Stochastic optimization algorithms 2018 Home problem 1

Martin SOGNO - Personal ID Number : 19970324-T233

September 2018

Problem 1.1, 3p, Penalty method

Question 1

The sum of f and the penalty term is equal to :

$$f_p(\mathbf{x}, \mu) = f_p(x_1, x_2, \mu) = (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu * max(0, x_1^2 + x_2^2 - 1)^2$$

Question 2

We define the Heaviside function :

$$H(x) = \begin{cases} 0 & if \quad x < 0 \\ 1 & if \quad x \geq 0 \end{cases}$$

We have

$$\nabla f_p(\mathbf{x}, \mu) = \frac{\partial f_p}{\partial x_1}, \frac{\partial f_p}{\partial x_2}, \frac{\partial f_p}{\partial \mu}$$

with :

$$\frac{\partial f_p}{\partial x_1} = 2(x_1 - 1) + 4 * \mu * H(x_1^2 + x_2^2 - 1) * (x_1^2 + x_2^2 - 1) * x_1$$

$$\frac{\partial f_p}{\partial x_2} = 4(x_2 - 2) + 4 * \mu * H(x_1^2 + x_2^2 - 1) * (x_1^2 + x_2^2 - 1) * x_2$$

$$\frac{\partial f_p}{\partial \mu} = \max(0, x_1^2 + x_2^2 - 1)^2$$

Question 3

To get the minimum of the unconstrained function, we to set the partial derivatives to 0.

$$\frac{\partial f_p}{\partial x_1} = 2(x_1 - 1) = 0 \Leftrightarrow x_1 = 1$$

$$\frac{\partial f_p}{\partial x_2} = 4(x_2 - 2) = 0 \Leftrightarrow x_2 = 2$$

So the minimum of the unconstrained function is $(1, 2)$.

Question 4

See Matlab files. The file to execute is PenaltyMethod.m

Question 5

I chose the values advised, so $\eta = 0.0001$, $T = 10^{-6}$, and $\mu = 1, 10, 100, 1000$. Here are the results :

μ	x_1^*	x_2^*
1	0.434	1.210
10	0.331	0.996
100	0.314	0.955
1000	0.312	0.951

The sequence of points seems to converge. It is only a supposition, but we can guess that $x_1 = \sqrt{0.1}$. Then we would have $x_2 = \sqrt{0.9}$ if we suppose that the result is on the border (which seems likely according to the previous points).

Problem 1.2, 3p, Constrained optimization

a) : analytical method

First, we compute the gradient of f. We have

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}$$

with

$$\frac{\partial f}{\partial x_1} = 8 * x_1 - x_2$$

$$\frac{\partial f}{\partial x_2} = -x_1 + 8 * x_2 - 6$$

The interior of s

Let's search a stationary point in the interior of s by setting the derivatives to 0.

$$\frac{\partial f}{\partial x_1} = 8 * x_1 - x_2 = 0 \Leftrightarrow x_1 = \frac{x_2}{8}$$

$$\frac{\partial f}{\partial x_2} = -x_1 + 8 * x_2 - 6 = 0 \Leftrightarrow x_2 = \frac{x_1 + 6}{8}$$

$$\left\{ \begin{array}{r} x_1 = \frac{x_2}{8} \\ x_2 = \frac{x_1+6}{8} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} 8 * x_1 = x_2 \\ 8 * x_1 = \frac{x_1+6}{8} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} 8 * x_1 = x_2 \\ \frac{63}{8} x_1 = \frac{3}{4} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} x_2 = \frac{16}{21} \\ x_1 = \frac{2}{21} \end{array} \right.$$

So $P_1 = (\frac{2}{21}, \frac{16}{21})$.

We have $0 < x1 = \frac{2}{21} < 1$, $0 < x2 = \frac{16}{21} < 1$, and $x1 = \frac{2}{21} < x2 = \frac{16}{21} < 1$, so P_1 is in S.

The borders of s

Let's search a stationary point in the interior of s by setting the derivatives to 0.

From (0,0) to (0,1), we have $x_1 = 0$. We only have to set $\frac{\partial f}{\partial x_2}$ to 0. We have

$$\frac{\partial f}{\partial x_2} = 8 * x_2 - 6 = 0 \Leftrightarrow x_2 = \frac{3}{4}$$

So $P_2 = (0, \frac{3}{4})$.

From (0,1) to (1,1), we have $x_2 = 1$. We only have to set $\frac{\partial f}{\partial x_1}$ to 0. We have

$$\frac{\partial f}{\partial x_1} = 8 * x_1 - 1 = 0 \Leftrightarrow x_1 = \frac{1}{8}$$

So $P_3 = (\frac{1}{8}, 1)$.

From (1,1) to (0,0), we have $x_1 = x_2$. We can express f only in function of this common x.

$$f(x) = 4 * x^2 - x^2 + 4 * x^2 - 6 * x = 7 * x^2 - 6 * x$$

$$f'(x) = 14 * x - 6 = 0 \Leftrightarrow x = \frac{3}{7}$$

So $P_4 = (\frac{3}{7}, \frac{3}{7})$.

The corners of s

Let's investigate the corner points of s. We have $P_5 = (0, 0)$, $P_6 = (0, 1)$, and $P_7 = (1, 1)$

Actual minimum

We have to check the values of the 7 points to know the actual minimum:

- $f(P_1) = f(\frac{2}{21}, \frac{16}{21}) = -\frac{16}{7}$
- $f(P_2) = f(0, \frac{3}{4}) = -\frac{9}{4}$
- $f(P_3) = f(\frac{1}{8}, 1) = -\frac{33}{16}$
- $f(P_4) = f(\frac{3}{7}, \frac{3}{7}) = -\frac{9}{7}$
- $f(P_5) = f(0, 0) = 0$
- $f(P_6) = f(0, 1) = -2$
- $f(P_7) = f(1, 1) = 1$

Finally, the global minimum is $P_1 = (\frac{2}{21}, \frac{16}{21})$.

b) : Lagrange Multiplier

Let's define

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2)$$

$$L(x_1, x_2, \lambda) = 2x_1 + 3x_2 + 15 + \lambda x_1^2 + \lambda x_1 x_2 + \lambda x_2^2 - 21\lambda$$

Setting the gradient of L equal to zero, we get :

$$\frac{\partial L}{\partial x_1} = 2\lambda x_1 + \lambda x_2 + 2 = 0$$

$$\frac{\partial L}{\partial x_2} = \lambda x_1 + 2\lambda x_2 + 3 = 0$$

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

$$\left\{\begin{array}{r} 2\lambda x_1 + \lambda x_2 + 2 = 0 \\ \lambda x_1 + 2\lambda x_2 + 3 = 0 \\ x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{r} -3\lambda x_2 = 4 \\ 3\lambda x_1 = -1 \\ x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{r} \lambda x_2 = -\frac{4}{3} \\ \lambda x_1 = -\frac{1}{3} \\ \lambda^2(x_1^2 + x_1 x_2 + x_2^2) = 21\lambda^2 \end{array}\right. \Leftrightarrow$$

$$\left\{\begin{array}{r} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ \frac{16}{9} + \frac{4}{9} + \frac{1}{9} = 21\lambda^2 \end{array}\right. \Leftrightarrow \left\{\begin{array}{r} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ \frac{21}{9} = 21\lambda^2 \end{array}\right. \Leftrightarrow \left\{\begin{array}{r} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ \lambda^2 = \frac{1}{9} \end{array}\right. \Leftrightarrow \left\{\begin{array}{r} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ \lambda = \frac{1}{3} \quad or \quad \lambda = -\frac{1}{3} \end{array}\right.$$

We have two solutions:

- $\lambda = \frac{1}{3}, x_1 = -1, x_2 = -4, P_1 = (-1, -4) \quad f(P_1) = 1$
- $\lambda = -\frac{1}{3}, x_1 = 1, x_2 = 4, P_2 = (1, 4) \quad f(P_2) = 29$

So the minimum of f is $P_1 = (-1, -4)$.

Problem 1.3, 4p, Basic GA program (Mandatory)

a)

See Matlab files. The file to execute is FunctionOptimizationA.m.
The algorithm gives $(x_1^*, x_2^*) = (-0.00045151, -1.0001)$.

We may suppose that $(x_1^*, x_2^*) = (0, -1)$, and we will try to prove it in section c.

b)

See Matlab files. The file to execute is FunctionOptimizationB.m. The algorithm gives :

Mutation rate	0	0.02	0.05	0.1
Median fitness	0.045956	0.33332	0.3324	0.3197

The best mutation rate is 0.02, which is equal to the inverse of the number of genes. It is coherent with what we calculated in class. We can also notice that the worst median is got when the mutation rate is zero. Is it because we cannot create new information by crossover.

c)

The Genetic Algorithms leads to $x_1 = 0, x_2 = -1$. Let's prove that $(0, -1)$ is a stationary point.

Let's set :
$u = 1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)$,
$v = 30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)$

$$\nabla g(\mathbf{x}) = \frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}$$

with

$$\frac{\partial g}{\partial x_1} = \frac{\partial u}{\partial x_1}v + u\frac{\partial v}{\partial x_1}$$

$$\frac{\partial g}{\partial x_2} = \frac{\partial u}{\partial x_2}v + u\frac{\partial v}{\partial x_2}$$

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

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

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

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

$$u(0,-1) = 1 + 0^2(19 + 14 + 3) = 1$$

$$v(0,1) = 30 + 3^2 * (18 - 48 + 27) = 30 + 9 * -3 = 3$$

$$\frac{\partial u}{\partial x_1}(0,-1) = 0 * (-14-6) + 2 * 0 * (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) = 0$$

$$\frac{\partial v}{\partial x_1}(0,-1) = 3^2(-32+36) + 4 * 3 * (18 - 48 + 27) = 9 * -4 + 12 * (-3) = 0$$

$$\frac{\partial u}{\partial x_2}(0,-1) = 0 * (-14-6) + 2 * 0 * (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) = 0$$

$$\frac{\partial v}{\partial x_2}(0,-1) = 3^2(48-54) - 6 * 3 * (18 - 48 + 27) = 9 * (-6) - 18 * -3 = 0$$

$$\frac{\partial g}{\partial x_1} = \frac{\partial u}{\partial x_1}v + u\frac{\partial v}{\partial x_1} = 0 * 3 + 1 * 0 = 0$$

$$\frac{\partial g}{\partial x_2} = \frac{\partial u}{\partial x_2}v + u\frac{\partial v}{\partial x_2} = 0 * 3 + 1 * 0 = 0$$

Finally, we have $\nabla g(0,-1) = (0,0)$, so (0,-1) is effectively a stationary point.