Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization I Spring 2020

Exercise set 7

1 Assume that $A \in \mathbb{R}^{m\times n}$ with $m < n$ is a matrix of full rank and that $b \in \mathbb{R}^m \setminus \{0\}$. Consider the optimization problem

$$\min_x \frac{1}{2}\|x\|^2 \qquad \text{subject to} \qquad Ax = b. \tag{1}$$

a) Formulate the KKT-conditions for this problem and show that the unique solution is given by

$$x^* = A^T(AA^T)^{-1}b.$$

b) Formulate the quadratic penalty method for this constrained optimization problem, and show that the unique minimizer with parameter $\mu > 0$ is given by

$$x_\mu := A^T\Big(\frac{1}{\mu}I + AA^T\Big)^{-1}b$$

with $I \in \mathbb{R}^{m\times m}$ denoting the identity matrix.

c) Now consider the optimization problem

$$\min_x \frac{1}{2}\|x\|^2 \qquad \text{subject to} \qquad \frac{1}{2}\|Ax-b\|^2 \le \varepsilon$$

for some $\varepsilon > 0$, and denote its solution by $\hat{x}_\varepsilon$. Show that either $\frac{1}{2}\|b\|^2 \le \varepsilon$ (in which case $\hat{x}_\varepsilon = 0$), or there exists $\mu > 0$ such that $\hat{x}_\varepsilon = x_\mu$.

2 Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function, and let x_k and x_{k+1} $\mathbb{R}^n$ be two iterates of some iterative algorithm with $x_k \neq x_{k+1}$.

Define

$$s_k := x_{k+1} - x_k, \qquad y_k := \nabla f(x_{k+1}) - \nabla f(x_k).$$

The BFGS update for $H_k \approx (\nabla^2 f(x_k))^{-1}$ is

$$H_{k+1} = (I - \rho_k s_k y_k^\top)H_k(I - \rho_k y_k s_k^\top) + \rho_k s_k s_k^\top,$$

where $\rho_k = \frac{1}{y_k^\top s_k}$.

a) Show that if H_k is symmetric positive definite and $y_k^\top s_k > 0$, then H_{k+1} is symmetric positive definite as well.
Hint: Try to show that $u^\top H_{k+1}u > 0$ for all nonzero u.

b) Show that if f is strongly convex[1] then $y_k^\top s_k > 0$ for any choice of nonequal x_k and x_{k+1}.

[1] A C^2 function f is strongly convex if there exists an $m > 0$ such that $p^\top \nabla^2 f(x) p \ge m\|p\|^2$ for all $x \in \mathbb{R}^n$ and $p \in \mathbb{R}^n$.