

メモ

大上由人

2025 年 3 月 4 日

目次

第 I 部	数学	3
1	位相空間	3
1.1	定義など	3
1.2	その他	7
2	線形代数	8
2.1	行列の積	8
2.2	行列式	9
2.3	ベクトル空間	18
2.4	固有値・固有ベクトル	32
2.5	対角化	34
2.6	ランク	40
2.7	その他	41
3	微分積分学	41
4	確率統計	41
4.1	確率分布のまとめ	41
4.2	漸近まわりの話	44
5	幾何学	49
6	ラプラス変換	49
7	その他	51
7.1	計算の類	51

7.2	ルジャンドル変換	52
7.3	ベクトル解析	53
7.4	参考にしたもの	61
第 II 部 物理		61
8	古典力学	61
9	電磁気学	61
10	熱力学	61
10.1	その他	64
11	統計力学	66
11.1	状態数	66
11.2	典型性	70
11.3	ミクロカノニカル分布	70
11.4	カノニカル分布	75
11.5	グランドカノニカル分布	83
11.6	モデルの等価性	83
11.7	量子統計	85
11.8	その他	86
11.9	母関数との対応	88
11.10	分布の他の導出方法	92
11.11	典型性に関する追記	94
12	非平衡統計力学	95
12.1	必要な数学	95
12.2	熱力学量の定義	98
12.3	ゆらぎの定理	102
12.4	NESS	106
13	量子力学	116
14	相対論	116
15	場の量子論	116
16	その他	116

第I部

数学

1 位相空間

1.1 定義など

見返しそうな定義などをまとめる。

Def: 位相空間

集合 X とその部分集合族 \mathfrak{O} が以下の条件を満たすとき、 (X, \mathfrak{O}) を位相空間という。

- (1) $\emptyset, X \in \mathfrak{O}$
- (2) $O_1, O_2 \in \mathfrak{O} \Rightarrow O_1 \cap O_2 \in \mathfrak{O}$
- (3) $\{O_i\} \subset \mathfrak{O} \Rightarrow \bigcup_i O_i \in \mathfrak{O}$

Thm: 相対位相

(X, \mathfrak{O}) を位相空間、 $A \subset X$ とする。このとき、

$$\mathfrak{O}_A = \{O \cap A \mid O \in \mathfrak{O}\} \quad (1.1)$$

は A 上の位相となる。

Prf

- (1) $\emptyset \cap A = \emptyset, X \cap A = A$ であり、 $\emptyset, A \in \mathfrak{O}$ であるから、 $\emptyset, A \in \mathfrak{O}_A$ である。
- (2) $O'_1, O'_2 \in \mathfrak{O}_A$ とする。このとき、 $O_1 = O'_1 \cap A, O_2 = O'_2 \cap A$ となる $O'_1, O'_2 \in \mathfrak{O}$ が存在する。したがって、

$$O_1 \cap O_2 = (O'_1 \cap A) \cap (O'_2 \cap A) = (O'_1 \cap O'_2) \cap A \in \mathfrak{O}_A \quad (1.2)$$

であるから、 \mathfrak{O}_A は交叉について閉じている。

- (3) $\{O_i\} \subset \mathfrak{O}_A$ とする。このとき、 $O_i = O'_i \cap A$ となる $O'_i \in \mathfrak{O}$ が存在する。したがって、

$$\bigcup_i O_i = \bigcup_i (O'_i \cap A) = \left(\bigcup_i O'_i \right) \cap A \in \mathfrak{O}_A \quad (1.3)$$

であるから、 \mathfrak{O}_A は和について閉じている。以上より、示された。 \square

Def: 閉集合

(X, \mathfrak{D}) を位相空間、 $A \subset X$ とする。このとき、 A が閉集合であるとは、 $X \setminus A$ が開集合であることをいう。

閉集合全体からなる集合族を \mathfrak{A} とする。

Thm: 閉集合の性質

X を位相空間とする。このとき、以下が成り立つ。

- (1) $\emptyset, X \in \mathfrak{A}$
- (2) $A_1, A_2 \in \mathfrak{A} \Rightarrow A_1 \cup A_2 \in \mathfrak{A}$
- (3) $\{A_i\} \subset \mathfrak{A} \Rightarrow \bigcap_i A_i \in \mathfrak{A}$

Prf

略 (開集合と同様に示す)。

□

Def: 内点/近傍

(X, \mathfrak{D}) を位相空間、 $U \subset X$ とする。このとき、(1) $x \in U$ が内点であるとは、 x を含む開集合 O が存在し、 $O \subset U$ であることをいう。

(2) x が U の内点であるとき、 U は x の近傍であるという。とくに、 U が X の開集合/閉集合であるとき、 U は x の開近傍/閉近傍であるという。

Def: 内部/外部/境界

(X, \mathfrak{D}) を位相空間、 $A \subset X$ とする。このとき、

(1) A の内部とは、 A に含まれる全ての内点からなる集合のことであり、

$$A^i = \bigcup \{O \subset A \mid O \text{ は開集合}\} \quad (1.4)$$

と定義される。

(2) A の外部とは、 A の補集合の内部のことであり、

$$A^o = (X \setminus A)^i \quad (1.5)$$

と定義される。

(3) A の境界とは、 A の内部でも外部でもない点全体のことであり、

$$\partial A = \overline{A^i \sqcup A^o} \quad (1.6)$$

と定義される。

すなわち、 A の内部とは、 A に含まれる全ての開集合のなかで最大のものである。

Def: 閉包/触点

(X, \mathfrak{O}) を位相空間、 $A \subset X$ とする。このとき、

(1) A の閉包とは、 A を含む最小の閉集合のことであり、

$$\overline{A} = \bigcap \{F \supset A \mid F \text{ は閉集合}\} \quad (1.7)$$

と定義される。

(2) $x \in \overline{A}$ であるとき、 x は A の触点であるという。

Prop: 内部/外部

(X, \mathfrak{O}) を位相空間、 $A \subset X$ とする。このとき、以下が成り立つ。

(1) A が開集合であるための必要う十分条件は、 $A = A^\circ$ であることである。

(2) A が閉集合であるための必要う十分条件は、 $A = \overline{A}$ であることである。

Prf

定義から明らかである。 □

Def: 連続写像

$(X, \mathfrak{O}_X), (Y, \mathfrak{O}_Y)$ を位相空間、 $f: X \rightarrow Y$ とする。このとき、 f が連続であるとは、 Y の任意の開集合 O に対して、 $f^{-1}(O)$ が X の開集合であることをいう。

Prop:

ユークリッド空間における連続写像の定義と、上での連続写像は等価である。

Prf

\Rightarrow

$f: (V, \mathfrak{O}_V) \rightarrow (W, \mathfrak{O}_W)$ をユークリッド空間における連続写像とする。 U を W の任意の開集合としたとき、 $f^{-1}(U)$ は V の開集合であることを示せばよい。

任意の点 $\mathbf{a} \in f^{-1}(U)$ をとると、逆像の定義より、 $f(\mathbf{a}) \in U$ である。 U が開集合なので、十分小さな $\epsilon > 0$ が存在して、 $B_\epsilon(f(\mathbf{a}); \mathbb{R}^n) \subset U$ となる。

また、 f は連続写像であるから、十分小さな $\delta > 0$ を選べば、

$$f(B_\delta(\mathbf{a}; \mathbb{R}^m)) \subset B_\epsilon(f(\mathbf{a}); \mathbb{R}^n) \quad (1.8)$$

となる。以上を組み合わせて、

$$f(B_\delta(\mathbf{a}; \mathbb{R}^m)) \subset U \quad (1.9)$$

である。ここで、逆像の性質から、

$$B_\delta(\mathbf{a}; \mathbb{R}^m) \subset f^{-1}(U) \quad (1.10)$$

であるから、 $f^{-1}(U)$ は開集合である。

⇐

任意の点 $\mathbf{a} \in V$ と、像 $f(\mathbf{a}) \in W$ の近傍 $B_\epsilon(f(\mathbf{a}); \mathbb{R}^n)$ をとる。このとき、 ϵ はどれほど小さくとっても良いので、

$$B_\epsilon(f(\mathbf{a}); \mathbb{R}^n) \subset W \quad (1.11)$$

としてよい。いま、開円板の定義から、

$$B_\epsilon(f(\mathbf{a}); \mathbb{R}^n) \in \mathfrak{D}_W \quad (1.12)$$

である。また、 f は連続写像であるから、

$$f^{-1}(B_\epsilon(f(\mathbf{a}); \mathbb{R}^n)) \in \mathfrak{D}_V \quad (1.13)$$

である。いま、

$$\mathbf{a} \in f^{-1}(B_\epsilon(f(\mathbf{a}); \mathbb{R}^n)) \quad (1.14)$$

であるから、 $\delta > 0$ を十分小さくとれば、

$$B_\delta(\mathbf{a}; \mathbb{R}^m) \subset f^{-1}(B_\epsilon(f(\mathbf{a}); \mathbb{R}^n)) \quad (1.15)$$

となる。以上より、

$$f(B_\delta(\mathbf{a}; \mathbb{R}^m)) \subset B_\epsilon(f(\mathbf{a}); \mathbb{R}^n) \quad (1.16)$$

であるから、 f は連続写像である。 □

Def: 商位相

(X, \mathfrak{D}) を位相空間、 Y を単なる集合であるとする。このとき、

$$f : X \rightarrow Y \quad (1.17)$$

が全射であるとする。このとき、

$$\mathfrak{D}_Y = \{O \subset Y \mid f^{-1}(O) \in \mathfrak{D}\} \quad (1.18)$$

と定義すると、 \mathfrak{D}_Y は Y 上の位相となる。これを商位相といい、 (Y, \mathfrak{D}_Y) を (X, \mathfrak{D}) の商空間という。

(\because) (\mathfrak{D}_Y が位相となること)

(1) $f^{-1}(\emptyset) = \emptyset, f^{-1}(Y) = X$ であるから、 $\emptyset, Y \in \mathfrak{D}_Y$ である。

(2) $O_1, O_2 \in \mathfrak{D}_Y$ とする。このとき、 $f^{-1}(O_1), f^{-1}(O_2) \in \mathfrak{D}$ であるから、

$$f^{-1}(O_1) \cap f^{-1}(O_2) = f^{-1}(O_1 \cap O_2) \in \mathfrak{D} \quad (1.19)$$

である。したがって、 $O_1 \cap O_2 \in \mathfrak{O}_Y$ である。

(3) $\{O_i\} \subset \mathfrak{O}_Y$ とする。このとき、 $f^{-1}(O_i) \in \mathfrak{O}$ であるから、

$$f^{-1}\left(\bigcup_i O_i\right) = \bigcup_i f^{-1}(O_i) \in \mathfrak{O} \quad (1.20)$$

である。したがって、 $\bigcup_i O_i \in \mathfrak{O}_Y$ である。以上より、示された。 \square

このとき、 Y に商位相を入れれば、 $f: X \rightarrow Y$ は連続写像となる。

1.2 その他

最大元と極大元の例

ex.

$(x, y) \in \mathbb{R}^2$ に対して、

$$(x_1, y_1) \leq (x_2, y_2) \Leftrightarrow x_1 \leq x_2 \quad \text{and} \quad y_1 \leq y_2 \quad (1.21)$$

と定義する。このとき、

$$A = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (2, 0)\} \quad (1.22)$$

について、

$$(0, 0) \leq (0, 1) \leq (0, 2) \quad (1.23)$$

$$(0, 0) \leq (1, 0) \leq (1, 1) \quad (1.24)$$

$$(0, 0) \leq (2, 0) \quad (1.25)$$

となり、極大元は 3 つ存在し、最大限は存在しない。

2 線形代数

2.1 行列の積

Thm. 行列の積の分割

行列 A, B の積は、ベクトルを用いることで以下のように分割できる。

1.

$$AB = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} (\mathbf{b}'_1 \quad \mathbf{b}'_2 \quad \cdots \quad \mathbf{b}'_n) = \begin{pmatrix} \mathbf{a}_1 \mathbf{b}'_1 & \mathbf{a}_1 \mathbf{b}'_2 & \cdots & \mathbf{a}_1 \mathbf{b}'_n \\ \mathbf{a}_2 \mathbf{b}'_1 & \mathbf{a}_2 \mathbf{b}'_2 & \cdots & \mathbf{a}_2 \mathbf{b}'_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \mathbf{b}'_1 & \mathbf{a}_m \mathbf{b}'_2 & \cdots & \mathbf{a}_m \mathbf{b}'_n \end{pmatrix} \quad (2.1)$$

2.

$$AB = (\mathbf{a}'_1 \quad \mathbf{a}'_2 \quad \cdots \quad \mathbf{a}'_m) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \mathbf{a}'_1 \mathbf{b}_1 & \mathbf{a}'_1 \mathbf{b}_2 & \cdots & \mathbf{a}'_1 \mathbf{b}_n \\ \mathbf{a}'_2 \mathbf{b}_1 & \mathbf{a}'_2 \mathbf{b}_2 & \cdots & \mathbf{a}'_2 \mathbf{b}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}'_m \mathbf{b}_1 & \mathbf{a}'_m \mathbf{b}_2 & \cdots & \mathbf{a}'_m \mathbf{b}_n \end{pmatrix} \quad (2.2)$$

3.

$$AB = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} B = \begin{pmatrix} \mathbf{a}_1 B \\ \mathbf{a}_2 B \\ \vdots \\ \mathbf{a}_m B \end{pmatrix} \quad (2.3)$$

4.

$$AB = A (\mathbf{b}'_1 \quad \mathbf{b}'_2 \quad \cdots \quad \mathbf{b}'_n) = (A\mathbf{b}'_1 \quad A\mathbf{b}'_2 \quad \cdots \quad A\mathbf{b}'_n) \quad (2.4)$$

2.2 行列式

2.2.1 置換

Def. 置換

n 個の文字 $1, 2, \dots, n$ からなる集合を

$$M_n = \{1, 2, \dots, n\} \quad (2.5)$$

とする。このとき、写像 $\sigma: M_n \rightarrow M_n$ が置換であるとは、この写像が全単射であることをいう。このとき、

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \quad (2.6)$$

と表す。また、 n 個の文字の置換全体の集合を S_n と書く。

Def. 置換の積

置換 σ, τ に対して、その積 $\tau\sigma$ を以下のように定義する。

$$\tau\sigma := \tau \circ \sigma \quad (2.7)$$

Def. 巡回置換/互換

$M_n = \{1, 2, \dots, n\}$ のうち、 i_1, i_2, \dots, i_m 以外を動かさないで、 $i_1 \mapsto i_2, i_2 \mapsto i_3, \dots, i_m \mapsto i_1$ のように一巡させる置換

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_m & i_{m+1} & \cdots & i_n \\ i_2 & i_3 & \cdots & i_1 & i_{m+1} & \cdots & i_n \end{pmatrix} \quad (2.8)$$

を巡回置換といい、

$$(i_1, i_2, \dots, i_m) \quad (2.9)$$

と表す。また、2 つの文字を入れ替える置換

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ i_2 & i_1 & \cdots & i_n \end{pmatrix} \quad (2.10)$$

を互換といい、

$$(i_1, i_2) \quad (2.11)$$

と表す。

Thm. 巡回置換による置換の分解

任意の置換は共通の文字を含まない巡回置換の積に分解できる。

Prf.

適当な置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \quad (2.12)$$

を考える。このとき、例えば 1 がどのように移るかを考える。1 を何回も移したときにもどって来ないと仮定する。いま、 M_n は有限だから、1 を移す操作を繰り返すと、1 以外のある数字が繰り返し出てくる。この数字を j_k とする。そうすると、ある二つの数字 j_{k-1}, j_m について、 $j_{k-1} \rightarrow j_k, j_m \rightarrow j_k$ となる。これは、置換が全単射なことに矛盾する。したがって、1 を何回も移してもとに戻る。同様に、他の数字についても同様の議論ができる。以上より、示された。 \square

Thm. 互換による巡回置換の分解

巡回置換 (i_1, i_2, \dots, i_m) は以下のように $m - 1$ 個の互換の積に分解できる。

$$(i_1, i_2, \dots, i_m) = (i_1, i_m)(i_1, i_{m-1}) \cdots (i_1, i_3)(i_1, i_2) \quad (2.13)$$

Prf.

手を動かせば明らかである。 \square

Cor. 互換による置換の分解

任意の置換は互換の積に分解できる。

Prf.

上二つの定理から明らかである。 \square

Def. 置換の符号

置換 σ に対して、その符号 $\text{sgn}(\sigma)$ を以下のように定義する。

$$\text{sgn}(\sigma) = (-1)^{N(\sigma)} \quad (2.14)$$

ただし、 $N(\sigma)$ は σ を互換の積に分解したときの互換の個数である。

置換の符号の定義が well-defined であることが以下で示される。

Thm. 置換の符号の不変性

互換への分解の仕方によらず、置換の符号は一意である。

Prf.

2.2.2 行列式の定義

Def. 行列式

n 次正方行列 A に対して、その行列式 $\det A$ は以下のように定義される。

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.15)$$

ただし、 $\operatorname{sgn}(\sigma)$ は置換 σ の符号であり、 σ が偶置換のとき $+1$ 、奇置換のとき -1 となる。

Ex.

2 次正方行列 A に対して、

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (2.16)$$

について、

$$\det A = 1 \cdot a_{11}a_{22} + (-1) \cdot a_{12}a_{21} \quad (2.17)$$

$$= a_{11}a_{22} - a_{12}a_{21} \quad (2.18)$$

である。

2.2.3 行列式の性質

Thm.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.19)$$

Prf.

行列式の定義は、

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.20)$$

であった。ここで、 $k > 1$ について、 $\sigma(k) = 1$ だとすると、

$$\operatorname{sgn} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} = 0 \quad (2.21)$$

となるから、和の中で考える必要がない。したがって、 $\sigma(1) = 1$ となる部分だけ和を考えればよい。このとき、

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.22)$$

$$= \sum_{\sigma \in S_n, \sigma(1)=1} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.23)$$

$$= a_{11} \sum_{\sigma \in S_n, \sigma(1)=1} \operatorname{sgn}(\sigma) a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.24)$$

$$= a_{11} \sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma) a_{22} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad (2.25)$$

$$= a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.26)$$

である。以上より、示された。 \square

Cor.

上三角行列 A の行列式は、対角成分の積に等しい。すなわち、

$$\det A = a_{11} a_{22} \cdots a_{nn} \quad (2.27)$$

Prf.

上の定理を繰り返し用いればよい。 \square

Thm.

A を r 次正方行列とし、 D を s 次正方行列とする。このとき、

$$\begin{vmatrix} A & B \\ O & D \end{vmatrix} = \det A \det D \quad (2.28)$$

が成り立つ。

Prf.

$$X = \begin{pmatrix} A & B \\ O & D \end{pmatrix} = (a)_{ij} \quad (2.29)$$

とし、 $n = r + s$ とする。このとき、行列式の定義から、

$$\det X = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) (a)_{1\sigma(1)} (a)_{2\sigma(2)} \cdots (a)_{n\sigma(n)} \quad (2.30)$$

である。上の定理を参考に、置換を分解して考えると、

Thm.

行列 A の一つのを c 倍すると、行列式も c 倍される。すなわち、

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.31)$$

が成り立つ。

Prf.

左辺を計算すると、行列式の定義より、

$$\sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots ca_{i\sigma(i)} \cdots a_{n\sigma(n)} = c \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} \quad (2.32)$$

となることがわかる。したがって、示された。 \square

Cor.

行列 A の一つの成分がすべて 0 であるとき、行列式は 0 である。

Prf.

$0 = 0 \times 0$ として上の定理を用いればよい。 \square

Thm.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{i1} + c_{i1} & b_{i2} + c_{i2} & \cdots & b_{in} + c_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.33)$$

Prf.

略。 \square

Thm. 二つの行の入れ替え

行列式の二つの行を入れ替えると、行列式は -1 倍される。すなわち、

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.34)$$

Prf.

$\tau := \sigma(i, j)$ とする。このとき、

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma(i, j)) \quad (2.35)$$

$$= \operatorname{sgn}(\sigma) \operatorname{sgn}(i, j) \quad (2.36)$$

$$= -\operatorname{sgn}(\sigma) \quad (2.37)$$

となる。これを用いて計算すると、

$$(\text{左辺}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} \quad (2.38)$$

$$(\sigma(i) = j, \sigma(j) = i) \text{ より、} \quad (2.39)$$

$$= - \sum_{\tau \in S_n} \operatorname{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{i\tau(i)} \cdots a_{j\tau(j)} \cdots a_{i\tau(i)} \cdots a_{n\tau(n)} \quad (2.40)$$

$$= (\text{右辺}) \quad (2.41)$$

である。したがって、示された。 \square

Thm.

二つの行が等しい行列の行列式は 0 である。

Prf.

上の定理を用いれば、

$$\det A = -\det A \quad (2.42)$$

となるから、 $\det A = 0$ である。したがって、示された。 \square

Thm.

行列の 1 つの行に任意の数をかけて、他の行に加えても、行列式の値は変わらない。すなわち、

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} + ca_{j1} & a_{i2} + ca_{j2} & \cdots & a_{in} + ca_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.43)$$

が成り立つ。

Prf.

$$\text{左辺} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + c \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.44)$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.45)$$

である。したがって、示された。 \square

Thm.

$$\det A = \det A^T \quad (2.46)$$

Prf.

2.2.4 行列式の展開

Def. 余因子

n 次正方行列 $A = (a_{ij})$ に対して、その第 i 行および第 j 列を取り除いた $n - 1$ 次正方行列 A_{ij} を考える。このとき、

$$\tilde{a}_{ij} = (-1)^{i+j} \det A_{ij} \quad (2.47)$$

を A の (a_{ij}) に対する余因子という。

Thm. 余因子展開

n 次正方行列 A に対して、

$$\det A = \sum_{j=1}^n a_{ij} \tilde{a}_{ij} \quad (2.48)$$

が成り立つ。これを、第 i 行に対する余因子展開という。また、

$$\sum_{j=1}^n a_{ij} \tilde{a}_{kj} = 0 \quad (i \neq k) \quad (2.49)$$

が成り立つ。

Prf.

与えられた行列 A の第 i 行は、

$$(a_{i1} \ a_{i2} \ \cdots \ a_{in}) = (a_{i1} \ 0 \ \cdots \ 0) + (0 \ a_{i2} \ \cdots \ 0) + \cdots + (0 \ 0 \ \cdots \ a_{in}) \quad (2.50)$$

と書ける。したがって、行列式は、

$$\det A = \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.51)$$

となる。行と列をうまく入れ替えて計算することで、

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij} \quad (2.52)$$

$$= \sum_{j=1}^n a_{ij} \tilde{a}_{ij} \quad (2.53)$$

が示される。第二の式は気が向いたら証明する。 \square

Thm.

n 次正方行列 A, B に対して、

$$\det(AB) = \det(A) \det(B) \quad (2.54)$$

が成り立つ。

Prf.

行列 A, B をそれぞれ $A = (a_{ij}), B = (b_{ij})$ と書くことにする。そして、行列 B のベクトルを b_1, b_2, \dots, b_n とする。すなわち、

$$\mathbf{b}_j = (b_{1j} \quad b_{2j} \quad \dots \quad b_{nj}) \quad (j = 1, 2, \dots, n) \quad (2.55)$$

ブロックに分けての計算法によって、

$$AB = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1i} \mathbf{b}_i \\ \vdots \\ \sum_{i=1}^n a_{ni} \mathbf{b}_i \end{pmatrix} \quad (2.56)$$

である。したがって、次の等式を得る。

$$\det(AB) = \sum_{j_1=1}^n \sum_{j_2=1}^n \dots \sum_{j_{n-1}=1}^n \sum_{j_n=1}^n a_{1j_1} \dots a_{nj_n} \begin{vmatrix} \mathbf{b}_{j_1} \\ \vdots \\ \mathbf{b}_{j_n} \end{vmatrix} \quad (2.57)$$

ここで、和は j_1, j_2, \dots, j_n がそれぞれ 1 から n まで動くので n^n 個の項になる。しかし、 b_{j_1}, \dots, b_{j_n} のうちに同じものがあれば、行列式

$$\begin{vmatrix} \mathbf{b}_{j_1} \\ \vdots \\ \mathbf{b}_{j_n} \end{vmatrix} = 0 \quad (2.58)$$

となるから、 j_1, j_2, \dots, j_n がすべて異なる場合の和を考えればよい。すなわち j_1, j_2, \dots, j_n の順列となる場合に他ならない。しかも、和はちょうど $n!$ の順列をわたる。ゆえに、

$$\det(AB) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \begin{vmatrix} \mathbf{b}_{1\sigma(1)} \\ \vdots \\ \mathbf{b}_{n\sigma(n)} \end{vmatrix} = \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \begin{vmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{vmatrix} \quad (2.59)$$

となるので、

$$\det(AB) = \det(A) \det(B) \quad (2.60)$$

が示された。 □

2.3 ベクトル空間

2.3.1 部分空間

Def. 部分空間

ベクトル空間 V の部分集合 W が次の条件を満たすとき、 W を V の部分空間という。

- $\mathbf{0} \in W$
- $\mathbf{a}, \mathbf{b} \in W \Rightarrow \mathbf{a} + \mathbf{b} \in W$
- $\mathbf{a} \in W, c \in \mathbb{R} \Rightarrow c\mathbf{a} \in W$

すなわち、 V における和とスカラー倍の演算によって閉じている部分集合が部分空間である。

ex. 生成される空間

$$S[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r] = \{c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_r\mathbf{a}_r \mid c_1, c_2, \dots, c_r \in \mathbb{R}\} \quad (2.61)$$

のことを、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ によって生成される部分空間という。このとき、 $S[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r]$ は V の部分空間である。(証明略)

ex. 解空間

A を $m \times n$ 行列とすると、

$$W = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\} \quad (2.62)$$

のことを、 A の解空間という。このとき、 W は \mathbb{R}^n の部分空間である。(証明略)

2.3.2 線形写像

Def. 線形写像

ベクトル空間 V, W に対して、写像 $f: V \rightarrow W$ が次の2つの条件を満たすとき、 f を V から W への線形写像という。

- $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}) \quad (\forall \mathbf{a}, \mathbf{b} \in V)$
- $f(c\mathbf{a}) = cf(\mathbf{a}) \quad (\forall \mathbf{a} \in V, c \in \mathbb{R})$

Def. 像空間/核空間

V, W をベクトル空間、 $f: V \rightarrow W$ を線形写像とする。このとき、

$$\text{Im}f = \{f(\mathbf{x}) \mid \mathbf{x} \in V\} \quad (2.63)$$

を f の像空間という。また、

$$\text{Ker}f = \{\mathbf{x} \in V \mid f(\mathbf{x}) = \mathbf{0}\} \quad (2.64)$$

を f の核空間という。

Thm.

V, W をベクトル空間、 $f: V \rightarrow W$ を線形写像とする。このとき、

1. $\text{Im}f$ は W の部分空間である。
2. $\text{Ker}f$ は V の部分空間である。

Prf.

(1)

$\mathbf{0} \in \text{Im}f$ であるから、 $\text{Im}f$ は空でない。

また、任意の $\lambda, \mu \in \mathbb{R}$ に対して、 $f(\mathbf{a}), f(\mathbf{b}) \in \text{Im}f$ とすると、

$$\lambda f(\mathbf{a}) + \mu f(\mathbf{b}) = f(\lambda \mathbf{a} + \mu \mathbf{b}) \in \text{Im}f \quad (2.65)$$

であるから、 $\text{Im}f$ は和とスカラー倍に対して閉じている。したがって、 $\text{Im}f$ は W の部分空間である。

(2)

$\mathbf{0} \in \text{Ker}f$ であるから、 $\text{Ker}f$ は空でない。

また、任意の $\lambda, \mu \in \mathbb{R}$ に対して、 $\mathbf{x}, \mathbf{y} \in \text{Ker}f$ とすると、

$$f(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) = \mathbf{0} \quad (2.66)$$

であるから、 $\text{Ker}f$ は和とスカラー倍に対して閉じている。したがって、 $\text{Ker}f$ は V の部分空間である。 \square

2.3.3 1 次独立と 1 次従属

Def.1 1 次独立

ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が次の条件を満たすとき、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立であるという。

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \mathbf{0} \quad (2.67)$$

となるとき、 $c_1 = c_2 = \dots = c_n = 0$ である。

また、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次独立でないとき、1 次従属であるという。

Thm.

ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ について、以下の条件は同値である。

1. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。
2. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ の線形結合で表される元の表し方は一意である。

Prf.

(1) \Rightarrow (2)

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = d_1 \mathbf{a}_1 + d_2 \mathbf{a}_2 + \dots + d_n \mathbf{a}_n \quad (2.68)$$

とする。このとき、

$$(c_1 - d_1) \mathbf{a}_1 + (c_2 - d_2) \mathbf{a}_2 + \dots + (c_n - d_n) \mathbf{a}_n = \mathbf{0} \quad (2.69)$$

である。 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立であるから、 $c_1 = d_1, c_2 = d_2, \dots, c_n = d_n$ である。したがって、線形結合で表される元の表し方は一意である。

(2) \Rightarrow (1)

$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \mathbf{0}$ とする。このとき、仮定より、 $c_1 = c_2 = \dots = c_n = 0$ 以外の方法で $\mathbf{0}$ を表すことはできない。したがって、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。 \square

Thm.

ベクトルの組 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次従属であることと、その組のうち、1 つのベクトルが残りのベクトルの線形結合で表されることは同値である。

Prf.

\mathbf{a}_1 が他の残りのベクトルの線形結合で表されるとする。すなわち、

$$\mathbf{a}_1 = c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n \quad (2.70)$$

とする。このとき、

$$\mathbf{a}_1 - c_2 \mathbf{a}_2 - \dots - c_n \mathbf{a}_n = \mathbf{0} \quad (2.71)$$

であるから、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次従属である。

逆に、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次従属であるとする。このとき、

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \mathbf{0} \quad (2.72)$$

となる。 $c_1 \neq 0$ とすると、

$$\mathbf{a}_1 = -\frac{c_2}{c_1} \mathbf{a}_2 - \dots - \frac{c_n}{c_1} \mathbf{a}_n \quad (2.73)$$

となるから、 \mathbf{a}_1 は残りのベクトルの線形結合で表される。したがって、示された。 \square

上の定理の対偶を考えることで、ベクトルの組 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次独立であることと、その組のうち、1 つのベクトルが残りのベクトルの線形結合で表されないことは同値であることがわかる。

Thm.

ベクトルの組 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ について、

1. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m (m < n)$ が 1 次従属であれば、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ も 1 次従属である。
2. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次独立であれば、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m (m < n)$ も 1 次独立である。(上の対偶)

Prf.

(1)

$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ が 1 次従属であるとする。このとき、自明でない関係式

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_m \mathbf{a}_m = \mathbf{0} \quad (2.74)$$

が存在する。これより、

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_m \mathbf{a}_m + 0 \mathbf{a}_{m+1} + \dots + 0 \mathbf{a}_n = \mathbf{0} \quad (2.75)$$

という自明でない関係式が存在する。したがって、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次従属である。

(2)

(1) の対偶である。 \square

Thm.

ベクトルの組 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が 1 次独立で、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{a}$ が 1 次従属であるとき、 \mathbf{a} は $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ の線形結合で表される。

Prf.

仮定より、自明でない関係式

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n + c \mathbf{a} = \mathbf{0} \quad (2.76)$$

が存在する。ここで、 $c = 0$ とすると、1 次独立の仮定と矛盾する。よって、 $c \neq 0$ である。したがって、

$$\mathbf{a} = -\frac{c_1}{c}\mathbf{a}_1 - \frac{c_2}{c}\mathbf{a}_2 - \cdots - \frac{c_n}{c}\mathbf{a}_n \quad (2.77)$$

であるから、 \mathbf{a} は $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ の線形結合で表される。また、一意性も簡単に示せる。 \square

2.3.4 行列式と 1 次独立性的関係

Thm.

A を n 次正方行列、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ を A の行ベクトル、 $\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_n$ を A の列ベクトルとする。このとき、次の条件は同値である。

1. $\det A \neq 0$
2. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。
3. $\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_n$ は 1 次独立である。

Prf.

(1) \Rightarrow (3)

$\det A \neq 0$ であるとする。このとき、 A は正則であるから、 $A\mathbf{x} = \mathbf{0}$ の解は $\mathbf{x} = \mathbf{0}$ のみである。したがって、

$$x_1\mathbf{a}'_1 + x_2\mathbf{a}'_2 + \cdots + x_n\mathbf{a}'_n = \mathbf{0} \quad (2.78)$$

となるとき、 $x_1 = x_2 = \cdots = x_n = 0$ である。したがって、 $\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_n$ は 1 次独立である。

(3) \Rightarrow (1)

対偶で示す。 $\det A = 0$ であるとする。このとき、 $A\mathbf{x} = \mathbf{0}$ は非自明な解をもつ。したがって、

$$x_1\mathbf{a}'_1 + x_2\mathbf{a}'_2 + \cdots + x_n\mathbf{a}'_n = \mathbf{0} \quad (2.79)$$

となる非自明な解が存在する。したがって、 $\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_n$ は 1 次従属である。

(1) \Leftrightarrow (2) は、上の議論を転置行列に対して行えばよい。 \square

Thm.

$n + 1$ 個の n 項数ベクトルは 1 次従属である。

Prf.

$n + 1$ 個の n 項数ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n+1}$ に対して、

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n+1,1} & a_{n+1,2} & \cdots & a_{n+1,n} & 0 \end{pmatrix} = 0 \quad (2.80)$$

であることから、 $(a_{11}, a_{12}, \dots, a_{1n}, 0), (a_{21}, a_{22}, \dots, a_{2n}, 0), \dots, (a_{n+1,1}, a_{n+1,2}, \dots, a_{n+1,n}, 0)$ は 1 次従属である。したがって、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n+1}$ は 1 次従属である。□

Thm.

$r > n$ とするとき、 r 個の n 項数ベクトルは 1 次従属である。

Prf.

一つ前の定理から $r = n + 1$ の場合が示せる。また、このことと、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m (m < n)$ が 1 次従属であれば、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ も 1 次従属であること (前のほうの定理) から、 $r > n$ の場合も示される。□

Thm.

$r < s$ とするとき、ベクトル空間の s 個のベクトル $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s$ がすべて r 個のベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ の線形結合で表されるとき、 $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s$ は 1 次従属である。

Prf.

$s = r + 1$ の場合について示せば十分である。仮定より、

$$\mathbf{b}_j = \sum_{i=1}^r c_{ij} \mathbf{a}_i \quad (2.81)$$

となる c_{ij} が存在する。これを行列で書くと、

$$(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{r+1}) = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r) C \quad (2.82)$$

となる行列 C が存在する。このとき、 C は $r + 1 \times r$ 行列であるから、 $r + 1$ 個の行ベクトルは 1 次従属である。したがって、

$$\sum_{j=1}^{r+1} d_j c_{ij} = 0 \quad (2.83)$$

となる d_1, d_2, \dots, d_{r+1} で、少なくとも一つが 0 でないものが存在する。ここで、

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{r+1} \end{pmatrix} \quad (2.84)$$

とおくと、

$$C \mathbf{d} = \mathbf{0} \quad (2.85)$$

であるから、

$$d_1 \mathbf{b}_1 + d_2 \mathbf{b}_2 + \cdots + d_{r+1} \mathbf{b}_{r+1} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{r+1}) \mathbf{d} \quad (2.86)$$

$$= (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r) C \mathbf{d} \quad (2.87)$$

$$= \mathbf{0} \quad (2.88)$$

となる。したがって、 $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{r+1}$ は 1 次従属である。 \square

2.3.5 ベクトル空間の基底

Def. ベクトル空間の基底

ベクトル空間 V のベクトルの組 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が次の 2 つの条件を満たすとき、 V の基底という。

- $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。
- 任意のベクトル $\mathbf{x} \in V$ は $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ の線形結合で表される。すなわち、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は V を生成する。

ex.

\mathbb{R}^3 において、ベクトルの組

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad (2.89)$$

は \mathbb{R}^3 の基底である。

Prf.

(1)

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + c_3 \mathbf{a}_3 = \mathbf{0} \quad (2.90)$$

とする。これを計算すると、

$$c_1 + c_3 = 0 \quad (2.91)$$

$$c_1 + c_2 = 0 \quad (2.92)$$

$$c_2 + c_3 = 0 \quad (2.93)$$

となる。これを解くと、 $c_1 = c_2 = c_3 = 0$ である。したがって、 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ は 1 次独立である。

(2)

任意のベクトル $\mathbf{b} \in \mathbb{R}^3$ に対して、

$$\mathbf{b} = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + c_3 \mathbf{a}_3 \quad (2.94)$$

となる c_1, c_2, c_3 が存在することを示す。これは、

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad (2.95)$$

を解けばよい。このとき、

$$\det \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 2 \neq 0 \quad (2.96)$$

であるから、逆行列が存在する。したがって、 c_1, c_2, c_3 が存在する。したがって、 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ は \mathbb{R}^3 を生成する。□

2.3.6 ベクトル空間の次元

ベクトル空間の次元を定義するために、2つの補題を示す。

Lem.1

$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s$ が共に 1 次独立なベクトルの組で、それぞれの生成する部分空間が一致するならば、 $r = s$ である。

Prf.

仮定より、各 \mathbf{a}_i は $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s$ の線形結合で表される。ところで、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ は 1 次独立であるから、前の前の定理の対偶より、 $r \leq s$ である。同様に、 $s \leq r$ であるから、 $r = s$ である。□

Lem.2

ベクトル空間 V の基底に含まれるベクトルの個数は、基底の取り方によらず一定である。

Prf.

$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ と $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s$ が共に V の基底であるとする。このとき、それぞれ 1 次独立であり、かつ V を生成する。したがって、前の定理より、 $r = s$ である。□

以上の補題から、ベクトル空間の次元を定義することができる。

Def. ベクトル空間の次元

ベクトル空間 V に対して、 V の基底に含まれるベクトルの個数を V の次元といい、 $\dim V$ と書く。

Thm.

ベクトル空間 V と、その部分空間 U について、 $\dim U \leq \dim V$ である。また、

$$U = V \Leftrightarrow \dim U = \dim V \quad (2.97)$$

が成り立つ。

Prf.

U の基底 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ は V のベクトルからなり、一次独立であるから、 $\dim U \leq \dim V$ である。

また、 $\dim U = \dim V$ であるとする。このとき、 U の基底 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ は V の 1 次独立なベクトルであり、 $\dim U = \dim V$ であるから、これらは V の基底である。したがって、 $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m)$ は V を生成する。したがって、 $U = V$ である。逆は自明。□

Thm. 基底の条件 1

ベクトル空間 V について、次の条件は同値である。

1. $\dim V = n$
2. V のベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ があって、 V の任意のベクトル \mathbf{a} は、

$$\mathbf{a} = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \cdots + c_n \mathbf{a}_n \quad (2.98)$$

と一意に表される。

Prf.

(1) \Rightarrow (2)

ベクトル空間 V の一つの基底を $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ とする。このとき、 V の任意のベクトル \mathbf{a} は、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ の線形結合で一意に表される。

(2) \Rightarrow (1)

仮定より、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は V を生成し、さらに一次結合の表し方が一意であることから、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。したがって、 $\dim V = n$ である。□

Thm. 基底の条件 2

V を n 次元ベクトル空間とする。 V の n 個のベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ について、以下の 3 条件は同値である。

1. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は V の基底である。
2. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は 1 次独立である。
3. $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は V を生成する。

Prf.

hoge

□

2.3.7 基底間の関係

Thm. 基底の変換

基底ベクトル間の変換は、2つの基底を $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ と $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ とすると、ある正則行列 P によって

$$\mathbf{b}_j = \sum_{i=1}^n P_{ij} \mathbf{a}_i \quad (2.99)$$

と表される。行列表記すると、

$$(\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n) = (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n) P \quad (2.100)$$

となる。

Prf.

hoge

□

このもとで、成分の変化を見る。今、

$$y_1 \mathbf{b}_1 + y_2 \mathbf{b}_2 + \dots + y_n \mathbf{b}_n = (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad (2.101)$$

$$= (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n) P \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad (2.102)$$

$$= (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (2.103)$$

となる。したがって、

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (2.104)$$

を得る。

2.3.8 線形写像の行列表現

線形写像の行列表現を考える。 V を n 次元ベクトル空間、 W を m 次元ベクトル空間とし、 V の基底を $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 、 W の基底を $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ とする。このとき、線形写像 $f: V \rightarrow W$ について、

$$f(\mathbf{a}_j) = \sum_{i=1}^m A_{ij} \mathbf{b}_i \quad (2.105)$$

により行列 A を定義する。このとき、

$$f(\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n) = (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_m) A \quad (2.106)$$

が成り立つ。この行列 A を線形写像 f の行列表現という。このもとで、

$$f(\mathbf{x}) = f(x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n) \quad (2.107)$$

$$= f\left(\sum_{j=1}^n x_j \mathbf{a}_j\right) \quad (2.108)$$

$$= \sum_{j=1}^n x_j f(\mathbf{a}_j) \quad (2.109)$$

$$= \sum_{j=1}^n x_j \sum_{i=1}^m A_{ij} \mathbf{b}_i \quad (2.110)$$

$$= \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} x_j\right) \mathbf{b}_i \quad (2.111)$$

$$= \sum_{i=1}^m y_i \mathbf{b}_i \quad (2.112)$$

となる。したがって、

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (2.113)$$

となる。

Thm.

$f: V \rightarrow W$ を線形写像とする。 V の基底 $\mathbf{a}_1, \dots, \mathbf{a}_n$ と W の基底 $\mathbf{b}_1, \dots, \mathbf{b}_m$ に関する f の表現行列を A とし、 V の基底 $\mathbf{a}'_1, \dots, \mathbf{a}'_n$ と W の基底 $\mathbf{b}'_1, \dots, \mathbf{b}'_m$ に関する f の表現行列を B とする。また、基底変換の行列をそれぞれ P, Q とする。すなわち、

$$(\mathbf{a}'_1, \dots, \mathbf{a}'_n) = (\mathbf{a}_1, \dots, \mathbf{a}_n)P, \quad (2.114)$$

$$(\mathbf{b}'_1, \dots, \mathbf{b}'_m) = (\mathbf{b}_1, \dots, \mathbf{b}_m)Q, \quad (2.115)$$

とおく。このとき、

$$B = Q^{-1}AP \quad (2.116)$$

が成り立つ。

Prf.

表現行列と基底の変換の行列の定義より

$$(f(\mathbf{a}'_1), \dots, f(\mathbf{a}'_n)) = (f(\mathbf{a}_1), \dots, f(\mathbf{a}_n))B = (\mathbf{b}'_1, \dots, \mathbf{b}'_m)QB \quad (2.117)$$

また、線形性から、

$$(f(\mathbf{a}'_1), \dots, f(\mathbf{a}'_n)) = (f(\mathbf{a}_1), \dots, f(\mathbf{a}_n))P = (\mathbf{b}_1, \dots, \mathbf{b}_m)AP \quad (2.118)$$

であるから、

$$(\mathbf{b}_1, \dots, \mathbf{b}_m)QB = (\mathbf{b}_1, \dots, \mathbf{b}_m)AP \quad (2.119)$$

が得られる。 $\mathbf{b}_1, \dots, \mathbf{b}_m$ は一次独立だから、定理 6.4.8 より $QB = AP$ が成り立つ。 Q は正則だから、

$$B = Q^{-1}AP \quad (2.120)$$

□

2.3.9 ベクトル空間の同型

Def. 同型写像/同型

ベクトル空間 V からベクトル空間 W への線形写像 $f: V \rightarrow W$ が全単射であるとき、 f を同型写像といい、 V と W は同型であるという。^a

このとき、 V と W は同型であるとき、 $V \simeq W$ と書く。

^a 群の場合と異なり、全単射しか定義に入れていないのは、もともとそれぞれのベクトル空間に演算が入っているからである。

Thm.

ベクトル空間の間の同型写像は、一次独立、一次従属、基底といった性質を保つ。

Prf.

$f : V \rightarrow W$ を同型写像とする。 V の一次独立なベクトル $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ に対して、一次関係 $a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \dots + a_n f(\mathbf{v}_n) = 0$ が成り立つとすると、

$$f(a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n) = 0 \quad (2.121)$$

である。ところで、 f は単射であるから、

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n = 0 \quad (2.122)$$

である。したがって、 $a_1 = a_2 = \dots = a_n = 0$ である。また、一次従属性については、そもそも一般の線形写像について成り立つ。

基底について、 f が全射なので、任意の $\mathbf{w} \in W$ について、ある $\mathbf{v} \in V$ が存在して、 $f(\mathbf{v}) = \mathbf{w}$ となる。 $\mathbf{v} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n$ と書けるとき、

$$\mathbf{w} = f(\mathbf{v}) = a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \dots + a_n f(\mathbf{v}_n) \quad (2.123)$$

である。したがって、 $f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)$ は W を生成する。また、前半の証明から、 $f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)$ は一次独立である。したがって、 $f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)$ は W の基底である。□

Thm.

ベクトル空間 V, W が同型であるための必要十分条件は、

$$\dim V = \dim W \quad (2.124)$$

である。

Prf.

(\Rightarrow)

$f : V \rightarrow W$ を同型写像とする。このとき、同型写像は基底を保つから、 $\dim V = \dim W$ である。

(\Leftarrow)

$\dim V = \dim W$ であるとする。 V, W の基底を一組ずつ取り、それぞれの基底を $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 、 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ とする。このとき、 V の任意のベクトル \mathbf{x} は、

$$\mathbf{x} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n \quad (2.125)$$

と書ける。そこで、

$$f(\mathbf{x}) = a_1 \mathbf{w}_1 + a_2 \mathbf{w}_2 + \dots + a_n \mathbf{w}_n \quad (2.126)$$

と定義する。このとき、 f は線形写像となる。実際、

$$f(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = f(c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2n}\mathbf{v}_n)) \quad (2.127)$$

$$= f((c_1a_{11} + c_2a_{21})\mathbf{v}_1 + (c_1a_{12} + c_2a_{22})\mathbf{v}_2 + \cdots + (c_1a_{1n} + c_2a_{2n})\mathbf{v}_n) \quad (2.128)$$

$$= (c_1a_{11} + c_2a_{21})\mathbf{w}_1 + (c_1a_{12} + c_2a_{22})\mathbf{w}_2 + \cdots + (c_1a_{1n} + c_2a_{2n})\mathbf{w}_n \quad (2.129)$$

$$= c_1(a_{11}\mathbf{w}_1 + a_{12}\mathbf{w}_2 + \cdots + a_{1n}\mathbf{w}_n) + c_2(a_{21}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{2n}\mathbf{w}_n) \quad (2.130)$$

$$= c_1f(\mathbf{x}_1) + c_2f(\mathbf{x}_2) \quad (2.131)$$

である。以後、 f が全単射であることを示す。

(f が単射であること)

$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ と $\mathbf{y} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n$ に対して、 $f(\mathbf{x}) = f(\mathbf{y})$ とすると、

$$a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n = b_1\mathbf{w}_1 + b_2\mathbf{w}_2 + \cdots + b_n\mathbf{w}_n \quad (2.132)$$

である。したがって、

$$(a_1 - b_1)\mathbf{w}_1 + (a_2 - b_2)\mathbf{w}_2 + \cdots + (a_n - b_n)\mathbf{w}_n = 0 \quad (2.133)$$

である。 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ は一次独立であるから、 $a_1 = b_1, a_2 = b_2, \dots, a_n = b_n$ である。したがって、 f は単射である。

(f が全射であること)

W の任意のベクトル \mathbf{w} は $a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n$ と書ける。したがって、

$$f(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n \quad (2.134)$$

となるので、 f は全射である。したがって、 f は同型写像である。 \square

したがって、抽象的な n 次元ベクトル空間と、その例として考えられる \mathbb{R}^n は、同型である。

Thm.

V, W を n 次元ベクトル空間とし、 V, W の基底をそれぞれ $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ とする。ここで、 $n = \dim V, m = \dim W$ とする。線形写像 $f: V \rightarrow W$ に対して、上記基底に対する f の行列表現をそれぞれ A とするとき、以下は同値である。

1. f が同型写像である。
2. $n = m$ であり、 A は正則行列である。さらに、 A の逆行列 A^{-1} は、 f の逆写像 f^{-1} に対応する行列である。

Prf.

(1 \Rightarrow 2)

f が同型写像であるとする。このとき、前の定理より、 $\dim V = \dim W$ である。したがって、 $n = m$ である。また、 $f^{-1} : W \rightarrow V$ の表現行列を B とすると、

$$AB = E = BA \quad (2.135)$$

である。したがって、 A は正則行列であり、 $A^{-1} = B$ である。

(2 \Rightarrow 1)

$n = m$ であり、 A は正則行列であるとする。このとき、 A の逆行列 A^{-1} は f の逆写像 f^{-1} に対応する行列である。したがって、 f は同型写像である。 \square

2.4 固有値・固有ベクトル

Def. 線形変換の固有値/固有ベクトル

ベクトル空間 V の元 $\mathbf{v} \neq 0_V$ と、スカラー $\alpha \in \mathbb{C}$ に対して、

$$f(\mathbf{v}) = \alpha \mathbf{v} \quad (2.136)$$

となるものが存在するとき、 α を f の固有値、 \mathbf{v} を f の固有ベクトルという。

また、 V の部分集合

$$V_\alpha = \{\mathbf{u} \in V | f(\mathbf{u}) = \alpha \mathbf{u}\} \quad (2.137)$$

を固有値 α に対する固有空間という。

Lem.

f の固有値 α に対する固有空間 V_α は、 V の部分空間である。

Prf.

$$f(c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2) = c_1 f(\mathbf{u}_1) + c_2 f(\mathbf{u}_2) \quad (2.138)$$

$$= c_1 \alpha \mathbf{u}_1 + c_2 \alpha \mathbf{u}_2 \quad (2.139)$$

$$= \alpha(c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2) \quad (2.140)$$

であるから、 V_α は部分空間である。 \square

とくに、 $V = \mathbb{C}^n$ のとき、線形変換は n 次複素正方行列 A によって

$$f(\mathbf{v}) = A\mathbf{v} \quad (2.141)$$

と表される。このとき、 f の固有値 α と固有ベクトル \mathbf{v} は、

$$A\mathbf{v} = \alpha \mathbf{v} \quad (2.142)$$

となる。

Prop.

P を n 次正則行列、 A を n 次正方行列とする。このとき、 $P^{-1}AP$ の固有値全体の集合は、 A の固有値全体の集合と一致する。

Prf.

$$\det(tE_n - P^{-1}AP) = \det(P^{-1}(tE_n - A)P) \quad (2.143)$$

$$= \det(P^{-1}) \det(tE_n - A) \det(P) \quad (2.144)$$

$$= \det(tE_n - A) \quad (2.145)$$

であるから、固有値は一致する。 \square

一般のベクトル空間 V の線形変換 T に対して、固有値と固有ベクトルを求めるには、まず、 V の基 $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ を取り、 T の表現行列 A を求める。そして、 A の固有値を求め、それに対応する固有ベクトル \mathbf{x} を求める。 A の各固有値は f の固有値となり、

$$\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)\mathbf{x} \quad (2.146)$$

が固有値 α に対する固有ベクトルである。

実際、

$$f(\mathbf{u}) = f(x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots + x_n\mathbf{u}_n) \quad (2.147)$$

$$= (f(\mathbf{u}_1), f(\mathbf{u}_2), \dots, f(\mathbf{u}_n))\mathbf{x} \quad (2.148)$$

$$= (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)A\mathbf{x} \quad (2.149)$$

$$= (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)\alpha\mathbf{x} \quad (2.150)$$

$$= \alpha\mathbf{u} \quad (2.151)$$

となる。

Prop.

どのような基をとって上の計算を行っても、固有値全体の集合とその固有空間は変化しない。

Prf.

基を $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ に取り換えた場合、基底変換の行列を P とする。このとき、

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)P \quad (2.152)$$

である。このとき、 f の表現行列は $P^{-1}AP$ である。ここで、 $P^{-1}AP$ と A の固有値全体の集合は同じである。したがって、 $A\mathbf{x} = \alpha\mathbf{x}$ であることと、 $P^{-1}AP(P\mathbf{x}') = \alpha P\mathbf{x}'$ ($\mathbf{x}' = P^{-1}\mathbf{x}$) であ

ることは同値であり、

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \mathbf{x}' = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) P P^{-1} \mathbf{x} \quad (2.153)$$

$$= (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \mathbf{x} \quad (2.154)$$

であるから、固有ベクトルも変わらない。□

2.5 対角化

Thm.

f を、ベクトル空間 V における線形変換とする。このとき、 f の相異なる固有値の固有ベクトルは一次独立である。

Prf.

略。□

Def. 対角化可能

複素 n 次正方行列 A が対角化可能であるとは、ある正則行列 P が存在して、

$$P^{-1} A P = D \quad (2.155)$$

となることである。このとき、 D は対角行列であり、 A は P によって対角化されるという。

Thm.

n 次正方行列 A に対して、以下の 3 条件は同値である。

1. A は対角化可能である。
2. A の固有ベクトルからなる \mathbb{C}^n の基底が存在する。
3. A の各固有値 α_i に対する固有空間の次元の和が n である。

Prf.

(1) \Rightarrow (2)

A が対角化可能であるとする。このとき、

$$P^{-1} A P = D \quad (2.156)$$

$$= \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} \quad (2.157)$$

となる。このとき、 P は正則なので、 P の列ベクトルを $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ とすると、これらは一次独

立である。さらに、これらは A の固有ベクトルである。実際、

$$A\mathbf{u}_i = (PDP^{-1})\mathbf{u}_i \quad (2.158)$$

$$= P D \mathbf{e}_i \quad (2.159)$$

$$= P \alpha_i \mathbf{e}_i \quad (2.160)$$

$$= \alpha_i P \mathbf{e}_i \quad (2.161)$$

$$= \alpha_i (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \mathbf{e}_i \quad (2.162)$$

$$= \alpha_i \mathbf{u}_i \quad (2.163)$$

であるから、 \mathbf{u}_i は α_i に対する固有ベクトルである。

$\dim(\mathbb{C}^n) = n$ で、 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ は一次独立であるから、これは \mathbb{C}^n の基底である。

(2) \Rightarrow (3)

$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ を、 A の固有ベクトルからなる \mathbb{C}^n の基底とする。このとき、各 \mathbf{u}_i の固有値を α_i とし、 A の固有値全体の集合を $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ を、重複する元をはずして整理して、 $\{\alpha_{k_1}, \alpha_{k_2}, \dots, \alpha_{k_m}\}$ とする。このとき、各 α_{k_i} に対して、 α_{k_i} の固有空間に属する $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ の個数を n_{ki} とし、 α_{k_i} の固有空間の次元を d_{ki} とする。このとき、 $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ は一次独立であるから、

$$n_{ki} \leq d_{ki} \quad (2.164)$$

である。したがって、

$$n = \sum_{i=1}^m n_{ki} \leq \sum_{i=1}^m d_{ki} \quad (2.165)$$

である。一方、異なる固有値に属する固有ベクトルは一次独立であるから、

$$\sum_{i=1}^m d_{ki} \leq n \quad (2.166)$$

である。したがって、

$$\sum_{i=1}^m d_{ki} = n \quad (2.167)$$

である。

(3) \Rightarrow (1)

固有値 α_i の固有空間の基を成すベクトルをとると、仮定より、 n 個のベクトル $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ が得られる。異なる固有値の固有ベクトルは一次独立なので、これらは \mathbb{C}^n の基底である。このとき、

$$P = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \quad (2.168)$$

とすると、

$$P^{-1}AP = P^{-1}A(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \quad (2.169)$$

$$= P^{-1}(\alpha_1 \mathbf{u}_1, \alpha_2 \mathbf{u}_2, \dots, \alpha_n \mathbf{u}_n) \quad (2.170)$$

$$= P^{-1}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} \quad (2.171)$$

$$= \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} \quad (2.172)$$

となる。したがって、 A は対角化可能である。 \square

Cor.

n 次正方行列 A の固有方程式が重根を持たないとき、 A は対角化可能である。

Prf.

固有方程式が重根を持たないとき、異なる固有値に対する固有空間の次元の和は n であるから、 A は対角化可能である。 \square

Def. 上三角行列

$i > j$ ならば、 $a_{ij} = 0$ であるような行列を上三角行列という。

対角行列は上三角行列である。また、上三角行列の行列式は対角成分の積である。

Thm.

複素ベクトル空間 V 上の任意の線形変換 f は、ある V の基に関する表現行列が上三角行列になえい、その対角成分に f の固有方程式の解が重複度を含めてすべて現れる。

Prf.

V の次元に関する帰納法で示す。

$n = 1$ のとき、明らか。

$n - 1$ まで成立すると仮定する。線形変換 f に対して、少なくとも一つの固有値 α が存在するので、その固有値に関する固有ベクトル \mathbf{v}_1 をとり、それを含む基底を $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ とする。このとき、 \mathbf{v}_1 は f の固有ベクトルであるから、 $f(\mathbf{v}_1) = \alpha \mathbf{v}_1$ であり、この基底に関する f の表現行

列は

$$\begin{pmatrix} \alpha & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad (2.173)$$

となる。

W を、 $\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ で張られる部分空間とし、 V から W への線形写像 pr を、

$$\text{pr} : V \ni b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \cdots + b_n \mathbf{v}_n \mapsto b_2 \mathbf{v}_2 + \cdots + b_n \mathbf{v}_n \in W \quad (2.174)$$

とする。ここで、合成写像 $f \circ \text{pr} : V \rightarrow W$ を W に制限することにより、 $\text{pr} \circ f|_W : W \rightarrow W$ を得る。

帰納法の仮定より、 W の基底 $\mathbf{w}_2, \dots, \mathbf{w}_n$ で、 $\text{pr} \circ f|_W$ のその基底に関する表現行列が上三角行列になるようなものが存在する。 $\mathbf{v}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ は V の基底となり、その基底に関する f の表現行列は上三角行列になる。

f の固有方程式は、この上三角行列 U により表現行列の固有方程式

$$\det(tE_n - U) = 0 \quad (2.175)$$

となる。その解は、 U の対角成分に一致する。したがって、 f の固有方程式の解は、重複度を含めてすべて現れる。 \square

Def. 正方行列の最小多項式

正方行列 A に対して、最高次係数 1 の多項式

$$F(t) = t^m + \sum_{k=0}^{m-1} c_k t^k \quad (2.176)$$

で、

$$F(A) = A^m + \sum_{k=1}^{m-1} c_k A^k + c_0 E = O \quad (2.177)$$

となるもののうち、次数が最小のものを A の最小多項式という。

n 次正方行列全体 $M(n, \mathbb{C})$ は、 \mathbb{C}^n 上の n^1 次元ベクトル空間なので、 $(n^1 + 1)$ 個の行列の集合 $\{E, A, A^2, \dots, A^{n^2}\}$ は線形従属である。したがって、

$$F(A) = A^m + \sum_{k=0}^{m-1} c_k A^k + c_0 E = O \quad (2.178)$$

となるような最小多項式 $F(t)$ が存在する。

また、 n 次正則行列 P に対して、

$$(P^{-1}AP)^n = P^{-1}A^nP \quad (2.179)$$

が成り立つので、

$$F(P^{-1}AP) = P^{-1}F(A)P = O \quad (2.180)$$

となる。これは、最小多項式が、基底の選び方に依らないことを示している。

したがって、最小多項式の定義を、線形写像まで拡張することができる。

Def. 線形写像の最小多項式

V を有限次元ベクトル空間、 $f: V \rightarrow V$ を線形写像とする。 f の最小多項式 $F(t)$ とは、最高次係数の 0 の多項式で、

$$F(f) = f^m + \sum_{k=0}^{m-1} c_k f^k + c_0 I = O \quad (2.181)$$

となるようなもののうち、次数が最小のものをいう。

Thm. ケーリーハミルトンの定理

線形変換 $f: V \rightarrow V$ に対して、 f の固有多項式 $F_f(t)$ に対して、 $F_f(f) = O$ が成り立つ。

Prf.

hoge

□

Thm.

線形変換 $f: V \rightarrow V$ に対する固有多項式 $F(t)$ とし、固有方程式 $F_f(t) = 0$ の相異なる解を $\lambda_1, \lambda_2, \dots, \lambda_r$ とする。このとき f の最小多項式 $M_f(t)$ は、 $(t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_r)$ で割り切れ、 $F_f(t)$ を割り切る。

Prf.

t の多項式 $F(t)$ に対して $F(f) = O$ ならば、 f の固有値 λ は $F(\lambda) = 0$ を満たすことを示す。

f の λ 固有ベクトルを \mathbf{v} とする。 $F(t) = t^n + \sum_{k=1}^{n-1} c_k t^k + c_0$ に対して、

$$\mathbf{0} = o(\mathbf{v}) \quad (2.182)$$

$$= F(f)(\mathbf{v}) \quad (\because F(f) = O) \quad (2.183)$$

$$= f^n(\mathbf{v}) + \sum_{k=1}^{n-1} c_k f^k(\mathbf{v}) + c_0 \mathbf{v} \quad (2.184)$$

$$= \lambda^n \mathbf{v} + \sum_{k=1}^{n-1} c_k \lambda^k \mathbf{v} + c_0 \mathbf{v} \quad (2.185)$$

$$= F(\lambda) \mathbf{v} \quad (2.186)$$

であるから、 $F(\lambda) = 0$ である。 よって、 $M_f(t)$ は f の固有値を解とする。

次に、 $F_f(t)$ が $M_f(t)$ で割り切れることを示す。 $F_f(t)$ を $M_f(t)$ で割った余りを $R(t)$ とすると、

$$F_f(t) = q(t)M_f(t) + R(t) \quad (2.187)$$

となる。 ただし、 $q(t), R(t) \in \mathbb{C}[t]$ で、 $\deg R(t) < \deg M_f(t)$ である。

個の両辺の t に f を代入すると、左辺は O であるから、右辺も O である。 また、 $M_f(f) = O$ なので、 $R(f) = O$ である。

最小多項式は次数が最小で最高次係数が 1 であるから、 $R(t) = 0$ でなくてはならない。 ($M(t)$ より次数が小さいのはまずい。) よって、 $R(t) = 0$ であり、 $F_f(t)$ は $M_f(t)$ で割り切れる。 \square

Def. 広義固有空間

線形変換 $f : V \rightarrow V$ の相異なる固有値全体を $\lambda_1, \lambda_2, \dots, \lambda_f$ とし、 f の最小多項式を、

$$M_f(t) = (t - \lambda_1)^{d_1} (t - \lambda_2)^{d_2} \dots (t - \lambda_f)^{d_f} \quad (2.188)$$

とする。 このとき、 f の固有値 λ_k の広義固有空間 V_{λ_k, d_k} を、

$$V_{\lambda_k, d_k} = \ker(f - \lambda_k I)^{d_k} \quad (2.189)$$

と定義する。 すなわち、

$$(f - \lambda_k I)^{d_k}(\mathbf{v}) = \mathbf{0} \quad (2.190)$$

となるベクトル \mathbf{v} からなる部分空間を V_{λ_k, d_k} とする。

気持ちとしては、広義固有空間に含まれる一般化固有ベクトルは、1 解 $(f - \lambda_k I)$ を作用させても 0 にならないかもしれないが、少なくとも d_k 回作用させると 0 になるということを意味する。 とくに、1 回作用させただけで 0 になる固有ベクトルは、通常の固有ベクトルと呼ばれる。

Thm.

線形変換 $f: V \rightarrow V$ の固有値 λ_k の広義固有空間 V_{λ_k, d_k} は、

$$f(V_{\lambda_k, d_k}) \subset V_{\lambda_k, d_k} \quad (2.191)$$

を満たす f の不変部分空間であり、 V はこれらの部分空間の直和

$$V = V_{\lambda_1, d_1} \oplus V_{\lambda_2, d_2} \oplus \cdots \oplus V_{\lambda_f, d_f} \quad (2.192)$$

と同型である。

Prf.

hoge

□

Cor.

線形変換 $f: V \rightarrow V$ に対して、以下の 3 条件は同値である。

1. V の基底で、 f の表現行列が対角行列になるものが存在する。
2. f の相異なる固有値 $\lambda_1, \lambda_2, \dots, \lambda_f$ に対して、

$$(f - \lambda_1 I)(f - \lambda_2 I) \cdots (f - \lambda_f I) = O \quad (2.193)$$

が成り立つ。

3. f の最小多項式は、重根を持たない。

Prf.

hoge

□

2.6 ランク

Def. ランク

$m \times n$ 行列 A に対して、 A のランクとは、

$$\text{rank} A = \dim \text{Im} A \quad (2.194)$$

で定義される。

Thm. ランクの特徴づけ

$m \times n$ 行列 A に対して、以下の (1) ~ (5) の数は一致する。

1. A のランク
2. A の列ベクトルのうち一次独立な最大個数
3. A の行ベクトルのうち一次独立な最大個数
4. A の t 次の小行列式の中には 0 でないものがあり、 $t+1$ 次以上の小行列式はすべて 0 であるような最大の t
5. A の行ベクトルが生成する部分空間の次元

Prf.

2.7 その他

Thm:

実 n 次元ベクトル空間 \mathbb{R}^n の任意の部分ベクトル空間 W_1, W_2 に対して、

$$W_1 \subset W_2 \Rightarrow \dim W_1 \leq \dim W_2 \quad (2.195)$$

が成立する。また、

$$W_1 \subset W_2 \quad \text{and} \quad \dim W_1 = \dim W_2 \quad \Rightarrow \quad W_1 = W_2 \quad (2.196)$$

が成立する。

Prf

3 微分積分学

4 確率統計

4.1 確率分布のまとめ

取り扱った確率分布をまとめる。

Def. 正規分布

平均を μ 、分散を σ^2 とする。確率分布関数が以下で与えられるとき、正規分布という。

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (4.1)$$

とくに、確率変数 X が正規分布に従うとき、

$$X \sim N(\mu, \sigma^2) \quad (4.2)$$

と書く。

Prop: 正規分布の標本平均

正規分布に従う確率変数 X_1, X_2, \dots, X_n の標本平均 $S_n = \frac{1}{n} \sum_{i=1}^n X_i$ は、正規分布に従う。

Prf

$$p(S_n = s) = \int_{\{\mathbf{x} \in \mathbb{R}^n : S_n(\mathbf{x}) = s\}} p(\mathbf{x}) \, d\mathbf{x} \quad (4.3)$$

$$= \int_{\mathbb{R}^n} \delta(S_n(\mathbf{x}) - s) p(\mathbf{x}) \, d\mathbf{x} \quad (4.4)$$

$$= \int_{\mathbb{R}^n} \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp(ik(S_n - s)) p(\mathbf{x}) \, d\mathbf{x} \, dk \quad (4.5)$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-iks} \left(\int_{\mathbb{R}^n} e^{ikS_n} p(\mathbf{x}) \, d\mathbf{x} \right) dk \quad (4.6)$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-iks} \left(\phi\left(\frac{k}{n}\right) \right)^n dk \quad (4.7)$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-iks} \exp\left(i\mu \frac{k}{n} - \frac{\sigma^2 k^2}{2n^2}\right)^n dk \quad (4.8)$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{\sigma^2}{2n} \left(k - n \frac{s - \mu}{\sigma^2}\right)^2 - \frac{n(s - \mu)^2}{2\sigma^2}\right) dk \quad (4.9)$$

$$= \sqrt{\frac{n}{2\pi\sigma^2}} \exp\left(-\frac{n(s - \mu)^2}{2\sigma^2}\right) \quad (4.10)$$

であるから、 S_n は正規分布 $N(\mu, \frac{\sigma^2}{n})$ に従う。 □

TODO ここに n 変数ガウス分布についての記述を追加する。

Def. 指数分布

指数分布とは、 $(0, \infty)$ に台を持つ確率分布で、確率密度関数が以下で与えられるものである。

$$p(x) = \lambda \exp(-\lambda x) \quad (4.11)$$

Def. アーラン分布

アーラン分布とは、 $(0, \infty)$ に台を持つ確率分布で、確率密度関数が以下で与えられるものである。

$$p(x) = \frac{1}{\Gamma(k)} \lambda^k x^{k-1} \exp(-\lambda x) \quad (4.12)$$

Prop: 指数分布とアーラン分布

n 個の独立な確率変数 X_1, X_2, \dots, X_n が指数分布に従うとき、その和 $Y = X_1 + X_2 + \dots + X_n$ はアーラン分布に従う。

Prf

数学的帰納法により示す。

(1) $n = 1$ のとき

このとき、

$$p(y) = \lambda \exp(-\lambda y) \quad (4.13)$$

であるから、指数分布に一致し、示された。

(2) $n = 2$ のとき

$$p_Y(y) = \int_0^y p_{X_1}(x) p_{X_2}(y-x) dx \quad (4.14)$$

$$= \int_0^y \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(y-x)} dx \quad (4.15)$$

$$= \int_0^y \lambda^2 e^{-\lambda y} \cdot e^{-\lambda x} \cdot e^{\lambda x} dx \quad (4.16)$$

$$= \lambda^2 e^{-\lambda y} \int_0^y dx \quad (4.17)$$

$$= \lambda^2 e^{-\lambda y} \cdot [x]_0^y \quad (4.18)$$

$$= \lambda^2 y e^{-\lambda y} \quad (4.19)$$

であるから、アーラン分布に一致し、示された。

(3) $n = k$ のとき成立すると仮定する。

$$p_Y(y) = \int_0^y p_{X_k}(x) p_{X_{k+1}}(y-x) dx \quad (4.20)$$

$$= \int_0^y \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \cdot \lambda e^{-\lambda(y-x)} dx \quad (4.21)$$

$$= \frac{\lambda^{k+1}}{(k-1)!} e^{-\lambda y} \int_0^y x^{k-1} e^{-\lambda x} \cdot e^{\lambda x} dx \quad (4.22)$$

$$= \frac{\lambda^{k+1} e^{-\lambda y}}{(k-1)!} \int_0^y x^{k-1} dx \quad (4.23)$$

$$= \frac{\lambda^{k+1} e^{-\lambda y}}{(k-1)!} \cdot \left[\frac{1}{k} x^k \right]_0^y \quad (4.24)$$

$$= \frac{\lambda^{k+1} y^k e^{-\lambda y}}{k \cdot (k-1)!} \quad (4.25)$$

$$= \frac{\lambda^{k+1} y^k e^{-\lambda y}}{k!} \quad (4.26)$$

であるから、 $n = k + 1$ のときも成立する。したがって、示された。 \square

4.2 漸近まわりの話

Thm:Markov の不等式

確率変数 \hat{X} に対して、任意の正数 $\epsilon > 0$ に対して、

$$P(|\hat{X}| \geq \epsilon) \leq \frac{\langle |\hat{X}| \rangle}{\epsilon} \quad (4.27)$$

が成立する。

Prf

新たに確率変数 θ を導入し、

$$\theta = \begin{cases} 1 & (|\hat{X}| \geq \epsilon) \\ 0 & (|\hat{X}| < \epsilon) \end{cases} \quad (4.28)$$

とする。このとき、明らかに、

$$\theta \leq \frac{|\hat{X}|}{\epsilon} \quad (4.29)$$

が成り立つ。^{*1}したがって、この関係式に p をかけて和をとることで、

$$\langle \theta \rangle \leq \frac{\langle |\hat{X}| \rangle}{\epsilon} \quad (4.30)$$

^{*1} グラフを書いてみると良い。

が成り立つ。ここで、左辺は、

$$\langle \theta \rangle = P(|\hat{X}| \geq \epsilon) \quad (4.31)$$

であるから、示された。 \square

Thm: チェビシェフの不等式

確率変数 \hat{X} に対して、任意の正数 $\epsilon > 0$ に対して、

$$P(|\hat{X} - \langle \hat{X} \rangle| \geq \epsilon) \leq \frac{\sigma^2(\hat{X})}{\epsilon^2} \quad (4.32)$$

が成立する。ただし、 $\sigma^2(\hat{X})$ は \hat{X} の分散である。

この結果は、確率分布によらず成り立つことに注意したい。すなわち、期待値と分散さえわかっているならば、確率変数の取りうる範囲が大体どの程度かがわかるということである。

また、ゆらぎが測定精度よりもはるかに小さいと仮定すると、

$$P(|\hat{X} - \langle \hat{X} \rangle| \geq \epsilon) \ll 1 \quad (4.33)$$

となる。これは、ゆらぎが小さいときは、確率変数を一回だけ測定した値が、期待値と一致することを表している。例えば、統計力学において、熱力学極限をとった時はこのような状況になる。

Prf

新たに確率変数 θ を導入し、

$$\theta_i = \begin{cases} 1 & (|\hat{X}_i - \langle \hat{X} \rangle| \geq \epsilon) \\ 0 & (|\hat{X}_i - \langle \hat{X} \rangle| < \epsilon) \end{cases} \quad (4.34)$$

とする。このとき、明らかに、

$$\theta_i \leq \frac{(\hat{X}_i - \langle \hat{X} \rangle)^2}{\epsilon^2} \quad (4.35)$$

が成り立つ。^{*2}したがって、この関係式に p_i をかけて和をとることで、

$$\langle \theta \rangle \leq \frac{\sigma^2(\hat{X})}{\epsilon^2} \quad (4.36)$$

が成り立つ。ここで、左辺は、

$$\langle \theta \rangle = P(|\hat{X} - \langle \hat{X} \rangle| \geq \epsilon) \quad (4.37)$$

^{*2} グラフを書いてみると良い。

であるから、示された。 □

Thm: 大数の法則

任意の正の実数 ϵ に対して、

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n \hat{X}_i - \langle \hat{X} \rangle \right| \geq \epsilon \right) = 0 \quad (4.38)$$

が成立する。

この定理は、数多くの系で同時に実験をしたときの結果の算術平均が、ほぼ確実に期待値に一致することを示している。

Prf

ある単独の系において、 \hat{X}, \hat{X}^2 の期待値をそれぞれ $\mu = \langle \hat{X} \rangle_{\mathbf{p}}, \nu = \langle \hat{X}^2 \rangle_{\mathbf{p}}$ とする。このとき、

$$\sigma_{\mathbf{p}}^2(\hat{X}) = \nu - \mu^2 \quad (4.39)$$

である。 N この系での \hat{X} の算術平均

$$\hat{m} = \frac{1}{N} \sum_{i=1}^N \hat{X}_i \quad (4.40)$$

を考える。このとき、線形性から、

$$\langle \hat{m} \rangle = \frac{1}{N} \sum_{i=1}^N \langle \hat{X}_i \rangle_{\tilde{\mathbf{p}}} \quad (4.41)$$

となる。また、独立性から、 $\langle \hat{X}_i \rangle_{\tilde{\mathbf{p}}} = \langle \hat{X} \rangle_{\mathbf{p}} = \mu$ である。したがって、

$$\langle \hat{m} \rangle = \mu \quad (4.42)$$

である。また、 \hat{m}^2 の期待値は、

$$\langle \hat{m}^2 \rangle_{\tilde{\mathbf{p}}} = \frac{1}{N^2} \sum_{j,k} \langle \hat{X}_j \hat{X}_k \rangle_{\tilde{\mathbf{p}}} \quad (4.43)$$

$$= \frac{1}{N^2} (N\nu + (N^2 - N)\mu^2) \quad (4.44)$$

となる。したがって、

$$\sigma_{\tilde{\mathbf{p}}}(\hat{m}) = \sqrt{\left\langle \frac{\nu - \mu^2}{N} \right\rangle} \quad (4.45)$$

となる。したがって、

$$\langle \hat{m} \rangle_{\mathbf{p}} = \mu, \quad \sigma_{\mathbf{p}}(\hat{m}) = \frac{\sigma_{\mathbf{p}}(\hat{X})}{\sqrt{N}} \quad (4.46)$$

となる。チェビシェフの不等式より、

$$P(|\hat{m} - \mu| \geq \epsilon) \leq \left(\frac{\sigma_{\mathbf{p}}(\hat{X})}{\epsilon} \right)^2 \frac{1}{N} \quad (4.47)$$

となる。したがって、極限をとって、

$$\lim_{N \rightarrow \infty} P(|\hat{m} - \mu| \geq \epsilon) = 0 \quad (4.48)$$

が示された。 \square

Def: モーメント母関数/キュムラント母関数

確率分布 $p(x)$ のモーメント母関数は、以下で定義される。

$$M(t) = \langle e^{tx} \rangle = \int e^{tx} p(x) dx \quad (4.49)$$

$$= \sum_{n=0}^{\infty} \frac{M_n}{n!} t^n \quad (4.50)$$

このとき、 M_n をモーメントという。

また、キュムラント母関数は、以下で定義される。

$$K(t) = \log M(t) = \log \langle e^{tx} \rangle \quad (4.51)$$

$$= \sum_{n=0}^{\infty} \frac{C_n}{n!} t^n \quad (4.52)$$

このとき、 C_n をキュムラントという。

このとき、確率分布と母関数は、一対一に対応することが示される。

Thm: モーメントとキュムラントの関係

モーメントとキュムラントは以下のように関係する。

$$C_1 = M_1 \quad (4.53)$$

$$C_2 = M_2 - M_1^2 \quad (4.54)$$

$$C_3 = M_3 - 3M_1M_2 + 2M_1^3 \quad (4.55)$$

Prf

キュムラント母関数を展開すると、

$$\begin{aligned}\ln \langle e^{s\hat{X}} \rangle &= \ln \left\langle 1 + \frac{1}{1!} s\hat{X} + \frac{1}{2!} s^2 \hat{X}^2 + \frac{1}{3!} s^3 \hat{X}^3 + O(s^4) \right\rangle \\ &= \frac{1}{1!} \langle \hat{X} \rangle s + \left(\frac{1}{2!} \langle \hat{X}^2 \rangle - \frac{1}{2} \langle \hat{X} \rangle^2 \right) s^2\end{aligned}\quad (4.56)$$

$$\begin{aligned}&+ \left(\frac{1}{3!} \langle \hat{X}^3 \rangle - \frac{1}{2} \cdot 2 \langle \hat{X} \rangle \cdot \frac{1}{2!} \langle \hat{X}^2 \rangle + \frac{1}{3!} \langle \hat{X} \rangle^3 \right) s^3 + O(s^4) \\ &= \frac{\langle \hat{X} \rangle}{1!} s + \left(\frac{\langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2}{2!} \right) s^2 + \left(\frac{\langle \hat{X}^3 \rangle - 3 \langle \hat{X} \rangle \langle \hat{X}^2 \rangle + 2 \langle \hat{X} \rangle^3}{3!} \right) s^3 + O(s^4)\end{aligned}\quad (4.57)$$

となることから、示された。 \square

上の関係からわかるように、キュムラント母関数を二階微分することで分散が出てくる。

Def: 大偏差原理

確率分布 $P_N(x)$ について、レート関数と呼ばれる関数 $I(x)$ が存在して、

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_N(x) = I(x) \quad (4.58)$$

を任意の x に対して満たすとき、大偏差原理が成り立つという。

Thm: ゲルトナー・エリスの定理

確率変数密度 x に対するスケールされたキュムラント母関数 $q(k)$ が存在し、それが実数 $k \in \mathbb{R}$ で微分可能であるとする。このとき、 x は大偏差原理を満たし、そのレート関数は以下のルジャンドル変換で与えられる。

$$I(x) = \sup_k \{kx - q(k)\} \quad (4.59)$$

また、逆ルジャンドル変換により、

$$q(k) = \sup_b \{kb - I(b)\} \quad (4.60)$$

が成り立つ。

Gärtner-Ellis の定理の直感的な理解を試みる。 A_n に対して大偏差原理が成り立つと仮定すると、

$$P(A_n \in da) \asymp e^{-nI(a)} da \quad (4.61)$$

となる。このとき、

$$\langle e^{nkA_n} \rangle = \int_{\mathbb{R}} e^{nka} e^{-nI(a)} da \quad (4.62)$$

$$\sim \sup_a \{ka - I(a)\} \quad \because \text{鞍点法} \quad (4.63)$$

したがって、

$$\lambda(k) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \langle e^{nkA_n} \rangle = \sup_a \{ka - I(a)\} \quad (4.64)$$

となる。逆ルジャンドル変換によって

$$I(a) = \sup_k \{ka - \lambda(k)\} \quad (4.65)$$

が得られる。

Thm. 中心極限定理

ある集団から n 個の標本を取り出し、その平均を \hat{X} とする。このとき、 n が十分大きいとき、 \hat{X} は、平均 μ 、分散 $\frac{\sigma^2}{n}$ の正規分布に従う。すなわち、

$$\hat{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad (4.66)$$

が成り立つ。

5 幾何学

6 ラプラス変換

Def: ラプラス変換

関数 $f(t)$ のラプラス変換は、以下で定義される。

$$F(s) = \mathcal{L}[f(t)] = \int_0^{\infty} f(t) e^{-st} dt \quad (6.1)$$

ラプラス変換には逆変換が存在する。

Thm: ラプラス逆変換

関数 $F(s)$ の逆ラプラス変換は、以下で定義される。

$$f(t) = \mathcal{L}^{-1}[F(s)] = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} F(s) e^{st} ds \quad (6.2)$$

ただし、 γ は右半平面における収束領域を表す。

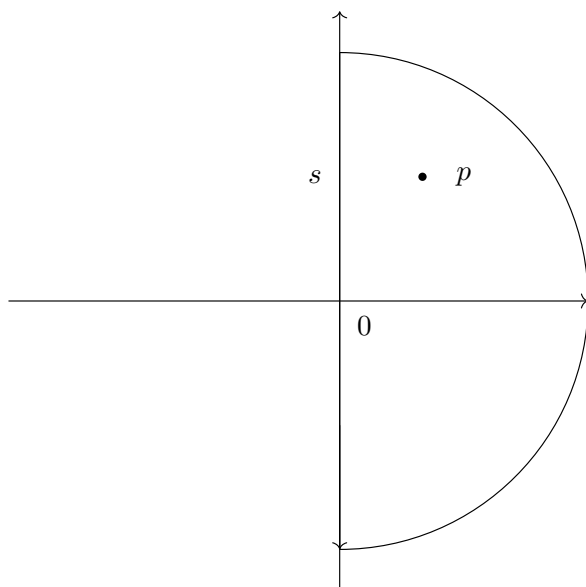


図 1 逆ラプラス変換の複素積分の積分経路 C .

Prf

コーシーの積分公式より、正則な領域の点 p について、

$$F(p) = \frac{1}{2\pi i} \oint_C \frac{F(s)}{s-p} ds \quad (6.3)$$

が成り立つ。いま、この線積分で生き残るのは、 $\text{Re}(s) = \gamma$ 上の経路のみである。ここで、

$$\frac{1}{s-p} = \int_0^\infty e^{-(p-s)t} dt \quad \because \text{Re}(p) > \text{Re}(s) \quad (6.4)$$

であるから、

$$F(p) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} F(s) \int_0^\infty e^{-(s-p)t} dt ds \quad (6.5)$$

$$= \int_0^\infty \left(\frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} F(s) e^{st} ds \right) e^{-pt} dt \quad (6.6)$$

$$= \int_0^\infty f(t) e^{-pt} dt \quad (6.7)$$

が成り立つ。したがって、

$$f(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} F(s) e^{st} ds \quad (6.8)$$

が示された。 □

7 その他

7.1 計算の類

ガウス積分

ガウス積分の変わったやり方を書く。

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = - \int_{-\infty}^{\infty} \frac{d}{da} e^{-ax^2} dx \quad (7.1)$$

$$= - \frac{d}{da} \int_{-\infty}^{\infty} e^{-ax^2} dx \quad (7.2)$$

$$= - \frac{d}{da} \sqrt{\frac{\pi}{a}} \quad (7.3)$$

$$= \frac{\sqrt{\pi}}{2a^{3/2}} \quad (7.4)$$

により計算できる。

負の二項定理

$$(1-x)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{n-1} x^k \quad (7.5)$$

が成り立つことが知られている。

Prf.

(1) $n = 1$ のとき

$$(1-x)^{-1} = \sum_{k=0}^{\infty} x^k \quad (7.6)$$

が成り立つ。

(2) $n = m$ のとき成立すると仮定する。このとき、

$$(1-x)^{-m} = \sum_{k=0}^{\infty} \binom{m+k-1}{m-1} x^k \quad (7.7)$$

が成り立つ。両辺微分することで、

$$m(1-x)^{-m-1} = \sum_{k=0}^{\infty} \binom{m+k}{m-1} (k+1) x^k \quad (7.8)$$

となる。ここで、

$$\frac{k+1}{m} \binom{m+k}{m-1} = \frac{k+1}{m} \frac{(m+k)!}{(k+1)!(m-1)!} \quad (7.9)$$

$$= \frac{(m+k)!}{k!m!} \quad (7.10)$$

$$= \binom{m+k}{m} \quad (7.11)$$

であるから、

$$(1-x)^{-m-1} = \sum_{k=0}^{\infty} \binom{m+k}{m} x^k \quad (7.12)$$

が成り立つ。したがって、 $n = m + 1$ のときも成立する。したがって、帰納的に示された。 \square

7.2 ルジャンドル変換

$f(x)$ を $x_{\min} < x < x_{\max}$ で定義された下に凸な関数とする。また、

$$\alpha_{\min} = \lim_{x \rightarrow x_{\min}} f'(x) \quad (7.13)$$

$$\alpha_{\max} = \lim_{x \rightarrow x_{\max}} f'(x) \quad (7.14)$$

とする。 $\alpha_{\min} < \alpha < \alpha_{\max}$ を固定する。任意の x_0 について、 $(x_0, f(x_0))$ を通る傾き α の直線を考える。直線の方程式は、

$$y = \alpha(x - x_0) + f(x_0) = \alpha(x - x_0) + y_0 \quad (7.15)$$

このもとで、 x_0 を動かすことにより、切片の最小値を探す。この最小値は α に依ってのみ決まる。この最小値を y_0^* とし、 $f^*(\alpha) = -y_0^*$ とする。このとき、ルジャンドル変換を以下によって定義する。

Def. ルジャンドル変換

関数 $f(x)$ のルジャンドル変換は、以下で表される。

$$f^*(\alpha) = -\min_x \{f(x) - \alpha x\} = \max_x \{\alpha x - f(x)\} \quad (7.16)$$

すなわち、 $f(x)$ のグラフに対して、とくに微分可能な場合についてはその接線を引き、その切片としてルジャンドル変換を導入する。また、このルジャンドル変換は接線の傾きの関数となる。

Thm. ルジャンドル変換の存在

任意の下に凸な関数 $f(x)$ に対して、ルジャンドル変換 $f^*(\alpha)$ が存在する。また、 $f^*(\alpha)$ は下に凸であり。連続である。

また、Young の不等式

$$f(x) + f^*(\alpha) \geq \alpha x \quad (7.17)$$

が成り立つ。

Prf.

前半の存在性については絵を書けばわかる。実際、微分可能な場合については接線を連続的に動かしていけば良い。また、微分可能な点があっても成り立つ。Gibbs 自由エネルギーなどで試すと良い。

Young の不等式は簡単に示せるので示しておく。ルジャンドル変換の定義より、

$$f(x) + f^*(\alpha) = \max_y \{f(x) + \alpha y - f(y)\} \geq \alpha x \quad (7.18)$$

が成り立つ。したがって、示された。 \square

Young の不等式より、

$$f(x) \geq \alpha x - f^*(\alpha) \quad (7.19)$$

が成り立つ。また、少なくともどこかで等号が成り立つことから、^{*3}

$$f(x) = \max_{\alpha} \{\alpha x - f^*(\alpha)\} \quad (7.20)$$

が成り立つ。この関係式は、ルジャンドル変換を二回施すと、元の関数に戻ることを示している。

7.3 ベクトル解析

7.3.1 準備

Levi-Civita 記号

^{*3} ルジャンドル変換が \sup などではなく \max で定義されていることに注意。

Def. Levi-Civita 記号

Levi-Civita 記号 $\epsilon_{\mu_1\mu_2\cdots\mu_n}$ は、以下で定義される。

$$\epsilon_{\mu_1\mu_2\cdots\mu_n} = \text{sgn}(\mu_1\mu_2\cdots\mu_n) \quad (7.21)$$

特に、三次元の場合は、スカラー三重積による定義も便利である。

$$\epsilon_{ijk} = \mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k) = \det \begin{pmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \mathbf{e}_3^\top \end{pmatrix} \quad (7.22)$$

である。

Prop. Levi-Civita 記号の積 (3 次元)

3 次元の Levi-Civita 記号について、以下が成り立つ。

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmn} &= \begin{vmatrix} \mathbf{e}_i^\top \\ \mathbf{e}_j^\top \\ \mathbf{e}_k^\top \end{vmatrix} \begin{vmatrix} \mathbf{e}_l & \mathbf{e}_m & \mathbf{e}_n \end{vmatrix} \\ &= \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix} \\ &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \end{aligned} \quad (7.23)$$

$$- \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kn} \quad (7.24)$$

Prf.

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmn} &= [(\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k] [(\mathbf{e}_l \times \mathbf{e}_m) \cdot \mathbf{e}_n] \\ &= \begin{vmatrix} \mathbf{e}_i^\top \\ \mathbf{e}_j^\top \\ \mathbf{e}_k^\top \end{vmatrix} \begin{vmatrix} \mathbf{e}_l & \mathbf{e}_m & \mathbf{e}_n \end{vmatrix} \\ &= \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix} \\ &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \end{aligned} \quad (7.25)$$

$$- \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kn} \quad (7.26)$$

であるから、示された。 \square

系として、以下の公式が手に入る。

Cor.

以下が成り立つ。

$$\sum_m \epsilon_{mij} \epsilon_{mkl} = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk} \quad (7.27)$$

$$\sum_{m,n} \epsilon_{mni} \epsilon_{mnj} = 2\delta_{ij} \quad (7.28)$$

Prf.

$$\begin{aligned} \epsilon_{mij} \epsilon_{mkl} &= \begin{vmatrix} \delta_{mm} & \delta_{mk} & \delta_{ml} \\ \delta_{im} & \delta_{ik} & \delta_{il} \\ \delta_{jm} & \delta_{jk} & \delta_{jl} \end{vmatrix} \\ &= \sum_{m=1}^3 \delta_{mm} \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} - \sum_{m=1}^3 \delta_{im} \begin{vmatrix} \delta_{mk} & \delta_{ml} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} + \sum_{m=1}^3 \delta_{jm} \begin{vmatrix} \delta_{mk} & \delta_{ml} \\ \delta_{ik} & \delta_{il} \end{vmatrix} \\ &= 3 \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} - \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} + \begin{vmatrix} \delta_{jk} & \delta_{ml} \\ \delta_{ik} & \delta_{il} \end{vmatrix} \\ &= \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} \\ &= \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk} \end{aligned} \quad (7.29)$$

$$\begin{aligned} \epsilon_{mni} \epsilon_{mnj} &= \begin{vmatrix} \delta_{mm} & \delta_{mn} & \delta_{mj} \\ \delta_{nm} & \delta_{nn} & \delta_{nj} \\ \delta_{im} & \delta_{in} & \delta_{ij} \end{vmatrix} \\ &= \sum_{m,n=1}^3 \delta_{mm} \begin{vmatrix} \delta_{nn} & \delta_{nj} \\ \delta_{in} & \delta_{ij} \end{vmatrix} - \sum_{m,n=1}^3 \delta_{nm} \begin{vmatrix} \delta_{mn} & \delta_{mj} \\ \delta_{in} & \delta_{ij} \end{vmatrix} + \sum_{m,n=1}^3 \delta_{im} \begin{vmatrix} \delta_{mn} & \delta_{mj} \\ \delta_{nn} & \delta_{nj} \end{vmatrix} \\ &= \sum_{n=1}^3 3(\delta_{nn} \delta_{ij} - \delta_{in} \delta_{nj}) - \sum_{n=1}^3 (\delta_{nn} \delta_{ij} - \delta_{in} \delta_{nj}) + \sum_{n=1}^3 (\delta_{in} \delta_{nj} - \delta_{nn} \delta_{ij}) \\ &= \sum_{n=1}^3 (\delta_{nn} \delta_{ij} - \delta_{in} \delta_{nj}) \\ &= 3\delta_{ij} - \delta_{ij} \\ &= 2\delta_{ij} \end{aligned} \quad (7.30)$$

□

7.3.2 微分形式

ベクトル解析の公式を真正面から計算するのは面倒なので、微分形式を用いて計算する例を以下に示す。

grad

$grad$ は、0-form の外微分によって得られる。

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \simeq \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \nabla f \quad (7.31)$$

rot rot は、1-form の外微分によって得られる。

$$dA = d(A_x dx + A_y dy + A_z dz) \quad (7.32)$$

$$= \left(\frac{\partial A_x}{\partial y} dy + \frac{\partial A_x}{\partial z} dz \right) \wedge dx + \left(\frac{\partial A_y}{\partial x} dx + \frac{\partial A_y}{\partial z} dz \right) \wedge dy + \left(\frac{\partial A_z}{\partial x} dx + \frac{\partial A_z}{\partial y} dy \right) \wedge dz \quad (7.33)$$

$$= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) dy \wedge dz + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) dz \wedge dx + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dx \wedge dy \quad (7.34)$$

$$\simeq \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \nabla \times A \quad (7.35)$$

div

div は、2-form の外微分によって得られる。

$$dA = d(A_x dy \wedge dz + A_y dz \wedge dx + A_z dx \wedge dy) \quad (7.36)$$

$$= \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx \wedge dy \wedge dz \quad (7.37)$$

$$\simeq \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \nabla \cdot A \quad (7.38)$$

以上の話から、容易に以下が示される。

Prop.

以下が成立する。

$$\operatorname{div}(\operatorname{rot} A) = 0 \quad (7.39)$$

$$\operatorname{rot}(\operatorname{grad} f) = 0 \quad (7.40)$$

Prf.

$$\operatorname{div}(\operatorname{rot} A) \simeq d(dA) = 0 \quad (7.41)$$

$$\operatorname{rot}(\operatorname{grad} f) \simeq d(df) = 0 \quad (7.42)$$

が成り立つ。したがって、示された。 \square

他にも様々な公式を示すことができる。

$$\operatorname{div}(A \times B) \simeq d(A \wedge B) \quad (7.43)$$

$$= dA \wedge B - A \wedge dB \quad (7.44)$$

$$\simeq \operatorname{rot} A \cdot B - A \cdot \operatorname{rot} B \quad (7.45)$$

などが成り立つ。

Def. ラプラシアン

ラプラシアンは、以下のように定義される。

$$\Delta = -(\delta d + d\delta) \quad (7.46)$$

ただし、 δ は余微分

$$\delta = (-1)^{n(p+1)+1} \star d \star \quad (7.47)$$

である。(p-form, n 次元)

特に、今は三次元について考えているので、余微分は、

$$\delta = (-1)^{3p+4} \star d \star = (-1)^p \star d \star \quad (7.48)$$

このとき、以下の公式を示すことができる。

Prop.

$$\Delta = \nabla \cdot (\nabla \mathbf{E}) - \nabla \times (\nabla \times \mathbf{E}) \quad (7.49)$$

Prf.

$$\Delta E = (-d \star d \star + \star d \star d) \mathbf{E} \quad (7.50)$$

を計算する。第一項について、

$$E = E_x dx + E_y dy + E_z dz \quad (7.51)$$

$$\star E = E_x dy \wedge dz + E_y dz \wedge dx + E_z dx \wedge dy \quad (7.52)$$

$$d \star E = \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) dx \wedge dy \wedge dz \quad (7.53)$$

$$\star d \star E = \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) \quad (7.54)$$

$$d \star d \star E = d \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) \quad (7.55)$$

$$\simeq \nabla(\nabla \cdot \mathbf{E}) \quad (7.56)$$

第二項について、

$$dE = (\text{rot} E)_x dy \wedge dz + (\text{rot} E)_y dz \wedge dx + (\text{rot} E)_z dx \wedge dy \quad (7.57)$$

$$\star dE = (\text{rot} E)_x dx + (\text{rot} E)_y dy + (\text{rot} E)_z dz \quad (7.58)$$

$$d \star dE = \text{rot}(\text{rot} E) \quad (7.59)$$

したがって、

$$\Delta E = \nabla \cdot (\nabla \mathbf{E}) - \nabla \times (\nabla \times \mathbf{E}) \quad (7.60)$$

が成り立つ。したがって、示された。 \square

p-form の内積 p-form の内積を定義しておく。

Def. p-form の内積

p-form の $\mathbf{u} = \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p$ と $\mathbf{v} = \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_p$ の内積は、

$$(\mathbf{u}, \mathbf{v}) = \begin{vmatrix} (\mathbf{x}_1, \mathbf{y}_1) & \cdots & (\mathbf{x}_1, \mathbf{y}_p) \\ \vdots & \ddots & \vdots \\ (\mathbf{x}_p, \mathbf{y}_1) & \cdots & (\mathbf{x}_p, \mathbf{y}_p) \end{vmatrix} \quad (7.61)$$

で定義される。

これを用いて、Hodge star 作用素を定義する。

Def. Hodge star 作用素

p-form の Hodge star 作用素は、

$$\mathbf{u} \wedge \mathbf{v} = (\star \mathbf{u}, \mathbf{v}) dx_1 \wedge \cdots \wedge dx_n \quad (7.62)$$

で定義される。ただし、 \mathbf{v} は、任意の $(n-p)$ -form である。

実際にこれを用いて、 \mathbb{R}^3 について考えてみる。

$$dx \wedge \mathbf{v} = (\star dx, \mathbf{v}) dx \wedge dy \wedge dz \quad (7.63)$$

について、特に $\mathbf{v} = dy \wedge dz$ について考えれば十分だから、

$$dx \wedge dy \wedge dz = (\star dx, dy \wedge dz) dx \wedge dy \wedge dz \quad (7.64)$$

となり、内積の部分が 1 になってほしいので、

$$\star dx = dy \wedge dz \quad (7.65)$$

となる。また、

$$(dx \wedge dy) \wedge \mathbf{v} = (\star(dx \wedge dy), \mathbf{v}) dx \wedge dy \wedge dz \quad (7.66)$$

について、特に、 $\mathbf{v} = dz$ について考えれば十分だから、

$$(dx \wedge dy) \wedge dz = (\star(dx \wedge dy), dz) dx \wedge dy \wedge dz \quad (7.67)$$

となり、内積の部分が 1 になってほしいので、

$$\star(dx \wedge dy) = dz \quad (7.68)$$

となる。0-form, 3-form についても同様に計算することで、

$$\star 1 = dx \wedge dy \wedge dz \quad (7.69)$$

$$\star(dx \wedge dy \wedge dz) = 1 \quad (7.70)$$

も得られる。

上と同様の定義だが

$$(\star A)_{\nu_1 \nu_2 \dots \nu_\ell} = \frac{\sqrt{|\det g|}}{k!} \epsilon_{\nu_1 \nu_2 \dots \nu_\ell}{}^{\mu_1 \mu_2 \dots \mu_k} g^{\mu_1 \rho_1} g^{\mu_2 \rho_2} \dots g^{\mu_k \rho_k} A_{\rho_1 \rho_2 \dots \rho_k} \quad (7.71)$$

としてもよい。(前野先生の pdf 参照)

電磁気学への応用

せっかくなので、Maxwell 方程式を微分形式の言葉で書いてみる。まずは、3+1 座標の立場で (すなわち、時間は単なるパラメータとみなして) 書き直してみる。Maxwell 方程式は、

$$\begin{cases} \operatorname{div} \mathbf{D} = \rho \\ \operatorname{div} \mathbf{B} = 0 \\ \operatorname{rot} \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \operatorname{rot} \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{j} \end{cases} \quad (7.72)$$

である。これを微分形式の言葉で書き直す。今、 \mathbf{E} および \mathbf{H} は 1-form とし、 \mathbf{D} および \mathbf{B} は 2-form とする。このとき、

$$\begin{cases} d\mathbf{D} = \rho \\ d\mathbf{B} = 0 \\ d\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\ d\mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{j} \end{cases} \quad (7.73)$$

が成り立つ。 div や rot がすべて d に置き換わって少し気分が良い形となる。特に、静的な場合については、

$$\begin{cases} d\mathbf{D} = \rho \\ d\mathbf{B} = 0 \\ d\mathbf{E} = 0 \\ d\mathbf{H} = \mathbf{j} \end{cases} \quad (7.74)$$

となる。ここで、ベクトルポテンシャルを

$$\mathbf{B} = d\mathbf{A} \quad (7.75)$$

と定義すると、

$$d\mathbf{B} = d d\mathbf{A} = 0 \quad (7.76)$$

が自明に成立する。また、ベクトルポテンシャルに $d\mathbf{A}$ だけの任意性があることもすぐ確かめられる。

以上の議論を次は Minkowski 空間に拡張し、時間を位置と対等に扱ってみる。ベクトルポテンシャルを以下のように拡張する。

$$\mathbf{A} = A_t dt + A_x dx + A_y dy + A_z dz \quad (7.77)$$

ただし、 $A_t = -\phi$ とする。このとき、

$$F = d\mathbf{A} \quad (7.78)$$

$$= \left(-\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t} \right) dx \wedge dt + \left(-\frac{\partial \phi}{\partial y} - \frac{\partial A_y}{\partial t} \right) dy \wedge dt + \left(-\frac{\partial \phi}{\partial z} - \frac{\partial A_z}{\partial t} \right) dz \wedge dt \quad (7.79)$$

$$+ \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) dy \wedge dz + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) dz \wedge dx + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dx \wedge dy \quad (7.80)$$

$$\mathbf{E} = \nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \text{ を用いて} \quad (7.81)$$

$$= E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt \quad (7.82)$$

$$B_x \wedge dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy \quad (7.83)$$

となる。また、

$$dF = d(E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt) \quad (7.84)$$

$$+ d(B_x \wedge dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy) \quad (7.85)$$

$$(7.86)$$

は恒等的に 0 となる。これについて、 $dx \wedge dy \wedge dt$ の項は、

$$\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} + \frac{\partial B_z}{\partial t} \right) dx \wedge dy \wedge dt \quad (7.87)$$

となる。したがって、 $\text{rot} E + \frac{\partial B}{\partial t} = 0$ が成り立つ。

また、 $dx \wedge dy \wedge dz$ の項は、

$$\left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) dx \wedge dy \wedge dz \quad (7.88)$$

となる。したがって、 $\nabla \cdot B = 0$ が成り立つ。

さらに、 F の Hodge star を考えると、

$$\star F = -B_x dx \wedge dt - B_y dy \wedge dt - B_z dz \wedge dt \quad (7.89)$$

$$+ E_x dy \wedge dz + E_y dz \wedge dx + E_z dx \wedge dy \quad (7.90)$$

となる。これの外微分を計算すると、

$$d \star F = \left(\frac{\partial B_x}{\partial y} - \frac{\partial B_y}{\partial x} + \frac{\partial E_z}{\partial t} \right) dx \wedge dy \wedge dt \quad (7.91)$$

$$+ \left(\frac{\partial B_y}{\partial z} - \frac{\partial B_z}{\partial y} + \frac{\partial E_x}{\partial t} \right) dy \wedge dz \wedge dt \quad (7.92)$$

$$+ \left(\frac{\partial B_z}{\partial x} - \frac{\partial B_x}{\partial z} + \frac{\partial E_y}{\partial t} \right) dz \wedge dx \wedge dt \quad (7.93)$$

$$+ \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) dx \wedge dy \wedge dz \quad (7.94)$$

となる。したがって、

$$j = \frac{\rho}{\epsilon} dx \wedge dy \wedge dz - \mu j_x dy \wedge dz \wedge dt - \mu j_y dz \wedge dx \wedge dt - \mu j_z dx \wedge dy \wedge dt \quad (7.95)$$

とすれば、 $\text{rot} H - \frac{\partial D}{\partial t} = j$ および $\nabla \cdot D = \rho$ が成り立つ。

以上より、Maxwell 方程式は、

$$d \star F = j \quad (7.96)$$

と書くことができる。

7.4 参考にしたもの

- <https://www.eng.niigata-u.ac.jp/~nomoto/download/df.pdf>
- <http://irobutsu.a.la9.jp/kougi/DFC.pdf>

第 II 部

物理

8 古典力学

9 電磁気学

10 熱力学

熱力学を再構成する。Lieb 流にエントロピーを導入して、残りは清水っぽくまとめたい。

Def: 平衡状態/状態空間

平衡状態とは、状態空間の点である。状態空間とは、マクロ変数の値を座標軸とする \mathbb{R}^n の部分空間であり、 Γ で表す。

Axiom1

系を長時間放置すると、系は平衡状態に達し、状態空間上の点で表される。

以下、平衡状態のことを単に状態と書くことにする。

Def: 断熱到達可能性

状態 $X \in \Gamma$ から状態 $Y \in \Gamma$ へ、断熱到達可能な時、

$$X \prec Y \quad (10.1)$$

と書く。ただし、状態 $X \in \Gamma$ から状態 $Y \in \Gamma$ へ、断熱到達可能であるとは、外部装置 D の状態が遷移前後で変化せず、その装置を動かす力学装置のみが変化すること、すなわち、

$$(X, D, h) \rightarrow (Y, D, h') \quad (10.2)$$

が成り立つことである。

この定義のもと、遷移の公理を以下に示す。

Axiom2

断熱到達可能性に、以下の性質を要請する。

(A1: 反射律) $X \overset{A}{\sim} X$

(A2: 推移律) $X \prec Y$ かつ $Y \prec Z$ ならば、 $X \prec Z$

(A3: 一貫性) $X \prec X'$ かつ $Y \prec Y'$ ならば、 $(X, Y) \prec (X', Y')$

(A4: スケーリング不変性) $X \prec Y$ ならば、 $tX \prec tY$

(A5: 分割と結合) $0 < t < 1$ に対して、 $X \overset{A}{\sim} (tX, (1-t)X)$

(A6: 安定性) 0 に近づく ϵ の列およびいくつかの状態 Z_0, Z_1 に対して、 $(X, \epsilon Z_0) \prec (Y, \epsilon Z_1)$ ならば、 $X \prec Y$

(A7: 凸性) 状態空間が凸であるとき、 $(tX, (1-t)Y) \prec tX + (1-t)Y$

いくつかの補足をする。

(A3) について、一般には逆向きの矢印は成り立たない。というのも、系を熱的に接触させることで、片方の系の温度を下げるのが可能であるが、そのような接触なしでは温度を下げることはできないからである。すなわち、複合系は、互いに相互作用させることができるという点で、取りうる操作の種類が多い。

(A6) について、この公理は、系に、小さな外部系がくっついているときは、その外部系を無視して、着目系のみを考えてよいということを示している。

(A7) について、この公理は、遷移前後で、エネルギーや体積の和が変化しないことを表して

いる。

Thm: キャンセル法則

X, Y, Z が系の状態のとき、

$$(X, Z) \prec (Y, Z) \Rightarrow X \prec Y \quad (10.3)$$

が成立する。

Prf

$\epsilon = \frac{1}{n}$ とする。このとき、

$$(X, \epsilon Z) \stackrel{A}{\sim} ((1 - \epsilon)X, \epsilon X, Z) \quad (10.4)$$

$$\prec ((1 - \epsilon)X, \epsilon Y, Z) \quad (10.5)$$

$$\stackrel{A}{\sim} ((1 - 2\epsilon)X, \epsilon X, \epsilon Y, Z) \quad (10.6)$$

$$\prec ((1 - 2\epsilon)Y, 2\epsilon Y, Z) \quad (10.7)$$

これを $n = \frac{1}{\epsilon}$ 回繰り返すと、 X の係数は 0 に、 Y の係数は 1 に収束する。

以上より、 $X \prec Y$ が示された。 □

Def: エントロピー

状態空間 Γ 上のある X_0, X_1 に対して、 $X_0 \prec\prec X_1$ であるとする。このとき、

$$S_\Gamma(X) = s^* \sup\{\lambda | ((1 - \lambda)X_0, \lambda X_1) \prec X\} \quad (10.8)$$

で定義される関数 S_Γ を、基準点 $X_0 \prec\prec X_1$ に関するエントロピーという。ただし、 s^* は単位エントロピーである。

このエントロピーの意味するところは、

1. まず、二つの基準系を $\lambda : 1 - \lambda$ の比で用意して、複合系とする。
2. 複合系全体に対して、断熱操作 (それぞれの系に独立に作用させても、熱的接触のように二つの系を相互作用させても良い) を行う。
3. このとき、複合系が状態 X に達するときの λ の最大値がエントロピーである。

となる。

Thm: エントロピー原理

以下の (1),(2) は同値である。

(1) 関係 \prec が、 Γ の多重スケーリングコピーにおいて、順序の公理 A1-A6 及び比較仮説 (CH) を満たす。

(2) ある関数 $S_\Gamma : \Gamma \rightarrow \mathbb{R}$ が存在して、以下の性質を満たす。

もし、 $(t_1 + \cdots + t_n) = t'_1 + \cdots + t'_m$ であるとき、

$$(t_1 Y_1, \cdots, t_n Y_n) \prec (t'_1 Y'_1, \cdots, t'_m Y'_m) \Leftrightarrow S_\Gamma(Y_1) + \cdots + S_\Gamma(Y_n) \leq S_\Gamma(Y'_1) + \cdots + S_\Gamma(Y'_m) \quad (10.9)$$

Prf

証明は気が向いたら書く。

10.1 その他

10.1.1 様々な操作

自分のメモ用に、操作についてまとめておく。

熱的接触

田崎熱問題 5.1 を解いたので、メモとして残しておく。

二つの系 $(T, X), (T', Y)$ を、示量変数を固定したまま接触させて、温度 \tilde{T} に達するとき、全系について、

$$U(T, X) + U(T', Y) = U(\tilde{T}, X) + U(\tilde{T}, Y) \quad (10.10)$$

が成立し、温度は一意に定まる。

また、このとき $T < \tilde{T} < T'$ を仮定すると、熱の移動は、

$$Q = U(\tilde{T}, X) - U(T, X) = -(U(\tilde{T}, Y) - U(T', Y)) \quad (10.11)$$

となる。

10.1.2 熱機関の効率の上限

Thm: カルノーの定理

任意の熱機関の効率の上限は、

$$\eta \leq 1 - \frac{T_L}{T_H} \quad (10.12)$$

と書くことができる。ただし、 η は熱効率、 T_L, T_H はそれぞれ低温熱浴の温度、高温熱浴の温度である。

Prf

着目系、高温熱浴、低温熱浴の複合系を考える。このとき、複合経緯の断熱遷移を、

$$((T, X), (T_H, X_H), (T_L, X_L)) \rightarrow ((T, X), (T'_H, X_H), (T'_L, X_L)) \quad (10.13)$$

とする。このとき、エントロピー原理より、

$$S((T, X), (T_H, X_H), (T_L, X_L)) \leq S((T, X), (T'_H, X_H), (T'_L, X_L)) \quad (10.14)$$

が成り立つ。ここで、エントロピーの相加性から、

$$S(T, X) + S(T_H, X_H) + S(T_L, X_L) \leq S(T, X) + S(T'_H, X_H) + S(T'_L, X_L) \quad (10.15)$$

が成り立つ。すなわち、

$$S(T_H, X_H) - S(T'_H, X_H) \leq S(T'_L, X_L) - S(T_L, X_L) \quad (10.16)$$

が成り立つ。ここで、左辺について、

$$S(T_H, X_H) - S(T'_H, X_H) = \int_{T'_H}^{T_H} \frac{\partial S}{\partial T} dT \quad (10.17)$$

$$= \int_{T'_H}^{T_H} \frac{1}{T} \frac{\partial U}{\partial T} dT \quad (10.18)$$

$$\geq \frac{1}{T_H} \int_{T'_H}^{T_H} \frac{\partial U}{\partial T} dT \quad (10.19)$$

$$= \frac{1}{T_H} (U(T_H) - U(T'_H)) \quad (10.20)$$

$$= \frac{Q_H}{T_H} \quad (10.21)$$

となる。ただし、 Q_H は高温熱浴からの熱量である。同様の手順により、

$$S(T_L, X_L) - S(T'_L, X_L) \leq \frac{Q_L}{T_L} \quad (10.22)$$

となる。したがって、

$$\frac{Q_H}{T_H} \leq \frac{Q_L}{T_L} \quad (10.23)$$

が成り立つ。すなわち、

$$\frac{T_L}{T_H} \leq \frac{Q_L}{Q_H} \quad (10.24)$$

となる。熱効率の定義は、

$$\eta = 1 - \frac{Q_L}{Q_H} \quad (10.25)$$

であるから、

$$\eta \leq 1 - \frac{T_L}{T_H} \quad (10.26)$$

が成り立つ。 \square

11 統計力学

主に田崎統計を参考にしながら、統計物理の一般的な議論をまとめる。
典型性からスタートして、その表現としての確率分布を導入し、その性質や分布間の関係を考える。

熱力学において、我々は、エントロピーなど完全な熱力学関数さえわかれば、諸々の熱力学量を導くことができることを学んだ。統計力学においては、その完全な熱力学関数の基本関係式を、微視的に導くことが目標である。

11.1 状態数

11.1.1 状態数

Def: 状態数

エネルギーが E_i 以下であるような固有状態の数を状態数といい、 $\Omega(E_i)$ で表す。

状態数を一般に計算することはできないが、例えば相互作用のない N 粒子理想気体など、簡単な系については計算が可能である。

ex:N 粒子理想気体

後に、温度計としての役割を理想気体に担ってもらうため、実際に理想気体の状態数を計算してみる。

3次元の箱の中での理想気体のシュレディンガー方程式は、

$$-\frac{\hbar^2}{2m}\nabla^2\psi = E\psi \quad (11.1)$$

である。このとき、シュレディンガー方程式を解くと、

$$E = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2) \quad (11.2)$$

となることがわかる。ただし、 n_x, n_y, n_z は自然数である。

いま、 $E_0 = \frac{\hbar^2\pi^2}{2mL^2}$ とおくと、

$$E_{n_\alpha^{(j)}} = E_0 \sum_{\alpha=x,y,z} \sum_{j=1}^N (n_\alpha^{(j)})^2 \quad (11.3)$$

と表すことができる。このとき、状態数は、

$$\Omega(E) = \frac{1}{N!} (E_{(n_\alpha^{(j)})} \leq E \text{ を満たす自然数の組 } (n_\alpha^{(j)}) \text{ の数}) \quad (11.4)$$

$$= \frac{1}{N!} \left(\sum_{\alpha=x,y,z} \sum_{j=1}^N (n_\alpha^{(j)})^2 \leq \frac{E}{E_0} \text{ を満たす自然数の組 } (n_\alpha^{(j)}) \text{ の数} \right) \quad (11.5)$$

と書くことができる。ただし、ここでは粒子を区別しない方法により状態数を定義している。

上の状態数を評価することを考える。このとき、状態数で数える自然数の条件を書き換えると、

$$\sqrt{\sum_{\alpha=x,y,z} \sum_{j=1}^N (n_\alpha^{(j)})^2} \leq \sqrt{\frac{E}{E_0}} \quad (11.6)$$

と書き換えることができる。これは、 $3N$ 次元空間での球の内部の格子点を数えることに相当する。

このとき、球の体積は、

$$V_{3N} = \frac{\pi^{3N/2}}{(3N/2)!} \left(\frac{E}{E_0} \right)^{3N/2} \quad (11.7)$$

なるので、状態数は、

$$\Omega(E) \sim \frac{1}{N!} \frac{1}{2^{3N}} \frac{\pi^{3N/2}}{(3N/2)!} \left(\frac{E}{E_0} \right)^{3N/2} \quad (11.8)$$

$$= \frac{1}{N!} \frac{1}{(3N/2)!} \tilde{\alpha}^N E^{3N/2} V^N \quad (11.9)$$

となる。ただし、第一式での 2^{3N} は、 $3N$ 次元でのある一つの象限に注目しているために出てきた因子である。また、 $\tilde{\alpha} = \left\{ \frac{m}{2\pi} \right\}^{3/2} \hbar^{-3}$ である。

ここで、スターリングの公式を用いて、

$$\Omega(E) \sim \left\{ \frac{e}{N} \right\}^N \left\{ \frac{e}{3N/2} \right\}^{3N/2} \tilde{\alpha}^N E^{3N/2} V^N \quad (11.10)$$

$$= \alpha^N \left\{ \frac{E}{V} \right\}^{3N/2} \left\{ \frac{V}{N} \right\}^{5N/2} \quad (11.11)$$

と書くことができる。ただし、 $\alpha = \{e^{5/2}(2/3)^{3/2}\tilde{\alpha}\}$ である。

対数をとると、

$$\log \Omega(E) \simeq N \log \left(\alpha \left\{ \frac{E}{V} \right\}^{3/2} \left\{ \frac{V}{N} \right\}^{5/2} \right) \quad (11.12)$$

$$= V(\rho \log(\alpha \epsilon^{3/2} \rho^{-5/2})) \quad (11.13)$$

のように体積をくくりだすことができる。ただし、

$$\rho = \frac{N}{V}, \epsilon = \frac{E}{N} \quad (11.14)$$

である。したがって、

$$\Omega(E) \sim \exp(V\sigma(\rho, \epsilon)) \quad (11.15)$$

と書くことができる。

ただし、

$$\sigma(\rho, \epsilon) = \rho \log(\alpha \epsilon^{3/2} \rho^{-5/2}) \quad (11.16)$$

である。

11.1.2 状態数の振る舞い

一般に、相互作用がある場合の状態数の振る舞いについて考える。ここでは厳密に示すことはしないが、“普通の”熱力学的な系について、以下が成り立つことが知られている。

Thm: 状態数の熱力学極限

$v(|\mathbf{r}|)$ を相互作用のポテンシャルとする。このとき、

1. ある定数 r_0 が存在し、任意の $|\mathbf{r}| > r_0$ に対して、

$$v(|\mathbf{r}|) \leq 0 \quad (11.17)$$

が成り立つ。

2. ある定数 $b > 0$ が存在して、任意の粒子数 N と任意の粒子の位置 $\mathbf{r}_1, \dots, \mathbf{r}_N$ に対して、

$$\sum_{i,j}^N v(|\mathbf{r}_i - \mathbf{r}_j|) \geq -bN \quad (11.18)$$

が成り立つ。

このとき、 $\frac{N}{V}$ を一定に保ちながら体積を大きくしたとき、基底エネルギー密度の極限

$$\epsilon_0(\rho) = \lim_{V \rightarrow \infty} \frac{E_0(V, N)}{V} \quad (11.19)$$

が存在する。

また、 $\epsilon_0(\rho) < \infty$ であるとき、任意の $\rho > 0, \epsilon > \epsilon_0(\rho)$ に対して、 ρ と ϵ を一定に保って体積を大きくしたときの極限

$$\sigma(\rho, \epsilon) = \lim_{V \rightarrow \infty} \frac{\log \Omega_{V,N}(E)}{V} \quad (11.20)$$

が存在する。このとき、関数 $\sigma(\rho, \epsilon)$ は、 ϵ の増加関数であり、 ρ, ϵ について上に凸である。

上の二つの条件を見ていく。一つ目の条件は、遠方で、引力相互作用が働くことを表している。無限遠で相互作用ポテンシャルが 0 に漸近することを考えると、漸近の仕方は 2 通り (正から近づくか負から近づくか) あるが、ここでは、負から近づくことを考えている。

そもそも遠方で斥力が働く系では、系全体が無限に広がってしまうのだから明らかに普通ではない。

二つ目の条件は、逆に、系のポテンシャルエネルギーの下限を定めている。あまりにポテンシャルエネルギーの極小値が小さいと、無数の粒子が強い引力で凝縮するというこれまた病的な状態になってしまう。

このようなとき、状態数は、

$$\Omega_{V,N}(E) \sim \exp(V\sigma(\rho, \epsilon)) \quad (11.21)$$

と表すことができる。同様の議論は、粒子数を大きくするときにもできて、

$$\Omega_{V,N}(E) \sim \exp(N\sigma(\rho, \epsilon)) \quad (11.22)$$

と表すことができる。

11.2 典型性

Def: "許された状態"

"許された状態"とは、マクロに見たときの系のエネルギーにほぼ等しいエネルギーをもつ状態およびその重ね合わせの状態のことである。^a

^a エネルギーを用いて定義しているわけだが、ほかの変数を用いて定義することはできないのかなと思う。

このことを用いて、以下の要請を考える。

このとき、マクロな量子系について、ある平衡状態に対応する「許される量子状態」のほとんどすべてが、マクロ物理量の測定により区別不可能である。このような状態を「典型的な状態」と呼ぶ。

Def: 典型的 (統計力学)

ある性質が、"許された状態"のうち、ほとんどの状態に対して成り立つとき、その性質は典型的であるという。

このもとで、平衡状態と典型性が以下の要請によって結びつく。

要請: 典型性と平衡状態

"許された状態"のうちから、ランダムに状態を選んでくると、ほとんど確実に平衡状態が選ばれる。すなわち、典型的な状態は、平衡状態である。

以上の要請を踏まえると、ある系の持つ平衡状態の持つ性質は、"許された状態"のうち、典型的な状態が共通に持っている性質のことであると言える。

典型性の要請から、系の状態を何らかの方法で平均すると、期待値として平衡状態における値が返ってくるはずである。我々が統計力学を組み立てていくうえですべき残りのことは、状態を適当な重みづけにより混合することで確率分布を構成し、その確率分布が平衡状態を記述するうえでふさわしいことを確かめることである。後者は特に、示量変数の密度のゆらぎが十分に小さいことを確かめればよい。以下、特に熱力学の知見との対応をとりやすいような混合の仕方を見ていく。

11.3 ミクロカノニカル分布

一番シンプルな重みづけの方法として、すべての状態を等確率に混ぜ合わせる手法が考えられる。

Def: ミクロカノニカル分布

エネルギー固有値が、

$$U - V\delta < E_i < U \quad (11.23)$$

であるような状態の数を、

$$W(U, \delta) = \Omega(U) - \Omega(U - \delta) \quad (11.24)$$

とかく。このとき、

$$p_i = \frac{1}{W(U, \delta)} \quad (U - \delta < E_i < U) \quad (11.25)$$

$$p_i = 0 \quad (\text{otherwise}) \quad (11.26)$$

で定義される分布をミクロカノニカル分布という。

注意されたいこととして、我々は、あくまでも物理量の期待値に関心があり、ある状態 i が選ばれる確率それ自体に物理的な意味はない。例えば、MC を少し加工したものとして、奇数番目の状態を選ぶ確率を 0 とし、偶数番目を選ぶ確率を $2p_i$ としても、物理量の期待値に影響は出ないであろう。したがって、あくまでも確率それ自体には物理的な意味はなく、単に混ぜ合わせているだけである。

状態数の振る舞いを考えると、実は、ミクロカノニカル分布を拡張することができる。いま、系が”許された状態”に入っていない確率を計算すると、

$$p = \frac{\Omega(U - \delta)}{\Omega(U)} \quad (11.27)$$

$$= \exp(V\sigma(U - \delta V, \epsilon) - V\sigma(U, \epsilon)) \quad (11.28)$$

$$= \exp\left(-V \frac{\partial \sigma}{\partial u} \delta\right) \ll 0 \quad (11.29)$$

となる。このことから、ミクロカノニカル分布を以下のように拡張する。

Def: 拡張されたミクロカノニカル分布

エネルギー固有値が U 以下である状態の数を $\Omega(U)$ とする。このとき、

$$p_i = \frac{1}{\Omega(U)} \quad (E_i \leq U) \quad (11.30)$$

$$p_i = 0 \quad (\text{otherwise}) \quad (11.31)$$

で定義される分布を拡張されたミクロカノニカル分布という。

これは、熱力学極限においては、状態数は、相空間の”表皮”の部分に集中することを意味している。

この拡張されたミクロカノニカル分布を用いて、以下のようにエントロピーを定義する。

Def:MC におけるエントロピー

MC におけるエントロピー S は、

$$S(U, V, N) = k_B \log \Omega_{V,N}(U) \quad (11.32)$$

により定まる。

注意されたいこととして、この定義から、MC は、熱力学でいうところの (U, V, N) 表示に対応することがわかる。

ところで、上では定義としたエントロピーだが、適切な仮定により導くことができる。

Thm: エントロピーの導出

エントロピー S が以下の条件を満たすとする:

1. S が、熱力学で得られたエントロピーと一致する。
2. 平衡状態が統計力学により記述される。
3. ミクロな力学によって現れる (U, V, N) は、それぞれマクロに見たエネルギー、体積、粒子数である。
4. 熱力学的な系への力学的操作は、ミクロな力学では、ハミルトニアンのパラメータ変化として表現できる。

このとき、エントロピーは (11.32) 式で表される。

Prf

熱力学のほうから準備をする。平衡状態 (U, V, N) から体積を $V + \Delta V$ まですばやく変化させ、平衡状態 $(U + \Delta U, V + \Delta V, N)$ に緩和させる。すなわち、

$$(U, V, N) \xrightarrow{a} (U + \Delta U, V + \Delta V, N) \quad (11.33)$$

という操作を考える。この操作が準静的に近いとすると、

$$S(U, V, N) = S(U + \Delta U, V + \Delta V, N) + O(\Delta V^2) \quad (11.34)$$

となる。ここで、 S は、熱力学的なエントロピーである。

同じ状況を統計力学的に考察してみる。今、4 つめの仮定から、操作者が外からする仕事はハミルトニアンのパラメータ変化によって記述され、

$$\Delta W = \left\langle \hat{H}(V + \Delta V) - \hat{H}(V) \right\rangle_{U,V}^{\text{MC}} \quad (11.35)$$

となる。ここで、MC を拡張したときの議論を思い出すと、あくまでも、エネルギー固有値が上辺の U に近いエネルギー固有状態のみが期待値に寄与するので、

$$U = \left\langle \hat{H} \right\rangle_{U,V}^{\text{MC}} \quad (11.36)$$

と書くことができる。今、断熱操作でのエネルギー変化が仕事 ΔW であることから、

$$U + \Delta U = \left\langle \hat{H}(V) \right\rangle_{U,V}^{\text{MC}} + \Delta W = \left\langle \hat{H}(V + \Delta V) \right\rangle_{U,V}^{\text{MC}} \quad (11.37)$$

となる。最右辺について、 ΔV について展開すると、

$$\left\langle \hat{H}(V + \Delta V) \right\rangle_{U,V}^{\text{MC}} = \left\langle \hat{H}(V) + \frac{\partial \hat{H}}{\partial V} \Delta V \right\rangle_{U,V}^{\text{MC}} + O((\Delta V)^2) \quad (11.38)$$

$$= \frac{1}{\Omega(\tilde{U}, V)_{V,N}(U)} \sum_i^{\Omega(\tilde{U}, V)_{V,N}(U)} (E_i(V) + \frac{\partial E_i}{\partial V} \Delta V) + O((\Delta V)^2) \quad (11.39)$$

$$= \frac{1}{\Omega(\tilde{U}, V)_{V,N}(U)} \sum_i^{\Omega(\tilde{U}, V)_{V,N}(U)} E_i(V + \Delta V) + O((\Delta V)^2) \quad (11.40)$$

となる。一方、

$$U + \Delta U = \left\langle \hat{H}(V + \Delta V) \right\rangle_{U, V + \delta V}^{\text{MC}} \quad (11.41)$$

$$= \frac{1}{\Omega(U, \tilde{V} + \delta V)_{V,N}(U)} \sum_i^{\Omega(U, \tilde{V} + \delta V)_{V,N}(U)} E_i(V + \Delta V) \quad (11.42)$$

である。右辺同士を比較して、

$$\Omega_{V,N}(U) = \Omega(\tilde{U}, V)_{V + \delta V, N}(U) + O((\Delta V)^2) \quad (11.43)$$

となる。したがって、エントロピーは、状態数の関数で表されと考えられる。

また、エントロピーが相加性を持つことから、関数形としては、状態数の対数をとればよいことがわかる*4。

*4 これは示すことができる。やりたいことは、

$$f(pq) = f(p) + f(q) \quad (11.44)$$

を満たすような関数 f が対数関数であることを示すことである。両辺 p で微分して、

$$qf'(pq) = f'(p) \quad (11.45)$$

となる。 $p = 1$ とすると、

$$f'(q) = \frac{1}{q} f'(1) \quad (11.46)$$

となる。これを両辺積分して、

$$f(q) = f'(1) \log q + C \quad (11.47)$$

となる。ここで、 $f(1) = 0$ なので、 $C = 0$ である。よって、

$$f(q) = a \log q \quad (11.48)$$

となる。

したがって、 $\log \Omega_{V,N}(U)$ は、エントロピーの持つべき資格を満たしていることになる。ところで、熱力学の知見より、エントロピーはこれの定数倍と定数の足し算になることがわかる。すなわち、

$$S(U, V, N) = a \log \Omega_{V,N}(U) + b \quad (11.49)$$

と表されることがわかる。一般性を失うことなく、(または、熱力学第三法則に適合するように)、 $b = 0$ としてよい。

ここで、 a の値を決めることを考える。ここで、上で準備した理想気体を用いる。

理想気体については、

$$\log \Omega_{V,N}(U) \simeq N \log \left(\alpha \left(\frac{N}{V} \right)^{-\frac{5}{2}} \left(\frac{U}{N} \right)^{\frac{3}{2}} \right) \quad (11.50)$$

であることがわかっている。また、

$$\frac{\partial S}{\partial U} = a \frac{3}{2} \frac{N}{U} \quad (11.51)$$

となる。ところで、熱力学の知見より、

$$\frac{\partial S}{\partial U} = \frac{1}{T} \quad (11.52)$$

である。したがって、

$$\frac{3}{2} a N \left(\frac{3}{2} N k T \right)^{-1} = \frac{1}{T} \quad (11.53)$$

となる。これを解いて、

$$a = k \quad (11.54)$$

となる。以上より、エントロピーは、

$$S(U, V, N) = k \log \Omega_{V,N}(U) \quad (11.55)$$

となることがわかる。 \square

また、エントロピーが示量性を持つことも示すことができる。

Thm: エントロピーの示量性

(11.32) 式で定義されるエントロピーは、熱力学極限において示量性を持つ。

Prf

状態数の熱力学極限を考えると、普通の系については、

$$s(u, \rho) = \lim_{V \rightarrow \infty} \frac{k}{V} \log \Omega_{V,N}(U) \quad (11.56)$$

が極限を持つ。この左辺はエントロピー密度を表しており、エントロピー密度の熱力学極限が定義できることから、たしかに熱力学極限において、エントロピーは示量性を持つことがわかる。 \square

また、このとき定義されたエントロピー密度が、その二つの変数に対して上に凸であることは、状態数の熱力学極限の話から直ちに得られる。

11.4 カノニカル分布

別の混合の仕方として、エネルギーに対して重みづけをする方法がある。

Def: カノニカル分布

カノニカル分布は、エネルギー E_i に対して、

$$p_i = \frac{e^{-\beta E_i}}{Z} \quad (11.57)$$

で定義される分布である。ただし、 Z は規格化定数であり、

$$Z = \sum_i^n e^{-\beta E_i} \quad (11.58)$$

である。この規格化定数を分配関数という。

カノニカル分布は、究極的には天から降ってくる重みづけの仕方としてもよいのだが^{*5}、ここでは、ミクロカノニカル分布を前提にして、カノニカル分布を導出する。

系のエネルギーを E_i 、熱浴のエネルギーを B_j とする。このとき、全系の内部エネルギーが、マクロにみて U であるときの”許された状態”は、

$$U - \delta V_R < E_i + B_j < U \quad (11.59)$$

である。これを变形して、

$$U - E_i - \delta V_R < B_j < U - E_i \quad (11.60)$$

となる。これを用いて、系がエネルギー状態 E_i にある状態数は、

$$\Omega_i = \Omega(U - E_i) - \Omega(U - E_i - \delta V_R) \quad (11.61)$$

により計算される。このとき、系がエネルギー状態 E_i にある確率は、

$$p_i = \frac{\Omega_i}{\Omega_{\text{total}}} \quad (11.62)$$

である。ここで、熱浴の状態数を、

$$\Omega_R(B) = \exp\left(V_R \sigma\left(\frac{B}{V_R}, \frac{N_R}{V_R}\right)\right) \quad (11.63)$$

^{*5} 典型性の話を思い出すとよい。

と書くことができる。 $\tilde{U} = U - E_i$ とおくと、

$$p_i = \frac{\Omega(\tilde{U}) - V_R \delta}{\Omega(\tilde{U})} \quad (11.64)$$

$$= \exp(V_R \{\sigma(\tilde{u} - \delta, \rho) - \sigma(\tilde{u}, \rho)\} + o(\delta)) \quad (11.65)$$

$$= \exp\left(-V_R \left(\frac{\partial}{\partial \tilde{u}} \sigma(\tilde{u}, \rho) \delta + o(\delta^2)\right) + o(V_R \delta)\right) \ll 1 \quad (11.66)$$

を満たす。ただし、 $\tilde{u} = \frac{U}{V_R}$ であり、熱浴の体積が十分大きいことを用いている。
このとき、

$$p_i \simeq \frac{\Omega_R(U - E_i)}{\sum_j^n \Omega_R(U - E_j)} \quad (11.67)$$

$$= \frac{\Omega_R(U - E_i)}{\Omega_R(U)} \left(\sum_j^n \frac{\Omega_R(U - E_j)}{\Omega_R(U)} \right)^{-1} \quad (11.68)$$

と書ける。ここで、

$$\log \frac{\Omega_R(U - E_i)}{\Omega_R(U)} = \log \Omega_R(U - E_i) - \log \Omega_R(U) \quad (11.69)$$

$$= -E_i \frac{\partial}{\partial U} \log \Omega_R(U) + \frac{E_i^2}{2} \frac{\partial^2}{\partial U^2} \log \Omega_R(U) + \dots \quad (11.70)$$

$$= -\frac{E_i}{V_R} \frac{\partial}{\partial u} \{V_R \sigma(u, \rho) + o(V_R)\} + \frac{E_i^2}{2V_R^2} \frac{\partial^2}{\partial u^2} \{V_R \sigma(u, \rho) + o(V_R)\} + \dots \quad (11.71)$$

$$= -E_i \frac{\partial}{\partial u} \sigma(u, \rho) + \frac{1}{V_R} \frac{E_i^2}{2} \frac{\partial^2}{\partial u^2} \sigma(u, \rho) + \dots + \frac{o(V_R)}{V_R} \quad (11.72)$$

$$\simeq -\beta(u, \rho) E_i \quad \because V_R \text{ が十分大きい} \quad (11.73)$$

となる。ここで、 $\beta(u, \rho) = \frac{\partial}{\partial u} \sigma(u, \rho) = \frac{\partial}{\partial U} \log \Omega_R(U)$ である。これにより、

$$\frac{\Omega_R(U - E_i)}{\Omega_R(U)} = \exp(-\beta(u, \rho) E_i) \quad (11.74)$$

と書くことができ、

$$p_i = \frac{\exp(-\beta(u, \rho) E_i)}{Z(\beta)} \quad (11.75)$$

と書くことができる。ただし、 Z は規格化定数であり、

$$Z(\beta) = \sum_i \exp(-\beta(u, \rho) E_i) \quad (11.76)$$

である。

また、 N この粒子を区別しないときには、エネルギー固有状態についての和を、 $N!$ で割ればよい。

すなわち、

$$Z'_{V,N}(\beta) = \frac{1}{N!} Z_{V,N}(\beta) \quad (11.77)$$

により新たな分配関数を定義すればよい。このとき、期待値の表式は、

$$\langle f \rangle_{\beta}^{\text{can}} = \frac{1}{Z'_{V,N}(\beta)} \frac{1}{N!} \sum_i f_i \exp(-\beta E_i) \quad (11.78)$$

となる。以下、このときの分配関数も Z と書くことにする。

パラメータ β の物理的な意味を考える。理想気体における状態数を温度計として使うと、定義から、

$$\beta(u, \rho) = \frac{\partial}{\partial u} \sigma(u, \rho) \quad (11.79)$$

$$= \frac{\partial}{\partial u} \left(\rho \log \left(\alpha u^{3/2} \rho^{-5/2} \right) \right) \quad (11.80)$$

$$= \frac{3\rho}{2u} \quad (11.81)$$

$$= \frac{3}{2} \frac{N_R}{U_{\text{total}}} \quad (11.82)$$

となる。ところで、熱力学の知見より、

$$U_{\text{total}} = \frac{3}{2} N_R kT \quad (11.83)$$

である。したがって、

$$\beta = \frac{1}{kT} \quad (11.84)$$

となる。この β は逆温度と呼ばれる。

以下、この分布を用いて、熱力学で登場する物理量を導出する。^{*6}

ex:N 粒子理想気体

後に、温度計としての役割を理想気体に担ってもらうため、理想気体の分配関数を計算してみる。

今、ハミルトニアンは、

$$H = \sum_j^N \frac{\mathbf{p}_j^2}{2m} \quad (11.85)$$

^{*6} ただし、これは熱力学極限をとらずとも定義できる。ただし、その整合性は熱力学極限をとったときにゆらぎが小さいことから確かめる。

である。このとき、分配関数は、

$$Z = \frac{1}{N!h^{3N}} \int d\mathbf{p}_1 \cdots \int d\mathbf{p}_N d\mathbf{q}_1 \cdots d\mathbf{q}_N \exp\left(-\frac{\beta}{2m} \sum_j^N \mathbf{p}_j^2\right) \quad (11.86)$$

である。ここで、

$$\int d\mathbf{q}_1 \cdots \int d\mathbf{q}_N = V^N \quad (11.87)$$

であり、また、ガウス積分を用いて、

$$\int d\mathbf{p}_1 \cdots \int d\mathbf{p}_N \exp\left(-\frac{\beta}{2m} \sum_j^N \mathbf{p}_j^2\right) = \left(\frac{2\pi m}{\beta}\right)^{3N/2} \quad (11.88)$$

であるから、

$$Z = \frac{V^N}{N!h^{3N}} \left(\frac{2\pi m}{\beta}\right)^{3N/2} \quad (11.89)$$

となる。

Prop: エネルギー期待値と分散/比熱

カノニカル分布において、エネルギーの期待値と分散は、

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln Z \quad (11.90)$$

$$\sigma^2 = \frac{\partial^2}{\partial \beta^2} \ln Z \quad (11.91)$$

で与えられる。また、比熱は、

$$C = \frac{1}{kT^2} \sigma^2 \quad (11.92)$$

で与えられる。

Prf

$$-\frac{\partial}{\partial \beta} \log Z = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \quad (11.93)$$

$$= -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_i e^{-\beta E_i} \quad (11.94)$$

$$= -\frac{1}{Z} \sum_i (-E_i) e^{-\beta E_i} \quad (11.95)$$

$$= \sum_i E_i p_i \quad (11.96)$$

$$= \langle E \rangle \quad (11.97)$$

$$\langle E^2 \rangle = \sum_i E_i^2 p_i \quad (11.98)$$

$$= \sum_i E_i^2 \frac{e^{-\beta E_i}}{Z} \quad (11.99)$$

$$= \frac{1}{Z} \frac{\partial^2}{\partial \beta^2} \sum_i e^{-\beta E_i} \quad (11.100)$$

$$= \frac{Z''}{Z} \quad (11.101)$$

となる。したがって、

$$\sigma^2 = \langle E^2 \rangle - \langle E \rangle^2 \quad (11.102)$$

$$= \frac{Z''}{Z} - \left(\frac{Z'}{Z} \right)^2 \quad (11.103)$$

$$= \frac{Z''Z - Z'^2}{Z^2} \quad (11.104)$$

$$= \frac{\partial^2}{\partial \beta^2} \log Z \quad (11.105)$$

となる。さらに、

$$C(T) = \frac{\partial \langle E \rangle}{\partial T} \quad (11.106)$$

$$= \frac{\partial \beta}{\partial T} \frac{\partial \langle E \rangle}{\partial \beta} \quad (11.107)$$

$$= \frac{1}{kT^2} \frac{\partial^2}{\partial \beta^2} \log Z \quad (11.108)$$

$$= \frac{1}{kT^2} (\langle E^2 \rangle - \langle E \rangle^2) \quad (11.109)$$

である。以上より示された。 □

Def: ヘルムホルツ自由エネルギー

カノニカル分布において、ヘルムホルツ自由エネルギーは、

$$F = -\frac{1}{\beta} \ln Z_{V,N}(\beta) \quad (11.110)$$

により定まる。

Thm: ヘルムホルツ自由エネルギーの導出

ヘルムホルツ自由エネルギーが以下の条件を満たすとする:

1. ヘルムホルツ自由エネルギーが、熱力学で得られたヘルムホルツ自由エネルギーと一致する。
2. 平衡状態が統計力学により記述される。
3. ミクロな力学によって現れる (U, V, N) は、それぞれマクロに見たエネルギー、体積、粒子数である。
4. 熱力学的な系への力学的操作は、ミクロな力学では、ハミルトニアンのパラメータ変化として表現できる。

このとき、ヘルムホルツ自由エネルギーは、(11.110) 式で表される。

Prf

熱力学のほうから準備する。平衡状態 (T, V, N) から体積を $V + \Delta V$ まですばやく変化させ、平衡状態 $(T, V + \Delta V, N)$ に緩和させる。すなわち、

$$(T, V, N) \xrightarrow{a} (T, V + \Delta V, N) \quad (11.111)$$

という操作を考える。この操作が準静的に近いとすると、

$$F(T, V, N) - F(T, V + \Delta V, N) = W_{\max} + O(\Delta V^2) \quad (11.112)$$

となる。ここで、 F は、熱力学的なヘルムホルツ自由エネルギーである。

同じ状況を統計力学的に考察してみる。今、4 つめの仮定から、操作者が外からする仕事はハミルトニアンのパラメータ変化によって記述され、

$$\Delta W = \left\langle \hat{H}(V + \Delta V) - \hat{H}(V) \right\rangle_{T,V}^{\text{can}} \quad (11.113)$$

と書くことができる。

ここで、準静的等温陸下について、内部エネルギーの変化が無視できると仮定すると^{*7}

$$\langle H(V + \Delta V) \rangle_{T,V+\Delta V}^{\text{can}} = \langle H(V) \rangle_{T,V}^{\text{can}} \quad (11.114)$$

となる。したがって、

$$\Delta W = \langle H(V + \Delta V) \rangle_{T,V}^{\text{can}} - \langle H(V + \Delta V) \rangle_{T,V+\Delta V}^{\text{can}} \quad (11.115)$$

となる。期待値を展開して、

$$\Delta W = \langle H(V) \rangle_{T,V}^{\text{can}} - \langle H(V) \rangle_{T,V+\Delta V}^{\text{can}} \quad (11.116)$$

$$= \left(\frac{1}{Z_{V,N}(\beta)} - \frac{1}{Z_{V+\Delta V,N}(\beta)} \right) \sum_i H_i e^{-\beta E_i} \quad (11.117)$$

^{*7} そんなわけない。この証明はダメそう。備忘録として書いておく。

と書くことができる。よって、自由エネルギーは、分配関数の関数となることがわかる。
また、エントロピーのときと同様に、ヘルムホルツ自由エネルギーは、

$$F(T, V, N) = a \log Z_{V,N}(\beta) + b \quad (11.118)$$

と表されることがわかる。 $b = 0$ としてよい。
理想気体と比較することで、

$$a = -\frac{1}{\beta} \quad (11.119)$$

となることがわかる。以上より示された。 \square

Prop: エントロピー

カノニカル分布において、エントロピーは、

$$S(\beta, V, N) = -k \sum_i p_i \log p_i - k \log N! \quad (11.120)$$

で与えられる。

Lem:

カノニカル分布において、エントロピーは、

$$S(T, V, N) = \frac{\langle \hat{H} \rangle_{T,V}^{\text{can}} - F(T, V, N)}{T} \quad (11.121)$$

と書くことができる。

Prf

熱力学の知見より、

$$S = -\frac{\partial F}{\partial T} \quad (11.122)$$

$$= -\frac{\partial F}{\partial \beta} \frac{\partial \beta}{\partial T} \quad (11.123)$$

$$= \frac{1}{kT^2} \left(\frac{1}{\beta^2} \log Z - \frac{1}{\beta} \frac{Z'}{Z} \right) \quad (11.124)$$

$$= k \log Z - k \frac{Z'}{Z} \quad (11.125)$$

$$= \frac{\langle \hat{H} \rangle_{T,V}^{\text{can}} - F(T, V, N)}{T} \quad (11.126)$$

となる。以上より示された。 \square

Prf(Prop)

$$-k \sum_i p_i \log p_i = -k \sum_i \frac{e^{-\beta E_i}}{Z} \log(\beta E_i - \log Z) \quad (11.127)$$

$$= k\beta \sum_i \left(\frac{E_i e^{-\beta E_i}}{Z} \right) + k \sum_i \left(\frac{e^{-\beta E_i}}{Z} \log Z \right) \quad (11.128)$$

$$= \frac{1}{T} \langle \hat{H} \rangle_{T,V}^{\text{can}} - \beta k \sum_i \left(\frac{e^{-\beta E_i}}{Z} F \right) \quad (11.129)$$

$$= \frac{\langle \hat{H} \rangle_{T,V}^{\text{can}} - F(T, V, N)}{T} \quad (11.130)$$

となる。これと補題より示された。 \square

次に、熱力学極限を考える。

Thm: ヘルムホルツ自由エネルギー密度

統計力学的に定めたヘルムホルツ自由エネルギーについて、

$$f(\beta, \rho) = \lim_{V \rightarrow \infty} \frac{1}{V} F(\beta, V, N) = -\frac{1}{V\beta} \log Z \quad (11.131)$$

の極限が存在する。この極限をヘルムホルツ自由エネルギー密度という。

Prf

状態数の熱力学極限と同様に従う。 \square

上の熱力学極限は、ヘルムホルツ自由エネルギーが、

$$F(\beta, V, N) = V f(\beta, \rho) + o(V) \quad (11.132)$$

と表されることを意味している。これを用いて、エネルギー密度の期待値やゆらぎを考える。

Prop: エネルギー密度の期待値とゆらぎ

カノニカル分布において、エネルギー密度の期待値とゆらぎは、

$$\langle \epsilon \rangle \simeq \frac{\partial}{\partial \beta} (\beta f(\beta, \rho)) \quad (11.133)$$

$$\sigma_{\beta}^{\text{can}}(\epsilon) \simeq \frac{1}{\sqrt{V}} \sqrt{-\frac{\partial^2}{\partial \beta^2} (\beta f(\beta, \rho))} \quad (11.134)$$

で与えられる。

Prf

$$\langle \epsilon \rangle = -\frac{1}{V} \frac{\partial}{\partial \beta} \log Z \quad (11.135)$$

$$= \frac{\partial}{\partial \beta} (\beta f(\beta, \rho)) + \frac{o(V)}{V} \quad (11.136)$$

となる。また、

$$\sigma_{\beta}^{\text{can}}(\epsilon) = \frac{1}{V} \sqrt{\frac{\partial^2}{\partial \beta^2} \log Z} \quad (11.137)$$

$$= \frac{1}{\sqrt{V}} \sqrt{-\frac{\partial^2}{\partial \beta^2} (\beta f(\beta, \rho)) + \frac{o(V)}{V}} \quad (11.138)$$

となる。以上より示された。 \square

したがって、たしかに、カノニカル分布による状態の混合は、熱力学極限においてゆらぎが十分小さく、混ぜ合わせ方として妥当であることがわかる。

11.5 グランドカノニカル分布

他の混合の仕方として、粒子数に対しても重みづけをする方法がある。

Def: グランドカノニカル分布

グランドカノニカル分布は、エネルギー E_i 、粒子数 N_i に対して、

$$p_i = \frac{e^{-\beta(E_i - \mu N_i)}}{\Xi} \quad (11.139)$$

で定義される分布である。ただし、 Z は規格化定数であり、

$$\Xi = \sum_i^n e^{-\beta(E_i - \mu N_i)} \quad (11.140)$$

である。この規格化定数を大分配関数という。

気が向いたら書くが、カノニカル分布から導出可能である。

11.6 モデルの等価性

これまで導いてきた分布が等価であることを示す。具体的な手続きとしては、まずミクロカノニカル分布が (U, V, N) 表示に対応することを示し、次に、ミクロカノニカル分布とカノニカル分布

がルジャンドル変換によって結ばれることを統計力学側の視点から示す。グランドカノニカル分布についても同様である。

11.6.1 ミクロカノニカル分布とカノニカル分布

体積 V 、粒子数 N を固定した系を考える。分配関数の和をとる時に、 $\Delta > 0$ を $\beta\Delta \ll 1$ となるような数として、粗視化する。すなわち、

$$Z_{V,N}(\beta) = \frac{1}{N!} \sum_i e^{-\beta E_i} \quad (11.141)$$

$$= \frac{1}{N!} \sum_{n=0}^{\infty} \sum_{\substack{i \\ (n\Delta < E_i < (n+1)\Delta)}} e^{-\beta E_i} \quad (11.142)$$

$$\simeq \frac{1}{N!} \sum_{n=0}^{\infty} \Omega_{V,N}(n\Delta) e^{-\beta n\Delta} \quad (11.143)$$

となる。すなわち、領域 $[n\Delta, (n+1)\Delta]$ のエネルギー固有状態を束としてまとめて和をとってから、その束に対して和をとることになる。このとき、

$$\frac{1}{N!} \sum_{\substack{i \\ (n\Delta < E_i < (n+1)\Delta)}} 1 = \Omega_{V,N}((n+1)\Delta) - \Omega_{V,N}(n\Delta) \quad (11.144)$$

$$= \Delta \frac{d\Omega_{V,N}(E)}{dE} \Big|_{E=n\Delta} \quad (11.145)$$

である。したがって、

$$Z_{V,N}(\beta) \simeq \sum_{n=0}^{\infty} \Delta \frac{d\Omega_{V,N}(E)}{dE} e^{-\beta n\Delta} \quad (11.146)$$

$$= \int_0^{\infty} dE \frac{d\Omega_{V,N}(E)}{dE} e^{-\beta E} \quad (11.147)$$

$$= \beta \int_0^{\infty} dE \Omega_{V,N}(E) e^{-\beta E} \quad (11.148)$$

となる。ここで、状態数の熱力学極限を考えると、

$$Z_{V,N}(\beta) \simeq \beta \int_0^{\infty} dE \exp(V\sigma(E/N, V/N)) e^{-\beta E} \quad (11.149)$$

$$= \beta V \int_0^{\infty} d\epsilon \exp(V(\sigma(\epsilon, \rho) - \beta\epsilon)) \quad (11.150)$$

となる。ここで、 $\epsilon = E/N$ とおいた。この積分を評価することを考える。いま、 $(\sigma(\epsilon, \rho) - \beta\epsilon)$ を、その最大値をとる点 ϵ^* の周りでテイラー展開すると、

$$\sigma(\epsilon, \rho) - \beta\epsilon = \sigma(\epsilon^*, \rho) - \beta\epsilon^* + \frac{1}{2} \frac{\partial^2 \sigma}{\partial \epsilon^2} + \cdots \quad (11.151)$$

となる。ここで、 $\alpha = -\frac{\partial^2 \sigma}{\partial \epsilon^2}$ とおき、これが正であると仮定する。^{*8}このとき、最大値周りだけが積分に寄与することから、

$$Z_{V,N}(\beta) \simeq \beta V \int_0^\infty d\epsilon \exp(V(\sigma(\epsilon^*, \rho) - \beta\epsilon^*)) \exp\left(-\frac{1}{2}\alpha V(\epsilon - \epsilon^*)^2\right) \quad (11.152)$$

$$= \beta V \exp(V(\sigma(\epsilon^*, \rho) - \beta\epsilon^*)) \sqrt{\frac{2\pi}{\alpha V}} \quad (11.153)$$

となる。これを用いて、ヘルムホルツ自由エネルギー密度を評価すると、

$$f(\beta, \rho)_V = -\frac{1}{V\beta} \log Z_{V,N}(\beta) \quad (11.154)$$

$$= -\frac{1}{\beta}(\sigma(\epsilon^*, \rho) - \beta\epsilon^*) - \frac{1}{\beta} \log \sqrt{\frac{2\pi}{\alpha V}} \quad (11.155)$$

$$\simeq -\frac{1}{\beta}(\sigma(\epsilon^*, \rho) - \beta\epsilon^*) \quad (11.156)$$

$$= \min_{\epsilon} \left(\epsilon - \frac{\sigma(\epsilon, \rho)}{\beta} \right) \quad (11.157)$$

となる。特に、体積を無限大に取ると、これまでの \simeq の部分が厳密になることが知られており、

$$f(\beta, \rho) = \min_u (u - Ts(u, \rho)) \quad (11.158)$$

となる。したがって、MC で導いたエントロピーと、C で導いたヘルムホルツ自由エネルギーがルジャンドル変換によって結ばれており、同じ熱力学極限を持つことがわかる。すなわち、2つの分布は等価であることがわかる。□

11.7 量子統計

11.7.1 多粒子系の量子力学

多粒子系の量子力学について、簡単にまとめておく。

要請: 粒子の区別

量子力学的な粒子は区別不可能である。

量子力学的な粒子が区別できないことから、

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \alpha \psi(\mathbf{r}_2, \mathbf{r}_1) \quad (11.159)$$

$$= \alpha^2 \psi(\mathbf{r}_1, \mathbf{r}_2) \quad (11.160)$$

となる。したがって、 $\alpha = \pm 1$ である。

^{*8} これは、より厳密な証明では必要ないことが知られている。

Def: フェルミ粒子/ボース粒子

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = -\psi(\mathbf{r}_2, \mathbf{r}_1) \quad (11.161)$$

を満たす粒子をフェルミ粒子、

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_2, \mathbf{r}_1) \quad (11.162)$$

を満たす粒子をボース粒子という。

また、 N 粒子系の波動関数は、ボース粒子の場合、

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \Psi(\mathbf{r}_{\sigma(1)}, \dots, \mathbf{r}_{\sigma(N)}) \quad (11.163)$$

を満たす。ここで、 σ は N 粒子の置換である。

また、フェルミ粒子の場合、

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = (-1)^\sigma \Psi(\mathbf{r}_{\sigma(1)}, \dots, \mathbf{r}_{\sigma(N)}) \quad (11.164)$$

を満たす。

11.8 その他

統計力学における諸注意や、ちょっとした議論をまとめておく。

11.8.1 独立な部分からなる系

11.8.2 分布間の関係

(11.147) 式から、

$$Z_{V,N}(\beta) = \int_0^\infty dE \frac{d\Omega_{V,N}(E)}{dE} e^{-\beta E} \quad (11.165)$$

である。これは、状態密度をラプラス変換することで、分配関数を得ているということになる。

逆に、逆ラプラス変換により、

$$\frac{d\Omega_{V,N}(E)}{dE} = \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta e^{\beta E} Z_{V,N}(\beta) \quad (11.166)$$

となる。以上より、

$$\begin{array}{ccc} D(E) & \longrightarrow & S(U, V, N) \\ \updownarrow \text{ラプラス変換} & & \updownarrow \text{ルジャンドル変換} \\ Z_{V,N}(\beta) & \longrightarrow & F(T, V, N) \end{array}$$

という関係があることがわかる。

ex: N 粒子理想気体

理想気体の分配関数は、(11.89) 式である。これをラプラス逆変換すると、

$$D(E) = \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta e^{\beta E} \frac{V^N}{h^{3N} N!} \left(\frac{2\pi m}{\beta} \right)^{3N/2} \quad (11.167)$$

となる。この複素積分は、以下の経路を考えると良い。

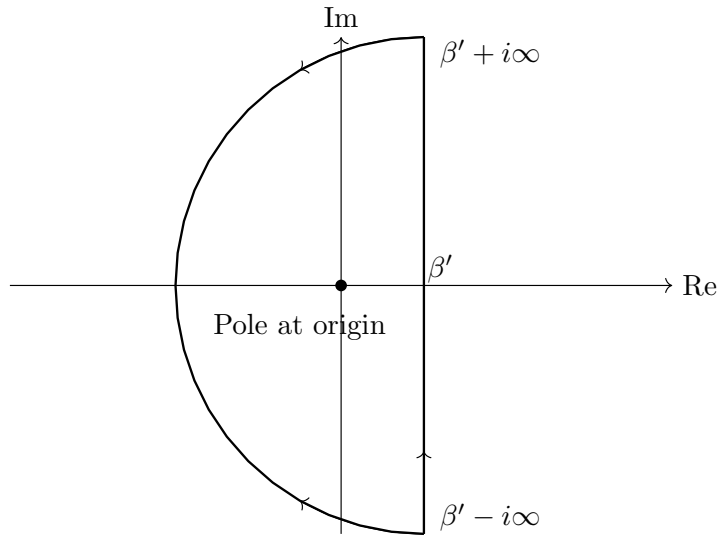


図2 逆ラプラス変換の複素積分の積分経路 C 。

このとき、原点が極となっていることがわかる。この経路に沿った線積分に等しいので、留数定理を用いて、

$$D(E) = \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta e^{\beta E} \frac{V^N}{h^{3N} N!} \left(\frac{2\pi m}{\beta} \right)^{3N/2} \quad (11.168)$$

$$= \frac{1}{2\pi i} 2\pi i \frac{V^N}{h^{3N} N!} (2\pi m)^{3N/2} \left(\frac{E^{3N/2-1}}{(3N/2-1)!} \right) \quad (11.169)$$

$$= \frac{V^N}{h^{3N} N!} (2\pi m)^{3N/2} \left(\frac{E^{3N/2-1}}{(3N/2-1)!} \right) \quad (11.170)$$

となる。また、(11.9) 式から、

$$\frac{d\Omega}{dE} = \frac{V^N}{h^{3N} N!} (2\pi m)^{3N/2} \left(\frac{E^{3N/2-1}}{(3N/2-1)!} \right) \quad (11.171)$$

である。したがって、たしかに、ラプラス変換逆変換によって状態密度を求めることができた。

11.9 母関数との対応

11.9.1 母関数との対応

各分布についてのモーメント母関数を考えてみる。

カノニカル分布を考える。カノニカル分布を再掲すると、

$$p_i = \frac{e^{-\beta E_i}}{Z(\beta)} \quad (11.172)$$

である。このとき、モーメント母関数は、

$$M(k) = \sum_i \frac{e^{-(\beta-k)E_i}}{Z(\beta)} \quad (11.173)$$

$$= \frac{Z(\beta - k)}{Z(\beta)} \quad (11.174)$$

とかくことができる。これを用いて、 n 次のモーメントは、

$$\langle E^n \rangle = \frac{\partial^n M(k)}{\partial k^n} \Big|_{k=0} \quad (11.175)$$

$$= \frac{1}{Z(\beta)} \frac{\partial^n Z(\beta - k)}{\partial k^n} \Big|_{k=0} \quad (11.176)$$

$$= (-1)^n \frac{1}{Z(\beta)} \frac{\partial^n Z(\beta - k)}{\partial k^n} \Big|_{k=0} \quad (11.177)$$

$$= (-1)^n \frac{Z^{(n)}(\beta)}{Z(\beta)} \quad (11.178)$$

となる。ここで、 $Z^{(n)}(\beta)$ は、 n 階微分である。とくに、1 次のモーメントは、

$$\langle E \rangle = -\frac{1}{Z(\beta)} \frac{\partial Z(\beta - k)}{\partial k} \Big|_{k=0} \quad (11.179)$$

$$= -\frac{\partial}{\partial \beta} \log Z(\beta) \quad (11.180)$$

となる。

この母関数との関連を見ることで、分配関数から熱力学量が出てくる理由が見えてくる。モーメント母関数を用いて計算したときは最後に $k = 0$ としたわけだが、逆に、 $\frac{e^{-\beta E_i}}{Z(\beta)}$ は、はじめから $k = 0$ としたモーメント母関数であると考えることができる。すなわち、

$$M(\beta) = \frac{e^{-\beta E_i}}{Z(\beta)} \quad (11.181)$$

のように考えることができる。ただし、このときのモーメント関数は、すでに $k = 0$ としているので、最後に値を代入する作業は必要ない。また、 β 微分に分配関数が巻き込まれないように、分配

関数を外に出してから計算する必要がある。すなわち、

$$\sum_i e^{-\beta E_i} = Z(\beta) \quad (11.182)$$

を微分して情報を取りだしてから、分配関数で割ってやるという手順が必要である。このとき、(11.182) 式は、モーメント母関数とみなすことができる。

実際に n 次のモーメントを計算すると、

$$\langle E^n \rangle = \frac{1}{Z(\beta)} \frac{\partial^n Z(\beta)}{\partial \beta^n} \quad (11.183)$$

$$= \frac{1}{Z(\beta)} (-1)^n Z^{(n)}(\beta) \quad (11.184)$$

となる。特に、1 次のモーメントは、

$$\langle E \rangle = -\frac{1}{Z(\beta)} \frac{\partial Z(\beta)}{\partial \beta} \quad (11.185)$$

$$= -\frac{\partial}{\partial \beta} \log Z(\beta) \quad (11.186)$$

となる。

また、分配関数をモーメント母関数とみなすということは、 $\log Z$ をキュムラント母関数とみなすことにもなる。実際、

$$\frac{\partial^2 \log Z}{\partial \beta^2} = \sigma^2(\beta) \quad (11.187)$$

である。(4.54) 式を参照)

まとめると、分配関数は、エネルギーのモーメント母関数とみなすことができ、確率分布の情報を取り出すことができるのである。

ところで、この話を MC にも持ち込めないかを考えてみる。C でみたこととしては、確率分布の規格化定数を母関数として取り扱うことができたということである。

11.9.2 母関数との関係 2

”許された状態”の密度は、

$$w = \sum_i \delta(E - E_i) \quad (11.188)$$

である。このとき、エントロピーが、

$$S = k \log W \quad (11.189)$$

である。これは、 W をモーメント母関数とみなしたときのキュムラント母関数である。^{*9}
また、

$$W = e^{\frac{S}{k}} \quad (11.190)$$

より、

$$p_i = \frac{1}{W} = \exp\left(-\frac{S}{k}\right) = \exp\left(-V \frac{s}{k}\right) \quad (11.191)$$

となる。よって、エネルギー分布 p_i のレート関数は $\frac{s}{k}$ である。

また、分配関数は、

$$Z_{V,N}(\beta) = \int_0^\infty dE w e^{-\beta E} \quad (11.192)$$

これと、モーメント母関数の定義 (4.49) 式を見比べてみると、分配関数は、 w を確率密度とみなしたときの (すなわち、 W を確率とみなしたときの)、エネルギーのモーメント母関数になっていることがわかる。^{*10}

モーメント母関数が求まったので、キュムラント母関数も求まる。キュムラント母関数の定義 (4.51) 式を見ると、

$$\frac{1}{k} \Phi = \log Z_{V,N}(\beta) \quad (11.193)$$

なる量がキュムラント母関数である。ここで、左辺の k はボルツマン定数、 Φ はマッシュー関数である。要するに、キュムラント母関数は、完全な熱力学関数に対応する。

キュムラント母関数は、ゆらぎと対応していたのであった。試しに、 $k\Phi$ を β で微分してみると、

$$\frac{\partial^2 \Phi}{\partial \beta^2} = k \sigma^2 \quad (11.194)$$

となる。この両辺 β^2 をかけると、

$$\beta^2 \frac{\partial^2 \Phi}{\partial \beta^2} = k \beta^2 \sigma^2 = C_V \quad (11.195)$$

となり、比熱が出てくる。^{*11}

^{*9} 何のモーメント母関数なのかは謎だが、カノニカル分布との対応を見てなんとなく書いてみた。

^{*10} ただし、状態密度が確率分布でないから、本当にモーメント母関数であるとはいえないのだが。

^{*11} この左辺は見慣れないかもしれないが、

$$C(T, V, N) = -T \frac{\partial^2 F}{\partial T^2} \quad (11.196)$$

のマッシュー関数版である。実際、

$$\frac{\partial^2}{\partial T^2} = k^2 \beta^2 \frac{\partial}{\partial \beta} \left(\beta^2 \frac{\partial}{\partial \beta} \right) \quad (11.197)$$

また、キュムラント母関数とレート関数はルジャンドル変換によって結ばれているのであった。
すなわち、

$$\phi = s - \frac{u}{T} \quad (11.201)$$

である。これは、エントロピーとマシュー関数がルジャンドル変換で結ばれていることをに対応する。

また、グランドカノニカル分布についても考えてみる。グランドカノニカル分布の大分配関数は、

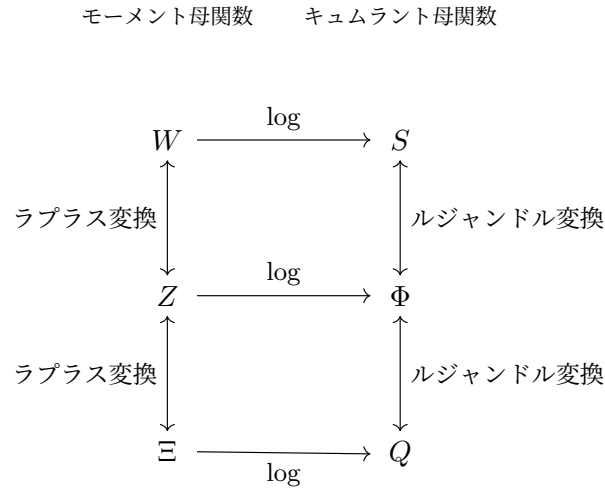
$$\Xi = \sum_{N=0}^{\infty} Z_{V,N}(\beta) \exp(\beta\mu N) \quad (11.202)$$

である。これは、 Z を確率分布と見たときの、粒子数に対するモーメント母関数である。このときのキュムラント母関数は、

$$kQ = \log \Xi \quad (11.203)$$

である。ただし、 Q はクラマース関数。したがって、グランドカノニカル分布の大分配関数は、粒子数のキュムラント母関数である。

以上より、



を用いると、

$$C(T, V, N) = -T \frac{\partial^2 F}{\partial T^2} \quad (11.198)$$

$$= \beta \frac{\partial}{\partial \beta} \left(\beta^2 \frac{\partial}{\partial \beta} \right) \left(\frac{\Phi}{\beta} \right) \quad (11.199)$$

$$= \beta^2 \frac{\partial^2 \Phi}{\partial \beta^2} \quad (11.200)$$

となる。

とも書くことができる。

11.10 分布の他の導出方法

カノニカル分布の導出方法はいろいろある。ここでは、その他の導出方法をいくつかみていく。

カノニカル分布 (状態数の利用)

熱浴の状態密度を D_{bath} 、系の状態密度を D_{sys} 、全系の状態密度を D とする。このとき、着目系のエネルギーが $[E - \Delta E, E]$ である状態数は、

$$D_{\text{sys}}(E)\Delta E D_{\text{bath}}(E_{\text{tot}} - E)\Delta E_{\text{bath}} \quad (11.204)$$

である。ここで、 E_{tot} は、全エネルギーである。このとき、すべての状態数は、

$$\int dE D_{\text{sys}}(E) D_{\text{bath}}(E_{\text{tot}} - E) \Delta E_{\text{bath}} \quad (11.205)$$

である。このとき、エネルギーが E である状態の確率は、

$$p(E) = \frac{D_{\text{sys}}(E) D_{\text{bath}}(E_{\text{tot}} - E) \Delta E_{\text{bath}}}{\int dE D_{\text{sys}}(E) D_{\text{bath}}(E_{\text{tot}} - E) \Delta E_{\text{bath}}} \quad (11.206)$$

である。ここで、

$$D_{\text{bath}}(E_{\text{bath}}) \Delta E_{\text{bath}} = \exp\left(\frac{S_{\text{bath}}(E_{\text{bath}})}{k}\right) \quad (11.207)$$

である。両辺対数をとって微分することで、

$$\frac{\partial D_{\text{bath}}}{\partial E_{\text{bath}}} = \frac{1}{kT} D_{\text{bath}} \quad (11.208)$$

となる。また、熱浴の温度が変化しないこと、すなわち、比熱が無限大になることを仮定して、

$$\frac{\partial T}{\partial E_{\text{bath}}} = 0 \quad (11.209)$$

となる。したがって、 $D_{\text{bath}}(E_{\text{tot}} - E)$ をテイラー展開すると、

$$D_{\text{bath}}(E_{\text{tot}} - E) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{1}{kT}\right)^n D_{\text{bath}}(E_{\text{tot}}) \quad (11.210)$$

$$= D_{\text{bath}}(E_{\text{tot}}) \exp(-\beta E) \quad (11.211)$$

となる。したがって、

$$p(E) = \frac{D_{\text{sys}}(E) e^{-\beta E}}{\int dE D_{\text{sys}}(E) e^{-\beta E}} \quad (11.212)$$

$$= \frac{D_{\text{sys}}(E) e^{-\beta E}}{Z} \quad (11.213)$$

となる。ただし、

$$Z = \int dE D_{\text{sys}}(E) e^{-\beta E} \quad (11.214)$$

である。これがカノニカル分布である。

シャノンエントロピーを、拘束条件のもと最大化する方法

エネルギー期待値を一定に保った時の、シャノンエントロピー最大化問題を考える。シャノンエントロピーは確率分布関数の汎関数であるから、変分によって停留点を探す。

$$\tilde{S} = -k_B \sum_i p_i \log p_i - \lambda \left(\sum_i p_i - 1 \right) - \rho \left(\sum_i p_i E_i - U \right) \quad (11.215)$$

である。ここで、 λ と ρ は未定乗数である。微小な確率変分を考えると、

$$\delta \tilde{S} = -k_B \sum_i (p_i + \delta p_i) \log(p_i + \delta p_i) - \lambda \left(\sum_i (p_i + \delta p_i) - 1 \right) - \rho \left(\sum_i (p_i + \delta p_i) E_i - U \right) - \tilde{S} \quad (11.216)$$

$$= -k_B \sum_i \left(\delta p_i \log p_i + (p_i + \delta p_i) \log \left(1 + \frac{\delta p_i}{p_i} \right) \right) - \lambda \left(\sum_i \delta p_i \right) - \rho \left(\sum_i \delta p_i E_i \right) \quad (11.217)$$

となる。ここで、 $\log(1+x) = x + O(x^2)$ であることを用いて、

$$\delta \tilde{S} = -k_B \sum_i (\delta p_i \log p_i + \delta p_i) - \lambda \left(\sum_i \delta p_i \right) - \rho \left(\sum_i \delta p_i E_i \right) \quad (11.218)$$

$$= \sum_i \delta p_i (-k_B \log p_i - k_B - \lambda - \rho E_i) + O(\delta p_i^2) \quad (11.219)$$

となる。したがって、

$$-k_B \log p_i - k_B - \lambda - \rho E_i = 0 \quad (11.220)$$

である。これを変形して、

$$p_i = \exp \left(-1 - \frac{\lambda}{k_B} - \frac{\rho E_i}{k_B} \right) \quad (11.221)$$

したがって、

$$p_i \propto \exp(-\beta E_i) \quad (11.222)$$

となる。ここで、 $\beta = \frac{1}{k_B T}$ である。 $(\rho = \frac{1}{T})$

あとは、規格化条件から、

$$p_i = \frac{1}{Z} \exp(-\beta E_i) \quad (11.223)$$

$$Z = \sum_i \exp(-\beta E_i) \quad (11.224)$$

であることがわかる。これがカノニカル分布である。

11.11 典型性に関する追記

12 非平衡統計力学

白石本や田崎非平衡の動画を参考にしてまとめる。

12.0.1 記号

記号をまとめておく。

時間反転

時間反転に関する記号をまとめておく。

$\bar{\omega}$ は時間反転状態を表す。例えば、ある粒子の状態が位相空間上で (x, p) であるとき、この時間反転は $(x, -p)$ である。

また、物理量について、 A^\dagger は逆向きの経路についての物理量である。また、逆向きの経路の確率や遷移レートについても、 P^\dagger や R^\dagger と表す。

12.1 必要な数学

必要な数学をまとめておく。

12.1.1 確率微分方程式

Markov 連鎖

状態も時間も離散的であるような確率過程を考える。このとき、状態間の遷移は以下のように行われる。

Def. 確率遷移行列

p^{n-1} から p^n への遷移を表す確率遷移行列を、

$$p_i^n = \sum_j T_{ij} p_j^{n-1} \quad (12.1)$$

により定義する。このとき、

$$T_{ij} \geq 0 \quad \text{非負性} \quad (12.2)$$

$$\sum_i T_{ij} = 1 \quad \text{規格化} \quad (12.3)$$

$$(12.4)$$

である。

このとき、規格化がうまくいっていることは以下のように確かめられる。

$$\sum_i p_i^n = \sum_i \sum_j T_{ij} p_j^{n-1} \quad (12.5)$$

$$= \sum_j p_j^{n-1} \sum_i T_{ij} \quad (12.6)$$

$$= \sum_j p_j^{n-1} \quad (12.7)$$

$$= 1 \quad (12.8)$$

Markov jump 過程

時間は連続だが状態は離散であるようなものを考える。

Def. 遷移レート行列

以下によって遷移レート行列を定義する:

$$T_{ij} = \begin{cases} R_{ij}(n\Delta)\Delta t & \text{if } i \neq j \\ 1 - \sum_k R_{kj}(n\Delta)\Delta t & \text{if } i = j \end{cases} \quad (12.9)$$

ただし、

$$R_{ij}(t) \geq 0 \quad (i \neq j) \quad (12.10)$$

$$\sum_i R_{ij} = 0 \quad (12.11)$$

である。

Thm. マスター方程式

Markov jump 過程において、確率分布は以下に従って時間発展する:

$$\frac{d}{dt}p_i(t) = \sum_j R_{ij}p_j(t) \quad (12.12)$$

Prf.

$$p_i^n = \sum_j T_{ij}p_j^{n-1} \quad (12.13)$$

のもとで、

$$p_i^n - p_i^{n-1} = \sum_j T_{ij}p_j^{n-1} - p_i^{n-1} \quad (12.14)$$

$$= \sum_{j(\neq i)} T_{ij}p_j^{n-1} + T_{ii}p_i^{n-1} - p_i^{n-1} \quad (12.15)$$

$$= \sum_{j(\neq i)} R_{ij}\Delta t p_j^{n-1} + (1 - \sum_k R_{ki}\Delta t)p_i^{n-1} - p_i^{n-1} \quad (12.16)$$

$$= - \sum_k R_{ki}\Delta t p_i^{n-1}\Delta t + \sum_{j(\neq i)} R_{ij}p_j^{n-1}\Delta t \quad (12.17)$$

$$= R_{jj}p_j^{n-1}\Delta t + \sum_{j(\neq i)} R_{ij}p_j^{n-1}\Delta t \quad (12.18)$$

$$= \sum_j R_{ij}p_j^{n-1}\Delta t \quad (12.19)$$

となる。したがって、

$$\frac{d}{dt}p_i(t) = \lim_{\Delta t \rightarrow 0} \frac{p_i^n - p_i^{n-1}}{\Delta t} = \sum_j R_{ij}p_j(t) \quad (12.20)$$

となる。 □

Def.escape rate

状態 i からそれ以外に遷移する遷移率を escape rate という。すなわち、

$$e_{i,t} = \sum_{j(\neq i)} R_{ij} \quad (12.21)$$

を escape rate という。

Req. escape rate と時間反転

escape rate は時間反転に対して不変である。すなわち、

$$e_{\omega,t} = e_{\bar{\omega},t}^{\dagger} \quad (12.22)$$

である。

収束定理

Def. 既約

Thm. 収束定理

12.1.2 経路積分

12.2 熱力学量の定義

まずは、熱力学第一法則を構成するエネルギー、熱、仕事についてみていく。
エネルギーは通常力学で用いるハミルトニアンで定義される。このとき、エネルギーについて以下の要請をする。

Req. エネルギーと時間反転

エネルギーは以下を満たす。

$$E_{\omega} = E_{\bar{\omega}} = E_{\omega}^{\dagger} \quad (12.23)$$

Def. 熱

状態 ω から ω' への遷移において、系が熱浴に放出する熱は以下で定義される:

$$\hat{Q}_{\omega \rightarrow \omega'} = E_{\omega} - E_{\omega'} = \frac{1}{\beta} \ln \frac{P_{\omega \rightarrow \omega'}}{P_{\omega' \rightarrow \omega}^{\dagger}} \quad (12.24)$$

Def. 仕事

系の状態が ω であるとする。このとき、系への仕事は制御パラメータ λ がポテンシャルを変化させることによって行われる。単位時間あたりに系が外部へ取り出す仕事の量は、以下で定義される:

$$\hat{W} = - \frac{dE_{\omega}^{\lambda(t)}}{d\lambda(t)} \frac{d\lambda(t)}{dt} \quad (12.25)$$

以上の定義から、熱力学第一法則は以下のように書くことができる:

$$W + Q = -\Delta E \quad (12.26)$$

ただし、 ΔE は系の平均のエネルギーである。

また、系のエントロピーは以下の Shanon エントロピーで定義される:

Def. エントロピー/自己情報量

系がつかう率分布 \mathbf{p} に従うとき、系の Shanon エントロピーは以下で定義される:

$$S(\mathbf{p}) = - \sum_{\omega} p_{\omega} \ln p_{\omega} \quad (12.27)$$

また、この量は、 $-\log p_{\omega}$ の期待値である。このとき、

$$s(\omega) = -\ln p_{\omega} \quad (12.28)$$

を自己情報量という。

非平衡統計力学において第二法則を考えると、エントロピーではなく、以下のエントロピー生成について考える。

Def. エントロピー生成/エントロピー生成率

時刻 0 から τ の間のエントロピー生成は以下で定義される:

$$\hat{\sigma} = \sum_{\nu} \beta_{\nu} \hat{Q}_{\nu} + \hat{s}(\omega(\tau)) - \hat{s}(\omega(0)) \quad (12.29)$$

また、時刻 0 から τ の間の平均のエントロピー生成は以下で定義される:

$$\sigma = \langle \hat{\sigma} \rangle \quad (12.30)$$

$$= \sum_{\nu} \beta_{\nu} Q_{\nu} + S(\mathbf{p}(\tau)) - S(\mathbf{p}(0)) \quad (12.31)$$

また、単位時間当たりのエントロピー生成をエントロピー生成率といい、以下のように書くことができる:

$$\dot{\sigma} = \sum_{\omega, \omega'} p_{\omega} P_{\omega \rightarrow \omega'} \ln \frac{p_{\omega} P_{\omega \rightarrow \omega'}}{p_{\omega'} P_{\omega' \rightarrow \omega}} \quad (12.32)$$

(\because (エントロピーの生成率の表式が (12.32) 式であることの証明))

この系と熱浴の項はそれぞれ次のように計算される。この系と熱浴の項はそれぞれ次のように計算される。

$$\frac{d}{dt} H(\mathbf{p}(t)) = - \frac{d}{dt} \sum_w p_w \ln p_w \quad (12.33)$$

$$= - \sum_w \frac{dp_w}{dt} \ln p_w \quad (12.34)$$

第 2 項は確率の規格化から消える。第 1 項にマスター方程式を代入して

$$= \sum_w \left(\sum_{w' (\neq w), \nu} p_w P_{w \rightarrow w'}^{\nu} - p_{w'} P_{w' \rightarrow w}^{\nu} \right) \ln p_w \quad (12.35)$$

和の記号は正確には熱浴の区別 ω, ω' 的なものが入るべきであるが、教科書にノーテーションを合わせる。

$$= \sum_{\omega, \omega'} (p_{\omega'} P_{\omega' \rightarrow \omega} - p_{\omega} P_{\omega \rightarrow \omega'}) \ln p_{\omega} \quad (12.36)$$

$$(\text{第 2 項の添字を入れ替えて}) \quad (12.37)$$

$$= \sum_{w, w'} p_w P_{w \rightarrow w'} \ln p_w - \sum_{w, w'} p_w P_{w \rightarrow w'} \ln p_{w'} \quad (12.38)$$

$$= \sum_{w, w'} p_w P_{w \rightarrow w'} \ln \frac{p_w}{p_{w'}} \quad (12.39)$$

$$\sum_{\nu=1}^k \beta_{\nu} \dot{Q}_{\nu} = \sum_{\nu} \beta_{\nu} \left(\frac{1}{\beta_{\nu}} \sum_w \frac{dp_w}{dt} \ln \frac{P_{w \rightarrow w'}^{\nu}}{P_{w' \rightarrow \bar{w}}^{\dagger, \nu}} \right) \quad (12.40)$$

$$= \sum_{\nu} \sum_{w, w'} p_w P_{w \rightarrow w'}^{\nu} \ln \frac{P_{w \rightarrow w'}^{\nu}}{P_{w' \rightarrow \bar{w}}^{\dagger, \nu}} \quad (12.41)$$

$$(12.42)$$

以上より示された。

このとき、KL ダイバージェンスの非負性から、エントロピー生成率は非負であることがわかる。

非平衡統計力学においては、以下のように平衡状態を定義する。

Def. 平衡状態

系が平衡状態にあるとは、

$$p_{\omega}^{\text{eq}} P_{\omega \rightarrow \omega'} = p_{\omega'}^{\text{eq}} P_{\omega' \rightarrow \bar{\omega}}^{\dagger} \quad (12.43)$$

が成り立つことである。

また、物理的に有効であるような確率過程においては、以下の条件が成り立つと考えられる。

Def. 詳細釣り合い条件

任意の状態 ω, ω' および定常状態 $p_{\omega}^{\text{ss}}, p_{\omega'}^{\text{ss}}$ に対して、以下の条件が成り立つとき、系は詳細釣り合い条件を満たすという：

$$p_{\omega}^{\text{ss}} P_{\omega \rightarrow \omega'} = p_{\omega'}^{\text{ss}} P_{\omega' \rightarrow \omega} \quad (12.44)$$

例えば、系が熱平衡にある時は上の条件が満たされていることが確認できる。また、上の定常状態が平衡状態であるときは、以下のようにして詳細釣り合い条件を示すことができる。

いくつかのポテンシャルの底が存在しているような状態を考える。各ポテンシャルの底を i, j などと表し、その周りの十分小さな (他の底を含まないような領域) を S_i, S_j などと表す。また、位相空間上の点を X で表す。このとき、制限されたカノニカル分布を以下によって導入する：

$$p_j(X) = \frac{1}{Z_i} \chi_j(X) e^{-\beta H(X)} \quad (12.45)$$

$$Z_i = \int dX \chi_j(X) e^{-\beta H(X)} \quad (12.46)$$

ただし、 $\chi_j(X)$ は、 X が S_j に含まれるとき 1、それ以外るとき 0 となる関数である。このとき、状態 j から状態 k への遷移確率は以下で表される：

$$P_{j \rightarrow k}^{\tau} = \int dX \chi_k(J_{\tau}(X)) p_j(X) \quad (12.47)$$

ただし、 $J_\tau(X)$ は、 X を τ だけ時間発展させる写像である。この遷移確率を、時間反転の表式に書き直していく。いま、

$$X' = (J_t(X))^* \quad (12.48)$$

$$dX' = dX \quad (12.49)$$

$$X^* = J_t(X') \quad (12.50)$$

$$H(X) = H(X') \quad (12.51)$$

である。これを用いて、

$$(12.47) = \frac{1}{Z_j} \int dX \chi_k(J_\tau(X)) \chi_j(X) e^{-\beta H(X)} \quad (12.52)$$

$$= \frac{1}{Z_j} \int dX' \chi_k(X') \chi_j(J_\tau(X')) e^{-\beta H(X')} \quad (12.53)$$

$$= \frac{Z_k}{Z_j} \frac{1}{Z_k} \int dX' \chi_k(X') \chi_j(J_\tau(X')) e^{-\beta H(X')} \quad (12.54)$$

$$= \frac{Z_k}{Z_j} P_{k \rightarrow j}^\tau \quad (12.55)$$

となる。したがって、

$$Z_j P_{j \rightarrow k}^\tau = Z_k P_{k \rightarrow j}^\tau \quad (12.56)$$

となり、両辺 Z で割ることで、詳細釣り合い条件が成り立つことがわかる。

また、環境が平衡状態でないときなどは、詳細釣り合い条件ではなく、以下の局所詳細釣り合い条件を用いる。

Def. 局所詳細釣り合い条件

系が複数の熱浴に接しているような状況を考える。このとき、熱浴を ν によりラベル付けし、その逆温度を β_ν とする。このとき、任意の状態 ω, ω' に対して、以下の条件が成り立つとき、系は局所詳細釣り合い条件を満たすという：

$$(E_\omega - E_{\omega'}) = \frac{1}{\beta_\nu} \ln \frac{P_{\omega \rightarrow \omega'}^\nu}{P_{\omega' \rightarrow \omega}^\nu} \quad (12.57)$$

12.3 ゆらぎの定理

Thm.:DFT

系が経路 Γ をたどるとき、その経路におけるエントロピー生成 $\hat{\sigma}(\Gamma)$ について、

$$\text{Prob}[\hat{\sigma}(\Gamma) = \Sigma] e^{-\Sigma} = \text{Prob}^\dagger[\hat{\sigma}(\Gamma) = -\Sigma] \quad (12.58)$$

が成り立つ。

Prf.

いま、

$$P(\Gamma)e^{-\hat{\sigma}(\Gamma)} = p_{\omega^0, t^0} \prod_{n=1}^N P_{\omega^{n-1} \rightarrow \omega^n} \prod_{n=0}^N \exp\left(-\int_{t_n}^{t_{n+1}} dt e_{\omega^n, t^n}\right) \exp(-\hat{\sigma}(\Gamma)) \quad (12.59)$$

$$= p_{\omega^0, t^0} \prod_{n=1}^N P_{\omega^{n-1} \rightarrow \omega^n} \prod_{n=0}^N \exp\left(-\int_{t_n}^{t_{n+1}} dt e_{\omega^n, t^n}\right) \frac{p_{\omega^N, t^N}}{p_{\omega^0, t^0}} \prod_{n=1}^N \frac{P_{\bar{\omega}^n \rightarrow \bar{\omega}^{n-1}}^\dagger}{P_{\omega^{n-1} \rightarrow \omega^n}} \quad (12.60)$$

$$= p_{\omega^N, t^N} \prod_{n=1}^N P_{\bar{\omega}^n \rightarrow \bar{\omega}^{n-1}}^\dagger \prod_{n=0}^N \exp\left(-\int_{t_n}^{t_{n+1}} dt e_{\bar{\omega}^n, t^n}^\dagger\right) \quad (12.61)$$

$$= P^\dagger(\Gamma^\dagger) \quad (12.62)$$

となる。したがって、

$$\text{Prob}[\hat{\sigma}(\Gamma) = \Sigma] e^{-\Sigma} = \int d\Gamma P(\Gamma) \delta(\hat{\sigma}(\Gamma) - \Sigma) e^{-\Sigma} \quad (12.63)$$

$$= \int d\Gamma^\dagger P^\dagger(\Gamma^\dagger) \delta(-\hat{\sigma}(\Gamma^\dagger) - \Sigma) \quad (12.64)$$

$$= \text{Prob}^\dagger[\hat{\sigma}(\Gamma) = -\Sigma] \quad (12.65)$$

となる。 □

DFT は、確率的に、エントロピー生成が負になる可能性があることを示している。実際、エントロピー生成が負になる可能性は以下のように評価することができる。

$$\text{Prob}\left[\frac{\hat{\sigma}}{V} < -\epsilon\right] = \int_{-\infty}^{-\epsilon V} d\Sigma \text{Prob}[\hat{\sigma} = \Sigma] \quad (12.66)$$

$$= \int_{-\infty}^{-\epsilon V} d\Sigma e^{\Sigma} \text{Prob}^\dagger[\hat{\sigma} = -\Sigma] \quad (12.67)$$

$$= \int_{\epsilon V}^{\infty} d\Sigma e^{-\Sigma} \text{Prob}^\dagger[\hat{\sigma} = \Sigma] \quad (12.68)$$

$$\leq e^{-\epsilon V} \int_{\epsilon V}^{\infty} d\Sigma \text{Prob}^\dagger[\hat{\sigma} = \Sigma] \quad (12.69)$$

$$\leq e^{-\epsilon V} \int_{-\infty}^{\infty} d\Sigma \text{Prob}^\dagger[\hat{\sigma} = \Sigma] \quad (12.70)$$

$$= e^{-\epsilon V} \quad (12.71)$$

したがって、熱力学極限においては、エントロピー生成が負になる可能性は 0 となるが、体積が小さいときにはエントロピー生成がそれなりの確率で負になることがわかる。

Thm.:IFT

DFT を満たすような系について、以下の等式が成り立つ:

$$\langle e^{-\hat{\sigma}} \rangle = 1 \quad (12.72)$$

Prf.

$$\langle e^{-\hat{\sigma}} \rangle = \int d\Gamma P(\Gamma) e^{-\hat{\sigma}(\Gamma)} \quad (12.73)$$

$$= \int d\Gamma^\dagger P^\dagger(\Gamma^\dagger) \quad (12.74)$$

$$= 1 \quad (12.75)$$

となる。 □

Thm.: 熱力学第二法則

IFT を満たすような家庭について、平均のエントロピー生成は非負である。すなわち、

$$\langle \hat{\sigma} \rangle \geq 0 \quad (12.76)$$

が成り立つ。

Prf.

Jensen の不等式より、

$$1 = \langle e^{-\hat{\sigma}} \rangle \quad (12.77)$$

$$\int d\Gamma P(\Gamma) e^{-\hat{\sigma}(\Gamma)} \quad (12.78)$$

$$\geq \int d\Gamma e^{-\int dt \hat{\sigma}(\Gamma)} \quad (12.79)$$

$$= e^{-\langle \hat{\sigma} \rangle} \quad (12.80)$$

となる。両辺対数をとることで、

$$\langle \hat{\sigma} \rangle \geq 0 \quad (12.81)$$

となる。 □

Thm.Jarzynski 等式

等温環境にある系を考える。

$$\langle e^{\beta \hat{W}} \rangle = e^{-\beta \Delta F} \quad (12.82)$$

Prf.

Cor. 最大仕事の原理

Jarzynski 等式を満たすような過程において、以下の不等式が成り立つ:

$$\langle \hat{W}_{\text{out}} \rangle = -\langle \hat{W} \rangle \geq \Delta F \quad (12.83)$$

Prf.

Jarzynski 等式に Jensen の不等式を適用することで、

$$e^{-\beta \Delta F} = \langle e^{\beta \hat{W}} \rangle \quad (12.84)$$

$$\geq e^{\beta \langle \hat{W} \rangle} \quad (12.85)$$

となる。両辺対数をとることで、

$$\langle \hat{W} \rangle \leq -\Delta F \quad (12.86)$$

となる。

□

簡単な例

N 個の独立なスピン系を考える。状態は、 $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ で表される。このとき、エネルギーは、

$$E_\sigma(t) = \begin{cases} 0 & t \in [0, \tau/2) \\ -h \sum_{i=1}^N \sigma_i & t \in [\tau/2, \tau] \end{cases} \quad (12.87)$$

とする。このとき、

$$Z(\beta) = 2^N \quad (12.88)$$

$$Z'(\beta) = (2 \cosh(\beta h))^N \quad (12.89)$$

である。(それぞれ、時刻 $[0, \tau/2)$ と $[\tau/2, \tau]$ における分配関数である。) このとき、自由エネルギーの差は、

$$\Delta F = F(\beta) - F'(\beta) \quad (12.90)$$

$$= \frac{N}{\beta} \ln \left(\frac{Z'(\beta)}{Z(\beta)} \right) \quad (12.91)$$

$$= \frac{N}{\beta} \ln(\cosh(\beta h)) \quad (12.92)$$

$$\geq 0 \quad (12.93)$$

となる。また、

$$\hat{W} = 0 - E_\sigma(\tau) \quad (12.94)$$

$$= h \sum_{i=1}^N \sigma_i \quad (12.95)$$

である。したがって、

$$\langle \hat{W} \rangle = 0 \quad (12.96)$$

である。したがって、 $\langle \hat{W} \rangle \leq \Delta F$ が成り立つ。

また、 $N = 1$ のとき、

$$\hat{W} = \begin{cases} h \geq \frac{1}{\beta} \ln(\cosh(\beta h)) \\ -h \leq \frac{1}{\beta} \ln(\cosh(\beta h)) \end{cases} \quad (12.97)$$

が確率 $1/2$ で成り立つ。したがって、確率的に最大仕事の原理が破れることがわかる。

12.4 NESS

12.4.1 揺動散逸定理

ゆらぎの定理を用いることで、系に微小なカレントが生じているときの応答を導くことができる。例えば、系に二つの熱浴が接しており、その熱浴の温度が微小に違うようなとき、系に熱流が流れることになる。このとき、この熱流によってエネルギーが運ばれる。他にも、化学ポテンシャルの異なる粒子浴を用いる方法なども考えられる。

記号

- X : 示量変数
- \hat{J}_X : 示量変数 X のカレント:

$$\hat{J}_X = \frac{\partial X}{\partial t} \quad (12.98)$$

- $\hat{\mathcal{J}}_X$: \hat{J}_X の時間積分:

$$\int_0^\tau dt \hat{J}_X \quad (12.99)$$

- $\Delta\Pi_X$: X に共役な示強変数の微小な差

Thm. 揺動散逸定理

$$\langle \hat{J}_X \rangle_{\Delta\Pi_i} = \lim_{\tau \rightarrow \infty} \sum_i \sum_j \frac{1}{2\tau} \langle \hat{\mathcal{J}}_{X_i} \hat{\mathcal{J}}_{X_j} \rangle_0 \Delta\Pi_{X_j} \quad (12.100)$$

この式の意味するところは、微小な流れが生じると、その応答として、平衡状態におけるカレントの揺らぎが生じるということである。

Prf.

とくにカレントが2つの場合を考える。このとき、DFTを用いて、

$$\langle \hat{\mathcal{J}}_X \rangle_{\Delta\Pi_X} = \int d\hat{\mathcal{J}}_X \int d\hat{\mathcal{J}}_Y \hat{\mathcal{J}}_X P(\hat{\mathcal{J}}_X, \hat{\mathcal{J}}_Y) \quad (12.101)$$

$$= \int d\hat{\mathcal{J}}_X \int d\hat{\mathcal{J}}_Y (\hat{\mathcal{J}}_X P(-\hat{\mathcal{J}}_X, -\hat{\mathcal{J}}_Y)) \exp(\hat{\mathcal{J}}_X \Delta\Pi_X + \hat{\mathcal{J}}_Y \Delta\Pi_Y) \quad (12.102)$$

$$= - \int d\hat{\mathcal{J}}_X \int d\hat{\mathcal{J}}_Y \hat{\mathcal{J}}_X P(\hat{\mathcal{J}}_X, \hat{\mathcal{J}}_Y) \exp(-\hat{\mathcal{J}}_X \Delta\Pi_X - \hat{\mathcal{J}}_Y \Delta\Pi_Y) \quad (12.103)$$

$$= - \langle \hat{\mathcal{J}}_X \rangle_{\Delta\Pi_X} + \langle \hat{\mathcal{J}}_X^2 \rangle_{\Delta\Pi} \Delta\Pi_X + \langle \hat{\mathcal{J}}_X \hat{\mathcal{J}}_Y \rangle_{\Delta\Pi} \Delta\Pi_Y \quad (12.104)$$

$$= - \langle \hat{\mathcal{J}}_X \rangle_{\Delta\Pi_X} + \langle \hat{\mathcal{J}}_X^2 \rangle_0 \Delta\Pi_X + \langle \hat{\mathcal{J}}_X \hat{\mathcal{J}}_Y \rangle_0 \Delta\Pi_Y \quad (12.105)$$

したがって、

$$\langle \hat{\mathcal{J}}_X \rangle_{\Delta\Pi_X} = \frac{1}{2} \langle \hat{\mathcal{J}}_X^2 \rangle_0 \Delta\Pi_X + \frac{1}{2} \langle \hat{\mathcal{J}}_X \hat{\mathcal{J}}_Y \rangle_0 \Delta\Pi_Y \quad (12.106)$$

となることがわかる。よって、

$$\langle \hat{\mathcal{J}}_X \rangle_{\Delta\Pi_X} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \langle \hat{\mathcal{J}}_X^2 \rangle_0 \Delta\Pi_X + \frac{1}{2\tau} \langle \hat{\mathcal{J}}_X \hat{\mathcal{J}}_Y \rangle_0 \Delta\Pi_Y \quad (12.107)$$

となる。これを一般化することで、揺動散逸定理が導かれる。□

Cor. Onsager の相反定理

カレントが2つの場合を考える。このとき、

$$\langle \hat{\mathcal{J}} \rangle_{\Delta\Pi_X} = L_{XX} \Delta\Pi_X + L_{XY} \Delta\Pi_Y \quad (12.108)$$

$$\langle \hat{\mathcal{J}} \rangle_{\Delta\Pi_Y} = L_{YX} \Delta\Pi_X + L_{YY} \Delta\Pi_Y \quad (12.109)$$

であるとき、

$$L_{XY} = L_{YX} \quad (12.110)$$

が成り立つ。

Prf.

揺動散逸定理から、

$$L_{XX} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \langle \hat{\mathcal{J}}_X^2 \rangle_0 \quad (12.111)$$

$$L_{XY} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \langle \hat{\mathcal{J}}_X \hat{\mathcal{J}}_Y \rangle_0 \quad (12.112)$$

$$L_{YX} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \langle \hat{\mathcal{J}}_Y \hat{\mathcal{J}}_X \rangle_0 \quad (12.113)$$

$$L_{YY} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \langle \hat{\mathcal{J}}_Y^2 \rangle_0 \quad (12.114)$$

であることがわかる。このとき、

$$L_{XY} = L_{YX} \quad (12.115)$$

となる。 □

12.4.2 パワーと効率のトレードオフ

エントロピー生成率の下限熱浴のエントロピー生成率は、

$$\theta_{j \rightarrow k}(t) = \log \left(\frac{\omega_{j \rightarrow k}(t)}{\omega_{k \rightarrow j}(t)} \right) \quad (12.116)$$

$$= \log \left(\frac{R_{kj}(t)}{R_{jk}(t)} \right) \quad (12.117)$$

である。このとき、熱浴と系の全エントロピー生成は、

$$\sigma(t) = \frac{d}{dt} S(\mathbf{p})(t) + \left\langle \hat{\theta}(t) \right\rangle_t \quad (12.118)$$

である。ここで、 $S(\mathbf{p})$ は系のシャノンエントロピーであり、 $\left\langle \hat{\theta}(t) \right\rangle_t$ は時刻 t における熱浴のエントロピー生成の期待値である。このとき、エントロピー生成の下限として、以下が知られている。

Thm: エントロピー生成の下限

エントロピー生成率は、以下の不等式を満たす:

$$\sigma(t) \geq \sum_{j,k(j \neq k)} \frac{(R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t))^2}{R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t)} \geq 0 \quad (12.119)$$

ただし、分子について、

$$j_{j \rightarrow k}(t) = R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t) \quad (12.120)$$

である。

この式は、カレントが nonzero であるとき、エントロピー生成が nonzero になることを表している。

Prf

$$\frac{d}{dt}S(\mathbf{p}(t)) = -\frac{d}{dt}\sum_k^{\Omega} p_k(t) \log p_k(t) \quad (12.121)$$

$$= -\sum_k^{\Omega} \frac{d}{dt} p_k(t) \log p_k(t) - \sum_k^{\Omega} p_k(t) \frac{d}{dt} \log p_k(t) \quad (12.122)$$

$$= -\sum_k^{\Omega} \frac{d}{dt} p_k(t) \log p_k(t) - \sum_k^{\Omega} \frac{d}{dt} p_k(t) \quad (12.123)$$

$$= -\sum_k^{\Omega} \frac{d}{dt} p_k(t) \log p_k(t) \quad (\because \text{規格化条件}) \quad (12.124)$$

$$= -\sum_{j,k}^{\Omega} R_{kj}(t) p_j(t) \log p_j(t) \quad (12.125)$$

$$= \sum_{j,k}^{\Omega} R_{kj}(t) p_j(t) \log \left(\frac{p_j(t)}{p_k(t)} \right) \quad (\because \sum_k^{\Omega} R_{kj}(t) = 0) \quad (12.126)$$

である。また、

$$\left\langle \hat{\theta}(t) \right\rangle_t = \sum_{j,k(j \neq k)}^{\Omega} R_{kj}(t) p_j(t) \log \left(\frac{R_{kj}(t)}{R_{jk}(t)} \right) \quad (12.127)$$

$$= \sum_{j,k}^{\Omega} R_{kj}(t) p_j(t) \log \left(\frac{R_{kj}(t) p_j(t)}{R_{jk}(t) p_k(t)} \right) \quad (\because \log 1 = 0) \quad (12.128)$$

である。これらを用いて、

$$\sigma(t) = \frac{d}{dt}S(\mathbf{p}(t)) + \left\langle \hat{\theta}(t) \right\rangle_t \quad (12.129)$$

$$= \sum_{j,k}^{\Omega} R_{kj}(t) p_j(t) \log \left(\frac{R_{kj}(t) p_j(t)}{R_{jk}(t) p_k(t)} \right) \quad (12.130)$$

$$= \sum_{j,k(j \neq k)}^{\Omega} R_{kj}(t) p_j(t) \log \left(\frac{R_{kj}(t) p_j(t)}{R_{jk}(t) p_k(t)} \right) \quad (12.131)$$

$$= \frac{1}{2} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t) p_j(t) - R_{jk}(t) p_k(t)) \log \left(\frac{R_{kj}(t) p_j(t)}{R_{jk}(t) p_k(t)} \right) \quad (\because \text{添え字に対する対称性}) \quad (12.132)$$

$$\geq \sum_{j,k(j \neq k)}^{\Omega} \frac{(R_{kj}(t) p_j(t) - R_{jk}(t) p_k(t))^2}{2(R_{kj}(t) p_j(t) + R_{jk}(t) p_k(t))} \quad \left(\because \frac{1}{2}(a-b) \log(a/b) \geq \frac{(a-b)^2}{2(a+b)} \right) \quad (12.133)$$

である。よって示された。□

Improved Shiraishi-Saito inequality

$g_{j \rightarrow k}(t)$ を、時刻 t における jump quantity とし、

$$g_{j \rightarrow k}(t) = -g_{k \rightarrow j}(t) \quad (j \neq k) \quad (12.134)$$

とする。このとき、以下の不等式が成立する。

Thm:Improved Shiraishi-Saito inequality

jump quantity の平均値は、以下の不等式を満たす:

$$|\langle \hat{g} \rangle_t| \leq \sqrt{\sigma(t) \Xi_g(t)} \quad (12.135)$$

ただし、 $\sigma(t)$ はエントロピー生成率であり、 $\Xi_g(t)$ は、

$$\Xi_g(t) = \frac{1}{4} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t))g_{j \rightarrow k}(t)^2 \quad (12.136)$$

である。

Prf

jump quantity の平均値は、

$$\langle \hat{g}(t) \rangle_t = \sum_{j,k(j \neq k)}^{\Omega} p_j(t) \omega_{j \rightarrow k}(t) g_{j \rightarrow k}(t) \quad (12.137)$$

$$= \sum_{j,k(j \neq k)}^{\Omega} R_{kj}(t) p_j(t) g_{j \rightarrow k}(t) \quad (12.138)$$

$$= \frac{1}{2} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t))g_{j \rightarrow k}(t) \quad (12.139)$$

$$= \frac{1}{2} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t))g_{j \rightarrow k}(t) \quad (\because \text{添え字に対する対称性}) \quad (12.140)$$

$$= \frac{1}{2} \sum_{j,k(j \neq k)}^{\Omega} \frac{R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t)}{\sqrt{R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t)}} \sqrt{R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t)} g_{j \rightarrow k}(t) \quad (12.141)$$

である。ここで、Cauchy-Schwarz の不等式を用いると、

$$|\langle \hat{g} \rangle_t| \leq \sqrt{\sum_{j,k(j \neq k)}^{\Omega} \frac{(R_{kj}(t)p_j(t) - R_{jk}(t)p_k(t))^2}{R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t)} \cdot \frac{1}{4} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t))g_{j \rightarrow k}(t)^2} \quad (12.142)$$

である。ここで、第一項は $\sigma(t)$ である。第二項を $\Xi_g(t)$ とすると、

$$|\langle \hat{g} \rangle_t| \leq \sqrt{\sigma(t) \Xi_g(t)} \quad (12.143)$$

である。よって示された。□

この不等式を変形することで、

$$\sigma(t) \geq \frac{|\langle \hat{g} \rangle_t|^2}{\Xi_g(t)} \quad (12.144)$$

となる。これは、jump quantity の平均値が nonzero であるとき、エントロピー生成が nonzero になることを表している。

このとき、

$$\Xi_g(t) = \frac{1}{4} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t) + R_{jk}(t)p_k(t))g_{j \rightarrow k}(t)^2 \quad (12.145)$$

$$= \frac{1}{2} \sum_{j,k(j \neq k)}^{\Omega} (R_{kj}(t)p_j(t))g_{j \rightarrow k}(t)^2 \quad (12.146)$$

は一般には有限であることが知られている。とくに、jump quantity が有限かつ、escape rate が有限であるとき、すなわち、

$$|g_{j \rightarrow k}(t)| \leq g_0, \quad \lambda_j \leq \lambda_0 \quad (12.147)$$

であるとき、

$$\Xi_g(t) \leq \frac{1}{2} g_0^2 \lambda_0 \quad (12.148)$$

である。

得られた不等式を変形すると、

$$\sigma(t) \geq \frac{|\langle \hat{g} \rangle_t|^2}{\Xi_g(t)} \quad (12.149)$$

である。これは、jump quantity の平均値が nonzero であるとき、エントロピー生成率が nonzero になることを表している。

また、時間平均をとることで、以下の不等式が成り立つことが知られている。

Thm:Improved Shiraishi-Saito inequality(時間平均)

jump quantity の時間平均値は、以下の不等式を満たす:

$$\overline{\Xi}_g^{-1} \left(\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{g} \rangle_t \right)^2 \leq \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{\theta} \rangle_t \quad (12.150)$$

ただし、 $\overline{\Xi}_g$ は、

$$\overline{\Xi}_g = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \Xi_g(t) \quad (12.151)$$

である。

Prf

普通の Shiraishi-Saito 不等式より、

$$\langle \hat{g} \rangle_t \leq \sqrt{\sigma(t) \Xi_g(t)} \quad (12.152)$$

である。Cauchy-Schwarz の不等式より、

$$\frac{1}{\tau} \int_0^\tau dt \langle \hat{g}(t) \rangle_t \leq \sqrt{\frac{1}{\tau} \int_0^\tau dt \sigma(t) \frac{1}{\tau} \int_0^\tau dt \Xi_g(t)} \quad (12.153)$$

である。ところで、

$$\frac{1}{\tau} \int_0^\tau dt \langle \sigma(t) \rangle_t = \frac{1}{\tau} (S(\mathbf{p}(\tau)) - S(\mathbf{p}(0))) + \frac{1}{\tau} \int_0^\tau dt \langle \hat{\theta}(t) \rangle_t \quad (12.154)$$

である。この右辺第一項は、 $\tau \rightarrow \infty$ で、 $S(\mathbf{p}(\tau)) - S(\mathbf{p}(0)) \rightarrow 0$ である。よって、

$$\frac{1}{\tau} \int_0^\tau dt \langle g(t) \rangle_t \leq \sqrt{\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \theta(t) \rangle_t \overline{\Xi}_g} \quad (12.155)$$

である。よって示された。□

この不等式は、jump quantity の平均値が nonzero であるとき、(カレントが nonzero であるとき) 平均の (熱浴での) 散逸が nonzero になることを表している。

ここで、長時間平均により Shanon エントロピーの項が消えることは重要である。というのも、シャノンエントロピーを知るためには粒子のミクロな状態を知らなくてはならないが、実際問題それが難しいためである。

**パワーと効率のトレードオフ
設定**

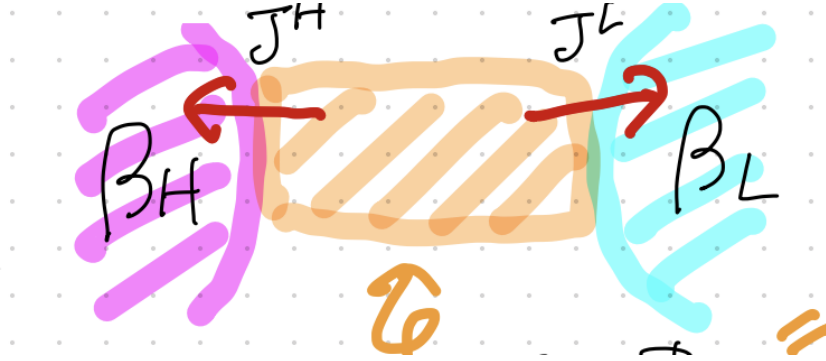


図3 設定

系が二つの熱浴に接触しているとする。 $\alpha = H, L$ で、それぞれの熱浴の逆温度を β_α 、遷移レートを $\omega_{j \rightarrow k}(t)$ 、系のエネルギーを $E_j(t)$ とする。このとき、局所詳細釣り合い条件は、

$$\log \frac{\omega_{j \rightarrow k}(t)}{\omega_{k \rightarrow j}(t)} = \beta_{\alpha(j,k)}(E_j(t) - E_k(t)) \quad (12.156)$$

である。 $(\alpha(j,k) = H, L)$

熱浴への熱の流れは、

$$J_{j \rightarrow k}^\alpha(t) = \begin{cases} E_j(t) - E_k(t) & \alpha(j,k) = \alpha \\ 0 & \text{otherwise} \end{cases} \quad (12.157)$$

である。また、エントロピー生成は、

$$\theta_{j \rightarrow k}(t) \beta_{\alpha(j,k)}(E_j(t) - E_k(t)) \quad (12.158)$$

$$= \beta_H J_{j \rightarrow k}^H(t) + \beta_L J_{j \rightarrow k}^L(t) \quad (12.159)$$

である。

仮定

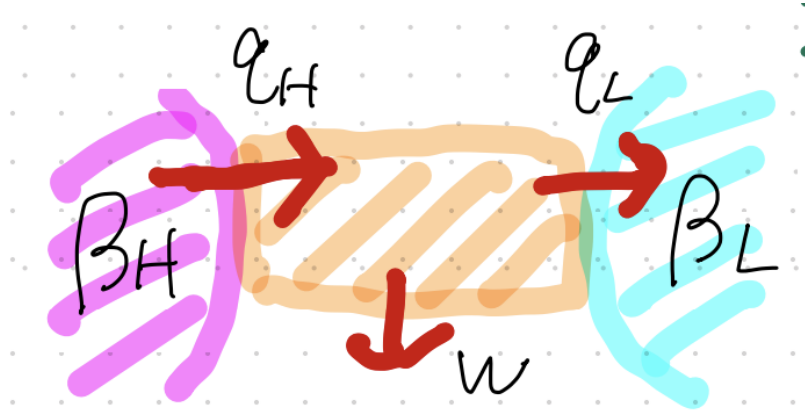


図4 仮定

以下の極限が存在し、その極限值が正であるとする:

$$q_H = - \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{J}^H(t) \rangle_t > 0 \quad (12.160)$$

$$q_L = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{J}^L(t) \rangle_t > 0 \quad (12.161)$$

これは、高温熱浴から系に入ってくる熱流および、低温熱浴に流れる熱流の時間平均である。また、このとき、平均のパワー (仕事率) は、

$$W = q_H - q_L \quad (12.162)$$

であり、効率は、

$$\eta = \frac{W}{q_H} = 1 - \frac{q_L}{q_H} \quad (12.163)$$

である。このとき、以下の不等式が成立する。

Thm: 熱機関の、パワーと効率のトレードオフ

熱機関のパワーと効率は、以下の不等式を満たす:

$$(q_H + q_L)^2 \leq \Xi \beta_L q_H (\eta_C - \eta) \quad (12.164)$$

Prf

今、

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{\theta}_t \rangle_t = -\beta_H q_H + \beta_L q_L \quad (12.165)$$

$$= -\beta_H q_H + \beta_L (q_H - W) \quad (12.166)$$

$$= \beta_L q_H \left(-\frac{\beta_H}{\beta_L} + 1 - \frac{W}{q_H} \right) \quad (12.167)$$

$$= \beta_L q_H (\eta_C - \eta) (\geq 0) \quad (12.168)$$

である。また、

$$g_{j \rightarrow k}(t) = \hat{J}_{j \rightarrow k}^L(t) - \hat{J}_{j \rightarrow k}^H(t) \quad (12.169)$$

とする。このとき、Shiraishi-Saito 不等式より、

$$\left(\frac{1}{\tau} \int_0^\tau dt \langle \hat{g}(t) \rangle_t \right)^2 \leq \bar{\Xi} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle \hat{\theta}_t \rangle_t \quad (12.170)$$

である。ここで、

$$\bar{\Xi} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \frac{1}{2} \sum_{j,k (j \neq k)}^\Omega R_{kj}(t) p_j(t) (E_j(t) - E_k(t))^2 \quad (12.171)$$

である。(これは、どれくらい active に熱のやり取りをするかを表している。) このとき、

$$\left(\frac{1}{\tau} \int_0^\tau dt \langle \hat{g}(t) \rangle_t \right)^2 = (q_H + q_L)^2 \quad (12.172)$$

である。よって、

$$(q_H + q_L)^2 \leq \bar{\Xi} \beta_L q_H (\eta_C - \eta) \quad (12.173)$$

である。よって示された。□

上で得られた不等式を変形すると、

$$\bar{\Xi} \beta_L (\eta_C - \eta) \geq \frac{(q_H + q_L)^2}{q_H} \geq q_H + q_L \quad (12.174)$$

となる。これは、効率がカルノーサイクルに近づくと、 q_H, q_L が 0 に近づくことを表している。このもとで、パワーは 0 に近づくことがわかる。

あるいは、以下のようにも変形できる:

$$W \leq \bar{\Xi} \beta_L \frac{q_H}{(q_H + q_L)^2} (\eta_C - \eta) W \quad (12.175)$$

$$= \bar{\Xi} \beta_L \frac{q_H}{(q_H + q_L)^2} \eta (\eta_C - \eta) \quad (\because W = q_H \eta) \quad (12.176)$$

$$\leq \bar{\Xi} \beta_L \eta (\eta_C - \eta) \quad (12.177)$$

これは、効率がカルノーサイクルに近づくと、パワーが 0 に近づくことを直接的に示している。

12.4.3 熱力学的不確定性関係

13 量子力学

14 相対論

15 場の量子論

16 その他