Statistics 185 - Midterm

October 17, 2019

Name(print):____

This is a closed-book exam so do not refer to your notes or any other books. There are 6 problems.

Write your solutions as clearly as possible and make sure it's easy to find your answers (circle them if necessary), since you will not receive credit for work that I cannot understand or find. Good Luck!

If you find a question confusing, please ask me to clarify it.

If you need additional sheets, please ask and I will supply you with blank printer paper.

Question	Points
1	
2	
3	
4	
5	
6	

Total	

1. Suppose that $\vec{x} \in \mathbb{R}^p$ has mean $\vec{0}$ and covariance I. Let $A \in \mathbb{R}^{q\times p}$ be a full rank matrix with $p > q$ and let $\vec{y} = A\vec{x} + \vec{\epsilon}$ be a vector in $\mathbb{R}^q$. where $\vec{\epsilon} \sim \mathcal{N}(0, \rho^2 I_q)$ is independent of $\vec{x}$.

   a. Find Σ_y and Σ_{xy}.

   b. Find the canonical correlations between $\vec{x}$ and $\vec{y}$. Express your answer in terms of ρ^2 and the singular values of A.

2. Suppose $\vec{q}_1, \ldots, \vec{q}_k$ are orthonormal vectors in $\mathbb{R}^d$ with $k < d$. Suppose

$$\vec{x} = a_1\vec{q}_1 + \cdots + a_k\vec{q}_k + \vec{\epsilon}$$

where $a_1, \ldots, a_k$ are uncorrelated, mean zero random variables with variance λ^2 and $\vec{\epsilon}$ is independent of $a_1, \ldots, a_k$ and follows a multivariate Normal distribution with mean $\vec{0}$ and covariance $\sigma^2 I$. Here, I is the $d-$dimensional identity matrix.

a. Find the mean of $\vec{x}$.

b. Find the covariance matrix, Σ, of $\vec{x}$.

c. Give the eigenvalues and respective eigenvectors of Σ. (Hint: you may assume there are orthonormal vectors $\vec{q}_{k+1}, \ldots, \vec{q}_d$ which are a basis for the orthogonal complement of the span of $\vec{q}_1, \ldots, \vec{q}_d$.)

3. Let $\vec{x}_1, \ldots, \vec{x}_N$ be vectors in $\mathbb{R}^d$. Assume a PCA of these data has loadings $\vec{w}_1, \ldots, \vec{w}_d$ with associated variances $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$. Let U be a $d \times d$ orthonormal matrix and set $\vec{y}_i = U\vec{x}_i$. Find expressions for the principal components loadings and variance of $\vec{y}_1, \ldots, \vec{y}_N$ in terms of U, $\vec{w}_1, \ldots, \vec{w}_d$ and $\lambda_1, \ldots, \lambda_d$.

4. Let

$$\|X - WH\|_F = \sqrt{\sum_{i,j} [X_{ij} - (WH)_{ij}]^2}$$

where $X \in \mathbb{R}^{N \times p}$, $W \in \mathbb{R}^{N \times k}$, and $H \in \mathbb{R}^{k \times p}$.

a. Find the partial derivative of $\frac{1}{2}\|X - WH\|_F^2$ with respect to W_{ij}.

b. Find the partial derivative of $\frac{1}{2}\|X - WH\|_F^2$ with respect to H_{jl}.

5. Suppose $\vec{x} \in \mathbb{R}^p$ and $\vec{y} \in \mathbb{R}^q$ are random vectors with covariance matrices Σ_x and Σ_y respectively. Suppose $\vec{x}$ and $\vec{y}$ have cross-covariance matrix Σ_{xy} with rank $k \leq \min\{p, q\}$ and canonical correlation variable $\eta_1, \ldots, \eta_k$ and $\xi_1, \ldots, \xi_k$ with canonical correlations $\sigma_1 \geq \cdots \geq \sigma_k > 0$.

   a. Find the covariance matrices of the random vectors $\vec{\eta} = (\eta_1, \ldots, \eta_k)^T$ and $\vec{\xi} = (\xi_1, \ldots, \xi_k)^T$.

   b. Find the cross-covariance of $\vec{\eta}$ and $\vec{\xi}$.

6. Let $\vec{x}_1, \ldots, \vec{x}_N$ be random vectors in $\mathbb{R}^p$ with sample covariance $\hat{\Sigma}_X$ and $\vec{y}_1, \ldots, \vec{y}_N$ be random vectors in $\mathbb{R}^q$ with covariance matrix $\hat{\Sigma}_Y$. Assume that $\hat{\Sigma}_X$ and $\hat{\Sigma}_Y$ are both full rank, i.e. the principal component variances of both $\vec{x}_1, \ldots, \vec{x}_n$ and the principal component variances of $\vec{y}_1, \ldots, \vec{y}_N$ are all positive.

   Let $\vec{z}_i^T = (\vec{x}_i^T, \vec{y}_i^T)$ for $i = 1, \ldots, N$ be random vectors in $\mathbb{R}^{p+q}$ with sample covariance $\hat{\Sigma}_Z$. Now assume $\hat{\Sigma}_Z$ is not full rank, i.e. at least one of the principal component variances of $\vec{z}_1, \ldots, \vec{z}_N$ is zero. Is the largest canonical correlation between $\{\vec{x}_i\}_{i=1}^N$ and $\{\vec{y}_i\}_{i=1}^N$ equal to 1? Explain your answer.

   Hint: If $A \in \mathbb{R}^{p \times p}$, $B \in \mathbb{R}^{p \times q}$, $C \in \mathbb{R}^{q \times p}$, and $D \in \mathbb{R}^{q \times q}$ are matrices and A is invertible then

$$det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \right) = det(A) det(D - CA^{-1}B).$$

   You do not need this hint. It just offers one route to a solution.