Lecture 3-4

A word about notation: given a function $f : \mathbb{R}^n \to \mathbb{R}$, the definition in the book for f to be differentiable at $\vec{a} \in \mathbb{R}^n$ is that $f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + o(\vec{h})$. This means exactly the same thing as my definition that $\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h})-f(\vec{a})-\nabla f(\vec{a})\cdot\vec{h}}{||\vec{h}||} = 0$, since $o(\vec{h})$ refers to a scalar quantity such that $\frac{o(\vec{h})}{||\vec{h}||} \to 0$ as $\vec{h} \to \vec{0}$.

Now we give an example. If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $f(x, y, z) = (xyz, xy, \sin z)$, then f is differentiable everywhere and in particular at the point $(1, 1, \pi)$. Its derivative there has matrix $\begin{pmatrix} \pi & \pi & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, since the first row of this matrix at any point (x_0, y_0, z_0) is the gradient of $g(x, y, z) = xyz$ at that point, or (y_0z_0, x_0z_0, x_0y_0), and similarly for the other rows. This matrix, regarded as a linear function from $\mathbb{R}^3$ to itself, sends the vector (x_1, x_2, x_3) to the vector with coordinates $\pi x_1 + \pi x_2 + x_3, x_1 + x_2, -x_3$, this vector being the one that you get on multiplying the matrix by $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and using the definition of the product of a matrix and a column vector.

Continuing now from last time, suppose we have a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. We consider the set S of points where this function takes a fixed value, say c, and call this a *level set* of the function f. Then the chain rule implies that if $\vec{r}(t)$ is any parametrized curve lying entirely in S, then $\nabla f(\vec{r}(t)) \cdot \vec{r}'(t) = 0$ for all t, since $f(\vec{r}(t))$ is constant. Thus *the tangent vector to any such curve at any point in the level set is orthogonal to the gradient of the function f defining the level set at that point*. Having previously defined hyperplanes in $\mathbb{R}^n$ as the sets of points defined by a single equation $\vec{a} \cdot (x_1, \ldots, x_n) = d$ for some constant d, we now define the *tangent hyperplane of the level set S at $\vec{a} \in S$ to be the set of points $(x_1, \ldots, x_n)$ specified by the equation $\nabla f(\vec{a}) \cdot (x_1, \ldots, x_n) = \nabla f(\vec{a}) \cdot \vec{a}$*, provided that $\nabla f(\vec{a}) \neq \vec{0}$; if $\nabla f(\vec{a}) = \vec{0}$, then we say that the tangent hyperplane os S at $\vec{a}$ is undefined. We call $\nabla f(\vec{a})$ the *normal vector* to the tangent hyperplane at that point. Thus the tangent to a level set at any point contains that point and also the tangent vector at that point to any curve lying in the level set and passing through that point. As an example, the level sets of the function $f(x, y, z) = x^2 + y^2 + z^2$ are either empty (if the constant c is negative, or one point (if it is 0), or spheres (if it is positive). The tangent plane to the unique sphere centered at $(0, 0, 0)$ and passing through $(x_0, y_0, z_0) \neq (0, 0, 0)$ has as normal vector $(2x_0, 2y_0, 2z_0)$, or any nonzero multiple of this vector, e.g. (x_0, y_0, z_0). We know from the geometry of $\mathbb{R}^3$ that the unique plane passing through $(x_0, y_0, z_0) \neq \vec{0}$ with normal vector pointing directly away from (or directly toward) the origin intersects this sphere containing this point exactly at this point and no other. (Note that the degenerate case of the point $(x_0, y_0, z_0) = (0, 0, 0)$, which lies in a level set consisting of just this point and no other, is the unique case where $\nabla f(x_0, y_0, z_0) = \vec{0}$.) We also know that if we are at some point (x_0, y_0, z_0) and we want to get as far away and as quickly from the origin as possible, then of course we should walk directly away from the origin. More generally, if we are on a level set of a function f and we want to increase the value of f as quickly as possible, then we should move orthogonally away from the level set containing the point where we are, in the direction where f increases.

If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a} \in \mathbb{R}^n$ and if $\nabla f(\vec{a}) \neq \vec{0}$, then the level set $S = \{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = f(\vec{a})\}$ is said to have a *smooth* or *nonsingular* point at $\vec{a}$; otherwise (if $\nabla f(\vec{a}) = \vec{0}$) we call $\vec{a}$ a *singular* point of the level set S. We just saw that the tangent hyperplane of $\vec{a} \in S$ is defined exactly when $\vec{a}$ is nonsingular. If all points of S are smooth then we call S itself smooth. An important example of a smooth level set is the graph of a differentiable function $x_n = f(x_1, \ldots, x_{n-1})$. This graph is defined by the equation $g(x_1, \ldots, x_n) = f(x_1, \ldots, x_{n-1}) - x_n = 0$; we have $\partial g/\partial x_n = -1 \neq 0$ at any point on this graph. Consequently the tangent hyperplane to this graph is always defined; it has equation $\partial f/\partial x_1(\vec{a})(x_1 - a_1) + \ldots + \partial f/\partial x_{n-1}(\vec{a})(x_{n-1} - a_{n-1}) - (x_n - a_n) = 0$ at the point $(a_1, \ldots, a_n)$, where $\vec{a} = (a_1, \ldots, a_{n-1})$ and $a_n = f(\vec{a})$. The vector whose first $n-1$ coordinates are given by $\nabla f(\vec{a})$ and whose last coordinate is -1 is called the *downward normal* of the graph at this point, regarding this graph as a level set; the negative of this vector is unsurprisingly called the upward normal.

Thus for example the tangent plane to the elliptic cone with equation $x^2 + 4y^2 = z^2$ at the point $(3, 2, 5)$ has normal vector $\nabla(x^2 + 4y^2 - z^2)(3, 2, 5) = (6, 16, -10)$ (Example 5 on p. 824 of the text). The equation of the tangent plane is $3(x_3) + 8(y-2) - 5(z-5) = 0$, which simplifies to $3x + 8y - = 5z = 0$; note that we divided the normal vector by 2 to write down this equation (this is harmless). Parametric equations for the normal line are given by $(3 + 3t, 2 + 8t, 5 - 5t)$. The only singular (=nonsmooth) point on this elliptic cone is its vertex $(0, 0, 0)$; it is clear even intuitively that the cone has no tangent plane at that point. The graph of $z = g(x, y) = \ln(x^2 + y^2)$ at the point $(-2, 1, \ln 5)$ has normal vector $(\partial g/\partial x, \partial g/\partial y, -1)(-2, 1, \ln 5) = (-4/5, 2/5, -1)$ and equation $z - \ln 5 = -(4/5)(x + 2) + (2/5)(y - 1)$ (Example 7 on p. 826 in the text).