LECTURES 5 + 6

LINEAR SYSTEMS CONTINUED

$\underline{A}\underline{x} = \underline{b}$

Rewrite out $\underline{M}_1\underline{A}\underline{x} = \underline{b}$

$(\underline{M}_n \cdots \underline{M}_1\underline{A})\underline{x} = \underbrace{(\underline{M}_n \cdots \underline{M}_1)}_{\underline{L}^{-1}}\underline{b}$

LU-decomposition

EX

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 4 & 2 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 4 & 2 & 1 \end{pmatrix}} \Rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & -6 & -11 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix} = \underline{U}.$$

NOTE:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 4 & 2 & 1 \end{pmatrix}.$$

NOTE also $\underline{L}^{-1} = \underline{M}_3\underline{M}_2\underline{M}_1 \Rightarrow \underline{L} = \underline{M}_1^{-1}\underline{M}_2^{-1}\underline{M}_3^{-1}$

How do we find $\underline{M}_i^{-1}$ if $\underline{M}_i^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$\underline{M}_i = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Check:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \checkmark$$

EX (cont).

Return to example and pull out $\underline{M}_1 \ldots \underline{M}_n$.

Compute $\underline{L}$. Verify $\underline{L}\,\underline{U} = \underline{A}$.

The Doolittle Algorithm:

for $k = 1$ to $n-1$

  if $a_{kk} = 0$ then stop ← There is a snag here.

  for $i = k+1 : n$

    $m_{ik} = a_{ik} / a_{kk}$

    $a_{i:} = a_{i:} + m_{ik} a_{k:}$

  end

end

Q: How would one fix this?

A: Row permutation.

M_{ik} - stores the operations from which we can recover $\underline{\underline{M}}_i$'s.

EX [redacted] (Get them to do this).

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & -3 \\ 4 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -5 \\ 4 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -5 \\ 0 & -3 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -5 \\ 0 & 0 & -21/2 \end{pmatrix} = \underline{\underline{U}}$$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \quad M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{3}{2} & 1 \end{pmatrix}$$

$$\underline{\underline{L}}^{-1} = \underline{\underline{M}}_3 \underline{\underline{M}}_2 \underline{\underline{M}}_1 \Rightarrow \underline{\underline{L}} = \underline{\underline{M}}_1^{-1} \underline{\underline{M}}_2^{-1} \underline{\underline{M}}_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \end{pmatrix}$$

$$\underline{\underline{L}} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & \frac{3}{2} & 1 \end{pmatrix}$$

$$\underline{\underline{L}}\,\underline{\underline{U}} = \begin{pmatrix} 1 & 0 & 0 \\ +2 & -1 & 0 \\ 4 & \frac{3}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -5 \\ 0 & 0 & -21/2 \end{pmatrix}$$

$$\frac{3}{2}\cdot 5 - \frac{21}{2} + 4 = \frac{15-21}{2} = -3+4 = 1$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & -3 \\ 4 & 1 & 1 \end{pmatrix} \quad \text{Good!}$$

Partial Pivoting

- Notice things go completely wrong if we hit an $a_{kk}=0$.
- However, we also lose information in rounding errors if $a_{kk}\approx 0$. Overflow!
- To avoid this numerical difficulty we use partial pivoting.

Partial pivoting is the switching of rows to ensure that the largest value on or below the diagonal occupies the diagonal.

Return to $\underline{EX}$

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 4 & 2 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 2 & 1 \\ 0 & \frac{1}{2} & \frac{5}{4} \\ 0 & \frac{5}{2} & \frac{13}{4} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 4 & 2 & 1 \\ 0 & \frac{5}{2} & \frac{13}{4} \\ 0 & \frac{1}{2} & \frac{5}{4} \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 2 & 1 \\ 0 & \frac{5}{2} & \frac{13}{4} \\ 0 & 0 & \frac{7}{4} \end{pmatrix} = \underline{\underline{U}}.$$

Notice: we now need to modify the string of transforms to include $\underline{P}_i$'s - permutation matrices.

In this example.

$$\underline{\underline{U}} = \ldots\ldots\ldots\ldots \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \underline{\underline{A}}$$

a permutation matrix.

Cholesky Factorisation (a special case of LU).

If $\underline{\underline{A}}$ is symmetric and positive definite the LU factorisation can be written as

$$\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}} = \underline{\underline{L}}\,\underline{\underline{L}}^T$$

Symmetric $\underline{\underline{A}} = \underline{\underline{A}}^T$

Positive definite $\underline{x}^T \underline{\underline{A}}\, \underline{x} > 0 \quad \forall \underline{x} \neq 0.$

~~Useful Because:~~

~~CHECK:~~

$$\cancel{\underline{\underline{L}}\,\underline{\underline{L}}^T = \underline{\underline{A}}}$$

$$\underline{\underline{A}} = \begin{pmatrix} a_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} l_{11} & L_{21}^T \\ 0 & L_{22}^T \end{pmatrix}$$

$$= \begin{pmatrix} l_{11}^2 & l_{11} L_{21}^T \\ l_{11} L_{21} & \cancel{} \end{pmatrix}$$

$L_{21} L_{21}^T + L_{22} L_{22}^T$

1) find $l_{11} = \sqrt{a_{11}}$

2) find $L_{21} = A_{21} / l_{11}$

3) find L_{22} from $L_{22} L_{22}^T = A_{22} - L_{21} L_{21}^T$

EX

$$\begin{pmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{pmatrix}$$

1) $l_{11} = \sqrt{25} = 5$

2) $L_{21}^T = \begin{pmatrix} 15 & -5 \end{pmatrix} / 5 = \begin{pmatrix} 3 & -1 \end{pmatrix}$

3) $\begin{pmatrix} 18 & 0 \\ 0 & 11 \end{pmatrix}$ $\begin{pmatrix} 18 & 0 \\ 0 & 11 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} \begin{pmatrix} 3 & -1 \end{pmatrix}$

$$= \begin{pmatrix} 18 & 0 \\ 0 & 11 \end{pmatrix} - \begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 3 \\ 3 & 10 \end{pmatrix} = L_{22} L_{22}^T$$

$$\cancel{L_{22} L_{22}^T = \begin{pmatrix} l_{22} & l_{23} \\ l_{23} & l_{33} \end{pmatrix}}$$

$$L_{22}L_{22}^T = \begin{pmatrix} 9 & 3 \\ 3 & 10 \end{pmatrix} = \begin{pmatrix} l_{22} & 0 \\ l_{23} & l_{33} \end{pmatrix}\begin{pmatrix} l_{22} & l_{23} \\ 0 & l_{33} \end{pmatrix}$$

1) $l_{22} = \sqrt{9} = 3$

2) $l_{23} = 3/3 = 1$

3) $l_{33}^2 = 10 - l_{23}^2 = 9$

$\Rightarrow l_{33} = 3$

Thus:

$$\begin{pmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{pmatrix}\begin{pmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

~~Useful because~~? Only ~~half the~~ matrix ~~is needed~~

~~RECTANGULAR SYSTEMS~~

EX

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 1 & 0 \\ 1 & 2 & 3 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 2 & 13 & 23 & 8 \\ 4 & 23 & 42 & 17 \\ 1 & 8 & 17 & 15 \end{pmatrix} = LL^T$$

ITERATIVE METHODS

1) Gauss-Seidel.
2) Jacobi

Gauss-Seidel:

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

Write $\underline{\underline{A}} = \underline{\underline{L}} + \underline{\underline{U}}$ $\Rightarrow (\underline{\underline{L}} + \underline{\underline{U}})\underline{x} = \underline{b}$

$\Rightarrow \underline{\underline{L}}\,\underline{x} = \underline{b} - \underline{\underline{U}}\,\underline{x}$

strictly upper.

NOTE ADDITIVE

GS iteration: $\underline{x}_{k+1} = \underline{\underline{L}}^{-1}(\underline{b} - \underline{\underline{U}}\,\underline{x}_k)$.

However, in practice usually easier to solve

$$\underline{\underline{L}}\,\underline{x}_{k+1} = \underline{y} \quad \text{where } \underline{y} = \underline{b} - \underline{\underline{U}}\,\underline{x}_k$$

by back substitution.

Jacobi

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

Write $\underline{\underline{A}} = \underline{\underline{D}} + \underline{\underline{R}}$

diagonal part

remainder

$$\underline{\underline{D}} = \begin{pmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & & \ddots & a_{nn} \end{pmatrix}, \quad \underline{\underline{R}} = \begin{pmatrix} 0 & a_{12} & \cdots \\ a_{21} & \ddots & \\ \vdots & & 0 \end{pmatrix}.$$

$$\Rightarrow (\underline{\underline{D}} + \underline{\underline{R}})\underline{x} = \underline{b}$$

Jacobi iteration $\Rightarrow \underline{x} = \underline{\underline{D}}^{-1}(\underline{b} - \underline{x}\underline{\underline{R}})$

$\underline{\underline{D}} = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{nn} \end{pmatrix}$ $\underline{\underline{D}}^{-1} = \begin{pmatrix} 1/a_{11} & 0 \\ 0 & 1/a_{nn} \end{pmatrix}$ $\underline{\underline{D}}$ - easy to invert

~~[illegible]~~

Gauss-Seidel	vs	Jacobi
Can overwrite $\underline{x}_k$ with $\underline{x}_{k+1}$ as we go - less stuff to store.		Finding $\underline{x}_{k+1}$ requires storage of all elements in $\underline{x}_k$
Cannot be parallelised.		Can be parallelised

Only convergence guaranteed if diagonally dominant or symmetric and pos. def.

EX

$$10x + 3y = 15$$
$$2x + 15y = 32$$

$x = 1 \quad y = 2.$

$$\begin{pmatrix} 10 & 3 \\ 2 & 15 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 15 \\ 32 \end{pmatrix}$$

~~GS~~ $\cancel{\underline{L} = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} \Rightarrow \underline{L}^{-1} = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}}$

$\cancel{\underline{U} = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}}$

$\cancel{\underline{x}_{k+1} = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} \left(\begin{pmatrix} 5 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \underline{x}_k \right)}$

$\underline{x}_0 = 2$

RECTANGULAR SYSTEMS

- Either overdetermined or undetermined probs.
- E.g. data fitting.

(7)

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \qquad \underline{\underline{A}} \in \mathbb{R}^{m \times n}$$

$$\underline{x} \in \mathbb{R}^{n}$$

$$\underline{b} \in \mathbb{R}^{m}$$

Overdet – $m > n$

Normal equations:

If $\underline{\underline{A}}$ has full column rank then $\underline{\underline{A}}^T\underline{\underline{A}}$ is nonsingular so write:

$$\underline{\underline{A}}^T\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}^T\underline{b}$$

Then $\underline{\underline{A}}^T\underline{\underline{A}}$ is symmetric positive definite and $\underline{\underline{A}}^T\underline{\underline{A}} = \underline{\underline{L}}^T\underline{\underline{L}}$ by Cholesky

Some issues with this:

1) Loss of information. E.g.:

$$A = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix} \qquad \epsilon^2 << 1 \qquad \epsilon^2 < \epsilon_{mach}$$

singular

$$\underline{\underline{A}}^T\underline{\underline{A}} = \begin{bmatrix} 1+\epsilon^2 & 1 \\ 1 & 1+\epsilon^2 \end{bmatrix} \text{ in floating point } = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

2) Sensitivity is worsened

$$\mathrm{cond}(\underline{\underline{A}}^T\underline{\underline{A}}) = \mathrm{cond}(\underline{\underline{A}}^2)$$

These issues of bad conditioning can be avoided using ~~ORTHOGONALISATION METHODS~~. e.g.

- QR factorisation.
- Householder transformation.
- Gram-Schmidt orthogonalisation.

orthogonalisation methods