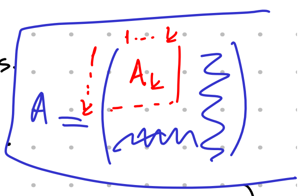
Thm Cholesky Theorem on LL^T factorizations

If A is a real, symmetric and pos. def. matrix, then it has a unique factorization, $A = LL^T$ in which L is lower triangular with positive diagonal entries.



Pf $A^T = A$ and $x^T A x > 0$ for all $x \neq 0$.

$\Rightarrow$ A is invertible (in fact all eigenvalues are positive see HW3)

In fact $A_k^T = A_k$ $\forall k$ and if $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,

$x^T A x = (x_1 \cdots x_k) A_k \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ So A_k is also pos. def.

In particular, all n leading principal minors of A are invertible.

Hence by prev. thm, it has an LU-decomposition.

Now, $LU = A = A^T = U^T L^T$ multiply from left by L^{-1} // // right by $(L^T)^{-1}$

$\Rightarrow$ $\underset{\substack{\uparrow \\ \text{upper} \\ \text{tri.}}}{U(L^T)^{-1}} = \underset{\substack{\uparrow \\ \text{lower} \\ \text{tri.}}}{L^{-1} U^T}$

Inverses and products of lower (upper) tri. are lower (upper) tri. see HW3.

Thus, we have $\underset{\text{tri}}{\text{lower}} = \underset{\text{tri}}{\text{upper}}$ so they are both diagonal.

Set $D = U(L^T)^{-1}$ so $U = DL^T$

Thus, $A = LU = LDL^T \Rightarrow D$ is pos. def! see HW3

So D also has all eigenvalues positive. Since it is diagonal, this means all diagonal entries $d_{ii} > 0$.

Set $D^{1/2}$ be the diag. matrix with entries $\sqrt{d_{ii}}$.

let $\hat{L} = LD^{1/2}$ Then $A = LDL^T = LD^{1/2}D^{1/2}L^T$
$= \hat{L}\hat{L}^T$

Uniqueness is exercise!

4.3 Pivoting and Constructing an Algorithm

Basic Gaussian Elimination

Example Solve $\begin{pmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ 34 \\ 27 \\ -38 \end{pmatrix}$

Apply $R_2 - 2R_1 \to R_2$, $R_3 - \frac{1}{2}R_1 \to R_3$ and $R_4 + R_1 \to R_4$ to get

$$\left(\begin{array}{cccc|c} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & -12 & 8 & 1 & 21 \\ 0 & 2 & 3 & -14 & -26 \end{array}\right)$$

$\leftarrow$ Pivot row (in the first step) and 6 is the pivot element.

$+2$, $+\frac{1}{2}$, -1 are called multipliers (for the first step)

Next $R_3 - 3R_2 \to R_3$ $R_4 + \frac{1}{2}R_2 \to R_4$

$$\left(\begin{array}{cccc|c} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & 0 & 2 & -5 & -9 \\ 0 & 0 & 4 & -13 & -21 \end{array}\right)$$

row 2 is the pivot row and -4 is the pivot element in the second step.

Multipliers are $3, -\frac{1}{2}$

Next $R_4 - 2R_3 \to R_4$

$$\left(\begin{array}{cccc|c} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & 0 & 2 & -5 & -9 \\ 0 & 0 & 0 & -3 & -3 \end{array}\right)$$

Multiplier is 2.

$\hookrightarrow$ $-3x_4 = -3 \Rightarrow x_4 = 1.$

$2x_3 - 5x_4 = 2x_3 - 5 = -9 \Rightarrow x_3 = -2$

$-4x_2 - 4 + 2 = 10 \Rightarrow x_2 = -3$

$6x_1 + 6 - 4 + 4 = 12 \Rightarrow x_1 = 1$

$x = \begin{pmatrix} 1 \\ -3 \\ -2 \\ 1 \end{pmatrix}$ is unique solution.

Recall that applying ERO $\Leftrightarrow$ multiplication by elementary matrices

Thus $E_6 \cdots E_2 E_1 A = U = \begin{pmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 0 & -3 \end{pmatrix}$

Hence, $A = E_1^{-1} E_2^{-1} \cdots E_6^{-1} U$

The inverse of $R_i + cR_j \to R_i$ is $R_i - cR_j \to R_i$

(This is why we took negatives to get multipliers.)

The last operation we applied was $R_4 - 2R_3 \to R_4$

So $E6 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & -2 & 1 \end{pmatrix} \Rightarrow E_6^{-1} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & 2 & 1 \end{pmatrix}$ notice last multiplier

The previous two operations were

$R_3 - 3R_2 \to R_3$ $R_4 + \frac{1}{2}R_2 \to R_4$

Thus, $E_5 = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & \frac{1}{2} & 0 & 1 \end{pmatrix}$ and $E_5^{-1} = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & -\frac{1}{2} & 0 & 1 \end{pmatrix}$

Note that $E_4^{-1} E_5^{-1} = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 3 & 1 & \\ 0 & -\frac{1}{2} & 0 & 1 \end{pmatrix}$ multipliers.

$E_4^{-1} E_5^{-1} E_6^{-1} = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 3 & 1 & \\ 0 & -\frac{1}{2} & 2 & 1 \end{pmatrix}$

In general, set $L = E_1^{-1} E_2^{-1} \cdots E_6^{-1}$

then
$$L = \begin{pmatrix} 1 & & & \\ 2 & 1 & & \\ \frac{1}{2} & 3 & 1 & \\ -1 & -\frac{1}{2} & 2 & 1 \end{pmatrix}$$
is a unit lower triangular matrix

and $A = LU$ is Doolittle's factorization.

Note that in their process, we only used ERO.s of the form $R_i + cR_j \to R_i$ where $i > j$. This is precisely why L is unit and lower triangular.

If any one of the pivot elements was a 0, then the process would have failed (unless we alter it).

Assuming none of the pivot elements is 0, we will write this algorithm in general now.

Say A is the given $n \times n$ matrix.

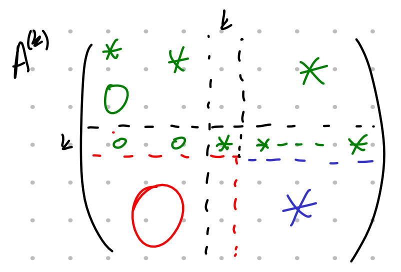
Set
$$A = A^{(1)} \sim A^{(2)} \sim A^{(3)} \sim \cdots \sim A^{(n)}$$

so that
$$A^{(k)} = \begin{pmatrix} a_{11}^{(k)} & \cdots & & & & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & \cdots & & & a_{2n}^{(k)} \\ \vdots & 0 & \ddots & & & \vdots \\ \vdots & \vdots & 0 & a_{k-1\,k-1}^{(k)} & & \\ & & & 0 & a_{kk}^{(k)} & a_{kn}^{(k)} \\ i & & & \vdots & a_{ik}^{(k)} & \\ 0 & & & 0 & a_{nk}^{(k)} \cdots & a_{nn}^{(k)} \end{pmatrix}$$

At the k^{th} step, starting with $A^{(k)}$, we apply the EROs given by $R_i - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} R_k \to R_i$ for $k < i \leq n$ to make all the elements below the pivot ($a_{kk}^{(k)} \neq 0$) 0 and that gives us $A^{(k+1)}$.

Therefore, starting with $A^{(1)} = A$ ($a_{ij}^{(1)} = a_{ij}$) we can define $A^{(k+1)}$ inductively by

$$a_{ij}^{(k+1)} = \begin{cases} a_{ij}^{(k)} & \text{if } i \leq k \\ a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)} & \text{if } i \geq k+1 \text{ and } j \geq k+1 \\ 0 & \text{if } j \leq k \text{ and } i \geq k+1 \end{cases}$$



Then we set

$U = A^{(n)}$ (which is upper triangular by construction)

and define L by
$$l_{ik} = \begin{cases} a_{ik}^{(k)} / a_{kk}^{(k)} & \text{if } i \geq k+1 \\ 1 & \text{if } i = k \\ 0 & \text{if } i \leq k-1 \end{cases}$$
$$L = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ \text{Multipliers} & \ddots & \\ & & 1 \end{pmatrix}$$

<u>Thm</u> (Theorem on Nonzero Pivots)

If all the pivot elements $a_{kk}^{(k)}$ are nonzero, then $A = LU$.

<u>Pivoting</u>

The above algorithm fails even on $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

or if we have $\begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ then $R_2 - \varepsilon^{-1} R_1 \to R_2$

$\varepsilon x_1 + x_2 = 1$

$$\begin{pmatrix} \varepsilon & 1 \\ 0 & 1 - \varepsilon^{-1} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 - \varepsilon^{-1} \end{pmatrix}$$
So $x_2 = \frac{2 - \varepsilon^{-1}}{1 - \varepsilon^{-1}}$ and $x_1 = \frac{1 - x_2}{\varepsilon}$

if ε is small enough, $fl(2 - \varepsilon^{-1}) = -\varepsilon^{-1}$ and $fl(1 - \varepsilon^{-1}) = -\varepsilon^{-1}$

thus $x_2 = 1$ and $x_1 = 0$. In this case, we have

$$\begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;!$$

Here the problem is not that ε is very small but rather that it is small relative to other elements in its row!

Let's consider

$$\begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 - \varepsilon \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 - 2\varepsilon \end{pmatrix} \to \begin{matrix} x_1 + x_2 = 2 \\ (1-\varepsilon) x_2 = (1 - 2\varepsilon) \end{matrix}$$

$\Rightarrow x_2 = \frac{1 - 2\varepsilon}{1 - \varepsilon} \approx 1$ and $x_1 = 2 - x_2 \approx 2 - 1 = 1$

Then $\begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \approx \begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ \varepsilon + 1 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

We see that even if we can't compute the exact answers, we still have a good approximation.

Q] How do we determine which permutations to use?

A] <u>Gaussian Elimination with Scaled Row Pivoting</u>

$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ example of a perm matrix

Say we want to solve $Ax = b$ and we somehow determined a permutation matrix P s.t. PA has an LU decomposition.

Then we have $PA = LU$ and <u>$PAx = Pb$</u>

Thus, $LUx = Pb$ or $L(Ux) = Pb$. Setting $Ux = z$

we can first solve $Lz = Pb$ by forward subs. and
then solve $Ux = z$ by back subs..

So we need to figure out how to determine P or
equivalently a permutation vector $(p_1\ p_2 \cdots p_n)$. e.g. (3.4.21)

Given A, set $s_i = \max\limits_{1 \le j \le n} |a_{ij}|$ (maximum modulus on ith row)

Start with perm. vect. $p = (1\ 2\ \cdots\ n)$ (representing no perm.)

Find the largest $|a_{i1}|/s_i$. The corresponding row will be our first pivot row. Say it is p_1^{th} row then swap 1 and p_1 in the perm vector.

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \begin{matrix} \to s_1 \\ \to s_2 \\ \to s_3 \end{matrix}$$

$$(1\ 2 \cdots p_1 \cdots n) \longrightarrow (p_1\ \underbrace{2\ 3 \cdots 1 \cdots n})$$

Using p_1^{th} row as pivot, make all other elements in the 1^{st} column 0.
(applying $R_i - \frac{a_{i1}}{a_{p_1 1}} R_{p_1} \to R_i$ for i in this set!)

Next, we look for largest $|a_{i2}|/s_i$ for i in the same set.

Then the perm vector becomes $(p_1\ p_2\ 3\ 4 \cdots 2 \cdots 1 \cdots n)$.

and we continue in this manner.

We will start with an example next time.