tutorial 7参考答案

gbl555

2019/4/16

1、For two-dimensional function $f(x, y) = 25x^2 + y^2$

a)Compute the gradient vector at initial point $(0.6, 4)$;

解：梯度表达式g=(50x,2y),所以初始点梯度(30,8)

b)Given the normalized gradient vector c at point $(0.6, 4)$;

解：标准化梯度向量c=g/—g—=(0.9662,0.2577)

c)If we decrease the initial point in the direction c by a step size of 0.5, what is value of the function at the new point?

解：(0.6,4)-0.5c=(0.1169,3.87115)

d)If we decrease the initial point in the direction $(1, 0)^T$ by a step size of 0.5, what is value of the function at the new point?

解：(0.6,4)-0.5(1,0)=(0.1,4)

2、As illustrated in following figure, is there any ideal to improve the optimization process for minimizing the function? [Hint: Robust gradient methods should re-scale the step size empirically depending on local properties of the function.]

解：No!梯度的方向是可信的，但梯度的大小什么都说明不了，和步长没有一毛钱的关系，鲁棒的梯度方法应该根据局部的性质来选择步长。根据经验，步长的选择应遵循以下两个原则：

1)如果走了这一步之后，函数值变大了，说明步子迈的太大，此时应该撤回这一步，并且减少步长。

2）如果走了一步之后，函数值变小了，则该步是成功的，说明步长再大一点也可能是ok的，此时应该增大步长。

3、1)解：

$$\frac{\partial}{\partial p_{uj}} \mathcal{J}(R; P, Q, b, d) = -\sum_{i:(u,i)\in\mathcal{K}} e_{ui} q_{ji} + \lambda p_{uj}$$

$$\frac{\partial}{\partial q_{ji}}\mathcal{J}(R; P, Q, b, d) = -\sum_{u:(u,i)\in\mathcal{K}} e_{ui}p_{uj} + \lambda q_{ji}$$

2)解：

$$\frac{\partial}{\partial b_u}\mathcal{J}(R; P, Q, b, d) = -\sum_{i:(u,i)\in\mathcal{K}} e_{ui} + \lambda b_u$$

$$\frac{\partial}{\partial d_i}\mathcal{J}(R; P, Q, b, d) = -\sum_{u:(u,i)\in\mathcal{K}} e_{ui} + \lambda d_i$$

3)解：更新规则如下

$$p_{uj} \leftarrow p_{uj} - t(-\sum_{i:(u,i)\in\mathcal{K}} e_{ui}q_{ji} + \lambda p_{uj})$$

$$q_{ji} \leftarrow q_{ji} - t(-\sum_{u:(u,i)\in\mathcal{K}} e_{ui}p_{uj} + \lambda q_{ji})$$

$$b_u \leftarrow b_u - t(-\sum_{i:(u,i)\in\mathcal{K}} e_{ui} + \lambda b_u)$$

$$d_i \leftarrow d_i - t(-\sum_{u:(u,i)\in\mathcal{K}} e_{ui} + \lambda d_i)$$

其中$e_{ui} = (r_{ui} - \mu - b_u - d_i - p_u^T q_i)$