ECS231 Problem Set #2 — due 4pm, April 30

1. Let $\|\cdot\|$ be a vector norm on $\mathbb{R}^m$ and assume that $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank}(A) = n$, prove that $\|x\|_A = \|Ax\|$ is a vector norm on $\mathbb{R}^n$.

2. Verify that $\|xy^H\|_F = \|xy^H\|_2 = \|x\|_2\|y\|_2$ for any $x, y \in \mathbb{C}^n$.

3. (a) Define $\|A\|_{\max} = \max\limits_{1\le i,j\le n}\{|a_{ij}|\}$. Show that $\|A\|_{\max}$ is a matrix norm.

   (b) Show by example that it is not a consistent norm, i.e., it does not satisfy the property $\|AB\|_{\max} \le \|A\|_{\max}\,\|B\|_{\max}$.

4. Given $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ sthe *trace* of A is $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$. Prove the following facts about the trace:

   (1) $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$.

   (2) $\operatorname{tr}(CE) = \operatorname{tr}(EC)$ for all $n\times m$ C and $m\times n$ E.

   (3) Show that if B is similar to A, i.e., $B = S^{-1}AS$ for a nonsingular matrix S, then $\operatorname{tr}(B) = \operatorname{tr}(A)$.

   (4) Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A, then $\operatorname{tr}(A) = \sum_{j=1}^{n}\lambda_j$.

5. Let Q and Z be orthogonal matrices.

   (a) Show that $\|QAZ\|_F = \|A\|_F$.

   (b) Show that $\|QAZ\|_2 = \|A\|_2$, where $\|A\|_2 = \max\limits_{x\ne 0}\dfrac{\|Ax\|_2}{\|x\|_2}$.

   Therefore, we can conclude that the Frobenius norm and 2-norm are invariant under orthogonal transformation.

6. Let A be an $n\times n$ nonsingular matrix, U and V be $n\times k$ matrices with $k \le n$, verify that if $T = I + V^{\mathrm{H}}A^{-1}U$ is nonsingular, then

$$(A + UV^{\mathrm{H}})^{-1} = A^{-1} - A^{-1}UT^{-1}V^{\mathrm{H}}A^{-1}.$$

   This formula is called the *Sherman-Morrison-Woodbury* formula, widely used in applications.

7. Given $n\times n$ nonsingular matrix A and $m\times m$ matrix D, show that

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\cdot\det(D - CA^{-1}B).$$

   The matrix $D - CA^{-1}B$ is called the Schur complement of A in $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

8. (1). Let A be symmetric positive definite and suppose that $A = LDL^T$, where L is unit lower triangular and D is diagonal. Prove that the main-diagonal entries of D are all positive.

   (2). Conversely, suppose $A = LDL^T$, where L is unit lower triangular and D is diagonal. Prove that if the main-diagonal entries of D are positive, then A is positive definite.

9. If A is a nonsingular symmetric matrix, and has the factorization $A = LDM^T$, where L and M are unit lower triangular matrices and D is a diagonal matrix, show that $L = M$.

10. Given an $n \times n$ nonsingular matrix A, how do you *efficiently* solve the following problems by using the LU with partial pivoting:

    (1) solve the linear system $A^k x = b$, where k is a positive integer, $b \in \mathrm{R}^n$.

    (2) solve the matrix equation $AX = B$, where $B \in \mathrm{R}^{n\times m}$.

    You should (i) describe your algorithms, (ii) present them in pseudo-code (Matlib-like language. You may use the LU with partial pivoting as a building block), and (iii) give the required flops (The flops of the LU with partial pivoting is $2/3n^3$).

11. Let A be nonsingular, and let $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ be its singular values.

    (1) Find the SVD of A^{-1} in terms of the SVD of A. What are the singular values and singular vectors of A^{-1}?

    (2) Deduce that $\|A^{-1}\|_2 = \sigma_n^{-1}$ and $\kappa(A) = \|A\|_2 \|A^{-1}\|_2 = \sigma_1/\sigma_n$. $\kappa(A)$ is referred to as the condition number of A.

12. (Mini-project) The *nuclear norm* of $A \in \mathbb{R}^{n\times n}$ is defined by

$$\|A\|_* = \sum_{i=1}^{n} \sigma_i$$

    where σ_i is the i-th singular value of A.

    (a) Show that $\|A\|_* = \mathrm{tr}(\sqrt{A^T A})$. Here the square root of a symmetric positive semidefinite matrix M is defined by $\sqrt{M} = Q\Lambda^{1/2}Q^T$, $M = Q\Lambda Q^T$ is a spectral decomposition of M and $\Lambda^{1/2} = \mathrm{diag}(\lambda_i^{1/2})$.

    (b) Show that $\|A\|_* = \max\limits_{Q^T Q = I} \mathrm{tr}(AQ)$.

    (c) Show that $\|A + B\|_* \leq \|A\|_* + \|B\|_*$

    (c) The nuclear norm is popular in machine learning. Provide an application of the nuclear norm.