

1. 用基于精确线搜索的最速下降法求解下述无约束优化问题

$$\min f(x_1, x_2) = x_1^2 + x_2^2 - x_1 + 2x_2$$

其中初始点为  $y_0 = (0, 0)^t$ , 精度为  $\varepsilon = 0.001$ .

**Answer.**  $\nabla f(x) = (2x_1 - 1, 2x_2 + 2)^t$  第一次迭代

- $d^{(0)} = -\nabla f(x^{(0)}) = (1, -2)^t$
- $\lambda = \arg \min_{\lambda \geq 0} \varphi(\lambda) = \arg \min_{\lambda \geq 0} f(x^{(0)} + \lambda d^{(0)}) = \frac{1}{2}$
- $x^{(1)} = (\frac{1}{2}, -1)^t$

$d^{(1)} = (0, 0)^t, \|d\| = 0 \leq \varepsilon$ , 所以  $x^* = (\frac{1}{2}, -1)^t, f^* = 0$ .

2. 考虑二次函数  $f(x) = \frac{1}{2}x^t A x + b^t x + c$ , 其中  $A \in \mathbb{R}^{n \times n}$  为对称正定矩阵, 设函数的极小值点为  $x^*$ , 并设  $x_0 = x^* + ts$ , 其中  $s$  为函数  $f$  的 Hesse 矩阵的某特征值的特征向量. 若以  $x_0$  为初始点, 试证明基于精确线搜索的最速下降法与牛顿法等价.

*Proof.*  $\nabla f(x) = Ax + b, \nabla^2 f(x) = A$

使用牛顿法:  $x^{(1)} = x^{(0)} - (\nabla^2 f(x))^{-1} \nabla f(x) = x^{(0)} - A^{-1}(Ax^{(0)} + b) = -A^{-1}b$

使用最速下降法:  $d^{(0)} = -\nabla f(x^{(0)}) = -Ax^{(0)} - b, f(x^{(1)}) = \frac{1}{2}x^{(1)t} A x^{(1)} + b^t x^{(1)} + c$ , 令  $f(x^{(1)})$  最小可得  $x^{(1)} = -A^{-1}b$

故最速下降法与牛顿法等价. □

3. 用共轭梯度法求解下列问题

$$\min \frac{1}{2}x_1^2 + x_2^2$$

取初始点  $x^{(1)} = (4, 4)^t$

**Answer.**  $f(x) = \frac{1}{2}x_1^2 + x_2^2, \nabla f(x) = (x_1, 2x_2)^t$ .

第一次迭代:  $g_1 = (4, 8)^t, d^{(1)} = -g_1 = (-4, -8)^t, x^{(2)} = x^{(1)} + \lambda_1 d^{(1)} = (\frac{16}{9}, -\frac{4}{9})^t$

第二次迭代:  $g_2 = (\frac{16}{9}, -\frac{8}{9})^t, d^{(2)} = -g_2 + \beta_1 d^{(1)} = (-\frac{160}{81}, \frac{40}{81})^t, x^{(3)} = x^{(2)} + \lambda_2 d^{(2)} = (0, 0)^t$

第三次迭代:  $g_3 = (0, 0)^t, \|g_3\| = 0$ , 故  $\bar{x} = (0, 0)^t, f^* = 0$ .

4. 总结最速下降法、牛顿法及共轭梯度法的基本思想和计算步骤.

**Answer.** 带精确线搜索的最速下降法

1. 给定初始点  $x^{(1)} \in E^n$ , 允许误差  $\varepsilon > 0$ , 置  $k = 1$ .
2. 取搜索方向:  $d^{(k)} = -\nabla f(x^{(k)})$

3. 若  $\|d^{(k)}\| \leq \varepsilon$ , 则停止计算; 否则, 从  $x^{(k)}$  出发, 沿  $d^{(k)}$  进行一维搜索, 求  $\lambda_k$ , 使

$$f(x^{(k)} + \lambda_k d^{(k)}) = \min_{\lambda \geq 0} f(x^{(k)} + \lambda d^{(k)})$$

4. 令  $x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$ , 置  $k := k + 1$ , 返回 2

牛顿法计算步骤:

1. 给定初始点  $x^{(0)} \in E^n$ , 允许误差  $\varepsilon > 0$ , 置  $k = 0$

2. 若  $\|\nabla f(x^{(k)})\| < \varepsilon$ , 则停止计算;

3.  $x^{(k+1)} = x^{(k)} - (\nabla^2 f(x^{(k)}))^{-1} \nabla f(x^{(k)})$ , 置  $k := k + 1$ , 返回 2.

FR 共轭梯度法 (二次凸函数)

1. 给定初始点  $x^{(1)}$ , 置  $k = 1$

2. 计算  $g_k = \nabla f(x^{(k)})$ , 若  $\|g_k\| = 0$ , 则停止计算, 否则进行下一步

3. 令  $d^{(k)} = -g_k + \beta_{k-1} d^{(k-1)}$ , 其中, 当  $k = 1$  时,  $\beta_0 = 0$ , 当  $k > 1$  时,  $\beta_{k-1} = \frac{d^{(k-1)t} A g_k}{d^{(k-1)t} A d^{(k-1)}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$

4. 令  $x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$ , 其中  $\lambda_k = \frac{g_k^t g_k}{d^{(k)t} A d^{(k)}}$

5. 若  $k = n$ , 则停止计算, 否则置  $k = k + 1$ , 返回步骤 2.

5.  $\min x_1^2 + 2x_2^2 - 2x_1x_2 + 2x_2 + 2$ , 取初始点  $x^{(1)} = (0, 0)^t$ .

**Answer.**  $\nabla f(x) = (2x_1 - 2x_2, 4x_2 - 2x_1 + 2)^t$

共轭梯度法:

第一次迭代:  $g_1 = (0, 2)^t, d^{(1)} = -g_1 = (0, -2)^t, x^{(2)} = x^{(1)} + \lambda_1 d^{(1)} = (0, -\frac{1}{2})^t$

第二次迭代:  $g_2 = (1, 0)^t, d^{(2)} = -g_2 + \beta_1 d^{(1)} = (-1, -\frac{1}{2})^t, x^{(3)} = x^{(2)} + \lambda_2 d^{(2)} = (-1, -1)^t$

第三次迭代:  $g_3 = (0, 0)^t, \|g_3\| = 0$ , 故  $\bar{x} = (-1, -1)^t, f^* = 1$

牛顿法:

$x^{(2)} = x^{(1)} - G_1^{-1} g_1 = (-1, -1)^t, \|g_2\| = 0$ , 故  $\bar{x} = (-1, -1)^t$ .

6.  $\min 2x_1^2 + 2x_1x_2 + 5x_2^2$ , 取初始点  $x^{(1)} = (2, -2)^t$ .

**Answer.**  $\nabla f(x) = (4x_1 + 2x_2, 2x_1 + 10x_2)^t$

共轭梯度法:

第一次迭代:  $g_1 = (4, -16)^t, d^{(1)} = -g_1 = (-4, 16)^t, x^{(2)} = x^{(1)} + \lambda d^{(1)} = (\frac{57}{37}, -\frac{6}{37})$

第二次迭代:  $g_2 = (\frac{216}{37}, \frac{54}{37})^t, d^{(2)} = -g_2 + \beta d^{(1)}, x^{(3)} = x^{(2)} + \lambda_2 d^{(2)} = (0, 0)^t$

第三次迭代:  $g_3 = (0, 0), \|g_3\| = 0$ , 故  $\bar{x} = (0, 0)^t$ .

牛顿法:

$x^2 = x^1 - G_1^{-1} g_1 = (0, 0), \|g_2\| = 0$ , 故  $\bar{x} = (0, 0)^t$ .

7.  $\min x_1^2 + 2x_2^2 + 2$ , 取  $x_0 = (1, 1)^t$ , 用共轭梯度法.

**Answer.**  $\nabla f(x) = (2x_1, 4x_2)^t$

第一次迭代:  $g_1 = (2, 4)^t, d^{(1)} = (-2, -4)^t, x^{(2)} = (\frac{4}{9}, -\frac{1}{9})^t$

第二次迭代:  $g_2 = (\frac{8}{9}, -\frac{4}{9})^t, d^{(2)} = -g_2 + \beta_1 d^{(1)} = \frac{20}{81}(-4, 1)^t, x^{(3)} = x^{(2)} + \lambda_2 d^{(2)} = (0, 0)^t$

第三次迭代:  $g_3 = (0, 0), \|g_3\| = 0$ , 故最优解  $\bar{x} = (0, 0)$ .

8. 证明: 如果非零向量  $p_0, \dots, p_l$  满足  $p_i^t A p_j = 0, \forall i \neq j$ , 其中  $A$  为对称正定矩阵, 则这些向量线性无关.

*Proof.* 设  $\alpha_0 p_0 + \dots + \alpha_l p_l = 0$

左乘  $p_i^t A$  可得  $\alpha_0 p_i^t A p_0 + \dots + \alpha_l p_i^t A p_l = 0$ , 得到  $\alpha_i = 0$ . 故  $p_0, \dots, p_l$  线性无关.  $\square$