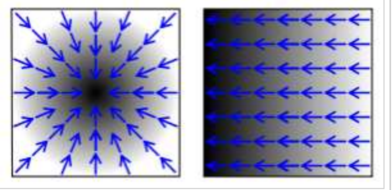
Gradient

From Wikipedia, the free encyclopedia



In the above two images, the values of the function are represented in black and white, black representing higher values, and its corresponding gradient is represented by blue arrows.

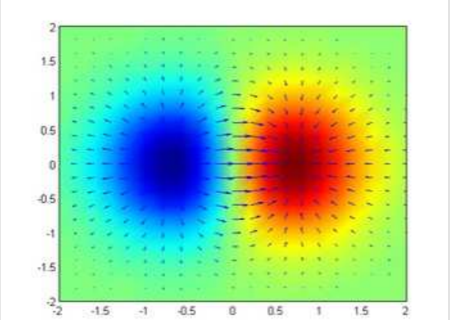
In mathematics, the **gradient** is a generalization of the usual concept of derivative of a function in one dimension to a function in several dimensions. If $f(x_1, ..., x_n)$ is a differentiable, scalar-valued function of standard Cartesian coordinates in Euclidean space, its gradient is the vector whose components are the *n* partial derivatives of *f*. It is thus a vector-valued function.

Similarly to the usual derivative, the gradient represents the slope of the tangent of the graph of the function. More precisely, the gradient points in the direction of the greatest rate of increase of the function and its magnitude is the slope of the graph in that direction. The components of the gradient in coordinates are the coefficients of the variables in the equation of the tangent space to the graph. This characterizing property of the gradient allows it to be defined independently of a choice of coordinate system, as a vector field whose components in a coordinate system will transform when going from one coordinate system to another.

The Jacobian is the generalization of the gradient for vector-valued functions of several variables and differentiable maps between Euclidean spaces or, more generally, manifolds. A further generalization for a function between Banach spaces is the Fréchet derivative.

Contents

Motivation



Gradient of the 2-d function $f(x, y) = xe^{-(x^2 + y^2)}$ is plotted as blue arrows over the pseudocolor plot of the function.

Consider a room in which the temperature is given by a scalar field, T, so at each point (x, y, z) the temperature is $T(x, y, z)$. (We will assume that the temperature does not change over time.) At each point in the room, the gradient of *T* at that point will show the direction the temperature rises most quickly. The magnitude of the gradient will determine how fast the temperature rises in that direction.

Consider a surface whose height above sea level at a point (x, y) is $H(x, y)$. The gradient of *H* at a point is a vector pointing in the direction of the steepest slope or grade at that point. The steepness of the slope at that point is given by the magnitude of the gradient vector.

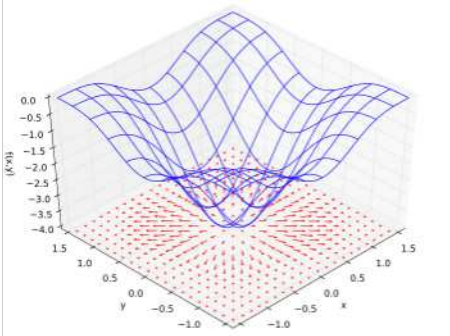
The gradient can also be used to measure how a scalar field changes in other directions, rather than just the direction of greatest change, by taking a dot product. Suppose that the steepest slope on a hill is 40%. If a road goes directly up the hill, then the steepest slope on the road will also be 40%. If, instead, the road goes around the hill at an angle, then it will have a shallower slope. For example, if the angle between the road and the uphill direction, projected onto the horizontal plane, is 60°, then the steepest slope along the road will be 20%, which is 40% times the cosine of 60°.

This observation can be mathematically stated as follows. If the hill height function *H* is differentiable, then the gradient of *H* dotted with a unit vector gives the slope of the hill in the direction of the vector. More precisely, when *H* is differentiable, the dot product of the gradient of *H* with a given unit vector is equal to the directional derivative of *H* in the direction of that unit vector.

Definition

The gradient (or gradient vector field) of a scalar function $f(x_1, x_2, x_3, ..., x_n)$ is denoted ∇f or $\vec{\nabla} f$ where ∇ (the nabla symbol) denotes the vector differential operator, del. The notation "grad(f)" is also commonly used for the gradient. The gradient of *f* is defined as the unique vector field whose dot product with any vector **v** at each point *x* is the directional derivative of *f* along **v**. That is,

$$(\nabla f(x)) \cdot \mathbf{v} = D_{\mathbf{v}} f(x).$$



The gradient of the function $f(x,y) = -(\cos^2 x + \cos^2 y)^2$ depicted as a projected vector field on the bottom plane.

In a rectangular coordinate system, the gradient is the vector field whose components are the partial derivatives of *f*:

$$\nabla f = \frac{\partial f}{\partial x_1}\mathbf{e}_1 + \cdots + \frac{\partial f}{\partial x_n}\mathbf{e}_n$$

where the $\mathbf{e}_i$ are the orthogonal unit vectors pointing in the coordinate directions. When a function also depends on a parameter such as time, the gradient often refers simply to the vector of its spatial derivatives only.

In the three-dimensional Cartesian coordinate system, this is given by

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

where **i**, **j**, **k** are the standard unit vectors. For example, the gradient of the function

$$f(x,y,z) = \ 2x + 3y^2 - \sin(z)$$

is:

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} = 2\mathbf{i} + 6y\mathbf{j} - \cos(z)\mathbf{k}.$$

In some applications it is customary to represent the gradient as a row vector or column vector of its components in a rectangular coordinate system.

Gradient and the derivative or differential

Linear approximation to a function

The gradient of a function *f* from the Euclidean space $\mathbb{R}^n$ to $\mathbb{R}$ at any particular point x_0 in $\mathbb{R}^n$ characterizes the best linear approximation to *f* at x_0. The approximation is as follows:

$$f(x) \approx f(x_0) + (\nabla f)_{x_0} \cdot (x - x_0)$$

for *x* close to x_0, where $(\nabla f)_{x_0}$ is the gradient of *f* computed at x_0, and the dot denotes the dot product on $\mathbb{R}^n$. This equation is equivalent to the first two terms in the multi-variable Taylor Series expansion of *f* at x_0.

Differential or (exterior) derivative

The best linear approximation to a function

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

at a point *x* in $\mathbb{R}^n$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$ which is often denoted by $\mathrm{d}f_x$ or $Df(x)$ and called the **differential** or (**total**) **derivative** of *f* at *x*. The gradient is therefore related to the differential by the formula

$$(\nabla f)_x \cdot v = \mathrm{d}f_x(v)$$

for any $v \in \mathbb{R}^n$. The function df, which maps x to df_x, is called the differential or exterior derivative of f and is an example of a differential 1-form.

If $\mathbb{R}^n$ is viewed as the space of (length n) column vectors (of real numbers), then one can regard df as the row vector with components

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right)$$

so that d$f_x(v)$ is given by matrix multiplication. The gradient is then the corresponding column vector, i.e.,

$$(\nabla f)_i = \mathrm{d}f_i^\mathsf{T}.$$

Gradient as a derivative

Let U be an open set in $\mathbf{R}^n$. If the function $f : U \to \mathbf{R}$ is differentiable, then the differential of f is the (Fréchet) derivative of f. Thus ∇f is a function from U to the space $\mathbf{R}$ such that

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - \nabla f(x) \cdot h\|}{\|h\|} = 0$$

where · is the dot product.

As a consequence, the usual properties of the derivative hold for the gradient:

Linearity

The gradient is linear in the sense that if f and g are two real-valued functions differentiable at the point $a \in \mathbf{R}^n$, and α and β are two constants, then αf + βg is differentiable at a, and moreover

$$\nabla\left(\alpha f + \beta g\right)(a) = \alpha \nabla f(a) + \beta \nabla g(a).$$

Product rule

If f and g are real-valued functions differentiable at a point $a \in \mathbf{R}^n$, then the product rule asserts that the product $(fg)(x) = f(x)g(x)$ of the functions f and g is differentiable at a, and

$$\nabla (fg)(a) = f(a)\nabla g(a) + g(a)\nabla f(a).$$

Chain rule

Suppose that $f : A \to \mathbf{R}$ is a real-valued function defined on a subset A of $\mathbf{R}^n$, and that f is differentiable at a point a. There are two forms of the chain rule applying to the gradient. First, suppose that the function g is a parametric curve; that is, a function $g : I \to \mathbf{R}^n$ maps a subset $I \subset \mathbf{R}$ into $\mathbf{R}^n$. If g is differentiable at a point $c \in I$ such that $g(c) = a$, then

$$(f\circ g)'(c) = \nabla f(a)\cdot g'(c),$$

where ∘ is the composition operator : (g ∘ f)(x) = g(f(x)). More generally, if instead $I \subset \mathbf{R}^k$, then the following holds:

$$\nabla (f\circ g)(c) = \big(Dg(c)\big)^\mathsf{T} \big(\nabla f(a)\big)$$

where $(Dg)^\mathrm{T}$ denotes the transpose Jacobian matrix.

For the second form of the chain rule, suppose that $h : I \to \mathbf{R}$ is a real valued function on a subset I of $\mathbf{R}$, and that h is differentiable at the point $f(a) \in I$. Then

$$\nabla (h\circ f)(a) = h'\big(f(a)\big)\nabla f(a).$$

Further properties and applications

Level sets

A level surface, or isosurface, is the set of all points where some function has a given value.

If f is differentiable, then the dot product $(\nabla f)_x \cdot v$ of the gradient at a point x with a vector v gives the directional derivative of f at x in the direction v. It follows that in this case the gradient of f is orthogonal to the level sets of f. For example, a level surface in three-dimensional space is defined by an equation of the form $F(x, y, z) = c$. The gradient of F is then normal to the surface.

More generally, any embedded hypersurface in a Riemannian manifold can be cut out by an equation of the form $F(P) = 0$ such that dF is nowhere zero. The gradient of F is then normal to the hypersurface.

Similarly, an affine algebraic hypersurface may be defined by an equation $F(x_1, ..., x_n) = 0$, where F is a polynomial. The gradient of F is zero at a singular point of the hypersurface (this is the definition of a singular point). At a non-singular point, it is a nonzero normal vector.

Conservative vector fields and the gradient theorem

The gradient of a function is called a gradient field. A (continuous) gradient field is always a conservative vector field: its line integral along any path depends only on the endpoints of the path, and can be evaluated by the gradient theorem (the fundamental theorem of calculus for line integrals). Conversely, a (continuous) conservative vector field is always the gradient of a function.

Riemannian manifolds

For any smooth function f on a Riemannian manifold (M,g), the gradient of f is the vector field ∇f such that for any vector field X,

$$g(\nabla f, X) = \partial_X f, \qquad \text{i.e.,} \quad g_x((\nabla f)_x, X_x) = (\partial_X f)(x)$$

where $g_x(\ ,\ )$ denotes the inner product of tangent vectors at x defined by the metric g and $\partial_X f$ (sometimes denoted $X(f)$) is the function that takes any point $x \in M$ to the directional derivative of f in the direction X, evaluated at x. In other words, in a coordinate chart φ from an open subset of M to an open subset of $\mathbf{R}^n$, $(\partial_X f)(x)$ is given by:

$$\sum_{j=1}^{n} X^j(\varphi(x)) \frac{\partial}{\partial x_j}(f \circ \varphi^{-1})\Big|_{\varphi(x)},$$

where X^j denotes the jth component of X in this coordinate chart.

So, the local form of the gradient takes the form:

$$\nabla f = g^{ik} \frac{\partial f}{\partial x^k} \frac{\partial}{\partial x^i}.$$

Generalizing the case $M = \mathbf{R}^n$, the gradient of a function is related to its exterior derivative, since

$$(\partial_X f)(x) = df_x(X_x)\ .$$

More precisely, the gradient ∇f is the vector field associated to the differential 1-form df using the musical isomorphism

$$\sharp = \sharp^g \colon T^*M \to TM$$

(called "sharp") defined by the metric g. The relation between the exterior derivative and the gradient of a function on $\mathbf{R}^n$ is a special case of this in which the metric is the flat metric given by the dot product.

Cylindrical and spherical coordinates

In cylindrical coordinates, the gradient is given by (Schey 1992, pp. 139–142):

$$\nabla f(\rho, \phi, z) = \frac{\partial f}{\partial \rho}\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial f}{\partial \phi}\mathbf{e}_\phi + \frac{\partial f}{\partial z}\mathbf{e}_z$$

where ϕ is the azimuthal angle, z is the axial coordinate, and $\mathbf{e}_\rho$, $\mathbf{e}_\varphi$ and $\mathbf{e}_z$ are unit vectors pointing along the coordinate directions.

In spherical coordinates (Schey 1992, pp. 139–142):

$$\nabla f(r, \theta, \phi) = \frac{\partial f}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\mathbf{e}_\phi$$

where ϕ is the azimuth angle and θ is the zenith angle.

For the gradient in other orthogonal coordinate systems, see Orthogonal coordinates (Differential operators in three dimensions).

Gradient of a vector

In rectangular coordinates, the gradient of a vector field **f** = (f_1, f_2, f_3) is defined by

$$\nabla \mathbf{f} = g^{jk}\frac{\partial f^i}{\partial x_j}\mathbf{e}_i\mathbf{e}_k$$

where the Einstein summation notation is used and the product of the vectors $\mathbf{e}_i$, $\mathbf{e}_k$ is a tensor of type (2,0), or the Jacobian matrix

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial (f_1, f_2, f_3)}{\partial (x_1, x_2, x_3)}.$$

In curvilinear coordinates, or more generally on a curved manifold, the gradient involves Christoffel symbols:

$$\nabla \mathbf{f} = g^{jk}\left(\frac{\partial f^i}{\partial x_j} + {\Gamma^i}_{jl} f^l\right)\mathbf{e}_i\mathbf{e}_k$$

where g^{jk} are the components of the metric tensor and the $\mathbf{e}_i$ are the coordinate vectors.

Expressed more invariantly, the gradient of a vector field **f** can be defined by the Levi-Civita connection and metric tensor:[1]

$$\nabla^a \mathbf{f}^b = g^{ac}\nabla_c \mathbf{f}^b$$

where ∇_c is the connection.

See also

- Curl
- Del
- Divergence
- Gradient theorem
- Hessian matrix
- Skew gradient

Notes

1. ^ Dubrovin, B. A.; Fomenko, A. T.; Novikov, S. P. (1991). *Modern Geometry--Methods and Applications: Part I: The Geometry of Surfaces, Transformation Groups, and Fields (Graduate Texts in Mathematics Vol. 93)* (2nd ed.). Springer. pp. 348–349. ISBN 978-0-387-97663-1.

References

- Korn, Theresa M.; Korn, Granino Arthur (2000), *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*, New York: Dover Publications, pp. 157–160, ISBN 0-486-41147-8, OCLC 43864234 (https://www.worldcat.org/oclc/43864234).
- Schey, H.M. (1992), *Div, Grad, Curl, and All That* (2nd ed.), W.W. Norton, ISBN 0-393-96251-2, OCLC 25048561 (https://www.worldcat.org/oclc/25048561).
- Dubrovin, B.A.; A.T. Fomenko, S.P. Novikov (1991), *Modern Geometry--Methods and Applications: Part I: The Geometry of Surfaces, Transformation Groups, and Fields (Graduate Texts in Mathematics)* (2nd ed.), Springer, pp. 14–17, ISBN 978-0-387-97663-1

External links

- "Gradient 1" (https://www.khanacademy.org/math/multivariable-calculus/partial_derivatives_topic/gradient/v/gradient-1) – Khan Academy
- Kuptsov, L.P. (2001), "Gradient" (http://www.encyclopediaofmath.org/index.php?title=G/g044680), in Hazewinkel, Michiel, *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Weisstein, Eric W., "Gradient" (http://mathworld.wolfram.com/Gradient.html), *MathWorld*.

Look up ***gradient*** in Wiktionary, the free dictionary.

Categories: Differential calculus | Generalizations of the derivative | Linear operators in calculus | Vector calculus