SI231b: Matrix Computations

Lecture 5: Solving Linear Equations (Squared Systems)

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Sept. 21, 2022

Solving Squared Linear System

- Forward substitution, backward substitution
- Row-oriented implementation
- LU factorization
- Existence and uniqueness of LU factorization

Consider the system of linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$ are given, and $\mathbf{x} \in \mathbb{R}^n$ is the solution to the system.

- $\mathbf{A}$ will be assumed to be nonsingular (unless specified)
- we consider the real case for convenience; extension to the complex case is simple

Problem: compute the solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ in a numerically efficient manner.

- the problem is easy if $\mathbf{A}^{-1}$ is known
  - but computing $\mathbf{A}^{-1}$ also costs computations...
  - do you know how to compute $\mathbf{A}^{-1}$ efficiently?
- $\mathbf{A}$ is assumed to be a general nonsingular matrix.
  - the problem may become easy in some special cases, e.g., orthogonal $\mathbf{A}$, or $\mathbf{A}$ is triangular.

Consider the following 2-by-2 triangular system

$$\begin{bmatrix} \ell_{11} & \\ \ell_{21} & \ell_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

If $\ell_{11}\ell_{22} \neq 0$, then the unknowns can be determined sequentially

$$\begin{aligned} x_1 &= b_1/\ell_{11}, \\ x_2 &= (b_2 - \ell_{21}x_1)/\ell_{22}. \end{aligned}$$

The general procedure of solving $\mathbf{Lx} = \mathbf{b}$

$$\begin{aligned} x_1 &= b_1/\ell_{11}, \\ x_i &= \left(b_i - \sum_{j=1}^{i-1} \ell_{ij}x_j \right) /\ell_{ii}, \quad i = 2, \ldots, n \end{aligned}$$

Consider the following 2-by-2 triangular system

$$\begin{bmatrix} u_{11} & u_{12} \\ & u_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

If $u_{11}u_{22} \neq 0$, then the unknowns can be determined sequentially

$$\begin{aligned} x_2 &= b_2/u_{22}, \\ x_1 &= (b_1 - u_{12}x_2)/u_{11}. \end{aligned}$$

The general procedure of solving $\mathbf{U}\mathbf{x} = \mathbf{b}$

$$\begin{aligned} x_n &= b_n/u_{nn}, \\ x_i &= \left(b_i - \sum_{j=i+1}^{n} u_{ij}x_j \right) /u_{ii}, \quad i = 1, \ldots, n-1 \end{aligned}$$

Forward substitution:

```
x(1)= b(1)/L(1,1);
for i = 2:n
     x(i)= (b(i) - L(i, 1:i-1)*x(1:i-1))/L(i,i);
end
```

Backward substitution:

```
x(n)= b(n)/U(n,n);
for i = n-1:-1:1
     x(i)= (b(i) - U(i, i+1:n)*x(i+1:n))/U(i,i);
end
```

Example

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \\ -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$$

We all know the Gaussian elimination from the linear algebra course,

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & -2 \\ 2 & -1 & 1 & 5 \\ -1 & 2 & 2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & -1 & -2 \\ 0 & 1 & -1 & -3 \\ 0 & 3 & 1 & -1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & -1 & -2 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

that gives $x = 1$, $y = -1$, $z = 2$.

Question: how to compute the solution while the right-hand side is changed to $[7 \quad 5 \quad 9]^T$?

LU Factorization given $\mathbf{A} \in \mathbb{R}^{n \times n}$, find two matrices $\mathbf{L}, \mathbf{U} \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{A} = \mathbf{LU},$$

where

- $\mathbf{L} \in \mathbb{R}^{n \times n}$ is lower triangular with unit diagonal elements (i.e., $\ell_{ii} = 1$),
- $\mathbf{U} \in \mathbb{R}^{n \times n}$ is upper triangular.

Suppose that $\mathbf{A}$ has an LU factorization. Then, solving $\mathbf{Ax} = \mathbf{b}$ can be made easy:

1. solve $\mathbf{Lz} = \mathbf{b}$ for $\mathbf{z}$,
2. solve $\mathbf{Ux} = \mathbf{z}$ for $\mathbf{x}$.

Question:

1. Does LU factorization exist?
2. How to perform $\mathbf{A} = \mathbf{LU}$?

Observation: given $\mathbf{x} \in \mathbb{R}^n$ with $x_k \neq 0$, $1 \leq k \leq n$,

$$\underbrace{\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\frac{x_{k+1}}{x_k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\frac{x_n}{x_k} & & & 1 \end{bmatrix}}_{\mathbf{M}_k} \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Outer-product form of $\mathbf{M}_k$:

$$\mathbf{M}_k = \mathbf{I} - \boldsymbol{\tau}^{(k)} \mathbf{e}_k^T, \qquad \boldsymbol{\tau}^{(k)} = [\ 0, \ldots, 0, x_{k+1}/x_k, \ldots, x_n/x_k\ ]^T.$$

Problem: find Gauss transformations $\mathbf{M}_1, \ldots, \mathbf{M}_{n-1} \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A} = \mathbf{U}, \quad \mathbf{U} \text{ being upper triangular.}$$

Step 1: choose $\mathbf{M}_1$ such that $\mathbf{M}_1 \mathbf{a}_1 = [\ a_{11}, 0, \ldots, 0\ ]^T$

- **if** $a_{11} \neq 0$, then we can choose

$$\mathbf{M}_1 = \mathbf{I} - \boldsymbol{\tau}^{(1)} \boldsymbol{e}_1^T, \qquad \boldsymbol{\tau}^{(1)} = [\ 0, a_{21}/a_{11}, \ldots, a_{n1}/a_{11}\ ]^T.$$

- result:

$$\mathbf{M}_1 \mathbf{A} = \begin{bmatrix} a_{11} & \times & \ldots & \times \\ 0 & \times & \ldots & \times \\ \vdots & \vdots & & \vdots \\ 0 & \times & \ldots & \times \end{bmatrix}$$

Step 2: let $\mathbf{A}^{(1)} = \mathbf{M}_1\mathbf{A}$. Choose $\mathbf{M}_2$ such that

$$\mathbf{M}_2\mathbf{a}_2^{(1)} = [\ a_{12}^{(1)},\ a_{22}^{(1)},\ 0,\ldots,0\ ]^T.$$

- **if** $a_{22}^{(1)} \neq 0$, then we can choose

$$\mathbf{M}_2 = \mathbf{I} - \boldsymbol{\tau}^{(2)}\boldsymbol{e}_2^T, \qquad \boldsymbol{\tau}^{(2)} = [\ 0,\ 0,\ a_{32}^{(1)}/a_{22}^{(1)},\ \ldots,\ a_{n,2}^{(1)}/a_{22}^{(1)}\ ]^T.$$

- result:

$$\mathbf{M}_2\mathbf{A}^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \times & \ldots & \times \\ 0 & a_{22}^{(1)} & \times & \ldots & \times \\ \vdots & 0 & \times & & \times \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \times & \ldots & \times \end{bmatrix}$$

Let $\mathbf{A}^{(k)} = \mathbf{M}_k \mathbf{A}^{(k-1)}$, $\mathbf{A}^{(0)} = \mathbf{A}$. Note $\mathbf{A}^{(k)} = \mathbf{M}_k \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A}$.

Step k: Choose $\mathbf{M}_k$ such that

$$\mathbf{M}_k \mathbf{a}_k^{(k-1)} = [\ a_{1k}^{(k-1)}, \dots, a_{kk}^{(k-1)}, 0, \dots, 0\ ]^T.$$

- **if** $a_{kk}^{(k-1)} \neq 0$, then

$$\mathbf{M}_k = \mathbf{I} - \boldsymbol{\tau}^{(k)} \mathbf{e}_k^T, \quad \boldsymbol{\tau}^{(k)} = [\ 0, \dots, 0, a_{k+1,k}^{(k-1)}/a_{kk}^{(k-1)}, \dots, a_{n,k}^{(k-1)}/a_{kk}^{(k-1)}\ ]^T,$$

- result:

$$\mathbf{A}^{(k)} = \mathbf{M}_k \mathbf{A}^{(k-1)} = \begin{bmatrix} a_{11}^{(k-1)} & \cdots & a_{1k}^{(k-1)} & \times & \cdots & \times \\ 0 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & & a_{kk}^{(k-1)} & \vdots & & \vdots \\ \vdots & & 0 & \times & & \times \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \times & \cdots & \times \end{bmatrix}$$

  - $\mathbf{A}^{(n-1)} = \mathbf{U}$ is upper triangular
  - $a_{kk}^{(k-1)}$ is called the **pivot**

How to Obtain **L**

We have seen that under the assumption of the pivot $a_{kk}^{(k-1)} \neq 0$ for all k,

$$\mathbf{U} = \mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A} \text{ is upper triangular.}$$

But where is **L**?

Suppose that every $\mathbf{M}_k$ is invertible. Then,

$$\mathbf{L} = \mathbf{M}_1^{-1} \mathbf{M}_2^{-1} \cdots \mathbf{M}_{n-1}^{-1}$$

satisfies $\mathbf{A} = \mathbf{L}\mathbf{U}$.

Questions:

1. Is $\mathbf{M}_k$ invertible for all k?
2. Is **L** lower triangular with unit diagonal entries?

Computations of $\mathbf{L}$

Is $\mathbf{M}_k$ invertible?

Fact: $\mathbf{M}_k^{-1} = \mathbf{I} + \boldsymbol{\tau}^{(k)} \boldsymbol{e}_k^T$.

Hint: applying the Woodbury matrix identity,

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}.$$

Using the fact that $\boldsymbol{e}_i^T \boldsymbol{\tau}^{(k)} = 0$ for $k \geq i$, we obtain

$$\mathbf{L} = \mathbf{M}_1^{-1} \dots \mathbf{M}_{n-1}^{-1} = \mathbf{I} + \sum_{k=1}^{n-1} \boldsymbol{\tau}^{(k)} \boldsymbol{e}_k^T$$

You can easily verify that $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal entries.

$\boldsymbol{L}$ is lower triangular with unit diagonal entries can also be verified using the following properties.

- 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.

  How to prove?

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

  Hands-on exercise

Existence and Uniqueness of LU Factorization

Theorem

The matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is nonsingular and has an LU factorization if every leading principal submatrix $\mathbf{A}_{\{1,\ldots,k\}}$ satisfies

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

for $k = 1, 2, \ldots, n-1$.

- the proof is essentially about when $a_{kk}^{(k-1)} \neq 0$.

Theorem

If the LU factorization of $\mathbf{A}$ exists, then $(\mathbf{L}, \mathbf{U})$ is unique.