Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 07–08–30
Examiner: Michael Patriksson

Question 1

(the Simplex method)

(2p) a) After changing sign of the second inequality and adding two slack variables s_1 and s_2, a BFS cannot be found directly. We create the phase I problem through an added artificial variable a_1 in the second linear constraint; the value of a_1 is to be minimized.

We use the BFS based on the variable pair (s_1, a_1) as the starting BFS for the phase I problem. In the first iteration of the Simplex method x_1 is the only variable with a negative reduced cost; hence x_1 is picked as the incoming variable. The minimum ratio test shows that s_1 should leave the basis. In the next iteration the reduced cost for varialbe x_3 is negative, and x_3 is picked as the incoming variable. The minimum ratio test shows that a_1 should leave the basis. We have found an optimal basis, $x_B = (x_1, x_3)^{\mathrm{T}}$, to the phase I problem. We proceed to phase II, since the basis is feasible in the original problem.

Starting phase II with this BFS, we see that all reduced costs are positive, $\tilde{c}_N = (3, 2, 3)^{\mathrm{T}} > 0$, and thus the BFS is optimal. $x_B = B^{-1}b = (2, 1)^{\mathrm{T}}$ so $x^* = (2, 0, 3)^{\mathrm{T}}$ and $z^* = c_B^{\mathrm{T}} x_B = 3$.

(1p) b) Yes. The reduced costs are positive.

(3p) Question 2

(strong duality in linear programming)

See Theorem 10.6 in The Book.

Question 3

(exterior penalty method)

(1p) a) Direct application of the KKT conditions yield that $\boldsymbol{x}^* = (\frac{3}{5}, \frac{2}{5})^{\mathrm{T}}$ and $\lambda^* = -1/5$ uniquely.

(1p) b) Letting the penalty parameter be $\nu > 0$, it follows that $\boldsymbol{x}_\nu = \frac{\nu}{1+5\nu}(3, 2)^{\mathrm{T}}$. Clearly, as $\nu \to \infty$ convergence to the optimal primal–dual solution follows.

(1p) c) From the stationarity conditions of the penalty function $\boldsymbol{x} \mapsto f(\boldsymbol{x}) + \lambda h(\boldsymbol{x}) + \nu |h(\boldsymbol{x})|^2$ follow that $\boldsymbol{x}_\nu$ fulfills $\nabla f(\boldsymbol{x}_\nu) + [2\nu h(\boldsymbol{x}_\nu)] \nabla h(\boldsymbol{x}_\nu) = 0^2$, and hence a proper Lagrange multiplier estimate comes out as $\lambda_\nu := 2\nu h(\boldsymbol{x}_\nu)$. Insertion from b) yields $\lambda_\nu = \frac{-\nu}{1+5\nu}$, which tends to $\lambda^* = -\frac{1}{5}$ as $\nu \to \infty$.

Question 4

(true or false claims in optimization)

(1p) a) True. $\nabla f(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{p} = -2$.

(1p) b) False. Suppose, for example, that the Hessian of f at $\boldsymbol{x}$ is negative definite, and that $\boldsymbol{x}$ is not a stationary point. Then the Newton direction is well-defined but it is an ascent direction.

(1p) c) True. The result follows rather immediately from the definition of descent direction.

(3p) Question 5

(least-squares minimization)

We wish to minimize $||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2$ or equivalently $f(\boldsymbol{x}) = ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2^2$ over $\boldsymbol{x} \in \mathbb{R}^n$, i.e. we have a unconstrained optimization problem. We rewrite

$$f(\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}$$

The hessian of $f(\boldsymbol{x})$ is $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ and is always positive semi-definite since $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v} = ||\boldsymbol{A}\boldsymbol{v}||^2 \geq 0 \quad \forall \boldsymbol{v}$. Thus, the minimization problem is convex and from the optimality conditions we know that stationarity is sufficient for a point to be optimal.

We have $\bigtriangledown f(\boldsymbol{x}^*) = \boldsymbol{0} \iff \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. If the rank of $\boldsymbol{A}$ is n then $||\boldsymbol{A}\boldsymbol{v}||^2 > 0 \quad \forall \boldsymbol{v} \neq \boldsymbol{0}$, the hessian is positive definite and therefore invertible and we get the wanted result.

(3p) Question 6

(modelling)

Introduce the variables:

x_{ij} Number of persons recruited in the beginning of month i to the end of month j, $i = 1, \dots, 24$, $j = i, \dots, 24$

y_t Value one if anyone is recruited month t, zero otherwise.

The objective is

$$\min \quad \sum_i \sum_j w x_{ij} + \sum_i y_i k + \sum_{t=1}^{24} \sum_{i \le t, j \ge t} x_{ij} r$$

and the constraints are

$$\begin{aligned}
\sum_{i \le t, j \ge t} x_{ij} &\ge d_t, && t = 1, \dots, 24, \\
x_{ij} &= 0, && \forall (i,j) : i = j, i = j+1, && (1) \\
M y_i &\ge \sum_j x_{ij}, && i = 1, \dots, 24 && (2) \\
y_t &\in \mathbb{B} \\
x_{ij} &\in \mathbb{Z}_+
\end{aligned}$$

where M is a big number. Constraint (1) sets the required work force. Constraint (1) sets the recruitments to more than 3 months. Constraint (2) is present for setting the auxiliary variable y.

Question 7

(duality in linear and nonlinear optimization)

(1p) a) The LP dual is to

$$\begin{array}{llllll}
\text{maximize} & w = \boldsymbol{b}_1^{\mathrm{T}} \boldsymbol{y}_1 & +\boldsymbol{b}_2^{\mathrm{T}} \boldsymbol{y}_2 & +a y_3 & \\
\text{subject to} & \boldsymbol{A}_1^{\mathrm{T}} \boldsymbol{y}_1 & +\boldsymbol{A}_2^{\mathrm{T}} \boldsymbol{y}_2 & & \le \boldsymbol{c}, \\
& \boldsymbol{B}^{\mathrm{T}} \boldsymbol{y}_1 & & +\boldsymbol{1}^{\ell} y_3 & \le \boldsymbol{d}, \\
& \boldsymbol{y}_1 \ge \boldsymbol{0}^{m_1}, & \boldsymbol{y}_2 \in \mathbb{R}^{m_2}, y_3 \in \mathbb{R}, & &
\end{array}$$

where $\boldsymbol{1}^{m_1}$ is the m_1-vector of ones.

(2p) b) With $g(\boldsymbol{x}) := -x_1 + 2x_2 - 4$, the Lagrange function becomes

$$\begin{aligned} L(\boldsymbol{x}, \mu) &= f(\boldsymbol{x}) + \mu g(\boldsymbol{x}) \\ &= 2x_1^2 + x_2^2 - 4x_1 - 6x_2 + \mu(-x_1 + 2x_2 - 4). \end{aligned}$$

Minimizing this function over $\boldsymbol{x} \in \mathbb{R}^2$ yields [since $L(\cdot, \mu)$ is a strictly convex quadratic function for every value of μ, it has a unique minimum for every value of μ] that its minimum is attained where its gradient is zero. This gives us that

$$\begin{aligned} x_1(\mu) &= (4 + \mu)/4; \\ x_2(\mu) &= 3 - \mu. \end{aligned}$$

Inserting this into the Lagrangian function, we define the dual objective function as

$$q(\mu) = L(\boldsymbol{x}(\mu), \mu) = \cdots = -2\left(\frac{4+\mu}{4}\right)^2 - (3-\mu)^2 - 4\mu.$$

This function is to be maximized over $\mu \geq 0$. We are done with task [1].

We attempt to optimize the one-dimensional function q by setting the derivative of q to zero. If the resulting value of μ is non-negative, then it must be a global optimum; otherwise, the optimum is $\mu^* = 0$.

We have that $q'(\mu) = \cdots = 1 - \frac{9\mu}{4}$, so the stationary point of q is $\mu = 4/9$. Since its value is positive, we know that the global maximum of q over $\mu \geq 0$ is $\mu^* = 4/9$. We are done with task [2].

Our candidate for the global optimum in the primal problem is $\boldsymbol{x}(\mu^*) = \frac{1}{9}(10, 23)^{\mathrm{T}}$. Checking feasibility, we see that $g(\boldsymbol{x}(\mu^*)) = 0$. Hence, without even evaluating the values of $q(\mu^*)$ and $f(\boldsymbol{x}(\mu^*))$ we know they must be equal, since $q(\mu^*) = f(\boldsymbol{x}(\mu^*)) + \mu^* g(\boldsymbol{x}(\mu^*)) = f(\boldsymbol{x}(\mu^*))$, due to the fact that we satisfy complementarity. We have proved that strong duality holds, and therefore task [4] is done.

By the Weak Duality Theorem 7.4 follows that if a vector $\boldsymbol{x}$ is primal feasible and $f(\boldsymbol{x}) = q(\mu)$ holds for some feasible dual vector μ, then $\boldsymbol{x}$ must be the optimal solution to the primal problem. (And μ must be optimal in the dual problem.) Task [4] is completed by the remark that this is exactly the case for the pair $(\boldsymbol{x}(\mu^*), \mu^*)$.