

Multivariate Differential Calculus

Dr Damien S. Eldridge

Australian National University

13 May 2021

Sources Part 1

- This material does not originate with me. It has been drawn from, or influenced by, a number of sources. Some of these include those listed below.
- Introductory Mathematical Economics Textbooks.
 - Bradley, T (2008), *Essential mathematics for economics and business (third edition)*, John Wiley and Sons, Great Britain: Chapter 7, Sections 1 to 3 (pp. 360-408).
 - Haeussler, EF Jr and RS Paul (1987), *Introductory mathematical analysis for business, economics, and the life and social sciences (fifth edition)*, Prentice-Hall International Edition, Prentice-Hall, USA: Chapter 17, Sections 1 to 7 and Sections 9 to 10 (pp. 668-706 and 714-723).
 - Shannon, J (1995), *Mathematics for business, economics and finance*, John Wiley and Sons, Brisbane: Chapter 10, Sections 1 to 6 (pp. 450-489).
 - Sydsaeter, K, P Hammond, A Strom, and A Carvajal (2016), *Essential mathematics for economic analysis (fifth edition)*, Pearson Education, United Kingdom: Chapters 11 and 12 (pp. 407-494).

- Intermediate Mathematical Economics Textbooks.

- Chiang, AC (1984), *Fundamental methods of mathematical economics (third edition)*, McGraw-Hill, Singapore: Chapters 6-8 and Chapter 10 (pp. 127-227 and 268-306).
- Chiang, AC and K Wainwright (2005), *Fundamental methods of mathematical economics (fourth edition)*, McGraw-Hill, Singapore: Chapters 6 to 10 (pp. 124-290).
- Simon, C and L Blume (1994), *Mathematics for economists*, WW Norton and Co, USA: Chapters 2 to 5 (pp. 10-103).

- Mathematics Textbooks.

- Spiegel, MR (1981a), *Theory and problems of advanced calculus (SI (metric) edition)*, Schaum's Outline Series, McGraw-Hill Book Company, Singapore (Chapters 4, 6, 7 and 8; pp. 57-79 and 101-179); and
- Spiegel, MR (1981b), *Theory and problems of vector analysis and an introduction to tensor analysis (SI (metric) edition)*, Schaum's Outline Series, McGraw-Hill Book Company, Singapore (Chapters 3 and 4; pp. 35-81).

Our Objective

- We want to extend our discussion of differential calculus to a setting in which there is more than one choice variable.
- In other words, we want to examine multivariate real-valued functions.
- These are functions of the form $f : X \rightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ and $n \in \mathbb{N} \setminus \{1\}$ (that is, n is a natural number other than the number one).
- Often we will write such functions as $f(x_1, x_2, \dots, x_n)$.
- We will be focussing on two concepts. These are partial derivatives and total differentials.

Single-Real-Valued Multi-Variate Functions

- Let $f : X \longrightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ and $n \in \mathbb{N} \setminus \{1\}$.
- This type of function is a single-real-valued multivariate function.
 - It is single-real-valued because, for any choice of inputs (independent variables), it produces a single real number as the output (dependent variable). The output might vary with the choice of inputs.
 - It is multi-variate because there is more than one input (independent variable).
- Recall the following.
 - $X \subseteq \mathbb{R}^n$ is the domain of f .
 - \mathbb{R} is the co-domain of f .
 - $f(X) = \{y \in \mathbb{R} : y = f(x) \text{ for some } x \in X\}$ is the image of X under f .
 - Sometimes we might refer to $f(X)$ as the range of the function f .
 - Clearly $f(X) \subseteq \mathbb{R}$.
 - Since f is a function, we know that it associates only one point in $f(X)$ with any particular point in X .

Partial Derivatives Part 1

- Consider a function $f : X \rightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$. This function can be written as $f(x_1, x_2, \dots, x_n)$.
- Suppose that we hold the values that are taken by $(n - 1)$ of the independent variables constant and only allow one of the independent variables to take on different values. Specifically, suppose that we fix the value of $x_i = \bar{x}_i$ for all $i \in \{1, 2, \dots, n\} \setminus \{k\}$ and allow only x_k to vary.
- To simplify notation, let

$$x_{-k} = (x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$$

and

$$\bar{x}_{-k} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{k-1}, \bar{x}_{k+1}, \dots, \bar{x}_n).$$

Partial Derivatives Part 2

- We can now think about the multivariate function as a univariate function, with x_k being the sole independent variable.
- To be precise, we have

$$g(x_k) = f(x_k; \bar{x}_{-k}).$$

- Suppose that the univariate function $g(x_k)$ is differentiable and that its derivative is given by

$$g'(x_k) = \frac{dg(x_k)}{dx_k}.$$

Partial Derivatives Part 3

- We use this derivative to define the partial derivative of the function $f(x_1, x_2, \dots, x_n)$ with respect to the variable x_k . To be precise, we define this partial derivative as

$$\frac{\partial f(x_k, x_{-k})}{\partial x_k} \equiv \frac{dg(x_k)}{dx_k}.$$

- As a matter of convenience will sometimes use the following notation for a partial derivative:

$$f_k(x_k, x_{-k}) = \frac{\partial f(x_k, x_{-k})}{\partial x_k}.$$

- Note that $f_k(x_k; \bar{x}_{-k})$ will potentially depend on the values taken by all of the independent variables, not just on the one that is allowed to vary.

Partial Derivatives Part 4

- The formal definition of the partial derivative of the function $f(x_1, x_2, \dots, x_n)$ with respect to the variable x_k is

$$\begin{aligned}\frac{\partial f(x_k, x_{-k})}{\partial x_k} &\equiv \lim_{h \rightarrow 0} \left\{ \frac{f(x_k + h, x_{-k}) - f(x_k, x_{-k})}{(x_k + h) - x_k} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{f(x_k + h, x_{-k}) - f(x_k, x_{-k})}{h} \right\}.\end{aligned}$$

- Note that if we think of $f(x_k, x_{-k})$ as a univariate function of x_k alone, then this definition collapses to the definition of the first-order derivative of a univariate function.

Partial Derivatives Part 5

- Let $S \subseteq \mathbb{R}^n$ be an open set.
- Consider a function $f : S \rightarrow \mathbb{R}$.
- This function can be written as $f(x_1, x_2, \dots, x_n)$.
- The vector of first-order partial derivatives of f at the point x , $\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T$, is sometimes known as the gradient vector of f at x . It is commonly denoted by one of $D_x f(x)$, $\nabla f(x)$, or $\text{grad } f(x)$.
- Theorem: If all n of the first-order partial derivatives of f exist and are continuous at the point $x \in X \subseteq \mathbb{R}^n$, then the function f is differentiable at the point $x \in X \subseteq \mathbb{R}^n$.
- Terminology: The function $f(x)$ is said to be “(at least) once continuously differentiable on X ” if and only if all of the first-order partial derivatives of f exist and are continuous on X .

Partial Derivatives Example 1 Part 1

Let $f(x, y) = 2x^2 - xy + y^2$. We will use the definition of a partial derivative find the partial derivative of f with respect to x . We have

$$\begin{aligned}\frac{\partial f}{\partial x} &= \lim_{h \rightarrow 0} \left\{ \frac{f(x+h, y) - f(x, y)}{(x+h) - x} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{f(x+h, y) - f(x, y)}{h} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{(2(x+h)^2 - (x+h)y + y^2) - (2x^2 - xy + y^2)}{h} \right\}\end{aligned}$$

Partial Derivatives Example 1 Part 2

$$\begin{aligned} &= \lim_{h \rightarrow 0} \left\{ \frac{2(x^2 + 2xh + h^2) - xy - hy}{h} \right. \\ &\quad \left. + \frac{y^2 - 2x^2 + xy - y^2}{h} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{2x^2 + 4xh + 2h^2 - xy - hy}{h} \right. \\ &\quad \left. + \frac{y^2 - 2x^2 + xy - y^2}{h} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{4xh + 2h^2 - hy}{h} \right\} \\ &= \lim_{h \rightarrow 0} \{4x + 2h - y\} \\ &= 4x + 2(0) - y \\ &= 4x - y. \end{aligned}$$

Partial Derivatives Example 2

Let $f(x, y) = x^3y + e^{xy^2} = x^3y + \exp(xy^2)$. We will use our knowledge of univariate calculus to find the partial derivative of f with respect to x and the partial derivative of f with respect to y . We have

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x} \{x^3y + \exp(xy^2)\} \\ &= \frac{\partial}{\partial x} \{x^3y\} + \frac{\partial}{\partial x} \{\exp(xy^2)\} \\ &= 3x^2y + y^2 \exp\{xy^2\},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y} \{x^3y + \exp(xy^2)\} \\ &= \frac{\partial}{\partial y} \{x^3y\} + \frac{\partial}{\partial y} \{\exp(xy^2)\} \\ &= x^3 + 2xy \exp\{xy^2\}.\end{aligned}$$

Some Economic Applications

- Marginal utilities;
- Marginal products;
- Various elasticities of demand;
- Cournot aggregation; and
- Engel aggregation.

Marginal Utilities Part 1

- We will provide some illustrations of an application of the concept of a partial derivative by using a variety of different two-good utility functions. These utility functions take the form $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$. We will think of this as a function that ranks bundles of two commodities according to their desirability to the consumer.
- The utility function may be written as

$$U = U(Q_1, Q_2),$$

where U is the level of utility (which is an ordinal concept, not a cardinal one, for people like me who worry about such things), Q_1 is the amount of good one that is consumed and Q_2 is the amount of good two that is consumed.

- We will consider a perfect substitutes utility function, a generalised perfect substitutes utility function, a Cobb-Douglas utility function and a constant elasticity of substitution utility function.
- We should emphasise that because utility is only an ordinal concept (only ranking is important), the concept of marginal utility is not a meaningful one. However, the related concept of a marginal rate of substitution is meaningful.

Marginal Utilities Part 3

- Suppose that we have a perfect substitutes utility function:

$$U(Q_1, Q_2) = Q_1 + Q_2.$$

- The marginal utility of good one is simply the first-order partial derivative of this utility function with respect to the quantity of good one that is consumed. Thus we have

$$MU_1(Q_1, Q_2) = \frac{\partial U(Q_1, Q_2)}{\partial Q_1} = \frac{\partial (Q_1 + Q_2)}{\partial Q_1} = 1.$$

- The marginal utility of good two is simply the first-order partial derivative of this utility function with respect to the quantity of good two that is consumed. Thus we have

$$MU_2(Q_1, Q_2) = \frac{\partial U(Q_1, Q_2)}{\partial Q_2} = \frac{\partial (Q_1 + Q_2)}{\partial Q_2} = 1.$$

Marginal Utilities Part 4

- Suppose that we have a generalised perfect substitutes utility function:

$$U(Q_1, Q_2) = \alpha Q_1 + \beta Q_2.$$

- The marginal utility of good one is simply the first-order partial derivative of this utility function with respect to the quantity of good one that is consumed. Thus we have

$$MU_1(Q_1, Q_2) = \frac{\partial U(Q_1, Q_2)}{\partial Q_1} = \frac{\partial (\alpha Q_1 + \beta Q_2)}{\partial Q_1} = \alpha.$$

- The marginal utility of good two is simply the first-order partial derivative of this utility function with respect to the quantity of good two that is consumed. Thus we have

$$MU_2(Q_1, Q_2) = \frac{\partial U(Q_1, Q_2)}{\partial Q_2} = \frac{\partial (\alpha Q_1 + \beta Q_2)}{\partial Q_2} = \beta.$$

- Suppose that we have a Cobb-Douglas utility function:

$$U(Q_1, Q_2) = Q_1^\alpha Q_2^{1-\alpha}.$$

- The marginal utility of good one is simply the first-order partial derivative of this utility function with respect to the quantity of good one that is consumed.
- Continued on the next slide.

Marginal Utilities Part 6

- Continued from the previous slide.
- Thus we have

$$\begin{aligned} MU_1(Q_1, Q_2) &= \frac{\partial U(Q_1, Q_2)}{\partial Q_1} \\ &= \frac{\partial Q_1^\alpha Q_2^{1-\alpha}}{\partial Q_1} \\ &= \alpha Q_1^{\alpha-1} Q_2^{1-\alpha} \\ &= \alpha Q_1^{-(1-\alpha)} Q_2^{1-\alpha} \\ &= \alpha \left(\frac{Q_2^{1-\alpha}}{Q_1^{1-\alpha}} \right) \\ &= \alpha \left(\frac{Q_2}{Q_1} \right)^{1-\alpha}. \end{aligned}$$

Marginal Utilities Part 7

- The marginal utility of good two is simply the first-order partial derivative of this utility function with respect to the quantity of good two that is consumed. Thus we have

$$\begin{aligned} MU_2(Q_1, Q_2) &= \frac{\partial U(Q_1, Q_2)}{\partial Q_2} \\ &= \frac{\partial Q_1^\alpha Q_2^{1-\alpha}}{\partial Q_2} \\ &= (1-\alpha) Q_1^\alpha Q_2^{1-\alpha-1} \\ &= (1-\alpha) Q_1^\alpha Q_2^{-\alpha} \\ &= (1-\alpha) \left(\frac{Q_1^\alpha}{Q_2^\alpha} \right) \\ &= (1-\alpha) \left(\frac{Q_1}{Q_2} \right)^\alpha. \end{aligned}$$

Marginal Utilities Part 8

- Suppose that we have a constant elasticity of substitution utility function:

$$U(Q_1, Q_2) = (\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{1}{\rho}}.$$

- The marginal utility of good one is simply the first-order partial derivative of this utility function with respect to the quantity of good one that is consumed. Thus we have

$$\begin{aligned} MU_1(Q_1, Q_2) &= \frac{\partial U(Q_1, Q_2)}{\partial Q_1} \\ &= \frac{\partial \left((\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{1}{\rho}} \right)}{\partial Q_1} \\ &= \left(\frac{1}{\rho} \right) (\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{1}{\rho}-1} (\rho \alpha Q_1^{\rho-1}) \\ &= \alpha Q_1^{\rho-1} (\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{(1-\rho)}{\rho}}. \end{aligned}$$

Marginal Utilities Part 9

- The marginal utility of good two is simply the first-order partial derivative of this utility function with respect to the quantity of good two that is consumed. Thus we have

$$\begin{aligned} MU_2(Q_1, Q_2) &= \frac{\partial U(Q_1, Q_2)}{\partial Q_2} \\ &= \frac{\partial \left((\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{1}{\rho}} \right)}{\partial Q_2} \\ &= \left(\frac{1}{\rho} \right) (\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{1}{\rho}-1} (\rho \beta Q_2^{\rho-1}) \\ &= \beta Q_2^{\rho-1} (\alpha Q_1^\rho + \beta Q_2^\rho)^{\frac{(1-\rho)}{\rho}}. \end{aligned}$$

Marginal Products Part 1

- We will provide some illustrations of an application of the concept of a partial derivative by using a variety of different two-input and one-output production technologies. These technologies can be represented by a production function of the form $f : \mathbb{R}_+^2 \longrightarrow \mathbb{R}_+$. We will think of this as a function that turns bundles of labour and capital into an amount of output.
- The production function may be written as

$$Q = f(L, K),$$

where $Q = f(L, K)$ is the quantity of output that is produced when L units of labour and K units of capital are employed.

- We will consider a perfect substitutes production function, a generalised perfect substitutes production function, a Cobb-Douglas production function and a constant elasticity of substitution production function.

Marginal Products Part 2

- Suppose that we have a perfect substitutes production function:

$$f(L, K) = L + K.$$

- The marginal product of labour is simply the first-order partial derivative of this production function with respect to the quantity of labour that is employed. Thus we have

$$MP_L(L, K) = \frac{\partial f(L, K)}{\partial L} = \frac{\partial (L + K)}{\partial L} = 1.$$

- The marginal product of capital is simply the first-order partial derivative of this production function with respect to the quantity of capital that is employed. Thus we have

$$MP_K(L, K) = \frac{\partial f(L, K)}{\partial K} = \frac{\partial (L + K)}{\partial K} = 1.$$

Marginal Products Part 3

- Suppose that we have a generalised perfect substitutes production function:

$$f(L, K) = \alpha L + \beta K.$$

- The marginal product of labour is simply the first-order partial derivative of this production function with respect to the quantity of labour that is employed. Thus we have

$$MP_L(L, K) = \frac{\partial f(L, K)}{\partial L} = \frac{\partial (\alpha L + \beta K)}{\partial L} = \alpha.$$

- The marginal product of capital is simply the first-order partial derivative of this production function with respect to the quantity of capital that is employed. Thus we have

$$MP_K(L, K) = \frac{\partial f(L, K)}{\partial K} = \frac{\partial (\alpha L + \beta K)}{\partial K} = \beta.$$

Marginal Products Part 4

- Suppose that we have a Cobb-Douglas production function:

$$f(L, K) = AL^{\alpha}K^{\beta}.$$

- The marginal product of labour is simply the first-order partial derivative of this production function with respect to the quantity of labour that is employed. Thus we have

$$\begin{aligned}MP_L(L, K) &= \frac{\partial f(L, K)}{\partial L} \\&= \frac{\partial (AL^{\alpha}K^{\beta})}{\partial L} \\&= \alpha AL^{\alpha-1}K^{\beta}.\end{aligned}$$

Marginal Products Part 5

- The marginal product of capital is simply the first-order partial derivative of this production function with respect to the quantity of capital that is employed. Thus we have

$$\begin{aligned}MP_K(L, K) &= \frac{\partial f(L, K)}{\partial K} \\&= \frac{\partial (AL^\alpha K^\beta)}{\partial K} \\&= \beta AL^\alpha K^{\beta-1}.\end{aligned}$$

Marginal Products Part 6

- Suppose that we have a constant elasticity of substitution production function:

$$f(L, K) = (\alpha L^\rho + \beta K^\rho)^{\frac{1}{\rho}}.$$

- The marginal product of labour is simply the first-order partial derivative of this production function with respect to the quantity of labour that is employed. Thus we have

$$\begin{aligned} MP_L(L, K) &= \frac{\partial f(L, K)}{\partial L} \\ &= \frac{\partial \left((\alpha L^\rho + \beta K^\rho)^{\frac{1}{\rho}} \right)}{\partial L} \\ &= \left(\frac{1}{\rho} \right) (\alpha L^\rho + \beta K^\rho)^{\frac{1}{\rho}-1} (\rho \alpha L^{\rho-1}) \\ &= \alpha L^{\rho-1} (\alpha L^\rho + \beta K^\rho)^{\frac{(1-\rho)}{\rho}}. \end{aligned}$$

Marginal Products Part 7

- The marginal product of capital is simply the first-order partial derivative of this production function with respect to the quantity of capital that is employed. Thus we have

$$\begin{aligned}MP_K(L, K) &= \frac{\partial f(L, K)}{\partial K} \\&= \frac{\partial \left((\alpha L^\rho + \beta K^\rho)^{\frac{1}{\rho}} \right)}{\partial K} \\&= \left(\frac{1}{\rho} \right) (\alpha L^\rho + \beta K^\rho)^{\frac{1}{\rho}-1} (\rho \beta K^{\rho-1}) \\&= \beta K^{\rho-1} (\alpha L^\rho + \beta K^\rho)^{\frac{(1-\rho)}{\rho}}.\end{aligned}$$

Elasticities of Demand Part 1

- Suppose that an individual's Marshallian (or “Walrasian” or “ordinary” or “uncompensated”) demand function for commodity k is given by

$$Q_k = D_k(p_1, p_2, \dots, p_n, y),$$

where p_i is the price of commodity i for each $i \in \{1, 2, \dots, n\}$ and y is the consumer's income.

- The own-price elasticity of demand for commodity k for this consumer is

$$\begin{aligned}\epsilon_k^k(p_1, p_2, \dots, p_n, y) &= \left(\frac{p_k}{Q_k} \right) \left(\frac{\partial D_k}{\partial p_k} \right) \\ &= \left(\frac{p_k}{D_k(p_1, p_2, \dots, p_n, y)} \right) \left(\frac{\partial D_k}{\partial p_k} \right).\end{aligned}$$

Elasticities of Demand Part 2

- The cross-price elasticity of demand for commodity k with respect to the price of commodity l for this consumer is

$$\begin{aligned}\varepsilon_l^k(p_1, p_2, \dots, p_n, y) &= \left(\frac{p_l}{Q_k} \right) \left(\frac{\partial D_k}{\partial p_l} \right) \\ &= \left(\frac{p_l}{D_k(p_1, p_2, \dots, p_n, y)} \right) \left(\frac{\partial D_k}{\partial p_l} \right).\end{aligned}$$

- The income elasticity of demand for commodity k for this consumer is

$$\begin{aligned}\varepsilon_y^k(p_1, p_2, \dots, p_n, y) &= \left(\frac{y}{Q_k} \right) \left(\frac{\partial D_k}{\partial y} \right) \\ &= \left(\frac{y}{D_k(p_1, p_2, \dots, p_n, y)} \right) \left(\frac{\partial D_k}{\partial y} \right).\end{aligned}$$

Elasticities of Demand Part 3

- We can use these elasticities to classify the types of commodities that are being considered.
 - If $\varepsilon_y^k > 0$, then commodity k is a normal good. If $\varepsilon_y^k < 0$, then commodity k is an inferior good.
 - If $\varepsilon_l^k > 0$, then commodities k and l are substitutes. If $\varepsilon_l^k < 0$, then commodities k and l are complements.
- The Marshallian demand curve for most commodities will usually slope down. As such, we would usually expect $\varepsilon_k^k < 0$.
 - However, there are circumstances in which the Marshallian demand curve for a commodity can slope up over some range of prices (at least in theory). Such commodities are known as Giffen goods. In such circumstances, we would have $\varepsilon_k^k > 0$ over the relevant range of prices. Note that a necessary, but not sufficient, condition for a commodity to be a Giffen good is that it be an inferior good.

Cournot Aggregation Part 1

- Suppose that a consumer's preferences over bundles of L commodities are locally non-satiated in the neighbourhood of any potentially feasible commodity bundle. Then we know that budget exhaustion (which is sometimes called Walras' law for the individual) must hold for the consumer.
- This ensures that

$$\sum_{l=1}^L p_l x_l(p, y) = y,$$

where $p = (p_1, p_2, \dots, p_n) = (p_k, p_{-k})$ is the price vector, y is the consumer's income and $x_l(p, y)$ is the consumer's Marshallian demand for good l .

- Note that this can be rewritten as

$$p_k x_k(p_k, p_{-k}, y) + \sum_{l \neq k} p_l x_l(p, y) = y.$$

Cournot Aggregation Part 2

- Partially differentiating both sides of this equation with respect to the price of commodity k , we obtain

$$\left\{ (1) x_k(p, y) + p_k \left(\frac{\partial x_k(p, y)}{\partial p_k} \right) \right\} + \sum_{l \neq k} p_l \left(\frac{\partial x_l(p, y)}{\partial p_k} \right) = 0.$$

- This can be simplified to obtain

$$x_k(p, y) + \sum_{l=1}^L p_l \frac{\partial x_l(p, y)}{\partial p_k} = 0.$$

- This can be rearranged to obtain

$$\sum_{l=1}^L p_l \frac{\partial x_l(p, y)}{\partial p_k} = -x_k(p, y).$$

Cournot Aggregation Part 3

- This can be rewritten as

$$\sum_{l=1}^L \left(\frac{p_k}{p_k} \right) \left(\frac{x_l(p, y)}{x_l(p, y)} \right) \left(\frac{y}{y} \right) p_l \frac{\partial x_l(p, y)}{\partial p_k} = -x_k(p, y).$$

- This can be rearranged to obtain

$$\sum_{l=1}^L \left(\frac{p_l x_l(p, y)}{y} \right) \left(\frac{p_k}{x_l(p, y)} \right) \left(\frac{\partial x_l(p, y)}{\partial p_k} \right) = - \left(\frac{p_k x_k(p, y)}{y} \right).$$

Cournot Aggregation Part 4

- This can be rewritten as

$$\sum_{l=1}^L s_l \varepsilon_k^l = -s_k,$$

where

$$s_l = \frac{p_l x_l(p, y)}{y}$$

is the budget share of commodity l and

$$\varepsilon_k^l = \left(\frac{p_k}{x_l(p, y)} \right) \left(\frac{\partial x_l(p, y)}{\partial p_k} \right)$$

is the k th commodity-price elasticity of demand for commodity l .

- The above formula is a result known as Cournot aggregation. It provides a relationship between the k th price elasticities of demand for the various commodities.

Engel Aggregation Part 1

- Suppose that a consumer's preferences over bundles of L commodities are locally non-satiated in the neighbourhood of any potentially feasible commodity bundle. Then we know that budget exhaustion (which is sometimes called Walras' law for the individual) must hold for the consumer.
- This ensures that

$$\sum_{l=1}^L p_l x_l(p, y) = y,$$

where $p = (p_1, p_2, \dots, p_n)$ is the price vector, y is the consumer's income and $x_l(p, y)$ is the consumer's Marshallian demand for good l .

- Partially differentiating both sides of this equation with respect to income, we obtain

$$\sum_{l=1}^L p_l \left(\frac{\partial x_l(p, y)}{\partial y} \right) = 1.$$

Engel Aggregation Part 2

- This can be rewritten as

$$\sum_{l=1}^L \left(\frac{y}{y} \right) \left(\frac{x_l(p, y)}{x_l(p, y)} \right) p_l \left(\frac{\partial x_l(p, y)}{\partial y} \right) = 1.$$

- This can be simplified to obtain

$$\sum_{l=1}^L \left(\frac{p_l x_l(p, y)}{y} \right) \left(\frac{y}{x_l(p, y)} \right) \left(\frac{\partial x_l(p, y)}{\partial y} \right) = 1.$$

Engel Aggregation Part 3

- This can be rewritten as

$$\sum_{l=1}^L s_l \varepsilon_y^l = 1,$$

where

$$s_l = \frac{p_l x_l(p, y)}{y}$$

is the budget share of commodity l and

$$\varepsilon_y^l = \left(\frac{y}{x_l(p, y)} \right) \left(\frac{\partial x_l(p, y)}{\partial y} \right)$$

is the income elasticity of demand for commodity l .

- The above formula is a result known as Engel aggregation. It provides a relationship between the income elasticities of demand for the various commodities.

Higher Order Derivatives Part 1

- Let $X \subseteq \mathbb{R}^n$ be an open set and $f : X \rightarrow \mathbb{R}$ be a function that is differentiable on X , with $D_x f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)^T$. The vector $D_x f(x)$ is called the gradient vector of f at the point $x \in S$.
- Note that $\frac{\partial f}{\partial x_k} : X \rightarrow \mathbb{R}$ is itself a function.
- Suppose that the function $\frac{\partial f}{\partial x_k}(x)$ is itself differentiable at the point $x \in X$. Often, we will want this to be the case for all $k \in \{1, 2, \dots, n\}$.
 - Consider one such partial derivative, say $\frac{\partial f}{\partial x_i}$.
 - For each $i \in \{1, 2, \dots, n\}$, we have n “second-order” partial derivatives of $\frac{\partial f}{\partial x_i}$. These take the form $\frac{\partial \left(\frac{\partial f}{\partial x_i} \right)}{\partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_i}$ for each $k \in \{1, 2, \dots, n\}$.
 - Note that there are n^2 such second-order partial derivatives of f in total.
 - Thus we have $D_x \tau \left(\frac{\partial f}{\partial x_i} \right) = \left(\frac{\partial^2 f}{\partial x_1 \partial x_i}, \frac{\partial^2 f}{\partial x_2 \partial x_i}, \dots, \frac{\partial^2 f}{\partial x_n \partial x_i} \right)$ for each $i \in \{1, 2, \dots, n\}$.

Higher Order Derivatives Part 2

- Let $X \subseteq \mathbb{R}^n$ be an open set and $f : X \rightarrow \mathbb{R}$ be a function that is differentiable on X , with $D_x f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)^T$.
- Suppose that each of the partial derivatives that make up $D_x f$ are themselves differentiable functions at the point $x \in X$.
- The derivative of the gradient vector $D_x f$ at the point $x \in X$ is given by the matrix

$$D_{x^T} (D_x f(x)) = D_{xx^T}^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix},$$

where $\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{\partial^2 f(x)}{\partial x_i \partial x_i}$.

Higher Order Derivatives Part 3

- The matrix $D_{xx}^2 f(x)$ is called the Hessian matrix of f at the point $x \in S$.
- If f is twice-differentiable at every point $x \in X$, then f is said to be twice-differentiable on X .
- If f is twice-differentiable on X and $\frac{\partial^2 f}{\partial x_k \partial x_i}$ is a continuous function on X for all $(i, k) \in \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$, then f is said to be twice continuously differentiable on X . This is denoted by $f \in C^2$ on X .
- Young's Theorem: If $f \in C^2$ on X , then $\frac{\partial^2 f}{\partial x_k \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_k}$ for all $i \neq k$.
- Young's Theorem ensures that the Hessian matrix for f will be symmetric when $f \in C^2$ on X .

Higher Order Derivatives Part 4

- It is possible to extend the process of differentiation for multivariate functions to even higher orders than second-order derivatives.
- Doing so for partial derivatives is relatively straight-forward.
- However, this will not be done in this course.

Second-Order Partial Derivatives Part 1

- Consider a function $f : X \longrightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ is an open set. This function can be written as $f(x_1, x_2, \dots, x_n)$. Assume that this function is at least twice continuously differentiable with respect to all of its arguments.
- Suppose that we want to find the second-order partial derivative of the underlying function $f(x_1, x_2, \dots, x_n)$ with respect to the variable x_i first and then with respect to the variable x_j second.
- This notation for this second-order derivative is

$$\frac{\partial^2 f(x_1, x_2, \dots, x_n)}{\partial x_j \partial x_i} = f_{ij}(x_1, x_2, \dots, x_n).$$

Second-Order Partial Derivatives Part 2

- Note that the first-order partial derivative of this function with respect to the variable x_i is a function in its own right:

$$\begin{aligned}\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} &= f_i(x_1, x_2, \dots, x_n) \\ &= g^i(x_1, x_2, \dots, x_n).\end{aligned}$$

- We can use this function $g^i(x_1, x_2, \dots, x_n)$ to define the second-order partial derivative of the underlying function $f(x_1, x_2, \dots, x_n)$ with respect to the variable x_i first and then with respect to the variable x_j second. To be precise, we have

$$f_{ij}(x_1, x_2, \dots, x_n) \equiv \frac{\partial g^i(x_1, x_2, \dots, x_n)}{\partial x_j}.$$

- Thus we have

$$\frac{\partial^2 f(x_1, x_2, \dots, x_n)}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left(\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} \right).$$

Young's Theorem Part 1

- Consider a function $f : X \rightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ is an open set. This function can be written as $f(x_1, x_2, \dots, x_n)$.
- Suppose that this function is at least twice continuously differentiable with respect to all of its arguments in a non-empty neighbourhood around the point

$$x^0 = (x_1^0, x_2^0, \dots, x_n^0).$$

- In this case, the order of differentiation will not affect the second-order partial derivatives evaluated at the point x^0 .

Young's Theorem Part 2

- More precisely, in this case we will have

$$\frac{\partial^2 f(x_1^0, x_2^0, \dots, x_n^0)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x_1^0, x_2^0, \dots, x_n^0)}{\partial x_i \partial x_j}$$

or, if you prefer,

$$f_{ij}(x_1^0, x_2^0, \dots, x_n^0) = f_{ji}(x_1^0, x_2^0, \dots, x_n^0).$$

- This result is known as Young's theorem.
- The result is trivially true when $i = j$. This is the case of second-order own partial derivatives.
- The result is also true, and much more interesting, when $i \neq j$. This is the case of second-order cross partial derivatives.

Young's Theorem Example 1 Part 1

- Recall the example in which

$$f(x, y) = x^3y + e^{xy^2} = x^3y + \exp(xy^2).$$

- We have already found that

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2y + y^2 \exp\{xy^2\}$$

and

$$f_y(x, y) = \frac{\partial f}{\partial y} = x^3 + 2xy \exp\{xy^2\}.$$

Young's Theorem Example 1 Part 2

- This means that

$$\begin{aligned} & f_{xx}(x, y) \\ = & \frac{\partial f_x(x, y)}{\partial x} \\ = & \frac{\partial (3x^2y + y^2 \exp\{xy^2\})}{\partial x} \\ = & (2)(3x^{2-1}y) + y^2 \exp\{xy^2\}(y^2) \\ = & 6xy + y^4 \exp\{xy^2\} \end{aligned}$$

Young's Theorem Example 1 Part 3

- It also means that

$$\begin{aligned} & f_{xy}(x, y) \\ = & \frac{\partial f_x(x, y)}{\partial y} \\ = & \frac{\partial (3x^2y + y^2 \exp\{xy^2\})}{\partial y} \\ = & 3x^2 + \{(2y)(\exp\{xy^2\}) \\ & + (\exp\{xy^2\})(2xy)(y^2)\} \\ = & 3x^2 + 2y \exp\{xy^2\} + 2xy^3 \exp\{xy^2\}. \end{aligned}$$

Young's Theorem Example 1 Part 4

- It also means that

$$\begin{aligned} & f_{yx}(x, y) \\ = & \frac{\partial f_y(x, y)}{\partial x} \\ = & \frac{\partial (x^3 + 2xy \exp \{xy^2\})}{\partial x} \\ = & 3x^2 + \{(2y) (\exp \{xy^2\}) \\ & + (y^2 \exp \{xy^2\}) (2xy)\} \\ = & 3x^2 + 2y \exp \{xy^2\} + 2xy^3 \exp \{xy^2\} \end{aligned}$$

Young's Theorem Example 1 Part 5

- It also means that

$$\begin{aligned} & f_{yy}(x, y) \\ = & \frac{\partial f_y(x, y)}{\partial y} \\ = & \frac{\partial (x^3 + 2xy \exp \{xy^2\})}{\partial y} \\ = & 0 + \{(2x) (\exp \{xy^2\}) + (2xy \exp \{xy^2\}) (2y)\} \\ = & 2x \exp \{xy^2\} + 4x^2 y^2 \exp \{xy^2\}. \end{aligned}$$

- Note that $f_{xy}(x, y) = f_{yx}(x, y)$ in this example.

Young's Theorem Example 2 Part 1

- Young's theorem is only guaranteed to hold when the underlying function f is at least twice continuously differentiable in the neighbourhood of the point of interest.
- An example in which Young's theorem cannot be applied because the underlying function is not twice continuously differentiable at the point of interest is provided by Spiegel (1981a, p. 126, Worked Problem 43). This example is presented below.
- Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} 0 & \text{if } (x, y) = (0, 0); \\ xy \left(\frac{x^2 - y^2}{x^2 + y^2} \right) & \text{if } (x, y) \neq (0, 0). \end{cases}$$

Young's Theorem Example 2 Part 2

- If $(x, y) = (0, 0)$, then the first-order partial derivatives are

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

and

$$f_y(0, 0) = \lim_{k \rightarrow 0} \frac{f(0, k) - f(0, 0)}{k} = \lim_{k \rightarrow 0} \frac{0}{k} = 0.$$

- If $(x, y) \neq (0, 0)$, then

$$f_x(x, y) = \frac{\partial}{\partial x} \left\{ xy \left(\frac{x^2 - y^2}{x^2 + y^2} \right) \right\} = xy \left(\frac{4xy^2}{(x^2 + y^2)^2} \right) + y \left(\frac{x^2 - y^2}{x^2 + y^2} \right)$$

and

$$f_y(x, y) = \frac{\partial}{\partial y} \left\{ xy \left(\frac{x^2 - y^2}{x^2 + y^2} \right) \right\} = xy \left(\frac{-4x^2y}{(x^2 + y^2)^2} \right) + x \left(\frac{x^2 - y^2}{x^2 + y^2} \right).$$

Young's Theorem Example 2 Part 3

- If $(x, y) = (0, 0)$, then the second-order partial derivatives are

$$f_{xx}(0, 0) = \lim_{h \rightarrow 0} \frac{f_x(h, 0) - f_x(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0,$$

$$f_{yy}(0, 0) = \lim_{k \rightarrow 0} \frac{f_y(0, k) - f_y(0, 0)}{k} = \lim_{k \rightarrow 0} \frac{0}{k} = 0,$$

$$f_{xy}(0, 0) = \lim_{k \rightarrow 0} \frac{f_x(0, k) - f_x(0, 0)}{k} = \lim_{k \rightarrow 0} \frac{-k}{k} = -1$$

and

$$f_{yx}(0, 0) = \lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1.$$

- Note that $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

Total Differentials Part 1

- Consider a function $f : S \rightarrow \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is an open set. This function can be written as $f(x_1, x_2, \dots, x_n)$. Suppose that this function is at least twice continuously differentiable with respect to all of its arguments.
- Suppose that we want to consider the impact of a small change in each of the independent variables on the value of the function.
- Imagine that the initial values of the independent variables are

$$x = (x_1, x_2, \dots, x_n).$$

- Following the change, the new values of the independent variables are

$$x + dx = (x_1 + dx_1, x_2 + dx_2, \dots, x_n + dx_n).$$

Total Differentials Part 2

- Note that the vector of changes in the independent variables is

$$\begin{aligned}dx &= x + dx - x \\&= (x_1 + dx_1, x_2 + dx_2, \dots, x_n + dx_n) \\&\quad - (x_1, x_2, \dots, x_n) \\&= (dx_1, dx_2, \dots, dx_n).\end{aligned}$$

- The actual change in the value of the function that is induced by these changes in the independent variables is

$$df = f(x + dx) - f(x).$$

- When the function is non-linear, this actual change can be rather complicated to calculate explicitly.

Total Differentials Part 3

- As such, we will often employ a first-order (or linear) differential approximation for df .
- This approximation is known as the (first-order) total differential of the function.
- It is given by

$$df \approx \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right) dx_i.$$

- If the function is twice continuously differentiable in all of its arguments and the change in each of the independent variables is sufficiently small, then this approximation for df will be reasonably accurate.

The Implicit Function Theorem Part 1

- This is Theorem 15.2 in Simon and Blume (1994, p. 341).
- Suppose that $G(x_1, x_2, \dots, x_n, y)$ is a function that is at least once continuously differentiable around the point $(x_1^*, x_2^*, \dots, x_n^*, y^*)$.
- Suppose also that

$$G(x_1^*, x_2^*, \dots, x_n^*, y^*) = c \text{ for some constant } c$$

and that

$$\frac{\partial G}{\partial y}(x_1^*, x_2^*, \dots, x_n^*, y^*) \neq 0.$$

- Continued on the next slide.

The Implicit Function Theorem Part 2

- Continued from the previous slide.
- Under these conditions, there is a function $y(x_1, x_2, \dots, x_n)$ defined on some open ball B around the point $(x_1^*, x_2^*, \dots, x_n^*)$ that is at least once continuously differentiable such that:

$$(a) \quad G(x_1, x_2, \dots, x_n, y(x_1, x_2, \dots, x_n)) = c$$

$$\text{for all } (x_1, x_2, \dots, x_n) \in B,$$

$$(b) \quad y^* = y(x_1^*, x_2^*, \dots, x_n^*); \text{ and}$$

$$(c) \quad \frac{\partial y}{\partial x_i}(x_1^*, x_2^*, \dots, x_n^*) = - \left(\frac{\frac{\partial G}{\partial x_i}(x_1^*, x_2^*, \dots, x_n^*, y^*)}{\frac{\partial G}{\partial y}(x_1^*, x_2^*, \dots, x_n^*, y^*)} \right)$$

$$\text{for each } i \in \{1, 2, \dots, n\}.$$

Derivatives of Implicit Functions Part 1

- Suppose that we know that some variable y is a function of n other variables (x_1, x_2, \dots, x_n) .
- Sometimes it is not easy to explicitly characterise this function in the form

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

- In some cases, we might only be able to characterise the function implicitly, along the lines of a relationship of the form

$$g(y, x_1, x_2, \dots, x_n) = b,$$

where $b \in \mathbb{R}$ is some fixed constant.

Derivatives of Implicit Functions Part 2

- Suppose that we want to obtain the partial derivative of y with respect to x_k in such a case. How would we do this?
- Since $b \in \mathbb{R}$ is some fixed constant, we must have $dg = 0$.
- Note that the total differential for g is

$$dg = \left(\frac{\partial g}{\partial y} \right) dy + \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i} \right) dx_i.$$

- Thus we must have

$$\left(\frac{\partial g}{\partial y} \right) dy + \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i} \right) dx_i = 0.$$

Derivatives of Implicit Functions Part 3

- Suppose that the only variables that are allowed to change are x_k and y . This means that $dx_k \neq 0$ and $dy \neq 0$. It also means that $dx_i = 0$ for all $i \neq k$. Substituting these values into the above equation yields

$$\left(\frac{\partial g}{\partial y}\right) dy + \left(\frac{\partial g}{\partial x_k}\right) dx_k = 0.$$

This can be rearranged to obtain

$$\left.\frac{dy}{dx_k}\right|_{dx_i=0 \text{ for all } i \neq k} = \frac{-\left(\frac{\partial g}{\partial x_k}\right)}{\left(\frac{\partial g}{\partial y}\right)}.$$

- Since we are holding $dx_i = 0$ for all $i \neq k$, this is really a partial derivative. Thus we have

$$\frac{\partial y}{\partial x_k} = \frac{-\left(\frac{\partial g}{\partial x_k}\right)}{\left(\frac{\partial g}{\partial y}\right)}.$$

Some Applications of Implicit Functions

- Some potential applications of the implicit function theorem include the following.
 - The slope of an indifference curve.
 - The slope of an iso-expenditure ("budget") line.
 - The slope of an isoquant.
 - The slope of an isocost.
- Before discussing these examples, we will consider the related concepts of level sets, upper contour sets and lower contour sets.

Level Sets and Contour Sets Part 1

- Consider a function $f : X \rightarrow Y$, where $X \subseteq \mathbb{R}^L$ and $Y \subseteq \mathbb{R}$.
- A level set for the function f defined with respect to a point in the domain $\hat{x} \in X$ is defined to be

$$f_{\hat{x}}^0 = \{x \in X : f(x) = f(\hat{x})\}.$$

- A weak upper contour set for the function f defined with respect to a point in the domain $\hat{x} \in X$ is defined to be

$$f_{\hat{x}}^+ = \{x \in X : f(x) \geq f(\hat{x})\}.$$

Level Sets and Contour Sets Part 2

- A strong upper contour set for the function f defined with respect to a point in the domain $\hat{x} \in X$ is defined to be

$$f_{\hat{x}}^{++} = \{x \in X : f(x) > f(\hat{x})\}.$$

- A weak lower contour set for the function f defined with respect to a point in the domain $\hat{x} \in X$ is defined to be

$$f_{\hat{x}}^{-} = \{x \in X : f(x) \leq f(\hat{x})\}.$$

- A strong lower contour set for the function f defined with respect to a point in the domain $\hat{x} \in X$ is defined to be

$$f_{\hat{x}}^{--} = \{x \in X : f(x) < f(\hat{x})\}.$$

The Slope of an Indifference Curve Part 1

- Suppose that $U : \mathbb{R}_+^L \rightarrow \mathbb{R}$ is a utility function.
- The level sets for a utility function are known as indifference curves. The indifference curve that passes through consumption bundle $\hat{x} \in \mathbb{R}_+^L$ is defined to be

$$U_{\hat{x}}^0 = \left\{ x \in \mathbb{R}_+^L : U(x) = U(\hat{x}) \right\}.$$

- The (weak) upper contour sets for a utility function could be called consumption requirement sets. The consumption requirement set that is associated with the reference consumption bundle $\hat{x} \in \mathbb{R}_+^L$ is defined to be

$$U_{\hat{x}}^+ = \left\{ x \in \mathbb{R}_+^L : U(x) \geq U(\hat{x}) \right\}.$$

The Slope of an Indifference Curve Part 2

- Note that we could also define the level set and the (weak) upper contour set for a utility function in terms of a reference utility level, \bar{U} . I will leave the modification of the definitions for the case in which a reference utility level is employed as an exercise for you to attempt.
- Note that any point $x \in \mathbb{R}_+^L$ that belongs to an indifference curve must yield the same level of utility.
- This means that the change in utility around any indifference curve is zero.
- In other words, $dU = 0$ along an indifference curve.

The Slope of an Indifference Curve Part 3

- Suppose that $L = 2$. We have

$$dU = \left(\frac{\partial U}{\partial x_1} \right) dx_1 + \left(\frac{\partial U}{\partial x_2} \right) dx_2.$$

- Along any indifference curve, we must have

$$dU = \left(\frac{\partial U}{\partial x_1} \right) dx_1 + \left(\frac{\partial U}{\partial x_2} \right) dx_2 = 0.$$

- This can be rearranged to obtain

$$\left. \frac{dx_2}{dx_1} \right|_{dU=0} = \frac{- \left(\frac{\partial U}{\partial x_1} \right)}{\left(\frac{\partial U}{\partial x_2} \right)} = -MRS_{12}(x),$$

which is the formula for the slope of an indifference curve when $U : \mathbb{R}_+^2 \rightarrow \mathbb{R}$.

The Slope of a Budget Line Part 1

- A consumer's expenditure is given by the linear function $E : \mathbb{R}_+^L \rightarrow \mathbb{R}$ that is defined by

$$E(x) = p^T x = \sum_{l=1}^L p_l x_l,$$

where $p \in \mathbb{R}_{++}^L$, so that $p_l > 0$ for all $l \in \{1, 2, \dots, L\}$.

- We will refer to this as the budget function, because the term "expenditure function" has a specific meaning in microeconomics that differs somewhat from this function.

The Slope of a Budget Line Part 2

- The level sets for a budget function are known budget lines (or, if you prefer, budget hyper-planes).
- The budget line that passes through consumption bundle $\omega \in \mathbb{R}_+^L$ is defined to be

$$B_{\omega}^0(p, \omega) = \left\{ x \in \mathbb{R}_+^L : E(x) = E(\omega) \right\}.$$

In this case, we might think of $\omega \in \mathbb{R}_+^L$ as an endowment bundle.

- The budget line that is consistent with an exogenous money income equal to $M \in \mathbb{R}_+$ is defined to be

$$B_M^0(p, M) = \left\{ x \in \mathbb{R}_+^L : E(x) = M \right\}.$$

The Slope of a Budget Line Part 3

- The (weak) lower contour sets for a budget function are known budget sets. The budget set that is associated with the endowment bundle $\omega \in \mathbb{R}_+^L$ is defined to be

$$B_{\omega}^{-}(p, \omega) = \left\{ x \in \mathbb{R}_+^L : E(x) \leq E(\omega) \right\}.$$

- The (weak) lower contour sets for a budget function are known budget sets. The budget set that is associated with an exogenous money income equal to $M \in \mathbb{R}_+$ is defined to be

$$B_M^{-}(p, M) = \left\{ x \in \mathbb{R}_+^L : E(x) \leq M \right\}.$$

The Slope of a Budget Line Part 4

- Note that any point $x \in \mathbb{R}_+^L$ that belongs to a budget line must require the same level of expenditure.
- This means that the change in expenditure along any budget line is zero.
- In other words, $dE = 0$ along a budget line.
- Suppose that $L = 2$. We have

$$\begin{aligned}dE &= \left(\frac{\partial E}{\partial x_1} \right) dx_1 + \left(\frac{\partial E}{\partial x_2} \right) dx_2 \\&= \left(\frac{\partial (p_1 x_1 + p_2 x_2)}{\partial x_1} \right) dx_1 \\&\quad + \left(\frac{\partial (p_1 x_1 + p_2 x_2)}{\partial x_2} \right) dx_2 \\&= p_1 dx_1 + p_2 dx_2.\end{aligned}$$

The Slope of a Budget Line Part 5

- Along any budget line, we must have

$$dE = p_1 dx_1 + p_2 dx_2 = 0.$$

- This can be rearranged to obtain

$$\left. \frac{dx_2}{dx_1} \right|_{dE=0} = \frac{-p_1}{p_2},$$

which is the formula for the slope of a budget line when the consumption set is given by \mathbb{R}_+^2 .

The Slope of an Isoquant Part 1

- Suppose that $F : \mathbb{R}_+^K \rightarrow \mathbb{R}_+$ is a single-product production function.
- The level sets for a production function are known as isoquants. The isoquant that is consistent with $y \in \mathbb{R}_+$ units of output being produced is defined to be

$$F_y^0 = \left\{ x \in \mathbb{R}_+^K : F(x) = y \right\}.$$

- The (weak) upper contour sets for a production function are known as input requirement sets. The input requirement set that is associated with $y \in \mathbb{R}_+$ units of output being produced is defined to be

$$F_y^+ = \left\{ x \in \mathbb{R}_+^K : F(x) \geq y \right\}.$$

The Slope of an Isoquant Part 2

- Note that we could also define the level set and the (weak) upper contour set for a production function in terms of a reference input bundle, \hat{x} . I will leave the modification of the definitions for the case in which a reference input bundle is employed as an exercise for you to attempt.
- Note that any point $x \in \mathbb{R}_+^K$ that belongs to an isoquant must yield the same level of output.
- This means that the change in output around any isoquant is zero.
- In other words, $dF = 0$ along an isoquant.

The Slope of an Isoquant Part 3

- Suppose that $K = 2$. We have

$$dF = \left(\frac{\partial F}{\partial x_1} \right) dx_1 + \left(\frac{\partial F}{\partial x_2} \right) dx_2.$$

- Along any isoquant, we must have

$$dF = \left(\frac{\partial F}{\partial x_1} \right) dx_1 + \left(\frac{\partial F}{\partial x_2} \right) dx_2 = 0.$$

- This can be rearranged to obtain

$$\left. \frac{dx_2}{dx_1} \right|_{dF=0} = \frac{- \left(\frac{\partial F}{\partial x_1} \right)}{\left(\frac{\partial F}{\partial x_2} \right)} = -MRTS_{12}(x),$$

which is the formula for the slope of an isoquant when $F : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$.

The Slope of an Isocost Part 1

- A firm's expenditure on inputs is given by the linear function $E : \mathbb{R}_+^K \rightarrow \mathbb{R}$ that is defined by

$$E(x) = w^T x = \sum_{k=1}^K w_k x_k,$$

where $w \in \mathbb{R}_{++}^K$ is the input price vector.

- Note that $w \in \mathbb{R}_{++}^K$ means that $w_k > 0$ for all $k \in \{1, 2, \dots, K\}$.
- We will refer to this as the budget function, because the term "cost function" has a specific meaning in microeconomics that differs somewhat from this function.

The Slope of an Isocost Part 2

- The level sets for a firm's budget function are known as isocosts.
- The isocost line that passes through input bundle $\hat{x} \in \mathbb{R}_+^K$ is defined to be

$$B_{\hat{x}}^0(w, \hat{x}) = \left\{ x \in \mathbb{R}_+^K : E(x) = E(\hat{x}) \right\}.$$

In this case, we might think of $\omega \in \mathbb{R}_+^L$ as an endowment bundle.

- The isocost line that is consistent with a particular level of expenditure on inputs equal to $C \in \mathbb{R}_+$ is defined to be

$$B_C^0(p, M) = \left\{ x \in \mathbb{R}_+^K : E(x) = C \right\}.$$

The Slope of an Isocost Part 3

- Note that any point $x \in \mathbb{R}_+^K$ that belongs to an isocost line must involve the same level of expenditure on inputs.
- This means that the change in expenditure on inputs along any isocost line is zero.
- In other words, $dE = 0$ along an isocost line.
- Suppose that $K = 2$. We have

$$\begin{aligned}dE &= \left(\frac{\partial E}{\partial x_1} \right) dx_1 + \left(\frac{\partial E}{\partial x_2} \right) dx_2 \\&= \left(\frac{\partial (w_1 x_1 + w_2 x_2)}{\partial x_1} \right) dx_1 \\&\quad + \left(\frac{\partial (w_1 x_1 + w_2 x_2)}{\partial x_2} \right) dx_2 \\&= w_1 dx_1 + w_2 dx_2.\end{aligned}$$

The Slope of an Isocost Part 4

- Along any isocost line, we must have

$$dE = w_1 dx_1 + w_2 dx_2 = 0.$$

- This can be rearranged to obtain

$$\left. \frac{dx_2}{dx_1} \right|_{dE=0} = \frac{-w_1}{w_2},$$

which is the formula for the slope of an isocost line for a firm with a production technology that involves two inputs.

The Inverse Function Theorem

- This is Theorem 15.9 in Simon and Blume (1994, p. 367).
- Consider a function of the form $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$.
- Suppose that:
 - (a) $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is at least once continuously differentiable;
 - (b) $f(x^*) = y^*$; and
 - (c) The Jacobian matrix $D_x f$ is non-singular at the point x^* .
- If this is the case, then:
 - (i) There exists an open ball $B_r(x^*)$ around the point x^* , and an open set V around the point y^* , such that $f : B_r(x^*) \longrightarrow V$ is both one-to-one and onto;
 - (ii) The inverse map $f^{-1} : V \longrightarrow B_r(x^*)$ is at least once continuously differentiable; and
 - (iii) The derivative of the inverse function at the point x^* is equal to the inverse of the derivative function at the point x^* (that is $(D_x f^{-1})(f(x^*)) = (D_x f(x^*))^{-1}$).

An Elasticity Application of the Inverse Function Theorem

Part 1

- Consider a function of the form $f : S \longrightarrow \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is an open set.
- Suppose that this function is at least once continuously differentiable in all of its arguments.
- The elasticity of f with respect to the variable x_i at the point $(x, f(x))$ is given by the formula

$$\varepsilon_i^f = \left(\frac{x_i}{f(x)} \right) \left(\frac{\partial f}{\partial x_i} \right).$$

- Another formula for the elasticity of f with respect to the variable x_i at the point $(x, f(x))$ that is sometimes useful is

$$\varepsilon_i^f = \frac{\partial \ln(f(x))}{\partial \ln(x_i)}.$$

An Elasticity Application of the Inverse Function Theorem

Part 2

- We will use the chain rule of differentiation and the inverse function theorem to show that these two formulas are equivalent.
- Note that

$$\frac{\partial \ln(f(x))}{\partial \ln(x_i)} = \left(\frac{\partial \ln(f(x))}{\partial f(x)} \right) \left(\frac{\partial f(x)}{\partial x_i} \right) \left(\frac{\partial x_i}{\partial \ln(x_i)} \right)$$

from the chain rule of differentiation.

- We know from the inverse function theorem that

$$\left(\frac{\partial x_i}{\partial \ln(x_i)} \right) = \frac{1}{\left(\frac{\partial \ln(x_i)}{\partial x_i} \right)}.$$

An Elasticity Application of the Inverse Function Theorem

Part 3

- Thus we have

$$\frac{\partial \ln(f(x))}{\partial \ln(x_i)} = \left(\frac{\partial \ln(f(x))}{\partial f(x)} \right) \left(\frac{\partial f(x)}{\partial x_i} \right) \left(\frac{1}{\left(\frac{\partial \ln(x_i)}{\partial x_i} \right)} \right)$$

- Note that $\frac{\partial \ln(f(x))}{\partial f(x)} = \frac{1}{f(x)}$ and $\frac{\partial \ln(x_i)}{\partial x_i} = \frac{1}{x_i}$.
- This means that

$$\frac{\partial \ln(f(x))}{\partial \ln(x_i)} = \left(\frac{1}{f(x)} \right) \left(\frac{\partial f(x)}{\partial x_i} \right) \left(\frac{1}{\frac{1}{x_i}} \right) = \left(\frac{1}{f(x)} \right) \left(\frac{\partial f(x)}{\partial x_i} \right) (x_i).$$

- This can be rearranged to obtain

$$\frac{\partial \ln(f(x))}{\partial \ln(x_i)} = \left(\frac{x_i}{f(x)} \right) \left(\frac{\partial f}{\partial x_i} \right).$$

Homogeneous Functions Part 1

- Consider a function $f : S \longrightarrow \mathbb{R}$, where $S \subseteq \mathbb{R}^n$.
- This function can be written as $f(x_1, x_2, \dots, x_n)$.
- Let $\lambda > 0$ be a positive real number.
- The function $f(x_1, x_2, \dots, x_n)$ is said to be homogeneous of degree r if

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^r f(x_1, x_2, \dots, x_n)$$

for all $(x_1, x_2, \dots, x_n) \in S$ and all $\lambda > 0$.

- A function that is homogeneous of degree one is said to be linearly homogeneous.

Homogeneous Functions Part 2

- Homogeneity is a cardinal property, not an ordinal one.
- This means that homogeneity will not necessarily be preserved under a strictly increasing transformation.
- An ordinal property that is somewhat similar to homogeneity is that of homotheticity.
- A discussion of homotheticity can be found in Simon and Blume (1994, pp. 483 and 500-504).

Euler's Theorem Part 1

- Suppose that the function $f(x_1, x_2, \dots, x_n)$ is homogeneous of degree r .
- In this case we have

$$\sum_{i=1}^n x_i \left(\frac{\partial f}{\partial x_i} \right) = r f(x_1, x_2, \dots, x_n).$$

- This result is known as Euler's theorem.

Euler's Theorem Part 2

- It is straightforward to explicitly derive this result from the definition of homogeneity of degree r . The following proof is based on the one in Haeussler and Paul (1987, p. 723).
- Since the function $f(x_1, x_2, \dots, x_n)$ is homogeneous of degree r , we know that

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^r f(x_1, x_2, \dots, x_n)$$

for any choice of $\lambda > 0$.

Euler's Theorem Part 3

- Consider the left-hand side of this definition first.
- Let $z_i = \lambda x_i$ for all $i \in \{1, 2, \dots, n\}$.
- Totally differentiating

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = f(z_1, z_2, \dots, z_n)$$

yields

$$df = \sum_{i=1}^n \left(\frac{\partial f}{\partial z_i} \right) dz_i.$$

- Dividing both sides of this total differential by $d\lambda \neq 0$ yields

$$\frac{df}{d\lambda} = \sum_{i=1}^n \left(\frac{\partial f}{\partial z_i} \right) \left(\frac{dz_i}{d\lambda} \right).$$

Euler's Theorem Part 4

- Suppose that we hold all of the x_i variables constant and only allow λ to vary. This means that $dx_i = 0$ for all $i \in \{1, 2, \dots, n\}$ and $d\lambda \neq 0$. This yields

$$\left. \frac{df}{d\lambda} \right|_{dx_i=0 \text{ for all } i} = \sum_{i=1}^n \left(\frac{\partial f}{\partial z_i} \right) \left(\left. \frac{dz_i}{d\lambda} \right|_{dx_i=0 \text{ for all } i} \right).$$

- Note that

$$\left. \frac{df}{d\lambda} \right|_{dx_i=0 \text{ for all } i} = \frac{\partial f}{\partial \lambda}$$

and

$$\left. \frac{dz_i}{d\lambda} \right|_{dx_i=0 \text{ for all } i} = \frac{\partial z_i}{\partial \lambda}.$$

Euler's Theorem Part 5

- Thus we have

$$\frac{\partial f}{\partial \lambda} = \sum_{i=1}^n \left(\frac{\partial f}{\partial z_i} \right) \left(\frac{\partial z_i}{\partial \lambda} \right).$$

- Note that

$$\frac{\partial z_i}{\partial \lambda} = \frac{\partial (\lambda x_i)}{\partial \lambda} = x_i$$

for all $i \in \{1, 2, \dots, n\}$.

- Thus we know that the partial derivative of

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n)$$

with respect to λ is

$$\frac{\partial f(\lambda x_1, \lambda x_2, \dots, \lambda x_n)}{\partial \lambda} = \sum_{i=1}^n \left(\frac{\partial f}{\partial z_i} \right) x_i,$$

where $z_i = \lambda x_i$ for all $i \in \{1, 2, \dots, n\}$.

Euler's Theorem Part 6

- Now consider the right hand-side of the following definition for a function that is homogeneous of degree r :

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^r f(x_1, x_2, \dots, x_n)$$

for any choice of $\lambda > 0$.

- Note that λ does not appear in the term (x_1, x_2, \dots, x_n) . As such, the partial derivative of $\lambda^r f(x_1, x_2, \dots, x_n)$ with respect to λ is

$$\frac{\partial \lambda^r f(x_1, x_2, \dots, x_n)}{\partial \lambda} = r \lambda^{r-1} f(x_1, x_2, \dots, x_n).$$

Euler's Theorem Part 7

- Thus we have

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^r f(x_1, x_2, \dots, x_n)$$

$$\iff$$

$$\frac{\partial f(\lambda x_1, \lambda x_2, \dots, \lambda x_n)}{\partial \lambda} = \frac{\partial \lambda^r f(x_1, x_2, \dots, x_n)}{\partial \lambda}$$

$$\iff$$

$$\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right) x_i = r \lambda^{r-1} f(x_1, x_2, \dots, x_n).$$

Euler's Theorem Part 8

- Note that when $\lambda = 1$ we have $z_i = 1x_i = x_i$ for all $i \in \{1, 2, \dots, n\}$. This means that

$$\left(\frac{\partial f}{\partial z_i} \right) = \left(\frac{\partial f}{\partial x_i} \right)$$

for all $i \in \{1, 2, \dots, n\}$ when $\lambda = 1$.

- Note also that when $\lambda = 1$ we have

$$\lambda^{r-1} = 1^{r-1} = 1.$$

- Now suppose that we evaluate both sides of this last equation at the point $\lambda = 1$. This yields

$$\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right) x_i = rf(x_1, x_2, \dots, x_n),$$

which is the result from Euler's theorem.

- Some applications of homogeneity include the following.
 - Returns to scale for production technologies in general.
 - Returns to scale for Cobb-Douglas production technologies.
 - Euler aggregation for a Marshallian demand function.
 - Product exhaustion under perfect competition and constant returns to scale.

Returns to Scale Part 1

- Suppose that an n input and one output production technology can be represented by a production function of the form

$$Q = f(L_1, L_2, \dots, L_n).$$

- What happens if we increase all of the inputs by the same proportion?
- Specifically, suppose that we move to an input bundle of the form $(\lambda L_1, \lambda L_2, \dots, \lambda L_n)$, where $\lambda > 0$.

Returns to Scale Part 2

- The production technology is said to display decreasing returns to scale if this increase in all of the inputs induces a less than proportionate change in the amount of output that can be produced.
- In other words, production technology is said to display decreasing returns to scale if

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) < \lambda Q,$$

which is equivalent to

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) < \lambda f(L_1, L_2, \dots, L_n).$$

- Thus the production function for a production technology that displays decreasing returns to scale is either homogeneous of degree less than one or it is not homogeneous.

Returns to Scale Part 3

- A simple replication argument would suggest that decreasing returns to scale do not make any sense.
- Studies that find decreasing returns to scale are probably really finding diminishing marginal products.
- There is probably some unobserved (or imprecisely observed) input that is implicitly being held fixed or, at least, not varying in the same proportion as all of the other inputs.

Returns to Scale Part 4

- The production technology is said to display constant returns to scale if this increase in all of the inputs induces a proportionate change in the amount of output that can be produced.
- In other words, production technology is said to display constant returns to scale if

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) = \lambda Q,$$

which is equivalent to

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) = \lambda f(L_1, L_2, \dots, L_n).$$

- Thus the production function for a production technology that displays constant returns to scale will be homogeneous of degree one. (Homogeneity of degree one is sometimes referred to as linear homogeneity.)

Returns to Scale Part 5

- Constant returns to scale production technologies are employed in the standard versions of the neoclassical model of economic growth.
- This model is often referred to as the Solow-Swan model of economic growth, because one of the seminal papers on this model was written by Robert Solow and the other was written by Trevor Swan. Note that the late Trevor Swan was an Australian economist who worked at the Australian National University.
- The relevant references are:
 - Solow, RM (1956), "A contribution to the theory of economic growth", *The Quarterly Journal of Economics* 70(1), February, pp. 65-94; and
 - Swan, TW (1956), "Economic growth and capital accumulation", *The Economic Record* 32, November, pp. 334-361.

Returns to Scale Part 6

- The production technology is said to display increasing returns to scale if this increase in all of the inputs induces a more than proportionate change in the amount of output that can be produced.
- In other words, production technology is said to display increasing returns to scale if

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) > \lambda Q,$$

which is equivalent to

$$f(\lambda L_1, \lambda L_2, \dots, \lambda L_n) > \lambda f(L_1, L_2, \dots, L_n).$$

- Thus the production function for a production technology that displays increasing returns to scale is either homogeneous of degree more than one or it is not homogeneous.

Returns to Scale Part 7

- The returns to scale displayed by some production technologies might vary with input vector.
- In such cases, the production technology will not necessarily display either constant returns to scale or increasing returns to scale over the entire set of possible values for the input vector.

Cobb-Douglas Production Functions Part 1

- Suppose that an two input and one output production technology can be represented by a production function of the form

$$Q = f(L, K) = AL^\alpha K^\beta.$$

- This type of production function is known as a Cobb-Douglas production function.

Cobb-Douglas Production Functions Part 2

- Suppose that $\lambda > 0$. Note that

$$\begin{aligned} f(\lambda L, \lambda K) &= A(\lambda L)^\alpha (\lambda K)^\beta \\ &= A\lambda^\alpha L^\alpha \lambda^\beta K^\beta \\ &= \lambda^{\alpha+\beta} A L^\alpha K^\beta \\ &= \lambda^{\alpha+\beta} f(L, K). \end{aligned}$$

- Thus we know that the Cobb-Douglas production function is homogeneous of degree $(\alpha + \beta)$.

Cobb-Douglas Production Functions Part 3

- If $(\alpha + \beta) < 1$, then the Cobb-Douglas production function displays decreasing returns to scale;
- If $(\alpha + \beta) = 1$, then the Cobb-Douglas production function displays constant returns to scale; and
- If $(\alpha + \beta) > 1$, then the Cobb-Douglas production function displays increasing returns to scale.

Euler Aggregation Part 1

- Suppose that a consumer's Marshallian demand for commodity k is given by the function

$$\begin{aligned}Q_k &= x_k(p_1, p_2, \dots, p_n, y) \\ &= x_k(p, y).\end{aligned}$$

- It can be shown that under certain circumstance, Marshallian demand functions are homogeneous of degree zero.

Euler Aggregation Part 2

- It can be shown that under certain circumstance, Marshallian demand functions are homogeneous of degree zero.
- We will not provide a formal proof of this proposition here. However, the result is fairly intuitive. If all prices and income increase by the same proportion, then the budget constraint for the consumer does not change. If his or her preferences remain the same, then the fact that the budget constraint has not changed would suggest that the choice problem that faces the consumer has not changed in any real sense. As such, the consumer's optimal consumption bundle should not have changed either.

Euler Aggregation Part 3

- If Marshallian demand functions are homogeneous of degree zero, then the Marshallian demand for commodity k will be homogeneous of degree zero.
- In such circumstances, we know from Euler's theorem that

$$\begin{aligned} & y \left(\frac{\partial x_k(p, y)}{\partial y} \right) + \sum_{i=1}^n p_i \left(\frac{\partial x_k(p, y)}{\partial p_i} \right) \\ &= (0) x_k(p_1, p_2, \dots, p_n, y). \end{aligned}$$

- This can be simplified to obtain

$$y \left(\frac{\partial x_k(p, y)}{\partial y} \right) + \sum_{i=1}^n p_i \left(\frac{\partial x_k(p, y)}{\partial p_i} \right) = 0.$$

- This can be rearranged to obtain

$$\sum_{i=1}^n p_i \left(\frac{\partial x_k(p, y)}{\partial p_i} \right) = -y \left(\frac{\partial x_k(p, y)}{\partial y} \right).$$

- Upon dividing both sides of this equation by $x_k(p, y)$, we obtain

$$\begin{aligned} & \sum_{i=1}^n \left(\frac{p_i}{x_k(p, y)} \right) \left(\frac{\partial x_k(p, y)}{\partial p_i} \right) \\ &= - \left(\frac{y}{x_k(p, y)} \right) \left(\frac{\partial x_k(p, y)}{\partial y} \right). \end{aligned}$$

Euler Aggregation Part 5

- Thus we have

$$\sum_{i=1}^n \varepsilon_i^k = -\varepsilon_y^k,$$

where

$$\varepsilon_i^k = \left(\frac{p_i}{x_k(p, y)} \right) \left(\frac{\partial x_k(p, y)}{\partial p_i} \right)$$

is the elasticity of demand for commodity k with respect to the price of good i and

$$\varepsilon_y^k = \left(\frac{y}{x_k(p, y)} \right) \left(\frac{\partial x_k(p, y)}{\partial y} \right)$$

is the income elasticity of demand for commodity k .

- As far as I am aware, this well known result does not have a name. But it could be referred to as Euler aggregation, because it makes use of Euler's theorem.

Product Exhaustion Part 1

- Consider a profit-maximising firm that has a single output (Q) and two inputs (L and K) production technology that displays constant returns to scale. This production technology can be represented by a production function of the form

$$Q = f(L, K).$$

- Suppose that this firm is a price-taker in both the output market and all input markets.
- Constant returns to scale imply that the production function is homogeneous of degree one. Thus we know from Euler's theorem that

$$\left(\frac{\partial f}{\partial L}\right) L + \left(\frac{\partial f}{\partial K}\right) K = f(L, K). \quad (1)$$

Product Exhaustion Part 2

- Recall that the first-order conditions for profit maximisation by a firm with these characteristics are

$$\left(\frac{\partial f}{\partial L}\right) = \frac{w}{p}$$

and

$$\left(\frac{\partial f}{\partial K}\right) = \frac{r}{p},$$

where p is the output price, w is the input price for labour (L) and r is the input price for capital (K).

- Upon substituting these FOCs into Equation (1), we obtain

$$\left(\frac{w}{p}\right) L + \left(\frac{r}{p}\right) K = f(L, K).$$

Product Exhaustion Part 3

- This can be rearranged to obtain

$$wL + rK = pf(L, K).$$

- In other words, the firm's revenue from sales of its output will be exactly equal to the firm's expenditure on inputs. This result, which implies zero profits for a perfectly competitive firm with a constant returns to scale production technology, is known as product exhaustion.
- (Note that profit here means economic profit, not accounting profit.)