Orthogonality *(Sec. 13)*

DEFINITION: $\vec{v}$ is *orthogonal* to a set $\mathcal{S}$ if $\vec{v}$ is orthogonal to every vector $\vec{u}$ in $\mathcal{S}$ $\quad(\vec{u}\cdot\vec{v} = 0)$.

DEFINITION: A set $\mathcal{S}$ (in particular, a basis) is *orthogonal* if $\vec{u}\cdot\vec{v} = 0$ for all $\vec{u}$, $\vec{v} \in \mathcal{S}$ with $\vec{u} \neq \vec{v}$.

DEFINITION: $\mathcal{S}$ is *orthonormal* if it is orthogonal and also $\|\vec{u}\| = 1$, $\forall \vec{u} \in \mathcal{S}$. (I.e., $\vec{u}_j \cdot \vec{u}_k = \delta_{jk}$.)

NOTE: An orthogonal (OG) set not containing $\vec{0}$ can be made orthonormal (ON) by replacing each $\vec{u}$ in it by $\hat{u} \equiv \frac{\vec{u}}{\|\vec{u}\|}$.

THEOREM 13.1′. *An orthogonal set not containing $\vec{0}$ is linearly independent.*

PROOF: $\sum \lambda^j \vec{u}_j = \vec{0} \Rightarrow 0 = \left(\sum \lambda^j \vec{u}_j\right) \cdot \vec{u}_k = \lambda^k \|\vec{u}_k\|^2$.

CONVENTION: From now on, when I speak of an orthogonal set, it will be tacitly understood that the set does not contain the zero vector.

TEMPORARY DEFINITION: An orthogonal set $\mathcal{S}$ (not containing $\vec{0}$) is *maximal* (also called *complete*) in $\mathcal{V}$ if one of these equivalent conditions is satisfied:

(A) $\mathcal{S}$ is not a proper subset of any larger orthogonal set [in $\mathcal{V}$].

(B) The only vector [in $\mathcal{V}$] orthogonal to $\mathcal{S}$ is $\vec{0}$.

THEOREM 13.2′. *If $\mathcal{V}$ is finite-dimensional, an orthogonal set (not containing $\vec{0}$) is a basis [for $\mathcal{V}$] iff it is maximal.*

PROOF:

(a) maximal $\Rightarrow$ basis: Will be a corollary of the Gram-Schmidt theorem, proved (noncircularly) below.

(b) basis $\Rightarrow$ maximal: Not maximal $\Rightarrow$ can create larger OG set $\Rightarrow$ larger independent set $\Rightarrow$ original set wasn't a basis.

Remark: In an infinite-dimensional Hilbert space, a maximal OG set is not a Hamel basis, but *is* a basis with respect to convergent infinite sums, convergence being defined with respect to the norm:

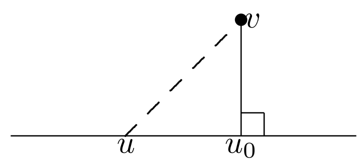
$$\sum_{j=1}^{\infty} \lambda^j \vec{u}_j = \vec{w} \quad \text{means} \quad \lim_{N\to\infty} \left\| \sum_{j=1}^{N} \lambda^j \vec{u}_j - \vec{w} \right\| = 0.$$

In this ∞-dim. context, the proof of Thm. 13.2′(a) needs to be replaced by a totally different argument:

Projection Theorem (Beppo-Levi's Theorem). *Let $\mathcal{U}$ be a (topologically)* **closed** *subspace of a Hilbert space $\mathcal{H}$. Let $\vec{v} \in \mathcal{H}$ but $\vec{v} \notin \mathcal{U}$. Then there is a unique $\vec{u}_0 \in \mathcal{U}$ which minimizes the distance from $\vec{v}$ to $\mathcal{U}$:*

$$0 < \inf_{u\in\mathcal{U}} \|\vec{u} - \vec{v}\| = \|\vec{u}_0 - \vec{v}\| \quad \bigl(= \min_{\vec{u}\in\mathcal{U}} \|\vec{u} - \vec{v}\|\bigr).$$

The vector $\vec{u}_0 - \vec{v}$ is orthogonal to $\mathcal{U}$.



Proof: See Milne, pp. 185–187, or Math. 642.

Corollary. *If $\mathcal{U}$ is a proper, closed subspace of a Hilbert space $\mathcal{H}$, then there exists in $\mathcal{H}$ a nonzero vector orthogonal to $\mathcal{U}$.*

The set of finite linear combinations of the vectors in an OG set is not a closed subspace; but that set together with all its limit points (the infinite linear combinations) is closed. The analogue of Thm. 13.2′ follows: an OG set is maximal iff it is a basis in the extended sense, allowing infinite sums.

Example **(Fourier cosine series):** $\mathcal{H} = \mathcal{L}^2(0,\pi)$. An OG basis (in the extended sense) is $\{1, \cos nx \ (n \in \mathbf{Z}_+)\}$:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx,$$

where

$$a_0 = \frac{1}{\pi}\int_0^\pi f(x)\,dx, \qquad a_n = \frac{2}{\pi}\int_0^\pi \cos nx \ f(x)\,dx;$$

$$a_0 = \frac{1\cdot f}{\|1\|^2}, \qquad \|1\|^2 = \int_0^\pi 1\,dx = \pi,$$

$$a_n = \frac{(\cos nx)\cdot f}{\|\cos nx\|^2}, \qquad \|\cos nx\|^2 = \int_0^\pi \cos^2 nx\,dx = \frac{\pi}{2}\,.$$

The corresponding ON basis is $\{\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}}\,\cos nx\}$. (Incidentally, this example shows how an OG basis may be more convenient than the related ON basis.)

Theorem. *Let $\mathcal{V}$ be a vector space of finite dimension M. If $\{\hat{e}_j\}_{j=1}^{M}$ is an ON basis, then every $\vec{v} \in \mathcal{V}$ can be expressed as*

$$\vec{v} = \sum_{j=1}^{M} \lambda^j \hat{e}_j\,, \qquad \lambda^j = \vec{v} \cdot \hat{e}_j.$$

If $\{\vec{e}_j\}$ is an OG basis, then

$$\vec{v} = \sum_{j=1}^{M} \lambda^j \vec{e}_j\,, \qquad \lambda^j = \frac{\vec{v} \cdot \vec{e}_j}{\|\vec{e}_j\|^2}.$$

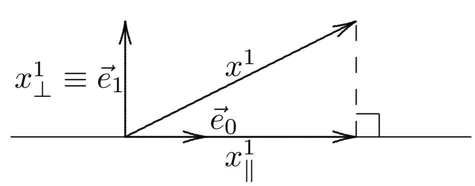
Proof: Implicit in proof of Thm. 13.1′. [Why is the denominator *squared* in the final equation?]

Gram-Schmidt process

Given a countable $\begin{Bmatrix} \text{basis} \\ \text{set} \end{Bmatrix}$ of vectors, one can construct an OG (in fact, ON) $\begin{Bmatrix} \text{basis} \\ \text{set with the same span} \end{Bmatrix}$.

Example and geometrical interpretation: Consider $L^2(-1,1)$. Let $\mathcal{V}$ = subspace of polynomials, restricted to $[-1,1]$ and equipped with the same inner product, $\int_{-1}^{1} f(x)\,\overline{g(x)}\,dx$. The functions may be either real- or complex-valued.

The obvious basis, $\{x^n\}_{n=0}^{\infty}$, is not OG. Let $\vec{e}_0 = x^0 = 1$. (To normalize: $\hat{e}_0 = \frac{\vec{e}_0}{\|\vec{e}_0\|} = \frac{1}{\sqrt{2}}$.) Now conceivably $(x^1) \cdot \vec{e}_0 \neq 0$. Break x^1 into components parallel and perpendicular to $\vec{e}_0$:



Claim: $x^1_\parallel = [(x^1) \cdot \vec{e}_0] \dfrac{\vec{e}_0}{\|\vec{e}_0\|^2}$; hence $x^1_\perp = x^1 - x^1_\parallel = \cdots$.

Verify: $\left\{ x^1 - [(x^1) \cdot \vec{e}_0] \dfrac{\vec{e}_0}{\|\vec{e}_0\|^2} \right\} \cdot \vec{e}_0 = (x^1) \cdot \vec{e}_0 - (x^1) \cdot e_0 = 0.$

In example: $(x^1)\cdot\vec{e}_0 = \int_{-1}^{1} x\,dx = \frac{x^2}{2}\Big|_{-1}^{1} = 0$ after all. Therefore, $x^1_\parallel = \vec{0}$, $\vec{e}_1 \equiv x^1_\perp = x^1 = x$.

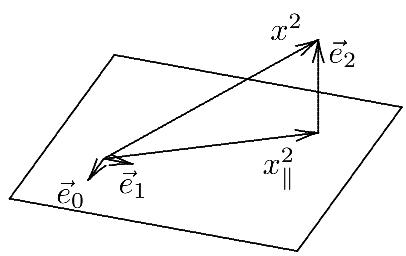
$$\|\vec{e}_1\|^2 = \int_{-1}^{1} x^2\,dx = \frac{x^3}{3}\bigg|_{-1}^{1} = \frac{2}{3} \;\Rightarrow\; \hat{e}_1 = \sqrt{\frac{3}{2}}\,x.$$

Next step: $\vec{e}_0$ and $\vec{e}_1$ span a plane.

$$x^2_\parallel = \text{projection of } x^2 \text{ onto that plane} = [(x^2)\cdot\vec{e}_0]\frac{\vec{e}_0}{\|\vec{e}_0\|^2} + [(x^2)\cdot\vec{e}_1]\frac{\vec{e}_1}{\|\vec{e}_1\|^2}\,.$$

$$\begin{aligned}\vec{e}_2 &\equiv x^2_\perp = x^2 - x^2_\parallel \\ &= x^2 - \frac{\langle x^2,1\rangle}{\langle 1,1\rangle}1 - \frac{\langle x^2,x\rangle}{\langle x,x\rangle}x \quad \text{in the more common bracket notation} \\ &= x^2 - \frac{1}{3}\,.\end{aligned}$$

Therefore $\hat{e}_2 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$.



General case: Projection of $\vec{v} \equiv \vec{v}_n$ onto $\mathcal{U} \equiv \text{span}\,\{\vec{e}_0,\ldots,\vec{e}_{n-1}\}$ is

$$\vec{v}_\parallel = \sum_{j=0}^{n-1} \frac{(\vec{v}\cdot\vec{e}_j)}{\|\vec{e}_j\|^2}\vec{e}_j = \sum_{j=0}^{n-1}(\vec{v}\cdot\hat{e}_j)\hat{e}_j.$$

Note that if $\vec{v}$ is *in* the span $\mathcal{U}$, then $\vec{v}_\parallel = \vec{v}$ by the (unnumbered) theorem above. (Otherwise, by the projection theorem, $\vec{v}_\parallel$ is the *best approximation* to $\vec{v}$ by a vector in $\mathcal{U}$, and $\vec{v}_\perp \equiv \vec{v} - \vec{v}_\parallel$ is the error left over (i.e., $\|\vec{v}_\perp\|$ is the shortest distance from $\vec{v}$ to $\mathcal{U}$).) If $\vec{v} \in \mathcal{U}$, it can be dropped from the list without changing the span. Otherwise, take $\vec{e}_j \equiv \vec{v}_\perp \neq \vec{0}$. Continuing inductively, we build up an OG set without changing the span.

The polynomials

$$\vec{e}_0 = 1, \qquad \vec{e}_1 = x, \qquad \vec{e}_2 = x^2 - \frac{1}{3}\,, \quad \ldots$$

are called *Legendre polynomials*. More precisely, $P_n(x) = N_n \vec{e}_n$ is the Legendre polynomial, when the normalization constant N_n is chosen by requiring $P_n(1) = 1$. The functions $P_n(\cos\theta)$ arise naturally in solving PDEs in spherical coordinates by separation of variables. Notice that the change of variables from θ to $x \equiv \cos\theta$ compensates for the factor $\sin\theta$ in the normalization integral arising from the Cartesian-to-polar Jacobian, $J = r^2 \sin\theta$:

$$\int_{-1}^{1} dx \quad \to \quad \int_{0}^{\pi} \sin\theta \, d\theta$$

GENERALIZATION: For any interval $[a, b] \subseteq \mathbf{R}$ and any positive "weight function" $w(x)$ we can define an inner product by

$$\|f\|_w^2 \equiv \int_a^b |f(x)|^2 \, w(x) \, dx.$$

If w is such that $\|x^n\|_w < \infty$ for all $n = 0, 1, 2, \ldots$, this is an inner product on the space of polynomials. We can use Gram-Schmidt to construct the corresponding *orthogonal polynomials*. Special cases with applications:

$$\int_{-\infty}^{\infty} |f(x)|^2 \, e^{-x^2} \, dx \ \Rightarrow \ \textit{Hermite polynomials} \text{ (harmonic oscillator)}$$

$$\int_{0}^{\infty} |f(x)|^2 \, e^{-x} x^{\alpha} \, dx \ \Rightarrow \ \textit{Laguerre polynomials} \text{ (hydrogen atom)}$$

$$\int_{-1}^{1} |f(x)|^2 \, \frac{dx}{\sqrt{1 - x^2}} \ \Rightarrow \ \textit{Chebyshev polynomials} \text{ (used in approximation theory)}$$

FINISH PROOF OF THM. 13.2′(a) (maximal OG set $\Rightarrow$ basis): OG $\Rightarrow$ independent $\Rightarrow$ can be extended to a basis, at least. Use Gram-Schmidt to orthogonalize the new vectors, if any (leaving the old ones alone). Thus if extension was necessary, we get a larger OG set, contradicting maximality.

ORTHOGONAL COMPLEMENTS

Definition: Let $\mathcal{U}$ be any set $\subset \mathcal{V}$ (not necessarily a subspace). The *orthogonal complement* of $\mathcal{U}$ is the set of all vectors orthogonal to $\mathcal{U}$:

$$\mathcal{U}^{\perp} \equiv \{\, \vec{v} : \vec{v} \cdot \vec{u} = 0 \ \forall \vec{u} \in \mathcal{U} \}.$$

Theorem 13.4′.

(a) *$\mathcal{U}^\perp$ is a subspace.*

(b) *If $\mathcal{U}$ is a subspace and is finite-dimensional, then $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$. (I.e., orthogonal complements are direct complements.) More precisely, every $\vec{v} \in \mathcal{V}$ has the decomposition*

$$\vec{v} = \vec{v}_\parallel + \vec{v}_\perp, \qquad \vec{v}_\parallel \in \mathcal{U}, \quad \vec{v}_\perp \in \mathcal{U}^\perp,$$

*and $\|\vec{v}\|^2 = \|\vec{v}_\parallel\|^2 + \|\vec{v}_\perp\|^2$ (**generalized Pythagorean theorem**).*

PROOF: See book. The proof of (b) amounts to the construction of $\vec{v}_\parallel$ and $\vec{v}_\perp$ for all $\vec{v} \in \mathcal{V}$, as in Gram-Schmidt.

REMARK: If $\mathcal{U}$ is an infinite-dimensional *subspace*, then $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$ **if** $\mathcal{U}$ is also *closed* in the topological sense (contains its limit points in the sense of convergence defined by the norm). This and all my similar infinite-dimensional remarks assume that $\mathcal{V}$ itself is *complete* (a Hilbert space) — i.e., all Cauchy sequences converge. (A sequence $\{\vec{v}_n\}$ is a *Cauchy sequence* if $\forall \epsilon > 0\, \exists M$ such that $\forall m, n > M, \quad \|\vec{v}_n - \vec{v}_m\| < \epsilon$.)

Coordinate expressions for the inner product

Let $\{\hat{e}_j\}$ be an ON basis, $\vec{v} = \sum \lambda^j \hat{e}_j, \quad \vec{u} = \sum \mu^j \hat{e}_j$. Then

$$\vec{v} \cdot \vec{u} = \sum_j \sum_k \lambda^j \overline{\mu^k}\, \hat{e}_j \cdot \hat{e}_k = \sum_j \lambda^j \overline{\mu^j},$$

$$\|\vec{v}\|^2 = \vec{v} \cdot \vec{v} = \sum_j |\lambda^j|^2.$$

[Cf. dot product in $\mathbf{R}^n$.]

If $\{\vec{d}_j\}$ is *any* basis, define $g_{jk} \equiv \vec{d}_j \cdot \vec{d}_k$. Then $\vec{v} = \sum \lambda^j \vec{d}_j, \quad \vec{u} = \sum \mu^j \vec{d}_j$ implies

$$\vec{v} \cdot \vec{u} = \sum_j \sum_k \lambda^j \overline{\mu^k}\, g_{jk}\,, \qquad \|\vec{v}\|^2 = \sum_j \sum_k \lambda^j \overline{\lambda^k}\, g_{jk}\,.$$

The Legendre polynomials, normalized so that $P_n(1) = 1$, provide an example of a basis for which $g_{jk} = 0$ if $j \neq k$, but $g_{jj} \neq 1$.

Generalized to manifolds, $\{g_{jk}\}$ becomes the *metric tensor*, which plays a central role in differential geometry and general relativity.

[Discussion of Sec. 14 postponed.]