Lecture 2-14

Last time Julie defined the dot product of the vectors $\vec{v} = (v_1, \ldots, v_n)$ and $\vec{w} = (w_1, \ldots, s_n)$ in $\mathbb{R}^n$ as the number $\vec{v} \cdot \vec{w} = \sum_{i=1}^n v_i w_i$; she also proved the Cauchy-Schwarz inequality, which asserts that $|\vec{v} \cdot \vec{w}| \leq ||\vec{v}||||\vec{w}||$. As a consequence we see that $\arccos \frac{\vec{v}\cdot\vec{w})}{||\vec{v}||||\vec{w}||}$ is always well defined; we take this angle (lying between 0 and π) to be the angle between $\vec{v}$ and $\vec{w}$, provided that both $\vec{v}$ and $\vec{w}$ are nonzero. In particular, the vectors $\vec{v}$ and $\vec{w}$ are orthogonal if and only if $\vec{v} \cdot \vec{w} = 0$.

We can now generalize familiar constructions from high-school geometry, which deals only with the plane, to higher dimensions, since given any two vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ there is always a plane in $\mathbb{R}^n$ containing both vectors. We define the *projection* $\vec{p}$ of $\vec{v}$ onto $\vec{w}$ to be the multiple $\frac{\vec{v}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\vec{w}$ of $\vec{w}$, if $\vec{w} \neq \vec{0}$; then you can easily check that $\vec{v} - \vec{p}$ is orthogonal to $\vec{w}$, so that projection in this context behaves like orthogonal projection in the plane. By the Pythagorean Theorem, the squared length of $\vec{v}$ is the sum of the squared lengths of $\vec{p}$ and $\vec{v} - \vec{p}$.

The Cauchy-Schwarz inequality implies in particular that for any nonzero vector $\vec{v}$, the unique unit vector $\vec{u}$ (having length 1) maximizing the dot product $\vec{v} \cdot \vec{u}$ is the unique unit vector making an angle of 0 with $\vec{v}$, namely $\vec{v}/||\vec{v}||$. The maximum value of $\vec{v} \cdot \vec{u}$ is thus $||\vec{v}||$. Similarly, the unique unit vector $\vec{u}$ minimizing $\vec{v} \cdot \vec{u}$ is the negative $-\vec{v}/||\vec{v}||$ of the first vector, for which $\vec{v} \cdot \vec{u} = -||\vec{v}||$. In this rather backhanded way we have managed to do a bit of calculus without even realizing it. This calculation will come into play later when we discuss directional derivatives of real-valued functions. Note also that by the Cauchy-Schwarz inequality we also have $||\vec{v}+\vec{w}||^2 = ||\vec{v}||^2 + 2\vec{v}\cdot\vec{w} + ||\vec{w}||^2 \leq ||\vec{v}||^2 + 2||\vec{v}||||\vec{w}|| + ||\vec{w}||^2 = (||\vec{v}|| + ||\vec{w}||)^2$, so that we get the triangle inequality $||\vec{v}+\vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$ mentioned earlier.

We indicated on Wednesday that there is also a *vector-valued* product of vectors $\vec{v}, \vec{w}$ in $\mathbb{R}^3$ (not defined on pairs of vectors in $\mathbb{R}^n$ unless $n = 3$) called the *cross product* and denoted $\vec{v} \times \vec{w}$. Given what was said about quaternions on Wednesday, the best way to give the definition of cross product is to note first that it is bilinear (i.e. satisfies the distributive law on both sides, just like the dot product), so that it is enough to say what the cross products of the unit coordinate vectors $\vec{i} = (1, 0, 0), \vec{j} = (0, 1, 0)$, and $\vec{k} = (0, 0, 1)$ are. The rule is that the cross product of any two of these vectors is the same (notationally) as their quaternion product if they are different, while the cross product of any of them (in fact of any vector in $\mathbb{R}^3$) with itself is $\vec{0}$. Thus we have $\vec{i} \times \vec{i} = \vec{j} \times \vec{j} = \vec{k} \times \vec{k} = \vec{0}$ and $\vec{i} \times \vec{j} = -\vec{j} \times \vec{i} = \vec{k}, \vec{j} \times \vec{k} = -\vec{k} \times \vec{j} = \vec{i}, \vec{k} \times \vec{i} = -\vec{i} \times \vec{k} = \vec{j}$. The general formula reads $(a_1, a_2, a_3) \times (b_1, b_2, b_3) = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$. The best way to remember this formula is to set it up as a 3×3 determinant. Recall first that the determinant of a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $ad - bc$. Given the vectors $(a_1, a_2, a_3), (b_1, b_2, b_3)$, form the 3×3 matrix $\begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$ whose first row consists of the unit coordinate vectors while the other two rows are filled out by the coordinates of the vectors being crossed. Compute the determinant as you would any 3×3 determinant; that is, multiply the first entry $\vec{i}$ of the first row by the determinant of the matrix obtained by removing its row and column, then multiply the second entry $\vec{j}$ of this row by the *negative* of the determinant of the matrix obtained by removing its row and column, then multiply the third entry $\vec{k}$ by the determinant of the matrix obtained by removing its row and column, and finally add the resulting vectors. You get exactly the formula given above.

There is a standard geometric interpretation of the cross product, paralleling another one for the dot product. If $\vec{v}, \vec{w}$ are two nonzero vectors in $\mathbb{R}^n$ then we could define their dot product $\vec{v} \cdot \vec{w}$ to be $||\vec{v}||||\vec{w}|| \cos \theta$, the product of their norms and the cosine of the angle θ between them; this is how we defined θ above. If $\vec{v}, \vec{w}$ happen to lie in $\mathbb{R}^3$ then it turns out that $\vec{v} \times \vec{w}$ has length $||\vec{v}||||\vec{w}|| \sin \theta$, while its direction is given by the *right-hand rule*: the direction of $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$, and if the fingers of the right hand curl from $\vec{v}$ to $\vec{w}$ then the right thumb points in the direction of $\vec{v} \times \vec{w}$. I will refer you to the text for the full justification of this interpretation, noting here only that the *triple scalar product* $\vec{u} \cdot (\vec{v} \times \vec{w})$ equals the determinant of the 3×3 matrix whose rows are the coordinates of $\vec{u}, \vec{v}$, and $\vec{w}$, in that order, as you can see immediately from the definition. From this it follows that if two of $\vec{u}, \vec{v}$, and $\vec{w}$ are equal then the triple scalar product $\vec{u} \cdot (\vec{v} \times \vec{w}) = 0$, since it is the determinant of a matrix with two equal rows. Thus the dot product of $\vec{v} \times \vec{w}$ with either $\vec{v}$ or $\vec{w}$ is 0, so that $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$, as claimed. Since $||\vec{v} \times \vec{w}||^2 = ||\vec{v}||^2 ||\vec{w}||^2 \sin^2 \theta$ and $(\vec{v} \cdot \vec{w})^2 = ||\vec{v}||^2 ||\vec{w}||^2 \cos^2 \theta$, we get $||\vec{v} \times \vec{w}||^2 + (\vec{v} \cdot \vec{w})^2 = ||\vec{v}||^2 ||\vec{w}||^2$; this identity is due to the Franco-Italian mathematician Lagrange.