**目 录**

[数学建模与最优化作业题 2](#_Toc150956563)

[最速下降法 2](#_Toc150956564)

[介绍 2](#_Toc150956565)

[求解步骤 2](#_Toc150956566)

[算例求解 3](#_Toc150956567)

[牛顿法 5](#_Toc150956568)

[介绍 5](#_Toc150956569)

[求解步骤 5](#_Toc150956570)

[算例求解 6](#_Toc150956571)

[DFP与BFGS 7](#_Toc150956572)

[拟Newton算法 7](#_Toc150956573)

[DFP 7](#_Toc150956574)

[BFGS 8](#_Toc150956575)

[算例求解 10](#_Toc150956576)

[方法比较 12](#_Toc150956577)

数学建模与最优化作业题

2102-U202011885-翁强

选题：E、按照课堂讲授的内容，对第 9 章中所介绍的算法（最速下降法、牛顿法、 DFP 算法、BFGS 算法、FR 算法中选取其中 3 个）编程或者调用库函数实现， 并给出具体算例。（最高可得 100 分）

本报告中算例均有采用

以便于不同方法间的比较，并对个别方法列举有额外算例，在此特别说明。

最速下降法

介绍

最速下降法是无约束最优化一种基本的迭代算法，以最速下降方向作为搜索方向。由A-L柯西(Cauchy,法国,1789～1857)于1847年提出，又称(负)梯度法或柯西方向法。

求解无约束问题:

的线搜索型算法的迭代格式如下:

式中为搜索方向；为步长。

最速下降法是无约束最优化一种基本的迭代算法，以最速下降方向作为搜索方向。最速下降方向是下面的问题的解:

不难发现，上面的问题的解为。因此，最速下降方向就是负梯度方向。

最速下降法步骤简单，总能产生下降方向，但算法收敛速度慢。最速下降法在每次迭代中只考虑了当前点的梯度信息，因此可能会出现在局部极小点附近振荡的情况。

求解步骤

最速下降法的具体步骤如下:

* 初始化起始点和迭代次数。
* 计算当前点的梯度，其中为目标函数。
* 沿着负梯度方向进行搜索，即计算搜索方向。
* 选择一个合适的步长，使得沿着搜索方向移动一定距离后能够使目标函数值下降。
* 更新当前点，并将迭代次数加1，即。
* 重复步骤5，直到满足终止条件（如目标函数值下降到一定程度或达到最大迭代次数)。

算例求解

代码解释

具体代码见附件steepest\_descent.py。

* f(x)函数与grad(x)函数分别表示目标函数与目标函数的梯度。
* amijo(x, pk, c1, beta)函数表示采用amijo线性搜索算法。
* steepest\_descent(x0, max\_iter, tol)函数是最速下降法的主体实现函数。

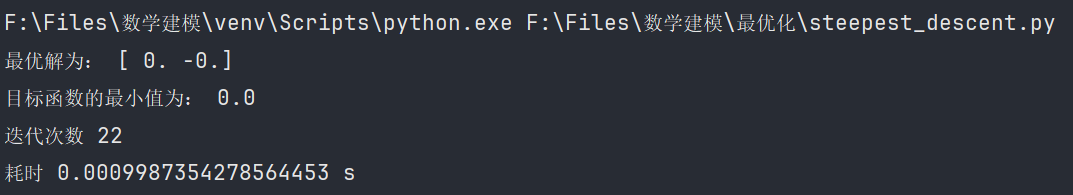


图1-1 最速下降法算例1运行结果

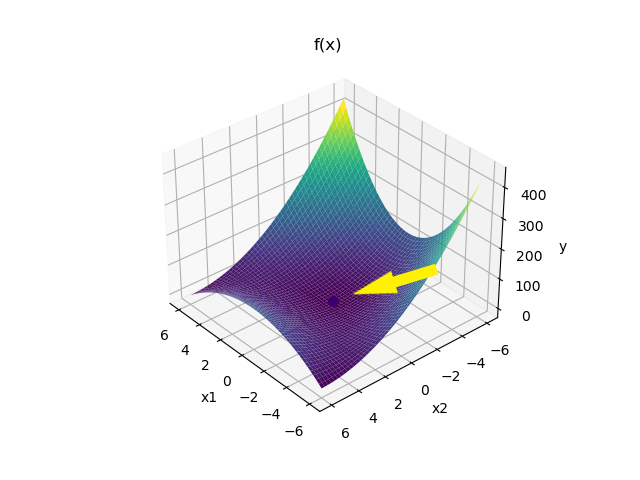


图1-2 最速下降法算例1运行结果可视化

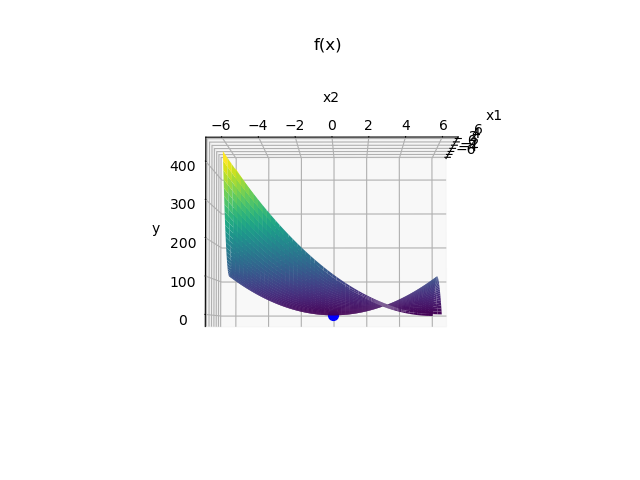


图1-3 最速下降法算例1运行结果可视化

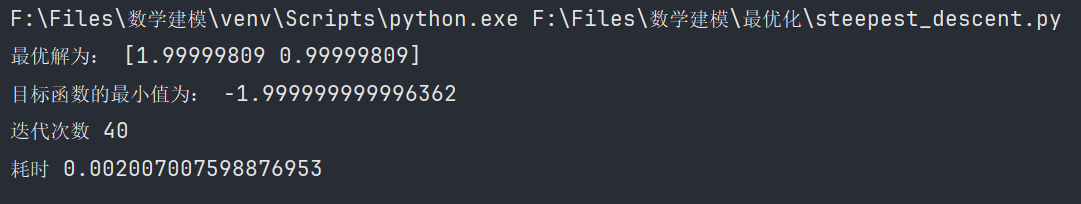


图1-4 最速下降法算例2运行结果

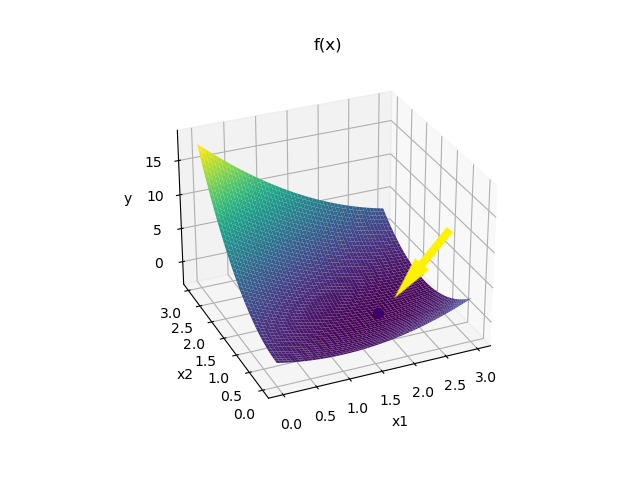


图1-5 最速下降法算例2运行结果可视化

牛顿法

介绍

考虑无约束最优化问题

其中为实值连续函数，且一般假定它具有二阶连续导数。Newton法的基本思想是利用目标函数在迭代点处的二次Taylor多项式作为二次函数，并用这个二次函数的极小点序列去逼近目标函数的极小点。

设目标函数在上具有连续的二阶导数，其Hesse矩阵 (简记为)正定。不妨设经过次迭代已获得迭代点，将在点处展成二阶泰勒公式，有

其中

显然 是二次函数. 由假设 是正定的, 则 有唯一的极小点, 将它取为的下一次近似 , 其中 的极小点.由极值点的必要条件知, 应满足 即.令, 其中 应满足,因此

称（9.3.2）式为Newton迷代公式.

对照基本迭代公式易知, 式 (9.3.2)中的搜索方向

步长因子. 换句话说, 从点 出发, 沿搜索方向, 并取步长 , 即可得 的极小点. 因此,为点 处近似二次函数 的极小点的方向. 此时称为从点 出发的Newton方向.

从初始点开始, 每一轮从当前迭代点出发,沿Newton方向, 并取步长 的算法称为 Newton 法.

求解步骤

已知目标函数 ，终止限 ，Newton法迭代步骤如下。

* 选定初始点 ，置 。
* 计算 。
* 若，则 , 结束；否则，计算 。
* 由方程 解出 。
* 令 ，置 ，转 (2)。

算例求解

代码解释

具体代码见附件Newton.py。

* get\_Hessian(data)计算Hessian矩阵。
* get\_Jacobian(data)计算Jacobian矩阵。
* plot\_contour()绘出等高线图。

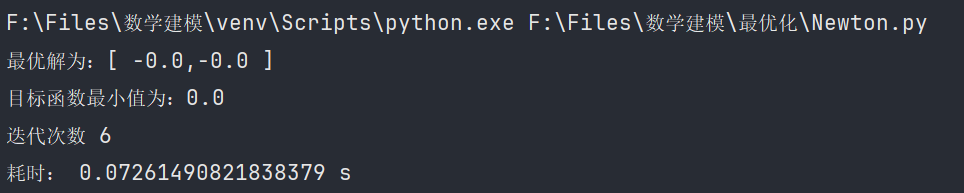


图2-1 牛顿法算例运行结果

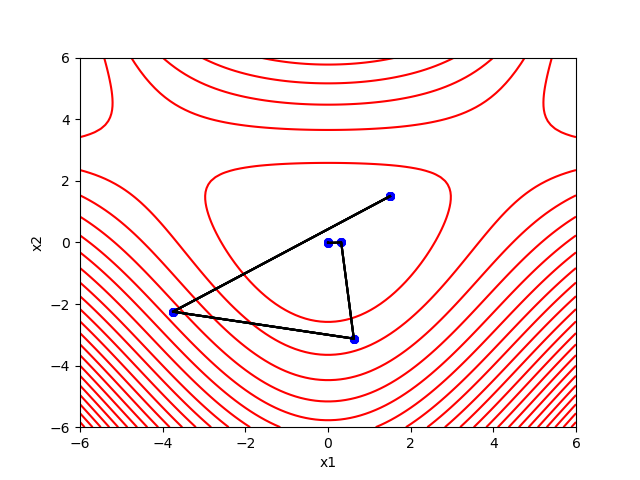
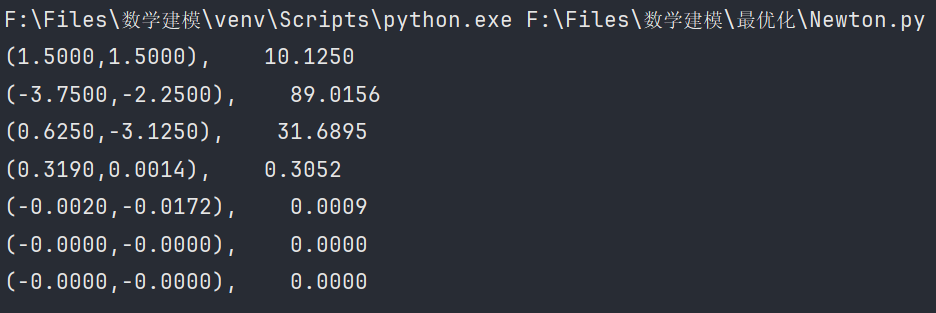


图2-2 牛顿法算例迭代过程 图2-3 牛顿法算例迭代过程可视化

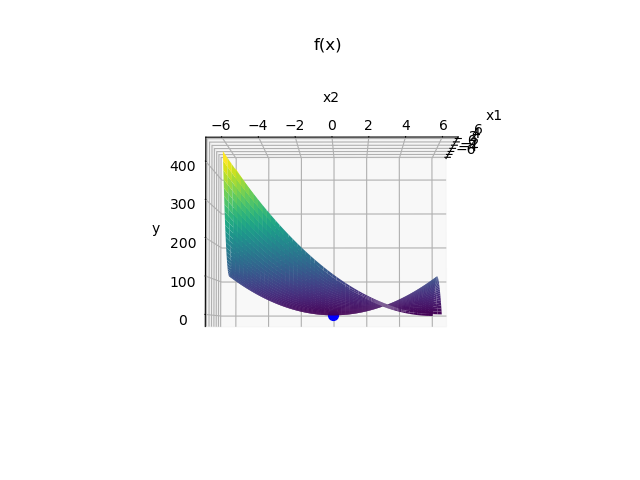
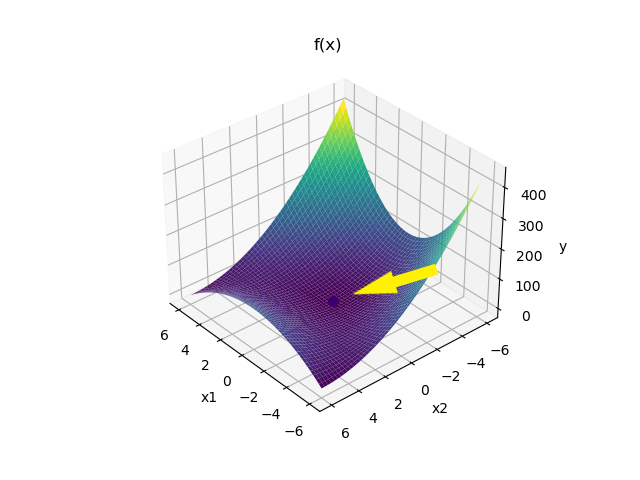


图2-4 牛顿法算例运行结果可视化 图2-5 牛顿法算例运行结果可视化

DFP与BFGS

拟Newton算法

在Newton法的迭代公式中，是初始点，和分别是目标函数在点的梯度和Hesse矩阵。为了消除这个迭代公式中的Hesse逆矩阵, 设想用矩阵来替换它，从而在第步，, 于是有

(9.3.3)

为了取得更大的灵活性，考虑更一般的迭代公式

其中步长因子通过从出发沿作一维搜索来确定，初始点和初始矩阵是预先给定的。

在迭代过程中利用已得的迭代点和目前函数的值，最多再利用一阶导数按某种规则求得。确定的一种自然想法是将作为的一种近似来构造，由于是对称矩阵，且

故

令，则,因此要求满足(1) 对称；(2) 条件 (2) 一般称为拟Newton条件。

进一步设想是由经过简单修正而得，即设

(9.3.5)

其中称为校正矩阵，应是对称矩阵，(9.3.5) 式称为校正公式。于是拟Newton条件 (2)变为(即

(9.3.6)

因满足 (9.3.6) 的对称矩阵有无穷多个，故拟Newton法是一族算法。

其中，DFP和BFPS是最常用的两个拟Newton算法。

DFP

介绍

Davidon-Fletcher-Powell方法或DFP 方法是一种拟牛顿方法，它使用一个公式(DFP 公式) 给出满足特定割线方程且最接近当前估计并满足曲率条件的解。以William C.Davidon, Roger Fletcher和Michael JD. Powell的名字命名。它是多维问题割线法的推广，是第一个拟牛顿法。如果使用这个公式更新Hessian矩阵，就可以保证对称性和正定性。

给定函数的泰勒展开式由其梯度定义 ，正定Hessian, 可以写成如下。

此外，梯度本身的泰勒展开式 (割线方程) 可以写成如下。

这用于更新。下面所示的 DFP 公式是对称且正定的。当前的近似值给出最接近的解决方案：

,

其中

是对称正定矩阵。对应的逆 Hessian 矩阵的近似值由以下等式给出。

假设是正定矩阵，所以和必须满足以下曲率条件。

求解步骤

已知目标函数及其梯度, 问题的维数, 终止限

* 选定初始点，置I
* 计算，若，则输出，结束； 否则转(3)
* 取，置，转(4)
* 一维搜索求，使得，令 =，转(5)
* 计算，若，则输出，结束； 否则转(6)
* 若，令，转(3)；否则，转(7)
* 计算

置*=*+ 1，转(4)

BFGS

介绍

在数值优化中， Broyden-Fletcher-Goldfarb-Shanno（BFGS）算法是一种求解无约束非线性优化问题的迭代算法。和相关的Davidon-Fletcher-Powell算法类似，BFGS算法通过利用曲率信息对梯度进行预处理来确定下降方向。曲率信息则是通过维护一个使用广义的割线法逐步近似的关于损失函数的Hessian矩阵来获得。

考虑如下形式的校正公式

(9.3.8)

式 (9.3.8) 中有一个参数, 它可以取任何实数，每取一个实数 就对应一种拟Newton算法. 容易验证，当取=0时就是DFP校正公式.

令

就转变为著名的BFGS校正公式

求解步骤

已知目标函数及其梯度, 问题的维数, 终止限

* 选定初始点，置。
* 计算，若，则输出，结束； 否则转(3)。
* 取，置，转(4)。
* 一维搜索求，使得，令 =，转 (5) 。
* 计算，若，则输出，结束； 否则转(6)。
* 若，令，转(3)；否则，转(7)。
* 计算

置,转(4)。

算例求解

代码解释

具体代码见附件DFPandBFGS.py。

* f(x)函数表示目标函数。
* getLosses(retall, target\_point, func)函数表示优化时的损失函数。
* fmin\_powell与fmin\_bfgs函数分别表示DFP方法与DFGS方法。

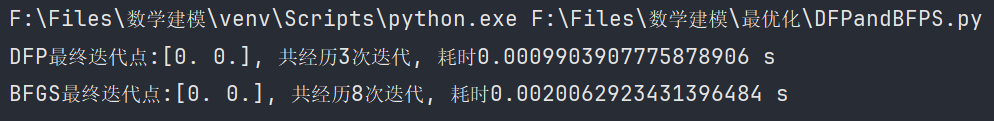


图3-1 DFP法与BFGS法算例1运行结果

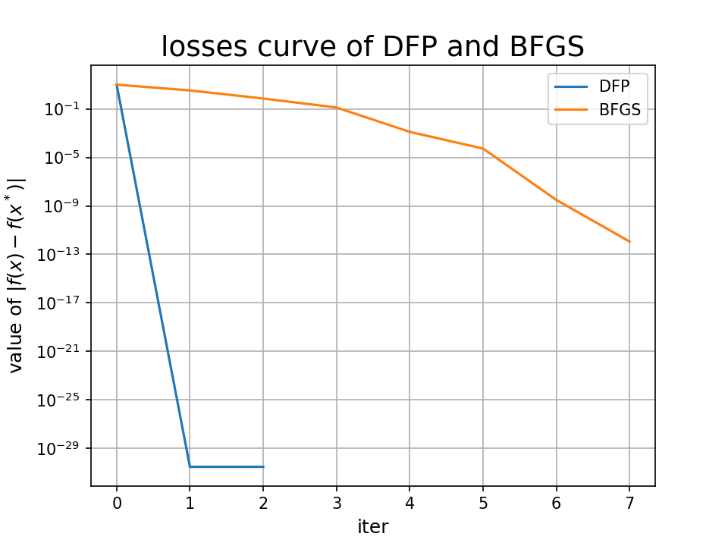


图3-2 DFP法与BFGS法算例1运行损失函数可视化结果

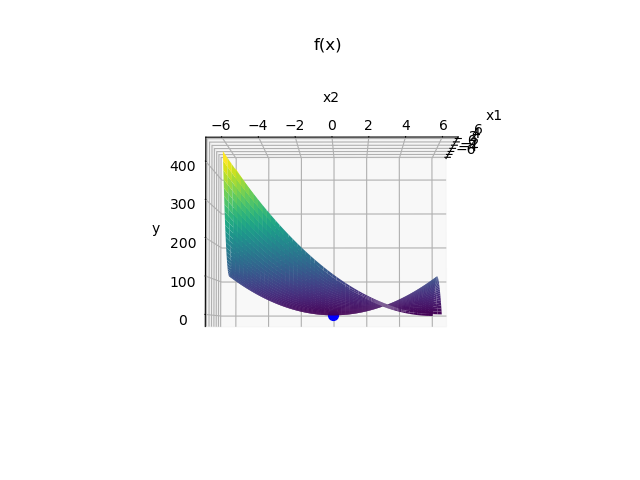
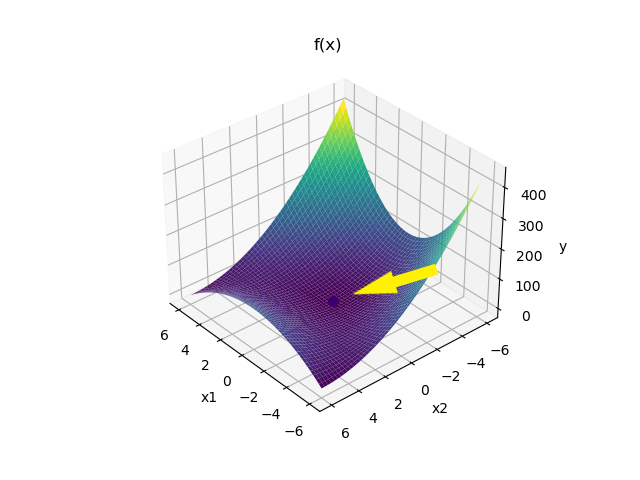


图3-3 DFP法与BFGS法算例1运行结果可视化图 3-4 牛顿DFP法与BFGS法算例1运行结果可视化

其中 , , 代表向量 的第 个维度上的元素。

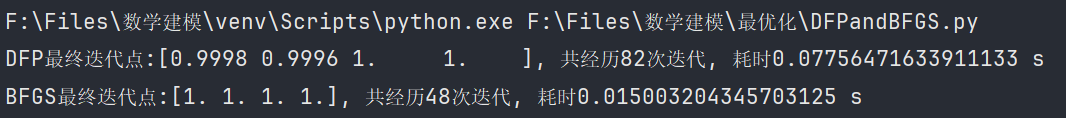


图3-5 DFP法与BFGS法算例2运行结果

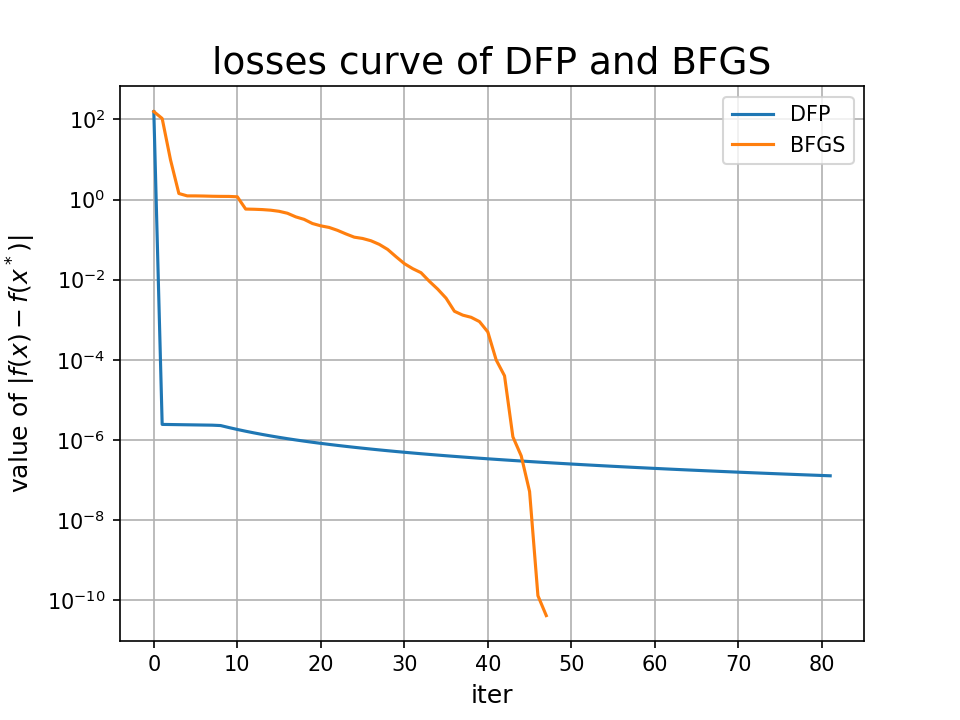


图3-6 DFP法与BFGS法算例2运行损失函数可视化

方法比较

针对同一算例：

各方法运行效率如下：

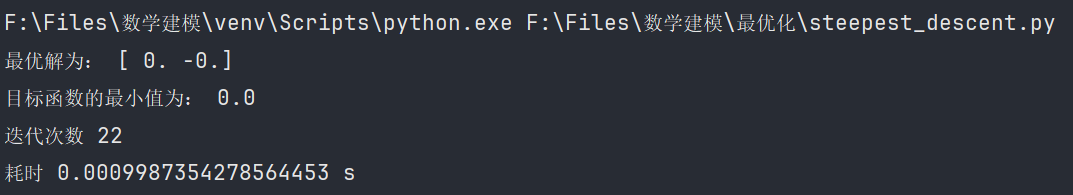
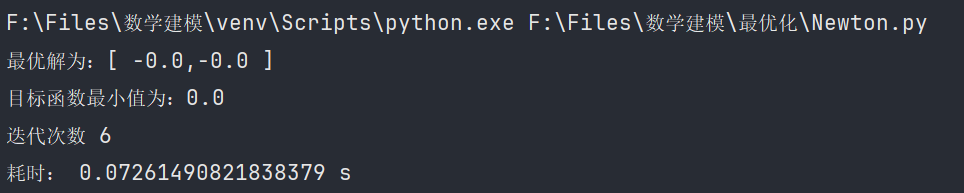
 

图4-1 最速下降法 图4-2 牛顿法

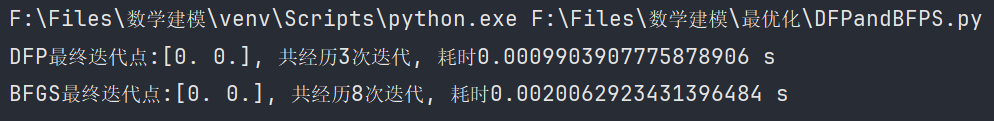


图4-3 DFP法与BFGS法

从各方法的运行时间可以看出，牛顿法运行速度最低，BFGS法运行速度略佳，而最速下降法与DFP法运行速度最佳。

从各方法的迭代次数可以看出，按照最速下降法、BFGS法、牛顿法、DFP法的次序，迭代次数以此下降。