TU Chemnitz
C. Helmberg

Chemnitz, October 5, 2017
due on October 17, 2017

Nonlinear Optimization
Exercises 1

1. (3) Compute the first and second derivative of
   (a) $f(x) = x^T Qx + b^T x$, with $Q^T = Q \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $x \in \mathbb{R}^n$,
   (b) $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$, $x \in \mathbb{R}^2$ (Rosenbrock function),
   (c) $f(Ax)$ for $f : \mathbb{R}^m \to \mathbb{R}^k$ and $A \in \mathbb{R}^{m\times n}$.

2. (3) Let A be a symmetric matrix of order n. Show that the following statements are equivalent:
   (i) A is positive semidefinite ($A \succeq 0$), i.e., $v^T Av \geq 0 \ \ \forall v \in \mathbb{R}^n$.
   (ii) $\lambda_i \geq 0$ where λ_i is an eigenvalue of A, $i = 1, 2, \ldots, n$.
   (iii) $\exists C \in \mathbb{R}^{m\times n}$ such that $A = C^T C$ (and $\operatorname{rank}(C) = \operatorname{rank}(A)$).

   What changes if A is positive definite ($A \succ 0$ or $v^T Av > 0 \ \forall v \in \mathbb{R}^n \setminus \{0\}$)?

4. (3) Using Matlab, have a look at the contour plot of the Rosenbrock function and write a routine that generates for an arbitrary point in $[-1, 2] \times [-\frac{1}{2}, 3]$ a mesh plot of the Rosenbrock function with its linear and quadratic model.

5. (6) Landau Symbols: For $g : \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}^n$ and $B_\varepsilon(a) = \{x : \|x - a\| \leq \varepsilon\}$,

$$\begin{aligned} O(g) &= \{f : \mathbb{R}^n \to \mathbb{R} \mid \exists\, c > 0\, \exists\, \varepsilon > 0\, \forall\, x \in B_\varepsilon(a) : \ |f(x)| \leq c|g(x)|\}, \\ \mathbf{o}(g) &= \{f : \mathbb{R}^n \to \mathbb{R} \mid \forall\, c > 0\, \exists\, \varepsilon > 0\, \forall\, x \in B_\varepsilon(a) : \ |f(x)| \leq c|g(x)|\}, \\ \Theta(g) &= \{f : \mathbb{R}^n \to \mathbb{R} \mid \exists\, c_1, c_2 > 0\, \exists\, \varepsilon > 0\, \forall\, x \in B_\varepsilon(a) : \ c_1|g(x)| \leq |f(x)| \leq c_2|g(x)|\}. \end{aligned}$$

   Given $f, g : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$ prove that for $x \to a$
   (a) $O(f) + O(g) = O(|f| + |g|)$,
   (b) $O(f)O(g) = O(fg)$,
   (c) if $f(x) > 0$ for $x \in B_\varepsilon(a)$ for some $\varepsilon > 0$ then $f \cdot O(g) = O(fg)$.

   Note, it is general practice to write $f = O(g)$ instead of $f \in O(g)$.

6. (2) Show that for $f \in C^2(\mathbb{R}^n)$ and $h \to 0$ it holds that

$$f(x + h) = f(x) + \nabla f(x)^T h + o(\|h\|),$$

$$f(x + h) = f(x) + \nabla f(x)^T h + O(\|h\|^2).$$

Send `*.m` files to `helmberg@mathematik.tu-chemnitz.de`