Why adjoint based least squares solving ought to be optimal

Andreas Griewank

Department of Mathematics, Humboldt-Universität zu Berlin, Germany
School of Information Sciences, Yachaytech, Ibarra, Ecuador

September 2, 2015

Numerical Methods for Large-Scale Nonlinear Problems and their Applications
ICERM, Brown University, Providence, RI

with thanks to
Andrea Walther(PAB) and Sebastian Schlenkrich(TUD)
Sandra Schneider(HUB) and Claudia Tutsch(CLU)

Setting

Problem

$$\min \varphi(x) \equiv \tfrac{1}{2}\|F(x)\|_2^2 \quad \text{for} \quad F : \mathbb{R}^n \mapsto \mathbb{R}^m \quad \text{with} \quad n \le m$$

First order optimality condition (necessary)

$$0 = \nabla\varphi(x_*) \equiv F(x_*)^\top F'(x_*) \ \in \mathbb{R}^n$$

Second order optimality condition (sufficient)

$$1 > \kappa_* \equiv \|R_*^{-\top} \sum_{i=1}^{m} F_i(x)\nabla^2 F_i(x) R_*^{-1}\|_2 \quad \text{with} \quad F'(x_*) = Q_* \mathbb{R}_*$$

Derivative availability and cost

$$\frac{\mathrm{OPS}\{\dot{y} \equiv F'(x)\dot{x}\}}{\mathrm{OPS}\{y \equiv F(x)\}} \ \le \ 3 \quad , \quad 4 \ \ge \ \frac{\mathrm{OPS}\{\bar{x}^\top \equiv \bar{y}^\top F'(x)\}}{\mathrm{OPS}\{y \equiv F(x)\}}$$

Gauss Adjoint Broyden Method

Tangent conditions for $B \approx F'$

$$B_+ s = y \equiv F'(x_+)\, s \in \mathbb{R}^m \quad \text{and} \quad B_+^\top \sigma = F'(x_+)^\top \sigma \in \mathbb{R}^n$$

Transposed Broyden Update

$$B_+ = B + \frac{\sigma\sigma^\top}{\sigma^\top\sigma}(F'(x_+) - B) \quad \textit{for} \quad \sigma = y \quad \text{and} \quad \sigma = r \equiv y - Bs$$

yields rank-two update, which can be implemented in $O(mn)$ operations.

Resulting Properties

Frobenius norm change minimality, domain transformation invariance, and heredity on affine systems $F(x) = Ax - b$.

Quasi-Gauss-Newton Iteration

$$x_+ = x - \alpha(B^\top B)^{-1}\nabla\varphi(x) \quad \text{with} \quad \alpha \quad \text{by Andersen(m=1)}$$

Provable Properties

Global convergence

$$0 = \inf_k \|\nabla\varphi(x_k)\| \quad \Leftarrow \quad x_0 \in \{\varphi(x) \le c\} \text{ compact and } \operatorname{rank}(F'(x)) = n$$

Asymptotic R-rate in overdetermined case $(m > n)$

$$0 = \inf_k \|x_k - x_*\| \;\Rightarrow\; \limsup_{k\to\infty} \|x_k - x_*\|^{\frac{1}{k}} \;\le\; \kappa_* \;<\; 1$$

Asymptotic order in consistent case $(m = n)$

$$\liminf_{k\to\infty} |\log(\|x_k - x_*\|)|^{\frac{1}{k}} \;\ge\; \rho_n \;\approx\; 1 + \frac{\log(n)}{n} \quad \text{with} \quad 1 = \rho_n^{n+1} - \rho_n^n$$

On affine problems

Finite termination in $\le n$ steps, (á la GMRES when $m = n$ and $B_0 = I$.)