
COM521500

Math. Methods for SP I

Lecture 7: Solving Square Linear Systems

Motivation

To solve

$$\mathbf{Ax} = \mathbf{b}$$

for \mathbf{x} , we can explicitly invert \mathbf{A} . Such a direct method, however, is not the best way from a computational complexity viewpoint.

There exists computationally faster methods for finding \mathbf{x} , which do not need explicit computation of \mathbf{A} .

LU Decomposition

A nonsingular $n \times n$ matrix \mathbf{A} may be decomposed as

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

where \mathbf{L} is a lower triangular matrix with $\text{diag}(\mathbf{L}) = [1, 1, \dots, 1]^T$, and \mathbf{U} is an upper triangular matrix.

The above decomposition is called **the LU decomposition** of \mathbf{A} .

To solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ for \mathbf{x} is equivalent to

solve $\mathbf{L}\mathbf{z} = \mathbf{b}$ for \mathbf{z} (forward substitution)

solve $\mathbf{U}\mathbf{x} = \mathbf{z}$ for \mathbf{x} (backward substitution)

Example: backward substitution

The solution to

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

is

$$x_2 = z_2 / u_{22}$$

$$x_1 = (z_1 - u_{12}x_2) / u_{11}$$

Backward Substitution Algorithm:

```

for  $i = n, n - 1, \dots, 1$ 
     $x_i := z_i$ 
    for  $j = i + 1, \dots, n$ 
         $x_i := x_i - u_{ij}x_j$ 
    end
     $x_i := x_i / u_{ii}$ 
end

```

Likewise, we have

Forward Substitution Algorithm:

```

for  $i = 1, 2, \dots, n$ 
   $z_i := b_i$ 
  for  $j = 1, \dots, i - 1$ 
     $z_i := z_i - \ell_{ij}z_j$ 
  end
   $z_i := x_i / \ell_{ii}$ 
end

```

Finding L & U by Gauss Elimination

We can find a sequence of $n \times n$ matrices $\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_{n-1}$ so that

$$\mathbf{M}_{n-1} \dots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A} = \mathbf{U}$$

where \mathbf{U} is upper triangular.

Let

$$\mathbf{A}^{(k)} = \mathbf{M}_k \mathbf{A}^{(k-1)}$$

with $\mathbf{A}^{(0)} = \mathbf{A}$, and $\mathbf{A}^{(n-1)} = \mathbf{U}$.

Each matrix \mathbf{M}_k is chosen so that

$$\mathbf{A}^{(k)} = \begin{bmatrix} \mathbf{A}_{11}^{(k)} & \mathbf{A}_{12}^{(k)} \\ \mathbf{0}_{n-k, n-k} & \mathbf{A}_{22}^{(k)} \end{bmatrix}$$

where $\mathbf{A}_{11}^{(k)} \in \mathbb{C}^{k \times k}$ is upper triangular.

Such an \mathbf{M}_k is given by

$$\mathbf{M}_k = \mathbf{I} - \boldsymbol{\tau}^{(k)} \mathbf{e}_k^T$$

where $[\mathbf{e}_k]_i = 1$ for $i = k$, and $[\mathbf{e}_k]_i = 0$ otherwise;

$$\boldsymbol{\tau}^{(k)} = [0, \dots, 0, \tau_{k+1}^{(k)}, \dots, \tau_n^{(k)}]^T$$

$$\tau_i^{(k)} = a_{ik}^{(k-1)} / a_{kk}^{(k-1)}$$

The element $a_{kk}^{(k-1)}$ is called the **pivot** element.

It is required that $a_{kk}^{(k-1)} \neq 0$. (If not, then a process called pivoting is needed)

The matrix \mathbf{M}_k has a structure

$$\mathbf{M}_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\tau_{k+1}^{(k)} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\tau_n^{(k)} & & & 1 \end{bmatrix}$$

and hence is lower triangular.

If $\mathbf{M}_{n-1} \dots \mathbf{M}_1$ were invertible, then by letting

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

we obtain the LU decomposition

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

Some useful properties:

Property 7.1 If $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ are lower triangular with unit diagonal entries, then \mathbf{AB} is lower triangular with unit diagonal entries.

Property 7.2 If \mathbf{A} is nonsingular and lower triangular, then \mathbf{A}^{-1} is lower triangular. In addition, \mathbf{A}^{-1} has unit diagonal entries if \mathbf{A} has unit diagonal entries.

It follows from Property 7.1 that $\mathbf{M}_{n-1} \dots \mathbf{M}_1$ is lower triangular.

Moreover, we see from Property 7.2 that if $\mathbf{M}_{n-1} \dots \mathbf{M}_1$ is invertible, then $\mathbf{L} = (\mathbf{M}_{n-1} \dots \mathbf{M}_1)^{-1}$ is lower triangular.

Is $\mathbf{M}_{n-1} \dots \mathbf{M}_1$ invertible?

By noting that

$$\begin{aligned} (\mathbf{I} + \tau^{(k)} \mathbf{e}_k^T) \mathbf{M}_k &= (\mathbf{I} + \tau^{(k)} \mathbf{e}_k^T) (\mathbf{I} - \tau^{(k)} \mathbf{e}_k^T) \\ &= \mathbf{I} + \tau^{(k)} \mathbf{e}_k^T - \tau^{(k)} \mathbf{e}_k^T + \underbrace{\tau^{(k)} \mathbf{e}_k^T \tau^{(k)} \mathbf{e}_k^T}_{=0} \\ &= \mathbf{I} \end{aligned}$$

we have

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

thereby showing that \mathbf{M}_k 's are invertible.

Subsequently, $\mathbf{M}_{n-1} \dots \mathbf{M}_1$ is invertible.

It can be further shown that \mathbf{L} can be easily computed (without inverting $\mathbf{M}_{n-1} \dots \mathbf{M}_1$):

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

From the above equation, it is clear that $\text{diag}(\mathbf{L}) = [1, 1, \dots, 1]^T$.

Having studied the construction of LU factors, we consider the existence of the LU decomposition:

Theorem 7.1 A matrix \mathbf{A} has an LU decomposition if

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

for $k = 1, 2, \dots, n - 1$. If the LU decomposition exists and \mathbf{A} is nonsingular, then the decomposition is unique.

A consequence of LU decomposition is that

Property 7.3 $\det(\mathbf{A}) = \prod_{i=1}^n u_{ii}$

This provides us with a numerically fast method of computing the determinant.

The inverse can also be numerically computed by using the LU decomposition, since $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$ & inverting lower/upper triangular matrices are rather simple.

Some remarks:

1. For real-valued \mathbf{A} , the LU decomposition requires $O(2n^3/3)$ flops.
2. Gauss elimination is numerically unstable when a pivot element $a_{kk}^{(k-1)}$ is zero or close to zero. In that case, **pivoting** is required. Pivoting works by interchanging the rows of $\mathbf{A}^{(k)}$ to obtain better pivot elements.

LDM Factorization

For a nonsingular \mathbf{A} , we can decompose

$$\mathbf{A} = \mathbf{LDM}^H$$

where \mathbf{L} is lower triangular with $\text{diag}(\mathbf{L}) = [1, \dots, 1]^T$, & \mathbf{M} is lower triangular with $\text{diag}(\mathbf{M}) = [1, \dots, 1]^T$.

Apparently LDM factorization is a variant of LU, where $\mathbf{U} = \mathbf{DM}^H$.

LDL Factorization for Hermitian Matrices

Theorem 7.2 If $\mathbf{A} = \mathbf{LDM}^H$ is the LDM factorization of a nonsingular Hermitian \mathbf{A} , then $\mathbf{L} = \mathbf{M}$.

For real-valued \mathbf{A} , LDL factorization requires $O(n^3/3)$ flops instead of $O(2n^3/3)$.

Cholesky Factorization for PD Matrices

Theorem 7.3 If \mathbf{A} is PD, then there exists a unique lower triangular $n \times n$ \mathbf{G} with positive diagonal entries, such that

$$\mathbf{A} = \mathbf{G}\mathbf{G}^H$$