Q1

(a) Let S be the row space of Φ

For a arbitrary weigh w,

$w = w_S + w_\perp$

WTS $J(w) \geq J(w_S)$

Case 1:

w is in the row space of Φ,

$\Rightarrow w_\perp$ is $\vec{0}$ $\Rightarrow w = w_S$

$\Rightarrow J(w) = J(w_S)$

Case 2:

w is not in the row space of Φ.

$\exists\, w_\perp,\ w_\perp \neq \vec{0} \wedge w = w_S + w_\perp$

$$J(w) = \frac{1}{N}\sum_{i=1}^{N} L(y^{(i)}, t^{(i)}) + \frac{\lambda}{2}\|w\|^2$$

$$= \frac{1}{N}\sum_{i=1}^{N} L(g(w^T\psi(x)^{(i)}), t^{(i)}) + \frac{\lambda}{2}\|w\|^2$$

$$\ldots^{T} \ldots^{(i)} \ldots w^{T} \psi(x)^{(i)} \ldots)\ t^{(i)}) + \frac{\lambda}{2}\|w\|^2$$

$= \frac{1}{N}\sum_{i=1}^{N} L(g(W_s^T\cdot\psi(x) + W_\perp^T\cdot\psi(x)), ...$

$= \frac{1}{N}\sum_{i=1}^{N} L(g(W_s^T\cdot\psi(x)^{(i)} + 0), t^{(i)}) + \frac{\lambda}{2}\|w\|^2$

Since $w_\perp^T$ is orthogonal to row space of Ψ

$= \frac{1}{N}\sum_{i=1}^{N} L(g(w_s^T\psi(x)^{(i)}), t^{(i)}) + \frac{\lambda}{2}\|w\|^2$

$> \frac{1}{N}\sum_{i=1}^{N} L(g(w_s^T\psi(x)^{(i)}), t^{(i)}) + \frac{\lambda}{2}\|w_s\|^2$

Since $w_\perp$ is not $\vec{0}$

$= J(w_s)$

Therefore, in both cases, $J(w) \geq J(w_s)$

$\Rightarrow$ optimal weights must lie in the row of Ψ.

(b) $$J(w) = \frac{1}{2N}\|t - \Phi w\|^2 + \frac{\lambda}{2}\|w\|^2$$

$$= \frac{1}{2N}\|t - \Phi\Phi^T\alpha\|^2 + \frac{\lambda}{2}\|\Phi^T\alpha\|^2$$

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

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

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

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

$$= \frac{1}{2N}t^Tt - \frac{1}{N}t^T\Phi\Phi^T\alpha + \frac{1}{2N}\alpha^T\Phi\Phi^T\Phi\Phi^T\alpha + \frac{\lambda}{2}\alpha^T\Phi\Phi^T\alpha$$

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

Q2

(a) $k_S(x, x') = k_1(x, x') + k_2(x, x')$

$$= \psi_1(x)^T \psi_1(x') + \psi_2(x)^T \psi_2(x')$$

$$= \left(\psi_1(x)^T, \psi_2(x)^T\right) \begin{pmatrix} \psi_1(x') \\ \psi_2(x') \end{pmatrix}$$

$$= \psi_S(x)^T \psi_S(x')$$

(b) Let $\psi_1(x) = \begin{pmatrix} \psi_{11}(x) \\ \psi_{12}(x) \\ \vdots \\ \psi_{1n}(x) \end{pmatrix}$

$$\psi_2(x) = \begin{pmatrix} \psi_{21}(x) \\ \psi_{22}(x) \\ \vdots \\ \psi_{2m}(x) \end{pmatrix}$$

$$k_p(x,x') = k_1(x,x')\,k_2(x,x')$$

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

$$= \sum_{i=1}^{N} \varphi_{1i}(x)\,\varphi_{1i}(x') \sum_{j=1}^{m} \varphi_{2j}(x)\,\varphi_{2j}(x')$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{m} \varphi_{1i}(x)\,\varphi_{2j}(x)\,\varphi_{1i}(x')\,\varphi_{2j}(x')$$

Let $\varphi_3(x) = \begin{pmatrix} \varphi_{11}(x)\cdot\varphi_{21}(x) \\ \varphi_{11}(x)\cdot\varphi_{22}(x) \\ \vdots \\ \varphi_{11}(x)\cdot\varphi_{2m}(x) \\ \varphi_{12}(x)\cdot\varphi_{21}(x) \\ \varphi_{12}(x)\cdot\varphi_{22}(x) \\ \vdots \\ \varphi_{12}(x)\cdot\varphi_{2m}(x) \\ \vdots \\ \varphi_{1n}(x)\cdot\varphi_{2m}(x) \end{pmatrix}$

$$\Rightarrow k_p(x,x') = \varphi_3(x)^T \cdot \varphi_3(x')$$

$$= \varphi_p(x)^T \cdot \varphi_p(x')$$