The fundamental theorem for line integrals

If $F = \nabla f$, then the integrand in the line integral

$$\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$$

is

$$f_1(x, y, z)\, dx + f_2(x, y, z)\, dy + f_3(x, y, z)\, dz, \qquad (*)$$

which the chain rule for several variables says is just $df(\mathbf{r}(t))$.

The fundamental theorem, part 2

The fundamental theorem of single variable calculus says that this integrates to the difference of the values of f at the points $\mathbf{r}(t)$ obtained from the endpoints of the interval in t over which you are integrating. These values of $\mathbf{r}(t)$ are just the endpoints of the arc $\mathcal{C}$. In particular, $f(x, y, z)$ can be determined up to an additive constant by integrating $\mathbf{F}$ along any path from a fixed base point (x_0, y_0, z_0) to (x, y, z).

Caution

Most integrals are not independent of path

Most integrals are not independent of path

Most integrals are not independent of path

Most integrals are not independent of path

Inverting the gradient

If a vector field $\mathbf{F}$ satisfies $\mathbf{F} = \nabla f$ for some **scalar function** f, the fundamental theorem suggests the f can be given by integrating $\mathbf{F}$. Indeed,

$$f(x, y, z) = \int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$$

where $\mathcal{C}$ is **any curve** joining a **base point** (x_0, y_0, z_0) to the general point (x, y, z).

A convenient path is one that follows a line parallel to the x axis, then a line parallel to the y axis, and finally a line parallel to the z axis.

Indefinite integration

This process of integrating with respect to one coordinate at a time suggests the process for finding integrals by guessing a function having the right derivative. If

$$\mathbf{F} = \langle P, Q, R \rangle,$$

the equation $\nabla f = \mathbf{F}$ requires

$$\frac{\partial f}{\partial x} = P(x, y, z), \ \frac{\partial f}{\partial y} = Q(x, y, z), \ \frac{\partial f}{\partial z} = R(x, y, z)$$

Indefinite integration, part 2

If we solve the first of these equations by integrating $P(x, y, z)$ with respect to x treating y and z as constants (a process that might be called **partial integration**), f is determined up to a "constant" that may depend on the quantities y and z that were treated as constants. Move on to the second equation by differentiating our current expression for f with respect to y and equating the result to $Q(x, y, z)$. Solve this for the derivative with respect to y of the "constant of integration" (depending on y and z) introduced in the previous step.

Indefinite integration, part 3

If this expression depends on x, the equation can't be solved, so $\mathbf{F}$ is not a gradient. Otherwise, integrate with respect to y treating z as a constant. Now $f(x, y, z)$ is known up to a function of z. Differentiate the current expression for f with respect to z and equate to $R(x, y, z)$. Simplify the result to get an equation saying that the derivative with respect to z of the second "constant of integration" is a known expression. This can be solved if the expression depends only on z.

An example

The vector field

$$\mathbf{F} = \langle\, yz\,,\ xz\,,\ xy + 2z\,\rangle$$

appears in exercise 15 of section 16.3. The process sketched in the previous slides says to begin by solving

$$\frac{\partial f}{\partial x} = yz$$

to obtain $f(x, y, z) = xyz + g(y, z)$. Equating the partial derivative of f with respect to y to the second component of $\mathbf{F}$ gives

$$xz + \frac{\partial g}{\partial y} = xz, \ \text{or}\ \frac{\partial g}{\partial y} = 0.$$

An example, part 2

Thus $g(y, z)$ is a function of only z. Equating the partial derivative of f with respect to z to the third component of $\mathbf{F}$ gives

$$xy + g'(z) = xy + 2z, \text{ or } g'(z) = 2z$$

Thus, $g(z) = z^2 + C$ and $f(x, y, z) = xyz + z^2 + C$

Another example

The vector field

$$\mathbf{F} = \langle x + y\,,\ 2x + y \rangle$$

is **not** a gradient. If $\nabla f = \mathbf{F}$, then

$$\frac{\partial f}{\partial x} = x + y,$$

so $f(x, y) = x^2/2 + xy + g(y)$. Then, $x + g'(y) = 2x + y$ or $g'(y) = x - y$, but $x - y$ is **not** independent of x

Maxima and minima, theory

Everyone's favorite problem of differential calculus is the determination of maxima and minima of a function on a region.

An important theoretical result asserts that, if the function is **continuous** and the region **closed** and **bounded** in some $\mathbb{R}^n$, then these extreme values exist and are attained at points of the domain.

Maxima and minima, method

To apply the methods of one-variable calculus to functions of several variables, we need the following

Secret Weapon. *If $A \subseteq B$ and if the maximum of a function f on B is taken on at a point $x \in A$, then $f(x)$ is also the maximum of f on A.*

Proof. Think about it!

Preparing to narrow the search

In applications, B will be the given region in $\mathbb{R}^n$ and A will be a curve lying in B. Composing the given function on B with the parameterization of A gives a real valued function of a real variable, and one-dimensional calculus applies.

You may protest that we don't know A, but that will be dealt with.

Narrowing the search

What one-variable calculus gives us is a **necessary condition** for a point to be a maximum or a minimum. If we can find **any** curve A through a point P for which this condition fails at P, then $f(P)$ is not an extreme value of f on B and we can look somewhere else.

It turns out that points on the boundary of B behave differently than interior points. Interior points are easier, so we do them first.

The key special case

What then do we take as A? The simplest examples turn out to suffice: let A be a line segment parallel to one of the coordinate axes and lying in B. Every interior point has segments through it parallel to each coordinate axis. The coordinate on the axis can be taken as the parameter on the segment. The derivative with respect to this parameter is exactly the partial derivative of the given function with respect to the selected coordinate.

The key special case, part 2

The **secret weapon** then implies

Theorem. *If f takes on its maximum (or minimum) value on the set B at an interior point P, then all partial derivatives of f are zero at P.*

Since the partial derivatives give all components of the gradient, such points P have $\nabla f(P) = 0$.

Boundary curves and vertices

If a portion of the boundary is a curve that can be parameterized, composing f with that parameterization produces a function of one variable that describes the behavior of f on this portion of the boundary. In particular, any extreme value of f on this curve leads to an extreme value of the composition. However, the composition can be studied by methods of single-variable calculus, so we can usually restrict attention to only a finite number of possible locations of extrema.

Corners are another matter. They are the endpoints of any parameterized curves on the boundary, so they need to be treated separately — just like endpoints in the single variable case.

Exercises 14.7

We just look at exercises that ask for maxima and minima of $f(x, y)$ on a closed bounded set $\mathcal{D}$.

#27. $f(x, y) = 5 - 3x + 4y$. $\mathcal{D}$ is the triangle with vertices (0, 0), (4, 0), (4, 5).

#31. $f(x, y) = 1 + xy - x - y$. $\mathcal{D}$ is the region bounded by $y = x^2$ and $y = 4$.

#33. $f(x, y) = 2x^3 + y^4$. $\mathcal{D}$ is the circular disk where $x^2 + y^2 \leq 1$.

The role of calculus

Calculus improves the existence theorem: if the function is **differentiable**, then the points at which the extreme values are taken on are **easy to find**. Note that this shifts the emphasis from the values of the function to the points in the domain where the function takes those values.

If a tangent plane to $z = f(x, y)$ at a point (x_0, y_0) is **not** horizontal, then it is easy to find a direction in which f takes larger values (it will take smaller values in the opposite direction). Thus — exactly as in the one-variable case — a maximum or minimum at an **interior point** of a domain can only occur at a point where all partial derivatives are zero.

The role of multivariable calculus

Since the partial derivatives are the components of the gradient, it is equivalent to have $\nabla f = \mathbf{0}$.

A function cannot take a maximum or minimum at an **interior point** of a domain that is a saddle point. A fairly complicated **second derivative test** is necessary in order to automatically exclude saddle points. Usually there aren't many possible extrema, so it may be better to include them in the list.

We now develop an approach to extrema on the boundary.

Lagrange multipliers, part 1

To study the function restricted to the boundary, B, the familiar method of assuming an explicit form and describing its properties in terms of the implicit definition will be used.

Thus, suppose there is a vector function $\mathbf{r}(t)$ whose domain contains a neighborhood of 0 and whose range lies entirely in B, with $\mathbf{r}(0) = P$, and $\mathbf{r}'(0) \neq \mathbf{0}$. The last part is the **smoothness** condition. Then the restriction of f to B can be represented as $f(\mathbf{r}(t))$, and the derivative with respect to t is $\nabla f \cdot \mathbf{r}'(t)$. At any max or min on this curve, we must then have that $\mathbf{r}'(t)$ is perpendicular to ∇f.

Lagrange multipliers, part 2

However, B has been given in the form

$$\{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}.$$

Hence, the composition $g(\mathbf{r}(t))$ is a constant function, so must have derivative zero. The chain rule, evaluated at $t = 0$, gives

$$\nabla g(P) \cdot \mathbf{r}'(0) = 0.$$

This says that $\mathbf{r}'(0)$ must be perpendicular to $\nabla g(P)$. However, we have already seen that

$$\nabla f(P) \cdot \mathbf{r}'(0) = 0.$$

Lagrange multipliers, part 3

For points at which the boundary is smooth, this forces the existence of a value traditionally called λ for which $\nabla f(P) = \lambda \nabla g(P)$. In $\mathbb{R}^2$ this gives 2 equations in λ and the 2 coordinate functions. The condition, $g(P) = 0$ gives one more equation. We have as many equations as we have variables, so we should be able to solve this system of equations. Unfortunately, the algebra for doing this is sometimes difficult. A systematic approach to this algebra is possible, but we don't have time to develop one. One practical consequence of this is that only simple examples will be considered.

Examples of Lagrange multipliers, part 1

There is one easy application of Lagrange multipliers that appears in many variants. There are two forms that reduce to the same algebra:

(1) Let x and y be nonnegative with $x + y = s$ (s constant). Find the maximum of xy.

(2) Let x and y be positive with $xy = p$ (p constant). Find the minimum of $x + y$.

Examples of Lagrange multipliers, part 2

Since our notation for gradients uses functions rather than expressions, introduce the functions $S(x, y) = x + y$ and $P(x, y) = xy$. Then the Lagrange multiplier method for either problem (1) or problem (2) says that extreme values occur when $\nabla S \parallel \nabla P$. Since $\nabla S = \langle 1, 1 \rangle$ and $\nabla P = \langle y, x \rangle$, this says $x = y$. In (1), the given constraint then gives $x = y = s/2$ and $xy = s^2/4$. In (2), we get $x = y = \sqrt{p}$ and $x + y = 2\sqrt{p}$.

Examples of Lagrange multipliers, part 3

In each problem, the method selects a unique point. The only other candidates for the location of extreme values of the function are the endpoints. In (1), the product is zero at both endpoints; in (2), there are no true endpoints, but $x + y \to \infty$ outside bounded parts of the curve. This explains how we knew that the interior point gave a maximum in (1) and a minimum in (2).

Higher dimensions

In higher dimensions, there are two new difficulties. First, the set B may have dimension greater than 1, so that it cannot be described by a single real parameter. This is a theoretical restriction that is met by considering all smooth curves through P that lie on B. For sets of **codimension** 1 described by $g(P) = 0$, this leads to $\nabla f \parallel \nabla g$ as before. Second, it may be necessary to consider B of lower dimension. In this case, there is more than one g and the condition is the ∇f be a linear combination of all ∇g. This requires more multipliers λ_i and a different mix of equations.

Exercises

14.8#5. $f(x, y) = x^2y$ on $x^2 + 2y^2 = 6$.

14.8#17. $f(x, y, z) = yz + xy$ on intersection of $xy = 1$ and $y^2 + z^2 = 1$.

and the examples from 14.7 discussed previously

$f(x, y) = 5 - 3x + 4y$. $\mathcal{D}$ is the triangle with vertices $(0, 0)$, $(4, 0)$, $(4, 5)$.

$f(x, y) = 1 + xy - x - y$. $\mathcal{D}$ is the region bounded by $y = x^2$ and $y = 4$.

$f(x, y) = 2x^3 + y^4$. $\mathcal{D}$ is the circular disk where $x^2 + y^2 \leq 1$.