1.6 Multivariable functions

Until now, we only considered functions $f(x)$ of one variable. In this section, we introduce functions of several variables, either taking values in the one-dimensional space $\mathbb{R}$, i.e., scalar valued multivariable functions, or taking values in the m-dimensional space $\mathbb{R}^m$, i.e., vector valued multivariable functions.

Scalar Valued Functions

Let $f : \mathbb{R}^n \to R$ be a function of n variables denoted by $x_1, x_2, ..., x_n$, and let $x = (x_1, x_2, ..., x_n)$.

Definition 1.2.

Let $f : \mathbb{R}^n \to \mathbb{R}$. The partial derivative of the function $f(x)$ with respect to the variable x_i is denoted by $\frac{\partial f}{\partial x_i}(x)$ and is defined as

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \frac{f(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - f(x_1, x_2, ..., x_n)}{h}$$

if the limit exists and is finite.

In practice the partial derivative $\frac{\partial f}{\partial x_i}(x)$ is computed by considering the variables $x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$ to be fixed, and differentiating $f(x)$ as a function of one variable x_i.

A compact formula for the above limit can be given as follows: Let e_i be the vector

- with all entries equal to 0
- with the exception of the i-th entry, which is equal to 1,
- i.e. $e(j) = 0$ for $j \neq i$, $1 \leq j \leq n$ and $e_{i(j)} = 1$.

Then,

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h}$$

Partial derivatives of higher order are defined similarly. For example, the second order partial derivative of $f(x)$ first with respect to x_i and then with respect to x_j, with $j \neq i$ is denoted by $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ and is equal to

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

While the second and third partial derivatives of $f(x)$ with respect to x_i are denoted by $\frac{\partial^2 f}{\partial x_i^2}$ and $\frac{\partial^2 f}{\partial x_i^3}$, respectively, and are given by

$$\frac{\partial^2 f}{\partial x_i^2}(x) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}(x)\right)$$

$$\frac{\partial^2 f}{\partial x_i^3}(x) = \frac{\partial}{\partial x_i}\left(\frac{\partial^2 f}{\partial x_i^2}(x)\right)$$

While the order in which the partial derivatives of a given function are computed might make a difference, i.e.,

- the partial derivative of $f(x)$ first with respect to x_i and then with respect to x_j, with $j \neq i$ is not necessarily equal to

- the partial derivative of $f(x)$ first with respect to x_j and then with respect to x_i, this is not the case if a function is smooth enough

Theorem 1.9.

If all the partial derivatives of order k of the function $f(x)$ exist and are continuous, then the order in which partial derivatives of $f(x)$ of order at most k is computed does not matter.

Definition 1.3.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables and assume that $f(x)$ is differentiable with respect to all variables $x_i, i = 1 : n$. The gradient $Df(x)$ of the function $f(x)$ is the following row vector of size n:

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x) \quad \frac{\partial f}{\partial x_2}(x) \quad \dots \quad \frac{\partial f}{\partial x_n}(x) \right)$$

Definition 1.4.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables. The Hessian of $f(x)$ is denoted by $D^2 f(x)$ and is defined as the following $n \times n$ matrix:

$$D^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

Another commonly used notations for the gradient and Hessian of $f(x)$ are $\nabla f(x)$ and $Hf(x)$, respectively. We will use $Df(x)$ and $D^2 f(x)$ unless otherwise specified.

Vector Valued Functions

A function that takes values in a multidimensional space is called a vector valued function. Let $F : \mathbb{R}^n \to R^m$ be a vector valued function given by

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{m(x)} \end{pmatrix}$$

Where $x = (x_1, x_2, ..., x_n)$.

Definition 1.5.

Let $F : \mathbb{R}^n \to R^m$ given by $F(x) = \left(f_{j(x)} \right)_{j=1:m}$, and assume that the functions $f_{j(x)}, j = 1 : m$ are differentiable with respect to all variables $x_i, i = 1 : m$. The gradient $DF(x)$ of the function $F(x)$ is the following matrix of size $m \times n$:

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

If $F : \mathbb{R}^n \to R^n$, then the gradient $DF(x)$ is a square matrix of size n.

The j-th row of the gradient matrix $DF(x)$ is equal to the gradient $Df_{j(x)}$ of the function $f_{j(x)}, j = 1 : m$. Therefore,

$$DF(x) = \begin{pmatrix} Df_1(x) \\ Df_2(x) \\ \ddots \\ Df_{m(x)} \end{pmatrix}$$

1.6.1 Functions of two variables

Functions of two variables are the simplest example of multivariable functions. To classify the definitions for partial derivatives and for the gradient and the Hessian of multivariable functions given in section 1.6, we present them againfor both scalar and vector valued functions of two variables.

Scalar Valued Functions

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables denoted by x and y. The partial derivatives of the function $f(x, y)$ with respect to the variables x and y are denoted by

- $\frac{\partial f}{\partial x}(x, y)$ and
- $\frac{\partial f}{\partial y}(x, y)$, respectively,

and defined as follows:

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

The gradient of $f(x, y)$ is

$$Df(x, y) = \left(\frac{\partial f}{\partial x}(x, y) \quad \frac{\partial f}{\partial y}(x, y) \right)$$

The Hessian of $f(x, y)$ is

$$D^2 f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix}$$

Vector Valued Functions

Let $F : \mathbb{R}^2 \to R^2$ given by

$$F(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix}$$

The gradient of $F(x, y)$ is

$$DF(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) \end{pmatrix}$$