

5. Coupling lemma

来嶋 秀治

滋賀大学 データサイエンス学部

COUPLING法

Total variation distance (総変動距離)

prob.

- μ と ν は Ω 上の分布とする. **Total variation distance**を

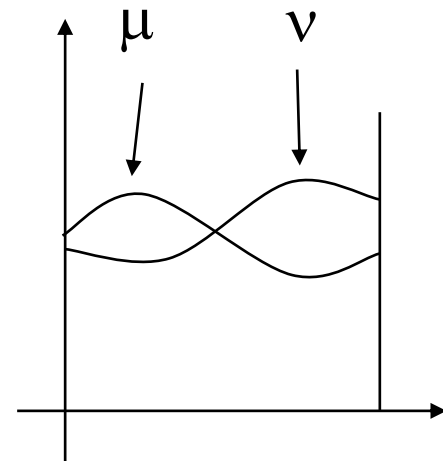
$$d_{\text{TV}}(\mu, \nu) = \frac{1}{2} \sum_{s \in \Omega} |\mu(s) - \nu(s)|$$

と定める.

- 明らかに $d_{\text{TV}}(\mu, \nu) = 0 \Leftrightarrow \mu \equiv \nu$.
- また

$$d_{\text{TV}}(\mu, \nu) = \max \sum_{s \in \Omega} (\mu(s) - \nu(s)) = \sum_{s \in S} (\mu(s) - \nu(s))$$

が成り立つ. ただし $S = \{s \in \Omega \mid \mu(s) - \nu(s) \geq 0\}$.



定理2を示すには $d_{\text{TV}}(\pi, xP^\infty) = 0$ を示せばよい.

補題5.1. (Coupling lemma)

μ, ν は共に Ω 上の分布とする. $X \sim \mu$ と $Y \sim \nu$ は

$$d_{\text{TV}}(\mu, \nu) \leq \Pr(X \neq Y)$$

を満たす.

証明(coupling lemma)

$S = \{s \in \Omega \mid \mu(s) - \nu(s) \geq 0\}$ とする. $s \in S$ について,

$$\begin{aligned} \Pr(X = s, Y \neq s) &= \Pr(X = s) - \Pr(X = s, Y = s) \\ &\geq \Pr(X = s) - \Pr(Y = s) \\ &= \mu(s) - \nu(s) \end{aligned}$$

が成り立つ. よって

$$\begin{aligned} \Pr(X \neq Y) &= \sum_{s \in \Omega} \Pr(X = s, Y \neq s) \geq \sum_{s \in S} \Pr(X = s, Y \neq s) \\ &\geq \sum_{s \in S} (\mu(s) - \nu(s)) = d_{\text{TV}}(\mu, \nu) \end{aligned}$$

Lemma 1 (Coupling Lemma)

							X^t
		0.12	0.17	0.2	0.23	0.26	μ_t
     Y^t	v_t	0.21					
	0.25						
	0.2						
	0.12						
	0.22						

定義: Coupling

μ, ν は共に Ω 上の分布とする. $\Omega \times \Omega$ 上の分布 ω が μ, ν を周辺分布としてもつ (すなわち $\forall x \in \Omega, \mu(x) = \sum_{y \in \Omega} \omega(x, y)$ かつ $\forall y \in \Omega, \nu(y) = \sum_{x \in \Omega} \omega(x, y)$) とき, ω を μ と ν の coupling という.

補題5.2. (Maximal coupling)

Ω 上の分布 μ, ν に対し, ある coupling ω が存在して, $(X, Y) \sim \omega$ は

$$d_{TV}(\mu, \nu) = \Pr(X \neq Y)$$

を満たす.

略証明 (maximal coupling)

任意の $x \in \Omega$ について $\omega(x, x) = \min\{\mu(x), \nu(x)\}$ とする.

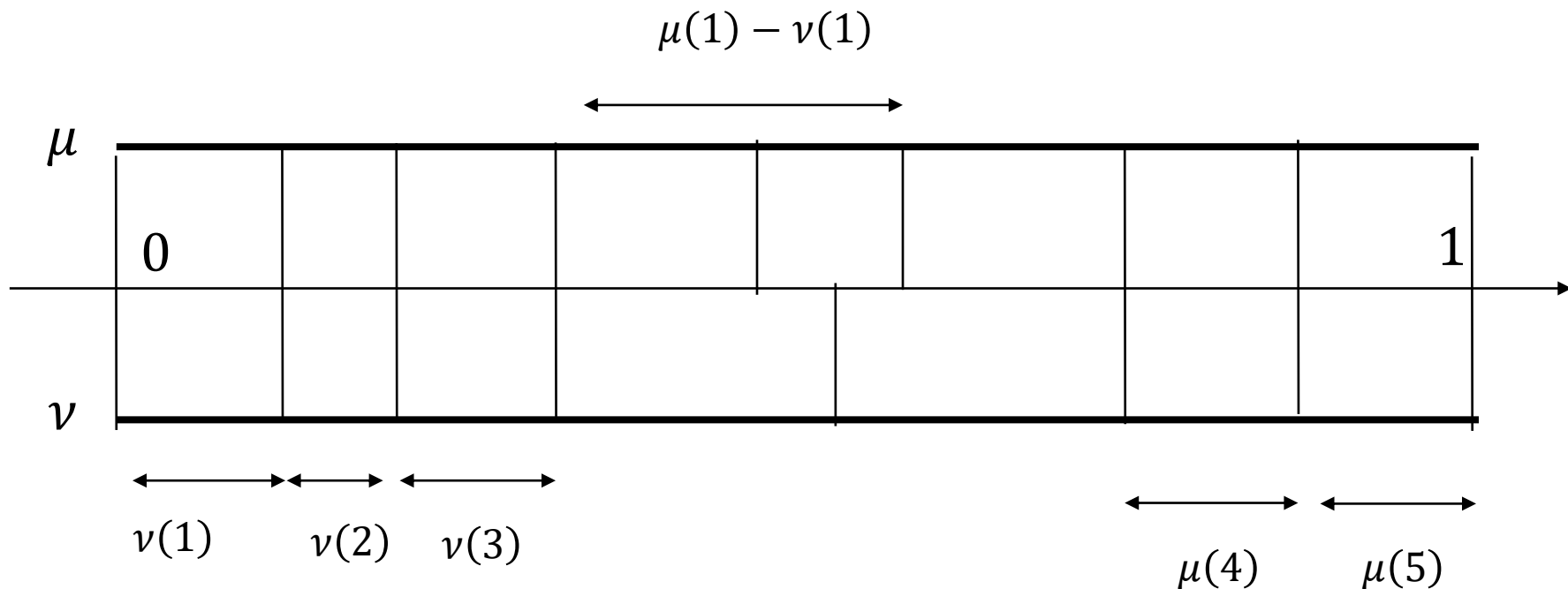
このとき, $x \neq y$ に対しても適切に $\omega(x, y)$ を割り当てること
 (「割り当て問題」 or 「最大マッチング」 参照)

ω を coupling にすることができる. $S = \{x \in \Omega \mid \mu(x) - \nu(x) \geq 0\}$ とすると, あきらかに $\Pr(X \neq Y) = \sum_{s \in S} (\mu(s) - \nu(s))$

Maximal coupling

$$S = \{s \in \Omega \mid \mu(s) \geq v(s)\}$$

$$\bar{S} = \{s \in \Omega \mid \mu(s) < v(s)\}$$



$$X = \psi_\mu(u) \sim \mu, Y = \psi_\nu(u) \sim \nu$$

$$\Pr(X \neq Y) = \sum_{s \in S} (\mu(s) - \nu(s))$$

$$\text{i.e., } \Pr(X = Y) = \sum_{s \in \Omega} \min\{\mu(s), \nu(s)\}$$

補題5.3. (coupling timeの劣乗法性(sub-multiplicativity))

$X_t \sim \mu P^t$, $Y_t \sim \nu P^t$ とし, coupling系列 $\{W_t\} = \{(X_t, Y_t)\}$ は

- i. $\{W_t\}$ は斉時マルコフ
- ii. $[X_t = Y_t] \Rightarrow [X_{t+1} = Y_{t+1}]$

をともに満たす. coupling系列 $\{W'_t\} = \{(X'_t, Y'_t)\}$ は $\{W_t\}$ と同じ遷移方法で, 初期分布が $\max_{x,y} \Pr(X'_c \neq Y'_c | X'_0 = x, Y'_0 = y) = \alpha$ を達成するものとする.

このとき, $\Pr(X_{kc} \neq Y_{kc}) \leq \alpha^k \quad k = 1, 2, \dots$

証明: 帰納法.

便宜のため, $A: (X'_0, Y'_0) = \arg \max_{x,y} \Pr(X'_c \neq Y'_c | X'_0 = x, Y'_0 = y)$ とする.

$k = 1$ の場合は $\Pr(X_c \neq Y_c) \leq \Pr(X'_c \neq Y'_c | A) = \alpha$ でOK.

k の場合の成立を仮定して、 $k + 1$ の場合を示す.

$$\begin{aligned}
 \Pr(X_{(k+1)c} \neq Y_{(k+1)c}) &= \Pr(X_{(k+1)c} \neq Y_{(k+1)c}, X_{kc} \neq Y_{kc}) && \text{by cond. ii} \\
 &= \Pr(X_{(k+1)c} \neq Y_{(k+1)c} \mid X_{kc} \neq Y_{kc}) \Pr(X_{kc} \neq Y_{kc}) \\
 &\leq \Pr(X_{(k+1)c} \neq Y_{(k+1)c} \mid X_{kc} \neq Y_{kc}) \alpha^k \\
 &= \Pr(X_c \neq Y_c \mid X_0 \neq Y_0) \alpha^k && \text{by cond. i} \\
 &\leq \Pr(X'_c \neq Y'_c \mid A) \alpha^k \\
 &= \alpha^{k+1}
 \end{aligned}$$

定理5.2.の証明

- $X_0, X_1, X_2, \dots$ は初期分布 μ に従うものとする. $X_t \sim \mu P^t$.
- $Y_0, Y_1, Y_2, \dots$ は初期分布 π に従うものとする. $Y_t \sim \pi$.
- X と Y は独立とし, $T = \min\{t \mid X_t = Y_t\}$ とすると, $\lim_{t \rightarrow \infty} \Pr(T > t) = 0$ を示す.
 - ✓ P はエルゴード的なので, ある c が存在して $\forall i, j, (P^c)_{ij} > 0$.
 $\alpha = \min\{(P^c)_{ij} \mid i \in \{1, \dots, k\}\}$ とする. この時 $\alpha > 0$ に注意.
 - ✓ $\Pr(X_c = s) = (\mu P^c)_s \geq \alpha, \Pr(Y_c = s) = (\pi P^c)_s \geq \alpha$ より
 $\Pr(T \leq M) \geq \Pr(X_c = Y_c) \geq \Pr(X_c = s, Y_c = s) = \Pr(X_c = s) \Pr(Y_c = s) \geq \alpha^2$
 - ✓ すなわち, $\Pr(T > c) \leq 1 - \alpha^2$
 - ✓ Sub-multiplicativityより $\Pr(T > kc) \leq (1 - \alpha^2)^k$.
 - ✓ よって $\Pr(T > \infty) \leq \lim_{k \rightarrow \infty} (1 - \alpha^2)^k = 0$ となり, 所望の主張を得る.
- いま $X'_0, X'_1, X'_2, \dots = X_0, X_1, \dots, X_T, Y_{T+1}, Y_{T+2}, \dots$ とする.
 - この時, $X'_t \sim x P^t$.
 - また, $[X'_t \neq Y_t] \Leftrightarrow [X_t \neq Y_t]$.
- Coupling lemmaより $d_{TV}(\mu P^t, \pi) \leq \Pr(X'_t \neq Y_t) = \Pr(X_t \neq Y_t) = \Pr(T > t)$
- 上記の議論から $d_{TV}(x P^\infty, \pi) \leq \lim_{t \rightarrow \infty} \Pr(T > t) = 0$.

MIXING TIME

Mixing time

定義

推移確率行列 P の **mixing time** を

$$\tau(\epsilon) = \min\{t \mid \forall t' \geq t, \forall \mu \in \mathcal{D}, d_{\text{TV}}(\mu P^{t'}, \pi) \leq \epsilon\}$$

と定める. ただし $\mathcal{D} = \{\mathbf{x} = (x_1, \dots, x_n) \mid \forall i, x_i \geq 0, \|\mathbf{x}\|_1 = 1\}$

Coupling法

定理 5.4.

$X_0, X_1, \dots$ と $Y_0, Y_1, \dots$ の期待カップリング時間を $E[T]$ とすると

$$\tau(\epsilon) \leq e \lceil \log \epsilon^{-1} \rceil E[T]$$

証明

- $\Pr(X_t \neq Y_t) = \Pr(T > t)$ (1)
- マルコフの不等式より $\Pr(T > eE[T]) \leq \frac{E[T]}{eE[T]} = \frac{1}{e}$
- 劣乗法性より $\Pr(T > keE[T]) \leq \frac{1}{e^k}$
- $k = \lceil \log \epsilon^{-1} \rceil$ を代入すると

$$\Pr(T > e \lceil \log \epsilon^{-1} \rceil E[T]) \leq \frac{1}{e^{\lceil \log \epsilon^{-1} \rceil}} \leq \frac{1}{e^{\log \epsilon^{-1}}} = \epsilon \quad (2)$$

- Coupling lemmaより

$$\begin{aligned} d_{\text{TV}} \left(P_x^{e \lceil \log \epsilon^{-1} \rceil E[T]}, \pi \right) &\leq \Pr \left(X_{e \lceil \log \epsilon^{-1} \rceil E[T]} \neq Y_{e \lceil \log \epsilon^{-1} \rceil E[T]} \right) \\ &= \Pr(T > e \lceil \log \epsilon^{-1} \rceil E[T]) \quad \text{by (1)} \\ &\leq \epsilon \end{aligned}$$

Mixing timeの例

推移確率行列 P は $2^d \times 2^d$

$\{0,1\}^d$ 上のlazyマルコフ連鎖

状態空間: $\{0,1\}^d$

推移確率: 現在の状態を $x = (x_1, \dots, x_d)$ とすると

- $i \in \{1, \dots, d\}$ を一様ランダムに選び,
- $b \in \{0,1\}$ を一様ランダムに選んで, x_i の値を b にする.

$d = 100$ で
 $\tau(10^{-10}) \leq 35,000$

命題 5.5.

$\{0,1\}^n$ 上のlazyマルコフ連鎖のmixing timeは $\tau(\epsilon) \leq e \lceil \log \epsilon^{-1} \rceil n(1 + \log n)$

証明

- 自然な更新関数 $\phi(x; i, b)$ を定める.
- coupling系列 $\{W_t\} = \{(X_t, Y_t)\}$ を, **共通**の乱数 $I_t \in \{1, \dots, d\}, B_t \in \{0,1\}$ を用いて $X_{t+1} = \phi(X_t; I_t, B_t), Y_{t+1} = \phi(Y_t; I_t, B_t)$ で定める.
- $(X_0, Y_0) = (x, y)$ からのcoupling timeの期待値を求める.
 - ✓ $I_t = i$ とすると, 任意の $t' \geq t + 1$ において $X_{t'}(i) = Y_{t'}(i)$.
 - ✓ すべての添え字が選ばれる期待時間は高々 $d(1 + \log d)$ (クーポン収集)
- Coupling法により $\tau(\epsilon) \leq e \lceil \log \epsilon^{-1} \rceil d(1 + \log d)$

EX: 0-1HYPERCUBE

マルコフ連鎖

状態空間 $\Omega = \{0,1\}^n$

$$\text{推移確率 } P(x, y) = \begin{cases} \frac{1}{2} & x = y \\ \frac{1}{2n} & |x \oplus y| = 1 \\ 0 & \text{それ以外} \end{cases}$$

定理 5.6.

$$\tau \leq n(1 + \log n)$$

証明

$|x \oplus y| = k$ から $|x \oplus y| = k - 1$ の時間を T_k とする.

$$E[T_k] = \frac{n}{k}$$

$$\text{よって } E[T] = \sum_{k=1}^n E[T_k] = nH_n \leq n(1 + \log n)$$

RANDOM COLORING

目標分布

標本空間： $\Omega = \{x \in \{1, \dots, q\}^V \mid x \text{は適切な彩色}\}$

確率関数： $f(x) = \frac{1}{c} \quad x \in \Omega$ (一様分布, $c = |\Omega|$)

マルコフ連鎖 (メトロポリス法)

状態空間： Ω

状態遷移：

1. $v \in V$ を一様ランダムに選ぶ
2. $k \in \{1, \dots, q\}$ を一様ランダムに選ぶ.
3. v の色を k として適切ならその状態に遷移する.
不適切なら元の状態を維持する.

定理 5.7.

任意のグラフに対して、 $q > \Delta + 1$ なら極限分布は一様分布

定理 5.8.

G の最大次数を Δ とする. $q \geq 4\Delta + 1$ のとき

$$\tau_{\text{mix}}(\epsilon) \leq \lceil qn(\log n + \log \epsilon^{-1}) \rceil = O(n \log n)$$

証明 1/2

彩色 $\{X_t\}$ と彩色 $\{Y_t\}$ をcouplingする. $d_t = |\{v | X_t(v) \neq Y_t(v)\}|$

□ 全体coupling (共通乱数法)

- $(v, k) \in V \times \chi$ を一様ランダムに選んで推移させる.
- 適切なら遷移, 不適切なら現状維持

□ 距離が減る確率 = $d_t \frac{q-2\Delta}{qn}$:

- $X_t(v) \neq Y_t(v)$ (確率 $\frac{d_t}{n}$) かつ $k \in \chi(X_t, v) \cap \chi(Y_t, v)$
- $|\chi(X_t, v)^c| \leq \Delta, |\chi(Y_t, v)^c| \leq \Delta$ より $|\chi(X_t, v)^c \cap \chi(Y_t, v)^c| \leq 2\Delta$ (union bound)
- 従って $|\chi(X_t, v) \cap \chi(Y_t, v)| = q - |\chi(X_t, v)^c \cap \chi(Y_t, v)^c| \geq 2\Delta$

□ 距離が増える確率 = $d_t \frac{2\Delta}{qn}$

- もともと一致していたのに, $k \in \chi(X_t, v) \oplus \chi(Y_t, v)$ に置き換えられた時
 - $|\chi(X_t, v) \oplus \chi(Y_t, v)| = 2$. $X_t(v) \neq Y_t(v)$ の近傍は高々 Δ .

□ 1回の遷移後の期待距離

$$\mathbb{E}[d_{t+1} | X_t, Y_t] \leq d_t - d_t \frac{q-2\Delta}{qn} + d_t \frac{2\Delta}{qn} = d_t \left(1 - \frac{q-4\Delta}{qn} \right)$$

- $\chi = \{1, \dots, q\}$: 色の集合
- $\chi(x, v) = \{c \in \chi \mid \forall u \in N(v), c \neq x(u)\}$
: 彩色 $x \in \Omega$ の頂点 $v \in V$ に塗れる色.

証明 2/2

□ (承前) 1回の遷移後の期待距離

$$\mathbb{E}[d_{t+1} | X_t, Y_t] \leq d_t \left(1 - \frac{q - 4\Delta}{qn} \right)$$

□ 再帰的に

$$\mathbb{E}[d_t | X_0, Y_0] \leq d_0 \left(1 - \frac{q - 4\Delta}{qn} \right)^t$$

$$(1 - x)^t \leq \exp(-xt)$$

□ $d_0 \leq n, q \geq 4\Delta + 1$ より

$$\mathbb{E}[d_t | X_0, Y_0] \leq d_0 \left(1 - \frac{q - 4\Delta}{qn} \right)^t \leq n \left(1 - \frac{1}{qn} \right)^t \leq n \exp \left(-\frac{t}{qn} \right)$$

□ カップリング(しない)確率は

$$\begin{aligned} \Pr[X_t \neq Y_t | X_0, Y_0] &= \Pr[d_t \geq 1 | X_0, Y_0] \\ &\leq \mathbb{E}[d_t | X_0, Y_0] \quad \text{マルコフの不等式} \\ &\leq n \exp \left(-\frac{t}{qn} \right) \end{aligned}$$

□ $t \geq qn(\ln n + \ln \epsilon^{-1})$ を代入すると

$$n \exp \left(-\frac{t}{qn} \right) \leq n \exp(-\ln n - \ln \epsilon^{-1}) = \epsilon$$

未解決問題

- $q \geq \Delta + 1$ の効率的なサンプリング法
- $q \geq \Delta + 2$ のGlauber dynamicsが $\tau_{mix} = O(n \log n)$

既知の事実

- $q \geq \Delta + 2$ ならエルゴードマルコフ連鎖
- $q \geq 2\Delta + 1$ から $q \geq (1 + \Omega(\log n))\Delta$ に改善
- $q = \Delta + 3$ が $\tau_{mix} = O(n \log n)$ [Chen et al. 2023]

Z. Chen, K. Liu, N. Mani, and A. Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. ArXiv preprint 2304.01954. 2023.



おわり
