Ecole Polytechnique Fédérale, Lausanne: Fall 2019
December 13, 2019

Problem Set 7 — *Not Fully Graded*

For the Exercise Sessions on Dec 13 and Dec 20

Last name	First name	SCIPER Nr	Points

Problem 1: Canonical Correlation Analysis

Let $\mathbf{X}$ and $\mathbf{Y}$ be zero-mean real-valued random vectors with covariance matrices $R_{\mathbf{X}}$ and $R_{\mathbf{Y}}$, respectively. Moreover, let $R_{\mathbf{XY}} = \mathbb{E}[\mathbf{X}\mathbf{Y}^T]$. Our goal is to find vectors $\mathbf{u}$ and $\mathbf{v}$ such as to maximize the correlation between $\mathbf{u}^T\mathbf{X}$ and $\mathbf{v}^T\mathbf{Y}$, that is,

$$\max_{\mathbf{u},\mathbf{v}} \frac{\mathbb{E}[\mathbf{u}^T\mathbf{X}\mathbf{Y}^T\mathbf{v}]}{\sqrt{\mathbb{E}[|\mathbf{u}^T\mathbf{X}|^2]}\sqrt{\mathbb{E}[|\mathbf{v}^T\mathbf{Y}|^2]}}. \tag{1}$$

Show how we can find the optimizing choices of the vectors $\mathbf{u}$ and $\mathbf{v}$ from the problem parameters $R_{\mathbf{X}}, R_{\mathbf{Y}}$, and $R_{\mathbf{XY}}$.

Hint: Recall for the singular value decomposition that

$$\max_{\mathbf{v}} \frac{\|A\mathbf{v}\|}{\|\mathbf{v}\|} = \max_{\|\mathbf{v}\|=1} \|A\mathbf{v}\| = \sigma_1(A), \tag{2}$$

where $\sigma_1(A)$ denotes the maximum singular value of the matrix A. The corresponding maximizer is the right singular vector $\mathbf{v}_1$ (i.e., eigenvector of A^TA) corresponding to $\sigma_1(A)$.

Problem 2: Some review problems on linear algebra

(a) (Frobenius norm) Prove that $\|A\|_F^2 = \operatorname{trace}(A^HA)$.

(b) (Singular Value Decomposition) Let $\sigma_i(A)$ denote the i^{th} singular value of an $m \times n$ matrix A. Prove that $\|A\|_F^2 = \sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A)$

(c) (Projection Matrices) Consider a set of k orthonormal vectors in $\mathbb{C}^n$, denoted by $\mathbf{u_1}, \mathbf{u_2}, \cdots, \mathbf{u_k}$. The projection matrix (that projects an arbitrary vector into the subspace spanned by these orthonormal vectors) is given by

$$P = \sum_{i=1}^{k} \mathbf{u}_i\mathbf{u}_i^H. \tag{3}$$

- Prove that this matrix is *Hermitian,* i.e., $P^H = P$.
- Prove that this matrix is *idempotent,* i.e., $P^2 = P$. (In words, projecting twice into the same subspace is the same as projecting only once.)
- Prove that $\operatorname{trace}(P) = k$, i.e., equal to the dimension of the subspace.
- Prove that the diagonal entries of P must be real-valued and non-negative. Then, prove that the diagonal entries of P cannot be larger than 1 (this is a little more tricky).

Problem 3: Eckart–Young Theorem

In class, we proved the converse part of the Eckart–Young theorem for the spectral norm. Here, you do the same for the case of the Frobenius norm.

(a) For any matrix A of dimension $m \times n$ and an arbitrary orthonormal basis $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ of $\mathbb{C}^n$, prove that

$$\|A\|_F^2 \quad = \quad \sum_{k=1}^{n} \|A\mathbf{x}_k\|^2. \tag{4}$$

(b) Consider any $m \times n$ matrix B with $\text{rank}(B) \leq p$. Clearly, its null space has dimension no smaller than $n-p$. Therefore, we can find an orthonormal set $\{\mathbf{x}_1, \cdots, \mathbf{x}_{n-p}\}$ in the null space of B. Prove that for such vectors, we have

$$\|A - B\|_F^2 \quad \geq \quad \sum_{k=1}^{n-p} \|A\mathbf{x}_k\|^2. \tag{5}$$

(c) (This requires slightly more subtle manipulations.) For any matrix A of dimension $m \times n$ and any *orthonormal* set of $n-p$ vectors in $\mathbb{C}^n$, denoted by $\{\mathbf{x}_1, \cdots, \mathbf{x}_{n-p}\}$, prove that

$$\sum_{k=1}^{n-p} \|A\mathbf{x}_k\|^2 \quad \geq \quad \sum_{j=p+1}^{r} \sigma_j^2. \tag{6}$$

Hint: Consider the case $m \geq n$ and the set of vectors $\{\mathbf{z}_1, \cdots, \mathbf{z}_{n-p}\}$, where $\mathbf{z}_k = V^H \mathbf{x}_k$. Express your formulas in terms of these and the SVD representation $A = U\Sigma V^H$.

(d) Briefly explain how *(a)*-*(c)* imply the desired statement.

Problem 4: Inner Products

Consider the standard n-dimensional vector space $\mathbb{R}^n$.

1. Characterize the set of matrices W for which $\mathbf{y}^T W \mathbf{x}$ is a valid inner product for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
2. Prove that *every* inner product $\langle \mathbf{x}, \mathbf{y} \rangle$ on $\mathbb{R}^n$ can be expressed as $\mathbf{y}^T W \mathbf{x}$ for an approriately chosen matrix W.
3. For a subspace of dimension $k < n$, spanned by the basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k \in \mathbb{R}^n$, express the orthogonal projection operator (matrix) with respect to the general inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T W \mathbf{x}$. *Hint:* For any vector $\mathbf{x} \in \mathbb{R}^n$, express its projection as $\widehat{\mathbf{x}} = \sum_{j=1}^{k} \alpha_j \mathbf{b}_j$.