HP1

Karin Hult, 950711-1368

September 2021

Problem 1.1

We have the function

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2 \tag{1}$$

subject to the constraint

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

The penalty term is thus

$$p(x, \mu) = \mu \sum_{i=1}^{m} (max\{g_i(x), 0\})^2 = \mu(max\{x_1^2 + x_2^2 - 1, 0\})^2 \tag{3}$$

which gives us

$$f_p(x_1, x_2, \mu) = (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu(max\{x_1^2 + x_2^2 - 1, 0\})^2 \tag{4}$$

$$= \begin{cases} (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu(x_1^2 + x_2^2 - 1)^2, & \text{if } x_1^2 + x_2^2 - 1 \geq 0. \\ (x_1 - 1)^2 + 2(x_2 - 2)^2, & \text{otherwise.} \end{cases} \tag{5}$$

Let us then calculate the gradient, whereof we have for $x_1^2 + x_2^2 - 1 \geq 0$

$$\frac{df_p}{dx_1} = 2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1), \tag{6}$$

$$\frac{df_p}{dx_2} = 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1), \tag{7}$$

$$\frac{df_p}{d\mu} = (x_1^2 + x_2^2 - 1)^2, \tag{8}$$

$$\implies \nabla f_p(\mathbf{x}) = [2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1), 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1)] \tag{9}$$

and otherwise, if the constraint is fulfilled

$$\frac{df_p}{dx_1} = 2(x_1 - 1), \tag{10}$$

$$\frac{df_p}{dx_2} = 4(x_2 - 2), \tag{11}$$

$$\frac{df_p}{d\mu} = 0, \tag{12}$$

$$\implies \nabla f_p(\mathbf{x}) = [2(x_1 - 1), 4(x_2 - 2)]. \tag{13}$$

It is worth noting that for e.g. gradient descent, only the gradient with respect to x is used, even though I wanted to calculate it for μ as well in case it could have been of use.
For $\mu = 0$, f_p have a stationary point at $(1, 2)$ since that's where $\frac{df_p}{dx_1} = \frac{df_p}{dx_2} = 0$. At this point we also have

$$\frac{d^2 f_p}{dx_1^2}(1) = 2 > 0$$
$$\frac{d^2 f_p}{dx_2^2}(1) = 4 > 0$$
$$\frac{d^2 f_p}{dx_1 dx_2} = 0$$

and thus it is a minimum of the function without constraints.
Upon running the program we received values of x for each μ as shown in table 1. As we can see in figure 2 both x_1 and x_2 seem to converge, and therefore these values are most likely reasonable.

μ	x_1*	x_2*
0	1	2
1	0,4338	1,2102
10	0,3314	0,9955
100	0,3137	0,9553
1000	0,3118	0,9507

Table 1: Values of x_1* for each iteration of μ.

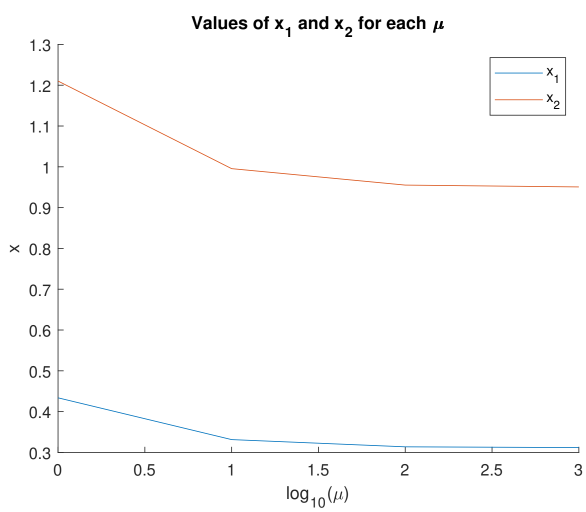

Figure 1: Values of x_1* and x_2* for different μ.

Problem 1.2

a)

We have the function

$$f(x_1, x_2) = 4x_1^2 - x_1x_2 + 4x_2^2 - 6x_2 \tag{14}$$

with the partial derivatives

$$\frac{df}{dx_1} = 8x_1 - x_2, \tag{15}$$

$$\frac{df}{dx_2} = 8x_2 - x_1 - 6. \tag{16}$$

To find the stationary points we need to find all points where $\frac{df}{dx_1} = \frac{df}{dx_2} = 0$.

$$8x_1 - x_2 = 0 \implies x_2 = 8x_1,$$

$$8x_2 - x_1 - 6 = 64x_1 - x_1 - 6 = 63x_1 - 6 = 0 \implies x_1 = \frac{6}{63} = \frac{2}{21}, \quad x_2 = \frac{16}{21},$$

whereof we can conclude the function only have one stationary point at $\mathbf{x} = (\frac{2}{21}, \frac{16}{21})^T$, which is within S.

We then need to consider any potential stationary points on the boundary of S, so we divide the function into the three cases where $x_1 = 0$, $x_2 = 1$ and $x_1 = x_2 = \tilde{x}$ and take the derivatives of these cases

$$f(0, x_2) = 4x_2^2 - 6x_2 \implies \frac{df}{dx_2} = 8x_2 - 6 = 0 \implies x_2 = \frac{3}{4}, \tag{17}$$

$$f(x_1, 1) = 4x_1^2 - x_1 \implies \frac{df}{dx_1} = 8x_1 - 1 = 0 \implies x_1 = \frac{1}{8}, \tag{18}$$

$$f(x, x) = 4x^2 - x^2 + 4x^2 - 6x = 7x^2 - 6x \implies \frac{df}{dx} = 14x - 6 = 0 \implies x_1 = x_2 = \frac{3}{7}. \tag{19}$$

Lastly we must consider the corners, i.e. $(0, 0)$, $(0, 1)$ and $(1, 1)$. Thus the candidates for a global minimum is as shown in table 2. Whereof we can observe from the value of f at each of these points that the global minimum is at $(x_1*, x_2*)^T = (\frac{2}{21}, \frac{16}{21})^T$

x_1	x_2	$f(x_1, x_2)$
2/21	16/21	-16/7 $\approx$ -2,29
0	3/4	-9/4 = -2,25
1/8	1	-33/16 $\approx$ -2,06
3/7	3/7	-9/7 $\approx$ -1.29
0	0	0
0	1	-2
1	1	1

Table 2: Values of f for each calculated candidate point.

b)

We have the function

$$f(x_1, x_2) = 15 + 2x_1 + 3x_2, \tag{20}$$

subject to the constraint

$$h(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 21 = 0. \tag{21}$$

To use Lagrange multiplier we will need a function L

$$\begin{aligned} L(x_1, x_2, \lambda) &= f(x_1, x_2) + \lambda h(x_1, x_2) \\ &= 15 + 2x_1 + 3x_2 + \lambda(x_1^2 + x_1x_2 + x_2^2 - 21), \end{aligned}$$

which can be used to calculate optima by computing the stationary points with $\nabla L = 0$.

$$\begin{aligned} \frac{dL}{dx_1} &= 2 + \lambda(2x_1 + x_2) = 0, \\ \frac{dL}{dx_2} &= 3 + \lambda(x_1 + 2x_2) = 0, \\ \frac{dL}{d\lambda} &= x_1^2 + x_1x_2 + x_2^2 - 21 = 0. \end{aligned}$$

By combining these equations we can calculate

$$\begin{gathered} \lambda = \frac{1}{x_1 - x_2}, \\ 2 + \frac{2x_1 + x_2}{x_1 - x_2} \implies x_2 = 4x_1, \\ x_1^2 + x_1x_2 + x_2^2 - 21 = x_1^2 + 4x_1^2 + 16x_1^2 - 21 = 21x_1^2 - 21 = 0, \implies x_1 = \pm 1. \end{gathered}$$

Since x_1 have two possible values we have now calculated the two stationary points to L $(1, 4, 1/3)$ and $(-1, -4, 1/3)$. We then investigate these points for $f(x_1, x_2)$ to determine the minimum.

$$f(1, 4) = 15 + 2 + 12 = 29, \tag{22}$$

$$f(-1, -4) = 15 - 2 - 12 = 3. \tag{23}$$

Thus the minimum is at $(x_1*, x_2*)^T = (-1, -4)^T$ with function value $f(-1, -4) = 3$.

Problem 1.3

a)

The selected parameters are shown in table 3. Whereof the 10 runs with these parameters yielded results shown in table 4.

$Size_{tour}$	p_{tour}	p_{Cross}	$p_{mutation}$	$N_{generations}$
2	0,8	0,8	0,02	2000

Table 3: Selected parameters for GA.

x_1*	x_2*	g(x_1, x_2)
3.0000045300	0.5000011772	3.351544e-12
3.0000072122	0.5000020713	1.017112e-11
3.0000307560	0.5000077337	1.516671e-10
3.0000030398	0.5000008792	1.841654e-12
2.9999970794	0.4999993891	1.661249e-12
2.9999982715	0.4999996871	7.869419e-13
2.9999988675	0.4999996871	2.290529e-13
2.9999923110	0.4999981970	9.710465e-12
2.9999988675	0.4999996871	2.290529e-13
3.0000033379	0.5000008792	1.844424e-12

Table 4: Values of x_1*, x_2* and g(x_1, x_2) .

b)

The statistics of fitness for different p_{Mut} is displayed in table 5, and figure 2 shows the median fitness as a function of p_{Mut}.

p_{Mut}	Median	Avarage	Standard Deviation
0,00	100,5653611655	144 891,5083162115	1 038 212,0261366601
0,01	10 401,4094155607	170 057 807 800,9904785156	843 019 123 917,5054931641
0,02	120 472 199,4221266806	64 652 110 722,7045898438	380 907 031 864,9739379883
0,03	1 687 399,5401257328	1 177 549 624,2165648937	9 192 052 615,2177734375
0,05	153 018,6461385706	36 782 149,8265300915	317 424 930,4626020193
0,07	47 877,5179952362	1 843 575,6805848465	13 471 449,9159224499
0,10	21 743,3073018005	173 939,3291741119	455 257,0586006998
0,20	2 590,9768679092	13 109,7427009996	349 12,13932320352
0,30	1 049,4237544862	3 431,6987519242	11 803,9303986041
0,40	708,4573640240	3 898,1838375287	11 702,2746910909
0,50	678,6477826176	2 577,9669104632	7 752,2735228678
0,60	667,4422714077	8 456,2458997346	49 387,5372451072
0,75	871,1215374605	3 837,7077346003	16 192,0772834256
0,90	2468,7888819455	9 039,4237182226	24 360,0421617138
1,00	102,5252700471	27 688,6747772135	147 506,7239781802

Table 5: Statistics of fitness for different values of p_{Mut}.

.

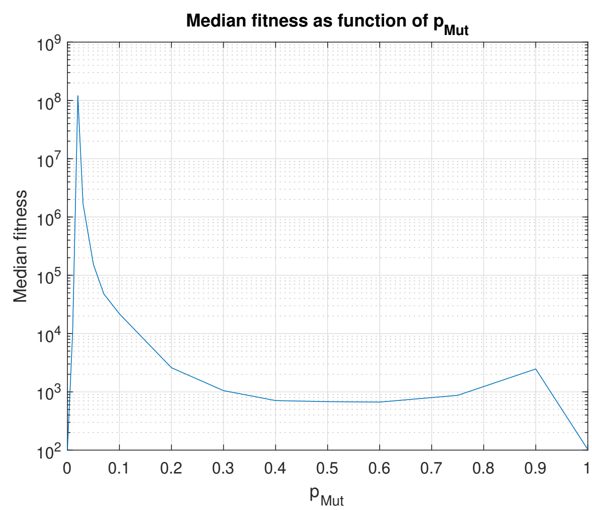

Figure 2: Median fitness for different p_{Mut}. Note the logarithmic y-axis.

These results seems to indicate a definite spike in fitness at $p_{\text{Mut}} = 0,02 = 1/n_{genes}$ with a slight but proportionally irrelevant increase at 0,9. When comparing the median and the mean it seems like some other values might in some cases obtain high fitness, but are less consistent and therefore less likely to yield desired results. Regardless it seems like some mutation is preferable to obtain a minimum.

c)

First we have our function g

$$g(x_1, x_2) = (1,5 - x_1 + x_1x_2)^2 + (2,25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2. \tag{24}$$

with its partial derivatives with respect to x_1 and x_2 being

$$\frac{dg}{dx_1} = 2\Big((1,5 - x_1 + x_1x_2)(x_2 - 1) + (2,25 - x_1 + x_1x_2^2)(x_2^2 - 1) + (2,625 - x_1 + x_1x_2^3)(x_2^3 - 1)\Big),$$

$$\begin{aligned}\frac{dg}{dx_2} &= 2\Big((1,5 - x_1 + x_1x_2)x_1 + (2,25 - x_1 + x_1x_2^2)2x_1x_2 + (2,625 - x_1 + x_1x_2^3)3x_1x_2^2\Big) \\ &= 2x_1\Big((1,5 - x_1 + x_1x_2) + 2x_2(2,25 - x_1 + x_1x_2^2) + 3x_2^2(2,625 - x_1 + x_1x_2^3)\Big).\end{aligned}$$

From the results in a) it seems likely that there is an exact minimum at $(x_1*, x_2*) = (3, 1/2)$. Putting these values into the partial derivatives of g and we obtain

$$\frac{dg}{dx_1}(3, \frac{1}{2}) = 2\Big((\frac{0}{2})(-\frac{1}{2}) + (\frac{0}{4})(-\frac{3}{4}) + (\frac{0}{8})(-\frac{7}{8})\Big) = 0 \tag{25}$$

$$\frac{dg}{dx_2} = 6\Big(0 + \frac{2}{2}0 + \frac{3}{4}0\Big) = 0 \tag{26}$$

Since both partial derivatives equals zero at this point, we have proven that $(x_1*, x_2*) = (3, 1/2)$ is indeed a stationary point. At this point the value of g(x_1, x_2) = 0, which is the smallest value possible for g(x_1, x_2) since it consists of a sum of squares. Of course this is assuming only real values of x are allowed.