Optimalisering og regulering TTK4135 - Assignment 3

Ola Kristoffer Hoff

7^{th} February, 2023

Problem 1: LP and KKT conditions

Consider the following LP in standard form:

$$\min_{x} c^T x \quad s.t. \quad Ax = b, \quad x \geq 0,\ c \in \mathbb{R}^n,\ x \in \mathbb{R}^n,\ b \in \mathbb{R}^m \tag{1}$$

State the KKT conditions for this problem (copy them from your last homework or the textbook).

$$\begin{aligned}
\nabla \mathcal{L}(x^*, \lambda^*, s^*) = c - A^T \lambda^* - s^* = 0 \\
Ax^* - b = 0, \\
x^* \geq 0 \\
s^* \geq 0 \\
{s^*}^T x = 0
\end{aligned}$$

a Show that the Newton direction (see p. 22) cannot be defined for problem (1).

The book defines Newton direction (N&W, page 22, 2.15) as:

$$p_k^N = -(\nabla^2 f_k)^{-1} \nabla f_k$$

In our LP we have that: $f_k = c^T x$. This is a function of degree one, meaning the double derivative will be equal to zero:

$$\begin{aligned}
f_k &= c^T x \\
\nabla f_k &= c \\
\nabla^2 f_k &= 0
\end{aligned}$$

This will then give us the inverse of zero in the first half, in other words a divide by zero, hence it cannot be defined.

$$\begin{aligned}
p_k^N &= -(\nabla^2 f_k)^{-1} \nabla f_k \\
p_k^N &= -(0)^{-1} c \\
p_k^N &= -\frac{1}{0} c
\end{aligned}$$

b Show that (1) is a convex problem by using the definition of a convex function and the definition of a convex optimization problem.

There are three conditions for convexity (N&W, page 8):

- the objective function is convex,
- the equality constraint functions $c_i(\cdot),\ i \in E$, are linear, and

- the inequality constraint functions $c_i(\cdot)$, $i \in I$, are concave.

"The term "convex" can be applied both to sets and to functions. A set $S \in \mathbb{R}^n$ is a convex set if the straight line segment connecting any two points in S lies entirely inside S. Formally, for any two points $x \in S$ and $y \in S$, we have $\alpha x + (1-\alpha)y \in S$ for all $\alpha \in [0,1]$. The function f is a convex function if its domain S is a convex set and if for any two points x and y in S, the following property is satisfied:"(N&W, page 8)

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \ \ \forall \alpha \in [0,1] \tag{2}$$

The objective function is obviously convex since it's linear:

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &\leq \alpha f(x) + (1-\alpha)f(y) \\ c^T(\alpha x + (1-\alpha)y) &\leq \alpha c^T x + (1-\alpha)c^T y \\ \alpha c^T x + (1-\alpha)c^T y &\leq \alpha c^T x + (1-\alpha)c^T y \\ 0 &\leq 0 \end{aligned}$$

Note it's also concave, since it is strictly equal. This makes sense since a straight line is right in the middle of "bending" to either a convex or concave shape.

The equality constraints have to be linear. These are in the standard form represented as: $Ax = b$ which is per definition a system of linear equations, hence the second condition is satisfied as well.

Lastly the inequality constraints must be concave. To check this we simply "flip" the direction of the inequality in equation (2).

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &\geq \alpha f(x) + (1-\alpha)f(y) \\ c &= x \\ \alpha x + (1-\alpha)y &\geq \alpha x + (1-\alpha)y \\ 0 &\geq 0 \end{aligned}$$

The constraint was linear, and as stated above, it's then concave too.

All three conditions are met and the problem is then convex.

c The dual problem for (1) is defined as

$$\underset{\lambda}{max}\ b^T\lambda \quad s.t. \quad A^T\lambda \leq c \tag{3}$$

Show that the KKT-conditions for the dual problem (3) equals the KKT-conditions for problem (1).

To show this we will first have to rewrite the dual problem to standard form:

$$\begin{aligned} \underset{\lambda}{max}\ b^T\lambda \quad s.t. \quad A^T\lambda \leq c \\ \underset{\lambda}{min}\ -b^T\lambda \quad s.t. \quad c - A^T\lambda \geq 0 \end{aligned}$$

We do this my negating the objective function turning the maximisation into an minimisation. Then we get the conditions rewritten to be greater than or equal to zero.

Then to define the KKT-conditions we need the Lagrangian.

$$\begin{aligned} \mathcal{L}(x,\lambda) &= f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x) \\ \bar{\mathcal{L}}(\hat{\lambda}, \hat{x}) &= -b^T\hat{\lambda} - \hat{x}^T(c - A^T\hat{\lambda}) \end{aligned}$$

Now that we have that we can use the derivative, with respect to λ, to find the KKT-conditions:

$$\nabla_\lambda \bar{\mathcal{L}}(\hat{\lambda}, \hat{x}) = -b + A\hat{x}$$

Using the definition for KKT-conditions for the dual problem from the book (N&W, page 346):

$$\begin{aligned}
\nabla f(\bar{x}) - \nabla c(\bar{x})\bar{\lambda} &= 0 \\
c(\bar{x}) &\geq 0 \\
\bar{\lambda} &\geq 0 \\
\bar{\lambda}_i c(\bar{x})_i &= 0, \quad i = 1, 2, \ldots, m
\end{aligned}$$

The first conditions is that the derivative above must be equal to zero. The second is that all constraints is greater than or equal to zero. Third, all Lagrangian multipliers are greater than or equal to zero. Fourth, all Lagrangian multipliers times constraints must be equal to zero.

$$\begin{aligned}
\nabla_\lambda \bar{\mathcal{L}}(\hat{\lambda}, \hat{x}) &= -b + A\hat{x} = 0 \\
A\hat{x} &= b \\
c - A^T\hat{\lambda} &\geq 0 \\
\hat{x} &\geq 0 \\
\hat{x}^T(c - A^T\hat{\lambda}) &= 0
\end{aligned}$$

As we can see be comparing these KKT-conditions with the ones at the start of the section, if we substitute: $x^* = c - A^T\bar{\lambda} \wedge s^* = \bar{x}$, we see that they are the same conditions.

d What is the relation between the optimal objective $c^T x^*$ of problem (1) and the optimal objective $b^T\lambda^*$ of problem (3)? (You do not have to derive the relation if you did so in the previous assignment.)

From the conclusion that the dual problem has the same KKT-conditions we can see the following (N&W, page 360):

$$\begin{aligned}
b^T\lambda^* &= (x^*)^T A^T \lambda^* \\
&= (x^*)^T(A^T\lambda^* - c) + c^T x^* \\
&\leq c^T x^* \\
&= b^T\hat{\lambda}
\end{aligned}$$

Therefore we can see that we have:

$$c^T x^* \geq b^T\lambda^*$$

e Define the term basic feasible point for problem (1).

It's hard to paraphrase definitions so here is the definition from the book:

"A vector x is a basic feasible point if it is feasible and if there exists a subset $\mathcal{B}$ of the index set $\{1, 2, ..., n\}$ such that

- $\mathcal{B}$ contains exactly m indices;
- $i \notin \mathcal{B} \Rightarrow x_i = 0$ (that i, the bound $x_i \geq 0$ can be inactive only if $i \in \mathcal{B}$);
- The $m \times m$ matrix B defined by

$$B = [A_i]_{i\in\mathcal{B}} \tag{4}$$

  is non-singular, where A_i is the ith column of A."(N&W, page 362)

f We always assume that A in (1) has full (row) rank (see page 362 in the textbook). What does this mean for satisfying the LICQ (Definition 12.4 in the textbook)?

The rank of a matrix is the number of rows (in this case) that are linearly independent. A full (row) rank matrix would then be a matrix where all rows are linearly independent. LICQ stands for Linear Independence Constraint Qualification, and is a check to see if all constraint gradients of the active set is linearly independent. Hence, we have a matrix of full rank, the LICQ is satisfied by definition.

Problem 2: LP

Two reactors, R_I and R_{II}, produce two products A and B. To make 1000 kg of A, 2 hours of R_I and 1 hour of R_{II} are required. To make 1000 kg of B, 1 hour of R_I and 3 hours of R_{II} are required. The order of R_I and R_{II} does not matter. R_I and R_{II} are available for 8 and 15 hours, respectively. The selling price of A is $\frac{3}{2}$ of the selling price of B (i.e., 50% higher). We want to maximize the total selling price of the two products.

a Formulate this problem as an LP in standard form.

The objective function must be the profit, which is the amount sold times the price.

$$
\begin{aligned}
f(x) &= A_{amount} * A_{price} + B_{amount} * B_{price} \\
f(x) &= x_1 * \frac{3}{2} + x_2 * 1 \\
f(x) &= 3x_1 + 2x_2 \\
f(x) &= c^T x \\
c^T x &= \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}
$$

Then we have some constraints, the total time used by R_I and R_{II} must be less than or equal to 8 and 15 respectively. Since they aren't strictly equal we must introduce new variables to convert the constraints into equality constraints.

$$
\begin{aligned}
c_1(x) &= R_{I\,A} * x_1 + R_{I\,B} * x_2 + x_3 = 8 \\
c_1(x) &= 2x_1 + 1x_2 + x_3 = 8
\end{aligned}
$$

$$
\begin{aligned}
c_2(x) &= R_{II\,A} * x_1 + R_{II\,B} * x_2 + x_4 = 15 \\
c_2(x) &= 1x_1 + 3x_2 + x_4 = 15
\end{aligned}
$$

$$
\begin{aligned}
Ax &= \begin{bmatrix} c_1(x) \\ c_2(x) \end{bmatrix} = \begin{bmatrix} 2x_1 & x_2 & x_3 & 0 \\ x_1 & 3x_2 & 0 & x_4 \end{bmatrix} \\
Ax &= \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix} = b
\end{aligned}
$$

We also have that $x \geq 0$ since we can't produce negative amounts of product.
We now have:

$$
\max_x c^T x = \begin{bmatrix} 3 & 2 & 0 & 0 \end{bmatrix} x, \quad \text{s.t. } Ax = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 8 \\ 15 \end{bmatrix} = b, \ \ x \geq 0
$$

This is a LP on standard form, but we want a minimisation problem, so we need to transform it.

$$
\max_x \ -c^T x = \begin{bmatrix} -3 & -2 & 0 & 0 \end{bmatrix} x, \quad \text{s.t. } Ax = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 8 \\ 15 \end{bmatrix} = b, \ \ x \geq 0
$$

b Make a contour plot (use the MATLAB functions contour and *meshgrid*) and sketch the constraints (i.e., use a pen for the constraints if you prefer).

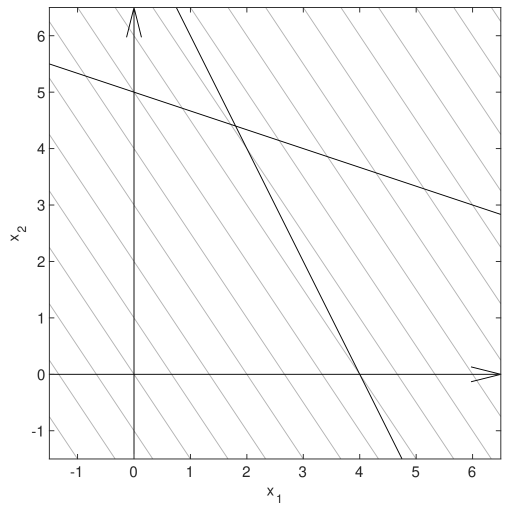

Figure 1: A contour plot of the objective function and the constraints in x_1 and x_2.

Not sure if we were to use the MATLAB scripts as a guide or an example of a potential solution, but I simply changed the file to show the contour and the constraints only.

c Calculate the production of A and B that maximizes the total selling price. Use the MATLAB function simplex published on Blackboard (an example of use is also published). Start the algorithm at $x_1 = x_2 = 0$. Is the solution at a point of intersection between the constraints? Are all constraints active? (DO NOT attach a printout of the algorithm output.)

The algorithm did three iterations starting at $x = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 15 \end{bmatrix}$ and ending in $x' = \begin{bmatrix} 1.8 \\ 4.4 \\ 0 \\ 0 \end{bmatrix}$

As we can see $x'_3 = x'_4 = 0$ meaning that our terms for turning the inequality into an equality is not being used, meaning the inequality is equal. In other words we have that the solution is at the intersection of the two constraints, meaning all constraints are active.

The total selling prince is: $c^T x = \begin{bmatrix} 3 & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1.8 \\ 4.4 \\ 0 \\ 0 \end{bmatrix} = 14.2$

d Mark all iterations on the plot made in b), as well as the iteration number.

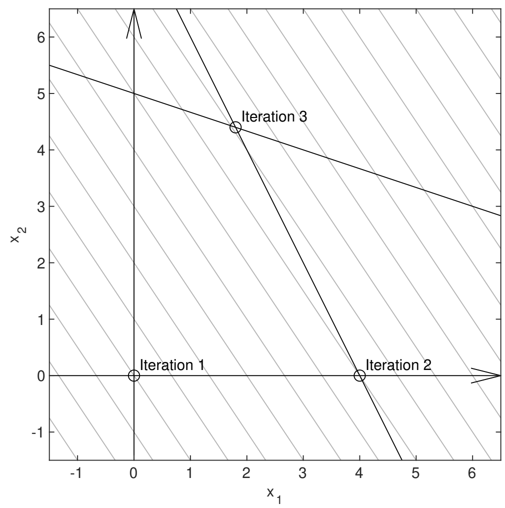

Figure 2: Plot from figure 1, with iterations from the Simplex algorithm marked.

Again I have just based myself on the example code. I did make some minor improvements to the placement of the text though.

e Look at the iterations on the plot and the algorithm output. Does everything agree with the theory in Chapter 13.3?

I cannot find anything that doesn't agree. It finds the solution, it approaches in the opposite direction of the gradient. I would say it agrees with the theory.

Problem 3: QP and KKT Conditions

A quadratic program (QP) can be formulated as

$$\begin{aligned} \min_{x}\ & q(x) = \frac{1}{2}x^T Gx + x^T c \\ s.t.\ & a_i^T x = b_i, \quad i \in \mathcal{E} \\ & a_i^T x \geq b_i, \quad i \in \mathcal{I} \end{aligned}$$

where G is a symmetric $n \times n$ matrix, $\mathcal{E}$ and $\mathcal{I}$ are finite set of indices and c, x and $\{a_i\}, i \in \mathcal{E} \cup \mathcal{I}$, are vectors in $\mathbb{R}^n$.

a Define the active set $\mathcal{A}(x^*)$ for problem 3.

The active set is the set of constraints (indices of the constraints) that are equal at the point x^*. Therefore, all equalities are in the set by definition and all inequalities which are equal is in it too. This gives us the following definition:

$$\mathcal{A}(x^*) = \mathcal{E} \cup \{i \in \mathcal{I} | a_i^T x^* = b_i\}$$

b Derive the KKT-conditions for problem 3, using the active set in the formulation.

As always we need the Lagrangian to start with:

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(x)$$
$$\mathcal{L}(x,\lambda) = \frac{1}{2}x^T G x + x^T c - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i (a_i^T x - b_i)$$

We can start by looking at the gradient of the Lagrangian with respect to x and the active set, $\mathcal{A}$, which should be equal to zero. This is the stationarity constraint. Since the non-active constraints per definition have zero multipliers we see they null out and we can just look at the remaining active set.

$$\nabla\mathcal{L}(x^*,\lambda^*) = Gx^* + c - \sum_{i\in\mathcal{A}^*} \lambda_i a_i = 0$$

For the primal feasibility constraints we can say that in terms of the active set it must be:

$$a_i^T x^* = b_i, \;\; i \in \mathcal{A}(x^*)$$

For all points in the active set this must hold per definition. The rest of the constraints, $\mathcal{I}\backslash\mathcal{A}(x^*)$, must then be greater than (or equal).

$$a_i^T x^* \geq b_i, \;\; i \in \mathcal{I}\backslash\mathcal{A}(x^*)$$

Then there is the dual feasibility conditions. The Lagrangian multipliers must be greater than or equal to zero for all points in the active set and all inequalities.

$$\lambda_i^* \geq 0, \;\; i \in \mathcal{A}(x^*) \cap \mathcal{I}$$

The complementary condition that states that all Lagrangian multipliers times their constraint must be equal to zero is enforced by the stationarity condition.

Hence, the KKT-conditions for the quadratic problem is:

$$\begin{aligned}
\nabla\mathcal{L}(x^*,\lambda^*) = Gx^* + c - \sum_{i\in\mathcal{A}^*} \lambda_i a_i &= 0 \\
a_i^T x^* &= b_i, \;\; i \in \mathcal{A}(x^*) \\
a_i^T x^* &\geq b_i, \;\; i \in \mathcal{I}\backslash\mathcal{A}(x^*) \\
\lambda_i^* &\geq 0, \;\; i \in \mathcal{A}(x^*) \cap \mathcal{I}
\end{aligned}$$