

10. MULTIVARIABLE FUNCTIONS: PARTIAL DERIVATIVES AND THE TOTAL DIFFERENTIAL

J.P. Bastos*

EPH 1100: Analytical Tools

Reading: Hoy et al., §11.1–11.3.

- **§11.1:** First-order partial derivatives; geometric interpretation; economic applications.
- **§11.2:** Second-order partial derivatives; Young’s theorem; the gradient.
- **§11.3:** The total differential.

1. The Limits of One-Variable Thinking

Every model we covered so far relied on a single input variable: cost as a function of quantity, revenue as a function of price, and so on. In this case, a single derivative was sufficient to show how the function responded.

Real economic decisions involve several variables at once. A firm chooses how much labor to hire and how much capital to rent. A consumer allocates spending across many goods simultaneously. In each case, the outcome we care about (output, utility, profit) is a function of several inputs.

To handle this, we extend our calculus toolkit to functions of several variables. A function of several variables can change in many directions at once, so there is no single slope that describes “the” rate of change. Instead, we proceed one variable at a time. First we ask how the function responds when any given variable changes and the others are held fixed. Then, by combining those one-variable answers, we can describe what happens when several variables move together.

2. Functions of Several Variables

Definition. A **function of several variables** assigns a single real number to every combination of input values. For two variables:

$$z = f(x, y)$$

where x and y are the **independent variables** and z is the **dependent variable**. More generally, for n variables: $z = f(x_1, x_2, \dots, x_n)$.

*Assistant Professor of Economics, University of Austin (UATX). jbastos@uastin.org

In economics, multivariable functions are the rule rather than the exception. In consumer theory, we typically describe a choice between two goods (x_1, x_2) , even when “good 2” can be described as “spending on everything else.” In production theory, we model the firm as employing two inputs, capital and labor. While both can be extended for N goods $(x_1, \dots, x_n)$, and any number of production inputs, we’ll see that we can get a lot of traction with only two inputs, hence why most models make that choice.¹

2.1 Level Curves

With one variable, we visualized functions as curves in a plane. With two variables, the full picture is a three-dimensional surface. The most economically useful two-dimensional representation is the **level curve**: the set of all input combinations (x, y) that produce the same output value $\bar{z}$.

Definition. A **level curve** of $f(x, y)$ at level $\bar{z}$ is the set

$$\{(x, y) : f(x, y) = \bar{z}\}$$

In consumer theory a common level curve is the **indifference curve**, which shows the combinations of (x_1, x_2) that provide the same level of utility $\bar{U}$. In production theory, the **isoquant** (“equal quantity”) shows combinations of capital and labor that produce some level of output $\bar{Q}$, and the **isocost** curve shows the multiple combinations that have the same cost ($\bar{C}$).

Example: an isoquant map. Take the Cobb-Douglas production function $Q = K^{0.5}L^{0.5}$. The isoquant at output level $\bar{Q}$ is the set of all (K, L) satisfying $K^{0.5}L^{0.5} = \bar{Q}$, i.e.,

$$K = \frac{\bar{Q}^2}{L}.$$

Figure 1 plots this curve for three output levels, $\bar{Q} = 3, 4, 5$.

Two features of this picture matter for how an isoquant should be read. First, it is downward sloping: producing the same output with less capital requires more labor, and vice versa. Second, and more important, an isoquant is a purely **technical** relationship. It describes every input combination capable of producing $\bar{Q}$ units of output, but it does not say *which* of those combinations is cheapest, or which one the firm should actually use to maximize profit (the **optimal** combination). That depends on the wage rate and the rental rate of capital, and is the subject of constrained optimization, which we will cover later.

¹In other areas of economics, it is common to create models that have two types of representative agents, or inputs with two types of quality, such as workers with low- and high-skill, investors that are either risk-averse or risk-takers, etc..

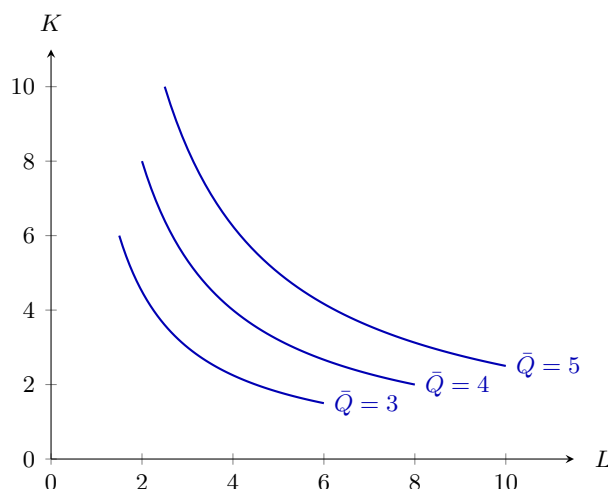


Figure 1: Isoquant Map for $Q = K^{0.5}L^{0.5}$.

Notes: Each curve is one isoquant: every (L, K) combination on it produces the same output $\bar{Q}$. Isoquants farther from the origin represent higher output. Each isoquant is downward sloping and convex to the origin: giving up capital requires progressively more labor to keep output at $\bar{Q}$.

3. Partial Derivatives

How do we differentiate a function of several variables? The key idea is to differentiate with respect to one variable at a time, treating all others as if they were constants. If z depends on (x, y) , to take the derivative with respect to x we would treat y as a constant and differentiate as usual. The computation rules are exactly identical to ordinary differentiation. The result is called a **partial derivative**.

Definition. The **partial derivative** of $f(x, y)$ with respect to x is:

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}$$

Notation. Alternative ways to write the partial derivatives of $z = f(x, y)$:

$$\frac{\partial f}{\partial x} = f_x = f_1 \qquad \frac{\partial f}{\partial y} = f_y = f_2$$

Note that we use the symbol ∂ (read “partial”) instead of d , the symbol for ordinary derivatives: $\partial f / \partial x$ is read “partial f , partial x .” The differential dy from Class 5 is the symbol for total differential. Since we assumed a single variable earlier, it could only mean one thing. Once y depends on several variables, dy alone is ambiguous. Section 6, below, defines the total differential, which resolves this.

Economic Interpretation: Marginal Products. In production theory, the partial derivatives of $F(K, L)$ are the **marginal products** of each input: the extra output obtained by adding one

more unit of that input while holding the other fixed.

$$\text{Marginal Product of Labor} = F_L = \frac{\partial F}{\partial L} \quad \text{Marginal Product of Capital} = F_K = \frac{\partial F}{\partial K}$$

Example: A firm's production function is $Q = F(K, L) = K^{0.5}L^{0.5}$. Find the marginal products of labor and capital.

MPL: Hold K constant and differentiate with respect to L :

$$F_L = K^{0.5} \cdot 0.5L^{-0.5} = \frac{1}{2}\sqrt{\frac{K}{L}}$$

MPK: Hold L constant and differentiate with respect to K :

$$F_K = 0.5K^{-0.5} \cdot L^{0.5} = \frac{1}{2}\sqrt{\frac{L}{K}}$$

Note that both F_L and F_K are > 0 for any positive quantities of K and L . The marginal product of labor, $\frac{1}{2}\sqrt{K/L}$, rises as capital increases, meaning a worker surrounded by more machines is more productive; on the other hand, each *additional* worker contributes less on the margin (although always positive). Symmetrically, capital is more valuable when there is more labor to operate it. The two inputs reinforce each other.

3.1 Evaluating a Partial Derivative at a Point

A partial derivative such as f_x is itself a function of (x, y) . It holds y fixed *symbolically* in the differentiation procedure, but the result still depends on the level of both variables. Often we want to evaluate it at one specific input combination, i.e., hold y fixed at a particular level y_0 and x at a particular level x_0 , and compute the resulting number. This is what we have been doing implicitly so far, but one can do it explicitly with proper notation.

Notation. The partial derivative of $f(x, y)$ with respect to x , evaluated at the point (x_0, y_0) , is written

$$f_x(x_0, y_0) \quad \text{or equivalently} \quad \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)}$$

Differentiate with respect to x holding y fixed, then plug in $x = x_0$ and $y = y_0$.

Example: For $F(K, L) = K^{0.5}L^{0.5}$, Section 3 found $F_L = 0.5 K^{0.5}L^{-0.5}$. To find the marginal product of labor at a specific bundle (K_0, L_0) , or holding only one of them fixed, we simply plug these values in the derivative. For instance, at $K = 4$ and $L = 1$:

$$F_L(4, 1) = F_L|_{(4,1)} = 0.5(4)^{0.5}(1)^{-0.5} = 0.5 \cdot 2 \cdot 1 = 1.00$$

4. Second-Order Partial Derivatives

Just as we took the second derivative in single-variable calculus to learn about curvature, we can differentiate the partial derivatives a second time. For $f(x, y)$, there are four second-order partial derivatives:

Second-order partial derivatives of $f(x, y)$:

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} \quad (\text{differentiate } f_x \text{ again with respect to } x)$$

$$f_{yy} = \frac{\partial^2 f}{\partial y^2} \quad (\text{differentiate } f_y \text{ again with respect to } y)$$

$$f_{xy} = \frac{\partial^2 f}{\partial y \partial x} \quad (\text{differentiate } f_x \text{ with respect to } y)$$

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y} \quad (\text{differentiate } f_y \text{ with respect to } x)$$

f_{xy} and f_{yx} are called **cross partial derivatives** (or **mixed partials**).

The cross partials almost always give the same answer, regardless of the order of differentiation. This is the content of **Young's theorem**. For all functions in this course, Young's theorem holds, so the order of differentiation never matters.

Young's Theorem. If the cross partial derivatives f_{xy} and f_{yx} are both continuous, then $f_{xy} = f_{yx}$.

Economic Interpretation: Diminishing Returns. The own second partials (F_{LL} , F_{KK}) measure diminishing or increasing marginal returns. In nearly every realistic production model, adding more of one input while holding the other fixed eventually yields smaller and smaller increases in output: **diminishing marginal returns**, which corresponds to $F_{LL} < 0$ and $F_{KK} < 0$. The intuition is straightforward: adding more machines to the factory floor barely raises output if there are not enough workers to run them. Each additional machine sits idle or under-used, so output grows more and more slowly.

The same logic explains **diminishing marginal utility**. Consider a utility function where the two goods are complements – say, $U(x, y) = x^{0.5}y^{0.5}$ for coffee (x) and cream (y), goods normally (but egregiously!) consumed together. Even though x and y are complements here (more cream raises the marginal utility of coffee, as the next subsection makes precise), coffee alone still exhibits diminishing marginal utility ($U_{xx} < 0$): giving the consumer more coffee while cream stays fixed raises satisfaction, but at a decreasing rate. Think that more coffee without more cream to go with it messes up the ratio, and thus the extra coffee is only partly useful.

4.1 Substitutes and Complements

The cross partial $f_{xy} = f_{yx}$ also has an economic interpretation. It measures how the two variables interact: how a change in y affects the marginal effect of x , or equivalently, by Young's theorem, how a change in x affects the marginal effect of y . This applies equally to production theory (inputs K and L) and consumer theory (goods x and y).

Definition. Two inputs (or goods) x and y in $f(x, y)$ are:

- **complements** if $f_{xy} > 0$: more of one raises the marginal product (or marginal utility) of the other;
- **substitutes** if $f_{xy} < 0$: more of one lowers the marginal product (or marginal utility) of the other.

By Young's theorem, this sign is unambiguous: it does not matter which variable is differentiated first.

Example (substitutes): A firm's total sales, S , depend on its spending on two advertising channels, online ads x and print ads y :

$$S(x, y) = 20x - x^2 + 20y - y^2 - 2xy.$$

The cross partial is $S_{xy} = -2 < 0$. Buying more print ads *lowers* the marginal payoff of online ads, because the two campaigns increasingly reach the same audience. Online and print advertising are substitutes.

Example (complements): Find all four second-order partial derivatives of $Q = F(K, L) = K^{0.5}L^{0.5}$.

From Section 3: $F_L = 0.5 K^{0.5} L^{-0.5}$ and $F_K = 0.5 K^{-0.5} L^{0.5}$.

Step 1: Own second partials.

$$F_{LL} = \frac{\partial}{\partial L} (0.5 K^{0.5} L^{-0.5}) = -0.25 K^{0.5} L^{-1.5} < 0$$

$$F_{KK} = \frac{\partial}{\partial K} (0.5 K^{-0.5} L^{0.5}) = -0.25 K^{-1.5} L^{0.5} < 0$$

Both are negative: diminishing marginal returns to each input. Hiring more workers, with capital held fixed, raises output at a decreasing rate.

Step 2: Cross partials.

$$F_{LK} = \frac{\partial}{\partial K} (0.5 K^{0.5} L^{-0.5}) = 0.25 K^{-0.5} L^{-0.5} = \frac{0.25}{\sqrt{KL}} > 0$$

$$F_{KL} = \frac{\partial}{\partial L} (0.5 K^{-0.5} L^{0.5}) = 0.25 K^{-0.5} L^{-0.5} = \frac{0.25}{\sqrt{KL}} > 0$$

As expected, $F_{LK} = F_{KL}$ (Young's theorem). The positive value confirms complementarity: adding capital raises the marginal product of labor, and vice versa.

5. The Gradient

When working with a function of several variables, it is useful to collect all the partial derivatives into a single object: one that records, for every coordinate direction, the payoff of moving in that direction.

Definition. The **gradient** of $f(x_1, x_2, \dots, x_n)$ is the vector of all first-order partial derivatives:

$$\nabla f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix}$$

For two variables: $\nabla f(x, y) = (f_x, f_y)$.

Think of the gradient as a map legend. Standing at a point (x_0, y_0) , there are two elementary moves available: a step in the x -direction (holding y fixed) or a step in the y -direction (holding x fixed). Each row of the gradient reports the payoff of one of those moves: $f_x(x_0, y_0)$ is how much z changes per unit step in x , and $f_y(x_0, y_0)$ is how much z changes per unit step in y . The gradient does not, by itself, say where you will end up – it is a table of exchange rates, one per direction, evaluated at your current position.

You get an actual destination once you commit to a specific move. If you step dx units in x and dy units in y , the resulting change in z is each exchange rate weighted by how far you moved in that direction: $f_x dx + f_y dy$. This is precisely the total differential of Section 6 below, and it is the same operation as a dot product, $\nabla f \cdot (dx, dy)$: the gradient supplies the per-direction rates, and the step (dx, dy) says how much of each direction to take.

In economics, the gradient of a profit function $\nabla \pi(\mathbf{x})$ reports, input by input, how much an extra unit of that input is worth to the firm; a firm choosing how to allocate a budget increase across inputs is choosing a move $(dx_1, dx_2, \dots)$, and the gradient tells it the payoff of each candidate move. The gradient of a utility function $\nabla U(\mathbf{x})$ plays the same role for a consumer deciding how to reallocate spending across goods.

Example: For $Q = K^{0.5}L^{0.5}$, find the gradient at $(K, L) = (4, 1)$.

Step 1: Find the partial derivatives (as in Section 3).

$$F_K = \frac{\partial F}{\partial K} = 0.5 K^{-0.5} L^{0.5} \qquad F_L = \frac{\partial F}{\partial L} = 0.5 K^{0.5} L^{-0.5}$$

Step 2: Evaluate each partial derivative at $(K, L) = (4, 1)$.

$$F_K(4, 1) = 0.5 (4)^{-0.5} (1)^{0.5} = 0.5 \cdot 0.5 \cdot 1 = 0.25 \qquad F_L(4, 1) = 0.5 (4)^{0.5} (1)^{-0.5} = 0.5 \cdot 2 \cdot 1 = 1.00$$

Step 3: Stack the two evaluated partials into the gradient vector.

$$\nabla F(4, 1) = \begin{pmatrix} F_K(4, 1) \\ F_L(4, 1) \end{pmatrix} = \begin{pmatrix} 0.25 \\ 1.00 \end{pmatrix}$$

5.1 Marginal Rates of Substitution

The ratio between any two components of the gradient is the **marginal rate of substitution** (MRS)² between those two inputs (or goods), which measures how much of one can be given up in exchange for one more unit of the other while keeping output (or utility) constant. For this example:

$$\text{MRS}_{KL} = \frac{F_L}{F_K} = \frac{1.00}{0.25} = 4$$

At this point, labor is four times as productive as capital per unit added. From a purely “engineering” perspective, the firm should invest more in labor: shifting one unit of budget from capital to labor produces a larger output gain at current input levels. However, from an *economic* point of view, the firm must also consider the relative prices of labor and capital to find the optimal allocation, which is the subject of constrained optimization. The general rule is that the firm should keep reallocating until the MRS equals the price ratio. In this example, the firm should substitute labor for capital unless capital is more than four times cheaper than labor, which would compensate for the lower productivity of capital. Identically, in consumer theory the consumer should keep reallocating spending until the MRS – in this case, the ratio between the marginal utilities – equals the price ratio of the two goods.

6. The Total Differential

In Class 5, we defined the differential $dy = f'(x) dx$ as the predicted change in y when x changes by a small amount dx . For a function of several variables, when every input changes simultaneously, we sum the contribution from each.

²Or sometimes referred to as the **marginal rate of technical substitution** (MRTS) in production theory.

Definition. The **total differential** of $z = f(x, y)$ is:

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = f_x dx + f_y dy$$

It approximates the total change in z when x changes by dx and y changes by dy simultaneously.

For n variables: $dz = \sum_{i=1}^n f_i dx_i = \nabla f \cdot d\mathbf{x}$.

The logic is the same as before: $f_x dx$ is the change in z due to the change dx in x alone (holding y fixed), and $f_y dy$ is the change due to dy in y alone. Adding them gives the combined effect. For small changes, this approximation is accurate because a differentiable function is nearly linear near any point.

Example: A firm produces output $Q = K^{0.5}L^{0.5}$. Currently $K = 4$ and $L = 1$. The firm increases capital by $\Delta K = 2$ and hires one more worker ($\Delta L = 1$). By how much does output increase?

Step 1: Compute the partial derivatives at $(K, L) = (4, 1)$.

$$F_K|_{(4,1)} = 0.5 (1/4)^{0.5} = 0.5 \cdot 0.5 = 0.25 \quad F_L|_{(4,1)} = 0.5 (4)^{0.5} = 0.5 \cdot 2 = 1.00$$

Step 2: Apply the total differential with $dK = 2$ and $dL = 1$.

$$dQ = F_K dK + F_L dL = 0.25(2) + 1.00(1) = 0.50 + 1.00 = 1.50$$

The total differential predicts an output increase of approximately 1.5 units.

Step 3: Compare with the exact change.

$$Q_{\text{new}} = 6^{0.5} \cdot 2^{0.5} = \sqrt{12} \approx 3.46 \quad Q_{\text{old}} = 4^{0.5} \cdot 1^{0.5} = 2.00 \quad \Delta Q \approx 1.46$$

The approximation (1.50) is close to the true change (1.46) and would improve for smaller changes.

The decomposition is informative: labor contributes 1.00 units to the output increase; capital contributes 0.50 units. Capital is being added in larger quantity (two units vs. one), but labor generates a larger output response because its marginal product at this point is four times higher.

Practice Problems

1. A firm has production function $Q = F(K, L) = K^{1/3}L^{2/3}$.

- Find the marginal product of labor F_L and the marginal product of capital F_K .
- Evaluate both marginal products at $(K, L) = (8, 27)$.
- Find the own second partials F_{LL} and F_{KK} . What is the sign of each? Interpret them economically.

- (d) Find both cross partials F_{LK} and F_{KL} and verify Young's theorem. Are they substitutes or complements?
2. A consumer has utility function $U(x, y) = x^{0.5}y^{0.5}$.
- (a) Find the **marginal utilities** U_x and U_y .
- (b) Find the cross partials U_{xy} and U_{yx} . What is the economic meaning of a positive cross partial in the context of utility?
- (c) Write the gradient ∇U at the bundle $(x, y) = (4, 9)$.
- (d) Which good has the higher marginal utility at this bundle? What does that suggest about where the consumer should shift their spending?
3. A firm's total cost is $C(w, r) = w^{2/3}r^{1/3}\bar{Q}$, where w is the wage rate, r is the rental rate of capital, and $\bar{Q} = 300$ is a fixed output target.
- (a) Find $\partial C/\partial w$ and $\partial C/\partial r$. Interpret each derivative.
- (b) Currently $w = 40$ and $r = 5$. Wages rise by $dw = 0.50$ and the rental rate rises by $dr = 1.00$. Use the total differential to approximate the change in total cost.
- (c) Which input price increase accounts for more of the cost increase? What determines this?
4. A firm's profit function is $\pi(p, a) = 10p - p^2 + 6a - a^2 + pa$, where p is its price and a is its advertising budget.
- (a) Find π_p and π_a . What do these partial derivatives represent economically? Do their signs make economic sense?
- (b) Find the cross partial π_{pa} . What does a positive cross partial between price and advertising mean? (Hint: how does a higher advertising budget change the return to raising price?)
- (c) Write the total differential $d\pi$ and interpret each term.