Universität Konstanz Sommersemester 2014
Fachbereich Mathematik und Statistik
Prof. Dr. Stefan Volkwein
Roberta Mancini, Carmen Gräßle, Laura Lippmann, Kevin Sieg

Optimierung

http://www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/

Sheet 5

Deadline for hand-in: 2014/07/02 at lecture

Exercise 15 (Classical Newton method)
Consider the functions $f, g : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = x^3 - 2x + 2, \quad \text{and} \quad g(x) = \sin(x).$$

1. Show that for the starting point $x_0 = 0$, the classical Newton iteration of f has two accumulation points which are both not roots of f. Find another initial point which does not lead to a convergence of the Newton method applied on f.

2. Find a starting point x_0 such that the Newton iteration for g tends to $+\infty$.

3. Explain why the methods do not converge to a root of the functions by a suitable plot.

Exercise 16 (Damped Newton)
Consider to use the classical Newton method for finding the root of $g(x) = \arctan(x)$. Let us be at a generic step ν of the algorithm and that $x_k = 5$. Determine the next iteration point x_{k+1} according to the Newton algorithm

$$\begin{aligned} g'(x_k)\Delta x_k &= -g(x_k); \\ x_{k+1} &= x_k + \Delta x_k. \end{aligned}$$

Then, let us modify the previous algorithm in the following manner (damped Newton algorithm)

$$\begin{aligned} g'(x_k)\Delta x_k &= -g(x_k); \\ x_{k+1} &= x_k + \alpha_k \Delta x_k, \end{aligned}$$

with $\alpha_k \in (0, 1)$, computed by a line search procedure wich returns an α_k such that

$$\|g(x_k + \alpha_k \Delta x_k)\| \leq (1 - \delta\alpha_k)\|g(x_k)\|.$$

Using $\delta = 0.5$ and starting from $\alpha = 1$ in the line search procedure and dividing it by 2 every time the condition is not satisfied, show what the next point x_{k+1} would be using the damped Newton algorithm and compare it to the one returned by the standard Newton method. (Hint: note that the root of $g(x)$ is $x = 0$.)

Exercise 17 (4 Points)

Let $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ be a quadratic function of the form

$$f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle + \gamma,$$

$Q \in \mathbb{R}^{n\times n}$ symmetric and positive definite, $c \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$. Let $x_0 \in \mathbb{R}^n$ and H be a symmetric, positive definite matrix.

Define $\tilde{f}(x) := f(H^{-\frac{1}{2}}x)$ and $\tilde{x}_0 = H^{\frac{1}{2}}x_0$. Let $(\tilde{x}_k)_{k\in\mathbb{N}}$ be a sequence generated by the steepest descent method,

$$\tilde{x}_{k+1} = \tilde{x}_k + \tilde{t}_k\tilde{d}_k \quad \text{with} \quad \tilde{d}_k = -\nabla\tilde{f}(\tilde{x}_k)$$

and $\tilde{t}_k = t(\tilde{d}_k)$ the optimal stepsize choice as determined in Exercise 5.

Let $(x_k)_{k\in\mathbb{N}}$ be generated by the gradient-like method with preconditioner H,

$$x_{k+1} = x_k + t_k d_k \quad \text{with} \quad d_k = H^{-1}(-\nabla f(x_k))$$

and $t_k = t(d_k)$ the optimal stepsize choice as determined in Exercise 5.

Show (by induction) that the two optimization methods are equivalent, i.e., for all $k \in \mathbb{N}$ it holds:

$$x_k = H^{-\frac{1}{2}}\tilde{x}_k.$$