Optimization and Multivariate Calculus

Joe Patten

August 6, 2019

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:

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

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:

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

Let the function $f : A \to \mathbb{R}$ be a C^2 function, where A is a convex and open set in $\mathbb{R}^n$.

- f is strictly convex iff its Hessian matrix is positive definite for any $x \in A$.
- f is strictly concave iff its Hessian matrix is negative definite for any $x \in A$.
- f is (weakly) convex iff its Hessian matrix is positive semidefinite for any $x \in A$.
- f is (weakly) concave iff its Hessian matrix is negative semidefinite for any $x \in A$.

2 Convexity and Concavity

2.1 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 $x_1, x_2 \in A$.

2.2 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:

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

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

2. f is convex iff:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{2}$$

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

3. f is strictly concave iff:

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

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

4. f is strictly convex iff:

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

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

Note

If a function is not convex, it does not mean that it is concave. Likewise, if a function is not concave, it does not mean that it is convex.

Practice

Consider $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ where A is a convex set in a vector space. If f and g are concave functions show that:

1. $f + g$ is a concave function.
2. cf is a concave function if $c > 0$, and a convex function if $c < 0$.

2.3 Jensen's Inequality

Let the function $f : A \to \mathbb{R}$ where A is a convex set in a vector space, then:

- f is concave iff

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq \sum_{i=1}^{n} \lambda_i f(x_i)$$

- f is convex iff

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

  for any $\lambda_1, ..., \lambda_n \in \mathbb{R}_+$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and $x_1, ..., x_n \in A$

2.4 Quasiconcave and Quasiconvex

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

1. f is quasiconcave iff:
$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \min\{f(x_1), f(x_2)\} \tag{5}$$
   for any $x_1, x_2 \in A$ and $\lambda \in [0, 1]$.
2. f is quasiconvex iff:
$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \max\{f(x_1), f(x_2)\} \tag{6}$$
   for any $x_1, x_2 \in A$ and $\lambda \in [0, 1]$.
3. f is strictly quasiconcave iff:
$$f(\lambda x_1 + (1 - \lambda)x_2) > \min\{f(x_1), f(x_2)\} \tag{7}$$
   for any $x_1, x_2 \in A$ and $\lambda \in [0, 1]$.
4. f is strictly quasiconvex iff:
$$f(\lambda x_1 + (1 - \lambda)x_2) < \max\{f(x_1), f(x_2)\} \tag{8}$$
   for any $x_1, x_2 \in A$ and $\lambda \in [0, 1]$.

Practice

1. Show that if a function f is concave, then f is also quasiconcave.
2. Show that if a function f is convex, then f is also quasiconvex.

2.5 Contour Sets

Let A be a convex set in vector space V. Consider the function $f : A \to \mathbb{R}$. An upper contour set of $a \in \mathbb{R}$ is defined as:

$$\{x \in A : f(x) \geq a\}$$

A lower contour set of $a \in A$ is defined similarly:

$$\{x \in A : f(x) \leq a\}$$

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

1. f is quasiconcave iff its upper contour set is convex for any $a \in \mathbb{R}$
2. f is quasiconvex iff its lower contour set is convex for any $a \in \mathbb{R}$

2.6 Graphs

Let the function $f : A \to \mathbb{R}$. The graph of f is defined as the following set:

$$G(f) = \{(x, y) \in A \times \mathbb{R} : y = f(x)\}$$

The epigraph is the set above the graph, and is defined as:

$$G^{+}(f) = \{(x, y) \in A \times \mathbb{R} : y \geq f(x)\}$$

The subgraph is the set below the graph, and is defined as:

$$G^{-}(f) = \{(x, y) \in A \times \mathbb{R} : y \leq f(x)\}$$

The following theorem follows:

1. $G^{-}(f)$ is a convex set iff f is concave.
2. $G^{+}(f)$ is a convex set iff f is convex.

3 Multivariate Calculus

3.1 Derivatives

Let $f(x)$ and $g(x)$ be differentiable functions, and $a, n \in \mathbb{R}$. Derivatives have following properties:

1. $(af)' = af'(x)$
2. $(f + g)' = f'(x) + g'(x)$
3. $(fg)' = f'g + fg'$
4. $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$
5. $\frac{d}{dx}(c) = 0$
6. $\frac{d}{dx}(f(g(x)) = f'(g(x))g'(x)$

3.2 Integrals

Integrals have the following properties:

1. $\int af(x)dx = a \int f(x)dx$
2. $\int (f + g)dx = \int f(x)dx + \int g(x)dx$

3.3 Integration by Parts

We can use integration by parts to integrate some more complex expressions. The formula for integration by parts is:

$$\int u(x) \cdot v'(x) dx = u(x) \cdot v(x) - \int u'(x) \cdot v(x) dx$$

Example

Using integration by parts, we can integrate the expression xe^{2x}:
Let $u(x) = x$, and $v'(x) = e^{2x}$. Thus $u'(x) = 1$ and $v(x) = \frac{1}{2}e^{2x}$. Using the integration by parts, we see that:

$$\begin{aligned} \int xe^{2x} dx &= x\frac{1}{2}e^{2x} - \int 1 \cdot \frac{1}{2}e^{2x} dx \\ &= \frac{1}{2}\left(xe^{2x} - \int e^{2x} dx\right) \\ &= \frac{1}{2}xe^{2x} - \frac{1}{4}e^{2x} + C \end{aligned}$$

where $C \in \mathbb{R}$.

Exercises

Let A be a convex subset of $\mathbb{R}^n$ where $f : A \to \mathbb{R}$. Let f be concave.

1. Show that f is quasiconcave.
2. Show that cf is a concave function when $c > 0$, and cf is a convex function when $c < 0$.