Algorithm Foundations of Data Science and Engineering
Welcome Tutorial :-)
Tutorial 7

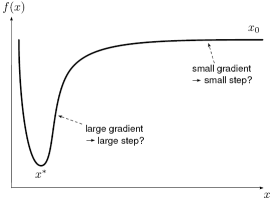
GAO Ming

DaSE @ ECNU

8 Apr., 2019

Tutorial 7

1. For two-dimensional function $f(x, y) = 25x^2 + y^2$
   a. Compute the gradient vector at initial point $(0.6, 4)$;
   b. Given the normalized gradient vector c at point $(0.6, 4)$;
   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?
   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?
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.]



Tutorial 7 Cont'd

3. Given the following cost function

$$\min_{P^*,Q^*,b_u^*,d_i^*} \mathscr{J}(R; P, Q, b_u, d_v)$$
$$= \frac{1}{2}[\sum_{(u,i)\in\mathscr{R}} (r_{ui} - \mu - b_u - d_i - \mathbf{p}_u^T \mathbf{q}_i)^2$$
$$+ \lambda(\|P\|_F^2 + \|Q\|_F^2 + \|\mathbf{b}\|_2^2 + \|\mathbf{d}\|_2^2)].$$

   a. Compute the gradient $\frac{\partial}{\partial p_{uj}} J$ and $\frac{\partial}{\partial q_{ji}} J$;
   b. Compute the gradient $\frac{\partial}{\partial b_u} J$ and $\frac{\partial}{\partial d_i} J$;
   c. By using gradient descent method, please give the update rules for estimating the parameters of the model;