Advanced Optimization and Decision Analyitcs: Assignment 1

Author: Simon Schmetz

The Document contains the solutions to the first Assignment of the Advanced Optimization Class at the Universidad Carlos III de Madrid in the Master of Statistics for Datascience. The first part contains the solutions to the two problems, manually found using matlab, the second part contains the solution to the same problems, but solved via Gurobi in Python.

Part 1: Solving manually Using Matlab

The following section contains the solutions solved using Matlab.

Problem 1

We try to optimize the following function defined $\mathbf{x} \in \mathbb{R}^3$,

$$f(\mathbf{x}) \triangleq 3x_1 + 5x_2 - 3x_3^2$$

s.t

$$g_1(\mathbf{x}) \triangleq 2x_1^2 - 37x_2 + 9x_3 = 18$$

$$g_1(\mathbf{x}) \triangleq 5x_1 + x_2 + 5x_3^2 = 24$$

We begin by defining the functions using the Symbolic Math Toolbox.

```
clear; clc;

% Define Variables
syms x1 x2 x3 lambda1 lambda2 real

x = [x1 x2 x3];
lambda = [lambda1 lambda2];
xlambda = [x1 x2 x3 lambda1 lambda2];

% Define functions
syms f(x1, x2, x3)
syms g1(x1, x2, x3)
syms g2(x1, x2, x3)
syms L(x1, x2, x3, lambda1, lambda2)

%% objective
f(x1, x2, x3) = 3*x1 + 5*x2 -3*x3^2; % define f

%% left side of constraints
g1(x1, x2, x3) = 2*x1^2 - 37*x2 +  9*x3;
g2(x1, x2, x3) = 5*x1 + x2 +  5*x3^2;

% right side of constraints
b1 = 18;
```

```
b2 = 24;
b = [b1, b2];

% define Lagrange Function:
L(x1, x2, x3, lambda1, lambda2) = f - lambda1*(g1-b1) - lambda2*(g2-b2)
```

$$\text{L(x1, x2, x3, lambda1, lambda2)} = 3\,x_1 + 5\,x_2 + \lambda_1\left(-2\,x_1^2 + 37\,x_2 - 9\,x_3 + 18\right) - \lambda_2\left(5\,x_3^2 + 5\,x_1 + x_2 - 24\right) - 3\,x_3^2$$

a) Analysis of the regularity of the feasible points

We than check that the the gradient of the constraint 1 can not be expressed as a linear combination of the the the gradient of constraint two, to verify that the problem is well defined, that the constraints are indipendent.

```
gradg1 = gradient(g1)
```

gradg1(x1, x2, x3) =
$$\begin{pmatrix} 4\,x_1 \\ -37 \\ 9 \end{pmatrix}$$

```
gradg2 = gradient(g2)
```

gradg2(x1, x2, x3) =
$$\begin{pmatrix} 5 \\ 1 \\ 10\,x_3 \end{pmatrix}$$

```
syms c1
xc = [x c1]
```

$$xc = \begin{pmatrix} x_1 & x_2 & x_3 & c_1 \end{pmatrix}$$

```
%assume(xc, 'real')
xnr = solve(gradg2 == c1 * gradg1, g1 == b(1), g2 == b(2), xc)
```

```
xnr = struct with fields:
    x1: [0×1 sym]
    x2: [0×1 sym]
    x3: [0×1 sym]
    c1: [0×1 sym]
```

As no solution exists, the constraints are independent and the feasible points are regular.

a) First order Necessary Conditions for local optimum

Define Gradients

We begin bei defining the gradient of the Lagrangian function L with respect to x

```
gradxL = simplify(gradient(L, x))
```

```
gradxL(x1, x2, x3, lambda1, lambda2) =
```

$$\begin{pmatrix} 3 - 4\lambda_1 x_1 - 5\lambda_2 \\ 37\lambda_1 - \lambda_2 + 5 \\ -9\lambda_1 - 6x_3 - 10\lambda_2 x_3 \end{pmatrix}$$

and the gradient of L with respect to both x and lambda.

```
gradL = simplify(gradient(L, xlambda))
```

```
gradL(x1, x2, x3, lambda1, lambda2) =
```

$$\begin{pmatrix} 3 - 4\lambda_1 x_1 - 5\lambda_2 \\ 37\lambda_1 - \lambda_2 + 5 \\ -9\lambda_1 - 6x_3 - 10\lambda_2 x_3 \\ -2x_1^2 + 37x_2 - 9x_3 + 18 \\ -5x_3^2 - 5x_1 - x_2 + 24 \end{pmatrix}$$

Identify Optima

With the Gradients defined, we solve for the points that fulfill the first oder NC of the gradient of L and yield a total of 4 candidate points.

```
stat_points = solve(gradL == 0, xlambda);
pe = [stat_points.x1 stat_points.x2 stat_points.x3 stat_points.lambda1
stat_points.lambda2];
pe = double(pe)
```

```
pe = 4×5
   -9.6690     5.4567     3.6576    -0.1504    -0.5630
  -10.1460     4.1641    -3.7568    -0.1523    -0.6365
    4.6563     0.7002     0.0607    -0.1080     1.0025
  -97.1627   509.8128    -0.0101     0.1080     8.9970
```

We then store the x and lambda values into seperate vectors.

```
xA = double([stat_points.x1(1) stat_points.x2(1) stat_points.x3(1)]);
lambdaA = double([stat_points.lambda1(1) stat_points.lambda2(1)]);

xB = double([stat_points.x1(2) stat_points.x2(2) stat_points.x3(2)]);
lambdaB = double([stat_points.lambda1(2) stat_points.lambda2(2)]);

xC = double([stat_points.x1(3) stat_points.x2(3) stat_points.x3(3)]);
lambdaC = double([stat_points.lambda1(3) stat_points.lambda2(3)]);

xD = double([stat_points.x1(4) stat_points.x2(4) stat_points.x3(4)]);
lambdaD = double([stat_points.lambda1(4) stat_points.lambda2(4)]);
```

c) Second order conditions for local optimum

We then want to identify wether the canidate points are maxima, minnima or saddle points. We beginn by calculating the Hessian of the Lagrangian. We find the Hessian to be indipendent of x values, meaning that if the Hessian for a given canidate point fulfills the second order conditions for the given lambdas, it does so for all values of x, e.g. globally.

```
HessxL = hessian(L, x)
```

HessxL(x1, x2, x3, lambda1, lambda2) =

$$\begin{pmatrix} -4\,\lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -10\,\lambda_2 - 6 \end{pmatrix}$$

```
syms d1 d2 d3 real
d = [d1 d2 d3].';
gradg1 = gradient(g1, x);
```

We thus calculate the Hessian for all Canidate Points.

Point A

For point A, we find the Hessian to be indefinite and thus a potential canidate for a local minimum.

```
HA = double(subs(HessxL, lambda, lambdaA));
double(eig(HA))
```

```
ans = 3×1
   -0.3700
         0
    0.6014
```

We continue by evaluating the tangential gradients

```
assume(d,'clear');
subs(gradg1, x, xA).' * d
```

ans(x1, x2, x3) =

$$9\,d_3 - 37\,d_2 - \frac{2721593429150539\,d_1}{70368744177664}$$

```
subs(gradg2, x, xA).' * d
```

ans(x1, x2, x3) =

$$5\,d_1 + d_2 + \frac{41180391572731175\,d_3}{1125899906842624}$$

where we find the term gradients to always be positive except for d = 0, indicating a local minimum with NC-1

```
assume(9*d3 - 37*d2 - (sym("2721593429150539")*d1)/sym("70368744177664") ==
0 &  5*d1 + d2 + (sym("41180391572731175")*d3)/sym("1125899906842624") ==0)
```

```
simplify(d.' * HA * d)
```

ans =

$$\frac{158167023807695756593662544485285069361110048086584 \, d_3{}^2}{30558320183165709706359994090657607176452145414144}$$

Solving the system of equations shows that the zero vector is the solution, concluding the SC2 Condition for a local minimum

```
solve(d3==0,d)
```

ans = struct with fields:
    d1: 0
    d2: 0
    d3: 0

Point B

We continue with point B, which we find to be positive semi-definite

```
HB = double(subs(HessxL, lambda, lambdaB));
double(eig(HB))
```

ans = 3×1
         0
    0.3650
    0.6094

We continue by evaluating the tangential gradients

```
assume(d,'clear');
subs(gradg1, x, xB).' * d
```

ans(x1, x2, x3) =

$$9 \, d_3 - 37 \, d_2 - \frac{713961626381871 \, d_1}{17592186044416}$$

```
subs(gradg2, x, xB).' * d
```

ans(x1, x2, x3) =

$$5 \, d_1 + d_2 - \frac{10574315830673315 \, d_3}{281474976710656}$$

where we find the term to always be positive except for d = 0, indicating a global minimum with NC-1

```
assume(9*d3 - 37*d2 - (sym("713961626381871")*d1)/sym("17592186044416") ==
0 &  5*d1 + d2 - (sym("10574315830673315")*d3)/sym("281474976710656") ==0)
simplify(d.' * HB * d)
```

ans =

$$\frac{834753253261976399630378265627635770309893143029677 \, d_3{}^2}{148833213438582847307369817371441958454426515537 92}$$

Again, solving the system of equations shows that the zero vector is the solution, and thus together with the initial observation of the Hessians independence of x, fullfilling the SC-2 of a global minimum

```
solve(d3==0,d)
```

```
ans = struct with fields:
    d1: 0
    d2: 0
    d3: 0
```

Point C

Like before, we evaluate the Hessian via the Eigenvalues and find it to be indefinite for Point C.

```
HC = double(subs(HessxL, lambda, lambdaC));
double(eig(HC))
```

```
ans = 3×1
  -16.0246
        0
    0.4322
```

In the tangent gradient analysis, we find the gradients to always be negative, indicating a NC1 Local maximum.

```
assume(d,'clear');
subs(gradg1, x, xC).' * d
```

```
ans(x1, x2, x3) =
```

$$\frac{5242499709344185 \, d_1}{281474976710656} - 37 \, d_2 + 9 \, d_3$$

```
subs(gradg2, x, xC).' * d
```

```
ans(x1, x2, x3) =
```

$$5 \, d_1 + d_2 + \frac{43724728865280405 \, d_3}{72057594037927936}$$

```
assume((sym("5242499709344185")*d1)/sym("281474976710656") - 37*d2 + 9*d3
== 0 &  5*d1 + d2 + (sym("43724728865280405")*d3)/sym("72057594037927936")
==0)
simplify(d.' * HC * d)
```

```
ans =
```

$$-\frac{345045516049093365191300552148188581147023401707487183 \, d_3^2}{21546136066498305225299114678351732534529874219499520}$$

With the only solution to the system the zero vector, the point thus fulfills the SC-2 condition for a Local Maximum.

```
solve(d3==0,d)
```

```
ans = struct with fields:
```

```
    d1: 0
    d2: 0
    d3: 0
```

Point D

For point D, we find the Hessian to be Negative Semidefinite, indicating a potential global maximum.

```
HD = double(subs(HessxL, lambda, lambdaD));
double(eig(HD))
```

```
ans = 3×1
  -95.9704
   -0.4321
        0
```

This is verified by the tangent gradient analysis with both the NC-1 of the following term always being negative fulfilled.

```
assume(d,'clear');
subs(gradg1, x, xD).' * d
```

```
ans(x1, x2, x3) =
```

$$9\,d_3 - 37\,d_2 - \frac{6837214120977123\,d_1}{17592186044416}$$

```
subs(gradg2, x, xD).' * d
```

```
ans(x1, x2, x3) =
```

$$5\,d_1 + d_2 - \frac{29199930463551635\,d_3}{288230376151711744}$$

```
assume(9*d3 - 37*d2 - (sym("6837214120977123")*d1)/sym("17592186044416") ==
0 &  5*d1 + d2 - (sym("29199930463551635")*d3)/sym("288230376151711744")
==0)
simplify(d.' * HD * d)
```

```
ans =
```

$$-\frac{2978375417854639444717415773763525572929657058755772987855\,d_3^{\,2}}{31034215142363974018918116680222741123184445232972300288}$$

As well as the only solution to the system the zero vector. Point D thus indeed is a global maximum.

```
solve(d3==0,d)
```

```
ans = struct with fields:
    d1: 0
    d2: 0
    d3: 0
```

We can also get the function values for the corresponding x coordinates as a quick sanity check to find that Point B indeed has the minimum function value out of all points, while Point D has the overall largest value.

```
fA = double(subs(f, x, xA));
```

```
fB = double(subs(f, x, xB));
fC = double(subs(f, x, xC));
fD = double(subs(f, x, xD));

fvals = [fA; fB; fC; fD]
```

```
fvals = 4×1
10³ ×
    -0.0419
    -0.0520
     0.0175
     2.2576
```

Problem 2

We try to optimize the following function defined $\mathbf{x} \in \mathbb{R}^2$,

$$f(\mathbf{x}) \triangleq x_1^2 + x_2^2 - x_3^2$$

s.t

$$h_1(\mathbf{x}) \triangleq 8x_1^2 + 24x_2 - 15x_3 \le 129$$

$$h_2(\mathbf{x}) \triangleq x_1^2 + 2x_2^2 + 4x_3^2 \ge 15$$

with $g_1(\mathbf{x})$ being reformulated to move the problem into standard form as

$$h_2(\mathbf{x}) \triangleq -x_1^2 - 2x_2^2 - 4x_3^2 \le -15$$

We then begin by again defining the functions and variables.

```
clear; clc; % clear all variables

% Define Variables
syms x1 x2 x3 mu1 mu2 real

x = [x1 x2 x3];
mu = [mu1 mu2];

xmu = [x mu];

% Define functions
syms f(x1, x2, x3)
syms h1(x1, x2, x3)
syms h2(x1, x2, x3)
syms L(x1, x2, x3, mu1, mu2)

%% objective
f(x1, x2, x3) = x1^2 + x2^2 - x3^2; % define f
```

```
%% left side of constraints
h1(x1, x2, x3) = 8*x1^2 + 24*x2 - 15*x3;
h2(x1, x2, x3) = -x1^2 + -2*x2^2 -4*x3^2;

% right side of constraints
b1 = 129;
b2 = -15;
b = [b1, b2];

% define Lagrange Function:
L(x1, x2, x3, mu1, mu2) = f - mu1*(h1-b1) - mu2*(h2-b2)
```

$$L(\text{x1, x2, x3, mu1, mu2}) = \mu_2 \left(x_1{}^2 + 2x_2{}^2 + 4x_3{}^2 - 15\right) - \mu_1 \left(8x_1{}^2 + 24x_2 - 15x_3 - 129\right) + x_1{}^2 + x_2{}^2 - x_3{}^2$$

a) Find candidate Points for a local maxima

As per assignment definition we can assume all points to be regular, we proceed to finding canidate points for Local Maxima under NC criterion. We start by calculating the Gradient of the Lagrangian for x and the Hessian

```
gradxL = simplify(gradient(L, x))
```

gradxL(x1, x2, x3, mu1, mu2) =

$$\begin{pmatrix} 2x_1 \left(\mu_2 - 8\mu_1 + 1\right) \\ 2x_2 - 24\mu_1 + 4\mu_2 x_2 \\ 15\mu_1 - 2x_3 + 8\mu_2 x_3 \end{pmatrix}$$

We then set the ktt criterion by assuming both ktt variables must to be non-negative and solve with complementary slackness. We get a total of 8 candidate points out of this approach.

```
assume(mu1 >= 0 & mu2 >= 0)
kkt = solve(gradxL == 0, mu1*(h1-b1) == 0, mu2*(h2-b2) == 0, xmu);
double([kkt.x1, kkt.x2, kkt.x3, kkt.mu1, kkt.mu2].')
```

```
ans = 5×8
  -3.5424    3.5424         0         0         0         0   -3.6583    3.6583
   0.9200    0.9200         0    8.8205         0         0    1.5000    1.5000
  -0.4356   -0.4356         0    5.5128   -1.9365    1.9365    0.9375    0.9375
   0.3383    0.3383         0    0.7350         0         0    0.1250    0.1250
   1.7061    1.7061         0         0    0.2500    0.2500         0         0
```

The candidate points are stored seperately like in Problem 1.

```
xA = double([kkt.x1(1), kkt.x2(1), kkt.x3(1)]);
muA = double([kkt.mu1(1), kkt.mu2(1)]);

xB = double([kkt.x1(2), kkt.x2(2), kkt.x3(2)]);
muB = double([kkt.mu1(2), kkt.mu2(2)]);

xC = double([kkt.x1(3), kkt.x2(3), kkt.x3(3)]);
muC = double([kkt.mu1(3), kkt.mu2(3)]);
```

```
xD = double([kkt.x1(4), kkt.x2(4), kkt.x3(4)]);
muD = double([kkt.mu1(4), kkt.mu2(4)]);

xE = double([kkt.x1(5), kkt.x2(5), kkt.x3(5)]);
muE = double([kkt.mu1(5), kkt.mu2(5)]);

xF = double([kkt.x1(6), kkt.x2(6), kkt.x3(6)]);
muF = double([kkt.mu1(6), kkt.mu2(6)]);

xG = double([kkt.x1(7), kkt.x2(7), kkt.x3(7)]);
muG = double([kkt.mu1(7), kkt.mu2(7)]);

xH = double([kkt.x1(8), kkt.x2(8), kkt.x3(8)]);
muH = double([kkt.mu1(8), kkt.mu2(8)]);
```

We then have to check the feasibility of the candidate points by introducing the x values of the candidate points into the constraints and check if they are fullfilled. Assuming very small values equal to zero, we yield only Point C unfeasable with a non-negative value (e.g. not fullfilling h2).

```
double(subs([h1 − b1, h2 − b2], x, xA))
```

ans = 1×2
10⁻¹⁵ ×
    0.5331   0.0042

```
double(subs([h1 − b1, h2 − b2], x, xB))
```

ans = 1×2
10⁻¹⁵ ×
    0.5331   0.0042

```
double(subs([h1 − b1, h2 − b2], x, xC))
```

ans = 1×2
  −129   15

```
double(subs([h1 − b1, h2 − b2], x, xD))
```

ans = 1×2
       0 −262.1677

```
double(subs([h1 − b1, h2 − b2], x, xE))
```

ans = 1×2
  −99.9526   0

```
double(subs([h1 − b1, h2 − b2], x, xF))
```

ans = 1×2
 −158.0474   0

```
double(subs([h1 − b1, h2 − b2], x, xG))
```

```
ans = 1×2
        0    -6.3984
```

```
double(subs([h1 - b1, h2 - b2], x, xH))
```

```
ans = 1×2
        0    -6.3984
```

b) Evaluate 2nd order Conditions to Evaluate for local or global maximum

With all points except Point C identified as feasible, we continue with evaluating which candidate points correspond to a local or global maximum. First, we set up the Hessian and notice its indipendence of x, which will give us later insights into the globality of maxima.

```
H = hessian(L, x)
```

H(x1, x2, x3, mu1, mu2) =

$$\begin{pmatrix} 2\,\mu_2 - 16\,\mu_1 + 2 & 0 & 0 \\ 0 & 4\,\mu_2 + 2 & 0 \\ 0 & 0 & 8\,\mu_2 - 2 \end{pmatrix}$$

We further define the gradients of the Constraints.

```
gradh1 = simplify(gradient(h1, x))
```

gradh1(x1, x2, x3) =

$$\begin{pmatrix} 16\,x_1 \\ 24 \\ -15 \end{pmatrix}$$

```
gradh2 = simplify(gradient(h2, x))
```

gradh2(x1, x2, x3) =

$$\begin{pmatrix} -2\,x_1 \\ -4\,x_2 \\ -8\,x_3 \end{pmatrix}$$

```
% variables for tangent direction analysis
syms d1 d2 d3 real
d = [d1 d2 d3].';
```

Point A

The Hessian of point A is positive semi-definite implying a potential mininmum. As we are only interested in Maximum, we do not pursue the candidate further.

```
HA = subs(H, mu, muA);
double(eig(HA))
```

```
ans = 3×1
        0
```

11

```
     8.8242
    11.6485
```

Point B

Like for Point A, the Hessian of point B is positive semi-definite implying a potential mininmum. As we are only interested in Maximum, we again do not pursue the candidate further.

```
HB = subs(H, mu, muB);
double(eig(HB))
```

```
ans = 3×1
         0
    8.8242
   11.6485
```

Point C

As point C is not feasible (shown in part a above), we do not investigate.

Point D

The Hessian of Point D is indefinite, giving a potential local optimum.

```
HD = subs(H, mu, muD);
double(eig(HD))
```

```
ans = 3×1
   -9.7607
   -2.0000
    2.0000
```

We thus apply a tangent direction analysis

```
assume(d,'clear');
subs(gradh1, x, xE).' * d
```

ans(x1, x2, x3) $= 24\,d_2 - 15\,d_3$

```
subs(gradh2, x, xE).' * d
```

ans(x1, x2, x3) $= 4\,\sqrt{15}\,d_3$

and find the gradient to be negative in all directions, implying a local maximum.

```
assume(24*d2 - 15*d3 == 0 &  4*sqrt(sym(15))*d3 ==0)
simplify(d.' * HD * d)
```

ans(x1, x2, x3, mu1, mu2) =

$$-\frac{1142\,d_1{}^2}{117}$$

With the only solution to the system the zero vector, the point thus fulfills the SC-2 condition for a Local Maximum.

```
solve(d3==0,d)
```

```
ans = struct with fields:
    d1: 0
    d2: 0
    d3: 0
```

Point E

The Hessian of candidate point E is positive semi-definite implies a potential mininmum. As we are only interested in maxima, we again do not pursue the candidate further.

```
HE = subs(H, mu, muE);
double(eig(HE))
```

```
ans = 3×1
         0
    2.5000
    3.0000
```

Point F

The Hessian of candidate point F is positive semi-definite implies a potential mininmum. As we are only interested in maxima, we again do not pursue the candidate further.

```
HF = subs(H, mu, muF);
double(eig(HF))
```

```
ans = 3×1
         0
    2.5000
    3.0000
```

Point G

The Hessian of candidate point G is indefinite, thus potentially giving a local optima.

```
HG = subs(H, mu, muG);
double(eig(HG))
```

```
ans = 3×1
    -2
     0
     2
```

We thus apply a tangent direction analysis

```
assume(d,'clear');
```

```
subs(gradh1, x, xG).' * d
```

ans(x1, x2, x3) = $24\,d_2 - 15\,d_3 - \sqrt{2}\ \sqrt{1713}\ d_1$

```
subs(gradh2, x, xG).' * d
```

ans(x1, x2, x3) =

$$\frac{\sqrt{2}\ \sqrt{1713}\ d_1}{8} - \frac{15\,d_3}{2} - 6\,d_2$$

and find the gradient to be positive in all directions, implying a local minimum. As we are only searching for maxima, candidate point G can thus be discarded.

```
assume(24*d2 - 15*d3 - sqrt(sym(2))*sqrt(sym(1713))*d1 == 0 &
(sqrt(sym(2))*sqrt(sym(1713))*d1)/8 - (15*d3)/2 - 6*d2 ==0)
simplify(d.' * HG * d)
```

ans(x1, x2, x3, mu1, mu2) =

$$\frac{561\,d_3{}^2}{32}$$

Point H

Like for pointG, the Hessian of candidate point G is indefinite, thus potentially giving a local optima.

```
HH = subs(H, mu, muH);
double(eig(HH))
```

ans = 3×1
    −2
     0
     2

We thus apply a tangent direction analysis

```
assume(d,'clear');
subs(gradh1, x, xH).' * d
```

ans(x1, x2, x3) = $24\,d_2 - 15\,d_3 + \sqrt{2}\ \sqrt{1713}\ d_1$

```
subs(gradh2, x, xH).' * d
```

ans(x1, x2, x3) =

$$-6\,d_2 - \frac{15\,d_3}{2} - \frac{\sqrt{2}\ \sqrt{1713}\ d_1}{8}$$

and find the gradient to be positive in all directions, implying a local minimum. As we are only searching for maxima, candidate point H like candidate point G previously, can thus be discarded.

```
assume(24*d2 - 15*d3 + sqrt(sym(2))*sqrt(sym(1713))*d1 == 0 &  - 6*d2 -
(15*d3)/2 - (sqrt(sym(2))*sqrt(sym(1713))*d1)/8 ==0)
```

```
simplify(d.' * HH * d)
```

ans(x1, x2, x3, mu1, mu2) =

$$\frac{561\, d_3{}^2}{32}$$

Conclusion

From all 8 candidate points, we thus conclude only Point D to fullfill the NC2 and SC2 Local Maxima Criteria. Point D's function value can be computed as shown below and can be later used to be compared with the gurbi python result.

```
fD = double(subs(f, x, xD))
```

fD =
47.4103