SO - Homework 1

Thibault Desjonquères 19990602-8890

September 20, 2022

1 Penalty method

In this problem, we shall use the penalty method to find the minimum of the function f (1), subject to the constraint g (2).

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

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

1.1

As only one constraint is applied to f, the penalty function takes the form :

$$p(\vec{x}, \mu) = \mu \sum_{1=1}^{m} max[(g_i(\vec{x}), 0)]^2 \tag{3}$$

Where $\mu > 0$,and $p(\vec{x}, \mu) = 0$ if the constraint is satisfied. The problem consists of minimizing equation 4 with respect to $\vec{x}$.

$$f_p(\vec{x}, \mu) = f(\vec{x}) + p(\vec{x}, \mu) \tag{4}$$

Which gives :

$$f_p(x_1, x_2) = \{ \begin{array}{ll} (x_1 - 1)^2 + 2 \cdot (x_2 - 2)^2 + \mu \cdot (x_1^2 + x_2^2 - 1) & \text{if } x \leq 0 \\ (x_1 - 1)^2 + 2 \cdot (x_2 - 2)^2 & \text{otherwise.} \end{array} \tag{5}$$

1.2

Taking the partial derivative of f_p with respect to x_1 and x_2, with $\mu = 0$ (equations 6,7) and $\mu = 0$ (equations 8,9) gives :

$$\frac{\partial f_p}{\partial x_1} = \frac{\partial}{\partial x_1}(\underbrace{(x_1 - 1)^2}_{=2 \cdot 1 \cdot (x_1 - 1)} + \underbrace{2 \cdot (x_2 - 2)^2}_{=0} + \mu \underbrace{(x_1^2 + x_2^2 - 1)^2}_{=2 \cdot 2 \cdot x_1 \cdot (x_1^2 + x_2^2 - 1)^2}) = 2(x_1 - 1) + 4\mu x_1 (x_1^2 + x_2^2 - 1) \tag{6}$$

$$\frac{\partial f_p}{\partial x_2} = \frac{\partial}{\partial x_2}(\underbrace{(x_1 - 1)^2}_{=0} + 2 \cdot \underbrace{(x_2 - 2)^2}_{=2 \cdot 1 \cdot (x_2 - 2)} + \mu \underbrace{(x_1^2 + x_2^2 - 1)^2)}_{=2 \cdot 2 \cdot x_2 \cdot (x_1^2 + x_2^2 - 1)^2} = 4(x_2 - 2) + 4\mu x_2 (x_1^2 + x_2^2 - 1) \tag{7}$$

$$\frac{\partial f_p}{\partial x_1} = \frac{\partial}{\partial x_1}(\underbrace{(x_1 - 1)^2}_{2 \cdot 1 \cdot (x_1 - 1)} + 2 \cdot \underbrace{(x_2 - 2)^2}_{=0} = 2(x_1 - 1) \tag{8}$$

$$\frac{\partial f_p}{\partial x_2} = \frac{\partial}{\partial x_2}(\underbrace{(x_2 - 2)^2}_{=0} + 2 \cdot \underbrace{(x_2 - 2)^2}_{=2 \cdot 1 \cdot (x_2 - 2)} = 4(x_2 - 2) \tag{9}$$

1.3

The unconstrained minimum of f_p ($\mu = 0$) is found by calculating $\nabla f_p(\vec{x}) = 0$ (equations (10),(11))

$$\frac{\partial f}{\partial x_1} = 2(x_1 - 1) = 0 \longleftrightarrow x_1 = 1 \tag{10}$$

$$\frac{\partial f}{\partial x_2} = 4(x_2 - 2) = 0 \longleftrightarrow x_2 = 2 \tag{11}$$

The chosen starting point is $P_0 = (1, 2)$.

X1	**X2**	**muValues**	**Eta**	**Tolerance**
0.4338	1.2102	1	0.0001	1E-4
0.3314	0.9956	10	0.0001	1E-4
0.3138	0.9552	100	0.0001	1E-4
0.3118	0.9507	1000	0.0001	1E-4

Table 1: Table of results from the first run.

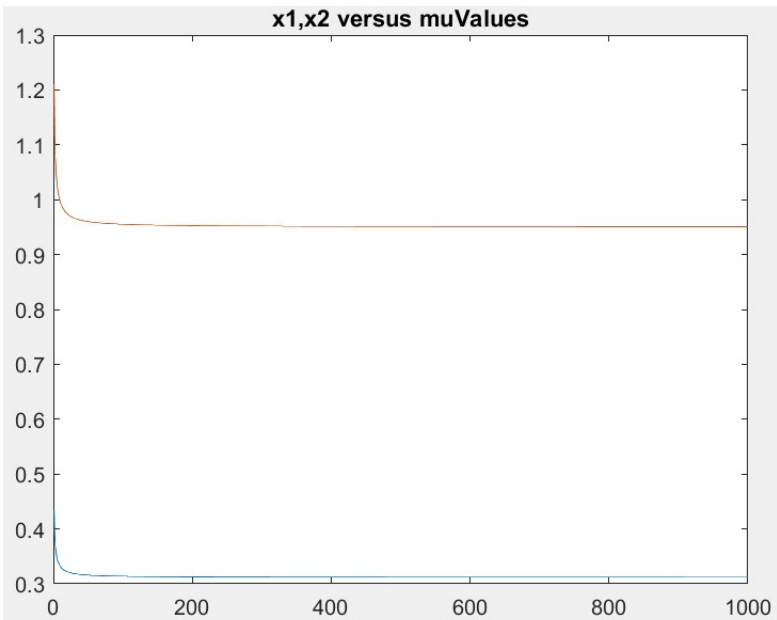

Figure 1: x_1 (blue),x_2 (red) VS parameter mu.

2

We consider the function f in equation (12), and the set shown in 2.

$$f(x_1, x_2) = 4x_1^2 - x_1x_2 + 4x_2^2 - 6x_2 \tag{12}$$

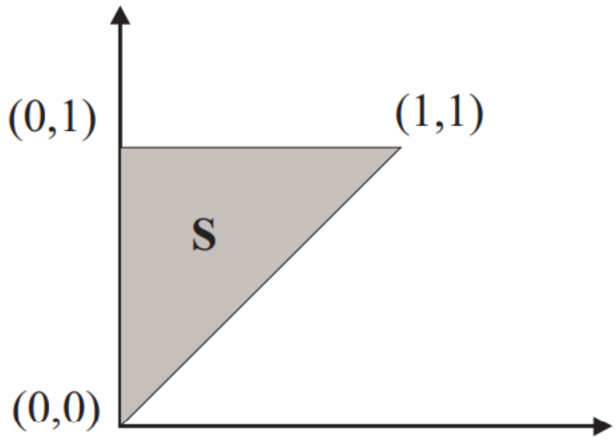

Figure 2: Set considered in task 2.

2.1

We first find the stationnary point of f in S by computing $\nabla f(\vec{x}) = 0$.

$$\frac{\partial f}{\partial x_1} = \frac{\partial}{\partial x_1}(\underbrace{4x_1^2}_{=4\cdot 2\cdot x_1} - \underbrace{x_1x_2}_{=x_2} + \underbrace{4x_2^2}_{=0} - \underbrace{6x_2}_{=0}) = 8x_1 - x_2 = 0 \leftrightarrow 8x_1 = x_2 \leftrightarrow x_1 = \frac{2}{21} \tag{13}$$

$$\frac{\partial f}{\partial x_2} = \frac{\partial}{\partial x_2}(\underbrace{4x_1^2}_{=0} - \underbrace{x_1x_2}_{=x_1} + \underbrace{4x_2^2}_{=4\cdot 2x_2} - \underbrace{6x_2}_{=6}) = -x_1 + 8x_2 - 6 = 0 \leftrightarrow x_1 = 8x_2 - 6 \leftrightarrow x_2 = \frac{16}{21} \tag{14}$$

A potential global minimum is found at $P_0=(\frac{2}{21}, \frac{16}{21})$, which lies in S. $f(\frac{2}{21}, \frac{16}{21})$ = -2.29.

Second, we look for the presence of other potential global minima on the boundary of S, which can be divided in 3 sections ;

a) $0 < x_2 < 1, x_1 = 0$

b) $0 < x_1 < 1, x_2 = 1$

c) $x_1 = x_2$

We compute $\nabla f(\vec{x}) = 0$, where $\vec{x}$ respects the conditions mentioned above.

a)

$$\frac{\partial f(0,x_2)}{\partial x_2} = \underbrace{\frac{\partial}{\partial x_2}(4x_2^2 - 6x_2)}_{=8x_2-6} = 0 \leftrightarrow x_2 = 3/4, \quad f(0,3/4) = 4\cdot(3/4)^2 - 6\cdot(3/4) = 9/4 - 18/4 = -9/4 = -2.25 \tag{15}$$

b)

$$\frac{\partial f(x_1,1)}{\partial x_1} = \underbrace{\frac{\partial}{\partial x_1}(4x_1^2 - x_1 + 4 - 6)}_{=8x_1-1} = 0 \leftrightarrow x_1 = 1/8, f(1/8,1) = 4\cdot(1/8)^2 - (1/8) + 4 - 6 = -2.06 \tag{16}$$

c)

$$\frac{\partial f(x_1,x_1)}{\partial x_1} = \underbrace{\frac{\partial}{\partial x_1}(8x_1 - 2x_1 + 8x_1 - 6)}_{=14x_1-6} = 0 \leftrightarrow x_1 = x_2 = 3/7,$$

f(3/7,3/7) $= 4 \cdot (3/7)^2 - (3/7)^2 + 4 \cdot (3/7)^2 - 6 \cdot 3/7 = -1.29$

(17)

On the boundary of S, there are three potential global minima : P_a=(0,3/4), P_b=(1/8,1), P_c=(3/7,3/7). In order to find the minima at the edges of S, we calculate f(0,0)=0, f(1,1)=1, f(0,1)=-2.
The global minimum is found at P_0=($\frac{6}{63}$, $\frac{48}{63}$), where f($\frac{6}{63}$, $\frac{48}{63}$) = -2.29.

2.2

We look for the global minimum of f($\vec{x}$) (equation 18), subjected to the constraint function h($\vec{x}$) = 0 (equation 19). We do so by using the Lagrange mulitplier L (equation 20)

$$f(x_1, x_2) = 15 + 2 \cdot x_1 + 3 \cdot x_2 \tag{18}$$

$$h(x_1, x_2) = x_1^2 + x_1 \cdot x_2 + x_2^2 - 21 = 0 \tag{19}$$

$$L(x_1, x_2) = f(x_1, x_2) + \lambda \cdot h(x_1, x_2) \tag{20}$$

First, we calculate $\nabla L(\vec{x}) = 0$

$$\frac{\partial L}{\partial x_1} = \frac{\partial}{\partial x_1}(\underbrace{15}_{=0} + \underbrace{2 \cdot x_1}_{=2} + \underbrace{3 \cdot x_2}_{=0}) + \lambda \cdot \frac{\partial}{\partial x_1}(\underbrace{x_1^2}_{=2x_1} + \underbrace{x_1 \cdot x_2}_{=x_2} + \underbrace{x_2^2}_{=0} - \underbrace{21}_{=0}) = 0 \leftrightarrow x_2 = \frac{-2}{\lambda} - 2x_1 \tag{21}$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial}{\partial x_2}(\underbrace{15}_{=0} + \underbrace{2 \cdot x_1}_{=0} + \underbrace{3 \cdot x_2}_{=3}) + \lambda \cdot \frac{\partial}{\partial x_2}(\underbrace{x_1^2}_{=0} + \underbrace{x_1 \cdot x_2}_{=x_1} + \underbrace{x_2^2}_{=2x_2} - \underbrace{21}_{=0}) = 3 + \lambda x_1 + 2\lambda x_2 = 3 + \lambda x_1 + 2\lambda(\frac{-2}{\lambda} - 2x_1) = 0$$

$\leftrightarrow x_1 = \frac{-1}{3\lambda}$(22)

Finally, $x_1 = \pm\frac{-1}{3\lambda}$ and $x_2 = \pm\frac{-4}{3\lambda}$. The constraint h$\vec{x} = 0$ is with both combination.

$$\frac{\partial L}{\partial \lambda} = \frac{\partial}{\partial \lambda}(\underbrace{f(\vec{x})}_{=0}) + \frac{\partial}{\partial \lambda}\underbrace{\lambda}_{=1} \cdot (x_1^2 + x_1 \cdot x_2 + x_2^2 - 21) = (x_1^2 + x_1 \cdot x_2 + x_2^2 - 21) = 0 \tag{23}$$

By plugging x_1 and x_2 :

$$\frac{\partial L}{\partial \lambda} = (\frac{-1}{3\lambda})^2 + \frac{-1}{3\lambda} \cdot \frac{-4}{3\lambda} + (\frac{-4}{3\lambda})^2 - 21 = (\frac{1}{9\lambda^2}) + \frac{4}{9\lambda^2} + (\frac{16}{9\lambda^2}) - 21 = \frac{21}{9\lambda^2} - 21 = 0 \leftrightarrow \lambda = \frac{1}{3} \tag{24}$$

So $x_1 = \pm 1$, $x_2 = \pm 4$, $\lambda = 1/3$.

$$f(-1, -4) = 15 + 2 \cdot (-1) + 3 \cdot (-4) = 1 \tag{25}$$

$$f(1, 4) = 15 + 2 + 3 \cdot 4 = 29 \tag{26}$$

Therefore, P=(-1,-4) is the global minimum of f.

3 Problem 1.3

3.1

Selected parameters in RunSingle.m :
populationSize = 100;
maximumVariableValue = 5;
numberOfGenes = 50;
numberOfVariables = 2;
tournamentSize = 2;
tournamentProbability = 0.1;
crossoverProbability = 0.8;
mutationProbability = 0.02;
numberOfGenerations = 2000;

x1	**x2**	**Fitness**
3.0001	0.5000	0.9999
2.9999	0.4999	0.9999
3.0000	0.5000	0.9999
3.0007	0.5001	0.9999
2.9980	0.4994	0.9999
2.9998	0.4999	0.9999
2.9995	0.4998	0.9999
2.9999	0.4999	0.9999
3.0002	0.5000	0.9999
2.9998	0.4999	0.9999

Table 2: Values of x1, x2, and their respective fitness (10 turns, Mutation Probability = 0.02).

3.2

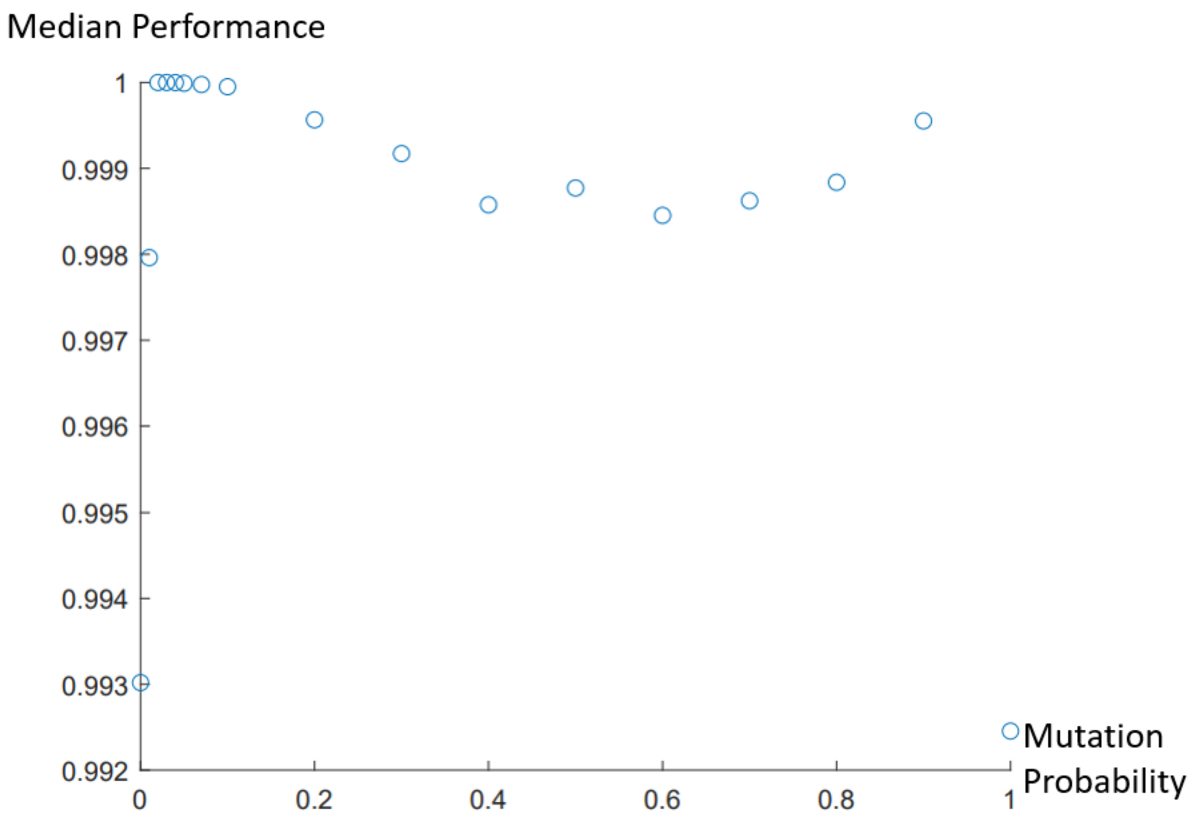

Figure 3: x-axis : Mutation Probability, y-axis : fitness value

DISCUSSION : For $\mu < 0.02$, we obtain better fitness value. This is due to the fact that at this stage, the mutations provide just enough noise to let some points explore new horizons, and potentially reach a better result. For $0.02 < \mu < 0,5$, the mutation probability becomes too big, it compromises too many points, including the ones that must not have been modified. Overall, the fitness decreases. The worst case scenario is reached when $\mu = 0.5$. At this point, we no longer introduce a bit of noise but we compromise our points radically. I do not get an idea why the code behaves the way it does when $0.5 < \mu < 1$, if I had more time I would repeat the plot with more values in this region to check whether or not the fitness values does increase. At $\mu = 1$, every ones have turned into zeros, and vice versa. Therefore, at the next iteration, I expect the fitness to increase.

Mutation Rate	**Fitness Values**
0	0.9926
0.01	0.9998
0.02	0.9999
0.03	0.9999
0.04	0.9999
0.05	0.9999
0.07	0.9999
0.1	0.9999
0.2	0.9996
0.3	0.9991
0.4	0.9988
0.5	0.9984
0.6	0.9985
0.7	0.9983
0.8	0.9989
0.9	0.9996
1	0.9895

Table 3: Chosen values of mutation rate and their corresponding fitness. (Results obtained from a different run, my program being extremely slow).

3.3

According to the values reported in 2, $x_1 \approx 3$,$x_2 \approx \frac{1}{2}$. Plugging those values into g(x_1,x_2) must return 0, if it is a stationary point.

We want to verify : g(3,$\frac{1}{2}$) $\stackrel{?}{=}$ 0.

$$g(x_1, x_2) = (1.5 - x_1 + x_1\dot{x}_2)^2 + (2.25 - x_1 + x_1 \cdot x_2^2)^2 + (2.625 - x_1 + x_1 \cdot x_2^3)^2 \tag{27}$$

In order to compute the result, one should know : $\frac{9}{4} = 2.25$, $\frac{21}{8} = 2.625$,

$$g(3, 1/2) = (\frac{3}{2} - 3 + 3 \cdot \frac{1}{2})^2 + (2.25 - 3 + 3 \cdot \frac{1}{4})^2 + (\frac{21}{8} - 3 + 3 \cdot \frac{1}{8}))^2 \tag{28}$$

$$g(3, 1/2) = (\frac{6}{2} - \frac{6}{2})^2 + (\frac{9}{4} - \frac{9}{4})^2 + (\frac{21}{8} - \frac{21}{8}))^2 = 0$$

Therefore, the function g has a stationary point at $x_1 = 3$, $x_2 = 0.5$.