Optimization - Exercise session 8
Optimality conditions in the constrained case

(I) Minimization problem with equality constraints:

$$\min_{x \in \mathbb{R}^n} f(x)$$

s.t.

$$h_i(x) = 0 \quad \forall i \in \mathcal{E}.$$

*To identify a set of feasible points including **all local minima and possibly other points that are not local minima** (any global minimum is necessarily found among these local minima) we need*

1. Write the Lagrangian function:

$$L(x, y, \lambda) = f(x) + \sum_{i \in \mathcal{E}} \lambda_i h_i(x).$$

2. Identify all the solutions of the following system of equations

$$\nabla_x L(x, \lambda) = 0$$

and

$$\nabla_\lambda L(x, \lambda) = 0.$$

3. Add to this the feasible solutions where the gradients of the constraints ($\nabla_x h_i(x)$) are linearly **dependent**.
4. If the problem consists of minimizing a continuous function over a compact set (***bounded and closed***) it follows by the Weierstrass theorem that it has a global optimal solution.

(II) Minimization problem with equality and inequality constraints:

$$\min_{x \in \mathbb{R}^n} f(x)$$

s.t.

$$h_i(x) = 0 \quad \forall i \in \mathcal{E} \quad \text{and} \quad h_i(x) \leq 0 \quad \forall i \in \mathcal{I}.$$

*To identify a set of feasible points including **all local minima and possibly other points that are not local minima** (any global minimum is necessarily found among these local minima) we need*

1. Write the Lagrangian function:

$$L(x, y, \lambda) = f(x) + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i h_i(x).$$

2. Identify all solutions satisfying KKT and initial conditions:

$$\begin{cases} \nabla_x L(x, \lambda) = \mathbf{0}; \\ \lambda_i \geq 0 \quad i \in \mathcal{I}; \\ \lambda_i h_i(x) = 0 \quad i \in \mathcal{I}; \\ h_i(x) \geq 0 \quad i \in \mathcal{I}; \\ h_i(x) = 0 \quad i \in \mathcal{E}. \end{cases}$$

3. Add to this the feasible solutions where the gradients of ***the active constraints*** are linearly dependent. A constraint i is said to be active at x^* if $h_i(x^*) = 0$.
4. If the problem consists of minimizing a continuous function over a compact set (***bounded and closed***) it follows by the Weierstrass theorem that it has a global optimal solution.

1. Among the points on the parabola with equation $y = \frac{1}{5}(x-3)^2$, we want to determine which is closest to the point $(x, y) = (3, 2)$, which we formulate as the following optimization problem

$$\min f(x, y),$$

such that

$$(x-3)^2 = 5y,$$

where $f(x, y) = (x-3)^2 + (y-2)^2$.

(a) Why is the *square* of the distance to the point $(3, 2)$ used as the objective function?

(b) Find all the points satisfying the (necessary) first-order optimality condition (taking care to check the gradient independence condition), and derive the solution to the problem. Answer again, this time using the parabola with equation $y = \frac{1}{3}(x-3)^2$.

(c) Could this problem be solved by eliminating x using the constraint and then solving the resulting unconstrained problem? And by eliminating y ?

Solution

(a) We use the *square* of the distance to the point $(3, 2)$ to facilitate the computation of the derivatives and it does not change the stationary points but make the computation much harder without the square.

(b_1) First of all we need to construct the Lagrangian:

$$L(x, y, \lambda) = (x-3)^2 + (y-2)^2 - \lambda[(x-3)^2 - 5y].$$

The first-order optimality condition is

$$\nabla L(x, y, \lambda) = \mathbf{0}.$$

Consequently,

$$\begin{cases} 2(x-3) - 2\lambda(x-3) = 0; \\ 2(y-2)^2 + 5\lambda = 0. \end{cases}$$

Using the constraint of optimization problem we get the following system of non linear equations

$$\begin{cases} (x-3)(1-\lambda) = 0; & (1) \\ 2(y-2)^2 = -5\lambda; & (2) \\ (x-3)^2 - 5y = 0. & (3) \end{cases}.$$

If $\lambda = 1$ then using (2) we have $y = -\frac{1}{2}$ and using (3) we get $(x-3)^2 = -\frac{5}{2}$. It is impossible for real numbers.

Hence, $\lambda \neq 1$. Consequently, $x = 3$. Takin into account (3) and (2) we obtain $y = 0, \lambda = \frac{4}{5}$.

Therefor, $(3, 0)$ satisfies all conditions above.

It is clear that

$$\nabla_{x,y}(x-3)^2 - 5y = (2(x-3), -5) \neq (0, 0).$$

Consequently, $\nabla_{x,y}(x-3)^2 - 5y$ is lin. independent for any x. Therefore, there is only one KKT point: $(3, 0)$.

(b_2) First of all we need to construct the Lagrangian:

$$L_(x, y, \lambda) = (x-3)^2 + (y-2)^2 - \lambda[(x-3)^2 - 3y].$$

A stationary point (or critical point) of L is a point x^* satisfying the first-order optimality condition:

$$\nabla L(x, y, \lambda) = \mathbf{0}.$$

Consequently,

$$\begin{cases} 2(x-3) - 2\lambda(x-3) = 0; \\ 2(y-2)^2 + 3\lambda = 0. \end{cases}$$

Using the constraint of optimization problem we get the following system of non linear equations

$$\begin{cases} (x-3)(1-\lambda) = 0; & (1) \\ 2(y-2) = -3\lambda; & (2) \\ (x-3)^2 - 3y = 0. & (3) \end{cases}.$$

If $\lambda = 1$ then using (2) we have $y = \frac{1}{2}$ and using (3) we get $(x-3)^2 = \frac{3}{2}$. Consequently, $x = 3 \pm \frac{\sqrt{6}}{2}$. Hence, we have

$$\left(3 - \frac{\sqrt{6}}{2}\right), \quad \left(3 - \frac{\sqrt{6}}{2}\right).$$

If $\lambda \neq 1$. Consequently, $x = 3$. Takin into account (3) and (2) we obtain $y = 0, \lambda = \frac{4}{3}$.
Therefor, we have one extreme points

$$\left(3 - \frac{\sqrt{6}}{2}, \frac{1}{2}\right), \quad \left(3 - \frac{\sqrt{6}}{2}, \frac{1}{2}\right), \quad (3,0)$$

Obviously,

$$f(3,0) = 4, \quad f\left(3 - \frac{\sqrt{6}}{2}\right) = \frac{15}{4} < f(3,0).$$

Consequently, at the following 2 points might be achieved minimum

$$\left(3 - \frac{\sqrt{6}}{2}, \frac{1}{2}\right), \quad \left(3 - \frac{\sqrt{6}}{2}, \frac{1}{2}\right)$$

(c) If $(x-3)^2 = 5y$ then we have

$$\min 5y + (y-2)^2.$$

The first order necessary condition is

$$(5y + (y-2)^2)' = 0.$$

Consequently,

$$y = -\frac{1}{2}.$$

It is impossible since $(x-3)^2 = 5y \geq 0$.
Hence, we have to solve the following problem

$$\min 5y + (y-2)^2.$$

s.t. $y \geq 0$.
If eliminating y, then we get

$$\min (x-3)^2 + \left(\frac{(x-3)^2}{5} - 2\right)^2.$$

The first order necessary condition is

$$\left((x-3)^2 + \left(\frac{(x-3)^2}{5} - 2\right)^2\right)' = 0.$$

Consequently, we obtain

$$x = 3, \quad y = 0.$$

2. On the $\mathbb{R}^2$ domain defined by the equations

$$\begin{cases} x_1 \geq -\frac{3}{2} \\ x_1 \geq x_2 \\ x_1^2 + x_2^2 \leq 4 \end{cases},$$

find *all* local minima (with justification) of the functions
(a) $f(x) = (x_1 - 1)^2$
(b) $f(x) = -\left((x_1 - 1)^2 + (x_2 - 1)^2\right)$
(c) $f(x) = (x_1 - 1)^2 + 2(x_2 - 1)^2$.

Solution

(a) First of all we need to construct the Lagrangian:

$$L(x, y, \lambda) = (x_1 - 1)^2 - \lambda_1 \left(x_1 + \frac{3}{2}\right) - \lambda_2(x_1 - x_2) - \lambda_3(4 - x_1^2 - x_2^2).$$

KKT conditions:

$$\begin{cases} \nabla L(x_1, x_2) = \begin{pmatrix} 2(x_1 - 1) - \lambda_1 - \lambda_2 + 2\lambda_3 x_1 \\ \lambda_2 + 2\lambda_3 x_2 \end{pmatrix} = \mathbf{0}; \\ \lambda_1(x_1 + 3/2) = 0; \\ \lambda_2(x_1 - x_2) = 0; \\ \lambda_3(4 - x_1^2 - x_2^2) = 0; \\ \lambda_1, \lambda_2, \lambda_3 \geq 0; \\ x_1 + 3/2 \geq 0, \quad x_1 - x_2 \geq 0, \quad 4 - x_1^2 - x_2^2 \geq 0. \end{cases}$$

If $\lambda_1 > 0$ then $x_1 = -\frac{3}{2}$.

$$4 - \left(-\frac{3}{2}\right)^2 - x_2^2 \leq 0.$$

Consequently,

$$\frac{7}{4} - x_2^2 \geq 0 \quad \Leftarrow$$

$$-\frac{3}{2} \leq -\frac{\sqrt{7}}{2} \leq x_2 \leq \frac{\sqrt{7}}{2}.$$

Since $x_1 \geq x_2$, $x_1 = -\frac{3}{2}$ is not admissible. Therefore, $\lambda_1 = 0$.
If $x_3 = 0$. Then $\lambda_2 = 0$, $x_1 = 1$, and $x_2 \leq 1, x_2^2 \leq 3$.
Consequently, $-\sqrt{3} \leq x_2 \leq 1$.
Hence, we have $\lambda_1 = \lambda_3 = 0$, $\{(1, \gamma) : \quad -\sqrt{3} \leq \gamma \leq 1\}$ and $\lambda_2 = 0$.
If $\lambda_3 \neq 0$ then

$$4 - x_1^2 - x_2^2 = 0 \quad \Rightarrow \quad x_1^2 + x_2^2 = 4.$$

If $\lambda_2 = 0$ then $\lambda_2 + 2\lambda_3 x_2 = 0 \Rightarrow x_2 = 0$.
Since $x_1^2 + x_2^2 = 4$ we get $x_1 = \pm 2$. Using that $2(x_1 - 1) - \lambda_1 - \lambda_2 + 2\lambda_3 x_1 = 0$ we obtain

$$\lambda_3 = -\frac{2(x_1 - 1)}{x_1} = -1 + \frac{1}{x_1}.$$

Consequently,

$$x_1 = 2, \lambda_3 = -\frac{1}{2} \quad \text{(impossible)}$$

$$x_1 = -2, \lambda_3 = -\frac{3}{2} \quad \text{(impossible)}$$

Therefore,

$$\lambda_1 = 0, \lambda_3 > 0, \lambda_2 = 0, s = \varnothing$$

If $\lambda_2 \neq 0$ ($\lambda_2 > 0$). Then

$$x_1 = x_2 \quad \text{and} \quad x_1^2 + x_2^2 = 4.$$

Consequently,

$$2x_1^2 = 4 \Rightarrow$$
$$x_1 = \pm\sqrt{2}.$$

Hence,

$$f(\sqrt{2}, \sqrt{2}) = 0.17, \quad f(-\sqrt{2}, -\sqrt{2}) = 5.8.$$

Therefore, local minima of the function might attend at the points from

$$\{(1,\gamma) : \quad -\sqrt{3} \leq \gamma \leq 1\} \cup \{\sqrt{2}, \sqrt{2}\}.$$

Since $f(x) \geq 0 \quad \forall x$ and $f(1,\gamma) = 0$ we get that

$$\{(1,\gamma) : \quad -\sqrt{3} \leq \gamma \leq 1\}$$

is a solution of the optimization problem.

(b) First of all we need to construct the Lagrangian:

$$L(x,y,\lambda) = -(x_1-1)^2 - (x_2-1)^2 - \lambda_1\left(x_1 + \frac{3}{2}\right) - \lambda_2(x_1 - x_2) - \lambda_3(4 - x_1^2 - x_2^2).$$

KKT conditions:

$$\begin{cases} \nabla L(x_1,x_2) = \begin{pmatrix} -2(x_1-1) - \lambda_1 - \lambda_2 + 2\lambda_3 x_1 \\ -2(x_2-1) + \lambda_2 + \lambda_3 x_2 \\ \lambda_2 + 2\lambda_3 x_2 \end{pmatrix} = \mathbf{0}; \\ \lambda_1(x_1 + 3/2) = 0; \\ \lambda_2(x_1 - x_2) = 0; \\ \lambda_3(4 - x_1^2 - x_2^2) = 0; \\ \lambda_1, \lambda_2, \lambda_3 \geq 0; \\ x_1 + 3/2 \geq 0, \quad x_1 - x_2 \geq 0, \quad 4 - x_1^2 - x_2^2 \geq 0. \end{cases}$$

If $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$ then $x_1 = 1$ and $x_2 = 1$.

$$4 - \left(-\frac{3}{2}\right)^2 - x_2^2 \leq 0.$$

If $\lambda_1 = 0$, $\lambda_2 > 0$, $\lambda_3 = 0$ then $x_1 = x_2$. Consequently,

$$\begin{cases} -2(x_1-1) - \lambda_2 = 0 \\ -2(x_2-1) + \lambda_2 = 0 \end{cases}$$

Hence, $\lambda_2 = 0$. (Impossible)

If $\lambda_1 = 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ then $x_1 = x_2$ and

$$x_1^2 + x_2^2 = 4.$$

Consequently, $x_1 = x_2 = \pm\sqrt{2}$.

If $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 > 0$ then

$$\begin{cases} -2(x_1-1) = -2\lambda_3 x_1 \\ -2(x_2-1) = -2\lambda_3 x_2 \end{cases}$$

Consequently, $x_1 = x_2$ and

$$x_1^2 + x_2^2 = 4.$$

Hence, $x_1 = x_2 = \pm\sqrt{2}$.

Therefore, we have 3 points:

$$(1,1), \quad (-\sqrt{2}, -\sqrt{2}) \quad f(\sqrt{2}, \sqrt{2}).$$

Since

$$f(1,1) = 0, \quad f(-\sqrt{2}, -\sqrt{2}) = -11,6567 \quad \text{and} \quad f(\sqrt{2}, \sqrt{2}) = -0.34$$

and problem consists of minimizing a continuous function over a compact set (***bounded and closed***) it follows by the Weierstrass theorem that it has a global optimal solution $(-\sqrt{2}, -\sqrt{2})$.

3. Consider the optimization problem

$$\min x_1 + x_2$$

s.t.

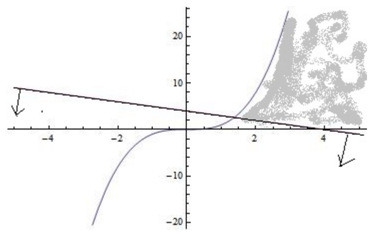
$$x_1^3 \geq x_2$$

$$x_2 \geq 0$$

(a) What is the optimal solution to this problem?
(b) Write and solve the KKT conditions for this problem. Comment.

Solution

(a) Optimal solution is $(0, 0)$



(b) First of all we need to construct the Lagrangian:

$$L(x_1, x_2, \lambda) = x_1 + x_2 - \lambda_1 \left(x_1^3 - x_2\right) - \lambda_2 x_2.$$

KKT conditions:

$$\begin{cases} \nabla L(x_1, x_2) = \begin{pmatrix} 1 - 3\lambda_1 x_1^2 \\ 1 + \lambda_1 - \lambda_2 \end{pmatrix} = \mathbf{0}; \\ \lambda_1(x_1^3 - x_2) = 0; \\ \lambda_2 x_2 = 0; \\ \lambda_1, \lambda_2 \geq 0; \\ x_1^3 - x_2 \geq 0, \quad x_2 \geq 0. \end{cases}$$

$\lambda_1 = 0$ and $\lambda_2 = 0$ (impossible)
$\lambda_1 = 0$, $\lambda_2 > 0$ (impossible)
If $\lambda_1 > 0$, $\lambda_2 = 0$ then

$$x_1^3 = x_2$$

$$\lambda_1 = -1.$$

(impossible)
If $\lambda_1 > 0$, $\lambda_2 > 0$ then

$$x_1^3 = x_2$$

$$x_2 = 0.$$

Consequently, $x_1 = 0$. (impossible)
Let's check LICQ. We have

$$\nabla h_1(x_1, x_2) = \nabla h_1(x_1^3 - x_2) = [3x_1^2, -1]^T$$

$$\nabla h_2(x_1, x_2) = \nabla x_2 = [0, 1]^T$$

Obviously,

$$\nabla h_1(0, 0) = [0, -1]^T$$

$$\nabla h_2(0, 0) = [0, 1]^T$$

are linear dependent. Consequently, $(0, 0)$ is might be extreme point. Note that $(0, 0)$ does not satisfy KKT conditions.

4. (a) Prove that if x^* is a point satisfying the KKT conditions and the gradient independence condition, then the corresponding vector of Lagrange multipliers λ^* is unique.

(b) Show that the reformulation of an equality $g(x) = 0$ in the equivalent form $g^2(x) = 0$ is not a good idea from the point of view of optimal conditions.

Solution

(a) Suppose that x^* is a point satisfying the KKT conditions and the gradient independence condition and we have two different Lagrange multipliers λ_i and λ_i'. Consequently,

$$\nabla_x f(x^*) = \sum_i \lambda_i \nabla_x h_i(x^*)$$

$$\nabla_x f(x^*) = \sum_i \lambda_i' \nabla_x h_i(x^*)$$

Hence,

$$\sum_i \lambda_i \nabla_x h_i(x^*) - \sum_i \lambda_i' \nabla_x h_i(x^*) = \sum_i (\lambda_i - \lambda_i') \nabla_x h_i(x^*).$$

Since $\{\nabla_x h_i(x^*)\}$ is LI, we have $\lambda_i - \lambda_i' = 0, \quad \forall i$. Therefor, $\lambda_i = \lambda_i' \quad \forall i$.

(b) If we replace $g(x) = 0$ by the equivalent form $g^2(x) = 0$ then we will have

$$\nabla(g^2(x)) = 2g(x)\nabla g(x) = 0.$$

Hence, the LICQ condition is not satisfied.

5. Optimality conditions for some standard problems. Derive, for each of the problem classes listed below, the conditions for first-order optimality. Specify any necessary assumptions.

When can these conditions be simplified? When can they be solved analytically? Where possible, formulate the conditions obtained in vector or matrix form.

These problem classes are defined on the basis of the following data:

c is a vector of $\mathbb{R}^n$,

b, l and u are vectors of $\mathbb{R}^m$ (and we suppose $l < u$)

A is a matrix of size $m \times n$ whose lines are named $\{a_i^T\}_{1 \le i \le m}$

Q is a symmetric square matrix of size n.

The functions f and g are defined everywhere on Rn and the n functions $h_1, h_2, \ldots, h_n$ are defined on $\mathbb{R}^n$; all are assumed to be sufficiently differentiable. Inequalities between vectors are to be interpreted component by component.

(a) Unconstrained quadratic optimization

$$\min x^2 + 4xy + 5y^2 + 3x - 5y$$

then in general

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$$

and quadratic optimization under linear equality constraints

$$\min x^2 + 4xy + 5y^2 + 3x - 5y$$

such that

$$x - y = 1$$

then in general

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$$

such that

$$Ax = b$$

(where $Ax = b$ is equivalent to m equality constraints $a_i^T x = b_i$ for $1 \le i \le m$).

Solution

$$\nabla f(x) = 0 \Rightarrow \begin{cases} 2x + 4y + 3 = 0; \\ 4x + 1 - y - 5 = 0. \end{cases} \Rightarrow$$

$$y = \frac{11}{2}, \quad x = -\frac{25}{2}.$$

Hence, we have the following extreme

$$\left(-\frac{25}{2}, \frac{11}{2}\right).$$

If

$$f(x) = \frac{1}{2} x^T Q x + c^T x$$

then

$$\nabla f(x) = Qx + c = 0.$$

Consequently,

$$x^* = -Q^{-1} c.$$

It easy to show that

$$x^2 + 4xy + 5y^2 + 3x - 5y = x(x + 2y) + y(2y + 5y) + [3, -5] \begin{matrix} x \\ y \end{matrix} = [x, y] \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} + [3, -5] \begin{matrix} x \\ y \end{matrix} = \frac{1}{2} x^T Q x + c^T x,$$

where

$$Q = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}.$$

Consequently,

$$x^* = -Q^{-1}c = \left(-\frac{25}{2}, \frac{11}{2}\right).$$

If

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}x^T Qx + c^T x$$

such that

$$Ax = b,$$

then

$$\begin{cases} \nabla_x L = 0; \\ Ax = b, \end{cases}$$

where

$$L = \frac{1}{2}x^T Qx + c^T x - \lambda^T(Ax - b), \quad \nabla_x L = Qx + c - A^T\lambda = 0.$$

Consequently,

$$Qx - A^T\lambda = -c.$$

Using the constraints, we get

$$\begin{cases} Qx - A^T\lambda = -c \\ Ax = b \end{cases}.$$

We have

$$\begin{pmatrix} Q & -A^T \\ A & 0 \end{pmatrix}$$

is invertible $\Leftrightarrow$ Q is invertible and $rgA = m$.

(b) Linear optimization on the boundary of a ball of radius R and center u.

$$\min 3x + 4y$$

such that

$$\|(x, y) - (2, -1)\| = 5$$

then in general

$$\min_{x\in\mathbb{R}^n} c^T x$$

such that

$$\|x - u\| = R$$

then on the whole ball

$$\min 3x + 4y$$

such that

$$\|(x, y) - (2, -1)\| \leq 5$$

then in general min

$$\min_{x\in\mathbb{R}^n} c^T x$$

such that

$$\|x - u\| = R$$

(a reformulation may be necessary).

Solution

Our problem is equivalent to

$$\min 3x + 4y$$

such that

$$\|(x, y) - (2, -1)\|^2 = 25$$

Necessary condition:

$$\begin{cases} 3 = 2\lambda(x-2) \\ 4 = 2\lambda(y+1) \\ (x-2)^2 + (y+1)^2 = 25 \end{cases}$$

We have

$$x - 2 = \frac{3}{2\lambda}, \quad y + 1 = \frac{2}{\lambda}.$$

Since $(x-2)^2 + (y+1)^2 = 25$ we get

$$4\lambda^2 = 1.$$

Consequently, $\lambda = \pm\frac{1}{2}$. Hence, we get 2 points

$$(\frac{1}{2}, 5, 3)$$

$$(-\frac{1}{2}, -1, -5)$$

and

$$f(5,3) = 27, \quad f(-1,5) = -23.$$

Consequently, solution is equal to $(-1,-5)$.

In general we will have

$$L(x, y, \lambda) = c^T x - \lambda(\|x-u\| - R).$$

Necessary conditions:

$$\begin{cases} c - \frac{\lambda(x-u)}{\|x-u\|} = 0 \\ \|x-u\| = R \end{cases}$$

Consequently,

$$\begin{cases} c - \frac{\lambda(x-u)}{R} = 0 \\ \|x-u\| = R \end{cases}$$

Hence,

$$x = \frac{1}{\lambda} cR + u$$

Using that $\|x-u\| = R$ we get

$$\frac{R}{|\lambda|}\|c\| = R.$$

Consequently, $\lambda = \pm\|c\|$.

Hence,

$$x_\pm = u \pm \frac{c}{\|c\|} R$$

and

$$f(x_+) = c^T u + \|c\|R, \quad f(x_-) = c^T u - \|c\|R.$$

Therefore, we get $c^T u - \|c\|R$.