Exam April 5, 2018 in SF1811 Optimization.

Examiner: Johan Karlsson, tel. 790 84 40.

Allowed utensils: Pen, paper, eraser and ruler. A formula-sheet is handed out. No calculator! No books or notes.

Language: Your solutions should be written in English or in Swedish.

Solution methods: All conclusions should be properly motivated. Unless otherwise stated in the problem statement, the problems should be solved using systematic methods that do not become unrealistic for large problems. Unless otherwise stated in the problem statement, known theorems can be used without proving them, as long as they are formulated correctly. Motivate all your conclusions carefully.

Note: Your personal number must be stated on the cover sheet. Number your pages and write your name on each sheet that you turn in!

Preliminary grades (Credit = exam credit + bonus from homeworks): 23-24 credits give grade Fx (contact examiner asap for further info), 25-27 credits give grade E, 28-32 credits give grade D, 33-38 credits give grade C, 39-44 credits give grade B, and 45 or more credits give grade A.

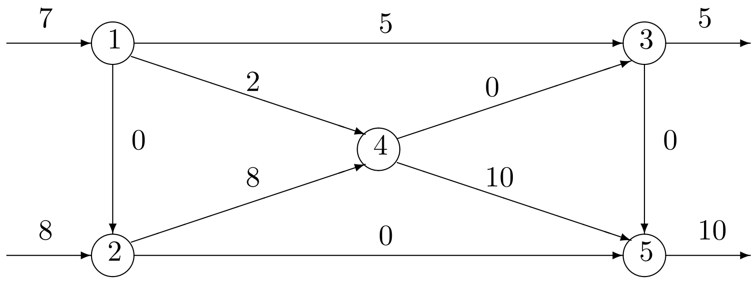
1. (a) Your first assignment is to design the working schedule for the pizza restaurant "Il Ristorante". At Il Ristorante the four employees Adrian, Bobby, Cecile and Davide work and to keep the restaurant going it is necessary to every week spend 40 manhours in the kitchen, 70 manhours at register selling the hamburgers and 40 manhours cleaning tables. The employees work at most 40 hours per week each, and they get payed only for the hours (or fraction of an hour) they work. At the last salary negotiations it was decided that Adrian gets 70 SEK/hour, Bobby 80 SEK/hour, Cecile 100 SEK/hour and Davide 110 SEK/hour.

   When designing the work schedule the following factors also has to be taken into account. If Adrian has to work with cleaning more hours than Bobby or Cecile, he is going to get very upset and this has to avoided. When Bobby is working at the register 30 SEK/hour usually "Disappears". Cecile is a terrible cook and is not allowed to work in the kitchen, while Davide has a contract that he will work at least 30 hours per week in the kitchen.

   You have to formulate the problem to find a working schedule that minimizes the cost for the restaurant while all the criteria mentioned above are satisfied.

   The answer should be a linear optimization problem with linear equality and/or inequality constraints, and the variables you use should be defined carefully. .... (5p)

(b) Consider the following network



Show that the flow suggested in the graph minimizes the cost of the flow if the costs in the links are given by

$c_{12} = 2, c_{13} = 2, c_{14} = 2, c_{24} = 1, c_{25} = 3, c_{43} = 1, c_{35} = 2, c_{45} = 1.$

.... (3p)

(c) Is the same solution still optimal if the cost c_{25} is changed to $c_{25} = 1$? Motivate your answer.
.... (2p)

2. (a) Consider the following linear programing problem

$$(P) \quad \left[\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \right]$$

where

$$A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 2 & 2 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad c = \begin{bmatrix} 2 & 1 & 2 & 1 \end{bmatrix}^T.$$

Show that the basic solution with x_1 and x_2 being basic variables is feasible. Use this basic solution as a starting point for solving the linear program (P) using the simplex method. ....(5p)

(b) Determine the dual linear program (D) to (P). .... (2p)

(c) Determine the optimal solution to the dual (D). It is ok to use the result from (a), or a graphic solution of the dual, to come up with a candidate optimal solution as long as optimality is verified using some known theorem. ....(3p)

3. First recall that $\mathcal{N}(A)$ and $\mathcal{R}(A)$ denotes the *nullspace* and the *range space* (also called the *column space*) of the matrix A.

   In this exercise $A = \begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ and $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$,

   where b_1, b_2, c_1 and c_2 are given numbers.

   (a) Determine the point among all points in $\mathcal{R}(A)$ which has the shortest distance to the given point b. The answer may include b_1 and b_2. .... (2p)

   (b) Determine the point among all points in $\mathcal{N}(A)^T$ which has the shortest distance to the given point b. The answer may include b_1 and b_2. .... (2p)

   (c) What are the relations between the solution in a), the solution in b), and the vector b? .... (2p)

   (d) Determine the point among all points in $\mathcal{R}(A)^T$ which has the shortest distance to the given point c. The answer may include c_1 and c_2. .... (2p)

   (e) Determine the point among all points in $\mathcal{N}(A)$ which has the shortest distance to the given point c. The answer may include c_1 and c_2. .... (2p)

4. Consider the following non-linear least-square problem in the variable vector $x \in \Re^2$:

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

   $$\begin{aligned} h_1(x) &= (x_1+1)^2 + x_2^2 - 3, \\ h_2(x) &= (x_1-1)^2 + x_2^2 + 2, \\ h_3(x) &= x_1^2 + (x_2-1)^2 - 4, \\ h_4(x) &= x_1^2 + (x_2+1)^2 + 1. \end{aligned}$$

   (a) Perform one iteration with Gauss-Newtons method starting from $x^{(1)} = (0,0)^T$.
   Check that the point $x^{(2)}$ you obtain satisfies $f(x^{(2)}) < f(x^{(1)})$. (6p)

   (b) An alternative would of course be to use Newton's method to minimize $f(x)$.
   Show that for this particular problem, and from this specific starting point $x^{(1)}$, the search direction is the same for Newton's method as for Gauss-Newton's method. .... (2p)

   (c) Is the objective function $f(x)$ a convex function on the whole $\Re^2$? Motivate your answer. .... (2p)

5. *Maximum Likelihood estimation of a probability distribution*

In statistics it is common to consider outcomes that can be classified into m different categories. We denote the probability to observe an outcome of category i with p_i, and we would like to use a number of independent observations to estimate these probabilities. Assume that we have obtained k_i observations of each category $i = 1, \cdots, m$. Assume that $p_1, p_2, \cdots, p_m$ are known, then

$$L(p) = \prod_{i=1}^{m} p_i^{k_i}$$

is the probability to obtain the ordered observation. The probabilities can then be estimated using the Maximum Likelihood method, which is based on the idea to choose the probabilities in such a way that the likelihood of obtaining the observed outcomes is maximized. It is often easier to consider the logarithm of the likelihood, so we consider the following problem

$$\begin{aligned} \text{maximize} \quad & \textstyle\sum_{i=1}^{m} k_i \log p_i \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{m} p_i \leq 1, \\ & p_i \geq 0, \quad i = 1, \cdots, m. \end{aligned}$$

The constraints say that the probabilities have to be non-negative and the sum of all the probabilities is less or equal to one. The last condition might seem strange, the natural thing would be to require equality constraint, but this problem is a relaxed version which will generate the same solution.

(a) Determine the dual problem. When you determine the dual, only the constraint on the sum of the probabilities should be relaxed, the positivity constraints should be kept implicit. .... (4p)

(b) Determine a stationary point to the dual problem. .... (2p)

(c) Use the solution in (b) to the dual problem to determine a solution to the primal problem .... (2p)

(d) Motivate global optimality for the primal and dual solutions. (2p)

Good luck!

1 Solutions

1. (a) Define variables x_A, x_B, x_C, x_D for how many manhours the respective persons are spending in the kitchen, y_A, y_B, y_C, y_D for

how many manhours they are working at the register, and z_A, z_B, z_C, z_D for how many manhours they are cleaning.
The cost for the restaurant is given by the objective function

$$f = (x_A + y_A + z_A) * 70 + (x_B + y_B + z_B) * 80 + (x_C + y_C + z_C) * 100 \\ + (x_D + y_D + z_D) * 110 + y_B * 40$$

The last term is the cost of Bobbys money-problems.
All the work that has to be done:

$$x_A + x_B + x_C + x_D = 40,$$

$$y_A + y_B + y_C + y_D = 70,$$

$$z_A + z_B + z_C + z_D = 40.$$

All working times has to be non-negative:

$$x_A \geq 0, x_B \geq 0, x_C \geq 0, x_D \geq 0$$

$$y_A \geq 0, y_B \geq 0, y_C \geq 0, y_D \geq 0$$

$$z_A \geq 0, z_B \geq 0, z_C \geq 0, z_D \geq 0$$

Furthermore, $x_C = 0$ and $x_D \geq 30$ are the constraints regulating the work of Cecile and Davide in the kitchen. These constraints make the constraints $x_C \geq 0$ and $x_D \geq 0$ above redundant, but the formulation as a linear program is OK with or without the redundant constraints. None are allowed to work more than 40 manhours per week:

$$x_A + y_A + z_A \leq 40, \quad x_B + y_B + z_B \leq 40,$$

$$x_C + y_C + z_C \leq 40, \quad x_D + y_D + z_D \leq 40.$$

Finally, to make sure that Adrian is not upset we require that $z_A \leq z_B$ and $z_A \leq z_C$.

(b) Setting the node potential $v_5 = 0$, and using that $c_{ij} = v_i - v_j$ for the i, j corresponding to basic variables, then the rest of the node potentials are given by

$$v_1 = 3, v_2 = 2, v_3 = 1, v_4 = 1.$$

Then the reduced costs for the other indices are given by $r_{ij} = c_{ij} - v_i + v_j$:

$$r_{12} = 2{-}3{+}2 = 1, r_{25} = 3{-}2{+}0 = 1, r_{34} = 1{-}1{+}1 = 1, r_{35} = 2{-}1{+}0 = 1.$$

Since all the reduced costs are non-negative the proposed solution is optimal.

(c) If $c_{25} = 1$, then the node potentials determined above are unchanged. The only thing that changes is the reduced cost r_{25} which is now given by $r_{25} = 1 - 2 + 0 = -1$. Since this reduced cost is negative, the value of the objective function decreases if x_{25} becomes a basic variable and increases from zero. The basic solution was non-degenarate so this implies that x_{25} can be increased and then the same solution can not be optimal.

2. a) Let $\beta = (1, 2)$ and $\nu = (3, 4)$. This gives

$$A_\beta = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad c_\beta = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A_\nu = \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix}, \quad c_\nu = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Hence we have

$$A_\beta \bar{b} = b = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \Rightarrow \bar{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which is a basic feasible solution. Next, we have

$$A_\beta^T y = c_\beta \Rightarrow y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore

$$r_\nu = c_\nu - A_\nu^T y = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}.$$

Since the basic feasible solution $\bar{b}$ is strictly positive and the second component of the reduced cost is negative the solution is not optimal. Therefore we introduce the variable $\nu_2 = 4$ as new active variable. Noting that $\bar{a}_4 = (4/3, 1/3)^T$ we get $t_{\max} = 3/4$ which gives $x_{\beta_1} = 0$ and hence the variable x_1 can be removed from the basic tuple.

Next we will do a second simplex iteration. This time for the basic tuple $\beta = (2, 4)$:

$$A_\beta = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}, \quad c_\beta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A_\nu = \begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix}, \quad c_\nu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Hence we have

$$A_\beta \bar{b} = b = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \Rightarrow \bar{b} = \begin{pmatrix} 3/4 \\ 3/4 \end{pmatrix},$$

which is a basic feasible solution. Next, we have

$$A_\beta^T y = c_\beta \Rightarrow y = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}.$$

Therefore

$$r_\nu = c_\nu - A_\nu^T y = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Since the the reduced cost is positive the solution $x = (0, 3/4, 0, 3/4)^T$ is the optimal solution.

b) The dual problem is given by

$$(D): \begin{cases} \max & 3y_1 + 3y_2 \\ \text{subject to} & 2y_1 + y_2 \leq 2 \\ & y_1 + 2y_2 \leq 1 \\ & 2y_1 + 2y_2 \leq 2 \\ & 3y_1 + 2y_2 \leq 1. \end{cases}$$

By complementarity $x^T(c - A^T y) = 0$, hence the second and forth component of $c - A^T y$ must be zero, i.e., $y_1 + 2y_2 = 1$ and $3y_1 + 2y_2 = 1$. This gives the optimal solution $y_1 = 0$, $y_2 = 1/2$. Also note that this is a feasible solution to the dual problem and that the corresponding values of the objective function is $3/2$ for both the primal and dual problem.

3. Note that the smallest norm problem

$$\min_{x=Zv} \|x - c\|^2 = \min_v \frac{1}{2} v^T (2Z^T Z) v - (2Z^T c)^T v + c^T c$$

has the uniques ($Z^T Z$ is positive definite) solution given by

$$(2Z^T Z)v - 2Z^T c = 0 \Rightarrow v = (Z^T Z)^{-1} Z^T c \Rightarrow x = Z(Z^T Z)^{-1} Z^T c$$

(a) $R(A) = \text{span} \begin{pmatrix} 1 \\ -3 \end{pmatrix}$, which gives closest solution

$$\hat{x}_{(\mathrm{a})} = \begin{pmatrix} 1 \\ -3 \end{pmatrix} (b_1 - 3b_2)/10.$$

(b) $N(A^T) = \text{span} \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, which gives closest solution

$$\hat{x}_{(\mathrm{b})} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} (3b_1 + b_2)/10.$$

(c) Note that $\hat{x}_{(\mathrm{a})}^T \hat{x}_{(\mathrm{b})} = 0$ and that $\hat{x}_{(\mathrm{a})} + \hat{x}_{(\mathrm{b})} = b$.

(d) $R(A^T) = \text{span}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$, which gives closest solution

$$\hat{x}_{(\mathrm{d})} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} (c_1 - 2c_2)/5.$$

(e) $N(A) = \text{span}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, which gives closest solution

$$\hat{x}_{(\mathrm{e})} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} (2c_1 + c_2)/5.$$

4. (a) Note that

$$h(x) = \begin{pmatrix} (x_1+1)^2 + x_2^2 - 3 \\ (x_1-1)^2 + x_2^2 + 2 \\ x_1^2 + (x_2-1)^2 - 4 \\ x_1^2 + (x_2+1)^2 + 1 \end{pmatrix}, \quad h(0) = \begin{pmatrix} -2 \\ 3 \\ -3 \\ 2 \end{pmatrix},$$

and

$$\nabla h(x) = \begin{pmatrix} 2(x_1+1) & 2x_2 \\ 2(x_1-1) & 2x_2 \\ 2x_1 & 2(x_2-1) \\ 2x_1 & 2(x_2+1) \end{pmatrix}, \quad \nabla h(x) = \begin{pmatrix} 2 & 0 \\ -2 & 0 \\ 0 & -2 \\ 0 & 2 \end{pmatrix}.$$

The Gauss-Newton search direction is given by

$$\nabla h(x^{(k)})^T \nabla h(x^{(k)}) d = -\nabla h(x^{(k)})^T h(x^{(k)}).$$

For $k = 0$ and $x^{(0)} = 0$, this yields

$$\begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix} d = -\begin{pmatrix} -10 \\ 10 \end{pmatrix} \Rightarrow d^{(0)} = \begin{pmatrix} 5/4 \\ -5/4 \end{pmatrix}.$$

Note that $f(x^{(1)} + d^{(0)}) > f(x^{(0)})$, hence we need to select $t_0 < 1$ in $x^{(1)} = x^{(0)} + t_0 d^{(0)}$. For example $t_0 = 1/2$ gives $f(x^{(1)}) < f(x^{(0)})$.

(b) The Newton iteration is given by

$$F(x^{(k)})d = -\nabla f(x^{(k)})^T.$$

For our case we have

$$F(x^{(0)}) = \nabla h(x^{(0)})^T \nabla h(x^{(0)}) + \sum_i h_i(x^{(0)}) H_i(x^{(0)}) = \nabla h(x^{(0)})^T \nabla h(x^{(0)}),$$

$$\nabla f(x^{(0)}) = h(x^{(k)})^T \nabla h(x^{(0)}).$$

since $H_i(x^{(0)}) = 0$ for all i. Therefore the Newton iteration is identical to the Gauss-Newton in the first iteration.

(c) The function f is convex since the Hessian $F(x)$ satisfies

$$\begin{aligned} F(x) &= \nabla h(x)^T \nabla h(x) + \sum_i h_i(x) H_i(x) \\ &= \nabla h(x)^T \nabla h(x) + 8x^T x I \geq 0, \end{aligned}$$

and is therefore positive definite for all x.