Solutions to Homework #2.

#1. (10 points.) Let a constant $p > 1$ and a vector $a = (a_1, a_2, \ldots, a_n)^T \in \mathbb{R}^n$ with components $a_j \geq 0$ be fixed. Find the maximum of the function

$$f(x) = f(x_1, x_2, \ldots, x_n) := (a, x) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

under the constraints

$$g(x) = g(x_1, x_2, \ldots, x_n) := x_1^p + x_2^p + \cdots + x_n^p = 1, \quad x_j \geq 0 \quad \text{for all} \quad j.$$

Solution. The constraints describe a bounded closed set (a compact) $K \subset \mathbb{R}^n$, hence the bounded continuous function $f(x)$ attains its maximum at some point $x = (x_1, x_2, \ldots, x_n)^T \in K$. First assume that $x_j > 0$ for all $j$, so that we have only one constraint $g(x) = 1$. As in Examples 2.3 and 2.4, from the relation $\nabla f = \lambda \nabla g$ it follows

$$a_j = \frac{\partial f}{\partial x_j} = \lambda \cdot \frac{\partial g}{\partial x_j} = \lambda p \cdot x_j^{p-1} \quad \text{for all} \quad j = 1, 2, \ldots, n. \quad (1.1)$$

It is convenient to introduce the notation for arbitrary $a \in \mathbb{R}^n$

$$||a||_q := \left( \sum_{j=1}^n |a_j|^q \right)^{1/q}, \quad (1.2)$$

where the given constant $p > 1$ and the new constant $q > 1$ are related as follows:

$$\frac{1}{p} + \frac{1}{q} = 1, \quad \text{i.e.} \quad q := \frac{p}{p-1} = 1 + \frac{1}{p-1} > 1. \quad (1.3)$$

Together with (1.1), these relations imply:

$$||a||_q^q = \sum_{j=1}^n a_j^q = (\lambda p)^q \cdot \sum_{j=1}^n x_j^{(p-1)q} = (\lambda p)^q \cdot \sum_{j=1}^n x_j^p = (\lambda p)^q,$$

$$\lambda p = ||a||_q, \quad x_j = \left( \frac{a_j}{\lambda p} \right)^{\frac{1}{p-1}} = ||a||_{q^{p-1}}^{\frac{1}{p-1}} \cdot a_j^{\frac{1}{p-1}},$$

$$f(x) = \sum_{j=1}^n a_j x_j = ||a||_q^{\frac{1}{p-1}} \cdot \sum_{j=1}^n a_j^{\frac{1}{p-1}} = ||a||_{q^{p-1}}^{\frac{1}{p-1}} \cdot ||a||_q^q = ||a||_q.$$
If we just drop $j \notin J$ from our consideration, i.e. consider only $j \in J$ in place of $j \in \{1, 2, \ldots, n\}$, then the corresponding value $f(x) = ||a||_q$ in (1.2) will be replaced by a similar expression, in which the sum is extended only over $j \in J$. Thus if we want to maximize the value of $f(x)$ we must restrict ourselves to $J := \{ j : a_j > 0 \}$. In any case, the answer $f(x) = ||a||_q$ remains the same.

**Remark 1.** The above solution was a demonstration of the method of Lagrange multipliers. There is a shorter way to do it based on Hölder’s inequality

\[
|(x, y)| \leq ||x||_p \cdot ||y||_q \quad \text{for all} \quad x, y \in \mathbb{R}^n.
\] (HE)

Here we use notations in (1.2), (1.3). In particular, the Cauchy–Schwarz inequality appears as a particular case $p = q = 2$. By linearity, the proof of (HE) is reduced to the case $||x||_p = ||y||_q = 1$, and then it follows from the elementary inequality

\[
ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \text{for all} \quad a \geq 0, b \geq 0.
\]

**#2.** (12 points.) Find the maximum of the function

\[
y = f(x_1, x_2, x_3) = \sin x_1 \cdot \sin x_2 \cdot \sin x_3
\]

over the set $\{g(x_1, x_2, x_3) = x_1 + x_2 + x_3 = \pi/2, \ x_1 > 0, \ x_2 > 0, \ x_3 > 0\}$.

**Solution.** From the Lagrange equality $\nabla f = \lambda \nabla g$ it follows

\[
\cos x_1 \cdot \sin x_2 \cdot \sin x_3 = \sin x_1 \cdot \cos x_2 \cdot \sin x_3 = \sin x_1 \cdot \sin x_2 \cdot \cos x_3 = \lambda.
\]

Then

\[
\frac{f(x_1, x_2, x_3)}{\lambda} = \tan x_1 = \tan x_2 = \tan x_3.
\]

From the given restriction it follows that $x_1, x_2, x_3$ belong to $(0, \pi/2)$. Since $\tan x$ is a one-to-one function on this interval, we must have $x_1 = x_2 = x_3 = \pi/6$, with the maximal value

\[
f_{max} = f(\pi/6, \pi/6, \pi/6) = \sin^3(\pi/6) = 1/8.
\]

**#3.** (10 points.) Using Newton’s iteration method, find the roots of the equation $x \ln x = 1$ correct to six decimal places. Start with the initial point $x_0 = 1$, and write all the successive approximations $x_1, x_2, \ldots$.

**Solution.** Following p.63 in the textbook, we have $F(x) = x \ln x - 1 = 0$. By Newton’s iteration method, the successive approximations are defined as follows:

\[
x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)} = x_n - \frac{x_n \ln x_n - 1}{\ln x_n + 1} = \frac{x_n + 1}{\ln x_n + 1}.
\]

Starting with $x_0 = 1$, we get

\[
x_1 = 2,
\]
\[
x_2 \approx 1.771\,848\,3,
\]
\[
x_3 \approx 1.763\,236\,2,
\]
\[
x_4 \approx 1.763\,222\,8,
\]
\[
x_5 \approx 1.763\,222\,8.
\]
#4. (12 points.) In the system
\[
\begin{align*}
  x_1 - 2x_2 & \leq 1 \\
  x_1 - x_2 & \leq 3 \\
  x_1 + 2x_2 & \leq 5 \\
  x_1 + x_2 & = 4,
\end{align*}
\]
determine which restriction are **binding** and which are **redundant**.

**Answer.** The restrictions (1), (3), and (4) are binding. The restriction (2) is redundant. One can either use the ERO method, or observe that
\[
(1), (3) \implies x_1 \leq 3, \quad (3), (4) \implies x_2 \leq 1,
\]
and then from (4) it follows \(x_1 = 3, \ x_2 = 1\).

#5. (16 points.) Using simplex method, find the maximum of the objective function
\[
y = f(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + x_3 + x_4 - 4x_5
\]
over the set
\[
\begin{align*}
  x_1 + 2x_2 + 3x_3 + 3x_4 - 7x_5 & = 10, \\
  x_1 + 3x_2 + x_3 - 6x_5 & = 7, \\
  x_1 - 3x_2 + x_3 - 6x_4 & = 1, \\
  x_i & \geq 0 \text{ for } i = 1, 2, 3, 4, 5;
\end{align*}
\]
starting from the initial corner point \(x^{(0)} = (2, 1, 2, 0, 0)\). Write down all the intermediate systems (or their matrices) and the corresponding corner points.

**Solution.** There is more than one way to approach this problem. The initial table for this system with the initial corner point \(x^{(0)} = (2, 1, 2, 0, 0)^T\) is
\[
T = \begin{bmatrix}
1 & -1 & -1 & -1 & 4 & 0 \\
0 & 1 & 2 & 3 & 3 & -7 & 10 \\
0 & 1 & 3 & 1 & 0 & -6 & 7 \\
0 & 1 & -3 & 1 & -6 & 0 & 1
\end{bmatrix}.
\]

Here the basic variables are \(x_j, \ j \in J := \{1, 2, 3\}\). In order to resolve this system with respect to these variable, we replace the columns \(A_0, A_1, A_2, \text{ and } A_3\) by standard unit vectors (see pages 12-13 in Notes), we need to multiply \(T\) by the inverse matrix composed of these columns. We will get an equivalent system with the same corner point \(x^{(0)}\):

\[
T^{(0)} = \begin{bmatrix}
1 & -1 & -1 & -1 \\
0 & 1 & 2 & 3 \\
0 & 1 & 3 & 1 \\
0 & 1 & -3 & 1
\end{bmatrix}^{-1} \cdot T = \begin{bmatrix}
1 & -1/2 & 2/3 & 1/3 \\
0 & 0 & 1/6 & -1/6 \\
0 & 1/2 & -5/12 & -1/12 \\
0 & 1 & -3 & 1
\end{bmatrix} \cdot T.
\]
One can take as a pivot element either \( a_{24} = 1 \), or \( a_{34} = 2 \). If we choose \( a_{24} = 1 \), then the next iteration produces the table

\[
T^{(1)} = \begin{bmatrix}
1 & 0 & -3 & 0 \\
0 & 1 & -5 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 2 & 1
\end{bmatrix}^{-1} \cdot T^{(0)} = \begin{bmatrix}
1 & 0 & 3 & 0 \\
0 & 1 & 5 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & -2 & 1
\end{bmatrix} \cdot T^{(0)}
\]

\[
= \begin{bmatrix}
1 & 0 & 3 & 0 & 0 & -3 & 8 \\
0 & 1 & 5 & 0 & 0 & -7 & 7 \\
0 & 0 & 1 & 0 & 1 & -1 & 1 \\
0 & 0 & -2 & 1 & 0 & 1 & 0
\end{bmatrix}
\]

For the corresponding corner point \( x^{(1)} \), the new non-basic variables \( x_2 = x_5 = 0 \), therefore,

\[
x^{(1)} = (7, 0, 0, 1, 0)^T.
\]

Now we can take as a pivot element \( a_{35} = 1 \). Then

\[
T^{(2)} = \begin{bmatrix}
1 & 0 & 0 & -3 \\
0 & 1 & 0 & -7 \\
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0
\end{bmatrix}^{-1} \cdot T^{(1)} = \begin{bmatrix}
1 & 0 & 0 & 3 \\
0 & 1 & 0 & 7 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix} \cdot T^{(1)}
\]

\[
= \begin{bmatrix}
1 & 0 & -3 & 3 & 0 & 0 & 8 \\
0 & 1 & -9 & 7 & 0 & 0 & 7 \\
0 & 0 & -1 & 1 & 1 & 0 & 1 \\
0 & 0 & -2 & 1 & 0 & 1 & 0
\end{bmatrix}
\]

with the corner point

\[
x^{(2)} = (7, 0, 0, 1, 0)^T.
\]

For \( k = 2 \), we have \( c_k = c_2 = 3 > 0 \), and the set \( P := \{i : a_{i2} > 0\} \) is empty. This is the case (II), in which \( y = f(x) \) is unbounded and can take arbitrarily large values. According to (28) on page 10 in Notes, we can choose

\[
J = \{1, 4, 5\}, \quad t = x_k = x_2 > 0, \quad \text{and} \quad x_j = 0 \quad \text{for all} \quad j \notin J, \quad \text{i.e.} \quad x_3 = 0.
\]

Then the system is reduced to the following one:

\[
\begin{align*}
x_1 - 9x_2 &= 7, \\
x_2 + x_4 &= 1, \\
-2x_2 + x_5 &= 0.
\end{align*}
\]

This non-homogeneous system has a solution \( x = x^{(2)} \), and the corresponding homogeneous system, with \( x_2 = t > 0 \) has solutions \( x = t \cdot (9, 1, 0, 1, 2)^T \). Then

\[
x = x(t) = (7, 0, 0, 1, 0)^T + t \cdot (9, 1, 0, 1, 2)^T \quad \text{for} \quad t > 0
\]

satisfies all the restrictions, and \( y = f(x(t)) = 8 + 3t \to +\infty \) as \( t \to +\infty \).