HOMEWORK PROBLEMS 02, ANLY 561, FALL 2017
DUE 09/22/17

Readings: Continue reading Lecture 01 Notes; Goodfellow and Bengio, Chapter 2; and Chapter 2 from `https://jakevdp.github.io/PythonDataScienceHandbook/`

Exercises:

1. Compute (by hand and showing your work) $x^{(1)}$ and $x^{(2)}$ of backtracking with steepest descent increments for $f(x) = x^2$ with $\alpha = \beta = 1/2$ and starting at $x^{(0)} = 1$. You may use the Python implementation of backtracking to check your work and that inequalities hold, but your answers should be exact.

2. For $f \in C^2(\mathbb{R})$, show that the Newton update $x^{(k+1)} = x^{(k)} - f'(x^{(k)})/f''(x^{(k)})$ coincides with the critical point of the second order Taylor approximation to f when $f''(x^{(k)}) \neq 0$:

$$g(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2} f''(x^{(k)})(x - x^{(k)})^2$$

   If f is also convex, what more can be said about this critical point?

3. Using the Python implementations of backtracking and accelerated backtracking provided in class, compute, display, and comment on the convergence of the function values and the iterates when minimizing the following functions using increments from both steepest descent and Newton's method:

   (a) $f(x) = e^x/x$

   (b) $f(x) = \log(x^2+1) + \log((x-1)^2+1) + \log\left(\left(x - \frac{3}{2}\right)^2 + 1\right)$

   (c) $f(x) = -\frac{1}{x^2+1} - \frac{1}{(x-1)^2+1} - \frac{1}{\left(x-\frac{3}{2}\right)^2+1}$

   Use three different starting points to initialize the different iterations for comparison: $x^{(0)} = 10, 100, 1000$, and run 50 iterations in every case.

4. For the next two parts, consider the function $f(x) = \frac{1}{4}x^2$ and the constrained program $\min f(x)$ subject to $x \in [0, 3]$.

   (a) Show that the optimal solution to the centering step for the log-barrier method is $x^{(0)} = 1$ by solving

$$\min f(x) - \log(x) - \log(3 - x)$$

   and assuming $\log(y) = \infty$ when $y \leq 0$.

   (b) Using the optimal centering initialization, compute (by hand and showing your work) $x^{(1)}_{\text{outer}}$ and $x^{(2)}_{\text{outer}}$ from the log-barrier method using 2 inner backtracking iterations with steepest descent increments for $f(x) = \frac{1}{4}x^2$ with $\alpha = \beta = 1/2$. This means that you will compute 4 backtracking steps by hand. Use $M = 2$ so that $x^{(1)}_{\text{outer}}$ is a numerical approximation to the solution of

$$\min f(x) - \frac{1}{2}\log(x) - \frac{1}{2}\log(3 - x)$$

   and then $x^{(2)}$ is a numerical approximation to the solution of

$$\min f(x) - \frac{1}{4}\log(x) - \frac{1}{4}\log(3 - x)$$

   You may use the Python implementations of the log-barrier method and backtracking to check your work and that inequalities hold, but your answers should be exact.

5. Suppose you are give the data

i	1	2	3	4	5	6	7	8	9	10
x_i	-1	-1	-1	0	0	0	1	1	1	1
y_i	-1	-1	-1	1	-1	1	1	1	1	-1

with the goal of fitting a model $P(y = 1|x;\alpha) = \text{logit}(\alpha x y) = \frac{1}{1+e^{-\alpha x y}}$ using the maximum likelihood principle. For the negative log likelihood, compute two steps of "stochastic gradient descent" by hand using the "random" subsets of indices $Q_1 = \{2,5,7,8\}$, $Q_2 = \{1,2,410\}$ and step sizes $\gamma^{(0)} = \frac{1}{2}$, $\gamma^{(1)} = \frac{2}{5}$, and using $\alpha^{(0)} = 1$. That is,

$$\alpha^{(1)} = \alpha^{(0)} - \gamma^{(0)}\Phi'_{Q_1}(\alpha^{(0)})$$

and

$$\alpha^{(2)} = \alpha^{(1)} - \gamma^{(1)}\Phi'_{Q_2}(\alpha^{(1)}),$$

where Φ_Q is as defined in the lecture notes.