

3. Concentration of measure

Yoji Tomita

May 12, 2021

Introduction

- 2章を前提として, この章では tail bound や concentration inequalities を求めるためのより上級の手法を紹介する.
- 3.1 : Concentration by entropic techniques
- 3.2 : A geometric perspective on concentration
- 3.3 : Wasserstein distances and information inequalities
- 3.4 : Tail bounds for empirical processes

3.1 Concentration by entropic techniques

- エントロピーと, 集中不等式導出のためのその関連テクニックに関する議論から始める.

3.1.1 Entropy and its properties

- 凸関数 $\phi: \mathbb{R} \rightarrow \mathbb{R}$ と, 確率変数 $X \sim \mathbb{P}$ に対して, ϕ -entropy を

$$\mathbb{H}_\phi(X) := \mathbb{E}[\phi(X)] - \phi(\mathbb{E}[X])$$

とする ($X, \phi(X)$ の有限期待値は仮定).

- Jensen の不等式より, ϕ -entropy は非負.
- これは X のばらつき加減を表す.
 - ▶ 極端な場合, X が a.s. で期待値と一致するなら, $\mathbb{H}_\phi(X) = 0$.

- 例 1 : $\phi(u) = u^2$ なら $\mathbb{H}_\phi(X)$ は分散.

$$\mathbb{H}_\phi(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \text{var}(X).$$

- 例 2 : $\phi(u) = -\log u$, $Z := e^{\lambda X}$ とすると,

$$\mathbb{H}_\phi(e^{\lambda X}) = -\lambda \mathbb{E}[X] + \log \mathbb{E}[e^{\lambda X}] = \log \mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}]$$

となり, centered cumulant generating function となる.

- この章では, 次の凸関数 $\phi : [0, \infty) \rightarrow \mathbb{R}$ に対する entropy を考える.

$$\phi(u) := \begin{cases} u \log u & \text{for } u > 0 \\ 0 & \text{for } u = 0 \end{cases}. \quad (3.1)$$

非負確率変数 Z に対して, ϕ -entropy は

$$\mathbb{H}(Z) = \mathbb{E}[Z \log Z] - \mathbb{E}[Z] \log \mathbb{E}[Z], \quad (3.2)$$

となる (ただし関連する期待値の存在は仮定).

- ▶ Shannon entropy や Kullback-Leibler divergence と関連がある (see Exercise 3.1).
 - ▶ 以後この entropy を考えるので, \mathbb{H}_ϕ の subscript ϕ は省略.
- $Z = e^{\lambda X}$ とすると, $\mathbb{H}(e^{\lambda X})$ は X のモーメント母関数 $\varphi_X(\lambda) = \mathbb{E}[e^{\lambda X}]$ とその導関数 $\phi'_X(\lambda)$ で表せる.

$$\mathbb{H}(e^{\lambda X}) = \lambda \phi'_X(\lambda) - \varphi_X(\lambda) \log \varphi_X(\lambda). \quad (3.3)$$

Example 3.1 (Entropy of a Gaussian random variable)

- X は 1 次元正規分布 $X \sim \mathcal{N}(0, \sigma^2)$ とすると, $\varphi_X(\lambda) = e^{\lambda^2 \sigma^2 / 2}$, $\varphi'_X(\lambda) = \lambda \sigma^2 \varphi_X(\lambda)$ なので,

$$\mathbb{H}(E^{\lambda X}) = \lambda^2 \sigma^2 \varphi_X(\lambda) - \frac{1}{2} \lambda^2 \sigma^2 \varphi_X(\lambda) = \frac{1}{2} \lambda^2 \sigma^2 \varphi_X(\lambda). \quad (3.4)$$

- この節の残りで, このエントロピー (3.3) と tail bounds との関連性を説明していく.