$$\hat{\beta} = X^T \boxed{(XX^T + I_n)^{-1} y} \quad \text{ⓐ}$$

$$(X^TX + I_p)^{-1} X^T y \quad \text{ⓑ}$$

---

1) A $p \times q$ B $q \times r$

AB : $(AB)_{ij} = A_i B^j$

(q) $(q-1)$

$pr(2q-1) \sim O(pqr)$

2) A $n \times n$ $x : n \times 1$

$y : n \times 1$

$Ax = y$ $O(n^3)$

(a) $X : n \times p$

$XX^T : n \times p \times n + n$

$\boxed{(XX^T + I_n) z = y}$ $O(n^3)$

$+ pn \cdot 1$

$= O(n^3 + pn^2)$

10.1 (a) —

10.1 (b) $f \in \mathcal{H}_K$

$$f(x) = \langle f, K(\cdot, x) \rangle$$

$$K(y, x) = y^T x$$

$$= y^T \left(\sum_{j=1}^{p} (x^T e_j) e_j \right)$$

$$= \sum_{j=1}^{p} (x^T e_j)(y^T e_j)$$

$$= \sum_{j=1}^{p} (x^T e_j) K(y, e_j)$$

$$K(\cdot, x) = \sum_{j=1}^{p} (x^T e_j) K(\cdot, e_j)$$

$$f(x) = \sum_{j=1}^{p} (x^T e_j) \langle f, K(\cdot, e_j) \rangle$$

$$= x^T \left(\sum_{j=1}^{p} \langle f, K(\cdot, e_j) \rangle e_j \right)$$

$$f(x) = x^T \tilde{f}$$

10.1 (c)

$$\mathcal{H}_K := \left\{ f : f(x) = x^T \tilde{f} \text{ where } \tilde{f} \in \mathbb{R}^p \right\}$$

$$\langle K(\cdot, x), K(\cdot, y) \rangle = x^T y$$

$$f \longleftrightarrow \tilde{f}$$

$$L : \mathcal{H}_K \to \mathbb{R}^p \qquad L(f) = \tilde{f}$$

1) L is bijective

2) linear

3) $[L(f)]^T [L(g)] = \langle f, g \rangle$

$$\mathcal{H}_K \qquad \mathbb{R}^p$$

10.2

$$\min_{f \in \mathcal{H}_K} \boxed{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2 + \lambda \|f\|^2_{\mathcal{H}_K}}$$

By the rep'n thm.

$$\boxed{\hat{f}(x) = \sum_{j=1}^{n} \hat{\alpha}_j K(x, x_j)}$$

$$(\hat{\alpha}_j) = \hat{\alpha} = (K + n\lambda I)^{-1} y$$

$$K = [K(x_i, x_j)]_{i,j}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{n}\alpha_j K(x_i, x_j)\right)^2 + \lambda \left\|\sum_{j=1}^{n}\alpha_j K(\cdot, x_j)\right\|^2_{\mathcal{H}}$$

$$\boxed{\sum_{j=1}^{n}\alpha_j K(x_i, x_j) = (K\alpha)_i}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - (K\alpha)_i\right)^2 = \|y - K\alpha\|^2_{\mathbb{R}^n}$$

$$\left\|\sum_{j=1}^{n} \alpha_j K(\cdot, x_j)\right\|_{\mathcal{H}_K}^2$$

$$= \left\langle \sum_{i=1}^{n} \alpha_i K(\cdot, x_i), \sum_{j=1}^{n} \alpha_j K(\cdot, x_j) \right\rangle_{\mathcal{H}_K}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j K(x_i, x_j) = \alpha^T K \alpha$$

$$\hat{\alpha} = \underset{\alpha \in \mathbb{R}^n}{\operatorname{argmin}} \left[\frac{1}{n} \|y - K\alpha\|^2 + \lambda \alpha^T K \alpha \right]$$

$$D = 2\frac{(-1)}{n} K^T (y - K\alpha) + 2\lambda K\alpha = 0$$

$$(K \cdot K + n\lambda K)\alpha + Ky = 0$$

$$\boxed{K^T = K} \qquad K(y - (K + n\lambda I)\alpha) = 0$$

$$\boxed{(K + n\lambda I)\alpha - y \in \ker K}$$

$$a = \hat{a} + h$$

$$\hat{a} = (K + n\lambda I)^{-1} y$$

$$\boxed{h = (K + n\lambda I)^{-1} h'}$$ where $h' \in \operatorname{Ker} K$

$$\hat{f}(x) = \sum_{i=1}^{n} a K(x, x_i)$$

$$= \boxed{\sum_{i=1}^{n} \hat{a} K(x, x_i)} + \sum_{i=1}^{n} h_i K(x, x_i)$$

$$= 0$$

$$\Big\| \sum_{i=1}^{n} h_i K(\cdot, x_i) \Big\|_{\mathcal{H}_K}^2 = \underline{h^T K h} = 0$$

$$h = (K + n\lambda I)^{-1} h'$$

$$= (K + n\lambda I)^{-1} \left[\frac{K h' + n\lambda h'}{n\lambda} \right]$$

$$= (K + n\lambda I)^{-1} \frac{(K + n\lambda I)}{n\lambda} \cdot h'$$

$$= \frac{1}{n\lambda} h' \in \underline{\operatorname{Ker} K}$$

(10.2) (b)

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \frac{1}{n} \|y - X\beta\|^2 + \lambda \|\beta\|_2^2$$

$$\boxed{K(x, y) = x^T y}$$

$$K\left(y - (K + n\lambda I)\alpha\right) = 0$$

$$y - (K + n\lambda I)\alpha \in \ker K$$

$$= h' \in \ker K$$

$$y = (K + n\lambda I)\alpha + h'$$

$$\alpha = (K + n\lambda I)^{-1} y - \underbrace{(K + n\lambda I)^{-1} h'}_{= h}$$

10.3

$$\hat{\beta} = X^T(XX^T + \lambda I_n)^{-1} y$$

$$\hat{\beta} = (X^TX + \lambda I_p)^{-1} X^T y$$

$$X^TXX^T + \lambda X^T$$

$$= X^T(XX^T + \lambda I_n)$$

$$= (X^TX + \lambda I_p) X^T$$

$$X^T(XX^T + \lambda I_n) = (X^TX + \lambda I_p) X^T$$

$(X^TX + \lambda I_p)^{-1}$ from left

$(XX^T + \lambda I_n)^{-1}$ from right

$$\boxed{(X^TX + \lambda I_p)^{-1} X^T = X^T(XX^T + \lambda I_n)^{-1}}$$

$\hat{\beta}$: $X^T(XX^T + n\lambda I_n)^{-1} y$ ✓

$(X^T X + n\lambda I_p)^{-1} X^T y$

---

1) $Ax = y$

$A : n \times n$

$x, y : n \times 1$

$O(n^3)$ $\quad x = A^{-1} y$

---

2) $B \quad p \times q \qquad C \quad q \times r$

$BC : (BC)_{ij} = B_{i\cdot} C^{j}$

$p \times r \qquad q \qquad q-1$

$O(q)$

BC — $O(pqr)$

$X \quad n \times p$

$XX^T : O(n \cdot p \cdot n) = O(n^2 p)$

$XX^T + I_n \qquad + O(n)$

$(XX^T + I_n)^{-1} y : O(n^3)$

$X^T (XX^T + I_n)^{-1} y : O(p \cdot n \cdot 1) = O(pn)$

$X : n \times p$

$XX^T \quad : \quad O(n \cdot p \cdot n)$

$XX^T + I_n \quad O(n)$

$(XX^T + I_n)^{-1} y \quad : \quad O(n^3)$

$X^T [(XX^T + I_n)^{-1} y] \quad : \quad O(p \cdot n \cdot 1)$

$$\boxed{O(n^3 + n^2 p)}$$

$X^T (XX^T + \lambda I_n)^{-1} y$

(10.4) $K(x,y) = \sum \gamma_j \phi_j(x) \phi_j(y)$

$\mathcal{H}_K = \left\{ f = \sum c_i \phi_i(x) \right\}$