Home problem 1

Contains solutions for:
Problem 1.1
Problem 1.3

Written by: Oscar Rosman
[991215-2791]

Problem 1.1: Penalty Method

Functions:

$$f(X_1, X_2) = (X_1 - 1)^2 + 2(X_2 - 2)^2$$

$$g(X_1, X_2) = X_1^2 + X_2^2 - 1 \leq 0$$

To add the penalty term which in this problem was only one inequality constraint the function g was simply squared and added to the function and no sum was necessary since there was only one constraint. This resulted in the two equations below depending on the value of the function g.

$$f_p(X_1, X_2; \mu) = (X_1 - 1)^2 + 2(X_2 - 2)^2 + \mu(X_1^2 + X_2^2 - 1)^2 \text{ } \mathbf{(If\ } \boldsymbol{g(X_1, X_2)} \geq \mathbf{0\,)}$$

$$f_p(X_1, X_2; \mu) = (X_1 - 1)^2 + 2(X_2 - 2)^2 \text{ } \mathbf{(otherwise)}$$

Gradient:

To find the gradient the partial derivative was taken on the function when the constraint was fulfilled for both X, see equations 1 and 2.

$$(1)\ \frac{\partial f_p(\mu = 0)}{\partial X_1} = (2)(1)(X_1 - 1)^{2-1} + 2(2)(0)(X_2 - 2)^{2-1} + 0 = 2X_1 - 2$$

$$(2)\ \frac{\partial f_p(\mu = 0)}{\partial X_1} = (2)(0)(X_1 - 1)^{2-1} + 2(2)(1)(X_2 - 2)^{2-1} + 0 = 4X_2 - 8$$

After finding the derivatives for when the constraint was fulfilled the partial derivatives were taken on the function when the constraint was unfulfilled, see equations 3 and 4.

$$(3)\ \frac{\partial f_p}{\partial X_1} = 2X_1 - 2 + \mu(2)(2X_1)(X_1^2 + X_2^2 - 1)^{2-1} = 2X_1 - 2 + 4\mu X_1(X_1^2 + X_2^2 - 1)$$

$$(4)\ \frac{\partial f_p}{\partial X_2} = 4X_2 - 8 + \mu(2)(2X_2)(X_1^2 + X_2^2 - 1)^{2-1} = 2X_1 - 2 + 4\mu X_2(X_1^2 + X_2^2 - 1)$$

Finally, the gradient can be determined by the vectors in (i) or (ii) below depending on if the constraint is fulfilled or not.

$$(\mathrm{i})\ \nabla f_p(\boldsymbol{X};\ \mu) = \left[\frac{\partial f_p}{\partial X_1}, \frac{\partial f_p}{\partial X_2}\right], \qquad (ii)\ \nabla f_p(\boldsymbol{X};\ \mu = 0) = \left[\frac{\partial f_p(\mu = 0)}{\partial X_1}, \frac{\partial f_p(\mu = 0)}{\partial X_2}\right]$$

Unconstrained minimum:

To find the unconstrained minimum equations (1) and (2) were set equal to zero to find the point P_0 containing $X_{1,0}$ and $X_{2,0}$ from where the search was later started.

$$\frac{\partial f_p(\mu=0)}{\partial X_1} = 0 \rightarrow 2X_1 - 2 = 0 \rightarrow X_1 = 1, \qquad \frac{\partial f_p(\mu=0)}{\partial X_2} = 0 \rightarrow 4X_2 - 8 = 0 \rightarrow X_2 = 2$$

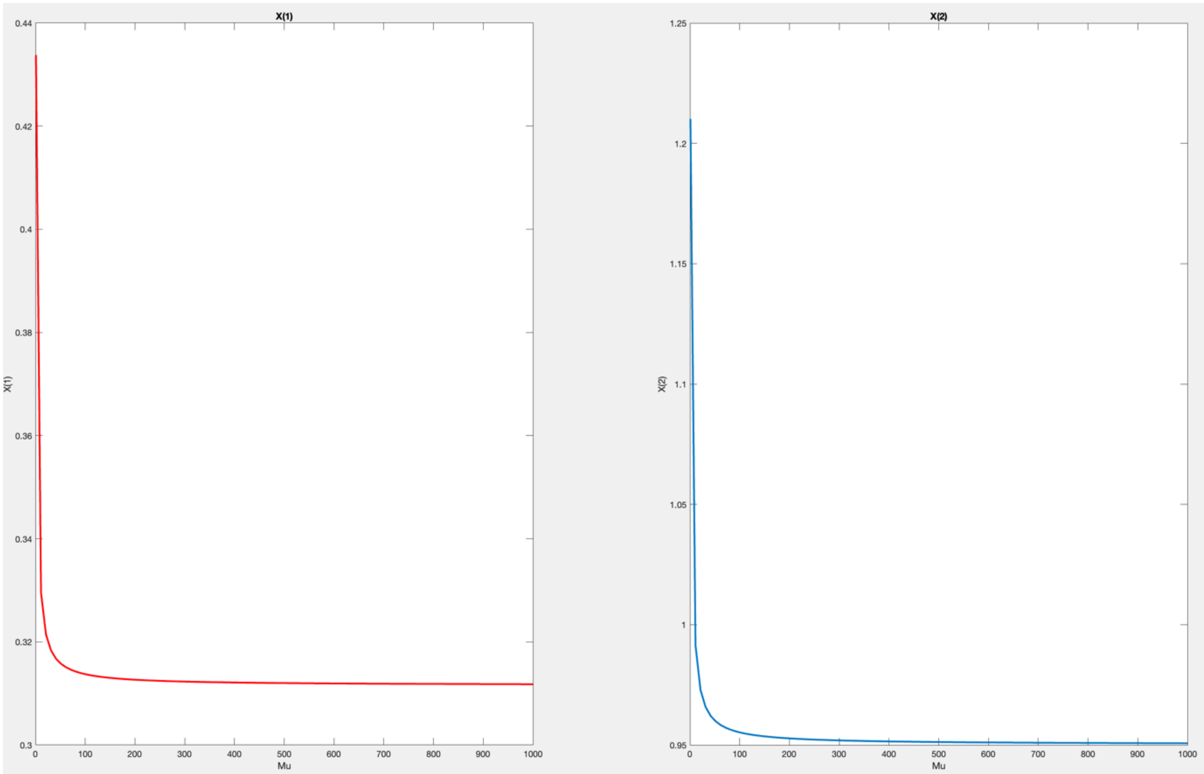
$$P_0 = \left(X_{1,0}, X_{2,0}\right) = (1, 2)$$

Results:

After running the Matlab script, the following table was the result.

μ	X_1	X_2	$f_p(\boldsymbol{X})$	$f_p(\boldsymbol{X};\ \mu)$	$g(\boldsymbol{X})$
1	0.4338	1.2102	1.5683	1.9943	0.6527
10	0.3314	0.9955	2.4650	2.5668	0.1009
100	0.3137	0.9553	2.6540	2.6659	0.0109
1000	0.3118	0.9507	2.6756	2.6768	0.0011

As mu or the penalty increases X_1 and X_2 moves away from the unconstrained minimum and converges to a point near (0.31, 0.95) where an increased function value balances a decreased value of the constraint. From the column furthest to the right the value of g(x) decreases as the penalty increase but it is still not fulfilled but is closing in to zero. The values of X_1 and X_2 were also plotted as a function of mu which shows convergence, see the figure below.



The parameters used to get these results was the same as the initial ones in the exercise. While making the program they were varied without significant effect on the result. Because of its small effect they were reverted to the original. Mu was also increased up to 10,000, because of the convergence the changes in the result were very small and the result omitted here in the report.

Problem 1.3 Basic GA program

a) Single Run

After experimenting with some different parameters, the parameters in the table below were chosen because they yielded a fitness value consistently close to one with an error of 10^{-10} or smaller. The results then in the ten runs were very similar.

Parameters used:

P_{Mut}	
Tournament size	5
Tournament probability	0.6
Crossover probability	0.8
Mutation probability	0.1
Number of generations	2000

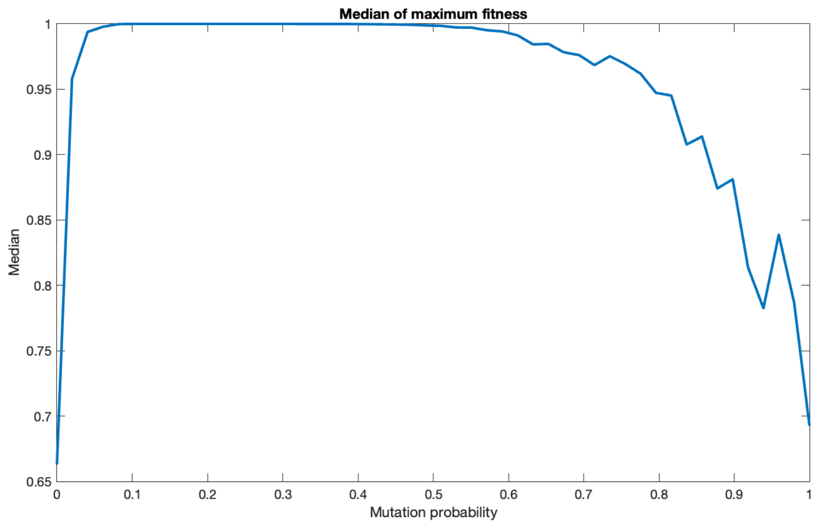
Result of 10 runs:

Run	$g(X_1, X_2)$	X_1	X_2
1	$9.7*10^{-12}$	2.999	0.499
2	$9.7*10^{-12}$	2.999	0.499
3	$9.7*10^{-6}$	3.008	0.502
4	$2.1*10^{-14}$	3.000	0.499
5	$2.1*10^{-14}$	3.000	0.499
6	$2.1*10^{-14}$	3.000	0.499
7	$7.9*10^{-13}$	3.000	0.500
8	$2.1*10^{-14}$	3.000	0.499
9	$2.1*10^{-14}$	3.000	0.499
10	$2.1*10^{-14}$	3.000	0.499

b) Varying P_{mut}

To get a high resolution when plotting the median performance as a function of the increasing mutation rate 50 mutation rates including the instructed ones were generated a run in "RunBatch.m". The result of ten of these are included in the table below and all are included in the graph.

index) P_{Mut}	Median fitness
1) 0.0	0.63117
2) 0.02	0.95800
10) 0.1837	0.99999
15) 0.2857	0.99998
20) 0.3878	0.99883
25) 0.4898	0.99881
30) 0.5918	0.99446
35) 0.6939	0.97629
40) 0.7959	0.95498
45) 0.8980	0.85355
50) 1.0	0.65222



In the graph and table, it is clear to see that the median performance increase quickly up until a mutation rate of 0.02 which corresponds well with the equation 1/m where m is the chromosome length which is normally used as an initial mutation rate. The median performance then increases further until 0.2 where it is steady for a while and then starts decreasing as the mutation rate grow closer to one. Judging by the table and the graph an optimal mutation rate for this problem is likely in the range [0.2, 0.3].

c) Finding true minimum

To find the true minimum of the function in the range [-5, 5] the function g(**X**) was partially differentiated to first find the stationary points.

$$(1)\ g(X_1, X_2) = (1.5 - X_1 + X_1X_2)^2 + (2.25 - X_1 + X_1X_2^2)^2 + (2.625 - X_1 + X_1X_2^3)^2$$

$$(1.1)\ \frac{\partial g}{\partial X_1} = 2(X_2 - 1)(1.5 - X_1 + X_1X_2) + 2(X_2^2 - 1)(2.25 - X_1 + X_1X_2^2) + 2(X_2^3 - 1)(2.625 - X_1 + X_1X_2^3)$$

$$(1.2)\ \frac{\partial g}{\partial X_2} = 2X_1(1.5 - X_1 + X_1X_2) + 4X_1X_2(2.25 - X_1 + X_1X_2^2) + 6X_1X_2^2(2.625 - X_1 + X_1X_2^3)$$

Instead of setting the derivatives to zero to find the stationary points an estimate was made from the points found in the ten single runs in section 1.3a to see if it corresponded to a stationary point. This point was estimated to (3.0, 0.5)

$$(1)\ g(3.0, 0.5) = 0.0$$

$$(1.1)\ \frac{\partial g}{\partial X_1}(3.0, 0.5) = 0.0$$

$$(1.2)\ \frac{\partial g}{\partial X_2}(3.0, 0.5) = 0.0$$

Since both derivatives were zero at this point the point is stationary. Investing points in all directions around this point proved it was a minimum. The fact that the GA consistently found this point in most runs suggests that it is the global minimum within the range [-5, 5].