R. Herzog, G. Müller, E. Herberg, M. Hashemi
Heidelberg University

Exercise 6

Date issued: 22nd May 2023
Date due: 31st May 2023

Homework Problem 6.1 (Truncated Newton CG) 6 Points

Implement the truncated Newton-CG method (Algorithm 5.44 with Algorithm 5.41), apply it for Rosenbrock's and/or Himmelblau's functions and compare its performance with the exact globalized Newton method.

Homework Problem 6.2 (Affine-Invariance of BFGS/DFP-updated quasi Newton) 6 Points

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be differentiable, $A \in \mathbb{R}^{n\times n}$ be invertible, $b \in \mathbb{R}^n$ and $g(y) := f(Ay+b)$.

(*i*) Let the sequences $\left(x^{(k)}\right)$ and $\left(y^{(k)}\right)$ be generated by applying full quasi Newton steps as in

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - {H_f^{(k)}}^{-1} f'(x^{(k)})^\mathsf{T} && \text{from } x^{(0)} = Ay^{(0)} + b && \text{with } H_f^{(0)} \text{ s. p. d.} \\ y^{(k+1)} &= y^{(k)} - {H_g^{(k)}}^{-1} g'(y^{(k)})^\mathsf{T} && \text{from } y^{(0)} && \text{with } H_g^{(0)} = A^\mathsf{T} H_f^{(0)} A. \end{aligned}$$

Show that $x^{(k)} = Ay^{(k)} + b$ for all $k \in \mathbb{N}$, when the BFGS or the DFP update are applied to update the model Hessians.

(*ii*) Let the sequences $\left(x^{(k)}\right)$ and $\left(y^{(k)}\right)$ be generated by applying full inverse quasi Newton steps as in

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - B_f^{(k)} f'(x^{(k)})^\mathsf{T} && \text{from } x^{(0)} = Ay^{(0)} + b && \text{with } B_f^{(0)} \text{ s. p. d.} \\ y^{(k+1)} &= y^{(k)} - B_g^{(k)} g'(y^{(k)})^\mathsf{T} && \text{from } y^{(0)} && \text{with } B_g^{(0)} = A^{-1} B_f^{(0)} A^{-\mathsf{T}}. \end{aligned}$$

Show that $x^{(k)} = Ay^{(k)} + b$ for all $k \in \mathbb{N}$, when the inverse BFGS or the inverse DFP update are applied to update the inverse of the model Hessians.

Hint: You can save yourselves some work using the connection of the updates of the Hessians and their inverses.

Note: The restriction to unit step length scalings in this exercise is to keep the required notation slim(er). Since we know that the Armijo and the curvature condition are affine invariant as well, we don't loose invariance when applying step lengths satisfying these conditions.

Homework Problem 6.3 (Inverse BFGS and DFP Updates) 6 Points

Derive the inverse BFGS and DFP update formulas

$$\Psi_{\mathrm{BFGS}}(B, s, y) = (\mathrm{Id} - \rho\, s\, y^{\mathsf{T}})\, B\, (\mathrm{Id} - \rho\, y\, s^{\mathsf{T}}) + \rho\, s\, s^{\mathsf{T}}, \tag{5.60}$$

$$\Psi_{\mathrm{DFP}}(B, s, y) = B - \frac{B\, y\, y^{\mathsf{T}} B}{y^{\mathsf{T}} B\, y} + \rho\, s\, s^{\mathsf{T}} \tag{5.59}$$

using the Sherman-Morrison-Woodbury formula from Lemma 5.50.

Homework Problem 6.4 (Two-Loop Recursion for Inverse BFGS Update) 6 Points

Show that Algorithm 5.53 in fact computes the action of the inverse BFGS updated matrix $B_{\mathrm{BFGS}}^{(k)}$.

Please submit your solutions as a single pdf and an archive of programs via moodle.