Métodos Matemáticos de Bioingeniería

Grado en Ingeniería Biomédica
Lecture 9

Marius A. Marinescu

Departamento de Teoría de la Señal y Comunicaciones
Área de Estadística e Investigación Operativa
Universidad Rey Juan Carlos

22 de marzo de 2021

Outline

Example 1 (One dimension)

- Let

$$\begin{aligned} f(x) &= \sin x \\ x(t) &= t^3 + t \end{aligned}$$

- We may then construct the **composite function**

$$(f \circ x)(t) = f(x(t)) = \sin(t^3 + t)$$

- The chain rule tells us how to find the derivative of $f \circ x$ with respect to t

$$(f \circ x)'(t) = \frac{d}{dt}(\sin(t^3 + t)) = (\cos(t^3 + t))(3t^2 + 1)$$

- Since $x = t^3 + t$, we can see it as

$$(f \circ x)'(t) = \frac{d}{dx}(\sin x) \cdot \frac{d}{dt}(t^3 + t) = f'(x) \cdot x'(t)$$

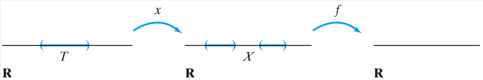
The Chain Rule in One Variable

- In general, suppose X and T are open subsets of $\mathbb{R}$.
- Suppose we define functions

$$f : X \subseteq \mathbb{R} \to \mathbb{R}$$
$$x : T \subseteq \mathbb{R} \to \mathbb{R}$$

- Suppose that the composite function makes sense

$$f \circ x : T \subseteq \mathbb{R} \to X \subseteq \mathbb{R} \to \mathbb{R}$$



This means that the range of the function x must be contained in X, the domain of f

Theorem 5.1: The Chain Rule in One Variable

- Let X and T be open subsets of $\mathbb{R}$
- We define functions

$$f : X \subseteq \mathbb{R} \to \mathbb{R}$$
$$x : T \subseteq \mathbb{R} \to \mathbb{R}$$
$$f \circ x : X \subseteq \mathbb{R} \to \mathbb{R}$$

- Suppose x is differentiable at $t_0 \in T$, and
- Suppose f is differentiable at $x_0 = x(t_0) \in X$
- Then, **the composite** $f \circ x$ is **differentiable** at t_0, and

$$(f \circ x)'(t_0) = f'(x_0)x'(t_0)$$

The Chain Rule in Two Variables

- Assume $f : X \subseteq \mathbb{R}^2 \to \mathbb{R}$ is a $\mathbf{C^1}$ **function** of two variables.
- Assume $\mathbf{x} : T \subseteq \mathbb{R} \to \mathbb{R}^2$ is a differentiable vector-valued function of a single variable and the range of $\mathbf{x}$ is contained in X.
- Then **the composition** $f \circ \mathbf{x} : T \to \mathbb{R}$ is differentiable at any point t_0, and,

$$\frac{df}{dt}(t_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0)\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(\mathbf{x}_0)\frac{dy}{dt}(t_0)$$

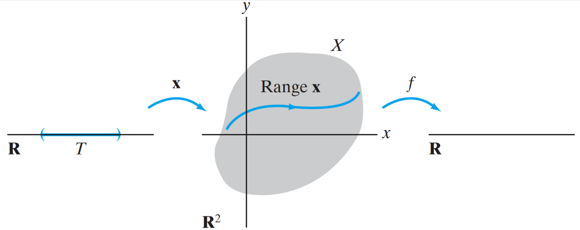
where $x_0 = x(t_0)$.

Notice the mixture of ordinary and partial derivatives appearing in the formula

The Chain Rule in Two Variables

- It helps to think of
  - $\mathbf{x}$ as describing a parametrized curve in $\mathbb{R}^2$, and
  - f as a sort of "temperature function" on X
- The composite $f \circ \mathbf{x}$ is then the **restriction** of f to the **curve**.

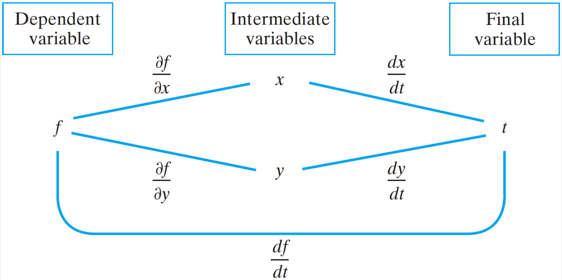
Is the function that measures the temperature along just the curve.



Proposition 5.2

$$\frac{df}{dt}(t_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0)\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(\mathbf{x}_0)\frac{dy}{dt}(t_0)$$

- We can construct an appropriate "variable hierarchy" diagram



- At the intermediate level, f depends on two variables, x and y.
- On the final or composite level, f depends on just a single independent variable t.

Example 2

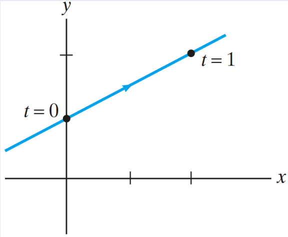
- Let,

$$f(x,y) = \frac{(x+y^2)}{(2x^2+1)}$$

- Suppose f is a temperature function on $\mathbb{R}^2$, and

$$\mathbf{x}(t) = (2t, t+1)$$

- that is a line given in parametric equations, $\mathbf{x}$:



Example 2

- Then

$$f \circ \mathbf{x}(t) = f(\mathbf{x}(t)) = \frac{2t + (t+1)^2}{8t^2+1} = \frac{t^2+4t+1}{8t^2+1}$$

- $f \circ \mathbf{x}$ is the temperature function along the line, and by the quotient rule the **rate of change** of the temperature (per unit change in t) is:

$$\frac{df}{dt} = \frac{4 - 14t - 32t^2}{(8t^2+1)^2}$$

Example 2 (board)

$$f(x, y) = \frac{(x + y^2)}{(2x^2 + 1)}$$

$$\mathbf{x}(t) = (2t, t + 1)$$

- On the other hand, all the hypotheses of Proposition 5.2 are satisfied, and so

$$\frac{\partial f}{\partial x} = \frac{1 - 2x^2 - 4xy^2}{(2x^2 + 1)^2}$$

$$\frac{\partial f}{\partial y} = \frac{2y}{2x^2 + 1}$$

$$\mathbf{x}'(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right) = (2, 1)$$

Example 2

$$
\begin{aligned}
f(x,y) &= \frac{(x+y^2)}{(2x^2+1)}, \quad \mathbf{x}(t) = (2t, t+1) \\
\frac{\partial f}{\partial x} &= \frac{1-2x^2-4xy^2}{(2x^2+1)^2}, \quad \frac{\partial f}{\partial y} = \frac{2y}{2x^2+1} \\
\mathbf{x}'(t) &= \left(\frac{dx}{dt}, \frac{dy}{dt}\right) = (2,1)
\end{aligned}
$$

- Therefore, applying the chain rule and substituting (x,y) by $(2t, t+1)$:

$$
\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \frac{1-2x^2-4xy^2}{(2x^2+1)^2} \cdot 2 + \frac{2y}{2x^2+1} \cdot 1
$$

$$
= \frac{2(1-8t^2-8t(t+1)^2)}{(8t^2+1)^2} + \frac{2(t+1)}{8t^2+1} = \frac{2(2-7t-16t^2)}{(8t^2+1)^2}
$$

The Chain Rule when $\mathbf{x}$ is a multidimensional path

- Proposition 5.2 is easy to generalize to the case where f is a function of n variables.
- Suppose $\mathbf{x} : T \subseteq \mathbb{R} \to \mathbb{R}^n$ and $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$.
- The appropriate chain rule formula in this case is,

$$\frac{df}{dt}(t_0) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0)\frac{dx_1}{dt}(t_0) + \frac{\partial f}{\partial x_2}(\mathbf{x}_0)\frac{dx_2}{dt}(t_0) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\frac{dx_n}{dt}(t_0)$$

- It can also be written by using matrix notation,

$$\frac{df}{dt}(t_0) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt}(t_0) \\ \frac{dx_2}{dt}(t_0) \\ \vdots \\ \frac{dx_n}{dt}(t_0) \end{bmatrix}$$

The Chain Rule when $\mathbf{x}$ is a multidimensional path

- It can also be written by using matrix notation

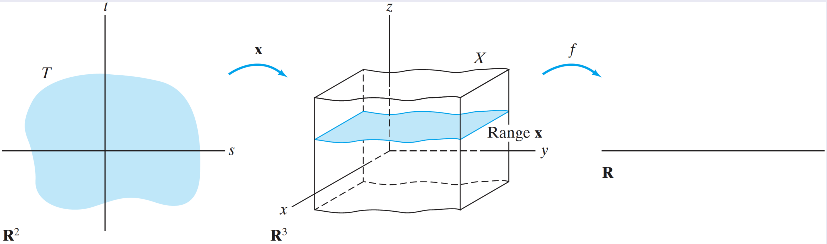
$$\frac{df}{dt}(t_0) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt}(t_0) \\ \frac{dx_2}{dt}(t_0) \\ \vdots \\ \frac{dx_n}{dt}(t_0) \end{bmatrix}$$

- Thus, we have shown

$$\frac{df}{dt}(t_0) = Df(\mathbf{x}_0)D\mathbf{x}(t_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{x}'(t_0)$$

The Chain Rule when $\mathbf{x}$ is a surface

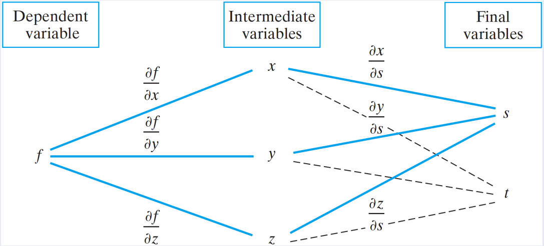
- Suppose X is open in $\mathbb{R}^3$ and T is open in $\mathbb{R}^2$.
- Suppose $f : X \subseteq \mathbb{R}^3 \to \mathbb{R}$ and $\mathbf{x} : T \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ are such that the range of $\mathbf{x}$ is contained in X.
- Then, the composite $f \circ \mathbf{x} : T \subseteq \mathbb{R}^2 \to \mathbb{R}$ can be formed:



- The range of $\mathbf{x}$, $\mathbf{x}(T)$, is just a surface in $\mathbb{R}^3$.
- So $f \circ \mathbf{x}$ can be thought of as an appropriate "temperature function" restricted to this surface.

The Chain Rule when $\mathbf{x}$ is a surface

- Let use $\mathbf{x} = (x, y, z)$ to denote the vector variable in $\mathbb{R}^3$ and $\mathbf{t} = (s, t)$ for the vector variable in $\mathbb{R}^2$.
- We can write a chain rule formula from the next hierarchy diagram:



The Chain Rule when **x** is a surface

- The following formulas holds:

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}$$

Example 3

- Suppose

$$f(x,y,z) = x^2 + y^2 + z^2 \quad \text{and} \quad \mathbf{x}(s,t) = (s\cos t, e^{st}, s^2 - t^2)$$

- Then

$$\begin{aligned} h(s,t) &= f\circ\mathbf{x}(s,t) = s^2\cos^2 t + e^{2st} + (s^2 - t^2)^2 \\ \frac{\partial h}{\partial s} &= \frac{\partial(f\circ\mathbf{x})}{\partial s} = 2s\cos^2 t + 2te^{2st} + 4s(s^2 - t^2) \\ \frac{\partial h}{\partial t} &= \frac{\partial(f\circ\mathbf{x})}{\partial t} = -2s^2\cos t\sin t + 2se^{2st} - 4t(s^2 - t^2) \end{aligned}$$

Example 3

$f(x, y, z) = x^2 + y^2 + z^2$ and $\mathbf{x}(s, t) = (s \cos t, e^{st}, s^2 - t^2)$

- On the other hand

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x, & \frac{\partial f}{\partial y} &= 2y, & \frac{\partial f}{\partial z} &= 2z \\ \frac{\partial x}{\partial s} &= \cos t, & \frac{\partial y}{\partial s} &= te^{st}, & \frac{\partial z}{\partial s} &= 2s \\ \frac{\partial x}{\partial t} &= -s \sin t, & \frac{\partial y}{\partial t} &= se^{st}, & \frac{\partial z}{\partial t} &= -2t \end{aligned}$$

- So for example,

$$\begin{aligned} \frac{\partial h}{\partial s} = \frac{\partial (f \circ \mathbf{x})}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s} \\ &= 2x(\cos t) + 2y(te^{st}) + 2z(2s) \\ &= 2s \cos t(\cos t) + 2e^{st}(te^{st}) + 2(s^2 - t^2)(2s) \\ &= 2s \cos^2 t + 2te^{2st} + 4s(s^2 - t^2) \end{aligned}$$

The Chain Rule in Multiple Variables

$$\mathbf{f} : X \subseteq \mathbb{R}^m \to \mathbb{R}^p, \quad \mathbf{x} : T \subseteq \mathbb{R}^n \to \mathbb{R}^m, \quad h = \mathbf{f} \circ \mathbf{x} : T \subseteq \mathbb{R}^n \to \mathbb{R}^p$$

Then, $\displaystyle \frac{\partial h_i}{\partial t_j} = \sum_{k=1}^{m} \frac{\partial f_i}{\partial x_k} \frac{\partial x_k}{\partial t_j}$, for $i = 1, 2, \ldots, p$ and $j = 1, \ldots, n$

- Knowing that:
  - the ijth entry of the matrix $D\mathbf{h}(\mathbf{t})$ is $\partial h_i / \partial t_j$
  - the ikth entry of the matrix $D\mathbf{f}(\mathbf{x})$ is $\partial f_i / \partial x_k$
  - the kjth entry of the matrix $D\mathbf{x}(\mathbf{t})$ is $\partial x_k / \partial t_j$
- We see that this formula expresses the following equation of matrices

$$D\mathbf{h}(\mathbf{t}) = D(\mathbf{f} \circ \mathbf{x})(\mathbf{t}) = D\mathbf{f}(\mathbf{x}) D\mathbf{x}(\mathbf{t})$$

A very similar expression to
the chain rule in one variable

Example 4

- Suppose $\mathbf{f}: \mathbb{R}^3 \to \mathbb{R}^2$ and $\mathbf{x}: \mathbb{R}^2 \to \mathbb{R}^3$ are given by

$$\begin{aligned} \mathbf{f}(x_1, x_2, x_3) &= (x_1 - x_2, x_1 x_2 x_3) \\ \mathbf{x}(t_1, t_2) &= (t_1 t_2, t_1^2, t_2^2) \end{aligned}$$

- Then $\mathbf{f} \circ \mathbf{x}: \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\mathbf{f} \circ \mathbf{x}(t_1, t_2) = (t_1 t_2 - t_1^2, t_1^3 t_2^3)$$

- So that

$$D(\mathbf{f} \circ \mathbf{x})(\mathbf{t}) = \begin{bmatrix} t_2 - 2t_1 & t_1 \\ 3t_1^2 t_2^3 & 3t_1^3 t_2^2 \end{bmatrix}$$

Example 4

- Suppose $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ and $\mathbf{x} : \mathbb{R}^2 \to \mathbb{R}^3$ are given by

$$\begin{aligned} \mathbf{f}(x_1, x_2, x_3) &= (x_1 - x_2, x_1 x_2 x_3) \\ \mathbf{x}(t_1, t_2) &= (t_1 t_2, t_1^2, t_2^2) \end{aligned}$$

- On the other hand

$$D\mathbf{f}(\mathbf{x}) = \begin{bmatrix} 1 & -1 & 0 \\ x_2 x_3 & x_1 x_3 & x_1 x_2 \end{bmatrix} \quad \text{and} \quad D\mathbf{x}(\mathbf{t}) = \begin{bmatrix} t_2 & t_1 \\ 2t_1 & 0 \\ 0 & 2t_2 \end{bmatrix}$$

- So that, after substituting for x_1, x_2, and x_3, the product matrix is

$$\begin{aligned} D\mathbf{f}(\mathbf{x}) D\mathbf{x}(\mathbf{t}) &= \begin{bmatrix} t_2 - 2t_1 & t_1 \\ x_2 x_3 t_2 + 2 x_1 x_3 t_1 & x_2 x_3 t_1 + 2 x_1 x_2 t_2 \end{bmatrix} \\ &= \begin{bmatrix} t_2 - 2t_1 & t_1 \\ t_1^2 t_2^3 + 2 t_1^2 t_2^3 & t_1^3 t_2^2 + 2 t_1^3 t_2^2 \end{bmatrix} \end{aligned}$$

Theorem 5.3: The (general) Chain Rule

- Suppose X is an open set in $\mathbb{R}^m$ and T is an open set in $\mathbb{R}^n$.
- Suppose functions $\mathbf{f} : X \subseteq \mathbb{R}^m \to \mathbb{R}^p$ and $\mathbf{x} : T \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are defined so that range $\mathbf{x} \subseteq X$.
- Suppose $\mathbf{x}$ is differentiable at $\mathbf{t}_0 \in T$ and $\mathbf{f}$ is differentiable at $\mathbf{x}_0 = \mathbf{x}(\mathbf{t}_0)$.
- Then, the composite $\mathbf{f} \circ \mathbf{x}$ is differentiable at $\mathbf{t}_0$, and

$$D(\mathbf{f} \circ \mathbf{x})(\mathbf{t}_0) = D\mathbf{f}(\mathbf{x}_0) D\mathbf{x}(\mathbf{t}_0)$$

Remark

- Theorem 5.3 requires $\mathbf{f}$ only to be differentiable at the point in question, not to be of class C^1.
- Theorem 5.3 includes all the special cases of the chain rule we have discussed.

Example 5

- Let $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$\mathbf{f}(x, y) = (x - 2y + 7, 3xy^2)$$

- Suppose that $\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^2$ is differentiable at $(0, 0, 0)$.
- Suppose also that we know that

$$\mathbf{g}(0,0,0) = (-2,1) \text{ and } D\mathbf{g}(0,0,0) = \begin{bmatrix} 2 & 4 & 5 \\ -1 & 0 & 1 \end{bmatrix}$$

- We use this information to determine $D(\mathbf{f} \circ \mathbf{g})(0, 0, 0)$.
- Regarding Theorem 5.3, $\mathbf{f} \circ \mathbf{g}$ must be differentiable at $(0, 0, 0)$, and

$$D(\mathbf{f} \circ \mathbf{g})(0,0,0) = D\mathbf{f}(\mathbf{g}(0,0,0))D\mathbf{g}(0,0,0) = D\mathbf{f}(-2,1)D\mathbf{g}(0,0,0)$$

Example 5

$$\begin{aligned} \mathbf{f}(x,y) &= (x-2y+7, 3xy^2) \\ \mathbf{g}(0,0,0) &= (-2,1) \text{ and } D\mathbf{g}(0,0,0) = \begin{bmatrix} 2 & 4 & 5 \\ -1 & 0 & 1 \end{bmatrix} \end{aligned}$$

- Since we know $\mathbf{f}$ completely, it is easy to compute that

$$D\mathbf{f}(x,y) = \begin{bmatrix} 1 & -2 \\ 3y^2 & 6xy \end{bmatrix} \text{ so that } D\mathbf{f}(-2,1) = \begin{bmatrix} 1 & -2 \\ 3 & -12 \end{bmatrix}$$

- Thus

$$D(\mathbf{f} \circ \mathbf{g})(0,0,0) = \begin{bmatrix} 1 & -2 \\ 3 & -12 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 3 \\ 18 & 12 & 3 \end{bmatrix}$$

We did not need to know anything about the differentiability of $\mathbf{g}$ other than at the point $(0,0,0)$