Homework set 2 — CSE 383C / CS 383C / M 383E / ME 397, Fall 2024

P.G. Martinsson, University of Texas at Austin

Hand in solutions to: 4.2, 5.3, 5.4, 6.2, 6.5 from the book. Problems 1 & 2 below.

Suggested problems: 4.1, 4.4, 5.1, 5.2, 6.1, 6.4 from the book. (Do not hand in solutions to these!)

Problem 1: Suppose that $\mathbf{A} \in \mathbb{C}^{m\times n}$ is a given matrix. Prove that if $\mathbf{U}$ is an $m \times m$ unitary matrix and $\mathbf{V}$ is an $n \times n$ unitary matrix, then

$$\|\mathbf{A}\|_{\mathrm{F}} = \|\mathbf{U}\mathbf{A}\|_{\mathrm{F}} = \|\mathbf{A}\mathbf{V}\|_{\mathrm{F}} = \|\mathbf{U}\mathbf{A}\mathbf{V}\|_{\mathrm{F}}.$$

Problem 2: Suppose that $\mathbf{A} \in \mathbb{C}^{m\times n}$ is a matrix with a factorization

$$\begin{array}{ccccccc} \mathbf{A} & = & \mathbf{U} & \mathbf{D} & \mathbf{V}^* & = & \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*, \\ m\times n & & m\times r & r\times r & r\times n & & \end{array}$$

where r denotes the rank of the matrix, where $\mathbf{D}$ is a diagonal matrix holding the nonzero singular values of $\mathbf{A}$, and where $\mathbf{U}$ and $\mathbf{V}$ are ON matrices. In other words, the factorization is an SVD that's been stripped so that involves only the *non-zero* singular values. In particular, observe that $\sigma_j > 0$ for $j \in \{1, 2, \dots, r\}$. Define

$$\mathbf{S} = \mathbf{V}\mathbf{D}^{-1}\mathbf{U}^* = \sum_{j=1}^{r} \frac{1}{\sigma_j} \mathbf{v}_j \mathbf{u}_j^* \in \mathbb{C}^{n\times m}.$$

The matrix $\mathbf{S}$ is called the *Moore-Penrose pseudoinverse* of $\mathbf{A}$ and is sometimes written $\mathbf{S} = \mathbf{A}^\dagger$. Now fix a vector $\mathbf{b} \in \mathbb{C}^m$, and consider the linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{1}$$

(a) Suppose that $m = n = r$ so that the matrix is square and non-singular. Prove that $\mathbf{S} = \mathbf{A}^{-1}$ so that the solution to (1) is $\mathbf{x}_\star = \mathbf{S}\mathbf{b}$.

(b) Now consider a general case (any positive m and n, and any $r \in \{1, 2, \dots, \min(m, n)\}$). Set $\mathbf{x}_\star = \mathbf{S}\mathbf{b}$. Prove that

$$\|\mathbf{A}\mathbf{x}_\star - \mathbf{b}\| = \inf_{\mathbf{y}\in\mathbb{C}^n} \|\mathbf{A}\mathbf{y} - \mathbf{b}\|.$$

Hint: It may help to use that $\inf_{\mathbf{y}\in\mathbb{C}^n} \|\mathbf{A}\mathbf{y} - \mathbf{b}\| = \|\mathbf{P}\mathbf{b} - \mathbf{b}\|$ *where* $\mathbf{P}$ *is the orthogonal projection onto the column space of* $\mathbf{A}$*.*

(c) *[Optional!]* Consider again a general case (any positive m and n, and any $r \in \{1, 2, \dots, \min(m, n)\}$). Set $\mathbf{x}_\star = \mathbf{S}\mathbf{b}$. Prove that if $\mathbf{y}$ is a vector such that

$$\|\mathbf{A}\mathbf{x}_\star - \mathbf{b}\| = \|\mathbf{A}\mathbf{y} - \mathbf{b}\|,$$

then

$$\|\mathbf{x}_\star\| \leq \|\mathbf{y}\|$$

with equality holding if and only if $\mathbf{x}_\star = \mathbf{y}$.

Comment: In problem (b), you prove that in the case where equation (1) has a solution (i.e. when $\mathbf{b} \in \mathrm{col}(\mathbf{A})$) the vector $\mathbf{x}_\star = \mathbf{S}\mathbf{b}$ is a solution. When the system is not consistent, the vector $\mathbf{x}_\star$ is as good of an approximate solution as is possible. In problem (c), you prove that in cases where the equation (1) has many solutions, the particular solution picked by the pseudoinverse is the unique one that has the shortest norm. In other words, $\mathbf{x}_\star$ is the point in the hyperplane of solutions that is closest to the origin.