Name: Yuhan Shao
Student Number: 1002281327
UTORid: shaoyuha

CSC411 Assignment 7

1. (a) • Let subspace S be $\text{span}\{\varphi(x^{(i)})\}$, then we have:

$w = w_s + w_\perp$

• Denote w^* as the optimal weight, then we have:

$w^* = w_s^* + w_\perp^*$

• Now, WTS: $w_\perp^* = 0$, so that $w^* = w_s^*$.

i.e: w^* lies in the row space of φ.

• For contradiction, assume $w_\perp^* \neq 0$.

(b) •
$$
\begin{aligned}
J(w) &= \frac{1}{2N}\|t - \varphi w\|^2 + \frac{\lambda}{2}\|w\|^2 \\
&\overset{w=\varphi^T\alpha}{=} \frac{1}{2N}\|t - \varphi(\varphi^T\alpha)\|^2 + \frac{\lambda}{2}\|\varphi^T\alpha\|^2 \\
&\overset{K=\varphi\varphi^T}{=} \frac{1}{2N}(t-K\alpha)^T(t-K\alpha) + \frac{\lambda}{2}(\varphi^T\alpha)^T(\varphi^T\alpha) \\
&= \frac{1}{2N}(t^Tt - t^TK\alpha - \alpha^TK^Tt + \alpha^TK^TK\alpha) + \frac{\lambda}{2}\alpha^TK\alpha \\
&= \frac{1}{2N}(t^Tt - 2t^TK\alpha + \alpha^TK^TK\alpha) + \frac{\lambda}{2}\alpha^TK\alpha
\end{aligned}
$$

• Let $A = \frac{1}{N}(K^TK) + \lambda K$

$b^T = -\frac{1}{N}t^TK$

$b = (b^T)^T = (-\frac{1}{N}t^TK)^T = -\frac{1}{N}K^Tt$.

• Then,
$$
\begin{aligned}
\alpha &= -A^{-1}b \\
&= -[\tfrac{1}{N}(K^TK) + \lambda K]^{-1} \cdot [-\tfrac{1}{N}K^Tt] \\
&= (\tfrac{1}{N}K^TK + \lambda K)^{-1} \cdot \tfrac{1}{N}K^Tt.
\end{aligned}
$$

2. (a) • Suppose: $K_1(x,x') = \varphi_1(x)^T\varphi_1(x')$

$K_2(x,x') = \varphi_2(x)^T\varphi_2(x')$

• Let
$$
\begin{aligned}
K_s(x,x') &= K_1(x,x') + K_2(x,x') \\
&= \varphi_1(x)^T\varphi_1(x') + \varphi_2(x)^T\varphi_2(x')
\end{aligned}
$$

• Let $\varphi_s(x) = \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}$, then $\varphi_s(x)^T = (\varphi_1(x)^T \quad \varphi_2(x)^T)$

then
$$
\begin{aligned}
K_s(x,x') &= \varphi_s(x)^T\varphi_s(x') \\
&= \varphi_1(x)^T\varphi_1(x') + \varphi_2(x)^T\varphi_2(x') \\
&= K_1(x,x') + K_2(x,x')
\end{aligned}
$$

(b) • WTS: $\exists$ kernel K_p s.t. $K_p(x,x') = \varphi_p(x)^T\varphi_p(x')$

• Let $\varphi_p = \begin{bmatrix} \varphi_1^1 & \varphi_2^1 \\ \varphi_1^1 & \varphi_2^2 \\ \vdots & \vdots \\ \varphi_1^1 & \varphi_2^m \\ \varphi_1^1 & \varphi_2^1 \\ \varphi_1^1 & \varphi_2^2 \\ \vdots & \vdots \\ \varphi_1^1 & \varphi_2^m \\ & \vdots & \\ \varphi_1^n & \varphi_2^1 \\ \varphi_1^n & \varphi_2^2 \\ \vdots & \vdots \\ \varphi_1^n & \varphi_2^m \end{bmatrix}$, then
$$
\begin{aligned}
\varphi_p(x)^T\varphi_p(x) &= \textstyle\sum_{i=1}^n\sum_{j=1}^m \varphi_1^i(x)\varphi_2^j(x)\varphi_1^i(x')\varphi_2^j(x') \\
&= [\textstyle\sum_{i=1}^n \varphi_1^i(x)\varphi_1^i(x')][\textstyle\sum_{j=1}^m \varphi_2^j(x)\varphi_2^j(x')] \\
&= \varphi_1^T(x)\varphi_1(x') \cdot \varphi_2^T(x)\varphi_2(x') \\
&= K_1(x,x') \cdot K_2(x,x').
\end{aligned}
$$