Home problems, set 1

Mats Högberg, 930615-1979, matsho@student.chalmers.se

September 23, 2018

Problem 1.1

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

$$f_p(\mathbf{x};\mu) = \begin{cases} (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu * (x_1^2 + x_2^2 - 1)^2 & , \ g(x_1, x_2) > 0 \\ (x_1 - 1)^2 + 2(x_2 - 2)^2 & , \text{ otherwise} \end{cases}$$

Taking the partial derivatives of this function we get

$$\frac{\partial f_p}{\partial x_1} = \begin{cases} 2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1) & , \ g(x_1, x_2) > 0 \\ 2(x_1 - 1) & , \text{ otherwise} \end{cases}$$

$$\frac{\partial f_p}{\partial x_2} = \begin{cases} 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1) & , \ g(x_1, x_2) > 0 \\ 4(x_2 - 2) & , \text{ otherwise} \end{cases}$$

Which gives us the following gradient

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

3. When $\mu = 0$ we have

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

We find the minimum by setting $\nabla f_p(\mathbf{x};\mu = 0) = 0$, which gives us two equations

$$\begin{cases} 2(x_1^* - 1) = 0 \\ 4(x_2^* - 2) = 0 \end{cases}$$

which, when solved, gives us the following values for x_1^* and x_2^*

$$x_1^* = 1$$
$$x_2^* = 2$$

If we insert these values into $f_p(\mathbf{x}; \mu = 0)$ we get

$$f_p(1, 2; \mu = 0) = (1 - 1)^2 + 2(2 - 2)^2 = 0$$

We can also easily see that $f_p(\mathbf{x}; \mu = 0)$ can never be < 0, since both $(x_1 - 1)^2$ and $2(x_2 - 2)^2$ are ≥ 0 for all x_1, x_2. Thus the smallest value f_p can take is 0, and the point $\mathbf{x}^* = (1, 2)^\top$ is the global minimum.

4. See code in `./Problem 1.1/`.
5. Result from running the program:

μ	x_1^*	x_2^*
0	1	2
1	0.434	1.21
10	0.331	0.996
100	0.314	0.955
1000	0.312	0.951
10000	0.312	0.95

Parameters used:

$$\begin{aligned} &\eta = 0.00001 \\ &T = 10^{-6} \\ &\mu \in \{1, 10, 100, 1000, 10000\} \end{aligned}$$

Problem 1.2

a) From Figure 1, we can see that we have the following constraints:

$$0 \leq x_1, x_2 \leq 1$$

$$x_1 \leq x_2$$

First we consider the stationary points of f. These are found by setting $\nabla f = 0$. We have the following partial derivatives

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 8x_1 - x_2 \\ \frac{\partial f}{\partial x_2} &= -x_1 + 8x_2 - 6 \end{aligned}$$

Setting the partial derivatives to 0, we get the following equations

$$\begin{cases} 8x_1 - x_2 = 0 & (1) \\ -x_1 + 8x_2 - 6 = 0 & (2) \end{cases}$$

If we rearrange (1) we get that $x_2 = 8x_1$. Replacing x_2 with $8x_1$ in (2) we get

$$-x_1 + 64x_1 - 6 = 63x_1 - 6 = 0$$

Solving for x_1 and using (1) to calculate x_2, we get the following values for x_1 and x_2

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

Thus our stationary point is

$$P_1 = \left(\frac{2}{21}, \frac{16}{21}\right)^\top$$

We know that $P_1 \in S$, since $0 < \frac{2}{21}, \frac{16}{21} < 1$ and $\frac{2}{21} < \frac{16}{21}$.

Next, we consider the boundary of S, ∂S, including the corner points. We get 4 cases:

Case 1: $x_1 = 0$, $0 < x_2 < 1$

$$f(0, x_2) = 4x_2^2 - 6x_2$$
$$f'(0, x_2) = 8x_2 - 6$$
$$f'(0, x_2) = 0 \Rightarrow 8x_2 - 6 = 0 \Leftrightarrow x_2 = \frac{6}{8} \Rightarrow P_2 = \left(0, \frac{6}{8}\right)^\top$$

$0 < \frac{6}{8} < 1$, so $P_2 \in S$.

Case 2: $0 < x_1 < 1$, $x_2 = 1$

$$f(x_1, 1) = 4x_1^2 - x_1 + 4 - 6 = 4x_1^2 - x_1 - 2$$
$$f'(x_1, 1) = 8x_1 - 1$$
$$f'(x_1, 1) = 0 \Rightarrow 8x_1 - 1 = 0 \Leftrightarrow x_1 = \frac{1}{8} \Rightarrow P_3 = \left(\frac{1}{8}, 1\right)^\top$$

$0 < \frac{1}{8} < 1$, so $P_3 \in S$.

Case 3: $x_1 = x_2 = x$, $0 < x_1, x_2 < 1$

$$f(x, x) = 4x^2 - x^2 + 4x^2 - 6x = 7x^2 - 6x$$
$$f'(x, x) = 14x - 6$$
$$f'(x, x) = 0 \Rightarrow 14x - 6 = 0 \Leftrightarrow x = \frac{3}{7} \Rightarrow P_4 = \left(\frac{3}{7}, \frac{3}{7}\right)^\top$$

$0 < \frac{3}{7} < 1$, so $P_4 \in S$.

Case 4: Corners

$$P_5 = (0, 0)^\top$$
$$P_6 = (0, 1)^\top$$
$$P_7 = (1, 1)^\top$$

Now we can compute the value of $f(P_i), i \in \{1,2,3,4,5,6,7\}$

$$\begin{aligned}
P_1 &= \left(\frac{2}{21}, \frac{16}{21}\right)^\top &, \; f(P_1) &= -\frac{16}{7} \approx -2.29 \\
P_2 &= \left(0, \frac{6}{8}\right)^\top &, \; f(P_2) &= -\frac{9}{4} \approx -2.25 \\
P_3 &= \left(\frac{1}{8}, 1\right)^\top &, \; f(P_3) &= -\frac{33}{16} \approx -2.06 \\
P_4 &= \left(\frac{3}{7}, \frac{3}{7}\right)^\top &, \; f(P_4) &= -\frac{9}{7} \approx -1.29 \\
P_5 &= (0,0)^\top &, \; f(P_5) &= 0 \\
P_6 &= (0,1)^\top &, \; f(P_6) &= -2 \\
P_7 &= (1,1)^\top &, \; f(P_7) &= 1
\end{aligned}$$

We see that $f(x_1, x_2)$ takes its smallest value in P_1, so we have that $(x_1^*, x_2^*)^\top = P_1 = \left(\frac{2}{21}, \frac{16}{21}\right)^\top$ and $f(x_1^*, x_2^*) = -\frac{16}{7}$.

b) First we define $L(x_1, x_2, \lambda)$ as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) = 2x_1 + 3x_2 + 15 + \lambda(x_1^2 + x_1x_2 + x_2^2 - 21)$$

Next, we take the gradient of L and set it to 0 to obtain three equations

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

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

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

From (2) we get that $x_1 = -\frac{3}{\lambda} - 2x_2$. Replacing x_1 with $-\frac{3}{\lambda} - 2x_2$ in (1) we get

$$2 + \lambda(2(-\frac{3}{\lambda} - 2x_2) + x_2) = 0 \;\Leftrightarrow\; x_2 = -\frac{4}{3\lambda}$$

Replacing x_2 with $-\frac{4}{3\lambda}$ in (1) we get the following expression for x_1

$$x_1 = -\frac{3}{\lambda} - 2(-\frac{4}{3\lambda}) = -\frac{1}{3\lambda}$$

Now we can replace x_1 and x_2 in (3) with $-\frac{1}{3\lambda}$ and $-\frac{4}{3\lambda}$ respectively to get an expression for λ

$$(-\frac{1}{3\lambda})^2 + (-\frac{1}{3\lambda})(-\frac{4}{3\lambda}) + (-\frac{4}{3\lambda})^2 - 21 \;\Leftrightarrow\; \lambda = \pm\frac{1}{3}$$

Finally, we can replace λ with $\pm\frac{1}{3}$ in our equations for x_1 and x_2 to get two points

$$\begin{aligned}
P_1 &= \left(-\frac{1}{3(1/3)}, -\frac{4}{3(1/3)}\right)^\top &= (-1,-4)^\top, \; f(P_1) = 1 \\
P_2 &= \left(-\frac{1}{3(-1/3)}, -\frac{4}{3(-1/3)}\right)^\top &= (1,4)^\top \quad, \; f(P_2) = 29
\end{aligned}$$

We see that P_1 is a minimum, so we have that $(x_1^*, x_2^*)^\top = P_1 = (-1,-4)^\top$, and $f(x_1^*, x_2^*) = 1$.

Problem 1.3

a) See code in `./Problem 1.3/`.

From running the GA, we get that $(x_1^*, x_2^*)^\top = (0, -1)^\top$, $f(x_1^*, x_2^*) = 3$.

b) The median fitness values achieved by running the GA 100 times for each $p_{mut} \in \{0.00, 0.02, 0.05, 0.10\}$ are:

p_{mut}	Median fitness value
0.00	0.0959
0.02	0.3333
0.05	0.3324
0.10	0.3203

We can see that the GA performs very poorly with $p_{mut} = 0.00$, i.e. with no mutations at all. Having no mutations, only selection and crossover, probably causes the GA to often get stuck in a local optimum. We can also see that the GA performs its best when $p_{mut} = 0.02 = \frac{1}{m}$, which is also the optimal mutation rate for the Onemax problem.

c) We have $(x_1^*, x_2^*)^\top = (0, -1)^\top$. To prove that $(x_1^*, x_2^*)^\top$ is a stationary point of g, we want to show that $\nabla g(x_1^*, x_2^*) = (0, 0)^\top$, i.e. that $\frac{\partial g}{\partial x_1} = 0$ and $\frac{\partial g}{\partial x_2} = 0$ for $x_1 = 0, x_2 = -1$.

First, let's define some functions.

$$
\begin{aligned}
f_1(x_1, x_2) &= (x_1 + x_2 + 1)^2 \\
f_2(x_1, x_2) &= 19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2 \\
f(x_1, x_2) &= 1 + f_1(x_1, x_2)f_2(x_1, x_2) \\
h_1(x_1, x_2) &= (2x_1 - 3x_2)^2 \\
h_2(x_1, x_2) &= 18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2 \\
h(x_1, x_2) &= 30 + h_1(x_1, x_2)h_2(x_1, x_2)
\end{aligned}
$$

We have that $g(x_1, x_2) = f(x_1, x_2)h(x_1, x_2)$. Using the product rule, we have that

$$\nabla g(x_1, x_2) = \nabla f(x_1, x_2) * h(x_1, x_2) + f(x_1, x_2) * \nabla h(x_1, x_2)$$

We also have that

$$
\begin{aligned}
\nabla f(x_1, x_2) &= \nabla f_1(x_1, x_2) * f_2(x_1, x_2) + f_1(x_1, x_2) * \nabla f_2(x_1, x_2) \\
\nabla h(x_1, x_2) &= \nabla h_1(x_1, x_2) * h_2(x_1, x_2) + h_1(x_1, x_2) * \nabla h_2(x_1, x_2)
\end{aligned}
$$

Let's start by computing the values of f_1, f_2, h_1 and h_2 for $x_1 = 0, x_2 = -1$

$$
\begin{aligned}
f_1(0, -1) &= (-1 + 1)^2 = 0 \\
f_2(0, -1) &= 19 - 14 * (-1) + 3 * (-1)^2 = 36 \\
h_1(0, -1) &= (-3 * (-1))^2 = 9 \\
h_2(0, -1) &= 18 + 48 * (-1) + 27 * (-1)^2 = -3
\end{aligned}
$$

We can see directly that since $f_1(0, -1) = 0$, $\nabla f(0, -1) = \nabla f_1(0, -1) * f_2(0, -1)$, so we don't need to compute ∇f_2.

We continue by taking the partial derivatives of f_1, h_1 and h_2 and computing them for $x_1 = 0, x_2 = -1$

$$\begin{aligned}
\frac{\partial f_1}{\partial x_1} &= 2(x_1 + x_2 + 1) && \Rightarrow \frac{\partial}{\partial x_1} f_1(0, -1) = 0 \\
\frac{\partial f_1}{\partial x_2} &= 2(x_1 + x_2 + 1) && \Rightarrow \frac{\partial}{\partial x_2} f_1(0, -1) = 0 \\
\frac{\partial h_1}{\partial x_1} &= 4(2x_1 - 3x_2) && \Rightarrow \frac{\partial}{\partial x_1} h_1(0, -1) = 12 \\
\frac{\partial h_1}{\partial x_2} &= -6(2x_1 - 3x_2) && \Rightarrow \frac{\partial}{\partial x_2} h_1(0, -1) = -18 \\
\frac{\partial h_2}{\partial x_1} &= -32 + 24x_1 - 36x_2 && \Rightarrow \frac{\partial}{\partial x_1} h_2(0, -1) = 4 \\
\frac{\partial h_2}{\partial x_2} &= 48 - 36x_1 + 54x_2 && \Rightarrow \frac{\partial}{\partial x_2} h_2(0, -1) = -6
\end{aligned}$$

Using these values in the formulae for ∇f and ∇h we get

$$\begin{aligned}
&\frac{\partial}{\partial x_1} f(0, -1) = 0 * f_2(0, -1) = 0 \\
&\frac{\partial}{\partial x_2} f(0, -1) = 0 * f_2(0, -1) = 0 \\
&\frac{\partial}{\partial x_1} h(0, -1) = 12 * (-3) + 9 * 4 = 0 \\
&\frac{\partial}{\partial x_2} h(0, -1) = (-18) * (-3) + 9 * (-6) = 0
\end{aligned}$$

And, finally, using these values in the formula for ∇g gives us that

$$\nabla g(0, -1) = (0, 0)^\top * h(0, -1) + f(0, -1) * (0, 0)^\top = (0, 0)^\top$$

Which shows that $(x_1^*, x_2^*)^\top = (0, -1)^\top$ is a stationary point of g.