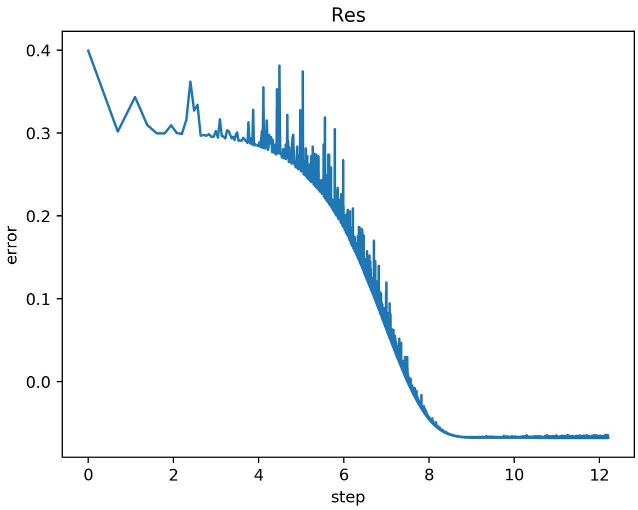
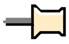
Homework 1
CS 259 @ SJTU
Prof. **David Bindel**
TA. **Zhou Fan**

Xun Gong(517020910141
ACM Class, Zhiyuan College, SJTU
Due Date: May 29, 2019
Submit Date: May 31, 2019

Problem 1: Stochastic gradient descent

(1)



Code is in attachment. 

(2)

Rule: Stochastic gradient descent is converge to 0.

Problem 2: Least 2p-norm regression

(1)

$$f(x) = \begin{pmatrix} \frac{1}{\sqrt{p}} r_1^p \\ \cdots \\ \frac{1}{\sqrt{p}} r_n^p \end{pmatrix}, \quad r_i = (b - Ax)_i = b_i - A_{i,:}x$$

$$J(x) = -\sqrt{p} \begin{pmatrix} r_1^{p-1} a_{1,1} & \cdots & r_1^{p-1} a_{1,n} \\ & \cdots & \\ r_n^{p-1} a_{n,1} & \cdots & r_b^{p-1} a_{n,n} \end{pmatrix}$$

(2)

Equal to show Gauss-Newton Step $\equiv min||D^k(Ap^k - r^k)||^2$

$$\mathcal{L} = ||D^k(Ap^k - b + Ax||^2$$
$$\delta\mathcal{L} = \delta(D^k(Ap^k - b + Ax)^T(D^k(Ap^k - b + Ax)$$
$$\delta\mathcal{L} = 2\delta x^T D^2(A^TAx - A^Tb + A^TAp)$$
$$\because \delta\mathcal{L} = 0$$
$$\therefore 0 = A^TAx - A^Tb + A^TAp$$
$$p = -(A^TA)^{-1}(A^TAx - A^Tb)$$

What is interesting, $p = -[f'^T f']^{-1}\delta\phi$

$\therefore$ These 2 question has same solution.

□