

Applied Linear Algebra in Data Analysis

Introduction to Constrained Optimization

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Constrained Optimization

- A general optimization problem can be fomulated as the following,

$$\begin{aligned} & \text{minimize } f(\mathbf{x}) \\ & \text{subject to } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}) \quad g_2(\mathbf{x}) \quad \cdots \quad g_p(\mathbf{x})]^\top \\ & \quad \quad \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}) \quad h_2(\mathbf{x}) \quad \cdots \quad h_q(\mathbf{x})]^\top \end{aligned}$$

where, $f(\mathbf{x})$ is the **objective function** and $\mathbf{g}(\mathbf{x})$ represents the set of **inequality constaints** and $\mathbf{h}(\mathbf{x})$ represents the set of **equality constraints**.

Constrained Optimization

The set of all values of \mathbf{x} that satisfy the constraints is called the **feasible set** and is denoted by \mathcal{F} .

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \text{ and } \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$$

Whatever the form, with constrained optimization the search for the minimizer of the function $f(\mathbf{x})$ is restricted to the feasible set \mathcal{F} .

Feasible direction - A direction \mathbf{d} is said to be a feasible direction at the point $\mathbf{x} \in \mathcal{F}$ if there exists $\alpha_0 > 0$ such that $\mathbf{x} + \alpha\mathbf{d} \in \mathcal{F}$ for all $\alpha \in [0, \alpha_0]$.

The conditions for a point to be a minimizer in the constrained optimization case are different from the unconstrained case.

Constrained Optimization

First Order Necessary Condition. Let $\mathcal{F} \in \mathbb{R}^n$, and let $f : \mathcal{F} \rightarrow \mathbb{R}$ be a differentiable function (at least once). If \mathbf{x}^* is a local minimizer of the function f over the set \mathcal{F} , then for any feasible direction \mathbf{d} at the point \mathbf{x}^* , we have,

$$\mathbf{d}^\top \nabla f(\mathbf{x}^*) \geq 0$$

This above conditions applies for both points at the boundary and interior to the set \mathcal{F} . For interior points, all directions are feasible, and so the above condition becomes,

$$\nabla f(\mathbf{x}^*) = 0$$

which is the same as the unconstrained case.

Constrained Optimization

Second Order Necessary Condition. Let $\mathcal{F} \in \mathbb{R}^n$, and let $f : \mathcal{F} \rightarrow \mathbb{R}$ be a differentiable function (at least twice), \mathbf{x}^* is a local minimizer of f over \mathcal{F} , and \mathbf{d} is a feasible direction at \mathbf{x}^* . If $\mathbf{d}^\top \nabla f(\mathbf{x}) = 0$, then

$$\mathbf{d}^\top \mathbf{H}(\mathbf{x}^*) \mathbf{d} \geq 0$$

where $\mathbf{H}(\mathbf{x})$ is the Hessian of f at \mathbf{x} .

Constrained Optimization: Equality Constraints

Consider the following optimization problem with only equality constraints,

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

where, $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}) \ h_2(\mathbf{x}) \ \cdots \ h_q(\mathbf{x})]^\top$, and we assume that each one of these constraint functions is differentiable.

The feasible set for this problem is,

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$$

Note that the feasibility set forms a $n - q$ dimensional surface or manifold in \mathbb{R}^n .

Constrained Optimization: Equality Constraints

Regular point – A point $\mathbf{x}^* \in \mathcal{F}$ is said to be a regular point if the gradients of the equality constraints are linearly independent at the point \mathbf{x}^* .

Note that the gradient of the equality constraints is a matrix of size $q \times n$.

$$\nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}) = \begin{bmatrix} \nabla_{\mathbf{x}} h_1(\mathbf{x}) \\ \nabla_{\mathbf{x}} h_2(\mathbf{x}) \\ \vdots \\ \nabla_{\mathbf{x}} h_q(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{q \times n}$$

The gradient $\nabla \mathbf{h}(\mathbf{x})$ has full rank at a regular point.

Constrained Optimization: Equality Constraints

The **tangent space** of a surface is the higher dimensional equivalent of a tangent line to a one dimensional curve.

The **tangent space** at a point $\mathbf{x}^* \in \mathcal{F}$ is defined as the following set,

$$T(\mathbf{x}^*) = \{\mathbf{y} \in \mathbb{R}^n \mid \nabla \mathbf{h}(\mathbf{x}) \mathbf{y} = \mathbf{0}\} = \mathcal{N}(\nabla \mathbf{h}(\mathbf{x}))$$

The **tangent plane** at the point \mathbf{x}^* on the surface \mathcal{F} is defined as,

$$TP(\mathbf{x}^*) = T(\mathbf{x}^*) + \mathbf{x}^* = \{\mathbf{x} + \mathbf{x}^* \mid \mathbf{x} \in T(\mathbf{x}^*)\}$$

Constrained Optimization: Equality Constraints

The **normal space** of a surface is the orthogonal complement of the tangent space.

The **normal space** at a point $\mathbf{x}^* \in \mathcal{F}$ is defined as the following set,

$$N(\mathbf{x}^*) = \left\{ \nabla \mathbf{h}(\mathbf{x})^\top \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^q \right\} = \mathcal{R} \left(\nabla \mathbf{h}(\mathbf{x})^\top \right)$$

The **normal plane** at the point \mathbf{x}^* on the surface \mathcal{F} is defined as,

$$NP(\mathbf{x}^*) = N(\mathbf{x}^*) + \mathbf{x}^* = \{ \mathbf{x} + \mathbf{x}^* \mid \mathbf{x} \in N(\mathbf{x}^*) \}$$

Constrained Optimization: Equality Constraints

Lagrange theorem: Let \mathbf{x}^* be a local minimizer (or maximizer) of the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^q$, $q \leq n$. Assume that \mathbf{x}^* is a regular point of the equality constraints. Then there exists a vector $\boldsymbol{\lambda}^* \in \mathbb{R}^q$ such that,

$$\nabla f(\mathbf{x}^*) + \nabla \mathbf{h}(\mathbf{x}^*)^\top \boldsymbol{\lambda}^* = \mathbf{0}$$

where, $\boldsymbol{\lambda}^*$ is called the **Lagrange multiplier vector**.

This theorem states that extremum of this constrained optimization problem occurs when the gradient of f is in the normal space of the surface \mathcal{F} , i.e. $\nabla f(\mathbf{x})$ is a linear combination of the gradients of the individual equality constraints $\nabla h_i(\mathbf{x})$, $1 \leq i \leq q$.

This is also equivalent to saying $\nabla f(\mathbf{x}) \in T(\mathbf{x})^\perp$ or $\nabla f(\mathbf{x}) \in N(\mathbf{x})$

Constrained Optimization: Equality Constraints

It is convenient to introduce the **Lagrange function**: $l : \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}$,

$$l(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{f}(\mathbf{x})$$

The Lagrange condition for \mathbf{x}^* to be a local minimizer is the following,

$$\nabla l(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$$

for some $\boldsymbol{\lambda}^*$, where the gradient is with respect to the combined argument $[\mathbf{x}^\top \quad \boldsymbol{\lambda}^\top]$. This is equivalent to the first order necessary condition for the unconstrained optimization problem with respect to $[\mathbf{x}^\top \quad \boldsymbol{\lambda}^\top]$. This condition is necessary but not sufficient.

Constrained Optimization: Equality Constraints

Let f and \mathbf{h} be both twice continuously differentiable,

$$\begin{aligned}l(\mathbf{x}, \boldsymbol{\lambda}) &= f(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{f}(\mathbf{x}) \\ &= f(\mathbf{x}) + \lambda_1 h_1(\mathbf{x}) + \lambda_2 h_2(\mathbf{x}) + \cdots + \lambda_q h_q(\mathbf{x})\end{aligned}$$

Taking the Hessian on both sides,

$$\mathbf{H}_l(\mathbf{x}) = \mathbf{H}_f(\mathbf{x}) + \sum_{i=1}^q \lambda_i \mathbf{H}_{h_i}(\mathbf{x})$$

where \mathbf{H}_l is the Hessian of $l(\mathbf{x}, \boldsymbol{\lambda})$ with respect to \mathbf{x} , and \mathbf{H}_{h_i} is the Hessian of the function $h_i(\mathbf{x})$ with respect to \mathbf{x} .

Constrained Optimization: Equality Constraints

Second Order Necessary Condition. Let \mathbf{x}^* be a local minimizer of f subject to the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^q$, $q \leq n$, and f, \mathbf{h} are twice differentiable, and \mathbf{x}^* is a regular point. Then, there exists a vector $\boldsymbol{\lambda}^* \in \mathbb{R}^q$ such that,

1. $\nabla f(\mathbf{x}^*) + \boldsymbol{\lambda}^{*\top} \nabla \mathbf{h}(\mathbf{x}) = \mathbf{0}$
2. For all $\mathbf{y} \in T(\mathbf{x}^*)$, we have $\mathbf{y}^\top \mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \mathbf{y} \geq 0$.

Second Order Sufficient Condition. \mathbf{x}^* is strict local minimizer of f subject to the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^q$,

1. $\nabla f(\mathbf{x}^*) + \boldsymbol{\lambda}^{*\top} \nabla \mathbf{h}(\mathbf{x}) = \mathbf{0}$
2. For all $\mathbf{y} \in T(\mathbf{x}^*)$, we have $\mathbf{y}^\top \mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \mathbf{y} > 0$.