Constrained Optimization

Maestría en Cómputo Estadístico

Centro de Investigación en Matemáticas A.C.

What is constrained optimization? I

- Roughly speaking, it is an optimization problem with either equallity or inequallity constrains
- Constrains can be linear or nonlinear

What is constrained optimization? II

- Mathematically, a nonlinear constrained optimization problem is defined as:

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & f(\mathbf{x}) \\
\text{subject to:} \quad & g_i(\mathbf{x}) = 0, i = 1, \ldots, m \\
& h_j(\mathbf{x}) \leq 0, j = 1, \ldots, l
\end{aligned}
\tag{1}
$$

CIMAT

What is constrained optimization? III

- Special cases are linear programming, quadratic programming, and equality constrained problems
- In linear programming problems, both objective function and constraints are linear
- In quadratic programming problems, the objective function is quadratic and constraints are linear
- In equality constrained problems, all constraints are equality
- In linearly constrained optimization problems, all constraints are linear

CIMAT

Feasible Solutions I

- In constrained programming, a point $\mathbf{x}$ is said to be feasible if and only if all constrains are satisfied

Feasible Solution

The point $\mathbf{x} \in \mathbb{R}^n$ is said to be a feasible point if and only if constraints in Equation (1) hold. The set of all feasible points is said to be a feasible set.

CIMAT

Feasible Solutions II

- The feasible set, $\mathcal{X}$, can be rewritten as:

$$\mathcal{X} = \{\mathbf{x} : g_i(\mathbf{x}) = 0 \forall i \in \{1, \ldots, m\} \cap h_j(\mathbf{x}) \leq 0 \forall j \in \{1, \ldots, l\}\} \quad (2)$$

- Thus, the optimization problem can be rewritten as:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & \mathbf{x} \in \mathcal{X} \end{aligned} \quad (3)$$

Recalling I

- A solution is a **global minimum** if and only if $f(\mathbf{x}^*) \leq f(\mathbf{x}), \forall \mathbf{x} \in \mathcal{X}$
- A solution is a **local minimum** if and only if $f(\mathbf{x}^*) \leq f(\mathbf{x}), \forall \mathbf{x} \in B(\mathbf{x}^*, \delta)$, where $B(\mathbf{x}^*, \delta) = \{\mathbf{x} : \| \mathbf{x} - \mathbf{x}^* \|_2 \leq \delta$

Active Constraints I

- Assume that $\mathbf{x}^*$ is a local minimizer of a constrained optimization problem, if there is an index $j_0 \in \{1, \ldots, j\}$ such that

$$h_{j_0}(\mathbf{x}^*) \leq 0 \tag{4}$$

- Then, if we delete the j_0^{th}-constraint, $\mathbf{x}^*$ is still the local minimizer of the constrained optimization problem
- In this case, the j_0^{th}-constraint is said to be inactive

CIMAT

Active Constraints II

Active constraints

Let $I(\mathbf{x})$ the set of inequality constraints such that $I(\mathbf{x}) = \{h_k : h_k(\mathbf{x}) = 0\}$ and $E(\mathbf{x})$ the set of equality constraints, the set of active constraints is defined as $A = E(\mathbf{x}) \cup IE(\mathbf{x})$

Active Constraints III

- Thus, we can rewrite the optimization problem as:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & c_i(\mathbf{x}) = 0, c_i \in A \end{aligned} \tag{5}$$

First-Order Optimality Conditions I

- The feasible directions play a very important role in deriving the optimality conditions

Feasible directions

Let $\mathbf{x}^* \in \mathcal{X}$, $d \in \mathbb{R}^n$ and $d \neq 0$. If there exists $\delta \geq 0$ such that $\mathbf{x}^* + td \in \mathcal{X}, \forall t \in [0, \delta]$, then d is said to be a feasible direction of $\mathcal{X}$ at $\mathbf{x}^*$. The set of all feasible directions of $\mathcal{X}$ at $\mathbf{x}^*$ is
$FD(\mathbf{x}^*, \mathcal{X}) = \{d : \mathbf{x}^* + td \in \mathcal{X}, \forall t \in [0, \delta]\}$

CIMAT

First-Order Optimality Conditions II

- For constrained optimization problems, only feasible directions are relevant
- For equality constraints problems, a necessary condition for a feasible point $\mathbf{x}^*$ to be a solution is that the negative of the gradient lie in the space spanned by constraints, i.e., $-\nabla f(\mathbf{x}^*) = \mathbf{J}_g^T \lambda$

First-Order Optimality Conditions III

- The Lagrangian function is defined as $\mathcal{L} = f(\mathbf{x}) - \sum_{i=1}^{l+m} \lambda_i c_i(\mathbf{x})$, such that:

$$\begin{aligned} c_i(\mathbf{x}) &= 0, c_i \in A \\ c_i(\mathbf{x}) &= 0, c_i \notin A \\ \lambda_i &\geq 0 \\ \lambda_i c_i &= 0 \end{aligned} \tag{6}$$

First-Order Optimality Conditions IV

- A necessary condition to consider $\mathbf{x}$ as an optimal solution can be stated as follows:

$$\nabla f(\mathbf{x}) - \sum_{i=1}^{l+m} \lambda_i \nabla c_i(\mathbf{x}) = 0 \tag{7}$$

- This is known as the Karush-Kuhn-Tucker (KKT) condition

Second-Order Optimality Conditions I

- Let $\mathcal{L}(\mathbf{x}, \gamma, \lambda) = f(\mathbf{x}) - \sum_{i=1}^{l+m} \lambda_i c_i(\mathbf{x})$
- A second-order condition should satisfy:

$$d^T \nabla^2 \mathcal{L}(\mathbf{x}, \gamma, \lambda) d \geq 0 \tag{8}$$

Questions?