Math 164: Optimization – Homework 6

Due: November 4, 2016

Exercise 1.

Exercises 8.18 and 8.20 from the book of Chong and Zak.

Exercise 2.

Exercises 8.22 and 9.1 (without (d-e)) from the book of Chong and Zak.

Exercise 3.

Exercises 9.3 and 10.5 from the book of Chong and Zak.

Exercise 4.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as $f(x) = \frac{1}{2}x^\top Qx$, where $Q = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$. We aim to construct some algorithms that approximate some local minimizers of f.

(1) Find all the local minimizers of f. Are they global? Why?

(2) Explain why one cannot use the usual Newton algorithm to find some of the minimizers from (1).

(3) Can we use the steepest descent algorithm with *optimal step size* to approximate some one of the local minimizers of f? Explain why or why not. If yes, compute 2 iterations using this algorithm. Here one supposes that x^0 is not a local minimizer of f.

(4) Can we use the steepest descent algorithm with *fixed step size* $\alpha > 0$ to approximate some one of the local minimizers of f? Explain why or why not. Is there some $\alpha > 0$ such that this algorithm would converge in one step to some of the local minimizers globally? If yes, what is the geometric relation between x^0 and x^1? Here one supposes that x^0 is not a local minimizer of f.

(5) In (2) one has seen that one cannot use the usual Newton method in this situation. So, let us consider the following modification of if: define the sequence of possible approximations as

$$x^{k+1} = x^k - \left(D^2 f(x^k) + \alpha_k I_2\right)^{-1} \nabla f(x^k),$$

for some sequence of positive real numbers α_k, that we have to construct as well. Why is this sequence now well-defined? Write down explicitly x^{k+1} in terms of x^k and α_k. What is happening if we use a sequence such that $\lim_{k\to+\infty} \alpha_k = 0$? Can we use a constant sequence $\alpha_k = \alpha > 0$ if one wishes to have convergence in one step? Here also, one supposes that x^0 is not a local minimizer of f. What is happening if α_k is very large?