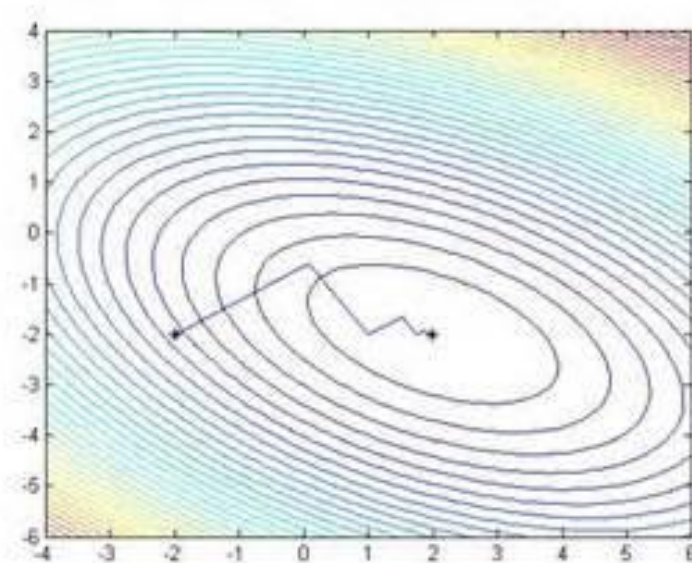
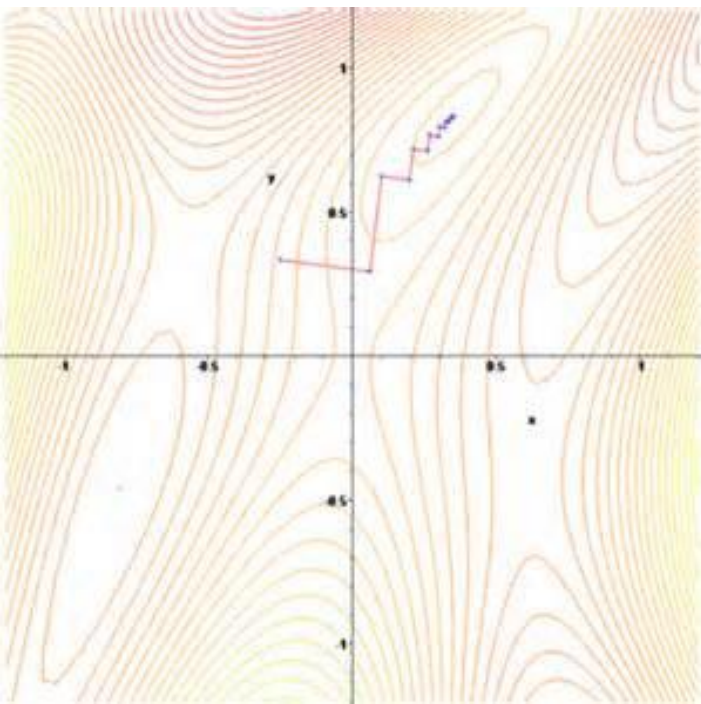


8.8 共轭梯度法



胡 茂 彬

<http://staff.ustc.edu.cn/~humaobin>
/

humaobin@ustc.edu.cn

共轭梯度法的优势

- 对于一个 N 阶的线性方程组，最多不超过 N 次迭代即可求得精确解！！！！
- 只需存贮非零元素，较消元法所需存贮量小得多

8.8.1 共轭梯度法的基本思想

- 1、方程组求解的等价定理
- 2、求解的思路

1 等价定理

- 求解方程组:

$$\mathbf{A}\mathbf{X}=\mathbf{b}$$

- 等价于求以下函数的极小值:

$$f(\mathbf{X}) = \frac{1}{2} \mathbf{X}^T \mathbf{A} \mathbf{X} - \mathbf{b}^T \mathbf{X}$$

证 明

$$f(\mathbf{X}) = \frac{1}{2} \mathbf{X}^T \mathbf{A} \mathbf{X} - \mathbf{b}^T \mathbf{X} = \frac{1}{2} \sum_{i,j=1}^N a_{i,j} x_i x_j - \sum_{i=1}^N b_i x_i$$

- 其极值点:

$$\frac{\partial f(\mathbf{X})}{\partial x_k} = \sum_{j=1}^N a_{k,j} x_j - b_k = 0, \quad k = 1, 2, \dots, N$$

AX分量

极值点

$$\frac{\partial^2 f(\mathbf{X})}{\partial x_i \partial x_j} = a_{i,j}, \quad i, j = 1, 2, \dots, N$$

- 在极值点附近:

$$f(\mathbf{X}^* + \mathbf{t}) - f(\mathbf{X}^*) = \sum_{i,j=1}^N \frac{\partial^2 f(\mathbf{X}^*)}{\partial x_i \partial x_j} t_i t_j = \sum_{i,j=1}^N a_{i,j} t_i t_j$$

A 正定

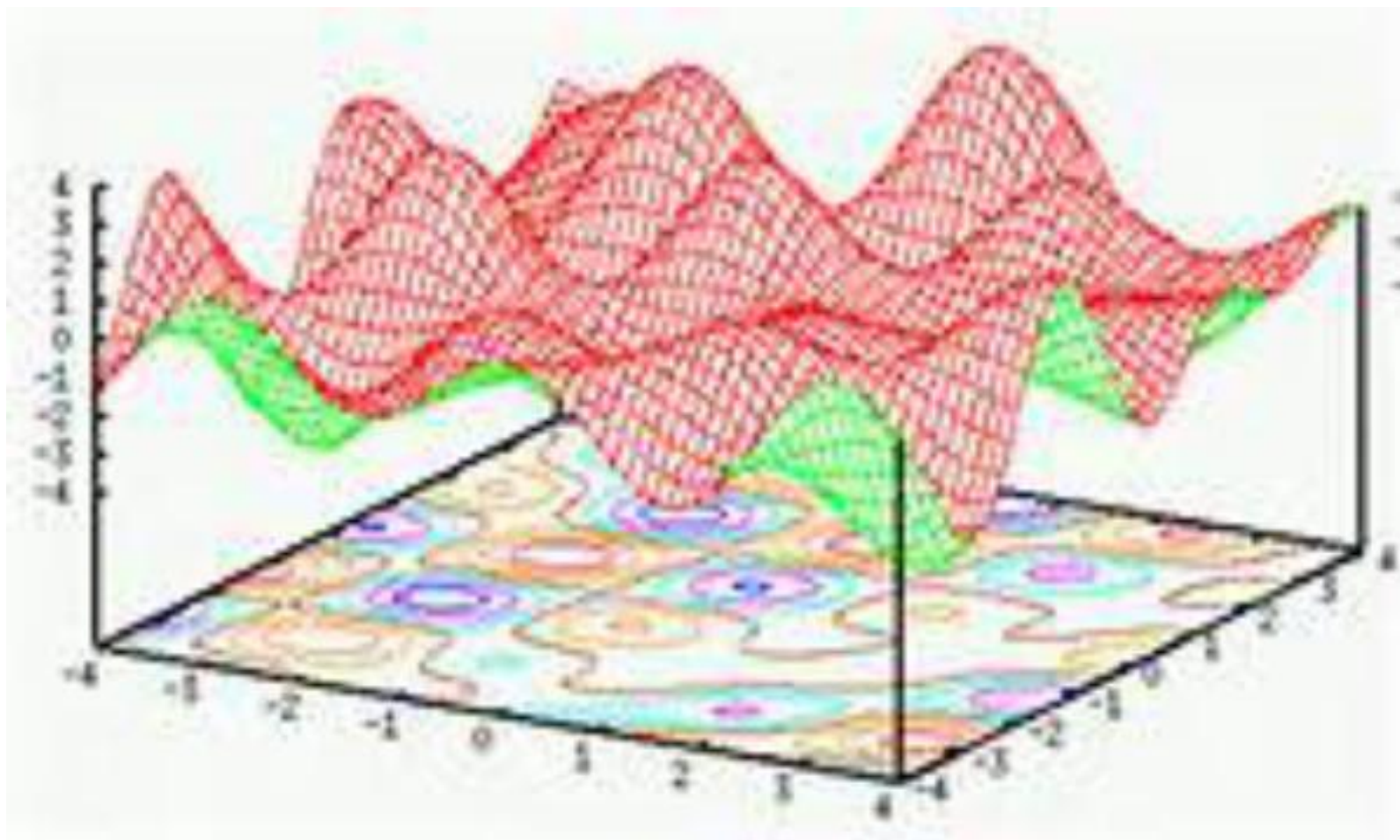
$$\mathbf{t}^T \mathbf{A} \mathbf{t} = \sum_{i,j=1}^N a_{i,j} t_i t_j > 0$$

极小值点！

$$f(\mathbf{X}^* + \mathbf{t}) - f(\mathbf{X}^*) > 0$$

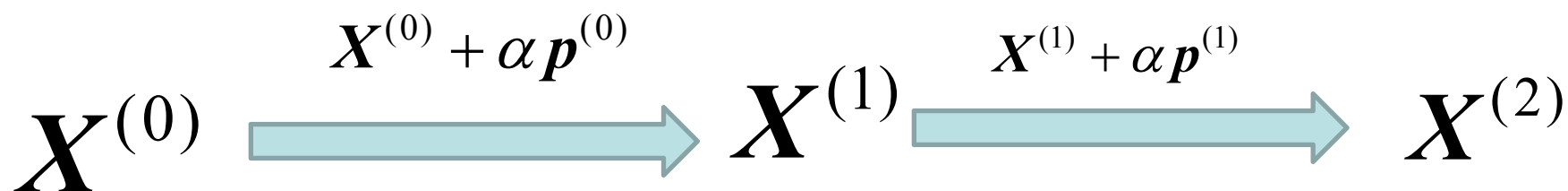
- 求得函数 $f(\mathbf{X})$ 的极小值点，就求得了代数方程组 $\mathbf{AX}=\mathbf{b}$ 的解

2. 求函数极小值



妥协方法

- $\{X^{(k)}\} \rightarrow f(X)$ 极小值



- 问题1: 确定搜索方向p
- 问题2: 求得下一个位置

问题2：求得下一个值

$$\begin{aligned}\varphi(\alpha) &= f(\mathbf{X}^{(k)} + \alpha \mathbf{p}^{(k)}) \\ &= \frac{1}{2} (\mathbf{X}^{(k)} + \alpha \mathbf{p}^{(k)})^T \mathbf{A} (\mathbf{X}^{(k)} + \alpha \mathbf{p}^{(k)}) - \mathbf{b}^T (\mathbf{X}^{(k)} + \alpha \mathbf{p}^{(k)})\end{aligned}$$

- 此方向上的极小值应满足：

$$\frac{d\varphi(\alpha)}{d\alpha} = \alpha (\mathbf{p}^{(k)})^T \mathbf{A} \mathbf{p}^{(k)} + (\mathbf{p}^{(k)})^T (\mathbf{A} \mathbf{X}^{(k)} - \mathbf{b}) = 0$$

该方向的极值点位置

$$\alpha_k = -\frac{\left(\mathbf{p}^{(k)}\right)^T \left(\mathbf{A}\mathbf{X}^{(k)} - \mathbf{b}\right)}{\left(\mathbf{p}^{(k)}\right)^T \mathbf{A}\mathbf{p}^{(k)}}$$



梯度(斜量)

$$\mathbf{g}^{(k)} = \mathbf{A}\mathbf{X}^{(k)} - \mathbf{b} = \nabla f(\mathbf{X}^{(k)})$$

$$\alpha_k = -\frac{\left(\mathbf{p}^{(k)}\right)^T \mathbf{g}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T \mathbf{A}\mathbf{p}^{(k)}}$$

极值近似点

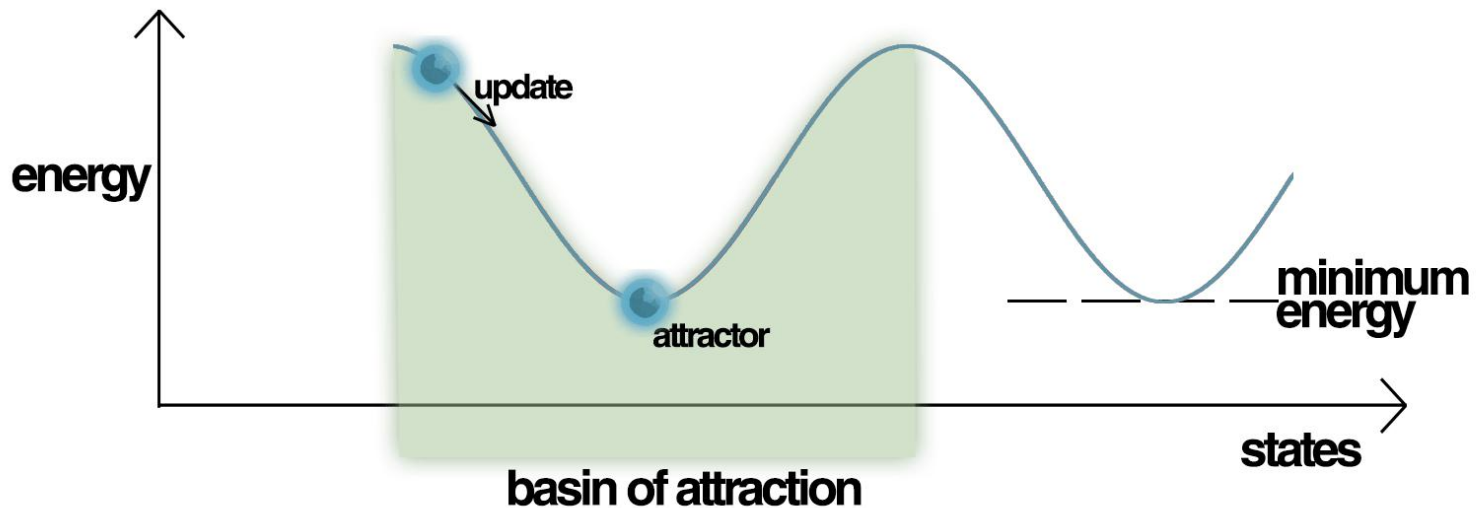
$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \alpha_k \mathbf{p}^{(k)}$$

问题1：确定搜索方向

- 梯度法(最速下降法):

取负梯度方向作为搜索方向

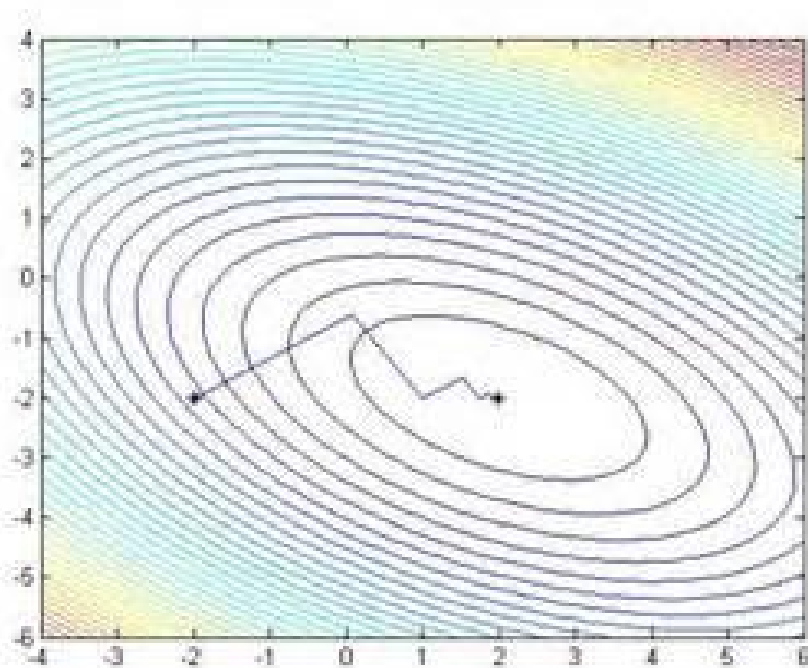
$$\mathbf{p}^{(k)} = -\nabla f(\mathbf{X}^{(k)})$$



共轭梯度法

- 采用正定矩阵 $\mathbf{A}$ 的共轭向量系:

$$\left(\mathbf{p}^{(i)}\right)^T \mathbf{A} \mathbf{p}^{(j)} = 0, \quad i \neq j$$



共轭向量系的性质

- 在函数定义域上**线性无关**
- 其向量**数不会超过 N**
- 以此方向得到的X满足梯度**正交关系**:

$$\underbrace{\left(\mathbf{g}^{(k)}\right)^T}_{\text{梯度(斜量)}} \underbrace{\mathbf{p}^{(i)}}_{\text{搜索方向}} = 0, \quad i = 0, 1, 2, \dots, k-1$$

由正交关系取得最终解的途径

- 以该组共轭向量系 $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \dots, \mathbf{p}^{(k-1)}$ 进行迭代

- 最后一个点必满足：

$$\mathbf{g}^{(N)} = \mathbf{A}\mathbf{X}^{(N)} - \mathbf{b} = \mathbf{0}$$

- 经过最多 N 次迭代就可得到解！

构造共轭向量系

- 起始: $\mathbf{p}^{(0)} = -\mathbf{g}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{X}^{(0)} = \mathbf{r}^{(0)}$
- 算得: $\mathbf{X}^{(1)}$
- 得到梯度: $\mathbf{g}^{(1)} = \mathbf{r}^{(1)} = \mathbf{b} - \mathbf{A}\mathbf{X}^{(1)}$
- 由 $\mathbf{p}^{(0)}$ 和 $\mathbf{g}^{(1)}$ 得到 $\mathbf{p}^{(1)}$ 下一方向

共轭向量的递推关系

上一个搜索方向

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

梯度(斜量)

- 求系数 **beta** (证明从略)

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

共轭梯度法的总结

- 一个正定系数矩阵 A 的方程组求解等价于一个二次函数的极小值求解问题
- 构造矩阵 A 的共轭向量系作为搜索极小值的方向
- 在最多 N 次迭代计算后，就可得某个极小值，即方程的解

8.8.2 共轭梯度法的实施

α_k 的等价简化形式

$$\begin{aligned}\alpha_k &= -\frac{\left(\mathbf{p}^{(k)}\right)^T \mathbf{g}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T A \mathbf{p}^{(k)}} = -\frac{\left(-\mathbf{g}^{(k)} + \beta_{k-1} \mathbf{p}^{(k-1)}\right)^T \mathbf{g}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T A \mathbf{p}^{(k)}} \\ &= \frac{\left(\mathbf{r}^{(k)}\right)^T \mathbf{r}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T A \mathbf{p}^{(k)}}\end{aligned}$$

β_k 的等价简化形式

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

$$\begin{aligned}\beta_k &= \frac{\left(\mathbf{g}^{(k+1)}\right)^T \mathbf{A} \mathbf{p}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T \mathbf{A} \mathbf{p}^{(k)}} = \frac{\left(\mathbf{g}^{(k+1)}\right)^T \left(\mathbf{A} \mathbf{X}^{(k+1)} - \mathbf{A} \mathbf{X}^{(k)}\right) / \alpha_k}{\left(\mathbf{p}^{(k)}\right)^T \mathbf{A} \mathbf{p}^{(k)}} \\ &= \frac{\left(\mathbf{r}^{(k+1)}\right)^T \mathbf{r}^{(k+1)}}{\left(\mathbf{r}^{(k)}\right)^T \mathbf{r}^{(k)}}\end{aligned}$$

共轭梯度法具体算法

- 假定初场 $\mathbf{X}^{(0)}$
- 计算初始余量: $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{X}^{(0)}$
- 令初始搜索方向: $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$

For $k = 0, 1, 2, \dots$ until $\mathbf{r}^{(k)} < \text{tolerance}$ (余量 < 允许值)

极值位置: $\alpha_k = \left(\mathbf{r}^{(k)}, \mathbf{r}^{(k)} \right) / \left(A\mathbf{p}^{(k)}, \mathbf{p}^{(k)} \right)$

函数值: $\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \alpha_k \mathbf{p}^{(k)}$

算余量: $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{p}^{(k)}$

算系数: $\beta_k = \left(\mathbf{r}^{(k+1)}, \mathbf{r}^{(k+1)} \right) / \left(\mathbf{r}^{(k)}, \mathbf{r}^{(k)} \right)$

下一迭代方向: $\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{p}^{(k)}$

