Home Problem 1

Konstantinos Zakkas
personnummer: 9206157795

September 2020

Problem 1.1, Penalty Method

1. The penalty function takes the form:

$$p(x) = \mu(max\{0,\, (x_1^2+x_2^2-1)\})^2 \quad (1.1)$$

Thus the function $f_p(x;\mu)$ for both the case where the constraints are not fullfilled and the case where they are(p(x; μ) = 0) is defined as:

$$f_p(x;\mu)=\begin{cases}(x_1-1)^2+2(x_2-2)^2+\mu(x_1^2+x_2^2-1)^2,\ if\ x_1^2+x_2^2-1\geqslant 0\\ (x_1-1)^2+2(x_2-2)^2,\ otherwise\end{cases} \quad (1.2)$$

2. In order to compute the gradient $\nabla f_p(x;\mu)$ we have to compute the derivative of $f_p(x;\mu)$ with respect to x_1 and the derivative with respect to x_2. The derivative with respect to x_1 for the case that the constraints are not fullfilled is the following:

$$2(x_1-1)+4\mu(x_1^3+x_1x_2^2-x_1) \quad (1.3)$$

The derivative with respect *to* x_1 when the constraints are fullfilled is:

$$2(x_1-1) \quad (1.4)$$

In the same way the two derivatives with respect to x_2 are:

$$4(x_2-2)+4\mu(x_2^3+x_1^2x_2-x_2) \quad (1.5) \quad \text{and} \quad 4(x_2-2) \quad (1.6)$$

Finally the gradient for the both cases is:

$$\nabla f_p(x;\mu)=\begin{cases}\begin{bmatrix}2(x_1-1)+4\mu(x_1^3+x_1x_2^2-x_1)\\ 4(x_2-2)+4\mu(x_2^3+x_1^2x_2-x_2)\end{bmatrix},\ if\ x_1^2+x_2^2-1\geqslant 0\\ \begin{bmatrix}2(x_1-1)\\ 4(x_2-2)\end{bmatrix},\ otherwise\end{cases} \quad (1.7)$$

3. Solving the equation system for which the constraints are fullfilled, therefore p(x; μ) = 0, $\begin{bmatrix}2(x_1-1)\\ 4(x_2-2)\end{bmatrix}$ we find that the unconstrained minimum of the function is:

$$(x_1,\, x_2)=(1,\, 2) \quad (1.8)$$

4. The main script is the file *RunPenaltyMethod.m* which calls the other two functions, generates and prints the output.

5. Writing *RunPenaltyMethod* in Matlab the program runs for the sequence of μ values: 1, 10, 100, 1000, for the step length η = 0.0001 and for the theshold value $T = 10^{-6}$. The output from the program is the following table:

μ	x_1^*	x_2^*
1	0.434	1.210
10	0.331	0.996
100	0.314	0.955
1000	0.312	0.951

Problem 1.2, Constrained optimization

a) First we find the stationary points of *f* on S. The partial derivatives of *f* are:

$$\frac{\partial f}{\partial x_1} = 8x_1 - x_2 \qquad (1.9)$$

$$\frac{\partial f}{\partial x_2} = -x_1 + 8x_2 - 6 \qquad (1.10)$$

So in order to find the stationary point we have to solve the equation system:

$$8x_1 - x_2 = 0 \qquad (1.11)$$

$$-x_1 + 8x_2 - 6 = 0 \qquad (1.12)$$

If we multiply equation (1.11) with 8 and add it to (1.12) we get the following equation:

$$63x_1 - 6 = 0 \qquad (1.13)$$

Therefore x_1 = 2/21 and the stationary point is $P_1 = (2/21, 16/21)^T$. This is the first potential location of the global minimum. Then we consider the boundary ∂S which is divided in three parts. Starting with the line $0 < x_1 < 1$, $x2 = 1$, we must determine the stationary points of $f(x_1, 1) = 4x_1^2 - x_1 - 2$. Taking the derivative we find:

$$8x_1 - 1 = 0 \qquad (1.14)$$

and therefore x_1 = 1/8. Thus the second point is $P_2 = (1/8, 1)^T$. We continue with the line $0 < x_2 < 1$, $x_1 = 0$ and we consider $f(0, x_2) = 4x_2^2 - 6x_2$. Taking the derivative we find:

$$8x_2 - 6 = 0 \qquad (1.15)$$

and therefore x_2 = 3/4, so another potential global minimum is P_3 = (0, 3/4). Finally, we consider the line $x_1 = x_2$ where $f(x_1, x_2) = 7x_1^2 - 6x_1$ and the derivative:

$$14x_1 - 6 = 0 \quad (1.16)$$

therefore $P_4 = (3/7, 3/7)^T$ emerges as a point to consider. We also consider the corners: $P_5 = (0, 0)^T$, $P_6 = (0, 1)^T$ and $P_7 = (1, 1)^T$. Finally we determine the function value of the seven points and we find: $f_{p1}(2/21, 16/21) = -2.2857$, $f_{p2}(1/8, 1) = -2.0625$, $f_{p3}(0, 3/4) = -2.25$, $f_{p4}(3/7, 3/7) = -1.2857$, $f_{p5}(0, 0) = 0$, $f_{p6}(0, 1) = -2$, $f_{p7}(1, 1) = 1$. Thus, the global minimum is located at the stationary point $\boldsymbol{P_1} \mathbf{= (2/21, 16/21)^T}$ for which f takes the value **-2.2857**.

b) Introducing the Lagrange multiplier λ, we generate the function $L(x, \lambda)$ as:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) = 15 + 2x_1 + 3x_2 + \lambda(x_1^2 + x_1 x_2 + x_2^2 - 21) \quad (1.17)$$

The stationary points of L occur where the following three equations hold:

$$\frac{\partial L}{\partial x_1} = 2 + 2\lambda x_1 + \lambda x_2 = 0 \qquad (1.18)$$
$$\frac{\partial L}{\partial x_2} = 3 + \lambda x_1 + 2\lambda x_2 = 0 \qquad (1.19)$$
$$\frac{\partial L}{\partial \lambda} = x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \quad (1.20)$$

Then we multiply (1.18) with -2 and add it to (1.19) and we get:

$$\lambda x_1 = -1/3 \quad (1.21)$$

Using (1.21) with (1.19) we get:

$$\lambda x_2 = -4/3 \quad (1.22)$$

From (1.21) and (1.22) we determine that $\lambda \neq 0$ so the two equations become:

$$x_1 = -1/3\lambda \quad (1.23)$$
$$x_2 = -4/3\lambda \quad (1.24)$$

Using (1.23) and (1.24) with (1.20) we find:

$$\lambda = \pm 1/3 \quad (1.25)$$

Thus from (1.23) and (1.24) we obtain two stationary points:

$$(-1, -4) \quad (1.26)$$
$$(1, 4) \qquad (1.27)$$

For which the function f takes the values $f(-1, -4) = 1$ and $f(1, 4) = 29$, thus the minima occurs at the point:

$$(-1, -4)^T \quad (1.27)$$

and the corresponding function value is $f(-1, -4) = 1$.

Problem 1.3, Constrained optimization

a) The main script of the program is the file *FunctionOptimization.m.* We set the population size to 100, the chromosome length to 50, the crossover probability to 0.8, the mutation probability to 0.02, the tournament selection parameter to 0.75. the tournament size to 2, the number of copies of best individual in the elitism step to 1 and the number of generations to 100. As a result, after one run, we get a point very close to (x_1, x_2) = (0, -1) and a function value very close to 3. Thus the minimum is at the location (0, -1) and the minimum value of the function is $g(0, -1) = 3$.

b) The median fitness values obtained for each value of the mutation rate are:

mutation rate	median fitness value
0.00	0.0759
0.02	0.3333
0.05	0.3321
0.10	0.3178

The median fitness values for all mutation rates we used, except for 0, are very close to the fitness of the minimum.The best result comes for a mutation rate 0.02 which, according to theory, is the optimal rate(1/chromosome length) and worse as we increase it. Therefore for small mutation rates the program converges very close to minimum but for mutation rate zero we get a much worse result. This is caused because without mutation there is a loss of diversity in the population and less fitted individuals are generated in the long run. Therefore the population will gather around a local optimum which dominates the population and this results to premature convergence. Using suitable values of mutation however will place individuals closer to other local optima or even the global optimum, giving better results. Also with zero mutation rate we can’t avoid re-evaluation of identical individuals which does not improve our results.

c) First we find the partial derivatives of the function $g(x_1, x_2)$:

$$\frac{\partial g}{\partial x_1} = ((2(x_1+x_2+1)(19-4x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2))+((x_1+x_2+1)^2(15+6x_1+6x_2))) \times ((4(2x_1-3x_2)(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2))+((2x_1-3x_2)^2(-32+24x_1-36x_2))) \quad (1.28)$$

$$\frac{\partial g}{\partial x_2} = ((2(x_1+x_2+1)(19-4x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2))+((x_1+x_2+1)^2(-14+6x_1+6x_2))) \times ((-6(2x_1-3x_2)(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2))+((2x_1-3x_2)^2(16-36x_1+54x_2))) \quad (1.29)$$

The points at which the two partial derivatives are equal to 0 are the stationary points. Therefore if the point (0, -1) we found is a solution to both (1.28), (1.29) when they are set equal to zero, it is a stationary point. First we check the first term of the larger multiplication for both derivatives. For $x_1 = 0$ and $x_2 = -1$ the $x_1+x_2+1 = 0$ and multiplies the two terms of the sum which forms the first term of the multiplication both in (1.28) and

(1.29), so both sums are equal to zero. Finally each sum(which now is zero) multiplies the second term of its corresponding derivative so both derivatives become zero, thus the point (0, -1) is a stationary point of $g(x_1, x_2)$.