Functions of More Than Two Variables: Extension of Clairaut's Theorem and Criteria for local Maximums and Minimums (selections from various Wkipedia sources)

(See Wikipedia:https://en.wikipedia.org/wiki/Symmetry_of_second_derivatives#Schwarz.27s_theorem):

Schwarz's theorem [edit]

In mathematical analysis, *Schwarz's theorem* (or *Clairaut's theorem*[3]) named after Alexis Clairaut and Hermann Schwarz, states that if

$$f\colon \mathbb{R}^n \to \mathbb{R}$$

has continuous second partial derivatives at any given point in $\mathbb{R}^n$, say, $(a_1, \ldots, a_n)$, then $\forall i, j \in \{1, 2, \ldots, n\}$,

$$\frac{\partial^2 f}{\partial x_i\, \partial x_j}(a_1, \ldots, a_n) = \frac{\partial^2 f}{\partial x_j\, \partial x_i}(a_1, \ldots, a_n).$$

The partial derivations of this function are commutative at that point. One easy way to establish this theorem (in the case where $n = 2$, $i = 1$, and $j = 2$, which readily entails the result in general) is by applying Green's theorem to the gradient of f.

Sufficiency of twice-differentiability [edit]

A weaker condition than the continuity of second partial derivatives (which is implied by the latter) which nevertheless suffices to ensure symmetry is that all partial derivatives are themselves differentiable.[4]

Hessian matrix

From Wikipedia, the free encyclopedia

In mathematics, the **Hessian matrix** or **Hessian** is a square matrix of second-order partial derivatives of a scalar-valued function, or scalar field. It describes the local curvature of a function of many variables. The Hessian matrix was developed in the 19th century by the German mathematician Ludwig Otto Hesse and later named after him. Hesse originally used the term "functional determinants".

Specifically, suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a function taking as input a vector $\mathbf{x} \in \mathbb{R}^n$ and outputting a scalar $f(\mathbf{x}) \in \mathbb{R}$; if all second partial derivatives of f exist and are continuous over the domain of the function, then the Hessian matrix $\mathbf{H}$ of f is a square $n \times n$ matrix, usually defined and arranged as follows:

$$\mathbf{H} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\, \partial x_n} \\[2.2ex] \dfrac{\partial^2 f}{\partial x_2\, \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\, \partial x_n} \\[2.2ex] \vdots & \vdots & \ddots & \vdots \\[2.2ex] \dfrac{\partial^2 f}{\partial x_n\, \partial x_1} & \dfrac{\partial^2 f}{\partial x_n\, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

or, component-wise:

$$\mathbf{H}_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

The determinant of the above matrix is also sometimes referred to as the Hessian.[1]

Mixed derivatives and symmetry of the Hessian [edit]

The **mixed derivatives** of f are the entries off the main diagonal in the Hessian. Assuming that they are continuous, the order of differentiation does not matter (Schwarz's theorem). For example,

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

In a formal statement: if the second derivatives of f are all continuous in a neighborhood D, then the Hessian of f is a symmetric matrix throughout D; see symmetry of second derivatives.

Critical points [edit]

If the gradient (the vector of the partial derivatives) of a function f is zero at some point $\mathbf{x}$, then f has a *critical point* (or *stationary point*) at $\mathbf{x}$. The determinant of the Hessian at $\mathbf{x}$ is then called the discriminant. If this determinant is zero then $\mathbf{x}$ is called a *degenerate critical point* of f, or a *non-Morse critical point* of f. Otherwise it is non-degenerate, and called a *Morse critical point* of f.

The Hessian matrix plays an important role in Morse theory and catastrophe theory, because its kernel and eigenvalues allow classification of the critical points.[2][3][4]

Second derivative test [edit]

Main article: Second partial derivative test

The Hessian matrix of a convex function is positive semi-definite. Refining this property allows us to test if a critical point x is a local maximum, local minimum, or a saddle point, as follows:

If the Hessian is positive definite at x, then f attains a local minimum at x. If the Hessian is negative definite at x, then f attains a local maximum at x. If the Hessian has both positive and negative eigenvalues then x is a saddle point for f. Otherwise the test is inconclusive. This implies that, at a local minimum (respectively, a local maximum), the Hessian is positive-semi-definite (respectively, negative semi-definite).

Note that for positive semidefinite and negative semidefinite Hessians the test is inconclusive (yet a conclusion can be made that f is locally convex or concave respectively). However, more can be said from the point of view of Morse theory.

The second derivative test for functions of one and two variables is simple. In one variable, the Hessian contains just one second derivative; if it is positive then x is a local minimum, and if it is negative then x is a local maximum; if it is zero then the test is inconclusive. In two variables, the determinant can be used, because the determinant is the product of the eigenvalues. If it is positive then the eigenvalues are both positive, or both negative. If it is negative then the two eigenvalues have different signs. If it is zero, then the second derivative test is inconclusive.

Equivalently, the second-order conditions that are sufficient for a local minimum or maximum can be expressed in terms of the sequence of principal (upper-leftmost) minors (determinants of sub-matrices) of the Hessian; these conditions are a special case of those given in the next section for bordered Hessians for constrained optimization—the case in which the number of constraints is zero. Specifically, the sufficient condition for a minimum is that all of these principal minors be positive, while the sufficient condition for a maximum is that the minors alternate in sign with the 1×1 minor being negative.

Local Maximums and Minimums for Functions of Many Variables

For a function f of more than two variables, there is a generalization of the rule above. In this context, instead of examining the determinant of the Hessian matrix, one must look at the eigenvalues of the Hessian matrix at the critical point. The following test can be applied at any critical point (a, b, ...) for which the Hessian matrix is invertible:

1. If the Hessian is positive definite (equivalently, has all eigenvalues positive) at (a, b, ...), then f attains a local minimum at (a, b, ...).
2. If the Hessian is negative definite (equivalently, has all eigenvalues negative) at (a, b, ...), then f attains a local maximum at (a, b, ...).
3. If the Hessian has both positive and negative eigenvalues then (a, b, ...) is a saddle point for f (and in fact this is true even if (a, b, ...) is degenerate).

In those cases not listed above, the test is inconclusive.[2]

Note that for functions of three or more variables, the *determinant* of the Hessian does not provide enough information to classify the critical point, because the number of jointly sufficient second-order conditions is equal to the number of variables, and the sign condition on the determinant of the Hessian is only one of the conditions. Note also that this statement of the second derivative test for many variables also applies in the two-variable and one-variable case. In the latter case, we recover the usual second derivative test.

In the two variable case, $D(a,b)$ and $f_{xx}(a,b)$ are the principal minors of the Hessian. The first two conditions listed above on the signs of these minors are the conditions for the positive or negative definiteness of the Hessian. For the general case of an arbitrary number *n* of variables, there are *n* sign conditions on the *n* principal minors of the Hessian matrix that together are equivalent to positive or negative definiteness of the Hessian (Sylvester's criterion): for a local minimum, all the principal minors need to be positive, while for a local maximum, the minors with an odd number of rows and columns need to be negative and the minors with an even number of rows and columns need to be positive. See Hessian matrix#Bordered Hessian for a discussion that generalizes these rules to the case of equality-constrained optimization.

Sylvester's criterion

From Wikipedia, the free encyclopedia

In mathematics, **Sylvester's criterion** is a necessary and sufficient criterion to determine whether a Hermitian matrix is positive-definite. It is named after James Joseph Sylvester.

Sylvester's criterion states that a Hermitian matrix *M* is positive-definite if and only if all the following matrices have a positive determinant:

- the upper left 1-by-1 corner of *M*,
- the upper left 2-by-2 corner of *M*,
- the upper left 3-by-3 corner of *M*,
- $\vdots$
- *M* itself.

In other words, all of the leading principal minors must be positive.

An analogous theorem holds for characterizing positive-semidefinite Hermitian matrices, except that it is no longer sufficient to consider only the *leading* principal minors: a Hermitian matrix *M* is positive-semidefinite if and only if all principal minors of *M* are nonnegative.[1]