EC5555
Economics Masters Refresher Course in Mathematics
September 2013

Lecture 4 – Optimization

Francesco Feri

Uses of Derivatives: Optimisation

Learning objectives.

- The use of the implicit function
- Know the definition of a convex function, a convex set and a concave function.
- Understand the relationship between a convex set and a convex function.
- Perform unconstrained Optimization

Implicit Functions

In most economic problems we have:

- exogenous variables (often called parameters)

  E.g. prices facing the consumer or wages and prices facing the competitive firm

- endogenous variables (determined within the system)

  E.g. demand for the consumer or supply and input demand for the firm

Often we want to know how the endogenous variables change as the exogenous variables alter

E.g. how does demand for oranges change if the price of apples rises?

The technique for finding the answer is called ***‘comparative statics’***

In comparative statics, we often may not need a specific number; just a direction of change. E.g. demand for oranges rises if price of apple rises

Sometimes solving this is straightforward because we have functions that directly relate the demand for apples to the price of oranges

But sometimes we don’t and that is where implicit functions can help

If a function is written in the form of

$$y = f(x) \quad \text{e.g., } y = 3x^2$$

then it is called an ***explicit function*** - because y is explicitly given as a function of x

Sometimes functions are given in the form

$$y - f(x) = 0 \quad \text{e.g., } y - 3x^2 = 0$$

in which case it is called an ***implicit function***

Implicit equations are written in a general form as $F(y, x) = 0$

(since the left hand side is a function of both variables)

While we can always change an explicit function into an implicit function (by taking f(x) to the other side of the equality) the reverse is not always true

Eg1: $x^2 + y^2 = 0$ does not define a function since it is only true at (0,0)

Eg 2: $x^2 + y^2 - 9 = 0$ is a ***relation*** not a function (there is no single value of y for a given value of x - it is the equation of a circle

For example, $x = 1$ gives $y = \pm\sqrt{8}$

The Implicit Function Theorem

It is a rule which allows us to see if an implicit function does imply an explicit function

If $F(y, x_1, x_2, \ldots. x_N)$ has continuous partial derivatives F_y, F_{x1}, … F_{xn}, at a point $(y_0, x_{01}, x_{02}, \ldots. x_{0n})$ and F_y is non-zero then there exists an implicit function at that point $y_0 = f(x_{01}, x_{02}, \ldots. x_{0n})$

Eg. $x^2 + y^2 – 9= 0$

$F_y = 2y$ (continuous and non-zero except at $y = 0$)

$F_x = 2x$ (continuous)

So there are some points on the circle that do define an explicit function $y = f(x)$ which give a unique value of y for a given value of x

The Implicit Function Rule

1) If two expressions are identical their respective total differentials must be equal

   Eg the identity $x^2 - y^2 = (x + y)(x - y)$

   total differential

   left: $2xdx - 2ydy$ right: $(x-y)d(x+y) + (x+y)d(x-y)$

   $(x-y)(dx+dy) + (x+y)(dx-dy)$

   $= 2xdx - 2ydy$

2) Differentiation of an expression that involves $y, x_1, x_2, \ldots, x_n$ gives an expression that involves the differentials $dy, dx_1, dx_2, \ldots, dx_n$

3) If so then we can solve for any partial $\delta y/\delta x_i$

   i.e. the implicit function $F(y, x_1, x_2, \ldots x_N) = 0$

   has total differential $dF = 0$

$$\text{So } dF = F_y dy + F_1 dx_1 + F_2 dx_2 + \ldots F_n dx_n = 0 \qquad (1)$$

Now, we know the explicit function $y = f(x_1, x_2, \ldots x_n)$ also has a total differential

$$dy = f_1 dx_1 + f_2 dx_2 + \ldots f_n dx_n \qquad (2)$$

Replacing (2) into (1)

$dF = F_y(f_1 dx_1 + f_2 dx_2 + \ldots. f_n dx_n) + F_1 dx_1 + F_2 dx_2 + \ldots. F_n dx_n = 0$

So

$(F_y f_1 + F_1) dx_1 + (F_y f_2 + F_2) dx_2 + \ldots. (F_y f_n + F_n) dx_n = 0$

Since the x variables can vary independently this is only equal to zero if ***each*** $(F_y f_i + F_i) dx_i = 0$

Dividing by F_y and solving for f_i gives the ***implicit function rule***

$$f_i = \frac{\partial y}{\partial x_i} = -\frac{F_i}{F_y}$$

So the value of the explicit partial can be obtained if we have information on the implicit function

Eg 1. $F(y,x) = y - 3x^4 = 0$, find dy/dx

$$f_i = \frac{\partial y}{\partial x_i} = -\frac{F_i}{F_y} = -\frac{F_x}{F_y} = \frac{12x^3}{1} = 12x^3$$

Eg 2. F(Q, K, L) = 0 implicitly defines a production function
The marginal products, for K and L (x_1 and x_2)

$$f_i = \frac{\partial Q}{\partial x_i}$$

$$\Rightarrow MPL = \frac{\partial Q}{\partial L} = -\frac{F_l}{F_Q}; MPK = \frac{\partial Q}{\partial K} = -\frac{F_k}{F_Q}$$

Another partial derivative gives

$$\frac{\partial K}{\partial L} = -\frac{F_l}{F_k}$$

Since a partial derivative implies that other variables in the function are held constant – in this case output – then this is the change in the capital-output ratio that keeps output constant (moving along an isoquant)
= slope of the isoquant
Also called the marginal technical rate of substitution

Now generalise to a simultaneous equation case where there are n implicit functions - one for each endogenous variable in the system (and m exogenous variables)

$$F^1(y_1, y_2, ..., y_n; x_1, x_2, ..., x_m) = 0$$
$$F^2(y_1, y_2, ..., y_n; x_1, x_2, ..., x_m) = 0$$
$$\vdots$$
$$F^n(y_1, y_2, ..., y_n; x_1, x_2, ..., x_m) = 0$$

Now the implicit function rule says that we can solve for the n unknowns (and their partial derivatives), even if we do not have the n explicit functions, if the following matrix of partial derivatives has a non-zero determinant

$$|J| = \left| \frac{\partial(F^1 ...., F^n)}{\partial(y_1, ..., y_n)} \right| = \begin{vmatrix} \frac{\partial F^1}{\partial y_1} & \frac{\partial F^1}{\partial y_2} & .. & \frac{\partial F^1}{\partial y_n} \\ \frac{\partial F^2}{\partial y_1} & \frac{\partial F^2}{\partial y_2} & .. & \frac{\partial F^2}{\partial y_n} \\ \vdots & . & .. & \vdots \\ \frac{\partial F^n}{\partial y_1} & \frac{\partial F^n}{\partial y_2} & .. & \frac{\partial F^n}{\partial y_n} \end{vmatrix} \neq 0$$

J is called a ***Jacobian matrix*** (a matrix of partial derivatives)

As with the single equation, substitution of partial derivatives from the explicit into the implicit functions (taking the x partials to the other side) gives

$$\frac{\delta F^1}{\delta y_1}dy_1 + \frac{\delta F^1}{\delta y_2}dy_2 + ..\frac{\delta F^1}{\delta y_n}dy_n = -\left(\frac{\delta F^1}{\delta x_1}dx_1 + ..\frac{\delta F^1}{\delta x_m}dx_m\right)$$

$$\frac{\delta F^2}{\delta y_1}dy_1 + \frac{\delta F^2}{\delta y_2}dy_2 + ..\frac{\delta F^2}{\delta y_n}dy_n = -\left(\frac{\delta F^2}{\delta x_1}dx_1 + ..\frac{\delta F^2}{\delta x_m}dx_m\right)$$

$$\vdots$$

$$\frac{\delta F^n}{\delta y_1}dy_1 + \frac{\delta F^n}{\delta y_2}dy_2 + ..\frac{\delta F^n}{\delta y_n}dy_n = -\left(\frac{\delta F^n}{\delta x_1}dx_1 + ..\frac{\delta F^n}{\delta x_m}dx_m\right)$$

and replacing in the total derivative of the explicit equations as before

$$dy_1 = \frac{\delta y_1}{\delta x_1}dx_1 + \frac{\delta y_1}{\delta x_2}dx_2 + ..\frac{\delta y_1}{\delta x_m}dx_m$$

$$dy_2 = \frac{\delta y_2}{\delta x_1}dx_1 + \frac{\delta y_2}{\delta x_2}dx_2 + ..\frac{\delta y_2}{\delta x_m}dx_m$$

$$\vdots$$

$$dy_n = \frac{\delta y_n}{\delta x_1}dx_1 + \frac{\delta y_n}{\delta x_2}dx_2 + ..\frac{\delta y_n}{\delta x_m}dx_m$$

It gives, if x_1 is the only variable to change (so the other variables do not change, i.e. $dx_i = 0$ for $i > 1$):

$$\frac{\delta F^1}{\delta y_1}\left(\frac{\delta y_1}{\delta x_{x1}}\right)+\frac{\delta F^1}{\delta y_2}\left(\frac{\delta y_2}{\delta x_{x1}}\right)+..\frac{\delta F^1}{\delta y_n}\left(\frac{\delta y_n}{\delta x_{x1}}\right)=-\frac{\delta F^1}{\delta x_1}$$

$$\vdots$$

$$\frac{\delta F^n}{\delta y_1}\left(\frac{\delta y_1}{\delta x_{x1}}\right)+\frac{\delta F^n}{\delta y_2}\left(\frac{\delta y_2}{\delta x_{x1}}\right)+..\frac{\delta F^n}{\delta y_n}\left(\frac{\delta y_n}{\delta x_{x1}}\right)=-\frac{\delta F^n}{\delta x_1}$$

(similar systems hold for the other x's)

The system can now be re-written in the form Ax= b

$$\begin{bmatrix} \frac{\delta F^1}{\delta y_1} & \frac{\delta F^1}{\delta y_2} & .. & \frac{\delta F^1}{\delta y_n} \\ \frac{\delta F^2}{\delta y_1} & \frac{\delta F^2}{\delta y_2} & & \frac{\delta F^2}{\delta y_n} \\ \vdots & & & \vdots \\ \frac{\delta F^n}{\delta y_1} & & & \frac{\delta F^n}{\delta y_n} \end{bmatrix} \begin{bmatrix} \frac{\delta y_1}{\delta x_1} \\ \frac{\delta y_2}{\delta x_1} \\ \vdots \\ \frac{\delta y_n}{\delta x_1} \end{bmatrix} = \begin{bmatrix} -\frac{\delta F^1}{\delta x_1} \\ -\frac{\delta F^2}{\delta x_1} \\ \vdots \\ -\frac{\delta F^n}{\delta x_1} \end{bmatrix}$$

The 1st term of which is a Jacobian matrix of partial derivatives of the implicit function F

So suppose we wish to find the partial derivative of one exogenous variable x_1 on one endogenous variable y_j

$$\begin{bmatrix} \frac{\delta F^1}{\delta y_1} & \frac{\delta F^1}{\delta y_2} & .. & \frac{\delta F^1}{\delta y_n} \\ \frac{\delta F^2}{\delta y_1} & \frac{\delta F^2}{\delta y_2} & & \frac{\delta F^2}{\delta y_n} \\ \vdots & & & \vdots \\ \frac{\delta F^n}{\delta y_1} & & & \frac{\delta F^n}{\delta y_n} \end{bmatrix} \begin{bmatrix} \frac{\delta y_1}{\delta x_1} \\ \frac{\delta y_2}{\delta x_1} \\ \vdots \\ \frac{\delta y_n}{\delta x_1} \end{bmatrix} = \begin{bmatrix} -\frac{\delta F^1}{\delta x_1} \\ -\frac{\delta F^2}{\delta x_1} \\ \vdots \\ -\frac{\delta F^n}{\delta x_1} \end{bmatrix}$$

Given this, simply use Cramer's rule $\frac{\delta y_j}{\delta x_1} = \frac{|J_j|}{|J|}$

Generalizing, suppose we wish to find the partial derivative of one exogenous variable x_i on one endogenous variable y_j

$$\begin{bmatrix} \frac{\delta F^1}{\delta y_1} & \frac{\delta F^1}{\delta y_2} & \cdots & \frac{\delta F^1}{\delta y_n} \\ \frac{\delta F^2}{\delta y_1} & \frac{\delta F^2}{\delta y_2} & & \frac{\delta F^2}{\delta y_n} \\ \vdots & & & \vdots \\ \frac{\delta F^n}{\delta y_1} & & & \frac{\delta F^n}{\delta y_n} \end{bmatrix} \begin{bmatrix} \frac{\delta y_1}{\delta x_i} \\ \frac{\delta y_2}{\delta x_i} \\ \vdots \\ \frac{\delta y_n}{\delta x_i} \end{bmatrix} = \begin{bmatrix} -\frac{\delta F^1}{\delta x_i} \\ -\frac{\delta F^2}{\delta x_i} \\ \vdots \\ -\frac{\delta F^n}{\delta x_i} \end{bmatrix}$$

Then

$$\frac{\delta y_j}{\delta x_j} = \frac{|J_j|}{|J|}$$

Example: 3 equation national income model

Explicit	Implicit
$Y = C + I + G$	$F^1 = Y - C - I - G = 0$
$C = a + b(Y - T)$	$F^2 = C - a - b(Y - T) = 0$
$T = g + dY$	$F^3 = T - g - dY = 0$

Find $\delta Y / \delta G$

Endogenous variables $(Y, C, T) = (y_1, y_2, y_3)$

Exogenous variables $(I, G) = (x_1, x_2)$
(a, b, g, d are constant parameters)

Explicit

$Y = C + I + G$

$C = a + b(Y - T)$

$T = g + dY$

Implicit

$F^1 = Y - C - I - G = 0$

$F^2 = C - a - b(Y - T) = 0$

$F^3 = T - g - dY = 0$

$$\begin{bmatrix} \frac{\delta F^1}{\delta y_1} & \frac{\delta F^1}{\delta y_2} & \cdots & \frac{\delta F^1}{\delta y_n} \\ \frac{\delta F^2}{\delta y_1} & \frac{\delta F^2}{\delta y_2} & & \frac{\delta F^2}{\delta y_n} \\ \vdots & & & \vdots \\ \frac{\delta F^n}{\delta y_1} & & & \frac{\delta F^n}{\delta y_n} \end{bmatrix} \begin{bmatrix} \frac{\delta y_1}{\delta x_i} \\ \frac{\delta y_2}{\delta x_i} \\ \vdots \\ \frac{\delta y_n}{\delta x_i} \end{bmatrix} = \begin{bmatrix} -\frac{\delta F^1}{\delta x_i} \\ -\frac{\delta F^2}{\delta x_i} \\ \vdots \\ -\frac{\delta F^n}{\delta x_i} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \frac{\delta F^1}{\delta Y} & \frac{\delta F^1}{\delta C} & \frac{\delta F^1}{\delta T} \\ \frac{\delta F^2}{\delta Y} & \frac{\delta F^2}{\delta C} & \frac{\delta F^2}{\delta T} \\ \frac{\delta F^3}{\delta Y} & \frac{\delta F^3}{\delta C} & \frac{\delta F^3}{\delta T} \end{bmatrix} \begin{bmatrix} \frac{\delta Y}{\delta G} \\ \frac{\delta C}{\delta G} \\ \frac{\delta T}{\delta G} \end{bmatrix} = \begin{bmatrix} -\frac{\delta F^1}{\delta G} \\ -\frac{\delta F^2}{\delta G} \\ -\frac{\delta F^3}{\delta G} \end{bmatrix}$$

Explicit
(1): Y = C + I + G
(2): C = a + b(Y – T)
(3): T = g + dY

Implicit
Y – C – I – G = 0 = F^1
C – a – b(Y – T) = 0 = F^2
T – g – dY = 0 = F^3

$$\frac{\delta y_j}{\delta x_1} = \frac{|J_j|}{|J|} \Rightarrow \frac{\delta Y}{\delta G} = \frac{|J_1|}{|J|} = \frac{\begin{vmatrix} -\frac{\delta F^1}{\delta G} & \frac{\delta F^1}{\delta C} & \frac{\delta F^1}{\delta T} \\ -\frac{\delta F^2}{\delta G} & \frac{\delta F^2}{\delta C} & \frac{\delta F^2}{\delta T} \\ -\frac{\delta F^3}{\delta G} & \frac{\delta F^3}{\delta C} & \frac{\delta F^3}{\delta T} \end{vmatrix}}{\begin{vmatrix} \frac{\delta F^1}{\delta Y} & \frac{\delta F^1}{\delta C} & \frac{\delta F^1}{\delta T} \\ \frac{\delta F^2}{\delta Y} & \frac{\delta F^2}{\delta C} & \frac{\delta F^2}{\delta T} \\ \frac{\delta F^3}{\delta Y} & \frac{\delta F^3}{\delta C} & \frac{\delta F^3}{\delta T} \end{vmatrix}} = \frac{\begin{vmatrix} -1 & -1 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -1 & 0 \\ -b & 1 & b \\ -d & 0 & 1 \end{vmatrix}} = \frac{1}{1-b+bd}$$

Check that we get the same answer if we re[lace (3) into (2) and then the result into (1) and compute the partial derivative for G in explicit function

Introduction to Optimisation

Many economic functions of interest (eg utility functions, production functions, profit functions, cost functions) are non linear

The idea behind optimisation is to choose the point where a function reaches a maximum or minimum value

Decision-makers are assumed to be "rational" i.e.

1. each decision-maker is assumed to have a preference ordering over the outcomes to which her actions lead
2. Each decision makers chooses the action, among those feasible, that leads to the most preferred outcome (according to this ordering).

We usually make assumptions that guarantee that a decision-maker's preference ordering is represented by a *payoff function* (sometimes called *utility function*), so the decision-maker's problem is:

$$\max_a u(a) \text{ subject to } a \in S$$

$$\max_a u(a) \text{ subject to } a \in S$$

Where: *u* is the decision-maker's payoff function over her actions

S is the set of her feasible actions.

classical consumer: ***a*** is a consumption bundle, ***u*** is the consumer's utility function, and ***S*** is the set of bundles of goods the consumer can afford.

Firm: ***a*** is an input-output vector, ***u*(*a*)** is the profit the action *a* generates, and ***S*** is the set of all feasible input-output vectors

In economic theory we sometimes need to solve a *minimization* problem of the form

$$\min_a u(a) \text{ subject to } a \in S$$

- we assume, for example, that firms choose input bundles to minimize the cost of producing any given output;
- an analysis of the problem of minimizing the cost of achieving a certain payoff greatly facilitates the study of a payoff-maximizing consumer.

Optimization: definitions

The optimization problems we study take the form

$$\max_x f(x) \text{ subject to } x \in S$$

where:

- f is a function,
- x is an n-vector (which we can also write as $(x_1, ..., x_n)$),
- S is a set of n-vectors.

We call:

- f the **objective function**,
- x the **choice variable**, and
- S the **constraint set** or **opportunity set**.

Definition

The value x^* of the variable x solves the problem

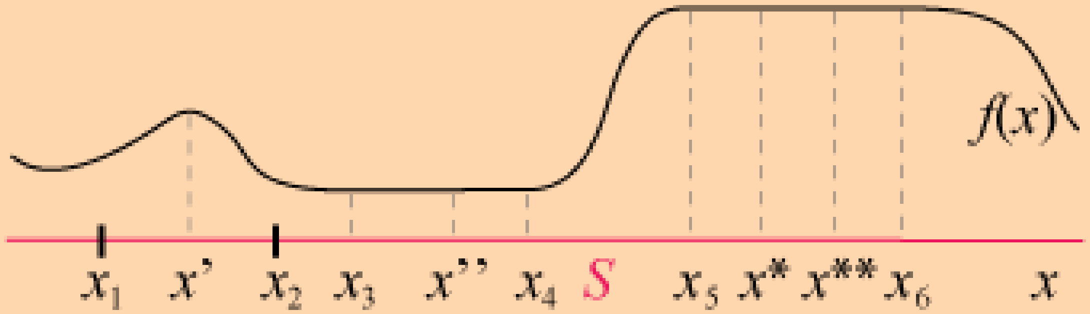
$$\max_x f(x) \text{ subject to } x \in S$$

if $f(x) \leq f(x^*)$ for all $x \in S$.

In this case we say that:

- x^* is a **maximizer** of the function f subject to the constraint $x \in S$
- that $f(x^*)$ is the **maximum** (or **maximum value**) of the function f subject to the constraint $x \in S$.

A **minimizer** is defined analogously



x* and ***x***** are maximizers of f subject to the constraint $x \in S$
x'' is a minimizer

What is ***x'*** ?

It is not a maximizer, because $f(x^*) > f(x')$,
It is not a minimizer, because $f(x'') < f(x')$

But it is a maximum *among the points close to it*. We call such a point a local maximizer

Definition

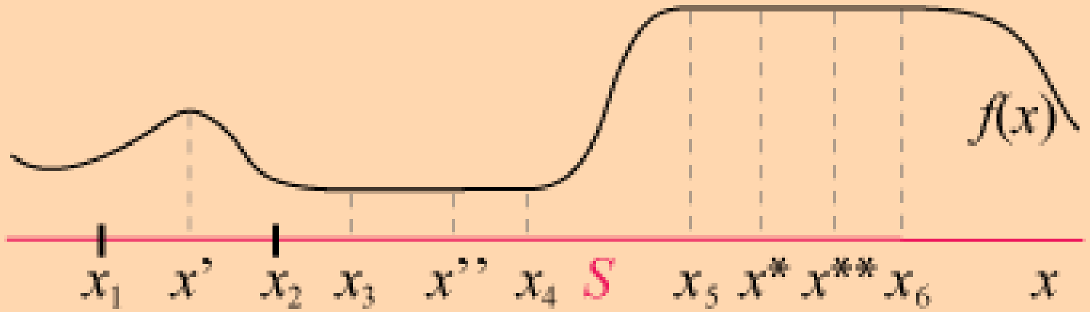
The variable x^* is a local maximizer of the function f subject to the constraint $x \in S$ if there is a number $\varepsilon > 0$ such that $f(x) \leq f(x^*)$ for all $x \in S$ for which the distance between x and x^* is at most ε.

Note: suppose that x and x' are vectors, then the *distance* between two points x and x' is the square root of $\sum_{i=1}^{n}(x_i - x_i')^2$

A **local minimizer** is defined analogously.

Sometimes we refer to a maximizer as a **global maximizer** to emphasize that it is not only a local maximizer.

Every global maximizer is, in particular, a local maximizer (ε can take any value), and every minimizer is a local minimizer.



$f\,(x) \le f\,(x')$ for all x between x_1 and x_2 (for example), where $|x_1 - x'| = |x_2 - x'|$ the point x' is a local maximizer of f (set $\varepsilon = x_1 - x'$).

But note that the point x'' is also a local maximizer of f, even though it is a global *minimizer*.

The function is constant between x_3 and x_4. The point x_4 is closer to x'' than is the point x_3, so we can take the ε in the definition of a local maximizer to be $x_4 - x''$. For every point x within the distance ε of x'', we have $f\,(x) = f\,(x'')$, so that in particular $f\,(x) \le f\,(x'')$.

Transforming the objective function

Let g be a strictly increasing function of a single variable.
i.e. if $z' > z$ then $g(z') > g(z)$

Then the set of solutions to the problem

$$\max_x f(x) \text{ subject to } x \in S$$

is identical to the set of solutions to the problem

$$\max_x g(f(x)) \text{ subject to } x \in S.$$

Proof:
If x^* is a solution to the first problem then by definition $f(x) \leq f(x^*)$ for all $x \in S$.
But if $f(x) \leq f(x^*)$ then $g(f(x)) \leq g(f(x^*))$, so that $g(f(x) \leq g(f(x^*))$ for all $x \in S$.
Hence x^* is a solution of the second problem.

Minimization problems

We concentrate on maximization problems. What about minimization problems?

Any minimization problem can be turned into a maximization problem by taking the negative of the objective function.

That is, the problem

$\min_x f(x)$ subject to $x \in S$

is equivalent (i.e. has the same set of solutions) to

$\max_x -f(x)$ subject to $x \in S$.

Thus we can solve any minimization problem by taking the negative of the objective function and apply the results for maximization problems.

Existence of an optimum

Let f be a function of n variables defined on the set S. The problems we consider take the form

$$\max_x f(x) \text{ subject to } x \in S \qquad \text{where } x = (x_1, ..., x_n).$$

Before we start to think about how to find the solution to a problem, we need to think about **whether the problem *has* a solution**

Some problems that do **not** have any solution.

1. $f(x) = x$, $S = [0, \infty)$

   In this case, f increases without bound, and never attains a maximum.

2. $f(x) = 1 - 1/x$, $S = [1, \infty)$.

   In this case, f converges to the value 1, but never attains this value.

3. $f(x) = x$, $S = (0, 1)$.

   In this case, the points 0 and 1 are excluded from S.

   As x approaches 1, the value of the function approaches 1, but this value I s never attained for values of x in S, because S excludes $x = 1$.

4. $f(x) = x$ if $x < 1/2$ and $f(x) = x - 1$ if $x \geq 1/2$; $S = [0, 1]$.

   In this case, as x approaches 1/2 the value of the function approaches 1/2, but this value is never attained, because at $x = 1/2$ the function jumps down to $-1/2$.

in the first two cases are that the set S is unbounded;

in the third case is that the interval S is open (does not contain its endpoints);

in the last case is that the function f is discontinuous.

Definition of Bounded set

For functions of many variables, we need to define the concept of a *bounded* set.

The set S is **bounded** if there exists a number k such that the distance of every point in S from the origin is at most k.

Example

The set $[-1, 100]$ is bounded, because the distance of any point in the set from 0 is at most 100. The set $[0, \infty)$ is not bounded, because for any number k, the number $2k$ is in the set, and the distance of $2k$ to 0 is $2k$ which exceeds k.

Example

The set $\{(x, y): x^2 + y^2 \leq 4\}$ is bounded, because the distance of any point in the set from (0, 0) is at most 2.

Example

The set $\{(x, y): xy \leq 1\}$ is not bounded, because for any number k the point $(2k, 0)$ is in the set, and the distance of this point from (0, 0) is $2k$, which exceeds k.

We say that a set that is closed and bounded is **compact.**

Proposition (Extreme value theorem) :

A continuous function on a compact set attains both a maximum and a minimum on the set

Note that the requirement of boundedness is on the *set*, not the *function*.

Note also that the result gives only a **sufficient** condition for a function to have a maximum.

If a function is continuous and is defined on a compact set **then** it definitely has a maximum and a minimum.

The result does **not** rule out the possibility that a function has a maximum and/or minimum if it is not continuous or is not defined on a compact set.

For each of the following functions, determine

(i) whether the extreme value theorem implies that the function has a maximum and a minimum and

(ii) (*ii*) if the extreme value theorem does not apply, whether the function does in fact have a maximum and/or a minimum.

- x^2 on the interval $[-1,1]$
- x^2 on the interval $(-1,1)$
- $|x|$ on the interval $[-1,\infty)$
- $f(x)$ defined by $f(x) = 1$ if $x < 0$ and $f(x) = x$ if $x \geq 0$, on the interval $[-1,1]$.
- $f(x)$ defined by $f(x) = 1$ if $x < 0$ and $f(x) = x$ if $x \geq 0$, on the interval $(-\infty,\infty)$.
- $f(x)$ defined by $f(x) = x^2$ if $x < 0$ and $f(x) = x$ if $x \geq 0$ on the interval $[-1,1]$.

UNCONSTRAINED OPTIMIZATION WITH ONE VARIABLE

Necessary conditions

Consider the following problem where $f(x)$ is a differentiable function defined on $\mathbb{R}$

$$\max_x f(x) \text{ subject to } x \in \mathbb{R}$$

(Remember: $\mathbb{R}$ is the set of real numbers, then x is a single variable)

A point x such that $f'(x) = 0$ is called *stationary point*

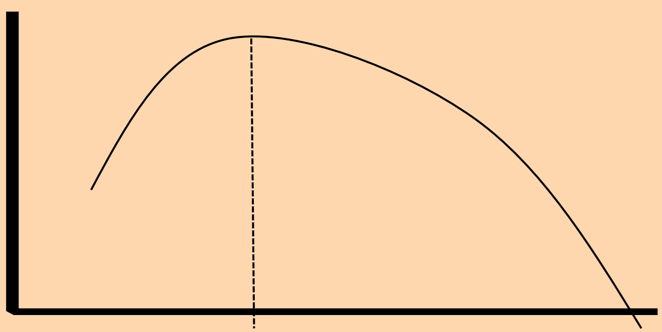
Consider the following figures

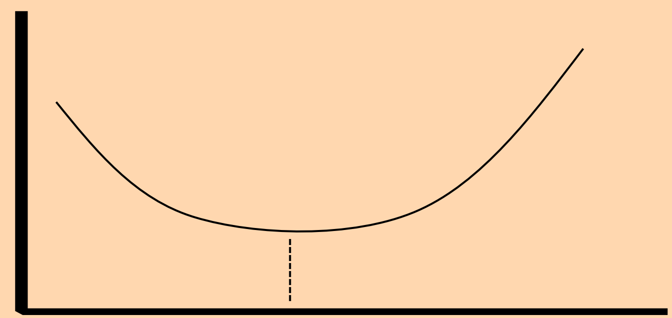
The stationary point is unique and a maximum in the first case.

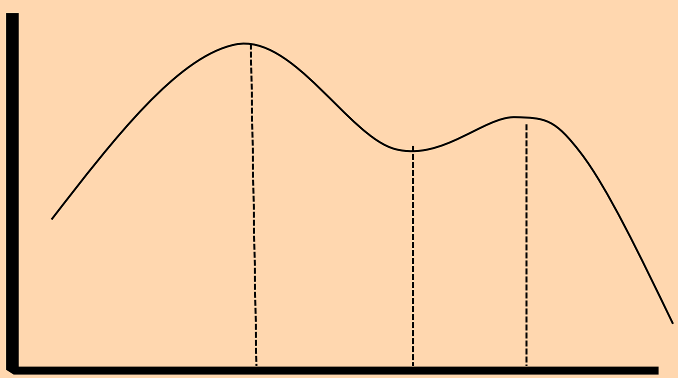
In the second case he stationary point is unique and is a minimum;

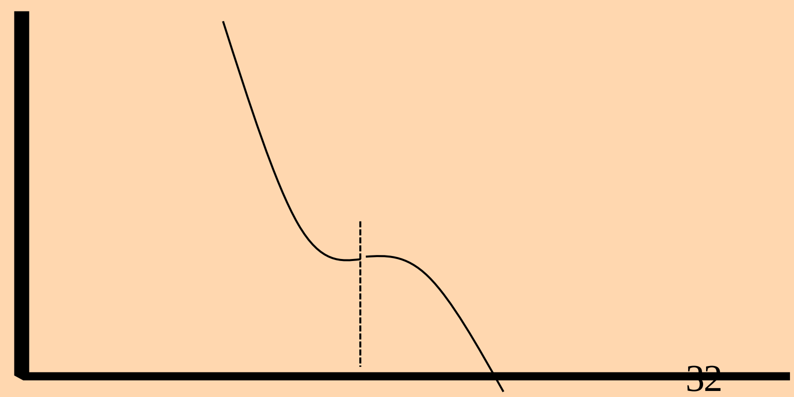
in the third case there are three stationary points is no unique solution to the first order conditions.

In the fourth case the stationary point is unique and is not a maximum, is not a minimum.









From the previous figures we see that:

a stationary point is not necessarily a global or local maximizer, or a global or local minimizer)

a global or local maximizer and a global or local minimizer is necessarily a stationary point

Proposition:

Let f be a differentiable function of a single variable defined on the set of real numbers.

If a point x is a local or global maximizer or minimizer of f then $f\,'(x) = 0$.

It is a *necessary* condition for *x* to be a maximizer (or a minimizer) of *f* : **if** it is a maximizer (or a minimizer) **then** *x* is stationary point of *f* .

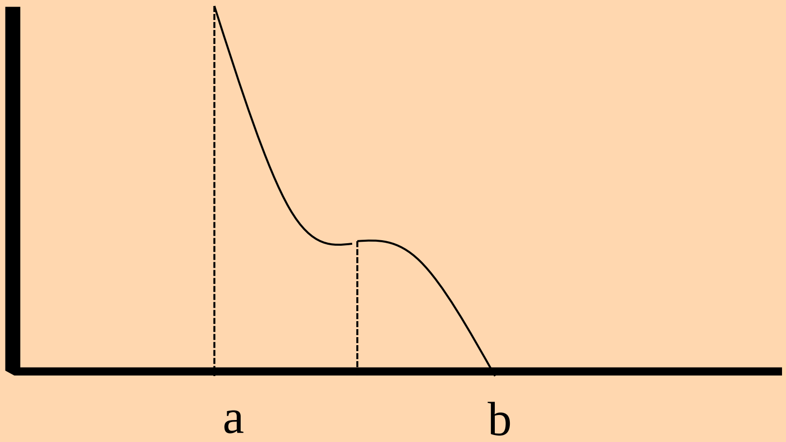
It is **not** *sufficient* for a point to be a maximizer—the condition is satisfied also, for example, at points that are minimizers.

We refer to this condition as a **first-order condition**

Disgression: interior optimum

if we consider the problem: $\max_x f(x)$ subject to $x \in I$
where I is an interval of real numbers, i.e. $I = [a, b]$

In this case a maximum is not necessarily a stationary point



In this case $f'(x) = 0$ is a necessary condition for maximizers and minimizers that are in the interior of I (it means that they are not on the boundaries of I)

If $I = \mathbb{R}$ we are in the previous case, all points are interiors because there is not boundaries.

Consider the problem

$\max_x x^2$ subject to $x \in [-1, 2]$.

This problem satisfies the conditions of the extreme value theorem, and hence has a solution.

Let $f(x) = x^2$. We have $f'(x) = 2x$, so the function has a single stationary point, $x = 0$, which is in the constraint set.

The value of the function at this point is $f(0) = 0$.

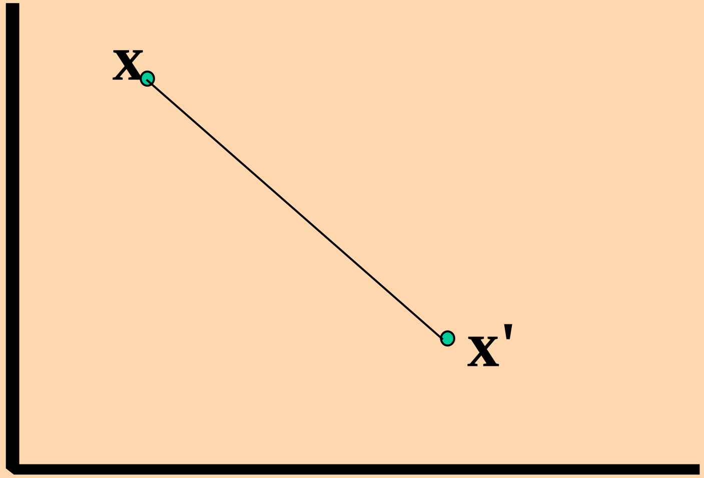
The values of f at the endpoints of the interval on which it is defined are $f(-1) = 1$ and $f(2) = 4$.

Thus the global maximizer of the function on $[-1, 2]$ is $x = 2$ and the global minimizer is $x = 0$.

Now we need to study the conditions that allow us to say if a stationary point is an optimizer or not.

Preliminaries: Convex combinations

A convex combination of two points is a point that lies on the line between them.



More formally, consider two points x and x', then define

1. A ***convex combination*** $x'' = \lambda x + (1-\lambda)x'$ for $\mathbf{0 \le \lambda \le 1}$

   i. So if $\lambda=0$ then $x'' = x'$; if $\lambda=1$ then $x'' = x$

$$e.g.\ x = \begin{pmatrix} 1 \\ 4 \end{pmatrix}; \quad x' = \begin{pmatrix} 5 \\ 0 \end{pmatrix}; \ x'' = \lambda x + (1-\lambda)x' = \begin{pmatrix} \lambda + (1-\lambda)5 \\ 4\lambda + (1-\lambda)0 \end{pmatrix}$$

2. A ***strictly convex combination:*** $x'' = \lambda x + (1-\lambda)x'$ for $\mathbf{0 < \lambda < 1}$

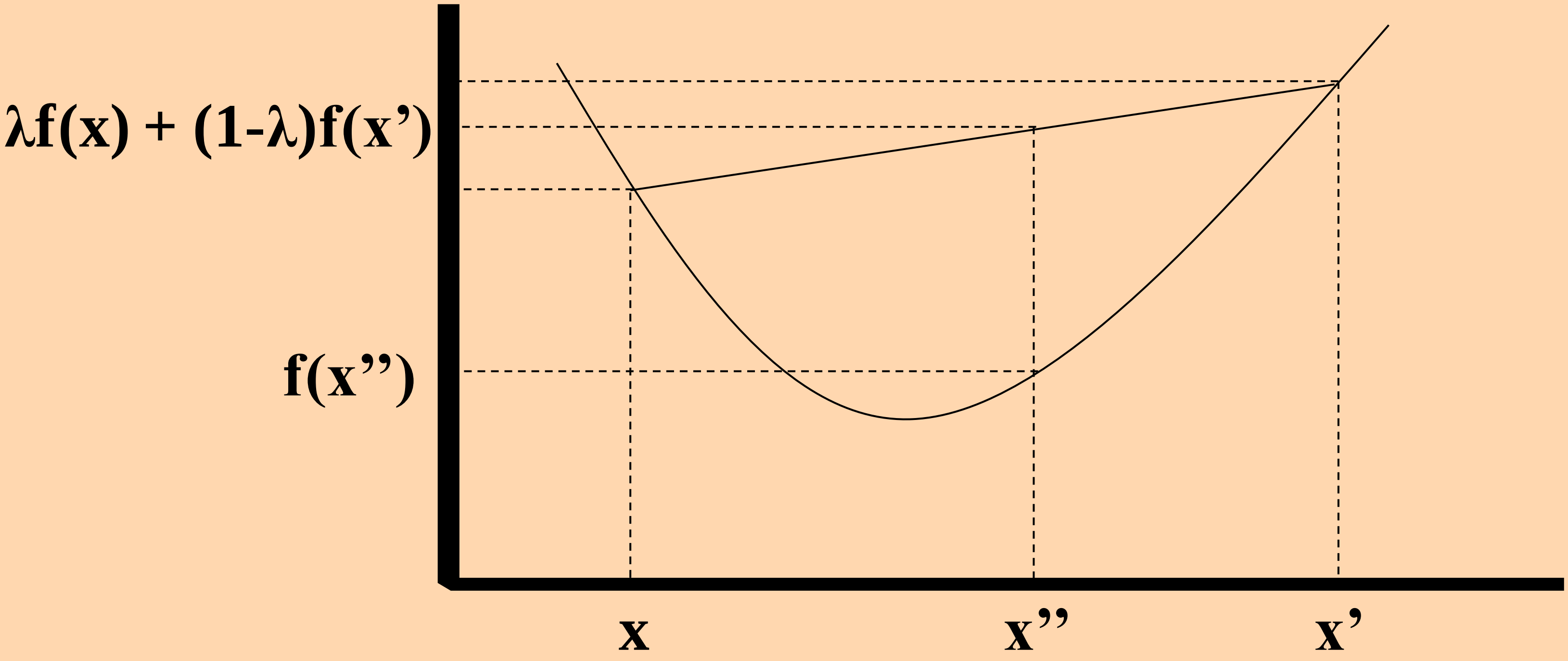
Convex functions

Convex ***functions*** are defined as follows

$f(x)$ is convex if given any x, x' and $x'' = \lambda x + (1-\lambda)x'$ where $0 \leq \lambda \leq 1$,

$$f(x'') = f(\lambda x + (1-\lambda)x') \leq \lambda f(x) + (1-\lambda)f(x')$$

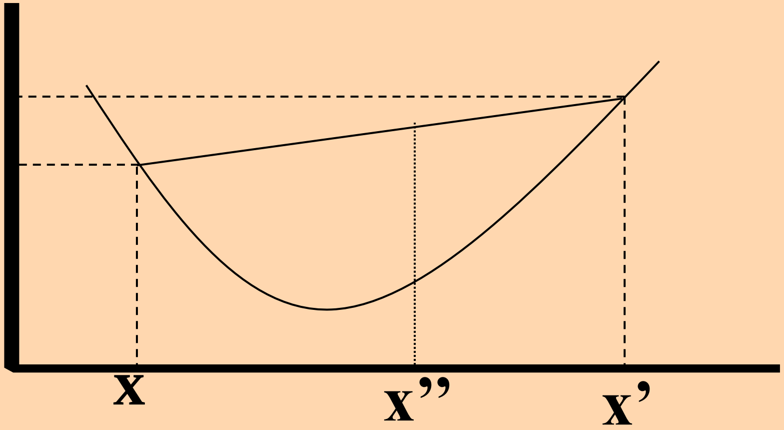
- the line joining x and x' lies above the function, f at any point between x and x'

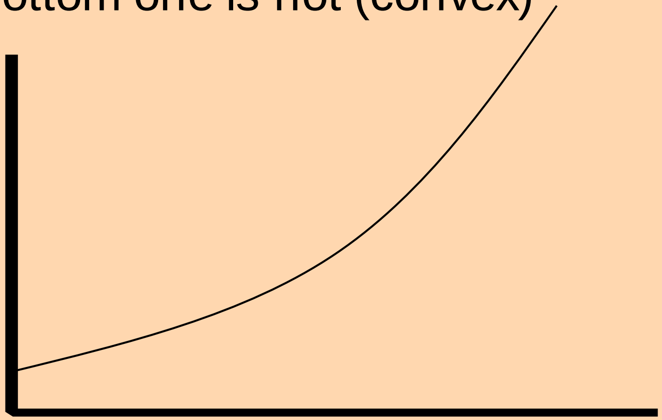


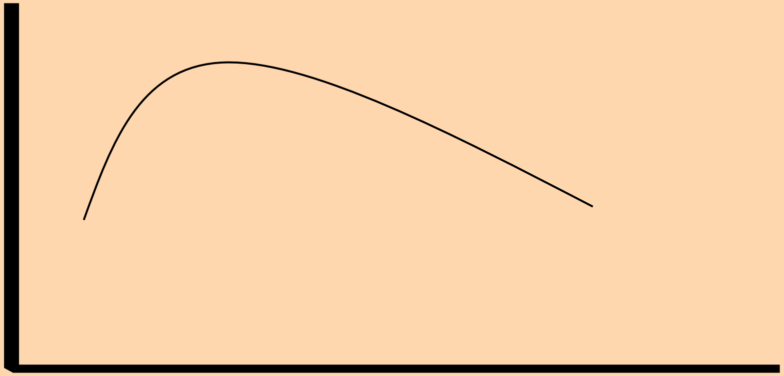
f(x) is convex if given any x, x', $x'' = \lambda x + (1-\lambda)x'$ where $0 \le \lambda \le 1$,

$$f(x'') = f(\lambda x + (1-\lambda)x') \le \lambda f(x) + (1-\lambda)f(x')$$

So the top 2 functions are convex but the bottom one is not (convex)





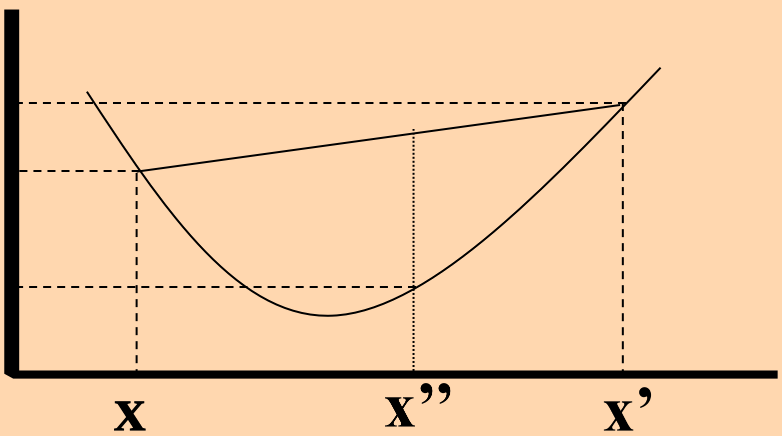


Strictly convex functions

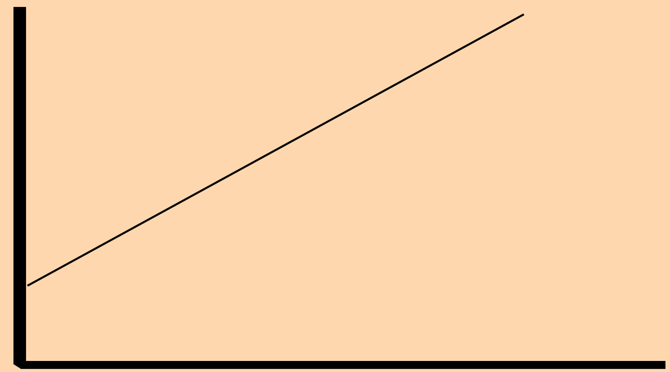
f(x) is ***strictly convex*** if given any x, x’, x″ = λx + (1-λ)x’ where 0 < λ < 1,

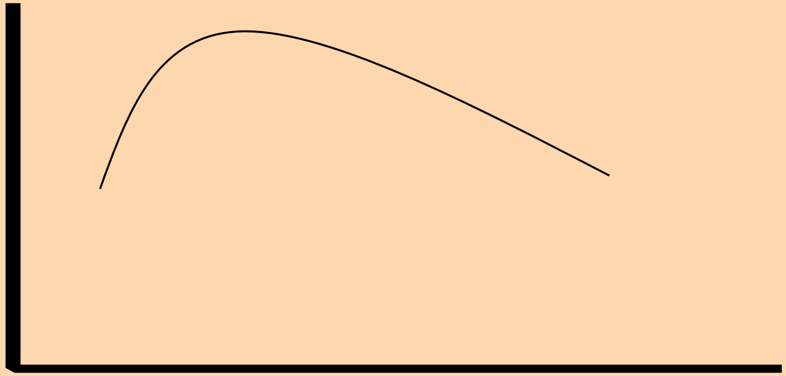
$$f(x'') < \lambda f(x) + (1-\lambda) f(x')$$

strictly convex and convex



convex, but not strictly convex





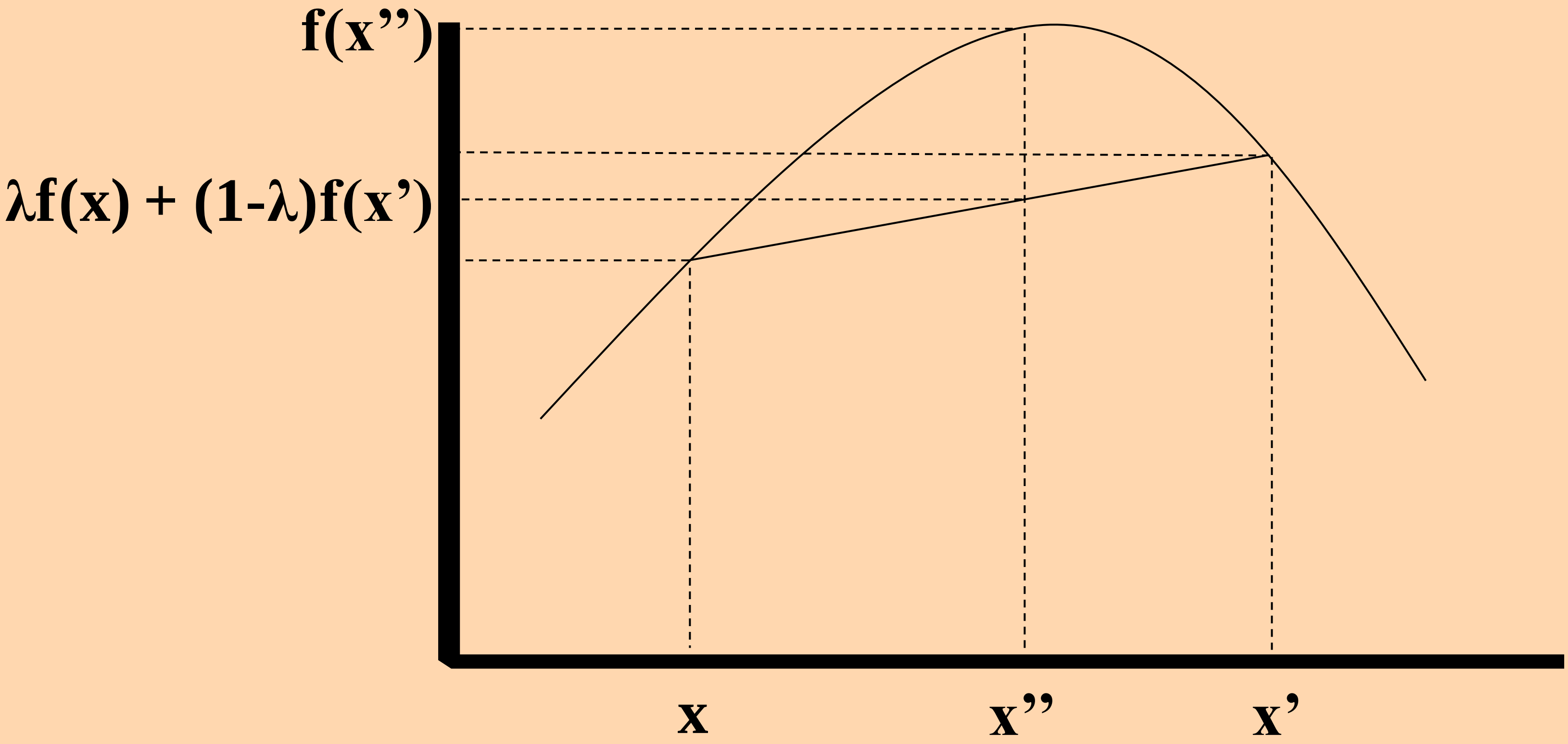
not convex or strictly convex

Concave functions

f(x) is ***concave*** if given any x, x', $x'' = \lambda x + (1-\lambda)x'$ where $0 \le \lambda \le 1$,

$$f(x'') = f(\lambda x + (1-\lambda)x') \ge \lambda f(x) + (1-\lambda)f(x')$$

*the line joining x and x' lies **below** the function, f* *(conCave)*

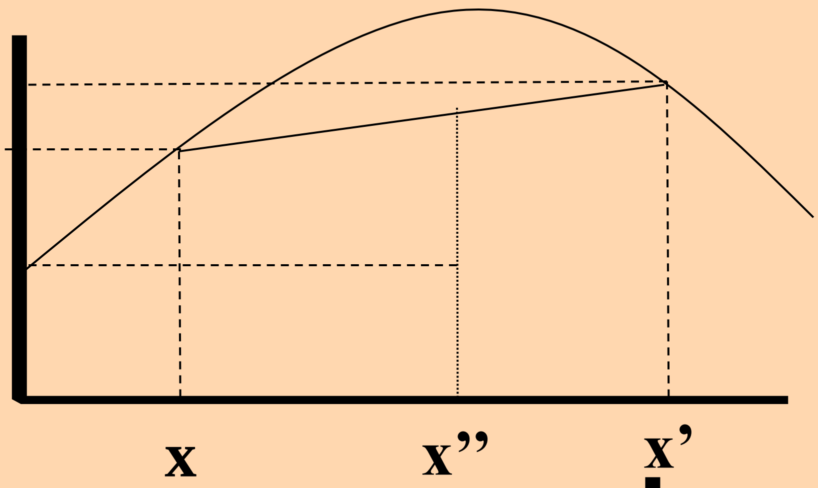


Strictly concave functions

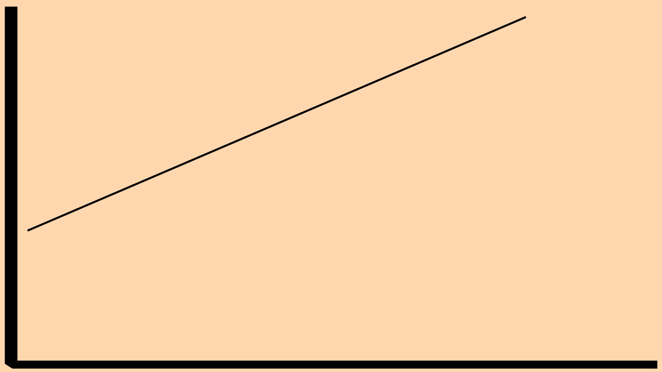
f(x) is ***strictly concave*** if given any x, x’, x″ = λx + (1-λ)x’ and 0 < λ < 1,

$$f(x'') = f(\lambda x + (1-\lambda)x') > \lambda f(x) + (1-\lambda)f(x')$$

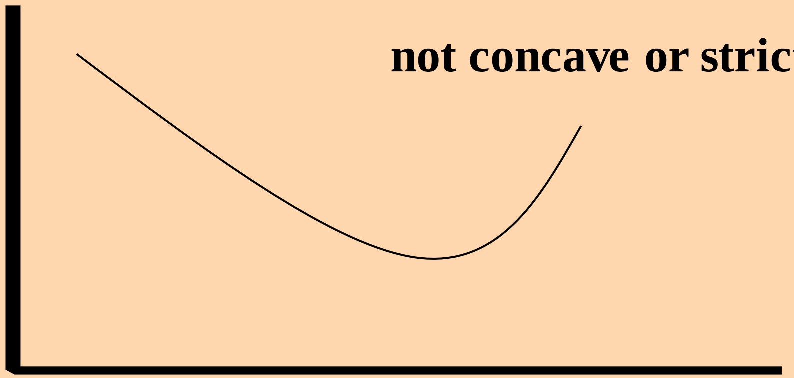
strictly concave and concave



concave, but not strictly concave



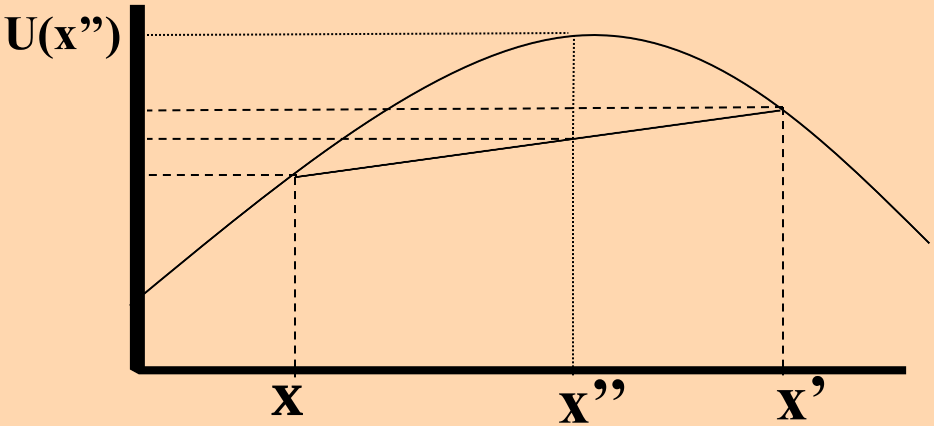
not concave or strictly concave



Example. Think of the function below as representing utility gained when faced with income combinations. Suppose utility function is U(x) and is concave.

Consider a "fair bet" – equal probability of each outcome

Expected utility = 0.5 * x + 0.5*x' = x'' (say)



"Risk averse" would get higher utility from having x'' with certainty than the fair bet

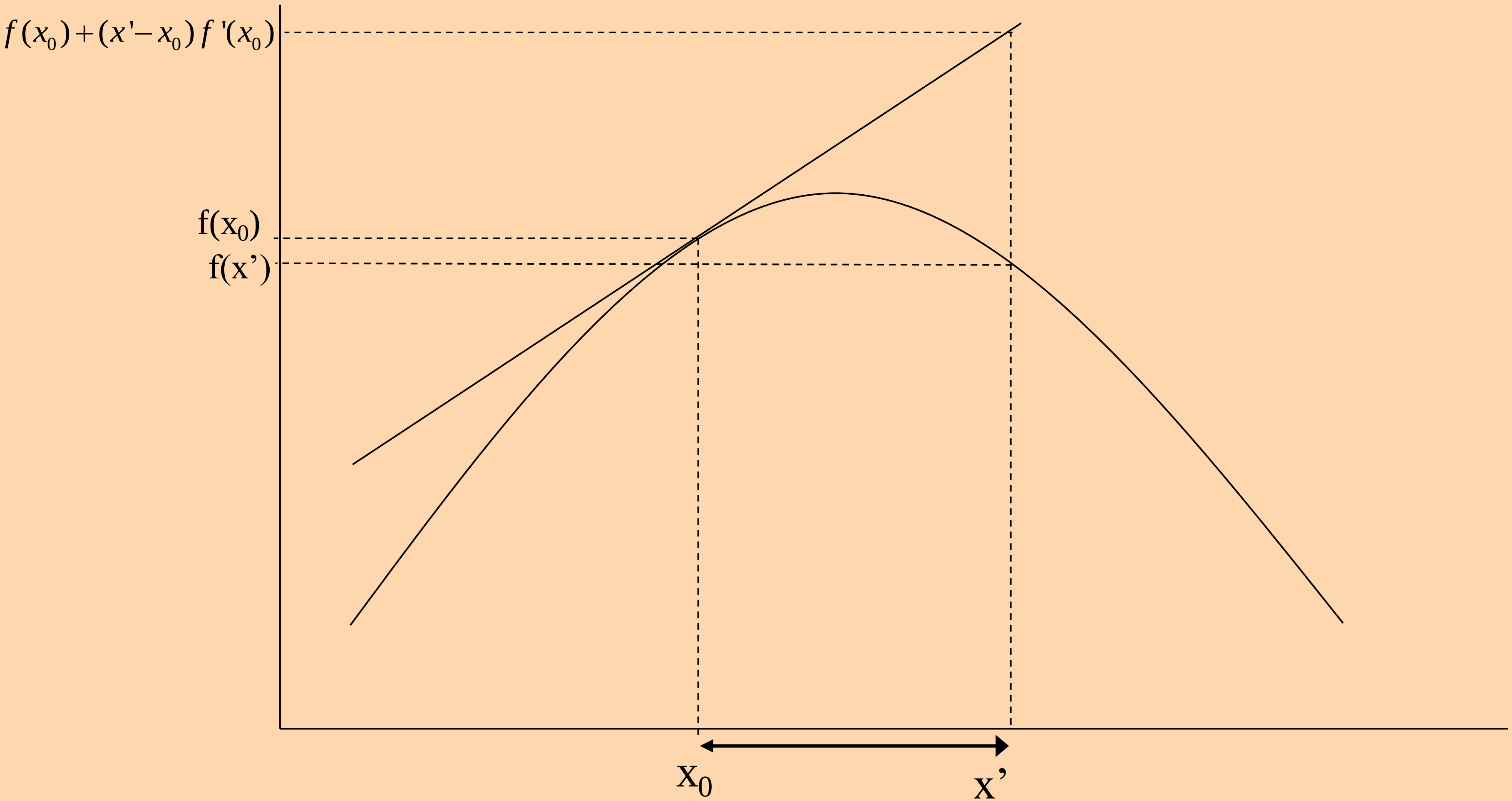
Quiz: prove that risk lovers are characterized by a strictly convex utility function

Finding if a function is concave

When f is differentiable, then we can check using an implication of concavity:

Given a straight line tangent to the function at x_0 it must be that for a concave function

$$f(x_0)+(x'-x_0)f'(x_0)\geq f(x')$$



First recall Taylor's theorem

$$f(x') \approx f(x_0) + (x' - x_0) f'(x_0) + \frac{1}{2}(x' - x_0)^2 f''(x_0) + \cdots$$

We now know that for a concave function:

$$f(x_0) + (x' - x_0) f'(x_0) \geq f(x')$$

So

$$f(x_0) + (x' - x_0) f'(x_0) \geq f(x_0) + (x' - x_0) f'(x_0) + \frac{1}{2}(x' - x_0)^2 f''(x_0) + \cdots$$

Cancelling terms and we get

$$f(x_0) \geq f(x_0) + \frac{1}{2}(x' - x_0)^2 f''(x_0) + \cdots$$

For inequality to hold 2nd term on right hand side must be non-positive

Since $(x\text{-}x_0)^2$ must be nonnegative, then it must mean that

$$0 \geq \frac{\partial^2 f}{\partial x^2}$$

i.e. for concavity to hold, the second derivative is **not** positive

(the slope of the function is getting smaller)

Strict concavity means that the second derivative must be negative

So back to $y = x^2$

concavity means $\frac{d^2 f}{dx^2} \leq 0$

(slope getting smaller)

convexity means $\frac{d^2 f}{dx^2} \geq 0$

(slope increasing)

So $\frac{dy}{dx} = 2x$

$$\frac{d^2 y}{dx^2} = 2 > 0$$

and function must be convex (check by sketching)

More properties

1. If f and g are both concave functions then f+g is also concave

Proof

f is concave so $f(x'') \geq \lambda f(x) + (1-\lambda)f(x')$

g is concave so $g(x'') \geq \lambda g(x) + (1-\lambda)g(x')$

So $f(x'')+g(x'') \geq \lambda f(x) + (1-\lambda)f(x') + \lambda g(x) + (1-\lambda)g(x')$

Or

$f(x'')+g(x'') \geq \lambda[f(x) + g(x)]+ (1-\lambda)[f(x') + g(x')]$

Similarly

2. If f and g are both convex functions then f+g is also convex
3. If f is concave then –f is convex
4. If f is convex then –f is concave
5. If f is concave and g is convex then f-g is concave
6. Linear functions are both concave and convex

Conditions under which a stationary point is a local optimum

Let f be a function of a single variable with continuous first and second derivatives, defined on the interval I. Suppose that x^* is a stationary point of f in the interior of I (so that $f'(x^*) = 0$).

- If $f''(x^*) < 0$ then x^* is a local maximizer.
- If x^* is a local maximizer then $f''(x^*) \leq 0$.
- If $f''(x^*) > 0$ then x^* is a local minimizer.
- If x^* is a local minimizer then $f''(x^*) \geq 0$.

Consider $f(x) = \frac{1}{3}x^3 - x$ defined on $\mathbb{R}$.

Stationary points are given by the solutions of $x^2 - 1 = 0$, i.e. $x = 1$and $x = -1$.

$$f''(x) = 2x$$

Consider the solution $x = 1$, $f''(1) = 2$, then $x = 1$ is a local minimizer

Consider the solution $x = -1$, $f''(-1) = -2$, then $x = 1$ is a local maximizer

Conditions under which a stationary point is a global optimum

Let f be a differentiable function defined on the interval I, and let x be in the interior of I. Then

- if f is concave then x is a global maximizer of f in I if and only if x is a stationary point of f
- if f is convex then x is a global minimizer of f in I if and only if x is a stationary point of f .

Note: a twice-differentiable function is concave if and only if its second derivative is nonpositive (and similarly for a convex function),

Then if f is a twice-differentiable function defined on the interval I and x is in the interior of I then:

- $f''(z) \leq 0$ for all $z \in I \Rightarrow$ [x is a global maximizer of f in I if and only if $f'(x) = 0$]
- $f''(z) \geq 0$ for all $z \in I \Rightarrow$ [x is a global minimizer of f in I if and only if $f'(x) = 0$].

Example 1: What output maximizes a monopolist's profit?

What quantity maximises consumer utility?

Suppose $f(x) = -x^2 + x - 10$

Three steps:

1. First order condition

2. Solve equation

3. Check that it is a maximum (second order condition < 0)

   1. $f_x = -2x + 1 = 0$

   2. Solving: $1 = 2x$ so $x = 0.5$

   3. $f_{xx} = -2 < 0$ so $x = 0.5$ gives a maximum

Example 2 Profit Maximisation

$$\pi(q) = R(q) - C(q)$$

1st Order Condition for a maximum

$$\frac{d\pi}{dq} = \pi'(q) = 0 = R'(q) - C'(q)$$

Or in words marginal revenue = marginal cost

2nd Order condition for a maximum

$$\frac{d^2\pi}{dq^2} = \frac{d\pi'(q)}{dq} < 0 = R''(q) - C''(q)$$

-which means that the rate of change (slope) of marginal revenue curve is less than the rate of change (slope) of the marginal cost curve at the point where MR=MC

-ie MR curve cuts the MC curve from above