线性方程组的迭代解法

- 直接解法的不足
  - 对大规模稀疏阵，由于“填入”导致巨大的计算时间、空间开销; 处理稠密阵情况更糟, O(n^3)复杂度
  - 不适合追求计算速度、允许一定准确度损失的场合
- 内容
  - 一阶定常迭代法基本理论
  - 三种经典迭代法(**Jacobi, G-S, SOR**)
  - 非固定格式迭代法(**最速下降法、共轭梯度法**)

迭代法基本概念

基本概念

- 求解线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$的迭代法
  - 通过向量序列$\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)},\cdots$逼近准确解$\boldsymbol{x}^*$
  - 类似于<u>不动点迭代法</u>的构造: $\boldsymbol{x}^{(k+1)}=\boldsymbol{g}(\boldsymbol{x}^{(k)})$

    $\boldsymbol{x}=\boldsymbol{g}(\boldsymbol{x})$为与原方程等价的方程
  - 希望每个迭代步计算量尽可能小，$\boldsymbol{g}$为线性函数:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f},\qquad k=0,1,\cdots,$$

  - $\boldsymbol{B},\boldsymbol{f}$为常数矩阵和向量，称为“<u>一阶定常迭代法</u>”
  - 具体的构造方法——矩阵分裂法

$$\boldsymbol{A}=\boldsymbol{M}-\boldsymbol{N}\ \Longrightarrow\ \boldsymbol{Mx}-\boldsymbol{Nx}=\boldsymbol{b}\Leftrightarrow\boldsymbol{x}=\boldsymbol{M}^{-1}\boldsymbol{Nx}+\boldsymbol{M}^{-1}\boldsymbol{b}$$

  - 怎么选$\boldsymbol{M}$? **1.** $\boldsymbol{M}$非奇异;　**2.**迭代法易收敛、收敛快;
    3. 求解以$\boldsymbol{M}$为系数矩阵的方程计算量小

迭代法的有关理论

- 判停准则
  - 类似“非线性方程求根”: 残差判据、误差判据
  - $\left\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)}\right\| \le \epsilon_1$并不意味着误差小, 其计算量较大
  - $\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\right\| \le \epsilon_2$ 也有缺陷, 计算简单

改造出相对量的判据

- 一阶定常迭代法的收敛性
  - $\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f}, k=0,1,\cdots,$考察误差$\boldsymbol{e}^{(k+1)}=\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^{*}$

    $\boldsymbol{x}^{*}=\boldsymbol{B}\boldsymbol{x}^{*}+\boldsymbol{f} \Longrightarrow \boldsymbol{e}^{(k+1)}=\boldsymbol{B}\boldsymbol{e}^{(k)} \Longrightarrow \boldsymbol{e}^{(k)}=\boldsymbol{B}^{k}\boldsymbol{e}^{(0)}$
  - 任意初始误差$\boldsymbol{e}^{(0)}\neq\boldsymbol{0}$，因此要保证一阶定常迭代法收敛需要迭代矩阵$\boldsymbol{B}$的幂序列$\{\boldsymbol{B}^k\}$: $\lim\limits_{k\to\infty}\boldsymbol{B}^{k}=\boldsymbol{O}$

迭代法的有关理论

- **定义4.1** $\boldsymbol{A}^{(k)} = \left(a_{ij}^{(k)}\right)$形成一矩阵序列, 若$\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}$, $(\forall i, j)$, 则称序列$\left\{\boldsymbol{A}^{(k)}\right\}$收敛于$\boldsymbol{A} = \left(a_{ij}\right)$
- **定理4.1** 矩阵算子范数的等价性: $c_1\|\boldsymbol{A}\|_\infty \leq \|\boldsymbol{A}\|_t \leq c_2\|\boldsymbol{A}\|_\infty$
- **定理4.2** $\lim\limits_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A} \Leftrightarrow \lim_{k\to\infty}\left\|\boldsymbol{A}^{(k)} - \boldsymbol{A}\right\|_t = 0$
- **定理4.3** $\lim\limits_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A} \Leftrightarrow \forall \boldsymbol{x}$都有$\lim\limits_{k\to\infty} \boldsymbol{A}^{(k)}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$
  - 一阶定常迭代法, 要$\lim\limits_{k\to\infty} \boldsymbol{B}^k = \boldsymbol{O} \Leftrightarrow \forall \boldsymbol{x}$都有$\lim\limits_{k\to\infty} \boldsymbol{B}^k\boldsymbol{x} = \mathbf{0}$

迭代法的有关理论

- 矩阵的谱半径$\rho(\boldsymbol{A}) = \max\limits_{1\le i\le n}|\lambda_i|$, λ_i为$\boldsymbol{A}$的特征值

例: $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$,

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0 \Rightarrow (\lambda - 1)[(\lambda - 1)^2 + 1] = 0$$

$$\Rightarrow \rho(\boldsymbol{A}) = \sqrt{2}$$

定理 若矩阵$\boldsymbol{A} \in \mathbb{R}^{n\times n}$，则$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$;
若$\boldsymbol{A}$为实对称矩阵，则$\|\boldsymbol{A}\|_2 = \rho(\boldsymbol{A})$

(A 的谱半径是 A 的任意一种范数的下界)。

证明 对于谱半径对应的特征值与特征向量λ和$\boldsymbol{x}$，$A\boldsymbol{x} = \lambda\boldsymbol{x}$，两边取范数并利用相容性,

$$\rho(A)\|\boldsymbol{x}\| = \|\lambda\boldsymbol{x}\| = \|A\boldsymbol{x}\| \le \|A\|\|\boldsymbol{x}\|$$

迭代法的收敛性

- **定理4.5** 设$\boldsymbol{B}=(b_{ij})\in\mathbb{R}^{n\times n}$，则$\lim\limits_{k\to\infty}\boldsymbol{B}^k=\boldsymbol{O}\Leftrightarrow\rho(\boldsymbol{B})<1$

- **定理4.6** $\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f},\ (k=0,1,\cdots)$, 且$\boldsymbol{I}-\boldsymbol{B}$非奇异, 则: 对任意初始$\boldsymbol{x}^{(0)}$迭代法收敛$\Longleftrightarrow\rho(\boldsymbol{B})<1$. 并且, 收敛值$\boldsymbol{x}^*$是方程$\boldsymbol{x}=\boldsymbol{B}\boldsymbol{x}+\boldsymbol{f}$的唯一解　　(迭代法基本定理)

证明: $\Longleftarrow$设迭代解误差$\boldsymbol{e}^{(k)}=\boldsymbol{x}^{(k)}-\boldsymbol{x}^*$, 则$\boldsymbol{e}^{(k)}=\boldsymbol{B}^k\boldsymbol{e}^{(0)}$,
由定理4.5, $\lim\limits_{k\to\infty}\boldsymbol{B}^k=\boldsymbol{O}$, 再由定理4.3, 对任意$\boldsymbol{x}^{(0)}$,
$\lim\limits_{k\to\infty}\boldsymbol{e}^{(k)}=\lim\limits_{k\to\infty}\boldsymbol{B}^k\boldsymbol{e}^{(0)}=\boldsymbol{0}$, 即迭代法收敛

$\Longrightarrow$ ($\boldsymbol{I}-\boldsymbol{B}$非奇异很重要, 即原方程解唯一)

即根据向量序列 $\{\boldsymbol{x}^{(k)}\}$ 对任意 $\boldsymbol{x}^{(0)}$ 都收敛, 证明 $\rho(\boldsymbol{B}) < 1$ 。

设 $\lim_{k\to\infty}\boldsymbol{x}^{(k)} = \boldsymbol{x}^*$, 易知 $\boldsymbol{x}^*$ 是方程 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$ 的唯一解 (由于 $\boldsymbol{I} - \boldsymbol{B}$ 非奇异), 因此对任意 $\boldsymbol{x}^{(0)}$ 及其对应的误差 $\boldsymbol{e}^{(0)}$, 都有 $\boldsymbol{e}^{(k)} = \boldsymbol{B}^k\boldsymbol{e}^{(0)} \to \boldsymbol{0}, (k \to \infty)$ 。根据定理 4.3 得出 $\lim_{k\to\infty}\boldsymbol{B}^k = \boldsymbol{O}$, 再根据定理 4.5 得 $\rho(\boldsymbol{B}) < 1$ 。

例(迭代法的收敛性): 迭代法解线性方程组的递推公式为

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f},\ (k = 0,1,2,\cdots)$$

其中, $\boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}, \boldsymbol{f} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$, 试判断该迭代法的收敛性。

【解】 首先, $\boldsymbol{I} - \boldsymbol{B}$ 为非奇异矩阵, 再看矩阵 $\boldsymbol{B}$ 的特征方程 $\det(\lambda - \boldsymbol{I}) = \lambda^2 - 6 = 0$, 因此 $\rho(\boldsymbol{B}) = \sqrt{6} > 1$ 。根据定理 4.6 知, 该迭代法对任意的初始值不一定收敛。

例 试判断以下迭代公式是否收敛：

(1) $$\begin{cases} x_1^{(k+1)} = \qquad -\dfrac{1}{3}x_2^{(k)} + \dfrac{5}{3} \\ x_2^{(k+1)} = -\dfrac{1}{2}x_1^{(k)} \qquad + \dfrac{5}{2} \end{cases}$$

(2) $$\begin{cases} x_1^{(k+1)} = -0.7x_1^{(k)} + 0.5x_2^{(k)} + 4 \\ x_2^{(k+1)} = -0.1x_1^{(k)} + 0.2x_2^{(k)} + 3 \end{cases}$$

解 (1)迭代矩阵 $\boldsymbol{B} = \begin{bmatrix} 0 & -1/3 \\ -1/2 & 0 \end{bmatrix}$ $\|\boldsymbol{B}\|_\infty = 1/2 < 1$ ，故迭代公式收敛。

(2)迭代矩阵 $\boldsymbol{B} = \begin{bmatrix} -0.7 & 0.5 \\ -0.1 & 0.2 \end{bmatrix}$ $\|\boldsymbol{B}\|_1 = \max\{0.8, 0.7\} = 0.8 < 1$

故迭代公式收敛。

迭代法的收敛性

- $\rho(\boldsymbol{B})<1$是一阶定常迭代法的收敛判据
  - 对任意初值$\boldsymbol{x}^{(0)}$, 迭代法都收敛 (全局收敛)
  - 若某种范数下$\|\boldsymbol{B}\|_t<1$, 则收敛 (充分条件)
- 收敛阶与收敛速度　　**1 阶定常迭代法为 1 阶收敛**
  - 若 $\lim\limits_{k\to\infty}\left\|\boldsymbol{e}^{(k+1)}\right\|/\left\|\boldsymbol{e}^{(k)}\right\|^p=c$, 则$p$阶收敛　　$\boldsymbol{e}^{(k+1)}=\boldsymbol{B}\boldsymbol{e}^{(k)}$
  - 设$\boldsymbol{B}$可对角化, 特征值按模递减为$\{\lambda_i\}$, 特征向量$\{\boldsymbol{u}_i\}$

  设$\boldsymbol{e}^{(0)}=\sum_{i=1}^{n}\alpha_i\boldsymbol{u}_i$, 各误差成份随迭代衰减情况如下

$\boldsymbol{e}^{(0)}$	$\alpha_1\boldsymbol{u}_1$	$\alpha_2\boldsymbol{u}_2$	$\cdots$	$\alpha_n\boldsymbol{u}_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{e}^{(k)}=\boldsymbol{B}^k\boldsymbol{e}^{(0)}$	$\lambda_1^k\alpha_1\boldsymbol{u}_1$	$\lambda_2^k\alpha_2\boldsymbol{u}_2$	$\cdots$	$\lambda_n^k\alpha_n\boldsymbol{u}_n$
	最大			最小

$$\lim_{k\to\infty}\frac{\left\|\boldsymbol{e}^{(k+1)}\right\|}{\left\|\boldsymbol{e}^{(k)}\right\|}=\rho(B)$$

1阶收敛, 且
$c=|\lambda_1|=\rho(\boldsymbol{B})$

经典迭代法

矩阵分裂法(splitting method)

- 求解线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$的迭代法
  - 通过向量序列$\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)},\cdots$逼近准确解$\boldsymbol{x}^*$

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f},\qquad k=0,1,\cdots,$$

  - $\boldsymbol{B},\boldsymbol{f}$为常数矩阵和向量，称为“<u>一阶定常迭代法</u>”

$\boldsymbol{A}=\boldsymbol{M}-\boldsymbol{N}\Longrightarrow\boldsymbol{Mx}-\boldsymbol{Nx}=\boldsymbol{b}\Leftrightarrow\boldsymbol{x}=\boldsymbol{M}^{-1}\boldsymbol{Nx}+\boldsymbol{M}^{-1}\boldsymbol{b}$

怎么选$\boldsymbol{M}$?

1. $\boldsymbol{M}$非奇异;
2. 迭代法要收敛且收敛快;
3. 求解以$\boldsymbol{M}$为系数矩阵的方程组计算量小

Jacobi迭代法

以3阶方程组为例说明

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$\begin{cases} x_1 = -\dfrac{1}{a_{11}}(\qquad a_{12}x_2 + a_{13}x_3) + \dfrac{b_1}{a_{11}} \\ x_2 = -\dfrac{1}{a_{22}}(a_{21}x_1 \qquad + a_{23}x_3) + \dfrac{b_2}{a_{22}} \\ x_3 = -\dfrac{1}{a_{33}}(a_{31}x_1 + a_{32}x_2 \qquad) + \dfrac{b_3}{a_{33}} \end{cases}$$

- 分量迭代计算公式

$$\begin{cases} x_1^{(k+1)} = -\dfrac{1}{a_{11}}\left(\qquad a_{12}x_2^{(k)} + a_{13}x_3^{(k)}\right) + \dfrac{b_1}{a_{11}} \\ x_2^{(k+1)} = -\dfrac{1}{a_{22}}\left(a_{21}x_1^{(k)} \qquad + a_{23}x_3^{(k)}\right) + \dfrac{b_2}{a_{22}} \\ x_3^{(k+1)} = -\dfrac{1}{a_{33}}\left(a_{31}x_1^{(k)} + a_{32}x_2^{(k)} \qquad\right) + \dfrac{b_3}{a_{33}} \end{cases}, \quad k = 0, 1, \cdots,$$

$$\boxed{\boldsymbol{x}^{(k)} \longrightarrow \boldsymbol{x}^{(k+1)}}$$

- 设$\boldsymbol{A} = \boldsymbol{D} - (\boldsymbol{D} - \boldsymbol{A})$, $\boldsymbol{D}$为对角阵

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

Jacobi迭代法

算法4.2 雅可比迭代法 $(\boldsymbol{A}=(a_{ij})\in\mathbb{R}^{n\times n})$

输入： $\boldsymbol{x},\boldsymbol{A},\boldsymbol{b}$; 输出： $\boldsymbol{x}$.
While 不满足判停准则 do
    $\boldsymbol{y}:=\boldsymbol{x}$; {上一步迭代解}
    For i=1, 2, …, n
        $x_i:=(b_i-\sum_{j=1}^{i-1}a_{ij}y_j-\sum_{j=i+1}^{n}a_{ij}y_j)/a_{ii}$;
    End
End

若$\boldsymbol{A}$为稀疏阵, 改为只遍历非零元

条件：矩阵$\boldsymbol{A}$的对角元不为零

- 每步迭代的计算量相当于一次矩阵与向量的乘法; 对于稀疏矩阵，计算量为N_{nz}次乘法
- 计算过程不改变矩阵$\boldsymbol{A}$

算法简单, 易实现

例 用雅可比 (Jacobi) 迭代法求解方程组

$$\begin{bmatrix} 5 & -1 & -1 & -1 \\ -1 & 10 & -1 & -1 \\ -1 & -1 & 5 & -1 \\ -1 & -1 & -1 & 10 \end{bmatrix} x = \begin{bmatrix} -4 \\ 12 \\ 8 \\ 34 \end{bmatrix}$$

(准确解是 $\boldsymbol{x} = [1 \quad 2 \quad 3 \quad 4]^{\mathrm{T}}$), 取 $\boldsymbol{x}^{(0)} = \mathbf{0}$.

迭代公式

$$\boldsymbol{x}^{(k+1)} = \begin{bmatrix} 0 & 0.2 & 0.2 & 0.2 \\ 0.1 & 0 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0 \end{bmatrix} \boldsymbol{x}^{(k)} + \begin{bmatrix} -0.8 \\ 1.2 \\ 1.6 \\ 3.4 \end{bmatrix},$$

由于 $\|\boldsymbol{B}_{\mathrm{J}}\|_{\infty} = 0.6 < 1$, 故 Jacobi 迭代法是收敛的.

k	0	1	2	3	4	5
$\boldsymbol{x}^{(k)}$	0	-0.800	0.440	0.716	0.883	0.948
	0	1.200	1.620	1.840	1.929	1.969
	0	1.600	2.360	2.732	2.880	2.948
	0	3.400	3.600	3.842	3.929	3.969

Gauss-Seidel迭代法

- 与Jacobi迭代法类似
- 求出$x^{(k+1)}$的第1个分量后，即用它求第2个分量$x_2^{(k+1)}$
- 分量迭代计算公式

$$\begin{cases} x_1 = -\dfrac{1}{a_{11}}(\qquad a_{12}x_2 + a_{13}x_3) + \dfrac{b_1}{a_{11}} \\ x_2 = -\dfrac{1}{a_{22}}(a_{21}x_1 \qquad + a_{23}x_3) + \dfrac{b_2}{a_{22}} \\ x_3 = -\dfrac{1}{a_{33}}(a_{31}x_1 + a_{32}x_2 \qquad) + \dfrac{b_3}{a_{33}} \end{cases}$$

$$\begin{cases} x_1^{(k+1)} = -\dfrac{1}{a_{11}}\left(\qquad a_{12}x_2^{(k)} + a_{13}x_3^{(k)}\right) + \dfrac{b_1}{a_{11}} \\ x_2^{(k+1)} = -\dfrac{1}{a_{22}}\left(a_{21}x_1^{(k+1)} \qquad + a_{23}x_3^{(k)}\right) + \dfrac{b_2}{a_{22}} \\ x_3^{(k+1)} = -\dfrac{1}{a_{33}}\left(a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)} \qquad \right) + \dfrac{b_3}{a_{33}} \end{cases}, k = 0, 1, \cdots,$$

注意计算顺序

- 设$\boldsymbol{A} = \boldsymbol{D} - \tilde{\boldsymbol{L}} - \tilde{\boldsymbol{U}}$, 则迭代公式为

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

$\Longrightarrow \boldsymbol{x}^{(k+1)} = \boldsymbol{L}^{-1}(\boldsymbol{L} - \boldsymbol{A})\boldsymbol{x}^{(k)} + \boldsymbol{L}^{-1}\boldsymbol{b}$, $\boldsymbol{L}$为$\boldsymbol{A}$的下三角阵 分裂法: $\boldsymbol{M} = \boldsymbol{L}$

Gauss-Seidel迭代法

- 算法4.3 高斯一赛德尔迭代法

输入： $\boldsymbol{x}, \boldsymbol{A}, \boldsymbol{b}$；输出： $\boldsymbol{x}$.
While 不满足判停准则 do
  For i=1, 2, …, n
    $x_i := \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j\right)/a_{ii}$;
  End
End

若A为稀疏阵, 改为只遍历非零元

- 每步迭代的计算量相当于一次矩阵与向量相乘; 不需保留上一步迭代解, 与Jacobi迭代法计算量一样

按从1到n的顺序计算解分量.

- 若从n到1更新解分量,则得“逆向G-S算法”;
- 对称高斯-赛德尔(SGS)迭代法

例 用高斯一赛德尔(Gauss-Seidel) 迭代法求解方程组

$$\begin{bmatrix} 5 & -1 & -1 & -1 \\ -1 & 10 & -1 & -1 \\ -1 & -1 & 5 & -1 \\ -1 & -1 & -1 & 10 \end{bmatrix} x = \begin{bmatrix} -4 \\ 12 \\ 8 \\ 34 \end{bmatrix},$$ (准确解是 $\boldsymbol{x} = [1 \quad 2 \quad 3 \quad 4]^{\mathrm{T}}$), 取 $\boldsymbol{x}^{(0)} = \boldsymbol{0}$.

迭代公式

$$\boldsymbol{x}^{(k+1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0 & 0 \\ 0.1 & 0.1 & 0.1 & 0 \end{bmatrix} \boldsymbol{x}^{(k+1)} + \begin{bmatrix} 0 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{x}^{(k)} + \begin{bmatrix} -0.8 \\ 1.2 \\ 1.6 \\ 3.4 \end{bmatrix}$$

$$\boldsymbol{B}_{\mathrm{GS}} = \begin{bmatrix} 0 & 0.2 & 0.2 & 0.2 \\ 0 & 0.02 & 0.12 & 0.12 \\ 0 & 0.044 & 0.064 & 0.264 \\ 0 & 0.0264 & 0.0384 & 0.0584 \end{bmatrix}$$,由于 $\| \boldsymbol{B}_{GS} \|_{\infty} = 0.6 < 1$, 迭代法收敛.

k	0	1	2	3	4	5
$\boldsymbol{x}^{(k)}$	0	-0.800	0.476	0.889	0.977	0.995
	0	1.120	1.774	1.956	1.990	1.998
	0	1.664	2.770	2.949	2.989	2.998
	0	3.598	3.902	3.929	3.996	3.999

SOR迭代法

- Successive over relaxation method
- 在G-S迭代法基础上引入松弛因子ω　可能超出[0, 1]范围
- 先按G-S方法由$\boldsymbol{x}^{(k)}$算得$\tilde{x}_1^{(k+1)}$, 再将它与$x_1^{(k)}$作加权平均得$x_1^{(k+1)}$… 类似地依次算出后续分量

3元方程的计算公式

ω=1时, 是G-S方法

ω=0时, 没有意义

$$
\begin{cases}
x_1^{(k+1)} = (1-\omega)x_1^{(k)} + \omega\left(-\dfrac{a_{12}}{a_{11}}x_2^{(k)} - \dfrac{a_{13}}{a_{11}}x_3^{(k)} + \dfrac{b_1}{a_{11}}\right) \\
x_2^{(k+1)} = (1-\omega)x_2^{(k)} + \omega\left(-\dfrac{a_{21}}{a_{22}}x_1^{(k+1)} - \dfrac{a_{23}}{a_{22}}x_3^{(k)} + \dfrac{b_2}{a_{22}}\right) \\
x_3^{(k+1)} = (1-\omega)x_3^{(k)} + \omega\left(-\dfrac{a_{31}}{a_{33}}x_1^{(k+1)} - \dfrac{a_{32}}{a_{33}}x_2^{(k+1)} + \dfrac{b_3}{a_{33}}\right)
\end{cases}
$$

- 一般地, 设$\boldsymbol{A}=\boldsymbol{D}-\tilde{\boldsymbol{L}}-\tilde{\boldsymbol{U}}$, 则SOR迭代公式为

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

SOR迭代法

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

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

- 要求矩阵A的对角元不为0　　对应于分裂法中, $\boldsymbol{M}=\frac{1}{\omega}\boldsymbol{D}-\tilde{\boldsymbol{L}}$
- 算法4.4 **SOR迭代法**

> **输入**： $\boldsymbol{x}, \boldsymbol{A}, \boldsymbol{b}, \omega$; **输出**： $\boldsymbol{x}$.
> **While 不满足判停准则 do**
> 　**For** i=1, 2, …, n
> 　　$x_i := (1-\omega)x_i + \omega\,(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j)/a_{ii}$;
> 　**End**
> **End**

  - 计算量与G-S迭代法差不多
  - 需按从1到n的顺序计算解分量

例子

例4.3 / 4.4: 解线性方程组，设$\boldsymbol{x}^{(0)} = [0,0,0]^T$

$$\begin{bmatrix} 10 & 3 & 1 \\ 2 & -10 & 3 \\ 1 & 3 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ -5 \\ 14 \end{bmatrix}$$

准确解$\boldsymbol{x}^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

- **Jacobi迭代**(解分量误差小于10^{-3})
  - $\boldsymbol{x}^{(8)} = [1.0001, 0.9991, 1.0001]^T$
- **G-S迭代**
  - $\boldsymbol{x}^{(5)} = [0.9998,\ 0.9998,\ 1.0001]^T$
- **SOR迭代**
  - 当$\omega = 0.95$, $\boldsymbol{x}^{(4)} = [1.0008,\ 0.9999,\ 0.9999]^T$
  - 当$\omega = 1.1$, $\boldsymbol{x}^{(6)} = [1.0005,\ 1.0005,\ 0.9997]^T$
  - 当$\omega = 0.6$, $\boldsymbol{x}^{(9)} = [1.0010,\ 1.0001,\ 0.9998]^T$

都收敛！

收敛快慢不同

三种迭代法的收敛条件

***A*都是非奇异阵**

- 一阶定常迭代法的收敛判据: 迭代矩阵的谱半径$\rho(\boldsymbol{B})<1$

Jacobi方法: $\boldsymbol{B}_J=\boldsymbol{D}^{-1}(\boldsymbol{D}-\boldsymbol{A})$

G-S方法: $\boldsymbol{B}_G=\boldsymbol{L}^{-1}(\boldsymbol{L}-\boldsymbol{A})$

SOR方法: $\boldsymbol{B}_S=(\boldsymbol{D}-\omega\tilde{\boldsymbol{L}})^{-1}\left[(1-\omega)\boldsymbol{D}+\omega\widetilde{\boldsymbol{U}}\right]$

很难计算迭代矩阵

- 若$\|\boldsymbol{B}_J\|_1<1$或$\|\boldsymbol{B}_J\|_\infty<1$，则**Jacobi**迭代法收敛，**G-S**迭代法也收敛

- SOR迭代法收敛的必要条件：松弛因子ω满足$0<\omega<2$

三种迭代法的收敛条件

- **定理4.8** 若$\boldsymbol{A}$为对称正定阵, 雅可比迭代法收敛$\Leftrightarrow$ $2\boldsymbol{D}-\boldsymbol{A}$正定
- **定理4.12** 对称正定矩阵$\boldsymbol{A}$
  - 求解$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的G-S迭代法收敛
  - 若$0<\omega<2$, 则相应的SOR迭代法收敛

三种迭代法的收敛条件

定理：若线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$的系数矩阵$A$按行严格对角占优，则**Jacobi**迭代法和**G-S**迭代法对任意给定初值均收敛。

证明：记$e_k=\max\limits_{1\le j\le n}\left|x_j^{(k)}-x_j^*\right|$为第$k$次近似值$\left(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)}\right)^T$的误差.

（1）**Jacobi**迭代法

$$\left|x_i^{(k+1)}-x_i^*\right|=\left|\sum_{\substack{j=1\\j\neq i}}^{n}\frac{a_{ij}}{a_{ii}}\left(x_j^{(k)}-x_j^*\right)\right|\le\sum_{\substack{j=1\\j\neq i}}^{n}\frac{|a_{ij}|}{|a_{ii}|}\left|x_j^{(k)}-x_j^*\right|$$

记 $L=\sum_{\substack{j=1\\j\neq i}}^{n}\left|\frac{a_{ij}}{a_{ii}}\right|$, 则有 $\left|x_i^{(k+1)}-x_i^*\right|\le L\cdot\max\limits_{1\le j\le n}\left|x_j^{(k)}-x_j^*\right|$

上式对 $i=1,2,\ldots,n$成立，故有 $e_{k+1}\le Le_k\le\cdots\le L^{k+1}e_0$

因为A严格对角占优，故$L<1$，从而有

$$\lim_{k\to\infty}\left|x_i^{(k+1)}-x_i^*\right|\le\lim_{k\to\infty}L^{k+1}\left|x_i^{(0)}-x_i^*\right|=0,$$ 即**Jacobi**方法收敛。

（2） **G-S**迭代法

考虑**G-S**方法的误差

$$\left|x_i^{(k+1)}-x_i^*\right| = \left|\frac{\sum_{j=1}^{i-1} a_{ij}\left(x_j^{(k+1)}-x_j^*\right)+\sum_{j=i+1}^{n} a_{ij}\left(x_j^{(k)}-x_j^*\right)}{a_{ii}}\right|$$

$$\leq \frac{\sum_{j=1}^{i-1}|a_{ij}|\left|x_j^{(k+1)}-x_j^*\right|}{|a_{ii}|}+\frac{\sum_{j=i+1}^{n}|a_{ij}|\left|x_j^{(k)}-x_j^*\right|}{|a_{ii}|}$$

记$L=\sum_{j=1}^{i-1}\frac{|a_{ij}|}{|a_{ii}|}$，$S=\sum_{j=i+1}^{n}\frac{|a_{ij}|}{|a_{ii}|}$

则$\left|x_i^{(k+1)}-x_i^*\right| \leq L\cdot \max_{1\leq j\leq i-1}\left|x_j^{(k+1)}-x_j^*\right|+S\cdot \underset{i+1\leq j\leq n}{mas}\left|x_j^{(k)}-x_j^*\right|$

从而 $e_{k+1}\leq Le_{k+1}+Se_k$，因为$A$严格对角占优，故$L<1$，$S+L<1$, 从而有

$$e_{k+1}\leq\frac{S}{1-L}e_k\leq\cdots\leq\left(\frac{S}{1-L}\right)^{k+1}\cdot e_0$$

而$\left(\frac{S}{1-L}\right)=\frac{S-L+L}{1-L}=\frac{S+L}{1-L}-\frac{L}{1-L}<\frac{1}{1-L}-\frac{L}{1-L}=1$, 所以$\lim_{k\to\infty}\left|x_i^{(k+1)}-x_i^*\right|=0$,

即**G-S**迭代法收敛。

将线性方程组中的方程适当组合仍是同解方程组,但能改变系数矩阵. 如果处理得当,可以使不收敛的变为收敛的,或使收敛慢的变成收敛快的.

例 对于线性方程组

$$\begin{cases} 2x_1 - 1.8x_2 + 0.4x_3 = 1, \\ 3x_1 + 2x_2 - 1.1x_3 = 0.2 \\ x_1 - x_2 + 7.3x_3 = 1.6. \end{cases}$$

它的系数矩阵不是按行严格对角占优的,因此不能直接套用常规的 Jacobi 迭代公式或 Seidel 迭代公式. 但是,将第 1 个方程与第 2 个方程相加能使 x_1 的系数增大,而用 2 乘第 2 个方程减去第 1 个方程的 3 倍能使 x_2 的系数增大,且 x_1 的系数为零,得到同解方程组

$$\begin{cases} 5x_1 + 0.2x_2 - 0.7x_3 = 1.2, \\ 9.4x_2 - 3.4x_3 = -2.6, \\ -x_2 + 7.3x_3 = 1.6. \end{cases}$$

这个方程组的系数矩阵是按行严格对角占优的,从而可用常规的 Jacobi 迭代法或 Gauss-Seidel 迭代法求解.

习题3. 设方程组

$$\begin{cases} 5x_1 + 2x_2 + x_3 & = -12 \\ -x_1 + 4x_2 + 2x_3 & = 20 \\ 2x_1 - 3x_2 + 10x_3 & = 3 \end{cases},$$

(1) 考查用雅可比迭代法、高斯-塞德尔迭代法解此方程组的收敛性。

(2) 取初始解为 $[0,0,0]^{\mathrm{T}}$, 用雅可比迭代法及高斯-赛德尔迭代法解此方程组, 要求当 $\|x^{(k+1)} - x^{(k)}\|_\infty < 10^{-2}$ 时终止迭代。

解

(1) 雅可比迭代法、高斯-赛德尔迭代法解此方程组时都收敛。

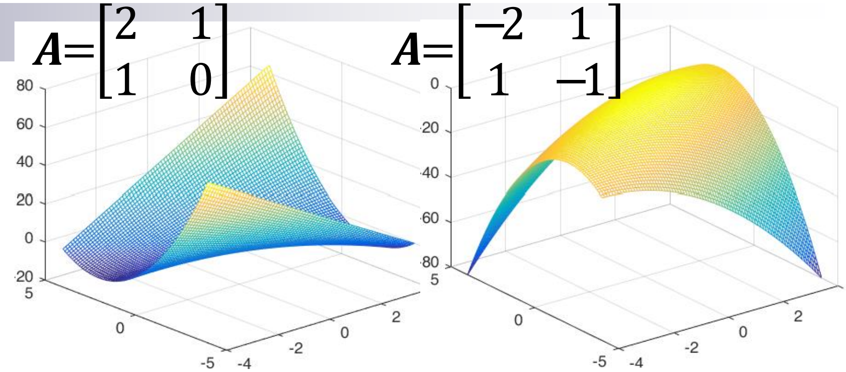
(2) 雅可比迭代法经过 17 步迭代, 得到近似解
$[-4.00002 \quad 3.00000 \quad 2.00000]^{\mathrm{T}}$;

高斯-赛德尔迭代法经过 8 步迭代, 得到近似解
$[-4.00002 \quad 3.00000 \quad 2.00000]^{\mathrm{T}}$ 。

最速下降法、共轭梯度法

最速下降法

若$\boldsymbol{A}$不是正定矩阵:

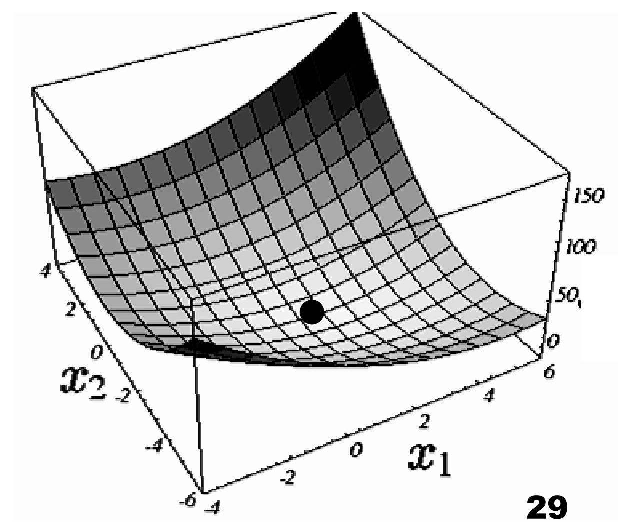


- 构造迭代法的新思路
  - 一阶定常迭代法(SOR)，对大规模矩阵收敛慢
  - “变分原理”：解线性方程组→在n维空间搜索极值点

设$\boldsymbol{A}$对称正定, 求n元二次函数$\varphi(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{x}$的最小值点

$$
\begin{cases}
\dfrac{\partial\varphi}{\partial x_1}=a_{11}x_1+\sum_{j\neq 1}a_{1j}x_j-b_1=0 \\
\quad\vdots \\
\dfrac{\partial\varphi}{\partial x_n}=a_{nn}x_n+\sum_{j\neq n}a_{nj}x_j-b_n=0
\end{cases}
\Longrightarrow \boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}=\boldsymbol{0}
$$

$\varphi(\boldsymbol{x})$最小值对应的$\boldsymbol{x}$就是$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的解



这个极值点一定是最小值?

$$\boldsymbol{Ax}-\boldsymbol{b}=\boldsymbol{0}\,,\varphi(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Ax}-\boldsymbol{b}^T\boldsymbol{x}=-\frac{1}{2}\boldsymbol{b}^T\boldsymbol{x}$$

$$\varphi(\boldsymbol{y})=\frac{1}{2}\boldsymbol{y}^T\boldsymbol{Ay}-\boldsymbol{b}^T\boldsymbol{y}=\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{x})-\frac{1}{2}\boldsymbol{b}^T\boldsymbol{x}$$

$$\varphi(\boldsymbol{y})-\varphi(\boldsymbol{x})=\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{x})\geq 0$$

最速下降法

- 直线搜索
  - 求$\varphi(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{x}$ 的最小值点
  - 无约束优化问题, 逐次搜索:$\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)},\cdots$
  - 在当前点$\boldsymbol{x}^{(k)}$, 设搜索方向$\boldsymbol{p}^{(k)}$, $\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)}$
  - 步长α_k应使$\varphi\left(\boldsymbol{x}^{(k+1)}\right)$最小

$$f(\alpha)=\varphi\left(\boldsymbol{x}^{(k+1)}\right)=\varphi(\boldsymbol{x}_k+\alpha\boldsymbol{p}_k)$$

$$=\frac{1}{2}(\boldsymbol{x}_k+\alpha\boldsymbol{p}_k)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}_k+\alpha\boldsymbol{p}_k)-\boldsymbol{b}^{\mathrm{T}}(\boldsymbol{x}_k+\alpha\boldsymbol{p}_k)$$

$$=\frac{1}{2}\alpha^2\boldsymbol{p}_k^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_k-\alpha\boldsymbol{r}_k^{\mathrm{T}}\boldsymbol{p}_k+\varphi(\boldsymbol{x}_k)$$，其中$\boldsymbol{r}^{(k)}=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)}$

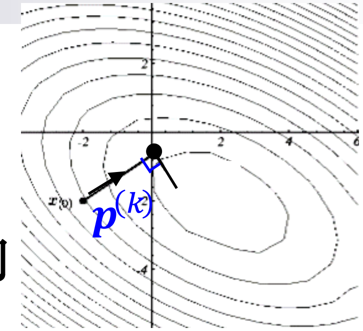
$f(\alpha)$ 有唯一的最小值 $\frac{\mathrm{d}f(\alpha)}{\mathrm{d}\alpha}=\alpha\boldsymbol{p}_k^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_k-\boldsymbol{r}_k^{\mathrm{T}}\boldsymbol{p}_k=0$

$$\boxed{\alpha_k=\frac{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{p}^{(k)}}{\left(\boldsymbol{p}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k)}}}$$

最速下降法

- 最速下降法
  - 沿负梯度方向$\boldsymbol{p}^{(k)}=\boldsymbol{r}^{(k)}$搜索，

$\boldsymbol{r}^{(k)}$沿$\varphi(\boldsymbol{x})$的梯度方向、等值面法向



$\varphi(\boldsymbol{x})$等值线(面)图

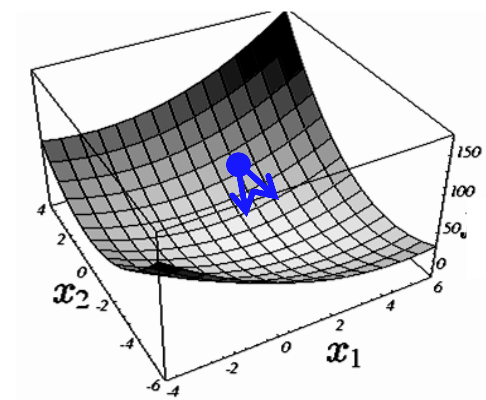
$$\left(\boldsymbol{p}^{(k)}\right)^T \boldsymbol{r}^{(k+1)}=0$$

搜索过程中相邻两步的搜索方向总是相互正交

$$\text{梯度 } \nabla\varphi\left(\boldsymbol{x}^{(k)}\right)=\left[\begin{matrix}\dfrac{\partial\varphi}{\partial x_1}\\ \vdots \\ \dfrac{\partial\varphi}{\partial x_n}\end{matrix}\right]\Bigg|_{\boldsymbol{x}=\boldsymbol{x}^{(k)}}=\boldsymbol{A}\boldsymbol{x}^{(k)}-\boldsymbol{b}=-\boldsymbol{r}^{(k)}$$

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

$$\alpha_k=\frac{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{p}^{(k)}}{\left(\boldsymbol{p}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k)}} \quad \text{得}\alpha_k=\frac{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{r}^{(k)}}{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{r}^{(k)}}$$



$$\boldsymbol{r}_{k+1}=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_{k+1}=\boldsymbol{b}-\boldsymbol{A}(\boldsymbol{x}_k+\alpha_k\boldsymbol{r}_k)=\boldsymbol{r}_k-\alpha_k\boldsymbol{A}\boldsymbol{r}_k$$

最速下降法

解线性方程组的最速下降算法 (steepest descent,SD)

输入: $\boldsymbol{x}, \boldsymbol{A}, \boldsymbol{b}$; 输出: $\boldsymbol{x}$.
$\boldsymbol{r} := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$;
While 不满足判停准则 **do**
$\alpha := \boldsymbol{r}^T\boldsymbol{r}/\boldsymbol{r}^T\boldsymbol{A}\boldsymbol{r}$;
$\boldsymbol{x} := \boldsymbol{x} + \alpha\boldsymbol{r}$;
$\boldsymbol{r} := \boldsymbol{r} - \alpha\boldsymbol{A}\boldsymbol{r}$;
End

用残差判据无额外代价
常用相对残差 $\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|/\|\boldsymbol{b}\|$

- 每步做一次矩阵与向量乘法, 运算量与SOR迭代法一样
- 需额外存储2个向量 $\boldsymbol{A}\boldsymbol{r}$ 和 $\boldsymbol{r}$
- A对称正定, 有唯一最小值点, 最速下降法一定收敛

最速下降法

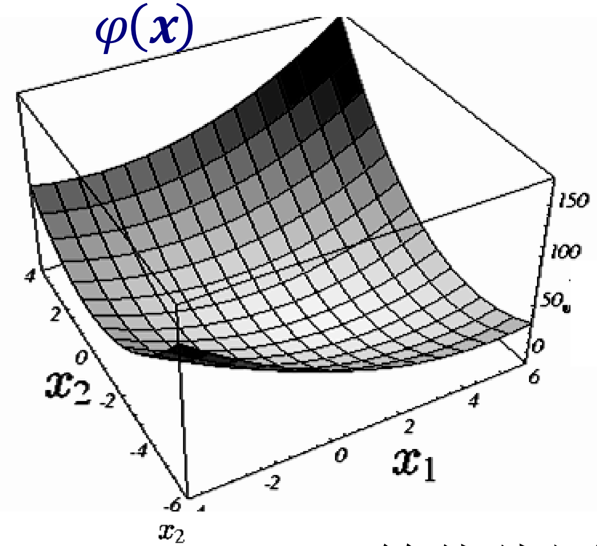
例: 解$\boldsymbol{Ax}=\boldsymbol{b}$, $\boldsymbol{A}=\begin{bmatrix}3 & 2\\ 2 & 6\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}2\\ -8\end{bmatrix}$

设初始解为$\boldsymbol{x}^{(0)}=[-2,-2]^T$

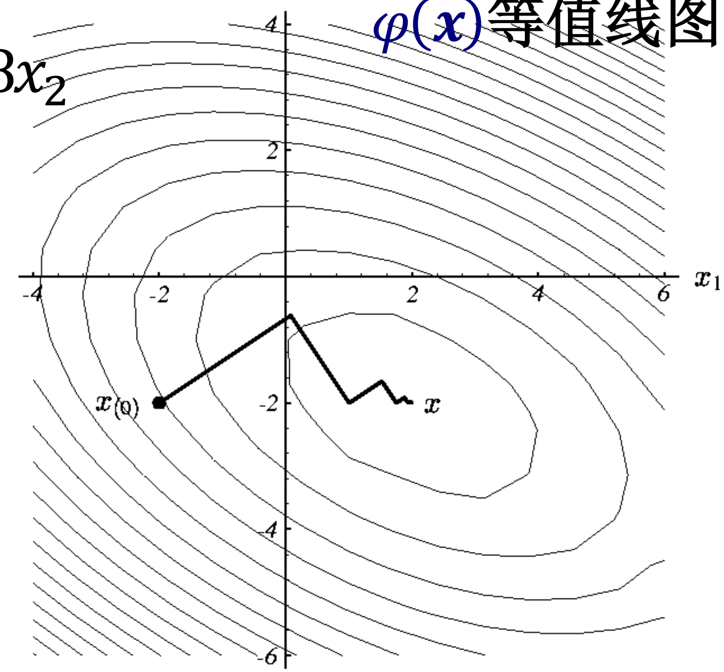
$$\boldsymbol{x}^{(1)}=\begin{bmatrix}0.08\\ -0.6133\end{bmatrix},\ldots,\boldsymbol{x}^{(9)}=\begin{bmatrix}1.9926\\ -1.9947\end{bmatrix}$$

$$\varphi(\boldsymbol{x})=\frac{3}{2}x_1^2+2x_1x_2+3x_2^2-2x_1+8x_2$$

$\varphi(\boldsymbol{x})$



$\varphi(\boldsymbol{x})$等值线图



- 最速下降法的缺点

  虽然每步找到局部最优解，
  但迭代法收敛速度慢

共轭梯度法

改进:

没必要沿最速下降方向搜索!

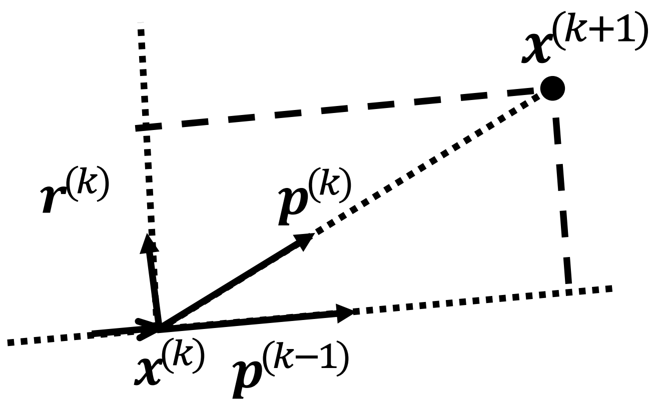
- Conjugate Gradient method (CG法)

给定初始向量 $\boldsymbol{x}_0$, 第一步仍选负梯度方向为搜索方向, 即 $\boldsymbol{p}_0=\boldsymbol{r}_0$, 有

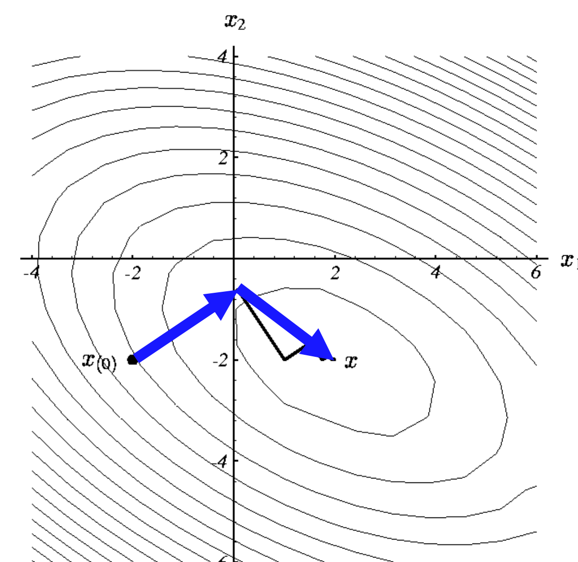
$$\alpha_0=\frac{\boldsymbol{r}_0^{\mathrm{T}}\boldsymbol{r}_0}{\boldsymbol{p}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_0},\ \boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{p}_0,\ \boldsymbol{r}_1=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_1$$

共轭梯度法

- 当前点$\boldsymbol{x}^{(k)}$, $\varphi(\boldsymbol{x})$的值减小最快的方向是$\boldsymbol{r}^{(k)}$, 前一个搜索方向是$\boldsymbol{p}^{(k-1)}$
- 在过$\boldsymbol{x}^{(k)}$,上述两方向<u>张成的平面</u>上找函数φ的最小值



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



$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\alpha_k\left(\boldsymbol{r}^{(k)}+\beta_{k-1}\boldsymbol{p}^{(k-1)}\right)=\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{r}^{(k)}+\alpha_k\beta_{k-1}\boldsymbol{p}^{(k-1)}$$

$$f(\alpha,\beta)=\varphi\left(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{r}^{(k)}+\alpha\beta\boldsymbol{p}^{(k-1)}\right)$$

$$=\frac{1}{2}\left(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{r}^{(k)}+\alpha\beta\boldsymbol{p}^{(k-1)}\right)^{\mathrm{T}}\boldsymbol{A}\left(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{r}^{(k)}+\alpha\beta\boldsymbol{p}^{(k-1)}\right)-\boldsymbol{b}^{\mathrm{T}}\left(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{r}^{(k)}+\alpha\beta\boldsymbol{p}^{(k-1)}\right)$$

$$\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{p}^{(k-1)}=0$$

$$\beta_{k-1}=-\frac{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k-1)}}{\left(\boldsymbol{p}^{(k-1)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k-1)}}$$

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

$$\alpha_k=\frac{\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{p}^{(k)}}{\left(\boldsymbol{p}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k)}}$$

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

公式可以进一步简化：

$\left(\boldsymbol{p}^{(k)}\right)^T\boldsymbol{A}\boldsymbol{p}^{(k-1)}=0$ (共轭正交)

残差向量相互正交，即 $\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{r}_j=0, i\neq j$

$\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{p}^{(k)}=\left(\boldsymbol{r}^{(k)}\right)^T\boldsymbol{r}^{(k)}$

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

共轭梯度法

- 几个重要的结论
  - $\boldsymbol{x}^{(k+1)} \in \boldsymbol{x}^{(0)} + span\{\boldsymbol{r}^{(0)}, \boldsymbol{A}\boldsymbol{r}^{(0)}, \ldots, \boldsymbol{A}^k\boldsymbol{r}^{(0)}\}$ **Krylov**子空间$\mathcal{K}(\boldsymbol{A}, \boldsymbol{r}^{(0)})$
  - $\boldsymbol{x}^{(k+1)}$不但在向量$\boldsymbol{r}^{(k)}$和$\boldsymbol{p}^{(k-1)}$所张平面上使$\varphi(\boldsymbol{x})$最小，且在超平面$\boldsymbol{x}^{(0)} + \mathcal{K}(\boldsymbol{A}, \boldsymbol{r}^{(0)})$上使$\varphi(\boldsymbol{x})$最小 $\mathbb{R}^{k+1}$维子空间
  - 若不考虑数值误差, $\boldsymbol{x}^{(n)}$为准确解

算法4.7:

$\boldsymbol{r} := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$; $\boldsymbol{p} := \boldsymbol{r}$;
While 不满足判停准则 **do**
  $\alpha := \boldsymbol{r}^T\boldsymbol{r}/\boldsymbol{p}^T\boldsymbol{A}\boldsymbol{p}$; {计算搜索步长}
  $\boldsymbol{x} := \boldsymbol{x} + \alpha\boldsymbol{p}$; {更新解}
  $\tilde{\boldsymbol{r}} := \boldsymbol{r}$; $\boldsymbol{r} := \boldsymbol{r} - \alpha\boldsymbol{A}\boldsymbol{p}$; {更新残差向量}
  $\beta := \boldsymbol{r}^T\boldsymbol{r}/\tilde{\boldsymbol{r}}^T\tilde{\boldsymbol{r}}$;
  $\boldsymbol{p} := \boldsymbol{r} + \beta\boldsymbol{p}$; {新的搜索方向}
End

一次矩阵向量乘法
两次向量内积
收敛所需步数≤n

定理(共轭向量一定线性无关) $\boldsymbol{S}$ 为 $n \times n$ 的对称正定矩阵, 如果 $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \cdots, \boldsymbol{d}^{(k)} \in \mathbb{R}^n, 0 < k \leqslant n-1$ 是关于 $\boldsymbol{S}$ 共轭的, 那么它们是线性无关的。

证明: 若存在一组标量 $\alpha_0, \alpha_1, \cdots, \alpha_k$, 使得

$$\alpha_0 \boldsymbol{d}^{(0)} + \alpha_1 \boldsymbol{d}^{(1)} + \cdots + \alpha_k \boldsymbol{d}^{(k)} = \mathbf{0}$$

两端左乘 $\boldsymbol{d}^{(j)\top}\boldsymbol{S}, 0 \leqslant j \leqslant k$, 根据 $\boldsymbol{S}$ 共轭的定义, 可知 $\boldsymbol{d}^{(j)\top}\boldsymbol{S}\boldsymbol{d}^{(i)} = 0, i \neq j$, 由此可得

$$\alpha_j \boldsymbol{d}^{(j)\top}\boldsymbol{S}\boldsymbol{d}^{(j)} = 0$$

由于 $\boldsymbol{S}$对称正定, $\boldsymbol{d}^{(j)} \neq \mathbf{0}$, 因此 $\alpha_j = 0, j = 0,1, \cdots, k$。

因此 $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \cdots, \boldsymbol{d}^{(k)}, 0 < k \leqslant n-1$ 是线性无关的。

推论 $\boldsymbol{S}$ 为 $n \times n$ 的对称正定矩阵, 关于 $\boldsymbol{S}$ 两两共轭的方向最多有n个。

例4.8 (与前例相同)

k	$\boldsymbol{Ap}_k$	α_k	$\boldsymbol{x}_{k+1}$	$\boldsymbol{r}_{k+1}$	β_k	$\boldsymbol{p}_{k+1}$
0	$[52,72]^{\mathrm{T}}$	0.1733	$[0.08,-0.6133]^{\mathrm{T}}$	$[2.9867,-4.48]^{\mathrm{T}}$	0.1394	$[4.6592,-3.365]^{\mathrm{T}}$
1	$[7.2476,-10.8715]^{\mathrm{T}}$	0.4121	$[2,-2]^{\mathrm{T}}$			

经过两步迭代, 共轭梯度法就得到了准确解。事实上, 由于 $\boldsymbol{x}_2$ 是在 $\boldsymbol{p}_1$ 和 $\boldsymbol{r}_1$ 所张成的二维平面上的 $\varphi(\boldsymbol{x})$ 的最小值点, 而本题的向量空间就是二维的, 因此 $\boldsymbol{x}_2$ 就是全局的最小值, 当然是方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的准确解了。

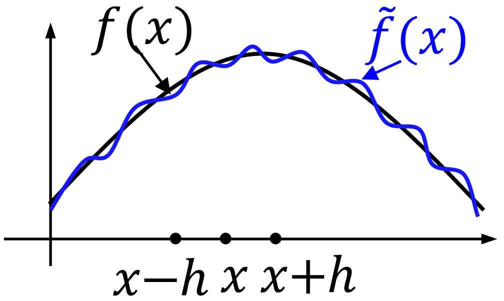
为什么共轭梯度法原则上是一种直接法? 但在实际计算中又将它作为迭代法?

在共轭梯度法中, 由于 $\{\boldsymbol{r}^{(k)}\}$ 互相正交, 所以在 $\boldsymbol{r}^{(0)},\boldsymbol{r}^{(1)},\cdots,\boldsymbol{r}^{(n)}$ 中至少有一个零向量, 若 $\boldsymbol{r}^{(k)}=0$, 则 $\boldsymbol{r}^{(k)}=\boldsymbol{x}^*$, 因而共轭梯度法求解 n 维对称正定线性方程组理论上最多 n 步可求得精确解, 从这个意义上讲, CG 算法是一种直接法.

但在实际计算中, 由于舍入误差的存在, 很难保证 $\{\boldsymbol{r}^{(k)}\}$ 的正交性. 另外当 n 很大时, 往往在实际计算步数 $k \ll n$ 时即可达到精度要求而不必计算 n 步, 所以实际计算中往往将 CG 算法作为迭代法.

数值微分

- 问题的描述
  - 近似计算函数的导数$f'(x)$, 其中$f(x)$的表达式未知
  - 使用若干函数值近似计算其导数



- 基本的有限差分公式(finite difference)
  - $f'(x) \approx D_f(h) = \dfrac{f(x+h)-f(x)}{h}$ (向前差分)
  - $f'(x) \approx D_b(h) = \dfrac{f(x)-f(x-h)}{h}$ (向后差分)
  - $f'(x) \approx D_c(h) = \dfrac{f(x+h)-f(x-h)}{2h}$ (中心差分)

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \cdots$$

Taylor展开

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \cdots$$

$$D_f(h) = \frac{f(x+h) - f(x)}{h} = f'(x) + O(h)$$

1 阶准确度

$$D_b(h) = \frac{f(x) - f(x-h)}{h} = f'(x) + O(h)$$

$$D_c(h) = \frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{f^{(3)}(\xi)}{3!}h^2 = f'(x) + O\left(h^2\right)$$

2 阶准确度

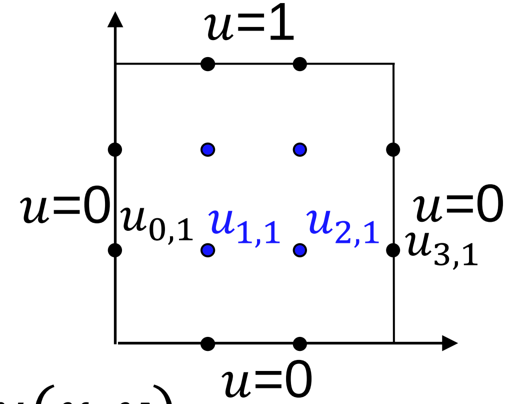
$$G_c(h) = 2f[x-h, x, x+h] = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f''(x) + O\left(h^2\right)$$

(二阶中心差分)

例 (正方形区域的拉普拉斯方程)

- **有限差分法**解偏微分方程
  - 正方形区域的二维Laplace方程

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \qquad x,y\in[0,1]\text{(除边界上)}$$



  - 已知$u(x,y)$在边界上的值, 求区域内的$u(x,y)$.

    每边均匀分n+1份得离散网格, 在网格点上设变量$u_{i,j}$

二阶中心差分 $$\frac{\partial^2 u(x_i,y_j)}{\partial x^2}\approx\frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{h^2}$$

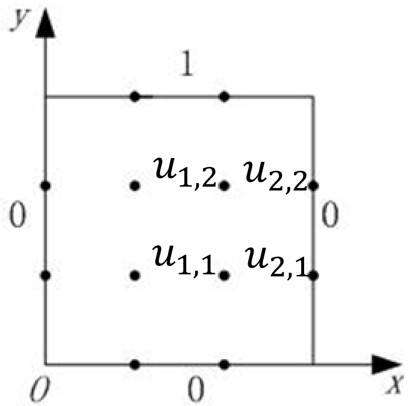
对每个格点列, 微分方程变为代数方程

$$\cdots\ \cdots$$

n=2时

$$\begin{cases}4u_{1,1}-u_{0,1}-u_{2,1}-u_{1,0}-u_{1,2}=0\\4u_{2,1}-u_{1,1}-u_{3,1}-u_{2,0}-u_{2,2}=0\\4u_{1,2}-u_{0,2}-u_{2,2}-u_{1,1}-u_{1,3}=0\\4u_{2,2}-u_{1,2}-u_{3,2}-u_{2,1}-u_{2,3}=0\end{cases}\Longrightarrow\begin{bmatrix}4&-1&-1&0\\-1&4&0&-1\\-1&0&4&-1\\0&-1&-1&4\end{bmatrix}\begin{bmatrix}u_{1,1}\\u_{2,1}\\u_{1,2}\\u_{2,2}\end{bmatrix}=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

k=2



$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$$

$$\begin{cases}4u_{1,1}-\cancel{u_{0,1}}-u_{2,1}-\cancel{u_{1,0}}-u_{1,2}=0\\ 4u_{2,1}-u_{1,1}-\cancel{u_{3,1}}-\cancel{u_{2,0}}-u_{2,2}=0\\ 4u_{1,2}-\cancel{u_{0,2}}-u_{2,2}-u_{1,1}-\cancel{u_{1,3}}=0\\ 4u_{2,2}-u_{1,2}-\cancel{u_{3,2}}-u_{2,1}-\cancel{u_{2,3}}=0\end{cases}$$

k=3

不同大小的一系列稀疏矩阵,每一行最多只有5个非零元素

方法比较

- 迭代法的比较
  - 收敛速度: **Jacobi < G-S < SOR < CG**
  - 随着k增大, 前三种方法的$\rho(\boldsymbol{B})$趋于**1**
  - 共轭梯度法的收敛显著好于这三种方法

- 直接法与迭代法比较
  - 直接法稳定、通用, 但处理大规模问题有困难
  - 迭代法能很好利用矩阵稀疏性, 但无通用的有效方法