Kernel Methods in Machine Learning
Homework Assignment 3
student: Donayorov Maksad
id: 571335

Question 1:

Assuming that kernel matrices are symmetric $K^T = K$ and the given constraints:

$\|K_a \alpha\|_2 = 1 \qquad \alpha^T K_a K_a \alpha - 1 = 0$

$\|K_b \beta\|_2 = 1 \qquad \beta^T K_b K_b \beta - 1 = 0$

we can write the lagrangian as:

$$L = \langle K_a \alpha, K_b \beta \rangle - \frac{1}{2} \cdot p_1 \cdot (\alpha^T \cdot K_a \cdot K_a - 1) - \frac{1}{2} \cdot p_2 \cdot (\alpha^T K_b \cdot K_b - 1)$$

taking the partial derivative we obtain:

$$\frac{\partial L}{\partial \alpha} = K_a K_b \beta - p_1 K_a^2 \alpha = 0 \qquad K_a^T K_b \beta = p_1 K_a^T K_a \alpha$$

$$\frac{\partial L}{\partial \alpha} = K_b K_a \alpha - p_2 K_b^2 \beta = 0 \qquad K_b^T K_a \alpha = p_2 K_b^T K_b \beta$$

applying the constraints we can derive the values of p_1 and p_2:

$$p_1 = \alpha^T K_a K_b \beta$$

$$p_2 = \beta^T K_b K_a \alpha$$

$$\beta^T K_b K_a = (\beta^T K_b K_a)^T = \alpha^T K_a K_b \beta \quad \Rightarrow p_1 = p_2$$

1. $\alpha^T K_a K_a \alpha - 1 = 0$
2. $\beta^T K_b K_b \beta - 1 = 0$

$$\alpha = K_a^{-1} \cdot K_a^{-1} K_a K_b \beta \cdot \frac{1}{p} = \frac{K_a^{-1} K_b \beta}{p}$$

$$\beta \Rightarrow K_a^{-1} \cdot K_a^{-1} K_a K_b \beta \cdot \frac{1}{p} - p K_b^2 \beta = 0$$

Question 2:

$$\widehat{cov}(\langle \phi_a(X_a), w_a\rangle, \langle \phi_b(X_b), w_b\rangle) = \frac{1}{n}\sum_{k=1}^{n} \langle \phi_a(X_a^k), w_a\rangle \langle \phi_b(X_b^k), w_b\rangle$$

$$= \frac{1}{n}\sum_{k=1}^{n} [\phi_a(x_a^k)(w_a^{\parallel} + w_a^{\perp})\cdot \phi_b(x_b^k)(w_b^{\parallel} + w_b^{\perp})]$$

$$= \frac{1}{n}\sum_{k=1}^{n} [(\phi_a(x_a^k) w_a^{\parallel} + \cancel{\phi(x_a^k) w_a^{\perp}})\cdot(\phi_b(x_b^k) w_b^{\parallel} + \cancel{\phi_b(x_b^k) w_b^{\perp}})]$$

$$= \frac{1}{n}\sum_{k=1}^{n} [\phi_a(x_a^k)(\sum_{j=1}^{n} \alpha_j \phi_a(x_a^j))\cdot(\phi_b(x_b^k)\cdot(\sum_{j=1}^{n} \beta_j \phi_b(x_b^j)))]$$

$$= \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{n} [(\alpha_j \phi_a(x_a^k)\phi_a(x_a^j))(\phi_b(x_b^k)\phi_b(x_b^j)\beta_j]$$

$$= \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{n} [\alpha_j K_a(x_a^k, x_a^j) K_b(x_b^k, x_b^j)\beta_j]$$

$$= \boxed{\frac{1}{n}\alpha^T K_a K_b^T \beta}$$