Part III: Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

Introduction to Part III

In Part I of the module we considered functions

$$\mathbf{r} : U \subseteq \mathbb{R} \to \mathbb{R}^n$$

and in Part II we considered functions

$$f : U \subseteq \mathbb{R}^n \to \mathbb{R}$$

In the remainder of the module we consider the general situation

$$f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$$

We shall focus particularly the following important cases.

- When $n = m$ the function f may be viewed as assigning an n-vector to each point in $\mathbb{R}^n$. This is called a vector field on $\mathbb{R}^n$.
- Another view of the case $n = m$ is that the function is a coordinate transformation, or change of coordinates, on R^n.
- In the case $n = 2$ and $m = 3$ the function provides a parametrisation of a two-dimensional surface in $\mathbb{R}^3$.

After introducing vector fields, we will focus on the second case this week. In Week 9 we will consider surfaces. Finally in Week 10 we return to parametrised curves and vector fields.

Week 8: Functions from $\mathbb{R}^n$ to $\mathbb{R}^n$

8.1 Vector Fields

There are many situations in which a vector is associated to each point in some region of space. Familiar examples come from fluid motion such as the wind or the motion of water in a river. Wind has both a magnitude and a direction (blowing North-East at 18 miles/hour) and hence is a vector quantity whose value generally varies with location.

The assignment of a vector to each point in a region of space is a function

$$\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$$

We typically use boldface to denote vector fields, (usually $\mathbf{F}$, $\mathbf{V}$, or $\mathbf{v}$), to emphasise that to each point in U the function $\mathbf{F}$ assigns a vector.

While the majority of physical examples are vector fields on $\mathbb{R}^3$, the general case is for any dimension n

$$\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$$

We will primarily be interested in vector fields on the plane ($n = 2$) or in space ($n = 3$).

A few examples

- The planar vector field,

  $$\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$$

  corresponds to vectors pointing counterclockwise around the origin.

  The vector field,

  $$\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

  corresponds to vectors in space pointing away from the origin.

- An example we have already seen is the gradient of a function of several variables $f(\mathbf{x})$. The gradient is a vector field since $\nabla f(\mathbf{x})$ is a vector whose value depends on the point $\mathbf{x}$. On the plane for example

  $$\mathbf{F}(\mathbf{x}) = \nabla f(\mathbf{x}) = \frac{\partial f}{\partial x}(x, y)\mathbf{i} + \frac{\partial f}{\partial y}(x, y)\mathbf{j}$$

- Systems of ordinary differential equations

$$\begin{aligned}\dot{x}_1 &= f_1(x_1, \cdots, x_n)\\ &\vdots\\ \dot{x}_n &= f_n(x_1, \cdots, x_n)\end{aligned}$$

  can be viewed as vector fields on the corresponding phase space. To each point in phase space $\mathbf{x} = (x_1, \cdots, x_n)$ there is an associated derivative vector,

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) = (f_1(x_1, \cdots, x_n), \cdots, f_n(x_1, \cdots, x_n))$$

- Vector fields arise frequently in applications. Fluid flows, gravitational fields, electric fields, and magnetic fields are all vector fields.

8.2 Coordinate Transformations

A different way in which mappings $\mathbb{R}^n \to \mathbb{R}^n$ arise is coordinate transformations. We motivate the subject by noting the similarity between what is called variable substitution in one-variable integration and changing to polar coordinates in two-variable integration. Both involve the same three steps.

Variable substitution

Consider integrating

$$\int_0^1 f(x)\,dx = \int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx$$

by variable substitution

1. Substitute $x = \sin u$ into the integrand $f(x)$.

$$f(x(u)) = \frac{1}{\sqrt{1-\sin^2 u}} = \frac{1}{\sqrt{\cos^2 u}}$$

2. Change the limits of integration to correspond to those of u

$$\int_{x=0}^{x=1} \quad \to \quad \int_{u=0}^{u=\frac{\pi}{2}}$$

3. Calculate the differential dx in terms of du

$$dx = \frac{dx}{du}du = \cos u du$$

   Thus

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \int_0^{\frac{\pi}{2}} \frac{\cos u}{\sqrt{\cos^2 u}}\,du = \frac{\pi}{2}$$

Polar coordinates

Now consider the double integral of $f(x, y) = \cos(x^2 + y^2)$ over the region

$$\Omega = \{(x, y)| -3 \le x \le 3, 0 \le y \le \sqrt{9 - x^2}\}$$

$$\iint_\Omega f\, dA = \int_{-3}^{3} \int_{0}^{\sqrt{9-x^2}} \cos(x^2 + y^2)\, dy\, dx$$

We can evaluate this double integral be changing to polar coordinates

1. Substitute $x = r\cos\theta$, $y = r\sin\theta$ into the integrand $f(x, y)$.

   $$f(x, y) = \cos(r^2\cos^2\theta + r^2\sin^2\theta) = \cos(r^2)$$

2. Change the limits of integration to those corresponding to Ω the polar coordinates

   $$\int_{-3}^{3} \int_{0}^{\sqrt{9-x^2}} \quad \rightarrow \quad \int_{\theta=0}^{\theta=\pi} \int_{r=0}^{r=3}$$

3. Use the correct area element for polar coordinates

   $$dA = r\, dr\, d\theta$$

   Thus

   $$\int_{-3}^{3} \int_{0}^{\sqrt{9-x^2}} \cos(x^2 + y^2)\, dy\, dx = \int_{\theta=0}^{\theta=\pi} \int_{r=0}^{r=3} \cos(r^2) r\, dr\, d\theta$$

In both cases we re-wrote the integrals in a more useful way by making a **change of variables**, or a **change of coordinates**. We say we **transformed the coordinates**.

8.3 Linear Transformations

You are familiar with linear transformations from Linear Algebra

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m : T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v})$$

This week we will concentrate on the case of the same dimension: $m = n$, where $n = 2$ or $n = 3$. Initially we focus on $n = 2$ later extend to $n = 3$.

T can be expressed as a matrix,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We also think of T in terms of **component functions**

$$x = x(u, v) \quad y = y(u, v)$$

where here

$$x(u, v) = au + bv \quad y(u, v) = cu + dv$$

We previously used this shorthand notation of naming functions by the corresponding variable name. We assume T is invertible and thus we can obtain the component functions of T^{-1}

$$u = u(x, y) \quad v = v(x, y)$$

although we often do not need such explicit expressions.

The transformation is viewed either as a mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ or as a change of coordinates in $\mathbb{R}^2$.

Suppose we want to write a double integral

$$\iint_\Omega f \, dA$$

in a form better suited for integration using a linear transformation T. Assume we figured out which transformation T will help us. The first two steps of the process are:

1. Substitute for x and y in terms of u and v in the integrand

   $$f(x, y) = f(x(u, v), y(u, v)) = g(u, v)$$

   where we have written $g(u, v)$ to emphasise that after the substitution we have a different function of (u, v) than we had of (x, y).

2. Find the region Γ in (u, v) coordinates corresponding to Ω in (x, y) coordinates:

   $$T : \Gamma \to \Omega$$

   This will determine the limits of integration of u and v.

What remains is to express the area element dA in terms of differentials du and dv. Consider a rectangle in the uv-plane with lower left corner at $\mathbf{u}_0 = (u_0, v_0)$, and with sides of length $\triangle u$ and $\triangle v$:

$$\{(u,v)|\, u_0 \le u \le u_0 + \triangle u, v_0 \le v \le v_0 + \triangle v\}$$

This gets mapped by T to a parallelogram in the xy-plane with vertices

$$\begin{array}{ll} \mathbf{x}_0 = T(u_0, v_0) & \mathbf{x}_1 = T(u_0 + \triangle u, v_0) \\ \mathbf{x}_2 = T(u_0, v_0 + \triangle v) & \mathbf{x}_3 = T(u_0 + \triangle u, v_0 + \triangle v) \end{array}$$

The area $\triangle A$ of this parallelogram is given by the cross product of two vectors corresponding to two adjacent sides.

$$\triangle A = \|\mathbf{r}_1 \times \mathbf{r}_2\|$$

where,

$$\mathbf{r}_1 = \mathbf{x}_1 - \mathbf{x}_0, \quad \mathbf{r}_2 = \mathbf{x}_2 - \mathbf{x}_0$$

(For the cross product these are viewed as vectors in $\mathbb{R}^3$ with zero third component.) The reader can show that

$$\mathbf{r}_1 = \big(a\mathbf{i} + c\mathbf{j}\big)\triangle u, \quad \mathbf{r}_2 = \big(b\mathbf{i} + d\mathbf{j}\big)\triangle v$$

Thus

$$\mathbf{r}_1 \times \mathbf{r}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & c & 0 \\ b & d & 0 \end{vmatrix} \triangle u \triangle v = \begin{vmatrix} a & c \\ b & d \end{vmatrix} \triangle u \triangle v\, \mathbf{k}$$

or transposing to bring into standard form

$$\mathbf{r}_1 \times \mathbf{r}_2 = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \triangle u \triangle v\, \mathbf{k}$$

The determinant in the above expression is called the **Jacobian** of the transformation T. In the present case where T is linear, it is just the determinant of the matrix representing the transformation. Denote this determinant by J. Then finally

$$\triangle A = \|\mathbf{r}_1 \times \mathbf{r}_2\| = |J|\, \triangle u \triangle v$$

For infinitesimally small du and dv we have the differential of area

$$dA = |J|\, du\, dv$$

Combining everything, we have the following formula for transforming the double integral

$$\iint_\Omega f dA = \iint_\Gamma f\,(x(u,v), y(u,v))\, |J|\, du\, dv$$

8.4 General transformations in $\mathbb{R}^2$

We now generalise this approach from linear mappings to arbitrary mappings. The main change will be that the formerly constant entries in the Jacobian will now depend on position and be given by partial derivatives of the mapping.

We continue to denote the mapping by T and we assume it maps some region Γ in uv-coordinates to some region of interest Ω in xy-coordinates. We focus on the component functions

$$x = x(u, v) \quad y = y(u, v)$$

which now are not assumed to be linear. We require that the partial derivatives of the component functions exist and are continuous. We also require that T is invertible except possibly on the boundary of Γ.

We proceed as in the linear case. We substitute the change of variables into $f(x, y)$ and then set limits of integration to correspond to the region Γ in (u, v). Finally, the thing we actually need to work out is the element of area dA in terms of $dudv$ for a general transformation.

Consider a rectangle in the uv-plane, which we now imagine to be small

$$\{(u, v) |\, u_0 \leq u \leq u_0 + \triangle u, v_0 \leq v \leq v_0 + \triangle v\}$$

This gets mapped by T to a slightly curved parallelogram in the xy-plane with vertices

$$\begin{array}{ll} \mathbf{x}_0 = T(u_0, v_0) & \mathbf{x}_1 = T(u_0 + \triangle u, v_0) \\ \mathbf{x}_2 = T(u_0, v_0 + \triangle v) & \mathbf{x}_3 = T(u_0 + \triangle u, v_0 + \triangle v) \end{array}$$

The x and y components of $\mathbf{x}_1$ can be approximated by the linear approximation

$$x(u_0 + \triangle u, v_0) \approx x(u_0, v_0) + \frac{\partial x}{\partial u}(u_0, v_0)\triangle u$$
$$y(u_0 + \triangle u, v_0) \approx y(u_0, v_0) + \frac{\partial y}{\partial u}(u_0, v_0)\triangle u$$

or

$$\mathbf{x}_1 \approx \mathbf{x}_0 + \frac{\partial x}{\partial u}\triangle u\mathbf{i} + \frac{\partial y}{\partial u}\triangle u\mathbf{j}$$

where it is understood that the partial derivatives are evaluated at (u_0, v_0). Similarly

$$\mathbf{x}_2 \approx \mathbf{x}_0 + \frac{\partial x}{\partial v}\triangle v\mathbf{i} + \frac{\partial y}{\partial v}\triangle v\mathbf{j}$$

We use the linear approximations to replace the sides of the sides of the slightly curved region by vectors $\mathbf{r}_1$ and $\mathbf{r}_2$,

$$\mathbf{r}_1 = \left(\frac{\partial x}{\partial u}\mathbf{i} + \frac{\partial y}{\partial u}\mathbf{j}\right) \triangle u, \quad \mathbf{r}_2 = \left(\frac{\partial x}{\partial v}\mathbf{i} + \frac{\partial y}{\partial v}\mathbf{j}\right) \triangle v$$

where,

$$\mathbf{r}_1 \approx \mathbf{x}_1 - \mathbf{x}_0, \quad \mathbf{r}_2 \approx \mathbf{x}_2 - \mathbf{x}_0$$

Recall that in the linear case, we had exactly the sides $\mathbf{r}_1$ and $\mathbf{r}_2$ of the of region in xy-coordinates, and these were given in terms of the constants a, b, c, and d of the linear transformation. For the general case, we instead use a linear approximation, i.e. first partial derivatives, to approximate the sides of the region. The partial derivatives then appear where before we had constants. Otherwise, the computation is very similar to the linear case.

The approximate area of the slightly curved region is

$$\triangle A \approx \|\mathbf{r}_1 \times \mathbf{r}_2\|$$

where

$$\mathbf{r}_1 \times \mathbf{r}_2 = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} \triangle u \triangle v \, \mathbf{k}$$

The transpose of the above determinant is denoted

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

This determinant is called the **Jacobian** of the transformation T. It gives the approximate area $\triangle A$ in terms of $\triangle u \triangle v$

$$\triangle A \approx \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \triangle u \triangle v$$

Unlike the linear case it is no longer constant but depends on the point (u, v) at which the partial derivative are evaluated. For infinitesimally small changes in u and v we have

$$\boxed{dA = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du \, dv}$$

Combining everything, we have the following formula for transforming a double integral with a general mapping

$$\iint_\Omega f dA = \iint_\Gamma f\left(x(u,v), y(u,v)\right) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv$$

8.5 Coordinate transformations in $\mathbb{R}^3$

The generalisation to three dimensions is straightforward. Our transformation T will now have three component functions depending on three variables (u, v, w)

$$x = x(u,v,w) \quad y = y(u,v,w) \quad z = z(u,v,w)$$

T will map some region $\Gamma \subseteq R^3$ to some region we care about $\Omega \subseteq R^3$. The first 2 steps in our transformation of the integral do not depend in any essential way on the dimensionality.

To derive the volume element consider a small rectangular box in uvw

$$\{(u,v,w)|\, u_0 \le u \le u_0 + \triangle u, v_0 \le v \le v_0 + \triangle v, \\ w_0 \le w \le w_0 + \triangle w\}$$

This get mapped by T to a slightly curved parallelepiped. The volume $\triangle V$ of this region can be approximated from three vectors approximating edges of the parallelepiped

$$\mathbf{r}_1 = \left(\frac{\partial x}{\partial u}\mathbf{i} + \frac{\partial y}{\partial u}\mathbf{j} + \frac{\partial z}{\partial u}\mathbf{k}\right)\triangle u$$
$$\mathbf{r}_2 = \left(\frac{\partial x}{\partial v}\mathbf{i} + \frac{\partial y}{\partial v}\mathbf{j} + \frac{\partial z}{\partial v}\mathbf{k}\right)\triangle v$$
$$\mathbf{r}_3 = \left(\frac{\partial x}{\partial w}\mathbf{i} + \frac{\partial y}{\partial w}\mathbf{j} + \frac{\partial z}{\partial w}\mathbf{k}\right)\triangle w$$

where

$$\mathbf{r}_1 \approx \mathbf{x}_1 - \mathbf{x}_0, \quad \mathbf{r}_2 \approx \mathbf{x}_2 - \mathbf{x}_0, \quad \mathbf{r}_3 \approx \mathbf{x}_3 - \mathbf{x}_0$$

with

$$\mathbf{x}_0 = T(u_0, v_0, w_0) \qquad \mathbf{x}_1 = T(u_0 + \triangle u, v_0, w_0)$$
$$\mathbf{x}_2 = T(u_0, v_0 + \triangle v, w_0) \quad \mathbf{x}_3 = T(u_0, v_0, w_0 + \triangle w)$$

The volume of the parallelepiped formed from these three vectors is given by their triple product (see Additional Material from Week 3),

$$\triangle V \approx (\mathbf{r}_1 \times \mathbf{r}_2) \cdot \mathbf{r}_3$$

The triple product gives us a determinant of a matrix built out of partial derivatives. The transpose of this is the Jacobian of T for this case and is denoted

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial w} \\ \dfrac{\partial z}{\partial u} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial w} \end{vmatrix}$$

This gives the approximate volume $\triangle V$ in terms of $\triangle u \triangle v \triangle w$

$$\triangle V \approx \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| \triangle u \triangle v \triangle w$$

For infinitesimally small du, dv, dw we have the differential of volume

$$dV = |J| \, du \, dv \, dw$$

Then finally we have

$$\iiint_\Omega f dV = \iiint_\Gamma f \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| du \, dv \, dw$$

where it is understood that one uses T to substitute for (x,y,z) in terms of (u,v,w) in f.

Additional Material

Why the absolute the value of the Jacobian ?

You will notice that in our expressions in generalised coordinates we took the absolute value of the Jacobian

$$|J|\, du\, dv, \qquad \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv, \qquad \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| du\, dv\, dw$$

This seems to be the convention used in text books at this level. It is not strictly necessary, but one must properly take into account limits for integration if one does not include the absolute value. We will use the absolute value together with the following rule about setting up integrals:

Always use limits of integration such that the lower limit is smaller than the upper limit..

For example, in

$$\iint_\Omega f dA = \iint_\Gamma f\left(x(u,v), y(u,v)\right) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv$$

if Γ is the rectangular region $0 \le u \le 1$, $-1 \le v \le 1$, then express $\iint_\Gamma$ as repeated integrals $\int_0^1 \int_{-1}^1$ $\int_{-1}^1 \int_0^1$ and not $\int_1^{-1} \int_0^1$ or $\int_{-1}^1 \int_1^0$.

Derivative matrix and the general chain rule

Consider now a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, where n and m are arbitrary. Express $\mathbf{f}$ in terms of m component functions each depending on n real values,

$$\begin{aligned} &f_1(x_1, x_2, \cdots, x_n), \\ &\vdots \\ &f_j(x_1, x_2, \cdots, x_n), \\ &\vdots \\ &f_m(x_1, x_2, \cdots, x_n) \end{aligned}$$

Now differentiate each component function with respect to each independent variable and arrange these in a matrix. The matrix is known as the **derivative matrix** of the function $\mathbf{f}$. It is commonly denoted $\mathbf{Df}$,

$$\mathbf{Df} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

Comments

- You will recognise special cases we have already encountered. A vector function $\mathbf{r}(t)$ corresponds to $\mathbf{f} : \mathbb{R} \to \mathbb{R}^m$. The derivative matrix has one column, the components of $\mathbf{r}'(t)$. (When a function depends on only one variable, the partial derivatives become ordinary derivatives.)

  A function of several variables corresponds to the case $f : \mathbb{R}^n \to \mathbb{R}$. The derivative matrix has one row, the elements of the gradient vector ∇f.

The general case includes also the derivative of a function of one variable, $f : \mathbb{R} \to \mathbb{R}$. The derivative matrix is just a 1×1 matrix.

In the context of transformations, we encountered functions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and we computed the elements of the derivative matrix. The Jacobian of the transformation we now see as the determinant of the derivative matrix. In the context of transformations the derivative matrix is often called the **Jacobian matrix** of the transformation.

- $\mathbf{Df}$ is itself a function of the independent variables and this may be indicated by $\mathbf{Df}(x_1, \cdots, x_n)$ or by $\mathbf{Df}(\mathbf{x})$.
- A function is differentiable if it can be approximated by a linear map (in a certain precise sense which you will learn in future modules). You know from Linear Algebra that there is a correspondence between linear maps and matrices. The derivative matrix is the matrix representation of the linear map corresponding to the derivative.

Finally, we return once again to the Chain Rule. Suppose now that we want to compose two functions $\mathbf{h}$ and $\mathbf{f}$ to form a function $\mathbf{g} = \mathbf{f} \circ \mathbf{h}$ and then take its derivative. We need agreement between the domain of $\mathbf{f}$ and the range of $\mathbf{h}$ and we need differentiability, but we will assume this. The question is, what is the derivative of $\mathbf{g}$ in terms of the derivatives of $\mathbf{f}$ and $\mathbf{h}$? The answer is given by the general Chain Rule.

Formally, let $\mathbf{h}$ be a differentiable function $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^p$ and let $\mathbf{f}$ be a differentiable function $\mathbf{f} : \mathbb{R}^p \to \mathbb{R}^m$. Compose these to obtain $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$, where $\mathbf{g} = \mathbf{f} \circ \mathbf{h}$. The derivative of $\mathbf{g}$ is given by

The **General Chain Rule**.

$$\mathbf{Dg}(\mathbf{x}) = \mathbf{Df}(\mathbf{h}(\mathbf{x}))\,\mathbf{Dh}(\mathbf{x})$$

where $\mathbf{Dg}(\mathbf{x})$ is an $n \times m$ derivative matrix, $\mathbf{Df}(\mathbf{h}(\mathbf{x}))$ is an $n \times p$ derivative matrix, and $\mathbf{Dh}(\mathbf{x})$ is an $p \times m$ derivative matrix.

The Chain Rule is expressed very simply as the product of derivative matrices. It is frequently expressed in component form using summation

$$\frac{\partial g_i}{\partial x_j} = \sum_{k=1}^{p} \frac{\partial f_i}{\partial h_k} \frac{\partial h_k}{\partial x_j}$$

where $\dfrac{\partial f_i}{\partial h_k}$ is understood to mean the derivative of f_i with respect to its k^{th} argument evaluated at the appropriate point $\mathbf{h}(\mathbf{x})$. *Recall, I warned you about possible confusion with the Chain Rule because of compact notation.*

In the case $n = m = 1$ the general Chain Rule reduces to the special case considered in Week 4. Using the notation from Week 4 where $g = f \circ \mathbf{r}$, this special case is the following matrix product

$$\underbrace{\begin{bmatrix} \dfrac{dg}{dt} \end{bmatrix}}_{1\times 1} = \underbrace{\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_p} \end{bmatrix}}_{1\times p} \underbrace{\begin{bmatrix} \dfrac{dx_1}{dt} \\ \vdots \\ \dfrac{dx_p}{dt} \end{bmatrix}}_{p\times 1}$$

(Previously n was used where here p appears. We have suppressed where the various derivatives are evaluated.) Which could be written as before as a sum or using the dot product

$$\frac{dg}{dt} = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt}, \qquad \frac{dg}{dt} = \nabla f \cdot \frac{d\mathbf{r}}{dt}$$