Lecture 5

CS 510

Linear systems

Want to solve

$Ax = b$

A $n \times n$ invertible

b $n \times 1$

x $n \times 1$ unknown

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & & & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$$

assume $a_{11}^{(1)} \neq 0$

$$A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & & & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$$

Set

$$m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \quad i = 2, \ldots, n$$

multiply row 1 by $-m_{i1}$ and add to row i. This gives:

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} \Bigg|$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1} a_{1j}^{(1)}$$

Set

$$m_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}, \quad i = 3, \ldots, n$$

multiply row 2 by $-m_{i2}$ and add to row i. This gives:

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & 0 & a_{33}^{(3)} \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & a_{nn}^{(2)} \end{pmatrix}$$

$$a_{ij}^{(3)} = a_{ij}^{(2)} - m_{i2}\, a_{2j}^{(2)}$$

Eventually we get

$$\begin{pmatrix} a_{11}^{(1)} & \cdots & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & & \ddots & \\ 0 & & & a_{nn}^{(n)} \end{pmatrix} = U$$

Let

$$L = \begin{pmatrix} 1 & & & 0 \\ m_{21} & 1 & & \\ m_{31} & m_{32} & 1 & \\ \vdots & \vdots & & \ddots \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & 1 \end{pmatrix}$$

Thm. If $a_{ii}^{(i)} \neq 0$, $i = 1, \ldots, n-1$

this produces L, U &

$$A = LU$$

Operation Counts:

Will count multiplication/Div

First step:

$(n-1)^2$

Second step

$(n-2)^2$ etc

overall

$$1^2 + 2^2 + \dots + (n-1)^2 = \frac{(n-1)(n)(2(n-1)+3)}{6}$$

$$\sim \frac{n^3}{3}$$

Solving $Ax = b$:

$LUx = b$

Let $y = Ux$, $Ly = b$

Solve for y, then solve for x.

These are done by forward + backward substitution respectively.

Each costs

$$1+2+\dots+(n-1) = \frac{n(n-1)}{2} \approx \frac{n}{2}$$

$\therefore$ to solve

$$Ax = b$$

Costs

$$\frac{n^3}{3} + \frac{n^2}{2} + \frac{n^2}{2} \text{ operations}$$

LU factorization with partial pivoting.

Pick pivot row to be row with largest $a_{i1}^{(1)}$ entry, $i=1,\ldots,n$

This makes absolute value of multipliers to be ≤ 1

DO this for each column.

This results in

$$LU = PA$$

where P is a permutation matrix

EX.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 2 \\ -x_1 + 2x_2 + x_3 &= 2 \\ 2x_1 - 4x_2 + x_3 &= -1 \end{aligned}$$

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & -4 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

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

$$m_{21} = -\frac{1}{2}$$

$$m_{31} = \frac{1}{2}$$

$$\begin{bmatrix} 2 & -4 & 1 \\ 0 & 0 & \frac{3}{2} \\ 0 & 1 & \frac{3}{2} \end{bmatrix}, \quad P = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -4 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & \frac{3}{2} \end{bmatrix} = U, \quad P = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -4 & 1 \\ m_{21} & 0 & \frac{3}{2} \\ m_{31} & 1 & \frac{3}{2} \end{bmatrix}$$

$$\begin{bmatrix} 2 & -4 & 1 \\ m_{31} & 1 & \frac{3}{2} \\ m_{21} & 0 & \frac{3}{2} \end{bmatrix} = U$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & \frac{7}{2} \end{bmatrix} = \begin{bmatrix} 2 & -4 & 1 \\ 1 & -1 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

To solve $Ax = b$, we solve

$$Ly = Pb = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

This gives $y_1 = -1$

$$\frac{1}{2}y_1 + y_2 = 2$$

$$y_2 = 2 + \frac{1}{2} = \frac{5}{2}$$

$$-\frac{1}{2}y_1 + y_3 = 2, \quad y_3 = \frac{3}{2}$$

Next we solve

$$Ux = y$$

$$\begin{bmatrix} 2 & -4 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{5}{2} \\ \frac{3}{2} \end{bmatrix}$$

$x_3 = 1$, $x_2 + \frac{3}{2} x_3 = \frac{5}{2}$, $x_2 = 1$, $2x_1 - 4x_2 + x_3 = -1$, $x_1 = 1$

In summary:
in LU factorization with partial pivoting we use the matrix A to store L e U and when necessary we permute rows but keep track using an array p.

Computing A^{-1}:

First compute LU

Then compute ith column of A^{-1}, say $x^{(i)}$

by solving

$$LUx^{(i)} = e^{(i)}$$

$$e^{(i)} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{bmatrix} \leftarrow 1 \text{ in } i\text{-th location}$$

Complexity:

$$\frac{n^3}{3} + n \cdot n^2$$

$$= \frac{4}{3} n^3$$

Iterative Method to Solve

$$Ax = b$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{2n} & \cdots & a_{nn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$A = D - L - U$$

$$D = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}, \quad L = \begin{bmatrix} 0 & & 0 \\ -a_{21} & \ddots & \vdots \\ \vdots & & \\ -a_{n1} & \cdots & 0 \end{bmatrix}, U = \begin{bmatrix} 0 & -a_{12} & \cdots \\ & \ddots & -a_{ij} \\ & & \ddots \\ & & 0 \end{bmatrix}$$

Jacobi Method

$$Ax = Dx - (L+U)x = b$$

so

$$Dx = (L+U)x + b$$

Now given $x^{(k)}$ as an approximation to solution of $Ax = b$, we compute $x^{(k+1)}$ from

$$D x^{(k+1)} = (L+U)x^{(k)} + b$$

or

$$x^{(k+1)} = D^{-1}(L+U)\,x^{(k)} + D^{-1}b$$

Let $T = D^{-1}(L+U)$, $c = D^{-1}b$

so

$$x^{(k+1)} = T x^{(k)} + c$$

This is Jacobi method

Clearly a_{ii} must be nonzero.

$$D^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{a_{nn}} \end{bmatrix}$$

Let's write

$$T_j = D^{-1}(L+U), \quad c_j = D^{-1}b$$

Each iteration takes $O(n^2)$ operations so if we don't need to iterate many times we gain over Gaussian method

Will it converge?

Suppose X_* is solution to

$$Ax = b$$

Then

$$X_* = T_j X_* + C \quad (1)$$

Also

$$x^{(k)} = T_j x^{(k-1)} + C \quad (2)$$

Subtracting (1) from (2)

$$x^{(k)} - X_* = T_j(x^{(k-1)} - X_*)$$

If we let

$$e^{(k)} = x^{(k)} - x^{*} \quad (\text{error})$$

then

$$e^{(k)} = T_j \, e^{(k-1)}$$

so $\| e^{(k)} \| \leq \| T_j \| \cdot \| e^{(k-1)} \|$

or

$$\|e^{(k)}\| \leq \|T_j\|^2 \cdot \|e^{(k-2)}\|$$

$$\vdots$$

$$\|e^{(k)}\| \leq \|T_j\|^k \|e^{(0)}\|$$

If $\rho(T_j) < 1$, then we can choose a matrix norm so that $\|T_j\| < 1 \Rightarrow$ convergence

We say A is diagonally dominant if

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|.$$

Thm. If A is diagonally dominant then Jacobi method converges.

Pf.

$$T_j = D^{-1}(L+U)$$

$$\|T_j\|_\infty < 1.$$

Recall $\| \ \|_\infty$ norm of a matrix is max $\| \ \|_1$ norm of its rows.

Gauss-Seidel

We start with

$$Ax = b$$

$$A = D - L - U$$

$$(D - L - U)x = b$$

We write

$$(D-L)x = Ux + b$$

iteratively:

$$(D-L)x^{(k)} = Ux^{(k-1)} + b$$

$$x^{(k)} = (D-L)^{-1}Ux^{(k-1)} + (D-L)^{-1}b$$

$$x^{(k)} = T_g x^{(k-1)} + C_g$$

$$T_g = (D-L)^{-1} U$$

$$C_g = (D-L)^{-1} b$$

As in Jacobi, if

X_* is solution to $AX=b$

$$X^{(k)} - X_* = T_g \left(X^{(k-1)} - X_*\right)$$

$$e^{(k)} = T_g e^{(k-1)} = T_g^k e^{(0)}$$

Successive Over-relaxation method

SOR

$$Ax = b$$

$$A = D - L - U$$

We have

$$Lx = (-D+U)x + b$$

Let w be any scalar

$$wLx = w(-D+U)x + wb$$

add

$$Dx = Dx$$

to both sides

we set

$$(D - \omega L)x = [(1-\omega)D + \omega U]x + \omega b$$

Write

$$(D - \omega L)x^{(k)} = [(1-\omega)D + \omega U]x^{(k-1)} + \omega b$$

Then

$$X^{(k)} = T_w X^{(k-1)} + C_w$$

where

$$T_w = (D - wL)^{-1}\left[(1-w)D + wU\right]$$

$$C_w = (D - wL)^{-1} wb$$

SOR $\therefore$ is

$$x^{(k)} = T_\omega x^{(k-1)} + C_\omega$$

We remark that to solve $(D-\omega L)^{-1} u = b$ can be done in $O(n^2)$ time.

In general convergence of SOR & the selection of ω is not so easy and may not converse. However, some results are known.

Thm (Kahan) If $a_{ii} \neq 0$ $\forall i$, then $\rho(T_w) \geq |w-1|$. So SOR cannot converge if $w \notin (0,2)$.

Thm (Ostrowski-Reich)

If A is PD (positive definite) $\omega \in (0,2)$, then SOR converges for any $x^{(0)}$.

Thm. If A is PD & tridiagonal, then $\rho(T_g) = \rho(T_j)^2 < 1$

and optimal w is

$$w = \frac{2}{1 + \sqrt{1 - \rho(T_j)^2}}$$

and with this choice of w

$$\rho(T_w) = w - 1.$$

Example:

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

Claim:

A is positive definite,

i.e. $x^T A x > 0$

$\forall x$

$$x^T A x = 4x_1^2 + 4x_2^2 + 4x_3^2 + 3x_1x_2 + 3x_2x_1 - x_2x_3 - x_2x_3$$

$$= 3(x_1 + x_2)^2 + (x_2 - x_3)^2 + x_1^2 + 3x_3^2$$

If $x \neq 0$

$x^T A x > 0$

$\therefore$ Positive definite

This is > 0 $\forall x$

So A is Positive definite.

So SOR converges

for any $w \in (0,2)$

and any initial $x^{(0)}$.

what is optimal w?

$$T_j = D^{-1}(-L-U)$$

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

$$T_j = \begin{bmatrix} 0 & -\frac{3}{4} & 0 \\ -\frac{3}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$$

What is $\rho(T_j) = ?$

We need to find eigenvalues of T

$$\det(T_j - \lambda I)$$

$$\begin{vmatrix} -\lambda & -\frac{3}{4} & 0 \\ -\frac{3}{4} & -\lambda & \frac{1}{4} \\ 0 & \frac{1}{4} & -\lambda \end{vmatrix}$$

$$= -\lambda\left(\lambda^2 - \frac{1}{16}\right)$$

so $\lambda = \frac{1}{4}$ is largest

$\therefore \rho(T_j) = \frac{1}{4}$

opt w is

$$w = \frac{2}{1+\sqrt{1-\frac{1}{4}}} = \frac{2}{1+\frac{\sqrt{3}}{2}} \approx 1.25$$

Accelerated SOR

AOR

$$X^{(k)} = T_{\sigma,\omega} X^{(k-1)} + C_{\sigma,\omega}$$

where

$$T_{\sigma,\omega} = (D-\sigma L)^{-1}\left[(1-\omega)D + (\omega-\sigma)L + \omega U\right]$$

$$c_{\sigma,\omega} = \omega(D-\sigma L)^{-1} b$$

If $\sigma = 0$, $w = 1$, Jacobi

If $\sigma = w = 1$, Gauss-Seidel

If $\sigma = w$, SOR