LU Factorization

Sunday, November 29, 2020 2:46 PM

LU factorization is how we actually do Gaussian Elimination on a computer
This is ch. 6.5 in Burden and Faires on Matrix Factorization

LU is an example of a $\underbrace{\text{matrix factorization}}$
means writing your matrix "A"
as the product of other matrices which have
special properties

Ex of matrix factorizations

covered in our class → LU $A = L \cdot U$ L is Lower triangular, U is Upper triangular

Eigenvalue $A = VDV^{-1}$ V is invertible, D is diagonal (only for square matrices)

"SVD"
Singular Value $A = V\Sigma U^T$ V, U are orthogonal ($V^TV = I$ ...), Σ is a diagonal matrix w/ nonzero entries

QR $A = QR$ Q is orthogonal, R is upper triangular (like Gram-Schmidt)

Schur, Polar not as common

NOT covered in our class

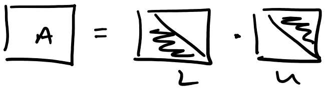
LDL^T, Cholesky variants of LU that we'll talk about briefly

LU decomposition to solve a linear system of equations

Want to solve $A\vec{x} = \vec{b}$, suppose we can write $A = LU$

n × n (square)

lower triag'l. upper trig.



so solve $L\underbrace{U\vec{x}}_{\vec{y}} = \vec{b}$

let $\vec{y} = U\vec{x}$

① Solve $L\vec{y} = \vec{b}$

This is "easy" ($O(n^2)$ flops not $O(n^3)$) since L is Lower Triangular

"Forward Substitution"

3×3 example:
$$\begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

1) $L_{11} \cdot y_1 = b_1$ so $y_1 = b_1 / L_{11}$

2) $L_{21} \overset{\text{known!}}{y_1} + L_{22} \underset{\text{unknown}}{y_2} = b_2$ so $y_2 = (b_2 - L_{21} y_1) / L_{22}$

3) $L_{31} \underset{\text{known}}{y_1} + L_{32} \underset{\text{known}}{y_2} + L_{33} \underset{\text{unknown}}{y_3} = b_3$

② Solve $U\vec{x} = \vec{y}$

Also "easy" ($O(n^2)$) since U is Upper triangular

"Back substitution"

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

3) $u_{33} x_3 = y_3$

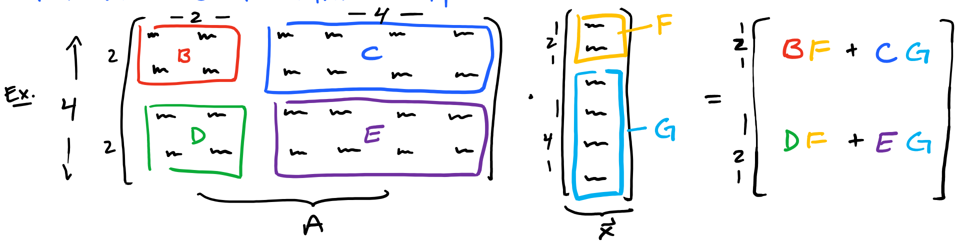
2) $u_{22} x_2 + u_{23} x_3 = y_2$

1) $u_{11} x_1 + u_{12} x_2 + u_{13} x_3 = y_1$

★ TRIANGULAR SYSTEMS OF EQUATIONS ARE EASY TO SOLVE IF YOU SOLVE THEM IN THE RIGHT ORDER

Interlude: block matrix multiplication

Ex.



Any size blocks work (as long as they're all compatible)

So, if A is square, and we want to keep the 1st block of $A\vec{x}$ the same as $\vec{x}$, meaning $F = BF + CG$

then choose $B = I$
$C = 0$

Finding the LU decomposition

We'll do Gaussian elimination, and it'll cost $O(n^3)$ flops, so this is the expensive part

"Gaussian Elimination = LU"

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

$$\left[\begin{array}{c|ccc|c} E_1 & a_{11} & a_{12} & a_{13} & b_1 \\ E_2 & a_{21} & a_{22} & a_{23} & b_2 \\ E_3 & a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

$E_2 \leftarrow E_2 - \overset{\alpha_2}{a_{21}/a_{11}} \cdot E_1$

$E_3 \leftarrow E_3 - \underset{\alpha_3}{a_{31}/a_{11}} \cdot E_1$

E_i = i^{th} row

to get

$$\left[\begin{array}{ccc|c}a_{11}&a_{12}&a_{13}&b_1\\0&\sim&\sim&\sim\\0&\sim&\sim&\sim\end{array}\right]$$

we "zeroed out" 1st column

just means "something"

Connect this step to matrix multiplication

$$\begin{bmatrix}1&0&0\\-\alpha_2&1&0\\-\alpha_3&0&1\end{bmatrix}\cdot\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}a_1\\-\alpha_2 a_1+a_2\\-\alpha_3 a_1+a_3\end{bmatrix}$$

Recall

$$A\cdot[\vec{y}_1,\vec{y}_2,\vec{y}_3]$$
$$=[A\vec{y}_1,A\vec{y}_2,A\vec{y}_3]$$

so

$$\underbrace{\begin{bmatrix}1&0&0\\-\alpha_2&1&0\\-\alpha_3&0&1\end{bmatrix}}_{M^{(1)}}\cdot\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&\sim&\sim\\0&\sim&\sim\end{bmatrix}$$

So that 1st step of Gaussian Elimination was equivalent to

$$M^{(1)}A\vec{x}=M^{(1)}\vec{b}$$

Now we do more steps, but ignore 1st row

so

$$M^{(2)}=\begin{bmatrix}\overset{=I}{\boxed{1}}&\overset{=0}{\boxed{0\quad 0}}\\\boxed{\begin{matrix}0\\0\end{matrix}}&\boxed{\begin{matrix}1&0\\\sim&1\end{matrix}}\end{bmatrix}$$

this has the block structure we just talked about in the interlude

$$M^{(2)}M^{(1)}A\vec{x}=M^{(2)}M^{(1)}\vec{b}$$

For a n×n matrix, end up with

$$\underbrace{\underbrace{M^{(n-1)}M^{(n-2)}\cdots M^{(1)}}_{\tilde{L}}A}_{U}\vec{x}=M^{(n-1)}M^{(n-2)}\cdots M^{(1)}\vec{b}$$

we know that via Gaussian elimination, we get

$$\left[\begin{array}{c|c}\underbrace{\begin{matrix}\sim&\sim&\sim\\0&\sim&\sim\\0&0&\sim\end{matrix}}_{\substack{\text{upper triangle}\\U}}&\begin{matrix}\sim\\\sim\\\sim\end{matrix}\end{array}\right]\leftarrow\tilde{L}\vec{b}$$

$\sim$ just means some number I don't want to give a variable name

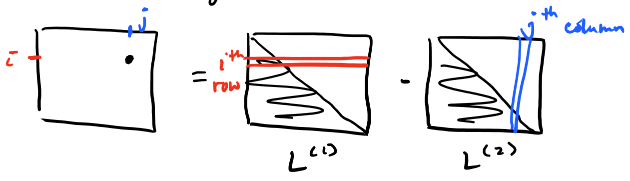
i.e. $\tilde{L}A=U$, so $A=LU$ if $L:=\tilde{L}^{-1}$

$\mathcal{U}$ is upper triangular (via design of Gaussian elimination)

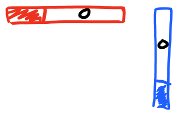
but is L really Lower triangular?

1) note $M^{(i)}$ are lower triangular. Is the product of two lower triangular matrices also lower triangular?

Yes!



$(L^{(1)}L^{(2)})_{ij}$ =



= 0 if $j > i$

So, $\underset{\text{lower trig}}{M^{(3)}} \underbrace{M^{(2)} M^{(1)}}_{\text{lower triang.}}$ is also lower triangular!

etc. $\Rightarrow$ $\tilde{L} = M^{(n-1)} M^{(n-2)} \cdots M^{(2)} M^{(1)}$ is lower triangular

2) if $\tilde{L}$ is lower triangular, is its inverse $L := \tilde{L}^{-1}$ also lower triangular?

yes! the inverse is the solution X to the equation $\tilde{L} X = I$

then the j^{th} column of X (i.e., of L) solves

$$\begin{bmatrix} \sim & 0 & 0 & \cdots & 0 \\ \sim & \sim & 0 & \cdots & \vdots \\ \sim & \sim & \sim & & 0 \\ \vdots & & & \ddots & \\ \sim & \sim & & \sim & \sim \end{bmatrix} \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{e_j} \begin{array}{l} \\ \\ \\ \leftarrow j^{th} \text{ spot} \\ \\ \\ \end{array}$$

So by forward substitution,

$L_{11} x_{1j} = 0 \quad \Rightarrow \quad x_{1j} = 0$

etc.

until we get to x_{jj}

So...

Gaussian elimination gave us $\tilde{L} A = \mathcal{U}$ which we can use to solve $(\tilde{L} A \vec{x} = \mathcal{U} \vec{x}) = \tilde{L} \vec{b}$ (ie. multiply $A\vec{x} = \vec{b}$ by $\tilde{L}$ on the left)

or, similarly, $L := \tilde{L}^{-1}$

$\underline{\underline{A = L\mathcal{U}}}$ and solve as we did earlier