Huỳnh Nguyễn Thế Dân 21110256

Bài tập lý thuyết
Kernel PCA

Bài 1):

a). Vì $(K_l)_{l \in N}$ là các kernel hợp lệ, nên

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K_l(x_i, x_j) \geq 0,$$

$$\forall x_1, x_2, \dots, x_n \in R^p$$
$$\forall c_1, \dots, c_n \in R$$

* Xét $K(x_i, x_j) = \sum_{l=1}^{n} \lambda_l K_l(x_i, x_j)$

Với mọi $x_1, \dots, x_n \in R^p$, $c_1, \dots, c_n \in R$, ta có:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K(x_i, x_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \left(\sum_{l=1}^{n} \lambda_l K_l(x_i, x_j)\right)$$

$$= \sum_{l=1}^{n} \lambda_l \overbrace{\left[\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K_l(x_i, x_j)\right]}^{A} \geq 0 \quad (\text{vì } A \geq 0)$$

Vậy tổng trên là một kernel hợp lệ

• Đầu tiên, ta cần chứng minh tích 2 kernel hợp lệ thì hợp lệ

Xét $k(x_i, x_j) = k_1(x_i, x_j) . k_2(x_i, x_j)$

Có:
$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j k(x_i, x_j) = \sum_{i=1}^{n} c_i k_1(x_i, x_j) . \sum_{j=1}^{n} c_j k_2(x_i, x_j)$$

Mà ở câu a ta đã chứng minh $\sum_{i=1}^{n} c_i k_1(x_i, x_j)$ là kernel hợp lệ nên tích trên ≥ 0

⇒ Tích 2 kernel hợp lệ cũng là kernel hợp lệ

Xét $k_1^{a_1} \dots k_n^{a_n}$

$= (k_1^{a_1} \dots k_n^{a_n - 1}) . k_n$

Có k_n là kernel hợp lệ

Ta xét tương tự đến khi còn chỉ còn $(k_1 . k_1) \dots$ thì ta được một kernel hợp lệ nhân với chuỗi kernel hợp lệ

Vậy tích $k_1^{a_1} \dots k_n^{a_n}$ là một kernel hợp lệ

- Ta biết rằng mỗi K_n là một kernel hợp lệ nên

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K_n(x_i, x_j) \geq 0$$

Ta có:

$$\lim_{n\to\infty} \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K_n(x_i, x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \lim_{n\to\infty} K_n(x_i, x_j)$$

Với x_i và x_j cụ thể, $\lim_{n\to\infty} K_n(x_i, x_j)$ là số thực. Đặt giới hạn này là $K(x_i, x_j)$, do đó:

$$\lim_{n\to\infty} \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K_n(x_i, x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K(x_i, x_j)$$

Vì $K(x_i, x_j)$ là một kernel hợp lệ nên:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K(x_i, x_j) \geq 0$$

Nên $K = \lim_{n\to\infty} K_n(x_i, x_i) = K(x_i, x_i)$ cũng là kernel hợp lệ

TITAN EDUCATION

b) Để chứng minh K là một kernel hợp lệ, ta cần chứng minh rằng mọi ma trận gram được tạo ra từ K ≥ 0

Giả sử $X = (x_1, \dots, x_n)$ và $Y = (y_1, \dots, y_n)$ là 2 tập hợp các điểm trong $R^{D_1} \times \dots \times R^{D_n}$. Khi đó ma trận Gram định bởi:

$$G(X, Y) = \begin{bmatrix} \prod_{i=1}^{n} K_i(x_i, y_1) & \cdots & \prod_{i=1}^{n} K_i(x_i, y_n) \\ \vdots & \ddots & \vdots \\ \prod_{i=1}^{n} K_i(x_n, y_1) & \cdots & \prod_{i=1}^{n} K_i(x_n, y_n) \end{bmatrix}$$

Vì K_i là kernel hợp lệ, nên $K_i(x_i, y_i) \geq 0$

$\Rightarrow$ Tích của các phần tử này luôn lớn hơn 0

Vậy K là kernel hợp lệ trên $R^{D_1} \times \dots \times R^{D_n}$

Tương tự với K(tổng), có ma trận Gram là

$$G(X, Y) = \begin{bmatrix} K(x_1, y_1) & \cdots & K(x_n, y_1) \\ \vdots & \ddots & \vdots \\ K(x_n, y_1) & \cdots & K(x_n, y_n) \end{bmatrix}$$

cũng là một kernel hợp lệ.

c) Để chứng minh hàm thu hẹp K_R của K-biến R là một kernel hợp lệ, ta làm tương tự câu b với ma trận Gram là:

$$G(X,Y) = \begin{bmatrix} K_R(x_1, y_1) & \cdots & K_R(x_1, y_n) \\ \vdots & \ddots & \vdots \\ K_R(x_n, y_n) & \cdots & K_R(x_n, y_n) \end{bmatrix}$$

d)

- Có q là một đa thức với hệ số không âm, khi đó $q(K(x,y))$ cũng là đa thức với hệ số không âm

Ta chọn các hệ số bằng 0 sao cho $K'(x,y) = K(x,y)$ là một kernel hợp lệ

- Với $K'(x,y) = e(K(x,y))$, xét $e(x) > 0$ với mọi z

Ta có thể xây dựng sao cho $K'(x,y) = K(x,y)$ thì ta cũng được kernel hợp lệ.

Câu 2

a)

$K(x,y) = (1 + x^T y)^2$ với $x, y \in R^3$

Đặt $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$

$K(x,y) = (1 + x_1 y_1 + x_2 y_2 + x_3 y_3)^2$

$= 1.1 + \sum_{i=1}^{n} x_i^2 y_i^2 + \sum_{i=1}^{n} \sqrt{2} x_i \sqrt{2} y_i + \sqrt{2} x_1 x_2 \sqrt{2} y_1 y_2$

$+ \sqrt{2} x_1 x_3 \sqrt{2} y_1 y_3$

$+ \sqrt{2} x_2 x_3 \sqrt{2} y_2 y_3$

Vậy công thức

$\phi(x) = (1, x_1^2, x_2^2, x_3^2, \sqrt{2} x_1, \sqrt{2} x_2, \sqrt{2} x_3, \sqrt{2} x_1 x_2$

$, \sqrt{2} x_2 x_3$

$, \sqrt{2} x_1 x_3) \in R^{10}$

và không gian thuộc tính có 10 chiều

Bài 3)

a) $K_1(x,y) = 1 + x^T y$ với $x, y \in R^D$

$K_1(x,y) = 1.1 + x_1 y_1 + x_2 y_2 + \dots + x_D y_D$

$\Rightarrow \phi_1(x) = (1, x_1, \dots, x_D)$

$K_2(x,y) = x^T y + \|x\| \, \|y\|$

$K_2(x,y) = (x_1 y_1 + \dots + x_D y_D) + \sqrt{x_1^2 + \dots + x_D^2} \cdot \sqrt{y_1^2 + \dots + y_D^2}$

$\Rightarrow \phi_2(x) = (x_1, \dots, x_D, \sqrt{x_1^2 + \dots + x_D^2})$

$K(x,y) = K_1(x,y) + K_2(x,y)$

$= 1 + 2x^T y + \|x\| \, \|y\|$

$K(x,y) = 1.1 + 2x_1 y_1 + 2x_2 y_2 + \dots + 2x_D y_D + \sqrt{x_1^2 + \dots + x_D^2} \cdot \sqrt{y_1^2 + \dots + y_D^2}$

$\phi(x) = (1, \sqrt{2}x_1, \sqrt{2}x_2, \dots, \sqrt{2}x_D, \sqrt{x_1^2 + \dots + x_D^2})$
$= (\phi_1(x), \phi_2(x))$

Bài 4)

a) Để chứng minh, giả sử α_j là 1 vectơ riêng của ma trận Kernel K tương ứng với λ_j, tức là:

$K a_j = \lambda_j \alpha_j$

Ta tìm α' sao cho $K^2 \alpha' = n \lambda_j K \alpha'$

Có

$\alpha' = a_j + \varepsilon\beta$

Ta thấy rằng $K(a_j + \beta) = \lambda_j(\alpha_j + \varepsilon\beta)$

Bây giờ ta tính $K^2\alpha'$:

$K^2\alpha' = K(K\alpha') = K(K(a_j + \varepsilon\beta))$
$= K(\lambda_j(a_j + \varepsilon\beta)) = \lambda_i K(\alpha_j + \varepsilon\beta)$

Do $K\beta = 0$, ta có:

$K^2\alpha' = \lambda_j K \alpha_j = n\lambda_j a_j$ (1)

Với n là hằng số

Do đó, $K^2\alpha' = n\lambda_i K\alpha'$, với α' đã cho, ta có nghiệm (1) với trị riêng λ_j

b) Phép chiếu lên thành phần chính được biểu diễn bởi

$$\phi(x)^T v_j = \sum_{i=1}^{n} a_{ji} K(x, x_j). \qquad (2)$$

Giả sử α và α' là hai vectơ thoả (2), tức $K\alpha = n\lambda\alpha$ và $K\alpha' = n\lambda\alpha'$, λ là trị riêng và α, α' là vectơ tương ứng

Ta muốn chứng minh
$$\sum_{i=1}^{n} a_{ji} K(x, x_j) = \sum_{i=1}^{n} a'_{ji} K(x, x_j)$$

Vì $K\alpha = n\lambda\alpha$ và $K\alpha' = n\lambda\alpha'$, ta có:

$$K(\alpha - \alpha') = n\lambda(\alpha - \alpha')$$

Từ đây:
$$K\alpha - K\alpha' = n\lambda\alpha - n\lambda\alpha'$$
$$n\lambda\alpha - n\lambda\alpha' = n\lambda(\alpha - \alpha')$$

Do đó, $K(\alpha - \alpha') = n\lambda(\alpha - \alpha')$, tức K giữ nguyên khác biệt giữa α và α' nên phép chiếu không ảnh hưởng

TITAN EDUCATION

b) Trong PCA thông thường, ta có:

$$Cov(X) = \frac{1}{n}(X-\bar{X})^T(X-\bar{X})$$

Trong Kernel PCA với hàm kernel tuyến tính

$$K = XX^T$$

Có ma trận trung tâm H:

$$H = I - \frac{1}{n}1.1^T$$

Sau đó ta tính ma trận trung tâm ở không gian chiều cao hơn F bằng cách sử dụng hàm Kernel:

$$\bar{K} = HKH$$

*So sánh:

PCA thông thường:

$$Cov(X) = \frac{1}{n}(X-\bar{X})^T(X-\bar{X})$$

Kernel PCA với hàm kernel tuyến tính:

$$\bar{K} = HKH = H(XX^T)H = \frac{1}{n}(X-\bar{X})(X-\bar{X})^T$$

Cả 2 đều là các phép biến đổi ~~tương đương~~ tuyến tính. Do đó vectơ riêng và giá trị riêng tương đương

Vậy PCA thông thường là trường hợp đặc biệt của kernel PCA khi dùng hàm kernel tuyến tính