AMSC/CMSC 660 Quiz 9 , Fall 2003

1. (10) You are asked to minimize a function of $n = 2000$ variables. Consider doing this by Newton's method, a quasi-Newton method, or Pattern search. Give the main advantages and disadvantages of each. Which would you choose? Why?

Answer:

- Newton: often converges with a quadratic rate when started close enough to a solution, but requires both first and second derivatives (or good approximations of them) as well as storage and solution of a linear system with a matrix of size 2000×2000.
- Quasi-Newton: often converges superlinearly when started close enough to a solution, but requires first derivatives (or good approximations of them) and storage of a matrix of size 2000×2000.
- Pattern search: converges only linearly, but has good global behavior and requires only function values, no derivatives.

If first derivatives (or approximations) were available, I would use QN, with updating of the matrix factorization (or a limited memory version, but we did not talk about this option). Otherwise, I would use pattern search.

2. (10) In Broyden's method for solving nonlinear equations, we need to solve a linear system involving the $n \times n$ matrix

$$B^{(k+1)} = B^{(k)} + \frac{(y - B^{(k)}s)s^T}{s^T s} .$$

Recall the Sherman-Morrison-Woodbury formula

$$(A - ZV^T)^{-1} = A^{-1} + A^{-1}Z(I - V^T A^{-1} Z)^{-1} V^T A^{-1} .$$

If we have a way to solve linear systems involving $B^{(k)}$ using p multiplications, how long would it take to solve a linear system involving $B^{(k+1)}$?

Answer: For some vector r, we need to form

$$(A - ZV^T)^{-1} r = A^{-1} r + A^{-1} Z (I - V^T A^{-1} Z)^{-1} V^T A^{-1} r, ,$$

where $A = B^{(k)}$, $Z = y - B^{(k)}s$, and $V = -s/(s^T s)$. Thus we need to

- form $t = A^{-1}r$ and $u = A^{-1}Z$, at a cost of $2p$ multiplications,
- form $\alpha = 1 - V^T u$ at a cost of n multiplications,
- form $y = ((V^T t)/\alpha)u$ at a cost of $2n$ multiplications and 1 division
- add t and y.

The total number of multiplications is $2p + 3n + 1$.