Exercise Sheet 4

Exercise 1

Team: RCSKSK	
Rumesha	3811249
Chen	381215
Shah	384426
Swamy	384418
Khalil	385402
Kao	381410

(i) In this exercise, we want to minimize the objective function

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

subject to constraint $\theta^T b = 0$

where $x_1, \ldots x_K, \theta, b \in \mathbb{R}^d$

According to Lagrangian function

$$L(\theta, \lambda) = J(\theta) + \lambda \theta^T b$$

$$= \sum_{k=1}^{n} \| \theta - x_k \|^2 + \lambda \theta^T b$$

We set gradient to zero to solve for λ and θ.

$$\frac{\partial L(\theta, \lambda)}{\partial \theta} = \Big[\sum_{k=1}^{n} 2(\theta - x_k) \Big] + \lambda b = 0$$

$\theta^T b$

$$\Big[\sum_{k=1}^{n} 2(\theta - x_k) \Big] + \lambda b = 0$$

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

$$2 \sum_{k=1}^{n} x_k - \lambda b = 2n\theta$$

$$\Rightarrow \theta = \frac{1}{n} \sum_{k=1}^{n} x_k - \frac{1}{2n} \lambda b \quad \ldots\ldots \quad (1)$$

Use this to solve for λ $\qquad \left[\bar{x} - \frac{\lambda b}{2n} \right]$

$$\theta^T b = 0$$

$$\Longleftrightarrow \left(\frac{1}{n} \sum_{k=1}^{n} x_k - \frac{\lambda}{2n} b \right)^T b = 0$$

$$\Longleftrightarrow \left(\frac{1}{n} \sum_{k=1}^{n} x_k^T - \frac{\lambda}{2n} b^T \right) b = 0$$

$$\Longleftrightarrow \frac{1}{n} \sum_{k=1}^{n} x_k^T b = \frac{\lambda}{2n} b^T b$$

$$\Longleftrightarrow \lambda = \frac{2}{b^T b} \Big[\sum_{k=1}^{n} x_k^T \Big] b \quad \ldots\ldots \quad (2)$$

Substituting (2) in (1)

$$\theta = \frac{1}{n}\sum_{k=1}^{n} x_k - \frac{1}{2n}\lambda b$$

$$\iff \theta = \frac{1}{n}\sum_{k=1}^{n} x_k - \frac{1}{2n}\left(\frac{2}{b^T b}\left(\left[\sum_{k=1}^{n} x_k^T\right] b\right) b\right.$$

$$\theta = \left(\frac{1}{n}\left[\sum_{k=1}^{n} x_k\right] - \frac{\sum_{k=1}^{n} x_k^T (b)}{b' b}\, b\right)$$

b) Repeating the same procedure for constraint

$\|\theta - c\|^2 = 1, \; c \in \mathbb{R}^d$

$$L(\theta,\lambda) = J(\theta) + \lambda\|\theta - c\|^2 - \lambda$$

$$\frac{\partial L(\theta,\lambda)}{\partial\theta} = \left(\sum_{k=1}^{n} 2(\theta - x_k)\right] + \lambda\, 2(\theta - c)\Big) = 0$$

Solving for θ

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

$$2n\theta - 2\left[\sum_{k=1}^{n} x_k\right] + 2\lambda\theta - 2\lambda c = 0$$

$$2n\theta + 2\lambda\theta = 2\sum_{k=1}^{n} x_k + 2\lambda c$$

$$\theta = \frac{1}{n+\lambda}\left(\sum_{k=1}^{n} x_k + \lambda c\right)$$

Substituting θ in eq

$$\|\theta - c\|^2 - 1 = 0$$

$$\iff \left\|\frac{1}{n+\lambda}\left(\sum_{k=1}^{n} x_k + \lambda c\right) - c\right\|^2 - 1 = 0$$

$$\left\|\frac{1}{n+\lambda}\sum_{k=1}^{n} x_k + \frac{\lambda}{n+\lambda} c - c\right\|^2 - 1 = 0$$

$$\iff \left\|\frac{1}{n+\lambda}\sum_{k=1}^{n} x_k + \frac{\lambda}{n+\lambda} c - \frac{n+\lambda}{n+\lambda} c\right\| = 1$$

$$\left\|\frac{1}{n+\lambda}\sum_{k=1}^{n} x_k - \frac{n}{n+\lambda} c\right\|^2 = 1$$

$$\left(\left(\frac{1}{n+\lambda}\right)\left(\sum_{k=1}^{n} x_k - nc\right]\right)^T\left(\frac{1}{n+\lambda}\left(\sum_{k=1}^{n} x_k - nc\right]\right) = 1$$

$$\frac{1}{(n+\lambda)^2}\left(\sum_{k=1}^{n} x_k - nc\right)^T\left(\sum_{k=1}^{n} x_k - nc\right) = 1$$

$$\Longleftrightarrow \left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\|^2 = (\lambda + n)^2 = \lambda^2 + 2n\lambda + n^2$$

$$\Longleftrightarrow \lambda^2 + 2n\lambda + n^2 - \left\| \left(\sum_{k=1}^{n} x_k - c \right) \right\|^2$$

$$\lambda_{1,2} = \frac{-2n}{2} \pm \sqrt{\frac{2n^2}{2} - \left(n^2 - \left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\|^2 \right)}$$

$$= -n \pm \sqrt{\left\| \sum_{k=1}^{n} x_k - c \right\|}$$

$$\lambda_1 = -n + \left\| \left(\sum_{k=1}^{n} x_k - c \right) \right\|$$

$$\lambda_2 = -n - \left\| \left(\sum_{k=1}^{n} x_k - c \right) \right\|$$

Now compute θ using λ_1 and λ_2

λ_1
$$\theta_1 = \frac{1}{n + \lambda_1} \left(\left[\sum_{k=1}^{n} x_k \right] + \lambda_1 c \right)$$

$$= \frac{1}{n + \left(-n + \left\| \sum_{k=1}^{n} x_k - c \right] \right\| \right)} \left(\left[\sum_{k=1}^{n} x_k \right] + \left(-n + \left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\| \right) c \right)$$

$$= \frac{1}{\left\| \sum_{k=1}^{n} x_k - c \right\|} \left(\left[\sum_{k=1}^{n} x_k \right] - nc + \left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\| c \right)$$

$$= \frac{1}{\left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\|} \left[\sum_{k=1}^{n} x_k - c \right] + c$$

λ_2
$$\theta_2 = \frac{1}{n + \lambda_2} \left(\sum_{k=1}^{n} x_k + \lambda_2 c \right)$$

$$= \frac{1}{n + \left(-n - \left\| \left(\sum_{k=1}^{n} x_k - c \right] \right\| \right)} \left(\left[\sum_{k=1}^{n} x_k \right] + \left(-n - \left\| \sum_{k=1}^{n} x_k - c \right] \right\| c \right)$$

$$= - \frac{1}{\left\| \sum_{k=1}^{n} x_k - c \right] \right\|} \left(\left[\sum_{k=1}^{n} x_k \right] - nc - \left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\| c \right)$$

$$= \frac{1}{\left\| \left[\sum_{k=1}^{n} x_k - c \right] \right\|} \left[\sum_{k=1}^{n} c - x_k \right] + c.$$

Exercise 4

2.

(a) $\sum_{i=1}^{d} S_{ii} = \mathrm{tr} S = \mathrm{tr}(W \Lambda W^T) = \mathrm{tr}(\underbrace{W^T W}_{I} \Lambda) = \mathrm{tr}(\Lambda) = \sum_{i=1}^{P} \lambda_i \geq \lambda_1$

Thus $\sum_{i=1}^{d} S_{ii}$ is an upper bound to λ_1

(b) The upper bound is tight, which means $\sum_{i=1}^{P} \lambda_i = \lambda_1$.

All the eigenvalues except λ_1 are zero.

It means w_1 includes all the variance of the original data, and all the data are in one line

(c) Assume $S_{jj} = \max_{i=1}^{d} S_{ii} > \lambda_1$, $j \in 1, \cdots, P$

Then there exist $w_{j'}^T = (0, \cdots 0, 1, 0, \cdots 0)$, so that $\|w_j'\| = 1$

$\qquad\qquad\qquad\qquad\quad$ 1 $\qquad\qquad$ j $\quad$ d

$w_{j'}^T S w_{j'} = S_{jj} > \lambda_1 \geq \lambda_i \quad i = 1, \cdots, P$

which is contradictory with $\lambda_i = \max_{w_i} w_i^T S w_i$

Thus $\max_{i=1}^{d} S_{ii}$ is a lower bound to λ_1

(d) The lower bound is tight which means $\lambda_1 = \max_{i=1}^{d} S_{ii}$

It means the corresponding x_1 is the first component

3

(a) $J(W) = \|SW\| - \frac{1}{2}W^TSW$

$$\frac{\partial J(W)}{\partial V} = \left[\frac{dJ}{dW^T} \cdot \frac{\partial W}{\partial V}\right]^T$$

$$= \left[\left(\frac{W^TS^T}{\|SW\|} - W^TS^T\right) \cdot S^{-\frac{1}{2}}\right]^T$$

$$= S^{-\frac{1}{2}}\left(\frac{SSW}{\|SW\|} - SW\right)$$

$$= \frac{S^{\frac{3}{2}}W}{\|SW\|} - S^{\frac{1}{2}}W$$

$$= \frac{SV}{\|SW\|} - V$$

$$V = V + \gamma\frac{\partial J}{\partial V}$$

$$= V + \gamma\frac{SV}{\|SW\|} - \gamma V$$

$$V = \frac{SV}{\|SW\|} \quad \text{where } \gamma = 1$$

power iteration method :

$$W = \frac{SW}{\|SW\|}$$

$$S^{\frac{1}{2}}W = \frac{S^{\frac{3}{2}}W}{\|SW\|}$$

$$V = \frac{SV}{\|SW\|}$$

(b) $$\frac{\partial J}{\partial W} = \frac{SSW}{\|SW\|} - SW = 0$$

since S is invertible

so $$\frac{SW}{\|SW\|} - W = 0$$

$$\|SW\| = \|SW\| \cdot \|W\|$$

$$\therefore \|W\| = 1$$