[MATH-17] Partial derivatives

Miguel-Angel Canela
Associate Professor, IESE Business School

Functions of several variables

I distinguish, in this and the following lectures, between n-dimensional points and vectors. I use capitals for points and boldface for vectors. So,

$$X = (x_1, x_2, \ldots, x_n), \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Remember that vectors can be identified with arrows, pointing from one point to another point, so that the difference of the coordinates of the endpoint minus the starting point gives the terms of the vector. $\mathbb{R}^n$ stands for the space of points of n coordinates, and E_n for the space of n-vectors.

The domain of a function of n variables is a subset of $\mathbb{R}^n$. The **graph** of a function f is the subset of $\mathbb{R}^{n+1}$ defined by the equation $x_{n+1} = f(x_1, \ldots, x_n)$. The graph is a curve for $n = 1$ and a surface for $n = 2$. For $n > 2$, it is a **hypersurface**.

For $n = 2$, we can use 3-D plots of surfaces as graphical representations. More appealing are the **isoquants** or contours. An isoquant is a curve whose equation has the form $f(x_1, x_2) = C$, with C constant. Thus, f takes the same value at all the points of an isoquant, and each point in the domain of f belongs to a unique isoquant.

Example 1. The isoquants of a linear function $f(x_1, x_2) = b_0 + b_1x_1 + b_2x_2$ form a set of parallel straight lines, with equation $b_1x_1 + b_2x_2 = C$. Each point of the plane belongs to one of these lines. □

Example 2. Quadratic functions are defined by polynomials of second degree. The isoquants of a quadratic function form a set of **conics**. The particular type of conic (ellipse, hyperbola, etc) depends on the coefficients of the polynomial that defines the function. □

Example 3. The general form of a **Cobb-Douglas function** is

$$f(x_1, x_2) = K\, x_1^{\alpha}\, x_2^{1-\alpha}, \quad x_1, x_2 \ge 0,$$

with $0 < \alpha < 1$ and $K > 0$. The constant K can be dropped after rescaling f. Cobb-Douglas functions are used in various applications in economics. Although the shape of the isoquants depends on the parameter α, a common trait is that they are asymptotic to both axes. Sometimes, the logarithm is preferred,

$$u(x_1, x_2) = \log f(x_1, x_2) = \alpha \log x_1 + (1 - \alpha) \log x_2,$$

since it is given by a linear expression. □

Partial derivatives

Let f be a function of two variables, and X^* a point in the domain of f. The partial derivative of f with respect to x_1 at X^* is the limit

$$\partial_1 f(X^*) = \lim_{h\to 0} \frac{f(x_1^* + h, x_2^*) - f(x_1^*, x_2^*)}{h}\,.$$

The derivative with respect to x_2 is the same, but summing h in the second coordinate. The extension to n variables is straightforward. $\partial f/\partial x_i$ is the usual notation, but I shorten it here to $\partial_i f$. Note that $\partial_i f$ is the derivative of the function of one variable obtained by fixing all the coordinates except x_i. For instance, for a function of two variables, calculating $\partial_1 f$ is equivalent to restricting f to the line $x_2 = x_2^*$ and taking the derivative.

We can calculate partial derivatives by using the rules of symbolic calculus discussed in lecture 9. Note that if the partial derivative of f with respect to x_i is positive (negative), the value of f increases (decreases) when x_i increases and all the other variables are kept constant.

The partial derivatives can be presented as the components of a vector, the **gradient** vector, usually denoted by $\nabla f(X)$. Thus,

$$\nabla f(X) = \begin{bmatrix} \partial_1 f(X) \\ \vdots \\ \partial_n f(X) \end{bmatrix}.$$

¶ When the partial derivatives are presented as a row matrix, such a matrix is called the Jacobian matrix of f.

Example 4. Let $f(x_1, x_2) = e^{x_1 x_2}$. Then

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_2\, e^{x_1 x_2} \\ x_1\, e^{x_1 x_2} \end{bmatrix}. \ \square$$

The gradient has an interesting geometric property. To keep it simple, I restrict the discussion to $n = 2$. At any point X^*, the gradient is perpendicular to the tangent to the isoquant of f that contains that point (each point belongs to a unique isoquant). This is sometimes interpreted by saying that the gradient marks the direction along which the variation of f is maximum.

How is this? Take an isoquant equation $f(x_1, x_2) = C$. This equation defines *implicitly* x_2 as a function of x_1, although finding an *explicit* expression for this function is not always easy, and sometimes even impossible. Nevertheless, an old trick, called **implicit derivative** allows to find the derivative dx_2/dx_1 without having a mathematical expression for this function. We take the derivative with respect to x_1 in the isoquant equation, assuming that x_2 is a function of x_1. Then

$$\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dx_1} = 0.$$

Then, the slope of the tangent is $-\partial_1 f(X^*)/\partial_2 f(X^*)$, and the vector

$$\begin{bmatrix} -\partial_2 f(X^*) \\ \partial_1 f(X^*) \end{bmatrix},$$

follows the direction of the tangent. This vector is orthogonal to the gradient.

The extension to $n > 2$ is not difficult, but the notation becomes cumbersome and the intuition is lost. Nevertheless, you can easily imagine that, for $n = 3$, $f(x_1, x_2, x_3) = C$ is the equation of a surface, that defines x_3 as an implicit function of x_1 and x_2. The gradient of f is orthogonal to the tangent plane at each point of the surface.

Example 5. For $f(x_1, x_2) = x_1 + 2x_2$, the isoquants are the parallel lines $x_1 + 2x_2 = C$. The gradient

$$\nabla f(x_1, x_2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is perpendicular to these lines. $\square$

Derivatives of higher order

We can take derivatives of higher order of functions of several variables. The problem is that, whereas for a function of one variable the first, the second, the third, etc, derivatives are just numbers, for functions of several variables they are objects of increasing complexity. This course does not go beyond second derivatives.

The notation $\partial_j\big(\partial_i f\big) = \partial^2_{ji} f$ is straightforward. A function of n variables has n derivatives of first order, and n^2 derivatives of second order. Nevertheless, it can be proved that, under certain constraints, the order does not matter for second derivatives, so that $\partial^2_{ij} f(X) = \partial^2_{ji} f(X)$. This result is called the **Schwarz theorem** in Mathematics textbooks.

The second derivatives can be packed in a square matrix, called the **Hessian** matrix,

$$\nabla^2 f(X) = \begin{bmatrix} \partial^2_{11} f(X) & \partial^2_{12} f(X) & \cdots & \partial^2_{1n} f(X) \\ \partial^2_{21} f(X) & \partial^2_{22} f(X) & \cdots & \partial^2_{nn} f(X) \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2_{n1} f(X) & \partial^2_{n2} f(X) & \cdots & \partial^2_{nn} f(X) \end{bmatrix}.$$

The Schwarz theorem tells us the Hessian matrix is symmetric.

Example 4 (continuation). Taking derivatives again, for $f(x_1, x_2) = e^{x_1 x_2}$,

$$\nabla^2 f(\mathbf{x}_1, x_2) = \begin{bmatrix} x_2^2\, e^{x_1 x_2} & (1 + x_1 x_2) e^{x_1 x_2} \\ (1 + x_1 x_2) e^{x_1 x_2} & x_2^2\, e^{x_1 x_2} \end{bmatrix}.$$

Homework

A. How are the isoquants of $f(x_1, x_2) = 2x_1^2 - x_2^2$?

B. For the unit circle $x_1^2 + x_2^2 = 1$, calculate the slope of the tangent line at an arbitrary point (x_1, x_2), with $x_2 > 0$. Check that the result is the same derived from the explicit function $x_2 = \sqrt{1 - x_1^2}$. Calculate the gradient of $f(x_1, x_2) = x_1^2 + x_2^2$ and check the orthogonality.