# Homework Set 4, CPSC 8420, Fall 2020

Last Name, First Name

**Due 11/18/2020, Wednesday, 11:59PM EST**

## Problem 1

Given $X, Y, Z$, now please follow the idea/method used in LDA/PCA to find the best solution to:

$$\begin{aligned} &\underbrace{arg\ max}_{a,b} \quad a^T Z b \\ &s.t. \quad a^T X a = 1,\ b^T Y b = 1 \end{aligned} \tag{1}$$

- *Let* $a = X^{-1/2}u, b = Y^{-1/2}v$, *and we have:*

$$\begin{aligned} a^T X a = 1 \Rightarrow u^T X^{-1/2} X X^{-1/2} u = 1 \Rightarrow u^T u = 1 \\ b^T Y b = 1 \Rightarrow v^T Y^{-1/2} Y Y^{-1/2} v = 1 \Rightarrow v^T v = 1 \end{aligned} \tag{2}$$

*therefore it is equivalent to solve:*

$$\begin{aligned} &\underbrace{arg\ max}_{u,v} \quad u^T X^{-1/2} Z Y^{-1/2} v \\ &s.t. \quad u^T u = 1,\ v^T v = 1 \end{aligned} \tag{3}$$

*Apparently, if we set* $[U, \Sigma, V] = svd(X^{-1/2} Z Y^{-1/2})$, *then* $u = U[:, 1]$ *and* $v = V[:, 1]$, $u^T X^{-1/2} Z Y^{-1/2} v = u^T U \Sigma V^T v = \sigma_1$, $a = X^{-1/2}u, b = S_{YY}^{-1/2}v$.

- *By utilizing Lagrangian Multipliers, we have:*

$$J(a, b) = a^T Z b - \frac{\lambda}{2}(a^T X a - 1) - \frac{\theta}{2}(b^T Y b - 1) \tag{4}$$

*taking the derivative w.r.t.* $a, b$ *and set to be 0, we obtain:*

$$\begin{aligned} Zb - \lambda X a = 0 \\ Za - \theta Y b = 0 \end{aligned} \tag{5}$$

*Multiplying* $a^T$ *and* $b^T$ *on the left of above equations and due to the fact that* $a^T X a = 1,\ b^T Y b = 1$, *we have* $\lambda = \theta = a^T Z b$.

*Again by multiplying $X^{-1}$ and $Y^{-1}$ on the left of above equations, we get:*

$$\begin{aligned} X^{-1}Zb &= \lambda a \\ Y^{-1}Za &= \lambda b \end{aligned} \tag{6}$$

*Plugin the second equation to the first we get $X^{-1}ZY^{-1}Za = \lambda^2 a$, and plugin the first to second we get $Y^{-1}ZX^{-1}Zb = \lambda^2 b$, where we can get the optimal $a, b$ respectively.*

## Problem 2

Recall vanilla SVM objective:

$$L(w, b, \alpha) = \frac{1}{2}||w||_2^2 - \sum_{i=1}^{m} \alpha_i [y_i(w^T x_i + b) - 1] \quad s.t. \quad \alpha_i \geq 0 \tag{7}$$

If we denote the margin as $\gamma$, and vector $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]$, now please show $\gamma^2 * \|\alpha\|_1 = 1$.

$$\begin{aligned} L(w, b, \alpha) = \frac{1}{2}||w||_2^2 - &\sum_{i=1}^{m} \alpha_i [y_i(w^T x_i + b) - 1] \quad s.t. \quad \alpha_i \geq 0 \\ \frac{\partial L}{\partial w} &= 0 \Rightarrow w = \sum_{i=1}^{m} \alpha_i y_i x_i \\ \frac{\partial L}{\partial b} &= 0 \Rightarrow \sum_{i=1}^{m} \alpha_i y_i = 0 \end{aligned} \tag{8}$$

*For any support vector $x_i$, since $\langle w, x_i \rangle + b = y_i$, thus $b$ can be obtained via $b = y_i - \sum_{j=1}^{m} \alpha_i y_i \langle x_i, x_j \rangle$. Multiplying both sides by $\alpha_i y_i$ and taking the sum will give:*

$$\sum_{i=1}^{m} \alpha_i y_i b = \sum_{i=1}^{m} \alpha_i y_i^2 - \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle$$

*By the fact $\sum_{i=1}^{m} \alpha_i y_i = 0, y_i^2 = 1, w = \sum_{i=1}^{m} \alpha_i y_i x_i$, the above equation is equivalent to:*

$$0 = \sum_{i=1}^{m} \alpha_i - \|w\|^2$$

*, which says $\|\alpha\|_1 = \|w\|^2 = \frac{1}{\gamma^2} \implies \gamma^2 * \|\alpha\|_1 = 1$.*