



planetmath.org

Math for the people, by the people.

gradient

Canonical name	Gradient
Date of creation	2013-03-22 11:59:20
Last modified on	2013-03-22 11:59:20
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	20
Author	CWoo (3771)
Entry type	Definition
Classification	msc 58A10
Classification	msc 26B12
Classification	msc 26B10
Related topic	Derivative2
Related topic	RiemannianMetric
Related topic	GradientInCurvilinearCoordinates
Related topic	NablaNabla
Related topic	DifferentialForms
Related topic	FirstOrderOperatorsInRiemannianGeometry
Related topic	VectorField
Related topic	HessianMatrix
Related topic	DerivativeNotation
Related topic	JacobianMatrix
Related topic	PartialDerivative
Related topic	TiltCurve
Defines	gradient

Summary. The *gradient* is a first-order differential operator that maps scalar functions to vector fields. It is a generalization of the ordinary derivative, and as such conveys information about the rate of change of a function relative to small variations in the independent variables. The gradient of a function f is customarily denoted by ∇f or by $\text{grad } f$.

Contents

1 Definition: Euclidean space

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be continuously differentiable. The *gradient* of f , denoted by ∇f , is defined by the property:

$$D_{\mathbf{v}} f = \nabla f \cdot \mathbf{v} \quad \text{for all vectors } \mathbf{v} \in \mathbb{R}^n. \quad (1)$$

The middle dot is the dot product, and $D_{\mathbf{v}}$ is the directional derivative with respect to \mathbf{v} .

If x^1, \dots, x^n are Euclidean coordinates, corresponding to the orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$, then

$$\nabla f = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \mathbf{e}_i. \quad (2)$$

The formula (??) is sometimes given as the definition of ∇f . We prefer to define ∇f by the coordinate-free formula (??) instead, because then the geometric interpretations (see below) become obvious, and (??) also indicates how we would go about calculating the gradient in other curvilinear coordinate systems. Formula (??) also makes it clear that the gradient is a physical vector, depending only on the inner product structure of \mathbb{R}^n , and not on the specific coordinate system used to calculate it.

There is the issue of whether the ∇f as defined by (??) exists; but this is proved easily enough, by substituting the concrete expression (??) and seeing that it satisfies (??).

The gradient can be considered to be a vector-valued differential operator, written as

$$\nabla = \sum_{i=1}^n \mathbf{e}_i \frac{\partial}{\partial x^i},$$

or, in the context of Euclidean 3-space, as

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

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the unit vectors lying along the positive direction of the x, y, z axes, respectively.

2 Geometric and physical interpretations

- (a) The direction of the vector ∇f is the direction of the greatest positive change, or increase, in f . The magnitude of ∇f is the magnitude of this increase. This follows immediately from (??):

$$D_{\mathbf{v}} f = \nabla f \cdot \mathbf{v} = \|\nabla f\| \|\mathbf{v}\| \cos \theta,$$

where θ is the angle between ∇f and \mathbf{v} . So among all unit directions \mathbf{v} of change, if \mathbf{v} is perpendicular to ∇f then the change $D_{\mathbf{v}} f$ is zero; if \mathbf{v} is parallel to ∇f then the change is maximized.

Similarly, $-\nabla f$ is the direction of the greatest negative change, or decrease, in f .

- (b) If M is the hypersurface in \mathbb{R}^n defined by

$$M = \{p \in \mathbb{R}^n : f(p) = 0, Df(p) \neq 0\},$$

then $\nabla f(p)$ is the normal to the hypersurface M at the point p . For $\ker Df(p)$ is the tangent space $T_p M$ to M at p , that is, $D_{\mathbf{v}} f(p) = 0$ for all $\mathbf{v} \in T_p M$, and by definition (??), $\nabla f(p)$ must be perpendicular to all $\mathbf{v} \in T_p M$.

Note that $Df \neq 0$ is equivalent to $\nabla f \neq 0$. Consequently, ∇f also gives an orientation to the hypersurface M .

For example, if $f(\mathbf{x}) = \|\mathbf{x}\| - 1$ for $\mathbf{x} \in \mathbb{R}^n$, M is the $(n-1)$ -dimensional sphere of unit radius, embedded in \mathbb{R}^n . Its normal, $\nabla f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$, as one would expect, points outward radially.

- (c) As a simple case of (b), consider the surface $z = f(x, y)$ in \mathbb{R}^3 , with Cartesian coordinates (x, y, z) . Think of this surface as describing a

hill, with height z . Then the direction of the gradient vector ∇f is the direction of steepest ascent of the hill, while its magnitude

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

is the slope or steepness in that direction.

If a ball is placed on the hill at a point (x, y, z) , theoretically it should roll down the hill in the direction of the gradient vector $-\nabla f(x, y)$. This may be easily derived by considering the mechanical forces on the ball. The direction of $-\nabla f(x, y)$ is, in fact, the projection to the xy -plane of an outward normal vector to the hill at (x, y, z) ; the normal vector is involved because the movement of the ball arises from the normal force from the hill.

- (d) Suppose the surface $z = f(x, y)$ in (c) describes a bowl instead of a hill, and we place a marble at any point (x, y, z) on this bowl. We would expect the marble to roll down to a local minimum point of $f(x, y)$. Since the marble should roll down in the direction of $-\nabla f$, we might hope that we can *find* local minima of a *given function* f by following the path mapped out by the gradients $-\nabla f$. Formally, this method of finding local extrema (with some modifications) is called *gradient descent*.
- (e) If U is the potential function corresponding to a conservative physical force, then $\mathbf{F} = -\nabla U$ is the corresponding force field.

Consequently, the gradient theorem,

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = - \int_{\gamma} \nabla U \cdot d\mathbf{s} = -U(\gamma(b)) + U(\gamma(a)), \quad \gamma: [a, b] \rightarrow \mathbb{R}^3$$

simply gives the formula for the change in the potential energy U when an object “does work” along a path γ in a conservative force field \mathbf{F} .

3 Definition: Riemannian geometry

It is obvious how (??) can be generalized to the setting of Riemannian manifolds: the dot product of \mathbb{R}^n must be replaced by the Riemannian metric, and the analogue of $D_{\mathbf{v}} f$ is the directional derivative $\mathbf{v}[f]$, for tangent vectors \mathbf{v}

on the Riemannian manifold. Thus for a smooth scalar-valued function f on a Riemannian manifold,

$$\mathbf{X} = \text{grad } f \Leftrightarrow df_p(\mathbf{v}) = \mathbf{v}[f] = \langle \mathbf{X}, \mathbf{v} \rangle_p. \quad (3)$$

We can calculate \mathbf{X} explicitly as follows. If x^i are local coordinates on the manifold (not necessarily orthonormal), set $\mathbf{X} = X^i \frac{\partial}{\partial x^i}$ (the Einstein summation convention is being used). Let g_{ij} and g^{ij} be the covariant and contravariant metric tensors, respectively. Then from (??),

$$\frac{\partial f}{\partial x^j} = \left\langle X^i \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = X^i \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = g_{ij} X^i,$$

and taking inverses,

$$X^i = g^{ij} \frac{\partial f}{\partial x^j}. \quad (4)$$

4 Duality with differential one-forms

Definitions (??) and (??) exhibit ∇f as the vector field dual to the differential form df . The isomorphism is given by applying the inner product or Riemannian metric. This isomorphism is, of course, linear; in particular it leads to the identity

$$\nabla f = \frac{\partial f}{\partial x^i} \nabla x^i, \quad (5)$$

which is the dual to the standard formula of differential one-forms:

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

Using (??) and (??), we have

$$\nabla x^i = g^{ij} \frac{\partial}{\partial x^j}, \quad \frac{\partial}{\partial x^i} = g_{ij} \nabla x^j. \quad (6)$$

So the isomorphism between vector fields and one-forms is expressed by changing the ∇ 's in (??) to d 's, and vice versa. That is,

$$dx^i \leftrightarrow g^{ij} \frac{\partial}{\partial x^j}, \quad \frac{\partial}{\partial x^i} \leftrightarrow g_{ij} dx^j. \quad (7)$$

It is commonly said that this isomorphism is expressed by “raising and lowering the indices of a tensor field, using contractions with g_{ij} and g^{ij} ”.

Notice that when x^i are orthonormal coordinates on \mathbb{R}^n , equation (??) reduces to equation (??), because $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ (Kronecker delta).

The formulae presented in this section are useful in the Euclidean setting as well, for deriving the formulae for the <http://planetmath.org/GradientInCurvilinearCoordinates> in various curvilinear coordinate systems.

5 Differential identities

Several properties of the one-dimensional derivative generalize to a multi-dimensional setting

$\nabla(af + bg) = a\nabla f + b\nabla g$	Linearity
$\nabla(fg) = f\nabla g + g\nabla f$	Product rule
$\nabla(f \circ \phi)(p) = D\phi(p)^* \nabla f(\phi(p))$	Chain rule
$\nabla(h \circ f)(p) = h'(f(p)) \nabla f(p)$	Another Chain rule

The function h is $h: \mathbb{R} \rightarrow \mathbb{R}$. The notation $(D\phi)^*$ denotes the transpose of the Jacobian matrix, in Euclidean coordinates, of $\phi: \mathbb{R}^m \rightarrow \mathbb{R}^n$. In the abstract setting, $(D\phi)^*$ is the adjoint to the tangent map $D\phi$ between the tangent bundles of two Riemannian manifolds.

These identities can be proved directly from the definition, but the first three are really just the duals of the following well-known identities for differential forms:

$$\begin{aligned} d(af + bg) &= adf + bdf \\ d(fg) &= fdf + gdf \\ d(\phi^* f) &= \phi^* df \end{aligned}$$

and so may be derived by changing the d 's here to ∇ 's! (Though the third identity may take a bit of thought.)

The following identity

$$\text{curl grad } f = \nabla \times \nabla f = 0,$$

is a special case of the differential forms identity $d^2 = 0$. Conversely, if $\text{curl } g = 0$ on a simply connected domain, then there is f such that $g = \text{grad } f$. See laminar field for details.

6 The ∇ symbolism

(This discussion does not really belong here, but should be moved to the nabla entry.)

Using the ∇ formalism, the divergence operator can be expressed as $\nabla \cdot$, the curl operator as $\nabla \times$, and the Laplacian operator as ∇^2 . To wit, for a given vector field

$$\mathbf{A} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k},$$

and a given function f we have

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \\ \nabla \times \mathbf{A} &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \mathbf{k} \\ \nabla^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.\end{aligned}$$

References

- [1] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, Volume I. Publish or Perish, 1979.