MSBD 5004 HOMEWORK 3
Name: Zheng Mingen
Stu id: 21126390

1. (a)

Not linear.

example:

$$\alpha=\beta=1,\quad \vec{x}=\begin{bmatrix}1\\0\end{bmatrix},\ \vec{y}=\begin{bmatrix}-1\\0\end{bmatrix}$$

then $f(\alpha\vec{x}+\beta\vec{y})=f\left(\begin{bmatrix}0\\0\end{bmatrix}\right)=0$

$$\alpha f(\vec{x})+\beta f(\vec{y})=(1-0)+(0-(-1))=2\neq f(\alpha\vec{x}+\beta\vec{y})$$

(b) Linear.

proof:
$$\begin{aligned} f(\alpha\vec{x}+\beta\vec{y}) &= (\alpha x_n+\beta y_n)-(\alpha x_1+\beta y_1) \\ &= \alpha(x_n-x_1)+\beta(y_n-y_1) \\ &= \alpha f(\vec{x})+\beta f(\vec{y}),\quad \forall \vec{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix},\ \vec{y}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}\in R^n \text{ and } \alpha,\beta\in R. \end{aligned}$$

inner product representation:

$$f(\vec{x})=x_n-x_1=\left\langle \begin{bmatrix}-1\\0\\ \vdots\\0\\1\end{bmatrix},\ \vec{x}\right\rangle=\vec{a}^T\vec{x},\ \text{for } \forall\vec{x}\in R^n,$$

where $\vec{a}=\begin{bmatrix}-1\\0\\ \vdots\\0\\1\end{bmatrix}$ $\to$ n-th entry

2. (a) False.

Reason: $y=\vec{x}^T\vec{a}+b=\left(\sum_{i=1}^{8}x_i\cdot a_i\right)+b$ is not merely depend on a_3 and x_3, even if $a_3>0$, $x_3>0$, if other terms are significantly negative, y can still be negative.

The counterexample is easy to find:

when $\vec{x}=\begin{bmatrix}0\\0\\1\\100\\0\\0\\0\\0\end{bmatrix}$, $\vec{a}=\begin{bmatrix}0\\0\\1\\-1\\0\\0\\0\\0\end{bmatrix}$, $b=-100$,

so $a_3>0$, $x_3>0$, but $y=\vec{x}^T\vec{a}+b=-199<0$.

(b) True.

Reason: $y=\vec{x}^T\vec{a}+b=\sum_{i=1}^{8}a_ix_i+b \xlongequal{\text{let } I=\{i\mid a_i\neq 0\}} \sum_{i\in I}a_ix_i+b$

thus y only depend on those feasures whose regression coefficient is not zero.

So, because a_2 (as the regression coefficient of the feature x_2) $=0$,

x_2 doesn't contribute any more to y.

(c) True.

Reason. As we discussed above,

$$y(x_1, x_2, \cdots, x_8) = \sum_{i=1}^{8} a_i x_i + b$$

if we denote the change of x_i, y as Δx_i, Δy,

then if we keep all other x is the same and only change x_6 to x_6'

we have $\Delta y = \left(\sum_{i \in \{1,2,\cdots,8\}-\{6\}} x_i a_i + x_6' \cdot a_6 + b\right) - \left(\sum_{i=1}^{8} x_i a_i + b\right)$

$$= a_6 \cdot (x_6' - x_6)$$

$$= a_6 \cdot \Delta x_6$$

if $a_6 = -0.8 < 0$ and $\Delta x_6 > 0$, then $\Delta y = -0.8 \Delta x_6 < 0$,

which means y is decreased as x_6 is increased.

[Remark: Indeed, $\frac{\partial y(x_1, x_2, \cdots, x_8)}{\partial x_6} = a_6$, (the partial derivative $a_6 < 0$) means that (y is monotonically decreasing in terms of x_6 increasing.)]

3.

(a) Set $L(\vec{\beta}) = \|X\vec{\beta} - \vec{y}\|_2^2$.

to solve $\min_{\vec{\beta} \in R^2} \|X\vec{\beta} - \vec{y}\|_2^2$ is an extremum problem.

$$\frac{\partial L}{\partial \vec{\beta}} = X^T[2(X\vec{\beta} - \vec{y})] = 2X^TX\vec{\beta} - 2X^T\vec{y}$$

Set $\frac{\partial L}{\partial \vec{\beta}} = 0$, we have $X^TX\vec{\beta} = X^T\vec{y}$

if $X = [\vec{x}^{(1)} \ \vec{x}^{(2)}]$ and $\vec{x}^{(1)} = 2\vec{x}^{(2)}$, then $r(X) \le 1$

so $r(X^TX) \le r(X) \le 1$,

which means X^TX is singular, and

$(X^TX)\vec{\beta} = X^T\vec{y}$ has infinite many solutions,

such that we cannot determine a unique $\vec{\beta}$ as our estimation.

(b) Suppose $f(\vec{\beta}) = \underbrace{\|X\vec{\beta}-\vec{y}\|_2^2}_{\downarrow} + \underbrace{\lambda\|\vec{\beta}\|_2^2}_{\downarrow}$

$$\frac{\partial f}{\partial \vec{\beta}} = \left(2X^TX\vec{\beta} - 2X^T\vec{y}\right) + 2\lambda\vec{\beta}$$

set $\frac{\partial f}{\partial \vec{\beta}} = 0$, we have $(X^TX+\lambda I)\vec{\beta} = X^T\vec{y}$

As X^TX is spd, λI is diagonal,

thus $X^TX+\lambda I$ is positive-definite, which means

$r(X^TX+\lambda I) = 2$, $X^TX+\lambda I$ is invertable.

We can get a unique $\vec{\beta}$ as the solution:

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

Suppose $X = \begin{bmatrix} 2a & a \\ 2b & b \end{bmatrix}$, a, b are arbitrary.

$$X^TX = \begin{bmatrix} 2a & 2b \\ a & b \end{bmatrix}\begin{bmatrix} 2a & a \\ 2b & b \end{bmatrix} = \begin{bmatrix} 4a^2+4b^2 & 2a^2+2b^2 \\ 2a^2+2b^2 & a^2+b^2 \end{bmatrix}$$

$$X^TX+\lambda I = \begin{bmatrix} 4a^2+4b^2+\lambda & 2a^2+2b^2 \\ 2a^2+2b^2 & a^2+b^2+\lambda \end{bmatrix}$$

$$(X^TX+\lambda I)^{-1} = \frac{1}{(4a^2+4b^2+\lambda)(a^2+b^2+\lambda)-(2a^2+2b^2)^2}\begin{bmatrix} a^2+b^2+\lambda & -(2a^2+2b^2) \\ -(2a^2+2b^2) & 4a^2+4b^2+\lambda \end{bmatrix}$$

$$X^T\vec{y} = \begin{bmatrix} 2a & 2b \\ a & b \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2ay_1+2by_2 \\ ay_1+by_2 \end{bmatrix} = (ay_1+by_2)\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

thus $$\hat{\beta} = \frac{ay_1+by_2}{(4a^2+4b^2+\lambda)(a^2+b^2+\lambda)-(2a^2+2b^2)^2}\cdot\begin{bmatrix} 2(a^2+b^2+\lambda)-(2a^2+2b^2) \\ -2(2a^2+2b^2)+(4a^2+4b^2+\lambda) \end{bmatrix}$$

$$= \frac{(ay_1+by_2)\cdot\lambda}{(4a^2+4b^2+\lambda)(a^2+b^2+\lambda)-(2a^2+2b^2)^2}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$ ✓

$\hat{\beta}_1 / \hat{\beta}_2 = \frac{2}{1} = 2$. ✓

(c) √ When $\lambda \to +\infty$, the punishment on $\vec{\beta}$ with greater norm ($\|\vec{\beta}\|_2^2$) is strong, which forces

$$\hat{\vec{\beta}} = \frac{(ay_1+by_2)\lambda}{(4a^2+4b^2+\lambda)(a^2+b^2+\lambda)-(2a^2+2b^2)^2}\begin{bmatrix}2\\1\end{bmatrix} \to \vec{0}$$ to adapt the optimization

(simply as $\underbrace{\text{the degree of } \lambda \text{ in numerator}}_{\text{is } 1}$ is less than $\underbrace{\text{that in denominator}}_{\text{is } 2}$.)

√ when $\lambda \to 0$, the regulazation term loses its contraint force, $\hat{\vec{\beta}}$ tend to be close to $\min_{\vec{\beta}\in R^2} \|X\vec{\beta}-\vec{y}\|_2^2$, which may make the solution unstable.

√ So in practice we should choose a proper λ, bigger λ can mitigate issues arising from linear dependency of columns of X, but extremly big λ can force $\vec{\beta}$ to be $\vec{0}$, extremely small λ will loosen the constraint and cause the dependency issues to reoccur.

---

4. Proof:

$$\min_{\vec{a}\in R^n, b\in R} \sum_{i=1}^{N} h(y_i(\langle\vec{a},\vec{x}_i\rangle+b)-1) + \lambda\|\vec{a}\|_2^2 \quad ①$$

$\boxed{h(t)=\max\{0,-t\}}$

$$\min_{\vec{a}\in R^n, b\in R} \sum_{i=1}^{N} \max\{0, 1-y_i(\langle\vec{a},\vec{x}_i\rangle+b)\} + \lambda\|\vec{a}\|_2^2$$

$\boxed{\text{denote } \max\{0, 1-y_i(\langle\vec{a},\vec{x}_i\rangle+b)\} \text{ as } \varphi_i}$

$$\min_{\substack{\vec{a}\in R^n, b\in R,\\ \vec{\varphi}\in R^n}} \sum_{i=1}^{N} \varphi_i + \lambda\|\vec{a}\|_2^2 . \quad ②$$

when $1-y_i(\langle\vec{a},\vec{x}\rangle+b) \le 0$, $\varphi_i = 0$

when $1-y_i(\langle\vec{a},\vec{x}\rangle+b) > 0$, $\varphi_i = 1-y_i(\langle\vec{a},\vec{x}\rangle+b) > 0$

$\Rightarrow \boxed{\varphi_i \geq 0}$

in ①, we try to find a solution s.t. $h(y_i(\langle\vec{a},\vec{x}\rangle+b)-1)=0$.

i.e. $y_i(\langle\vec{a},\vec{x}\rangle+b) \geq 1$

$\Updownarrow$

in ②, we try to find a solution s.t. $\varphi_i = 0$

i.e. $1-y_i(\langle\vec{a},\vec{x}_i\rangle+b) \leq 0$ ③

Since $\varphi_i \geq 0$, ③ can be expressed as

$1-y_i(\langle\vec{a},\vec{x}_i\rangle+b) \leq 0 \leq \varphi_i$

i.e. $\boxed{y_i(\langle\vec{a},\vec{x}_i\rangle+b) \geq 1-\varphi_i}$.

Overall, $$\min_{\substack{\vec{a}\in R^n \\ b\in R}} \sum_{i=1}^{N} h\left(y_i(\langle \vec{a}, \vec{x}_i \rangle + b) - 1\right) + \lambda \|\vec{a}\|_2^2 .$$

$$\Updownarrow$$

$$\min_{\substack{\vec{a}\in R^n \\ b\in R \\ \vec{\varphi}\in R^n}} \sum_{i=1}^{N} \varphi_i + \lambda \|\vec{a}\|_2^2$$

$$\text{s.t. } y_i\left(\langle \vec{a}, \vec{x}_i \rangle + b\right) \geqslant 1 - \varphi_i \text{ and } \varphi_i \geqslant 0 ,$$

$$i = 1, 2, \cdots, N.$$

$\boxtimes$

(in next page)

5.

(a) Proof: $\forall \vec{x}, \vec{z} \in S_1 \cap S_2$, we have

① $\langle \vec{a}_1, (1+t)\vec{z} - t\vec{x} \rangle = \langle \vec{a}_1, \vec{z} \rangle + t \langle \vec{a}_1, \vec{z} - \vec{x} \rangle$

$$= \langle \vec{a}_1, \vec{z} \rangle + t(\langle \vec{a}_1, \vec{z} \rangle - \langle \vec{a}_1, \vec{x} \rangle)$$

$$\overset{\vec{x},\vec{z} \in S_1}{=} b_1 + t(b_1 - b_1)$$

$$= b_1$$

$\Rightarrow (1+t)\vec{z} - t\vec{x} \in S_1$ for $\forall \vec{x}, \vec{z} \in S_1 \cap S_2$ ✓

② $\langle \vec{a}_2, (1+t)\vec{z} - t\vec{x} \rangle = \langle \vec{a}_2, \vec{z} \rangle + t \langle \vec{a}_2, \vec{z} - \vec{x} \rangle$

$$= \langle \vec{a}_2, \vec{z} \rangle + t(\langle \vec{a}_2, \vec{z} \rangle - \langle \vec{a}_2, \vec{x} \rangle)$$

$$\overset{\vec{x},\vec{z} \in S_2}{=} b_2 + t(b_2 - b_2)$$

$$= b_2$$

$\Rightarrow (1+t)\vec{z} - t\vec{x} \in S_2$ for $\forall \vec{x}, \vec{z} \in S_1 \cap S_2$

Therefore, by both ① & ② we can conclude that

$(1+t)\vec{z} - t\vec{x} \in S_1 \cap S_2$ for $\forall \vec{x}, \vec{z} \in S_1 \cap S_2$

(b) Proof:

① We first prove:

$\vec{z}$ is a solution of $\min_{\vec{x} \in S_1 \cap S_2} \|\vec{x} - \vec{y}\| \Rightarrow \vec{z} \in S_1 \cap S_2$ and $\langle \vec{z} - \vec{y}, \vec{z} - \vec{x} \rangle = 0$, $\forall \vec{x} \in S_1 \cap S_2$

As we have proved in (a), $(1+t)\vec{z} - t\vec{x} \in S_1 \cap S_2$.

Since $\vec{z}$ is the closest to $\vec{y}$ on $S_1 \cap S_2$, we have

$$\|\vec{z} - \vec{y}\|^2 \le \|[(1+t)\vec{z} - t\vec{x}] - \vec{y}\|^2$$

$$= \|(\vec{z} - \vec{y}) + t(\vec{z} - \vec{x})\|^2$$

(as V is a Hilbert space, $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$ for $\forall \vec{x} \in V$)

$$= \|\vec{z} - \vec{y}\|^2 + t^2 \|\vec{z} - \vec{x}\|^2 + 2t \langle \vec{z} - \vec{y}, \vec{z} - \vec{x} \rangle.$$

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

- if we choose $t > 0$, then we have

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

letting $t \to 0^+$, then $\langle \vec{z} - \vec{y}, \vec{z} - \vec{x} \rangle \ge 0$

- if we choose $t<0$, then

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

letting $t \to 0^-$, then $\langle \vec{z}-\vec{y}, \vec{z}-\vec{x} \rangle \le 0$

Altogether, $\underbrace{\langle \vec{z}-\vec{y}, \vec{z}-\vec{x} \rangle = 0 \text{ for } \forall \vec{x} \in S_1 \cap S_2}$

and of course since $\vec{z}$ is a solution, $\underbrace{\vec{z} \in S_1 \cap S_2}$

② We then prove:

$$\boxed{\vec{z} \in S_1 \cap S_2 \text{ and } \langle \vec{z}-\vec{y}, \vec{z}-\vec{x} \rangle = 0 \text{ for } \forall \vec{x} \in S_1 \cap S_2 \Rightarrow \vec{z} \text{ is a solution of } \min_{\vec{x} \in S_1 \cap S_2} \|\vec{x}-\vec{y}\|}$$

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

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

in Hilbert space, $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$

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

$$\overset{\langle \vec{z}-\vec{x}, \vec{z}-\vec{y} \rangle = 0}{=\!=\!=} \|\vec{z}-\vec{x}\|^2 + \|\vec{z}-\vec{y}\|^2$$

$$\overset{\text{as } \|\vec{z}-\vec{x}\|^2 \ge 0}{\geqslant} \|\vec{z}-\vec{y}\|^2 \quad \text{for } \forall \vec{x} \in S_1 \cap S_2$$

Which means $\vec{z}$ is the closest vector to $\vec{y}$ on $S_1 \cap S_2$,

i.e. $\vec{z}$ is a solution of $\min_{\vec{x} \in S_1 \cap S_2} \|\vec{x}-\vec{y}\|$

(C) Solution: We are finding a solution of an omptimization problem with 2 linear constraints, which is an extremum problem that can be solved by Lagrange multiplier.

Note that $\min_{\vec{x} \in S_1 \cap S_2} \|\vec{x}-\vec{y}\| = \min_{\vec{x} \in S_1 \cap S_2} \|\vec{x}-\vec{y}\|^2$, so:

Construct the Lagrange function:

$$L(\vec{x}, \lambda_1, \lambda_2) = \|\vec{x}-\vec{y}\|^2 + \lambda_1(\langle \vec{a}_1, \vec{x} \rangle - b_1) + \lambda_2(\langle \vec{a}_2, \vec{x} \rangle - b_2)$$

$$L'_{\vec{x}} = 2\vec{x} - 2\vec{y} + \lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 = \vec{0} \Leftrightarrow 2\vec{x} = 2\vec{y} - \lambda_1 \vec{a}_1 - \lambda_2 \vec{a}_2 \quad ①$$

$$L'_{\lambda_1} = \langle \vec{a}_1, \vec{x} \rangle - b_1 = 0 \Leftrightarrow \langle \vec{a}_1, 2\vec{x} \rangle = 2b_1 \quad ②$$

$$L'_{\lambda_2} = \langle \vec{a}_2, \vec{x} \rangle - b_2 = 0 \Leftrightarrow \langle \vec{a}_2, 2\vec{x} \rangle = 2b_2 \quad ③$$

put ① into ②, ③ respectively, we get:

$$\begin{cases} \langle \vec{a}_1, \vec{a}_1 \rangle \lambda_1 + \langle \vec{a}_1, \vec{a}_2 \rangle \lambda_2 = \langle \vec{a}_1, 2\vec{y} \rangle - 2b_1 & ④ \\ \langle \vec{a}_2, \vec{a}_1 \rangle \lambda_1 + \langle \vec{a}_2, \vec{a}_2 \rangle \lambda_2 = \langle \vec{a}_2, 2\vec{y} \rangle - 2b_2 & ⑤ \end{cases}$$

$④ - \frac{\langle \vec{a}_1, \vec{a}_2 \rangle}{\langle \vec{a}_2, \vec{a}_2 \rangle} \times ⑤$, we have:

$$\left[\langle \vec{a}_1, \vec{a}_1 \rangle - \frac{\langle \vec{a}_1, \vec{a}_2 \rangle^2}{\langle \vec{a}_2, \vec{a}_2 \rangle} \right] \lambda_1 = \left(\langle \vec{a}_1, 2\vec{y} \rangle - 2b_1 \right) - \frac{\langle \vec{a}_1, \vec{a}_2 \rangle}{\langle \vec{a}_2, \vec{a}_2 \rangle} \left(\langle \vec{a}_2, 2\vec{y} \rangle - 2b_2 \right)$$

$$\Rightarrow \lambda_1 = \frac{\langle \vec{a}_2, \vec{a}_2 \rangle \left(\langle \vec{a}_1, 2\vec{y} \rangle - 2b_1 \right) - \langle \vec{a}_1, \vec{a}_2 \rangle \left(\langle \vec{a}_2, 2\vec{y} \rangle - 2b_2 \right)}{\langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle - \langle \vec{a}_1, \vec{a}_2 \rangle^2}$$

similarly, by $④ - \frac{\langle \vec{a}_1, \vec{a}_1 \rangle}{\langle \vec{a}_2, \vec{a}_1 \rangle} \times ⑤$, we have:

$$\lambda_2 = \frac{\langle \vec{a}_1, \vec{a}_2 \rangle \left(\langle \vec{a}_1, 2\vec{y} \rangle - 2b_1 \right) - \langle \vec{a}_1, \vec{a}_1 \rangle \left(\langle \vec{a}_2, 2\vec{y} \rangle - 2b_2 \right)}{\langle \vec{a}_1, \vec{a}_2 \rangle^2 - \langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle}$$

put λ_1, λ_2 into ①, we get the soluetion of $\min\limits_{\vec{x} \in S_1 \cap S_2} \|\vec{x} - \vec{y}\|$:

$$\begin{aligned} \vec{z} = \min_{\vec{x} \in S_1 \cap S_2} \|\vec{x} - \vec{y}\| &= \vec{y} - \frac{\lambda_1}{2} \vec{a}_1 - \frac{\lambda_2}{2} \vec{a}_2 \\ &= \vec{y} - \underbrace{\frac{\langle \vec{a}_2, \vec{a}_2 \rangle \left(\langle \vec{a}_1, \vec{y} \rangle - b_1 \right) - \langle \vec{a}_1, \vec{a}_2 \rangle \left(\langle \vec{a}_2, \vec{y} \rangle - b_2 \right)}{\left(\langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle - \langle \vec{a}_1, \vec{a}_2 \rangle^2 \right)}}_{\text{denoted as } \rho_1} \vec{a}_1 \\ &\quad - \underbrace{\frac{\langle \vec{a}_1, \vec{a}_1 \rangle \left(\langle \vec{a}_2, \vec{y} \rangle - b_2 \right) - \langle \vec{a}_1, \vec{a}_2 \rangle \left(\langle \vec{a}_1, \vec{y} \rangle - b_1 \right)}{\underbrace{\left(\langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle - \langle \vec{a}_1, \vec{a}_2 \rangle^2 \right)}_{\text{denoted as } K}}}_{\text{denoted as } \rho_2} \vec{a}_2 \end{aligned}$$

Proof: $\boxed{\text{Part 1}}$ First we prove $\vec{z} \in S_1 \cap S_2$.

P1.1

$$\langle \vec{a}_1, \vec{z} \rangle = \langle \vec{a}_1, \vec{y} \rangle - \rho_1 \cdot \langle \vec{a}_1, \vec{a}_1 \rangle - \rho_2 \langle \vec{a}_1, \vec{a}_2 \rangle$$

$$= \langle \vec{a}_1, \vec{y} \rangle - \frac{\langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle \left(\langle \vec{a}_1, \vec{y} \rangle - b_1 \right) - \cancel{\langle \vec{a}_1, \vec{a}_2 \rangle \langle \vec{a}_1, \vec{a}_1 \rangle \left(\langle \vec{a}_2, \vec{y} \rangle - b_2 \right)} + \cancel{\langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_1, \vec{a}_2 \rangle \left(\langle \vec{a}_2, \vec{y} \rangle - b_2 \right)} - \langle \vec{a}_1, \vec{a}_2 \rangle^2 \left(\langle \vec{a}_1, \vec{y} \rangle - b_1 \right)}{K}$$

$$= \langle \vec{a}_1, \vec{y} \rangle - \frac{\left(\langle \vec{a}_1, \vec{y} \rangle - b_1 \right) \cdot K}{K}$$

$$= b_1$$

so, $\vec{z} \in S_1$

$\boxed{P1.2}$ similarly:

$$\langle \vec{a}_2, \vec{z} \rangle = \langle \vec{a}_2, \vec{y} \rangle - p_1 \langle \vec{a}_2, \vec{a}_1 \rangle - p_2 \langle \vec{a}_2, \vec{a}_2 \rangle$$

$$= \langle \vec{a}_2, \vec{y} \rangle - \frac{\cancel{\langle \vec{a}_2, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle (\langle \vec{a}_1, \vec{y} \rangle - b_1)} - \langle \vec{a}_1, \vec{a}_2 \rangle^2 (\langle \vec{a}_2, \vec{y} \rangle - b_2) + \langle \vec{a}_1, \vec{a}_1 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle (\langle \vec{a}_2, \vec{y} \rangle - b_2) \cancel{- \langle \vec{a}_1, \vec{a}_2 \rangle \langle \vec{a}_2, \vec{a}_2 \rangle (\langle \vec{a}_1, \vec{y} \rangle - b_1)}}{K}$$

$$= \langle \vec{a}_2, \vec{y} \rangle - \frac{K \cdot (\langle \vec{a}_2, \vec{y} \rangle - b_2)}{K}$$

$$= b_2 .$$

so $\vec{z} \in S_2$

Overall, $\vec{z} \in S_1 \cap S_2$.

$\boxed{\text{Part 2}}$ Then we prove $\langle \vec{z} - \vec{y}, \vec{z} - \vec{x} \rangle = 0$ for $\forall \vec{x} \in S_1 \cap S_2$.

$$\langle \vec{z} - \vec{x}, \vec{z} - \vec{y} \rangle = -\frac{\lambda_1}{2} \langle \vec{a}_1, \vec{z} - \vec{x} \rangle - \frac{\lambda_2}{2} \langle \vec{a}_2, \vec{z} - \vec{x} \rangle$$

$$= -\frac{\lambda_1}{2} \left(\langle \vec{a}_1, \vec{z} \rangle - \langle \vec{a}_1, \vec{x} \rangle \right) - \frac{\lambda_2}{2} \left(\langle \vec{a}_2, \vec{z} \rangle - \langle \vec{a}_2, \vec{x} \rangle \right)$$

$$= -\frac{\lambda_1}{2} (b_1 - b_1) - \frac{\lambda_2}{2} (b_2 - b_2) = 0$$

(d) Proof: Suppose $\vec{z}_1$, $\vec{z}_2$ are both solutions, then we have:

$$\langle \vec{z}_1 - \vec{y}, \vec{z}_1 - \vec{z}_2 \rangle = 0 \quad ①$$

$$\langle \vec{z}_2 - \vec{y}, \vec{z}_2 - \vec{z}_1 \rangle = 0 \iff \langle \vec{z}_2 - \vec{y}, \vec{z}_1 - \vec{z}_2 \rangle = 0 \quad ②$$

① - ②, we have $\langle \vec{z}_1 - \vec{z}_2, \vec{z}_1 - \vec{z}_2 \rangle = 0 \iff \vec{z}_1 - \vec{z}_2 = \vec{0} \iff \vec{z}_1 = \vec{z}_2$.

so $\vec{z}$ is unique.