二变量二次规划问题收敛速度的讨论与证明

一、问题背景

考虑以下两个变量的二次规划问题：

$$\min f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2)\ ,\quad \gamma > 0$$

易知，其最优解为$x^* = (0,0)$，其 Hessian 矩阵$\nabla^2 f(x)$是常值，其特征值为1和γ，则强凸性常数满足此关系 $\frac{M}{m} = \frac{\max\{1,\gamma\}}{\min\{1,\gamma\}} = \max\left\{\gamma, \frac{1}{\gamma}\right\}$ ，采用精确直线搜索的梯度下降方法，选取初始点$x^0 = (\gamma, 1)$，可以推导出下述闭式表达式：

$$x^k = \left(\frac{\gamma - 1}{\gamma + 1}\right)^k \begin{bmatrix} \gamma \\ (-1)^k \end{bmatrix}$$

由此可知：

① $\gamma = 1$时，一次迭代可得到最优解；

② 对于γ离1不远的情况，如$\frac{1}{3}$和3之间，收敛速度很快

③ 如果$\gamma \gg 1$或$\gamma \ll 1$时，收敛速度将会非常慢

本报告将使用编程讨论以上收敛速度快慢问题。

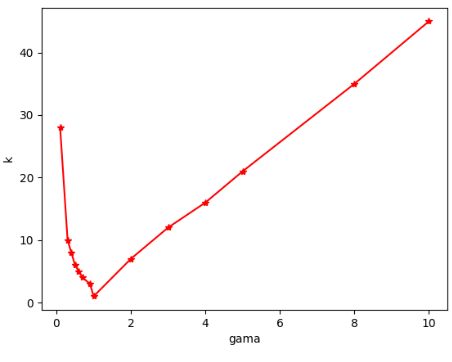
二、问题讨论

（1）基本设定

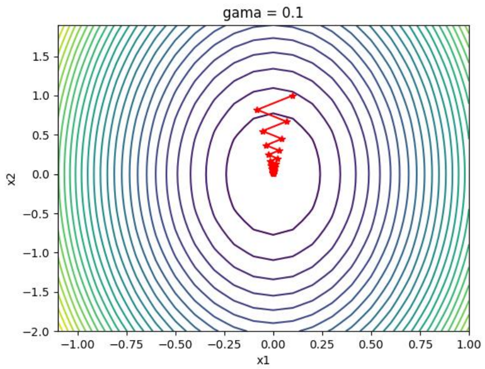
本实验使用问题背景中的x^k的递推式，由于问题的最优值$p^* = 0$，本实验使用的停止迭代条件为$\left|f(x^k) - p^*\right| < \epsilon$,其中设定$\epsilon = 1e-6$，统计对于不同的$\gamma$（范围从 0.1 到 10）收敛时的迭代次数$k$。

（2）实验结果

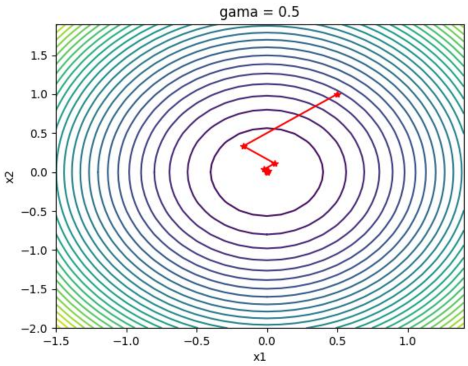
实验发现在$\gamma = 1$的时候只需要一次迭代即可收敛，在γ距离 1 不远的时候，收敛速度很快，大约 10 次迭代即可收敛，$\gamma \gg 1$或$\gamma \ll 1$时，收敛速度非常慢，需要 30 次左右的迭代才会收敛，这是由于当$\gamma \gg 1$或$\gamma \ll 1$时，$\frac{M}{m} = \max\left\{\gamma, \frac{1}{\gamma}\right\}$，会变得非常大，使得收敛变慢，而$\gamma$在 1 附近时，$\frac{M}{m}$较小，使得收敛较快。具体实验图表如下所示。



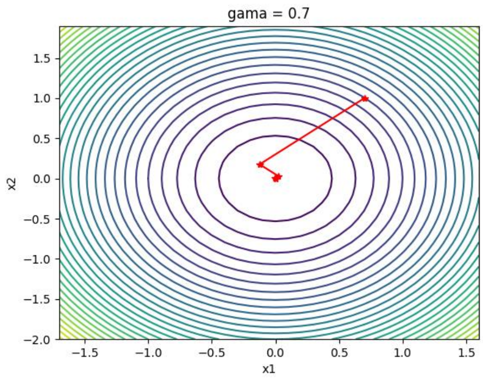
图一：收敛次数k随γ的变化曲线



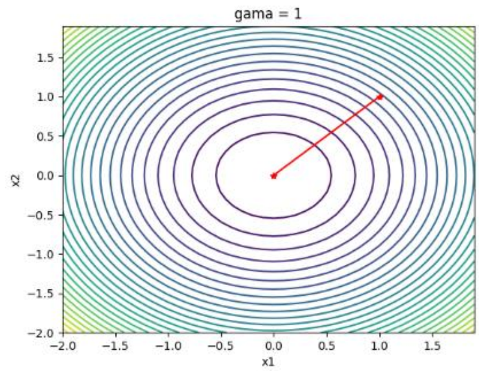
(a) $\gamma = 0.1$时的收敛过程，需要多次迭代



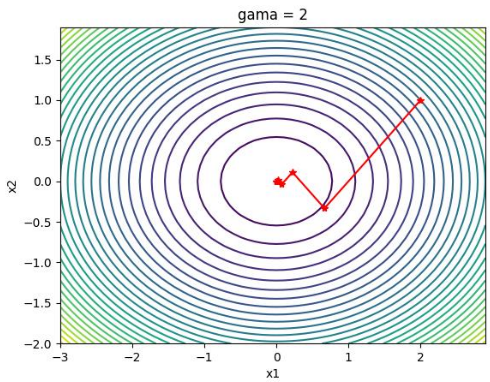
(b) $\gamma = 0.5$时的收敛过程，迭代次数较少



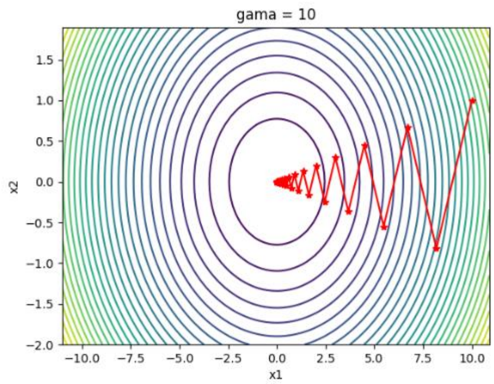
(c) $\gamma = 0.7$时的收敛过程，迭代次数较少



(d) $\gamma = 1$的收敛过程，仅需要一次迭代



(e) $\gamma = 2$的收敛过程，迭代次数较少



(f) $\gamma = 10$收敛过程，需要较多的迭代次数

图二：γ不同取值下的收敛过程