# Recitation 1 - Math Tools *

Jonathan Cohen

14.03 Fall 2022

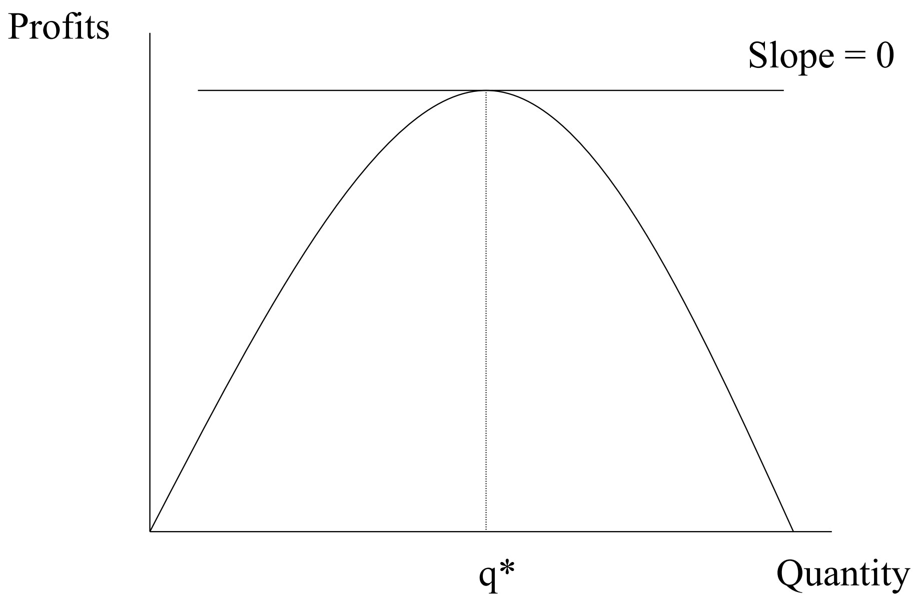
**Agenda**

- Optimization
  - single variable
  - multi-variable
- Implicit Function Theorem and comparative statics
- Envelope Theorem: constrained and unconstrained
- Constrained optimization (Lagrangian method)
- Duality

*Thanks to Professor Autor, Basil Halperin, and Andrea Manera for sharing materials from previous years.

## 1 Single Variable Optimization

Suppose that Elon Musk can produce $q$ Teslas at a cost of $q^2$. He sells Teslas at price $p > 0$. His profits are $\pi(q) = pq - q^2$ – this is the profit function – and he choose $q^*$ to maximize $\pi(q)$

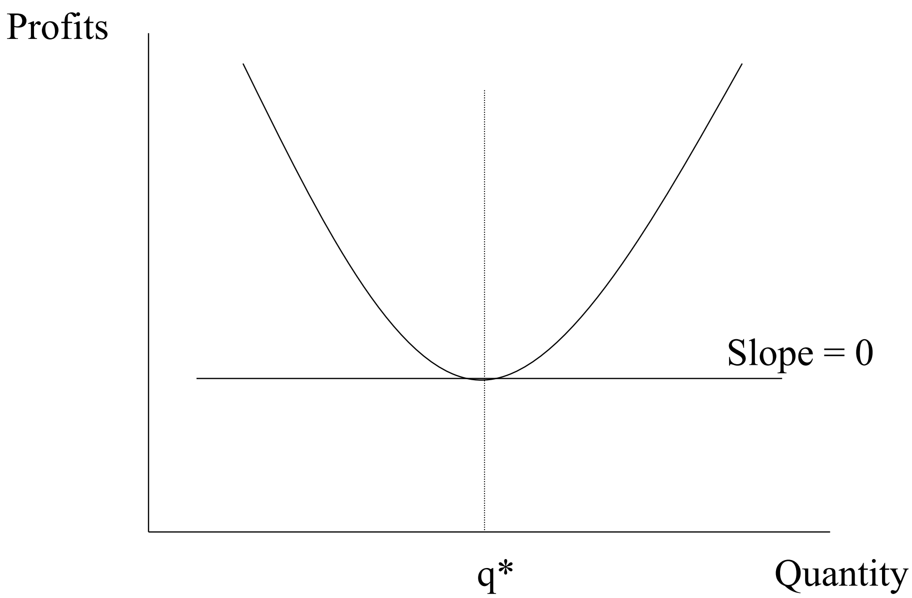



First Order Condition (FOC):

$$\begin{aligned} \left.\frac{\partial \pi}{\partial q}\right|_{q^*} &= 0 \\ p - 2q &= 0 \\ q^*(p) &= \frac{p}{2} \end{aligned}$$

Q: Is $q^*$ necessarily the profit max?

A: No, FOC is necessary, but not sufficient for profit maximization. The point that satisfies FOC could well be a minimum or an inflection point.

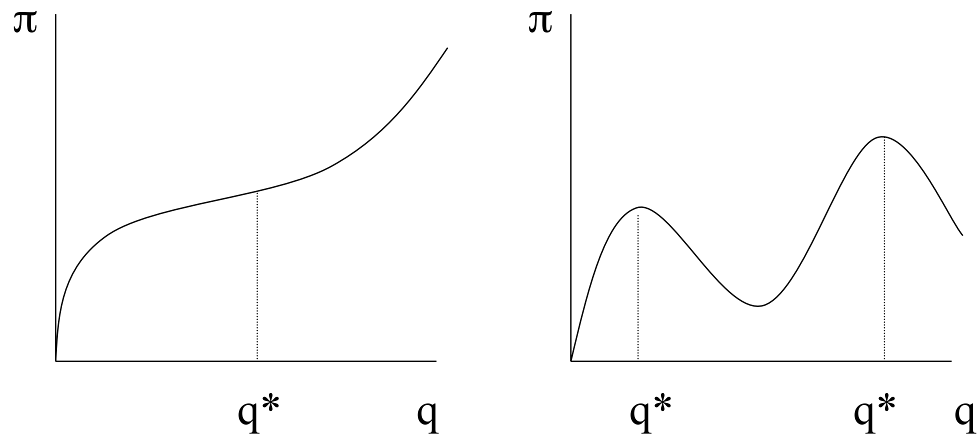



$$\left.\frac{\partial \pi}{\partial q}\right|_{q^*} = 0$$

but $q^*$ is a profit minimum.
We have to look at second order condition (SOC):

$$\left.\frac{\partial^2 \pi}{(\partial q)^2}\right|_{q^*} < 0$$
$$-2 < 0$$

This guarantees that $q^*$ is a local maximum.
This method doesn't help when the function is not well behaved.
Examples:




We'll generally work with "well-behaved" functions: continuous, differentiable, concave. Hence we will typically not focus on SOC.

## 2 Multivariate Optimization

Given a function:

$$y = f(x_1, x_2, ..., x_n)$$

and given all partial derivatives:

$$\frac{\partial f}{\partial x_1} \equiv f_1, \;\; \frac{\partial f}{\partial x_2} \equiv f_2, \; ..., \; \frac{\partial f}{\partial x_n} \equiv f_n$$

First order condition (FOC) for maximum (or minimum):

$$f_1 = f_2 = ... = f_n = 0$$

For multivariable functions the sufficient conditions for a critical point to be a minimum or a maximum turn out to be whether the Hessian matrix is positive definite or negative definite (all negative or all positive eigenvalues). Fortunately, we won't worry about this much and when we do it'll be in a case with two variables in which case there's a shortcut.

The Second Order Condition (SOC) for a maximum when there is more than one variable has the following form:

$$\begin{aligned} f_{11} &< 0 \\ f_{11}f_{22} - (f_{12})^2 &> 0 \end{aligned}$$

Note that $f_{22} < 0$ is implied by the above, since if $f_{11}$ and $f_{22}$ have opposite signs there is no way to get $f_{11}f_{22} - (f_{12})^2 > 0$. Also note that it is not sufficient to have just $f_{11} < 0$ and $f_{22} < 0$.

For a minimum, the SOC takes the form

$$\begin{aligned} f_{11} &> 0 \\ f_{11}f_{22} - (f_{12})^2 &> 0 \end{aligned}$$

If $f_{11}f_{22} - (f_{12})^2 < 0$, then the point in question is neither a minimum nor a maximum.

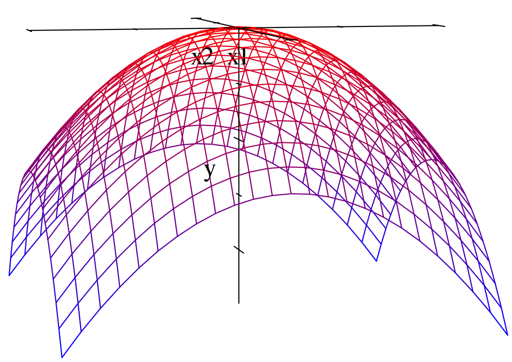
### 2.1 Concave Functions

To simplify this we will often work with concave functions because they have the nice property that the FOC is sufficient for a global max.

**Definition 1** *A concave function is a function that always lies below any hyperplane that is tangent to it.*

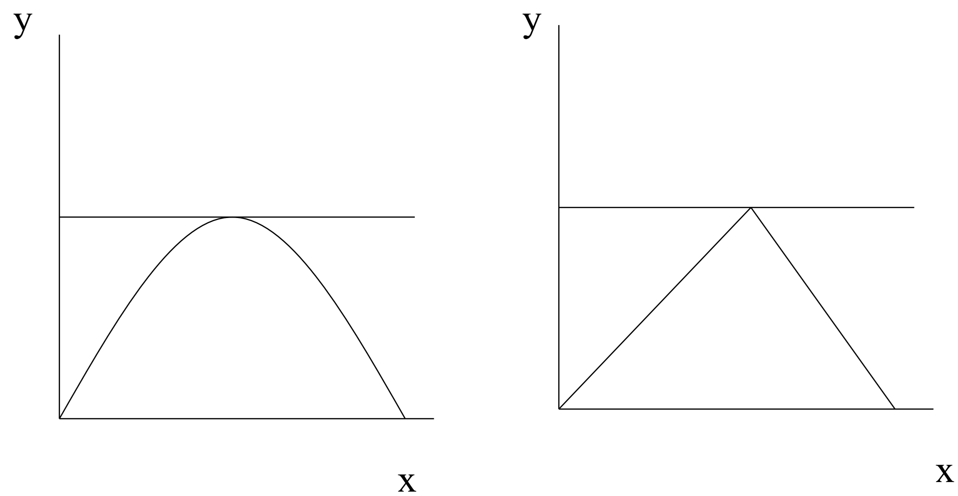
For example a function of one variable is concave if it always lies below any line tangent to it.

$$-1000x^2 - 1000y^2 - xy + 200$$

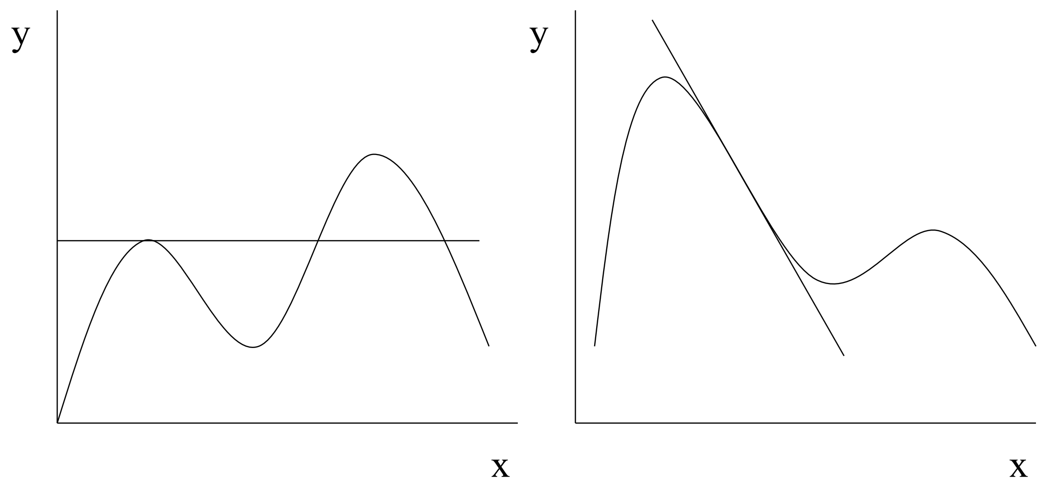



Another useful definition for concave functions is that for any two points $a, b$ and $\alpha \in (0,1)$ $f(\alpha a + (1-\alpha)b) \geq \alpha f(a) + (1-\alpha)f(b)$ that is that the graph lies above any line that connects the two.

This set of functions satisfies the condition for concavity:




This set of functions doesn't satisfy the condition for concavity:




### 2.2 Example

For example, now suppose we need both capital and labor to produce a good so our profit function becomes $\pi(l,k) = \ln(l) + \ln(k) - wl - rk$. We set both derivatives to zero:

$$
\begin{aligned}
\frac{\partial \pi}{\partial l} &= 0 \implies \frac{1}{l} - w = 0 \\
l^* &= \frac{1}{w} \\
\frac{\partial \pi}{\partial k} &= 0 \implies \frac{1}{k} - r = 0 \\
k^* &= \frac{1}{r}
\end{aligned}
$$

And we can check our two variable conditions:

$$
\begin{bmatrix} -\frac{1}{l^2} & 0 \\ 0 & -\frac{1}{k^2} \end{bmatrix} \implies f_{xx} < 0 \text{ and } f_{xx}f_{yy} - f_{xy}^2 > 0
$$

Or we can note that $\pi(l, k)$ is concave and thus the FOC are sufficient.

We often don't want to just know what the max is. We often want to know the how changes in one variable around the maximum affect the other optimized values (comparative statics). The implicit function theorem and the envelope theorem help us with this.

# 3 Implicit Functions

Many times in economics we end up with implicit functions where exogenous and endogenous variables are all mixed together but we still want to know how the change in one variable (say $x$) affects another say ($y$). If we could solve for $y^*(a)$ explicitly we could just differentiate and we'd be fine, but the fact that they're all mixed up means we sometimes can't solve explicitly for $y^*$. However, the derivative $\frac{\partial y}{\partial x}$ may still exist. The implicit function theorem helps us find *IF it exists*.

Functions can be written in their implicit form or in their explicit form.
Examples:

1. $y = mx + b$ Explicit
2. $y - mx - b = 0$ Implicit
3. $f(y, x; m, b) = 0$ Implicit

Functions 2 and 3 are called implicit because the relationship between the variables is implicitly present rather than explicitly shown as $y = f(x)$. One common source of implicit functions are first order conditions.

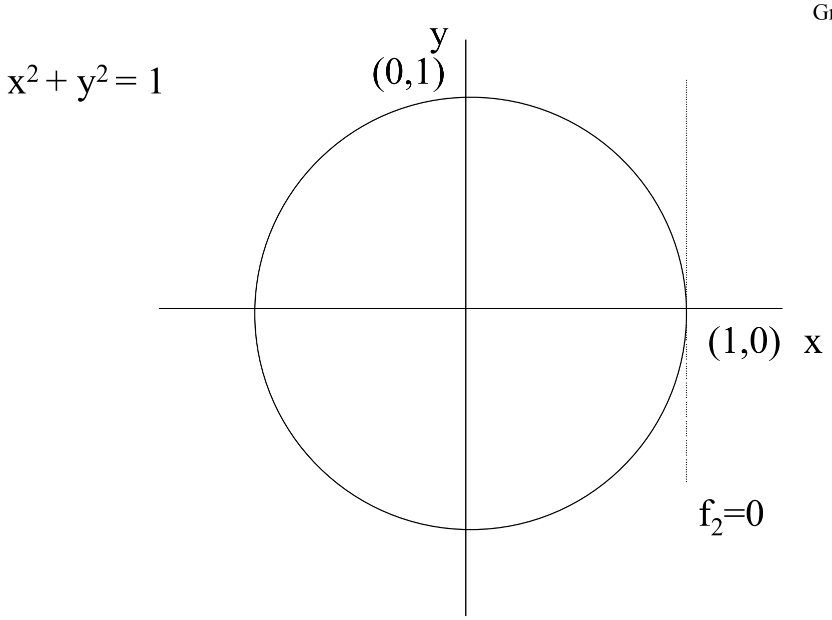
It's easy to work with implicit functions.

$$\begin{aligned} f(x,y) &= 0 \\ f(x,y(x)) &= 0 \\ f_x dx + f_y dy &= 0 \\ \frac{dy}{dx} &= -\frac{f_x}{f_y} \end{aligned}$$

Caveat: $\frac{dy}{dx}$ may not exist...

### 3.1 Example

Graph 21




We can write the equation for this function as follows:

$$f(x^*, y(x^*)) = 0$$

which yield the following equation:

$$x^2 + y^2 - 1 = 0$$

We can differentiate and find the derivative:

$$\begin{aligned} 2x dx + 2y dy &= 0 \\ \frac{dy}{dx} &= -\frac{f_1}{f_2} = -\frac{x}{y} \end{aligned}$$

This derivative is not defined when $y = 0$.

Q: What is the intuition for the non-existence of $\frac{dy}{dx}$ at $(x, y) = (1, 0)$?
A: $\frac{dy}{dx}$ could be positive or negative here. Undefined.
You can see how this works formally.
Suppose there is a continuous solution $y = y(x)$ for the equation $F(x, y) = c \Rightarrow F(x, y(x)) = c$
We want to know $\frac{\partial y}{\partial x}$ for some $(x_0, y(x_0))$.
Use the chain rule to differentiate.

$$
\begin{aligned}
\frac{dF(x_0, y(x_0))}{dx} &= \frac{\partial F(x_0, y(x_0))}{\partial x}\frac{dx}{dx} + \frac{\partial F(x_0, y(x_0))}{\partial y}\frac{dy(x_0)}{dx} = 0 \\
0 &= \frac{\partial F(x_0, y(x_0))}{\partial x} + \frac{\partial F(x_0, y(x_0))}{\partial y} y'(x_0) \\
y'(x_0) &= -\frac{\frac{\partial F(x_0, y(x_0))}{\partial x}}{\frac{\partial F(x_0, y(x_0))}{\partial y}}
\end{aligned}
$$

Necessary condition for $y'(x_0)$ to exist is that $\frac{\partial F(x_0, y(x_0))}{\partial y} \neq 0$ (Implicit function theorem)
It turns out that this is sufficient also.
In the multivariate case this condition can be written as $\frac{\partial F(x_1^*, \ldots, x_n^*, y^*)}{\partial y} \neq 0$

### 3.2 Example

Given the following function:

$$
\begin{aligned}
2x^2 + y^2 &= 225, \\
y &\geq 0
\end{aligned}
$$

we want to find $\frac{dy}{dx}$.

1. One way to find the derivative we are looking for is to find $y$ as a function of $x$

$$
\begin{aligned}
y &= \sqrt{225 - 2x^2} \\
\frac{dy}{dx} &= \frac{1}{2}\left(225 - 2x^2\right)^{-\frac{1}{2}}(-4x) = \\
&= \frac{-4x}{2\sqrt{225 - 2x^2}} = -\frac{2x}{y}
\end{aligned}
$$

2. The other way to find it is to use the implicit function method:

- Write: $2x^2 + y^2 - 225 = 0$
- Find the total differential: $4xdx + 2ydy = 0$
- Rearrange:

$$\frac{dy}{dx} = -\frac{2x}{y}$$

### 3.3 Example

Take a more complicated example:

$$g(x,y) = x^2 - 3xy + y^3 - 7 = 0$$

What is $\left.\frac{dy}{dx}\right|_{x=4,y=3}$?

In this case making use of the implicit function theorem is the only way to find the derivative we are interested in.

1. Find total differential:

$$\begin{aligned} 2xdx - 3ydx - 3xdy + 3y^2dy &= 0 \\ 3y^2dy - 3xdy &= -2xdx + 3ydx \\ \frac{dy}{dx} &= \frac{3y - 2x}{3y^2 - 3x} = -\frac{g_x(x,y)}{g_y(x,y)} \end{aligned}$$

2. To find the derivative at the point we simply substitute:

$$\left.\frac{dy}{dx}\right|_{x=4,y=3} = \frac{9-8}{27-12} = \frac{1}{15}$$

3. What is $y$ at $x = 4.3$?

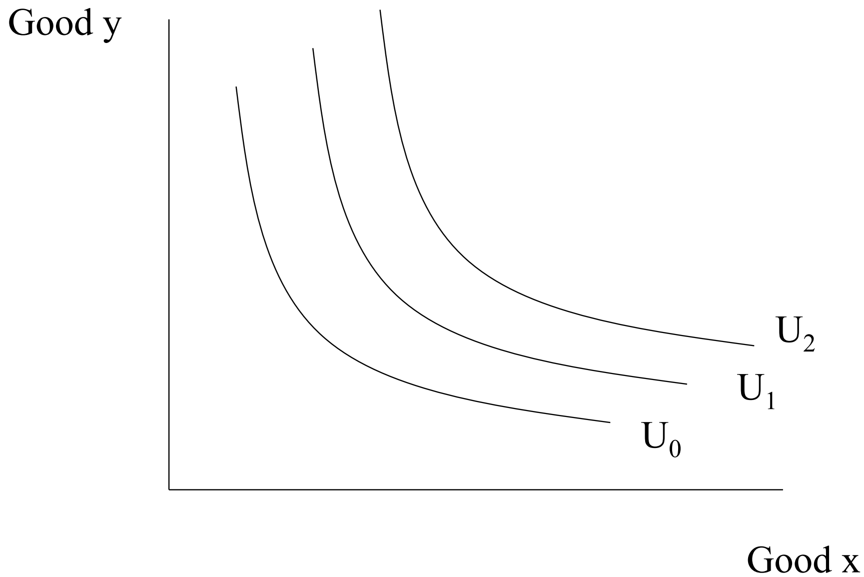
   We can approximate it by:

$$\begin{aligned} & y(4) + \left.\frac{dy}{dx}\right|_{x=4}(0.3) \\ y(4.3) &\approx 3 + 0.3 \times \frac{1}{15} = 3.02 \end{aligned}$$

   If we solve numerically for $y$ at $x = 4.3$ we get 3.01475.

### 3.4 Applications of Implicit Functions

Being able to handle implicit functions turns out to be useful when one deals with utility functions and indifference curves.




Along an indifference curve, we have $U(x,y) = \overline{U}$

The implicit function $U(x^*, y^*(x^*)) = \overline{U}$ tells how much $y$ we'd give up for a little more $x$ (at the margin) while holding total utility constant.

$$\begin{aligned} U(x^*, y^*(x^*)) &= \overline{U} \\ \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy &= 0 \\ \frac{dy}{dx} &= -\frac{U'(x)}{U'(y)} \end{aligned}$$

## 4 Envelope Theorems

A shortcut for taking derivatives of optimized functions with respect to their parameters.

**Theorem 2** *(Envelope Theorem for the unconstrained case). Let $f(x,a)$ be a $C^1$ function of $x \in R^n$ and the scalar $a$. For each $a$ consider the unconstrained maximization:*

$$\max_x \; f(x;a)$$

*Let $x^*(a)$ be a solution of this problem. Suppose that $x^*(a)$ is a $C^1$ function of $a$. Then,*

$$\underbrace{\frac{d}{da}f(x^*(a),a)}_{\textit{Total derivative}} = \underbrace{\frac{\partial}{\partial a}f(x^*(a),a)}_{\textit{Partial derivative}}$$

**Proof.**

$$\begin{aligned} \frac{d}{da}f(x^*(a),a) &= \underbrace{\sum \frac{\partial f}{\partial x_i}(x^*(a),a)\frac{\partial x_i^*(a)}{\partial a}}_{=0} + \frac{\partial f}{\partial a}(x^*(a),a) = \\ &= \frac{\partial f}{\partial a}(x^*(a),a) \end{aligned}$$

where the first term of the derivative is zero because:

$$\frac{\partial f}{\partial x_i}(x^*(a),a) = 0 \ \ \forall i = 1,2,...,n$$

These are the FOC of the maximization problem to obtain $x^*$. ■

Note: much more intuitive - <u>and useful</u> than it looks

**Example:**

Take the following function

$$y = -x^2 + ax$$

We want to know $\frac{dy^*}{da}$ where $y^*$ is the maximized value of the above function. We can proceed in the two ways:

1. Find $x^*$ through single variable optimization and then substitute.

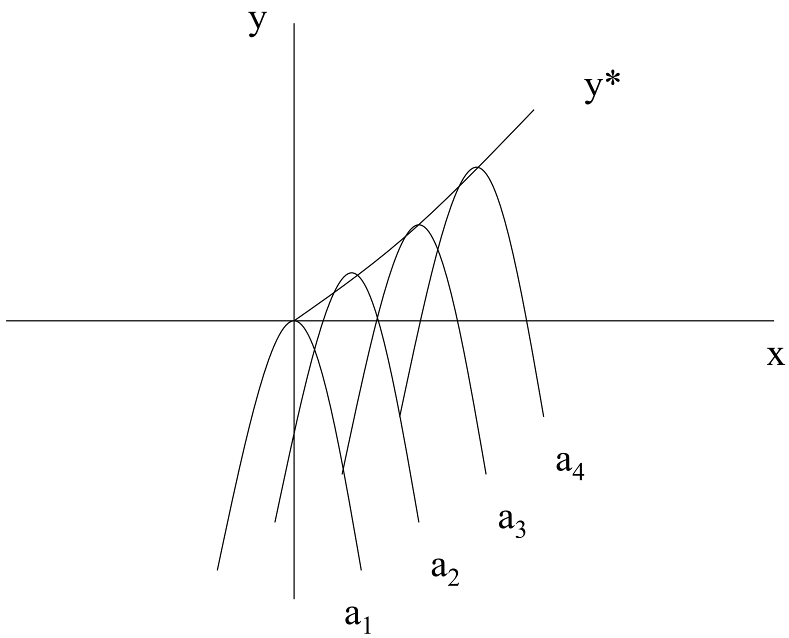
$$\begin{aligned} \frac{dy}{dx} &= -2x + a = 0 \\ x^* &= \frac{a}{2} \\ y^* &= -\left(\frac{a}{2}\right)^2 + a\left(\frac{a}{2}\right) = \frac{a^2}{4} \\ \frac{dy^*}{da} &= \frac{a}{2} = x^* \end{aligned}$$

2. Find the derivative using the envelope theorem:

$$\begin{aligned} y^* &= -(x^*)^2 + ax^* \\ \left.\frac{\partial y}{\partial a}\right|_{x=x^*} &= x^* = \frac{dy^*}{da} \end{aligned}$$

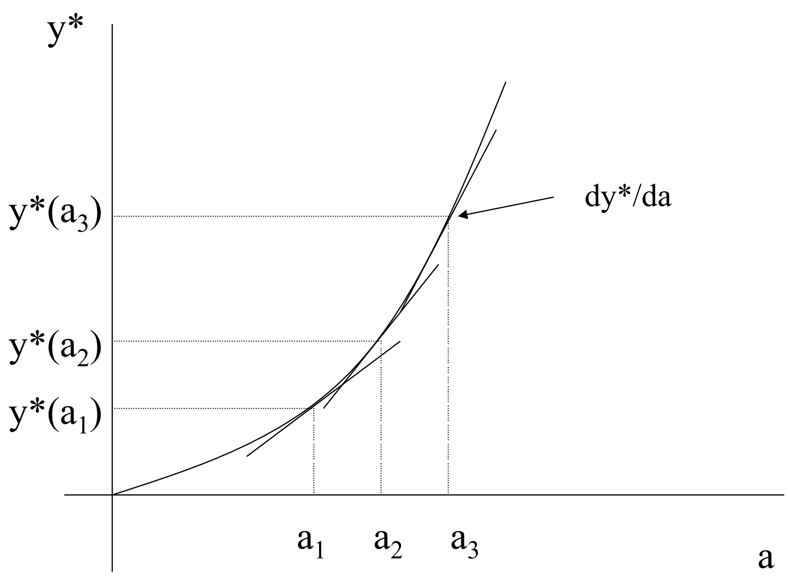
### 4.1 Visual Explanation of the Envelope Theorem

Remember that $y = -x^2 + ax$ and $y^* = f(a, x^*(a)) = \frac{a^2}{4}$




Note: Envelope Theorem is a linear approximation and hence holds in an "envelope" surrounding $x^*(a)$.

It is called Envelope Theorem because we are evaluating the upper envelope of a function.




The function drawn in this graph is such that its derivative at each point is

$$\frac{\partial y^*}{\partial a} = x^*(a)$$

*Remember:* Envelope Theorem is multi-variate

$$\begin{aligned}
y^* &= f\left[x_1^*(a), x_2^*(a), ..., x_n^*(a); a\right] \\
\frac{dy^*}{da} &= \underbrace{\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial a} + ... + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial a}}_{=0} + \frac{\partial f}{\partial a} \\
\frac{dy^*}{da} &= \frac{\partial f}{\partial a}
\end{aligned}$$

# 5 Constrained Maximization

Most maximization problems in economics are subject to constraints:

- maximize utility subject to budget constraint
- maximize social welfare subject to a resource constraint
- maximize profits subject to a technological constraint

The tool for maximizing constrained functions is the Lagrangian Method.
This a “trick" that turns out to have very useful economic content.

## 5.1 Lagrangian Method

*Problem:*

$$\begin{aligned}
\max\ y &= f(x_1, x_2, ..., x_n) \\
s.t.\ g(x_1, x_2, ..., x_n) &= 0
\end{aligned}$$

*Setup:*

$$\mathcal{L} = f(x_1, x_2, ..., x_n) + \lambda g(x_1, x_2, ..., x_n)$$

FOC's

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1} &= f_1 + \lambda g_1 = 0 \\
&... \\
\frac{\partial \mathcal{L}}{\partial x_n} &= f_n + \lambda g_n = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= g = 0
\end{aligned}$$

This way we obtain as many equations as unknowns, since we introduced another unknown, $\lambda$.

One solves simultaneously for $x_1^*, x_2^*, ..., x_n^*$ and $\lambda$.

$\lambda$ has a special interpretation that we will discuss.

### 5.2 Example: Optimal fence dimensions

Given a fencing perimeter of length $p$ how do we maximize the fenced area (provided that the area must have a rectangular shape)?

So the problem can be summarized as follows:

$$\begin{aligned}
&\max\ xy \\
s.t.\ 2x + 2y &= p
\end{aligned}$$

The Lagrangian for this problem is:

$$\begin{aligned}
\mathcal{L} &= xy + \lambda(p - 2x - 2y) \\
\frac{\partial \mathcal{L}}{\partial x} &= y - 2\lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial x} &= x - 2\lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= p - 2x - 2y = 0 \\
\frac{y}{2} &= \frac{x}{2} = \lambda \\
x &= y = \frac{p}{4} \\
\lambda &= \frac{p}{8}
\end{aligned}$$

- We can conclude that the optimal fence is square $(x = y)$.
- What is the interpretation of $\lambda = \frac{p}{8}$?

Observe that:

$$\frac{f_1}{-g_1} = \frac{f_2}{-g_2} = \lambda$$

where $f_1$ is the marginal gain to the lagrangian from adding one more unit of $x$ and $g_1$ is the marginal cost of adding more $x$ in terms of tightening the constraint and hence reducing feasible $y$.

This ratio, $\lambda$, is called the "shadow price" of the constraint.

$\lambda$ is, in other words, the opportunity cost of the constraint at the margin expressed in units of the maximand. It is the gain in terms of maximand obtained by relaxing the constraint by one unit.

In our example $\lambda$ tells us the increase in area we can obtain by increasing the size of the perimeter by one unit.

$\lambda = \frac{p}{8}$ implies that relaxing the constraint that $2x + 2y = p$ by one unit would allow us to increase the maximand area by $\frac{p}{8}$.

Let's check this:

Let $p = 40 \;\Rightarrow\; x = y = 10,\; A = 100$

Now let $p = 41 \;\Rightarrow\; x = y = 10.25,\; A = 105.06$ which confirms that $\Delta A = 5.06 \approx \frac{40}{8}$

The multiplier $\lambda$ is quite close to the actual change in A for a one-unit change in the constraint (and it would be exactly correct for a small enough change in the perimeter).

## 5.3 Example

$$\max \; U = x^{\frac{1}{2}} y^{\frac{1}{2}}$$

$$\begin{aligned}
s.t.\; x + y &= I \\
\mathcal{L} &= x^{\frac{1}{2}} y^{\frac{1}{2}} + \lambda\left(I - x - y\right) \\
\frac{\partial \mathcal{L}}{\partial x} &= \frac{1}{2} x^{-\frac{1}{2}} y^{\frac{1}{2}} - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial y} &= \frac{1}{2} x^{\frac{1}{2}} y^{-\frac{1}{2}} - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= I - x - y = 0 \\
x &= y = \frac{I}{2},\;\; \lambda = \frac{1}{2}
\end{aligned}$$

Let's just check the multiplier's implication:

$$
\begin{aligned}
f(x,y,4) &= 2^{\frac{1}{2}}2^{\frac{1}{2}} = 2 \\
f(x,y,5) &= 2.5^{\frac{1}{2}}2.5^{\frac{1}{2}} = 2.5 \\
\Delta f &= \frac{1}{2} = \lambda
\end{aligned}
$$

## 5.4 Envelope Theorem for Constrained Problems

Let $x^*(a)$ denote the solution to the following problem:

$$
\begin{aligned}
&\max\ y = f(x) \\
s.t.\ g(x;a) &= 0
\end{aligned}
$$

Let $\lambda$ be the Lagrange multiplier for the constraint in this problem.
Then:

$$
\underbrace{\frac{d}{da}f(x^*(a))}_{\substack{Total\ derivative\ of \\ the\ original\ function\ f}} = \underbrace{\lambda\frac{\partial g(x;a)}{\partial a} = \frac{\partial}{\partial a}\mathcal{L}(x^*(a),\lambda(a),a)}_{\substack{Partial\ derivative \\ of\ Lagrangian}}
$$

Why is this true? First use the chain rule:

$$
\underbrace{\frac{d}{da}f(x^*(a))} = \sum_i \frac{\partial f(x^*(a),a)}{\partial x_i}\frac{dx_i^*}{da}
$$

The FOC for maximizing the Lagrangean $\mathcal{L}(x,\lambda,a) = f(x) + \lambda g(x;a)$ are

$$
\frac{\partial \mathcal{L}(x,\lambda,a)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} + \lambda^*\frac{\partial g}{\partial x_i} = 0
$$

So

$$
\frac{d}{da}f(x^*(a)) = -\lambda\sum_i \frac{\partial g}{\partial x_i}\frac{dx_i^*}{da}
$$

But by taking the derivative of $g(x;a) = 0$ with respect to $a$ we get

$$
\sum_i \frac{\partial g}{\partial x_i}\frac{dx_i^*}{da} + \frac{\partial g}{\partial a} = 0
$$

And this gives us our result.

This is much more obvious than it looks: it says that the marginal gain from increasing $a$ is the value of relaxing the constraint ($\lambda$) times the amount by which $a$ relaxes the constraint ($\frac{\partial g}{\partial a}$).

Consider our previous problem:

$$\begin{aligned} & \max \ x^{\frac{1}{2}}y^{\frac{1}{2}} \\ s.t.\ x+y &= I \end{aligned}$$

We found:

$$x^* = y^* = \frac{I}{2}, \ \ \lambda^* = \frac{1}{2}$$

What is:

1. $\frac{\partial f(x^*(a),y^*(a),a)}{\partial x^*}$?
2. $\frac{\partial f(x^*(a),y^*(a),a)}{\partial y^*}$?
3. $\frac{dx^*(a)}{da}$?
4. $\frac{dy^*(a)}{da}$?

# 6 Duality

Every *primal* maximization problem subject to a constraint has a corresponding *dual* problem that minimizes the constrained function subject to the original objective function being equal to its optimal value in the original problem.

Primal:

$$\begin{aligned} \max\ z &= f(x,y) \\ s.t.\ x+y &= \bar{k} \\ z^* &= f(x^*,y^*) \end{aligned}$$

Dual:

$$\begin{aligned} \min\ k &= x+y \\ s.t.\ f(x,y) &= z^* \\ k^* &= \bar{k} \end{aligned}$$

The two problems will yield the same optimal values:

$$\begin{aligned} x_P^* &= x_D^* \\ y_P^* &= y_D^* \\ z_P^* &= z_D^* \end{aligned}$$

where $P$ stands for primal and $D$ stands for dual.

## 6.1 Example

*Primal* problem:

$$\begin{aligned} \max\ z &= x^{\frac{1}{2}}y^{\frac{1}{2}} \\ s.t.\ x+y &= 4 \\ \mathcal{L} &= x^{\frac{1}{2}}y^{\frac{1}{2}} + \lambda(4-x-y) \\ x^* &= y^* = 2,\ \ \lambda^* = \frac{1}{2},\ z^* = 2 \end{aligned}$$

*Dual* problem:

$$\begin{aligned} \min\ k &= x+y \\ s.t.\ 2 &= x^{\frac{1}{2}}y^{\frac{1}{2}} \\ \mathcal{L}^D &= x+y+\lambda^D(2-x^{\frac{1}{2}}y^{\frac{1}{2}}) \\ x_D^* &= y_D^* = 2,\ \ \lambda_D^* = 2,\ z^* = 2,\ k = 4 \end{aligned}$$

Notice the value of the multipliers in the two problems.
Recall that in the primal problem:

$$\lambda_P = -\frac{f_i}{g_i}$$

In the dual problem we invert the two functions therefore the multiplier will be:

$$\lambda_D = -\frac{g_i}{f_i}$$

Therefore:

$$\lambda_P = \frac{1}{\lambda_D}$$

Why should we care about duality?

- cost minimization is the dual problem of profit maximization
- expenditure minimization is the dual problem of utility maximization

We will be relying on these duality relationships all semester. Furthermore the dual problem often has useful economic interpretation and so it may be more informative to solve and interpret than the primal problem.