3. Consider the function

$$f(x_1, x_2) = -\frac{x_1}{2} + \frac{x_2}{3} - \log(x_1^2) + \log(x_2^2)$$

(a) Let $x^0 = (2, 2)^T$. Apply one step of the gradient (steepest descent) method. To save time, do not do a line search, take the step length of $\lambda = 0.5$, but verify that it satisfies Goldstein-Armijo conditions with $\mu_1 = 0.2,\ \mu_2 = 0.8$.

(b) Let $x^0 = (2, 2)^T$. Apply a full Newton step and give x^1.

(c) Let $x^0 = (2, 2)^T$. Calculate the Trust Region search direction with the initial value $\alpha = 1$. Let choose $\mu = 0.2,\ \eta = 0.9,\ \gamma_1 = 0.5,\ \gamma_2 = 2.5$. Would you accept this step in the Trust Region algorithm or α should be changed? If it needs to be changed, should α be increased or decreased?

(d) Compare the function values $f(x^1)$ computed after one iteration of the gradient method in (a), Newton's method in (b) and Trust Region method in (c). Based on that comparison and properties of the function $f(x_1, x_2)$, which algorithm is more advantageous for minimizing the function $f(x_1, x_2)$. Explain why.

8p

Solution

(a)

$$\nabla f(2,2) = \begin{bmatrix} -1/2 - 2\,{x_1}^{-1} \\ 1/3 + 2\,{x_2}^{-1} \end{bmatrix}_{(2,2)} = (-3/2, 4/3)$$

Therefore the descent direction is $(3/2, -4/3)$

$$x_1 = (2,2) + 0.5(3/2, -4/3) = (11/4, 4/3)$$

$$f(x^0) = -1.0000, \text{ and } f(x^1) = -2.3863$$

$$f(x^0) - f(x^1) = 1.3863$$

$$-\lambda_0 \mu_1 \nabla f(x^0)^T s_0 = 1 * 0.2 * ((-3/2)^2 + (4/3)^2) = 0.4000$$

$$-\lambda_0 \mu_2 \nabla f(x^0)^T s_0 = 1 * 0.8 * ((-3/2)^2 + (4/3)^2) = 1.6000$$

$$-\lambda_0 \mu_1 \nabla f(x^0)^T s_0 < f(x^0) - f(x^1) < -\lambda_0 \mu_2 \nabla f(x^0)^T s_0$$

The Goldstein-Armijo conditions are satisfied.

(b) Newton step:

$$x^1 = x^0 + s^0$$

$$s^0 = -[\nabla^2 f(2,2)]^{-1} \nabla f(2,2)$$

$$\nabla f(2,2) = (-3/2, 4/3)^T$$

$$\nabla^2 f(2,2) = \begin{bmatrix} 2\,x_1{}^{-2} & 0 \\ 0 & -2\,x_2{}^{-2} \end{bmatrix}_{(2,2)} = \begin{bmatrix} 1/2 & 0 \\ 0 & -1/2 \end{bmatrix}$$

$$s^0 = -[\nabla^2 f(2,2)]^{-1} \nabla f(2,2) = (3, 8/3)^T$$

$$x^1 = (2,2)^T + (3, 8/3)^T = (5, 14/3)^T$$

$$f(x^0) - f(x^1) = 1.4463$$

(c) The trust-region step:

$$H = (\nabla^2 f + \alpha I)$$

$$H(2,2) = \begin{bmatrix} 3/2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

$$s^0 = -H(2,2)^{-1} \nabla f(2,2) = (1, -8/3)^T$$

$$x^1 = (2,2)^T + (1, -8/3)^T = (3, -2/3)^T$$

To determine if the step is good enough, let's check the ratio

$$\rho_0 = \frac{f(x^0) - f(x^1)}{f(x^0) - q(x^1)}$$

$$\text{actual decrease} = f(x^0) - f(x^1) = f(2,2) - f(3, -2/3) = -1.0000 - (-5.1972) = 4.1972$$

$$\text{predicted decrease} = f(x^0) - q(x^1) = -\nabla f(x^0)^T (x^1 - x^0) - 1/2 (x^1 - x^0)^T \nabla^2 f(x)(x^1 - x^0) = 2.5278$$

$\rho_0 =$ actual decrease / predicted decrease $= 1.6604$ and as $\rho_0 > \eta$ the step is very good and we accept it. We update $\alpha^1 = 0.5\alpha^0 = 0.5$.