Optimal Theory and Method

最优化理论与方法

翟春杰

杭州电子科技大学 自动化学院

科技馆512

Email: cjzhai@hdu.edu.cn

概述

- **考虑无约束优化问题**

$$\min_{x\in R^n} f(x)$$

- **无约束优化问题的优化算法**

➢ **用到了目标函数的导数——使用导数的优化方法**

最速下降法，牛顿法，共轭梯度法，信赖域法，拟牛顿法，最小二乘法

➢ **只用到了目标函数值——直接法**

模式搜索法，单纯形搜索法，**Powell**方法，**Rosenbrock**法

概述

■ 直接算法的六大共同特点

- 不需要计算导数
- 收敛地比较慢
- 不要求目标函数的导数存在
- 算法迭代比较简单
- 比较容易编制程序
- 对于变量不多的优化问题，效果较好

无约束最优化的直接方法

- **模式搜索法**
- **Rosenbrock法**
- **单纯形搜索法**
- **Powell方法**

模式搜索法

■ 基本思想

◆ 1961年由Hooke和Jeeves提出，又称**Hooke-Jeeves法**

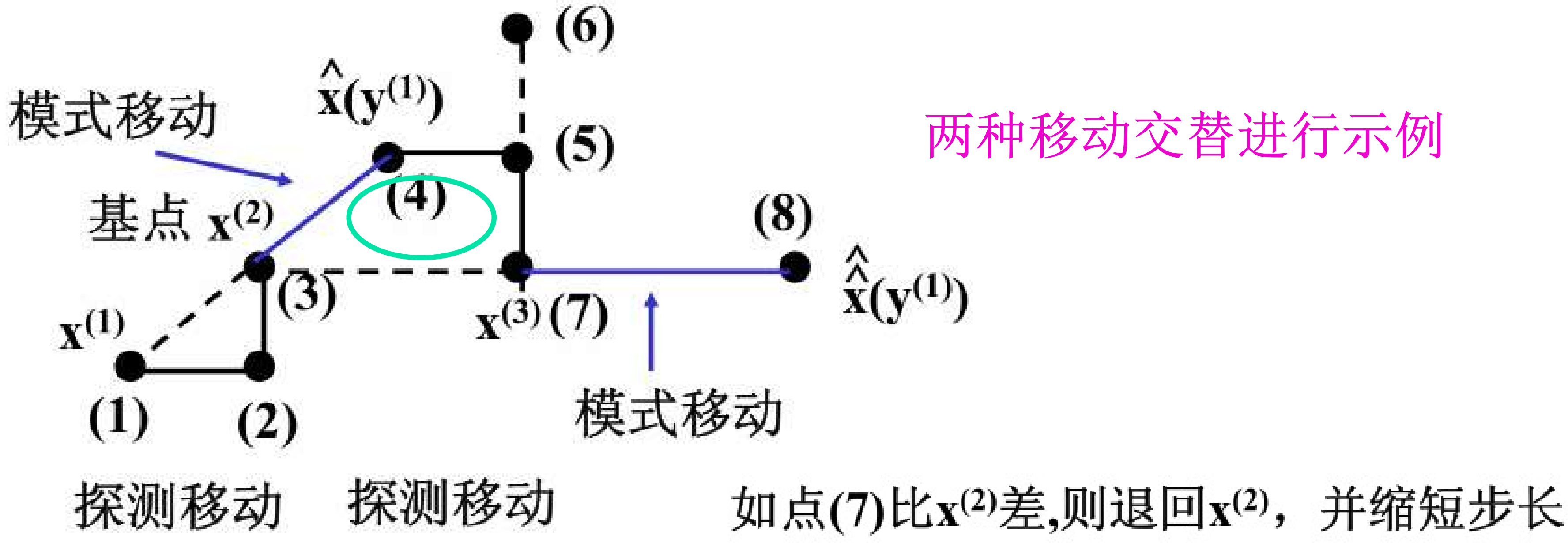
◆ 从**几何意义**上讲，模式搜索法就是寻找**具有较小函数值的"山谷"**，力图使迭代产生的序列**沿"山谷"走向**逼近极小点。

◆ 算法从初始**基点**开始，包括两种类型的移动，即**探测移动**和**模式移动**。

模式搜索法

■ 基本思想

- ◆ **探测移动：依次沿n个坐标轴进行，用以确定新的基点和有利于函数值下降的方向。**
- ◆ **模式移动：沿相邻两个基点连线方向进行，试图顺着“山谷”使函数值更快地减小。**



两种移动交替进行示例

模式搜索法

■ 基本思想

- 设目标函数为$f(\boldsymbol{x}), \boldsymbol{x} \in R^{\boldsymbol{n}}$；
- 坐标方向 $\boldsymbol{e}_j = (0,\cdots,0,1,0,\cdots,0)^T, j = 1,\cdots,n$ ；
- 给定初始步长δ，加速因子α， 任取初始点$\boldsymbol{x}^{(1)}$作为第1个基点，$x^{(j)}$表示第j个基点；
- 在每轮探测中，自变量用$y^{(j)}$表示，即： $y^{(j)}$是沿$\boldsymbol{e}_j$探测的出发点， $y^{(j+1)}$是沿$\boldsymbol{e}_j$探测得到的点。

模式搜索法

■ 基本思想

首先，从$y^{(1)}=x^{(1)}$出发，进行探测移动。

◆ 沿着e_1进行探测

如果（if）$f\left(y^{(1)}+\delta e_1\right)<f(y^{(1)})$，则沿$e_1$探测成功，令

$$y^{(2)}=y^{(1)}+\delta e_1$$

并从$y^{(2)}$出发，沿着e_2进行探测；

否则（else），沿着e_1方向探测失败，再沿着$-e_1$方向进行探测。

◆ 沿着$-e_1$进行探测

如果（if）$f\left(y^{(1)}-\delta e_1\right)<f(y^{(1)})$，则沿$-e_1$探测成功，令

$$y^{(2)}=y^{(1)}-\delta e_1$$

并从$y^{(2)}$出发，沿着e_2进行探测；

模式搜索法

■ 基本思想

否则（**else**），沿着$-e_1$方向也探测失败，令

$$y^{(2)}=y^{(1)}$$

然后，从$y^{(2)}$出发，沿着进行探测移动。

- ◆ 沿着e_2进行探测
- ◆ 沿着$-e_2$进行探测

继续，从$y^{(3)}$出发，沿着进行探测移动。

⋮

如此方式继续下去，直至沿**n**个坐标方向探测完毕，得到$y^{(n+1)}$。

模式搜索法

■ 基本思想

如果 $\mathbf{f}(\boldsymbol{y}^{(n+1)}) < \mathbf{f}(\boldsymbol{y}^{(1)})$,则$\boldsymbol{y}^{(n+1)}$作为新的基点，记为

$$\boldsymbol{x}^{(2)} = \boldsymbol{y}^{(n+1)}$$

则方向$\boldsymbol{d} = \boldsymbol{x}^{(2)} - \boldsymbol{x}^{(1)}$是有利于函数值减小的方向。

下一步，沿方向$\boldsymbol{d}$进行模式移动

令新的$\boldsymbol{y}^{(1)}$为 $\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(2)} + \boldsymbol{\alpha d}$

模式移动后，以新的$\boldsymbol{y}^{(1)}$为起点，进行探测移动，找到$\boldsymbol{y}^{(n+1)}$

如果(**if**) $\mathbf{f}(\boldsymbol{y}^{(n+1)}) < \mathbf{f}(\boldsymbol{x}^{(2)})$,则表明此次模式移动成功，取新的基点$\boldsymbol{x}^{(3)} = \boldsymbol{y}^{(n+1)}$，再沿沿方向$\boldsymbol{d} = \boldsymbol{x}^{(3)} - \boldsymbol{x}^{(2)}$进行模式移动

模式搜索法

■ 基本思想

否则（**else**），则表明模式移动及此次模式移动之后的探测移动均无效。

退回到基点$\boldsymbol{x}^{(2)}$，减小步长$\boldsymbol{\delta}$，再从$\boldsymbol{x}^{(2)}$出发，进行探测移动。如此继续下去，直到步长$\boldsymbol{\delta}$小于事先给定的某个正数$\boldsymbol{\varepsilon}$。

模式搜索法

■ 课堂提问

问题1：模式搜索法包括哪两部分?

探测移动和模式移动

问题2：若模式移动后，探测失败，该如何做?

退回基点，减小步长，继续探测

问题3：模式搜索法的迭代截止条件是什么?

步长小于事先给定的某个正数

模式搜索法

■ 算法步骤

步骤1：给定初始点$\boldsymbol{x}^{(1)} \in R^n$，n个坐标方向$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$，初始步长$\delta$，加速因子$\alpha \geq 1$，缩减率$\beta \in (0,1)$,允许误差$\varepsilon > 0$，置$\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(1)}, k = 1, j = 1$

步骤2：如果$f\left(\boldsymbol{y}^{(j)} + \delta \boldsymbol{e}_j\right) < f(\boldsymbol{y}^{(j)})$，则令$\boldsymbol{y}^{(j+1)} = \boldsymbol{y}^{(j)} + \delta \boldsymbol{e}_j$，进行步骤4；否则，进行步骤3

步骤3：如果$f\left(\boldsymbol{y}^{(j)} - \delta \boldsymbol{e}_j\right) < f(\boldsymbol{y}^{(j)})$，则令$\boldsymbol{y}^{(j+1)} = \boldsymbol{y}^{(j)} - \delta \boldsymbol{e}_j$，进行步骤4；否则，令$\boldsymbol{y}^{(j+1)} = \boldsymbol{y}^{(j)}$，进行步骤4

步骤4：如果$j < n$，则置$j := j + 1$，转步骤2；否则，进行步骤5

模式搜索法

■ 算法步骤

步骤5：如果$\mathbf{f}\left(\boldsymbol{y}^{(n+1)}\right) < \mathbf{f}\left(\boldsymbol{x}^{(k)}\right)$，则进行步骤6；否则，进行步骤7

步骤6：置 $\boldsymbol{x}^{(k+1)} = \boldsymbol{y}^{(n+1)}$，令 $\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(k+1)} + \alpha(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})$，置$k := k+1, j = 1$，转步骤2

步骤7：如果$\delta \leq \epsilon$，则停止迭代，得点$\boldsymbol{x}^{(k)}$； 否则，置$\delta := \beta\delta, \boldsymbol{y}^{(1)} = \boldsymbol{x}^{(k)}, \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$，置$k := k+1, j = 1$，转步骤2

模式搜索法

■ 算法求解问题示例

$$\min f(\boldsymbol{x}) = (1-x_1)^2 + 5(x_2 - x_1^2)^2$$

初始点$\boldsymbol{x}^{(1)} = (2,0)^T$，
坐标方向$\boldsymbol{e}_1 = (1,0)^T, \boldsymbol{e}_2 = (0,1)^T$，
步长$\delta = 1/2$，加速因子$\alpha = 1$，缩减率$\beta = 1/2$，

先在$\boldsymbol{x}^{(1)}$周围进行探测移动，令$\boldsymbol{y}^{(1)} = (2,0)^T$，探测情况如下：

$$f\left(\boldsymbol{y}^{(1)}\right) = \mathbf{81}$$

$$f\left(\boldsymbol{y}^{(1)} + \delta \boldsymbol{e}_1\right) = 197\frac{9}{16} > f(\boldsymbol{y}^{(1)})$$ 失败

$$f\left(\boldsymbol{y}^{(1)} - \delta \boldsymbol{e}_1\right) = 25\frac{9}{16} < f(\boldsymbol{y}^{(1)})$$ 成功

$$\Rightarrow \boldsymbol{y}^{(2)} = \boldsymbol{y}^{(1)} - \delta \boldsymbol{e}_1 = \begin{bmatrix} \frac{3}{2} \\ 0 \end{bmatrix}$$

模式搜索法

■ 算法求解问题示例

从$\boldsymbol{y}^{(2)}$出发，沿着$\boldsymbol{e}_2$探测的情况：

$$f\left(\boldsymbol{y}^{(2)}+\delta\boldsymbol{e}_2\right)=15\frac{9}{16}<f\left(\boldsymbol{y}^{(2)}\right)$$

$$\Rightarrow\quad \boldsymbol{y}^{(3)}=\boldsymbol{y}^{(2)}+\delta\boldsymbol{e}_2=\begin{bmatrix}\frac{3}{2}\\ \frac{1}{2}\end{bmatrix}$$

第1轮探测完成，由于$f\left(\boldsymbol{y}^{(3)}\right)<f\left(\boldsymbol{x}^{(1)}\right)$，得到第2个基点：

$$\boldsymbol{x}^{(2)}=\boldsymbol{y}^{(3)}=\begin{bmatrix}\frac{3}{2}\\ \frac{1}{2}\end{bmatrix}$$

模式搜索法

■ 算法求解问题示例

再沿方向$\boldsymbol{x}^{(2)}-\boldsymbol{x}^{(1)}$进行模式移动，令

$$\boldsymbol{y}^{(1)}=\boldsymbol{x}^{(2)}+\alpha\left(\boldsymbol{x}^{(2)}-\boldsymbol{x}^{(1)}\right)=[1,1]^{\boldsymbol{T}}$$

在$\boldsymbol{y}^{(1)}$周围进行第2轮探测移动，探测情况如下：

$$f\left(\boldsymbol{y}^{(1)}\right)=\mathbf{0}$$

$$f\left(\boldsymbol{y}^{(1)}+\delta\boldsymbol{e}_{\mathbf{1}}\right)=8\frac{1}{16}>f\left(\boldsymbol{y}^{(1)}\right)$$ 失败

$$f\left(\boldsymbol{y}^{(1)}-\delta\boldsymbol{e}_{\mathbf{1}}\right)=3\frac{1}{16}>f\left(\boldsymbol{y}^{(1)}\right)$$ 失败

$$\Longrightarrow\quad \boldsymbol{y}^{(2)}=\boldsymbol{y}^{(1)}=\begin{bmatrix}1\\1\end{bmatrix}$$

模式搜索法

■ 算法求解问题示例

从$\boldsymbol{y}^{(2)}$出发，沿着$\boldsymbol{e}_2$探测的情况：

$$f\left(\boldsymbol{y}^{(2)}\right)=\mathbf{0}$$

$$f\left(\boldsymbol{y}^{(2)}+\delta\boldsymbol{e}_2\right)=1\frac{1}{4}>f(\boldsymbol{y}^{(2)})$$ 失败

$$f\left(\boldsymbol{y}^{(2)}-\delta\boldsymbol{e}_2\right)=1\frac{1}{4}>f(\boldsymbol{y}^{(2)})$$ 失败

$$\boldsymbol{y}^{(3)}=\boldsymbol{y}^{(2)}=\begin{bmatrix}1\\1\end{bmatrix}$$

$$f\left(\boldsymbol{y}^{(3)}\right)=0<f\left(\boldsymbol{x}^{(2)}\right)=15\frac{9}{16}$$

模式移动成功，得到新基点$\boldsymbol{x}^{(3)}=\boldsymbol{y}^{(3)}$

模式搜索法

■ 算法求解问题示例

从$\boldsymbol{x}^{(3)}$出发，沿方向$\boldsymbol{x}^{(3)}-\boldsymbol{x}^{(2)}$进行模式移动.令

$$\boldsymbol{y}^{(1)}=\boldsymbol{x}^{(3)}+\alpha\left(\boldsymbol{x}^{(3)}-\boldsymbol{x}^{(2)}\right)=\left[\frac{1}{2},\ \ \frac{3}{2}\right]^{\boldsymbol{T}}$$

从$\boldsymbol{y}^{(1)}$出发，进行探测移动，会发现，模式移动失败。退回到基点$\boldsymbol{x}^{(3)}$。减小步长，令

$$\delta:=\beta\delta=\frac{1}{4}$$

再从$\boldsymbol{y}^{(1)}=\boldsymbol{x}^{(3)}$，依次沿$\boldsymbol{e}_{\mathbf{1}}$和$\boldsymbol{e}_{\mathbf{2}}$探测，探测失败，必须继续缩减步长，探测移动也是失败的。

$\boldsymbol{x}^{(3)}$即是局部最优解

模式搜索法

■ 讨论

- 不同坐标方向**可以给定不同的步长**；
- 模式移动的方向**可以看成**是**最速下降方向的近似**；
- **收敛速度比较慢**；
- 编制程序比较简单，**适用于变量不多的问题**。

n个变量就是n维，n个探测方向

无约束最优化的直接方法

- **模式搜索法**
- **Rosenbrock法**
- **单纯形搜索法**
- **Powell方法**

Rosenbrock方法

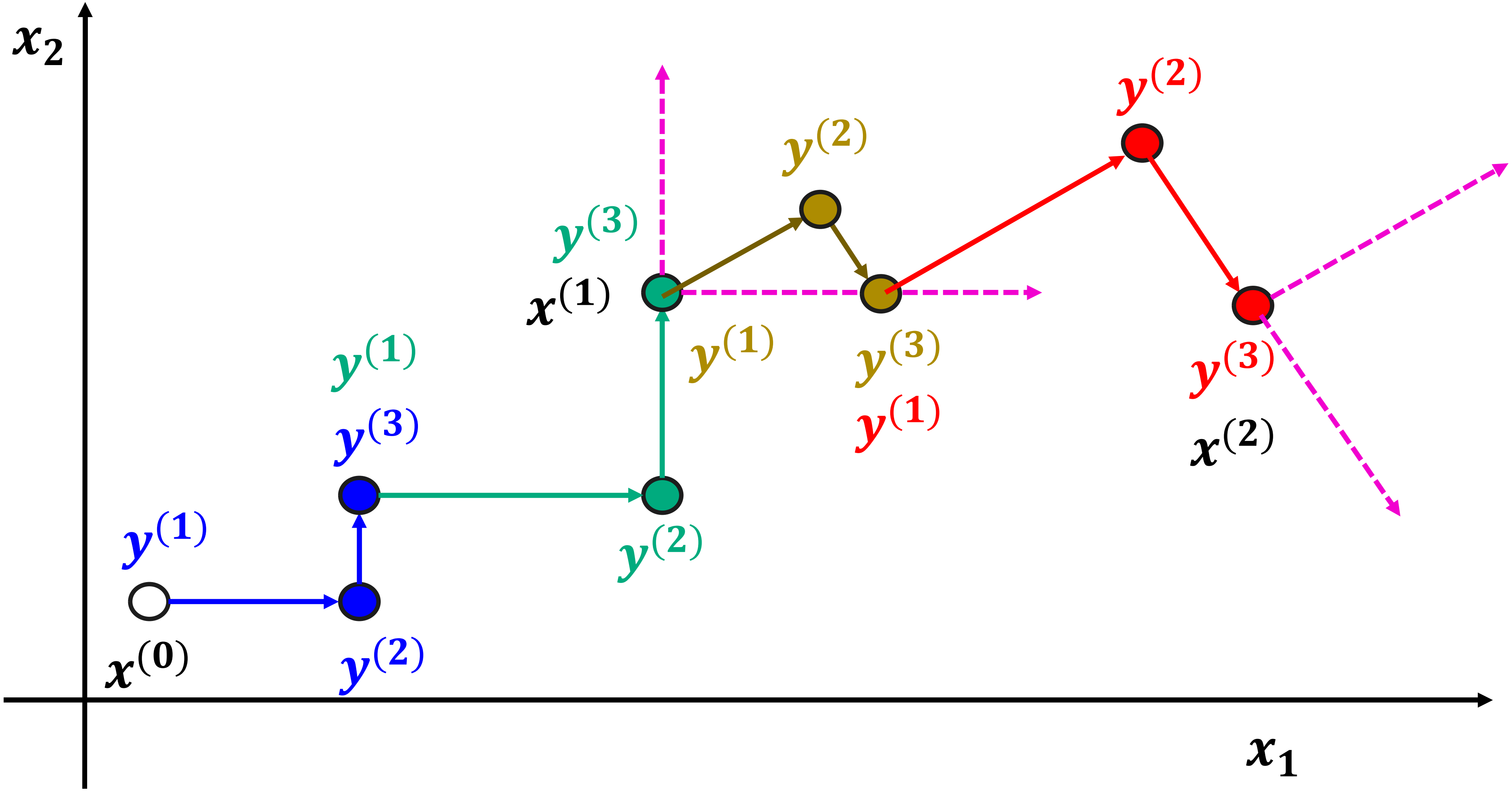
■ 基本思想

◆ 又称转轴法，与模式搜索法类似，也是设法顺着“山谷”求函数的极小点。

◆ 基本思路：每次迭代包括探测阶段和构造搜索方向两部分内容。探测阶段中，从一个点出发，依次沿n个单位正交方向进行探测移动，一轮探测之后，再从第1个方向开始继续探测。经过若干轮探测移动，完成一个探测阶段。然后，构造一组新的单位正交方向，称之为转轴，在下一次迭代中，将沿这些方向进行探测。

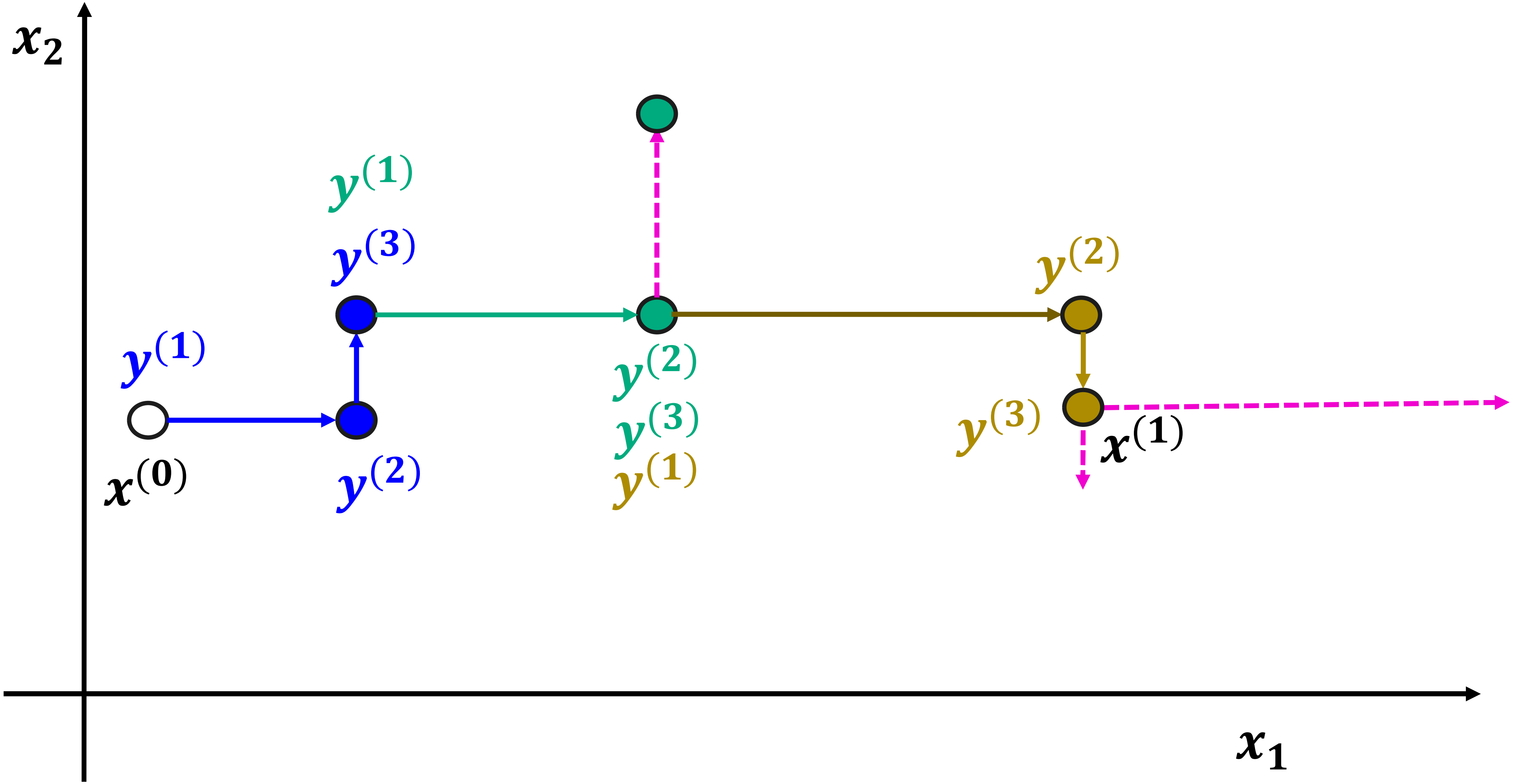
Rosenbrock方法

■ 基本思想



Rosenbrock方法

■ 基本思想



Rosenbrock方法

■ 探测阶段

给定初始点$x^{(1)}$, 放大因子$\alpha > 1$, 缩减因子$\beta \in (-1,0)$
给定初始搜索方向和步长.

设第k次迭代的初始点为$x^{(k)}$, 搜索方向

$$d^{(1)}, d^{(2)}, \dots, d^{(n)}$$

它们是单位正交方向, 沿各方向的步长为

$$\delta_1, \delta_2, \dots, \delta_n$$

每轮探测的起点和终点用$y^{(1)}$ 和$y^{(n+1)}$ 表示.

Rosenbrock方法

■ 探测阶段

令 $y^{(1)} = x^{(k)}$，开始第1轮探测移动

先从 $y^{(1)}$ 出发，沿 $d^{(1)}$ 探测

若 $f(y^{(1)}+\delta_1 d^{(1)}) < f(y^{(1)})$，则探测成功，令

$$y^{(2)} = y^{(1)} + \delta_1 d^{(1)} \qquad (2.1)$$

令 $\delta_1 := \alpha\delta_1$（以备下一轮探测时，沿 $d^{(1)}$ 方向增大步长）

若 $f(y^{(1)}+\delta_1 d^{(1)}) \ge f(y^{(1)})$，则探测失败，令

$$y^{(2)} = y^{(1)} \qquad (2.2)$$

令 $\delta_1 := \beta\delta_1$（下一轮探测时，用 δ_1 乘 $d^{(1)}$，缩短步长）

Rosenbrock方法

■ 探测阶段

再从$y^{(2)}$出发，沿$d^{(2)}$作探测移动. 得$y^{(3)}$. 按此方式探测下去，直至沿$d^{(n)}$探测，得到$y^{(n+1)}$

完成一轮探测后，再来一轮，令 $y^{(1)}= y^{(n+1)}$

进行下一轮探测. 往复下去，至某一轮沿n个方向的探测均失败.

第k次迭代的探测结束时，得到的点记为 $x^{(k+1)}= y^{(n+1)}$

Rosenbrock方法

■ 构造新的搜索方向

用当前的搜索方向

$$d^{(1)},d^{(2)},\ldots,d^{(n)}$$

迭代中得到的数据

$$x^{(k+1)}=x^{(k)}+\sum_{i=1}^{n}\lambda_i d^{(i)}$$

λ_j是沿方向$\boldsymbol{d}^{(i)}$移动的距离

构造线性无关的方向

$$p^{(1)},p^{(2)},\ldots,p^{(n)}$$

Rosenbrock方法

■ 构造新的搜索方向

$$x^{(k+1)} - x^{(k)} = \sum_{i=1}^{n} \lambda_i d^{(i)} \qquad (2.5)$$

记 $p = x^{(k+1)} - x^{(k)}$，有利于函数下降的方向

构造的方向应当包括这个下降方向

定义 $p^{(1)}, p^{(2)}, \ldots, p^{(n)}$ 如下

$$p^{(j)} = \begin{cases} d^{(j)} & \text{当}\lambda_j = 0 \\ \sum_{i=j}^{n} \lambda_i d^{(i)} & \text{当}\lambda_j \neq 0 \end{cases} \qquad (2.6)$$

Rosenbrock方法

■ 构造新的搜索方向

将其正交化

$$q^{(j)} = \begin{cases} \boldsymbol{p}^{(j)}, & j=1 \\ \boldsymbol{p}^{(j)} - \sum_{i=1}^{j-1} \dfrac{\boldsymbol{q}^{(i)T}\boldsymbol{p}^{(j)}}{\boldsymbol{q}^{(i)T}\boldsymbol{q}^{(i)}}\boldsymbol{q}^{(i)}, & j \geq 2 \end{cases} \qquad (2.7)$$

再单位化

$$\bar{\boldsymbol{d}}^{(j)} = \frac{\boldsymbol{q}^{(j)}}{\|\boldsymbol{q}^{(j)}\|} \qquad (2.8)$$

Rosenbrock方法

■ Rosenbrock方法步骤

1, 给定初始点$x^{(1)} \in E^n$, 单位正交方向

$$d^{(1)}, d^{(2)}, \ldots, d^{(n)}$$

步长$\delta_1, \delta_2, \ldots, \delta_n$, 放大因子$\alpha > 1$, 缩减因子$\beta \in (-1, 0)$,

允许误差$\varepsilon > 0$, 置$y^{(1)} = x^{(1)}$, $\delta_i = \delta_i^0, i = 1, 2, \ldots, n$

$k = 1, j = 1$.

2, 若$f(y^{(j)} + \delta_j d^{(j)}) < f(y^{(j)})$, 则令

$$y^{(j+1)} = y^{(j)} + \delta_j d^{(j)}$$

$$\delta_j := \alpha \delta_j$$ 下次探测，步长迈大一些

Rosenbrock方法

■ Rosenbrock方法步骤

若$f(y^{(j)}+\delta_j d^{(j)})\geq f(y^{(j)})$,则令

$$y^{(j+1)}=y^{(j)}$$

$$\delta_j:=\beta\delta_{\mathrm{j}}$$

反方向，步长变小

3. 若j<n, 则置j:=j+1, 转步2, 否则, 进行步4.

4. 若$f(y^{(n+1)})<f(y^{(1)})$, 则令

$$y^{(1)}=\ y^{(\mathrm{n+1})}$$

置j=1, 转步2. ⟶ 再来一轮探测

若$f(y^{(n+1)})=f(y^{(1)})$,则进行步5. 本轮探测没有改善解

Rosenbrock方法

■ Rosenbrock方法步骤

探测失败，步长小于允许值

5.若$f(y^{(n+1)}) < f(x^{(k)})$,则进行步6;否则,若对每个$j$成立$|\delta_j| \le \varepsilon$,则停止计算,$x^{(k)}$作为估计,若不满足终止准则,则令$y^{(1)} = y^{(n+1)}$,置$j=1$,转步2

6. 令$\boldsymbol{x}^{(k+1)} = \boldsymbol{y}^{(n+1)}$ 若

$$\boxed{\left\|x^{(k+1)} - x^{(k)}\right\| \le \varepsilon}$$

则取$\boldsymbol{x}^{(k+1)}$作为极小点的估计,停止计算;否则计算 $\lambda_1, \lambda_2, \ldots, \lambda_n$,

利用公式(2.6)—(2.8)构造新的正交方向

Rosenbrock方法

■ Rosenbrock方法步骤

$$\bar{d}^{(1)}, \bar{d}^{(2)}, \ldots, \bar{d}^{(n)}$$

令 $\boxed{d^{(j)} = \bar{d}^{(j)}, j = 1, 2, \ldots, n}$ 新的单位正交方向

置 $\delta_j = \delta_j^{(0)}, j = 1, 2, \ldots, n$ 步长再次初始化

置$\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(\boldsymbol{k}+1)}$, $\boldsymbol{k} := \boldsymbol{k} + \mathbf{1}$, $\boldsymbol{j} = \mathbf{1}$, 转步2.

Rosenbrock方法

■ 提问环节

1. **Rosenbrock**方法主要包括哪两部分？

2. **Rosenbrock**方法有没有模式移动？

3. **Rosenbrock**方法的搜索方向有什么特点？

4. **Rosenbrock**方法的探测终止条件？

5. **Rosenbrock**方法的迭代搜索方向变化吗？

6. 如何构造**Rosenbrock**方法的搜索方向？

Rosenbrock方法

■ 算法求解问题示例

$$\min f(\boldsymbol{x}) = (x_1 - 3)^2 + 2(x_2 + 2)^2$$

初始点$\boldsymbol{x}^{(1)} = (0,0)^T$，
初始搜索方向$d^{(1)} = (1,0)^T, d^{(2)} = (0,1)^T$，
步长$\delta_1^{(0)} = \delta_2^{(0)} = 1$，$\alpha$ =3，$\beta = -1/2$，
把沿方向$d^{(i)}$进行第j次探测所用步长记作δ_{ij}，这样第1次探测时，$\delta_{11} = \delta_{21} = 1$，下面进行探测移动。

第1轮探测：

初点取为$\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(1)} = [0\ \ 0]^T, f(\boldsymbol{y}^{(1)}) = 17$

先从$\boldsymbol{y}^{(1)}$出发沿$d^{(1)}$探测移动：

$$\boldsymbol{y}^{(1)} + \delta_{11}\,\boldsymbol{d}^{(1)} = [\mathbf{1}\ \ \mathbf{0}]^T$$

$$f\left(\boldsymbol{y}^{(1)} + \delta_{11}\,\boldsymbol{d}^{(1)}\right) = 12 < f(\boldsymbol{y}^{(1)})$$ 探测成功

Rosenbrock方法

■ 算法求解问题示例

令$\delta_{12} = \alpha\delta_{11} = 3$，得$\boldsymbol{y}^{(2)} = \boldsymbol{y}^{(1)} + \delta_{11}\,\boldsymbol{d}^{(1)}$

先从$\boldsymbol{y}^{(2)}$出发沿$d^{(2)}$探测移动：

$$\boldsymbol{y}^{(2)} + \delta_{21}\,\boldsymbol{d}^{(2)} = [\mathbf{1}\ \ \mathbf{1}]^{\boldsymbol{T}}$$

$$f\left(\boldsymbol{y}^{(2)} + \delta_{21}\,\boldsymbol{d}^{(2)}\right) = 22 > f(\boldsymbol{y}^{(2)}) \quad \text{探测失败}$$

令$\delta_{22} = \beta\delta_{21} = -0.5$，得$\boldsymbol{y}^{(3)} = \boldsymbol{y}^{(2)} = [\mathbf{1}\ \ \mathbf{0}]^{\boldsymbol{T}}$

由于$f\left(\boldsymbol{y}^{(3)}\right) = 12 < f\left(\boldsymbol{y}^{(1)}\right) = 17$，因此继续探测。

第2轮探测：

初点取为$\boldsymbol{y}^{(1)} = [1\ \ 0]^{T}$, $f\left(\boldsymbol{y}^{(1)}\right) = 12$，$\delta_{12} = 3$，$\delta_{22} = -0.5$

先从$\boldsymbol{y}^{(1)}$出发沿$d^{(1)}$探测移动：

$$\boldsymbol{y}^{(1)} + \delta_{12}\,\boldsymbol{d}^{(1)} = [\mathbf{4}\ \ \mathbf{0}]^{\boldsymbol{T}}$$

Rosenbrock方法

■ 算法求解问题示例

$$f\left(\boldsymbol{y}^{(1)}+\delta_{12}\,\boldsymbol{d}^{(1)}\right)=9<f(\boldsymbol{y}^{(1)})$$ 探测成功

令$\delta_{13}=\alpha\delta_{12}=9$，得$\boldsymbol{y}^{(2)}=\boldsymbol{y}^{(1)}+\delta_{12}\,\boldsymbol{d}^{(1)}$ $f\left(\boldsymbol{y}^{(2)}\right)=9$

先从$\boldsymbol{y}^{(2)}$出发沿$d^{(2)}$探测移动：

$$\boldsymbol{y}^{(2)}+\delta_{22}\,\boldsymbol{d}^{(2)}=[\mathbf{4}\quad-\mathbf{0.5}]^{T}$$

$$f\left(\boldsymbol{y}^{(2)}+\delta_{22}\,\boldsymbol{d}^{(2)}\right)=5.5<f(\boldsymbol{y}^{(2)})$$ 探测成功

令$\delta_{23}=\alpha\delta_{22}=-1.5$，得$\boldsymbol{y}^{(3)}=\boldsymbol{y}^{(2)}+\delta_{22}\,\boldsymbol{d}^{(2)}=[\mathbf{4}\quad-\mathbf{0.5}]^{T}$

由于$f\left(\boldsymbol{y}^{(3)}\right)=5.5<f\left(\boldsymbol{y}^{(1)}\right)=12$，因此继续探测。

第3轮探测：

初点取为$\boldsymbol{y}^{(1)}=[4\quad-0.5]^{T}$, $f\left(\boldsymbol{y}^{(1)}\right)=5.5$， $\delta_{12}=9$， $\delta_{22}=-1.5$

Rosenbrock方法

■ 算法求解问题示例

先从$\boldsymbol{y}^{(1)}$出发沿$d^{(1)}$探测移动：

$$f\left(\boldsymbol{y}^{(1)}+\delta_{13}\,\boldsymbol{d}^{(1)}\right)=104.5>f(\boldsymbol{y}^{(1)})$$ 探测失败

令$\delta_{14}=\beta\delta_{13}=-4.5$，得$\boldsymbol{y}^{(2)}=\boldsymbol{y}^{(1)}=[\mathbf{4}\quad -\mathbf{0.5}]^{T}$

先从$\boldsymbol{y}^{(2)}$出发沿$d^{(2)}$探测移动：

$$\boldsymbol{y}^{(2)}+\delta_{23}\,\boldsymbol{d}^{(2)}=[\mathbf{4}\quad -\mathbf{2}]^{T}$$

$$f\left(\boldsymbol{y}^{(2)}+\delta_{23}\,\boldsymbol{d}^{(2)}\right)=1<f(\boldsymbol{y}^{(2)})$$ 探测成功

令$\delta_{24}=\alpha\delta_{23}=-4.5$，得$\boldsymbol{y}^{(3)}=\boldsymbol{y}^{(2)}+\delta_{22}\,\boldsymbol{d}^{(2)}=[\mathbf{4}\quad -\mathbf{0.5}]^{T}$

由于$f\left(\boldsymbol{y}^{(3)}\right)=1<f\left(\boldsymbol{y}^{(1)}\right)=5.5$，因此继续探测。

Rosenbrock方法

■ 算法求解问题示例

第4轮探测：

初点取为$\boldsymbol{y}^{(1)} = [4 \quad -2]^T$，$f(\boldsymbol{y}^{(1)}) = 1$， $\delta_{14} = -4.5$，$\delta_{24} = -4.5$

先从$\boldsymbol{y}^{(1)}$出发沿$d^{(1)}$探测移动：

$$\boldsymbol{y}^{(1)} + \delta_{14}\, \boldsymbol{d}^{(1)} = [-\mathbf{0.5} \quad -\mathbf{2}]^T$$

$$f(\boldsymbol{y}^{(1)} + \delta_{14}\, \boldsymbol{d}^{(1)}) = 12.25 > f(\boldsymbol{y}^{(1)})$$ 探测失败

令$\delta_{15} = \beta\delta_{14} = 2.25$，得$\boldsymbol{y}^{(2)} = \boldsymbol{y}^{(1)} = [\mathbf{4} \quad -\mathbf{2}]^T$

先从$\boldsymbol{y}^{(2)}$出发沿$d^{(2)}$探测移动：

$$\boldsymbol{y}^{(2)} + \delta_{24}\, \boldsymbol{d}^{(2)} = [\mathbf{4} \quad -\mathbf{6.5}]^T$$

$$f(\boldsymbol{y}^{(2)} + \delta_{24}\, \boldsymbol{d}^{(2)}) = 41.5 > f(\boldsymbol{y}^{(2)}) = 1$$ 探测失败

令$\delta_{25} = \beta\delta_{24} = 2.25$，得$\boldsymbol{y}^{(3)} = \boldsymbol{y}^{(2)} = [\mathbf{4} \quad -\mathbf{2}]^T$

Rosenbrock方法

算法求解问题示例

第4轮探测在两个方向均失败.比较$\boldsymbol{y}^{(3)}$和本次迭代初始点$\boldsymbol{x}^{(1)}$处的函数值，$f\left(\boldsymbol{y}^{(3)}\right)=1<f\left(\boldsymbol{x}^{(1)}\right)=17$

经4轮探测完成这次迭代的探测阶段，令

$$\boldsymbol{x}^{(2)}=\boldsymbol{y}^{(3)}=[4\quad -2]^T$$

$\boldsymbol{x}^{(2)}$是下次迭代的初始点。

下面构造一组新的单位正交方向，先求步长的代数和：

$$\lambda_1=\delta_{11}+\delta_{12}=1+3=4$$

$$\lambda_2=\delta_{22}+\delta_{23}=-0.5-1.5=-2$$

然后，令

$$\boldsymbol{p}^{(1)}=\lambda_1 d^{(1)}+\lambda_2\, d^{(2)}=[4\quad -2]^T\quad \boldsymbol{p}^{(2)}=\lambda_2\, d^{(2)}=[0\quad -2]^T$$

Rosenbrock方法

■ 算法求解问题示例

将$\boldsymbol{p}^{(1)}$和$\boldsymbol{p}^{(2)}$正交化，令

$$\boldsymbol{q}^{(1)} = \boldsymbol{p}^{(1)} = [4 \quad -2]^T$$

$$\boldsymbol{q}^{(2)} = \boldsymbol{p}^{(2)} - \frac{\boldsymbol{q}^{(1)T}\boldsymbol{p}^{(2)}}{\boldsymbol{q}^{(1)T}\boldsymbol{q}^{(1)}} = [-4/5 \quad -8/5]^T$$

将$\boldsymbol{q}^{(1)}$和$\boldsymbol{q}^{(2)}$正交化，令

$$\boldsymbol{d}^{(1)} = \left[2/\sqrt{5} \quad -1/\sqrt{5}\right]^T$$

沿新构造的方向继续探索

$$\boldsymbol{d}^{(2)} = \left[-1/\sqrt{5} \quad -2/\sqrt{5}\right]^T$$

$$\boldsymbol{x}^* = [3 \quad -2]^T$$

无约束最优化的直接方法

- 模式搜索法
- Rosenbrock法
- 单纯形搜索法
- Powell方法

单纯形搜索法

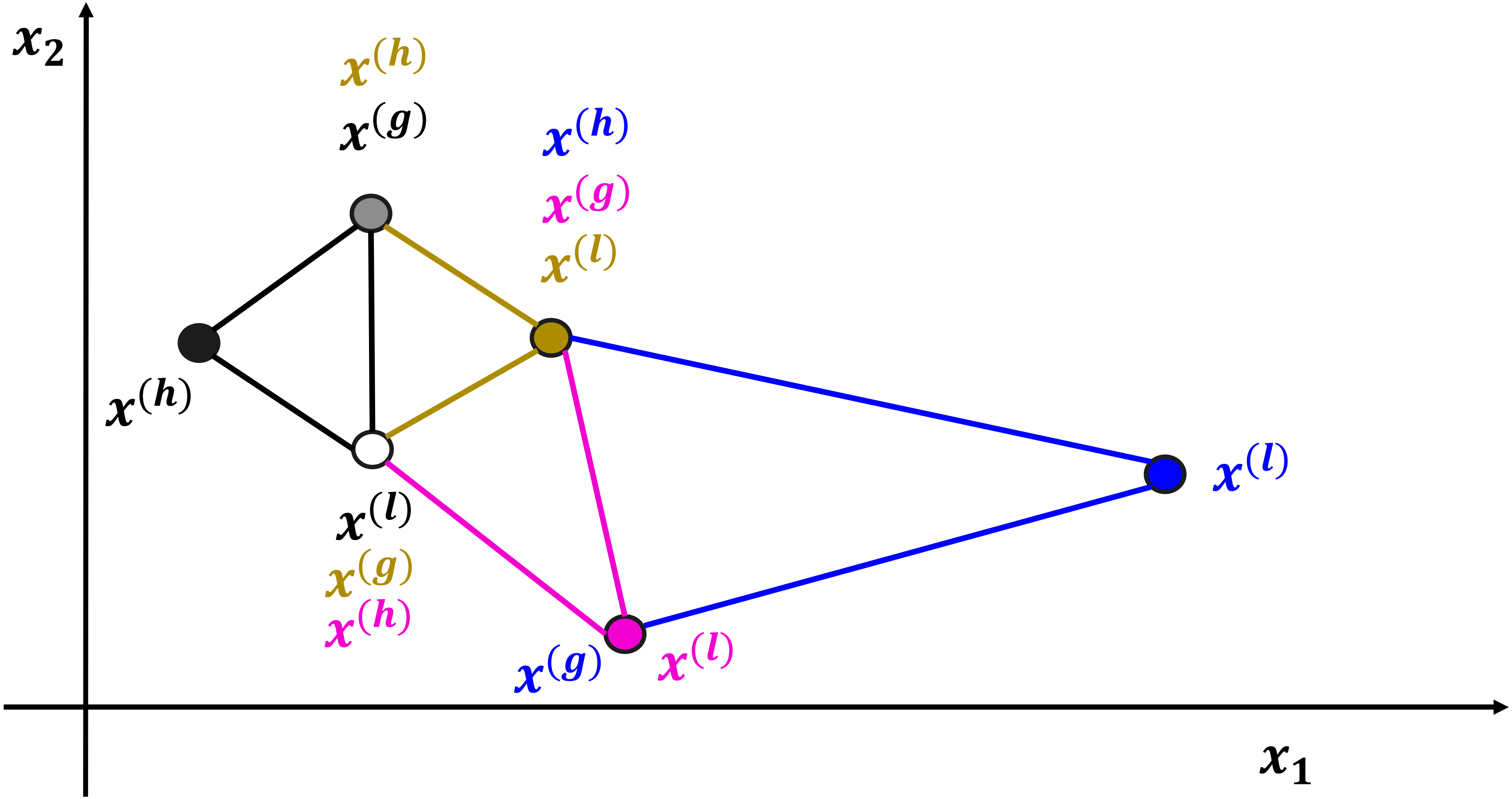
■ 基本思想

◆ **单纯形**：n维空间中具有**n+1个顶点的凸多面体**。比如一维空间的线段，二维空间的三角形，三维空间的四面体。

◆ **基本思路**：给定R^n中一个单纯形后，**求出n+1个顶点上的函数值**，确定出有最大函数值的点（称为**最高点**）和有最小函数值的点（称为**最低点**），然后通过**反射**、**扩展**、**压缩等**方法（几种方法**不一定同时使用**），求出一个最好点，用它取代最高点，构成新的单纯形，或者**向最低点收缩**形成新的单纯形，用如此方法逼近极小点。

单纯形搜索法

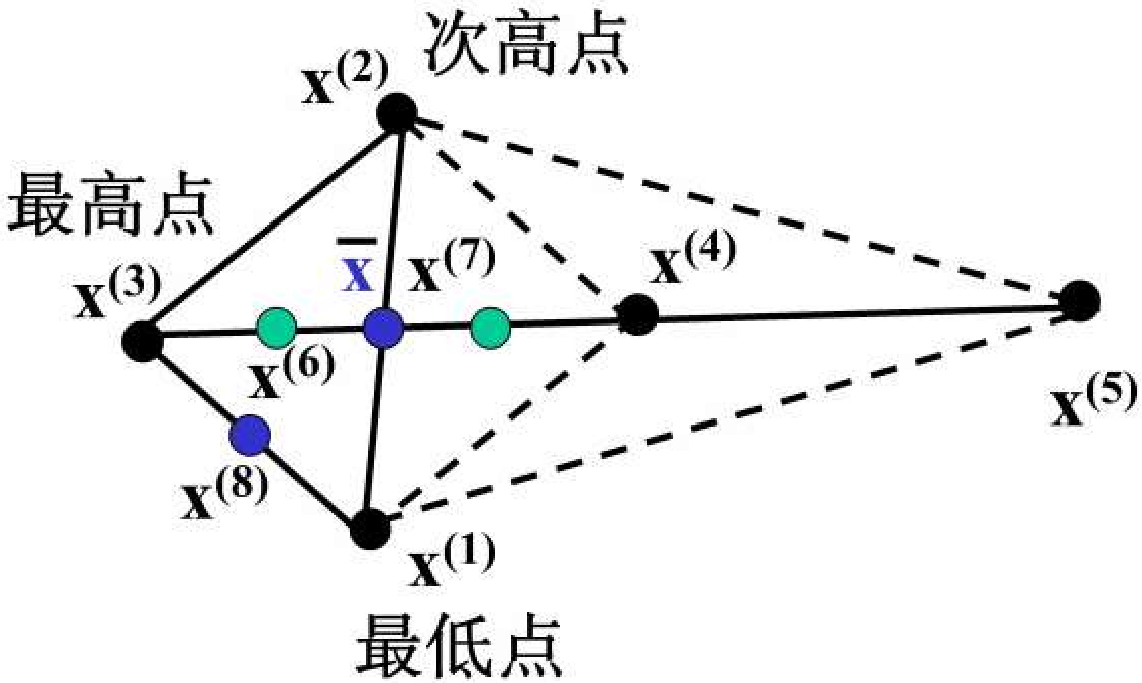
■ 基本思想



单纯形搜索法

■ 单纯形转换示例

◆ 以**极小化二元函数**$f(x_1, x_2)$为例，说明如何实现单纯形转换

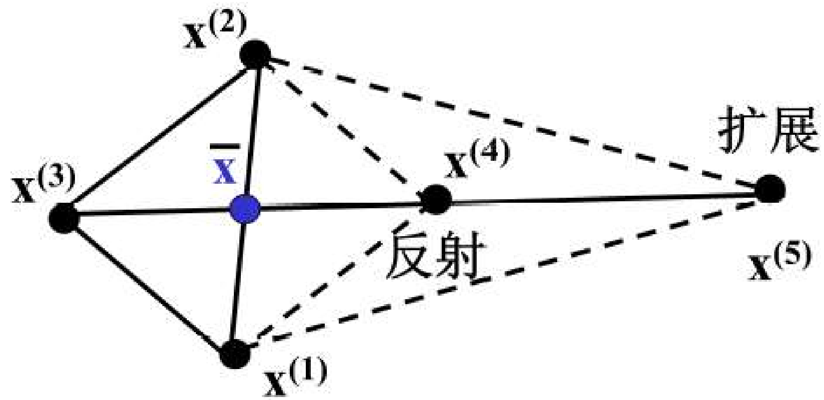


平面上取不共线的三点$x^{(1)}, x^{(2)}, x^{(3)}$构成单纯形，$x^{(3)}$为最高点，$x^{(1)}$为最低点。

单纯形搜索法

■ 单纯形转换示例

反射：将最高点经过其余点的形心进行反射。



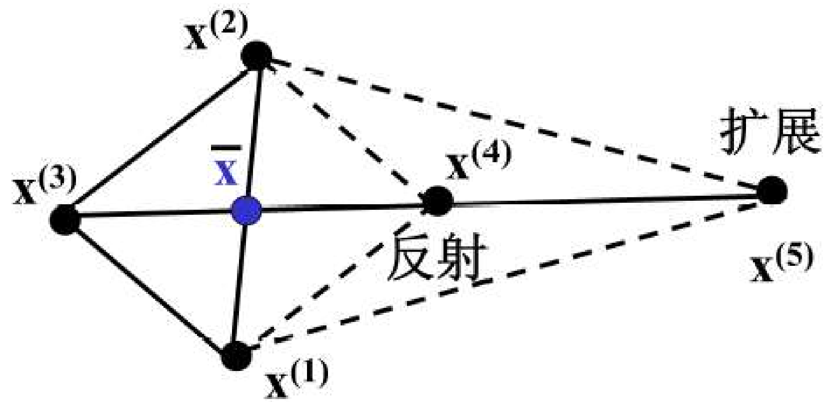
对于本问题，就是将$x^{(3)}$经过线段$x^{(1)}$ $x^{(2)}$的中点

$\bar{x}=\frac{x^{(1)}+x^{(2)}}{2}$进行反射。

得到反射点$x^{(4)}=\bar{x}+\alpha(\bar{x}-x^{(3)})$，正数$\alpha$为反射系数，一般取$\alpha=1$

单纯形搜索法

■ 单纯形转换示例



反射后有三种可能的情形

(1) $f(x^{(4)}) < f(x^{(1)})$

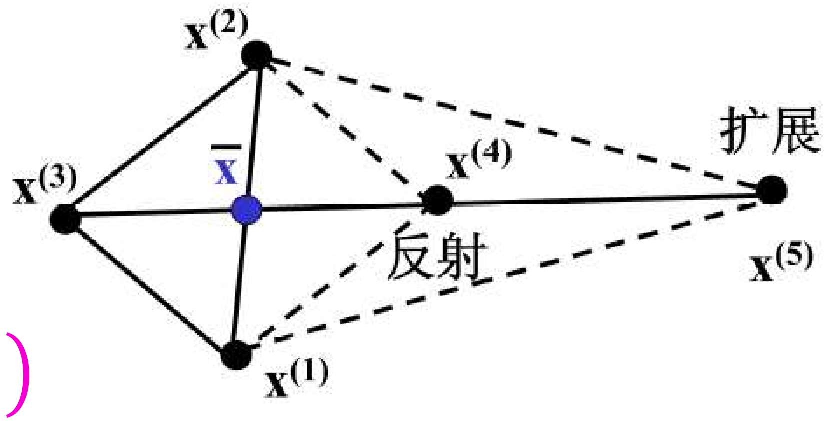
表明方向$d = x^{(4)} - \bar{x}$有利于减少函数值，于是沿该方向进行扩展。令$x^{(5)} = \bar{x} + \gamma(x^{(4)} - \bar{x}), \gamma > 1$为扩展系数。

若$f(x^{(5)}) < f(x^{(4)})$，则用$x^{(5)}$取代$x^{(3)}$，得到以$x^{(1)}$、$x^{(2)}$、$x^{(5)}$为顶点的新的单纯形。

若$f(x^{(5)}) \geq f(x^{(4)})$，扩展失败，则用$x^{(4)}$取代$x^{(3)}$，得到以$x^{(1)}$、$x^{(2)}$、$x^{(4)}$为顶点的新的单纯形。

单纯形搜索法

■ 单纯形转换示例



(2) $f\left(x^{(1)}\right) \leq f\left(x^{(4)}\right) \leq f\left(x^{(2)}\right)$

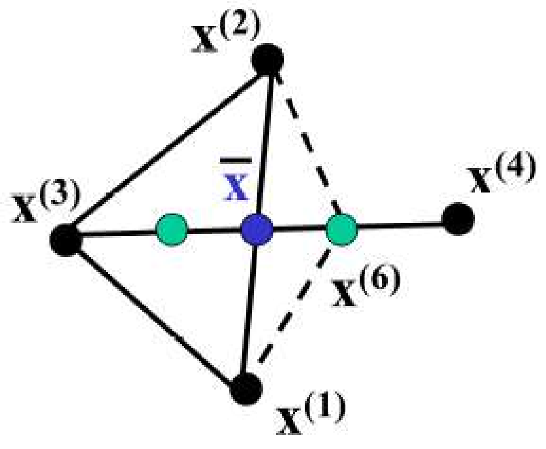
表明$f\left(x^{(4)}\right)$不小于最低点处的函数值，不大于次高点的函数值，用$x^{(4)}$取代$x^{(3)}$，得到以$x^{(1)}$、$x^{(2)}$、$x^{(4)}$为顶点的新的单纯形.

(3) $f\left(x^{(4)}\right) > f\left(x^{(2)}\right)$

即$f\left(x^{(4)}\right)$大于次高点处的函数值，则进行压缩步骤。为此，在$x^{(4)}$取代$x^{(3)}$中选择函数值最小的点，令

单纯形搜索法

■ 单纯形转换示例



压缩

(3) $f(x^{(4)}) > f(x^{(2)})$

$$f(x^{(h')}) = \min\{f(x^{(3)}), f(x^{(4)})\}, x^{(h')} \in \{x^{(3)}, x^{(4)}\}$$

令$x^{(6)} = \bar{x} + \beta(x^{(h')} - \bar{x}), \beta \in (0,1)$为**压缩**系数，于是$x^{(6)}$位于$\bar{x}$和$x^{(h')}$之间。

若$f(x^{(6)}) \le f(x^{(h')})$，则用$x^{(6)}$取代$x^{(3)}$，得到以$x^{(1)}$、$x^{(2)}$、 $x^{(6)}$为顶点的新的单纯形。

若$f(x^{(6)}) > f(x^{(h')})$，则进行**收缩**。最低点$x^{(1)}$不动，$x^{(2)}$和$x^{(3)}$均向$x^{(1)}$移近一半距离。

单纯形搜索法

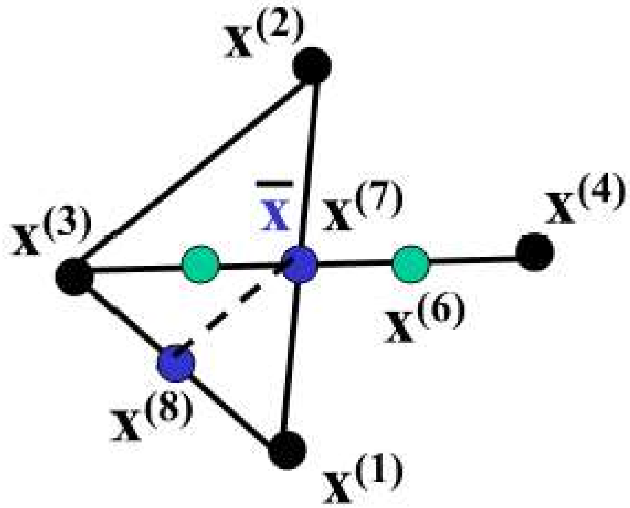
■ 单纯形转换示例

(3) $f\left(x^{(4)}\right) > f\left(x^{(2)}\right)$

令

$$x^{(7)} = x^{(2)} + \frac{1}{2}\left(x^{(1)} - x^{(2)}\right),$$

$$x^{(8)} = x^{(2)} + \frac{1}{2}\left(x^{(1)} - x^{(3)}\right),$$



收缩边长

得到以$x^{(1)}$、$x^{(7)}$、$x^{(8)}$为顶点的新的单纯形。

以上三种情况得到的新的单纯形，必有一个顶点的函数值小于或等于原单纯形各顶点上的函数值。如上，不断得到新的单纯形，直至满足收敛准则。

单纯形搜索法

■ 课堂提问

问题1：什么叫单纯形?

n维空间中具有n+1个顶点的凸多面体

问题2：单纯形搜索法有哪几种模式?

反射、扩展、压缩、收缩

问题3：单纯形搜索法如何进行收缩?

最低点$x^{(l)}$不动，$x^{(g)}$和$x^{(h)}$均向$x^{(l)}$移近一半距离

单纯形搜索法

■ 算法步骤

步骤1：给定初始点单纯形，其顶点$\boldsymbol{x}^{(i)} \in E^n, i = 1,2,\cdots,n+1$，反射系数$\alpha \geq 0$，扩展系数$\gamma > 1$，压缩系数$\beta \in (0,1)$，允许误差$\varepsilon > 0$，计算$f(\boldsymbol{x}^{(i)})$，置$k = 1$

步骤2：确定最高点$\boldsymbol{x}^{(h)}$，次高点$\boldsymbol{x}^{(g)}$，最低点$\boldsymbol{x}^{(l)}$，$h, g, l \in \{1,2,\cdots,n+1\}$，使得

$$f\left(\boldsymbol{x}^{(h)}\right) = \max\{f\left(\boldsymbol{x}^{(1)}\right), f\left(\boldsymbol{x}^{(2)}\right), \cdots, f\left(\boldsymbol{x}^{(n+1)}\right)\}$$

$$f\left(\boldsymbol{x}^{(g)}\right) = \max\{f\left(\boldsymbol{x}^{(i)}\right), \boldsymbol{x}^{(i)} \neq \boldsymbol{x}^{(h)}\}$$

$$f\left(\boldsymbol{x}^{(l)}\right) = \min\{f\left(\boldsymbol{x}^{(1)}\right), f\left(\boldsymbol{x}^{(2)}\right), \cdots, f\left(\boldsymbol{x}^{(n+1)}\right)\}$$

计算$\boldsymbol{x}^{(h)}$外的**n**个点的形心$\bar{\boldsymbol{x}} = \frac{\mathbf{1}}{\boldsymbol{n}}\left[\sum_{i=1}^{n+1} \boldsymbol{x}^{(i)} - \boldsymbol{x}^{(h)}\right]$和$f(\bar{\boldsymbol{x}})$

单纯形搜索法

算法步骤

步骤3：进行反射，令$x^{(n+2)}=\bar{x}+\alpha(\bar{x}-x^{(h)})$，计算$f\left(x^{(n+2)}\right)$

步骤4：若 $f\left(x^{(n+2)}\right)<f\left(x^{(l)}\right)$，则进行扩展，令$x^{(n+3)}=\bar{x}+\gamma(x^{(n+2)}-\bar{x})$，计算$f\left(x^{(n+3)}\right)$，转步骤5

若$f\left(x^{(l)}\right)\le f\left(x^{(n+2)}\right)\le f\left(x^{(g)}\right)$，则置$x^{(h)}=x^{(n+2)}$，$f\left(x^{(h)}\right)=f\left(x^{(n+2)}\right)$， 转步骤7

若$f\left(x^{(n+2)}\right)>f\left(x^{(g)}\right)$，则进行压缩，令

$$f\left(x^{(h')}\right)=\min\{f\left(x^{(h)}\right),f\left(x^{(n+2)}\right)\},h'\in\{h,n+2\}$$

令$x^{(n+4)}=\bar{x}+\beta(x^{(h')}-\bar{x})$，计算$f\left(x^{(n+4)}\right)$，转步骤6

单纯形搜索法

■ 算法步骤

步骤5：若$f\left(x^{(n+3)}\right) < f\left(x^{(n+2)}\right)$,则置$x^{(h)} = x^{(n+3)}$，$f\left(x^{(h)}\right) = f\left(x^{(n+3)}\right)$，转步骤7；否则，置$x^{(h)} = x^{(n+2)}$，$f\left(x^{(h)}\right) = f\left(x^{(n+2)}\right)$，转步骤7

步骤6：若$f\left(x^{(n+4)}\right) \le f\left(x^{(h')}\right)$,则置$x^{(h)} = x^{(n+4)}$，$f\left(x^{(h)}\right) = f\left(x^{(n+4)}\right)$，转步骤7；否则，进行收缩，令$x^{(i)} := x^{(i)} + \frac{1}{2}\left(x^{(l)} - x^{(i)}\right), i = 1, 2, \cdots, n+1$，计算$f\left(x^{(i)}\right)$

步骤7：检验是否满足收敛准则。若$\left\{\frac{1}{n+1}\sum_{i=1}^{n+1}\left[f\left(x^{(i)}\right) - f(\overline{x})\right]^2\right\}^{\frac{1}{2}} < \epsilon$ 停止计算；现行最好点最好点即是极小点的近似，否则，$k := k+1$，转步骤2

单纯形搜索法

■ 算法求解问题示例

$$\min f(\boldsymbol{x}) = (x_1 - 3)^2 + 2(x_2 + 2)^2$$

取初始单纯形的顶点为$\boldsymbol{x}^{(1)} = (0,0)^T$， $\boldsymbol{x}^{(2)} = (1,0)^T$，$\boldsymbol{x}^{(3)} = (0,1)^T$

$\alpha = 1$，$\gamma = 1$，$\beta = 1/2$，$\epsilon = 1$

第一次迭代

$$f\left(\boldsymbol{x}^{(1)}\right) = 17,\quad f\left(\boldsymbol{x}^{(2)}\right) = 12,\quad f\left(\boldsymbol{x}^{(3)}\right) = 27$$

有$\boldsymbol{x}^{(h)} = \boldsymbol{x}^{(3)}$， $\boldsymbol{x}^{(g)} = \boldsymbol{x}^{(1)}$， $\boldsymbol{x}^{(l)} = \boldsymbol{x}^{(2)}$

除$\boldsymbol{x}^{(3)}$外的形心$\overline{\boldsymbol{x}} = \frac{1}{2}\left(\boldsymbol{x}^{(1)} + \boldsymbol{x}^{(2)}\right) = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$， $f(\overline{\boldsymbol{x}}) = \frac{57}{4}$

单纯形搜索法

■ 算法求解问题示例

将$\boldsymbol{x}^{(3)}$经形心进行反射，令

$$\boldsymbol{x}^{(4)}=\overline{\boldsymbol{x}}+\alpha\left(\overline{\boldsymbol{x}}-\boldsymbol{x}^{(3)}\right)$$

$$=2\overline{\boldsymbol{x}}+\alpha\boldsymbol{x}^{(3)}=\begin{bmatrix}1\\-1\end{bmatrix}\qquad f\left(\boldsymbol{x}^{(4)}\right)=\boldsymbol{6}$$

$f\left(\boldsymbol{x}^{(4)}\right)<\boldsymbol{f}(\boldsymbol{x}^{(l)})$，因此进行扩展，令

$$\boldsymbol{x}^{(5)}=\overline{\boldsymbol{x}}+\boldsymbol{\gamma}\left(\boldsymbol{x}^{(4)}-\overline{\boldsymbol{x}}\right)=2\boldsymbol{x}^{(4)}-\overline{\boldsymbol{x}}=\begin{bmatrix}\frac{3}{2}\\-2\end{bmatrix}f\left(\boldsymbol{x}^{(5)}\right)=\frac{\boldsymbol{9}}{\boldsymbol{4}}$$

$f\left(\boldsymbol{x}^{(5)}\right)<f\left(\boldsymbol{x}^{(4)}\right)$，因此用$\boldsymbol{x}^{(5)}$替代$\boldsymbol{x}^{(\mathrm{h})}$ 最高点

得新的单纯形

$\boldsymbol{x}^{(1)}=(0,0)^T$， $\boldsymbol{x}^{(2)}=(1,0)^T$， $\boldsymbol{x}^{(3)}=(3/2,-2)^T$

单纯形搜索法

■ 算法求解问题示例

$$\left\{\frac{1}{3}\sum_{i=1}^{3}\left[f\left(x^{(i)}\right)-f(\overline{x})\right]^{2}\right\}^{\frac{1}{2}}=7.23>\epsilon$$

第二次迭代

$$x^{(h)}=x^{(1)},\quad x^{(g)}=x^{(2)},\quad x^{(l)}=x^{(3)}$$

$$f\left(x^{(1)}\right)=17,\quad f\left(x^{(2)}\right)=12,\quad f\left(x^{(3)}\right)=\frac{9}{4}$$

将$x^{(1)}$经形心进行反射，令

$$\overline{x}=\frac{1}{2}\left(x^{(2)}+x^{(3)}\right)=\begin{bmatrix}\frac{5}{4}\\-1\end{bmatrix},\quad f(\overline{x})=\frac{81}{16}$$

反射点

$$x^{(4)}=\overline{x}+\alpha\left(\overline{x}-x^{(h)}\right)=2\overline{x}-x^{(1)}=\begin{bmatrix}\frac{5}{2}\\-2\end{bmatrix}\quad f\left(x^{(4)}\right)=\frac{1}{4}$$

单纯形搜索法

■ 算法求解问题示例

$f\left(x^{(4)}\right) < f(x^{(l)})$，因此进行扩展，令

$$x^{(5)} = \bar{x} + \gamma\left(x^{(4)} - \bar{x}\right) = 2x^{(4)} - \bar{x} = \begin{bmatrix} \frac{15}{4} \\ -3 \end{bmatrix}$$

$f\left(x^{(5)}\right) > f\left(x^{(4)}\right)$，因此用$x^{(4)}$替代$x^{(h)}$ 最高点

$$\left\{\frac{1}{3}\sum_{i=1}^{3}\left[f\left(x^{(i)}\right) - f(\bar{x})\right]^2\right\}^{\frac{1}{2}} = 5.14 > \epsilon$$

第三次迭代

$$x^{(h)} = x^{(2)}, \quad x^{(g)} = x^{(3)}, \quad x^{(l)} = x^{(1)}$$

$$\bar{x} = \frac{1}{2}\left(x^{(1)} + x^{(3)}\right) = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad f(\bar{x}) = 1$$

单纯形搜索法

■ 算法求解问题示例

反射点$\boldsymbol{x}^{(4)}=\overline{\boldsymbol{x}}+\boldsymbol{\alpha}\left(\overline{\boldsymbol{x}}-\boldsymbol{x}^{(2)}\right)=\begin{bmatrix}3\\-4\end{bmatrix}$

$f\left(\boldsymbol{x}^{(4)}\right)>f\left(\boldsymbol{x}^{(3)}\right)$，因此进行压缩

$$\boldsymbol{f}\left(\boldsymbol{x}^{(4)}\right)=\min\left\{\boldsymbol{f}\left(\boldsymbol{x}^{(h)}\right),\boldsymbol{f}\left(\boldsymbol{x}^{(4)}\right)\right\}$$

令$\boldsymbol{x}^{(6)}=\overline{\boldsymbol{x}}+\boldsymbol{\beta}\left(\boldsymbol{x}^{(4)}-\overline{\boldsymbol{x}}\right)=\begin{bmatrix}\frac{5}{2}\\-3\end{bmatrix}$

$f\left(\boldsymbol{x}^{(6)}\right)<f\left(\boldsymbol{x}^{(4)}\right)$，因此用$\boldsymbol{x}^{(6)}$替代$\boldsymbol{x}^{(h)}$ 最高点

得新的单纯形 $\boldsymbol{x}^{(h)}=\boldsymbol{x}^{(6)}$， $\boldsymbol{x}^{(g)}=\boldsymbol{x}^{(3)}$ ， $\boldsymbol{x}^{(l)}=\boldsymbol{x}^{(1)}$

$$\left\{\frac{1}{3}\sum_{i=1}^{3}\left[\boldsymbol{f}\left(\boldsymbol{x}^{(i)}\right)-\boldsymbol{f}(\overline{\boldsymbol{x}})\right]^2\right\}^{\frac{1}{2}}=\mathbf{1.11}<\boldsymbol{\epsilon}$$

得近似解$\boldsymbol{x}^{(1)}=\left(\frac{5}{2},-2\right)^{T}$，而$\boldsymbol{x}^{*}=(3,-2)^{T}$

单纯形搜索法

■ 讨论

- **最初的单纯形方法称为正规单纯形方法；**

  正规单纯形：n+1个顶点中任意两点间的距离都相等

- **以上单纯形搜索法是对正规单纯形方法的改进；**

- **对于变量较多的情形，比如自变量超过10个时，单纯形搜索法效果很差。**

无约束最优化的直接方法

- **模式搜索法**
- **Rosenbrock法**
- **单纯形搜索法**
- **Powell方法**

Powell方法

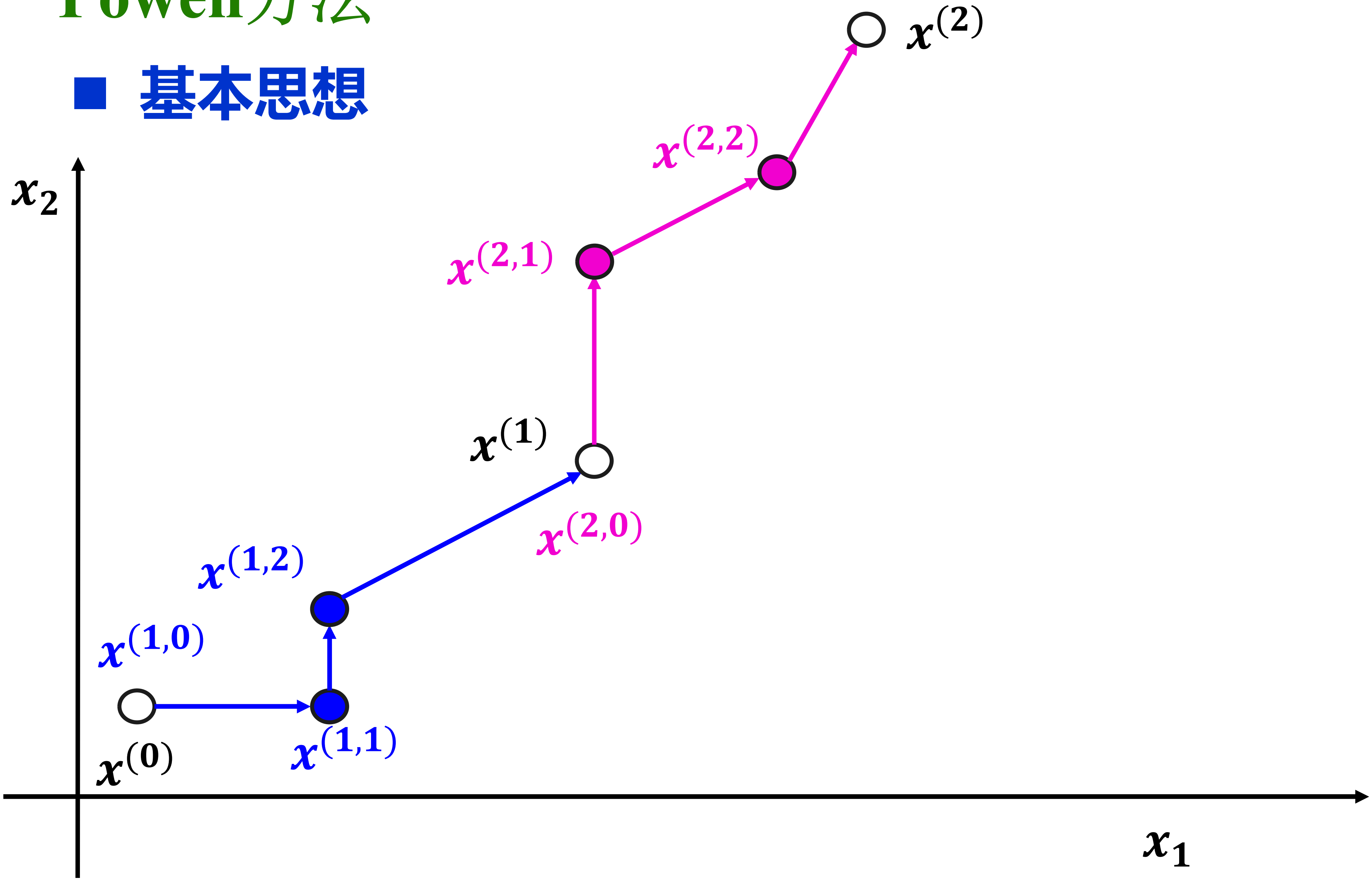
■ 基本思想

◆ 是一种有效的直接搜索方法，本质上是共轭方向法。

◆ 基本思路：把整个计算过程分成若干个阶段，每一阶段（一轮迭代）由n+1次一维搜索组成。在算法的每一阶段中，先依次沿着已知的n个方向搜索，得一个最好点，然后沿本阶段的初点与该最好点连线方向进行搜索，求得这一阶段的最好点。再用最后的搜索方向取代前n个方向之一，开始下一阶段的迭代。

Powell方法

■ 基本思想



Powell方法

■ 算法步骤

步骤1：给定初始点$x^{(0)}$，n个线性无关的方向$\boldsymbol{d}^{(1,1)},\boldsymbol{d}^{(1,2)},\cdots,\boldsymbol{d}^{(1,n)}$，允许误差$\varepsilon>0$，置$k=1$

步骤2：置$x^{(k,0)}=x^{(k-1)}$，从$x^{(k,0)}$出发，依次沿方向$\boldsymbol{d}^{(k,1)},\boldsymbol{d}^{(k,2)},\cdots,\boldsymbol{d}^{(k,n)}$进行搜索，得到点$\boldsymbol{x}^{(k,1)},\boldsymbol{x}^{(k,2)},\cdots,\boldsymbol{x}^{(k,n)}$，再从$\boldsymbol{x}^{(k,n)}$出发，沿着方向

$$d^{(k,n+1)}=x^{(k,n)}-x^{(k,0)}$$

作一维搜索，得到点$x^{(k)}$。

步骤3：若$\left\|x^{(k)}-x^{(k-1)}\right\|<\epsilon$，则停止计算，得点$x^{(k)}$；否则，令

$$d^{(k+1,j)}=d^{(k,j+1)},j=1,\cdots,n.$$

置$k:=k+1$，返回步骤2

Powell法

■ 课堂提问

问题1：Powell方法对搜索方向有什么要求?

线性无关

问题2：Powell方法需要用的一维搜索吗?

需要

问题3： Powell方法的迭代截止条件是什么?

$$\|x^{(k)} - x^{(k-1)}\| < \epsilon$$

Powell方法

■ 算法求解问题示例

$$\min f(\boldsymbol{x}) = (x_1 + x_2)^2 + (x_1 - 1)^2$$

取初始点和初始搜索方向分别为

$$\boldsymbol{x}^{(0)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \boldsymbol{d}^{(1,1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{d}^{(1,2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

第一轮迭代

置 $\boldsymbol{x}^{(1,0)} = \boldsymbol{x}^{(0)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

先沿$\boldsymbol{d}^{(1,1)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(1,0)} + \lambda \boldsymbol{d}^{(1,1)})$$

$$\boldsymbol{x}^{(1,0)} + \lambda \boldsymbol{d}^{(1,1)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 + \lambda \\ 1 \end{bmatrix}$$

Powell方法

■ 算法求解问题示例

令 $\varphi(\lambda) = f\left(\boldsymbol{x}^{(1,0)} + \lambda \boldsymbol{d}^{(1,1)}\right) = (3+\lambda)^2 + (1+\lambda)^2$

$$\frac{d\varphi}{d\lambda} = 2(3+\lambda) + 2(1+\lambda)=0$$

得到$\lambda_1 = -2$，$\boldsymbol{x}^{(1,1)} = (0,1)^T$.

再从$\boldsymbol{x}^{(1,1)}$出发，沿$\boldsymbol{d}^{(1,2)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(1,1)} + \lambda \boldsymbol{d}^{(1,2)})$$

$$\boldsymbol{x}^{(1,1)} + \lambda \boldsymbol{d}^{(1,2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1+\lambda \end{bmatrix}$$

令 $\varphi(\lambda) = f\left(\boldsymbol{x}^{(1,1)} + \lambda \boldsymbol{d}^{(1,2)}\right) = (1+\lambda)^2 + 1$

$$\frac{d\varphi}{d\lambda} = 2(1+\lambda)=0$$

得到$\lambda_2 = -1$，$x^{(1,2)} = (0,0)^T$.

Powell方法

■ 算法求解问题示例

令 方向$\boldsymbol{d}^{(1,3)} = x^{(1,2)} - x^{(1,0)} = (-2,-1)^T$

再从$\boldsymbol{x}^{(1,2)}$出发，沿$\boldsymbol{d}^{(1,3)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(1,2)} + \lambda \boldsymbol{d}^{(1,3)}) \Rightarrow \lambda_3 = -\frac{2}{13}$$

得到第一轮迭代的最好点：$\boldsymbol{x}^{(1)} = \left(\frac{4}{13}, \frac{2}{13}\right)^T$

第二轮迭代

第二轮的搜索方向为

$$\boldsymbol{d}^{(2,1)} = \boldsymbol{d}^{(1,2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\boldsymbol{d}^{(2,2)} = \boldsymbol{d}^{(1,3)} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

替换掉第一个搜索方向

Powell方法

■ 算法求解问题示例

初始点$x^{(2,0)} = x^{(1)} = (4/13, 2/13)^T$

沿$\boldsymbol{d}^{(2,1)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(2,0)} + \lambda \boldsymbol{d}^{(2,1)}) \Longrightarrow \lambda_1 = -\frac{6}{13} \quad x^{(2,1)} = \left(\frac{4}{13}, -\frac{4}{13}\right)^T$$

沿$\boldsymbol{d}^{(2,2)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(2,1)} + \lambda \boldsymbol{d}^{(2,2)}) \Longrightarrow \lambda_2 = -\frac{18}{169} \quad x^{(2,2)} = \left(\frac{36}{169}, -\frac{60}{169}\right)^T$$

令方向$\boldsymbol{d}^{(2,3)} = \boldsymbol{x}^{(2,2)} - x^{(2,0)} = (36/169, -60/169)^T$

最后，从$\boldsymbol{x}^{(2,2)}$出发，沿$\boldsymbol{d}^{(2,3)}$作一维搜索：

$$\min_{\lambda} f(\boldsymbol{x}^{(2,2)} + \lambda \boldsymbol{d}^{(2,3)}) \Longrightarrow \lambda_3 = -\frac{9}{4}$$

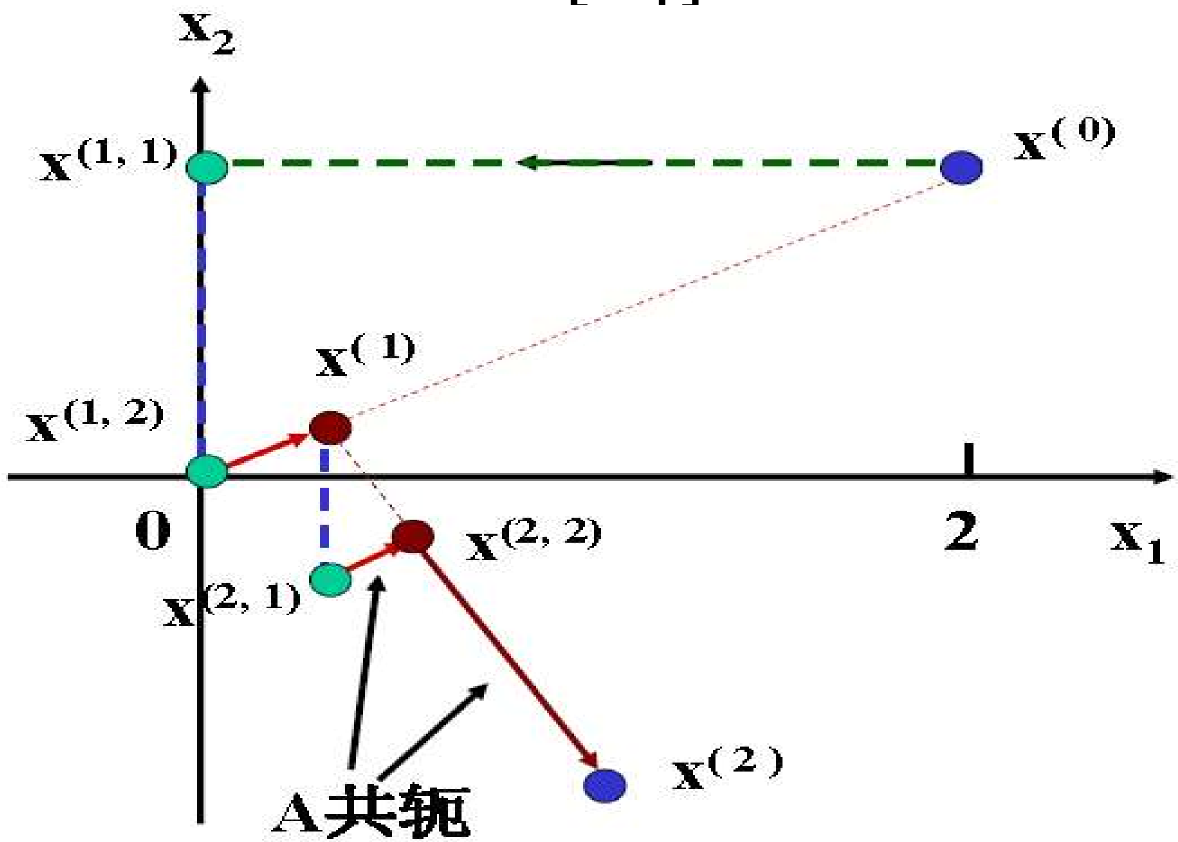
Powell方法

■ 算法求解问题示例

得到第二轮迭代的极小点

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(2,2)} + \lambda_3\ \boldsymbol{d}^{(2,3)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

搜索过程示意图



Powell方法

■ 二次终止性

- 当极小化正定二次函数时，如果每轮迭代中**n** 个搜索方向均线性无关，那么**Powell**方法至多经**n**轮迭代达到极小点；

- 在某轮迭代中，**n**个搜索方向线性相关，由此导致即使对正定二次函数经**n**轮迭代也达不到极小点，甚至任意迭代下去，永远达不到极小点；

- 在**Powell**方法中，保持**n**个搜索方向线性无关对于找到算法的最优点至关重要，然而**Powell**方法可能选取接近线性相关的方向，尤其是变量很多的情况。

Powell方法

■ 改进的Powell方法

◆ 改进的Powell方法与原来方法的主要区别在于**替换方向的规则不同**。

◆ **基本思路**：当**初始搜索方向线性无关**时，需要保证每轮迭代中以**搜索方向为列的行列式不为零**，这样搜索方向就是线性无关的。此外，随着迭代的延续，**搜索方向接近共轭**的程度逐渐增加。

◆ 改进的Powell方法**不具有二次终止性**，但是其计算效果仍然**令人满意**。

Powell方法 第k阶段相邻搜索点目标函数值差的最大值

■ 改进的Powell方法计算步骤如下

1．给定初始点$x^{(0)}$，线性无关的方向

$d^{(1,1)}, \ldots, d^{(1,n)}$

允许误差 $\varepsilon>0$。置k=1

2．$x^{(k,0)}=x^{(k-1)}$，从$x^{(k,0)}$出发，依次

沿方向 $d^{(1,1)}, \ldots, d^{(1,n)}$ 作一维搜索

得到点 $x^{(k,1)}, \ldots, x^{(k,n)}$

求指标m，使得

$$f(x^{(k,m-1)})-f(x^{(k,m)}) = \max_{j=1,\ldots,n}\{f(x^{(k,j-1)})-f(x^{(k,j)})\}$$

令 $d^{(k,n+1)}=x^{(k,n)}-x^{(k,0)}$

如$\|x^{(k,n)}-x^{(k,0)}\|\le\varepsilon$，则停止计算；否则转3

Powell方法 第k阶段的基点与上一阶段基点目标函数值差

■ 改进的Powell方法计算步骤如下

3. 求 $\lambda_{n+1} \rightarrow \min\limits_{\lambda} f(x^{(k,0)}+\lambda d^{(k,n+1)})$

令 $x^{(k+1,0)}=x^{(k)}=x^{(k,0)}+\lambda_{n+1} d^{(k,n+1)}$

如 $\| x^{(k)}-x^{(k-1)} \| \leqslant \varepsilon$，则停止计算，得点 $x^{(k)}$；否则转4

4. 如：$$|\lambda_{n+1}| > \left[\frac{f(x^{(k,0)})-f(x^{(k+1,0)})}{f(x^{(k,m-1)})-f(x^{(k,m)})}\right]^{\frac{1}{2}}$$

令 $d^{(k+1,j)}=d^{(k,j)}$, $j=1,\ldots,m-1$

$d^{(k+1,j)}=d^{(k,j+1)}$, $j=m,\ldots,n$ 替换掉第m个

置k:=k+1，转2；否则，令

$d^{(k+1,j)}=d^{(k,j)}$, $j=1,\ldots,n$ 不替换

置k:=k+1，转2

Powell方法

■ 讨论

- Powell极小化正定二次函数时，如果每轮搜索方向都线性无关，那么至多n轮迭代，达到极小点；
- Powell方法在应用于变量很多的极小化问题时，所选择的搜索方向可能接近线性相关，不利于收敛；
- 改进的Powell方法不具有二次终止性，但计算效果仍令人满意。