Exercise (10)

Ex.1

$$\min_{w} \sum_{i=1}^{m} (w^T x_i - y_i)^2 + l \|w\|_2^2$$

a) $$\mathcal{E}(w) = \sum_{i=1}^{m} (w^T x_i - y_i)^2 + l \|w\|_2^2$$
$$= (X^T w - y)^T (X^T w - y) + l \|w\|_2^2$$

<u>Note:</u> $X := [x_1, \dots, x_m] \in \mathbb{R}^{d \times m}$
$y := [y_1, \dots, y_m]^T \in \mathbb{R}^m$

$$\frac{\partial \mathcal{E}(w)}{\partial w} = 2X(X^T w - y) + 2lw = 0$$

$$\Rightarrow XX^T w - Xy + lw = 0$$
$$XX^T w + lw = Xy$$
$$(XX^T + l I_d) w = Xy$$
$$w = (XX^T + l I_d)^{-1} X y$$

b)
$x \to \phi(x)$, $\phi : \mathbb{R}^d \to \mathbb{R}^h$ | $h \gg d$
$y = w^T x$ $\quad w \in \mathbb{R}^d$
$y = \beta^T \phi(x)$
$\beta = (\phi(X) \phi(X)^T + l I_h)^{-1} \phi(X) y$

$$\hat{y} = \phi(x_*)^T \beta = \phi(x_*)^T (\phi(X)\phi(X)^T + l I)^{-1} \phi(X) y =$$
$$= \underbrace{\phi(x_*)^T \phi(X)}_{k^*} \underbrace{(\phi(X)^T \phi(X)}_{k} + l I)^{-1} y$$

$$\hat{y} = k^* \underbrace{(k + l I)^{-1} y}_{\alpha} = \sum_{i=1}^{m} k_\phi^*(x_*, x_i) \alpha_i$$

Where:
$k_{ij} = k_\phi^*(x_i, x_j)$, $k_\phi(x_i, x_i) = \langle \phi(x_i), \phi(x_i) \rangle = \phi(x_i)^T \phi(x_i)$

$\boxed{\text{Ex.2}}$

$$\min_{\xi, \omega} \sum_{i=1}^{m} \xi_i^2$$

$$\text{S.T. } \xi_i = \omega^T x_i - Y_i \quad \forall i$$

$$\text{AND } \|\omega\|_2^2 \leq C$$

$$\mathcal{L}(\xi_i, \omega, \beta, l) = \frac{1}{2} \sum_{i=1}^{m} \xi_i^2 + \sum_{i=1}^{m} \beta_i (\xi_i - \omega^T x_i + Y_i) + \frac{l}{2} (\|\omega\|_2^2 - C)$$

$$\max_{\beta, l} \min_{\xi, \omega} \mathcal{L}(\xi_i, \omega, \beta, l)$$

$$\beta_i \in \mathbb{R}$$

$$l \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial \xi_i} = \xi_i + \beta_i = 0 \Rightarrow \xi_i = -\beta_i$$

$$\frac{\partial \mathcal{L}}{\partial \omega} = -\sum_{i=1}^{m} \beta_i x_i + l\omega \Rightarrow \omega = \frac{\sum_{i=1}^{m} \beta_i x_i}{l}$$

$$\mathcal{L}(\beta, l) = \frac{1}{2} \sum_{i=1}^{m} \beta_i^2 - \sum_{i=1}^{m} \beta_i^2 - \sum_{i=1}^{m} \frac{\beta_i \left(\sum_{j=1}^{m} \beta_j x_j\right)^T x_i}{l} +$$

$$+ \sum_{i=1}^{m} \beta_i Y_i + \sum_{ij} \frac{1}{2} \frac{\beta_i \beta_j x_i x_j}{l} - \frac{lC}{2} =$$

$$= -\frac{1}{2} \sum_i \beta_i^2 - \frac{1}{2l} \sum_{ij} \beta_i \beta_j x_i^T x_j + \sum \beta_i Y_i - \frac{lC}{2} =$$

$$= -\frac{1}{2} \beta^T \beta - \frac{1}{2l} \beta^T k \beta + \beta^T Y - \frac{lC}{2}$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = -\beta - \frac{k\beta}{l} + Y = 0 \Rightarrow \left(\frac{k}{l} + I\right) \beta = Y$$

$$\beta = \left(\frac{k}{l} + I\right)^{-1} Y$$

$$w = \frac{\sum_i \beta_i x_i}{l} \qquad \gamma = \frac{\sum_i \beta_i \phi(x_i)}{l}$$

$$\hat{y} = w^T x = \gamma^T \phi(x) = \phi^T(x) \gamma$$

$$= \phi(x)^T . \frac{\phi(X) \beta}{l} = \phi(x) . \phi(X) . (K + lI)^{-1} . Y$$