Adaptive Low Complexity Algorithms for Unconstrained Minimization

Carmine Di Fiore, Stefano Fanelli, Paolo Zellini
Dipartimento di Matematica, Università di Roma "Tor Vergata"
Via della Ricerca Scientifica, 1 – 00133 Roma, Italy
emailto: difiore@mat.uniroma2.it

September 20 – 24, 2004, Cortona, Italia

The $\mathcal{L}QN$ methods, recently introduced in [1], [2], are minimization algorithms particularly competitive in solving large scale problems. In order to minimize a function $f : \mathbb{R}^n \to \mathbb{R}$, the $\mathcal{L}QN$ algorithms use a quasi-Newton iterative scheme having $O(n \log n)$ complexity per step. The Hessian approximation B_{k+1} in $\mathbf{x}_{k+1} = \mathbf{x}_k - \lambda_k B_k^{-1} \nabla f(\mathbf{x}_k)$ is defined in terms of a suitable structured matrix $\mathcal{L}_{B_k}$, by the updating formula $B_{k+1} = \varphi(\mathcal{L}_{B_k}, \mathbf{s}_k, \mathbf{y}_k)$, where

$$\varphi(A, \mathbf{s}, \mathbf{y}) = A + \frac{1}{\mathbf{s}^T \mathbf{y}} \mathbf{y}\mathbf{y}^T - \frac{1}{\mathbf{s}^T A \mathbf{s}} A\mathbf{s}\mathbf{s}^T A,$$

$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$, and $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$. The matrix $\mathcal{L}_{B_k}$ is chosen in a fixed matrix algebra $\mathcal{L}$ (f.i. $\mathcal{L}$ =Hartley algebra) and is defined as the best least squares fit to B_k in $\mathcal{L}$. The $\mathcal{L}QN$ algorithms work since $\mathcal{L}_{B_k}$ inherites positive definiteness from B_k.

Here we propose to change the structure of $\mathcal{L}_{B_k}$ by an adaptive procedure. More precisely, for each step k we consider the matrix $\mathcal{L}_{\mathbf{sy}} \in \mathcal{L}$ solving the secant equation $X\mathbf{s}_{k-1} = \mathbf{y}_{k-1}$ and we test if it is positive definite (pd) (like $\mathcal{L}_{B_k}$). If no, then we redefine $\mathcal{L}$ so that $\mathcal{L}_{\mathbf{sy}}$ is pd.

Our aim (see also [3],[4]) is in fact to make $\mathcal{L}_{B_k}$ in the $\mathcal{L}QN$ updating formula close to $\mathcal{L}_{\mathbf{sy}}$ during the minimization procedure, so to assign to the φ-updated matrix both the spectral information and the secant property of the matrix B_k used in the original well known $BFGS$ method.

References

1 C. Di Fiore, S. Fanelli, F. Lepore, P. Zellini, Matrix algebras in quasi-Newton methods for unconstrained minimization, *Numerische Mathematik*, 94 (2003), pp.479-500.

2 A. Bortoletti, C. Di Fiore, S. Fanelli, P. Zellini, A new class of quasi-Newtonian methods for optimal learning in MLP-networks, *IEEE Transactions on Neural Networks*, 14 (2003), pp.263-273.

3 C. Di Fiore, S. Fanelli, P. Zellini, Low complexity minimization algorithms, *Numerical Linear Algebra with Applications*, to appear.

4 C. Di Fiore, S. Fanelli, P. Zellini, On the best least squares fit to a matrix and its applications, in *Advances in Pure and Applied Algebra*, Nova Science Publishers, Inc., to appear.