

1. Consider the minimization problem

$$\min \frac{1}{2}x^tAx + b^tx + c$$

where

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \quad c = 2$$

Let $x^{(0)} = (1, 1, 1, 1)^t$, $d^{(0)} = -\nabla f(x^{(0)})$, $\sigma = 1.0e - 4$, $\beta = 1$, $\gamma = 0.5$. Use Armijo, Goldstein, and Wolfe line search to find the next iterate point $x^{(1)}$.

Answer. $\nabla f(x) = Ax + b$, we get $d^{(0)} = -\nabla f(x^{(0)}) = (-5, -7, -3, -3)^t$.

- Armijo:

- the minimum non-negative integer $m = 2$, and $x^{(1)} = (-0.25, -0.75, 0.25, 0.25)^t$.

$$f(x^{(0)} + \beta\gamma^m d^{(0)}) \leq f(x^{(0)}) + \sigma\beta\gamma^m \nabla f(x^{(0)})^T d^{(0)}$$

- Goldstein:

- λ satisfy the flowing, then we get $\lambda = 0.2441$.

$$\begin{aligned} f(x^{(k)} + \lambda d^{(k)}) &\leq f(x^{(k)}) + \sigma\lambda \nabla f(x^{(k)})^T d^{(k)} \\ f(x^{(k)} + \lambda d^{(k)}) &> f(x^{(k)}) + (1 - \sigma)\lambda \nabla f(x^{(k)})^T d^{(k)} \end{aligned}$$

so the $x^{(1)} = (-0.2207, -0.7090, 0.2676, 0.2676)^t$

- Wolfe:

- λ satisfy the flowing, $0 < \sigma_1 < \sigma_2 < 1$, let $\sigma_1 = 0.2$, $\sigma_2 = 0.5$, then we get $\lambda = 0.2441$.

$$\begin{aligned} f(x^{(k)} + \lambda d^{(k)}) &\leq f(x^{(k)}) + \sigma_1\lambda \nabla f(x^{(k)})^T d^{(k)} \\ \nabla f(x^{(k)} + \lambda d^{(k)})^T d^{(k)} &\geq \sigma_2 \nabla f(x^{(k)})^T d^{(k)} \end{aligned}$$

so the $x^{(1)} = (-0.2207, -0.7090, 0.2676, 0.2676)^t$

2. 对严格凸二次函数 $f(x) = \frac{1}{2}x^tAx + b^tx$, 试求在 x_k 点沿下降方向 d_k 的最优步长。若取 $d_k = -g_k$, 试计算目标函数在每一迭代步的下降量。

解.

$$\begin{aligned}
 f(x + \lambda_k d_k) &= \frac{1}{2}(x + \lambda_k d_k)^t A(x + \lambda_k d_k) + b^t(x + \lambda_k d_k) \\
 &= \frac{1}{2}(x^t A x + x^t A d_k \lambda_k + d_k^t A x \lambda_k + d_k^t A d_k \lambda_k^2) + b^t x + b^t d_k \lambda_k \\
 &= \frac{1}{2} d_k^t A d_k \lambda_k^2 + \left(\frac{1}{2} x^t A d_k + \frac{1}{2} x^t A^t d_k + b^t d_k \right) \lambda_k + b^t x
 \end{aligned}$$

$$\min_{\lambda} f(x + \lambda_k d_k) \text{ 对应的最优步长 } \lambda^* = -\frac{x^t A d_k + x^t A^t d_k + 2b^t d_k}{2d_k^t A d_k}.$$

3. 给出算法线性收敛、超线性收敛和二次收敛的定义。

解. 设序列 $\{\gamma^{(k)}\}$ 收敛于 γ^* , 定义满足

$$\lim_{k \rightarrow +\infty} \frac{\|\gamma^{(k+1)} - \gamma^*\|}{\|\gamma^{(k)} - \gamma^*\|^p} = \beta < \infty$$

的非负数 p 的上确界为序列 $\{\gamma^{(k)}\}$ 的收敛级。

- 线性收敛: 收敛级 $p = 1$ 且 $0 < \beta < 1$
- 超线性收敛: 收敛级 $p > 1$ 或者 $p = 1, \beta = 0$ 。
- 二次收敛: 收敛级 $p = 2$