STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2022
PROBLEM SET 5

1. Consider the block matrix
$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{n\times n}$$
where $A \in \mathbb{R}^{p\times p}$, $B \in \mathbb{R}^{p\times q}$, $C \in \mathbb{R}^{q\times p}$, $D \in \mathbb{R}^{q\times q}$ and $n = p+q$. The Schur complements of A and D are
$$S = D - CA^{\dagger}B \quad \text{and} \quad T = A - BD^{\dagger}C$$
respectively.

(a) Verify that if A and S are nonsingular, then
$$X^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{bmatrix}$$
and if D and T are nonsingular, then
$$X^{-1} = \begin{bmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}CT^{-1} & D^{-1} + D^{-1}CT^{-1}BD^{-1} \end{bmatrix}.$$

(b) Show that
$$\det X = \begin{cases} \det(A)\det(D - CA^{-1}B) & \text{if } A \text{ nonsingular,} \\ \det(D)\det(A - BD^{-1}C) & \text{if } D \text{ nonsingular.} \end{cases}$$
Deduce that
$$\det(A + BC) = \det(A)\det(I + CA^{-1}B)$$
and use it to find the determinants of the following matrices
$$\begin{bmatrix} \frac{1+\lambda_1}{\lambda_1} & 1 & \cdots & 1 \\ 1 & \frac{1+\lambda_2}{\lambda_2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & \frac{1+\lambda_n}{\lambda_n} \end{bmatrix}, \quad \begin{bmatrix} 1+\lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1 & 1+\lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 & \lambda_2 & \cdots & 1+\lambda_n \end{bmatrix}, \quad \begin{bmatrix} \lambda & \mu & \mu & \cdots & \mu \\ \mu & \lambda & \mu & \cdots & \mu \\ \mu & \mu & \lambda & \cdots & \mu \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu & \mu & \mu & \cdots & \lambda \end{bmatrix}.$$

(c) Show that if A has all principal matrices nonsingular so that we may perform Gaussian elimination without pivoting to A, then applying the first p steps of that to X yields
$$X = \begin{bmatrix} L_{11} & 0 \\ L_{21} & I_q \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & I_q \end{bmatrix}$$
where $A = L_{11}U_{11}$ is the LU factorization of A. What are L_{21} and U_{12} in terms of L_{11}, U_{11} and the blocks of X?

(d) Suppose X is symmetric (so $C = B^{\mathsf{T}}$) and A is positive definite. Show that applying the first p steps of Cholesky factorization to X yields
$$X = \begin{bmatrix} R_{11}^{\mathsf{T}} \\ R_{12}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & S \end{bmatrix}$$
where $A = R_{11}^{\mathsf{T}}R_{11}$ is the Cholesky factorization. What is R_{12} in terms of R_{11} and the blocks of X?

Date: November 23, 2022 (Version 1.0); due: December 2, 2022.

2. 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

$$\min \|A\mathbf{x} - \mathbf{b}_1\|_2, \dots\dots, \min \|A\mathbf{x} - \mathbf{b}_p\|_2 \tag{2.1}$$

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 (2.1) 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

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

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.

3. 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.,

$$\min \left\| \begin{bmatrix} A \\ \mathbf{c}^{\mathsf{T}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b} \\ d \end{bmatrix} \right\|_2. \tag{3.2}$$

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 (3.2) 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 (3.2). 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., $\min \|A\mathbf{x} - \mathbf{b}\|_2$.

4. 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

$$\begin{aligned} \text{minimize} \quad & \|A\mathbf{x} - \mathbf{b}\|_2 \\ \text{subject to} \quad & C\mathbf{x} = \mathbf{d}. \end{aligned} \tag{4.3}$$

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

$$f_r(\mathbf{x}) \coloneqq \|A\mathbf{x} - \mathbf{b}\|_2^2 + r\|C\mathbf{x} - \mathbf{d}\|_2^2 \tag{4.4}$$

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

(a) For any fixed value $r > 0$, show that a minimizer $\mathbf{x}_r$ of (4.4) 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

$$(I + X)^{-1} = X^{-1} - X^{-2} + X^{-3} - X^{-4} + \cdots \tag{4.5}$$

(c) Applying the Sherman–Morrison–Woodbury formula and (4.5), show that

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

where $\mathbf{x}_*$ is the solution to (4.3) and also deduce that

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

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

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

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

(a) Verify that

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

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 Problem **1**]

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

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

or otherwise, show that

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