Home Problem 1

Simon Rödén 0101225837

September 2025

1 Problem 1.1

1.1 1

We want to minimize $f(x_1, x_2)$ with the boundary $g(x_1, x_2)$. They are stated as

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

and

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

For the function we write

$$f_p(x, \mu) = f(x_1, x_2) + p(x_1, x_2; \mu)$$

where

$$p(x_1, x_2; \mu) = \begin{cases} \mu\left(x_1^2 + x_2^2 - 1\right)^2, & \text{if } x_1^2 + x_2^2 \leq 1, \\ 0, & \text{else} \end{cases}$$

is the penalty function.

Which gives us

$$f_p(x_1, x_2; \mu) = \begin{cases} (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu\left(x_1^2 + x_2^2 - 1\right)^2, & \text{if } x_1^2 + x_2^2 \leq 1, \\ (x_1 - 1)^2 + 2(x_2 - 2)^2, & \text{else.} \end{cases}$$

1.2 2

The gradient is given by

$$\nabla f_p(x_1, x_2; \mu) = \begin{bmatrix} \frac{\partial f_p}{\partial x_1} \\ \frac{\partial f_p}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2(x_1 - 1) + 4x_1\mu(x_1^2 + x_2^2 - 1) \\ 4(x_2 - 2) + 4x_2\mu(x_1^2 + x_2^2 - 1) \end{bmatrix}, \text{if } x_1^2 + x_2^2 \leq 1,$$

else

$$\nabla f_p(x_1, x_2; \mu) = \begin{bmatrix} \frac{\partial f_p}{\partial x_1} \\ \frac{\partial f_p}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2(x_1 - 1) \\ 4(x_2 - 2) \end{bmatrix}$$

1.3 3

For $\mu = 0$ the global minima is the same as the stationary point. Solving

$$\nabla f_p(x_1, x_2; 0) = \begin{bmatrix} 2(x_1 - 1) \\ 4(x_2 - 2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

gives us $x_1 = 1, x_2 = 2$, this will be our starting point for the penealty method.

1.4 5

x_1	x_1	μ
0.4340	1.2101	1.0
0.3315	0.9955	10.0
0.3139	0.9552	100.0
0.3116	0.9508	1000.0

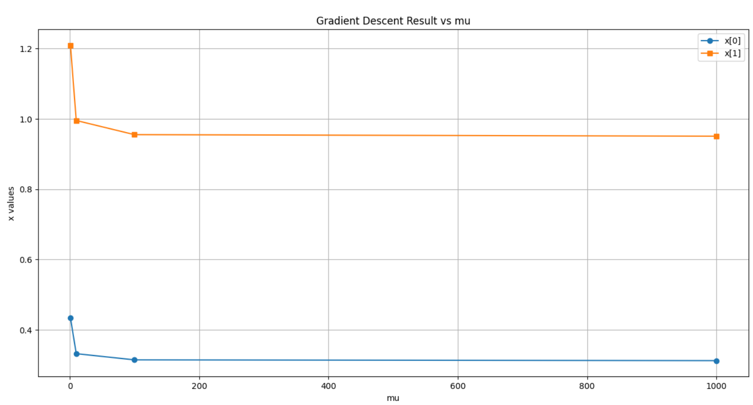


Figure 1: Plot with μ on the x-axis and the x values on the y-axis.

We can see in the plot 1 that the axis values decreases quickly, but the change from 100 to 1000 is almost nothing.

2 1.2

For the second problem we want to solve

$$\text{Max/min} f(x_1, x_2) = 4x_1^2 + 2x_2^3,$$

subject to the closed set S defined by

$$x_1^2 + x_2^2 \leq 4.$$

The method consist of three parts:

1. Find all stationary points inside S
2. Find stationary points on the boundary of S
3. Check the points and find the minima and maxima

For the first part we want to find were the gradient

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 8x_1 \\ 6x_2^2 \end{bmatrix}$$

is zero. This only happens when $x_1 = 0, x_2 = 0$. For the boundary $h(x_1, x_2) = 4x_1^2 + 2x_2^3$ we get the lagrange multiplier

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

Setting the gradient of L equal to zero, we obtain the system of equations:

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0,$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0,$$

$$\frac{\partial L}{\partial \lambda} = h(x_1, x_2) = 0.$$

which give

$$8x_1 + \lambda 2x_1 = 0, \tag{1}$$

$$6x_2^2 + \lambda 2x_2 = 0, \tag{2}$$

$$x_1^2 + x_2^2 - 4 = 0. \tag{3}$$

For equation 1 to be fulfiled we can solve for λ

$$x_1(4 + \lambda) = 0 => \lambda = -4.$$

Inserting $\lambda = -4$ into equation 2 gives

$$2x_2(3x_2 - 4) = 0 => x_2 = \frac{4}{3}, x_2 = 0.$$

Finally we can solve for x_1 using $x_2 = \frac{4}{3}$ and $x_2 = 0$ in equation 3

$$x_1^2 + (\frac{4}{3})^2 - 4 = 0 => x_1 = \pm\sqrt{\frac{20}{9}}$$

and

$$x_1^2 + 0^2 - 4 = 0 => x_1 = \pm 2$$

However, Equation 1 is also met when $x_1 = 0$ and then $x_2 = \pm 2$ for equation 3 to be fulfiled. The points we get are

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{\sqrt{20}}{3} \\ \frac{4}{3} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{20}}{3} \\ \frac{4}{3}. \end{bmatrix}$$

The minimal value of $f = -16$ is

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}.$$

The maximum value is $f = 16$ for

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}.$$

3 1.3

3.1 a

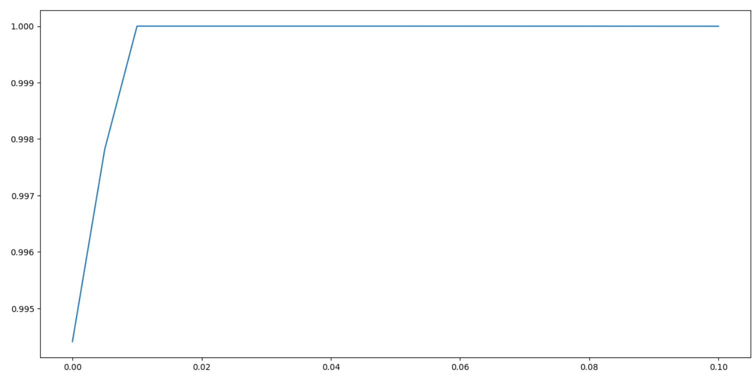
The chosen parameters were chosen tournament size = 4, tournament selection probability $p_{\text{tour}} = 0.65$, crossover probability $p_{\text{cross}} = 0.7$, mutation probability $p_{\text{mut}} = 0.02$, and number of generations = 3000. Outputing the following table:

$g(x_1, x_2)$	x_1	x_2
0.04383	3.750	0.6386
8.931e-6	2.993	0.4980
0.04383	3.750	0.6386
0.04383	3.750	0.6386
9.003e-6	2.993	0.4980
0.0001619	2.969	0.4922
2.089e-14	3.000	0.5000
0.04383	3.750	0.6386
0.002184	3.125	0.5294
2.089e-14	3.000	0.5000

3.2 b

For the different values of p_{mut}

p_{mut}	median performance
0.0	0.9944087643614843
0.005	0.9978174480438465
0.01	0.9999999999999791
0.02	0.9999999999781383
0.05	0.9999999382320441
0.1	0.9999985381028577



The performance increases rapidly and then gets worse slowly. As we can see from the table the best performance is 0.01. Decreasing makes the model worse quiker than increasing, therefore we can assume optimal p_{mut} is somewhere between 0.01 and 0.02.

3.3 c

The point we converge to seems to be

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0.5 \end{bmatrix}.$$

Taking the gradient of

$$g(x_1, x_2) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$$

gives

$$\nabla g(x_1, x_2) = \begin{bmatrix} 2(x_1x_2 - x_1 + 1.5)(x_2 - 1) + 2(x_1x_2^2 - x_1 + 2.25)(x_2^2 - 1) + 2(x_1x_2^3 - x_1 + 2.625)(x_2^3 - 1) \\ \\ 2(x_1x_2 - x_1 + 1.5)(x_1) + 2(x_1x_2^2 - x_1 + 2.25)(2x_1x_2) + 2(x_1x_2^3 - x_1 + 2.625)(3x_1x_2^2) \end{bmatrix}$$

Inserting $(3, 0.5)^T$ gives

$$(1.5 - 3 + 3 \cdot \tfrac{1}{2}) = -1.5 + 1.5 = 0$$

$$(2.25 - 3 + 3 \cdot (\tfrac{1}{2})^2) = -0.75 + 0.75 = 0$$

$$(2.625 - 3 + 3 \cdot (\tfrac{1}{2})^3) = -0.375 + 0.375 = 0$$

$\nabla g(3, 0.5) = (0, 0)^T$ which means that the point is a stationary point.