STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2021
PROBLEM SET 5

1. Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b}_1, \dots, \mathbf{b}_p \in \mathbb{R}^m$. Write $B \coloneqq [\mathbf{b}_1, \dots, \mathbf{b}_p] \in \mathbb{R}^{m\times p}$.

(a) Show that $\mathbf{x}_1, \dots, \mathbf{x}_p \in \mathbb{R}^n$ are respectively solutions to

$$\min \|A\mathbf{x} - \mathbf{b}_1\|_2, \dots\dots, \min \|A\mathbf{x} - \mathbf{b}_p\|_2 \tag{1.1}$$

if and only if $X \coloneqq [\mathbf{x}_1, \dots, \mathbf{x}_p] \in \mathbb{R}^{n\times p}$ is a solution to

$$\min \|AX - B\|_{\mathsf{F}}.$$

Hence write down a normal equation and a minimum norm solution (i.e., $\|X\|_{\mathsf{F}}$ is minimum) for the ordinary least squares problem with multiple right-hand sides (1.1) in terms of A and B. Prove your results. You should get Homework **2**, Problem **2**(c) as a special case.

(b) Generalize our proof of total least squares solution in the lectures to the case with multiple right-hand sides

$$\min\bigl\{\|[E, R]\|_{\mathsf{F}} : (A + E)X = B + R\bigr\}$$

and show that if

$$[A, B] = [U_1, U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^{\mathsf{T}} \\ V_2^{\mathsf{T}} \end{bmatrix}$$

is a singular value decomposition, then the solution is given by

$$X = -V_{12}V_{22}^{-1}, \qquad V_2 = \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix},$$

where we have assumed that $V_{22} \in \mathbb{R}^{p\times p}$ is nonsingular.

2. Let $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = n \le m$ and $\mathbf{b} \in \mathbb{R}^m$. Suppose we solved the least squares problem $\min \|A\mathbf{x} - \mathbf{b}\|_2$ using normal equation (assuming we are in one of these exceptional regimes where it is acceptable to use normal equation) and that we saved the Cholesky factor of $A^{\mathsf{T}}A$.

(a) Given a new row vector $\mathbf{c} \in \mathbb{R}^n$ and an additional value $d \in \mathbb{R}$ and that we want $|\mathbf{c}^{\mathsf{T}}\mathbf{x} - d|$ to be simultaneously minimized, i.e.,

$$\min \left\| \begin{bmatrix} A \\ \mathbf{c}^{\mathsf{T}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b} \\ d \end{bmatrix} \right\|_2. \tag{2.2}$$

Show how we may take advantage of our solution of the earlier least squares problem to solve this new one. [*Hint*: Sherman–Morrison formula]

(b) Show how this process can be reversed. Assuming that that we have already obtained the least-squares solution to (2.2) using some *unspecified* method, i.e., you are not supposed to assume what method this is, only that you have a least squares solution of (2.2). Show how we may take advantage of this solution to obtain a solution of the least squares problem where the last row is deleted, i.e., $\min \|A\mathbf{x} - \mathbf{b}\|_2$.

Date: November 26, 2021 (Version 1.3); due: December 5, 2021.

3. Let $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = n \le m$, $\mathbf{b} \in \mathbb{R}^m$, $C \in \mathbb{R}^{p\times n}$ with $\operatorname{rank}(C) = p \le n$, and $\mathbf{d} \in \mathbb{R}^p$. Consider the linearly constrained least squares problem

$$\begin{aligned} \text{minimize} \quad & \|A\mathbf{x} - \mathbf{b}\|_2 \\ \text{subject to} \quad & C\mathbf{x} = \mathbf{d}. \end{aligned} \tag{3.3}$$

An alternative way to solve this is to consider the *penalty function*

$$f_r(\mathbf{x}) := \|A\mathbf{x} - \mathbf{b}\|_2^2 + r\|C\mathbf{x} - \mathbf{d}\|_2^2 \tag{3.4}$$

for a sequence of values $r \to \infty$.

(a) For any fixed value $r > 0$, show that a minimizer $\mathbf{x}_r$ of (3.4) is a solution to an ordinary least squares problem with normal equation

$$(A^\mathsf{T}A + rC^\mathsf{T}C)\mathbf{x} = A^\mathsf{T}\mathbf{b} + rC^\mathsf{T}\mathbf{d}.$$

(b) Show that if $X \in \mathrm{GL}(n)$ and $\|X^{-1}\|_2 < 1$, then

$$(I + X)^{-1} = X^{-1} - X^{-2} + X^{-3} - X^{-4} + \cdots \tag{3.5}$$

(c) Applying the Sherman–Morrison–Woodbury formula and (3.5), show that

$$\lim_{r\to\infty} \mathbf{x}_r = \mathbf{x}_*$$

where $\mathbf{x}_*$ is the solution to (3.3) and also deduce that

$$\mathbf{x}_* = (A^\mathsf{T}A)^{-1}C^\mathsf{T}[C(A^\mathsf{T}A)^{-1}C^\mathsf{T}]^{-1}[\mathbf{d} - C(A^\mathsf{T}A)^{-1}A^\mathsf{T}\mathbf{b}] + (A^\mathsf{T}A)^{-1}A^\mathsf{T}\mathbf{b}.$$

4. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix partitioned into

$$A = \begin{bmatrix} A_{11} & A_{21}^\mathsf{T} \\ A_{21} & A_{22} \end{bmatrix}$$

with $A_{11} \in \mathbb{R}^{k\times k}$ invertible and $S = A_{22} - A_{21}A_{11}^{-1}A_{21}^\mathsf{T}$ the Schur complement.

(a) Verify that

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{21}^\mathsf{T} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{21}^\mathsf{T} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}.$$

Deduce that A is positive definite if and only if A_{11} and S are both positive definite.

(b) Show that

$$\|A_{ij}\|_2 \le \|A\|_2, \quad i, j \in \{1, 2\}$$

and deduce that if A is positive definite, then

$$\kappa_2(S) \le \kappa_2(A).$$

[*Hint:* Use Homework **4**, Problem **6**]

(c) Suppose A is positive definite, by considering the block Cholesky factorization

$$A = \begin{bmatrix} R_{11}^\mathsf{T} & 0 \\ R_{12}^\mathsf{T} & R_{22}^\mathsf{T} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

or otherwise, show that

$$\|A_{21}A_{11}^{-1}\|_2 \le \kappa_2(A)^{1/2}.$$

5. Let $p \in [1, \infty]$ and $\|\cdot\|_p$ denote the matrix p-norm. Let $M \in \mathbb{C}^{n\times n}$ and $\mathbf{b} \in \mathbb{C}^n$.

(a) Suppose $\|M\|_p < 1$. Show that $(I - M)\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x} \in \mathbb{C}^n$. Show that for any $\mathbf{x}^{(0)} \in \mathbb{C}^n$, the iterates

$$\mathbf{x}^{(k+1)} = M\mathbf{x}^{(k)} + \mathbf{b}, \quad k = 0, 1, 2, \ldots$$

converge to $\mathbf{x}$. [*Hint:* Use Homework **1**, Problem **2**]

(b) Show that

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_p \leq \frac{\|M\|_p^k}{1 - \|M\|_p} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_p$$

where $\mathbf{x}$ is the solution of $(I - M)\mathbf{x} = \mathbf{b}$ for all $k \in \mathbb{N}$.

(c) Let $A \in \mathbb{C}^{n \times n}$ and define

$$\gamma_1 := \max_{j=1,\dots,n} \sum_{\substack{i=1 \\ i \neq j}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad \gamma_\infty := \max_{i=1,\dots,n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad \gamma_2 := \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right|^2 \right)^{1/2} < 1.$$

Show that if $\gamma_p < 1$ for $p = 1$, 2, or ∞, then Jacobi method

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left[b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_i^{(k)} \right], \quad i = 1, \dots, n, \quad k = 0, 1, 2, \dots$$

converges for any $\mathbf{x}^{(0)} \in \mathbb{C}^n$ to the unique solution of $A\mathbf{x} = \mathbf{b}$ (note that the $p = \infty$ case we have covered in lectures). Furthermore,

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_p \leq \frac{\gamma_p^k}{1 - \gamma_p} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_p$$

for $p = 1, 2, \infty$.