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 $\boldsymbol{a} \in \mathbb{R}^n$ for which $f(x) = \langle \boldsymbol{a}, \boldsymbol{x} \rangle$ for all $\boldsymbol{x} \in \mathbb{R}^n$. If it is not linear, give specific $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ for which superposition fails, i.e., $f(\alpha \boldsymbol{x} + \beta \boldsymbol{x}) \neq \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$.
   (a) The spread of values of the vector, defined as $f(\boldsymbol{x}) = \max_k x_k - \min_k x_k$.
   (b) The difference of the last element and the first, $f(\boldsymbol{x}) = x_n - x_1$.
   (c) The median of a vector, where we will assume $n = 2k+1$ is odd. The median of the vector $\boldsymbol{x}$ is defined as the $(k+1)$-st largest number among the entries of $\boldsymbol{x}$. 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
$$S_1 = \{\boldsymbol{x} \in V \mid \langle \boldsymbol{a}_1, \boldsymbol{x} \rangle = b_1\}, \quad S_2 = \{\boldsymbol{x} \in V \mid \langle \boldsymbol{a}_2, \boldsymbol{x} \rangle = b_2\}.$$
   Let $\boldsymbol{y} \in V$ be given. We consider the projection of $\boldsymbol{y}$ onto $S_1 \cap S_2$, i.e., the solution of
$$\min_{\boldsymbol{x} \in S_1 \cap S_2} \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{1}$$
   (a) Prove that $S_1 \cap S_2$ is a plane, i.e., if $\boldsymbol{x}, \boldsymbol{z} \in S_1 \cap S_2$, then $(1+t)\boldsymbol{z} - t\boldsymbol{x} \in S_1 \cap S_2$ for any $t \in \mathbb{R}$.
   (b) Prove that $\boldsymbol{z}$ is a solution of (1) if and only if $\boldsymbol{z} \in S_1 \cap S_2$ and
$$\langle \boldsymbol{z} - \boldsymbol{y}, \boldsymbol{z} - \boldsymbol{x} \rangle = 0, \quad \forall \boldsymbol{x} \in S_1 \cap S_2. \tag{2}$$
   (c) Find an explicit solution of (1).
   (d) Prove the solution found in part (c) is unique.

3. Let $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^N$ be given with $\boldsymbol{x}_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$. Assume $N < n$, and $\boldsymbol{x}_i$, $i = 1, 2, \ldots, N$, are linearly independent. Consider the ridge regression
$$\min_{\boldsymbol{a} \in \mathbb{R}^n} \sum_{i=1}^N (\langle \boldsymbol{a}, \boldsymbol{x}_i \rangle - y_i)^2 + \lambda \|\boldsymbol{a}\|_2^2,$$
   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 $\boldsymbol{a} = \sum_{i=1}^N c_i \boldsymbol{x}_i$ for some $\boldsymbol{c} = [c_1, c_2, \ldots, c_N]^T \in \mathbb{R}^N$. (*Hint: Similar to the proof of the representer theorem.*)

(b) Re-express the minimization in terms of $\boldsymbol{c} \in \mathbb{R}^N$, 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 demonstraction:
     `https://www.mathworks.com/matlabcentral/fileexchange/63122-kernel-ridge-regression`
   - Application in business forecasting:
     `http://businessforecastblog.com/kernel-ridge-regression-example-spreadsheet/`

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 $\boldsymbol{a} \in \mathbb{R}^n$ for which $f(x) = \langle \boldsymbol{a}, \boldsymbol{x} \rangle$ for all $\boldsymbol{x} \in \mathbb{R}^n$. If it is not linear, give specific $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ for which superposition fails, i.e., $f(\alpha \boldsymbol{x} + \beta \boldsymbol{x}) \neq \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$.
   (a) The spread of values of the vector, defined as $f(\boldsymbol{x}) = \max_k x_k - \min_k x_k$.
   (b) The difference of the last element and the first, $f(\boldsymbol{x}) = x_n - x_1$.
   (c) The median of a vector, where we will assume $n = 2k + 1$ is odd. The median of the vector $\boldsymbol{x}$ is defined as the $(k+1)$-st largest number among the entries of $\boldsymbol{x}$. 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}.)

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

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

So $f(x) = 1 - 0 = 1$, $f(y) = 1 - 0 = 1$

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

all terms are 0 all terms are 0

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

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

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

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

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

$\Rightarrow f$ is not linear.

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

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

$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ & $x = (x_1, x_2, \cdots, x_n)$

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

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

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

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

c). let $x = (1, 2, 6, 8, 3)$

$y = (2, 3, 10, 5, 1)$

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

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

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

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

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

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

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

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

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

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

$\Rightarrow f$ is not linear.

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

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

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

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

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

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

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

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

$f(x) = 2x_n - x_{n-1}$

$f(x) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -1 \\ 2 \end{pmatrix} (x_1\ x_2 \cdots x_n)$

$\therefore a = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -1 \\ 2 \end{pmatrix} \in \mathbb{R}^n$.

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

$$\min_{x \in S_1 \cap S_2} \|x - y\|. \tag{1}$$

(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

$$\langle z - y, z - x\rangle = 0, \quad \forall x \in S_1 \cap S_2. \tag{2}$$

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

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

2a). 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, $\forall x, z \in S_1 \cap S_2$,

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

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

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

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

Consider

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

$= b_1 + t(0)$

$= b_1$

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

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

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

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

$= b_2$

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

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

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

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

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 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 $t \in \mathbb{R}$,

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

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

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

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

Since z is closest to y on K, we have

$$\|z-y\|^2 \le \|(1+t)z - tx - y\|^2$$
$$= \|(z-y) + t(z-x)\|^2$$
$$= \|z-y\|^2 + t^2\|z-x\|^2 + 2t\langle z-y, z-x\rangle$$

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

if we choose $t>0$,

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

let $t \to 0^+$ gives $\langle z-y, z-x\rangle \geq 0$

if we choose $t<0$,

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

let $t \to 0$, gives $\langle z-y, z-x\rangle \geq 0$

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

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

② We show that $z \in K$ satisfies $\langle z-y, z-x\rangle = 0$, then z is a solution of $\min_{x \in K} \|x-y\|^2$.

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

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

$\forall x \in K$

This, together with $z \in K$, implies
z minimizes $\|x - y\|^2$ in $x \in K$.

c).

$\langle z - y, z - x \rangle = 0$

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 $c_1, c_2 \in \mathbb{R}$.

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

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

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

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

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

$$\langle z - y, z - x \rangle = 0$$

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

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

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

$$A \qquad X = b$$

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

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 $X = A^{-1} B$.

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

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

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

$\therefore z = y - c_1 a_1 - c_2 a_2$ where

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

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

c). Using the technique from Gram-Schmidt process,

$$v_1 = a_1$$

$$v_2 = a_2 - \frac{a_2 \cdot v_1}{v_1 \cdot v_1} v_1 = a_2 - \frac{\langle a_2, a_1\rangle}{\|a_1\|^2} a_1, \quad a_2 = v_2 + \frac{\langle a_2, a_1\rangle}{\|a_1\|^2} a_1$$

$$v_3 = z = y - \frac{\langle a_1, y\rangle - b_1}{v_1 \cdot v_1} v_1 - \frac{\langle a_2, y\rangle - \frac{\langle a_1, y\rangle - b_1}{v_1 \cdot v_1}\langle a_1, a_2\rangle - b_2}{v_2 \cdot v_2} v_2,$$

$$y \in S_1 \cap S_2.$$

$$\therefore z = y - \frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2} v_1 - \frac{\langle a_2, y\rangle - \frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2} v_2$$

d). Suppose we have two solution z_1 and z_2.

Then, 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 $\langle z_1 - z_2, z_1 - z_2 \rangle = 0$

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

$\therefore$ z is unique

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

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

By solving the system

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

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 $b_s \in \mathbb{R}^n$.

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

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

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

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

$$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

$$+2\sum_{i=1}^{N}\langle b_s, x_i\rangle + \langle b_s, b_s\rangle\Big)$$

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

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

Let $K = \begin{bmatrix} \langle x_1, x_1\rangle & \langle x_1, x_2\rangle & \cdots & \langle x_1, x_N\rangle \\ \langle x_2, x_1\rangle & \langle x_2, x_2\rangle & \cdots & \langle x_2, x_N\rangle \\ \vdots & \vdots & & \vdots \\ \langle x_N, x_1\rangle & \langle x_N, x_2\rangle & \cdots & \langle x_N, x_N\rangle \end{bmatrix}$

$$f(c, b_s) = \min_{c\in\mathbb{R}^N, b_s\in\mathbb{R}^n} \underbrace{\|Kc - y\|^2 + \lambda c^T K c}_{\text{depends on } c \text{ only}} + \underbrace{\lambda\|b_s\|_2^2}_{\text{depends on } b_s \text{ only.}}$$

$\Rightarrow$ split into two minimization problems:

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

For $h(b_s)$,

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

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). From (a), $f(a) = g(c)$,

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

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

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