Numerical Analysis

Lusine Poghosyan

AUA

November 24, 2018

Constrained Optimization

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

where $\Omega \subset \mathbb{R}^n$.

$x^*$is called a global minimizer for problem (1) if

$$f(x^*) \leq f(x), \quad \forall x \in \Omega.$$

The points of Ω are called feasible points.

Here, we are going to consider minimization problems, for which the constraint set Ω is given by

$$\Omega = \{x \in \mathbb{R}^n : \; h_i(x) = 0, \text{ for } i = 1, \dots m\},$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots m$ are given functions and $m \le n$.

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

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 $\nabla h_1(x^*), \ldots, \nabla h_m(x^*)$ 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)

Let x^* be a local minimizer of

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

where f, h_i for $i = 1, \dots m$ are continuously differentiable functions on $\mathbb{R}^n$. Assume x^* is a regular point. Then, there exists $\lambda^* \in \mathbb{R}^m$ such that

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla h_j(x^*) = 0.$$

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

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j h_j(x).$$

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

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

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

Example

Consider the optimization problem

$$\begin{aligned} \textit{minimize} \quad & x_1^2 + x_2^2 \\ \textit{subject to} \quad & x_1^2 + 2x_2^2 = 2. \end{aligned}$$

Use Lagrange's theorem to find all possible local minimizers.

The Penalty Method

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

where $\Omega \subset \mathbb{R}^n$.

We now discuss a method for solving this problem using techniques from unconstrained optimization problem. Specifically we consider the following unconstrained minimization problem

$$\text{minimize } f(x) + \gamma P(x),$$

where $\gamma \in \mathbb{R}$ is a positive constant and $P : \mathbb{R}^n \to \mathbb{R}$ is a given function. The constant γ is called the penalty parameter and the function $P(x)$ is called the penalty function.

Definition

A function $P : \mathbb{R}^n \to \mathbb{R}$ is called a penalty function for the constrained minimization problem above, if it satisfies the following conditions

1. P is continuous.
2. $P(x) \geq 0$ for all $x \in \mathbb{R}^n$.
3. $P(x) = 0$ if and only if $x \in \Omega$.

We take the solution of

$$\text{minimize } f(x) + \gamma P(x),$$

as approximation to the solution of original problem. Of course, it may not be exactly equal to the true solution but we expect that the larger the value of the penalty parameter γ, the closer the approximated solution will be to the true solution.

Assume $\{\gamma_k\}$ is a sequence of positive numbers. Let's denote by $x^{(k)}$ a solution of

$$\text{minimize}\, f(x) + \gamma_k P(x).$$

Theorem

Suppose the f is a continuous function, $\{\gamma_k\}$ is a sequence of positive numbers such that $\lim_{k\to\infty} \gamma_k = +\infty$. If $\{x^{(k)}\}$ is a convergent sequence, then its limit is a solution to the original problem.

How can we choose the penalty function?

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

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

$$P(x) = \sum_{i=1}^{m} |h_i(x)|$$

or

$$P(x) = \sum_{i=1}^{m} (h_i(x))^2 .$$

Example

Consider the following constrained optimization problem: find the minimum of $f(x_1, x_2) = 2x_1 + x_2$ subject to $x_2 - x_1^2 = 1$.

a. Solve this constrained minimization problem by using the Lagrange Multipliers Method;

b. Consider the following Penalty function: for large $\gamma > 0$,

$$g(x_1, x_2) = f(x_1, x_2) + \gamma \cdot \left(x_2 - x_1^2 - 1\right)^2 .$$

Assuming γ is fixed, find the minimum point $x^{(\gamma)}$ of g;

c. Prove that $x^{(\gamma)}$ tends to the solution obtained by the LMM, as $\gamma \to +\infty$.