

信号解析の数理

線型代数で信号を理解するために

calamari_dev



はじめに

本書を読み始めるには、行列の計算規則さえ知っていれば十分である。そして、数学的な厳密さにあまり興味がないのなら、ほどよく読み飛ばせばそれ以上の知識を要さず読了できる。背後にある数理に興味がある読者は、付録 A を読むことでしくみをだいたい把握できるだろう。

本書はクリエイティブ・コモンズ 表示-非営利-継承 4.0 国際ライセンスの下に配布している。また、最新版は <https://github.com/calamari-dev/sigproc> から入手できる。

2023 年 3 月 24 日

calamari_dev

目次

はじめに	iii
記号について	vii
第 1 章 準備と前提知識	1
1.1 ベクトル空間と行列	1
ベクトル空間／基底／内積／線型写像と表現行列／核と像／固有値と固有空間／対角化	
1.2 微分積分学	10
上限と下限／数列の極限	
第 2 章 数ベクトル空間	13
2.1 イントロダクション	13
2.2 直交射影	14
直交射影／分析と合成	
2.3 最小 2 乗法	21
最小 2 乗法／スペクトル分解と特異値分解／擬似逆行列と解の構造／ $\mathbb{K} = \mathbb{R}$ の場合／曲線あてはめ	
2.4 離散フーリエ変換	31
離散フーリエ変換／エイリアシング／曲線あてはめと DFT／巡回畳み込み	
2.5 短時間フーリエ変換	41
短時間フーリエ変換	
2.6 多重解像度解析	43
2.7 補遺	43
スペクトル定理の証明／主成分分析	
2.8 演習問題	45

第 3 章	ヒルベルト空間	47
3.1	イントロダクション	47
3.2	無限次元のベクトル空間 距離空間／ノルム空間	48
3.3	ヒルベルト空間	52
3.4	直交射影 直交射影／正規直交系	53
3.5	l^p 空間	57
3.6	フーリエ級数展開	58
3.7	多重解像度解析	59
3.8	演習問題	59
第 4 章	確率空間	61
4.1	イントロダクション	61
4.2	確率空間	61
4.3	ウィーナー解	61
4.4	カルマンフィルタ	61
4.5	演習問題	61
付録 A	測度空間	63
A.1	イントロダクション	63
A.2	測度論の基本概念 σ -加法族／ボレル測度とルベーグ測度	64
A.3	ルベーグ積分 ルベーグ積分／収束定理	68
A.4	確率論の基本概念 確率変数が定める量／条件つき期待値	71
A.5	l^p 空間	75
付録 B	演習問題の略解	79
	参考文献	80
	索引	82

記号について

書籍ごとに異なることが多い記号について，記号と定義の組を示す．表になる記号については，巻末の索引を参照のこと．

記号	定義
\mathbb{N}	自然数の全体集合 $\{1, 2, \dots\}$
\mathbb{Z}	整数の全体集合 $\{\dots, -2, -1, 0, 1, 2, \dots\}$
\mathbb{K}	実数の全体集合 \mathbb{R} か複素数の全体集合 \mathbb{C}
S^c	集合 S の補集合
$\text{cl } S$	集合 S の閉包
δ_{ij}	クロネッカーのデルタ
$\langle \mathbf{u}, \mathbf{v} \rangle$	ベクトル \mathbf{u} , \mathbf{v} の内積
$\ \mathbf{v}\ $	ベクトル \mathbf{v} のノルム
\mathbf{I}	単位行列
\mathbf{O}	零行列
\mathbf{A}^\top	行列 \mathbf{A} の転置行列
\mathbf{A}^H	行列 \mathbf{A} のエルミート転置
$\ \mathbf{A}\ _F$	行列 \mathbf{A} のフロベニウスノルム
$\mathcal{F}_{\mathbb{Z}/N}$	離散フーリエ変換
$\mathcal{F}_{\mathbb{Z}}$	離散時間フーリエ変換
$\mathcal{F}_{\mathbb{T}}$	フーリエ係数列
\mathcal{F}	フーリエ変換

準備と前提知識

1.1 ベクトル空間と行列

ここでは簡単に（特に有限次元の）ベクトル空間が持つ性質を確かめる．省略した証明については，たとえば斎藤 [11] を参照するとよい．

1.1.1 ベクトル空間

以下，集合 \mathbb{K} は実数の全体集合 \mathbb{R} か，複素数の全体集合 \mathbb{C} であるとする． \mathbb{K} 上のベクトル空間とは次のように定義される，加法とスカラー乗法が備わった集合のことである．

定義 1.1.1 (ベクトル空間) V を空でない集合とする．また，任意の $x, y \in V$, $s \in \mathbb{K}$ について，和 $x + y \in V$ とスカラー倍 $sx \in V$ が定義されているとする．任意の $x, y, z \in V$, $s, t \in \mathbb{K}$ に対する以下の条件を満たすとき， V は \mathbb{K} 上の**ベクトル空間** (vector space) であるという．

1. $(x + y) + z = x + (y + z)$
2. $x + y = y + x$
3. ある $0 \in V$ が存在し，任意の $v \in V$ に対して $v + 0 = v$ を満たす．
4. 各 $v \in V$ に対し， $w \in V$ が一意に存在して $v + w = 0$ を満たす．
5. $(s + t)x = sx + tx$
6. $s(x + y) = sx + sy$
7. $(st)x = s(tx)$
8. $1x = x$

しばしば V の元を**ベクトル** (vector)， \mathbb{K} の元を**スカラー** (scalar) と呼

ぶ. また, 定義 1.1.1 の $\mathbf{0}$ を**零ベクトル** (zero vector), \mathbf{w} を \mathbf{v} の**加法逆元** (additive inverse) という. 通常, \mathbf{v} の加法逆元は $-\mathbf{v}$ と表される.

ついで, ベクトル空間にかかわる概念を 2 つ定義する.

定義 1.1.2 (線型結合) V を \mathbb{K} 上のベクトル空間, $\mathbf{v}_1, \dots, \mathbf{v}_n$ を V の元とする. $c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$ ($c_1, \dots, c_n \in \mathbb{K}$) という形をした V の元を, $\mathbf{v}_1, \dots, \mathbf{v}_n$ の**線型結合** (linear combination) という.

定義 1.1.3 (部分空間) V を \mathbb{K} 上のベクトル空間, W を V の空でない部分集合とする. W が V の加法とスカラー乗法について定義 1.1.1 の条件をすべて満たすとき, W は V の**部分ベクトル空間** (vector subspace), あるいは単に**部分空間** (subspace) であるという.

ある部分集合が部分空間かどうか調べるには, 命題 1.1.4 を使うとよい.

命題 1.1.4 V を \mathbb{K} 上のベクトル空間, W を V の空でない部分集合とする. このとき, 次の命題は同値である.

1. W は V の部分空間である.
2. 任意の $s, t \in \mathbb{K}$, $\mathbf{x}, \mathbf{y} \in W$ に対して $s\mathbf{x} + t\mathbf{y} \in W$ である.

例 1.1.5 V がベクトル空間なら, V 自身と $\{\mathbf{0}\}$ は V の部分空間である. ◇

例 1.1.6 集合 $\mathbb{K}^n = \{(s_1 \ \cdots \ s_n)^T \mid s_1, \dots, s_n \in \mathbb{K}\}$ は, 通常の加法とスカラー乗法によって, \mathbb{K} 上のベクトル空間になる. ただし, \mathbf{A}^T は行列 \mathbf{A} の転置行列を意味する. ◇

また, 2 つの部分空間 $W_1, W_2 \subseteq V$ があれば, それらを含むより大きな部分空間を作れる.

定義 1.1.7 (部分空間の和) V を \mathbb{K} 上のベクトル空間, $W_1, W_2 \subseteq V$ を部分空間とする. このとき, 集合 $W = \{\mathbf{w}_1 + \mathbf{w}_2 \mid \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2\}$ は V の部分空間になる. W を W_1 と W_2 の**和** (sum) といい, $W_1 + W_2$ と表記する.

$W_1 \cap W_2 = \{\mathbf{0}\}$ であるとき, $W_1 + W_2$ を W_1 と W_2 の**直和** (direct sum) という. 直和であることを強調したいときは, 和 $W_1 + W_2$ を $W_1 \oplus W_2$ とも書く.

1.1.2 基底

任意のベクトル $\mathbf{x} = (x_1 \ \cdots \ x_n)^T \in \mathbb{K}^n$ は, 第 i 成分が 1, 他の成分が 0 のベクトル \mathbf{e}_i を用いて $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ と表せる. すなわち, 集合 $\mathcal{S}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ は「 \mathbb{K}^n のすべての元を \mathcal{S}_n の元の線型結合で書ける」という性質を持つ.

一般に, ベクトル空間 V の部分集合 S に対して, S の元の線型結合で書けるベクトルの全体集合を S が**生成する部分空間** (generated subspace) といい, $\text{span } S$ と表記する. この記法を使えば, 先述した \mathcal{S}_n が持つ性質を「 $\text{span } \mathcal{S}_n = \mathbb{K}^n$ が成り立つ」と書き表せる.

$\text{span } S = \mathbb{K}^n$ を満たす集合 $S \subseteq \mathbb{K}^n$ は, \mathcal{S}_n 以外にも無数にある. たとえば $\mathbb{K}^n = \mathbb{R}^2$ のとき, 集合 $T = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ ($\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$) が生成する部分空間は \mathbb{R}^2 である. しかし, $\mathcal{S}_2 = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ の元の線型結合で \mathbb{R}^2 の元を表す方法はただ一通りであるのに対して, T はこの性質を持たない. 実際, $\mathbf{x} = \begin{pmatrix} 3/2 \\ 0 \end{pmatrix}$ とすると $\mathbf{x} = (1/2)\mathbf{v}_1 + (1/2)\mathbf{v}_2 = (-3/2)\mathbf{v}_3$ である.



S の元の線型結合で $\text{span } S$ の元を一意に表せるとき, 任意の $a_i, b_i \in \mathbb{K}$, $\mathbf{v}_i \in S$ について

$$\sum_{i=1}^k a_i \mathbf{v}_i = \sum_{i=1}^k b_i \mathbf{v}_i \implies (a_1 \ \cdots \ a_k) = (b_1 \ \cdots \ b_k)$$

が成立する. $b_1 = \cdots = b_k = 0$ とすると

$$a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = \mathbf{0} \implies a_1 = \cdots = a_k = 0 \quad (1.1)$$

が得られる.

任意の $a_1, \dots, a_k \in \mathbb{K}$ に対して式(1.1)が成立するとき, $\mathbf{v}_1, \dots, \mathbf{v}_k$ は線型独立であるという. 特に, $V = \text{span } S$ かつ, S の元からなる有限個のベクトルの組が常に線型独立であるとき, S は V の基底であるという.

定義 1.1.8 (生成系・線型独立・線型従属) V を \mathbb{K} 上のベクトル空間, S を V の部分集合とする. また, $\mathbf{v}_1, \dots, \mathbf{v}_k$ を V の元とする.

1. $V = \text{span } S$ であるとき, S を V の**生成系** (generating set) という.
2. $\sum_{i=1}^k c_i \mathbf{v}_i = \mathbf{0}$ を満たす $c_1, \dots, c_k \in \mathbb{K}$ の組が $c_1 = \dots = c_k = 0$ しかないとき, $\mathbf{v}_1, \dots, \mathbf{v}_k$ は**線型独立** (linearly independent) であるという.
3. $\mathbf{v}_1, \dots, \mathbf{v}_k$ が線型独立でないとき, $\mathbf{v}_1, \dots, \mathbf{v}_k$ は**線型従属** (linearly dependent) であるという.

定義 1.1.9 (基底) V を \mathbb{K} 上のベクトル空間, \mathcal{B} を V の部分集合とする. \mathcal{B} が V の生成系かつ, \mathcal{B} に属する有限個のベクトル $\mathbf{v}_1, \dots, \mathbf{v}_k$ が常に線型独立であるとき, \mathcal{B} は V の**基底** (basis) であるという.

例 1.1.10 (標準基底) \mathcal{S}_n は \mathbb{K}^n の基底である. \mathcal{S}_n を \mathbb{K}^n の**標準基底** (standard basis) という. ◇

さきほどの議論によれば, S の元の線型結合で $\text{span } S$ の元を一意に表せるとき, 任意の $a_1, \dots, a_k \in \mathbb{K}$ について式 (1.1) が成立する. すなわち, S は $\text{span } S$ の基底である. この逆も成り立つので, 次の命題が成立する.

命題 1.1.11 V を \mathbb{K} 上のベクトル空間, S を V の部分集合とする. このとき, 次の命題は同値である.

1. S の元の線型結合で $\text{span } S$ の元を一意に表せる.
2. S は $\text{span } S$ の基底である.

V の基底で有限集合のものがあるとき, V は**有限次元** (finite-dimensional) であるという. V が有限次元なら, V の基底はすべて有限集合で, その元の個数は等しい. すなわち, 元の個数 $\#\mathcal{B}$ は基底 \mathcal{B} のとりかたによらず定まる. $\#\mathcal{B}$ を V の**次元** (dimension) といい, $\dim V$ と表記する¹⁾.

基底に関連して, 次の命題が成り立つ.

1) 任意のベクトル空間は基底を持つ (雪江 [16] に証明がある) が, 有限集合とは限らない.

命題 1.1.12 $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{K}^n$ とする. このとき, 次の命題は同値である.

1. 集合 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ は \mathbb{K}^n の基底である.
2. 行列 $(\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$ は正則である.

命題 1.1.13 (基底の延長) V を \mathbb{K} 上の n 次元ベクトル空間とする. $k < n$ 個のベクトル $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ が線型独立なら, 集合 $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ が V の基底になる $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n \in V$ が存在する.

1.1.3 内積

\mathbb{R}^3 において, ベクトルの長さとなす角はドット積 $(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = \sum_{i=1}^3 x_i y_i$ から計算できた. 次の定義はドット積を抽象化したものである.

定義 1.1.14 (内積) V を \mathbb{K} 上のベクトル空間とする. $\langle _, _ \rangle$ が V の内積 (inner product) であるとは, 任意の $\lambda \in \mathbb{K}$, $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ に対し, $\langle _, _ \rangle$ が以下の条件を満たすことをいう.

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \in \mathbb{K}$
2. $\langle \lambda \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \lambda \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
3. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, $(\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0})$

内積が備わっているベクトル空間のことを**内積空間** (inner product space) という. また, $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ であるときベクトル \mathbf{v} と \mathbf{w} は**直交**するという.

例 1.1.15 (標準内積) $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1^T \mathbf{v}_2$ ($\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{K}^n$) とすると, $\langle _, _ \rangle$ は \mathbb{K}^n の内積になる. $\langle _, _ \rangle$ を \mathbb{K}^n の**標準内積**という. ◇

定義 1.1.16 (正規直交系, 正規直交基底) V を内積空間, \mathcal{B} を V の空でない部分集合とする. \mathcal{B} の相異なる 2 元が常に直交し, すべての $\mathbf{e} \in \mathcal{B}$ が $\langle \mathbf{e}, \mathbf{e} \rangle = 1$ を満たすとき, \mathcal{B} は**正規直交系** (orthonormal system; ONS) であるという. また, 正規直交系である基底を**正規直交基底** (orthonormal basis; ONB) という.

\mathcal{B} が正規直交系なら、有限個の \mathcal{B} の元からなる組はすべて線型独立である。よって、 \mathcal{B} が基底であることを見るには $V = \text{span } \mathcal{B}$ だけ確かめればよい。

また、内積空間に属する線型独立なベクトルの組があれば、それらから正規直交系を作れる。

命題 1.1.17 (グラム・シュミットの直交化法) V を内積空間とする。

$\mathbf{a}_1, \dots, \mathbf{a}_n \in V$ が線型独立なら

$$\mathbf{u}_1 = \mathbf{a}_1, \quad \mathbf{u}_k = \mathbf{a}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{a}_k, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \quad (k \geq 2)$$

で \mathbf{u}_k を定義すると、集合 $\{\mathbf{u}_k / \sqrt{\langle \mathbf{u}_k, \mathbf{u}_k \rangle} \mid k = 1, \dots, n\}$ は V の正規直交系になる。正規直交系を作るこの方法を **グラム・シュミットの直交化法** (Gram-Schmidt orthogonalization) という。

1.1.4 線型写像と表現行列

V は有限次元とする。命題 1.1.11 によれば、 V の基底 $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ を選ぶと、任意の $\mathbf{x} \in V$ を

$$\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_m \mathbf{v}_m \quad (c_1, \dots, c_m \in \mathbb{K}) \quad (1.2)$$

の形で一意に表せる。言い換えれば、 V の各元 \mathbf{x} に式 (1.2) の $(c_1 \ \dots \ c_m)^T$ を割り当てる写像 $\phi: V \rightarrow \mathbb{K}^m$ を定義でき、それは単射²⁾である。この写像 ϕ は、次に定義する「線型写像」の一例である。

定義 1.1.18 (線型写像) V と W を \mathbb{K} 上のベクトル空間とする。写像 $f: V \rightarrow W$ が以下の条件を満たすとき、 f は**線型写像** (linear mapping) であるという。

1. 任意の $\mathbf{x}, \mathbf{y} \in V$ に対して $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ である。
2. 任意の $c \in \mathbb{K}$, $\mathbf{x} \in V$ に対して $f(c\mathbf{x}) = cf(\mathbf{x})$ である。

2) 写像 f の定義域に属する任意の \mathbf{x}, \mathbf{y} について $f(\mathbf{x}) = f(\mathbf{y}) \implies \mathbf{x} = \mathbf{y}$ が成立するとき、 f は**単射** (injection) であるという。

W を \mathbb{K} 上の有限次元ベクトル空間とする. W の基底 $\mathcal{B}' = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ ($n = \dim W$) をとると, ϕ と同様

$$\mathbf{y} = d_1 \mathbf{w}_1 + \dots + d_n \mathbf{w}_n \iff \psi(\mathbf{y}) = (d_1 \ \dots \ d_n)^\top$$

を満たす線型写像 $\psi: W \rightarrow \mathbb{K}^n$ が定義できる.

ϕ と ψ を利用すると, V から W への任意の線型写像 f を, 対応する行列によって表現できる. $\mathbf{x} \in V$ を任意にとる. $\phi(\mathbf{x}) = (c_1 \ \dots \ c_m)^\top$ とおくと

$$\begin{aligned} f(\mathbf{x}) &= f\left(\sum_{i=1}^m c_i \mathbf{v}_i\right) = \sum_{i=1}^m c_i f(\mathbf{v}_i), \\ \psi(f(\mathbf{x})) &= \sum_{i=1}^m c_i \psi(f(\mathbf{v}_i)) = (\psi(f(\mathbf{v}_1)) \ \dots \ \psi(f(\mathbf{v}_m))) \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \end{aligned}$$

であるから, $\mathbf{A} = (\psi(f(\mathbf{v}_1)) \ \dots \ \psi(f(\mathbf{v}_m)))$ とおくと

$$\psi(f(\mathbf{x})) = T_A(\phi(\mathbf{x})) \quad (T_A(\mathbf{x}) = \mathbf{A}\mathbf{x}) \quad (1.3)$$

が成り立つ.

ここまでの議論をまとめると, 次のようになる. V の基底 \mathcal{B} と, W の基底 \mathcal{B}' をとるごとに, $n \times m$ 行列 $\mathbf{A} = (\psi(f(\mathbf{v}_1)) \ \dots \ \psi(f(\mathbf{v}_m)))$ を定義でき, \mathbf{A} は式 (1.3) を満たす. この \mathbf{A} を, 基底 \mathcal{B} と \mathcal{B}' に関する f の**表現行列** (representation matrix) という.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow \phi & & \downarrow \psi \\ \mathbb{K}^m & \xrightarrow{T_A} & \mathbb{K}^n \end{array}$$

なお, \mathcal{B} の元を並べる順序に応じて, 式 (1.2) の c_1, \dots, c_n の順序も変化するので, ϕ は \mathcal{B} に対して一意ではない. ϕ は \mathcal{B} の元を並べる順序を決めて初めて定まる. 本書では, $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ のような書き方をした場合, \mathcal{B} の元を $\mathbf{v}_1, \mathbf{v}_2, \dots$ の順に並べると決めておく.

例 1.1.19 (形式的な微分) 次数が n 未満の 1 変数多項式全体 $V_n = \{c_0 + c_1 x + \dots + c_{n-1} x^{n-1} \mid c_0, \dots, c_{n-1} \in \mathbb{R}\}$ は, \mathbb{R} 上の n 次元ベクトル空間である. また, 写像 $D: V_3 \rightarrow V_2$ を $D(c_0 + c_1 x + c_2 x^2) = c_1 + 2c_2 x$ で定義すると, これは線型写像になる. V_n の基底として $\mathcal{B}_n = \{1, x, \dots, x^{n-1}\}$ をとったとき, 基底 \mathcal{B}_3 と \mathcal{B}_2 に関する D の表現行列は $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ である. \diamond

1.1.5 核と像

線型写像に付随して, 重要なベクトル空間が 2 つ定まる.

定義 1.1.20 (核, 像) $f: V \rightarrow W$ を線型写像とする.

1. V の部分空間 $f^{-1}[\{\mathbf{0}\}]$ を f の **核** (kernel) といい, $\ker f$ と表す.
2. W の部分空間 $f[V]$ を f の **像** (image) といい, $\operatorname{im} f$ と表す.

ノート 写像 $f: X \rightarrow Y$ と集合 $A \subseteq X$, $B \subseteq Y$ に関して $f[A] = \{f(x) \mid x \in A\}$, $f^{-1}[B] = \{x \in X \mid f(x) \in B\}$ である. $f[A]$ を f による A の **像** (image), $f^{-1}[B]$ を f による B の **逆像** (inverse image) という.

命題 1.1.21 $f: V \rightarrow W$ を線型写像とする. このとき, f が単射であることと, $\ker f = \{\mathbf{0}\}$ が成立することは同値である.

証明 f が単射なら, $f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}) + f(\mathbf{0})$ より $f(\mathbf{0}) = \mathbf{0}$, $\ker f = \{\mathbf{0}\}$ である. 逆に $\ker f = \{\mathbf{0}\}$ なら, $\mathbf{v}_1, \mathbf{v}_2 \in V$ が $f(\mathbf{v}_1) = f(\mathbf{v}_2)$ を満たすとき $f(\mathbf{v}_1 - \mathbf{v}_2) = f(\mathbf{v}_1) - f(\mathbf{v}_2) = \mathbf{0}$, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$ である. よって, このとき f は単射である. \square

1.1.6 固有値と固有空間

対角化に向けて, 固有値に関連する事項を整理する.

定義 1.1.22 (固有値, 固有ベクトル) A を n 次正方行列とする. 複素数 λ と $\mathbf{0}$ でないベクトル $\mathbf{x} \in \mathbb{C}^n$ が式 $A\mathbf{x} = \lambda\mathbf{x}$ を満たすとき, λ を A の **固有値** (eigenvalue) という. また, \mathbf{x} を A の (固有値 λ に属する) **固有ベクトル** (eigenvector) という.

定義 1.1.23 (固有空間) 定義 1.1.22 の A , λ について, 集合 $E_\lambda(A) = \{\mathbf{x} \in \mathbb{C}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$ は \mathbb{C}^n の部分空間になる. 部分空間 $E_\lambda(A)$ を, A の (固有値 λ に属する) **固有空間** (eigenspace) という.

例 1.1.24 $\mathbf{x}_1 = (1 + i \ 2)^\top$, $\mathbf{x}_2 = (1 - i \ 2)^\top$ は $A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$ の固有ベクトルである. 実際 $A\mathbf{x}_1 = i\mathbf{x}_1$, $A\mathbf{x}_2 = -i\mathbf{x}_2$ である. \diamond

以下, A を任意の n 次正方行列, $\operatorname{spec} A$ を A の固有値の全体集合とする.

n 次多項式 $p(\lambda) = \det(\lambda I - A)$ を A の**固有多項式** (characteristic polynomial) という. 命題 1.1.25 から, A の固有値を求めるには方程式 $p(\lambda) = 0$ を解けばよい.

命題 1.1.25 $\text{spec } A = \{\lambda \in \mathbb{C} \mid \det(\lambda I - A) = 0\}$ である.

また, 固有空間は次の性質を持つ.

命題 1.1.26 $\lambda_1, \lambda_2 \in \text{spec } A$, $\lambda_1 \neq \lambda_2$ ならば $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$ である.

証明 $\mathbf{x} \in E_{\lambda_1}(A) \cap E_{\lambda_2}(A)$ とすると, $A\mathbf{x} = \lambda_1\mathbf{x} = \lambda_2\mathbf{x}$ より $(\lambda_1 - \lambda_2)\mathbf{x} = \mathbf{0}$ であり, $\lambda_1 \neq \lambda_2$ なので $\mathbf{x} = \mathbf{0}$ である. よって $E_{\lambda_1}(A) \cap E_{\lambda_2}(A)$ は $\mathbf{0}$ 以外に元を持たない. \square

1.1.7 対角化

適当な n 次正則行列 P , 対角行列 Λ の組を見つけて, n 次正方行列 A を $A = P\Lambda P^{-1}$ の形で書くことを A の**対角化** (diagonalization) という. A が対角化可能である必要十分条件は, 次の命題 1.1.27 で与えられる.

命題 1.1.27 n 次正方行列 A に関して, 以下の条件は同値である.

1. A の固有ベクトルのみからなる \mathbb{K}^n の基底が存在する.
2. $\mathbb{K}^n = \bigoplus_{\lambda \in \text{spec } A} E_{\lambda}(A)$ が成立する.
3. n 次正則行列 P , 対角行列 Λ が存在して $A = P\Lambda P^{-1}$ を満たす.

以下, 対角行列 $\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ を $\text{diag}(a_1, \dots, a_n)$ と略記する.

証明 1 と 3 の同値性のみ示す. A の固有ベクトルのみからなる \mathbb{K}^n の基底 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ があるとき, A は対角化できることを示す. $P = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)$ とおく. このとき, 各 \mathbf{v}_i に対応する固有値を λ_i とおくと $AP = (A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n) = (\lambda_1\mathbf{v}_1 \ \cdots \ \lambda_n\mathbf{v}_n)$ だから, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ とおくと $AP = P\Lambda$, $A = P\Lambda P^{-1}$ となる. ただし, P の逆行列が存在することは命題 1.1.12 による.

逆に, $A = PAP^{-1}$ を満たす n 次正則行列 P , 対角行列 A が存在したとする. $P = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)$, $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ とおく. すると $(A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n) = AP = PA = (\lambda_1\mathbf{v}_1 \ \cdots \ \lambda_n\mathbf{v}_n)$ なので, 各 λ_i , \mathbf{v}_i は $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ を満たす. また, P は正則だから $\mathbf{v}_i \neq \mathbf{0}$ である. よって命題 1.1.12 より, 集合 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ は A の固有ベクトルからなる \mathbb{K}^n の基底である. \square

1.2 微分積分学

ここでは ε - N 論法による極限の定義を既知として, 実数の性質からしたがう事実をいくつか挙げる. なお, 紙幅の都合で証明はほぼ省略した. 興味があれば杉浦 [12] を参照するとよい.

1.2.1 上限と下限

定義 1.2.1 (上界, 下界) X を \mathbb{R} の部分集合とする.

1. 実数 a が X の **上界** (upper bound) であるとは, 任意の $x \in X$ に対して $x \leq a$ が成立することをいう.
2. 実数 b が X の **下界** (lower bound) であるとは, 任意の $x \in X$ に対して $x \geq b$ が成立することをいう.

X の上界が存在するとき, X は **上に有界** であるという. 同様に, X の下界が存在するとき, X は **下に有界** であるという. X が上にも下にも有界——つまり集合 $\{x \mid x \in X\}$ が上に有界——なときは, 単に **有界** であるという.

定義 1.2.2 (上限, 下限) X を \mathbb{R} の空でない部分集合とする.

1. X が上に有界であれば, 上界の全体集合は最小元 m を持つ. m を X の **上限** (supremum) といい, $\sup X$ と書く.
2. X が下に有界であれば, 下界の全体集合は最大元 M を持つ. M を X の **下限** (infimum) といい, $\inf X$ と書く.

便宜上, X が上に有界でないときは $\sup X = +\infty$, X が下に有界でないときは $\inf X = -\infty$ と決めておく. 上限と下限を用いて議論するときは, 次の命

題 1.2.3 が便利である.

命題 1.2.3 X を \mathbb{R} の部分集合とする. このとき, 実数 s に関する以下の条件は同値であり, 同様のことが $\inf X$ についても成り立つ.

1. $s = \sup X$ である.
2. 任意の $\varepsilon > 0$ に対し, $x \in X$ が存在して $x + \varepsilon > s$ を満たす.

1.2.2 数列の極限

数列の各項からなる集合が有界であるとき, その数列は有界であるという.

命題 1.2.4 実数列 $\{a_n\}_{n \in \mathbb{N}}$ は上に有界とする. このとき, $a_1 \leq a_2 \leq \dots$ なら $\{a_n\}_{n \in \mathbb{N}}$ は収束する.

証明 $\alpha = \sup S$ とする. このとき $a_n \rightarrow \alpha$ ($n \rightarrow \infty$) であることを示す. 任意に $\varepsilon > 0$ をとる. 命題 1.2.3 より, $x + \varepsilon > \alpha$ となる $x \in S$ がある. $x = a_N$ を満たす N について, $n \geq N$ なら $a_N \leq a_n \leq \alpha$, $|a_n - \alpha| = \alpha - a_n \leq \alpha - a_N < \varepsilon$ である. よって $a_n \rightarrow \alpha$ ($n \rightarrow \infty$) である. \square

定義 1.2.5 (上極限, 下極限) 実数列 $\{a_n\}_{n \in \mathbb{N}}$ に対して, $\limsup_{n \rightarrow \infty} a_n$ と $\liminf_{n \rightarrow \infty} a_n$ を

$$\begin{aligned}\limsup_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \sup\{a_n, a_{n+1}, \dots\}, \\ \liminf_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \inf\{a_n, a_{n+1}, \dots\}\end{aligned}$$

で定義する. $\limsup_{n \rightarrow \infty} a_n$ を数列 $\{a_n\}_{n \in \mathbb{N}}$ の**上極限** (limit superior), $\liminf_{n \rightarrow \infty} a_n$ を数列 $\{a_n\}_{n \in \mathbb{N}}$ の**下極限** (limit inferior) という.

命題 1.2.4 より $\liminf_{n \rightarrow \infty} a_n < +\infty \iff \sup\{a_1, a_2, \dots\} < +\infty$ であり, 同様のことが $\limsup_{n \rightarrow \infty} a_n$ についても成り立つ. また, 次のことが知られている.

命題 1.2.6 $\{a_n\}_{n \in \mathbb{N}}$ を実数列とする。このとき、実数 α に関する以下の条件は同値である。

1. $a_n \rightarrow \alpha$ ($n \rightarrow \infty$) である。
2. $\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \alpha$ である。

第 3 章以降では、望ましい性質を持つ収束列を定義して、その極限によって命題を示すことが多くなる。極限值が予想できる場合を除き、数列が収束することを示すには、それがコーシー列であることを示すのがよい。

定義 1.2.7 (コーシー列) $\{x_n\}_{n \in \mathbb{N}}$ を実数列とする。 $\{x_n\}_{n \in \mathbb{N}}$ が**コーシー列** (Cauchy sequence) であるとは、任意の $\varepsilon > 0$ に対し、 $N \in \mathbb{N}$ が存在して $(m, n > N \implies |x_m - x_n| < \varepsilon)$ を満たすことをいう。このことを次のように表す。

$$|x_m - x_n| \rightarrow 0 \quad (m, n \rightarrow \infty), \quad \lim_{m, n \rightarrow \infty} |x_m - x_n| = 0$$

実数列について、数列がコーシー列であることと収束列であることは同値である。ただし、コーシー列であることは極限值を使わずに確かめられる。

例 1.2.8 $a_n = \sum_{i=0}^n (-1)^i / (i!)$ とする。このとき、 $m > n \geq N$ なら

$$|a_m - a_n| \leq \sum_{i=n+1}^m \frac{1}{i!} \leq \sum_{i=n+1}^m \frac{1}{i(i-1)} = \sum_{i=n+1}^m \left(\frac{1}{i-1} - \frac{1}{i} \right) = \frac{1}{n} - \frac{1}{m}$$

であり、 $1/n - 1/m \leq 1/N \rightarrow 0$ ($N \rightarrow \infty$) だから $\{a_n\}_{n \in \mathbb{N}}$ はコーシー列である。よって、級数 $\sum (-1)^n / (n!)$ は収束する。 \diamond

ノート 実は $a_n \rightarrow 1/e$ ($n \rightarrow \infty$) である。しかし、 $a_n \rightarrow 1/e$ ($n \rightarrow \infty$) を示すよりも $|a_m - a_n| \rightarrow 0$ ($m, n \rightarrow \infty$) を示すほうがずっとやさしい。

数ベクトル空間

2.1 イントロダクション

第 2 章では、数ベクトル空間 \mathbb{K}^n ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) に関する理論を扱う。信号解析において、この理論は

1. 離散時間信号の時系列分析
2. 観測値をモデルに対応づける回帰・判別分析

という、2 つの方向に応用される。

音声信号処理は前者の主要な例である。音声信号を計算機で処理するには、時々刻々と値が変わる信号を有限長のデータで表現しなければならない。たとえば、CD では音声信号の瞬時値を 1 秒あたり 44100 個記録している。すなわち、時刻 t 秒における瞬時値を $x(t)$ 、収録時間を T 秒とおくと、CD には数列 $\{x(n/44100)\}_{n=0}^{44100T-1}$ が記録されている。そこで、収録されたデータを \mathbb{K}^{44100T} の元とみなせば、 \mathbb{K}^n に関する理論に基づいて音声を解析できる。

後者の主要な例は最小 2 乗法である。実験で得られた標本を理論と見比べるとき、理論から得られる式へのあてはめ（回帰）がしばしば試される。あてはまりのよさを示す指標はいろいろあるが、最もポピュラーなのは 2 乗誤差を指標にする最小 2 乗法である。本書ではこの最小 2 乗法を、内積と関連づけ幾何的に説明する。

2.2 直交射影

本節では、あるベクトルを他のベクトルの線型結合で近似する手法を説明する。第2章において \mathbb{K} は \mathbb{R} か \mathbb{C} を意味し、 $\langle _, _ \rangle$ は \mathbb{K}^n の標準内積を意味する。また、 $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ を \mathbf{x} のノルム (norm) と呼ぶ。

2.2.1 直交射影



\mathbb{K}^n のベクトル \mathbf{x} 、部分空間 V が与えられたとき、 V の元で \mathbf{x} に最も近いベクトル、すなわち、距離 $\|\mathbf{x} - \mathbf{m}\|$ を最小にする $\mathbf{m} \in V$ について考えよう。

\mathbb{K}^n が平面 \mathbb{R}^2 で、 V があるベクトル $\mathbf{v} \neq \mathbf{0}$ により生成される直線 $\text{span}\{\mathbf{v}\}$ の場合、 \mathbf{m} は図の位置にある。図を見ると、 $\mathbf{x} - \mathbf{m}$ は \mathbf{v} と直交しているのが分かる。

一般の部分空間 $V \subseteq \mathbb{K}^n$ においても、直交性と最良近似には密接な関係がある。その証明へと入る前に、便利な記法を2つ定義しておく。

定義 2.2.1 (argmin, argmax) 集合 S を定義域に含む実数値関数 f に対して、集合 $\arg \min_{x \in S} f(x)$, $\arg \max_{x \in S} f(x)$ を以下の通り定義する。

$$\arg \min_{x \in S} f(x) = \{x \in S \mid \text{任意の } y \in S \text{ に対して } f(y) \geq f(x)\},$$

$$\arg \max_{x \in S} f(x) = \{x \in S \mid \text{任意の } y \in S \text{ に対して } f(y) \leq f(x)\}$$

定義 2.2.1 からただちに、次のことが分かる。

命題 2.2.2 S の元 a に関する以下の条件は同値であり、同様のことが $\arg \max$ についても成り立つ。

1. $a \in \arg \min_{x \in S} f(x)$ である。
2. $f(a)$ は集合 $f[S] = \{f(x) \mid x \in S\}$ の最小元である。

図は e^{-x} と $|\sin x|$ のグラフである。 $e^{-x} \rightarrow 0$ ($x \rightarrow \infty$) であるが、 $e^{-x} = 0$ となる実数 x は存在しない。そのため

$$\begin{aligned}\arg \min_{x \in [0, +\infty)} e^{-x} &= \emptyset \quad (\text{空集合}), \\ \arg \min_{x \in [0, +\infty)} |\sin x| &= \{0, \pi, 2\pi, \dots\}\end{aligned}$$



である。このように、 $\arg \min_{x \in S} f(x)$ は空になることも、無限集合になることもある。

$\mathbb{K} = \mathbb{R}$ の場合も同様に証明できるので、命題 2.2.6 まで証明では $\mathbb{K} = \mathbb{C}$ を仮定する。また、部分空間が $\{\mathbf{0}\}$ でないことも仮定する。

補題 2.2.3 $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2 \operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$ ($\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$) である。

証明 $\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle$ の右辺を展開すれば示せる。 □

命題 2.2.4 $\mathbf{x} \in \mathbb{K}^n$ かつ、 V は \mathbb{K}^n の部分空間とする。このとき、 $\arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ はただ一つ元からなる集合である。

証明 本証明に限り、 $\sum_{i=1}^m$ ($m = \dim V$) を Σ と略記する。任意に $\mathbf{y} \in V$ をとる。 $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ を V の正規直交基底とすると、 \mathbf{y} は \mathcal{B} の元の線型結合で $\mathbf{y} = \sum z_i \mathbf{e}_i$ と書ける。また $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$ だから

$$\left\| \sum z_i \mathbf{e}_i \right\|^2 = \left\langle \sum_{i=1}^m z_i \mathbf{e}_i, \sum_{j=1}^m z_j \mathbf{e}_j \right\rangle = \sum_{i=1}^m z_i \sum_{j=1}^m \bar{z}_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum z_i \bar{z}_i = \sum |z_i|^2$$

である。よって、補題 2.2.3 より $s_i = \operatorname{Re} z_i$, $t_i = \operatorname{Im} z_i$ とおくと

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|^2 &= \left\| \mathbf{x} - \sum z_i \mathbf{e}_i \right\|^2 = \|\mathbf{x}\|^2 - 2 \operatorname{Re} \langle \mathbf{x}, \sum z_i \mathbf{e}_i \rangle + \left\| \sum z_i \mathbf{e}_i \right\|^2 \\ &= \|\mathbf{x}\|^2 + \sum (-2 \operatorname{Re}(\bar{z}_i \langle \mathbf{x}, \mathbf{e}_i \rangle) + |z_i|^2) \\ &= \|\mathbf{x}\|^2 + \sum (-2 \operatorname{Re}((s_i - it_i) \langle \mathbf{x}, \mathbf{e}_i \rangle) + s_i^2 + t_i^2) \\ &= \|\mathbf{x}\|^2 + \sum (-2(s_i \operatorname{Re} \langle \mathbf{x}, \mathbf{e}_i \rangle + t_i \operatorname{Im} \langle \mathbf{x}, \mathbf{e}_i \rangle) + s_i^2 + t_i^2) \\ &= \|\mathbf{x}\|^2 + \sum ((s_i - \operatorname{Re} \langle \mathbf{x}, \mathbf{e}_i \rangle)^2 + (t_i - \operatorname{Im} \langle \mathbf{x}, \mathbf{e}_i \rangle)^2 - |\langle \mathbf{x}, \mathbf{e}_i \rangle|^2)\end{aligned}$$

なので、次式が成立する.

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \sum_{i=1}^m |z_i - \langle \mathbf{x}, \mathbf{e}_i \rangle|^2 - \sum_{i=1}^m |\langle \mathbf{x}, \mathbf{e}_i \rangle|^2 \quad (2.1)$$

右辺の値が最も小さくなるのは $z_i = \langle \mathbf{x}, \mathbf{e}_i \rangle$ のときであり, そのとき $\mathbf{y} = \sum \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i$ だから $\arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\| = \{\sum \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i\}$ である. \square

命題 2.2.5 $\mathbf{x} \in \mathbb{K}^n$ かつ, V は \mathbb{K}^n の部分空間とする. V のある元 \mathbf{m} が任意の $\mathbf{v} \in V$ に対し $\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle = 0$ を満たすとき, $\mathbf{m} \in \arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ である.

証明 任意に $\mathbf{y} \in V$ をとり, $\Delta \mathbf{y} = \mathbf{m} - \mathbf{y}$ とおく. すると, $\langle \mathbf{x} - \mathbf{m}, \Delta \mathbf{y} \rangle = 0$ より $\|\mathbf{x} - \mathbf{y}\|^2 = \|(\mathbf{x} - \mathbf{m}) + \Delta \mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{m}\|^2 + \|\Delta \mathbf{y}\|^2$ が成立する. よって $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{m}\|$ だから, $\mathbf{m} \in \arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ である. \square

命題 2.2.6 $\mathbf{x} \in \mathbb{K}^n$ かつ, V は \mathbb{K}^n の部分空間とする. このとき, V の元 \mathbf{m} に関する以下の条件は同値であり, 条件を満たす \mathbf{m} はただ一つ存在する.

1. $\mathbf{m} \in \arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ である.
2. 任意の $\mathbf{v} \in V$ に対して $\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle = 0$ である.

証明 「2 ならば 1」は命題 2.2.5 で示した. $\mathbf{m} \in \arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ のとき, すべての $\mathbf{v} \in V$ に対して $\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle = 0$ が成り立つことを示す. それには $\|\mathbf{v}\| = 1$ のときについて示せば十分である. $\mathbf{m} \in \arg \min_{\mathbf{y} \in V} \|\mathbf{x} - \mathbf{y}\|$ なので, 関数 $\delta(z) = \|\mathbf{x} - (\mathbf{m} + z\mathbf{v})\|^2 - \|\mathbf{x} - \mathbf{m}\|^2$ ($z \in \mathbb{C}$) は負の値をとらない. 一方, $x = \operatorname{Re} z$, $y = \operatorname{Im} z$ とおくと

$$\begin{aligned} \delta(z) &= \|(\mathbf{x} - \mathbf{m}) - z\mathbf{v}\|^2 - \|\mathbf{x} - \mathbf{m}\|^2 = -2 \operatorname{Re}(\bar{z}\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle) + |z|^2 \|\mathbf{v}\|^2 \\ &= -2(x \operatorname{Re}\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle + y \operatorname{Im}\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle) + x^2 + y^2 \\ &= (x - \operatorname{Re}\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle)^2 + (y - \operatorname{Im}\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle)^2 - |\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle|^2 \\ &= |z - \langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle|^2 - |\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle|^2 \end{aligned}$$

より $|\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle|^2 = -\delta(\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle) \leq 0$, よって $\langle \mathbf{x} - \mathbf{m}, \mathbf{v} \rangle = 0$ である. \mathbf{m} が存在し一意であることは命題 2.2.4 による. \square

定義 2.2.7 (直交射影) 命題 2.2.6 の m を x の V への直交射影 (orthogonal projection) といい, $\text{proj}_V x$ と表す.



図 2.1 x の $V = \text{span}\{v_1, v_2\}$ への直交射影 $m = \text{proj}_V x$ の模式図.

例 2.2.8 $V = \text{span}\{i, j\}$ ($i = (1 \ 0 \ 0)^T$, $j = (0 \ 1 \ 0)^T$) で \mathbb{R}^3 の部分空間 V を定める. このとき, 集合 $\{i, j\}$ は V の正規直交基底なので $\text{proj}_V r = \langle r, i \rangle i + \langle r, j \rangle j = (x \ y \ 0)^T$ ($r = (x \ y \ z)^T$) である. ◇

命題 2.2.9 \mathbb{K}^n の任意の部分空間 V について, 写像 $\text{proj}_V: \mathbb{K}^n \rightarrow V$ は線型写像である.

証明 $s, t \in \mathbb{K}$, $x, y \in \mathbb{K}^n$ を任意にとり, $m = s \text{proj}_V(x) + t \text{proj}_V(y)$, $z = sx + ty$ とおく. このとき, 各 $v \in V$ に対し $\langle z - m, v \rangle = s \langle x - \text{proj}_V x, v \rangle + t \langle y - \text{proj}_V y, v \rangle = s0 + t0 = 0$ なので $\text{proj}_V z = m$ である. □

定義 2.2.10 (直交補空間) V を \mathbb{K}^n の部分空間とする. W が V の部分空間なら, 集合

$$\{v \in V \mid \text{任意の } w \in W \text{ に対して } \langle v, w \rangle = 0\}$$

も V の部分空間になる. この部分空間を (V における) W の直交補空間 (orthogonal complement) といい, $W^{\perp V}$, W^{\perp} などと書く.

例 2.2.11 $W = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ を \mathbb{R}^3 の 2 次元部分空間とする. このとき, W^\perp は \mathbf{e}_1 と \mathbf{e}_2 に直交し $\mathbf{0}$ でないベクトル \mathbf{e}_3 が生成する直線 $\text{span}\{\mathbf{e}_3\}$ である. $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = 0$ なら集合 $\{\mathbf{e}_i / \|\mathbf{e}_i\| \mid i = 1, 2, 3\}$ は \mathbb{R}^3 の正規直交基底である. ◇

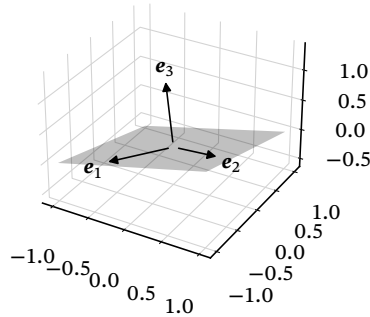


図 2.2 W と $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ の様子.

命題 2.2.12 V は \mathbb{K}^n の部分空間で, W は V の部分空間とする. このとき $V = W \oplus W^\perp$ である.

証明 $\mathbf{x} \in W \cap W^\perp$ なら $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ より $\mathbf{x} = \mathbf{0}$ である. また, 任意の $\mathbf{x} \in W$ に対して $\mathbf{x} - \text{proj}_W \mathbf{x} \in W^\perp$, $\mathbf{x} = \text{proj}_W(\mathbf{x}) + (\mathbf{x} - \text{proj}_W \mathbf{x}) \in W + W^\perp$ である. よって $V = W \oplus W^\perp$ である. □

2.2.2 分析と合成

命題 2.2.4 の証明では, $\text{proj}_V \mathbf{x}$ の存在を示すために V の正規直交基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ を 1 つ選び, $\text{proj}_V \mathbf{x}$ を $\sum_{i=1}^m \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i$ と表した. 一方で, \mathbf{x} の性質を調べるのに使いたい \mathbb{C}^n の正規直交基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ がさきにあって, そこから部分空間 $V_m = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ ($m = 1, \dots, n$) への直交射影 $\text{proj}_{V_m} \mathbf{x}$ を作ることも多い. そのような場合, 直交射影は 3 つの操作に分解できる.

定義 2.2.13 (エルミート転置) A を $m \times n$ 複素行列とする. $n \times m$ 行列 \bar{A}^T を A のエルミート転置 (Hermitian transpose) といい¹⁾, A^H と表す.

$U = (\mathbf{e}_1 \ \cdots \ \mathbf{e}_n)$, $\Delta = \begin{pmatrix} I_m & \\ & \mathbf{0}_{n-m} \end{pmatrix}$ とおく (I_m は m 次単位行列, $\mathbf{0}_{n-m}$ は $n-m$ 次零行列). このとき, 任意の $\mathbf{x} = (x_1 \ \cdots \ x_n)^\top \in \mathbb{C}^n$ に対して

$$U^H \mathbf{x} = \begin{pmatrix} \langle \mathbf{e}_1^H, \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{e}_n^H, \mathbf{x} \rangle \end{pmatrix} = \begin{pmatrix} \langle \mathbf{x}, \mathbf{e}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{e}_n \rangle \end{pmatrix}, \quad \Delta \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \\ \mathbf{0} \end{pmatrix}, \quad U \mathbf{x} = U \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i \mathbf{e}_i$$

であるから

$$U \Delta U^H \mathbf{x} = U \Delta \begin{pmatrix} \langle \mathbf{x}, \mathbf{e}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{e}_n \rangle \end{pmatrix} = U \begin{pmatrix} \langle \mathbf{x}, \mathbf{e}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{e}_m \rangle \\ \mathbf{0} \end{pmatrix} = \sum_{i=1}^m \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i = \text{proj}_{V_m} \mathbf{x}$$

であり, $\text{proj}_{V_m} \mathbf{x} = U \Delta U^H \mathbf{x}$ が成立する. よって, proj_{V_m} は線型写像 $T(\mathbf{x}) = U^H \mathbf{x}$, $D(\mathbf{x}) = \Delta \mathbf{x}$, $T^*(\mathbf{x}) = U \mathbf{x}$ を用いて $\text{proj}_{V_m} = T^* D T$ と表せる.

$T(\mathbf{x})$ の第 i 成分 $\langle \mathbf{x}, \mathbf{e}_i \rangle$ は, \mathbf{x} に含まれる \mathbf{e}_i の「成分」を表すと考えられる. その理由は 2 つある. 1 つめの理由は, $\|\text{proj}_{\text{span}\{\mathbf{e}_i\}} \mathbf{x}\| = \|\langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i\| = |\langle \mathbf{x}, \mathbf{e}_i \rangle|$ なので, $|\langle \mathbf{x}, \mathbf{e}_i \rangle|$ が \mathbf{e}_i のスカラー倍で \mathbf{x} を最もよく近似するベクトルの長さを表すことである. もう一つの理由は, \mathcal{B} は \mathbb{K}^n の正規直交基底であるから

$$\mathbf{x} = \text{proj}_{V_n} \mathbf{x} = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i \quad (2.2)$$

が成立し, $\langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i$ の和で \mathbf{x} が表されることである.

以上の理由から, 本書では線型写像 $T(\mathbf{x}) = (\langle \mathbf{x}, \mathbf{e}_1 \rangle \ \cdots \ \langle \mathbf{x}, \mathbf{e}_n \rangle)^\top$ を分解作用素, $T^*(\mathbf{x}) = \sum_{i=1}^n x_i \mathbf{e}_i$ を合成作用素と呼ぶ.

定義 2.2.14 (分解作用素, 合成作用素) $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ を \mathbb{K}^n の正規直交基底とする.

1. 線型写像 $T: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $T(\mathbf{x}) = (\langle \mathbf{x}, \mathbf{e}_1 \rangle \ \cdots \ \langle \mathbf{x}, \mathbf{e}_n \rangle)^\top$ を \mathcal{B} に関する**分解作用素** (analysis operator) という.
2. 線型写像 $T^*: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $T^*((x_1 \ \cdots \ x_n)^\top) = \sum_{i=1}^n x_i \mathbf{e}_i$ を \mathcal{B} に関する**合成作用素** (synthesis operator) という.

1) エルミート転置は**随伴行列** (adjoint matrix) と呼ばれることも多い.

式 (2.2) より, 合成作用素は分解作用素の逆写像である. また, 分解作用素と合成作用素が持つ性質は, 表現行列に関する条件へと言い換えられる.

定義 2.2.15 (正規行列, ユニタリ行列) A を n 次複素正方行列とする.

1. $A^H A = A A^H$ であるとき, A を **正規行列** (normal matrix) という.
2. $A^H A = A A^H = I$ (つまり $A^H = A^{-1}$) であるとき, A を **ユニタリ行列** (unitary matrix) という.

正規行列については第 2.3 節で詳述するとして, ここでは次の命題を示す.

命題 2.2.16 (ユニタリ行列の特徴づけ) $U = (u_1 \cdots u_n)$ を n 次複素正方行列とする. このとき, U に関する以下の条件は同値である.

1. U はユニタリ行列である.
2. 集合 $\{u_1, \dots, u_n\}$ は \mathbb{C}^n の正規直交基底である.
3. 任意の $x, y \in \mathbb{C}^n$ に対して $\langle Ux, Uy \rangle = \langle x, y \rangle$ である.

証明 まず, 1 と 2 の同値性を示す.

$$U^H U = \begin{pmatrix} u_1^H \\ \vdots \\ u_n^H \end{pmatrix} (u_1 \cdots u_n) = \begin{pmatrix} u_1^H u_1 & \cdots & u_1^H u_n \\ \vdots & \ddots & \vdots \\ u_n^H u_1 & \cdots & u_n^H u_n \end{pmatrix}$$

なので, $\langle u_i, u_j \rangle = u_j^H u_i = \delta_{ij}$ がすべての $i, j \in \{1, \dots, n\}$ で成り立つことは, $U^H U = I$ と同値である.

次に, 3 と 1 の同値性を示す. $\langle Ux, Uy \rangle = (Uy)^H Ux = (U^H Uy)^H x = \langle x, U^H Uy \rangle$ なので $\langle x, y \rangle - \langle Ux, Uy \rangle = \langle x, Ey \rangle$ ($E = I - U^H U$) である. 任意の $x \in \mathbb{C}^n$ に対し $\langle x, Ey \rangle = 0$ なら, $x = Ey$ とすれば $\|Ey\|^2 = 0$, すなわち $Ey = 0$ が得られる. $Ey = 0$ がすべての $y \in \mathbb{C}^n$ で成り立つとき, $E = O$ だから U はユニタリ行列である. 逆に U がユニタリ行列なら, $E = O$ より $\langle x, y \rangle - \langle Ux, Uy \rangle = \langle x, Ey \rangle = 0$ である. \square

ノート 対象がある条件を満たし, しかも条件を満たす対象が他にないとき, その条件は対象を**特徴づける** (characterize) という. 命題 2.2.16 から, ユニタリ行列の全体集合は複数の方法で特徴づけられる. しかし, どの方法を採用しても指し示す集合は変わらない.

系 2.2.17 $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ を線型写像とする。このとき、 T に関する以下の条件は同値である。

1. T は \mathbb{C}^n のある正規直交基底に関する分解作用素である。
2. 標準基底に関する T の表現行列はユニタリ行列である。

証明 T を正規直交基底 $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ に関する分解作用素とすると、[2.2.2 小節](#)冒頭の議論から $T(\mathbf{x}) = \mathbf{U}^H \mathbf{x}$ ($\mathbf{U} = (\mathbf{u}_1 \ \cdots \ \mathbf{u}_n)$) である。また、 \mathcal{B} は正規直交基底なので、[命題 2.2.16](#) より \mathbf{U} は——よって \mathbf{U}^H も——ユニタリ行列である。逆の証明は省略する。 \square

2.3 最小 2 乗法

本節では、直交射影の理論を近似へと応用する。

2.3.1 最小 2 乗法

いまから考えるのは、変数 $\mathbf{x} = (x_1 \ \cdots \ x_p)^T$ に関する連立 1 次方程式

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad (2.3)$$

の解き方である。 $\mathbf{A} = (a_{ij})$, $\mathbf{b} = (b_1 \ \cdots \ b_n)^T$ とおくと、式(2.3)は $\mathbf{Ax} = \mathbf{b}$ と書ける。もし $n = p$ で \mathbf{A} が正則なら、解は $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ ただ一つである。

\mathbf{A} が正則でないとき、式(2.3)の厳密解があるかどうかは \mathbf{b} 次第である。**最小 2 乗法** (least squares method) は、厳密解の代わりに**残差** (residual) $\boldsymbol{\epsilon}(\mathbf{x}) = \mathbf{b} - \mathbf{Ax}$ のノルム $\|\boldsymbol{\epsilon}(\mathbf{x})\|$ を最小にする近似解 \mathbf{x} を見つける手法である。次の命題から、所望の近似解 \mathbf{x} は常に存在する。

命題 2.3.1 $\mathbf{b} \in \mathbb{K}^n$ かつ、 \mathbf{A} は \mathbb{K} 上の $n \times p$ 行列とする。このとき、集合 $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\boldsymbol{\epsilon}(\mathbf{x})\|$ ($\boldsymbol{\epsilon}(\mathbf{x}) = \mathbf{b} - \mathbf{Ax}$) は空でない。

証明 線型写像 $T_A: \mathbb{K}^p \rightarrow \mathbb{K}^n$ を $T_A(\mathbf{x}) = \mathbf{Ax}$ で定義する。関数 $\|\boldsymbol{\epsilon}(\mathbf{x})\|$ の最小値は $\min\{\|\mathbf{b} - \mathbf{Ax}\| \mid \mathbf{x} \in \mathbb{K}^p\} = \min\{\|\mathbf{b} - \mathbf{z}\| \mid \mathbf{z} \in \text{im } T_A\}$ であり、

$\|\mathbf{b} - \mathbf{z}\|$ の値を最小にする $\mathbf{z} \in \text{im } T_A$ は $\text{proj}_{\text{im } T_A} \mathbf{b}$ ただ一つである. よって, $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\epsilon(\mathbf{x})\|$ は方程式 $\mathbf{Ax} = \text{proj}_{\text{im } T_A} \mathbf{b}$ の解集合である. この方程式は明らかに解をもつから, $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\epsilon(\mathbf{x})\|$ は空でない. \square

上の証明における $\text{im } T_A$, つまり部分空間 $\{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{K}^p\}$ を \mathbf{A} の列空間 (column space) といい, $\text{col } \mathbf{A}$ と表す. $\mathbf{A} = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_p)$ とおくと

$$\mathbf{Ax} = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_p) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \sum_{j=1}^p x_j \mathbf{a}_j \quad (\mathbf{x} = (x_1 \ \cdots \ x_p)^T)$$

だから, $\text{col } \mathbf{A}$ は \mathbf{A} の列ベクトル全体が生成する部分空間でもある.

例 2.3.2 $\mathbf{A} = \begin{pmatrix} 4 & 1 & 2 \\ 0 & 3 & 2 \end{pmatrix}^T$, $\mathbf{b} = (-4 \ -6 \ 4)^T$ とする. $\mathbf{a}_1 = (4 \ 1 \ 2)^T$, $\mathbf{a}_2 = (0 \ 3 \ 2)^T$ とおくと $\text{col } \mathbf{A} = \text{span}\{\mathbf{a}_1, \mathbf{a}_2\}$ である.

$P = \text{proj}_{\text{col } \mathbf{A}}$ とおく. $\{\mathbf{a}_1, \mathbf{a}_2\}$ にグラム・シュミットの直交化法を適用すると $\mathbf{e}_1 = (4 \ 1 \ 2)^T / \sqrt{21}$, $\mathbf{e}_2 = (-1 \ 2 \ 1)^T / \sqrt{6}$ が得られるので, $P(\mathbf{b}) = \langle \mathbf{b}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{b}, \mathbf{e}_2 \rangle \mathbf{e}_2 = -2(1 \ 1 \ 1)^T$ である. よって, $\arg \min_{\mathbf{x} \in \mathbb{R}^2} \|\mathbf{b} - \mathbf{Ax}\|$ は方程式 $\mathbf{Ax} = -2(1 \ 1 \ 1)^T$ の解集合 $\{(-1/2 \ -1/2)^T\}$ である. \diamond

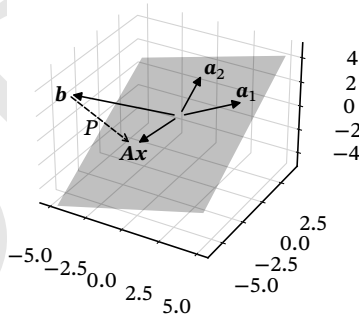


図 2.3 \mathbf{a}_1 , \mathbf{a}_2 と $P(\mathbf{b}) = \mathbf{Ax}$ の様子.

例 2.3.2 から分かるよう, $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\epsilon(\mathbf{x})\|$ を求めるには $\tilde{\mathbf{b}} = \text{proj}_{\text{col } \mathbf{A}} \mathbf{b}$ を計算して, 方程式 $\mathbf{Ax} = \tilde{\mathbf{b}}$ を解けばよい. 一方で, $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\epsilon(\mathbf{x})\|$ をより機械的に求める方法もある.

命題 2.3.3 $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\epsilon(\mathbf{x})\| = \{\mathbf{x} \in \mathbb{K}^p \mid \mathbf{A}^H \mathbf{Ax} = \mathbf{A}^H \mathbf{b}\}$ である.

証明 関数 $\epsilon(\mathbf{x}) = \|\epsilon(\mathbf{x})\|^2$ が点 \mathbf{a} で最小値をとることと, $\mathbf{g} = \mathbf{A}^H \epsilon(\mathbf{a}) = \mathbf{A}^H \mathbf{b} - \mathbf{A}^H \mathbf{A} \mathbf{a}$ が $\mathbf{0}$ であることが同値なのを示す. \mathbf{x} を \mathbf{a} から $\Delta \mathbf{x}$ だけ動かしたとき, $\epsilon(\mathbf{x})$ の増分 $\Delta \epsilon(\mathbf{a}) = \epsilon(\mathbf{a} + \Delta \mathbf{x}) - \epsilon(\mathbf{a})$ は

$$\begin{aligned} \Delta \epsilon(\mathbf{a}) &= \|\epsilon(\mathbf{a}) - \mathbf{A} \Delta \mathbf{x}\|^2 - \|\epsilon(\mathbf{a})\|^2 = -2 \operatorname{Re} \langle \epsilon(\mathbf{a}), \mathbf{A} \Delta \mathbf{x} \rangle + \|\mathbf{A} \Delta \mathbf{x}\|^2 \\ &= -2 \operatorname{Re} \langle \mathbf{A}^H \epsilon(\mathbf{a}), \Delta \mathbf{x} \rangle + \|\mathbf{A} \Delta \mathbf{x}\|^2 = -2 \operatorname{Re} \langle \mathbf{g}, \Delta \mathbf{x} \rangle + \|\mathbf{A} \Delta \mathbf{x}\|^2 \end{aligned}$$

である. よって, $\mathbf{g} = \mathbf{0}$ ならば $\Delta \mathbf{x}$ の値によらず $\Delta \epsilon(\mathbf{a}) = \|\mathbf{A} \Delta \mathbf{x}\|^2 \geq 0$ だから, ϵ は \mathbf{a} で最小値をとる. また, $\Delta \mathbf{x} = h \mathbf{g}$ ($h > 0$) のとき $\Delta \epsilon(\mathbf{a}) = -2h \|\mathbf{g}\|^2 + h^2 \|\mathbf{A} \mathbf{g}\|^2$ なので, $\|\mathbf{g}\| \neq 0$ なら h の値が十分小さいとき $\Delta \epsilon(\mathbf{a}) < 0$ である. つまり, $\mathbf{g} \neq \mathbf{0}$ ならば $\epsilon(\mathbf{a})$ は ϵ の最小値ではない. \square

\mathbb{K}^p の元 \mathbf{x} に関する方程式 $\mathbf{A}^H \mathbf{A} \mathbf{x} = \mathbf{A}^H \mathbf{b}$ を, 方程式 $\mathbf{A} \mathbf{x} = \mathbf{b}$ に関する **正規方程式** (normal equation) という. 命題 2.3.1, 2.3.3 より, もとの方程式が解をもつかどうかによらず, 正規方程式の解は存在する.

正規方程式は命題 2.2.6 から導ける. \mathbb{K}^p の元 \mathbf{x} が $\mathbf{A} \mathbf{x} = \operatorname{proj}_{\operatorname{col} \mathbf{A}} \mathbf{b}$ を満たすとき, $\mathbf{a}_j \in \operatorname{col} \mathbf{A}$ なので $\langle \mathbf{b} - \operatorname{proj}_{\operatorname{col} \mathbf{A}} \mathbf{b}, \mathbf{a}_j \rangle = \mathbf{a}_j^H (\mathbf{b} - \mathbf{A} \mathbf{x}) = 0$ である. よって $\mathbf{A}^H (\mathbf{b} - \mathbf{A} \mathbf{x}) = \mathbf{0}$ である.

2.3.2 スペクトル分解と特異値分解

$\|\epsilon(\mathbf{x})\|$ についてさらに詳しく調べるには, 以下の定理が必要である.

定理 2.3.4 (スペクトル定理) \mathbf{A} を n 次複素正方行列とする. このとき, \mathbf{A} に関する以下の条件は同値である. これを **スペクトル定理** (spectral theorem) という.

1. \mathbf{A} は正規行列である.
2. \mathbf{A} はユニタリ行列で対角化できる. すなわち, $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ を満たす n 次ユニタリ行列 \mathbf{U} , 対角行列 $\mathbf{\Lambda}$ が存在する.

証明 第 2.7.1 小節を参照せよ (ひとまず認めてもかまわない). \square

ユニタリ行列による対角化は **スペクトル分解** (spectral decomposition) と呼ばれる. 定理 2.3.4 から, 正規行列であることとスペクトル分解できることは同値である.

ノート 正規行列のスペクトル分解を計算するには、固有空間一つ一つの正規直交基底を求め、それらを並べて1つのユニタリ行列を作ればよい。

系 2.3.5 (特異値分解) A を任意の $n \times p$ 複素行列とする. このとき, 以下の条件すべてを満たす行列 U, V, Σ が存在する. 式 $A = U\Sigma V^H$ を A の**特異値分解** (singular value decomposition; SVD) という.

1. $A = U\Sigma V^H$ である.
2. U は n 次ユニタリ行列, V は p 次ユニタリ行列である.
3. Σ は $n \times p$ 行列であり $\Sigma = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_k) & 0 \end{pmatrix}$ と書ける. ただし $k = \min\{n, p\}$, $\sigma_1 \geq \dots \geq \sigma_k \geq 0$ である.

証明 $A \neq O$, $n \geq p$ のときのみ示す. $A^H A$ は p 次正規行列なので, 定理 2.3.4 よりスペクトル分解 $A^H A = V\Lambda V^H$ が存在する. $V = (v_1 \ \dots \ v_p)$, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ とおく. $\delta_{ij} = \langle v_i, v_j \rangle = v_j^H v_i$, $\lambda_i v_i = (A^H A)v_i$ より

$$\lambda_i \delta_{ij} = v_j^H (\lambda_i v_i) = v_j^H (A^H A)v_i = (Av_j)^H Av_i = \langle Av_i, Av_j \rangle \quad (2.4)$$

だから, 各 $\lambda_i = \|Av_i\|^2$ はすべて非負である. また

$$A^H A = (v_1 \ \dots \ v_p) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \begin{pmatrix} v_1^H \\ \vdots \\ v_p^H \end{pmatrix} = \sum_{i=1}^p \lambda_i v_i v_i^H$$

であり, 和の順序を変えても値は変わらない. そこで, $\lambda_1 \geq \dots \geq \lambda_p$ が成り立つように $\lambda_1, \dots, \lambda_p$ と v_1, \dots, v_p を並べ替えておく.

$\sigma_i = \sqrt{\lambda_i} = \|Av_i\|$, $r = \max\{i \mid Av_i \neq 0\}$ とおく. 式 (2.4) より, $u_i = (1/\sigma_i)Av_i$ とおくと集合 $\{u_1, \dots, u_r\}$ は \mathbb{C}^n の正規直交系である. $r < n$ のときは $u_{r+1}, \dots, u_n \in \mathbb{C}^n$ を適切に補うことで, 集合 $\{u_1, \dots, u_n\}$ を \mathbb{C}^n の正規直交基底にできる. すると, $U = (u_1 \ \dots \ u_n)$ はユニタリ行列であり

$$U^H(AV) = U^H(\sigma_1 u_1 \ \dots \ \sigma_r u_r \ 0) = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_r) & 0 \end{pmatrix}$$

が成り立つ. $\sigma_{r+1} = \dots = \sigma_k = 0$ なので, 右辺の行列は $\begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_k) & 0 \end{pmatrix}$ と表せる. この行列を Σ とおくと $A = U\Sigma V^H$ である. \square

A の特異値分解が $A = U \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_k) & 0 \end{pmatrix} V^H$ であるとき, 各 σ_i を A の第 i 特異値 (singular value) という.

特異値分解に基づいて $\arg \min_{\mathbf{x} \in \mathbb{C}^p} \|\boldsymbol{\epsilon}(\mathbf{x})\|$ を求めよう. $\mathbf{A} \neq \mathbf{O}$ を仮定し, \mathbf{A} の特異値分解を $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ とおく. 命題 2.2.16 から, ユニタリ行列を掛けてもノルムは変わらないので $\|\boldsymbol{\epsilon}(\mathbf{x})\|^2 = \|\mathbf{b} - \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H\mathbf{x}\|^2 = \|\mathbf{U}^H\mathbf{b} - \boldsymbol{\Sigma}\mathbf{V}^H\mathbf{x}\|^2$ である. $\mathbf{y} = (y_1 \ \cdots \ y_p)^T = \mathbf{V}^H\mathbf{x}$, $\mathbf{U} = (\mathbf{u}_1 \ \cdots \ \mathbf{u}_n)$, $\boldsymbol{\Sigma} = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_r) & \\ & \mathbf{0} \end{pmatrix}$ ($\sigma_r > 0$) とおくと

$$\|\boldsymbol{\epsilon}(\mathbf{x})\|^2 = \|\mathbf{U}^H\mathbf{b} - \boldsymbol{\Sigma}\mathbf{y}\|^2 = \sum_{i=1}^r |\mathbf{u}_i^H\mathbf{b} - \sigma_i y_i|^2 + \sum_{i=r+1}^p |\mathbf{u}_i^H\mathbf{b}|^2$$

となる. 右辺の値が最も小さくなるのは $y_i = \mathbf{u}_i^H\mathbf{b}/\sigma_i$ ($1 \leq i \leq r$) のときであり, y_{r+1}, \dots, y_p の値はなんであっても構わない. $\mathbf{x} = \mathbf{V}\mathbf{y}$ だから, $\|\boldsymbol{\epsilon}(\mathbf{x})\|$ が最小値をとるのは \mathbf{x} が

$$\mathbf{x} = \mathbf{V}(\mathbf{u}_1^H\mathbf{b}/\sigma_1 \ \cdots \ \mathbf{u}_r^H\mathbf{b}/\sigma_r \ z_{r+1} \ \cdots \ z_p)^T \quad (z_i \in \mathbb{C})$$

と表せるときである. $\mathbf{z} = (z_1 \ \cdots \ z_p)^T$, $\boldsymbol{\Sigma}^+ = \begin{pmatrix} \text{diag}(1/\sigma_1, \dots, 1/\sigma_r) & \\ & \mathbf{0} \end{pmatrix}$, $\mathbf{A}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H$ において整理すると, $\boldsymbol{\Sigma}^+\boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{I}_r & \\ & \mathbf{0}_{p-r} \end{pmatrix}$ なので

$$\begin{aligned} \mathbf{x} &= \mathbf{V} \begin{pmatrix} (1/\sigma_1)\mathbf{u}_1^H\mathbf{b} \\ \vdots \\ (1/\sigma_r)\mathbf{u}_r^H\mathbf{b} \\ \mathbf{0}_{p-r} \end{pmatrix} + \mathbf{V} \begin{pmatrix} \mathbf{0}_r \\ z_{r+1} \\ \vdots \\ z_p \end{pmatrix} = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H\mathbf{b} + \mathbf{V}(\mathbf{I} - \boldsymbol{\Sigma}^+\boldsymbol{\Sigma})\mathbf{z} \\ &= \mathbf{A}^+\mathbf{b} + (\mathbf{I} - \mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{V}^H)\mathbf{V}\mathbf{z} = \mathbf{A}^+\mathbf{b} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{w} \end{aligned}$$

である (ただし $\mathbf{w} = \mathbf{V}\mathbf{z}$). \mathbf{A}^+ を \mathbf{A} の擬似逆行列という.

定義 2.3.6 (擬似逆行列) \mathbf{A} を $n \times p$ 複素行列とし, \mathbf{A} の特異値分解を $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ とおく. $p \times n$ 行列 $\boldsymbol{\Sigma}^+$ を次のように定義する. $\mathbf{A} = \mathbf{O}$ のときは $\boldsymbol{\Sigma}^+ = \mathbf{O}$ とする. $\mathbf{A} \neq \mathbf{O}$ のときは, $\boldsymbol{\Sigma} = \begin{pmatrix} \Delta & \\ & \mathbf{0} \end{pmatrix}$ を満たす正則な対角行列 Δ を用いて $\boldsymbol{\Sigma}^+ = \begin{pmatrix} \Delta^{-1} & \\ & \mathbf{0} \end{pmatrix}$ とする. $p \times n$ 行列 $\mathbf{A}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H$ を \mathbf{A} の **ムーア・ペンローズ逆行列** (Moore–Penrose inverse), あるいは**擬似逆行列** (pseudoinverse) という.

ノート 普通, 特異値分解はもとの行列に対して一意ではない. しかし, 擬似逆行列はもとの行列に対して一意に定まる. このことは第 2.3.3 小節で示す.

さきほどの議論から分かることは, 擬似逆行列によって次のようにまとめられる.

命題 2.3.7 任意の $\mathbf{b} \in \mathbb{C}^n$, $n \times p$ 複素行列 \mathbf{A} に対して、次の式が成り立つ.

$$\arg \min_{\mathbf{x} \in \mathbb{C}^p} \|\mathbf{b} - \mathbf{A}\mathbf{x}\| = \{\mathbf{A}^+\mathbf{b} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{w} \mid \mathbf{w} \in \mathbb{C}^p\}$$

2.3.3 擬似逆行列と解の構造

これまでと同様、 \mathbf{A} は任意の $n \times p$ 複素行列、 \mathbf{b} は \mathbb{C}^n の任意の元とする. ここでは擬似逆行列の性質と $\arg \min_{\mathbf{x} \in \mathbb{C}^p} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ の構造について説明する.

命題 2.3.8 (擬似逆行列の特徴づけ) 以下の条件すべてを満たす $p \times n$ 行列 \mathbf{X} は、 \mathbf{A} に対し一意である. また、 \mathbf{A} の擬似逆行列は条件を満たす. つまり、以下の条件は $\mathbf{X} = \mathbf{A}^+$ と同値である.

1. $\mathbf{AXA} = \mathbf{A}$
2. $\mathbf{XAX} = \mathbf{X}$
3. $(\mathbf{AX})^H = \mathbf{AX}$
4. $(\mathbf{XA})^H = \mathbf{XA}$

証明 一意性のみ示す. $p \times n$ 行列 \mathbf{X}, \mathbf{Y} が条件をすべて満たすとき、 i 番目の条件から分かる等号を $\stackrel{i}{=}$ と書くと

$$\begin{aligned} \mathbf{Y} &\stackrel{2,3}{=} \mathbf{Y}(\mathbf{AY})^H = \mathbf{Y}\mathbf{Y}^H\mathbf{A}^H \stackrel{1}{=} \mathbf{Y}\mathbf{Y}^H(\mathbf{AXA})^H = \mathbf{Y}(\mathbf{AY})^H(\mathbf{AX})^H \\ &\stackrel{2,3}{=} \mathbf{YAX} \stackrel{2,4}{=} (\mathbf{YA})^H\mathbf{XAX} \stackrel{4}{=} \mathbf{A}^H\mathbf{Y}^H(\mathbf{XA})^H\mathbf{X} = (\mathbf{AYA})^H\mathbf{X}^H\mathbf{X} \\ &\stackrel{1}{=} \mathbf{A}^H\mathbf{X}^H\mathbf{X} = (\mathbf{XA})^H\mathbf{X} \stackrel{2,4}{=} \mathbf{X} \end{aligned}$$

なので $\mathbf{Y} = \mathbf{X}$ である. □

命題 2.3.9 集合 $S = \arg \min_{\mathbf{x} \in \mathbb{C}^p} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ の元でノルムが最も小さいものは、 $\mathbf{A}^+\mathbf{b}$ ただ一つである. すなわち $\arg \min_{\mathbf{x} \in S} \|\mathbf{x}\| = \{\mathbf{A}^+\mathbf{b}\}$ である.

証明 命題 2.3.7 より S の任意の元は $\mathbf{x} = \mathbf{A}^+\mathbf{b} + \mathbf{n}$ ($\mathbf{n} = (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{w}$) と書けて、命題 2.3.8 から $\langle \mathbf{A}^+\mathbf{b}, \mathbf{n} \rangle = \langle (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{A}^+\mathbf{b}, \mathbf{w} \rangle = 0$ である. よって $\|\mathbf{x}\|^2 = \|\mathbf{A}^+\mathbf{b}\|^2 + \|\mathbf{n}\|^2$ なので、 $\arg \min_{\mathbf{x} \in S} \|\mathbf{x}\| = \{\mathbf{A}^+\mathbf{b}\}$ である. □

実は, 命題 2.3.9 において $\arg \min_{\mathbf{x} \in S} \|\mathbf{x}\|$ の元が唯一であることは擬似逆行列を使わず証明できる. 第 2.3 節の初めに述べた通り, $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ は方程式 $\mathbf{A}\mathbf{x} = \tilde{\mathbf{b}}$ ($\tilde{\mathbf{b}} = \text{proj}_{\text{col } \mathbf{A}} \mathbf{b}$) の解集合である. この方程式の解を 1 つ選び \mathbf{a} とおくと, 他の解 \mathbf{x} は $\mathbf{A}(\mathbf{x} - \mathbf{a}) = \tilde{\mathbf{b}} - \tilde{\mathbf{b}} = \mathbf{0}$, つまり $\mathbf{x} - \mathbf{a} \in \text{nul } \mathbf{A}$ ($\text{nul } \mathbf{A} = \{\mathbf{n} \in \mathbb{C}^p \mid \mathbf{A}\mathbf{n} = \mathbf{0}\}$) を満たす.

逆に, \mathbb{K}^p の元 \mathbf{x} が $\mathbf{x} - \mathbf{a} \in \text{nul } \mathbf{A}$ を満たせば $\mathbf{A}\mathbf{x} = \tilde{\mathbf{b}}$ も成立する. つまり $\arg \min_{\mathbf{x} \in \mathbb{K}^p} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ は, 部分空間 $\text{nul } \mathbf{A}$ を \mathbf{a} によって平行移動した集合 $\{\mathbf{a} + \mathbf{n} \mid \mathbf{n} \in \text{nul } \mathbf{A}\}$ である. このような集合をアフィン部分空間という.

定義 2.3.10 (アフィン部分空間) S をベクトル空間 V の部分集合とする. ある $\mathbf{a} \in S$ に対して集合 $W = \{\mathbf{x} - \mathbf{a} \mid \mathbf{x} \in S\}$ が V の部分ベクトル空間であるとき, S は V の **アフィン部分空間** (affine subspace) であるという. このとき S を $S = \mathbf{a} + W$ と表す.



図 2.4 アフィン部分空間 $\mathbf{a} + W$ の模式図.

任意の $\mathbf{x} \in \mathbb{K}^p$ とアフィン部分空間 $S \subseteq \mathbb{K}^p$ に対し, $\arg \min_{\mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|$ は S が部分ベクトル空間のときと同様, ただ一つの元を持つ. 実際, $S = \mathbf{a} + W$ ($\mathbf{a} \in S$) とおくと $\arg \min_{\mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\| = \arg \min_{\mathbf{y} - \mathbf{a} \in W} \|(\mathbf{x} - \mathbf{a}) - (\mathbf{y} - \mathbf{a})\| = \{\mathbf{a} + \text{proj}_W(\mathbf{x} - \mathbf{a})\}$ である.

よって, 集合 $\arg \min_{\mathbf{x} \in S} \|\mathbf{x}\|$ ($S = \mathbf{a} + \text{nul } \mathbf{A}$) も元をただ一つ持つ. そして命題 2.3.9 から, その元とは $\mathbf{A}^+ \mathbf{b}$ のことである.

2.3.4 $\mathbb{K} = \mathbb{R}$ の場合

第 2.3.2 小節から第 2.3.3 小節の結果は、スペクトル分解と特異値分解を次の形で適当に読み替えれば $\mathbb{K} = \mathbb{R}$ のときも適用できる。証明は $\mathbb{K} = \mathbb{C}$ のときとほとんど変わらない。

定理 2.3.11 (スペクトル定理) A を n 次実正方行列とする。このとき、 A に関する以下の条件は同値である。

1. A は正規行列である。
2. A は直交行列²⁾で対角化できる。すなわち、 $A = U\Lambda U^T$ を満たす n 次直交行列 U 、対角行列 Λ が存在する。

系 2.3.12 (特異値分解) A を任意の $n \times p$ 実行列とする。このとき、以下の条件すべてを満たす行列 U , V , Σ が存在する。

1. $A = U\Sigma V^T$ である。
2. U は n 次直交行列、 V は p 次直交行列である。
3. Σ は $n \times p$ 行列であり $\Sigma = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_k) & 0 \\ 0 & 0 \end{pmatrix}$ と書ける。ただし $k = \min\{n, p\}$, $\sigma_1 \geq \dots \geq \sigma_k \geq 0$ である。

2.3.5 曲線あてはめ

ここまで説明した最小 2 乗法の理論は、データを曲線にあてはめるとき威力を発揮する。

変数 x と変数 y の間には、多項式で表される関係 $y = c_0 + c_1x + \dots + c_px^p$ があると仮定する。次数 p は既知であり、各 x^k の係数 c_k を y の観測値から推定したい。

たとえば、おもりを自由落下させると落下時間 t と移動距離 y の間には $y = (g/2)t^2$ という関係がある (g は重力加速度と呼ばれる定数である)。そこで、 t の値を変えながら y の値をくり返し測定し、得られたデータを式 $y = at^2$ にあてはめれば、 $g = 2a$ の値を間接的に測定できる。

2) $QQ^T = Q^TQ = I$ を満たす実正方行列 Q を直交行列 (orthogonal matrix) という。

以下では、 $x = x_i$ に対応する観測値 y_i の系列 y_1, \dots, y_n を y の標本と呼ぶ。次の定理から、 c_0, \dots, c_p を決定するには最低限 $p + 1$ 個の標本があればよい。

定理 2.3.13 点列 $(x_1, y_1), \dots, (x_{p+1}, y_{p+1}) \in \mathbb{K}^2$ は、相異なる任意の $i, j \in \{1, \dots, p+1\}$ に対して $x_i \neq x_j$ を満たすとする。このとき、すべての i で $f(x_i) = y_i$ を満たす、次数が p 以下の \mathbb{K} 係数多項式 $f(x)$ がただ一つ存在する³⁾。

証明 まず $f(x)$ が存在することを示す。

$$l_k(x) = (x - x_1)(x - x_2) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_{p+1})$$

とおくと $l_k(x_k) \neq 0$, $l_k(x_i) = 0$ ($i \neq k$) より $l_k(x_i)/l_k(x_k) = \delta_{ki}$ である。よって、 $f(x) = \sum_{k=1}^{p+1} y_k l_k(x)/l_k(x_k)$ は条件を満たす \mathbb{K} 係数多項式である。

次に $f(x)$ が一意であることを示す。多項式 $f_1(x)$ と $f_2(x)$ はどちらも条件を満たすとする。このとき、多項式 $\delta(x) = f_1(x) - f_2(x)$ は $\delta(x_i) = y_i - y_i = 0$ を満たすので、 $\delta(x) = q(x)(x - x_1) \cdots (x - x_{p+1})$ と因数分解できる。 $\delta(x)$ の次数は最大でも p だから $q(x) = 0$ 、すなわち $f_1(x) = f_2(x)$ である。□

実験で得られるデータは誤差を含むものだから、普通は標本数 n を $p + 1$ より十分に大きくとり、多項式関数 $f(\mathbf{c}; x) = \sum_{k=0}^p c_k x^k$ のあてはまりが平均的によくなる係数 $\mathbf{c} = (c_0 \ \cdots \ c_p)^T$ を求める。図は $p = 2$ における標本と $f(\mathbf{c}; x) = c_0 + c_1 x + c_2 x^2$ の例である。

あてはまりのよさを評価する指標はいろいろ考えられるが、ここでは関数



$$\epsilon(\mathbf{c}) = \sum_{i=1}^n |y_i - f(\mathbf{c}; x_i)|^2 \quad (2.5)$$

を採用する。各 $f(\mathbf{c}; x_i)$ の値が y_i の値に近いほど $\epsilon(\mathbf{c})$ の値は小さくなるので、 $\epsilon(\mathbf{c})$ の値が小さいほど $f(\mathbf{c}; x)$ は標本によくあてはまっているといえる。

3) この多項式を**補間多項式** (interpolating polynomial) という。

ノート 式 (2.5) の指標を**残差平方和** (residual sum of squares; RSS) という。RSS であてはまりのよさを測るのが妥当かどうかは、データの性質や分析する目的にもよる。しかし、ほかの指標に関して $\arg \min$ を計算するのは——空集合でないことを保証するだけでも——難しいことが多い。まずは RSS が使えないか考えるべきだろう。

$\phi_i = (x_i^0 \ \cdots \ x_i^p)^T$, $\Phi = (\phi_1 \ \cdots \ \phi_n)^T$ とおくと, $f(\mathbf{c}; x_i) = \phi_i^T \mathbf{c}$ だから

$$\epsilon(\mathbf{c}) = \left\| \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \phi_1^T \\ \vdots \\ \phi_n^T \end{pmatrix} \mathbf{c} \right\|^2 = \|\mathbf{y} - \Phi \mathbf{c}\|^2 \quad (\mathbf{y} = (y_1 \ \cdots \ y_n)^T)$$

である。よって, $f(\mathbf{c}; x)$ が標本に最もよくあてはまる \mathbf{c} を求めるには, さきほど示した方法で $\arg \min_{\mathbf{c} \in \mathbb{K}^{p+1}} \|\mathbf{y} - \Phi \mathbf{c}\|$ を計算すればよい。

例 2.3.14 $p = 1$, $n = 3$, $(x_1, y_1) = (1, 0)$, $(x_2, y_2) = (2, 3)$, $(x_3, y_3) = (3, 1)$ とする。 $\Phi = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}^T$ とおくと, 方程式 $\Phi \mathbf{c} = \mathbf{y}$ ($\mathbf{y} = (0 \ 3 \ 1)^T$) に関する正規方程式は $\Phi^T \Phi \mathbf{c} = \Phi^T \mathbf{y}$, $\begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \mathbf{c} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$ である。この解は $\mathbf{c} = \begin{pmatrix} 1/3 \\ 1/2 \end{pmatrix}$ なので, $f(\mathbf{c}; x) = 1/3 + (1/2)x$ のとき $f(\mathbf{c}; x)$ は標本に最もよくあてはまる。◇

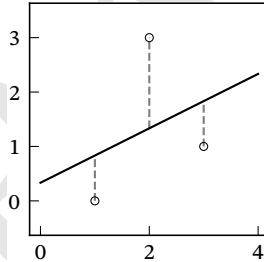


図 2.5 例 2.3.14 の標本と $f(\mathbf{c}; x)$ の様子。 $\epsilon(\mathbf{c})$ は破線の長さの平方和である。

同じ考え方は, y が x の多項式関数でなくても使える。変数 y は, 変数 x を既知の関数 $\phi(x) = (\phi_1(x) \ \cdots \ \phi_p(x))^T$ で写像した線型結合によって

$$y = c_1 \phi_1(x) + \cdots + c_p \phi_p(x) = \mathbf{c}^T \phi(x) \quad (2.6)$$

と書けるとする。関数 $\epsilon(\mathbf{c}) = \sum_{i=1}^n |y_i - f(\mathbf{c}; x_i)|^2$ ($f(\mathbf{c}; x) = \sum_{k=1}^p c_k \phi_k(x)$) が最小値をとる $\mathbf{c} \in \mathbb{K}^p$ 全体は, 集合 $\arg \min_{\mathbf{c} \in \mathbb{K}^p} \|\mathbf{y} - \Phi \mathbf{c}\|$ である。ただし, Φ は $n \times p$ 行列 $\Phi = (\phi(x_1) \ \cdots \ \phi(x_n))^T$ である。

なお, ϕ_1, \dots, ϕ_p が線型従属だと式 (2.6) を満たす c_1, \dots, c_p の値は一意に定まらない。そのため, しばしば ϕ_1, \dots, ϕ_p は線型独立になるよう設計される。

2.4 離散フーリエ変換

第 2.4 節から、本書の主題である信号解析に入る。その手始めに、いまから扱う信号とは何か明確にしておく。

ごくごく一般的に言うと、信号とは入力に対して出力が観測されるもの全般を指し、入力全体 I から出力全体 O への写像でモデル化できる。たとえば、対象が音声信号なら I は時刻全体、 O は瞬時値全体である。また、対象がモノクロ画像なら I は画素の位置全体、 O は白黒の階調全体である。

始域が X 、終域が Y である写像の全体集合は通常 Y^X と書かれる。 I と O を決めたとき、考えられる信号の最も広い枠組みは O^I であろう。しかし、 O^I の元すべてに共通する特徴は普通そう多くない。そこで我々は、 O^I の中でも性質がよい元に限って、それらが持つ特徴を調べる。

第 2 章で扱うのは、 I が整数かその組からなる信号である。こうした信号を**離散信号** (discrete signal) という。反対に、 I が区間 $[-\pi, \pi]$ や \mathbb{R} のように連続的である信号は**連続信号** (continuous signal) という。

時間 N が有限な離散信号の全体 $\mathbb{K}^{\{0, \dots, N-1\}}$ は、数ベクトル空間 \mathbb{K}^N と同一視できる。これは対応 $x = \{x_n\}_{n=0}^{N-1} \longleftrightarrow \mathbf{x} = (x_0 \ \cdots \ x_{N-1})^T$ を考えれば明らかであろう。また、数列全体 $\mathbb{K}^{\mathbb{Z}}$ ——これを**数列空間** (sequence space) という——も、次の演算でベクトル空間になる。

$$\{x_n\} + \{y_n\} = \{x_n + y_n\}, \quad \lambda\{x_n\} = \{\lambda x_n\} \quad (\{x_n\}, \{y_n\} \in \mathbb{K}^{\mathbb{Z}}, \lambda \in \mathbb{K})$$

以上を踏まえて、まずは \mathbb{C}^N に属する信号の理論を構成しよう。

2.4.1 離散フーリエ変換

定義 2.4.1 (離散フーリエ変換) $\mathbf{x} = (x_0 \ \cdots \ x_{N-1})^T \in \mathbb{C}^N$ とする。

$$\mathcal{F}_{\mathbb{Z}/N} \mathbf{x} = (\hat{x}_0 \ \cdots \ \hat{x}_{N-1})^T, \quad \hat{x}_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}$$

で定義される線型写像 $\mathcal{F}_{\mathbb{Z}/N}: \mathbb{C}^N \rightarrow \mathbb{C}^N$ を**離散フーリエ変換** (discrete Fourier transform; DFT) という。

以下では $e^{2\pi i/N} = \cos(2\pi/N) + i \sin(2\pi/N)$ を単に ζ と書く。

命題 2.4.2 $\mathbf{w}_k = N^{-1/2}(\zeta^{k \cdot 0} \dots \zeta^{k(N-1)})^\top$ とする。このとき、集合 $\{\mathbf{w}_0, \dots, \mathbf{w}_{N-1}\}$ は \mathbb{C}^N の正規直交基底である。

証明 $\bar{\zeta} = \zeta^{-1}$ だから、 $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \mathbf{w}_i^\top \bar{\mathbf{w}}_j$ は

$$\sum_{n=0}^{N-1} \frac{\zeta^{in}}{\sqrt{N}} \frac{\bar{\zeta}^{jn}}{\sqrt{N}} = \frac{1}{N} \sum_{n=0}^{N-1} \zeta^{(i-j)n} = \begin{cases} (\zeta^{(i-j)N} - 1)/(N(\zeta^{i-j} - 1)) & (i \neq j), \\ 1 & (i = j) \end{cases}$$

と変形できる。 $\zeta^N = 1$ なので $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \delta_{ij}$ である。 □

命題 2.4.2 から、 $\mathcal{F}_{\mathbb{Z}/N}$ は正規直交基底 $\mathcal{W} = \{\mathbf{w}_0, \dots, \mathbf{w}_{N-1}\}$ に関する分解作用素である。分解作用素の逆写像は合成作用素なので、 $\mathcal{F}_{\mathbb{Z}/N}$ の逆変換は

$$\mathbf{x} = \sum_{k=0}^{N-1} \hat{x}_k \mathbf{w}_k, \quad x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k n / N} \quad (2.7)$$

と書ける。

他の諸性質を導く前に、離散フーリエ変換の工学的重要性を見ておこう。

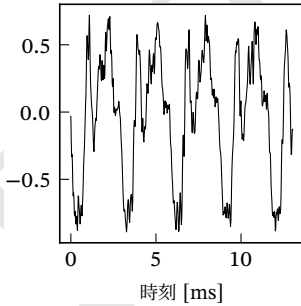


図 2.6 「あ」の波形。



図 2.7 e^{it} のグラフ。

図 2.6 は「あ」という音声の波形である⁴⁾。冒頭で述べたように、図 2.6 のデータは数ベクトル $\mathbf{x} \in \mathbb{C}^N$ と見なせる。そして、分解作用素に関する考察によれば、 $\hat{x}_k = \langle \mathbf{x}, \mathbf{w}_k \rangle$ は \mathbf{x} に含まれる \mathbf{w}_k の成分に相当する。つまり、 \hat{x}_k

4) カノン [2] による波音リツ単独音バージョン 1.5.1 を使用した。

の絶対値 $|\hat{x}_k|$ は音声 \mathbf{x} に含まれる \mathbf{w}_k の量を表すと考えられる。では、偏角 $\arg \hat{x}_k$ はどういう意味を持つのだろうか？ 極形式 $\hat{x}_k = |\hat{x}_k|e^{i\arg \hat{x}_k}$ を式 (2.7) に代入すると

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k n / N} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |\hat{x}_k| e^{i(2\pi k n / N + \arg \hat{x}_k)}$$

となるから、 $\arg \hat{x}_k$ は音声 \mathbf{x} に含まれる周波数 k/N の波 $\sqrt{N}\mathbf{w}_{kn} = e^{2\pi i k n / N}$ の初期位相を表している。この波は図 2.7 のような螺旋形を描く（矢印は時間軸）。

実際に $|\hat{x}_k|$ を計算すると、図 2.8 実線部のようになる。



図 2.8 「あ」の離散フーリエ変換。

図 2.8 を見ると、350 Hz 周辺に 1 つめのピークが現れている。これは**基本周波数** (fundamental frequency) と呼ばれる量で、人間が知覚する声の高さとかかなりよく対応する。

ノート 人間が知覚する声の高さを**ピッチ** (pitch) という。基本周波数とピッチはしばしば同一視されるが、厳密には定義からして異なる。詳しくは柏野 [6] など。

また、図 2.8 の破線部 \hat{h}_k は、基本周波数に起因する細かな変動を $|\hat{x}_k|$ から除いた曲線である。この曲線は**スペクトル包絡** (spectral envelope) といい、音声の音色とよく対応する。つまり、離散フーリエ変換を使うことで、音声を持つ基本周波数（≡声の高さ）由来の性質と、スペクトル包絡（≡音色）由来の性質を分離して解析できる。

2.4.2 エイリアシング

さきほど「 $\sqrt{N}w_{kn} = e^{2\pi i kn/N}$ は周波数 k/N の波である」と述べたが³、この表現には少し語弊がある。

周波数 f Hz の波 $e^{2\pi i f t}$ について、瞬時値を 1 秒あたり f_s 回記録すると数列 $\{e^{2\pi i f n/f_s}\}_{n \in \mathbb{Z}}$ ができる。一方、周波数 $(f - f_s)$ Hz の波 $e^{2\pi i (f - f_s)t}$ について同じ方法で数列を作ると、その一般項は $e^{2\pi i (f - f_s)n/f_s} = e^{2\pi i f n/f_s} (e^{-2\pi i})^n = e^{2\pi i f n/f_s}$ となる。つまり、周波数が f Hz でも $(f - f_s)$ Hz でも、できる数列は変わらない。

ノート 波が普通と逆向きに進むのを、周波数に負号をつけて表すことがある。特に、複素指数関数 $e^{-2\pi i f t}$ ($f > 0$) は周波数 $-f$ の波とみなされる。

言い換えると、周波数が f_s だけ異なる波は数列から区別できない。図 2.9 は $f_s = 2.5$ Hz のとき、周波数 2 Hz の波 $\sin(4\pi t)$ と -0.5 Hz の波 $\sin(-\pi t)$ が³、時刻 n/f_s 秒 ($n \in \mathbb{Z}$) では同じ瞬時値を持つことを示している。

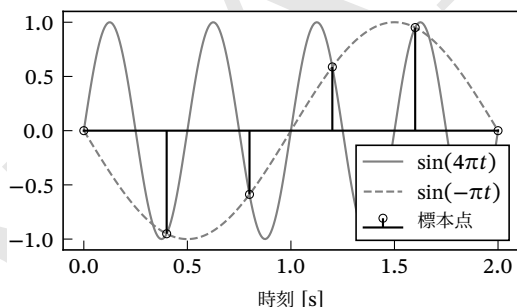


図 2.9 エイリアシングの様子。

一般に、連続信号から離散信号を得る操作を**標本化** (sampling) といい、標本化によって信号が区別できなくなる現象を**エイリアシング** (aliasing) という。また、 f_s を**標本化周波数** (sampling frequency) という。

図 2.9 の場合、標本化後の信号は周波数の絶対値が小さい (つまり低周波である) -0.5 Hz の波を表すと考えるのが自然だろう。周波数 f の波を標本化すると、 f より低周波の波と区別できなくなるのは $|f| \geq f_s/2$ のときである。 $f_s/2$ のことを**ナイキスト周波数** (Nyquist frequency) という。

話を離散フーリエ変換に戻すと、 $\sqrt{N}w_{kn} = e^{2\pi i kn/N}$ は $f = k$ の波が $f_s = N$ で標本化されたものとみなせる。そのため、 $k \geq N/2$ のとき $\sqrt{N}w_{kn}$ は周波数 k/N の波ではなく、周波数 $(k - N)/N$ の波を表すとみなすのが普通である。

2.4.3 曲線あてはめと DFT

離散フーリエ変換は曲線あてはめの観点からも理解できる。周期 2π の関数 f を、複素指数関数の線型結合

$$f(\mathbf{c}; t) = c_0\phi_0(t) + \cdots + c_{K-1}\phi_{K-1}(t) \quad (\phi_k(t) = e^{ikt})$$

で近似しよう。標本化周波数 N で f を標本化すると、離散信号 $x[n] = f(t_n)$ ($t_n = 2\pi n/N$) が得られる⁵⁾。

N 個の標本 $f(t_n) = x[n]$ ($0 \leq n < N$) を用いて f を $f(\mathbf{c}; t)$ にあてはめる場合、 K を N より大きくとると、 $\phi_0, \dots, \phi_{K-1}$ の中に標本化後は区別できなくなる組が現れてしまう（たとえば ϕ_0 と ϕ_N ）。そこで $K = N$ とし、残差平方和

$$\epsilon(\mathbf{c}) = \sum_{n=0}^{N-1} |x[n] - f(\mathbf{c}; t_n)|^2 \quad (t_n = 2\pi n/N)$$

の値を最も小さくする $\mathbf{c} \in \mathbb{C}^N$ を求める。それには正規方程式 $\Phi^H \Phi \mathbf{c} = \Phi^H \mathbf{x}$ を解けばよい。ただし $\mathbf{x} = (x[0] \ \cdots \ x[N-1])^T$ かつ

$$\Phi = (\phi_{i-1}(t_{j-1}))^T = \begin{pmatrix} e^{i \cdot 0(2\pi \cdot 0/N)} & \cdots & e^{i(N-1)(2\pi \cdot 0/N)} \\ \vdots & \ddots & \vdots \\ e^{i \cdot 0(2\pi(N-1)/N)} & \cdots & e^{i(N-1)(2\pi(N-1)/N)} \end{pmatrix}$$

である。ところが、 Φ は命題 2.4.2 の正規直交基底を用いて $\sqrt{N}\mathbf{W}$ ($\mathbf{W} = (\mathbf{w}_0 \ \cdots \ \mathbf{w}_{N-1})$) と書けるので $\Phi^H \Phi = \mathbf{N}\mathbf{W}^H \mathbf{W} = \mathbf{N}\mathbf{I}$ である。つまり、正規方程式は初めから解かれているも同然であり、解は $\mathbf{c} = (1/N)\Phi^H \mathbf{x} = (1/\sqrt{N})\mathbf{W}^H \mathbf{x} = (1/\sqrt{N})\mathcal{F}_{\mathbb{Z}/N} \mathbf{x}$ である。

まとめると、区間 $[0, 2\pi]$ を N 等分した左端における f の値を標本とし、関数 $f(\mathbf{c}; t) = \sum_{k=0}^{N-1} c_k e^{ikt}$ で f を最もよく近似しようとするとき、係数 \mathbf{c} は $\mathcal{F}_{\mathbb{Z}/N} \mathbf{x}$ の定数倍になる。また、各 $\phi_0, \dots, \phi_{N-1}$ を \sqrt{N} で割っておけば \mathbf{c} は完全に $\mathcal{F}_{\mathbb{Z}/N} \mathbf{x}$ と等しくなる。

5) 信号処理では慣習的に、整数値などの離散的な引数を大かっこで書く。

ノート 標準化周波数 N の値を大きくすれば、より多くの標本、より高周波の複素指数関数を近似に使える。そこで $N \rightarrow \infty$ の極限をとってみよう。エイリアシングを考慮すると、形式的には

$$f(c; t) = \sum_{k=-N/2}^{N/2-1} c_k e^{ikt} \rightarrow \sum_{k=-\infty}^{\infty} c_k e^{ikt},$$

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} f(t_n) e^{-ikt_n} = \frac{1}{2\pi} \sum_{n=0}^{N-1} f(t_n) e^{-ikt_n} \frac{2\pi}{N} \rightarrow \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt$$

となる。無限に高い周波数まで近似に含めたのだから、 $c_k = (1/(2\pi)) \int_0^{2\pi} f(t) e^{-ikt} dt$ のとき $f(t) = \sum c_k e^{ikt}$ が成り立つと期待するのは自然だろう。詳しくは第3章で説明するが、式 $f(t) = \sum c_k e^{ikt}$ を f のフーリエ級数展開といい、 f に適切な仮定を課せば実際に等式が成り立つことを証明できる。

2.4.4 巡回畳み込み

以下に式 (2.7) を再掲する。

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k n / N}$$

本来なら、式 (2.7) は $0 \leq n < N$ でしか意味をなさない。しかし、そのことを忘れて n が \mathbb{Z} 全体を動くとする、 x_n は周期 N の無限数列に拡張される。そこで、周期 N の数列 $x = \{x[n]\}$ に対し $\hat{x} = \mathcal{F}_{\mathbb{Z}/N} x$ を次のように定義する。

定義 2.4.3 N を周期に持つ \mathbb{K} 上の数列全体を $\mathbb{K}^{\mathbb{Z}/N}$ と書く。線型写像 $\mathcal{F}_{\mathbb{Z}/N}: \mathbb{C}^{\mathbb{Z}/N} \rightarrow \mathbb{C}^{\mathbb{Z}/N}$ を次式で定義する。

$$(\mathcal{F}_{\mathbb{Z}/N} x)[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-2\pi i k n / N} \quad (k \in \mathbb{Z})$$

明らかに $\mathbb{C}^{\mathbb{Z}/N}$ は数列空間の部分空間である。 $\mathbb{C}^{\mathbb{Z}/N}$ 上の**ラグ作用素** (lag operator) L を $Lx = \{x[n-1]\}$ で定義し、数列 $\mathbb{III} \in \mathbb{C}^{\mathbb{Z}/N}$ を

$$\mathbb{III}[n] = \begin{cases} 1 & (n \text{ は } N \text{ で割り切れる}), \\ 0 & (\text{otherwise}) \end{cases}$$

と定める⁶⁾。

6) \mathbb{III} (sha) はキリル文字。

命題 2.4.4 集合 $\mathcal{D} = \{L^m \text{III} \mid 0 \leq m < N\}$ は $\mathbb{C}^{\mathbb{Z}/N}$ の基底である。ただし $L^0 x = x$ とする。

証明 $x \in \mathbb{C}^{\mathbb{Z}/N}$, $m, n \in \{0, \dots, N-1\}$ を任意にとる。 $|n-m| < N$ なので、 $n-m$ が N で割り切れるのは $n-m=0$ のときだけである。よって $x[m](L^m \text{III})[n] = x[m]\text{III}[n-m] = x[n]\delta_{mn}$ だから

$$\sum_{m=0}^{N-1} x[m](L^m \text{III})[n] = x[n] \sum_{m=0}^{N-1} \delta_{mn} = x[n] \quad (2.8)$$

である。よって $x = \sum_{m=0}^{N-1} x[m] L^m \text{III}$ なので、 \mathcal{D} は基底である。 \square

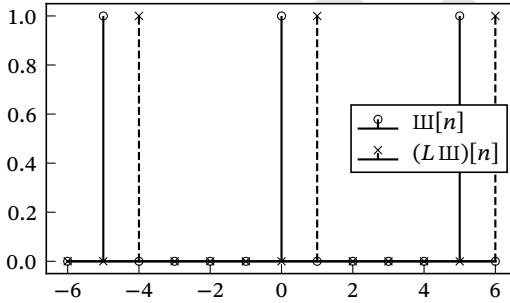


図 2.10 $N=5$ のときの III と $L \text{III}$ の様子。

命題 2.4.5 任意の $x \in \mathbb{C}^{\mathbb{Z}/N}$ に対して $(\mathcal{F}_{\mathbb{Z}/N} Lx)[k] = \zeta^{-k} \hat{x}[k]$ である。

証明 \sum の添字 n を $m = n+1$ に置き換えると

$$\zeta^{-k} \hat{x}[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-2\pi i k(n+1)/N} = \frac{1}{\sqrt{N}} \sum_{m=1}^N x[m-1] e^{-2\pi i k m/N}$$

となる。右辺は周期数列の 1 周期に渡る和だから、 $\sum_{m=1}^N$ を $\sum_{m=0}^{N-1}$ に変えても値は変わらない。よって $\zeta^{-k} \hat{x}[k] = (\mathcal{F}_{\mathbb{Z}/N} Lx)[k]$ である。 \square

命題 2.4.5 を使うと、数列の積 $x \cdot y = \{x[n]y[n]\}$ と離散フーリエ変換の間にある関係を示せる。積は $\mathcal{F}_{\mathbb{Z}/N}$ で、次の 2 項演算へと写される。

定義 2.4.6 (巡回畳み込み) $x, y \in \mathbb{C}^{\mathbb{Z}/N}$ とする. 次式で定義される $x * y \in \mathbb{C}^{\mathbb{Z}/N}$ を x と y の**巡回畳み込み** (circular convolution) という.

$$(x * y)[n] = \sum_{m=0}^{N-1} x[m]y[n-m] \quad (n \in \mathbb{Z})$$

$y[n-m] = (L^m y)[n]$ だから, 巡回畳み込みは次のように表せる.

$$x * y = \sum_{m=0}^{N-1} x[m] L^m y = \left(\sum_{m=0}^{N-1} x[m] L^m \right) y \quad (2.9)$$

命題 2.4.7 任意の $x, y \in \mathbb{C}^{\mathbb{Z}/N}$ に対して次式が成立する.

$$\mathcal{F}_{\mathbb{Z}/N}(x * y) = \sqrt{N} \hat{x} \cdot \hat{y}, \quad \mathcal{F}_{\mathbb{Z}/N}(x \cdot y) = \frac{1}{\sqrt{N}} \hat{x} * \hat{y}$$

証明 1 つめのみ示す. $z = x * y$ とすると, 式 (2.9) から

$$\hat{z} = \mathcal{F}_{\mathbb{Z}/N} \left(\sum_{m=0}^{N-1} x[m] L^m y \right) = \sum_{m=0}^{N-1} x[m] \mathcal{F}_{\mathbb{Z}/N} L^m y$$

であり, 命題 2.4.5 より $(\mathcal{F}_{\mathbb{Z}/N} L^m y)[k] = (\zeta^{-k})^m \hat{y}[k]$ なので

$$\hat{z}[k] = \sum_{m=0}^{N-1} x[m] (\zeta^{-km} \hat{y}[k]) = \left(\sum_{m=0}^{N-1} x[m] \zeta^{-km} \right) \hat{y}[k] = \sqrt{N} \hat{x}[k] \hat{y}[k]$$

となる. よって $\hat{z} = \sqrt{N} \hat{x} \cdot \hat{y}$ である. □

系 2.4.8 任意の $x, y, z \in \mathbb{C}^{\mathbb{Z}/N}$ に対して, 以下の式が成立する.

1. $x * y = y * x$
2. $(x * y) * z = x * (y * z)$
3. $x * (y + z) = (x * y) + (x * z)$

証明 2 つめのみ示す. $\mathcal{F}_{\mathbb{Z}/N}(x * (y * z)) = \sqrt{N} \hat{x} \cdot \mathcal{F}_{\mathbb{Z}/N}(y * z) = N \hat{x} \cdot (\hat{y} \cdot \hat{z})$ より $x * (y * z) = N \mathcal{F}_{\mathbb{Z}/N}^{-1}(\hat{x} \cdot \hat{y} \cdot \hat{z})$ であり, 同じ計算によって $(x * y) * z = N \mathcal{F}_{\mathbb{Z}/N}^{-1}(\hat{x} \cdot \hat{y} \cdot \hat{z})$ も確かめられる. □

ノート 式 (2.8), (2.9) より, 多項式 $X(T) = \sum x[m]T^m$, $Y(T) = \sum y[m]T^m$ に関して $x * y = X(T)Y(T)\text{III}$ が成立する. 実は, $X(T)Y(T)$ を $T^N - 1$ で割ったあまりは $\sum (x * y)[m]T^m$ である. 実際, 商を $Q(T)$, あまりを $R(T) = \sum r_m T^m$ とすれば, $X(T)Y(T) = Q(T)(T^N - 1) + R(T)$ と $L^N = 1$ から $X(L)Y(L) = R(L) = \sum r_m L^m$ であり, $(x * y)[n] = r_n$ ($0 \leq n < N$) が分かる.

$\mathbb{C}^{\mathbb{Z}/N}$ で得られた結果を \mathbb{C}^N で使える形に書き換えよう. \mathbb{C}^N の標準基底を $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ とおく. 各 $x \in \mathbb{C}^{\mathbb{Z}/N}$, $\mathbf{x} = (x[0] \ \cdots \ x[N-1])^\top$ に対して

$$\begin{aligned} \mathbf{x}^\top(\mathbf{e}_N \quad \mathbf{e}_1 \quad \cdots \quad \mathbf{e}_{N-1}) &= (x[N-1] \quad x[0] \quad \cdots \quad x[N-2]) \\ &= (x[-1] \quad x[0] \quad \cdots \quad x[N-2]) \end{aligned}$$

なので，基底 \mathcal{D} に関する L の表現行列 \mathbf{L} は $\mathbf{L} = (\mathbf{e}_N \ \mathbf{e}_1 \ \cdots \ \mathbf{e}_{N-1})^T = (\mathbf{I}_{N-1}^1)$ と書ける．同様に計算すれば， $0 < m < N$ のとき L^m の表現行列は $\mathbf{L}^m = \begin{pmatrix} & \mathbf{I}_m \\ \mathbf{I}_{N-m} & \end{pmatrix}$ であることも分かる．

$h \in \mathbb{C}^{\mathbb{Z}/N}$ を任意にとり, $\mathbb{C}^{\mathbb{Z}/N}$ 上の線型写像 H を $H(x) = h * x$ で定義する. 式(2.9) から $H = \sum_{m=0}^{N-1} h_m L^m$ ($h_m = h[m]$) なので, H の表現行列 \mathbf{H} は

$$\mathbf{H} = h_0 \mathbf{I} + \sum_{m=1}^{N-1} h_m (\mathbf{I}_{N-m} \mathbf{I}_m) = \begin{pmatrix} h_0 & h_{N-1} & \cdots & h_2 & h_1 \\ h_1 & h_0 & h_{N-1} & \ddots & h_2 \\ h_2 & h_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_0 & h_{N-1} \\ h_{N-1} & \cdots & h_2 & h_1 & h_0 \end{pmatrix}$$

である. この形の行列を巡回行列 (circulant matrix) という.

基底 \mathcal{D} に関する $\mathcal{F}_{\mathbb{Z}/N}$ の表現行列を \mathbf{W}^H とおく. \mathbf{W}^H は \mathbb{C}^N 上で定義した $\mathcal{F}_{\mathbb{Z}/N}$ の, 標準基底に関する表現行列でもある.

系 2.4.9 巡回行列は離散フーリエ変換で対角化される. すなわち, 任意の N 次巡回行列 H について $W^H H W$ は対角行列である.

証明 H の第 1 列を $(h_0 \ \dots \ h_{N-1})^T$ とおき, $h[n] = h_n$ ($0 \leq n < N$) となるように $\mathbb{C}^{Z/N}$ の元 h を定める. $\mathbb{C}^{Z/N}$ 上の線型写像 H を $H(x) = h * x$ で定義すると, 命題 2.4.7 より $(\mathcal{F}_{Z/N} H \mathcal{F}_{Z/N}^{-1})x = \mathcal{F}_{Z/N}(h * \mathcal{F}_{Z/N}^{-1}x) = \hat{h} \cdot x$ となる. よって, $\mathcal{F}_{Z/N} H \mathcal{F}_{Z/N}^{-1}$ の表現行列は $\text{diag}(\hat{h}[0], \dots, \hat{h}[N-1])$ である.

一方, H の表現行列は \mathbf{H} だから $\mathcal{F}_{\mathbb{Z}/N} H \mathcal{F}_{\mathbb{Z}/N}^{-1}$ の表現行列は $\mathbf{W}^H \mathbf{H} \mathbf{W}$ でもある. したがって $\mathbf{W}^H \mathbf{H} \mathbf{W}$ は対角行列である. \square

例 2.4.10 $N = 32$ とし, $\mathbb{C}^{\mathbb{Z}/N}$ の元 h を $h[n] = 1/3$ ($0 \leq n < 3$), $h[n] = 0$ ($3 \leq n < N$) で定義する. このとき

$$(h * x)[n] = \sum_{m=0}^{N-1} h[m]x[n-m] = \frac{x[n] + x[n-1] + x[n-2]}{3}$$

だから, $(h * x)[n]$ は時刻 $n, n-1, n-2$ における x の値の平均である.

平均をとると細かい変動が均されるので, 高周波の成分は減るだろう. 実際 $|\hat{h}[k]| = (1/3)|\zeta^{-k \cdot 0} + \zeta^{-k \cdot 1} + \zeta^{-k \cdot 2}|$ のグラフは図 2.11 のようになり, $\mathcal{F}_{\mathbb{Z}/N}(h * x) = \hat{h} \cdot \hat{x}$ だから h を畳み込むと高周波は削られる. ◇

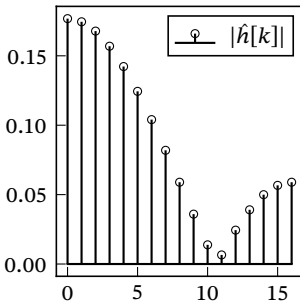


図 2.11 $|\hat{h}[k]|$ の様子.

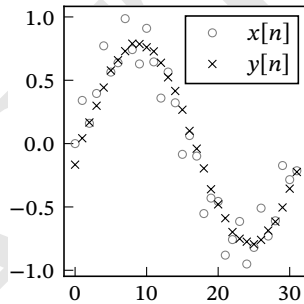


図 2.12 x と $y = h * x$ の一例.

なお, $k \geq 16$ で図 2.11 がどうなっているかという, 次の命題から直線 $k = 16$ に関して鏡写しになっている.

命題 2.4.11 $x \in \mathbb{R}^{\mathbb{Z}/N}$ ならば $\hat{x}[-k] = \overline{\hat{x}[k]}$ である.

証明 実際

$$\hat{x}[-k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \zeta^{-(-k)n} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \overline{x[n]} \zeta^{-kn} = \overline{\hat{x}[k]}$$

である. □

系 2.4.12 $x \in \mathbb{R}^N$ ならば $\hat{x}_{N-k} = \bar{\hat{x}}_k$ ($0 < k < N$) である.

証明 $x = (x_0 \ \cdots \ x_{N-1})^T$, $s = \sum_{m=0}^{N-1} x_m L^m$ III とおくと, 命題 2.4.11 と $\hat{s} \in \mathbb{C}^{\mathbb{Z}/N}$ より $\hat{x}_{N-k} = \hat{s}[N-k] = \hat{s}[-k] = \overline{\hat{s}[k]} = \bar{\hat{x}}_k$ である. □

2.5 短時間フーリエ変換

第 2.5 節, 第 2.6 節では, 時刻と周波数にまたがって現れる信号の特徴を調べる, **時間周波数解析** (time-frequency analysis) の理論を説明する. この理論は旋律・地震動・景気など, 時が経つにつれ波の様子が移り変わる信号を解析する分野で活用されている.

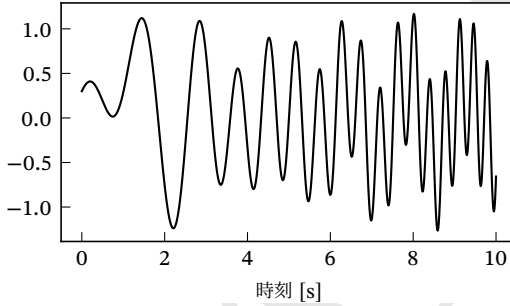


図 2.13 周波数が時間変化して見える信号の例.

たとえば, 図 2.13 の信号を見て「周波数が時とともに変化している」と感じるのは自然なことだろう. 時間周波数解析の技法を用いると, 周波数成分の時間変化をより定量的に追いかけられる.

2.5.1 短時間フーリエ変換

まず, 記号をいくつか定義する.

2 つの整数を引数にとる \mathbb{K} 上の数列 (2 重数列) 全体を $\mathbb{K}^{\mathbb{Z} \times \mathbb{Z}}$ と書く. また, $\mathbb{K}^{\mathbb{Z} \times \mathbb{Z}}$ の元 $x = \{x[m, n]\}$ で, m について周期 M , n について周期 N であるものの全体を $\mathbb{K}^{\mathbb{Z}^2/(M, N)}$ と表す. m の値を固定してできる数列 $\{x[m, n]\}_{n \in \mathbb{Z}} \in \mathbb{K}^{\mathbb{Z}}$ を $x[m, _]$ と表し, n についても同様に $x[_, n]$ を定めれば, $\mathbb{K}^{\mathbb{Z}^2/(M, N)}$ は

$$\mathbb{K}^{\mathbb{Z}^2/(M, N)} = \{x \in \mathbb{K}^{\mathbb{Z} \times \mathbb{Z}} \mid x[_, n] \in \mathbb{K}^{\mathbb{Z}/M}, x[m, _] \in \mathbb{K}^{\mathbb{Z}/N}\}$$

と書ける.

短時間フーリエ変換は, 数列 (信号) x に重みを掛けて離散フーリエ変換する手法である.

定義 2.5.1 (短時間フーリエ変換) L を自然数, w を $\mathbb{C}^{\mathbb{Z}/L}$ の元とする.
各 $x \in \mathbb{C}^{\mathbb{Z}/L}$ を

$$(\mathcal{V}_w^{(L)} x)[k, n] = \sum_{m=0}^{L-1} x[m] \overline{w[m-n]} e^{-2\pi i k m / L}$$

に写す写像 $\mathcal{V}_w^{(L)}: \mathbb{C}^{\mathbb{Z}/L} \rightarrow \mathbb{C}^{\mathbb{Z}^2/(L,L)}$ を**短時間フーリエ変換** (short-time Fourier transform; STFT) という.

w を $\mathcal{V}_w^{(L)}$ の**窓関数** (window function) という. 文脈から L の値が明らかであれば, $\mathcal{V}_w^{(L)}$ を単に \mathcal{V}_w と書く.

$x, y \in \mathbb{C}^{\mathbb{Z}/L}$ の内積を $\langle x, y \rangle = \sum_{n=0}^{L-1} x[n] \overline{y[n]}$ で定義する.

命題 2.5.2 $\mathbb{C}^{\mathbb{Z}/L}$ 上の写像 T_m, M_k を

$$(T_m x)[n] = x[n-m], \quad (M_k x)[n] = e^{2\pi i k n / L} x[n]$$

で定義する. このとき, 任意の $w, x \in \mathbb{C}^{\mathbb{Z}/L}$ に対して次式が成立する.

$$(\mathcal{V}_w x)[k, n] = \langle x, M_k T_n w \rangle \quad (k, n \in \mathbb{Z})$$

証明 実際

$$\begin{aligned} \langle x, M_k T_n w \rangle &= \sum_{m=0}^{L-1} x[m] \overline{e^{2\pi i k m / L} (T_n w)[m]} \\ &= \sum_{m=0}^{L-1} x[m] e^{-2\pi i k m / L} \overline{w[m-n]} = (\mathcal{V}_w x)[k, n] \end{aligned}$$

である. □

命題 2.5.2 から, 次の系が得られる.

系 2.5.3 \mathcal{V}_w は線型写像である.

定義 2.5.4 (離散ガボール変換) (K, N, L) を自然数の 3 つ組, w を $\mathbb{C}^{\mathbb{Z}/L}$ の元とする. 各 $x \in \mathbb{C}^{\mathbb{Z}/L}$ を

$$\begin{aligned} (\mathcal{G}_w^{(K, N, L)} x)[i, j] &= (\mathcal{V}_w^{(L)} x)[iK, jN] \\ &= \sum_{m=0}^{L-1} x[m] \overline{w[m - jN]} e^{-2\pi i(iK)m/L} \end{aligned}$$

に写す線型写像 $\mathcal{G}_w^{(K, N, L)}: \mathbb{C}^{\mathbb{Z}/L} \rightarrow \mathbb{C}^{\mathbb{Z}^2/(L/K, L/N)}$ を **離散ガボール変換** (discrete Gabor transform; DGT) という.

定義 2.5.5 (ε -集中) $\varepsilon > 0$, $x \in \mathbb{C}^{\mathbb{Z}/N}$ とする. 条件

$$\sum_{n \in I \setminus A} |x[n]|^2 \leq \varepsilon^2 \sum_{n \in I} |x[n]|^2 \quad (I = \{0, \dots, N-1\})$$

を満たすとき, x は集合 $A \subseteq I$ に **ε -集中** (ε -concentrated) であるという.

定理 2.5.6 (Donoho-Stark の不確定性原理) $\varepsilon_T, \varepsilon_F > 0$ とする. $\mathbb{C}^{\mathbb{Z}/N}$ の元 x が集合 T に ε_T -集中, $\mathcal{F}_{\mathbb{Z}/N} x$ が集合 F に ε_F -集中であるとき

$$(\#T)(\#F) \geq N(1 - \varepsilon_T - \varepsilon_F)^2$$

である. これを **Donoho-Stark の不確定性原理** (Donoho-Stark uncertainty principle) という.

2.6 多重解像度解析

2.7 補遺

2.7.1 スペクトル定理の証明

定理 2.3.4 を証明する前に, 次の定理を示そう.

定理 2.7.1 (シュア分解) 任意の複素正方行列はユニタリ行列で上三角化できる. すなわち, 任意の n 次複素正方行列 A に対し, n 次ユニタリ行列 U が存在して, $A' = U^H A U$ は上三角行列⁷⁾になる. 式 $A = U A' U^H$ を A のシュア分解 (Schur decomposition) という.

証明 n に関する帰納法で示す. u_1 を A のある固有値 λ に属する, ノルム 1 の固有ベクトルとする. また, 集合 $\{u_2, \dots, u_n\}$ を $(\text{span}\{u_1\})^\perp$ の正規直交基底とする. このとき, 集合 $\{u_1, \dots, u_n\}$ は \mathbb{C}^n の正規直交基底だから, $U = (u_1 \ \dots \ u_n)$ はユニタリ行列である.

\mathbb{C}^n の元 e_1 を $e_1 = (1 \ 0 \ \dots \ 0)^T$ で定めると, $U^H A u_1 = U^H (\lambda u_1) = \lambda e_1$ なので $U^H A U = (\lambda e_1 \ U^H A u_2 \ \dots \ U^H A u_n)$ である. 右辺を $\begin{pmatrix} \lambda & b^T \\ 0 & C \end{pmatrix}$ とおく (C は $n-1$ 次正方行列, $b \in \mathbb{C}^{n-1}$). 帰納法の仮定から C はシュア分解できる. $C = V C' V^H$ をシュア分解とし, n 次ユニタリ行列 W を $W = \begin{pmatrix} 1 & \\ & V \end{pmatrix}$ で定義する. このとき

$$\begin{pmatrix} \lambda & b^T \\ 0 & C \end{pmatrix} = \begin{pmatrix} \lambda & b^T \\ 0 & V C' V^H \end{pmatrix} = \begin{pmatrix} 1 & \\ & V \end{pmatrix} \begin{pmatrix} \lambda & b^T V \\ 0 & C' \end{pmatrix} \begin{pmatrix} 1 & \\ & V^H \end{pmatrix}$$

なので, $A' = \begin{pmatrix} \lambda & b^T V \\ 0 & C' \end{pmatrix}$ とおくと $U^H A U = W A' W^H$ である. よって, UW をあらためて U とすれば $A = U A' U^H$ で, A' は上三角行列だから, これは A のシュア分解になっている. \square

定理 2.7.1 を利用すると, 定理 2.3.4 は次のようにして示せる.

スペクトル定理の証明 ユニタリ行列で対角化できれば正規行列であることはすぐ分かる. A を n 次正規行列とする. A のシュア分解を $A = U B U^H$ とおくと, $A^H A - A A^H = U (B^H B - B B^H) U^H = O$ より B も正規行列である.

$B = (b_{ij})$ とおく. $B^H = (\bar{b}_{ji})$, $(i > j \implies b_{ij} = 0)$ なので, $B^H B = B B^H$ の d 行 d 列にある成分を比較すれば

$$\sum_{i=1}^n \bar{b}_{id} b_{id} = \sum_{j=1}^n b_{dj} \bar{b}_{dj}, \quad \sum_{i=1}^d |b_{id}|^2 = \sum_{j=d}^n |b_{dj}|^2$$

が分かる. $d = 1$ とすると, $|b_{11}|^2 = \sum_{j=1}^n |b_{1j}|^2$ より $b_{12} = \dots = b_{1n} = 0$ である. 次に $d = 2$ とすると, $|b_{12}|^2 + |b_{22}|^2 = |b_{22}|^2 = \sum_{j=2}^n |b_{2j}|^2$ より $b_{23} =$

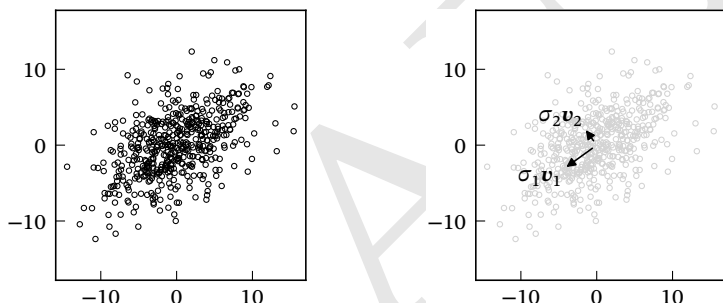
7) $a_{ij} = 0$ ($i > j$) である正方行列 (a_{ij}) を上三角行列 (upper triangular matrix) という.

$\cdots = b_{2n} = 0$ である. 同様に $d = n$ まで計算すれば ($i < j \implies b_{ij} = 0$) が分かる. \mathbf{B} は上三角行列だから, これは \mathbf{B} が対角行列であることを意味する. つまり, シューア分解 $\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{U}^H$ は \mathbf{A} のスペクトル分解になっている. \square

2.7.2 主成分分析

最小 2 乗法のもう一つの応用例として, **主成分分析** (principal component analysis; PCA) という手法を紹介する.

変数 $\mathbf{x} = (x_1 \ \cdots \ x_p)^T$ の観測値 $\mathbf{x}_1, \dots, \mathbf{x}_n$ を眺めても, \mathbf{x} が高次元⁸⁾だとその特徴が分かりにくいことがある. 主成分分析の目的は, この問題を解決するために \mathbf{x} の特徴を保ったまま次元を減らすことである.



2.8 演習問題

- 系 2.3.5 の証明を完成させよ.
- 任意の $n \times p$ 複素行列 \mathbf{A} , $\mathbf{x} \in \mathbb{C}^p$, $\mathbf{y} \in \mathbb{C}^n$ に対して, 以下の式が成立することを示せ.

$$\text{proj}_{\text{nul } \mathbf{A}} \mathbf{x} = (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{x}, \quad \text{proj}_{(\text{nul } \mathbf{A})^\perp} \mathbf{x} = \mathbf{A}^+ \mathbf{A} \mathbf{x},$$

$$\text{proj}_{\text{col } \mathbf{A}} \mathbf{y} = \mathbf{A} \mathbf{A}^+ \mathbf{y}, \quad \text{proj}_{(\text{col } \mathbf{A})^\perp} \mathbf{y} = (\mathbf{I} - \mathbf{A} \mathbf{A}^+) \mathbf{y}$$

- n 次複素正方行列 \mathbf{A} に対して $\exp \mathbf{A}$ を $\exp \mathbf{A} = \sum_{k=0}^{\infty} \mathbf{A}^k / (k!)$ で定義する (この極限は各成分の極限で定義され, $\exp \mathbf{A}$ を定義する級数は常に収束することが分かっている). \mathbf{A} が正規行列なら, そのスペクトル分解を

8) p の値が非常に大きいこと.

$\mathbf{A} = \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^H$ とすると $\exp \mathbf{A} = \mathbf{U} \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \mathbf{U}^H$ であることを示せ.

4. $\mathcal{F}_{\mathbb{Z}/N} \mathbf{x} = (\hat{x}_0 \ \cdots \ \hat{x}_{N-1})^T$ ($\mathbf{x} = (x_0 \ \cdots \ x_{N-1})^T$) のとき

$$\sum_{k=0}^{N-1} |\hat{x}_k|^2 = \sum_{n=0}^{N-1} |x_n|^2 \quad (2.10)$$

であることを示せ. 信号処理ではしばしば式 (2.10) を **パーセヴァルの定理** (Parseval's theorem), あるいは **プランシュレルの定理** (Plancherel's theorem) と呼ぶ.

3.1 イントロダクション

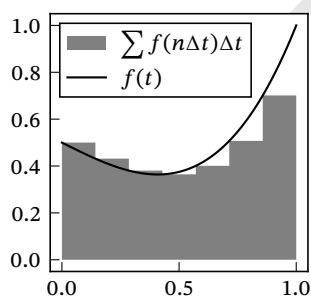


図 3.1 $f(t)$ と $\sum f(n\Delta t)\Delta t$ の比較.

$$\sum_{n=0}^{N-1} x(n\Delta t) \overline{y(n\Delta t)} \Delta t \rightarrow \int_0^1 x(t) \overline{y(t)} dt \quad (N \rightarrow \infty)$$

3.2 無限次元のベクトル空間

3.2.1 距離空間

定義 3.2.1 (距離) S を集合とする. d が S 上の**距離** (metric) であるとは, 任意の $x, y, z \in S$ に対して, d が以下の条件を満たすことをいう.

1. $d(x, y) \geq 0$, $(d(x, y) = 0 \iff x = y)$
2. $d(x, y) = d(y, x)$
3. $d(x, y) + d(y, z) \geq d(x, z)$

集合と距離の組 (S, d) を**距離空間** (metric space) という.

例 3.2.2 $S = \mathbb{C}$, $d(z, w) = |z - w|$ とすると, (S, d) は距離空間になる. ◇

例 3.2.3 (離散距離) 集合 S は空でないとする. また, 各 $x, y \in S$ に対して, $x = y$ のとき $d(x, y) = 0$, $x \neq y$ のとき $d(x, y) = 1$ とする. このとき d は S 上の距離になる. 距離 d を**離散距離** (discrete metric), 距離空間 (S, d) を**離散空間** (discrete space) という. ◇

定義 3.2.1 のように抽象的な形で距離を定義する利点の 1 つは, \mathbb{K}^n 以外の集合に対しても, 点列の極限を定義できることである.

定義 3.2.4 (点列の収束) (S, d) を距離空間とする. S 上の点列 $\{x_n\}_{n \in \mathbb{N}}$ が $\alpha \in S$ に**収束する** (converge) とは, 任意の $\varepsilon > 0$ に対し, $N \in \mathbb{N}$ が存在して $(n > N \implies d(x_n, \alpha) < \varepsilon)$ を満たすことをいう. このことを次のように表す.

$$x_n \rightarrow \alpha \quad (n \rightarrow \infty)$$

ノート 次の命題が成り立つことに注意.

$$x_n \rightarrow \alpha \quad (n \rightarrow \infty) \iff \lim_{n \rightarrow \infty} d(x_n, \alpha) = 0$$

$\{x_n\}_{n \in \mathbb{N}}$ が α に収束するとき, α を $\{x_n\}_{n \in \mathbb{N}}$ の**極限点** (limit point) という. 定義 3.2.4 は要するに「 N の値を十分に大きくとれば, 点 x_{N+1}, x_{N+2}, \dots が点

α から距離 ε 以上離れないようにできる」ことを意味する。

例 3.2.5 (S, d) を例 3.2.2 の距離空間とする。 S 上の点列 $\{z_n\}_{n \in \mathbb{N}}$ を $z_n = (\sqrt{3}+i)/(2n)$ で定義すると、 $\{z_n\}_{n \in \mathbb{N}}$ は定義 3.2.4 の意味で $z_n \rightarrow 0$ ($n \rightarrow \infty$) を満たす。 ◇

例 3.2.6 (一様収束) 閉区間 I は有界とする。連続関数 $f: I \rightarrow \mathbb{R}$ の全体集合を $C^0(I)$ とおくと、 $d(f, g) = \max\{|f(t) - g(t)| \mid t \in I\}$ は $C^0(I)$ 上の距離になる。 $C^0(I)$ 上の関数列 $\{f_n\}_{n \in \mathbb{N}}$ が定義 3.2.4 の意味で $f \in C^0(I)$ に収束するとき、 $\{f_n\}_{n \in \mathbb{N}}$ は f に**一様収束する** (converge uniformly) という。

たとえば $I = [0, 1]$, $f_n(t) = (1/n)|\sin(n\pi t)|$ のとき、 $\{f_n\}_{n \in \mathbb{N}}$ は定数関数 $\phi(t) = 0$ に一様収束する。実際 $d(f_n, \phi) = \max\{|f_n(t)| \mid t \in I\} = 1/n$ なので、 n の値を十分大きくとれば $d(f_n, \phi)$ の値を限りなく小さくできる (図 3.3)。 ◇

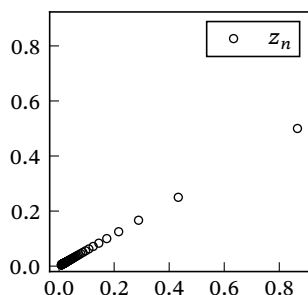


図 3.2 $z_n \rightarrow 0$ ($n \rightarrow \infty$) の様子。

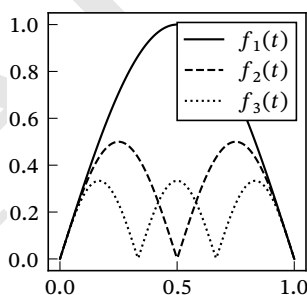


図 3.3 $f_n \rightarrow \phi$ ($n \rightarrow \infty$) の様子。

ノート 例 3.2.5 において $|z_n| = 1/n$ であるから、 $d(f_n, \phi) = |z_n|$ である。よって、図 3.2 は (z_n を f_n に書き換えれば) $f_n \rightarrow \phi$ ($n \rightarrow \infty$) の様子を描いた図とも考えられる。このように、関数などの一見「点」とは思えないような対象を点とみなして考察するのは、しばしば理解の助けになる。

命題 3.2.7 極限点は存在すれば一意である。すなわち、距離空間 (S, d) 上の点列 $\{x_n\}_{n \in \mathbb{N}}$ が $\alpha, \beta \in S$ に収束するなら、 $\alpha = \beta$ である。

証明 $0 \leq d(\alpha, \beta) \leq d(\alpha, x_n) + d(x_n, \beta) = d(x_n, \alpha) + d(x_n, \beta)$ なので、 $d(x_n, \alpha) \rightarrow 0$, $d(x_n, \beta) \rightarrow 0$ ($n \rightarrow \infty$) なら $d(\alpha, \beta) = 0$, $\alpha = \beta$ である。 □

命題 3.2.7 から、各収束列 $\{x_n\}_{n \in \mathbb{N}}$ に対して、その極限点は一意に定まる。そのため、以降は収束列 $\{x_n\}_{n \in \mathbb{N}}$ の極限点を

$$\lim_{n \rightarrow \infty} x_n$$

と書く。

定義 3.2.8 (閉包・閉集合・稠密) (S, d) を距離空間、 A を S の部分集合とする。

1. A 上の収束列すべての極限点からなる集合を A の**閉包** (closure) といい、 $\text{cl}A$ と書く¹⁾。
2. $A = \text{cl}A$ であるとき、 A は**閉集合** (closed set) であるという。
3. 集合 $B \subseteq A$ が $\text{cl}B = A$ を満たすとき、 B は A において**稠密** (dense) であるという。

例 3.2.9 $\text{cl}(0, 1] = [0, 1]$, $\text{cl}\mathbb{Q} = \mathbb{R}$ である。◇

定義 3.2.10 (コーシー列) (S, d) を距離空間とする。 S 上の点列 $\{x_n\}_{n \in \mathbb{N}}$ が**コーシー列** (Cauchy sequence) であるとは、任意の $\varepsilon > 0$ に対し、 $N \in \mathbb{N}$ が存在して $(m, n > N \implies d(x_m, x_n) < \varepsilon)$ を満たすことをいう。このことを次のように表す。

$$d(x_m, x_n) \rightarrow 0 \quad (m, n \rightarrow \infty), \quad \lim_{m, n \rightarrow \infty} d(x_m, x_n) = 0$$

また、 S 上の任意のコーシー列が収束列でもあるとき、 (S, d) は**完備距離空間** (complete metric space) であるという。一般に収束列はコーシー列でもあるから、完備距離空間において収束列とコーシー列は同値な概念である。

例 3.2.11 $S = \mathbb{Q}$, $d(x, y) = |x - y|$ とすると、 (S, d) は距離空間になるが完備距離空間にはならない。◇

1) 本書では閉包を $\text{cl}A$, 補集合を A^c で表す。

3.2.2 ノルム空間

定義 3.2.12 (ノルム) V を \mathbb{K} 上のベクトル空間とする. $\|\cdot\|$ が V の **ノルム** (norm) であるとは, 任意の $\lambda \in \mathbb{K}$, $x, y \in V$ に対して, $\|\cdot\|$ が以下の条件を満たすことをいう.

1. $\|x\| \geq 0$, $(\|x\| = 0 \iff x = 0)$
2. $\|\lambda x\| = |\lambda| \|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$

ノルムが備わっているベクトル空間のことを **ノルム空間** (normed space) という. V がノルム空間であれば, $d(x, y) = \|x - y\|$ ($x, y \in V$) により V 上の距離 d が定義される. (V, d) が完備距離空間であるとき, V は **バナッハ空間** (Banach space) であるという.

例 3.2.13 V が \mathbb{K} 上の内積空間なら, V の内積 $\langle \cdot, \cdot \rangle$ から V のノルムを $\|x\| = \sqrt{\langle x, x \rangle}$ で定義できる. つまり, 内積空間はノルム空間でもある. ◇

例 3.2.14 (ℓ^p 空間) 各 $x = \{x[n]\}_{n \in \mathbb{Z}} \in \mathbb{K}^{\mathbb{Z}}$ に対して $\|x\|_p \in [0, +\infty]$ ($p \geq 1$) を

$$\|x\|_p = \left(\sum_{n=-\infty}^{\infty} |x[n]|^p \right)^{1/p}$$

で定義する. このとき, 集合 $\ell^p(\mathbb{Z}) = \{x \in \mathbb{K}^{\mathbb{Z}} \mid \|x\|_p < +\infty\}$ は $\mathbb{K}^{\mathbb{Z}}$ の部分空間かつ, $\|\cdot\|_p$ をノルムとするバナッハ空間である. バナッハ空間 $\ell^p(\mathbb{Z})$ を **ℓ^p 空間** (ℓ^p space) という. ◇

例 3.2.15 例 3.2.6 の集合 $C^0(I)$ は, ノルム $\|f\|_{\infty} = \max\{|f(t)| \mid t \in I\}$ についてバナッハ空間になる. ただし, 関数の和 $\phi = f + g$ とスカラー倍 $\psi = \lambda f$ はそれぞれ $\phi(t) = f(t) + g(t)$, $\psi(t) = \lambda \cdot (f(t))$ で定義する. ◇

3.3 ヒルベルト空間

定義 3.3.1 (ヒルベルト空間) 内積空間 H がヒルベルト空間 (Hilbert space) であるとは, H の内積 $\langle _, _ \rangle$ から定まるノルム $\|x\| = \sqrt{\langle x, x \rangle}$ について, H がバナッハ空間であることをいう。

もう少し定義をさかのぼると, ノルム空間 H がバナッハ空間であるとは, 距離 $d(x, y) = \|x - y\|$ について (H, d) が完備距離空間であることをいうのであった。したがって, 完備距離空間・ノルム空間・バナッハ空間・内積空間が有する性質はすべて, ヒルベルト空間にも引き継がれる。

ノート 以下に述べる命題は, 内積空間であればすべて成立する。内積空間がヒルベルト空間であるための条件「完備性」は, 条件を満たす点列に対して, 極限点の存在を保証するものである。そのため, ヒルベルト空間でないと成立しない定理は, 存在を主張する定理であることが多い。本書においても, 存在定理である定理 3.4.2 で初めて, 完備性が本質的に効いてくる。

定理 3.3.2 (中線定理) V を内積空間とすると, 任意の $x, y \in V$ に対して $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ が成立する。

証明 式 $\|x \pm y\|^2 = \|x\|^2 \pm 2 \operatorname{Re}\langle x, y \rangle + \|y\|^2$ (複合同順) から分かる。 □

定理 3.3.3 (コーシー・シュワルツの不等式) V を内積空間とする。このとき, 任意の $a, b \in V$ について $|\langle a, b \rangle| \leq \|a\| \|b\|$ が成立する。これを**コーシー・シュワルツの不等式** (Cauchy-Schwarz inequality) という。

証明 $b \neq 0$ のときについて示す。 $\lambda = \langle a, b \rangle / \|b\|^2$, $a_\perp = a - \lambda b$ とおくと, $\langle a_\perp, b \rangle = 0$ より $\|a_\perp\|^2 = \langle a_\perp, a_\perp \rangle = \langle a - \lambda b, a - \lambda b \rangle = \langle a, a_\perp \rangle - \lambda \langle b, a_\perp \rangle = \langle a, a_\perp \rangle = \langle a, a - \lambda b \rangle = \|a\|^2 - \bar{\lambda} \langle a, b \rangle = \|a\|^2 - |\langle a, b \rangle|^2 / \|b\|^2$ である。よって $(\|a\| \|b\|)^2 - |\langle a, b \rangle|^2 = (\|a_\perp\| \|b\|)^2 \geq 0$ である。 □

命題 3.3.4 (ノルムの連続性) V がノルム空間なら, V 上の任意の収束列 $\{x_n\}$ について次式が成立する.

$$\lim_{n \rightarrow \infty} \|x_n\| = \left\| \lim_{n \rightarrow \infty} x_n \right\|$$

証明 $\{x_n\}$ を V 上の収束列とし, 極限点を a とおく. このとき $\|x_n\| \leq \|x_n - a\| + \|a\|$, $\|a\| \leq \|a - x_n\| + \|x_n\|$ なので $\|x_n\| - \|a\| \leq \|x_n - a\| \rightarrow 0$ ($n \rightarrow \infty$), よって $\|x_n\| \rightarrow \|a\|$ ($n \rightarrow \infty$) である. \square

命題 3.3.5 (内積の連続性) V が内積空間なら, V 上の任意の収束列 $\{x_n\}$, $\{y_n\}$ について次式が成立する.

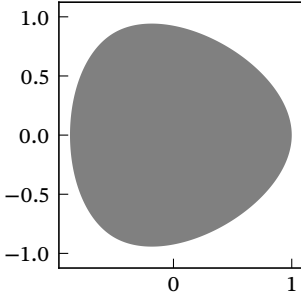
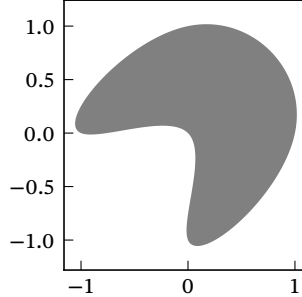
$$\lim_{k \rightarrow \infty} \langle x_k, y_k \rangle = \left\langle \lim_{m \rightarrow \infty} x_m, \lim_{n \rightarrow \infty} y_n \right\rangle$$

証明 $x_n \rightarrow a$, $y_n \rightarrow b$ ($n \rightarrow \infty$) とする. $\langle x_n, y_n \rangle = \langle x_n - a, y_n \rangle + \langle a, y_n - b \rangle + \langle a, b \rangle$ だから, [コーシー・シュワルツの不等式](#) より $|\langle x_n, y_n \rangle - \langle a, b \rangle| \leq |\langle x_n - a, y_n \rangle| + |\langle a, y_n - b \rangle| \leq \|x_n - a\| \|y_n\| + \|a\| \|y_n - b\|$ である. 命題 3.3.4 より $\|x_n - a\| \|y_n\| \rightarrow 0 \|b\|$, $\|y_n - a\| \rightarrow 0$ ($n \rightarrow \infty$) なので, $\langle x_n, y_n \rangle \rightarrow \langle a, b \rangle$ ($n \rightarrow \infty$) である. \square

3.4 直交射影

3.4.1 直交射影

定義 3.4.1 (線分, 凸集合) V をベクトル空間とする. 2 点 $x, y \in V$ に対し, 集合 $\{(1-t)x + ty \mid t \in [0, 1]\}$ を x と y を結ぶ**線分** (line segment) という. また, 集合 $S \subseteq V$ に属する任意の 2 点を結ぶ線分が S に含まれるとき, S は**凸集合** (convex set) であるという.

図 3.4 \mathbb{R}^2 の凸集合.図 3.5 \mathbb{R}^2 の凸集合でない部分集合.

定理 3.4.2 (凸射影定理) H をヒルベルト空間とする. また, $x \in H$ かつ, 集合 $C \subseteq H$ は空でない閉凸集合とする. このとき, $\arg \min_{y \in C} \|x - y\|$ はただ一つの元からなる集合である.

証明 $\delta = \inf\{\|x - y\| \mid y \in C\}$ とおくと, 集合 $A_n = \{y \in C \mid \|x - y\| \leq \delta + 1/n\}$ ($n \in \mathbb{N}$) は n の値によらず空でない. そこで, 各 n に対して A_n の元 a_n を 1 つずつ選べる. $\delta \leq \|x - a_n\| \leq \delta + 1/n$ なので, $\{a_n\}$ の極限点 a_∞ が存在すれば $\|x - a_\infty\| = \lim_{n \rightarrow \infty} \|x - a_n\| = \delta$ である. また, C は閉集合だから $a_\infty \in C$, よって $a_\infty \in \arg \min_{y \in C} \|x - y\|$ である.

要するに, a_∞ が存在する——つまり $\{a_n\}$ がコーシー列である——ならば, $\arg \min_{y \in C} \|x - y\|$ は空でない. $m, n \in \mathbb{N}$ を任意にとる. 中線定理より

$$\begin{aligned} \|(x - a_m) + (x - a_n)\|^2 + \|a_m - a_n\|^2 &= 2(\|x - a_m\|^2 + \|x - a_n\|^2), \\ \|a_m - a_n\|^2 &= 2\|x - a_m\|^2 + 2\|x - a_n\|^2 - 4\left\|x - \frac{a_m + a_n}{2}\right\|^2 \end{aligned}$$

である. $a_m \in A_m$, $a_n \in A_n$ かつ, C は凸集合だから $(a_m + a_n)/2 \in C$ で

$$\|a_m - a_n\|^2 \leq 2\left(\delta + \frac{1}{m}\right)^2 + 2\left(\delta + \frac{1}{n}\right)^2 - 4\delta^2 \rightarrow 0 \quad (m, n \rightarrow \infty)$$

である. よって $\{a_n\}$ はコーシー列なので, $\arg \min_{y \in C} \|x - y\|$ は空でない.

次に, $\arg \min_{y \in C} \|x - y\|$ の元は 1 つしかないことを示す. $y_1, y_2 \in \arg \min_{y \in C} \|x - y\|$ とする. このとき $\|x - y_1\| = \|x - y_2\| = \delta$, $(y_1 + y_2)/2 \in C$ だから $\|y_1 - y_2\|^2 = 2\delta^2 + 2\delta^2 - 4\|x - (y_1 + y_2)/2\|^2 \leq 0$, したがって $y_1 = y_2$ である. \square

定理 3.4.3 (射影定理) H をヒルベルト空間とする。また、 $x \in H$ かつ、 V は H の閉部分空間とする。このとき、 V の元 m に関する以下の条件は同値であり、条件を満たす m はただ一つ存在する。

1. $m \in \arg \min_{y \in V} \|x - y\|$ である。
2. 任意の $v \in V$ に対して $\langle x - m, v \rangle = 0$ である。

証明 閉部分空間は閉凸集合だから、**凸射影定理**より $m \in \arg \min_{y \in V} \|x - y\|$ を満たす m が一意に定まる。あとは命題 2.2.6 と同様に示せる。 \square

定義 3.4.4 (直交射影) 定理 3.4.3 の m を x の V への**直交射影**(orthogonal projection) といい、 $\text{proj}_V x$ と表す。

命題 3.4.5 H はヒルベルト空間で、 V は H の閉部分空間とする。このとき $H = V \oplus V^\perp$ である。

3.4.2 正規直交系

射影定理は直交射影 $\text{proj}_V x$ の存在を示す定理であり、具体的な式を与えるものではない。しかし、 V が正規直交系によって生成される空間（正確にはその閉包）であれば、 $\text{proj}_V x$ の具体的な式が得られる。

定義 3.4.6 (正規直交系) H をヒルベルト空間、 $\{\phi_n\}$ を H 上の点列とする。 $\langle \phi_i, \phi_j \rangle = \delta_{ij}$ ($i, j \in \mathbb{N}$) であるとき、 $\{\phi_n\}$ は**正規直交系** (orthonormal system; ONS) であるという。

定理 3.4.7 (ベッセルの不等式) H をヒルベルト空間とする。 H 上の点列 $\{\phi_n\}$ が正規直交系なら、任意の $x \in H$ に対して次式が成立する。これを**ベッセルの不等式** (Bessel's inequality) という。

$$\sum_{n=1}^{\infty} |\langle x, \phi_n \rangle|^2 \leq \|x\|^2 \quad (3.1)$$

証明 命題 2.2.4 の式 (2.1) と同様に計算すると、任意の $z_1, \dots, z_m \in \mathbb{C}$ に対して次式が成り立つと分かる。

$$\left\| x - \sum_{k=1}^m z_k \phi_k \right\|^2 = \|x\|^2 + \sum_{k=1}^m |z_k - \langle x, \phi_k \rangle|^2 - \sum_{k=1}^m |\langle x, \phi_k \rangle|^2$$

したがって、特に $z_k = \langle x, \phi_k \rangle$ なら

$$\|x\|^2 = \left\| x - \sum_{k=1}^m \langle x, \phi_k \rangle \phi_k \right\|^2 + \sum_{k=1}^m |\langle x, \phi_k \rangle|^2 \geq \sum_{k=1}^m |\langle x, \phi_k \rangle|^2$$

である。よって、級数 $\sum |\langle x, \phi_n \rangle|^2$ は上に有界な正項級数だから収束し、級数の和は式 (3.1) を満たす。 \square

定理 3.4.7 の状況で、点列 $\{x_n\}$ を $x_n = \sum_{k=1}^n \langle x, \phi_k \rangle \phi_k$ で定義すると、 $\{x_n\}$ は収束列になる。実際、 $m > n$ なら

$$\|x_m - x_n\|^2 = \left\| \sum_{k=n+1}^m \langle x, \phi_k \rangle \phi_k \right\|^2 = \sum_{k=n+1}^m |\langle x, \phi_k \rangle|^2$$

となるので、 $\{x_n\}$ がコーシー列であることと、級数 $\sum |\langle x, \phi_n \rangle|^2$ がコーシー列であることは同値である。そして、式 (3.1) の級数は収束しているから、 $\{x_n\}$ はコーシー列である。

命題 3.4.8 H をヒルベルト空間とする。 H 上の点列 $\{\phi_n\}$ が正規直交系なら、任意の $x \in H$ について次式が成立する。

$$\text{proj}_{\text{cl } V} x = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n \quad (V = \text{span}\{\phi_1, \phi_2, \dots\})$$

証明 $v \in \text{cl } V$ を任意にとる。 $\epsilon = x - \sum \langle x, \phi_n \rangle \phi_n$ が $\langle \epsilon, v \rangle = 0$ を満たせば、射影定理から $\text{proj}_{\text{cl } V} x = \sum \langle x, \phi_n \rangle \phi_n$ といえる。

$v \in \text{cl } V$ なので、 $\tilde{v}_n \rightarrow v$ ($n \rightarrow \infty$) を満たす V 上の点列 $\{\tilde{v}_n\}$ がある。また、 $\langle \phi_i, \phi_j \rangle = \delta_{ij}$ より

$$\langle \epsilon, \phi_j \rangle = \left\langle x - \sum_{i=1}^{\infty} \langle x, \phi_i \rangle \phi_i, \phi_j \right\rangle = \langle x, \phi_j \rangle - \sum_{i=1}^{\infty} \langle x, \phi_i \rangle \langle \phi_i, \phi_j \rangle = 0$$

だから, $V = \text{span}\{\phi_1, \phi_2, \dots\}$ の元はすべて ϵ と直交する. よって $\langle \epsilon, \tilde{v}_n \rangle = 0$, $\langle \epsilon, v \rangle = \langle \epsilon, v - \tilde{v}_n \rangle$ である. $n \rightarrow \infty$ とすれば $\langle \epsilon, v \rangle = \langle \epsilon, 0 \rangle = 0$ が分かる. \square

命題 3.4.8 より, $\text{cl } V = H$ であれば任意の $x \in H$ に対して

$$x = \text{proj}_H x = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$$

が成立する. そのような正規直交系 $\{\phi_n\}$ は完全正規直交系と呼ばれる.

定義 3.4.9 (完全正規直交系) H をヒルベルト空間, $\{\phi_n\}$ を H 上の正規直交系とする. $\text{span}\{\phi_1, \phi_2, \dots\}$ が H において稠密であるとき, $\{\phi_n\}$ は**完全正規直交系** (complete orthonormal system; CONS) であるという.

3.5 L^p 空間

定義 3.5.1 (L^p 空間) 集合 $\Omega \subseteq \mathbb{R}$ はルベグ可測とする. 各 $p \in [1, +\infty)$ に対し, 可測関数 $f: \Omega \rightarrow \mathbb{K}$ で

$$\|f\|_p = \left(\int_{\Omega} |f(t)|^p dt \right)^{1/p}$$

の値が有限であるものの全体集合を $L^p(\Omega)$ とおく. このとき, ほとんどいたるところ等しい関数を同一視すれば, $L^p(\Omega)$ は $\|\cdot\|_p$ をノルムとしてバナッハ空間になる. このバナッハ空間を **L^p 空間** (L^p space) という.

命題 3.5.2 (L^2 空間の性質) $p = 2$ のときのみ $L^p(\Omega)$ はヒルベルト空間になり, 内積は次の式で表される.

$$\langle f, g \rangle = \int_{\Omega} f(t) \overline{g(t)} dt \quad (f, g \in L^2(\Omega))$$

3.6 フーリエ級数展開

$\mathbb{T} = [-\pi, \pi]$ とする.

定理 3.6.1 (リース・フィッシャーの定理) $L^2(\mathbb{T})$ 上の関数列 $\{\phi_n\}_{n \in \mathbb{Z}}$ を

$$\phi_n(t) = \frac{e^{int}}{\sqrt{2\pi}}$$

で定義する. このとき $\{\phi_n\}_{n \in \mathbb{Z}}$ は完全正規直交系である. これを **リース・フィッシャーの定理** (Riesz–Fischer theorem) という.

リース・フィッシャーの定理によれば

$$\hat{f}_n = \frac{\langle f, \phi_n \rangle}{\sqrt{2\pi}} = \frac{1}{2\pi} \int_{\mathbb{T}} f(t) e^{-int} dt \quad (n \in \mathbb{Z})$$

とおくと, \mathbb{T} 上の関数列 $S_N(t) = \sum_{n=-N}^N \hat{f}_n e^{int}$ は f に L^2 収束する. このことを形式的に

$$f(t) = \text{l.i.m.}_{N \rightarrow \infty} \sum_{n=-N}^N \hat{f}_n e^{int} \quad (t \in \mathbb{T})$$

と表そう²⁾.

定義 3.6.2 (フーリエ級数展開) 各 $f \in L^2(\mathbb{T})$ に対して, 次式を f の **フーリエ級数展開** (Fourier series expansion) という.

$$f(t) = \text{l.i.m.}_{K \rightarrow \infty} \sum_{k=-K}^K \hat{f}_k e^{ikt} \quad (t \in \mathbb{T})$$

ノート 定義 3.6.2 で「形式的に」と断りを入れたのは, $f(t) = \text{l.i.m.}_{N \rightarrow \infty} S_N(t)$ ($t \in \mathbb{T}$) であっても, 各 $a \in \mathbb{T}$ に対して数列 $\{S_N(a)\}$ が $f(a)$ に収束するとは言えないからである. あくまで l.i.m. は, 関数列の L^2 収束

$$f(t) = \text{l.i.m.}_{N \rightarrow \infty} S_N(t) \quad (t \in \mathbb{T}) \iff \lim_{N \rightarrow \infty} \int_{\mathbb{T}} |f(t) - S_N(t)|^2 dt = 0$$

2) l.i.m. は limit in the mean (平均収束) の略.

で定義される.

定理 3.6.3 (カールソンの定理) $f \in L^2(\mathbb{T})$ とする. このとき

$$\lim_{K \rightarrow \infty} \sum_{k=-K}^K \hat{f}_k e^{ikt} = f(t) \text{ a.e. } t \in \mathbb{T}$$

である. これを**カールソンの定理** (Carleson's theorem) という.

3.7 多重解像度解析

定義 3.7.1 (多重解像度解析) $L^2(\mathbb{R})$ の閉部分空間の列 $\{V_n\}_{n \in \mathbb{Z}}$ が以下の条件を満たすとき, $\{V_n\}_{n \in \mathbb{Z}}$ は**多重解像度解析** (multiresolution analysis; MRA) をなすという.

1. $\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots$
2. $\bigcap_{n \in \mathbb{Z}} V_n = \{0\}$, $\text{cl}(\bigcup_{n \in \mathbb{Z}} V_n) = L^2(\mathbb{R})$
3. $f(\cdot) \in V_n \iff f(2\cdot) \in V_{n+1}$, ただし n は任意の整数.
4. $\{\phi(\cdot - n)\}_{n \in \mathbb{Z}}$ が V_0 の完全正規直交系となる $\phi \in V_0$ が存在する.

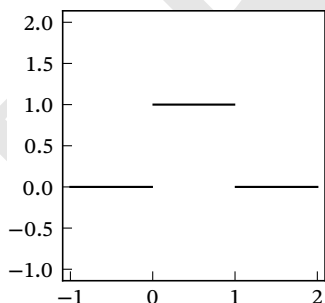


図 3.6 Haar のスケーリング関数.

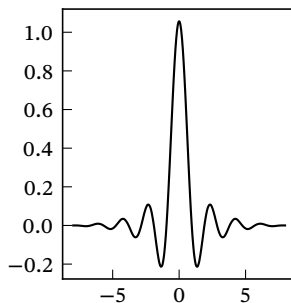


図 3.7 Meyer のスケーリング関数.

3.8 演習問題

確率空間

4.1 イントロダクション

4.2 確率空間

4.3 ウィーナー解

4.4 カルマンフィルタ

4.5 演習問題

付録 A

測度空間

A.1 イントロダクション

第 3 章, 第 4 章では, 関数空間での直交射影を積分に基づいて論じた. 関数空間でも直交射影を定義できる理由は, $L^2(\Omega) = \{f | \int_{\Omega} |f(x)|^2 dx < +\infty\}$ がヒルベルト空間, つまり, 完備な内積空間だからである. そして, $L^2(\Omega)$ の完備性には, 積分がリーマン積分ではなくルベーグ積分であることが効いている.

また, 確率変数の期待値はルベーグ積分で定義される. これは, ルベーグ積分を使うと関数の定義域をより自由に選べることの恩恵である. まとめると, ルベーグ積分はリーマン積分に次の 2 点で優る.

1. 積分と極限の相性がよい.
2. 関数の定義域が \mathbb{R}^n の部分集合でなくともよい.

これらの強みは主に, ルベーグ積分では面積の測り方 (測度) を自由に決められること, そして, 測度の観点から区別できないものは同一視できることに由来する. そのため, ルベーグ積分の利点を享受するには

1. 面積が定義される集合のあつまり (**σ -加法族**) を用意する.
2. 各集合に面積を割り当てる写像 (**測度**) を用意する.
3. 積分を定義できる関数 (**可測関数**) を規定する.

という, 3 つの段階を要する. 本章では, この 3 段階をどう踏んでいけばよいのかを, かけ足で概観する.

A.2 測度論の基本概念

A.2.1 σ -加法族

集合 S の部分集合全体を, S の**べき集合** (powerset) という. 以後, S のべき集合を 2^S と表す.

定義 A.2.1 (σ -加法族) Ω を集合とする. 2^Ω の部分集合 \mathcal{F} が以下の条件を満たすとき, \mathcal{F} は Ω 上の **σ -加法族** (σ -algebra) であるという.

1. $\Omega \in \mathcal{F}$ である.
2. 任意の $A \in \mathcal{F}$ に対して $A^c = \Omega \setminus A \in \mathcal{F}$ である.
3. 任意の $A_1, A_2, \dots \in \mathcal{F}$ に対して $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$ である.

組 (Ω, \mathcal{F}) を**可測空間** (measurable space) という.

例 A.2.2 $\{\emptyset, \Omega\}$ と 2^Ω は Ω 上の σ -加法族である. ◇

例 A.2.3 $\Omega = \{\square, \square, \boxplus, \boxtimes, \boxminus, \boxdiv\}$ を 6 元集合とし, $O = \{\square, \boxplus, \boxtimes\}$ とおく. このとき, $\mathcal{G} = \{\emptyset, O, O^c, \Omega\}$ は Ω 上の σ -加法族である. ◇

定義 A.2.4 (生成する σ -加法族) Ω を集合, \mathcal{S} を 2^Ω の部分集合とする. また, \mathcal{S} を包含する Ω 上の σ -加法族全体を $\Sigma(\mathcal{S})$ とおく. このとき, 集合

$$\sigma(\mathcal{S}) = \bigcap_{\mathcal{F} \in \Sigma(\mathcal{S})} \mathcal{F} = \{A \in 2^\Omega \mid \Sigma(\mathcal{S}) \text{ の元すべてに } A \text{ は属する}\}$$

は $\Sigma(\mathcal{S})$ の元である. $\sigma(\mathcal{S})$ を \mathcal{S} が**生成する σ -加法族** (generated σ -algebra) という.

例 A.2.5 (ボレル集合族) 集合 $\mathcal{S} = \{(-\infty, a] \mid a \in \mathbb{R}\}$ により生成される σ -加法族を \mathbb{R} 上の**ボレル集合族** (Borel algebra) といい, $\mathcal{B}(\mathbb{R})$ と表記する. ◇

ノート $\mathcal{B}(\mathbb{R})$ は \mathbb{R} の開集合系が生成する σ -加法族でもある. 実はより一般に, 位相空間の開集合系が生成する σ -加法族のことをボレル集合族という.

例 A.2.6 例 A.2.3 の σ -加法族 \mathcal{G} は $\mathcal{G} = \sigma(\{O\})$ と書ける. 実際, \mathcal{F} が $\{O\}$ を包含する σ -加法族なら $O^c \in \mathcal{F}$, $\mathcal{G} \subseteq \mathcal{F}$ である. つまり, \mathcal{G} は $\{O\}$ を包含する最小の σ -加法族だから, $\mathcal{G} = \sigma(\{O\})$ である. \diamond

A.2.2 ボレル測度とルベーグ測度

定義 A.2.7 (拡大実数) \mathbb{R} に正負の無限大 $+\infty, -\infty \notin \mathbb{R}$ を加えた集合 $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ を **拡大実数** (extended real number) という. 各 $a \in \mathbb{R}$, $x \in \bar{\mathbb{R}}$ に対し, $\pm\infty$ との演算を以下の通り定義する (複合同順).

$$\begin{aligned} a + (\pm\infty) &= (\pm\infty) + a = \pm\infty, & a - (\pm\infty) &= \mp\infty, \\ (\pm\infty) + (\pm\infty) &= \pm\infty, & (\pm\infty) - (\mp\infty) &= \pm\infty, \\ \frac{a}{\pm\infty} &= 0, & x \cdot (\pm\infty) &= (\pm\infty) \cdot x = \begin{cases} \pm\infty & (x > 0), \\ 0 & (x = 0), \\ \mp\infty & (x < 0) \end{cases} \end{aligned}$$

ノート $(\pm\infty) + (\mp\infty)$ や $(\pm\infty) - (\pm\infty)$ などは定義されない. 要するに, $\bar{\mathbb{R}}$ 上では不定形でない式だけが計算できる. また, 普通 $0 \cdot (\pm\infty)$ は不定形だが, ここでは $0 \cdot (\pm\infty) = 0$ と定義した.

$\{A_n\}_{n \in \mathbb{N}}$ を集合列とする. $\{A_n\}_{n \in \mathbb{N}}$ が**互いに素** (disjoint) であるとは, 任意の $i, j \in \mathbb{N}$ に対して $(i \neq j \implies A_i \cap A_j = \emptyset)$ であることをいう. 互いに素な集合の和であることを強調したいときは, 和集合 $\bigcup_n A_n$ を $\bigsqcup_n A_n$ とも書く.

定義 A.2.8 (測度) \mathcal{F} を σ -加法族とする. 写像 $\mu: \mathcal{F} \rightarrow [0, +\infty]$ が \mathcal{F} 上の**測度** (measure) であるとは, μ が以下の条件を満たすことをいう.

1. $\mu(\emptyset) = 0$ である.
2. \mathcal{F} 上の集合列 $\{A_n\}_{n \in \mathbb{N}}$ が互いに素ならば $\mu(\bigsqcup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ である.

3つ組 $(\Omega, \mathcal{F}, \mu)$ を**測度空間** (measure space) という. 特に $\mu(\Omega) = 1$ であるとき, μ を**確率測度** (probability measure), $(\Omega, \mathcal{F}, \mu)$ を**確率空間** (probability space) という. また, 確率空間の Ω に属する各元は**根源事象** (elementary event), \mathcal{F} に属する各元は**事象** (event) と呼ばれる.

例 A.2.9 (計数測度) \mathcal{F} を σ -加法族とする. 各 $A \in \mathcal{F}$ に対して, A が有限集合なら $\mu(A) = \#A$, 無限集合なら $\mu(A) = +\infty$ とすると, μ は \mathcal{F} 上の測度になる. μ を **計数測度** (counting measure) という. \diamond

例 A.2.10 (ディラック測度) (Ω, \mathcal{F}) を可測空間とし, $x \in \Omega$ を 1 つ選ぶ. 各 $A \in \mathcal{F}$ に対して, $x \in A$ なら $\delta_x(A) = 1$, $x \notin A$ なら $\delta_x(A) = 0$ とすると, δ_x は \mathcal{F} 上の測度になる. δ_x を **ディラック測度** (Dirac measure) という. \diamond

例 A.2.11 Ω を例 A.2.3 と同じにし, $\mathcal{F} = 2^\Omega$ とする. このとき, 写像 $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$, $\mathbb{P}(A) = (1/6)\#A$ は \mathcal{F} 上の確率測度である. \mathbb{P} は「どの目が出るのも同様に確からしい (公平な) 6 面ダイスに関する確率」を表すと解釈できる. たとえば, 奇数の目が出る確率は $\mathbb{P}(O) = 1/2$ である. \diamond

例 A.2.11 のように Ω が有限集合のときは, 各 $\omega \in \Omega$ に対して $\mu(\{\omega\})$ の値を決定することで, 測度 μ を直に構成できる. しかし, たとえば $\mathcal{B}(\mathbb{R})$ 上の測度を定める場合, この方法は使えない. そうした状況では, σ -加法族よりも考えやすい集合族の上で測度の雛形を作り, それを σ -加法族全体へと拡張する.

定義 A.2.12 (有限加法族) Ω を集合とする. 2^Ω の部分集合 \mathcal{A} が以下の条件を満たすとき, \mathcal{A} を Ω 上の **有限加法族** (finitely additive class) であるという.

1. $\Omega \in \mathcal{A}$ である.
2. 任意の $A \in \mathcal{A}$ に対して $A^c = \Omega \setminus A \in \mathcal{A}$ である.
3. 任意の $A, B \in \mathcal{A}$ に対して $A \cup B \in \mathcal{A}$ である.

定義 A.2.13 (有限加法的測度) \mathcal{A} を有限加法族とする. 写像 $m: \mathcal{A} \rightarrow [0, +\infty]$ が \mathcal{A} 上の **有限加法的測度** (finitely additive measure) であるとは, m が以下の条件を満たすことをいう.

1. $m(\emptyset) = 0$ である.
2. $A, B \in \mathcal{A}$ が $A \cap B = \emptyset$ を満たすとき $m(A \sqcup B) = m(A) + m(B)$ である.

ノート infinite は「インフィニット」と読むが, finite は「ファイナイト」と読む.

定理 A.2.14 (ホップの拡張定理) \mathcal{A} を集合 Ω 上の有限加法族, m を \mathcal{A} 上の有限加法的測度とする. このとき, 以下の命題は同値である.

1. \mathcal{A} 上の集合列 $\{A_n\}_{n \in \mathbb{N}}$ が互いに素で $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ を満たすとき,
 $m(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} m(A_n)$ である.
2. $\sigma(\mathcal{A})$ 上の測度 μ で, 任意の $A \in \mathcal{A}$ に対して $\mu(A) = m(A)$ を満たすものが存在する.

さらに, \mathcal{A} 上の集合列 $\{A_n\}_{n \in \mathbb{N}}$ で $m(A_n) < +\infty$, $\bigcup_{n \in \mathbb{N}} A_n = \Omega$ を満たすものが存在するとき, μ は一意である. これを**ホップの拡張定理** (Hopf extension theorem) という.

略証 各 $S \in 2^\Omega$ に対し, \mathcal{A} 上の集合列 $\{A_n\}_{n \in \mathbb{N}}$ で $S \subseteq \bigcup_{n \in \mathbb{N}} A_n$ を満たすものの全体を $\text{cover}_{\mathcal{A}} S$ とおく. そして, 写像 $\mu^*: 2^\Omega \rightarrow [0, +\infty]$ を

$$\mu^*(S) = \inf \left\{ \sum_{n=1}^{\infty} m(A_n) \mid \{A_n\}_{n \in \mathbb{N}} \in \text{cover}_{\mathcal{A}} S \right\}$$

で定義する. すると, 集合

$$\{A \in 2^\Omega \mid \text{任意の } E \in 2^\Omega \text{ に対し } \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)\} \quad (\text{A.1})$$

は完全加法族であり, \mathcal{A} を包含する. この集合を \mathcal{F} とおくと, μ^* の始域を \mathcal{F} へと制限した写像 $\bar{\mu}$ は, \mathcal{F} 上の測度であることが示せる. したがって, $\bar{\mu}$ の $\sigma(\mathcal{A})$ への制限 μ は $\sigma(\mathcal{A})$ 上の測度である. \square

ノート \mathcal{A} 上の集合列 $\{A_n\}_{n \in \mathbb{N}}$ で $m(A_n) < +\infty$, $\bigcup_{n \in \mathbb{N}} A_n = \Omega$ を満たすものが存在するとき, m は **σ -有限** (σ -finite) であるという. 本書が扱う (有限加法的) 測度はすべて σ -有限なので, 定理 A.2.14 から定まる拡張された測度は常に一意である. また, 集合 $A \subseteq \Omega$ が式 (A.1) の \mathcal{F} に属する——つまり, 任意の $E \in 2^\Omega$ に対して $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)$ である——とき, A は μ^* に関して**カラテオドリ可測** (Carathéodory-measurable) であるという.

\mathcal{E} を左半開区間 $(a, b] \cap \mathbb{R}$ ($-\infty \leq a < b \leq +\infty$) の全体集合とすると, 集合 $\mathcal{A} = \{\emptyset\} \cup \{\bigcup_{k=1}^n I_k \mid I_1, \dots, I_n \in \mathcal{E} \text{ は互いに素}\}$ は有限加法族をなし, $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R})$ が成り立つ. \mathcal{A} 上の有限加法的測度 vol を

$$\text{vol} \left(\bigcup_{k=1}^n (a_k, b_k] \right) = \sum_{k=1}^n (b_k - a_k)$$

で定義する. このとき vol は σ -有限かつ**ホップの拡張定理**の条件を満たすので, $\mathcal{B}(\mathbb{R})$ 上の測度 μ へと一意に拡張できる. この $\mathcal{B}(\mathbb{R})$ 上の測度 μ を**ボレル測度** (Borel measure) という. また, $m = \text{vol}$ のときの $\bar{\mu}$ を**ルベグ測度** (Lebesgue measure) という.

A.3 ルベグ積分

以下, $(\Omega, \mathcal{F}, \mu)$ を測度空間とする.

A.3.1 ルベグ積分

定義 A.3.1 (可測関数) 関数 $f: \Omega \rightarrow \mathbb{R}$ が**可測関数** (\mathcal{F} -可測, measurable function) であるとは, 任意の $A \in \mathcal{B}(\mathbb{R})$ に対して $f^{-1}[A] \in \mathcal{F}$ が成立することをいう.

特に, 確率空間における可測関数は (実数値) **確率変数** (random variable) とも呼ばれる.

命題 A.3.2 集合 $\mathcal{S} \subseteq 2^{\Omega}$ が $\sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R})$ を満たすとき, 任意の $A \in \mathcal{S}$ に対して $f^{-1}[A] \in \mathcal{F}$ が成立すれば, f は可測関数である. 特に, 任意の $a \in \mathbb{R}$ に対して $f^{-1}[(a, \infty)] \in \mathcal{F}$ なら, f は可測関数である.

証明 集合 $\mathcal{G} = \{A \in 2^{\Omega} \mid f^{-1}[A] \in \mathcal{F}\}$ は σ -加法族であり, 仮定から $\mathcal{S} \subseteq \mathcal{G}$ なので $\sigma(\mathcal{S}) \subseteq \mathcal{G}$ である. よって, f は可測関数である. \square

定義 A.3.3 (単関数) $a_1, \dots, a_n \in \mathbb{R}$ と互いに素な $A_1, \dots, A_n \in \mathcal{F}$ により, 次の形で表せる関数 $\phi: \Omega \rightarrow \mathbb{R}$ を**単関数** (simple function) という.

$$\phi(x) = \sum_{k=1}^n a_k \mathbb{1}_{A_k}(x), \quad \mathbb{1}_{A_k}(x) = \begin{cases} 1 & (x \in A_k), \\ 0 & (x \notin A_k) \end{cases}$$

非負値単関数 $\phi = \sum_{k=1}^n a_k \mathbb{1}_{A_k}$ については, 積分 $\int \phi d\mu = \int \phi(x) \mu(dx)$ を

$$\int \phi d\mu = \sum_{k=1}^n a_k \mu(A_k)$$

で定義できる. このことを利用して, 非負値可測関数 f の積分を

$$\int f \, d\mu = \sup \left\{ \int \phi \, d\mu \mid \phi \in \text{SF}(\mathcal{F}), 0 \leq \phi \leq f \right\}$$

で定義する. ただし, $\text{SF}(\mathcal{F})$ は単関数の全体集合であり, $\phi \leq f$ は任意の $x \in \Omega$ に対し $\phi(x) \leq f(x)$ であることを意味する.

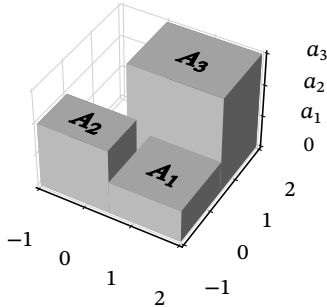


図 A.1 $\phi = a_1 \mathbb{1}_{A_1} + a_2 \mathbb{1}_{A_2} + a_3 \mathbb{1}_{A_3}$ の模式図.

より具体的に, $\int f \, d\mu$ を非負値単関数の積分に関する極限で表すこともできる. $E_{nk} = f^{-1}[[2^{-n}k, 2^{-n}(k+1)))$ とする. このとき

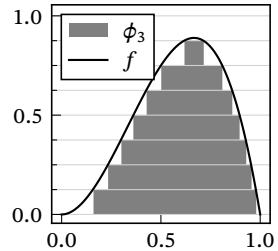
$$\phi_n = n \mathbb{1}_{f^{-1}[[n, +\infty))} + \sum_{k=0}^{2^n n - 1} \frac{k}{2^n} \mathbb{1}_{E_{nk}}$$

は非負値単関数で, $\int \phi_n \, d\mu \rightarrow \int f \, d\mu$ ($n \rightarrow \infty$) である.

f が負の値をとる可測関数のときは, $f^+(x) = \max\{f(x), 0\}$ と $f^-(x) = \max\{-f(x), 0\}$ が非負値可測関数であることを利用して

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

とする. ただし, $\int f^- \, d\mu = \int f^+ \, d\mu = +\infty$ のとき $\int f \, d\mu$ は定義されない. $\int f^- \, d\mu$ と $\int f^+ \, d\mu$ がともに有限——つまり $\int |f(x)| \mu(dx) < +\infty$ ——のとき, f は**可積分** (integrable) であるという. また, 集合 $S \in \mathcal{F}$ 上での積分は



$$\int_S f \, d\mu = \int_S f(x) \mu(dx) = \int f(x) \mathbb{1}_S(x) \mu(dx)$$

で定義される.

$(\Omega, \mathcal{F}, \mathbb{P})$ が確率空間のとき, $\int X \, d\mathbb{P}$ を確率変数 X の**期待値** (expected value) といい, $E[X]$ と書く. また, $\int_A X \, d\mathbb{P}$ を $E[X, A]$ と表記する.

例 A.3.4 例 A.2.11 の確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ において, 各 $\omega \in \Omega$ に対し $X(\omega)$ を ω の目で定義する. すなわち $X(\square) = 1$, $X(\square) = 2$ のようにする. このとき $X = 1\mathbb{1}_{\{\square\}} + 2\mathbb{1}_{\{\square\}} + \cdots + 6\mathbb{1}_{\{\square\}}$ だから

$$E[X] = 1\mathbb{P}(\{\square\}) + 2\mathbb{P}(\{\square\}) + \cdots + 6\mathbb{P}(\{\square\}) = \frac{1+2+\cdots+6}{6} = \frac{7}{2}$$

であり, これは公平な 6 面ダイスの出目の期待値に相当する. ◇

以上で定義した積分を**ルベグ積分** (Lebesgue integration) という. ルベグ積分とリーマン積分の間には, 次の関係がある.

命題 A.3.5 実数値関数 f は有界閉区間 $[a, b]$ 上で定義され有界とする. このとき, f がリーマン積分できれば f はルベグ測度に関して可積分で

$$\int_a^b f(x) \, dx = \int_{[a,b]} f(x) \lambda(dx) \quad (\lambda \text{ はルベグ測度})$$

が成立する.

ノート 積分区間が非有界なときは注意が要る. 実際, **ディリクレ積分** (Dirichlet integral) と呼ばれる広義積分

$$\int_0^\infty \frac{\sin x}{x} \, dx = \lim_{\varepsilon \downarrow 0} \int_\varepsilon^1 f(x) \, dx + \lim_{R \rightarrow \infty} \int_1^R f(x) \, dx \quad (f(x) = (1/x) \sin x)$$

の値は $\pi/2$ だが $\int f^+ \, d\lambda = \int f^- \, d\lambda = +\infty$ であり, $\int f \, d\lambda$ は定義されない.

最後に, 複素数値可測関数の積分を定義しよう. 関数 $\operatorname{Re} f(x)$, $\operatorname{Im} f(x)$ がどちらも \mathcal{F} -可測であるとき, 関数 $f: \Omega \rightarrow \mathbb{C}$ は \mathcal{F} -可測であるという. f の積分は

$$\int f \, d\mu = \int (\operatorname{Re} f(x)) \mu(dx) + i \int (\operatorname{Im} f(x)) \mu(dx)$$

で定義される.

A.3.2 収束定理

極限と積分を交換したいときは、以下の定理が非常に強力である。定理 A.3.6 から A.3.8 はすべて、任意の測度空間 $(\Omega, \mathcal{F}, \mu)$ 上で成立する。

定理 A.3.6 (単調収束定理) 可測関数列 $\{f_n\}_{n \in \mathbb{N}}$ が $0 \leq f_1 \leq f_2 \leq \dots$ を満たすとき

$$\lim_{n \rightarrow \infty} \int f_n(x) \mu(dx) = \int \left(\lim_{n \rightarrow \infty} f_n(x) \right) \mu(dx)$$

である。これを**単調収束定理** (monotone convergence theorem; MCT) という。

定理 A.3.7 (ファトゥの補題) 任意の可測関数列 $\{f_n\}_{n \in \mathbb{N}}$ に対して

$$\int \left(\liminf_{n \rightarrow \infty} f_n(x) \right) \mu(dx) \leq \liminf_{n \rightarrow \infty} \int f_n(x) \mu(dx)$$

である。これを**ファトゥの補題** (Fatou's lemma) という。

定理 A.3.8 (優収束定理) 複素数値可測関数列 $\{f_n\}_{n \in \mathbb{N}}$ が各点収束し、すべての $n \in \mathbb{N}$, $x \in \Omega$ で $|f_n(x)| \leq g(x)$ を満たす可積分関数 g があれば

$$\lim_{n \rightarrow \infty} \int f_n(x) \mu(dx) = \int \left(\lim_{n \rightarrow \infty} f_n(x) \right) \mu(dx)$$

である。これを**優収束定理** (dominated convergence theorem; DCT) という。

A.4 確率論の基本概念

本節では、確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上で定義される諸概念を見ていく。

確率論ではしばしば、実数 x に関する条件 $C(x)$ と確率変数 X に対して、集合 $\{\omega \in \Omega \mid C(X(\omega))\}$ を $\{C(X)\}$ と書く。たとえば $\{X = 1\} = X^{-1}[\{1\}]$, $\mathbb{P}(X = Y) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = Y(\omega)\})$ である。

A.4.1 確率変数が定める量

個々の確率変数の様子は、以下の関数から解析できる。

定義 A.4.1 X を確率変数とする。

1. 関数 $F_X(x) = \mathbb{P}(X \leq x)$ を X の**累積分布関数**(cumulative distribution function; CDF) という。
2. 像 $X[\Omega]$ が有限または可算集合のとき、関数 $p_X(x) = \mathbb{P}(X = x)$ を X の**確率質量関数**(probability mass function; PMF) という。
3. F_X が導関数 $f_X(x) = F'_X(x)$ を持つとき、 f_X を X の**確率密度関数**(probability density function; PDF) という¹⁾。

写像 $X\#\mathbb{P}: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ を $(X\#\mathbb{P})(A) = \mathbb{P}(X \in A)$ で定義すると、 $X\#\mathbb{P}$ は $\mathcal{B}(\mathbb{R})$ 上の確率測度になる。 $X\#\mathbb{P}$ を X の**確率分布**(probability distribution)、もしくは X による \mathbb{P} の**像測度**(image measure) という。 X が確率質量関数 p_X を持つとき、任意の $A \in \mathcal{B}(\mathbb{R})$ に対して

$$(X\#\mathbb{P})(A) = \sum_{x \in A} p_X(x)$$

が成立する。同様に、 X が確率密度関数 f_X を持つとき

$$(X\#\mathbb{P})(A) = \int_A f_X(x) \lambda(dx) \quad (\lambda \text{ はルベーグ測度})$$

が成立する。

これらの関数は、期待値を計算するとき重用される。 \mathbb{R} 上のボレル可測関数 f に関して $\mathbb{E}[f(X)] = \int f(X(\omega)) \mathbb{P}(d\omega)$ だが、この値を求めるには $X(\omega) \rightarrow x$, $\mathbb{P}(d\omega) \rightarrow f_X(x) \lambda(dx)$ と置き換えて、 $\int f(x) f_X(x) \lambda(dx)$ を計算すればよい。

A.4.2 条件つき期待値

例 A.2.3, A.2.11 において、確率変数 X の期待値は $7/2$ であった。これは、公平な 6 面ダイスをなんども投げ続けたとき、出目は平均的に $7/2$ くらいの値

1) 測度論を活用すると、さらに多くの確率変数に対して PDF を定義できる。詳しくは船木 [5] を参照せよ。

をとることを意味する。しかし、出目が奇数のとき、すなわち O に属するときだけ勘定に入れると条件づけた場合、 X の期待値は変わると予想される。直感的には、この条件の下で $\square, \boxplus, \boxtimes$ が出る確率は等しく $1/3$ であり、 X の期待値は $(1 + 3 + 5)/3 = 3$ になるだろう。

以上の考え方に基づいて、事象 A の下での事象 B の**条件つき確率** $\mathbb{P}(B|A)$ 、確率変数 X の**条件つき期待値** $\mathbb{E}[X|A]$ を、それぞれ

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}, \quad \mathbb{E}[X|A] = \frac{\mathbb{E}[X, A]}{\mathbb{P}(A)}$$

と定義する ($\mathbb{P}(A) = 0$ のときはどちらも定義しない)。

出目の偶奇に関する事象全体は、 σ -加法族 $\mathcal{G} = \{\emptyset, O, O^c, \Omega\}$ である。 \mathcal{G} -可測な確率変数 $\mathbb{E}[X|\mathcal{G}]$ を

$$\mathbb{E}[X|\mathcal{G}](\omega) = \sum_{A \in \{O, O^c\}} \mathbb{E}[X|A] \mathbb{1}_A(\omega) = \begin{cases} \mathbb{E}[X|O] & (\omega \in O), \\ \mathbb{E}[X|O^c] & (\omega \in O^c) \end{cases} \quad (\text{A.2})$$

と定める。このとき、 $\mathbb{E}[X|\mathcal{G}]$ は各 $\omega \in \Omega$ に対し、 ω が属する \mathcal{G} 上の事象で条件づけたときの、 X の条件つき期待値を与える確率変数である。確率変数 $\mathbb{E}[X|\mathcal{G}]$ を σ -加法族 \mathcal{G} の下での X の条件つき期待値という。もう少し一般に、 \mathcal{G} が Ω の有限または可算な分割 $\Omega = \bigsqcup_{n \in I} A_n$ ($A_n \in \mathcal{F}$, $\mathbb{P}(A_n) > 0$) から生成される場合は、 $\mathbb{E}[X|\mathcal{G}]$ を式 (A.2) と同様に

$$\mathbb{E}[X|\mathcal{G}](\omega) = \sum_{n \in I} \mathbb{E}[X|A_n] \mathbb{1}_{A_n}(\omega) \quad (\mathcal{G} = \sigma(\{A_n | n \in I\}))$$

で定義できる。

以下では、一般の σ -加法族 \mathcal{G} に対して通用する $\mathbb{E}[X|\mathcal{G}]$ の定義を考える。 X と Y を実数値確率変数とする。 X と Y の値にはなんらかの関係があり、 Y の値に関する条件の下で、 X の期待値が持つ性質を解析したい²⁾。

Y の値に関する事象全体は、 σ -加法族 $\mathcal{G} = \{Y^{-1}[A] | A \in \mathcal{B}(\mathbb{R})\}$ である。また、 $y \equiv n\Delta y$, $0 < \Delta y \ll 1$ なら $\mathbb{E}[X|\delta Y_n]$ ($\delta Y_n = \{(n-1)\Delta y < Y \leq n\Delta y\}$) はおおむね、条件 $Y = y$ の下での X の期待値を表すとみられる。

$\Omega = \bigsqcup_{n \in \mathbb{Z}} \delta Y_n$ であるから、各 $\mathbb{P}(\delta Y_n)$ の値が正なら $\mathbb{E}[X|\mathcal{G}_{\Delta y}]$ を

$$\mathbb{E}[X|\mathcal{G}_{\Delta y}](\omega) = \sum_{n \in \mathbb{Z}} \mathbb{E}[X|\delta Y_n] \mathbb{1}_{\delta Y_n}(\omega) \quad (\mathcal{G}_{\Delta y} = \sigma(\{\delta Y_n | n \in \mathbb{Z}\}))$$

2) たとえば、 Y は気温、 X はアイスクリームの売上と考えるとよい。

と定義できる. $\mathbb{E}[X | \mathcal{G}_{\Delta y}](\omega)$ の値は集合 $\{Y = y\}$ 上で一定であり, 条件 δY_n (つまり $Y \equiv y$) の下での X の期待値を表す. そこで, いったん $\mathbb{E}[X | \mathcal{G}]$ を

$$\mathbb{E}[X | \mathcal{G}](\omega) \stackrel{?}{=} \lim_{\Delta y \rightarrow 0} \mathbb{E}[X | \mathcal{G}_{\Delta y}](\omega) = \lim_{\Delta y \rightarrow 0} \sum_{n \in \mathbb{Z}} \mathbb{E}[X | \delta Y_n] \mathbb{1}_{\delta Y_n}(\omega)$$

と“定義”してみよう. 任意の $B \in \mathcal{B}(\mathbb{R})$ に対し, \mathbb{E} と $\lim \sum$ が交換できれば

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}], Y \in B] = \lim_{\Delta y \rightarrow 0} \sum_{n \in \mathbb{Z}} \frac{\mathbb{E}[X, \delta Y_n]}{\mathbb{P}(\delta Y_n)} \mathbb{E}[\mathbb{1}_{\delta Y_n}, Y \in B] \quad (\text{A.3})$$

$$= \lim_{\Delta y \rightarrow 0} \sum_{n \in \mathbb{Z}} \mathbb{E}[X, \delta Y_n] \frac{\mathbb{P}(\{Y \in B\} \cap \delta Y_n)}{\mathbb{P}(\delta Y_n)} \quad (\text{A.4})$$

$$= \lim_{\Delta y \rightarrow 0} \sum_{n \in \mathbb{Z}} \mathbb{E}[X, \delta Y_n] \mathbb{P}(Y \in B | \delta Y_n) \quad (\text{A.5})$$

となる. 事象 $\{Y = y\}$ の下での $\{Y \in B\}$ の条件つき確率は, $y \in B$ のとき 1, $y \notin B$ のとき 0 と考えるのが妥当だろう. つまり, 大雑把に言って式 (A.5) は, $y \in B$ となるすべての $y \in \mathbb{R}$ に関して $\mathbb{E}[X, Y = y]$ を足しているから

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}], A] \stackrel{?}{=} \mathbb{E}[X, A] \quad (A = \{Y \in B\})$$

と考えられる.

実は逆に, $\mathbb{E}[X | \mathcal{G}]$ は「任意の $A \in \mathcal{G}$ に対し $\mathbb{E}[\mathbb{E}[X | \mathcal{G}], A] = \mathbb{E}[X, A]$ を満たす」確率変数として定義される.

定義 A.4.2 (条件つき期待値) $(\Omega, \mathcal{F}, \mathbb{P})$ は確率空間で, Ω 上の σ -加法族 \mathcal{G} は $\mathcal{G} \subseteq \mathcal{F}$ を満たすとする. また, X は $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数とする. このとき, \mathcal{G} -可測な確率変数 Y で, 任意の $A \in \mathcal{G}$ に対し $\mathbb{E}[Y, A] = \mathbb{E}[X, A]$ を満たすものが存在する³⁾. Y を \mathcal{G} の下での X の**条件つき期待値** (conditional expectation) といい, $\mathbb{E}[X | \mathcal{G}]$ と書く.

定義 A.4.2 の $\mathbb{E}[X | \mathcal{G}]$ は, 測度 0 の集合上での違いを除いて一意に定まる. すなわち, \mathcal{G} -可測な確率変数 Y と Y' がともに $\mathbb{E}[X | \mathcal{G}]$ の条件を満たすとき, $\mathbb{P}(Y \neq Y') = 0$ である. このことを $Y = Y'$ a.s. と表す⁴⁾.

また, $\mathbb{1}_A$ ($A \in \mathcal{F}$) の条件つき期待値 $\mathbb{E}[\mathbb{1}_A | \mathcal{G}]$ を \mathcal{G} の下での A の**条件つき確率** (conditional probability) といい, $\mathbb{P}(A | \mathcal{G})$ と書く.

3) これはラドン・ニコディムの定理 (Radon-Nikodým theorem) から保証される.

4) a.s. は almost surely (ほとんど確実に) の略.

A.5 L^p 空間

以下, $(\Omega, \mathcal{F}, \mu)$ を測度空間, \mathbb{K} を \mathbb{R} または \mathbb{C} とする.

定義 A.5.1 (零集合) A を Ω の部分集合とする. $\mu(N) = 0$, $A \subseteq N$ を満たす集合 $N \in \mathcal{F}$ があるとき, A は μ -**零集合** (μ -null set) であるという.

定義 A.5.2 (ほとんどいたるところ)

1. $C(\omega)$ を Ω の元 ω に関する条件とする. 集合 $\{\omega \mid C(\omega)\}^c$ が μ -零集合であるとき, $C(\omega)$ は**ほとんどいたるところ** (almost everywhere; a.e.) 成立するという. このことを $C(\omega)$ μ -a.e. $\omega \in \Omega$ と書く.
2. $C(z)$ を \mathbb{K} の元 z に関する条件とする. 関数 $f: \Omega \rightarrow \mathbb{K}$ が $C(f(\omega))$ μ -a.e. $\omega \in \Omega$ を満たすとき, これを $C(f)$ μ -a.e. と表す.

ノート すべての μ -零集合が \mathcal{F} に属するとき, μ を**完備測度** (complete measure) という. $\mathcal{B}(\mathbb{R})$ 上のボレル測度は完備測度でないが, ルベグ測度は完備測度である.

$1 \leq p < +\infty$, $\|f\|_{\mathcal{L}^p} = (\int |f(x)|^p \mu(dx))^{1/p}$ とする. 第 3 章で見たように, $\|f\|_{\mathcal{L}^p} < +\infty$ を満たす可測関数 $f: \Omega \rightarrow \mathbb{K}$ からなる空間は, 応用的にも非常に重要である. この空間を $\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ としよう. **ミンコフスキーの不等式** (Minkowski inequality) $\|f + g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}$ から, \mathcal{L}^p はベクトル空間である. また, $\|\cdot\|_{\mathcal{L}^p}$ はノルムの条件のうち $\|af\|_{\mathcal{L}^p} = |a|\|f\|_{\mathcal{L}^p}$, $\|f + g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}$ を満たす.

しかし厄介なことに, 多くの場合 $\|f\|_{\mathcal{L}^p} = 0 \iff f = 0$ は成立しない. 一方で, 次の事実がある.

命題 A.5.3 f を可測関数とする. このとき, f に関する以下の条件は同値である.

1. $\int |f(x)| \mu(dx) = 0$ である.
2. $f = 0$ μ -a.e. である.

命題 A.5.3 から, $\|f\|_{\mathcal{L}^p} = 0$ となるのは $|f(x)|^p = 0$ a.e. x , すなわち $f = 0$ a.e. のときに限られる. そのため, ほとんどいたるところで等しい値をとる関数を同じものとみなせば, \mathcal{L}^p はノルム空間になる. つまり, f と同一視される可測関数全体を $[f] = \{\varphi \mid \varphi = f \text{ a.e.}\}$ とおくと, 集合 $IP = \{[f] \mid f \in \mathcal{L}^p\}$ はノルム空間である.

実は, IP はバナッハ空間でもある. バナッハ空間 IP を IP 空間という.

定義 A.5.4 (IP 空間) 任意の可測関数 f について, 集合 $[f]$ を $[f] = \{\varphi \mid \varphi \text{ は可測関数で } \varphi = f \text{ a.e.}\}$ とする. このとき, 各 $p \in [1, +\infty)$ に対して集合 $IP = \{[f] \mid \int |f(x)|^p \mu(dx) < +\infty\}$ は, 加法 $[f] + [g] = [f + g]$, スカラー乗法 $a[f] = [af]$ に関するベクトル空間である. また, ノルム $\| \cdot \|_p$ を

$$\|[f]\|_p = \left(\int |f(x)|^p \mu(dx) \right)^{1/p}$$

で定義すると, IP はバナッハ空間になる. バナッハ空間 $IP = IP(\Omega, \mathcal{F}, \mu)$ を IP 空間 (IP space) という.

ノート IP の L は数学者アンリ・ルベグ (Lebesgue, Henri L., 1875–1941) にちなむ. なお, 細かいことを言うと定義 A.5.4 はいわゆる「well-definedness」を検証すべき定義だが, 気にしないことにする.

通常は \mathcal{L}^p と IP の違いをあまり意識せず, $IP(\Omega, \mathcal{F}, \mu)$ を関数からなる集合とみなして, $\|[f]\|_p$ を単に $\|f\|_p$ と書く.

例 A.5.5 $\ell^p(\mathbb{N})$ は計数測度 $\mu: 2^{\mathbb{N}} \rightarrow [0, +\infty]$ に関する IP 空間である. 実際, 関数 $f: \mathbb{N} \rightarrow \mathbb{R}$ が \mathcal{F} -可測なら, 関数 $\phi_n^+ = \sum_{k=1}^n f^+(k) \mathbb{1}_{\{k\}}$ は非負値単関数であり, 単調収束定理から

$$\int f^+ d\mu = \lim_{n \rightarrow \infty} \int \phi_n^+ d\mu = \lim_{n \rightarrow \infty} \sum_{k=1}^n f^+(k) \mu(\{k\}) = \sum_{k=1}^{\infty} f^+(k)$$

となる. f^- についても同じように考えれば

$$\int f d\mu = \left(\sum_{n=1}^{\infty} f^+(n) \right) - \left(\sum_{n=1}^{\infty} f^-(n) \right) = \sum_{n=1}^{\infty} f(n)$$

である. よって $\|f\|_p = (\sum |f(n)|^p)^{1/p}$, $\ell^p(\mathbb{N}) = IP(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ である. \diamond

命題 A.5.6 $p = 2$ のときのみ $L^p(\Omega, \mathcal{F}, \mu)$ はヒルベルト空間になり、内積は次の式で表される.

$$\langle f, g \rangle = \int f(x) \overline{g(x)} \mu(dx)$$

命題 A.5.7 $(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間, $\mathcal{G} \subseteq \mathcal{F}$ を Ω 上の σ -加法族とする. このとき, $\mathbb{E}[|X|^2] < +\infty$ を満たす任意の確率変数 X に対して, 次の式が成立する.

$$\mathbb{E}[X | \mathcal{G}] = \text{proj}_V X \text{ a.s. } (V = L^2(\Omega, \mathcal{G}, \mathbb{P}))$$

演習問題の略解

第 2 章

1. $A = O$ のときは $\Sigma = O$ とすればよい. $A \neq O$, $n < p$ のとき, A^H の特異値分解を $A^H = V\Sigma U^H$ とおくと $A = U\Sigma^H V^H$ である. Σ^H をあらためて Σ とすれば $A = U\Sigma V^H$ であり, これは A の特異値分解である.
2. 左側を示す. 命題 2.3.8 より, 任意の $n \in \text{nul } A$ に対して $\langle (I - A^+A)x - x, n \rangle = -\langle x, (A^+A)^H n \rangle = -\langle x, A^+An \rangle = 0$ である. よって $\text{proj}_{\text{nul } A} x = (I - A^+A)x$ である. 同様に, 任意の $Aw \in \text{col } A$ に対して $\langle AA^+y - y, Aw \rangle = \langle A^H(AA^+)^H y - A^H y, w \rangle = \langle ((AA^+A)^H - A^H)y, w \rangle = 0$ である. したがって $\text{proj}_{\text{col } A} y = AA^+y$ である.
3. $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ とおくと, $A^k = (U\Lambda U^H)^k = U\Lambda^k U^H$ より

$$\exp A = \sum \frac{A^k}{k!} = \sum \frac{U\Lambda^k U^H}{k!} = U \left(\sum \frac{\text{diag}(\lambda_1^k, \dots, \lambda_n^k)}{k!} \right) U^H$$

であり, 右辺は $U \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) U^H$ に等しい.

4. 命題 2.2.16 と命題 2.4.2 からしたがう.

参考文献

- [1] 新井仁之. ウェーブレット. 共立出版, 2010, 463p., (共立叢書 現代数学の潮流, 10).
- [2] カノン. “波音リツ音源配布所”. カノンの落ちる城. <http://www.canon-voice.com/index.html>, (参照 2022-11-25).
- [3] Casazza, Peter G. et al. *Finite Frames: Theory and Applications*. Birkhäuser Boston, 2013, 485p., (online), available from SpringerLink, (accessed 2022-08-09).
- [4] Donoho, David L.; Stark, Philip B. Uncertainty Principles and Signal Recovery. *SIAM Journal on Applied Mathematics*. 1989, vol. 49, no. 3, p. 906-931. <https://epubs.siam.org/doi/10.1137/0149053>, (accessed 2023-03-20).
- [5] 舟木直久. 確率論. 朝倉書店, 2022, 261p., (講座 数学の考え方, 20).
- [6] 柏野牧夫. “ピッチと基本周波数はどう違うのですか.”. 日本音響学会. <https://acoustics.jp/qanda/answer/101.html>, (参照 2022-12-18).
- [7] 黒田成俊. 関数解析. 共立出版, 2021, 339p., (共立数学講座, 15).
- [8] Luenberger, David G. *Optimization by Vector Space Methods*. Wiley, 1969, 326p.
- [9] 松坂和夫. 集合・位相入門. 岩波書店, 2018, 329p., (松坂和夫 数学入門シリーズ, 1).
- [10] 森勢将雅. 日本音響学会編. 音声分析合成. コロナ社, 2018, 272p., (音響テクノロジーシリーズ, 22).
- [11] 齋藤正彦. 線型代数入門. 東京大学出版会, 2020, 274p., (基礎数学, 1).
- [12] 杉浦光夫. 解析入門 I. 東京大学出版会, 2018, 442p., (基礎数学, 2).
- [13] Yanai, Haruo. et al. *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*. Springer New York, 2011, 243p., (online), available from SpringerLink, (accessed 2022-08-22).
- [14] 矢田部浩平ほか. 小特集, 位相情報を考慮した音声音響信号処理: 位相変換による複素スペクトログラムの表現. 日本音響学会誌. 2019, vol. 75, no. 3, p. 147-155. <https://www.jstage.jst.go.jp/article/jasj/>

- 75/3/75_147/_article/-char/ja/, (参照 2023-03-15).
- [15] 矢田部浩平. 第三回: 短時間フーリエ変換. 日本音響学会誌. 2021, vol. 77, no. 6, p. 396-403. https://www.jstage.jst.go.jp/article/jasj/77/6/77_396/_article/-char/ja/, (参照 2023-03-17).
- [16] 雪江明彦. 環と体とガロア理論. 日本評論社, 2019, 300p., (代数学, 2).

索引

【記号】		【D】		【M】	
$V_1 + V_2$	2	DCT	→ 優収束定理	MCT	→ 単調収束定理
$V_1 \oplus V_2$	3	DFT	→ 離散フーリエ変換	MRA	→ 多重解像度解析
$\langle _, _ \rangle$	5	DGT	→ 離散ガボール変換	【N】	
$f[S]$	8			$\text{nul } \mathbf{A}$	27
$f^{-1}[S]$	8			【O】	
$\ _ \ $	14, 51	$\text{diag}(a_1, \dots, a_n)$	9	ONB	→ 正規直交基底
\emptyset	15	$\dim V$	4	ONS	→ 正規直交系
$W^{\perp V}$	17	Donoho-Stark の不確定性原理	43	【P】	
V^{\perp}	17	【E】		$\mathbb{P}(A \mathcal{G})$	74
\mathbf{A}^+	25	$E_{\lambda}(\mathbf{A})$	8	$X \# \mathbb{P}$	72
Y^X	31	$\mathbb{E}[X \mathcal{G}]$	74	PCA	→ 主成分分析
$\mathbb{K}^{\mathbb{Z}/N}$	36	$\mathbb{E}[X, A]$	70	PDF	→ 確率密度関数
III	36	$\mathbb{E}[X]$	70	PMF	→ 確率質量関数
$x * y$	38	【F】		proj_V	17, 55
$\mathbb{K}^{\mathbb{Z}^2}$	41	$\mathcal{F}_{\mathbb{Z}/N}$	31, 36	【R】	
$\mathbb{K}^{\mathbb{Z}^2/(M,N)}$	41	【G】		\mathbb{R}	65
2^S	64	\mathcal{G}_w	43	RSS	→ 残差平方和
$\sigma(S)$	64	【H】		【S】	
$A \sqcup B$	65	\mathbf{A}^H	18	$\text{span } S$	3
$\mathbb{1}_A$	68	【I】		$\text{spec } \mathbf{A}$	8
【A】		$\text{im } f$	8	STFT	→ 短時間フーリエ変換
a.e.	75	$\inf S$	10	$\sup S$	10
$\arg \max$	14	【K】		SVD	→ 特異値分解
$\arg \min$	14	$\ker f$	8	【T】	
a.s.	74	【L】		\mathbf{A}^T	2
【B】		l.i.m.	58	【V】	
$\mathcal{B}(\mathbb{R})$	64	\liminf	11	\mathcal{V}_w	42
【C】		\limsup	11	【あ】	
S^c	50	ℓ^p 空間	51	アフィン部分空間	27
CDF	→ 累積分布関数	L^p 空間	57, 76		
$\text{cl } S$	50				
$\text{col } \mathbf{A}$	22				
CONS	→ 完全正規直交系				

【い】		直交——	28	収束, 点列の	48
一様収束	49	表現——	7	ε -集中	43
【う】		ユニタリ——	20	主成分分析	45
上三角行列	44	極限点	48	巡回行列	39
【え】		距離	48	巡回畳み込み	38
エイリアシング	34	距離空間	48	上界	10
エルミート転置	18	完備——	50	上極限	11
【か】		【く】		上限	10
カールソンの定理	59	グラム・シュミットの直		条件つき確率	74
下界	10	交化法	6	条件つき期待値	74
下極限	11	【け】		信号	
核	8	計数測度	65	離散——	31
拡大実数	65	【こ】		連続——	31
確率空間	65	コーシー・シュワルツの		【す】	
確率質量関数	72	不等式	52	随伴行列 → エルミート	
確率測度	65	コーシー列	12, 50	転置	
確率分布	72	合成作用素	19	数列空間	31
確率変数	68	固有空間	8	スカラー	1
確率密度関数	72	固有多項式	9	スペクトル定理	23
下限	10	固有値	8	実行列	28
可積分	69	固有ベクトル	8	スペクトル分解	23
可測関数	68	根源事象	65	スペクトル包絡	33
可測空間	64	【さ】		【せ】	
加法逆元	2	最小 2 乗法	21	正規行列	20
カラテオドリ可測	67	残差	21	正規直交基底	5
完全正規直交系	57	残差平方和	30	正規直交系	5, 55
完備距離空間	50	サンプリング → 標本化		完全——	57
完備測度	75	【し】		正規方程式	23
【き】		時間周波数解析	41	生成	
擬似逆行列	25	σ -加法族	64	σ -加法族	64
期待値	70	生成する——	64	部分空間	3
条件つき——	74	σ -有限	67	零集合	75
基底	4	次元	4	零ベクトル	2
基本周波数	33	指示関数	68	線型結合	2
逆像	8	事象	65	線型写像	6
行列		射影定理	55	線型従属	4
上三角——	44	凸——	54	線型独立	4
巡回——	39	シュア分解	44	線分	53
正規——	20			【そ】	
				像	8

像測度	72
測度	65
確率——	65
完備——	75
計数——	65
像——	72
ディラック——	66
ボレル——	68
有限加法的——	66
ルベグ——	68
測度空間	65
【た】	
対角化	9
互いに素	65
多重解像度解析	59
畳み込み	
巡回	38
単関数	68
短時間フーリエ変換	42
単射	6
単調収束定理	71
【ち】	
中線定理	52
稠密	50
直和, 部分空間の	3
直交	5
直交行列	28
直交射影	17, 55
直交補空間	17
【て】	
ディラック測度	66
ディクレ積分	70
【と】	
特異値	24
特異値分解	24
実行列	28
特徴づける	20
凸射影定理	54
凸集合	53

【な】	
ナイキスト周波数	34
内積	5
内積空間	5
【の】	
ノルム	14, 51
ノルム空間	51
【は】	
パーセヴァルの定理	46
バナッハ空間	51
【ひ】	
ピッチ	33
表現行列	7
標準基底	4
標準内積	5
標本化	34
標本化周波数	34
ヒルベルト空間	52
【ふ】	
フーリエ級数展開	36, 58
ファトゥの補題	71
不確定性原理	
Donoho-Stark	43
部分空間	2
アフィン——	27
生成する——	3
直和	3
和	2
プランシュレルの定理	46
分解作用素	19
【へ】	
閉集合	50
閉包	50
べき集合	64
ベクトル	1
ベクトル空間	1
ベッセルの不等式	55

【ほ】	
補間多項式	29
ホップの拡張定理	67
ほとんどいたるところ	75
ほとんど確実に	74
ボレル集合族	64
ボレル測度	68
【ま】	
窓関数	42
【み】	
ミンコフスキーの不等式	75
【む】	
ムーア・ペンローズ逆行列	25
【ゆ】	
有界	10
有限加法族	66
有限加法的測度	66
有限次元	4
優収束定理	71
ユニタリ行列	20
【ら】	
ラグ作用素	36
ラドン・ニコディムの定理	74
【り】	
リース・フィッシャーの定理	58
離散ガボール変換	43
離散距離	48
離散空間	48
離散信号	31
離散フーリエ変換	
数ベクトル	31
周期数列	36

【る】	ルベーク測度	68	連続信号	31
累積分布関数	【れ】		【わ】	
ルベーク積分	列空間	22	和, 部分空間の	2