Classical Numerical Analysis, Chapter 03

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

Chapter 03
Systems of Linear Equations

Solving Linear Systems

In this chapter, we will be concerned with the following problem: Given the matrix $\mathsf{A} = [a_{i,j}] \in \mathbb{C}^{n\times n}$ and the vector $\mathbf{f} = [f_i] \in \mathbb{C}^n$, find $\mathbf{x} = [x_i] \in \mathbb{C}^n$ such that

$$\mathsf{A}\mathbf{x} = \mathbf{f}.$$

Of course, this is short–hand for the following system of linear equations

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = f_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = f_2, \\ \qquad\qquad\qquad\qquad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = f_n. \end{cases}$$

We call A the *coefficient matrix*. First of all, we need to make sure that a solution exists and is unique. The following result is nothing but a recapitulation of statements that the reader will have encountered before.

Theorem

The system $\mathsf{A}\mathbf{x} = \mathbf{f}$ *has a unique solution iff* $\det(\mathsf{A}) \neq 0$ *iff* $\mathsf{A}\mathbf{x} = \mathbf{0}$ *has only the trivial solution iff* A^{-1} *exists.*

Simple Systems

Triangular Systems

Let us, to be definite, consider the case when A is upper triangular. The system of equations reads

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = f_1, \\ \qquad\qquad a_{2,2}x_2 + \cdots + a_{2,n}x_n = f_2, \\ \qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad a_{n,n}x_n = f_n. \end{cases}$$

A unique solution exists iff $a_{i,i} \neq 0$, for all $i = 1, \ldots n$. In this case, the solution can be easily found by first computing the value of the last variable

$$x_n = f_n / a_{n,n}$$

and, after that, recursively computing

$$x_k = \frac{1}{a_{k,k}} \left(f_k - \sum_{j=k+1}^{n} a_{k,j} x_j \right), \quad k = n-1, n-2, \ldots, 2, 1.$$

The order of execution of this algorithm is vital: one must start with $k = n - 1$, and proceed in reverse order, finishing with $k = 1$. This algorithm is known as *back substitution*.

Tridiagonal Systems

Definition (tridiagonal matrix)

Let $\mathsf{A} = [a_{i,j}] \in \mathbb{C}^{n\times n}$. *We say that* A *is* **tridiagonal** *iff when* $i, j \in \{1, \dots, n\}$ *and* $|i-j| > 1$ *then* $a_{i,j} = 0$.

A tridiagonal system can be conveniently expressed as

$$a_k x_{k-1} + b_k x_k + c_k x_{k+1} = f_k, \quad k = 1, \dots, n,$$

with $a_1 = c_n = 0$. This can be visualized as

$$\begin{bmatrix} b_1 & c_1 & 0 & \cdots & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & & 0 \\ 0 & a_3 & b_3 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & b_{n-1} & c_{n-1} \\ 0 & 0 & \cdots & 0 & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{n-2} \\ f_{n-1} \\ f_n \end{bmatrix}.$$

See Listing 3.1 for an implementation of the Thomas algorithm for computing the solution to tridiagonal linear system.

LU Factorization

Properties of Triangular Matrices

Theorem

Let the matrices $\mathsf{T}, \mathsf{T}_k \in \mathbb{C}^{n\times n}$, *for* $k = 1, 2$, *be lower (upper) triangular. Then the following are true:*

1. *The product* $\mathsf{T}_1\mathsf{T}_2$ *is lower (upper) triangular.*
2. *If, in addition* $[\mathsf{T}_k]_{i,i} = 1$, *for* $k = 1, 2$ *and* $i = 1, \ldots, n$ *— i.e.,* $\mathsf{T}_1, \mathsf{T}_2 \in \mathbb{C}^{n\times n}$ *are unit lower (unit upper) triangular — then the product* $\mathsf{T}_1\mathsf{T}_2$ *is unit lower (unit upper) triangular.*
3. *The matrix* T *is non–singular iff* $[\mathsf{T}]_{i,i} \neq 0$, *for all* $i = 1, \ldots, n$.
4. *If* T *is non–singular,* $\mathsf{T}^{-1} \in \mathbb{C}^{n\times n}$ *is lower (upper) triangular.*
5. *If* T *is unit lower (unit upper) triangular, then it is invertible and* T^{-1} *is unit lower (unit upper) triangular.*
6. *If* $[\mathsf{T}]_{i,i} > 0$, *for* $i = 1, \ldots, n$, *then* $[\mathsf{T}^{-1}]_{i,i} = \frac{1}{[\mathsf{T}]_{i,i}} > 0$, *for* $i = 1, \ldots, n$.

Proof.

Homework exercise(s). □

Sub–Matrix of a Matrix

Definition

Suppose that $\mathsf{A} \in \mathbb{C}^{n \times n}$ *and* $S \subseteq \{1, 2, \ldots, n\}$ *is non–empty with cardinality* $k = \#(S) > 0$. *The* **sub–matrix** $\mathsf{A}(S) \in \mathbb{C}^{k \times k}$ *is that matrix obtained by deleting the columns and rows of* A *whose indices are not in* S. *In symbols,*

$$[\mathsf{A}(S)]_{i,j} = [\mathsf{A}]_{m_i, m_j}, \quad i, j = 1, \ldots, k,$$

where

$$S = \{m_1, \ldots, m_k\} \quad \text{and} \quad 1 \leq m_1 < m_2 < \cdots < m_k \leq n.$$

Definition

Let $\mathsf{A} \in \mathbb{C}^{n \times n}$ *and* $S = \{1, 2, \ldots, k\}$, *with* $k \leq n$. *Then, we define*

$$\mathsf{A}^{(k)} = \mathsf{A}(S) \in \mathbb{C}^{k \times k}$$

and we call $\mathsf{A}^{(k)}$ *the* **leading principal sub–matrix** *of* A *of order* k.

Example Sub–Matrices

Example

Suppose that

$$A = \begin{bmatrix} 1 & -7 & 12 & 4 \\ 6 & 9 & -3 & -4 \\ 1 & -6 & 8 & 9 \\ 4 & 4 & -11 & 17 \end{bmatrix}, \quad S = \{2,4\}.$$

Then $m_1 = 2$, $m_2 = 4$, and

$$A(S) = \begin{bmatrix} 9 & -4 \\ 4 & 17 \end{bmatrix}.$$

If $S = \{1,2\}$, then $m_1 = 1$, $m_2 = 2$, and

$$A(S) = \begin{bmatrix} 1 & -7 \\ 6 & 9 \end{bmatrix}.$$

Existence of the LU Factorization

Theorem

Let $n \geq 2$ and $\mathsf{A} \in \mathbb{C}^{n\times n}$. Suppose that all the leading principal sub–matrices of A are non–singular, that is, $\det(\mathsf{A}^{(k)}) \neq 0$, for all $k = 1, \ldots, n-1$. Then, there exists a unit lower triangular matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $\mathsf{U} \in \mathbb{C}^{n\times n}$ such that

$$\mathsf{A} = \mathsf{LU}.$$

Proof.

The proof is by induction on n, the size of the matrix.

($n = 2$): Consider

$$\mathsf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

with $a \neq 0$, by assumption.

Proof, Cont.

Define

$$\mathsf{L} = \begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}, \qquad \mathsf{U} = \begin{bmatrix} u & v \\ 0 & \eta \end{bmatrix},$$

where

$$u = a, \quad v = b, \quad m = \frac{c}{a}, \quad \eta = d - b\frac{c}{a}.$$

Then

$$mu = c, \qquad mv + \eta = d,$$

and consequently $\mathsf{A} = \mathsf{LU}$, as is easily confirmed.

$(n = m)$: The induction hypothesis is as follows: suppose that the result is valid for any $\mathsf{A} \in \mathbb{C}^{m\times m}$, provided $\mathsf{A}^{(k)}$ is non–singular for all $k = 1, \ldots, m-1$.

$(n = m+1)$: Suppose that $\mathsf{A}^{(k)}$ is non-singular for $k = 1, \ldots, m$. Set

$$\mathsf{A} = \begin{bmatrix} \mathsf{A}^{(m)} & \mathbf{b} \\ \mathbf{c}^{\mathsf{T}} & d \end{bmatrix} \in \mathbb{C}^{(m+1)\times(m+1)}.$$

Proof, Cont.

From the induction hypothesis, there is a unit lower triangular matrix $\mathsf{L}^{(m)}$ and an upper triangular matrix $\mathsf{U}^{(m)}$, such that $\mathsf{A}^{(m)} = \mathsf{L}^{(m)}\mathsf{U}^{(m)}$, where $\mathsf{A}^{(m)}$ is the leading principal sub–matrix of A of order m. Define

$$\mathsf{L} = \begin{bmatrix} \mathsf{L}^{(m)} & \mathbf{0} \\ \mathbf{m}^\intercal & 1 \end{bmatrix}, \quad \mathsf{U} = \begin{bmatrix} \mathsf{U}^{(m)} & \mathbf{v} \\ \mathbf{0}^\intercal & \eta \end{bmatrix},$$

where $\mathbf{b}, \mathbf{c}, \mathbf{m}, \mathbf{v}, \mathbf{0} \in \mathbb{C}^m$. Then

$$\mathsf{L}\mathsf{U} = \begin{bmatrix} \mathsf{L}^{(m)}\mathsf{U}^{(m)} & \mathsf{L}^{(m)}\mathbf{v} \\ \mathbf{m}^\intercal\mathsf{U}^{(m)} & \mathbf{m}^\intercal\mathbf{v} + \eta \end{bmatrix}.$$

Let us set this equal to A and determine whether or not the resulting equations are solvable. It is easy to see that $\mathsf{A} = \mathsf{L}\mathsf{U}$ iff

$$\mathsf{L}^{(m)}\mathsf{U}^{(m)} = \mathsf{A}^{(m)}, \qquad \mathsf{L}^{(m)}\mathbf{v} = \mathbf{b},$$

$$\mathbf{m}^\intercal\mathsf{U}^{(m)} = \mathbf{c}^\intercal, \qquad \mathbf{m}^\intercal\mathbf{v} + \eta = d.$$

Proof, Cont.

The last three equations are uniquely solvable, as we now show: since $\mathsf{L}^{(m)}$ is invertible,

$$\mathbf{v} = \left(\mathsf{L}^{(m)}\right)^{-1} \mathbf{b}.$$

The matrix $\mathsf{U}^{(m)}$ is invertible since

$$0 \neq \det(\mathsf{A}^{(m)}) = \det(\mathsf{L}^{(m)}\mathsf{U}^{(m)}) = \det(\mathsf{U}^{(m)}).$$

Hence,

$$\mathbf{m}^{\intercal} = \mathbf{c}^{\intercal} \left(\mathsf{U}^{(m)}\right)^{-1} \quad \text{or} \quad \mathbf{m} = \left(\mathsf{U}^{(m)}\right)^{-\intercal} \mathbf{c}.$$

Finally,

$$\eta = d - \mathbf{m}^{\intercal}\mathbf{v}.$$

The proof by induction is complete. $\square$

Using the LU Factorization to Solve Problems

Before we go any further, we ought to say why it is that an LU factorization of a matrix is useful. Suppose that we want to solve the indexed family of problems

$$\mathsf{A}\mathbf{x}^{(k)} = \mathbf{f}^{(k)}, \quad k = 1, \ldots, K,$$

and that there exists a unit lower triangular matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $\mathsf{U} \in \mathbb{C}^{n\times n}$ such that $\mathsf{A} = \mathsf{L}\mathsf{U}$. To find the solutions $\mathbf{x}^{(k)}$, we solve the follow equivalent family:

$$\mathsf{L}\mathbf{y}^{(k)} = \mathbf{f}^{(k)}, \qquad \mathsf{U}\mathbf{x}^{(k)} = \mathbf{y}^{(k)}, \quad k = 1, \ldots, K.$$

The vector $\mathbf{y}^{(k)}$ can be obtained easily and cheaply via forward substitution. Subsequently, the vector $\mathbf{x}^{(k)}$ can be obtained by back substitution.

Remember Gaussian Elimination?

Example

Consider the following system of linear equations

$$\begin{cases} x_1 + x_2 + x_3 = 6, \\ 2x_1 + 4x_2 + 2x_3 = 16, \\ -x_1 + 5x_2 - 4x_3 = -3. \end{cases}$$

Of course, we can represent this as a matrix–vector equation $A\mathbf{x} = \mathbf{f}$. We write this as an augmented matrix and perform Gaussian elimination to put the system into so–callled row echelon form.

$$[A|\mathbf{f}] = \left[\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 6 \\ 2 & 4 & 2 & 16 \\ -1 & 5 & -4 & -3 \end{array}\right] \xrightarrow{-2R_1+R_2\to R_2} \xrightarrow{1R_1+R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & \boxed{2} & 0 & 4 \\ 0 & 6 & -3 & 3 \end{array}\right]$$

$$\xrightarrow{-3R_2+R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -3 & -9 \end{array}\right].$$

Example (Cont.)

The boxed entries indicate the so called pivot elements. The values of the pivot elements help to determine the row multipliers in the algorithm. As long as these are non–zero, the algorithm can run to completion.

Let us focus on the left hand side of the augmented system, as this will be the important part with respect to the LU factorization. We have

$$L^{(3,2)}L^{(3,1)}L^{(2,1)}A = U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix},$$

where $L^{(2,1)}, L^{(3,1)}, L^{(3,2)}$ are elementary matrices encoding the elementary row operations performed in our Gaussian elimination process.

Example (Cont.)

To produce the matrix representations of these operations, recall that we need only to apply the corresponding elementary row operations on the identity matrix:

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{-2R_1+R_2\to R_2} \begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = \mathsf{L}^{(2,1)}.$$

Likewise,

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{1R_1+R_3\to R_3} \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix} = \mathsf{L}^{(3,1)},$$

and

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{-3R_2+R_3\to R_3} \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -3 & 1\end{bmatrix} = \mathsf{L}^{(3,2)}.$$

Example (Cont.)

Then, it is easy to see that

$$L^{(3,1)}L^{(2,1)}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 6 & -3 \end{bmatrix} \quad \text{and} \quad L^{(3,2)}L^{(3,1)}L^{(2,1)}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Observe that the order of application of $L^{(2,1)}$ and $L^{(3,1)}$ does not matter:

$$L^{(3,1)}L^{(2,1)} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = L^{(2,1)}L^{(3,1)}.$$

This will be shown to be true in general.

Gaussian Elimination and Elementary Matrices

Suppose $A \in \mathbb{C}^{n\times n}$, where $n \geq 2$. If Gaussian elimination for A proceeds to completion without encountering any zero pivots, then one gets

$$L^{(n,n-1)} \cdots L^{(n,2)} \cdots L^{(3,2)} L^{(n,1)} \cdots L^{(2,1)} A = U,$$

where U is square and upper triangular. Moreover, since we assume that no zero pivot entries are encountered, $[U]_{i,i} \neq 0$, for $i = 1, \cdots, n-1$. However, it is possible that $[U]_{n,n} = 0$. We can group the elementary operations into column operations as follows:

$$\begin{aligned}
&\text{(column 1):} && L_1 = L^{(n,1)} \cdots L^{(2,1)}, \\
&\text{(column 2):} && L_2 = L^{(n,2)} \cdots L^{(3,2)}, \\
& && \vdots \\
&\text{(column } n-2\text{):} && L_{n-2} = L^{(n,n-2)} L^{(n-1,n-2)}, \\
&\text{(column } n-1\text{):} && L_{n-1} = L^{(n,n-1)},
\end{aligned}$$

so that

$$L_{n-1} L_{n-2} \cdots L_2 \, L_1 A = U.$$

The Inverse of an Elementary Matrix

What is the inverse of an elementary matrix? Suppose that

$$\mathsf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathsf{L}' = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then, it is easy to see that $\mathsf{L}'\mathsf{L} = \mathsf{I}_3$. Or, in other words, $\mathsf{L}' = \mathsf{L}^{-1}$.

Rigorous Definition of an Elementary Matrix

Definition

A matrix $\mathsf{E} \in \mathbb{C}^{n\times n}$ *is called* **elementary** *iff* $\mathsf{E} = \mathsf{I} + \mu_{r,s}\mathsf{M}^{(r,s)}$ *for some* $\mu_{r,s} \in \mathbb{C}$*, and for some* $1 \leq s < r \leq n$*, where*

$$\mathsf{M}^{(r,s)} = \mathbf{e}_r\mathbf{e}_s^{\mathsf{T}},$$

that is,

$$\left[\mathsf{M}^{(r,s)}\right]_{i,j} = \delta_{i,r}\delta_{j,s}.$$

Example

$$\mathsf{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Here $r = 3$, $s = 2$, and $\mu_{3,2} = 3$.

Basic Properties of Elementary Matrices

Proposition

Suppose that

$$\mathsf{E}_k = \mathsf{I} + \mu_{r_k,s}\mathsf{M}^{(r_k,s)}, \quad k = 1, 2,$$

are two elementary matrices with $r_1 \neq r_2$*. Then, both matrices are invertible, the inverses are elementary, and the matrices commute. Furthermore,*

$$\mathsf{E}_k^{-1} = \mathsf{I} - \mu_{r_k,s}\mathsf{M}^{(r_k,s)},$$

$$(\mathsf{E}_1\mathsf{E}_2)^{-1} = \mathsf{E}_2^{-1}\mathsf{E}_1^{-1} = \mathsf{E}_1^{-1}\mathsf{E}_2^{-1} = \mathsf{I} - \mu_{r_1,s}\mathsf{M}^{(r_1,s)} - \mu_{r_2,s}\mathsf{M}^{(r_2,s)},$$

and

$$\mathsf{E}_1\mathsf{E}_2 = \mathsf{E}_2\mathsf{E}_1 = \mathsf{I} + \mu_{r_1,s}\mathsf{M}^{(r_1,s)} + \mu_{r_2,s}\mathsf{M}^{(r_2,s)},$$

Proof.

Homework exercise. □

Column Complete Elementary Matrix

Definition

Suppose $n \geq 2$ and let the index $s \in \{1, 2, \ldots, n-1\}$ be given. The matrix $\mathsf{F} \in \mathbb{C}^{n \times n}$ is called a **column–s complete elementary matrix** *iff*

$$\mathsf{F} = \mathsf{I} + \sum_{r=s+1}^{n} \mu_{r,s} \mathsf{M}^{(r,s)},$$

for some scalars $\mu_{r,s} \in \mathbb{C}$, $r = s+1, \ldots, n$. In other words, F *is a unit lower triangular matrix of the form*

$$\mathsf{F} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \mu_{s+1,s} & 1 & & \\ & & \vdots & & \ddots & \\ & & \mu_{n,s} & & & 1 \end{bmatrix}.$$

Rigorous Definition of Gaussian Elimination

Definition

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ *be given with* $n \geq 2$*. We define the* **Gaussian elimination** *algorithm recursively as follows. Suppose that* k *stages of Gaussian elimination have been completed, where* $k \in \{0, \ldots, n-1\}$*, such that no zero pivots have been encountered, producing the matrix factorization*

$$\mathsf{L}_k \cdots \mathsf{L}_1 \mathsf{A} = \mathsf{A}^{(k)}, \quad k = 1, \ldots, n-1,$$

Definition (Cont.)

where, $\mathsf{A}^{(0)} = \mathsf{A}$, *and, for* $k = 1, \ldots, n-1$,

$$\mathsf{A}^{(k)} = \begin{bmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & a_{1,4}^{(0)} & \cdots & a_{1,k+1}^{(0)} & \cdots & a_{1,n}^{(0)} \\ 0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & a_{2,4}^{(1)} & \cdots & a_{2,k+1}^{(1)} & \cdots & a_{2,n}^{(1)} \\ 0 & 0 & a_{3,3}^{(2)} & a_{3,4}^{(2)} & \cdots & a_{3,k+1}^{(2)} & \cdots & a_{3,n}^{(2)} \\ \vdots & & \ddots & \ddots & \ddots & \vdots & & \vdots \\ 0 & & & 0 & a_{k,k}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{k,n}^{(k-1)} \\ 0 & & & 0 & 0 & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ \vdots & & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & a_{n,k+1}^{(k)} & \cdots & a_{n,n}^{(k)} \end{bmatrix}.$$

If $k = n-1$, *we are done, and we set* $\mathsf{U} = \mathsf{A}^{(n-1)}$. *Otherwise, if the* $(k+1)$*–st pivot entry,* $a_{k+1,k+1}^{(k)}$, *is not equal to zero, the algorithm may proceed.*

Definition (Cont.)

Construct the column–$(k+1)$ complete elementary matrix

$$\mathsf{L}_{k+1} = \mathsf{I} + \sum_{r=k+2}^{n} \mu_{r,k+1} \mathsf{M}^{(r,k+1)}, \quad \mu_{r,k+1} = -\frac{a_{r,k+1}^{(k)}}{a_{k+1,k+1}^{(k)}}, \quad r = k+2, \ldots, n.$$

Then, set $\mathsf{L}_{k+1}\mathsf{A}^{(k)} = \mathsf{A}^{(k+1)}$, *obtaining*

$$\mathsf{A}^{(k+1)} = \begin{bmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & a_{1,4}^{(0)} & \cdots & a_{1,k+1}^{(0)} & a_{1,k+2}^{(0)} & \cdots & a_{1,n}^{(0)} \\ 0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & a_{2,4}^{(1)} & \cdots & a_{2,k+1}^{(1)} & a_{2,k+2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ 0 & 0 & a_{3,3}^{(2)} & a_{3,4}^{(2)} & \cdots & a_{3,k+1}^{(2)} & a_{3,k+2}^{(2)} & \cdots & a_{3,n}^{(2)} \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & & 0 & a_{k,k}^{(k-1)} & a_{k,k+1}^{(k-1)} & a_{k,k+2}^{(k-1)} & \cdots & a_{k,n}^{(k-1)} \\ 0 & & & 0 & 0 & a_{k+1,k+1}^{(k)} & a_{k+1,k+2}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ 0 & & & 0 & 0 & 0 & a_{k+2,k+2}^{(k+1)} & \cdots & a_{k+2,n}^{(k+1)} \\ \vdots & & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & a_{n,k+2}^{(k+1)} & \cdots & a_{n,n}^{(k+1)} \end{bmatrix}.$$

The Result of Gaussian Elimination

Theorem

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$. *If Gaussian elimination proceeds to completion without encountering any zero pivots, then there are column–k complete elementary matrices* $\mathsf{L}_k \in \mathbb{C}^{n\times n}$, *for* $k = 1, \ldots, n-1$, *such that*

$$\mathsf{L}_{n-1} \cdots \mathsf{L}_2 \mathsf{L}_1 \mathsf{A} = \mathsf{U},$$

where $\mathsf{U} \in \mathbb{C}^{n\times n}$ *is upper triangular and*

$$[\mathsf{U}]_{i,i} \neq 0, \quad i = 1, \ldots, n-1,$$

since no zero pivots are encountered. Furthermore,

$$\mathsf{A} = \mathsf{L}_1^{-1} \cdots \mathsf{L}_{n-1}^{-1} \mathsf{U} = \mathsf{L}\mathsf{U},$$

where L *is unit lower triangular.*

Theorem (Cont.)

Writing

$$\mathsf{L}_k = \mathsf{I} + \sum_{r=k+1}^{n} \mu_{r,s} \mathsf{M}^{(r,k)},$$

it follows that

$$\mathsf{L} = \mathsf{I} - \sum_{k=1}^{n-1} \sum_{r=k+1}^{n} \mu_{r,k} \mathsf{M}^{(r,k)}.$$

In other words,

$$\mathsf{L} = \begin{bmatrix} 1 & & & & \\ -\mu_{2,1} & 1 & & & \\ -\mu_{3,1} & -\mu_{3,2} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ -\mu_{n,1} & -\mu_{n,2} & \cdots & -\mu_{n,n-1} & 1 \end{bmatrix}.$$

Theorem (uniqueness)

Suppose that $n \geq 2$ and suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ is invertible. Suppose that there is a unit lower triangular matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $\mathsf{U} \in \mathbb{C}^{n\times n}$ such that $\mathsf{A} = \mathsf{L}\mathsf{U}$. Then, this LU factorization is unique.

Proof.

Suppose that there are two factorizations with the desired properties:

$$\mathsf{L}_1\mathsf{U}_1 = \mathsf{A} = \mathsf{L}_2\mathsf{U}_2.$$

Since A is invertible U_1 and U_2 must be invertible, i.e., there are no zeros on their diagonals. Furthermore,

$$\mathsf{L}_2^{-1}\mathsf{L}_1 = \mathsf{U}_2\mathsf{U}_1^{-1} = \mathsf{D},$$

where D is by necessity diagonal. Therefore,

$$\mathsf{L}_1 = \mathsf{L}_2\mathsf{D}.$$

But, it must be that $\mathsf{D} = \mathsf{I}_n$, since the diagonal elements of L_1 and L_2 are all ones. □

Complexity of LU

Proposition

Let $A \in \mathbb{C}^{n\times n}$, then the LU factorization algorithm requires, to leading order, $\frac{2}{3}n^3$ operations.

Proof.

We only care about the leading order of operations, which, from the Matlab code at the end of the chapter, are in number, roughly

$$
\begin{aligned}
\sum_{k=1}^{n-1}\sum_{j=k}^{n}\sum_{t=k}^{n} 2 &= 2\sum_{k=1}^{n-1}(n-k)\sum_{j=k+1}^{n} 1 \\
&\approx 2\sum_{k=1}^{n-1}(n-k)^2 \\
&= \frac{1}{3}(n-1)n(2n-1) \\
&\approx \frac{2}{3}n^3.
\end{aligned}
$$

□

Column Pivoting

Example

Suppose that $A \in \mathbb{C}^{3\times 3}$ is given by

$$A = \begin{bmatrix} 0 & 1 & 5 \\ -2 & 1 & 1 \\ 4 & -2 & 6 \end{bmatrix}.$$

In the following algorithm, let us agree to interchange rows so that the largest in modulus element in the column at or below the pivot position moves into the pivot position. This is called Gaussian elimination with maximal column pivoting:

$$A = \begin{bmatrix} \boxed{0} & 1 & 5 \\ -2 & 1 & 1 \\ 4 & -2 & 6 \end{bmatrix} \xrightarrow{R_1 \leftrightarrows R_3} \begin{bmatrix} \boxed{4} & -2 & 6 \\ -2 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_1 + R_2 \to R_2} \xrightarrow{0R_1 + R_3 \to R_3} \begin{bmatrix} 4 & -2 & 6 \\ 0 & \boxed{0} & 4 \\ 0 & 1 & 5 \end{bmatrix}$$

Example (Cont.)

$$\xrightarrow{R_2 \leftrightarrows R_3} \begin{bmatrix} 4 & -2 & 6 \\ 0 & \boxed{1} & 5 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\xrightarrow{0R_2 + R_3 \to R_3} \begin{bmatrix} 4 & -2 & 6 \\ 0 & 1 & 5 \\ 0 & 0 & 4 \end{bmatrix}.$$

Our procedure may be expressed as

$$\mathsf{L}_2\mathsf{P}_2\mathsf{L}_1\mathsf{P}_1\mathsf{A} = \mathsf{U},$$

where P_1 and P_2 are simple permutation matrices, and L_1 and L_2 are column-1 and column–2 complete elementary matrices, respectively.

Permutation Matrices

Definition (permutation)

A matrix $\mathsf{P} \in \mathbb{C}^{n\times n}$ *is called a* **simple permutation** *matrix iff it is obtained from the* $n \times n$ *identity matrix* I *by interchanging exactly two rows of* I*.* P *is called a* **regular permutation** *(or just a* **permutation***) matrix iff* P *is the product of simple permutation matrices.*

Proposition (action of permutations)

Let $n \geq 2$*. Suppose that* $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is any matrix and* $\mathsf{P} \in \mathbb{C}^{n\times n}$ *is a simple permutation matrix obtained by interchanging rows* r *and* s *of the identity matrix, with* $1 \leq r < s \leq n$*. Then* PA *is identical to* A*, except with rows* r *and* s *interchanged. Furthermore,* AP *is identical to* A*, except with columns* r *and* s *interchanged.*

Properties of Permutation Matrices

Lemma

Suppose that $\mathsf{P}, \mathsf{Q} \in \mathbb{C}^{n\times n}$, *with* $n \geq 2$, *are permutation matrices. Then*

1. *The product* PQ *is a permutation matrix.*
2. $\det(\mathsf{P}) = \pm 1$ *according to whether* P *is the product of an even* ($\det(\mathsf{P}) = 1$) *or an odd* ($\det(\mathsf{P}) = -1$) *number of simple permutation matrices.*
3. *The inverse of a simple permutation matrix is itself. Any regular permutation matrix* P *is invertible, and, if*
$$\mathsf{P} = \mathsf{P}_1 \mathsf{P}_2 \cdots \mathsf{P}_k,$$
where P_i *is a simple permutation matrix, for* $1 \leq i \leq k$, *then*
$$\mathsf{P}^{-1} = \mathsf{P}_k \cdots \mathsf{P}_2 \mathsf{P}_1 = \mathsf{P}^{\intercal}.$$

Gaussian Elimination with Column Pivoting

Example

Let us continue with our 3×3 example, but in general terms. We have

$$\mathsf{L}_2\mathsf{P}_2\mathsf{L}_1\mathsf{P}_1\mathsf{A} = \mathsf{U},$$

where P_j, $j = 1, 2$ are simple permutation matrices or the identity matrix (in the case that no row interchange took place) and L_j are column–j complete elementary matrices, $j = 1, 2$. Now, observe that

$$\mathsf{L}_2\mathsf{P}_2\mathsf{L}_1\mathsf{P}_2\mathsf{P}_2\mathsf{P}_1\mathsf{A} = \mathsf{U}.$$

Therefore,

$$\hat{\mathsf{L}}_2\hat{\mathsf{L}}_1\mathsf{P}\mathsf{A} = \mathsf{U},$$

where

$$\hat{\mathsf{L}}_2 = \mathsf{L}_2, \quad \hat{\mathsf{L}}_1 = \mathsf{P}_2\mathsf{L}_1\mathsf{P}_2, \quad \mathsf{P} = \mathsf{P}_2\mathsf{P}_1.$$

Gaussian Elimination with Column Pivoting

Example

Suppose that Gaussian elimination with maximal column pivoting is applied to $A \in \mathbb{C}^{4\times 4}$. Then, it should be clear that one obtains

$$L_3 P_3 L_2 P_2 L_1 P_1 A = U,$$

which can be rewritten as

$$\hat{L}_3 \hat{L}_2 \hat{L}_1 P A = U,$$

where

$$\hat{L}_3 = L_3, \quad \hat{L}_2 = P_3 L_2 P_3, \quad \hat{L}_1 = P_3 P_2 L_1 P_2 P_3 \quad P = P_3 P_2 P_1.$$

Permutations and Elementary Matrices

Proposition

Suppose that $\mathsf{L}_k \in \mathbb{C}^{n\times n}$ is a column–k complete elementary matrix,

$$\mathsf{L}_k = \mathsf{I}_n + \sum_{r=k+1}^{n} \mu_{r,k}\mathsf{M}^{(r,k)},$$

for some constants $\mu_{r,k} \in \mathbb{C}$, for $k = k+1, \ldots, n$. Assume that $\mathsf{Q} \in \mathbb{C}^{n\times n}$ is a simple permutation matrix encoding the interchange of rows r' and s', where $k < r' < s' \leq n$. Then the matrix $\mathsf{QL}_k\mathsf{Q}$ is a column–k complete matrix. In particular, $\mathsf{QL}_k\mathsf{Q}$ is identical to L_k, except that entries $\mu_{r',k}$ and $\mu_{s',k}$ are interchanged.

Proof.

It follows that

$$\mathsf{QL}_k\mathsf{Q} = \mathsf{QI}_n\mathsf{Q} + \sum_{r=k+1}^{n} \mu_{r,k}\mathsf{Q}\mathbf{e}_r\mathbf{e}_k^{\intercal}\mathsf{Q}.$$

Proof, Cont.

But, observe that $\mathbf{e}_k^\intercal Q = \mathbf{e}_k^\intercal$ and

$$Q\mathbf{e}_r = \begin{cases} \mathbf{e}_{s'}, & r = r', \\ \mathbf{e}_{r'}, & r = s', \\ \mathbf{e}_r & r \in \{1, \ldots n\} \setminus \{r', s'\}. \end{cases}$$

Therefore,

$$QL_kQ = I_n + \sum_{\substack{r=k+1 \\ r \neq r', s'}}^{n} \mu_{r,k}\mathbf{e}_r\mathbf{e}_k^\intercal + \mu_{s',k}\mathbf{e}_{r'}\mathbf{e}_k^\intercal + \mu_{r',k}\mathbf{e}_{s'}\mathbf{e}_k^\intercal.$$

In other words, QL_kQ is a column–k complete elementary matrix that is identical to L_k, except that the positions of $\mu_{r',k}$ are $\mu_{s',k}$ are swapped. □

Theorem (LU factorization with pivoting)

Suppose that $n \geq 2$ *and* $\mathsf{A} \in \mathbb{C}^{n\times n}$. *The Gaussian elimination with maximal column pivoting algorithm always proceeds to completion to yield an upper triangular matrix* U. *In particular, there are matrices* $\mathsf{L}_j, \mathsf{P}_j \in \mathbb{C}^{n\times n}$, $j = 1, \ldots, n-1$, *such that*

$$\mathsf{L}_{n-1}\mathsf{P}_{n-1} \cdots \mathsf{L}_2\mathsf{P}_2\mathsf{L}_1\mathsf{P}_1\mathsf{A} = \mathsf{U},$$

where L_j *is a column–*j *complete elementary matrix, and* P_j *is either the* $n \times n$ *identity or a simple permutation matrix. Furthermore, there are column–*j *complete elementary matrices* $\hat{\mathsf{L}}_j$, *for* $j = 1, \ldots, n-1$, *and a permutation matrix* P, *such that*

$$\hat{\mathsf{L}}_{n-1} \cdots \hat{\mathsf{L}}_1\mathsf{P}\mathsf{A} = \mathsf{U},$$

where

$$\mathsf{P} = \mathsf{P}_{n-1} \cdots \mathsf{P}_1,$$

$$\hat{\mathsf{L}}_j = \mathsf{P}_{n-1} \cdots \mathsf{P}_{j+1}\mathsf{L}_j\mathsf{P}_{j+1} \cdots \mathsf{P}_{n-1}, \quad j = 1, \ldots, n-2, \quad \hat{\mathsf{L}}_{n-1} = \mathsf{L}_{n-1}.$$

Finally, there is a unit lower triangular matrix L *such that*

$$\mathsf{P}\mathsf{A} = \mathsf{L}\mathsf{U}.$$

Special Matrices

Diagonal Dominance

Definition

A matrix $\mathsf{A} = [a_{i,j}] \in \mathbb{C}^{n\times n}$ *is called* **diagonally dominant** *iff*

$$|a_{i,i}| \geq \sum_{\substack{k=1 \\ k\neq i}}^{n} |a_{i,k}| \,, \quad \forall\, i = 1, \ldots, n.$$

A *is* **strictly diagonally dominant** *(SDD) iff*

$$|a_{i,i}| > \sum_{\substack{k=1 \\ k\neq i}}^{n} |a_{i,k}| \,, \quad \forall\, i = 1, \ldots, n.$$

A *is called* **strictly diagonally dominant of dominance** δ *iff there is a* $\delta > 0$ *such that*

$$|a_{i,i}| \geq \delta + \sum_{\substack{k=1 \\ k\neq i}}^{n} |a_{i,k}| \,, \quad \forall\, i = 1, \ldots, n.$$

Theorem (properties of a strictly diagonally dominant matrix)

If $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is strictly diagonally dominant, then* A *is invertible. If* A *is SDD of dominance* $\delta > 0$*, then*

$$\left\|\mathsf{A}^{-1}\right\|_\infty < \frac{1}{\delta}.$$

Proof.

Suppose that A is singular. If that is the case there is an $\mathbf{x} = [x_i] \in \mathbb{C}^n_\star$, such that $\mathsf{A}\mathbf{x} = \mathbf{0}$. Suppose that $k \in \{1, \dots, n\}$ is an index for which $|x_k| = \|\mathbf{x}\|_\infty$. Since $\mathsf{A}\mathbf{x} = \mathbf{0}$, we must have that, for each $i = 1, \dots, n$

$$\sum_{j=1}^{n} a_{i,j}x_j = 0.$$

In particular, $\sum_{j=1}^{n} a_{k,j}x_j = 0$. Then, from the triangle inequality,

$$|a_{k,k}| \cdot \|\mathbf{x}\|_\infty = |a_{k,k}x_k| = \left| - \sum_{\substack{j=1\\ j\neq k}}^{n} a_{k,j}x_j \right| \le \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{k,j}| \cdot |x_j| \le \|\mathbf{x}\|_\infty \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{k,j}|.$$

Proof, Cont.

Since $\|\mathbf{x}\|_\infty > 0$, we have

$$|a_{k,k}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{k,j}|.$$

This proves that A is not SDD, a contradiction.

Next, suppose that A has dominance $\delta > 0$. Let $\mathbf{x}$ be arbitrary and set $\mathsf{A}\mathbf{x} = \mathbf{f}$. Assume that $\|\mathbf{x}\|_\infty = |x_k|$, for some $k = 1, \ldots, n$. Then

$$a_{k,1}x_1 + \cdots + a_{k,k}x_k + \cdots + a_{k,n}x_n = f_k,$$

and, using the reverse triangle inequality,

$$|f_k| \geq |a_{k,k}||x_k| - \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{j,k}||x_j| \geq \left(|a_{k,k}| - \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{j,k}| \right) |x_k| \geq \delta \|\mathbf{x}\|_\infty.$$

Proof, Cont.

This shows that

$$\frac{\left\|\mathsf{A}\mathbf{x}\right\|_{\infty}}{\left\|\mathbf{x}\right\|_{\infty}} \geq \delta, \quad \forall\, \mathbf{x} \in \mathbb{C}^n,$$

which is equivalent to

$$\frac{1}{\delta} \geq \frac{\left\|\mathsf{A}^{-1}\mathbf{w}\right\|_{\infty}}{\left\|\mathbf{w}\right\|_{\infty}}, \quad \forall\, \mathbf{w} \in \mathbb{C}^n_{\star}.$$

This, in turn, implies that

$$\frac{1}{\delta} \geq \sup_{\mathrm{w} \in \mathbb{C}^n_{\star}} \frac{\left\|\mathsf{A}^{-1}\mathbf{w}\right\|_{\infty}}{\left\|\mathbf{w}\right\|_{\infty}} = \left\|\mathsf{A}^{-1}\right\|_{\infty},$$

as we intended to show. □

Gaussian elimination and SDD Matrices

Theorem

Let $\mathsf{A} = [a_{i,j}] \in \mathbb{C}^{n\times n}$ *be strictly diagonally dominant (SDD), and assume it is represented as*

$$\mathsf{A} = \begin{bmatrix} \alpha & \mathbf{v}^{\intercal} \\ \mathbf{p} & \hat{\mathsf{A}} \end{bmatrix},$$

where $\alpha \in \mathbb{C}$, $\mathbf{p}, \mathbf{v} \in \mathbb{C}^{n-1}$, *and* $\hat{\mathsf{A}} = [\hat{a}_{i,j}] \in \mathbb{C}^{(n-1)\times(n-1)}$. *After one step of Gaussian elimination (without pivoting),* A *will be reduced to the matrix*

$$\begin{bmatrix} \alpha & \mathbf{v}^{\intercal} \\ \mathbf{0} & \mathsf{B} \end{bmatrix},$$

where $\mathsf{B} = [b_{i,j}] \in \mathbb{C}^{(n-1)\times(n-1)}$ *is SDD.*

Proof.

Let us construct a matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ such that, if possible,

$$\mathsf{LA} = \begin{bmatrix} \alpha & \mathbf{v}^{\intercal} \\ \mathbf{0} & \mathsf{B} \end{bmatrix}.$$

Proof, Cont.

Consider

$$\mathsf{L} = \begin{bmatrix} 1 & \mathbf{0}^\intercal \\ \mathbf{m} & \mathsf{I}_{n-1} \end{bmatrix}.$$

Then

$$\mathsf{L}\mathsf{A} = \begin{bmatrix} 1 & \mathbf{0}^\intercal \\ \mathbf{m} & \mathsf{I}_{n-1} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{v}^\intercal \\ \mathbf{p} & \hat{\mathsf{A}} \end{bmatrix} = \begin{bmatrix} \alpha & \mathbf{v}^\intercal \\ \alpha\mathbf{m} + \mathbf{p} & \mathbf{m}\mathbf{v}^\intercal + \hat{\mathsf{A}} \end{bmatrix}.$$

Note that since A is SDD, $\alpha \neq 0$ and $\hat{\mathsf{A}}$ is SDD. Choosing $\mathbf{m} = -\alpha^{-1}\mathbf{p}$, we have

$$\mathsf{L}\mathsf{A} = \begin{bmatrix} \alpha & \mathbf{v}^\intercal \\ \mathbf{0} & \hat{\mathsf{A}} - \alpha^{-1}\mathbf{p}\mathbf{v}^\intercal \end{bmatrix}.$$

Having successfully constructed L, we find that $\mathsf{B} = \hat{\mathsf{A}} - \alpha^{-1}\mathbf{p}\mathbf{v}^\intercal$. All that remains is to show that B is SDD. To see that this is the case, consider for row i of B

Proof, Cont.

$$
\begin{aligned}
\sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|b_{i,j}\right| &= \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|\hat{a}_{i,j} - \alpha^{-1} p_i v_j\right| \\
&\leq \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|\hat{a}_{i,j}\right| + \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|\alpha^{-1} p_i v_j\right| \\
&= \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|\hat{a}_{i,j}\right| + \frac{|p_i|}{|\alpha|} \sum_{\substack{j=1\\ j\neq i}}^{n-1} |v_j| \\
&= \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|a_{i+1,j+1}\right| + \frac{|a_{i+1,1}|}{|a_{1,1}|} \sum_{\substack{j=1\\ j\neq i}}^{n-1} \left|a_{1,j+1}\right| \\
&= \sum_{\substack{j=2\\ j\neq i+1}}^{n} \left|a_{i+1,j}\right| + \frac{|a_{i+1,1}|}{|a_{1,1}|} \sum_{\substack{j=2\\ j\neq i+1}}^{n} \left|a_{1,j}\right| .
\end{aligned}
$$

Proof, Cont.

Thus,

$$\sum_{\substack{j=1\\ j\neq i}}^{n-1} |b_{i,j}| \leq \sum_{\substack{j=1\\ j\neq i+1}}^{n} |a_{i+1,j}| - |a_{i+1,1}| + \frac{|a_{i+1,1}|}{|a_{1,1}|} \sum_{j=2}^{n} |a_{1,j}| - \frac{|a_{i+1,1}| \cdot |a_{1,i+1}|}{|a_{1,1}|}.$$

Now, since A is SDD, we can continue this string of inequalities to obtain

$$\begin{aligned}
\sum_{\substack{j=1\\ j\neq i}}^{n-1} |b_{i,j}| &< |a_{i+1,i+1}| - |a_{i+1,1}| + |a_{i+1,1}| - \frac{|a_{i+1,1}| \cdot |a_{1,i+1}|}{|a_{1,1}|} \\
&= |a_{i+1,i+1}| - \frac{|a_{i+1,1}| \cdot |a_{1,i+1}|}{|a_{1,1}|} \\
&\leq \left| a_{i+1,i+1} - \frac{a_{i+1,1} a_{1,i+1}}{a_{1,1}} \right| \\
&= |b_{i,i}|,
\end{aligned}$$

where we used the reverse triangle inequality. Thus, as claimed, B is SDD. □

Gaussian Elimination of SDD Matrices

Corollary (SDD of magnitude δ)

If $\mathsf{A} \in \mathbb{C}^{n\times n}$ is SDD of magnitude $\delta > 0$, then $\mathsf{B} \in \mathbb{C}^{(n-1)\times(n-1)}$, introduced above, is SDD of magnitude δ.

Corollary (Gaussian elimination and SDD)

If $\mathsf{A} \in \mathbb{C}^{n\times n}$ is SDD, then Gaussian elimination without pivoting applied to A proceeds to completion without encountering any zero pivot elements.

HPD Matrices

Definition

A matrix $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is called* **Hermitian positive semi–definite** *(HPSD) iff* $\mathsf{A} = \mathsf{A}^{\mathsf{H}}$ *and*

$$\mathbf{x}^{\mathsf{H}}\mathsf{A}\mathbf{x} \geq 0, \quad \forall \mathbf{x} \in \mathbb{C}^n.$$

A *is called* **Hermitian positive definite** *(HPD) iff* $\mathsf{A} = \mathsf{A}^{\mathsf{H}}$ *and*

$$\mathbf{x}^{\mathsf{H}}\mathsf{A}\mathbf{x} > 0, \quad \forall \mathbf{x} \in \mathbb{C}^n_{\star}.$$

Properties of HPD Matrices

Theorem

Suppose that $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$ *is HPD. Then*

1. $a_{i,i} > 0$, *for all* $i = 1, \dots, n$.
2. $\sigma(A) \subset (0, \infty)$.
3. $\det(A) > 0$ *and* $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{i,i} > 0$.
4. *For all* $\emptyset \neq S \subseteq \{1, \dots, n\}$ *we have that* $A(S)$ *is HPD.*
5. $|a_{i,j}|^2 \leq a_{i,i} a_{j,j}$, *for all* $i \neq j$.
6. $\max_{1 \leq i,j \leq n} |a_{i,j}| \leq \max_{1 \leq i \leq n} |a_{i,i}|$.

Theorem (factorization of HPD matrices)

Let $n \geq 2$ and suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ is HPD. There exists a unit lower triangular matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ and a diagonal matrix $\mathsf{D} \in \mathbb{R}^{n\times n}$ with positive diagonal entries such that

$$\mathsf{A} = \mathsf{L}\mathsf{D}\mathsf{L}^{\mathsf{H}}.$$

Proof.

Since A is HPD, for $k = 1, \ldots, n-1$ the principal sub–matrices $\mathsf{A}^{(k)} \in \mathbb{C}^{k\times k}$, are invertible. By a theorem, there exists a unit lower triangular matrix L and an upper triangular matrix U such that $\mathsf{A} = \mathsf{L}\mathsf{U}$.

Next, we claim that all of the diagonal elements of U are real and positive. This follows by an induction argument and the fact that $\mathbf{det}(\mathsf{A}) = \mathbf{det}(\mathsf{U})$. The details are left to the reader.

Now, set $\mathsf{D} = \mathbf{diag}(u_{1,1}, \ldots, u_{n,n})$ and $\tilde{\mathsf{U}} = \mathsf{D}^{-1}\mathsf{U}$. Then

$$\mathsf{A} = \mathsf{L}\mathsf{D}\mathsf{D}^{-1}\mathsf{U} = \mathsf{L}\mathsf{D}\tilde{\mathsf{U}}.$$

Proof, Cont.

It follows that, for $i = 1, \ldots, n$, $\tilde{u}_{i,i} = 1$. Taking the conjugate transpose, we have

$$\tilde{\mathsf{U}}^{\mathsf{H}}\mathsf{D}\mathsf{L}^{\mathsf{H}} = \mathsf{A}^{\mathsf{H}} = \mathsf{A} = \mathsf{L}\mathsf{D}\tilde{\mathsf{U}}.$$

Thus,

$$\mathsf{L}^{-1}\tilde{\mathsf{U}}^{\mathsf{H}}\mathsf{D} = \mathsf{D}\tilde{\mathsf{U}}\mathsf{L}^{-\mathsf{H}}. \tag{1}$$

Recall that, by a theorem, $\mathsf{L}^{-1}\tilde{\mathsf{U}}^{\mathsf{H}}$ must be unit lower triangular, and $\tilde{\mathsf{U}}\mathsf{L}^{-\mathsf{H}}$ must be unit upper triangular. The only way for (1) to hold is for $\mathsf{L}^{-1}\tilde{\mathsf{U}}^{\mathsf{H}}$ and $\tilde{\mathsf{U}}\mathsf{L}^{-\mathsf{H}}$ to be diagonal. But as these products must be unit triangular, they are both equal to the identity. In other words,

$$\mathsf{L} = \tilde{\mathsf{U}}^{\mathsf{H}},$$

and we have proven that

$$\mathsf{A} = \mathsf{L}\mathsf{D}\mathsf{L}^{\mathsf{H}}. \qquad \square$$

Cholesky Factorization

Corollary

Let $n \geq 2$ and suppose that $\mathsf{A} \in \mathbb{C}^{n \times n}$ is HPD. Then, there is a lower triangular matrix $\mathsf{L} \in \mathbb{C}^{n \times n}$ such that

$$\mathsf{A} = \mathsf{L}\mathsf{L}^{\mathsf{H}}.$$

This is known as the Cholesky factorization.

Proof.

From the last theorem, there is a unit lower triangular matrix $\tilde{\mathsf{L}}$ and a diagonal matrix D with positive real diagonal entries, such that

$$\mathsf{A} = \tilde{\mathsf{L}}\mathsf{D}\tilde{\mathsf{L}}^{\mathsf{H}}.$$

Suppose that $\mathsf{D} = \mathbf{diag}(d_1, \ldots, d_n)$, and define $\tilde{\mathsf{D}} = \mathbf{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n})$. Then, setting $\mathsf{L} = \tilde{\mathsf{L}}\tilde{\mathsf{D}}$, we see that

$$\mathsf{A} = \mathsf{L}\mathsf{L}^{\mathsf{H}}.$$

The proof is complete. $\square$

Theorem (uniqueness of Cholesky factorization)

Let $n \geq 2$ and suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ is HPD. Then, there is a unique lower triangular matrix L such that the diagonal entries of L are positive real numbers and

$$\mathsf{A} = \mathsf{L}\mathsf{L}^{\mathsf{H}}.$$

In other words, the Cholesky factorization is unique.

Proof.

Suppose that there are two lower triangular matrices L_1 and L_2 with positive real diagonal entries and

$$\mathsf{L}_1\mathsf{L}_1^{\mathsf{H}} = \mathsf{A} = \mathsf{L}_2\mathsf{L}_2^{\mathsf{H}}.$$

Then

$$\mathsf{L}_2^{-1}\mathsf{L}_1 = \mathsf{L}_2^{\mathsf{H}}\mathsf{L}_1^{-\mathsf{H}},$$

and, by a theorem, $\mathsf{L}_2^{-1}\mathsf{L}_1$ is lower triangular and $\mathsf{L}_2^{\mathsf{H}}\mathsf{L}_1^{-\mathsf{H}}$ is upper triangular. Thus, there is a diagonal matrix D such that

$$\mathsf{L}_2^{-1}\mathsf{L}_1 = \mathsf{D} = \mathsf{L}_2^{\mathsf{H}}\mathsf{L}_1^{-\mathsf{H}}.$$

Proof, Cont.

Therefore,

$$L_1 = L_2 D, \tag{2}$$

and

$$D L_1^H = L_2^H,$$

or, equivalently,

$$L_2 = L_1 D^H. \tag{3}$$

Combining (2) and (3), we have

$$L_1 = L_2 D = L_1 D^H D.$$

Since L_1 is invertible, the cancellation property holds and $D^H D = I_n$. But, since L_1 and L_2 have positive diagonal entries, so must D have positive diagonal entries. It follows that $D = I_n$, which implies $L_1 = L_2$. □

Theorem (HPD and spectrum)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is Hermitian. Then* A *is Hermitian positive definite (HPD) iff* $\sigma(\mathsf{A}) \subset (0,\infty)$.

Proof.

We only prove one direction here, as the other has already been proven. Since $\mathsf{A} \in \mathbb{C}^{n\times n}_{\mathrm{Her}}$, there exists a unitary matrix U and a diagonal matrix $\mathsf{D} = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, where $\lambda_i \in \sigma(\mathsf{A})$, for $i = 1, \ldots, n$, such that $\mathsf{A} = \mathsf{U}^{\mathsf{H}}\mathsf{D}\mathsf{U}$. Let $\mathbf{x} \in \mathbb{C}^n_\star$. Set $\mathbf{y} = \mathsf{U}\mathbf{x}$, and note that $\mathbf{y} \neq \mathbf{0}$. Then

$$\mathbf{x}^{\mathsf{H}}\mathsf{A}\mathbf{x} = (\mathsf{U}\mathbf{x})^{\mathsf{H}}\mathsf{D}\mathsf{U}\mathbf{x} = \mathbf{y}^{\mathsf{H}}\mathsf{D}\mathbf{y} = \sum_{i=1}^{n} \lambda_i |y_i|^2 \geq \min_{1\leq i\leq n} \lambda_i \mathbf{y}^{\mathsf{H}}\mathbf{y} = \min_{1\leq i\leq n} \lambda_i \|\mathbf{y}\|_2^2 > 0.$$

This proves that A is HPD. $\square$

Gil Strang's Favorite Matrix

Theorem

Define $\mathsf{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ *via*

$$\mathsf{A} = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & -1 & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}.$$

Then A *is SPD. Let* $h = 1/n$. *The set*

$$S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n-1}\},$$

where the i–th component of $\mathbf{w}_k$ *is defined via*

$$[\mathbf{w}_k]_i = \sin(k\pi ih),$$

is an orthogonal set of eigenvectors of A.

Proof.

Note that for $k = 1, \ldots, n-1$,

$$\begin{aligned}
[A\mathbf{w}_k]_i &= -\sin(k\pi(i-1)h) + 2\sin(k\pi ih) - \sin(k\pi(i+1)h) \\
&= 2\sin(k\pi ih) - 2\cos(k\pi h)\sin(k\pi ih) \\
&= (2 - 2\cos(k\pi h))\sin(k\pi ih) \\
&= 2(1-\cos(k\pi h))[\mathbf{w}_k]_i.
\end{aligned}$$

Hence the distinct eigenvalues are $\lambda_k = 2 - 2\cos(k\pi h)$. To see that these are strictly positive for $k = 1, \ldots, n-1$, note that

$$1 > \cos(k\pi h) > -1$$

which implies

$$-2 < -2\cos(k\pi h) < 2$$

which implies

$$0 = 2 - 2 < 2 - 2\cos(k\pi h) < 2 + 2 = 4.$$

Since A is symmetric, the eigenvectors associated to distinct eigenvalues are orthogonal. A is SPD since its eigenvalues are strictly positive. □

HPD Matrices and Similarity Transformations

Proposition

Let $\mathsf{A} \in \mathbb{C}^{m\times m}$ be HPD and $\mathsf{X} \in \mathbb{C}^{m\times n}$ with $m \geq n$ have full rank. Then $\mathsf{X}^{\mathsf{H}}\mathsf{A}\mathsf{X} \in \mathbb{C}^{n\times n}$ is HPD.

Proof.

Notice that

$$(\mathsf{X}^{\mathsf{H}}\mathsf{A}\mathsf{X})^{\mathsf{H}} = \mathsf{X}^{\mathsf{H}}\mathsf{A}^{\mathsf{H}}\mathsf{X} = \mathsf{X}^{\mathsf{H}}\mathsf{A}\mathsf{X}.$$

Suppose $\mathbf{x} \in \mathbb{C}^n_\star$ is arbitrary. Since X is full rank, it follows that $\mathbf{y} = \mathsf{X}\mathbf{x} \neq \mathbf{0}$, i.e., $\mathbf{y} \in \mathbb{C}^m_\star$. Then,

$$\mathbf{x}^{\mathsf{H}}\mathsf{X}^{\mathsf{H}}\mathsf{A}\mathsf{X}\mathbf{x} = \mathbf{y}^{\mathsf{H}}\mathsf{A}\mathbf{y} > 0,$$

since A is HPD. $\square$

Theorem (HPD and Gaussian elimination)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ *be Hermitian positive definite (HPD), and represented as*

$$\mathsf{A} = \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{p} & \hat{\mathsf{A}} \end{bmatrix},$$

where $\alpha \in \mathbb{C}$, $\mathbf{p} \in \mathbb{C}^{n-1}$, *and* $\hat{\mathsf{A}} \in \mathbb{C}^{(n-1)\times(n-1)}$. *After one step of Gaussian elimination (without pivoting),* A *will be reduced to the matrix*

$$\begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{0} & \mathsf{B} \end{bmatrix},$$

where $\mathsf{B} \in \mathbb{C}^{(n-1)\times(n-1)}$. *Then* B *is HPD, and the corresponding diagonal elements of* B *are smaller than those of* $\hat{\mathsf{A}}$.

Proof.

Let us construct a matrix $\mathsf{L} \in \mathbb{C}^{n\times n}$ such that, if possible,

$$\mathsf{L}\mathsf{A} = \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{0} & \mathsf{B} \end{bmatrix}.$$

Proof, Cont.

Consider

$$\mathsf{L} = \begin{bmatrix} 1 & \mathbf{0}^\intercal \\ \mathbf{m} & \mathsf{I}_{n-1} \end{bmatrix}.$$

Then

$$\mathsf{L}\mathsf{A} = \begin{bmatrix} 1 & \mathbf{0}^\intercal \\ \mathbf{m} & \mathsf{I}_{n-1} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{p} & \hat{\mathsf{A}} \end{bmatrix} = \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \alpha\mathbf{m} + \mathbf{p} & \mathbf{m}\mathbf{p}^{\mathsf{H}} + \hat{\mathsf{A}} \end{bmatrix}.$$

Note that since A is HPD, $\alpha > 0$ and $\hat{\mathsf{A}}$ is HPD. Choosing $\mathbf{m} = -\alpha^{-1}\mathbf{p}$, we have

$$\mathsf{L}\mathsf{A} = \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{0} & \hat{\mathsf{A}} - \alpha^{-1}\mathbf{p}\mathbf{p}^{\mathsf{H}} \end{bmatrix}.$$

Having successfully constructed L, we find $\mathsf{B} = \hat{\mathsf{A}} - \alpha^{-1}\mathbf{p}\mathbf{p}^{\mathsf{H}}$. Notice that this is not the only way to find the matrix B.

Proof, Cont.

Now let $\mathbf{x} \in \mathbb{C}_\star^{n-1}$ be arbitrary. Define $\mathbf{y} \in \mathbb{C}_\star^n$ via

$$\mathbf{y} = \begin{bmatrix} \gamma \\ \mathbf{x} \end{bmatrix},$$

where $\gamma \in \mathbb{C}$ is arbitrary. Then, since A is HPD,

$$\begin{aligned} 0 &< \mathbf{y}^{\mathsf{H}} \mathsf{A} \mathbf{y} \\ &= \begin{bmatrix} \overline{\gamma} & \mathbf{x}^{\mathsf{H}} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{p}^{\mathsf{H}} \\ \mathbf{p} & \hat{\mathsf{A}} \end{bmatrix} \begin{bmatrix} \gamma \\ \mathbf{x} \end{bmatrix} \\ &= \begin{bmatrix} \overline{\gamma} & \mathbf{x}^{\mathsf{H}} \end{bmatrix} \begin{bmatrix} \alpha\gamma + \mathbf{p}^{\mathsf{H}}\mathbf{x} \\ \gamma\mathbf{p} + \hat{\mathsf{A}}\mathbf{x} \end{bmatrix} \\ &= \alpha|\gamma|^2 + \overline{\gamma}\mathbf{p}^{\mathsf{H}}\mathbf{x} + \gamma\mathbf{x}^{\mathsf{H}}\mathbf{p} + \mathbf{x}^{\mathsf{H}}\hat{\mathsf{A}}\mathbf{x}. \end{aligned}$$

Proof, Cont.

Now we set $\gamma = -\alpha^{-1}\mathbf{p}^{\mathsf{H}}\mathbf{x}$. From the last calculation

$$
\begin{aligned}
0 &< \mathbf{y}^{\mathsf{H}}\mathsf{A}\mathbf{y} \\
&= \alpha^{-1}|\mathbf{p}^{\mathsf{H}}\mathbf{x}|^2 - \alpha^{-1}|\mathbf{p}^{\mathsf{H}}\mathbf{x}|^2 - \alpha^{-1}|\mathbf{p}^{\mathsf{H}}\mathbf{x}|^2 + \mathbf{x}^{\mathsf{H}}\hat{\mathsf{A}}\mathbf{x} \\
&= \mathbf{x}^{\mathsf{H}}\hat{\mathsf{A}}\mathbf{x} - \alpha^{-1}|\mathbf{p}^{\mathsf{H}}\mathbf{x}|^2 \\
&= \mathbf{x}^{\mathsf{H}}\mathsf{B}\mathbf{x}.
\end{aligned}
$$

This proves that B is HPD.

Now, the diagonal elements of B, which must be positive since B is HPD, are precisely $[\mathsf{B}]_{i,i} = [\hat{\mathsf{A}}]_{i,i} - \alpha^{-1}|[\mathbf{p}]_i|^2$. Hence $0 < [\mathsf{B}]_{i,i} \le [\hat{\mathsf{A}}]_{ii}$, since $\alpha^{-1}|[\mathbf{p}]_i|^2 \ge 0$. □

HPD Matrices and Gaussian Elimination

Corollary

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is HPD. Then Gaussian elimination without pivoting proceeds to completion to produce a unit lower triangular matrix* $\mathsf{L} \in \mathbb{C}^{n\times n}$ *and an upper triangular matrix* $\mathsf{U} \in \mathbb{C}^{n\times n}$ *with positive diagonal elements such that* $\mathsf{A} = \mathsf{LU}$*.*

Proof.

Apply recursively the previous result. □

HPD Criterion

Theorem

$A \in \mathbb{C}^{n\times n}$ *is Hermitian positive definite (HPD) if and only if* $A = LL^H$, *where* $L \in \mathbb{C}^{n\times n}$ *is invertible.*

Proof.

Suppose that $A = LL^H$, where L is invertible. Let $\mathbf{x} \in \mathbb{C}^n$ be arbitrary, and set $\mathbf{y} = L^H\mathbf{x}$. Since L is invertible, L^H is invertible; and $\mathbf{y} = \mathbf{0}$ if and only if $\mathbf{x} = \mathbf{0}$. Then

$$\mathbf{x}^H A\mathbf{x} = \mathbf{x}^H LL^H\mathbf{x} = \left(L^H\mathbf{x}\right)^H L^H\mathbf{x} = \mathbf{y}^H\mathbf{y} = \|\mathbf{y}\|_2^2 \geq 0,$$

with equality if and only if $\mathbf{x} = \mathbf{0}$. This proves that A is HPD.

The converse direction follows from the Cholesky factorization proved earlier. □

HPD Block Matrices

Theorem

Let $k, m \in \mathbb{N}$, and set $n = k + m$. Suppose $\mathsf{A} \in \mathbb{C}^{n\times n}$ has the decomposition

$$\mathsf{A} = \begin{bmatrix} \mathsf{B} & \mathsf{C}^{\mathsf{H}} \\ \mathsf{C} & \mathsf{D} \end{bmatrix},$$

where $\mathsf{B} \in \mathbb{C}^{k\times k}$, $\mathsf{C} \in \mathbb{C}^{m\times k}$, and $\mathsf{D} \in \mathbb{C}^{m\times m}$.

1. *If A is HPD, then B, D and $\mathsf{S} = \mathsf{D} - \mathsf{C}\mathsf{B}^{-1}\mathsf{C}^{\mathsf{H}}$ are HPD. S is called the Schur complement of B in A.*
2. *If A is HPD, the Cholesky factorization of A may be expressed in terms of the matrix C and the Cholesky factorizations of B and S.*

Proof.

1. Since A is HPD, $\mathbf{x}^{\mathsf{H}} A \mathbf{x} > 0$ for any $\mathbf{x} \in \mathbb{C}^n_\star$. Let $\mathbf{y} \in \mathbb{C}^k_\star$ and $\mathbf{w} \in \mathbb{C}^m_\star$ be arbitrary. Setting $\mathbf{x} = \left[\mathbf{y}^{\mathsf{H}}, \mathbf{0}^{\mathsf{H}}\right]^{\mathsf{H}} \in \mathbb{C}^n_\star$, we have

$$0 < \mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{y}^{\mathsf{H}} B \mathbf{y},$$

and, on the other hand, setting $\mathbf{x} = \left[\mathbf{0}^{\mathsf{H}}, \mathbf{w}^{\mathsf{H}}\right]^{\mathsf{H}} \in \mathbb{C}^n_\star$, we have

$$0 < \mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{w}^{\mathsf{H}} D \mathbf{w}.$$

Thus B and D are HPD. More generally, suppose $\mathbf{x} = \left[\mathbf{x}_1^{\mathsf{H}}, \mathbf{x}_2^{\mathsf{H}}\right]^{\mathsf{H}} \in \mathbb{C}^n_\star$. Then,

$$0 < \mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{x}_1^{\mathsf{H}} B \mathbf{x}_1 + \mathbf{x}_1^{\mathsf{H}} C^{\mathsf{H}} \mathbf{x}_2 + \mathbf{x}_2^{\mathsf{H}} C \mathbf{x}_1 + \mathbf{x}_2^{\mathsf{H}} D \mathbf{x}_2.$$

Proof, Cont.

Now, pick $\mathbf{x}_1 = -B^{-1}C^H\mathbf{x}_2$ and suppose $\mathbf{x}_2 \neq \mathbf{0}$. Then

$$\begin{aligned} 0 &< \mathbf{x}^H A\mathbf{x} \\ &= \mathbf{x}_2^H CB^{-H}BB^{-1}C^H\mathbf{x}_2 - \mathbf{x}_2^H CB^{-H}C^H\mathbf{x}_2 - \mathbf{x}_2^H CB^{-1}C^H\mathbf{x}_2 + \mathbf{x}_2^H D\mathbf{x}_2 \\ &= \mathbf{x}_2^H \left(D - CB^{-1}C^H\right)\mathbf{x}_2, \end{aligned}$$

where we used the fact that $B^{-1} = B^{-H}$ — since B is HPD — on the last step. It follows that S is HPD.

2. Since A is HPD there is a unique lower triangular matrix $L_A \in \mathbb{C}^{n\times n}$, with positive diagonal entries, such that $A = L_A L_A^H$. Likewise there are unique lower triangular matrices $L_B \in \mathbb{C}^{k\times k}$ and $L_S \in \mathbb{C}^{m\times m}$, both with positive diagonal entries, such that $B = L_B L_B^H$ and $S = L_S L_S^H$. Suppose

$$L_A = \begin{bmatrix} L_1 & O \\ M & L_2 \end{bmatrix},$$

where $L_1 \in \mathbb{C}^{k\times k}$ is lower triangular with positive diagonal entries; $M \in \mathbb{C}^{m\times k}$; and $L_2 \in \mathbb{C}^{m\times m}$ is lower triangular with positive diagonal entries.

Proof, Cont.

Then

$$L_A L_A^H = \begin{bmatrix} L_1 & O \\ M & L_2 \end{bmatrix} \begin{bmatrix} L_1^H & M^H \\ O^H & L_2^H \end{bmatrix} = \begin{bmatrix} L_1 L_1^H & L_1 M^H \\ M L_1^H & M M^H + L_2 L_2^H \end{bmatrix} = \begin{bmatrix} B & C^H \\ C & D \end{bmatrix}.$$

Comparing entries we conclude that

$$\begin{aligned} L_1 L_1^H = B \quad &\Longrightarrow \quad L_1 = L_B, \\ M L_1^H = C \quad &\Longrightarrow \quad M = C L_B^{-H}, \\ M M^H + L_2 L_2^H = D \quad &\Longrightarrow \quad L_2 L_2^H = D - C L_B^{-H} L_B^{-1} C^H \\ &\Longrightarrow \quad L_2 L_2^H = S \\ &\Longrightarrow \quad L_2 = L_S, \end{aligned}$$

and the result is proven. □