The chain rule for scalar functions 30/05

We shall study the geometry of a function $f: D \to \mathbb{R}$ where $D = \mathrm{dom}(f)$ is a subset of the plane $\mathbb{R}^2$. Next time we shall do the same when $D \subseteq \mathbb{R}^3$.

Consider a curve in D parametrized by

$$\gamma: [a,b] \to D, \qquad \gamma(t) = (x(t), y(t))$$

with nowhere-zero tangent (velocity) vector $\gamma'(t)$. To define the line integral $\int_\gamma f$, we considered the composition

$$(f \circ \gamma)(t) = f(x(t), y(t)) = f(t)$$

and multiplied by $\|\gamma'(t)\|$ before integrating to get a number.

This time we shall differentiate $f(t)$ to obtain a new function defined along the curve, namely

$$f'(t) = \frac{d}{dt} f(x(t), y(t)).$$

The ***chain rule*** states that this equals

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = f_x\, x'(t) + f_y\, y'(t) = (\nabla f) \cdot \gamma'(t).$$

Example 1. If $f(x,y) = p(x)q(y)$ then

$$\begin{aligned} f'(t) &= (p'(x(t))x'(t))q(y(t)) + p(x(t))(q'(y(t))y'(t)) \\ &= x'(t) f_x + y'(t) f_y. \end{aligned}$$

For the chain rule to work, all derivatives must be continuous, and the gradient is calculated at the point $\gamma(t) = (x(t), y(t))$.

Example 2. If $\gamma(t) = (x_0 + t, y_0)$ is a line parallel to the x-axis,

$$f'(0) = \lim_{t \to 0} \frac{f(x_0 + t, y_0)}{t} = f_x(x_0, y_0)$$

measures the rate of change of f in the x-direction.

If $\gamma(t) = (x_0 + tA, y_0 + tB)$ is a line through (x_0, y_0),

$$f'(0) = \lim_{t\to 0} \frac{f(x_0 + tA, y_0 + tB)}{t} = Af_x(x_0, y_0) + Bf_y(x_0, y_0)$$

is the ***directional derivative*** relative to the vector $\mathbf{p} = (A, B)$, sometimes written

$$\nabla_{\mathbf{p}} f = \frac{df}{d\mathbf{p}} = (\nabla f) \cdot \mathbf{p}.$$

In general,

$$f'(t) = (\nabla f) \cdot \gamma'(t) = (f_x \;\; f_y)\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Once we fix t so that f_x, f_y, x', y' are numbers, we can regard $f'(t)$ as a composition of ***linear*** mappings:

$$\mathbb{R} \xleftarrow{(f_x \;\; f_y)} \mathbb{R}^2 \xleftarrow{\begin{pmatrix} x' \\ y' \end{pmatrix}} \mathbb{R},$$

derived from the composition

$$\mathbb{R} \xleftarrow{f} D \xleftarrow{\gamma} [a, b].$$

and the choice of starting point t.

Corollary. If the scalar function f is ***constant*** along γ then $(\nabla f)(\gamma(t))$ is orthogonal to $\gamma'(t)$. The gradient vector field ∇f is in fact orthogonal to the level curves of C at all points where it is non-zero.

Example 3. If $f(x, y) = x^2 + y^2$, the level curves $f = c$ are circles (providing $c > 0$). The vector $\nabla f = 2(x, y)$ points radially, everywhere orthogonal to the circles. Now let

$$f(x, y) = (x \;\; y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2$$

be a quadratic form. Then

$$\nabla f = 2(ax + by,\; bx + cy) = 2(x \;\; y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

The eigenvectors of the 2×2 matrix determine the axes of symmetry for which ∇f is perpendicular to the level curves.