Lecture 23

Last time we analyzed error of Newton and Secant method.

We proved that in Newton's method

$e_{n+1} \approx c\, e_n^2$ $\qquad$ c constant

while in Secant method

$e_{n+1} \approx c\, e_n e_{n-1}$

We then defined order of convergence

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^p} = \text{non-zero constant}$$

→ For most fixed pt iteration order of convergence $= 1$

→ For Newton's method, order $= 2$

→ For secant method order $= \frac{1+\sqrt{5}}{2} = 1.618\ldots$

Today we learn iterative methods to solve linear equations.

It is based on fixed pt iteration

We want to solve

$$A\bar{x} = \bar{b}$$

It is equivalent to find zero of

$$f(x) = A\bar{x} - \bar{b}$$

We write it in equivalent form

$$\bar{x} = g(\bar{x})$$

A fixed pt ξ of $g(\bar{x})$ will be root of $f(x)$. i.e $A\bar{\xi} = \bar{b}$.

<u>Iteration</u> initial guess $x^{(0)}$

and then we generate the sequence

$$x^{(m+1)} = g(x^{(m)}) \qquad m = 0, 1, 2 \ldots$$

For convergence we require the following theorem

$S \subseteq \mathbb{R}^n$ is closed if any cgt seq $\{x_n\}$ $x_n \in S$ $\Rightarrow \lim_{n\to\infty} x_n \in S$

Example :- 1) $(0,1]$ is not closed
2) $[a,b] \times [c,d]$ is closed in $\mathbb{R}^2$

Theorem :- Suppose $g: S \to S$, S closed set in $\mathbb{R}^n$. Suppose further that g is contraction on S, i.e.,

$$\|g(x) - g(y)\| \leq K \|x-y\|$$

for all $x, y \in S$ and $0 \leq K < 1$

Then

1) g has a unique fixed pt ξ in S

2) Fixed pt iteration starting with any $x^{(0)}$ in S converges to ξ. i.e., $\lim_{m\to\infty} \|\xi - x^{(m)}\| = 0$

for such a seq $x^{(m+1)} = g(x^{(m)})$, $m = 0, 1, 2, \ldots$

More explicitly

$$(*) \quad \|\xi - x^{(m)}\| \leq \frac{K}{1-K} \|x^{(m)} - x^{(m-1)}\|$$

$$(**) \quad \|\xi - x^{(m)}\| \leq \frac{K^m}{1-K} \|x^{(1)} - x^{(0)}\|$$

S is closed set, $g: S \to S$ contractive
$\Rightarrow g$ has a fixed pt.

<u>non-trivial theorem</u>

derivation of $(*)$ and $(**)$

$$\|\xi - x^{(m)}\| = \|g(\xi) - g(x^{(m-1)})\|$$
$$\leq K \|\xi - x^{(m-1)}\|$$

by Δ^{le} ineq

$$\|\xi - x^{(m-1)}\| \leq \|\xi - x^{(m)}\| + \|x^{(m)} - x^{(m-1)}\|$$
$$\leq K\|\xi - x^{(m-1)}\| + \|x^{(m)} - x^{(m-1)}\|$$

$$(1-K)\|\xi - x^{(m-1)}\| \leq \|x^{(m)} - x^{(m-1)}\|$$

So $$\|\xi - x^{(m)}\| \leq \frac{K}{1-K} \|x^{(m)} - x^{(m-1)}\|$$

This proves $(*)$

$$\|x^{(m)} - x^{(m-1)}\| = \|g(x^{(m-1)}) - g(x^{(m-2)})\|$$
$$\leq K \|x^{(m-1)} - x^{(m-2)}\|$$

So $$\|\xi - x^{(m)}\| \leq \frac{K^2}{1-K} \|x^{(m-1)} - x^{(m-2)}\|$$
$$\leq \frac{K^m}{1-K} \|x^{(1)} - x^{(0)}\|$$

All iteration schemes to solve $Ax = b$ are based on the notion of approximate inverse,

by this we mean any matrix C for which

$$\| I - CA \| < 1$$

in some matrix norm.

Lemma If C is an approximate inverse for the matrix A, i.e., if $\| I - CA \| < 1$ in some matrix norm, then both C and A are invertible

Pf If C OR A is not invertible then so would the matrix CA.

Thus we can find $\bar{x} \neq 0$ s/t $CA\bar{x} = 0$

But then

$$
\begin{aligned}
0 \neq \|x\| &= \|(I - CA)\bar{x}\| \\
&\leq \| I - CA \| \|\bar{x}\| \\
&< \|x\|
\end{aligned}
$$

this is a contradiction

Iteration function

Let C be an approximate inverse to A

$$g(\bar{x}) = C\bar{b} + (I - CA)\bar{x}$$
$$= \bar{x} + C(\bar{b} - A\bar{x})$$

If $g(\bar{\xi}) = \bar{\xi}$

then $\bar{\xi} = \bar{\xi} + C(\bar{b} - A\bar{\xi})$

$$C(\bar{b} - A\bar{\xi}) = 0$$

C invertible. So $A\bar{\xi} = \bar{b}$

$$g(x) - g(y) = C\bar{b} + (I - CA)\bar{x} - \left(C\bar{b} + (I - CA)y\right)$$
$$= (I - CA)(\bar{x} - \bar{y})$$

So $\|g(x) - g(y)\| = \|(I - CA)(\bar{x} - \bar{y})\|$
$$\leq \|I - CA\| \, \|\bar{x} - \bar{y}\|$$

So g is contractive with $K = \|I - CA\| < 1$

$S = \mathbb{R}^n$

Therefore the fixed pt iteration

$$x^{(m+1)} = x^{(m)} + C(b - Ax^{(m)}) \quad m = 0, 1, 2, \ldots$$

starting from any $x^{(0)}$ will converge to the unique solution ξ

Two most common choices of C

$$A = (a_{ij})$$
$$= \hat{L} + \hat{D} + \hat{U}$$

$\hat{L} = (l_{ij})$ $\qquad l_{ij} = \begin{cases} a_{ij} & i > j \\ 0 & i \le j \end{cases}$

$\hat{D} = (d_{ij})$ $\qquad d_{ij} = \begin{cases} a_{ij} & \text{if } i = j \\ 0 & i \ne j \end{cases}$

$\hat{U} = (U_{ij})$ $\qquad u_{ij} = \begin{cases} 0 & \text{if } i \ge j \\ a_{ij} & \text{if } i < j \end{cases}$

<u>example</u>

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ 6 & 7 & 0 \end{pmatrix}}_{\hat{L}} + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix}}_{\hat{D}} + \underbrace{\begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{pmatrix}}_{\hat{U}}$$

we assume all diagonal entries of A are non-zero

(if this is not so at the outset, we first rearrange the equations so that this condition is satisfied. This can always be done since A is invertible)

<u>Jacobi iteration</u>

$$C = \hat{D}^{-1}.$$

This converges if $\| I - \hat{D}^{-1} A \| < 1$ for some norm.

$$x^{(m+1)} = x^{(m)} + \hat{D}^{-1} (b - A x^{(m)})$$

$$\boxed{x_i^{(m+1)} = \frac{b_i - \sum_{j \neq i} a_{ij} x_j^{(m)}}{a_{ii}}}$$

$i = 1, 2, \ldots n$

Example

$$10x_1 - x_2 = 9$$
$$-x_1 + 10x_2 - 2x_3 = 7$$
$$-2x_2 + 10x_3 = 6$$

$$\begin{cases} x_1 = 0.1\,x_2 + 0.9 \\ x_2 = 0.1\,x_1 + 0.2x_3 + 0.7 \\ x_3 = 0.2x_2 + 0.6 \end{cases}$$

Jacobi iteration

initial $x_1^{(0)} = 0$
$x_2^{(0)} = 0$
$x_3^{(0)} = 0$

$$\boxed{\begin{array}{ll} \text{iteration scheme} & x_1^{(m+1)} = 0.1\,x_2^{(m)} + 0.9 \\ & x_2^{(m+1)} = 0.1\,x_1^{(m)} + 0.2\,x_3^{(m)} + 0.7 \\ & x_3^{(m+1)} = 0.2\,x_2^{(m)} + 0.6 \end{array}}$$

Iteration

m	0	1	2	3	4
$x_1^{(m)}$	0	0.9	0.97	0.991	0.9941
$x_2^{(m)}$	0	0.7	0.91	0.941	0.9555
$x_3^{(m)}$	0	0.6	0.72	0.782	0.7564

m	5	6	7	8	9
$x_1^{(m)}$	0.99555	0.995069	0.995778	0.995115	0.995789
$x_2^{(m)}$	0.95069	0.957775	0.951151	0.957889	0.951174
$x_3^{(m)}$	0.7911	0.758220	0.791555	0.758311	0.791578

converging nicely to exact sol^n (in 6 sig digits)

$$x_1 = 0.995789$$

$$x_2 = 0.957895$$

$$x_3 = 0.791579$$

$$A = \hat{L} + \hat{D} + \hat{U}$$

for Jacobi method

$$C = \hat{D}^{-1}$$

<u>For Gauss - Seidel iteration</u>

$$C = (\hat{L} + \hat{D})^{-1}$$

$$x^{(m+1)} = x^{(m)} + (\hat{L} + \hat{D})^{-1}(b - Ax^{(m)})$$

$$(\hat{L} + \hat{D})\, x^{(m+1)} = (\hat{L} + \hat{D} - A)\, x^{(m)} + b$$

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

This gives the formulas

$$x_i^{(m+1)} = \frac{-\sum_{j<i} a_{ij}\, x_j^{(m+1)} - \sum_{j>i} a_{ij}\, x_j^{(m)} + b_i}{a_{ii}}$$

$i = 1, 2, \ldots n$

This converges if for some matrix norm

$$\| I - CA \| < 1$$

$$I - CA = I - (\hat{L}+\hat{D})^{-1}(\hat{L}+\hat{D}+\hat{U})$$

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

i.e if $\| C\hat{U} \| < 1$ then Gauss-Seidel method converges

<u>Example</u> (Same example as before

$$10x_1 - x_2 = 9$$
$$-x_1 + 10x_2 - 2x_3 = 7$$
$$-2x_2 + 10x_3 = 6$$

$$x_1 = 0.1x_2 + 0.9$$
$$x_2 = 0.1x_1 + 0.2x_3 + 0.7$$
$$x_3 = 0.2x_2 + 0.6$$

Iteration scheme

$$x_1^{(m+1)} = 0.1\, x_2^{(m)} + 0.9$$

$$x_2^{(m+1)} = \underline{0.1\, x_1}^{(m+1)} + 0.2\, x_3^{(m)} + 0.7$$

$$x_3^{(m+1)} = \underline{0.2\, x_2}^{(m+1)} + 0.6$$

(note the diff from Jacobi-method)

(in Gauss-Siedel while computing $x_i^{(m+1)}$ we can use $x_j^{(m+1)}$ for $j = 0, 1, \ldots, i-1$)

initial approx. $= (\bar{0}, \bar{0}, \bar{0})$

m	0	1	2	3	4
$x_1^{(m)}$	0	0.9	0.979	0.99495	0.995748
$x_2^{(m)}$	0	0.79	0.9495	0.957475	0.957874
$x_3^{(m)}$	0	0.758	0.7899	0.791495	0.791575

	5	6
$x_1^{(m)}$	0.995787	0.995789
$x_1^{(m)}$	0.957894	0.957895
$x_3^{(m)}$	0.791579	0.791579

Gauss-Seidel iteration is usually faster than Jacobi iteration (not always but!)

It is also easier to compute since in Jacobi iteration <u>two</u> iterates must be kept in memory while in Gauss Sieder only one vector is required to be kept in memory.

Gauss-Seidel iteration can be shown to be convergent if the coefficient matrix A is strictly row diagonally-dominant. It also converges if A is +ve-definite i.e if A is symmetric and for all non-zero vectors y, $y^t A y > 0$.

Proposition

Suppose A is strictly row-diagonally dominant. Then the Jacobi-iteration converges

Proof $A = (a_{ij})$ is strictly row diagonally dominant

So $$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \qquad i = 1, 2, \ldots n$$

$\hat{D} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ be the diagonal of A

$$\left.\begin{aligned} \|x\|_\infty &= \max_i |x_i| \\ \|B\|_\infty &= \max_{1 \le i \le n} \sum_{1 \le j \le n} |b_{ij}| \end{aligned}\right\}$$

$$\|I - \hat{D}^{-1} A\|_\infty = \max_i \left\{ 1 - \sum_{j=1}^{n} \frac{|a_{ij}|}{|a_{ii}|} \right\}$$

$$= \max_i \left\{ \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{|a_{ij}|}{|a_{ii}|} \right|$$

$$< 1$$

Thus $\hat{D}^{-1}$ is an approximate inverse to A.

Therefore the Jacobi-Method converges

Remark

Recall that C is an approximate inverse to A if in some matrix norm

$$\| I - CA \| < 1.$$

Note that in some other matrix norm it is possible that

$$\| I - CA \|' > 1$$

Example

$$B = \begin{bmatrix} 0.9 & 0.9 \\ 0 & 0 \end{bmatrix}$$

$$\| B_\infty \| = 1.8 > 1$$

$$\| B \|_1 = 0.9 < 1$$

This makes it important to find ways of telling whether $\|B\| < 1$ in some matrix norm (without having to try out all possible matrix norms)

def^n

$$\rho(B) = \text{spectral radius of } B$$
$$= \max\{\, |\lambda| \mid \lambda \text{ is an eigenvalue of } B \,\}$$

Theorem There exists, for any $\varepsilon > 0$, a vector norm for which the associated matrix norm of B satisfies

$$\|B\| \le \rho(B) + \varepsilon$$

Corollary C is an approximate inverse for A iff $\rho(I - CA) < 1$

Remark Iterative methods are usually applied to large linear systems with sparse coefficient matrix.