20. Homogeneous and Homothetic Functions

11/10/20

Homogeneous and homothetic functions are closely related, but are used in different ways in economics. We start with a look at homogeneity when the numerical values themselves matter.

This happens with production functions. You should be familiar with the idea of returns to scale. Let's focus on constant returns to scale. A production function exhibits *constant returns to scale* if, whenever we change all inputs in a given proportion, output changes by the same proportion.

In equations, let $f\colon \mathbb{R}^m_+ \to \mathbb{R}$ be a production function. It exhibits *constant returns to scale* if

$$f(t\mathbf{x}) = tf(\mathbf{x})$$

for every $\mathbf{x} \in \mathbb{R}^m_+$ and $t > 0$.

20.1 Homogeneous Functions

Homogeneous Function. On $\mathbb{R}^m_+$, a real-valued function is *homogeneous of degree* γ if

$$f(t\mathbf{x}) = t^\gamma f(\mathbf{x})$$

for every $\mathbf{x} \in \mathbb{R}^m_+$ and $t > 0$.

The degree of homogeneity can be negative, and need not be an integer. Constant returns to scale functions are homogeneous of degree one. If $\gamma > 1$, homogeneous functions of degree γ have increasing returns to scale, and if $0 < \gamma < 1$, homogeneous functions of degree γ have decreasing returns to scale.

► **Example 20.1.1: Cobb-Douglas Production.** Consider the Cobb-Douglas production function defined on $\mathbb{R}^2_+$ by $f(x, y) = Ax^\alpha y^\beta$ with $A, \alpha, \beta > 0$.

Then

$$f(tx, ty) = A(tx)^\alpha (ty)^\beta = t^{\alpha+\beta} Ax^\alpha y^\beta = t^{\alpha+\beta} f(x, y)$$

showing that f is homogeneous of degree $\alpha + \beta$. If $\alpha + \beta < 1$, f has decreasing returns to scale, if $\alpha + \beta = 1$, f has constant returns to scale, and if $\alpha + \beta > 1$, f has increasing returns to scale. ◄

20.2 Examples of Homogeneous Functions

Homogeneous functions arise in both consumer's and producer's optimization problems. The cost, expenditure, and profit functions are homogeneous of degree one in prices. Indirect utility is homogeneous of degree zero in prices and income. Further, homogeneous production and utility functions are often used in empirical work.

The Cobb-Douglas functions $u(\mathbf{x}) = A\prod_{i=1}^{m} x_i^{\gamma_i}$ with $\gamma_i > 0$ are homogeneous of degree $\sum_i \gamma_i$ on $\mathbb{R}^m_+$. The constant elasticity of substitution function $f(\mathbf{x}) = [\delta x_1^{-\rho} + (1-\delta)x_2^{-\rho}]^{\nu/\rho}$ for $\nu > 0$ and $\rho > -1$, $\rho \neq 0$ is homogeneous of degree ν. The Leontief function $u(\mathbf{x}) = \min x_i$ is homogeneous of degree 1.

Any monomial,

$$a\prod_{i=1}^{m} x_i^{\gamma_i}$$

is homogeneous of degree $\gamma = \sum_{i=1}^{m} \gamma_i$ on $\mathbb{R}^m_+$. A sum of monomials of degree γ is homogeneous of degree γ, the sum of monomials of differing degrees is not homogeneous.

Additional examples of homogeneous functions include:

$$\mathbf{p}\cdot\mathbf{x}, \quad x^{3/2} + 3x^{1/2}y, \quad \sqrt{x+y+z}$$
$$x^8 + 5x^4y^4 + 10x^3y^5 + 7xy^7 + 13y^8, \quad Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}.$$

The following are not homogeneous:

$$\sin xyz, \quad x^2 + y^3, \quad 3xyz + 3x^2z - 3xy$$
$$\frac{xy + yz + xz}{x + z^2}, \quad \exp(x^2 + y^2).$$

Most quasi-linear utility functions, such as $u(\mathbf{x}) = x_1 + x_2^{1/2}$ are not homogeneous of any degree. The linear term means that they can only be homogeneous of degree one, meaning that the function can only be homogeneous if the non-linear term is also homogeneous of degree one. Non-linear cases that **are** homogeneous of degree one require at least three goods. One example is

$$x_1 + (x_2x_3)1/2$$

which is homogeneous of degree one.

20.3 Cones and Homogeneity

Homogeneous functions require us to know $f(t\mathbf{x})$ when we know $f(\mathbf{x})$. That means that $t\mathbf{x} \in \operatorname{dom} f$ whenever $\mathbf{x} \in \operatorname{dom} f$. Sets with that property are called cones, and cones are the natural domain of homogeneous functions.

Cone. Let V be a vector space. A set $C \subset V$ is a *cone* if for every $\mathbf{x} \in C$ and $t > 0$, $t\mathbf{x} \in C$. Equivalently, C is a *cone* if $tC \subset C$ for all $t > 0$.

Cones include the positive orthant $\mathbb{R}^m_+$, the strictly positive orthant $\mathbb{R}^m_{++}$, any vector subspace of $\mathbb{R}^m$, any ray in $\mathbb{R}^m$ and any set of non-negative linear combinations of a collection of vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_i\}$. The set $\{(x, y) : x, y \geq 0, y \leq x\}$ is as example of the last as it can also be written $\{\mathbf{x} = t_1(1, 0) + t_2(1, 1) : t_1, t_2 \geq 0\}$.

We can redefine homogeneity to apply to functions defined on any cone.

Homogeneous Function. Let C be a cone in a vector space V. A function $f\colon C \to \mathbb{R}$ is *homogeneous of degree* γ if

$$f(t\mathbf{x}) = t^\gamma f(\mathbf{x})$$

for every $\mathbf{x} \in \mathbb{R}^m$ and $t > 0$.

Restricting the domain of a homogeneous function so that it is not all of $\mathbb{R}^m$ allows us to expand the notation of homogeneous functions to negative degrees by avoiding division by zero.

E.g., $1/\|\mathbf{x}\|_2$ is homogeneous of degree -1 on the cone $\mathbb{R}^m_{++}$. Restricting the domain also allows us to consider $f(\mathbf{x})/g(\mathbf{x})$ where f is homogeneous of degree γ_1 and g is homogeneous of degree γ_2, both on $\mathbb{R}^m_{++}$. The quotient is homogeneous of degree $\gamma_1 - \gamma_2$ on $\mathbb{R}^m_{++}$.

20.4 Derivatives of Homogeneous Functions

When a function is homogeneous, homogeneity is inherited by the derivative, but with the degree reduced by one.

Theorem 20.4.1. *Let* $C \subset \mathbb{R}^m$ *be an open cone and* $f\colon C \to \mathbb{R}$ *is* $\mathcal{C}^1$ *and homogeneous of degree* γ. *Then the derivative* Df *is homogeneous of degree* $(\gamma - 1)$.

Proof. By homogeneity, $f(t\mathbf{x}) = t^\gamma f(\mathbf{x})$. Taking the $\mathbf{x}$ derivative of that equation, we obtain $t\,Df(t\mathbf{x}) = t^\gamma\,Df(\mathbf{x})$. Dividing by t, we find that $Df(t\mathbf{x}) = t^{(\gamma-1)}\,Df(\mathbf{x})$. ∎

We assumed C was open so that Df could be defined at all points of C.

The converse fails. If Df is homogeneous of degree β, we cannot conclude that f is homogeneous of degree $(\beta + 1)$. For example, let $m = 2$ and consider $f(\mathbf{x}) = 1 + x_1 x_2$, which is not homogeneous of any degree.

A quick calculation shows that $(D_\mathbf{x} f)(\mathbf{x}) = (x_2, x_1)$ which is homogeneous of degree one, even though f is not homogeneous. Fortunately, addition of a constant is the main thing that goes wrong with the converse when Df is homogeneous of degree β for $\beta \neq -1$.

The case $\beta = -1$ can suffer from two other types of complications. The first involves logarithmic functions. Suppose $f(\mathbf{x}) = b \ln \phi(\mathbf{x})$ where ϕ is homogeneous of degree one with $\phi > 0$. Then

$$Df = \frac{b}{\phi(\mathbf{x})} D\phi(\mathbf{x}),$$

which is homogeneous of degree minus one.

The $\beta = -1$ case has a second type of complication when there is more than one variable. In $\mathbb{R}^m$ with $m > 1$, functions can be homogeneous of degree zero without being constant. One such function is $g(\mathbf{x}) = x_1/(x_1 + x_2)$. Its derivative is

$$Dg = \frac{1}{(x_1 + x_2)^2} \left(x_2, -x_1\right),$$

which is clearly homogeneous of degree minus one.

20.5 Homogeneity and Indifference Curves

The fact that derivatives of homogeneous functions are also homogeneous has consequences for the shape of indifference curves and isoquants. It implies that the marginal rate of substitution and marginal rate of technical substitution are constant along rays through the origin. That is, they are homogeneous of degree zero.

To see this let u be homogeneous of degree γ and consider $MRS_{ij} = (\partial u/\partial x_i)\big/(\partial u/\partial x_j)$. We calculate

$$MRS_{ij}(t\mathbf{x}) = \frac{\frac{\partial u}{\partial x_i}(t\mathbf{x})}{\frac{\partial u}{\partial x_j}(t\mathbf{x})} = \frac{t^{\gamma-1}\frac{\partial u}{\partial x_i}(\mathbf{x})}{t^{\gamma-1}\frac{\partial u}{\partial x_j}(\mathbf{x})} = \frac{\frac{\partial u}{\partial x_i}(\mathbf{x})}{\frac{\partial u}{\partial x_j}(\mathbf{x})} = MRS_{ij}(\mathbf{x}).$$

Of course, the calculation for the marginal rate of technical substitution is essentially the same.

Consequences of this include that facts that income expansion paths and scale expansion paths are rays through the origin whenever the original production or utility function is homogeneous.

20.6 Euler's Theorem

The second important property of homogeneous functions is given by Euler's Theorem.

Euler's Theorem. *Let* $f\colon \mathbb{R}^m_{++} \to \mathbb{R}$ *be* $\mathcal{C}^1$. *Then* f *is homogeneous of degree* γ *if and only if* $\big[D_{\mathbf{x}} f(\mathbf{x})\big]\mathbf{x} = \gamma f(\mathbf{x})$, *that is*

$$\sum_{i=1}^{m} x_i \frac{\partial f}{\partial x_i}(\mathbf{x}) = \gamma f(\mathbf{x}).$$

Proof. Define $\varphi(t) = f(t\mathbf{x})$. By the chain rule, $d\varphi/dt = \big[Df(t\mathbf{x})\big]\mathbf{x}$.

If f is homogeneous of degree γ, we can also write $\varphi(t) = t^\gamma f(\mathbf{x})$ so that $d\varphi/dt = \gamma t^{\gamma-1} f(\mathbf{x})$. Setting $t = 1$ we obtain $(Df)\cdot\mathbf{x} = \gamma f(\mathbf{x})$.

Conversely, if $\big[D_{\mathbf{x}} f(\mathbf{x})\big]\mathbf{x} = \gamma f(\mathbf{x})$,

$$\frac{d\varphi}{dt} = \big[D_{\mathbf{x}} f(t\mathbf{x})\big]\mathbf{x} = \frac{1}{t}\big[D_{\mathbf{x}} f(t\mathbf{x})\big](t\mathbf{x}) = \frac{\gamma f(t\mathbf{x})}{t} = \frac{\gamma\varphi(t)}{t}.$$

It follows that

$$\frac{d(\ln\varphi)}{dt} = \frac{1}{\varphi}\frac{d\varphi}{dt} = \frac{\gamma}{t}.$$

Integrating from $t = 1$ to t, we obtain $\ln\varphi(t) - \ln\varphi(1) = \gamma\ln t$. Taking the exponential yields $\varphi(t) = \varphi(1)t^\gamma$. This can be rewritten $f(t\mathbf{x}) = t^\gamma f(\mathbf{x})$, showing that f is homogeneous of degree γ. ∎

Gradients and Euler's Theorem. Euler's Theorem is sometimes written using the gradient, $\nabla f = Df^{\mathsf{T}}$, which is a column vector. In that case, the condition can be written $\mathbf{x}\cdot\nabla_{\mathbf{x}} f = \gamma f(\mathbf{x})$. Either way, it means that

$$\sum_i x_i \frac{\partial f}{\partial x_i} = \gamma f(\mathbf{x}).$$

20.7 Wicksteed's Theorem

Suppose f is a constant returns to scale production function. If the firm is a price-taker in both input and output markets, profits are $pf(\boldsymbol{x}) - \boldsymbol{w}\cdot\boldsymbol{x}$ where p is the output price and $\boldsymbol{w}$ the vector of factor prices.

To maximize profit we set its derivative equal to zero.

$$pDf(\boldsymbol{x}^*) = \boldsymbol{w}^T.$$

This tells us that for each good i, the value of marginal profit must be equal to the factor price:

$$p\frac{\partial f}{\partial x_i} = w_i.$$

What about profits? Euler's Theorem provides the answer. By Euler's Theorem,

$$f(\boldsymbol{x}) = \sum_i x_i \frac{\partial f}{\partial x_i}.$$

We multiply by p to obtain

$$pf(\boldsymbol{x}) = \sum_i x_i p \frac{\partial f}{\partial x_i}.$$

Then substitute the first order conditions to find

$$pf(\boldsymbol{x}) = \sum_i x_i w_i = \boldsymbol{w}\cdot\boldsymbol{x},$$

showing that revenue equals cost and that profit is zero.

If the production function was homogeneous of degree γ, $0 < \gamma < 1$, the same argument would show profit is positive, while if $\gamma > 1$, profit would be negative.

20.8 Homogeneity and Monotonicity

We define three types of monotonicity for real-valued functions of m variables.

Monotonicity. A real-valued function f defined on a subset of $\mathbb{R}^m$ is *monotonic* if $f(\mathbf{x}) \geq f(\mathbf{y})$ whenever $\mathbf{x} \geq \mathbf{y}$. It is *strictly monotonic* is $f(\mathbf{x}) > f(\mathbf{y})$ whenever $\mathbf{x} \gg \mathbf{y}$. Finally, it is *strongly monotonic* if $f(\mathbf{x}) > f(\mathbf{y})$ whenever $\mathbf{x} > \mathbf{y}$.

For functions on $\mathbb{R}$, strong and strict monotonicity are the same.

We saw earlier that the homogeneous of degree zero function $g(\mathbf{x}) = x_1/(x_1 + x_2)$ is not monotonic. This is normal for such functions. Indeed, Euler's Theorem can be used to show that functions that are homogeneous of degree zero cannot be monotonic when there are two or more variables.

Theorem 20.8.1. *Any function $f \in \mathcal{C}^1(\mathbb{R}^m_{++})$ for $m > 1$ that is homogeneous of degree zero is not monotonic.*

Proof. We prove this by applying Euler's Theorem,

$$0 = \sum_i x_i \frac{\partial f}{\partial x_i}.$$

If every $\partial f/\partial x_i > 0$, the right-hand side is positive. This is impossible as it must be zero, so there has to be j with $\partial f/\partial x_j < 0$,. The function f cannot be monotonic. ∎

However, it is possible to combine a logarithmic form with a homogeneous of degree zero form to get an increasing function with a derivative that is homogeneous of degree minus one. The function

$$f(x_1, x_2) = \frac{x_1}{x_1 + x_2} + \ln(x_1 + x_2)$$

is such a case.

20.9 Cardinal Functions vs. Ordinal Functions

The terms *cardinal* and *ordinal* are often used when discussing utility functions. However, they are not properties of the functions themselves, but of our interpretation of them. When we say a production function is cardinal, we mean that the output numbers have direct meaning. Two tons of steel and three tons of steel are different things.

In contrast, with utility functions what is important is not the particular level of utility, but rather comparisons of utility levels—which is more and which is less.

To that end, we say that utility functions $u, v\colon \mathfrak{X} \to \mathbb{R}$ are *equivalent utility functions* if $u(\mathbf{x}) \geq u(\mathbf{y})$ if and only if $v(\mathbf{x}) \geq v(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$.

Theorem 20.9.1. *Equivalence between utility functions from $\mathfrak{X} \to \mathbb{R}$ is an equivalence relation.*

Proof. The relation is reflexive. A function $u\colon \mathfrak{X} \to \mathbb{R}$ is equivalent to itself because $u(\mathbf{x}) \geq u(\mathbf{y})$ if and only if $u(\mathbf{x}) \geq u(\mathbf{y})$. Yes, that was pretty trivial.

The relation is symmetric. If u is equivalent to v, then v is trivially equivalent to u. Here $u(\mathbf{x}) \geq u(\mathbf{y})$ if and only if $v(\mathbf{x}) \geq v(\mathbf{y})$, so $v(\mathbf{x}) \geq v(\mathbf{y})$ if and only if $u(\mathbf{x}) \geq u(\mathbf{y})$.

Finally, the relation is transitive. This is barely harder to show. If u is equivalent to v and v is equivalent to a function w then $u(\mathbf{x}) \geq u(\mathbf{y})$ if and only if $v(\mathbf{x}) \geq v(\mathbf{y})$ if and only if $w(\mathbf{x}) \geq w(\mathbf{y})$, so $u(\mathbf{x}) \geq u(\mathbf{y})$ if and only if $w(\mathbf{x}) \geq w(\mathbf{y})$, showing that u is equivalent to w. This proves transitivity. ∎

A property of a function $f\colon \mathfrak{X} \to \mathbb{R}$ is *ordinal* if it shared by all equivalent utility functions.

20.10 Two Types of Equivalence

You'll notice that this treatment is slightly different from that in Simon and Blume (section 20.3) with an inequality rather than an equals sign. The reason is that if we used equality, as they do, the utility functions $u(x, y) = x + y$ and $v(x, y) = -x - y$ would be equivalent functions. However, as utility functions, they are completely different. For u, more is better, for v, more is worse. Our definition maintains the sense of utility.

If two utility functions are equivalent in our sense, they are equivalent in the Simon and Blume sense. As the above example shows, the converse is false.

Theorem 20.10.1. *If u and v are equivalent utility functions on $\mathfrak{X}$, then $u(\mathbf{x}) = u(\mathbf{y})$ if and only if $v(\mathbf{x}) = v(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$.*

Proof. Now $u(\mathbf{x}) = u(\mathbf{y})$ if and only if both $u(\mathbf{x}) \geq u(\mathbf{y})$ and $u(\mathbf{y}) \geq u(\mathbf{x})$. By equivalence, that holds if and only if both $v(\mathbf{x}) \geq v(\mathbf{y})$ and $v(\mathbf{y}) \geq v(\mathbf{x})$. The last pair of statements hold if and only if $v(\mathbf{x}) = v(\mathbf{y})$. ∎

20.11 Monotonic Transformations

Monotonic transformations will allow us to characterize all equivalent utility functions.

Monotonic Transformation. Let I be an interval in the real line. A function $\phi\colon I \to \mathbb{R}$ is a *monotonic transformation of* I if ϕ is strictly increasing on I. Moreover, if f is a real-valued function of m variables and ϕ is a monotonic transformation on its image, we say

$$\phi \circ f\colon \mathbf{x} \mapsto \phi\big(f(\mathbf{x})\big)$$

is a monotonic transformation of f.

If ϕ is differentiable with $\phi' > 0$, then ϕ is a monotonic transformation.

20.12 Monotonic Transformations and Utility Equivalence

One important result is that two utility functions are equivalent if and only if there are monotonic transformations of each into the other.

Theorem 20.12.1. *Let $u, v\colon \mathfrak{X} \to \mathbb{R}$. Then u and v are equivalent utility functions on $\mathfrak{X}$ if and only if there is a monotonic transformation ϕ with $u = \phi \circ v$.*

Proof. If: Suppose $u = \phi \circ v$. Because ϕ is strictly increasing, $v(\mathbf{x}) \geq v(\mathbf{y})$ if and only if $u(\mathbf{x}) = \phi\big(v(\mathbf{x})\big) \geq \phi\big(v(\mathbf{y})\big) = u(\mathbf{x})$, showing that the two functions are equivalent.

Only If: Suppose u and v are equivalent. Let $\bar{u} \in \operatorname{ran} u$. There is $\mathbf{x} \in \mathfrak{X}$ with $\bar{u} = u(\mathbf{x})$. Define $\phi(\bar{u}) = v(\mathbf{x})$.

The function v is well-defined.[1] If we had picked any other $\mathbf{y}$ with $u(\mathbf{y}) = \bar{u}$, we would have $v(\mathbf{x}) = v(\mathbf{y})$ by Theorem 20.10.1. This shows that the definition of $\phi(\bar{u})$ depends only on $\bar{u}$, not on our choice of $\mathbf{x}$ with $u(\mathbf{x}) = \bar{u}$.

Now $v(\mathbf{x}) = \phi\big(u(\mathbf{x})\big)$ for all $\mathbf{x} \in \mathfrak{X}$. We need only show that ϕ is increasing. Suppose $\bar{u}_0 < \bar{u}_1$. Take $\mathbf{x}_i$ with $u(\mathbf{x}_i) = \bar{u}_i$, $i = 0, 1$. Then $u(\mathbf{x}_1) > u(\mathbf{x}_0)$. By the definition of equivalence, $v(\mathbf{x}_1) \geq v(\mathbf{x}_0)$. By Theorem 20.10.1, $v(\mathbf{x}_1) = v(\mathbf{x}_0)$ if and only if $u(\mathbf{x}_0) = u(\mathbf{x}_1)$. Since $u(\mathbf{x}_0) \neq u(\mathbf{x}_1)$, $v(\mathbf{x}_0) \neq v(\mathbf{x}_1)$, showing that $v(\mathbf{x}_1) > v(\mathbf{x}_0)$. Then $\phi(\bar{u}_1) = v(\mathbf{x}_1) > v(\mathbf{x}_0) = \phi(\bar{u}_0)$. This shows that ϕ is increasing and completes the proof. ∎

Thus on $\mathbb{R}_+$, the functions x, x^2, x^7, e^x, and $\ln x$ are all monotonic transformations of $f(x) = x$, and so all equivalent. Similarly, xy, $xy + (xy)^4, \log xy$, and $(xy)^{1/2}$ are all equivalent on $\mathbb{R}^2_+$.

When utility is differential, one property preserved by monotonic differentiable transformations is the marginal rate of substitution. Suppose $v(\mathbf{x}) = \phi\big(u(\mathbf{x})\big)$. Then by the Chain Rule,

$$\mathrm{MRS}^v_{ij} = \frac{\frac{\partial v}{\partial x_i}}{\frac{\partial v}{\partial x_j}} = \frac{\phi' \frac{\partial u}{\partial x_i}}{\phi' \frac{\partial u}{\partial x_j}} = \frac{\frac{\partial u}{\partial x_i}}{\frac{\partial u}{\partial x_j}} = \mathrm{MRS}^u_{ij}\,.$$

[1] "Well-defined" is a mathematics term meaning that the definition is unambiguous.

20.13 Homothetic Functions

Homothetic functions are the ordinal equivalent of homogeneous functions.

Homothetic Function. Let C be a cone. A function $f\colon C \to \mathbb{R}$ is *homothetic* if for every $\mathbf{x}, \mathbf{y} \in C$ and $t > 0$, $f(\mathbf{x}) \geq f(\mathbf{y})$ if and only if $f(t\mathbf{x}) \geq f(t\mathbf{y})$.

One consequence of the definition of homotheticity is that f is equivalent to g defined by $g(\mathbf{x}) = f(t\mathbf{x})$.

Any homogeneous utility function is also homothetic. And since homotheticity is an ordinal property, any increasing transformation of a homogeneous utility function is homothetic too.

However, not all homothetic preferences have a homogeneous utility representation. Lexicographic preferences are homothetic, but cannot be represented by any utility function—homogeneous or otherwise. If we require that preferences are also monotonic and continuous, we can represent homothetic preferences by a homogeneous utility function.

For a utility function, homotheticity means that preferences are invariant under scalar multiplication in the sense that the set of indifference curves is unchanged when all consumption bundles are multiplied by the same positive number. More precisely, preferences are invariant under homothetic transformations centered on the origin.

Homothetic preferences include commonly used functional forms such as Cobb-Douglas utility and constant elasticity of substitution utility.

One special type of homothetic utility is homogeneous utility, where multiplying the consumption bundle by a scalar multiplies utility by some power of that scalar. The Homothetic Representation Theorem will show that monotonic continuous and homothetic preferences can be represented by a homogeneous utility function.

Homotheticity in economics is based on comparing positive scalar multiples of vectors. By restricting our attention to consumption sets that are cones, we ensure that scalar multiplication is always possible. Such scaling preserves the shapes of objects, including indifference surfaces. It only changes their scale.

20.14 Homothetic and Homogeneous Functions

One remarkable fact is that every continuous homogeneous utility functions is a monotonic transformation of a homogeneous utility function.

The key part of the theorem is separated in a lemma, as it is of general use.

Lemma 20.14.1. *Suppose* $\mathbf{x}^* \in \mathbb{R}^m_{++}$ *and* u *is continuous and increasing on* $\mathbb{R}^m_+$. *If* $\mathbf{x} \in \mathbb{R}^m_+$, *there is a unique* $\alpha \geq 0$ *with* $u(\alpha\mathbf{x}^*) = u(\mathbf{x})$.

Proof. Let $\mathbf{x} \in \mathbb{R}^m_+$. Choose $\bar{\alpha}$ such that $\bar{\alpha}\mathbf{x}^* \gg \mathbf{x} \geq \mathbf{0}$ and define $A^+ = \{\alpha \in [0, \bar{\alpha}] : u(\alpha\mathbf{x}^*) \geq u(\mathbf{x})\}$ and $A^- = \{\alpha \in [0, \bar{\alpha}] : u(\mathbf{x}) \geq u(\alpha\mathbf{x}^*)\}$. By continuity, both A^+ and A^- are closed sets. Further, they obey $A^+ \cup A^- \supset [0, \bar{\alpha}]$. Now $0 \in A^-$ and $\bar{\alpha} \in A^+$, so both are non-empty. Since intervals are connected, $A^+ \cap A^-$ is non-empty. By construction, $u(\mathbf{x}) = u(\alpha\mathbf{x}^*)$ whenever $\alpha \in (A^+ \cap A^-)$. Transitivity and the fact that u is increasing imply there is only one such α. ∎

20.15 Homothetic Representation Theorem

We are now ready to prove the Homothetic Representation Theorem.

Homothetic Representation Theorem. *Suppose* $u\colon \mathbb{R}^m_+ \to \mathbb{R}$ *is increasing, homothetic, and continuous on* $\mathbb{R}^m_+$. *Then* $u(\mathbf{x})$ *is a monotonic transformation of a homogeneous of degree one function* v.

Proof of Theorem. Apply Lemma 20.14.1 to $\mathbf{x}^* = \boldsymbol{e}$. Then for every $\mathbf{x} \in \mathbb{R}^m_+$, there is a $\alpha \geq 0$ with $u(\alpha\boldsymbol{e}) = u(\mathbf{x})$. In that case, define $v(\mathbf{x}) = \alpha$, so that $u\big(v(\mathbf{x})\boldsymbol{e}\big) = u(\mathbf{x})$.

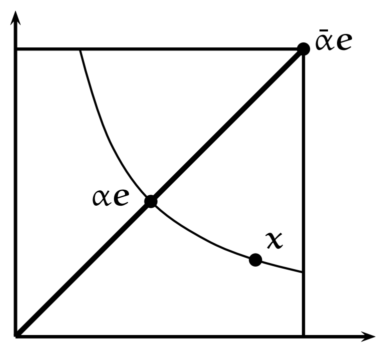

Figure 20.15.1: The position of the intersection of the indifference curve with the diagonal determines the unique α with $u(\mathbf{x}) = u(\alpha\boldsymbol{e})$.

Now $v(\mathbf{x}) \geq v(\mathbf{y})$ if and only if

$$u(\mathbf{x}) = u\big(v(\mathbf{x})\boldsymbol{e}\big) \geq u\big(v(\mathbf{y})\boldsymbol{e}\big) = v(\mathbf{y}).$$

It follows that v is utility equivalent to u.

We know $u(\mathbf{x}) = u\big(v(\mathbf{x})\boldsymbol{e}\big)$. By homotheticity of u, $u(t\mathbf{x}) = u\big(tv(\mathbf{x})\boldsymbol{e}\big)$. But $u(t\mathbf{x}) = u\big(v(t\mathbf{x})\boldsymbol{e}\big)$, so

$$u\big(tv(\mathbf{x})\boldsymbol{e}\big) = u\big(v(t\mathbf{x})\boldsymbol{e}\big). \tag{20.15.1}$$

so $tv(\mathbf{x}) = v(t\mathbf{x})$ because Lemma 20.14.1 shows that equation (20.15.1) has a unique solution. It follows that v is homogeneous of degree one.

Since u and v are equivalent, by Theorem 20.12.1, there is a monotonic transformation ϕ with $u(\mathbf{x}) = (\phi \circ v)(\mathbf{x})$. ∎

Although homothetic functions are related to homogeneous functions, there are differences. Nothing like Euler's Theorem need hold for functions that are merely homothetic. To see this, consider $f(\mathbf{x}) = \sum_\ell \alpha_\ell \ln x_\ell$. Then $\big[D_x f(\mathbf{x})\big]\mathbf{x} = \sum_\ell \alpha_\ell$, which cannot be written in the desired form.

20.16 Homotheticity and Marginal Rates of Substitution

When preferences are described by a smooth utility function, we can also describe homotheticity by saying that the marginal rate of substitution is the same anywhere on a given ray through the origin. The marginal rates of substitution are constant along any ray through the origin. This means that the shape of the indifference curves are preserved under scalar multiplication. All slopes remain unchanged.

Theorem 20.16.1. *Let* $C \subset \mathbb{R}^m$ *be an open cone and* $f\colon C \to \mathbb{R}$ *with* $Df(\mathbf{x}) \gg 0$ *on* C. *Suppose* $f\colon C \to \mathbb{R}$ *is homothetic and differentiable. Then* $MRS_{k\ell}(\mathbf{x}) = MRS_{k\ell}(t\mathbf{x})$ *for all* $t > 0$ *and* $\mathbf{x} \in C$.

The requirement that the marginal rate of substitution exists rules out cases where we are dividing by zero.

The proof is easy if we can write $f = F \circ \phi$, F monotonic, ϕ homogeneous of degree one, and both F and ϕ differentiable. Use the Chain Rule to compute $MRS_{k\ell}(\mathbf{x}) = (\partial\phi/\partial x_k)/(\partial\phi/\partial x_\ell)$. By Theorem 20.4.1, both derivatives are homogeneous of degree zero, and so is the marginal rate of substitution. Below, we prove this in a slightly more general fashion.

Proof. Consider the upper contour sets $U(\mathbf{x}) = \{\mathbf{y} : f(\mathbf{y}) \ge f(\mathbf{x})\}$. By homotheticity, $U(t\mathbf{x}) = tU(\mathbf{x})$. Homogeneity of the inner product then shows that if $\mathbf{p}$ supports $U(\mathbf{x})$, it also supports $U(t\mathbf{x})$.

Since f is differentiable, $\mathbf{p}$ and $Df(\mathbf{x})$ are proportional by Theorem 21.22.1. Thus $Df(\mathbf{x})$ and $Df(t\mathbf{x})$ are also proportional. The factors of proportionality cancel when constructing the marginal rates of substitution, so $MRS_{k\ell}(\mathbf{x}) = MRS_{k\ell}(t\mathbf{x})$. ∎

There is a converse, which we will state, but not prove.

Theorem 20.16.2. *Suppose* $f\colon \mathbb{R}^m_{++} \to \mathbb{R}$ *is* $\mathcal{C}^2$, $Df \gg \mathbf{0}$, *and* $MRS_{k\ell}$ *is homogeneous of degree zero in* $\mathbf{x}$ *for every* k *and* ℓ. *Then* f *is homothetic.*

Proof. A proof may be found in Lau.[2] Lau uses the slightly weaker assumption that there is some j with $\partial f/\partial x_j \neq 0$, in which case the MRS condition must be restated to work around the fact that $MRS_{k\ell}$ may not be defined for all pairs k and ℓ. Lau does not do this, but the replacement for the MRS condition is that for all k and ℓ, there are homogeneous of degree zero functions $g_{k\ell}$ such that $\partial f/\partial x_k = g_{k\ell} \times (\partial f/\partial x_\ell)$. When $Df \gg \mathbf{0}$, this is equivalent to the marginal rates of substitution being homogeneous of degree zero in $\mathbf{x}$. ∎

[2] Lemma 1 in Lau, Lawrence J. (1969) Duality and the structure of utility functions *J. Econ. Theory*, **1**, 374–396.

21. Concave and Quasiconcave Functions

11/12/20

NB: Problems 14, 18 and 22 from Chapter 19 and problems 1, 11, and 18 from Chapter 20 are due on Thursday, November 19.

Convex, concave, and similar functions arise naturally in economics. They include the indirect utility function, cost function, expenditure function, and profit function. Moreover, concavity is usually assumed of utility as it ensures a diminishing (or at least non-increasing) marginal rate of substitution. It often applies to production.

21.1 Convex and Concave Functions

Recall that in any vector space, the *line segment between* $\mathbf{x}$ *and* $\mathbf{y}$ is given by $\ell(\mathbf{x},\mathbf{y}) = \{(1-t)\mathbf{x} + t\mathbf{y} : 0 \le t \le 1\}$, and that a set S is *convex* if it contains $\ell(\mathbf{x},\mathbf{y})$ whenever $\mathbf{x},\mathbf{y} \in S$.

Convex and Concave Functions. Let S be a convex set.

- A function $f\colon S \to \mathbb{R}$ is *convex* if for all $\mathbf{x},\mathbf{y} \in S$ and $0 \le t \le 1$, $f\big(t\mathbf{x} + (1-t)\mathbf{y}\big) \le tf(\mathbf{x}) + (1-t)f(\mathbf{y})$.
- A function $f\colon S \to \mathbb{R}$ is *concave* if for all $\mathbf{x},\mathbf{y} \in S$ and $0 \le t \le 1$, $f\big(t\mathbf{x} + (1-t)\mathbf{y}\big) \ge tf(\mathbf{x}) + (1-t)f(\mathbf{y})$.

When the inequalities are strict for $0 < t < 1$, we say the function is *strictly convex* or *strictly concave*.

One consequence is that for convex functions, every chord of the graph of the function lies on or above the graph. For concave functions, every chord lies on or below the graph.

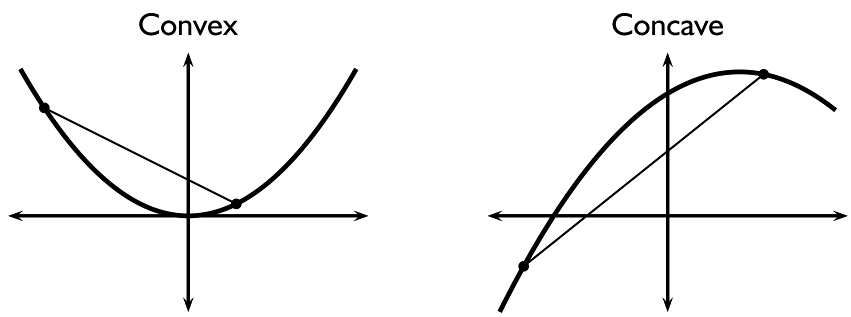

Figure 21.1.1: The left panel shows a convex function, where every chord connecting any two points of the graph lies above the graph.

The right panel illustrates a concave function, and every chord connecting any two points of the graph lies below the graph.

► **Example 21.1.2: Convex Functions.** Any linear function $f(\mathbf{x}) = \mathbf{p}\cdot\mathbf{x}$ is both concave and convex, as is the generic affine function $f(\mathbf{x}) = a + \mathbf{p}\cdot\mathbf{x}$. The function $f(\mathbf{x}) = \sum_{\ell=1}^{m} x_\ell^2$ is convex while $f(\mathbf{x}) = \sum_{\ell=1}^{m} x_\ell^{1/2}$ is concave. The function $f(x) = e^x$ is convex, while $f(x) = \ln x$ is concave. ◄

21.2 Basic Properties of Concave and Convex Functions

Several easily established properties of convex and concave functions are collected together without proof in Theorem 21.2.1.

Theorem 21.2.1.

1. *A function* f *is (strictly) convex if and only if* $-f$ *is (strictly) concave.*
2. *A positive scalar multiple of a concave (convex) function is concave (convex).*
3. *The sum of two concave (convex) functions is concave (convex).*
4. *If* ϕ *is concave (convex) and weakly increasing on* $\mathbb{R}$ *and* f *is a concave (convex) function, then* $\phi \circ f$ *is concave (convex).*
5. *The pointwise limit of a sequence of concave (convex) functions is concave (convex).*
6. *The infimum (supremum) of a sequence of concave (convex) functions is concave (convex).*

Proof. You should be able to prove these yourself.

21.3 Support Property

Supporting Hyperplane. We say a vector $\mathbf{p}$ *supports a set* S at $\mathbf{x}_0$ if either $\mathbf{p}\cdot\mathbf{x} \geq \mathbf{p}\cdot\mathbf{x}_0$ for every $\mathbf{x} \in S$ or $\mathbf{p}\cdot\mathbf{x} \leq \mathbf{p}\cdot\mathbf{x}_0$ for every $\mathbf{x}$ in S.

Of course, the set $H = \{\mathbf{x} : \mathbf{p}\cdot\mathbf{x} = \mathbf{p}\cdot\mathbf{x}_0\}$ is a hyperplane. So supporting the set means that the set is on one side of the hyperplane. Moreover, they necessarily touch at $\mathbf{x}_0$.

That means we can restate the definition in terms of half-spaces. A vector $\mathbf{p}$ supports S at $\mathbf{x}_0$ if and only if S is contained in the one of the two half-spaces $H^+(\mathbf{p}, \mathbf{p}\cdot\mathbf{x}_0)$ and $H^-(\mathbf{p}, \mathbf{p}\cdot\mathbf{x}_0)$.

Let $f\colon S \to \mathbb{R}$ be a function where $S \subset \mathbb{R}^m$. The *upper contour set* is defined by $U(\mathbf{y}) = \{\mathbf{x} \in S : f(\mathbf{x}) \geq f(\mathbf{y})\}$. The *lower contour set* is defined by $L(\mathbf{y}) = \{\mathbf{x} \in S : f(\mathbf{x}) \leq f(\mathbf{y})\}$.

► Example 21.3.1: A Convex Upper Contour Set. Figure 21.14.2 illustrates an upper contour set for $u\colon \mathbb{R}^2_+ \to \mathbb{R}$ given by $u(x, y) = xy$.

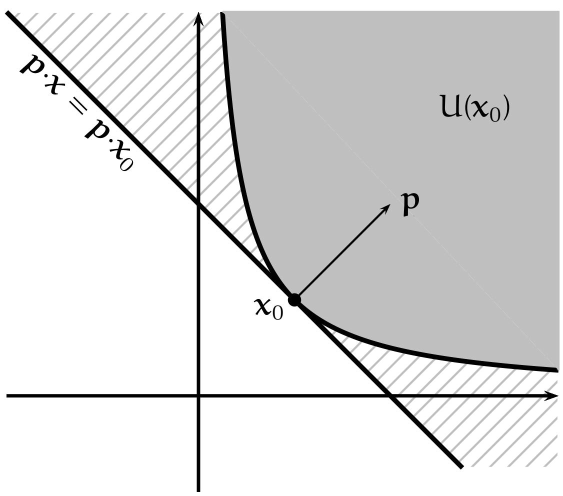

Figure 21.3.2: The shaded area is the upper contour set $U(\mathbf{x}_0)$ for $u(x, y) = xy$. The half-space $H^+(\mathbf{p}, \mathbf{x}_0)$ is the hatched area when $\mathbf{x}_0 = (1, 1)$ and $\mathbf{p} = (1, 1)$. The vector $\mathbf{p}$ supports $U(\mathbf{x}_0)$ at $\mathbf{x}_0$.

◄

21.4 Support Property Theorem

The Support Property Theorem shows that a differentiable function f is concave if and only if $Df(x_0)$ supports the upper contour set at x_0 for every $x_0 \in \operatorname{dom} f$. It also shows that f is convex if and only if $Df(x_0)$ supports the lower contour set at x_0 for every $x_0 \in \operatorname{dom} f$.

Support Property Theorem. *Suppose $f\colon U \to \mathbb{R}$ is $\mathcal{C}^1$ where U is an open convex set $U \subset \mathbb{R}^m$. The function f is concave if and only if*

$$f(y) \le f(x) + Df(x)(y - x) \qquad (21.4.1)$$

for all $x, y \in U$. The function f is convex if and only if

$$f(y) \ge f(x) + Df(x)(y - x) \qquad (21.4.2)$$

for all $x, y \in U$.

Proof (Only If). First suppose f is concave and take ε with $0 < \varepsilon < 1$.

$$f\big(x + \varepsilon(y - x)\big) = f\big((1-\varepsilon)x + \varepsilon y\big) \ge \varepsilon f(y) + (1-\varepsilon)f(x).$$

We can rearrange to obtain

$$f\big(x + \varepsilon(y - x)\big) - f(x) \ge \varepsilon\big[f(y) - f(x)\big].$$

Dividing by $\varepsilon > 0$ and letting $\varepsilon \to 0$ yields

$$Df(x)(y - x) \ge f(y) - f(x).$$

When f is convex, the only change to this part of the proof is that all three inequalities must be reversed.

Proof Continues...

21.5 Support Property, continued

Proof (If). Now suppose the inequality

$$f(\mathbf{y}) \le f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{y} - \mathbf{x}) \tag{21.4.1}$$

is satisfied for all $\mathbf{x}, \mathbf{y} \in U$. Replace $\mathbf{x}$ by $\mathbf{x}' = \mathbf{x} + (1-\alpha)(\mathbf{y} - \mathbf{x}) = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ where $0 \le \alpha \le 1$. Then

$$f(\mathbf{y}) \le f(\mathbf{x}') + \alpha\, Df(\mathbf{x}')(\mathbf{y} - \mathbf{x}) \le f(\mathbf{x}') - \alpha\, Df(\mathbf{x}')(\mathbf{x} - \mathbf{y}). \tag{21.5.3}$$

Rewrite the support equation (21.4.1) by replacing $\mathbf{x}$ with $\mathbf{x}'$ and $\mathbf{y}$ with $\mathbf{x}$. This yields

$$f(\mathbf{x}) \le f(\mathbf{x}') + (1-\alpha)\, Df(\mathbf{x}')(\mathbf{x} - \mathbf{y}). \tag{21.5.4}$$

Then multiply equation (21.5.3) by $(1-\alpha)$, equation (21.5.4) by α, and add them together. The $Df(\mathbf{x}')$ terms cancel, leaving

$$\alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) \le f(\mathbf{x}') = f\big(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}\big)$$

establishing concavity. The proof for the convex case is the same, but with every inequality reversed. ∎

One consequence of the Support Property Theorem is that if f is concave, the $Df(\mathbf{x}_0)$ supports the upper contour set $U(\mathbf{x}_0)$ at $\mathbf{x}_0$. To see this, suppose $f(\mathbf{y}) \ge f(\mathbf{x}_0)$. Then equation (21.4.1) implies

$$0 \le f(\mathbf{y}) - f(\mathbf{x}_0) \le Df(\mathbf{x}_0)(\mathbf{y} - \mathbf{x}_0),$$

so $Df(\mathbf{x}_0)\mathbf{y} \ge Df(\mathbf{x}_0)\mathbf{x}_0$. This implies $U(\mathbf{x}_0) \subset H^+\big(Df(\mathbf{x}_0), \mathbf{x}_0\big)$ as in Figure Fig:SupportedSet.

21.6 The Support Property in $\mathbb{R}$

By the Support Property Theorem, a differentiable function is concave if and only if equation (21.4.1) holds. This important inequality can be taken as the definition of concavity for differentiable functions.

When f is a function on a subset of the real line, we can use the right-hand side of equation (21.4.1) to define a line, $y = f(x_0) + f'(x_0)(x - x_0)$. This line is tangent to the graph of f at the point $(x_0, f(x_0))$, and the graph of the function is in the lower half-space that the tangent line defines. The tangent line supports the graph of f in the sense that the graph lies within one of the half-spaces defined by the tangent.

The tangent line supports both the graph and subgraph of f at $(x_0, f(x_0))$. The subgraph is $\text{sub}\, f\{(\mathbf{x}, y) : f(\mathbf{x}) \le y\}$ at $(x_0, f(x_0))$.

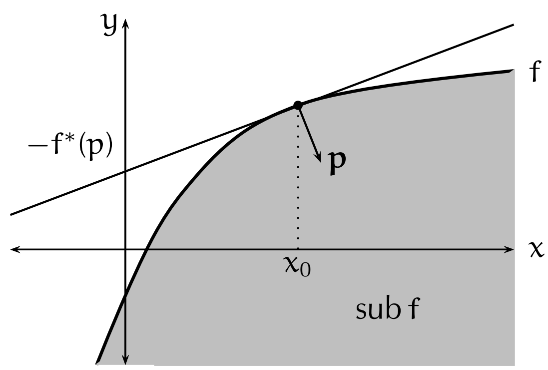

Figure 21.6.1: The tangent line at x_0 has the equation $y = f(x_0) + f'(x_0)(x - x_0)$. Because f is concave, the tangent line supports the subgraph of f. The graph is never above the tangent line and touches it at $(x_0, f(x_0))$. Let $p = f'(x_0)$. The vector $\mathbf{p} = (p, -1)$ is perpendicular to the tangent line. Its vertical intercept is the negative of the concave conjugate function $f^*(p) = px_0 - f(x_0)$.

21.7 Support Property in $\mathbb{R}^m$

In Figure 21.6.1 $p = f'(x_0)$ is the slope of the tangent line. We now rewrite the equation of the tangent in a way that that shows it is a hyperplane in $\mathbb{R}^2$. Thus

$$(p, -1)\begin{pmatrix} x \\ y \end{pmatrix} = px_0 - f(x_0). \tag{21.7.5}$$

The vector $\mathbf{p} = (p, -1)$ is perpendicular to the tangent line.

The right-hand side of equation (21.7.5) is not zero unless tangent goes through the origin. It tells us how much the tangent line is offset from the origin. That value is called the *concave conjugate function* and is denoted $f^*(p) = px_0 - f(x_0)$ when $p = f'(x_0)$. In fact, $f^*(p)$ is the negative of the vertical intercept of the tangent line. The equation of the tangent line then becomes

$$(p, -1)\begin{pmatrix} x \\ y \end{pmatrix} = f^*(p)$$

and the supergradient inequality is

$$(p, -1)\begin{pmatrix} x \\ f(x) \end{pmatrix} \geq f^*(p). \tag{21.7.6}$$

More generally, if f is a function of m variables, we can consider its graph in $\mathbb{R}^{m+1}$, and the picture is much the same.

$$(Df(\mathbf{x}_0), -1)\begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix} \leq (Df(\mathbf{x}_0), -1)\begin{pmatrix} \mathbf{x}_0 \\ f(\mathbf{x}_0) \end{pmatrix}$$

where $\mathbf{p} = Df(\mathbf{x}_0)$.

21.8 Supporting Non-Differentiable Functions

We can often support concave functions in the same way even if they aren't differentiable. However, there may be more than one vector $\mathbf{p}$ that supports them in such a case.

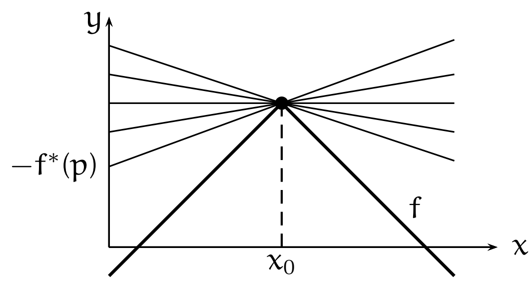

Figure 21.8.1: As you can see, there are many lines that support the function f at $(x_0, f(x_0))$. The function has slope $+1$ to the left of x_0 and -1 to the right. Any line with a slope between -1 and $+1$ that is tangent to the graph of f at $(x_0, f(x_0))$ will satisfy the support property.

Supergradient and Subgradient. If f is a concave function on a convex set $U \subset \mathbb{R}^m$, and $\mathbf{p} \in \mathbb{R}^m$ satisfies

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \mathbf{p}\cdot(\mathbf{y} - \mathbf{x}), \tag{21.8.7}$$

for all $\mathbf{x}, \mathbf{y} \in U$, we call $\mathbf{p}$ a *supergradient* of f at $\mathbf{x}$. It is a type of generalized derivative. There's also one for convex functions, called a subgradient. A vector $\mathbf{p}$ is a *subgradient* of f at $\mathbf{x}$ if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{p}\cdot(\mathbf{y} - \mathbf{x}), \tag{21.8.8}$$

for all $\mathbf{x}, \mathbf{y} \in U$.

The Support Function Theorem tells us that when f is differentiable and concave (convex) $Df(\mathbf{x}_0)$ is the only supergradient (subgradient) of f at $\mathbf{x}_0$.

21.9 Supergradients and Optimization

When f is concave (convex) the first order conditions are not only necessary for an optimum, they are sufficient too.

Theorem 21.9.1. *Let $U \subset \mathbb{R}^m$ be convex.*

(a) *Suppose $\mathbf{0}$ is a supergradient of f at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a global maximizer of f over U. In particular, this applies if f is differentiable and $Df(\mathbf{x}^*) = \mathbf{0}$.*

(b) *Suppose $\mathbf{0}$ is a subgradient of f at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a global minimizer of f over U. In particular, this applies if f is differentiable and $Df(\mathbf{x}^*) = \mathbf{0}$.*

Proof. Setting $\mathbf{p} = \mathbf{0}$ in the supergradient inequality (21.8.7), yields $f(\mathbf{y}) \leq f(\mathbf{x}^*)$ for all $\mathbf{y} \in U$. ∎

This even works on some concave functions that aren't differentiable, as in Figure 21.8.1 where a horizontal line ($p = 0$) supports the function at x_0. That means x_0 is a maximum, as is also clear from the graph.

This result is extremely useful in economics, as we are faced with a convex feasible set U and a concave utility function we wish to maximize or convex function such as $\boldsymbol{w}\cdot\boldsymbol{z}$ needing to be minimized. In either case, the first order necessary conditions become sufficient for optimization.

We can go a bit further than this for differentiable functions.

Theorem 21.9.2. *Let $U \subset \mathbb{R}^m$ be convex.*

(a) *Suppose f is a concave $\mathcal{C}^1$ function on U and $\mathbf{x}^*$ obeys $Df(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \leq 0$ for all $\mathbf{y} \in U$. Then $\mathbf{x}^*$ is a global maximizer of f on U.*

(b) *Suppose f is a convex $\mathcal{C}^1$ function on U and $\mathbf{x}^*$ obeys $Df(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \geq 0$ for all $\mathbf{y} \in U$. Then $\mathbf{x}^*$ is a global minimizer of f on U.*

Proof. For the concave case, if $Df(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \leq 0$, then $f(\mathbf{y}) \leq f(\mathbf{x}^*) + Df(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \leq f(\mathbf{x}^*)$. The convex case is similar. ∎

► **Example 21.9.3: Decreasing Functions.** Suppose f is a decreasing $\mathcal{C}^1$ function of one variable on an interval $[a, b]$. Then $f'(a)(x - a) \leq 0$ for all $x \in [a, b]$ because $x - a \geq 0$ and $f'(a) \leq 0$. It follows that a is a maximizer of f on $[a, b]$ ◄

It works in $\mathbb{R}^m$ too. Here's an example for $\mathbb{R}^2$.

► **Example 21.9.4: Global Minimum via Support.** Let $f(x, y) = -x^2 - y^2$ on the set $[0, 4] \times [0, 3] \subset \mathbb{R}^2$. Then $Df = (-2x, -2y)$. We will show that $\mathbf{x}^* = (4, 3)$ is a minimizer. Now $Df(\mathbf{x}^*) = (-8, -6)$ and $Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = (-8, -6)\cdot(x - 4, y - 3)$. Here both $x \leq 4$ and $y \leq 3$, so $Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \geq 0$, showing that $\mathbf{x}^*$ is a global minimum. ◄

21.10 Hessian Convexity Tests: Necessity

The support property can be used to show that the Hessian $D^2f(x)$ is negative semidefinite when f is concave and positive semidefinite when f is convex.

Theorem 21.10.1. *Suppose $f\colon U \to \mathbb{R}$ is $\mathcal{C}^2$ on an open convex set $U \subset \mathbb{R}^m$. If f is concave, then the Hessian $D^2f(x)$ is negative semidefinite for all $x \in U$. If f is convex, then the Hessian $D^2f(x)$ is positive semidefinite for all $x \in U$.*

Proof. We will prove the concave case, the convex case is similar, with inequalities reversed. By the Support Property Theorem (equation 21.4.1)

$$f(y) \le f(x) + Df(x)(y - x) = f(x) - Df(x)(x - y)$$

reversing x and y in equation (21.4.1) yields

$$f(x) \le f(y) + Df(y)(x - y).$$

Adding the equations together and simplifying, we obtain

$$0 \le \big(Df(y) - Df(x)\big)(x - y).$$

Now set $y = x + h$. We then have

$$\big(Df(x + h) - Df(x)\big)h \le 0.$$

Divide by $\|h\|$ and let $\|h\| \to 0$, obtaining $h^T D^2f(x)h \le 0$. In other words, $D^2f(x)$ is negative semidefinite for all $x \in U$. ∎

21.11 Hessian Convexity Tests: Necessity and Sufficiency

When f is $\mathcal{C}^2$, we can use the Hessian matrix and Taylor's formula to determine whether f is concave or convex.

Theorem 21.11.1. *Suppose* $f\colon U \to \mathbb{R}$ *is* $\mathcal{C}^2$ *on an open convex set* $U \subset \mathbb{R}^m$.

1. *The function* f *is concave if and only if* $D^2f(x)$ *is negative semidefinite for all* $x \in U$.
2. *If* $D^2f(x)$ *is negative definite for all* $x \in U$, f *is strictly concave.*
3. *The function* f *is convex if and only if* $D^2f(x)$ *is positive semidefinite for all* $x \in U$.
4. *If* $D^2f(x)$ *is positive definite for all* $x \in U$, f *is strictly convex.*

Proof. Taylor's formula tells us that

$$f(y) = f(x) + Df(x)(y - x) + (y - x)^T D^2f(z)(y - x). \tag{21.11.9}$$

for some z on the line segment between x and y.

(1) If D^2f is negative semidefinite on U, this implies the support property, so f is concave by the Support Property Theorem. Conversely, if f is concave, Proposition 21.10.1 shows that $D^2f(x)$ is negative semidefinite on U.

(2) If f is negative definite on U, equation (21.11.9) yields $f(y) < f(x) + Df(x)(y - x)$ for all $y \neq x$. Repeating the calculations in the Support Property Theorem for $y \neq x$ and $0 < \alpha < 1$, shows that f is strictly concave.

Parts (3) and (4) are similar. ∎

21.12 Convexity and Concavity: Determinant Tests

Recall the following definitions

Definite and Semidefinite Matrices. Recall that an $L \times L$ symmetric matrix $\mathbf{A}$ is

(a) *Positive semidefinite* if $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^L$,
(b) *Positive definite* if $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
(c) *Negative semidefinite* if $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^L$,
(d) *Negative definite* if $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$, and
(e) *Indefinite* if there are $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ and $\mathbf{y}^T\mathbf{A}\mathbf{y} < 0$.

The support property can be used to show that $D^2f(\mathbf{x})$ is negative semidefinite when f is concave and positive semidefinite when f is convex.

If $\mathbf{A}$ is a matrix, we will use $\tilde{\mathbf{A}}_k$ to denote a generic k^{th}-order principal minor of $\mathbf{A}$. One example of a non-leading principal minor for $m = 6$ and $k = 3$ is

$$\begin{vmatrix} a_{11} & a_{13} & a_{16} \\ a_{31} & a_{33} & a_{36} \\ a_{61} & a_{63} & a_{66} \end{vmatrix}$$

where the 2^{nd}, 4^{th}, and 5^{th} rows and columns have been deleted. Keep in mind that there are usually many k^{th}-order principal minors.

We can rewrite our results relating convexity and definiteness by applying the usual determinant tests to the Hessian. That yields the following theorem.

Theorem 21.12.1. *Suppose $f\colon U \to \mathbb{R}$ is $\mathcal{C}^2$ on an open convex set $U \subset \mathbb{R}^m$. Let $\mathbf{H}(\mathbf{x}) = D^2f(\mathbf{x})$ denote the Hessian of f.*

1. *The function f is convex if and only if every k^{th}-order principal minor obeys $\tilde{\mathbf{H}}_k(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in U$.*
2. *The function f is concave if and only if every k^{th}-order principal minor obeys $(-1)^k\tilde{\mathbf{H}}_k(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in U$.*
3. *Suppose the leading principal minors obey $H_k(\mathbf{x}) > 0$ for $k = 1, \ldots, m$ and all $\mathbf{x} \in U$. Then f is strictly convex.*
4. *Suppose the leading principal minors obey $(-1)^kH_k(\mathbf{x}) > 0$ for $k = 1, \ldots, m$ and all $\mathbf{x} \in U$. Then f is strictly concave.*

21.13 Using the Determinant Test

▶ **Example 21.13.1: Cobb-Douglas Utility.** Consider the Cobb-Douglas utility function on $\mathbb{R}^2_{++}$ defined by $u(\mathbf{x}) = x_1^\alpha x_2^{1-\alpha}$ with $0 < \alpha < 1$. The Hessian is

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^\alpha x_2^{-\alpha-1} \end{bmatrix}.$$

Clearly, $\det \mathbf{H} = 0$. The two first-order minors are $\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} < 0$ and $(1-\alpha)(-\alpha)x_1^\alpha x_2^{-\alpha-1} < 0$, which means that u is concave on $\mathbb{R}^2_{++}$. Concavity on $\mathbb{R}^2_+$ then follows from continuity.

Note that u is not strictly concave because $\alpha u(\mathbf{0}) + (1-\alpha)u(\boldsymbol{e}) = 1-\alpha = u\big((1-\alpha)\boldsymbol{e}\big)$. ◀

▶ **Example 21.13.2: CES Utility.** Another example is the constant elasticity of substitution utility function $u(\mathbf{x}) = [x_1^{-\rho} + x_2^{-\rho}]^{-1/\rho}$ where $\rho > -1$, $\rho \neq 0$. The Hessian is

$$\mathbf{H}(\mathbf{x}) = (1+\rho)\frac{[x_1^{-\rho} + x_2^{-\rho}]^{-(\frac{1+2\rho}{\rho})}}{x_1^{2+\rho}x_2^{2+\rho}} \begin{pmatrix} -x_2^2 & x_1x_2 \\ x_1x_2 & -x_1^2 \end{pmatrix}.$$

The two first-order minors are both negative, while the second-order minor is 0. Thus u is concave on $\mathbb{R}_{++}$. ◀

21.14 Example: Upper and Lower Contour Sets

Upper contour sets are convex for concave functions while lower contour sets are convex for convex functions. It is often the case that the lower contour set of a concave function is not convex. This happens in the figure below.

► Example 21.14.1: Upper and Lower Contour Sets. The left side Figure 21.14.2 illustrates an upper contour set for $u\colon \mathbb{R}^2_+ \to \mathbb{R}$ given by $u(x, y) = x^{1/3}y^{2/3}$. The right side shows the lower contour set for the same function. It is not convex.

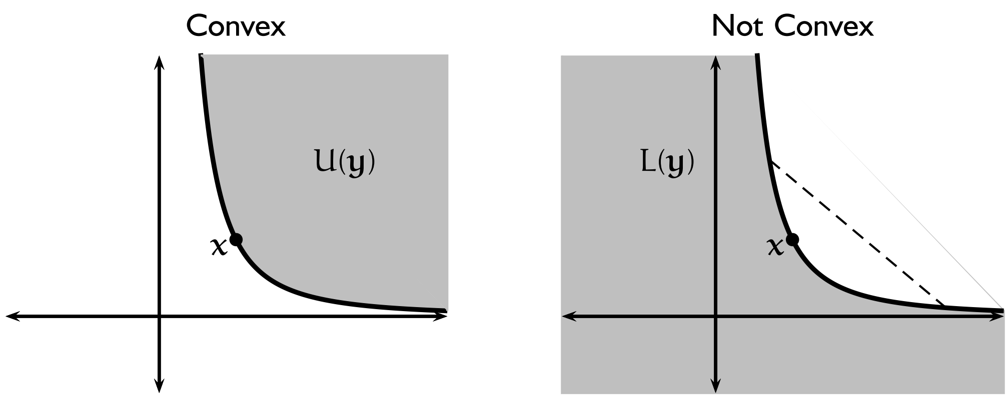

Figure 21.14.2: The shaded area in the left panel is the upper contour set $U(\mathbf{y})$ for $u(x, y) = x^{1/3}y^{2/3}$ and $\mathbf{y} = (1, 1)$.

The shaded area in right panel is the lower contour set for the same function. The dashed line segment would be in $L(\mathbf{y})$ if $L(\mathbf{y})$ were convex. The line segment goes outside $L(\mathbf{y})$, so the lower contour set is not convex.

◄

21.15 Convexity of Upper and Lower Contour Sets

In Example 21.14.2, the upper contour set was convex. This was no accident. Rather, it is a consequence of using a concave function.

Theorem 21.15.1. *Let $U \subset \mathbb{R}^m$ be a convex set.*

1. *If $f\colon U \to \mathbb{R}$ is a concave function, then the upper contour set $U(y)$ is convex.*
2. *If $f\colon U \to \mathbb{R}$ is a convex function, then the lower contour set $L(y)$ is convex.*

Proof. Let $x, x' \in U(y)$. Then $f(x), f(x') \geq f(y)$. By concavity of f

$$f\big((1-t)x + tx'\big) \geq (1-t)f(x) + tf(x') \geq (1-t)f(y) + tf(y) = f(y),$$

showing that the convex combination $(1-t)x + tx' \in U(y)$ whenever $x, x' \in U(y)$.

The proof of (2) is similar. ∎

Contour Sets are Ordinal. Notice that $U(x)$ and $L(x)$ remain unchanged when a monotonic transformation is applied to f. This is because $f(x) \geq f(y)$ if and only if $(\phi \circ f)(x) \geq (\phi \circ f)(y)$ when ϕ is a monotonic transformation.

21.16 Quasiconvex and Quasiconcave Functions

Quasiconvexity and quasiconcavity are the ordinal versions of convexity and concavity. They are defined in terms of the upper and lower contour sets, respectively. We just saw that upper and lower contour sets are unaffected by monotonic transformations, and hence ordinal.

Quasiconvexity and Quasiconcavity. Let $f\colon S \to \mathbb{R}$.

1. The function f is *quasiconcave* on S if and only if the upper contour set $U(\mathbf{y}) = \{\mathbf{x} \in S : f(\mathbf{x}) \geq f(\mathbf{y})\}$ is a convex set for every $\mathbf{y} \in S$.
2. The function f is *quasiconvex* on S if and only if the lower contour set $L(\mathbf{x}) = \{\mathbf{x} \in S : f(\mathbf{x}) \leq f(\mathbf{y})\}$ is a convex set for every $\mathbf{y} \in S$.

From the definition, it is easy to see that f is quasiconvex if and only if $-f$ is quasiconcave.

There are other ways to characterize quasiconvexity and quasiconcavity. One is the following.

Theorem 21.16.1. *Let $f\colon S \to \mathbb{R}$ where S is a convex set.*

1. *The function f is quasiconcave if and only if for all t obeying $0 \leq t \leq 1$, $f(t\mathbf{x} + (1-t)\mathbf{y}) \geq \min\{f(\mathbf{x}), f(\mathbf{y})\}$.*
2. *The function f is quasiconvex if and only if for all t obeying $0 \leq t \leq 1$, $f(t\mathbf{x} + (1-t)\mathbf{y}) \leq \max\{f(\mathbf{x}), f(\mathbf{y})\}$.*

Of course, if $f\colon S \to \mathbb{R}$ is quasiconcave or quasiconvex on S, S must be a convex set because if $\mathbf{x}', \mathbf{x}'' \in S$, convex combinations of $\mathbf{x}'$ and $\mathbf{x}''$ are in either $\{\mathbf{x} \in S : f(\mathbf{x}) \geq \min[f(\mathbf{x}'), f(\mathbf{x}'')]\}$ (f quasiconcave) or in $\{\mathbf{x} \in S : f(\mathbf{x}) \leq \max[f(\mathbf{x}'), f(\mathbf{x}'')]\}$ (f quasiconvex).

In Theorem 21.15.1 we saw that convex functions are quasiconvex and concave functions are quasiconcave. The relationship is only one-way. Indeed, when $f\colon \mathbb{R} \to \mathbb{R}$, every monotone function is simultaneously quasiconvex and quasiconcave! In contrast, the only functions that are both convex and concave on $\mathbb{R}$ are the affine functions.

21.17 Quasiconvexity on the Real Line

For f defined on a convex subset of $\mathbb{R}$, we can completely characterize the quasiconcave and quasiconvex functions.

Theorem 21.17.1. *Let* I *be an interval in* $\mathbb{R}$. *Suppose* $f\colon I \to \mathbb{R}$.

1. *The function* f *is quasiconcave if and only* f *is either monotone or there exists a number* x^* *such that* f *is weakly increasing when* $x \le x^*$ *and weakly decreasing when* $x \ge x^*$.
2. *The function* f *is quasiconvex if and only if* f *is either monotone or there exists a number* x^* *such that* f *is weakly decreasing when* $x \le x^*$ *and weakly increasing when* $x \ge x^*$.

Proof. It is enough to prove the quasiconcave case as $-f$ is quasiconcave if f is quasiconvex.

We first prove sufficiency. As noted above, monotone functions are quasiconcave. We need only consider the case where x^* exists. Let $y \in I$ be given and let x_1 and x_2 be arbitrary points in the upper contour set $U(y)$ with $x_1 < x_2$. We must show that the interval $[x_1, x_2] \subset U(y)$.

If x^* is not in $[x_1, x_2]$, then f is either always weakly decreasing or always weakly increasing on $[x_1, x_2]$. Either way, for any $z \in [x_1, x_2]$, $f(z) \ge \min\{f(x_1), f(x_2)\} \ge f(y)$, so $z \in U(y)$. This shows that $[x_1, x_2] \subset U(y)$.

If $x^* \in [x_1, x_2]$, f is weakly increasing on $[x_1, x^*]$ and weakly decreasing on $[x^*, x_2]$. Again, for any $z \in [x_1, x_2]$, $f(z) \ge \min\{f(x_1), f(x_2)\} \ge f(y)$, so $z \in U(y)$ and hence $[x_1, x_2] \subset U(y)$. In other words, f is quasiconcave.

We now turn to necessity. Suppose that f is quasiconcave on I, but not monotone.

We prove this part by contradiction. If x^* does not exist, we can find x_1, x_2, and x_3 in I such that $x_1 < x_2 < x_3$ and $f(x_2) < \min\{f(x_1), f(x_3)\}$. If $f(x_1) \le f(x_3)$, set $y = x_1$, otherwise take $y = x_3$. Then the upper contour set $U(y)$ includes x_1 and x_3, but not x_2. It is not convex, contradicting the fact that f is quasiconcave. ∎

To sum up, quasiconcave functions defined on a real interval are either monotonic or single-peaked, while quasiconcave functions are either monotonic or single-troughed. Either the peak or trough may be an interval.

21.18 Strict Quasiconvexity and Quasiconcavity

One consequence of Proposition 21.16.1 is that a function is quasiconcave if for every $\mathbf{x}, \mathbf{y}$ with $f(\mathbf{x}) \geq f(\mathbf{y})$ and every t, $0 \leq t \leq 1$, we have $f\big((1-t)\mathbf{x} + t\mathbf{y}\big) \geq f(\mathbf{y})$. We can use this criterion for quasiconcavity to define *strict quasiconcavity*.

Strict Quasiconcavity and Strict Quasiconvexity. Let $f\colon S \to \mathbb{R}$ where S is a convex set.

1. The function f is *strictly quasiconcave* on S if and only if for every $\mathbf{x}, \mathbf{y} \in S$ with $f(\mathbf{x}) \geq f(\mathbf{y})$ and every t with $0 < t < 1$, we have $f\big((1-t)\mathbf{x} + t\mathbf{y}\big) > f(\mathbf{y})$.
2. The function $f\colon S \to \mathbb{R}$ is *strictly quasiconvex* if and only if for every $\mathbf{x}, \mathbf{y} \in S$ with $f(\mathbf{x}) \geq f(\mathbf{y})$ and every t with $0 < t < 1$, we have $f\big((1-t)\mathbf{x} + t\mathbf{y}\big) < f(\mathbf{y})$.

It is easy to see that any increasing transformation of a convex (or quasiconvex) function is quasiconvex and any increasing transformation of a concave (or quasiconcave) function is quasiconcave.

Not all quasiconcave functions are transformations of concave functions. Quasiconcave functions that are a monotonic transformation of a concave function are called *concavifiable*. Similarly, a quasiconvex function that is a monotonic transformation of a convex function is called *convexifiable*. There are quasiconcave functions that are not concavifiable.[1]

► **Example 21.18.1: Quasiconvex does Not Mean Convex.** Consider $f(x, y) = (x^2 + y^2)^{1/3}$. The function $g(x, y) = x^2 + y^2$ is convex (the Hessian is $2\mathbf{I}_2$). Raising it to the $1/3$-power is an increasing transformation. Thus f is quasiconvex.

However, f is not convex. To see this, we compute $f(0, 0) = 0$ and $f(1, 1) = 2$. Then $f(t, t) = 2t^2 < 2t = (1-t)f(0, 0) + tf(1, 1)$ for $0 < t < 1$. If it were convex, the left-hand side would be larger than the right. Instead, it is smaller. ◄

A strictly quasiconcave (quasiconvex) function can have at most one maximizer (minimizer).

Theorem 21.18.2. *Let $f\colon S \to \mathbb{R}$ where S is a convex set. Suppose f is strictly quasiconcave there are $\mathbf{x}_0, \mathbf{x}_1 \in S$ so that for all $\mathbf{y} \in S$, $f(\mathbf{y}) \leq f(\mathbf{x}_i)$, $i = 0, 1$. Then $\mathbf{x}_0 = \mathbf{x}_1$.*

Proof. Since both $\mathbf{x}_i$ maximize f over S, $f(\mathbf{x}_1) = f(\mathbf{x}_0)$. If $\mathbf{x}_0 \neq \mathbf{x}_1$, define $\mathbf{x}_2 = \frac{1}{2}\mathbf{x}_0 + \frac{1}{2}\mathbf{x}_1$. By strict quasiconcavity, $f(\mathbf{x}_2) > f(\mathbf{x}_0)$, showing that $\mathbf{x}_0$ is not a maximum. This contradiction shows that $\mathbf{x}_0 = \mathbf{x}_1$. ∎

A similar result holds for minima of a strictly quasiconvex function.

[1] This has been known since de Finetti (1949). The problem of finding whether a given quasiconcave function is an increasing transformation of a concave function was posed (in terms of level sets) by Fenchel (1953). Although some results are known, it still does not have a completely satisfactory solution.

21.19 Support Property Theorem II

There are also support properties that characterize quasiconvex and quasiconcave functions.

Support Property Theorem II. *Suppose* f *is* $\mathcal{C}^1$ *on* U*, an open convex subset of* $\mathbb{R}^m$*. The function* f *is quasiconcave if and only if for all* $x, y \in U$

$$f(y) \ge f(x) \textit{ implies } \big[Df(x)\big](y - x) \ge 0. \tag{21.19.10}$$

The function f *is quasiconvex if and only if for all* $x, y \in U$

$$f(y) \ge f(x) \textit{ implies } \big[Df(x)\big](y - x) \le 0.$$

then f *is quasiconcave.*

Proof (Only if). Here f is quasiconcave. Suppose $f(y) \ge f(x)$. Consider $(1-\varepsilon)x + \varepsilon y = x + \varepsilon(y - x)$ for $0 < \varepsilon < 1$. By quasiconcavity,

$$f\big(x + \varepsilon(y - x)\big) \ge f(x) \quad \text{so} \quad \frac{f\big(x + \varepsilon(y - x)\big) - f(x)}{\varepsilon} \ge 0.$$

Let $\varepsilon \to 0$ to obtain $[Df(x)](y - x) \ge 0$, which establishes the result. ∎ (Only If)

Proof Continues...

21.20 Support Property Theorem II, Part II

Proof (If). Now suppose that for all $x, y \in U$; $f(y) \geq f(x)$ implies $[Df(x)](y - x) \geq 0$. We must show that f is quasiconcave.

We do this by contradiction. If f is not quasiconcave, there are x_0, x_1 and $0 < t_0 < 1$ with $f(x_1) \geq f(x_0)$ and $f\big((1 - t_0)x_0 + t_0 x_1\big) < f(x_0)$. Define

$$x_t = (1 - t)x_0 + tx_1.$$

We can then write $f(x_{t_0}) < f(x_0)$.

Take an interval $I = [t_1, t_2] \subset [0, 1]$ with $t_0 \in I$ and $f(x_t) \leq f(x_0)$ for all $t \in I$, and $f(x_{t_1}) = f(x_{t_2}) = f(x_0)$.

Now if $t \in I$ with $0 < t < 1$, $f(x_1) \geq f(x_0) \geq f(x_t)$. But then by equation (21.19.10),

$$Df(x_t)(x_0 - x_t) \geq 0 \quad \text{and} \quad Df(x_t)(x_1 - x_t) \geq 0. \tag{21.20.11}$$

Of course, $x_0 - x_t = t(x_0 - x_1)$ and $x_1 - x_t = (1 - t)(x_1 - x_0)$. We substitute into equation (21.20.11) to obtain

$$-tDf(x_t)(x_1 - x_0) \geq 0 \quad \text{and} \quad (1 - t)Df(x_t)(x_1 - x_0) \geq 0.$$

Since $-t < 0$ and $(1 - t) > 0$, we must have

$$Df(x_t)(x_1 - x_0) = 0 \tag{21.20.12}$$

for all $t \in I$, $0 < t < 1$.

Now $0 < f(x_0) - f(x_{t_0}) = f(x_{t_1}) - f(x_{t_0})$, so by the Mean Value Theorem, there is $t_3 \in (t_1, t_0)$ with

$$0 < f(x_0) - f(x_{t_0}) = Df(x_{t_3})(x_{t_1} - x_{t_0}) = (t_0 - t_1)Df(x_{t_3})(x_1 - x_0).$$

But this contradicts equation (21.20.12). That contradiction shows that f is quasiconcave. ∎

21.21 Maximization and Quasiconcavity

We can use Support Property II for f to show that any point solves a maximization problem.

Theorem 21.21.1. *Suppose* f *is quasiconcave and* $\mathcal{C}^1$ *on a convex open set* U *and that* $\mathbf{p} = Df(\bar{\mathbf{x}})$*. Then* $\bar{\mathbf{x}}$ *maximizes* $f(\mathbf{x})$ *under the constraints* $\mathbf{p}\cdot\mathbf{x} \le \mathbf{p}\cdot\bar{\mathbf{x}}$ *and* $\mathbf{x} \in U$.

Proof. We prove this by contradiction. Suppose to the contrary there is $\mathbf{x} \in U$ with $\mathbf{p}\cdot\mathbf{x} \le \mathbf{p}\cdot\bar{\mathbf{x}}$ and $f(\mathbf{x}) > f(\bar{\mathbf{x}})$. Then for $\varepsilon > 0$ small enough $\mathbf{x} - \varepsilon\mathbf{p} \in U$ (because U is open) and $f(\mathbf{x} - \varepsilon\mathbf{p}) > f(\bar{\mathbf{x}})$ (by continuity). By Support Property II, $\mathbf{p}\cdot(\mathbf{x} - \varepsilon\mathbf{p}) \ge \mathbf{p}\cdot\bar{\mathbf{x}}$. But then, $\mathbf{p}\cdot\mathbf{x} \ge \mathbf{p}\cdot\bar{\mathbf{x}} + \varepsilon\|\mathbf{p}\|^2 > \mathbf{p}\cdot\bar{\mathbf{x}}$. This contradicts our original supposition and so proves the theorem. ∎

Since the constraint $\mathbf{x} \in U$ cannot bind when U is open, $\bar{\mathbf{x}}$ maximizes f over all $\mathbf{x}$ with $\mathbf{p}\cdot\mathbf{x} \le \mathbf{p}\cdot\bar{\mathbf{x}}$. By Theorem , $\mathbf{h}^T[D^2f(\bar{\mathbf{x}})]\mathbf{h} \le 0$ for all $\mathbf{h}$ obeying $\mathbf{p}\cdot\mathbf{h} = 0$ or equivalently the bordered Hessian

$$\mathbf{B} = \begin{bmatrix} 0 & \mathbf{p}^T \\ \mathbf{p} & D^2f \end{bmatrix}$$

has bordered principal minors that obey $(-1)^{n-1}\mathbf{A}_n \ge 0$ for $n = 3, \ldots, m+1$. This provides a second-derivative test for quasiconcavity.

21.22 Uniqueness of Supports

When f is differentiable, we can show that the derivative is the only vector (up to scalar multiplication) that can support the upper (or lower) contour set. More precisely, we say $\mathbf{p}$ *supports a set* S at $\mathbf{x}_0$ if $\mathbf{p}\cdot\mathbf{x} \ge \mathbf{p}\cdot\mathbf{x}_0$ whenever $\mathbf{x} \in S$. We apply this idea to the upper contour set $\{\mathbf{x} : f(\mathbf{x}) \ge f(\mathbf{x}_0)\}$ at $\mathbf{x}_0$. It is important to realize that this result does not require quasiconcavity. However, if the function is not quasiconcave, the upper contour set may not have supports.

Theorem 21.22.1. *Suppose* f *is differentiable at* $\mathbf{x}_0$ *with* $Df(\mathbf{x}_0) \ne \mathbf{0}$ *and that* $\mathbf{p}$ *supports* $\{\mathbf{x} : f(\mathbf{x}) \ge f(\mathbf{x}_0)\}$ *at* $\mathbf{x}_0$. *Then* $\mathbf{p} = \alpha\, Df(\mathbf{x}_0)$ *for some* $\alpha \ne 0$.

Proof. Let $\mathbf{z} = Df - (\mathbf{p}\cdot Df/\|\mathbf{p}\|^2)\mathbf{p}$. Then $\mathbf{p}\cdot\mathbf{z} = 0$ and $Df\cdot\mathbf{z} = \|Df\|^2 - |\mathbf{p}\cdot Df|^2/\|\mathbf{p}\|^2$. By the Cauchy-Schwartz inequality, either $\mathbf{p}$ is proportional to Df (and we are done) or $Df\cdot\mathbf{z} > 0$.

In the case where $Df\cdot\mathbf{z} > 0$, the first-order Taylor approximation $f(\mathbf{x}_0) + \alpha\, Df\cdot\mathbf{z} > f(\mathbf{x}_0)$ shows

$$f(\mathbf{x}_0 + \alpha\mathbf{z}) > f(\mathbf{x}_0)$$

for small $\alpha > 0$. Continuity of f then yields

$$f(\mathbf{x}_0 + \alpha\mathbf{z} - \varepsilon\mathbf{p}) > f(\mathbf{x}_0)$$

for $\varepsilon > 0$ small. Apply Support Property II to obtain

$$\begin{aligned}\mathbf{p}\cdot\mathbf{x}_0 &\le \mathbf{p}\cdot(\mathbf{x}_0 + \alpha\mathbf{z} - \varepsilon\mathbf{p})\\ &= \mathbf{p}\cdot\mathbf{x}_0 - \varepsilon\|\mathbf{p}\|^2\\ &< \mathbf{p}\cdot\mathbf{x}_0.\end{aligned}$$

This contradiction shows $Df\cdot\mathbf{z} > 0$ is impossible. Therefore $\mathbf{p}$ is proportional to $Df(\mathbf{x}_0)$. ∎

21.23 Bordered Hessian Test for Quasiconcavity

Since $\lambda\mathbf{p} = Df$, we can instead use the bordered Hessian with Df. The bordered principal minors are multiplied by λ^2, so the signs are unchanged. This yields the following theorem.

Theorem 21.23.1. *Suppose $f\colon U \to \mathbb{R}$ is a $\mathcal{C}^2$ function on some open convex set $U \subset \mathbb{R}^m$ with $m > 1$. Consider the bordered Hessian*

$$H = \begin{bmatrix} 0 & Df \\ Df^T & D^2f \end{bmatrix} = \begin{bmatrix} 0 & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_m} \\ \frac{\partial f}{\partial x_1} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_m} & \frac{\partial^2 f}{\partial x_1 \partial x_m} & \cdots & \frac{\partial^2 f}{\partial x_m^2}. \end{bmatrix}$$

1. *If the bordered leading principal minors H_k obey $(-1)^{n-1}H_n > 0$ on U for $n = 3, \ldots, m+1$, then f is quasiconcave on U.*
2. *If all non-trivial bordered leading principal minors are negative on U, then f is quasiconvex on U.*
3. *If f is quasiconcave on U, then every k^{th} order bordered principal minor $\tilde{H}_k$ obeys $(-1)^{n-1}\tilde{H}_n \geq 0$ on U for $n = 3, \ldots, m+1$.*
4. *If f is quasiconvex on U, then all non-trivial bordered principal minors are non-positive on U.*

Proof. For the quasiconcave case, see Arrow and Enthoven (1961). Applying the result to $-f$ yields the quasiconvex case. ∎

November 13, 2020