Modern Optimization Techniques - Exercise Sheet 6

Lydia Voß

voss@ismll.de

December 1, 2015

Solutions need to be handed in until **Tuesday, December 8th, 2015 at 10:15**

Exercise 1: Exact Newton Method (10P)

Let us (theoretically) optimize the following function using a Newton Descent approach:

$$f(x, y) = \exp(x^2 + y^2)$$

a) Compute the gradient $\nabla f(x, y)$ and the Hessian $\nabla^2 f(x, y)$!

b) Compute $\nabla^2 f(x, y)^{-1}$, using Cramer's Rule:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Longrightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

c) Compute the update step of the Newton Algorithm:

$$\Delta_{x,y} = -\nabla^2 f(x, y)^{-1} \nabla f(x, y)$$

Exercise 2: Quasi-Newton Method: BFGS (10P)

Learn a logistic regression using the BFGS Quasi-Newton Method from the lecture, for the following data:

$$A = \begin{pmatrix} 1 & 1 & 5 & 2 & 3 \\ 1 & 2 & 1 & 1 & 2 \\ 1 & 3 & 9 & -2 & 6 \\ 1 & 1 & 4 & 0 & -1 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix} \qquad y = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

For a given model the prediction is:

$$\hat{y}(a) = \begin{cases} 1 & \text{if} \quad \beta^T a \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Start with a random initialization and show the convergence of the algorithm! Also start with a diagonal Hessian in the first iteration!