Modern Optimization Techniques – Group 01

Exercise Sheet 09

Submitted by: Muhammad Inaam Ashraf (Matrikel-Nr: 307524)

Semester 2 MSc. Data Analytics

Question 1: Barrier and Penalty methods

(a). Explain in your own words how the barrier and penalty methods work for a general constrained minimization problem and how they differ..

Barrier Methods	**Penalty Methods**
We try to find the optimum solution while staying inside the interior of the feasible area.	We try to find the optimum solution in unconstrained manner.
A barrier function is added to the objective function such that the barrier function goes to infinity as we approach the border of the feasible area i.e. $f(x) + c.B(x)$	A penalty function is added to the objective function such that the penalty function becomes zero as we approach the optimum solution i.e. $f(x) + c.P(x)$
The value of the weight c is iteratively reduced such that it becomes 0 as we approach the border.	The value of the weight c is iteratively increased such that it approaches infinity as we approach the optimum.
Only work for inequality constraints, however equality constraints can be passed through to the inner problem.	Work for both inequality and equality constraints, but there are no inequality constraints in most cases. In cases where there are inequality constraints, these are represented as equality constraints and the resulting function is not differentiable at the border. So, subgradients are used.

(b). $\boldsymbol{minimze\ x_1{}^2 + x_2{}^2, subject\ to: x_1 \geq 0,\ \ x_2 \leq 0,\ \ x_1 - x_2 \geq 2}$

Standard form is given as

$$argmin_{x \in R} f(x),$$

$$subject\ to\ g_p(x) = 0\ \ for\ p = 1, \dots, \dots, P$$

$$and\ h_q(x) \leq 0\ \ for\ q = 1, \dots, \dots, Q$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad p = 0,\ \ so\ g_p(x)\ does\ not\ exist\ , \qquad q = 3$$

So

$$h_1(x) = -x_1\ , \qquad h_2(x) = x_2, \qquad h_3(x) = -x_1 + x_2 + 2$$

OR

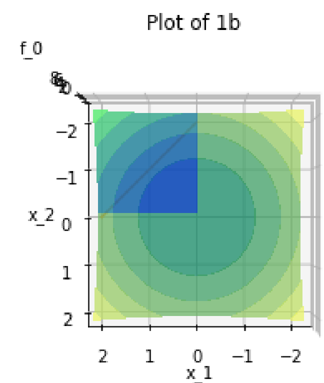
$$argmin_{x\in R} f(x), subject\ to\ Ax - a = 0\ and\ Bx - b \leq 0$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad A = 0, \qquad a = 0,$$

$$B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Plot is shown on the right. The area to the left of the line segment is the feasible region.



The objective function using the quadratic penalty function can be written as

$$f(x) + c.P(x) = f(x) + c.\sum_{p=1}^{P} \left(g_p(x)\right)^2$$

Since we have inequality constraints, we have to write these in terms of equality constraints as follows:

$$h_q^+(x) = \max\{0, h_q(x)\}$$

So,

$$P(x) = \sum_{q=1}^{Q} \left(h_q^+(x)\right)^2$$

And the objective function becomes

$$f(x) + c.P(x) = f(x) + c.\sum_{q=1}^{Q} \left(h_q^+(x)\right)^2$$

$$f(x) + c.P(x) = x_1^2 + x_2^2 + c.((\max\{0, -x_1\})^2 + (\max\{0, x_2\})^2 + (\max\{0, -x_1 + x_2 + 2\})^2)$$

Taking the derivative

w.r.t to x_1:

$$\begin{aligned}\nabla_{x_1} f(x) + c.\nabla_{x_1} P(x) \\ &= 2x_1 \\ &+ c\big((\max\{0, -x_1\}(-1) + (-1)\max\{0, -x_1\}) + 0 \\ &+ (\max\{0, -x_1 + x_2 + 2\}(-1) + (-1)\max\{0, -x_1 + x_2 + 2\})\big)\end{aligned}$$

$$\nabla_{x_1} f(x) + c.\nabla_{x_1} P(x) = 2x_1 + c\big((-2\max\{0, -x_1\}) + 0 + (-2\max\{0, -x_1 + x_2 + 2\})\big)$$

$$\nabla_{x_1} f(x) + c.\nabla_{x_1} P(x) = 2x_1 - 2c(\max\{0, -x_1\} + \max\{0, -x_1 + x_2 + 2\})$$

w.r.t to x_2:

$$\nabla_{x_2} f(x) + c.\nabla_{x_2} P(x)$$
$$= 2x_2$$
$$+ c\big(0 + (\max\{0, x_2\}\,(1) + (1)\max\{0, x_2\})$$
$$+ (\max\{0, -x_1 + x_2 + 2\}\,(1) + (1)\max\{0, -x_1 + x_2 + 2\})\big)$$

$$\nabla_{x_2} f(x) + c.\nabla_{x_2} P(x) = 2x_2 + c\big(0 + (2\max\{0, x_2\}) + (2\max\{0, -x_1 + x_2 + 2\})\big)$$

$$\nabla_{x_2} f(x) + c.\nabla_{x_2} P(x) = 2x_2 + 2c(\max\{0, x_2\} + \max\{0, -x_1 + x_2 + 2\})$$

Finding unconstrained minimum

$$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = 0$$

This gives

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Taking

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla f\big(x^{(0)}\big) + c.\nabla P\big(x^{(0)}\big) = \begin{bmatrix} 2x_1 - 2c(\max\{0, -x_1\} + \max\{0, -x_1 + x_2 + 2\}) \\ 2x_2 + 2c(\max\{0, x_2\} + \max\{0, -x_1 + x_2 + 2\}) \end{bmatrix} = \begin{bmatrix} -4c \\ 4c \end{bmatrix}$$

So

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \mu\left(\nabla f\big(x^{(0)}\big) + c.\nabla P\big(x^{(0)}\big)\right)$$

With $\mu = 1$

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 1\left(\begin{bmatrix} -4c \\ 4c \end{bmatrix}\right) = \begin{bmatrix} 4c \\ -4c \end{bmatrix}$$

Next iteration

$$\nabla f\big(x^{(1)}\big) + c.\nabla P\big(x^{(1)}\big) = \begin{bmatrix} 8c - 2c(\max\{0, -4c\} + \max\{0, -4c - 4c + 2\}) \\ -8c + 2c(\max\{0, -4c\} + \max\{0, -4c - 4c + 2\}) \end{bmatrix}$$

I don’t know what to do next.

So, solving another way. Taking derivative and putting it equal to 0.

$$\begin{bmatrix} 2x_1 - 2c(\max\{0, -x_1\} + \max\{0, -x_1 + x_2 + 2\}) \\ 2x_2 + 2c(\max\{0, x_2\} + \max\{0, -x_1 + x_2 + 2\}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We have

$$x_1 = c(\max\{0, -x_1\} + \max\{0, -x_1 + x_2 + 2\})$$

$$x_2 = -c(\max\{0, x_2\} + \max\{0, -x_1 + x_2 + 2\})$$

Or

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} c(\max\{0,-x_1\} + \max\{0,-x_1+x_2+2\}) \\ -c(\max\{0,x_2\} + \max\{0,-x_1+x_2+2\}) \end{bmatrix}$$

Putting

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} c(\max\{0,-x_1\} + \max\{0,-x_1+x_2+2\}) \\ -c(\max\{0,x_2\} + \max\{0,-x_1+x_2+2\}) \end{bmatrix} = \begin{bmatrix} \mathbf{2c} \\ \mathbf{-2c} \end{bmatrix}$$

(c). $\boldsymbol{minimze\ x_1{}^2 + x_2{}^2, subject\ to: x_1 \geq 1, x_1 \leq 3,\ \ x_2 \geq -3, x_2 \leq -1}$

Standard form is given as

$$argmin_{x\in R} f(x),$$

$$subject\ to\ g_p(x) = 0\ \ for\ p = 1, \dots, \dots, P$$

$$and\ h_q(x) \leq 0\ \ for\ q = 1, \dots, \dots, Q$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad P = 0,\ \ so\ g_p(x)\ does\ not\ exist\ , \qquad Q = 4$$

So

$$h_1(x) = -x_1 + 1\ , \qquad h_2(x) = x_1 - 3, \qquad h_3(x) = -x_2 - 3, \qquad h_4(x) = x_2 + 1$$

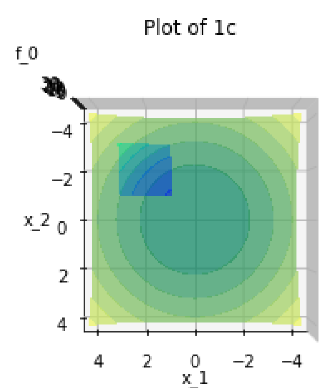
OR

$$argmin_{x\in R} f(x), subject\ to\ Ax - a = 0\ and\ Bx - b \leq 0$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad A = 0, \qquad a = 0,$$

$$B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} -1 \\ 3 \\ 3 \\ -1 \end{bmatrix}$$



Plot is shown on the right. The shaded box is the feasible region.

The objective function using the log barrier function can be written as

$$f(x) + c.B(x) = f(x) + c.(-\textstyle\sum_{q=1}^{Q} \log(-h_q(x)))$$

$$f(x) + c.B(x) = x_1^2 + x_2^2 + c.(-\log(x_1 - 1) - \log(-x_1 + 3) - \log(x_2 + 3) - \log(-x_2 - 1))$$

Taking the derivative

w.r.t to x_1:

$$\nabla_{x_1} f(x) + c.\nabla_{x_1} B(x) = 2x_1 + c.\left(-\frac{1}{x_1 - 1} + \frac{1}{-x_1 + 3}\right)$$

w.r.t to x_2:

$$\nabla_{x_2} f(x) + c.\nabla_{x_2} B(x) = 2x_2 + c.\left(-\frac{1}{x_2 + 3} + \frac{1}{-x_2 - 1}\right)$$

Putting it equal to 0.

$$\begin{bmatrix} 2x_1 + c.\left(-\frac{1}{x_1 - 1} + \frac{1}{-x_1 + 3}\right) \\ 2x_2 + c.\left(-\frac{1}{x_2 + 3} + \frac{1}{-x_2 - 1}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 = -\frac{c}{2}\left(-\frac{1}{x_1 - 1} + \frac{1}{-x_1 + 3}\right) = -\frac{c}{2}\left(\frac{2x_1 - 4}{-x_1^2 + 4x_1 - 3}\right) = \frac{cx_1 - 2c}{x_1^2 - 4x_1 + 3}$$

$$x_1^3 - 4x_1^2 + 3x_1 = cx_1 - 2c$$

$$x_2 = -\frac{c}{2}\left(-\frac{1}{x_2 + 3} + \frac{1}{-x_2 - 1}\right) = -\frac{c}{2}\left(\frac{2x_2 + 4}{-x_2^2 - 4x_2 - 3}\right) = \frac{cx_2 + 2c}{x_2^2 + 4x_2 + 3}$$

$$x_2^3 + 4x_2^2 + 3x_2 = cx_2 + 2c$$

Purring c=0 in both equations

$$x_1^3 - 4x_1^2 + 3x_1 = 0$$

$$x_1(x_1^2 - 4x_1 + 3) = 0$$

$$x_1(x_1^2 - x_1 - 3x_1 + 3) = 0$$

$$x_1(x_1 - 1)(x_1 - 3) = 0$$

So

$$x_1 = 0, 1, 3$$

$$x_2^3 + 4x_2^2 + 3x_2 = 0$$

$$x_2(x_2^2 + 4x_2 + 3) = 0$$

$$x_2(x_2^2 + x_2 + 3x_2 + 3) = 0$$

$$x_2(x_2 + 1)(x_2 + 3) = 0$$

So

$$x_2 = 0, -1, -3$$

Ignoring $x_1, x_2 = (0,0)$ since it is not in the feasible region. Putting all other values in the function to see at which point the function is minimum:

For $x_1, x_2 = (1, -1)$:

$$f(x) = x_1^2 + x_2^2 = 1 + 1 = 2$$

For $x_1, x_2 = (1, -3)$:

$$f(x) = x_1^2 + x_2^2 = 1 + 9 = 10$$

For $x_1, x_2 = (3, -1)$:

$$f(x) = x_1^2 + x_2^2 = 9 + 1 = 10$$

For $x_1, x_2 = (3, -3)$:

$$f(x) = x_1^2 + x_2^2 = 9 + 9 = 18$$

Hence

$$\boldsymbol{x^* = (1, -1)}$$

Question 1: Applying the Barrier Method

$$\boldsymbol{minimze\ {x_1}^2 + {x_2}^2, subject\ to: x_1 - x_2 \leq 0, \qquad -x_2 \leq 0}$$

Standard form is given as

$$argmin_{x \in R} f(x),$$

$$subject\ to\ g_p(x) = 0\ \ for\ p = 1, \dots, \dots, P$$

$$and\ h_q(x) \leq 0\ \ for\ q = 1, \dots, \dots, Q$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad P = 0,\ \ so\ g_p(x)\ does\ not\ exist\ , \qquad Q = 2$$

So

$$h_1(x) = x_1 - x_2\ , \qquad h_2(x) = -x_2$$

OR

$$argmin_{x \in R} f(x), subject\ to\ Ax - a = 0\ and\ Bx - b \leq 0$$

Here,

$$f(x) = x_1^2 + x_2^2, \qquad A = 0, \qquad a = 0,$$

$$B = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The objective function using the log barrier function can be written as

$$tf(x) + B(x) = tf(x) + \left(-\sum_{q=1}^{Q} \log(-h_q(x))\right)$$

$$tf(x) + c.B(x) = tx_1^2 + tx_2^2 + (-\log(-x_1 + x_2) - \log(x_2))$$

Taking the derivative

w.r.t to x_1:

$$t\nabla_{x_1} f(x) + \nabla_{x_1} B(x) = 2tx_1 + \left(\frac{1}{-x_1 + x_2} - 0\right) = 2tx_1 + \frac{1}{-x_1 + x_2}$$

w.r.t to x_2:

$$t\nabla_{x_2} f(x) + \nabla_{x_2} B(x) = 2tx_2 + \left(-\frac{1}{-x_1+x_2} - \frac{1}{x_2}\right)$$

Computing Hessian

$$\begin{bmatrix} 2t + \frac{1}{(-x_1+x_2)^2} & -\frac{1}{(-x_1+x_2)^2} \\ \frac{1}{(-x_1+x_2)^2} & 2t + \left(\frac{1}{(-x_1+x_2)^2} + \frac{1}{x_2^2}\right) \end{bmatrix}$$

With

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}$$

$$\mathrm{t}\nabla f\left(x^{(0)}\right) + \nabla \mathrm{B}\left(x^{(0)}\right) = \begin{bmatrix} 2tx_1 + \frac{1}{-x_1+x_2} \\ 2tx_2 + \left(-\frac{1}{-x_1+x_2} - \frac{1}{x_2}\right) \end{bmatrix} = \begin{bmatrix} -10 + \frac{2}{3} \\ 20 + \left(-\frac{2}{3} - 1\right) \end{bmatrix} = \begin{bmatrix} -\frac{28}{3} \\ \frac{55}{3} \end{bmatrix}$$

and

$$\mathrm{t}\nabla^2 f\left(x^{(0)}\right) + \nabla^2 \mathrm{B}\left(x^{(0)}\right) = \begin{bmatrix} 2t + \frac{1}{(-x_1+x_2)^2} & -\frac{1}{(-x_1+x_2)^2} \\ \frac{1}{(-x_1+x_2)^2} & 2t + \left(\frac{1}{(-x_1+x_2)^2} + \frac{1}{x_2^2}\right) \end{bmatrix}$$

$$\mathrm{t}\nabla^2 f\left(x^{(0)}\right) + \nabla^2 \mathrm{B}\left(x^{(0)}\right) = \begin{bmatrix} 20 + \frac{4}{9} & -\frac{4}{9} \\ \frac{4}{9} & 20 + \left(\frac{4}{9} + 1\right) \end{bmatrix} = \begin{bmatrix} \frac{184}{9} & -\frac{4}{9} \\ \frac{4}{9} & \frac{193}{9} \end{bmatrix}$$

Newton Update step is given by:

$$x^{(t+1)} = x^{(t)} + \mu^{(t)} \Delta x^{(t)}, where\ \Delta x^{(t)} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

$$\Delta x^{(0)} = -\begin{bmatrix} \frac{184}{9} & -\frac{4}{9} \\ \frac{4}{9} & \frac{193}{9} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{28}{3} \\ \frac{55}{3} \end{bmatrix} = \begin{bmatrix} 0.0489 & 0.001 \\ -0.001 & 0.0466 \end{bmatrix} \begin{bmatrix} -9.33 \\ 18.33 \end{bmatrix} = \begin{bmatrix} 0.4377 \\ -0.864 \end{bmatrix}$$

Now

$$x^{(1)} = \begin{bmatrix} -0.5 \\ 1 \end{bmatrix} + 0.05 \begin{bmatrix} 0.4377 \\ -0.864 \end{bmatrix} = \begin{bmatrix} -0.4781 \\ 0.9568 \end{bmatrix}$$

Iteration 2:

$$\mathrm{t}\nabla f\left(x^{(1)}\right) + \nabla \mathrm{B}\left(x^{(1)}\right) = \begin{bmatrix} 2tx_1 + \frac{1}{-x_1+x_2} \\ 2tx_2 + \left(-\frac{1}{-x_1+x_2} - \frac{1}{x_2}\right) \end{bmatrix} = \begin{bmatrix} -9.562 + 0.6969 \\ 19.136 - 1.742 \end{bmatrix} = \begin{bmatrix} -8.865 \\ 17.4 \end{bmatrix}$$

and

$$\mathrm{t}\nabla^2 f\left(x^{(1)}\right) + \nabla^2 \mathrm{B}\left(x^{(1)}\right) = \begin{bmatrix} 2t + \dfrac{1}{(-x_1 + x_2)^2} & -\dfrac{1}{(-x_1 + x_2)^2} \\ \dfrac{1}{(-x_1 + x_2)^2} & 2t + \left(\dfrac{1}{(-x_1 + x_2)^2} + \dfrac{1}{x_2^2}\right) \end{bmatrix}$$

$$\mathrm{t}\nabla^2 f\left(x^{(1)}\right) + \nabla^2 \mathrm{B}\left(x^{(1)}\right) = \begin{bmatrix} 20 + 0.4857 & -0.4857 \\ 0.4857 & 20 + 1.578 \end{bmatrix} = \begin{bmatrix} 20.4857 & -0.4857 \\ 0.4857 & 21.578 \end{bmatrix}$$

$$\Delta x^{(1)} = -\begin{bmatrix} 20.4857 & -0.4857 \\ 0.4857 & 21.578 \end{bmatrix}^{-1} \begin{bmatrix} -8.865 \\ 17.4 \end{bmatrix} = \begin{bmatrix} 0.0488 & 0.001 \\ -0.001 & 0.0463 \end{bmatrix} \begin{bmatrix} -8.865 \\ 17.4 \end{bmatrix} = \begin{bmatrix} 0.4134 \\ -0.8157 \end{bmatrix}$$

Now

$$x^{(2)} = \begin{bmatrix} -0.4781 \\ 0.9568 \end{bmatrix} + 0.05 \begin{bmatrix} 0.4134 \\ -0.8157 \end{bmatrix} = \begin{bmatrix} -0.457 \\ 0.916 \end{bmatrix}$$

x is changing very slowly, so I don’t know how many iterations will be required.