Numerical Optimization 2024 - Homework 6

Aral Cimcim, k11720457

Artificial Intelligence, JKU Linz

May 15, 2024

Problem 1.

SR1 algorithm:

Given starting point x_0, initial Hessian approx. B_0,
trust region radius Δ_0, tolerance $\epsilon > 0$,
parameters $\eta \in (0, 10^{-3})$ and $r \in (0, 1)$;
$k \leftarrow 0$;
while $\|\nabla f(x_k)\| > \epsilon$;
Compute s_k by solving the subproblem

$$\min_s \nabla f_k^T s + \frac{1}{2} s^T B_k s \quad \text{s.t.} \quad \|s\| \leq \Delta_k; \tag{1}$$

Compute

$$y_k = \nabla f(x_k + s_k) - \nabla f_k,$$
$$\text{ared} = f(x_k) - f(x_k + s_k),$$
$$\text{pred} = -\Big(\nabla f_k^T s_k + \tfrac{1}{2} s_k^T B_k s_k\Big),$$

if ared/pred $> \eta$;
else $x_{k+1} = x_k$;
end (if)
if ared/pred > 0.75
if $\|s_k\| \leq 0.8\Delta_k$
$\Delta_{k+1} = \Delta_k$;
else
$\Delta_{k+1} = 2\Delta_k$;
end (if)
elif $0.1 <$ ared/pred ≤ 0.75
$\Delta_{k+1} = \Delta_k$;
else
$\Delta_{k+1} = 0.5\Delta_k$;
end (if)

if (6.26) holds
Use (6.24) to compute B_{k+1} (even if $x_{k+1} = x_k$);
else
$B_{k+1} \leftarrow B_k$;
end (if)
$k \leftarrow k + 1$;
end (while)

(6.26):

$$|s_k^T(y_k - B_k s_k)| \geq r\|s_k\|\|y_k - B_k s_k\| \tag{2}$$

(6.24):

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k} \tag{3}$$

Function to minimize:

$$f(x) = \frac{1}{2}x^T Q x - b^T x \tag{4}$$

$B_0 = H_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$p_0 = -H_0 \nabla f(x_0) = (1, 1)^T$

$x_1 = x_0 + p_0 = (1, 1)^T$

$\nabla f(x_1) = (0.5, -0.16)^T$

$y_0 = (1.5, 0.83)^T$

$s_0 = (1, 1)^T$

$B_1 = B_0 + \frac{(y_0 - B_0 s_0)(y_0 - B_0 s_0)^T}{(y_0 - B_0 s_0)^T s_0}$

$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} 0.25 & -0.083 \\ -0.083 & 0.027 \end{bmatrix}}{0.334}$

$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.75 & -0.25 \\ -0.25 & 0.08 \end{bmatrix}$

$B_1 = \begin{bmatrix} 1.75 & -0.25 \\ -0.25 & 1.08 \end{bmatrix}$

$H_1 = B_1$

$p_1 = -H_1 \nabla f(x_1)$

$p_1 = \begin{bmatrix} -1.75 & 0.25 \\ 0.25 & -1.08 \end{bmatrix} (0.5, -0.16)^T$

$$p_1 = \begin{bmatrix} -0.915 \\ 0.298 \end{bmatrix}$$

$$x_2 = x_1 + p_1$$

$$x_2 = (1,1)^T + \begin{bmatrix} -0.915 \\ 0.298 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} 0.085 \\ 1.298 \end{bmatrix}$$

$$\nabla f(x_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.085 \\ 1.298 \end{bmatrix} - \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\nabla f(x_2) = \begin{bmatrix} 0.085 \\ 1.298 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\nabla f(x_2) = \begin{bmatrix} -0.915 \\ 0.298 \end{bmatrix}$$

$$y_1 = \begin{bmatrix} -0.915 \\ 0.298 \end{bmatrix} + \begin{bmatrix} 0.085 \\ 1.298 \end{bmatrix}$$

$$y_1 = \begin{bmatrix} -0.83 \\ 1.596 \end{bmatrix}$$

$$s_1 = \begin{bmatrix} -0.915 \\ 0.298 \end{bmatrix}$$

$$B_2 = B_1 + \frac{(y_1 - B_1 s_1)(y_1 - B_1 s_1)^T}{(y_1 - B_1 s_1)^T s_1}$$

$$B_2 = \begin{bmatrix} 1.75 & -0.25 \\ -0.25 & 1.08 \end{bmatrix} + \frac{\begin{bmatrix} 0.64 & 0.8 \\ 0.8 & 1 \end{bmatrix}}{-0.434}$$

$$B_2 = \begin{bmatrix} 1.75 & -0.25 \\ -0.25 & 1.08 \end{bmatrix} + \begin{bmatrix} -1.5 & -1.8 \\ -1.8 & -2.3 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 0.25 & -2.05 \\ -2.05 & -1.22 \end{bmatrix}$$