Home Problem 1

Anton Hedén 20001222-5157

Problem 1.1 Penalty method

Given the function f and and the constrains of g find the minimum of f by using the penalty method.

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

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

$$f_p(\boldsymbol{x}; \mu) = f(x) + p(\boldsymbol{x}; \mu)$$

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

Because we only have one constrain g described in (2)

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

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

then finding the the gradient $\nabla f_p(\boldsymbol{x}; \mu)$ is done by taking partial derivative at x_1 and x_2 when constrains are fulfilled

$$\nabla f_p(\boldsymbol{x}; \mu) = \nabla f_p(x_1, x_2) = \begin{bmatrix} \frac{\delta f}{\delta x_1} \\ \frac{\delta f}{\delta x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 - 2 \\ 4x_2 - 8 \end{bmatrix}$$

and when constrains are not fulfilled

$$\nabla f_p(\boldsymbol{x};\mu) = \nabla f_p(x_1,x_2) = \begin{bmatrix} \frac{\delta f}{\delta x_1} \\ \frac{\delta f}{\delta x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 - 2 + \mu(4x_1^3 + 4x_1x_2 - 4x_1) \\ 4x_2 - 8 + \mu(4x_2^3 + 4x_1x_2 - 4x_2) \end{bmatrix} \quad (3)$$

The unconstrained minimum are calculated by searching for where the derivative equal zero

$$\frac{\delta f}{\delta x_1} = 2x_1 - 2 = 0 \Rightarrow x_1 = 1$$
$$\frac{\delta f}{\delta x_2} = 4x_2 - 8 = 0 \Rightarrow x_2 = 2$$

then use this result $(x_1, x_2) \to (1, 2)$ as a starting point for finding the constrained minimum by writing a program that use the gradient decent method to calculate the constrained method.

Gradient decent program

When implementing gradient decent method to the program the gradient from equation (3) was used. This gradient was then multiplied with a small scalar in this case $\eta = 0.0001$ make the step length shorter. The program was then run until the gradient was smaller then a gradient tolerance T that was chosen to $t = 10^{-6}$. Then The output from varying impact of the penalty term by varying μ is calculated and show in table 1.

The difference between the next (x_1, x_2) and previous (x_1, x_2) in table 1 is decreasing with higher μ and therefore it shows that (x_1, x_2) is converging

$\mu\backslash$	x1	x2
1	0.4338	1.2102
10	0.3314	0.9955
100	0.3137	0.9553
1000	0.3118	0.9507

Table 1: Table that shows the values of the output of the gradient decent with varying with 4 decimal precision. With $\eta = 0.0001$, $T = 10^{-6}$ and varying μ.

Problem 1.2 Constrained optimization

1.2a) Analytical method

Determine the global minimum $(x_1^*, x_2^*)^T$ in the closed set S with corners of a triangle located at $(0,0), (0,1)$ and $(1,1)$ by using analytical method. Determine also corresponding value of function$f(x_1^*, x_2^*)$ in of function in equation (4).

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

Interior

First check for local minimum in the interior of S by looking for stationary points where the gradient of f is zero .

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\delta f}{\delta x_1} \\ \frac{\delta f}{\delta x_2} \end{bmatrix} = \begin{bmatrix} 8x_1 - x_2 \\ 8x_2 - x_1 - 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

by using substitution

$$\frac{\delta f}{\delta x_1} : 8x_1 - x_2 = 0 \Leftrightarrow 8x_1 = x_2 \Rightarrow \frac{\delta f}{\delta x_2} : 64x_1 - x_1 - 6 = 0$$

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

Stationary point found in $(\frac{2}{21}, \frac{16}{21})$ which is contained in set S. Then determine the corresponding function value.

$$f(\frac{2}{21}, \frac{16}{21}) = -2.1406$$

Boundary

Then the boundary of S is analysed analytical for finding stationary points

First the value of the function in the corners $(0,0), (0,1)$ and $(1,1)$ is determine

$$f(0,0) = 0 \quad (5)$$

$$f(0,1) = -2 \quad (6)$$

$$f(0,0) = 1 \quad (7)$$

The stationary points along the boundaries is evaluated.

First stationary points across the boundary $(0,0) \rightarrow (0,1)$ where $x_1 = 0$ is a constant

$$0 < x_2 < 1, x_1 = 0 \Rightarrow f(0, x_2) = 4x_2^2 - 6x_2 \quad (8)$$

$$\frac{\delta}{\delta x_2} f(0, x_2) = 8x_2 - 6 = 0 \Rightarrow x_2 = \frac{3}{4} \Rightarrow 0 < x_2 < 1 \quad (9)$$

The stationary point at line $(0,0) \rightarrow (0,1)$ is at $(0, \frac{3}{4})$. Corresponding value of function f is determined by

$$f(0, \frac{3}{4}) = -2.25 \quad (10)$$

stationary points across the boundary $(0,1) \rightarrow (1,1)$ where $x_2 = 1$ is a constant

$$0 < x_1 < 1, x_2 = 1 \Rightarrow f(x_1, 1) = 4x_1^2 - x_1 - 2 \quad (11)$$

$$\frac{\delta}{\delta x_1} f(x_1, 1) = 8x_1 - 1 = 0 \Rightarrow x_2 = \frac{1}{8} \Rightarrow 0 < x_1 < 1$$

The stationary point at line $(0,1) \rightarrow (1,1)$ is at $(\frac{1}{8},1)$. Corresponding value of function f is determined by

$$f(\frac{1}{8},1) = -2.065 \tag{12}$$

Stationary points across the third boundary $(1,1) \rightarrow (0,0)$ where $x_1 = x_2$

$$0 < x_1, x_2 < 1, x_1 = x_2 = x_* \Rightarrow f(x_*, x_*) = f(x_*) = 9x_*^2 - 6x_*$$

$$\frac{\delta}{\delta x_*} f(x_*) = 18x_* - 6 = 0 \Rightarrow x_* = \frac{1}{3} \Rightarrow 0 < x_* < 1$$

$$x_1 = x_2 = x_* = \frac{1}{3}$$

The stationary point at line $(0,0) \rightarrow (1,1)$ is at $(\frac{1}{3},\frac{1}{3})$. Corresponding value of function f is determined by

$$f(\frac{1}{3},\frac{1}{3}) = -1.2222 \tag{13}$$

Finding minimum

Then to find the minimum all stationary points and their corresponding value is show in the table 2 and the smallest function value and their corresponding points gives the minimum.

then taking all stationary points ant her corresponding function value as shown i table 2 the global minimum in the set S is at $(0,\frac{3}{4})$ when $f(0,\frac{3}{4}) = -2.25$

(x_1, x_2)	$f(x_1, x_2)$
$(\frac{2}{21}, \frac{16}{21})$	-2.1406
$(0, 0)$	0
$(0, 1)$	-2
$(1, 1)$	1
$(0, \frac{3}{4})$	-2.25
$(\frac{1}{8}, 1)$	-2.065
$(\frac{1}{3}, \frac{1}{3})$	-1.2222

Table 2: Table of the stationary points and there corresponding function value

1.2b) Lagrange multiplier

Determine the minimum $(x_1^*, x_2^*)^T$ of function $f(x_1, x_2)$ shown in equation 14 subject to the constrain $h(x_1, x_2)$ in equation 15 by using the Lagrange multiplier.

$$f(x_1, x_2) = 15 + 2x_1 + 3x_2 \tag{14}$$

$$h(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 21 = 0 \tag{15}$$

This is done by finding the point where ∇ft is perpendicular to the level curve and therefore fulfills this criteria of

$$\nabla f(x_1^*, x_2^*) = -\lambda \nabla h(x_1^*, x_2^2) \tag{16}$$

the Lagrange multiplier is defined by equation 17 and in this special case is represented as shown in equation 18

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) \tag{17}$$

$$L(x_1, x_2, \lambda) = 15 + 2x_1 + 3x_2 + \lambda(x_1^2 + x_1x_2 + x_2^2 - 21) \tag{18}$$

Then setting the gradient to zero

$$\nabla L(x_1, x_2, \lambda) = \begin{bmatrix} \frac{\delta L}{\delta x_1} \\ \frac{\delta L}{\delta x_2} \\ \frac{\delta L}{\delta \lambda} \end{bmatrix} = \begin{bmatrix} 2 + \lambda(2x_1 + x_2) \\ 3 + \lambda(2x_2 + x_1) \\ x_1^2 + x_1 x_2 + x_2^2 - 21 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{19}$$

Which solves when $x_1 = 1, x_2 = 4, \lambda = -\frac{1}{3}$ or $x_1 = -1, x_2 = -4, \lambda = \frac{1}{3}$. Therefore stationary point is at $(1, 4)$ and $(-1, -4)$

$$f(1, 4) = 29 > f(-1, -4) = 1 \tag{20}$$

Therefore minima of f with constrain h is at $f(-1, -4)$.

Problem 1.3 Basic GA program

1.3a) Run single

When running single GA parameters from table 3 was used and the result from the 10 runs is shown i table 4.

Parameter	Explanation	Value
$Size_{population}$	Population Size	100
V_{max}	Maximum variable value	5
N_{genes}	Number of genes	50
N_{var}	Number of variables	2
$Size_{tour}$	Tournament size	3
P_{tour}	Tournament probability	0.75
P_{cross}	Crossover probability	0.8
P_{mut}	Mutation probability	0.02
$N_{generations}$	Number of generations	500

Table 3: Used parameters for run single

Run	x_1	x_2	$g(x_1, x_2)$
1	3.0000	0.5000	$8.1046 * 10^{-13}$
2	2.9980	0.4995	$6.1189 * 10^{-7}$
3	3.1250	0.5294	0.0021
4	3.0287	0.5078	$1.4231 * 10^{-4}$
5	3.0287	0.5078	$1.4231 * 10^{-4}$
6	2.9687	0.4921	$1.6186 * 10^{-4}$
7	2.9980	0.4995	$6.1189 * 10^{-7}$
8	2.9929	0.4980	$8.9311 * 10^{-6}$
9	3.0106	0.5029	$2.0053 * 10^{-5}$
10	2.9998	0.4999	$2.3846 * 10^{-9}$

Table 4: Table over the different runs with

1.3b) Run batch

The optimal value of P_{mut} is according to the table 5 is $P_{mut} = 0.03$. This is not the case all the time because its dependent on the function, where that algorithm starts and the other parameters. However a good estimate is to choose $P_{mut} \approx \frac{1}{N_{genes}}$ which is similar to what the plot in figure 1 with the given $N_{genes} = 50$.

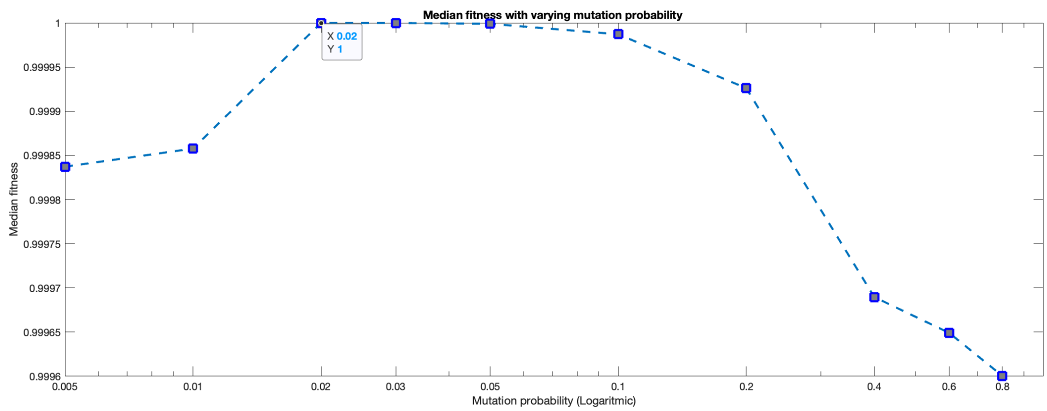

Figure 1: Plot over median of the best fitness values from 100 runs as a function of mutation probability (P_{mut})

1.3c) Prove analytically that the point found is a stationary point of the function g

Stationary point is given at (x_1^*, x_2^*) if $\nabla g(x_1^*, x_2^*) = 0$

$$\nabla g(x_1, x_2) = \begin{bmatrix} \frac{\delta g}{\delta x_1} \\ \frac{\delta g}{\delta x_2} \end{bmatrix} \tag{21}$$

Using the chain rule

$$\frac{\delta g}{\delta x_1} = 2(1.5 - x_1 + x_1 x_2)(x_2 - 1) + 2(2.25 - x_1 + x_1 x_2^2)(x_2^2 - 1) + 2(2.625 - x_1 + x_1 x_2^3)(x_2^3 - 1) \tag{22}$$

$$\frac{\delta g}{\delta x_2} = 2(1.5 - x_1 + x_1 x_2)x_1 + 2(2.25 - x_1 + x_1 x_2^2)(2x_1 x_2) + 2(2.625 - x_1 + x_1 x_2^3)(3x_1 x_2^2) \tag{23}$$

Then taking (x_1^*, x_2^*) that gives the smallest $g(x_1^*, x_2^*)$ from table 4 as a guess. In this case when $x_1^* = 3$ and $x_2^* = 0.5$.

P_{mut}	Median
0	0.9916688036
0.005	0.9978672205
0.01	0.9984745318
0.02	0.9999986518
0.03	0.9999996343
0.05	0.9999966390
0.1	0.9999707802
0.2	0.9998600877
0.4	0.9996896869
0.6	0.9996491041
0.8	0.9996005179

Table 5: Result from taking median at the best fitness value over 100 runs with different P_{mut}

$$\nabla f(3, 0.5) = \begin{bmatrix} \frac{\delta g}{\delta x_1} \\ \frac{\delta g}{\delta x_2} \end{bmatrix} = \begin{bmatrix} 2(0)(0.5-1) + 2(0)(0.5^2-1) + 2(0)(0.5^3-1) \\ 2(0)3 + 2(0)(3) + 2(0)(3 * 3 * 0.5^2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{24}$$

which fulfills the criteria for stationary points and $x_1^* = 3$ and $x_2^* = 0.5$ is stationary point to g