Nonlinear Constrained Optimization

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & x \in \Omega, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\Omega \subset \mathbb{R}^n$. Here, we are going to consider minimization problems, for which the constraint set Ω is given by

$$\Omega = \left\{ x \in \mathbb{R}^n : \; h_i(x) = 0, \text{ for } i = 1, \ldots m, \; g_j(x) \leq 0, \text{ for } j = 1, \ldots p \right\},$$

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.

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

or

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \\ & g(x) \leq 0, \end{aligned}$$

where $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$ and $g : \mathbb{R}^n \to \mathbb{R}^p$.

Example

Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & (x_1 - 1)^2 + x_2 - 2 \\ \text{subject to} \quad & x_2 - x_1 = 1, \\ & x_1 + x_2 \leq 2. \end{aligned}$$

Problems with equality constraints

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

We will assume that f, h_i for $i = 1, \ldots m$ are continuously differentiable functions on $\mathbb{R}^n$.

Definition

A point x^* satisfying the constraints $h_i(x^*) = 0$, $i = 1, \ldots m$ is said to be a regular point of the constraints, if the gradient vectors $\nabla h_1(x^*), \ldots, \nabla h_1(x^*)$ are linearly independent. When $m = 1$, this means $\nabla h_1(x^*) \neq 0$

Example

Consider following constraints $h_1(x) = x_1$ and $h_2(x) = x_2 - x_3^2$ on $\mathbb{R}^3$. Show that all feasible points are regular points.

Theorem (Lagrange's Theorem, First Order Necessary Condition)

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

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0.$$

$\lambda_1^*, \ldots, \lambda_m^*$ are called Lagrange multipliers.

Example

Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \end{aligned}$$

where $f(x) = x$ and $h(x) = \begin{cases} x^2 & \text{if } x < 0, \\ 0 & \text{if } 0 \leq x \leq 1, \\ (x-1)^2 & \text{if } x > 1. \end{cases}$

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

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

The necessary condition for x^* to be a local minimizer will be

$$\nabla \mathcal{L}(x^*, \lambda^*) = 0$$

for some $\lambda^* \in \mathbb{R}^m$.

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 $h_i(x) = 0$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots m$, $m \leq n$. Assume f, h_i for $i = 1, \ldots m$ 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* $TS(x^*) = \left\{ y \in \mathbb{R}^n : y^T \nabla h_i(x^*) = 0, \text{for } i = 1, \ldots, m \right\}$, *then for all* $y \in TS(x^*)$ *we have that* $y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y \geq 0$ $(y^T \nabla^2 \mathcal{L}(x^*, \lambda^*) y \leq 0)$.

Theorem (Second Order Sufficient Condition SOSC)

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

- $\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*) = 0$ *for* $i = 1, \ldots n$,
- *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 $h_i(x) = 0$, $i = 1, \ldots, m$.

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

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