quiz3

2020年11月5日 19:05

1 a). Consider the set $S_1 \cap S_2 = \{x \in V \mid \langle a_1, x\rangle = b_1$ 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 $t \in \mathbb{R}$,

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

$\langle a_2, (1+t)z - tx\rangle = b_2$ (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. $t\langle z-y, z-x\rangle \geqslant -\frac{t^2}{2}\|z-x\|^2$

If we choose $t > 0$,

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

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

If we choose $t < 0$,

as $t \to 0$ ...

If we choose $t<0$,

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

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

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

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

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

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

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

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

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

$\forall x \in K$

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

c). Using the technique from Gram-Schmidt process

$v_1 = a_1$

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

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

$y \in S_1 \cap S_2$.

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

① $\langle a_1, z\rangle = \langle a_1, y\rangle - \left(\frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\right)\langle a_1, v_1\rangle -$

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

$= b_1 - (\langle a_1, y\rangle - b_1) - \underbrace{\frac{b_2 - \frac{b_1 - b_1}{\|a_1\|_2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2} v_2}_{0}$

$= b_1 - \langle a_1, y\rangle + b_1 = b_1 \quad \therefore z \in S_1$

② $\langle a_2, z\rangle = \langle a_2, y\rangle - \left(\frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\right)(\langle a_2, v_1\rangle)$

$- \left(\frac{\langle a_2, y\rangle - \frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2}\right)(\langle a_2, v_1\rangle)$

$= b_2 - (\cancel{b_1} - b_1)\langle a_1, a_2\rangle - \left(b_2 - \frac{b_1 - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2\right)\langle v_2 + \frac{\langle a_2, a_1\rangle}{\|a_1\|^2} a_1, v_2\rangle$

$$\ldots \frac{\ldots}{\|v_2\|^2}\Big)(\langle a_2, v_1\rangle)$$

$$= b_2 - \Big(\frac{\cancel{b_1 - b_1}^{\,0}}{\|a_1\|^2}\Big)\langle a_2, a_1\rangle - \Big(\frac{b_2 - \frac{b_1 - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2}\Big)\Big\langle v_2 + \frac{\langle a_2, a_1\rangle}{\|a_1\|^2} a_1, v_2\Big\rangle$$

$$= b_2 - \Big(\frac{b_2 - b_2}{\|v_2\|^2}\Big)\Big(\underbrace{\langle v_2, v_2\rangle}_{\|v_2\|^2} + \frac{\langle a_2, a_1\rangle}{\|a_1\|^2}\cancel{\langle a_1, v_2\rangle}\Big)$$

$a_1 = v_1 \perp v_2$, $\therefore \langle a_1, v_2\rangle = 0$

$$= b_2 - (b_2 - b_2) = b_2$$

$\therefore z \in S_2$

From ①, ②, $z \in S_1 \cap S_2$.

③ $$\langle z - y, z - x\rangle = \Big\langle -\Big(\frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\Big)v_1 - \Big(\frac{\langle a_2, y\rangle - \frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2}\Big)v_2,\ z - x\Big\rangle$$

$$= -\Big(\frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\Big)\langle a_1, z - x\rangle - \Big(\frac{\langle a_2, y\rangle - \frac{\langle a_1, y\rangle - b_1}{\|a_1\|^2}\langle a_1, a_2\rangle - b_2}{\|v_2\|^2}\Big)\langle v_2, z - x\rangle$$

Here, $\langle a_1, z - x\rangle = \langle a_1, z\rangle - \langle a_1, x\rangle = b_1 - b_1 = 0$

And, $\langle v_2, z - x\rangle = \langle a_2, z - x\rangle - \Big\langle \frac{\langle a_2, a_1\rangle}{\|a_1\|^2} a_1, z - x\Big\rangle$

$$= \langle a_2, z\rangle - \langle a_2, x\rangle - \Big(\frac{\langle a_2, a_1\rangle}{\|a_1\|^2}\Big)\big[\langle a_1, z\rangle - \langle a_1, x\rangle\big]$$

$$= 0$$

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

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

d). Suppose we have two solution z_1 an[redacted] 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.

Q2.

2. (20 pts) Let $\{(x_i, y_i)\}_{i=1}^{N}$ be given with $x_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$. Assume $N < n$ and x_i are linearly independent. Give the closed form solution to the ridge regression

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

In other words, suppose we write $X = [x_1, \ldots, x_N]^T$ and $y = [y_1, \ldots, y_N]^T$, represent a using the matrix X and vector y and λI.

Q2.

Let L be the objective function

$$\frac{\partial L}{\partial a_k} = \sum_{i=1}^{N} 2\left(\sum_{j=1}^{M} a_j x_{ij} - y_i\right) x_{ik} + 2\lambda a_k$$

Note that $\sum_{i=1}^{N} x_{ik} \sum_{j=1}^{M} a_j x_{ij} = \sum_{i=1}^{N} x_{ik} (X \cdot a)_i = (X^T \cdot X \cdot a)_k$

and $\sum_{i=1}^{N} x_{ik} y_i = (X^T y)_k$

Putting this together,

$$(X^T X a)_k - (X^T y)_k + \lambda a_k = 0 \quad \forall k.$$

Rewrite it as a vector equation, we have:

$$X^T y = (X^T X + \lambda I) a$$

$$x^T y = (x^T X + \lambda I)\, a$$

$$\therefore \quad a = (x^T X + \lambda I)^{-1} x^T y$$

$\therefore$

with $\langle \Phi(u), \Phi(v) \rangle = 2\langle u, v\rangle + 5\langle u, \ldots$

der $H = \mathbb{R}$ ...

and

3a).

$$2\langle u,v\rangle^2 + 5\langle u,v\rangle^3 \qquad u[$$

$$= \cancel{2(u^T v)^2 + 5(u^T v)^3}$$

let $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in \mathbb{R}^3$, $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3$.

$$2\langle u,v\rangle^2 + 5\langle u,v\rangle^3$$

$$= 2(u_1v_1 + u_2v_2 + \ldots + u_nv_n)^2 + 5(u_1v_1 + u_2v_2 + \ldots + u_nv_n)^3$$

$$= 2(u_1^2 v_1^2 +$$

$= 2(u_1^2 v_1^2 +$

$= 2(u_1v_1 + u_2v_2 + u_3v_3)^2 + 5(u_1v_1 + u_2v_2 + u_3v_3)^3$

$= 2(u_1^2v_1^2 + u_1v_1u_2v_2 + u_1v_1u_3v_3 + u_2v_2u_1v_1 + u_2^2v_2^2 + u_2v_2u_3v_3 +$
$u_3v_3u_1v_1 + u_3v_3u_2v_2 + u_3^2v_3^2) + 5(u_1v_1 + u_2v_2 + u_3v_3)^3$

$= 2(u_1^2v_1^2 + 2u_1v_1u_2v_2 + 2u_1v_1u_3v_3 + 2u_2v_2u_3v_3 + u_2^2v_2^2 + u_3^2v_3^2) +$
$\underbrace{5(u_1v_1 + u_2v_2 + u_3v_3)^3}$

For this term, we have:

$u_1^3v_1^3 + u_2^3v_2^3 + u_3^3v_3^3 + 2u_1^2v_1^2u_2v_2 + 2u_1^2v_1^2u_3v_3 + 6u_1u_2u_3v_1v_2v_3 +$
$2u_1v_1u_2^2v_2^2 + 2u_2^2v_2^2u_3v_3 + 2u_1v_1u_3^2v_3^2 + 2u_2v_2u_3^2v_3^2$

$\therefore$ we can conclude that $(\langle x, z\rangle)^2 = (x_1^2z_1^2 + x_2^2z_2^2 + \dots + x_n^2z_n^2) + 2$
$(x_1z_1x_2z_2 + x_1z_1x_3z_3 + \dots + x_{n-1}z_{n-1}x_nz_n)$

and similarly for $(\langle x, z\rangle)^3 = (x_1^3z_1^3 + x_2^3z_2^3 + \dots + x_n^3z_n^3) +$

$2\left(\sum_{\substack{i,j\le n \\ i\ne j}} u_i^2v_i^2u_jv_j\right) + 6\, u_1u_2u_3\dots u_nv_1v_2\dots v_n$

a). ~~And then we can have~~

Since we know that we ~~know~~ don't need to find the explicit transformation Φ

Adding the previous result, we have

$$2\langle u,v\rangle^2 + 5\langle u,v\rangle^3 = 2(u_1^2v_1^2 + u_2^2v_2^2 + u_3^2v_3^2) + 4(u_1u_2v_1v_2 + u_1u_3v_1v_3 + u_2u_3v_2v_3)$$
$$+ 5(u_1^3v_1^3 + u_2^3v_2^3 + u_3^3v_3^3) + 10\Big(\sum_{\substack{i,j\le n \\ i\ne j}} u_i^2v_i^2u_jv_j\Big) + 30\,u_1u_2u_3v_1v_2v_3$$

$\therefore$ we can construct a Φ ~~s.t.~~ ~~$\Phi \in \mathbb{R}^n$~~, $\Phi: \mathbb{R}^n \to H$ s.t.

~~st~~ $\langle \Phi(u), \Phi(v)\rangle = 2\langle u,v\rangle^2 + 5\langle u,v\rangle^3$ for $H = \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$

b). Consider the set of Cauchy sequences (x_n) in ~~$\mathbb{R}$~~ $\mathbb{R}^n$. Two Cauchy sequences (x_n) and (y_n) are defined to be equivalent

if $\lim_{n\to\infty} \|x_n - y_n\| = 0$

~~=~~ This establishes an equivalent relationship among all cauchy

~~This~~ This establishes an equivalent relationship among all cauchy ~~seq~~ sequence in $\mathbb{R}^n$.

Then we define

H with a vector space structure:

$$\alpha[x_n] + \beta[y_n] = [\alpha x_n + \beta y_n]$$

let (x_n) and (y_n) be the cauchy sequence in $\mathbb{R}^n$,

Consider the series

$$\lim_{n\to\infty} (x_n, y_n)_{\mathbb{R}^n}$$

~~ex~~

this limit exists since:

$$\left|(x_n, y_n)_{\mathbb{R}^n} - (x_m, y_m)_{\mathbb{R}^n}\right| = \left|(x_n, y_n)_{\mathbb{R}^n} - (x_n, y_m)_{\mathbb{R}^n} + (x_n, y_m)_{\mathbb{R}^n} - (x_m, y_m)_{\mathbb{R}^n}\right|$$

(triangle ineq.) $\leq \left|(x_n, y_n - y_m)_{\mathbb{R}^n}\right| + \left|(x_n - x_m, y_m)_{\mathbb{R}^n}\right|$

(cauchy-schwarz) $\leq \|x_n\|_{\mathbb{R}^n} \|y_n - y_m\|_{\mathbb{R}^n} + \|x_n - x_m\| \, \|y_m\|_{\mathbb{R}^n}$

$\therefore$ ~~th~~ since the cauchy sequences are bounded, $\exists M > 0$ s.t.

$\forall m, n$, $\left|(x_n, y_n)_{\mathbb{R}^n} - (x_m, y_m)_{\mathbb{R}^n}\right| \leq M\left(\|y_n - y_m\|_{\mathbb{R}^n} + \|x_n - x_m\|_{\mathbb{R}^n}\right)$

$\Rightarrow$ $(x_n, y_n)_{\mathbb{R}^n}$ is a Cauchy sequence in f, hence converges.

($\because$ f with non-negative coefficients)

36)

3b) Cont'd

Let $\Phi : \mathbb{R}^n \to H$. For $x \in \mathbb{R}^n$ let

$$\Phi(x) = [(x, x, \dots)]_H$$

maps every vector in $\mathbb{R}^n$ into the equivalent class of a constant sequence.

By the definition of linear structure on H, Φ is linear.

$\Rightarrow$ the inner product on H is

$$\langle \Phi(x), \Phi(y) \rangle_H = \lim_{n\to\infty} (\Phi(x)_n, \Phi(y)_n)_{\mathbb{R}^n} = \lim_{n\to\infty} (x, y)_{\mathbb{R}^n} = (x, y)_{\mathbb{R}^n}.$$

So, $\langle \Phi(u), \Phi(v) \rangle = f(u, v)$ for all $u, v \in \mathbb{R}^n$ //

More explanation:

Let $\Phi(v) \in \mathbb{R}^{n \times n}$ where $(i j) = v_i \times v_j$

Then $\langle \Phi(u), \Phi(v) \rangle = \langle u, v \rangle^2$

So, we can extend Φ to higher rank ~~of highe~~ or arrays of higher dimensions.

$\Rightarrow$ We can get $\langle u, v \rangle^d$ for any nonnegative integer d.

and using the hint and result of (a),

we can ~~construct~~ get the result of $f(\langle u, v \rangle)$.

$\therefore$ H would be $\mathbb{R} / \mathbb{R}^n \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n \times n} \times \dots \times \mathbb{R}^{\dots}$

$\dots (\|y_n - y_m\|_{\mathbb{R}^n} + \|x_n - x_m\|_{\mathbb{R}^n})$

... cauchy sequence in f, hence converges

($\because$ f with non-negati...

coefficients)