KTH Mathematics

Solutions for the exam in SF1811/SF1831/SF1841 Optimization for F.
monday June 8, 2009, time. 14.00–19.00

Instructor: Per Enqvist, tel. 790 62 98

There may be alternative solutions to the problem.

---

1. (a) The problem (P) can be written on standard form

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

where

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

We start with x_3 and x_4 in the basis, giving the solution $x = (0, 0, 3, 1)$ from (a).

Basic and non-basic variable indices are given by $\beta = \{3, 4\}$ and $\eta = \{1, 2\}$, and

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$$

Then the equations $B^T y = c_B$ and $\hat{c}_N^T = c_N^T - y^T N$ gives

$$y = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \hat{c}_N^T = \begin{bmatrix} -1 & -2 \end{bmatrix}.$$

Let x_2 enter the basis. Which one should exit ?

From $B\hat{a}_2 = a_2$, we get $\hat{a}_2 = (3, -1)^T$, and since the second element is negative, x_3 exits the basis.

Update the basis and nonbasis matrices: Basic and non-basic variable indices are given by $\beta = \{2, 4\}$ and $\eta = \{1, 3\}$, and

$$B = \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Then the equations $B^T y = c_B$ and $\hat{c}_N^T = c_N^T - y^T N$ gives

$$y = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}, \quad \hat{c}_N^T = \begin{bmatrix} -8/3 & -2/3 \end{bmatrix}.$$

Let x_1 enter the basis. Which one should exit ?

From $B\hat{a}_1 = a_1$, we get $\hat{a}_1 = (1/3, 4/3)^T$. $\hat{b} = (1, 2)^T$ satisfies $B\hat{b} = b$, and since $\hat{b}_1/(\hat{a}_1)_1 = 3 > 4/3 = \hat{b}_2/(\hat{a}_1)_2$, x_4 exits the basis.

Update the basis and nonbasis matrices: Basic and non-basic variable indices are given by $\beta = \{2,1\}$ and $\eta = \{4,3\}$, and

$$B = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}, N = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Then the equations $B^T y = c_B$ and $\hat{c}_N^T = c_N^T - y^T N$ gives

$$y = \begin{bmatrix} -3/4 \\ -1/4 \end{bmatrix}, \quad \hat{c}_N^T = \begin{bmatrix} 1/4 & 3/4 \end{bmatrix}.$$

Since all reduced costs are nonnegative, the current bfs $\hat{x} = (3/2, 1/2, 0, 0)^T$ is optimal.

(b) The dual linear programing problem is

$$(D') \quad \begin{bmatrix} \max\limits_{y} & -b^T y \\ \text{s.t.} & -A^T y \le c \\ & y \ge 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} \max\limits_{y} & -3y_1 - y_2 \\ \text{s.t.} & -y_1 - y_2 \le -1 \\ & -3y_1 + y_2 \le -2 \\ & y_1 \ge 0, y_2 \ge 0. \end{bmatrix}.$$

Since $x_1^{(a)}$ and $x_2^{(a)}$ are both non-zero, it follows by complementarity that $-y_1 - y_2 = -1$ and $-3y_1 + y_2 = -2$, i.e. $y_1 = 3/4$ and $y_2 = 1/4$. This solution also satisfies the positivity constraints, so it is feasible for the dual.
The dual solution can also be obtained from the last simplex iteration in (b).

(c) In problem (P_2) the constraint that $x_2 = 1 - 2x_1 \ge 0$ has disappeared, but it still has to be valid for (P_1) and (P_2) to be equivalent.

2. (a) The problem can be written as $\min_z \|Az - b\|^2$ with

$$A = \begin{bmatrix} x_1^2 & x_1 \\ x_2^2 & x_2 \\ \vdots & \vdots \\ x_m^2 & x_m \end{bmatrix}, \quad z = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

(b) For given data we get

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 9 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}.$$

The equation $(A^T A)\bar{z} = A^T b$ is then

$$\begin{bmatrix} 1 & 4 & 9 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 9 & 3 \end{bmatrix} \bar{z} = \begin{bmatrix} 1 & 4 & 9 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix},$$

i.e.

$$\begin{bmatrix} 98 & 36 \\ 36 & 14 \end{bmatrix} \bar{z} = \begin{bmatrix} 35 \\ 15 \end{bmatrix}, \quad \Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{98 \cdot 14 - 36^2} \begin{bmatrix} 14 & -36 \\ -36 & 98 \end{bmatrix} \begin{bmatrix} 35 \\ 15 \end{bmatrix}.$$

3. (a) For the optimization problem to be convex, it is necessary that the feasible region is convex and that the objective function is convex on the whole feasible region. The feasible region is convex since it is given by a linear equality constraint.

The objective function is convex on the feasible region if $Z^T HZ$ is positive semidefinite for some matrix Z whos columns spans the nullspace of A. With

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

then

$$Z^T HZ = \begin{bmatrix} 7 & 5 \\ 5 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 5/7 & 1 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 10/7 \end{bmatrix} \begin{bmatrix} 1 & 5/7 \\ 0 & 1 \end{bmatrix}$$

and from the LDL^T-factorization we see that it is positive definite.

(b) Evaluating the gradient at $\bar{x}$ we get

$$H\bar{x} + c = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix} = A^T(-2),$$

so $H\bar{x} + c$ is in the range space of A^T, thus there exists no feasible descent directions at $\bar{x}$.

Since the problem is convex and there exists no feasible descent directions we know that $\bar{x}$ is a global optimum.

The directional derivative in the direction $d = (0, -1, 0)^T$ is negative, so it is a descent direction. (but the direction d is not feasible.)

(c) Using LDL^T-factorization

$$H+2I = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & -1/2 \\ 0 & -1/2 & 19/4 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

we get a zero diagonal element on a non-zero row, hence the matrix $H + 2I$ is not positive definite.

Using LDL^T-factorization

$$H+3I = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2/5 & 1 & 0 \\ 1/5 & -1/3 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 6/5 & 0 \\ 0 & 0 & 17/3 \end{bmatrix} \begin{bmatrix} 1 & 2/5 & 1/5 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

and since all diagonal element are positive the matrix $H+3I$ is positive definite.

The first optimization problem has no finite optimal solution, while the second one has a unique one which is $x = (0, 0, 0)^T$.

4. (a) The gradient of f is given by

$$\nabla f(x) = \begin{bmatrix} (x_2 - 1/3)(x_3 - 1/4) & (x_1 - 1/2)(x_3 - 1/4) & (x_2 - 1/3)(x_1 - 1/2) \end{bmatrix}.$$

The gradient is zero if two of the three conditions $x_1 = 1/2$, $x_2 = 1/3$ and $x_3 = 1/4$ holds. This holds true in for example the three points

$$(1/2, 1/3, \epsilon) \qquad (\epsilon, 1/3, 1/4) \qquad (1/2, \epsilon, 1/4),$$

where $\epsilon = 1/10$, or some other value in $(0, 1)$.

(b) The hessian of f is given by

$$\nabla^2 f(x) = \begin{bmatrix} 0 & x_3 - 1/4 & x_2 - 1/3 \\ x_3 - 1/4 & 0 & x_1 - 1/2 \\ x_2 - 1/3 & x_1 - 1/2 & 0 \end{bmatrix},$$

and for the particular $x^{(0)}$ given, the matrix is all zero and therefore positive semidefinite, hence the point satisfies the second order necessary conditions.

However, the point is not a local minimum. We know that $f(x^{(0)}) = 0$. Consider the point $x_\epsilon = (1/2 - \epsilon)(1/3 - \epsilon)(1/4 - \epsilon)$, then $f(x_\epsilon) = -\epsilon^3 < 0$ for any $\epsilon > 0$, which demonstrates that there is no neighborhood of $x^{(0)}$ such that the minimal value in that set is zero.

(c) Let

$$g_0(x) = 1 - (x_1^2 + x_2^2 + x_3^2)$$
$$g_1(x) = x_1, \quad g_2(x) = x_2, \quad g_3(x) = x_3.$$

Then, the problem can be written $\min f(x)$ subject to $g_i(x) \geq 0$ for $i = 0, 1, 2, 3$. At $x^{(1)}$, constraints 0,1 and 2 are active.

$$\nabla f(x^{(1)})^T = \begin{pmatrix} -3/12 \\ -3/8 \\ 1/6 \end{pmatrix},$$

$$\nabla g_0(x^{(1)})^T = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}, \quad \nabla g_1(x^{(1)})^T = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \nabla g_2(x)^T = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

For the KKT-conditions to be satisfied we need to find non-negative Lagrange parameters such that

$$\begin{pmatrix} -1/4 \\ -3/8 \\ 1/6 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} \lambda_0 - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \lambda_1 - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \lambda_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now $\lambda_0 = -1/12, \lambda_1 = -1/4, \lambda_2 = -3/8$ and $\lambda_3 = 0$.
So the KKT-conditions are not satisfied.
Then we can say that $x^{(1)}$ is not a local minimum.

(d) At $x^{(2)}$, constraints 1,2 and 3 are active.

$$\nabla f(x^{(2)})^T = \begin{pmatrix} 1/12 \\ 1/8 \\ 1/6 \end{pmatrix},$$

$$\nabla g_1(x^{(1)})^T = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \nabla g_2(x)^T = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \nabla g_3(x^{(2)})^T = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

For the KKT-conditions to be satisfied we need to find non-negative Lagrange parameters such that

$$\begin{pmatrix} 1/12 \\ 1/8 \\ 1/6 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \lambda_1 - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \lambda_2 - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \lambda_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now $\lambda_0 = 0$, $\lambda_1 = 1/12, \lambda_2 = 1/8$, and $\lambda_3 = 1/6$.
So the KKT-conditions are satisfied.
But this is not enough to say that $x^{(2)}$ is a local minimum.

5. (a) The tangent at $(x_1^{(k)}, x_2^{(k)})$ is given by $\cos(v_k)x_1 + \sin(v_k)x_2 = 1$.
A linear approximative solution is then given by

$$\begin{array}{rl} \text{minimize} & -x_1 - 2x_2 \\ \text{s.t.} & \cos(v_k)x_1 + \sin(v_k)x_2 + y_k = 1, \quad k = 0, 1, \cdots, N \\ & x_1 \geq 0, x_2 \geq 0, y_k \geq 0, \quad k = 0, 1, \cdots, N. \end{array}$$

where y_k are slack-variables.

(b) Since the feasible region to the linear approximative problem is larger than the original nonlinear problem, the optimal value of the linear problem is lower or equal to the nonlinear one. The optimal point for the nonlinear problem will always be feasible also for the approximative linear one.

(c) The Lagrange function is

$$L(x, \lambda) = -x_1 - 2x_2 - \lambda \left(1 - x_1^2 - x_2^2\right)$$

It is well defined for $x_i \geq 0$, and separates into two independent convex quadratic minimization problems that are minimized for the $x_i(\lambda)$ such that

$$\frac{\partial L}{\partial x_1}(x, \lambda) = -1 + \lambda 2x_1 = 0, \quad \frac{\partial L}{\partial x_2}(x, \lambda) = -2 + \lambda 2x_2 = 0,$$

i.e., $x_1(\lambda) = 1/(2\lambda)$ and $x_2(\lambda) = 1/\lambda$.
The dual function is obtained by inserting this x_i in the Lagrange function

$$J(\lambda) = -1/(2\lambda) - 2/\lambda - \lambda \left(1 - \frac{1}{4\lambda^2} - \frac{1}{\lambda^2}\right) = -\lambda - \frac{5}{4\lambda}$$

(d) The derivative of the dual objective function is given by $f'(\lambda) = \frac{5}{4\lambda^2} - 1$, and the tangent at λ_k is given by $y = K\lambda + m$ for $K = \frac{5}{4\lambda_k^2} - 1$ and m such that

$$-\lambda_k - \frac{5}{4\lambda_k} = K\lambda_k + m$$

that is $m = -5/(2\lambda_k)$.
The tangent to the dual objective function is then given by

$$y = \left(\frac{5}{4\lambda_k^2} - 1\right)\lambda - \frac{5}{2\lambda_k},$$

and the piecewise linear approximation is given by

$$y(\lambda) = \min_{k=0,1,\cdots,N} \left\{ \frac{5}{4\lambda_k^2} - 1\lambda - \frac{5}{2\lambda_k} \right\}.$$

Finally, the linear optimization problem

$$\begin{aligned} \text{maximize} \quad & t \\ \text{s.t.} \quad & t \le \left(\tfrac{5}{4\lambda_k^2} - 1\right)\lambda - \tfrac{5}{2\lambda_k}, \quad k = 0, 1, \cdots, N \end{aligned}$$

approximates the dual optimization problem.

(e) All the tangents of the concave dual objective function lies above the the function and therefore the optimal value of the linear problem in (d) gives a larger value than the nonlinear dual optimization problem.