ECON2125/4021/8013

Lecture 19

John Stachurski

Semester 1, 2015

Introduction

In this lecture we study topics such as

- Convexity / concavity
  - and uniqueness in optimization
  - sufficient conditions for optimality
  - how to detect these properties?

- Zeros of functions
  - solving nonlinear equations
  - existence of solutions
  - applications

Convex Sets

Uniqueness of optima often connected to convexity / concavity

- Convexity is a shape property for sets
- Convexity and concavity are shape properties for functions

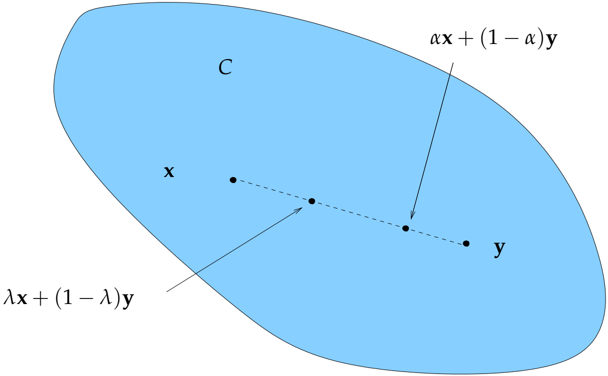
However, only one fundamental concept: convex sets

A set $C \subset \mathbb{R}^K$ is called **convex** if

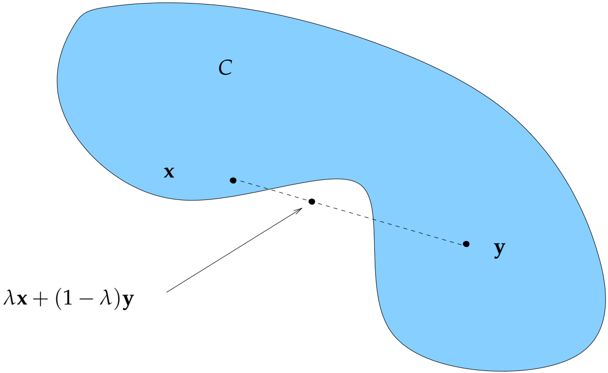
$$\mathbf{x}, \mathbf{y} \text{ in } C \text{ and } 0 \leq \lambda \leq 1 \implies \lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in C$$

Remark: This is vector addition and scalar multiplication

Convexity $\iff$ line between any two points in C lies in C



A non-convex set



Example. The “positive cone” $P := \{\mathbf{x} \in \mathbb{R}^K : \mathbf{x} \geq \mathbf{0}\}$ is convex

To see this, pick any $\mathbf{x}$, $\mathbf{y}$ in P and any $\lambda \in [0, 1]$

Let $\mathbf{z} := \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ and let $z_k := \mathbf{e}_k'\mathbf{z}$

Since

- $z_k = \lambda x_k + (1 - \lambda) y_k$
- $x_k \geq 0$ and $y_k \geq 0$

It is clear that $z_k \geq 0$ for all k

Hence $\mathbf{z} \in P$ as claimed

Example. Every ϵ-ball is convex

Proof: Fix $\mathbf{a} \in \mathbb{R}^K$, $\epsilon > 0$ and let $B_\epsilon(\mathbf{a})$ be the ϵ-ball

Pick any $\mathbf{x}$, $\mathbf{y}$ in $B_\epsilon(\mathbf{a})$ and any $\lambda \in [0, 1]$

The point $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}$ lies in $B_\epsilon(\mathbf{a})$ because

$$\begin{aligned}
\|\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} - \mathbf{a}\| &= \|\lambda \mathbf{x} - \lambda \mathbf{a} + (1 - \lambda)\mathbf{y} - (1 - \lambda)\mathbf{a}\| \\
&\leq \|\lambda \mathbf{x} - \lambda \mathbf{a}\| + \|(1 - \lambda)\mathbf{y} - (1 - \lambda)\mathbf{a}\| \\
&= \lambda \|\mathbf{x} - \mathbf{a}\| + (1 - \lambda)\|\mathbf{y} - \mathbf{a}\| \\
&< \lambda \epsilon + (1 - \lambda)\epsilon \\
&= \epsilon
\end{aligned}$$

Example. Let $\mathbf{p} \in \mathbb{R}^K$ and let M be the "half-space"

$$M := \{\mathbf{x} \in \mathbb{R}^K : \mathbf{p}'\mathbf{x} \leq m\}$$

The set M is convex

Proof: Let $\mathbf{p}$, m and M be as described

Fix $\mathbf{x}$, $\mathbf{y}$ in M and $\lambda \in [0, 1]$

Then $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in M$ because

$$\mathbf{p}'[\lambda\mathbf{x} + (1-\lambda)\mathbf{y}] =$$

$$\lambda\mathbf{p}'\mathbf{x} + (1-\lambda)\mathbf{p}'\mathbf{y} \leq \lambda m + (1-\lambda)m = m$$

Hence M is convex

Fact. If A and B are convex, then so is $A \cap B$

Proof: Let A and B be convex and let $C := A \cap B$

Pick any $\mathbf{x}$, $\mathbf{y}$ in C and any $\lambda \in [0, 1]$

Set

$$\mathbf{z} := \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$$

Since $\mathbf{x}$ and $\mathbf{y}$ lie in A and A is convex we have $\mathbf{z} \in A$

Since $\mathbf{x}$ and $\mathbf{y}$ lie in B and B is convex we have $\mathbf{z} \in B$

Hence $\mathbf{z} \in A \cap B$

Example. Let $\mathbf{p} \in \mathbb{R}^K$ be a vector of prices and consider the budget set

$$B(m) := \{\mathbf{x} \in \mathbb{R}^K : \mathbf{x} \geq \mathbf{0} \text{ and } \mathbf{p}'\mathbf{x} \leq m\}$$

The budget set $B(m)$ is convex

To see this, note that $B(m) = P \cap M$ where

$$P := \{\mathbf{x} \in \mathbb{R}^K : \mathbf{x} \geq \mathbf{0}\} \qquad M := \{\mathbf{x} \in \mathbb{R}^K : \mathbf{p}'\mathbf{x} \leq m\}$$

We already know that

- P and M are convex, intersections of convex sets are convex

Hence $B(m)$ is convex

Convex Functions

Let $A \subset \mathbb{R}^K$ be a convex set and let f be a function from A to $\mathbb{R}$

f is called **convex** if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$ and all $\lambda \in [0,1]$

f is called **strictly convex** if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$ with $\mathbf{x} \neq \mathbf{y}$ and all $\lambda \in (0,1)$

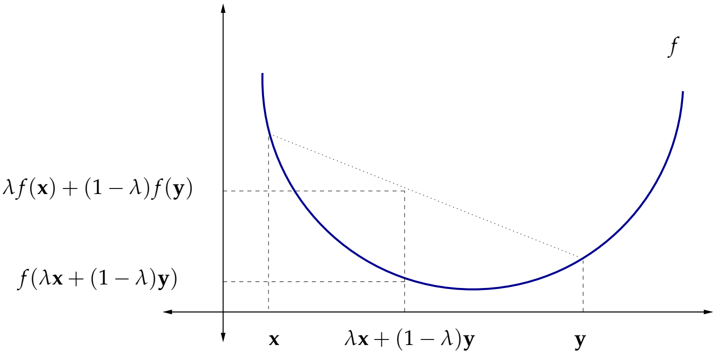
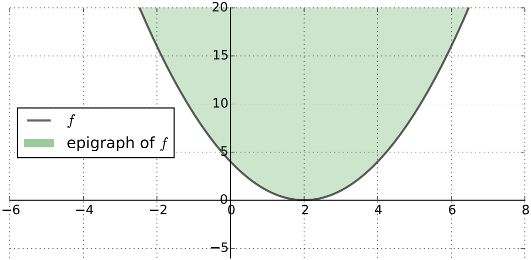


Figure : A strictly convex function on a subset of $\mathbb{R}$

Fact. $f \colon A \to \mathbb{R}$ is convex if and only if its **epigraph**

$$E_f := \{(\mathbf{x}, y) \in A \times \mathbb{R} : f(\mathbf{x}) \leq y\}$$

is a convex subset of $\mathbb{R}^K \times \mathbb{R}$



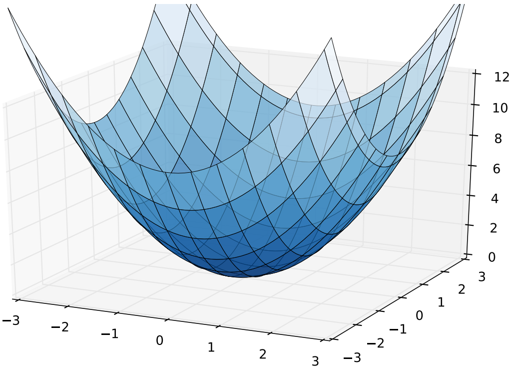


Figure : A strictly convex function on a subset of $\mathbb{R}^2$

Example. $f(\mathbf{x}) = \|\mathbf{x}\|$ is convex on $\mathbb{R}^K$

To see this recall that, by the properties of norms,

$$\begin{aligned} \|\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\| &\leq \|\lambda \mathbf{x}\| + \|(1-\lambda)\mathbf{y}\| \\ &= \lambda\|\mathbf{x}\| + (1-\lambda)\|\mathbf{y}\| \end{aligned}$$

That is,

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Example. $f(x) = \cos(x)$ is not convex on $\mathbb{R}$ because

$$1 = f(2\pi) = f(\pi/2 + 3\pi/2) > f(\pi)/2 + f(3\pi)/2 = -1$$

Fact. If $\mathbf{A}$ is $K \times K$ and positive definite, then

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} \qquad (\mathbf{x} \in \mathbb{R}^K)$$

is strictly convex on $\mathbb{R}^K$

Proof: Fix $\mathbf{x}, \mathbf{y} \in \mathbb{R}^K$ with $\mathbf{x} \neq \mathbf{y}$ and $\lambda \in (0,1)$

Ex. Show that

$$\lambda Q(\mathbf{x}) + (1-\lambda)Q(\mathbf{y}) - Q(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = \lambda(1-\lambda)(\mathbf{x}-\mathbf{y})'\mathbf{A}(\mathbf{x}-\mathbf{y})$$

Since $\mathbf{x} - \mathbf{y} \neq \mathbf{0}$ and $0 < \lambda < 1$, the right hand side is > 0

Hence

$$\lambda Q(\mathbf{x}) + (1-\lambda)Q(\mathbf{y}) > Q(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$$

Concave Functions

Let $A \subset \mathbb{R}^K$ be a convex and let f be a function from A to $\mathbb{R}$

f is called **concave** if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$ and all $\lambda \in [0,1]$

f is called **strictly concave** if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) > \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$ with $\mathbf{x} \neq \mathbf{y}$ and all $\lambda \in (0,1)$

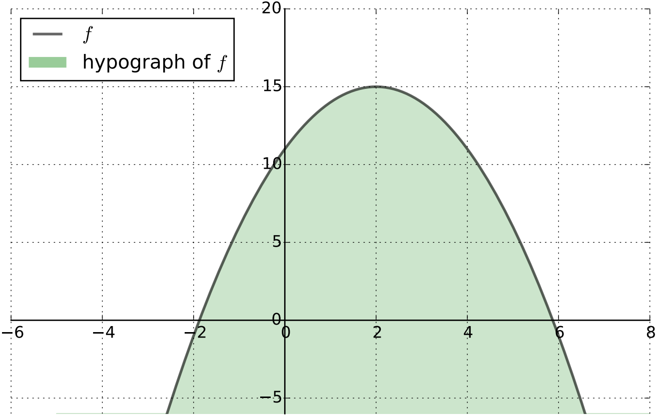
Ex. Show that

1. f is concave if and only if $-f$ is convex
2. f is strictly concave if and only if $-f$ is strictly convex

Fact. $f\colon A \to \mathbb{R}$ is concave if and only if its **hypograph**

$$H_f := \{(\mathbf{x}, y) \in A \times \mathbb{R} : f(\mathbf{x}) \geq y\}$$

is a convex subset of $\mathbb{R}^K \times \mathbb{R}$



f
hypograph of f
20
15
10
5
0
−5
−6
−4
−2
0
2
4
6
8

Preservation of Shape

Let $A \subset \mathbb{R}^K$ be convex and let f and g be functions from A to $\mathbb{R}$

Fact. If f and g are convex (resp., concave) and $\alpha \geq 0$, then

- αf is convex (resp., concave)
- $f + g$ is convex (resp., concave)

Fact. If f and g are strictly convex (resp., strictly concave) and $\alpha > 0$, then

- αf is strictly convex (resp., strictly concave)
- $f + g$ is strictly convex (resp., strictly concave)

Let's prove that f and g convex $\implies h := f + g$ convex

Pick any $\mathbf{x}, \mathbf{y} \in A$ and $\lambda \in [0, 1]$

We have

$$\begin{aligned} h(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &= f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) + g(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \\ &\leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) + \lambda g(\mathbf{x}) + (1-\lambda) g(\mathbf{y}) \\ &= \lambda [f(\mathbf{x}) + g(\mathbf{x})] + (1-\lambda)[f(\mathbf{y}) + g(\mathbf{y})] \\ &= \lambda h(\mathbf{x}) + (1-\lambda) h(\mathbf{y}) \end{aligned}$$

Hence h is convex

Derivative Conditions

The i, j-th cross partial of $f \colon A \to \mathbb{R}$ at $\mathbf{x} \in A$ is

$$f_{ij}(\mathbf{x}) := \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}) \qquad (1 \leq i, j \leq K)$$

We say that f is **a C^2 function** if these partials are all continuous in $\mathbf{x}$ for all $\mathbf{x} \in A$

The **Hessian matrix** of f at $\mathbf{x}$ is the matrix of cross partials

$$H(\mathbf{x}) := \begin{pmatrix} f_{11}(\mathbf{x}) & \cdots & f_{1K}(\mathbf{x}) \\ & \vdots & \\ f_{K1}(\mathbf{x}) & \cdots & f_{KK}(\mathbf{x}) \end{pmatrix}$$

Fact. If $f\colon A \to \mathbb{R}$ is a C^2 function where $A \subset \mathbb{R}^K$ is open and convex, then

1. $H(\mathbf{x})$ nonnegative definite for all $\mathbf{x} \in A \iff f$ convex
2. $H(\mathbf{x})$ nonpositive definite for all $\mathbf{x} \in A \iff f$ concave

In addition,

1. $H(\mathbf{x})$ positive definite for all $\mathbf{x} \in A \implies f$ strictly convex
2. $H(\mathbf{x})$ negative definite for all $\mathbf{x} \in A \implies f$ strictly concave

Proof: Omitted

Example. Let $A := (0, \infty) \times (0, \infty)$ and let $U \colon A \to \mathbb{R}$ be the utility function

$$U(c_1, c_2) = \alpha \ln c_1 + \beta \ln c_2$$

Assume that α and β are both strictly positive

Ex. Show that the Hessian at $\mathbf{c} := (c_1, c_2) \in A$ has the form

$$H(\mathbf{c}) := \begin{pmatrix} -\frac{\alpha}{c_1^2} & 0 \\ 0 & -\frac{\beta}{c_2^2} \end{pmatrix}$$

Ex. Show that any diagonal matrix with strictly negative elements along the principle diagonal is negative definite

Conclude that U is strictly concave on A

Uniqueness of Maximizers and Minimizers

Let $A \subset \mathbb{R}^K$ be convex and let $f \colon A \to \mathbb{R}$

Facts

1. If f is strictly convex, then f has at most one minimizer on A
2. If f is strictly concave, then f has at most one maximizer on A

Interpretation, strictly concave case:

- we don't know in general if f has a maximizer
- but if it does, then it has exactly one
- in other words, we have uniqueness

Proof for the case where f is strictly concave:

Suppose to the contrary that

- $\mathbf{a}$ and $\mathbf{b}$ are distinct points in A
- both are maximizers of f on A

By the def of maximizers, $f(\mathbf{a}) \geq f(\mathbf{b})$ and $f(\mathbf{b}) \geq f(\mathbf{a})$

Hence we have $f(\mathbf{a}) = f(\mathbf{b})$

By strict concavity, then

$$f\left(\frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b}\right) > \frac{1}{2}f(\mathbf{a}) + \frac{1}{2}f(\mathbf{b}) = \frac{1}{2}f(\mathbf{a}) + \frac{1}{2}f(\mathbf{a}) = f(\mathbf{a})$$

This contradicts the assumption that $\mathbf{a}$ is a maximizer

A Sufficient Condition

We can now restate more precisely optimization results stated in the introductory lectures

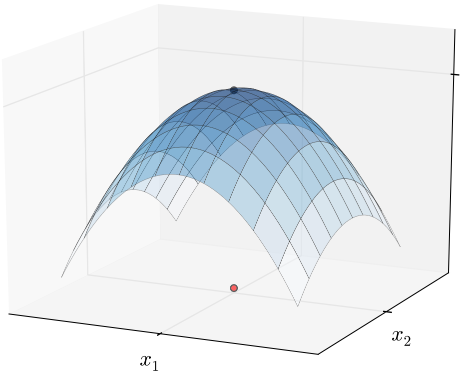
Let $f \colon A \to \mathbb{R}$ be a C^2 function where $A \subset \mathbb{R}^K$ is open, convex

Recall that $\mathbf{x}^* \in A$ is a stationary point of f if

$$\frac{\partial}{\partial x_i} f(\mathbf{x}^*) = 0 \quad \text{for all } i \text{ in } 1, \ldots, K$$

Fact. If f and A are as above and $\mathbf{x}^* \in A$ is stationary, then

1. f strictly concave $\implies$ $\mathbf{x}^*$ is the unique maximizer of f on A
2. f strictly convex $\implies$ $\mathbf{x}^*$ is the unique minimizer of f on A


x_1
x_2

Example. In an introductory lecture we studied the problem

$$\max_{k,\ell} \pi(k,\ell) := pk^\alpha \ell^\beta - w\ell - rk$$

where all parameters are > 0 and $\alpha + \beta < 1$

Points on the boundary (either $k = 0$ or $\ell = 0$) generate ≤ 0 profits and hence are never maximal

Hence we concentrate on interior points:

$$\max_{(k,\ell) \in A} \pi(k,\ell) \quad \text{where} \quad A := (0,\infty) \times (0,\infty)$$

Ex. Show that A is open and convex

We already showed that π is strictly concave, so any stationary point is a unique maximizer

Algorithms

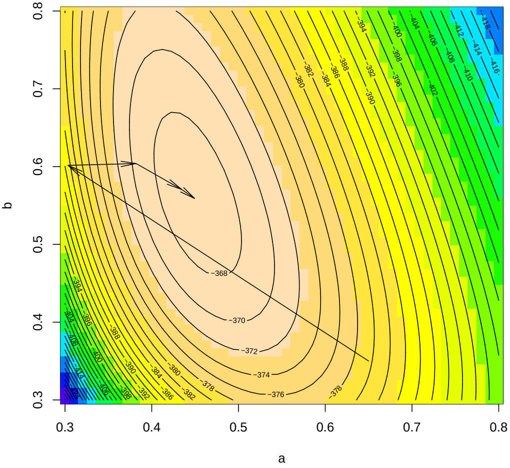
Another benefit of concavity / convexity for optimization: finding optima on computers is much easier

A sample algorithm might be

1. Start at some $\mathbf{x}$
2. Evaluate the slope of f at $\mathbf{x}$
3. Take a step "uphill" to a new point $\mathbf{y}$
4. Set $\mathbf{x}$ to $\mathbf{y}$ and go to step 2

For more details look up "hill climbing" or "steepest ascent"

If f is concave then this procedure typically converges


b
0.3
0.4
0.5
0.6
0.7
0.8
a
0.3
0.4
0.5
0.6
0.7
0.8
−368
−370
−372
−374
−376
−378

Zeros of Functions

Let $f\colon A \to \mathbb{R}$ where $A \subset \mathbb{R}$

A point $\bar{x} \in A$ is called a **zero** or **root** of f if $f(\bar{x}) = 0$

Example. If $f\colon \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = |x|$ then 0 is the unique zero of f

Example. If $f\colon \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x - b$ then b is the unique zero of f

Example. If $f\colon \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = (x-1)(x+1)$ then -1 and 1 are both zeros of f

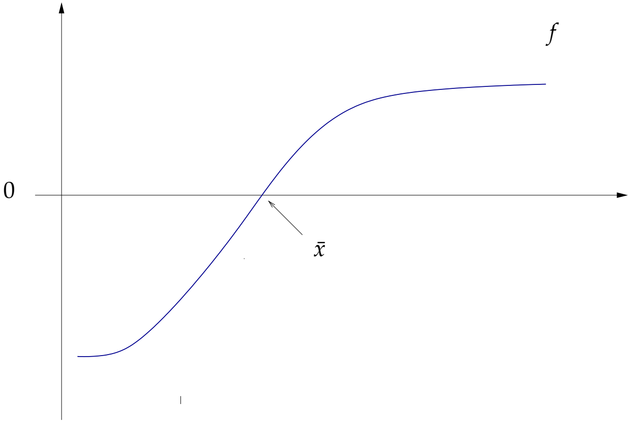


Figure : Zero of a function

The problem of finding zeros is important for many reasons

One example is finding stationary points of functions

Another is solving nonlinear equations

Example. Suppose we want to find all x such that

$$g(x) = b \tag{$\star$}$$

We can recast this as a problem of finding zeros by defining

$$f(x) := g(x) - b$$

Now x is a zero of $f \iff x$ solves $(\star)$

Example. The McCall job search model

Features an agent who decides when to accept a job offer

In a simplified version of the model, the agent

- receives offer w_t in period t where $\{w_t\}$ is IID
- accepts this offer at time t or remains unemployed
  - if unemployed receives compensation $c > 0$
  - if accepts then works indefinitely at this wage
- discounts the future at rate $\beta \in (0, 1)$

Optimal strategy: set a reservation wage $\bar{w}$

- Accept the first offer w_t such that $w_t \geq \bar{w}$

It can be shown (details omitted) that $\bar{w}$ should satisfy

$$\frac{\bar{w}}{1-\beta} = c + \frac{\beta}{1-\beta} \sum_{k=1}^{K} \max\left\{w_k, \bar{w}\right\} p_k \qquad (\star)$$

- $w_1, \ldots, w_K$ are possible wage values with pmf $p_1, \ldots, p_K$

Does there exists a $\bar{w} \in [0, \infty)$ that solves $(\star)$?

To study this problem, let

$$f(x) = \frac{x}{1-\beta} - c - \frac{\beta}{1-\beta} \sum_{k=1}^{K} \max\left\{w_k, x\right\} p_k$$

We seek a zero of f on $[0, \infty)$

Existence of Zeros

Of course zeros can fail to exist

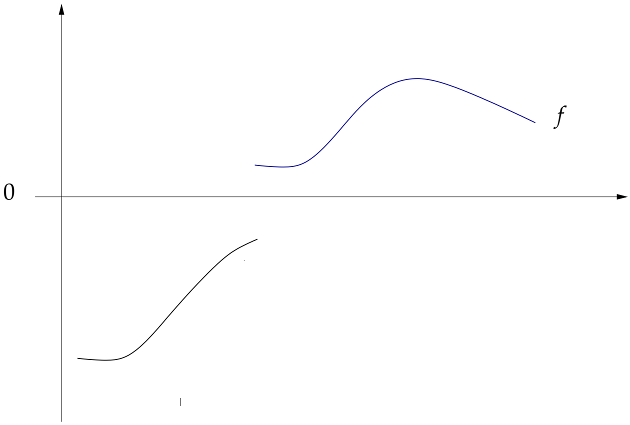
Example. If $f(x) > 0$ on its domain then f has no zero

Example. If $f(x) < 0$ on its domain then f has no zero

A more interesting case is when

- $f(x) \leq 0$ for some x
- $f(x) \geq 0$ for some x

But even then we don't always have a zero


0
f

Let $f\colon [a,b] \to \mathbb{R}$

Fact. (Intermediate Value Theorem) If $f(a) < 0 < f(b)$ and f is continuous, then f has a zero in $[a,b]$

Sketch of proof: Let

- $N := \{x \in [a,b] : f(x) < 0\}$
- $\bar{x} := \sup N$

It can be shown from the hypotheses that $f(\bar{x}) = 0$

Details will be given in the solved exercises

Ex. Using the IVT, show that the same result holds if f is continuous and $f(b) < 0 < f(a)$

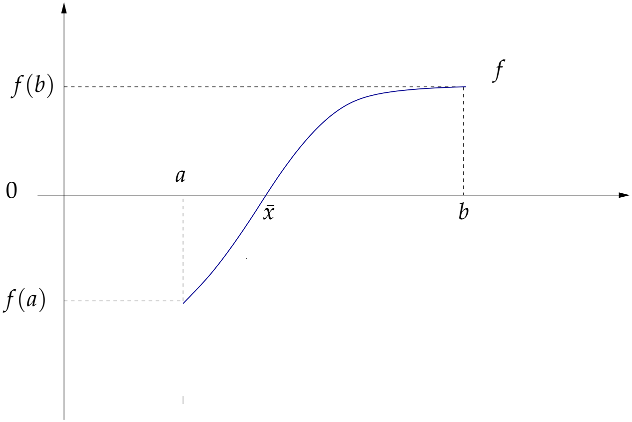


Figure : Existence of a root

Example. Let $f\colon [0,1] \to \mathbb{R}$ be defined by

$$f(x) = \sin(4(x - 1/4)) + x + x^{20} - 1$$

This function is continuous on $[0,1]$

Moreover,

- $f(0) = \sin(-1) - 1 < 0$
- $f(1) = \sin(3) + 1 > 0$

Hence f has at least on zero on $[0,1]$

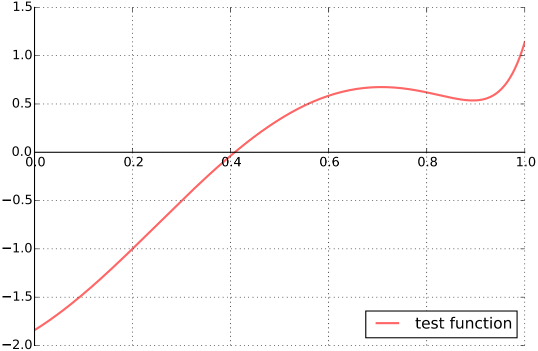
Obtaining the zero using a bisection algorithm:

```
In [3]: import numpy as np

In [4]: from scipy.optimize import bisect

In [5]: def f(x):
   ...:     return np.sin(4*(x - 0.25)) + x + x**20 - 1
   ...:

In [6]: bisect(f, 0, 1)
Out[6]: 0.4082935042797544
```


1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0.0
0.2
0.4
0.6
0.8
1.0
test function

Example. Recall that in solving the McCall model we sought a zero of

$$f(x) = \frac{x}{1-\beta} - c - \frac{\beta}{1-\beta} \sum_{k=1}^{K} \max\{w_k, x\} \, p_k$$

where

- $p_1, \ldots, p_K$ is a pmf and $0 < w_k < \infty$
- $c > 0$ and $\beta \in (0, 1)$

This function is continuous — details omitted (but not hard)

We claim that $f(0) < 0 < f(\hat{x})$ when

$$\hat{x} := \max\{c, w_1, \ldots, w_K\} + 1$$

To show that $f(\hat{x}) > 0$, note that $\hat{x} > w_k$ for all k

Hence $\max\{w_k, \hat{x}\} = \hat{x}$, and

$$f(\hat{x}) = \frac{\hat{x}}{1-\beta} - c - \frac{\beta}{1-\beta} \sum_{k=1}^{K} \max\{w_k, \hat{x}\} \, p_k$$

$$= \frac{\hat{x}}{1-\beta} - c - \frac{\beta \hat{x}}{1-\beta} = \hat{x} - c$$

By construction, $\hat{x} > c$

Hence $f(\hat{x}) > 0$ as claimed

Ex. Show that $f(0) < 0$ also holds

Conclusion: f has at least one solution on $[0, \hat{x}]$