Homework 8

due Tues, April 4, 11:59 PM

(Upload your solutions to Gradescope)

---

1. (5 points) Suppose [illegible] matrix of rank r. In lecture 23, we stated that the two "defining properties" of [illegible] A^+ are that

(1) AA^+ is the orthog[illegible] $\mathcal{R}(A)$.

(2) $\mathcal{R}(A^+) \subset \mathcal{N}(A)^\perp$.

Derive from these two properties that A^+A is the orthogonal projection onto $\mathcal{N}(A)^\perp$. *(Hint: consider A^+Ax for $x \in \mathcal{N}(A)$ and $x \in \mathcal{N}(A)^\perp$ separately.)*

---

2. (5 points) (Variant of I.9.6, page 80, Strang.) Suppose $A = Q\Lambda Q^T$ is a real, symmetric $n \times n$ matrix with $Q = (q_1, \dots, q_n)$ orthogonal and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Suppose $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_n|$. For what values of a_1 is

$$B = q_1 a_1 q_1^T$$

a closest rank-1 approximation of A in the 2-norm? Justify your answer, and keep in mind that λ_1 can be positive or negative.

---

3. (5 points) Compute the polar decomposition $A = Q|A|$ of the 3×3 matrix

$$A = \begin{pmatrix} -2 & 3 & -2 \\ 0 & 2 & -4 \\ -2 & 2 & 0 \end{pmatrix}.$$

Here Q is a partial isometry with the same nullspace as A and $|A|$ is positive semi-definite. *Notes: to compute a fully reduced SVD by hand, one can either (1) solve the eigenvalue problem $A^TA = V\Lambda V^T$, discard columns of V and rows/columns of Λ corresponding to zero eigenvalues, and set $S = \sqrt{\Lambda}$, $U = AVS^{-1}$; or (2) start with a CR factorization, which I'll call CF^T here, and then compute the QR factorizations of C and F:*

$$A = CF^T = (Q_1R_1)(Q_2R_2)^T = Q_1(R_1R_2^T)Q_2^T.$$

You can now compute the SVD of $A_3 = R_1R_2^T = U_3S_3V_3^T$, which involves a smaller eigenvalue problem. Eventually, $A = Q_1U_3S_3V_3^TQ_2^T$. In this problem, both approaches are somewhat tedious, so feel free to use a computer to compute the SVD of A or carry out any of the steps described above. Express your answer with rational numbers in the matrix entries of Q and $|A|$.

---

4. (5 points) (Variant of I.11.2, page 96, Strang.) Prove the Cauchy-Schwarz inequality

$$|\langle u, v\rangle| \le \|u\|\|v\|, \qquad u, v \in \mathbb{C}^n,$$

where $\langle u, v\rangle = \overline{u}^T v$ and $\|u\| = \sqrt{\langle u, u\rangle}$. Hint: if $v \neq 0$, the projection of u onto $\operatorname{span}(v)^\perp$ is

$$[illegible] = u - \frac{\langle v, u\rangle}{\langle v, v\rangle} v.$$

Simplify the equation $\langle$[illegible] the Cauchy-Schwarz inequality when $v \neq 0$. (Be careful: $\langle v, u\rangle$ is a complex num[illegible]0 case separately.

---

5. (5 points) In problem 2 of homework 2, we proved that the 1-norm of an $m \times n$ matrix A is the maximum absolute column sum of A. Now show that the ∞-norm of an $m \times n$ matrix is the maximum absolute row sum, i.e., $\|A\|_\infty = C$ with $C = \max_{1\le i\le m} \sum_{j=1}^n |a_{ij}|$. You just have to (i) show that $\|Ax\|_\infty \le C\|x\|_\infty$ for all $x \in \mathbb{C}^n$; and (ii) produce an $x \neq 0$ (tailored to A) such that $\|Ax\|_\infty \ge C\|x\|_\infty$.

---

6. (5 points) The space of real $m \times n$ matrices $M^{m\times n}$ is commonly given the operator 2-norm, $\|B\|_2 = \sigma_1(B)$, where $\sigma_1(B)$ is the largest singular value of B. The most general linear functional on this space is obtained by taking a "double dot product" of B with a fixed matrix A, namely

$$f_A(B) = A : B = \sum_{k=1}^m \sum_{j=1}^n a_{kj} b_{kj} = \operatorname{tr}(A^T B).$$

Note that $f_A : M^{m\times n} \to \mathbb{R}$ is a linear mapping. Show that the operator norm of f_A is given by $\|f_A\| = \|A\|_N$, where

$$\|A\|_N = \sum_{j=1}^{\mu} \sigma_j(A), \qquad \big(\mu = \min(m, n)\big).$$

$\|A\|_N$ is known as the **nuclear norm** of A. At this point it is just a formula involving the singular values of A; this exercise shows that it is actually a norm. (The only difficult property to check for the nuclear norm is the triangle inequality, which automatically holds for operator norms: $\|A_1 + A_2\|_N = \|f_{A_1+A_2}\| = \|f_{A_1} + f_{A_2}\| \le \|f_{A_1}\| + \|f_{A_2}\| = \|A_1\|_N + \|A_2\|_N$.) Fixing A, you just have to (i) show that $|f_A(B)| \le \|A\|_N\|B\|_2$ for all $m \times n$ matrices B, and (ii) find a particular matrix $B \neq 0$ such $|f_A(B)| = \|A\|_N\|B\|_2$.

(Hint: for part (i), let $A = USV^T$ be the full SVD of A. Then

$$|\operatorname{tr}(A^T B)| = |\operatorname{tr}(VS^TU^TB)| = |\operatorname{tr}(S^T\widetilde{B})| \le \sum_{j=1}^{\mu} \sigma_j(A)|\tilde{b}_{jj}|,$$

where $\widetilde{B} = U^TBV$. Explain the last inequality and show that $\max_j |\tilde{b}_{jj}| \le \sigma_1(B)$. Once you understand part (i), part (ii) is straightforward.)