§2 Solving systems of equations

Start with linear case

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

$A \in \mathbb{R}^{n \times n}$ $\quad$ $A = (a_{ij})$

$\underline{b} \in \mathbb{R}^n$

A: invertible $\quad$ $AB = BA = I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$

$B = A^{-1}$

§2.1 Iterative method

$A\underline{x} = \underline{b}$ $\quad$ $\underline{x}^{(k)} \to \underline{x}^{(k+1)}$

Example 1:

$$\begin{cases} 3u + v = 5 & \cdots \text{ 1st eqn} \\ u + 2v = 5 & \cdots \text{ 2nd eqn} \end{cases}$$

u: 1st
v: 2nd

matrix - vector form

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

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

unknown

Start with an initial

$$\underline{x}^{(0)} = \begin{bmatrix} u^{(0)} \\ v^{(0)} \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Jacobi method

idea: Solve the ith unknown from ith equation.

$$u = \frac{5 - v}{3} \qquad v = \frac{5 - u}{2}$$

$$\begin{cases} u^{(1)} = \dfrac{5 - v^{(0)}}{3} = \dfrac{5}{3} \\[2ex] v^{(1)} = \dfrac{5 - u^{(0)}}{2} = \dfrac{5}{2} \end{cases}$$

then $\underline{x}^{(1)} = \begin{pmatrix} \frac{5}{3} \\ \frac{5}{2} \end{pmatrix}$

$$\begin{cases} u^{(2)} = \dfrac{5 - v^{(1)}}{3} = \dfrac{5 - \frac{5}{2}}{3} = \dfrac{5}{6} \\[2ex] v^{(2)} = \dfrac{5 - u^{(1)}}{2} = \dfrac{5 - \frac{5}{3}}{2} = \dfrac{5}{3} \end{cases}$$

then $\underline{x}^{(2)} = \begin{pmatrix} \frac{5}{6} \\ \frac{5}{3} \end{pmatrix}$

$\underline{x}^{(k)} \to \underline{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as $k \to \infty$

"converge"

Example 2:

$$u + 2v = 5 \qquad \text{1st eqn}$$

$$3u + v = 5 \qquad \text{2nd eqn}$$

Apply Jacobi method

with $\underline{x}^{(0)} = \begin{pmatrix} u^{(0)} \\ v^{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$u = 5 - 2v$, $\quad v = 5 - 3u$

$$\Rightarrow \underline{x}^{(1)} = \begin{pmatrix} u^{(1)} \\ v^{(1)} \end{pmatrix} = \begin{bmatrix} 5 - 2v^{(0)} \\ 5 - 3u^{(0)} \end{bmatrix}$$

$$= \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$\underline{x}^{(2)} = \begin{pmatrix} u^{(2)} \\ v^{(2)} \end{pmatrix} = \begin{bmatrix} 5 - 2v^{(1)} \\ 5 - 3u^{(1)} \end{bmatrix}$$

$$= \begin{bmatrix} 5 - 2 \cdot 5 \\ 5 - 3 \cdot 5 \end{bmatrix} = \begin{pmatrix} -5 \\ -10 \end{pmatrix}$$

$\underline{x}^{(k)}$ diverges as $k \to \infty$.

matrix- vector form

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \qquad \underline{b} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

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

Observation.

Example 1: matrix is
diagonally dominant.

a reason to make
the method
work

Definition, $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, is **strictly diagonally dominant** if for each $1 \le i \le n$,

$$|a_{ii}| > \sum_{j \ne i} |a_{ij}|$$

That is, each diagonal entry dominates its **row** in the sense that it is greater in magnitude than the sum of magnitude of the remainder of the entries in its row.

Jacobi method converges if A is strictly diagonally dominant.

"Sufficient condition"

Express Jacobi method in matrix - vector form

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \underline{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\underline{x} = \begin{pmatrix} u \\ v \end{pmatrix}$$

Scheme :

$$\begin{cases} a_{11} u^{(k)} + a_{12} v^{(k-1)} = b_1 \\ a_{21} u^{(k-1)} + a_{22} v^{(k)} = b_2 \end{cases}$$

$$\Leftrightarrow \underbrace{\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix}}_{\underline{x}^{(k)}} = - \underbrace{\begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix}}_{L+U} \underbrace{\begin{bmatrix} u^{(k-1)} \\ v^{(k-1)} \end{bmatrix}}_{\underline{x}^{(k-1)}} + \underline{b}$$

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

$$A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} + \begin{pmatrix} 0 & a_{12} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ a_{21} & 0 \end{pmatrix}$$

$$= D + U + L$$

$$A\underline{x} = \underline{b} \quad \leftarrow \cdots \text{originally} \ldots$$

$$\Leftrightarrow (D + U + L)\underline{x} = \underline{b}$$

$$\Leftrightarrow \quad D\underline{x} = \underline{b} - (U+L)\underline{x}$$

Assume D is invertible

$$\Rightarrow \quad \underline{x} = D^{-1}(\underline{b} - (U+L)\underline{x})$$

$$= G(\underline{x}) \quad \leftarrow \text{fixed point problem}$$

Fixed point iteration

$$\underline{x}(k+1) = \underline{G}(\underline{x}(k))$$

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

↳ Jacobi method is a fixed point iteration

Jacobi method (general form)

To solve $A\underline{x} = \underline{b}$ for $\underline{x} \in \mathbb{R}^n$,

where $A \in \mathbb{R}^{n\times n}$, $\underline{b} \in \mathbb{R}^n$.

and $A = D + L + U$,

diagonal (D), upper-triangle (U)

Jacobi method is

$$D\underline{x}^{(k)} = \left(\underline{b} - (L+U)\underline{x}^{(k-1)}\right).$$

or

...

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

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

when D is invertible.

Idea: to get the k-th iterate $\underline{x}^{(k)}$,

Solve the ith unknown from the ith equation, based on $\underline{x}^{(k-1)}$

Example

Apply Jacobi method to $A\underline{x} = \underline{b}$

where $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -5 & 2 \\ 1 & 6 & 8 \end{bmatrix}$

$$\underline{x} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \underline{b} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} u^{(k+1)} \\ v^{(k+1)} \\ w^{(k+1)} \end{pmatrix} = \begin{pmatrix} \dfrac{4 - v^{(k)} + w^{(k)}}{3} \\ \dfrac{1 - 2u^{(k)} - 2w^{(k)}}{-5} \\ \dfrac{1 - u^{(k)} - 6v^{(k)}}{8} \end{pmatrix}$$

Example

Apply Gauss-Seidel method to $A\underline{x} = \underline{b}$ where

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -5 & 2 \\ 1 & 6 & 8 \end{bmatrix}$$

$$\underline{x} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \underline{b} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{pmatrix} u^{(k+1)} \\ v^{(k+1)} \\ w^{(k+1)} \end{pmatrix} = \begin{bmatrix} \dfrac{4 - v^{(k)} + w^{(k)}}{3} \\ \dfrac{1 - 2u^{(k+1)} - 2w^{(k)}}{-5} \\ \dfrac{1 - u^{(k+1)} - 6v^{(k+1)}}{8} \end{bmatrix}$$

Gauss-Seidel method

idea: solve ith unknown from the ith equation, using the most recently updated values of the unknowns.

Revisit Example 1

$$\begin{cases} 3u+v=5 \\ u+2v=5 \end{cases} \qquad u=\frac{5-v}{3} \qquad v=\frac{5-u}{2}$$

$$\begin{pmatrix} u^{(0)} \\ v^{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} u^{(1)} = \dfrac{5-v^{(0)}}{3} = \dfrac{5-0}{3} = \dfrac{5}{3} \\ v^{(1)} = \dfrac{5-u^{(1)}}{2} = \dfrac{5-\frac{5}{3}}{2} = \dfrac{5}{3} \end{cases}$$

$$\begin{cases} u^{(2)} = \dfrac{5-v^{(1)}}{3} = \dfrac{5-\frac{5}{3}}{3} = \dfrac{10}{9} \\ v^{(2)} = \dfrac{5-u^{(2)}}{2} = \dfrac{5-10/9}{2} = 35/18 \end{cases}$$

matrix-vector form of GS

$$\begin{cases} a_{11} u^{(k)} + a_{12} v^{(k-1)} = b_1 \\ a_{21} u^{(k)} + a_{22} v^{(k)} = b_2 \end{cases}$$

$$\Leftrightarrow \underbrace{\begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}}_{D+L} \underbrace{\begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix}}_{\underline{x}^{(k)}}$$

$$= \underline{b} - \underbrace{\begin{bmatrix} 0 & a_{12} \\ 0 & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} u^{(k-1)} \\ v^{(k-1)} \end{bmatrix}}_{\underline{x}^{(k-1)}}$$

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

Fixed point problem

its FPI leads to GS.

$$\left\{ \begin{aligned} (D+L)\underline{x} &= \underline{b} - U\underline{x} \\ \underline{x} &= (D+L)^{-1}(\underline{b} - U\underline{x}) \\ &= G(\underline{x}) \end{aligned} \right.$$

Gauss-Seidel method

To solve $A\underline{x} = \underline{b}$ for $\underline{x} \in \mathbb{R}^n$,

where $A \in \mathbb{R}^{n\times n}$, $\underline{b} \in \mathbb{R}^n$

and $A = D + L + U$.

Gauss-Seidel method is

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

when D is invertible

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

$$\underset{-}{\uparrow}$$

Idea: to solve the ith unknown from the ith equation, using the most recently updated unknowns.

$A\underline{x} = \underline{b}$ $\quad A \in \mathbb{R}^{n\times n}$, $\underline{b} \in \mathbb{R}^n$
are given

$A = M - N$ $\quad M$: invertible

a splitting of A

$$A\underline{x} = \underline{b} \Leftrightarrow M\underline{x} = N\underline{x} + \underline{b}$$
$$\Leftrightarrow \underline{x} = M^{-1}(N\underline{x} + \underline{b})$$
$$= G(\underline{x})$$

For Jacobi method

$M = D$, $\quad N = -(L+U)$

Gauss-Seidel

$M = D+L$ $\quad N = -U$

fixed point iteration for
the related $\underline{x} = G(\underline{x})$.

General form of these two method,

problem $\underline{x} = B\underline{x} + \underline{d}$ ...... (1)

numerical method

$$\underline{x}^{(k)} = B\underline{x}^{(k-1)} + \underline{d} \quad \cdots (2)$$

Question on convergence..

$$\underline{x}^{(k)} \overset{?}{\rightarrow} \underline{x} \text{ as } k \to \infty$$

(2) - (1)

$$\underbrace{\underline{x}^{(k)} - \underline{x}} = B(\underline{x}^{(k-1)} - \underline{x})$$

$$\underline{z}^{(k)} = B \quad \underline{z}^{(k-1)}$$

$$= B^2 \underline{z}^{(k-2)}$$

$$= B^k \underline{z}^{(0)}$$

$$\underbrace{\underline{x}^{(k)} - \underline{x}} = B(\underbrace{\underline{x}^{(k-1)} - \underline{x}})$$

$$\underline{z}^{(k)} = B\,\underline{z}^{(k-1)}$$

$$\vdots$$

$$= B^k \underline{z}^{(0)} \quad \cdots\cdots (3)$$

A special case

$$B = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$B^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}$$

$$(3) \Rightarrow \begin{pmatrix} z_1^{(k)} \\ z_2^{(k)} \\ \vdots \\ z_n^{(k)} \end{pmatrix} = \begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix} \begin{pmatrix} z_1^{(0)} \\ z_2^{(0)} \\ \vdots \\ z_n^{(0)} \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} z_1^{(k)} = \lambda_1^k z_1^{(0)} \\ z_2^{(k)} = \lambda_2^k z_2^{(0)} \\ \quad\vdots \\ z_n^{(k)} = \lambda_n^k z_n^{(0)} \end{cases}$$

$z_i^{(k)} \to 0$ as $k \to \infty$

if and if only

$|\lambda_i| < 1$

That is, when $B = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.

then $\underline{x}^{(k)} = B\underline{x}^{(k)} + \underline{d}$

converges if and if only

$|\lambda_i| < 1$ for any i

In general: for general $B \in \mathbb{R}^{n \times n}$

$$\underline{x}(k) = B\,\underline{x}(k-1) + \underline{d}$$

it converges to $\underline{x}$

(the solution of

$\underline{x} = B\underline{x} + \underline{d}$)

if all eigenvalues of B are bounded by 1 in their absolute values.

---

spectral radius of B

$$= \max_{\substack{1 \le i \le n \\ \lambda_i \text{ is an eigenvalue of } B}} |\lambda_i|$$

$= \rho(B)$

convergence $\Longleftrightarrow$ $\rho(B) < 1$

Theorem. Let $B \in \mathbb{R}^{n \times n}$.

then $\underline{x}^{(k+1)} = B\,\underline{x}^{(k)} + \underline{d}$

converges to a solution of

$$\underline{x} = B\underline{x} + \underline{d}$$

for any given $\underline{d}$ and any initial $\underline{x}^{(0)}$ if and only if the spectral radius

$$\rho(B) < 1$$

$$\rho(B) = \max_{1 \le j \le n} |\lambda_j|$$

λ_j is eigenvalue of B

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

Theorem Given $A \in \mathbb{R}^{n\times n}$.

If A is strictly diagonally dominant, then

1) A is invertible

2) Jacobi and Gauss-Seidel methods applied to $A\underline{x}=\underline{b}$ converges to the unique solution for any $\underline{b} \in \mathbb{R}^n$, and with any initial $\underline{x}^{(0)}$

For the related: $\underline{x}^{(k+1)} = B\,\underline{x}^{(k)} + \underline{d}$

$\rho(B) < 1$

Review: matrix multiplication

$A, B \in \mathbb{R}^{n \times n}$

$A = (a_{ij})$

$B = (b_{ij})$

$AB = C = (c_{ij})$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

---

a special case

$$B = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$B^2 = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{pmatrix}$$

$$B^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}$$

Review: Given.

$B \in \mathbb{R}^{n \times n}$

if you can find a pair

$\lambda \in \mathbb{R}$

$0 \neq \underline{x} \in \mathbb{R}^n$ such that

$B \underline{x} = \lambda \underline{x}$

then λ is an eigenvalue of A

and $\underline{x}$ is an eigenvector