Univ.-Prof. Dr. Radu Ioan Boţ

Vienna, December 13, 2019
To be discussed on January 10, 2020

# Nonlinear Optimization

## Exercise session 6

38. (Frobenius norm) Let be $A = (a_{ij}) \in \mathbb{R}^{n \times n}$.

   (a) Prove that

   $$\|A\|_F := \left( \sum_{i,j=1}^{n} a_{ij}^2 \right)^{1/2}$$

   defines a matrix norm and that this matrix norm is not induced by a vector norm.

   (b) We denote by

   $$\operatorname{Spur}(A) := \sum_{i=1}^{n} a_{ii}$$

   the trace of the matrix $A$. Prove that

   $$\operatorname{Spur}(AB) = \operatorname{Spur}(BA) \ \forall B \in \mathbb{R}^{n \times n},$$

   $$\operatorname{Spur}(S^{-1}AS) = \operatorname{Spur}(A) \ \forall S \in \mathbb{R}^{n \times n} \text{ with } S \text{ regular},$$

   and

   $$\|A\|_F^2 = \operatorname{Spur}(A^T A).$$

   (c) If $\{v^1, ..., v^n\}$ is an orthonormal basis of $\mathbb{R}^n$, then it holds

   $$\|A\|_F^2 = \sum_{i=1}^{n} \|Av^i\|^2.$$

(4 points)

39. Let be $u, v \in \mathbb{R}^n$. Show that the following statements are true:

   (a) $\|uv^T\| = \|uv^T\|_F = \|u\|\|v\|$.

   (b) $det(I + uv^T) = 1 + u^T v$.

(4 points)

40. (Lemma 9.10 in the lecture course)

    Let be $s, y \in \mathbb{R}^n$ with $s \neq 0$. Prove that $s^T y > 0$ if and only if there exists a symmetric and positive definite matrix $Q \in \mathbb{R}^{n \times n}$ such that $Qs = y$.

    *Hint.* "$\Rightarrow$" For $v := \sqrt{\frac{y^T s}{s^T s}}\, s$, prove that the matrix

    $$Q := \left( I + \frac{1}{v^T v}(y - v)v^T \right) \left( I + \frac{1}{v^T v}(y - v)v^T \right)^T$$

    has the desired properties. (4 points)

41. Let $H \in \mathbb{R}^{n \times n}$ be a regular matrix and $u, v \in \mathbb{R}^n$. Prove that the matrix $H + uv^T$ is regular if and only if $1 + v^T H^{-1} u \neq 0$. Show that under these assumptions the so-called *Sherman-Morrison formula* holds:

    $$(H + uv^T)^{-1} = \left( I - \frac{1}{1 + v^T H^{-1} u} H^{-1} u v^T \right) H^{-1}.$$

    *Hint.* Use Exercise 39(b). (3 points)

42. Let be $s, y \in \mathbb{R}^n$ with $s^T y > 0$ and $H \in \mathbb{R}^{n \times n}$ a symmetric and positive definite matrix. The *direct BFGS update formula* reads

    $$H_+^{BFGS} := H + \frac{1}{s^T y} y y^T - \frac{1}{s^T H s} H s s^T H.$$

    Show that $B = H^{-1}$ implies $B_+^{BFGS} = (H_+^{BFGS})^{-1}$, whereby $B_+^{BFGS}$ is given by the inverse BFGS update formula (Theorem 9.12 in the lecture course). (3 points)

43. Let $M \in \mathbb{R}^{n \times n}$ be a regular matrix and $\{M_k\}_{k \geq 0} \in \mathbb{R}^{n \times n}$ a sequence of matrices which converges to $M$ as $k \to +\infty$. Show that there exists $k_0 \geq 0$ such that $M_k$ is regular for all $k \geq k_0$, and that the sequence $\{M_k^{-1}\}_{k \geq k_0}$ converges to $M^{-1}$. (3 points)

44. Consider the quadratic optimization problem

    $$(P) \qquad \begin{array}{rl} \min & f(x) := \gamma + c^T x + \frac{1}{2} x^T Q x, \\ \text{s.t.} & h(x) := b^T x = 0 \end{array}$$

    with $Q \in \mathbb{R}^{n \times n}$ a symmetric and positive definite matrix, $b, c \in \mathbb{R}^n$, $b \neq 0$, and $\gamma \in \mathbb{R}$. For given $\alpha > 0$, find the minimum $x^*(\alpha)$ of the penalty function

    $$P(x; \alpha) := f(x) + \frac{\alpha}{2}(h(x))^2,$$

    determine $x^* := \lim\limits_{\alpha \to +\infty} x^*(\alpha)$, and prove that $x^*$ is the unique optimal solution of the optimization problem $(P)$.

    *Hint.* Use the *Sherman-Morrison formula.* (3 points)

45. Implement the local inverse BFGS algorithm in MATLAB. To do this, create the file `quasi_newton_method.m` with

```
function [X, Iter] = quasi_newton_method(fhandle, x0, B0, kmax, eps)
```

as first line. Here, `func` denotes a function handle, `x0` the starting vector, `B0` the starting matrix, `kmax` the maximal number of allowed iterations, and `epsilon` the parameter for the stopping criterion. A matrix $X = [x0, x1, x2, ..]$ containing the iteration history and the number of performed iterations `Iter` should be returned.

The implemented algorithm should be tested for `epsilon`$=10^{-6}$, `kmax`$=500$ and `B0` the identity matrix on the following functions and starting values:

(a) $f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$ (Rosenbrock function) for `x0`$=(-1.2, 1)^T$.

(b) $f(x_1, x_2, x_3, x_4) = \sum_{i=1}^4 \left(4 - \sum_{j=1}^4 \cos x_j + i(1 - \cos x_i) - \sin x_i\right)^2$ (trigonometric function) for `x0`$=\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)^T$.

(c) $f(x_1, x_2) = (x_1 - 10^6)^2 + (x_2 - 2 \cdot 10^6)^2 + (x_1 x_2 - 2)^2$ (Brown function) for `x0`$=(1, 1)^T$.

(d) $f(x_1, x_2, x_3, x_4) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2 + 10(x_2 + x_4 - 2)^2 + \frac{1}{10}(x_2 - x_4)^2$ (Wood function) for `x0`$=(-3, -1, -3, -1)^T$.

The generated files should be sent by January 17, 2020 to `radu.bot@univie.ac.at`.