Exercise 5

group: RCSKSK

1.

(a) $Var(X) = \Sigma_{XX}$ $(d_X \times d_X)$

$Var(Y) = \Sigma_{YY}$ $(d_Y \times d_Y)$

$Cov(X,Y) = \Sigma_{XY} = \Sigma_{YX}^T$ $(d_X \times d_Y)$

$\hat{x} = W_X^T X$ $\hat{y} = W_Y^T Y$ where $W_X \in \mathbb{R}^{d_X}$ $W_Y \in \mathbb{R}^{d_Y}$ are linear projection

$Cov(\hat{x},\hat{y}) = Cov(W_X^T X, W_Y^T Y) = W_X^T Cov(X,Y) W_Y = W_X^T \Sigma_{XY} W_Y$

$Var(\hat{x}) = Var(W_X^T X) = W_X^T \Sigma_{XX} W_X$

$Var(\hat{y}) = Var(W_Y^T Y) = W_Y^T \Sigma_{YY} W_Y$

$$\Rightarrow Corr(\hat{x},\hat{y}) = \frac{Cov(\hat{x},\hat{y})}{\sqrt{Var(\hat{x}) Var(\hat{y})}} = \frac{W_X^T \Sigma_{XY} W_Y}{(W_X^T \Sigma_{XX} W_X)^{1/2} (W_Y^T \Sigma_{YY} W_Y)^{1/2}}$$

The optimization problem is

$$\max_{W_X, W_Y} \frac{W_X^T \Sigma_{XY} W_Y}{(W_X^T \Sigma_{XX} W_X)^{1/2} (W_Y^T \Sigma_{YY} W_Y)^{1/2}}$$

(b) Let $V_X = \Sigma_{XX}^{1/2} W_X$ $V_Y = \Sigma_{YY}^{1/2} W_Y$

$$Corr(\hat{x},\hat{y}) = \frac{V_X^T \Sigma_{XX}^{-1/2} \cdot \Sigma_{XY} \Sigma_{YY}^{-1/2} V_Y}{\sqrt{V_X^T V_X} \cdot \sqrt{V_Y^T V_Y}}$$

By the cauchy-Schwarz inequality, we have

$$(V_X^T \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}) V_Y \leq (V_X^T \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2} \cdot \Sigma_{YY}^{-1/2} \cdot \Sigma_{YX} \Sigma_{XX}^{-1/2} V_X)^{1/2} \cdot (V_Y^T V_Y)^{1/2}$$

$$Corr(\hat{x},\hat{y}) \leq \underbrace{\frac{(V_X^T \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2} V_X)^{1/2}}{(V_X^T V_X)^{1/2}}}_{*}$$

There is equality if V_Y and $\Sigma_{XY}^{-1/2} \Sigma_{YX} \Sigma_{XY}^{-1/2} V_X$ are collinear

Similar to PCA, the maximum of correlation (*) is attained if V_X is the eigen vector with the maximum eigenvalue for the matrix $\Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2}$

Similarly, we can also get V_Y is an eigenvector of $\Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$

Reversing the chang of coordinates, we have

W_x is an eigen vector of $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$

W_y is an eigenvector of $\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$

(c) from (*), we can get $\max \mathrm{Corr}(\hat{X}, \hat{y})_i = \sqrt{\lambda_i}$

where λ_i is the ith eigenvalue of $\Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{1/2}$

2 (a) uniform distribution

(b) $\phi(X) = \langle w^*(X), X \rangle$

where $\Sigma_{w^*}(X) = \langle X, (M_2 - M_1) \rangle \Sigma_1 + (1 - \langle X, (M_2 - M_1) \rangle) \Sigma_2$

and $w^*(X) = \Sigma_{w^*}^{-1}(X)(M_2 - M_1)$