TMA4180 Optimization Theory
Spring 2013

Exercise set 9

Norwegian University of Science and Technology
Department of Mathematical Sciences

Tutorial: Thursday 21 16:15-17:00 in Kjl 4.

$\boxed{1}$ (Exam May 2008)

a) What is the content of the duality theorem of linear programming?

Solution: The dual and primal problems have equivalent KKT equations, and variables and Lagrange multipliers switch place. The duality theorem states that if any of the problems are unbounded, the other is infeasible. Moreover, the optimal objective values are equal, and since one is a minimum and the other a maximum problem, objectives values for the two are separated by the optimal objective value on the real line.

b) Show that the following two problems are dual problems (A has full row rank):

$$
\begin{array}{cc}
(\mathcal{P}) & (\mathcal{D}) \\
\min_x c^{\mathrm{T}} x & \max_\lambda b^{\mathrm{T}} \lambda \\
Ax \geq b, x \geq 0 & A^{\mathrm{T}} \lambda \leq c, \lambda \geq 0.
\end{array}
$$

(*Hint*: Consider the KKT equations and start by using (λ, s) as Lagrange multipliers for $(\mathcal{P})$ and (x, u) for $(\mathcal{D})$.)

Solution: In order to establish the statement, we need to look at the KKT equations. For the primal problem $(\mathcal{P})$, using that $\mathcal{L}_{\mathcal{P}}(x, \lambda, s) = c^{\mathrm{T}} x - \lambda^{\mathrm{T}}(Ax - b) - s^{\mathrm{T}} x$, we obtain

$$
\begin{aligned}
\nabla_x \mathcal{L}_{\mathcal{P}} = c - A^{\mathrm{T}} \lambda - s &= 0, \\
Ax - b &\geq 0, \\
\lambda^{\mathrm{T}}(Ax - b) &= 0, \\
s^{\mathrm{T}} x &= 0, \\
\lambda, s, x &\geq 0.
\end{aligned}
$$

Solving for s,

$$
\begin{aligned}
\lambda^{\mathrm{T}}(Ax - b) &= 0, \\
(c - A^{\mathrm{T}} \lambda)^{\mathrm{T}} x &= 0, \\
Ax - b &\geq 0, \\
c - A^{\mathrm{T}} \lambda &\geq 0, \\
\lambda, x &\geq 0.
\end{aligned} \tag{1}
$$

For the dual problem we use $\mathcal{L}_{\mathcal{D}}(\lambda, x, u) = -b^{\mathrm{T}}\lambda - x^{\mathrm{T}}(c - A^{\mathrm{T}}\lambda) - u^{\mathrm{T}}\lambda$*:*

$$
\begin{gathered}
\nabla_\lambda \mathcal{L}_{\mathcal{D}} = -b + Ax - u = 0, \\
x^{\mathrm{T}}(c - A^{\mathrm{T}}\lambda) = 0, \\
c - A^{\mathrm{T}}\lambda \geq 0 \\
u^{\mathrm{T}}\lambda = 0, \\
\lambda, u, x \geq 0.
\end{gathered}
$$

Eliminating u *leads directly to* (1)*. This establishes the correspondence.*

c) Find the minimum value of

$$2x_1 + 2x_2 + 3x_3 + 2x_4$$

when

$$
\begin{gathered}
2x_1 + x_2 + x_3 + 0x_4 \geq 3 \\
x_1 + 2x_2 + 0x_3 + x_4 \geq 1 \\
x_i \geq 0, i = 1, \dots, 4.
\end{gathered}
$$

Also, find x.

Solution: We consider the dual problem, which may be read directly from ***b)****:*

$$
\max_{\lambda}(3\lambda_1 + \lambda_2),
$$

$$
\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda \leq \begin{bmatrix} 2 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \quad \lambda \geq 0.
$$

This is easy to solve graphically, see Fig. 1. Since $\min c^{\mathrm{T}}x = \max \lambda^{\mathrm{T}}b = 3 \cdot 1 + 0 = 3$*, the minimum we seek is* 3*.*

Finding x *is easy: Since* λ *is known, we apply* $(c - A^{\mathrm{T}}\lambda)^{\mathrm{T}}x = 0$ *to show that only* x_1 *may be different from* 0*. Looking at the original problem it follows that*

$$x^* = (3/2, 0, 0, 0)^{\mathrm{T}}.$$

$\boxed{2}$ Solve the problem

$$
\begin{gathered}
\min_x(2x_1 + 3x_2 + 4x_1^2 + 2x_1x_2 + x_2^2) \\
x_1 - x_2 \geq 0 \\
x_1 + x_2 \leq 4 \\
x_1 \leq 3 \\
x_2 \geq 0
\end{gathered}
$$

graphically. Verify the KKT conditions at the solution.

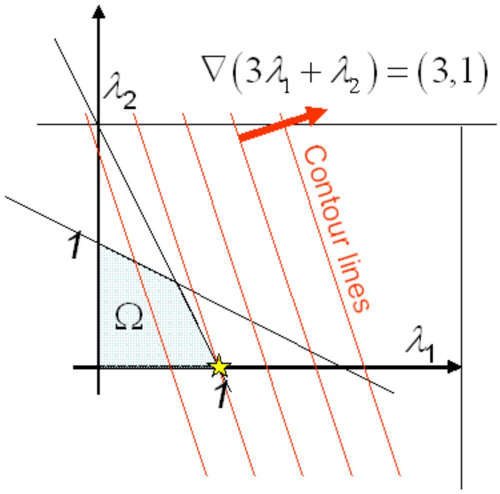

Figure 1: Solution of the dual problem is found at $\lambda_1 = 1$, $\lambda_2 = 0$.

Solution: The function f is convex since

$$\nabla^2 f(x) = \begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix} > 0.$$

A feasible point x^* will be a global minimum if $\nabla f(x^*) = 0$. Now,

$$\nabla f(x) = \begin{bmatrix} 2 + 8x_1 + 2x_2 \\ 3 + 2x_1 + 2x_2 \end{bmatrix},$$

which is 0 for $x_1 = 1/6$ and $x_2 = -5/3$. Since we should have $x_2 \geq 0$, $(1/6, -5/3)$ is not feasible.

The easiest is now to sketch Ω and use what we already know about the global minimum, see Fig. 2. We also know that the solution is unique (since f is strictly *convex!), and situated on the boundary of Ω, since $\nabla f(x) \neq 0$ for all $x \in \Omega$. The KKT conditions state that $\nabla f(x)$ is parallel to the gradient of the corresponding constraint along each boundary, and a (positive) sum of two gradients at each corner.*

The first point one tries is probably the origin, where $\nabla f(0,0) = (2,3)^{\mathrm{T}}$. It is obvious that

$$\nabla f(0,0)^{\mathrm{T}} d > 0$$

for all feasible directions from the origin, and this shows at once the origin is a local (and hence global) minimum. We verify that the KKT equation hold at the origin by noting that

$$\nabla f(0,0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \lambda_I \nabla c_I + \lambda_{IV} \nabla c_{IV} = 5 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Thus I and IV are active with positive Lagrange multipliers, whereas the rest are non-active and $\lambda_{II} = \lambda_{III} = 0$.

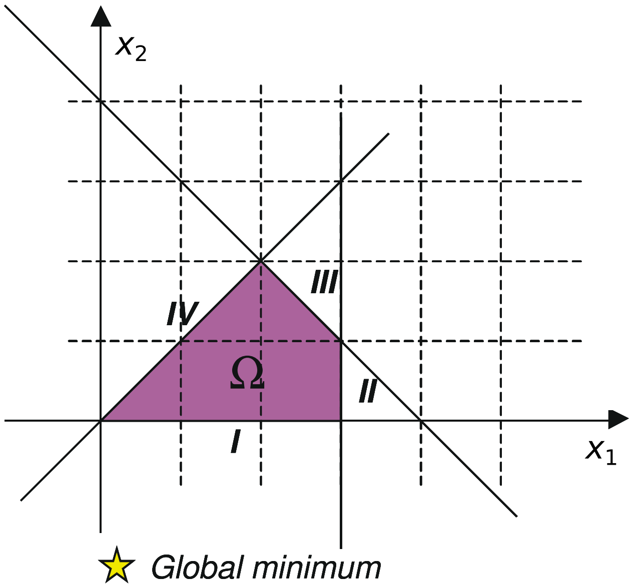

Figure 2: The feasible domain with four constraints.

$\boxed{3}$ N&W 2nd edition, exercise 16.2, p. 493.

Solution: This is a classic problem in Hilbert space. With our background in optimization, we introduce the Lagrange function

$$\begin{aligned}\mathcal{L}(x,\lambda) &= \tfrac{1}{2}(x-x_0)^{\mathrm{T}}(x-x_0) - \lambda^{\mathrm{T}}(Ax-b)\\ &= \tfrac{1}{2}x^{\mathrm{T}}x - x_0^{\mathrm{T}}x + \tfrac{1}{2}x_0^{\mathrm{T}}x_0 - \lambda^{\mathrm{T}}(Ax-b).\end{aligned}$$

The KKT equations become

$$\begin{aligned}\nabla_x\mathcal{L}(x,\lambda) = x - x_0 - A^{\mathrm{T}}\lambda &= 0,\\ Ax &= b,\end{aligned} \tag{2}$$

with the solution

$$\lambda^* = (AA^{\mathrm{T}})^{-1}(b - Ax_0),$$

since A has full rank. The solution then follows from (2):

$$x^* = x_0 + A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(b - Ax_0).$$

We may recall the expression for $b = 0$ as the projection *of x_0 onto the null space of A, $\mathcal{N}(A)$.*

If $A = a$ is a row vector, then

$$x^* = x_0 + a^{\mathrm{T}}\frac{1}{\|a\|^2}(b - ax_0)$$

and

$$\|x^* - x_0\| = \frac{|b - ax_0|}{\|a\|}$$

$\boxed{4}$ (Exam June 2006) Given the following problem:

$$\min_{x\in\mathbb{R}^n} q(x) \tag{3}$$
$$\text{where } a_i^{\mathrm{T}} x \geq b_i, \qquad i = 1, 2, \ldots, m,$$

with $q(x) := \frac{1}{2}x^{\mathrm{T}}Gx + x^{\mathrm{T}}d$ and G a symmetric, positive definite matrix.

Solution: You can find the solution to this exercise in the "Old Exams" section on the website. Look for the 2006 exam file.

a) Write down the Karush-Kuhn-Tucker conditions for the problem (but do not solve them). Assume that $x^\star$ is a solution to the KKT conditions. Is $x^\star$ then a global minimum of (3)?

b) Solve the problem

$$\min\,(x_1 - 1)^2 + (x_2 - 1)^2, \tag{4}$$

where

$$\begin{aligned} -x_1 - 2x_2 &\geq -2 \\ -2x_1 &\geq -3 \\ x_1, x_2 &\geq 0. \end{aligned}$$

Hint: Start with a sketch of the admissible domain and the level curves of $q(x)$.

The remainder of this problem is about constructing an iterative method for solving the general quadratic problem (3).

Given an admissible point x_0, let $\mathcal{W}$ be a given set of *active constraints* in x_0, such that $\mathcal{W} \subset \mathcal{A}(x_0)$.

c) Find a solution p of the reduced problem

$$\min_p q(x_0 + p)$$
$$\text{where } a_i^{\mathrm{T}}(x_0 + p) = b_i, \qquad i \in \mathcal{W},$$

assuming that the a_i are linearly independent for $i \in \mathcal{W}$.
Hint: Show first that this is equivalent to finding a minimum of $p^{\mathrm{T}}Gp/2 + p^{\mathrm{T}}(Gx_0 + d)$ with the constraints $a_i^{\mathrm{T}}p = 0$ for $i \in \mathcal{W}$.

d) Assume that the solution p of c) is not 0. Find an expression for the maximum value α such that $x_0 + \alpha p$ is an admissible point. The next step in the iteration is then given by

$$x_1 := x_0 + \min\{\alpha, 1\} \cdot p.$$

Explain why.

e) Execute one iteration step (points c) and d)) for problem (4). Start with $x_0 := [3/2, 0]^{\mathrm{T}}$. Choose $\mathcal{W}$ yourself.
Note: Even if you have not found formal solutions of the general problem in c) and d) it can well happen that you will find a solution to the specific problem (4).

f) The points c) and d) are part of an *active set method* for quadratic problems. To complete the algorithm, the following questions have to be answered:

- Is x_1 a solution?
- If not, how should we choose $\mathcal{W}$ in the next iteration?

Explain how these questions can be answered.

$\boxed{5}$ The following problem is copied from T.L. Saaty and J. Bram: *Nonlinear Mathematics*, p. 144. This is a little challenge!

Solve

$$\begin{aligned} \min f(x), \quad f(x) &= x_1^3 - 6x_1^2 + 11x_1 + x_3, \\ -x_1^2 - x_2^2 + x_3^2 &\geq 0, \\ x_1^2 + x_2^2 + x_3^2 - 4 &\geq 0, \\ -x_3 + 5 &\geq 0, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

Observe that f is independent of x_2.

a) Show that $(0, \sqrt{2}, \sqrt{2})$ is a KKT point.

Solution: At $x_0 = (0, \sqrt{2}, \sqrt{2})^{\mathrm{T}}$ we have

$$\begin{aligned} \nabla f(x_0) &= (11, 0, 1)^{\mathrm{T}}, \\ \nabla c_1(x_0) &= (0, -2\sqrt{2}, 2\sqrt{2})^{\mathrm{T}}, \\ \nabla c_2(x_0) &= (0, 2\sqrt{2}, 2\sqrt{2})^{\mathrm{T}}, \\ \nabla c_3(x_0) &= (0, 0, -1)^{\mathrm{T}}, \\ \nabla c_4(x_0) &= (1, 0, 0)^{\mathrm{T}}, \\ \nabla c_5(x_0) &= (0, 1, 0)^{\mathrm{T}}, \\ \nabla c_6(x_0) &= (0, 0, 1)^{\mathrm{T}}. \end{aligned}$$

The constraints c_1, c_2 and c_4 are active, and then $\lambda_3 = \lambda_5 = \lambda_6 = 0$. The rest of the Lagrange multipliers are found from

$$\begin{bmatrix} 11 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \lambda_1 - \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \lambda_2 - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \lambda_4 = 0,$$

which gives $\lambda_1 = \lambda_2 = 1/2$ and $\lambda_4 = 11$; in other words, $(0, \sqrt{2}, \sqrt{2})$ is a KKT point.

b) Do some numerical experiments using the MATLAB Optimization Toolbox function `fmincon`. Suggested code:

```
x0 = [2.1, 0, 2.1]';
x = fmincon(@SBfunction, x0, [], [], [], [], [], [], @constraints);
```

Function:

```
function f = SBfunction(x)
    f = x(1)^3 - 6*x(1)^2 + 11*x(1) + x(3);
end
```

x_0	x_{sol}
$(2.1, 0, 2.1)$	$(2.0069, 0, 2.0069)$
$(1, 1, 2)$	$(0.0000, 1.4142, 1.4142)$
$(0, 0, 4)$	$(0, 0, 2.0000)$
$(2.02, 0, 2.02)$	$(2.0200, 0, 2.0200)$

Table 1: Some start vectors and "solutions", using the standard parameters in `fmincon`.

Constraints:

```
function [c, ceq] = constraints(x)
    % Nonlinear inequalities (note sign-convention!)
    c(1) = x(1)^2 + x(2)^2 - x(3)^2;
    c(2) = -x(1)^2 - x(2)^2 - x(3)^2 + 4;
    c(3) = x(3) - 5;
    c(4) = -x(1);
    c(5) = -x(2);
    c(6) = -x(3);
    ceq = [];
end
```

Solution: The data obtained in the MATLAB *experiments is shown in Table 1. The KKT point from* **a)** *is recovered, but also some other "solutions" like* $(0, 0, 2)$ *and maybe,* $(2, 0, 2)$.

c) The book states that *"This problem actually has another local solution"*. Is this really true? Use MATLAB for experiments, but try to prove your claims. (This seems to be a challenge!)

Solution: The book claims that there is another local minimum, and this also seems to be supported by MATLAB. *The feasible domain is* Ω *bounded by the coordinate planes, the plane* $x_3 = 5$, *the sphere* $x_1^2 + x_2^2 + x_3^2 = 2^2$, *and the cone* $x_3 = \sqrt{x_1^2 + x_2^2}$. *The domain is shown in Fig. 3.*

Since the function f *is independent of* x_2, *it is constant on lines parallel to the* x_2-*axis. A part of the function* $s(x_1) = x_1^3 - 6x_1^2 + 11x_1$ *is plotted in Fig. 4. The function* s *has a local maximum for* $x_1 = 2 - 1/\sqrt{3} \approx 1.4226 > \sqrt{2}$, *and a local minimum for* $x_1 = 2 + 1/\sqrt{3}$.

Where is the other local minimum?

1. *Certainly not in the interior of* Ω *since* $\frac{\partial f}{\partial x_3} \equiv 1 \neq 0$.
2. *Not in a boundary point from which it is possible to move a small distance parallel to the* x_2-*axis and then a little down in the negative* x_3-*direction.(Why?)*
3. *Not on* $x_3 = 5$.

This leaves only the intersection between the sphere and the cone, and the line segment where $x_1 = x_3$.

Apart from the KKT point x_0, *there are no other local minima on the intersection between the sphere and the cone: The function* $s(x_1)$ *(which is the only part of* f *that changes) is strictly decreasing when we move on this intersection from* $(\sqrt{2}, 0, \sqrt{2})$ *over to the KKT point* $(0, \sqrt{2}, \sqrt{2})$.

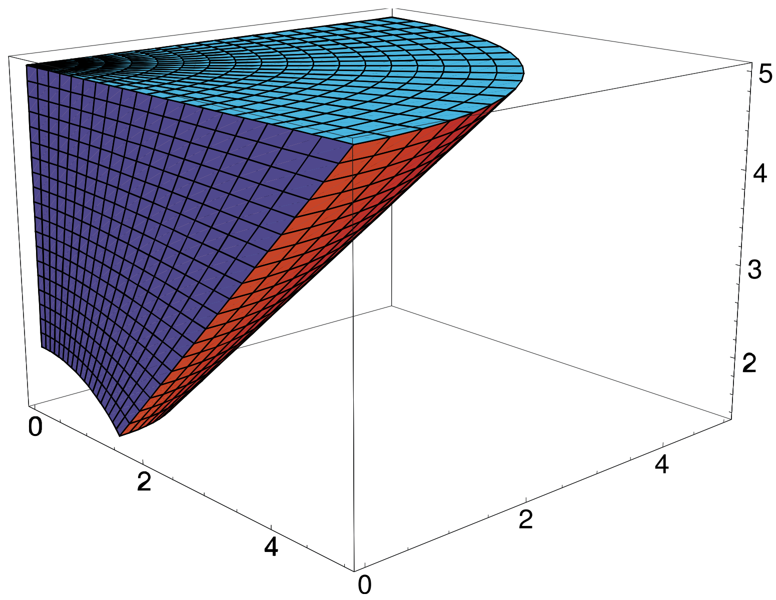

Figure 3: The feasible domain Ω.

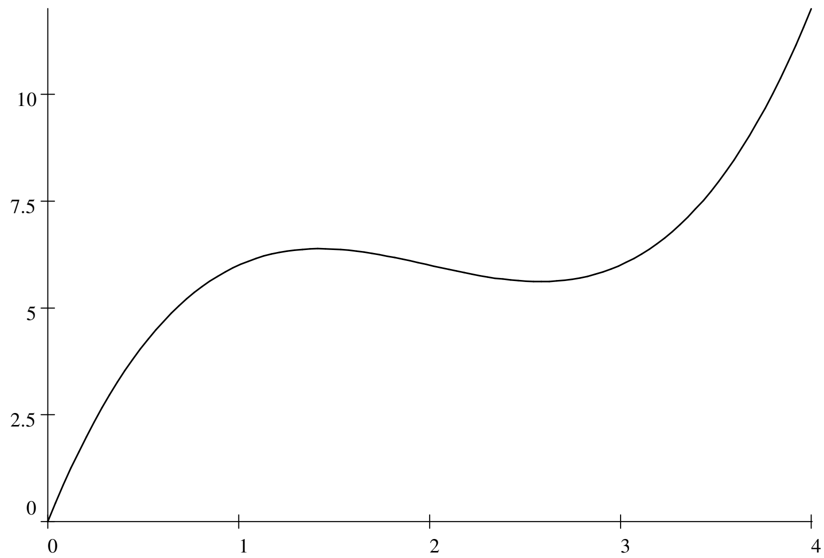

Figure 4: Part of the s-function.

On the line $x_1 = x_3$, the function f may be written

$$f(x_1, 0, x_1) = x_1^3 - 6x_1^2 + 12x_1,$$

and we leave to the reader to show that it is strictly increasing for all values of x. However, it is really to be expected that the point $(2, 0, 2)$ cheats the algorithm. (Why?) In summary, a quite tricky analysis!