And now we have a theorem

Given $A \in \mathbb{R}^{n\times n}$, $\exists\, P, L, U \in \mathbb{R}^{n\times n}$,

P is a PM

L is a (unit) lower Δ matrix

U is an upper Δ matrix

s.t. $\boxed{PA = LU}$

Pf. Induction on n.

$n=2$ | $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Case 1 | $a \neq 0$, last thm $\Rightarrow A = LU$, $P = I$.

Case 2 | $a = 0$, $c \neq 0$. Interchange rows!

$$\Rightarrow P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow PA = \begin{bmatrix} c & d \\ 0 & b \end{bmatrix} \leftarrow \text{upper } \Delta$$

$$= \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} c & d \\ 0 & b \end{bmatrix}}_{U}$$

Case 3 | $a = 0$, $c = 0$

$$A = \underbrace{\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}}_{\text{upper } \Delta} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}}_{U}$$

Induction step

Given A, choose PM $P^{(1,r)}$ to move the largest (in absolute value) of the 1^{st} column into the 1^{st} spot:

$$\begin{bmatrix} a_{11} & \\ \vdots & \\ a_{r1} = \alpha & \\ \vdots & \end{bmatrix} \qquad |a_{r1}| \geq |a_{i1}| \quad i = 1(1)n$$

$$P^{(1,r)} A = \left[\begin{array}{c|c} \alpha & w^T \\ \hline p & B \end{array}\right] \overset{\text{Ansatz}}{=} \left[\begin{array}{c|c} 1 & 0^T \\ \hline m & I \end{array}\right] \left[\begin{array}{c|c} \alpha & v^T \\ \hline 0 & C \end{array}\right]$$

Determine appropriate m, v, C

~~Case 1 | $\alpha = 0 \Rightarrow p = 0$~~

$v = w, \alpha m = p, \quad m v^T + C = B$

Case 1 | $\alpha = 0 \Rightarrow p = 0$. Take $m = 0 \Rightarrow B = C$

Case 2 | $\alpha \neq 0 \Rightarrow m = p/\alpha$ w. $m_i \leq 1 \quad i = 1(1)(n-1)$

$\cancel{B = C + m v^T} \quad C = B - m v^T$ ✓ fine decomposition.

$C \in \mathbb{R}^{(n-1)\times(n-1)}$

Induction Hypothesis

$\exists\ P^*, L^*, U^* \ni P^* C = L^* U^*$ $\quad \left[B = m v^T + (P^*)^T L^* U^* \right.$

$$\Rightarrow P^{(1,r)} A = \left(\begin{array}{c|c} 1 & 0^T \\ \hline 0 & (P^*)^T \end{array}\right) \left(\begin{array}{c|c} 1 & 0^T \\ \hline P^* m & L^* \end{array}\right) \left(\begin{array}{c|c} \alpha & v^T \\ \hline 0 & U^* \end{array}\right)$$

Check this!

Check

$$= \left(\begin{array}{c|c} 1 & 0^T \\ \hline m & (P^*)^T L^* \end{array}\right) \left(\begin{array}{c|c} \alpha & v^T \\ \hline 0 & U^* \end{array}\right) = \left(\begin{array}{c|c} \alpha & v^T = w^T \\ \hline \cancel{\alpha p} \\ \alpha m = p & \begin{array}{c} m v^T + (P^*)^T L^* U^* \\ = B \end{array} \end{array}\right) \checkmark$$

$$\text{Set } \underline{P} = \left(\begin{array}{c|c} 1 & 0^T \\ \hline 0 & P^* \end{array}\right) P^{(1,r)}$$

$$PA = \left(\begin{array}{c|c} 1 & 0^T \\ \hline P^* m & L^* \end{array}\right) \left(\begin{array}{c|c} \alpha & v^T \\ \hline 0 & U^* \end{array}\right) = LU \checkmark$$

Computational Cost of ~~LU d~~ solving

$$Ax = b$$

w. an LU decomposition.

For simplicity, assume $P = I$ (no pivoting)

<u>Recall</u> $l_{ij} = \frac{1}{u_{jj}} \left\{ a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right\}$ $\begin{array}{l} i = (2)(1)n \\ j = 1(1)(i-1) \end{array}$

$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$ $\begin{array}{l} i = 1(1)n \\ j = i(1)(n) \end{array}$

How many flops?
to find L & U $\sim \frac{2}{3} n^3 - \frac{1}{2} n^2$
for moderately large n.

Forward / backward ~~sub~~ substitution
$\sim n^2$ each

~~Backward substitution~~

<u>In total</u> $\frac{2}{3} n^3 - \frac{1}{2} n^2$ LU

$2n^2$ F/B subst.

Norms on $\mathbb{R}^n$ & $\mathbb{R}^{n\times m}$

Defn $\|\cdot\| : V \to \mathbb{R}_+$ has the following properties.

[V is a linear space over $\mathbb{R}$, i.e. $\mathbb{R}^n, \mathbb{R}^{n\times m}$]

(1) $\|v\| \geq 0 \quad \forall v \in V$ and $\|v\| = 0$ iff $v = 0$.

(2) $\lambda \in \mathbb{R} \quad \|\lambda v\| = |\lambda| \|v\|$

(3) $\|v + w\| \leq \|v\| + \|w\|$; $\forall v, w \in V$

Known as Δ inequality.

Vector norms on $\mathbb{R}^n$

L_p - norms, $p \in \mathbb{Z}_+$ (positive integers)

$$\|v\|_p = \left[\sum_{i=1}^n |v_i|^p\right]^{1/p}$$

i.e. $\|v\|_2 = \left[\sum_{i=1}^n v_i^2\right]^{1/2} = [v^T v]^{1/2}$ Euclidean norm L_2 -norm

distance from origin in $\mathbb{R}^n$.

$$\|v\|_1 = \sum_{i=1}^n |v_i|$$

L_1 - norm or "Manhattan" norm

why?

max - norm

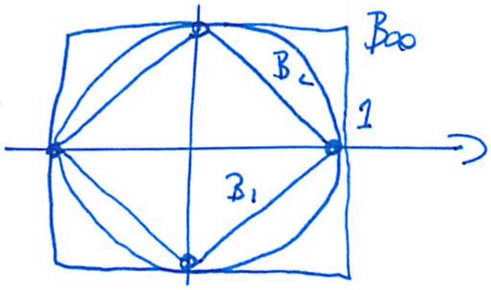
or L_∞ - norm $\|v\|_\infty = \max_i |v_i|$

Useful ?

Unit "balls" for $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$

$$B_p = \{ v \in \mathbb{R}^n \mid \|v_d\|_p \le 1 \}$$

$\mathbb{R}^2$



Fact: L_p norms are ordered.

$\|v\|_p \le \|v\|_q$

$p \le q$

Norms on matrices, ~~i.e. on $\mathbb{R}$~~

Treat A as a big vector

On $\mathbb{R}^{n\times m}$, the Frobenius norm

$$A \in \mathbb{R}^{n\times m}, \quad \|A\|_F = \Big[\sum_{i,j} a_{ij}^2\Big]^{1/2} \quad \text{like } \|\cdot\|_2$$

$$= \left(\mathrm{tr}(A^T A)\right)^{1/2} = (A : A)^{1/2}$$

Induced norms on $\mathbb{R}^{n\times n}$

Given a vector norm $\|\cdot\|$ on $\mathbb{R}^n$, define a norm on $\mathbb{R}^{n\times n}$ by

$$\text{For } A \in \mathbb{R}^{n\times n}, \quad \|A\| = \max_{\substack{v \in \mathbb{R}^n \\ v \neq 0}} \frac{\|Av\|}{\|v\|}$$

$$\|A+B\| = \max_{\substack{v \in \mathbb{R}^n \\ v \neq 0}} \frac{\|Av + Bv\|}{\|v\|}$$

use ~~max~~

$$\frac{\|Av + Bv\|}{\|v\|} \le \frac{\|Av\| + \|Bv\|}{\|v\|}$$

By the defn (Important property of induced norm)

$$\|A\| \geq \frac{\|Aw\|}{\|w\|} \quad \forall\, w \neq 0$$

$$\Rightarrow \|Aw\| \leq \|A\| \cdot \|w\|$$

i.e. the induced matrix norm is "compatible" w. the vector norm.

Equivalent definition $\quad \|A\|_* = \max_{\|v\|=1} \|Av\|$

Pf: $\|A\| = \max_{v \neq 0} \frac{\|Av\|}{\|v\|} \geq \max_{\|v\|=1} \frac{\|Av\|}{\|v\|} = \|A\|_*$

For $v \neq 0$

$$\frac{\|Av\|}{\|v\|} = \left\|A \frac{v}{\|v\|}\right\| \leq \max_{\|w\|=1} \|Aw\| \quad \forall\, v$$

$$\Rightarrow \cancel{\max_{v \neq 0}} \; \|A\| = \max_{v \neq 0} \frac{\|Av\|}{\|v\|} \leq \max_{\|w\|=1} \|Aw\| = \|A\|_*$$

i.e. $\|A\| = \|A\|_*$ ✓

Some more explicit statements

$A \in \mathbb{R}^{n\times n}$

(1) $\|A\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |a_{ij}|$ i.e. the "biggest row sum"

(2) $\|A\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |a_{ij}|$ the "biggest column sum"

(3) $\|A\|_2 = \max\limits_{1\le i\le n} \sqrt{\lambda_i}$ where λ_i are the ~~non~~ eigenvalues (all non-negative) of A^TA.

If $A^T = A \Rightarrow A^TA = A^2 \Rightarrow \lambda_i = \sigma_i^2$

$\|A\|_2 = \max |\sigma_i|$