Variable precision trust region methods

Anuj Bajaj, Matt Menickelly, Paul Hovland, Jan Hueckelheim, and Sri Hari Krishna Narayanan

May 6, 2021

1 Pseudocode

Input: $x_0 \in \mathbb{R}^n$ (initial point), $\delta_0 > 0$ (initial trust-region radius), $\epsilon > 0$ (final gradient accuracy), $N \in \mathbb{N}$ (budget constraint)

Initialize: $0 < \eta_1 \leq \eta_2 < 1$, $0 < \gamma_1 < 1 \leq \gamma_2$, $\kappa > 0$ (large enough), $\kappa_{\text{low}}, \epsilon_{\text{tol}} > 0$ (small enough), $B_0 \in \mathbb{R}^{n \times n}$, $\omega \in \{\omega_h = 10^{-4}, \omega_s = 10^{-8}$, or $\omega_d = 0\}$ (initial precision), counter $= 0$, max_counter > 0;

for $k = 0, 1, 2, \ldots, N$ **do**

Compute $f_k = f(x_k; \omega)$, $g_k = g(x_k; \omega)$

Terminate:

if $\|g_k\| \leq \epsilon$ **then**

- $\kappa_{\text{low}} = 0$,
- **if** $\omega = \omega_d$ **then**
  - **terminate**;
  - **else if** $\omega = \omega_s$ **then**
    - Compute $f_k = f(x_k; \omega_d)$, $g_k = g(x_k; \omega_d)$
  - **else if** $\omega = \omega_h$ **then**
    - Compute $f_k = f(x_k; \omega_s)$, $g_k = g(x_k; \omega_s)$

Step calculation: Compute s_k such that $\|s_k\| \leq \delta_k$ which sufficiently decreases the model

$$q(x_k, s) = g_k^T s + \frac{1}{2} s^T B_k s$$

where $g_k \approx \nabla f_k$, and B_k is an L-SR1 approximation to $\nabla^2 f_k$.

Evaluate the function: Compute $\tilde{f}_k = f(x_k + s_k; \omega), \tilde{g}_k = g(x_k + s_k; \omega)$

Acceptance of trial point: Define

$$\rho_k = \frac{f_k - \tilde{f}_k}{q(x_k, 0) - q(x_k, s_k)} = \frac{f_k - \tilde{f}_k}{-q(x_k, s_k)}.$$

(Failure) **if** $\rho_k < \eta_1$ or $q(x_k, s_k) > 0$ **then**

- $x_{k+1} = x_k$
- **if** $\omega = \omega_h$ ($= \omega_s$ resp.) **then**
  - counter = counter+1;
  - **if** counter=max_counter **then**
    - counter $= 0$; $\tilde{\omega} = \omega_s$ ($= \omega_d$ resp.)
    - Compute $f(x_k + s_k; \tilde{\omega}), g(x_k + s_k; \tilde{\omega})$
    - **if** $|f(x_k + s_k; \tilde{\omega}) - (f_k + q(x_k, s_k))| > \kappa \delta_k^2$ **then**
      - $f_k = f(x_k; \tilde{\omega}), g_k = g(x_k; \tilde{\omega})$, and $\tilde{g}_k = g(x_k + s_k; \tilde{\omega})$
      - **else**
      - $\delta_{k+1} = \gamma_1 \delta_k$
    - **if** $\tilde{\omega} = \omega_d$ and $|f(x_k + s_k; \tilde{\omega}) - (f(x_k; \omega) + q(x_k, s_k))| < \kappa_{\text{low}} \delta_k^2$ **then**
      - $f_k = f(x_k; \omega_h), g_k = g(x_k; \omega_h)$, $\tilde{g}_k = g(x_k + s_k; \tilde{\omega})$
    - **else**
    - $\delta_{k+1} = \gamma_1 \delta_k$
- **if** $\omega = \omega_d$ **then**
  - $\delta_{k+1} = \gamma_1 \delta_k$

(Success) **if** $\rho_k > \eta_1$ or $q(x_k, s_k) < 0$ **then**

- $x_{k+1} = x_k + s_k$
- **if** $|f_k + q(x_k, s_k - \tilde{f}_k)| < \kappa_{\text{low}} \delta_k^2$ **then**
  - **if** $\omega = \omega_d$ ($= \omega_s$ resp.) **then**
    - $f_k = f(x_{k+1}; \omega_s)$ ($=f(x_{k+1}; \omega_h)$ resp.), $g_k = g(x_{k+1}; \omega_s)$ ($=g(x_{k+1}; \omega_h)$ resp.)
- **if** $|f(x_k; \omega) + q(x_k, s_k - f(x_k + s_k; \omega))| > \kappa_{\text{low}} \delta_k^2$ **then**
  - $f_k = f(x_{k+1}; \omega), g_k = g(x_{k+1}; \omega)$

Radius update (Expand):

if $\rho_k \geq \eta_2$ and $\|s_k\| \geq 0.8 \delta_k$ **then**

- $\delta_{k+1} = \gamma_2 \delta_k$

No Progress:

if $\delta_k \leq \epsilon_{\text{tol}}$ **then**

- break;