Concave functions in economics

1. Preliminaries 1
2. Concave function of one variable 4
3. Concave function of more than one variable 7
4. Necessary and sufficient conditions for a maximum 10
5. When is a function concave 11
6. The gains to diversifying 15
7. Production plans and supporting prices 17
8. Constrained maximization 22

25 pages
Maximization with concave functions

Elsewhere module we have discussed necessary conditions for a maximum for the following problem:

\[ \max_x \{ f(x) \mid x \in \mathbb{R}^n \} \]

As long as the set of vectors satisfying the necessary conditions is small, it is in principle possible to solve by computing the value of \( f \) for each such vector and hence solve for the one that is the global maximizer.

With concave functions, solving maximization problems is so much easier. If you can find a vector satisfying the first order conditions for a maximum, then you have found the solution.

It is therefore very important to have a strong understanding of concave functions.

1. Preliminaries

Line through \( a \) in the direction of \( b \)

Consider the vector \( x = a + \theta b \). This is depicted below for three values of \( \theta \). As is clear from the figure, the graph of \( x(\theta) \) is a line in \( \mathbb{R}^2 \) through the vector \( a \) in the direction of \( b \).

Similarly, if \( a \) and \( b \in \mathbb{R}^3 \), then \( x(\theta) = a + \theta b \) is a line in \( \mathbb{R}^3 \). For higher dimensions we keep this same terminology.

![Fig.1-1: Line through \( a \) in the direction of \( b \)](image-url)
Line through $a^0$ and $a^1$

As argued above, $x(\theta) = a^0 + \theta(a^1 - a^0)$ is a line through $a^0$ in the direction of $a^1 - a^0$. Since $x(1) = a^1$, this line passes through $a^1$.

Convex combination of $a^0$ and $a^1$

$x(\theta) = a^0 + \theta(a^1 - a^0)$ where $0 < \theta < 1$

Note that this is a line segment in the direction of $a^1 - a^0$ with one boundary point $a^0$. Note that the other boundary point is $a^1$. 

Fig. 1.2: Line through $a^0$ and $a^1$

Fig. 1.3: Convex combinations of $a^0$ and $a^1$
The convex combinations of two vectors are most commonly written as follows:

$$x(\theta) = (1-\theta)a^0 + \theta a^1$$  where $0 < \theta < 1$.

While it is not general notation, I find it helpful to write a particular convex combination of the vectors $x^0$ and $x^1$ as follows:

$$x^\lambda = (1-\lambda)x^0 + \lambda x^1$$  where $0 < \lambda < 1$.

**Convex set**

A set $X$ of $n$-vectors is convex if, for every pair of vectors $x^0$ and $x^1$ that are in $X$, all convex combinations are also in $X$.

If $n = 1$ a convex set is an interval. It may of may not contain its boundary points. An interval with a lower boundary point $a$ and upper boundary point $b$ is written as $X = [a,b]$. Then $x \in X$ if and only if $a \leq x \leq b$. This called a closed interval. If the set contains neither of its boundary points it is written as $X = (a,b)$. Then $x \in X$ if and only if $a < x < b$. This is called an open interval.

Four examples of convex sets when $n = 2$ are depicted below. The star set is not convex since the convex combinations of any two neighboring vertexes are not in $X$.
2. Concave functions of one variable

Consider a function $f(x)$ with a graph as depicted below. Pick any two points $(x^0, y^0)$ and $(x^1, y^1)$ on the graph of the function. The dotted line is the set of convex combinations of these two points.

![Concave function](image)

**Figure 2.1: Concave function**

**Definition: Concave function**

The function $f$ is concave on $X$ if, for any $x^0, x^1 \in X$, all the convex combinations of these vectors lie below the graph of $f$. That is,

$$f(x^1) \geq (1 - \lambda) f(x^0) + \lambda f(x^1) \quad \text{for all } \lambda \in (0,1)$$

---

1 This figure was created in EXCEL. To download right click on (to be added) Take a look at Sheet1.
If you consider the definition, you should be able to convince yourself that a function must be continuous in order for it to be concave. However the definition makes no assumption about differentiability. Try drawing the graph of

$$f(x) = \begin{cases} 
2x, & x < 1 \\
1 + x, & x \geq 1
\end{cases}.$$ 

This is a concave function which is not differentiable at $x = 1$.

However, for maximization problems, assuming differentiability is very helpful, as it simplifies the characterization of the maximizer.

In the figure above, the line tangent to the graph of $f$ at $(x^0, f(x^0))$ is depicted. This is the line

$$y = f(x^0) + f'(x^0)(x - x^0).$$

Note that the graph of this line has the same value and gradient as $f(x)$ at $x^0$.

Intuitively if a function is concave and differentiable, then any such tangent line must lie above the graph of the function.
We have the following alternative definition.

**Definition: Concave function**

The differentiable function $f$ is concave on $X$ if, for any $x^0 \in X$, the tangent line through $(x^0, f(x^0))$ is above the graph of $f$. That is

$$f(x) \leq f(x^0) + f'(x^0)(x-x^0)$$

---

**Proof:** (for those who like proofs)

We show that the first definition implies the second definition\(^2\). From the first definition

$$f(x^1) - f(x^0) \geq \lambda(f(x^1) - f(x^0)) .$$

Therefore

$$\frac{(x^1-x^0)(f(x^1)-f(x^0))}{\lambda(x^1-x^0)} \geq f(x^1) - f(x^0) .$$

Also $x^1 = x^0 + h$ where $h = \lambda(x^1-x^0)$.

Substituting this into the inequality we can rewrite it as follows:

$$\frac{(x^1-x^0)f(x^0+h)-f(x^0)}{h} \geq f(x^1) - f(x^0) .$$

This holds for all $h > 0$. Since the function is differentiable, the limit of the ratio is the derivative at $x^0$. Therefore

$$(x^1-x^0)f'(x^0) \geq f(x^1) - f(x^0) .$$

QED

\(^2\) From the second definition, $f(x^0) - f(x^1) - f'(x^0)(x^0-x^1) \leq 0$ and $f(x^1) - f(x^0) - f'(x^1)(x^1-x^0) \leq 0$. Multiply the first inequality by $1-\lambda$ and the second by $\lambda$ to show that the second definition implies the first.
From Figure 2.2, it is intuitively clear that a differentiable function can only be concave if the slope of the function, \( f'(x) \), is decreasing. To see that this really must be the case, consider the following figures.

The shaded area under the graph of \( f''(x) \) is the integral of the derivative. Therefore the shaded area is

\[
f(x^1) - f(x^0) = \int_{x^0}^{x^1} f'(x)dx.
\]

Then the second definition can be rewritten as follows:

\[
\int_{x^0}^{x^1} f'(x)dx \leq f''(x^0)(x^1 - x^0).
\]  \hspace{1cm} (2.1)

The right-hand side of this inequality is the area of the rectangle marked with a heavy boundary. In the left-hand figure, where the slope is decreasing the rectangle is larger than the shaded area so (2.1) holds. This is not the case in the right-hand figure. Thus we have the third equivalent definition of a concave function.

**Definition: Concave function**

The differentiable function \( f : \mathbb{R} \to \mathbb{R} \) is concave on \( X \) if the derivative of the function \( f'(x) \) is decreasing on \( X \).
3. Concave functions of more than one variable

With more than one variable, the first definition of a concave function is exactly the same as in the one variable case except that the convex combinations are now combinations of two vectors.

**Definition: Concave function**

The function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is concave on \( X \) if, for any vectors \( x^0, x^1 \in X \subset \mathbb{R}^n \)

\[
f(x^\lambda) \geq (1-\lambda)f(x^0) + \lambda f(x^1)
\]

for every convex combination \( x^\lambda = (1-\lambda)x^0 + \lambda x^1 \), where \( \lambda \in (0,1) \)

For the two variable case we can illustrate using “surface diagrams.”  The left hand diagram in Figure 3.1 depicts the function

\[
l(x) = l(0) + a_1x_1 + a_2x_2
\]

![Fig. 3.1: Concave functions of two variables](image)
This is a plane. The gradient (slope of the plane) in the \( x_1 \) direction is \( a_1 \) and the gradient in the \( x_2 \) direction is \( a_2 \). The vector \( a = (a_1, a_2) \) is then called the gradient vector. From the definition it can be shown in a few steps that

\[
l(x^2) = (1 - \lambda)l(x^0) + \lambda l(x^1) .
\]

Therefore a linear function is concave.

Now consider the right hand figure. Consider any two points \((x^0, y^0)\) and \((x^1, y^1)\) on the graph of the function. Viewed from above, the surface is “bowed out”. Then all the convex combinations lie below the surface. Thus this function is also concave.

**Tangent plane at** \( x^0 \).

The gradient of the function \( f(x) \) in the direction of \( x_1 \) is the partial derivative \( \frac{\partial f}{\partial x_1} (x^0) \) and in the direction of \( x_2 \) is \( \frac{\partial f}{\partial x_2} (x^0) \). Consider the function

\[
l(x) = f(x^0) + \frac{\partial f}{\partial x_1} (x^0)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2} (x^0)(x_2 - x_2^0)
\]

Note that \( l(x^0) = f(x^0) \). Therefore this is a linear function which has the same value and the same gradient vector as \( f(x) \) at \( x^0 \). It is called the tangent plane at \( x^0 \). (In higher dimensions the function is called the tangent hyper-plane.) Intuitively, for a concave function the tangent plane must lie on or above the surface of the function.

Therefore we have the following second definition of a concave function.

**Definition: Concave function**

The differentiable function \( f \) is concave on \( X \) if, for any \( x^0 \in X \), the tangent (hyper-)plane \((x^0, f(x^0))\) is above the graph of \( f \). That is
The proof of the equivalence of the two definitions is not terribly difficult. It builds on the proof for $n = 1$. But what is most important at this stage is to have a good intuitive understanding of the different definitions and their equivalence.

4. **Necessary and sufficient conditions for a maximum**

For a concave function, maximization is especially easy. Find a vector $x^0$ satisfying the first order conditions (FOC) for a maximum. This is the global maximizer. That is, the FOC are both necessary and sufficient for a maximum.

**Functions of one variable**

Consider the following maximization problem, where $f$ is a differentiable concave function.

$$\max_x \{ f(x) \mid x \in \mathbb{R}_+ \}$$

For $x^0$ to be maximizing there are two possibilities. These are depicted below.

**Fig. 4.1: First Order conditions for a maximum**

If $x^0 > 0$ (as in the left-hand diagram), the gradient of the function at $x = x^0$ must be zero. For a solution $x^0 = 0$ on the boundary of the feasible set (as in the right-hand diagram), the gradient of the
function at $x = x^0$ can be zero or negative, i.e. $f'(x^0) \leq 0$. Since a feasible $x$ is positive and $x^0$ is zero, $x - x^0$ must be positive. It follows that in both cases

$$f''(x^0)(x - x^0) \leq 0$$

Appealing to the second definition for a concave function,

$$f(x) \leq f(x^0) + f'(x^0)(x - x^0) \leq f(x^0) \quad \text{for all } x \in \mathbb{R}^+$$

Thus the necessary condition is also sufficient.

**Functions of $n$ variables**

Consider the following maximization problem, where $f$ is a differentiable concave function.

$$\max_x \{ f(x) \mid x \in \mathbb{R}^n \}$$

With $n$ variables the First Order Conditions (FOC) are obtained by considering the maximization problem one variable at a time. Holding all other variables constant, for $x_i^0$ to be maximizing there are two possibilities just as in Figure 4.1. If $x_i^0 > 0$, the gradient of the function at $x_i = x_i^0$ must be zero. If $x_i^0 = 0$, the gradient of the function at $x_i = x_i^0$ cannot be strictly increasing, i.e. $\frac{\partial f}{\partial x_i}(x^0) \leq 0$. Since a feasible $x_i$ must be positive, it follows that $x_i - x_i^0$ is positive. Therefore in both cases

$$\frac{\partial f}{\partial x_i}(x^0)(x_i - x_i^0) \leq 0$$

Exactly the same argument holds for each variable. From the second definition of a concave function

$$f(x) \leq f(x^0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x^0)(x_j - x_j^0)$$

But we have just argued that each term in the summation must be negative. Therefore $f(x) \leq f(x^0)$

**5. When is a function concave?**
For the one variable case, checking whether a function is concave is easy. The gradient of the function must be decreasing. For two or more variables we can consider one variable at a time. From the one variable case we know that for each \( j \) the gradient \( \frac{\partial f}{\partial x_j} \) must be decreasing. But these \( n \) conditions are only necessary conditions for concavity.

Example: \( f(x) = 10 - 3x_1^2 + 10x_1x_2 - 3x_2^2 \)

This function is depicted below.

![Function graph](image)

Fig. 5.1: Saddle point

Note that \( \frac{\partial f}{\partial x_1} = -6x_1 + 10x_2 \) and \( \frac{\partial f}{\partial x_2} = 10x_1 - 6x_2 \). Therefore the gradient vector is zero \( x^0 = 0 \).

As is clear from the figure, \( \frac{\partial f}{\partial x_1}(x_1^0, x_2^0) \) and \( \frac{\partial f}{\partial x_2}(x_1^0, x_2^0) \) are both strictly decreasing. However \( x^0 \) is a saddle point rather than a maximum. To prove this consider \( x = (z, z) \). Then \( f(x_1, x_2) = f(z, z) = 10 + z^2 \).
To check for concavity the following three propositions are extremely helpful.\footnote{The first proposition follows directly from the first definition of a concave function. The proof of the second proposition is a little more complicated. The proof of the third requires some cunning.}

**Proposition 5.1:** A function is concave if it is the sum of concave functions.

**Proposition 5.2:** A function $h(x) = g(f(x))$ is concave if $f(x)$ is concave and $g(\cdot)$ is strictly increasing and concave.

**Proposition 5.3:** A function $f(x)$ is concave if (i) $g(\cdot)$ is increasing, (ii) $h(x) = g(f(x))$ is concave, (iii) $f(x)$ is homogeneous of degree 1, (i.e. $f(\theta x) = \theta f(x)$ for all $\theta > 0$).

As the following examples show, using these three propositions to prove that a function is concave can require some cunning. Unless you are doing a Ph. D level exam, you will not be asked to complete such proofs. If asked to prove concavity as an exercise there will be multiple hints.

**Example 1:** $U(x) = \sum_{j=1}^{n} \alpha_j u(x_j)$ where $u(x_j) = \ln x_j$ for all $x \gg 0$

Note that $u'(x_j) = \frac{1}{x_j}$ is decreasing so $U(x)$ is the sum of $n$ concave functions. Therefore $U(x)$ is concave.

**Example 2:** $U(x) = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$ where $\alpha \gg 0$ and $\sum_{j=1}^{n} \alpha_j = 1$.

We appeal to Proposition 5.3

$$U(\theta x) = (\theta x_1)^{\alpha_1} (\theta x_2)^{\alpha_2} \ldots (\theta x_n)^{\alpha_n} = \theta x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} = \theta U(x).$$
Therefore (iii) holds.

\[ h(x) = g(U(x)) = \ln U(x) = \sum_{j=1}^{n} \alpha_j \ln x_j. \]

This is concave (by Example 1) so (ii) holds. Finally \( g(\cdot) \) is increasing so (i) holds. Therefore \( U(x) \) is concave.

**Example 3:** \( f(x) = x_1^{\beta_1} x_2^{\beta_2} \) where \( \beta \gg 0 \) and \( \beta_1 + \beta_2 < 1 \).

Define \( \alpha_1 = \frac{\beta_1}{\theta} \) and \( \alpha_2 = \frac{\beta_2}{\theta} \) and choose \( \theta < 1 \) so that \( \alpha_1 + \alpha_2 = 1 \). Then

\[ f(x) = x_1^{\beta_1} x_2^{\beta_2} = x_1^{\alpha_1 \theta} x_2^{\alpha_2 \theta} = (x_1^{\alpha_1})^\theta (x_2^{\alpha_2})^\theta = (x_1^{\alpha_1} x_2^{\alpha_2})^\theta = U(x)^\theta \]

where \( U(x) \) is defined in Example 2. Note that \( U^\theta \) is a strictly increasing, concave function of \( U \), since \( 0 < \theta < 1 \). We know from Example 2 that \( U(x) \) is concave. Appealing to Proposition 5.2 it follows that \( f(x) \) is concave.

**Example 4:** \( U(x) = \left( \frac{1}{x_1} + \frac{1}{x_2} \right)^{-1} \).

\[ U(\theta x) = \left( \frac{1}{\theta x_1} + \frac{1}{\theta x_2} \right)^{-1} = \left( \frac{1}{\theta} \left( \frac{1}{x_1} + \frac{1}{x_2} \right) \right)^{-1} = \theta \left( \frac{1}{x_1} + \frac{1}{x_2} \right)^{-1} = \theta U(x) \]

Therefore \( U(x) \) is homogeneous of degree 1.

Consider

\[ h(x) = g(U(x)) \text{ where } g(U) = -U^{-1} \]

Note that \( g(U) \) is increasing and \( h(x) = g(U(x)) = -\frac{1}{x_1} - \frac{1}{x_2} \).

Appealing to Proposition 5.1, this is a concave function since it is the sum of concave functions.
6. The gains to diversifying

Suppose a consumer is indifferent between the two consumption bundles $x^0$ and $x^1$ depicted below. Let $U(x)$ be a “utility function” representing the consumer’s preferences. A “level set” of a function is the set of vectors yielding the same value of the function.\(^4\) One such level set is depicted below. It is the boundary of the shaded region. Both $x^0$ and $x^1$ are in this level set, indicating that the consumer’s utility is the same. The shaded region is the set of bundles for which $U(x) \geq U(x^0)$. This is called a “superlevel set”.

![Figure 6.1: Convex supersets and the value of diversifying](image)

In the figure, each of the two consumption bundles is highly concentrated in one commodity. In almost all such situations the consumer will strictly prefer any positively weighted average of the two bundles. With two commodities this is usually explained as follows. At $x^0$, since the consumption bundle is heavily concentrated in commodity 2, the consumer is willing to give up a relatively large quantity of commodity 2 in order to substitute

---

\(^4\) In the 2 commodity case economists call such a level set an indifference curve and we will sometimes do so as well.
more of commodity 1 into his consumption bundle. In the figure, this “marginal rate of substitution is the slope of the consumer’s level set at \( x^0 \).

Next consider \( x^1 \) where consumption bundle is highly concentrated in commodity 1. Now the consumer is only willing to give up a little of commodity 2 in order to substitute more of commodity 1 into his consumption bundle. This is called the law of diminishing marginal rates of substitution.

As should be clear from the figure, if a consumer’s preferences exhibit diminishing marginal rates of substitution, then for any two bundles \( x^0 \) and \( x^1 \) in the same level set, every weighted average or “convex combination”

\[
x^\lambda = (1-\lambda)x^0 + \lambda x^1,
\]

will be preferred. Note from the figure that this is equivalent to the assumption that the superlevel set \( S^+ = \{ x \mid U(x) \geq U(x^0) \} \) is convex.

When there are more than two commodities, we capture the preference for diversity by assuming that superlevel sets are convex.

This is a second reason why concave functions play such a central role in economics. As we shall see, the following proposition is very easy to prove.

**Proposition: The superlevel sets of a concave function are convex**

**Proof:** Consider any superlevel set \( S^+ = \{ x \mid U(x) \geq \bar{U} \} \). Consider two vectors \( x^0 \) and \( x^1 \) in \( S^+ \), that is (i) \( U(x^0) \geq \bar{U} \) and (ii) \( U(x^1) \geq \bar{U} \). From the definition of a concave function,

\[
U(x^\lambda) \geq (1-\lambda)U(x^0) + \lambda U(x^1)
\]

for all \( \lambda \in (0,1) \).

Appealing to (i) and (ii),

\[
(1-\lambda)U(x^0) + \lambda U(x^1) \geq (1-\lambda)\bar{U} + \lambda \bar{U} = \bar{U}.
\]

Therefore \( U(x^\lambda) \geq \bar{U} \) and so the convex combination lies in the superlevel set.
In the theory of the firm it is also usually assumed that there is value in diversifying.

We can reinterpret $U(x)$ as a “production” function. If a firm has input bundle $x$, then the maximum output achievable is $U(x)$. Typically the law of diminishing marginal rates of technical substitution holds for production. Arguing exactly as above, with two inputs it is equivalent to assume that the superlevel sets of the production function are convex. With more than two inputs it is almost always assumed that superlevel sets continue to be convex.

7. Production plans and supporting prices

Consider a firm that chooses each component of a vector of activities $x = (x_1, x_2, ..., x_n)$. This “activity vector” determines both the firm’s output $f(x)$ and the inputs required, $g(x)$.

Let $X$ be the set of feasible activity vectors. We shall assume that $X$ is convex.

In many economic applications each of the “activities” must be some positive number. Then $X = \mathbb{R}_+^n$.

If the firm wants to sell $q$ units of output it must choose $x$ so that

$$f(x) \geq q$$

The firm must also purchase $b$ units of the input where

$$b \geq g(x).$$

An input-output vector $(b,q)$ is called a feasible plan if $x \in X$, for some for some activity vector $x$. We can then define the set of feasible plans as follows:

$$Y = \{(b,q) \mid f(x) - q \geq 0, \ b - g(x) \geq 0, \ x \in X \}. \quad (7.1)$$
Technical remark on the convexity of $Y$

Note that if $f(x)$ is a concave function of $x$ then $f(x) - q$ is a convex function of $x$ and $q$ since this is the sum of concave functions. By the same argument, if $-g(x)$ is concave then $b - g(x)$ is a concave function of $x$ and $b$. From the previous section the superlevel sets of a concave function are concave. Therefore $Y$ is the intersection of convex sets.

From the definition of a convex set it follows directly that the intersection of convex sets is convex. It follows that the concavity of $f$ and $-g$ implies that the set of feasible plans is convex.

Figure 7.1 depicts a case in which $Y$ is convex.

![Figure 7.1: Maximizing output](image)

Suppose that the supply of inputs is fixed at $b^0$. Then the maximum feasible output is $q^0$ on the boundary of $Y$.

That is,

$$q^0 = \max_{x \in X} \{ q \mid f(x) - q \geq 0, \ b^0 - g(x) \geq 0 \} .$$
Let \( x^0 \) be an activity vector that solves this maximization problem. Note that if \( f(x) - q > 0 \) it is possible to increase output. Therefore we can replace the first inequality by the equality constraint \( f(x) = q \) and so

\[
q^0 = \max_{x \in X} \{ q \mid f(x) - q = 0, \ b^0 - g(x) \geq 0 \}.
\]

Equivalently, \( x^0 \) solves the following maximization problem.

\[
q^0 = \max_{x \in X} \{ f(x) \mid b^0 - g(x) \geq 0 \}
\]

This constrained maximization problem plays a central role for much of microeconomic theory. Therefore developing a deep understanding of the problem is critical.

We continue to interpret the problem as the output maximization problem of a firm. To give the problem a stronger economic flavor, pick an output price \( p > 0 \) and an input price \( r \geq 0 \). Then we can rewrite the output maximization problem as the following profit-maximization problem.

\[
\Pi^0 = \max_{x \in X} \{ pq - rb^0 \mid f(x) - q \geq 0, \ b^0 - g(x) \geq 0 \}.
\]

Since the input is fixed, this profit-maximization is really a revenue maximization problem.

**Relaxed problem**

Given an output price \( p > 0 \) and input price \( r \geq 0 \), suppose we now allow the firm to choose its input as well as its output. In Fig. 7.2, the set of feasible input-output vectors is now the shaded region and not just the vertical line.

The profit of the firm for any feasible plan is

\[
\Pi = pq - rb
\]

The level set of input-output vectors for any \( \Pi \) is the line

\[
pq - rb = \Pi, \text{ i.e. } q = \frac{r}{p} b + \frac{\Pi}{p}.
\]
The slope of a level set is therefore the price ratio \( \frac{r}{p} \).

If the input price is really high, so that the level set is very steep, the firm maximizes by purchasing \( b < b^0 \) units of the input. And if the input price is really low, the firm maximizes by purchasing \( b > b^0 \) units of the input.

But suppose, as depicted in Fig. 7-2, there is some input price such that the solution of the relaxed problem is \((b^0, q^0)\). That is, even with the option of buying a different quantity of the input the firm’s profit is maximized by choosing \( b^0 \) units. It then follows that the solution to the relaxed problem is also the solution to the original problem.

But can we always find such an input price? To see that this is not possible consider Fig. 7.3. As depicted, the input price is chosen so that the profit maximizing plans are \((b^0, q^0)\) and \((b^1, q^1)\). For any higher input price the profit-maximizing input-output vector is to the left of \((b^0, q^0)\). For any lower input price the profit-maximizing input-output vector is to the right of \((b^1, q^1)\). Therefore none of the boundary points with \( b \in (b^0, b^1) \) are profit-maximizing for any input price.
The problem is that the set of feasible plans is not convex. While both \((b^0, q^0)\) and \((b^1, q^1)\) are feasible, none of the convex combinations are in the set of feasible plans.

However if \(Y\) is convex, it is intuitively clear from Fig. 7.2 that for any point \((b^0, q^0)\) on the boundary, there is a line through this point such that all the points in \(Y\) lie below this line. Let

\[ pq - rb = pq^0 - rb^0 = \Pi^0 \]

be the equation of this line. Then

\[ \Pi^0 = pq^0 - rb^0 = \text{Max}_{(b^0, q^0) \in Y} \{ pq - rb \} \]

The crucial observation is that if \(Y\) is convex, then for every input-output vector \(y^0 = (b^0, q^0)\) on the boundary of \(Y\) there is a vector of prices \((p, r)\) for which \(y^0\) is profit-maximizing. Therefore we can solve the constrained maximization by considering the relaxed maximization problem.
In fact this is true, whether there as a single input \( b \) or a vector of inputs.

This result is summarized in the following proposition.

**Proposition: Supporting prices**

For a convex set \( X \), let \( Y = \{(b, q) \mid f(x) - q \geq 0, \ b_i - g_i(x) \geq 0, \ i = 1,\ldots,m, \ x \in X\} \) be the set of feasible plans.

For any \( b^0 \), the maximum feasible output \( q^0 \) is the solution to the following maximization problem.

\[
q^0 = \max_{x \in X} \{q \mid f(x) - q \geq 0, \ h_i(x) \equiv b^0_i - g_i(x) \geq 0, \ i = 1,\ldots,m\}.
\]  

(7.3)

If \( Y \) is convex (as is the case when \( f \) and \( h_i, \ i = 1,\ldots,m \) are concave functions of \( x \)) then there is supporting price vector \((p, r)\) where \( p > 0 \) and \( r \geq 0 \) such that \((b^0, q^0)\) is profit-maximizing.

**8. Constrained maximization**

We now consider the following constrained maximization problem

\[
q^0 = \max_{x \in X} \{f(x) \mid h_i(x) \equiv b^0_i - g_i(x) \geq 0, \ i = 1,\ldots,m\}
\]  

(8.1)

From the supporting prices proposition, if the problem is concave (i.e. \( f(x) \) and \( h_i(x), \ i = 1,\ldots,m \) are concave functions of \( x \)), and \( q^0 \) is the solution to this maximization problem, then there exists prices such that

\[
pq^0 - rb^0 = \max_{x \in X} \{pq - \sum_{i=1}^m rb_i \mid f(x) \geq q, \ g_i(x) \geq b_i, \ i = 1,\ldots,m\}
\]

Since \( p > 0 \), profit can be increased if \( q < f(x) \). Therefore \( f(x) = q \) and we can rewrite the maximization problem as follows:
John Riley

\[ pq^0 - rb^0 = \max_{x \in X} \left\{ pf(x) - \sum_{i=1}^{m} r_i b_i \mid g_i(x) \geq b_i, \ i = 1, \ldots, m \right\} \]

If \( r_i > 0 \), profit can be increased by decreasing \( b_i \) if \( g_i(x^0) < b_i \). Then in such cases

\[ r^0 = r_i g_i(x^0) \]

If \( r_i = 0 \) this equality again holds. Therefore we can rewrite the maximization problem as follows:

\[ pq^0 - rb^0 = \max_{x \in X} \left\{ pf(x) - \sum_{i=1}^{m} r_i g_i(x) \right\} \quad (8.2) \]

Define \( r_i / p = \lambda_i \). Then

\[ pq^0 - rb^0 = p \max_{x \in X} \left\{ f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) \right\} \]

We have therefore shown that the solution to the constrained maximization problem must also be the solution to an unconstrained maximization problem in which each of the inputs has a “shadow price”.

The FOC for \( x^0 \) to be a solution to this unconstrained maximization problem has already been examined in section 4. There is a simple way of remembering these conditions.

Step 1: Write down each constraint as follows: \( h_i(x) \equiv b_i - g_i(x) \geq 0 \).

Step 2: Write down the Lagrangian

\[ L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) . \]

Step 3: Write down the gradient of this function

\[ \frac{\partial L}{\partial x_j}(x, \lambda) = \frac{\partial f}{\partial x_j}(x) + \sum_{i=1}^{m} \lambda_i \frac{\partial h_i}{\partial x_j}(x) . \]

Step 4: FOC (from section 4):
If \( \lambda_{j} > 0 \) then \( \frac{\partial L}{\partial \lambda_{j}} (x^{0}, \lambda) = 0 \). If \( \lambda_{j} = 0 \) then \( \frac{\partial L}{\partial \lambda_{j}} (x^{0}, \lambda) \leq 0 \).

Example: Consumer choice

A consumer has utility function \( U(x) = \sum_{j=1}^{n} \alpha_{j} \ln x_{j} \). The price vector is \( p \gg 0 \) and her income is \( I \). She maximizes her utility subject to her budget constraint, that is, she solves the following problem.

\[
\text{Max} \{ U(x) = \sum_{j=1}^{n} \alpha_{j} \ln x_{j} | h(x) = I - \sum_{j=1}^{n} p_{j} x_{j} \geq 0 \}
\]

Since \( \ln x \) is concave the utility function is concave as it is the sum of concave functions. The constraint function \( h(x) \) is also concave because it is the sum of linear (and hence concave) functions. Thus if we can find \( x^{0} \) satisfying the FOC, it is a maximizer.

Suppose that \( x^{0} \gg 0 \) is a solution. Note first that since utility is increasing, the budget will be satisfied with equality at the maximum \( \left( \sum_{j=1}^{n} p_{j} x_{j}^{0} = I \right) \). The Lagrangian is

\[
L = \sum_{j=1}^{n} \alpha_{j} \ln x_{j} + \lambda (I - \sum_{j=1}^{n} p_{j} x_{j}^{0})
\]

At \( x^{0} \) the FOC are

\[
\frac{\partial L}{\partial \lambda_{j}} (x^{0}, \lambda) = \frac{\alpha_{j}}{x_{j}^{0}} - \lambda p_{j} = 0 , \ j = 1,\ldots, n.
\]

Therefore

\[
\alpha_{j} = \lambda p_{j} x_{j}^{0} , \ j = 1,\ldots, n.
\] (8.3)

Summing over the commodities,
\[
\sum_{j=1}^{n} \alpha_j = \lambda \sum_{j=1}^{n} p_j x_j^0 = \lambda I .
\]

Therefore \( \lambda = \frac{1}{I} \sum_{j=1}^{n} \alpha_j \).

Appealing to (8.3),

\[
x_j^0 = \frac{\alpha_j}{\lambda p_j} = \frac{\alpha_j}{\sum_{j=1}^{n} \alpha_j} \frac{I}{p_j} .
\]