Exercises in SF1811 Optimization

2014

Division of Optimization and Systems Theory
Department of Mathematics
Kungliga Tekniska Högskolan
Stockholm, Sweden

Contents

Exercises

1. Linear programming

1.1 (20070601-nr.1a)

A company manufactures the three products: A,B and C.
The manufacturing process consists of the moments *cutting* and *pressing*.
Every product has to pass *both* moments.

The department of cutting, which can be used 8 hours per day has the following capacity:
2000 units per hour of product A or
1600 units per hour of product B, or
1100 units per hour of product C.
The production can be switched from one product to another without problems.

The department of pressing, which can be used 8 hours per day has the following capacity:
1000 units per hour of product A, or
1500 units per hour of product B, or
2400 units per hour of product C.
The production can be switched from one product to another without problems.

The coverage contribution (income minus moving cost) per manufactured unit of the product are:
12 SEK for A, 9 SEK for B and 8 SEK for C.

The company now wants to decide how many units of each product that (in average) should be produced to make the total coverage contribution as large as possible, without breaking the capacity constraints of the sections.

Your assignment is to *formulate* the company's problem as a LP-problem. You don't have to calculate the optimal solution of this LP-problem.

1.2 (20070601-nr.2)

Consider the following LP-problem::

$$
\begin{array}{rrcrcrcrcl}
\text{minimize} & -3x_1 & + & 4x_2 & - & 2x_3 & + & 5x_4 & & \\
\text{s.t.} & x_1 & + & x_2 & - & x_3 & - & x_4 & \leq & 8, \\
 & x_1 & - & x_2 & + & x_3 & - & x_4 & \leq & 4, \\
 & x_1 & & & & & & & \geq & 0, \\
 & & & x_2 & & & & & \geq & 0, \\
 & & & & & x_3 & & & \geq & 0, \\
 & & & & & & & x_4 & \geq & 0.
\end{array}
$$

(a) Transform the problem to standard form using the two slack variables x_5 and x_6. Then solve the problem using the simplex method. Start with the introduced slack variables as basic variables. .... (4p)

(b) Assume that the objective function coefficient corresponding to x_4 is changed from 5 to 2. Use the simplex method on this modified problem. It is allowed to start from the final solution from the (a)-task. Explain what happens. .... (4p)

(c) Formulate the dual LP-problem corresponding to the problem on standard form (with 6 variables) from the (a)-task. Visualize the feasible region to this dual problem in a figure with the dual variables y_1 and y_2 on the axis. .... (2p)

(d) What does the corresponding figure look like for the dual LP-problem on standard form from the (b)-task? Comment on the figure!

1.3 (20070307-nr.1a)

The Cidermans family business produce and sell four different types of cider: Apple cider, Pear cider, Mixed Cider and Standard Cider.
Every hectoliter cider requires p working hours for the production and q working hours for packaging. The economic profit of the cider for the Cidermans is v SEK/hektoliter.

p, q and v has different values for the different types of cider according to the following table:

	p	q	v
Apple cider	1.6	1.2	196
Pear cider	1.8	1.2	210
Mixed Cider	3.2	1.2	280
Standard Cider	5.4	1.8	442

A normal week the company has 80 hours (two family members) to spend on production and 40 hours (one family member) on packaging.

Further, they have decided that the Apple cider shall be *at least* 20% of the produced volume and that the Pear cider shall be *at most* 30% of the total volume produced cider.

The question is how much of each sort of cider that is to be produced so that the profit of the Cidermans is maximized given the constraints above. The cider is very popular and they can sell everything that they produce without problems.

Your assignment is to *formulate* the Cidermans problem as a LP-problem. However you do *not* have to compute the optimal solution.

1.4 (20070307-nr.2)

(a) The following system consists of two linear equations and four linear inequalities.

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 10, \\ 2x_1 & + & 3x_2 & + & 4x_3 & + & 5x_4 & = & 12, \\ x_1 & & & & & & & \geq & 0, \\ & & x_2 & & & & & \geq & 0, \\ & & & & x_3 & & & \geq & 0, \\ & & & & & & x_4 & \geq & 0. \end{array}$$

To investigate systematically whether this system has a feasible solution or not, you can form the following LP-problem with the two "artificial" variables x_5 and x_6.

$$\begin{array}{lrcrcrcrcrcrcl} \text{minimize} & & & & & & & & & x_5 & + & x_6 & & \\ \text{s.t.} & x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & + & x_5 & & & = & 10, \\ & 2x_1 & + & 3x_2 & + & 4x_3 & + & 5x_4 & & & + & x_6 & = & 12, \\ & x_1 & & & & & & & & & & & \geq & 0, \\ & & & x_2 & & & & & & & & & \geq & 0, \\ & & & & & x_3 & & & & & & & \geq & 0, \\ & & & & & & & x_4 & & & & & \geq & 0, \\ & & & & & & & & & x_5 & & & \geq & 0, \\ & & & & & & & & & & & x_6 & \geq & 0. \end{array}$$

Your task is to first solve this LP-problem with the simplex method and then answer the question whether there exist a solution to the original problem above. Motivate your answer carefully! ....(6p)

(b) The following two LP-problems (which have their origin in a certain two-person Zero-sum game) are each-others duals. This you do not have to show.

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{s.t.} & -x_1 + 2x_2 + x_3 \geq 0 \\ & 3x_1 - 4x_2 + x_3 \geq 0 \\ & x_1 + x_2 \qquad = 1 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \text{ free.} \end{array}$$

$$\begin{array}{ll} \text{maximize} & y_3 \\ \text{s.t.} & -y_1 + 3y_2 + y_3 \leq 0 \\ & 2y_1 - 4y_2 + y_3 \leq 0 \\ & y_1 + y_2 \qquad = 1 \\ & y_1 \geq 0,\ y_2 \geq 0,\ y_3 \text{ free.} \end{array}$$

The primal problem has been solved and the optimal solution
$\hat{x}_1 = 0.6$, $\hat{x}_2 = 0.4$, $\hat{x}_3 = -0.2$
has been obtained. Use this information to (with optional method) obtain an optimal solution $\hat{\mathbf{y}}$ to the dual problem.

1.5 (20060603-nr.1)

Consider the following LP-problem on standard form:

$$\begin{array}{rrcrcrcrcrcl}
\text{minimize} & 4x_1 & + & 3x_2 & + & 2x_3 & + & 3x_4 & + & 4x_5 & & \\
\text{s.t.} & 4x_1 & + & 3x_2 & + & 2x_3 & + & x_4 & & & = & 5, \\
 & & & x_2 & + & 2x_3 & + & 3x_4 & + & 4x_5 & = & 3, \\
 & x_1 & & & & & & & & & \geq & 0, \\
 & & & x_2 & & & & & & & \geq & 0, \\
 & & & & & x_3 & & & & & \geq & 0, \\
 & & & & & & & x_4 & & & \geq & 0, \\
 & & & & & & & & & x_5 & \geq & 0.
\end{array}$$

(a) Use the Simplex method to determine an optimal solution to the problem. Start with x_1 and x_5 as basic variables. .... (5p)

(b) The optimal solution is not unique. Determine another optimal basic solution than the one you obtained in the (a)-task. .... (2p)

(c) Formulate the corresponding dual problem and state an optimal solution to it. Illustrate also the dual problem graphically in a figure with y_1 and y_2 on the coordinate axis. .... (3p)

1.6 (20060308-nr.1)

Consider the following linear optimization problem:

$$\begin{array}{rl}
\text{minimize} & x_1 + 5x_2 + 2x_3 \\
\text{s.t.} & x_1 + x_2 \geq 2, \\
 & x_1 + x_3 \geq 2, \\
 & x_2 + x_3 \geq 2, \\
 & x_j \geq 0, \quad j = 1, 2, 3.
\end{array}$$

(a) Use the simplex method to determine an optimal solution. You must start from the basic solution $x_1 = x_2 = x_3 = 1$ (which is feasible but not optimal). .... (7p)

(b) Formulate the corresponding dual LP-problem and state an optimal solution to it. Verify in particular that the objective values are equal. (3p)

Optional help: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$

1.7 (20060308-nr.4b)

Assume that $\mathbf{a}_1, \ldots, \mathbf{a}_m$ are given vectors in $I\!R^3$ and that $b_1, \ldots, b_m$ are given positive numbers (i.e. $b_i > 0$).

Let $\Omega = \{\, \mathbf{x} \in I\!R^3 \mid \mathbf{a}_i^{\mathsf{T}} \mathbf{x} \leq b_i, \text{ for } i = 1, \ldots, m \,\}$.

The set Ω is a region (in $I\!R^3$) whose "walls" are formed by planes on the form $\mathcal{P}_i = \{\, \mathbf{x} \in I\!R^3 \mid \mathbf{a}_i^{\mathsf{T}} \mathbf{x} = b_i \,\}$.

Now assume that you want to determine the center point and radius to the *biggest sphere* that is contained in the set Ω.

Formulate this as a LP-problem! .... (6p)

1.8 (20051024-nr.1)

Consider the following LP-problem. (Observe that it is a maximization problem).

$$\begin{aligned} \text{maximize} \quad & \mathbf{q}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{P}\mathbf{x} \le \mathbf{b}, \\ & \mathbf{x} \ge \mathbf{0}, \end{aligned}$$

where $\mathbf{P} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{q} = (1,\, 1,\, 2)^{\mathsf{T}}$.

(a) Solve the problem with the simplex method. Start with the slack variables in the basis. .... (7p)

(b) Determine two vectors $\mathbf{x}_0 \in I\!R^3$ and $\mathbf{d} \in I\!R^3$ such that if you let $\mathbf{x}(t) = \mathbf{x}_0 + t \cdot \mathbf{d}$, with $t \in I\!R$, then it holds both that $\mathbf{x}(t)$ is a feasible solution to the problem for every $t > 0$, and that $\mathbf{q}^{\mathsf{T}}\mathbf{x}(t) \to +\infty$ s.t. $t \to +\infty$. (1p)

(c) Formulate the corresponding dual LP-problem and determine with optional method (e.g. graphically) if it has any feasible solutions. Comment. (2p)

1.9 (20051024-nr.3)

In a certain city there is a subway-line with 12 stations. One year ago they did a careful investigation about how many commuters that are going between different pairs of stations. They computed r_{ij} = the average number of commuters per day that use the subway to go between station i to station j (i.e. enters at station i and leaves at station j). This was done for every pair (i, j) with $i \in \{1, \ldots, 12\}$, $j \in \{1, \ldots, 12\}$ and $i \neq j$.

But now one year has passed since the investigation, and many have changed residence and/or place where they work, and they are to update the the above mentioned investigation about traveling in the subway. However, they don't want to do this investigation as extensive as one year ago, but only for every $i \in \{1, \ldots, 12\}$ measure p_i = the average number of commuters per day that *enters* the subway at station i, and for every $j \in \{1, \ldots, 12\}$ measure q_j = the average number of commuters per day that *leaves* the subway at station j.

Surely the result will be that the one year old numbers r_{ij} are not consistent with these measured p_i and q_j, and hence one should change the r_{ij}:s to new numbers x_{ij} that are both consistent to the numbers p_i and q_j, and differ "as little as possible" from the old numbers r_{ij}.

Your task is to formulate the problem to determine such estimations x_{ij} (of the average number of commuters per day that goes from station i to station j) as an optimization problem on suitable form!

Note that there is no unambiguous solution to this exercise since it is not self-evident what is meant by "as little as possible".

To obtain full point you shall select a *LP-formulation* that catches the problem in a reasonable way. A convex QP-formulation can also give a decent number of points, as can also other reasonable formulations as convex optimization problems. ....(10p)

1.10 (20050331-nr.2)

It is well known that the following two LP-problems P and D are each-others duals.

$$\text{P:} \quad \begin{array}{rl} \text{minimize} & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} \geq \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{array} \qquad \text{D:} \quad \begin{array}{rl} \text{maximize} & \mathbf{b}^{\mathsf{T}}\mathbf{y} \\ \text{s.t.} & \mathbf{A}^{\mathsf{T}}\mathbf{y} \leq \mathbf{c}, \\ & \mathbf{y} \geq \mathbf{0}. \end{array}$$

In this exercise the connection between the optimal values of the problems is illustrated (which sometimes can be $+\infty$ or $-\infty$).

Suppose that $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Solve *graphically* the above problems P and D, state their optimal values and, if there are any, their optimal solutions.

(a) When $\mathbf{b} = (\,1,\, -1\,)^{\mathsf{T}}$ and $\mathbf{c} = (-2,\; 2\,)^{\mathsf{T}}$. ....(2p)

(b) When $\mathbf{b} = (\,1,\, -1\,)^{\mathsf{T}}$ and $\mathbf{c} = (\,2,\; 2\,)^{\mathsf{T}}$. ....(2p)

(c) When $\mathbf{b} = (-1,\, -1\,)^{\mathsf{T}}$ and $\mathbf{c} = (-2,\; 2\,)^{\mathsf{T}}$. ....(2p)

(d) When $\mathbf{b} = (-1,\; 1\,)^{\mathsf{T}}$ and $\mathbf{c} = (\,2,\; 2\,)^{\mathsf{T}}$. ....(2p)

(e) When $\mathbf{b} = (-1,\, -1\,)^{\mathsf{T}}$ and $\mathbf{c} = (\,2,\, -2\,)^{\mathsf{T}}$. ....(2p)

To solve a problem graphically means in this case to draw the feasible region and the curvature of the objective function, and from this make conclusions about the optimal value of the problem and the, in case there is one, optimal solution. Since both P and D are to be solved in each part of the exercise it will in total be 10 figures (but each individual figure is fast to draw).

1.11 (20050307-nr.2)

Consider the following LP-problem which we denote P.

$$\begin{array}{rl} \text{minimize} & z = \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{array}$$

where $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 2 & 2 & 1 & 0 & 2 \\ 2 & 0 & 2 & 1 & 2 & 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$, $\mathbf{c}^{\mathsf{T}} = \begin{pmatrix} 2 & 1 & 1 & 1 & 2 & 1 \end{pmatrix}$.

(a) Suppose that x_1, x_3 and x_5 are chosen as basic variables. Determine the corresponding basic solution, and show that this is a feasible basic solution to P. .... (2p)

(b) The basic solution above is not optimal (right?). Perform an iteration with the simplex method and determine a new, (better) feasible basic solution. Then verify that you found an optimal solution to the problem. (6p)

(c) Formulate the dual problem D corresponding to P and denote an optimal solution to D. Verify specifically that the optimal values of P and D are equal. .... (2p)

1.12 (20050307-nr.5)

Given are two sets of points $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \ldots, \mathbf{q}_\ell\}$ in $I\!R^n$. The $k + \ell$ points $\mathbf{p}_i \in I\!R^n$ and $\mathbf{q}_j \in I\!R^n$ are hence given, and we assume that $\mathbf{p}_1 = \mathbf{0}$.

Now you want to determine whether there is any vector $\mathbf{a} \in I\!R^n$ with non-negative components (i.e. $\mathbf{a} \geq \mathbf{0}$) such that the hyper-plane $\{\mathbf{x} \in I\!R^n \mid \mathbf{a}^\mathsf{T}\mathbf{x} = 1\}$ *separates* the sets of points (so that all points in $\mathcal{P}$ are on one side of the plane while all points in $\mathcal{Q}$ are on the other side of the plane).

(a) First assume that both points in $\mathcal{P}$ and in $\mathcal{Q}$ are permitted to lie *on* the separating plane, i.e. we seek a plane that "weakly" separates $\mathcal{P}$ and $\mathcal{Q}$.

Formulate the problem to determine a weakly separating plane on the form above (i.e. $\mathbf{a}^\mathsf{T}\mathbf{x} = 1$ with $\mathbf{a} \geq \mathbf{0}$) as a LP-problem on standard form. State specifically a feasible starting basic solution to your LP-problem. Also describe how you from the optimal solution to the LP-problem can determine whether there is any weakly separating plane on the form above. .... (5p)

(b) Now suppose that neither points in $\mathcal{P}$ nor points in $\mathcal{Q}$ are allowed to be on the separating plane, i.e. we search a plane that "'strictly"' separates $\mathcal{P}$ and $\mathcal{Q}$.

Formulate the problem to determine a strictly separating plane on the form above (i.e. $\mathbf{a}^\mathsf{T}\mathbf{x} = 1$ with $\mathbf{a} \geq \mathbf{0}$) as a LP-problem on standard form. State specifically a feasible starting basic solution to your LP-problem. Also describe how you from the optimal solution to the LP-problem can determine whether there is any strictly separating plane on the form above. .... (5p)

1.13 (20041016-nr.2)

The five following vectors in $I\!R^3$ are given:

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \ \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{a}_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \ \mathbf{a}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}.$$

You want to determine whether there are any *non-negative* scalars x_j such that
$\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 + \mathbf{a}_4 x_4$.

Therefore you form the following LP-problem in the seven variables $\mathbf{x} = (x_1, x_2, x_3, x_4)^\mathsf{T}$ and $\mathbf{v} = (v_1, v_2, v_3)^\mathsf{T}$:

$$\begin{aligned} \text{minimize} \quad & \mathbf{e}^\mathsf{T}\mathbf{v} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{I}\,\mathbf{v} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \ \text{ and } \ \mathbf{v} \geq \mathbf{0}, \end{aligned}$$

where $\mathbf{A}$ is a 3×4 matrix with the vectors above, $\mathbf{a}_j$, as columns,
$\mathbf{I}$ is a 3×3 unitary matrix and $\mathbf{e}^\mathsf{T} = (1,\ 1,\ 1)$.

(a) Show that an optimal solution to the LP-problem above is given by $\mathbf{x} = (2,\ 5,\ 0,\ 0)^\mathsf{T}$ and $\mathbf{v} = (0,\ 0,\ 1)^\mathsf{T}$. .... (5p)

(b) Are there any scalars $x_j \geq 0$ such that $\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 + \mathbf{a}_4 x_4$? Motivate your answer carefully. .... (2p)

(c) Is there any vector $\mathbf{y} \in I\!\!R^3$ such that $\mathbf{b}^\mathsf{T}\mathbf{y} > 0$ and $\mathbf{a}_j^\mathsf{T}\mathbf{y} \leq 0$ for all j ? Determine in that case such a vector $\mathbf{y}$. .... (3p)
(*Help:* Consider the dual problem to the LP-problem above.)

1.14 (20041504-nr.2)

You have downloaded a program from the net of unknown quality to solve LP-problems of the form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

You test the program with the following data:

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 & 3 & 3 & 2 \\ 2 & 4 & 2 & 1 & 2 & 1 \\ 1 & 2 & 3 & 2 & 3 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 14 \\ 16 \\ 10 \end{pmatrix}, \quad \mathbf{c}^\mathsf{T} = (\,2,\ 3,\ 2,\ 2,\ 3,\ 2\,).$$

Then the program prints the following:

"An optimal solution to the problem is $\mathbf{x} = (\,3,\ 2,\ 1,\ 0,\ 0,\ 0\,)^\mathsf{T}$, and an optimal
solution to the corresponding dual problem is $\mathbf{y} = (0.25,\ 0.50,\ 0.25)^\mathsf{T}$".

(a) Verify that the result of the program is correct. .... (6p)

(b) Assume that the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ above are changed to the constraints $\mathbf{A}\mathbf{x} \geq \mathbf{b}$.
Determine an optimal solution to this new problem. .... (2p)

(c) Now suppose that the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ above are changed to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$.
Determine an optimal solution to this new problem. .... (2p)

1.15 (20040310-nr.5)

Consider the following LP-problem with 101 variables:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 100 & 99 & 98 & \cdots & 51 & 50 & 49 & \cdots & 2 & 1 & 0 \\ 0 & 1 & 2 & \cdots & 49 & 50 & 51 & \cdots & 98 & 99 & 100 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 100 \\ 200 \end{pmatrix}$

and $\mathbf{c}^{\mathsf{T}} = (\ 50 \quad 49 \quad 48 \quad \cdots \quad 1 \quad 0 \quad 1 \quad \cdots \quad 48 \quad 49 \quad 50\ )$.

The j:th column in $\mathbf{A}$ is hence given by $\mathbf{a}_j = (101–j\ ,\ j–1)^{\mathsf{T}}$, for $j = 1, \ldots, 101$, while the j:th component in $\mathbf{c}$ is given by $c_j = |\,j - 51\,|$ (absolute norm).

(a) Start with the variables x_1 and x_{101} in the basis, and perform one iteration with
the simplex method. Determine whether the new basic solution you obtain
is an optimal solution to the problem or not. .... (4p)

(b) How many feasible basic solutions does this LP-problem have? .... (3p)

(c) How many of these basic solutions are optimal solutions? .... (3p)

2. Network problems

2.1 (20070601-nr.1b)

Here a so called balanced transportation problem, with four factories and four costumers, is studied. I.e. a problem on the form

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij}x_{ij} \\
\text{s.t.} \quad & \sum_{j=1}^{4} x_{ij} = s_i\,, \quad \text{for } i = 1,\dots,4 \\
& \sum_{i=1}^{4} x_{ij} = d_j\,, \quad \text{for } j = 1,\dots,4 \\
& x_{ij} \geq 0\,, \quad \text{for all } i \text{ and } j,
\end{aligned}$$

where
s_i = given supply at factory i,
d_j = given demand at customer j,
c_{ij} = given transportation cost per unit from the factory i to customer j,
x_{ij} = number of units transported from factory i to customer j.

Assume that the supply at the factories and the demand at the customers is given by
$s_1 = 40,\ s_2 = 30,\ s_3 = 20,\ s_4 = 10,\ d_1 = 10,\ d_2 = 20,\ d_3 = 30,\ d_4 = 40,$
and the transportation costs are given by the table:

c_{ij}	customer 1	customer 2	customer 3	customer 4
factory 1	116	125	136	149
factory 2	109	116	125	136
factory 3	104	109	116	125
factory 4	101	104	109	116

With the help of the "North West Corner"-rule the following feasible basic solution is obtained:

x_{ij}	customer 1	customer 2	customer 3	customer 4	s_i
factory 1	10	20	10	0	40
factory 2	0	0	20	10	30
factory 3	0	0	0	20	20
factory 4	0	0	0	10	10
d_j	10	20	30	40	

Decide whether this is an optimal solution to the problem or not.

2.2 (20070307-nr.1b)

Consider the LP-problem

$$\begin{aligned}
\text{minimize} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\
& \mathbf{x} \geq \mathbf{0},
\end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 3 \\ 5 \\ 7 \\ -2 \\ -4 \\ -9 \end{pmatrix}$

and $\mathbf{c}^{\mathsf{T}} = \begin{pmatrix} 2 & 3 & 4 & 3 & 3 & 4 & 3 & 2 & 4 \end{pmatrix}$.

From the special form of the matrix $\mathbf{A}$ it follows that the problem in reality is a minimum cost flow problem. Draw the corresponding network and verify thereafter that $\hat{\mathbf{x}} = (2\,,0\,,1\,,0\,,0\,,5\,,0\,,4\,,3)^{\mathsf{T}}$ is an optimal solution to the problem. ....(5p)

2.3 (20060603-nr.5)

A company has agreed on supplying p_1, p_2 and p_3 tonnes of a certain product (to an important costumer) at the end of each of the following three months. p_1, p_2 and p_3 are given constants.

Every month the company can manufacture at most a ton to the cost of c SEK/tonne. By using overtime they can manufacture another at most b tonnes per month to the cost of d SEK/tonne. a, b, c and d are given constants with $a > b$ and $d > c$.

The quantities of the product that is manufactured of the product, but are not needed for delivery the same month, can be stored for delivery another month. The storage cost is ℓ SEK per tonne and month that you store. ℓ is a given constant.

If the company does not supply the agreed quantity a certain month, they can instead deliver the missing quantity at a later month, but no later than the third and last month. The agreed fee for being late is f SEK per tonne and month that you are late. f is a given constant.

In the beginning of month 1 the storage is empty, and you don't want anything in the storage after the three months. We can assume that $p_1+p_2+p_3 < 3a+3b$.

Formulate the company's planning problem, in which the costs of the company is to be minimized, as an optimization problem of adequate form.

A totally correct formulation gives 8 points.
A totally correct formulation as a minimum cost flow problem gives 10 points.

2.4 (20060308-nr.3)

Consider a balanced transportation problem with 4 suppliers and 4 customers:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij}x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{4} x_{ij} \;=\; s_i\,, \quad \text{for } i = 1,\dots,4 \\ & \sum_{i=1}^{4} x_{ij} \;=\; d_j\,, \quad \text{for } j = 1,\dots,4 \\ & x_{ij} \;\geq\; 0\,, \qquad \text{for all } i \text{ and } j, \end{aligned}$$

where s_i = supply at supplier i, d_j = demand at customer j,
c_{ij} = the transportation cost per unit from supplier i to customer j.

Suppose supply and demand are given by

$s_1 = 80,\ \ s_2 = 60,\ \ s_3 = 40,\ \ s_4 = 20,$

$d_1 = 20,\ \ d_2 = 40,\ \ d_3 = 60,\ \ d_4 = 80,$

and that the transportation costs are given by the table:

c_{ij}	customer 1	customer 2	customer 3	customer 4
supp 1	16	25	36	49
supp 2	9	16	25	36
supp 3	4	9	16	25
supp 4	1	4	9	16

(a) Determine a feasible basic solution with the “North West Corner”–rule. (1p)

(b) Show that the solution you obtained in the (a)-task above turns out to be an optimal solution. (If you do not know the “North West Corner”–rule you may solve the problem starting from an optional basic solution.) (5p)

(c) Assume that both s_4 and d_1 are changed from 20 to 40. Determine an optimal solution to this new problem. Motivate your answer. ......(2p)

(d) Restore s_4 and d_1 to 20. Assume that c_{22} is decreased from 16 to $16-\delta_{22}$, while the other c_{ij} are unchanged. For which values on δ_{22} is it valid that the optimal solution from the (b)-task above is still optimal? ......(2p)

2.5 (20051024-nr.2)

A given directed network has the node set $\mathcal{N} = \{1,2,3,4,5,6\}$ and edge set $\mathcal{B} = \{(1,2),\ (1,3),\ (2,3),\ (2,4),\ (3,4),\ (3,5),\ (4,5),\ (4,6),\ (5,6)\}$ (directed edges).
The network has two source nodes, node 1 with the supply 25 units and node 2 with the supply 10 units, and two sink nodes, node 5 with the demand 15 units and node 6 with the demand 20 units. The nodes 3 and 4 are intermediate nodes, with neither supply nor demand. The flow cost c_{ij}, in kSEK per unit flow, for respective edge (i,j) in the network are according to the following:
$c_{12} = 3,\ \ c_{13} = 2,\ \ c_{23} = 1,\ \ c_{24} = 4,\ \ c_{34} = 4,\ \ c_{35} = 4,\ \ c_{45} = 1,\ \ c_{46} = 2,$
$c_{56} = 3.$

(a) Determine a flow of minimal cost that fulfills the constraints on supply and demand according to above. Start from the following (natural) feasible basic solution:
$x_{12} = 10,\ \ x_{13} = 15,\ \ x_{24} = 20,\ \ x_{35} = 15,\ \ x_{46} = 20$, the other $x_{ij} = 0$. (8p)

(b) Formulate the dual LP-problem corresponding to the minimum cost flow problem above. Determine a solution to this dual problem and verify that the optimal values are equal (for the dual and the primal). ....(2p)

2.6 (20050331-nr.1)

A company has two factories, here called F1 and F2, and three big customers here called K1, K2 and K3. All transports from plants to customers are through any of the company's reloading terminals, called T1 and T2.

Since factories, terminals and customers are spread out over the country, the costs for the transports between different units are different.

The transportation costs from the factories to the terminals and from the terminals to the customers, in the unit 100 SEK per tonne, are given by the following table:

	$T1$	$T2$
$F1$	7	6
$F2$	4	5

	$K1$	$K2$	$K3$
$T1$	6	7	7
$T2$	6	9	5

A specific week the demand by each of the three customers is 200 tonnes of the company's product. The company's supply of the product the specific week is 300 tonnes in each of the two plants.

The head of the company's transport division has proposed the following transport plan, in the unit tonnes.

	$T1$	$T2$
$F1$	0	300
$F2$	200	100

	$K1$	$K2$	$K3$
$T1$	0	200	0
$T2$	200	0	200

Your exercise, as a contracted optimization consultant, is to decide whether the proposed plan is optimal from the transportation point of view. If that is not the case, you should produce an optimal plan.

Note that you must use a general method that is useful also on bigger problems of this type. .... (10p)

2.7 (20050307-nr.1)

(a) A network with capacities has the set of nodes $\mathcal{N} = \{1,2,3,4,5,6,7,8\}$ and the set of edges $\mathcal{B} = \{(1,2),(1,3),(1,4),(2,6),(3,5),(3,7),(4,6),(5,8),(6,8),(7,8)\}$. Each of these 10 edges in $\mathcal{B}$ has the capacity $k_{ij} = 1$.

Determine a maximum flow from node 1 to node 8 in this network! Use the maximum flow algorithm and start from the zero flow (i.e. all $x_{ij} = 0$).

Determine also a cut of minimum capacity that separates node 1 and 8. Verify that the capacity of the minimum cut = the value of the maximum flow. .... (5p)

(b) This (b)-task is independent of the (a)-task above.

The owner of a street-stand is about to make a weekly schedule for his seven part-time employed technology students. Every day of the week two students shall work in the stand, and every student shall work two times a week. To meet the wishes of the employees each of the employees has stated which days they want to work. These are the wishes:

Student 1: Monday, Saturday, Sunday.
Student 2: Tuesday, Saturday, Sunday.
Student 3: Wednesday, Saturday, Sunday.
Student 4: Thursday, Saturday, Sunday.
Student 5: Friday, Saturday, Sunday.
Student 6: Monday, Tuesday, Wednesday, Thursday, Friday.
Student 7: Monday, Tuesday, Wednesday, Thursday, Friday.

The owner shall now intent to make a schedule where each student gets to work two of the days he/she stated.

Your task is to *formulate* the scheduling problem of the owner as a maximum flow problem in a certain network. Explain and motivate carefully. You do *not* need to compute the maximum flow, but you should describe how you given the maximum flow determine whether there exists a schedule that fulfills the wishes or not. . . . . (5p)

2.8 (20041016-nr.1)

An airport operator will under a period of time maintain four airports with fuel. The demand for fuel is estimated to the following:

Airport	1	2	3	4
Demand (tonnes)	300	300	300	300

They have received offers from three different fuel distributing companies regarding the total transport capacity and prices for delivery to the respective airports for the period:

Supplier	1	2	3
Capacity (tonnes)	400	400	400

Airport	1	2	3	4
Sup. 1 (SEK/tonne)	5	4	4	5
Sup. 2 (SEK/tonne)	7	4	4	7
Sup. 3 (SEK/tonne)	8	6	5	7

They want to determine a plan for the purchases that minimizes the fuel costs of the airport operator during the period.

(a) Formulate this as a linear programming problem.
State carefully what your variables stand for. . . . . (2 p)

(b) As you surely have noted the problem is a typical “transportation problem”.
Use therefore the transport algorithm to solve it. . . . . (6 p)

(c) Formulate the corresponding *dual* LP-problem and state an optimal solution to it.
(2 p)

2.9 (20040415-nr.1)

A company has four factories, in the cities A, B, C and D, and five big customers, in the cities P, Q, R, S and T.

At a specific time the customers demand the following quantities of the company's product: Customer in P 50 tonnes, Customer in Q 120 tonnes, Customer in R 110 tonnes, Customer in S 70 tonnes and Customer in T 90 tonnes.

The company's supply of the product at the time is 90 tonnes at the factory in A, 100 tonnes at the factory in B, 60 tonnes at the factory in C and 190 tonnes at the factory in D.

The transportation costs from factories to customers, in 100 SEK/tonne, is given by the following table:

	P	Q	R	S	T
A	6	6	5	6	3
B	9	8	7	8	6
C	8	8	5	7	5
D	8	9	6	7	5

(a) Determine how much the company should transport from each factory to each customer to make the total transportation cost minimized under the constraints that the customers demand are satisfied and the supply of the factories are not exceeded. Use the transport algorithm. .... (7p)

(b) The solution you obtained is not unique, right? Determine therefore also another optimal solution. .... (3p)

2.10 (20040310-nr.1)

The linear optimization problem stated below in the variables x_{ik} and z_{kj} can be interpreted as a minimum cost problem with I source-nodes, K intermediate nodes, J sink-nodes, an edge from every source-node to every intermediate node (corresponding to the variables x_{ik}) and an edge from every intermediate node to every sink-node (corresponding to the variables z_{kj}).
All flow from the source-nodes (factories) to the sink-nodes (stores) must always go through intermediate nodes (trans-shipment terminals).

$$
\begin{array}{lll}
\text{minimize} & \displaystyle\sum_{i=1}^{I}\sum_{k=1}^{K} p_{ik}x_{ik} + \sum_{k=1}^{K}\sum_{j=1}^{J} q_{kj}z_{kj} & \\
\text{s.t.} & \displaystyle\sum_{k=1}^{K} x_{ik} \;=\; s_i\,, & \text{for } i = 1,\dots,I \\
 & \displaystyle -\sum_{i=1}^{I} x_{ik} \;+\; \sum_{j=1}^{J} z_{kj} \;=\; 0\,, & \text{for } k = 1,\dots,K \\
 & \displaystyle -\sum_{k=1}^{K} z_{kj} \;=\; -d_j\,, & \text{for } j = 1,\dots,J \\
 & x_{ik} \ge 0\,,\; z_{kj} \ge 0\,, & \text{for all } i,k,j
\end{array}
$$

Here s_i, d_j, p_{ik} and q_{kj} are given positive numbers such that $\sum_{i=1}^{I} s_i = \sum_{j=1}^{J} d_j$.

Assume specifically that we have the following data given:
$I = K = J = 2$, $s_1 = 30$, $s_2 = 20$, $d_1 = 40$, $d_2 = 10$, $p_{11} = 5$, $p_{12} = 2$, $p_{21} = 3$, $p_{22} = 2$, $q_{11} = 5$, $q_{12} = 5$, $q_{21} = 7$, $q_{22} = 6$.

(a) Show that the following solution is an optimal solution to the problem:
$x_{11} = 0$, $x_{12} = 30$, $x_{21} = 20$, $x_{22} = 0$, $z_{11} = 20$, $z_{12} = 0$, $z_{21} = 20$, $z_{22} = 10$.
(7p)

(b) Formulate the dual LP-problem corresponding to the problem above and state an optimal solution to this dual problem. Verify specifically that the optimal values of the problems (primal and dual) are equal. (3p)

3. Convexity

3.1 Let C and D be two convex sets, $\alpha \in I\!R$ and $f : C \to I\!R$ a convex function. Show that the following sets are convex.

(a) $\alpha C = \{\alpha x | x \in C\}$.

(b) $C \bigcap D$.

(c) $C + D = \{x + y | x \in C, y \in D\}$.

(d) $\{x \in C | f(x) \leq \alpha\}$.

(e) $\text{epi}\, f = \{(x, \mu) \in C \times I\!R | f(x) \leq \mu\}$.

3.2 Let C_α denote a convex set in $I\!R^n$ for each α in the index set A. Show that $\bigcap_{\alpha \in A} C_\alpha$ is a convex set.

3.3 Let f and g be convex functions on a convex set C, and let α be a positive constant. Show that the following functions are convex on C.

(a) $f + g$.

(b) αf.

(c) $\max\{f, g\}$.

3.4 Let f_α be convex functions (defined on the same convex set C) for each α in the index set A. Show that $\sup_{\alpha \in A} f_\alpha$ is a convex function.

3.5 Let $f : I \to I\!R$ be a nondecreasing convex function on the interval $I \subseteq I\!R$, and let $g : C \to I$ be a convex function on the convex set $C \subseteq I\!R^n$. Show that $f(g(x))$ is a convex function on C.

3.6 Which of the following functions are convex?

(a) $f(x) = \ln(e^{x_1} + e^{x_2})$.

(b) $f(x) = \ln(\sum_{i=1}^n e^{a_i x_i})$.

(c) $f(x) = \sqrt{\sum_{i=1}^n x_i^2}$.

(d) $f(x) = x_1^2 / x_2$, for $x_2 > 0$.

(e) $f(x) = -\sqrt{x_1 x_2}$, for $x_1, x_2 > 0$.

(f) $f(x) = -(\prod_{i=1}^n x_i)^{1/n}$, for $x_i > 0$.

3.7 Show the inequality between the arithmetic and the geometric mean, i.e., show that for $x_i > 0$, it holds that $(\prod_{i=1}^n x_i)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n x_i$.

3.8 **(a)** Let $x_1, \ldots, x_m \in I\!R^n$ be given. Show that

$$C = \{x \in I\!R^n |\text{ There exist } t_i \geq 0 \text{ such that } x = \sum_{i=1}^m t_i x_i, \sum_{i=1}^m t_i = 1\}$$

is a convex set.

(b) Suppose that X is a convex subset of $I\!R^n$ and that $x_1, \ldots, x_m \in X$. Show that $\sum_{i=1}^m t_i x_i \in X$, if $t_i \geq 0$ and $\sum_{i=1}^m t_i = 1$.

3.9 (20060308-nr.5)

Let f be a one-variable function ($f : I\!R \to I\!R$) which is twice continuously differentiable on the entire $I\!R$ and fulfills that $f(x) > 0$ for all $x \in I\!R$.

Let the one-variable function g be defined by $g(x) = (f(x))^2$, for all $x \in I\!R$.

Determine which of the following statements that are true (proof or counterexample).

Note that it is given that $f(x) > 0$ for all $x \in I\!R$.

(a) If f is convex on $I\!R$ then g is convex on $I\!R$. .... (2p)

(b) If f is not convex on $I\!R$ then g is not convex on $I\!R$. .... (2p)

(c) If $\hat{x}$ is a local minimizer to f then $\hat{x}$ is a local minimizer to g. ..... (2p)

(d) If $\hat{x}$ is not a local minimizer to f then $\hat{x}$ is not a local minimizer to g. (2p)

(e) Assume that in a given point $x_0 \in I\!R$ is holds that
$f'(x_0) \neq 0$, $f''(x_0) > 0$, $g'(x_0) \neq 0$ and $g''(x_0) > 0$.
Let x_1 be the point you obtain if you perform *one* iteration with Newton's (one-variable)-method to minimize $f(x)$ starting from the point x_0, and let $\bar{x}_1$ be the point you obtain if you perform *one* iteration with Newton's (one-variable)-method to minimize $g(x)$ starting from x_0.
Show that $|\bar{x}_1 - x_0| < |x_1 - x_0|$. .... (2p)

Known theorems may be used without proof if they are properly formulated.

3.10 (20051024-nr.5)

Let f and g be two given real valued functions that are *convex* on whole $I\!R^n$, and consider the following convex optimization problem, which we denote P0, in the variable vector $\mathbf{x} \in I\!R^n$:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g(\mathbf{x}) \leq 0. \end{aligned}$$

This exercise is about how you can determine an overestimation of the optimal value to P0 with help of linear programming.

Let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(K)}$ be K given points in $I\!R^n$ and form the following LP-problem, which we denote LP1, in the variables $w_1, \ldots, w_K$:

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{K} w_k f(\mathbf{x}^{(k)}) \\ \text{s.t.} \quad & \sum_{k=1}^{K} w_k g(\mathbf{x}^{(k)}) \leq 0 \\ & \sum_{k=1}^{K} w_k = 1 \\ & w_k \geq 0, \qquad \text{for } k = 1, \ldots, K. \end{aligned}$$

Suppose that $\hat{\mathbf{x}}$ is an optimal solution to P0 and that $\hat{\mathbf{w}} = (\hat{w}_1, \ldots, \hat{w}_K)^\mathsf{T}$ is an optimal basic solution to LP1 (obtained with the simplex method).

(a) Show that $\sum_{k=1}^{K} \hat{w}_k f(\mathbf{x}^{(k)}) \geq f(\hat{\mathbf{x}})$, i.e. that the optimal value to LP1 is an overestimation of the optimal value of P0. .... (7p)

(b) Assume that one of the four given points $\mathbf{x}^{(k)}$ is an optimal solution to P0. Determine an optimal solution to LP1 and show that the optimal solutions
to P0 and LP1 now are identical. .... (3p)

4. Lagrange relaxations and duality

4.1 Solve the following problem, and motivate global optimality.

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} x_j^2 \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_j x_j \geq b \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned}$$

where a_j, $j = 1, \dots, n$ and b are constants.

4.2 Solve the following problem, and motivate global optimality.

$$\begin{aligned} \max \quad & \sum_{j=1}^{n} \ln x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_j x_j \leq b, \\ & x_j > 0, \quad j = 1, \dots, n, \end{aligned}$$

where a_j, $j = 1, \dots, n$ and b are positive constants.

4.3 Solve the following problem, and motivate global optimality.

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} a_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} \frac{b_j}{x_j} \leq b_0, \\ & x_j > 0, \quad j = 1, \dots, n, \end{aligned}$$

where a_j, $j = 1, \dots, n$, and b_j, $j = 0, \dots, n$ are positive constants.

4.4 Solve the following problem, and motivate global optimality.

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} \frac{a_j}{x_j} \\ \text{s.t.} \quad & \sum_{j=1}^{n} b_j x_j = b_0, \\ & x_j > 0, \quad j = 1, \dots, n, \end{aligned}$$

where a_j, $j = 1, \dots, n$, and b_j, $j = 0, \dots, n$ are positive constants.

4.5 Solve the following problem, and motivate global optimality.

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} e^{c_j x_j} \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_j x_j \geq b, \end{aligned}$$

where a_j, $j = 1, \dots, n$, c_j, $j = 1, \dots, n$ and b are positive constants.

4.6 At an exam, a "new" method is used for answering multiple-choice questions. The answer is given by denoting the probability of each answer to be correct.

The given probabilities have to sum up to 1 (and be nonnegative of course).

If choice k is correct and you have put the probability q_k you are given the score $\ln q_k$ (which is normally negative).

Suppose you are to answer such a multiple-choice question with N alternatives. You judge that the probability is p_n for alternative n to be correct. The probabilities q_n which you give as answers do not have to be identical to the p_n.

(a) Formulate the problem to determine the probabilities q_n so that the expected total score is maximized.

(b) Determine the optimal choice of $q_n, n = 1, \ldots, N$.

Remark: You use subjective probabilities in the same way you would use regular probabilities.

4.7 Consider the problem (P) defined as

$$
(P) \quad \begin{array}{ll}
\min & x_1^4 + 2x_1x_2 + x_2^2 + x_3^8 \\
\text{s.t.} & (x_1 - 2)^2 + (x_2 - 2)^2 + (x_3 - 3)^2 \leq 6, \\
& x_1x_2x_3 \leq 10, \\
& x_1 \geq 1, \\
& x_2 \geq 0, \\
& x_3 \geq 0.
\end{array}
$$

Use Lagrangean relaxation to show that $\hat{x} = (\ 1\ 1\ 1\ )^T$ is a global minimizer to (P).

4.8 Determine the dual problem to

$$
(P) \quad \begin{array}{ll}
\min & \displaystyle\sum_{i=1}^{n} x_i^2 \\
\text{s.t.} & \displaystyle\sum_{i=1}^{n} a_i x_i = b.
\end{array}
$$

4.9 Determine a suitable dual problem to

$$
(P) \quad \begin{array}{ll}
\min & \displaystyle\sum_{i=1}^{n} \frac{a_i}{x_i} \\
\text{s.t.} & \displaystyle\sum_{i=1}^{n} b_i x_i = b_0, \\
& l_i \leq x_i \leq u_i, \quad i = 1, \ldots, n,
\end{array}
$$

where $a_i > 0$, $0 < l_i < u_i$ for $i = 1, \ldots, n$ and $b_0 > 0$.

4.10 Determine a suitable dual problem to

$$(P)\quad \begin{array}{ll} \min & \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln x_{ij} \\ \text{s.t.} & \displaystyle\sum_{j=1}^{n} x_{ij} = a_i, \quad i = 1,\dots,n, \\ & \displaystyle\sum_{i=1}^{n} x_{ij} = b_j, \quad j = 1,\dots,n, \\ & x_{ij} \ge 0, \quad i = 1,\dots,n, j = 1,\dots,n, \end{array}$$

where $a_i > 0$ for $i = 1,\dots,n$ and $b_j > 0$ for $j = 1,\dots,n$.

4.11 Consider the problem

$$(P)\quad \begin{array}{ll} \min & \displaystyle\sum_{j=1}^{n} x_j^3 \\ \text{s.t.} & \displaystyle\sum_{j=1}^{n} a_j x_j \ge b, \\ & x_j \ge 0, \quad j = 1,\dots,n, \end{array}$$

where $a_j > 0, j = 1,\dots,n$ and $b > 0$.

(a) Determine (D), the (Lagrange-) dual problem to (P) which is created when the sum constraint is relaxed.

(b) Determine an optimal solution to (D). (It is not necessary to use a systematic method.)

5. Quadratic programming

5.1 (20070601-nr.3)

This problem concerns a small electrical network with resistances on the links. The network has the set of nodes $\mathcal{N} = \{1, 2, 3, 4\}$ (i.e. in total 4 nodes) and the set of links $\mathcal{B} = \{(1,3),\ (1,4),\ (2,3),\ (2,4)\}$. (Hence there are no link between the nodes 1 and 2 and no link between the nodes 3 and 4, but there is a link between any other node pair.) Every link $(i,j) \in \mathcal{B}$ has a given resistance R_{ij} Ohm. Now suppose that the current 500 mA is fed into node 1 and 100 mA into node 2, while 500 mA is taken out in node 3 and 100 mA is taken out in node 4. The total (heat) effect in the network is then given by

$$R_{13}x_{13}^2 + R_{14}x_{14}^2 + R_{23}x_{23}^2 + R_{24}x_{24}^2,$$

where x_{ij} = current in the link (i,j). (If $x_{ij} > 0$ the current in the link is going from node i to node j, while if $x_{ij} < 0$ the current in the link is going from node j to node i.)

Nature decides the currents x_{ij} in such a way that the above sum is minimized under the constraints of balance of the currents in the first three nodes. The current balance in the fourth node, $-x_{14} - x_{24} = -100$, follows from the current balances in the other three nodes, as for all problems regarding balanced networks or flows.

Your task is now to compute the currents in the links x_{ij} by solving the optimization problem above, which has a convex quadratic objective function and linear equality constraints. Use a general method for this type of quadratic optimization.
For simplicity you may assume that $R_{ij} = 1$ for all links.

5.2 (20070307-nr.3)

In the entire exercise

$$f(\mathbf{x}) = (x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2, \quad \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 10 \\ 14 \end{pmatrix}$$

is valid.

(a) Determine a symmetric 3×3–matrix $\mathbf{H}$ such that $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$. . (1p)

(b) Determine *one* solution $\bar{\mathbf{x}}$ to the system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$. . . . . . (1p)

(c) Determine a basis of $\mathcal{N}(\mathbf{A})$ (= Null-space to $\mathbf{A}$). . . . . (2p)

(d) Use the results from (a)–(c) to determine an optimal solution $\hat{\mathbf{x}}$ to the problem: minimize $f(\mathbf{x})$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$. . . . . (2p)

(e) Now let $\mathbf{c} \in I\!R^3$ be a given vector and consider the problem to minimize $f(\mathbf{x}) + \mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{x} \in I\!R^3$, i.e. a quadratic optimization problem *without* constraints. For some choices of the vector $\mathbf{c}$ it turns out that this problem have at least one optimal solution, which therefore is a minimizer to $f(\mathbf{x}) + \mathbf{c}^\mathsf{T}\mathbf{x}$, while it for other choices of the vector $\mathbf{c}$ it turns out that $f(\mathbf{x}) + \mathbf{c}^\mathsf{T}\mathbf{x}$ has no lower bound and hence does not have a minimizer.

Your task is to show that there exist a vector $\mathbf{a} \in I\!R^3$, which you should state, such that the following equivalence is valid:

$f(\mathbf{x}) + \mathbf{c}^\mathsf{T}\mathbf{x}$ has at least one minimizer $\Leftrightarrow$ $\mathbf{a}^\mathsf{T}\mathbf{c} = 0$. ... (4p)

5.3 (20040310-nr.4)

In this exercise $\mathbf{c} \in I\!R^n$ and $\mathbf{a} \in I\!R^n$ are two given vectors that fulfill that

$\mathbf{a}^\mathsf{T}\mathbf{a} = 1$, $\mathbf{c}^\mathsf{T}\mathbf{c} = 1$ and $\mathbf{a}^\mathsf{T}\mathbf{c} = 0$.

Now consider the following problem in the variable vector $\mathbf{x} \in I\!R^n$:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x} - \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x} - \mathbf{a}^\mathsf{T}\mathbf{x} \le 0\,. \end{aligned}$$

(a) Make a Lagrange relaxation of the problem and determine an explicit expression for the dual objective function φ. .... (4p)

(b) Determine an optimal solution $\hat{y}$ to the dual problem. ....(3p)

(c) Determine an optimal solution $\hat{\mathbf{x}}$ to the original (primal) problem above and verify that $(\hat{\mathbf{x}}, \hat{y})$ fulfill the global optimality conditions. (3p)

5.4 (20060603-nr.2)

Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$ and $\mathbf{q} = \begin{pmatrix} 4 \\ 2 \\ 0 \\ -2 \end{pmatrix}$.

(a) First assume that you want to determine the vector $\mathbf{x}$ in the null-space to $\mathbf{A}$
which is closest to the vector $\mathbf{q}$, i.e. you want to solve the problem

$$\text{P1:} \quad \text{minimize } |\,\mathbf{x} - \mathbf{q}\,|^2 \ \text{ s.t. } \mathbf{x} \in \mathcal{N}(\mathbf{A}),$$

where $|\cdot|$ means the standard Euclidean norm in $I\!R^4$, i.e. $|\,\mathbf{x}-\mathbf{q}\,|^2 = (\mathbf{x}-\mathbf{q})^\mathsf{T}(\mathbf{x}-\mathbf{q})$,
and $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in I\!R^4 \,|\, \mathbf{A}\mathbf{x} = \mathbf{0}\}$. Determine the optimal $\mathbf{x}$. .... (5p)

(b) Now suppose you want to determine the vector $\mathbf{x}$ in the range space of $\mathbf{A}^\mathsf{T}$
which is the closest to the vector $\mathbf{q}$, i.e. you want to solve the problem

$$\text{P2:} \quad \text{minimize } |\,\mathbf{x} - \mathbf{q}\,|^2 \ \text{ s.t. } \mathbf{x} \in \mathcal{R}(\mathbf{A}^\mathsf{T}),$$

where $\mathcal{R}(\mathbf{A}^\mathsf{T}) = \{\mathbf{x} \in I\!R^4 \,|\, \mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{v} \text{ for some } \mathbf{v} \in I\!R^2\}$. Determine optimal $\mathbf{x}$. .... (5p)

5.5 (20060308-nr.4)

(a) Let $\mathbf{a} \in I\!R^3$ be a given vector and $b \in I\!R$ a given constant.
Then $\mathcal{P} = \{\, \mathbf{x} \in I\!R^3 \mid \mathbf{a}^\mathsf{T}\mathbf{x} = b \,\}$ is a plane in $I\!R^3$.
Let $\bar{\mathbf{x}} \in I\!R^3$ be a given point that fulfills $\mathbf{a}^\mathsf{T}\bar{\mathbf{x}} < b$.

Determine the point $\hat{\mathbf{x}} \in \mathcal{P}$ that is closest to $\bar{\mathbf{x}}$ among all points in $\mathcal{P}$, i.e. optimal solution to the problem to minimize $|\mathbf{x}-\bar{\mathbf{x}}|^2$ s.t. $\mathbf{a}^\mathsf{T}\mathbf{x} = b$. (The answer will of course contain $\bar{\mathbf{x}}$, $\mathbf{a}$ and b.)
Also show that the shortest distance d from $\bar{\mathbf{x}}$ to the plane $\mathcal{P}$, i.e. $|\hat{\mathbf{x}}-\bar{\mathbf{x}}|$, is given by the expression $d = (b - \mathbf{a}^\mathsf{T}\bar{\mathbf{x}})/|\mathbf{a}|$. .... (4p)

5.6 (20050331-nr.4)

Let $\mathcal{U}$ and $\mathcal{V}$ be the following two subsets of $I\!R^4$:

$$\mathcal{U} = \{\mathbf{u} \in I\!R^4 \mid \mathbf{Ru} = \mathbf{p}\}, \text{ where } \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{p} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

$$\mathcal{V} = \{\mathbf{v} \in I\!R^4 \mid \mathbf{Sv} = \mathbf{q}\}, \text{ where } \mathbf{S} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{q} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Determine the shortest (Euclidean) distance d between $\mathcal{U}$ and $\mathcal{V}$. Determine also
the two points $\hat{\mathbf{u}} \in \mathcal{U}$ and $\hat{\mathbf{v}} \in \mathcal{V}$ between which the distance is the smallest. (10p)

5.7 (20050307-nr.3)

In this exercise $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, where b_1 and b_2 are given numbers.

The Euclidean norm of a vector $\mathbf{x}$ are as usual denoted by $|\mathbf{x}|$, so that $|\mathbf{x}|^2 = \mathbf{x}^\mathsf{T}\mathbf{x}$ and $|\mathbf{Ax}-\mathbf{b}|^2 = (\mathbf{Ax}-\mathbf{b})^\mathsf{T}(\mathbf{Ax}-\mathbf{b})$.

(a) Determine *all* optimal solutions $\mathbf{x}$ to the following problem:

$$\begin{aligned} \text{P1:} \qquad & \text{minimize} \quad |\mathbf{Ax}-\mathbf{b}|^2 \\ & \text{s.t.} \quad \mathbf{x} \in I\!R^2. \end{aligned}$$

The answer can of course contain b_1 and b_2. ....(2p)

(b) Let $X(\mathbf{b})$ = be the set of optimal solutions $\mathbf{x}$ to the problem P1 above. Determine the unique optimal solution to the following problem:

$$\begin{aligned} \text{P2:} \qquad & \text{minimize} \quad |\mathbf{x}|^2 \\ & \text{s.t.} \quad \mathbf{x} \in X(\mathbf{b}). \end{aligned}$$

The answer may of course contain b_1 and b_2. .... (2p)

(c) Let $\hat{\mathbf{x}}(\mathbf{b})$ denote the optimal solution to the problem P2 above. Show that $\hat{\mathbf{x}}(\mathbf{b}) = \mathbf{A}^{+}\mathbf{b}$ for a certain matrix $\mathbf{A}^{+}$. State $\mathbf{A}^{+}$. .... (1p)

(d) Now let ε be a given number that fulfills $\varepsilon > 0$. Determine the unique optimal solution to the following problem:

$$\begin{aligned} \text{P3:} \qquad & \text{minimize} \quad |\mathbf{Ax}-\mathbf{b}|^2 + \varepsilon|\mathbf{x}|^2 \\ & \text{s.t.} \quad \mathbf{x} \in I\!R^2. \end{aligned}$$

The answer may of course contain b_1, b_2 and ε. .... (3p)

(e) Let $\tilde{\mathbf{x}}_\varepsilon(\mathbf{b})$ denote the optimal solution to the problem P3 above.
Show that $\tilde{\mathbf{x}}_\varepsilon(\mathbf{b}) = \tilde{\mathbf{A}}_\varepsilon \mathbf{b}$ for a certain matrix $\tilde{\mathbf{A}}_\varepsilon$ which depends on ε. Show also that if ε goes to 0, then every element of the matrix $\tilde{\mathbf{A}}_\varepsilon$ goes to the corresponding element in the matrix $\mathbf{A}^+$. .... (2p)

5.8 (20041016-nr.3)

Let L_1 and L_2 be two given lines on parameter form in $I\!R^3$:

$L_1 = \{\, \mathbf{x} \in I\!R^3 \mid \mathbf{x} = \mathbf{a} + \alpha\cdot\mathbf{u}, \;\text{ for } \alpha \in I\!R\}$ and

$L_2 = \{\, \mathbf{y} \in I\!R^3 \mid \mathbf{y} = \mathbf{b} + \beta\cdot\mathbf{v}, \;\text{ for } \beta \in I\!R\}$,

where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$ and $\mathbf{v}$ are given vectors in $I\!R^3$.

We assume that the direction vectors of the lines $\mathbf{u}$ and $\mathbf{v}$ are normalized, i.e. $\mathbf{u}^\mathsf{T}\mathbf{u} = \mathbf{v}^\mathsf{T}\mathbf{v} = 1$, and that they are *not* parallel.

You want to connect the lines with a thread. The question is between which two points $\hat{\mathbf{x}} \in L_1$ and $\hat{\mathbf{y}} \in L_2$ the thread should be tied to make it as short as possible.

(a) Formulate this as an optimization problem in two variables (α and β). Determine specifically if your formulation is convex! .... (4p)

(b) Solve the problem you formulated above analytically, with the help of well-known optimality conditions, for the special case that the given vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, i.e. fulfills $\mathbf{u}^\mathsf{T}\mathbf{v} = 0$. Denote the optimal points $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ expressed in $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$ and $\mathbf{v}$. .... (6p)

5.9 (20040415-nr.3)

Consider the quadratic optimization problem to minimize $\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$ when $\mathbf{A}\mathbf{x} = \mathbf{b}$,

$$\text{where } \mathbf{H} = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

A feasible (but not optimal) solution to the problem is $\bar{\mathbf{x}} = (\,1,\ 1,\ 1,\ 1,\ 1\,)^\mathsf{T}$.

(a) Show that the vectors $\mathbf{z}_1 = (\,0,\ -1,\ 0,\ 1,\ 0\,)^\mathsf{T}$ and $\mathbf{z}_2 = (\,1,\ 0,\ -1,\ 0,\ 1\,)^\mathsf{T}$ is a basis to $\mathcal{N}(\mathbf{A})$ (= the null-space of $\mathbf{A}$). .... (2p)

(b) Use the above information to solve the considered optimization problem with a null-space method. .... (6p)

(c) Let $\hat{\mathbf{x}}$ denote the optimal solution to the problem. Determine a vector $\hat{\mathbf{u}}$ that together with $\hat{\mathbf{x}}$ satisfies the optimality conditions to the problem. (2p)

5.10 (20040310-nr.2)

This exercise is about an electrical network with resistances on the links.
The network consists of 5 nodes, and there is a link between every pair of nodes. Hence there are in total 10 links (i, j) with $1 \le i < j \le 5$.

For simplicity every link (i, j) is assumed to have the resistance $r_{ij} = 1$ Ohm.

Assume that you send in the current 1 Ampere in a node, say node 1, and take it out in another node, say node 5. The total (heat-)effect in the network is then given by $\sum r_{ij}x_{ij}^2$, where x_{ij} = the current in the link (i, j) and where the sum goes over to all the other 10 links in the network. (If $x_{ij} > 0$, the current in the link goes from node i to node j, whereas if $x_{ij} < 0$, the current goes in the link from node j to node i.)

Nature chooses the currents x_{ij} in a way that minimizes the mentioned sum $\sum r_{ij}x_{ij}^2$ under constraints on current balances in the nodes, i.e. 4 linearly independent flow constraints. (Balance in the fifth node follows from balance in the other four nodes, as for all network flow problems.)

Your task is now to compute the currents in the links x_{ij} by solving the minimization problem above, which has a quadratic objective function and linear constraints. State also the total effect. .... (10p)

The following might be useful for the computations:

$$\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}^{-1} = \frac{1}{5} \cdot \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}.$$

6. Nonlinear programming

6.1 (20070601-nr.5)

In the following optimization problem c_1 is a constant.

$$\begin{array}{rl} \text{minimize} & c_1x_1 - 4x_2 - 2x_3 \\ \text{s.t.} & x_1^2 + x_2^2 \le 2, \\ & x_1^2 + x_3^2 \le 2, \\ & x_2^2 + x_3^2 \le 2. \end{array}$$

(a) Decide whether it is a convex optimization problem or not.
Motivate your answer carefully.. .... (1p)

(b) Write down the KKT-conditions for the problem. .... (1p)

(c) Are there any values for the constant c_1 which make the point $\mathbf{x} = (1.4,\ 0.2,\ 0.2)^\mathsf{T}$ an optimal solution to the problem?
If there are any such values, determine *all* of these values for c_1. . . (4p)

(d) Are there any values for the constant c_1 which make the point $\mathbf{x} = (1,\ 1,\ 1)^\mathsf{T}$ an optimal solution to the problem?
If there are any such values, determine *all* of these values for c_1. . . (4p)

Optional help: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$

(e) In this final task we will throughout assume that $c_1 = -6$.
The dual objective function $\varphi(\mathbf{y})$ of the problem above is, as known, defined with Lagrangian relaxation. Your task is to compute the value of the dual objective function $\varphi(\hat{\mathbf{y}})$ in the point $\hat{\mathbf{y}} = (1, 1, 1)^\mathsf{T}$, and determine whether $\hat{\mathbf{y}}$ is an optimal solution to the dual problem. .... (4p)

6.2 (20070307-nr.4)

Let δ_1, δ_2, δ_3 and δ_4 be four given numbers (which typically are pretty "small") and consider the following non-linear least squares problem in the variable-vector $\mathbf{x} \in I\!R^2$:

$$\text{minimize} \quad f(\mathbf{x}) = \tfrac{1}{2}(h_1(\mathbf{x})^2 + h_2(\mathbf{x})^2 + h_3(\mathbf{x})^2 + h_4(\mathbf{x})^2),$$

where the functions h_i are given by

$$\begin{aligned} h_1(\mathbf{x}) &= x_1^2 - x_2 - \delta_1\,, \\ h_2(\mathbf{x}) &= x_1^2 + x_2 - \delta_2\,, \\ h_3(\mathbf{x}) &= x_2^2 - x_1 - \delta_3\,, \\ h_4(\mathbf{x}) &= x_2^2 + x_1 - \delta_4\,. \end{aligned}$$

(a) First assume that $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$.
Show that then $\hat{\mathbf{x}} = (0, 0)^\mathsf{T}$ is a global minimizer to $f(\mathbf{x})$.
(This motivates that we use this point as starting point below.) . . . (1p)

(b) Now assume that $\delta_1 = -0.1,\ \delta_2 = 0.1,\ \delta_3 = -0.2$ and $\delta_4 = 0.2$.
Perform one iteration with the Gauss-Newton method starting from $\mathbf{x}^{(1)} = (0,0)^\mathsf{T}$.
Make sure that your obtained point $\mathbf{x}^{(2)}$ satisfies $f(\mathbf{x}^{(2)}) < f(\mathbf{x}^{(1)})$.
Then decide whether $\mathbf{x}^{(2)}$ is a local minimizer to $f(\mathbf{x})$. .... (5p)

6.3 (20060603-nr.3)

In the following QP-problem with inequality constraints the number c_3 is a constant.

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 + \tfrac{1}{2}x_3^2 - x_1 - x_2 + c_3 x_3 \\ \text{s.t.} \quad & x_1 + x_2 \geq 4 \\ & x_1 + x_3 \geq 4 \\ & x_2 + x_3 \geq 4 \end{aligned}$$

(a) Are there any values of the constant c_3 which makes the point $\mathbf{x} = (2,2,2)^\mathsf{T}$ an optimal solution to the problem?
If that is the case, then determine *all* such values on c_3. .... (4p)

(b) Is there any values of the constant c_3 which makes the point $\mathbf{x} = (2,2,4)^\mathsf{T}$ an optimal solution to the problem?
If that is the case, then determine *all* such values on c_3. .... (3p)

(c) Is there any value of the constant c_3 such that the point $\mathbf{x} = (3,3,1)^\mathsf{T}$ is an optimal solution to the problem?
If that is the case, then determine *all* such values on c_3. .... (3p)

Optional help: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \dfrac{1}{2} \cdot \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$

6.4 (20060603-nr.4)

In this exercise $f(\mathbf{x}) = x_1^2 x_2^4 x_3^6$, where $\mathbf{x} = (x_1,\, x_2,\, x_3)^\mathsf{T} \in I\!R^3$.

(a) Determine whether $\hat{\mathbf{x}} = (0,0,0)^\mathsf{T}$ is a *global* optimal solution to the problem to
minimize $f(\mathbf{x})$ under the constraint $x_1^2 + x_2^2 + x_3^2 \leq 1$. .... (1p)

(b) Determine whether $\hat{\mathbf{x}} = (0,0,0)^\mathsf{T}$ is a *local* optimal solution to the problem to
minimize $f(\mathbf{x})$ under the constraint $x_1^2 + x_2^2 + x_3^2 \leq 1$. .... (1p)

(c) Determine whether f is a convex function on $I\!R^3$ or not. .... (3p)

(d) Determine all *global* optimal solutions to the problem to
maximize $f(\mathbf{x})$ under the constraint $x_1^2 + x_2^2 + x_3^2 \leq 1$.
Observe that now, in contrary to above, it is a maximization problem! (5p)

6.5 (20051024-nr.4)

Consider the following quadratic optimization problem with four variables and six linear inequality constraints (of which four are in the form of non-negativity constraints on the variables).

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 2 & -2 & 1 & 1 \\ 1 & 1 & 2 & -2 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 20 \\ 30 \end{pmatrix}$.

Use for example the optimality constraints to answer the following questions:

(a) Is there for the above problem any *optimal* solution $\hat{\mathbf{x}}$ that satisfies both that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that all four variables are strictly positive, i.e. $\hat{x}_j > 0$? (5p)

(b) Is there to the above problem any *optimal* solution $\hat{\mathbf{x}}$ that satisfies both that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that $\hat{x}_3 = \hat{x}_4 = 0$? .... (5p)

6.6 (20050307-nr.4)

(a) Determine whether $\hat{\mathbf{x}} = (2,\, 1,\, 0)^\mathsf{T}$ is a global optimal solution or not to the following problem in the variable vector $\mathbf{x} = (x_1,\, x_2,\, x_3)^\mathsf{T} \in I\!R^3$. (5p)

$$\begin{aligned} \text{minimize} \quad & (x_1+x_2)^2 + (x_2+x_3)^2 + (x_3+x_1)^2 - 12x_1 - 8x_2 - 4x_3 \\ \text{s.t.} \quad & 0 \leq x_j \leq 2, \quad \text{for } j = 1, 2, 3. \end{aligned}$$

(b) Determine values on the three constants k_1, k_2 and k_3 so that $\hat{\mathbf{x}} = (2,\, 1,\, 0)^\mathsf{T}$ becomes a global optimal solution to the following problem. (5p)

$$\begin{aligned} \text{minimize} \quad & (x_1+x_2)^2 + (x_2+x_3)^2 + (x_3+x_1)^2 - 12x_1 - 8x_2 - 4x_3 \\ \text{s.t.} \quad & (x_1-k_1)^2 + (x_2-k_2)^2 + (x_3-k_3)^2 \leq 1. \end{aligned}$$

6.7 (20041016-nr.5)

Consider the following non-linear optimization problem.

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, 2 \\ & \mathbf{x} \in X, \end{aligned}$$

where X is given by $X = \{\, \mathbf{x} \in I\!R^n \,|\, -1 \leq x_j \leq 1,\ j = 1, \ldots, n\}$,

while the functions $f_i(\mathbf{x})$ are given by the following expression, for $i = 0, 1, 2$:

$f_i(\mathbf{x}) = \sum_{j=1}^{n} \left(\dfrac{p_{ij}}{2 - x_j} + \dfrac{q_{ij}}{x_j + 2} \right) + r_i$. Note that this holds also for $i = 0$, i.e. the objective function.

Here p_{ij} and q_{ij} are given strictly positive constants for $i = 0, 1, 2$ and all j, while r_i are given constants such that $r_i < -\sum_{j=1}^{n} \left(\dfrac{p_{ij}}{2} + \dfrac{q_{ij}}{2} \right)$ for $i = 1, 2$.

In the following subproblems known theorems may be used without proof if they
first are correctly formulated.

(a) Determine whether the functions f_i are convex on X. .... (2p)

(b) Determine whether the feasible region is non-empty. .... (1p)

(c) Determine whether the feasible region to the problem is a convex set. (1p)

(d) Formulate carefully the KKT-conditions to the problem. .... (3p)

(e) Make a Lagrangian relaxation to the problem with respect to the constraints $f_i(\mathbf{x}) \leq 0,\ i = 1, 2,$
and determine an explicit expression of the dual objective function φ. (3p)

6.8 (20040415-nr.4)

Three components are placed out on a board with the coordinates:

$$(1,\ 2)\,,\ \ (-2,\ -1)\ \text{ and }\ (1,\ -1).$$

Now you want to connect these components with a fourth one, so that from each of the first three there is a (direct) thread to the fourth. The question is *where* on the board this fourth component should be placed. One could think of at least two different criteria, represented in each of the following to sub-exercises.

(a) First assume that you want to place the fourth component in such a way that the *sum* of the three distances from the fourth component to each of the original three components is as small as possible.

Formulate this problem on mathematical form (with variables, objective function and possible constraints) and decide whether the point $(0, 0)$ fulfills the optimality conditions for your formulated problem. ..... (5p)

(b) Now assume that you want to place the fourth component in such a way that the *biggest* of the three distances from the fourth component to each of the original three components is as *small* as possible.

Formulate this problem on mathematical form (with variables, objective function and possible constraints) and decide whether the point $(0, 0)$ fulfills the optimality conditions for your formulated problem. ..... (5p)

7. Mixed examples

7.1 Consider the following approximation problem.

Given is a set $T \in I\!R^k$, and the continuous functions f and f_j, $j = 1, \ldots, n$ on T. The aim is to approximate f by a linear combination of the functions f_j, $j = 1, \ldots, n$.

This leads to the following optimization problem.

$$(P) \qquad \min_{x_j, j=1,\ldots,n} \max_{t \in T} |f(t) - \sum_{j=1}^{n} x_j f_j(t)|$$

For numerical solution of (P), a set of points $t_1, \ldots, t_m$ are selected in T and the following problem is solved.

$$(P') \qquad \min_{x_j, j=1,\ldots,n} \max_{i=1,\ldots,m} |f(t_i) - \sum_{j=1}^{n} x_j f_j(t_i)|$$

(a) What conclusions can be drawn concerning the optimal value of (P) from an optimal solution to (P')?

(b) (P') looks complex, with an inner maximization and an outer minimization. Reformulate (P') as a mathematical programming problem of simplest possible kind.

(c) Determine the dual problem to the problem formulated in Exercise 7.1b and simplify as far as possible.

7.2 Throughout this exercise, let

$$\begin{aligned} f(x) &= x_1^2 - x_1 x_2 + x_2^2 + x_3^2 - 2x_1 + 4x_2, \\ g_1(x) &= -x_1 - x_2, \\ g_2(x) &= 1 - x_3. \end{aligned}$$

(a) Consider the problem

$$(P_d) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in I\!R^3. \end{array}$$

Determine a *global* minimum to (P_d). (Motivate the answer.)

(b) Consider the problem

$$(P_c) \qquad \begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \le 0, & i = 1, 2, \\ & x \in I\!R^3. & \end{array}$$

Show that $x^* = (1\ {-1}\ 1)^T$ fulfills the KT-conditions for (P_c).

(c) Show that the dual problem (D_c) corresponding to (P_c) is

$$(D_c) \qquad \begin{array}{ll} \max & -\lambda_1^2 + 2\lambda_1 - \dfrac{\lambda_2^2}{4} + \lambda_2 - 4 \\ \text{s.t.} & \lambda_1 \ge 0, \\ & \lambda_2 \ge 0. \end{array}$$

(d) Determine globally optimal solutions to (P_c) and (D_c). (Motivate global optimality.)

Hint: The results from 7.2a and 7.2b can be used in conjunction with weak duality.

7.3 Consider the problem (P), defined as

$$(P) \quad \begin{array}{ll} \min & -2x_1^2 + 12x_1x_2 + 7x_2^2 - 8x_1 - 26x_2 \\ \text{s.t.} & x_1 + 2x_2 \leq 6, \\ & 0 \leq x_1 \leq 3, \\ & x_2 \geq 0. \end{array}$$

(a) Find *all* points that satisfy the KT-conditions for (P).

(b) Find a global minimizer to (P).

(Note that the amount of computation required would be very large using this strategy on larger problems.)

7.4 (20070601-nr.4)

Let n be a given (big) integer, and let the n-variable function f be given by

$$f(\mathbf{x}) = \sum_{j=1}^{n}(x_j^4 - x_j^3 + x_j^2 - x_j), \quad \text{where } \mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T} \in I\!R^n.$$

(a) Decide whether f is a convex function on whole $I\!R^n$. .... (4p)

(b) Now suppose you want to minimize $f(\mathbf{x})$ without any constraints.
Your task is to make *one* complete iteration of the Newton method from the starting point $\mathbf{x}^{(1)} = (1, \ldots, 1)^\mathsf{T}$.
Make sure that your obtained point $\mathbf{x}^{(2)}$ fulfills $f(\mathbf{x}^{(2)}) < f(\mathbf{x}^{(1)})$. . (4p)

7.5 (20050331-nr.5)

Let $g_1, \ldots, g_m$ and f be given nonlinear n-variable functions which are all convex and continuously differentiable on the entire $I\!R^n$ and consider the following problem:

$$\text{P:} \quad \begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \leq 0\,, \quad i = 1, \ldots, m \\ & \mathbf{x} \in I\!R^n\,. \end{array}$$

Assume that the problem is regular, i.e. that there is at least one point $\mathbf{x}$ which fulfills all constraints with strict inequality.

Assume furthermore that you in some way have found a point $\hat{\mathbf{x}} \in I\!R^n$ which you *believe* is an optimal solution to P. One way of investigating whether $\hat{\mathbf{x}}$ really *is* optimal is to linearize the $m+1$ nonlinear functions $g_1, \ldots, g_m$ and f in the point $\hat{\mathbf{x}}$ (i.e. compute the first order Taylor polynomial in $\hat{\mathbf{x}}$ for each function) and then consider the *LP-problem* that is formed, if you in the problem P replace each of the nonlinear functions with its linearization. Your task is to prove the following statement:

The point $\hat{\mathbf{x}}$ is an optimal solution to the problem P *if and only if* the same point $\hat{\mathbf{x}}$ is an optimal solution to the above mentioned LP-problem (which is formed by approximating all functions in P with the first order Taylor polynomials in $\hat{\mathbf{x}}$).

Theorems from the course may be used without proof if they are correctly formulated. .... (10p)

7.6 (20041016-nr.4)

In this exercise $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$.

(a) First consider the following quadratic minimization problem in the variable vector $\mathbf{x} \in I\!R^2$:

$$\text{minimize} \quad \tfrac{1}{2}\,|\mathbf{A}\mathbf{x} - \mathbf{b}|^2 = \tfrac{1}{2}\,(\mathbf{A}\mathbf{x} - \mathbf{b})^{\mathsf{T}}(\mathbf{A}\mathbf{x} - \mathbf{b}).$$

Determine an optimal solution to this problem. .... (4p)

(b) Assume that you add the constraint that all components in the "error vector"
$\mathbf{A}\mathbf{x} - \mathbf{b}$ must be non-negative, you instead obtain the problem

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\,|\mathbf{A}\mathbf{x} - \mathbf{b}|^2 \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}. \end{aligned}$$

Now show that $\mathbf{x} = \left(\frac{13}{6}, \frac{11}{6}\right)^{\mathsf{T}}$ is an optimal solution. .... (6p)

7.7 (20040415-nr.5)

Let f and g be two given non-linear n-variable functions which both are *convex* and continuously differentiable on the entire $I\!R^n$, and denote the following problem NLP.

$$\text{NLP:} \quad \begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g(\mathbf{x}) \leq 0\,, \\ & \mathbf{x} \in I\!R^n\,. \end{aligned}$$

Let $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^K$ be a number of given points in $I\!R^n$, and denote the following LP-problem in the variables $\mathbf{x} \in I\!R^n$ and $z \in I\!R$ by LP.

$$\begin{aligned} \text{LP:} \quad \text{minimize} \quad & z \\ \text{s.t.} \quad & z - \nabla f(\mathbf{x}^k)\,\mathbf{x} \;\geq\; f(\mathbf{x}^k) - \nabla f(\mathbf{x}^k)\,\mathbf{x}^k\,, \quad k = 1, \ldots, K, \\ & -\nabla g(\mathbf{x}^k)\,\mathbf{x} \;\geq\; g(\mathbf{x}^k) - \nabla g(\mathbf{x}^k)\,\mathbf{x}^k\,, \quad k = 1, \ldots, K. \end{aligned}$$

Assume that you solve this LP-problem and obtain an optimal solution $(\hat{\mathbf{x}}, \hat{z})$.

(a) Show that $\hat{z}$ is a lower bound of the optimal value to NLP,
i.e. that $f(\mathbf{x}) \geq \hat{z}$ for every feasible solution $\mathbf{x}$ to NLP. .... (5p)

(b) Assume that the above mentioned $\hat{\mathbf{x}}$ happens to be one of the given points $\mathbf{x}^k$.
Show that $\hat{\mathbf{x}}$ is an optimal solution to NLP. .... (5p)

Answers/Solutions to the examples

8. Linear programming

8.1 (20070601-nr.1a)

Introduce the following variables:

x_1 = number of units of product A manufactured per day
x_2 = number of units of product B manufactured per day
x_3 = number of units of product C manufactured per day

The coverage contribution per day is then given by $12x_1 + 9x_2 + 8x_3$.

The capacity constraint in the cutting department can be written as $\dfrac{x_1}{2000} + \dfrac{x_2}{1600} + \dfrac{x_3}{1100} \leq 8$.

The capacity constraint in the pressing department can be written as $\dfrac{x_1}{1000} + \dfrac{x_2}{1500} + \dfrac{x_3}{2400} \leq 8$.

That gives us the following problem formulation:

$$\begin{aligned} \text{maximize} \quad & 12x_1 + 9x_2 + 8x_3 \\ \text{s.t.} \quad & \frac{x_1}{2000} + \frac{x_2}{1600} + \frac{x_3}{1100} \leq 8, \\ & \frac{x_1}{1000} + \frac{x_2}{1500} + \frac{x_3}{2400} \leq 8, \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0. \end{aligned}$$

8.2 (20070601-nr.2)

(a)

If we introduce slack variables x_5 and x_6, to transform the inequality constraints to equality constraints, we obtain a LP-problem on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -1 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (-3,\ 4,\ -2,\ 5,\ 0,\ 0)$.

The initial solution should have the basic variables x_5 and x_6, i.e. $\beta = (5, 6)$ and $\delta = (1, 2, 3, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the equation system
$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, with solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The reduced costs of the non-basic variables are obtained from

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (-3,\, 4, -2,\, 5) - (0,\, 0) \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} = (-3,\, 4, -2,\, 5)$.

Since $r_{\delta_1} = r_1 = -3$ is the smallest, and < 0, we should let x_1 become the new basic variable.

We then need to compute the vector $\bar{\mathbf{a}}_1$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_1 = \mathbf{a}_1$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{11} \\ \bar{a}_{21} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_1 = \begin{pmatrix} \bar{a}_{11} \\ \bar{a}_{21} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

The biggest value that the new basic variable x_1 can be assigned to is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i1}} \mid \bar{a}_{i1} > 0 \right\} = \min \left\{ \frac{8}{1},\, \frac{4}{1} \right\} = \frac{4}{1} = \frac{\bar{b}_2}{\bar{a}_{21}}.$$

The minimizing index is $i = 2$, and hence $x_{\beta_2} = x_6$ will leave the set of basic variables.
Its place will be taken by x_1.

Hence $\beta = (5, 1)$ and $\delta = (6, 2, 3, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the equation system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$.

The reduced costs of the non-basic variables are given by

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (0,\, 4, -2,\, 5) - (0, -3) \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} = (3,\, 1,\, 1,\, 2)$.

Since $\mathbf{r}_\delta \geq \mathbf{0}$ the current basic solution is optimal.

Hence the point $x_1 = 4,\ x_2 = 0,\ x_3 = 0,\ x_4 = 0$ is optimal to the original problem. The optimal value is $z = -12$.

(b)

Now suppose $\mathbf{c}^{\mathsf{T}} = (-3,\ 4, -2,\ 2,\ 0,\ 0)$ instead of $(-3,\ 4, -2,\ 5,\ 0,\ 0)$.

If we start from the final solution above, with $\beta = (5, 1)$ and $\delta = (6, 2, 3, 4)$,

$$\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},\ \mathbf{A}_\delta = \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix},\ \bar{\mathbf{b}} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

is still valid.

But the reduced costs of the non-basic variables are now given by

$$\mathbf{r}_\delta^{\mathsf{T}} = \mathbf{c}_\delta^{\mathsf{T}} - \mathbf{y}^{\mathsf{T}}\mathbf{A}_\delta = (0,\ 4, -2,\ 2) - (0, -3) \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} = (3,\ 1,\ 1, -1).$$

Since $r_{\delta_4} = r_4 = -1$ is the smallest, and < 0, we should let x_4 become the new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_4$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$,

i.e. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

Since $\bar{\mathbf{a}}_4 \leq \mathbf{0}$, x_4 can increase without constraint, the value of the objective function goes to $-\infty$.
Hence the problem lacks finite optimal solution and the algorithm is canceled.

Extra comments (that are not required):

If you set $x_4 = t$ and let t increase from 0, while the other non-basic variables stays at 0, the objective function is changed according to $z = \bar{z} + r_4 t = -12 - t$, while the values of the basic variables are affected according to $\mathbf{x}_\beta = \bar{\mathbf{b}} - \bar{\mathbf{a}}_4 t$,

i.e. $\begin{pmatrix} x_5 \\ x_1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \end{pmatrix} t\,.$

This can be written as $\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \\ x_6(t) \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \\ 4 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{x}_0 + t \cdot \mathbf{d}.$

Then $\mathbf{A}\mathbf{x}(t) = \mathbf{b}$ and $\mathbf{x}(t) \geq \mathbf{0}$ for all $t \geq 0$, i.e. $\mathbf{x}(t)$ is a feasible solution for every $t \geq 0$,

while $\mathbf{c}^{\mathsf{T}}\mathbf{x}(t) = \mathbf{c}^{\mathsf{T}}\mathbf{x}_0 + t \cdot \mathbf{c}^{\mathsf{T}}\mathbf{d} = -12 - t \to -\infty$ when $t \to +\infty$.

(c)

If the primal problem is on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

then the dual problem is on the form maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ s.t. $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$, which here becomes:

$$\begin{array}{rrcrcr}
\text{maximize} & 8y_1 & + & 4y_2 & & \\
\text{s.t.} & y_1 & + & y_2 & \leq & -3, \\
 & y_1 & - & y_2 & \leq & 4, \\
 & -y_1 & + & y_2 & \leq & -2, \\
 & -y_1 & - & y_2 & \leq & 5, \\
 & y_1 & & & \leq & 0, \\
 & & & y_2 & \leq & 0.
\end{array}$$

If you draw the feasible region to this problem in a figure with y_1 and y_2 on the axis, you see that it is a pentagon with corners in $(-0.5\,,\,-2.5)$, $(\,0\,,\,-3)$, $(\,0\,,\,-4)$, $(-0.5\,,\,-4.5)$ and $(-1.5\,,\,-3.5)$.

(d)

In the figure above we will now replace the constraint $-y_1 - y_2 \leq 5$ with $-y_1 - y_2 \leq 2$.
But then we see that there is no $\mathbf{y}$ that fulfills both $y_1 + y_2 \leq -3$ and $-y_1 - y_2 \leq 2$. (Which is also understood when summing these inequalities.) Hence the dual problem has no feasible solutions, which is what we expected, since the primal problem had feasible solutions, but did not have a (finite) optimal solution.

8.3 (20070307-nr.1a)

(a) Let:

X_A = number of hectoliters Applecider produced each week.
X_P = number of hectoliters Pearcider produced each week.
X_B = number of hectoliters Mixed cider produced each week.
X_C = number of hectoliters Original cider produced each week.

We get the problem formulation

$$\begin{array}{rl}
\text{maximize} & 196X_A + 210X_P + 280X_B + 442X_C \\
\text{s.t.} & 1.6X_A + 1.8X_P + 3.2X_B + 5.4X_C \leq 80 \\
 & 1.2X_A + 1.2X_P + 1.2X_B + 1.8X_C \leq 40 \\
 & -0.8X_A + 0.2X_P + 0.2X_B + 0.2X_C \leq 0 \\
 & -0.3X_A + 0.7X_P - 0.3X_B - 0.3X_C \leq 0 \\
 & X_A \geq 0,\;\; X_P \geq 0,\;\; X_B \geq 0,\;\; X_C \geq 0.
\end{array}$$

8.4 (20070307-nr.2)

(a)

The current LP-problem on standard form is

$$\begin{array}{rl}
\text{minimize} & \mathbf{c}^\mathsf{T}\mathbf{x} \\
\text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\
 & \mathbf{x} \geq \mathbf{0}\,,
\end{array}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 & 0 \\ 2 & 3 & 4 & 5 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (0,\,0,\,0,\,0,\,1,\,1)$.

The natural starting basic solution has the basic variables x_5 and x_6, which means that $\beta = (5, 6)$ and $\delta = (1, 2, 3, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system of equations $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

The reduced costs for the non-basic variables is given by

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (0,\ 0,\ 0,\ 0) - (1,\ 1) \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix} = (-3,\ -5,\ -7,\ -9)$.

Since $r_{\delta_4} = r_4 = -9$ is the smallest, and < 0, we let x_4 become new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_4$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$.

The biggest value that the new basic variable x_4 can be incremented to is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \min \left\{ \frac{10}{4},\ \frac{12}{5} \right\} = \frac{12}{5} = \frac{\bar{b}_2}{\bar{a}_{24}}.$$

The minimizing index is $i = 2$ and hence $x_{\beta_2} = x_6$ can no longer be a basic variable.
Its place is taken by x_4.

Hence, now $\beta = (5, 4)$ and $\delta = (1, 2, 3, 6)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 4 \\ 0 & 5 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 1 \end{bmatrix}$.

The values of the basic variables in the basic solution is given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system of equations $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 4 \\ 0 & 5 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 2.4 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 0 \\ 4 & 5 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.8 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by
$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (0,\ 0,\ 0,\ 1) - (1,\ -0.8)\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 1 \end{bmatrix} = (0.6,\ 0.4,\ 0.2,\ 0.8).$

Since $\mathbf{r}_\delta \geq \mathbf{0}$ the current basic solution is optimal.

Hence the point $x_1 = 0,\ x_2 = 0,\ x_3 = 0,\ x_4 = 2.4,\ x_5 = 0.4,\ x_6 = 0$ is optimal.

The optimal value is given by $x_5 + x_6 = 0.4 > 0$.

If there is a feasible solution $(\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)^\mathsf{T}$ to the original system, then $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4,\ 0,\ 0)^\mathsf{T}$ is a feasible solution to the LP-problem with the objective function value $= 0$.
Since the objective function value never can be negative $\hat{\mathbf{x}}$ is an optimal solution to the LP-problem.

Hence: If there is a feasible solution to the original system, then the optimal solution to the LP-problem $= 0$. But above we found that the optimal value of the LP-problem is 0.4, which means that the original problem can not have a solution.

(b)

Suppose that $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \hat{y}_3)^\mathsf{T}$ is an optimal solution to the dual problem.

From the duality theorem it follows that $\hat{y}_3 = \hat{x}_3 = -0.2$.

From the complementarity theorem follows that since $\hat{x}_1 > 0$, then it must hold that
$-\hat{y}_1 + 3\hat{y}_2 + \hat{y}_3 = 0$ and since $\hat{x}_2 > 0$ then it must hold that $2\hat{y}_1 - 4\hat{y}_2 + \hat{y}_3 = 0$.

Together this gives that $\hat{y}_1 = 0.7$ and $\hat{y}_2 = 0.3$, which also fulfills the other constraints in the dual problem.

Hence: $\hat{y}_1 = 0.7$, $\hat{y}_2 = 0.3$ and $\hat{y}_3 = -0.2$ is an optimal solution to the dual problem (in reality the unique optimal solution).

8.5 (20070603-nr.1)

(a)

We here have a LP-problem on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 4 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (4,\ 3,\ 2,\ 3,\ 4)$.

The given starting solution should have the basic variables x_1 and x_5, which implies that
$\beta = (1, 5)$ and $\delta = (2, 3, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix}$.

The values of the basic variables in in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system of equations $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1.25 \\ 0.75 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

The reduced costs for the non-basic variables is given by

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (3,\ 2,\ 3) - (1,\ 1) \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix} = (-1,\ -2,\ -1)$.

Since $r_{\delta_2} = r_3 = -2$ is the smallest, and < 0, we will let x_3 become new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_3$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_3 = \mathbf{a}_3$,

i.e. $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

The biggest value to which the new basic variable x_3 can be incremented to is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{1.25}{0.5}, \frac{0.75}{0.5} \right\} = \frac{0.75}{0.5} = \frac{\bar{b}_2}{\bar{a}_{23}}.$$

The minimizing index is $i = 2$, and therefore $x_{\beta_2} = x_5$ is removed as basic variable.
Its place is taken by x_3.

Hence, now $\beta = (1, 3)$ and $\delta = (2, 5, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 4 & 3 \end{bmatrix}$.

The values of the basic variables in the basic solution is given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ in computed from the equation system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 4 & 0 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta = (3,\ 4,\ 3) - (1,\ 0) \begin{bmatrix} 3 & 0 & 1 \\ 1 & 4 & 3 \end{bmatrix} = (0,\ 4,\ 2)$.

Since $\mathbf{r}_\delta \geq \mathbf{0}$, the current basic solution is optimal.

Hence the point $x_1 = 0.5$, $x_2 = 0$, $x_3 = 1.5$, $x_4 = 0$, $x_5 = 0$ is optimal.

The optimal value is $\mathbf{c}^\mathsf{T}\mathbf{x} = 5$.

(b)

Since $r_{\delta_1} = r_2 = 0$ we can let x_2 become new basic variable (and increase from zero)
without that the objective function value is changed.

Then we need to compute the vector $\bar{\mathbf{a}}_2$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_2 = \mathbf{a}_2$,

i.e. $\begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_2 = \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$.

The biggest value that the new basic variable x_2 can be incremented to is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i2}} \mid \bar{a}_{i2} > 0 \right\} = \min \left\{ \frac{0.5}{0.5}, \frac{1.5}{0.5} \right\} = \frac{0.5}{0.5} = \frac{\bar{b}_1}{\bar{a}_{12}}.$$

The minimizing index is $i = 1$, and therefore $x_{\beta_1} = x_1$ is no longer kept as basic variable.
Its place is taken by x_2.

Hence, now $\beta = (2, 3)$ and $\delta = (1, 5, 4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$, while $\mathbf{A}_\delta = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 4 & 3 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Hence the new basic solution is $x_1 = 0$, $x_2 = 1$, $x_3 = 1$, $x_4 = 0$, $x_5 = 0$.

The value of the objective function is (of course) still $\mathbf{c}^\mathsf{T}\mathbf{x} = 5$, so also this basic solution is optimal.

Here you can stop, but as a control we can go on:

The vector $\mathbf{y}$ with the values of the simplex multipliers are now obtained from the system $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by

$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta = (4,\ 4,\ 3) - (1,\ 0) \begin{bmatrix} 4 & 0 & 1 \\ 0 & 4 & 3 \end{bmatrix} = (0,\ 4,\ 2)$.

Since $\mathbf{r}_\delta \geq \mathbf{0}$ the new basic solution is optimal.

(c)

If the primal problem is on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

then the dual problem is on the form: maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ s.t. $\mathbf{A}^\mathsf{T}\mathbf{y} \le \mathbf{c}$, which written out becomes:

$$\begin{array}{rrcrcl}
\text{maximize} & 5y_1 & + & 3y_2 & & \\
\text{s.t.} & 4y_1 & & & \le & 4, \\
& 3y_1 & + & y_2 & \le & 3, \\
& 2y_1 & + & 2y_2 & \le & 2, \\
& y_1 & + & 3y_2 & \le & 3, \\
& & & 4y_2 & \le & 4.
\end{array}$$

It is well-known that an optimal solution to this dual problem is given by the vector $\mathbf{y}$ with
"the simplex multipliers" in the optimal basic solution in the (a)-task, i.e. $\mathbf{y} = (1,\ 0)^\mathsf{T}$.
You can quickly verify that this is a feasible solution of the dual problem above. Further the dual objective function value is $\mathbf{b}^\mathsf{T}\mathbf{y} = 5y_1 + 3y_2 = 5 - 0 = 5 =$ the optimal value to the primal problem in the (a)-task.

8.6 (20060308-nr.1)

(a)

Introduce slack variables x_4, x_5 and x_6 so that the problem is on standard form

$$\begin{array}{rl}
\text{minimize} & \mathbf{c}^\mathsf{T}\mathbf{x} \\
\text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b}, \\
& \mathbf{x} \ge \mathbf{0},
\end{array}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$ and $\mathbf{c} = (1,\ 5,\ 2,\ 0,\ 0,\ 0)^\mathsf{T}$.

The given starting solution corresponds to that x_1, x_2 and x_3 are basic variables, i.e. that $\beta = (1, 2, 3)$ and $\delta = (4, 5, 6)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

This is the basic solution we were supposed to start from.

The vector $\mathbf{y}$ with the values of the simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by $\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta =$
$= (0,\ 0,\ 0) - (2,\ -1,\ 3)\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = (2,\ -1,\ 3).$

Since $r_{\delta_2} = r_5 = -1$ is the smallest, and < 0, we will let x_5 become new basic variable.

Then we need to calculate the vector $\bar{\mathbf{a}}_5$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_5 = \mathbf{a}_5$,

i.e. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \\ \bar{a}_{35} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_5 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \\ \bar{a}_{35} \end{pmatrix} =$
$\begin{pmatrix} -0.5 \\ 0.5 \\ -0.5 \end{pmatrix}$.

The biggest value that the new basic variable x_5 can be incremented to is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \frac{\bar{b}_2}{\bar{a}_{25}} = \frac{1}{0.5}.$$

The minimizing index is $i = 2$, and hence $x_{\beta_2} = x_2$ can no longer be a basic variable.

Hence $\beta = (1, 5, 3)$ and $\delta = (4, 2, 6)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ in computed from the system of equations
$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$,

i.e. $\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$.

The vector $\mathbf{y}$ with values of the simplex multipliers is obtained from the system
$\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, with the solution $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by $\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta =$
$= (0,\ 5,\ 0) - (1,\ 0,\ 2)\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} = (1,\ 2,\ 2).$

Since $\mathbf{r}_\delta \geq \mathbf{0}$ the current basic solution is optimal.
Hence the point $x_1 = 2,\ x_2 = 0,\ x_3 = 2,\ x_4 = 0,\ x_5 = 2,\ x_6 = 0$ is optimal.
The optimal value is $\mathbf{c}^\mathsf{T}\mathbf{x} = 6$.

(b)
If the primal problem is on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

then the dual problem

$$\text{maximize } \mathbf{b}^\mathsf{T}\mathbf{y} \ \text{ s.t. } \ \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c},$$

which written out becomes:

$$\begin{aligned} \text{maximize} \quad & 2y_1 + 2y_2 + 2y_3 \\ \text{s.t.} \quad & y_1 + y_2 \leq 1, \\ & y_1 + y_3 \leq 5, \\ & y_2 + y_3 \leq 2, \\ & -y_1 \leq 0, \\ & -y_2 \leq 0, \\ & -y_3 \leq 0. \end{aligned}$$

It is well known that an optimal solution to this dual problem is given by the vector $\mathbf{y}$ with the simplex multipliers in the optimal basic solution in the (a)-task, i.e. $\mathbf{y} = (1,\ 0,\ 2)^\mathsf{T}$.
You can quickly verify that this is a feasible solution to the dual problem.
Further the optimal value $\mathbf{b}^\mathsf{T}\mathbf{y} = 6 =$ the optimal value to the primal problem above.

8.7 (20060308-nr.4b)

(b).

First a comment. Since according to the conditions all $b_i > 0$, for example $\mathbf{x} = \mathbf{0}$ is a point which is both in Ω, and do not touch any of the walls in Ω. Hence there is place for at least a small sphere in Ω.

Now to the formulation. Let $\mathbf{x} \in \Omega$ be the mean point to our searched sphere and let r be its radius. We can assume that $\mathbf{x}$ does not touch any of the walls in Ω, i.e. that $\mathbf{a}_i^\mathsf{T}\mathbf{x} < b_i$ for all i. Then the distance $d_i(\mathbf{x})$ from the point $\mathbf{x}$ to the plane $\mathcal{P}_i$ can be written $d_i(\mathbf{x}) = (b_i - \mathbf{a}_i^\mathsf{T}\mathbf{x})/|\mathbf{a}_i|$.

If the sphere has the mean point in $\mathbf{x}$ it fits in Ω if and only if its radius r fulfills that $r \leq d_i(\mathbf{x})$ for all $i = 1, \ldots, m$.

We hence obtain the following LP-problem in the variables $\mathbf{x} \in I\!R^3$ and $r \in I\!R$ (four variables):

$$\begin{aligned} \text{maximize} \quad & r \\ \text{s.t.} \quad & r \leq d_i(\mathbf{x}),\ i = 1, \ldots, m. \end{aligned}$$

or, equivalently

$$\begin{aligned} \text{maximize} \quad & r \\ \text{s.t.} \quad & |\mathbf{a}_i|\, r + \mathbf{a}_i^\mathsf{T}\mathbf{x} \leq b_i,\ i = 1, \ldots, m. \end{aligned}$$

8.8 (20051024-nr.1)

(a)

Introduce the slack variables x_4 and x_5, and change the sign of the objective function so that the problem is on standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{c} = (-1,\ -1,\ -2,\ 0,\ 0)^\mathsf{T}$.

That x_4 and x_5 are basic variables corresponds to that $\beta = (4,5)$ and $\delta = (1,2,3)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = [\mathbf{a}_4\ \ \mathbf{a}_5] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system

$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

This is a *feasible* basic solution since $\bar{\mathbf{b}} \geq \mathbf{0}$.

The values of the simplex multipliers are given by the system $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, with the solution $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta =$

$= (-1,\ -1,\ -2) - (0,\ 0) \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} = (-1,\ -1,\ -2)$.

Since $\bar{c}_{\delta_3} = \bar{c}_3 = -2$ is the smallest, and < 0, we will let x_3 become the new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_3$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_3 = \mathbf{a}_3$,

i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

The biggest value that the new basic variable x_3 can be incremented to is given by

$$x_3^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \frac{\bar{b}_1}{\bar{a}_{13}} = \frac{1}{1}.$$

The minimizing index is $i = 1$, and hence $x_{\beta_1} = x_4$ can no longer be a basic variable.

Hence, now $\beta = (3,5)$ and $\delta = (1,2,4)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = [\mathbf{a}_3 \;\; \mathbf{a}_5] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$.

The values of the variables of the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system

$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

This is, as expected, a feasible basic solution.

The values of the simplex multipliers are given by the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$, with the solution $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta =$
$= (-1,\; -1,\; 0) - (-2,\; 0) \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = (1,\; -3,\; 2)$.

Since $\bar{c}_{\delta_2} = \bar{c}_2 = -3$ is the smallest, and < 0, we will let x_2 become new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_2$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_2 = \mathbf{a}_2$,

i.e. $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_2 = \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Since $\bar{\mathbf{a}}_2 \leq \mathbf{0}$, x_2 can increment without limit, and the objective function value (for the minimization problem) goes to $-\infty$. Hence the problem lacks finite optimal value and the algorithm is canceled.

(b) When x_2 is increased from 0 the values of the basic variables are affected according to $\mathbf{x}_\beta = \bar{\mathbf{b}} - \bar{\mathbf{a}}_2 x_2$,

i.e. $\begin{pmatrix} x_3 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \end{pmatrix} x_2$, while x_1 and x_4 stays at 0. Expressed in the original variables x_1, x_2 and x_3 this corresponds to that $x_2 = t$, $x_3 = 1 + t$ and $x_1 = 0$,

which can be written as $\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \mathbf{x}_0 + t \cdot \mathbf{d}$.

Then $\mathbf{Px}(t) \leq \mathbf{b}$ and $\mathbf{x}(t) \geq \mathbf{0}$ for all $t \geq 0$, while $\mathbf{q}^\mathsf{T}\mathbf{x}(t) = 1 + 3t \to +\infty$ when $t \to +\infty$.

(c) If the primal problem is

$$\text{maximize } \mathbf{q}^\mathsf{T}\mathbf{x} \text{ s.t. } \mathbf{Px} \leq \mathbf{b} \text{ and } \mathbf{x} \geq \mathbf{0},$$

then the corresponding dual problem

$$\text{minimize } \mathbf{b}^\mathsf{T}\mathbf{y} \text{ s.t. } \mathbf{P}^\mathsf{T}\mathbf{y} \geq \mathbf{q} \text{ and } \mathbf{y} \geq \mathbf{0},$$

which written out becomes

minimize y_1+y_2 s.t. $y_1+y_2 \geq 1,\ -y_1+y_2 \geq 1,\ y_1-y_2 \geq 2,\ y_1 \geq 0$ and $y_2 \geq 0$.

With the help of a figure you immediately see that there is no $\mathbf{y}$ that fulfills both $-y_1+y_2 \geq 1$ and $y_1-y_2 \geq 2$. (Which is also understood if you add these inequalities.) Hence the dual problem lacks feasible solutions, which is exactly what we expected, since the primal problem had feasible solutions, but lacked finite optimal solution.

8.9 (20051024-nr.3)

Let $S = \{1,2,3,4,5,6,7,8,9,10,11,12\}$ and $B = \{(i,j) \mid i \in S,\ j \in S \text{ and } i \neq j\}$.

We have the following given:

12 constants p_i, for $i \in S$, 12 constants q_j, for $j \in S$,
and 132 constants r_{ij}, for $(i,j) \in B$.

Further we have 132 unknowns x_{ij}, for $(i,j) \in B$.
These x_{ij} become our variables in the optimization formulation.

The demand on consistency implies that

$\sum_{j \in J(i)} x_{ij} = p_i$, for $i \in S$, and that $\sum_{i \in I(j)} x_{ij} = q_j$, for $j \in S$,

where $J(i) = \{j \in S \mid j \neq i\}$ and $I(j) = \{i \in S \mid i \neq j\}$.

Furthermore we demand that $x_{ij} \geq 0$ for $(i,j) \in B$.

Given that the variables x_{ij} fulfill these consistency demands we want to choose them as close to the constants r_{ij} as possible. There are several ways to define what you mean by "close". Here are some possible measures of the "distance" between the variables x_{ij} and the constants r_{ij}:

(A1): $\sum_{(i,j) \in B} (x_{ij} - r_{ij})^2$.

(A2): $\sum_{(i,j) \in B} |x_{ij} - r_{ij}|$.

(A3): $\max_{(i,j) \in B} \{|x_{ij} - r_{ij}|\}$.

(A1) leads to a convex QP-problem, while (A2) and (A3) both leads to LP-problems. If you for instance use (A3) you get the following LP-problem in the variables x_{ij} and z:

$$
\begin{aligned}
\text{minimize} \quad & z \\
\text{s.t.} \quad & x_{ij} + z \geq r_{ij}, && \text{for } (i,j) \in B, \\
& x_{ij} - z \leq r_{ij}, && \text{for } (i,j) \in B, \\
& \sum_{j \in J(i)} x_{ij} = p_i, && \text{for } i \in S, \\
& \sum_{i \in I(j)} x_{ij} = q_j, && \text{for } j \in S, \\
& x_{ij} \geq 0, && \text{f0r } (i,j) \in B.
\end{aligned}
$$

The first two constraints says that $z \geq |x_{ij} - r_{ij}|$ for all $(i, j) \in B$.
Hence in the optimal it will hold that $z = \max\limits_{(i,j)\in B} \{|x_{ij} - r_{ij}|\}$.

8.10 (20050331-nr.2)

Every sub-exercise can be handled by the following methodology.

First draw the feasible region to P in a coordinate system with x_1 and x_2 on the axes.
Then draw (in the same figure) the curvature of the objective function $\mathbf{c}^\mathsf{T}\mathbf{x}$.
Then make your conclusion about the optimal solution of P.

Draw the feasible region of D in a coordinate system with y_1 and y_2 on the axes.
Then draw (in the same figure) the curvature of the objective function $\mathbf{b}^\mathsf{T}\mathbf{y}$.
The make conclusions about the optimal solution to D.

(a) Here $\mathbf{b} = (\,1,\,-1\,)^\mathsf{T}$ and $\mathbf{c} = (-2,\,2\,)^\mathsf{T}$.
Optimal solution to P is $\mathbf{x} = (1, 1)^\mathsf{T}$, with the optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = 0$.
Optimal solution to D is $\mathbf{y} = (2, 2)^\mathsf{T}$, with the optimal value $\mathbf{b}^\mathsf{T}\mathbf{y} = 0$.
Both P and D hence have the the optimal value $= 0$.

(b) Here $\mathbf{b} = (\,1,\,-1\,)^\mathsf{T}$ and $\mathbf{c} = (\,2,\,2\,)^\mathsf{T}$.
Optimal solution to P is $\mathbf{x} = (0, 1)^\mathsf{T}$, with the optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = 2$.
Optimal solution to D is $\mathbf{y} = (2, 0)^\mathsf{T}$, with the optimal value $\mathbf{b}^\mathsf{T}\mathbf{y} = 2$.
Both P and D hence have the the optimal value $= 2$.

(c) Here $\mathbf{b} = (-1,\,-1\,)^\mathsf{T}$ and $\mathbf{c} = (-2,\,2\,)^\mathsf{T}$.
Optimal solution to P is $\mathbf{x} = (1, 0)^\mathsf{T}$, with the optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = -2$.
Optimal solution to D is $\mathbf{y} = (0, 2)^\mathsf{T}$, with the optimal value $\mathbf{b}^\mathsf{T}\mathbf{y} = -2$.
Both P and D hence have the the optimal value $= -2$.

(d) Here $\mathbf{b} = (-1,\,1\,)^\mathsf{T}$ and $\mathbf{c} = (\,2,\,2\,)^\mathsf{T}$.
The problem P now lacks feasible solutions.
For the problem D holds that if you let $\mathbf{y}(t) = t \cdot (0, 1)^\mathsf{T}$, then $\mathbf{y}(t)$ is feasible to D for
all $t \geq 0$, and $\mathbf{b}^\mathsf{T}\mathbf{y}(t) \longrightarrow \infty$ where $t \longrightarrow \infty$. This means that D lacks a finite optimal solution.
In this case you say that both P and D have the optimal value $+\infty$.

(e) Here $\mathbf{b} = (-1,\,-1\,)^\mathsf{T}$ and $\mathbf{c} = (\,2,\,-2\,)^\mathsf{T}$.
The problem D now lacks feasible solutions.
For the problem P holds that if you let $\mathbf{x}(t) = t \cdot (0, 1)^\mathsf{T}$ then $\mathbf{x}(t)$ is feasible to P for
all $t \geq 0$, and $\mathbf{c}^\mathsf{T}\mathbf{x}(t) \longrightarrow -\infty$ where $t \longrightarrow \infty$. This means that P lacks a finite optimal solution.
In this case you say that both P and D have the optimal value $-\infty$.

8.11 (20050307-nr.2)

(a) That x_1, x_3 and x_5 are basic variables corresponds to that $\beta = (1, 3, 5)$ and $\delta = (2, 4, 6)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = [\mathbf{a}_1\ \ \mathbf{a}_3\ \ \mathbf{a}_5] = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$.

This matrix is non- singular, i.e. a basic matrix, since it is triangular with nonzero elements on the diagonal.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is computed from the system

$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \\ 1 \end{pmatrix}$.

This is a *feasible* basic solution Since $\bar{\mathbf{b}} \geq \mathbf{0}$.

(b) The values of the simplex multipliers are given by the system $\mathbf{A}_\beta^{\mathsf{T}} \mathbf{y} = \mathbf{c}_\beta$,

i.e. $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$, with the solution $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.5 \\ 1 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^{\mathsf{T}} = \mathbf{c}_\delta - \mathbf{y}^{\mathsf{T}} \mathbf{A}_\delta =$

$= (1,\ 1,\ 1) - (1,\ -0.5,\ 1) \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} = (1,\ -0.5,\ 1).$

Since $\bar{c}_{\delta_2} = \bar{c}_4 = -0.5 < 0$ we will let x_4 become new basic variable.

Then we need to compute the vector $\bar{\mathbf{a}}_4$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$,

i.e. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \\ \bar{a}_{34} \end{pmatrix} = \begin{pmatrix} 1 \\ -0.5 \\ 0 \end{pmatrix}$.

The biggest value that the new basic variable x_4 can be incremented to is given by

$$x_4^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i4}} \mid \bar{a}_{i4} > 0 \right\} = \frac{\bar{b}_1}{\bar{a}_{14}} = \frac{1}{1}.$$

The minimizing index is $i = 1$, and hence $x_{\beta_1} = x_1$ will no longer be basic variable.

Hence now $\beta = (4, 3, 5)$ and $\delta = (2, 1, 6)$.

The corresponding basic matrix is given by $\mathbf{A}_\beta = [\mathbf{a}_4\ \ \mathbf{a}_3\ \ \mathbf{a}_5] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$.

The values of the basic variables in the basic solution are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ in computed from the system

$$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}, \text{ i.e. } \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

This is as expected a feasible basic solution.

The values of the simplex multipliers are given by the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$,

$$\text{i.e. } \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \text{ with the solution } \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -0.5 \\ 1 \end{pmatrix}.$$

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T} \mathbf{A}_\delta =$

$$= (1,\ 2,\ 1) - (0.5,\ -0.5,\ 1) \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix} = (1.5,\ 0.5,\ 1).$$

Since $\bar{\mathbf{c}}_\delta \geq \mathbf{0}$ the current basic solution is optimal,
i.e. $x_4 = x_3 = x_5 = 1$ and $x_2 = x_1 = x_6 = 0$, with the optimal value $z = 4$.

(c) The dual problem can be written as:

$$\text{maximize } \mathbf{b}^\mathsf{T} \mathbf{y} \text{ s.t. } \mathbf{A}^\mathsf{T} \mathbf{y} \leq \mathbf{c},$$

with $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$ as above.

Hence we have three variables and six inequality constraints in the dual problem D.
The optimal solution to D is given by $\mathbf{y} = (0.5,\ -0.5,\ 1)^\mathsf{T}$ from the (b)-task above.
You can quickly confirm that this is a feasible solution to D.
Furthermore $\mathbf{b}^\mathsf{T} \mathbf{y} = 4 = \mathbf{c}^\mathsf{T} \mathbf{x}$, with $\mathbf{x}$ according to above.

8.12 (20050307-nr.5)

(a) We look for a vector $\mathbf{a} \in \mathbb{R}^n$ such that $\mathbf{a}^\mathsf{T} \mathbf{p}_i \leq 1$, $\mathbf{a}^\mathsf{T} \mathbf{q}_j \geq 1$ and $\mathbf{a} \geq \mathbf{0}$.
(Since $\mathbf{p}_1 = \mathbf{0}$ there exist no vector $\mathbf{a} \in \mathbb{R}^n$ such that $\mathbf{a}^\mathsf{T} \mathbf{p}_1 \geq 1$.)

With slack variables u_i and v_j, the above can be written as:
$\mathbf{p}_i^\mathsf{T} \mathbf{a} + u_i = 1,\ i = 1 \dots k,\ \ \mathbf{q}_j^\mathsf{T} \mathbf{a} - v_j = 1,\ j = 1 \dots \ell,\ \ \mathbf{a} \geq \mathbf{0},\ \mathbf{u} \geq \mathbf{0},\ \mathbf{v} \geq \mathbf{0}$.

To determine whether this system has a solution we form a Phase1-problem, with artificial variables $w_j,\ j = 1 \dots \ell$.

$$\begin{aligned} \text{minimize} \quad & w_1 + \cdots + w_\ell \\ \text{s.t.} \quad & \mathbf{p}_i^\mathsf{T} \mathbf{a} + u_i = 1, \quad i = 1 \dots k \\ & \mathbf{q}_j^\mathsf{T} \mathbf{a} - v_j + w_j = 1, \quad j = 1 \dots \ell \\ & \mathbf{a} \geq \mathbf{0},\ \mathbf{u} \geq \mathbf{0},\ \mathbf{v} \geq \mathbf{0},\ \mathbf{w} \geq \mathbf{0}. \end{aligned}$$

A feasible starting basic solution is obtained with $u_1 \dots u_k$ and $w_1 \dots w_\ell$ as basic variables.

If and only if the optimal solution to this LP-problem has all $w_j = 0$ there is a weakly separating plane, whose coefficients are given by optimal $\mathbf{a}$.

(b) Now we are looking for a strictly separating plane.

Such a plane exists if there is a number $\delta > 0$ and an $\mathbf{a} \in I\!R^n$ such that
$\mathbf{a}^\mathsf{T}\mathbf{p}_i \le 1-\delta,\ i = 1\dots k,\ \ \mathbf{a}^\mathsf{T}\mathbf{q}_j \ge 1+\delta,\ j = 1\dots\ell,\ \ \mathbf{a} \ge \mathbf{0}$.

With slack variables u_i and v_j the above can be written as:
$\mathbf{p}_i^\mathsf{T}\mathbf{a}+u_i+\delta = 1,\ i = 1\dots k,\ \mathbf{q}_j^\mathsf{T}\mathbf{a}-v_j-\delta = 1,\ j = 1\dots\ell,\ \mathbf{a} \ge \mathbf{0},\ \mathbf{u} \ge \mathbf{0},\ \mathbf{v} \ge \mathbf{0}$.

Now treat δ as a non-negative variable that shall be maximized. This leads to the following LP-problem.

$$\begin{aligned} \text{maximize}\quad & \delta \\ \text{s.t.}\quad & \mathbf{p}_i^\mathsf{T}\mathbf{a} + u_i + \delta = 1, \quad i = 1\dots k \\ & \mathbf{q}_j^\mathsf{T}\mathbf{a} - v_j - \delta = 1, \quad j = 1\dots\ell \\ & \mathbf{a} \ge \mathbf{0},\ \mathbf{u} \ge \mathbf{0},\ \mathbf{v} \ge \mathbf{0},\ \delta \ge 0. \end{aligned}$$

If and only if $\delta > 0$ in the optimal solution to this LP-problem there exists a strictly separating plane, whose coefficients are given by optimal $\mathbf{a}$.

A feasible starting basic solution is for example the optimal basic solution to the LP-problem in the a)-task above.

8.13 (20041016-nr.2)

(a). The cost vector to the LP-problem is obviously $\mathbf{c}^\mathsf{T} = (\,0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 1\,)$,
while the constraint matrix is given by $\mathbf{G} = \begin{bmatrix} 1 & 0 & -1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 & 1 \end{bmatrix}$.

The proposed solution corresponds to that the columns 1, 2 and 7 are the basic columns,

so that we obtain the basic matrix $\mathbf{G}_\beta = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. The values of the basic variables are given by the system

$\mathbf{G}_\beta \mathbf{x}_\beta = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$, with the solution $\begin{pmatrix} x_1 \\ x_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$, **OK!**

The values of the simplex multipliers are given by the system $\mathbf{G}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$, i.e.

$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, with the solution $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

The reduced costs for the non-basic variables are given by

$$\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T}\mathbf{G}_\delta = (\,0,\ 0,\ 1,\ 1\,) - (-1, -1,\ 1\,)\begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 1 & 0 & 0 \end{bmatrix} = (\,0,\ 1,\ 2,\ 2\,).$$

Since $\bar{\mathbf{c}}_\delta \geq \mathbf{0}$, the proposed feasible basic solution is optimal.

We can also conclude that the optimal value of the stated LP-problem is $= 1$.

(b). The answer is *NO* because of the following:

Suppose there were scalars $x_j \geq 0$ such that $\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 + \mathbf{a}_4 x_4$. These scalars would then together with $\mathbf{v} = \mathbf{0}$ be a feasible solution with the objective function value $\mathbf{e}^\mathsf{T}\mathbf{v} = 0$ to the (a)-task's LP-problem. But this is not possible, since we have already seen that the optimal value of this LP-problem is $= 1$.

(c). The dual problem corresponding to the LP-problem above can be written as

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{0}, \\ & \mathbf{I}\mathbf{y} \leq \mathbf{e}. \end{aligned}$$

Since the primal problem had the optimal value $= 1$ also the dual problem has the optimal value $= 1$ (the duality theorem). Hence every optimal solution to the dual problem fulfills that $\mathbf{b}^\mathsf{T}\mathbf{y} = 1$ and $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{0}$, and therefore $\mathbf{b}^\mathsf{T}\mathbf{y} > 0$ and $\mathbf{a}_j^\mathsf{T}\mathbf{y} \leq 0$ for all j.
But an optimal solution to the dual problem is given by the vector of multipliers in the optimal basic solution to the primal problem, i.e. $\mathbf{y} = (-1, -1,\ 1\,)^\mathsf{T}$. You can quickly verify that this vector $\mathbf{y}$ really fulfills the inequalities.

8.14 (20040415-nr.2)

(a)

If the primal problem is on the form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

then the corresponding dual problem is

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}\,. \end{aligned}$$

It is well-known that if
(i) $\mathbf{x}$ is a feasible solution to the primal problem,
(ii) $\mathbf{y}$ is a feasible solution to the dual problem, and
(iii) $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$,
then $\mathbf{x}$ and $\mathbf{y}$ are optimal solutions to their respective problem.
But the from the program proposed $\mathbf{x}$ and $\mathbf{y}$ fulfill

$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$ and $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$.
Hence they are optimal to P and D respectively.

(b) If the primal problem is on the form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

then the corresponding dual problem is

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}\,, \\ & \mathbf{y} \geq \mathbf{0}\,. \end{aligned}$$

But $\mathbf{x}$ and $\mathbf{y}$ from the (a)-task above fulfill that
$\mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c},\ \mathbf{y} \geq \mathbf{0}$ and $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$.
Hence they are optimal also to these two problems.

(c) If the problem is on the form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

with the given $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$, then it is realized by "inspection" that $\hat{\mathbf{x}} = (0,0,0,0,0,0)^\mathsf{T}$ is the unique optimal solution. This is a feasible solution with the objective function value $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = 0$, and for every other feasible solution $\mathbf{x}$ it holds that $\mathbf{c}^\mathsf{T}\mathbf{x} > 0$, since $\mathbf{c} > \mathbf{0}$, $\mathbf{x} \geq \mathbf{0}$ and at least one $x_j > 0$.
As an alternative to "inspection" you can introduce slack variables and let these be starting basic variables in the simplex method. You then immediately realize that the starting basic solution is optimal.

8.15 (20040310-nr.5)

(a): With $\beta = (1, 101)$ and $\delta = (2, 3, \ldots, 99, 100)$ it is obtained that $\mathbf{A}_\beta = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$.

The vector $\bar{\mathbf{b}}$ with values of the basic variables in the basic solution is given by the equation system $\mathbf{A}_\beta\bar{\mathbf{b}} = \mathbf{b}$, with the solution $\bar{\mathbf{b}} = (1,\ 2)^\mathsf{T}$ (i.e. $x_1 = 1$ and $x_{101} = 2$ in the current basic solution).

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta$, where $\mathbf{y}$ is given by the equation system $\mathbf{y}^\mathsf{T}\mathbf{A}_\beta = \mathbf{c}_\beta^\mathsf{T} = (50,\ 50)$, with the solution $\mathbf{y}^\mathsf{T} = (0.5,\ 0.5)$.

For every non-basic index j we hence have
$\bar{c}_j = c_j - \mathbf{y}^\mathsf{T}\mathbf{a}_j = |\,51 - j\,| - 0.5(101 - j) - 0.5(j-1) = |\,51 - j\,| - 50.$

The least reduced cost is obviously obtained for $j = 51$ so we set $k = 51$.
Then $\bar{c}_k = -50 < 0$ and the non-basic variable $x_k = x_{51}$ shall be new basic variable.

We compute $\bar{\mathbf{a}}_k$ from the equation system $\mathbf{A}_\beta \bar{\mathbf{a}}_k = \mathbf{a}_k$, where $\mathbf{a}_k = \mathbf{a}_{51} = (50,\, 50)^\mathsf{T}$, which gives that $\bar{\mathbf{a}}_k = (0.5,\, 0.5)^\mathsf{T}$.

Then we compare the quotas $\dfrac{\bar{b}_1}{\bar{a}_{1k}} = \dfrac{1}{0.5}$ and $\dfrac{\bar{b}_2}{\bar{a}_{2k}} = \dfrac{2}{0.5}$.

The first quota is the smallest, so $x_{\beta_1} = x_1$ shall be removed from the basis.

Now $\beta = (51, 101)$ and $\mathbf{A}_\beta = \begin{bmatrix} 50 & 0 \\ 50 & 100 \end{bmatrix}$.

The vector $\bar{\mathbf{b}}$ with the values of the basic variables in the basic solution is given by the equation system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, with the solution $\bar{\mathbf{b}} = (2,\, 1)^\mathsf{T}$ (i.e. $x_{51} = 2$ and $x_{101} = 1$ in the current basic solution).

The reduced costs for the non-basic variables are given by $\bar{\mathbf{c}}_\delta^\mathsf{T} = \mathbf{c}_\delta^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\delta$, where $\mathbf{y}$ is given by the equation system $\mathbf{y}^\mathsf{T}\mathbf{A}_\beta = \mathbf{c}_\beta^\mathsf{T} = (0,\, 50)$, with the solution $\mathbf{y}^\mathsf{T} = (-0.5,\, 0.5)$.

For every non-basic index j we hence have that
$\bar{c}_j = c_j - \mathbf{y}^\mathsf{T}\mathbf{a}_j = |\,51-j\,| + 0.5(101-j) - 0.5(j-1) = |\,51-j\,| + (51-j) \geq 0$,
with equality for all $j \geq 51$.

The current basic solution $x_{51} = 2$, $x_{101} = 1$ and the other $x_j = 0$ is hence optimal.

The optimal value to the problem is $c_{51}x_{51} + c_{101}x_{101} = 50$.

(b): Suppose that $\beta = (p,\, q)$ with $p < q$, then $\mathbf{A}_\beta = \begin{bmatrix} 101-p & 101-q \\ p-1 & q-1 \end{bmatrix}$.

The vector $\bar{\mathbf{b}}$ with the values of the basic variables in the basic solution is given by the equation system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$.

Computations and simplifications give that $x_p = \bar{b}_1 = \dfrac{3q-203}{q-p}$ and $x_q = \bar{b}_2 = \dfrac{203-3p}{q-p}$.

The basic solution is hence feasible if and only if $3q - 203 \geq 0$ and $203 - 3p \geq 0$, i.e. if and only if $q \in \{68, 69, 70, \ldots, 100, 101\}$ and $p \in \{1, 2, 3, \ldots, 66, 67\}$.

This gives in total $34 \cdot 67 = 2278$ feasible basic solutions.

(c): Suppose that $\beta = (p,\, q)$ with $p \in \{1, 2, 3, \ldots, 66, 67\}$ and $q \in \{68, 69, 70, \ldots, 100, 101\}$.

Then $x_p = \dfrac{3q-203}{q-p} > 0$ and $x_q = \dfrac{203-3p}{q-p} > 0$.

The objective function value of the basic solution is given by

$\bar{z} = c_p x_p + c_q x_q = |\,51-p\,|\, x_p + |\,51-q\,|\, x_q = |\,51-p\,|\, x_p + (q-51)\, x_q$, since $q \geq 68 > 51$.

We obtain two cases, depending on whether $p \geq 51$ or $p < 51$ (while of course $q \geq 68$ in both cases).

If $p \geq 51$, then $|\,51-p\,| = p - 51$, and then

$$\bar{z} = (p-51) \cdot \frac{3q-203}{q-p} + (q-51) \cdot \frac{203-3p}{q-p} = \;\ldots\; = 50.$$

If $p < 51$, then $|\,51 - p\,| = 51 - p$, and then

$\bar{z} = (51-p)\cdot\dfrac{3q-203}{q-p} + (q-51)\cdot\dfrac{203-3p}{q-p} \;=\; \ldots \;=\; 50 + \dfrac{2(51-p)(3q-203)}{q-p} \;>$ 50.

The basic solution is hence optimal if and only if
$q \in \{68, 69, \ldots, 100, 101\}$ and $p \in \{51, 52, \ldots, 66, 67\}$.

This gives in total $34 \cdot 17 = 578$ optimal feasible basic solutions.

9. Network problems

9.1 (20070601-nr.1b)

Given the feasible basic solution we get the corresponding simplex multipliers u_i and v_j, calculated from the relation $c_{ij} = u_i - v_j$ for the basic variables and $v_4 = 0$.

c_{ij}	customer 1	customer 2	customer 3	customer 4	u_i
fac 1	116	125	136		147
fac 2			125	136	136
fac 3				125	125
fac 4				116	116
v_j	31	22	11	0	

From that we get the following reduced cost r_{ij} for the non-basic variables computed from the relation $r_{ij} = c_{ij} - u_i + v_j$.

r_{ij}	customer 1	customer 2	customer 3	customer 4	u_i
fac 1				2	147
fac 2	4	2			136
fac 3	10	6	2		125
fac 4	16	10	4		116
v_j	31	22	11	0	

Since all $r_{ij} \geq 0$ the proposed basic solution is optimal.

9.2 (20070307-nr.1b)

(b)

That the problem is a minimum cost flow problem follows from the fact that every column in $\mathbf{A}$ consists of one element $+1$, one element -1, and the rest zeros. Every row in $\mathbf{A}$ then corresponds to a node in the network problem and every column in $\mathbf{A}$ corresponds to an edge in the network, i.e. an edge *from* the node that corresponds to the row with $+1$ *to* the node which corresponds to the row with -1.
Hence the network consists of 6 nodes and the set of edges
$\mathcal{B} = \{(1,4),\ (1,5),\ (1,6),\ (2,4),\ (2,5),\ (2,6),\ (3,4),\ (3,5),\ (3,6)\}$.
Node 1, 2 and 3 are source nodes, while node 4, 5 and 6 are sink nodes.

In the following we denote the variables x_{ij} and the corresponding costs c_{ij}, i.e. $\mathbf{x} = (x_{14}, x_{15}, x_{16}, x_{24}, x_{25}, x_{26}, x_{34}, x_{35}, x_{36})^\mathsf{T}$
and $\mathbf{c} = (c_{14},\ c_{15},\ c_{16},\ c_{24},\ c_{25},\ c_{26},\ c_{34},\ c_{35},\ c_{36})^\mathsf{T}$.

The proposed solution fulfills $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\hat{\mathbf{x}} \geq \mathbf{0}$.
Furthermore it corresponds to a spanning tree in the network with the basic edges
$\mathcal{B}_\beta = \{(1,4),\ (1,6),\ (2,6),\ (3,5),\ (3,6)\}$.
The proposed solution $\hat{\mathbf{x}}$ is hence a feasible basic solution.
The reduced costs are now given from the formula

$$r_{ij} = c_{ij} - y_i + y_j \quad \text{for all non-basic edges,}$$

where the scalars (simplex multipliers) y_i are given from the formula

$$y_i - y_j = c_{ij} \text{ for all basic edges and } y_6 = 0.$$

The scalars y_i can for example be computed in the following order:
First $y_6 = 0$, which holds per definition.
The basic edge $(3,6)$ then gives that $y_3 - y_6 = c_{36}$, i.e. $y_3 = c_{36} = 4$.
The basic edge $(2,6)$ then gives that $y_2 - y_6 = c_{26}$, i.e. $y_2 = c_{26} = 4$.
The basic edge $(1,6)$ then gives that $y_1 - y_6 = c_{16}$, i.e. $y_1 = c_{16} = 4$.
The basic edge $(1,4)$ then gives that $y_1 - y_4 = c_{14}$, i.e. $y_4 = y_1 - c_{14} = 4 - 2 = 2$.
The basic edge $(3,5)$ then gives that $y_3 - y_5 = c_{35}$, i.e. $y_5 = y_3 - c_{35} = 4 - 2 = 2$.

Next step is to compute the reduced costs for the non-basic variables, which gives
$r_{15} = c_{15} - y_1 + y_5 = 3 - 4 + 2 = 1$,
$r_{24} = c_{24} - y_2 + y_4 = 3 - 4 + 2 = 1$,
$r_{25} = c_{25} - y_2 + y_5 = 3 - 4 + 2 = 1$,
$r_{34} = c_{34} - y_3 + y_4 = 3 - 4 + 2 = 1$.

Since all $r_{ij} \geq 0$ the given basic solution is optimal.

Comment: The problem can also be solved as a transportation problem!

9.3 (20060603-nr.5)

First comes an example of a LP-formulation.

First we choose the following variables:

Let x_j denote the number of tonnes that are manufactured on normal working time in month j.

Let y_j denote the number of tonnes that are manufactured on overtime in month j.

Let z_j denote the number of tonnes that are delivered to the customer in the end of each month j.

Let s_j denote the number of tonnes that are stored in the storage during month j.

Let u_j denote the number of tonnes that you are "owing" the customer in the beginning of month j
(Not including p_j, p_{j+1} etc.).

Then the objective function, which should be minimized, can be written as
$$\sum_{j=1}^{3}(c\,x_j + d\,y_j + \ell\,s_j + f u_j).$$

The constraints that have to be fulfilled are the following:

$s_1 = 0, \;\; u_1 = 0,$

$x_1 + y_1 - z_1 - s_2 = 0,$

$x_2 + y_2 - z_2 + s_2 - s_3 = 0,$

$x_3 + y_3 - z_3 + s_3 = 0,$

$z_1 + u_2 = p_1,$

$z_2 + u_3 - u_2 = p_2,$

$z_3 - u_3 = p_3,$

$x_j \leq a$ and $y_j \leq b$ for $j = 1, 2, 3,$

and all variables must be ≥ 0.

If you want you can eliminate the variables z_j, and then the constraints have the following form:

$s_1 = 0, \;\; u_1 = 0,$

$x_1 + y_1 + u_2 - s_2 = p_1,$

$x_2 + y_2 + u_3 - u_2 + s_2 - s_3 = p_2,$

$x_3 + y_3 - u_3 + s_3 = p_3,$

$u_2 \leq p_1,$

$u_3 - u_2 \leq p_2,$

$x_j \leq a$ and $y_j \leq b$ for $j = 1, 2, 3,$

and all variables must be ≥ 0.

9.4 (20060308-nr.5)

(a).

We obtain the following feasible basic solution with help of the NWC-method:

x_{ij}	customer1	customer2	customer3	customer4	s_i
sup 1	20	40	20		80
sup 2			40	20	60
sup 3				40	40
sup 4				20	20
d_j	20	40	60	80	

(b). Corresponding to the feasible basic solution above we obtain the following simplex multipliers u_i and v_j, computed with help of the relation $c_{ij} = u_i - v_j$ for basic variables and $v_4 = 0$.

c_{ij}	customer1	customer2	customer3	customer4	u_i
sup 1	16	25	36		47
sup 2			25	36	36
sup 3				25	25
sup 4				16	16
v_j	31	22	11	0	

Then we obtain the following reduced costs r_{ij} for the non-basic variables, computed with the relation $r_{ij} = c_{ij} - u_i + v_j$.

r_{ij}	customer1	customer2	customer3	customer4	u_i
sup 1				2	47
sup 2	4	2			36
sup 3	10	6	2		25
sup 4	16	10	4		16
v_j	31	22	11	0	

All $r_{ij} \geq 0$, which implies that this basic solution is optimal.

(c). Try to use the same collection of basic variables as in the optimal solution above. That gives the new solution:

x_{ij}	customer1	customer2	customer3	4	s_i
sup 1	40	40	0		80
sup 2			60	0	60
sup 3				40	40
sup 4				40	40
d_j	40	40	60	80	

All x_{ij} were ≥ 0, so it is still a feasible basic solution (but degenerated since it has basic variables with the value 0). Hence the computation of u_i, v_j and r_{ij} will be identical with the one above, so all r_{ij} will still be ≥ 0. Hence the solution in the table is optimal to the new problem.

(d). x_{22} is a non-basic variable in the optimal solution, so if c_{22} is decreased with δ_{22}, then r_{22} is decreased with δ_{22} while the other r_{ij} not are affected. Hence it follows that the given solution is still optimal if and only if $\delta_{22} \leq 2$.

9.5 (20051024-nr.2)

The proposed solution in the exercise corresponds to a spanning tree in the network, i.e. to a basic solution of the problem. Further it is a feasible basic solution, since the balance equations are fulfilled in all nodes and no variables are negative.

You compute the simplex multipliers y_i from the constraints $y_i - y_j = c_{ij}$ for basic variables (i.e. edges of the tree) and $y_6 = 0$. This gives that $\mathbf{y} = (9, 6, 7, 2, 3, 0)^\mathsf{T}$.
After that the reduced costs for the non-basic variables are computed from $\bar{c}_{ij} = c_{ij} - y_i + y_j$.
This gives that $\bar{c}_{23} = 2$, $\bar{c}_{34} = -1$, $\bar{c}_{45} = 2$ and $\bar{c}_{56} = 0$.
Since $\bar{c}_{34} = -1$ we shall let x_{34} become new basic variable. The corresponding edge $(3, 4)$ forms a loop in the network together with the tree edges $(2, 4)$ (backwards), $(1, 2)$ (backwards) and $(1, 3)$ (forward). The flow in the edge $(3, 4)$, i.e. x_{34}, can increment to 10 before one of the backward edges, x_{12}, has reached 0. Hence x_{34} becomes new basic variable instead of x_{12}. The new feasible basic solution becomes:
$x_{13} = 25,\ x_{24} = 10,\ x_{34} = 10,\ x_{35} = 15,\ x_{46} = 20$, other $x_{ij} = 0$.

You compute the simplex multipliers y_i again from the constraints $y_i - y_j = c_{ij}$ for basic variables (i.e. edges of the trees), and $y_6 = 0$. This gives that $\mathbf{y} = (8, 6, 6, 2, 2, 0)^\mathsf{T}$.

After that the reduced costs of the non-basic variables are computed from $\bar{c}_{ij} = c_{ij} - y_i + y_j$.
It gives that $\bar{c}_{12} = 1$, $\bar{c}_{23} = 1$, $\bar{c}_{45} = 1$ and $\bar{c}_{56} = 1$.
Since all $\bar{c}_{ij} \geq 0$ the current feasible basic solution is optimal.
The optimal value is $\sum c_{ij} x_{ij} = 230$.

A LP-formulation of the problem above is

$$\begin{array}{rll} \text{minimize} & \displaystyle\sum_{(i,j)\in\mathcal{B}} c_{ij}x_{ij} & \\ \text{s.t.} & \displaystyle\sum_{j\in\mathcal{J}(i)} x_{ij} - \sum_{k\in\mathcal{K}(i)} x_{ki} \;=\; b_i\,, & \text{for } i \in \mathcal{N}, \\ & x_{ij} \;\geq\; 0\,, & \text{for } (i,j) \in \mathcal{B}, \end{array}$$

where $\mathcal{J}(i) = \{j \in \mathcal{N} \mid (i,j) \in \mathcal{B}\}$, $\mathcal{K}(i) = \{k \in \mathcal{N} \mid (k,i) \in \mathcal{B}\}$ and $(b_1, b_2, b_3, b_4, b_5, b_6) = (25, 10, 0, 0, -15, -20)$.

The corresponding dual LP-problem is then

$$\begin{array}{rll} \text{maximize} & \displaystyle\sum_{i\in\mathcal{N}} b_i y_i & \\ \text{s.t.} & y_i - y_j \leq c_{ij}, & \text{for } (i,j) \in \mathcal{B}. \end{array}$$

An optimal solution is given by the simplex multipliers corresponding to the optimal basic solution above, i.e. $\mathbf{y} = (8, 6, 6, 2, 2, 0)^\mathsf{T}$.
This is a feasible solution to the dual problem (since $y_i - y_j = c_{ij}$ for basic edges and $y_i - y_j = c_{ij} - 1$ for non-basic edges). Further $\sum b_i y_i = 25 \cdot 8 + 10 \cdot 6 - 15 \cdot 2 - 20 \cdot 0 = 230$, which corresponds to the optimal value of the primal problem above.

9.6 (20050331-nr.1)

The problem can be modeled as a minimum cost flow problem (MCFP) in a network with 7 nodes and 10 edges. The of the head of the transport division suggested solution is a feasible basic solution, since it corresponds to a spanning tree in the network. Applying the MCFP-algorithm (i.e. the simplex method applied on MCFP) gives that there is a negative reduced cost in the proposal, and hence you should change basis, etc.

It is even easier to model the problem as a transportation problem (TP) with two source nodes (factories) and three sink nodes (customers). The transportation cost from for example F2 to K2 is given by the least of the the two numbers 4+7=11 resp 5+9=14, where 4+7 is the transportation cost via T1 whereas 5+9 is the transportation cost via T2. This gives us transportation costs according to the left table beneath.

	$K1$	$K2$	$K3$
$F1$	12	14	11
$F2$	10	11	10

	$K1$	$K2$	$K3$
$F1$	100		200
$F2$	100	200	

The TP-algorithm gives an unique optimal solution in accordance with the right table above.

If you translate this optimal solution to the same format as the head of the transportations had, with help of the cheapest ways between the current pairs of plants and customers, you get the following tables:

	$T1$	$T2$
$F1$	0	300
$F2$	300	0

	$K1$	$K2$	$K3$
$T1$	100	200	0
$T2$	100	0	200

This solution is some 100 SEK-bills cheaper that the original proposal.

9.7 (20050307-nr.1)

(a) Only the answers are given here.

A maximum flow is given by (for example) $x_{12} = x_{26} = x_{68} = 1$, $x_{13} = x_{35} = x_{58} = 1$, the rest $x_{ij} = 0$. Its value is $= 2$. A minimum cut is given by the edges $(S, \bar{S})$, where $S = \{1, 2, 4, 6\}$ and $\bar{S} = \{3, 5, 7, 8\}$, i.e. $(S, \bar{S}) = \{(1, 3),\ (6, 8)\}$. Its capacity is $= 1 + 1 = 2$.

(b).

The problem can be modeled as a maximum flow problem in the network beneath, where you want to determine the flow from node A to node B.

The edges most to the left, from the node A to the student-nodes, all have the capacity 2 (since all students work 2 days a week).

The edges most to the right, from the 7 day-nodes to the node B, also have the capacity 2 (since every day 2 students are required).

The edges in the middle, from the student-nodes to the day-nodes, have capacity one, (since for example student one can work at most one Monday a week). Which middle edges that are in the network is given by the list of wishes.

If and only if the maximum flow from node A to node B has the value 14, there is a working schedule that fulfills the wishes of the students.

	S1	Mon	
	S2	Tue	
	S3	Wed	
A	S4	Thu	B
	S5	Fri	
	S6	Sat	
	S7	Sun	

(According to the exercise it is not necessary to solve the formulated maximum flow problem, but if you still want to do it, you easily find a feasible flow of value 13. There are several. But then there are no way from A to B in the increment network, so the maximum flow algorithm stops. Hence it is *not* possible to work out a schedule that fulfills all wishes.)

9.8 (20041016-nr.1)

(a) Let:
x_{ij} = number of tonnes of fuel that the air-company orders from supplier i to airport j,
s_i = the capacity at supplier i ($s_1 = s_2 = s_3 = 400$),
d_j = the demand at airport j ($d_1 = d_2 = d_3 = d_4 = 300$),
c_{ij} = cost per tonne from supplier i to airport j ($c_{11} = 5,\ c_{12} = 4$, etc.).

Since $s_1 + s_2 + s_3 = d_1 + d_2 + d_3 + d_4$ the problem can be formulated as:

$$
\begin{array}{lll}
\text{TP:} \quad & \text{minimize} & \displaystyle\sum_{i=1}^{3}\sum_{j=1}^{4} c_{ij}x_{ij} \\
& \text{s.t.} & \displaystyle\sum_{j=1}^{4} x_{ij} \ = \ s_i\,, \qquad \text{for } i = 1,\ldots,3 \\
& & \displaystyle-\sum_{i=1}^{3} x_{ij} \ = -d_j\,, \quad \text{for } j = 1,\ldots,4 \\
& & x_{ij} \ \geq \ 0\,, \qquad \text{for all } i \text{ and } j.
\end{array}
$$

(b) Starting basic solution, with the “Northwest”-method:

x_{ij}	A1	A2	A3	A4	s_i
S1	300	100			400
S2		200	200		400
S3			100	300	400
d_j	300	300	300	300	

Corresponding to this basic solution we obtain the following simplex multipliers u_i and v_j, computed with the relation $c_{ij} = u_i - v_j$ for basic variables and $v_4 = 0$, and the following reduced costs $\bar{c}_{ij}$ for non-basic variables, computed with the relation $\bar{c}_{ij} = c_{ij} - u_i + v_j$.

$\bar{c}_{ij}$	A1	A2	A3	A4	u_i
S1			0	-1	6
S2	2			1	6
S3	2	1			7
v_j	1	2	2	0	

Let x_{14} become new basic variable. That gives the following new basic solution:

x_{ij}	A1	A2	A3	A4	s_i
S1	300			100	400
S2		300	100		400
S3			200	200	400
d_j	300	300	300	300	

Corresponding to this basic solution we get the following simplex multipliers u_i and v_j, and the reduced costs $\bar{c}_{ij}$.

$\bar{c}_{ij}$	A1	A2	A3	A4	u_i
S1		1	1		5
S2	1			1	6
S3	1	1			7
v_j	0	2	2	0	

Now all $\bar{c}_{ij} \geq 0$, which imply that this basic solution is optimal. The optimal value becomes

$$\sum_{i=1}^{3}\sum_{j=1}^{4} c_{ij}x_{ij} = 5\cdot 300 + 5\cdot 100 + 4\cdot 300 + 4\cdot 100 + 5\cdot 200 + 7\cdot 200 = 6000 \text{ kSEK}.$$

(c) The dual problem in the variables u_i and v_j can be written:

$$\begin{aligned} \text{D:} \qquad & \text{maximize} \quad \sum_{i=1}^{3} s_i u_i - \sum_{j=1}^{4} d_j v_j \\ & \text{s.t.} \quad u_i - v_j \leq c_{ij}\ , \quad \text{for all } i \text{ and } j. \end{aligned}$$

An optimal solution to this dual problem is given by the simplex multipliers in the optimal basic solution to the primal problem above, i.e. $u_1 = 5,\ u_2 = 6,\ u_3 = 7,\ v_1 = 0,\ v_2 = 2,\ v_3 = 2,\ v_4 = 0.$ The optimal value becomes

$$\sum_{i=1}^{3} s_i u_i - \sum_{j=1}^{4} d_j v_j = 400\cdot 5 + 400\cdot 6 + 400\cdot 7 - 300\cdot 2 - 300\cdot 2 = 6000 \text{ kSEK}.$$

OK!

9.9 (20040415-nr.1)

The starting basic solution, with the “Northwest”-method, becomes:

x_{ij}	P	Q	R	S	T
A	50	40			
B		80	20		
C			60		
D			30	70	90

Corresponding to this basic solution we get the following simplex multipliers u_i and v_j, computed with help of the relation $c_{ij} = u_i - v_j$ for basic variables and $v_5 = 0$, and the following reduced costs $\bar{c}_{ij}$ for the non-basic variables, computed with help of the relation $\bar{c}_{ij} = c_{ij} - u_i + v_j$.

$\bar{c}_{ij}$	P	Q	R	S	T	u_i
A			0	0	−1	4
B	1			0	0	6
C	2	2		1	1	4
D	1	2				5
v_j	−2	−2	−1	−2	0	

Let x_{15} become the new basic variable. This gives the new basic solution:

x_{ij}	P	Q	R	S	T
A	50	20			20
B		100			
C			60		
D			50	70	70

Corresponding to this basic solution we get the following simplex multipliers u_i and v_j, and the following reduced costs $\bar{c}_{ij}$ for the non-basic variables:

$\bar{c}_{ij}$	P	Q	R	S	T	u_i
A			1	1		3
B	1		1	1	1	5
C	1	1		1	1	4
D	0	1				5
v_j	-3	-3	-1	-2	0	

All $\bar{c}_{ij} \geq 0$, which implies that this basic solution is optimal.
Since $\bar{c}_{41} = 0$, x_{41} can become new basic variable without that the value of the objective function is affected.
This gives the following new basic solution:

x_{ij}	P	Q	R	S	T
A		20			70
B		100			
C			60		
D	50		50	70	20

This basic solution is also optimal.

9.10 (20040310-nr.1)

Our network have 6 nodes. Let node 1 and node 2 be the source-nodes, node 3 and node 4 the intermediate nodes, and node 5 and node 6 the sink-nodes. The set of links is then $\mathcal{B} = \{(1,3),\ (1,4),\ (2,3),\ (2,4),\ (3,5),\ (3,6),\ (4,5),\ (4,6)\}$.

The minimum cost flow problem corresponding to the given network can be written on the following form:

$$\text{MCF}: \quad \text{minimize} \ \mathbf{c}^\mathsf{T}\mathbf{v} \quad \text{s.t.} \ \mathbf{A}\mathbf{v} = \mathbf{b}, \quad \mathbf{v} \geq \mathbf{0},$$

where $\mathbf{v} = (v_{13},\ v_{14},\ v_{23},\ v_{24},\ v_{35},\ v_{36},\ v_{45},\ v_{46})^\mathsf{T} = (x_{11},\ x_{12},\ x_{21},\ x_{22},\ z_{11},\ z_{12},\ z_{21},\ z_{22})^\mathsf{T}$,

$$\mathbf{c} = (5,\ 2,\ 3,\ 2,\ 5,\ 5,\ 7,\ 6)^\mathsf{T},$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 30 \\ 20 \\ 0 \\ 0 \\ -40 \end{pmatrix},$$

where we have ignored the redundant balance equation for the last node.

The proposed solution corresponds to a spanning tree (and is hence a basic solution) consisting of the links $\mathcal{B}_\beta = \{(1,4),\ (2,3),\ (3,5),\ (4,5),\ (4,6)\}$.

The values of the basic variables can be determined in for example the following order:
$v_{14} = x_{12} = 30$, because of the flow balance condition in node 1.
$v_{23} = x_{21} = 20$, because of the flow balance condition in node 2.
$v_{35} = z_{11} = 20$, because of the flow balance condition in node 3.
$v_{45} = z_{21} = 20$, because of the flow balance condition in node 5.
$v_{46} = z_{22} = 10$, because of the flow balance condition in node 4.
This agrees perfectly with the proposed solution.

It remains to check if this feasible solution is optimal.

The vector $\mathbf{y}$ is computed from the equations $y_i - y_j = c_{ij}$ for all $(i,j) \in \mathcal{B}_\beta$, where $y_6 = 0$.
The basic link $(4,6)$ gives that $y_4 - y_6 = c_{46}$, i.e. $y_4 = c_{46} = 6$.
The basic link $(4,5)$ then gives that $y_4 - y_5 = c_{45}$, i.e. $y_5 = y_4 - c_{45} = 6-7 = -1$.
The basic link $(3,5)$ then gives that $y_3 - y_5 = c_{35}$, i.e. $y_3 = y_5 + c_{35} = -1+5 = 4$.
The basic link $(2,3)$ then gives that $y_2 - y_3 = c_{23}$, i.e. $y_2 = y_3 + c_{23} = 4+5 = 7$.
The basic link $(1,4)$ then gives that $y_1 - y_4 = c_{14}$, i.e. $y_1 = y_4 + c_{14} = 6+2 = 8$.

The next step is to compute the reduced costs from the formula
$\bar{c}_{ij} = c_{ij} - y_i + y_j$ for all $(i,j) \in \mathcal{B}_\delta$ (i.e. for all non-basic link). We get that $\bar{c}_{13} = c_{13} - y_1 + y_3 = 5 - 8 + 4 = 1$, $\bar{c}_{24} = c_{24} - y_2 + y_4 = 2 - 7 + 6 = 1$ and $\bar{c}_{36} = c_{36} - y_3 + y_6 = 5 - 4 + 0 = 1$.

Since all $\bar{c}_{ij} \geq 0$ the proposed optimal solution is optimal.
The optimal value is $\sum c_{ij} v_{ij} = 2 \cdot 30 + 3 \cdot 20 + 5 \cdot 20 + 7 \cdot 20 + 6 \cdot 10 = 420$.

(b) The corresponding dual LP-problem is

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{6} b_i y_i \\ \text{s.t.} \quad & y_i - y_j \leq c_{ij}, \quad \text{for all } (i,j) \in \mathcal{B}. \end{aligned}$$

An optimal solution is given by the scalars y_i computed above, i.e. $\mathbf{y} = (8, 7, 4, 6, -1, 0)^T$.
This is a feasible solution to the dual problem (since $y_i - y_j = c_{ij}$ for basic links and $y_i - y_j < c_{ij}$ for non-basic links). Further $\sum b_i y_i = 30 \cdot 8 + 20 \cdot 7 + 40 \cdot 1 = 420$, which is the same as the optimal value of the primal problem above.

10. Convexity

10.1 **(c)** Take two points a and b in $C+D$. Show that $\lambda a + (1-\lambda)b$ belongs to $C+D$ for all $\lambda \in [0,1]$.

$$\begin{array}{lcll} a \in C+D & \Longrightarrow & a = x_a + y_a & \text{for some } x_a \in C \text{ and } y_a \in D \\ b \in C+D & \Longrightarrow & b = x_b + y_b & \text{for some } x_b \in C \text{ and } y_b \in D \end{array}$$

Using the above we get

$$\begin{aligned} \lambda a + (1-\lambda)b &= \lambda(x_a + y_a) + (1-\lambda)(x_b + y_b) = \\ &= \lambda x_a + \lambda y_a + (1-\lambda)x_b + (1-\lambda)y_b = \\ &= \lambda x_a + (1-\lambda)x_b + \lambda y_a + (1-\lambda)y_b \in C+D \end{aligned}$$

since C and D are convex. This shows that $C+D$ is convex.

10.2 Take two points a and b in $\bigcap_{\alpha\in A} C_\alpha$. Show that $\lambda a + (1-\lambda)b$ belongs to $\bigcap_{\alpha\in A} C_\alpha$ for all $\lambda \in [0,1]$. This is true for if a and b are in $\bigcap_{\alpha\in A} C_\alpha$, then for each $\alpha \in A$ we have $a \in C_\alpha$ and $b \in C_\alpha$. Since C_α is a convex set, $\lambda a+(1-\lambda)b$ belongs to C_α. Since this is true for every $\alpha \in C_\alpha$, we have $\lambda a + (1-\lambda)b$ in $\bigcap_{\alpha\in A} C_\alpha$. This shows that $\bigcap_{\alpha\in A} C_\alpha$ is a convex set.

10.3 **(a)** Let x and y be two points in C. Show that $(f+g)(\lambda y + (1-\lambda)x) \leq \lambda(f+g)(y)(1-\lambda)(f+g)(x)$ for all $\lambda \in [0,1]$.

$$\begin{array}{lcl} (f+g)(\lambda y + (1-\lambda)x) & = & f(\lambda y + (1-\lambda)x) + g(\lambda y + (1-\lambda)x) \leq \\ \{f \text{ and } g \text{ are convex}\} & \leq & \lambda(f(y)+g(y)) + (1-\lambda)(f(x)+g(x)) = \\ & = & \lambda(f+g)(y) + (1-\lambda)(f+g)(x) \end{array}$$

This shows that $f+g$ is convex.

10.4 Let x and y be two points in C. Show that $\sup\limits_{\alpha\in A} f_\alpha(\lambda y + (1-\lambda)x) \leq \lambda \sup\limits_{\alpha\in A} f_\alpha(y) + (1-\lambda)\sup\limits_{\alpha\in A} f_\alpha(x)$ for all $\lambda \in [0,1]$.

$$\begin{array}{lcl} \lambda \sup\limits_{\alpha\in A} f_\alpha(y) + (1-\lambda)\sup\limits_{\alpha\in A} f_\alpha(x) & \geq & \lambda f_\beta(y) + (1-\lambda)f_\beta(x) \geq \\ \{f_\beta,\ \beta \in A \text{ is convex }\} & \geq & f_\beta(\lambda y + (1-\lambda)x) \end{array}$$

Since the inequality above holds for all $\beta \in A$ it must hold that

$$\sup_{\beta\in A} f_\beta(\lambda y + (1-\lambda)x) \quad \leq \quad \lambda \sup_{\alpha\in A} f_\alpha(y) + (1-\lambda)\sup_{\alpha\in A} f_\alpha(x)$$

This shows that $\sup\limits_{\alpha\in A} f_\alpha$ is convex.

10.5 Let x and y be two points in C, and let $\lambda \in [0,1]$. Since g is convex on C, it follows that

$$g((1-\lambda)x + \lambda y) \leq (1-\lambda)g(x) + \lambda g(y).$$

Hence, since f is a nondecreasing function on I, we have

$$f(g((1-\lambda)x + \lambda y)) \leq f((1-\lambda)g(x) + \lambda g(y)).$$

Finally, since f is a convex function on I, we have

$$f((1-\lambda)g(x) + \lambda g(y)) \leq (1-\lambda)f(g(x)) + \lambda f(g(y)).$$

The required result now follows by combining the last two inequalities.

10.6 **(a)** Convex.

(b) Convex.

(c) Convex.

(d) $f(x) = x_1^2/x_2$, $x_2 > 0$ is convex if and only if the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite for all $x_2 > 0$, i.e., if all the eigenvalues are non-negative. The eigenvalues can be determined by solving the equation $\det(\nabla^2 f(x) - I\lambda) = 0$.

$$\nabla^2 f(x) \quad = \quad \begin{pmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2} \\ -\frac{2x_1}{x_2} & \frac{2x_1^2}{x_2^3} \end{pmatrix}$$

The equation becomes

$$\det(\nabla^2 f(x) - I\lambda) = (\frac{2}{x_2} - \lambda)(\frac{2x_1^2}{x_2^3} - \lambda) - \frac{4x_1^2}{x_2^4} = \lambda(\lambda - 2\frac{x_1^2 + x_2^2}{x_2^3}) = 0$$

This yields $\lambda_1 = 0$ and $\lambda_2 = 2\frac{x_1^2 + x_2^2}{x_2^3} > 0$ for $x_2 > 0$. This shows that f is convex for $x_2 > 0$. (For 2×2-matrices there are easier ways of checking if they are positive semidefinite, for example you can use Sylvester's criterion.)

(e) Convex.

(f) Convex.

10.7 Since ln is an increasing function which is well-defined for positive arguments, it holds that

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} x_i \quad \Longleftrightarrow \quad \frac{1}{n}\sum_{i=1}^{n} \ln x_i \leq \ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right),$$

for $x_i > 0$, $i = 1, \dots, n$.

The proof is by induction. Consider the inequality

$$\frac{1}{n}\sum_{i=1}^{n} \ln x_i \leq \ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right).$$

for $x_i > 0$, $i = 1, \dots, n$. $\ln x$ is a concave function for $x > 0$, and hence the inequality holds for $n = 2$. Now, suppose the inequality holds for $n = k$. We want to show that it holds also for $n = k + 1$.

For some $x_i > 0$, $i = 1, \dots, n$, consider the identity

$$\frac{1}{k+1}\sum_{i=1}^{k+1} \ln x_i = \frac{k}{k+1}\left(\frac{1}{k}\sum_{i=1}^{k} \ln x_i\right) + \frac{1}{k+1}\ln x_{k+1}.$$

Since the inequality in question is assumed to be valid for $k = n$, it follows that

$$\frac{1}{k+1}\sum_{i=1}^{k+1} \ln x_i \le \frac{k}{k+1}\ln\left(\frac{1}{k}\sum_{i=1}^{k} x_i\right) + \frac{1}{k+1}\ln x_{k+1}.$$

The concavity of ln now gives

$$\frac{1}{k+1}\sum_{i=1}^{k+1} \ln x_i \le \ln\left(\frac{k}{k+1}\frac{1}{k}\sum_{i=1}^{k} x_i + \frac{1}{k+1}x_{k+1}\right) = \ln\left(\frac{1}{k+1}\sum_{i=1}^{k+1} x_i\right),$$

as required.

10.8 **(a)** Let x and y be arbitrary points in C and let $\lambda \in [0, 1]$. Since $x \in C$, there exist $t_i \ge 0$, $i = 1, \dots, m$ such that

$$x = \sum_{i=1}^{m} t_i x_i \quad \text{and} \quad \sum_{i=1}^{m} t_i = 1.$$

Similarly, since $y \in C$, there exist $s_i \ge 0$, $i = 1, \dots, m$ such that

$$y = \sum_{i=1}^{m} s_i x_i \quad \text{and} \quad \sum_{i=1}^{m} s_i = 1.$$

But then,

$$(1-\lambda)x + \lambda y = (1-\lambda)\sum_{i=1}^{m} t_i x_i + \lambda\sum_{i=1}^{m} s_i x_i = \sum_{i=1}^{m}((1-\lambda)t_i + \lambda s_i)x_i.$$

Hence, if we define $r_i = (1-\lambda)t_i + \lambda s_i$, $i = 1, \dots, m$, we have

$$(1-\lambda)x + \lambda y = \sum_{i=1}^{m} r_i x_i,$$

and the properties of s_i and t_i give $r_i \ge 0$, $i = 1, \dots, m$ and $\sum_{i=1}^{m} r_i = 1$. Consequently, $(1-\lambda)x + \lambda y \in C$, as required.

(b) The proof is by induction.

For $m = 2$, the statement is that if $x_1 \in X$ and $x_2 \in X$, $t_1 \ge 0$, $t_2 \ge 0$ and $t_1 + t_2 = 1$, then $t_1 x_1 + t_2 x_2 \in X$. This is true from the convexity of X.

Suppose that the statement is true for $m = k$, i.e., if $x_1, \dots, x_k \in X$, $t_i \ge 0$, $i = 1, \dots, k$ and $\sum_{i=1}^{k} t_i = 1$, then $\sum_{i=1}^{k} t_i x_i \in X$. We want to show that the statement is true also for $m = k + 1$.

Let $x_1, \ldots, x_{k+1} \in X$, $t_i \geq 0$, $i = 1, \ldots, k+1$ and $\sum_{i=1}^{k+1} t_i = 1$. We want to show that $\sum_{i=1}^{k+1} t_i x_i \in X$. If $t_{k+1} = 1$, then we must have $t_i = 0$, $i = 1, \ldots, k$ and the statement is true since $\sum_{i=1}^{k+1} t_i x_i = x_{k+1} \in X$. Now consider the case when $t_{k+1} < 1$. Then,

$$\sum_{i=1}^{k+1} t_i x_i = \sum_{i=1}^{k} t_i x_i + t_{k+1} x_{k+1} = (1 - t_{k+1}) \sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} x_i + t_{k+1} x_{k+1}.$$

Since $t_i \geq 0$, $i = 1, \ldots, k+1$ and $\sum_{i=1}^{k+1} t_i = 1$, it follows that

$$\frac{t_i}{1 - t_{k+1}} \geq 0,\ i = 1, \ldots k \quad \text{and} \quad \sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} = 1.$$

Hence, since it is assumed that the statement is true for $m = k$, it holds that

$$\sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} x_i \in X.$$

But then, since $t_{k+1} \in [0, 1]$, the convexity of X ensures that

$$\sum_{i=1}^{k+1} t_i x_i = (1 - t_{k+1}) \sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} x_i + t_{k+1} x_{k+1} \in X,$$

and the induction proof is complete.

10.9 (20060308-nr.5)

The chain rule gives that $g'(x) = 2f(x)f'(x)$ and $g''(x) = 2f(x)f''(x) + 2(f'(x))^2$.

(a). That f is twice continuously differentiable implies that according to above also g is twice continuously differentiable. Hence we obtain: f convex on $I\!R$ $\Rightarrow f''(x) \geq 0$ for all $x \in I\!R \Rightarrow$
$\Rightarrow g''(x) \geq 0$ for all $x \in I\!R$ (since $f(x) > 0$ and $(f'(x))^2 \geq 0$) $\Rightarrow g$ convex on $I\!R$.

(b). Take for example $f(x) = (x^2 + 1)^{2/5}$, and then $g(x) = (x^2 + 1)^{4/5}$.
Straightforward computations show that $g''(x) \geq 0$ for all $x \in I\!R$, while $f''(x) < 0$ for $x \in I\!R$ big enough. g is hence convex on $I\!R$ although f is not!

(c). If $\hat{x}$ is a local minimizer to $f(x)$ then there is a number $\delta > 0$ such that $f(x) - f(\hat{x}) \geq 0$ for all $x \in I\!R$ such that $|x - \hat{x}| < \delta$. But for all those x it also holds that $h(x) - h(\hat{x}) = f(x)^2 - f(\hat{x})^2 = (f(x) - f(\hat{x}))(f(x) + f(\hat{x})) \geq 0$. Hence $\hat{x}$ is a local minimizer also to $h(x)$.

(d). If $\hat{x}$ is a local minimizer to $h(x)$ then there is a number $\delta > 0$ such that $h(x) - h(\hat{x}) \geq 0$ for all $x \in I\!R$ such that $|x - \hat{x}| < \delta$. But for all these x it also holds that $f(x) - f(\hat{x}) = (f(x)^2 - f(\hat{x})^2)/(f(x) + f(\hat{x})) = (h(x) - h(\hat{x}))/(f(x) + f(\hat{x})) \geq 0$.
Hence $\hat{x}$ is a local minpoint also to $f(x)$.

(e). We have that $x_1 = x_0 - \dfrac{f'(x_0)}{f''(x_0)}$ and $\bar{x}_1 = x_0 - \dfrac{g'(x_0)}{g''(x_0)} = x_0 - \dfrac{f(x_0)f'(x_0)}{f(x_0)f''(x_0) + (f'(x_0))^2}$,

such that $\dfrac{\bar{x}_1 - x_0}{x_1 - x_0} = \dfrac{f(x_0)f''(x_0)}{f(x_0)f''(x_0) + (f'(x_0))^2}$ which is both > 0 and < 1 (since $f(x_0)f''(x_0) > 0$ and $(f'(x_0))^2 > 0$).

Hence $\bar{x}_1 - x_0$ and $x_1 - x_0$ has the same sign, and $|\bar{x}_1 - x_0| < |x_1 - x_0|$.

10.10 (20051024-nr.5)

(a). Let $\tilde{\mathbf{x}} = \sum_k \hat{w}_k \mathbf{x}^{(k)}$.
Since LP1 only has two constraints (except the non-negativity demands), and $\hat{\mathbf{w}}$ is an optimal *basic solution*, it follows that at most two components in $\hat{\mathbf{w}}$ (the basic variables) are > 0, say $\hat{w}_p$ and $\hat{w}_q$, which then fulfill that $\hat{w}_p + \hat{w}_q = 1$. We hence have that $\tilde{\mathbf{x}} = \hat{w}_p \mathbf{x}^{(p)} + \hat{w}_q \mathbf{x}^{(q)}$.

Since f and g are convex it holds that

$f(\tilde{\mathbf{x}}) = f(\hat{w}_p \mathbf{x}^{(p)} + \hat{w}_q \mathbf{x}^{(q)}) \leq \hat{w}_p f(\mathbf{x}^{(p)}) + \hat{w}_q f(\mathbf{x}^{(q)}) = \sum_k \hat{w}_k f(\mathbf{x}^{(k)})$ and

$g(\tilde{\mathbf{x}}) = g(\hat{w}_p \mathbf{x}^{(p)} + \hat{w}_q \mathbf{x}^{(q)}) \leq \hat{w}_p g(\mathbf{x}^{(p)}) + \hat{w}_q g(\mathbf{x}^{(q)}) = \sum_k \hat{w}_k g(\mathbf{x}^{(k)}) \leq 0$,

where the last inequality follows from that $\hat{\mathbf{w}}$ is optimal, and hence feasible to LP1.

Hence the point $\tilde{\mathbf{x}}$ is a feasible solution to P0, with the objective function value $f(\tilde{\mathbf{x}}) \leq \sum_k \hat{w}_k f(\mathbf{x}^{(k)})$, and hence it follows that the optimal value $f(\hat{\mathbf{x}})$ to P0 must be $\leq \sum_k \hat{w}_k f(\mathbf{x}^{(k)})$.

(An option is to use the *Jensen inequality* which directly gives that $f(\sum_k \hat{w}_k \mathbf{x}^{(k)}) \leq \sum_k \hat{w}_k f(\mathbf{x}^{(k)})$, without need for using that only two components in the vector $\hat{\mathbf{w}}$ are > 0.)

(b)
Assume that $\hat{\mathbf{x}} = \mathbf{x}^{(k)}$ for a given k, say for simplicity that $\hat{\mathbf{x}} = \mathbf{x}^{(1)}$ (where $\hat{\mathbf{x}}$ as before is an optimal solution to P0).

Let $\hat{\mathbf{w}} = (1, 0, \ldots, 0)^\mathsf{T}$. Then:

$\sum_k \hat{w}_k g(\mathbf{x}^{(k)}) = \hat{w}_1 g(\mathbf{x}^{(1)}) = g(\mathbf{x}^{(1)}) = g(\hat{\mathbf{x}}) \leq 0$ (Since $\hat{\mathbf{x}}$ is optimal and hence feasible to P0).

$\sum_k \hat{w}_k f(\mathbf{x}^{(k)}) = \hat{w}_1 f(\mathbf{x}^{(1)}) = f(\mathbf{x}^{(1)}) = f(\hat{\mathbf{x}})$.

This shows that $\hat{\mathbf{w}}$ is a feasible solution to LP1 (since $\sum_k \hat{w}_k g(\mathbf{x}^{(k)}) \leq 0$) with the objective function value $\sum_k \hat{w}_k f(\mathbf{x}^{(k)}) = f(\hat{\mathbf{x}})$.

But each feasible solution $\mathbf{w}$ to LP1 fulfills according to the a)-task that $\sum_k w_k f(\mathbf{x}^{(k)}) \geq f(\hat{\mathbf{x}})$, i.e. that $\sum_k w_k f(\mathbf{x}^{(k)}) \geq \sum_k \hat{w}_k f(\mathbf{x}^{(k)})$.

This shows that $\hat{\mathbf{w}} = (1, 0, \ldots, 0)^\mathsf{T}$ is an optimal solution to LP1. Furthermore the optimal values of P0 and LP1 are now equal (Since $\sum_k \hat{w}_k f(\mathbf{x}^{(k)}) = f(\hat{\mathbf{x}})$).

11. Lagrange relaxations and duality

11.1 We first note that if $b \le 0$, then the optimal solution is $\hat{x} = 0$ and that if $a_j \le 0$, then it is optimal to set $\hat{x}_j = 0$. This index j can then be removed from the problem. We can hence in the following assume that $a_j > 0$, $j = 1, \dots, n$ and that $b > 0$.

We relax the constraint with a multiplier $\lambda \ge 0$ and solve the problem

$$\begin{array}{ll} \min & \sum_{j=1}^n x_j^2 + \lambda(b - \sum_{j=1}^n a_j x_j) \\ & x_j \ge 0, \;\; j = 1, \dots, n \end{array}$$

This problem has the solution $x_j(\lambda) = \lambda a_j/2$. We thereafter adjust λ so that $\sum_{j=1}^n a_j x_j(\lambda) = b$, i.e. $\sum_{j=1}^n \lambda a_j^2/2 = b$. I.e., we choose $\lambda = 2b/\sum_{j=1}^n a_j^2$. The optimal solution is hence given by $\hat{x}_j = b a_j/\sum_{j=1}^n a_j^2$, with the objective function value $b^2/\sum_{j=1}^n a_j^2$.

11.2 Switch to minimization and Lagrangean relax the first constraint to obtain the problem, $\lambda \ge 0$:

$$(P_L) \quad \begin{array}{ll} \min & \sum_{j=2}^n -\ln x_j + \lambda(\sum_{j=1}^n a_j x_j - b) = \\ & -\lambda b + \min \sum_{j=1}^n (-\ln x_j + \lambda a_j x_j) \\ \text{s.t.} & x > 0 \end{array}$$

This problem can be separated into n independent problems, one for each x_j. Let $f_j(x_j) = -\ln x_j + \lambda a_j x_j$. The problems can then be written

$$(P_{Lj}) \quad \begin{array}{ll} \min & f_j(x_j) \\ \text{s.t.} & x_j > 0 \end{array}$$

Since the objective function is convex f_j assumes its global minimum when $f_j'(x_j) = -\frac{1}{x_j} + \lambda a_j = 0$. This yields that $\hat{x}_j(\lambda) = \frac{1}{\lambda a_j}$. If the objective function of the Lagrangean relaxed problem is to yield the same value as the objective function of the unrelaxed problem it must hold that $\sum_{j=1}^n a_j \hat{x}_j(\lambda) = b$ since $\lambda \ne 0$, i.e. $b = \sum_{j=1}^n \frac{a_j}{\lambda a_j} = \frac{1}{\lambda}\sum_{j=1}^n 1 = \frac{n}{\lambda}$. This yields that $\lambda = \frac{n}{b}$ and $\hat{x}_j(\frac{n}{b}) = \hat{x}_j = \frac{b}{n a_j} > 0$. This $\hat{x}$ is also a feasible solution to the unrelaxed problem. Thus $\hat{x}_j = \frac{b}{n a_j}$, $j = 1, \dots, n$ is the optimal solution to the unrelaxed problem.

11.3 $\hat{x}_j = \dfrac{\sum_{k=1}^n \sqrt{a_k b_k}}{b_0} \sqrt{\dfrac{b_j}{a_j}}$, $j = 1, \dots, n$.

11.4 $\hat{x}_j = \dfrac{b_0}{\sum_{k=1}^n a_k b_k} \sqrt{\dfrac{a_j}{b_j}}, \quad j = 1, \dots, n.$

11.5 $$\hat{x}_j = \frac{\left(b - \displaystyle\sum_{k=1}^n \frac{a_k}{c_k} \ln \frac{a_k}{c_k}\right)}{c_j \displaystyle\sum_{i=1}^n \frac{a_i}{c_i}} + \frac{1}{c_j} \ln \frac{a_j}{c_j}, \quad j = 1, \dots, n.$$

11.6 **(a)** The problem can be formulated as

$$\begin{array}{lll} \max & \sum_{n=1}^{N} p_n \ln q_n & \\ \text{s.t.} & \sum_{n=1}^{N} q_n = 1, & \\ & q_n \geq 0, & n = 1, \ldots, N. \end{array}$$

(b) $\hat{q}_n = p_n, \quad n = 1, \ldots, N.$

11.7 Let

$$\begin{array}{rcl} f(x) & = & x_1^4 + 2x_1x_2 + x_2^2 + x_3^8, \\ g_1(x) & = & (x_1-2)^2 + (x_2-2)^2 + (x_3-3)^2 - 6, \\ g_2(x) & = & x_1x_2x_3 - 10, \\ g_3(x) & = & 1 - x_1, \\ g_4(x) & = & -x_2, \\ g_5(x) & = & -x_3, \end{array}$$

so that the problem is on standard form

$$(P) \quad \begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0, & i = 1, \ldots, 5, \\ & x \in I\!R^3. & \end{array}$$

Since $g_1(\hat{x}) = 0$, $g_2(\hat{x}) = -9$, $g_3(\hat{x}) = 0$, $g_4(\hat{x}) = -1$ and $g_5(\hat{x}) = -1$, it follows that $\hat{x}$ is feasible to (P). For $\hat{x}$ to be globally optimal to (P), it must in particular be locally optimal. Hence, we try to satisfy the KT conditions at $\hat{x}$, i.e., find $\hat{\lambda} \in I\!R^5$ such that $\hat{\lambda}_1 \geq 0$, $\hat{\lambda}_2 = 0$, $\hat{\lambda}_3 \geq 0$, $\hat{\lambda}_4 = 0$, $\hat{\lambda}_5 = 0$ and $\nabla f(\hat{x}) + \hat{\lambda}_1 \nabla g_1(\hat{x}) + \hat{\lambda}_3 \nabla g_3(\hat{x}) = 0$. Differentiation gives

$$\nabla f(x)^T = \begin{pmatrix} 4x_1^3 + 2x_2 \\ 2x_1 + 2x_2 \\ 8x_3^7 \end{pmatrix}, \quad \nabla g_1(x)^T = \begin{pmatrix} 2(x_1-2) \\ 2(x_2-2) \\ 2q(x_3-3) \end{pmatrix}, \quad \nabla g_3(x)^T = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

so $\hat{\lambda}_1$ och $\hat{\lambda}_3$ must satifsy

$$\begin{pmatrix} 6 \\ 4 \\ 8 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ -4 \end{pmatrix} \hat{\lambda}_1 + \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \hat{\lambda}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This system of equations has a unique solution $\hat{\lambda}_1 = 2$, $\hat{\lambda}_3 = 2$. With $\hat{\lambda}$ as above, i.e., $\hat{\lambda} = (2\ 0\ 2\ 0\ 0)^T$, let $f_{\hat{\lambda}}(x) = f(x) + \sum_{i=1}^{5} \hat{\lambda}_i g_i(x)$ and consider the Lagrangean-relaxed problem

$$(P_{\hat{\lambda}}) \quad \begin{array}{ll} \min & f_{\hat{\lambda}}(x) \\ \text{s.t.} & x \in I\!R^3. \end{array}$$

We have

$$\begin{aligned}
f_{\hat{\lambda}}(x) &= x_1^4 + 2x_1x_2 + x_2^2 + x_3^8 + 2((x_1-2)^2 + (x_2-2)^2 + (x_3-3)^2 - 6) + 2(1-x_1), \\
\nabla f_{\hat{\lambda}}(x)^T &= \begin{pmatrix} 4x_1^3 + 2x_2 + 4(x_1-2) - 2 \\ 2x_1 + 2x_2 + 4(x_2-2) \\ 8x_3^7 + 4(x_3-3) \end{pmatrix}, \\
\nabla^2 f_{\hat{\lambda}}(x) &= \begin{pmatrix} 12x_1^2 + 4 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 56x_3^6 + 4 \end{pmatrix}.
\end{aligned}$$

The matrix $\nabla^2 f_{\hat{\lambda}}(x)$ is positive definite for all x, and hence $(P_{\hat{\lambda}})$ is a convex problem. Since $\hat{\lambda}$ is chosen so that $\nabla f_{\hat{\lambda}}(\hat{x}) = 0$, it follows that $\hat{x}$ is globally optimal to $(P_{\hat{\lambda}})$. In addition, it holds that

(i) $g_i(\hat{x}) \leq 0, \;\; i = 1, \ldots, 5.$

(ii) $\hat{\lambda}_i g_i(\hat{x}) = 0, \;\; i = 1, \ldots, 5.$

(iii) $\hat{\lambda}_i \geq 0, \;\; i = 1, \ldots, 5.$

The main theorem of Lagrangean relaxation now guarantees that $\hat{x}$ is globally optimal to (P).

11.8 The dual problem is given by

$$(D) \quad \begin{array}{ll} \max & -\lambda^2 \displaystyle\sum_{i=1}^{n} \frac{a_i^2}{4} - \lambda b \\ \text{s.t.} & \lambda \in \mathbb{R}. \end{array}$$

11.9 For each fixed value of λ minimize the Lagrangean with respect to x, i.e. solve the problem

$$\begin{array}{ll} \min\limits_{x} & l(x, \lambda) = \sum_{i=1}^{n} (\frac{a_i}{x_i} + \lambda b_i x_i) - \lambda b_0 \\ \text{s.t.} & l_i \leq x_i \leq u_i \end{array}$$

This is a separable problem so the minimization can be carried out for each x_i separately. Let $f_i(x_i) = \frac{a_i}{x_i} + \lambda b_i x_i$. Then $f_i(x_i)$ is a convex function and a necessary and sufficient condition for optimality is that $f_i'(x_i) = -a_i \frac{1}{{x_i}^2} + \lambda b_i = 0$. This yields that ${x_i}^2 = \frac{a_i}{\lambda b_i}$.

Now considering the constraint $l_i \leq x_i \leq u_i$ the optimal, feasible choice of x_i is given by

$$\begin{array}{llll}
1) & \lambda b_i \leq 0 & \Longrightarrow & \hat{x}_i = u_i \\
2) & \lambda b_i > 0 \;\text{ and }\; l_i \leq \sqrt{\frac{a_i}{\lambda b_i}} \leq u_i & \Longrightarrow & \hat{x}_i = \sqrt{\frac{a_i}{\lambda b_i}} \\
3) & \lambda b_i > 0 \;\text{ and }\; \sqrt{\frac{a_i}{\lambda b_i}} \leq l_i & \Longrightarrow & \hat{x}_i = l_i \\
4) & \lambda b_i > 0 \;\text{ and }\; \sqrt{\frac{a_i}{\lambda b_i}} \geq u_i & \Longrightarrow & \hat{x}_i = u_i
\end{array}$$

With $\hat{x}_1, \hat{x}_2, \hat{x}_3, \ldots, \hat{x}_n$ determined according to the above the dual objective function becomes $\phi(\lambda) = \sum_{i=0}^{n} (\frac{a_i}{x_i} + \lambda b_i \hat{x}_i) - \lambda b_0$ for the chosen λ. In the same way $\phi(\lambda)$ can be determined for every λ. The dual problem is

$$(D) \qquad \max_{\lambda} \phi(\lambda)$$

11.10 Lagrangean relax the first two constraints and determine the dual objective function

$$\begin{aligned}
\phi(\lambda) &= \min_{x\geq 0}\left\{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}\ln(x_{ij}) + \sum_{j=1}^{n}\lambda_j\left(b_j - \sum_{i=1}^{m} x_{ij}\right) + \sum_{i=1}^{m}\mu_i\left(a_i - \sum_{j=1}^{n} x_{ij}\right)\right\} = \\
&= \sum_{j=1}^{n}\lambda_j b_j + \sum_{i=1}^{m}\mu_i a_i + \min_{x\geq 0}\left\{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}(\ln(x_{ij}) - \lambda_j - \mu_i)\right\} = \\
&= \sum_{j=1}^{n}\lambda_j b_j + \sum_{i=1}^{m}\mu_i a_i + \min_{x\geq 0}\quad h(x)
\end{aligned}$$

$h(x)$ is a convex function. Therefore if $\nabla h(\hat{x}) = 0$ and $\hat{x} \geq 0$ then $\hat{x}$ is a feasible optimum.

$$\begin{aligned}
\frac{dh(x)}{dx_{ij}} &= \quad \ln(x_{ij}) - \lambda_j - \mu_i + 1 \quad = \quad 0 \\
&\Longrightarrow \quad \hat{x}_{ij} = \exp(\lambda_j - \mu_i - 1) \quad \geq \quad 0
\end{aligned}$$

The dual problem is given by

$$(D) \qquad \begin{array}{ll} \max & \displaystyle\sum_{j=1}^{n} b_j\lambda_j + \sum_{i=1}^{m} a_i\mu_i - \sum_{i=1}^{m}\sum_{j=1}^{n} e^{\lambda_j+\mu_i-1} \\ \text{s.t.} & \lambda \in I\!R^n, \quad \mu \in I\!R^m. \end{array}$$

11.11 **(a)** For a fixed λ ($\lambda \geq 0$), we obtain

$$(P_\lambda) \qquad \begin{array}{ll} \min & f_\lambda(x) \\ \text{s.t.} & x_j \geq 0, \quad j = 1,\ldots,n. \end{array}$$

with

$$f_\lambda(x) = \sum_{j=1}^{n} x_j^3 + \lambda(b - \sum_{j=1}^{n} a_j x_j).$$

Since f_λ is a separable function which is convex on the positive orthant, the minimizing $x(\lambda)$ can be determined analytically as

$$x_j(\lambda) = \sqrt{\frac{\lambda a_j}{3}}, \quad j = 1,\ldots,n.$$

Hence, the dual objective function is given by

$$\varphi(\lambda) = f_\lambda(x(\lambda)) = \ldots = \lambda b - 2\left(\frac{\lambda}{3}\right)^{3/2}\sum_{j=1}^{n} a_j^{3/2}.$$

Finally, we obtain the dual problem as

$$\begin{array}{ll} \max & \varphi(\lambda) \\ \text{s.t.} & \lambda \geq 0, \end{array}$$

with $\varphi(\lambda)$ as above.

(b) Assume that the constraint $\lambda \geq 0$ is inactive. Then, since φ is a concave function for $\lambda \geq 0$, the maximizing $\hat{\lambda}$ is given by

$$0 = \varphi'(\hat{\lambda}) = b - \sqrt{\frac{\hat{\lambda}}{3}} \sum_{j=1}^{n} a_j^{3/2},$$

i.e.,

$$\hat{\lambda} = \frac{3b^2}{\sum_{j=1}^{n} a_j^{3/2}}.$$

Since this value of $\hat{\lambda}$ is positive, it was correct to assume that the constraint $\lambda \geq 0$ is inactive, and we have found the maximizer.

11.12 (20040310-nr.4)

The Lagrange function to the problem, with $\mathbf{x} \in I\!R^n$ and $y \in I\!R$, is given by

$$L(\mathbf{x}, y) = \tfrac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{x} - \mathbf{c}^\mathsf{T}\mathbf{x} + y\,(\tfrac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{x} - \mathbf{a}^\mathsf{T}\mathbf{x}) = \frac{1+y}{2}\,\mathbf{x}^\mathsf{T}\mathbf{x} \;-\; (\mathbf{c} + \mathbf{a}y)^\mathsf{T}\mathbf{x}\,.$$

The Lagrange relaxed problem KPR_y consists in that, for a given $y \geq 0$, minimize $L(\mathbf{x}, y)$ with respect to $\mathbf{x} \in I\!R^n$.

The optimal solution to KPR_y becomes in our case $\mathbf{x}(y) = \dfrac{1}{1+y}\,(\mathbf{c} + \mathbf{a}y)$.

Then the dual objective function becomes

$$\varphi(y) = L(\mathbf{x}(y), y) = -\,\frac{(\mathbf{c} + \mathbf{a}y)^\mathsf{T}(\mathbf{c} + \mathbf{a}y)}{2\,(1+y)} = -\,\frac{1+y^2}{2\,(1+y)}\;,$$

where we have used the given relations $\mathbf{a}^\mathsf{T}\mathbf{a} = 1$, $\mathbf{c}^\mathsf{T}\mathbf{c} = 1$ and $\mathbf{a}^\mathsf{T}\mathbf{c} = 0$.

The dual problem consists in maximizing $\varphi(y)$ with respect to $y \geq 0$.

Straightforward differentiation gives that $\varphi'(y) = \dfrac{1 - 2y - y^2}{2\,(1+y)^2}$.

In particular $\varphi'(0) = \frac{1}{2} > 0$, which means that $y = 0$ does not maximize $\varphi(y)$.

That y is optimal to the dual problem is therefore equivalent to that $\varphi'(y) = 0$ and $y > 0$, which in turn is equivalent to that $1 - 2y - y^2 = 0$ and $y > 0$,
which is fulfilled for (and only for) $\hat{y} = \sqrt{2} - 1$.

Let $\hat{\mathbf{x}} = \mathbf{x}(\hat{y}) = \dfrac{1}{\sqrt{2}}\,(\mathbf{c} + \mathbf{a}(\sqrt{2} - 1)) = \mathbf{a} + \sqrt{0.5} \cdot (\mathbf{c} - \mathbf{a})$.

Then the following is fulfilled:

(1): $\hat{\mathbf{x}}$ minimizes $L(\mathbf{x}, \hat{y})$ with respect to $\mathbf{x}$, since $\hat{\mathbf{x}} = \mathbf{x}(\hat{y})$.

(2): $\hat{\mathbf{x}}$ is a feasible solution to the primal problem, since $\frac{1}{2}\,\hat{\mathbf{x}}^\mathsf{T}\hat{\mathbf{x}} - \mathbf{a}^\mathsf{T}\hat{\mathbf{x}} = \cdots = 0$.

(3): $\hat{y} \geq 0$.

(4): $\hat{y} \cdot (\frac{1}{2}\,\hat{\mathbf{x}}^\mathsf{T}\hat{\mathbf{x}} - \mathbf{a}^\mathsf{T}\hat{\mathbf{x}}) = 0$.

Hence $(\hat{\mathbf{x}},\, \hat{y})$ fulfills the global optimality conditions,
which implies that $\hat{\mathbf{x}}$ is an optimal solution to the primal problem.

12. Quadratic programming

12.1 (20070601-nr.3)

Let $\mathbf{x} = (x_{13},\, x_{14},\, x_{23},\, x_{24})^\mathsf{T} \in I\!R^4$.

Since all $R_{ij} = 1$ the effect minimizing problem is equivalent to the QR-problem

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{I}\,\mathbf{x} \quad (= \text{half the heat effect}) \\ \text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \end{array}$$

where $\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & -1 & 0 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 500 \\ 100 \\ -500 \end{pmatrix}$.

This QP-problem is equivalent to the following linear system of equations

$$\begin{array}{rcrcl} \mathbf{I}\mathbf{x} & - & \mathbf{A}^\mathsf{T}\mathbf{u} & = & \mathbf{0} \\ \mathbf{A}\mathbf{x} & & & = & \mathbf{b} \end{array}$$

From $\mathbf{I}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{0}$ is obtained that $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u}$, which in $\mathbf{A}\mathbf{x} = \mathbf{b}$ gives the equation system $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{b}$.

In our case $\mathbf{A}\mathbf{A}^\mathsf{T} = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 500 \\ 100 \\ -500 \end{pmatrix}$.

The Gauss-Jordan method (or gauss elimination) applied on the system

$\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 500 \\ 100 \\ -500 \end{pmatrix}$ gives the solution $\mathbf{u} = \begin{pmatrix} 150 \\ -50 \\ -200 \end{pmatrix}$.

Then $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} 150 \\ -50 \\ -200 \end{pmatrix} = \begin{pmatrix} 350 \\ 150 \\ 150 \\ -50 \end{pmatrix}$,

i.e. $x_{13} = 350$, $x_{14} = 150$, $x_{23} = 150$ and $x_{24} = -50$.
The current in link $(2,4)$ is hence going from node 4 to node 2!

12.2 (20070307-nr.3)

$f(\mathbf{x}) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2 = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 - 2x_3x_1 =$

$= \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$ with $\mathbf{H} = \begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix}$.

Since $f(\mathbf{x})$ is a sum of three squares, $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in I\!R^3$, which implies that $\mathbf{H}$ is positive semidefinite. (But not positive definite, since $f(\mathbf{x}) = 0$ if $x_1 = x_2 = x_3$.)

Gauss-Jordan gives after some elementary row operations that the system $\mathbf{Ax} = \mathbf{b}$ is equivalent to the system $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$. From this follows that *one* feasible solution to the system is $\bar{\mathbf{x}} = (2, 4, 0)^{\mathsf{T}}$.

By setting the right-hand side to zero we obtain that the system $\mathbf{Az} = \mathbf{0}$ is equivalent to the system $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \mathbf{z} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. From this follows that a basis to $\mathcal{N}(\mathbf{A})$ is given by the single vector $\mathbf{z} = (1, -2, 1)^{\mathsf{T}}$. ($z_3 = 1 \Rightarrow z_1 = 1$ and $z_2 = -2$.)

We now search an optimal solution to the QP-problem

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{Hx} \\ \text{s.t.} & \mathbf{Ax} = \mathbf{b}\,, \end{array}$$

We know that $\mathbf{Ax} = \mathbf{b}$ is equivalent to $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{z}\,v$ for $v \in I\!R$.

Insertion of this expression in the objective function leads the following optimization problem in the single variable v:

minimize $\frac{1}{2}\,v\mathbf{z}^{\mathsf{T}}\mathbf{Hz}v + \mathbf{z}^{\mathsf{T}}\mathbf{H}\bar{\mathbf{x}}v + \frac{1}{2}\,\bar{\mathbf{x}}^{\mathsf{T}}\mathbf{H}\bar{\mathbf{x}} = 18v^2 - 36v + 24.$

This is a convex quadratic function which is minimized by $\hat{v} = 1$.

Hence the optimal solution to the QP-problem is $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = (3,\ 2,\ 1)^{\mathsf{T}}$.

Since $\mathbf{H}$ according to above is positive semidefinite $\hat{\mathbf{x}}$ is a minimizer to $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{Hx} + \mathbf{C}^{\mathsf{T}}\mathbf{x}$
if and only if $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{0}$. Hence there exists at least one minimizer to $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{Hx} + \mathbf{C}^{\mathsf{T}}\mathbf{x}$ if and only if the system of equations $\mathbf{Hx} = -\mathbf{c}$ has at least one solution.
But this system has at least one solution if and only if $\mathbf{c} \in \mathcal{R}(\mathbf{H}) = \mathcal{N}(\mathbf{H}^{\mathsf{T}})^{\perp}$.

Gauss-Jordan gives after some simple row operations that the system $\mathbf{H}^{\mathsf{T}}\mathbf{z} = \mathbf{0}$ is equivalent to the system $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{z} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. From this follows that a basis for $\mathcal{N}(\mathbf{H}^{\mathsf{T}})$ is given by $\mathbf{z} = (1,\ 1,\ 1)^{\mathsf{T}}$.

Let us call this single basis vector $\mathbf{a}$, i.e. $\mathbf{a} = (1,\ 1,\ 1)^{\mathsf{T}}$.

Then it holds that $\mathbf{c} \in \mathcal{N}(\mathbf{H}^{\mathsf{T}})^{\perp}$ if and only if $\mathbf{c}$ is orthogonal to this basis vector $\mathbf{a}$.

Hence: $\frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{Hx} + \mathbf{C}^{\mathsf{T}}\mathbf{x}$ has at least one minimizer if and only if $\mathbf{a}^{\mathsf{T}}\mathbf{c} = 0$.

12.3 (20060603-nr.2)

(a)

For simplicity we multiply the objective function with the factor $\frac{1}{2}$, which do not influence the optimal solution $\mathbf{x}$ to the problem. Then the objective function becomes

$\frac{1}{2}\,(\mathbf{x}-\mathbf{q})^{\mathsf{T}}(\mathbf{x}-\mathbf{q}) = \frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} - \mathbf{q}^{\mathsf{T}}\mathbf{x} + \frac{1}{2}\,\mathbf{q}^{\mathsf{T}}\mathbf{q},$

so the problem P1 is equivalent to a quadratic optimization problem with linear equality constraints on the form

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} + c_0 \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \end{aligned}$$

where $\mathbf{H} = \mathbf{I}$, $\mathbf{c} = -\mathbf{q}$, $c_0 = \frac{1}{2}\mathbf{q}^\mathsf{T}\mathbf{q}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The matrix $\mathbf{H} = \mathbf{I}$ is positive definite, so we have a *convex* QP-problem.

The optimality constraints are then given by $\mathbf{H}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{c}$ and $\mathbf{A}\mathbf{x} = \mathbf{b}$, i.e. $\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{q}$ and $\mathbf{A}\mathbf{x} = \mathbf{0}$.

The first equations give that $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u} + \mathbf{q}$, which inserted in $\mathbf{A}\mathbf{x} = \mathbf{0}$ gives the equations system

$\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{A}\mathbf{q}$, i.e. $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -8 \\ -4 \end{pmatrix}$, with the solution $\hat{\mathbf{u}} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$.

Then the optimal solution $\hat{\mathbf{x}}$ is $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{u}} + \mathbf{q} = \begin{pmatrix} -3 \\ -1 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

(b)

Insert $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{v}$ to the objective function, which then becomes the following quadratic function, after multiplication with the factor $\frac{1}{2}$:

$f(\mathbf{v}) = \frac{1}{2}(\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q})^\mathsf{T}(\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q}) = \frac{1}{2}\mathbf{v}^\mathsf{T}(\mathbf{A}\mathbf{A}^\mathsf{T})\mathbf{v} - (\mathbf{A}\mathbf{q})^\mathsf{T}\mathbf{v} + \frac{1}{2}\mathbf{q}^\mathsf{T}\mathbf{q}$.

$\mathbf{A}\mathbf{A}^\mathsf{T} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ is positive definite.

$f(\mathbf{v})$ is hence a strictly convex quadratic function that is minimized when its gradient is the null vector i.e. when $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{A}\mathbf{q} = \mathbf{0}$.

This gives the system $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{v} = \mathbf{A}\mathbf{q}$, i.e. $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$,

with the solution $\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

The optimal solution $\hat{\mathbf{x}}$ is then given by $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{v}} = \begin{pmatrix} 3 \\ 1 \\ -1 \\ -3 \end{pmatrix}$.

You can note that the optimal $\mathbf{x}$ to P1 and optimal $\mathbf{x}$ to P2 are orthogonal and sum up to $\mathbf{q}$. This is of course not a surprise since the two subspaces $\mathcal{N}(\mathbf{A})$ and $\mathcal{R}(\mathbf{A}^\mathsf{T})$ are each others orthogonal complements.

12.4 (20060308-nr.4a)

The problem to minimize $\frac{1}{2}|\mathbf{x}-\bar{\mathbf{x}}|^2$ s.t. $\mathbf{a}^\mathsf{T}\mathbf{x}=b$ is a QP-problem with equality constraints.

It can be written on the form

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}+\mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x}=\mathbf{b}\,, \end{aligned}$$

where $\mathbf{H}=\mathbf{I}$, $\mathbf{c}=-\bar{\mathbf{x}}$, $\mathbf{A}=\mathbf{a}^\mathsf{T}$ and $\mathbf{b}=b$. (Since $\frac{1}{2}\,|\,\mathbf{x}-\bar{\mathbf{x}}\,|^2=\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x}-\bar{\mathbf{x}}^\mathsf{T}\mathbf{x}+\frac{1}{2}\bar{\mathbf{x}}^\mathsf{T}\bar{\mathbf{x}}$).

Since $\mathbf{H}$ is positive definite, the following optimality conditions are both necessary and sufficient:

$\mathbf{H}\mathbf{x}+\mathbf{c}=\mathbf{A}^\mathsf{T}\mathbf{u}$ and $\mathbf{A}\mathbf{x}=\mathbf{b}$, which can be written as $\mathbf{x}-\bar{\mathbf{x}}=\mathbf{a}\,u$ and $\mathbf{a}^\mathsf{T}\mathbf{x}=b$.

We obtain that $\mathbf{x}=\bar{\mathbf{x}}+\mathbf{a}\,u$, which inserted in $\mathbf{a}^\mathsf{T}\mathbf{x}=b$ gives that $u=(b-\mathbf{a}^\mathsf{T}\bar{\mathbf{x}})/|\mathbf{a}|^2$.

An optimal solution to our problem is hence given by $\hat{\mathbf{x}}=\bar{\mathbf{x}}+\mathbf{a}\,u$, with $u=(b-\mathbf{a}^\mathsf{T}\bar{\mathbf{x}})/|\mathbf{a}|^2$.

According to the prerequisites $\mathbf{a}^\mathsf{T}\bar{\mathbf{x}}<b$, and hence $u>0$.

The shortest distance is then $d=|\,\hat{\mathbf{x}}-\bar{\mathbf{x}}\,|=|\mathbf{a}\,u|=|\mathbf{a}|\,|u|=|\mathbf{a}|\,u=(b-\mathbf{a}^\mathsf{T}\bar{\mathbf{x}})/|\mathbf{a}|$.

12.5 (20050331-nr.4)

The problem can be formulated as:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}|\mathbf{u}-\mathbf{v}|^2=\tfrac{1}{2}(\mathbf{u}-\mathbf{v})^\mathsf{T}(\mathbf{u}-\mathbf{v}) \\ \text{s.t.} \quad & \mathbf{R}\mathbf{u}=\mathbf{p}\,, \\ & \mathbf{S}\,\mathbf{v}=\mathbf{q}\,. \end{aligned}$$

With “Lagrange multipliers” $\mathbf{y}\in I\!R^3$ and $\mathbf{z}\in I\!R^3$ the optimality conditions become

$\mathbf{u}-\mathbf{v}-\mathbf{R}^\mathsf{T}\mathbf{y}=\mathbf{0}$, $\mathbf{v}-\mathbf{u}-\mathbf{S}^\mathsf{T}\mathbf{z}=\mathbf{0}$, $\mathbf{R}\mathbf{u}=\mathbf{p}$ and $\mathbf{S}\,\mathbf{v}=\mathbf{q}$.

This is just a linear system of equations, but since there are 14 unknowns and there is no obvious way to simplify the system (and we have no computer on the exam) we try a null-space method instead!

Simple computations give that $\mathbf{R}\mathbf{u}=\mathbf{p} \;\Leftrightarrow\; \mathbf{u}=\mathbf{u}_0+\mathbf{g}x_1$, where $\mathbf{u}_0=(1,1,1,0)^\mathsf{T}$, $\mathbf{g}=(-1,-1,-1,\,1\,)^\mathsf{T}$ and x_1 is an arbitrary real number.

Just as simple computations give that $\mathbf{S}\,\mathbf{v}=\mathbf{q} \;\Leftrightarrow\; \mathbf{v}=\mathbf{v}_0+\mathbf{h}x_2$, where $\mathbf{v}_0=(0,2,2,2)^\mathsf{T}$, $\mathbf{h}=(\,1,-1,-1,-1\,)^\mathsf{T}$ and x_2 is an arbitrary real number.

Note that $\mathbf{g}$ is a basis to $\mathcal{N}(\mathbf{R})$ while $\mathbf{h}$ is a basis to $\mathcal{N}(\mathbf{S})$.

Our original problem can then be written as the following problem in the variables x_1 and x_2:

minimize $\frac{1}{2}|\mathbf{u}_0-\mathbf{v}_0+\mathbf{g}x_1-\mathbf{h}x_2|^2$,

which is equivalent to the MC-problem to minimize $\frac{1}{2}|\mathbf{A}\mathbf{x}-\mathbf{b}|^2$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{A} = \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \mathbf{v}_0 - \mathbf{u}_0 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

The normal equations $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}$ become $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$,

with the solution $\mathbf{x} = (x_1,\ x_2)^\mathsf{T} = (1/4,\ 5/4)^\mathsf{T}$.

Hence $\hat{\mathbf{u}} = (3/4,\ 3/4,\ 3/4,\ 1/4)^\mathsf{T}$ and $\hat{\mathbf{v}} = (5/4,\ 3/4,\ 3/4,\ 3/4)^\mathsf{T}$.

The shortest distance between the two sets (lines) is hence $d = |\hat{\mathbf{u}} - \hat{\mathbf{v}}| = 1/\sqrt{2}$.

Now that we know $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, we can conclude that the optimality conditions that we relaxed before are fulfilled with $\mathbf{y} = (1/2,\ 0,\ 0)^\mathsf{T}$ and $\mathbf{z} = (0,\ 0,\ -1/2)^\mathsf{T}$.

12.6 (20050307-nr.3)

(a) A vector $\mathbf{x}$ is optimal to the MC-problem P1 if and only if $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}$.

In our case $\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$ and $\mathbf{A}^\mathsf{T}\mathbf{b} = \begin{pmatrix} b_1 - b_2 \\ b_2 - b_1 \end{pmatrix}$, so the solutions to the normal equations above, and hence also the optimal solutions to P1, are given by

$x_1 = b_1/2 - b_2/2 + t$ and $x_2 = t$, where t is an arbitrary real number.

(b) Let $X(\mathbf{b}) = \{\ \mathbf{x} \in I\!R^2 \mid x_1 = b_1/2 - b_2/2 + t \text{ and } x_2 = t, \text{ for } t \in I\!R\ \}$.

For $\mathbf{x} \in X(\mathbf{b})$ it holds that $\mathbf{x}^\mathsf{T}\mathbf{x} = (b_1/2 - b_2/2 + t)^2 + t^2$, which is minimized by

$t = b_2/4 - b_1/4$, and hence $x_1 = b_1/4 - b_2/4$ and $x_2 = b_2/4 - b_1/4$.

This is hence the optimal solution $\hat{\mathbf{x}}(\mathbf{b})$ to P2.

(c) From the above we see that $\hat{\mathbf{x}}(\mathbf{b}) = \mathbf{A}^+\mathbf{b}$, with $\mathbf{A}^+ = \begin{bmatrix} 1/4 & -1/4 \\ -1/4 & 1/4 \end{bmatrix}$.

(d) A vector $\mathbf{x}$ is optimal to the problem P3 if and only if $(\mathbf{A}^\mathsf{T}\mathbf{A} + \varepsilon\,\mathbf{I})\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}$.

In our case $\mathbf{A}^\mathsf{T}\mathbf{A} + \varepsilon\,\mathbf{I} = \begin{bmatrix} 2+\varepsilon & -2 \\ -2 & 2+\varepsilon \end{bmatrix}$ and $\mathbf{A}^\mathsf{T}\mathbf{b} = \begin{pmatrix} b_1 - b_2 \\ b_2 - b_1 \end{pmatrix}$, so the solution to this system of equations, and hence also the optimal solution to P3, is given by

$x_1 = (b_1 - b_2)/(4+\varepsilon)$; and $x_2 = (b_2 - b_1)/(4+\varepsilon)$;.

This is hence the optimal solution $\tilde{\mathbf{x}}_\varepsilon(\mathbf{b})$ to P3.

(e) From the above follows that $\tilde{\mathbf{x}}_\varepsilon(\mathbf{b}) = \tilde{\mathbf{A}}_\varepsilon\mathbf{b}$, with $\tilde{\mathbf{A}}_\varepsilon = \begin{bmatrix} 1/(4+\varepsilon) & -1/(4+\varepsilon) \\ -1/(4+\varepsilon) & 1/(4+\varepsilon) \end{bmatrix}$.

We see that $\tilde{\mathbf{A}}_\varepsilon \longrightarrow \begin{bmatrix} 1/4 & -1/4 \\ -1/4 & 1/4 \end{bmatrix} = \mathbf{A}^+$ if $\varepsilon \to 0$.

12.7 (20041016-nr.3)

(a). Let $\mathbf{x}(\alpha) = \mathbf{a} + \alpha \cdot \mathbf{u} \in L_1$ and $\mathbf{y}(\beta) = \mathbf{b} + \beta \cdot \mathbf{v} \in L_2$.
The quadratic distance between $\mathbf{x}(\alpha)$ and $\mathbf{y}(\beta)$ is given, with $\mathbf{c} = \mathbf{b} - \mathbf{a}$ and $\mathbf{z} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, by

$f(\mathbf{z}) = (\mathbf{y}(\beta) - \mathbf{x}(\alpha))^\mathsf{T}(\mathbf{y}(\beta) - \mathbf{x}(\alpha)) = (\mathbf{c} + \beta \cdot \mathbf{v} - \alpha \cdot \mathbf{u})^\mathsf{T}(\mathbf{c} + \beta \cdot \mathbf{v} - \alpha \cdot \mathbf{u}) =$
$= \mathbf{c}^\mathsf{T}\mathbf{c} - 2(\mathbf{c}^\mathsf{T}\mathbf{u})\alpha + 2(\mathbf{c}^\mathsf{T}\mathbf{v})\beta + (\mathbf{u}^\mathsf{T}\mathbf{u})\alpha^2 + (\mathbf{v}^\mathsf{T}\mathbf{v})\beta^2 - 2(\mathbf{u}^\mathsf{T}\mathbf{v})\alpha\beta =$
$= \mathbf{c}^\mathsf{T}\mathbf{c} - 2\,\mathbf{g}^\mathsf{T}\mathbf{z} + \mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$, where $\mathbf{g} = \begin{pmatrix} \mathbf{c}^\mathsf{T}\mathbf{u} \\ -\mathbf{c}^\mathsf{T}\mathbf{v} \end{pmatrix}$ and $\mathbf{H} = \begin{bmatrix} \mathbf{u}^\mathsf{T}\mathbf{u} & -\mathbf{u}^\mathsf{T}\mathbf{v} \\ -\mathbf{u}^\mathsf{T}\mathbf{v} & \mathbf{v}^\mathsf{T}\mathbf{v} \end{bmatrix} =$
$\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$.

Since $\mathbf{u}$ and $\mathbf{v}$ according to the prerequisites not are parallel, $|\rho| < 1$, which implies that the matrix $\mathbf{H}$ is positive definite, which in turn implies that f is a strictly *convex* quadratic function.

(b). Global minimum is obtained when $\nabla f(\mathbf{z})^\mathsf{T} = \mathbf{0}$, i.e. when $\mathbf{H}\mathbf{z} = \mathbf{g}$.

Especially when $\rho = \mathbf{u}^\mathsf{T}\mathbf{v} = 0$ then $\mathbf{H} = \mathbf{I} =$ is the unitary matrix, and the optimal solution is given by $\mathbf{z} = \mathbf{g}$, i.e. $\alpha = \mathbf{c}^\mathsf{T}\mathbf{u}$ and $\beta = -\mathbf{c}^\mathsf{T}\mathbf{v}$.

Insertion of this in $\mathbf{x}(\alpha)$ and $\mathbf{y}(\beta)$ gives that $\hat{\mathbf{x}} = \mathbf{a} + (\mathbf{c}^\mathsf{T}\mathbf{u}) \cdot \mathbf{u}$ and $\hat{\mathbf{y}} = \mathbf{b} - (\mathbf{c}^\mathsf{T}\mathbf{v}) \cdot \mathbf{v}$.

Especially then $\hat{\mathbf{y}} - \hat{\mathbf{x}} = \mathbf{c} - \mathbf{u}\,\mathbf{c}^\mathsf{T}\mathbf{u} - \mathbf{v}\,\mathbf{c}^\mathsf{T}\mathbf{v}$, and the quadratic length of the thread in is given by

$|\,\hat{\mathbf{y}} - \hat{\mathbf{x}}\,|^2 = |\,\mathbf{c} - \mathbf{u}\,\mathbf{c}^\mathsf{T}\mathbf{u} - \mathbf{v}\,\mathbf{c}^\mathsf{T}\mathbf{v}\,|^2 = \quad \cdots \quad = \mathbf{c}^\mathsf{T}\mathbf{c} - (\mathbf{c}^\mathsf{T}\mathbf{u})^2 - (\mathbf{c}^\mathsf{T}\mathbf{v})^2.$

12.8 (20040415-nr.3)

(a)

You quickly realize (for example with Gauss-Jordan) that the matrix $\mathbf{A}$ has the rank $r = 3$.
Hence $\mathcal{N}(\mathbf{A})$ has the dimension $n - r = 5 - 3 = 2$, so a basis to $\mathcal{N}(\mathbf{A})$ consists of two linearly independent vectors that both are in $\mathcal{N}(\mathbf{A})$. But the in the text given vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ are linearly independent and fulfill $\mathbf{A}\mathbf{z}_1 = \mathbf{A}\mathbf{z}_2 = \mathbf{0}$. Hence they are a basis for $\mathcal{N}(\mathbf{A})$.
You can also conclude that the in the text given vector $\bar{\mathbf{x}}$ fulfills $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$.

(b)

We have a QP-problem on the form: minimize $\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{H}$, $\mathbf{A}$ and $\mathbf{b}$ is the given from the text, while $\mathbf{c} = \mathbf{0}$.

A feasible solution $\bar{\mathbf{x}}$ and a null-space matrix $\mathbf{Z}$ is given according to above by

$$\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{Z} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Every feasible solution $\mathbf{x}$ can now be written on the form $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, for $\mathbf{v} \in I\!R^2$.

The optimal $\mathbf{v}$ is obtained by solving the system $(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$,

which becomes $\begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$, with the solution $\hat{\mathbf{v}} = \begin{pmatrix} 0 \\ -1/3 \end{pmatrix}$.

Then the optimal solution to the original problem is $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}} = \begin{pmatrix} 2/3 \\ 3/3 \\ 4/3 \\ 3/3 \\ 2/3 \end{pmatrix}$.

The optimal value of the problem is $\frac{1}{2}\hat{\mathbf{x}}^\mathsf{T}\mathbf{H}\hat{\mathbf{x}} = \dfrac{26}{3}$.

(c)

The wanted vector $\hat{\mathbf{u}}$ shall together with the above computed optimal solution $\hat{\mathbf{x}}$ fulfill that $\begin{bmatrix} \mathbf{H} & -\mathbf{A}^\mathsf{T} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} -\mathbf{c} \\ \mathbf{b} \end{pmatrix}$.

But $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\mathbf{c} = \mathbf{0}$, so we obtain the constraint that $\mathbf{A}^\mathsf{T}\hat{\mathbf{u}} = \mathbf{H}\hat{\mathbf{x}}$, i.e.

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \end{pmatrix} = \begin{pmatrix} 7/3 \\ 12/3 \\ 14/3 \\ 12/3 \\ 7/3 \end{pmatrix}$, which is fulfilled by (and only by) $\hat{\mathbf{u}} = \begin{pmatrix} 7/3 \\ 12/3 \\ 7/3 \end{pmatrix}$.

12.9 (20040310-nr.2)

Let $\mathbf{x} = (x_{12}, x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25}, x_{34}, x_{35}, x_{45})^\mathsf{T} \in I\!R^{10}$.

Since all $r_{ij} = 1$, the effect minimizing problem is equivalent with the QP-problem

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{I}\,\mathbf{x} \quad (= \text{half the heat effect}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 0 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

This QP-problem is in turn equivalent to the linear equation system

$$\begin{aligned} \mathbf{I}\,\mathbf{x} \;-\; \mathbf{A}^\mathsf{T}\mathbf{u} &= \mathbf{0} \\ \mathbf{A}\mathbf{x} \qquad\quad &= \mathbf{b} \end{aligned}$$

From $\mathbf{I}\,\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{0}$ it is obtained that $\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{u}$, which inserted in $\mathbf{A}\mathbf{x} = \mathbf{b}$ gives the equation system $\mathbf{A}\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{b}$.

In our case $\mathbf{A}\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$,

so the given computational help gives that $\mathbf{u} = (\mathbf{A}\mathbf{A}^{\mathsf{T}})^{-1}\mathbf{b} = (0.4,\, 0.2,\, 0.2,\, 0.2)^{\mathsf{T}}$.

This gives in turn that the optimal solution $\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{u} = (0.2,\, 0.2,\, 0.2,\, 0.4,\, 0,\, 0,\, 0.2,\, 0,\, 0.2,\, 0.2)^{\mathsf{T}}$.

Hence the heat-effect $\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} = 0.4$ (= the total resistance between node 1 and node 5.)

13. Nonlinear programming

13.1 (20070601-nr.5)

(a) The problem can be written as: minimize $f(\mathbf{x})$ s.t. $g_i(\mathbf{x}) \le 0$, $i = 1, 2, 3$, where

$f(\mathbf{x}) = c_1x_1 - 4x_2 - 2x_3$, $g_1(\mathbf{x}) = x_1^2 + x_2^2 - 2$, $g_2(\mathbf{x}) = x_1^2 + x_3^2 - 2$, $g_3(\mathbf{x}) = x_2^2 + x_3^2 - 2$.

The objective function is linear and hence convex. The constraint functions have the second derivative matrices $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, which are positive semidefinite for all $\mathbf{x}$.

Hence also the constraint functions are convex, which makes the whole problem convex. Furthermore for example $\mathbf{x} = (0, 0, 0)^\mathsf{T}$ fulfills all constraints with strict inequality, so the studied problem is a *regular* convex problem. This implies that a point $\hat{\mathbf{x}}$ is a global optimal solution to the problem if and only if $\hat{\mathbf{x}}$ is a KKT-point.

(b) The Lagrangian can be written as $L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \sum_{i=1}^{3} y_i g_i(\mathbf{x}) =$

$= c_1x_1 - 4x_2 - 2x_3 + y_1(x_1^2 + x_2^2 - 2) + y_2(x_1^2 + x_3^2 - 2) + y_3(x_2^2 + x_3^2 - 2).$

The KKT-constraints can be divided into four groups in the following way.

(KKT–1) $\partial L/\partial x_j = 0$ for $j = 1, 2, 3$:
$c_1 + 2x_1(y_1 + y_2) = 0,$
$-4 + 2x_2(y_1 + y_3) = 0,$
$-2 + 2x_3(y_2 + y_3) = 0.$

(KKT–2) Feasible point, i.e. $g_i(\mathbf{x}) \le 0$ for $i = 1, 2, 3$:
$x_1^2 + x_2^2 - 2 \le 0,$
$x_1^2 + x_3^2 - 2 \le 0,$
$x_2^2 + x_3^2 - 2 \le 0.$

(KKT–3) Lagrangian multipliers non-negative:
$y_1 \ge 0,$
$y_2 \ge 0,$
$y_3 \ge 0.$

(KKT–4) Complementarity, i.e $y_i g_i(\mathbf{x}) = 0$ for $i = 1, 2, 3$:
$y_1(x_1^2 + x_2^2 - 2) = 0,$
$y_2(x_1^2 + x_3^2 - 2) = 0,$
$y_3(x_2^2 + x_3^2 - 2) = 0.$

(c) First suppose $\mathbf{x} = (1.4,\ 0.2,\ 0.2)^\mathsf{T}$.

Then $x_1^2 + x_2^2 - 2 = 0$, $x_1^2 + x_3^2 - 2 = 0$, $x_2^2 + x_3^2 - 2 < 0$.

The complementarity constraints the give that $y_3 = 0$, and the KKT–1 constraints can be written as

$$c_1 + 2.8(y_1 + y_2) = 0,$$
$$-4 + 0.4(y_1 + 0) = 0,$$
$$-2 + 0.4(y_2 + 0) = 0.$$

We see that this system has no solution if $c_1 \neq -42$.
If $c_1 = -42$ then $\mathbf{x} = (1.4,\ 0.2,\ 0.2)^\mathsf{T}$, together with $\mathbf{y} = (10,\ 5,\ 0)^\mathsf{T}$ fulfills all KKT-constraints, and $\mathbf{x}$ is a global optimal solution to the problem.

(d)

Now suppose $\mathbf{x} = (1,\ 1,\ 1)^\mathsf{T}$.

Then $x_1^2 + x_2^2 - 2 = 0,\ x_1^2 + x_3^2 - 2 = 0,\ x_2^2 + x_3^2 - 2 = 0.$

The KKT–1 constraints can now be written as

$$\begin{aligned} y_1 + y_2 &= -0.5\, c_1, \\ y_1 + y_3 &= \ \ 2, \\ y_2 + y_3 &= \ \ 1. \end{aligned}$$

The solution of this system of equations in $\mathbf{y}$ becomes, with help of the given help,

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} -0.5\, c_1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{pmatrix} -0.5\, c_1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{4} \cdot \begin{pmatrix} -c_1 + 2 \\ -c_1 - 2 \\ c_1 + 6 \end{pmatrix}.$$

We see that the KKT–3 constraints will be fulfilled if and only if $-6 \leq c_1 \leq -2$.

For these values of the constant c,
$\mathbf{x} = (1,\ 1,\ 1)^\mathsf{T}$ together with
$\mathbf{y} = \left(\frac{2 - c_1}{4},\ \frac{-2 - c_1}{4},\ \frac{6 + c_1}{4} \right)^\mathsf{T}$, fulfills all KKT-constraints, and hence $\mathbf{x}$ is a global optimal solution.

(e) The Lagrangian for the problem now becomes, with $c_1 = -6$,

$$L(\mathbf{x}, \mathbf{y}) = -6x_1 - 4x_2 - 2x_3 + y_1(x_1^2 + x_2^2 - 2) + y_2(x_1^2 + x_3^2 - 2) + y_3(x_2^2 + x_3^2 - 2) =$$
$$= ((y_1 + y_2)x_1^2 - 6x_1) + ((y_1 + y_3)x_2^2 - 4x_2) + ((y_2 + y_3)x_3^2 - 2x_3) - 2(y_1 + y_2 + y_3).$$

To obtain the dual objective function value $\varphi(\hat{\mathbf{y}})$, where $\hat{\mathbf{y}} = (1, 1, 1)^\mathsf{T}$,

one should minimize $L(\mathbf{x}, \hat{\mathbf{y}})$ with respect to $\mathbf{x} \in I\!R^3$.

But $L(\mathbf{x}, \hat{\mathbf{y}}) = 2x_1^2 - 6x_1 + 2x_2^2 - 4x_2 + 2x_3^2 - 2x_3 - 6$,

so the minimized values on x_j are given by

$x_1(\hat{\mathbf{y}}) = 1.5,\ x_2(\hat{\mathbf{y}}) = 1.0,\ x_3(\hat{\mathbf{y}}) = 0.5,$

and the dual objective function value is given by

$\varphi(\hat{\mathbf{y}}) = L(\mathbf{x}(\hat{\mathbf{y}}), \hat{\mathbf{y}}) = 4.5 - 9 + 2 - 4 + 0.5 - 1 - 6 = -13.$

The (d)-exercise makes us guess that $\tilde{\mathbf{y}} = (2, 1, 0)^\mathsf{T}$ is an optimal solution for the dual problem.

Now $L(\mathbf{x}, \tilde{\mathbf{y}}) = 3x_1^2 - 6x_1 + 2x_2^2 - 4x_2 + x_3^2 - 2x_3 - 6$,

so the minimized values of x_j are given by

$x_1(\tilde{\mathbf{y}}) = 1,\ x_2(\tilde{\mathbf{y}}) = 1,\ x_3(\tilde{\mathbf{y}}) = 1,$

and the dual objective function value is given by

$\varphi(\hat{\mathbf{y}}) = L(\mathbf{x}(\hat{\mathbf{y}}), \hat{\mathbf{y}}) = 3 - 6 + 2 - 4 + 1 - 2 - 6 = -12.$

Since $\varphi(\hat{\mathbf{y}}) < \varphi(\tilde{\mathbf{y}})$, $\hat{\mathbf{y}}$ can not be an optimal solution to the dual problem (which consists of maximizing $\varphi(\mathbf{y})$ s.t. $\mathbf{y} \geq \mathbf{0}$).

13.2 (20070307-nr.4)

(a)

We have that $\mathbf{h}(\mathbf{x}) = \begin{pmatrix} x_1^2 - x_2 - \delta_1 \\ x_1^2 + x_2 - \delta_2 \\ x_2^2 - x_1 - \delta_3 \\ x_2^2 + x_1 - \delta_4 \end{pmatrix}$ and $f(\mathbf{x}) = \frac{1}{2}\mathbf{h}(\mathbf{x})^\mathsf{T}\mathbf{h}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in I\!R^2$.

As a special case, if all $\delta_i = 0$ and $\hat{\mathbf{x}} = (0,\, 0)^\mathsf{T}$ then $\mathbf{h}(\hat{\mathbf{x}}) = (0,\, 0,\, 0,\, 0)^\mathsf{T}$ and $f(\hat{\mathbf{x}}) = 0$.

Then $f(\hat{\mathbf{x}}) \leq f(\mathbf{x})$ for all $\mathbf{x} \in I\!R^2$, which means that $\hat{\mathbf{x}}$ is a global minimizer to $f(\mathbf{x})$.

(b)

Now $\delta_1 = -0.1,\ \delta_2 = 0.1,\ \delta_3 = -0.2,\ \delta_4 = 0.2$ and $\mathbf{x}^{(1)} = (0,0)^\mathsf{T}$.

Differentiation gives

$\nabla h_1(\mathbf{x}) = (2x_1,\, -1),\ \ \nabla h_2(\mathbf{x}) = (2x_1,\ 1),\ \ \nabla h_3(\mathbf{x}) = (-1,\, 2x_2),\ \ \nabla h_4(\mathbf{x}) = (\,1,\, 2x_2).$

Hence $\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2x_1 & -1 \\ 2x_1 & 1 \\ -1 & 2x_2 \\ 1 & 2x_2 \end{bmatrix}$, such that $\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 0 \\ 1 & 0 \end{bmatrix}$ and $\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} 0.1 \\ -0.1 \\ 0.2 \\ -0.2 \end{pmatrix}$.

In the Gauss-Newton method you should solve the system of equations

$\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)})$

In our case $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -0.4 \\ -0.2 \end{pmatrix}$,

so the system of equations becomes $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.2 \end{pmatrix}$, with the solution $\mathbf{d}^{(1)} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$.

We test $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$.

Then

$h_1(\mathbf{x}^{(2)}) = 0.04 - 0.1 + 0.1 = 0.04,$

$h_2(\mathbf{x}^{(2)}) = 0.04 + 0.1 - 0.1 = 0.04,$
$h_3(\mathbf{x}^{(2)}) = 0.01 - 0.2 + 0.2 = 0.01,$
$h_4(\mathbf{x}^{(2)}) = 0.01 + 0.2 - 0.2 = 0.01,$

such that $f(\mathbf{x}^{(2)}) = 0.0017 < 0.05 = f(\mathbf{x}^{(1)})$. Hence the step $t_1 = 1$ was fine. Hence we have performed a complete iteration with the Gauss-Newton method and ended up in the point $\mathbf{x}^{(2)} = (0.2,\ 0.1)^\mathsf{T}$.

The gradient of the objective function in this point $\mathbf{x}^{(2)}$ is given by

$$\nabla f(\mathbf{x}^{(2)})^\mathsf{T} = \nabla \mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(2)}) = \begin{bmatrix} 0.4 & 0.4 & -1 & 1 \\ -1 & 1 & 0.2 & 0.2 \end{bmatrix} \begin{pmatrix} 0.04 \\ 0.04 \\ 0.01 \\ 0.01 \end{pmatrix} = \begin{pmatrix} 0.032 \\ 0.004 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since the gradient is not the null vector, $\mathbf{x}^{(2)}$ can not be a local minimizer.

13.3 (20060603-nr.3)

There are several ways of solving the problem. Here follows one of them.

The problem is a convex QP-problem, and hence the KKT-constraints are both necessary and sufficient constraints for a global optimal solution.

The Lagrangian function to the problem can be written as:

$L(\mathbf{x}, \mathbf{y}) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2 - x_1 - x_2 + c_3x_3 + y_1(4 - x_1 - x_2) + y_2(4 - x_1 - x_3) + y_3(4 - x_2 - x_3)$.

The KKT-constraints then becomes:

$$\begin{array}{rll}
x_1 - y_1 - y_2 = 1 & \partial L/\partial x_1 = 0 & \text{(KKT1)} \\
x_2 - y_1 - y_3 = 1 & \partial L/\partial x_2 = 0 & \text{(KKT2)} \\
x_3 - y_2 - y_3 = -c_3 & \partial L/\partial x_3 = 0 & \text{(KKT3)} \\
x_1 + x_2 \geq 4 & \text{primal feasibility} & \text{(KKT4)} \\
x_1 + x_3 \geq 4 & \text{primal feasibility} & \text{(KKT5)} \\
x_2 + x_3 \geq 4 & \text{primal feasibility} & \text{(KKT6)} \\
y_1 \geq 0 & \text{dual feasibility} & \text{(KKT7)} \\
y_2 \geq 0 & \text{dual feasibility} & \text{(KKT8)} \\
y_3 \geq 0 & \text{dual feasibility} & \text{(KKT9)} \\
y_1(x_1 + x_2 - 4) = 0 & \text{complementarity} & \text{(KKT10)} \\
y_2(x_1 + x_3 - 4) = 0 & \text{complementarity} & \text{(KKT11)} \\
y_3(x_2 + x_3 - 4) = 0 & \text{complementarity} & \text{(KKT12)}
\end{array}$$

(a). With $\mathbf{x} = (2, 2, 2)^\mathsf{T}$ (KKT4)–(KKT6) is fulfilled with equality, and hence (KKT10)–(KKT12) are fulfilled. (KKT1)–(KKT3) then gives (after solving a system of equations) that $y_1 = -c_3/2$ and $y_2 = y_3 = 1 + c_3/2$. In order to have (KKT7)–(KKT9) fulfilled it is required that $-2 \leq c_3 \leq 0$.
The KKT-constraints are hence fulfilled if and only if $c_3 \in [-2, 0]$.

(b). With $\mathbf{x} = (2, 2, 4)^\mathsf{T}$ (KKT4) is fulfilled with equality and (KKT5)–(KKT6) with strict inequality. Hence (KKT10) is fulfilled, while (KKT11)–(KKT12) requires that $y_2 = y_3 = 0$. (KKT1)–(KKT2) then gives that $y_1 = 1$, and to fulfill (KKT3), $c_3 = -4$ must be fulfilled. Then also (KKT7)–(KKT9)

are fulfilled.
The KKT-constraints are hence fulfilled if and only if $c_3 = -4$.

(c). With $\mathbf{x} = (3,3,1)^\mathsf{T}$, (KKT5)–(KKT6) are fulfilled with equality and (KKT4) with strict inequality. Hence (KKT11)–(KKT12) are fulfilled, while (KKT10) requires that $y_1 = 0$. (KKT1)–(KKT2) then gives that $y_2 = y_3 = 2$, so to make (KKT3) fulfilled $c_3 = 3$ is required. Then also (KKT7)–(KKT9) are fulfilled.
The KKT-constraints are hence fulfilled if and only if $c_3 = 3$.

13.4 (20060603-nr.4)

(a)

Since $f(\mathbf{x}) = (x_1 x_2^2 x_3^3)^2$, $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$. But $\hat{\mathbf{x}}$ is a feasible solution with $f(\hat{\mathbf{x}}) = 0$. Hence $f(\hat{\mathbf{x}}) \leq f(\mathbf{x})$ for all feasible solutions $\mathbf{x}$, which per definition implies that $\hat{\mathbf{x}}$ is a global optimal solution to the minimization problem.

(b)

Every global optimal solution is also a local optimal solution, so $\hat{\mathbf{x}}$ is a local optimal solution to the minimization problem.

(c)

f is convex if $f(t\mathbf{u} + (1-t)\mathbf{v}) \leq t f(\mathbf{u}) + (1-t) f(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and $t \in [0,1]$.
Let (for example) $\mathbf{u} = (1,0,1)^\mathsf{T}$, $\mathbf{v} = (1,1,0)^\mathsf{T}$, $t = 0.5$.
Then $f(\mathbf{u}) = f(\mathbf{v}) = 0$ while $f(t\mathbf{u} + (1-t)\mathbf{v}) > 0$. f is hence not convex.

(d)

You realize that the biggest value that $f(x)$ can obtain in the unit circle is strictly positive, i.e. all three variables are separated from zero in the maximum point/points.
Since $(-x_1)^2 = (+x_1)^2$, $(-x_2)^4 = (+x_2)^4$ and $(-x_3)^6 = (+x_3)^6$ we can assume that all three variables are strictly positive without loss of generalization.

To maximize $f(x)$ is then equivalent to maximizing $\ln(f(x)) = 2\ln(x_1) + 4\ln(x_2) + 6\ln(x_3)$ (which is a concave function!) under the constraint $x_1^2 + x_2^2 + x_3^2 \leq 1$ and the implicit requirement that the variables should be positive.

This is equivalent to *minimizing* $-2\ln(x_1) - 4\ln(x_2) - 6\ln(x_3)$ (which is a convex function!) under the constraint $x_1^2 + x_2^2 + x_3^2 \leq 1$ and the implicit requirement that the variables should be positive.

The KKT-constraints, or Lagrange relaxation, gives that $\mathbf{x} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}\right)^\mathsf{T}$ is optimal.

Hence we have 8 maximum points to $f(x)$ in the unit sphere: $\mathbf{x} = \left(\pm\frac{1}{\sqrt{6}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{2}}\right)^\mathsf{T}$.

13.5 (20051024-nr.4)

The problem can be formulated in the following way, where $\mathbf{I}$ is the identity matrix:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{I}\,\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \\ & \mathbf{I}\,\mathbf{x} \geq \mathbf{0}. \end{aligned}$$

$\hat{\mathbf{x}} \in I\!R^4$ is an optimal solution to this convex QP-problem if and only if there are vectors $\mathbf{u} \in I\!R^2$ and $\mathbf{v} \in I\!R^4$ such that all the following constraints are fulfilled:

(1) $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\mathbf{u} + \mathbf{I}\mathbf{v}$,
(2) $\mathbf{A}\hat{\mathbf{x}} \geq \mathbf{b}$,
(3) $\hat{\mathbf{x}} \geq \mathbf{0}$,
(4) $\mathbf{u} \geq \mathbf{0}$,
(5) $\mathbf{v} \geq \mathbf{0}$,
(6) $u_i(\mathbf{A}\hat{\mathbf{x}} - \mathbf{b})_i = 0$, for $i = 1, 2$,
(7) $v_j\hat{x}_j = 0$, for $j = 1, 2, 3, 4$.

(a)
Here we will assume that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that all $\hat{x}_j > 0$.
Then (7) gives that $\mathbf{v} = \mathbf{0}$, where-after (1) gives that $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\mathbf{u}$. If this is compared with the given assumption that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$, then the following system of equations is obtained $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{b}$, with the unique solution $\mathbf{u} = (2,\ 3)^\mathsf{T}$. But then it must hold that $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\mathbf{u} = (7,\ -1,\ 8,\ -4)^\mathsf{T}$, which contradicts that all $\hat{x}_j > 0$.
Hence there is no optimal solution $\hat{\mathbf{x}}$ that fulfills both that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that all $\hat{x}_j > 0$.

(b)
Now we will assume that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that $\hat{x}_3 = \hat{x}_4 = 0$.
The only solution to this is $\hat{\mathbf{x}} = (20,\ 10,\ 0,\ 0)^\mathsf{T}$.
The constraint (7) then gives that $v_1 = v_2 = 0$, and then the first two constraints in (1) gives that $20 = 2u_1 + u_2$ and $10 = -2u_1 + u_2$, with the unique solution $\mathbf{u} = (2.5,\ 15)^\mathsf{T}$.
The last two constraints in (1) then gives that $0 = u_1 + 2u_2 + v_3$ and $0 = u_1 - 2u_2 + v_4$, i.e. $v_3 = -32.5$ and $v_4 = 27.5$. But $v_3 < 0$ contradicts constraint (5).
Hence there is no optimal solution $\hat{\mathbf{x}}$ that fulfills both that $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and that $\hat{x}_3 = \hat{x}_4 = 0$.

13.6 (20050307-nr.4)

(a) The problem can be written on the form: minimize $\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} \geq \mathbf{b}$,

$$\text{where } \mathbf{H} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix},\ \mathbf{c} = \begin{pmatrix} -12 \\ -8 \\ -4 \end{pmatrix},\ \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix},\ \mathbf{b} = \begin{pmatrix} 0 \\ -2 \\ 0 \\ -2 \\ 0 \\ -2 \end{pmatrix}.$$

The matrix $\mathbf{H}$ is positive definite (can be controlled by for example $\mathbf{LDL}^\mathsf{T}$–factorization), so $\hat{\mathbf{x}}$ is optimal if and only if there is a vector $\mathbf{y} \in I\!R^6$ such that

KKT-1: $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{A}^\mathsf{T}\mathbf{y}$,
KKT-2: $\mathbf{A}\hat{\mathbf{x}} \geq \mathbf{b}$,
KKT-3: $\mathbf{y} \geq \mathbf{0}$,
KKT-4: $y_i(\mathbf{A}\hat{\mathbf{x}} - \mathbf{b})_i = 0, \ i = 1, \ldots, 6$.

With $\hat{\mathbf{x}} = (2,\ 1,\ 0)^\mathsf{T}$, $\mathbf{A}\hat{\mathbf{x}} - \mathbf{b} = (2, 0, 1, 1, 0, 2)^\mathsf{T}$, so KKT-2 is fulfilled,

while KKT-4 gives that $y_1 = y_3 = y_4 = y_6 = 0$.

Since $\mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = (-2,\ 0,\ 2)^\mathsf{T}$, KKT-1 leads to the following system of equations in y_2 and y_5:

$-1y_2 + 0y_5 = -2, \ \ 0y_2 + 0y_5 = 0, \ \ 0y_2 + 1y_5 = 2, \quad (\text{since } y_1 = y_3 = y_4 = y_6 = 0)$,

with the solution $y_2 = y_5 = 2$, which also fulfills KKT-3.

All the KKT-constraints are now fulfilled, and $\hat{\mathbf{x}}$ is hence a global optimal solution.

(b) Now we have a problem on the form: minimize $f_0(\mathbf{x})$ s.t. $f_1(\mathbf{x}) \leq 0$,

with $f_0(\mathbf{x}) = (x_1 + x_2)^2 + (x_2 + x_3)^2 + (x_3 + x_1)^2 - 12x_1 - 8x_2 - 4x_3$

and $f_1(\mathbf{x}) = (x_1 - k_1)^2 + (x_2 - k_2)^2 + (x_3 - k_3)^2 - 1$.

Differentiation gives that

$\nabla f_0(\mathbf{x}) = (\ 4x_1 + 2x_2 + 2x_3 - 12,\ 2x_1 + 4x_2 + 2x_3 - 8,\ \ 2x_1 + 2x_2 + 4x_3 - 4\ )$,
$\nabla f_1(\mathbf{x}) = (\ 2(x_1 - k_1),\ 2(x_2 - k_2),\ 2(x_3 - k_3)\ )$,

$$\nabla^2 f_0(\mathbf{x}) = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}, \quad \nabla^2 f_1(\mathbf{x}) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since both $\nabla^2 f_0(\mathbf{x})$ and $\nabla^2 f_1(\mathbf{x})$ are positive definite for all $\mathbf{x}$, both f_0 and f_1 are convex functions. Furthermore for example $\mathbf{x} = (k_1,\ k_2,\ k_3)^\mathsf{T}$ fulfills that $f_1(\mathbf{x}) < 0$, which means that the problem is regular.

The KKT-conditions are then both necessary and sufficient for $\hat{\mathbf{x}}$ to be a global optimal solution.

These conditions say that there is a scalar y_1 so that

KKT-1: $\nabla f_0(\hat{\mathbf{x}})^\mathsf{T} + y_1 \nabla f_1(\hat{\mathbf{x}})^\mathsf{T} = \mathbf{0}$.
KKT-2: $f_1(\hat{\mathbf{x}}) \leq 0$,
KKT-3: $y_1 \geq 0$,
KKT-4: $y_1 f_1(\hat{\mathbf{x}}) = 0$.

KKT-1 gives that $(-2,\ 0,\ 2) + y_1(\ 2(2 - k_1),\ 2(1 - k_2),\ 2(0 - k_3)\ ) = (0,\ 0,\ 0)$.

We can here directly exclude that $y_1 = 0$. Hence it must hold that $y_1 > 0$ (according to KKT-3).

But then it must hold that $2 - k_1 = 1/y_1$, $\ 1 - k_2 = 0$ and $\ 0 - k_3 = -1/y_1$.

KKT-4 gives that $f_1(\hat{\mathbf{x}}) = 0$ (Since $y_1 > 0$), i.e. $(2 - k_1)^2 + (1 - k_2)^2 + (0 - k_3)^2 = 1$,

which hence can be written as $(1/y_1)^2 + 0^2 + (-1/y_1)^2 = 1$.
Since $y_1 > 0$ this equation has the unique solution $y_1 = \sqrt{2}$,
and $k_1 = 2 - 1/\sqrt{2}$, $k_2 = 1$ and $k_3 = 1/\sqrt{2}$.
With these values of the constants all KKT-conditions will be fulfilled,
and hence $\hat{\mathbf{x}}$ is a global optimal solution.

13.7 (20041016-nr.5)

(a) The Hessian $\nabla^2 f_i(\mathbf{x})$ is a $n \times n$ diagonal matrix with the diagonal elements

$$[\nabla^2 f_i(\mathbf{x})]_{jj} = \frac{2p_{ij}}{(2-x_j)^3} + \frac{2q_{ij}}{(x_j+2)^3}\ ,\ j = 1,\ldots,n.$$

Since all these diagonal elements are > 0 for all $\mathbf{x} \in X$, $\nabla^2 f_i(\mathbf{x})$ is positive definite for all $\mathbf{x} \in X$, and hence the function f_i is strictly convex on X.

(b) That the feasible region in non-empty is equivalent to that there are at least one $\mathbf{x} \in X$ such that $f_i(\mathbf{x}) \leq 0$ for $i = 1, 2$.
But $\mathbf{0} = (0,0)^{\mathsf{T}} \in X$ and $f_i(\mathbf{0}) = \sum_{j=1}^{n} \left(\frac{p_{ij}}{2} + \frac{q_{ij}}{2} \right) + r_i < 0$ according to the prerequisites.
Hence the feasible region contain at least the point $\mathbf{x} = \mathbf{0}$.

(c) Let $S = \{\mathbf{x} \in X \,|\, f_i(\mathbf{x}) \leq 0,\ i = 1,\ldots,m\}$. We will show that S is convex.

Take arbitrary $\mathbf{u} \in S$ and $\mathbf{v} \in S$, and take an arbitrary $t \in (0,1)$.
Set $\mathbf{x} = (1-t)\mathbf{u} + t\mathbf{v}$. We will show that $\mathbf{x} \in S$.
That x_j is between -1 and 1 follows from that x_j is between u_j and v_j which both is between -1 and 1. Further it holds that $f_i(\mathbf{x}) = f_i((1-t)\mathbf{u} + t\mathbf{v}) \leq (1-t)f_i(\mathbf{u}) + tf_i(\mathbf{v}) \leq 0$, where the first inequality follows from that f_i is convex and the second inequality from that $t < 1$, $f_i(\mathbf{u}) \leq 0$, $t > 0$ and $f_i(\mathbf{v}) \leq 0$.

Hence it holds that $\mathbf{x} \in S$, which shows that S is convex.

(d) The KKT-conditions become the following, with the Lagrangian multipliers y_i, ξ_j, η_j, and with the notation

$\mathbf{y} = (y_1,\, y_2)^{\mathsf{T}}$, $p_j(\mathbf{y}) = p_{0j} + p_{1j}y_1 + p_{2j}y_2$ and $q_j(\mathbf{y}) = q_{0j} + q_{1j}y_1 + q_{2j}y_2$.

$$\begin{array}{rll}
\dfrac{p_j(\mathbf{y})}{(2-x_j)^2} - \dfrac{q_j(\mathbf{y})}{(x_j+2)^2} - \xi_j + \eta_j = 0, & j = 1,\ldots,n & (\partial L/\partial x_j = 0) \\
f_i(\mathbf{x}) \leq 0, & i = 1,2 & \text{(primal feasibility)} \\
-1 \leq x_j \leq 1, & j = 1,\ldots,n & \text{(primal feasibility)} \\
y_i \geq 0, & i = 1,2 & \text{(dual feasibility)} \\
\xi_j \geq 0 \ \text{ and } \ \eta_j \geq 0, & j = 1,\ldots,n & \text{(dual feasibility)} \\
y_i f_i(\mathbf{x}) = 0, & i = 1,2 & \text{(complementary slackness)} \\
\xi_j(x_j + 1) = 0, & j = 1,\ldots,n & \text{(complementary slackness)} \\
\eta_j(1 - x_j) = 0, & j = 1,\ldots,n & \text{(complementary slackness)}
\end{array}$$

(e) The Lagrangian function to the problem, with the vector of multipliers $\mathbf{y} = (y_1,\, y_2)^\mathsf{T}$, is given by

$$L(\mathbf{x},\mathbf{y}) = f_0(\mathbf{x})+y_1 f_1(\mathbf{x})+y_2 f_2(\mathbf{x}) = \sum_{j=1}^{n}\left(\frac{p_j(\mathbf{y})}{2-x_j}+\frac{q_j(\mathbf{y})}{x_j+2}\right)+r_0+y_1r_1+y_2r_2,$$

where as above $p_j(\mathbf{y}) = p_{0j} + p_{1j}y_1 + p_{2j}y_2$ and $q_j(\mathbf{y}) = q_{0j} + q_{1j}y_1 + q_{2j}y_2$.

The Lagrange relaxed problem KPR_y consists in, for a given $\mathbf{y} \geq \mathbf{0}$, minimize $L(\mathbf{x},\mathbf{y})$ with respect to $\mathbf{x} \in X$. This falls apart to a problem for each j, minimizing

$\dfrac{p_j(\mathbf{y})}{2-x_j}+\dfrac{q_j(\mathbf{y})}{x_j+2}$ with respect to $x_j \in [-1,\, 1]$.

Since both $p_j(\mathbf{y}) > 0$ and $q_j(\mathbf{y}) > 0$ this is a convex one-variable problem which unique optimal solution $x_j(\mathbf{y})$ is given by the following:

If $p_j(\mathbf{y}) \geq 9\, q_j(\mathbf{y})$, then $x_j(\mathbf{y}) = -1$.

If $q_j(\mathbf{y}) \geq 9\, p_j(\mathbf{y})$, then $x_j(\mathbf{y}) = 1$.

If both $p_j(\mathbf{y}) < 9\, q_j(\mathbf{y})$ and $q_j(\mathbf{y}) < 9\, p_j(\mathbf{y})$ then $x_j(\mathbf{y}) = 2\,\dfrac{\sqrt{q_j(\mathbf{y})}-\sqrt{p_j(\mathbf{y})}}{\sqrt{p_j(\mathbf{y})}+\sqrt{q_j(\mathbf{y})}}$.

The dual objective function is then given by

$$\varphi(\mathbf{y}) = L(\mathbf{x}(\mathbf{y}),\mathbf{y}) = \sum_{j=1}^{n}\left(\frac{p_j(\mathbf{y})}{2-x_j(\mathbf{y})}+\frac{q_j(\mathbf{y})}{x_j(\mathbf{y})+2}\right)+r_0+y_1r_1+y_2r_2.$$

13.8 (20040415-nr.4)

(a)

Let the three given points be denoted (p_i, q_i), $i = 1,2,3$, and let (x,y) be the wanted coordinates for the fourth point. Then the problem can be formulated as:

minimize $f(x,y) = \displaystyle\sum_{i=1}^{3}\sqrt{(x-p_i)^2+(y-q_i)^2}$, without constraints.

A necessary condition for a point (x,y) in which f has continuous derivatives to be optimal is that $\dfrac{\partial f}{\partial x}(x,y) = \dfrac{\partial f}{\partial y}(x,y) = 0$.

Differentiation gives the following, where $r_i(x,y) = \sqrt{(x-p_i)^2+(y-q_i)^2}$, and where we suppose that $(x,y) \neq (p_i,q_i)$ for $i = 1,2,3$.

$$\frac{\partial f}{\partial x}(x,y) = \sum_{i=1}^{3}\frac{x-p_i}{r_i(x,y)} \quad\text{and}\quad \frac{\partial f}{\partial y}(x,y) = \sum_{i=1}^{3}\frac{y-q_i}{r_i(x,y)}\ .$$

Hence $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{-1}{\sqrt{5}}+\dfrac{2}{\sqrt{5}}+\dfrac{-1}{\sqrt{2}} < 0$ and $\dfrac{\partial f}{\partial y}(0,0) = \dfrac{-2}{\sqrt{5}}+\dfrac{1}{\sqrt{5}}+\dfrac{1}{\sqrt{2}} > 0$.

Hence the point $(x,y) = (0,0)$ does not fulfill the mentioned necessary optimality conditions above and is hence *not* an optimal solution.

(b)

The problem can (for example) be formulated as the following in the variables x, y and z, where z denotes the squared distance from (x, y) to the point (p_i, q_i) which is the furthest away from (x, y):

$$\begin{aligned} \text{minimize} \quad & z \\ \text{s.t.} \quad & z \geq (x - p_i)^2 + (y - q_i)^2, \quad i = 1, 2, 3 \end{aligned}$$

or, equivalently,

$$\begin{aligned} \text{minimize} \quad & z \\ \text{s.t.} \quad & -z + (x-1)^2 + (y-2)^2 \leq 0, \\ & -z + (x+2)^2 + (y+1)^2 \leq 0, \\ & -z + (x-1)^2 + (y+1)^2 \leq 0. \end{aligned}$$

The problem is regular and convex, and hence the KKT-conditions are both necessary and sufficient conditions for global optimality.

The Lagrange function (with the Lagrangian multipliers λ_1, λ_2 and λ_3) to this problem is

$$\begin{aligned} L(x, y, z, \lambda_1, \lambda_2, \lambda_3) = z &+ \lambda_1((x-1)^2 + (y-2)^2 - z) + \\ &+ \lambda_2((x+2)^2 + (y+1)^2 - z) + \\ &+ \lambda_3((x-1)^2 + (y+1)^2 - z). \end{aligned}$$

The KKT-conditions then become:

$$\begin{array}{rll} 2\lambda_1(x-1) + 2\lambda_2(x+2) + 2\lambda_3(x-1) = 0 & \partial L/\partial x = 0 & \text{(KKT1)} \\ 2\lambda_1(y-2) + 2\lambda_2(y+1) + 2\lambda_3(y+1) = 0 & \partial L/\partial y = 0 & \text{(KKT2)} \\ 1 - \lambda_1 - \lambda_2 - \lambda_3 = 0 & \partial L/\partial z = 0 & \text{(KKT3)} \\ (x-1)^2 + (y-2)^2 - z \leq 0 & \text{primal feasibility} & \text{(KKT4)} \\ (x+2)^2 + (y+1)^2 - z \leq 0 & \text{primal feasibility} & \text{(KKT5)} \\ (x-1)^2 + (y+1)^2 - z \leq 0 & \text{primal feasibility} & \text{(KKT6)} \\ \lambda_1 \geq 0 & \text{dual feasibility} & \text{(KKT7)} \\ \lambda_2 \geq 0 & \text{dual feasibility} & \text{(KKT8)} \\ \lambda_2 \geq 0 & \text{dual feasibility} & \text{(KKT9)} \\ \lambda_1((x-1)^2 + (y-2)^2 - z) = 0 & \text{complementarity} & \text{(KKT10)} \\ \lambda_2((x+2)^2 + (y+1)^2 - z) = 0 & \text{complementarity} & \text{(KKT11)} \\ \lambda_3((x-1)^2 + (y+1)^2 - z) = 0 & \text{complementarity} & \text{(KKT12)} \end{array}$$

With $x = y = 0$ you obtain the constraints

$$\begin{array}{rl} -2\lambda_1 + 4\lambda_2 - 2\lambda_3 = 0 & \text{(KKT1)} \\ -4\lambda_1 + 2\lambda_2 + 2\lambda_3 = 0 & \text{(KKT2)} \\ 1 - \lambda_1 - \lambda_2 - \lambda_3 = 0 & \text{(KKT3)} \\ 5 - z \leq 0 & \text{(KKT4)} \\ 5 - z \leq 0 & \text{(KKT5)} \\ 2 - z \leq 0 & \text{(KKT6)} \\ \lambda_1 \geq 0 & \text{(KKT7)} \\ \lambda_2 \geq 0 & \text{(KKT8)} \\ \lambda_2 \geq 0 & \text{(KKT9)} \\ \lambda_1(5 - z) = 0 & \text{(KKT10)} \\ \lambda_2(5 - z) = 0 & \text{(KKT11)} \\ \lambda_3(2 - z) = 0 & \text{(KKT12)} \end{array}$$

(KKT1)–(KKT3) are fulfilled if and only if $(\lambda_1, \lambda_2, \lambda_3) = (1/3,\ 1/3,\ 1/3)$.

But then (KKT10) and (KKT12) give that $5 - z = 0$ *and* $2 - z = 0$, which is impossible!

The KKT-conditions can hence not be fulfilled if $x = y = 0$, and hence the point $(0, 0)$ is *not* an optimal location of the fourth component.

14. Mixed examples

14.1 Let

$$\begin{aligned} f(x) &= \max_{t\in T} |f(t) - \sum_{J=1}^{n} x_j f_j(t)| \\ f_L(x) &= \max_{t\in T_m} |f(t) - \sum_{J=1}^{n} x_j f_j(t)| \end{aligned}$$

where $T_m = \{t_1, t_2, \ldots, t_m\}$. Since T_m is a subset of T if holds that $f_L(x) \le f(x)$, therefore (P') is a relaxation of (P). Furthermore let $\hat{x}_L$ be the optimal solution to (P') and p the optimal value to (P).

(a) Using the above notation we have

$$f_L(\hat{x}_L) \le p \le f(\hat{x}_L) = \max_{t\in T} |f(t) - \sum_{j=1}^{n} \hat{x}_{Lj} f_j(t)|$$

(b) (P') can be reformulated as

$$(P'') \quad \begin{array}{ll} \min & x_0 \\ \text{s.t.} & x_0 + \sum_{j=1}^{n} x_j f_j(t_i) \ge f(t_i) \quad i = 1, \ldots, m, \\ & x_0 - \sum_{j=1}^{n} x_j f_j(t_i) \ge -f(t_i) \quad i = 1, \ldots, m, \\ & x_0 \ge 0. \end{array}$$

(c) The dual of the problem formulated above is

$$(D'') \quad \begin{array}{ll} \max & \sum_{i=1}^{m} (\mu_i - \nu_i) f(t_i) \\ \text{s.t.} & \sum_{i=1}^{m} (\mu_i + \nu_i) \le 1 \\ & \sum_{i=1}^{m} (\mu_i - \nu_i) f_j(t_i) = 0 \quad j = 1, \ldots, n \\ & \mu_i,\ \nu_i \ge 0 \qquad i = 1, \ldots, m. \end{array}$$

Since it is not optimal to have both μ_i and $\nu_i > 0$, we can simplify (D'') by introducing the variables $\lambda_i = \mu_i - \nu_i$ with $|\lambda_i| = \mu_i + \nu_i$. We then obtain the following problem

$$(\tilde{D}'') \quad \begin{array}{ll} \max & \sum_{i=1}^{m} \lambda_i f(t_i) \\ \text{s.t.} & \sum_{i=1}^{m} |\lambda_i| \le 1 \\ & \sum_{i=1}^{m} \lambda_i f_j(t_i) = 0 \quad j = 1, \ldots, n \end{array}$$

14.2 Differentiation gives

$$\nabla f(x)^T = \begin{pmatrix} 2x_1 - x_2 - 2 \\ -x_1 + 2x_2 + 4 \\ 2x_3 \end{pmatrix}, \ \nabla g_1(x)^T = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}, \ \nabla g_2(x)^T = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

(a) Differentiation of f a second time gives

$$\nabla^2 f(x) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The symmetric matrix $\nabla^2 f(x)$ has positive eigenvalues (3, 2 and 1). Hence $\nabla^2 f(x)$ is positive definite, independent of x, implying that (P_d) is a convex problem. The solution $\hat{x}$ to the system of equations $\nabla f(\hat{x}) = 0$ is thus a globally optimal solution to (P_d). We have

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 0 \end{pmatrix},$$

with solution $\hat{x} = (0 \ -2 \ 0)$.

(b) The point x^* is feasible to (P_b) with constraints 1 and 2 binding. To fulfil the KT-conditions we must find nonnegative λ_1^* and λ_2^* such that

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \lambda_1^* + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \lambda_2^* = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is fulfilled for $\lambda_1^* = 1$ and $\lambda_2^* = 2$. Hence, the KT-conditions are satisfied at x^*.

(c) For nonnegative λ_1 and λ_2 we obtain the Lagrangean-relaxed problem

$$(P_\lambda) \quad \begin{array}{ll} \min & f_\lambda(x) \\ \text{s.t.} & x \in \mathbb{R}^3, \end{array}$$

where

$$f_\lambda(x) = x_1^2 - x_1 x_2 + x_2^2 + x_3^2 - 2x_1 + 4x_2 + \lambda_1(-x_1 - x_2) + \lambda_2(1 - x_3).$$

Differentiation of f_λ gives

$$\begin{aligned} \nabla f_\lambda(x)^T &= \begin{pmatrix} 2x_1 - x_2 - 2 - \lambda_1 \\ -x_1 + 2x_2 + 4 - \lambda_1 \\ 2x_3 - \lambda_2 \end{pmatrix} \\ \nabla^2 f_\lambda(x) &= \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned}$$

Since all constraints are linear, $\nabla^2 f_\lambda(x)$ is identical to $\nabla^2 f(x)$, independent of λ. Hence, (P_λ) is a convex problem, and the minimizing $x(\lambda)$ is uniquely determined by $\nabla f_\lambda(x(\lambda)) = 0$, i.e.,

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1(\lambda) \\ x_2(\lambda) \\ x_3(\lambda) \end{pmatrix} = \begin{pmatrix} 2+\lambda_1 \\ -4+\lambda_1 \\ \lambda_2 \end{pmatrix}.$$

The solution is

$$x(\lambda) = \begin{pmatrix} \lambda_1 \\ -2+\lambda_1 \\ \dfrac{\lambda_2}{2} \end{pmatrix}$$

and we obtain the dual objective function as

$$\varphi(\lambda) = f_\lambda(x(\lambda)) = \ldots = -\lambda_1^2 + 2\lambda_1 - \frac{\lambda_2^2}{4} + \lambda_2 - 4.$$

Hence, we may write

$$(D_c) \quad \begin{array}{ll} \max & -\lambda_1^2 + 2\lambda_1 - \dfrac{\lambda_2^2}{4} + \lambda_2 - 4 \\ \text{s.t.} & \lambda_1 \geq 0, \\ & \lambda_2 \geq 0. \end{array}$$

(d) Take x^* and λ^* as above, i.e., $x^* = (1 \; –1 \; 1)^T$ and $\lambda^* = (1 \; 2)^T$. Then x^* is feasible to (P_c) and λ^* is feasible to (D_c). It also holds that $f(x^*) = \varphi(\lambda^*) = -2$. It now follows from weak duality that x^* is globally optimal to (P_c) and λ^* is globally optimal to (D_c).

14.3 **(a)** Let

$$\begin{aligned} f(x) &= -2x_1^2 + 12x_1x_2 + 7x_2^2 - 8x_1 - 26x_2, \\ g_1(x) &= x_1 + 2x_2 - 6, \\ g_2(x) &= -x_1, \\ g_3(x) &= x_1 - 3, \\ g_4(x) &= -x_2, \end{aligned}$$

so that the problem is on standard form

$$(P) \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad i = 1, \ldots, 4, \\ & x \in \mathbb{R}^2. \end{array}$$

Differentiation gives

$$\nabla f(x)^T = \begin{pmatrix} -4x_1 + 12x_2 - 8 \\ 12x_1 + 14x_2 - 26 \end{pmatrix}, \; \nabla g_1(x)^T = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \; \nabla g_2(x)^T = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

$$\nabla g_3(x)^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \nabla g_4(x)^T = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

We may now try all combinations of active constraints to find all KT points. The following three combinations of active constraints give the KT points. With no active constraints we get

$$\begin{aligned} -4x_1 + 12x_2 - 8 &= 0, \\ 12x_1 + 14x_2 - 26 &= 0, \end{aligned}$$

yielding the KT point $x^1 = (1\ 1)^T$ with $\lambda^1 = (0\ 0\ 0\ 0)^T$. With constraint 2 active we get

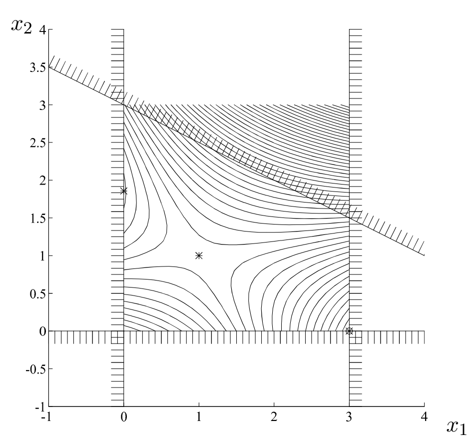
$$\begin{aligned} -4x_1 + 12x_2 - 8 - \lambda_2 &= 0, \\ 12x_1 + 14x_2 - 26 &= 0, \\ -x_1 &= 0, \end{aligned}$$

yielding the KT point $x^2 = (0\ 13/7)^T$ with $\lambda^2 = (0\ 100/7\ 0\ 0)^T$. With constraints 3 and 4 active we get

$$\begin{aligned} -4x_1 + 12x_2 - 8 + \lambda_3 &= 0, \\ 12x_1 + 14x_2 - 26 - \lambda_4 &= 0, \\ x_1 - 3 &= 0, \\ -x_2 &= 0, \end{aligned}$$

yielding a KT point $x^3 = (3\ 0)^T$ with $\lambda^3 = (0\ 0\ 20\ 10)^T$.

(b) Since we are minimizing a continuous function over a closed and bounded set, a global minimizer exists. In addition, linear constraints is a constraint qualification, implying that all local minimizers must satisfy the KT conditions. The global minimizer is thus obtained as the KT point with the minimum objective function value. We have $f(x^1) = -17$, $f(x^2) = -24\frac{1}{7}$, $f(x^3) = -42$, and hence conclude that $x^3 = (3\ 0)^T$ is the global minimizer.



14.4 (20070601-nr.4)

(a)

The gradient is $\nabla f(\mathbf{x}) = \left(\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n}\right)$, where $\dfrac{\partial f}{\partial x_j} = 4x_j^3 - 3x_j^2 + 2x_j - 1$.

The Hessian $\mathbf{F}(\mathbf{x})$ is in this example a diagonal matrix with the diagonal elements
$\dfrac{\partial^2 f}{\partial x_1^2}, \ldots, \dfrac{\partial^2 f}{\partial x_n^2}$, where $\dfrac{\partial^2 f}{\partial x_j^2} = 12x_j^2 - 6x_j + 2$.

f is convex on $I\!R^n$ if and only if $\mathbf{F}(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in I\!R^n$.
A diagonal matrix is positive semidefinite if and only if all diagonal elements are ≥ 0.

But $\dfrac{\partial^2 f}{\partial x_j^2} = 12(x_j^2 - \frac{1}{2}x_j + \frac{1}{6}) = 12((x_j - \frac{1}{4})^2 - \frac{1}{16} + \frac{1}{6}) > 0$ for all values of x_j.

Hence $\mathbf{F}(\mathbf{x})$ is positive definite for all $\mathbf{x} \in I\!R^n$, and hence f is (strictly) convex on $I\!R^n$.

(b)

The Newton direction $\mathbf{d}^{(1)}$ is determined from the equation system $\mathbf{F}(\mathbf{x}^{(1)})\,\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$, given that $\mathbf{F}(\mathbf{x}^{(1)})$ is positive definite, which we already confirmed.

In our case $\mathbf{F}(\mathbf{x}^{(1)})$ is a diagonal matrix, and hence the solution of the equation system is

$$d_j^{(1)} = -\frac{\partial f}{\partial x_j}(x^{(1)}) \Big/ \frac{\partial^2 f}{\partial x_j^2}(x^{(1)}) = -\frac{4(x_j^{(1)})^3 - 3(x_j^{(1)})^2 + 2x_j^{(1)} - 1}{12(x_j^{(1)})^2 - 6x_j^{(1)} + 2}, \; j = 1, \ldots, n.$$

Since $x_j^{(1)} = 1$ for all j, $d_j^{(1)} = -0.25$ for all j.

We try with the step $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = (0.75, \ldots, 0.75)^\mathsf{T}$.

Then $f(\mathbf{x}^{(2)}) = -\dfrac{75n}{256} < 0 = f(\mathbf{x}^{(1)})$, so the step $t_1 = 1$ was fine.

Hence we have made a complete iteration with Newton's method.

14.5 (20070307-nr.5)

(a) We have a QP-problem on the form

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \end{aligned}$$

$$\text{where} \;\; \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \;\; \mathbf{c} = \begin{pmatrix} 4 \\ 2 \\ 6 \end{pmatrix}, \;\; \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \;\text{and}\; \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

First Iteration: In the given starting point the constraints 1, 3 and 4 are fulfilled with equality. Hence we start with $\alpha = (1, 3, 4)$ and $\gamma = (2)$.

Then $\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 5 \\ 2 \\ 6 \end{pmatrix}$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\mathbf{A}_\alpha^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

We obtain the answer "Yes" in Step 1, since $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_\alpha^\mathsf{T}\bar{\mathbf{u}}$ with $\bar{\mathbf{u}} = (5,\ -3,\ 1)^\mathsf{T}$, so we go on to Step 2.

Here we conclude that $\bar{u}_2 < 0$ (and the smallest), and hence $\alpha_2 = 3$ is moved to the γ–vector.

Then we move to Step 3 with $\alpha = (1,4)$, $\gamma = (2,3)$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{A}_\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

In step 3 we are to minimize $\frac{1}{2}\mathbf{d}^\mathsf{T}\mathbf{H}\mathbf{d} + (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})^\mathsf{T}\mathbf{d}$ under the constraint $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$,

The optimality conditions for this convex QP-problem with equality constraints is given by

$\mathbf{H}\mathbf{d} - \mathbf{A}_\alpha^\mathsf{T}\mathbf{u} = -(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ and $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$.

Since $\mathbf{H} = \mathbf{I}$, the first equations give that $\mathbf{d} = \mathbf{A}_\alpha^\mathsf{T}\mathbf{u} - \bar{\mathbf{x}} - \mathbf{c}$,

which inserted in $\mathbf{A}_\alpha\mathbf{d} = \mathbf{0}$ gives the equations system $\mathbf{A}_\alpha\mathbf{A}_\alpha^\mathsf{T}\mathbf{u} = \mathbf{A}_\alpha(\bar{\mathbf{x}} + \mathbf{c})$, i.e.

$\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 6 \end{pmatrix}$, with the solution $\mathbf{u} = \begin{pmatrix} 3.5 \\ 2.5 \end{pmatrix}$, where-after $\hat{\mathbf{d}} = \begin{pmatrix} -1.5 \\ 1.5 \\ 0 \end{pmatrix}$.

Since $\bar{\mathbf{x}} + \hat{\mathbf{d}} = (-0.5,\ 1.5,\ 0)^\mathsf{T}$ does not fulfill all constraints we compute

$\mathbf{s} = \mathbf{A}_\gamma\bar{\mathbf{x}} - \mathbf{b}_\gamma = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{g} = \mathbf{A}_\gamma\hat{\mathbf{d}} = \begin{pmatrix} -1.5 \\ 1.5 \end{pmatrix}$ and $\hat{t} = \min_i \left\{ \dfrac{s_i}{-g_i} \mid g_i < 0 \right\} = \dfrac{1}{1.5} = \dfrac{s_2}{-g_2}$.

Then $\bar{\mathbf{x}}$ is changed to $\bar{\mathbf{x}} + \hat{t} \cdot \hat{\mathbf{d}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \dfrac{1}{1.5} \cdot \begin{pmatrix} -1.5 \\ 1.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, while γ_2 is moved to the α–vector.

New Iteration. Now $\alpha = (1,2,4)$, $\gamma = (3)$. Further

$\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix}$, $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\mathbf{A}_\alpha^\mathsf{T} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

We obtain the answer "Yes" in Step 1, since $\mathbf{H}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{A}_\alpha^\mathsf{T}\bar{\mathbf{u}}$ with $\bar{\mathbf{u}} = (3,\ 1,\ 3)^\mathsf{T}$, so we move on to Step 2.

Here we conclude that $\bar{\mathbf{u}} \geq \mathbf{0}$, which means that the current iteration point is optimal, and therefore the algorithm stops. Hence an optimal solution is $\hat{\mathbf{x}} = (0,1,0)^\mathsf{T}$.

(b)

The problem is a convex QP-problem, and hence the KKT-conditions are both necessary and sufficient conditions for global optimality.

The Lagrange function to the problem can be written as:

$L(\mathbf{x},\mathbf{y}) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2 + 4x_1 + 2x_2 + 6x_3 + y_1(1 - x_1 - x_2 - x_3) - y_2x_1 - y_3x_2 - y_4x_3$.

The KKT-constraints then becomes:

$$
\begin{array}{rll}
x_1 - y_1 - y_2 = -4 & \partial L/\partial x_1 = 0 & \text{(KKT1)} \\
x_2 - y_1 - y_3 = -2 & \partial L/\partial x_2 = 0 & \text{(KKT2)} \\
x_3 - y_1 - y_4 = -6 & \partial L/\partial x_3 = 0 & \text{(KKT3)} \\
x_1 + x_2 + x_3 \geq 1 & \text{primal feasibility} & \text{(KKT4)} \\
x_1 \geq 0 & \text{primal feasibility} & \text{(KKT5)} \\
x_2 \geq 0 & \text{primal feasibility} & \text{(KKT6)} \\
x_3 \geq 0 & \text{primal feasibility} & \text{(KKT7)} \\
y_1 \geq 0 & \text{dual feasibility} & \text{(KKT8)} \\
y_2 \geq 0 & \text{dual feasibility} & \text{(KKT9)} \\
y_3 \geq 0 & \text{dual feasibility} & \text{(KKT10)} \\
y_4 \geq 0 & \text{dual feasibility} & \text{(KKT11)} \\
y_1(1 - x_1 - x_2 - x_3) = 0 & \text{complementarity} & \text{(KKT12)} \\
y_2x_1 = 0 & \text{complementarity} & \text{(KKT13)} \\
y_3x_2 = 0 & \text{complementarity} & \text{(KKT14)} \\
y_4x_3 = 0 & \text{complementarity} & \text{(KKT15)}
\end{array}
$$

First suppose that $\mathbf{x} = (1,0,0)^\mathsf{T}$. Then (KKT13) gives that $y_2 = 0$, and then (KKT1)-(KKT3) gives that $y_1 = 5$, $y_3 = -3$ and $y_4 = 1$. But this contradicts (KKT10).
Hence the KKT-constraints can not be fulfilled with $\mathbf{x} = (1,0,0)^\mathsf{T}$.

Now suppose that $\mathbf{x} = \hat{\mathbf{x}} = (0,1,0)^\mathsf{T}$. Then (KKT13) gives that $y_3 = 0$, and then (KKT1)-(KKT3) gives that $y_1 = 3$, $y_2 = 1$ and $y_4 = 3$. A quick check gives that now all the KKT-conditions are fulfilled. The KKT-conditions are hence fulfilled by $\hat{\mathbf{x}} = (0,1,0)^\mathsf{T}$ and $\hat{\mathbf{y}} = (3,1,0,3)^\mathsf{T}$.

(c)

Now we consider the problem to minimize $f(\mathbf{x})$ s.t. $g_1(\mathbf{x}) \leq 0$ and $\mathbf{x} \in X$, where $g_1(\mathbf{x}) = 1 - x_1 - x_2 - x_3$ and $X = \{\mathbf{x} \in I\!R^3 \mid \mathbf{x} \geq \mathbf{0}\}$.

The Lagrangian function to the problem can now be written as:

$$
\begin{aligned}
L(\mathbf{x}, y_1) &= \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 + \tfrac{1}{2}x_3^2 + 4x_1 + 2x_2 + 6x_3 + y_1(1 - x_1 - x_2 - x_3) = \\
&= y_1 + \tfrac{1}{2}x_1^2 + (4 - y_1)x_1 + \tfrac{1}{2}x_2^2 + (2 - y_1)x_2 + \tfrac{1}{2}x_3^2 + (6 - y_1)x_3.
\end{aligned}
$$

To obtain the dual objective function $\varphi(y_1)$ you should minimize $L(\mathbf{x}, y_1)$ with respect to $\mathbf{x} \in X$. As can be seen, this minimization can be performed on each variable x_j at a time.

Note that if the expression $\frac{1}{2}x_j^2 + (c_j - y_1)x_j$ is to be minimized under the condition that $x_j \geq 0$, the minimizing value of x_j is given by:

$x_j(y_1) = y_1 - c_j$ if $y_1 > c_j$ and $x_j(y_1) = 0$ if $y_1 \leq c_j$.

This can be written short as $x_j(y_1) = \max\{\,0\,,\; y_1 - c_j\,\} = (y_1 - c_j)_+$

where the last equality is the definition of $(y_1 - c_j)_+$.

The dual objective function is now given by

$\varphi(y_1) = \min_{\mathbf{x} \in X} L(\mathbf{x}, y_1) = L(\mathbf{x}(y_1), y_1) = y_1 +$

$+\frac{1}{2}(y_1-4)_+^2 + (4-y_1)(y_1-4)_+ + \frac{1}{2}(y_1-2)_+^2 + (2-y_1)(y_1-2)_+ + \frac{1}{2}(y_1 - 6)_+^2 + (6-y_1)(y_1-6)_+ =$

$= y_1 - \frac{1}{2}(y_1-4)_+^2 - \frac{1}{2}(y_1-2)_+^2 - \frac{1}{2}(y_1-6)_+^2$.

From (a) and (b) we guess that $\hat{y}_1 = 3$.

Insertion above gives that $\varphi(3) = 3 - \frac{1}{2}(3-4)_+^2 - \frac{1}{2}(3-2)_+^2 - \frac{1}{2}(3-6)_+^2 = 2.5$.

According to (a) and (b) $\hat{\mathbf{x}} = (0,1,0)^\mathsf{T}$ with $f(\hat{\mathbf{x}}) = 2.5$, i.e. $f(\hat{\mathbf{x}}) = \varphi(\hat{y}_1)$.

This shows that the guess $\hat{y}_1 = 3$ was correct.

14.6 (20050331-nr.5)

The LP-problem that is obtained after linearization of all functions is

$$\begin{aligned} \text{minimize} \quad & f(\hat{\mathbf{x}}) + \nabla f(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) \\ \text{s.t.} \quad & g_i(\hat{\mathbf{x}}) + \nabla g_i(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) \leq 0\,, \quad i = 1, \ldots, m \\ & \mathbf{x} \in I\!\!R^n. \end{aligned}$$

Let $\mathbf{c}^\mathsf{T} = \nabla f(\hat{\mathbf{x}})$, let $\mathbf{A}$ be a $m{\times}n$ matrix with the rows $-\nabla g_i(\hat{\mathbf{x}}),\ i = 1, \ldots, m$, and let the vector $\mathbf{b} \in I\!\!R^m$ have the components $b_i = g_i(\hat{\mathbf{x}}) - \nabla g_i(\hat{\mathbf{x}})\,\hat{\mathbf{x}}$.

Then the LP-problem above can be written as

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}\,, \\ & \mathbf{x} \in I\!\!R^n. \end{aligned}$$

(Except from an in the context not important constant in the objective function that doesn't affect the optimal solution to the LP-problem.)

The corresponding dual LP-problem can be written as

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{c}\,, \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned}$$

According to the duality theorem it holds that $\hat{\mathbf{x}} \in I\!\!R^n$ is an optimal solution to the primal LP-problem if and only if there is a (dual) vector $\hat{\mathbf{y}} \in I\!\!R^m$ such that $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ together fulfill the following four optimality conditions:

1. $\mathbf{A}^\mathsf{T}\hat{\mathbf{y}} = \mathbf{c}$.

2. $\mathbf{A}\hat{\mathbf{x}} \geq \mathbf{b}$.

3. $\hat{\mathbf{y}} \geq \mathbf{0}$.

4. $\mathbf{c}^{\mathsf{T}}\hat{\mathbf{x}} = \mathbf{b}^{\mathsf{T}}\hat{\mathbf{y}}$.

If you put in what $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$ really stand for, you obtain that these four optimality conditions can be equivalently written as

1. $-\sum_i \hat{y}_i \nabla g_i(\hat{\mathbf{x}})^{\mathsf{T}} = \nabla f(\hat{\mathbf{x}})^{\mathsf{T}}$.

2. $-\nabla g_i(\hat{\mathbf{x}})\,\hat{\mathbf{x}} \geq g_i(\hat{\mathbf{x}}) - \nabla g_i(\hat{\mathbf{x}})\,\hat{\mathbf{x}}$, for $i = 1, \ldots, m$.

3. $\hat{\mathbf{y}} \geq \mathbf{0}$.

4. $\nabla f(\hat{\mathbf{x}})\,\hat{\mathbf{x}} = \sum_i \hat{y}_i (g_i(\hat{\mathbf{x}}) - \nabla g_i(\hat{\mathbf{x}})\,\hat{\mathbf{x}})$.

Some re-formulations give the following equivalent formulations of these constraints.

1. $\nabla f(\hat{\mathbf{x}}) + \sum_i \hat{y}_i \nabla g_i(\hat{\mathbf{x}}) = \mathbf{0}^{\mathsf{T}}$.

2. $g_i(\hat{\mathbf{x}}) \leq 0$, for $i = 1, \ldots, m$.

3. $\hat{y}_i \geq 0$, for $i = 1, \ldots, m$.

4. $\sum_i \hat{y}_i g_i(\hat{\mathbf{x}}) = 0$.

Using 2 and 3 gives that 4 is equivalent to

4. $\hat{y}_i g_i(\hat{\mathbf{x}}) = 0$, for $i = 1, \ldots, m$.

But now the conditions 1-4 are the well-known KKT-conditions for the NLP-problem!

Since our original non-linear optimization problem NLP, according to the prerequisites, is a regular convex problem with continuously differentiable functions, $\hat{\mathbf{x}}$ is an optimal solution to this NLP if and only if there is a vector $\hat{\mathbf{y}} \in I\!R^m$ that together with $\hat{\mathbf{x}}$ fulfills the KKT-conditions above.

But according to above this is equivalent to that $\hat{\mathbf{x}}$ is an optimal solution to the LP-problem which you obtain if you in P replace all functions with their linearizations (computed in $\hat{\mathbf{x}}$).

14.7 (20041016-nr.4)

(a). The problem can be written as:

minimize $f(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} - (\mathbf{A}^{\mathsf{T}}\mathbf{b})^{\mathsf{T}}\mathbf{x} + \frac{1}{2}\,\mathbf{b}^{\mathsf{T}}\mathbf{b}$, where $\mathbf{A}^{\mathsf{T}}\mathbf{A} = \begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix}$ and $\mathbf{A}^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}$.

Since $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ is positive definite, f is a (strictly) convex quadratic function. Therefore $\mathbf{x}$ is a global minpoint to f if and only if $\nabla f(\mathbf{x})^{\mathsf{T}} = \mathbf{0}$, i.e. if and only if $\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{b} = \mathbf{0}$. (The Normal Equations.)

The unique solution to this simple system of equations is $\mathbf{x} = \left(\dfrac{12}{6}\,,\,\dfrac{10}{6}\right)^{\mathsf{T}}$.

(b). The problem can be written on the following form:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = \tfrac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} - (\mathbf{A}^{\mathsf{T}}\mathbf{b})^{\mathsf{T}}\mathbf{x} + \tfrac{1}{2}\,\mathbf{b}^{\mathsf{T}}\mathbf{b} \\ \text{s.t.} \quad & \mathbf{A}\,\mathbf{x} \geq \mathbf{b}. \end{aligned}$$

This is a convex QP-problem with linear inequality constraints.
$\mathbf{x} \in I\!\!R^2$ is a global minpoint to this problem if and only if there is a vector $\mathbf{y} \in I\!\!R^3$ which together with $\mathbf{x}$ fulfills the following four optimality conditions:

$\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{b} = \mathbf{A}^\mathsf{T}\mathbf{y}, \quad \mathbf{A}\mathbf{x} \geq \mathbf{b}, \quad \mathbf{y} \geq \mathbf{0}$ and $\mathbf{y}^\mathsf{T}(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0.$

Insertion of the proposed point $\mathbf{x} = \left(\frac{13}{6}\,,\ \frac{11}{6}\right)^\mathsf{T}$ gives that $\mathbf{A}\mathbf{x} - \mathbf{b} = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$.

The second optimality condition is hence fulfilled and the fourth gives that $y_1 = y_2 = 0$.
The first optimality condition can then be written as:

$$\mathbf{0} = \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{b} - \mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{A}^\mathsf{T}(\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{y}) = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ -y_3 \end{pmatrix} = \begin{pmatrix} 1/2 - y_3 \\ 1/2 - y_3 \end{pmatrix},$$

with the solution $y_3 = 1/2$, which is ≥ 0.

All optimality conditions are hence fulfilled with $\mathbf{x} = \left(\frac{13}{6}\,,\ \frac{11}{6}\right)^\mathsf{T}$ and $\mathbf{y} = \left(0\,,\ 0\,,\ \frac{1}{2}\right)^\mathsf{T}$.

14.8 (20040415-nr.5)

(a)

Let $\mathbf{x}$ be an arbitrary feasible solution to NLP, i.e. $g(\mathbf{x}) \leq 0$, and let $z = f(\mathbf{x})$.

That f and g are convex and continuously differentiable implies that

$f(\mathbf{x}) \geq f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k)$ and $g(\mathbf{x}) \geq g(\mathbf{x}^k) + \nabla g(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k)$, for $k = 1, \ldots, K$.

Since $z = f(\mathbf{x})$ and $g(\mathbf{x}) \leq 0$ this gives that

$z \geq f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k)$ and $0 \geq g(\mathbf{x}^k) + \nabla g(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k)$, for $k = 1, \ldots, K$,

which equivalently can be written as

$z - \nabla f(\mathbf{x}^k)\,\mathbf{x} \;\geq\; f(\mathbf{x}^k) - \nabla f(\mathbf{x}^k)\,\mathbf{x}^k$ and $-\nabla g(\mathbf{x}^k)\,\mathbf{x} \;\geq\; g(\mathbf{x}^k) - \nabla g(\mathbf{x}^k)\,\mathbf{x}^k$, for $k = 1, \ldots, K$.

Hence $(\mathbf{x}, z)$ is a *feasible* solution to the problem LP.

But since $(\hat{\mathbf{x}}, \hat{z})$ is an *optimal* solution to LP it follows that $\hat{z} \leq z$, i.e. $f(\mathbf{x}) \geq \hat{z}$.

(b)

Suppose that for example $(\mathbf{x}^1, \hat{z})$, where $\mathbf{x}^1$ is the first of the given points $\mathbf{x}^k$, is an optimal solution to LP. We will show that $\mathbf{x}^1$ then is an optimal solution to NLP.

That $(\mathbf{x}^1, \hat{z})$ is an optimal solution to LP implies that $(\mathbf{x}^1, \hat{z})$ is a feasible solution to LP, i.e.

$\hat{z} - \nabla f(\mathbf{x}^k)\,\mathbf{x}^1 \;\geq\; f(\mathbf{x}^k) - \nabla f(\mathbf{x}^k)\,\mathbf{x}^k$ and $-\nabla g(\mathbf{x}^k)\,\mathbf{x}^1 \;\geq\; g(\mathbf{x}^k) - \nabla g(\mathbf{x}^k)\,\mathbf{x}^k$, for $k = 1, \ldots, K$.

For $k = 1$ this gives that

$\hat{z} - \nabla f(\mathbf{x}^1)\,\mathbf{x}^1 \;\geq\; f(\mathbf{x}^1) - \nabla f(\mathbf{x}^1)\,\mathbf{x}^1$ and $-\nabla g(\mathbf{x}^1)\,\mathbf{x}^1 \;\geq\; g(\mathbf{x}^1) - \nabla g(\mathbf{x}^1)\,\mathbf{x}^1$,

i.e. $f(\mathbf{x}^1) \leq \hat{z}$ and $g(\mathbf{x}^1) \leq 0$.

$\mathbf{x}^1$ is hence a feasible solution to NLP with $f(\mathbf{x}^1) \leq \hat{z}$.

But according to the (a)-task above $f(\mathbf{x}) \geq \hat{z}$ for every feasible solution $\mathbf{x}$ to NLP.

Hence $f(\mathbf{x}^1) \leq f(\mathbf{x})$ for every feasible solution $\mathbf{x}$ to NLP, which implies that $\mathbf{x}^1$ is an optimal solution to NLP.

(Further is follows that $f(\mathbf{x}^1) = \hat{z}$, which implies that the optimal values to NLP and LP are equal in this specific case.)