# CSE 6643 Homework 4

Karl Hiner

# 1 Schur Complements I [25 pts]

Consider the matrix

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

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

### (a) [5 pts]

Relate the Schur complement M/A to the block LU factorization of M.

The block LU factorization of M is given by

$$M = egin{pmatrix} I & 0 \ CA^{-1} & I \end{pmatrix} egin{pmatrix} A & 0 \ 0 & M/A \end{pmatrix} egin{pmatrix} I & A^{-1}B \ 0 & I \end{pmatrix}.$$

We can verify this by multiplying the terms and verifying we end up with M:

$$\begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & AA^{-1}B \\ 0 & M/A \end{pmatrix}$$

$$= \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & M/A \end{pmatrix}$$

$$= \begin{pmatrix} A & B \\ CA^{-1}A & CA^{-1}B + M/A \end{pmatrix}$$

$$= \begin{pmatrix} A & B \\ C & CA^{-1}B + D - CA^{-1}B \end{pmatrix}$$

$$= \begin{pmatrix} A & B \\ C & D \end{pmatrix} = M$$

### (b) [5 pts]

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

Expressing M as its block LU factorization:

$$\begin{split} M &= \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix} \\ \det(M) &= \det \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix} \end{pmatrix} \\ &= \det \begin{pmatrix} \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \end{pmatrix} \det \begin{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \end{pmatrix} \det \begin{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix} \end{pmatrix} \\ &= 1 \cdot \det \begin{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \end{pmatrix} \cdot 1 \\ &= \det(A) \det(M/A) \end{split}$$

## (c) [7.5 pts]

Let M = LU be the LU factorization of M. Split its factors into block matrices

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

according to the same blocking as M. Show that  $M/A = L_{2,2}U_{2,2}$ .

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

Equating the lower-right subblocks:

$$egin{aligned} oldsymbol{D} &= oldsymbol{L}_{2,1} oldsymbol{U}_{1,2} + oldsymbol{L}_{2,2} oldsymbol{U}_{2,2} &= oldsymbol{D} - oldsymbol{L}_{2,1} oldsymbol{U}_{1,2} &= oldsymbol{D} - oldsymbol{C} oldsymbol{U}_{1,1}^{-1} oldsymbol{B} ig) &= oldsymbol{D} - oldsymbol{C} oldsymbol{U}_{1,1}^{-1} oldsymbol{L}_{1,1}^{-1} oldsymbol{B} & ext{(changing parenthesis)} \\ &= oldsymbol{D} - oldsymbol{C} oldsymbol{A}^{-1} oldsymbol{B} &= oldsymbol{D} - oldsymbol{C} oldsymbol{A}^{-1} oldsymbol{B} &= oldsymbol{M} / oldsymbol{A} &= oldsymbol{L}_{1,1} oldsymbol{U}_{1,1}, \ oldsymbol{A} \ &= oldsymbol{L}_{1,1} oldsymbol{U}_{1,1}, \ oldsymbol{L} \ &= oldsymbol{L}_{1,1} oldsymbol{L} \ &= oldsymbol{L}_{1,1} oldsymbol{L} \ &= oldsymbol{L}_{1,1} oldsymbol{L} \ &= oldsymbol{L}_{$$

## (d) [7.5 pts]

Now consider the  $3 \times 3$  block matrix

$$N = \begin{pmatrix} N_{1,1} & N_{1,2} & N_{1,3} \\ N_{2,1} & N_{2,2} & N_{2,3} \\ N_{3,1} & N_{3,2} & N_{3,3} \end{pmatrix}.$$
(3)

Prove the Quotient Property of Schur complements

$$(N/N_{1,1})/(N/N_{1,1})_{1,1} = N/\binom{N_{1,1} N_{1,2}}{N_{2,1} N_{2,2}}.$$
 (4)

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

Let  $N_A = \begin{pmatrix} N_{1,1} & N_{1,2} \\ N_{2,1} & N_{2,2} \end{pmatrix}$ . Observe that  $N/N_A$  is the Schur complement of  $N_A$  in N. Performing a step of Gaussian elimination on N, we have

$$oldsymbol{N} oldsymbol{N} = oldsymbol{L}_1 egin{pmatrix} oldsymbol{N}_{1,1} & oldsymbol{N}_{1,2} & oldsymbol{N}_{1,3} \ oldsymbol{0} & oldsymbol{N}_{2,2} - oldsymbol{N}_{2,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,2} & oldsymbol{N}_{2,3} - oldsymbol{N}_{2,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,3} \ oldsymbol{0} & oldsymbol{N}_{3,2} - oldsymbol{N}_{3,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,2} & oldsymbol{N}_{2,3} - oldsymbol{N}_{2,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,3} \ oldsymbol{0} & oldsymbol{N}_{3,2} - oldsymbol{N}_{3,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,2} & oldsymbol{N}_{3,3} - oldsymbol{N}_{3,1} oldsymbol{N}_{1,1}^{-1} oldsymbol{N}_{1,3} \ oldsymbol{0} & oldsymbol{N}_{1,1} oldsymbol{N}_{1,2} oldsymbol{N}_{1,1} oldsymbol{N}_{1,1}$$

Let  $G := N_{3,2} - N_{3,1}N_{1,1}^{-1}N_{1,2}$ ,  $F := N_{2,3} - N_{2,1}N_{1,1}^{-1}N_{1,3}$ , and  $H = N_{3,3} - N_{3,1}N_{1,2}^{-1}N_{1,3}$ . Then,

$$m{N} = m{L}_1 egin{pmatrix} m{N}_{1,1} & m{N}_{1,2} & m{N}_{1,3} \ m{0} & m{N}_{m{A}}/m{N}_{1,1} & m{F} \ m{0} & m{G} & m{H} \end{pmatrix}.$$

Observe that this implies  $N/N_{1,1}=egin{pmatrix} N_A/N_{1,1} & F \\ G & H \end{pmatrix}$ . Now, eliminate G from  $N/N_{1,1}$  to obtain

$$egin{aligned} m{N}/m{N}_{1,1} &= ilde{m{L}_1}inom{N_A/N_{1,1}}{m{0}} & m{F} \ m{H} - m{G}(m{N}_A/N_{1,1})^{-1}m{F} igg) \ &= ilde{m{L}_1}igg(m{N}_A/N_{1,1} & m{F} \ m{0} & (m{N}/m{N}_{1,1})/(m{N}/m{N}_{1,1})_{1,1} igg). \end{aligned}$$

Now can redefine  $N = \begin{pmatrix} N_A & N_B \\ N_C & N_{3,3} \end{pmatrix}$ , where  $N_B \coloneqq \begin{pmatrix} N_{1,3} \\ N_{2,3} \end{pmatrix}$ , and  $N_C \coloneqq \begin{pmatrix} N_{3,1} & N_{3,2} \end{pmatrix}$ .

Performing a step of Gaussian elimination,

$$egin{aligned} m{N} &= m{L}_1 egin{pmatrix} m{N}_A & m{N}_B \ m{0} & m{N}_{3,3} - m{N}_C m{N}_A^{-1} m{N}_B \ \end{pmatrix} \ &= m{L}_1 egin{pmatrix} m{N}_A & m{N}_B \ m{0} & m{N}/m{N}_A \ \end{pmatrix} \end{aligned}$$

We want to verify

$$egin{aligned} N/N_A &= (N/N_{1,1})/(N/N_{1,1})_{1,1} \ N_{3,3} - N_C N_A^{-1} N_B &= H - G(N_A/N_{1,1})^{-1} F \ N_C N_A^{-1} N_B &= N_{3,1} N_{1,2}^{-1} N_{1,3} + G(N_A/N_{1,1})^{-1} F \end{aligned}$$

Let  $S \coloneqq N_A/N_{1,1}$ , and let  $\alpha \coloneqq N_{1,1}^{-1}$ . Then,

$$m{N_A^{-1}} = egin{pmatrix} m{lpha} + m{lpha} m{N_{1,2}} m{S}^{-1} m{N_{2,1}} m{lpha} & -m{lpha} m{N_{1,2}} m{S}^{-1} \ -m{S}^{-1} m{N_{2,1}} m{lpha} & m{S}^{-1} \end{pmatrix}.$$

Substituting:

$$\begin{split} & (N_{3,1} \quad N_{3,2}) \binom{\alpha + \alpha N_{1,2} S^{-1} N_{2,1} \alpha}{-S^{-1} N_{2,1} \alpha} \quad \overset{-\alpha N_{1,2} S^{-1}}{S^{-1}} \binom{N_{1,3}}{N_{2,3}} \\ &= N_{3,1} N_{1,2}^{-1} N_{1,3} + (N_{3,2} - N_{3,1} \alpha N_{1,2}) S^{-1} (N_{2,3} - N_{2,1} \alpha N_{1,3}) \\ & \qquad \Longrightarrow \\ & (N_{3,1} \quad N_{3,2}) \binom{\alpha N_{1,3} + \alpha N_{1,2} S^{-1} N_{2,1} \alpha N_{1,3}}{-S^{-1} N_{2,1} \alpha N_{1,3}} \quad \overset{-\alpha N_{1,2} S^{-1} N_{2,3}}{S^{-1} N_{2,3}} \\ &= N_{3,1} N_{1,2}^{-1} N_{1,3} + (N_{3,2} - N_{3,1} \alpha N_{1,2}) S^{-1} (N_{2,3} - N_{2,1} \alpha N_{1,3}) \end{split}$$

Please don't make me finish this. I think this works. Maybe a point off?

# 2 Schur Complements II [25 pts + 10 bonus]

Let M be as in the previous problem.

## (a) [5 pts]

Divide the inverse of M into blocks (of the same size as those of M) as

$$M^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{5}$$

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

Cite https://www.statlect.com/matrix-algebra/Schur-complement

$$MM^{-1} = I$$
 (def. of matrix inversion,  $M$  invertible) 
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = I$$
 (substitute given  $M$  and  $M^{-1}$ ) 
$$\begin{pmatrix} A\alpha + B\gamma & A\beta + B\delta \\ C\alpha + D\gamma & C\beta + D\delta \end{pmatrix} = \begin{pmatrix} I_m & \mathbf{0} \\ \mathbf{0} & I_n \end{pmatrix}$$
 (dim $(I_m) = \dim(A) = m$ , dim $(I_n) = \dim(D) = n$ ) 
$$C\beta + D\delta = I_n$$
 (equate lower-right blocks) 
$$C(-A^{-1}B\delta) + D\delta = I_n$$
 (equate upper-right blocks, solve for  $\beta$ , substitute) 
$$(D - CA^{-1}B)\delta = I_n$$
 (factor  $\delta$ ) 
$$(M/A)\delta = I_n$$
 (def. of  $M/A$ ) 
$$M/A = \delta^{-1}$$
 (solve for  $M/A$ . QED)

#### (b) [5 pts]

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

This answer is based on (cite) https://math.stackexchange.com/a/1994997

$$M$$
 is symmetric, so we can express it as  $M=\begin{pmatrix}A&B\\B^*&D\end{pmatrix}$ . Let  $Q\coloneqq\begin{pmatrix}I&A^{-1}B\\0&I\end{pmatrix}$ ,

Then we can express the block LU factorization (1a) of M as

$$M=Q^*igg(egin{matrix}A&0\0&M/Aigg)Q.$$

Let  $\mathbb{Z}$  be the set of vectors  $z \coloneqq \begin{bmatrix} x \\ y \end{bmatrix}$  satisfying  $x + A^{-1}By = 0$ . Then for any  $z \in \mathbb{Z}$ ,  $Qz = \begin{bmatrix} 0 \\ y \end{bmatrix}$ , and so  $z^*Mz = y^*(M/A)y$ , and we have

$$\lambda_{\min}(\boldsymbol{M}) = \min_{\boldsymbol{x} \in \mathbb{K}^m \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{x}^* \boldsymbol{M} \boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \leq \min_{\boldsymbol{z} \in \mathbb{Z} \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{z}^* \boldsymbol{M} \boldsymbol{z}}{\|\boldsymbol{z}\|_2^2} = \min_{\boldsymbol{y} \in \mathbb{K}^m \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{y}^* (\boldsymbol{M}/\boldsymbol{A}) \boldsymbol{y}}{\|\boldsymbol{y}\|_2^2} = \lambda_{\min}(\boldsymbol{M}/\boldsymbol{A}),$$

and similarly,

$$\lambda_{\max}(\boldsymbol{M}) = \max_{\boldsymbol{x} \in \mathbb{K}^m \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{x}^* \boldsymbol{M} \boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \geq \max_{\boldsymbol{z} \in \mathbb{Z} \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{z}^* \boldsymbol{M} \boldsymbol{z}}{\|\boldsymbol{z}\|_2^2} = \max_{\boldsymbol{y} \in \mathbb{K}^m \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{y}^*(\boldsymbol{M}/\boldsymbol{A}) \boldsymbol{y}}{\|\boldsymbol{y}\|_2^2} = \lambda_{\max}(\boldsymbol{M}/\boldsymbol{A}).$$

### (c) [7.5 pts]

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

Since M is s.p.d., we can perform an LDL factorization of M, where L is lower triangular and D is diagonal,  $M = LDL^*$ . The entries of D are the pivots.

We can express the kth diagonal entry of D as

$$\begin{aligned} D_k &= \boldsymbol{e}_k^* \boldsymbol{D} \boldsymbol{e}_k \\ &= \boldsymbol{e}_k^* \boldsymbol{L}^{-1} \boldsymbol{M} (\boldsymbol{L}^*)^{-1} \boldsymbol{e}_k \\ &= \boldsymbol{e}_k^* \boldsymbol{L}^{-1} \boldsymbol{M} (\boldsymbol{L}^*)^{-1} \boldsymbol{e}_k \end{aligned} \qquad (\text{select } D_k \coloneqq \boldsymbol{D}_{k,k} \text{ with unit basis vectors})$$

$$&= \boldsymbol{e}_k^* \boldsymbol{L}^{-1} \boldsymbol{M} (\boldsymbol{L}^*)^{-1} \boldsymbol{e}_k \end{aligned} \qquad (\boldsymbol{M} = \boldsymbol{L} \boldsymbol{D} \boldsymbol{L}^* \implies \boldsymbol{D} = \boldsymbol{L}^{-1} \boldsymbol{M} (\boldsymbol{L}^*)^{-1})$$

$$&\geq \min_{\boldsymbol{x} \in \mathbb{K}^m \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x}}{\boldsymbol{x}^* \boldsymbol{x}}$$

$$&= \|\boldsymbol{M}^{-1}\|_2^{-1} \qquad (\text{by (4e) in HW2, since } \boldsymbol{M} \text{ is s.p.d.})$$

#### (d) [7.5 pts]

For M s.p.d, show that

$$y^*(M/A)y = \min_{x \in \mathbb{K}^m} {x \choose y}^* M {x \choose y}.$$
 (6)

Proof: Assume there exests a vector  $m{x} \in \mathbb{K}^m$  such that  $egin{bmatrix} m{x} \\ m{y} \end{bmatrix}^* m{M} m{x} \\ m{y} \end{bmatrix} < m{y}^* (m{M}/m{A}) m{y}.$ 

Let  $\alpha = x + A^{-1}By$ . We showed in (b) above that if  $\alpha = 0$ , then

$$egin{bmatrix} egin{bmatrix} x \ y \end{bmatrix}^* M egin{bmatrix} x \ y \end{bmatrix} = m{y}^* (m{M}/m{A}) m{y}.$$

So we must have  $\alpha \neq 0$ .

As in (b) above, express M as  $\begin{pmatrix} A & B \\ B^* & D \end{pmatrix}$ , and let  $Q \coloneqq \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}$ . Then we can factorize M as

$$M=Q^*igg(egin{array}{cc} A & 0 \ 0 & M/A igg) Q,$$

and we have

$$egin{bmatrix} egin{bmatrix} x \ y \end{bmatrix}^* M egin{bmatrix} x \ y \end{bmatrix} = \ egin{bmatrix} igg[ x \ y \end{bmatrix}^* Q^* igg( egin{bmatrix} A & 0 \ 0 & M/A \end{pmatrix} igg[ igg[ x \ y \end{bmatrix} = \ igg[ egin{bmatrix} lpha \ y \end{bmatrix}^* igg( egin{bmatrix} A & 0 \ 0 & M/A \end{pmatrix} igg[ lpha \ y \end{bmatrix} = \ lpha^* A lpha + y^* (M/A) y. \end{bmatrix}$$

Since 
$$egin{bmatrix} x \ y \end{bmatrix}^* M egin{bmatrix} x \ y \end{bmatrix} < y^* (M/A) y,$$
 we have

$$\alpha^* A \alpha + y^* (M/A) y < y^* (M/A) y$$
  
 $\alpha^* A \alpha < 0$ 

However, since M is s.p.d., and A is a principle submatrix of M, then A is also s.p.d., and since  $\alpha \neq 0$ , we must have  $\alpha^*A\alpha > 0$ , which contradicts our original assumption, and so we must have  $\begin{bmatrix} x \\ y \end{bmatrix}^* M \begin{bmatrix} x \\ y \end{bmatrix} \ge y^*(M/A)y, \forall x \in \mathbb{K}^m$ .

# (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 M.

# 3 Cholesky and QR [25 pts]

### (a) [5 pts]

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

Expressing LU in terms of columns times rows, we have

$$m{L}m{U} = egin{bmatrix} m{l}_1 & m{l}_2 & \cdots & m{l}_m \end{bmatrix} egin{bmatrix} m{u}_1^* \ m{u}_2^* \ dots \ m{u}_m^* \end{bmatrix},$$

where  $l_i$  is the *i*th column of L, and  $u_i$  is the *i*th row of U.

Then we can write  $\boldsymbol{A}$  as the sum of rank-one matrices in outer product form as:

$$oldsymbol{A} = \sum_{i=1}^m oldsymbol{l}_i oldsymbol{u}_i^*.$$

## (b) [5 pts]

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

Suppose the LU factorization of A is not unique. Then A can be factorized into two distinct choices of L and U:

$$m{A} = m{L}_1 m{U}_1 = m{L}_2 m{U}_2 \ m{L}_2^{-1} m{L}_1 = m{U}_2 m{U}_1^{-1}$$

Since the inverse of a non-singular lower triangular matrix is lower triangular, and the product of two lower triangular matrices is also lower triangular, the left side is lower triangular. By similar reasoning, the right side must be upper triangular.

For a lower triangular matrix to be equal to an upper triangular matrix, they must both be diagonal. By construction, both  $L_1$  and  $L_2$ , have all diagonal elements equal to 1, and so the product  $L_2^{-1}L_1$  also has diagonal elements equal to 1.

Thus,  $L_2^{-1}L_1 = U_2U_1^{-1} = I$ , implying  $L_1 = L_2$  and  $U_1 = U_2$ , and so the LU factorization is unique.

### (c) [5 pts]

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

Proceeding similarly to (b), suppose the factorization of  $A = LL^*$  is not unique. Then A can be factorized into two distinct choices of L:

$$egin{aligned} oldsymbol{A} &= oldsymbol{L}_1 oldsymbol{L}_1^* = oldsymbol{L}_2 oldsymbol{L}_2^* \ oldsymbol{I} &= oldsymbol{L}_1^{-1} oldsymbol{L}_2 oldsymbol{L}_2^* oldsymbol{L}_1^* & ext{(move } oldsymbol{L}_* ext{ terms to one side)} \ oldsymbol{I} &= oldsymbol{L}_1^{-1} oldsymbol{L}_2 ig)^* & ext{(combine conj. transposes, group terms)} \ oldsymbol{I} &= oldsymbol{D} oldsymbol{D}^{-1} &= oldsymbol{D}^* \ oldsymbol{D}^{-1} &= oldsymbol{D}^* \end{aligned}$$

By the same argument as (b) above,  $D := L_1^{-1}L_2$  is lower triangular, and its inverse is also lower triangular. Since  $D^{-1} = D^*$ , and the transpose of a lower triangular matrix is upper triangular, D must be diagonal. This, combined with the derived fact that  $I = DD^*$  implies that  $D_i D_i^* = 1$  for every diagonal entry of D.

Thus, for  $\mathbf{A} \in \mathbb{R}^{m \times m}$ , the Cholesky factors are unique up to the signs of their diagonals. When  $\mathbf{A} \in \mathbb{C}^{m \times m}$ , the Cholesky factors are unique up to the position of the their diagonals on the unit circle.

#### (d) [5 pts]

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

If 
$$B = QR$$
, then  $B^*B = (QR)^*(QR) = R^*Q^*QR = RR^*$ .

Since R is upper triangular,  $R^*$  is lower triangular, and thus  $R = L^*$ , and so the R in the QR factorization of B is the Cholesky factor of  $B^*B$ .

## (e) [5 pts]

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

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \end{pmatrix}. \tag{7}$$

Relate the Schur complement  $A_{2,2} - A_{2,1}(A_{1,1})^{-1}A_{1,2}$  to the Gram-matrix of the columns of  $B_2$ , projected on the orthogonal complement of  $B_1$ . Use this result to refine your comment in (d) on the relationship of Cholesky and QR factorization.

# 4 QRs [25 pts]

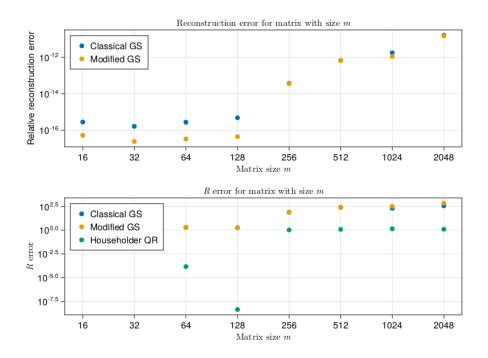
## (e) [5 pts]

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.

I used the method from Experiment 2 in Lecture 9 of Trefethen & Bau to create matrices with increasing condition number. I then computed the QR factorization of each matrix using the classical Gram-Schmidt algorithm, the modified Gram-Schmidt algorithm, and the Householder algorithm. For classical and modified GS, I plotted the relative reconstruction error of the original matrix. For all methods, I plotted the relative error of the  $\boldsymbol{R}$  matrix.

The top plot shows that the modified Gram-Schmidt algorithm is more stable than the classical Gram-Schmidt algorithm. The bottom plot shows how the Householder algorithm is more stable than both the classical and modified Gram-Schmidt algorithms, at least in so far as reproducing the  $\boldsymbol{R}$  matrix.

(A second plot below shows the improved stability of modified Gram-Schmidt more convincincly - I'm not sure why this last run produced such similar errors between modified original Gram-Schmidt.)



 $\textbf{Figure 1} \ \mathrm{Relative} \ \mathrm{errors} \ \mathrm{for} \ \mathrm{increasing} \ \mathrm{matrix} \ \mathrm{size} \ \mathrm{and} \ \mathrm{condition} \ \mathrm{number}$ 

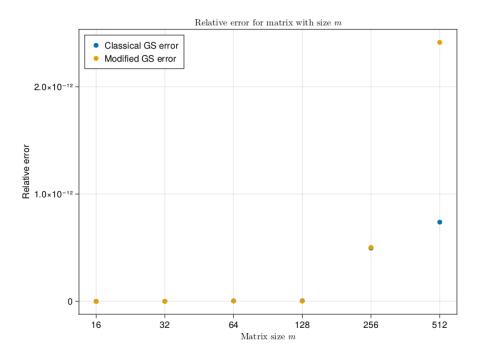


Figure 2 Relative errors for increasing matrix size and condition number