Stochastic Optimisation Algorithms - Home Problem 1

Name: Peizheng Yang ID: 19960728-1418

September 24, 2019

Problem 1.1

The function $f(x_1, x_2)$ is to be minimized under the constraint $g(x_1, x_2) \leq 0$.

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2 \tag{1}$$
$$g(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0 \tag{2}$$

The following steps solve this problem.

1. Define $f_{\mathrm{p}}(\mathbf{x}; \mu)$, where $\mathbf{x} \in \mathbb{R}^2$, $\mu > 0$.

$$f_{\mathrm{p}}(\mathbf{x}; \mu) = f(x_1, x_2) + \mu(\max(g(x_1, x_2), 0))^2 \tag{3}$$

It can be written in the form of piecewise function.

$$f_{\mathrm{p}}(\mathbf{x}; \mu) = \begin{cases} (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu(x_1^2 + x_2^2 - 1)^2 & x_1^2 + x_2^2 > 1 \\ (x_1 - 1)^2 + 2(x_2 - 2)^2 & \text{otherwise} \end{cases} \tag{4}$$

2. Step 2: Analytically compute the gradient $\nabla f_{\mathrm{p}}(\mathbf{x}; \mu)$.

$$\nabla f_{\mathrm{p}}(\mathbf{x}; \mu) = (\frac{\partial f_{\mathrm{p}}}{\partial x_1}, \frac{\partial f_{\mathrm{p}}}{\partial x_2})^{\mathrm{T}} \tag{5}$$

If $x_1^2 + x_2^2 > 1$ (the constraint does not hold),

$$\frac{\partial f_{\mathrm{p}}}{\partial x_1} = 2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1) \tag{6}$$
$$\frac{\partial f_{\mathrm{p}}}{\partial x_2} = 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1) \tag{7}$$

otherwise (the constraint holds),

$$\frac{\partial f_{\mathrm{p}}}{\partial x_1} = 2(x_1 - 1) \tag{8}$$
$$\frac{\partial f_{\mathrm{p}}}{\partial x_2} = 4(x_2 - 2) \tag{9}$$

3. Step 3: Find and report the unconstrained minimum of $f(x_1, x_2)$.
   The Hessian matrix of $f(x_1, x_2)$, $\forall x_1, x_2 \in \mathbb{R}$ is positive-definite. Thus $f(x_1, x_2)$ is convex.

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}. \tag{10}$$

   The unconstrained global minimum is therefore the only local minimum $(1,\ 2)^{\mathrm{T}}$ where $\nabla f(1, 2) = (0,\ 0)^{\mathrm{T}}$.

4. Step 4: Write MATLAB code. See *.m files. Note that in the script "RunPenaltyMethod.m". The step length may vary according to μ, i.e. larger μ requires smaller step length to make the newton iterations converge.

5. Step 5: Run the program.
   The series of μ and corresponding η is in Table 1. For $\mu \geq 10^5$, the corresponding η decreases as μ increases to make x* converge.

μ	η
1	10^{-4}
10	10^{-4}
10^2	10^{-4}
10^3	10^{-4}
10^4	10^{-5}
10^5	10^{-6}
10^6	10^{-7}

Table 1: The series of μ and corresponding η

The table printed out in MATLAB program is shown in Table 2.

μ	x_1^*	x_2^*
1	0.434	1.210
10	0.331	0.996
10^2	0.314	0.956
10^3	0.312	0.951
10^4	0.312	0.950
10^5	0.312	0.950
10^6	0.312	0.950

Table 2: Different values of μ and the corresponding values of x_1^*, x_2^*

According to the table, when $\mu \to \infty$, $x_1^* \to 0.312$ and $x_2^* \to 0.950$. The sequence of (x_1^*, x_2^*) appears to be converge. Thus (x_1^*, x_2^*) is a reasonable result.

Problem 1.2

1.2a

1. Find the stationary points in the interior of S.

$$\nabla f(x_1, x_2) = (8x_1 - x_2,\ -x_1 + 8x_2 - 6)^{\mathrm{T}} \tag{11}$$

   Solving $\nabla f(x_1, x_2) = 0$ yields $p_1 = (\frac{2}{21},\ \frac{16}{21})^{\mathrm{T}}$.

2. Find the stationay points or critical points on the boundary.

   - $0 < x_2 < 1,\ x_1 = 0$

$$f(0, x_2) = 4x_2^2 - 6x_2 \tag{12}$$

$$\frac{\partial f}{\partial x_2} = 8x_2 - 6 \tag{13}$$

     Solving $\frac{\partial f}{\partial x_2} = 0$ yields $p_2 = (0,\ \frac{3}{4})^{\mathrm{T}}$.

   - $0 < x_1 < 1,\ x_2 = 1$

$$f(x_1, 1) = 4x_1^2 - x_1 - 2 \tag{14}$$

$$\frac{\partial f}{\partial x_2} = 8x_1 - 1 \tag{15}$$

     Solving $\frac{\partial f}{\partial x_1} = 0$ yields $p_3 = (\frac{1}{8},\ 1)^{\mathrm{T}}$.

   - $x_1 = x_2 = x,\ 0 < x < 1$

$$f(x) = 7x^2 - 6x \tag{16}$$

$$\frac{\partial f}{\partial x} = 14x - 6 \tag{17}$$

     Solving $\frac{\partial f}{\partial x} = 0$ yields $p_4 = (\frac{3}{7},\ \frac{3}{7})^{\mathrm{T}}$.

   - Points at the corner of the boundary.
     $p_5 = (0,\ 0)^{\mathrm{T}}$, $p_6 = (0,\ 1)^{\mathrm{T}}$, $p_7 = (1,\ 1)^{\mathrm{T}}$.

3. Investigate $f(p_i)$ for $i = 1, ..., 7$.

$$f(p_1) = -2.286 \tag{18}$$

$$f(p_2) = -2.250 \tag{19}$$

$$f(p_3) = -2.063 \tag{20}$$

$$f(p_4) = -1.286 \tag{21}$$

$$f(p_5) = 0 \tag{22}$$

$$f(p_6) = -2 \tag{23}$$

$$f(p_7) = 1 \tag{24}$$

   The global minimum is at $p_1 = (\frac{2}{21},\ \frac{16}{21})^{\mathrm{T}}$ and the corresponding value is $f(p_1) = -2.286$.

1.2b

The Lagrange function is

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2). \tag{25}$$

Solving the following equations yields points $p_1 = (1, 4)^{\mathrm{T}}$ and $p_2 = (-1, -4)^{\mathrm{T}}$.

$$\frac{\partial L}{x_1} = 2 + \lambda(2x_1 + x_2) = 0 \tag{26}$$

$$\frac{\partial L}{x_2} = 3 + \lambda(x_1 + 2x_2) = 0 \tag{27}$$

$$\frac{\partial L}{\lambda} = x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \tag{28}$$

Then investigate each p_i, $i = 1, 2$.

$$f(p_1) = 29 \tag{29}$$

$$f(p_2) = 1 \tag{30}$$

The minimum is at $p_2 = (-1, -4)^{\mathrm{T}}$ with $f(p_2) = 1$.

Problem 1.3

1.3a

See MATLAB functions in the report. The minimum found is $\mathbf{x}^* = (0, -1)^{\mathrm{T}}$. $g(\mathbf{x}^*) = 3$.

1.3b

The median fitness value table is in Table 3.

mutation rate	median fitness value
0.00	0.115
0.02	0.333
0.05	0.332
0.10	0.317

Table 3: Median fitness value for different mutation rate

The conclusion is that when the mutation rate is set as $\frac{1}{m}$ (where m denotes the length of chromosome, in this case 0.02), the performance of GA is better than other mutation rate.

1.3c

The found best point is $\mathbf{x}^* = (0, -1)^{\mathrm{T}}$. Let $g(x_1, x_2) = g_1(x_1, x_2)\, g_2(x_1, x_2)$, where

$$g_1(x_1, x_2) = 1 + (x_1 + x_2 + 1)^2\,(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) \tag{31}$$

$$g_2(x_1, x_2) = 30 + (2x_1 - 3x_2)^2\,(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2). \tag{32}$$

The gradient of $g(x_1, x_2)$ is

$$\nabla g(\mathbf{x}) = \left(\frac{\partial g_1}{\partial x_1} g_2 + \frac{\partial g_2}{\partial x_1} g_1,\ \frac{\partial g_1}{\partial x_2} g_2 + \frac{\partial g_2}{\partial x_2} g_1\right)^{\mathrm{T}}. \tag{33}$$

The partial derivatives of g_1, g_2 is

$$\frac{\partial g_1}{\partial x_1} = 2(x_1 + x_2 + 1)(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) + (x_1 + x_2 + 1)^2(-14 + 6x_1 + 6x_2) \tag{34}$$

$$\frac{\partial g_1}{\partial x_2} = 2(x_1 + x_2 + 1)(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) + (x_1 + x_2 + 1)^2(-14 + 6x_1 + 6x_2) \tag{35}$$

$$\frac{\partial g_2}{\partial x_1} = 4(2x_1 - 3x_2)(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2) + (2x_1 - 3x_2)^2(-32 + 24x_1 - 36x_2) \tag{36}$$

$$\frac{\partial g_2}{\partial x_2} = -6(2x_1 - 3x_2)(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2) + (2x_1 - 3x_2)^2(48 - 36x_1 + 54x_2) \tag{37}$$

The calculation of $\nabla g(\mathbf{x}^*)$ is shown below.

$$\left.\frac{\partial g_1}{\partial x_1}\right|_{x=\mathbf{x}^*} = 0 + 0 = 0 \tag{38}$$

$$\left.\frac{\partial g_1}{\partial x_2}\right|_{x=\mathbf{x}^*} = 0 + 0 = 0 \tag{39}$$

$$\left.\frac{\partial g_2}{\partial x_1}\right|_{x=\mathbf{x}^*} = 4 \cdot 3 \cdot (-3) + 3^2 \cdot 4 = 0 \tag{40}$$

$$\left.\frac{\partial g_2}{\partial x_2}\right|_{x=\mathbf{x}^*} = -6 \cdot 3 \cdot (-3) + 3^2 \cdot (-6) = 0 \tag{41}$$

It shows that $\nabla g(\mathbf{x}^*) = 0$, thus $\mathbf{x}^*$ is a stationary point.