Lecture 17

Last time we did

- Cholesky's factorization of a positive definite matrix A

$$A = LL^t$$ where L is lower triangular.

This requires only 50% of the calculation needed for LU decomposition.

- If GE has row changes then we can factor A into

$$PA = LU$$ where P is a permutation matrix.

we solve $Ax = b$

by $PAx = Pb = b'$

$LUx = b'$

set $y = Ux$

solve $Ly = b'$

and then solve $Ux = y$

Matrix factorization is useful if we have to solve $Ax = b$ for many different values of b

we then turned our attention to round-off errors that can occur while doing GE.

we learn scaled partial pivoting for GE.

Today we study Matrix Norms

We first study norms in $\mathbb{R}^n$

What is a norm?
We want to measure "size" of a vector $\bar{a}$ in $\mathbb{R}^n$

Defn a norm assign to each vector $\bar{a}$ in $\mathbb{R}^n$ a number $\|\bar{a}\|$, called the norm of $\bar{a}$ subject to the following restriction

1) $\|\bar{a}\| \geq 0$ for all $\bar{a} \in \mathbb{R}^n$ and
   $\|\bar{a}\| = 0$ if and only if $\bar{a} = 0$

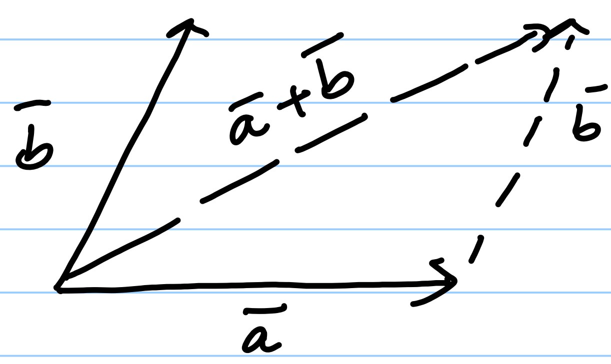
2) $\|r\bar{a}\| = |r| \, \|\bar{a}\|$ for all $\bar{a} \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

3) (triangle inequality)
   $\|\bar{a} + \bar{b}\| \leq \|\bar{a}\| + \|\bar{b}\|$ for all $\bar{a}, \bar{b} \in \mathbb{R}^n$

Note the first restriction forces all n-vectors but the zero vector to have positive "length".

The second restriction states for example that $\bar{a}$ and its negative $-\bar{a}$ have the same length and that the $3\bar{a}$ has length 3 times that of $\bar{a}$

The third restriction is called triangle inequality



Examples of norms

1) <u>Euclidean norm</u>

$\bar{a} = (a_1, a_2, \ldots, a_n)$

$$\|\bar{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

Euclidean norm is also called the 2-norm

Easy to verify
that $\|\bar{a}\|_2 \geq 0$ and $= 0$ iff $\bar{a} = \bar{0}$

Also $\|r\bar{a}\| = |r|\,\|\bar{a}\|$ is clear

To prove triangle inequality we need the "Cauchy - Schwarz inequality

<u>Theorem</u> For $\bar{x}, \bar{y} \in \mathbb{R}^n$

$$\bar{x} \cdot \bar{y} \leq \|x\|_2 \|y\|_2$$

Pf $y = \bar{0}$ or $x = \bar{0}$ then nothing to prove

So suppose $\bar{x} \neq 0$ and $\bar{y} \neq 0$

For $\lambda \in \mathbb{R}$

$$0 \leq \|x - \lambda y\|_2^2 = \sum_{i=1}^{n} (x_i - \lambda y_i)^2$$

$$= \sum_{i=1}^{n} x_i^2 - 2\lambda \sum_{i=1}^{n} x_i y_i + \lambda^2 \sum_{i=1}^{n} y_i^2$$

$$2\lambda \sum_{i=1}^{n} x_i y_i \leq \|x\|_2^2 + \lambda^2 \|y\|_2^2$$

Since $\|x\|_2 > 0$ and $\|y\|_2 > 0$

take $\lambda = \dfrac{\|x\|_2}{\|y\|_2}$

$$\left(2\frac{\|x\|_2}{\|y\|_2}\right)\left(\sum_{i=1}^{n} x_i y_i\right) \leq \|x\|_2^2 + \frac{\|x\|_2^2}{\|y\|^2}\|y\|^2$$

$$= 2\|x_2\|^2$$

$$\therefore \sum_{i=1}^{n} x_i y_i \leq \|x\|_2 \|y\|_2$$

We now prove triangle inequality

$$\|x+y\|_2^2 = \sum_{i=1}^{n} (x_i + y_i)^2$$

$$= \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$$

$$\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2$$

$$= (\|x\|_2 + \|y\|_2)^2$$

Thus $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$

1-norm

$\bar{a} = (a_1, a_2, -, a_n)$

$\|\bar{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n|$

It is easy to see that

$\|\bar{a}\|_1 \geq 0$ and $= 0$ iff $\bar{a} = 0$

also $\|r\bar{a}\|_1 = |r| \, \|\bar{a}\|_1$

$$\begin{aligned} \|a+b\|_1 &= |a_1+b_1| + |a_2+b_2| + \cdots + |a_n+b_n| \\ &\leq |a_1| + |b_1| + |a_2| + |b_2| + \cdots + |a_n| + |b_n| \\ &= \|a\|_1 + \|b\|_1 \end{aligned}$$

Thus $\|\bar{a}+\bar{b}\|_1 \leq \|\bar{a}\|_1 + \|\bar{b}\|_1$

∞-norm

$\bar{a} = (a_1, \ldots, a_n)$

$$\|\bar{a}\|_\infty = \max_{1 \le i \le n} |a_i|$$

One can easily verify that $\|\bar{a}\|_\infty$ is also a norm on $\mathbb{R}^n$.

If $\| \ \|$ is a norm in $\mathbb{R}^n$ then we can define distance between two vectors $\bar{x}, \bar{y}$ as

$$\text{dist}(\bar{x}, \bar{y}) = \|\bar{x} - \bar{y}\|$$

Matrix norm

A matrix norm on the set of all $n \times n$ matrices is a real valued function, $\| \ \|$ defined on this set satisfying for all $n \times n$ matrices A, B and all real numbers α

(i) $\|A\| \geq 0$

$\|A\| = 0$ iff $A = 0$

(ii) $\|\alpha A\| = |\alpha| \, \|A\|$

(iii) $\|A+B\| \leq \|A\| + \|B\|$

(iv) $\|AB\| \leq \|A\| \, \|B\|$

Theorem If $\| \ \|$ is a norm in $\mathbb{R}^n$

then

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

is a matrix norm.

Proof Clearly $\|A\| \geq 0$

1) $\|A\| = 0 \implies \|Ax\| = 0$ for all x with $\|x\| = 1$

$\implies Ax = 0$ for all x with $\|x\| = 1$

Claim $A = 0$

otherwise some column say i is non-zero

e_i be i^{th} co-ordinate vector

$Ae_i \neq 0$

$u = \dfrac{e_i}{\|e_i\|}$ has norm 1

$$Au = \frac{1}{\|e_i\|} Ae_i \neq 0$$

a contradiction

(2) $$\|\alpha A\| = \max_{\|x\|=1} \|\alpha Ax\|$$

$$= \max_{\|x\|=1} |\alpha| \, \|Ax\|$$

$$= |\alpha| \max_{\|x\|=1} \|Ax\|$$

$$= |\alpha| \, \|A\|$$

(3) $$\|A+B\| = \max_{\|x\|=1} \|(A+B)(x)\|$$

for x with $\|x\|=1$

$$\|Ax+Bx\| \leq \|Ax\| + \|Bx\|$$

$$\leq \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\|$$

$$= \|A\| + \|B\|$$

Thus $\max_{\|x\|=1} \|Ax+Bx\| \leq \|A\| + \|B\|$

So $\|A+B\| \leq \|A\| + \|B\|$

(4) $\|AB\| = \max_{\|x\|=1} \|ABx\|$

$(AB)x = A(Bx)$

if $Bx \neq 0$ $\quad u = \dfrac{Bx}{\|Bx\|}$ has norm 1

$\left\| A\left(\dfrac{Bx}{\|Bx\|}\right) \right\| \leq \|A\|$

$\dfrac{1}{\|Bx\|} \|ABx\| \leq \|A\|$

$$\|ABx\| \leq \|A\| \, \|Bx\|$$
$$\leq \|A\| \, \|B\|$$

Thus $\|AB\| \leq \|A\| \, \|B\|$

Remark

for $\bar{z} \neq 0$ $\quad \bar{x} = \dfrac{\bar{z}}{\|\bar{z}\|}$ is a unit vector

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \max_{z \neq 0} \left\| A\left(\frac{z}{\|z\|}\right)\right\|$$

$$= \max_{z \neq 0} \frac{\|Az\|}{\|z\|}$$

Thus for any vector $\bar{z}$

$$\|Az\| \leq \|A\| \, \|z\|$$

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$$ Euclidean norm

$$\|A\|_1 = \max_{\|x\|_1 = 1} \|Ax\|_1$$ 1-norm

$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$ ∞-norm

Euclidean norm is difficult to calculate

However one can calculate both the 1-norm and the ∞-norm

Theorem $$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

Exercise $$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$$

$\|A\|_\infty$ sums rows

$\|A\|_1$ sums columns

Example

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

$$\|A\|_\infty = \max\{8, 6, 6\} = 8$$

$$\|A\|_1 = \max\{6, 7, 7\} = 7$$

Proof of Theorem

we want to show

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$$

we first show

$$\|A\|_\infty \le \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$$

Let $\bar{x} \in \mathbb{R}^n$ with $\|\bar{x}\|_\infty = 1$

$$\|\bar{x}\|_\infty = \max_{1 \le i \le n} |x_i|$$

$$\|Ax\|_\infty = \max_{1 \le i \le n} |(Ax)_i|$$

$$= \max_{1 \le i \le n} \left| \sum_{j=1}^{n} a_{ij} x_j \right|$$

$$\left|\sum_{j=1}^{n} a_{ij} x_j\right| \leq \sum_{j=1}^{n} |a_{ij}|\,|x_j| \leq \sum_{j=1}^{n} |a_{ij}|$$

since $|x_j| \leq 1$

Thus $$\|Ax\|_\infty = \max_{1\leq i\leq n} \left|\sum_{j=1}^{n} a_{ij} x_j\right|$$
$$\leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$$

Thus $$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty$$
$$\leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$$

We will now show the opposite inequality

Let p be an integer with
$$\sum_{j=1}^{n} |a_{pj}| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$$

Let $\bar{u}$ be the vector with components

$$u_j = \begin{cases} 1 & \text{if } a_{pj} \geq 0 \\ -1 & \text{if } a_{pj} < 0 \end{cases}$$

$$\|u\|_\infty = \max_{1 \leq j \leq n} |u_j| = 1$$

$$a_{pj} u_j = |a_{pj}| \quad \text{for } j = 1, 2, \ldots, n$$

$$\|Au\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} a_{ij} u_j \right|$$

$$\geq \left| \sum_{j=1}^{n} a_{pj} u_j \right| = \left| \sum_{j=1}^{n} |a_{pj}| \right|$$

$$= \sum_{j=1}^{n} |a_{pj}|$$

$$= \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$

$$\geq \|Au\|_\infty$$

$$\geq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

Thus $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$

Frobenius norm

$$A = (a_{ij})$$

$$\|A\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}$$

One can prove that the Frobenius norm is a matrix norm

Reasons for studying matrix norms

if $\hat{x}$ is computed solution of

$$Ax = b$$

then its error is the difference

$$\bar{e} = \bar{x} - \hat{x}$$

This error is unknown to us since we do not know exact sol^n $\bar{x}$

But we can always compute the "residual error"

$$\bar{r} = Ax - A\hat{x} = b - A\hat{x}$$

If $\bar{r} = \bar{0}$ then $\hat{x}$ is exact sol^n.

One would expect $\bar{r}$ to be small if

$\hat{x}$ is close to x

This is not so as the following example shows

Example

1)

$$1.01 x_1 + 0.99 x_2 = 2$$
$$0.99 x_1 + 1.01 x_2 = 2$$

unique sol^n $x_1 = x_2 = 1$

$\hat{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ has error $\bar{e} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

but residual error

$\bar{r} = \begin{bmatrix} -0.02 \\ 0.02 \end{bmatrix}$ is small

2) By taking a diff right side we can achieve the opp effect

$$1.01x_1 + 0.99x_2 = 2$$

$$1.01x_1 + 1.01x_2 = -2$$

exact answer $x_1 = 100$, $x_2 = -100$.

The approximate sol^n

$\hat{x} = \begin{bmatrix} 101 \\ -99 \end{bmatrix}$ has small error

$\bar{e} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

but residual

$\bar{r} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$ is large

Thus the residual $\bar{r} = \bar{b} - A\hat{x}$ is not always a reliable indicator of the size of the error $\bar{e} = \bar{x} - \hat{x}$.

It depends on the size of the matrix A and A^{-1}.

$$\bar{r} = Ax - A\hat{x} = A(x - \hat{x})$$
$$= A\bar{e}$$

$$\bar{e} = A^{-1}\bar{r}$$

<u>remark</u> If matrix A is invertible

$$\bar{u} = A^{-1}(Au)$$
$$\|\bar{u}\| \leq \|A^{-1}\| \, \|Au\|$$

Thus $$\frac{\|u\|}{\|A^{-1}\|} \leq \|Au\| \leq \|A\| \, \|u\|$$

applying this to $\bar{e} = A^{-1}\bar{r}$ we get

$$\frac{\|\bar{r}\|}{\|A\|} \leq \|e\| = \|A^{-1}r\| \leq \|A^{-1}\| \, \|\bar{r}\| \qquad (*)$$

This gives an upper and a lower bound

on the relative error $\frac{\|e\|}{\|x\|}$ in

terms of relative residual $\frac{\|r\|}{\|b\|}$

$$\frac{\|b\|}{\|A\|\,\|x\|} \cdot \frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|} \le \frac{\|A^{-1}\|\,\|b\|}{\|x\|} \frac{\|r\|}{\|b\|}$$

$$\frac{\|b\|}{\|A\|} \le \|x\| \le \|A^{-1}\|\,\|b\|$$

(* Applied to $\hat{x} = 0$)

we put this in above equations

to get

$$\frac{1}{\|A\|\,\|A^{-1}\|} \frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|} \le \|A^{-1}\|\,\|A\| \frac{\|r\|}{\|b\|}$$

$$\text{cond}(A) = \|A\|\,\|A^{-1}\|$$

is called condition number of A

$$\frac{1}{\text{cond}(A)} \frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \text{cond}(A) \frac{\|r\|}{\|b\|}$$

Remarks $\quad \text{cond}(A) \geq 1$

$$I = A^{-1}A$$

$$\|I\| \leq \|A^{-1}\| \, \|A\|$$

note $\quad \|I\| = \max_{\|x\|=1} \|Ix\| = \max_{\|x\|=1} \|x\|$

$$= 1$$

thus $\quad \|A^{-1}\| \, \|A\| \geq 1$