LECTURE 15

Return to usual participation policy starting next Monday, 5/8. (Week 6)

Review of LU decomposition (§3.1)

Gaussian elimination: you are modifying the bases so that the matrix becomes more "compact" (eg. diagonal or triangular).

Review: Gauss Elimination (w/ partial pivoting)

difference: Minimize the # calculations you do like

$1000 + 0.0001 \quad (\approx 1000 - 0.0001!)$

ie try to do meaningful, robust arith. operations

eg. $\begin{bmatrix} 1 & 1 \\ 300 & -200 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 400 \end{bmatrix}$

In small example, we humans may see that you could just multiply the top row by 10^2. But we have checked that the magnitudes across the whole row are ~1! For large matrix, just want an algorithm to run in every scenario!

$$\downarrow$$

$$\left[\begin{array}{cc|c} 1 & 1 & 3 \\ 300 & -200 & 400 \end{array}\right]$$

$$\Big\downarrow \; |300| > |1|. \text{ pivot}$$

$$\left[\begin{array}{cc|c} 300 & -200 & 400 \\ 1 & 1 & 3 \end{array}\right]$$

$$\Big\downarrow \; \text{G.E} \quad r_2 \rightsquigarrow r_2 - \frac{r_1}{300}$$

$$\left[\begin{array}{cc|c} 300 & -200 & 400 \\ 1-1 & 1+\frac{200}{300} & 3-\frac{400}{300} \end{array}\right] = \left[\begin{array}{cc|c} 300 & -200 & 400 \\ 0 & 5/3 & 5/3 \end{array}\right]$$

$$\Rightarrow \boxed{x_2 = 1},$$

$$300x_1 - 200\cancel{x_2}^{\,1} = 400$$

$$\Rightarrow \boxed{x_1 = 2}$$

We used:

① pivoting $\Longrightarrow$ acheived by permutation matrix

$$P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

② Subtracted multiple of first row from second
(and in general, all other rows):

$$L_1 = \begin{pmatrix} 1 & 0 \\ -\frac{1}{300} & 1 \end{pmatrix}$$ $\rightsquigarrow$ note: $\begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -m & 1 \end{pmatrix}$

In general:

$$L_1 = \left(\begin{array}{c|c} 1 & 0 \\ \hline m_1 & I \end{array}\right) \quad \text{where} \quad m = \begin{pmatrix} m_{21} \\ m_{31} \\ \vdots \\ m_{n1} \end{pmatrix}$$

These are all lower triangular.

<u>thm</u> (LU decomposition)

Any nonsingular $A \in \mathbb{R}^{n\times n}$ can be decomposed into

$$\underbrace{P}_{\text{permutation}} A = \underbrace{L}_{\text{lower } \Delta}\ \underbrace{U}_{\text{upper } \Delta}.$$

eg. In our example,

$$A \longrightarrow PA \longrightarrow L^{-1}PA = A^{(1)}$$

$$\begin{pmatrix} 1 & 1 \\ 300 & -200 \end{pmatrix} \mapsto \begin{pmatrix} 300 & -200 \\ 1 & 1 \end{pmatrix} \longmapsto \begin{pmatrix} 300 & -200 \\ -1+1 & \frac{2}{3}+1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{300} & 1 \end{pmatrix}\begin{pmatrix} 300 & -200 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 300 & -200 \\ -1+1 & \frac{2}{3}+1 \end{pmatrix}$$

We now have better operation than just row ops:

- Givens rotations
- Householder reflections

We'll use this to factor any matrix A into QR

$$\underbrace{Q}_{\text{orthogonal}}\ \underbrace{R}_{\text{(upper) triangular, rectangular!}} \qquad \begin{pmatrix} R \\ 0 \end{pmatrix}$$

* This is the difference b/w solving the exactly determined linear system $Ax=b$ where $A \in \mathbb{R}^{n\times n}$

and solving the least squares problem for overdetermined "$Ax=b$" where $A \in \mathbb{R}^{m\times n}$ where $n \le m$.

Let's first state the theorem and see how this is useful:

<u>thm</u> (QR decomp) Any matrix $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ can be transformed into upper triangular form by an orthogonal matrix. The transformation is equiv to a decomposition

$$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix} \qquad \boxed{A} = \boxed{Q}\ \begin{bmatrix} R \\ \hline 0 \end{bmatrix}$$

Moreover, if A is $\underbrace{\text{nonsingular}}_{\text{full col rank}}$, then R is $\underbrace{\text{also}}_{\text{invertible! (b/c square)}}$! (This should be clear from linear transformation perspective)

Since the last $m-n$ cols of Q will be multiplied by 0,

$$A = (Q_1\ Q_2)\begin{pmatrix} R \\ 0 \end{pmatrix} = Q_1 R$$

we can also define:

The <u>thin QR decomp</u> of A is $Q_1 R$.

$$\overset{a_j}{\begin{bmatrix} A \end{bmatrix}} = \underbrace{\begin{bmatrix} Q_1 \end{bmatrix}}_{\text{orthogonal basis for the range}(A).} \overset{\text{coordinates of } a_j \text{ in the ortho basis.}}{\begin{bmatrix} R \end{bmatrix}}$$

We'll use QR decomposition to solve least squares without forming the normal equations

This answers the question: why is the shortest residual vector the normal one?

Problem: $A \in \mathbb{R}^{m\times n}$, $n \leq m$, full rank. Find $x \in \mathbb{R}^n$ such that the residual vector $r = b - Ax$ $(b \in \mathbb{R}^m)$ is minimized.

Solution: Suppose we have the QR decomp of A.

$$\|r\|_2^2 = \|b - Ax\|_2^2$$

$$= \left\|b - Q\begin{pmatrix} R \\ 0 \end{pmatrix} x\right\|_2^2 \quad \text{by QR-decomp. of } A$$

$$= \left\|QQ^T\left(b - Q\begin{pmatrix} R \\ 0 \end{pmatrix} x\right)\right\|_2^2 \quad QQ^T = Id_m$$

$$= \left\|Q\left(Q^Tb - Q^TQ\begin{pmatrix} R \\ 0 \end{pmatrix} x\right)\right\|_2^2 \quad \text{distribute}$$

$$= \left\|Q^Tb - \begin{pmatrix} R \\ 0 \end{pmatrix} x\right\|_2^2 \quad \|Qy\|_2 = \|y\|_2 \quad \forall y \in \mathbb{R}^m$$

$$= \left\|\begin{pmatrix} Q_1^Tb \\ Q_2^Tb \end{pmatrix} - \begin{pmatrix} Rx \\ 0 \end{pmatrix}\right\|_2^2 \quad \text{rewrite}$$

$$= \|Q_1^Tb - Rx\|_2^2 + \|Q_2^Tb\|_2^2 \quad \sum_{i=1}^m y_i^2 = \sum_{i=1}^n y_i^2 + \sum_{j=n+1}^m y_j^2$$

This is minimized when $\|Q_1^Tb - Rx\|_2^2$ is minimized

$\|Q_2^Tb\|$ is non-negotiable as it doesn't change as we change x.

This is actually the portion of b normal to colspace(A).

Since R is invertible, we can actually achieve $Q^Tb - Rx = 0$:

$$x = R^{-1}Q^Tb.$$

Next time: The algorithm for obtaining the QR decomp