Optimization

Lusine Poghosyan

YSU

December 10, 2020

Example

Assume we want to find the extremum points of $f(x_1, x_2) = x_1^2 + x_2^2$ subject to $x_1^2 + 2x_2^2 = 2$. Use Lagrange's theorem to find all possible local extremum points.

Theorem (Second Order Necessary Condition SONC)

Let x^* be a local minimizer (maximizer) of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $h_i(x) = 0$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots m$, $m \leq n$. Assume f, h_i for $i = 1, \dots m$ are twice continuously differentiable functions and x^* is a regular point. Then, there exists $\lambda^* \in \mathbb{R}^m$ such that:

- $\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla h_j(x^*) = 0$ *or* $\nabla \mathcal{L}(x^*, \lambda^*) = 0$,
- *if* $TS(x^*) = \left\{y \in \mathbb{R}^n : y^T \nabla h_i(x^*) = 0, \text{for } i = 1, \dots, m\right\}$, *then for all* $y \in TS(x^*)$ *we have that* $y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y \geq 0$ ($y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y \leq 0$).

Example

Consider the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & -x_2(x_1 + x_3) \\ \text{subject to} \quad & x_1 + x_2 = 0, \\ & x_2 + x_3 = 0. \end{aligned}$$

Is it possible that $x^* = [0,0,0]^T$ is a maximum point of the problem.

Theorem (Second Order Sufficient Condition SOSC)

Suppose f, h_i, for $i = 1, \dots m$ are twice continuously differentiable functions on $\mathbb{R}^n$ and there exists a point $x^* \in \mathbb{R}^n$ satisfying $h_i(x^*) = 0$, $i = 1, \dots, m$ and $\lambda^* \in \mathbb{R}^m$ such that:

- $\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*) = 0$ *for* $i = 1, \dots n$,
- *for all* $y \in TS(x^*)$, $y \neq 0$ *we have that* $y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y > 0$ ($y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y < 0$), *i.e.* $\nabla^2 \mathcal{L}(x^*, \lambda^*)$ *is a positive (negative) definite matrix on the tangent space* $TS(x^*)$,

then x^* is a strict local minimizer (maximizer) of f subject to $h_i(x) = 0$, $i = 1, \dots, m$.

Note: If $\nabla^2 \mathcal{L}(x^*, \lambda^*)$ is a positive (negative) definite matrix, then you don't need to consider the tangent space $TS(x^*)$.

Example

Find the minimizers and maximizers of the function

$$f(x_1, x_2, x_3) = \left(a^T x\right)\left(b^T x\right),$$

subject to

$$x_1 + x_2 = 0,$$

$$x_2 + x_3 = 0,$$

where

$$a = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Problems with inequality constraints

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h_i(x) = 0, \quad i = 1, \ldots m, \\ & g_j(x) \leq 0, \quad j = 1, \ldots p \end{aligned}$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots m$, $m \leq n$ and $g_j : \mathbb{R}^n \to \mathbb{R}$ for $j = 1, \ldots p$ are given functions.

Definition

Assume x^* is a feasible point. An inequality constraint $g_j(x) \leq 0$ is said to be active at x^* if $g_j(x^*) = 0$. It is inactive at x^* if $g_j(x^*) < 0$.

Definition

Let x^* be a feasible point, i.e. $h_i(x^*) = 0$, $i = 1, \dots m$ and $g_j(x^*) \leq 0$, $j = 1, \dots p$ and let $J(x^*)$ be the index set of active inequality constraints

$$J(x^*) = \left\{ j : g_j(x^*) = 0 \right\}.$$

We will say that x^* is a regular point if the vectors

$$\nabla h_i(x^*),\ i = 1, \dots m, \nabla g_j(x^*),\ j \in J(x^*)$$

are linearly independent.

Example

Consider the following constraints on $\mathbb{R}^2$

$$h(x_1, x_2) = x_1 - 2 = 0 \quad \text{and} \quad g(x_1, x_2) = (x_2 + 1)^3 \leq 0.$$

Find the set of feasible points. Are the feasible points regular?

Let's introduce the Lagrangian function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, given by

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x).$$

Theorem (Karush-Kuhn-Tucker (KKT) Theorem)

Let x^* be a local minimizer of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $h(x) = 0$, $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, $g(x) \leq 0$, $g : \mathbb{R}^n \to \mathbb{R}^p$. Assume f, h_i, $i = 1, \ldots m$ and g_j, $j = 1, \ldots p$ are continuously differentiable functions and x^* is a regular point. Then, there exists $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that

- $\mu_j^* \geq 0$, *for* $j = 1, \ldots p$,
- $\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*, \mu^*) = 0$, *for* $i = 1, \ldots n$,
- $\mu_j^* g_j(x^*) = 0$, *for* $j = 1, \ldots p$.

Theorem (Karush-Kuhn-Tucker (KKT) Theorem)

Let x^* be a local maximizer of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $h(x) = 0$, $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, $g(x) \leq 0$, $g : \mathbb{R}^n \to \mathbb{R}^p$. Assume f, h_i, $i = 1, \dots m$ and g_j, $j = 1, \dots p$ are continuously differentiable functions and x^* is a regular point. Then, there exists $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that

- $\mu_j^* \leq 0$, *for* $j = 1, \dots p$,
- $\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*, \mu^*) = 0$, *for* $i = 1, \dots n$,
- $\mu_j^* g_j(x^*) = 0$, *for* $j = 1, \dots p$.