Multivariate calculus

Partial derivatives

If $f : R^n \to R$ with $(x_1, x_2, \cdots, x_n) \to f(x_1, x_2, \cdots, x_n)$ then we define its partial derivatives (if they exist) as follows:

$$\frac{\partial f}{\partial x_i} = \frac{df}{dx_i}$$

assuming all variables other than x_i are fixed

If $f : R^n \to R$ has partial derivatives, the vector

$$\nabla f = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n})$$

is called the gradient of f at x

The gradient gives **the direction of greatest growth** of f at x

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}(3x^2y^2) = 6xy^2$$

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

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}(2x^3y) = 6x^2y$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}(3x^2y^2) = 6x^2y$$

In general, the order of taking the mixed derivatives matters

Proposition 39
Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has, for some $i \neq j$, first partial derivatives $\frac{\partial f}{\partial x_i}$ and $\frac{\partial f}{\partial x_j}$ in some open disk $\{x \in \mathbb{R}^n : |x - y| < r\}$, where $y \in \mathbb{R}^n$ is a given point, $|x - y| = \sqrt{\sum_{m=1}^{n}(x_m - y_m)^2}$ and $r > 0$. If the second partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ are continuous at y, then $\frac{\partial^2 f}{\partial x_i \partial x_j}(y) = \frac{\partial^2 f}{\partial x_j \partial x_i}(y)$.

From now on we assume we deal with well-behaved functions for which **all mixed derivatives are always the same**

Critical points of a multivariate function

A **critical point** of a function $f : R^n \to R$ is a point $x \in R^n$ such that $\nabla f(x) = 0$

Let’s recall how we determine the nature of a critical point of a function of a single variable as minimum, maximum or point of inflection

We assume the function is analytic, so it has derivatives of all order

If x is a critical point of $f : R \to R$, i.e., $f'(x) = 0$, we can use the Taylor series to find the value of f at point y near x:

$$f(y) = f(x) + f'(x)(y-x) + f''(x)(y-x)^2/2 + \cdots + (y-x)^n f^{(n)}(x)/(n)! + \cdots$$

Let n be the smallest integer such that $f^{(n)}(x) \neq 0$

Since f is analytic, such n always exists

As $f'(x) = 0$ we obtain near x:

$$f(y) \approx (y-x)^{(n)} f^{(n}(x)/(n)!$$

So the behaviour of the function at y near the critical point x is determined by the term $f^{(n)}(x)(y-x)^n$

If n is even, and $f^{(n)}(x) > 0$ it follows that the function is increasing at x_0 since the term $(y-x)^n$ is positive; so x is a minimum

If n is even, and $f^{(n)}(x) < 0$ it follows that the function is decreasing at x, so x is a maximum

If n is odd, it has a point of inflection

Characterising Critical Points

In higher dimensions, a generic or non-degenerate critical point of a multivariate function is either a minimum, a maxima or a saddle

We will define below what we mean by “generic” or “non-degenerate”

We will present the analysis for $n = 2$ as the formulation for $n > 2$ is entirely similar

$$a := \frac{\partial^2 f}{\partial x_1^2}, \quad b := \frac{\partial^2 f}{\partial x_2 \partial x_1} \quad c := \frac{\partial^2 f}{\partial x_2^2}$$

$a > 0 \land b^2 - ac < 0 \implies$ Critical Point is a minimum
$c > 0 \land b^2 - ac < 0 \implies$ CP is a minimum
$a > 0 \land c > 0 \land b^2 - ac < 0 \implies$ CP is a minimum
$(a + c) > 0 \land b^2 - ac < 0 \implies$ CP is a minimum

$a < 0 \land b^2 - ac < 0 \implies$ CP is a maximum
$c < 0 \land b^2 - ac < 0 \implies$ CP is a maximum
$a < 0 \land c < 0 \land b^2 - ac < 0 \implies$ CP is a maximum
$a + c < 0 \land b^2 - ac < 0 \implies$ CP is a maximum

$b^2 - ac > 0 \implies$ CP is a saddle point

$b^2 - ac = 0 \implies$ CP is degenerate (we don't know what it is)

Extension to $n > 2$:

If $f : R^n \to R$, has a critical point at $(x_i)_{1 \le i \le n}$ then we calculate its $n \times n$ Hessian matrix at this point, whose ij entry is given by:

$$(H(x_1, \ldots, x_n))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

All Eigenvalues of $H(x_1, \ldots, x_n)$ are positive $\implies$ CP is a minimum

All Eigenvalues of $H(x_1, \ldots, x_n)$ are negative $\implies$ CP is a maximum

$H(x_1, \ldots, x_n)$ has both positive and negative eigenvalues with no zero eigenvalues $\implies$ CP is a saddle point