SECTION C: CONTINUOUS OPTIMISATION
PROBLEM SET 2

HONOUR SCHOOL OF MATHEMATICS, OXFORD UNIVERSITY
HILARY TERM 2006, DR RAPHAEL HAUSER

Instructions: All four problems are intended as a homework assignments. Please hand in solutions by Monday noon and clearly mark them for "Nicolas Jeannequin". Hand in at the Comlab reception desk. Thanks!

***Problem 1.**

(i) Consider the Sherman-Morrison-Woodbury formula

$$(B + UV^{\mathrm{T}})^{-1} = B^{-1} - B^{-1}U(\mathrm{I} + V^{\mathrm{T}}B^{-1}U)^{-1}V^{\mathrm{T}}B^{-1},$$

where B is a $n \times n$ invertible matrix, U and V are $n \times m$ matrices with $m \leq n$, I is the $m \times m$ identity matrix, and where we assume that the matrix $\mathrm{I} + V^{\mathrm{T}}B^{-1}U$ is invertible. Show that the formula holds true by multiplying it with $(B + UV^{\mathrm{T}})$.

(ii) Can it happen that B and $B + UV^{\mathrm{T}}$ are invertible but $I + V^{\mathrm{T}}B^{-1}U$ is not, so that the formula is not applicable? (Hint: did you ever use the condition $m \leq n$ in part (i)?)

***Problem 2.** Let $f \in C^2(\mathbb{R}^n, \mathbb{R})$, and let $x_k \in \mathbb{R}^n$ be such that $D^2 f(x_k) \succ 0$.

(i) Write down the 2nd order Taylor approximation $q(x)$ of f around x_k.
(ii) Find the global minimiser x^* of $q(x)$.
(iii) How does $x^* - x_k$ compare to the Newton-Raphson step for f applied at x_k?

***Problem 3.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function

$$x \mapsto \frac{1}{2} x^{\mathrm{T}} \begin{pmatrix} \kappa & 0 \\ 0 & 1 \end{pmatrix} x.$$

(i) Show that f is strictly convex and has the global minimiser $x^* = 0$.
(ii) Consider applying the steepest descent method with exact line searches (see Lectures 2 and 3) to f starting at the point $x_0 = (1, \kappa)^{\mathrm{T}}$. Show that

$$x_k = \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \begin{pmatrix} (-1)^k \\ \kappa \end{pmatrix}.$$

(iii) Show that this sequence converges Q-linearly to x^*.
(iv) Describe what you observe if the problem is solved in the new coordinates

$$y = \begin{pmatrix} \kappa^{\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix} x.$$

(v) Do the same again in the coordinates

$$z = \begin{pmatrix} \cos\sqrt{\kappa^2+\kappa^3} & -\sin\sqrt{\kappa^2+\kappa^3} \\ \sin\sqrt{\kappa^2+\kappa^3} & \cos\sqrt{\kappa^2+\kappa^3} \end{pmatrix} y,$$

and use part (ii) and (iii) to analyse the convergence of the sequence $(z_k)_{\mathbb{N}_0}$.

***Problem 4.** Consider a quasi-Newton formula in which $B^{(k)}$ approximates the Hessian $D^2 f(x_k)$ and is updated by

$$B^{(k+1)} = B^{(k)} + \frac{\eta\delta^{\mathrm{T}} + \delta\eta^{\mathrm{T}}}{\delta^{\mathrm{T}}\delta} - \frac{\eta^{\mathrm{T}}\delta}{(\delta^{\mathrm{T}}\delta)^2}\delta\delta^{\mathrm{T}},$$

where $\eta = \gamma - B^{(k)}\delta$, $\gamma = \nabla f(x^{(k+1)}) - \nabla f(x^{(k)})$ and $\delta = x^{(k+1)} - x^{(k)}$.

(i) Show that this update satisfies the secant equation.

(ii) Show that when applied to a quadratic function with Hessian G, we have

$$B^{(k+1)} - G = \left(\mathrm{I} - \frac{\delta\delta^{\mathrm{T}}}{\delta^{\mathrm{T}}\delta}\right)(B-G)\left(\mathrm{I} - \frac{\delta\delta^{\mathrm{T}}}{\delta^{\mathrm{T}}\delta}\right).$$

(iii) Show that if $B^{(k)} = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, $\delta = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\gamma = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, then $\eta = -\delta$ and hence, $B^{(k+1)} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$.

(iv) What implication do (ii) and (iii) have for use of the updating formula in quasi-Newton methods?