FFR 105 - Report Home Problem 1

Name : Pierre-Eliot Jourdan
Civic registration number : 20001219-T136

Problem 1.1, 3p, Penalty method (Mandatory)

In this first part of the problem, we want to solve the following problem using the penalty method :

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2 \quad (1)$$

$$subject\ to\ the\ constraint: \ g(x_1, x_2) = x_1^2 + x_2^2 - 1 \le 0. \quad (2)$$

1) *According to the penalty method, minimizing f subject to g is equivalent to minimizing* : $f_p(\boldsymbol{x}; \mu) = f(\boldsymbol{x}) + p(\boldsymbol{x}; \mu)$ *without constraints.* $p(\boldsymbol{x}; \mu)$ *is the penalty term with p the penalty function* :

$$p(\boldsymbol{x}; \mu) = \mu \times (\sum_{i=1}^{m} (max\{g_i(\boldsymbol{x}), 0\})^2), \ \boldsymbol{x} = (x_1, x_2)$$

Thus, here is the expression of $f_p(\boldsymbol{x}; \mu)$:

$$\boxed{f_p(\boldsymbol{x}; \mu) = (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu \times (max\{x_1^2 + x_2^2 - 1, 0\})^2, \quad \boldsymbol{x} = (x_1, x_2)}$$

2) *Let's now compute analytically the gradient* $\nabla f_p(\boldsymbol{x}; \mu)$:

* *1st case : the constraints are fulfilled, so* $g(x_1, x_2) = x_1^2 + x_2^2 - 1 \le 0$. *Therefore* :

$$f_p(\boldsymbol{x}; \mu) = f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2$$

$$\boxed{So\ \nabla f_p(\boldsymbol{x}; \mu) = (\frac{df_p(\boldsymbol{x};\mu)}{dx_1}, \frac{df_p(\boldsymbol{x};\mu)}{dx_2}) = (\frac{df(x_1, x_2)}{dx_1}, \frac{df(x_1, x_2)}{dx_2}) = (2(x_1 - 1); 4(x_2 - 2)).}$$

* $2nd\ case$: $the\ constraints\ are\ not\ fulfilled,\ so\ g(x_1, x_2) = x_1^2 + x_2^2 - 1 > 0.\ Therefore:$

$f_p(\boldsymbol{x}; \mu) = f(x_1, x_2) + \mu(g(x_1, x_2))^2 = (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu(x_1^2 + x_2^2 - 1)^2$

$Let's\ compute\ both\ components\ of\ \nabla f_p(\boldsymbol{x}; \mu) = (\frac{df_p(\boldsymbol{x};\mu)}{dx_1}, \frac{df_p(\boldsymbol{x};\mu)}{dx_2}):$

$$\boxed{\nabla f_p(\boldsymbol{x}; \mu) = (2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1); 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1)).}$$

3) $Let's\ find\ out\ the\ unconstrained\ minimum\ i.e.\ the\ point\ \boldsymbol{x}\ such\ that\ \nabla f_p(\boldsymbol{x}; \mu) = 0$ $and\ we\ are\ considering\ that\ \mu = 0,\ which\ leads\ us\ to\ the\ same\ expression\ as\ the\ 1st\ case:$

$$f_p(\boldsymbol{x}; \mu) = f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2.$$

$Thus\ we\ are\ looking\ for\ the\ point\ \boldsymbol{x}\ such\ that\ (\frac{df(x_1, x_2)}{dx_1}, \frac{df(x_1, x_2)}{dx_2}) = (0, 0):$

$$\frac{df(x_1, x_2)}{dx_1} = 0 \Leftrightarrow 2(x_1 - 1) = 0 \Leftrightarrow x_1 = 1$$

$$\frac{df(x_1, x_2)}{dx_2} = 0 \Leftrightarrow 4(x_2 - 2) = 0 \Leftrightarrow x_2 = 2$$

$$\boxed{Thus\ the\ unconstrained\ minimum\ of\ the\ function\ is\ \boldsymbol{x} = (1, 2).}$$

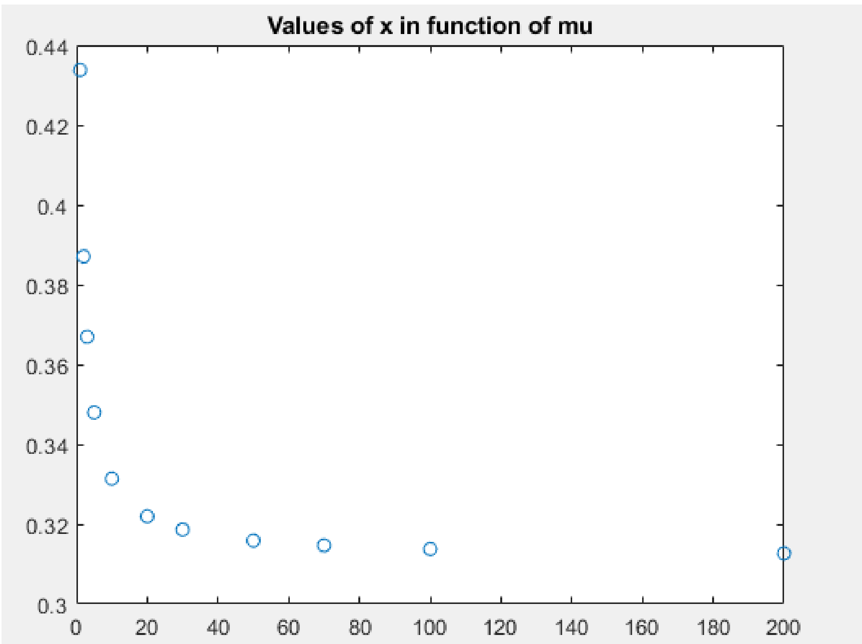
$(We\ will\ use\ this\ point\ as\ the\ starting\ point\ for\ gradient\ descent).$

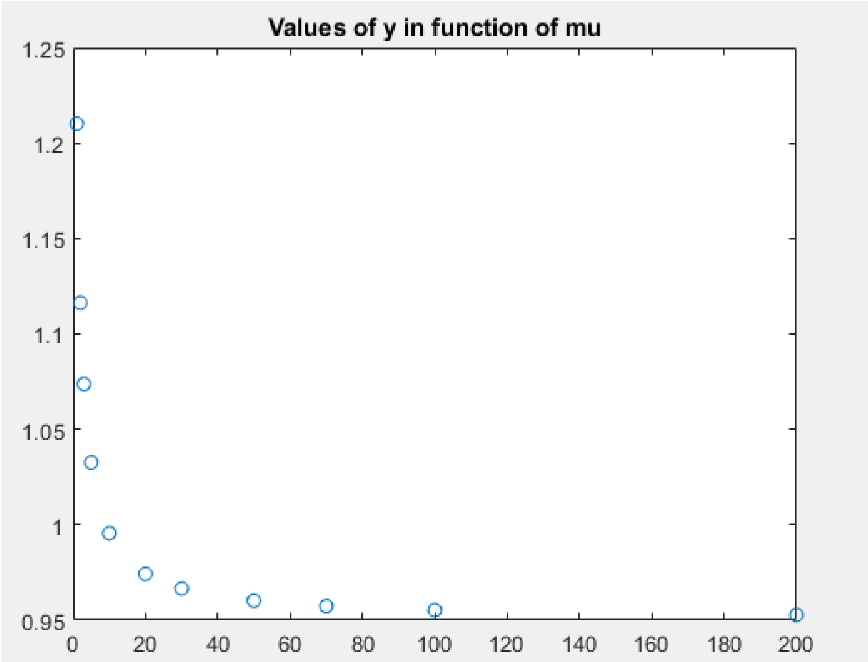
4) $See\ the\ Matlab\ scripts\ RunPenaltyMethod.m,\ ComputeGradient.m\ and\ RunGradientDescent.$

5) $Let's\ take\ a\ look\ at\ the\ results\ of\ our\ programm.\ I\ chose\ for\ the\ values\ of\ the\ parameters:\ T = 10^{-6},\ \eta = 0.0001\ (small\ value)\ and\ for\ the\ different\ values\ of\ \mu\ (see\ the\ table\ below),\ here\ is\ the\ computation\ of\ the\ values\ of\ x_1\ and\ x_2:$

$Values\ of\ \mu$	$Values\ of\ x_1$	$Values\ of\ x_2$
$0\ (Starting\ point)$	1	2
1	0.4337	1.2101
5	0.3479	1.0326
10	0.3313	0.9955
50	0.3158	0.9601
100	0.3137	0.9952
500	0.3120	0.9512
1000	0.3117	0.9507

To prove that our results are reasonable, we are going to plot the values of x_1 *and* x_2 *as function of* μ. *For more visibility, I chose not to plot the starting point (for* $\mu = 0$*) and to plot the values of* x_1 *and* x_2 *for* $\mu \in [1; 2; 3; 5; 10; 20; 30; 50; 70; 100; 200]$ *and we kept the same values for the parameters :* $T = 10^{-6}$, $\eta = 0.0001$. *Here we plot the 2 figures :*





Conclusion : *We can observe on these* 2 *figures that both sequences of values of* x_1 *and* x_2 *are convergent.* x_1 *seem to converge towards around* 0.31 *and* x_2 *towards around* 0.95. *We also notice that the lower the value of the tolerance parameter* T *is, the faster* x_1 *and* x_2 *will converge.*

Problem 1.2, 3p, Constrained optimization (Voluntary)

a)

In this part $1.2.a$ *of the problem, we want to solve the following constrained optimization problem using the analytical method seen in the course. Let's find the minimum of the function:*

$f(x_1, x_2) = 4{x_1}^2 - x_1 x_2 + 4{x_2}^2 - 6x_2$ *on the closed set* S *shown in the figure below.*

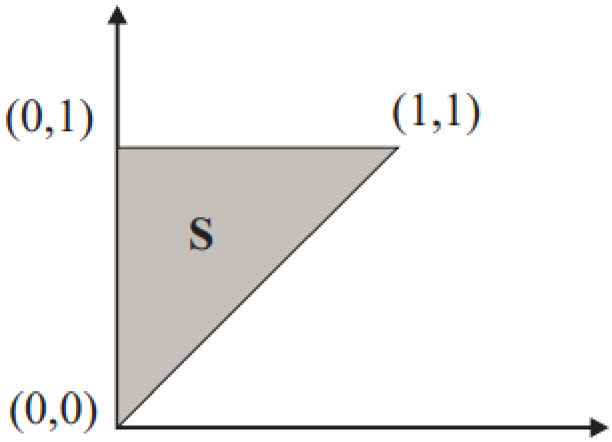


Figure 1: The set S used in Problem 1.2a.

* $1st$ *step: Find the stationary points in the interior of* S

We solve the equation $\nabla f(x_1, x_2) = 0$ *and verify that the solutions are in the interior of* S:

$$\frac{df(x_1,x_2)}{dx_1} = 0 \Leftrightarrow 8x_1 - x_2 = 0 \Leftrightarrow x_2 = 8x_1 \ (1)$$

$\frac{df(x_1,x_2)}{dx_2} = 0 \Leftrightarrow - x_1 + 8x_2 - 6 = 0$ (2) *we now use* (1) *to determine* x_2 *and then* x_1:

$$(2) \Leftrightarrow 63x_1 = 6 \Leftrightarrow x_1 = \frac{6}{63} = \frac{2}{21} \text{ and then } (1) \Leftrightarrow x_2 = \frac{16}{21}.$$

We, thus, notice that $0 < \frac{2}{21} < \frac{16}{21} < 1$ *so the point* $P_1 = (\frac{2}{21}, \frac{16}{21})^T$ *is in the interior of* S.

Conclusion: $P_1 = (\frac{2}{21}, \frac{16}{21})^T$ *is the only stationary point in the interior of* S.

* $2nd$ *step: Find the stationary points on the boundary of* S

In this case, the boundary consists of three parts, which will be considered in sequence:

1) $1st$ *part*: $x_1 = 0$ *and* $0 < x_2 < 1$. *Taking the derivative, we get*:

$$f'(0, x_2) = 8x_2 - 6. \ So\ f'(0, x_2) = 0 \Leftrightarrow 8x_2 = 6 \Leftrightarrow x_2 = \frac{6}{8} = \frac{3}{4}.$$

$Yet,\ 0 < \frac{3}{4} < 1\ so\ the\ point\ P_2 = (0, \frac{3}{4})^T\ is\ on\ the\ boundary\ of\ S.$

$2)\ 2nd\ part:\ x_2 = 1\ and\ 0 < x_1 < 1.\ Taking\ the\ derivative,\ we\ get:$

$$f'(x_1, 1) = 8x_1 - 1. \ So\ f'(x_1, 1) = 0 \Leftrightarrow 8x_1 = 1 \Leftrightarrow x_1 = \frac{1}{8}.$$

$Yet,\ 0 < \frac{1}{8} < 1\ so\ the\ point\ P_3 = (\frac{1}{8}, 1)^T\ is\ on\ the\ boundary\ of\ S.$

$3)\ 3rd\ part:\ 0 < x_1 = x_2 < 1.\ Here\ is\ the\ form\ of\ f\ in\ this\ case:$

$$f(x_1, x_1) = 4x_1^{\ 2} - x_1^{\ 2} + 4x_1^{\ 2} - 6x_1 = 7x_1^{\ 2} - 6x_1.$$

$$So\ f'(x_1, x_1) = 0 \Leftrightarrow 14x_1 - 6 = 0 \Leftrightarrow x_1 = \frac{6}{14} = \frac{3}{7}.$$

$Yet,\ 0 < \frac{3}{7} < 1\ so\ the\ point\ P_4 = (\frac{3}{7}, \frac{3}{7})^T\ is\ on\ the\ boundary\ of\ S.$

$Conclusion:\ There\ are\ 3\ stationary\ points\ on\ the\ boundary\ of\ S:$ $P_2 = (0, \frac{3}{4})^T,\ P_3 = (\frac{1}{8}, 1)^T\ and\ P_4 = (\frac{3}{7}, \frac{3}{7})^T$

$Finally,\ the\ corners\ must\ be\ considered\ as\ well,\ contributing\ the\ points:$ $P_5 = (0, 0)^T,\ P_6 = (1, 1)^T,\ and\ P_7 = (0, 1)^T$

$*\ \ 3rd\ step: Examining\ these\ seven\ points\ one\ by\ one\ (computing\ f(P_i)\ for\ i = [1, 7]):$

$Starting\ with\ the\ corners:\ f(P_5) = 0,\ f(P_6) = 4 - 1 + 4 - 6 = 1\ \ and$

$f(P_7) = 4 - 6 = -2.$

$Then:\ f(P_3) = 4 \times \frac{1}{64} - \frac{1}{8} + 4 - 6 = \frac{1}{16} - \frac{1}{8} - 2 = -2 - \frac{1}{16} = -\frac{33}{16} = -2.0625.$

$f(P_4) = 4 \times (\frac{3}{7})^2 - (\frac{3}{7})^2 + 4 \times (\frac{3}{7})^2 - 6 \times \frac{3}{7} = \frac{9}{7} - \frac{18}{7} = -\frac{9}{7} \simeq -1.29.$

$f(P_2) = 4 \times (\frac{3}{4})^2 - 6 \times \frac{3}{4} = \frac{9}{4} - \frac{9}{2} = -\frac{9}{4} = -2.25.$

$f(P_1) = 4 \times (\frac{2}{21})^2 - \frac{2}{21} \times \frac{16}{21} + 4 \times (\frac{16}{21})^2 - 6 \times \frac{16}{21} = \frac{16}{21^2} - 2 \times \frac{16}{21^2} + 4 \times \frac{16}{21^2} - 6 \times \frac{16}{21^2}$

$= -\frac{16}{7} \simeq -2.29.$

$$\boxed{\begin{array}{c} Thus,\ the\ global\ minimum\ of\ f\ on\ the\ set\ S\ occurs\ at\ the\ point\ \ P_1 = (\frac{2}{21}, \frac{16}{21})^T, \\ where\ the\ function\ takes\ the\ value\ \ -\frac{16}{7}. \end{array}}$$

b)

$In\ this\ part\ of\ the\ problem,\ we\ are\ going\ to\ use\ the\ Lagrange\ multiplier\ method\ to$ $determine\ the\ minimum\ (x_1^*, x_2^*)^T\ \ of\ the\ function:$

$f(x_1, x_2) = 15 + 2x_1 + 3x_2\ \ subject\ to\ the\ constraint\ h(x_1, x_2) = x_1^{\ 2} + x_1x_2 + x_2^{\ 2} - 21 = 0.$

$Let's\ consider\ the\ function:$

$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \times h(x_1, x_2) = 15 + 2x_1 + 3x_2 + \lambda \times (x_1^{\ 2} + x_1x_2 + x_2^{\ 2} - 21).$

$The\ local\ optima\ of\ f\ subject\ to\ the\ equality\ constraint(s)\ h(x_1, x_2) = \ 0\ can\ thus\ be\ found$ $by\ computing\ the\ stationary\ points\ of\ L.\ We\ will\ find\ both\ minima\ and\ maxima.$

$Then,\ let's\ consider\ the\ equation\ \nabla L(x_1, x_2, \lambda) = 0:$

$$\frac{dL(x_1,x_2,\lambda)}{dx_1} = 2 + \lambda \times (2x_1 + x_2) = 0\ \ (1)$$

$$\frac{dL(x_1,x_2,\lambda)}{dx_2} = 3 + \lambda \times (x_1 + 2x_2) = 0\ \ (2)$$

$$\frac{dL(x_1,x_2,\lambda)}{d\lambda} = x_1^{\ 2} + x_1x_2 + x_2^{\ 2} - 21 = 0\ \ (3)$$

$We\ notice\ with\ equations\ (1)\ and\ (2),\ that\ \lambda \neq 0\ (otherwise\ 2 = 0\ and\ 3 = 0,\ absurd\ !).\ Then:$

$$(1) \Leftrightarrow x_2 = \ -\frac{2}{\lambda} - 2x_1\ (we\ replace\ this\ expression\ of\ x_2\ in\ (2))$$

$$(2) \Leftrightarrow 3 + \lambda \times (x_1 - \frac{4}{\lambda} - 4x_1) = 0 \Leftrightarrow \lambda x_1 - 4 - 4\lambda x_1 = \ -3 \Leftrightarrow \ -3\lambda x_1 = 1 \Leftrightarrow x_1 = -\frac{1}{3\lambda}$$

$Thus\ we\ deduce\ the\ expression\ of\ x_2\ in\ function\ of\ \lambda:\ (1) \Leftrightarrow x_2 = \ -\frac{4}{3\lambda}.$

$We\ now\ can\ use\ the\ expression\ of\ x_1\ and\ \ x_2\ in\ the\ equation\ (3):$

$$(3) \Leftrightarrow (-\frac{1}{3\lambda})^2 + (-\frac{1}{3\lambda}) \times (-\frac{4}{3\lambda}) + (-\frac{4}{3\lambda})^2 = 21 \Leftrightarrow \frac{1}{9\lambda^2} + \frac{4}{9\lambda^2} + \frac{16}{9\lambda^2} = 21 \Leftrightarrow 9\lambda^2 = \frac{21}{21} = 1.$$

$$Thus,\ (3) \Leftrightarrow \lambda^2 = \frac{1}{9} \Leftrightarrow \lambda = \ \pm\frac{1}{3}.$$

$We\ can\ now\ find\ 2\ couples\ (x_1^*, x_2^*)^T\ \ solution\ of\ the\ equation\ \nabla L(x_1, x_2, \lambda) = 0:$

$$If\ \lambda = \frac{1}{3},\ x_1^* = \ -\frac{1}{3\lambda} = \ -\frac{1}{3\times\frac{1}{3}} = \ -1\ and\ \ x_2^* = \ -\frac{4}{3\lambda} = \ -\frac{4}{3\times\frac{1}{3}} = \ -4.$$

$$In\ that\ case,\ f(-1, -4) = 15 - 2 - 3 \times 4 = 1.$$

$If\ \lambda = -\frac{1}{3},\ x_1^* = -\frac{1}{3\lambda} = -\frac{1}{3\times(-\frac{1}{3})} = 1\ and\ x_2^* = -\frac{4}{3\lambda} = -\frac{4}{3\times(-\frac{1}{3})} = 4.$

$In\ that\ case,\ f(1, 4) = 15 + 2 + 3 \times 4 = 29 > f(-1, -4).$

$Thus,\ according\ to\ the\ Lagrange\ multiplier\ method,\ the\ maximum\ of\ f\ subject\ to\ h,\ occurs\ at\ the\ point\ (1, 4)^T\ and\ the\ function\ takes\ the\ value\ 29\ and\ its\ minimum\ occur\ at\ the\ point\ (-1, -4)^T and\ the\ function\ takes\ the\ value\ 1.$

$$\boxed{\begin{array}{c} The\ minimum\ of\ the\ function\ f\ subject\ to\ the\ constraint\ h\ occurs\ at\ the\ point \\ (x_1^*, x_2^*)^T = (-1, -4)^T where\ the\ function\ f\ takes\ the\ value\ 1. \end{array}}$$

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

In this third part of the problem, we want to implement a genetic algorithm (GA) for finding the minimum of the following function:

$$g(x_1, x_2) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^{\,2})^2 + (2.625 - x_1 + x_1 x_2^{\,3})^2 \quad (4)$$

a) *Let's programm the genetic algorithm on Matlab (see Matlab scripts).*

Let's select a set of parameters (see below) and run the file RunSingle.m 10 times. Then, we collect the values of x_1, x_2 *and* $g(x_1, x_2)$ *(i.e. the inverse of the fitness value). Here is the table of the results obtained:*

Set of parameters: $tournamentSize = 2$, $tournamentProbability = 0.75$, $crossoverProbability = 0.8$, $mutationProbability = 0.02$, $numberOfGenerations = 10000$.

Run n°:	*Values of* x_1	*Values of* x_2	*Values of* $g(x_1, x_2)$
1	3.9148	0.6594	5.7×10^{-2}
2	3.9148	0.6594	5.7×10^{-2}
3	3.9148	0.6594	5.7×10^{-2}
4	3.9148	0.6594	5.7×10^{-2}
5	3.9148	0.6594	5.7×10^{-2}
6	3.9148	0.6594	5.7×10^{-2}
7	3.9148	0.6594	5.7×10^{-2}
8	3.9148	0.6594	5.7×10^{-2}
9	3.9148	0.6594	5.7×10^{-2}
10	3.9148	0.6594	5.7×10^{-2}

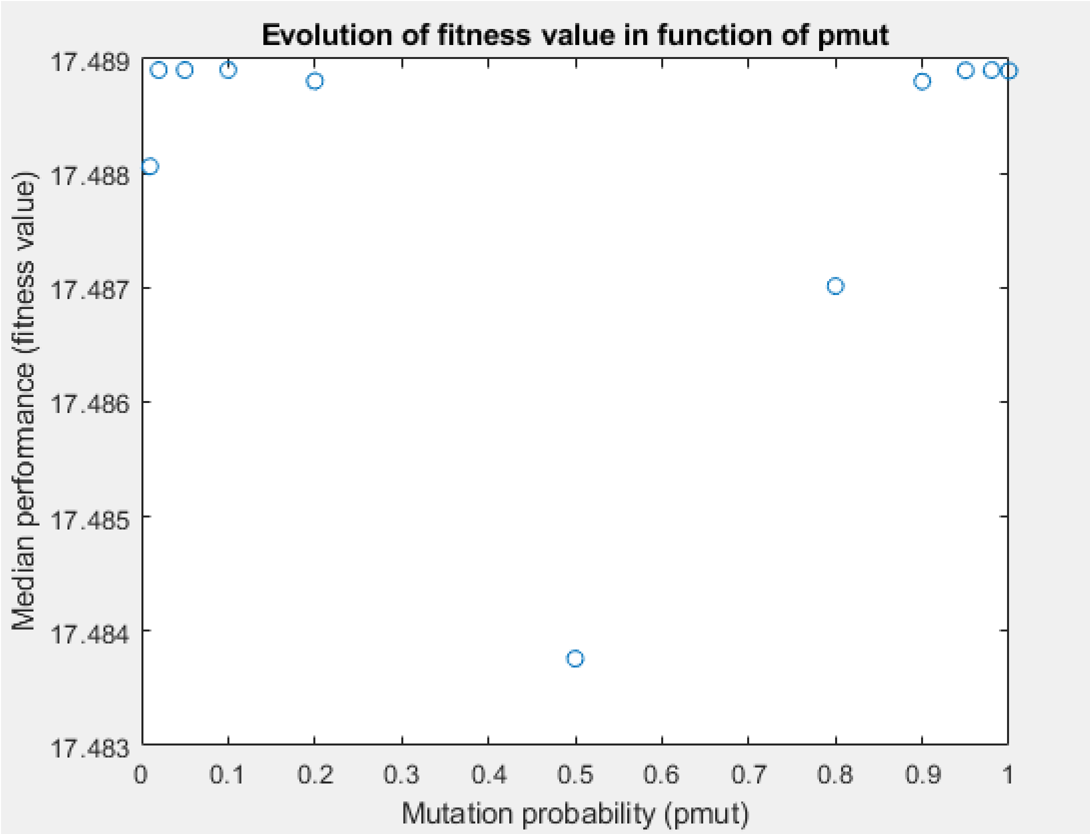
Note: The point found is a bit too far from the actual one (see question c)).

b) *Our goal, in this question, is to make a parameter search in order to find the optimal value of the mutation probability* p_{mut} *for this problem (corresponding to the highest fitness value).*

Let's collect the values of the median performance (= fitness value), over a 100 *runs for different values of* p_{mut}. *We chose* $p_{mut} = [0,\ 0.02,\ 0.1,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 0.9,\ 0.98,\ 1]$.
The other parameters remain unchanged (same as the previous question a)).

Values of p_{mut}	*Values of median performance*
0	7.97
0.01	17.4881
0.02	17.4889036034
0.05	17.4889036034
0.1	17.4889033
0.2	17.4888
0.5	17.484
0.8	17.487
0.9	17.4888
0.95	17.488899
0.98	17.488904
1	17.488895

Here is the figure of the median performance (= fitness value), over the 100 *runs, as a function of the mutation probability* p_{mut}:



$I\ chose\ not\ to\ plot\ the\ first\ fitness\ value\ (\ p_{mut}\ =\ 0)\ for\ more\ visibility.$

$By\ observing\ the\ results\ we\ obtained,\ we\ can\ say\ that\ the\ optimal\ value\ of\ p_{mut}$
$for\ this\ problem\ is\ around\ p_{mut\ optimal}\ =\ 0.02.$

c) $In\ this\ last\ part,\ we\ want\ to\ show\ that\ the\ minimum\ we\ can\ deduce\ from\ the\ runs\ of\ the\ file\ RunSingle\ is\ indeed\ a\ stationary\ point.\ The\ minimum\ found\ was\ the\ point:$

$$(3.9148, 0.6594)$$

$Thus\ we\ can\ suppose\ the\ actual\ minimum\ should\ be:\ (x_1^{\ *}, x_2^{\ *})^T\ =\ (3, 0.5)$

$Let's\ compute\ the\ gradient\ \nabla g(x_1, x_2)\ =\ (\frac{dg(x_1,x_2)}{dx_1}, \frac{dg(x_1,x_2)}{dx_2})^T\ to\ determine\ if\ the\ point$
$(x_1^{\ *}, x_2^{\ *})^T\ =\ (3, 0.5)\ is\ actually\ a\ stationary\ point\ of\ the\ function\ g:$

$$\frac{dg(x_1,x_2)}{dx_1} = 2(x_2 - 1)(1.5 - x_1 + x_1x_2) + 2(x_2^{\ 2} - 1)(2.25 - x_1 + x_1x_2^{\ 2})$$
$$+ 2(x_2^{\ 3} - 1)(2.625 - x_1 + x_1x_2^{\ 3})$$
$$\frac{dg(x_1,x_2)}{dx_2} = 2x_1(1.5 - x_1 + x_1x_2) + 4x_1x_2(2.25 - x_1 + x_1x_2^{\ 2}) + 6x_1x_2^{\ 2}(2.625 - x_1 + x_1x_2^{\ 3})$$

$Let's\ now\ compute\ \nabla g(x_1^{\ *}, x_2^{\ *}).\ We\ notice\ that:$

$$1.5 - x_1^{\ *} + x_1^{\ *}x_2^{\ *} = 1.5 - 3 + 3 \times 0.5 = 0.$$
$$2.25 - x_1^{\ *} + x_1^{\ *}(x_2^{\ *})^2 = 2.25 - 3 + 3 \times (0.5)^2 = 0.$$
$$2.625 - x_1^{\ *} + x_1^{\ *}(x_2^{\ *})^3 = 2.625 - 3 + 3 \times (0.5)^3 = 0.$$

$Thus,\ according\ to\ the\ expression\ of\ \frac{dg(x_1,x_2)}{dx_1}\ and\ \frac{dg(x_1,x_2)}{dx_2}\ computed\ previously:$

$$\frac{dg(x_1^{\ *},x_2^{\ *})}{dx_1} = 0 \quad and \quad \frac{dg(x_1^{\ *},x_2^{\ *})}{dx_2} = 0.$$

$Thus,\ \nabla g(x_1^{\ *}, x_2^{\ *}) = (0, 0)^T\ so\ we\ can\ say\ that\ the\ true\ minimum\ deduced\ from\ question\ a)\ is\ a\ stationary\ point\ of\ the\ function\ g.$