

Lecture 5: 2.5 The Chain rule cont..

The **Chain rule in several variables**: Suppose that $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$, $f : \mathbf{R}^m \rightarrow \mathbf{R}^p$ and let $h = f \circ g : \mathbf{R}^n \rightarrow \mathbf{R}^p$ (i.e. $h(x) = f(g(x))$). Then

$$Dh(x_0) = Df(y_0) Dg(x_0), \quad \text{where} \quad y_0 = g(x_0)$$

and the right hand side is the $p \times n$ matrix formed by the matrix product of the $p \times m$ matrix $Df(y_0)$ by the $m \times n$ matrix $Dg(x_0)$.

The **Chain rule case 1** If $h(t) = f(\mathbf{c}(t))$ where $f : \mathbf{R}^3 \rightarrow \mathbf{R}$, $\mathbf{c}(t) = (x(t), y(t), z(t))$ is a path or curve, then by the chain rule

$$\frac{\partial h}{\partial t} = Df D\mathbf{c} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{bmatrix} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}$$

The **gradient** of a function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ given by

$$\mathbf{grad} f = \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

This can also be expressed with the gradient notation and dot product

$$\frac{dh}{dt}(t) = \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$$

where

$$\mathbf{c}'(t) = (x'(t), y'(t), z'(t)).$$

Ex. If $z = x^2 + y^2$, $x = \cos t$ and $y = \sin t$ find dz/dt .

Sol. 1 $\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} = 2x(-\sin t) + 2y \cos t = 2 \cos t(-\sin t) + 2 \sin t \cos t = 0$

Sol. 2 $z = x^2 + y^2 = \cos^2 t + \sin^2 t = 1, \quad \frac{dz}{dt} = 0$

Let $\mathbf{c}(t) = (x(t), y(t))$ be a curve in the plane and let $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$. Then $\mathbf{p}(t) = f(\mathbf{c}(t))$ is another curve in the plane and by the chain rule

$$\mathbf{p}'(t) = Df(\mathbf{c}(t)) \mathbf{c}'(t)$$

where the right is matrix multiplication between the 2×2 matrix Df and the column vector \mathbf{c}' .

Ex Let $f(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$, $\mathbf{c}(t) = (1 + t, 3t)$ and $\mathbf{p}(t) = f(\mathbf{c}(t))$. Find $\mathbf{p}'(t)$.

Sol

$$\mathbf{p}'(t) = Df(\mathbf{c}(t)) \mathbf{c}'(t) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

2.6 The gradient and directional derivative.

The **gradient** of a function $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ is the vector

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

i.e. it is the matrix of derivatives written as a vector.

Consider the equation of a line in space $\ell(t) = \mathbf{x} + t\mathbf{v}$, $-\infty < t < \infty$.

The function $h(t) = f \circ \ell(t) = f(\mathbf{x} + t\mathbf{v})$ represents the function f restricted to the line.

The **directional derivative** of f at \mathbf{x} in the direction of unit vector \mathbf{v} is given by

$$\left. \frac{d}{dt} f(\mathbf{x} + t\mathbf{v}) \right|_{t=0} = \nabla f(\mathbf{x}) \cdot \mathbf{v}$$

Here the equality follows from the chain rule: $h' = Dh = Df D\ell = \nabla f \cdot \ell' = \nabla f \cdot \mathbf{v}$.

The reason we choose \mathbf{v} to be a unit vector is that we want the directional derivative to represent the rate of change in different directions.

Suppose that f represents the temperature at different points in space. Suppose that a fly flies along the line above at unit speed then the change of temperature per unit time or distance is the directional derivative.

The gradient points in the direction along which f increases the fastest.

In fact $\nabla f \cdot \mathbf{v} = |\nabla f| |\mathbf{v}| \cos \theta$, where θ is the angle between ∇f and \mathbf{v} , and the max is when $\cos \theta = 1$.

Suppose we are lost in wood and we want to reach a high hill top to see where we are. However, we can only see a few feet in front of us because of the high trees. In which direction shall we go in order to reach a hill-top fast. The answer is that if we go in the direction of the grade likely to reach a hill-top fast.

The gradient is normal to the tangent plane of the level surface:

Let $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ and let (x_0, y_0, z_0) be a point on the level surface S defined by $f(x, y, z) = k$, for some constant k . Then $\nabla f(x_0, y_0, z_0)$ is normal to the level surface in the following sense: If \mathbf{v} is a tangent vector to a path $\mathbf{c}(t)$ in S with $\mathbf{c}(0) = (x_0, y_0, z_0)$, then $\nabla f(x_0, y_0, z_0) \cdot \mathbf{v} = 0$.

In fact, since $f(x(t), y(t), z(t)) = k$ it follows that

$$0 = \frac{d}{dt} f(x(t), y(t), z(t)) = \nabla f(x(t), y(t), z(t)) \cdot \mathbf{c}'(t)$$

Let S be a level surface $f(x, y, z) = k$. The **tangent plane** of S at a point (x_0, y_0, z_0) of S is defined by the equation

$$\nabla f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0$$

In fact (x, y, z) is in the tangent plane if $(x, y, z) - (x_0, y_0, z_0)$ is parallel to the plane and hence perpendicular to the normal.