SI231B: Matrix Computations, 2024 Spring

Homework Set #1

Acknowledgements:

1) Deadline: **2024-03-21 23:59:59**
2) You have 5 "free days" in total for late submission.

Problem 1. (Range Space and Rank) (15 points)

Given $\mathbf{A} \in \mathbb{R}^{m \times k}$ and $\mathbf{B} \in \mathbb{R}^{k \times n}$, prove that

1) $\dim \mathcal{R}(\mathbf{AB}) \leq \dim \mathcal{R}(\mathbf{A})$. (3 points)
2) $\dim \mathcal{R}(\mathbf{AB}) = \dim \mathcal{R}(\mathbf{A})$ if $\mathbf{B}$ has full row rank. (3 points)
3) Based on the above two results, show that

$$\mathsf{rank}(\mathbf{AB}) \leq \min\{\mathsf{rank}(\mathbf{A}), \mathsf{rank}(\mathbf{B})\}$$

   and the equality attains if the columns of $\mathbf{A}$ are linearly independent or the rows of $\mathbf{B}$ are linearly independent. (5 points)

   Hint: It suffices to show $\dim \mathcal{R}(\mathbf{AB}) \leq \dim \mathcal{R}(\mathbf{A})$ and $\dim \mathcal{R}(\mathbf{B}^T\mathbf{A}^T) \leq \dim \mathcal{R}(\mathbf{B}^T)$.
4) $\dim \mathcal{R}(\mathbf{AB}) = k$ if $\mathbf{A}$ has full column rank and $\mathbf{B}$ has full row rank. (4 points)

Problem 2. (Vector Norms) (15 points)

Recall that we talked about the vector norm in class. For any $\mathbf{x} = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$, please prove the following arguments:

1) The maximum norm is defined to be
$$\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}.$$
Show that
$$\|\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} |x_i|.$$
(5 points)

2) Verify the following inequality chain
$$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2 \le \|\mathbf{x}\|_1 \le \sqrt{n}\,\|\mathbf{x}\|_2 \le n\,\|\mathbf{x}\|_\infty.$$
(5 points)

3) Show that
$$\|\mathbf{x}\|_1 \|\mathbf{x}\|_\infty \le \frac{1+\sqrt{n}}{2}\|\mathbf{x}\|_2^2.$$
(5 points)

Problem 3. (Direct Sum of Subspaces) (20 points)

A vector space V is the *(internal) direct sum* of a family $\mathcal{F} = \{S_1, \cdots, S_n\}$ of subspaces of V, written $V = \oplus_{i=1}^n S_i = S_1 \oplus \cdots \oplus S_n$ if the following two conditions hold:

- V is the sum of the family $\mathcal{F}$: $V = \sum_{i=1}^n S_i$,
- for each i, $S_i \cap \left(\sum_{j\neq i} S_j\right) = \{0\}$.

For example, a vector space V is the direct sum of a subspace S and its orthogonal complement $S^\perp$: $V = S \oplus S^\perp$. $\mathbb{R}^3$ is the direct sum of any three non-coplanar lines.

1) Determine which of the following sums are direct sums. Briefly explain why.
   a) $\mathbb{R}^3$ is the sum of any two distinct planes. (1 point)
   b) $\mathbb{R}^{n\times n}$ is the sum of the subspace of upper-triangular matrices and the subspace of lower-triangular matrices. (2 points)
   c) $\mathbb{R}^{n\times n}$ is the sum of the subspace of symmetric matrices and the subspace of skew-symmetric matrices. (A square matrix $\mathbf{A}$ is called skew-symmetric if $\mathbf{A} = -\mathbf{A}^T$.) (2 points)
2) Let $\mathcal{F} = \{S_1, \cdots, S_n\}$ be a family of distinct subspaces of a finite-dimensional vector space V such that $V = \sum_{i=1}^n S_i$. Prove that the following statements are equivalent:
   a) **(Independence of the family)** For each i, $S_i \cap \left(\sum_{j\neq i} S_j\right) = \{0\}$.
   b) **(Uniqueness of expression for 0)** The zero vector 0 cannot be written as a sum of nonzero vectors from distinct subspaces of $\mathcal{F}$.
   c) **(Uniqueness of expression)** Every nonzero $v \in V$ has a unique, except for order of terms, expression as a sum $v = s_1 + \cdots + s_n$ of nonzero vectors from distinct subspaces in $\mathcal{F}$.

   (So a sum $V = \sum_{i=1}^n S_i$ is direct if and only if any one of the above statements holds.)

   Hint: to prove the equivalence, you could show that, for example, a) implies b) and b) implies c) and c) implies a). (15 points)

Problem 4. (Grassmann Formula) (10 points)

Let S and T be subspaces of a finite-dimensional vector space V. Please prove the following formula

$$\dim(S) + \dim(T) = \dim(S + T) + \dim(S \cap T).$$

In particular, if T is any complement of S in V, then

$$\dim(S) + \dim(T) = \dim(V),$$

that is, the dimensions of vector spaces are additive in a direct sum:

$$\dim(S \oplus T) = \dim(S) + \dim(T).$$

Hint: to prove the first formula, you might start by considering a basis for $S \cap T$, and extending it to bases for S and T separately. With these bases, then consider how to construct a basis for $S + T$ and prove that it is a basis indeed.

Problem 5. (Block Matrix Computations) (20 points)

Let $\mathbf{S}$ as the sample covariance matrix of n independent observed samples of a p-variate Gaussian random variable with zero mean and covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$. Denote σ_{ij} as the ij-th entry in $\mathbf{\Sigma}$. Suppose four elements $\{\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}\}$ in $\mathbf{\Sigma}$ are missing as

$$\mathbf{\Sigma} = \begin{bmatrix} ? & ? & \cdots \\ ? & ? & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

Our goal is to estimate the missing entries based on the observed ones in $\mathbf{\Sigma}$ and the sample covariance matrix $\mathbf{S}$. Given

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{AA} & \mathbf{\Sigma}_{AB} \\ \mathbf{\Sigma}_{AB} & \mathbf{\Sigma}_{BB} \end{bmatrix}, \quad \text{with} \quad \mathbf{\Sigma}_{AA} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}.$$

We are interested in the following Gaussian maximum likelihood estimation problem for $\mathbf{\Sigma}_{AA}$:

$$\min_{\mathbf{\Sigma}_{AA}} \quad \mathrm{tr}\left(\mathbf{S}\mathbf{\Sigma}^{-1}\right) + \log\det\left(\mathbf{\Sigma}\right). \tag{1}$$

1) If $\mathbf{\Sigma}_{AA}$ is invertible, it easy to verify that the above 2×2 partitioned $\mathbf{\Sigma}$ has the following factorization form

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{\Sigma}_{BA}\mathbf{\Sigma}_{AA}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{AA} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{BB} - \mathbf{\Sigma}_{BA}\mathbf{\Sigma}_{AA}^{-1}\mathbf{\Sigma}_{AB} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{\Sigma}_{AA}^{-1}\mathbf{\Sigma}_{AB} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{2}$$

(Note: $\mathbf{\Sigma}_{BB} - \mathbf{\Sigma}_{BA}\mathbf{\Sigma}_{AA}^{-1}\mathbf{\Sigma}_{AB}$ is named *Schur complement* of $\mathbf{\Sigma}_{AA}$ in the matrix literature.)

a) Based on the above factorization result, compute the inverse and determinant of $\mathbf{\Sigma}$. (5 points)

b) Write the objective function in (1) explicitly as a function of variable $\mathbf{\Sigma}_{AA}$. (5 points)

Hint: You may need to partition $\mathbf{S}$ as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{AA} & \mathbf{S}_{AB} \\ \mathbf{S}_{BA} & \mathbf{S}_{BB} \end{bmatrix},$$

where $\mathbf{S}_{AA}$, $\mathbf{S}_{AB}$, $\mathbf{S}_{BA}$ and $\mathbf{S}_{BB}$ take the same dimension as $\mathbf{\Sigma}_{AA}$, $\mathbf{\Sigma}_{AB}$, $\mathbf{\Sigma}_{BA}$, and $\mathbf{\Sigma}_{BB}$, respectively.

2) Directly solving for $\mathbf{\Sigma}_{AA}$ based on the objective function derived above is difficult, one alternative way is to solve for an intermediate variable $\tilde{\mathbf{\Sigma}}_{AA} = \mathbf{\Sigma}_{AA} - \mathbf{\Sigma}_{AB}\mathbf{\Sigma}_{BB}^{-1}\mathbf{\Sigma}_{BA}$.

a) Mimicking the form of (2), decompose $\mathbf{\Sigma}$ into the product of three matrices, where $\tilde{\mathbf{\Sigma}}_{AA}$ shows up. (4 points)

b) Try to derive the objective function in terms of $\tilde{\mathbf{\Sigma}}_{AA}$ and discuss why the resulting problem is easier than the one you derived in 1.b). (6 points)

Hint: $\tilde{\mathbf{\Sigma}}_{AA}$ is the Schur complement of $\mathbf{\Sigma}_{BB}$. Deriving the objective function in terms of $\tilde{\mathbf{\Sigma}}_{AA}$ may require a similar proof as that in the previous question.

Problem 6. (Determinant) (20 points)

Consider the following matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 6 & 2 & 1 \\ -2 & 2 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 0 & -2 & 3 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & 1 \\ 0 & 2 & -3 & 0 \end{bmatrix}.$$

1) Use a cofactor expansion to evaluate the determinants $\det(\mathbf{A})$ and $\det(\mathbf{B})$. (10 points)
2) Use determinants and adjugate matrices to compute the inverses $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$. (10 points)