

MATH-329 Nonlinear optimization

Exercise session 10: KKT

Instructor: Nicolas Boumal

TAs: Andreea Musat, Andrew McRae

Document compiled on November 20, 2024

1. KKT. Consider the set

$$S = \{(x, y) \mid 0 \leq x \leq 2\pi \text{ and } 0 \leq y \leq \sin(x) + 2\}$$

and the cost function $f(x, y) = -y$ that we want to minimize.

1. Draw the search space.
2. Is the problem convex?
3. Do CQs hold globally?
4. Write down the KKT conditions.
5. Find all KKT points and all stationary points.
6. Find all local and global minima.
7. How does your answer change if the set is

$$S = \{(x, y) \mid 0 \leq x \leq 2\pi \text{ and } 0 \leq y \leq \sin(x) + 1\}$$

instead? In particular, is LICQ or MFCQ holding at the point $(3\pi/2, 0)$?

To train you may want to revisit this exercise for a general linear cost function $f(x, y) = x \cos(\theta) + y \sin(\theta)$ with parameter $\theta \in [0, 2\pi[$.

Answer.

1. See Figure 1.
2. The problem is not convex because S is not convex.
3. CGs hold globally. We can check this using LICQ.
4. We rewrite the set as $S = \{(x, y) \in \mathbb{R}^2 \mid g_i(x, y) \leq 0 \text{ for } i = 1, 2, 3, 4\}$ where

$$\begin{aligned} g_1(x, y) &= -x & g_2(x, y) &= x - 2\pi \\ g_3(x, y) &= -y & g_4(x, y) &= y - \sin(x) - 2. \end{aligned}$$

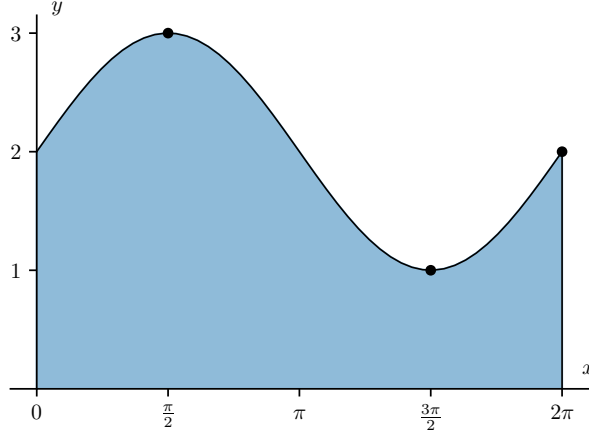


Figure 1: Search space S .

We find

$$\begin{aligned} \nabla g_1(x, y) &= \begin{bmatrix} -1 \\ 0 \end{bmatrix} & \nabla g_2(x, y) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \nabla g_3(x, y) &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} & \nabla g_4(x, y) &= \begin{bmatrix} \cos(x) \\ 1 \end{bmatrix}. \end{aligned}$$

A point $(x, y) \in S$ is KKT if there exists $\lambda \geq 0$ such that

$$-\nabla f(x, y) = \lambda_1 \nabla g_1(x, y) + \lambda_2 \nabla g_2(x, y) + \lambda_3 \nabla g_3(x, y) + \lambda_4 \nabla g_4(x, y)$$

and $\lambda_i g_i(x, y) = 0$ for $i = 1, \dots, 4$.

5. Stationary points and KKT points coincide because CQs hold globally.

For all $(x, y) \in S$ we have $\nabla f(x, y) = [0 \ 1]^\top$. In the interior of S there is no KKT point because the gradient is always non-zero. For a point to be KKT we find that the fourth constraint must be active. This yields 3 KKT points on the boundary:

$$p_1 = \begin{bmatrix} \frac{\pi}{2} \\ 3 \end{bmatrix}, \quad p_2 = \begin{bmatrix} \frac{3\pi}{2} \\ 1 \end{bmatrix}, \quad \text{and} \quad p_3 = \begin{bmatrix} 2\pi \\ 2 \end{bmatrix}.$$

6. Local (hence global) minima are stationary. The point p_1 is clearly the unique global minimum.

The point p_3 is a local minimum. There is a neighborhood $\mathcal{U}$ of p_3 such that for all $(x, y) \in \mathcal{U} \cap S$ we have $f(x, y) \geq f(p_3)$.

The point p_2 is a saddle point (stationary but not a local minimum). To see this we define the curve $c(t) = (\frac{3\pi}{2} + t, \sin(\frac{3\pi}{2} + t) + 2)$. We find that $c(0) = p_2$, $c(t) \in S$ for t small enough, and $f(c(t)) < f(c(0))$ for small enough $t > 0$.

7. See Figure 2. Nothing changes except for the point $[\frac{3\pi}{2} \ 0]^\top$. We cannot apply LICQ and MFCQ to show that CQs hold. However if we compute the tangent cone and the cone of linearized feasible directions we can check that they are the same.

■

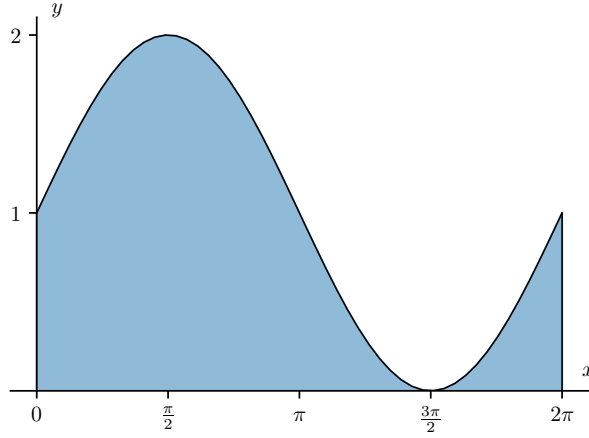


Figure 2: Search space S .

2. KKT (bis). Consider the problem

$$\min_{x,y \in \mathbb{R}} -2x + y \quad \text{subject to} \quad \begin{cases} (1-x)^3 - y \geq 0 \\ y + \frac{1}{4}x^2 - 1 \geq 0. \end{cases}$$

The optimal solution is $x^* = (0, 1)^\top$, where both constraints are active.

1. Do CQs hold at x^* ?
2. Is x^* stationary?
3. Is x^* a KKT point?

Answer the same questions for the optimization problem

$$\min_{x,y \in \mathbb{R}} y \quad \text{subject to} \quad x^2 - y^3 \leq 0,$$

where the global minimum is $x^* = (0, 0)^\top$.

Answer. See Figure 3 for the search space of the first optimization problem.

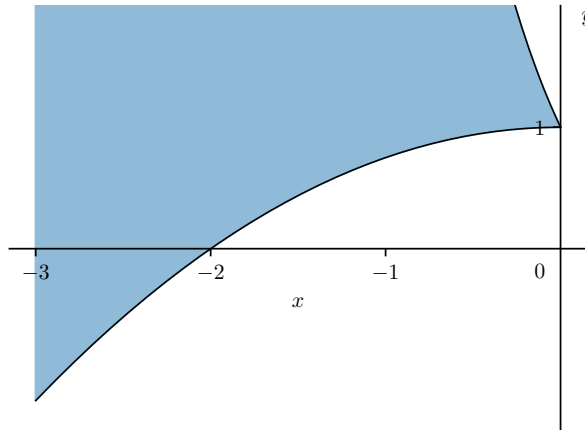


Figure 3: Search space $\{(x, y) \mid (1-x)^3 - y \geq 0 \text{ and } y + \frac{1}{4}x^2 - 1 \geq 0\}$.

1. LICQ holds at x^* .
2. The point x^* is stationary because it is a global minimum.
3. The point x^* is KKT because CQs hold at x^* .

See Figure 4 for the search space of the second optimization problem.

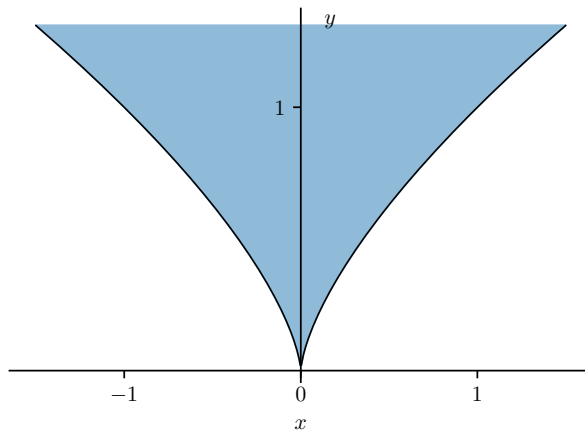


Figure 4: Search space $\{(x, y) \mid x^2 - y^3 \leq 0\}$.

1. We define $g(x, y) = x^2 - y^3$ and $S = \{(x, y) \mid g(x, y) \leq 0\}$. Then $\nabla g(x, y) = (2x, -3y^2)^\top$. The gradient at x^* is zero so we deduce that $F_{x^*}S = \mathbb{R}^2$. Clearly $T_{x^*}S$ is not $\mathbb{R}^2$ so CQs do not hold.
2. The point x^* is stationary because it is a minimum.
3. The point x^* is not KKT because $\nabla g(x^*) = 0$ and $\nabla f(x^*) \neq 0$.

■

3. Convex constraints. In the lecture we showed that the set

$$S = \{x \in \mathcal{E} \mid h_i(x) = 0, i = 1, \dots, p, \text{ and } g_i(x) \leq 0, i = 1, \dots, m\}$$

is convex if the functions $h_1, \dots, h_p$ are affine and the functions $g_1, \dots, g_m$ are convex.

1. Provide an example of a non-convex set S where all the functions $g_1, \dots, g_m$ are convex but one function h_i is not affine.
2. Provide an example of a non-convex set S where all the functions $h_1, \dots, h_p$ are affine but one function g_i is non-convex.

Answer.

1. We define

$$S = \{x \in \mathcal{E} \mid h(x) = 0 \text{ and } g(x) \leq 0\}$$

where $h(x) = \|x\|^2 - 1$ and $g(x) = x_1 + \dots + x_n$. The set S is not convex.

2. We define

$$S = \{x \in \mathcal{E} \mid h(x) = 0 \text{ and } g(x) \leq 0\}$$

where $h(x) = x_1 + \dots + x_n$ and $g(x) = 1 - \|x\|^2$. The set S is not convex. ■

4. Quadratic constraints. Sometimes quadratically constrained programs can be solved efficiently.

1. We let $S = \{x \in \mathbb{R}^n \mid x^\top Ax + b^\top x + c \leq 0\}$ where $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n, c \in \mathbb{R}$. Show that S is convex if $A \succeq 0$.
2. Suppose $A \succ 0$ and let $d \in \mathbb{R}^n$. Find a solution for the minimization problem

$$\min_{x \in S} d^\top x$$

where $S = \{x \in \mathbb{R}^n \mid x^\top Ax \leq 1\}$ and $A \succ 0$.

Answer.

1. The function $x \mapsto x^\top Ax + b^\top x + c$ is convex because $A \succeq 0$. The set S is the zero sublevel set of that map so we conclude that it is convex.
2. We find that

$$\begin{aligned} S &= \{x \in \mathbb{R}^n \mid x^\top Ax \leq 1\} \\ &= \{A^{-1/2}y \mid y \in \mathbb{R}^n, y^\top y \leq 1\} \end{aligned}$$

so the optimization problem is equivalent to

$$\min_{y \in S'} d^\top A^{-1/2}y$$

where $S' = \{y \in \mathbb{R}^n \mid y^\top y \leq 1\}$. The gradient is $g = A^{-1/2}d$ and since we minimize a linear function on the unit ball the solution is

$$y^* = -\frac{g}{\|g\|}.$$
■