

Lecture 12: Recall:

Goal: Develop iterative method: find a sequence $\vec{x}_0, \vec{x}_1, \dots$ such that $\vec{x}_k \rightarrow \vec{x}^*$ = solution of $A\vec{x} = \vec{f}$ as $k \rightarrow \infty$.

Remark: We can stop when error is small enough.

Method: Splitting method

Consider a linear system $A\vec{x} = \vec{f}$ where $A \in M_{n \times n}$ (n is BIG)

Split A as follows: $A = N + (A - N) = N - \underbrace{(N - A)}_P = N - P$

Then: $A\vec{x} = \vec{f} \Leftrightarrow (N - P)\vec{x} = \vec{f} \Leftrightarrow N\vec{x} = P\vec{x} + \vec{f}$

Develop an iterative scheme as follows:

$$(*) N\vec{x}^{n+1} = P\vec{x}^n + \vec{f}$$

If $\{\vec{x}^n\}_{n=1}^{\infty}$ converges, then it converges to the sol. \vec{x}^* of $A\vec{x} = \vec{f}$

Remark:

- N must be simple enough so that $(*)$ can be solved easily
- Will $\{\vec{x}^n\}_{n=1}^{\infty}$ converge? Suitable N .

Splitting choice 1: Jacobi method

Take $N = D =$ diagonal part of A

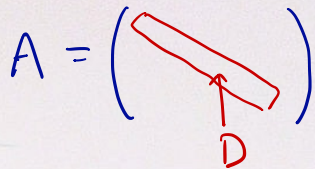
Split A as $A = D - (D - A)$

$$\begin{aligned} \text{Then: } A\vec{x} = \vec{f} &\Leftrightarrow D\vec{x} - (D - A)\vec{x} = \vec{f} \\ &\Leftrightarrow D\vec{x} = (D - A)\vec{x} + \vec{f} \end{aligned}$$

We can consider an iterative scheme such that:

$$\begin{aligned} D\vec{x}^{k+1} &= (D - A)\vec{x}^k + \vec{f} \\ \Leftrightarrow \vec{x}^{k+1} &= D^{-1}(D - A)\vec{x}^k + D^{-1}\vec{f} \end{aligned}$$

Remark: All diagonal entries of A must be non-zero, such that D is non-singular.

$$A = \begin{pmatrix} & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \end{pmatrix}$$


This is equivalent to solving: (assume $A = (a_{ij})_{1 \leq i, j \leq n}$) $\vec{x}^k = \begin{pmatrix} x_1^k \\ \vdots \\ x_n^k \end{pmatrix}$

$$\left\{ \begin{array}{l} a_{11} x_1^{k+1} + a_{12} x_2^k + a_{13} x_3^k + \dots + a_{1n} x_n^k = f_1 \quad (\text{for } x_1^{k+1}) \\ a_{21} x_1^k + a_{22} x_2^{k+1} + a_{23} x_3^k + \dots + a_{2n} x_n^k = f_2 \quad (\text{for } x_2^{k+1}) \\ \vdots \\ a_{n1} x_1^k + a_{n2} x_2^k + a_{n3} x_3^k + \dots + a_{nn} x_n^{k+1} = f_n \quad (\text{for } x_n^{k+1}) \end{array} \right.$$

Example: Consider: $\begin{pmatrix} 5 & -2 & 3 \\ -3 & 9 & 1 \\ 2 & -1 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$

Then: Jacobi method gives:

$$\vec{x}^{k+1} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -7 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 2 & -3 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{pmatrix} \vec{x}^k + \begin{pmatrix} 5 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -7 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

Start with $\vec{x}^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. The sequence almost converge in 7

iteration to get: $\vec{x}_7 = \begin{pmatrix} -0.186 \\ 0.331 \\ -0.423 \end{pmatrix}$

Question:

- Does it always converge?
- Is the initialization important?

Splitting choice 2: Gauss-Seidel method

Split A as $A = L + D + U$
 ↑ ↑ ↑
 lower tri diagonal upper tri

Develop an iterative scheme: $L\vec{x}^{k+1} + D\vec{x}^{k+1} + U\vec{x}^k = \vec{f}$

(So, take $N = L + D$ and $P = -U$)

$\Leftrightarrow \vec{x}^{k+1} = -(L+D)^{-1}U\vec{x}^k + (L+D)^{-1}\vec{f}$. (Solve lower-triangular linear system)

This is equivalent to:

$$\begin{cases} a_{11}x_1^{k+1} + a_{12}x_2^k + a_{13}x_3^k + \dots + a_{1n}x_n^k = f_1 & (\text{for } x_1^{k+1}) \\ a_{21}x_1^{k+1} + a_{22}x_2^{k+1} + a_{23}x_3^k + \dots + a_{2n}x_n^k = f_2 & (\text{for } x_2^{k+1}) \\ \vdots \\ a_{n1}x_1^{k+1} + a_{n2}x_2^{k+1} + a_{n3}x_3^{k+1} + \dots + a_{nn}x_n^{k+1} = f_n & (\text{for } x_n^{k+1}) \end{cases}$$

Remark: Again, all diagonal entries of A must be non-zero, in order that $L+D$ is non-singular.

Example 2: Continue with Example 1.

$$\vec{x}^{k+1} = -\begin{pmatrix} 5 & 0 & 0 \\ -3 & 9 & 0 \\ 2 & -1 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{x}^k + \begin{pmatrix} 5 & 0 & 0 \\ -3 & 9 & 0 \\ 2 & -1 & 7 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

Start with $\vec{x}^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. After 7 iterations, we get: $\vec{x}^7 = \begin{pmatrix} -0.186 \\ 0.331 \\ -0.423 \end{pmatrix}$

Question: Does Jacobi / Gauss-Seidel method always converge?

Example 3: Consider: $\begin{pmatrix} 1 & -5 \\ 7 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \end{pmatrix}$

Jacobi: $\vec{x}^{k+1} = \begin{pmatrix} 1 & -1 \\ -7 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 5 \\ -7 & 0 \end{pmatrix} \vec{x}^k + \begin{pmatrix} 1 & -1 \\ -7 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -4 \\ 6 \end{pmatrix}$

$\vec{x}^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \dots \vec{x}^7 = \begin{pmatrix} -2143 & 74 \\ -300127 \end{pmatrix}$ (Doesn't converge)

Gauss-Seidel also doesn't converge !!

Analysis of convergence

Let $A = N - P$

Goal: Solve $A\vec{x} = \vec{f} \Leftrightarrow (N - P)\vec{x} = \vec{f}$

Consider the iterative scheme: $N\vec{x}^{m+1} = P\vec{x}^m + \vec{f}$, $m = 0, 1, 2, \dots$

Let \vec{x}^* = sol of $A\vec{x} = \vec{f}$. Define error: $\vec{e}_m = \vec{x}^m - \vec{x}^*$, $m = 0, 1, 2, \dots$

Now, (1) $N\vec{x}^{m+1} = P\vec{x}^m + \vec{f}$

(2) $N\vec{x}^* = P\vec{x}^* + \vec{f}$

($\because A\vec{x}^* = \vec{f} \Leftrightarrow (N - P)\vec{x}^* = \vec{f}$)

(1) - (2): $N(\vec{x}^{m+1} - \vec{x}^*) = P(\vec{x}^m - \vec{x}^*)$

$\Leftrightarrow N\vec{e}^{m+1} = P\vec{e}^m \Leftrightarrow \vec{e}^{m+1} = \overbrace{N^{-1}P}^M \vec{e}^m$

Let $M = N^{-1}P$. We get: $\vec{e}^m = M^m \vec{e}^0$

Assume a simple case: let $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ be the set of linearly independent eigenvectors of M (\vec{u}_i can be complex-valued vectors)
(In other words, assume diagonalizable)

Let $\vec{e}^0 = \sum_{i=1}^n a_i \vec{u}_i$. Then:

$$\vec{e}^m = M^m \vec{e}^0 = \sum_{i=1}^n a_i M^m \vec{u}_i = \sum_{i=1}^n a_i \lambda_i^m \vec{u}_i$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are corresponding eigenvalues.

WLOG, we can assume:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

Assume $|\lambda_1| < 1$. Then:

$$\vec{e}^m = \lambda_1^m \left\{ a_1 \vec{u}_1 + \sum_{i=2}^n a_i \left(\frac{\lambda_i}{\lambda_1} \right)^m \vec{u}_i \right\} \rightarrow 0 \text{ as } m \rightarrow \infty$$

Remark: • In order to reduce error by a factor of 10^{-m} , then we need about k iterations such that $|\lambda_1|^k < 10^{-m}$

$$\text{That is, } k > \frac{m}{-\log_{10}(\rho(M))} := \frac{m}{R}$$

Here, we call $\rho(M) = |\lambda_1|$ the asymptotic convergence factor, or the spectral radius.

$\therefore \rho(M) \stackrel{\text{def}}{=} \max_k \{ |\lambda_k| : \lambda_k = \text{eigenvalue of } M \}$ is a good indicator for convergence.

• Finding $\rho(M)$ is difficult \Rightarrow Numerically (next topic)