Norwegian University of Science and Technology
Department of Engineering Cybernetics

Exercise 4
Solution

Problem 1 (50 %) Quadratic Programming

a A quadratic program

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

is convex if the Hessian matrix G is positive semidefinite. If G is positive definite, the QP is strictly convex.

Convexity is important because in the case of convex QPs, the problem is often similar in difficulty to a linear program, and the problem has only one solution. A nonconvex QP can be more challenging due to the possibility of several local minima and stationary points.

b If the condition in Theorem 16.2 is changed from $Z^\top GZ > 0$ to $Z^\top GZ \geq 0$, all we can say about x^* is that it is *a (local) solution* of the equality constrained QP (16.3).

The proof is identical up until the equation

$$q(x) = \frac{1}{2}u^\top Z^\top GZu + q(x^*) \tag{2}$$

Now, since we have $Z^\top GZ \geq 0$, we can only conclude $q(x) \geq q(x^*)$, which shows that x^* is a local minimizer (not strict!).

c We will apply Algorithm 16.3 to the quadratic program with

$$G = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad c^\top = \begin{bmatrix} -2 & -5 \end{bmatrix} \tag{3a}$$
$$a_1^\top = \begin{bmatrix} 1 & -2 \end{bmatrix} \qquad a_2^\top = \begin{bmatrix} -1 & -2 \end{bmatrix} \tag{3b}$$
$$a_3^\top = \begin{bmatrix} -1 & 2 \end{bmatrix} \qquad a_4^\top = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{3c}$$
$$a_5^\top = \begin{bmatrix} 0 & 1 \end{bmatrix} \qquad b^\top = \begin{bmatrix} -2 & -6 & -2 & 0 & 0 \end{bmatrix} \tag{3d}$$

and $\{1, 2, 3, 4, 5\} \in \mathcal{I}$. Note that the objective function also has a constant element 7.25, but this has no effect on the solution. We start the algorithm at $x = [2, 0]^\top$, where constraints 3 and 5 are active. However, we set the initial working set $\mathcal{W}_0$ to only contain constraint 3.

Let $x_k,\ \lambda_k,\ p_k$ denote the variables $x,\ \lambda,\ p$ at iteration k; let A_k the matrix $[A_i]_{i\in\mathcal{W}_k}$, that is, containing the vectors a_i which are in the working set at iteration k.

Iteration $k = 0$

Since only constraint 3 is active, A_0 will contain a_3. To find the direction p, we solve the equality-constrained QP

$$\min_p \quad q(x) = \frac{1}{2}p^\top G p + g_0^\top p \tag{4a}$$

$$\text{s.t.} \quad a_3^\top p = 0 \tag{4b}$$

where $g_0 = Gx_0 + c$. (See equation (16.39) in the textbook.) From Chapter 16.1, we know that the solution to this problem can be found by solving the equation set

$$\begin{bmatrix} G & -A_0^\top \\ A_0 & 0 \end{bmatrix} \begin{bmatrix} p_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} -g_0 \\ 0 \end{bmatrix} \tag{5a}$$

which gives

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -2 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} p_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 0 \end{bmatrix} \tag{5b}$$

This equation set has the solution $p_0 = [0.2,\ 0.1]^\top$ and $\lambda_0 = -2.4$. With this direction we get

$$x_1 = x_0 + p_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 2.2 \\ 0.1 \end{bmatrix} \tag{6}$$

This point is feasible with respect to all constraints, so we do not need to find an α_0 at this iteration. Note that there also are no blocking constraints. Hence, the working set at the next iteration will be the same as in this iteration. We set $k = 1$ and proceed to the next iteration.

Iteration $k = 1$

We now solve the same QP as above, except that we now have $g_1 = Gx_1 + c = [2.4,\ -4.8]^\top$. This gives the equation set

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -2 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} -2.4 \\ 4.8 \\ 0 \end{bmatrix} \tag{7}$$

with the solution $p_1 = [0,\ 0]^\top$ and $\lambda_1 = -2.4$. Since $p_1 = 0$ and $\lambda_1 < 0$, we remove the constraint with the most negative multiplier from the working set. With only one constraint in the working set, the working set will now be empty. Furthermore, $x_2 = x_1 + 0 = x_1$. We set $k = 2$ and proceed to the next iteration.

Iteration $k = 2$

Now that we have an empty working set, so A_2 is an empty matrix. The direction p is then found from the unconstrained QP

$$\min_p \quad q(x) = \frac{1}{2}p^\top G p + g_2^\top p \tag{8}$$

where $g_2 = Gx_2 + c = [2.4,\ -4.8]^\top$. Solving the equation set

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} p_2 = \begin{bmatrix} -2.4 \\ 4.8 \end{bmatrix} \tag{9}$$

gives the solution $p_2 = [-1.2,\ 2.4]^\top$. (There are no multipliers due to the absence of constraints in the working set.) Note that the solution to the unconstrained problem is $p_2 = G^{-1}(-g_2)$, that $\nabla^2 q(x_2) = G$, and that $\nabla q(x_2) = Gx_2 + c$. Hence, we can write $p_2 = -(\nabla^2 q(x_2))^{-1}\nabla q(x_2)$, which is the Newton direction (see equation (2.15) in the textbook). Setting $x_3 = x_2 + p_2 = [1,\ 2.5]^\top$ gives an infeasible point (constraint 1 is violated), so a step-length parameter α_2 must be found. $a_i^\top p_2 < 0$ for $i = 1,\ 2$, and 4, so

$$\begin{aligned} \alpha_2 &= \min\left(1,\ \frac{b_1 - a_1^\top x_2}{a_1^\top p_2},\ \frac{b_2 - a_2^\top x_2}{a_2^\top p_2},\ \frac{b_4 - a_4^\top x_2}{a_4^\top p_2}\right) \\ &= \min\left(1,\ \frac{2}{3},\ 1,\ \frac{11}{6}\right) = \frac{2}{3} \end{aligned} \tag{10}$$

As the minimum value corresponds to corresponds to constraint 1 (a blocking constraint), this constraint is added to the working set. That is, $\mathcal{W}_3 = \{1\}$. x at the next iteration is then found from

$$x_3 = x_2 + \alpha_2 p_2 = \begin{bmatrix} 1.4 \\ 1.7 \end{bmatrix} \tag{11}$$

We set $k = 3$ and proceed to the next iteration.

Iteration $k = 3$

We now have $g_3 = Gx_3 + c = [0.8,\ -1.6]^\top$ and $A_3 = a_1$. Then, the KKT system that solves the QP for the direction p_3 becomes

$$\begin{bmatrix} G & -A_3^\top \\ A_3 & 0 \end{bmatrix} \begin{bmatrix} p_3 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} -g_3 \\ 0 \end{bmatrix} \tag{12a}$$

which gives

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 2 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} p_3 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 1.6 \\ 0 \end{bmatrix} \tag{12b}$$

The solution to this equation set is $p_3 = [0,\ 0]^\top$ and $\lambda_0 = 0.8$. Hence, we can conclude that we have found the solution. That is, $x^* = [1.4,\ 1.7]^\top$.

The problem and iteration sequence is illustrated in Figure 1.

d We write the problem in Example 16.4 as

$$\begin{aligned} \min_x \quad & q(x) = \frac{1}{2}x^\top G x + x^\top c && \text{(13a)} \\ \text{s.t.} \quad & Ax - b \geq 0 && \text{(13b)} \end{aligned}$$

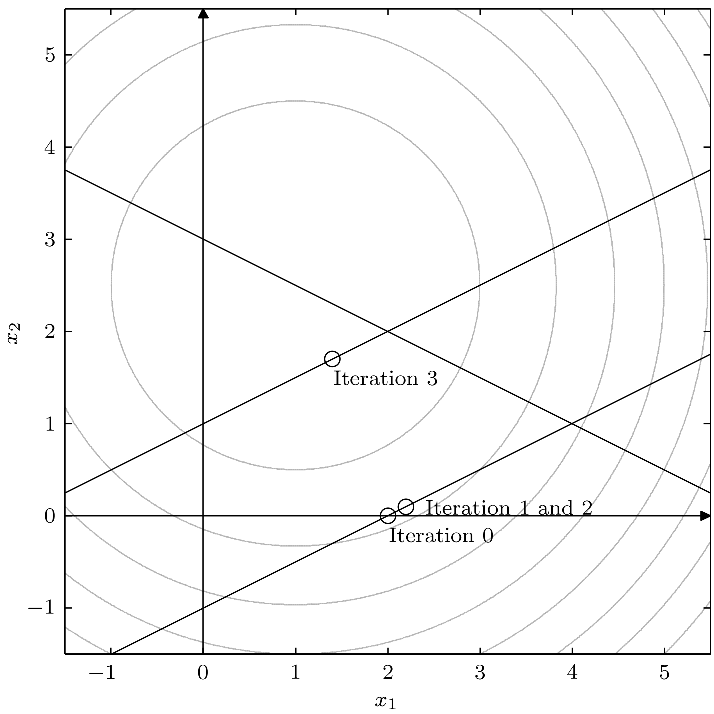

Figure 1: Contour plot with constraints and iterations for Problem 1.

where

$$G = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad c = \begin{bmatrix} -2 \\ -5 \end{bmatrix} \tag{14a}$$

$$A = \begin{bmatrix} 1 & -2 \\ -1 & -2 \\ -1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ -6 \\ -2 \\ 0 \\ 0 \end{bmatrix} \tag{14b}$$

The Lagrangian for this problem is

$$\mathcal{L}(x, \lambda) = \frac{1}{2} x^\top G x + x^\top c - \lambda^\top (Ax - b) \tag{15}$$

Hence, the dual objective is

$$f(\lambda) = \inf_x \mathcal{L}(x, \lambda) = \inf_x \frac{1}{2} x^\top G x + x^\top c - \lambda^\top (Ax - b) \tag{16}$$

Since $G > 0$ and the Lagrangian is a strictly convex quadratic function, the infimum with respect to x is where $\nabla_x \mathcal{L}(x, \lambda) = 0$. That is:

$$\begin{aligned} \nabla_x \mathcal{L}(x, \lambda) &= \frac{1}{2} G^\top x + \frac{1}{2} G x + c - (\lambda^\top A)^\top, \qquad G = G^\top \\ &= Gx + c - A^\top \lambda = 0 \end{aligned} \tag{17}$$

We can now write the the dual problem in three different forms. One form is obtained by writing x as $x = G^{-1}(A^\top\lambda - c)$ and substitute in the dual objective to get

$$\begin{aligned} f(\lambda) = \frac{1}{2}&(A^\top\lambda - c)^\top G^{-1}GG^{-1}(A^\top\lambda - c) \\ &+ (A^\top\lambda - c)^\top G^{-1}c - \lambda^\top(AG^{-1}(A^\top\lambda - c) - b) \end{aligned} \tag{18}$$

After some rearranging, we can write

$$\begin{aligned} f(\lambda) &= \frac{1}{2}(A^\top\lambda - c)^\top G^{-1}(A^\top\lambda - c) \\ &\quad + A^\top\lambda G^{-1}c - c^\top G^{-1}c \\ &\quad - \lambda^\top AG^{-1}A^\top\lambda + \lambda^\top AG^{-1}c + \lambda^\top b \\ \Rightarrow f(\lambda) &= \frac{1}{2}(A^\top\lambda - c)^\top G^{-1}(A^\top\lambda - c) \\ &\quad - (A^\top\lambda - c)^\top G^{-1}(A^\top\lambda - c) + \lambda^\top b \\ \Rightarrow f(\lambda) &= -\frac{1}{2}(A^\top\lambda - c)^\top G^{-1}(A^\top\lambda - c) + \lambda^\top b \end{aligned} \tag{19}$$

The dual problem can then be formulated as

$$\max_{\lambda} \quad f(\lambda) = -\frac{1}{2}(A^\top\lambda - c)^\top G^{-1}(A^\top\lambda - c) + \lambda^\top b \tag{20a}$$

$$\text{s.t.} \quad \lambda \geq 0 \tag{20b}$$

Alternatively, the dual problem can be formulated in both x and λ as

$$\max_{x,\lambda} \quad \frac{1}{2}x^\top Gx + x^\top c - \lambda^\top(Ax - b) \tag{21a}$$

$$\text{s.t.} \quad Gx + c - A^\top\lambda = 0 \tag{21b}$$

$$\lambda \geq 0 \tag{21c}$$

A third option is to reformulate the constraint as $(c - A^\top\lambda)^\top x = -x^\top Gx$, and substitute into the objective function to get

$$\max_{x,\lambda} \quad -\frac{1}{2}x^\top Gx + \lambda^\top b \tag{22a}$$

$$\text{s.t.} \quad Gx + c - A^\top\lambda = 0 \tag{22b}$$

$$\lambda \geq 0 \tag{22c}$$

e The dual optimization problem can be used to give an over-estimate of $q(\bar{x}) - q(x^*)$ (q is the primal objective), when x^* is not known. Any feasible $\bar{x}$ and any $\bar{\lambda} \geq 0$ will give $f(\bar{\lambda}) \leq q(x^*)$ (f is the dual objective). Therefore,

$$q(\bar{x}) - q(x^*) \leq q(\bar{x}) - f(\bar{\lambda}) \tag{23}$$

Problem 2 (50 %) Production Planning and Quadratic Programming

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. We want to maximize the total profit from the two products.

The profit now depends on the production rate:

- the profit from A is $3 - 0.4x_1$ per tonne produced,
- the profit from B is $2 - 0.2x_2$ per tonne produced,

where x_1 is the production of product A and x_2 is the production of product B (both in tonnes).

a The total profit from selling the two products is

$$p(x) = (3 - 0.4x_1)x_1 + (2 - 0.2x_2)x_2 \tag{24}$$

Since we want to maximize $p(x)$, we minimize the negative of $p(x)$. Hence, the objective function for the minimization problem is

$$\begin{aligned} f(x) = -p(x) = -(3 - 0.4x_1)x_1 - (2 - 0.2x_2)x_2 = \\ \frac{1}{2}x^\top \underbrace{\begin{bmatrix} 0.8 & 0 \\ 0 & 0.4 \end{bmatrix}}_{G} x + \underbrace{\begin{bmatrix} -3 & -2 \end{bmatrix}}_{c^\top} x \end{aligned} \tag{25}$$

Note that $G > 0$. The availability constraints can be formulated as $2x_1 + x_2 \leq 8$ and $x_1 + 3x_2 \leq 15$; nonnegative production rates are formulated $x_1 \geq 0$ and $x_2 \geq 0$. We then have the optimization problem

$$\min_x \quad q(x) = \frac{1}{2}x^\top G x + x^\top c \tag{26a}$$

$$\text{s.t.} \quad -2x_1 - x_2 \geq -8 \tag{26b}$$

$$-x_1 - 3x_2 \geq -15 \tag{26c}$$

$$x_1 \geq 0 \tag{26d}$$

$$x_2 \geq 0 \tag{26e}$$

b A contour plot with constraints and iterations indicated is illustrated in Figure 2. Code for producing a contour plot in MATLAB is attached at the end of this solution.

c The modifications that have to be made to the file qp_prodplan.m so that it solves the problem formulated in a) is attached at the end of this solution. We see in Figure 2 that as opposed to the linear case in exercise 3, the solution is not at a point of intersection between constraints. The solution and all iterations are indicated in Figure 2.

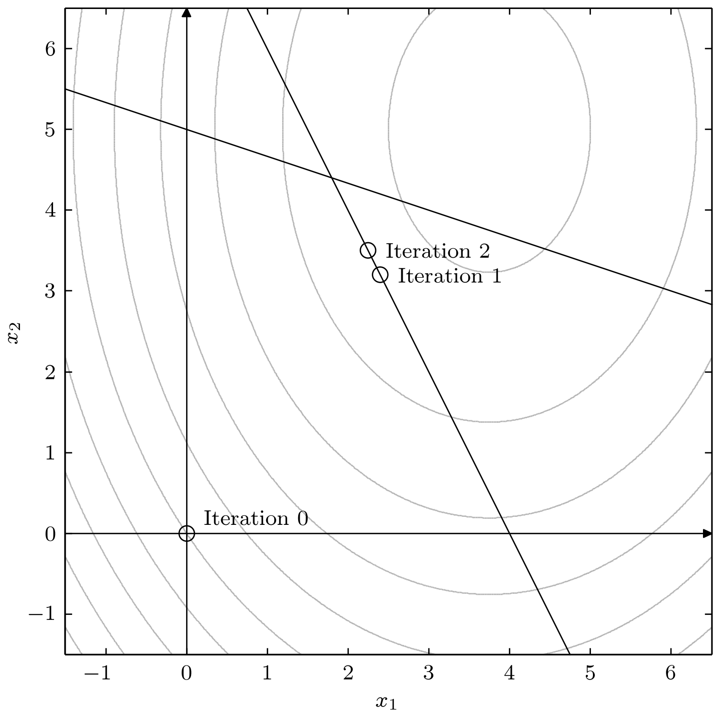

Figure 2: Contour plot with constraints and iterations for Problem 2.

d See Problem 1 c) and the textbook for an explanation of the active set algorithm.

e As mentioned, the solution is not found at a corner point of the feasible area. For linear programs, the solution is always at a point of intersection between constraints. For quadratic programs, the solution may be at any point in the feasible area, including the interior (where no constraints are active).

```
% Code for making a contour plot for Problem 2.
x1_l = -1.5; x1_h = 6.5;
x2_l = -1.5; x2_h = 6.5;
res = 0.01;
[x1, x2] = meshgrid(x1_l:res:x2_h, x1_l:res:x2_h);
f = -(3-0.4*x1).*x1 - (2-0.2*x2).*x2;
levels = (-12:2:8)';
[C, h] = contour(x1, x2, f, levels, 'Color', .7*[1 1 1]);
set(h, 'ShowText', 'on', 'LabelSpacing', 300); % For text labels
```

```
% Changes made to the file qp_prodplan.m
% Quadratic objective (MODIFY THESE)
G = [0.8  0   ;
     0    0.4]; % Remember the factor 1/2 in the objective
c = [-3 ; -2];
% Linear constraints (MODIFY THESE)
A = [2 1 ;
     1 3];
b = [8 ; 15];
```