1 Recap

- Positive-definiteness.
- Cauchy-Schwarz inequality, norms.

2 Orthogonality

In $\mathbb{R}^2$, the basis $\hat{i}, \hat{j}$ is special because $\hat{i}.\hat{i} = \hat{j}.\hat{j} = 1$ and $\hat{i}.\hat{j} = 0$. Thus $v.w = v_1 w_1 + v_2 w_2$. Motivated by this observation, we defined the notion of *Orthogonality*: In an inner product space V, $v, w \in V$ are said to be orthogonal to each other if $\langle v, w \rangle = 0$. A subset $S \subseteq V$ is said to be orthogonal if *any* pair of distinct elements are orthogonal to each other. A subset $S \subseteq V$ is said to be *orthonormal* if it is orthogonal and each element has unit norm.

An important result is: In an inner product space $(V, \langle , \rangle)$, an orthogonal set of nonzero elements is linearly independent. In particular, if V is f.d with $dim(V) = n$, any orthogonal set of nonzero elements of size n forms a basis.

Proof: Suppose $\sum_k c_k v_k = 0$ where $v_k \in S$. Then $\langle \sum_k c_k v_k, v_l \rangle = 0 = \sum_k c_k \langle v_k, v_l \rangle$. Thus, $c_l \|v_l\|^2 = 0$ and hence $c_l = 0$ for all l. $\square$

Examples and non-examples:

- The 0 vector is orthogonal to every vector. (As we shall see, it is the only vector with such properties. If we do not require positivity for all vectors in the inner product, then this property is false. Such "non-positive inner products" are useful in Relativity.)
- The standard basis vectors in $\mathbb{R}^n, \mathbb{C}^n$ with the usual inner products are orthonormal bases.
- The elements e^{ikx} in the space of continuous complex-valued functions on $[0, 2\pi]$ are orthogonal under the integration inner product. Alternatively, $u_0 = 1, u_{2n-1} = \cos(nx), u_{2n} = \sin(nx)$ are orthogonal but not orthonormal. Instead, $\frac{u_0}{\sqrt{2\pi}}, \frac{u_n}{\sqrt{\pi}}$ are orthonormal.
- The set $\{1, x, x^2\}$ is not orthogonal under the integration inner product.

Let V be a f.d. inner product space of dim n (over $\mathbb{R}$ or $\mathbb{C}$ as usual). Suppose $e_1, \ldots, e_n$ is an orthogonal basis. Then the components of a vector $x = \sum_k c_k e_k$ are: $c_j = \frac{\langle x, e_j \rangle}{\langle e_j, e_j \rangle}$. In particular, if e_j are orthonormal, then $c_j = \langle x, e_j \rangle$.

Proof: $\langle x, e_j \rangle = \sum_j c_k \langle e_k, e_j \rangle = c_j \langle e_j, e_j \rangle$. $\square$

In other words, on f.d. space with an orthonormal basis, $x = \sum_k \langle x, e_k \rangle e_k$. Let V be a f.d. inner product space and $e_1, \ldots, e_n$ is an orthonormal basis. Then $\langle x, y \rangle = \sum_k \langle x, e_k \rangle \overline{\langle y, e_k \rangle}$. In particular, $\|x\|^2 = \sum_{i=1}^n |\langle x, e_i \rangle|^2$. The proof is straightforward.

It turns out that in a certain function space (larger than continuous functions), e^{ikx} form an orthonormal "basis" of sorts. The analogue of the theorem above was discovered by Fourier and Parseval. It forms the basis for Fourier's technique of solving certain differential equations.

3 Gram-Schmidt algorithm/procedure/process

Suppose in $\mathbb{R}^2$, we were given $(1,1)$ and $(1,2)$ as a basis. If we were asked to construct an orthonormal basis, what would we have done ? Of course, we would have replaced $(1,2)$ with the piece that is orthogonal to $(1,1)$. This idea leads to the Gram-Schmidt procedure:
Let $x_1, \ldots$ be a finite or infinite sequence of vectors in $(V, \langle, \rangle)$. Let $L(x_1, \ldots, x_k)$ be the span of the first k elements. Then there is another collection $y_1, \ldots,$ in V such that

- y_k is orthogonal to every element in $L(y_1, \ldots, y_{k-1})$.
- $L(y_1, \ldots, y_k) = L(x_1, \ldots, x_k)$.
- The sequence $y_1, \ldots$ satisfying the above properties is unique upto scaling factors.

An example:
On the real vector space of say, continuous real-valued functions on $[-1,1]$, define the inner product $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$. Consider the linearly independent set $\{x_t = t^n\}$. As we saw earlier, this set is not orthogonal. Let's apply the GS procedure to this set to get an orthogonal set $y_0, y_1 \ldots$. The resulting polynomials (upto scaling factors) were obtained by earlier by Legendre in the context of differential equations. The (scaled versions) of these polynomials are called Legendre polynomials.
The polynomials $\phi_n = \frac{y_n}{\|y_n\|}$ are orthonormal and called the normalised Legendre polynomials. Here are a few : $y_0 = x_0 = 1$. $y_1 = x_1 - \frac{\int_{-1}^{1} x_1 x_0}{\int_{-1}^{1} x_0^2} = t$. $y_2 = x_2 - \frac{\int_{-1}^{1} x_2 y_1}{\int_{-1}^{1} y_1^2} - \frac{\int_{-1}^{1} x_2 y_0}{\int_{-1}^{1} y_0^2}$ which equals $t^2 - 0 - \frac{1}{3}$.
More generally, it turns out that $y_n = \frac{n!}{(2n)!}\frac{d^n(t^2-1)^n}{dt^n}$. The Legendre polynomials are $P_n(t) = \frac{(2n)!}{2^n(n!)^2} y_n(t)$.