Operations Research Techniques and Algorithms (620-361)

Dr Yao-ban Chan

y.chan@ms.unimelb.edu.au
Telephone: 8344 9073
Office: Room 198, Richard Berry Building

Christina Burt
c.burt@ms.unimelb.edu.au
Telephone: 8344 1797
Office: 139 Barry St

Monday 8th April, 2008

Today's Lecture

Unimodal n-D unconstrained optimisation
- BFGS

We show that the BFGS direction does in fact satisfy condition **??**.

$$
\begin{aligned}
& H_{k+1}(\nabla f(x^{k+1}) - \nabla f(x^k)) \\
&= H_{k+1} g^k \\
&= \left[H_k + \frac{1 + \langle r^k, g^k \rangle}{\langle s^k, g^k \rangle} s^k (s^k)^T - s^k (r^k)^T - r^k (s^k)^T \right] g^k \\
&= H_k g^k + \frac{1 + \langle r^k, g^k \rangle}{\langle s^k, g^k \rangle} s^k (s^k)^T g^k - s^k (r^k)^T g^k \\
&\quad - \frac{1}{\langle s^k, g^k \rangle} H_k g^k (s^k)^T g^k \\
&= H_k g^k + (1 + \langle r^k, g^k \rangle) s^k - \langle r^k, g^k \rangle s^k - H_k g^k \\
&= s^k \\
&= x^{k+1} - x^k.
\end{aligned}
$$

It follows from Lemma 7, that if H_k is symmetric and positive definite and x^k is not stationary, then the quasi-Newton direction will be a descent direction.
It can easily be shown that H_k is symmetric.

It follows from the next result that, if we use the line search procedure satisfying the Armijo-Goldstein and Wolff conditions to determine the step size t_k, the BFGS update of the quasi-Newton matrix H_k preserves positive definiteness.

Lemma 9: *Let H_0 be a symmetric positive definite matrix in $\Re^{n \times n}$. If, for each k, $d^k := -H_k \nabla f(x^k)$ and the step size t_k satisfies the Wolff condition (W), then H_k is symmetric and positive definite.*

So the BFGS method is a well defined descent method. It is also a 'good' algorithm, as shown in the next convergence result due to Powell (1976).

Theorem: *Let* $H_0 \in \Re^{n\times n}$ *be symmetric positive definite,* $x^0 \in \Re^n$*,* $\sigma \in (0, 1/2)$*, and* $\mu \in (\sigma, 1)$*. If* f *is convex and* C^2*, and the lower level set*

$$\{x \in \Re^n \,:\, f(x) \leq f(x^0)\}$$

is bounded, then the sequence produced by BFGS has a cluster point x^* *which is a (local and) global minimum of* f*. If, in addition,* $\nabla^2 f(x^*)$ *is positive definite, then* $x^k \to x^*$ *superlinearly.*