Optimization and Multivariate Calculus

Joe Patten

August 7, 2018

1 Derivatives

Recall from single-variable calculus, the derivative of a function f with respect to x at point x_0 is defined as:

$$\frac{df(x_0)}{dx} = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}$$

If this limit exists, then we say that f is differentiable at x_0. We can extend this definition to talk about derivatives of multivariate functions.

1.1 Partial Derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$. The partial derivative of f with respect to variable x_i at $\mathbf{x}^0$ is given by:

$$\frac{\partial f(\mathbf{x}^0)}{\partial x_i} = \lim_{h\to 0} \frac{f(x_1, x_1, ..., x_i + h, ..., x_n) - f(x_1, x_2, ..., x_i, ..., x_n)}{h}$$

Notice that in this definition, the ith variable is affected. To take the partial derivative of variable x_i, we treat all the other variables as constants.

Example

Consider the function: $f(x, y) = 4x^2y^5 + 3x^3y^2 + 6y + 10$.

$\frac{\partial f(x,y)}{\partial x} = 8xy^5 + 9x^2y^2$

$\frac{\partial f(x,y)}{\partial y} = 20x^2y^4 + 6x^3y + 6$

1.2 Gradient Vector

We can put all of the partials of the function $F : \mathbb{R}^n \to \mathbb{R}$ at x^* (which we call the derivative of F) in a row vector:

$$DF_{x^*} = \begin{bmatrix} \frac{\partial F(x^*)}{\partial x_1} & \dots & \frac{\partial F(x^*)}{\partial x_n} \end{bmatrix}$$

This can also be referred to as the Jacobian derivative of F.

We can express the derivative in a column vector:

$$\nabla F_{x^*} = \begin{bmatrix} \frac{\partial F(x^*)}{\partial x_1} \\ \vdots \\ \frac{\partial F(x^*)}{\partial x_n} \end{bmatrix}$$

This representation is usually referred to as the gradient vector.

Example

The gradient vector of our previous example would be:

$$\nabla F = \begin{bmatrix} 8xy^5 + 9x^2y^2 \\ 20x^2y^4 + 6x^3y + 6 \end{bmatrix}$$

1.3 Jacobian Matrix

We won't always be working with functions of the form $F : \mathbb{R}^n \to \mathbb{R}$. We might work with functions of the form $F : \mathbb{R}^n \to \mathbb{R}^m$. A common example example in economics is a production function that has n inputs and m outputs. Considering the production function example, notice that we can write this function as m functions:

$$\begin{aligned} q_1 =& f_1(x_1, x_2, ..., x_n) \\ q_2 =& f_2(x_1, x_2, ..., x_n) \\ &\vdots \\ q_m =& f_1(x_1, x_2, ..., x_n) \end{aligned}$$

We can put the functions and their respective partials in a matrix in order to get the Jacobian Matrix:

$$DF(x^*) = \begin{bmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \frac{\partial f_1(x^*)}{\partial x_2} & \cdots & \frac{\partial f_1(x^*)}{\partial x_n} \\ \frac{\partial f_2(x^*)}{\partial x_1} & \frac{\partial f_2(x^*)}{\partial x_2} & \cdots & \frac{\partial f_2(x^*)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(x^*)}{\partial x_1} & \frac{\partial f_m(x^*)}{\partial x_2} & \cdots & \frac{\partial f_m(x^*)}{\partial x_n} \end{bmatrix}$$

1.4 Hessian Matrix

Recall that for an function of n variables, there are n partial derivatives. We can take partial derivatives of each partial derivative. The partial derivative of a partial derivative is called the second order partial derivative.

Example

The second order partial derivatives for the example above are defined as:

$\frac{\partial^2 f(x,y)}{\partial x^2} = 8y^5 + 18xy^2$

$\frac{\partial^2 f(x,y)}{\partial y^2} = 80x^2y^3 + 6x^3$

$\frac{\partial^2 f(x,y)}{\partial y \partial x} = 40xy^4 + 18x^2y$

$\frac{\partial^2 f(x,y)}{\partial x \partial y} = 40xy^4 + 18x^2y$

The second order partial derivatives of the form $\frac{\partial^2 f(x,y)}{\partial x \partial y}$ where $x \neq y$ are called the cross partial derivatives. Notice from our example, that $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$. This is always the case with cross partials. We see that:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$$

We can put all of these second order partials into a matrix, which is referred to as the Hessian Matrix:

$$\begin{bmatrix} \frac{\partial^2 f(x^*)}{\partial x_1^2} & \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x^*)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x^*)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x^*)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x^*)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x^*)}{\partial x_n^2} \end{bmatrix}$$

2 Homogeneity

We say that a function $f(x_1, x_2, ..., x_n)$ is homogeneous of degree k (commonly referred to as HDk) if:

$$f(\alpha x_1, \alpha x_2, ..., \alpha x_n) = \alpha^k f(x_1, x_2, ..., x_n) \text{ for all } \mathbf{x} \text{ and all } \alpha > 0 \tag{1}$$

In Economics, when a production is homogeneous of degree 1, it is said to have constant returns to scale (CRS). If $k > 1$, the production function has increasing returns to scale, and if $k < 1$, the production function has decreasing returns to scale.

Example

Consider the function: $f(x, y) = 5x^2y^3 + 6x^6y^{-1}$. To determine if this function if homogeneous, we need to multiply each input by α:

$$\begin{aligned} f(\alpha x, \alpha y) &= 5(\alpha x)^2(\alpha y)^3 + 6(\alpha x)^6(\alpha y)^{-1} \\ &= \alpha^{2+3}5x^2y^3 + \alpha^{6-1}6x^6y^{-1} \\ &= \alpha^5(5x^2y^3 + 6x^6y^{-1}) \\ &= \alpha^5(f(x, y)) \end{aligned}$$

This function is homogeneous of degree 5 (HD5).

2.1 Euler's Theorem

If we take the derivative of both sides of equation (1) by x_i, we get the following:

$$\begin{aligned} \frac{\partial f(\alpha x_1, \alpha x_2, ..., \alpha x_n)}{\partial x_i} \cdot \alpha &= \alpha^k \frac{\partial f(x_1, x_2, ..., x_n)}{\partial x_i} \\ \frac{\partial f(\alpha x_1, \alpha x_2, ..., \alpha x_n)}{\partial x_i} &= \alpha^{k-1} \frac{\partial f(x_1, x_2, ..., x_n)}{\partial x_i} \end{aligned} \tag{2}$$

We can use the result from equation (2) to derive Euler's Thereom:

Theorem 1. If f is a C^1, homogeneous of degree k function on $\mathbb{R}^n_+$, then it follows:

$$x_1 \frac{\partial f(x)}{\partial x_1} + x_2 \frac{\partial f(x)}{\partial x_2} + \ldots + x_n \frac{\partial f(x)}{\partial x_n} = kf(x)$$

3 Convexity and Concavity

3.1 Convex Sets

A set A, in a real vector space V, is convex iff:

$$\lambda x_1 + (1 - \lambda)x_2 \in A$$

for any $\lambda \in [0, 1]$ and any $x_1, x_2 \in A$.

3.2 Function Concavity and Convexity

Let A be a convex set in vector space V. Consider the function $f : A \to \mathbb{R}$.

1. f is concave iff:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \tag{3}$$

   for any $x_1, x_2 \in A$ and $\lambda \in [0,1]$.

2. f is convex iff:

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2) \tag{4}$$

   for any $x_1, x_2 \in A$ and $\lambda \in [0,1]$.

3. f is strictly concave iff:

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda)f(x_2) \tag{5}$$

   for any $x_1, x_2 \in A$ and $\lambda \in [0,1]$.

4. f is strictly convex iff:

$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda)f(x_2) \tag{6}$$

   for any $x_1, x_2 \in A$ and $\lambda \in [0,1]$.