Biostat 216 Homework 7

Due Nov 18 Friday @ 11:59pm

Submit a PDF (scanned/photographed from handwritten solutions, or converted from RMarkdown or Jupyter Notebook) to Gracescope on BruinLearn.

SVD

- Q1. Find the closest rank-1 approximation (in Frobenius norm or spectral norm) to these matrices
$$\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 & 3 \\ 2 & 0 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

- Q2. Moore-Penrose inverse from SVD.
    1. With singular value decomposition $\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}'$, verify that
    $$\mathbf{X}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}' = \mathbf{V}_r\boldsymbol{\Sigma}_r^{-1}\mathbf{U}_r' = \sum_{i=1}^r \sigma_i^{-1}\mathbf{v}_i\mathbf{u}_i',$$
    where $\boldsymbol{\Sigma}^+ = \text{diag}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0)$ and $r = \text{rank}(\mathbf{X})$, satifies the four properties of the [Moore-Penrose inverse (https://ucla-biostat-216.github.io/2022fall/slides/06-matinv/06-matinv.html#Generalized-inverse-and-Moore-Penrose-inverse-(optional))](https://ucla-biostat-216.github.io/2022fall/slides/06-matinv/06-matinv.html#Generalized-inverse-and-Moore-Penrose-inverse-(optional)).
    2. Show that $\text{rank}(\mathbf{X}^+) = \text{rank}(\mathbf{X})$.
    3. Show that $\mathbf{X}^+\mathbf{X}$ is the orthogonal projector into $\mathcal{C}(\mathbf{X}')$ and $\mathbf{X}\mathbf{X}^+$ is the orthogonal projector into $\mathcal{C}(\mathbf{X})$.
    4. Show that $\boldsymbol{\beta}^+ = \mathbf{X}^+\mathbf{y}$ is a minimizer of the least squares criterion $f(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2$. Hint: check $\boldsymbol{\beta}^+$ satisfies the normal equation $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{y}$.
    5. Show that $\boldsymbol{\beta}^+ \in \mathcal{C}(\mathbf{X}')$.
    6. Show that if another $\boldsymbol{\beta}^\star$ minimizes $f(\boldsymbol{\beta})$, then $\|\boldsymbol{\beta}^\star\| \ge \|\boldsymbol{\beta}^+\|$. This says that $\boldsymbol{\beta}^+ = \mathbf{X}^+\mathbf{y}$ is the least squares solution with smallest L2 norm. Hint: since both $\boldsymbol{\beta}^\star$ and $\boldsymbol{\beta}^+$ satisfy the normal equation, $\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^\star = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}^+$ and deduce that $\boldsymbol{\beta}^\star - \boldsymbol{\beta}^+ \in \mathcal{N}(\mathbf{X})$.

- Q3. Let $\mathbf{B}$ be a submatrix of $\mathbf{A} \in \mathbb{R}^{m\times n}$. Show that the largest singular value of $\mathbf{B}$ is always less than or equal to the largest singular value of $\mathbf{A}$.

- Q4. Show that all three matrix norms (ℓ_2, Frobenius, nuclear) are invariant under orthogonal transforms. That is
$$\|\mathbf{Q}_1\mathbf{A}\mathbf{Q}_2'\| = \|\mathbf{A}\| \text{ for orthogonal } \mathbf{Q}_1 \text{ and } \mathbf{Q}_2.$$

Optimization and multivariate calculus

- Q5.
    1. Explain why the intersection $K_1 \cap K_2$ of two convex sets is a convex set.
    2. Prove that the maximum F_3 of two convex functions F_1 and F_2 is a convex function. Hint: What is the set above the graph of F_3?

- Q6. Show that these functions are convex:
    1. Entropy $x \log x$.
    2. $\log(e^x + e^y)$.
    3. ℓ_p norm $\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p)^{1/p}$, $p \geq 1$.
    4. $\lambda_{\max}(\mathbf{S})$ as a function of the symmetric matrix $\mathbf{S}$. Hint: HW6 Q10.6 and Q5.2.

- Q7. Minimize $f(x_1, x_2) = \frac{1}{2}\mathbf{x}'\mathbf{S}\mathbf{x} = \frac{1}{2}x_1^2 + 2x_2^2$ subject to the constraint $\mathbf{A}'\mathbf{x} = x_1 + 3x_2 = b$.

    1. What is the Lagrangian $L(\mathbf{x}, \lambda)$ for this problem.
    2. What are the three equations "derivative of L=zero"?
    3. Solve these equations to find $\mathbf{x}^\star = (x_1^\star, x_2^\star)$ and the multiplier $\lambda^\star$.
    4. Verify that the derivative of the minimum cost is $\partial f^\star / \partial b = -\lambda^\star$.