

最优化作业 2

第 6 周周 5(3 月 31 日)交到 570533048@qq.com (注意: 实现 8 个基本的测试函数, 除下面的函数外, 尝试各个方法在 8 个测试函数上的性能表现, 只需要在代码中体现即可)

1. 写代码, 用共轭梯度法求解: $\min f(x) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ 假设初始点 $x^{(0)} = (0,0)^T, \epsilon = 10^{-6}$ 。
2. 写代码, 用黄金分割法, 斐波那契亚数列法, 二分法, **Dichotomous**, **Shubert-Piyavskii** 分别求解问题:

(a) $\min f(x) = 2x^2 - x - 1$, 初始区间 $[a_0, b_0] = [-1, 1]$, 区间精度 $\delta = 0.06$;

(b) $\min f(x) = 3x^2 - 21.6x - 1$, 初始区间 $[a_0, b_0] = [0, 25]$, 区间精度 $\delta = 0.08$;

3. 写代码, 用不精确一维搜索中的 Goldstein 方法, Goldstein-Price 法, Wolfe-Powell 方法来计算:

$$\min f(x + \lambda d)$$

其中 $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$, 取 $x^{(0)} = (-1, 1)^T, d = (1, 1)^T$, 并进行比较 (可同时采用 2 中的精确搜索方法进行比较)。

4. 请计算下面凸函数在指定点的次微分:

(a) $f(x_1, x_2, x_3) = \max\{|x_1|, |x_2|, |x_3|\}$, 在点 $(x_1, x_2, x_3) = (0, 0, 0)$

(b) $f(x) = e^{|x|}$, 在点 $x = 0$ 处 (这是标量)

(c) $f(x_1, x_2) = \max\{x_1 + x_2 - 1, x_1 - x_2 + 1\}$ 在点 $(x_1, x_2) = (1, 1)$

5. 用 DFP 方法求解:

$$\min 10x_1^2 + x_2^2$$

取初始 H 矩阵为单位阵, 即 $H^{(0)} = I, x^{(0)} = \left(\frac{1}{10}, 1\right)^T$, 做精确一维搜索。

6. 用 BFGS 法求解问题: $\min x_1^2 + 4x_2^2 - 4x_1 - 8x_2$

取初始 H 矩阵为单位阵, 即 $H^{(0)} = I, x^{(0)} = (0, 0)^T$, 做精确一维搜索。

7. 请用精确一维搜索来求解:

$$\min x_1^2 + x_2^2 - x_1x_2 - 10x_1 - 4x_2 + 60$$

取 $x^{(0)} = (0, 0)^T$,

(a) DFP 方法, $H^{(0)} = I$;

(b) BFGS 方法, $H^{(0)} = I$;

(c) FR 共轭梯度法, $\epsilon = 10^{-6}$, 初始点可设为 $x^{(0)} = (0, 0)^T$, 或自己设置。

比较各算法的迭代点, 各有什么特点?

8. 总结无约束优化求解的基本思想, 思考如何将非凸优化问题转化为凸优化问题, 如何将有约束问题转化为无约束问题。
9. 请将最速下降法、牛顿法、修正牛顿法的计算公式进行统一描述, 并体会变尺度法的基本思想。
10. 尝试采用深度学习中的随机梯度下降算法中各种一阶改进的优化算法来求解上述习题中的典型问题。
11. 自己查看已有的深度学习模型, 训练中其参数的变化是否服从低维特性? 尝试画出这些变化的分布。
12. 查阅一阶优化算法的最新加速算法思想, 并总结其基本思路。(可查阅北大林宙辰老师的英文版书籍)。
13. 自己查阅资料, 思考共轭函数和对偶性的联系。