KTH Mathematics

Solutions for the exam in Optimization.
Wednesday May 29, 2013, time. 8.00–13.00

Instructor: Per Enqvist, tel. 790 62 98

There may be alternative solutions to the problem.

---

1. (a) The NW solution is given by

x_{ij}	$j=1$	$j=2$	$j=3$	s_i
$i=1$	4	0	0	4
$i=2$	3	1	0	4
$i=3$	0	7	5	12
d_j	7	8	5	

The basic variables are $x = (x_{11}, x_{21}, x_{22}, x_{32}, x_{33}) = (4, 3, 1, 7, 5)$, which corresponds to a spanning tree in the graph.

Put the node potential v_3 at node sink 3 to be 0.
Then $u_3 - v_3 = c_{33} = 2$ gives $u_3 = 2$.
Then $u_3 - v_2 = c_{32} = 1$ gives $v_2 = 1$.
Then $u_2 - v_2 = c_{33} = 2$ gives $u_2 = 3$.
Then $u_2 - v_1 = c_{21} = 1$ gives $v_1 = 2$.
Then $u_1 - v_1 = c_{11} = 2$ gives $u_1 = 4$.

c_{ij}, r_{ij}	$j=1$	$j=2$	$j=3$	u_i
$i=1$	2	0	-2	4
$i=2$	1	2	-1	3
$i=3$	2	1	2	2
v_j	2	1	0	

The reduced costs are now $r_{12} = c_{12} - u_1 + v_2 = 0$, $r_{13} = c_{13} - u_1 + v_3 = -2$, $r_{23} = c_{23} - u_2 + v_3 = -1$ and $r_{31} = c_{31} - u_3 + v_1 = 2$. Since the reduced cost r_{13} is most negative the flow in x_{13} should be increased.

x_{ij}	$j=1$	$j=2$	$j=3$	s_i
$i=1$	$4-t$	0	$+t$	4
$i=2$	$3+t$	$1-t$	0	4
$i=3$	0	$7+t$	$5-t$	12
d_j	7	8	5	

Increasing the flow in x_{13} to t a cycle in the graph is created and we must compensate to get $x_{33} = 5-t$, $x_{32} = 7+t$, $x_{22} = 1-t$, $x_{21} = 3+t$, and $x_{11} = 4-t$. So t can become at most 1 and then x_{22} becomes zero and exits the basis.

x_{ij}	$j=1$	$j=2$	$j=3$	s_i
$i=1$	3	0	1	4
$i=2$	4	0	0	4
$i=3$	0	8	4	12
d_j	7	8	5	

In the new flow $x = (x_{11}, x_{21}, x_{13}, x_{32}, x_{33}) = (3, 4, 1, 8, 4)$.
Put the node potential v_3 at node sink 3 to be 0.
Then $u_3 - v_3 = c_{33} = 2$ gives $u_3 = 2$.
Then $u_3 - v_2 = c_{32} = 1$ gives $v_2 = 1$.
Then $u_1 - v_3 = c_{13} = 2$ gives $u_1 = 2$.
Then $u_1 - v_1 = c_{21} = 2$ gives $v_1 = 0$.
Then $u_2 - v_1 = c_{21} = 1$ gives $u_2 = 1$.

c_{ij}, r_{ij}	$j=1$	$j=2$	$j=3$	u_i
$i=1$	2	0	2	2
$i=2$	1	2	-1	1
$i=3$	2	1	2	2
v_j	0	1	0	

The reduced costs are now $r_{12} = c_{12} - u_1 + v_2 = 2$, $r_{22} = c_{22} - u_2 + v_2 = 2$, $r_{23} = c_{23} - u_2 + v_3 = 1$ and $r_{31} = c_{31} - u_3 + v_1 = 0$. Since the reduced costs are all non-negative the optimal value can not be increased. However, the optimal solution is not unique, since one reduced cost is zero and the solution is not degenerate.

(b) Perform row operations on A,

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix},$$

to get it into staircase form.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

Columns 1,2,3,4 and 7 are unit vectors, so the corresponding columns in A span the range space, *i.e.*, the columns of

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

is a basis for the range space of A.
For the nullspace, let
$x_1 = x_5 + x_6 + x_8 + x_9$, $x_4 = -x_5 - x_6$, $x_7 = -x_8 - x_9$, $x_2 = -x_5 - x_8$ and $x_3 = -x_6 - x_9$.
so an arbitrary vector x in the nullspace of A can be written

$$x = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_5 \\ x_6 \\ x_8 \\ x_9 \end{pmatrix}$$

2. (a) The standard form is

$$(P_s) \quad \begin{bmatrix} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{bmatrix}$$

where $c = (-2\ -3\ -3\ 1\ 1\ -2)^T$. The constraints are defined by

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 2 & 1 \\ 1 & -1 & 0 & 1 & 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

We start with x_1 and x_3 as basic variables. I.e. basic and non-basic variable indices are $\beta = \{1, 3\}$ and $\eta = \{2, 4, 5, 6\}$, so

$$A_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, A_\nu = \begin{bmatrix} 1 & 0 & 2 & 1 \\ -1 & 1 & 0 & 2 \end{bmatrix}$$

and $\bar{b} = A_\beta^{-1} b = [3\ 0]^T$ which gives the starting basic solution $x = (3, 0, 0, 0, 0, 0)$. This is a degenerate basic solution.
From the equations $A_\beta^T y = c_\beta$ and $\hat{c}_\nu^T = c_\nu^T - y^T A_\nu$ we get

$$y = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \quad r_\nu^T = \begin{bmatrix} 1 & 0 & 7 & -1 \end{bmatrix}.$$

Let x_6 enter the basis. Which one should exit ?

From $A_\beta \hat{a}_6 = a_6$, we get that $\hat{a}_6 = (2, -1)^T$, and since the second element is negative, x_1 exits the basis.

Update basic and non-basic matrices; The basic and non-basic variable indices are given by $\beta = \{3, 6\}$ och $\eta = \{1, 2, 4, 5\}$, and

$$A_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, A_\nu = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & -1 & 1 & 0 \end{bmatrix}$$

The equations $A_\beta^T y = c_\beta$ and $\hat{c}_\nu^T = c_\nu^T - y^T A_\nu$ gives

$$y = \begin{bmatrix} -3 \\ 0.5 \end{bmatrix}, \quad r_\nu^T = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 7 \end{bmatrix}.$$

Since all reduced costs are non-negative, $\hat{x} = (0, 0, 3/2, 0, 0, 3/2)^T$ is optimal.

(b) The dual is

$$(D) \quad \begin{bmatrix} \max\limits_{y} & 3y_1 + 3y_2 \\ \text{s.t.} & y_1 + y_2 \leq -2 \\ & y_1 - y_2 \leq -3 \\ & y_1 \leq -3 \\ & y_2 \leq 1 \\ & 2y_1 \leq 1 \\ & y_1 + 2y_2 \leq -2 \\ & y_i \text{ free.} \end{bmatrix}.$$

(c) The optimal dual solution is

$$\hat{y} = \begin{bmatrix} -3 \\ 0.5 \end{bmatrix},$$

from the last step of the simplex method. It is easy to check that all constraints are satisifed and that the objective function value is -7.5.

The starting simplex multiplicators are

$$y_s = \begin{bmatrix} -3 \\ 1 \end{bmatrix},$$

and it is easy to see that the last constraint is not satisfied and that the objective function value is -6.

According to the weak duality Theorem it should hold that $c^T x \geq b^T y$ for all x feasible for the primal and y feasible for the dual.

Here, $b^T y_s = -6 \not\leq -9 = b^T \hat{y} = c^T \hat{x}$ but y_s is not feasible to the dual, so it does not contradict the Theorem.

3. (a) This a least-squares problem with

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then

$$A^T A = \begin{bmatrix} 2 & 2 & 4 \\ 2 & 4 & 6 \\ 4 & 6 & 10 \end{bmatrix}, \quad A^T B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

and $\bar{x} = [0\ 1/2\ 0]$ is a solution to $A^T A x = A^T B$ and there are infinitely many others since A only has rank 2.
Therefore, we want to find the minimum norm solution:

$$AA^T = \begin{bmatrix} 2 & 4 \\ 4 & 14 \end{bmatrix}, \quad A\bar{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

then $u = [-1/3\ 1/6]^T$ solve $AA^T u = A\bar{x}$.
Finally

$$\hat{x}^{(1)} = A^T u = \begin{bmatrix} -1/6 \\ 1/3 \\ 1/6 \end{bmatrix}$$

(b) We add the linear constraint $Ax = b$, where $A = [1\ 1\ 1]$ and $B = 1$.
For the Lagrange method, the following equation system must be solved

$$\begin{bmatrix} H & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix},$$

That is:

$$\begin{bmatrix} 2 & 2 & 4 & -1 \\ 2 & 4 & 6 & -1 \\ 4 & 6 & 10 & -1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix},$$

from which we see that $\hat{u} = 0$ and $\hat{x}^{(2)} = [1/2\ 1\ \ -1/2]$.
In both cases, the objective function values are zero, *i.e.*,

$$f(\hat{x}^{(i)}) = \frac{1}{2}\hat{x}^{(i)T} A^T A \hat{x}^{(i)} - (A^T B)^T \hat{x}^{(i)} + \frac{1}{2} B^T B = 0$$

for $i = 1, 2$. In both cases there is no matching error, *i.e.*, perfect matches are possible, in addition two other criteria are used to obtain unique solutions.

4. (a) The gradient and hessian are given by

$$\nabla f(x, y) = \begin{bmatrix} y + x^2 & x + y^2 \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} 2x & 1 \\ 1 & 2y \end{bmatrix}.$$

The first order optimality conditions are satisfied when $\nabla f(x, y) = 0$, i.e., when $y^2 + x = (-x^2)^2 + x = x(1 + x^3) = 0$. That is when $x = 0$ or $x = -1$, which gives us two points $x^{(1)} = (0, 0)$ and $x^{(2)}(-1, -1)$.

The second order optimality conditions depends on the definiteness of the Hessian.

$$\nabla^2 f(x^{(1)}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \nabla^2 f(x^{(2)}) = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

The first point corresponds to a saddle point and is not a local optimum and the second point corresponds to a local maximum since the Hessian is negative definite.

The problem is unbounded from below, so there is no global optimum for the problem.

(b) We consider the Lagrange optimality conditions

$$\nabla f(x) + \lambda \nabla h(x) = \begin{bmatrix} y + x^2 & x + y^2 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 1 \end{bmatrix} = 0,$$

which is satisfied for $\lambda = 0$ and $x/y = 1/e$. Using the constraint $x + y = 1$ gives $y = 1 - x$. Then $x + y^2 + \lambda = x^2 - x + (1 + \lambda) =$

Then $x = \frac{1}{2} \pm \sqrt{\frac{1}{4} - 1 - \lambda}$ and $y = \frac{1}{2} \mp \sqrt{\frac{1}{4} - 1 - \lambda}$.

Since $t = \sqrt{\frac{1}{4} - 1 - \lambda} \geq 0$ can be chosen arbitrary, any x and y such that $x + y = 1$ satifies the equation system.

In fact, $f(x, 1 - x) = \frac{1}{3}$, so all points on the line are optimal and have the same optimal value.

5. (a) The feasible set is convex since the functions f_i are linear, and therefore also convex.

The function e^{-x} is convex, since $\nabla^2 e^{-x} = e^{-x} > 0$. Since f_0 is a sum of positive konstant times convex functions it is convex on ^{2}, hence convex on the feasible set.

(P) is a convex optimization problem.

(b) The gradients are given by

$$\nabla f_0(x) = \begin{bmatrix} -e^{-x_1} & -2e^{-x_2} & -3e^{-x_3} \end{bmatrix},$$

$$\nabla f_1(x) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad \nabla f_2(x) = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}.$$

$$\nabla f_3(x) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad \nabla f_2(x) = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

The KKT conditions are

$$\begin{bmatrix} -e^{-x_1} + y_1 + y_4 & -2e^{-x_2} + y_2 + y_4 & -3e^{-x_3} + y_3 + y_4 \end{bmatrix} = 0.$$

$$y_1 \geq 0, y_2 \geq 0, y_3 \geq 0, y_4 \geq 0,$$

$$x_1 \leq 1, \quad x_2 \leq 1, \quad x_3 \leq 1, \quad x_1 + x_2 + x_3 \leq 2.5,$$

$$(x_1 - 1)y_1 = 0, \ (x_2 - 1)y_2 = 0, \ (x_3 - 1)y_3 = 0, \ (x_1 + x_2 + x_3 - 2.5)y_4 = 0,$$

Since the problem is strictly convex, there will be at most one point that satisfies the KKT conditions.

The objective function is decreasing in each of the variables so we want all variables as small as possible and especially those with large index. Not all constraints can be active at the same time. Assume that constraints 2,3 and 4 are active. Then $x_2 = 1$, $x_3 = 1$ and $x_1 + x_2 + x_3 = 2.5$, *i.e.*, $x_1 = 0.5$. Can the other KKT conditions be satisfied?

With $y_1 = 0$ then KKT4 are satisfied.

For KKT1 we have

$$\begin{bmatrix} -e^{-0.5} + y_4 & -2e^{-1} + y_2 + y_4 & -3e^{-1} + y_3 + y_4 \end{bmatrix} = 0.$$

This holds if $y_4 = e^{-0.5} = 0.61 \geq 0$, and $y_2 = 2e^{-1} - y_4 = 2 \cdot 0.37 - 0.61 = 0.13 \geq 0$, and $y_3 = 3e^{-1} - y_4 = 3 \cdot 0.37 - 0.13 = 0.98 \geq 0$. So KKT2 is also satisfied and since the problem is convex and regular this is the optimal point.

(c) The objective function is lower bounded by $\sum_{k=1}^{n} ke^{-1} = 0.37\frac{n(n+1)}{2}$ and therefore there exists an optimal point, which must be unique due to strict convexity.

If $m \geq n$ then $x_i = 1$ for $i = 1, \cdots, n$, and the lower bound above is attained.

It is clear that if some $x_\mu = 1$ at the optimal solution then x_k has to be 1 also for all $k \geq \mu$. It is important to understand that the optimal solution has the following structure; If $m \leq n$, then there is a μ such that $x_i = 1$ for $i = n - \mu + 1, \cdots, n$ and $y_i = 0$ for $i = 1, \cdots, n - \mu$.

It is a bit complicated to prove that the KKT solution has this structure, and a complete analysis is not needed for full points.

To prove optimality we need to show that the KKT conditions has such a solution.

KKT1 leads to $ke^{-x_k} - y_{n+1} = 0$ for $k = 1, \cdots, n - \mu$, hence

$$x_k = \log(k/y_{n+1}), \qquad k = 1, \cdots, n - \mu.$$

The condition $x_1 + \cdots + x_{n-\mu} = m - \mu$, then determines y_{n+1} from

$$\sum_{k=1}^{n-\mu} \log(k/y_{n+1}) = \sum_{k=1}^{n-\mu} \log(k) - (n - \mu) \log y_{n+1} = m - \mu.$$

which is

$$y_{n+1} = \exp\left\{ \frac{\sum_{k=1}^{n-\mu} \log(k) - (m - \mu)}{n - \mu} \right\} \geq 0.$$

It remains to show that for this choice of x_k there are positive $y_{n-\mu+1}, \cdots, y_n$ such that the rest of the KKT1 conditions are satisfied. Now $-ke^{-1} + y_k + y_{n+1} = 0$, so $y_k = ke^{-1} - y_{n+1} \geq 0$ should hold, and it is enough to check $(n - \mu + 1)e^{-1} \geq y_{n+1}$. This inequality determines the value of μ.