DMA 4007 | Sep 24, 2024 | Lecture 7

Outline

1 LU for solving $Ax=b$

2 Why?

- repeatedly
- tridiagonal, sparse, large
- finite difference

3. Permutation matrix

4. Transpose, symmetric

The last 2 lectures have looked at the inverse of a square matrix A. The inverse matrix A^{-1} satisfies

$$AA^{-1} = A^{-1}A = I$$

and provides the unique sol$^{\underline{n}}$ of $Ax=b$, $x = A^{-1}b$. Whether a square matrix has an inverse can be decided by inspection for diagonal matrices and for diagonally dominant matrices (matrices whose diagonal entries exceed the sum of the absolute values of off diagonal entries). In general, elimination + back substitution in the form of G-J is how to check that an inverse exists* (n pivots) & compute it.

$$[A|I] \xrightarrow[\text{eliminate below pivots}]{} [\overset{\text{upper}}{U}|?] \xrightarrow[\text{eliminate above pivots}]{} [\overset{\text{diagonal}}{D}|??] \xrightarrow[\text{divide by pivots}]{} [I|A^{-1}]$$

This closely related to the fact that the j-th column of A^{-1} is the sol$^{\underline{n}}$ of $Ax = \vec{e}_j$ j-th col. of I

The same procedure as in G-J can be used to solve $Ax=b$ for 2 rhs $[A|b_1\ b_2] \longrightarrow [I|x_1\ x_2]$ if b_1 & b_2 are known in advance.

* We can decide that an inverse does <u>not</u> exist if A is not square, has row/column of zeros etc.

There are applications (details about one later), where $Ax=b$ must be solved repeatedly w/o knowing the rhs in advance.

1. First solve $A\vec{x}_1 = \vec{b}_1$
2. Then solve $A\vec{x}_2 = \vec{b}_2$

Two options

1. Compute & save A^{-1}
2. Compute & save $A = LU$

Example. do both

$$A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} \boxed{-2} & 1 & 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} \boxed{-2} & 1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{-3/2} & 1 & 1/2 & 1 & 0 \\ 0 & 0 & \boxed{-4/3} & 1/3 & 2/3 & 1 \end{array}\right]$$

$$U = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -3/2 & 1 \\ 0 & 0 & -4/3 \end{bmatrix}$$, L undoes the elimination so $A = LU$

check

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 0 & -3/2 & 1 \\ 0 & 0 & -4/3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \checkmark$$

Continue & zero above pivots   start at bottom & go up.

$$\left[\begin{array}{ccc|ccc} \boxed{-2} & 1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{-3/2} & 1 & 1/2 & 1 & 0 \\ 0 & 0 & \boxed{-4/3} & 1/3 & 2/3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccccc} -2 & 0 & 0 & 3/2 & 1 & 1/2 \\ 0 & -3/2 & 0 & 3/4 & 3/2 & 3/4 \\ 0 & 0 & -4/3 & 1/3 & 2/3 & 1 \end{array}\right]$$

- add 3/4 row 3 to row 2
- add 2/3 row 2 to row 1

Divide by pivots

$$A^{-1} = \begin{bmatrix} -3/4 & -1/2 & -1/4 \\ -1/2 & -1 & -1/2 \\ -1/4 & -1/2 & -3/4 \end{bmatrix} = -\frac{1}{4}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

Knowing A^{-1}, we solve ~~find the sol^n~~ of $Ax=b$ by $x = A^{-1}b$.

Know $A = LU$, we can solve $Ax=b$ how?

$$Ax = b$$

$$L\underbrace{\boxed{Ux}}_{y} = b$$

1. Solve $\boxed{Ly = b \text{ for } y.}$
2. Solve $\boxed{Ux = y \text{ for } x.}$

} relatively easy b/c L & U are triangular

which means only back sub. required, no elimination.

Previous example

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3/2 & 1 \\ 0 & 0 & -4/3 \end{bmatrix}$$

take $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$Ax = b$$

$$L\underbrace{\boxed{Ux}}_{y} = b$$

1. Solve $Ly = b$

top down

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

$y_1 = 1$

$y_2 = 2 + 1/2 = 5/2$

$y_3 = 3 + \frac{2}{3}\, 5/2 = \frac{14}{3}$

2. Solve $Ux = y$

bottom up

$$\left[\begin{array}{ccc|c} 2 & 1 & 0 & 1 \\ 0 & -3/2 & 1 & 5/2 \\ 0 & 0 & -4/3 & 14/3 \end{array}\right]$$

$x_3 = -\frac{3}{4}\,\frac{14}{3} = -7/2$

$-\frac{3}{2}x_2 = 5/2 + \cancel{\frac{3}{4}}\; 7/2 = \frac{12}{2} = 6$

$x_2 = -\frac{2}{3}\, 6 = -4$

$2x_1 = 1 - 6 = -5$

$x_1 = -5/2$

1. It works.

2. Q: Better than storing A^{-1}?

3. Q: Why repeatedly solving?

4. Q: ~~B~~ Is this matrix famous?

An example / application: repeatedly solving

suppose $\vec{y}(t)$ is a vector of functions of t

$$\vec{y}(t) = \begin{bmatrix} y_1(t) \\ | \\ y_n(t) \end{bmatrix}$$

and $y(t)$ satisfies $\frac{d}{dt}\vec{y}(t) = \underset{n\times n}{B}\,\underset{n\times 1}{\vec{y}(t)}$

given $\vec{y}(0)$ find $\vec{y}(\Delta t)$ (approximately)

finite difference

$$\frac{1}{\Delta t}\left[\vec{y}(\Delta t) - \vec{y}(0)\right] = B\,y(\Delta t)$$

by choosing to put Δt here instead of 0 we make this an implicit method

$$\underbrace{(I - \Delta t\, B)}_{A}\,\underset{\substack{\text{unknown}\\ x_1}}{\vec{y}(\Delta t)} = \underset{\substack{\text{known}\\ b_1}}{\vec{y}(0)}$$

solve for $y(\Delta t)$.

then to advance the sol^{n} another Δt

$$\underset{A}{(I - \Delta t B)}\,\underset{x_2}{\vec{y}(2\Delta t)} = \underset{b_2}{y(\Delta t)}$$

$\Rightarrow$ $\left\{\begin{array}{l}\text{repeatedly solving} \\ \text{the rhs not known in advance}\end{array}\right.$ ✓

Better than storing A^{-1} etc? Sometimes, especially if A is ~~For~~ large & sparse

Make the example A "larger"

5×5

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

A is tridiagonal with same pattern as before. Requires storing $\sim 3n$ numbers instead of n^2. Lots of zeros = sparse. Fast to multiply etc.

It turns out

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 & 0 \\ 0 & -2/3 & 1 & 0 & 0 \\ 0 & 0 & -3/4 & 0 & 0 \\ 0 & 0 & 0 & -4/5 & 1 \end{bmatrix} \quad \& \quad U = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & -3/2 & 1 & 0 & 0 \\ 0 & 0 & -4/3 & 1 & 0 \\ 0 & 0 & 0 & -5/4 & 1 \\ 0 & 0 & 0 & 0 & -6/5 \end{bmatrix}$$

* U & L are bidiagonal & the cost of storing them is also $\sim 3n$ (why not $4n$?) same as A.

Where A is sparse there is a good chance that L & U are sparse. Not so for A^{-1}.

not spare $\longrightarrow$

dense

$$A^{-1} = -\frac{1}{6} \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

(7)

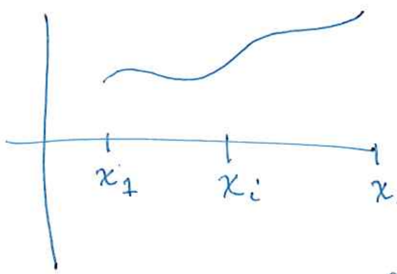
In particular,

tridiagonal has both

1. When a row of A starts w/ zeros, so does the same ~~corresponding~~ row of L b/c nothing to eliminate.

2. When a column of A starts (from the top) w/ zeros so does that ~~row~~ row of U b/c each row of U is L.C. of the rows of A above it.

"Is this matrix famous; how does it get larger?"



Suppose we wish to approximate the derivative of a function $f(x)$ known at points $x_1, \dots, x_n$

$x_{i+1} - x_i = h$ = spacing

the (forward) finite difference is $\dfrac{f(x_{i+1}) - f(x_i)}{h}$

the (backward) finite difference is $\dfrac{f(x_i) - f(x_{i-1})}{h}$

and an approximation of f'' is given by their difference $\times \frac{1}{h}$

$$\frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{h^2} \approx f''(x_i)$$

If $\frac{d^2f}{dx^2} = g(x)$ and we evaluate at $x = x_1, \dots x_n$

$$f''(x_i) = g(x_i) \qquad i = 1, \dots, n$$

and use finite difference

n linear equation in n unknowns $f(x_i)$

$$\frac{1}{h^2}\left(f(x_{i+1}) - 2f(x_i) + f(x_{i-1})\right) = g(x_i)$$

in matrix form w/ $n=4$

$$\frac{1}{h^2} \underbrace{\begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}}_{K} \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_4) \end{bmatrix} = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_4) \end{bmatrix}$$

Note that the 1st & last equations are missing $f(x_0)$ & $f(x_{n+1})$ which means they are taken to be zero, and solving

$$\frac{1}{h^2} K \vec{f} = \vec{g} \quad \text{is solving}$$

$$f'' = g \quad \text{w/ zero boundary condition}$$

$f =$ zero on boundaries.

Other BCs would change the 1st & last rows. In weather/climate applications can be $n \sim 10^9$, 10^{18} single precision $= 4$ exabytes not good

Not every square matrix $A = LU$ b/c row exchanges. But $PA = LU$ works where P is a permutation matrix that does the row exchanges.

Permutation facts $\underset{n \times n}{P}$

1. Each row of P is a row of I, no repeats
2. There are $n!$ permutation matrices b/c $n!$ permutations
3. The product of permution matrices is a permutation, b/c in the end its a permutation
4. Permutation matrices are invertible.
   - square
   - n pivots (re-order the rows to get I)

What is P^{-1}? Consider the 3×3 case

unknown Known

$$\underset{P^{-1}}{\begin{bmatrix} \text{row 1} \\ \text{2} \\ \text{3} \end{bmatrix}} \underset{P}{\begin{bmatrix} \text{col 1} & 2 & 3 \end{bmatrix}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

row 1 of P^{-1} is $\perp$ to columns 2 & 3 of P

and so on

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

P^{-1} is the matrix whose rows are the columns of P

In other words

$P^{-1} = P^T$ transpose = switch columns and rows

$(P^T)_{ij} = P_{ji}$ switch indices

Sometimes a permutation matrix is it own inverse

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

why? B/C $P = P^T$. A matrix that is equal it its transpose is called ~~symmet~~ symmetric

$S = S^T$ - application PDE, stats
- nice properties

The rules of transpose - (square)

0. $(A^T)^T = A$

1. $(A+B)^T = A^T + B^T$

2. $(AB)^T = B^T A^T$

3. $[A^T]^{-1} = [A^{-1}]^T$ if A^{-1} exists so does $[A^T]^{-1}$.

Comments on 2. & 3.

$$\underset{\substack{\uparrow \\ \text{L.C. of columns of } A}}{(Ax)^T} = \overbrace{x^T A^T}^{\text{L.C. of } \cancel{\text{columns}} \text{ rows of } A^T \text{ (which are columns of } A)}$$

$$\underset{\substack{\uparrow \\ \text{col. 3 of } A}}{\left(A\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)^T} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} A^T = \text{row 3 of } A^T$$

$$= \text{col. 3 of } A$$

3.

$$AA^{-1} = I \quad \& \quad [AA^{-1}]^T = I \quad \cancel{= A^T [A^{-1}]^T}$$

$$\cancel{\boxed{A^{-1}}} \quad \cancel{= Q}$$

$$[A^{-1}]^T A^T = I$$

$\uparrow$ inverse from left

$$[A^{-1}A] = I$$

$$[A^{-1}A]^T = I$$

$$A^T [A^{-1}]^T = I$$

$\uparrow$ from right

✓