

熱力学的不確定性関係

大上由人

2024 年 11 月 25 日

1 TUR

1.1 Main Claim

記号

- $\hat{\mathcal{J}}_{ij} = \sum_n (\delta_{\omega_j \rightarrow \omega_i}(\omega^{n-1} \rightarrow \omega^n) - \delta_{\omega_i \rightarrow \omega_j}(\omega^{n-1} \rightarrow \omega^n))$
→ j から i への jump が起こった回数 - i から j への jump が起こった回数 (time-integrated empirical probability current)
- $\hat{\mathcal{J}}_d = \sum_{(i,j)} d_{ij} \hat{\mathcal{J}}_{ij}$ → 示量変数の積算カレント (cumulative thermodynamic current)
- $d_{ij} = X_i - X_j$ (X は任意の示量変数)

ex. 熱流

このとき、積算熱流は、 $X_i = E_i$ とすることで、以下のように定義される:

$$\hat{\mathcal{J}}_d = \sum_{(i,j)} (E_i - E_j) \hat{\mathcal{J}}_{ij} \quad (1.1)$$

また、jump quantity としては以下のように定義される

$$J_{ij}(t) = R_{ij}p_j(t) - R_{ji}p_i(t) \quad (1.2)$$

$$J_d(t) = \sum_{(i,j)} d_{ij} J_{ij}(t) \quad (1.3)$$

また、以下では、 $\langle \cdot \rangle$ で時間積分された量の平均も、jump quantity の平均も表すこととする。ただし、後者は、

$$\langle X \rangle = \sum_{i,j} X_{ij} R_{ij} p_j(t) \quad (1.4)$$

である。

このとき、以下の定理が成り立つことが知られている：

Thm. 熱力学的不確定性関係

定常 Markov 過程が局所詳細釣り合い条件を満たすとき、以下の関係が成り立つことが知られている：

$$\frac{\text{Var}(\mathcal{J}_d)}{(\mathcal{J}_d^{\text{ss}})^2} \sigma \geq 2 \quad (1.5)$$

ただし、

$$\text{Var}(\mathcal{J}_d) = \left\langle \left(\hat{\mathcal{J}}_d - \langle \hat{\mathcal{J}}_d \rangle \right)^2 \right\rangle \quad (1.6)$$

である。

すなわち、カレントのゆらぎを小さくするためには、エントロピー生成を大きくする必要がある。エントロピー生成が不可逆性の指標となることを踏まえると、カレントのゆらぎと不可逆性の間には、一定の関係があることがわかる。また、この定理はゆらぎの定理からの帰結ではない。その意味で、ゆらぎの定理とは異なるタイプの不可逆性の指標であることがわかる。

1.2 Cramér-Rao の不等式による証明

Cramér-Rao の不等式を用いて、熱力学的不確定性関係を証明する。^{*1} まず、準備として、Fisher 情報量を以下のように定義する。

Def: Fisher 情報量

パラメータ θ についての確率分布 $P_\theta(x)$ が与えられたとき、Fisher 情報量 $F(\theta)$ は以下のように定義される：

$$F(\theta) = - \left\langle \frac{\partial^2}{\partial \theta^2} \log P_\theta(x) \right\rangle = \left\langle \left(\frac{\partial}{\partial \theta} \log P_\theta(x) \right)^2 \right\rangle \quad (1.7)$$

$P_\theta(x)$ をパラメータ θ をもつ確率変数 x の確率分布であるとする。また、引数 θ の関数 $f(\theta)$ が与えられていたとする。ここでの目的は、 x を $g(x)$ として測定したとき、 $f(\theta)$ を推定することである。ここで、 $g(x)$ が $f(\theta)$ の不偏推定量であると仮定する。すなわち、

$$f(\theta) = \langle g(x) \rangle \quad (1.8)$$

が成り立つとする。このとき、 $g(x)$ がどれほど正確な推定量なのかの限界を表すのが Cramér-Rao の不等式である。

^{*1} ϵ を初めてタイプした。

Thm. 一般化 Cramér-Rao の不等式

$g(x)$ がパラメータ $f(\theta)$ の不偏推定量であるとき、

$$\text{Var}_\theta(g(x)) \geq \frac{(f'(\theta))^2}{F(\theta)} \quad (1.9)$$

が成り立つ。

Prf.

$$\text{Var}_\theta g(x) F(\theta) = \left(\int dx (g(x) - f(\theta))^2 P_\theta(x) \right) \left(\int dx \left(\frac{\partial}{\partial \theta} \log P_\theta(x) \right)^2 P_\theta(x) \right) \quad (1.10)$$

$$\geq \left(\int dx (g(x) - f(\theta)) \frac{\partial}{\partial \theta} (\ln P_\theta(x)) P_\theta(x) \right)^2 \quad \because \text{Cauchy-Schwarz の不等式} \quad (1.11)$$

$$= \left(\int dx g(x) \frac{\partial}{\partial \theta} P_\theta(x) - \int dx f(\theta) \frac{\partial}{\partial \theta} P_\theta(x) \right)^2 \quad (1.12)$$

$$= \left(\frac{\partial}{\partial \theta} \int dx g(x) P_\theta(x) \right)^2 \quad \because \text{第二項の積分は規格化条件より 0} \quad (1.13)$$

$$= \left(\frac{\partial}{\partial \theta} f(\theta) \right)^2 \quad \because \text{不偏推定量の仮定より} \quad (1.14)$$

$$(1.15)$$

□

Cor. Cramér-Rao の不等式

$g(x)$ がパラメータ θ の不偏推定量であるとき、

$$\text{Var}_\theta(g(x)) \geq \frac{1}{F(\theta)} \quad (1.16)$$

が成り立つ。

Prf.

$f(\theta) = \theta$ として、 $f'(\theta) = 1$ とすればよい。

□

Prf(熱力学的不確定性関係)

$$R_{ij}^\theta = R_{ij} e^{\theta Z_{ij}} \quad (i \neq j) \quad (1.17)$$

$$R_{ii}^\theta = - \sum_{j(\neq i)} R_{ij}^\theta e^{\theta Z_{ij}} \quad (1.18)$$

$$(1.19)$$

とおく。ただし、

$$Z_{ij} = \frac{R_{ij}p_j^{\text{ss}} - R_{ji}p_i^{\text{ss}}}{R_{ij}p_j^{\text{ss}} + R_{ji}p_i^{\text{ss}}} \quad (1.20)$$

である。このとき、

$$\begin{aligned} F(0) &= - \int_0^\tau dt \left[\sum_{i \neq j} R_{ij}p_j(t) \frac{\partial^2}{\partial \theta^2} \ln R_{ij}^\theta \Big|_{\theta=0} + \sum_j p_j(t) \frac{\partial^2}{\partial \theta^2} \ln e^{R_{jj}^\theta} \Big|_{\theta=0} \right] \\ &= \int_0^\tau dt \sum_j p_j(t) \frac{\partial^2}{\partial \theta^2} \sum_{k(\neq j)} R_{kj} e^{\theta Z_{kj}} \Big|_{\theta=0} \end{aligned} \quad (1.21)$$

$$\because \text{第一項は微分で消える。第二項について、} R_{jj} = - \sum_{k(\neq j)} R_{kj} \quad (1.22)$$

$$= \int_0^\tau dt \sum_{k \neq j} R_{kj} p_j(t) Z_{kj}^2 \quad \because \text{微分を実行} \quad (1.23)$$

$$= \left\langle \int_0^\tau dt \hat{Z}^2 \right\rangle \quad (1.24)$$

はじめの状態を、定常状態に取る。(定理の仮定より) このとき、 $R_{ij}^\theta = R_{ij}(1 + \theta Z_{ij}) + O(\theta^2)$ であることを用いて、

$$\sum_i R_{ij}^\theta p_j^{\text{ss}} - R_{ji}^\theta p_i^{\text{ss}} \quad (1.25)$$

$$= \sum_i R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}} + \theta(R_{ij} p_j^{\text{ss}} Z_{ij} - R_{ji} p_i^{\text{ss}} Z_{ji}) + O(\theta^2) \quad (1.26)$$

$$= \sum_i R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}} + \theta X_{ij} (R_{ij} p_j^{\text{ss}} + R_{ji} p_i^{\text{ss}}) + O(\theta^2) \quad (1.27)$$

$$= \sum_i R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}} + \theta X_{ij} (R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}}) + O(\theta^2) \quad (1.28)$$

$$= O(\theta^2) \quad (1.29)$$

となる。したがって、 θ 系の定常状態は、 $O(\theta)$ までの近似で、 $\theta = 0$ の定常状態と一致する。このときの定常カレントは、

$$\langle \hat{\mathcal{J}}_{ij}^\theta \rangle = R_{ij}^\theta p_j^{\text{ss}} - R_{ji}^\theta p_i^{\text{ss}} = O(\theta^2) \quad (1.30)$$

である。したがって、

$$\left. \frac{\partial \langle \hat{\mathcal{J}}_{ij} \rangle_{\theta}^{\text{ss}}}{\partial \theta} \right|_{\theta=0} \quad (1.31)$$

$$= \frac{\partial}{\partial \theta} (R_{ij}^{\theta} p_j^{\text{ss}} - R_{ji}^{\theta} p_i^{\text{ss}}) \Big|_{\theta=0} \quad (1.32)$$

$$= R_{ij} Z_{ij} p_j^{\text{ss}} - R_{ji} Z_{ji} p_i^{\text{ss}} \quad (1.33)$$

$$= Z_{ij} (R_{ij} p_j^{\text{ss}} + R_{ji} p_i^{\text{ss}}) \quad (1.34)$$

$$= R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}} \quad (1.35)$$

$$= \langle \hat{\mathcal{J}}_{ij} \rangle_0^{\text{ss}} \quad (1.36)$$

となる。このことから、

$$\left. \frac{\partial \langle \hat{\mathcal{J}}_d \rangle_{\theta}^{\text{ss}}}{\partial \theta} \right|_{\theta=0} = \langle \hat{\mathcal{J}}_d \rangle_0^{\text{ss}} \quad (1.37)$$

であることもわかる。また、

$$\langle \hat{Z}^2 \rangle^{\text{ss}} \quad (1.38)$$

$$= \frac{1}{2} \sum_{i,j} (Z_{ij}^2 p_j^{\text{ss}} R_{ij} + Z_{ji}^2 p_i^{\text{ss}} R_{ji}) \quad (1.39)$$

$$= \frac{1}{2} \sum_{i,j} Z_{ij}^2 (R_{ij} p_j^{\text{ss}} + R_{ji} p_i^{\text{ss}}) \quad (1.40)$$

$$= \frac{1}{2} \sum_{i,j} \frac{(R_{ij} p_j^{\text{ss}} - R_{ji} p_i^{\text{ss}})^2}{R_{ij} p_j^{\text{ss}} + R_{ji} p_i^{\text{ss}}} \quad (1.41)$$

$$\leq \frac{\sigma}{2} \quad \because 16 \text{ 章で示した不等式} \quad (1.42)$$

が得られる。

以上の関係式を一般化 Cramér-Rao の不等式

$$\text{Var}_{\theta}(g(x)) \geq \frac{(f'(\theta))^2}{F(\theta)} \quad (1.43)$$

の $\theta = 0$ の場合に適用すると、*2

$$\left(\langle \hat{\mathcal{J}}_d \rangle_0^{\text{ss}} \right)^2 \leq \left(\frac{\partial \langle \hat{\mathcal{J}}_d \rangle_{\theta}^{\text{ss}}}{\partial \theta} \right)^2 \Big|_{\theta=0} \leq \text{Var}_0(\hat{\mathcal{J}}_d) \int_0^{\tau} dt \langle \hat{Z}^2 \rangle \leq \frac{\sigma}{2} \text{Var}_0(\hat{\mathcal{J}}_d) \quad (1.44)$$

*2 ここで、 $x = \Gamma, g(\Gamma) = \hat{\mathcal{J}}_d, f(\theta) = \langle g(\Gamma) \rangle_{\theta}^{\text{ss}} = \langle \hat{\mathcal{J}}_d \rangle_{\theta}^{\text{ss}}$ とする。

が得られる。したがって、

$$\frac{\text{Var}(\hat{\mathcal{J}}_d)}{\left(\langle \hat{\mathcal{J}}_d \rangle^{\text{ss}}\right)^2} \sigma \geq 2 \quad (1.45)$$

が成り立つ。 \square

補足: $F(0)$ の表式

経路にわたっての期待値を計算するのだが、一旦時間を離散化したほうがわかりやすいので、離散化してマルコフ連鎖として考える。このとき、経路 Γ の実現確率は

$$P_\theta(\Gamma) = \prod_{n=1}^N T_{w^n w^{n-1}}^\theta p_{w^0}^0 \quad (1.46)$$

と書かれる。このもとで、m-step 目における確率分布は、

$$p_{w_m}^m = \sum_{w_0, \dots, w_{m-1}} \prod_{n=1}^m T_{w^n w^{n-1}}^\theta p_{w^0}^0 \quad (1.47)$$

と書かれる。これを下に、Fisher 情報量を計算する。

$$\int d\Gamma P(\Gamma) \frac{\partial^2}{\partial \theta^2} \ln P_\theta(\Gamma) = \sum_{w_0, \dots, w_N} \prod_{n=1}^N T_{w^n w^{n-1}}^\theta p_{w^0}^0 \frac{\partial^2}{\partial \theta^2} \left(\sum_{n=1}^N \ln T_{w^n w^{n-1}}^\theta + \ln p_{w^0}^0 \right) \quad (1.48)$$

$$= \sum_{w_0, \dots, w_N} \prod_{n=1}^N T_{w^n w^{n-1}}^\theta p_{w^0}^0 \frac{\partial^2}{\partial \theta^2} \left(\sum_{n=1}^N \ln T_{w^n w^{n-1}}^\theta \right) \quad (1.49)$$

$$(1.50)$$

一番最後の Σ について考える。和の k 番目の項について、

$$(\text{第 } k \text{ 項}) = \sum_{w_0, \dots, w_N} \prod_{n=1}^N T_{w^n w^{n-1}}^\theta p_{w^0}^0 \frac{\partial^2}{\partial \theta^2} \ln T_{w^k w^{k-1}}^\theta \quad (1.51)$$

(w^0 から w^{k-2} まで計算して、)

$$= \sum_{w_{k-1}, w_k, \dots, w_N} \prod_{n=k-1}^N T_{w^n w^{n-1}}^\theta p_{w^{k-1}}^{k-1} \frac{\partial^2}{\partial \theta^2} \ln T_{w^k w^{k-1}}^\theta \quad (1.52)$$

(w^{k+1} から w^N まで計算すると、規格化条件より)

$$= \sum_{w_{k-1}, w_k} T_{w^k w^{k-1}}^\theta p_{w^{k-1}}^{k-1} \frac{\partial^2}{\partial \theta^2} \ln T_{w^k w^{k-1}}^\theta \quad (1.53)$$

となる。したがって、

$$F(\theta) = - \sum_{n=1}^N \sum_{w_{n-1}, w_n} T_{w^n w^{n-1}}^\theta p_{w^{n-1}}^{n-1} \frac{\partial^2}{\partial \theta^2} \ln T_{w^n w^{n-1}}^\theta \quad (1.54)$$

となる。これを連続極限に持っていくと、 $w^n = w^{n-1}$ の場合と $w^n \neq w^{n-1}$ の場合に分けて、

$$F(\theta) = - \int dt \left[\sum_{i \neq j} R_{ij} p_j(t) \frac{\partial^2}{\partial \theta^2} \ln R_{ij}^\theta + \sum_j p_j(t) \frac{\partial^2}{\partial \theta^2} \ln e^{R_{jj}^\theta} \right] \quad (1.55)$$

となる。ただし、第二項については、 $T_{ii} = 1 - \sum_{j(\neq i)} R_{ij} \Delta t = 1 + R_{ii} \Delta t$ であることを用いて、

$$(1 + R_{ii} \Delta t) \ln(1 + R_{ii} \Delta t) = (1 + R_{ii} \Delta t) \ln \exp(R_{ii} \Delta t) + O(\Delta t^2) \quad (1.56)$$

$$= (1 + R_{ii} \Delta t) R_{ii} \Delta t + O(\Delta t^2) \quad (1.57)$$

$$= R_{ii} \Delta t + O(\Delta t^2) \quad (1.58)$$

$$= \ln e^{R_{ii} \Delta t} + O(\Delta t^2) \quad (1.59)$$

となることを用いている。

2 TUR-Type Inequalities

2.1 TUR の一般化

先程は、初期状態も含めて常に系の状態が定常状態であると仮定した。ここでは、遷移レートが固定されていて (すなわち、外部操作がない状況で) 系の状態が平衡状態とは限らないときに成り立つ TUR を考える。

すなわち、緩和過程における TUR を考える。

Thm. 緩和過程における TUR

連続時間 ($0 \leq t \leq \tau$) における Markov 過程について考える。ただし、遷移レートは固定されているとする。このとき、

$$\frac{\text{Var}(\mathcal{J}_d)}{(\tau J_d(\tau))^2} \sigma \geq 2 \quad (2.1)$$

が成り立つ。ただし、 $J_d(\tau)$ は、時刻 $t = \tau$ における平均のカレントである。

Prf.

$$Z_{ij} = \frac{R_{ij} p_j(t) - R_{ji} p_i(t)}{R_{ij} p_j(t) + R_{ji} p_i(t)} \quad (2.2)$$

$$(2.3)$$

とおく。(前の定理と同じ) 初期状態を、 θ とは独立に $\mathbf{p}(0)$ ととる。このとき、前の定理と同様に、

$$\int_0^\tau dt \langle \hat{Z}(t)^2 \rangle \leq \frac{\sigma}{2} \quad (2.4)$$

$$Z_{ij}(t)R_{ij}\mathbf{p}_j(t) - Z_{ji}(t)R_{ji}\mathbf{p}_i(t) = R_{ij}\mathbf{p}_j(t) - R_{ji}\mathbf{p}_i(t) \quad (2.5)$$

$$R^\theta(t)\mathbf{p}(t) = (1 + \theta)R\mathbf{p}(t) + O(\theta^2) \quad (2.6)$$

となる。3つ目の式を観察すると、 R^θ によって時間発展する確率分布は、

$$\mathbf{p}'(t) = \mathbf{p}((1 + \theta)t) + O(\theta^2) \quad (2.7)$$

となるのではないかと考えられる。すなわち、タイムスケールを $(1 + \theta)$ 倍にしたものになるということである。これを確かめるには、 $\mathbf{p}'(t)$ が R^θ によって時間発展するかを確かめればよい。実際、

$$\frac{d}{dt}\mathbf{p}'(t) = (1 + \theta) \frac{d}{dt'}\mathbf{p}(t') \Big|_{t'=(1+\theta)t} \quad (2.8)$$

$$= (1 + \theta) \frac{d}{dt'}\mathbf{p}(t') \Big|_{t'=t} + \frac{d}{dt''} \left(\frac{d}{dt'}\mathbf{p}(t') \right) \Big|_{t'=t''} \cdot \theta t + O(\theta^2) \quad (2.9)$$

$$\because t' \text{ について展開して、} (1 + \theta) \cdot \theta t = \theta t + O(\theta^2)$$

$$= R^\theta(t)\mathbf{p}(t) + \frac{d}{dt'}(R\mathbf{p}(t')) \Big|_t \cdot \theta t + O(\theta^2) \quad \because \text{master e.q.} \quad (2.10)$$

$$= R^\theta(t)\mathbf{p}((1 + \theta)t) + O(\theta^2) \quad \because R - R^\theta = O(\theta) \quad (2.11)$$

$$= R^\theta(t)\mathbf{p}'(t) + O(\theta^2). \quad (17.26)$$

したがって、 $\mathbf{p}'(t)$ は、 R^θ によって時間発展することがわかり、 R^θ によって時間発展する確率分布は、もとの系のタイムスケールを $(1 + \theta)$ 倍にしたものであることがわかる。

これを踏まえると、

$$\langle \hat{X} \rangle_\theta - \langle \hat{X} \rangle_0 = \int_0^{(1+\theta)\tau} dt J_d(t) - \int_0^\tau dt J_d(t) = \theta \tau J_d(\tau) + O(\theta^2).$$

である。以上のことと、一般化 Cramér-Rao の不等式

$$\text{Var}_\theta(g(x))F(\theta) \geq (f'(\theta))^2 \quad (2.12)$$

で $\theta = 0$ の場合を考えると、

$$(\tau J_d(\tau))^2 \leq \text{Var}_0(J_d) \cdot \frac{\sigma}{2} \quad (2.13)$$

が得られる。したがって、

$$\frac{\text{Var}(J_d)}{(\tau J_d(\tau))^2} \sigma \geq 2 \quad (2.14)$$

が成り立つ。 □

3 Cramér-Rao の不等式

統計学の文脈での Cramér-Rao の不等式について述べる。

$P_\theta(x)$ をパラメータ θ をもつ確率変数 x の確率分布であるとする。また、引数 θ の関数 $f(\theta)$ が与えられていたとする。ここでの目的は、 x を $g(x)$ として測定したとき、 $f(\theta)$ を推定することである。ここで、 $g(x)$ が $f(\theta)$ の不偏推定量であると仮定する。すなわち、

$$f(\theta) = \langle g(x) \rangle \quad (3.1)$$

が成り立つとする。

例えば、誤差が毎回でるような秤を用いて物の重さを測定することを考える。このとき、各測定値が x であり、真の体重が θ に対応する。仮に、測定値が以下のような θ まわりのガウス分布に従うとする：

$$P_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) \quad (3.2)$$

このとき、Cramér-Rao の不等式は以下のように表される：

$$\text{Var}_\theta(g(x)) \geq \frac{(f')^2}{F(\theta)} \quad (3.3)$$

ただし、 $F(\theta)$ は Fisher 情報量であり、以下のように定義される：

$$F(\theta) = -\left\langle \frac{\partial^2}{\partial \theta^2} \log P_\theta(x) \right\rangle = \left\langle \left(\frac{\partial}{\partial \theta} \log P_\theta(x) \right)^2 \right\rangle \quad (3.4)$$

Cramér-Rao の不等式の左辺は、どれほどこの推定が正確かを表している。右辺のフィッシャー情報量は、 θ に対する鋭敏性を表している。もし、 $F(\theta)$ が大きいとき、 θ の変化に対して、確率分布 $P_\theta(x)$ が大きく変化することになる。また、 $F(\theta)$ が小さい場合は、 θ の変化に対して、確率分布 $P_\theta(x)$ があまり変化しないことを意味しており、これは、測定値から θ を測定するのが難しいということの意味する。

例えば、上のガウス分布の例だと、

$$F(\theta) = \frac{1}{\sigma^2} \quad (3.5)$$

となる。このとき、 σ が大きいと、 $F(\theta)$ が小さくなり、 θ の推定が難しくなることがわかる。

Cramér-Rao の不等式は、正確な推定の限界を示すものである。すなわち、不偏推定量を変えることで推定の精度を向上させたいときに、どこまで精度を向上させることができるかを示すものである。

ex.

hoge(計算例を書く、白石の計算はガウス分布を計算すればどうとでもなる。)