Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 13–04–02
Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. We rewrite $x_2 = x_2^+ - x_2^-$ and introduce slack variables s_1 and s_2.

$$\begin{array}{llllllll} \text{minimize} & x_1 - & 2x_2^+ + & 2x_2^- & & & \\ \text{subject to} & -x_1 + & x_2^+ - & x_2^- & + s_1 & & = 1, \\ & 2x_1 + & x_2^+ - & x_2^- & & + s_2 & = 4, \\ & & & & & x_1, x_2^+, x_2^-, s_1, s_2 & \geq 0. \end{array}$$

Phase I

If we start with basis (s_1, s_2), we have a unit basis matrix, and the right-hand side then is $(1, 4)^{\mathrm{T}} \geq (0, 0)^{\mathrm{T}}$, which is therefore a basic feasible solution.

Phase II

Calculating the reduced costs, we obtain $\tilde{\boldsymbol{c}}_N = (1, -2, 2)^T$, meaning that x_2^+ should enter the basis. From the minimum ratio test, we get that the outgoing variable is s_1. Updating the basis we now have (x_2^+, s_1) in the basis.

Calculating the reduced costs, we obtain $\tilde{\boldsymbol{c}}_N = (-1, 0, 2)^T$, meaning that x_1 should enter the basis. From the minimum ratio test, we get that the only eligible outgoing variable is s_2. Updating the basis we now have (x_1, x_2^+) in the basis.

Calculating the reduced costs, we obtain $\tilde{\boldsymbol{c}}_N \geq 0$, meaning that the current basis is optimal. The optimal solution is thus $(x_1, x_2^+, x_2^-, s_1, s_2)^{\mathrm{T}} = (1, 2, 0, 0, 0)^{\mathrm{T}}$, which in the original variables means $(x_1, x_2) = (1, 2)^{\mathrm{T}}$, with optimal objective value $f^* = -3$.

(1p) b) We have that the optimal dual variables are $\boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} = -\frac{1}{3}(5, 1)^{\mathrm{T}}$. So a $\varepsilon > 0$ change in the first constraint would mean that the optimal objective value would change to $f^* = -3 - \frac{5}{3}\varepsilon$.

Question 2

(descent methods in unconstrained optimization)

(1p) a) Assuming that the optimal step length is positive—otherwise the algorithm would have stopped in the search direction phase with the verdict that the gradient of f at $\boldsymbol{x}_k$ is zero—the optimality conditions for the problem to minimize $f(\boldsymbol{x}_k + \ell \boldsymbol{p}_K)$ with respect to $\ell \geq$ is simply that the derivative of $f(\boldsymbol{x}_k + \ell \boldsymbol{p}_k)$ with respect to $\ell \geq$ is zero. With $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \ell \boldsymbol{p}_k$ this is expressed precisely as $\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}} \boldsymbol{p}_k = 0$.

(2p) b) At $\boldsymbol{x} \in \mathbb{R}^2$, the gradient of f equals

$$\begin{pmatrix} -400x_1(x_2 - x_1^2) + 2(x_1 - 1) \\ 200(x_2 - x_1^2) \end{pmatrix}.$$

Hence, the Hessian of f at $\boldsymbol{x} \in \mathbb{R}^2$ equals

$$\begin{pmatrix} 100(12x_1^2 - 4x_2) + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}.$$

At $\boldsymbol{x}^* = (1,1)^{\mathrm{T}}$, then, $\nabla f(\boldsymbol{x}^*) = (0,0)^{\mathrm{T}}$, and

$$\nabla^2 f(\boldsymbol{x}^*) = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}.$$

The eigenvalues of $\nabla^2 f(\boldsymbol{x}^*)$ are both positive; hence, $\nabla^2 f(\boldsymbol{x}^*)$ is positive definite.

Investigating the eigenvalues of $\nabla^2 f(\boldsymbol{x})$ we arrive at the conclusion that the Hessian matrix is singular when $x_1^2 - 2x_2 = 0.005$ and positive definite when $x_1^2 - x_2 > 0.005$.

(3p) Question 3

(separation and projection)

Let $\hat{\boldsymbol{z}} = \mathrm{proj}_S(\boldsymbol{z})$. Then,

$$\begin{aligned} \|\boldsymbol{z} - \boldsymbol{x}\|^2 &= \|\boldsymbol{z} - \hat{\boldsymbol{z}} + \hat{\boldsymbol{z}} - \boldsymbol{x}\|^2 \\ &= \|\boldsymbol{z} - \hat{\boldsymbol{z}}\|^2 + \|\hat{\boldsymbol{z}} - \boldsymbol{x}\|^2 + 2(\boldsymbol{z} - \hat{\boldsymbol{z}})^{\mathrm{T}}(\hat{\boldsymbol{z}} - \boldsymbol{x}). \end{aligned}$$

But the hyperplane $(\boldsymbol{z}-\hat{\boldsymbol{z}})^{\mathrm{T}}(\hat{\boldsymbol{z}}-\boldsymbol{x})$ is a hyperplane separating $\boldsymbol{z}$ from S, i.e. $(\boldsymbol{z}-\hat{\boldsymbol{z}})^{\mathrm{T}}(\hat{\boldsymbol{z}}-\boldsymbol{y}) \geq 0$ for all $\boldsymbol{y} \in S$. In particular, $(\boldsymbol{z}-\hat{\boldsymbol{z}})^{\mathrm{T}}(\hat{\boldsymbol{z}}-\boldsymbol{x}) \geq 0$. Since $\|\boldsymbol{z}-\hat{\boldsymbol{z}}\|^2 \geq$, we find that

$$\|\boldsymbol{z}-\boldsymbol{x}\|^2 = \|\boldsymbol{z}-\hat{\boldsymbol{z}}\|^2 + \|\hat{\boldsymbol{z}}-\boldsymbol{x}\|^2 + 2(\boldsymbol{z}-\hat{\boldsymbol{z}})^{\mathrm{T}}(\hat{\boldsymbol{z}}-\boldsymbol{x}) \geq \|\hat{\boldsymbol{z}}-\boldsymbol{x}\|^2.$$

Question 4

(true or false claims in optimization)

(1p) a) *True.*

Motivation: The problem setting is such that we may convert the problem to an unconstrained optimization problem over a subspace of $\mathbb{R}^n$; the local optimality conditions then imply, in fact, that the objective function is convex, whence the local minimum is a global one.

(1p) b) *False.*

Motivation: The case of $f(x) := x^3$ at $x = 0$ serves as an example. The direction $p = -1$ is a direction of descent with respect to f at $x = 0$, yet $f'(x) = 0$.

(1p) c) *False.*

Motivation: The case of maximizing x_1 subject to $x_1 \geq 0$ is a simple LP example where there exist feasible solutions, but no optimal solution.

(3p) Question 5

(sufficiency of the KKT conditions under convexity)

This is Theorem 5.45 in The Book.

Question 6

(Lagrangian duality)

(2p) a) Denote the Lagrangian dual function with respect to relaxation of both constraints as $q(\boldsymbol{\mu})$. We have that $q(\mathbf{0}) = \min_{x_1,x_2}(x_1-2)^2 + (x_2-1)^4 = 0$. Further we have that $\boldsymbol{x} = [-3,-1]^{\mathrm{T}}$ is feasible with respect to the constraints, with objective value $f(\boldsymbol{x}) = 41$. Thus, by weak Lagrangian duality, the optimal value f^* lies in the interval $[0, 41]$.

(1p) b) The problem is clearly convex, and the point $\boldsymbol{x} = [-3,-1]^{\mathrm{T}}$ is *strictly* feasible. Thus the Slater CQ holds, so strong Lagrangian duality must hold.

(3p) Question 7

(modelling)

The decision variables in the model are k_1, k_2, k_3. In order to formulate a linear program, we introduce the following auxiliary variables:

$$\begin{aligned} s_i &= \text{the error of the velocity at point } i = 1, \dots, n, \\ h_i &= \text{the error of the acceleration at point } i = 1, \dots, n. \end{aligned}$$

We have that $v'(t) = k_2 \cos(t) - k_3 \sin(t)$. The model can then be formulated as that to

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} (s_i + h_i), \\ \text{subject to} \quad & y_i - (k_1 + k_2 \sin(t_i) + k_3 \cos(t_i)) \leq s_i, \quad && i = 1, \dots, n, \\ & y_i - (k_1 + k_2 \sin(t_i) + k_3 \cos(t_i)) \geq -s_i, \quad && i = 1, \dots, n, \\ & a_i - (k_2 \cos(t_i) - k_3 \sin(t_i)) \leq h_i, \quad && i = 1, \dots, n, \\ & a_i - (k_2 \cos(t_i) - k_3 \sin(t_i)) \geq -h_i, \quad && i = 1, \dots, n, \\ & s_i, h_i \geq 0, \quad && i = 1, \dots, n. \\ & k_1, k_2, k_3 \in \mathbb{R}. \end{aligned}$$
