# Modern Optimization Techniques – Group 01

# Exercise Sheet 08


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

## Semester 2 MSc. Data Analytics


### Question 1: Inequality Constrained Minimization Problems

**(a). For inequality constrained optimization problem, can we find the search direction in active set method without using gradient projection? If so, please explain one method. Given function to be minimized is quadratic and the constraints are affine.**

We can find the search direction using one of the following active set methods (Luenberger & Ye, 2008):

- **Reduced Gradient Method**

  Given the problem

  $$minimze\ f(x), subject\ to\colon Ax = b, x \geq 0\ and\ h(x) = 0, c \leq x \leq d$$

  We divide the set of variables $x$ in to $(y, z)$, where $y$ and $z$ are termed as vectors of dependent and independent variables respectively. Then the reduced gradient of the new objective function $f(y, z)$ is computed with respect to $z$ and we move along the direction of the negative reduced gradient.

- **Convex Simplex Method**

  The Convex Simplex Method is a variation of the Reduced Gradient Method such that instead of moving all of the independent variables in the direction of the negative reduced gradient, only one variable is changed at a time. This variable is selected in a similar manner as the ordinary simplex method.

**(b).** $\boldsymbol{A} = \begin{bmatrix} \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \mathbf{-1} & \mathbf{2} & \mathbf{3} \end{bmatrix}$

The Null Space of a matrix is a set of all vectors $\boldsymbol{x}$ such that

$$\boldsymbol{Ax = 0}$$

The Nullity is defined as the number of all these vectors $\boldsymbol{x}$ that exist in the Null Space. The Rank-Nullity Theorem states that

Nullity of A + Rank of A = Number of columns of A

Rank of A is the number of independent rows of the matrix A i.e. we cannot get a row of zeros by subtracting any row from any other row. This is true for the given matrix, hence the rank is 2. Thus

Nullity of A + 2 = 3, i.e. Nullity of A = 1. This means that there is one vector in the null space of matrix A. Let's compute that:

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 7 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 14 & 0 & -2 \\ 0 & 7 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

Hence

$$14x_1 - 2x_3 = 0 \implies x_1 = \frac{x_3}{7} = 0.14x_3$$
$$7x_2 + 10x_3 = 0 \implies x_2 = -\frac{10x_3}{7} = -1.428x_3$$

Now

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1x_3}{7} \\ -\frac{10x_3}{7} \\ x_3 \end{bmatrix} = 0$$

$x_3$ can have any value. So putting it equal to 1 gives

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{7} \\ -\frac{10}{7} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Hence, proved.

**In gradient projection method, we compute a projection onto the null space of the equality constraints matrix A. Then this projection is multiplied with the gradient so that the gradient does not move away from the constraints during optimization.**

## Question 2: Gradient Projection Method

### (a). $\boldsymbol{minimze\ x_1{}^2 + x_2{}^2, subject\ to: x_1 - 2x_2 = 2}$

Here, we do not have inequality constraints. So,

$$\hat{A} = A = [1 \quad -2], \qquad \hat{a} = a = 2 \qquad and\ \hat{P} = P = 1$$

$$\hat{F} = I - \hat{A}^T\left(\hat{A}\hat{A}^T\right)^{-1}\hat{A} = I - \begin{bmatrix} 1 \\ -2 \end{bmatrix}\left(\begin{bmatrix} 1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & -2 \end{bmatrix}$$

$$\hat{F} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -2 \end{bmatrix}\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}$$

$$\Delta x^{(k-1)} = -\hat{F}^T \nabla f\left(x^{(k-1)}\right)$$

$$\nabla f\left(x^{(k-1)}\right) = \begin{bmatrix} 2x_1^{(k-1)} \\ 2x_2^{(k-1)} \end{bmatrix}$$

For $k = 1$

$$\nabla f\left(x^{(0)}\right) = \begin{bmatrix} 2x_1^{(0)} \\ 2x_2^{(0)} \end{bmatrix} = \begin{bmatrix} 2(2) \\ 2(0) \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\Delta x^{(0)} = -\hat{F}^T \nabla f\left(x^{(0)}\right) = -\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{16}{5} \\ -\frac{8}{5} \end{bmatrix}$$

Since there are no inequalities i.e. $|Q| = 0$, we use the normal backtracking line search, starting with $\mu = 1$

$$\Delta f^0 = \alpha \nabla f\left(x^{(0)}\right)^T \Delta x^{(0)} = 0.4\begin{bmatrix} 4 & 0 \end{bmatrix}\begin{bmatrix} -\frac{16}{5} \\ -\frac{8}{5} \end{bmatrix} = -5.12$$

Checking for condition

$$f\left(x^{(0)} + \mu \Delta x^{(0)}\right) > f\left(x^{(0)}\right) + \mu \Delta f^0$$

$$f\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix} + 1\begin{bmatrix} -\frac{16}{5} \\ -\frac{8}{5} \end{bmatrix}\right) > f\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) + 1(-5.12)$$

$$f\left(\begin{bmatrix} -1.2 \\ -1.6 \end{bmatrix}\right) > f\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) - 5.12$$

$$(-1.2)^2 + (-1.6)^2 > 2^2 + 0^2 - 5.12$$

$$4 > -1.12$$

So

$$\mu = \beta\mu = 0.5(1) = 0.5$$

Again

$$f\left(x^{(0)} + \mu \Delta x^{(0)}\right) > f\left(x^{(0)}\right) + \mu \Delta f^0$$

$$f\left(\begin{bmatrix}2\\0\end{bmatrix}+0.5\begin{bmatrix}-\frac{16}{5}\\-\frac{8}{5}\end{bmatrix}\right)>f\left(\begin{bmatrix}2\\0\end{bmatrix}\right)+0.5(-5.12)$$

$$f\left(\begin{bmatrix}0.4\\-0.8\end{bmatrix}\right)>f\left(\begin{bmatrix}2\\0\end{bmatrix}\right)-2.56$$

$$(0.4)^2+(-0.8)^2>2^2+0^2-2.56$$

$$0.8 \ngtr 1.44$$

Next,

$$x^{(1)}=x^{(0)}+\mu\Delta x^{(0)}=\begin{bmatrix}2\\0\end{bmatrix}+(0.5)\begin{bmatrix}-\frac{16}{5}\\-\frac{8}{5}\end{bmatrix}=\begin{bmatrix}\frac{2}{5}\\-\frac{4}{5}\end{bmatrix}=\begin{bmatrix}0.4\\-0.8\end{bmatrix}$$

For $k=2$

$$\nabla f\left(x^{(1)}\right)=\begin{bmatrix}2x_1^{(1)}\\2x_2^{(1)}\end{bmatrix}=\begin{bmatrix}2(0.4)\\2(-0.8)\end{bmatrix}=\begin{bmatrix}0.8\\-1.6\end{bmatrix}$$

$$\Delta x^{(1)}=-\hat{F}^T\nabla f\left(x^{(1)}\right)=-\begin{bmatrix}\frac{4}{5}&\frac{2}{5}\\\frac{2}{5}&\frac{1}{5}\end{bmatrix}\begin{bmatrix}0.8\\-1.6\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}, Hence\ converged$$

$$\boldsymbol{x}^*=\begin{bmatrix}\mathbf{0.4}\\\mathbf{-0.8}\end{bmatrix}$$

**(b).** $\boldsymbol{minimze\ x_1{}^2+x_2{}^2, subject\ to: -2x_1+x_2=0.5,\ x_1+x_2\leq 1}$

Checking if the inequality constraint is active with $x^{(0)}=\begin{bmatrix}0\\0.5\end{bmatrix}$

$$x_1+x_2\leq 1$$

$$0+0.5\leq 1$$

$$0.5\leq 1$$

Since it is not active which means $|Q|=0$ i.e. we will only use the equality constraint for solving. Therefore:

$$\hat{A}=A=[-2\quad 1],\qquad \hat{a}=a=2\qquad and\ \hat{P}=P=1$$

$$\hat{F}=I-\hat{A}^T\left(\hat{A}\hat{A}^T\right)^{-1}\hat{A}=I-\begin{bmatrix}-2\\1\end{bmatrix}\left([-2\quad 1]\begin{bmatrix}-2\\1\end{bmatrix}\right)^{-1}[-2\quad 1]$$

$$\hat{F} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \end{bmatrix} \begin{bmatrix} -\frac{2}{5} & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$$

$$\Delta x^{(k-1)} = -\hat{F}^T \nabla f\left(x^{(k-1)}\right)$$

$$\nabla f\left(x^{(k-1)}\right) = \begin{bmatrix} 2x_1^{(k-1)} \\ 2x_2^{(k-1)} \end{bmatrix}$$

For $k = 1$

$$\nabla f\left(x^{(0)}\right) = \begin{bmatrix} 2x_1^{(0)} \\ 2x_2^{(0)} \end{bmatrix} = \begin{bmatrix} 2(0) \\ 2(0.5) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Delta x^{(0)} = -\hat{F}^T \nabla f\left(x^{(0)}\right) = -\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ -\frac{4}{5} \end{bmatrix}$$

Since there are no inequalities i.e. $|Q| = 0$, we use the normal backtracking line search, starting with $\mu = 1$

$$\Delta f^0 = \alpha \nabla f\left(x^{(0)}\right)^T \Delta x^{(0)} = 0.4 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{2}{5} \\ -\frac{4}{5} \end{bmatrix} = -0.32$$

Checking for condition

$$f\left(x^{(0)} + \mu \Delta x^{(0)}\right) > f\left(x^{(0)}\right) + \mu \Delta f^0$$

$$f\left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix} + 1 \begin{bmatrix} -\frac{2}{5} \\ -\frac{4}{5} \end{bmatrix}\right) > f\left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix}\right) + 1(-0.32)$$

$$f\left(\begin{bmatrix} -0.4 \\ -0.3 \end{bmatrix}\right) > f\left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix}\right) - 0.32$$

$$(-0.4)^2 + (-0.3)^2 > 0^2 + 0.5^2 - 0.32$$

$$0.25 > -0.07$$

So

$$\mu = \beta \mu = 0.5(1) = 0.5$$

Again

$$f\left(x^{(0)} + \mu \Delta x^{(0)}\right) > f\left(x^{(0)}\right) + \mu \Delta f^0$$

$$f\left(\begin{bmatrix}0\\0.5\end{bmatrix}+0.5\begin{bmatrix}-\frac{2}{5}\\-\frac{4}{5}\end{bmatrix}\right)>f\left(\begin{bmatrix}0\\0.5\end{bmatrix}\right)+0.5(-0.32)$$

$$f\left(\begin{bmatrix}-0.2\\0.1\end{bmatrix}\right)>f\left(\begin{bmatrix}0\\0.5\end{bmatrix}\right)-0.16$$

$$(-0.2)^2+(0.1)^2>0^2+0.5^2-0.16$$

$$0.05 \ngtr 0.09$$

Next,

$$x^{(1)}=x^{(0)}+\mu\Delta x^{(0)}=\begin{bmatrix}0\\0.5\end{bmatrix}+(0.5)\begin{bmatrix}-\frac{2}{5}\\-\frac{4}{5}\end{bmatrix}=\begin{bmatrix}-0.2\\0.1\end{bmatrix}$$

For $k=2$

$$\nabla f\left(x^{(1)}\right)=\begin{bmatrix}2x_1^{(1)}\\2x_2^{(1)}\end{bmatrix}=\begin{bmatrix}2(-0.2)\\2(0.1)\end{bmatrix}=\begin{bmatrix}-0.4\\0.2\end{bmatrix}$$

$$\Delta x^{(1)}=-\hat{F}^T\nabla f\left(x^{(1)}\right)=-\begin{bmatrix}\frac{1}{5}&\frac{2}{5}\\\frac{2}{5}&\frac{4}{5}\end{bmatrix}\begin{bmatrix}-0.4\\0.2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}, Hence\ converged$$

$$\boldsymbol{x}^*=\begin{bmatrix}\mathbf{-0.2}\\\mathbf{0.1}\end{bmatrix}$$

## References


Luenberger, G. D., & Ye, Y. (2008). *Linear and Nonlinear Programming* (Fourth ed.). Springer.