STAT 309: MATHEMATICAL COMPUTATIONS I FALL 2022 PROBLEM SET 4

1. Let $A \in \mathbb{R}^{m \times n}$ where $m \ge n$ and $\operatorname{rank}(A) = n$. Suppose GECP is performed on A to get

$$\Pi_1 A \Pi_2 = L U$$

where $L \in \mathbb{R}^{m \times n}$ is unit lower triangular, $U \in \mathbb{R}^{n \times n}$ is upper triangular, and $\Pi_1 \in \mathbb{R}^{m \times m}$, $\Pi_2 \in \mathbb{R}^{n \times n}$ are permutation matrices.

(a) Show that U is nonsingular and that L is of the form

$$L = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}$$

where $L_1 \in \mathbb{R}^{n \times n}$ is nonsingular.

(b) We will see how the LU factorization may be used to solve the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x}-\mathbf{b}\|_2$$

(i) Show that the problem may be solved via

$$U\widetilde{\mathbf{x}} = \mathbf{y}, \quad L^{\mathsf{T}}L\mathbf{y} = L^{\mathsf{T}}\mathbf{b},$$

where $\widetilde{\mathbf{b}} = \Pi_1 \mathbf{b}$ and $\widetilde{\mathbf{x}} = \Pi_2^\mathsf{T} \mathbf{x}$.

(ii) Describe how you would compute the solution \mathbf{y} in

$$L^{\mathsf{T}}L\mathbf{y} = L^{\mathsf{T}}\mathbf{b}$$

2. Let $\varepsilon > 0$. Consider the matrix

$$A = \begin{bmatrix} 1 & 1\\ 1 & 1+\varepsilon\\ 1 & 1-\varepsilon \end{bmatrix}$$

(a) Why is it a bad idea to solve the normal equation associated with A, i.e.

$$A^{\mathsf{T}}A\mathbf{x} = A^{\mathsf{T}}\mathbf{b}$$

when ε is small?

(b) Show that the condensed LU factorization of A is

$$A = LU = \begin{bmatrix} 1 & 0\\ 1 & 1\\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1\\ 0 & \varepsilon \end{bmatrix}$$

(c) Why is it a much better idea to solve the normal equation associated with L, i.e.

$$L^{\mathsf{T}}L\mathbf{y} = L^{\mathsf{T}}\widetilde{\mathbf{b}}?$$

This shows that the method in Problem 1 is a more stable method than using the normal equation in (a) directly.

(d) Show that the Moore–Penrose pseudoinverse of A is

$$A^{\dagger} = \frac{1}{6} \begin{bmatrix} 2 & 2 - 3\varepsilon^{-1} & 2 + 3\varepsilon^{-1} \\ 0 & 3\varepsilon^{-1} & -3\varepsilon^{-1} \end{bmatrix}.$$

Date: November 12, 2022 (Version 1.0); due: November 22, 2022.

- (e) Describe a method to compute A^{\dagger} given L and U. Verify that your method is correct by checking it against the expression in (d).
- 3. We will now discuss an alternative method to solve the least squares problem in Problem 1 that is more efficient when m n < n.
 - (a) Show that the least squares problem in Problem 1 is equivalent to

$$\min_{\mathbf{z}\in\mathbb{R}^n} \left\| \begin{bmatrix} I_n\\S \end{bmatrix} \mathbf{z} - \widetilde{\mathbf{b}} \right\|_2$$

where $S = L_2 L_1^{-1}$ and $L_1 \mathbf{y} = \mathbf{z}$. Here and below, I_n denotes the $n \times n$ identity matrix. (b) Write

$$\widetilde{\mathbf{b}} = \begin{bmatrix} \widetilde{\mathbf{b}}_1 \\ \widetilde{\mathbf{b}}_2 \end{bmatrix}$$

where $\widetilde{\mathbf{b}}_1 \in \mathbb{R}^n$ and $\widetilde{\mathbf{b}}_2 \in \mathbb{R}^{m-n}$. Show that the solution \mathbf{z} is given by

$$\mathbf{z} = \widetilde{\mathbf{b}}_1 + S^{\mathsf{T}} (I_{m-n} + SS^{\mathsf{T}})^{-1} (\widetilde{\mathbf{b}}_2 - S\widetilde{\mathbf{b}}_1).$$

- (c) Explain why when m n < n, the method in (a) is much more efficient than the method in Problem 1. For example, what happens when m = n + 1?
- 4. Let $\mathbf{c} \in \mathbb{R}^n$ and consider the linearly constrained least squares problem/minimum norm linear system

minimize
$$\|\mathbf{w}\|_2$$

subject to $A^{\mathsf{T}}\mathbf{w} = \mathbf{c}$.

(a) If we write $\widetilde{\mathbf{c}} = \Pi_2^\mathsf{T} \mathbf{c}$ and $\widetilde{\mathbf{w}} = \Pi_1 \mathbf{w}$, show that

$$\widetilde{\mathbf{w}} = L(L^{\mathsf{T}}L)^{-1}U^{-\mathsf{T}}\widetilde{\mathbf{c}}$$

where $U^{-\intercal} = (U^{-1})^{\intercal} = (U^{\intercal})^{-1}$, a standard notation that we will also use below. (*Hint*: You'd need to use something that you've already determined in an earlier part).

(b) Write

$$\widetilde{\mathbf{w}} = \begin{bmatrix} \widetilde{\mathbf{w}}_1 \\ \widetilde{\mathbf{w}}_2 \end{bmatrix}$$

where $\widetilde{\mathbf{w}}_1 \in \mathbb{R}^n$ and $\widetilde{\mathbf{w}}_2 \in \mathbb{R}^{m-n}$. Show that

$$\widetilde{\mathbf{w}}_1 = L_1^{-\mathsf{T}} U^{-\mathsf{T}} \widetilde{\mathbf{c}} - S^{\mathsf{T}} \widetilde{\mathbf{w}}_2.$$

(c) Write $\mathbf{d} = L_1^{-\mathsf{T}} U^{-\mathsf{T}} \widetilde{\mathbf{c}}$. Deduce that $\widetilde{\mathbf{w}}_2$ may be obtained either as a solution to

$$\min_{\widetilde{\mathbf{w}}_{2} \in \mathbb{R}^{m-n}} \left\| \begin{bmatrix} S^{\mathsf{T}} \\ I_{m-n} \end{bmatrix} \widetilde{\mathbf{w}}_{2} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_{2}$$

or as

$$\widetilde{\mathbf{w}}_2 = (I_{m-n} + SS^{\mathsf{T}})^{-1}S\mathbf{d}.$$

Note that when m - n < n, this method is advantageous for the same reason in Problem 3.

5. So far we have assumed that A has full column rank. Suppose now that $\operatorname{rank}(A) = r < \min\{m, n\}$.

(a) Show that the LU factorization obtained using GECP is of the form

$$\Pi_1 A \Pi_2 = L U = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}$$

where $L_1, U_1 \in \mathbb{R}^{r \times r}$ are triangular and nonsingular.

(b) Show that the above equation may be rewritten in the form

$$\Pi_1 A \Pi_2 = \begin{bmatrix} I_r \\ S_1 \end{bmatrix} L_1 U_1 \begin{bmatrix} I_r & S_2^{\mathsf{T}} \end{bmatrix}$$

for some matrices S_1 and S_2 .

(c) Hence show that the Moore–Penrose inverse of A is given by

$$A^{\dagger} = \Pi_2 \begin{bmatrix} I_r & S_2^{\mathsf{T}} \end{bmatrix}^{\dagger} U_1^{-1} L_1^{-1} \begin{bmatrix} I_r \\ S_1 \end{bmatrix}^{\dagger} \Pi_1.$$

(d) Using the general formula (derived in the lectures) for the Moore–Penrose inverse of a rank-retaining factorization, what do you get for A^{\dagger} ?