《数值分析》第六章

- Jacobi迭代与Seidel迭代
- 超松驰(SOR)迭代算法
- 迭代矩阵与收敛性分析
- 共轭梯度法

6.1迭代法基本概念

向量序列的收敛

矩阵序列的收敛

向量序列的极限

定义：设向量序列 $\left\{x^{(k)}\right\}_{k=0}^{\infty}$，$x^{(k)}=\left[x_1^{(k)},x_2^{(k)},\ldots,x_n^{(k)}\right]^T$，

若存在向量 $x=\left[x_1,x_2,\ldots,x_n\right]^T$，使得

$$\lim_{k\to\infty} x_i^{(k)} = x_i \qquad i = 1, 2, \ldots, n$$

则称向量序列 $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ 收敛到 x，记作 $\lim\limits_{k\to\infty} x^{(k)} = x$

- 每个分量组成的序列都收敛
- 称 x 是向量序列 $x^{(k)}$ 的**极限**

$$\lim_{k\to\infty} x^{(k)} = x \iff \lim_{k\to\infty} \left\|x^{(k)} - x\right\| = 0$$

矩阵序列的极限

定义： 设矩阵序列 $A_k=\left[a_{ij}^{(k)}\right]_{n\times n}$，若存在矩阵 $A=\left[a_{ij}\right]_{n\times n}$，使得

$$\lim_{k\to\infty} a_{ij}^{(k)} = a_{ij} \qquad i, j = 1, 2, \dots, n$$

则称矩阵序列 $\{A_k\}$ 收敛到 A，记作 $\lim\limits_{k\to\infty} A_k = A$

例： 设 $|a|<1$，考察下面矩阵序列的极限

$$A=\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix},\ A^2=\begin{bmatrix} a^2 & 2a \\ 0 & a^2 \end{bmatrix},\ \dots,\ A^k=\begin{bmatrix} a^k & ka^{k-1} \\ 0 & a^k \end{bmatrix},\ \dots$$

矩阵极限判断

定理： $\lim\limits_{k\to\infty} A_k = A \Leftrightarrow \lim\limits_{k\to\infty} \|A_k - A\| = 0$

(其中 $\|\cdot\|$ 为任一算子范数)

$$\lim_{k\to\infty} A_k = 0 \Leftrightarrow \lim_{k\to\infty} \|A_k\| = 0$$

$$\lim_{k\to\infty} A^k = 0 \Leftrightarrow \lim_{k\to\infty} \|A^k\| = 0$$

矩阵极限判断

定理： $\lim\limits_{k\to\infty} A_k = 0 \iff \lim\limits_{k\to\infty} A_k x = 0, \quad \forall x \in \mathrm{R}^n$

定理： $\| B \| < 1 \Rightarrow \lim\limits_{k\to\infty} B^k = 0$

矩阵极限判断

定理： $\lim\limits_{k\to\infty} B^k = 0 \Longleftrightarrow \rho(B) < 1$

定理： $\lim\limits_{k\to\infty} B^k = 0 \Longleftrightarrow$ 存在某算子范数 $\|\bullet\|_\varepsilon$，使得

$$\|B\|_\varepsilon < 1$$

定理： $\rho(B) = \lim\limits_{k\to\infty} \left\|B^k\right\|^{\frac{1}{k}}$ (其中 $\|\cdot\|$ 为任一矩阵范数)

6.1.3 线性方程组迭代法

- **直接法的缺点**
  - 运算量大，不适合大规模的线性方程组求解
  - 无法充分利用系数矩阵的稀疏性

- **迭代法**

从一个初始向量出发，按照一定的迭代格式，构造出一个趋向于真解的向量序列。

  - 只需存储系数矩阵中的非零元素
  - 运算量不超过 $O(kn^2)$，其中 k 为迭代步数

迭代法是目前求解大规模线性方程组的主要方法

迭代法的构造

- **矩阵分裂迭代法基本思想**

A 的一个矩阵分裂

$A = M - N$

$$Ax = b \quad \Longleftrightarrow \quad Mx = Nx + b$$

M 非奇异

$$\Updownarrow$$

$$x = M^{-1}Nx + M^{-1}b$$

给定一个初始向量 $x^{(0)}$，可得 迭代格式

$$\boxed{x^{(k+1)} = Bx^{(k)} + f} \qquad k = 0, 1, 2, \ldots$$

其中 $B = M^{-1}N$ 称为迭代矩阵

矩阵分裂迭代法

$$\boxed{x^{(k+1)} = Bx^{(k)} + f} \qquad k = 0, 1, 2, \ldots$$

定义：若 $\lim\limits_{k\to\infty} x^{(k)}$ 存在，则称该迭代法收敛，否则称为发散

性质：若 $\lim\limits_{k\to\infty} x^{(k)} = x_*$，则 x_* 为原方程组 $Ax = b$ 的解

$$\boxed{Ax_* = b \iff x_* = Bx_* + f}$$

注：这里用 x_* 表示真解，教材上用 x^*

迭代法的收敛性

$$x^{(k+1)} = Bx^{(k)} + f$$

定理： 对任意初始向量 $x^{(0)}$，上述迭代格式收敛的充要条件是

$$\rho(B) < 1$$

定理： 若存在算子范数 $\|\cdot\|$，使得 $\|B\| < 1$，对任意的初始向量 $x^{(0)}$，上述迭代格式收敛。

例： 考虑迭代法 $x^{(k+1)} = Bx^{(k)} + f$ 的收敛性，其中

$$B = \begin{bmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{bmatrix}$$

收敛性分析

$$Ax = b \Rightarrow (M-N)x = b \Rightarrow Mx = Nx + b$$

计算格式: $x^{(k+1)} = Bx^{(k)} + f \quad (B = M^{-1}N)$

设方程组的精确解为 x^*,则有

$$x^* = Bx^* + f \quad \Rightarrow$$

$$x^{(k+1)} - x^* = B(x^{(k)} - x^*)$$

记 $\varepsilon^{(k)} = x^{(k)} - x^* \quad (k = 0, 1, 2, 3, \cdots\cdots)$

则有 $\varepsilon^{(k+1)} = B\,\varepsilon^{(k)}$

$$\varepsilon^{(k)} = B\,\varepsilon^{(k-1)} \quad (k = 1, 2, 3, \cdots\cdots)$$

命题 若$\|B\|<1$,则迭代法 $x^{(k+1)} = B x^{(k)} + f$ 收敛

证：由$\varepsilon^{(k)} = B\,\varepsilon^{(k-1)}$,得

$$\|\varepsilon^{(k)}\| \le \|B\|\,\|\varepsilon^{(k-1)}\| \qquad (k = 1, 2, 3, \cdots\cdots)$$

$$\Rightarrow \quad \|\varepsilon^{(k)}\| \le \|B\|^{k}\,\|\varepsilon^{(0)}\|$$

$$\|B\| < 1 \quad \Rightarrow \quad \lim_{k\to\infty}\|\varepsilon^{(k)}\| \le \lim_{k\to\infty}\|B\|^{k}\,\|\varepsilon^{(0)}\| = 0$$

所以 $$\lim_{k\to\infty}\varepsilon^{(k)} = 0$$

迭代法的收敛性

$$x^{(k+1)} = Bx^{(k)} + f$$

$B = M^{-1}N$

定理： 若存在算子范数 $\|\cdot\|$，使得 $\|B\| = q < 1$，则

(1) 迭代法收敛

(2) $$\left\|x^{(k)} - x_*\right\| \le q^k \left\|x^{(0)} - x_*\right\|$$

(3) $$\left\|x^{(k)} - x_*\right\| \le \frac{q}{1-q}\left\|x^{(k)} - x^{(k-1)}\right\|$$

(4) $$\left\|x^{(k)} - x_*\right\| \le \frac{q^k}{1-q}\left\|x^{(1)} - x^{(0)}\right\|$$

收敛速度

- **迭代法的收敛速度**

第 k 步的误差： $\varepsilon^{(k)} \triangleq x^{(k)} - x_* = B^k \varepsilon^{(0)}$

$$\Rightarrow \quad \frac{\left\|\varepsilon^{(k)}\right\|}{\left\|\varepsilon^{(0)}\right\|} \le \left\|B^k\right\|$$

平均每次迭代后的误差压缩率约为： $\boxed{\left\|B^k\right\|^{\frac{1}{k}}}$

- 若要求 k 步迭代后上述误差比值不超过 σ，则

$$\left\|B^k\right\| \le \sigma \quad \Rightarrow \quad \left\|B^k\right\|^{\frac{1}{k}} \le \sigma^{\frac{1}{k}}$$

$$\Rightarrow \quad \ln\left\|B^k\right\|^{\frac{1}{k}} \le \frac{1}{k}\ln\sigma \quad \Rightarrow \quad \boxed{k \ge \frac{-\ln\sigma}{-\ln\left\|B^k\right\|^{\frac{1}{k}}}}$$

收敛速度

- **迭代法的收敛速度**

定义：迭代法 $x^{(k+1)} = Bx^{(k)} + f$ 的平均收敛速度定义为

$$R_k(B) = -\ln\left(\left\|B^k\right\|^{\frac{1}{k}}\right)$$

渐进收敛速度为

$$R(B) = \lim_{x\to\infty} R_k(B) = -\ln \rho(B) = \ln \frac{1}{\rho(B)}$$

- $\rho(B)$ 越小，迭代收敛越快

几类基本迭代法

- Jacobi 迭代法
- Gauss-Seidel 迭代法
- SOR（超松弛）迭代法
- 迭代法收敛性分析

6.2.1 Jacobi 迭代方法

- 对应的矩阵分裂
- 迭代格式（矩阵形式，分量形式）

Jacobi 迭代

- **Jacobi 迭代法**

  考虑线性方程组 $Ax = b$

  其中 $A=[a_{ij}]_{n\times n}$ 非奇异，且对角线元素全不为 0

  - 将 A 分裂成 $A = D - L - U$， 其中

$$D = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$$

$$L = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ -a_{21} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{bmatrix}, \qquad U = \begin{bmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -a_{n-1,n} \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Jacobi 迭代法

取 $M=D$，$N=L+U$，可得 雅可比 (Jacobi) 迭代方法

$$\boxed{x^{(k+1)} = D^{-1}(L+U)x^{(k)} + D^{-1}b} \qquad k=0,1,2,\ldots$$

- 迭代矩阵记为：$\boxed{J = D^{-1}(L+U)}$
- 分量形式：

$$\boxed{x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) \Big/ a_{ii}}$$

$$i=1,2,\ldots,n, \quad k=0,1,2,\ldots$$

$$
\begin{cases}
x_1^{(k+1)} = \left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)}\right)/a_{11} \\
x_2^{(k+1)} = \left(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)}\right)/a_{22} \\
\qquad\vdots \\
x_n^{(k+1)} = \left(b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)}\right)/a_{nn}
\end{cases}
$$

- **在计算 $x_i^{(k+1)}$ 时，如果用 $x_1^{(k+1)},\cdots,x_{i-1}^{(k+1)}$ 代替 $x_1^{(k)},\cdots,x_{i-1}^{(k)}$ ，则可能会得到更好的收敛效果。**

$$
\begin{cases}
x_1^{(k+1)} = \left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)}\right)/a_{11} \\
x_2^{(k+1)} = \left(b_2 - \boxed{a_{21}x_1^{(k+1)}} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)}\right)/a_{22} \\
\qquad\vdots \\
x_n^{(k+1)} = \left(b_n - \boxed{a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)}}\right)/a_{nn}
\end{cases}
$$

例 $\begin{cases} 9x_1 - x_2 - x_3 = 7 \\ -x_1 + 10x_2 - x_3 = 8 \\ -x_1 - x_2 + 15x_3 = 13 \end{cases}$

特点：系数矩阵主对角元均不为零

$$\Longleftrightarrow \begin{cases} x_1 = (7 + x_2 + x_3)/9 \\ x_2 = (8 + x_1 + x_3)/10 \\ x_3 = (13 + x_1 + x_2)/15 \end{cases}$$

取 $X^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

计算格式 $X^{(1)} = B\,X^{(0)} + f$

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1/9 & 1/9 \\ 1/10 & 0 & 1/10 \\ 1/15 & 1/15 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} 7/9 \\ 8/10 \\ 13/15 \end{bmatrix}$$

计算格式: $X^{(k+1)}=BX^{(k)}+f$

$X^{(0)}$	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	········
0	0.7778	0.9630	0.9929	0.9987	
0	0.8000	0.9644	0.9935	0.9988	
0	0.8667	0.9778	0.9952	0.9991	

准确解 ➔ X^*

1.0000
1.0000
1.0000

6.2.2Gauss-Seidel 迭代方法

- 对应的矩阵分裂
- 迭代格式（矩阵形式，分量形式）

Gauss-Seidel 迭代

写成矩阵形式:

$$x^{(k+1)} = D^{-1}\left(b + Lx^{(k+1)} + Ux^{(k)}\right)$$

可得

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

$k = 0, 1, 2, \ldots$

此迭代方法称为 高斯-塞德尔 (Gauss-Seidel) 迭代法

- 迭代矩阵记为: $\boxed{G = \left(D-L\right)^{-1} U}$

例 $\begin{cases} 9x_1 - x_2 - x_3 = 7 \\ -x_1 + 10x_2 - x_3 = 8 \\ -x_1 - x_2 + 15x_3 = 13 \end{cases} \Leftrightarrow \begin{cases} x_1 = (7 + x_2 + x_3)/9 \\ x_2 = (8 + x_1 + x_3)/10 \\ x_3 = (13 + x_1 + x_2)/15 \end{cases}$

$$x_1^{(k+1)} = (7 + x_2^{(k)} + x_3^{(k)})/9$$

$$x_2^{(k+1)} = (8 + x_1^{(k+1)} + x_3^{(k)})/10$$

$$x_3^{(k+1)} = (13 + x_1^{(k+1)} + x_2^{(k+1)})/15$$

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1/10 & 1 & 0 \\ -1/15 & -1/15 & 1 \end{bmatrix} \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{bmatrix} = \begin{bmatrix} 0 & 1/9 & 1/9 \\ 0 & 0 & 1/10 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix} + \begin{bmatrix} 7/9 \\ 8/10 \\ 13/15 \end{bmatrix}$$

迭代法解线性方程组

$$\begin{cases} 9x_1 - x_2 - x_3 = 7 \\ -x_1 + 10x_2 - x_3 = 8 \\ -x_1 - x_2 + 15x_3 = 13 \end{cases}$$

雅可比迭代法实验数据

0.7778	0.8000	0.8667
0.9630	0.9644	0.9719
0.9929	0.9935	0.9952
0.9987	0.9988	0.9991
0.9998	0.9998	0.9998
1.0000	1.0000	1.000

赛德尔迭代法实验数据

0.7778	0.8778	0.9770
0.9839	0.9961	0.9987
0.9994	0.9998	0.9999
1.0000	1.0000	1.0000
1.0000	1.0000	1.0000

6.3.1 SOR 迭代

G-S 迭代法的分量形式

$$x_i^{(k+1)} = \left(b_i - a_{i1}x_1^{(k+1)} - \cdots - a_{i,i-1}x_{i-1}^{(k+1)} - a_{i,i+1}x_{i+1}^{(k)} - \cdots - a_{i,n}x_n^{(k)} \right) \Big/ a_{ii}$$

$$= x_i^{(k)} + \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)} \right) \Big/ a_{ii}$$

- 为了得到更好的收敛效果，可在修正项前乘以一个 **松弛因子** ω，于是可得迭代格式

$$x_i^{(k+1)} = x_i^{(k)} + \omega \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)} \right) \Big/ a_{ii}$$

6.3.1 SOR 迭代方法

- 迭代格式（矩阵形式，分量形式）
- 迭代矩阵

SOR 迭代

写成矩阵形式:

$$x^{(k+1)} = x^{(k)} + \omega D^{-1}\left(b + Lx^{(k+1)} + Ux^{(k)} - Dx^{(k)}\right)$$

可得

$$\boxed{x^{(k+1)} = \left(D-\omega L\right)^{-1}\left[(1-\omega)D+\omega U\right]x^{(k)} + \omega\left(D-\omega L\right)^{-1} b}$$

—— SOR (Successive Over-Relaxation) 迭代方法

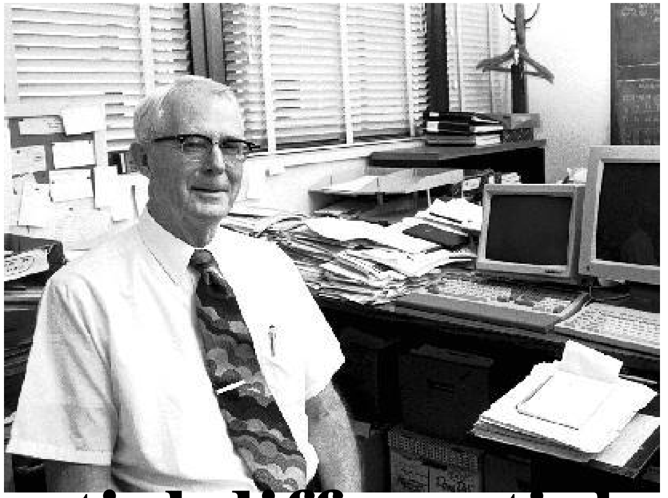
- 迭代矩阵记为: $\boxed{L_\omega = \left(D-\omega L\right)^{-1}\left[(1-\omega)D+U\right]}$

- SOR 的优点: 通过选取合适的 ω，可获得更快的收敛速度
- SOR 的缺点: 最优参数 ω 的选取比较困难

successive overrelaxation

Prof. David M. Young

1954 美国数学科学学报



Iterative methods for solving partial differential equations of elliptic

定理 若 A 是对称正定矩阵, 则当 $0<\omega<2$ 时

SOR 迭代法解方程组 $Ax=b$ 是收敛的

最佳松驰因子选取 $$\omega=\frac{2}{1+\sqrt{1-\rho(B_J)^2}}$$

$\rho(B_J)$为Jacobi迭代谱半径

- **Jacobi 迭代** $M = D,\ \ N = L + U$

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

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} \right) \Big/ a_{ii}$$

- **G-S 迭代** $M = D - L,\ \ N = U$

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

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} \right) \Big/ a_{ii}$$

- **SOR 迭代** $M = \dfrac{1}{\omega}\left(D - \omega L\right),\ \ N = \dfrac{1}{\omega}\left[(1-\omega)D + \omega U\right]$

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

$$x_i^{(k+1)} = x_i^{(k)} + \omega\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)} \right) \Big/ a_{ii}$$

举例

例： 分别用 Jacobi、G-S、SOR 迭代解线性方程组

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ -5 \end{bmatrix}$$

$$x_* = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

取初始向量 $x^{(0)} = [0, 0, 0]^T$，迭代过程中小数点后保留 4 位。

解：

Jacobi 迭代：

$$\begin{cases} x_1^{(k+1)} = \left(1 + x_2^{(k)}\right)/2 \\ x_2^{(k+1)} = \left(8 + x_1^{(k)} + x_3^{(k)}\right)/3 \\ x_3^{(k+1)} = \left(-5 + x_2^{(k)}\right)/2 \end{cases}$$

迭代可得：

$$x^{(1)} = [\ 0.5000, 2.6667, -2.5000\ ]^T$$

$$\vdots$$

$$x^{(21)} = [\ 2.0000, 3.0000, -1.0000\ ]^T$$

举例

G-S 迭代：

$$\begin{cases} x_1^{(k+1)} = \left(1 + x_2^{(k)}\right)/2 \\ x_2^{(k+1)} = \left(8 + x_1^{(k+1)} + x_3^{(k)}\right)/3 \\ x_3^{(k+1)} = \left(-5 + x_2^{(k+1)}\right)/2 \end{cases}$$

迭代可得：

$$x^{(1)} = [\ 0.5000,\ 2.8333,\ -1.0833\ ]^T$$

$$\vdots$$

$$x^{(9)} = [\ 2.0000,\ 3.0000,\ -1.0000\ ]^T$$

举例

SOR 迭代：

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \omega\left(1 - 2x_1^{(k)} + x_2^{(k)}\right)/2 \\ x_2^{(k+1)} = x_2^{(k)} + \omega\left(8 + x_1^{(k+1)} - 3x_2^{(k)} + x_3^{(k)}\right)/3 \\ x_3^{(k+1)} = x_3^{(k)} + \omega\left(-5 + x_2^{(k+1)} - 2x_3^{(k)}\right)/2 \end{cases}$$

取 $\omega = 1.1$，迭代可得

$$x^{(1)} = [\ 0.5500, 3.1350, -1.0257\ ]^T$$

$$\vdots$$

$$x^{(7)} = [\ 2.0000, 3.0000, -1.0000\ ]^T$$

如何确定 SOR 的最优松弛因子是一件非常困难的事

6.2.3 迭代方法的收敛性

- Jacobi 迭代收敛的充要条件 $\rho(J)<1$
- G-S 迭代收敛的充要条件 $\rho(G)<1$
- SOR 迭代收敛的充要条件 $\rho(L_\omega)<1$
- Jacobi 迭代收敛的充分条件 $\|J\|<1$
- G-S 迭代收敛的充分条件 $\|G\|<1$
- SOR 迭代收敛的充分条件 $\|L_\omega\|<1$

- 对角占优、不可约矩阵
- 对称正定矩阵

对角占优矩阵

定义：设 $A \in \mathrm{R}^{n\times n}$，若

$$|a_{ii}| \geq \sum_{j=1,\, j\neq i}^{n} |a_{ij}| \quad (i = 1, 2, \ldots, n)$$

且至少有一个不等式严格成立，则称 A 为 **弱对角占优**；

若所有不等式都严格成立，则称 A 为 **严格对角占优**。

命题： 如果A是严格主对角占优矩阵, 则 $\det(A) \neq 0$.

证: 用反证法。设$\det(A) = 0$, 则齐次方程组$Ax=0$有非零解 $u = [u_1, u_2, \cdots, u_n]^T$.

设 $\|u\|_\infty = |u_k|$ 考虑$Au = 0$的第k个等式

$$a_{k1}u_1 + \cdots + a_{kk}u_k + \cdots + a_{kn}u_n = 0 \quad \Rightarrow$$

$$|a_{kk}| \cdot |u_k| = \left|\sum_{\substack{j=1 \\ j\neq k}}^{n} a_{kj}u_j\right| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}u_j| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}| \cdot |u_k|$$

两边约去 $|u_k|$，得 $$|a_{kk}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}|$$

这与主对角占优矛盾, 故$\det(A) \neq 0$。

可约矩阵与不可约矩阵

定义： 设 $A\in\mathrm{R}^{n\times n}$，若存在置换矩阵 P 使得

$$P^T AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

$A_{11}\in R^{r\times r}$， $A_{22}\in R^{(n-r)\times(n-r)}$

$1\le r\le n-1$

则称 A 为 **可约矩阵**；否则称为 **不可约矩阵** 。

- 如果 A 是可约矩阵，则方程组 $Ax=b$ 等价于

$$\left(P^T AP\right)\underbrace{\boxed{\left(P^T x\right)}}_{y} = \underbrace{\boxed{P^T b}}_{f} \iff \begin{cases} A_{11}y_1 + A_{12}y_2 = f_1 \\ A_{22}y_2 = f_2 \end{cases}$$

即可以把原方程组化成两个低阶的方程组来处理。

可约矩阵的几个简单判别方法

性质： 设 $A\in\mathrm{R}^{n\times n}$，若 A 的所有元素都非零，则 A 不可约。

性质： 设 $A\in\mathrm{R}^{n\times n}$，且 $n\geq 2$。若 A 可约，则 A 的零元素个数大于等于 n-1。

性质： 设 $A\in\mathrm{R}^{n\times n}$ 是三对角矩阵，且三条对角线元素均非零，则 A 不可约。

$$A=\begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix}$$

思考： 如 A 是三对角矩阵，上、下次对角线元素均非零，则 A 是不是不可约？

Jacobi、G-S 收敛性

■ 对角占优情形

定理： 若 A 严格对角占优或不可约弱对角占优，则 A 非奇异

定理： 若 A 严格对角占优或不可约弱对角占优，则 Jacobi 迭代和 G-S 迭代均收敛

定义 $A=(a_{ij})_{n\times n}$, 如果 $$|a_{ii}|>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|$$

则称A为（行）严格对角占优阵.

定理 若$Ax=b$的系数矩阵 A 是（行）严格对角占优矩阵,则Jacobi迭代收敛.

证: 由于矩阵 A 严格对角占优

$$|a_{ii}|>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}| \quad \Rightarrow \quad \frac{1}{|a_{ii}|}\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|<1$$

而 $$B_J=D^{-1}(D-A)=\begin{bmatrix} 0 & -a_{12}/a_{11} & \cdots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & \cdots & -a_{2n}/a_{22} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \cdots & 0 \end{bmatrix}$$

故Jacobi迭代矩阵$B_J = D^{-1}(D - A)$ 第 i 行绝对值求和

$$\frac{1}{|a_{ii}|}\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| < 1 \qquad (i = 1,2,\cdots n)$$

所以 $$\|B_J\|_\infty = \max_{1\leq i\leq n}\{\frac{1}{|a_{ii}|}\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|\} < 1$$

故Jacobi迭代 $X^{(k+1)} = B_J X^{(k)} + f$ 收敛.

定理 若 A 是（行）严格主对角占优矩阵，则求解方程组$Ax=b$ 的高斯-赛德尔迭代法收敛。

证：高斯-赛德尔迭代矩阵为$(D-L)^{-1}U$, 该矩阵的特征方程为

$$|\lambda(D-L)-U|=0$$

行列式对应的矩阵为

$$C(\lambda)=\begin{bmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{bmatrix}$$

当$|\lambda|\geq 1$时，利用A矩阵的主对角占优性质，得

$$|\lambda a_{ii}| = |\lambda| \times |a_{ii}| > |\lambda| \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| = \sum_{\substack{j=1 \\ j\neq i}}^{n} |\lambda| \times |a_{ij}| \geq \sum_{j=1}^{i-1} |\lambda a_{ij}| + \sum_{j=i+1}^{n} |a_{ij}|$$

故$C(\lambda)$是严格主对角占优矩阵, 故矩阵的行列式$|C(\lambda)|$不为零，所以满足$|\lambda| \geq 1$的λ不是特征方程

$$|C(\lambda)| = |\lambda(D-L) - U| = 0$$

的根。当 A 是严格主对角占优矩阵时, $(D-L)^{-1}U$的特征值必然满足：$|\lambda| < 1$，从而高斯–赛德尔迭代矩阵谱半径小于1，迭代法收敛。

Jacobi、G-S 收敛性

■ 对称正定情形

定理： 设 A 对称且 D 正定，则 Jacobi 迭代收敛的**充要条件**是 A 和 $2D-A$ 都正定。

定理： 设 A 对称且 D 正定，则 G-S 迭代收敛的**充要条件**是 A 正定。

SOR 收敛性

- **SOR 收敛的必要条件**

> **定理：** 若 SOR 迭代收敛，则 $0 < \omega < 2$。

- **SOR 收敛的充分条件**

> **定理：** 若 A 对称正定，且 $0 < \omega < 2$，则 SOR 迭代收敛。

> **定理：** 若 A 严格对角占优或不可约弱对角占优，且 $0 < \omega \le 1$，则 SOR 迭代收敛。

收敛性小结

A 对称且对角线元素为正，则

(1) Jacobi 迭代收敛 $\Longleftrightarrow$ A 正定且 $2D\text{-}A$正定

(2) G-S 迭代收敛 $\Longleftrightarrow$ A 正定

A 对称正定，则 SOR 迭代收敛 $\Longleftrightarrow$ $0 < \omega < 2$

A 严格对角占优或不可约弱对角占优

$\Longrightarrow$ Jacobi 和 Gauss-Seidel 迭代收敛

A 严格对角占优或不可约弱对角占优且 $0 < \omega \leq 1$

$\Longrightarrow$ SOR 迭代收敛

举例

例： 设 $A=\begin{bmatrix}1 & a & a\\ a & 1 & a\\ a & a & 1\end{bmatrix}$，给出 Jacobi 和 G-S 收敛的充要条件

解： A 对称，且对角线元素均大于 0，故

(1) Jacobi 收敛的充要条件是 A 和 2D-A 均正定

(2) G-S 收敛的充要条件是 A 正定

A 正定 $\Leftrightarrow D_1=1>0, D_2=1-a^2>0, D_3=(1-a)^2(1+2a)>0$

$\Leftrightarrow -0.5<a<1$

$2D\text{-}A$ 正定 $\Leftrightarrow D_1=1>0, D_2=1-a^2>0, D_3=(1+a)^2(1-2a)>0$

$\Leftrightarrow -0.5<a<0.5$

$\Rightarrow$

Jacobi 收敛的充要条件是：$-0.5<a<0.5$

G-S 收敛的充要条件是：$-0.5<a<1$

举例

解法二： Jacobi 的迭代矩阵为 $J=\begin{bmatrix} 0 & -a & -a \\ -a & 0 & -a \\ -a & -a & 0 \end{bmatrix}$

设 λ 是 J 的特征值，则由 $\det(\lambda I - J) = 0$ 可得

$$(\lambda - a)^2 (\lambda + 2a) = 0$$

Jacobi 收敛的充要条件是 $\rho(J)<1 \Leftrightarrow |\lambda|<1$，即 $-0.5<a<0.5$

补：算法的实施

停机准则： $\dfrac{\left\|b-Ax^{(k)}\right\|}{\left\|b-Ax^{(0)}\right\|}<\varepsilon$

算法： **Jacobi** **（G-S 和 SOR 的实施类似）**

(1) 给定初值 $x^{(0)}$ 和精度要求 ε

(2) 计算 $r^{(0)}=b-Ax^{(0)}$

(3) 对 $k=0,1,2,\dots$，计算

$$x_i^{(k+1)}=\left(b_i-\sum_{j=1}^{i-1}a_{ij}x_j^{(k)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}\right)\Big/a_{ii}\quad,\quad i=1,2,\dots,n$$

若 $\dfrac{\left\|b-Ax^{(k)}\right\|}{\left\|r^{(0)}\right\|}<\varepsilon$，则输出 $x_*\approx x^{(k+1)}$，停止计算。

6.4 CG 方法

- 线性方程组与极值问题
- 最速下降法
- CG 方法（共轭梯度法）

6.4.1 线性方程组与极值问题

设 A 对称正定

定理： 向量 $x_* \in R^n$ 是 $\varphi(x)$ 的极小值当且仅当

$\varphi'(x_*)=0$ 且 Hessen 矩阵 $H_\varphi(x_*)$ 对称正定

$Ax = b$ 的解 $\Longleftrightarrow$ $\varphi(x) \triangleq \frac{1}{2}(Ax, x) - (b, x)$ 的最小值

6.4.2 梯度下降法

求解线性方程组 $Ax = b$

求 $\varphi(x)$ 的最小值点(极小值点)

基本思想:

(1) 选取一个初始向量 $x^{(0)}$

(2) 确定一个下降方向 $p^{(0)}$，从 $x^{(0)}$ 出发，沿方向 $p^{(0)}$ 找极小点 $x^{(1)}$

(3) 以 $x^{(1)}$ 点为新的出发点，确定一个新的下降方向 $p^{(1)}$，从 $x^{(1)}$ 出发，沿 $p^{(1)}$ 找极小点 $x^{(2)}$

(4) 以此类推，不断寻找新的点 $x^{(k)}$，直到找出最小值点 x_*

两个基本问题

问题一：如何确定 下降方向?

问题二：如何确定 步长?

最速下降法

基本思想：

任取一个迭代初始向量 $x^{(0)}$，构造迭代序列 $x^{(0)}, x^{(1)}, x^{(2)}, \ldots,$ 使得 $\varphi(x^{(0)}) > \varphi(x^{(1)}) > \varphi(x^{(2)}) > \ldots$，且每一步都以最快的速度下降到 $\varphi(x)$ 的极小值。

具体作法：

设 $x^{(k)}$ 已经求得，计算 $x^{(k+1)}$：$\varphi(x)$ 沿 $x^{(k)}$ 处的最速下降方向，即负梯度方向 $r^{(k)} = -(Ax^{(k)} - b)$ 的最小值点，即

$$\varphi(x^{(k+1)}) = \min_{\alpha \in R} \varphi\left(x^{(k)} + \alpha r^{(k)}\right)$$

最速下降法

计算 α 的值

$$\varphi(x^{(k)}+\alpha r^{(k)})=\frac{\alpha^2}{2}\left(Ar^{(k)},r^{(k)}\right)+\alpha\left(Ax^{(k)}-b,r^{(k)}\right)+\varphi(x^{(k)})$$

$$\Downarrow$$

$$\alpha_k=\arg\min_{\alpha\in R}\varphi\left(x^{(k)}+\alpha r^{(k)}\right)=-\frac{\left(Ax^{(k)}-b,r^{(k)}\right)}{\left(Ar^{(k)},r^{(k)}\right)}=\frac{\left(r^{(k)},r^{(k)}\right)}{\left(Ar^{(k)},r^{(k)}\right)}$$

最速下降法

算法：（最速下降法）

(1) 给定初值 $x^{(0)}$ 和精度要求 ε

(2) 计算 $r^{(0)} = b - Ax^{(0)}$

(3) 对 $k = 0, 1, 2, \ldots,$ 计算

$$\alpha_k = \left(r^{(k)}, r^{(k)}\right) \Big/ \left(Ar^{(k)}, r^{(k)}\right)$$

$$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - \alpha_k Ar^{(k)}$$

若 $\dfrac{\left\|r^{(k+1)}\right\|}{\left\|r^{(0)}\right\|} < \varepsilon$ ，则输出 $x_* \approx x^{(k+1)}$，停止计算。

最速下降法的收敛性

定理： 设 A 对称正定，其特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$

则由最速下降法产生的序列满足

$$\lim_{k\to\infty} x^{(k)} = x_* = A^{-1}b$$

且有

$$\left\| x^{(k)} - x_* \right\|_A \leq \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \right)^k \left\| x^{(0)} - x_* \right\|_A$$

缺点：

(1) 若 $\lambda_1 >> \lambda_n$ ，则收敛很慢，并可能出现不稳定现象 (舍入误差)

(2) 每次的下降方向为局部下降最快，并非全局最快

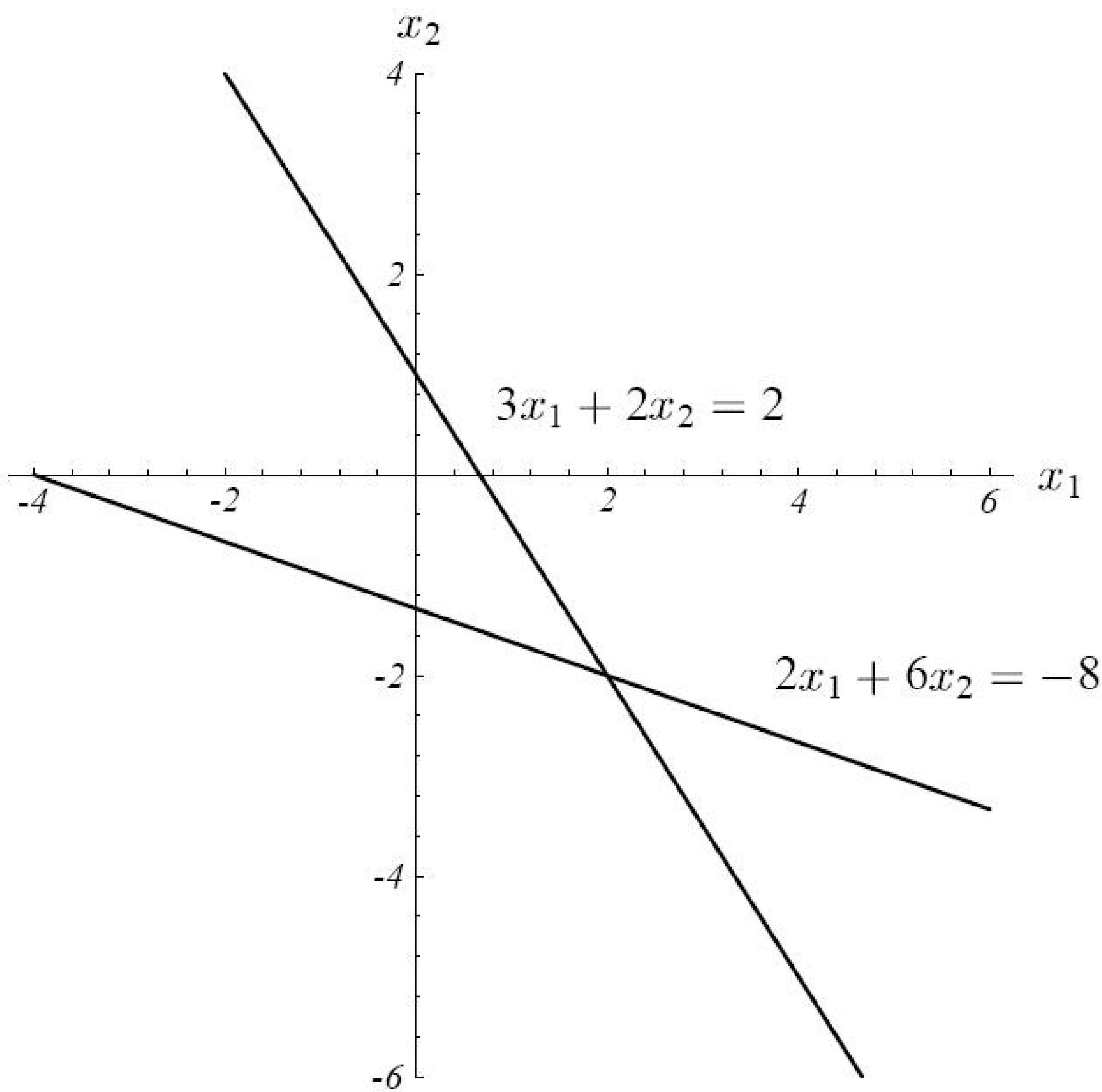
举例

$$\boxed{\begin{cases} 3x_1 + 2x_2 = 2 \\ 2x_1 + 6x_2 = -8 \end{cases}}$$

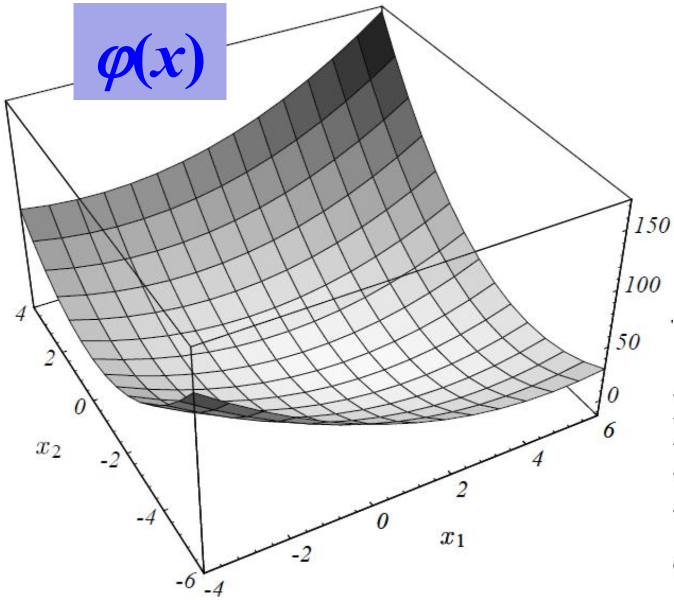
$$A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \quad \boxed{x_* = A^{-1}b = \begin{bmatrix} 2 \\ -2 \end{bmatrix}}$$

$$\boxed{\varphi(x) = \frac{1}{2}x^T Ax - b^T x = \frac{3}{2}x_1^2 + 2x_1x_2 + 3x_2^2 - 2x_1 + 8x_2}$$

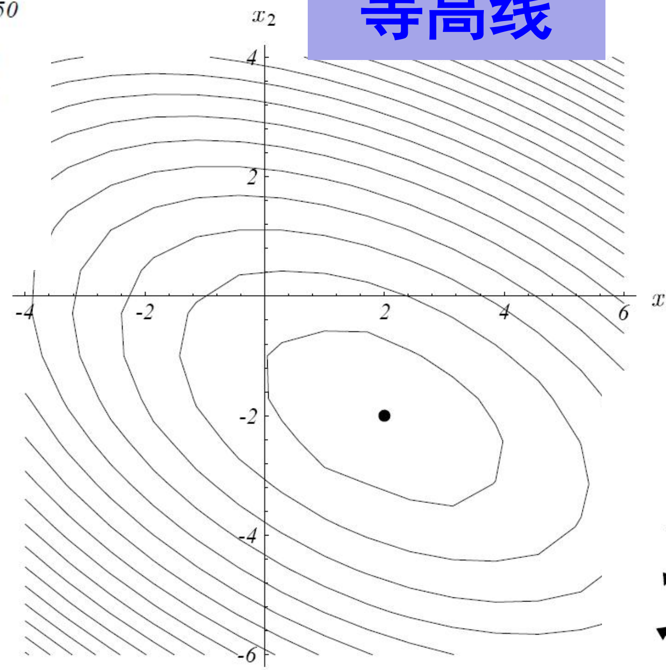
举例



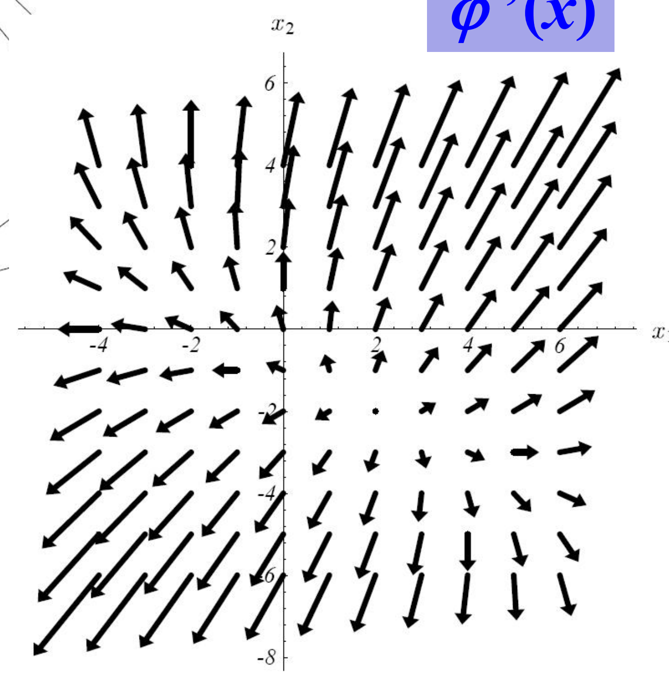
$\varphi(x)$



等高线

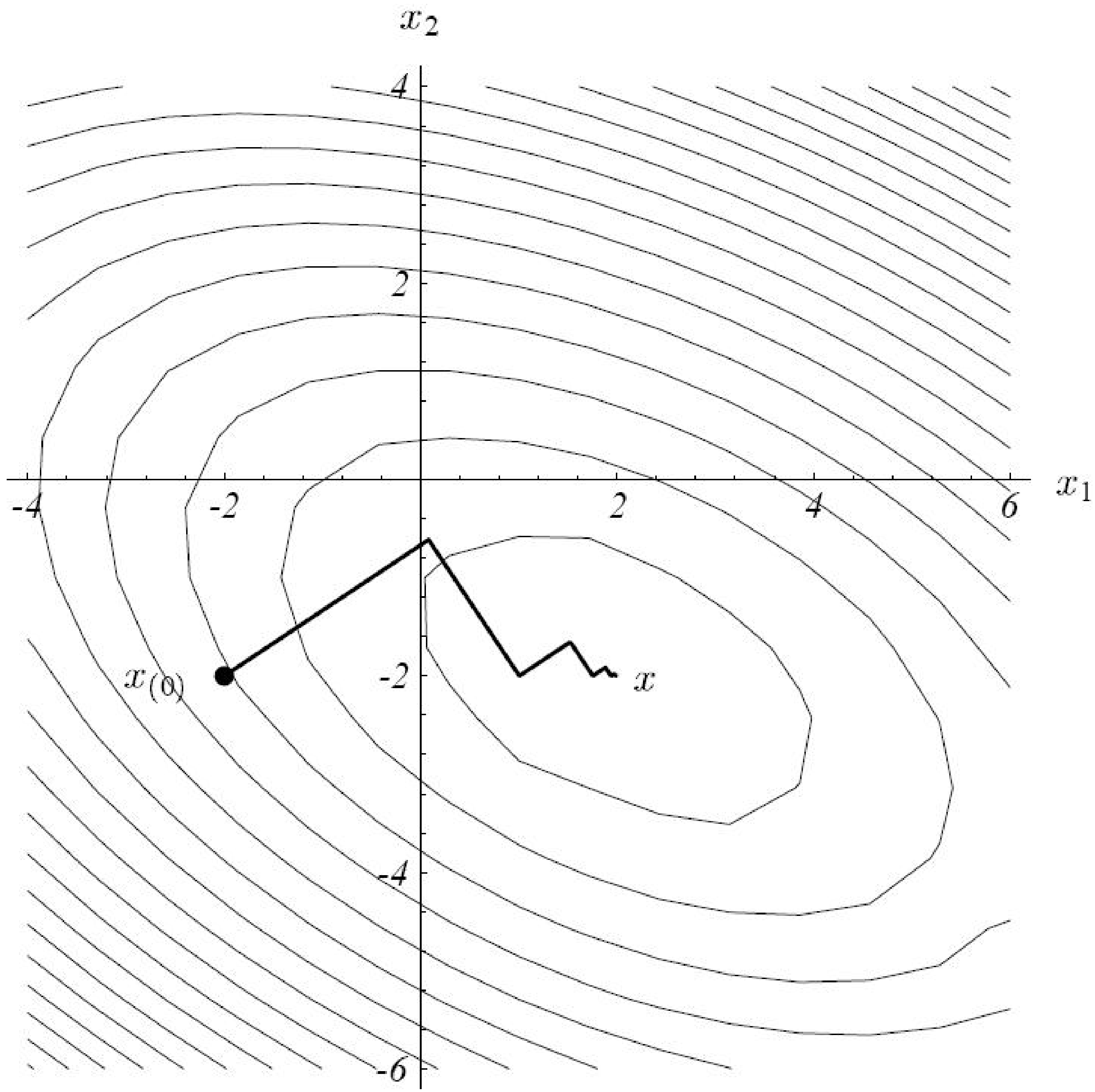


$\varphi'(x)$



举例

最速下降法



共轭梯度法

出发点:

在确定 $x^{(k+1)}$ 时，不沿负梯度方向取极小，而是寻找一个更好的方向 $p^{(k)}$，使得 $\varphi(x)$ 下降得更快！

定义: 设 A 对称正定，若 $(x, Ay) = 0$，则称 x, y 关于 A 正交 (A-正交) 或 A-共轭，若 $z_1, z_2, \ldots, z_n$ 相互 A-共轭，则称 $z_1, z_2, \ldots, z_n$ 构成 A-正交向量组或A-共轭向量组。

共轭梯度法

具体作法：

令 $p^{(0)} = r^{(0)}$， 设 $x^{(k)}$ 已经求得，则 $x^{(k+1)}$ 由下面的公式确定：

$$\boxed{x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}}$$

其中

$$p^{(k)} = r^{(k)} + \beta_{k-1} p^{(k-1)}$$

$$\beta_{k-1} = -\left(r^{(k)}, Ap^{(k-1)}\right) \Big/ \left(p^{(k-1)}, Ap^{(k-1)}\right)$$

$$\alpha_k = \left(r^{(k)}, p^{(k)}\right) \Big/ \left(p^{(k)}, Ap^{(k)}\right)$$

共轭梯度法

定理： 设 A 对称正定，则由 CG 算法产生的序列满足

(1) 当 $i \neq j$ 时，$(r^{(i)}, r^{(j)}) = 0$ ， 即 $r^{(0)}, r^{(1)}, r^{(2)}, \ldots$ 相互正交

(2) 当 $i \neq j$ 时，$(p^{(i)}, Ap^{(j)}) = 0$ ， 即 $p^{(0)}, p^{(1)}, p^{(2)}, \ldots$ 相互 A-共轭

参数的新计算公式：

$$\alpha_k = \frac{\left(r^{(k)}, r^{(k)}\right)}{\left(p^{(k)}, Ap^{(k)}\right)}, \qquad \beta_k = \frac{\left(r^{(k+1)}, r^{(k+1)}\right)}{\left(r^{(k)}, r^{(k)}\right)}$$

共轭梯度法

算法： （共轭梯度法）

(1) 给定初值 $x^{(0)}$ 和精度要求 ε

(2) 计算 $r^{(0)} = b - Ax^{(0)}$，令 $p^{(0)} = r^{(0)}$

(3) 对 $k = 0, 1, 2, \dots$，计算

$$\alpha_k = \left(r^{(k)}, r^{(k)}\right) \Big/ \left(Ap^{(k)}, p^{(k)}\right)$$

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}$$

若 $\left\|r^{(k+1)}\right\| \Big/ \left\|r^{(0)}\right\| < \varepsilon$，则输出 $x_* = x^{(k+1)}$，停止计算。

$$\beta_k = \left(r^{(k+1)}, r^{(k+1)}\right) \Big/ \left(r^{(k)}, r^{(k)}\right)$$

$$p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)}$$

共轭梯度法的收敛性

定理：设 A 对称正定，则共轭梯度法至多 n 步就能找到精确解。

定理：设 A 对称正定，x_* 为精确解， $x^{(k)}$ 为共轭梯度法的数值解，则有

$$\left\|x^{(k)}-x_*\right\|_A \le 2\left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right)^k \left\|x^{(0)}-x_*\right\|_A$$

其中 $K=\mathrm{cond}(A)_2=\left\|A\right\|_2\left\|A^{-1}\right\|_2$

总结

- Jacobi迭代与Seidel迭代
- 超松驰(SOR)迭代算法
- 迭代矩阵与收敛性分析
- 共轭梯度