Norwegian University of Science and Technology
Department of Engineering Cybernetics

Exercise 3
Solution

Problem 1 (35 %) LP and KKT conditions (Exam August 2000)

We consider the following LP on standard form:

$$\min_x c^\top x \qquad \text{s.t.} \qquad Ax = b, \quad x \geq 0 \tag{1}$$

with $c,\ x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. We know that the KKT conditions for this problem are

$$A^\top \lambda^* + s^* = c \tag{2a}$$
$$Ax^* = b \tag{2b}$$
$$x^* \geq 0 \tag{2c}$$
$$s^* \geq 0 \tag{2d}$$
$$s_i^* x_i^* = 0, \quad i = 1, \dots, n \tag{2e}$$

a The Newton direction is $p_k^N = (-\nabla^2 f_k)^{-1} \nabla f_k$ (see equation (2.15) in the textbook). Here, $f(x) = c^\top x$, so that $\nabla f_k = c$ and $\nabla^2 f_k = 0$. Hence, $(-\nabla^2 f_k)^{-1}$ does not exist, and $p_k^N = (-\nabla^2 f_k)^{-1} \nabla f_k$ is therefore not defined.

b For an optimization problem to be convex, the following must hold:

- the objective function is convex,
- the equality constraint functions $c_i(\cdot),\ i \in \mathcal{E}$ are linear, and
- the inequality constraint functions $c_i(\cdot),\ i \in \mathcal{I}$ are concave.

(See page 8 in the textbook.)

Using the definition of convexity (equation (1.4) in the textbook) on the objective function, we have

$$\begin{aligned}\alpha f(x_1) + (1-\alpha) f(x_2) &= \alpha c^\top x_1 + (1-\alpha) c^\top x_2 \\ &= c^\top (\alpha x_1 + (1-\alpha) x_2) \\ &= f(\alpha x_1 + (1-\alpha) x_2)\end{aligned} \tag{3}$$

which shows the convexity of the objective function. The equality constraint $Ax = b$ is linear, which means the second requirement is satisfied. The inequality constraint is $x \geq 0$, so $c(x) = x$ must be a concave function if the problem is to be convex. Using the definition of concavity (the opposite of convexity), we have that

$$c\big(\alpha x + (1-\alpha) y\big) = \alpha x + (1-\alpha) y \tag{4a}$$

and that

$$\alpha c\big(x\big) + (1-\alpha)c\big(y\big) = \alpha x + (1-\alpha)y \tag{4b}$$

That is, the function is both convex and concave at the same time (this is true for all linear functions). Hence, problem (1) is a convex optimization problem since all three requirements are satisfied.

c The dual problem for (1) is defined as

$$\max_{\lambda} b^\top \lambda \qquad \text{s.t.} \qquad A^\top \lambda \leq c \tag{5}$$

We rewrite the dual problem as

$$\min_{\lambda} -b^\top \lambda \qquad \text{s.t.} \qquad c - A^\top \lambda \geq 0 \tag{6}$$

and define the Lagrangean for the problem as

$$\bar{\mathcal{L}}(\lambda, x) = -b^\top \lambda - x^\top (c - A^\top \lambda) \tag{7}$$

where $x \in \mathbb{R}^n$ are multipliers for the constraints. Differentiating with respect to λ and requiring the derivative to be zero gives

$$\nabla_\lambda \bar{\mathcal{L}} = -b + (x^\top A^\top)^\top = Ax - b = 0 \tag{8}$$

This, together with the general KKT conditions, gives the following KKT conditions for the dual problem:

$$Ax^* = b \tag{9a}$$
$$A^\top \lambda^* \leq c \tag{9b}$$
$$x^* \geq 0 \tag{9c}$$
$$x_i^*(c - A^\top \lambda^*)_i = 0, \quad i = 1, \dots, n \tag{9d}$$

Defining $s = c - A^\top \lambda$ shows that the KKT-conditions for the dual problem (5) equals the KKT-conditions for problem (1).

d The optimal objective $c^\top x^*$ of problem (1) and the optimal objective $b^\top \lambda^*$ of problem (5) have identical values.

e A *basic feasible point* x for problem (1) is defined by the following:

- A subset $\mathcal{B} \subseteq \{1, \dots, n\}$ can be defined as containing exactly m indices,
- $i \notin \mathcal{B} \Rightarrow x_i = 0$,
- the $m \times m$ matrix B defined by $B = [A_i]_{i \in \mathcal{B}}$, where A_i is the ith column of A, is nonsingular.

An equivalent definition is that a basic feasible point x has $n - m$ elements set to zero, and nonzero elements given by the constraint $Ax = b$.

f The equality constraints $Ax = b$ can be written

$$a_i^\top x = b_i, \qquad i \in \mathcal{E} \tag{10a}$$

or

$$c_i(x) = a_i^\top x - b_i = 0, \qquad i \in \mathcal{E} \tag{10b}$$

where $a_i^\top$ is the ith row of the matrix A. We then have that the gradients of the equality constraints are

$$\nabla c_i(x) = a_i, \qquad i \in \mathcal{E} \tag{11}$$

Since A has full (row) rank, all rows $a_i^\top$ are linearly independent. This means that all equality constraint gradients $\nabla c_i(x) = a_i$ are linearly independent.

Problem 2 (40 %) 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 This problem can be formulated as the following standard-form LP:

$$\begin{aligned} \min \quad & \begin{bmatrix} -3 & -2 & 0 & 0 \end{bmatrix} x \\ \text{s.t.} \quad & \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 8 \\ 15 \end{bmatrix} \\ & x \geq 0 \end{aligned} \tag{12}$$

b Figure 1 shows a contour plot with the constraints indicated in the space of x_1 and x_2.

c By modifying the example file provided to fit this problem, we find the following iteration sequence:

	Iteration number 1	2	3
x	$\begin{bmatrix} 0 \\ 0 \\ 8 \\ 15 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 0 \\ 0 \\ 11 \end{bmatrix}$	$\begin{bmatrix} 1.8 \\ 4.4 \\ 0 \\ 0 \end{bmatrix}$

From Figure 1 we see that the the solution is at the intersection between the availability constraints $2x_1 + x_2 \leq 8$ and $x_1 + 3x_2 \leq 15$. The non-negativity constraints are not active.

d The iterations along with iteration numbers are indicated in Figure 1.

e Looking at the algorithm output (the report given at each iteration) shows that the iteration sequence agrees with the theory in Chapter 13.3.

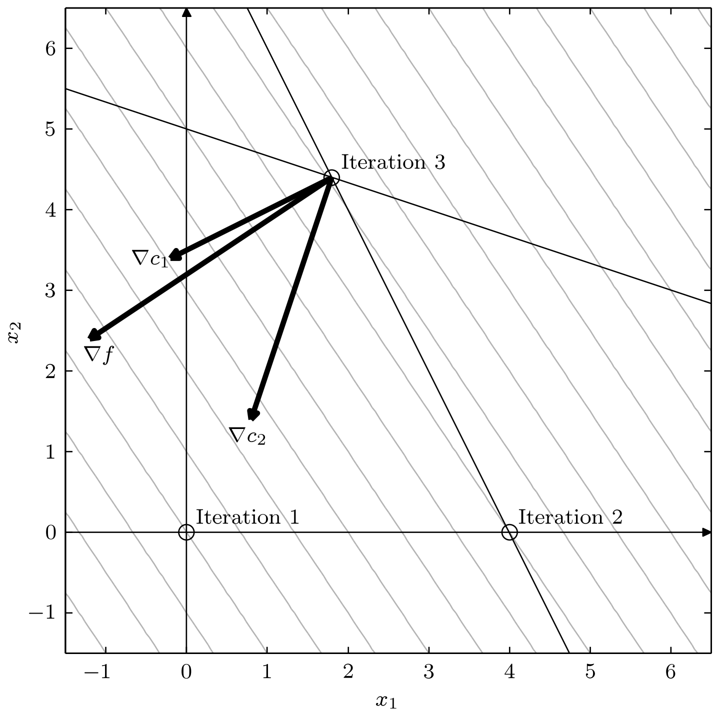

Figure 1: Contour plot with constraints for LP problem (12).

Problem 3 (25 %) QP and KKT Conditions (Exam May 2000)

A quadratic program (QP) can be formulated as

$$\min_{x} \quad q(x) = \frac{1}{2}x^\top Gx + x^\top c \tag{13a}$$
$$\text{s.t.} \quad a_i^\top x = b_i, \qquad i \in \mathcal{E} \tag{13b}$$
$$a_i^\top x \geq b_i, \qquad i \in \mathcal{I} \tag{13c}$$

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

a The active set $\mathcal{A}(x^*)$ for problem (13) is defined as the set of indices of the constraints for which equality holds at x^*. That is,

$$\mathcal{A}(x^*) = \{i \in \mathcal{E} \cup \mathcal{I} | a_i^\top x^* = b_i\} \tag{14}$$

b Defining the Lagrangian for problem (13) as

$$\begin{aligned}\mathcal{L}(x, \lambda) &= q(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x) \\ &= \frac{1}{2}x^\top Gx + x^\top c - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i (a_i^\top x - b_i)\end{aligned} \tag{15}$$

From the complementarity condition $\lambda_i^* c_i(x^*) = 0,\ i \in \mathcal{E} \cup \mathcal{I}$, we know that for constraints that are not active at the solution, the corresponding multipliers are zero. Hence, the derivative with respect x of the Lagrangian at the solution can be written as

$$\begin{aligned}\nabla_x \mathcal{L}(x^*, \lambda^*) &= \frac{1}{2}(Gx^* + G^\top x^*) + c - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* a_i, \quad G^\top = G \\ &= Gx^* + c - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* a_i \qquad (16)\end{aligned}$$

For all constraints i in the active set $(i \in \mathcal{A}(x^*))$,

$$a_i^\top x^* = b_i \qquad (17)$$

must hold at the solution. For inequality constraints i that are not active at the solution $(i \in \mathcal{I} \setminus \mathcal{A}(x^*))$,

$$a_i^\top x^* \geq b_i \qquad (18)$$

must hold. For active inequality constraints i $(i \in \mathcal{I} \cap \mathcal{A}(x^*))$, the multipliers must be non-negative. That is,

$$\lambda_i^* \geq 0 \qquad (19)$$

The KKT conditions for problem (13) can now be summarized as

$$Gx^* + c - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* a_i = 0 \qquad (20a)$$

$$a_i^\top x^* = b_i, \quad i \in \mathcal{A}(x^*) \qquad (20b)$$

$$a_i^\top x^* \geq b_i, \quad i \in \mathcal{I} \setminus \mathcal{A}(x^*) \qquad (20c)$$

$$\lambda_i^* \geq 0, \quad i \in \mathcal{I} \cap \mathcal{A}(x^*) \qquad (20d)$$