Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 04–08–23

Examiner: Michael Patriksson

Question 1

(linear programming duality)

a) The LP dual is to

$$\begin{array}{rrcrcrcl} \text{minimize} & w = 2y_1 & + & 7y_2 & + & 3y_3, & & \\ \text{subject to} & -2y_1 & - & y_2 & + & y_3 & \geq & 1, \\ & y_1 & + & 2y_2 & & & \geq & 2, \\ & y_1 & , & y_2 & , & y_3 & \geq & 0. \end{array}$$

The complementarity conditions are next investigated. First, we check the conditions of the type

$$y_i^* \cdot \left(\sum_{j=1}^{n} a_{ij} x_j^* - b_i \right) = 0, \qquad i = 1, \dots, m.$$

Checking the primal constraints reveals that the first constraint is fulfilled strictly while the remaining two have no slack. This is implies that $y_1^* = 0$ must hold. Next, we investigate the second type of complementarity conditions:

$$x_j^* \cdot \left(\sum_{i=1}^{m} a_{ij} y_i^* - c_j \right) = 0, \qquad j = 1, \dots, n.$$

Since $\boldsymbol{x} = (3, 5)^{\mathrm{T}}$ is strictly positive, both dual constraints are active. Together with the fact that $y_1^* = 0$ leaves the following system of linear equations:

$$\begin{array}{rcrcl} -y_2^* & + & y_3^* & = & 1; \\ 2y_2^* & & & = & 2; \end{array}$$

its unique solution is that $y_2^* = 1$; $y_3^* = 2$.

It remains to check that all dual constraints are satisfied, that is, to also check the sign conditions. Non-negativity is clearly satisfied, so $\boldsymbol{y}^* = (0, 1, 2)^{\mathrm{T}}$ is the unique dual optimal solution. We therefore know from the complementarity theorem that $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ are optimal in their respective problem. But we check nevertheless that strong duality is fulfilled: $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^* = 13 = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^*$.

b) The proof is by contradiction. Suppose that (D) has a feasible solution. Since (P) has a feasible solution we can apply the Strong Duality Theorem and conclude that both (P) and (D) have finite optimal solutions which moreover have the same objective value. But this contradicts the fact that (P) has an unbounded solution. Therefore, the claim that (D) has a feasible solution is false. We are done.

Question 2

(convexity)

A proof is found in the course notes (Theorem 3.26).

Question 3

(modeling)

Variable declaration:

- c_i = number of barrels of crude oil of type i bought ($i = 1, 2, 3$);
- b_{ij} = number of barrels of crude oil of type i used to produce gas of type j ($i = 1, 2, 3,\ j = 1, 2, 3$);
- a_j = number of dollars spent on advertising for gas type i ($i = 1, 2, 3$);
- g_i = number of barrels of gas of type j produced ($j = 1, 2, 3$).

Objective function: maximize the difference between the income of selling oil and the cost of producing it (the latter including buying crude oil, transforming crude oil to gas, and advertizing), that is:

maximize $(70 - 4)g_1 + (60 - 4)g_2 + (50 - 4)g_3 - (45c_1 + 35c_2 + 25c_3 + a_1 + a_2 + a_3)$.

Constraints:

- For each type of oil:
  - definition of product;
  - minimum octane rating; and
  - maxiumum lead content;
- All crude oil bought is used;
- Maximum purchase of crude oil;
- Maximum capacity of production;

- Demand of products;
- Physical constraints.

In the same order:

$$\begin{aligned}
b_{11} + b_{21} + b_{31} &= g_1, \\
12b_{11} + 6b_{21} + 8b_{31} &\geq 10g_1, \\
0.5b_{11} + 2b_{21} + 3b_{31} &\leq g_1, \\
b_{12} + b_{22} + b_{32} &= g_2, \\
12b_{12} + 6b_{22} + 8b_{32} &\geq g_2, \\
0.5b_{12} + 2b_{22} + 3b_{32} &\leq 2g_2, \\
b_{13} + b_{23} + b_{33} &= g_3, \\
12b_{13} + 6b_{23} + 8b_{33} &\geq 6g_3, \\
0.5b_{13} + 2b_{23} + 3b_{33} &\leq g_3,
\end{aligned}$$

$$\begin{aligned}
b_{11} + b_{12} + b_{13} &= c_1, \\
b_{21} + b_{22} + b_{23} &= c_2, \\
b_{31} + b_{32} + b_{33} &= c_3,
\end{aligned}$$

$$c_j \leq 5,000, \quad j = 1, 2, 3,$$

$$c_1 + c_2 + c_3 \leq 14,000,$$

$$\begin{aligned}
g_1 &\geq 3,000 + 10a_1, \\
g_2 &\geq 2,000 + 10a_1, \\
g_3 &\geq 3,000 + 10a_1,
\end{aligned}$$

$$c_i, b_{ij}, a_j, g_j \geq 0, \quad i = 1, 2, 3; \; j = 1, 2, 3.$$

Question 4

(on the Armijo step length rule in unconstrained optimization)

a) We have that $x_{k+1} = x_k(1 - \alpha_k x_k^2)$.

The requirements of linear convergence imply that $1-\alpha_k x_k^2$ must be bounded away from 1, that is, that $\alpha_k x_k^2$ must be bounded away from zero. But since $x^* = 0$ this requires that α_k tends to infinity faster than x_k^2 tends to zero; there is obviously no finite value of α_0 that can produce such step lengths.

b) We have that $x_{k+1} = x_k(1 - \alpha_k/3)$.

According to the Newton formula above, if we can ensure that $\alpha_k = 1$ is always going to be accepted by the Armijo rule, then we have linear convergence with rate $q := 2/3$. In this case, we than have that $x_{k+1} = (2/3)x_k$. With $\alpha_k = 1$ the Armijo rule requires that $1 - (2/3)^4 \geq (4/3)\mu$ which clearly is satsified as long as the value of μ is small enough. ($\mu \leq 0.6$ will do.)

Question 5

(nonlinear programming optimality)

a) Let us first rewrite the LP problem into the following equivalent form, and note that $h_j(\bar{\boldsymbol{x}}) = 0$ for all j, since $\bar{\boldsymbol{x}}$ is feasible:

$$\begin{aligned} \underset{p}{\text{minimize}} \quad & \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{p}, \\ \text{subject to} \quad & -\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{p} \geq g_i(\bar{\boldsymbol{x}}), \qquad i = 1, \dots, m, \\ & -\nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{p} = 0, \qquad j = 1, \dots, \ell. \end{aligned}$$

Letting $\boldsymbol{\mu} \geq \mathbf{0}^m$ and $\boldsymbol{\lambda} \in \mathbb{R}^\ell$ be the dual variable vector for the inequality and equality constraints, respectively, we obtain the following dual program:

$$\begin{aligned} \underset{(\boldsymbol{\mu},\boldsymbol{\lambda})}{\text{maximize}} \quad & \sum_{i=1}^m \mu_i g_i(\bar{\boldsymbol{x}}), \\ \text{subject to} \quad & -\sum_{i=1}^m \mu_i \nabla g_i(\bar{\boldsymbol{x}}) - \sum_{j=1}^\ell \lambda_j \nabla h_j(\bar{\boldsymbol{x}}) = \nabla f(\bar{\boldsymbol{x}}), \\ & \mu_i \geq 0, \qquad i = 1, \dots, m. \end{aligned}$$

LP duality now establishes the result sought: First, suppose that the optimal value of the above primal problem over $\boldsymbol{p}$ is zero. Then, the same is true for the dual problem. Hence, by the sign conditions $\mu_i \geq 0$ and $g_i(\boldsymbol{x}) \leq 0$, each term in the sum must be zero. Hence, we established that *complementarity* holds. Next, the two constraints in the dual problem are precisely the *dual feasibility* conditions, which hence are fulfilled. Finally, *primal feasibility* of $\bar{\boldsymbol{x}}$ was assumed. It follows that this vector indeed is a KKT point.

Conversely, if $\bar{\boldsymbol{x}}$ is a KKT point, then the dual problem above has a feasible solution given by any KKT multiplier vector $(\boldsymbol{\mu}, \boldsymbol{\lambda})$. The dual objective is upper bounded by zero, since each term in the sum is non-positive. On the other hand, there is a feasible solution with the objective value 0, namely any KKT point! So, each KKT point must constitute an optimal solution to this dual LP problem! It then follows by duality theory that the dual of this problem, which is precisely the primal problem in $\boldsymbol{p}$ above, has a finite optimal solution, whose optimal value must then be zero. We are done.

The LP problem given in the exam is essentially the subproblem in the *Sequential Linear Programming* (SLP) algorithm. By the above analysis, the optimal value must be negative if $\bar{\boldsymbol{x}}$ is not a KKT point, and it must therefore also be negative (since a zero value is given by setting $\boldsymbol{p} = \boldsymbol{0}^n$). The optimal value of $\boldsymbol{p}$, if one exists, is therefore a descent direction with respect to f at $\bar{\boldsymbol{x}}$. A convergent SLP method introduces additional box constraints on $\boldsymbol{p}$ in the LP subproblem to make sure that the solution is finite, and the update is made according to a line search with respect to some penalty function.

b) The problem is convex if f and the functions g_i $(i = 1, \dots, m)$ are convex, and the functions h_j $(j = 1, \dots, \ell)$ are affine. A proof that every KKT point is globally optimal is found in the course notes (Theorem 6.45).

Question 6

(linear programming geometry)

We prove first the result in the direction "$\Longleftarrow$". So we assume that such a vector $\boldsymbol{\mu}$ exists. Let $\boldsymbol{x} \in X$. Then,

$$\boldsymbol{d}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{b} \leq d_0$$

holds, which establishes that the inequality is redundant: it is always fulfilled on X.

We next prove the result in the direction "$\Longrightarrow$". So we assume that the inequality is redundant. An implication of that is that the following LP problem must have an optimal value which is less than or equal to d_0, because otherwise we would reach a contradiction:

$$\begin{aligned}
\text{maximize} \quad & \boldsymbol{d}^{\mathrm{T}}\boldsymbol{x}, \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \\
& \boldsymbol{x} \geq \boldsymbol{0}^n.
\end{aligned}$$

Since the primal problem has a finite optimal solution, so does the dual problem to

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\mu} \geq \boldsymbol{d}, \\ & \boldsymbol{\mu} \geq \boldsymbol{0}^m. \end{aligned}$$

This solution is in particular feasible, and its optimal value must also be less than or equal to d_0. We are done.

Question 7

(Lagrangian duality)

(2p) a) The Slater's CQ is clearly verified since the problem is convex (even linear), and there is a strictly feasible point [e.g., $(x, y)^{\mathrm{T}} = (3, 1)^{\mathrm{T}}$].

Introducing Lagrange multipliers μ_1 and μ_2 we calculate the Lagrangian dual function q:

$$\begin{aligned} q(\mu_1, \mu_2) &= \min_{(\mu_1,\mu_2)\in\mathbb{R}^2_+} \{x - 0.5y + \mu_1(-x + y + 1) + \mu_2(-2x + y + 2)\} \\ &= \mu_1 + 2\mu_2 + \min_{x\geq 0} (1 - \mu_1 - 2\mu_2)x + \min_{y\geq 0} (-0.5 + \mu_1 + \mu_2)y \\ &= \begin{cases} \mu_1 + 2\mu_2, & \text{if } \mu_1 + 2\mu_2 \leq 1 \text{ and } \mu_1 + \mu_2 \geq 0.5, \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Thus the set of optimal Lagrange multipliers is $\{ (\mu_1, \mu_2) \mid \mu_1 \geq 0, \ \mu_2 \geq 0, \ \mu_1 + 2\mu_2 = 1 \}$, which is clearly convex and bounded (e.g., you may illustrate this graphically) as it should be in the presence of Slater's CQ.

(1p) b) Subgradients of the Lagrangian dual function are calculated as follows:

1. At $(\mu_1, \mu_2)^{\mathrm{T}} = (1, 0)^{\mathrm{T}}$ the set of optimal solutions to the Lagrangian relaxed problem is the singleton $\{ (0, 0)^{\mathrm{T}} \}$. Hence, the Lagrangian function is differentiable at this point and its gradient equals the value of the vector of constraint functions evaluated at the optimal solution to the relaxed problem, i.e., $(-0 + 0 + 1, -2 \cdot 0 + 0 + 2)^{\mathrm{T}} = (1, 2)^{\mathrm{T}}$. Alternatively, we may directly differentiate q at a given point to obtain the same result.
2. At $(\mu_1, \mu_2)^{\mathrm{T}} = (1/4, 1/3)^{\mathrm{T}}$ the set of optimal solutions to the Lagrangian relaxed problem is not a singleton: it equals $\{ (x, 0)^{\mathrm{T}} \mid x \geq$

$0\,\}$. Hence, the dual function is not differentiable, and the set of subgradients is obtained by evaluating the constraint functions at the optimal solutions to the relaxed problem, i.e., $\partial q(1/4, 1/3) = \{\,(-x+1, -2x+2)^{\mathrm{T}} \mid x \geq 0\,\}$.