SECTION C: CONTINUOUS OPTIMISATION
REVISION CLASS 1, PART 2

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

Problem 1 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x) = x_1 x_2$ where $x = (x_1, x_2)$, and consider the trust region $B_\Delta(\widehat{x}) = \{x : \|x - \widehat{x}\| \leq \Delta\}$ around $\widehat{x} = (1, 1/2)$, where $\|\cdot\|$ denotes the Euclidean norm. We are interested in solving the trust-region subproblem

$$\text{(TR)} \qquad \min_{x \in B_\Delta(\widehat{x})} f(x).$$

(i) Show that the Cauchy point is $y^c = (3/8, -3/4)$.
(ii) Find an explicit parameterisation of the dogleg path.
(iii) Show that the objective function is decreasing on the first segment of the dogleg path and increasing on the second segment. Why does this phenomenon occur?
(iv) Conclude that there exists a critical value $\Delta_c > 0$ such that the dogleg-solution to (TR) is given by the Cauchy point for all $\Delta \geq \Delta_c$.
(v) Apply Steihaug's method with exact line-searches and find a formula for its solution $y^s(\Delta)$ as a function of Δ.
(vi) Show that $f(y^s(\Delta))$ is strictly decreasing as a function of Δ.

Problem 2 Consider the BFGS updating formula

$$B_{k+1} = B_k - \frac{B_k \delta_k \delta_k^{\mathrm{T}} B_k}{\delta_k^{\mathrm{T}} B_k \delta_k} + \frac{\gamma_k \gamma_k^{\mathrm{T}}}{\gamma_k^{\mathrm{T}} \delta_k}$$

(i) Explain the meaning of each of the terms appearing in this formula and explain their connection with the so-called "secant-condition".

Now let f be a strictly convex quadratic function in $x \in \mathbb{R}^n$,

$$f(x) = \frac{1}{2} x^{\mathrm{T}} A x + b^{\mathrm{T}} x + c,$$

where $b \in \mathbb{R}^n$ is a vector and $A \in \mathbb{R}^{n^2}$ a symmetric positive definite matrix. Starting with $B_0 = \mathrm{I}$, we will apply BFGS steps with exact line searches.

(ii) Let S_k be the linear subspace of $\mathbb{R}^n$ such that $s \in S_k$ if $B_k s = f''(x_k) s$ and $\nabla f(x_k)^{\mathrm{T}} s = 0$. Show that if $s \in S_k$ then $\delta_k^{\mathrm{T}} B_k s$ and $\gamma_k s$ are both zero.
(iii) Conclude from this that $s \in S_k$ implies $\nabla f(x_{k+1})^{\mathrm{T}} s = 0$.
(iv) Show that $\delta_k^{\mathrm{T}} B_k s = \gamma_k^{\mathrm{T}} s = 0$ implies $B_{k+1} s = B_k s$ and conclude that

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \cdots \subseteq S_{k+1}.$$

(v) Show that δ_k is not in S_k.
(vi) Show that $\delta_k \in S_{k+1}$, and conclude that $\dim S_k < \dim S_{k+1}$.
(vii) Assuming that we stop the procedure when $\nabla f(x_k) = 0$, show that the algorithm terminates after at most n steps.