Exercise (4)

Ex. 1

a) $$\min_{\theta} J(\theta) = \sum_{k=1}^{n} \|\theta - x_k\|^2$$

$$\text{S.T. } \theta^T b = 0 \quad \text{where } b \in \mathbb{R}^d$$

$$\mathcal{L}(\theta, l) = \sum_{k=1}^{n} \|\theta - x_k\|^2 + l\theta^T b$$

$\nabla \mathcal{L}(\theta, l) = 0$:

$$\frac{\partial \mathcal{L}}{\partial l} = \theta^T b = 0 \overset{*1}{=} \left(\bar{x} - \frac{lb}{2n}\right)^T b \Rightarrow \bar{x}^T b = \frac{l}{2n} b^T b$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = 2\sum_{k=1}^{n} (\theta - x_k) + lb = 0 \quad \Rightarrow \underbrace{\frac{l}{2n} = \frac{\bar{x}^T b}{b^T b}}_{*2}$$

$$= 2 \cdot n \cdot \theta - 2\sum_{k=1}^{n} x_k + lb = 0$$

$$\Rightarrow \theta = \frac{1}{n} \sum_{k=1}^{n} x_k - \frac{lb}{2 \cdot n} = \overset{*1}{\bar{x} - \frac{lb}{2n}} =$$

$$\overset{*2}{=} \underbrace{\bar{x} - \frac{\bar{x}^T b}{b^T b} \cdot b}_{\mathrm{Proj}_{\vec{b}} \bar{x} \ (!)}$$

Geometric Example:

b)

$$\min_{\theta} J(\theta) = \sum_{k=1}^{n} \|\theta - x_k\|^2$$

$$\text{S.T. } \|\theta - c\|^2 = 1 \quad \text{where } c \in \mathbb{R}^d$$

$$\mathcal{L}(\theta, l) = \sum_{k=1}^{n} \|\theta - x_k\|^2 + l(\|\theta - c\|^2 - 1)$$

$\nabla \mathcal{L}(\theta, l) = 0$:

$$\frac{\partial \mathcal{L}}{\partial \theta} = 2 \sum_{k=1}^{n} (\theta - x_k) + 2l \sum_{k=1}^{n} (\theta - c) = 0$$

$$\Leftrightarrow 2 \cdot n \cdot \theta - 2 \sum_{k=1}^{n} x_k + 2 \cdot l \cdot n(\theta - c) = 0$$

$$\Leftrightarrow \theta + l\theta = \bar{x} - lc \Leftrightarrow \theta = \frac{\bar{x} + lc}{1 + l} \; *$$

$$\frac{\partial \mathcal{L}}{\partial l} = \|\theta - c\|^2 - 1 = 0 \Leftrightarrow \left\| \frac{\bar{x} + lc}{1 + l} - c \right\|^2 = 1 \Leftrightarrow$$

$$\Leftrightarrow \left\| \frac{\bar{x} - c}{(1 + l)} \right\|^2 = 1 \Leftrightarrow \|\bar{x} - c\|^2 = (1 + l)^2 \Leftrightarrow$$

$$\Leftrightarrow l^2 + 2l + 1 - \|\bar{x} - c\|^2 = 0$$

$$\Rightarrow l = \frac{-2 \pm \sqrt{2^2 - 4 \cdot (1 - \|\bar{x} - c\|^2)}}{2}$$

$$= \frac{-\cancel{2} \pm \sqrt{\cancel{4} \cdot \|\bar{x} - c\|^{\cancel{2}}}}{\cancel{2}} = -1 \pm \|\bar{x} - c\|$$

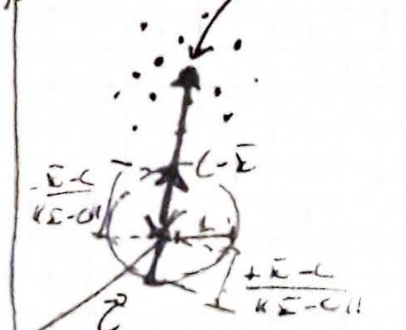
$$* \quad \theta = \frac{\bar{x} + (-1 \pm \|\bar{x} - c\|)c}{1 - 1 \pm \|\bar{x} - c\|} = \frac{\bar{x} - c \pm \|\bar{x} - c\| c}{\pm \|\bar{x} - c\|} =$$

$$= c \pm \frac{\bar{x} - c}{\|\bar{x} - c\|}$$

GEOMETRIC INTERPRETATION:

$$\|\theta - x_k\|^2 = \left\| c \pm \frac{\bar{x} - c}{\|\bar{x} - c\|} - \bar{x} \right\|^2 =$$

$$= \left\| c - \bar{x} \pm \frac{\bar{x} - c}{\|\bar{x} - c\|} \right\|^2$$



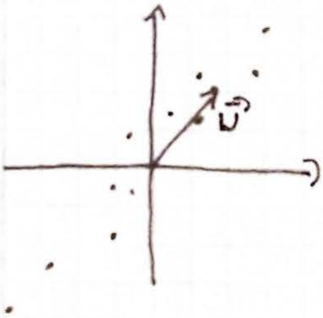
Ex. 2

a) $W^* = \underset{w}{\mathrm{argmax}}\ w^T X X^T w$ $\geq$ $X X^T w = l w$

s.t. $\|w\|^2 = 1$

AND: $l_1 = \max_w \|\sum_k w^T x_k\|^2 = w^T X X^T w = w^T S w$

$= \sum_k (w^T x_k)(x_k^T w) = w^T (\sum x_k x_k^T) w$



$\sum^d S_{ii} = \mathrm{Tr}(S) = \sum^d l_i \geq l_1$

since $l_i \geq 0 \quad \forall i = 1, \dots, d$

b) This upper bound is tight if $l_2, \dots, l_d = 0$.

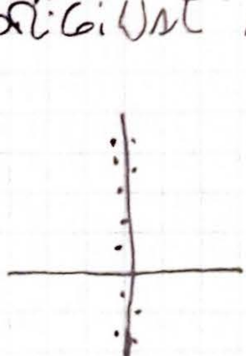
c) $S_{ii} = \sum_k (x_k^i - m)(x_k^i - m) = \sigma_i^2$

$l_1 \geq \sigma_i^2 \quad \forall i = 1, \dots, d \quad \Rightarrow \max_{i=1}^d S_{ii} \leq l_1$

Also:

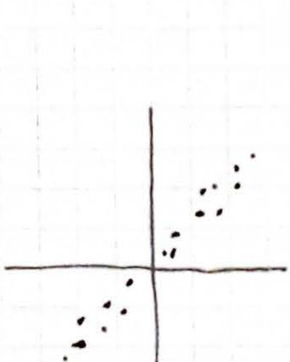
$l_1 = \max_{w \in W} w^T S w \geq \max_{w \in E} w^T S w = \max_{i=1} e_i^T S e_i = \max_{i=1} S_{ii}$

d) The bound $l_1 \geq \sigma_i^2$ is tight when the data spreads across one of the original axis:



Tight

Tight



Not Tight

$\boxed{\text{Ex. 3}}$

a) $v = S^{0.5} w \implies w = S^{-0.5} v$

$$J(w(v)) = \|Sw\| - \frac{1}{2} w^T S w =$$
$$= \|S S^{-0.5} v\| - \frac{1}{2} v^T \underbrace{S^{-0.5} S S^{-0.5}}_{I} v =$$
$$= \|S^{0.5} v\| - \frac{1}{2} v^T v$$

$$\frac{\partial J}{\partial v} = \frac{S^{0.5}(S^{0.5} v)}{\|S^{0.5} v\|} - v$$

For $\gamma = 1$: $v \leftarrow v + \frac{\partial J}{\partial v} = \frac{S^{0.5}(S^{0.5} v)}{\|S^{0.5} v\|}$

$$\cancel{S^{0.5}} w \leftarrow \frac{S^{0.5} \cdot S^{0.5} \cdot S^{0.5} w}{\|S^{0.5} S^{0.5} w\|} = \frac{\cancel{S^{0.5}} S w}{\|Sw\|}$$

$$w \leftarrow \frac{Sw}{\|Sw\|}$$

b) $\frac{\partial J}{\partial w} = S \frac{(Sw)}{\|Sw\|} - Sw = 0$

$$\iff \frac{SSw}{\|Sw\|} = Sw \implies \frac{Sw}{\|Sw\|} = w$$

$$\implies \left\| \frac{Sw}{\|Sw\|} \right\| = \|w\| = 1$$