

# HW3

2020年11月1日 18:20



HW3(1)

## MATH 3332 Data Analytic Tools Homework 3

Due date: 2 November, 6pm, Monday

1. Determine whether each of the following functions of vectors in  $\mathbb{R}^n$  is linear. If it is a linear function, give its inner product representation, i.e., an vector  $\mathbf{a} \in \mathbb{R}^n$  for which  $f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle$  for all  $\mathbf{x} \in \mathbb{R}^n$ . If it is not linear, give specific  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  and  $\alpha, \beta \in \mathbb{R}$  for which superposition fails, i.e.,  $f(\alpha\mathbf{x} + \beta\mathbf{y}) \neq \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$ .
  - (a) The spread of values of the vector, defined as  $f(\mathbf{x}) = \max_k x_k - \min_k x_k$ .
  - (b) The difference of the last element and the first,  $f(\mathbf{x}) = x_n - x_1$ .
  - (c) The median of vector, where we will assume  $n = 2k + 1$  is odd. The median of the vector  $\mathbf{x}$  is defined as the  $(k + 1)$ -st largest number among the entries of  $\mathbf{x}$ . For example, the median of  $(7.1, 3.2, 1.5)$  is 1.5.
  - (d) Vector extrapolation, defined as  $x_n + (x_n - x_{n-1})$ , for  $n \geq 2$ . (This is a simple prediction of what  $x_{n+1}$  would be, based on a straight line drawn through  $x_n$  and  $x_{n-1}$ .)
2. Let  $V$  be a Hilbert space. Let  $S_1$  and  $S_2$  be two hyperplanes in  $V$  defined by
$$S_1 = \{\mathbf{x} \in V \mid \langle \mathbf{a}_1, \mathbf{x} \rangle = b_1\}, \quad S_2 = \{\mathbf{x} \in V \mid \langle \mathbf{a}_2, \mathbf{x} \rangle = b_2\}.$$
Let  $\mathbf{y} \in V$  be given. We consider the projection of  $\mathbf{y}$  onto  $S_1 \cap S_2$ , i.e., the solution of

$$\min_{\mathbf{x} \in S_1 \cap S_2} \|\mathbf{x} - \mathbf{y}\|. \quad (1)$$

- (a) Prove that  $S_1 \cap S_2$  is a plane, i.e., if  $\mathbf{x}, \mathbf{z} \in S_1 \cap S_2$ , then  $(1+t)\mathbf{z} - t\mathbf{x} \in S_1 \cap S_2$  for any  $t \in \mathbb{R}$ .
  - (b) Prove that  $\mathbf{z}$  is a solution of (1) if and only if  $\mathbf{z} \in S_1 \cap S_2$  and
$$\langle \mathbf{z} - \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle = 0, \quad \forall \mathbf{x} \in S_1 \cap S_2. \quad (2)$$
  - (c) Find an explicit solution of (1).
  - (d) Prove the solution found in part (c) is unique.
3. Let  $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$  be given with  $\mathbf{x}_i \in \mathbb{R}^n$  and  $y_i \in \mathbb{R}$ . Assume  $N < n$ , and  $\mathbf{x}_i$ ,  $i = 1, 2, \dots, N$ , are linearly independent. Consider the ridge regression
$$\min_{\mathbf{a} \in \mathbb{R}^n} \sum_{i=1}^N (\langle \mathbf{a}, \mathbf{x}_i \rangle - y_i)^2 + \lambda \|\mathbf{a}\|_2^2,$$
where  $\lambda \in \mathbb{R}$  is a regularization parameter, and we set the bias  $b = 0$  for simplicity.

- (a) Prove that the solution must be in the form of  $\mathbf{a} = \sum_{i=1}^N c_i \mathbf{x}_i$  for some  $\mathbf{c} = [c_1, c_2, \dots, c_N]^T \in \mathbb{R}^N$ .  
*(Hint: Similar to the proof of the representer theorem.)*

- (b) Re-express the minimization in terms of  $c \in \mathbb{R}^N$ , which has fewer unknowns than the original formulation.
4. (You don't need to do anything.) You can find many resources on Kernel Ridge Regression online.
- A matlab implementation and demonstration:  
<https://www.mathworks.com/matlabcentral/fileexchange/63122-kernel-ridge-regression>
  - Application in business forecasting:  
<http://businessforecastblog.com/kernel-ridge-regression-example-spreadsheet/>

2

- Determine whether each of the following functions of vectors in  $\mathbb{R}^n$  is linear. If it is a linear function, give its inner product representation, i.e., a vector  $a \in \mathbb{R}^n$  for which  $f(x) = \langle a, x \rangle$  for all  $x \in \mathbb{R}^n$ . If it is not linear, give specific  $x, y \in \mathbb{R}^n$  and  $\alpha, \beta \in \mathbb{R}$  for which superposition fails, i.e.,  $f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y)$ .
  - The spread of values of the vector, defined as  $f(x) = \max_k x_k - \min_k x_k$ .
  - The difference of the last element and the first,  $f(x) = x_n - x_1$ .
  - The median of a vector, where we will assume  $n = 2k + 1$  is odd. The median of the vector  $x$  is defined as the  $(k + 1)$ -st largest number among the entries of  $x$ . For example, the median of  $(7.1, 3.2, 1.5)$  is 1.5.
  - Vector extrapolation, defined as  $x_n + (x_n - x_{n-1})$ , for  $n \geq 2$ . (This is a simple prediction of what  $x_{n+1}$  would be, based on a straight line drawn through  $x_n$  and  $x_{n-1}$ .)

$$(a). \quad f(x) = \max_k x_k - \min_k x_k$$

$$\text{Let } x = (\underbrace{1, 0, 0, \dots}_{\text{all terms are 0}}), y = (\underbrace{0, 1, 0, 0, \dots}_{\text{all terms are 0}}) \in \mathbb{R}^n$$

$$\text{So } f(x) = 1 - 0 = 1, \quad f(y) = 1 - 0 = 1$$

$$\text{Now } \alpha x + \beta y = \alpha(\underbrace{1, 0, 0, \dots}_{\text{all terms are 0}}) + \beta(\underbrace{0, 1, 0, 0, \dots}_{\text{all terms are 0}}) = (1, \beta, 0, 0, \dots)$$

$$\text{Now } \alpha x + \beta y = \underbrace{\alpha(1, 0, 0, \dots)}_{\substack{\text{all terms} \\ \text{are 0}}} + \underbrace{\beta(0, 1, 0, \dots)}_{\substack{\text{all terms are 0}}} = (1, \alpha, 0, \dots)$$

$$f(\alpha x + \beta y) = f(\underbrace{\alpha, \beta, 0, \dots}_{\text{all terms}})$$

and now let  $\alpha = 1$  &  $\beta = 1$

$$\therefore f(\alpha x + \beta y) = f(1, 1, 0, \dots) = 1$$

$$\alpha f(x) = \alpha, \beta f(y) = \beta, \alpha f(x) + \beta f(y) = \alpha + \beta = 2$$

$$\therefore f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y)$$

$\Rightarrow f$  is not linear.

b).  $f(x) = x_n - x_1$

$$\text{Let } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$f(x) = x_n - x_1, f(y) = y_n - y_1$$

$$\alpha x + \beta y = \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{pmatrix}$$

$$f(\alpha x + \beta y) = (\alpha x_n + \beta y_n) - (\alpha x_1 + \beta y_1)$$

$$f(\alpha x + \beta y) = \alpha(x_n - x_1) + \beta(y_n - y_1)$$

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$$\Rightarrow f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$\therefore$  we can find a vector  $a \in \mathbb{R}^n$ , for which  $f(x) = \langle a, x \rangle$ ,  $\forall x \in \mathbb{R}^n$ .

$$a = (a_1, a_2, \dots, a_n) \& x = (x_1, x_2, \dots, x_n)$$

$$f(x) = x_n - x_1$$

$$f(x) = (-1, \underbrace{0, 0, \dots, 0}_{\substack{\text{all terms} \\ \text{are 0}}}, 1)^T (x_1, x_2, \dots, x_n)$$

$$\therefore f(x) = \langle a, x \rangle$$

where  $a = \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^n$ .

c). Let  $x = (1, 2, 6, 8, 3)$   
 $y = (2, 3, 1, 5, 1)$

$$f(x) = 3, f(y) = 3.$$

$$\therefore a = 2 \quad n=5 \quad a=3$$

$$f(x) = 3, f(y) = 3.$$

Now let  $\alpha = 3$  and  $\beta = 2$ .

$$\alpha x = (3, 6, 18, 24, 9)$$

$$\beta y = (4, 6, 20, 10, 2)$$

$$\alpha x + \beta y = (7, 12, 38, 34, 11)$$

$$f(\alpha x + \beta y) = 12$$

$$3f(x) + 2f(y) = 3(3) + 2(3) = 15 \neq 12$$

$$\therefore f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y)$$

$\Rightarrow f$  is not linear.

Q1.  $f(x) = x_n + (x_n - x_{n-1})$  for  $n \geq 2$ .

$$\text{Let } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$f(x) = x_n + (x_n - x_{n-1})$$

$$f(y) = y_n + (y_n - y_{n-1})$$

$$\alpha x + \beta y = \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{pmatrix}$$

$$f(\alpha x + \beta y) = (\alpha x_n + \beta y_n) + (\alpha x_n + \beta y_n - \alpha x_{n-1} - \beta y_{n-1})$$

$$\text{Now } \alpha f(x) = \alpha x_n + \alpha (x_n - x_{n-1})$$

$$\beta f(y) = \beta y_n + \beta (y_n - y_{n-1})$$

$$\alpha f(x) + \beta f(y) = \alpha x_n + \beta y_n + \alpha (x_n - x_{n-1}) + \beta (y_n - y_{n-1})$$

$$\alpha f(x) + \beta f(y) = (\alpha x_n + \beta y_n) + (\alpha x_n + \beta y_n - \alpha x_{n-1} - \beta y_{n-1})$$

$$\alpha f(x) + \beta f(y) = f(\alpha x + \beta y)$$

$\Rightarrow f$  is linear

And we can find a vector  $a \in \mathbb{R}^n$ .

for which  $f(x) = \langle a, x \rangle \quad \forall x \in \mathbb{R}^n$ .

$$\Rightarrow f(x) = y_n + (x_n - x_{n-1})$$

$$f(x) = \alpha x_n - x_{n-1}$$

$$f(x) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} (x_1, x_2, \dots, x_n)$$

$$tx = \left( \begin{array}{c} a_1 \\ \vdots \\ a_n \end{array} \right) (x_1, x_2, \dots, x_n)$$

$$\therefore a = \left( \begin{array}{c} a_1 \\ \vdots \\ a_n \end{array} \right) \in \mathbb{R}^n.$$

2. Let  $V$  be a Hilbert space. Let  $S_1$  and  $S_2$  be two hyperplanes in  $V$  defined by

$$S_1 = \{x \in V \mid \langle a_1, x \rangle = b_1\}, \quad S_2 = \{x \in V \mid \langle a_2, x \rangle = b_2\}.$$

Let  $y \in V$  be given. We consider the projection of  $y$  onto  $S_1 \cap S_2$ , i.e., the solution of

$$\min_{x \in S_1 \cap S_2} \|x - y\|. \quad (1)$$

(a) Prove that  $S_1 \cap S_2$  is a plane, i.e., if  $x, z \in S_1 \cap S_2$ , then  $(1+t)z - tx \in S_1 \cap S_2$  for any  $t \in \mathbb{R}$ .

(b) Prove that  $z$  is a solution of (1) if and only if  $z \in S_1 \cap S_2$  and

$$\langle z - y, z - x \rangle = 0, \quad \forall x \in S_1 \cap S_2. \quad (2)$$

(c) Find an explicit solution of (1).

(d) Prove the solution found in part (c) is unique.

a). Consider the set  $S_1 \cap S_2 = \{x \in V \mid \langle a_1, x \rangle = b_1 \text{ and } \langle a_2, x \rangle = b_2\}$

Then,  $\forall x, z \in S_1 \cap S_2$ ,

$$\langle a_1, x \rangle = b_1, \quad \langle a_2, x \rangle = b_2$$

$$\langle a_1, z \rangle = b_1, \quad \langle a_2, z \rangle = b_2$$

$$\langle a_1, x-z \rangle = 0, \quad \langle a_2, x-z \rangle = 0$$

Now, let  $t \in \mathbb{R}$ ,

Consider

$$\langle a_1, z \rangle + t \langle a_1, z-x \rangle$$

$$= b_1 + t(0)$$

$$= b_1$$

$$\text{and } \langle a_1, z \rangle + t \langle a_1, z-x \rangle$$

$$= \langle a_1, z + t(z-x) \rangle$$

$$= \langle a_1, (1+t)z - tx \rangle = b_1$$

$$\text{Consider } \langle a_2, z \rangle + t \langle a_2, z-x \rangle$$

$$= b_2$$

$$\text{and } \langle a_2, z \rangle + t \langle a_2, z-x \rangle$$

$$= \langle a_2, z + t(z-x) \rangle$$

$$= \langle a_2, (1+t)z - tx \rangle = b_2$$

$\therefore \forall x, z \in S_1 \cap S_2, (1+t)z - tx \in S_1 \cap S_2 \quad \forall t \in \mathbb{R}$ .

b). Let  $P_K y$  be the solution of (1),

where  $K = S_1 \cap S_2$ .

First, we prove that if  $z \in K$  is a

First, we prove that if  $z \in K$  is a solution of  $\min_{x \in K} \|x - y\|^2$ , then  $\langle z - y, z - x \rangle = 0$

$\forall x \in K$

Since  $z$  is a solution,  $z \in K$ , i.e.  $\langle a_1, z \rangle = b_1$ ,  
 $\langle a_2, z \rangle = b_2$ .

$\forall x \in K$  and  $t \in \mathbb{R}$ ,

$$\begin{aligned}\langle a_1, (1+t)z - tx \rangle &= b_1 \\ \langle a_2, (1+t)z - tx \rangle &= b_2 \quad (\text{proved in (a)})\end{aligned}$$

Therefore,  $(1+t)z - tx \in K$ .

Since  $z$  is closest to  $y$  on  $K$ , we have

$$\begin{aligned}\|z - y\|^2 &\leq \|(1+t)z - tx - y\|^2 \\ &= \|(z - y) + t(z - x)\|^2 \\ &= \|z - y\|^2 + t^2\|z - x\|^2 + 2t\langle z - y, z - x \rangle\end{aligned}$$

$$\text{i.e. } t\langle z - y, z - x \rangle \geq -\frac{t^2}{2}\|z - x\|^2$$

if we choose  $t > 0$ ,

$$\langle z - y, z - x \rangle \geq -\frac{t^2}{2}\|z - x\|^2$$

$$\text{let } t \rightarrow 0^+, \text{ then } \langle z - y, z - x \rangle \geq 0$$

if we choose  $t < 0$ ,

$$\langle z - y, z - x \rangle \leq -\frac{t^2}{2}\|z - x\|^2$$

$$\text{let } t \rightarrow 0^-, \text{ gives } \langle z - y, z - x \rangle \geq 0$$

$\therefore \forall t \in \mathbb{R}, \langle z - y, z - x \rangle \text{ can only be } 0$ .

$z$  satisfies  $\langle z - y, z - x \rangle = 0 \quad \forall x \in K$ .

② We show that  $z \in K$  satisfies  $\langle z - y, z - x \rangle = 0$ ,

then  $z$  is a solution of  $\min_{x \in K} \|x - y\|^2$ .

$\therefore \langle z - y, z - x \rangle = 0 \quad \forall x \in K$ .

$$\therefore \langle z-y, z-x \rangle = 0 \quad \forall x \in K,$$

$$\|x-y\|^2 = \| (z-x) - (z-y) \|^2$$

$$\|x-y\|^2 = \|z-x\|^2 + \|z-y\|^2 - 2 \langle z-x, z-y \rangle$$

$$\|x-y\|^2 = \|z-y\|^2 + \|z-x\|^2 \geq \|z-y\|^2$$

$$\forall x \in K$$

this, together with  $z \in K$ , implies

$z$  minimizes  $\|x-y\|^2$  in  $x \in K$ .

7.  $z$  should satisfy  $\begin{cases} \langle a_1, z \rangle = b_1 \\ \langle a_2, z \rangle = b_2 \end{cases}$  from (b), we know that  $z$  is in the form of  $z = y - c_1 a_1 - c_2 a_2$ , and  $z \in S_1 \cap S_2$ , where  $c_1, c_2 \in \mathbb{R}$ .

$$\langle a_1, z \rangle = \langle a_1, y \rangle - \langle a_1, c_1 a_1 - c_2 a_2 \rangle$$

$$\langle a_1, z \rangle = \langle a_1, y \rangle - \langle a_1, c_1 a_1 \rangle + \langle a_1, c_2 a_2 \rangle$$

$$b_1 = \langle a_1, y \rangle - c_1 \langle a_1, a_1 \rangle + c_2 \langle a_1, a_2 \rangle - (1)$$

$$\langle a_2, z \rangle = \langle a_2, y \rangle - \langle a_2, c_1 a_1 - c_2 a_2 \rangle$$

$$b_2 = \langle a_2, y \rangle - c_1 \langle a_2, a_1 \rangle + c_2 \langle a_2, a_2 \rangle - (2)$$

$$\text{From (1), } c_1 \langle a_1, a_1 \rangle + c_2 \langle a_1, a_2 \rangle = \langle a_1, y \rangle - b_1$$

$$\text{From (2), } c_1 \langle a_2, a_1 \rangle + c_2 \langle a_2, a_2 \rangle = \langle a_2, y \rangle - b_2$$

$$\begin{bmatrix} \|a_1\|^2 & \langle a_1, a_2 \rangle \\ \langle a_2, a_1 \rangle & \|a_2\|^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \langle a_1, y \rangle - b_1 \\ \langle a_2, y \rangle - b_2 \end{bmatrix}$$

$$A \quad X = b$$

$$\text{Let } A = \begin{bmatrix} \|a_1\|^2 & \langle a_1, a_2 \rangle \\ \langle a_2, a_1 \rangle & \|a_2\|^2 \end{bmatrix}$$

$$\text{Let } B = \begin{bmatrix} \langle a_1, y \rangle - b_1 \\ \langle a_2, y \rangle - b_2 \end{bmatrix}$$

For solving  $x$ , we can have  $x = A^{-1}B$ .

$$\therefore \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{\|a_1\|^2 \|a_2\|^2 - \langle a_1, a_2 \rangle^2} \begin{bmatrix} \|a_2\|^2 - \langle a_1, a_2 \rangle \\ -\langle a_2, a_1 \rangle & \|a_1\|^2 \end{bmatrix} \begin{bmatrix} \langle a_1, y \rangle - b_1 \\ \langle a_2, y \rangle - b_2 \end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{\|a_1\|^2 \|a_2\|^2 - \langle a_1, a_2 \rangle^2} \begin{bmatrix} \|a_2\|^2 \langle a_1, y \rangle - \|a_2\|^2 b_1 - \langle a_1, a_2 \rangle \langle a_2, y \rangle + b_2 \langle a_1, a_2 \rangle \\ b_1 \langle a_2, a_1 \rangle - \langle a_1, a_2 \rangle \langle a_1, y \rangle + \|a_1\|^2 \langle a_2, y \rangle - b_2 \|a_1\|^2 \end{bmatrix}$$

$$\therefore z = y - c_1 a_1 - c_2 a_2 \quad \text{where}$$

$$c_1 = \frac{\|a_2\|^2 \langle a_1, y \rangle - \|a_2\|^2 b_1 - \langle a_1, a_2 \rangle \langle a_2, y \rangle + b_2 \langle a_1, a_2 \rangle}{\|a_1\|^2 \|a_2\|^2 - \langle a_1, a_2 \rangle^2}$$

$$c_2 = \frac{\|a_1\|^2 \langle a_2, y \rangle - \|a_1\|^2 b_2 - \langle a_1, a_2 \rangle \langle a_1, y \rangle + b_1 \langle a_1, a_2 \rangle}{\|a_1\|^2 \|a_2\|^2 - \langle a_1, a_2 \rangle^2}$$

1). Suppose we have two solution  $z_1$  and  $z_2$ .

$z_1$  is a solution,  $\Rightarrow \langle z_1 - y, z_1 - z_2 \rangle = 0$

$z_2$  is a solution,  $\Rightarrow \langle z_2 - y, z_2 - z_1 \rangle = 0$

Taking the difference leads to  $\langle z_1 - z_2, z_1 - z_2 \rangle = 0$

$$\|z_1 - z_2\|^2 = 0 \Rightarrow z_1 = z_2$$

3. (a) Prove that the solution must be in the form of  $a = \sum_{i=1}^N c_i x_i$  for some  $c = [c_1, c_2, \dots, c_N]^T \in \mathbb{R}^N$ .  
*(Hint: Similar to the proof of the representer theorem.)*

Given any vector  $a \in \mathbb{R}^n$ , and  $x_i, i=1, 2, \dots, N$

Given any vector  $a \in \mathbb{R}^n$ , and  $x_i, i=1, 2, \dots, N$

By solving the system

$$\begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \dots & \langle x_1, x_N \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \dots & \langle x_2, x_N \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_N, x_1 \rangle & \langle x_N, x_2 \rangle & \dots & \langle x_N, x_N \rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} \langle a, x_1 \rangle \\ \langle a, x_2 \rangle \\ \vdots \\ \langle a, x_N \rangle \end{bmatrix}$$

The solution is always unique because the vectors are linearly independent.

Let residual be  $b_s = a - \sum_{i=1}^N c_i x_i$  where  $b_s \in \mathbb{R}^n$ .

$$\Rightarrow a = \sum_{i=1}^N c_i x_i + b_s$$

for some  $c_i \in \mathbb{R}$  and residual  $b_s \in \mathbb{R}^n$  where  $\langle b_s, x_i \rangle = 0, \forall i$ .

$$\text{def } f(a) = \min_{a \in \mathbb{R}^n} \sum_{i=1}^N (\langle a, x_i \rangle - y_i)^2 + \lambda \|a\|_2^2$$

$$f(c, b_s) = \min_{\substack{c \in \mathbb{R}^N \\ b_s \in \mathbb{R}^n}} \sum_{i=1}^N \left( \left\langle \sum_{j=1}^N c_j x_j + b_s, x_i \right\rangle - y_i \right)^2 + \lambda \left\langle \sum_{j=1}^N c_j x_j + b_s, \sum_{j=1}^N c_j x_j + b_s \right\rangle$$

$$f(c, b_s) = \min_{\substack{c \in \mathbb{R}^N \\ b_s \in \mathbb{R}^n}} \sum_{i=1}^N \left( \sum_{j=1}^N c_j \langle x_j, x_i \rangle + \langle b_s, x_i \rangle - y_i \right)^2 + \lambda \left( \sum_{i=1}^N \sum_{j=1}^N c_i c_j \langle x_i, x_j \rangle + 2 \sum_{i=1}^N \langle b_s, x_i \rangle + \langle b_s, b_s \rangle \right)$$

$$\because \langle b_s, x_i \rangle = 0$$

$$f(c, b_s) = \min_{c \in \mathbb{R}^N} \sum_{i=1}^N \left( \sum_{j=1}^N c_j \langle x_j, x_i \rangle - y_i \right)^2 + \lambda \sum_{i=1}^N \sum_{j=1}^N c_i c_j \langle x_i, x_j \rangle + \lambda \langle b_s, b_s \rangle$$

$$f(c, bs) = \min_{c \in \mathbb{R}^n, bs \in \mathbb{R}^n} \sum_{i=1}^N \left( \sum_{j=1}^n c_j \langle x_j, x_i \rangle - y_i \right)^2 + \lambda \sum_{i=1}^N \sum_{j=1}^n b_i b_j \langle x_i, x_j \rangle + \lambda \langle bs, bs \rangle$$

let  $K = \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \dots & \langle x_1, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \dots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \dots & \langle x_n, x_n \rangle \end{bmatrix}$

$$f(c, bs) = \min_{c \in \mathbb{R}^n, bs \in \mathbb{R}^n} \|Kc - y\|^2 + \lambda c^T K c + \lambda \|bs\|_2^2$$

$\underbrace{\quad}_{\text{depends on } c \text{ only}}$        $\underbrace{\quad}_{\text{depends on } bs \text{ only.}}$

$\Rightarrow$  split into two minimization problems:

$$g(c) = \min_{c \in \mathbb{R}^n} \|Kc - y\|^2 + \lambda c^T K c \quad \text{and} \quad h(bs) = \min_{bs \in \mathbb{R}^n} \lambda \|bs\|_2^2$$

For  $h(bs)$ ,

$$\min_{bs \in \mathbb{R}^n} \lambda \|bs\|_2^2 = bs = 0$$

$$\text{so, } f(a) = \min_{a \in \mathbb{R}^n} \sum_{i=1}^N (\langle a, x_i \rangle - y_i)^2 + \lambda \|a\|_2^2$$

must be in the form of

$$a = \sum_{i=1}^n c_i x_i$$

b). Form  $(a)$ ,  $f(a) = g(c)$ ,

$$f(a) = g(c) = \min_{c \in \mathbb{R}^n} \|Kc - y\|^2 + \lambda c^T K c$$

and  $a$  can be obtained by  $a = \sum_{i=1}^n c_i x_i$