ENGG 5781: Matrix Analysis and Computations	2022-23 First Term

Lecture 7: Linear Systems

Instructor: Wing-Kin Ma

This note shows the proof of the properties and theorems in the main lecture slides.

1 Proof of Properties 7.1–7.3

Recall Properties 7.1–7.3 in the main lecture slides:

Property 7.1 *Let* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ *be lower triangular. Then,* $\mathbf{AB}$ *is lower triangular. Also, if* $\mathbf{A}$, $\mathbf{B}$ *have unit diagonal entries, then* $\mathbf{AB}$ *has unit diagonal entries.*

Property 7.2 *If* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *is lower triangular, then* $\det(\mathbf{A}) = \prod_{i=1}^n a_{ii}$.

Property 7.3 *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be nonsingular lower triangular. Then,* $\mathbf{A}^{-1}$ *is lower triangular with* $[\mathbf{A}^{-1}]_{ii} = 1/a_{ii}$.

Their proofs are shown as follows.

1.1 Proof of Property 7.1

Property 7.1 can be shown by examining the matrix product $\mathbf{AB}$ in an element-by-element fashion. I also show you an alternative proof using unit vector representations. For convenience, let $\mathbf{C} = \mathbf{A}^T$, and $\mathbf{D} = \mathbf{AB} = \mathbf{C}^T\mathbf{B}$. The (k, l)th entry of $\mathbf{D}$ is

$$d_{kl} = \mathbf{c}_k^T \mathbf{b}_l.$$

Since $\mathbf{B}$ is lower triangular, its columns can be represented by

$$\mathbf{b}_l = \sum_{j=l}^n b_{jl}\mathbf{e}_j, \quad l = 1, \dots, n,$$

where we recall that $\mathbf{e}_k$'s are unit vectors. Also, since $\mathbf{C} = \mathbf{A}^T$ is upper triangular, we can employ a similar representation

$$\mathbf{c}_k = \sum_{i=1}^k a_{ki}\mathbf{e}_i, \quad i = 1, \dots, n.$$

Using the above representations, d_{kl} can be expressed as

$$\begin{aligned} d_{kl} &= \left(\sum_{i=1}^k a_{ki}\mathbf{e}_i\right)^T \left(\sum_{j=l}^n b_{jl}\mathbf{e}_j\right) \\ &= \sum_{i=1}^k \sum_{j=l}^n a_{ki} b_{jl} \mathbf{e}_i^T \mathbf{e}_j \end{aligned}$$

By noting that $\mathbf{e}_i^T\mathbf{e}_j = 0$ for all $i \neq j$, and $\mathbf{e}_i^T\mathbf{e}_i = 1$, the above expression can be simplified to

$$d_{kl} = \begin{cases} 0, & k < l \\ \sum_{i=k}^{l} a_{ki}b_{il}, & k \geq l \end{cases}$$

It follows that $\mathbf{D}$ is lower triangular. The above formula also indicates that if $a_{kk} = b_{kk} = 1$ for all $1 \leq k \leq n$, then $d_{kk} = a_{kk}b_{kk} = 1$ for all $1 \leq k \leq n$.

1.2 Proof of Property 7.2

Recall the cofactor expansion formula for the determinant of a general $\mathbf{A} \in \mathbb{R}^{n\times n}$:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij}c_{ij}, \qquad c_{ij} = (-1)^{i+j}\det(\mathbf{A}_{ij}),$$

for any $i = 1, \ldots, n$, where $\mathbf{A}_{ij}$ is a submatrix obtained by deleting the ith row and jth column of $\mathbf{A}$. Now, consider a lower triangular $\mathbf{A}$. Let us choose $i = 1$ for the above cofactor expansion formula

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{1j}c_{1j} = a_{11}\det(\mathbf{A}_{11}).$$

By repeatedly applying the same cofactor expansion on the cofactors, we obtain $\det(\mathbf{A}) = a_{11}a_{22}\cdots a_{nn}$.

1.3 Proof of Property 7.3

Consider the following system

$$\mathbf{A}\mathbf{x} = \mathbf{e}_k,$$

where $1 \leq k \leq n$, and $\mathbf{A}$ is lower triangular. Let us examine the first k equations of the system:

$$a_{11}x_1 = 0, \tag{1a}$$
$$a_{21}x_1 + a_{22}x_2 = 0, \tag{1b}$$
$$\vdots \tag{1c}$$
$$a_{k-1,1}x_1 + \ldots + a_{k-1,k-1}x_{k-1} = 0, \tag{1d}$$
$$a_{k,1}x_1 + \ldots + a_{kk}x_k = 1. \tag{1e}$$

By applying forward substitution w.r.t. (1a)–(1e), we obtain

$$x_1 = \ldots = x_{k-1} = 0, \quad x_k = \frac{1}{a_{kk}}.$$

Here, we make an assumption that $a_{kk} \neq 0$. This assumption is satisfied if $\mathbf{A}$ is nonsingular; cf. Property 7.2

Now, we show that the inverse of a lower triangular $\mathbf{A}$ is also lower triangular. Let $\mathbf{B}$ be the inverse of $\mathbf{A}$. The identity $\mathbf{AB} = \mathbf{I}$ can be decomposed into n linear systems:

$$\mathbf{A}\mathbf{b}_k = \mathbf{e}_k, \qquad k = 1, \ldots, n.$$

Using the previously proven result, the solution $\mathbf{b}_k$ has $[\mathbf{b}_k]_l = 0$ for $l = 1, \ldots, k-1$. Consequently, $\mathbf{B}$ takes a lower triangular structure. In addition, we have $[\mathbf{b}_k]_k = 1/a_{kk}$.

2 Proof of Theorem 7.1

Let us recapitulate the theorem:

Theorem 7.1 *A matrix* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *has an LU decomposition if every principal submatrix* $\mathbf{A}_{\{1,\ldots,k\}}$ *satisfies*

$$\det(\mathbf{A}_{\{1,\ldots,k\}}) \neq 0,$$

for $k = 1, 2, \ldots, n-1$. *If the LU decomposition of* $\mathbf{A}$ *exists and* $\mathbf{A}$ *is nonsingular, then* $(\mathbf{L}, \mathbf{U})$ *is unique.*

From the development of Gauss elimination shown in the main slides, we see that the LU decomposition of a given $\mathbf{A}$ exists (or can be constructed) if the pivots $a_{kk}^{(k-1)}$'s are all nonzero. In the following, we show that if every principal submatrix $\mathbf{A}_{\{1,\ldots,k\}}$, $1 \leq k \leq n-1$, is nonsingular, then $a_{kk}^{(k-1)}$ is nonzero. Consider the matrix equation

$$\mathbf{A}^{(k-1)} = \mathbf{M}_{k-1}\cdots\mathbf{M}_2\mathbf{M}_1\mathbf{A}$$

for any $1 \leq k \leq n-1$. For convenience, let $\mathbf{W} = \mathbf{M}_{k-1}\cdots\mathbf{M}_2\mathbf{M}_1$. By Properties 7.1 and 7.3, $\mathbf{W}$ is lower triangular with unit diagonal elements. By denoting $\mathbf{A}_{i:j,k:l}$ be a submatrix of $\mathbf{A}$ obtained by keeping $i, i+1, \ldots, j$ rows and $k, k+1, \ldots, l$ columns of $\mathbf{A}$, we can expand $\mathbf{A}^{(k-1)} = \mathbf{W}\mathbf{A}$ as

$$\begin{bmatrix} \mathbf{A}_{1:k,1:k}^{(k-1)} & \mathbf{A}_{1:k,k+1:n}^{(k-1)} \\ \mathbf{A}_{k+1:n,1:k}^{(k-1)} & \mathbf{A}_{k+1:n,k+1:n}^{(k-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{1:k,1:k} & \mathbf{0} \\ \mathbf{W}_{k+1:n,1:k} & \mathbf{W}_{k+1:n,k+1:n} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{1:k,1:k} & \mathbf{A}_{1:k,k+1:n} \\ \mathbf{A}_{k+1:n,1:k} & \mathbf{A}_{k+1:n,k+1:n} \end{bmatrix}$$

From the above equation, we see that

$$\mathbf{A}_{1:k,1:k}^{(k-1)} = \mathbf{W}_{1:k,1:k}\mathbf{A}_{1:k,1:k}.$$

Consequently, we have

$$\det(\mathbf{A}_{1:k,1:k}^{(k-1)}) = \det(\mathbf{W}_{1:k,1:k})\det(\mathbf{A}_{1:k,1:k}).$$

Note that $\mathbf{A}_{1:k,1:k}^{(k-1)}$ is upper triangular, and $\mathbf{W}_{1:k,1:k}$ is lower triangular with unit diagonal elements. Thus, by Property 7.2, their determinants are

$$\det(\mathbf{A}_{1:k,1:k}^{(k-1)}) = \prod_{i=1}^{k} a_{ii}^{(k-1)}, \qquad \det(\mathbf{W}_{1:k,1:k}) = 1,$$

respectively. It follows that if $\mathbf{A}_{1:k,1:k}$ is nonsingular, then $a_{kk}^{(k-1)} \neq 0$.

We are also interested in proving the uniqueness of the LU decomposition. Suppose that $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1$ and $\mathbf{A} = \mathbf{L}_2\mathbf{U}_2$ are two LU decompositions of $\mathbf{A}$. Also, assume that $\mathbf{A}$ is nonsingular. Then, we claim that $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{U}_1$ and $\mathbf{U}_2$ are all nonsingular, and thus invertible: the nonsingularity of $\mathbf{L}_1$ and $\mathbf{L}_2$ follows from Property 7.2, as well as the fact that $\mathbf{L}_1$ and $\mathbf{L}_2$ are lower triangular with unit diagonal elements; the nonsingularity of $\mathbf{U}_1$ and $\mathbf{U}_2$ can be deduced from the nonsingularity of $\mathbf{L}_1$ and $\mathbf{L}_2$ and the nonsingularity of $\mathbf{A}$ (how?). From $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$, we can write

$$\mathbf{L}_2^{-1}\mathbf{L}_1 = \mathbf{U}_2\mathbf{U}_1^{-1}. \tag{2}$$

Note that the left-hand side of the above equation is a lower triangular matrix, while the right-hand side a upper triangular matrix (see Properties 7.1 and 7.3). Hence, (2) can only be satisfied when $\mathbf{L}_2^{-1}\mathbf{L}_1$ and $\mathbf{U}_2\mathbf{U}_1^{-1}$ are diagonal. Also, since $\mathbf{L}_2^{-1}\mathbf{L}_1$ has unit diagonal elements, we further obtain $\mathbf{L}^{-2}\mathbf{L}_1 = \mathbf{I}$, and consequently, $\mathbf{L}_1 = \mathbf{L}_2$. Moreover, from the above result, we also have $\mathbf{U}^{-2}\mathbf{U}_1 = \mathbf{I}$, and then $\mathbf{U}_1 = \mathbf{U}_2$. Thus, we have shown that the LU decomposition of a nonsingular $\mathbf{A}$, if exists, has to be unique.

3 Proof of Theorem 7.2

Theorem 7.2 *If $\mathbf{A} = \mathbf{LDM}^T$ is the LDM decomposition of a nonsingular symmetric $\mathbf{A}$, then $\mathbf{L} = \mathbf{M}$.*

Let $\mathbf{A} = \mathbf{LDM}^T$ be the LDM decomposition of $\mathbf{A}$, and consider

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M}^{-T} = \mathbf{M}^{-1}\mathbf{LD};$$

(note that any lower triangular $\mathbf{M}$ (or $\mathbf{L}$) with unit diagonal elements is invertible, as we have discussed in the proof of Theorem 7.2). Since $\mathbf{A}$ is symmetric, $\mathbf{M}^{-1}\mathbf{A}\mathbf{M}^{-T}$ is also symmetric. It follows that $\mathbf{M}^{-1}\mathbf{LD}$ is symmetric. By noting that $\mathbf{M}^{-1}\mathbf{L}$ is lower triangular with unit diagonal elements, the only possibility for $\mathbf{M}^{-1}\mathbf{LD}$ to be symmetric is that $\mathbf{M}^{-1}\mathbf{LD}$ is diagonal. Also, if $\mathbf{A}$ is nonsingular, then it can be verified from $\mathbf{A} = \mathbf{LDM}^T$ that the diagonal matrix $\mathbf{D}$ is nonsingular. As a result, $\mathbf{M}^{-1}\mathbf{L}$ must be diagonal. Since $\mathbf{M}^{-1}\mathbf{L}$ has unit diagonal elements, we further conclude that $\mathbf{M}^{-1}\mathbf{L} = \mathbf{I}$, or equivalently, $\mathbf{L} = \mathbf{M}$.

4 Proof of Theorem 7.3

Theorem 7.3 *If $\mathbf{A} \in \mathbb{S}^n$ is PD, then there exists a unique lower triangular $\mathbf{G} \in \mathbb{R}^{n\times n}$ with positive diagonal elements such that $\mathbf{A} = \mathbf{GG}^T$.*

If $\mathbf{A}$ is PD, then any principal submatrix of $\mathbf{A}$ is PD—and nonsingular; see Lecture 4, page 15. Hence, by Theorem 7.1, the LU or LDM decomposition of a PD $\mathbf{A}$ always exists in a unique sense. Also, by Theorem 7.2, the LDM decomposition can be simplified to the LDL decomposition $\mathbf{A} = \mathbf{LDL}^T$, where $\mathbf{L}, \mathbf{D}$ is unique. It can be verified that for a PD $\mathbf{A}$, we have $d_i > 0$ for all i (I leave this as an exercise for you). By constructing $\mathbf{G} = \mathbf{LD}^{\frac{1}{2}}$, we get $\mathbf{A} = \mathbf{GG}^T$.