OPTIMIZATION - OPTIMALISATIE

EXERCISE SESSION 5: OPTIMALITY CONDITIONS FOR UNCONSTRAINED AND CONSTRAINED OPTIMIZATION

Exercise 1 (**Optimality Conditions Unconstrained Optimization).**

Let the function

$$f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$$

be given.

(a) Compute the gradient of f.

Solution:

$$\nabla f(x) = \begin{bmatrix} 8 + 2x_1 \\ 12 - 4x_2 \end{bmatrix}$$

(b) Compute the Hessian of f.

Solution:

$$\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$$

(c) Show that f has only one stationary point.

Solution: We require $\nabla f(x) \stackrel{!}{=} 0$. We immediately see that the only stationary point is $x^* = (-4, 3)^\top$.

(d) Show that the stationary point is neither a maximum nor a minimum, but a saddle point.

Solution: A necessary optimality condition for a minimizer at the solution is that $\nabla^2 f(x)$ is positive semi-definite. Using $x = (0, 1)^\top$ yields

$$x^\top \nabla^2 f x = -4,$$

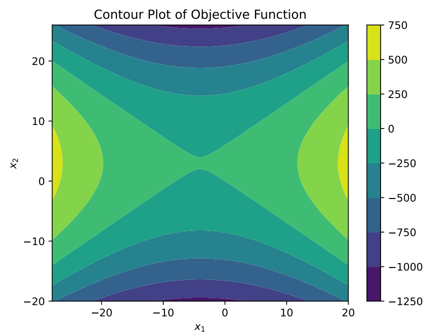
i.e., the matrix is not positive definite. A necessary optimality condition for a maximizer is that $\nabla^2 f(x)$ is negative semi-definite. Using $x = (1, 0)^\top$ yields

$$x^\top \nabla^2 f x = 2,$$

i.e., the matrix is not negative definite. We deduce that the Hessian matrix is indefinite and therefore we do not have a maximum nor a minimum. This point is called a saddle point.

(e) Sketch the contour lines of f.

Solution:



Exercise 2 (Optimality Conditions Constrained Optimization).

Regard the following minimization problem:

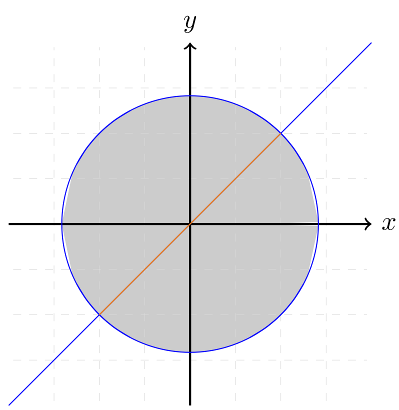
$$\min_{x\in\mathbb{R}^2} x_2^4 + (x_1+2)^4 \;\text{ s.t. }\; \begin{cases} -x_1^2 - x_2^2 \geq -8 \\ x_1 - x_2 = 0. \end{cases}$$

(2a) How many scalar decision variables, how many equality, and how many inequality constraints does this problem have?

Solution: 2 scalar decision variables, 1 equality constraint, 1 inequality constraint.

(2b) Sketch the feasible set $\Omega \subset \mathbb{R}^2$.

Solution:



The orange line shows the feasible set.

(2c) Bring this problem into the NLP standard form

$$\min_{x\in\mathbb{R}^n} f(x) \text{ s.t. } \begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases}$$

by defining the functions f, g, h appropriately.

Solution:

$$\begin{aligned} f(x) &= x_2^4 + (x_1+2)^4 \\ g(x) &= x_1^2 + x_2^2 - 8 \\ h(x) &= x_1 - x_2 \end{aligned}$$

(2d) Is this optimization problem convex? Justify.

Solution: $f(x)$ is convex, $h(x)$ is affine, $g(x)$ is convex, so the problem is convex.

(2e) Write down the Lagrangian function of this optimization problem.

Solution:

$$\begin{aligned} L(x,\lambda,\mu) &= f(x) + \lambda^\top h(x) + \mu^\top g(x) \\ &= x_2^4 + (x_1+2)^4 + \lambda\,(x_1 - x_2) + \mu\left(x_1^2 + x_2^2 - 8\right) \end{aligned}$$

where $\lambda \in \mathbb{R}, \mu \in \mathbb{R}_{\geq 0}$.

(2f) A feasible point of the problem is $\bar{x} = (2,2)^\top$. What is the active set $\mathcal{A}(\bar{x})$ at this point?

Solution: $g(\bar{x}) = 2^2 + 2^2 - 8 = 0 \Rightarrow$ the constraint is active, $\mathcal{A}(\bar{x}) = \{1\}$ (This notation interprets $g(x)$ as a vector valued function with only one dimension, i.e. a "scalar vector")

(2g) Is the linear independence constraint qualification (LICQ) satisfied at $\bar{x}$? Justify.

Solution: Check linear independence of $\nabla h(\bar{x})$ and $\nabla g_i(\bar{x}), i \in \mathcal{A}$ or whether $[\nabla h(\bar{x}) \quad \nabla g_1(\bar{x})]$ is full rank.

$$\nabla h(x) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \nabla h(\bar{x}) \quad \nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \nabla g_1(\bar{x}) = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\mathbf{det}\begin{bmatrix}\nabla g(\bar{x}) & \nabla h_1(\bar{x})\end{bmatrix} = \mathbf{det}\begin{bmatrix} 1 & 4 \\ -1 & 4 \end{bmatrix} = 8 \neq 0 \Rightarrow \text{full rank} \Rightarrow \text{LICQ satisfied.}$$

(2h) An optimal solution of the problem is $x^* = (-1,-1)^\top$. What is the active set $\mathcal{A}(x^*)$ at this point?

Solution:

$$g(x^*) = -6 < 0 \Rightarrow \mathcal{A}(x^*) = \emptyset \text{ (no active inequality constraints).}$$

(2i) Is the linear independence constraint qualification (LICQ) satisfied at x^*? Justify.

Solution:

$$\nabla h(x^*) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \text{ full rank } \Rightarrow \text{ LICQ satisfied.}$$

(2j) Describe the tangent cone $T_\Omega(x^*)$ (the set of feasible directions) to the feasible set at this point x^*, by a set definition formula with explicitly computed numbers.

Solution: LICQ holds at x^*, so the tangent cone and the linearized feasible cone coincide:

$$T_\Omega(x^*) = L_\Omega(x^*) = \left\{p \in \mathbb{R}^n \mid \nabla h_i(x^*)^\top p = 0, i = 1, \ldots, m \wedge \nabla g_i(x^*)^\top p \leq 0, i \in \mathcal{A}(x^*)\right\}$$

Here:

$$\begin{aligned} L_\Omega(x^*) &= \left\{p \in \mathbb{R}^2 \mid \nabla h(x^*)^\top p = 0\right\} = \left\{p \in \mathbb{R}^2 \mid \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}p_1\\p_2\end{bmatrix} = 0\right\} \\ &= \left\{p \in \mathbb{R}^2 \mid p_1 = p_2\right\} = \left\{t\begin{bmatrix}1\\1\end{bmatrix} \mid t \in \mathbb{R}\right\} \end{aligned}$$

(2k) Find the multiplier vectors λ^*, μ^* such that the above point x^* satisfies the KKT conditions.

Solution: General KKT conditions for inequality constraint optimization

$$\begin{aligned} \nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* &= 0 \\ h(x^*) &= 0 \\ g(x^*) &\leq 0 \\ \mu^* &\geq 0 \\ \mu_i^* g_i(x^*) &= 0, \quad i = 1, \ldots, q \end{aligned}$$

Here:

$$\begin{aligned} & h(x^*) = 0 \quad \checkmark \\ & g(x^*) < 0 \Rightarrow \underline{\mu^* = 0} \\ & \nabla_x L(x, \lambda, \mu) = \begin{bmatrix}4(x_1+2)^3 \\ 4x_2^3\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix}\lambda + \begin{bmatrix}2x_1\\2x_2\end{bmatrix}\mu \\ & \nabla_x L(x^*, \lambda^*, \mu^*) = \begin{bmatrix}4+\lambda^* \\ -4-\lambda^*\end{bmatrix} = 0 \Rightarrow \underline{\lambda^* = -4} \end{aligned}$$

(2l) Describe the critical cone $\mathcal{C}(x^*, \mu^*)$ at the point (x^*, λ^*, μ^*) in a set definition using explicitly computed numbers.

Solution:

$$\mathcal{C}(x^*, \mu^*) = \left\{ p \in \mathbb{R}^n \;\middle|\; \begin{array}{rl} & \nabla h_i(x^*)^\top p = 0, i = 1, \ldots, m \\ \wedge & \nabla g_i(x^*)^\top p = 0, i \in \mathcal{A}_+(x^*) \\ \wedge & \nabla g_i(x^*)^\top p \leq 0, i \in \mathcal{A}_0(x^*) \end{array} \right\}$$

Here ($\mathcal{A}(x^*) = \emptyset$) :

$$\mathcal{C}(x^*, \mu^*) = \left\{p \in \mathbb{R}^2 \mid \nabla h(x^*)^\top p = 0\right\} = L_\Omega(x^*) = \left\{t\begin{bmatrix}1\\1\end{bmatrix} \mid t \in \mathbb{R}\right\}$$

(2m) Formulate the second order necessary conditions for optimality (SONC) for this problem and test if they are satisfied at (x^*, λ^*, μ^*). Can you prove whether x^* is a local or even global minimizer?

Solution: SONC: Regard x^* with LICQ. If x^* is a local minimizer of the NLP, then

i. $\exists \lambda^*, \mu^*$ such that KKT conditions hold;

ii. $\forall p \in \mathcal{C}(x^*, \mu^*)$ holds $p^\top \nabla_x^2 L(x^*, \lambda^*, \mu^*) p \geq 0$.

Here:

$$\nabla_x^2 L(x, \lambda, \mu) = \begin{bmatrix} 12(x_1+2)^2 + 2\mu & 0 \\ 0 & 12x_2^2 + 2\mu \end{bmatrix}, \quad \Lambda^* := \nabla_x^2 L(x^*, \lambda^*, \mu^*) = \begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix}$$

Check SONC

i. Holds due to task Task (2k);

ii. $\Lambda^* \succ 0 \Rightarrow \forall p \in \mathbb{R}^n : p^\top \Lambda^* p \geq 0$, therefore this specifically holds also for $\forall p \in C(x^*, \mu^*)$.

$\Rightarrow$ SONC are satisfied.
Due to $\Lambda^* \succ 0$ we furthermore have $\forall p \in \mathbb{R}^n \setminus \{0\} : p^\top \Lambda^* p > 0$, and therefore specifically $\forall p \in C(x^*, \mu^*) \setminus \{0\}$. Thus SOSC also holds, and x^* is a local minimizer. Due to convexity of the NLP this is equivalent to x^* being a global minimizer.

Exercise 3 (**Degenerate Optimization Problem).** Regard the following minimization problem:

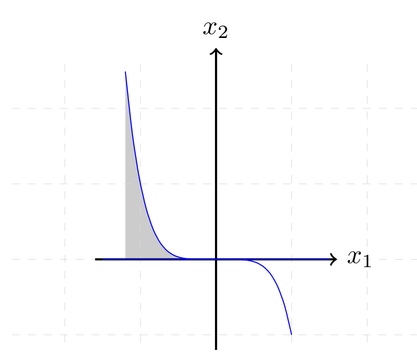
$$\min_{x \in \mathbb{R}^2} -x_1 \text{ s.t. } \begin{cases} -x_2 \leq 0 \\ x_2 + x_1^5 \leq 0. \end{cases}$$

(3a) How many scalar decision variables, how many equality, and how many inequality constraints does this problem have?

Solution: 2 scalar decision variables, 0 equality constraints, 2 inequality constraints.

(3b) Sketch the feasible set $\Omega \subset \mathbb{R}^2$.

Solution:



(3c) Determine the unique solution of the problem.

Solution: The optimal solution is $x^* = (0, 0)^\top$.

(3d) Bring this problem into the NLP standard form

$$\min_{x \in \mathbb{R}^n} f(x) \text{ s.t. } \quad \begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases}$$

by defining the functions f, g, h appropriately.

Solution:

$$f(x) = -x_1$$
$$g(x) = \begin{bmatrix} -x_2 \\ x_2 + x_1^5 \end{bmatrix}$$

(3e) Write down the Lagrangian function of this optimization problem.

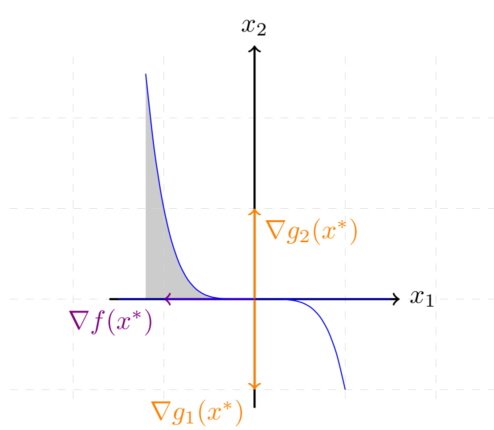
Solution:

$$\begin{aligned} L(x, \lambda, \mu) &= f(x) + \mu^\top g(x) \\ &= -x_1 - \mu_1 x_2 + \mu_2 \left(x_2 + x_1^5\right) \end{aligned}$$

where $\mu \in \mathbb{R}^2_{\geq 0}$.

(3f) Draw the gradients of the objective function and of the active constraints at the solution x^* in your sketch.

Solution:



(3g) Write down the Lagrange gradient. Does a Lagrange multiplier vector exist to satisfy $\nabla_x L(x^*, \mu^*) = 0$? Justify.

Solution: Here:

$$\nabla_x L(x^*, \mu^*) = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} \mu_1^* + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu_2^*.$$

The vector $\nabla f(x^*)$ is orthogonal to $\nabla g_1(x^*)$ and $\nabla g_2(x^*)$. Therefore, there cannot exist an optimal Lagrange multiplier fulfilling the stationarity condition.

(3h) Can a constraint qualification hold at the optimal solution? Justify.

Solution: The tangent cone at x^* is the negative x_1-axis,

$$T_\Omega(x^*) = \left\{ t \begin{bmatrix} -1 \\ 0 \end{bmatrix} \;\middle|\; t \in \mathbb{R}_{\geq 0} \right\}.$$

The cone of linearized feasible directions is the complete x_1-axis,

$$L_\Omega(x^*) = \left\{ t \begin{bmatrix} -1 \\ 0 \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}.$$

Note that in general, we have $T_\Omega(x^*) \subseteq L_\Omega(x^*)$, but in this case we do not have equality.

By definition, constraint qualifications are sufficient conditions for quasiregularity. Since $T_\Omega(x^*) \neq L_\Omega(x^*)$, x^* is not quasiregular, and no constraint qualifications can hold.

Alternatively, you could argue that x^* is a local minimizer that does not satisfy the KKT conditions, so no constraint qualifications can hold at x^*, because:

$$\left.\begin{array}{r} \text{LICQ} \Rightarrow \text{MFCQ} \Rightarrow \text{CQ} \overset{\text{(lem. 9.2)}}{\Rightarrow} x^* \text{ quasiregular} \\ x^* \text{ local minimizer} \end{array}\right\} \overset{\text{(thm. 9.3)}}{\Rightarrow} \begin{array}{c} \exists \lambda^*, \mu^* \text{ satisfying} \\ \text{KKT conditions.} \end{array}$$

Either motivation is valid, but one may be easier to verify than the other.

(3i) Bonus: Cancel the objective gradient out of the Lagrange gradient optimality condition and try to find multipliers such that the sum of the constraint gradients vanishes. Both multipliers are not allowed to be zero!
This describes a generalization of the KKT conditions the so-called Fritz-John conditions.

Solution:

$$\nabla_x L(x^*, \mu^*) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \mu_1^* + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu_2^*.$$

The multipliers are real numbers, and thus $\mu_1^* = \mu_2^* > 0$.