ECEn 671: Mathematics of Signals and Systems

Randal W. Beard

Brigham Young University

September 1, 2023

Section 1

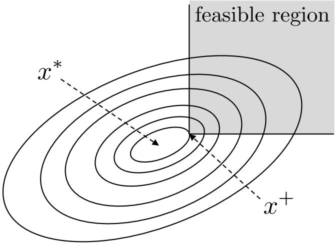
Constrained Optimization

Constrained Optimization

In Chapter 14 we studied unconstrained minimization of continuously differentiable functions.

In Chapter 18 we focus on constrained optimization problems.

For example, given the level curves,



Note that the constrained optimum x^+ does not equal the unconstrained optimum x^*.

The unconstrained optimum is x^*; the constained optimum is x^+.

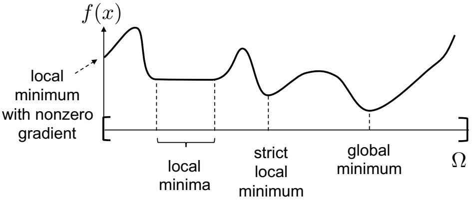
Constrained Optimization

Definition

Let $\Omega \subseteq \mathbb{R}^n$ be the feasible region. Then $x^* \in \Omega$ is a <u>local minimum</u> of $f : \mathbb{R}^n \to \mathbb{R}$ over Ω if $\exists \epsilon > 0$ such that

$$x \in \Omega \cap \{y \in \mathbb{R}^n : |u - x^*| < \epsilon\} \quad \Longrightarrow \quad f(x) \geq f(x^*).$$

If $f(x) > f(x^*)$ then x^* is a <u>strict local minimum</u>. If true for all $\epsilon > 0$ then x^* is a global minimum.



Constrained Optimization

Definition

Let $x \in \Omega$ and $d \in \mathbb{R}^n$, then

$$y = x + \alpha d$$

is a feasible point if $y \in \Omega$.

Definition

The vector d is a feasible direction at x, if $\exists \epsilon_0 > 0$ such that

$$x + \epsilon d \in \Omega$$

for every $0 \leq \epsilon \leq \epsilon_0$.

Constrained Optimization

Recall, if $f : \mathbb{R}^n \to \mathbb{R}$ then the gradient vector is

$$\frac{\partial f}{\partial x} = \nabla_x f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

and the Hessian matrix is

$$\frac{\partial^2 f}{\partial x^2} = \nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

If $\Omega = \mathbb{R}^n$ then a necessary condition for x^* to be a local minima is that $\nabla_x f(x^*) = 0$. What about constrained optimization problems?

Constrained Optimization

Theorem (Moon Theorem 18.1)

Let $\Omega \subseteq \mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ (continuously differentiable) on Ω.

1. *If x^* is a local minimum of f over Ω, then for any feasible direction $d \in \mathbb{R}^n$ at x^**

$$[\nabla_x f(x^*)]^\top d \geq 0$$

2. *If x^* is an interior point of Ω, then*

$$\nabla f(x^*) = 0.$$

3. *If in addition, $f \in \mathcal{C}^2$ and $\nabla_x f(x^*)^\top d = 0$, then*

$$d^\top \nabla^2 f(x^*) d \geq 0$$

Note that this is a weaker condition than psd Hessian.

Proof of Theorem 18.1

1. By Taylor series expansion,

$$\begin{aligned} f(x^* + \epsilon d) &= f(x^*) + \epsilon \nabla_x f(x^*)^\top d + O(\epsilon) \\ \implies f(x^* + \epsilon d) - f(x^*) &= \epsilon \nabla_x f(x^*)^\top d + O(\epsilon) \end{aligned}$$

Since x^* is a local minimum, for ϵ sufficiently small we must have that

$$\begin{aligned} &\implies f(x^* + \epsilon d) - f(x^*) \geq 0 \\ &\implies \nabla_x f(x^*)^\top d \geq 0. \end{aligned}$$

Proof of Theorem 18.1, cont.

2. If x^* is an interior point then every $d \in \mathbb{R}^n$ is feasible at x^*, i.e.

$$\langle \nabla_x f(x^*), d \rangle_{\mathbb{R}^n} = 0, \qquad \forall d \in \mathbb{R}^n.$$

Therefore,

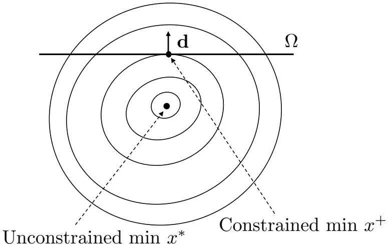
$$\begin{aligned} & \nabla_x f(x^*)^\top d \geq 0 \quad \text{and} \quad \nabla_x f(x^*)^\top (-d) \geq 0 \\ \Longrightarrow\ & \nabla_x f(x^*)^\top d = 0, \quad \forall d \in \mathbb{R}^n \\ \Longrightarrow\ & \nabla_x f(x^*) = 0 \end{aligned}$$

since R^n is a finite dimensional vector space .

Proof of Theorem 18.1, cont.

3. If $\nabla_x f(x^*)^\top d = 0$ then the Taylor series for f is

$$
\begin{aligned}
& f(x^* + \epsilon d) = f(x^*) + \epsilon^2 d^\top \nabla^2 f(x^*) d + O(\epsilon^2) \\
\implies & 0 \le f(x^* + \epsilon d) - f(x^*) = \epsilon^2 d^\top \nabla^2 f(x^*) d + O(\epsilon^2) \\
\implies & d^\top \nabla^2 f(x^*) d \ge 0.
\end{aligned}
$$



Note: Any feasible d points uphill.

Note: The function is concave in feasible region.

Constrained Optimization: Sufficient Conditions

Are there sufficient conditions?

First, suppose that the constraints are not active, i.e. x^* is an interior point of Ω. (We will consider the active constraint case later.)

Theorem (Moon Theorem 18.2)

Let $f \in \mathcal{C}^2$ on Ω and let x^* be an interior point of Ω. If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite then x^* is a strict local minimum of f.

Constrained Optimization: Sufficient Conditions: Proof

Proof.

Let d be any unit vector in $\mathbb{R}^n$ then

$$\begin{aligned} & f(x^* + \epsilon d) = f(x^*) + \epsilon \nabla f(x^*)^\top d + \epsilon^2 d^\top \nabla^2 f(x^*) d + O(\epsilon^2) \\ \implies & f(x^* + \epsilon d) - f(x^*) = \epsilon^2 d^\top \nabla^2 f(x^*) d + O(\epsilon^2) \end{aligned}$$

Since $\nabla^2 f(x^*)$ is positive definite, it follows that for ϵ sufficiently small

$$f(x^* + \epsilon d) - f(x^*) > 0,$$

which implies that x^* is a strict local minimum. $\square$

Note: we cannot generalize this theorem to the case when $\nabla^2 f(x^*)$ is p.s.d.. Why?

Section 2

General Constrained Optimization

Constrained Optimization

In general we have two types of constraints:

1. Equality constraints of the form

$$h_i(x) = 0$$

   For example:

$$h_1(x) \triangleq x_1^2 + x_1 x_2 x_3 + \tan(x_3)\cos(x_2) = 0$$

2. Inequality constraints of the form

$$g_i(x) \leq 0$$

   For example

$$\begin{aligned} x_1 &\geq 0, & x_2 &\geq 0 \\ \implies g_1(x) &\triangleq -x_1 \leq 0, & g_2(x) &\triangleq -x_2 \leq 0 \end{aligned}$$

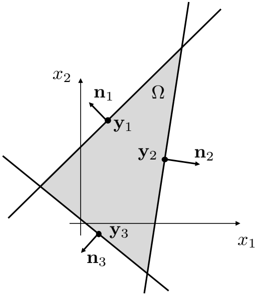
Constrained Optimization

In fact a region $\Omega \subset \mathbb{R}^n$ can always be described by inequality constraints.

Example

Feasible Region Ω:

$$(x - \mathbf{y}_1)^\top \mathbf{n}_1 \leq 0$$
$$(x - \mathbf{y}_2)^\top \mathbf{n}_2 \leq 0$$
$$(x - \mathbf{y}_3)^\top \mathbf{n}_3 \leq 0$$



Where $\mathbf{n}_i$ is a vector normal to the linear constraint.

Constrained Optimization

A general constrained optimization problem can be written as

$$
\begin{aligned}
\min_{x \in \Omega} \quad & f(x) \\
\text{s.t.} \quad & h_1(x) = 0, \\
& \vdots, \\
& h_m(x) = 0, \\
& g_1(x) \leq 0, \\
& \vdots, \\
& g_p(x) \leq 0
\end{aligned}
$$

Constrained Optimization

Letting

$$\mathbf{h} = (h_1 \ldots h_m)^\top$$
$$\mathbf{g} = (g_1 \ldots g_p)^\top,$$

we have

$$\begin{aligned} \min_{x \in \Omega} \quad & f(x) \\ \text{s.t.} \quad & \mathbf{h}(x) = 0, \\ & \mathbf{g}(x) \le 0 \end{aligned}$$

Equality constraints are easier to deal with than inequality constraints.

We will first treat equality constraints, then inequality constraints.