Lecture 12: Gradient

1 Gradient

1.1 Introduce Gradient

According to the chain rule, suppose that there is a function $w = w(x, y, z)$, where $x = x(t)$, $y = y(t)$, and $z = z(t)$, then

$$\begin{aligned}
\frac{dw}{dt} &= w_x\frac{dx}{dt} + w_y\frac{dy}{dt} + w_z\frac{dz}{dt} \\
&= < w_x, w_y, w_z > \cdot < \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} > \\
&= \nabla w \cdot \frac{d\vec{r}}{dt} \\
\nabla w &= < w_x, w_y, w_z > \\
\frac{d\vec{r}}{dt} &= < \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} >
\end{aligned}$$

The vector $< w_x, w_y, w_z >$ is called gradient, denoted as ∇w. Gradients on different positions can have different directions and magnitude.

1.2 A Property of Gradient

Theorem. *The gradient of a function is always perpendicular to the level surfaces of the function, where a level surface means the set of points on the function domain whose function values are equal to a constant.*

Notice that the level surfaces refer to points in the function domain, rather than on the function graph. It doesn't include the dimension of the function value.

Example 1. *Examine the relation between the gradient and the level surfaces for the function $w = a_1x + a_2y + a_3z$.*

$$
\begin{aligned}
\frac{\partial w}{\partial x} &= a_1 \\
\frac{\partial w}{\partial y} &= a_2 \\
\frac{\partial w}{\partial z} &= a_3 \\
\nabla w &= < a_1, a_2, a_3 >
\end{aligned}
$$

The level surfaces of the function $w = a_1x + a_2y + a_3z = c$*, where* c *is any constant, are a series of planes parallel to each other. For any level surfaces of the function, the gradient* ∇w *is a normal vector to it. Therefore, the gradient is always perpendicular to the level surfaces for the function* $w = a_1x + a_2y + a_3z$*.*

Example 2. *Examine the relation between the gradient and the level surfaces for the function* $w = x^2 + y^2$*.*

$$
\begin{aligned}
\frac{\partial w}{\partial x} &= 2x \\
\frac{\partial w}{\partial y} &= 2y \\
\nabla w &= < 2x, 2y > = < x, y >
\end{aligned}
$$

The level surfaces of the function $w = x^2 + y^2 = c$*, where* c *is any constant, are a series of circles whose centers are the origin. At any point* (x_0, y_0) *in the function domain, the gradient vector is* $< x_0, y_0 >$*, which has the same direction as the radius through* (x_0, y_0) *on the level surface* $x^2 + y^2 = x_0^2 + y_0^2$*. A circle's radius is always perpendicular to the circle, hence the gradient vector is perpendicular to the level surface. Therefore, the gradient is always perpendicular to the level surfaces for the function* $w = x^2 + y^2$*.*

Proof of the theorem:
Suppose there is a curve $\vec{r} = \vec{r}(t)$ that always stays on a level surface of a function f. The velocity vector $\frac{d\vec{r}}{dt}$ is tangent to the curve, and therefore tangent to the level surface on which the curve stays.
By chain rule, on any level surfaces it satisfies the following equation:

$$\frac{df}{dt} = \nabla f \cdot \frac{d\vec{r}}{dt} = 0$$

Therefore, at any points of a curve that stays on a level surface:

$$\nabla f \perp \frac{d\vec{r}}{dt}$$

The same reasoning applies to any curves on a level surface, so at any points on a level surface, the gradient ∇f is perpendicular to velocity vectors of every direction, which are tangent to the level surface. Therefore, at any points on a level surface, the gradient ∇f is perpendicular to the tangent plane to the level surface.

1.3 Application of Gradient

We can use gradients to find the equation of the tangent line of a function graph at any points.

Example 3. *Find the equation of the tangent plane to the surface $x^2+y^2-z^2=4$ at the point (2, 1, 1).*
Solution:
The surface $x^2+y^2-z^2=4$ is a level surface of the function $w=x^2+y^2-z^2$.

$$\frac{\partial w}{\partial x} = 2x$$
$$\frac{\partial w}{\partial y} = 2y$$
$$\frac{\partial w}{\partial z} = -2z$$
$$\nabla w =< 2x, 2y, -2z >=< x, y, -z >$$

Hence the gradient of the function at the point (2, 1, 1) is $< 2, 1, -1 >$. According to the property of gradients, the gradient is perpendicular to the tangent plane of the corresponding level surface at the point. Therefore, the equation of the tangent plane is

$$2x + y - z = c, \text{where c is a constant}$$

Then we can substitute the coordinate of the point (2, 1, 1) into the plane equation:

$$c = 2 \times 2 + 1 \times 1 - 1 \times 1 = 4$$

Therefore, the equation of the tangent plane to the surface $x^2+y^2-z^2=4$ at the point (2, 1, 1) is

$$2x + y - z = 4$$

2 Directional Derivatives

2.1 Introduce Directional Derivatives

How to calculate the partial derivative towards any direction $\vec{u}$, rather than only along the x axis and y axis?

Suppose there is a multivariable function $f(x, y)$, and an unit vector $\hat{\vec{u}} =< a, b >$. The partial derivative along a direction is the rate of change of the function value $f(x, y)$ over the arclength, hence we need the differential of the arclength along the direction of $\hat{\vec{u}}$, which is denoted by ds.
By chain rule, we have

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$
$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds}$$
$$\frac{df}{ds} =< \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} > \cdot < \frac{dx}{ds}, \frac{dy}{ds} >$$
$$\frac{df}{ds} = \nabla f \cdot \frac{d\vec{r}}{ds}$$

Therefore, we need to describe the position vector $\vec{r}$ along the straight line trajectory with the direction $\vec{u}$ using the parameter s, the arclength.

$$\begin{cases} x(s) = x_0 + as \\ y(s) = y_0 + bs \end{cases}$$
$$\frac{d\vec{r}}{ds} =< a, b >$$
$$\frac{d\vec{r}}{ds} = \vec{u}$$

Therefore,

$$\begin{aligned} \frac{df}{ds} &= \nabla f \cdot \frac{d\vec{r}}{ds} \\ &= \nabla f \cdot \vec{u} \end{aligned}$$

which is the formula for directional derivatives.

Definition. *The instantaneous rate of change of a multivariable function f at a given direction $\vec{u}$ is called the directional derivative, whose notation is $\frac{df}{ds}|_{\vec{u}}$. The formula of directional derivatives is*

$$\frac{df}{ds}|_{\vec{u}} = \nabla f \cdot \vec{u}$$

2.2 Geometric Interpretation of Directional Derivatives

For a multivariable function f with two independent variables, the geometric interpretation of the function toward a direction $\vec{u}$ is the slope of the slice of the function graph by a vertical plane parallel to $\vec{u}$.

Example 4. *Verify the formula of directional derivatives by finding the partial derivative along the x axis.*
Suppose there is a multivariable function $f(x, y)$.

$$\begin{aligned}\frac{\partial f}{\partial x} &= \nabla f \cdot \vec{i} \\ &=< \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} > \cdot \vec{i} \\ &= \frac{\partial f}{\partial x}\end{aligned}$$

2.3 Max/Min of Directional Derivatives

According to the geometric interpretation of dot product of vectors:

$$\begin{aligned}\frac{df}{ds}|_{\vec{u}} &= \nabla f \cdot \vec{u} \\ &= |\nabla f| \times |\vec{u}| \times \cos\theta \\ &= \cos\theta \times |\nabla f|\end{aligned}$$

We can see that the directional derivative changes as the angle θ between its direction and the gradient vector changes. It is obvious that at a given point:

- The directional derivative with the maximum value has the same direction as the gradient, and the maximum value is the magnitude of the gradient.
- The directional derivative with the minimum value has the opposite direction to the gradient, and the minimum value is the negative value of the magnitude of the gradient.