Lecture 3-12

We now return to vector algebra and calculus. We should all know by now how to add and subtract vectors in $\mathbb{R}^n$ and multiply vectors by scalars; this term we learned how to compute the dot product $\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i b_i$ of two vectors $\vec{a} = (a_1, \ldots, a_n), \vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$, and the cross product $\vec{a} \times \vec{b} = (a_2, b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$ if $n = 3$. Dot products are the fundamental tool we need to do geometry in $\mathbb{R}^n$, as we have the formula $\cos\theta = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}||||\vec{b}||}$ for the angle θ between nonzero vectors $\vec{a}, \vec{b}$ (for any n), taking $\theta \in [0, \pi]$; this formula makes sense by the Cauchy-Schwarz inequality, which asserts that the fraction occurring in it indeed has absolute value at most 1. In particular the vectors $\vec{a}, \vec{b}$ are orthogonal if and only if their dot product is 0; one of the most useful properties of the cross product $\vec{a} \times \vec{b}$ if $n = 3$ is that it is automatically orthogonal to both $\vec{a}$ and $\vec{b}$.

A particularly simple (but also particularly important and useful) parametrized curve in $\mathbb{R}^n$ is a line $\vec{p} + t\vec{v}$; here $\vec{p}, \vec{v}$ are constant vectors in $\mathbb{R}^n$ with $\vec{v} \neq \vec{0}$. In this setup $\vec{p}$ is a point on the line (any point; $\vec{p}$ is not uniquely determined by the line) and $\vec{v}$ is a direction vector for the line (it too is not unique and could be replaced by any nonzero multiple of itself). We can determine whether two lines or more generally two parametrized curves $\vec{r}(t), \vec{s}(u)$ intersect by setting $\vec{r}(t) = \vec{s}(u)$ and determining whether there are two values of the parameters t, u such that all coordinates of $\vec{r}(t)$ equal their counterparts in $\vec{s}(u)$. Since the vector equation $\vec{r}(t) = \vec{s}(u)$ amounts to a system of n equations in just two variables t, u, we see that most pairs of parametrized curves in $\mathbb{R}^n$ do not intersect if $n > 2$, but some do. If moreover we have $\vec{r}(t) = \vec{s}(t)$, so that the curves $\vec{r}(t), \vec{s}(u)$ intersect at a point with the same values of the parameters t, u, then we say the curves collide at this point; of course it is rare even for two intersecting curves to collide. We also looked at planes in $\mathbb{R}^3$; it is most convenient to give equations for these rather than parametrizing them. Any plane in $\mathbb{R}^3$ (or more generally a hyperplane in $\mathbb{R}^n$ is completely determined by a point $\vec{p}$ lying in the plane together with a normal vector $\vec{n}$ for it; as with direction vectors for lines the normal vector $\vec{n}$ cannot be 0, but can be equivalently replaced by any nonzero multiple of itself. The equation of the plane passing through $\vec{p}$ with normal $\vec{n}$ is then $\vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{p}$, or, writing it out in coordinates, $\sum a_i x_i = \sum n_i p_i$, where $x_1, \ldots, x_m$ are the variables (and the coordinates of $\vec{x}$), $\vec{p} = (p_1, \ldots, p_m)$, and $\vec{n} = (n_1, \ldots, n_m)$. The angle between a (hyper)plane and a line is $\pi/2$ minus the angle between a normal vector for the (hyper)plane and the line.

The product rule carries over to dot and cross products of parametrized curves, so that $(d/dt)(\vec{r}(t) \cdot \vec{s}(t)) = \vec{r}'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}'(t)$ for $\vec{r}(t), \vec{s}(t) \in \mathbb{R}^n$, and similarly with the dot product replaced by the cross product, if $n = 3$. In particular if $||\vec{r}(t)||$ is constant (so that $\vec{r}(t)$ has constant speed) then the dot product $\vec{r}(t) \cdot \vec{r}'(t)$ is constantly equal to 0, so that the tangent vector of any such curve is always orthogonal to its position vector.

We wind up our review with real-valued functions of several variables, the last topic of the course. Given a function $f : \mathbb{R}^n \to \mathbb{R}$ its ith partial derivative $f_i = \partial f / \partial x_i$ is obtained by differentiating f with respect to x_i in the usual way, treating other variables as constants. If all partials of f exist at a point $\vec{a} \in \mathbb{R}^n$ then we set $\nabla f(\vec{a}) = (\partial f/\partial x_1(\vec{a}), \ldots, \partial f/\partial x_n(\vec{a}))$ and call this vector the gradient of f at $\vec{a}$. Thus if $\vec{f}$ is differentiable at $\vec{a}$ then its gradient must exist there; but this necessary condition is not sufficient. If ∇f is defined in a neighborhood of $\vec{a}$ and in addition is continuous at $\vec{a}$ then f is always differentiable (and in fact continuously differentiable) at $\vec{a}$. In this case the directional derivative $D_{\vec{u}} f(\vec{a})$ of f at $\vec{a}$ in the direction of $\vec{u}$ (a unit vector, by definition) is defined by $\lim_{t \to 0} \frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t}$ and computed by $D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}$. In particular, the gradient $\nabla f(\vec{a})$ has the direction of maximal rate of increase of f at $\vec{a}$ (if this rate of increase is nonzero in at least one direction) and its magnitude equal the rate of increase of f in that direction.

The formula for the directional derivative admits a massive generalization, to the chain rule in several variables. This asserts that if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a} \in \mathbb{R}^n$ and if $\vec{r}(t)$ is a parametrized curve with $\vec{r}(t_0) = \vec{a}$, then the composite function sending t to $f(\vec{r}(t))$ is differentiable at $t = t_0$ and its derivative there is $\nabla f(\vec{a}) \cdot \vec{r}'(t_0)$. In formulas, if a variable z depends differentiably on variables $y_1, \ldots, y_n$ and if each y_i depends differentiably on $x_1, \ldots, x_m$, then z (now regarded as a composite function) also depends differentiably on each x_i and $\partial z/\partial x_i = \sum_{j=1}^{n} (\partial z/\partial y_j)(\partial y_j/\partial x_i)$. Note that the sum on the right hand side can be interpreted either as a sum of products or as a single dot product.