# Modern Optimization Techniques – Group 01

# Exercise Sheet 07


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

## Semester 2 MSc. Data Analytics


## Question 1: Constrained Minimization: Primal and Dual problems

**(a).** $\boldsymbol{minimze\ \ f_0(x_1, x_2) = x_1{}^2 + x_2{}^2, subject\ to\ h_1(x_1, x_2) = x_1 + 2x_2 = 3}$

The plot is shown on the right.
KKT conditions are given by

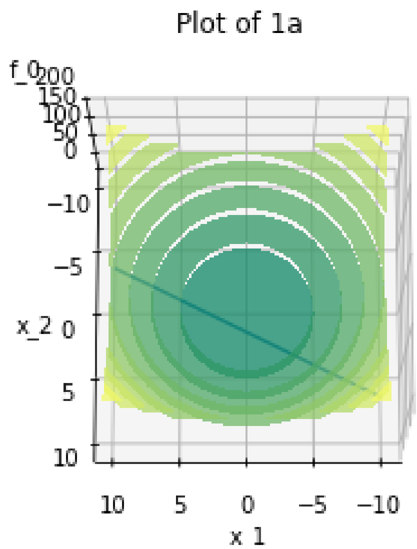
$$1. Primal\ feasibility: \qquad Ax = a$$

$$2. Stationarity: \qquad \nabla f(x) + A^T v^* = 0$$

i.e.

$$1. Primal\ feasibility: \qquad x_1 + 2x_2 = 3$$

$$2. Stationarity: \qquad \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} v = 0$$




Simplifying

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

Converting to row echelon form

$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 4 & -1 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & -5 & 6 \end{bmatrix}$$

Thus

$$2x_1 + 0x_2 + v = 0$$
$$0x_1 + 2x_2 + 2v = 0$$
$$0x_1 + 0x_2 - 5v = 6$$

Hence,

$$v = -\frac{6}{5}$$

$$2x_2 + 2\left(-\frac{6}{5}\right) = 0 \implies 2x_2 = \frac{12}{5} \implies x_2 = \frac{6}{5}$$

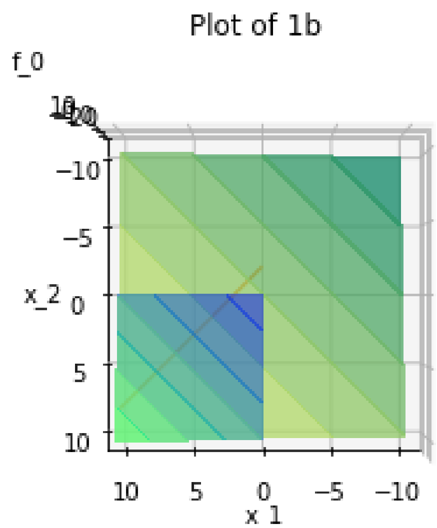
$$2x_1 + \left(-\frac{6}{5}\right) = 0 \implies 2x_1 = \frac{6}{5} \implies x_1 = \frac{3}{5}$$

Thus

$$\boldsymbol{x^* = \left(\frac{3}{5}, \frac{6}{5}\right), and\ v^* = -\frac{6}{5}}$$

**(b).** $\boldsymbol{minimze\ \ f_0(x_1, x_2) = x_1 + x_2,}$
$\boldsymbol{subject\ to\ h_1(x_1, x_2) = x_1 - x_2 = 2,}$ $\quad$ $\boldsymbol{f_1(x_1, x_2) = x_1 \geq 0,}$ $\quad$ $\boldsymbol{f_2(x_1, x_2) = x_2 \geq 0}$

Plotting the function gives the graph on the right side.




Since $x_2 \geq 0$. For $x_2 = 0$

$$x_1 - x_2 = 2 \implies x_1 = 2$$

Hence, $x_1 \geq 2$. Therefore,

$$\boldsymbol{x^* = (2, 0)}$$

Now, let's try to write the dual problem. The Langrangian is given by

$$L(x, v, \lambda) = x + v^T(Ax - b) + \lambda^T(Cx - d)$$

$$L(x, v, \lambda) = x_1 + x_2 + v(x_1 - x_2 - 2) + \lambda_1(x_1) + \lambda_2(x_2)$$

Dual langrangian is given by

$$g(v, \lambda) = \inf_{x \epsilon D}(x_1 + x_2 + v(x_1 - x_2 - 2) + \lambda_1(x_1) + \lambda_1(x_2))$$

Minimizing $L\ over\ x$

$$\nabla_x L(x, v, \lambda) = 1 + A^T v + C^T \lambda = 0$$

**As can be seen, we cannot write $x$ in terms of $v$ and $\lambda$ and thus cannot substitute it in dual langrangian. Hence, we cannot compute the dual problem for a linear program as this one.**

Moreover, Professor Lars also mentioned in the duality lecture (slide 3) that there is no analytical solution possible for a constrained linear program. However, there are certain specialized algorithms like the Simplex Tableau to solve such problems.

**(c).** $\boldsymbol{minimze\ \ f_0(x_1, x_2) = {x_1}^2 + {x_2}^2,}$
$\boldsymbol{subject\ to\ f_1(x_1, x_2) = x_1 + x_2 \leq 1,}$ $\quad$ $\boldsymbol{h(x_1, x_2) = x_2 - 2x_1 = \frac{1}{2}}$

Langrangian is given by

$$L(x, v, \lambda) = x^T x + v^T(Ax - b) + \lambda^T(Cx - d)$$

$$L(x,\nu,\lambda) = {x_1}^2 + {x_2}^2 + \nu\left(x_2 - 2x_1 - \frac{1}{2}\right) + \lambda(x_1 + x_2 - 1)$$

Dual langrangian is given by

$$g(\nu,\lambda) = \inf_{x\epsilon D}\left({x_1}^2 + {x_2}^2 + \nu\left(x_2 - 2x_1 - \frac{1}{2}\right) + \lambda(x_1 + x_2 - 1)\right)$$

Minimizing $L\ over\ x$

$$\nabla_x L(x,\nu,\lambda) = 2x + A^T\nu + C^T\lambda = 0$$

Hence,

$$x = \frac{-A^T\nu - C^T\lambda}{2}$$

Substituting $x\ in\ L$, we get

$$g(\nu,\lambda) = \left(\frac{-\nu^T A - \lambda^T C}{2}\right)\left(\frac{-A^T\nu - C^T\lambda}{2}\right) + \nu^T A\left(\frac{-A^T\nu - C^T\lambda}{2}\right) - \nu^T b + \lambda^T C\left(\frac{-A^T\nu - C^T\lambda}{2}\right) - \lambda^T d$$

$$g(\nu,\lambda) = \left(\frac{\nu^T AA^T\nu + \nu^T AC^T\lambda + \lambda^T CA^T\nu + \lambda^T CC^T\lambda}{4}\right) + \left(\frac{-\nu^T AA^T\nu - \nu^T AC^T\lambda}{2}\right) - \nu^T b + \left(\frac{-\lambda^T CA^T\nu - \lambda^T CC^T\lambda}{2}\right) - \lambda^T d$$

$$g(\nu,\lambda) = \left(\frac{\nu^T AA^T\nu + \nu^T AC^T\lambda + \lambda^T CA^T\nu + \lambda^T CC^T\lambda - 2\nu^T AA^T\nu - 2\nu^T AC^T\lambda - 2\lambda^T CA^T\nu - 2\lambda^T CC^T\lambda}{4}\right) - \nu^T b - \lambda^T d$$

$$g(\nu,\lambda) = -\frac{1}{4}(\nu^T AA^T\nu + \nu^T AC^T\lambda + \lambda^T CA^T\nu + \lambda^T CC^T\lambda) - \nu^T b - \lambda^T d$$

We have

$$A = [-2 \quad 1], \qquad b = \frac{1}{2}$$

$$C = [1 \quad 1], \qquad d = 1$$

$$\nu^T = \nu, \qquad \lambda^T = \lambda$$

So

$$g(\nu,\lambda) = -\frac{1}{4}\left(\nu^2[-2 \quad 1]\begin{bmatrix}-2\\1\end{bmatrix} + \nu\lambda[-2 \quad 1]\begin{bmatrix}1\\1\end{bmatrix} + \nu\lambda[1 \quad 1]\begin{bmatrix}-2\\1\end{bmatrix} + \lambda^2[1 \quad 1]\begin{bmatrix}1\\1\end{bmatrix}\right) - \frac{1}{2}\nu - \lambda$$

$$g(\nu,\lambda) = -\frac{1}{4}(5\nu^2 - \nu\lambda - \nu\lambda + 2\lambda^2) - \frac{1}{2}\nu - \lambda$$

$$g(\nu,\lambda) = -\frac{5}{4}\nu^2 + \frac{1}{2}\nu\lambda - \frac{1}{2}\lambda^2 - \frac{1}{2}\nu - \lambda$$

The function is plotted on the next page. As can be seen, it is clearly a concave function. It can also be proved by showing that the Hessian is negative semi-definite as follows:

$$\nabla g(\nu, \lambda) = \begin{bmatrix} -\frac{5}{2}\nu + \frac{1}{2}\lambda - \frac{1}{2} \\ -\lambda + \frac{1}{2}\nu - 1 \end{bmatrix}$$

$$\nabla^2 g(\nu, \lambda) = \begin{bmatrix} -\frac{5}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix}$$

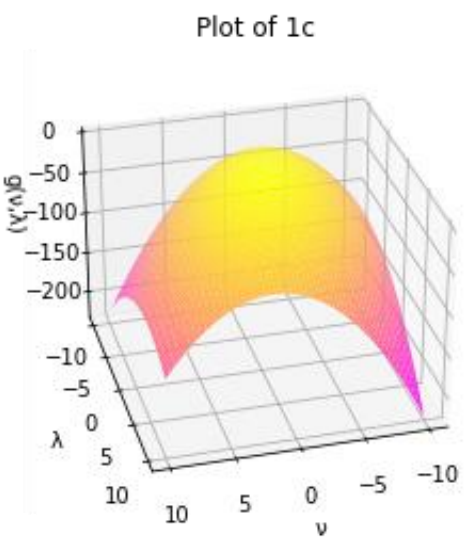
Now

$$[a \quad b] \begin{bmatrix} -\frac{5}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -\frac{5}{2}a + \frac{1}{2}b & \frac{1}{2}a - b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

$$= -\frac{5}{2}a^2 + \frac{1}{2}ab + \frac{1}{2}ab - b^2$$

$$= -\left(\frac{5}{2}a^2 + b^2 - ab\right)$$

Now, no matter what values of $a$ $and$ $b$ we chose, the term $\frac{5}{2}a^2 + b^2$ will always be positive and greater than the term $ab$. This means that the term inside the brackets i.e. $\frac{5}{2}a^2 + b^2 - ab$ will always be positive and the overall result will always be negative. Hence, the Hessian is negative semi-definite and the function is concave.




## Question 2: Newton Algorithm for Equality Constrained Problems

$$\boldsymbol{minimze\ \ f_0(x_1, x_2) = x_1{}^2 + x_2{}^2, subject\ to\ h_1(x_1, x_2) = x_1 + 2x_2 = 3}$$

Newton Algorithm for Equality Constrained Problems is computed by solving the following:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ v \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

Here,

$$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \qquad \nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad A = [1 \quad 2]$$

**Starting with $\boldsymbol{x^0 = (0, 1.5)}$**

Iteration 1:

$$\begin{bmatrix} \Delta x^0 \\ v^0 \end{bmatrix} = -\begin{bmatrix} \nabla^2 f(x^0) & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla f(x^0) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \Delta x^0 \\ v^0 \end{bmatrix} = -\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = -\begin{bmatrix} 0.4 & -0.2 & 0.2 \\ -0.2 & 0.1 & 0.4 \\ 0.2 & 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.6 \\ -0.3 \\ -1.2 \end{bmatrix}$$

Checking convergence using $\epsilon = 10^{-6}$

$$-\nabla f(x^0)^T \Delta x^0 = -[0 \quad 3] \begin{bmatrix} 0.6 \\ -0.3 \end{bmatrix} = 0.9 \nless \epsilon$$

So, we continue

$$x^1 = x^0 + \mu \Delta x^0 = \begin{bmatrix} 0 \\ 1.5 \end{bmatrix} + 1 \begin{bmatrix} 0.6 \\ -0.3 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 1.2 \end{bmatrix}$$

Iteration 2:

$$\begin{bmatrix} \Delta x^1 \\ v^1 \end{bmatrix} = -\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1.2 \\ 2.4 \\ 0 \end{bmatrix} = -\begin{bmatrix} 0.4 & -0.2 & 0.2 \\ -0.2 & 0.1 & 0.4 \\ 0.2 & 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} 1.2 \\ 2.4 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1.2 \end{bmatrix}$$

Checking convergence using $\epsilon = 10^{-6}$

$$-\nabla f(x^1)^T \Delta x^1 = -[1.2 \quad 2.4] \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 < \epsilon$$

Hence, **converged**.

Also, we can see that there will be no change in values even if we continue

$$x^2 = x^1 + \mu \Delta x^1 = \begin{bmatrix} 0.6 \\ 1.2 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 1.2 \end{bmatrix}$$

**Starting with $\boldsymbol{x^0 = (0, 5)}$**

Iteration 1:

$$\begin{bmatrix} \Delta x^0 \\ v^0 \end{bmatrix} = -\begin{bmatrix} \nabla^2 f(x^0) & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla f(x^0) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \Delta x^0 \\ v^0 \end{bmatrix} = -\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix} = -\begin{bmatrix} 0.4 & -0.2 & 0.2 \\ -0.2 & 0.1 & 0.4 \\ 0.2 & 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -4 \end{bmatrix}$$

Checking convergence using $\epsilon = 10^{-6}$

$$-\nabla f(x^0)^T \Delta x^0 = -[0 \quad 10] \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 10 \nless \epsilon$$

So, we continue

$$x^1 = x^0 + \mu \Delta x^0 = \begin{bmatrix} 0 \\ 5 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Iteration 2:

$$\begin{bmatrix} \Delta x^1 \\ v^1 \end{bmatrix} = -\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix} = -\begin{bmatrix} 0.4 & -0.2 & 0.2 \\ -0.2 & 0.1 & 0.4 \\ 0.2 & 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -4 \end{bmatrix}$$

Checking convergence using $\epsilon = 10^{-6}$

$$-\nabla f(x^1)^T \Delta x^1 = -[4 \quad 8] \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 < \epsilon$$

Hence, **converged**.

Also, we can see that there will be no change in values even if we continue

$$x^2 = x^1 + \mu \Delta x^1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

**Comments:**

In the first case, the algorithm converged and we had the correct solution as can be seen:

$$f_0(0,1.5) = 0^2 + 1.5^2 = 2.25, \qquad subject\ to\ h_1(0,1.5) = 0 + 2(1.5) = 3$$

**This is because we started at a feasible point since the equality constraint was satisfied.**

In the second case, the algorithm converged but we did not have the correct solution as can be seen:

$$f_0(0,1.5) = 0^2 + 5^2 = 25, \qquad subject\ to\ h_1(0,5) = 0 + 2(5) = 10 \neq 3$$

**This is because we started at an infeasible point since the equality constraint was not satisfied.** Here, we can use Newton Step at Infeasible Points Algorithm.