Home Problem 1

Lisa Samuelsson, 000401-0047, lisasam@student.chalmers.se

September 19, 2022

Problem 1.1

1. We are given the function

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

and the constraint

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

The function $f_p(\mathbf{x}; \mu)$ is then defined as

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

2. We get two different cases:

Case 1: $x_1^2 + x_2^2 - 1 \leq 0$

$$\nabla f_p(\mathbf{x}; \mu) = (2(x_1 - 1),\ 4(x_2 - 2))$$

Case 2: $x_1^2 + x_2^2 - 1 > 0$

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

3. $\mu = 0$ gives us

$$\nabla f_p(\mathbf{x}; \mu = 0) = (2(x_1 - 1),\ 4(x_2 - 2))$$

We find the stationary points by setting the gradient equal to zero:

$$\nabla f_p(\mathbf{x}; \mu = 0) = (2(x_1 - 1),\ 4(x_2 - 2)) = 0$$

Which gives us:

$$\begin{cases} 2(x_2 - 1) = 0 \\ 4(x_2 - 2) = 0 \end{cases}$$

This system of equations has the solution $x_1 = 1, x_2 = 2$. This is the unconstrained minimum of the function.

4. See code in `./Problem1.1`.

5. Running the program gives following results:

μ	x_1^*	x_2^*
1	0.4338	1.2102
10	0.3314	0.9955
100	0.3137	0.9553
1000	0.3118	0.9507

Parameters used:
$\eta = 0.0001$
$\mu = \{1, 10, 100, 1000\}$
$T = 10^{-6}$

When $\mu \to \infty$, the stationary points of f_p should converge to the solution of the original function f. And, as we can see by looking at the result from running the Matlab code (presented in the table above) the values of x_1 and x_2 seems to converge for larger μ.

Problem 1.2

a. The aim is to find the global minimum $(x_1^*, x_2^*)^\top$ and the function value at this point for the function

$$f(x_1, x_2) = 4x_2^2 - x_1x_2 + 4x_2^2 - 6x_2$$

on the closed set S.

To do this, the analytical method is used. We should then evaluate the function in a number of points, namely:

1. The stationary points in the interior of the set S.
2. The stationary points at the boundary of the set S.
3. The corners of the set S.

The corner points are given, so we have:

$$P_1 = (0, 1)^\top$$

$$P_2 = (1, 1)^\top$$

$$P_3 = (0, 0)^\top$$

To find the stationary points in the interior of S, we take the partial derivate of $f(x_1, x_2)$ with respect to x_1 as well as x_2:

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

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

Setting these equal to 0 gives us:

$$\begin{cases} 8x_1 - x_2 = 0 \\ 8x_2 - x_1 - 6 = 0 \end{cases}$$

The first equation gives us :

$$x_2 = 8x_1$$

And plugging this into the second equation we get:

$$63x_1 = 6 \Rightarrow x_1 = \frac{2}{21},\ x_2 = 8 * \frac{2}{21} = \frac{16}{21}$$

We have found the stationary point in the interior of S:

$$P_4 = (\frac{2}{21}, \frac{16}{21})^\top$$

For the boundary, there are three cases. The first one is $0 < x_1 < 1$, $x_2 = 1$, and we shall therefore determine the stationary points of $f(x_1, 1) = 4x_1^2 - x_1 - 2$. We set the derivative to 0 and solve for x_1:

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

This gives us the point:

$$P_5 = (\frac{1}{8},\ 1)^\top$$

The second boundary is $0 < x_2 < 1$, $x_1 = 0$, and we shall therefore determine the stationary points of $f(0, x_2) = 4x_2^2 - 6x_2$. We set the derivative to 0 and solve for x_2:

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

This gives us the point:

$$P_6 = (0,\ \frac{3}{4})^\top$$

Finally, the third boundary is $x_1 = x_2,\ 0 < x_1 < 1$. We determine the stationary point of $f(x_1, x_1) = 7x_1^2 - 6x_1$ by setting the derivative to 0:

$$f'(x_1,\ x_1) = 14x_1 - 6 = 0 \implies x_1 = x_2 = \frac{3}{7}$$

Which gives us the last point:

$$P_7 = (\frac{3}{7},\ \frac{3}{7})^\top$$

Evaluating the function in the points gives the following results:

$$f(P_1) = f(0, 1) = -2$$

$$f(P_2) = f(1, 1) = 1$$

$$f(P_3) = f(1, 1) = 0$$

$$f(P_4) = f(\frac{2}{21}, \frac{16}{21}) = -\frac{16}{7} \approx -2.2857$$

$$f(P_5) = f(\frac{1}{8}, 1) = -\frac{33}{16} = -2.0625$$

$$f(P_6) = f(0, \frac{3}{4}) = -\frac{9}{4} = -2.25$$

$$f(P_7) = f(\frac{3}{7}, \frac{3}{7}) = -\frac{9}{7} \approx -1.2857$$

And we have now found that the function takes a minimum value $f(x_1^*, x_2^*) = -\frac{16}{7}$ at $P_4 = (\frac{2}{21}, \frac{16}{21})^\top$.

b. The goal is to find the point $(x_1^*, x_2*)^\top$ that gives us a minimum value for $f(x_1, x_2)$, subject to the constraint $h(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 21 = 0$.

We define $L(x_1, x_2, \lambda)$ as

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

We then compute ∇L and set it equal to 0:

$\frac{\partial L}{\partial x_1} = 2 + 2\lambda x_1 + \lambda x_2 = 0$ **(1)**

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

$\frac{\partial L}{\partial \lambda} = x_1^2 + x_1x_2 + x_2^2 - 21 = 0$ **(3)**

From **(1)**, we get that $x_2 = -\frac{2}{\lambda} - 2x_1$. Plugging this into **(2)**, we get the following equation:

$$3 + \lambda(x_1 + 2(-\frac{2}{\lambda} - 2x_1) = 0$$

Solving this will give us an expression for x_1:

$$x_1 = -\frac{1}{3\lambda}$$

If we plug this into the previous expression for x_2, we will get an expression for x_2 that is now only dependent on λ:

$$x_2 = -\frac{2}{\lambda} - 2(-\frac{1}{3\lambda}) = -\frac{4}{3\lambda}$$

By plugging our x_1 and x_2 into **(3)**, we get the equation:

$$(-\frac{1}{3\lambda})^2 + (-\frac{1}{3\lambda})(-\frac{4}{3\lambda}) + (-\frac{4}{3\lambda})^2 - 21 = 0.$$

Solving for λ gives us:

$$\frac{21}{9\lambda^2} = 21 \implies \lambda = \pm\frac{1}{3}.$$

Now, we have what we need to find the stationary points. We plug in λ in the expressions for x_1 and x_2 to find the points:

$$x_{1,1} = \frac{1}{3(-\frac{1}{3})} = -1 \qquad x_{1,2} = \frac{1}{3(\frac{1}{3})} = 1$$

$$x_{2,1} = \frac{4}{3(-\frac{1}{3})} = -4 \qquad x_{2,2} = \frac{4}{3(\frac{1}{3})} = 4$$

$$P_1 = (-1, -4)^\top \qquad P_2 = (1, 4)^\top$$

$$f(P_1) = 1 \qquad f(P_2) = 29$$

We find that the minimum function value is $f(P_1) = 1$, at the point $P_1 = (x_1^*, x_2^*)^\top = (-1, -4)^\top$.

Problem 1.3

a. See code in `./Problem1.3`. The following parameters were used:

Tournament size	4
Tournament probability	0.75
Crossover probability	0.8
Mutation probability	0.02
Number of generations	100

From running the program 10 times, the following results where obtained for x_1, x_2 and $g(x_1, x_2)$:

Run number	x_1	x_2	$g(x_1, x_2)$
1	2.9687	0.49215	0.0001
2	2.8125	0.44922	0.0070
3	2.8939	0.4687	0.0023
4	2.9930	0.4980	0.0000
5	3.2031	0.5469	0.0053
6	3.0469	0.5115	0.0003
7	3.0001	0.5001	0.0000
8	2.8906	0.4687	0.0023
9	3.0000	0.5000	0.0000
10	2.9930	0.4980	0.0000

b. Presented in the table and plot below are the median fitness values over 100 runs for different mutation rates. Although the median fitness doesn't vary by a lot, the highest median fitness occur when the mutation rate is $p_{mut} = 0.02 = \frac{1}{m}$. This means that the GA performs best when on average one gene in each chromosome mutates.

Mutation rate	**Median fitness value**
0	0.9934237903
0.01	0.9998576980
0.02	0.9999999645
0.05	0.9999937549
0.075	0.9999756747
0.1	0.9999510491
0.3	0.9991478387
0.5	0.9983462124
0.75	0.9987528955
0.9	0.9996482436

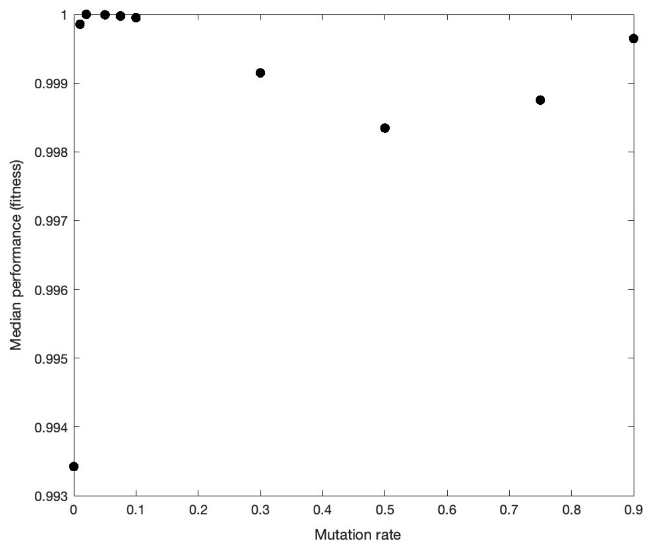

Figure 1: The median performance as a function of p_{mut}

c. Looking at the results obtained by running the program, as presented in **(a)**, it seems as if the true minimum of $g(x_1, x_2)$ is located at the point $(3, 0.5)^\top$, where the minimum value is $g(x_1^*, x_2^*) = 0$.

I now want to prove analytically that $(x_1^*, x_2^*)^\top$ actually is a stationary point of the function g. We can prove that this is the case by examining if $\nabla g(3, 0.5) = (0, 0)^\top$, as in a stationary point we know that $\frac{\partial f}{\partial x_2} = 0$ and $\frac{\partial f}{\partial x_1} = 0$.

$$\nabla g(x_1, x_2) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}]$$

$\frac{\partial f}{\partial x_1} = 2x_1x_2^6 + 2x_1x_2^4 + 5.25x_2^3 - 4x_1x_2^3 + 4.5x_2^2 - 2x_1x_2^2 + 3x_2 - 4x_1x_2 + 6x_1 - 12.75$

$\frac{\partial f}{\partial x_2} = 6x_1^2x_2^5 + 4x_1^2x_2^3 - 6x_1^2x_2^2 - 2x_1^2x_2 - 2x_1^2 + 15.75x_1x_2^2 + 9x_1x_2 + 3x_1$

We can now simply plug in our assumed optimal point, $(3, 0.5)^\top$, and this gives us the result:

$\frac{\partial f}{\partial x_1}(3, 0.5) = 2 \cdot 3 \cdot 0.5^6 + 2 \cdot 3 \cdot 0.5^4 + 5.25 \cdot 0.5^3 - 4 \cdot 3 \cdot 0.5^3 + 4.5 \cdot 0.5^2 - 2 \cdot 3 \cdot 0.5^2 + 3 \cdot 0.5 - 4 \cdot 3 \cdot 0.5 + 6 \cdot 3 - 12.75 = 0$

$\frac{\partial f}{\partial x_2}(3, 0.5) = 6 \cdot 3^2 \cdot 0.5^5 + 4 \cdot 3^2 \cdot 0.5^3 - 6 \cdot 3^2 \cdot 0.5^2 - 2 \cdot 3^2 \cdot 0.5 - 2 \cdot 3^2 + 15.75 \cdot 3 \cdot 0.5^2 + 9 \cdot 3 \cdot 0.5 + 3 \cdot 3 = 0$

Which means that

$$\nabla g(3, 0.5) = (0, 0)^{\top}.$$

We can thereby conclude that this point is indeed a stationary point of the function g.