Hessian matrix

Canonical name	HessianMatrix
Date of creation	2013-03-22 12:59:41
Last modified on	2013-03-22 12:59:41
Owner	cvalente (11260)
Last modified by	cvalente (11260)
Numerical id	31
Author	cvalente (11260)
Entry type	Definition
Classification	msc 26B12
Related topic	Gradient
Related topic	PartialDerivative
Related topic	SymmetricMatrix
Related topic	ComplexHessianMatrix
Related topic	HessianForm
Related topic	DirectionalDerivative
Defines	Hessian

Let $x \in \mathbb{R}^n$ and let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a real-valued function having 2nd-order partial derivatives in an open set U containing x. The *Hessian matrix* of f is the matrix of second partial derivatives evaluated at x:

$$\mathbf{H}(x) := \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}. \tag{1}$$

If f is in $C^2(U)$, $\mathbf{H}(x)$ is `http://planetmath.org/SymmetricMatrix`symmetric because of the equality of mixed partials. Note that $\mathbf{H}(x) = \mathbf{J}(\nabla f)$, the Jacobian of the gradient of f.

Given a vector $\boldsymbol{v} \in \mathbb{R}^n$, the *Hessian* of f at $\boldsymbol{v}$ is:

$$\mathbf{H}(x)(\boldsymbol{v}) := \frac{1}{2}\boldsymbol{v}\mathbf{H}(x)\boldsymbol{v}^{\mathrm{T}}. \tag{2}$$

Here we view $\boldsymbol{v}$ as a 1 by n matrix so that $\boldsymbol{v}^{\mathrm{T}}$ is the transpose of $\boldsymbol{v}$.

Remark. The Hessian of f at $\boldsymbol{v}$ is a quadratic form, since $\mathbf{H}(x)(r\boldsymbol{v}) = r^2\mathbf{H}(x)(\boldsymbol{v})$ for any $r \in \mathbb{R}$.

If f is further assumed to be in $C^2(U)$, and x is a critical point of f such that $\mathbf{H}(x)$ is `http://planetmath.org/PositiveDefinite`positive definite, then x is a strict local minimum of f.

This is not difficult to show. Since $\mathbf{H}(x)$ is `http://planetmath.org/PositiveDefinite`positive definite, the Rayleigh-Ritz theorem shows that there is a $c > 0$ such that for all $h \in \mathbb{R}^n$, $h^T\mathbf{H}(x)h \geq 2c\|h\|^2$. Thus by `http://planetmath.org/TaylorPolynomialsInBanachS` theorem (form)

$$f(x+h) = f(x) + \frac{1}{2}h^T\mathbf{H}(x)h + o(\|h\|^2) \geq c\|h\|^2 + o(\|h\|^2).$$

For small $\|h\|$ the first on the the second, so that both sides are positive for small $\|h\|$.