Chalmers/GU
Mathematics

EXAM

TMA947/MAN280
APPLIED OPTIMIZATION

Date:	04–03–08
Time:	House V, morning
Aids:	Text memory-less calculator
Number of questions:	7; passed on one question requires 2 points of 3. Questions are *not* numbered by difficulty. To pass requires 10 points and three passed questions.
Examiner:	Michael Patriksson
Teacher on duty:	Anton Evgrafov (0740-459022)
Result announced:	04–03–22 Short answers are also given at the end of the exam on the notice board for optimization in the MD building.

Exam instructions

When you answer the questions

State your methodology carefully.
Use generally valid methods and theory.

Only write on one page of each sheet. Do not use a red pen.
Do not answer more than one question per page.

At the end of the exam

Sort your solutions by the order of the questions.
Mark on the cover the questions you have answered.
Count the number of sheets you hand in and fill in the number on the cover.

Question 1

(the Simplex method)

Consider the linear program

$$\begin{array}{llrrrl} \text{minimize} & z = & x_1 & +3x_2 & +x_3 & \\ \text{subject to} & & 2x_1 & -5x_2 & +x_3 & \leq -5, \\ & & 2x_1 & -x_2 & +2x_3 & \leq \ 4, \\ & & x_1, & x_2, & x_3 & \geq \ 0. \end{array}$$

(2p) a) Solve the problem by using Phase I & II of the Simplex method.

Hint: Some matrix inverses that can be useful when solving the problem are:

$$\begin{pmatrix} 2 & -5 \\ 2 & -1 \end{pmatrix}^{-1} = \frac{1}{8}\begin{pmatrix} -1 & 5 \\ -2 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 5 & 0 \\ -1 & 1 \end{pmatrix}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 0 \\ 1 & 5 \end{pmatrix},$$

$$\begin{pmatrix} -1 & -1 \\ 2 & 0 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -2 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 5 & -1 \\ -1 & 2 \end{pmatrix}^{-1} = \frac{1}{9}\begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}.$$

(1p) b) Is the solution obtained unique? *Motivate!*

Question 2

(optimality conditions)

Consider the optimization problem to

$$\begin{aligned} \text{minimize} \quad & f(x,y) := \frac{1}{2}(x-2)^2 + \frac{1}{2}(y-1)^2, \\ \text{subject to} \quad & x - y \geq 0, \\ & y \geq 0, \\ & y(x-y) = 0, \end{aligned} \tag{1}$$

where $x, y \in \mathbb{R}$.

(1p) a) Find *all* points of global and local minimum, as well as *all* KKT-points.

Hint: Draw the problem graphically!

Is this a convex problem?

(1p) b) Demonstrate that the linear independence constraint qualification (LICQ) is violated at *every feasible point* of the problem (1).

The problem (1) can be solved as follows. Based on the original problem (1), we can formulate *two* optimization problems, both of which are convex and have one linear constraint. Having solved the two problems, the solution to the problem (1) is the solution with the best objective value.

Show *which* two problems should be solved.

Hint: Use the graphics used in a)!

(1p) c) In part b) we devised a "procedure" for solving the problem (1) in which two problems are solved and their respective optimal solutions compared. Generalize this procedure to the more general optimization problem to

$$\begin{aligned} &\text{minimize} \quad g(\boldsymbol{x}), \\ &\text{subject to} \quad & \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} \geq b_i, \quad & i = 1, \dots, n, \\ && x_i \geq 0, \quad & i = 1, \dots, n, \\ && x_i(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} - b_i) = 0, \quad & i = 1, \dots, n, \end{aligned}$$

where $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, $\boldsymbol{a}_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \dots, n$, and $g : \mathbb{R}^n \to \mathbb{R}$ is a convex differentiable function.

How many problems do we need to solve, and what are their forms?

(3p) Question 3

(modelling)

Figure 1 below describes a production process by which we can make three products, A, B, and C, from the raw materials D, E, and F.

The numbers at the top of the figure provide the maximum sales and unit revenues for the three products. The numbers at the bottom indicate the raw materials used and the unit costs of the raw materials. We assume that the supply of the raw materials is unlimited.

The network structure shows the processing requirements for the products. The nodes represent operations that the intermediate products must pass through. Each node is labeled with the corresponding operation number. The arcs (links) represent the inputs to the operations. Product A requires one unit from operation 1. Product B requires one unit from operation 2. Product C requires one unit from operation 3.

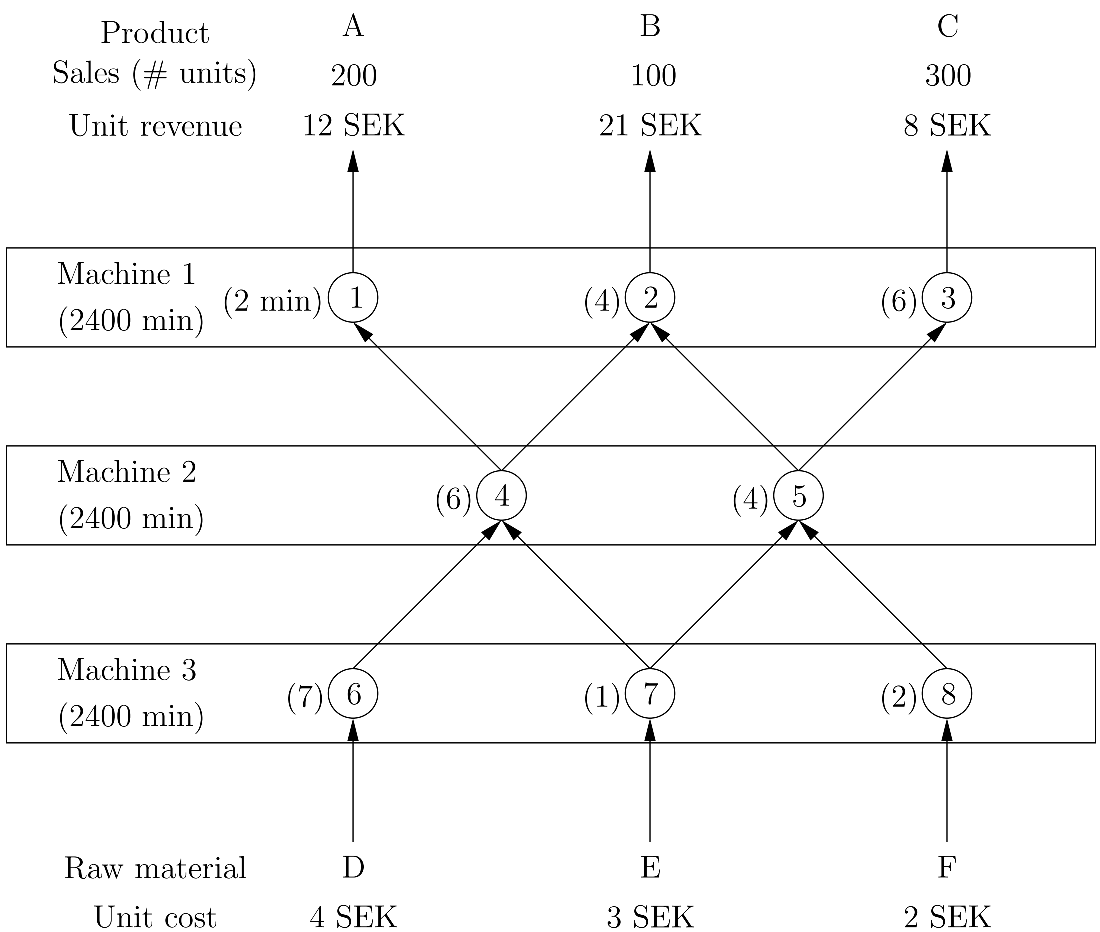

Figure 1: The production process.

The product at operation 1 is made from one unit passing out of operation 4. The product at operation 2 is made from one unit *each* of the products passing out of operations 4 and 5. The product at operation 3 is made from one unit passing out of operation 5.

A unit at operation 4 is made from one unit *each* from operations 6 and 7. A unit at operation 5 is made from one unit *each* from operations 7 and 8.

Finally, operation 6 requires one unit of raw material D. Operation 7 requires one unit of raw material E. Operation 8 requires one unit of raw material F.

The numbers adjacent to the nodes are the operation times in minutes. For example, operation 6 requires 7 minutes in order to produce one unit of output. The operations use time on Machine 1, Machine 2, and Machine 3. Each machine has a weekly capacity of 2400 minutes. The products share the capacities of the machines. For example, if 100 units of each product were produced, 1200 minutes would be used on machine 1. Note that one unit of product B needs 10 minutes of machine 2 because both operations 4 and 5 are needed for one unit of B.

Formulate a *linear integer programming model* (that is, if the integer requirements are relaxed we shall end up with an ordinary linear program) for finding the weekly production quantities that maximize the total income!

Hint: Introduce one variable for each arc (link) in the figure.

Question 4

(applications of the Newton algorithm)

(1p) a) Let a be a positive real number.

Consider the optimization problem to

$$\begin{aligned} &\text{minimize} \quad f(x) = ax - \log(x), \\ &\text{subject to} \quad x > 0. \end{aligned} \tag{1}$$

Prove that there exists a globally optimal solution to the problem (1), which is furthermore unique and equal to a^{-1}.

Motivate every step!

If you wish, you may replace the condition $x > 0$ with $x \geq 0$ in your analysis, as long as you are aware of the fact that $\log(0) = -\infty$.

(1p) b) Show that Newton's method with unit steps as applied to the problem (1) gives a computationally viable procedure for computing $x = a^{-1}$. That is, show that every iteration of Newton's method requires only additions (or subtractions) and multiplications to be performed; thus, we never need to perform divisions in order to compute the next iterate.

(*Note:* This idea is in fact used in the Intel Itanium processor!)

Construct an example (that is, choose some appropriate $a > 0$ and a starting point $x_0 > 0$) that satisfies the following requirements:

(i) Newton's method converges, that is, $\infty \neq \bar{x} = \lim_{k\to\infty} x_k$, but

(ii) $\bar{x} \neq a^{-1}$.

(*Note:* This illustrates the *local* nature of Newton's method, which is guaranteed to converge only when we start "near enough" to an optimal solution.)

(1p) c) Similarly to the previous parts, construct a convex optimization problem that can be used to calculate $x = a^{-1/2}$.

[That is, Newton's method with unit steps applied to your problem should not contain any other operations than additions (or subtractions) and multiplications.]

Question 5

(optimality conditions)

Farkas' Lemma can be stated as follows:

Let $\boldsymbol{A}$ *be an* $m \times n$ *matrix and* $\boldsymbol{b}$ *an* $m \times 1$ *vector. Then exactly one of the systems*

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{x} &\geq \mathbf{0}^n, \end{aligned} \tag{I}$$

and

$$\begin{aligned} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} &\leq \mathbf{0}^n, \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} &> 0, \end{aligned} \tag{II}$$

has a feasible solution, and the other system is inconsistent.

(2p) a) Prove Farkas' Lemma.

(1p) b) Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 1)^2, \\ \text{subject to} \quad & 2x_1 - x_2 = 0, \\ & 0 \leq x_1 \leq 2, \\ & 0 \leq x_2 \leq 2. \end{aligned}$$

Geometrically, it is not difficult to see that the vector $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$ cannot be an optimal solution to this problem. Your task is to prove this fact rigorously by using Farkas' Lemma, namely, prove that $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$ is not optimal to the above problem, by using Farkas' Lemma to show that there must exist a feasible descent direction with respect to f at $\bar{\boldsymbol{x}}$.

(3p) Question 6

(convexity)

Carathéodory's Theorem can be stated as follows:

Let $\boldsymbol{x} \in \operatorname{conv} V$, where $V \subseteq \mathbb{R}^n$. Then $\boldsymbol{x}$ can be expressed as a convex combination of $n+1$ or fewer points of V.

Prove Carathéodory's Theorem.

When you prove this result, you may make reference, without proof, to the following proposition:

Let $V \subseteq \mathbb{R}^n$. Then, $\operatorname{conv} V$ is the set of all convex combinations of points of V. ■

Question 7

(duality in linear and nonlinear optimization)

(1p) a) Consider the LP problem to

$$\begin{array}{lrrl} \text{minimize} & z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} & +\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} & \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x} & +\boldsymbol{B}\boldsymbol{v} & \geq \boldsymbol{b}_1, \\ & \boldsymbol{A}_2\boldsymbol{x} & & = \boldsymbol{b}_2, \\ & & \displaystyle\sum_{k=1}^{\ell} v_k & = a, \\ & \boldsymbol{x} & & \geq \boldsymbol{0}^n, \\ & & \boldsymbol{v} & \geq \boldsymbol{0}^\ell, \end{array}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{v} \in \mathbb{R}^\ell$, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{d} \in \mathbb{R}^\ell$, $\boldsymbol{A}_1 \in \mathbb{R}^{m_1 \times n}$, $\boldsymbol{A}_2 \in \mathbb{R}^{m_2 \times n}$, $\boldsymbol{B} \in \mathbb{R}^{m_1 \times \ell}$, $\boldsymbol{b}_1 \in \mathbb{R}^{m_1}$, $\boldsymbol{b}_2 \in \mathbb{R}^{m_2}$, and $a \in \mathbb{R}$.

State its LP dual problem.

(2p) b) Consider the strictly convex quadratic optimization problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := 2x_1^2 + x_2^2 - 4x_1 - 6x_2, && \text{(1a)} \\ \text{subject to} \quad & -x_1 + 2x_2 \leq 4. && \text{(1b)} \end{aligned}$$

For this problem, do the following:

[1] Explicitly state its Lagrangian dual function q and its Lagrangian dual problem, associated with the Lagrangian relaxation of the constraint (1b);

[2] Solve this Lagrangian dual problem and provide the optimal Lagrange multiplier μ^*;

[3] Provide the globally optimal solution $\boldsymbol{x}^*$ to the problem (1);

[4] Prove that strong duality holds, that is, prove that $q(\mu^*) = f(\boldsymbol{x}^*)$ holds.

Good luck!

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 04–03–08
Examiner: Michael Patriksson

Question 1

(the Simplex method)

a) By introducing slack variables we get the problem in standard form:

$$\begin{array}{lrrrrrl} \text{minimize} & z = & x_1 & +3x_2 & +x_3 & & & \qquad \text{(P)} \\ \text{subject to} & & -2x_1 & +5x_2 & -x_3 & -x_4 & & = 5, \\ & & 2x_1 & -x_2 & +2x_3 & & +x_5 & = 4, \\ & & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0. \end{array}$$

The Phase I problem becomes

$$\begin{array}{lrrrrrrl} \text{minimize} & w = & & & & & & a \\ \text{subject to} & & -2x_1 & +5x_2 & -x_3 & -x_4 & & +a = 5, \\ & & 2x_1 & -x_2 & +2x_3 & & +x_5 & = 4, \\ & & x_1, & x_2, & x_3, & x_4, & x_5, & a \geq 0. \end{array}$$

Start with the basis defined by $\boldsymbol{x}_B = (a, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, x_2, x_3, x_4)^{\mathrm{T}}$. The reduced costs of $\boldsymbol{x}_N$ become $(2, -5, 1, 1)$, so x_2 is the entering variable. The leaving variable becomes a. The new basis is given by $\boldsymbol{x}_B = (x_2, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, a, x_3, x_4)^{\mathrm{T}}$, and the reduced costs of $\boldsymbol{x}_N$ are $(0, 1, 0, 0)$, which means that the current basis is optimal to the Phase I problem and since $w^* = 0$ it follows that $\boldsymbol{x}_B = (x_2, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, x_3, x_4)^{\mathrm{T}}$ define a BFS to the Phase II problem (P). The reduced costs of $\boldsymbol{x}_N$ becomes $(2.2, 1.6, 0.6)^{\mathrm{T}} \geq \mathbf{0}^3$, which means that an optimal solution to (P) is given by

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} x_2 \\ x_5 \\ x_1 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence an optimal solution to the original problem is given by

$$\boldsymbol{x}^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

b) Since the reduced costs of $\boldsymbol{x}_N$ are all strictly positive, it follows that the BFS found is the unique optimal solution (see Proposition 10.9 in the course notes).

Question 2

(optimality conditions)

a) Drawing the figure one can verify that the problem is non-convex, because the feasible set is not convex (even though the objective function is). The optimization problem amounts to finding the shortest distance from the point $(x, y)^{\mathrm{T}} = (2, 1)^{\mathrm{T}}$ to the feasible set, and the geometrical considerations give us one local minimum $(x, y)^{\mathrm{T}} = (2, 0)^{\mathrm{T}}$ with the objective value $f((2, 0)^{\mathrm{T}}) = 1/2$ and a global minimum $(x, y)^{\mathrm{T}} = (3/2, 3/2)^{\mathrm{T}}$ with objective value $f((3/2, 3/2)^{\mathrm{T}}) = 1/4$.

Introducing the KKT-multipliers μ_1 and μ_2 for the inequality constraints, as well as λ for the equality constraint, the KKT system for this problem can be stated as follows:

$$\begin{cases} \begin{pmatrix} x-2 \\ y-1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} \mu_1 + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \mu_2 + \begin{pmatrix} y \\ x-2y \end{pmatrix} \lambda = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \hfill y - x \leq 0, \\ \hfill -y \leq 0, \\ \hfill y(x-y) = 0, \\ \hfill \mu_1, \mu_2 \geq 0, \\ \hfill \mu_1(x-y) = 0, \\ \hfill \mu_2 y = 0. \end{cases}$$

As it can be verified, this system gives two [in the space $(x, y)^{\mathrm{T}}$] KKT-points:

- The point of local minimum: $(x, y)^{\mathrm{T}} = (2, 0)^{\mathrm{T}}$, $\mu_1 = 0$, $\mu_2 \geq 0$, $2\lambda = 1 + \mu_2$.
- The point of global minimum: $(x, y)^{\mathrm{T}} = (3/2, 3/2)^{\mathrm{T}}$, $\mu_1 \geq 0$, $\mu_2 = 0$, $3\lambda = 1 + 2\mu_1$.

b) A simple calculation shows that the gradients of the free constraints are: $\nabla g_1(x, y) = (1, -1)^{\mathrm{T}}$, $\nabla g_2(x, y) = (0, 1)^{\mathrm{T}}$, $\nabla g_3(x, y) = (y, x - 2y)^{\mathrm{T}}$. At every feasible point we have either $y = 0$, which results in $\nabla g_2(x, y) = x \nabla g_3(x, y)$, or $x = y$, which results in $\nabla g_1(x, y) = y \nabla g_3(x, y)$. In either case, the LICQ is violated.

Again, from either geometrical or analytical considerations, we can split the feasible set of the original problem into two (non-disjoint) parts defined by linear constraints:

$$\mathcal{F}_1 = \{\, (x, y) \in \mathbb{R}^2 \mid y = 0, x - y \geq 0 \,\},$$

and

$$\mathcal{F}_2 = \{\, (x, y) \in \mathbb{R}^2 \mid y \geq 0, x - y = 0 \,\}.$$

We can therefore solve two convex linearly constrained optimization problems:

$$\begin{aligned} &\text{minimize} \;\; f(x, y), \\ &\text{subject to} \;\; (x, y) \in \mathcal{F}_1, \end{aligned}$$

and

$$\begin{aligned} &\text{minimize} \;\; f(x, y), \\ &\text{subject to} \;\; (x, y) \in \mathcal{F}_2, \end{aligned}$$

and choose the best solution among the two.

c) The procedure in the previous part can be generalized for problems with several complementarity constraints as follows. The feasible set can be split into 2^n parts $\mathcal{F}_I$, $I \subseteq \{1, \dots, n\}$, where

$$\begin{aligned} \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} = b_i, \text{and } x_i \geq 0, \quad i \in I, \\ \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} \geq b_i, \text{and } x_i = 0, \quad i \notin I. \end{aligned}$$

Therefore, instead of solving the origial non-convex problem, which violates the LICQ, one can (in principle) solve 2^n convex problems with linear constraints.

Question 3

(modelling)

Introduce variables according to Figure 1.

Introduce constraints according to the following list:

Maximum sales:

$$x_1 \leq 200, \quad x_2 \leq 100, \quad x_3 \leq 300. \tag{1}$$

Process balances, Machine 1:

$$y_1 \geq x_1, \quad y_2 \geq x_2, \quad y_3 \geq x_2, \quad y_4 \geq x_3. \tag{2}$$

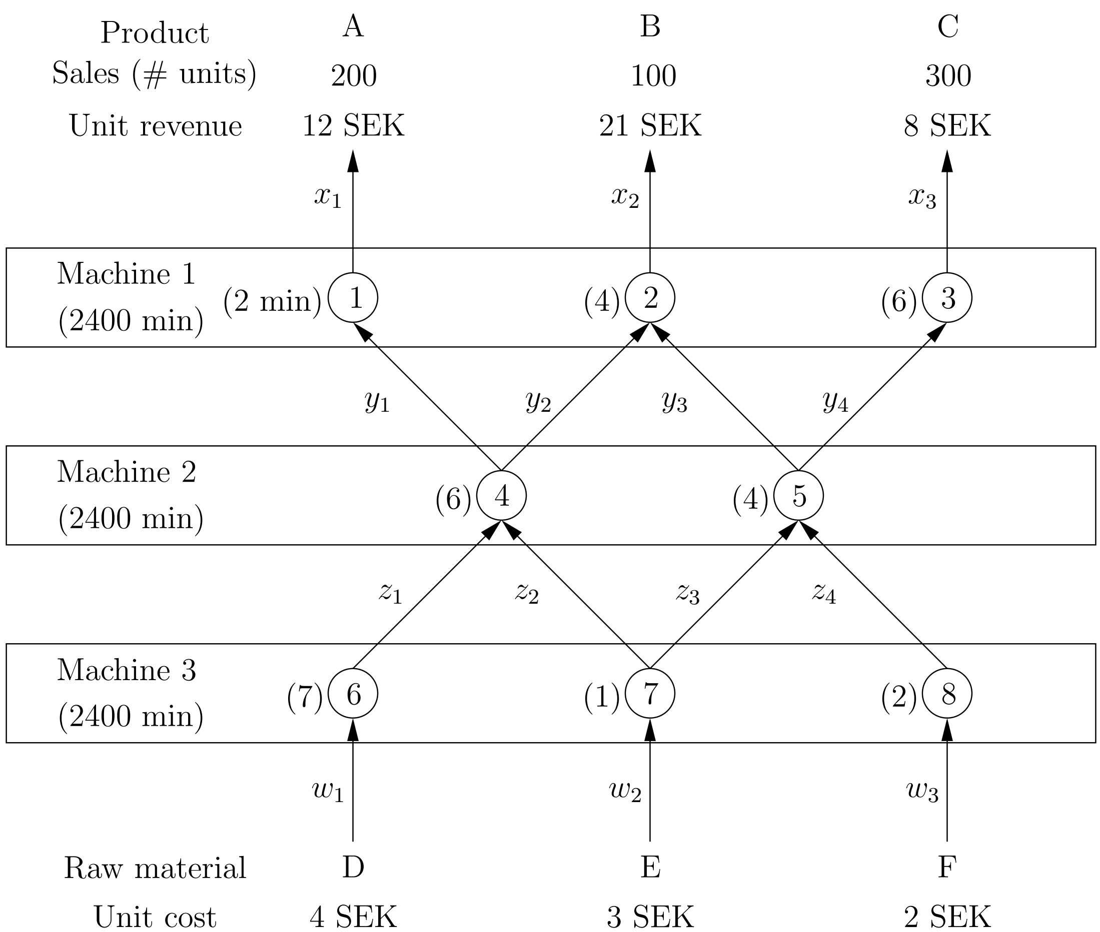

Figure 1: Variable definitions.

Process balances, Machine 2:

$$z_1 \geq y_1 + y_2, \quad z_2 \geq y_1 + y_2, \quad z_3 \geq y_3 + y_4, \quad z_4 \geq y_3 + y_4. \tag{3}$$

Process balances, Machine 3:

$$w_1 \geq z_1, \quad w_2 \geq z_2 + z_3, \quad w_3 \geq z_4. \tag{4}$$

Weekly capacity, Machine 1:

$$2x_1 + 4x_2 + 6x_3 \leq 2400. \tag{5}$$

Weekly capacity, Machine 2:

$$6(y_1 + y_2) + 4(y_3 + y_4) \leq 2400. \tag{6}$$

Weekly capacity, Machine 3:

$$7z_1 + (z_2 + z_3) + 2z_4 \leq 2400. \tag{7}$$

Objective function:

$$f(\boldsymbol{x}, \boldsymbol{w}) = 12x_1 + 21x_2 + 8x_3 - 4w_1 - 3w_2 - 2w_3.$$

We end up with the linear integer program

$$\begin{aligned} \text{maximize} \quad & f(\boldsymbol{x}, \boldsymbol{w}), \\ \text{subject to} \quad & (1) - (7), \\ & \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w} \geq \boldsymbol{0} \text{ and integer.} \end{aligned}$$

Question 4

(applications of the Newton algorithm)

a) The objective function $f(x) = ax - \log(x)$ is strictly convex inside the feasible set $\{ x \in \mathbb{R} \mid x > 0 \}$, since $f''(x) = 1/x^2 > 0$ there; therefore, every local minimum in this problem is also a global one, and the global minimum is unique, provided any exists. Now we can test the necessary (and sufficient in this case, owing to the convexity) optimality conditions

$$\begin{aligned} f'(x) = a - x^{-1} &= 0, \\ x &> 0, \end{aligned}$$

which is uniquely solvable, giving us $x^* = a^{-1} > 0$.

b) Direct calculations show that

$$x_{k+1} = x_k - f'(x_k)/f''(x_k) = x_k(2 - ax_k),$$

which does not involve any divisions.

Assuming that $x_k \to \bar{x}$ (and thus also $x_{k+1} \to \bar{x}$) gives us

$$\bar{x} = \bar{x}(2 - a\bar{x}),$$

which has two solutions: $\bar{x}_1 = a^{-1}$ or $\bar{x}_2 = 0$. It is the latter solution that is not a global/local optimum of the original problem (it is not even feasible, to start with). One can easily obtain this solution by starting from the point $x_0 = 2/a > 0$, which generates $x_1 = 0$, and thus $x_k = 0$ for all $k \geq 1$.

c) One can for example start from the optimality conditions

$$\begin{aligned} g'(x) = a - x^{-2} = 0, \\ x > 0, \end{aligned}$$

to end up with the strictly convex minimization problem to

$$\begin{aligned} &\text{minimize} \;\; g(x) = ax + x^{-1}, \\ &\text{subject to} \;\; x > 0. \end{aligned}$$

It is verified as in b) that Newton's method for this problem involves only simple operations (additions/subtractions and multiplications).

Question 5

(optimality conditions)

Farkas' Lemma can be stated as follows:

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{b}$ an $m \times 1$ vector. Then exactly one of the systems

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ \boldsymbol{x} \geq \boldsymbol{0}^n, \end{aligned} \tag{I}$$

and

$$\begin{aligned} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{0}^n, \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} > 0, \end{aligned} \tag{II}$$

has a feasible solution, and the other system is inconsistent.

a) Farkas' Lemma is proved in Theorem 11.10.

b) At $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$, the cone of feasible directions is

$$\begin{aligned} R_S(\bar{\boldsymbol{x}}) &= \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid 2p_1 - p_2 = 0;\ \boldsymbol{p} \geq \boldsymbol{0}^2 \,\} \\ &= \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid 2p_1 - p_2 \leq 0;\ -2p_1 + p_2 \leq 0;\ -p_1 \leq 0;\ -p_2 \leq 0 \,\}. \end{aligned}$$

At $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$, the cone of descent directions is

$$\overset{\circ}{F}(\bar{\boldsymbol{x}}) = \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{p} < 0 \,\} = \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid p_1 + p_2 > 0 \,\}.$$

To prove that the set $R_S(\bar{\boldsymbol{x}}) \cap \overset{\circ}{F}(\bar{\boldsymbol{x}})$ is non-empty (that is, that there exists a feasible descent direction), we define

$$\boldsymbol{A} := \begin{pmatrix} 2 & -2 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The consistency of the system (II) then is equivalent to the existence of a feasible descent direction (with $\boldsymbol{p} = \boldsymbol{y}$). We therefore need to establish that the system (I) is inconsistent. The consistency of this system is equivalent to the possibility to choose a non-negative $\boldsymbol{x} \in \mathbb{R}^4$ such that

$$\begin{pmatrix} 2 & -2 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This is however impossible. (One way to check this is via Phase I in the Simplex method.)

We are done.

Question 6

(convexity)

The proof of Carathéodory's Theorem can be found in Theorem 3.8 in the Course Notes.

Question 7

(duality in linear and nonlinear optimization)

a) The LP dual is to

$$\begin{array}{llllll} \text{maximize} & w = \boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1 & +\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2 & +ay_3 & \\ \text{subject to} & \boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{y}_1 & +\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{y}_2 & & \leq \boldsymbol{c}, \\ & \boldsymbol{B}^{\mathrm{T}}\boldsymbol{y}_1 & & +\boldsymbol{1}^{\ell}y_3 & \leq \boldsymbol{d}, \\ & \boldsymbol{y}_1 \geq \boldsymbol{0}^{m_1}, & \boldsymbol{y}_2 \in \mathbb{R}^{m_2}, y_3 \in \mathbb{R}, & & \end{array}$$

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

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 $(\boldsymbol{x}(\mu^*), \mu^*)$.