3. Homotheticity and Separability

1/10/23

Economic models based on consumers often make a number of assumptions about consumer preferences. Chapter 2 examined many of such properties—including continuity, smoothness, convexity, and monotonicity. All these properties are routinely assumed of utility representations, usually without seriously compromising generality.

This chapter considers two properties that are more substantive, that have economic consequences. It is sometimes appropriate to use them in your models, and sometimes not. The two properties we focus on are homotheticity and separability.

Chapter Outline

3.0.1 More about Chapter Three

This chapter focuses on how restrictions on utility translate into restrictions on preferences, and vice-versa. Section one covers homothetic preferences, preferences that are independent of the scale of consumption. Homothetic preferences can generally be represented by homogeneous functions. Accordingly, we also study some of the more important properties of homogeneous functions. Much of this material was covered in the fall Math Methods course (ECO 7405).

Before section one is quite done, we take time to learn some math connected with Stokes' Theorem. We then do the rest of section one, proving a converse to Proposition 3.1.5, which uses Stokes' Theorem.

The third section (numbered 3.2) takes up the widely used additive separable utility, and characterizes utility functions that are ordinally equivalent to additive separable utility. The fourth section introduces the concept of induced preferences orders on commodity groups, where preferences over commodity bundles in one group are independent of consumption of goods from other commodity groups. Section five (numbered 3.4) shows how notions of separability may be applied directly to preference orderings, and relates it to the existence of an additive utility representation. Finally, the last section looks at the representation of preferences that are both homothetic and separable.

3.1 Homothetic Preferences REVIEW

Homothetic preferences are invariant under scalar multiplication, meaning that the preference map is unchanged when all consumption bundles are multiplied by the same number. More precisely, preferences are invariant under homothetic transformations centered on the origin.

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. This means that the shape of the indifference curves are preserved under scalar multiplication. All slopes remain unchanged.

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 shows 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 such scalar multiplication is always possible. Such scaling preserves the shapes of objects, including indifference surfaces. It only changes their scale.[1]

[1] This scaling is isotropic, the same in all directions. It is also possible to consider the effects of anisotropic scaling. This has not seen much use in utility theory, but is sometimes useful when homogeneous production is involved. See Boyd (1990a) for some applications.

3.1.1 Cones

REVIEW

To define homotheticity, we focus on cones, sets where positive scalar multiplication is always possible.

Cone. A set A is a *cone* if for every $\mathbf{x} \in A$ and $t > 0$, $t\mathbf{x} \in A$. Equivalently, A is a cone if $tA \subset A$ for every $t > 0$.

Cones are the natural setting for defining homotheticity. Examples of 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 an example of the last as it can also be written $\{\mathbf{x} = t_1(1, 0) + t_2(1, 1) : t_1, t_2 \geq 0\}$.

Cones can also be spiky. The set

$$\left\{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} = t(1, 2) \text{ or } \mathbf{x} = t(1, 1) \text{ or } \mathbf{x} = t(3, 1) \text{ for } t \geq 0\right\}$$

is also a cone, even though it consists of three unrelated rays from the origin, as in Figure 3.1.1.

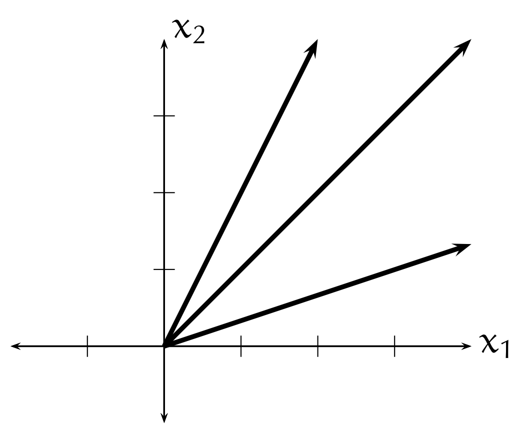

Figure 3.1.1: A spiky cone, made up by the union of three rays through the origin.

3.1.2 Homothetic Preferences

We are now ready to define homothetic preferences.

Homothetic Preferences and Functions. Preferences defined on a cone $\mathfrak{X}$ are *homothetic* if for every $t > 0$, $\mathbf{x} \succsim \mathbf{y}$ if and only if $t\mathbf{x} \succsim t\mathbf{y}$.

Similarly, we say a function f defined on a cone is *homothetic* if for every $t > 0$, $f(\mathbf{x}) \geq f(\mathbf{y})$ if and only if $f(t\mathbf{x}) \geq f(t\mathbf{y})$.

As you may already know, homothetic preferences yield Marshallian demands that are proportional to income. In case you don't know, we show it in Theorem 3.1.2.

Theorem 3.1.2. *Let prices and income be strictly positive, $\mathbf{p} \gg \mathbf{0}$ and $m > 0$. Suppose $\mathbf{x}(\mathbf{p}, m)$ is the set of points that maximize homothetic preferences $\succsim$ over the budget set $B(\mathbf{p}, m) = \{\mathbf{x} \in \mathbb{R}^m_+ : \mathbf{p}\cdot\mathbf{x} \leq m\}$. Then $\mathbf{x}(\mathbf{p}, m) = m\mathbf{x}(\mathbf{p}, 1)$*

Proof. Suppose $\mathbf{x}^* \in B(\mathbf{p}, m)$ with $\mathbf{x}^* \succsim \mathbf{x}$ for all $\mathbf{x} \in B(\mathbf{p}, m)$. Now $\mathbf{x}^*/m \in B(\mathbf{p}, 1)$. By homotheticity, $\mathbf{x}^*/m \succsim \mathbf{x}/m$ for all $\mathbf{x} \in B(\mathbf{p}, m)$. But $\mathbf{x}/m \in B(\mathbf{p}, 1)$ if and only if $\mathbf{x} \in B(\mathbf{p}, m)$, so $\mathbf{x}^*/m \succsim \mathbf{x}'$ for all $\mathbf{x}' \in B(\mathbf{p}, 1)$. It follows that $\mathbf{x}^*/m \in \mathbf{x}(\mathbf{p}, 1)$ if and only if $\mathbf{x}^* \in \mathbf{x}(\mathbf{p}, m)$. In other words, $\mathbf{x}(\mathbf{p}, m) = m\mathbf{x}(\mathbf{p}, 1)$. $\square$

3.1.3 Homogeneous Functions REVIEW

Homogeneous functions are one type of homothetic function.

Homogeneous Function. Let A be a cone in $\mathbb{R}^m$, a real-valued function is *homogeneous of degree* γ if

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

for every $\mathbf{x} \in A$ and $t > 0$.

The degree of homogeneity, γ, can be either positive or negative, and need not be an integer. 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.

The function $1/\|\mathbf{x}\|_2$ is homogeneous of degree -1 on the cone $\mathbb{R}^m_{++}$, the strictly positive orthant, but not defined on all of $\mathbb{R}^m_{+}$. Restricting the domain also allows us to consider $f(\mathbf{x})/g(\mathbf{x})$ where f is homogeneous of degree $\gamma_1 > 0$ and g is homogeneous of degree $\gamma_2 > 0$, both on $\mathbb{R}^m_{++}$. The quotient is homogeneous of degree $\gamma_1 - \gamma_2$ on $\mathbb{R}^m_{++}$.

When dealing with production rather than consumption, the degree of homogeneity determines returns to scale. 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. Returns to scale will be considered further when we study production.

3.1.4 Examples of Homogeneous Functions **REVIEW**

Functions such as $f(x_1, x_2) = x_1^2 + x_2^2$ and $f(x_1, x_2) = x_1 x_2$ are homogeneous of degree 2.

Many commonly used utility functions are homogeneous. The Cobb-Douglas utility 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 utility

$$u(\mathbf{x}) = [\delta x_1^{-\rho} + (1 - \delta) x_2^{-\rho}]^{\nu/\rho}$$

for $\nu > 0$ and $\rho > -1$, $\rho \neq 0$ is homogeneous of degree ν. The Leontief utility

$$u(\mathbf{x}) = \min_i \{x_i\}$$

is homogeneous of degree 1 on $\mathbb{R}_+^m$. In contrast, the quasi-linear utility

$$u(\mathbf{x}) = x_1 + x_2^{1/2}$$

is **not** homogeneous of any degree on $\mathbb{R}_+^2$.

3.1.5 Homothetic Preferences: Representation REVIEW

Any homogeneous utility function yields homothetic preferences. And since homotheticity is an ordinal property, any increasing transformation of a homogeneous utility function also defines a homothetic preference order. 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.

Homothetic Representation Theorem. *Suppose $\succsim$ is monotonic, homothetic, and continuous on $\mathbb{R}^m_+$. Then $\succsim$ has a utility representation ϕ that is homogeneous of degree 1. Moreover, any utility representation of $\succsim$ is then an increasing function F of ϕ, so utility $u = F \circ \phi$ where F is increasing and ϕ is homogeneous of degree one.*

3.1.6 Marginal Rates of Substitution and Homotheticity I REVIEW

One interesting property of homothetic functions is that marginal rate of substitutions are constant along rays through the origin.

Theorem 3.1.2. *Suppose* $f\colon \mathbb{R}^m_{++} \to \mathbb{R}$ *is homothetic and differentiable. If* $\mathrm{MRS}_{ij}(\mathbf{x})$ *exists, then* $\mathrm{MRS}_{ij}(\mathbf{x}) = \mathrm{MRS}_{ij}(t\mathbf{x})$ *for all* $t > 0$ *and* $\mathbf{x} \in \mathbb{R}^m_{++}$.

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

The fact that marginal rates of substitution are constant along rays through the origin has consequences for the shape of indifference curves. That is, they are homogeneous of degree zero. We illustrate this in Figure 3.1.3.

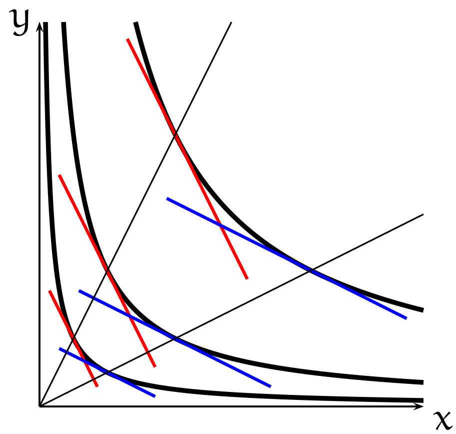

Figure 3.1.3: Here are three indifference curves for the Cobb-Douglas utility function $u(x, y) = \sqrt{xy}$. I've also draw two rays from the origin. Notice how the slope of the indifference curves remains the same along each ray. The red tangent lines all have slope -2, while the blue tangents have slope -0.5.

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.

3.1.7 Marginal Rates of Substitution and Homotheticity II REVIEW

There is a converse to Theorem 3.1.2, which we will state, but not prove.

Theorem 3.1.4. *Suppose* $f\colon \mathbb{R}^m_{++} \to \mathbb{R}$ *is* $\mathcal{C}^2$, $Df \gg 0$, *and* MRS_{ij} *is homogeneous of degree zero in* $\mathbf{x}$ *for every* i *and* j. *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_{ij} may not be defined for all pairs i and j. Lau does not do this, but the replacement for the MRS condition is that for all i and j, there are homogeneous of degree zero functions g_{ij} such that $\partial f/\partial x_i = g_{ij} \times (\partial f/\partial x_j)$. When $Df \gg 0$, this is equivalent to the marginal rates of substitution being homogeneous of degree zero in $\mathbf{x}$. □

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

3.1.8 Homogeneous Functions: Euler's Theorem REVIEW

When functions are not merely homothetic, but homogeneous, they have some important additional properties. The first one is given by Euler's Theorem, which relates homogeneous functions and their derivatives.

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* $\big[D_x f(\mathbf{x})\big]\mathbf{x} = \gamma f(\mathbf{x})$, *that is*

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

Nothing like Euler's Theorem need hold for functions that are merely homothetic. To see this, consider the homothetic function

$$f(\mathbf{x}) = \sum_i \alpha_i \ln x_i.$$

Then $Df = (\alpha_i/x_i)$, so

$$\big[D_x f(\mathbf{x})\big]\mathbf{x} = \left(\frac{\alpha_1}{x_1}, \dots, \frac{\alpha_m}{x_m}\right)\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \sum_i \alpha_i,$$

which cannot be written in the desired form because

$$\frac{\big[D_x f(\mathbf{x})\big]\mathbf{x}}{f(\mathbf{x})} = \frac{\sum_i \alpha_i}{f(\mathbf{x})} = \frac{\sum_i \alpha_i}{\sum_i \alpha_i \ln x_i}$$

is not constant.

3.1.9 Homogeneous Functions and Their Derivatives **REVIEW**

Homogeneous functions have another interesting property. Their derivatives are also homogeneous, with degree reduced by one.

Proposition 3.1.5. *Suppose* f *is a* $\mathcal{C}^1$ *function on* $\mathbb{R}^m_{++}$ *that is homogeneous of degree* γ *and obeys* $Df \neq 0$*. Then* Df *is a homogeneous function of degree* $(\gamma - 1)$.

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. Although the function f is homothetic, it 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 $\beta \neq -1$.

3.1.10 Indefinite Integrals of Homogeneous Functions REVIEW

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 $Df = bD\phi(\mathbf{x})/\phi(\mathbf{x})$, which is homogeneous of degree minus one.

The $\beta = -1$ case has a second type of complication when $m > 1$. This allows functions to be homogeneous of degree zero without being constant. One such function is $g(\mathbf{x}) = x_1/(x_1 + x_2)$. Its derivative is

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

which is clearly homogeneous of degree minus one.

Once we take the above possibilities into account, we can obtain a partial converse to Proposition 3.1.5.

3.1.11 Degree Zero Homogeneity and Monotonicity **REVIEW**

But first, we want to consider whether such functions can be utility functions. The logarithmic form f above poses no problem as such utility functions are always used. The function g is a problem though. It is not monotonic, and such functions are rarely used to describe utility. The problem is not peculiar to the function g.

Proposition 3.1.6. *Any function $f \in \mathcal{C}^1(\mathbb{R}^m_{++})$ with $m > 1$ that is homogeneous of degree zero is not 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. See Exercise 3.1.5.

29.8 Differential Forms and Stokes' Theorem

We will use Stokes' Theorem to prove the converse to Proposition 3.1.5. To make sense of Stokes' Theorem, we need to learn a bit about differential forms.[3]

Stokes' Theorem is an important result that originated in vector analysis, later generalized to an advanced calculus version, differential geometry, and eventually to geometric measure theory. It relates the integral of a differential form to an integral of its derivative. It is an extremely powerful generalization of the Fundamental Theorem of Calculus and of Green's Theorem. Before examining Stokes' Theorem, we need to learn some basics about differential forms.

[3] This section is based on section 29.8 of the text, as reflected in the numbering.

29.8.1 Derivatives vs. Differentials

When writing higher derivatives as linear functions, we found ourselves using k-tensors to write the k^{th} derivative. These could be represented using tensor products $dx_1 \otimes dx_2 \otimes dx_3$. If we tried to express these higher derivatives as higher total derivatives, we might even write them as ordinary products $dx_1\,dx_2\,dx_3$.

There's another place where we see products such as $dx_1\,dx_2\,dx_3$—integrals. Such expressions appear under the integral sign when we are integrating with respect to multiple variables. So we have to ask, do they mean the same thing? The answer is **no**!

What appears under the integral is not your ordinary tensor, it is a differential form. These products are properly regarded as exterior products rather than general tensor products. Differential forms are functions mapping into the exterior algebra generated by the basis $\{dx_1, dx_2, \dots, dx_m\}$ and its exterior products. The exterior product is designed to study lengths in $\mathbb{R}^1$, areas in $\mathbb{R}^2$, volumes in $\mathbb{R}^3$, and hyper-volumes in $\mathbb{R}^m$.

29.8.2 One-Forms and k-Forms

First derivatives are often written as differential forms. One simple type of *differential form* is the 1-*form*, where each term includes a single differential, one of the dx_i's. Any 1-form $\boldsymbol{\omega}$ can be written as

$$\boldsymbol{\omega} = \sum_{i=1}^{m} f_i(\mathbf{x})\, dx_i \tag{29.8.1}$$

where the f_i are functions.

A differential expression where each term involves products of k differentials, is called a k-form. Thus the form $\boldsymbol{\omega}$ given in Equation 29.8.1 is a 1-form because each term of the sum contains only a single dx_ℓ.

This definition also applies when $k = 0$, meaning that a 0-*form* involves no differentials. In other words, it is a real-valued function $f(\mathbf{x})$. We'll see that its exterior derivative coincides with the Fréchet derivative, re-imaged as a one-form.

We will primarily use differential forms inside integrals, although they have other uses. A k-dimensional integral will only involve k-forms. It is possible to combine different types of forms, creating a graded algebra, we will have no need to do so.

29.8.3 First Derivatives as Differential Forms

What do the differentials dx_i mean? Consider the i^{th} coordinate function given by $x_i(\mathbf{x}) = x_i$. Then the differential of x_i is its derivative, the row vector

$$\begin{aligned} dx_i &= Dx_i \\ &= \big(\delta_{ij}\big)_{j=1}^{m} \\ &= \big(0, 0, \cdots, 0, 1, 0, \cdots, 0\big) \end{aligned}$$

where the 1 occurs in the i^{th} coordinate. This means that $dx_i = \mathbf{e}_i^*$, so the differentials dx_i are a basis for the space of m-dimensional covectors, $(\mathbb{R}^m)^*$.

Now suppose $u\colon \mathbb{R}^m \to \mathbb{R}$ is a $\mathcal{C}^1$ function, its exterior derivative (differential) can be written[4]

$$du = \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_m}\right) = \sum_i \frac{\partial u}{\partial x_i}\, dx_i.$$

We have written Du as a differential 1-form, du. Differential forms that are derivatives of real-valued functions are called *exact differentials*.

[4] For functions mapping into $\mathbb{R}$, the differential 1-form and Fréchet derivative are the same row vector for such functions. That's generally not the case for higher derivatives.

29.8.4 Bivectors

When dealing with k-dimensional integrals, we need to use k-forms, which involve a special type of product of k differentials, the *exterior product*. The simplest exterior product involves two vectors or covectors. The exterior product can be regarded as a special type of tensor product that is *alternating*—the sign of the exterior product flips when we reverse the order of the vectors.

The *exterior* or *wedge product* of two vectors, $x \wedge y$ is called a *bivector*.[5] Similarly, exterior products can be applied to covectors, yielding *bicovectors*.[6]

The exterior product has three important properties.

Exterior Products. The *exterior or wedge product* is

1. Associative: $x \wedge (y \wedge z) = (x \wedge y) \wedge z$.
2. Alternating: $x \wedge y = -y \wedge x$, and
3. Two-linear: $x \wedge y$ is separately linear in both x and y.

The associative law tells us that exterior products of 3 vectors, 4 vectors, etc. do not depend on how we associate them.

Because the wedge product is alternating, $x \wedge x = 0$. In this context 0 refers to the zero bivector, not the zero vector.

[5] The term 2-*blade* is sometimes used.

[6] Mixed exterior products involving vectors and covectors are also possible.

29.8.5 Exterior Product of Linearly Dependent Vectors is Zero

The wedge product of any linearly dependent set of vectors is zero.

Theorem 29.8.1. *Suppose* $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ *is a linearly dependent set. Then*

$$\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_n = \mathbf{0}$$

where $\mathbf{0}$ *is the zero* n*-vector.*

Proof. Without loss of generality, we can assume that we can write $\mathbf{x}_1$ as a linear combination of the other $\mathbf{x}_i$. That is, $\mathbf{x}_1 = \sum_{i=2}^{n} \alpha_i \mathbf{x}_i$ for some α_i. Then

$$\begin{aligned} \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_n &= \sum_{i=2}^{n} \alpha_i \mathbf{x}_i \wedge (\mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_n) \\ &= \mathbf{0} \end{aligned}$$

since for each $i = 2, \dots, m$, the i$^{\text{th}}$ term of the sum is zero due to the repeated vector $\mathbf{x}_i$ in the wedge product $\mathbf{x}_i \wedge (\mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_n)$. $\square$

29.8.6 The Bivectors form a Vector Space

The set of bivectors in $\mathbb{R}^m$ is a vector space because the exterior product is bilinear. We denote space of bivectors by $\mathbb{R}^m \wedge \mathbb{R}^m$. Since it is a vector space, it must have a basis. We can build one based on the standard basis.

If we were dealing with covectors rather than vectors, we could use $\{dx_i\}$ to form our basis, giving us differential forms.

We can also form trivectors, etc. Each set of k-forms are a vector space. However, this only works for $k = 1, \dots, m$. There is no point to forming $(m+1)$-fold exterior products of vectors in $\mathbb{R}^m$. Such vectors must make a linearly dependent set, and by Theorem 29.8.1, their exterior product is zero.

29.8.7 Grassman's Idea

Now that we know how to find a basis for the bivectors (or bicovectors) we still have some questions concerning their use and meaning. Why are bicovectors used in integrals? Can we represent bivectors and bicovectors as geometric objects?

We deal with the geometric interpretation first as it will answer the other question for us. Grassmann's original idea (1844, 1861) was to construct higher dimensional shapes by extension from lower dimensional shapes. The exterior product accomplishes this extension. The bivector extends two vectors into a two dimensional shape, specifically, a parallelogram. Moreover, the bivector includes both the area of the parallelogram and its orientation.[7]

One way to think about this is that $\mathbf{x}\wedge\mathbf{y}$ starts with the vectors $\mathbf{x}$ and $\mathbf{y}$ at zero, and then sweeps $\mathbf{y}$ along $\mathbf{x}$, creating a parallelogram at the origin, as illustrated in Figure 29.8.2. The parallelogram represents $\mathbf{x}\wedge\mathbf{y}$.[8]

[7] The process was described in similar terms by Grassman himself. Grassman (1844, quoted in Crowe 1985, pg. 70) stated "We go from the vector to a spatial form of higher order when we allow the entire vector, that is each point of the vector, to describe another vector which is heterogeneous to the first, so that all points construct an equal vector. The surface area produced this way has the form of a parallelogram. Two such surface areas which belong to the same plane are designated as equal if the direction of the moved vector lies on the same side...of the motion. When in the two cases the corresponding vectors lie on opposite sides, then the surface areas are designated unequal."

[8] The same principle is used in higher dimensions. As Crowe puts it (1985, pg. 72) "Grassmann stated that a product such as $\mathfrak{a}.\mathfrak{b}.\mathfrak{c}...$ was to mean that the vector $\mathfrak{a}$ first moves along $\mathfrak{b}$ (as before), then the resultant oriented area would move along $\mathfrak{c}$, and so on though orders higher than the third."

29.8.8 Bivectors Illustrated

Below, we show how the vector y slides along x to sweep out an area (light blue).

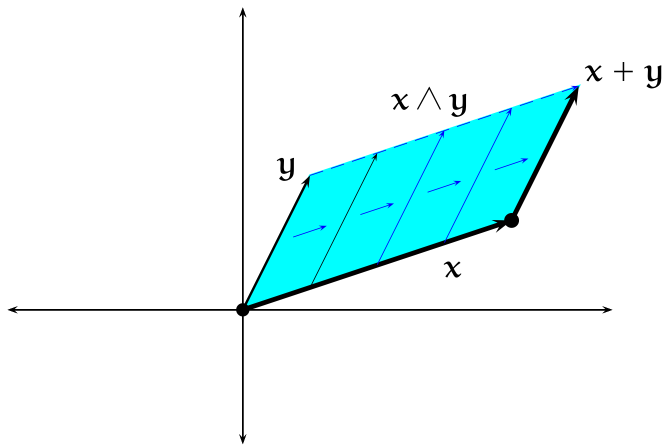

Figure 29.8.2: The parallelogram swept out has area $\|x \wedge y\|$ (shaded).

29.8.9 Bivector Area and Norm

The area of the associated parallogram is the norm of the bivector. This is easiest to see in $\mathbb{R}^2$. Consider any two vectors $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^2$. We can write $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$. Then

$$\begin{aligned}
\mathbf{x} \wedge \mathbf{y} &= x_1y_1\, \mathbf{e}_1 \wedge \mathbf{e}_1 + x_1y_2\, \mathbf{e}_1 \wedge \mathbf{e}_2 \\
&\qquad + x_2y_1\, \mathbf{e}_2 \wedge \mathbf{e}_1 + x_2y_2\, \mathbf{e}_2 \wedge \mathbf{e}_2. \\
&= (x_1y_2 - x_2y_1)\, \mathbf{e}_1 \wedge \mathbf{e}_2
\end{aligned}$$

where we used the fact that the exterior product is alternating to simplify the expression.

The coefficient on $\mathbf{e}_1 \wedge \mathbf{e}_2$ is

$$x_1y_2 - x_2y_1 = \det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

As is well-known, the absolute value of this determinant is the area of the parallelogram generated by $\mathbf{x}$ and $\mathbf{y}$. The determinant here is positive, but if we reversed the order of $\mathbf{x}$ and $\mathbf{y}$, its sign would become negative.

Since $\mathbb{R}^2 \wedge \mathbb{R}^2$ is two-dimensional, we can write

$$\|\mathbf{x} \wedge \mathbf{y}\| = \left|\det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\right|.$$

29.8.10 Orientation and Bivectors

The sign of $x \wedge y$, carries information about its orientation, whether it is in standard orientation (counter-clockwise or right-handed) or reversed orientation (clockwise or left-handed). Counter-clockwise orientation is indicated by a positive sign, clockwise orientation by a negative sign. In $\mathbb{R}^2$ the bivector $x \wedge y$ will be one of $+\|x \wedge y\|\, e_1 \wedge e_2$ or $-\|x \wedge y\|\, e_1 \wedge e_2$, depending on the orientation.

We think of the former as indicating a positive orientation, the latter a negative orientation. The orientations are associated with rotations, and the usual convention (the right-hand rule) associates counter-clockwise rotations with a positive orientation, and clockwise rotations with a negative orientation. The orientation will be important when using multi-covectors to define integrals.

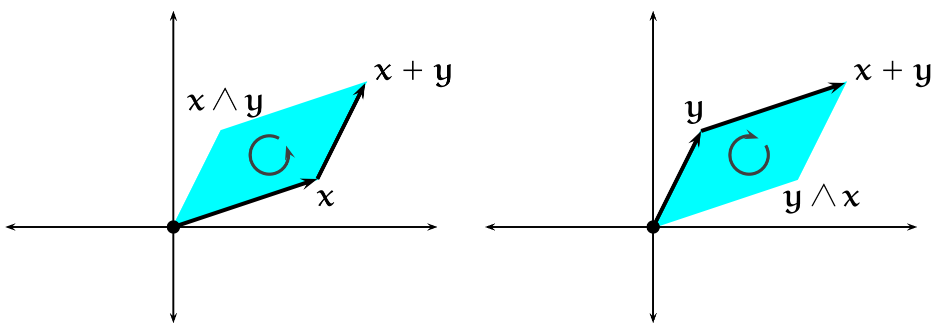

Figure 29.8.3: In the left panel, the bivector $x \wedge y$ defines a parallelogram in $\mathbb{R}^2$ with area $\|x \wedge y\|$ (shaded) and in standard (positive, counter-clockwise) orientation. Here the vector y bends from x in a counter-clockwise direction.

In the right panel, the bivector $x \wedge y$ defines the same parallelogram, but with reversed (negative, clockwise) orientation, reflecting the fact that $y \wedge x = -x \wedge y$. Here x bends from y in clockwise fashion.

29.8.11 Integrals and Bivectors

The norm of a bivector tells us the area of its associated parallelogram just as the norm of a vector gives us the length of its line segment from the origin. We using covectors, dx is used to measure the length of line segments, while $dx \wedge dy$ measures parallelograms.

The connection with integration is that to integrate a function over a line segment, Riemann taught us to divide the segement into smaller and smaller pieces. The integral is then the limit of the Riemann sums formed from the pieces as the pieces shrink to zero, if the limit exists. If not, the function is not Riemann integrable.

Because parallelograms fit together without gaps or overlaps, we can divide the region of integration into smaller and smaller parallelograms, or bivectors. We form the analogous Riemann sums and take the limit, which exists if the function is Riemann integrable. We can also use this strategy on functions defined over two-dimensional differentiable manifolds.

29.8.12 Integrals and Trivectors, etc.

To integrate over a three-dimensional region, we use trivectors, which define parallelepipeds. In that case, the norm of the trivector is the volume of the associated parallelepiped.

This generalizes to exterior products of k vectors, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, which define k-dimensional *parallelotopes* by

$$\left\{\mathbf{x} : \mathbf{x} = \sum_{i=1}^{k} t_i\mathbf{x}_i, 0 \le t_i \le 1\right\}.$$

The k-dimensional volume is again the norm of the wedge product. This is why we use k-forms to integrate over k-dimensional manifolds.

Keep in mind that we're using signed areas, so the integrals depend on the direction we use to traverse the region we integrate over. You've seen this when you took calculus. Definite integrals on $[a, b] \subset \mathbb{R}$ change sign when you reverse the direction in which you traverse the interval:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

29.8.13 A Basis for Bivectors

Any bivector in $\mathbb{R}^2 \wedge \mathbb{R}^2$ can be written

$$\begin{aligned}
\boldsymbol{x} \wedge \boldsymbol{y} &= (x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2) \wedge (y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2) \\
&= x_1\boldsymbol{e}_1 \wedge (y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2) + x_2\boldsymbol{e}_2 \wedge (y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2) \\
&= x_1y_1\, \boldsymbol{e}_1 \wedge \boldsymbol{e}_1 + x_1y_2\, \boldsymbol{e}_1 \wedge \boldsymbol{e}_2 + x_2y_1\, \boldsymbol{e}_2 \wedge \boldsymbol{e}_1 + x_2y_2\, \boldsymbol{e}_2 \wedge \boldsymbol{e}_2 \\
&= (x_1y_2 - x_2y_1)\, \boldsymbol{e}_1 \wedge \boldsymbol{e}_2 \\
&= \det(\boldsymbol{x}, \boldsymbol{y})\, \boldsymbol{e}_1 \wedge \boldsymbol{e}_2
\end{aligned}$$

where the fact that the wedge product is alternating has been used to simplify the product. This implies that $\mathbb{R}^2 \wedge \mathbb{R}^2$ is a one-dimensional vector space. Similarly, the alternating property means that $\mathbb{R}^m \wedge \mathbb{R}^m$ is an $m(m-1)/2$-dimensional vector space.

More generally, $\wedge^k\mathbb{R}^m$ has dimension

$$\binom{m}{k} = \frac{m!}{(m-k)!k!},$$

the number of combinations of m basis vectors taken k at a time.

It follows that $\wedge^m\mathbb{R}^m$ is a one-dimensional vector space, as we saw with $\mathbb{R}^2 \wedge \mathbb{R}^2$. That means that $\boldsymbol{e}_1 \wedge \cdots \wedge \boldsymbol{e}_m$ (for vectors) or $dx_1 \wedge \cdots \wedge dx_m$ (for covectors) can be used as its only basis element.

29.8.14 Bivector Products are not Unique

There is a complication when trying to show a set of bivectors is linearly independent. There may be many ways to write the same bivector. Suppose $\mathbf{a}$ and $\mathbf{b}$ are linearly independent. Form the exterior product

$$\begin{aligned}
&\big[(1-s)\mathbf{a}+s\mathbf{b}\big]\wedge\big[(1-t)\mathbf{a}+t\mathbf{b}\big]\\
&\qquad\qquad = (1-s)\mathbf{a}\wedge t\mathbf{b} + (1-t)\mathbf{b}\wedge s\mathbf{a}\\
&\qquad\qquad = \big[t-st-s+st\big]\mathbf{a}\wedge\mathbf{b}\\
&\qquad\qquad = \big[t-s\big]\mathbf{a}\wedge\mathbf{b}.
\end{aligned}$$

So whenever $t-s=1$, the exterior product of $(1-s)\mathbf{a}+s\mathbf{b}$ and $(1-t)\mathbf{b}\wedge s\mathbf{a}$ is $\mathbf{a}\wedge\mathbf{b}$.

This sort of thing could cause trouble when trying to define a basis for the space of bivectors in $\mathbb{R}^m$. However, it is not a problem when building our basis from a basis for $\mathbb{R}^m$. Here's how it works for the standard basis. Suppose we have a bivector

$$\mathbf{x} = \sum_{i<j} x_{ij}\,\mathbf{e}_i\wedge\mathbf{e}_j = \mathbf{0}.$$

We restrict ourselves to $i<j$ because $\mathbf{e}_i\wedge\mathbf{e}_j = -\mathbf{e}_j\wedge\mathbf{e}_i$. If we take the exterior product of the this with the any $\mathbf{e}_k$ with $k\neq i,j$, all of the terms other than the ij term will have a wedge product containing repeated basis vectors. Those must be zero. That leaves us with $\pm x_{ij}\,\mathbf{e}_k\wedge\mathbf{e}_i\wedge\mathbf{e}_j$, which is zero since the sum is zero. This allows us to show that $\{\mathbf{e}_i\wedge\mathbf{e}_j : i<j\}$ is a linearly independent set.

29.8.15 k-fold Exterior Products

We noted earlier that dimension of the k-fold exterior product of $\mathbb{R}^m$ is given by $\dim \bigwedge^k \mathbb{R}^m = m!/k!(m-k)!$. This means that

$$\begin{aligned} \dim \textstyle\bigwedge^0 \mathbb{R}^m &= 1 = \dim \textstyle\bigwedge^m \mathbb{R}^m \\ \dim \textstyle\bigwedge^1 \mathbb{R}^m &= m = \dim \textstyle\bigwedge^{m-1} \mathbb{R}^m \\ \dim \textstyle\bigwedge^2 \mathbb{R}^m &= m(m-1) = \dim \textstyle\bigwedge^{m-2} \mathbb{R}^m \\ &\text{etc.} \end{aligned}$$

This suggests these pairs of spaces are isomorphic. There is a natural isomorphism that depends on the inner product and the basis. For simplicity, we restrict our attention to $\mathbb{R}^m$ with the Euclidean inner product and standard basis.

We introduce a shorthand notation for k-fold wedge products of vectors. For any ordered index set $I = \{i_1 < \cdots i_k\} \subset \{1, \dots, m\}$, define the complementary ordered index set $I' = \{i'_1 < \cdots < i'_{m-k}\}$ such that $I \cup I' = \{1, \dots, m\}$ and define $\#I = k$ (so $\#I' = m - k$). Then denote the k-fold exterior product of $\{\boldsymbol{e}_{i_k} : i_k \in I\}$ by

$$\boldsymbol{e}_I = \boldsymbol{e}_{i_1} \wedge \boldsymbol{e}_{i_2} \cdots \wedge \boldsymbol{e}_{i_k}$$

29.8.16 The Hodge Star Operator

The *Hodge star operator* is the linear operator from each $\bigwedge^k \mathbb{R}^m$ to $\bigwedge^{m-k} \mathbb{R}^m$ such that

$$\star \boldsymbol{e}_I = \operatorname{sgn} \sigma(II')\, \boldsymbol{e}_{I'}$$

where $\operatorname{sgn} \sigma(II')$ is the sign of the permutation $II' = i_1 i_2 \cdots i_k i'_1 \cdots i'_{m-k}$. That is, $\sigma(II')$ is -1 for permutations using an odd number of interchanges and $+1$ for permutations using an even number of interchanges.

This means that in $\mathbb{R}^3$,

$$\begin{aligned}
\star \boldsymbol{e}_1 &= \boldsymbol{e}_2 \wedge \boldsymbol{e}_3, \\
\star \boldsymbol{e}_2 &= -\boldsymbol{e}_1 \wedge \boldsymbol{e}_3, \\
\star \boldsymbol{e}_3 &= \boldsymbol{e}_1 \wedge \boldsymbol{e}_2.
\end{aligned}$$

Similarly, $\star(\boldsymbol{e}_1 \wedge \boldsymbol{e}_2) = \boldsymbol{e}_3$, $\star(\boldsymbol{e}_1 \wedge \boldsymbol{e}_3) = -\boldsymbol{e}_2$, and $\star(\boldsymbol{e}_2 \wedge \boldsymbol{e}_3) = \boldsymbol{e}_1$.

29.8.17 Inner Products of k-vectors **SKIPPED**

If we have an inner product $\mathbf{x} \cdot \mathbf{y}$ on $\mathbb{R}^m$, we can extend the inner product to $\bigwedge^k \mathbb{R}^m$ as follows. If $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 \wedge \cdots \wedge \boldsymbol{\alpha}_k$ and $\boldsymbol{\beta} = \boldsymbol{\beta}_1 \wedge \cdots \wedge \boldsymbol{\beta}_k$, we define $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = \det\left(\boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j\right)$. Since the linear product is linear, this is easily extended to all of $\bigwedge^k \mathbb{R}^m$.

The *Hodge star operator* can be defined generally in terms of the inner product. If we work in terms of the Euclidean inner product and standard orthonormal basis on $\mathbb{R}^m$, $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ and set $\boldsymbol{\omega} = \boldsymbol{e}_1 \wedge \cdots \wedge \boldsymbol{e}_m$, then given $\boldsymbol{\beta} \in \bigwedge^k \mathbb{R}^m$, the *Hodge star* $\star\boldsymbol{\beta}$ is the unique element of $\bigwedge^{m-k} \mathbb{R}^m$ such that

$$\boldsymbol{\alpha} \wedge (\star\boldsymbol{\beta}) = (\boldsymbol{\alpha} \cdot \boldsymbol{\beta})\, \boldsymbol{\omega}$$

for every $\boldsymbol{\alpha} \in \bigwedge^k \mathbb{R}^m$.

29.8.18 The Vector Cross Product SKIPPED

The Hodge star allows us to write the 3-dimensional vector cross product in terms of the wedge product.

Recall that the *vector cross product* is defined on $\mathbb{R}^3$ by

$$\boldsymbol{x} \times \boldsymbol{y} = (x_2y_3 - y_2x_3)\,\boldsymbol{e}_1 - (x_1y_3 - y_3x_1)\,\boldsymbol{e}_2 + (x_1y_2 - y_1x_2)\,\boldsymbol{e}_3.$$

Now compute

$$\begin{aligned}\boldsymbol{x} \wedge \boldsymbol{y} &= x_1y_2\,\boldsymbol{e}_1 \wedge \boldsymbol{e}_2 + x_1y_3\,\boldsymbol{e}_1 \wedge \boldsymbol{e}_3 + x_2y_1\,\boldsymbol{e}_2 \wedge \boldsymbol{e}_1 + x_2y_3\,\boldsymbol{e}_2 \wedge \boldsymbol{e}_3\\ &\qquad + x_3y_1\,\boldsymbol{e}_3 \wedge \boldsymbol{e}_1 + x_3y_2\,\boldsymbol{e}_3 \wedge \boldsymbol{e}_2\\ &= (x_1y_2 - x_2y_1)\,\boldsymbol{e}_1 \wedge \boldsymbol{e}_2 + (x_1y_3 - x_3y_1)\,\boldsymbol{e}_1 \wedge \boldsymbol{e}_3\\ &\qquad + (x_2y_3 - x_3y_2)\,\boldsymbol{e}_2 \wedge \boldsymbol{e}_3.\end{aligned}$$

Then apply the Hodge star operator, obtaining

$$\begin{aligned}\star(\boldsymbol{x} \wedge \boldsymbol{y}) &= (x_1y_2 - x_2y_1)\,\boldsymbol{e}_3 - (x_1y_3 - x_3y_1)\,\boldsymbol{e}_2\\ &\qquad + (x_2y_3 - x_3y_2)\,\boldsymbol{e}_1\\ &= \boldsymbol{x} \times \boldsymbol{y}.\end{aligned}$$

Although the vector cross product only exists in $\mathbb{R}^3$, we can consider $\star(\boldsymbol{x} \wedge \boldsymbol{y})$ in other dimensions. For example, in $\mathbb{R}^2$,

$$\begin{aligned}\star(\boldsymbol{x} \wedge \boldsymbol{y}) &= \star\left(\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\,\boldsymbol{e}_1 \wedge \boldsymbol{e}_2\right)\\ &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.\end{aligned}$$

29.8.19 Differential Forms

We can now give differential forms a proper definition. Let $(\mathbb{R}^m)^*$ denote the space of m-dimensional covectors, the dual of $\mathbb{R}^m$.

Differential Form. Let U be an open subset of $\mathbb{R}^m$. A $\mathcal{C}^k$ *differential form of order* ℓ *on* U is a $\mathcal{C}^k$ mapping from U to $\bigwedge^\ell(\mathbb{R}^m)^*$, the ℓ-fold exterior product of the dual of $\mathbb{R}^m$. We can write it as

$$\sum_{\{I:\#I=\ell\}} f_I(\mathbf{x})\, dx_I$$

where each $f_I\colon U \to \mathbb{R}$ is $\mathcal{C}^k$.[9]

The simplest such differential form is a zero form, where I is empty and the form is a real-valued function. We have also encountered 1-forms, where every set I has a single element. E.g., I is successively $\{1\}, \dots, \{m\}$, yielding

$$\boldsymbol{\omega} = f_1(\mathbf{x})\, dx_1 + f_2(\mathbf{x})\, dx_2 + \cdots + f_m(\mathbf{x})\, dx_m.$$

[9] The functions f_I are functions of all m variables, not just those in I.

29.8.20 The Exterior Derivative 01/12/23

Differential forms have derivatives. To compute them, we use a special type of derivative called the *exterior derivative*. The exterior derivative of a k-form is a $(k+1)$-form. The exterior derivative of a 0-form, a smooth function f mapping $\mathbb{R}^m$ into $\mathbb{R}$, is just its derivative, which we write as[10]

$$df = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}\, dx_i.$$

Exterior Derivative. The *exterior derivative* is the unique linear mapping from k-forms to $(k+1)$-forms that has the following properties:

1. df is the differential of f for any 0-form f.
2. $d(df) = 0$ for any 0-form f.
3. If α is a ℓ-form and β a differential form, then $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^m(\alpha \wedge d\beta)$.

[10] Recall that df is a covector, and that the dx_i form a basis for covectors.

29.8.21 Calculating Exterior Derivatives I

Let's calculate the exterior derivative of a 1-form. Let

$$\omega = \sum_{i=1}^{m} f_i \, dx_i$$

where the f_i are all $\mathcal{C}^1$ functions. To take the exterior derivative of such a form, think of the i^{th} term as the 0-form f_i times the 1-form dx_i. We focus on the term $d(f_i \, dx_i)$.

$$d(f_i \wedge dx_i) = df_i \wedge dx_i + (-1)^m f_i d(dx_i) = df_i \wedge dx_i$$

since $d(dx_i) = 0$. Now use the fact that $df_i = \sum_{j=1}^{m} (\partial f_i / \partial x_j) \, dx_j$ to obtain

$$d(f_i \wedge dx_i) = \left(\sum_{j=1}^{m} \frac{\partial f_i}{\partial x_j} dx_j \right) \wedge dx_i.$$

29.8.22 Calculating Exterior Derivatives II

To find $d\omega$, we add up the $d(f_i \wedge dx_i)$ terms.

$$
\begin{aligned}
d\omega &= \sum_{i=1}^{m} df_i \wedge dx_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{m}\frac{\partial f_i}{\partial x_j}dx_j\right)\wedge dx_i \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial f_i}{\partial x_j}\,dx_j \wedge dx_i \\
&= \sum_{i<j}\frac{\partial f_i}{\partial x_j}\,dx_j \wedge dx_i + \sum_{i>j}\frac{\partial f_i}{\partial x_j}\,dx_j \wedge dx_i \\
&= \sum_{i<j}\frac{\partial f_i}{\partial x_j}\,dx_j \wedge dx_i + \sum_{j>i}\frac{\partial f_j}{\partial x_i}\,dx_i \wedge dx_j \\
&= \sum_{i<j}\frac{\partial f_i}{\partial x_j}\,dx_j \wedge dx_i - \sum_{i<j}\frac{\partial f_j}{\partial x_i}\,dx_j \wedge dx_i \\
&= \sum_{j=1}^{m}\sum_{i=1}^{j-1}\left(\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i}\right)dx_j \wedge dx_i
\end{aligned}
\tag{29.8.2}
$$

The antisymmetry made the $i=j$ terms disappear in the third line, allowing us to divide the sum into two parts. We switched indicies in the second sum of the fourth line, and applied antisymmetry again in the fifth line.

Applying the same procedure to 2-forms, 3-forms, etc. allows us to inductively define the exterior derivatives of any k-form.

29.8.23 Some Exterior Derivatives

For example, consider the 2-form

$$\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy.$$

We think of the terms of ω as products of 0-forms and 2-forms, e.g., $\alpha = x$ and $\beta = dy \wedge dz$. The derivative of the 2-form always involves a d^2 term and so is zero. We than have

$$\begin{aligned} d\omega &= dx \wedge dy \wedge dz + dy \wedge dz \wedge dx + dz \wedge dx \wedge dy \\ &= dx \wedge dy \wedge dz - dy \wedge dx \wedge dz - dx \wedge dz \wedge dy \\ &= dx \wedge dy \wedge dz + dx \wedge dy \wedge dz + dx \wedge dy \wedge dz \\ &= 3\,dx \wedge dy \wedge dz. \end{aligned}$$

Different types of differential forms can be distinguished by the number of wedge products involved. The Fréchet derivative of a function from $\mathbb{R}^m \to \mathbb{R}$ can be regarded as creating a *1-form*.

Of course, the wedge product only allows products to go as far as m-forms. Any exterior derivative past that is 0, reflecting the fact that in $\mathbb{R}^m$, it is impossible to obtain any $(m+1)$-dimensional forms.

Another useful fact is that $d(dx_I) = 0$ for any index set I. You can show this by repeatedly applying rule (3) of the exterior derivative to any dx_I. For example, when $m \geq 3$,

$$\begin{aligned} d\big(dx_1 \wedge dx_2 \wedge dx_3\big) &= d^2x_1 \wedge dx_2 \wedge dx_3 - dx_1 \wedge d(dx_2 \wedge dx_3) \\ &= -dx_1 \wedge (d^2x_2 \wedge dx_3 - dx_2 \wedge d^2x_3) \\ &= 0. \end{aligned}$$

29.8.24 Closed and Exact Differentials

Differential forms ω with $d\omega = 0$, are called *closed differentials*.. An important fact is that exact differentials—exterior derivatives of functions—are closed differential forms if the original function is $\mathcal{C}^2$.

Theorem 29.8.4. *Let $u \in \mathcal{C}^2$. Then du is a closed differential form, that is, $d(du) = 0$.*

Proof. Suppose u is a $\mathcal{C}^2$ function and consider the differential du. We use equation (29.8.2) with $f_j = \partial u/\partial x_j$ to obtain

$$d(du) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} \left[\frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_j} \right) - \frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_i} \right) \right] dx_i \wedge dx_j. \qquad (29.8.3)$$

Now u is twice continuously differentiable, so the cross-partial terms are equal:

$$\frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_i} \right).$$

But then they cancel out when we substitute them in equation (29.8.3). That means $d(du) = 0$, establishing that du is closed. □

29.8.25 Wedge Products and Determinants **SKIPPED**

The wedge product has some other uses. Suppose $\mathbf{df}_j = \sum_{k=1}^{m} f_{jk}\,dx_k$ for $j = 1, \dots, m$. We can then write

$$\bigwedge_{i=1}^{m} \mathbf{df}_i = A(\mathbf{f}_1, \dots, \mathbf{f}_m)\,dx_1 \wedge dx_2 \cdots \wedge dx_m.$$

for some multilinear function A. Moreover, if we interchange any two distinct $\mathbf{df}_j$ and $\mathbf{df}_k$, the sign flips (A is alternating). Finally, if $\mathbf{df}_j = \sum_{i=1}^{m} dx_i$, some calculation reveals that $\wedge_{j=1}^{m}\mathbf{df}_j = \wedge_{j=1}^{m} dx_j$, so $A(\mathbf{e}_1, \dots, \mathbf{e}_m) = 1$. In other words, $A(\mathbf{f}_1, \dots, \mathbf{f}_m)$ must be the determinant:

$$A(\mathbf{f}_1, \dots, \mathbf{f}_m) = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ \vdots & & \ddots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mm} \end{vmatrix}$$

where $f_{ij} = \frac{\partial f_i}{\partial x_j}$.

29.8.26 Stokes' Theorem

Stokes' Theorem is an important result from advanced calculus and geometric measure theory that relates integrals of differential forms and of their derivatives. It is an extremely powerful generalization of the Fundamental Theorem of Calculus and of Green's Theorem.

We now consider Stokes' Theorem.[11]

Stokes' Theorem. *If* M *is a* n*-dimensional manifold with boundary* ∂M *and* $\boldsymbol{\omega}$ *is a* $\mathcal{C}^1$ $(n-1)$*-form with compact support on* M*, then* $\int_M d\boldsymbol{\omega} = \int_{\partial M} \boldsymbol{\omega}$.

There are more general versions of Stoke's Theorem, where the manifold is actually a bunch of manifolds pasted together. We will apply Stokes' Theorem to curves. More general versions of Stokes' Theorem allow us to apply it to manifolds where the boundary consists of piecewise smooth curves with 90° corners. Dealing with all of the details would take us far afield, so we won't. Nonetheless, we will use Stokes' Theorem on manifolds with corners.[12]

[11] See Buck and Buck (1978, sec. 9.4), Fleming (1965, pg. 273) or Spivak (1965, pg. 102).

[12] A version that properly handles corners can be found in sec. III.14 of Whitney (1957). The formula is essentially the same.

29.8.27 Stokes and the Fundamental Theorem of Calculus

Among other things, Stokes' Theorem generalizes the Fundamental Theorem of Calculus. Let $\omega = f$ and M be the interval $[a, b]$. Then its boundary is $\partial M = \{a, b\}$. Normally, when we integrate f over a finite set, we add the values of f at the points in the set, with a weight of $+1$ on the ending point b and -1 on the starting point a.

Here we must take the orientation into account. When the orientation is negative, we take b as the starting point and a as the ending point. Equivalently, we assign a sign of $+1$ to the integral when the interval is oriented from a to b and -1 when it is oriented from b to a.

Figure 29.8.1: The left interval has postive orientation and the right interval has negative orientation.

Since M has a positive orientation, we assign a weight of $+1$ to the ending point b and -1 to the starting point a. The integral then becomes

$$\int_{\partial M} f(x) = f(b) - f(a)$$

Since $d\omega = f'\,dx$, Stokes' Theorem then tells us that

$$\int_a^b f'\,dx = \int_M f'\,dx = \int_{\partial M} f = f(b) - f(a).$$

Notice that our orientation convention ensures that

$$\int_a^b f'(x)\,dx = -\int_b^a f'(x)\,dx.$$

29.8.28 Stokes' Theorem and Line Integrals I

In our application, we will have a $\mathcal{C}^2$ function and consider its differential du. As an exact differential, du is closed by Theorem 29.8.4.

We will be interested in computing integrals of the form $\int_\gamma du$ where γ is a path in $\mathbb{R}^2$. We first show that the integral doesn't depend on the path taken. Let γ_1 and γ_2 be two non-intersecting paths between $\mathbf{x}$ and $\mathbf{y}$. Define γ by following γ_1 from $\mathbf{x}$ to $\mathbf{y}$, and then γ_2 in reverse back to $\mathbf{x}$. In other words, $\gamma\colon [0,1] \to \mathbb{R}^2$ is defined by

$$\gamma(t) = \begin{cases} \gamma_1(2t) & \text{for } 0 \le t \le 1/2 \\ \gamma_2(2-2t) & \text{for } 1/2 \le t \le 1. \end{cases}$$

Because $\gamma_1(1) = \mathbf{y} = \gamma_2(1)$, the curve γ is unambiguously defined at $t = 1/2$. Further $\gamma(0) = \gamma_1(0) = \mathbf{x}$ and $\gamma(1) = \gamma_2(0) = \mathbf{x}$, so the curve circles back to its starting point.

Let M be the area enclosed by γ. For reasonable γ, the path γ is the boundary of M.[13]

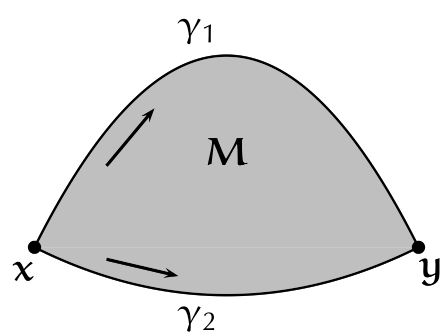

Figure 29.8.5: Both γ_1 and γ_2 are curves from $\mathbf{x}$ to $\mathbf{y}$. We paste them together by traversing γ_1 and then travelling in reverse along γ_2 (reversing its orientation). The arrows indicate the original orientation of γ_1 and γ_2. The combined curve enclosed M, a manifold with boundary γ.

[13] The term "reasonable" glosses over a lot of details. It includes the case where γ is piecewise smooth.

29.8.29 Stokes' Theorem and Line Integrals II

Stokes' Theorem now tells us that

$$\int_\gamma du = \int_M d(du) = 0$$

since du is closed. But now $\int_\gamma du = \int_{\gamma_1} du - \int_{\gamma_2} du = 0$, showing that both paths from x to y have the same integral:

$$\int_{\gamma_1} du = \int_{\gamma_2} du$$

Thus if $u \in \mathcal{C}^2$ and γ is a well-behaved path from x to y, we now know that $\int_\gamma du$ depends only on the endpoints, not how we get between them. We can use Stokes' Theorem a second time to calculate $\int_\gamma du$ by treating the path γ itself as a manifold with boundary. Its boundary is $\{x, y\}$.

Then the integral over $\{x, y\}$ becomes

$$\int_{\partial\gamma} u = u(y) - u(x).$$

This extends our previous calculation showing that Stokes' Theorem includes the Fundamental Theorem of Calculus. Here it is further extended to include line integrals between two points x and y.[14]

[14] We restricted our attention to $\mathbb{R}^2$ to avoid complications that might occur if the two paths were not in the same plane. Eliminating that restriction requires additional work.

29.8.30 Coordinate Transformations

The integrals in Stokes' Theorem are ultimately based on Lebesgue or Riemann integrals, which give standard notions of length, area, volume, etc. in Euclidean space.

What happens if we use different coordinates in our integrals? Let's transform our usual rectilinear coordinates (x, y) to polar coordinates. Here $x = r\cos\theta$ and $y = r\sin\theta$. Of course, under the integral sign, $dx\,dy$ is another way of writing the 2-form $dx \wedge dy$. We can use the theory of differential forms to do the change of coordinates. We have $dx = \cos\theta\,dr - r\sin\theta\,d\theta$ and $dy = \sin\theta\,dr + r\cos\theta\,d\theta$. Then

$$dx \wedge dy = r\cos^2\theta\,dr \wedge d\theta - r\sin^2\theta\,d\theta \wedge dr = r\,dr \wedge d\theta$$

which is the area element in 2-dimensional Euclidean space.

In fact, the area element is also the Jacobian determinant of the transformation times the polar area element. The relevant transformation is

$$\mathbf{f}(r, \theta) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta & r\sin\theta \end{pmatrix}.$$

The Jacobian determinant is

$$\det \mathbf{J} = \det D_{(r,\theta)}\mathbf{f} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r,$$

so $r\,dr \wedge d\theta$ is the area element in polar coordinates.

This same technique can be used to handle non-Euclidean transformations and transformations from one manifold to another. Basically, one substitutes coordinates in both the function and the differential form.

3.1.12 Converse of Proposition 3.1.5

We now provide a near converse to Proposition 3.1.5.

Theorem 3.1.7. *Suppose* $f \in \mathcal{C}^2$ *on* $\mathbb{R}^m_{++}$ *and* $Df \neq 0$ *is homogeneous of degree* β *in* $\mathbf{x}$ *on* $\mathbb{R}^m_{++}$.

1. *If* $\beta \neq -1$*, there is a constant* c *and a homogeneous of degree one function* $\nu(\mathbf{x})$ *such that* $f(\mathbf{x}) = c + \big(\nu(\mathbf{x})\big)^{1+\beta}$.
2. *If* $\beta = -1$*, there is a constant* b*, a homogeneous of degree zero function* $\phi(\mathbf{x})$*, and a homogeneous of degree one function* $\nu(\mathbf{x})$ *with* $f(\mathbf{x}) = \phi(\mathbf{x}) + b \ln \nu(\mathbf{x})$. *Either* b *or* ϕ *may be zero.*

Proof. Since $f \in \mathcal{C}^2$, df (not Df!) is a closed form by Theorem 29.8.4. Stokes' Theorem then tells us that $\int_\alpha df = f(\mathbf{x}) - f(\mathbf{x}_0)$ for any path α in $\mathbb{R}^m_{++}$ running from $\mathbf{x}_0$ to $\mathbf{x}$.

We consider the effect of the transformation $\mathbf{x} \to t\mathbf{x}$ on the integral $\int_\alpha df$. The transformed interval is t times a long, so we multiply the dx_i terms by t to accomodate the new length as well as substituting $t\mathbf{x}$ in the integrand. That makes $df = \sum_i (\partial f/\partial x_i)\, dx_i$ homogeneous of degree $1 + \beta$. Then

$$f(t\mathbf{x}) - f(t\mathbf{x}_0) = t^{1+\beta}[f(\mathbf{x}) - f(\mathbf{x}_0)]. \tag{3.1.1}$$

To learn more about $f(t\mathbf{x}_0)$, define $g(t) = f(t\mathbf{x}_0)$. Then $g'(t) = \big[Df(t\mathbf{x}_0)\big]\mathbf{x}_0 = t^\beta \big[Df(\mathbf{x}_0)\big]\mathbf{x}_0$ is homogeneous of degree β. Since g' is a function of one variable, we can write $g'(t) = bt^\beta$.

Proof continues on next page . . .

3.1.13 Part (2) of Proposition 3.1.5

Part (2) of Proof. There are now two cases to consider: (a) $\beta \neq -1$ and (b) $\beta = -1$.

In case (a), $\beta \neq -1$. We next show $b \neq 0$ for some $\mathbf{x}_0$. Suppose, by way of contradiction, that $b = \big[Df(\mathbf{x}_0)\big]\mathbf{x}_0$ is zero for all $\mathbf{x}_0$, f must be homogeneous of degree zero by Euler's Theorem. But in that case Df must be homogeneous of degree minus one, contradicting $\beta \neq -1$.

As a result, there must be a $\mathbf{x}_0$ with $b \neq 0$. We now integrate g' to obtain $g(t) = c + b't^{1+\beta}$ for some constant c and $b' = b/(1+\beta)$. Then

$$\begin{aligned} f(t\mathbf{x}) &= f(t\mathbf{x}_0) + t^{1+\beta}[f(\mathbf{x}) - f(\mathbf{x}_0)] \\ &= g(t) + t^{1+\beta}[f(\mathbf{x}) - g(1)] \\ &= c + b't^{1+\beta} + t^{1+\beta}f(\mathbf{x}) - t^{1+\beta}(c + b') \\ &= c(1 - t^{1+\beta}) + t^{1+\beta}f(\mathbf{x}). \end{aligned}$$

It follows that $f(t\mathbf{x}) - c = t^{1+\beta}[f(\mathbf{x}) - c]$. Thus $f(\mathbf{x}) - c$ is homogeneous of degree $(1+\beta)$. Take the $(1+\beta)$ root to obtain $v(\mathbf{x}) = [f(\mathbf{x}) - c]^{1/(1+\beta)}$. The function v is homogeneous of degree one and $f(\mathbf{x}) = c + \big(v(\mathbf{x})\big)^{1+\beta}$, establishing part (a).

Proof concludes on next page . . .

3.1.14 Part (3) of Proposition 3.1.5

Proof of Case (b). In case (b), $\beta = -1$. Now $g'(t) = b/t$. As we saw above, b might be zero. We again integrate g', now obtaining $g(t) = c + b \ln t$ where c is a constant of integration.

Equation 3.1.1 now simplifies to

$$f(t\mathbf{x}) - f(t\mathbf{x}_0) = f(\mathbf{x}) - f(\mathbf{x}_0). \tag{3.1.2}$$

Substitute $f(t\mathbf{x}_0) = g(t) = c + b \ln t$ and $f(\mathbf{x}_0) = g(1) = c$ to obtain

$$f(t\mathbf{x}) = c + b \ln t + f(\mathbf{x}) - c = f(\mathbf{x}) + b \ln t.$$

Notice that the constant term c disappeared.

If $b = 0$, this shows that f is homogeneous of degree zero. We can then write $f(\mathbf{x}) = \phi(\mathbf{x})$ with ϕ homogeneous of degree zero.

If $b \neq 0$, define $\nu(\mathbf{x}) = \exp[f(\mathbf{x})/b]$. An easy calculation shows that $\nu(t\mathbf{x}) = \exp[f(t\mathbf{x})/b] = t \exp[f(\mathbf{x})/b]$. Thus ν is homogeneous of degree one and setting $\phi(\mathbf{x}) = 0$ yields the result. This completes the proof of part (b) and of the theorem. □

In the case $b \neq 0$, and ϕ is homogeneous of degree zero, we can also write

$$f(\mathbf{x}) = \phi(\mathbf{x}) + b \ln \left[e^{-\phi(\mathbf{x})/b} \nu(\mathbf{x}) \right].$$

As $-\phi(\mathbf{x})\nu(\mathbf{x})/b$ is homogeneous of degree 1, this provides a unified way to write f when $\beta = -1$.

3.2 Additive Separability

Our study of separability begins with the most basic type of separability: additive separability. We will derive a condition that is equivalent to additive separability for smooth utility functions. This allows us to characterize the monotonic transformations that preserve separability. This lays the groundwork for considering more general separability of both utility and of preferences in the next section.

3.2.1 Additive Separable Utility

Just as homothetic utility naturally lives on cones, separable utility naturally lives on product spaces. To simplify the presentation, we restrict our attention to two cases: the positive orthant $\mathbb{R}^m_+$ and the strictly positive orthant $\mathbb{R}^m_{++}$.[15]

We often use the strictly positive orthant to avoid continuity or differentiability issues that may arise if some $x_i = 0$. For example, even though the Cobb-Douglas utility $u(x_1, x_2) = \sqrt{x_1 x_2}$ is defined and continuous on all of $\mathbb{R}^2_+$, it is not differentiable if either $x_1 = 0$ or $x_2 = 0$.

Let the consumption set $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and u be a $\mathcal{C}^2$ utility function, $u\colon \mathfrak{X} \to \mathbb{R}$. We say utility u is *additive separable* on $\mathfrak{X}$ if there are functions $u_i\colon \mathbb{R}_+ \to \mathbb{R}$ or $u_i\colon \mathbb{R}_{++} \to \mathbb{R}$, as appropriate, so that

$$u(\mathbf{x}) = \sum_i u_i(x_i). \tag{3.2.3}$$

We refer to the functions u_i as *subutility functions*.

It is easily verified that additive separability of utility is not preserved by arbitrary increasing transformations. For example, if we square $u(x_1, x_2) = x_1 + x_2$, we obtain $(x_1 + x_2)^2$, which is a utility function that cannot be rewritten in an additive separable form. That tells us that additive separability is a **cardinal** property, not an **ordinal** property.

The only increasing transformations that obviously preserve the additive separable form are increasing affine transformations, defined by $\varphi(u) = au + b$ for some $a > 0$. Part (2b) of Theorem 3.2.1 shows that the affine transformations are the only transformations that preserve additive separability.

[15] For a treatment of more general product spaces, see Fishburn (1970, chapters 4-5). Fishburn does not cover the differentiable case.

3.2.2 Characterization of Additive Separability I

Together, Theorems 3.2.1 and 3.2.3 characterize additive separable utility functions.

Theorem 3.2.1. *Let $\mathfrak{X}$ be $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and $u\colon \mathfrak{X} \to \mathbb{R}$ be a $\mathcal{C}^2$ additive separable utility function.*

1. *Then whenever $i \neq j$,*

$$\frac{\partial^2 u}{\partial x_j\,\partial x_i} = 0.$$

2. *Suppose further that $u'_k \neq 0$ for at least two goods k. If φ is a $\mathcal{C}^2$ monotonic increasing transformation such that $v = \varphi \circ u$ is also additive separable, then $\varphi = au + b$ for some $a > 0$.*

Proof. By additive separability, we can write $u(\mathbf{x}) = \sum u_i(x_i)$.

Part (1). The marginal utility of good i is $\partial u/\partial x_i = u'_i(x_i)$. As this marginal utility does not depend on x_j whenever $i \neq j$,

$$\frac{\partial^2 u}{\partial x_j\,\partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial u}{\partial x_i}\right) = 0.$$

(Proof concludes on next page...)

3.2.3 Characterization of Additive Separability II

Part (2). Now suppose $\varphi \in \mathcal{C}^2$ preserves additive separability. Since $\varphi(u)$ is also separable, part (1) tells us that

$$\begin{aligned}
0 &= \frac{\partial^2 \varphi\big(u(\mathbf{x})\big)}{\partial x_j\,\partial x_i} = \frac{\partial}{\partial x_j}\left(\varphi'\big(u(\mathbf{x})\big)\frac{\partial u}{\partial x_i}\right) \\
&= \varphi''\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} + \varphi'\frac{\partial^2 u}{\partial x_j\,\partial x_i} \\
&= \varphi''\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} + 0 \\
&= \varphi'' u_i'(x_i) u_j'(x_j).
\end{aligned}$$

By hypothesis at least two goods have $u_i' > 0$, allowing us to conclude that $\varphi'' = 0$.

Integrating φ'' twice, we find $\varphi = au + b$ for some constants a and b. In other words, φ must be an affine transformation. Moreover, since φ is increasing, $a > 0$. $\square$

3.2.4 Separability with a Single Good

The requirement that at least two commodities have non-zero derivatives is needed in part (2) of Theorem 3.2.1. If there is only one such good, the utility function depends only on one good. In that case additive separability becomes a trivial condition and part (2) of Theorem 3.2.1 may fail as in the following example.

Example 3.2.2: Separability with One Good Consider the utility function on $\mathbb{R}^2_+$ given by $u(\mathbf{x}) = x_1$. This is additive separable, as is the equivalent representation $(x_1)^2$. In both cases the subutility for good two is zero. Here the transformation $\varphi(u) = u^2$ preserves the preference order and even gives us a new additive separable form, even though the transformation is not affine. ◀

3.2.5 Characterization of Additive Separability III

You may wonder whether the condition $\partial^2 u/\partial x_j \partial x_i = 0$ for all i and j with $i \neq j$ characterizes $\mathcal{C}^2$ additive separable utility functions. In fact, it does.

Theorem 3.2.3. *Let $\mathfrak{X} = \mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is $\mathcal{C}^2$ with*

$$\frac{\partial^2 u}{\partial x_j\,\partial x_i} = 0$$

for all i and j with $i \neq j$. Then there are functions $u_i\colon \mathbb{R}_+ \to \mathbb{R}$ or $u_i\colon \mathbb{R}_{++} \to \mathbb{R}$, respectively, such that $u(\mathbf{x}) = \sum_i u_i(x_i)$.

Proof. We will prove the theorem in the case when $\mathfrak{X} = \mathbb{R}^m_+$. Without loss of generality, we may presume $u(\mathbf{0}) = 0$.[16] Consider the differential

$$du = \sum_i \frac{\partial u}{\partial x_i}\, dx_i.$$

As the derivative of a $\mathcal{C}^2$ function, du is a closed differential form (Theorem 29.8.4). By Stokes' Theorem, the integral of du along any path from $\mathbf{0}$ to $\mathbf{x}$ will be the same.

Now consider a path γ from $\mathbf{0}$ to $\mathbf{x}$. By Stokes' Theorem, $\int_\gamma du = u(\mathbf{x}) - u(\mathbf{0}) = u(\mathbf{x})$. By choosing our path γ in the right way, we will end up with an additive separable expression.

(Proof concludes on next page...)

[16] In the case that $\mathfrak{X} = \mathbb{R}^m_{++}$, we replace $\mathbf{0}$ by any $\mathbf{x}^0 \in \mathfrak{X}$. In all the following equations, each 0 is then replaced by the appropriate x^0_j.

3.2.6 Characterization of Additive Separability IV

Rest of Proof. Start γ at $\mathbf{0}$, run along the x_1-axis to $(x_1, 0, \ldots, 0)$, next run parallel to the x_2-axis to $(x_1, x_2, 0, \ldots, 0)$, then on to $(x_1, x_2, x_3, 0, \ldots)$, et cetera, until we finally reach $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ along a path parallel to the x_m-axis. This yields

$$u(\mathbf{x}) = \sum_{i=1}^{m} \int_0^{x_i} \frac{\partial u}{\partial x_i}(x_1, x_2, \ldots, x_{i-1}, x_i', 0, \ldots, 0)\, dx_i' \tag{3.2.4}$$

Now consider what happens to the integral along each segment. We have

$$\begin{aligned}
&\int_0^{x_i} \frac{\partial u}{\partial x_i}(x_1, x_2, \ldots, x_{i-1}, x_i', 0, \ldots, 0)\, dx_i' \\
&= \int_0^{x_i} \frac{\partial u}{\partial x_i}(0, 0, \ldots, 0, x_i', 0, \ldots, 0)\, dx_i'
\end{aligned}$$

since $\partial u/\partial x_i$ depends only on x_i (zero cross-partials). We define $u_i(x_i)$ to be the latter integral.

$$u_i(x_i) = \int_0^{x_i} \frac{\partial u}{\partial x_i}(0, 0, \ldots, 0, x_i', 0, \ldots, 0)\, dx_i'$$

Substituting in equation 3.2.4, we obtain

$$u(\mathbf{x}) = \sum_{i=1}^{m} u_i(x_i).$$

□

3.2.7 Additive Separability and the MRS

In other words, u is a $\mathcal{C}^2$ additive separable utility representation if and only if $\partial^2 u/\partial x_j \partial x_i = 0$ for all $i \neq j$. This can sometimes be used to find a transformation that yields an additive separable representation.

Before showing how this works, it is useful to note another property that additive separable utility has that is shared by all differentially equivalent utility functions. The marginal rates of substitution between any pair of distinct goods depends only on the consumption of those two goods. It is unaffected by the consumption level of any other good.

Proposition 3.2.4. *Let $\mathfrak{X} = \mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ with $m > 2$. Suppose u and v are $\mathcal{C}^1$ utility functions on $\mathfrak{X}$ with $du, dv \gg \mathbf{0}$ and that there is an increasing $\mathcal{C}^1$ function φ with $v = \varphi \circ u$. If either u or v is additive separable, then both MRS^u_{ij} and MRS^v_{ij} are independent of x_k for every triple of distinct goods, i, j, and k.*

Proof. When utility u is additive separable, the marginal rate of substitution is

$$\mathrm{MRS}^u_{ij} = \frac{u_i'(x_i)}{u_j'(x_j)}.$$

It depends solely on the consumption levels x_i and x_j. It is independent of the consumption level of any other good.

By Theorem 1.3.8, $\mathrm{MRS}^u_{ij} = \mathrm{MRS}^v_{ij}$, so if one of them is additive separable, both are independent of x_j for every $k \neq i, j$.

The argument of Theorem 1.3.8 is simple.

$$\mathrm{MRS}^v_{ij} = \frac{\partial v/\partial x_i}{\partial v/\partial x_j} = \frac{\varphi'\,\partial u/\partial x_i}{\varphi'\,\partial u/\partial x_j} = \frac{u_i'(x_i)}{u_j'(x_j)} = \mathrm{MRS}^u_{ij}\,.$$

□

3.2.8 Restriction to 3 or More Goods

Three Goods. Proposition 3.2.4 requires that $m > 2$, that there are at least three goods. This is necessary because the condition that MRS_{ij} be independent of x_k for $k \neq i, j$ with $i \neq j$ requires at least three goods in its definition.

When there are three or more goods, Proposition 3.2.4 tells us that if a utility function is differentially equivalent to an additive separable utility function, then the marginal rate of substitution between any two goods depends only on the quantities consumed of those two goods. The Smooth Separability Theorem from section 3.4 shows that this condition is also sufficient for a utility function to be differentially equivalent to an additive separable utility function, at least provided there are at least three goods.

3.2.9 Testing for Additive Separability

Proposition 3.2.4 allows us test whether a utility function is equivalent to an additive separable utility function by examining the marginal rates of substitution. If the utility function passes the test, we can then use our second derivative condition (Theorems 3.2.1 and 3.2.3) to attempt to find how to transform it into additive separable form. We illustrate this with Cobb-Douglas utility.

Example 3.2.5: Separability and Cobb-Douglas Utility Consider the Cobb-Douglas utility function $u(\mathbf{x}) = \prod_{i=1}^m x_i^{\gamma_i}$ where each $\gamma_i > 0$. The marginal rates of substitution are $\mathrm{MRS}_{ij} = \gamma_i x_j/\gamma_j x_i$. As these depend only on x_j and x_i, there is a way to transform u into an additive separable form.

Suppose φ is an increasing transformation that yields an additive separable $v = \varphi \circ u$. We compute the second partial derivatives to obtain a differential equation for φ.

Now

$$\begin{aligned} 0 = \frac{\partial^2 \varphi\big(u(\mathbf{x})\big)}{\partial x_j\,\partial x_i} &= \varphi'' \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} + \varphi' \frac{\partial^2 u}{\partial x_j\,\partial x_i} \\ &= \gamma_i \gamma_j \varphi'' \frac{u^2}{x_j x_i} + \gamma_i \gamma_j \varphi' \frac{u}{x_j x_i}. \end{aligned}$$

Thus $\varphi'' u + \varphi' = 0$ where $\varphi' = d\varphi/du$ and $\varphi'' = d^2\varphi/du^2$.

We solve this differential equation by using the substitution $\psi = \varphi'$ to find $\psi' u + \psi = 0$. This is easily integrated, obtaining $\psi(u) = C/u$ for some constant C. That means $\varphi' = C/u$. A second integration yields $\varphi(u) = C \ln u + D$. The requirement that φ is increasing means $C > 0$.

Any transformation of the form $\varphi(u) = C \ln u + D$ with $C > 0$ will transform u into an additive separable form. If we take $C = 1$ and $D = 0$, we obtain $\varphi\big(u(\mathbf{x})\big) = \sum_{i=1}^m \gamma_i \ln x_i$. ◀

3.2.10 What about Two Goods? SKIPPED

Two Goods. Unfortunately, the characterization of additive separable utility via the marginal rate of substitution only applies when there are three or more goods. Indeed, the condition that MRS_{ij} is independent of the consumption of a third good is meaningless when there are only two goods.

Another approach is necessary in the two good case. Suppose $u\colon \mathbb{R}^2_+ \to \mathbb{R}$ is additive separable. Then $u(\mathbf{x}) = u_1(x_1) + u_2(x_2)$. It follows that $\mathrm{MRS}_{12} = u_1'(x_1)/u_2'(x_2)$. Now consider the logarithm,

$$\ln \mathrm{MRS}_{12} = \ln u_1'(x_1) - \ln u_2'(x_2).$$

Here the logarithm of the marginal rate of substitution **is** additive separable. Is that enough to guarantee that u is additive separable? The following proposition shows that it is.

Proposition 3.2.6. *Let $\mathfrak{X}$ be $\mathbb{R}^2_+$ or $\mathbb{R}^2_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is a $\mathcal{C}^2$ utility function with $du \gg \mathbf{0}$ and that $\ln \mathrm{MRS}^u_{12}$ is additive separable. Then u is equivalent to an additive separable utility function.*

Proof. When $\ln \mathrm{MRS}^u_{12}$ is additive separable, we can write $\ln \mathrm{MRS}^u_{12} = \phi_1(x_1) - \phi_2(x_2)$. Define

$$\varphi_i(x_i) = \int e^{\phi_i(x_i')}\,dx_i'$$

and set $v(x_1, x_2) = \varphi_1(x_1) + \varphi_2(x_2)$. By the Fundamental Theorem of Calculus,

$$\mathrm{MRS}^v_{12} = \frac{\varphi_1'}{\varphi_2'} = \frac{e^{\phi_1(x_1)}}{e^{\phi_2(x_2)}}.$$

This implies $\ln \mathrm{MRS}^v_{12} = \phi_1 - \phi_2 = \ln \mathrm{MRS}^u_{12}$. It follows that $\mathrm{MRS}^v_{12}(\mathbf{x}) = \mathrm{MRS}^u_{12}(\mathbf{x})$ for all $\mathbf{x} \in \mathfrak{X}$. By Corollary 2.6.9, u and v are equivalent. $\square$

3.2.11 Two Goods, the Complete Solution SKIPPED

By combining Proposition 3.2.6 and Theorem 3.2.3, we get necessary and sufficient conditions for functions on $\mathbb{R}^2_+$ to be additive separable.

Theorem 3.2.7. *Let $\mathfrak{X}$ be $\mathbb{R}^2_+$ or $\mathbb{R}^2_{++}$ and suppose $u\colon \mathbb{R}^2_+ \to \mathbb{R}$ or $u\colon \mathbb{R}^2_{++} \to \mathbb{R}$ is a $\mathcal{C}^3$ utility function. Then u is equivalent to an additive separable utility function if and only if*

$$\frac{\partial^2}{\partial x_1\,\partial x_2} \ln \mathrm{MRS}_{12} = 0.$$

Proof. Suppose $\partial^2 \ln \mathrm{MRS}_{12} / \partial x_1 \partial x_2 = 0$. Then $\ln \mathrm{MRS}_{12}$ is additive separable by Theorem 3.2.3. Proposition 3.2.6 then shows that u is equivalent to an additive separable utility function.

Conversely, if u is equivalent to an additive separable utility function, $\ln \mathrm{MRS}_{12}$ is additive separable, and so has a zero cross partial derivative. $\square$

3.2.12 Group Additive Separable Utility

We do not need to restrict ourselves to subutility functions that depend only on a single good. We can also consider subutility functions that depend on multiple goods. For example, we might have $u(\mathbf{x}) = x_3 + (x_1 x_2 x_4)^{1/2} = u_3(x_3) + u_1(x_1, x_2, x_4)$. This utility function also exhibits a limited type of additive separability.

To formalize this, we introduce commodity groups.

3.2.13 Commodity Groups

In $\mathbb{R}^m_+$, the set of goods or commodities is $\mathcal{G} = \{1, ..., m\}$.

Commodity Group. Let $\mathcal{G}$ be the set of commodities. A *commodity group* is a subset A of $\mathcal{G}$.

Given a commodity group A, define the vectors $\mathbf{x}_A = (x_i)_{i \in A}$ and $\mathbf{x}_{\sim A} = (x_i)_{i \notin A}$. We slightly abuse notation to write $\mathbf{x} = (\mathbf{x}_A, \mathbf{x}_{\sim A})$.

In the utility function above, we might take $\{3\}$ or $\{1, 2, 4\}$ as A. If taken literally, writing $\mathbf{x} = (\mathbf{x}_A, \mathbf{x}_{\sim A})$ would mean that we would be writing $\mathbf{x}$ as $\big(x_3, (x_1, x_2, x_4)\big)$ or $\big((x_1, x_2, x_4), x_3\big)$, respectively, instead of (x_1, x_2, x_3, x_4). We don't intend $\mathbf{x} = (\mathbf{x}_A, \mathbf{x}_{\sim A})$ to be taken literally, and always interpret $(\mathbf{x}_A, \mathbf{x}_{\sim A})$ as the original $\mathbf{x}$. We use the notation $\mathbb{R}^A$ to denote the set of all $\mathbf{x}_A$.

3.2.14 Partitioning Commodities into Groups

A *partition* of the space of goods $\mathcal{G} = \{1, \ldots, m\}$ (or of any set) is a collection $\mathcal{P}$ of disjoint subsets $P \subset \mathcal{G}$ whose union is the entire space of goods, $\mathcal{G}$. That is, when $P_i \cap P_j = \emptyset$ for $P_i \neq P_j$ and $\cup_{P \in \mathcal{P}} P = \mathcal{G}$. Any partition of the set of commodities divides the commodities into distinct commodity groups.

Example 3.2.8: Partitions The set $\{1\}$ has only one partition, $\big\{\{1\}\big\}$. The set $\{1, 2\}$ can be partitioned in two ways. Each singleton has its own set in the partition, $\mathcal{P} = \big\{\{1\}, \{2\}\big\}$, or we can consider the partition consisting only of the set $\{1, 2\}$ itself. This type of partition is always possible and is called the *trivial partition*.

When there are three elements, $\{1, 2, 3\}$, the set can be partitioned in 5 ways. There is the trivial partition $\{1, 2, 3\}$, the partition of singletons $\big\{\{1\}, \{2\}, \{3\}\big\}$, and the three partitions into a singleton and a doubleton: $\big\{\{1\}, \{2, 3\}\big\}$, $\big\{\{2\}, \{1, 3\}\big\}$, and $\big\{\{3\}, \{1, 2\}\big\}$.

The number of ways a set of size n may be partitioned is the *Bell number* B_n.[17]

The Bell numbers obey the recursion

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$$

where $\binom{n}{k} = n!/(n-k)!k!$ are the binomial coefficients and $B_0 = 1$. Starting at $n = 0$, the first several Bell numbers are $1, 1, 2, 5, 15, 203, 877, 4140, \ldots$. ◄

[17] Although the Bell numbers (Bell, 1938) are named after Scottish-American mathmetician and science fiction writer Eric Temple Bell (1883–1960), he did not invent them. The earliest investigation of the number of partitions appears to date to medieval Japan. The first formula for them seems to be that of Dobiński in 1877.

Bell is also known for his work on generating functions. Today, he's best remembered for the book *Men of Mathematics* and his science fiction, both under his own name and as John Taine.

3.2.15 Group Additive Separability

Partitions allow us to define group separability. Let $\mathfrak{X}$ be $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$. We say that $u\colon \mathfrak{X} \to \mathbb{R}$ is *group additive separable* or *additive separable relative to the commodity partition* $\mathcal{P}$ if there are subutility functions u_P for every commodity group $P \in \mathcal{P}$ where $u_P\colon \mathbb{R}^P_+ \to \mathbb{R}$ or $u_P\colon \mathbb{R}^P_{++} \to \mathbb{R}$, respectively, and u is the sum of subutility functions

$$u(\mathbf{x}) = \sum_{P \in \mathcal{P}} u_P(\mathbf{x}_P). \tag{3.2.5}$$

Thus $u(\mathbf{x}) = x_3 + (x_1 x_2 x_4)^{1/2}$ is additively separable with respect to the partition $\big\{\{3\},\{1,2,4\}\big\}$.

3.2.16 Characterizing Group Additive Separable Utility I

We can generalize Theorem 3.2.1 to encompass utility functions that are additively separable relative to a partition $\mathcal{P}$ as follows.

Theorem 3.2.9. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is a $\mathcal{C}^2$ utility function that is additively separable relative to a commodity partition $\mathcal{P}$, so $u(\mathbf{x}) = \sum_{P\in\mathcal{P}} u_P(\mathbf{x}_P)$.*

1. *Then $\partial^2 u/\partial x_i \partial x_j = 0$ whenever i and j are members of different commodity groups.*
2. *Suppose further that there are goods in at least two commodity groups with $\partial u_P/\partial x_i > 0$. If φ is a $\mathcal{C}^2$ monotonic increasing transformation such that $v = \varphi \circ u$ is also additive separable relative to $\mathcal{P}$, then $\varphi = au + b$ for some $a > 0$.*

Proof. Mimic the proof of Theorem 3.2.1. $\square$

3.2.17 Characterizing Group Additive Separable Utility II

We also have a version of Theorem 3.2.3.

Theorem 3.2.10. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is $\mathcal{C}^2$ and there is a commodity partition $\mathcal{P}$ of the list of goods $\{1, \ldots, m\}$ with $\frac{\partial^2 u}{\partial x_j\, \partial x_i} = 0$ whenever i and j are in different elements of the partition $\mathcal{P}$. Then there are $u_P\colon \mathbb{R}^P_+ \to \mathbb{R}$ such that $u(\mathbf{x}) = \sum_{P \in \mathcal{P}} u_P(x_P)$.*

Proof. Omitted, see manuscript. □

Theorem 3.2.10 does a nice job of characterizing $\mathcal{C}^2$ utility functions that are group additive separable. We saw in Example 3.2.5 that this characterization can sometimes be used to find an additive separable representation of a given utility function, as we did with Cobb-Douglas. What the theorem does not do is tell whether a given utility function (or even preference order) has an additive separable representation. One clue that this theorem is not enough is that the criterion for additive separability (cross-partials are zero) is not invariant under monotonic transformations.

3.3 Induced Orders on Commodity Groups 01/17/23

We've studied additive separable preferences over both individual goods and commodity groups. Although we have characterized additive separable utility, we still lack a way of characterizing utility that is a monotonic transformation of additive separable utility. This section is a step toward that goal. However, we will not reach that goal in this section.

This section focuses on induced orders, where preferences over one commodity group are independent of consumption of other goods. It culminates in the concept of Sono separability and conditions that allow us to find subutility functions. However, those subutilities need not be related in an additive way.

3.3.1 Additive Separability and Commodity Groups

To understand how a preference order might be separable, we turn to additive separable utility to develop some intuition. One important fact about additive separable utility is that it unambiguously defines utility over each commodity group. The following example shows how this works.

Example 3.3.1: Separability and Commodity Groups If u is an additive separable utility function on $\mathbb{R}^m_+$, and A is a commodity group. There is a natural way to define preferences on $\mathbb{R}^A_+$. Simply restrict the additive form to $\mathbb{R}^A$. If $u(\mathbf{x}) = \sum_i u_i(x_i)$, we can do this by defining

$$u_A(\mathbf{x}_A) = \sum_{i \in A} u_i(x_i).$$

Then $u(\mathbf{x}) = u_A(\mathbf{x}_A) + u_{\sim A}(\mathbf{x}_{\sim A})$. Now suppose that for **some** $\mathbf{y}_{\sim A} \in \mathbb{R}^{\sim A}$,

$$\begin{aligned} u(\mathbf{x}'_A, \mathbf{y}_{\sim A}) &= u_A(\mathbf{x}'_A) + u_{\sim A}(\mathbf{y}_{\sim A}) \\ &\geq u_A(\mathbf{x}_A) + u_{\sim A}(\mathbf{y}_{\sim A}) \\ &= u(\mathbf{x}_A, \mathbf{y}_{\sim A}). \end{aligned}$$

Then $u_A(\mathbf{x}'_A) \geq u_A(\mathbf{x}_A)$. Conversely, if $u_A(\mathbf{x}'_A) \geq u_A(\mathbf{x}_A)$, we have $u(\mathbf{x}'_A, \mathbf{y}_{\sim A}) \geq u(\mathbf{x}_A, \mathbf{y}_{\sim A})$ for **every** $\mathbf{y}_{\sim A} \in \mathbb{R}^{\sim A}_+$.

This means the preference order given by the derived utility function u_A is the same as we would by defining $\mathbf{x}'_A \succsim_A \mathbf{x}_A$ if and only if $(\mathbf{x}'_A, \mathbf{y}_{\sim A}) \succsim (\mathbf{x}_A, \mathbf{y}_{\sim A})$ for all $\mathbf{y}_{\sim A} \in \mathbb{R}^{\sim A}_+$. This definition of $\succsim_A$ is equivalent to using the utility function u_A. ◀

3.3.2 Ordinal Utility and Commodity Groups

In fact, the idea of Example 3.3.1 applies to any utility function that is equivalent to an additive separable utility function. One way to implement it is to fix the consumption of goods that are **not** in the commodity group A. That is, we fix consumption of goods in $\sim A$. Proposition 3.3.2 shows us that the resulting order on $\mathbb{R}^A_+$ is independent of the consumption levels of goods outside group A.

Proposition 3.3.2. *Let A be a commodity group. If $u\colon \mathbb{R}^m_+ \to \mathbb{R}$ is equivalent to an additive separable utility function, then $u_A(\mathbf{x}_A) = u(\mathbf{x}_A, \bar{\mathbf{x}}_{\sim A})$ defines a preference order on $\mathbb{R}^A_+$ that is independent of the choice of $\bar{\mathbf{x}}$.*

Proof. We can write $u = \phi \circ \nu$ where ν is additive separable and ϕ is increasing. Now

$$u_A(\mathbf{x}_A) = \phi\big(\nu(\mathbf{x}_A, \bar{\mathbf{x}}_{\sim A})\big) = \phi\left(\sum_{i\in A} \nu_i(x_i) + \sum_{i\notin A} \nu_i(\bar{x}_i)\right).$$

Since ϕ is increasing, this is equivalent to $\sum_{i\in A} \nu_i(x_i) + \sum_{i\notin A} \nu_i(\bar{x}_i)$. The second term is constant, so the preference order defined by u_A on $\mathbb{R}^A_+$ is equivalent to $\sum_{i\in A} \nu_i(x_i)$, which is independent of $\bar{\mathbf{x}}$. □

3.3.3 Cardinal and Ordinal Properties

Before delving into the implications of this idea, we first stop to see how it is connected to the characterization of additive separable utility in Theorems 3.2.1 and 3.2.3. Those theorems involved the condition that $\frac{\partial^2 u}{\partial x_j \, \partial x_i} = 0$ for every i, j with $i \neq j$. Equivalently, the marginal utility of each good is independent of consumption of any other good.

This may seem a somewhat unsatisfactory condition to build our intuition on because it depends on the utility representation, not the underlying preference order. Marginal utility is just **not** an **ordinal** property. However, the closely related marginal rate of substitution **is ordinal**.

Proposition 3.2.4 showed that the independence of marginal utility immediately implies that the marginal rate of substitution between i and j, MRS_{ij}, depends only on x_i and x_j. This is true not only of additive separable utility, but of any increasing transformation of an additive separable utility function. It is an ordinal property, and an eminently suitable foundation for intuition.

Since the i-j indifference curves of such utility functions have the same slope regardless of the consumption of other goods, the i-j indifference map is unaffected by the consumption of other goods. We can define utility in i-j space independently of consumption of other goods by $u_{ij}(x_i, x_j) = u_i(x_i) + u_j(x_j)$. This is not only true of i and j, but applies to any commodity group.

3.3.4 Induced Orders

So if we have an arbitrary commodity group A, separability allows us to define preferences over goods in A without regard to the consumption of goods not in A. This is the key to an ordinal definition of separability.[18] One requirement is that it define a preference order on every commodity group A. We can extend this method of defining utility on a commodity group A to more general preferences.

Induced Order. Let $A \subset \{1, \ldots, m\}$ be any commodity group. A preference order $\succsim$ on $\mathfrak{X} \subset \mathbb{R}^m_+$ *induces an order* on group A if $(\mathbf{x}_A, \mathbf{x}_{\sim A}) \succsim (\mathbf{y}_A, \mathbf{x}_{\sim A})$ implies $(\mathbf{x}_A, \mathbf{z}_{\sim A}) \succsim (\mathbf{y}_A, \mathbf{z}_{\sim A})$ for all $\mathbf{z} \in \mathfrak{X}$. We denote this preference order by $\succsim_A$.

Saying that we can define an induced order means that we can rank $\mathbf{x}_A$ and $\mathbf{y}_A$ unambiguously. In that case we write $\mathbf{x}_A \succsim_A \mathbf{y}_A$. When $\succsim$ induces an order on commodity group A, it means that $\succsim$ unambiguously defines preferences over commodity bundles in $\mathbb{R}^A$.[19] Proposition 3.3.2 showed that when preferences are defined by an additive separable utility function (or equivalent utility function), they induce an order on **every commodity group**.

[18] Sono (1943) seems to have been the first to realize this.

[19] In Sono's (1943) definition, a group A is separable from commodity j if the consumption of j does not affect the ranking of bundles of goods from A. That is, A is separable from j in Sono's sense if there is an induced order $\succsim_A$ on $\mathfrak{X}^A$.

3.3.5 Additive Separability and Induced Orders

There is also a version of Proposition 3.3.2 for group additive separable utility.

Theorem 3.3.3. *Let $\mathcal{P}$ be a commodity partition and A a union of commodity groups in $\mathcal{P}$. Suppose $u\colon \mathbb{R}^m_+ \to \mathbb{R}$ is equivalent to a utility function that is additive separable relative to $\mathcal{P}$. Then $u_A(\mathbf{x}_A) = u(\mathbf{x}_A, \bar{\mathbf{x}}_{\sim A})$ defines a preference order on $\mathbb{R}^A_+$ that is independent of the choice of $\bar{\mathbf{x}}$.*

Proof. We can write $u = \phi \circ v$ where v is additive separable and ϕ is increasing. Now for any commodity group in $\mathcal{P}$, either P is one of the commodity groups that make up A, in which case $P \subset A$ or P is not one of those groups, when $P \cap A = \emptyset$. This lets us write

$$u_A(\mathbf{x}_A) = \phi\big(v(\mathbf{x}_A, \bar{\mathbf{x}}_{\sim A})\big) = \phi\left(\sum_{P\subset A} v_P(x_P) + \sum_{P\cap A=\emptyset} v_P(\bar{x}_P)\right).$$

Since ϕ is increasing, this is equivalent to $\sum_{P\subset A} v_P(x_P) + \sum_{P\cap A=\emptyset} v_P(\bar{x}_P)$. The second term is constant, so the preference order defined by u_A on $\mathbb{R}^A_+$ is equivalent to $\sum_{P\subset A} v_P(x_P)$, which is independent of $\bar{\mathbf{x}}$. $\square$

3.3.6 Monotonicity and Induced Orders

Additive separable utility is not the only way to induce an order. Other types of preference orders can also induce orders on particular commodity groups. One of the simplest types of induced orders occurs when preferences are monotonic.

Proposition 3.3.4. *Suppose* $\succsim$ *is weakly monotonic in* x_i. *Define the commodity group* $P = \{i\}$. *Then* $\succsim$ *induces an order on* P.

Proof. Now $(\boldsymbol{x}_P, \boldsymbol{x}_{\sim P}) \succsim (\boldsymbol{y}_P, \boldsymbol{x}_{\sim P})$ if and only if $x_i \geq y_i$. But this holds if and only if $(\boldsymbol{x}_P, \boldsymbol{z}_{\sim P}) \succsim (\boldsymbol{y}_P, \boldsymbol{z}_{\sim P})$ for any $\boldsymbol{z}$, thus $\succsim$ induces an order on $P = \{i\}$. $\square$

Monotonicity is not enough to guarantee additive separability.

Example 3.3.5: Monotonic but not Additive Separable Utility Consider the utility function $u(\boldsymbol{x}) = (x_1 + x_2)(x_2 + x_3)$. Since $du \gg \mathbf{0}$ on $\mathbb{R}^3_{++}$, this function is strongly monotonic there. As a result, it induces an order on the singleton commodity groups $\{1\}$, $\{2\}$, and $\{3\}$.

This utility function is not additive separable because each marginal rate of substitution depends on all three variables, not just the two involved variables. This means it cannot even be transformed into an additive separable form (Problem 3.3.1). ◀

3.3.7 Utility with No Non-trivial Induced Orders I

When u is not monotonic, it need not induce an order on any commodity group other than $\{1, \ldots, m\}$.

Example 3.3.6: Utility without Induced Orders To see that utility need not induce orders on anything other than the entire consumption set, consider the utility function $u(x_1, x_2) = (x_1 - x_2)^2$. It does not induce an order on either singleton, $\{1\}$ or $\{2\}$.

This is based on the observation that $u(2,0) = 4$, $u(0,0) = 0$, $u(2,2) = 0$, and $u(0,2) = 4$. It follows that although $(2,0) \succ (0,0)$, changing the amount of good two to 2 reverses the relation: $(2,2) \prec (0,2)$. Similarly, although $(0,2) \succ (0,0)$, changing the amount of good one reverses the relation: $(2,0) \prec (2,2)$. Thus u does not induce an order on either commodity group $\{1\}$ or $\{2\}$.

Figures 3.3.7 and 3.3.8 illustrate this.

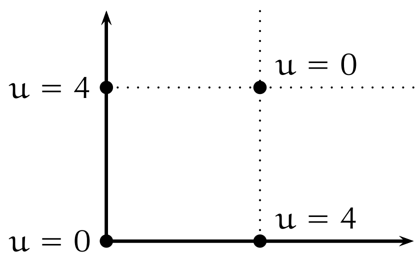

Figure 3.3.7: Inducing an order on $\{1\}$ would mean that either $x_1 = 1$ is always better than $x_1 = 0$ or always worse, regardless of how much of good two we have. But the preference order reverses as we move from the horizontal axis to the dotted horizontal line.

A similar thing happens with good two, as we see by comparing preferences along the vertical axis and the dotted vertical line.

3.3.8 Utility with No Non-trivial Induced Orders II

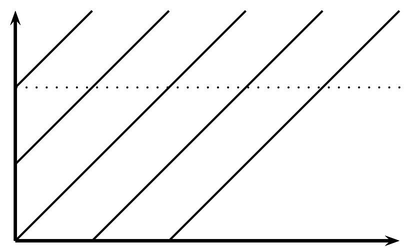

Figure 3.3.8: The diagonal lines are the indifference curves for $u = 0, 1, 4$. For $u > 0$, the indifference curve consists of two parallel lines. As we move right along the dotted line, utility decreases until we hit the 45° line, and then increases. It defines a different preference order on good one for every value of x_2. These preferences do not induce any fixed order on good one. Considering vertical lines shows the same thing for good two.

◀

3.3.9 Induced Orders and MRS

Proposition 3.3.4 and Example 3.3.5 drive home the point that merely inducing an order on some commodity group, or even every singleton, is insufficient for additive separability. It is necessary, but insufficient. In fact, Theorem 3.2.1 shows that additive separability requires that orders are induced on every commodity group. We will return to this idea later, after we explore induced orders a bit more.

When dealing with larger groups of goods, utility functions of the form

$$u(\mathbf{x}) = \psi\big(\phi(\mathbf{x}_A), \mathbf{x}_{\sim A}\big) \tag{3.3.6}$$

induce an order on A when ψ is strictly increasing. Here ϕ is a subutility function on $\mathbb{R}^A_+$ and it represents the induced order on $\mathbb{R}^A_+$. If A contains at least two elements and ψ and ϕ are both $\mathcal{C}^2$, we can consider $\mathrm{MRS}_{ij} = (\partial\phi/\partial x_i)\big/(\partial\phi/\partial x_j)$ for $i, j \in A$. In that case

$$\frac{\partial}{\partial x_k}\big(\,\mathrm{MRS}_{ij}\,\big) = 0 \text{ for every } k \notin A. \tag{3.3.7}$$

Ever since Sono (1943), equation 3.3.7 has been used to define a type of separability. Nonetheless, we consider Sono separability to be a case of an induced order, and reserve the term "separable" for cases where goods in A and the goods not in A are treated on an equal footing—where orders are induced on both A and its complement.

An example of this type of separability is $\phi(x_1, x_2) = (1 + x_2)(x_1 + x_2)$ and $u(x_1, x_2, x_3) = -e^{-x_3} + x_3\phi(x_1, x_2)$. When considered on $\mathbb{R}^3_{++}$, this function induces an order on (x_1, x_2). Verification that u is not equivalent to an additive separable utility function is straightforward and left as Exercise 3.3.5. This example shows that we can have a limited type of separability without attaining additive separability.

3.3.10* Leontief's Subutility Theorem

We are now ready to state Leontief's Subutility Theorem.

Subutility Theorem (Leontief). *Suppose $u \in \mathcal{C}^2$ is strictly increasing on the strictly positive orthant $\mathbb{R}^m_{++}$ and let A be a commodity group where u obeys $\frac{\partial}{\partial x_k}(\mathrm{MRS}_{ij}) = 0$ for every $i, j \in A$ and every $k \notin A$. Given $\mathbf{x}^0 \in \mathbb{R}^m_{++}$, there are $\mathcal{C}^2$ functions ψ and ϕ such that $u(\mathbf{x}) = \psi\big(\phi(\mathbf{x}_A), \mathbf{x}_{\sim A}\big)$ on a neighborhood of $\mathbf{x}^0$. Moreover, ψ is strictly increasing in $(\phi, \mathbf{x}_{\sim A})$.*

One weakness of Leontief's Separability Theorem is that the representation is local, not global. In general, it is a non-trivial problem to extend the representation to all of $\mathbb{R}^m_+$.

3.4 Separable Preference Orders

Why do we care about separability of preferences as well as of utility? One answer is that by studying preferences directly, by breaking separability loose from utility, we clarify the meaning and significance of separability.

The problem with using utility as the basis for definition is that conditions imposed directly on utility may result in unknown assumptions concerning preferences themselves. For this reason, many economic theorists prefer to impose hypotheses concerning preferences directly on the preference ordering. When dealing directly with preferences, the extra complication of a plethora of equivalent forms does not arise. The ideal situation is to have theorems relating properties of preferences and of utility functions. That is what we do for separability in this section.

When can a utility function be transformed into an additive separable form? When can preferences be represented by such a function? To fully answer these questions, we have to consider separability of preference orders rather than separability of utility. The resulting framework applies not just to preferences with additive separable representations, but also allows us to consider other types of preferences, such as those with a quasi-linear representation.

3.4.1 Weak, Strong, and Complete Separability

Now suppose we have a bunch of commodity groups. Can we write utility in terms of subutilities for each group?

To answer this, we introduce three distinct types of separability for preference orders. Whether a given preference order will have a separable utility representation will depend on the amount of separability it has.

Types of Separability. Let $\mathcal{P}$ be a non-trivial partition of goods $\mathcal{G} = \{1, \ldots, m\}$ into commodity groups. If $\succsim$ is a preference order on either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$, we say $\succsim$ is *weakly separable* relative to $\mathcal{P}$ if $\succsim$ induces an order on each commodity group $P \in \mathcal{P}$. Weakly separable preferences are sometimes referred to as *separable*.

A preference order $\succsim$ is *strongly separable* relative to $\mathcal{P}$ if it induces an order not only on each commodity group in $\mathcal{P}$, but also on each **union** of commodity groups in $\mathcal{P}$.

Finally, we say $\succsim$ is *completely separable* if it is strongly separable relative to the partition of singletons, $\mathcal{P}_s = \big\{\{1\}, \{2\}, \ldots, \{m\}\big\}$.

Unlike Sono's separability condition from equation 3.3.7, weak, strong, and complete separability treat all commodity groups symmetrically.

Complete separability is equivalent to requiring that $\succsim$ induce an order on every partition, or that it be both weakly and strongly separable on every partition of $\{1, \ldots, m\}$, that it induce an order on every possible commodity group. As we noted before, Proposition 3.3.2 shows that any additive separable utility induces an order on every commodity group. In other words, every additive separable utility defines a completely separable preference order.

Similarly, Theorem 3.3.3 implies that any group additive separable utility is strongly separable relative to its partition.

3.4.2 Separability with Two Goods

When there are only two goods the different types of separability are not distinct. There is only one non-trivial partition, the partition of singletons, $\mathcal{P}_s = \{\{1\},\{2\}\}$. Weak separability and strong separability are the same because the only union is set of all goods, where the original preference order is the induced order. When there are only two goods, weak and strong separability on $\mathcal{P}_s$ are the same as complete separability. We need at least three goods for the different types of separability to be distinct.

The same sort of problem can occur if there are three goods, but one of them does not affect utility. In that case, its presence or absence will not affect the ability to induce an order. There are only two relevant goods. Once again, the definitions all coincide.

Rather than requiring that all commodities affect preference, it is enough that there be a sufficient variety of goods that affect preferences. Consider a commodity group A. If there are x_A, y_A and $z_{\sim A}$ with $(x_A, z_{\sim A}) \not\sim (y_A, z_{\sim A})$, we say group A is *essential* for $\succsim$. In other words, commodity group A is essential if there are consumption bundles where changes in consumption of the goods in A can affect preference. Inessential commodity groups will not show up in any additive separable representation. We will sometimes require that at least three commodity groups are essential.

The case where there are only two essential commodity groups has to be handled separately. We will deal with this later when we consider the two good case, and introduce another criterion for separability there.

3.4.3 Additivity and Complete Separability

To get a little more feel for the definitions, if u is a group additive separable utility function relative to a partition $\mathcal{P}$, as in Equation 3.2.5, then u is strongly separable relative to $\mathcal{P}$. This means that ordinary additive separable utility is always completely separable, as shown in the following proposition.

Proposition 3.4.1. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is equivalent to an additive separable utility function. Then u is completely separable.*

Proof. Recall the partition of singletons: $\mathcal{P}_s = \big\{\{1\},\{2\},\dots,\{m\}\big\}$. By Proposition 3.3.2, u induces an order on every commodity group. Thus it not only induces an order on every singleton (is weakly separable relative to $\mathcal{P}_s$), but it also induces an order on every union of singletons, and hence on every commodity group. Then u is strongly separable on $\mathcal{P}_s$ and so is completely separable. □

Theorem 3.3.3 tells us that group additive utility is strongly separable relative to its defining partition. As in the example below, such utility functions need not be strongly separable relative to finer partitions.

Example 3.4.2: Group Additive Utility and Separability On $\mathbb{R}^6_+$, the utility function $u(\mathbf{x}) = u_1(x_1, x_2, x_3) + u_4(x_4, x_5, x_6)$ is both weakly and strongly separable relative to the partition $\big\{\{1,2,3\},\{4,5,6\}\big\}$. There are many functions u_1 and u_4, where it is not completely separable. One such case is when $u_1(x_1, x_2, x_3) = x_1x_3 + x_2x_3$, which has $\mathrm{MRS}_{13} = x_3/(x_1 + x_2)$. Since this marginal rate of substitution depends on x_2, u is not completely separable. ◀

3.4.4 Monotonicity and Weak Separability

One easy result is that monotonic preferences are always weakly separable with respect the partition of singletons.[20]

Proposition 3.4.3. *Suppose $\succsim$ is weakly monotonic. Then $\succsim$ is weakly separable relative to the partition of singletons.*

Proof. In this case Proposition 3.3.4 tells us that u induces an order on each $P \in \mathcal{P}_s$. Therefore u is weakly separable relative to $\mathcal{P}_s$. □

Combining Proposition 3.4.3 and Example 3.3.5 tells us that weak separability is insufficient to yield an additive separable representation. The point is that any weakly monotonic utility is weakly separable, but Example 3.3.5 shows that even strongly monotonic utility need not be equivalent to additive separable utility.

Although monotonically increasing utility is sufficient for weak separability relative to the partition of singletons $\mathcal{P}_s$, it is not necessary. Any monotonically decreasing utility function also generates a weakly separable preference order. Example 3.4.4 shows there are also weakly separable preferences that are not monotonic.

Example 3.4.4: Weakly Separable, Non-Monotonic Preferences On $\mathbb{R}^2_+$, the utility function $u(\mathbf{x}) = [1 + (x_1 - 2)^2](x_2 + 1)$ is separable relative to the partition of singletons $\mathcal{P}_s = \big\{\{1\},\{2\}\big\}$, although is not monotonic.

To see that u is separable, note that for any x_2, $(x_1, x_2) \succsim (y_1, x_2)$ if and only if $(x_1 - 2)^2 \geq (y_1 - 2)^2$. This condition does not depend on x_2, u defines a preference order on $P = \{1\}$. Since u is increasing in x_2, Proposition 3.3.4 implies it also defines a preference order on $P = \{2\}$. It follows that u is separable on $\mathcal{P}_s$. ◄

[20] If preferences are only weakly monotonic, some goods may not be essential.

3.4.5 Weak Separability and the MRS

Just as additive separability puts certain restrictions on the marginal rates of substitution (that MRS_{ij} depends only on x_i and x_j), weak separability relative to a partition also places restrictions on the marginal rates of substitution.

The following proposition is established in Problem 3.4.8.

Proposition 3.4.5. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u\colon \mathfrak{X} \to \mathbb{R}$ is a differentiable utility function with $du \gg 0$. Let $\mathcal{P}$ be a partition of goods.*

1. *If u is weakly separable relative to the partition $\mathcal{P}$ and the commodity group $P \in \mathcal{P}$ contains at least two distinct goods i and j, then the marginal rate of substitution MRS_{ij} is independent of x_k for each $k \notin P$.*
2. *If u is strongly separable relative to the partition $\mathcal{P}$, then for every $P_1 \neq P_2 \in \mathcal{P}$ goods $i \in P_1$ and $j \in P_2$, the marginal rate of substitution MRS_{ij} is independent of x_k for every $k \notin P_1 \cup P_2$.*

The assumption that $du \gg 0$ ensures that MRS_{ij} is defined for all goods.

Proposition 3.4.5 has consequences for both weakly and strongly separable utility functions. For $u \in \mathcal{C}^2$ that is weakly separable with respect to a partition $\mathcal{P}$, this implies that for every $P \in \mathcal{P}$ containing at least two goods,

$$\frac{\partial}{\partial x_k}(MRS_{ij}) = 0 \tag{3.4.1}$$

for all distinct $i, j \in P$ and $k \notin P$. If instead u is strongly separable over $\mathcal{P}$, it implies that equation 3.4.9 holds for every $i \in P_1$, $j \in P_2$, and $k \notin (P_1 \cup P_2)$.

3.4.6 Utility that is Weakly but not Strongly Separable

We can sometimes use Proposition 3.4.5 to show that weakly separable utility functions need not be strongly separable, as in the following example.

Example 3.4.6: Weakly but not Strongly Separable Utility This example shows that requiring an induced order on every collection of commodity groups (strong separability) is stronger than merely having an induced order on each commodity group by itself. Consider the utility function $u(x_1, x_2, x_3) = (x_1 + x_1x_2)x_3 + x_3^2$ and use the partition of singletons $\mathcal{P}_s = \{\{1\}, \{2\}, \{3\}\}$. The utility function is separable relative to the partition $\mathcal{P}_s$, as are all increasing utility functions (see Proposition 3.4.3). However, it is not strongly separable on $\mathcal{P}_s$.

By Proposition 3.4.5 with $\mathcal{P}$, if u were strongly separable, MRS_{13} would not depend on x_2. A simple calculation shows that $\text{MRS}_{13} = x_3(1 + x_2)/(x_1 + x_1x_2 + 2x_3)$. As this depends on x_2 (e.g., $\text{MRS}_{13}(1, 1, 1) = 1/2$ while $\text{MRS}_{13}(1, 2, 1) = 3/5$), u cannot be strongly separable. In this case, strong separability and complete separability coincide.

It is also possible to show this directly from the definitions. By trading away from $(1, 1, 1)$ and $(1, 2, 1)$ at the marginal rates of substitution, we can show that preferences cannot be defined on the pair of commodity groups $\{\{1\}, \{3\}\}$. Consider $u(0, 1, 3/2) = 9/4 > u(2, 1, 2/5) = 44/25$ and $u(0, 2, 3/2) = 9/4 < u(2, 2, 2/5) = 64/25$. The amount of good 2 consumed affects the ranking of $(x_1, x_3) = (0, 3/2)$ and $(y_1, y_3) = (2, 2/5)$.

Note that the x_k could be replaced by subutilities to get an example with multiple goods in each commodity group. ◀

3.4.7 The Smooth Separability Theorem

Converses to Proposition 3.4.5 are due to Goldman and Uzawa (1964) and Gorman (1968), showing that preferences with appropriately independent marginal rates of substitution have an additive separable representation. Gorman's version of the theorem is a bit more general. It is based on earlier work by Leontief (1947a, b). This is the version we state. The proof below is that of Goldman and Uzawa, which is a bit simpler, but requires that utility be thrice continuously differentiable.

Smooth Separability Theorem. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose $u \in \mathcal{C}^2$ with $du \gg 0$ on $\mathfrak{X}$. Let $\mathcal{P}$ be a partition of commodities with at least 3 commodity groups. If for every pair P, P' of distinct commodity groups in $\mathcal{P}$, $\frac{\partial}{\partial x_k}\big(\mathrm{MRS}_{ij}\big) = 0$ for every $i \in P$, $j \in P'$, and $k \notin (P \cup P')$, then there is an strictly increasing $\mathcal{C}^2$ function F and strictly increasing $\mathcal{C}^2$ functions v_P on $\mathfrak{X}$ with*

$$u(\mathbf{x}) = F\left(\sum_{P \in \mathcal{P}} v_P(\mathbf{x}_P)\right).$$

3.4.8 The Weak Separability Theorem

The first results characterizing weakly separable functions were obtained by Sono (1943) and Leontief (1947a, b). Sono focused on the case of two commodity groups, while Leontief considered the general case. As Sono found, an additional property was needed for the case of two commodity groups. Goldman and Uzawa (1964) provided a complete characterization of weak separability for well-behaved smooth functions with three or more goods. In general, there must be at least three essential commodities for this to work. We assume $du \gg 0$, which enures that all commodities are essential and all marginal rates of substitution exist.

3.4.9 Debreu's Separability Theorem

For smooth utility functions, we now have a characterization of weak and strong separability, provided there are sufficient essential goods. In particular, additive separable utility functions correspond to completely separable utility. But what if we start with a completely separable preference order? We may not have a utility function, much less a smooth one. Even in this case we can still get an additive separable representation. The key result is a deep theorem of Debreu which we state without proof (see Debreu, 1960 or Fishburn, 1970, Chapter 5).

Separability Theorem (Debreu). *Suppose $\succsim$ is a continuous preference order on $\mathbb{R}^m_+$ that is strongly separable relative to a partition $\mathcal{P}$. If at least three of the commodity groups P are essential, then there are continuous functions u_P for each $P \in \mathcal{P}$ such that $u(\mathbf{x}) = \sum_{P\in\mathcal{P}} u_P(x_P)$ represents $\succsim$.*

Among other things, Debreu's Separability Theorem allows us to characterize all preferences that can be represented by an additive separable utility function. Like the Smooth Separability Theorem, Debreu's Separability Theorem still requires at least 3 essential goods.

Corollary 3.4.7. *Suppose at least 3 goods are essential. A preference order $\succsim$ on $\mathbb{R}^m_+$ is continuous and completely separable if and only if there is a continuous additive separable u that represents $\succsim$.*

Proof. Note that each good is a commodity group in this case. The only if part is an immediate consequence of Debreu's Separability Theorem and the if part follows easily using the separable representation. □

3.4.10 The Importance of at least Three Commodity Groups

In both the Smooth Separability Theorem and Debreu's Separability Theorem, it is important that there are at least three essential commodity groups. Moreover, even the Weak Separability Theorem requires at least three essential goods. The following example shows that when there are only two essential commodity groups, it may not be possible to find an additive separable representation even of strongly separable utility.

Example 3.4.8: Strongly Separable Preferences without Separable Representation Let utility on $\mathbb{R}^2_+$ be defined by $u(x_1, x_2) = x_1 + x_2 \ln(2 + x_1)$. Since u is strictly increasing in each variable, it is weakly separable relative to the partition $\{\{1\}, \{2\}\}$. The only union of commodity groups is $\{1, 2\}$, so u is also strongly separable. This means it is completely separable. However, there are at most two commodity groups, and the Smooth Separability Theorem does not apply.

This utility function cannot be represented in additive separable form. To see that, suppose there is a smooth transformation φ that converts this to an additive separable form $v = \varphi \circ u$ Then $\partial v/\partial x_\ell = \varphi' \partial u/\partial x_\ell$ and

$$\frac{\partial^2 v}{\partial x_1\,\partial x_2} = \varphi' \frac{\partial^2 u}{\partial x_1\,\partial x_2} + \varphi'' \frac{\partial u}{\partial x_1}\frac{\partial u}{\partial x_2}.$$

As this must equal zero to have an additive separable representation, we find

$$\frac{1}{2 + x_1}\varphi' + \varphi'' \frac{2 + x_1 + x_2}{2 + x_1} \ln(2 + x_1) = 0.$$

Substituting $\psi = \varphi'$ and clearing the denominator yields

$$\psi + \psi'(2 + x_1 + x_2) \ln(2 + x_1) = 0. \tag{3.4.14}$$

Now both ψ and ψ' depend only on $u = x_1 + x_2 \ln(2 + x_1)$. Equation 3.4.14 must hold as we vary $\mathbf{x}$ while holding $u(\mathbf{x})$ constant.

Example dontinues . . .

3.4.11 Rest of Example 3.4.8

We are free to vary $(2 + x_1 + x_2)\ln(2 + x_1)$ while holding u fixed. To hold u constant, we set $x_2 = (u - x_1)/\ln(2 + x_1)$ when $x_1 > 0$. Then $(2 + x_1 + x_2)\ln(2 + x_1) = (2 + x_1)\ln(2 + x_1) + u - x_1$, which clearly varies as x_1 varies. Consider equation 3.4.14 at the same value of u with $x_1 = 1$ and $x_1 = 2$. Subtracting the two equations yields $\psi' = 0$, implying $\psi = 0$, which contradicts the assumption that φ is increasing. It follows that there is no transformation of u into an additive separable form. ◀

3.4.12 Separability with Two Commodity Groups

The forgoing makes it pretty clear that there are important differences in separability with three or more commodity groups and separability with two commodity groups. We saw a hint of that earlier when we noticed that weak and strong separability coincide on $\mathbb{R}^2_+$. We will focus on the case of two goods rather than considering all cases with two commodity groups.

When we have only two goods, the usual marginal rate of substitution condition, that MRS_{12} depend only on goods one and two, has become meaningless. All utility functions satisfy that, regardless of whether they are additive separable. Separability of preferences is also too weak. All increasing utility functions are strongly separable relative to the partition of singletons, $\mathcal{P} = \big\{\{1\},\{2\}\big\}$. This type of separability is irrelevant as far as additive separability is concerned.

Additive separable preferences on $\mathbb{R}^2_+$ must have some extra property. But what is it? Start with an additive separable utility function u on $\mathbb{R}^2_+$. We can write $u(x,y) = u_1(x) + u_2(y)$. Suppose $u(x_1,y_1) \ge u(x_2,y_2)$ and $u(x_2,y_3) \ge u(x_3,y_1)$. We can rewrite these as

$$\begin{aligned} u_1(x_1) + u_2(y_1) &\ge u_1(x_2) + u_2(y_2) \\ u_1(x_2) + u_2(y_3) &\ge u_1(x_3) + u_2(y_1). \end{aligned}$$

Adding, we obtain

$$\begin{aligned} u_1(x_1) + {\color{red}u_2(y_1)} + {\color{green}u_1(x_2)} + u_2(y_3) &\ge \\ {\color{green}u_1(x_2)} + u_2(y_2) + u_1(x_3) + {\color{red}u_2(y_1)}&. \end{aligned}$$

Cancelling the terms that are equal (the red and green terms) yields

$$u_1(x_1) + u_2(y_3) \ge u_1(x_3) + u_2(y_2). \tag{3.4.2}$$

3.4.13 The Double Cancellation Property

Equation (3.4.15) tells us

$$u(x_1, y_3) = u(x_3, y_2)$$

whenever $u(x_1, y_1) \geq u(x_2, y_2)$ and $u(x_2, y_3) \geq u(x_3, y_1)$.

In terms of preference, whenever $\succsim$ is equivalent to an additive seple utility function on $\mathbb{R}^2_+$, then whenever $(x_1, y_1) \succsim (x_2, y_2)$ and $(x_2, y_3) \succsim (x_3, y_1)$, we also have $(x_1, y_3) \succsim (x_3, y_2)$.

With this in mind, the *double cancellation condition* holds for $\succsim$ on $\mathbb{R}^2_+$ if whenever $(x_1, y_1) \succsim (x_2, y_2)$ and $(x_2, y_3) \succsim (x_3, y_1)$, we also have $(x_1, y_3) \succsim (x_3, y_2)$.

3.4.14 Double Cancellation Theorem

Theorem 3.4.9. *Let $\mathfrak{X} = \mathbb{R}^2_+$ or $\mathbb{R}^2_{++}$ and suppose $\succsim$ is continuous on $\mathfrak{X}$. Then $\succsim$ obeys the double cancellation condition if and only if $\succsim$ has an additive separable representation.*

Proof. We already showed above that additive separable preferences on $\mathfrak{X}$ obey the double cancellation condition.

If neither good is essential, $u_1 = u_2 = 0$. Now let u represent $\succsim$. If only one good is essential, then either u_1 or $u_2 = u$, with the other function being zero. That leaves the case where both goods are essential.

Theorem 1 of Debreu (1960) can be adapted to prove this (see Theorem 5.4 of Fishburn, 1970, or Wang, 2015). All that is necessary is to remove the symmetry part of the Debreu's Assumption 3, when it becomes the double cancellation condition. Without symmetry, the transformations of the two axes are no longer identical, allowing $u_1 \neq u_2$ (unlike Debreu's Theorem 1).[21] □

Double Cancellation with Commodity Groups. The double cancellation condition can also be written in terms of commodity groups. Fishburn (1970, Theorem 5.4) shows in considerable, but not complete, detail how the proof of Debreu's Separability Theorem applies in that case. Fishburn also considers the multi-group case in the same framework (Fishburn, Theorem 5.5).

[21] The key to Debreu's proof draws on the theory of web geometry as developed by Thomsen (1927), Blaschke (1928), and Blaschke and Bol (1938).

3.4.15 Sono Independence

Sono (1943) had already realized that something more than separability was needed when goods are partitioned into two commodity groups, a condition he called "independence".

Sono Independence. A commodity group P_1 is *Sono independent* of commodity group P_2 if there exist functions $\psi^{ji}(\mathbf{x}_{P_1})$ such that

$$\frac{\partial}{\partial x_i}\big(\ln \mathrm{MRS}_{kj}\big) = \frac{\partial}{\partial x_i}\big(\ln \mathrm{MRS}_{\ell j}\big) = \psi^{ji}(\mathbf{x}_{P_1})$$

for all $i, j \in P_1$ and $k, \ell \in P_2$.

Blackorby, Primot, and Russell (1978) proved the following theorem, which I have recast into our framework.

Theorem 3.4.10. *Let $\mathfrak{X}$ be either $\mathbb{R}^m_+$ or $\mathbb{R}^m_{++}$ and suppose that a utility function u is $\mathcal{C}^2$ on $\mathfrak{X}$ with $du \gg 0$ and that commodities are partitioned into two groups, $\mathcal{P} = \{P_1, P_2\}$. If P_1 is Sono independent of P_2 and $\partial\, \mathrm{MRS}_{k\ell}/\partial x_j = 0$ for all $j \in P_1$ and $k, \ell \in P_2$, then u is equivalent to a group additive separable utility relative to $\mathcal{P}$.*

3.4.16 Quasi-Linear Utility

Theorem 3.4.10 is the type of theorem we need for dealing with quasi-linear utility. Setting $A_j = \sim\{j\} = \{1, \dots, m\} \setminus \{j\}$, we can write utility as $u(\mathbf{x}) = ax_j + \varphi(\mathbf{x}_{A_j})$ where $\varphi\colon \mathbb{R}_+^{A_j} \to \mathbb{R}$.

It is clear that these preferences are separable relative to $\mathcal{P} = \big\{\{j\}, A_j\big\}$, but they have an additional property. If we compute the marginal rate of substitution $\mathrm{MRS}_{\ell j}$ for $\ell \neq j$, we find it is independent of x_j. But then

$$\frac{\partial \ln \mathrm{MRS}_{kj}}{\partial x_j} = \frac{\partial \ln \mathrm{MRS}_{\ell j}}{\partial x_j}$$

for all $k, \ell \in A_j$. It follows that $\{j\}$ is Sono independent of A_j.

Quasi-Linear Representation Theorem. *Let $\mathfrak{X}$ be either $\mathbb{R}_+^m$ or $\mathbb{R}_{++}^m$ and suppose that a utility function u is $\mathcal{C}^2$ on $\mathfrak{X}$ with $\mathbf{d}u \gg \mathbf{0}$. Then $\mathrm{MRS}_{\ell k}$ is independent of x_k for all $\ell \neq k$. if and only if u has an equivalent representation in quasi-linear form.*

Proof. Part I (only if): We showed above that if u is quasi-linear in x_j, it then $\{j\}$ is Sono independent of A_j. It also obeys the separability condition $\partial\,\mathrm{MRS}_{k\ell}/\partial x_j = 0$.

Part II (if): Under these assumptions, Theorem 3.4.10 yields an additive representation of u, $v_j(x_j) + v_{A_j}(\mathbf{x}_{A_j})$.

The marginal rate of substitution $\mathrm{MRS}_{\ell k} = (\partial v_{A_j}/\partial x_\ell)/v_j'(x_j)$ is independent of x_j. Thus, its x_j-derivative is zero. That means

$$0 = \frac{1}{v_j'}\frac{\partial^2 v_{A_j}}{\partial x_j\,\partial x_\ell} - \frac{v_j''}{(v_j')^2}.$$

Since x_j is not part of $\mathbf{x}_{A_j}$, the first term is zero, implying $v_j'' = 0$. It follows that $v_j(x_j) = ax_j + b$ for some a and b, which is the required quasi-linear form. $\square$

3.5 Bergson's Theorem

If we combine additive separability and homotheticity, we obtain further restrictions on the form of the utility function. This was first investigated by Bergson (1936), who consider the additive separable case where the subutilities must either have the Bergson form or the Bernoulii (1738) form.

Rader (1981) extended the result to allow for utility that was not smooth, and considered group additive separable preferences. In order to include the non-differentiable cases, Rader restricted his attention to monotonic and quasi-concave utility. We present a version of the theorem with monotonicity, but without requiring quasi-concavity, at the cost of assuming utility is $\mathcal{C}^2$.

The following form of the theorem requires smoothness, as in Bergson (1936), but does not require concavity and does allow for multiple variables, as in Rader (1981).

3.6 Statement of Bergson's Theorem

Separability Theorem (Bergson). *Suppose* u *is a* $\mathcal{C}^2$ *utility function on* $\mathbb{R}^m_{++}$ *with* $du \gg 0$*. If* u *is both homothetic and group additive separable relative to a partition* $\mathcal{P}$ *with at least two commodity groups, then* u *has one of the following two forms:*

1. *(Bergson) For some* $\gamma \neq 0$*, there are constants* c *and* b_P *and functions* v_P*, homogeneous of degree one on* $\mathbb{R}^P_{++}$ *and* b_P *obeying* $b_P \gamma dv_P \gg 0$ *with*

$$u(\mathbf{x}) = c + \sum_{P \in \mathcal{P}} b_P \big(v_P(x_P)\big)^{\gamma}.$$

2. *(Modified Bernoulli) There are constants* b_P *obeying* $b_P \, dv_P \geq 0$*, and functions* ϕ_P *and* v_P*, homogeneous of degree zero and one on* $\mathbb{R}^P_{++}$*, respectively, with* $v_P > 0$*, such that*

$$u(\mathbf{x}) = \sum_{P \in \mathcal{P}} \phi_P(\mathbf{x}_P) + \sum_{P \in \mathcal{P}} b_P \ln v_P(x_P).$$

3.6.1 Proof of Bergson's Theorem I

Proof. Group additive separability tells us that $u(\mathbf{x}) = \sum_{P\in\mathcal{P}} u_P(x_P)$ for some functions u_P. These functions must be $\mathcal{C}^2$ since $u \in \mathcal{C}^2$. Take k in commodity group P and ℓ in $P' \neq P$. Because $du \gg \mathbf{0}$, we can form the marginal rate of substitution,

$$\mathrm{MRS}_{k\ell} = \frac{\big(\partial u_P/\partial x_k\big)(\mathbf{x}_P)}{\big(\partial u_{P'}/\partial x_\ell\big)(\mathbf{x}_{P'})}.$$

Let φ_k denote the marginal utility $\varphi_k(\mathbf{x}_P) = \big(\partial u_P/\partial x_k\big)(\mathbf{x}_P)$. Then $\mathrm{MRS}_{k\ell} = \varphi_k(\mathbf{x}_P)/\varphi_\ell(\mathbf{x}_{P'})$. By the homotheticity of u and Theorem 3.1.2, $\mathrm{MRS}_{k\ell}$ is homogeneous of degree zero, yielding

$$\frac{\varphi_k(t\mathbf{x}_P)}{\varphi_\ell(t\mathbf{x}_{P'})} = \frac{\varphi_k(\mathbf{x}_P)}{\varphi_\ell(\mathbf{x}_{P'})}.$$

for all $t > 0$. Now use the chain rule to take the t-derivative of both sides and rearrange, to obtain

$$\varphi_\ell(t\mathbf{x}_{P'})\big[D_{\mathbf{x}}\varphi_k(t\mathbf{x}_P)\big]\mathbf{x}_P = \varphi_k(t\mathbf{x}_P)\big[D_{\mathbf{x}}\varphi_\ell(t\mathbf{x}_{P'})\big]\mathbf{x}_{P'}.$$

Then set $t = 1$ and collect the P terms on the left and P' terms on the right to find

$$\frac{\big[D_{\mathbf{x}}\varphi_k(\mathbf{x}_P)\big]\mathbf{x}_P}{\varphi_k(\mathbf{x}_P)} = \frac{\big[D_{\mathbf{x}}\varphi_\ell(\mathbf{x}_{P'})\big]\mathbf{x}_{P'}}{\varphi_\ell(\mathbf{x}_{P'})} \tag{3.5.16}$$

Proof concludes on next page . . .

3.6.2 Proof of Bergson's Theorem II

Rest of Proof. The left side of equation 3.5.16 depends only on x_P and the right side depends only on $x_{P'}$. That means that both must be constant. We call the common value β. Notice that since P and P′ were arbitrary, the same constant β applies to all such ratios. We then write

$$\beta = \frac{\big[D_{\mathbf{x}}\varphi_k(\mathbf{x}_P)\big]\mathbf{x}_P}{\varphi_k(\mathbf{x}_P)} = \frac{\big[D_{\mathbf{x}}\varphi_\ell(\mathbf{x}_{P'})\big]\mathbf{x}_{P'}}{\varphi_\ell(\mathbf{x}_{P'})}$$

But then

$$\beta\varphi_k(\mathbf{x}_P) = \big[D_{\mathbf{x}}\varphi_k(\mathbf{x}_P)\big]\mathbf{x}_P \tag{3.5.17}$$

Equation 3.5.17 applies to any P and $k \in P$. By Euler's Theorem, each φ_k is homogeneous of degree β. Since this holds for every $k \in P$, $\varphi_P = D_{\mathbf{x}_P}u_P$ is homogeneous of degree β in $\mathbf{x}_P$ because each component of φ_P is homogeneous of degree β in $\mathbf{x}_P$.

An appeal to Theorem 3.1.7 shows that either for all $P \in \mathcal{P}$, we can write $u_P(\mathbf{x}_P) = c_P + b_P\big(\nu_P(\mathbf{x}_P)\big)^\gamma$ where $\gamma = 1 + \beta \neq 0$, or that for all $P \in \mathcal{P}$, we can write $u_P(\mathbf{x}_P) = \phi_P + b_P \ln \nu_P(\mathbf{x}_P)$. In both cases, $\nu_P > 0$ is homogeneous of degree one. The functions ϕ_P are homogeneous of degree zero.

The fact that $du \gg \mathbf{0}$ is responsible for the requirement $b_P\gamma\, d\nu_P \gg 0$. $\square$

January 13, 2023