CALCULUS
DEGREE IN SOFTWARE ENGINEERING
CHAPTER 23. DIFFERENTIABILITY

DEFINITION OF DIFFERENTIABILITY

A function $z = f(x, y)$ is differentiable at (x_0, y_0) if both $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ exist at (x_0, y_0) and

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

with $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as Δx and Δy tend to zero. ϵ_1 and ϵ_2 are functions of the point and the increments of the variables. We call the function differentiable if it is differentiable at all points of its domain. The graph of a differentiable function is a smooth surface. You can see that if we only consider the two first terms on the right-hand side, we have the equation of the tangent plane. This means that we can approximate the surface by its tangent plane and the error produced by the other two terms goes faster to zero than the linear part.
If we take the limit on both sides of the definition as Δx and Δy approach 0, we can prove that if f is differentiable at a point, it is continuous at that point too.

EXAMPLE

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

We calculate $f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$

$$(x_0 + \Delta x)^2 + (y_0 + \Delta y)^2 - (x_0^2 + y_0^2) = 2x_0\Delta x + 2y_0\Delta y + (\Delta x)^2 + (\Delta y)^2$$

Then, if we compare with the definion of differentiability, we see that the function is differentiable with $\epsilon_1 = \Delta x$ and $\epsilon_2 = \Delta y$
It would be complicated to establish the differentiability of a function by applying the definition. Luckily, there is a simpler way stated in the following theorem

A SUFFICIENT CONDITION OF DIFFERENTIABILITY

If the partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ of a function $f(x, y)$ are continuous throughout an open region R, then f is differentiable at every point of R.

Thus, to prove the differentiability of a function we must only see that its partial derivatives are continuous. This theorem and the mixed derivative theorem are proved in Appendix 9 of Thomas' Calculus (twelfth edition).
An interesting fact is that a differentiable function is also continuous as expected from the single variable case. But take into account that we need an extended definition of differentiablity- not the mere existence of partial derivatives- to prove this property.

DIFFERENTIABILITY IMPLIES CONTINUITY

If a function $f(x, y)$ is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

If we write the definition of differentiability

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

and take the limit as Δx and Δy tend to zero, we see that

$$\lim_{(\Delta x, \Delta y) \to (0,0)} f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0)$$

since ϵ_1 and ϵ_2 approach zero as Δx and Δy approach zero.
Then, if a function f is C^2 on its domain, i.e. it has continuous second-order partial derivatives, its first-order partial derivatives will be differentiable and then continuous. This makes the function differenciable and also continuous. To sum this up

$$C^2(D) \subseteq C^1(D) \subseteq C^0(D)$$

DIRECTIONAL DERIVATIVES

Let $P_0 = (x_0, y_0)$ be an interior point of the domain D of a function $f(x, y)$. If $\vec{u} = (u_1, u_2)$ is a unit vector that indicates a certain direction on the plane, we can define

$$(D_{\vec{u}} f)_{P_0} = \lim_{h \to 0} \frac{f(x_0 + hu_1, y_0 + hu_2) - f(x_0, y_0)}{h}$$

and we call this limit the directional derivative of f at P_0 in the direction of the unit vector $\vec{u}$. This means the local variation of the function in that direction divided by the distance traveled: the variation of the function per distance unit in that direction. You can observe that this concept is a generalization of the partial derivatives, which are the directional derivatives along the x-axis, $\dfrac{\partial f}{\partial x}$ and the y-axis, $\dfrac{\partial f}{\partial y}$.
Can we calculate the directional derivatives in a simple way ? In general, we could define

$$g(h) = f(x_0 + hu_1, y_0 + hu_2)$$

and then

$$(D_{\vec{u}}f)_{P_0} = g'(0)$$

where we have fixed P_0 and $\vec{u}$ and taken the derivative with respect to h at $h = 0$. Besides, if the function is differentiable at P_0, we can write

$$f(x_0 + hu_1, y_0 + hu_2) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\, hu_1 + \frac{\partial f}{\partial y}(x_0, y_0)\, hu_2 + \epsilon_1\, hu_1 + \epsilon_2\, hu_2$$

If we divide both terms by h and take the limit as h approaches 0, then

$$(D_{\vec{u}}f)_{P_0} = \frac{\partial f}{\partial x}(x_0, y_0)\, u_1 + \frac{\partial f}{\partial y}(x_0, y_0)\, u_2 = \vec{\nabla} f.\, \vec{u}$$

that is, the directional derivative is the scalar product of the gradient times the unit vector.
A geometrical interpretation of the directional derivative can be easily done: at a point $P_0 = (x_0, y_0)$ we draw the plane containing the vector $\vec{u}$ and perpendicular to the XY-plane; the intersection of this plane with the surface $z = f(x, y)$ is a curve $z = f(x_0 + hu_1, y_0 + hu_2) = g(h)$ and the directional derivative $(D_{\vec{u}}f)_{P_0}$ is the slope of this curve at $h = 0$, where $z = f(x_0, y_0)$. See Figure 14.27 in Presentation 10.

EXAMPLE

We will calculate the directional derivative of $f(x, y) = x^2y^3$ at the point $(2, 1)$ in the direction of the vector $\vec{v} = (1, 1)$. Since the function is differentiable, we find the gradient at the point.

$$\vec{\nabla} f(x, y) = (2xy^3, 3x^2y^2)$$

$$\vec{\nabla} f(2, 1) = (4, 12)$$

we obtain the unit vector in the direction of $\vec{v}$

$$\vec{u} = (1, 1)/\sqrt{2}$$

and, finally

$$\vec{\nabla} f(2, 1).\vec{u} = (4 + 12)/\sqrt{2} = 16/\sqrt{2}$$

Thus, for differentiable functions, the calculation of directional derivatives is as simple or difficult as the calculation of partial derivatives.
Of course, the extension of these concepts to functions of more than two variables is trivial. For instance, a differentiable function of three variables has a directional derivative

$$(D_{\vec{u}}f)_{(x_0,y_0,z_0)} = \vec{\nabla} f(x_0, y_0, z_0).\vec{u}$$

where $\vec{u}$ is a unit vector (u_1, u_2, u_3). You can figure out the formula for a function of n variables. To finish this chapter, we will present some properties of the gradient.

PROPERTIES OF THE GRADIENT

A differentiable function of two variables has a directional derivative

$$(D_{\vec{u}}f)_{(x_0,y_0)} = |\vec{\nabla} f(x_0, y_0)||\vec{u}| \cos\alpha = |\vec{\nabla} f(x_0, y_0)| \cos\alpha$$

taking into account the definition of the scalar product and with α the angle between the gradient and the unit vector.
Now, we can analyze different cases:

1) if $\alpha = 0$, gradient and unit vector are in the same direction, then the directional derivative is maximum $\quad (\cos\alpha = 1)$

$$(D_{\vec{u}}f)_{(x_0,y_0)} = |\vec{\nabla} f(x_0, y_0)|$$

and the gradient indicates the direction along which there is a fastest increase of the function (a highest positive slope)

2) if $\alpha = \pi$, gradient and unit vector are in opposite directions, then the directional derivative is minimum $\quad (\cos\alpha = -1)$

$$(D_{\vec{u}}f)_{(x_0,y_0)} = -|\vec{\nabla} f(x_0, y_0)|$$

and the gradient indicates the direction along which there is a fastest decrease of the function (the most negative slope)

3) if $\alpha = \pi/2$, gradient and unit vector are in orthogonal directions, then the directional derivative is zero $(\cos\alpha = 0)$

$$(D_{\vec{u}}f)_{(x_0,y_0)} = 0$$

and the gradient indicates the direction along which the local change in the function is zero. This means that the gradient is perpendicular to the level curve that passes though that point- curve where the function remains constant-.
Therefore, the gradient of temperature is orthogonal to the isotherms on a weather map. The gradient indicates the direction of maximum increase of the temperature. If we take the direction opposite to the gradient, we move along the direction of fastest decrease. This method, moving along this direction, is called steepest descent and is commonly used for finding the minimum of a function in many applications, such as neural networks.
For a function of three variables, we also have the three previous cases and now the gradient is perpendicular to the level surfaces. For instance, minus the gradient of

the gravitational potencial, Φ, the gravitational field, $\vec{E} = -\vec{\nabla}\Phi$, is perpendicular to the equipotencial surfaces and indicates the direction of maximum decrease of the potential.

ALGEBRA RULES FOR THE GRADIENT

1. Sum Rule

$$\vec{\nabla}(f+g) = \vec{\nabla}f + \vec{\nabla}g$$

2. Difference Rule

$$\vec{\nabla}(f-g) = \vec{\nabla}f - \vec{\nabla}g$$

3. Constant Multiple Rule

$$\vec{\nabla}(kf) = k\vec{\nabla}f$$

4. Product Rule

$$\vec{\nabla}(f.g) = f\vec{\nabla}g + g\vec{\nabla}f$$

5. Quotient Rule

$$\vec{\nabla}(f/g) = \frac{f\vec{\nabla}g - g\vec{\nabla}f}{g^2}$$

   with $g \neq 0$

These are the basic algebraic rules for the gradients. These rules are often applied in physics and other sciences.