Numerical analysis
System of linear equations

Kazuaki Tanaka

$1 = \|I\| = \|A \cdot A^{-1}\|$
$\le \|A\| \|A^{-1}\|$
$= \mathrm{cond}(A)$

1 $\boxed{\text{Condition number}}$

Definition 1.1. *The condition number* cond(A) (n×n matrix) *of nonsingular matrix A is defined by*

$$\mathrm{cond}(A) = \|A\|\|A^{-1}\|. \ge 1 \tag{1}$$

Note: Condition numbers depend on the norm imposed on spaces (e.g., l^1, l^2 or l^∞).

1.1 Impact on the stability of linear equations

$$Ax = b$$

$$A(x + \underbrace{\Delta x}_{\text{small?}}) = b + \underbrace{\Delta b}_{\text{small error}}$$

$$\|b\| = \|Ax\| \le \|A\| \|x\|$$
$$\Leftrightarrow \|x\| \ge \|A\|^{-1} \|b\| \cdots ①$$

$$A \Delta x = \Delta b \Leftrightarrow \Delta x = A^{-1} \Delta b$$
$$\Rightarrow \|\Delta x\| \le \|A^{-1}\| \|\Delta b\| \cdots ②$$

By $\frac{②}{①}$

$$\underbrace{\frac{\|\Delta x\|}{\|x\|}}_{\text{Can be large}} \le \underbrace{\|A\| \|A^{-1}\|}_{\substack{\mathrm{cond}(A) \\ \gg 1}} \underbrace{\frac{\|\Delta b\|}{\|b\|}}_{\text{small}}$$

2 Stationary Iterative methods

initial

$x^0 \to x^1 \to x^2 \to \cdots \to x^n \sim x^*$ (exact)

give　　relation　　stop

$Ax = b$

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

D　U　L

$A = D + U + L$

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

$\Leftrightarrow Dx = -(U+L)x + b$

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

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

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

Gauss-Seidel

Jacobi method

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b, \quad k = 0, 1, 2, \cdots \tag{2}$$

This can be written in the component form

$$x^k = \begin{pmatrix} x_1^k \\ \vdots \\ x_n^k \end{pmatrix}$$

$$x_i^{k+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^{n} a_{ij} x_j^k \right), \quad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots \tag{3}$$

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \qquad \begin{pmatrix} a_{ji} \end{pmatrix}$$

Gauss-Seidel method

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b, \quad k = 0, 1, 2, \cdots \tag{4}$$

This can be written in the component form

$$x_i^{k+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{k+1} - \sum_{j=i+1}^{n} a_{ij} x_j^k \right), \quad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots \tag{5}$$

2.1 Stopping criterion

For small $\varepsilon > 0$,

- $\dfrac{\|x^{k+1} - x^k\|}{\|x^k\|} < \varepsilon$ then stop
  Very easy, but it sometimes does not converge enough.
- $\dfrac{\|Ax^k - b\|}{\|b\|} < \varepsilon$ then stop　　$\varepsilon = 1e-10$
  Stable, but computational cost is higher then the above.

2.2 Convergence analysis

The spectral radius of matrices is important for analyzing convergence property of iterative methods.

Definition 2.1. *The spectral radius* $\rho(A)$ *of an* $n \times n$ *matrix* A *is defined by*

$$\rho(A) := \max_{1 \le i \le n} |\lambda_i|, \tag{6}$$

where λ_i *is the* i*-th eigenvalue of* A.

Lemma 2.2. *For any matrix norm* $\|A\|$ *induced from by the vector norm, we have*

$$\rho(A) \le \|A\|. \tag{7}$$

Proof. $Ax = \lambda x, \ x \neq 0$

Therefore

$|\lambda| \cancel{\|x\|} = \|\lambda x\| = \|Ax\| \le \|A\| \cancel{\|x\|}$

So, we have $|\lambda| \le \|A\|$ for all eigenvalues λ. □

Both Jacobi and Gauss-Seidel methods can be written in the general form

$$x^{k+1} = Mx^k + c, \quad k = 0, 1, 2, \cdots \tag{8}$$

for finding a solution of

$$x = Mx + c, \tag{9}$$

where

$$\begin{aligned} &[\text{Jacobi method}] \quad M = -D^{-1}(L+U), \quad c = D^{-1}b, \\ &[\text{Gauss-Seidel method}] \quad M = -(L+D)^{-1}U, \quad c = (L+D)^{-1}b. \end{aligned}$$

Theorem 2.3. *Equation* (9) *has a unique solution and the sequence* $\{x^k\}$ *defined by* (8) *converges to the unique solution from any starting point* x^0 **if and only if** $\rho(M) < 1$.

Proof. See, for example, Section 8.2 in [J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, 2002]. □

Corollary 2.4. *If* $\|M\| < 1$, *then the sequence* $\{x^k\}$ *defined by* (8) *converges to the unique solution of* (9) *from any starting point* x^0.

Proof. This follows from Lemma 2.2 and Theorem 2.3. □

$\rho(M) \le \|M\| < 1$

Definition 2.5. *An $n \times n$ matrix A is called* **strictly diagonally dominant** *if*

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}| \tag{10}$$

for all $i = 1, 2, \cdots, n$.

$$\begin{pmatrix} 10 & -1 & 2 \\ -2 & 10 & 3 \\ 1 & -5 & 10 \end{pmatrix} \begin{array}{l} \to \ 10 > 3 \quad OK \\ \to \ 10 > 5 \quad OK \\ \to \ 10 > 6 \quad OK \end{array}$$

$$\begin{pmatrix} \boxed{10} & \overbrace{1 \; 1 \; 0 \; 0 \; 0 \; 0 \; 0}^{n-1} \\ & 10 \\ & & \ddots \\ & & & 10 \end{pmatrix}$$

Theorem 2.6. *If A is strictly diagonally dominant, both Jacobi and Gauss-Seidel methods generate sequences $\{x^k\}$ that converge to unique solution of $Ax = b$ for any starting point x^0.*

Proof. [Jacobi] We want to prove $\|M\|_\infty = \|D^{-1}(L+U)\|_\infty < 1$.

$$M = \begin{pmatrix} 0 & & \frac{a_{ij}}{a_{ii}} \\ & \ddots & \\ \frac{a_{ij}}{a_{ii}} & & 0 \end{pmatrix}, \quad \|M\|_\infty = \max_{1\le i\le n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| = \max_{1\le i\le n} \left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \right) < 1$$

[Gauss-Seidel] Ex 6.1

Hint : proving by induction $\quad My =: z$

$$\|M\|_\infty = \|(L+D)^{-1}U\|_\infty = \max_{\|y\|_\infty = 1} \|\underbrace{(L+D)^{-1}U}_{M} y\|_\infty \quad (< 1)$$

we want

Prove, by induction, that

(i) $|z_1| = ? < 1$

(ii) Prove $|z_{i+1}| < 1$ when $|z_i| < 1$

$(x, y) = x \cdot y$

$:= x^T A y = (x, y)_A$

when $A = I$

$(x, y) = x^T y$

Remark 2.7. *In fact, if A is a symmetric positive-definite matrix ($A = A^T$ and $x^T Ax > 0$ for all $x \neq 0$), Gauss-Seidel method~~s~~ generates sequences $\{x^k\}$ that converge to unique solution of $Ax = b$ for any starting point x^0.*

$$Ax = b$$

$$\underbrace{A^T A}_{A'} x = \underbrace{A^T b}_{b'}, \quad \underbrace{A'}_{\text{symmetric positive-definite}} x = b'$$

$$\text{cond}(A') = \text{cond}(A)^2$$

3 Nonstationary iterative methods — Conjugate gradient method (CG method)

One of the top 10 algorithms in the 20th century for solving linear equation $Ax = b$, where $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^n \times \mathbb{R}^n$ is a symmetric positive-definite matrix ($A = A^T$ and $x^T Ax > 0$ for all $x \neq 0$).

Algorithm

Compute $p_0 = r_0 = b - Ax_0$.

For $k = 0, 1, 2, \cdots$

> If r_k is small enough, break.
>
> $$\alpha_k = \frac{(r_k, p_k)}{(Ap_k, p_k)} \left(= \frac{\|r_k\|_2^2}{(Ap_k, p_k)} \right)$$
>
> $$x_{k+1} = x_k + \alpha_k p_k$$
>
> $$r_{k+1} = r_k - \alpha_k A p_k$$
>
> $$\beta_k = -\frac{(Ap_k, r_{k+1})}{(Ap_k, p_k)} \left(= \frac{\|r_{k+1}\|_2^2}{\|r_k\|_2^2} \right)$$
>
> $$p_{k+1} = r_{k+1} + \beta_k p_k$$

End

Theorem 3.1. *The sequence generated by the conjugate gradient method converges to the solution of $Ax = b$ at most n steps.*

Theorem 3.2. *Let*

$$\kappa := \frac{\sqrt{\rho} - 1}{\sqrt{\rho} + 1}, \quad \rho = \mathrm{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2$$

Then, we have

exact sol. of $Ax = b$

$$\|x^* - x_k\|_2 \leq \frac{2\rho\kappa^k}{1 + \kappa^{2k}} \|x^* - x_0\|_2, \quad k = 0, 1, 2, \cdots, n.$$

3.1 Preconditioned conjugate gradient method (PCG method)

Theorem 3.3. *If A is a symmetric positive-definite matrix, then there exists a unique real lower triangular matrix L with positive diagonals such that*

$$A = LL^T. \tag{11}$$

This factorization is called **Cholesky's factorization or Cholesky's decomposition**.

$$Ax = b, \quad \mathrm{cond}(A) \gg 1 \text{ very large}$$

$$\underbrace{(L^{-1} A L^{-T})}_{A'} \underbrace{(L^{T} x)}_{x'} = \underbrace{L^{-1} b}_{b'}$$

$$A' x' = b' \quad (\Longleftrightarrow Ax = b)$$

$$A' = L^{-1} A L^{-T} = L^{-1} L L^{T} L^{-T} = I$$

$$A = LL^{T} + \underset{\text{error}}{E} \Rightarrow A' \sim I$$

Incomplete cholesky's factorization to avoid $\boxed{\text{fill-in}}$

$$A \sim LL^{T}$$