知っているとちょっと得する（？）確率順序とその応用

三好 直人

東京工業大学 大学院情報理工学研究科

確率順序って？

確率変数や確率過程，またそれらの分布の間に（何らかの）大小関係を定める

Ex.) 平均やモーメントの大きさ，ばらつきの大きさ，依存性の大きさ

どんなときに使うの？

- 性能比較，最適化（ex. どんな処理の仕方が好ましいか？）
- 評価量の上界・下界
- 漸近解析（分布の裾の評価）
  - $X_{-\epsilon} \leq_{st} X \leq_{st} X_{+\epsilon}$
    $$\iff \quad \mathrm{P}(X_{-\epsilon} > x) \leq \mathrm{P}(X > x) \leq \mathrm{P}(X_{+\epsilon} > x) \quad \forall x \in \mathbb{R}$$
  - $\mathrm{P}(X_{-\epsilon} > x) \sim a_{-\epsilon}(x)$, $\mathrm{P}(X_{+\epsilon} > x) \sim a_{+\epsilon}(x) \quad x \to \infty$
  - $a_{-\epsilon}(x) \to a(x)$, $a_{+\epsilon}(x) \to a(x) \quad \epsilon \to 0$
    $$\Longrightarrow \quad \boxed{\mathrm{P}(X > x) \sim a(x) \quad x \to \infty}$$

確率順序の長所・短所

長所

- 使う分にはそれほど難しくない
- 定性的な性質については，かなり一般的な結果が得られる
- 知っていると ちょっと得する (··· 時がもしかしたら来るかもしれない)

短所

- まじめに勉強しようとすると，それなりに難しい
- 定量的なことは分からない
- 知らなくても別に困らない (··· かもしれない)

本講演で扱う確率順序

確率変数とその分布の間の確率順序

平均・モーメント・ばらつきの大きさを比較する

1. 通常の確率順序
2. 凸順序と増加凸順序

確率ベクトル・確率過程とその分布の間の確率順序

（上記に加えて）依存性の強さを比較する

1. 優モジュラ順序と方位性凸順序

通常の確率順序 $(\leq_{st})$

X, Y: 実数値確率変数. それぞれの分布関数 F_X, F_Y

$X \leq_{st} Y$, $F_X \leq_{st} F_Y$

X（または F_X）が 通常の確率順序に関して [or 確率的に] Y（または F_Y）より小さい

$\iff$ **(i)** $\mathrm{P}(X > t) \leq \mathrm{P}(Y > t) \quad \forall t \in \mathbb{R}$

$\iff$ **(ii)** $F_X(t) \geq F_Y(t) \quad \forall t \in \mathbb{R}$

$\iff$ **(iii)** それぞれが分布関数 F_X, F_Y に従い, $\mathrm{P}(\hat{X} \leq \hat{Y}) = 1$ が成り立つような確率変数 $\hat{X}, \hat{Y}$ を定義できる

$\iff$ **(iv)** $\mathrm{E}[f(X)] \leq \mathrm{E}[f(Y)] \quad \forall f \in \{$ 非減少関数 $\} \cap \{$ 期待値が存在 $\}$

同値性の証明

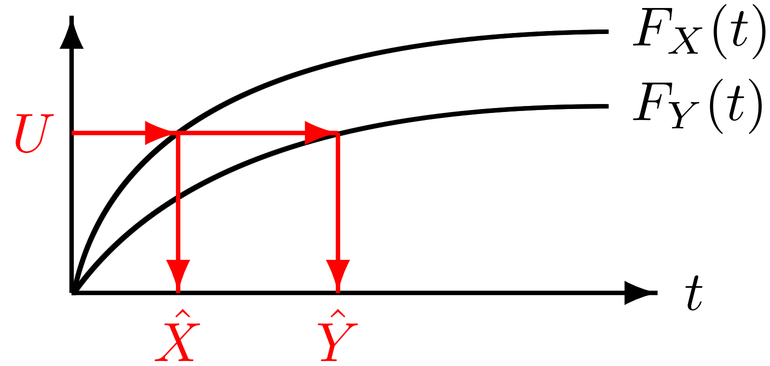
(i)⇔(ii) 明らか

(ii)⇒(iii) $F_X^{-1}(u) := \inf\{t \in \mathbb{R} : F_X(t) \geq u\},\ u \in (0,1)$ (F_Y^{-1} も同様)

$$F_X(t) \geq F_Y(t),\ \forall t \in \mathbb{R} \iff F_X^{-1}(u) \leq F_Y^{-1}(u),\ \forall u \in (0,1)$$

U: $(0,1)$ 上の一様分布に従う確率変数 $\Longrightarrow$ $\hat{X} = F_X^{-1}(U)$, $\hat{Y} = F_Y^{-1}(U)$ は

それぞれ F_X, F_Y に従い $\hat{X} \leq \hat{Y}$ w.p.1



(iii)⇒(iv) $\hat{X} \leq \hat{Y}$ w.p.1 $\Longrightarrow f(\hat{X}) \leq f(\hat{Y})$ w.p.1 $\Longrightarrow \mathrm{E}[f(X)] \leq \mathrm{E}[f(Y)]$

(iv)⇒(i) $f(X) = 1\{X > t\}$ $(t \in \mathbb{R})$ とする

$\Longrightarrow f(x)$ は非減少で $\mathrm{E}[f(X)] = \mathrm{P}(X > t)$ (Y についても同様)

$\leq_{st}$ の性質

(i) $X \leq_{st} Y \Longrightarrow \mathrm{E}[X^n] \leq \mathrm{E}[Y^n] \ \ \forall n \in \{$ 奇数 $\}$

X, Y が非負のとき $\mathrm{E}[X^t] \leq \mathrm{E}[Y^t] \ \ \forall t \geq 0$

(ii) $X \leq_{st} Y \Longrightarrow \mathrm{E}[e^{tX}] \leq \mathrm{E}[e^{tY}] \ \ \forall t \geq 0$

(iii) $X \leq_{st} Y$ かつ $\mathrm{E}(X) = \mathrm{E}(Y) \Longrightarrow X =_{st} Y$ $(F_X = F_Y)$

(iv) $X \leq_{st} Y \Longrightarrow f(X) \leq_{st} f(Y) \ \ \forall f \in \{$ 非減少関数 $\}$

(v) $(X_1, \ldots, X_n)$ と $(Y_1, \ldots, Y_n)$: ともに独立な確率変数の組

$X_i \leq_{st} Y_i, \ \ i = 1, \ldots, n$

$\Longrightarrow \phi(X_1, \ldots, X_n) \leq_{st} \phi(Y_1, \ldots, Y_n)$ $\ \forall \phi$: $\mathbb{R}^n \to \mathbb{R} \in \{$ 非減少関数 $\}$

$\le_{st}$ の応用例: カップリングによる性能比較

広い意味でのカップリング (確率変数 [または確率過程] X, Y のカップリング)

$(\hat{X}, \hat{Y})$: $\hat{X} =_{st} X$, $\hat{Y} =_{st} Y$ を満たし，同じ確率空間上で定義された $\hat{X}$, $\hat{Y}$ の組

狭い意味でのカップリング (確率過程 $X(t)$, $Y(t)$ のカップリング)

ある確率変数 T が存在して，$X(t) = Y(t)$ w.p.1 ($\forall t \ge T$) となること

($X(t)$, $Y(t)$ は広い意味でのカップリング)

サンプルパス・カップリングによる性能比較

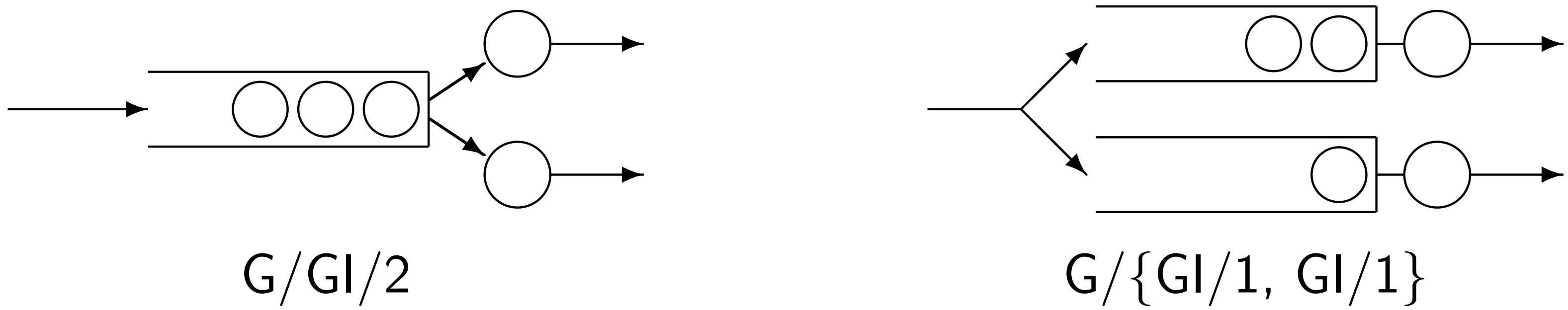
2 つの確率過程 $X(t)$, $Y(t)$

(i) $X(t)$, $Y(t)$ の (広い意味での) カップリング $(\hat{X}(t), \hat{Y}(t))$ を考える

(ii) サンプルパス上で，どんな場合でも $\hat{X}(t) \le \hat{Y}(t) \Longrightarrow X(t) \le_{st} Y(t)$

カップリングによる性能比較の応用例

G/GI/2 vs. G/{GI/1, GI/1}



システム X: G/GI/2: 先着順サービス

システム Y: G/{GI/1, GI/1}: 到着客は短い行列へ．各行列では先着順サービス

T_n^X, T_n^Y $(n = 1, 2, \ldots)$ 到着時刻の列

S_n^X, S_n^Y $(n = 1, 2, \ldots)$ サービス時間の列

$Q^X(t)$, $Q^Y(t)$ $(t \geq 0)$ システム内客数

[定理] $\{T_n^X\} =_{st} \{T_n^Y\}$, $S_n^X =_{st} S_n^Y$, $n = 1, 2, \ldots$, $Q^X(0) = Q^Y(0) = 0$

$\Longrightarrow Q^X(t) \leq_{st} Q^Y(t)$, $\forall t \geq 0$

G/GI/2 vs. G/{GI/1, GI/1}　(証明の概略)

Point: 2 つのサンプルパスを平行して作っていく(広い意味でのカップリング)

(i) $T_n^X = T_n^Y = T_n$ $(n = 1, 2, \ldots)$ (到着時刻をそろえる)

(ii) $\tilde{S}_n^X$, $\tilde{S}_n^Y$ $(n = 1, 2, \ldots)$ サービス時間の列をサービス開始の順に並べ替え
$\tilde{S}_n^X = \tilde{S}_n^Y = S_n$ $(n = 1, 2, \ldots)$ (サービス時間は i.i.d. で到着時刻と独立)

(iii) U_n^X, U_n^Y $(n = 1, 2, \ldots)$ サービス開始時刻の列

$$0 \le U_1^X \le U_2^X \le \cdots, \qquad 0 \le U_1^Y \le U_2^Y \le \cdots$$

(iv) D_n^X, D_n^Y $(n = 1, 2, \ldots)$ 退去時刻の列

$$0 \le D_1^X \le D_2^X \le \cdots, \qquad 0 \le D_1^Y \le D_2^Y \le \cdots$$

$$D_k^X = \min\Big(\{U_n^X + S_n,\ n = 1, \ldots, \underline{k+1}\} \setminus \{D_1^X, \ldots, D_{k-1}^X\}\Big)$$

(D_k^Y も同様)　　$\uparrow$ ($k+1$ 番目のサービス開始までを見れば十分)

(v) $U_n^X \le U_n^Y$ w.p.1 $(n = 1, 2, \ldots)$ を示せばよい

$\Longrightarrow D_n^X \le D_n^Y$ w.p.1 $\Longrightarrow Q^X(t) \le Q^Y(t)$ $(t \ge 0)$ w.p.1 (到着時刻は同じ)

G/GI/2 vs. G/{GI/1, GI/1} (証明のつづき)

$U_n^X \le U_n^Y$ w.p.1 $(n = 1, 2, \ldots)$ を示す

(i) 明らかに $U_1^X = U_1^Y = T_1$, $U_2^X = U_2^Y = T_2$ (到着 = サービス開始)

(ii) $n \ge 3$ に対して $\boxed{U_n^X = \max\{T_n, D_{n-2}^X\}}$, $\boxed{U_n^Y \ge \max\{T_n, D_{n-2}^Y\}}$

割り込みのないすべての 2 サーバ待ち行列で成立

(最大 2 人までしか同時にサービスされない)

(iii) $U_n^X \le U_n^Y$ w.p.1, $n = 1, \ldots, m \ (\ge 2)$ を仮定 (帰納法)

$$
\begin{aligned}
D_{m-1}^X &= \min\Big(\{U_n^X + S_n,\ n = 1, \ldots, m\} \setminus \{D_1^X, \ldots, D_{m-2}^X\}\Big) \\
&\le \min\Big(\{U_n^Y + S_n,\ n = 1, \ldots, m\} \setminus \{D_1^Y, \ldots, D_{m-2}^Y\}\Big) \\
&= D_{m-1}^Y \quad \text{w.p.1}
\end{aligned}
$$

(iv) $\Longrightarrow U_{m+1}^X = \max\{T_{m+1}, D_{m-1}^X\} \le \max\{T_{m+1}, D_{m-1}^Y\} \le U_{m+1}^Y$

サンプルパス・カップリングの注意点

- サンプルパス上で起こり得るすべての場合を考えなければならない
  場合分けの嵐になることも!!
- 最低限度の仮定から始めて，必要であれば仮定を加えていく
  - 確率変数列の並べ替え $\longrightarrow$ 独立同一分布
  - 確率変数の無記憶性 $\longrightarrow$ 指数分布

$\Longrightarrow$ どんな条件が本質的であるかが分かる!?

凸順序 ($\leq_{cx}$) と増加凸順序 ($\leq_{icx}$)

変動（ばらつき）の大きさを比較する

X, Y: 実数値確率変数，分布関数 F_X, F_Y

$X \leq_{cx} Y$， $F_X \leq_{cx} F_Y$

X（または F_X）が 凸順序に関して Y（または F_Y）より小さい

$\iff$ $\mathrm{E}[f(X)] \leq \mathrm{E}[f(Y)]$ $\forall f \in \{\underline{凸関数}\} \cap \{$ 期待値が存在 $\}$

$X \leq_{icx} Y$， $F_X \leq_{icx} F_Y$

X（または F_X）が 増加凸順序に関して Y（または F_Y）より小さい

$\iff$ $\mathrm{E}[f(X)] \leq \mathrm{E}[f(Y)]$ $\forall f \in \{\underline{非減少凸関数}\} \cap \{$ 期待値が存在 $\}$

$\le_{st}$，$\le_{cx}$，$\le_{icx}$ の関係と性質

(i) $X \le_{st} Y \Longrightarrow X \le_{icx} Y$ ({ 非減少関数 } ⊃ { 非減少凸関数 })

(ii) $X \le_{cx} Y \Longrightarrow X \le_{icx} Y$ （$X \le_{cx} Y \Longleftrightarrow X \le_{icx} Y$ かつ $\mathrm{E}[X] = \mathrm{E}[Y]$）

(iii) $X \le_{cx} Y \Longrightarrow \mathrm{E}[X^n] \le \mathrm{E}[Y^n]\ \ \forall n \in$ { 偶数 } とくに $\boxed{\mathrm{Var}[X] \le \mathrm{Var}[Y]}$

(iv) $X \le_{icx} Y$ かつ X, Y が非負 $\Longrightarrow \mathrm{E}[X^t] \le \mathrm{E}[Y^t]\ \ \forall t \ge 1$

(v) $X \le_{cx} Y \Longrightarrow \mathrm{E}[e^{tX}] \le \mathrm{E}[e^{tY}]\ \ \forall t \in \mathbb{R}$

(vi) $X \le_{icx} Y \Longrightarrow \mathrm{E}[e^{tX}] \le \mathrm{E}[e^{tY}]\ \ \forall t \ge 0$

(vii) $\boxed{X \le_{cx} Y \text{ かつ } \mathrm{Var}[X] = \mathrm{Var}[Y] \Longrightarrow X =_{st} Y}$

$\le_{st}$, $\le_{cx}$, $\le_{icx}$ の関係と性質 (つづき)

(viii) $X \le_{icx} Y \iff \boxed{\mathrm{E}[\max(X-t,0)] \le \mathrm{E}[\max(Y-t,0)]}$ $\forall t \in \mathbb{R}$

- X, Y が非負のとき X, Y の 残余寿命分布 に従う確率変数 X_e, Y_e:

$$\mathrm{P}(X_e > t) = \frac{1}{\mathrm{E}[X]} \int_t^\infty (1 - F_X(s))\,\mathrm{d}s = \frac{\mathrm{E}[\max(X-t,0)]}{\mathrm{E}[X]}, \quad t \ge 0$$

$$\boxed{X \le_{cx} Y \implies X_e \le_{st} Y_e}$$

応用: 定常な M/GI/1 の系内仕事量の単調性

サービス時間の分布関数: $F(t) = \mathrm{P}(X \le t)$

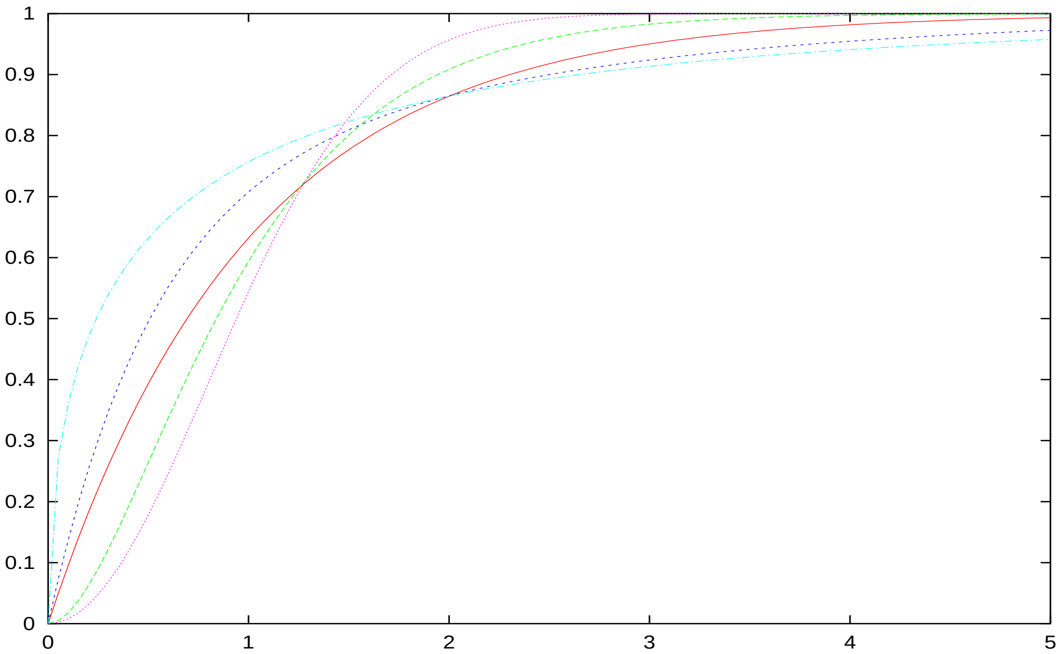
サービス時間の残余寿命の分布関数: $F_e(t) = \mathrm{P}(X_e \le t)$

M/GI/1 の系内仕事量 V: $\boxed{\mathrm{P}(V \le x) = (1-\rho) \sum_{n=0}^{\infty} \rho^n F_e^{*n}(x)}$ Beneš の公式

サービス時間が凸順序の意味で大きい

$\implies$ 系内仕事量が通常の確率順序の意味で大きい ($\le_{st}$ の性質 (v))

$\leq_{cx}$ による確率分布の比較例



平均 1

- 指数分布
- アーラン (2 次)
- 超指数 $(2, 1/2)$
- ワイブル $(r = 2)$
- ワイブル $(r = 1/2)$

ワイブル $(r=2)$ $\leq_{cx}$ アーラン (2 次) $\leq_{cx}$ 指数分布

$\leq_{cx}$ 超指数 $(2,1/2)$ $\leq_{cx}$ ワイブル $(r=1/2)$

X, Y の平均が同じとき

$\exists t_0 \in \mathbb{R},\ \ F_X(t) \leq F_Y(t)\ (t \leq t_0)$ かつ $F_X(t) \geq F_Y(t)\ (t > t_0) \Longrightarrow F_X \leq_{cx} F_Y$

Strassen の定理

(i) $\boxed{X \leq_{cx} Y}$ $\Longleftrightarrow$ それぞれ X, Y と同じ分布に従い $\boxed{\hat{X} = \mathrm{E}[\hat{Y} \mid \hat{X}] \text{ w.p.1}}$ となるような確率変数 $\hat{X}, \hat{Y}$ を定義できる

(ii) $\boxed{X \leq_{icx} Y}$ $\Longleftrightarrow$ それぞれ X, Y と同じ分布に従い $\boxed{\hat{X} \leq \mathrm{E}[\hat{Y} \mid \hat{X}] \text{ w.p.1}}$ が成り立つような確率変数 $\hat{X}, \hat{Y}$ を定義できる

$\Longleftarrow$ の証明 ($\Longrightarrow$ の証明は難しい!)

(i) $\forall f \in \{$ 凸関数 $\}$

$$
\begin{aligned}
\mathrm{E}[f(X)] = \mathrm{E}[f(\hat{X})] &= \mathrm{E}[f(\mathrm{E}[\hat{Y} \mid \hat{X}])] \\
&\leq \mathrm{E}[\mathrm{E}[f(\hat{Y}) \mid \hat{X}]] = \mathrm{E}[f(\hat{Y})] = \mathrm{E}[f(Y)]
\end{aligned}
$$

(ii) $\forall f \in \{$ 非減少凸関数 $\}$

$$
\begin{aligned}
\mathrm{E}[f(X)] = \mathrm{E}[f(\hat{X})] &\leq \mathrm{E}[f(\mathrm{E}[\hat{Y} \mid \hat{X}])] \\
&\leq \mathrm{E}[\mathrm{E}[f(\hat{Y}) \mid \hat{X}]] = \mathrm{E}[f(\hat{Y})] = \mathrm{E}[f(Y)]
\end{aligned}
$$

(Jensen の不等式)

Strassen の定理の応用: GI/GI/1 待ち行列の単調性

待ち行列 X と待ち行列 Y

τ_n^X, τ_n^Y $(n = 0, 1, 2, \ldots)$ 到着時間間隔の列

S_n^X, S_n^Y $(n = 0, 1, 2, \ldots)$ サービス時間の列

W_n^X, W_n^Y $(n = 0, 1, 2, \ldots)$ 待ち時間の列

$$\underline{W_{n+1}^X = \max(W_n^X + S_n^X - \tau_n^X, 0)} \quad (W_n^Y \text{ も同様})$$

[定理] $X_n = S_n^X - \tau_n^X$, $Y_n = S_n^Y - \tau_n^Y$, $W_0^X = W_0^Y = 0$ とする

$$X_n \le_{icx} Y_n \ (n = 0, 1, 2, \ldots) \implies W_n^X \le_{icx} W_n^Y \ (n = 1, 2, \ldots)$$

たとえば . . .

- $\tau_n^X =_{st} \tau_n^Y$, $S_n^X \le_{icx} S_n^Y \implies W_n^X \le_{icx} W_n^Y$
- $\underline{\tau_n^X \le_{cx} \tau_n^Y}$, $S_n^X =_{st} S_n^Y \implies W_n^X \le_{icx} W_n^Y$
  ($f(x)$: 非減少凸 $\implies g(x) = f(-x)$: 非増加凸)

GI/GI/1 待ち行列の単調性の証明

(i) $\underline{X_n \le \mathrm{E}[Y_n \mid X_n] \text{ w.p.1}}$ $(n = 0, 1, 2, \ldots)$ とする (Strassen の定理)

(ii) $\underline{W_n^X \le \mathrm{E}[W_n^Y \mid X_0, \ldots, X_{n-1}] \text{ w.p.1}}$ $(n = 1, 2, \ldots)$ を示す （帰納法）

(ii-a) $n = 1$ のとき

$$
\begin{aligned}
W_1^X &= \max(X_0, 0) \le \max(\underline{\mathrm{E}[Y_0 \mid X_0]}, 0) \\
&\le \mathrm{E}[\max(Y_0, 0) \mid X_0] = \mathrm{E}[W_1^Y \mid X_0] \ \text{ w.p.1}
\end{aligned}
$$

↑ Jensen の不等式

(ii-b) n での成立を仮定

$$
\begin{aligned}
W_{n+1}^X &= \max(W_n^X + X_n, 0) \\
&\le \max(\underline{\mathrm{E}[W_n^Y \mid X_0, \ldots, X_{n-1}]} + \underline{\mathrm{E}[Y_n \mid X_n]}, 0) \\
&\le \mathrm{E}[\max(W_n^Y + Y_n, 0) \mid X_0, \ldots, X_n] = \mathrm{E}[W_{n+1}^Y \mid X_0, \ldots, X_n]
\end{aligned}
$$

(iii) $\forall f \in \{$ 非減少凸関数 $\}$

$$
\begin{aligned}
\mathrm{E}[f(W_n^X)] &\le \mathrm{E}[f(\mathrm{E}[W_n^Y \mid X_0, \ldots, X_{n-1}])] \\
&\le \mathrm{E}[\mathrm{E}[f(W_n^Y) \mid X_0, \ldots, X_{n-1}]] = \mathrm{E}[f(W_n^Y)]
\end{aligned}
$$

(Jensen の不等式)

多次元確率ベクトル・確率過程に対する $\le_{st}$, $\le_{cx}$, $\le_{icx}$

定義は同じ

$X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$

$X \le_{st} Y$ X が 通常の確率順序に関して Y より小さい

$\iff$ $\mathrm{E}[f(X)] \le \mathrm{E}[f(Y)]$ (*)

$\forall f \in \{$ 非減少関数 $\} \cap \{$ 期待値が存在 $\}$

$X \le_{cx} Y$ X が 凸順序に関して Y より小さい

$\iff$ (*) $\forall f \in \{$ 凸関数 $\} \cap \{$ 期待値が存在 $\}$

$X \le_{icx} Y$ X が 増加凸順序に関して Y より小さい

$\iff$ (*) $\forall f \in \{$ 非減少凸関数 $\} \cap \{$ 期待値が存在 $\}$

注意

$n \ge 2$, $\forall (t_1, \ldots, t_n) \in \mathbb{R}^n$

$$X \le_{st} Y \;\underset{\cancel{\Longleftarrow}}{\Longrightarrow}\; \begin{cases} \mathrm{P}(X_1 > t_1, \ldots, X_n > t_n) \le \mathrm{P}(Y_1 > t_1, \ldots, Y_n > t_n) \\ \mathrm{P}(X_1 \le t_1, \ldots, X_n \le t_n) > \mathrm{P}(Y_1 \le t_1, \ldots, Y_n \le t_n) \end{cases}$$

依存性を比較する確率順序

- 優モジュラ順序 (Supermodular Order; $\le_{sm}$)
- 増加優モジュラ順序 (Increasing Supermodular Order; $\le_{ism}$)
- 方位性凸順序 (Directionally Convex Order; $\le_{dcx}$)
- 増加方位性凸順序 (Increasing Directionally Convex Order; $\le_{idcx}$)
- $\cdots$

優モジュラ関数と方位性凸関数

- 関数 $f\colon \mathbb{R}^n$ [または $\mathbb{Z}^n$] $\to \mathbb{R}$ が 優モジュラ (Supermodular)

$$\iff \boxed{f(x)+f(y) \le f(x \wedge y) + f(x \vee y)} \quad \forall x,\ y \in \mathbb{R}^n \text{ [または } \mathbb{Z}^n]$$

ここで $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ に対して

$$x \wedge y = (\min(x_1, y_1), \ldots, \min(x_n, y_n))$$

$$x \vee y = (\max(x_1, y_1), \ldots, \max(x_n, y_n))$$

($n = 1$ のときは任意の関数)

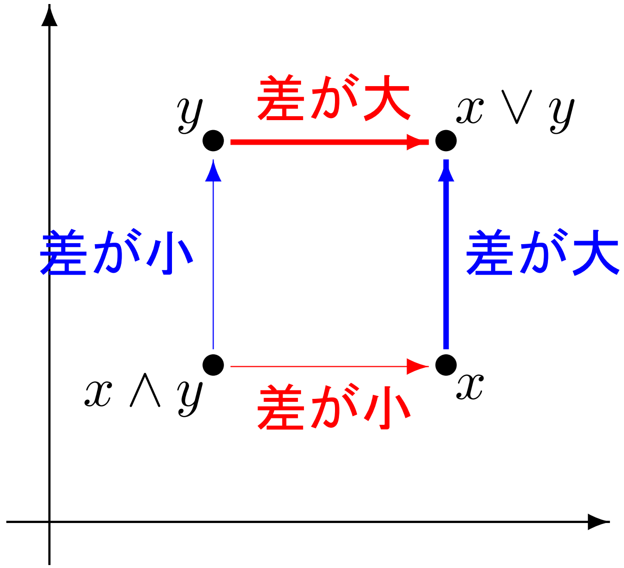
- 関数 $f\colon \mathbb{R}^n$ [または $\mathbb{Z}^n$] $\to \mathbb{R}$ が 方位性凸 (Directionally Convex)

$$\iff \boxed{f(x_1+y) - f(x_1) \le f(x_2+y) - f(x_2)}$$

$$\forall x_1, x_2, y \in \{\mathbb{R}^n \text{[または } \mathbb{Z}^n] \mid \underline{x_1 \le x_2,\ y \ge 0}\}$$

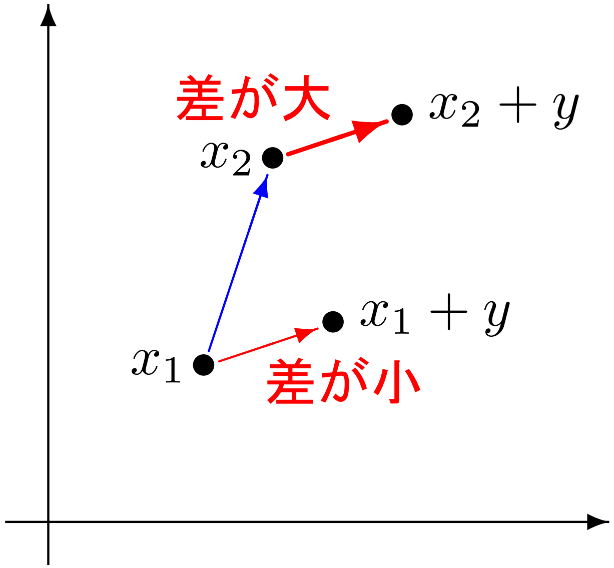
($n = 1$ のときは通常の凸関数)

優モジュラ関数と方位性凸関数（つづき）



$$\color{red}{f(x) - f(x \wedge y) \leq f(x \vee y) - f(y)}$$
$$\color{blue}{f(y) - f(x \wedge y) \leq f(x \vee y) - f(x)}$$

優モジュラ関数



$$\color{red}{f(x_1 + y) - f(x_1) \leq f(x_2 + y) - f(x_2)}$$

方位性凸関数

- $n \geq 2$ のとき，方位性凸 $\not\Longrightarrow$, $\not\Longleftarrow$ （通常の）凸
- 方位性凸 $\Longrightarrow$ 優モジュラ (方位性凸 $\Longleftrightarrow$ 優モジュラ かつ 成分ごとに凸)

実際にチェックするには...

- 関数 $f\colon \mathbb{R}^n$ [または $\mathbb{Z}^n$] $\to \mathbb{R}$ が **優モジュラ**

$$\iff \forall x = (x_1, \ldots, x_n),\ (s,t) \gtreqless (0,0),\ i \neq j\ (i,j = 1,\ldots,n)$$

$$\boxed{\begin{aligned} &f(\ldots, x_i + s, \ldots, x_j + t, \ldots) - f(\ldots, x_i + s, \ldots, x_j, \ldots) \\ &\qquad - f(\ldots, x_i, \ldots, x_j + t, \ldots) + f(\ldots, x_i, \ldots, x_j, \ldots) \geq 0 \end{aligned}}$$

- 関数 $f\colon \mathbb{R}^n$ [または $\mathbb{Z}^n$] $\to \mathbb{R}$ が **方位性凸**

$\iff$ f は 優モジュラ かつ

$$\forall x = (x_1, \ldots, x_n),\ (s,t) \gtreqless (0,0),\ i = 1,\ldots,n$$

$$\boxed{\begin{aligned} &f(\ldots, x_i + s + t, \ldots) - f(\ldots, x_i + s, \ldots) \\ &\qquad - f(\ldots, x_i + t, \ldots) + f(\ldots, x_i, \ldots) \geq 0 \end{aligned}}$$

f が 2 階微分可能の場合

- f が優モジュラ $\iff$ $\boxed{\dfrac{\partial^2 f(x)}{\partial x_i \, \partial x_j} \geq 0}$ $\forall x \in \mathbb{R}^n,\ i \neq j\ (i,j = 1,\ldots,n)$

- f が方位性凸 $\iff$ $\boxed{\dfrac{\partial^2 f(x)}{\partial x_i \, \partial x_j} \geq 0}$ $\forall x \in \mathbb{R}^n,\ i,j = 1,\ldots,n$

優モジュラ関数と方位性凸関数の例

(i) $f(x_1, \ldots, x_n) = \min(x_1, \ldots, x_n) \in \{$ 非減少優モジュラ $\}$

(ii) $\phi : \mathbb{R} \to \mathbb{R} \in \{$非減少凸$\}$

$f : \mathbb{R}^n \to \mathbb{R} \in \{$非減少 [or 非増加] 優モジュラ／方位性凸$\}$

$\Longrightarrow$ $\phi\big(f(x_1, \ldots, x_n)\big) \in \{$ 非減少 [or 非増加] 優モジュラ／方位性凸 $\}$

(iii) $f : \mathbb{R}^n \to \mathbb{R} \in \{$優モジュラ$\}$

$\{\phi_i : \mathbb{R} \to \mathbb{R}, i = 1, \ldots, n\} \in \{$すべて非減少 or すべて非増加$\}$

$\Longrightarrow$ $f\big(\phi_1(x_1), \ldots, \phi_n(x_n)\big) \in \{$ 優モジュラ $\}$

また

$f : \mathbb{R}^n \to \mathbb{R} \in \{$非減少方位性凸$\}$

$\phi_i : \mathbb{R} \to \mathbb{R} \in \{$非減少凸$\}$, $i = 1, \ldots, n$

$\Longrightarrow$ $f\big(\phi_1(x_1), \ldots, \phi_n(x_n)\big) \in \{$ 非減少方位性凸 $\}$

優モジュラ関数と方位性凸関数の例 (つづき)

(iv) $f : \mathbb{R}^n \to \mathbb{R} \in \{\underline{\text{優モジュラ}}\}$

$\{g_i : \mathbb{R}^{m_i} \to \mathbb{R}, i = 1, \dots, n\} \in \{\underline{\text{すべて非減少 or すべて非増加}}\}$

$$\Longrightarrow \boxed{f\big(g_1(x_1^1, \dots, x_{m_1}^1), \dots, g_n(x_1^n, \dots, x_{m_n}^n)\big) \in \{\text{ 優モジュラ }\}}$$

とくに

$f : \mathbb{R}^n \to \mathbb{R} \in \{\underline{\text{非減少方位性凸}}\}$

$g_i : \mathbb{R}^{m_i} \to \mathbb{R} \in \{\underline{\text{非減少方位性凸}}\}, \ i = 1, \dots, n$

$$\Longrightarrow \boxed{f\Big(g_1(x_1^1, \dots, x_{m_1}^1), \dots, g_n(x_1^n, \dots, x_{m_n}^n)\Big) \in \{\text{ 非減少方位性凸 }\}}$$

(v) $f : \mathbb{R}^n \to \mathbb{R} \in \{\underline{\text{非減少方位性凸}}\}$

$g_i : \mathbb{R}^m \to \mathbb{R} \in \{\underline{\text{非減少方位性凸}}\}, \ i = 1, \dots, n$

$$\Longrightarrow \boxed{f\Big(g_1(x_1, \dots, x_m), \dots, g_n(x_1, \dots, x_m)\Big) \in \{\text{ 非減少方位性凸 }\}}$$

優モジュラ順序 ($\leq_{sm}$) と方位性凸順序 ($\leq_{dcx}$) (I)

X, Y: n 次元実数値確率ベクトル

$\boxed{X \leq_{sm} Y}$ X は 優モジュラ順序に関して Y より小さい

$\iff$ $\boxed{\mathrm{E}[f(X)] \leq \mathrm{E}[f(Y)]}$ (*)

$\forall f \in$ {<u>優モジュラ関数</u>} $\cap$ { 期待値が存在 }

$\boxed{X \leq_{ism} Y}$ X は 増加優モジュラ順序に関して Y より小さい

$\iff$ (*) $\forall f \in$ {<u>非減少優モジュラ関数</u>} $\cap$ { 期待値が存在 }

$\boxed{X \leq_{dcx} Y}$ X は 方位性凸順序に関して Y より小さい

$\iff$ (*) $\forall f \in$ {<u>方位性凸関数</u>} $\cap$ { 期待値が存在 }

$\boxed{X \leq_{idcx} Y}$ X は 増加方位性凸順序に関して Y より小さい

$\iff$ (*) $\forall f \in$ {<u>非減少方位性凸関数</u>} $\cap$ { 期待値が存在 }

優モジュラ順序 ($\leq_{sm}$) と方位性凸順序 ($\leq_{dcx}$) (II)

確率過程の場合

$\{X(t), t \in \mathbb{R}\}$, $\{Y(t), t \in \mathbb{R}\}$: 実数値確率過程

$\{X(t)\} \leq_{sm} \{Y(t)\}$ $\{X(t)\}$ は 優モジュラ順序に関して $\{Y(t)\}$ より小さい

$$\iff \forall n = 1, 2, \ldots, \forall t_1, \ldots, t_n \in \mathbb{R}$$

$$\big(X(t_1), \ldots, X(t_n)\big) \leq_{sm} \big(Y(t_1), \ldots, Y(t_n)\big)$$

$$\left\{ \begin{array}{l} \{X(t)\} \leq_{ism} \{Y(t)\} \\ \{X(t)\} \leq_{dcx} \{Y(t)\} \\ \{X(t)\} \leq_{idcx} \{Y(t)\} \end{array} \right\}$$ の定義も同様

離散時間型確率過程の場合も同様

$\le_{sm}$, $\le_{ism}$, $\le_{dcx}$, $\le_{idcx}$ の性質

$X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_n)$

- $\begin{cases} X \le_{sm} Y \Longrightarrow X \le_{dcx} Y \Longrightarrow X \le_{idcx} Y \\ X \le_{sm} Y \Longrightarrow X \le_{ism} Y \Longrightarrow X \le_{idcx} Y \end{cases}$ { 優モジュラ関数 } ⊃ { 方位性凸関数 }

- $X \le_{sm} Y \Longrightarrow \begin{cases} \mathrm{P}(X_1 > t_1, \dots, X_n > t_n) \le \mathrm{P}(Y_1 > t_1, \dots, Y_n > t_n) \\ \mathrm{P}(X_1 \le t_1, \dots, X_n \le t_n) \le \mathrm{P}(Y_1 \le t_1, \dots, Y_n \le t_n) \\ \forall (t_1, \dots, t_n) \in \mathbb{R}^n \end{cases}$ ($n = 2$ のとき ($\Longleftarrow$) も成立)

- $X \le_{sm} Y \Longrightarrow \mathrm{Cov}(X_i, X_j) \le \mathrm{Cov}(Y_i, Y_j)$ $i, j = 1, \dots, n,\ i \ne j$

- $X \le_{sm} Y \Longrightarrow X_i =_{st} Y_i$, $i = 1, \dots, n$ (周辺分布が等しい)

$\le_{sm}$, $\le_{ism}$, $\le_{dcx}$, $\le_{idcx}$ の性質 (つづき)

$X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$

- $X \le_{ism} Y \Longrightarrow \mathrm{P}(X_1 > t_1, \ldots, X_n > t_n) \le \mathrm{P}(Y_1 > t_1, \ldots, Y_n > t_n)$ $\forall (t_1, \ldots, t_n) \in \mathbb{R}^n$
- $X \le_{ism} Y \Longrightarrow X_i \le_{st} Y_i$, $i = 1, \ldots, n$
- $\begin{cases} X \le_{sm} Y \Longleftrightarrow X \le_{ism} Y \text{ かつ } X_i =_{st} Y_i, \ i = 1, \ldots, n \\ X \le_{sm} Y \Longleftrightarrow X \le_{ism} Y \text{ かつ } \mathrm{E}X_i = \mathrm{E}Y_i, \ i = 1, \ldots, n \end{cases}$
- $X \le_{dcx} \ [\le_{idcx}] \, Y \Longrightarrow X_i \le_{cx} \ [\le_{icx}] \, Y_i \quad i = 1, \ldots, n$

正の依存性との関係 (I)

Lorentz の不等式

$(X_1, \ldots, X_n)$: 周辺分布関数 $F_i(t) = \mathrm{P}(X_i \leq t)$ $(i = 1, \ldots, n)$
U: $(0,1)$ 上の一様分布に従う確率変数
$\Longrightarrow$ $\boxed{(X_1, \ldots, X_k) \leq_{sm} \left(F_1^{-1}(U), \ldots, F_k^{-1}(U)\right)}$
ただし $F_i^{-1}(u) = \inf\{x \in \mathbb{R} : F_i(x) \geq u\}$, $u \in [0,1)$

注意

- $\left(F_1^{-1}(U), \ldots, F_k^{-1}(U)\right)$ は周辺分布関数 $F_i(t)$ $(i = 1, \ldots, n)$ をもつ確率ベクトルの中で最も正の依存性が強い.
- その結合分布関数は $\boxed{F^+(t_1, \ldots, t_n) = \min\{F_1(t_1), \ldots, F_n(t_n)\}}$
- 周辺分布関数 $F_i(t)$ $(i = 1, \ldots, n)$ をもつ任意の確率ベクトルの結合分布関数 $F(t_1, \ldots, t_n) \Longrightarrow \boxed{F(t_1, \ldots, t_n) \leq F^+(t_1, \ldots, t_n)}$ $\forall t_1, \ldots, t_n \in \mathbb{R}$

正の依存性との関係 (II)

いろいろな正の依存性の概念

$X = (X_1, \ldots, X_n)$

(i) 条件付き増加 (Conditionally Increasing)

$\iff \forall f \in$ { 非減少関数 }, $\forall J \subset \{1, \ldots, n\}$

$\mathrm{E}[f(X_i) \mid X_j = t_j, j \in J]$ が t_j $(j \in J)$ に関して非減少

(ii) 関連している (Associated)

$\implies \forall f, g:\ \mathbb{R}^n \to \mathbb{R} \in$ { ともに非減少 or ともに非増加 }

$$\boxed{\mathrm{E}\big[f(X)\,g(X)\big] \geq \mathrm{E}[f(X)]\,\mathrm{E}[g(X)]} \iff \boxed{\mathrm{Cov}\big(f(X),\, g(X)\big) \geq 0}$$

(iii) 弱く関連している (Weakly Associated)

$\implies \forall I, J \subset \{1, \ldots, n\},\ I \cap J = \emptyset$

$\forall f:\ \mathbb{R}^{|I|} \to \mathbb{R},\ g:\ \mathbb{R}^{|J|} \to \mathbb{R} \in$ { ともに非減少 or ともに非増加 }

$$\boxed{\mathrm{E}\big[f(X_i, i \in I)\,g(X_j, j \in J)\big] \geq \mathrm{E}[f(X_i, i \in I)]\,\mathrm{E}[g(X_j, j \in J)]}$$

正の依存性との関係 (II のつづき)

(iv) 正の優モジュラ依存性をもつ (Positive Supermodular Dependent)

$\Longleftrightarrow X^{\perp} = (X_1^{\perp}, \ldots, X_n^{\perp})$: 互いに独立 かつ $X_i^{\perp} =_{st} X_i,\ i = 1, \ldots, n$

$$X^{\perp} \leq_{sm} X$$

(v) 正象限依存性をもつ (Positive Orthant Dependent)

$\Longleftrightarrow \forall t_1, \ldots, t_n \in \mathbb{R}$

$$\begin{cases} \mathrm{P}(X_1 > t_1) \cdots \mathrm{P}(X_n > t_n) \leq \mathrm{P}(X_1 > t_1, \ldots, X_n > t_n) \\ \mathrm{P}(X_1 \leq t_1) \cdots \mathrm{P}(X_n \leq t_n) \leq \mathrm{P}(X_1 \leq t_1, \ldots, X_n \leq t_n) \end{cases}$$

「条件付き増加」 $\Longrightarrow$ 「関連している」 $\Longrightarrow$ 「弱く関連している」 $\Longrightarrow$ 「正の優モジュラ依存性をもつ」 $\Longrightarrow$ 「正象限依存性をもつ」

$\leq_{dcx}$, $\leq_{idcx}$ の応用例: G/G/1 待ち行列の単調性

待ち行列 X と待ち行列 Y

τ_n^X, τ_n^Y $(n = 0, 1, 2, \ldots)$ 到着時間間隔の列

S_n^X, S_n^Y $(n = 0, 1, 2, \ldots)$ サービス時間の列

$\underline{W_{n+1}^X = \max(W_n^X + S_n^X - \tau_n^X, 0)}$ $(W_n^Y$ も同様): 待ち時間の列

[定理] $W_0^X = W_0^Y = 0$, $X_n = S_n^X - \tau_n^X$, $Y_n = S_n^Y - \tau_n^Y$ とする

$$(X_0, \ldots, X_n) \leq_{idcx} (Y_0, \ldots, Y_n), \quad n = 1, 2, \ldots$$

$$\Longrightarrow \underline{(W_1^X, \ldots, W_n^X) \leq_{idcx} (W_1^Y, \ldots, W_n^Y)}, \quad n = 1, 2, \ldots$$

とくに $\Longrightarrow$ $\underline{W_n^X \leq_{icx} W_n^Y}$, $\quad n = 1, 2, \ldots$

たとえば ...

- $(\tau_0^X, \ldots, \tau_n^X) \leq_{sm} (\tau_0^Y, \ldots, \tau_n^Y)$, $(S_0^X, \ldots, S_n^X) =_{st} (S_0^Y, \ldots, S_n^Y)$
$\Longrightarrow (W_1^X, \ldots, W_n^X) \leq_{idcx} (W_1^Y, \ldots, W_n^Y) \Longrightarrow W_n^X \leq_{icx} W_n^Y$

G/G/1 待ち行列の単調性: 証明の概略

(i) $W_n = \max(W_{n-1} + X_{n-1}, 0) = f_n(X_0, \dots, X_{n-1})$

$f_n : \mathbb{R}^n \to \mathbb{R} \in \{$ 非減少方位性凸 $\}$ を示せばよい

方位性凸関数の例 (v) より

(ii) $W_1 = \max(X_0, 0) \Longrightarrow f_1(x) = \max(x, 0) \in \{$ 非減少凸 $\}$

(iii) $f_n \in \{$ 非減少方位性凸 $\}$ を仮定 (帰納法)

(a) $W_{n+1} = \max(W_n + X_n, 0) = \max(f_n(X_0, \dots, X_{n-1}) + X_n, 0)$

$\Longrightarrow \boxed{f_{n+1}(x_0, \dots, x_n) = \max\big(f_n(x_0, \dots, x_{n-1}) + x_n, 0\big)}$

(b) $x + y : \mathbb{R}^2 \to \mathbb{R} \in \{$ 非減少方位性凸 $\}$ (定義から明らか)

$\Longrightarrow f_n(x_0, \dots, x_{n-1}) + x_n \in \{$ 非減少方位性凸 $\}$

方位性凸関数の例 (iv) より

$\Longrightarrow \max\big(f_n(x_0, \dots, x_{n-1}) + x_n,\ 0\big) \in \{$ 非減少方位性凸 $\}$

方位性凸関数の例 (ii) より

$\leq_{sm}$, $\leq_{ism}$ の応用例: マルコフ連鎖の単調性

$P = \{p_{i,j}\}$: マルコフ連鎖の推移確率行列（π: P の定常分布）

推移行列 P をもつマルコフ連鎖 $\{X_0, X_1, X_2, \ldots\}$ が 確率的に単調

$\Longleftrightarrow \forall f \in \{$ 非減少関数 $\}$

$\mathrm{E}[f(X_{n+1}) \mid X_n = i] = \sum_j f(j)\, p_{i,j}$ が i に関して非減少

マルコフ連鎖 $\{X_0, X_1, X_2, \ldots\}$ とその 逆過程 がともに確率的に単調

$\Longrightarrow$ $r_1 \leq r_2 \leq \cdots \leq r_n$, $s_1 \leq s_2 \leq \cdots \leq s_n$

$r_{i+1} - r_i \geq s_{i+1} - s_i \Longrightarrow (X_{r_1}, \ldots, X_{r_n}) \leq_{sm} (X_{s_1}, \ldots, X_{s_n})$

$\Longrightarrow$ $P_c := (1-c)\, I + c\, P$, $c \in (0, 1]$ (I: 単位行列)

他の状態への推移確率を $c\,(\leq 1)$ 倍, 定常分布は同じ ($P_c \mathbf{1} = \mathbf{1}$, $\pi\, P_c = \pi$)

$\{X_n^{(c)}, n = 0, 1, 2, \ldots\}$: 推移行列 P_c をもつ離散時間型マルコフ連鎖

$$c > c' \Longrightarrow (X_0^{(c)}, \ldots, X_n^{(c)}) \leq_{sm} (X_0^{(c')}, \ldots, X_n^{(c')})$$

おわりに

- ここで紹介したのはごく一部．他にもたくさんの確率順序の概念と豊富な応用がある
- あくまで道具である
  - 解析の途中で知っていると手間を省けることが（たまに）ある
  - 対象としているモデルが直接は解けないとき，解ける上下限を探してみよう
  - 数値結果に何らかの傾向が見られるなら，何らかの確率順序による単調性が成り立っているかも . . . ?
    $\Longrightarrow$ より一般的な結果が（意外と簡単に）得られるかもしれない

参考文献 (I)

最近の本

A. Müller & D. Stoyan (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley. (確率順序についての決定版．たくさんの確率順序の概念を扱っている．待ち行列への応用例も豊富．ただし応用に入るのは 4 章から．1〜3 章の理論パートが結構大変．)

S. Asmussen (2003). *Applied Probability and Queues, Second Edition*. Springer. ($\le_{cx}$, $\le_{icx}$ の待ち行列への応用のほか，$\le_{st}$ の応用例がちらほら)

F. Baccelli & P. Brémaud (2003). *Elements of Queueing Theory: Palm Martingale Calculus and Stochastic Recurrences, Second Edition*. Springer. (主にパルム理論に関する本だが，最後の章で確率順序を扱っている．$\le_{st}$, $\le_{cx}$, $\le_{icx}$ の待ち行列への応用例がいろいろ．)

H. Chen & D. D. Yao (2001). *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization*. Springer. (尤度比順序 (Likelihood Ratio Order; $\le_{lr}$) とその応用例が少し)

参考文献 (II)

他に参考にした論文

L. E. Meester & J. G. Shanthikumar (1993). Regularity of stochastic processes: A theory based on directional convexity. *Probab. Engrg. Inform. Sci.* **7**. 343–360. (方位性凸順序とその応用を紹介)

R. Kulik & R. Szekli (2005). Dependence orderings for some functionals of multivariate point processes. *J. Multivariate Anal.* **92**. 145–173. (Appendix に優モジュラ関数，方位性凸関数の例がたくさん)

T. C. Christofides & E. Vaggelatou (2004). A connection between supermodular ordering and positve/negative association. *J. Multivariate Anal.* **88**. 138–151. (「弱い関連」$\Longrightarrow$「正の優モジュラ依存性」を証明)

T. Hu, A. Müller & M. Scarsini (2004). Some counterexamples in positive dependence. *J. Statist. Plann. Inference*. **124**. 153–158. (「関連」$\Longrightarrow$「弱い関連」$\Longrightarrow$「正の優モジュラ依存性」$\Longrightarrow$「正象限依存性」の逆が成り立たない例を紹介)