Gauss-Newton methods with approximate projections for solving constrained nonlinear least squares problems

Max L. N. Gonçalves & Tiago C. Menezes

Goiás Federal University

maxlng@ufg.br and tiagmenezes1@gmail.com

XIII celebrating Alfredo Iusem's 70th birthday

Brazilian Workshop on Continuous Optimization

IMPA, Rio de Janeiro, Brazil

September 23rd to 27th, 2019

Introduction

Let $\Omega \subseteq \mathbb{R}^n$ an open set and $F : \Omega \to \mathbb{R}^m$ a continuously differentiable nonlinear function. Consider the following constrained nonlinear least squares problem

$$\min_{x \in C} \; G(x) := \frac{1}{2}\|F(x)\|^2, \tag{1}$$

where $C \subseteq \Omega$ is a nonempty convex closed set. In order to solve the problem, when $F'(x)$ is injective for all $x \in \Omega$, a proximal Gauss-Newton method for solving a more general class of constrained nonlinear least squares problems was proposed in [3]. Under the assumption that F' is Lipschitz continuous, local convergence results of the proximal Gauss-Newton method were established in [3]. Depending on the application, the exact projection used in [3] can be extremely difficult to obtain. Consequently, the first goal of this paper is to propose an extension of the algorithm in [3] in which inexact projections can be admitted. Hence, the first method to be proposed here basically consists of computing the unconstrained Gauss-Newton step, and then an approximate projection of it, with respect to the metric defined onto C. Our analysis is done by using a majorant condition. Another issue in [3] is that no globalization strategy was considered. Therefore, the second goal of this paper is to propose a global version of our first method. Our globalization technique is based on the efficient nonmonotone line search in [1]. In order to illustrate the robustness and effectiveness of the new schemes, we report some numerical experiments on a set of box- and polyhedral-constrained nonlinear systems and compare their performances with the proximal Gauss-Newton method in [3].

Basic assumptions

We suppose that the following assumptions hold:

(A1) $\langle F'(x_*)^*F(x_*), x - x_*\rangle \geq 0$, for all $x \in C$, and $F'(x_*)$ is injective;

(A2) The sequence $\{\theta_k\}$ satisfies $\theta_k \leq \bar{\theta}$ for all $k \geq 0$, where $\bar{\theta} \in [0, 1)$.

For simplicity, let us consider the following constants: $c := \|F(x_*)\|$, $\beta := \|F'(x_*)^\dagger\|$, $\kappa := \beta\,\|F'(x_*)\|$ and $\delta := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$, where $R > 0$ is a given scalar.

The method and its local convergence

The GNM-AP method is formally described as follows.

GNM-AP

Step 0 (Initialization). Let $x_0 \in C$, $\{\theta_k\} \subset [0, \infty)$ be given, and set $k = 0$.

Step 1 (Gauss-Newton step). Compute $s_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^n$ such that

$$F'(x_k)^*F'(x_k)s_k = -F'(x_k)^*F(x_k), \qquad y_k = x_k + s_k.$$

Step 2 (Computation of new iterative). Define $H_k = F'(x_k)^*F'(x_k)$. Compute $x_{k+1} \in C$ such that

$$\langle y_k - x_{k+1}, x - x_{k+1}\rangle_{H_k} \leq \varepsilon_k := \theta_k^2\|x_{k+1} - x_k\|^2_{H_k}, \qquad \forall\, x \in C,$$

i.e. x_{k+1} is an ε_k–approximate projection of y_k onto C.

Step 3 (Termination criterion and update). If $x_{k+1} = x_k$, then **stop**; Otherwise, set $k \leftarrow k+1$ and go to step 1.

end

In the following, we state our local convergence result satisfying a Lipschitz condition.

Theorem 1. *Suppose that there exists a $L > 0$ such that*

$$\lambda = \frac{[(1+\sqrt{2})\kappa + 1]\, c\,\beta L + \kappa\bar{\theta}}{(1-\bar{\theta})} < 1, \qquad \beta\left\|F'(x) - F'(x_* + \tau(x - x_*))\right\| \leq L(1-\tau)\sigma(x),$$

where $x \in B(x_*, \delta)$, $\tau \in [0, 1]$ and $\sigma(x) = \|x - x_*\|$. Let be given the positive constant

$$r := \min\left\{\frac{4 + \kappa - 2\bar{\theta}(1+\kappa) + 2c(1+\sqrt{2})\beta - \sqrt{\left[4 + \kappa - 2\bar{\theta}(1+\kappa) + 2c(1+\sqrt{2})\beta\right]^2 - 8(1-\lambda)(1-\bar{\theta})}}{2L}, \delta\right\}.$$

Then the GNM-AP with starting point $x_0 \in C \cap B(x_*, r)\backslash\{x_*\}$ is well-defined, the generated $\{x_k\}$ is contained in $B(x_*, r) \cap C$, converges to x_* and satisfies

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|$$

and

$$\|x_{k+1} - x_*\| \leq \frac{\kappa L + L^2\sigma(x_0)}{2(1-\theta_k)[1 - L\sigma(x_0)]^2}\|x_k - x_*\|^2 +$$
$$\frac{[(1+\sqrt{2})\kappa + 1]c\beta L + c(1+\sqrt{2})\beta L^2\sigma(x_0)}{(1-\theta_k)[1 - L\sigma(x_0)]^2}\|x_k - x_*\| + \frac{\theta_k(L\sigma(x_0) + k)}{(1-\theta_k)(1 - L\sigma(x_0))}\|x_k - x_*\|, \quad \forall k = 0, 1, \ldots$$

Globalized method

We now present a globalized version of the GNM-AP.

Global GNM-AP (G-GNM-AP)

Step 0 (Initialization). Let $x_0 \in C$, $\tau \in (0, 1)$, $\eta_1, \eta_2 > 0$, an integer $M \geq 1$, and $\{\theta_k\} \subset [0, \infty)$ be given, and set $k = 0$.

Step 1 (Inexact projected Gauss-Newton direction). If $F'(x_k)^*F'(x_k)$ is non-singular, then compute $y_k \in \mathbb{R}^n$ and $\tilde{P}_C^{H_k}(y_k) \in C$ such that

$$y_k = x_k - H_k^{-1}F'(x_k)^*F(x_k), \quad \langle y_k - \tilde{P}_C^{H_k}(y_k), x - \tilde{P}_C^{H_k}(y_k)\rangle_{H_k} \leq \varepsilon_k, \quad \forall\, x \in C, \tag{2}$$

where $H_k = F'(x_k)^*F'(x_k)$ and $\varepsilon_k := \theta_k^2\|\tilde{P}_C^{H_k}(y_k) - x_k\|^2_{H_k}$. If

$$\langle \tilde{P}_C^{H_k}(y_k) - x_k, F'(x_k)^*F(x_k)\rangle \leq -\eta_1\|\tilde{P}_C^{H_k}(y_k) - x_k\|^2, \quad \|\tilde{P}_C^{H_k}(y_k) - x_k\| \leq \eta_2\|F'(x_k)^*F(x_k)\|, \tag{3}$$

then set $d_k = \tilde{P}_C^{H_k}(y_k) - x_k$ and go to **Step 3**.

Step 2 (Inexact projected gradient direction). Compute $y_k \in \mathbb{R}^n$ and $\tilde{P}_C(y_k) \in C$ such that

$$y_k = x_k - F'(x_k)^*F(x_k), \quad \langle y_k - \tilde{P}_C(y_k), x - \tilde{P}_C(y_k)\rangle \leq \varepsilon_k := \theta_k^2\|\tilde{P}_C(y_k) - x_k\|^2, \quad \forall\, x \in C,$$

and set $d_k = \tilde{P}_C(y_k) - x_k$.

Step 3 (Backtracking). Define $G_{max} = \max\{G(x_{k-j});\ 0 \leq j \leq \min\{k, M-1\}\}$. Set $\alpha \leftarrow 1$.

Step 3.1 Set $x_+ = x_k + \alpha d_k$.

Step 3.2 If

$$G(x_+) \leq G_{max} + \tau\alpha\langle F'(x_k)^*F(x_k), d_k\rangle,$$

then $\alpha_k = \alpha$, $x_{k+1} = x_+$, and go to step 4. Otherwise, set $\alpha \leftarrow \alpha/2$ and go to step 3.2.

Step 4 (Termination criterion and update). If $x_{k+1} = x_k$, then **stop**; otherwise, set $k \leftarrow k+1$ and go to step 1.

end

In the following, we state our main global convergence result for the G-GNM-AP method.

Theorem 2. *Assume that the level set $\Omega_0 := \{x \in C : G(x) \leq G(x_0)\}$ is bounded and the sequence $\{\theta_k\}$ satisfies $\theta_k \leq \bar{\theta}$ for all $k \geq 0$, where $\bar{\theta} \in [0, 1)$. Then, either the G-GNM-AP stops at some stationary point x_k, or every limit point of the generated sequence is stationary.*

Numerical experiments

We took $\theta_k = 1/3$, for every k, in both algorithms. The ε_k–approximate projection of point y_k onto C was computed by the conditional gradient method. In order to avoid an excessive number of inner iterations, the input ε_k was replaced by $\max\{\theta_k^2\|x_{k+1} - x_k\|^2_{H_k}, 10^{-2}\}$. The other initialization parameters of the G-GNM-AP method were set $\tau = 10^{-4}$ and $M = 10$. For a comparison purpose, we also run the proximal Gauss-Newton (Prox-GN) method in [3], which, applied to (1), corresponds to our GNM-AP method with exact projections. In the latter method, the exact projections were computed by the MATLAB command *quadprog*. For the GNM-AP, G-GNM-AP and Prox-GN methods, we used the same termination condition $\|x_{k+1} - x_k\|_{H_k} < 10^{-4}$. For all algorithms, a failure was declared if the number of iterations was greater than 300 or no progress was detected.

Box-constrained nonlinear least squares problems

We consider a set of 23 problems of the form (1) with $C = \{x \in \mathbb{R}^n; c \leq x \leq d\}$, where $c, d \in R^n$. We firstly chose 10 initial points of the form $x_0(\gamma) = c + (\gamma/11)(d - c)$ for $\gamma = 1, 2, \ldots, 10$.

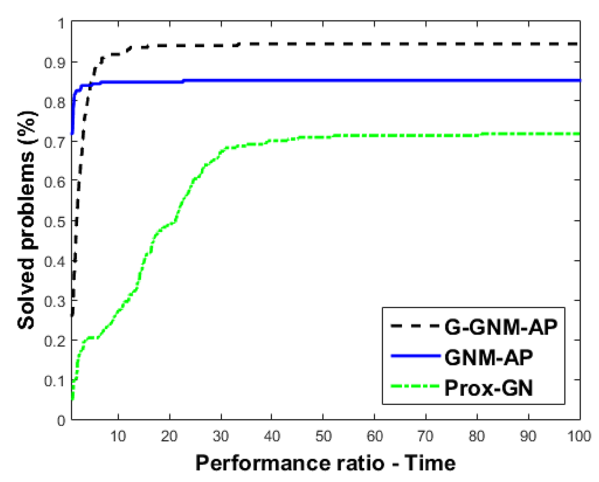

Figure 1: Performance of the G-GNM-AP, GNM-AP and Prox-GN methods

Polyhedral-constrained nonlinear least squares problems

We consider a set of 23 test problems of the form (1) with $C = \{x \in \mathbb{R}^n; c \leq x \leq d,\ Ax \leq b\}$, where $c, d \in R^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$.

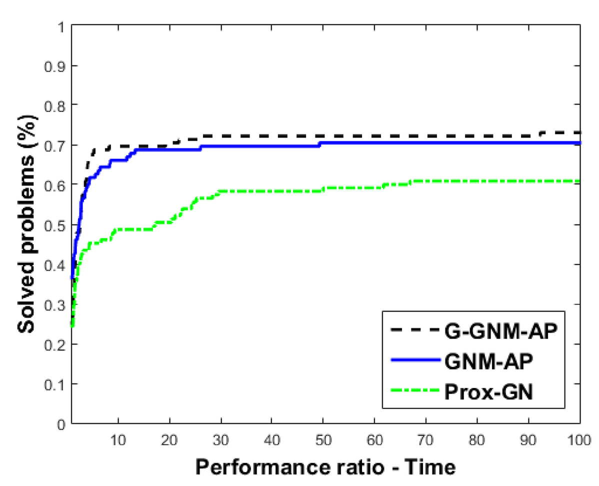

Figure 2: Performance of the G-GNM-AP, GNM-AP and Prox-GN methods

References

[1] L. Grippo, F. Lampariello, and S. Lucidi. A nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.*, 23(4):707–716, 1986.

[2] Margherita Porcelli. On the convergence of an inexact Gauss-Newton trust-region method for nonlinear least-squares problems with simple bounds. *Optim. Lett.*, 7(3):447–465, 2013.

[3] Saverio Salzo and Silvia Villa. Convergence analysis of a proximal Gauss-Newton method. *Comput. Optim. Appl.*, 53(2):557–589, 2012.

Acknowledgements

CAPES

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico