$$y_i = \langle \phi(x_i), w \rangle + \varepsilon_i$$

$$= \phi^T(x_i)\, w + \varepsilon_i$$

$$\underset{\in \mathbb{R}^{n\times 1}}{Y} = \begin{bmatrix} \underset{\in \mathbb{R}^{d\times 1}}{\phi^T(x_1)} \\ \vdots \\ \vdots \\ \phi^T(x_n) \end{bmatrix} \underset{\in \mathbb{R}^{d\times 1}}{w} + \varepsilon$$

$$\lambda \|w\|^2 + \| Y - \phi^T(X)\, w \|^2 = f(w)$$

$$\frac{\partial f}{\partial w} = 2\, \phi(X)\phi^T(X)\, w - 2\,\phi(X)\, Y$$

$$w^* = \big(\underset{d\times n}{\phi(X)}\ \underset{n\times d}{\phi^T(X)} + \lambda I_d\big)^{-1} \phi(X)\, Y$$

$$= \phi(X)\big(\phi^T(X)\,\phi(X) + \lambda I_n\big)^{-1} Y$$

$$\langle w^*, \phi(x_q) \rangle$$

$$= Y^T \big(\phi^T(X)\phi(X) + \lambda I\big)^{-1} \phi^T(X)\phi(x_q)$$

$$\left(\lambda I_d + \phi(X)\phi^T(X)\right)^{-1}$$

$$= \frac{1}{\lambda} I_d^{-1} - \frac{1}{\lambda}\phi(X)\left(I + \frac{1}{\lambda}\phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\frac{1}{\lambda}$$

$$= \frac{1}{\lambda}\left[I_d - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\right]$$

$$\left(\phi(X)\phi^T(X) + \lambda I_d\right)^{-1}\phi(X)$$

$$= \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1} Y$$

$$\phi(X)\phi^T(X) = \phi(X^{-i})\phi^T(X^{-i}) + \phi(x_i)\phi^T(x_i)$$

$$\left(\phi(X^{-i})\phi^T(X^{-i}) + \lambda I_d\right)^{-1}$$

$$= \Big(\underbrace{\lambda I_d + \phi(X)\phi^T(X)}_{A} - \underbrace{\phi(x_i)}_{u}\underbrace{\phi^T(x_i)}_{v^T}\Big)^{-1}$$

$$= \left(\lambda I_d + \phi(X)\phi^T(X)\right)^{-1} \quad ①$$

$$+ \frac{A^{-1} u v^T A^{-1}}{1 - \underbrace{\phi^T(x_i)}_{v^T}\underbrace{\left(\lambda I_d + \phi(X)\phi^T(X)\right)^{-1}}_{A}\underbrace{\phi(x_i)}_{u}} \quad ②$$

$\phi(X^{-i})Y^{-i}$

$= \phi(X)Y - y_i \cdot \phi(x_i)$

③ ④

①, ③ trivial

①, ④

$(\lambda I_d + \phi(X)\phi^T(X))^{-1}\phi(X_i)$

$= \frac{1}{\lambda}\left[I_d - \phi(X)(\lambda I_n + \phi^T(X)\phi(X))^{-1}\phi^T(X)\right]\phi(X_i)$

$= \frac{1}{\lambda}\left[\phi(X_i) - \phi(X)(\lambda I_n + \phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(X_i)\right]$

for ②:

$\phi^T(X_i)(\lambda I_d + \phi(X)\phi^T(X))^{-1}\phi(X_i)$

$= \frac{1}{\lambda}\left[\phi^T(X_i)\phi(X_i) - \phi^T(X_i)\phi(X)(\lambda I_n + \phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(X_i)\right]$

②, ③

$A^{-1}\phi(X_i)\phi^T(X_i)A^{-1}\phi(X)Y$

$$= \frac{1}{\lambda}\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]$$

$$\cdot\, \phi^T(x_i)\phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1} Y$$

②④

$$A^{-1}\phi(x_i)\phi^T(x_i)A^{-1}\phi(x_i)$$

$$= \frac{1}{\lambda^2}\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]$$

$$\phi^T(x_i)\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]$$

−①④ + ②③ − ②④

$$= -\frac{y_i}{\lambda}\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]$$

$$+ \frac{1}{1 - \frac{1}{\lambda}\left[\phi^T(x_i)\phi(x_i) - \phi^T(x_i)\phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]}$$

$$\cdot\, \frac{1}{\lambda}\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]$$

$$\cdot \Big\{\phi^T(x_i)\phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1} Y$$

$$- \frac{y_i}{\lambda}\phi^T(x_i)\left[\phi(x_i) - \phi(X)\left(\lambda I_n + \phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]\Big\}$$

$$\Big\{\phi^T(x_i)\phi(X)\left(\lambda I_n+\phi^T(X)\phi(X)\right)^{-1}Y$$

$$-\frac{y_i}{\lambda}\phi^T(x_i)\left[\phi(x_i)-\phi(X)\left(\lambda I_n+\phi^T(X)\phi(X)\right)^{-1}\phi^T(X)\phi(x_i)\right]\Big\}$$

$$=-\frac{y_i}{\lambda}\phi^T(x_i)\phi(x_i)$$

$$+\phi^T(x_i)\phi(X)\left(\lambda I_n+\phi^T(X)\phi(X)\right)^{-1}\left(Y+\frac{y_i}{\lambda}\phi^T(X)\phi(x_i)\right)$$

On a new query x_q:

prediction difference:

$$Y^T(\phi^T(X)\phi(X)+\lambda I)^{-1}\phi^T(X)\phi(x_q)$$

$$-Y^{-i T}(\phi^T(X^{-i})\phi(X^{-i})+\lambda I)^{-1}\phi^T(X^{-i})\phi(x_q)$$

$$=[\phi^T(x_i)\phi(x_q)-\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(x_q)]$$

$$\cdot\Big\{-\frac{y_i}{\lambda}+$$

$$\frac{1}{\lambda-[\phi^T(x_i)\phi(x_i)-\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(x_i)]}$$

$$\cdot\Big[-\frac{y_i}{\lambda}\phi^T(x_i)\phi(x_i)$$

$$+\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}\Big(Y+\frac{y_i}{\lambda}\phi^T(X)\phi(x_i)\Big)\Big]\Big\}$$

$$=[\phi^T(x_i)\phi(x_q)-\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(x_q)]$$

$$\cdot\frac{1}{\lambda-[\phi^T(x_i)\phi(x_i)-\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(x_i)]}$$

$$[\phi^T(x_i)\phi(X)(\lambda I_n+\phi^T(X)\phi(X))^{-1}Y-y_i]$$

$$\lambda - \left[\phi^T(x_i)\phi(x_i) - \phi^T(x_i)\phi(X)(\lambda I_n + \phi^T(X)\phi(X))^{-1}\phi^T(X)\phi(x_i)\right]$$

$$+ \phi^T(x_i)\phi(x_i) - \phi^T(x_i)\phi(X)(\lambda I_n + \phi^T(X)\phi(X))^{-1}$$

$$= \lambda$$