CNAc: Continuous Optimization
Problem set 2 — linesearch methods

Honour School of Mathematics, Oxford University
Hilary Term 2006, Dr Nick Gould

Instructions: Asterisked problems are intended as a homework assignment while nonasterisked problems are not compulsory but can further help you understand the material. Please put your solutions in Denis Zuev's pigeon hole at the Maths Institute by 9AM on Monday of 4th week.

***Problem 1†.**

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function

$$f(x) = \tfrac{1}{2}\kappa x_1^2 + \tfrac{1}{2}x_2^2$$

for some $\kappa > 0$.

(i) Show that f has its global minimiser at $x^* = 0$.

(ii) Consider applying the method of steepest descent with an exact line search to f, starting at the point $x_0 = (1, \kappa)^T$. Show that

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

(iii) A sequence x_k is said to converge Q-linearly (with factor ρ) to x_* if

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq \rho$$

for some $0 \leq \rho < 1$. Show that the sequence obtained in (ii) converges Q-linearly. What is the factor ρ? Roughly how many iterations will it take to get the first component of x_* correct to one decimal place if $\kappa = 1000$? What if $\kappa = 10^6$?

(iv) Describe what you observe if the problem is solved in the new coordinates $(y_1, y_2) = (\kappa^{\frac{1}{2}} x_1, x_2)$.

***Problem 2.**

(i) Show that if $g(x)$ is the gradient of the quadratic function $f(x) = g^T x + \frac{1}{2}x^T Bx$, then

$$Bs = y, \quad \text{where } \; y = g(x + s) - g(x).$$

(ii) Suppose that B_k is a given symmetric matrix, and that $B_{k+1} = B_k + \beta vv^T$ is a rank-one correction for which the secant condition

$$B_{k+1}s_k = y_k, \quad \text{where } \; s_k = x_{k+1} - x_k \; \text{ and } \; y_k = g_{k+1} - g_k \tag{1}$$

is required to hold. Show that the resulting B_{k+1} is unique, and give its form; state any assumptions you need to make

(iii) Now suppose that

$$B_{k+1} = B_k + \theta y_k y_k^T + \beta v v^T \tag{2}$$

is a rank-two correction to a given symmetric B_k for which (1) holds. Find an expression for θ, β and v in terms of

$$\phi = 1 - \theta y_k^T s_k;$$

once again state any assumptions you require.

(iv) What special cases occur in the formula you have derived in part (iii) when $\phi = 0$ and $\phi = 1$?

Problem 3.

Show that if B is symmetric positive definite and $s^T y > 0$, then the BFGS updsate

$$B^+ = B + \frac{yy^T}{y^T s} - \frac{Bss^T B}{s^T Bs}$$

is also positive definite.

***Problem 4[†].**

Consider applying the conjugate gradient algorithm to the problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad q(x) = g^T x + \tfrac{1}{2} x^T B x$$

where B is symmetric, positive definite.

(i) Using the fact that $x_1 = x_0 + \alpha_0 d_0$, show that

$$g(x_1) \in \text{span}\{g(x_0), Bg(x_0)\}.$$

where $g(x) = \nabla f(x)$

(ii) Generalise this result and show by induction that

$$g(x_k) \in \mathcal{K}_k := \text{span}\{g(x_0), Bg(x_0), \ldots, B^k g(x_0)\}$$

for $k = 0, \ldots, n$. The spaces $\mathcal{K}_k$ are called **Krylov subspaces**. [Hint: in the lectures we showed that $\text{span}\{d_0, \ldots, d_k\} = \text{span}\{g(x_0), \ldots, g(x_k)\}$.]

(iii) Now let $B =_{\mathcal{I}} +A$ where $\text{rank}(A) = r$. Show that

$$\mathcal{K}_k = \text{span}\{g(x_0), Ag(x_0), \ldots, A^k g(x_0)\}.$$

(iv) Show that $\mathcal{K}_k$ can be at most $r + 1$ dimensional.

(v) Using part (iv) of this exercise and the identity $g(x_k)^T g(x_j)$ for all $j \neq k$ derived in the Lecture notes, show that the conjugate gradient algorithm must terminate after at most $r + 1$ iterations at an iterate that corresponds to the minimiser of f.

[†] Thanks to Raphael Hauser for these examples