Fundamentals of Linear Algebra and Optimization

Extrema of Real-Valued Functions

Jean Gallier and Jocelyn Quaintance

CIS Department
University of Pennsylvania

jean@cis.upenn.edu

May 8, 2020

Extrema of Real-Valued Functions

This lesson deals with extrema of real-valued functions. In most optimization problems we need to find necessary conditions for a function $J\colon \Omega \to \mathbb{R}$ to have a local extremum with respect to a subset U of Ω (where Ω is open). This can be done in two cases:

Extrema of Real-Valued Functions

This lesson deals with extrema of real-valued functions. In most optimization problems we need to find necessary conditions for a function $J\colon \Omega \to \mathbb{R}$ to have a local extremum with respect to a subset U of Ω (where Ω is open). This can be done in two cases:

(1) The set U is defined by a set of equations,

$$U = \{x \in \Omega \mid \varphi_i(x) = 0,\ \ 1 \leq i \leq m\},$$

where the functions $\varphi_i\colon \Omega \to \mathbb{R}$ are continuous and differentiable.

Extrema of Real-Valued Functions

This lesson deals with extrema of real-valued functions. In most optimization problems we need to find necessary conditions for a function $J\colon \Omega \to \mathbb{R}$ to have a local extremum with respect to a subset U of Ω (where Ω is open). This can be done in two cases:

(1) The set U is defined by a set of equations,

$$U = \{x \in \Omega \mid \varphi_i(x) = 0,\ \ 1 \leq i \leq m\},$$

where the functions $\varphi_i\colon \Omega \to \mathbb{R}$ are continuous and differentiable.

(2) The set U is defined by a set of inequalities,

$$U = \{x \in \Omega \mid \varphi_i(x) \leq 0,\ \ 1 \leq i \leq m\},$$

where the functions $\varphi_i\colon \Omega \to \mathbb{R}$ are continuous and differentiable.

Equality Constraints

In (1), the equations $\varphi_i(x) = 0$ are called *equality constraints*, and in (2), the inequalities $\varphi_i(x) \leq 0$ are called *inequality constraints*. The case of equality constraints is much easier to deal with and is treated in this lesson.

Equality Constraints

In (1), the equations $\varphi_i(x) = 0$ are called *equality constraints*, and in (2), the inequalities $\varphi_i(x) \leq 0$ are called *inequality constraints*. The case of equality constraints is much easier to deal with and is treated in this lesson.

In the case of equality constraints, a necessary condition for a local extremum with respect to U can be given in terms of *Lagrange multipliers*. In the case of inequality constraints, there is also a necessary condition for a local extremum with respect to U in terms of generalized Lagrange multipliers and the *Karush–Kuhn–Tucker* conditions.

Definition of a Local Minimum

Let $J \colon E \to \mathbb{R}$ be a real-valued function defined on a normed vector space E. Ideally we would like to find where the function J reaches a minimum or a maximum value, at least locally.

Definition of a Local Minimum

Let $J\colon E \to \mathbb{R}$ be a real-valued function defined on a normed vector space E. Ideally we would like to find where the function J reaches a minimum or a maximum value, at least locally.

Definition. If $J\colon E \to \mathbb{R}$ is a real-valued function defined on a normed vector space E, we say that J has a *local minimum* (or *relative minimum*) at the point $u \in E$ if there is some open subset $W \subseteq E$ containing u such that

$$J(u) \leq J(w) \quad \text{for all } w \in W.$$

Definition of a Local Maximum

Similarly, we say that J has a *local maximum* (or *relative maximum*) at the point $u \in E$ if there is some open subset $W \subseteq E$ containing u such that

$$J(u) \geq J(w) \quad \text{for all } w \in W.$$

Definition of a Local Maximum

Similarly, we say that J has a *local maximum* (or *relative maximum*) at the point $u \in E$ if there is some open subset $W \subseteq E$ containing u such that

$$J(u) \geq J(w) \quad \text{for all } w \in W.$$

In either case, we say that J has a *local extremum* (or *relative extremum*) at u. We say that J has a *strict local minimum* (resp. *strict local maximum*) at the point $u \in E$ if there is some open subset $W \subseteq E$ containing u such that

$$J(u) < J(w) \quad \text{for all } w \in W - \{u\}$$

(resp.

$$J(u) > J(w) \quad \text{for all } w \in W - \{u\}).$$

Necessary Condition for Local Extrema

We begin with a *necessary condition* for a local extremum.

Necessary Condition for Local Extrema

We begin with a *necessary condition* for a local extremum.

Proposition. Let E be a normed vector space and let $J\colon \Omega \to \mathbb{R}$ be a function, with Ω some open subset of E. If the function J has a local extremum at some point $u \in \Omega$ and if J is differentiable at u, then

$$dJ_u = J'(u) = 0.$$

Necessary Condition for Local Extrema

Proof. Pick any $v \in E$. Since Ω is open, for t small enough we have $u + tv \in \Omega$, so there is an open interval $I \subseteq \mathbb{R}$ such that the function φ given by

$$\varphi(t) = J(u + tv)$$

for all $t \in I$ is well-defined. By applying the chain rule, we see that φ is differentiable at $t = 0$, and we get

$$\varphi'(0) = dJ_u(v).$$

Necessary Condition for Local Extrema

Without loss of generality, assume that u is a local minimum. Then we have

$$\varphi'(0) = \lim_{t \mapsto 0_-} \frac{\varphi(t) - \varphi(0)}{t} \leq 0$$

and

$$\varphi'(0) = \lim_{t \mapsto 0_+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0,$$

which shows that $\varphi'(0) = dJ_u(v) = 0$. As $v \in E$ is arbitrary, we conclude that $dJ_u = 0$. $\square$

Critical Point

Definition. A point $u \in \Omega$ such that $J'(u) = 0$ is called a *critical point* of J.

Critical Point

Definition. A point $u \in \Omega$ such that $J'(u) = 0$ is called a *critical point* of J.

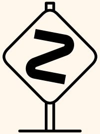
If $E = \mathbb{R}^n$, then the condition $dJ_u = 0$ is equivalent to the system

$$\frac{\partial J}{\partial x_1}(u_1, \ldots, u_n) = 0$$

$$\vdots$$

$$\frac{\partial J}{\partial x_n}(u_1, \ldots, u_n) = 0.$$

Necessary Condition for Local Extrema



The condition of the preceding proposition is only a *necessary* condition for the existence of an extremum, but *not* a sufficient condition.

Necessary Condition for Local Extrema

The condition of the preceding proposition is only a *necessary* condition for the existence of an extremum, but *not* a sufficient condition.

Here are some counter-examples.

Necessary Condition for Local Extrema

The condition of the preceding proposition is only a *necessary* condition for the existence of an extremum, but *not* a sufficient condition.

Here are some counter-examples.
If $f\colon \mathbb{R} \to \mathbb{R}$ is the function given by $f(x) = x^3$, since $f'(x) = 3x^2$, we have $f'(0) = 0$, but 0 is neither a minimum nor a maximum of f as evidenced by the graph shown in Figure 1.

Illustration of a Cubic Curve

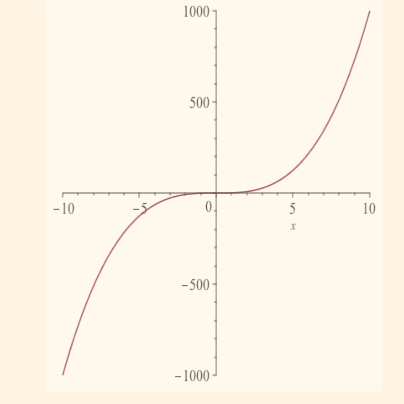


Figure 1: The graph of $f(x) = x^3$. Note that $x = 0$ is a saddle point and not a local extremum.

Necessary Condition for Local Extrema

If $g\colon \mathbb{R}^2 \to \mathbb{R}$ is the function given by $g(x, y) = x^2 - y^2$, then $g'_{(x,y)} = (2x \;\; -2y)$, so $g'_{(0,0)} = (0 \;\; 0)$, yet near $(0, 0)$ the function g takes negative and positive values. See Figure 2.

Illustration of a Hyperbolic Paraboloid

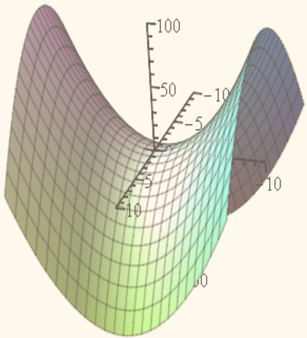
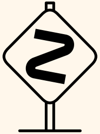


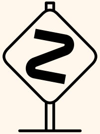
Figure 2: The graph of $g(x, y) = x^2 - y^2$. Note that $(0, 0)$ is a saddle point and not a local extremum.

Necessary Condition for Local Extrema



It is very important to note that the hypothesis that Ω *is open* is crucial for the validity of the preceding proposition.

Necessary Condition for Local Extrema



It is very important to note that the hypothesis that Ω *is open* is crucial for the validity of the preceding proposition.

For example, if J is the identity function on $\mathbb{R}$ and $U = [0, 1]$, a closed subset, then $J'(x) = 1$ for all $x \in [0, 1]$, even though J has a minimum at $x = 0$ and a maximum at $x = 1$.