# Assignment 1

Anthon Odengard
19950114-3953

September 2023

## 1 Problem 1

### 1.1

As described in the coursebook the function we opt to minimize is defined as

$$f_p(\mathbf{x};\mu) = f(\mathbf{x}) + p(\mathbf{x};\mu) \tag{1}$$

where $f(\mathbf{x})$ is the objective function and $p(\mathbf{x};\mu)$ the penalty function. Furthermore, the penalty function is defined as

$$p(\mathbf{x};\mu) = \mu\left(\sum_{i=1}^{m}(\max\{g_i(\mathbf{x}),0\})^2 + \sum_{i=1}^{k}(h_i(\mathbf{x}))^2\right) \tag{2}$$

where $g_i(\mathbf{x})$ represents an inequality constraint and $h_i(\mathbf{x})$ an inequality constraint. In our case we are dealing with the function

$$f(\mathbf{x};\mu) = (x_1-1)^2 + 2(x_2-2)^2 \tag{3}$$

with the penalty term

$$g(\mathbf{x};\mu) = x_1^2 + x_2^2 - 1 \leq 0. \tag{4}$$

This yeilds that the function we opt to minimize equals

$$f_p(\mathbf{x};\mu) = \begin{cases} (x_1-1)^2 + 2(x_2-2)^2 + \mu(x_1^2+x_2^2-1)^2 & \text{if } g(\mathbf{x}) \geq 0 \\ (x_1-1)^2 + 2(x_2-2)^2 & \text{otherwise} \end{cases} \tag{5}$$

### 1.2

with the corresponding gradient

$$\nabla f_p(\mathbf{x};\mu) = \begin{cases} \big(2(x_1-1) + 4\mu x_1(x_1^2+x_2^2-1),\ 4(x_2-2) + 4\mu x_2(x_1^2+x_2^2-1)\big) & \text{if } g(\mathbf{x}) \geq 0 \\ \big(2(x_1-1),\ 4(x_2-2)\big) & \text{otherwise} \end{cases} \tag{6}$$

### 1.3

For the unconstrained case ($\mu = 0$), it's evident that the only stationary point that exist is the point $(1,2)$, which is a minima since $f(\mathbf{x}) \geq 0$ due to the squared nature of the function.

### 1.4

Running the penalty method for various $\mu$ yields

| $\mu$ | $x_1^*$ | $x_2^*$ |
|---|---|---|
| 1 | 0.4337 | 1.2101 |
| 10 | 0.3313 | 0.9955 |
| 100 | 0.3137 | 0.9552 |
| 1000 | 0.3117 | 0.9507 |

These results seem plausible, since $x_1^*$ and $x_2^*$ starts outside, but tends to the boundary of $S$ as the penalty term grows ($\mu$ increases). The parameters used was $\eta = 0.0001$ and $T = 10^{-6}$ as prescribed in the labb pm.

# 2 Problem 3

## 2.1 a)

The parameters used while running the GA was

| | |
|---|---|
| Population size | 100 |
| Maximum variable value | 5 |
| Number of genes | 50 |
| Number of variables | 2 |
| Number of generations | 5000 |
| Tournament size | 2 |
| Tournament probability | 0.75 |
| Crossover probability | 0.8 |

and the values returned by the algorithm was

| fitness | $x_1^*$ | $x_2^*$ |
|---|---|---|
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |
| 1 | 3 | 0.5 |

$$\mathbf{x}^* = (3, 0.5)$$

## 2.2 b)

While iterating for various mutation probabilities, the probabilities used were

$$\mathbf{p}_{\text{mut}} = [0, 0.02, 0.04, 0.06, 0.12, 0.20, 0.28, 0.32, 0.36, 0.4]$$

It was evident that the best mutation probability is when on average one gene per chromosome is mutated. For probabilities larger then this, median fitness is rapidly decreasing. The worst case was when there were no mutation at all, which is not surprising since this makes the algorithm prone to getting stuck at local maxima/minima.

| $p_{\mathrm{mut}}$ | Median fitness |
|---|---|
| 0 | 0.9949 |
| 0.02 | 0.9999 |
| 0.04 | 0.9999 |
| 0.06 | 0.9999 |
| 0.12 | 0.9998 |
| 0.20 | 0.9996 |
| 0.28 | 0.9994 |
| 0.32 | 0.9990 |
| 0.36 | 0.9989 |
| 0.40 | 0.9988 |

Table 1: Mutation probability in relation to median fitness.

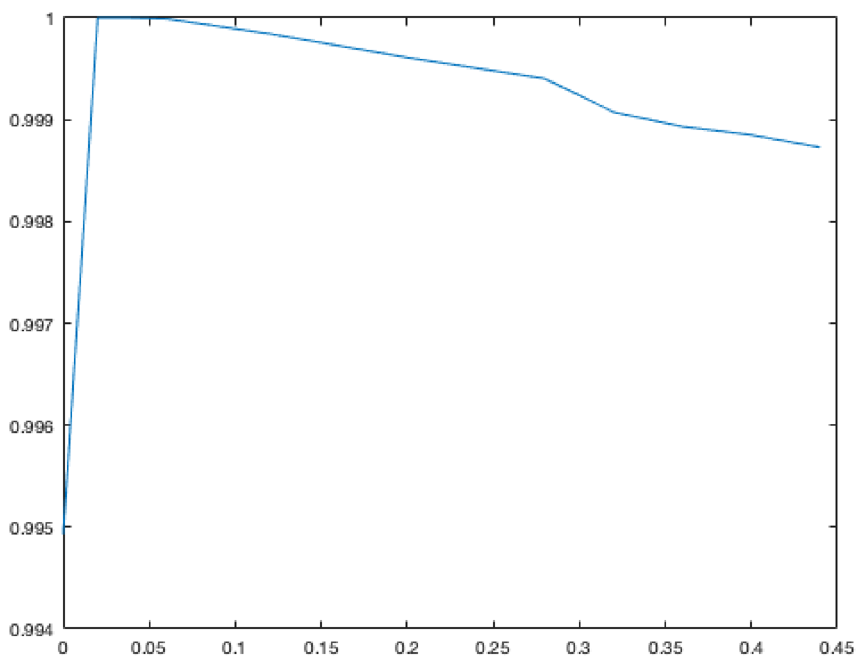


Figure 1: Plot showing fitness as a function of mutation probability

### 2.3 c)

From our numerical estimation of the minimum yielded that $(x_1^*, x_2^*) = (3, 0.5)$ seems to be the minimum. In order to verify that this is a stationary point, we want to investigate whether $\nabla g(x_1^*, x_2^*) = 0$. First we compute the gradient.

$$\nabla g(x_1^*, x_2^*) = \begin{cases} \dfrac{dg}{dx_1} =& 2 \cdot \big((1.5 - x_1 + x_1 x_2) \cdot (-1 + x_2) + (2.25 - x_1 + x_1 x_2^2) \cdot (-1 + x_2^2) \\ & + (2.625 - x_1 + x_1 x_2^3) \cdot (-1 + x_2^3)\big) \\ \dfrac{dg}{dx_2} =& 2 \cdot \big((1.5 - x_1 + x_1 x_2) \cdot x_1 + (2.25 - x_1 + x_1 x_2^2) \cdot (2x_1 x_2) + \\ & (2.625 - x_1 + x_1 x_2^3) \cdot (3x_1 x_2^2)\big) \end{cases} \tag{7}$$

Inserting the point $(x_1^*, x_2^*) = (3, 0.5)$ into the gradient results in

$$\nabla g(3, 0.5) = \begin{cases} \dfrac{dg}{dx_1} =& 2 \cdot \big((1.5 - 3 + 0.5 \cdot 3) \cdot (-1 + 0.5) + (2.25 - 3 + 3 \cdot 0.5^2) \cdot (-1 + 0.25) \\ & + (2.625 - 3 + 0.5^3 \cdot 3) \cdot (-1 + 0.125)\big) = 0 \\ \dfrac{dg}{dx_2} =& 2 \cdot \big((1.5 - 3 + 3 \cdot 0.5) \cdot 0.5 + (2.25 - 3 + 3 \cdot 0.5^2) \cdot (2 \cdot 0.5 \cdot 3) + \\ & (2.625 - 3 + 3 \cdot 0.5^3) \cdot (3 \cdot 3 \cdot 0.5^2)\big) = 0 \end{cases} \tag{8}$$

This shows that $\nabla g(x_1^*, x_2^*) = 0$, which proves that the point $(x_1^*, x_2^*) = (3, 0.5)$ is a stationary point for the function $g(x_1, x_2)$.