CSE 6643 Homework 4

Karl Hiner

1 Schur Complements I [25 pts]

Consider the matrix

$$\boldsymbol{M} \coloneqq \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}, \tag{1}$$

where $\boldsymbol{A} \in \mathbb{K}^{m \times m}$ and $\boldsymbol{D} \in \mathbb{K}^{n \times n}$ are square matrices. The matrix $\boldsymbol{M}/\boldsymbol{A} \coloneqq \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}$ is called the Schur complement of $\boldsymbol{A}$ in $\boldsymbol{M}$.

(a) [5 pts]

Relate the Schur complement $\boldsymbol{M}/\boldsymbol{A}$ to the block LU factorization of $\boldsymbol{M}$.

The block LU factorization of $\boldsymbol{M}$ is given by

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}. \tag{2}$$

We can verify this by multiplying the terms and verifying we end up with $\boldsymbol{M}$:

$$\begin{aligned}
&\begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{A} & \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \boldsymbol{M}
\end{aligned}$$

(b) [5 pts]

Show that the determinant of $\boldsymbol{M}$ is given by $\det(\boldsymbol{A})\det(\boldsymbol{M}/\boldsymbol{A})$.

Expressing $\boldsymbol{M}$ as its block LU factorization:

$$\begin{aligned}
\boldsymbol{M} &= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \\
\det(\boldsymbol{M}) &= \det\left(\begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \right) \\
&= \det\left(\begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix} \right) \det\left(\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \right) \det\left(\begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \right) \\
&= 1 \cdot \det\left(\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}/\boldsymbol{A} \end{pmatrix} \right) \cdot 1 \\
&= \det(\boldsymbol{A}) \det(\boldsymbol{M}/\boldsymbol{A})
\end{aligned}$$

(c) [7.5 pts]

Let $\boldsymbol{M} = \boldsymbol{L}\boldsymbol{U}$ be the LU factorization of $\boldsymbol{M}$. Split its factors into block matrices

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{L}_{1,1} & \boldsymbol{0} \\ \boldsymbol{L}_{2,1} & \boldsymbol{L}_{2,2} \end{pmatrix}, \quad \boldsymbol{U} = \begin{pmatrix} \boldsymbol{U}_{1,1} & \boldsymbol{U}_{1,2} \\ \boldsymbol{0} & \boldsymbol{U}_{2,2} \end{pmatrix} \tag{3}$$

according to the same blocking as $\boldsymbol{M}$. Show that $\boldsymbol{M}/\boldsymbol{A} = \boldsymbol{L}_{2,2}\boldsymbol{U}_{2,2}$.

$$\begin{aligned}
\boldsymbol{M} = \boldsymbol{L}\boldsymbol{U} &= \begin{pmatrix} \boldsymbol{L}_{1,1} & \boldsymbol{0} \\ \boldsymbol{L}_{2,1} & \boldsymbol{L}_{2,2} \end{pmatrix} \begin{pmatrix} \boldsymbol{U}_{1,1} & \boldsymbol{U}_{1,2} \\ \boldsymbol{0} & \boldsymbol{U}_{2,2} \end{pmatrix} \\
\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{L}_{1,1}\boldsymbol{U}_{1,1} & \boldsymbol{L}_{1,1}\boldsymbol{U}_{1,2} \\ \boldsymbol{L}_{2,1}\boldsymbol{U}_{1,1} & \boldsymbol{L}_{2,1}\boldsymbol{U}_{1,2} + \boldsymbol{L}_{2,2}\boldsymbol{U}_{2,2} \end{pmatrix}
\end{aligned}$$

Equating the lower-right subblocks:

$$\begin{aligned}
\boldsymbol{D} &= \boldsymbol{L}_{2,1}\boldsymbol{U}_{1,2} + \boldsymbol{L}_{2,2}\boldsymbol{U}_{2,2} \\
\boldsymbol{L}_{2,2}\boldsymbol{U}_{2,2} &= \boldsymbol{D} - \boldsymbol{L}_{2,1}\boldsymbol{U}_{1,2} \\
&= \boldsymbol{D} - \left(\boldsymbol{C}\boldsymbol{U}_{1,1}^{-1}\right)\left(\boldsymbol{L}_{1,1}^{-1}\boldsymbol{B}\right) && \text{(solve for } \boldsymbol{L}_{2,1} \text{ and } \boldsymbol{U}_{1,2} \text{ by equating subblocks)} \\
&= \boldsymbol{D} - \boldsymbol{C}\left(\boldsymbol{U}_{1,1}^{-1}\boldsymbol{L}_{1,1}^{-1}\right)\boldsymbol{B} && \text{(changing parenthesis)} \\
&= \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} && (\boldsymbol{A} = \boldsymbol{L}_{1,1}\boldsymbol{U}_{1,1},\ \boldsymbol{A} \text{ invertible}) \\
&= \boldsymbol{M}/\boldsymbol{A}
\end{aligned}$$

(d) [7.5 pts]

Now consider the 3×3 block matrix

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{N}_{1,1} & \boldsymbol{N}_{1,2} & \boldsymbol{N}_{1,3} \\ \boldsymbol{N}_{2,1} & \boldsymbol{N}_{2,2} & \boldsymbol{N}_{2,3} \\ \boldsymbol{N}_{3,1} & \boldsymbol{N}_{3,2} & \boldsymbol{N}_{3,3} \end{pmatrix}. \tag{4}$$

Prove the *Quotient Property* of Schur complements

$$(\boldsymbol{N}/\boldsymbol{N}_{1,1})/(\boldsymbol{N}/\boldsymbol{N}_{1,1})_{1,1} = \boldsymbol{N}\Big/\begin{pmatrix} \boldsymbol{N}_{1,1} & \boldsymbol{N}_{1,2} \\ \boldsymbol{N}_{2,1} & \boldsymbol{N}_{2,2} \end{pmatrix}. \tag{5}$$

Here, $(\boldsymbol{N}/\boldsymbol{N}_{1,1})_{1,1}$ is the top left block of $\boldsymbol{N}/\boldsymbol{N}_{1,1}$.

2 Schur Complements II [25 pts + 10 bonus]

Let $\boldsymbol{M}$ be as in the previous problem.

(a) [5 pts]

Divide the inverse of $\boldsymbol{M}$ into blocks (of the same size as those of $\boldsymbol{M}$) as

$$\boldsymbol{M}^{-1} = \begin{pmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ \boldsymbol{\gamma} & \boldsymbol{\delta} \end{pmatrix}. \tag{6}$$

Show that the Schur complement $\boldsymbol{M}/\boldsymbol{A}$ is the inverse of $\boldsymbol{\delta}$.

(b) [5 pts]

For $\boldsymbol{M}$ symmetric positive definite, show that if λ is an eigenvalue of $\boldsymbol{M}/\boldsymbol{A}$, then $\lambda_{\min}(\boldsymbol{M}) \leq \lambda \leq \lambda_{\max}(\boldsymbol{M})$.

(c) [7.5 pts]

Show that the size of the pivots occuring in the Cholesky or LU factorization of a s.p.d. matrix $\boldsymbol{M}$ is lower bounded by $\left\|\boldsymbol{M}^{-1}\right\|^{-1}$.

(d) [7.5 pts]

For $\boldsymbol{M}$ s.p.d, show that

$$\boldsymbol{y}^*(\boldsymbol{M}/\boldsymbol{A})\boldsymbol{y} = \min_{x \in \mathbb{K}^m} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^* \boldsymbol{M} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}. \tag{7}$$

(e) [10 bonus pts]

By replacing the min with a min max over suitable variables, derive a version of the results of (d) that is valid for general $\boldsymbol{M}$.

3 Cholesky and QR [25 pts]

(a) [5 pts]

Assume that the invertible matrix $\boldsymbol{A} \in \mathbb{K}^{m \times m}$ satisfies $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$. Derive a way to write $\boldsymbol{A}$ as the sum of rank-one matrices in outer product form.

(b) [5 pts]

Show that the LU factorization of an invertible matrix $\boldsymbol{A} \in \mathbb{K}^{m \times m}$ is unique.

(c) [5 pts]

Assume that we factor $\boldsymbol{A} \in \mathbb{K}^{m \times m}$ as $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^*$ for lower triangular $\boldsymbol{L}$. Up to which choices is this factorization unique? Prove your results (as always) and make sure to treat the case $\mathbb{K} = \mathbb{C}$.

(d) [5 pts]

Using these results, prove and explain the relationship between the Cholesky factor of $\boldsymbol{B}^*\boldsymbol{B}$ and the QR factorization of $\boldsymbol{B}$, for $\boldsymbol{B} \in \mathbb{K}^{m \times n}$ and $m \geq n$.

(e) [5 pts]

For $\boldsymbol{A} = \boldsymbol{B}^*\boldsymbol{B}$ and $\boldsymbol{B}$ as in (d), write

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{1,1} & \boldsymbol{A}_{1,2} \\ \boldsymbol{A}_{2,1} & \boldsymbol{A}_{2,2} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 \end{pmatrix}. \tag{8}$$

Relate the Schur complement $\boldsymbol{A}_{2,2} - \boldsymbol{A}_{2,1}(\boldsymbol{A}_{1,1})^{-1}\boldsymbol{A}_{1,2}$ to the Gram-matrix of the columns of $\boldsymbol{B}_2$, projected on the orthogonal complement of $\boldsymbol{B}_1$. Use this result to refine your comment in (d) on the relationship of Cholesky and QR factorization.

4 QRs [25 pts]

(a) [5 pts]

Go to section (a) of the file `HW4_your_code.jl` and implement a function that uses the classical Gram-Schmidt algorithm for computing a reduced QR factorization.

(b) [5 pts]

Go to section (b) of the file `HW4_your_code.jl` and implement a function that uses the modified Gram-Schmidt algorithm for computing a reduced QR factorization.

(c) [5 pts]

Go to section (c) of the file `HW4_your_code.jl` and implement a function that computes the QR factorization using Householder reflections. Your algorithm should operate in place, overwriting the input matrix and not allocating additional memory.

(d) [5 pts]

Go to section (d) of the file `HW4_your_code.jl` and implement functions that use the QR factorization matrix computed in (c) to multiply a vector with the $\boldsymbol{QR}$ (multiplication) or $\boldsymbol{RQ}^*$ (solving overdetermined least squares problem). Your functions should take a preallocated output vector as input in which to store the result. No additional allocation should be performed.

(e) [5 pts]

Using what you learned in class, design an experiment that compares the numerical stability of the different methods. For instance, compute the accuracy of the different approaches over matrices of increasing size or condition number and provide a plot of your results. Using what you learned in class to make sure that you include examples highlighting the lack of stability of classical Gram-Schmidt.