Lecture 2 – Optimization with equality constraints

Constrained optimization

The idea of constrained optimisation is that the choice of one variable often affects the amount of another variable that can be used:

- if a firm employs more labour, this may affect the amount of capital it can rent if it is restricted (constrained) by how much it can spend on inputs
- when a consumer maximizes utility, income provides the constraint.
- when a government sets expenditure levels, it faces constraints set by its income from taxes

Note that the optimal quantities obtained under a constraint may be different from the quantities obtained without constraint

Binding and non-binding constraints

In the solution we say that a constraint is ***binding*** if the constraint function holds with equality (sometimes called an *equality constraint*)

Otherwise the constraint is non-binding or ***slack*** (sometimes called an *inequality constraint*)

When the constraint is binding we can use the ***Lagrangean*** technique

In general cases we do not know whether a constraint will be binding.

Sometime we can use our economic understanding to tell us if a constraint is binding

– Example: a non-satiated consumer will always spend all her income, so the budget constraint will be satisfied with equality

When we are not able to say if constraints are binding we use a technique which is related to the Lagrangean, but which is slightly more general (***Kuhn-Tucker programming***)

Objectives and constraints - example

A firm chooses output x to maximize a profit function

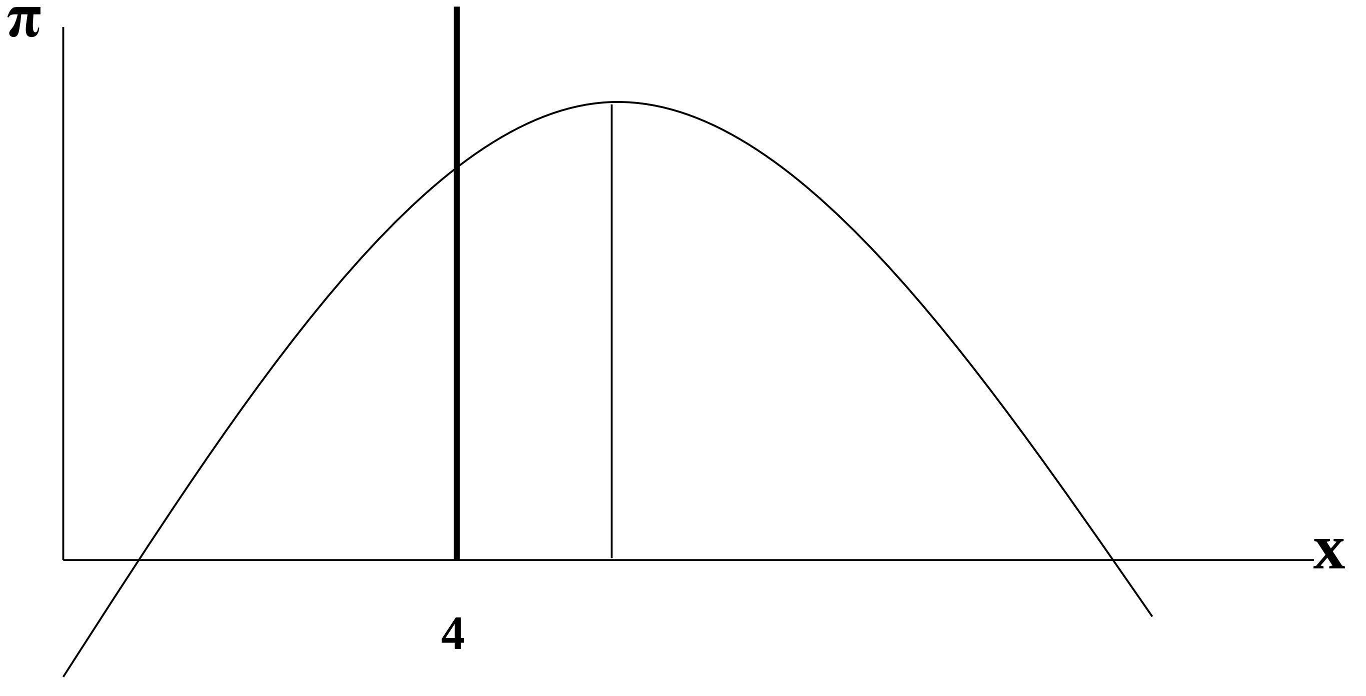
$\pi = -x^2 + 10x - 6$

Because of a staff shortage, it cannot produce an output higher than x = 4

What are the objective and constraint functions?

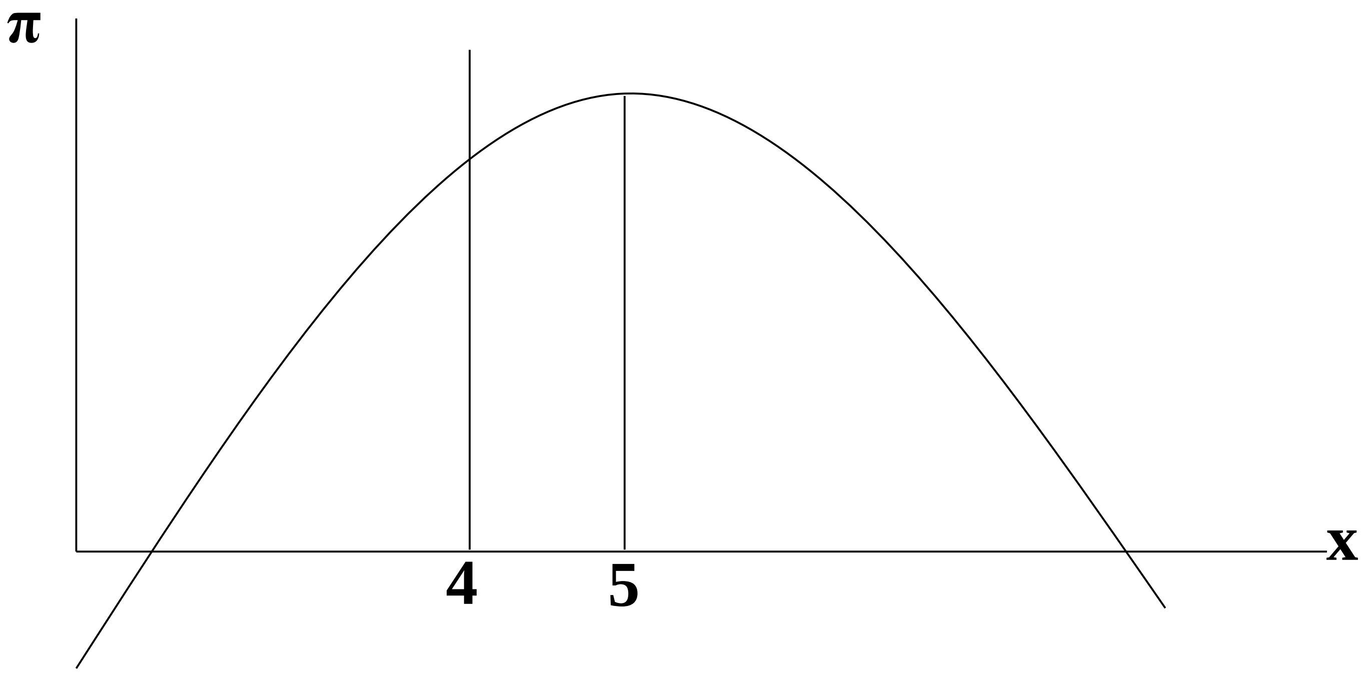
The objective function: $\pi = -x^2 + 10x-6$

The constraint: $x \leq 4$



Note that without the constraint the optimum is x = 5

So the constraint is binding (but a constraint of, say, x ≤ 6 would not be)



Sometime in the following (and in the textbook) we denote:

$$\frac{df(x,y)}{dx} = f_1(x,y)$$

$$\frac{df(x,y)}{dy} = f_2(x,y)$$

more in general

$$\frac{df\left(x_1, x_2, \dots x_j \dots x_n\right)}{dx_j} = f_j\left(x_1, x_2, \dots x_j \dots x_n\right)$$

Constrained optimization with two variables and one constraint

The problem is:

$$\max_{\{x,y\}} f(x,y)$$

$$s.t\ g(x,y) = c \qquad x,y \in S$$

To get the solution we have to write the *Lagrangean:*

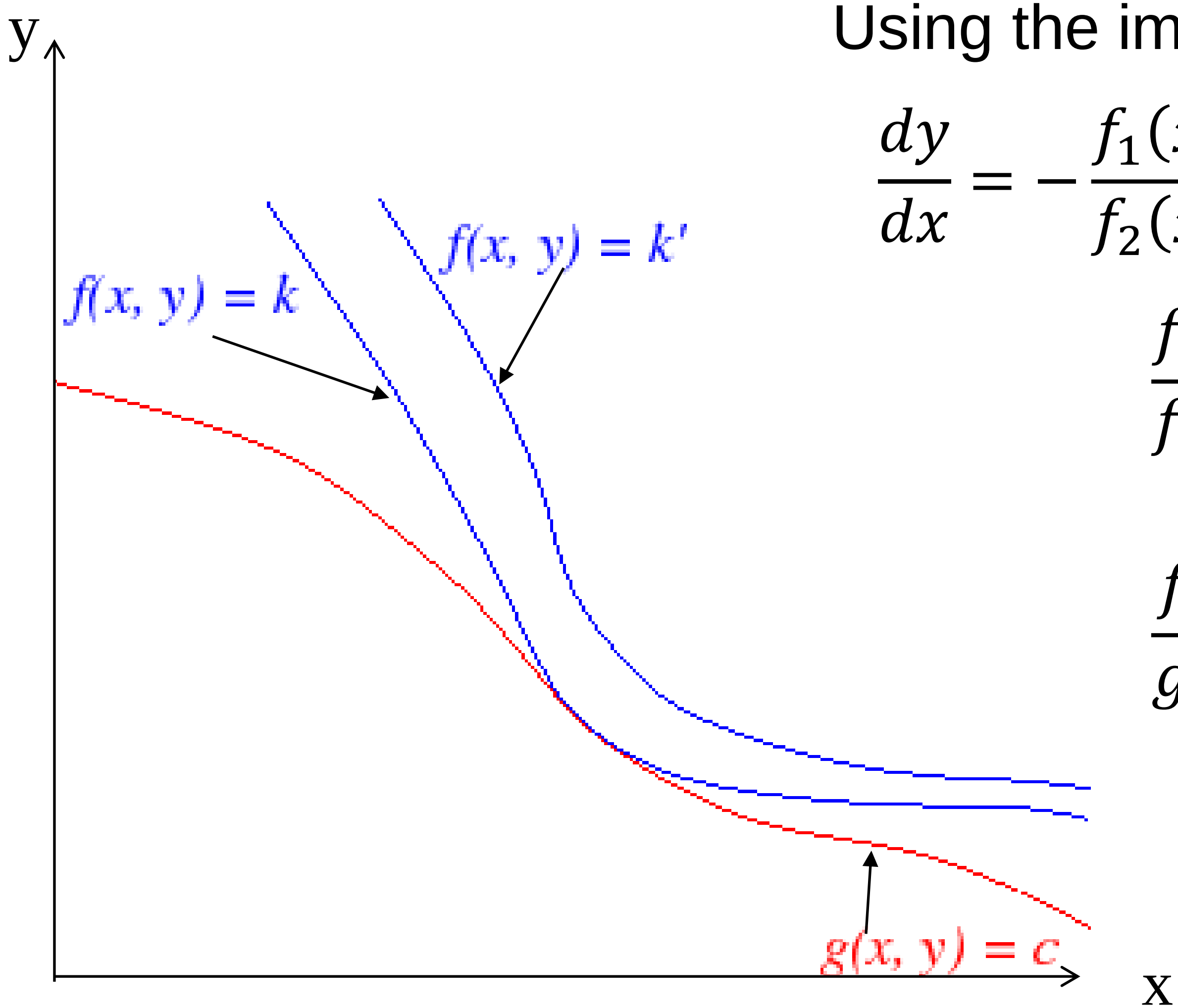
$$L(x,y,\lambda) = f(x,y) - \lambda(g(x,y) - c)$$

where λ is a new variable

The candidates to the solution are the stationary points of the Lagrangean, i.e. all points that satisfy the following system of equations:

$$\begin{cases} f_1(x,y) - \lambda g_1(x,y) = 0 \\ f_2(x,y) - \lambda g_2(x,y) = 0 \\ g(x,y) = c \end{cases}$$

Intuition about Lagrangean



Using the implicit function theorem

$$\frac{dy}{dx} = -\frac{f_1(x^*, y^*)}{f_2(x^*, y^*)} = -\frac{g_1(x^*, y^*)}{g_2(x^*, y^*)}$$

$$\frac{f_1(x^*, y^*)}{f_2(x^*, y^*)} = \frac{g_1(x^*, y^*)}{g_2(x^*, y^*)}$$

$$\frac{f_1(x^*, y^*)}{g_1(x^*, y^*)} = \frac{f_2(x^*, y^*)}{g_2(x^*, y^*)}$$

$$\frac{f_1(x^*,y^*)}{g_1(x^*,y^*)} = \frac{f_2(x^*,y^*)}{g_2(x^*,y^*)} = \lambda$$

Using $\frac{f_2(x^*,y^*)}{g_2(x^*,y^*)} = \lambda$ we get $f_1(x^*,y^*) - \lambda g_1(x^*,y^*) = 0$

Using $\frac{f_1(x^*,y^*)}{g_1(x^*,y^*)} = \lambda$ we get $f_2(x^*,y^*) - \lambda g_2(x^*,y^*) = 0$

Moreover the solution has to satisfy the constraint $g(x^*,y^*) = c$

Then the solution has to satisfy the following three equations:

$$f_1(x^*,y^*) - \lambda g_1(x^*,y^*) = 0$$

$$f_2(x^*,y^*) - \lambda g_2(x^*,y^*) = 0$$

$$g(x^*,y^*) = c$$

These equations are the derivatives of the **Lagrangean**

$$L(x,y,\lambda) = f(x,y) - \lambda(g(x,y) - c)$$

with respect to x, y and λ to be zero

The first two equations are know as the **first order conditions**

Proposition (necessary conditions)

Let f and g be continuously differentiable functions of two variables defined on the set *S*, c be a number.

Suppose that:

- (x^*, y^*) is an interior point of *S* that solves the problem

$$\max_{\{x,y\}} f(x,y) \quad s.t\ g(x,y) = c \qquad x, y \in S$$

- $either\ g_1(x^*, y^*) \neq 0\ or\ g_2(x^*, y^*) \neq 0.$

Then there is a unique number λ such that (x^*, y^*) is a stationary point of the Lagrangean

$$L(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

That is, (x^*, y^*) satisfies the first-order conditions.

$$L_1(x^*, y^*) = f_1(x^*, y^*) - \lambda g_1(x^*, y^*) = 0$$

$$L_2(x^*, y^*) = f_2(x^*, y^*) - \lambda g_2(x^*, y^*) = 0.$$

and $g(x^*, y^*) = c$.

Procedure for the solution

1. Find all stationary points of the Lagrangean
2. Find all points (x, y) that satisfy $g_1(x, y) = \ 0 \,, g_2(x, y) = 0$ and $g(x, y) = c$
3. If the set S has boundary points, find all boundary points (x, y) that satisfy $g(x, y) = c$
4. The points you have found at which $f(x, y)$ is largest are the maximizers of $f(x, y)$

Example 1

$$\max_{\{x,y\}} x^a y^b \;\; s.t\; x + y = c$$

where $a, b > 0$ and $x^a y^b$ is defined for $x \geq 0\; y \geq 0$.

1.

$$L(x, y, \lambda) = x^a y^b - \lambda(x + y - c)$$

$$a x^{a-1} y^b - \lambda = 0$$

$$b x^a y^{b-1} - \lambda = 0$$

$$x + y = c$$

$$x = \frac{c\,a}{a+b} \qquad y = \frac{c\,b}{a+b} \qquad \lambda = \frac{a^a b^b}{(a+b)^{a+b-1}} \cdot c^{a+b-1}$$

The value of the objective function at the stationary point is:

$$x^a y^b = \frac{a^a b^b}{(a+b)^{a+b}} \cdot c^{a+b} > 0$$

2. $g_1(x, y) = 1$, $g_2(x, y) = 1$ then no values for which $g_1(x, y) = 0$, $g_2(x, y) = 0$

3. The boundary points of the set on which the objective function is defined is the set of points (x, y) with either $x = 0$ or $y = 0$. At every such point the value of objective function is 0

4. Then the solution of the problem is $x = \frac{10a}{a+b}$ $y = \frac{10b}{a+b}$

Interpretation of λ

$$\frac{\partial f^*(c)}{\partial c} = \lambda^*(c)$$

the value of the Lagrange multiplier at the solution of the problem is equal to the rate of change in the maximal value of the objective function as the constraint is relaxed.

Example 2:

$$\max_{\{x\}} x^2$$

$$s.t. \quad x = c$$

solution is $x = c$ so the maximized value of the objective function is c^2.

Its derivative respect to c is $2c$

Now consider the Lagrangean

$$L(x) = x^2 - \lambda(x - c)$$

The FOC is $2x - \lambda = 0$.

Then $x = c$ and $\lambda = 2c$ satisfy FOC and the constraint.

Note that λ is equal to the derivative of the maximized value of the function with respect to c

From example 1:

$$x = \frac{c\ a}{a+b} \quad y = \frac{c\ b}{a+b} \quad \lambda = \frac{a^a b^b}{(a+b)^{a+b-1}} \cdot c^{a+b-1}$$

the maximized value of the objective function is:

$$x^a y^b = \frac{a^a b^b}{(a+b)^{a+b}} \cdot c^{a+b}$$

its derivative respect to c is $\frac{a^a b^b}{(a+b)^{a+b-1}} \cdot c^{a+b-1}$, i.e. λ

Conditions under which a stationary point is a local optimum

$$\max_{\{x,y\}} f(x,y)$$
$$s.t\ g(x,y) = c$$

$$L(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

Borderd Hessian of the Lagrangean

$$H_b(x,y,\lambda) =$$

$$= \begin{matrix} 0 & g_1(x,y) & g_2(x,y) \\ g_1(x,y) & f_{11}(x,y) - \lambda g_{11}(x,y) & f_{12}(x,y) - \lambda g_{12}(x,y) \\ g_2(x,y) & f_{21}(x,y) - \lambda g_{21}(x,y) & f_{22}(x,y) - \lambda g_{22}(x,y) \end{matrix}$$

Suppose that it exists a value λ^* such that (x^*, y^*) is a stationary point of the Lagrangean.

To check if it is a local maximum

1) Compute the bordered Hessian at the values (x^*, y^*, λ^*) , i.e. $H_b(x^*, y^*, \lambda^*)$

2) Compute its determinant, i.e $|H_b(x^*, y^*, \lambda^*)|$

3) If $|H_b(x^*, y^*, \lambda^*)| > 0 \ then \ (x^*, y^*)$ is a **local maximizer**

Note, if $|H_b(x^*, y^*, \lambda^*)| < 0 \ then \ (x^*, y^*)$ is a **local minimizer**

Example 3

$$\max_{\{x,y\}} x^3 y \ \ subject\ to\ x + y = 6 \qquad x, y > 0$$

We simplify the problem using a log transformation

$$\max_{\{x,y\}} \ 3\ln x + \ln y \ \ subject\ to\ x + y = 6.$$

$$L(x, y) = 3\ln x + \ln y - \lambda(x + y - 6)$$

FOC are:

$$\frac{3}{x} - \lambda = 0, \qquad \frac{1}{y} - \lambda = 0 \qquad x + y = 6$$

The solution is $x = 4.5, y = 1.5, \ \lambda = \frac{2}{3}$

Borderd Hessian of the Lagrangean is

$$H_b(x,y,\lambda) = \begin{matrix} 0 & 1 & 1 \\ 1 & -\frac{3}{x^2} & 0 \\ 1 & 0 & -\frac{1}{y^2} \end{matrix} \quad H_b\left(4.5, 1.5, \frac{2}{3}\right) = \begin{matrix} 0 & 1 & 1 \\ 1 & -\frac{3}{4.5^2} & 0 \\ 1 & 0 & -\frac{1}{1.5^2} \end{matrix}$$

The determinant is $\frac{3}{4.5^2} + \frac{1}{1.5^2} > 0$, then the solution is a **local maximizer**

Conditions under which a stationary point is a global optimum

- Suppose that *f* and *g* are continuously differentiable functions defined on an open convex subset *S* of two-dimensional space and
- suppose that there exists a number λ^* such that (x^*, y^*) is an interior point of *S* that is a stationary point of the Lagrangean

$$L(x, y) = f(x, y) - \lambda^*(g(x, y) - c).$$

- Suppose further that $g(x^*, y^*) = c$.
- **Then if L is concave then (x^*, y^*) solves the problem**

$$\max_{\{x,y\}} f(x, y) \quad s.t\ g(x, y) = c$$

Example 4

Consider example 3

$$\max_{\{x,y\}} x^3 y \ \ subject\ to\ x + y = 6 \quad x, y > 0$$

We found that the solution of the FOC $x = 4.5, y = 1.5, \lambda = \frac{2}{3}$ is a local maximizer. **Is it a global maximizer?**

For a global maximizer we need that Lagrangean is concave

$$L(x, y) = 3\ln x + \ln y - \lambda(x + y - 6)$$

Given that constraint is linear we need to check the objective function

The Hessian of the objective function is

$$H = \begin{pmatrix} -\frac{3}{x^2} & 0 \\ 0 & -\frac{1}{y^2} \end{pmatrix}$$

$$H = \begin{pmatrix} -\frac{3}{x^2} & 0 \\ 0 & -\frac{1}{y^2} \end{pmatrix}$$

The leading principal minors are:

$|D_1| = -\frac{3}{x^2} < 0$ and $|D_2| = \frac{3}{x^2}\frac{1}{y^2} > 0 \; for \; all \; x, y > 0$

Then the Hessian is negative definite, so the objective function is strictly concave, and the point $x = 4.5, y = 1.5$ is a global maximum.

Optimization with equality constraints: *n* variables, *m* constraints: necessary conditions

Let f and $g_1, \ldots, g_m$ be continuously differentiable functions of n variables defined on the set S,

let c_j for $j = 1, \ldots, m$ be numbers, and suppose that x^* is an interior point of S that solves the problem:

$$\max_{\{x\}} f(x) \ subject\ to \ \ g_j(x) = c_j \ for\ j = 1, \ldots, m$$

Suppose also that the rank of the Jacobian matrix is m

$$J = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

Then there are unique numbers $\lambda_1, \ldots, \lambda_m$ such that x^* is a stationary point of the Lagrangean function L defined by

$$L(x) = f(x) - \sum_{j=1}^{m} \lambda_j \left(g_j(x) - c_j \right)$$

That is, x^* satisfies the first-order conditions:

$$L_i'(x^*) = f_i'(x^*) - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(x^*) = 0 \; for \; i = 1, \ldots, n.$$

In addition, $g_j(x^*) = c_j \; for \; j = 1, \ldots, m$.

Conditions under which necessary conditions are sufficient

Suppose that f and g_j for $j = 1, \ldots, m$ are continuously differentiable functions defined on an open convex subset S of *n*-dimensional space and let $x^* \in S$ be an interior stationary point of the Lagrangean:

$$L(x) = f(x) - \sum_{j=1}^{m} \lambda_j^* \left(g_j(x) - c_j\right)$$

suppose further that $g_j(x^*) = c_j \; for \; j = 1, \ldots, m$.

Then if L is concave then x^* solves the constrained maximization problem

Interpretation of λ

Consider the problem

$$\max_{\{x\}} f(x) \; subject \; to \; g_j(x) = c_j \; for \; j = 1, \ldots, m$$

Let $x^*(c)$ be the solution of this problem, where $c = (c_1, \ldots, c_m)$ and let $f^*(c) = f(x^*(c))$.

Then we have

$$\frac{\partial f^*(c)}{\partial c_j} = \lambda_j(c) \; for \; j = 1, \ldots, m,$$

where λ_j is the value of the Lagrange multiplier on the *j*th constraint at the solution of the problem.

Interpretation of λ

The value of the Lagrange multiplier on the *j*th constraint at the solution of the problem is equal to the rate of change in the maximal value of the objective function as the *j*th constraint is relaxed.

If the *j*th constraint arises because of a limit on the amount of some resource, then we refer to $\lambda_j(c)$ as the **shadow price** of the *j*th resource.

Quasi-concave functions

Let $f(x)$ be defined on the set S.

Then for all pairs $x, x' \in S$ $x, \neq x'$ and for all $\lambda \in (0, 1)$:

- If $f(x'') = f(\lambda x + (1-\lambda)x') \geq \min\big(f(x), f(x')\big)$ then $f(x)$ is **quasi concave**
- If $f(x'') = f(\lambda x + (1-\lambda)x') > \min(f(x), f(x'))$ then $f(x)$ is **strictly quasi concave**

Note these conditions hold even if $f(x) = f(x')$

A concave function is also quasi-concave, but the opposite is not true

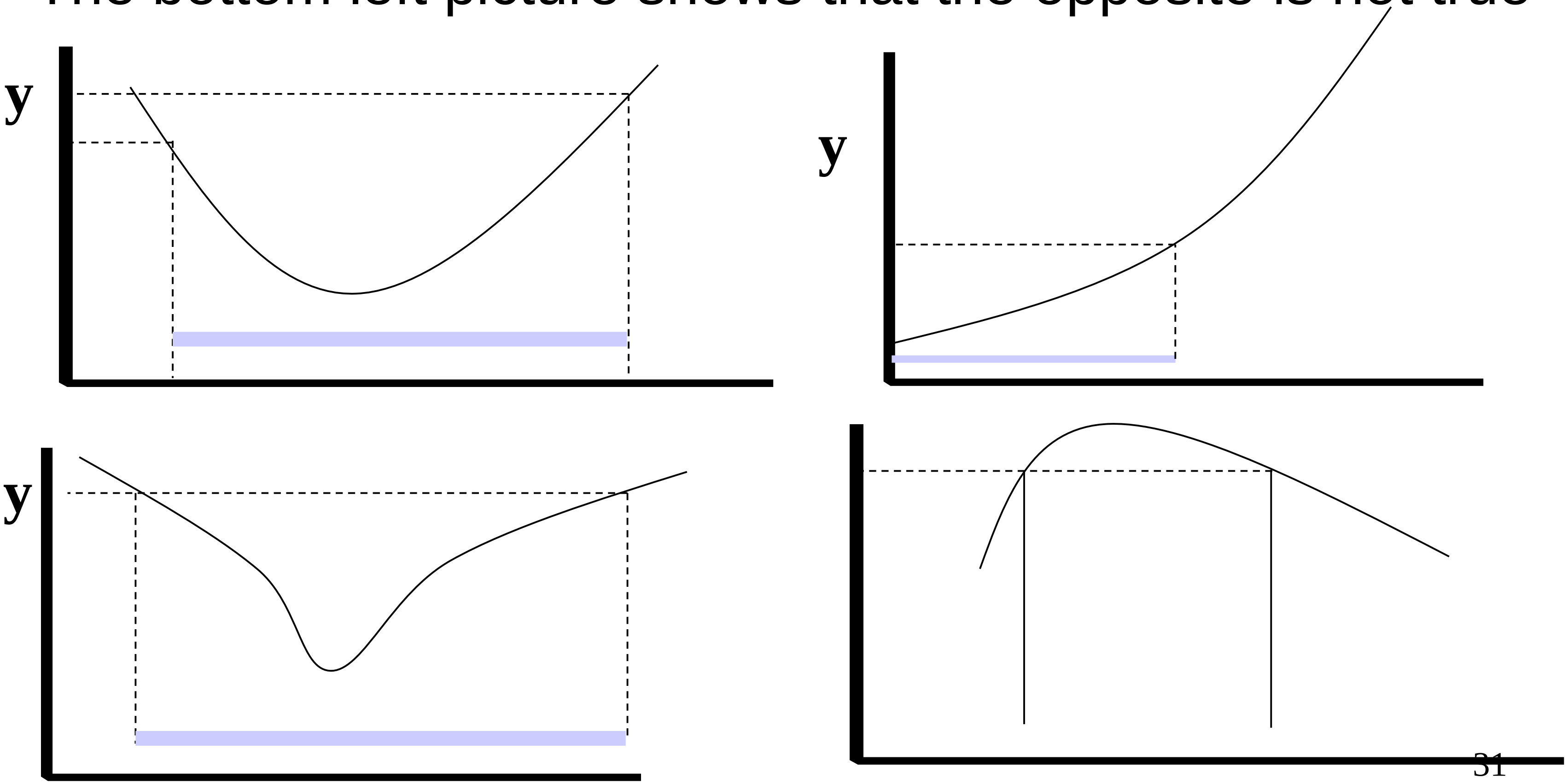
If $f(x) > f(x')$ and $f(\lambda x + (1-\lambda)x') > f(x')$ the function is **explicitly quasi concave**

It is **quasiconvex** if:

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

Note that a convex function is also quasi-convex

The bottom left picture shows that the opposite is not true



The importance of concavity and quasi-concavity

Consider the problem

$$\max_{\{x\}} f(x) \ subject\ to\ g_j(x) = c_j\ for\ j = 1, \ldots, m$$

and let x^* be a stationary point of the lagrangean

If

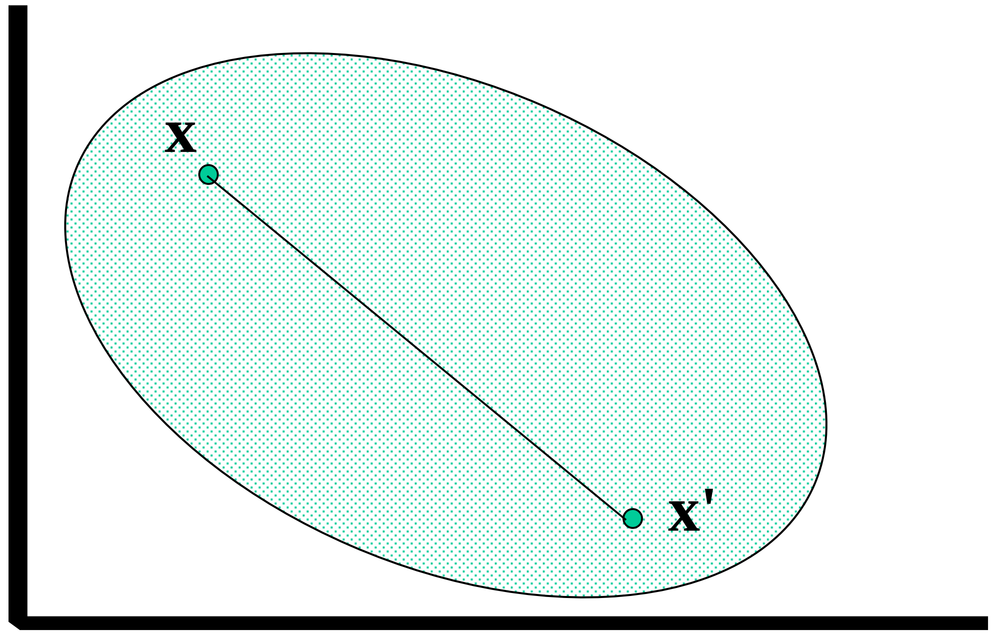
1. $f(x)$ is explicitly quasi-concave
2. The constrained set is convex
3. then x^* is a global maximum

If

1. $f(x)$ is strictly quasi-concave
2. The constrained set is convex
3. then x^* is the unique global maximum

Convex sets.

A convex set, X, is such that for any two elements of the set, x and x’ any convex combination of them is also a member of the set.



More formally, X is convex if for all x and $x' \in X$, and $0 \leq \lambda \leq 1$,

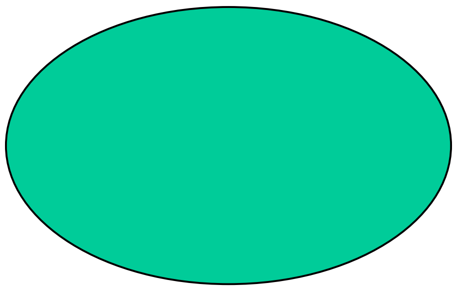
$x'' = \lambda x + (1 - \lambda)x' \in X$.

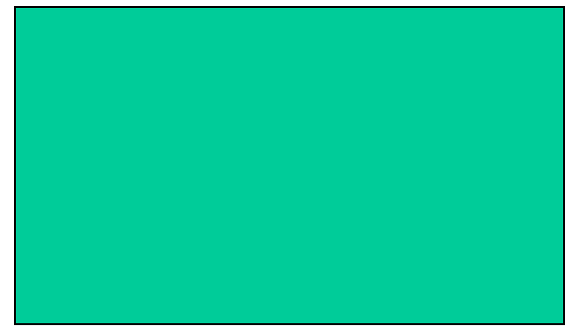
Sometimes X is described as strictly convex if for *any* $0 < \lambda < 1$, x'' is in the interior of X (i.e. not on the edges)

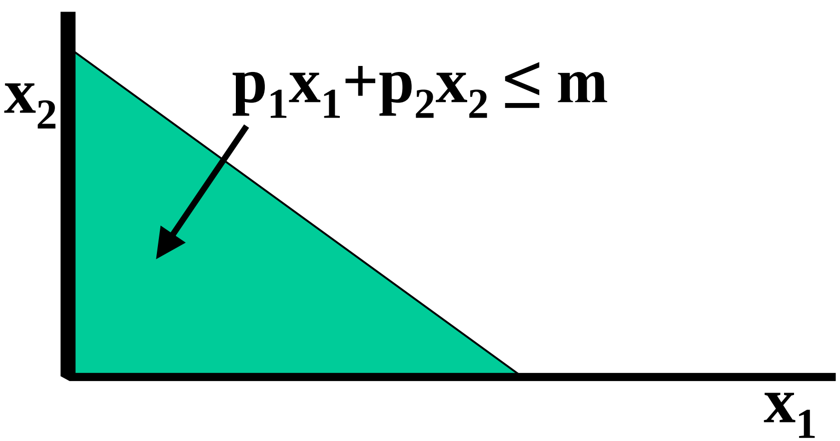
e.g. convex but not strictly convex

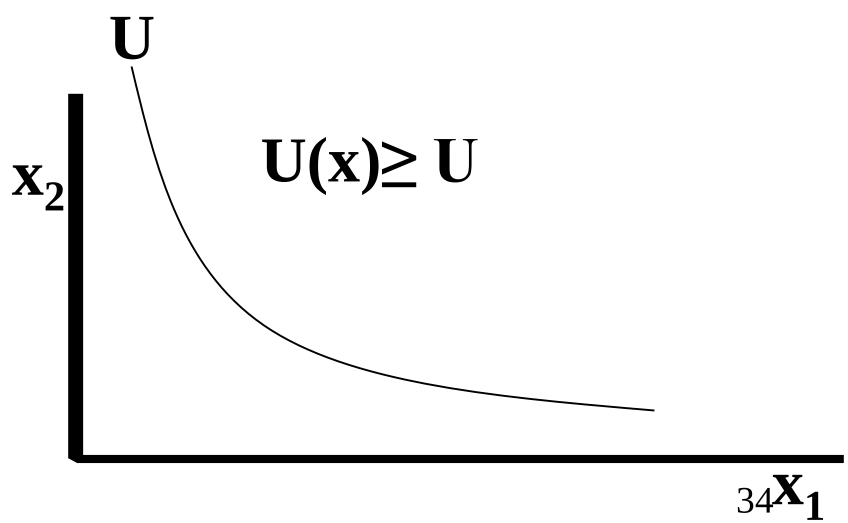
Convex sets.

If, for any two points in the set S, the line segment connecting these two points lie entirely in S, then S is a convex set.

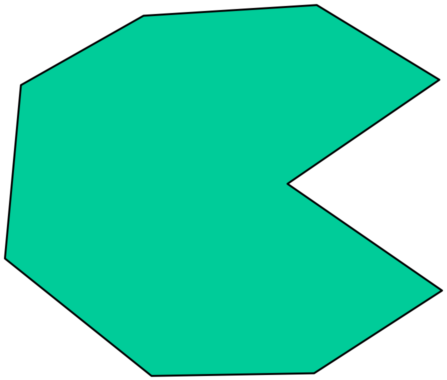


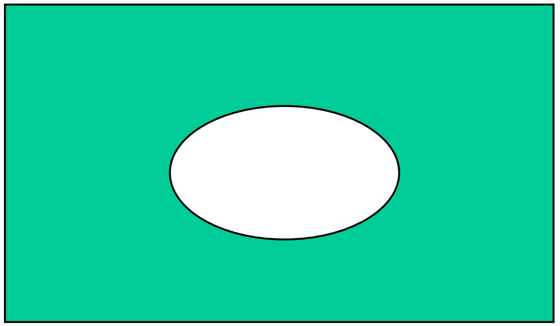


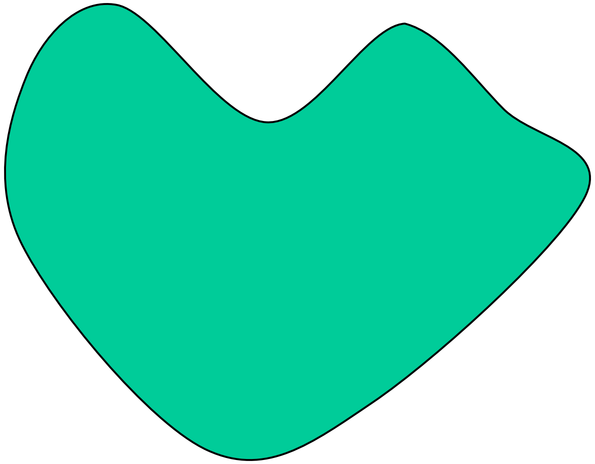


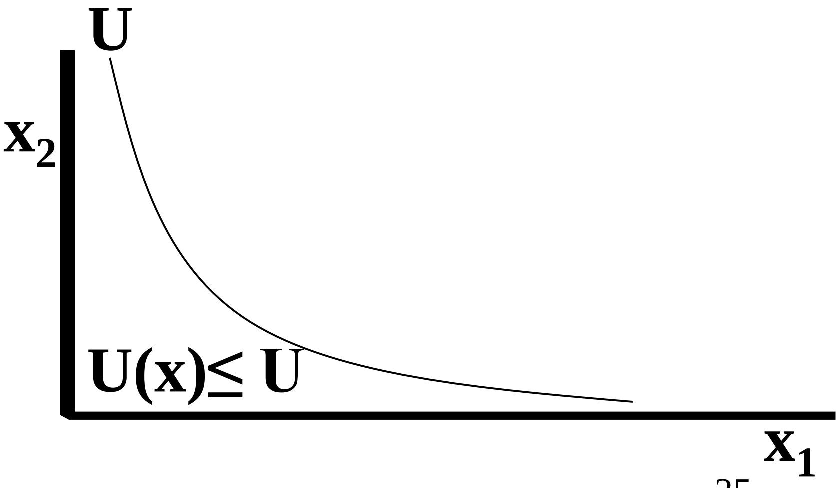


Non-Convex sets.









A different definition of quasi concavity

Let f be a multivariate function defined on the set S.

f is *quasi concave* if, for any number a, the set of points for which $f(x) \geq a$ is convex.

For any real number a, the set $P_a = \{x \in S\colon f(x) \geq a\}$ is called the **upper level set** of f for a.

The multivariate function f defined on a convex set S is **quasiconcave** if **every** upper level set of f is convex.

That is, $P_a = \{x \in S\colon f(x) \geq a\}$ is convex for every value of a.

Example 5

1. $f\ (x, y)\ =\ x^2\ +\ y^2$.

The upper level set of f for a is the set of pairs (x, y) such that $x^2\ +\ y^2\ \geq\ a$.

Thus for $a\ >\ 0$ it the set of point out of a disk of radius a, then the upper level set is not convex

2. $f\ (x, y)\ =\ -x^2\ -\ y^2$.

The upper level set of *f* for *a* is the set of pairs (x, y) such that $-x^2\ -\ y^2\ \geq\ a$, or $x^2\ +\ y^2\ \leq\ -a$.

Thus for $a > 0$ the upper level set P_a is empty

for $a < 0$ it is the set of points inside a disk of radius a.

Checking quasi concavity

To determine whether a twice-differentiable function is quasi concave or quasi convex, we can examine the determinants of the **bordered Hessians** of the function, defined as follows:

$$B = \begin{matrix} 0 & f_1(x) & f_2(x) & & f_n(x) \\ f_1(x) & f_{11}(x) & f_{12}(x) & & f_{1n}(x) \\ f_2(x) & f_{21}(x) & f_{22}(x) & & f_{2n}(x) \\ & & & & \\ f_n(x) & f_{n1}(x) & f_{n2}(x) & & f_{nn}(x) \end{matrix}$$

We have to compute the determinants of the leading principal minors

$$|B_1| = \begin{vmatrix} 0 & f_1(x) \\ f_1(x) & f_{11}(x) \end{vmatrix}, |B_2| = \begin{vmatrix} 0 & f_1(x) & f_2(x) \\ f_1(x) & f_{11}(x) & f_{12}(x) \\ f_2(x) & f_{21}(x) & f_{22}(x) \end{vmatrix}, \ldots\ldots\ldots\ldots\ldots$$

If f is quasi concave then
$|B_x| \geq 0$ if x is even and
$|B_x| \leq 0$ if x is odd

If then $|B_x| > 0$ if x is even and
$|B_x| < 0$ if x is odd
then f is strictly quasi concave

If f is quasi convex then $|B_x| \leq 0$ are negative
If$|B_x| < 0$ then f is strictly quasi convex

for all x in the set where function f is defined

Envelope theorem: unconstrained problem

Let $f(x,r)$ be a continuously differentiable function where x is an *n*-vector of variables and r is a *k*-vector of parameters.

The maximal value of the function is given by $f(x^*(r),r)$

where $x^*(r)$, is the vector of variables x that maximize f and that are function of r.

Note that we can write $f(x^*(r),r)$ as $f^*(r)$

(because in this function only parameters appear)

If the solution of the maximization problem is a continuously differentiable function of *r* then:

$$\frac{df^*(r)}{dr_i} = \frac{df(x,r)}{dr_i} \quad \text{evaluated in } x^*(r),$$

the change in the maximal value of the function as a parameter changes is the change caused by the direct impact of the parameter on the function, holding the value of *x* fixed at its optimal value;

the indirect effect, resulting from the change in the optimal value of *x* caused by a change in the parameter, is zero

Example 6

$$\max \; p \ln x - cx$$

FOC is $\frac{p}{x} - c = 0$

then $x^* = \frac{p}{c}$

and $f^*(p, c) = p \ln \frac{p}{c} - p$

The effect of a change of parameter c on the maximized value is:

$$\frac{df^*(p, c)}{dc} = -\frac{p}{c}$$

Consider the derivative of the objective function evaluated at the solution x^*

$$\frac{dp \ln x - cx}{dc} = -x$$

Evaluating it in $x^* = \frac{p}{c}$ we get $-\frac{p}{c}$

Envelope theorem: constrained problems

Let $f(x, r)$ be a continuously differentiable function where x is an *n*-vector of variables and r is a *k*-vector of parameters.

The maximal value of the function is given by $f(x^*(r), r)$

where $x^*(r)$, is the vector of variables x that maximize f and that are function of r.

Note that we can write $f(x^*(r))$, as $f^*(r)$,

Then

$$\frac{df^*(r)}{dr_i} = \frac{dL\,(\mathrm{x}, r)}{dr_i} \text{ evaluated at the solution } x^*(r),$$

where the function L is the Lagrangean of the problem

Example 7 $\max_{\{x,y\}} xy \;\; s.t\; x + y = B$

$$L(x, y, \lambda) = xy - \lambda(x + y - B)$$

we solve: $y - \lambda = 0$

$x - \lambda = 0$

$x + y = B$

then $x^* = y^* = \lambda^* = \frac{B}{2}$ and $f^*(B) = \frac{B^2}{4}$

The effect of a change of parameter c on the maximized value is:

$$\frac{df^*(B)}{dB} = \frac{B}{2}$$

Consider the derivative of the Lagrangean evaluated at the solution x^*

$$\frac{d\big(xy - \lambda(x + y - B)\big)}{dB} = \frac{B}{2}$$