2.5 Matrix Factorizations

Sometimes we can <u>factor</u> A into the product of 2 (or more) matrices.

We say that A $(m \times n)$ has (or admits) an LU factorization if it can be written in the form $A = LU$ where L is $m \times m$, is lower triangular, and has 1's on the diagonal and U is $m \times n$ and is upper triangular. <u>L is invertible (why?)</u>.

Suppose $A = LU$. Then the system $A\vec{x} = \vec{b}$ becomes $LU\vec{x} = \vec{b}$. Now set $U\vec{x} = \vec{y}$ so we have

$$L\vec{y} = \vec{b}$$
$$U\vec{x} = \vec{y}$$

Why do this? Why trade the solution of one system for solutions of 2 systems? To save work.

Ex $\quad A = \begin{bmatrix} 3 & 1 & 0 \\ 6 & 4 & -1 \\ 3 & 7 & -2 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 5 \\ 5 \\ -7 \end{bmatrix}$

To solve, $A\vec{x} = \vec{b}$ we form $[A \ \vec{b}]$ and row reduce to get $[I \ \vec{x}]$.

$$\left[\begin{array}{ccc:c} 3 & 1 & 0 & 5 \\ 6 & 4 & -1 & 5 \\ 3 & 7 & -2 & -7 \end{array}\right] \sim \left[\begin{array}{ccc:c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

6 DIVISIONS
8 MULTIPLICATIONS
8 ADDITIONS

$$\vec{x} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

However, if we knew $A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 3 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{U}$

Then to solve $L\vec{y} = \vec{b}$ we solve

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 2 & 1 & 0 & 5 \\ 1 & 3 & 1 & -7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

0 DIV.
3 MULT
3 ADD.

Now solve $U\vec{x} = \vec{y}$

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

3 DIV
3 MULT
3 ADD.

The savings are <u>much</u> more pronounced for larger systems

Solving Triangular systems is easy (inexpensive).

Solving $L\vec{y} = \vec{b}$ is sometimes called a forward solve since

$$L\vec{y} = \vec{b} \Rightarrow \begin{aligned} &l_{11} y_1 &&= b_1 \quad (1) \\ &l_{21} y_1 + l_{22} y_2 &&= b_2 \quad (2) \\ &l_{31} y_1 + l_{32} y_2 + l_{33} y_3 &&= b_3 \quad (3) \\ &\vdots \end{aligned}$$

find y_1 from (1), sub into (2) and solve for y_2, sub these into (3) and find y_3, etc.

(Forward Solve)

Similarly, $U\vec{x}=\vec{y}$ is called a backward solve

$$\begin{aligned}
u_{11}x_1 + u_{12}x_2 + \cdots \quad & + u_{1n}x_n = y_1 \quad (1) \\
u_{22}x_2 + \cdots \quad & + u_{2n}x_n = y_2 \quad (2) \\
& \vdots \\
& \quad (n-1) \\
& u_{nn}x_n = y_n \quad (n)
\end{aligned}$$

Solve (n) for x_n, use this in $(n-1)$ to find x_{n-1}, etc.

Given A, how do we find LU?

Assuming A has no zero elements on the diagonal (and none appear) we can reduce A to an upper triangular matrix U using only row replacement operations (no row interchanges or scalings).

Recall that these operations can be affected by multiplications on the left by elementary matrices

$$E_p \cdots E_1 A = U$$

where all $E_p, \dots, E_1$ are lower triangular and invertible.

$$(E_p \cdots E_1)A = U \rightarrow A = (E_p \cdots E_1)^{-1}U$$

Note that $E_p \cdots E_1$ is lower triangular and the inverse of a lower triangular matrix is also lower triangular so $(E_p \cdots E_1)^{-1} = L$ is lower triangular.

Also, we know $(E_p \cdots E_1)L = I$ since

$$(E_p \cdots E_1)L = (E_p \cdots E_1)(E_p \cdots E_1)^{-1} = I$$

So L is a lower triangular matrix which has the property that when the same row operation which reduce A to upper triangular U are applied to L results in I.

<u>Ex</u>

$$A = \begin{bmatrix} 3 & 1 & 2 & 0 \\ 6 & 4 & 5 & -1 \\ 3 & 7 & 5 & -2 \end{bmatrix}$$

our first step will be to reduce the first column of A to $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$. This is exactly what we want to do to the first column of L, except we also want $\underline{1}$ in the first position.

So, it makes sense to take as the first column of L the first column of A, scaled by the first element.

$$L = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ 1 & 3 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} \boxed{3} & 1 & 2 & 0 \\ \boxed{6} & 4 & 5 & -1 \\ \boxed{3} & 7 & 5 & -2 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 2 & 0 \\ 0 & \boxed{2} & 1 & -1 \\ 0 & \boxed{6} & 3 & -2 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 2 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

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

Costs

Solving $A\vec{x} = \vec{b}$ $\quad A\ (n \times n)$

① Row reduction $[A\ \vec{b}] \to [I\ \vec{x}]$: $\sim \frac{n^3}{3} + \frac{n^2}{2}$ operations

② $\vec{x} = A^{-1}\vec{b}$: $\sim n^3$ ops/inverse $+\ n^2$ ops for mult. $\quad \sim n^3 + n^2$ operations

③ $L\vec{y} = \vec{b}$; $U\vec{x} = \vec{y}$: $\sim n^2$ operations

	$n = 10$	10^3	10^6	
①	~ 383	3.34×10^8	3.33×10^{17}	# operations
②	~ 1100	1.00×10^9	1.00×10^{18}	
③	~ 100	1.00×10^6	1.00×10^{12}	

of course, the work to find LU is not included in ③.

If we assume that a particular computer can do 200 MFLOPS then we have

	$n = 10$	10^3	10^6	
①	1.9×10^{-6} sec	1.67 sec	52.7 years	time
②	5.5×10^{-6} sec	5 sec	158.4 years	
③	0.5×10^{-6} sec	0.005 sec	83.3 minutes	