Warning: These notes are not meant to be a replacement for lecture. They are incomplete, likely contain many typos, and do not contain the explanatory pictures I draw on the board during lecture.

1 Curves, tangent vectors, the chain rule—Wednesday, Sep 12

1.1 Parameterized Curves

A parameterized curve in $\mathbb{R}^n$ is a continuous function from an interval to $\mathbb{R}^n$.

$$r: (a, b) \to \mathbb{R}^n, \ t \mapsto (r_1(t), \ldots, r_n(t)).$$

It can be convenient to think of the output as a vector.

Example. Consider the curve $r(t) = (\cos t, \sin t, t)$ in $\mathbb{R}^3$ (I have not specified the exact interval for the domain). Try drawing this curve (it is a helix). To help visualize, it can be helpful to project the graph onto the coordinate planes.

If $r_1(t), \ldots, r_n(t)$ are differentiable at $t_0$, then the tangent vector of $r$ at $t_0$ is

$$r'(t_0) = \lim_{h \to 0} \frac{r(t+h) - r(t)}{h} = \langle r'_1(t_0), \ldots, r'_n(t_0) \rangle.$$

If $r(t)$ is thought of as the position of an object at time $t$, then $r'(t)$ can be thought of as a velocity vector, and $|r'(t)|$ is the speed at time $t$.

Example. If $y = f(x)$ is the graph of a differentiable function in $\mathbb{R}^2$, then $r(t) = (t, f(t))$ is a parameterization of this graph, and $r'(t) = (1, f'(t))$ is a vector with slope $f'(t)$ at the point $(t, f(t))$.

The unit tangent vector

$$T(t) = \frac{r'(t)}{|r'(t)|},$$

defined where $r'(t)$ does not vanish, is independent (up to sign) of the parameterization $r$ of the curve. We will see this more later when we integrate along curves.

What is the connection between $r'(t)$ and the derivative $df_t$ that we discussed last time? Since

$$\lim_{h \to 0} \frac{|r(t-h) - r(t) - hr'(t)|}{|h|} = \lim_{h \to 0} \frac{r(t-h) - r(t)}{h} - r'(t) \bigg| = 0,$$

we see that $hr'(t) = dr_t(h)$. By linearity of $dr_t$ this is equal to $hdr_t(1)$, and so $r'(t) = dr_t(1)$.

1.2 The Chain Rule

Let $r(t) = (r_1(t), \ldots, r_n(t))$ be a parameterized curve in $\mathbb{R}^n$, and $F: \mathbb{R}^n \to \mathbb{R}^m$ a differentiable function. Then $F \circ r$ is also a parameterized curve in $\mathbb{R}^m$.

Theorem 1 (Chain Rule). The function $F \circ r$ is differentiable, and

$$(F \circ r)'(t) = dF_{r(t)}(r'(t)) = dF_{r(t)}dr_t(1).$$

You can think of this as saying that “$dF$ sends tangent vectors to tangent vectors.”
Proof sketch. Since $F$ is differentiable, the function $v \mapsto F(x + v) - F(x)$ is approximated to first order by the linear transformation $v \mapsto dF_x(v)$. So

$$\frac{(F \circ r)(t + h) - (F \circ r)(t)}{h} = \frac{F \left( r(t) + h \frac{r(t+h) - r(t)}{h} \right) - F(r(t))}{h} \approx dF_{r(t)} \left( \frac{r(t + h) - r(t)}{h} \right)$$

when $h$ is small. As we let $h \to 0$, this becomes $dF_{r(t)}(r'(t))$.

\[\square\]

1.3 Tangent planes

The tangent plane of the graph $z = f(x, y)$ through $P = (a, b, f(a, b))$ is given by

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Note that a normal vector to this plane is $n = \langle f_x(a, b), f_y(a, b), -1 \rangle$.

Let us justify this formula. What should the tangent plane be? It should be the plane containing $P$ which is parallel to any tangent vectors at $P$.

For any $v = \langle v_1, v_2 \rangle$, $r(t) = \langle a + tv_1, b + tv_2 \rangle$ is a curve in $\mathbb{R}^2$, then $F \circ r$ is a curve in the graph of $f$, with $(F \circ r)(0) = (a, b, f(a, b))$ and $(F \circ r)'(0)$ tangent to $z = f(x, y)$ at this point. From the chain rule, this tangent vector is

$$(F \circ r)'(0) = dF_{r(0)}(r'(0)) = dF_{r(0)}(v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$  

The normal vector $n$ above is precisely the thing orthogonal to these two vectors.

Note that the tangent plane consist of all points of the form $P + dF_P(v)$ for $v \in \mathbb{R}^n$.

**Theorem 2** (Chain rule, general). If $F$ is differentiable at $a$, $G$ is differentiable at $F(a)$, and $G \circ F$ is defined near $a$, then $G \circ F$ is differentiable at $a$ and

$$d(G \circ F)_a = dG_{F(a)}dF_a$$

Proof. For any $v$ in $\mathbb{R}^n$, and $a$ in the domain of $F$, let $r$ be such that $r(0) = a$ and $r'(0) = v$. For example, $r(t) = a + tv$ would work. Then

$$d(G \circ F)_a(v) = d(G \circ F)_{r(0)}(r'(0))$$
$$= (G \circ F \circ r)'(0)$$
$$= dG_{F(r(0))}(F \circ r)'(0)$$
$$= dG_{F(a)}(dF_a(v)) = (dG_{F(a)}dF_a)(v).$$  

\[\square\]