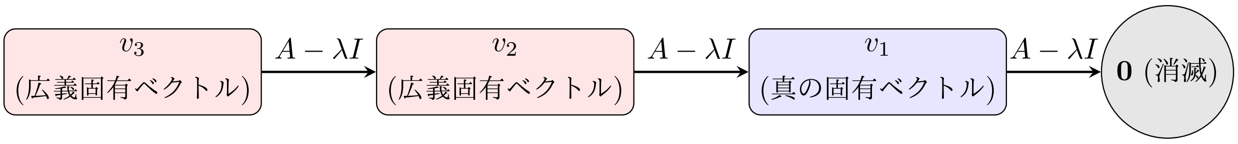
2. ジョルダン細胞の世界（現実）

固有ベクトルが足りない（縮退している）場合，空間は「伸び縮みする軸」だけでは支えきれません．そこで，**「隣のベクトルに寄りかかるベクトル（広義固有ベクトル）」**が登場します．

ジョルダン細胞 $J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$ に対応する基底 v_1, v_2, v_3 の関係を見てみましょう．

$$
\begin{aligned}
(A-\lambda I)v_1 &= \mathbf{0} \quad &&（v_1 \text{ は真の固有ベクトル：ここで止まる}） \\
(A-\lambda I)v_2 &= v_1 \quad &&（v_2 \text{ は変換されると } v_1 \text{ に成分を渡す}） \\
(A-\lambda I)v_3 &= v_2 \quad &&（v_3 \text{ は変換されると } v_2 \text{ に成分を渡す}）
\end{aligned}
$$

3. 空間のイメージ図：バケツリレー



情報の連鎖（チェーン構造）
右上の「1」は「隣へ渡す」作用を表す

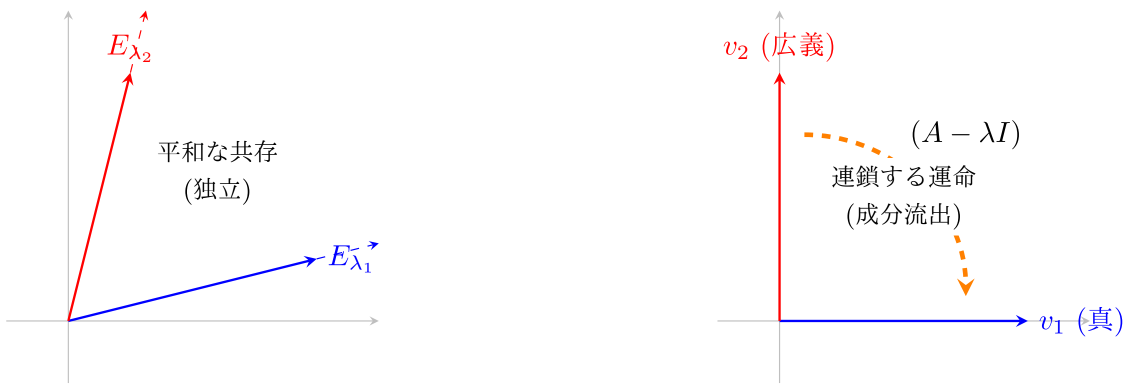
図 7.3: ジョルダン鎖のイメージ：作用するたびに情報が隣へシフトしていく．

- v_3: 3 番手のランナー．変換を受けると，自分の成分を v_2 に渡します．
- v_2: 2 番手のランナー．受け取った成分を v_1 に渡します．
- v_1: アンカー．受け取った成分を吸収して消滅（ゼロ化）させます．

このように，ジョルダン標準形の空間は，**「情報の連鎖（チェーン）」**構造を持っています．行列の「右上の 1」は，この**「隣へ渡す（Shift）」**という作用を表しているのです．

4. 図解：対角化可能 vs 対角化不可能

ジョルダン細胞の意義を，対角化可能な場合との比較で視覚化します．



(a) 対角化可能 (Diagonalizable)
$V = E_{\lambda_1} \oplus E_{\lambda_2}$

(b) 対角化不可能 (Non-Diagonalizable)
$V = W_\lambda$ (広義固有空間)

図 7.4: 空間構造の違い．左：対角化可能なら各軸は独立して伸び縮みする．右：対角化不可能だと，広義固有ベクトル (v_2) から真の固有ベクトル (v_1) へ成分が「漏れ出す」ため，軸を分離できない．

7.1　定理と構造

定理 7.1 (ジョルダン標準形)**.** 任意の正方行列 A は，適当な基底変換により，以下の形のブロック対角行列に変換できる．

$$J = \begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{pmatrix}, \quad J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

この形は，ブロックの順序を除いて一意である．

コメント 7.1 (連鎖反応とバケツリレー)．ジョルダンブロック $J_m(\lambda)$ の意味を考えます．基底ベクトルを $v_1, \ldots, v_m$ とすると，この行列は以下の関係を表しています．

$$\begin{aligned} (A-\lambda I)v_1 &= \mathbf{0} \quad \text{(真の固有ベクトル)} \\ (A-\lambda I)v_2 &= v_1 \\ (A-\lambda I)v_3 &= v_2 \\ &\vdots \\ (A-\lambda I)v_m &= v_{m-1} \end{aligned}$$

- v_m に加えられた入力は，変換のたびに $v_{m-1} \to \cdots \to v_1$ とバケツリレーのように伝搬し，最後に v_1 で吸収されて消えます．
- これは，社会ネットワークにおける「インフルエンサー（v_1）とフォロワーの連鎖」や，化学反応における「多段階反応プロセス」のモデルそのものです．

コラム：ケーリー・ハミルトンと広義固有空間 〜全体と部分〜

ケーリー・ハミルトンの定理 $\Phi_A(A) = O$ は，広義固有空間分解を使うと「当たり前」に見えてきます．

1. 行列式の分解

固有多項式は，固有値ごとの成分に因数分解できます．

$$\Phi_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\ldots(t-\lambda_k)^{n_k}$$

2. 空間の分解（広義固有空間）

全空間 V は，広義固有空間の直和に分解されます．

$$V = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_k}$$

ここで，各 W_{λ_i} 上のベクトル $\boldsymbol{x}$ に対して，$(A - \lambda_i I)^{n_i}\boldsymbol{x} = \mathbf{0}$ が成り立ちます（広義固有空間の定義）．

3. 「全滅」のメカニズム

行列 $\Phi_A(A)$ を任意のベクトル $\boldsymbol{v}$ に掛けてみましょう．$\boldsymbol{v}$ は各空間の成分に分解できます（$\boldsymbol{v} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k$）．

- $\boldsymbol{w}_1$ に対して: $\Phi_A(A)$ の中には因子 $(A - \lambda_1 I)^{n_1}$ が含まれています．これが $\boldsymbol{w}_1$ にヒットすると，即座に $\mathbf{0}$ になります．他の因子が何であろうと，結果は $\mathbf{0}$ です．
- $\boldsymbol{w}_2$ に対して: 今度は因子 $(A - \lambda_2 I)^{n_2}$ が $\boldsymbol{w}_2$ を消滅させます．

結論

$\Phi_A(A)$ は，**「どの空間の住人が来ても，必ずどれか一つの因子（トラップ）が作動して消してしまう」**ような，完璧な包囲網になっているのです．だから，空間全体のどのようなベクトルに対しても $\Phi_A(A)\boldsymbol{v} = \mathbf{0}$，つまり $\Phi_A(A) = O$ となるのです．

この解説により，「ケーリー・ハミルトン＝計算テクニック」という認識から，「ケーリー・ハミルトン＝空間全体の構造的性質」という深い理解へと導くことができます．

【直感的イメージ】最小多項式とジョルダン細胞の「背比べ」

第 7 章の最後に，**「最小多項式 (Minimal Polynomial)」**と広義固有空間（ジョルダン細胞）の関係を，視覚的なメタファーで整理します．

1. 図解イメージ：「ビルの高さ制限」

広義固有空間 W_λ を，いくつかのビル（ジョルダン細胞）が建っている**「敷地」**だと想像してください．

- **固有多項式の指数** n_i **（代数的重複度）**: 敷地の**「総面積（次元の合計）」**です．すべてのビルの床面積を足した値です．例えば，3 階建てと 2 階建てがあれば，総面積は $3 + 2 = 5$ です．
- **最小多項式の指数** m_i: 敷地内で**「最も高いビルの階数」**です．他の低いビル（小さい細胞）がいくつあろうと関係ありません．一番高いビルの高さだけで決まります．

2. 数学的意味：全滅させるための最小手数

行列 $(A - \lambda I)$ を掛けるという操作は，ビルの階数を 1 つ減らす（上から潰す）攻撃です．

- 3 階建てのビル（3×3 ジョルダン細胞）を消滅させるには，攻撃を 3 回繰り返す必要があります（$(A - \lambda I)^3$）．
- 隣に平屋（1×1 細胞）があっても，それは 1 回目の攻撃で既に消えています．
- したがって，**「一番しぶとい奴（最大細胞）」**を倒すのに必要な回数が，そのまま最小多項式の次数 m_i になるのです．

まとめ表

指標	記号	意味（ビルの例）	数学的対応
代数的重複度	n_i	総床面積	固有多項式の次数 $(t-\lambda)^{n_i}$
幾何学的重複度	$\dim E_\lambda$	ビルの棟数	固有空間の次元（独立な固有ベクトルの数）
最小多項式の次数	m_i	最高階数	最大のジョルダン細胞のサイズ $(t-\lambda)^{m_i}$

表 7.1: 固有値にまつわる 3 つの指標の直感的意味

この表とメタファーを使えば，「なぜ最小多項式を考えるのか？」という問いに対し，**「最大派閥（ボトルネック）のサイズを知りたいから」**と即答できるようになります．

7.2　例題：微分方程式とジョルダン標準形

数学的な証明の集大成として，実用的な微分方程式を解きます．

例題 7.1. 次の微分方程式を解け．

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{初期値} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

【解答】

1. **固有値の計算**: $A = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}$．$\det(A-\lambda I) = (\lambda-3)(\lambda-5)-(-1)(1) = \lambda^2 - 8\lambda + 16 = (\lambda-4)^2$．固有値は $\lambda = 4$（重解）．
2. **固有ベクトルの計算**: $(A-4I) = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$．$(A-4I)v_1 = \mathbf{0} \Rightarrow -x+y=0$．$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ が唯一の固有ベクトル．対角化不可能．
3. **広義固有ベクトルの計算**: $(A-4I)v_2 = v_1$ を解く．$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow -x+y=1$．例えば $x=0, y=1$ として $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を選べる．
4. **ジョルダン標準形による解**: 変換行列 $P = (v_1 v_2) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ とすると，$P^{-1}AP = J = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$．微分方程式の解は $u = P^{-1}x$ と置くと $\frac{du}{dt} = Ju$．J の指数関数行列は $e^{Jt} = e^{4t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$．元の座標に戻すと，

$$\boldsymbol{x}(t) = Pe^{Jt}P^{-1}\boldsymbol{x}(0)$$

計算すると，

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{4t}\begin{pmatrix} 1-t \\ 1-t \end{pmatrix}$$

このように，解に te^{4t} という項が現れます．これが「臨界状態」特有の挙動です．もし対角化可能なら $e^{\lambda t}$ の和にしかなりません．

第 IV 部

データの「距離」と「情報量」の幾何学

第 8 章

内積空間と直交性 (Inner Product Spaces)

8.1 内積の定義とコーシー・シュワルツの不等式

ベクトル空間に「内積」という新たなルールを導入することで，空間に「物差し（ノルム）」と「分度器（角度）」を与えます．

定義 8.1 (内積 (Inner Product))**.** 実ベクトル空間 V において，2 つのベクトル u, v から実数 (u, v) を対応させる写像が，以下の公理を満たすとき，これを内積と呼ぶ．

1. **対称性**: $(u, v) = (v, u)$
2. **線形性**: $(au + bw, v) = a(u, v) + b(w, v)$
3. **正定値性**: $(u, u) \geq 0$ であり，$(u, u) = 0 \iff u = \mathbf{0}$

この内積を用いて，ベクトルの**長さ（ノルム）**を $\|u\| = \sqrt{(u, u)}$，**角度** θ を $\cos\theta = \frac{(u,v)}{\|u\|\|v\|}$ と定義できる．

定理 8.1 (コーシー・シュワルツの不等式)**.** 任意の $u, v \in V$ に対して以下が成り立つ．

$$|(u, v)| \leq \|u\|\|v\|$$

等号成立は u, v が線形従属（平行）のときに限る．

Proof.

1. $v = \mathbf{0}$ のときは両辺 0 で成立．$v \neq \mathbf{0}$ とする．
2. 任意の実数 t に対して，正定値性より $\|u + tv\|^2 \geq 0$ である．
3. 展開すると，
$$(u + tv, u + tv) = \|u\|^2 + 2t(u, v) + t^2\|v\|^2 \geq 0$$
4. これは t に関する 2 次不等式である．常に 0 以上であるためには，この 2 次方程式の判別式 D が 0 以下でなければならない．
$$D/4 = (u, v)^2 - \|v\|^2\|u\|^2 \leq 0$$
5. 移項してルートをとれば $|(u, v)| \leq \|u\|\|v\|$ が得られる．

□

例題 8.1 (コサイン類似度)**.** 自然言語処理（NLP）において，単語や文書をベクトル化（Embedding）した際，その類似度を測るために内積が使われます．

- $\cos\theta \approx 1$（平行）: 意味が近い（例 :「猫」と「子猫」）
- $\cos\theta \approx 0$（直交）: 無関係（例 :「猫」と「経済」）

コーシー・シュワルツの不等式は，**「相関係数は必ず -1 から 1 の間に収まる」**という統計学の基本原理の数学的保証です．

8.2 直交射影と最小二乗法

データにノイズが含まれているとき，それを「最も近い」きれいなデータで近似したい．これが射影の考え方です．

定理 8.2 (射影定理 (Projection Theorem))**.** V を有限次元内積空間，W をその部分空間とする．任意のベクトル $v \in V$ は，

$$v = w + z \quad (w \in W, z \in W^\perp)$$

の形に一意に分解できる．ここで $W^\perp$（直交補空間）は，W 内の全てのベクトルと直交するベクトルの集合である．この w を，v の W への**正射影**と呼ぶ．

Proof.

1. **グラム・シュミットの直交化**: W の基底 $\{u_1, \ldots, u_k\}$ から，正規直交基底 $\{e_1, \ldots, e_k\}$ を構成できる（証明略）．
2. **射影の構成**: $w = \sum_{i=1}^k (v, e_i)e_i$ と置く．明らかに $w \in W$．
3. **直交性の確認**: 残差ベクトル $z = v - w$ を考える．任意の基底 e_j との内積をとると，

$$(z, e_j) = (v, e_j) - \sum_{i=1}^k (v, e_i)(e_i, e_j)$$

正規直交性より $(e_i, e_j) = \delta_{ij}$ なので，右辺の和は $i = j$ の項だけ残り，

$$(z, e_j) = (v, e_j) - (v, e_j) = 0$$

よって z は全ての基底と直交するため，$z \in W^\perp$．

4. **一意性**: $v = w + z = w' + z'$ と 2 通りに書けたとすると，$w - w' = z' - z$．左辺は W に属し，右辺は $W^\perp$ に属する．$W \cap W^\perp = \{\mathbf{0}\}$（自分自身と直交するのはゼロのみ）なので，$w - w' = \mathbf{0}$．よって一意．

□

直感的理解：正射影＝「真上からの影」

直交射影（正射影）のイメージは，**「真上からの光による影」**です．

1. 図解：光と影の関係

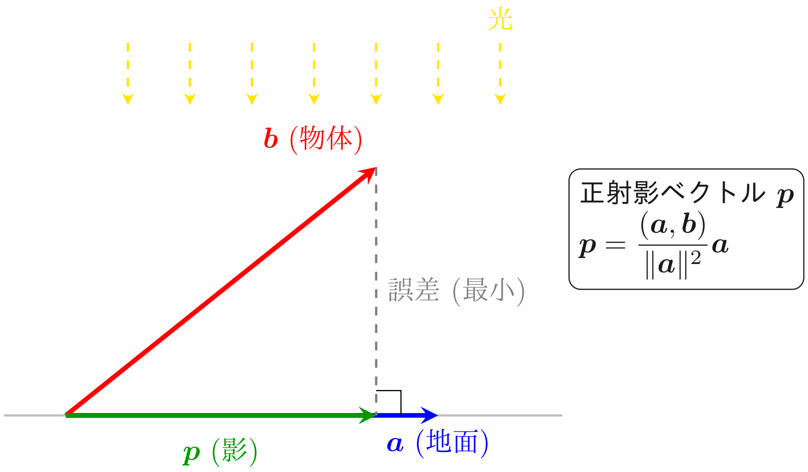


図 8.1: 正射影の幾何学的意味．$\boldsymbol{b}$ に対して「真上（$\boldsymbol{a}$ と垂直な方向）」から光を当てたときにできる「影」が $\boldsymbol{p}$ である．

2. 図の見方

- **地面**（青い矢印 $\boldsymbol{a}$）：スクリーンや部分空間を表します．
- **物体**（赤い矢印 $\boldsymbol{b}$）：影を落とす元のベクトルです．
- **光**（黄色い矢印）：地面に対して**垂直に**降り注ぐ光です．これが「直交（Orthogonal）」の意味です．
- **影**（緑の矢印 $\boldsymbol{p}$）：これが**直交射影ベクトル**です．$\boldsymbol{b}$ の成分のうち，$\boldsymbol{a}$ 方向の成分だけを取り出したものです．

3. 最短距離の定理

図の「灰色の点線」は物体から影までの距離（誤差）を表しています．直交射影の最大の特徴は，**「この点線が一番短くなる」**ことです．もし影の位置（緑の矢印の先端）が左右にズレたら，点線（誤差）は斜辺になって長くなってしまいます．

$$\|\boldsymbol{b} - \boldsymbol{p}\| = \text{最小}$$

これが，最小二乗法で直交射影が使われる理由（誤差を最小にする近似）です．

4. 式の直感的意味

$$\boldsymbol{p} = \frac{(\boldsymbol{a}, \boldsymbol{b})}{\|\boldsymbol{a}\|^2}\boldsymbol{a}$$

この式は，**「内積（重なり具合）を，土台の長さで正規化したもの」**と読めます．

- $(\boldsymbol{a}, \boldsymbol{b})$: 影の長さ × 土台の長さ
- $\|\boldsymbol{a}\|$: 土台の長さ

割り算して、「影の長さ」だけを取り出し，それを土台 $\boldsymbol{a}$ の向きに乗せているのです．

8.3　グラム・シュミットの直交化：影を引き算して「純化」する

グラム・シュミットの直交化法の本質は，**「自分の影を引き算して，純粋な自分になる」**というプロセスです．

コンセプト：影落とし (Shadow Casting)

与えられたベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots$ は，最初は斜めに向き合っていて，直交していません．ここから，互いに直交するきれいな基底 $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$ を作り出します．

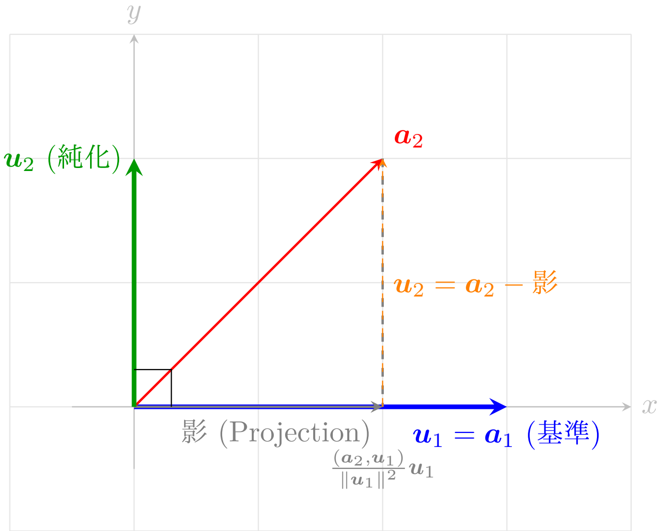


図 8.2: グラム・シュミットの直交化の幾何学的意味．$\boldsymbol{a}_2$ から「$\boldsymbol{u}_1$ 方向の成分（影）」を引き算することで，$\boldsymbol{u}_1$ と直交する新しいベクトル $\boldsymbol{u}_2$ を生み出す．

ステップごとの操作

1. **Step 1: 最初の基準を決める** $\boldsymbol{a}_1$ をそのまま最初の基底 $\boldsymbol{u}_1$ にします（青い矢印）．これがすべての基準（土台）になります．
2. **Step 2: 影（Projection）を見つける**次のベクトル $\boldsymbol{a}_2$（赤い矢印）を見ます．このベクトルは，少しだけ $\boldsymbol{u}_1$ の方向成分を持っています．図の「灰色の矢印（Projection)」が，$\boldsymbol{a}_2$ が $\boldsymbol{u}_1$ に落とした影（正射影）です．この影の部分は，$\boldsymbol{u}_1$ と情報が被っている「邪魔な成分」です．
3. **Step 3: 影を引き算する（純化）** $\boldsymbol{a}_2$ から，この影を引き算します．

$$\boldsymbol{u}_2 = \boldsymbol{a}_2 - (\text{影}) = \boldsymbol{a}_2 - \frac{(\boldsymbol{a}_2, \boldsymbol{u}_1)}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1$$

すると，$\boldsymbol{u}_1$ と被っている成分がきれいになくなり，**垂直に立ち上がるベクトル $\boldsymbol{u}_2$（緑の矢印）** だけが残ります．

これを繰り返すことで，次々と新しいベクトルを持ってきては，**「先輩たち（既存の基底）の上に落ちる影をすべて引き算する」**ことで，純粋に新しい方向（直交成分）だけを取り出していきます．

式のイメージ化

$$\boldsymbol{u}_{\text{new}} = \boldsymbol{a}_{\text{target}} - \sum(\text{先輩たちへの射影})$$

「自分の中から，先輩たちと被ってるキャラ（成分）を全部捨てれば，独自性（直交性）が出る」と覚えれば忘れません．

例題 8.2 (ノイズ除去と回帰分析)．

- **信号処理**: 観測データ v にはノイズが乗っている．信号が存在すべき部分空間 W（例：低周波成分）へ射影することで，ノイズ成分 z（高周波成分）を除去できる．

- **最小二乗法**: 連立方程式 $Ax = b$ が解を持たないとき，$\|Ax - b\|^2$ を最小にする x を求めることは，b を $\mathrm{Im}\ A$ という部分空間へ射影することと同義です.

第 9 章

実対称行列とスペクトル定理 (Real Symmetric Matrices)

ジョルダン標準形が必要になるような「複雑な行列」とは対照的に，実社会で最も頻出する「対称行列（$A^T = A$)」は，驚くほど美しい性質を持っています．

9.1　対称行列の固有値はすべて実数

定理 9.1. 実対称行列 A の固有値は，すべて実数である．

Proof.
1. λ を固有値，x を固有ベクトルとする（複素数の範囲で考える）．$Ax = \lambda x$.
2. x の随伴ベクトル（転置して複素共役をとったもの）を x^* とする．内積（エルミート形式） x^*Ax を考える．
3. 一方から計算：$x^*(Ax) = x^*(\lambda x) = \lambda(x^*x) = \lambda\|x\|^2$.
4. 他方から計算：$(x^*A)x = (Ax)^*x$ （A は実対称なので $A^* = A^T = A$)．$= (\lambda x)^*x = \bar{\lambda}x^*x = \bar{\lambda}\|x\|^2$.
5. よって $\lambda\|x\|^2 = \bar{\lambda}\|x\|^2$．$x \neq \mathbf{0}$ より $\|x\|^2 \neq 0$ なので，$\lambda = \bar{\lambda}$．これは λ が実数であることを意味する．

□

【直感的イメージ】実対称行列＝「現実世界の演算子」

実対称行列の固有値がすべて実数になる理由は，**「対称行列は〈実ベクトル同士の内積〉に対して，正しく振る舞う“エネルギー演算子”になっているから」**と見るとイメージしやすくなります．

1. 内積を保つ「実世界の演算子」

実対称行列 A は，任意の実ベクトル x, y に対して次を満たします．

$$(Ax, y) = (x, Ay)$$

これは，「**A を片側に動かしても，内積の値は変わらない**」という意味で，物理でいう**自己共役演算子（エルミート演算子）**と同じ性質です．エネルギーや長さのような「現実の量」を表す演算子は，この性質を満たす必要があります．

2. 固有値は「エネルギーの測定値」

固有値・固有ベクトルの式 $Ax = \lambda x$ を，内積を使って覗き込みます．複素共役を含めた内積 $(u, v) = u^* v$ を考えると，

$$(Ax, x) = (\lambda x, x) = \bar{\lambda}(x, x) \quad \text{(第 1 成分から出したので共役)}$$

一方，対称性（自己共役性）から

$$(Ax, x) = (x, Ax) = (x, \lambda x) = \lambda(x, x)$$

も成り立ちます．左辺は同じ量なので，右辺同士が等しくなければいけません．

$$\bar{\lambda}(x, x) = \lambda(x, x)$$

ここで $(x, x) = \|x\|^2 > 0$（ゼロベクトルでない限り）なので割り算でき，

$$\lambda = \bar{\lambda}$$

となります．**自分と共役が等しい数は実数しかありません．**よって，λ は必ず実数（虚数成分を持たない）になります．

3. イメージまとめ

- 対称行列は「内積（エネルギー）」と相性がよく，左右を入れ替えても矛盾が生じません．
- 内積を通して左右から眺めても同じ値になるため，その測定値は**虚数成分を持てず，必ず実数**になります．

一般の行列
(固有空間は歪んでいる)

対称行列
(固有空間は直交する)

図 9.1: 対称行列の幾何学的性質．一般の行列の固有ベクトルは斜交することもあるが，対称行列の異なる固有値に対応する固有ベクトルは必ず直交する．これは「空間を歪ませずに伸縮させる」ことを意味する．

このイメージを押さえておくと，「実対称行列＝現実世界の量を表す“安全な”行列」という感覚で，その後の直交対角化やスペクトル分解にも自然につながります．

このイメージを押さえておくと，「実対称行列＝現実世界の量を表す“安全な”行列」という感覚で，その後の直交対角化やスペクトル分解にも自然につながります．

【直感的イメージ】直交対角化 ～歪んだ世界から直角の世界へ～

実対称行列の美しい性質 $P^T AP = D$（直交対角化）は，幾何学的には**「歪んだ格子を，回転させるだけできれいな長方形格子に直す」**操作に対応します．

(a) 元の世界

行列 $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$

$\xrightarrow{P^T}$

(b) 回転後の世界

行列 $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

図 9.2: 直交対角化のイメージ．左：元の基底では，空間の軸が歪んでいる（非対角成分 b の存在）．右：固有ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2$ を新しい軸に選ぶと，軸は直交し，行列の作用は各軸ごとの単純な伸縮（固有値 λ_1, λ_2）になる．

変換の意味と「実対称行列だけの特権」

1. **「歪み」の解消**: 非対角行列 $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ の成分 b は，座標軸間の「相関（歪み）」を表しています．対角化とは，この b を消去して，互いに無関係な（直交した）軸を取り直すことです．
2. **なぜ実対称行列だけがこれを許されるのか？**
   - **固有ベクトルが直交する ($P^TP = I$)**: これにより，座標変換が単なる「回転」で済みます．形を変えずに視点を変えるだけです．
   - **固有値が実数**: 変換後の世界の伸縮率（スケーリング）が現実的な値になります．

   この 2 つの性質が揃うのは，実対称行列特有の「魔法」です．そのため，データ解析（PCA など）で「まず相関行列（対称行列）を作る」という手順が定石となっているのです．

9.2　スペクトル定理（対称行列の対角化）

定理 9.2 (スペクトル定理)**.** 実対称行列 A は，直交行列 P（$P^TP = I$）を用いて対角化可能である．

$$P^{-1}AP = P^TAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Proof. 証明の概略は以下の通りです．

1. 固有値 λ_1（実数）と，対応する単位固有ベクトル u_1 を取る．
2. u_1 を含む正規直交基底を作り，行列 P_1 を構成する．
3. $P_1^TAP_1$ を計算すると，1 列目は $(\lambda_1, 0, \ldots, 0)^T$ となる．
4. さらに対称性より，1 行目も $(\lambda_1, 0, \ldots, 0)$ となる．

$$P_1^TAP_1 = \begin{pmatrix} \lambda_1 & O \\ O & A' \end{pmatrix}$$

5. 残りのブロック A' も対称行列なので，同様の操作を繰り返す（数学的帰納法）ことで，最終的に対角行列になる．

□

例題 9.1 (PCA とエネルギー地形).

- **主成分分析 (PCA)**: 分散共分散行列（対称行列）を対角化することは，データの分布の形を解析し，最も情報量が多い軸（主成分）を見つける操作そのものです．直交行列 P で回転しても，データの相対的な位置関係は壊れません．
- **最適化問題**: ニューラルネットワークの損失関数の「谷底」付近の形状は，ヘッセ行列（対称行列）で近似できます．固有値がすべて正なら極小値（安定），負が混ざっていれば鞍点（不安定）です．スペクトル定理は，この「多次元の曲面」を単純な放物面の足し合わせに分解できることを保証しています．

第Ｖ部

現代データ解析の核心

第 10 章

特異値分解 (Singular Value Decomposition)

対称行列のスペクトル分解（対角化）を，一般の $m \times n$ 行列に拡張します．これは，データを「回転・伸縮・回転」という単純な幾何学操作に分解することを意味します．

10.1 準備：$A^T A$ の性質

任意の行列 A に対し，$A^T A$ という行列を作ると，非常に扱いやすい性質が現れます．

補題 10.1. A を実 $m \times n$ 行列とする．対称行列 $S = A^T A$ （$n \times n$）について，以下が成り立つ．

1. S は半正定値である（固有値 $\lambda \geq 0$）．
2. $Ker(S) = Ker(A)$．したがって $rank(S) = rank(A)$．

Proof. 1. 任意の $x \in \mathbb{R}^n$ に対し，$x^T S x = x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$．対称行列のレイリー商が常に非負であるため，固有値も非負である．

2. $x \in \mathrm{Ker}(A) \implies Ax = \mathbf{0} \implies A^T A x = \mathbf{0} \implies x \in \mathrm{Ker}(S)$．逆に，$x \in \mathrm{Ker}(S) \implies A^T A x = \mathbf{0} \implies x^T A^T A x = 0 \implies \|Ax\|^2 = 0 \implies Ax = \mathbf{0}$．よって集合として等しい．

□

この非負の固有値の平方根 $\sigma_i = \sqrt{\lambda_i(A^T A)}$ を，行列 A の**特異値 (Singular Values)** と呼びます．

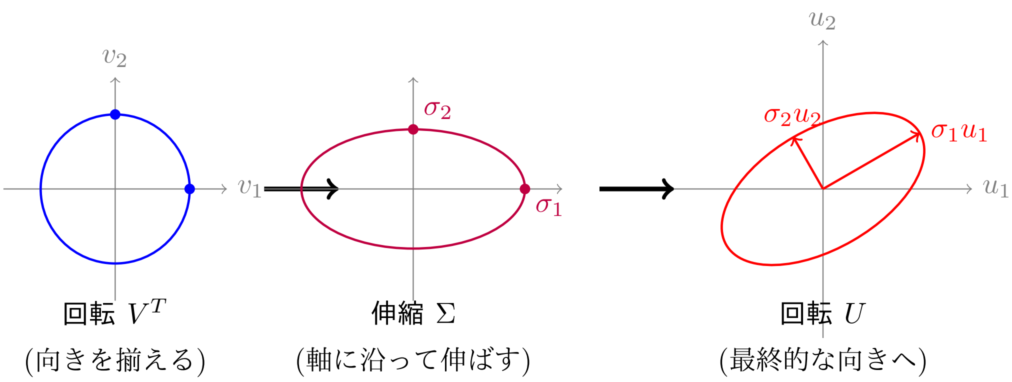


図 10.1: SVD の幾何学的意味：$A = U\Sigma V^T$．任意の行列変換は，1. 回転 (V^T)，2. 軸方向への伸縮 (Σ)，3. 回転 (U) の 3 ステップに分解できる．

10.2　特異値分解定理 (SVD Theorem)

定理 10.1 (特異値分解)**.** 任意の $m \times n$ 実行列 A は，以下のように分解できる．

$$A = U\Sigma V^T$$

ここで，

- U: m 次直交行列（左特異ベクトル）
- V: n 次直交行列（右特異ベクトル）
- Σ: $m \times n$ 対角行列（対角成分に特異値 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ が並び，残りは 0）

Proof.

1. **右側の直交行列 V の構成**: A^TA は n 次実対称行列なので，スペクトル定理（第 9 章）により直交行列 $V = (v_1, \dots, v_n)$ で対角化できる．$V^T(A^TA)V = \mathrm{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0)$．ここで $\sigma_1 \geq \cdots \geq \sigma_r > 0$ と並べておく．
2. **左側の正規直交系 $\{u_i\}$ の構成**: $i = 1, \dots, r$ に対し，ベクトル u_i を以下で定義する．

$$u_i = \frac{1}{\sigma_i} A v_i$$

これらが正規直交であることを確認する．

$$(u_i, u_j) = \frac{1}{\sigma_i \sigma_j}(Av_i)^T(Av_j) = \frac{1}{\sigma_i \sigma_j} v_i^T(A^TAv_j)$$

v_j は A^TA の固有ベクトル（固有値 σ_j^2）なので，

$$= \frac{1}{\sigma_i \sigma_j} v_i^T(\sigma_j^2 v_j) = \frac{\sigma_j}{\sigma_i} v_i^T v_j$$

V は直交行列なので $v_i^T v_j = \delta_{ij}$．よって $(u_i, u_j) = \delta_{ij}$．

3. **左側の直交行列 U の完成**: $\{u_1, \dots, u_r\}$ に，適当な正規直交基底 $\{u_{r+1}, \dots, u_m\}$ を付け加えて（基底の延長定理），m 次の直交行列 U を作る．
4. **分解の確認**: 行列 U^TAV の (i, j) 成分を計算する．これは $u_i^T A v_j$ に等しい．
   - **Case** $j \leq r$: $Av_j = \sigma_j u_j$ なので，$u_i^T(\sigma_j u_j) = \sigma_j \delta_{ij}$．つまり，対角成分は σ_j，それ以外は 0．
   - **Case** $j > r$: v_j は A^TA のゼロ固有値に対応する（核に含まれる）．$v_j \in \mathrm{Ker}(A)$ なので $Av_j = \mathbf{0}$．よって成分はすべて 0．

以上より，$U^TAV = \Sigma$ となり，$A = U\Sigma V^T$ が示された．

□

例題 10.1 (データの「真の姿」)**.** A を顧客ごとの購買データ行列とします．

- V（右特異ベクトル）: 商品の「カテゴリ」軸（野菜系、肉系...）．
- U（左特異ベクトル）: 顧客の「嗜好」軸（健康志向、節約志向...）．
- Σ（特異値）: そのパターンの「重要度」．

SVD は，膨大な購買データから「誰が (U)」「どんなカテゴリを (V)」「どれくらいの強さで (Σ)」買っているかという**潜在構造**をあぶり出します．

例題 10.2 (手計算で SVD)**.** 行列 $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ を特異値分解せよ．

1. A^TA の固有値を求める: $A^TA = \begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$. 固有値は $\lambda = 2, 0$. 特異値は $\sigma_1 = \sqrt{2}, \sigma_2 = 0$. $\Sigma = \begin{pmatrix}\sqrt{2} & 0\\ 0 & 0\end{pmatrix}$.
2. 右特異ベクトル V を求める: A^TA の固有ベクトルを正規化すると，$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}, v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix}$. よって $V = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$.
3. 左特異ベクトル U を求める: $u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} = \begin{pmatrix}1\\ 0\end{pmatrix}$. u_2 は u_1 に直交する単位ベクトルなので $\begin{pmatrix}0\\ 1\end{pmatrix}$. よって $U = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$.
4. 結論: $A = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{2} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{pmatrix}^T$.

第 11 章

低ランク近似とエッカート・ヤングの定理

SVD の真価は，計算できることではなく，**「情報を捨てられる」**ことにあります．複雑な行列を「重要な成分（信号）」と「どうでもいい成分（ノイズ）」に分解し，ノイズを切り捨てることで，データの**本質的な構造**を抽出します．

11.1 【直感的イメージ】SVD によるノイズ除去

行列 A を「信号」と「ノイズ」の和として捉えます．特異値 σ_i の大きさは，その成分の重要度（エネルギー）を表します．

- **大きな特異値** $(\sigma_1, \dots, \sigma_k)$: 本質的な情報（信号）.
- **小さな特異値** $(\sigma_{k+1}, \dots, \sigma_r)$: 誤差やノイズ.

$$\underbrace{A}_{\text{元データ}} = \underbrace{\sum_{i=1}^{k} \sigma_i u_i v_i^T}_{\text{信号（情報の要約）}} + \underbrace{\sum_{i=k+1}^{r} \sigma_i u_i v_i^T}_{\text{ノイズ（捨てる部分）}}$$

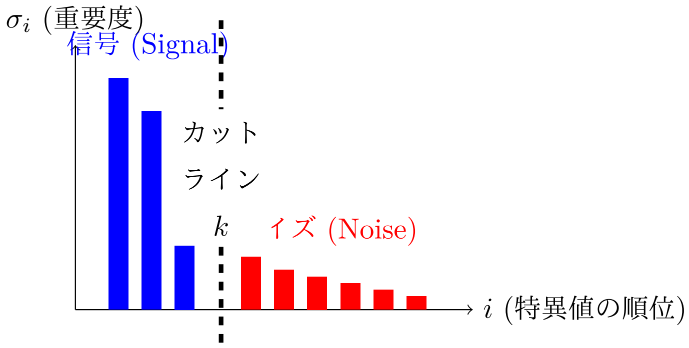


図 11.1: 特異値の分布とカットライン．大きな特異値だけを残すことで，元の行列の「骨格」だけを取り出すことができる．

11.2 エッカート・ヤングの定理 (Eckart-Young Theorem)

「特異値を大きい順に k 個残す」という単純な操作が，数学的に**「最良の近似」**であることを保証する強力な定理です．

定理 11.1 (エッカート・ヤングの定理). 行列 A の特異値分解を $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ とする．このとき，ラン

ク k $(<r)$ の行列の中で，フロベニウスノルム $\|\cdot\|_F$ の意味で A に最も近い行列 A_k は，

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T \quad (SVD\text{ の上位 } k \text{ 項})$$

で与えられる．また，そのときの最小誤差は，切り捨てられた特異値の二乗和の平方根である．

$$\|A - A_k\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2}$$

なぜ「最良」なのか？（証明の鍵）

この定理の美しさは，行列の空間における**「直交性（三平方の定理）」**にあります．

フロベニウスノルムの二乗 $\|A\|_F^2$ は，全特異値の二乗和 $\sum \sigma_i^2$ と等しくなります（SVD の性質）．したがって，誤差行列 $E = A - A_k$ のノルムは，単純に捨てた特異値の二乗和になります．

$$\|A - A_k\|_F^2 = \| \sum_{i=k+1}^{r} \sigma_i u_i v_i^T \|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2$$

特異値は大きい順 $(\sigma_1 \geq \sigma_2 \geq \dots)$ に並んでいるため，誤差（捨てた部分の和）を最小にするには，**「小さい方から順に捨てる」**のが唯一の正解となります．これが数学的な保証の正体です．

11.3　応用分野：低ランク近似の威力

低ランク近似は，現代のデータサイエンスの至る所で使われています．

分野	**低ランク近似の役割**	**定理の恩恵**
画像圧縮	画像のノイズ除去・容量削減	見た目を保ったままデータ激減
推薦システム	ユーザー・商品行列の潜在構造抽出	スパースなデータから好みを予測
信号処理	観測信号からのノイズ分離	SN 比（信号対雑音比）の最大化
主成分分析	高次元データの次元削減	情報損失を最小化して可視化
自然言語処理	単語の意味空間の抽出 (LSA)	類義語や文脈の数学的表現

表 11.1: 低ランク近似の主な応用分野

11.4　数値例：2 × 2 行列での確認

簡単な例で，定理が成り立っていることを確認しましょう．

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

この行列はすでに対角化されており，特異値は $\sigma_1 = 4, \sigma_2 = 1$ です．

- **ランク 1 近似** $(k = 1)$: 小さい特異値 $\sigma_2 = 1$ を捨てて，$\sigma_1 = 4$ だけを残します．

$$A_1 = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$$

- **誤差の計算**:
$$A - A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\|A - A_1\|_F = \sqrt{0^2 + 0^2 + 0^2 + 1^2} = 1$$
- **定理の確認**: 定理によれば，誤差は $\sqrt{\sum_{i=2}^{2} \sigma_i^2} = \sqrt{1^2} = 1$．確かに一致しています．

もし他のランク 1 行列，例えば $B = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}$ （これはランク 1）で近似しようとすると，

$$\|A - B\|_F = \left\| \begin{pmatrix} 4 & 0 \\ 0 & -4 \end{pmatrix} \right\|_F = \sqrt{16 + 16} = \sqrt{32} \approx 5.65$$

となり，SVD による近似 A_1 （誤差 1）の方が圧倒的に優れていることがわかります．

11.5　数値例：3 × 3 行列での手計算と Python 検証

もう少し複雑な行列で，実際に近似がどう行われるかを見てみましょう．

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & 13 \\ 4 & 5 & 5 \end{pmatrix}$$

これを SVD 分解すると，特異値は以下のようになります（小数点第 2 位まで）．

$$\sigma_1 \approx 25.73, \quad \sigma_2 \approx 3.96, \quad \sigma_3 \approx 1.77$$

ランク 1 近似 ($k = 1$)

最大の特異値 σ_1 とその特異ベクトルだけを使って復元します．

$$A_1 \approx \begin{pmatrix} 6.44 & 9.64 & 13.82 \\ 5.91 & 8.85 & 12.68 \\ 2.83 & 4.24 & 6.08 \end{pmatrix}$$

誤差の検証

- **実際の誤差**: 元行列との差分 $A - A_1$ のフロベニウスノルムを計算すると，
$$\|A - A_1\|_F \approx 4.34$$
となります．
- **理論値（エッカート・ヤングの定理）**: 切り捨てた特異値 (σ_2, σ_3) の二乗和の平方根と一致するはずです．
$$\sqrt{\sigma_2^2 + \sigma_3^2} = \sqrt{3.96^2 + 1.77^2} \approx \sqrt{15.68 + 3.13} \approx 4.34$$

理論通り，**「捨てた部分のエネルギー」**が**そのまま誤差**になっていることが確認できます．

再現率（情報の保持率）

元の行列が持っていた情報量のうち，どれくらいを維持できたかは，特異値の二乗和（エネルギー）の比で計算できます．

$$\text{再現率} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} \approx \frac{662.0}{662.0 + 15.7 + 3.1} \approx 92.5\%$$

たった 1 つの成分（ランク 1）だけで，全体の情報の **9 割以上**を表現できていることになります．これがデータ圧縮の威力です．

11.6　QR 分解による高速近似（SVD の代替）

SVD は「最良」の近似を与えますが，計算コストが高いという欠点があります．数万・数億規模の巨大な行列に対しては，より高速な **QR 分解（列ピボット付き）**を用いた低ランク近似が，エンジニアリングの現場ではよく使われます．

基本原理：列ピボット付き QR (CPQR)

行列の列を「ノルム（重要度）の大きい順」に並べ替えながら QR 分解を行います．

$$AP = QR$$

ここで R は上三角行列ですが，対角成分が概ね降順（$|r_{11}| \geq |r_{22}| \geq \dots$）に並びます．下の方の小さな成分（$R_{22}$ ブロック）をゼロとみなして切り捨てることで，SVD に近い近似を超高速に実現します．

SVD vs QR 分解：使い分けガイド

項目	**SVD** (特異値分解)	**QR** (列ピボット付き)	勝者
理論的最適性	最良（証明済み）	近似的に最良	SVD
計算速度	遅い $O(mn^2)$	非常に速い $O(mnk)$	**QR**
メモリ効率	全要素が必要	ストリーミング処理が可能	**QR**
ランク推定	正確	高速に推定可能	用途次第
構造保持	密行列になる	疎行列のゼロ配置を維持	**QR**

表 11.2: 精度重視なら SVD，速度・規模重視なら QR を選ぶのが定石です．

表 11.3: 精度重視なら SVD，速度・規模重視なら QR を選ぶのが定石です．

図 11.2: SVD と QR の構造的違い．SVD は成分を完全に「独立した軸」に分解するのに対し，QR（グラム・シュミット）は「前の成分に依存する」三角形状の分解を行う．低ランク近似においては，SVD が楕円の軸を正確に捉えるのに対し，QR はその近似形（平行四辺形など）を与える．

- **研究・分析**: 理論的な保証が必要な場合は **SVD.**
- **リアルタイム処理**: 速度が命（例：動画の背景除去）の場合は **QR.**
- **超巨大データ**: メモリに乗り切らない場合は **QR**（またはランダム化 SVD）.

コラム：SVD は「万能近似器」である

エッカート・ヤングの定理は，「行列を近似するなら SVD が最強である」と断言しています．ニューラルネットワークなどの複雑なモデルを使わなくても，単なる線形代数の演算（SVD）だけで，データの圧縮・ノイズ除去・特徴抽出が**理論的限界（最適解）**で行える．これが，特異値分解が「線形代数の頂点」と呼ばれる所以です．

第 12 章

一般化逆行列 (Moore-Penrose Pseudoinverse)

「正則でない行列」や「長方形の行列」に対して，逆行列のような働きをする行列を定義します．これは，逆行列が存在しない（ランク落ちした）連立方程式 $Ax = b$ を，「可能な限り解け」という命令に対する数学的な答えです．

12.1　【直感的イメージ】一般化逆行列＝「影を追いかける」

1. 通常の逆行列との違い

- **通常の逆行列:** $Ax = b \implies x = A^{-1}b$ （完璧解）
- **一般化逆行列:** $Ax = b \implies x = A^{+}b$ （最善解）

可逆行列なら「完璧な解」が求まりますが，非可逆行列では解が存在しなかったり無数にあったりします．一般化逆行列は，その中で**「誤差を最小にする解（最小二乗解）」**かつ**「ノルムが最小の解（最短解）」**を選び出します．

2. 幾何学的イメージ：影の再現

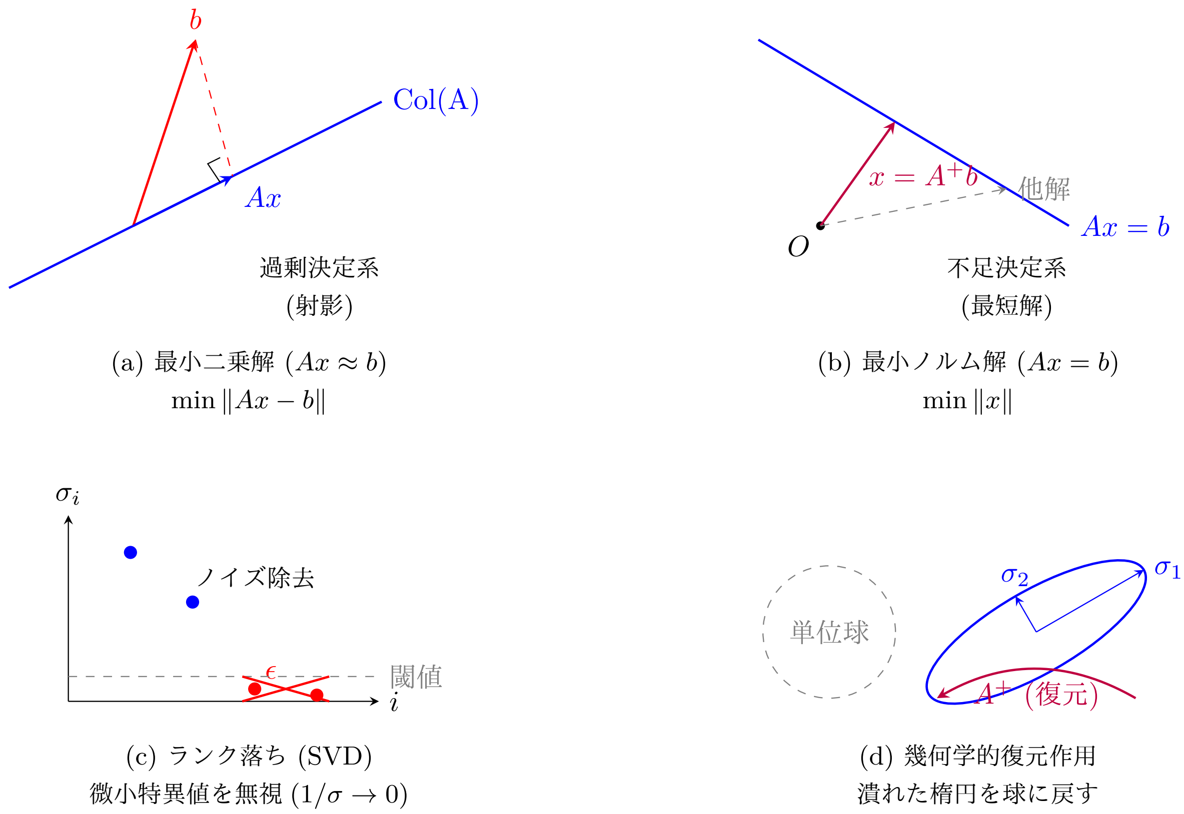
連立方程式 $Ax = b$ を，「影を落とす」操作と考えます．

- A: 影を落とす変換（射影行列など）
- b: 壁に映った影（観測データ）
- x: 元の物体（推定したいもの）

一般化逆行列 A^{+} は，**「観測された影 b から，元の物体 x を復元しようとする」**操作です．情報が失われていて完全には復元できない場合でも，「観測された影と最も矛盾しない（最も近い）物体」を再構成します．

3. 図解：一般化逆行列の 4 つの顔

一般化逆行列は，状況に応じて異なる「最適化」を自動的に行っています．



(a) 最小二乗解 ($Ax \approx b$)
$\min \|Ax - b\|$

(b) 最小ノルム解 ($Ax = b$)
$\min \|x\|$

(c) ランク落ち (SVD)
微小特異値を無視 ($1/\sigma \to 0$)

(d) 幾何学的復元作用
潰れた楕円を球に戻す

図 12.1: 一般化逆行列の統一的理解．左上：列空間への射影（過剰決定）．右上：原点からの最短距離（不足決定）．左下：SVD によるノイズカット．右下：変換の引き戻し作用．これらすべてが「一般化逆行列」という一つの枠組みで記述される．

12.2　定義と計算方法

定義 12.1 (ムーア・ペンローズ型一般化逆行列)**.** $A = U\Sigma V^T$ に対し，A の擬似逆行列 A^+ を以下で定義する．

$$A^+ = V\Sigma^+U^T$$

ここで Σ^+ は，Σ の非ゼロ成分 σ_i を逆数 $1/\sigma_i$ に置き換え，転置したものである．

12.3　一般化逆行列の計算アルゴリズム

ムーア・ペンローズ逆行列の手計算アルゴリズムを，**3 つのケース別に体系化**して提示します．SVD を使わない場合と使う場合の両方をカバーし，実用的な手順を整理します．

アルゴリズム分類表

ケース	行列形状・条件	計算式	手順
1	縦長フルランク ($m > n$)	$A^+ = (A^TA)^{-1}A^T$	通常の最小二乗
2	横長フルランク ($m < n$)	$A^+ = A^T(AA^T)^{-1}$	最小ノルム解
3	一般形（ランク欠損含む）	$A = U\Sigma V^T \to A^+ = V\Sigma^+U^T$	SVD 必須

統一手順（判定ツリー）

$$\boxed{\text{入力 } A(m \times n)}$$

$$\downarrow \quad \text{ランク判定：}\mathrm{rank}(A) = \min(m, n)?$$

$$\begin{cases} \textbf{YES (フルランク)} \to \begin{cases} m > n \text{ (縦長)} & \to \text{公式 } (A^TA)^{-1}A^T \text{ を使用} \\ m < n \text{ (横長)} & \to \text{公式 } A^T(AA^T)^{-1} \text{ を使用} \\ m = n \text{ (正方)} & \to \text{逆行列 } A^{-1} \text{ と一致} \end{cases} \\ \textbf{NO (ランク落ち)} \to \text{SVD 分解必須 } (A = U\Sigma V^T \to V\Sigma^+U^T) \end{cases}$$

確認アルゴリズム（4 性質チェック）

計算結果 A^+ が正しいかどうかは，以下の 4 式をチェックすれば確実です．

- $AA^+A = A$ （再生性 1）
- $A^+AA^+ = A^+$ （再生性 2）
- $(AA^+)^T = AA^+$ （Hermitian 性：列空間射影）
- $(A^+A)^T = A^+A$ （Hermitian 性：行空間射影）

応用：掃き出し法補足（参考）

小さな行列の場合，以下の掃き出し法が使えることもありますが，ランク落ちの判定がややこしいため，基本的には上記の SVD ルートを推奨します．

$$[A \mid I] \xrightarrow{\text{行基本変形}} [I \mid A^+] \quad \text{(正方可逆の場合のみ)}$$

12.4 演習：手計算で理解する 3 つのパターン

ここでは，手計算で一般化逆行列を導出するプロセスを 3 つのパターン（過剰決定，不足決定，ランク落ち）で実践します．これらを追うことで，理論がどのように具体的な数値計算に落ちていくかを体感してください．

例題 1：過剰決定系（縦長・フルランク）

問題: $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ に対し $x = A^+b$ を求めよ．

解法: A は列フルランクなので A^TA が正則です．$A^+ = (A^TA)^{-1}A^T$ を使います．

1. **A^TA の計算**:

$$A^TA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}$$

行列式は $\det = 3 \times 14 - 6 \times 6 = 42 - 36 = 6 \neq 0$.

2. **逆行列 $(A^TA)^{-1}$**:

$$(A^TA)^{-1} = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}$$

3. **$A^+ = (A^TA)^{-1}A^T$**:

$$A^+ = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{pmatrix}$$

（※注：途中計算 $14-6=8, 14-12=2, 14-18=-4\dots$）

4. **解 $x = A^+ b$**:

$$x = \frac{1}{6}\begin{pmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 8+4-12 \\ -3+0+9 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

確認: $Ax = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = b$．なんと，今回は誤差ゼロで解けてしまいました（$b$ が偶然 A の列空間内にあったため）．通常はここには残差が残ります．

例題 2：不足決定系（横長・フルランク）

問題: $A = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}$, $\quad b = \begin{pmatrix} 4 \end{pmatrix}$ に対し一般化逆行列を求めよ．

解法: A は行フルランクなので AA^T が正則です．$A^+ = A^T(AA^T)^{-1}$ を使います．

1. **AA^T の計算**:

$$AA^T = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 1+4+0 = 5$$

2. **逆行列 $(AA^T)^{-1}$**: 単なる数なので，逆数は 1/5.

3. **$A^+ = A^T(AA^T)^{-1}$**:

$$A^+ = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \cdot \frac{1}{5} = \begin{pmatrix} 0.2 \\ 0.4 \\ 0 \end{pmatrix}$$

解: $x = A^+ b = \begin{pmatrix} 0.2 \\ 0.4 \\ 0 \end{pmatrix} \times 4 = \begin{pmatrix} 0.8 \\ 1.6 \\ 0 \end{pmatrix}$．これは $x_1 + 2x_2 = 4$ を満たす無数の解の中で，原点からの距離 $\sqrt{0.8^2 + 1.6^2}$ が最小になるものです．

例題 3：ランク落ち正方行列（SVD 必須）

問題: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ のムーア・ペンローズ逆行列を求めよ．

解法: 行列式が 0 なので，公式は使えません．定義通り SVD を行います．

1. **固有値分解**: 対称行列なので対角化可能です．固有値 $\lambda_1 = 2$（固有ベクトル $v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$）固有値 $\lambda_2 = 0$（固有ベクトル $v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$）

2. **SVD の要素**: $V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ $U = V$（対称行列かつ固有値が正なので $U = V$ とできる）$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$（特異値は固有値の絶対値だが，ここは正なのでそのまま）

3. **Σ^+ の作成**: 非ゼロ成分 2 の逆数を取り，ゼロ成分は放置します．

$$\Sigma^+ = \begin{pmatrix} 0.5 & 0 \\ 0 & 0 \end{pmatrix}$$

4. **復元**:

$$A^+ = V\Sigma^+U^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0.5 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}0.5 & 0.5\\ 0 & 0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}0.5 & 0.5\\ 0.5 & 0.5\end{pmatrix} = \begin{pmatrix}0.25 & 0.25\\ 0.25 & 0.25\end{pmatrix}$$

答え: $A^+ = \frac{1}{4}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$.

まとめ表

例題	行列形状	計算方法	解の特徴
1	縦長 (3 × 2)	$(A^TA)^{-1}A^T$	最小二乗（最も尤もらしい解）
2	横長 (1 × 3)	$A^T(AA^T)^{-1}$	最小ノルム（最も無駄のない解）
3	正方 (2 × 2) ランク 1	SVD (定義)	一般化逆（ランク落ち対応）

12.5 実務への応用：最小二乗問題の決定版

最小二乗問題は，一般化逆行列（ムーア・ペンローズ逆行列 A^+）の**最も重要な応用**です．連立方程式 $Ax = b$ が解を持たない場合（過剰決定系），$x = A^+b$ が**残差ノルム** $\|Ax - b\|_2$ **を最小にする解**を自動的に与えます．

最小二乗問題の定式化

- **問題**: $\min_x \|Ax - b\|_2^2$ を満たす x を求めよ．
- **解**: $x = A^+b$ （一般化逆行列を掛けるだけ）．
- **幾何学的意味**: b を列空間 $\mathrm{Col}(A)$ へ直交射影した点 $P_Ab = Ax$ を求めていることになります．

応用例 5 選（実務直結）

例 1：直線回帰（データフィッティング）

散布データ (x_i, y_i) に直線 $y = ax + b$ を当てはめる場合．

$$A = \begin{pmatrix}x_1 & 1\\ x_2 & 1\\ \vdots & \vdots\\ x_n & 1\end{pmatrix}, \quad b = \begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$$

解は $\begin{pmatrix}a\\ b\end{pmatrix} = A^+b$ で一発で求まります．これは統計学の「最小二乗法」そのものです．

例 2：実験データから二次関数フィッティング

時間 t_i と観測値 y_i から $y = at^2 + bt + c$ を推定する場合．

$$A = \begin{pmatrix}t_1^2 & t_1 & 1\\ t_2^2 & t_2 & 1\\ \vdots & \vdots & \vdots\\ t_n^2 & t_n & 1\end{pmatrix}$$

解は $\begin{pmatrix}a\\ b\\ c\end{pmatrix} = A^+y$．モデルが複雑になっても，行列 A の列を増やすだけで対応できます．

例 3：信号処理（ノイズ除去）

ノイズ混入信号 $y = Ax + e$ から真の信号 x を復元する場合．（A: 測定行列（縦長），x: 真の信号，e: ノイズ）

$$\hat{x} = A^{+}y$$

これにより，観測値 y に含まれるノイズ成分（A の列空間と直交する成分）が自動的に除去され，最も確からしい信号が抽出されます．

例 4：画像圧縮後の復元

SVD で低ランク近似した画像行列 A_k の逆変換．

$$\text{画像 } B \approx A_k x \implies x = A_k^{+} b$$

圧縮されたデータから，オリジナルの画像を推定する際にも使われます．

例 5：センサーデータ融合（ロボット）

複数センサーから位置推定を行う場合（過剰測定）．

$$A = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 & d_1 \\ \cos\theta_2 & \sin\theta_2 & d_2 \\ \vdots & \vdots & \vdots \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ \vdots \end{pmatrix}$$

解 $(x, y) = A^{+}b$ は，複数のセンサー情報を統合した「最尤推定値」を与えます．

理論的根拠と Ridge 回帰

定理 12.1. 任意の A, b に対し，$x = A^{+}b$ は以下を満たす唯一の解である．

$$\begin{cases} Ax \text{ が } b \text{ への直交射影} \quad (\min \|Ax - b\|) \\ x \text{ が最小ノルム（不足決定の場合）} \quad (\min \|x\|) \end{cases}$$

コメント 12.1 (Ridge 回帰への拡張). 実務では，逆行列計算を安定させるために「正則化項」を入れることがあります（Ridge 回帰）．

$$x = (A^T A + \lambda I)^{-1} A^T b$$

これも一般化逆行列の一種とみなせます．A^{+} は $\lambda \to 0$ の極限に対応します．

12.6　ムーア・ペンローズの 4 条件と幾何学的意味

ムーア・ペンローズ逆行列（一般化逆行列）A^{+} は，SVD を使わずとも，以下の **4 つの代数的性質（ムーア・ペンローズ条件）**をすべて満たす唯一の行列として定義することもできます．この 4 条件は，「逆行列の性質 $AA^{-1} = I$ などが成り立たない世界で，何を以て『逆』とするか」という哲学を表しています．

定義 12.2 (ムーア・ペンローズの 4 条件)**.** 行列 G が A の一般化逆行列 A^{+} であるための必要十分条件は，以下の 4 式をすべて満たすことである．

1. $AGA = A$　（A の再生性）
2. $GAG = G$　（A^{+} の再生性）
3. $(AG)^T = AG$　（AG の対称性：列空間への射影）
4. $(GA)^T = GA$　（GA の対称性：行空間への射影）

各条件の幾何学的意味

- **条件 1:** $AGA = A$「A で行って，G（逆）で戻って，もう一度 A で行く」と，元の A と同じ結果になる．これは，「可逆な部分空間（行空間 $\to$ 列空間）」においては，完全な逆行列として振る舞うことを要求しています．
- **条件 2:** $GAG = G$「G（逆）で行って，A で戻って，もう一度 G で行く」と，元の G と同じ結果になる．条件 1 と対になっています．
- **条件 3:** $(AG)^T = AG$ $P = AG$ と置くと，$P^2 = A(GA)G = AG = P$ かつ $P^T = P$ なので，AG は**直交射影行列**になります．具体的には，「像空間（列空間） $\mathrm{Col}(A)$ への直交射影」を表します．これが，過剰決定系で「影（射影）」を求める動作の正体です．
- **条件 4:** $(GA)^T = GA$ 同様に $Q = GA$ も直交射影行列になります．こちらは「定義域の行空間 $\mathrm{Col}(A^T)$ への直交射影」を表します．これが，不足決定系で「最短解」を選ぶ（行空間の成分だけを残し，核の成分を消す）動作に対応します．

主な性質一覧

性質	数式	意味・用途
等価性	A が正則	$A^+ = A^{-1}$ （一致する）
左逆	A が縦長フルランク	$A^+ = (A^TA)^{-1}A^T$ （最小二乗法）
右逆	A が横長フルランク	$A^+ = A^T(AA^T)^{-1}$ （最小ノルム解）
ランク	$\mathrm{rank}(A^+) = \mathrm{rank}(A)$	情報量は保存される
転置	$(A^T)^+ = (A^+)^T$	転置しても関係は変わらない

表 12.1: 一般化逆行列の基本性質と特殊ケース

巻末：線形代数を学ぶあなたへ

これで，本テキストの数学的な旅は終わります．高校数学の「矢印」から始まり，空間の歪み，次元の呪い，スペクトル分解を経て，最後は特異値分解という「データ解析の万能ナイフ」を手に入れました．

もしあなたが今後，

- ディープラーニングの重み行列を解析するとき
- 量子力学で演算子を扱うとき
- 統計学で多重共線性に悩むとき

迷わず $A = U\Sigma V^T$ と書いてみてください．複雑に見える現象も，適切な座標系（基底）を選べば，驚くほど単純な「対角行列」の姿をしているはずです．

線形代数とは，複雑な世界をシンプルに見るための「視点」の学問なのです．