Modern Optimization Final Exam

Guanhua Huang

December 8, 2020

1 Problem 12

12. We again consider the weakened ROSENBROCK function $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $f(x) := 10\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2$.

a) Plot the level curves of f and the level curves of the quadratic approximation of f at the point $x_0 := (0, -1)^T$.

b) Calculate the solutions to the quadratic minimization problem

$$\begin{cases} \varphi(x_0 + d) := f(x_0) + \nabla f(x_0)^T d + \frac{1}{2} d^T \nabla^2 f(x_0) d \longrightarrow \min \\ \|d\| \le \Delta_0 \end{cases}$$

for the trust regions with the radii $\Delta_0 := 0.25, 0.75, 1.25$.

c) Repeat *a)* and *b)* for $x_0 := (0, 0.5)^T$.

1.1 Part a

1.2 Part b

1.3 Part c

2 Problem 15

2.1 Part 1

First, let's write the matrix format of the function,

15. Let $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined by $f(x) := -12x_2 + 4x_1^2 + 4x_2^2 - 4x_1x_2$. Write f as

$$f(x) = \tfrac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle$$

with a positive definite symmetric matrix $A \in \mathbb{M}_2$ and $b \in \mathbb{R}^2$. To $d_1 := (1,0)^T$ find all the vectors $d_2 \in \mathbb{R}^2$ such that the pair (d_1, d_2) is A-conjugate.

Minimize f starting from $x^{(0)} := (-\frac{1}{2}, 1)^T$ in the direction of d_1 and from the thus obtained point in the direction of $d_2 := (1,2)^T$.

Sketch the situation (level curves; $x^{(0)}, x^{(1)}, x^{(2)}$).

$$(x_1, x_2)\begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (b_1, b_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (a_1x_1 + a_2x_2, a_2x_1 + a_3x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + x_1b_1 + x_2b_2$$

$$(a_1x_1 + a_2x_2)x_1 + (a_2x_1 + a_3x_2)x_2 + x_1b_1 + x_2b_2$$

$$= a_1x_1^2 + 2a_2x_1x_2 + a_3x_2^2 + b_1x_1 + b_2x_2$$

$$= -12x_2 + 4x_1^2 + 4x_2^2 - 4x_1x_2$$

By comparing parameters, we got the following,

$$b_2 = -12, b_1 = 0, a_2 = -2, a_1 = a_3 = 4$$

$$A = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}, b = \begin{pmatrix} 0 \\ -12 \end{pmatrix}$$

Based on A-conjugate definition (on page 120), $\langle d_1, Ad_2\rangle = 0, \quad d_1 = (1,0)^T$, then, let's write out the problem as the following:

$$(1,0)\begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = (4,-2)\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = 4\delta_1 - 2\delta_2 = 0$$

So, as long as $2\delta_1 = \delta_2$, the condition $\langle d_1, Ad_2\rangle = 0$ is satisfied

$$d_2 = \begin{pmatrix} \gamma \\ 2\gamma \end{pmatrix}, \quad \forall \gamma \in \mathbb{R}$$

2.2 Part 2

Let's calculate the gradient of the function,

$$f(x) = -12x_2 + 4x_1^2 + 4x_2^2 - 4x_1x_2, \quad \nabla f(x) = \begin{pmatrix} 8x_1 - 4x_2 \\ -12 + 8x_2 - 4x_1 \end{pmatrix} = g(x)$$

Then, we can evaluate $g_0 = \begin{pmatrix} -\frac{1}{2} \times 8 - 4 \times -\frac{1}{2} \\ -12 + 8 + 2 \end{pmatrix} = \begin{pmatrix} -8 \\ -2 \end{pmatrix}$ by given $x_0 = \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}$

In this problem, d_1 is given so we don't have to calculate it by g_0. So, the next stuff we need is the $\lambda_0 = -\frac{\langle g_0, d_0 \rangle}{\langle Ad_0, d_0 \rangle}$. I would calculate the numerator and enumerator separately.

$$-\langle g_0, d_0 \rangle = -(-8, -2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 8$$

$$\langle Ad_0, d_0 \rangle = (1, 0) \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (4, -2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 4$$

For this reason, $\lambda_0 = \frac{8}{4} = 2$, and $x^{(1)}$ is,

$$x^{(1)} = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix}$$

In the next iteration, we calculate $x^{(2)}$. To accomplish that, we calculate $g_1 = \begin{pmatrix} \frac{3}{2} \times 8 - 4 \times 1 \\ -12 + 8 - 4 \times 1.5 \end{pmatrix} = \begin{pmatrix} 8 \\ -10 \end{pmatrix}$

and by given in the problem, we know $d_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, I follow the same routine to calculate the λ_1

$$-\langle g_1, d_1 \rangle = -(8, -10) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 12$$

$$\langle Ad_0, d_0 \rangle = (1, 2) \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (0, 6) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 12$$

For this reason, $\lambda_1 = 1$ and we can evaluate $x^{(2)}$ as

$$x^{(2)} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \\ 3 \end{pmatrix}$$

2.3 Part 3

The visualization has been done in python via matplotlib. Please refer to the source code to see details. Here is the plot.