2019年04月21日

CSC411 Homework 7

1. a) $F = \frac{1}{N}\sum_{i=1}^{N} L(g(w^T\varphi(x^i)), t^i) + \frac{\lambda}{2}w^Tw$

$w^Tw = w_1^2 + \quad + w_D^2$

$$\frac{dF}{dw} = \frac{1}{N}\sum_{i=1}^{N} \frac{dL}{dg}\frac{dg}{dz}\frac{dz}{w} + \frac{\lambda}{2}2w^T$$

$$0 = \frac{1}{N}\sum_{i=1}^{N} \frac{dL}{dg}\frac{dg}{dz}\varphi(x^i)^T + \lambda w^T$$

$Z = w^T\varphi(x) = w_1\varphi_1(x) + \quad + w_D\varphi_D$

$$w^T = -\frac{1}{\lambda N}\sum_{i=1}^{N} \frac{dL}{dg}\frac{dg}{dz}\varphi(x^i)^T$$

↳ scalar

↳ scalar

∴ Clearly the optimal weights is a linear combination of the row space of $\Psi = \begin{bmatrix} -\varphi(x^1)^T - \\ \vdots \\ -\varphi(x^N)^T - \end{bmatrix}$

b)

$$F = \frac{1}{2N}\|t - \Psi\Psi^T\alpha\|^2 + \frac{\lambda}{2}(\alpha^T\Psi)(\Psi^T\alpha)$$

$$= \frac{1}{2N}(t-K\alpha)^T(t-K\alpha) + \frac{\lambda}{2}\alpha^TK\alpha$$

$$= \frac{1}{2N}(t^T - \alpha^TK^T)(t-K\alpha) + \frac{\lambda}{2}\alpha^TK\alpha$$ — all scalars. Can transpose te

$$= \frac{1}{2N}(t^Tt - t^TK\alpha - \alpha^TK^Tt + \alpha^TK^TK\alpha) + \frac{\lambda}{2}\alpha^TK\alpha$$

$$= \frac{1}{2N}(t^Tt - t^TK\alpha - t^TK\alpha + \alpha^TK^TK\alpha) + \frac{\lambda}{2}\alpha^TK\alpha$$

$$= \frac{1}{2N}(t^Tt - 2t^TK\alpha + \alpha^TK^TK\alpha) + \frac{\lambda}{2}\alpha^TK\alpha$$

$$= \frac{1}{2}\alpha^T\left(\frac{1}{N}K^TK + \lambda K\right)\alpha - \frac{1}{N}t^TK\alpha + \frac{1}{2N}t^Tt$$

$$= \frac{1}{2}\alpha^T \; A \; \alpha + b^T\alpha + \frac{1}{2N}t^Tt$$

$\boxed{A = \frac{1}{N}K^TK + \lambda K}$, $b^T = -\frac{1}{N}t^TK \Rightarrow \boxed{b = -\frac{1}{N}K^Tt}$

Optimal α: $\boxed{\alpha = -A^{-1}b}$, A, b as above

2. a) $K_S = K_1 + K_2 = \varphi_1(x)^T \varphi_1(x') + \varphi_2(x)^T \varphi_2(x')$

$\therefore$ Define $\varphi_S(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \end{bmatrix}$

$$\therefore K_S = \varphi_S(x)^T \varphi_S(x') = \begin{bmatrix} \varphi_1^T(x) & \varphi_2^T(x) \end{bmatrix} \begin{bmatrix} \varphi_1(x') \\ \varphi_2(x') \end{bmatrix}$$

$$= \varphi_1^T(x)\,\varphi_1(x') + \varphi_2^T(x)\,\varphi_2(x')$$

as required.

b) $K_P(x,x') = K_1(x,x')\,K_2(x,x')$

$= \varphi_1(x)^T\varphi_1(x')\,\varphi_2(x)^T\varphi_2(x')$ $\quad$ want. $= \varphi_P^T(x)\,\varphi_P(x')$

$$= \left(\varphi_1^1(x)\varphi_1^1(x') + \varphi_1^2(x)\varphi_1^2(x') + \ldots + \varphi_1^{D_1}(x)\varphi_1^{D_1}(x')\right)\left(\varphi_2^1(x)\varphi_2^1(x') + \ldots + \varphi_2^{D_2}(x)\varphi_2^{D_2}(x')\right)$$

assuming $\varphi_1(x) = \begin{bmatrix} \varphi_1^1(x) \\ \vdots \\ \varphi_1^{D_1}(x) \end{bmatrix}$ $\qquad \varphi_2(x) = \begin{bmatrix} \varphi_2^1(x) \\ \vdots \\ \varphi_2^{D_2}(x) \end{bmatrix}$

$$K_P = \varphi_1^1(x)\varphi_1^1(x')\left(\varphi_2^1(x)\varphi_2^1(x') + \varphi_2^2(x)\varphi_2^2(x') + \ldots + \varphi_2^{D_2}(x)\varphi_2^{D_2}(x')\right)$$
$$+ \quad \vdots$$
$$\varphi_1^{D_1}(x)\varphi_1^{D_1}(x')\left(\varphi_2^1(x)\varphi_2^1(x') + \ldots + \varphi_2^{D_2}(x)\varphi_2^{D_2}(x')\right)$$

Clearly we have every comb^n of $\varphi_1^i \varphi_2^j$ $\quad i \in (1, D_1)$ $\; j \in (1, D_2)$

$$\therefore \varphi_P = \varphi_1 \times \varphi_2$$

↳ cartesian product.