第7回課題

33B21009 情報科学研究科 情報数理学専攻
羽田充宏

⑦

(a) => (b)

$^\exists B \in \mathbb{R}^{n \times n}$ $A = B^T B$ と表せるので

$\forall x \in \mathbb{R}^n$, $x^T A x = x^T B^T B x$

$= \|Bx\|^2 \geq 0$ ▨

(b) => (c)

$\forall x \in \mathbb{R}^n$, $x^T A x \geq 0$ より

Aの固有ベクトル u として, 対応する固有値λとすると,

$$u^T A u = \lambda u^T u$$
$$= \lambda \|u\|^2 \geq 0$$

従って任意の固有値 $\lambda \geq 0$ である ▨

(c) => (a)

Aのすべての固有値が非負より

ある直交行列 $P \in \mathbb{R}^{n \times n}$ が存在して

$$P^T A P = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \quad (\lambda_i \text{はAの固有値})$$

$$A = P\,\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})\,\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})\,P^T$$

$B^T = P\,\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ とすれば

$A = B^T B$ と表せる. ▨

③

$A\in\mathbb{R}^{3\times3}$ の固有値 $\lambda_1, \lambda_2, \lambda_3$ として

Aの固有多項式は

$$\Phi_A(\lambda) = \det(I\lambda - A) = (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$$

と表せる

よって
$$\det(A) = (-1)^3\Phi_A(0) = -(-\lambda_1)(-\lambda_2)(-\lambda_3) = \lambda_1\lambda_2\lambda_3$$

また, $\det(A) < 0$ であれば $\lambda_1, \lambda_2, \lambda_3 \neq 0$ で,

少なくとも1つは負の固有値となるので A は非負定値とはならない

□

④

$A, B \succeq O \in \mathbb{R}^{n\times n}$

この時 $\circ$ を Hadamard 積として $A\circ B \succeq O$ を示す.

A, B を固有値分解して

$\lambda_1,\dots,\lambda_n \geq 0$　　対応する固有ベクトル　$v_1,\dots,v_n \in\mathbb{R}^n$

$\nu_1,\dots,\nu_n \geq 0$　　　　　　　　　　　　　$u_1,\dots,u_n \in\mathbb{R}^n$

として

$$A = \sum_{i=1}^{n}\lambda_i v_i v_i^T \qquad B = \sum_{j=1}^{n}\nu_j u_j u_j^T$$ と表せる.

ここで

$$v_i v_i^T \circ u_j u_j^T = \begin{pmatrix} v_{i1}v_{i1} & \cdots & v_{i1}v_{in} \\ \vdots & & \\ v_{in}v_{i1} & \cdots & v_{in}v_{in}\end{pmatrix} \circ \begin{pmatrix} u_{j1}u_{j1} & \cdots & u_{j1}u_{jn} \\ \vdots & & \vdots \\ u_{jn}u_{j1} & \cdots & u_{jn}u_{jn}\end{pmatrix}$$

$$= \begin{pmatrix} v_{i1}u_{j1}v_{i1}u_{j1} & \cdots & v_{i1}u_{j1}v_{in}u_{jn} \\ \vdots & & \vdots \\ v_{in}u_{jn}v_{i1}u_{j1} & \cdots & v_{in}u_{jn}v_{in}u_{jn}\end{pmatrix}$$

$$= (v_i\circ u_j)(v_i\circ u_j)^T$$

従って

$$A \circ B = \sum_i \sum_j \lambda_i \nu_j (v_i \circ u_j)(v_i \circ u_j)^T$$

よって $A \circ B \succeq O$ ■

また正定値カーネル $k_1, k_2 : E \times E \to \mathbb{R}$ に対して $u, v \in E$

$k_1 k_2 (u, v) := k_1(u, v) k_2(u, v)$ で積を定義すると.

$k_1 k_2$ のグラム行列 $K_{12} \in \mathbb{R}^{n \times n}$ は

k_1 のグラム行列 $K_1 \in \mathbb{R}^{n \times n}$

k_2 のグラム行列 $K_2 \in \mathbb{R}^{n \times n}$ を用いて

$$K_{12} = \begin{bmatrix} k_1(x_1, x_1) k_2(x_1, x_1) & \cdots & k_1(x_1, x_n) k_2(x_1, x_n) \\ \vdots & & \\ k_1(x_n, x_1) k_2(x_n, x_1) & \cdots & k_1(x_n, x_n) k_2(x_n, x_n) \end{bmatrix}$$

$$= K_1 \circ K_2$$

$K_1 \succeq O,\ K_2 \succeq O$ より $K_{12} \succeq O$ ■

⑤ $\mathbb{R} \ni a > 0$ として $\begin{pmatrix} a & \cdots & a \\ \vdots & & \vdots \\ a & \cdots & a \end{pmatrix} \succeq O \in \mathbb{R}^{n \times n}$ を示す.

$$\begin{pmatrix} a & \cdots & a \\ \vdots & & \vdots \\ a & \cdots & a \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{a}{n}} & \cdots & \sqrt{\frac{a}{n}} \\ \vdots & & \vdots \\ \sqrt{\frac{a}{n}} & \cdots & \sqrt{\frac{a}{n}} \end{pmatrix}^T \begin{pmatrix} \sqrt{\frac{a}{n}} & \cdots & \sqrt{\frac{a}{n}} \\ \vdots & & \vdots \\ \sqrt{\frac{a}{n}} & \cdots & \sqrt{\frac{a}{n}} \end{pmatrix}$$ と書ける.

よって命題が成り立つ

また, $\forall x, y \in E,\ k(x, y) \mapsto a > 0$ となるようなカーネル関数のグラム行列は上での定数行列と一致するため

このようなカーネル関数は正定値カーネルとなる ■

⑥ 多項式カーネル $k_{3,2}(x,y)=(x^\top y+1)^3,\ x,y\in\mathbb{R}^2$

の特徴写像

$$\Phi_{3,2}(x_1,x_2)=[1,\sqrt{3}x_1,\sqrt{3}x_2,\sqrt{6}x_1x_2,\sqrt{3}x_1^2,\sqrt{3}x_2^2,$$
$$x_1^3,\sqrt{3}x_1^2x_2,\sqrt{3}x_1x_2^2,x_2^3]$$

$\langle\Phi_{3,2}(x),\Phi_{3,2}(y)\rangle$

$$=1+3x_1y_1+3x_2y_2+6x_1x_2y_1y_2$$
$$+3x_1^2y_1^2+3x_2^2y_2^2+x_1^3y_1^3$$
$$+3x_1^2x_2y_1^2y_2+3x_1x_2^2y_1y_2^2$$
$$+x_2^3y_2^3$$

m_0 +	m_1 +	m_2 = 3
3	0	0
2	1	0
2	0	1
1	1	1
1	2	0
1	0	2
0	3	0
0	2	1
0	1	2
0	0	3

$-\frac{1}{\text{P}}$

$$(1+x_1y_1+x_2y_2)^3$$
$$=(1+x_1y_1+x_2y_2)(1+x_1^2y_1^2+x_2^2y_2^2+2x_1y_1+2x_2y_2+2x_1y_1x_2y_2)$$
$$=1+x_1^2y_1^2+x_2^2y_2^2+2x_1y_1+2x_2y_2+2x_1y_1x_2y_2$$
$$+x_1y_1+x_1^3y_1^3+x_1y_1x_2^2y_2^2+2x_1^2y_1^2+2x_1y_1x_2y_2+2x_1^2y_1^2x_2y_2$$
$$+x_2y_2+x_1^2y_1^2x_2y_2+x_2^3y_2^3+2x_1y_1x_2y_2+2x_2^2y_2^2+2x_1y_1x_2^2y_2^2$$
$$=1+3x_1^2y_1^2+3x_2^2y_2^2+3x_1y_1+3x_2y_2+6x_1y_1x_2y_2$$
$$+x_1^3y_1^3+3x_1y_1x_2^2y_2^2+3x_1^2y_1^2x_2y_2+x_2^3y_2^3$$

従って $\langle \Phi_{3,2}(x), \Phi_{3,2}(y) \rangle = (1+x^\top y)^3$ 四

⑦ 指数カーネル

$$k_m(x,y) = 1 + \beta x^\top y + \frac{\beta^2}{2}(x^\top y)^2 + \cdots + \frac{\beta^m}{m!}(x^\top y)^m \quad (m \geq 1)$$

を考える.

ここで $(x^\top y)^k = \underbrace{(x^\top y)\cdots(x^\top y)}_{k\text{コ}}$ は,正定値カーネル(内積)の k 回の積で表せるので

これもまた正定値カーネル

正定値カーネルの非負係数の1次結合も正定値カーネルより

k_m は正定値カーネル よって 命題4(3) より 指数カーネル $\lim_{m\to\infty} k_m(x,y)$

は正定値カーネル

Gaussカーネル　$x, y \in \mathbb{R}^d$,

$$k(x,y) := \exp\left(-\frac{1}{2\sigma^2}\|x-y\|_2^2\right), \quad \sigma > 0$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left(\|x\|^2 - 2x^\top y + \|y\|^2\right)\right\}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\|x\|^2\right)\exp\left(\frac{1}{2\sigma^2}x^\top y\right)\exp\left(-\frac{1}{2\sigma^2}\|y\|^2\right)$$

ここで

$$\exp\left(\frac{1}{2\sigma^2}x^\top y\right) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{x^\top y}{2\sigma^2}\right)^k$$

よって 命題4 (3) より $k(x,y) = \exp\left(\frac{1}{2\sigma^2}x^\top y\right)$

は正定値カーネルなので

命題4 (5) より Gaussカーネルは正定値カーネル.

線形カーネル

$$k_{m,d}(x,y) = (1+x^\top y)^m = \underbrace{(1+x^\top y)\cdots(1+x^\top y)}_{m\text{コ}}$$

正定値カーネルの積より

線形カーネルも正定値カーネル 四

正定値カーネル $k : E \times E \to \mathbb{R}$ に対して

$$\tilde{k}(x, y) = \frac{k(x, y)}{\sqrt{k(x,x)k(y,y)}}$$ で定める $\tilde{k}$ も

$$= \frac{1}{\sqrt{k(x,x)}} k(x,y) \frac{1}{\sqrt{k(y,y)}}$$

$$= f(x) \quad k(x,y) \quad f(y)$$ の形で表せるので

正定値カーネル

Gaussカーネル正規化

$$k(x,x) = \exp\left(-\frac{1}{2\sigma^2}\|x-x\|^2\right) = 1$$ より

正規化したカーネル $\tilde{k}$ として

$\tilde{k}(x, y) = k(x, y)$ で変化しない

指数カーネル正規化

$$\tilde{k}(x,y) = \frac{\exp(\beta x^T y)}{\sqrt{\exp(\beta x^T x)\exp(\beta y^T y)}}$$

$$= \exp\left\{-\frac{\beta}{2}(x^T x - 2x^T y + y^T y)\right\}$$

$$= \exp\left\{-\frac{\beta}{2}\|x-y\|^2\right\}$$

よって指数カーネルを正規化するとGaussカーネルとなる.

(9)

$(E, \mathcal{F}, \mu)$ 確率sp. $E = \{1,2,3,4,5,6\}$

$X : E \to \mathbb{R}$

$$X(e) = \begin{cases} 1 & (e = 1, 3, 5) \\ 0 & (\text{others}) \end{cases}$$

$\mathcal{F} = \{\{1,2,3\}, \{4,5,6\}, \phi, E\}$ の時.

$X^{-1}(\{1\}) = \{1, 3, 5\} \notin \mathcal{F}$ より

X は $\mathcal{F}$-可測でないので確率変数でない □

(10) 正規分布の特性関数は

$$\varphi(t) = E[\exp(itX)]$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} - itx\right) dx$$

肩の部分について

$$-\frac{(x-\mu)^2}{2\sigma^2} - \frac{2\sigma^2 itx}{2\sigma^2} = -\frac{x^2 - 2\mu x + 2\sigma^2 itx + \mu^2}{2\sigma^2}$$

$$= -\frac{x^2 - 2(\mu + \sigma^2 it)x + \mu^2}{2\sigma^2}$$

$$= -\frac{(x - \mu + \sigma^2 it)^2 - (\mu + \sigma^2 it)^2 + \mu^2}{2\sigma^2}$$

$$= -\frac{(x - \mu + \sigma it)^2 - 2\mu\sigma^2 it + \sigma^4 t^2}{2\sigma^2}$$

$$= \frac{(x - \mu + \sigma it)^2}{2\sigma^2} + \mu it - \frac{\sigma^2 t^2}{2}$$

よって $\varphi(t) = \exp\left(-\frac{\sigma^2 t^2}{2} + \mu it\right)$

$\mu = 0$ の時 $\forall t \in \mathbb{R}$, $\varphi(t)$ は実数値をとる

ラプラス分布の特性関数は、

$$\varphi(t) = \int_{-\infty}^{\infty} \frac{\alpha}{2} \exp(-\alpha|x| + itx)\, dx$$

$$= \int_{-\infty}^{0} \frac{\alpha}{2} \exp((it+\alpha)x)\, dx + \int_{0}^{\infty} \frac{\alpha}{2} \exp((it-\alpha)x)\, dx$$

$$= \frac{\alpha}{2}\left\{\left[\frac{e^{(it+\alpha)x}}{it+\alpha}\right]_{-\infty}^{0} + \left[\frac{e^{(it-\alpha)x}}{it-\alpha}\right]_{0}^{\infty}\right\}$$

$$= \frac{\alpha}{2}\left(\frac{1}{it+\alpha} - \frac{1}{it-\alpha}\right) = \frac{\alpha}{2} \frac{it - \alpha - (it+\alpha)}{-t^2 - \alpha^2} = \frac{\alpha^2}{\alpha^2 + t^2}$$

$\forall \alpha > 0$, $\varphi(t)$ は実数値をとる. □

(13)

文字列カーネル

E_1, E_3 : 有限長の文字列の集合

E_2 : 長さ p の文字列の集合　　とする.

$$R : E_1 \times E_2 \times E_3 \to E$$
$$(x_1, x_2, x_3) \longmapsto \underset{\sim}{x}$$

x_1, x_2, x_3 を連結した文字列

として写像 R を定義する.

この時　$k_1 : E_1 \times E_1 \to \mathbb{R}$

$k_2 : E_2 \times E_2 \to \mathbb{R}$

$k_3 : E_3 \times E_3 \to \mathbb{R}$ として

$$k_1(x_1, y_1) = 1$$
$$k_3(x_3, y_3) = 1$$
$$k_2(x_2, y_2) = \begin{cases} 1 & (\text{if } x_2 = y_2) \\ 0 & (\text{others}) \end{cases}$$

k_1, k_2, k_3 はそれぞれ正定値カーネルであり,

文字列カーネルは

$x, y \in E$

$$k(x, y) = \sum_{R(x_1, x_2, x_3) = x} \sum_{R(y_1, y_2, y_3) = y} 1 \cdot k_2(x_2, y_2) \cdot 1$$

以上より文字列カーネルはたたみ込みカーネルであり,

正定値カーネルの和で書けているので正定値カーネルである.

木カーネル

木 x, y に対して

$$k(x, y) := \sum_t C_t(x) C_t(y)$$

ここで $C_t(x)$：木 x における部分木 t の頻度

t をとり出す集合としては x と y の共通の木を考えれば十分である。

ここで
$$I_t(a) = \begin{cases} 1 & t \text{ が } a \text{ をルートにもつ部分木} \\ 0 & \text{others} \end{cases}$$

によって $I_t(\cdot)$ を定義すると、

$$C_t(x) = \sum_{a \in V_x} I_t(a)$$ と表せるので

$$k(x, y) = \sum_t \sum_{a \in V_x} \sum_{b \in V_y} I_t(a) I_t(b)$$

$$\sum_i \sum_j z_i z_j k(x_i, x_j) = \sum_t \sum_i \sum_j \sum_{a \in x_i} \sum_{b \in x_j} z_i z_j I_t(a) I_t(b)$$

$$= \sum_t \sum_i \sum_j z_i z_j C_t(x_i) C_t(x_j)$$

$$= \sum_t \left\{ \sum_i z_i C_t(x_i) \right\}^2 \geq 0$$

よって木カーネルは正定値カーネル。

周辺化カーネル

確率変数 X, Y についてそのとりうる値の集合 E_X, E_Y で、
$P(y|x)$ $(x \in E_X, y \in E_Y)$ で条件付き確率が与えられ、
$k_{XY} : (E_X \times E_Y) \times (E_X \times E_Y) \to \mathbb{R}$ が与えられている時、

$$k(x, x') := \sum_{y \in E_Y} \sum_{y' \in E_Y} k_{XY}((x, y), (x', y')) P(y|x) P(y'|x')$$

$x, x' \in E_X$

を周辺化カーネルとよぶ

この時

k_{XY} が正定値カーネルとすれば

$\exists \Phi : E_X \times E_Y \to \mathcal{H}$,

$$k_{XY}((x,y),(x',y')) = \langle \Phi(x,y), \Phi(x',y') \rangle$$ と表せる.

$$k(x,x') = \sum_{y \in E_Y} \sum_{y' \in E_Y} P(y|x) P(y'|x') \langle \Phi(x,y), \Phi(x',y') \rangle$$

$$= \left\langle \sum_{y \in E_Y} P(y|x) \Phi(x,y), \sum_{y' \in E_Y} P(y'|x') \Phi(x',y') \right\rangle$$

新たに $\Phi' : E_X \to \mathcal{H}$

$$x \mapsto \Phi'(x) = \sum_{y \in E_Y} P(y|x) \Phi(x,y)$$ と定めれば

$$k(x,x') = \langle \Phi'(x), \Phi'(x') \rangle_{\mathcal{H}}$$ と書くことができ,

$$\sum_i \sum_j z_i z_j \langle \Phi'(x_i), \Phi'(x_j) \rangle_{\mathcal{H}} = \left\langle \sum_i z_i \Phi'(x_i), \sum_j z_j \Phi'(x_j) \right\rangle_{\mathcal{H}}$$

$$= \left\| \sum_i z_i \Phi'(x_i) \right\|_{\mathcal{H}}^2 \geq 0$$

で周辺化カーネルは正定値カーネル ☐

グラフカーネル

グラフ G_1, G_2 に対して以下でカーネルを定める.

$$k(G_1, G_2) := \sum_{\pi_1} \sum_{\pi_2} P(\pi_1 | G_1) P(\pi_2 | G_2) I(L(\pi_1) = L(\pi_2))$$

上の周辺化カーネルの定義で

$$k_{XY}((G_1,\pi_1),(G_2,\pi_2)) = I[L(\pi_1) = L(\pi_2)]$$

とすれば

グラフカーネルは周辺化カーネル

☐

14

(a) $3 \to 1 \to 4 \to 3 \to 5$

$$\frac{1}{5}\left(\frac{\cancel{2}}{3}\cdot\frac{1}{\cancel{2}}\right)\left(\frac{\cancel{2}}{3}\cdot\frac{1}{\cancel{2}}\right)\left(\frac{2}{3}\cdot 1\right)\left(\frac{\cancel{2}}{3}\cdot\frac{1}{\cancel{2}}\right)\cdot\frac{1}{3} = \frac{1}{5}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}$$

$$= \frac{2}{5\cdot 3^5}$$

(b) $1 \to 2 \to 4 \to 1 \to 2$

$$\frac{1}{5}\left(\frac{2}{3}\cdot\frac{1}{2}\right)\left(\frac{2}{3}\cdot 1\right)\left(\frac{2}{3}\cdot 0\right)\left(\frac{2}{3}\cdot\frac{1}{2}\right)\cdot\frac{1}{3} = 0$$

(c) $3 \to 5 \to 3 \to 5$

$$\frac{1}{5}\left(\frac{\cancel{2}}{3}\cdot\frac{1}{\cancel{2}}\right)\left(\frac{2}{3}\cdot 1\right)\left(\frac{\cancel{2}}{3}\cdot\frac{1}{\cancel{2}}\right)\frac{1}{3} = \frac{1}{5}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}$$

$$= \frac{2}{5\cdot 3^4}$$