PS #2 — Matrix norms and conditioning

Problem 1. *Part (a):* Using the definition of matrix-matrix multiplication, show that

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}) \tag{1}$$

for square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$.

Part (b): Explain why this property implies that for any set of square matrices $\{\mathbf{A}_i\}_{i=1}^n$ the trace satisfies the cyclic property

$$\mathrm{tr}(\mathbf{A}_{\sigma(1)} \cdots \mathbf{A}_{\sigma(n)}) = \mathrm{tr}(\mathbf{A}_1 \cdots \mathbf{A}_n) \tag{2}$$

where $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ is cyclic permutation of the indices.

Problem 2. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Use the definition of matrix-matrix multiplication to show that

$$\|\mathbf{A}\|_F^2 = \mathrm{tr}(\mathbf{A}^T \mathbf{A}). \tag{3}$$

Problem 3. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Show that the Frobenius norm satisfies

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^r \sigma_i^2} \tag{4}$$

where $\{\sigma_i\}$ are the singular values of $\mathbf{A}$ and $r = \mathrm{rank}(\mathbf{A})$. *Hint: use the results of Problem 1 and Problem 2.*

Problem 4. Show that the Frobenius and induced matrix 2-norms satisfy

$$\|\mathbf{A}\|_2 \le \|\mathbf{A}\|_F \le \sqrt{\mathrm{rank}(\mathbf{A})}\, \|\mathbf{A}\|_2 \tag{5}$$

for all $\mathbf{A} \in \mathbb{R}^{m \times n}$, i.e., these norms are equivalent. *Bonus:* For what class of matrices does $\|\mathbf{A}\|_2 = \|\mathbf{A}\|_F$?

Problem 5. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be an invertible matrix. Let $\mathbf{A} = \mathbf{U\Sigma V}^T$ be its SVD with the diagonal entries of $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$ given in descending order. The Eckart-Young theorem states that

$$\operatorname*{arg\,min}_{\mathbf{Z} \in \mathbb{R}^{n \times n} \,:\, \mathrm{rank}(\mathbf{Z}) = k} \|\mathbf{A} - \mathbf{Z}\|_2 = \mathbf{A}_k, \tag{6}$$

where $\mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ is the truncated SVD of rank k.

Part (a): Use this theorem to show that

$$\min_{\mathbf{X} \in \mathbb{R}^{n \times n} \,:\, \mathbf{X} \text{ singular}} \|\mathbf{A} - \mathbf{X}\|_2 = \sigma_n, \tag{7}$$

i.e., that the smallest singular value measures the absolute distance from $\mathbf{A}$ to the nearest singular matrix.

Part (b): Use this theorem to show that

$$\min_{\mathbf{X}\in\mathbb{R}^{n\times n}\,:\,\mathbf{X}\text{ singular}} \frac{\|\mathbf{A}-\mathbf{X}\|_2}{\|\mathbf{A}\|_2} = \frac{1}{\kappa(\mathbf{A})}, \tag{8}$$

i.e., that the reciprocal of the condition number measures the relative distance from $\mathbf{A}$ to the nearest singular matrix.

Problem G1. Prove the Eckart-Young theorem.

Problem G2. Prove Lemma 3.1 of *Accuracy and Stability of Numerical Algorithms* by Nicholas Higham (2002). Then, use this result to prove backward stability of the inner product operation.