Chapter 14

Extrema

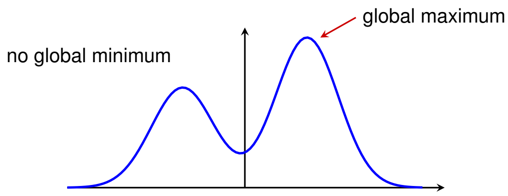
Global Extremum (Optimum)

A point x^* is called **global maximum** (*absolute maximum*) of f, if for all $x \in D_f$,

$$f(x^*) \geq f(x) \; .$$

A point x^* is called **global minimum** (*absolute minimum*) of f, if for all $x \in D_f$,

$$f(x^*) \leq f(x) \; .$$



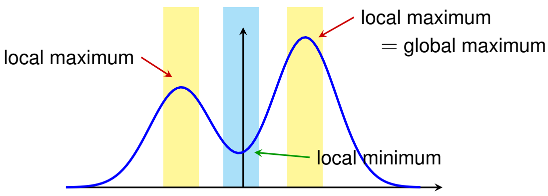
Local Extremum (Optimum)

A point x_0 is called **local maximum** (*relative maximum*) of f, if for all x in some *neighborhood* of x_0,

$$f(x_0) \geq f(x) \; .$$

A point x_0 is called **local minimum** (*relative minimum*) of f, if for all x in some neighborhood of x_0,

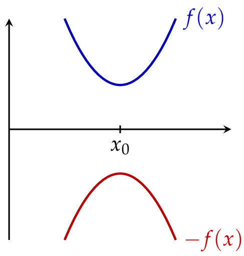
$$f(x_0) \leq f(x) \; .$$



Minima and Maxima

Notice!
Every minimization problem can be transformed into a maximization problem (and vice versa).

Point x_0 is a minimum of $f(x)$, if and only if x_0 is a maximum of $-f(x)$.



Critical Point

At a (local) maximum or minimum the first derivative of the function must vanish (i.e., must be equal to 0).

A point x_0 is called a **critical point** (or *stationary point*) of function f, if

$$f'(x_0) = 0$$

Necessary condition for differentiable functions:

Each extremum of f is a critical point of f.

Global Extremum

Sufficient condition:

Let x_0 be a critical point of f.
If f is **concave**, then x_0 is a **global maximum** of f.
If f is *convex*, then x_0 is a *global minimum* of f.

If f is **strictly** concave (or convex), then the extremum is *unique*.

This condition immediately follows from the properties of (strictly) concave functions. Indeed, we have for all $\mathbf{x} \neq \mathbf{x}_0$,

$$f(\mathbf{x}) - f(\mathbf{x}_0) \leq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \cdot (\mathbf{x} - \mathbf{x}_0) = 0$$

and thus

$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \; .$$

Example – Global Extremum / Univariate*

Let $f(x) = e^x - 2x$.

Function f is strictly convex:

$$f'(x) = e^x - 2$$
$$f''(x) = e^x \; > 0 \quad \text{for all } x \in \mathbb{R}$$

Critical point:

$$f'(x) = e^x - 2 = 0 \quad \Rightarrow \quad x_0 = \ln 2$$

$x_0 = \ln 2$ is the (unique) global minimum of f.

Example – Global Extremum / Multivariate

Let $f\colon D = [0,\infty)^2 \to \mathbb{R},\; f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - x - y$

Hessian matrix at $\mathbf{x}$:

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}\,x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$$

Leading principle minors:

$H_1 = -\frac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} \; < 0$

$H_2 = \frac{1}{2}\,x^{-\frac{3}{2}}y^{-\frac{3}{2}} \; > 0$

f is strictly concave in D.

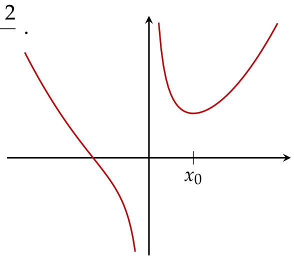
critical point: $\nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1, x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1) = 0$

$$\begin{aligned} f_x &= x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1 = 0 \\ f_y &= x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1 = 0 \end{aligned} \quad \Rightarrow \quad \mathbf{x}_0 = (1,1)$$

$\mathbf{x}_0$ is the global maximum of f.

Sources of Errors

Find all global minima of $f(x) = \dfrac{x^3+2}{3x}$.



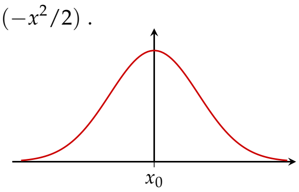
1. $f'(x) = \frac{2(x^3-1)}{3x^2}$,
   $f''(x) = \frac{2x^3+4}{3x^3}$.
2. critical point at $x_0 = 1$.
3. $f''(1) = 2 > 0$
   $\Rightarrow$ global minimum **???**

However, looking *just* at $f''(\mathbf{1})$ is not sufficient as we are looking for *global* minima!

Beware! We have to look at $f''(x)$ *at all* $x \in D_f$.
However, $f''(-1) = -\frac{2}{3} < 0$.
Moreover, domain $D = \mathbb{R} \setminus \{0\}$ is not an interval.
So f is not convex and we cannot apply our theorem.

Sources of Errors

Find all global maxima of $f(x) = \exp(-x^2/2)$.



1. $f'(x) = x\exp(-x^2)$,
   $f''(x) = (x^2-1)\exp(-x^2)$.
2. critical point at $x_0 = 0$.
3. However,
   $f''(0) = -1 < 0$ but $f''(2) = 2e^{-2} > 0$.
   So f is not concave and thus there cannot be a global maximum.
   Really ???

Beware! We are checking a *sufficient* condition.
Since an assumption does not hold (f is not concave),
we simply **cannot apply** the theorem.
We *cannot* conclude that f does not have a global maximum.

Global Extrema in $[a,b]$*

Extrema of $f(x)$ in **closed** interval $[a,b]$.

Procedure for differentiable functions:

(1) Compute $f'(x)$.

(2) Find all stationary points x_i (i.e., $f'(x_i) = 0$).

(3) Evaluate $f(x)$ for all *candidates*:
- all stationary points x_i,
- boundary points a and b.

(4) Largest of these values is **global maximum**,
smallest of these values is **global minimum**.

It is *not* necessary to compute $f''(x_i)$.

Global Extrema in $[a,b]$*

Find all *global* extrema of function

$$f\colon [0{,}5; 8{,}5] \to \mathbb{R},\ x \mapsto \frac{1}{12}x^3 - x^2 + 3x + 1$$

(1) $f'(x) = \frac{1}{4}x^2 - 2x + 3$.

(2) $\frac{1}{4}x^2 - 2x + 3 = 0$ has roots $x_1 = 2$ and $x_2 = 6$.

(3) $f(0.5) = 2.260$
$f(2) = 3.667$
$f(6) = 1.000 \quad \Rightarrow$ global minimum
$f(8.5) = 5.427 \quad \Rightarrow$ global maximum

(4) $x_2 = 6$ is the global minimum and
$b = 8.5$ is the global maximum of f.

Global Extrema in (a,b)*

Extrema of $f(x)$ in **open** interval (a,b) (or $(-\infty,\infty)$).

Procedure for differentiable functions:

(1) Compute $f'(x)$.

(2) Find all stationary points x_i (i.e., $f'(x_i) = 0$).

(3) Evaluate $f(x)$ for all *stationary* points x_i.

(4) Determine $\lim_{x\to a} f(x)$ and $\lim_{x\to b} f(x)$.

(5) Largest of these values is **global maximum**,
smallest of these values is **global minimum**.

(6) A global extremum exists **only if** the largest (smallest) value is obtained in a *stationary point*!

Global Extrema in (a,b)*

Compute all *global* extrema of

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto e^{-x^2}$$

(1) $f'(x) = -2x\,e^{-x^2}$.

(2) $f'(x) = -2x\,e^{-x^2} = 0$ has unique root $x_1 = 0$.

(3) $f(0) = 1 \quad \Rightarrow$ global maximum
$\lim_{x\to-\infty} f(x) = 0 \quad \Rightarrow$ no global minimum
$\lim_{x\to\infty} f(x) = 0$

(4) The function has a global maximum in $x_1 = 0$,
but no global minimum.

Existence and Uniqueness

- A function need not have maxima or minima:

$$f\colon (0,1) \to \mathbb{R},\ x \mapsto x$$

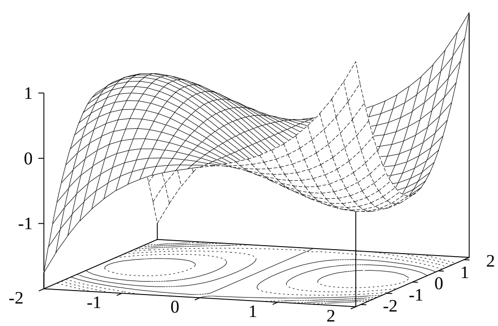
(Points 0 and 1 are not in domain $(0,1)$.)

- (Global) maxima need not be unique:

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x^4 - 2x^2$$

has two global minima at -1 and 1.

Example – Local Extrema



$$f(x,y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Local Extremum

A point x_0 is a **local maximum** (or *local minimum*) of f, if

- x_0 is a **critical point** of f,
- f is **locally concave** (and *locally convex*, resp.) around x_0.

Sufficient condition for two times differentiable functions:

> Let x_0 be a critical point of f. Then
>
> - $f''(x_0)$ *negative definite* $\Rightarrow$ x_0 is local maximum
> - $f''(x_0)$ *positive definite* $\Rightarrow$ x_0 is local minimum

It is sufficient to evaluate $f''(x)$ at the critical point x_0.
(In opposition to the condition for global extrema.)

Necessary and Sufficient

We again want to explain two important concepts using the example of local minima.

Condition "$f'(x_0) = 0$" is **necessary** for a local minimum:
Every local minimum must have this properties.
However, not every point with such a property is a local minimum
(e.g. $x_0 = 0$ in $f(x) = x^3$).
Stationary points are *candidates* for local extrema.

Condition "$f'(x_0) = 0$ *and* $f''(x_0)$ *is positive definite*" is **sufficient** for a local minimum.
If it is satisfied, then x_0 is a local minimum.
However, there are local minima where this condition does not hold
(e.g. $x_0 = 0$ in $f(x) = x^4$).
If it is *not* satisfied, we cannot draw *any conclusion*.

Procedure – Univariate Functions*

Sufficient condition
for local extrema of a differentiable function in *one* variable:

> 1. Compute $f'(x)$ and $f''(x)$.
> 2. Find all roots x_i of $f'(x_i) = 0$ (critical points).
> 3. If $f''(x_i) < 0$ $\Rightarrow$ x_i is a *local maximum*.
>   If $f''(x_i) > 0$ $\Rightarrow$ x_i is a *local minimum*.
>   If $f''(x_i) = 0$ $\Rightarrow$ *no conclusion possible*!

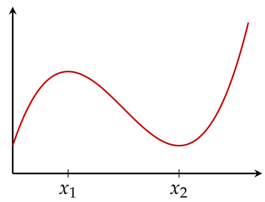
If $f''(x_i) = 0$ we need more sophisticated methods!
(E.g., terms of higher order of the Taylor series expansion around x_i.)

Example – Local Extrema*

Find all local extrema of

$$f(x) = \frac{1}{12}x^3 - x^2 + 3x + 1$$

1. $f'(x) = \frac{1}{4}x^2 - 2x + 3$,
   $f''(x) = \frac{1}{2}x - 2$.
2. $\frac{1}{4}x^2 - 2x + 3 = 0$
   has roots
   $x_1 = 2$ and $x_2 = 6$.
3. $f''(2) = -1$ $\Rightarrow$ x_1 is a local maximum.
   $f''(6) = 1$ $\Rightarrow$ x_2 is a local minimum.



Example – Critical Points

Compute all critical points of

$$f(x,y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Partial derivatives:

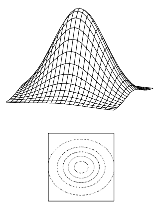
$$\begin{aligned} (I) &\quad f_x = \tfrac{1}{2}x^2 - 1 + \tfrac{1}{4}y^2 = 0 \\ (II) &\quad f_y = \tfrac{1}{2}xy = 0 \end{aligned}$$

$$\begin{array}{lllcl} (II) & \Rightarrow & x = 0 & \text{or} & y = 0 \\ (I) & \Rightarrow & -1 + \tfrac{1}{4}y^2 = 0 & \Big| & \tfrac{1}{2}x^2 - 1 = 0 \\ & & y = \pm 2 & \Big| & x = \pm\sqrt{2} \end{array}$$

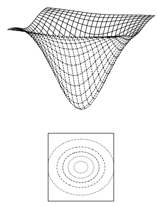
Critical points:

$$\mathbf{x}_1 = (0, 2) \qquad \mathbf{x}_3 = (\sqrt{2}, 0)$$
$$\mathbf{x}_2 = (0, -2) \qquad \mathbf{x}_4 = (-\sqrt{2}, 0)$$

Critical Point – Local Extrema

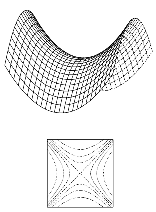


local maximum

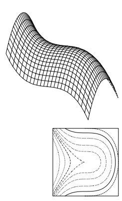


local minimum

Critical Point – Saddle Point



saddle point



example for higher order

Procedure – Local Extrema

1. Compute gradient $\nabla f(x)$ and Hessian matrix $\mathbf{H}_f$.
2. Find all $\mathbf{x}_i$ with $\nabla f(\mathbf{x}_i) = 0$ (critical points).
3. Compute leading principle minors H_k for all *critical points* $\mathbf{x}_i$:
   (a) All leading principle minors $H_k > 0$
   $\Rightarrow$ $\mathbf{x}_0$ is a **locale minimum** of f.
   (b) For all leading principle minors, $(-1)^k H_k > 0$
   [i.e., $H_1, H_3, \ldots < 0$ and $H_2, H_4, \ldots > 0$]
   $\Rightarrow$ $\mathbf{x}_0$ is a **locale maximum** of f.
   (c) $\det(\mathbf{H}_f(\mathbf{x}_i)) \neq 0$ but neither (a) nor (b) is satisfied
   $\Rightarrow$ $\mathbf{x}_0$ is a **saddle point** of f.
   (d) Otherwise *no conclusion* can be drawn,
   i.e., $\mathbf{x}_i$ may or may not be an extremum or saddle point.

Procedure – Bivariate Function

1. Compute gradient $\nabla f(x)$ and Hessian matrix $\mathbf{H}_f$.
2. Find all $\mathbf{x}_i$ with $\nabla f(\mathbf{x}_i) = 0$ (critical points).
3. Compute leading principle minors H_k for all *critical points* $\mathbf{x}_i$:

   (a) $H_2 > 0$ and $H_1 > 0$
   $\Rightarrow$ $\mathbf{x}_0$ is a **locale minimum** of f.

   (b) $H_2 > 0$ and $H_1 < 0$
   $\Rightarrow$ $\mathbf{x}_0$ is a **locale maximum** of f.

   (c) $H_2 < 0$
   $\Rightarrow$ $\mathbf{x}_0$ is a **saddle point** of f.

   (d) $H_2 = \det(\mathbf{H}_f(\mathbf{x}_0)) = 0$
   $\Rightarrow$ *no conclusion* can be drawn,
   i.e., $\mathbf{x}_i$ may or may not be an extremum or saddle point.

Example – Bivariate Function

Compute all local extrema of of

$$f(x,y) = \frac{1}{6}x^3 - x + \frac{1}{4}x\,y^2$$

1. $\nabla f = (\frac{1}{2}x^2 - 1 + \frac{1}{4}y^2, \frac{1}{2}x\,y)$

   $$\mathbf{H}_f(x,y) = \begin{pmatrix} x & \frac{1}{2}y \\ \frac{1}{2}y & \frac{1}{2}x \end{pmatrix}$$

2. Critical points:

   $$\mathbf{x}_1 = (0,2),\ \mathbf{x}_2 = (0,-2),\ \mathbf{x}_3 = (\sqrt{2},0),\ \mathbf{x}_4 = (-\sqrt{2},0)$$

Example – Bivariate Function / Cont.

3. Leading principle minors:

   $$\mathbf{H}_f(\mathbf{x}_1) = \mathbf{H}_f(0,2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

   $H_2 = -1 < 0$ $\Rightarrow$ $\mathbf{x}_1$ is a *saddle point*

   $$\mathbf{H}_f(\mathbf{x}_2) = \mathbf{H}_f(0,-2) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

   $H_2 = -1 < 0$ $\Rightarrow$ $\mathbf{x}_2$ is a *saddle point*

Example – Bivariate Function / Cont.

3. Leading principle minors:

   $$\mathbf{H}_f(\mathbf{x}_3) = \mathbf{H}_f(\sqrt{2},0) = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

   $H_2 = 1 > 0$ and $H_1 = \sqrt{2} > 0$
   $\Rightarrow$ $\mathbf{x}_3$ is a *local minimum*

   $$\mathbf{H}_f(\mathbf{x}_4) = \mathbf{H}_f(-\sqrt{2},0) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

   $H_2 = 1 > 0$ and $H_1 = -\sqrt{2} < 0$
   $\Rightarrow$ $\mathbf{x}_4$ is a *local maximum*

Derivative of Optimal Value

Let $p, r > 0$ and $f\colon D = [0,\infty)^2 \to \mathbb{R}$, $f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$

Hessian matrix: $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$

Leading principle minors:

$H_1 = -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} < 0$

$H_2 = \frac{1}{2}x^{-\frac{3}{2}}y^{-\frac{3}{2}} > 0$

f is strictly concave in D.

Critical point: $\nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - p, x^{\frac{1}{4}}y^{-\frac{3}{4}} - r) = 0$

$f_x = x^{-\frac{3}{4}}y^{\frac{1}{4}} - p = 0$

$f_y = x^{\frac{1}{4}}y^{-\frac{3}{4}} - r = 0$

$\Rightarrow$ $\mathbf{x}_0 = \left(\sqrt{\frac{1}{r\,p^3}}, \sqrt{\frac{1}{r^3 p}}\right)$

$\mathbf{x}_0$ is the global maximum of f.

Question:
What is the derivative of optimal value $f^* = f(\mathbf{x}_0)$ w.r.t. r or p?

Envelope Theorem

We are given function

$\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{r})$

$\mathbf{x} = (x_1, \dots, x_n)$ ... variable (endogeneous)

$\mathbf{r} = (r_1, \dots, r_k)$ ... parameter (exogeneous)

with extremum $\mathbf{x}^*$.

This extremum depends on parameter $\mathbf{r}$:

$\mathbf{x}^* = \mathbf{x}^*(\mathbf{r})$

and so does the optimal value f^*:

$f^*(\mathbf{r}) = f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})$

We have:

$$\boxed{\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}}$$

Envelope Theorem / Proof Idea

$$\begin{aligned}
\frac{\partial f^*(\mathbf{r})}{\partial r_j} &= \left.\frac{\partial f(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \qquad \text{[chain rule]} \\
&= \left.\sum_{i=1}^{n} \underbrace{f_{x_i}(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}_{\substack{=0 \\ \text{as } \mathbf{x}^* \text{ is a critical point}}} \cdot \frac{\partial x_i^*(\mathbf{r})}{\partial r_j} + \frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \\
&= \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}
\end{aligned}$$

Example – Envelope Theorem

The (unique) maximum of

$$f\colon D = [0,\infty)^2 \to \mathbb{R},\quad f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$$

is $\mathbf{x}^*(p,r) = (x^*(p,r), y^*(p,r)) = \left(\sqrt{\frac{1}{r\,p^3}}, \sqrt{\frac{1}{r^3 p}}\right)$.

Question:
What is the derivative of optimal value $f^* = f(\mathbf{x}_0)$ w.r.t. r or p?

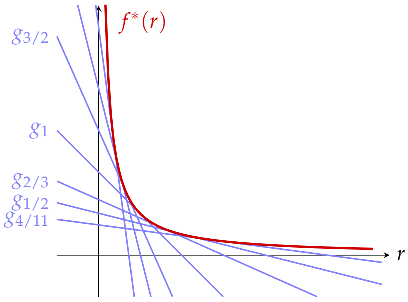
$$\frac{\partial f^*(p,r)}{\partial p} = \left.\frac{\partial f(\mathbf{x};p,r)}{\partial p}\right|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -x\Big|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -\sqrt{\frac{1}{r\,p^3}}$$

$$\frac{\partial f^*(p,r)}{\partial r} = \left.\frac{\partial f(\mathbf{x};p,r)}{\partial r}\right|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -y\Big|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -\sqrt{\frac{1}{r^3 p}}$$

A Geometric Interpretation

Let $f(x,r) = \sqrt{x} - rx$. We want $f^*(r) = \max_x f(x,r)$.

Graphs of $g_x(r) = f(x,r)$ for various values of x.



Summary

- global extremum
- local extremum
- minimum, maximum and saddle point
- critical point
- hessian matrix and principle minors
- envelope theorem