## Master in Mathematics and Applications - Técnico, Lisbon

Numerical Functional Analysis and Optimization - Fall Semester 2016

## $5^{\text {th }}$ Problem set

1. Consider the quadratic functional $f(x)=\frac{1}{2} x^{T} A x-b^{T} x+c$, where $A \in \mathbb{R}^{N \times N}$ is a symmetric and positive semi-definite matrix, $b \in \mathbb{R}^{N}$ and $c \in \mathbb{R}$.
a) Show that every stationary point of $f$ is a global minimizer of $f$.
b) Prove that if $A$ positive definite then $f$ has a unique global minimizer.
2. Let $f: \mathbb{R}^{N} \rightarrow \mathbb{R}$ be twice continuously differentiable in an open convex set $K \subset \mathbb{R}^{N}$ and assume that the Hessian matrix of $f$ satisfies the following Lipschitz condition at $x \in K$ :

$$
\exists \gamma \geq 0 \text { tal que }\left\|H_{f}(y)-H_{f}(x)\right\| \leq \gamma\|y-x\| \quad \forall y \in K
$$

Show that

$$
\left|f(x+p)-f(x)-\nabla f(x)^{T} p-\frac{1}{2} p^{T} H_{f}(x) p\right| \leq \frac{\gamma}{6}\|p\|^{3} \quad \forall p \in \mathbb{R}^{N} \text { such that } x+p \in K .
$$

3. Consider the linear least squares problem

$$
\min _{x \in \mathbb{R}^{N}}\|A x-b\|_{2}^{2}
$$

where $A \in \mathbb{R}^{M \times N}, M>N$, and $b \in \mathbb{R}^{M}$. Use the necessary and sufficient conditions of optimality for unconstrained minimization problems to analyse the existence and uniqueness of solutions to the least squares problem.
4. Consider the unconstrained minimization problem $\min _{x \in \mathbb{R}^{2}} f(x)$, where $f(x)=\frac{1}{2} x_{1}^{2}+x_{1} \cos x_{2}$.
a) Determine the stationary and saddle points as well as the local and global minimizers of $f$.
b) Choosing $x^{(0)}=[10]^{T}$ and $B_{0}=H_{f}\left(x^{(0)}\right)$, compute the matrix $B_{1}$ from the BFGS method's update formula.
5. Given $F: \mathbb{R}^{N} \rightarrow \mathbb{R}^{N}$, consider the nonlinear system $F(x)=0$ and the unconstrained minimization problem

$$
\begin{equation*}
\min _{x \in \mathbb{R}^{N}} f(x) \tag{1}
\end{equation*}
$$

where $f(x)=\frac{1}{2} F(x)^{T} F(x)$.
a) Relate the zeros of $F$ with the local and global minimizers of $f$.
b) Compute $\nabla f(x)$ and $H_{f}(x)$.
c) Show that the direction

$$
p_{k}=-J_{F}\left(x_{k}\right)^{-1} F\left(x_{k}\right),
$$

i.e. Newton's direction for the nonlinear system $F(x)=0$, is a descent direction for $f$ at $x_{k}$, provided $J_{F}\left(x_{k}\right)$ is non singular and $F\left(x_{k}\right) \neq 0$.
d) Consider the quadratic model

$$
m_{k}(s)=\frac{1}{2}\left(F\left(x_{k}\right)+J_{F}\left(x_{k}\right) s\right)^{T}\left(F\left(x_{k}\right)+J_{F}\left(x_{k}\right) s\right)
$$

Show that $p_{k}$ is the (unique) global minimizer of $m_{k}$ and that $p_{k}$ is a descent direction for $m_{k}$.
6. Considere the quadratic function $f: \mathbb{R}^{N} \rightarrow \mathbb{R}, f(x)=\frac{1}{2} x^{T} A x-b^{T} x+c$, where $A \in \mathbb{R}^{N \times N}$ is a symmetric and positive semi-definite matrix, $b \in \mathbb{R}^{N}$ and $c \in \mathbb{R}$.
a) Show that the solution of the one-dimensional minimization problem

$$
\min _{\alpha_{k}>0} f\left(x_{k}-\alpha_{k} \nabla f\left(x_{k}\right)\right)
$$

is given by

$$
\alpha_{k}=\frac{\left\|A x_{k}-b\right\|_{2}^{2}}{\left\|A x_{k}-b\right\|_{A}^{2}},
$$

where $\|x\|_{A}=\sqrt{(x, A x)}$. Conclude that the method of steepest descent reduces to:
Given $x_{0} \in \mathbb{R}^{N}$, compute

$$
x_{k+1}=x_{k}-\frac{\left\|A x_{k}-b\right\|_{2}^{2}}{\left\|A x_{k}-b\right\|_{A}^{2}}\left(A x_{k}-b\right), \quad k=0,1, \ldots .
$$

b) Write the method of steepest descent as

$$
x_{k+1}=x_{k}+\alpha_{k} r_{k}, \quad k=0,1, \ldots,
$$

where $r_{k}=b-A x_{k}$ and $\alpha_{k}=\frac{\left\|r_{k}\right\|_{2}^{2}}{\left\|r_{k}\right\|_{A}^{2}}$, let $x_{*} \in \mathbb{R}^{N}$ be the global minimizer of $f$, i.e. $A x_{*}=b$, and define the error vector $e_{k}=x_{*}-x_{k}$.
Prove that

$$
\begin{equation*}
\left\|e_{k+1}\right\|_{A}^{2}=\left(1-\frac{\left\|r_{k}\right\|_{2}^{4}}{r^{T} A r_{k} r_{k}^{T} A^{-1} r_{k}}\right)\left\|e_{k}\right\|_{A}^{2}, \tag{2}
\end{equation*}
$$

where $\|x\|_{A}=\sqrt{(x, A x)}$.
c) Kantorovich Lemma: Let $A \in \mathbb{R}^{N \times N}$ be symmetric and positive semi-definite and assume that $0<$ $\lambda_{1} \leq \lambda_{2} \leq \ldots \leq \lambda_{N}$ are the eigenvalues of $A$. Then

$$
\frac{\left(x^{T} x\right)^{2}}{x^{T} A x x^{T} A^{-1} x} \geq \frac{4 \lambda_{1} \lambda_{N}}{\left(\lambda_{1}+\lambda_{N}\right)^{2}} \quad \forall x \in \mathbb{R}^{N} \backslash\{0\} .
$$

Use the Kantorovich lemma in the estimate (2) to conclude that

$$
\left\|e_{k+1}\right\|_{A} \leq \frac{\lambda_{N}-\lambda_{1}}{\lambda_{N}+\lambda_{1}}\left\|e_{k}\right\|_{A}
$$

d) Verify that

$$
f\left(x_{k}\right)-f\left(x_{*}\right)=\frac{1}{2}\left\|e_{k}\right\|_{A}^{2} .
$$

7. Consider a quasi-Newton line search method $x_{k+1}=x_{k}+\alpha_{k} p_{k}$, with the search direction

$$
p_{k}=-B_{k}^{-1} \nabla f\left(x_{k}\right),
$$

where $B_{k} \in \mathbb{R}^{N \times N}$ is symmetric and positive definite matrix. Assume that the step length $\alpha_{k}>0$ satisfies the Wolfe conditions and that $\operatorname{cond}_{2}\left(B_{k}\right) \leq M$. Suppose that the objective function $f: \mathbb{R}^{N} \rightarrow \mathbb{R}$ is bounded below in $\mathbb{R}^{N}$ and that $f$ is continuously differentiable in a convex open set $D \subset \mathbb{R}^{N}$ containing the level set $L\left(x_{0}\right)=\left\{x: f(x) \leq f\left(x_{0}\right)\right\}$ where $x_{0}$ is the starting point of the iteration. Assume also that $\nabla f$ is Lipschitz continuous in $D$.

Show that

$$
\frac{-\nabla f\left(x_{k}\right)^{T} p_{k}}{\left\|\nabla f\left(x_{k}\right)\right\|_{2}\left\|p_{k}\right\|_{2}} \geq 1 / M,
$$

and conclude that the line search method is globally convergent.
8. Let $f: \mathbb{R}^{N} \rightarrow \mathbb{R}$ be twice continuously differentiable and assume that the Hessian matrix $H_{f}$ is Lipschitzian in a neighborhood of $x_{*}$, a point where the sufficient conditions of optimality are satisfied. Show that the sequence $\left\{\left\|\nabla f\left(x_{k}\right)\right\|_{2}\right\}$, generated by the Newton's iteration $x_{k+1}=x_{k}-H_{f}^{-1}\left(x_{k}\right) \nabla f\left(x_{k}\right)$, converges to zero with a quadratic rate.
9. Consider the unconstrained minimization problem

$$
\begin{equation*}
\min _{x \in \mathbb{R}^{N}} f(x) \tag{3}
\end{equation*}
$$

for the quadratic objective function $f(x)=\frac{1}{2} x^{T} A x-b^{T} x+c$, where the matrix $A \in \mathbb{R}^{N \times N}$ is symmetric and positive definite, $b \in \mathbb{R}^{N}$ and $c \in \mathbb{R}$. Let $x_{*} \in \mathbb{R}^{N}$ the global minimizer of $f$.
Show that the method of steepest descent, for the numerical solution of (3), converges in one iteration, if $x_{0}=x_{*}+\beta u$, where $u \in \mathbb{R}^{N}$ is an eigenvector associated with one of the eigenvalues of $A$ and $\beta \in \mathbb{R}$.
10. Consider the following unconstrained minimization problem

$$
\begin{equation*}
\min _{x \in \mathbb{R}^{N}}\|x\|^{3} \tag{4}
\end{equation*}
$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^{N}$.
Show that the pure Newton's method, applied to problem (4) reduces to

$$
x^{(k+1)}=\frac{1}{2} x^{(k)}, \quad k=0,1, \ldots
$$

Determine the rate of convergence of Newton's method in this case.
Hint: You may find the following formula useful:

$$
\left(I+\frac{x x^{T}}{\|x\|^{2}}\right)^{-1}=\left(I-\frac{x x^{T}}{2\|x\|^{2}}\right)
$$

11. Let $A \in \mathbb{R}^{N \times N}$ be a non singular matrix, $U, V \in \mathbb{R}^{N \times M}$, with $1 \leq M \leq N$, and assume that $\operatorname{det}\left(I+V^{T} A^{-1} U\right) \neq 0$.
a) Establish the Sherman-Morrison-Woodbury formula

$$
\left(A+U V^{T}\right)^{-1}=A^{-1}-A^{-1} U\left(I+V^{T} A^{-1} U\right)^{-1} V^{T} A^{-1}
$$

b) Let $U=\left[u_{1}, u_{2}\right], V=\left[v_{1}, v_{2}\right]$, with $u_{j}, v_{j} \in \mathbb{R}^{N}$. Show that

$$
\left(A+\sum_{j=1}^{2} u_{j} v_{j}^{T}\right)^{-1}=A^{-1}-A^{-1} U C^{-1} V^{T} A^{-1}
$$

where

$$
C=\left[\begin{array}{cc}
1+v_{1}^{T} A^{-1} u_{1} & v_{1}^{T} A^{-1} u_{2} \\
v_{2}^{T} A^{-1} u_{1} & 1+v_{2}^{T} A^{-1} u_{2}
\end{array}\right]
$$

12. Consider the following update formula for the BFGS method which generates symmetric and positive definite approximations of the inverse of the Hessian matrix

$$
\begin{equation*}
H_{k+1}^{-1}=\left(I-\rho_{k} s_{k} y_{k}^{T}\right) H_{k}^{-1}\left(I-\rho_{k} y_{k} s_{k}^{T}\right)+\rho_{k} s_{k} s_{k}^{T}, \quad \rho_{k}=\frac{1}{y_{k}^{T} s_{k}}>0 \tag{5}
\end{equation*}
$$

a) Use the Sherman-Morrison-Woodbury formula to derive expression (5) from the BFGS update of matrix $H_{k}$.
b) Show that $H_{k+1}^{-1}=H_{k}^{-1}+E$, with $\operatorname{rank}(E)=2$.
c) Confirm that $H_{k+1}^{-1}$ is symmetric if $H_{k}^{-1}$ is symmetric.
d) Show that $H_{k+1}^{-1}$ satisfies the secant equation $H_{k+1}^{-1} y_{k}=s_{k}$.
e) Show that $H_{k+1}^{-1}$ is positive definite if $H_{k}^{-1}$ is positive definite,

