Iterative method for Linear systems

So far we have discussed direct methods

$A = LU$ (PLU with pivoting)
$A = LL^T$
$A = QR$
$A = U\Sigma V^T$

Generally, these all cost $O(n^3)$

The challenge is how to solve $Ax = b$ when A is a very large matrix.

Maybe we can forego finding solutions in finitely many steps and focus on "approximation"

Most large systems in applications are sparse. However, if the sparsity is unstructured, direct methods do not exploit the sparsity for computation.

Iterative Method: sequence of approximations $x^{(0)}\ x^{(1)}\ x^{(2)}$

$x^{(0)} \equiv$ Initial guess
$x^{(k)} \equiv$ k-th iterate

Typically, $O(n)$ or $O(n^2)$ per iteration
useful when # of iterations $= O(1)$

Basic Idea

Formulate $Ax = b$ as $x = Bx + d$

Define $x^{(k)}$ as follows

$x^{(k)} = Bx^{(k)} + d \quad k = 1, 2, \ldots$

Goal i) Does sequence converge?
ii) How to pick $x^{(0)}$?

Stopping criterion

- $\dfrac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k)}\|} < \varepsilon$
- $\|b - Ax^{(k)}\| < \varepsilon$

Residual at the k-th iteration $\equiv b - Ax^{(k)} = r^{(k)}$

Jacobi iteration

$A \in \mathbb{R}^{m \times m}$ invertible

$Ax = b$

$D \equiv$ Diagonal matrix

$$D = \begin{pmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{mm} \end{pmatrix} \qquad d_{ii} = A_{ii}$$

Assume D is invertible.

True solution: $A^{-1}b$

Approximate solution at $(k-1)$-th iteration : $x^{(k-1)}$

$$\text{Error} = e^{(k-1)} = A^{-1}b - x^{(k-1)}$$

Assume "magically" I know the error

$$x^{(k)} = x^{(k-1)} + e^{(k-1)}$$
$$= x^{(k-1)} + A^{-1}b - x^{(k-1)}$$

Could we estimate the error?

$$Ae^{(k-1)} = AA^{-1}b - Ax^{(k-1)}$$
$$Ae^{(k-1)} = b - Ax^{(k-1)} = r^{(k-1)}$$

To find the error, we solve

$$Ae^{(k-1)} = r^{(k-1)} \qquad \text{Assume } A \approx D$$
$$De^{(k-1)} = r^{(k-1)}$$
$$e^{(k-1)} = D^{-1} r^{(k-1)}$$

$$x^{(k)} = x^{(k-1)} + e^{(k-1)}$$
$$= x^{(k-1)} + D^{-1} r^{(k-1)}$$
$$= x^{(k-1)} + D^{-1}(b - Ax^{(k-1)}) \qquad \text{Jacobi iteration}$$

Proposition If the limit $\lim_{k\to\infty} x^{(k)}$ exists, it solves $Ax = b$. That is $A \lim_{k\to\infty} x^{(k)} = b$

proof

$$\underbrace{\lim_{k\to\infty} x^{(k)}} = \underbrace{\lim_{k\to\infty} x^{(k-1)}} + D^{-1}\left(b - A \lim_{k\to\infty} x^{(k-1)}\right)$$

The same

Therefore, $D^{-1}\big(\underbrace{b - A \lim_{k\to\infty} x^{(k-1)}}_{z}\big) = 0$

$$D^{-1} z = 0 \Rightarrow z = 0$$

$\therefore \; A \lim_{k\to\infty} x^{(k-1)} = b$ as desired $\square$

Jacobi iteration computation

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

$$\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_m^{(k)} \end{bmatrix} = \begin{bmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ \vdots \\ x_m^{(k-1)} \end{bmatrix} + \begin{bmatrix} 1/a_{11} & & & \\ & 1/a_{22} & & \\ & & \ddots & \\ & & & 1/a_{nn} \end{bmatrix} \begin{bmatrix} b_1 - [A x^{(k-1)}]_1 \\ b_2 - [A x^{(k-1)}]_2 \\ \vdots \\ b_m - [A x^{(k-1)}]_m \end{bmatrix}$$

$$x_i^{(k)} = x_i^{(k-1)} + \frac{1}{a_{ii}}\left(b_i - [A x^{(k-1)}]_i\right)$$

$$x_i^{(k)} = x_i^{(k-1)} + \frac{1}{a_{ii}}\left(b_i - \sum_j a_{ij} x_j^{(k-1)}\right)$$

$$= \underbrace{x_i^{(k-1)}}_{1} + \frac{1}{a_{ii}}\left(b_i - \sum_{i\neq j} a_{ij} x_j^{(k-1)} - \underbrace{a_{ii} x_i^{(k-1)}}_{4}\right)$$

Note $1 - \frac{1}{a_{ii}} 4 = 0$

$$x_i^{(k)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j\neq i} a_{ij} x_j^{(k-1)}\right]$$

Iterative Methods $\quad x^{(k)} = Bx^{(k-1)} + d$

Jacobi: $\quad x^{(k)} = x^{(k-1)} + D^{-1}(b - Ax^{(k-1)})$

$$= \underbrace{[I - D^{-1}A]}_{B} x^{(k-1)} + \underbrace{D^{-1}b}_{d}$$

Explicit form: $\quad x_i^{(k)} = \dfrac{b_i - \sum_{j\neq i} a_{ij} x_j^{(k-1)}}{a_{ii}} \qquad 1 \le i \le m$

Exercise: What is the computational complexity of Jacobi?

For general $A \Rightarrow O(m^2)$
If A is sparse with $O(1)$ non-zero entries in each row $\Rightarrow O(m)$

Algorithm

Input: $x^{(0)} \in \mathbb{R}^m$
For $k = 1$: max_iterations
  for $i = 1:m$
    $x_i^{(k)} = \dfrac{b_i - \sum_{j\neq i} a_{ij} x_j^{(k-1)}}{a_{ii}}$

Convergence of Jacobi iteration

A is row diagonally dominant if

$$|a_{ii}| > \sum_{\substack{j=1 \\ i\neq j}}^{m} |a_{ij}| \qquad i = 1, 2, \dots m$$

Theorem: If A is row-diagonally dominant, Jacobi iteration converges for any arbitrary choice of $x^{(0)}$.

Example: Is $A = \begin{pmatrix} -3 & 1 & 0 \\ 2 & -7 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ row diagonally dominant?

(Yes!)

Numerical convergence
$$\frac{\| x^{(k)} - x^{(k-1)} \|}{\| x^{(k)} \|} < \varepsilon \qquad \varepsilon, \delta \text{ tolerances}$$
$$\| b - A x^{(k)} \| < \delta$$

Example
$$5x_1 - 2x_2 + 3x_3 = -1$$
$$-3x_1 + 9x_2 + x_3 = 2$$
$$2x_1 - x_2 - 7x_3 = 3$$

solve for x_1 $\Rightarrow$ $x_1 = \dfrac{-1 + 2x_2 - 3x_3}{5}$

solve for x_2 $\quad x_2 = \dfrac{2 + 3x_1 - x_3}{9}$

solve for x_3 $\quad x_3 = \dfrac{3 - 2x_1 + x_2}{7}$

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

First approximation
$$x_1^{(1)} = -1/5$$
$$x_2^{(1)} = 2/9$$
$$x_3^{(1)} = -3/7$$

Jacobi iteration $\quad A \approx D$
$$x^{(k)} = x^{(k-1)} + D^{-1}(b - A x^{(k-1)})$$

Decompose A as $A = L + D + U$ — upper triangular

Lower triangular; Diagonal entries are zero

$A \approx L + D$

$$x^{(k)} \approx x^{(k-1)} + (L+D)^{-1}[b - A x^{(k-1)}] \qquad \text{Gauss-Seidel iteration}$$

Assume diagonal entries of A are non-zero

$\Downarrow$

$(L+D)$ is invertible

Explicit form
$$x_i^{(k)} = \frac{b_i - \sum_{j<i} a_{ij} x_j^{(k)} - \sum_{j>i} x_j^{(k-1)}}{a_{ii}} \qquad 1 \le i \le m$$

Iterative methods: Analysis

Theorem If A is row-diagonally dominant, Gauss-Seidel converges for any choice of $x^{(0)}$.

Theorem also holds for positive definite matrix.

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

$$x^{(k)} = \underbrace{[I - (L+D)^{-1}A]}_{B} x^{(k-1)} + \underbrace{(L+D)^{-1} b}_{d}$$

Is this a good approximation of I?

Example

$$5x_1 - 2x_2 + 3x_3 = -1$$
$$-3x_1 + 9x_2 + x_3 = 2$$
$$2x_1 - x_2 - 7x_3 = 3$$

Gauss-Seidel $\quad x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$

$$x_1^{(1)} = \frac{-1 + 2x_2^{(0)} - 3x_3^{(0)}}{5} = -1/5$$

$$x_2^{(1)} = \frac{2 + 3x_1^{(1)} - x_3^{(0)}}{9} = \frac{2 - 3/5}{9} = \frac{7}{45}$$

$$x_3^{(1)} = \frac{3 - 2x_1^{(1)} + x_2^{(1)}}{7}$$

$$= \frac{3 + \frac{2}{5} + \frac{7}{45}}{7} = \frac{32}{63}$$

Repeat above until convergence

Diagonal Dominance $\Longrightarrow$ invertibility

$A x = \lambda x$

Consider an eigenvector $\bar{x} \in \mathbb{R}^m$

choose an index as follows: $i = \underset{1 \le j \le m}{\arg\max} (\bar{x}_j)$

Rescale $\bar{x}$ as follows:

$$x = \frac{x}{x_i}$$

Not x is an eigenvector

$Ax = \lambda x$

$(Ax)_i = (\lambda x)_i$

$\sum_j A_{ij} x_j = \lambda x_i$

$\Rightarrow \sum_j A_{ij} x_j = \lambda$

$\sum_{j=i} A_{ii} X_i + \sum_{\substack{j \\ i \neq j}} A_{ij} X_j = \lambda$

one term

$A_{ii} + \sum_{i \neq j} A_{ij} X_j = \lambda$

$\lambda - A_{ii} = \sum_{i \neq j} A_{ij} X_j$

$|\lambda - A_{ii}| \leq \sum_{i \neq j} |A_{ij}| \, |x_j| \leq \sum_{i \neq j} |A_{ij}| = R_i$

Every eigenvalue of A lies within atleast one of the discs $D(a_{ii}, R_i)$

Def B_K, $k = 1, 2, 3 \ldots$ matrices in $\mathbb{R}^{m \times m}$

$\lim_{k \to \infty} B_k = 0$

(All entries of B_K converge to zero)

proposition $\lim_{k \to \infty} B_k = 0$ iff $\lim_{k \to \infty} B_K u = 0 \ \forall u$

proof $\Rightarrow$ Easy $\Leftarrow$ use "good" u's

consider convergence of $M, M^2, M^3, \ldots$ what is $\lim_{k \to \infty} M^k = ?$

- $\lim_{k \to \infty} M^k = 0 \Rightarrow \lim_{k \to \infty} M^k u = 0 \quad \forall u$

Let u be an eigenvector of M with $|\lambda| > 1$

$M^k u = \lambda^k u$

$\lim_{k \to \infty} \lambda^k u \not\to 0$

- M has eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_m$

$|\lambda_i| < 1$

Now, consider $M^k x$ for any x

CASE 1 Diagonalizable

$$M^k x = c_1 \lambda_1^k u_1 + c_2 \lambda_2^k u_2 + \ldots + c_m \lambda_m^k u_m$$

$\lim_{k \to \infty} M^k x = 0$ since $|\lambda_i| < 1$

Conclusion $\lim_{k \to \infty} M^k = 0$ iff all eigenvalues are strictly smaller than 1

Exercise $M = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$ Find $\lim_{k \to \infty} M^k$

$\lambda_1 = -1$ $\lambda_2 = -2$

$\therefore \lim_{k \to \infty} M^k \neq 0$

Exercise $M = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/4 & 0 \\ -11 & 7 & 0 \end{pmatrix}$ Find $\lim_{k \to \infty} M^k$

$\lambda_1 = 1/2$; $\lambda_2 = -1/4$; $\lambda_3 = 0$

$\lim_{k \to \infty} M^k = 0$

Spectral radius $\rho(M) = \max \{ |\lambda| ; \lambda \text{ is eigenvalue of } M \}$

Need to find all eigenvalues of M

Lemma If $\|\cdot\|$ is an induced matrix norm, $\rho(A) \le \|A\|$

proof $A \in \mathbb{R}^{m \times m}$ Eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$

Eigenvectors $x_1, x_2, \ldots, x_m$

$Ax_1 = \lambda_1 x_1$ $Ax_2 = \lambda_2 x_2$ $\cdots$ $Ax_m = \lambda_m x_m$

$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & & | \end{bmatrix}$ $AX_1 = \lambda_1 X_1 \equiv Ax_1 = \lambda x_1$

$|\lambda_1| \|x\| = \|\lambda_1 x\| = \|Ax\| \le \|A\| \|x\| \rightarrow |\lambda_1| \le \|A\|$

Argue similarly for all eigenvalues to get

$$\rho(A) \le \|A\|$$

Jacobi: $x^{(k)} = x^{(k-1)} + D^{-1}(b - Ax^{(k-1)})$ ①

Exact solution:

$$Ax = b$$
$$D^{-1}Ax = D^{-1}b$$
$$D^{-1}(b - Ax) = 0$$
$$x = x + D^{-1}(b - Ax) \quad ②$$

② - ①

$$\underbrace{x - x^{(k)}}_{e^{(k)}} = \underbrace{x - x^{(k-1)}}_{e^{(k-1)}} - D^{-1}A\underbrace{(x - x^{(k-1)})}_{e^{(k-1)}}$$

$$e^{(k)} = (I - D^{-1}A)\, e^{(k-1)}$$

$$\Downarrow$$

Linear

Similarily for Gauss Seidel, we have

$$e^{(k)} = [I - (L+D)^{-1}A]\, e^{(k-1)}$$