Home problem 1

FFR105 Stochastic optimization algorithms

Rebecca Svensson
960427-3061
rebsven@student.chalmers.se

September 23, 2019

1 Problem 1.1

For this problem the object function $f(x_1, x_2)$ is defined in equation 1 and the constraint $g(x_1, x_2)$ is defined in equation 2.

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

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

To use the penalty, method a new function $f_p(\boldsymbol{x}, \mu)$ is defines as given by equation 3. In this case the function $f_p(\boldsymbol{x}, \mu)$ is as presented in equation 4.

$$\begin{aligned} f_p(\boldsymbol{x}, \mu) &= f(\boldsymbol{x}) + p(\boldsymbol{x}, \mu) \\ p(\boldsymbol{x}, mu) &= \mu(\sum_{i=1}^{m}(max\{g_i(\boldsymbol{x}), 0\})^2 + \sum_{i=1}^{k}(h_i(\boldsymbol{x}))^2) \end{aligned} \tag{3}$$

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

The gradient of this is different within the constraint and outside of it. In equation 5 both cases are given.

$$\begin{aligned} &\text{Case 1: } x_1^2 + x_2^2 - 1 \leq 0 \\ &\quad \nabla f_p = (2(x_1 - 1), 4(x_2 - 2)) \\ &\text{Case 2: } x_1^2 + x_2^2 - 1 > 0 \\ &\quad \nabla f_p = (2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1), 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1) \end{aligned} \tag{5}$$

By putting the gradient for the unconstrained case equal to 0 it is apparent that a stationary point is located at $(x_1, x_2) = (1, 2)$.

When performing the penalty method using a Matlab program, the attained values for x_1 and x_2 differed depending on the value of μ. The attained values are presented in Table 1. As $\mu \longrightarrow \infty$, the stationary points of f_p should approach the stationary points of f. Therefore is is not surprising that the values for x_1 and x_2 when $\mu = 1$ differs a fair bit from the values attained for higher values of μ. The values of x_1 and x_2 seems to converge as μ increases as the difference between the values decreases.

μ	x_1	x_2
1	0.434	1.210
10	0.332	0.995
100	0.314	0.955
1000	0.311	0.951

Table 1: Minima of f_p located by using the penalty method with different values of μ.

Matlab program information: The main file is RunPenaltyMethod.m. The output is a table of the found minimum for each value of μ. No input parameters are required to run the main file.

2 Problem 1.2

2.1 a)

The aim is to find the global minimum of $f(x_1, x_2)$ defined in equation 6 on the closed set S. S is a triangle with the corners located at (0,0), (0,1) and (1,1).

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

The first step is to locate the minimum within the boundary. This is done by simply taking the partial derivatives of $f(x_1, x_2)$, as shown in equation 7, and the stationary point is concluded to be $(\frac{2}{21}, \frac{16}{21})^T$.

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 8x_1 - x_2 = 0 \\ \frac{\partial f}{\partial x_2} &= -x_1 + 8x_2 - 6 = 0 \end{aligned} \tag{7}$$

The remaining stationary points that may be the global minimum has to be located on the boundary. In equation 8 the three different constraints and functions are shown, along with the derivative for each case. From this the three stationary points on the boundary are concluded to be $(0, \frac{3}{4})^T$, $(\frac{1}{8}, 1)^T$ and $(\frac{3}{7}, \frac{3}{7})^T$.

$$\begin{aligned} &\text{Line 1 - (0,0) to (0,1):} \\ &x_1 = 0, 0 < x_2 < 1 \longrightarrow f(0, x_2) = 4x_2^2 - 6x_2 \\ &\frac{\partial f}{\partial x_2} = 8x_2 - 6 = 0 \\ \\ &\text{Line 2 - (0,1) to (1,1):} \\ &0 < x_1 < 1, x_2 = 1 \longrightarrow f(x_1, 1) = 4x_1^2 - x_1 - 2 \\ &\frac{\partial f}{\partial x_1} = 8x_1 - 1 = 0 \\ \\ &\text{Line 3 - (1,1) to (0,0):} \\ &x_1 = x_2 \longrightarrow f(x) = 7x^2 - 6x \\ &\frac{\partial f}{\partial x_1} = 14x - 6 = 0 \end{aligned} \tag{8}$$

Since these are the only possible locations for a global minimum to exist on this boundary, it is possible to identify which of these values of $(x_1, x_2)^T$ corresponds to the global minimum by evaluating $f(x_1, x_2)$ for each and every pair. This is done in equation 9.

$$\begin{aligned} f(\frac{2}{21}, \frac{16}{21}) &= -\frac{16}{7} \approx -2.29 \\ f(0, \frac{3}{4}) &= -\frac{9}{4} \approx -2.25 \\ f(\frac{1}{8}, 1) &= -\frac{33}{16} \approx -2.06 \\ f(\frac{3}{7}, \frac{3}{7}) &= -\frac{9}{7} \approx -1.29 \end{aligned} \tag{9}$$

From this it is concluded that $(\frac{2}{21}, \frac{16}{21})^T$ is the global minimum with the function value $f(\frac{2}{21}, \frac{16}{21}) = -\frac{16}{7}$.

2.2 b)

The minimum of a function $f(x_1, x_2)$ subject to the constraint of a function $h(x_1, x_2)$ is to be determined by using the Lagrange multiplier method. The two functions are defined in 10.

$$\begin{aligned} f(x_1, x_2) &= 15 + 2x_1 + 3x_2 \\ h(x_1, x_2) &= x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \end{aligned} \tag{10}$$

A Lagrange multiplier, λ, is introduced by defining a new function $L(x_1, x_2, \lambda)$ as shown in equation 11.

$$\begin{aligned} L(x_1, x_2, \lambda) &= f(x_1, x_2) = \lambda h(x_1, x_2) \\ L(x_1, x_2, \lambda) &= 15 + 2x_1 + 3x_2 + \lambda(x_1^2 + x_1 x_2 + x_2^2 - 21) \end{aligned} \tag{11}$$

To find the minimum of $f(x_1, x_2)$ the first step is to locate the stationary points of $L(x_1, x_2, \lambda)$ since the minimum we are looking for will occur at one of these stationary points. The partial equations of $L(x_1, x_2, \lambda)$ are found in equation 12.

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 2 + \lambda(2x_1 + x_2) = 0 \\ \frac{\partial L}{\partial x_2} &= 3 + \lambda(x_1 + 2x_2) = 0 \\ \frac{\partial L}{\partial \lambda} &= h(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \end{aligned} \tag{12}$$

By solving this equation system it is concluded that the stationary points are $(-1, -4, \frac{1}{3})^T$ and $(1, 4, -\frac{1}{3})^T$. The minimum is received by evaluation $f(x_1, x_2)$ for $(-1, -4)^T$ and $(1, 4)^T$ which is done in equation 13.

$$\begin{aligned} f(-1, -4) &= 1 \\ f(1, 4) &= 29 \end{aligned} \tag{13}$$

It is clear that the minimum is located at $(-1, -4)^T$.

3 Problem 1.3

3.1 Part a)

The minimum value of the function $g(x_1, x_2)$ is determined to be $g = 3.000$ in the specified range [-10,10]. This is located at $\boldsymbol{x^*} = (0.000, -1.000)$. For this optimisation the population size was set to 30, with each chromosome containing 40 genes, iterating over 500 generations. The crossover probability was set to 0.8, the mutation probability to 0.025 and the probability used in the tournament selection was set to 0.75 and the size of the tournament being 4. Each generation, the best individual was inserted 2 times in the new population. This is done 10 times and the individual with the best fitness value is chosen as the minimum. This is done since due to stochastic reasons, the program will sometimes converge prematurely.

Matlab program information: The main file is FunctionOptimization.m. This is run with no input and the output is the values of (x_1^*, x_2^*) for the minimum and the function value for this point.

3.2 Part b)

It is clear that mutation contribute to a much better optimisation and consistent results. When the mutation probability was set to 0, the median fitness value varied a lot between runs, while and increased mutation probability yielded a more consistent median fitness value. For a mutation probability larger than $\frac{1}{m}$ the median fitness value starts to decline slowly. It is evident that the genetic algorithm yields the best result when the mutation probability is such that, on average, one gene will mutate in each chromosome.

Mutation probability	Median fitness value
0.00	0.0824
0.02	0.3333
0.05	0.3322
0.1	0.3073

Matlab program information: The main file is FunctionOptimization.m. This is run with no input and the output is the median fitness values for the four specified mutation probabilites.

3.3 Part c)

The function to be considered, $g(x_1, x_2)$ is defined by equation 14. To make the computation easier to follow $g(x_1, x_2)$ is split in to pieces, $f(x_1, x_2)$ and $h(x_1, x_2)$ which are split one more time in to $f_1(x_1, x_2)$, $f_2(x_1, x_2)$, $h_1(x_1, x_2)$ and $h_2(x_1, x_2)$. These are defined in equation 15.

$$\begin{aligned} g(x_1, x_2) =&(1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2))\cdot \\ &(30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)) \end{aligned} \tag{14}$$

$$\begin{aligned}
f(x_1,x_2) &:= (1+(x_1+x_2+1)^2(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2))\\
&f_1(x_1,x_2) := (x_1+x_2+1)^2\\
&f_2(x_1,x_2) := (19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)\\
h(x_1,x_2) &:= (30+(2x_1-3x_2)^2(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2))\\
&h_1(x_1,x_2) := (2x_1-3x_2)^2\\
&h_2(x_1,x_2) := (18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2)
\end{aligned} \tag{15}$$

To prove that the point $(x_1^*, x_2^*) = (0, -1)$ is in fact a stationary point it is required that the gradient of $g(0, -1) = 0$. The definition of the gradient can be found in equation 16. To evaluate the derivatives of $g(x_1, x_2)$ with respect to x_1 and x_2 the product rule for derivatives is used, which yields the expressions found in equation 17.

$$\nabla g = (\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}) \tag{16}$$

$$\begin{aligned}
\frac{\partial g}{\partial x_1} &= \frac{\partial f}{\partial x_1}h + f\frac{\partial h}{\partial x_1} = (\frac{\partial f_1}{\partial x_1}f_2 + f_1\frac{\partial f_2}{\partial x_1})h + f(\frac{\partial h_1}{\partial x_1}h_2 + h_1\frac{\partial h_2}{\partial x_1})\\
\frac{\partial g}{\partial x_2} &= \frac{\partial f}{\partial x_2}h + f\frac{\partial h}{\partial x_2} = (\frac{\partial f_1}{\partial x_2}f_2 + f_1\frac{\partial f_2}{\partial x_2})h + f(\frac{\partial h_1}{\partial x_2}h_2 + h_1\frac{\partial h_2}{\partial x_2})
\end{aligned} \tag{17}$$

$$\begin{aligned}
\frac{\partial f_1}{\partial x_1}f_2 + f_1\frac{\partial f_2}{\partial x_1} =& 2(x_1+x_2+1)(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)+\\
&(x_1+x_2+1)^2(-14+6x_1+6x_2) = \{\ x_1 = 0 \text{ and } x_2 = -1\} = 0\\
\frac{\partial h_1}{\partial x_1}h_2 + h_1\frac{\partial h_2}{\partial x_1} =& 4(2x_1-3x_2)(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2)+\\
&(2x_1-3x_2)^2(-32+24x_1-36x_2) = \{\ x_1 = 0 \text{ and } x_2 = -1\} = 0\\
\frac{\partial f_1}{\partial x_2}f_2 + f_1\frac{\partial f_2}{\partial x_2} =& 2(x_1+x_2+1)(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)+\\
&(x_1+x_2+1)^2(-14+6x_1+6x_2) = \{\ x_1 = 0 \text{ and } x_2 = -1\} = 0\\
\frac{\partial h_1}{\partial x_2}h_2 + h_1\frac{\partial h_2}{\partial x_2} =& -6(2x_1-3x_2)((18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2))+\\
&(2x_1-3x_2)^2(48-36x_1+54x_2) = \{\ x_1 = 0 \text{ and } x_2 = -1\} = 0
\end{aligned} \tag{18}$$

By evaluating $f(x_1, x_2)$ and $h(x_1, x_2)$ along with equation 18 for $(x_1^*, x_2^*) = (0, -1)$ equation 19 is received. This shows that $(0, -1)$ is a stationary point.

$$\frac{\partial g}{\partial x_1} = (\frac{\partial f_1}{\partial x_1}f_2 + f_1\frac{\partial f_2}{\partial x_1})h + f(\frac{\partial h_1}{\partial x_1}h_2 + h_1\frac{\partial h_2}{\partial x_1}) = 0 \cdot h + f \cdot 0 = 0 \tag{19}$$