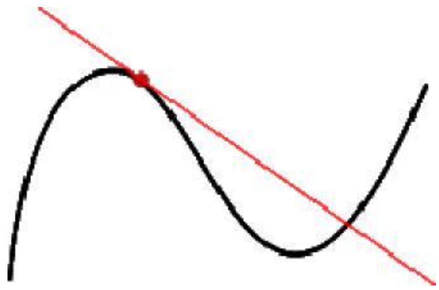
MODULE 3 TOTAL DERIVATIVES OF FUNCTION OF SEVERAL VARIABLES

UNIT 1 DERIVATIVE

CONTENTS

1.0 INTRODUCTION

This article is an overview o the term as used in calculus. For a less technical overview of the subject, see Differential calculus. For other uses, see Derivative (disambiguation).



The graph of a function, drawn in black, and a tangent line to that function, drawn in red. The slope of the tangent line is equal to the derivative of the function at the ma ked point.

In calculus, a branch of mathematics, the **derivative** is a measure of how a function changes as its input changes. Loosely speaking, a derivative can be thought of as how much one quantity is changing in response to changes in some other quantity; for example, the derivative of the position of a moving object with respect to time is the object's instantaneous velocity.

The derivative of a function at a chosen input value describes the best linear approximation of the function near that input value. For a real-valued function of a single real variable, the derivative at a point equals the slope of the tangent line to the graph of the function at that point. In higher dimensions, the derivative of a function at a point is a linear transformation called the linearization. A closely related notion is the differential of a function.

The process of finding a derivative is called **differentiation**. The reverse process is called **anti-differentiation**. The fundamental theorem of calculus states that anti-differentiation is thesame as integration. Differentiation and integration constitute the two fundamental operations in single-variable calculus.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- Solve directional derivatives;
- Use derivative to solve total derivative, total differential; and Jacobian matrix.

3.0 MAIN CONTENT

Directional derivatives

If f is a real-valued function on $\mathbf{R}^n$, then the partial derivatives of f measure its variation in the direction of the coordinate axes. For example, if f is a function of x and y, then its partial derivatives measure the variation in f in the x direction and the y direction. They do not, however, directly measure the variation of f in any other direction, such as along the diagonal line $y = x$. These are measured using directional derivatives. Choose a vector

$$\mathbf{v} = (v_1, \ldots, v_n).$$

The **directional derivative** of f in the direction of **v** at the point **x** is the limit

$$D_{\mathbf{v}} f(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}.$$

In some cases it may be easier to compute or estimate the directional derivative after changing the length of the vector. Often this is done to turn the problem into the computation of a directional derivative in the direction of a unit vector. To see how this works, suppose that $v = \lambda u$. Substitute $h = k/\lambda$ into the difference quotient. The difference quotient becomes:

$$\frac{f(\mathbf{x} + (k/\lambda)(\lambda\mathbf{u})) - f(\mathbf{x})}{k/\lambda} = \lambda \cdot \frac{f(\mathbf{x} + k\mathbf{u}) - f(\mathbf{x})}{k}.$$

This is λ times the difference quotient for the directional derivative of Furthermore, taking the limit as h tends to zero is the same as taking zero because h and k are multiples of each other. Therefore Dv(f) = λDu (rescaling property, directional derivatives are frequently considered only for unit vector.

If all the partial derivatives of f exist and are continuous at **x**, then they determine the directional derivative of f in the direction v by the formula:

$$D_{\mathbf{v}} f(\boldsymbol{x}) = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}.$$

This is a consequence of the definition of the total derivative. It follows that the directional derivative is linear in v, meaning that Dv + w(f) = Dv(f) + Dw(f).

The same definition also works when f is a function with values in R^m. The above definition is applied to each component of the vectors. In this case, the directional derivative is a vector in R^m.

Total derivative, total differential and Jacobian matrix

When f is a function from an open subset of $\mathbf{R}^n$ to $\mathbf{R}^m$, then the directional derivative of f in a chosen direction is the best linear approximation to f at that point and in that direction. But when $n > 1$, no single directional derivative can give a complete picture of the behavior of f. The total derivative, also called the (**total**) **differential**, gives a complete picture by considering all directions at once. That is, for any vector **v** starting at **a**, the linear approximation formula holds:

$$f(\mathbf{a} + \mathbf{v}) \approx f(\mathbf{a}) + f'(\mathbf{a})\mathbf{v}.$$

Just like the single-variable derivative, $f\ '$(**a**) is chosen so that the error in this approximation is as small as possible.

If n and m are both one, then the derivative $f\ '$(**a**) is a number and the expression $f\ '$(**a**)**v** is the product of two numbers. But in higher dimensions, it is impossible for $f\ '$(**a**) to be a number. If it were a number , then $f\ '$(**a**)**v** would be a vector in $\mathbf{R}^n$while the other terms would be vectors in $\mathbf{R}^m$, an therefore the formula would not make sense. For the linear approximation formula to ma ke sense, $f\ '$(**a**) must be a function that sends vectors in R^n to vectors in R^m, and $f\ '$**(a)v** must denote this function evaluated at v.

To determine what kind of function it is, notice that the linearapproximation formula can berewritten as

$$f(\mathbf{a} + \mathbf{v}) - f(\mathbf{a}) \approx f'(\mathbf{a})\mathbf{v}.$$

Notice that if we choose another vector w, then this approximate equation determines another approximate equation by substituting **w** for **v**. It determines a third approximate equation by substituting both **w** for **v** and **a** + **v** for a. By subtracting these two new equations, we get

$$f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}+\mathbf{w}) + f(\mathbf{a}) \approx f'(\mathbf{a}+\mathbf{v})\mathbf{w} - f'(\mathbf{a})\mathbf{w}.$$

If we assume that **v** is small and that the derivative varies continuously in **a**, then $f'(\mathbf{a}+\mathbf{v})$ is approximately equal to $f'(\mathbf{a})$, and therefore the right-hand side is approximately zero. The left-hand side can be rewritten in a different way using the linear approximation formula with v + w substituted for v. The linear approximation formula implies:

$$\begin{aligned} 0 &\approx f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}+\mathbf{w}) + f(\mathbf{a}) \\ &= (f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a})) - (f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a})) - (f(\mathbf{a}+\mathbf{w}) - f(\mathbf{a})) \\ &\approx f'(\mathbf{a})(\mathbf{v}+\mathbf{w}) - f'(\mathbf{a})\mathbf{v} - f'(\mathbf{a})\mathbf{w}. \end{aligned}$$

This suggests that $f'(\mathbf{a})$ is a linear transformation from the vector space R^n to the vector space R^m. In fact, it is possible to make this a precise derivation by measuring the error in the approximations. Assume that the error in these linear approximation formula is bounded by a constant times ||v||, where the constant is independent of **v** but depends continuously on a. Then, after adding an appropriate error term, all of the above approximate equalities can be rephrased as inequalities. In particular, $f'(\mathbf{a})$ is a linear transformation up to a small error term. In the limit as **v** and **w** tend to zero, it must therefore be a linear transformation. Since we define the total derivative by taking a limit as v goes to zero, $f'(\mathbf{a})$ must be a linear transformation.

In one variable, the fact that the derivative is the best linear approximation is expressed by the fact that it is the limit of difference quotients. However, the usual difference quotient does not make sense in higher dimensions because it is not usually possible to divide vectors. In particular, the numerator and denominator of the difference quotient are not even in the same vector space: The numerator lies in the co domain $\mathbf{R}^m$ while the denominator lies in the domain $\mathbf{R}^n$. Furthermore, the derivative is a linear transformation, a different type of object from both the numerator and denominator. To make precise the idea that $f'(\mathbf{a})$ is the best linear approximation, it is necessary to adapt a different formula for the one-variable derivative in which these problems disappear. If f: $\mathbf{R} \to \mathbf{R}$, then the usual definition of the derivative may be manipulated to show that the derivative of f at a is the unique number $f'(a)$ such that

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0.$$

This is equivalent to

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - f'(a)h|}{|h|} = 0$$

because the limit of a function tends to zero if and only if the limit of the absolute value of the function tends to zero. This last formula can be adapted to the many-variable situation by replacing the absolute values with norms.

The definition of the total derivative of f at a, therefore, is that it is the unique linear transformation $f'(a) : \mathbf{R}^n - \mathbf{R}^m$ such that

$$\lim_{\|\mathbf{h}\|\to 0} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - f'(\mathbf{a})\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Here **h** is a vector in $\mathbf{R}^n$, s the norm in the denominator is the standard length on $\mathbf{R}^n$. However, $f'(\mathbf{a})\mathbf{h}$ is a vector in $\mathbf{R}^m$, and the norm in the numerator is the standard length on $\mathbf{R}^m$. If v is a vector starting at a, then $f'(\mathbf{a})\mathbf{v}$ is called the push forward of **v** by f and is sometimes written fv.

If the total derivative exists at a, then all the partial derivatives and directional derivatives of f exist at a, and for all **v**, $f'(\mathbf{a})\mathbf{v}$ is the directional derivative of f in the direction **v**. If we write f using coordinate functions, so that $f = (f_1, f_2, ..., f_m)$, then the total derivative can be expressed using the partial derivatives as a matrix. This matrix is called the Jacobian matrix of f at **a**:

$$f'(\mathbf{a}) = \mathrm{Jac}_{\mathbf{a}} = \left(\frac{\partial f_i}{\partial x_j}\right)_{ij}.$$

The existence of the total derivative $f'(a)$ is strictly stronger than the existence of all the partial derivatives, but if the partial derivatives exist and are continuous, then the total derivative exists, is given by the Jacobian, and depends continuously on a.

The definition of the total derivative subsumes the definition of the derivative in one variable. That is, if f is a real-valued function of a real variable, then the total derivative exists if and only if the usual derivative exists. The Jacobian matrix reduces to a 1×1 matrix whose only entry is the derivative $f'(x)$. This 1×1 matrix satisfies the property that $f(a + h) - f(a) - f'(a)h$ is approximately zero , in other words that

$$f(a+h) \approx f(a) + f'(a)h.$$

Up to changing variables, this is the statement that the function $x \mapsto f(a) + f'(a)(x-a)$ is the best linear approximation to f at a.

The total derivative of a function does not give another function in the same way as the one-variable case. This is because the total derivative of a multivariable function has to record much more information than the derivative of a single-variable function. Instead, the total derivative gives a function from the tangent bundle of the source to the tangent bundle of the target.

The natural analog of second, third, and higher-order total derivatives is not a linear transformation, is not a function on the tangent bundle, and is not built by repeatedly taking the total derivative. The analog of a higher-order derivative, called a jet, cannot be a linear transformation because higher-order derivatives reflect subtle geometric information, such as concavity, which cannot be described in terms of linear data such as vectors. It cannot be a function on the tangent bundle because the tangent bundle only has room for the base space and the directional derivative s. Because jets capture higher-order information, they take as arguments additional coordinates representing higher-order changes in direction. The space determined by these additional coordinates is called the jet bundle. The relation between the total derivative and the partial derivatives of a function is paralleled in the relation between the kth order jet of a function and its partial derivatives of order less than or equal to k.

4.0 CONCLUSION

In this unit, you have used derivative to solve problems on directional derivatives and have also solve problems on total derivative, total differentiation and Jacobian matrix.

5.0 SUMMARY

In this unit you have studied:

- Solve directional derivatives
- Use derivative to solve problems on total derivative, total differentiation d Jacobian matrix.

6.0 TUTOR-MARKED ASSIGNMENT

1. Evaluate the derivative of $F(x,y,z) = 3(x^3 + y) \sin (z^2)$
2. Find the derivative of $F(x,y,z) = xy^3 + z^4$
3. Let $F(x,y,z) = x^5 + y^4z^3 + \sin z^2$,find the derivative.
4. Evaluate the derivatives of $F(x,y,z) = x^2 - xy + z^4$
5. Find the derivative of $F(x,y,z) = \frac{\sin x + COS^2 X}{\tan \quad x}$

7.0 REFERENCES/FURTHER READING

Anton, Howard; Bivens, Irl; Davis, Stephen (February 2, 2005), Calculus: Early Transcendentals Single and Multivariable (8th ed.), New York: Wiley, ISBN 978-0-47147244-5

Apostol, Tom M. (June 1967), Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra, 1 (2nd ed.), Wiley, ISBN 978-0-471-00005-1

Apostol, Tom M. (June 1969), Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications, 1 (2nd ed.), Wiley, ISBN 978-0-471-00007-5

Courant, Richard; John, Fritz (December 22, 1998), Introduction to Calculus and Analysis, Vol. 1, Springer-Verlag, ISBN 978-3-540-65058-4

Eves, Howard (January 2, 1990), An Introduction to the History of Mathematics (6th ed.), Brooks Cole, ISBN 978-0-03-029558-4

Larson, Ron; Hostetler, Robert P.; Edwards, Bruce H. (February 28, 2006), Calculus: Early Transcendental Functions (4th ed.), Houghton Mifflin Company, ISBN 978-0-61860624-5

Spivak, Michael (September 1994), Calculus (3rd ed.), Publish or Perish, ISBN 978-0-914098-89-8

Stewart, James (December 24, 2002), Calculus (5th ed.), Brooks Cole, ISBN 978-0-534-39339-7

Thompson, Silvanus P. (September 8, 1998), Calculus Made Easy (Revised, Updated, Expanded ed.), New York: St. Martin's Press, ISBN 978-0-312-18548-0

UNIT 2 TOTAL DERIVATIVE

CONTENTS

1.0 INTRODUCTION

In the mathematical field of differential calculus, the term **total derivative** has a number of closely related meanings.

The total derivative (full derivative) of a function *f*, of several variables, e.g., *t*, *x*, *y*, etc., with respect to one of its input variables, e.g., *t*, is different from the partial derivative (∂).Calculation of the total derivative of *f* with respect to *t* does not assume that the other arguments are constant while *t* varies; instead, it allows the other arguments to depend on *t*. The total derivative adds in these *indirect dependencies* to find the overall dependency of *f* on *t*. For example, the total derivative of *f*(*t*,*x*,*y*) with respect to *t* is

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}.$$

Consider multiplying both sides of the equation by the differential :

$$\mathrm{d}f = \frac{\partial f}{\partial t}\mathrm{d}t + \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y.$$

The result will be the differential change ***df*** in the function *f*. Because *f* depends on *t*, some of that change will be due to the partial derivative of *f* with respect to *t*. However, some of that change will also be due to the partial derivatives of *f* with respect to the variables *x* and y. So, the differential ***dt*** is applied to the total derivatives of *x* and *y* to find differentials ***d****x* and ***dy,*** which can then be used to find the contribution to ***df.***

- It refers to a differential operator such as

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \sum_{j=1}^{k} \frac{dyj}{dx} \frac{\partial f}{\partial yj},$$

  which computes the total derivative of a function (with respect to *x* in this case).

- It refers to the (total) differential *df* of a function, either in the traditional language of infinitesimals or the modern language of differential forms.

- A differential of the form

$$\sum_{j=1}^{k} f_j(x_1, \ldots, x_k)\, \mathbf{d}\, x_j$$

  is called a total differential or an exact differential if it is the differential of a function. Again this can be interpreted infinitesimally, or by using differential forms and the exterior derivative.

- It is another name for the derivative as a linear map, i.e., if f is a differentiable function from $\mathbf{R}^n$ to $\mathbf{R}^m$, then the total derivative (or differential) of *f* at $\mathbf{x} \in \mathbf{R}^n$ is the linear map from $\mathbf{R}^n$ to $\mathbf{R}^m$ whose matrix is the Jacobian matrix of *f* at *x*.

- It is a synonym for the gradient, which is essentially the derivative of a function from $\mathbf{R}^n$

- It is sometimes used as a synonym for the material derivative, $\frac{D\mathbf{u}}{\mathrm{Dt}}$, in fluid mechanics.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- differentiate with indirect dependent;
- find the derivative via differentials;
- solve total derivative as a linear map; and
- explain total differential equation.

3.0 MAIN CONTENT

Differentiation with indirect dependencies

Suppose that *f* is a function of two variables, *x* and *y*. Normally these variables are assumed to be independent. However, in some situations they may be dependent on each other. For example *y* could be a function of *x*, constraining the domain of *f* to a curve in R^2. In this case the partial derivative of *f* with respect to *x* does not give the true rate of change of *f* with respect to changing *x* because changing *x* necessarily changes *y*. The **total derivative** takes such dependencies into account.

For example, suppose

$f(x,y) = xy.$

The rate of change of *f* with respect to x is usually the partial derivative of f with respect to x; in this case,

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

However, if y depends on x, the partial derivative does not give the true rate of change of *f* as *x* changes because it holds *y* fixed.

Suppose we are constrained *t* the line

$y = x$

then

$f(x,y) = f(x,x) = x^2.$

In that case, the total derivative of f with respect to x is

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 2x$$

Notice that this is not equal to the partial derivative:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 2x \neq \frac{\partial f}{\partial x} = y = x$$

While one can often perform substitutions to eliminate indirect dependencies, the chain rule provides for a more efficient and general technique. Suppose $M(t, p_1, ..., p_n)$ is a function of time t and n variables p_i which themselves depend on time. Then, the total time derivative of M is

$$\frac{\mathrm{d}\, M}{\mathrm{d}\, t} = \frac{\mathrm{d}}{\mathrm{d}\, t} M\big(t, p_1(t), \ldots, p_n(t)\big).$$

This expression is often used in physics for a gauge transformation of the Lagrangian, as two Lagrangians that differ only by the total time derivative of a function of time and t generalized coordinates lead to the same equations of motion. The operator in brackets (in the final expression) is also called the total derivative operator (with respect to t).

For example, the total derivative of f(x(t), y(t)) is

$$\frac{\mathrm{d}\, f}{\mathrm{d}\, t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}\, x}{\mathrm{d}\, t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}\, y}{\mathrm{d}\, t}.$$

Here there is no ∂f / ∂t term since f itself does not depend on the independent variable *t*directly

The total derivative via differentials

Differentials provide a simple way to understand the total derivative. For instance, suppose $\boldsymbol{M}(t, p_1, \ldots, p_n)$ is a function of time t and n variables p^i as in the previous section. Then, the differential of M is

$$\mathrm{d}\, M = \frac{\partial M}{\partial t}\mathrm{d}\, t + \sum_{i=1}^{n} \frac{\partial M}{\partial p_i}\mathrm{d}\, p_i.$$

This expression is often interpreted heuristically as a relation between infinitesimals.However, if the variables $\boldsymbol{t}$ and $\boldsymbol{p_l}$ are interpreted as functions, and $M(t, p_1, \ldots, p_n)$ interpreted to mean the composite of M with these functions, then the above expression makes perfect sense as an equality of differential 1-forms, and is immediate from the chain rule for the exterior derivative. The advantage of this point of view is that it takes into account arbitrary dependencies between the variables. For example, if $p_1^2 = p_2 p_3$ then $2p_1 dp_1 = p_3 dp_2 + p_2 dp_3$ in particular, if the variables p_l are all functions of t, as in the previous section, then

$$\mathrm{d}\, M = \frac{\partial M}{\partial t}\mathrm{d}\, t + \sum_{i=1}^{n} \frac{\partial M}{\partial p_i}\frac{\partial p_i}{\partial t}\,\mathrm{d}\, t.$$

The total derivative as a linear map

Let $U \subseteq \mathcal{R}^n$ be an open subset. Then a function $f: U \subseteq \mathbb{R}^n$ is said to be (**totally**) differentiable at a point $p \in U$, if there exists a linear map $df_p: \mathbb{R}^n \to \mathbb{R}^m$ denoted $D_p f$ or $Df(p)$) such that

$$\lim_{x \to p} \frac{\|f(x) - f(p) - \mathrm{d}\, f_p(x-p)\|}{\|x-p\|} = 0.$$

The linear map df_p is called the (total) derivative or (total) differential of f at p. A function is (totally) differentiable if its total derivative exists at every point in its domain.

Note that f is differentiable if and only if each of its components $f_i : U \to \mathbb{R}$ is differentiable. For this it is necessary, but not sufficient, that the partial derivatives of each function f_j exist. However, if these partial derivatives exist and are continuous, then f is differentiable and its differential at any point is the linear map determined by the Jacobian matrix of partial derivatives at that point.

Total differential equation

A total differential equation i s a differential equation expressed in terms of total derivatives. Since the exterior derivative is a natural operator, in a sense that can be given a technical meaning, such equations are intrinsic and *geometric*.

4.0 CONCLUSION

In this unit, you have know how to differentiate with indirect dependent. You have used total derivative via differentials and have known the total derivative as a linear map.

5.0 SUMMARY

In this unit, you have studied the following:

- Differentiation with indirect dependent
- The total derivative via differentials
- The total derivative as a linear map
- The total differential equation

6.0 TUTOR – MARK ASSIGNMENT

1. Find the total derivative for the second – order of the function
   $F(x,y,z) = x^3 + y^4 - z^3$
2. Find the total derivative for the function
   $F(x,y,z) = x^2 y^3 + z^3$
3. Solve the total derivative to the third - order of the function
   $F(x,y,z) = x^3 y^4 + x^2 y + y^3 x^4 z^4$

7.0 REFERENCES/FURTHER READING

A. D. Polyanin and V. F. Zaitsev, Handbook of Exact Solutions for Ordinary Differential Equations (2nd edition), Chapman & Hall/CRC Press, Boca Raton, 2003. ISBN 1-58488297-2

nton, Howard; Bivens, Irl; Davis, Stephen (February 2, 2005), Calculus: Early Transcendentals Single and Multivariable (8th ed.), New York: Wiley, ISBN 978-0-47147244-5

Apostol, Tom M. (June 1967), Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra, 1 (2nd ed.), Wiley, ISBN 978-0-471-00005-1

Apostol, Tom M. (June 1969), Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications, 1 (2nd ed.), Wiley, ISBN 978-0-471-00007-5

Courant, Richard; John, Fritz (December 22, 1998), Introduction to Calculus and Analysis, Vol. 1, Springer-Verlag, ISBN 978-3-540-65058-4

Eves, Howard (January 2, 1990), An Introduction to the History of Mathematics (6th ed.), Brooks Cole, ISBN 978-0-03-029558-4

Larson, Ron; Hostetler, Robert P.; Edwards, Bruce H. (February 28, 2006), Calculus: Early Transcendental Functions (4th ed.), Houghton Mifflin Company, ISBN 978-0-618-60624-5

Spivak, Michael (September 1994), Calculus (3rd ed.), Publish or Perish, ISBN 978-0-914098-89-8

Stewart, James (December 24, 2002), Calculus (5th ed.), Brooks Cole, ISBN 978-0-534-39339-7

Thompson, Silvanus P. (September 8, 1998), Calculus Made Easy (Revised, Updated, Expanded ed.), New York: St. Martin's Press, ISBN 978-0-312-18548-0

UNIT 3 APPLICATION OF TOTAL DERIVATIVE OF A FUNCTION

CONTENTS

1.0 INTRODUCTION

Let us consider a function

1) $u = f(x, y, z, p, q, ...)$

of several variables. Such a function can be studied by holding all variables except one constant and observing its variation with respect to one single selected variable. If we consider all the variables except x to be constant, then

$$\frac{du}{dx} = \frac{df(x, \hat{y}, \hat{z}, \hat{p}, \hat{q}, \dots)}{dx}$$

represents the partial derivative of $f(x, y, z, p, q, ...)$ with respect to x (the hats indicating variables held fixed). The variables held fixed are viewed as parameters.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- apply total derivative on chain rule for functions of functions;
- apply total derivative to find directional derivative;
- apply total derivative to solve differentiation under integral sign; and
- apply total derivative on lebnitz rule.

3.0 MAIN CONTENT

APPLICATION OF TOTAL DERIVATIVES

Chain rule for functions of functions

If w = f(x, y, z, ...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives $dw/dx\,,dw/dy\,,dw/dz$,... and if *x, y, z,* ... are differentiable functions $x = x(t)$,$y = y(t)$, $z = z(t)$, etc. of a variable t, then the total derivative of w with respect to t is given by

2) $$\frac{dw}{dt}=\frac{dw}{dx}\frac{dx}{dt}+\frac{dw}{dy}\frac{dy}{dt}+\frac{dw}{dz}\frac{dz}{dt}$$

This rule is called the chain rule for the partial derivatives of functions of functions.

Similarly, if w = f(x, y, z,...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives $dw/dx\,,dw/dy\,,dw/dz$,... and if x, y, z, ... are differentiable functions of m independent variables *r, s, t* ..., then

$$\frac{dw}{dr}=\frac{dw}{dx}\frac{dx}{dr}+\frac{dw}{dy}\frac{dy}{dr}+\frac{dw}{dz}\frac{dz}{dr}+\cdots$$

$$\frac{dw}{ds}=\frac{dw}{dx}\frac{dx}{ds}+\frac{dw}{dy}\frac{dy}{ds}+\frac{dw}{dz}\frac{dz}{ds}+\cdots \text{etc}$$

This rule is called the chain rule for the partial derivatives of functions of functions.

Similarly, if w = f(x, y, z, ...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives$dw/dx\,,dw/dy\,,dw/dz$,... and if *x, y, z,* ... are differentiable functions of m independent variables *r, s, t*..., then

$$\frac{dw}{dr}=\frac{dw}{dx}\frac{dx}{dr}+\frac{dw}{dy}\frac{dy}{dr}+\frac{dw}{dz}\frac{dz}{dr}+\cdots$$

$$\frac{dw}{ds}=\frac{dw}{dx}\frac{dx}{ds}+\frac{dw}{dy}\frac{dy}{ds}+\frac{dw}{dz}\frac{dz}{ds}+\cdots \text{etc}$$

Note the similarity between total differentials and total derivatives. The total derivative above can be obtained by dividing the total differential

$$dw=\frac{dw}{dx}dx+\frac{dw}{dy}dy+\frac{dw}{dz}+\cdots \text{by } dt$$

As a special application of the chain rule let us consider the relation defined by the two equations

$z = f(x, y)$; $y = g(x)$

Here, z is a function of x and y while y in turn is a function of x. Thus z is really a function of the single variable x. If we apply the chain rule we get

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}$$

which is the total derivative of z with respect to x.

Definition of Scalar point function

A scalar point function is a function that assigns a real number (i.e. a scalar) to each point of some region of space. If to each point (x, y, z) of a region R in space there is assigned a real number u = $\Phi(x, y, z)$, then Φ is called a scalar point function.

Examples. 1. The temperature distribution within somebody at a particular point in time. 2. The density distribution within some fluid at a particular point in time.

Directional derivatives. Let $\Phi(x, y, z)$ be a scalar point function defined over some region R of space. The function $\Phi(x, y, z)$ could, for example, represent the temperature distribution within some body. At some specified point $P(x, y, z)$ of R we wish to know the rate of change of Φ in a particular direction. The rate of change of a function Φ at a particular point P, in a specified direction, is called the directional derivative of Φat P in that direction. We specify the direction by supplying the direction angles or direction cosines of a unit vector e pointing in the desired direction.

Theorem. The rate of change of a function CD(x, y, z) in the direction of a vector with direction angles (α, β, γ) is given by

3) $$\frac{d\Phi}{ds} = \frac{\partial\Phi}{dx}\cos\alpha + \frac{\partial\Phi}{dy}\cos\beta + \frac{\partial\Phi}{dz}\cos\gamma$$

where s corresponds to distance in the metric of the coordinate system. That direction for which the function Φ at point P has its maximum value is called the gradient of Φ at P.

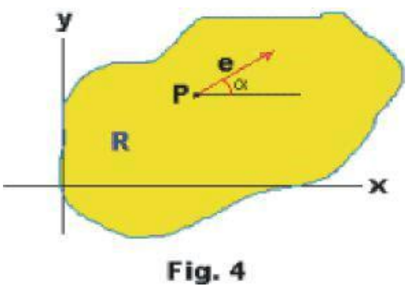


Fig. 4

We shall prove the theorem shortly. First let us consider the same problem for two dimensional space.

Let Φ(x, y) be a scalar point function defined over some region R of the plane. At some specified point P(x, y) of R we wish to know the rate of change of Φ in a particular direction. We specify the direction by supplying the angle a that a unit vector e pointing in the desired direction makes with the positive x direction. See Fig. 4. The rate of change of function Φat point P in the direction of **e** corresponding to angle a is given by

4) $$\frac{d\Phi}{ds} = \frac{\partial\Phi}{dx}\cos\alpha + \frac{\partial\Phi}{dy}\sin\beta$$

where s corresponds to distance in the metric of the coordinate system. We show this as follows:

Let

T = f(x, y)

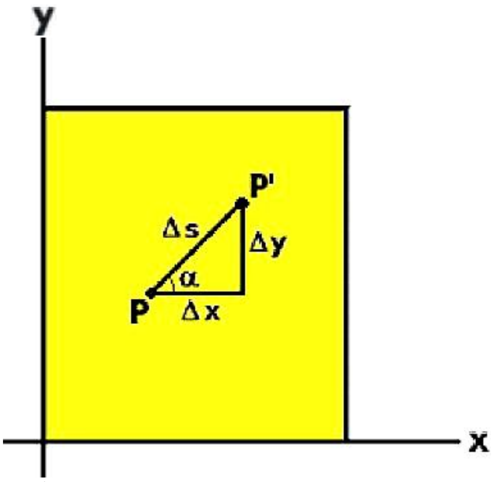


Fig. 5

where T is the temperature at any point of the plate shown in Fig. 5. We wish to derive expression 4) above. In other words, we wish to derive the expression for the rate of change of T with respect to the distance moved in any selected direction. Suppose we move from point P to point P'. This represents a displacement Δx in the x-direction and Δy in the y-direction. The distance moved along the plate is

$$PP' = \Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

The direction is given by the angle α that PP' makes with the positive x-direction. The change in the value of T corresponding to the displacement from P to P' is

$$\Delta T = \frac{\partial T}{\partial x}\Delta x + \frac{\partial T}{\partial y}\Delta y + \varepsilon\sqrt{(\Delta x)^2 + (\Delta y)^2}$$

where ε is a quantity that approaches 0 when Δx and Δy approach 0.
If we divide ΔT by the distance moved along the plate, we have

$$\frac{\Delta T}{\Delta s}=\frac{\partial T}{\partial x}\frac{\Delta x}{\Delta s}+\frac{\partial T}{\partial y}\frac{\Delta y}{\Delta s}+\varepsilon\sqrt{\left(\frac{\Delta x}{\Delta s}\right)^2+\left(\frac{\Delta y}{\Delta s}\right)^2}\ .$$

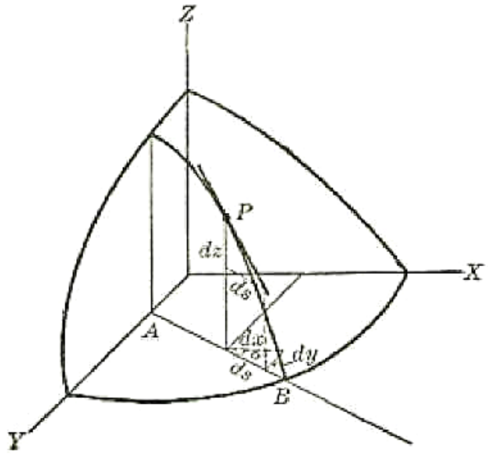


Fig. 6

From Fig. 5 we observe that $\Delta x/\Delta s = \cos\alpha$ and $\Delta y/\Delta s = \sin\alpha$. Making these substitutions and letting P' approach P along line PP', we have

$$\frac{dT}{ds}=\frac{\partial T}{\partial x}\cos\alpha+\frac{\partial T}{\partial y}\sin\alpha$$

This is the directional derivative of T in the direction a.
A geometric interpretation of a directional derivative in the case of a function $z = f(x, y)$ is that of a tangent to the surface at point P as shown in Fig. 6.

Def. Directional derivative. The directional derivative of a scalar point function $\Phi(x, y, z)$ is the rate of change of the function $\Phi(x, y, z)$ at a particular point $P(x, y, z)$ as measured in a specified direction.

Tech. Let $\Phi(x, y, z)$ be a scalar point function possessing first partial derivatives throughout some region R of space. Let $P(x_0, y_0, z_0)$ be some point in R at which we wish to compute the directional derivative and let $P'(x_1, y_1, z_1)$ be a neighboring point. Let the distance from P to P' be As. Then the directional derivative of Φ in the direction PP' is given by

5) $$\frac{d\Phi}{ds}=\lim_{p'\to p}\frac{\Phi[P'(x_1,y_1,z_1)]-\Phi[P(x_0,y_0,z_0)]}{\Delta s}$$

where P' approaches P along the line PP' and Δs approaches 0.

Using this definition, let us now derive 3) above. In moving from P to P' the function Φ will change by an amount

$$\Delta\Phi = \frac{\partial\Phi}{\partial x}\Delta x + \frac{\partial\Phi}{\partial y}\Delta y + \frac{\partial\Phi}{\partial z}\Delta z + \varepsilon_1\Delta x + \varepsilon_2\Delta y + \varepsilon_3\Delta z$$

whereε_1, ε_2, ε_3 are higher order infinitesimals which approach zero as P' approaches P i.e. as Δx, Δy and Δz approach zero. If we divide the change ΔØ by the distance Δs we obtain a measure of the rate at which Øchanges as we move from P to P':

6) $$\frac{\Delta\Phi}{\Delta s} = \frac{\partial\Phi}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial\Phi}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial\Phi}{\partial z}\frac{\Delta z}{\Delta s} + \varepsilon_1\frac{\Delta x}{\Delta s} + \varepsilon_2\frac{\Delta y}{\Delta s} + \varepsilon_3\frac{\Delta z}{\Delta s}$$

We now observe that Δx/Δs, Δy/Δs, Δz/Δs are the direction cosines of the line segment PP'. They are also the direction cosines of a unit vector **e** located at P pointing in the direction of '. If the direction angles of **e** are α, β, γ, then Δx/Δs, Δy/Δs, Δz/Δs are equal to cos α, cosβ, and cosγ, respectively. Thus 6) becomes

$$\frac{d\Phi}{ds} = \frac{\partial\Phi}{\partial x}\cos\alpha + \frac{\partial\Phi}{\partial y}\cos\beta + \frac{\partial\Phi}{\partial z}\cos\gamma + \varepsilon_1\frac{\Delta x}{\Delta s} + \varepsilon_2\frac{\Delta y}{\Delta s} + \varepsilon_3\frac{\Delta z}{\Delta s}$$

and

7) $$\frac{d\Phi}{ds} = \lim_{\Delta x \to 0}\frac{\partial\Phi}{\partial s}\frac{\partial\Phi}{\partial x}\cos\alpha + \frac{\partial\Phi}{\partial y}\cos\beta + \frac{\partial\Phi}{\partial z}\cos\gamma$$

Let us note that 7) can be written in vector form as the following dot product:

8) $$\frac{d\Phi}{ds} = \left[\frac{\partial\Phi}{\partial x}\frac{\partial\Phi}{\partial y}\frac{\partial\Phi}{\partial z}\right].[\cos\alpha\cos\beta\cos\gamma] = \left[\frac{\partial\Phi}{\partial x}\frac{\partial\Phi}{\partial y}\frac{\partial\Phi}{\partial z}\right].e$$

The vector

$$\left[\frac{\partial\Phi}{\partial x}\frac{\partial\Phi}{\partial y}\frac{\partial\Phi}{\partial z}\right]$$

is called the gradient of Φ. Thus the directional derivative of Φ is equal to the dot product of the gradient of Φand the vector **e**. In other words,

$$\left(\frac{d\Phi}{ds}\right)_e = grad\Phi.e$$

where

$$\left(\frac{d\Phi}{ds}\right)_e$$

is the directional derivative of Φ in the direction of unit vector e.

If the vector **e** is pointed in the same direction as the gradient of Φ then the directional derivative of Φ is equal to the gradient of Φ.

Differentiation under the integral sign. Leibnitz's rule. We now consider differentiation with respect to a parameter that occurs under an integral sign, or in the limits of integration, or in both places.

Theorem 1. Let

$$F(x) = \int_a^x f(t)dt$$

where $a \leq x \leq b$ and f is assumed to be integrable on [a, b]. Then the function *F(x)* is continuous and *F'(x) = f(x)* at each point where *f*(*x*) is continuous.

Theorem 2. Let f(x, α) and ∂f/∂α be continuous in some region
R: ($a \leq x \leq b$, $c \leq \alpha \leq d$) of the x-α plane. Let

9) $$G(\alpha) = \int_a^b f(x,\alpha)dx \qquad c \leq \alpha \leq d$$

Then

10) $$\frac{dG}{d\alpha} = \int_a^b \frac{\partial f(x,\alpha)}{\partial \alpha} dx$$

Theorem 3. Leibnitz's rule. Let

11) $$G(\alpha) = \int_{u_1}^{u_2} f(x,\alpha)\,dx \qquad c \leq \alpha \leq d$$

where u1 and u2 are functions of the parameter α i.e.

$u_1 = u_1(\alpha)$
$u_2 = u_2(\alpha)$.

Let f(x, α) and ∂f/∂α be continuous in both x and α in a region **R** of the x-α plane that includes the region $u_1 \leq x \leq u_2$, $c \leq \alpha \leq d$. Let u_1 and u_2 be continuous and have continuous derivatives for $c \leq \alpha \leq d$. Then

12) $$\frac{dG}{d\alpha} = \int_{u_1}^{u_2} \frac{\partial f(x,\alpha)}{\partial \alpha} dx + f(u_2,\alpha)\frac{du_2}{d\alpha} - f(u_1,\alpha)\frac{du_1}{d\alpha}$$

where f(u1, α) is the expression obtained by substituting the expression $u_1(\alpha)$ for *x* in *f*(*x*, *α*). Similarly for $f(u_2, \alpha)$. The quantities $f(u_1, \alpha)$ and
$f(u_2, \alpha)$ correspond to $\partial G/u_1$ and $\partial G/u_2$ respectively and 12) represents the chain rule.

Order of differentiation. For most functions that one meets

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

However, in some cases it is not true. Under what circumstances is it true? It is true if both functions f_{yx} and f_{xy} are continuous at the point where the partials are bein taken.

Theorem. Let the function f(x, y) be defined in some neighborhood of the point (a, b). Let the partial derivatives *fx*, *fy*, *fxy*, and *fyx* also be defined in this neighborhood. Then if *fxy* and *fyx* are both continuous at (a,), *fxy*(a, b) = *fyx*(a, b).

EXAMPLE

Given $u = x^2 + 2y$ where $x = r\sin(t)$ and $y = \sin^2(t)$, determine the value of $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial t}$ using the chain rule.

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = (2x)(\sin(t)) + (2)(0) = 2r\sin^2(t)$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} = (2x)(r\cos(t)) + (2)(2\sin(t)\cos(t))$$
$$= 2(r\sin(t))r\cos(t) + 4\sin(t)\cos(t) = 2\left(r^2+2\right)\sin(t)\cos(t).$$

4.0 CONCLUSION

In this unit, you have applied total derivative on chain rule. You have solved problems directional derivatives using total derivative. You have used total derivative to solve differentiation under integral sign and leibnitz rule.

5.0 SUMMARY

In this unit, you have studied the following:

- The application of total derivative on chain rule
- The application of total derivative on directional derivative
- The application of total derivative on differentiation under integral sign
- The application of total derivative on leibnitz rule.

6.0 TUTOR - MARKED ASSIGNMENT

1. Find all directional derivatives of the function
   F(x,y) – $(3\,x^2 + \,y^4)^{\frac{1}{4}}$

   where $(x,y) \in R_2$), in the point (0,0)

2. Find the integral of the function
   F(x,y,z) = $3x^2 + 2xyz$
   In the point (0,1)

3. Find the total derivative of the function
   F(xy) = $3xy + 4y^2$

7.0 REFERENCES/FURTHER READING

James & James. Mathematics Dictionary

Middlemiss. Differential and Integral Calculus

Spiegel. Advanced Calculus

Taylor. Advanced Calculus

Spiegel. Vector Analysis

Bernard R. Gelbaum and John M. H. Olmsted.Counterexamples in Analysis.

Dover, 2003.