1. The big picture: why study maps $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$?

Although in this course we will focus on the calculus of functions of two and three variables (i.e we study $f : \mathbb{R}^n \to \mathbb{R}$ for n=2,3), along the way we will need to understand more general maps $f : \mathbb{R}^n \to \mathbb{R}^m$. Let me try to explain a little why this is the case and some of the most important examples.

Take for example the case $n = 1$ and $m = 2\,, 3$. Then we are studying $\vec{f}(t)$, $a \leq t \leq b$ and the image of f is a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$. For general m, the image is a curve in $\mathbb{R}^m$. We will say that the image curve is regular if $\vec{f}'(t) \neq 0$. This means that the image curve always has a well defined tangent direction.

Now let $n = 2$ and $m = 3$ and $\vec{X}(u, v) : D \subset \mathbb{R}^2 \to \mathbb{R}^3$. We can think of the image $\vec{X}(D)$ as a "surface patch" (i.e a piece of a surface in $\mathbb{R}^3$) and by choosing different patches we are "parametrizing" (representing by good coordinates pieces of a larger surface. For example we may think of the unit sphere S^2 in $\mathbb{R}^3$ which is the nicest example of a compact surface in $\mathbb{R}^3$ without boundary. Then as early cartographers understood, it is topologically impossible to represent S^2 by a single patch. We will also want $\vec{X}$ to be continuously differentiable, i.e the partial derivatives $\vec{X}_u$ and $\vec{X}_v$ exist and are continuous. We say that the patch $\vec{X}(D)$ is regular if $\vec{X}_u$ and $\vec{X}_v$ are linearly independent vectors in $\mathbb{R}^3$. This condition means that the surface patch has a well defined tangent plane that changes continuously as the parameters (u, v) changes continuously.

One of the main interpretations of the case $n = m = 2$ is that of a vector field in $\mathbb{R}^2$ and of $n = m = 3$ is a vector field in $\mathbb{R}^3$. We visualize a vector field as a vector (arrow) emanating from the point where it is defined. For example, we have already seen the vector field $\vec{F}(x, y) = -y\hat{i} + x\hat{j}$ which is associated with a rotation. Vector fields in $\mathbb{R}^3$ may represents forces or velocity vectors of a steady fluid, or many other fields of physical or geometric interest.

Another reason to study the case $n = m = 2$ or $n = m = 3$ is to introduce a change of coordinates that better suits the geometry of the problem we are studying. For example we may want to introduce polar coordinates $x = r\cos\theta$, $y = r\sin\theta$. Then $\vec{f}(r,\theta) =< r\cos\theta, r\sin\theta >$ is the polar coordinate mapping. For example let D be the rectangle $0 \le r \le R$, $0 \le \theta \le \frac{\pi}{4}$. Then the image $\vec{f}(D)$ is a sector of the disk of radius R of angle $\frac{\pi}{4}$ in the first quadrant.

When $n = m = 3$ we may want to introduce cylindrical coordinates $x = r\cos\theta$, $y = r\sin\theta\,, z = z$ which are analogous to the polar coordinates in the plane. Then we may define a cylindrical box $B = \{(r,\theta,z) : 0 \le r \le R,\, 0 \le \theta_1 \le \theta \le \theta_2 \le 2\pi,\, z_1 \le z \le z_2\}$ and the cylindrical coordinate mapping

$$\vec{f}(r,\theta,z) =< r\cos\theta, r\sin\theta, z >$$

with image $\vec{f}(B)$ is the portion of the standard cylinder of radius R in $\mathbb{R}^3$ whose cylindrical coordinates satisfy $0 \le r \le R$, $0 \le \theta_1 \le \theta \le \theta_2 \le 2\pi$, $z_1 \le z \le z_2\}$. In other words, it is that part of the cylinder of radius R that is contained between the planes $z = z_1$ and $z = z_2$ and lying in the sector defined by $0 \le r$, $0 \le \theta_1 \le \theta \le \theta_2$ (this is a wedge shaped region).

Another interesting geometry (again $n = m = 3$) is that defined by spherical coordinates

$$x = \rho\sin\phi\cos\theta,\; y = \rho\sin\phi\sin\theta,\; z = \rho\cos\phi\ .$$

Here $\rho \ge 0$ is $|\vec{X}|$ where $\vec{X} =< x, y, z >$ is the position vector, $0 \le \phi \le \pi$ is the angle between $\vec{X}$ and $\hat{k}$ and $0 \le \theta < 2\pi$ is the usual polar angle of (x, y) in the plane $z = 0$. These coordinates are closely related to longitude and latitude with the difference being that the latitude angle $\phi' = \phi - \frac{\pi}{2}$ varies between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ (i.e ϕ is the “co-latitude”.

We will study all of these changes of coordinates in more detail when we study double and triple integrals. In particular we need to understand how area and volume change infinitesimally, that is up to first order. That is where the derivative map $D\vec{f}$ comes in as the linear map which gives the first order Taylor approximation to the "nonlinear map" $\vec{f}$. We will define what we mean by $D\vec{f}$ in another installment of these notes. Instead we now discuss in more detail what we mean by a linear transformation and how such a linear transformation changes volume.

2. Linear transformations

We say that a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping if

$$T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v}) \ ,$$

for all scalars a and b. Necessarily, $T(\vec{0}) = \vec{0}$.

Example 2.1. Fix standard coordinate systems in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and represent vectors in these spaces by column vectors. Let $T(\vec{X}) = A\vec{X}$ where A is an mxn matrix and $A\vec{X}$ is matrix multiplication, i.e the ith component of the column vector $A\vec{X}$ is $(A\vec{X})_i = \sum_{j=1}^n a_{ij}x_j$, $1 \leq i \leq m$ where $x_1, \ldots, x_n$ are the components of $\vec{X}$. Because matrix multipication as defined above is linear, it is easily verified that T is linear.

Note however that it is not necessary to choose coordinate systems in the spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ to define linear mappings.

Example 2.2. Let $\{\vec{v}_1, \ldots, \vec{v}_m\}$ be m vectors in $\mathbb{R}^n$. Define

$$T(\vec{X}) = (\vec{X} \cdot \vec{v}_1, \ldots, \vec{X} \cdot \vec{v}_m) \ .$$

(Here we think of $\mathbb{R}^n$ and $\mathbb{R}^m$ as ordered n and m tuple respectively and define the dot product in the usual way). Then because dot product is a linear operation, T is linear.

However, once we choose arbitrary bases $\vec{v}_1, \ldots, \vec{v}_n$ for $\mathbb{R}^n$ and $\vec{w}_1, \ldots, \vec{w}_m$ for $\mathbb{R}^m$ (not necessarily orthonormal), then any $\vec{v} \in \mathbb{R}^n$ has a unique representation

$$\vec{v} = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$$

and

$$T(\vec{v}) = T(c_1\vec{v}_1 + \ldots + c_n\vec{v}_n) = c_1T(\vec{v}_1) + \ldots c_nT(\vec{v}_n) \ .$$

Therefore T is uniquely determined by $T(\vec{v}_1), \ldots, T(\vec{v}_n)$ and

$$T(\vec{v}_j) = a_{1j}\vec{w}_1 + \ldots + a_{mj}\vec{w}_m \ .$$

Thus T is uniquely determined by the mxn matrix (a_{ij}), $1 \leq i \leq m$, $1 \leq j \leq n$ and $T(\vec{X}) = A\vec{X}$.

3. Change of volume for linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, $n = 2, 3$.

Now let $\vec{v}$ and $\vec{w}$ be linearly independent vectors in $\mathbb{R}^3$ and consider the parallelogram $P = \{s\vec{v} + t\vec{w},\ 0 \leq s \leq 1,\ 0 \leq t \leq 1\}$ that they span. Suppose that $T\mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation. Then since $T(s\vec{v} + t\vec{w}) = sT(\vec{v}) + tT(\vec{w})$, the image $T(P)$ is again a parallelogram in $\mathbb{R}^3$ and its area is $|T(\vec{v}) \times T(\vec{w})|$.

Example 3.1. Given vectors $\vec{A} =< a_1, a_2 >$ and $\vec{B} =< b_1, b_2 >$ in the plane, there exists a unique linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ such that $L(\hat{i}) = \vec{A}$ and $L(\hat{j}) = \vec{B}$, with associated matrix

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} .$$

We say $\det(L) = \det(\vec{A}, \vec{B}) = a_1 b_2 - a_2 b_1$ is the determinant of the associated matrix. Then if S is the unit square spanned by $\hat{i}$ and $\hat{j}$, and if P is the parallelogram spanned by $\vec{A}$ and $\vec{B}$, then $L(S) = P$ and

$$\text{Area(P)} = |\det L| \ .$$

Now let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be another linear map. Then the composition $L_1 = T \circ L$ is again a linear map (check this) and

$$T(P) = T(L(S)) = L_1(S) \ .$$

Moreover,

$$\text{Area T(P)} = |\det L_1| = |\det T \det L| = |\det L|\,\text{Area(P)}$$

since the matrix representing L_1 is the composition (matrix multiplication) of the matrix representing T and L and the determinant is multiplicative under composition.

Now consider a parallelpiped P in $\mathbb{R}^3$ spanned by linearly independent vectors $\vec{A}$, $\vec{B}$, $\vec{C}$. Then we have seen that the volume of P is

$$\text{Vol(P)} = |\det(\vec{C}, \vec{A}, \vec{B})| = |\det(\vec{A}, \vec{B}, \vec{C}) = |\vec{A} \times \vec{B} \cdot \vec{C}| \ .$$

Now let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear map and agin define $\det L$ to be the determinant of the associated matrix. Now let S denote the unit cube spanned by the units vectors $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ and let L be the unique linear map defined by

$$L(\hat{\imath}) = \vec{A},\ L(\hat{\jmath}) = \vec{B},\ L(\hat{k}) = \vec{C}\ .$$

Then $L(S) = P$ and Vol(P) $= \det L$. Moreover,

Theorem 3.2. *Vol(T(P)* $= |\det T|$ *Vol(P).*

Later on, we will define $D\vec{f}$, the derivative of a nonlinear map and show that it is a linear map T. This will allow us to define the associated matrix to $D\vec{f}$, namely the Jacobian matrix and show that its determinant gives the infinitesimal volume distortion.