Vector Spaces & Bases

Recap => A vector space is a nonempty set V where $\underbrace{\alpha x_1 + \beta x_2}_{\text{any scalars}} \in V$ when $x_1, x_2 \in V$.

=> A basis for V is a set $\{x_1, \ldots, x_n\}$ that
a) is linearly independent
b) spans V (the whole space).

=> Any vector $x \in V$ has unique coords s.t.

coordinates of x

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

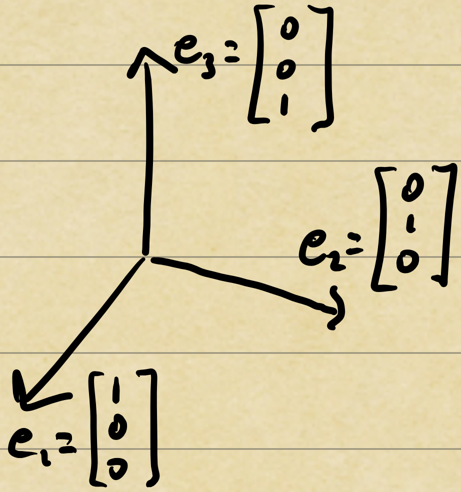
basis $\{x_1, \ldots, x_n\}$

=> The dimension of V is the # of vectors in any basis for V (always the same).

=> Two key structures on a vector space:
i) norm $\|\cdot\| : V \to [0,\infty)$ measures "size"
ii) inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ compares "size + direction"

Orthogonal Bases

In an inner product space, orthogonal bases are the gold standard. They generalize the usual Cartesian coordinate system



$$e_i^T e_j = 0 \qquad i \neq j$$
$$e_i^T e_i = 1$$

An orthogonal basis $S = \{x_1, \ldots, x_n\}$ satisfies

$$\langle x_i, x_j \rangle = 0 \quad \text{when } i \neq j$$

All basis vectors are orthogonal to each other

S is an orthonormal basis if, in addition,

$$\langle x_i, x_i \rangle = \|x_i\|^2 = 1$$

All basis vectors have unit length

Coordinates of $x \in V$ are easy to compute in an orthogonal basis $S = \{x_1, \ldots, x_n\}$

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

Calculate α_j:

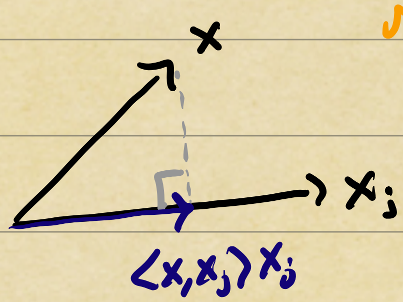
all inner products $= 0$ except jth

$$\langle x, x_j \rangle = \alpha_1 \overset{=0}{\langle x_1, x_j \rangle} + \cdots + \alpha_j \overset{\neq 0}{\langle x_j, x_j \rangle} + \cdots + \alpha_n \overset{=0}{\langle x_n, x_j \rangle}$$
$$= \alpha_j \langle x_j, x_j \rangle$$

$$\Rightarrow \quad \alpha_j = \frac{\langle x, x_j \rangle}{\langle x_j, x_j \rangle} = \frac{\langle x, x_j \rangle}{\|x_j\|^2}$$

Therefore, x has the unique representation

$$x = \frac{\langle x, x_1 \rangle}{\langle x_1, x_1 \rangle} x_1 + \frac{\langle x, x_2 \rangle}{\langle x_2, x_2 \rangle} x_2 + \cdots + \frac{\langle x, x_n \rangle}{\langle x_n, x_n \rangle} x_n$$



orthogonal projections onto jth coordinate

We'll return to the idea of orthogonal proj. when we discuss best approximations for functions.

Q: Suppose I start with a basis for $\mathbb{P}_n$

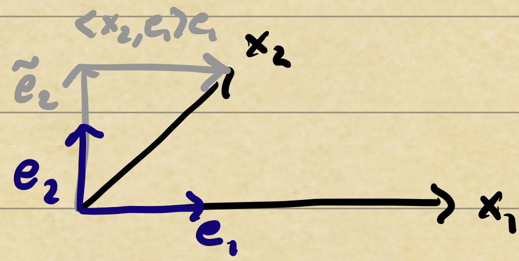
$$M = \{1, x, x^2, \ldots, x^n\}.$$

Is M orthonormal? How can we construct an orthonormal basis for $\mathbb{P}_n$?

Gram-Schmidt Orthogonalization

Given linearly independent vectors $\{x_1, \ldots, x_n\}$ in a vector space V with inner product $\langle \cdot, \cdot \rangle$, Gram-Schmidt systematically constructs an ONB.

$$e_1 = \frac{x_1}{\|x_1\|}$$



orthogonalize $\Rightarrow$ $\tilde{e}_2 = x_2 - \underbrace{\langle x_2, e_1 \rangle e_1}_{\text{orthogonal proj. onto } e_1}$

normalize $\Rightarrow$ $e_2 = \dfrac{\tilde{e}_2}{\|\tilde{e}_2\|}$

repeat for each vector $\Rightarrow$ $\tilde{e}_k = x_k - \overbrace{\langle x_k, e_{k-1} \rangle e_{k-1} - \cdots - \langle x_k, e_1 \rangle e_1}^{\text{orthogonal projection onto span}\{e_1, e_2, \ldots, e_{k-1}\}}$

$$e_k = \tilde{e}_k / \|\tilde{e}_k\|$$

The result is an orthonormal basis with

$$\langle e_i, e_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad i, j = 1, 2, \ldots, n$$

and note that e_1 is parallel to x_1.

<u>Example:</u> Let's build an orthonormal basis for $\mathbb{P}_n$, starting with the monomial basis

$$M = \{1, x, x^2, \ldots, x^n\}.$$

We'll use the inner product

$$\langle p, q \rangle = \int_{-1}^{+1} p(x) q(x) \, dx$$

Starting with $x_1 = 1$, we normalize:

$$\|x_1\| = \left(\int_{-1}^{+1} (1)^2 dx\right)^{1/2} = \sqrt{2} \Rightarrow e_1 = 1/\sqrt{2}$$

Now, we orthogonalize and normalize:

orthogonalize

$$\tilde{e}_2 = x_2 - \langle x_2, e_1 \rangle e_1 \quad \text{where } e_1 = 1/\sqrt{2}, \; x_2 = x$$

$$\langle x_2, e_1 \rangle = \int_{-1}^{+1} x \, (1/\sqrt{2}) \, dx = \frac{1}{\sqrt{2}} \left. \frac{x^2}{2} \right|_{-1}^{+1} = 0$$

normalize

$$\| \tilde{e}_2 \| = \left(\int_{-1}^{+1} x^2 \, dx \right)^{1/2} = \left(\left. \frac{x^3}{3} \right|_{-1}^{+1} \right)^{1/2} = \sqrt{\frac{2}{3}}$$

so that $e_2 = \sqrt{\frac{3}{2}}\, x$

So far we have $e_1 = 1/\sqrt{2}$ and $e_2 = \sqrt{3/2}\, x$.
For e_3, we orthogonalize by computing

$$\tilde{e}_3 = x^2 - \left[\int_{-1}^{+1} x^2 \left(\sqrt{\tfrac{3}{2}}\, x \right) dx \right] \left(\sqrt{\tfrac{3}{2}}\, x \right) - \left[\int_{-1}^{+1} x^2 \left(\tfrac{1}{\sqrt{2}} \right) dx \right] \left(\tfrac{1}{\sqrt{2}} \right)$$

$$x_3 \quad - \quad \langle x_3, e_2 \rangle \quad e_2 \quad - \quad \langle x_3, e_1 \rangle \quad e_1$$

$$\int_{-1}^{+1} x^2 \left(\sqrt{\tfrac{3}{2}}\, x \right) dx = \sqrt{\tfrac{3}{2}} \int_{-1}^{+1} x^3 \, dx = \sqrt{\tfrac{3}{2}} \left. \frac{x^4}{4} \right|_{-1}^{+1} = \sqrt{\tfrac{3}{2}}\,(0) = 0$$

$$\int_{-1}^{+1} x^2 (1/\sqrt{2}) \, dx = \frac{1}{\sqrt{2}} \int_{-1}^{+1} x^2 \, dx = \frac{1}{\sqrt{2}} \left. \frac{x^3}{3} \right|_{-1}^{+1} = \frac{1}{\sqrt{2}} \, \frac{2}{3} = \frac{2}{3\sqrt{2}}$$

$$\tilde{e}_3 = x^2 - 1/3$$

$$e_3 = \frac{3}{2} \sqrt{\frac{5}{2}} \left(x^2 - 1/3 \right) = \sqrt{\frac{5}{2}}\, \frac{1}{2} \left(3x^2 - 1 \right)$$

Repeating this process up to e_n, we obtain the Legendre polynomials. Here are the first 6:

$l_0(x) = c_0$

$l_1(x) = c_1 x$

$l_2(x) = \frac{c_2}{2}(3x^2-1)$

$l_3(x) = \frac{c_3}{8}(5x^3-3x)$

$l_4(x) = \frac{c_4}{8}(35x^4-30x^2+3)$

$l_5(x) = {}^{c_5}/_{16}(231x^6-315x^4+105x^2-5)$

$\vdots$

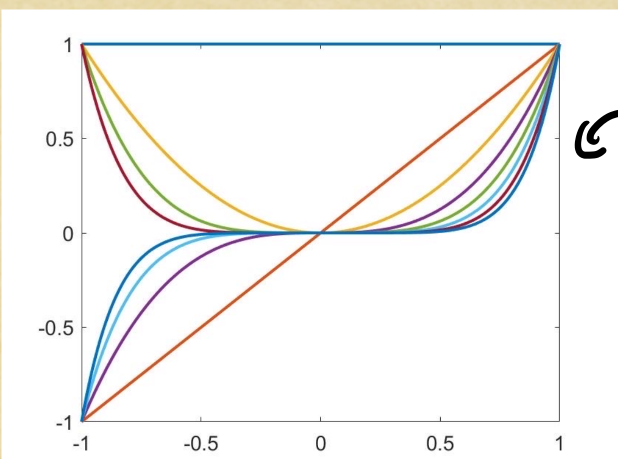
Normalization constants

$$c_k = \sqrt{\frac{2k+1}{2}}$$

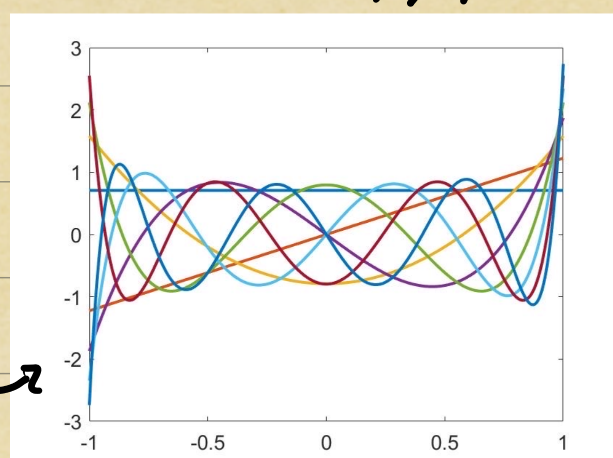
$$l_n(x) = c_n \, 2^n \sum_{k=0}^{n} \binom{n}{k}\binom{\frac{n+k-1}{2}}{n} x^k$$

(General formula)

Classically, Legendre polynomials are presented w/out the c_k's, i.e. $p_k(x) = \frac{l_k(x)}{c_k}$. These classical orthogonal polynomials are not orthonormal, but are scaled so that $p_k(1) = 1$ for $k = 0, 1, 2, \dots$



monomials



Legendre $(\|l_n\| = 1)$

Change-of-Basis

Converting back & forth between different coordinate systems is a key task. Given

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

To compute x from its coefficients, we write

$$x = \underbrace{\overset{X}{\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix}}}_{\text{If } x_1,\ldots,x_n \text{ are functions, this is a quasimatrix}} \overset{\underline{\alpha}}{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}$$

If $\{x_1, \ldots, x_n\}$ is an ONB in inner product $\langle , \rangle$:

$$\underline{\alpha} = X^T x = \underbrace{\begin{bmatrix} \langle x, x_1 \rangle \\ \langle x, x_2 \rangle \\ \vdots \\ \langle x, x_n \rangle \end{bmatrix}}_{\text{Same orthogonal projection as before!}}$$

This works for any space/ONB, even functions!

If $\{x_1, \dots, x_n\}$ is not ONB, then

$$x = \underline{X}\,\underline{\alpha} \implies \underline{X}^T x = \underbrace{\underline{X}^T \underline{X}}_{\langle x_i, x_j \rangle = (i,j)^{th} \text{ entry}} \underline{\alpha}$$

$$\underline{\alpha} = \underbrace{(\overbrace{\underline{X}^T \underline{X}}^{\text{"Gram matrix"}})^{-1} \underline{X}^T}_{\text{pseudo-inverse}} x$$

Note: the Gram matrix is invertible, whenever the columns of $\underline{X}$ are linearly indep.

Now suppose we have another basis

$$x = \beta_1 y_1 + \beta_2 y_2 + \cdots + \beta_n y_n$$

How are the coordinates of x in the two bases $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$ related?

$$x = \underbrace{\begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}}_{\underline{X}} \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}}_{\underline{\alpha}} = \underbrace{\begin{bmatrix} | & & | \\ y_1 & \cdots & y_n \\ | & & | \end{bmatrix}}_{\underline{Y}} \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}}_{\underline{\beta}}$$

$$X^TX\underline{\alpha} = X^TY\underline{\beta} \Rightarrow \underline{\alpha} = (X^TX)^{-1}\underbrace{X^TY}_{\substack{(i,j)^{th}\text{ entry}\\ = \langle y_i, x_j\rangle}}\underline{\beta}$$

Linear Transformations

Given vector spaces V, W, a map $T: V \to W$ is a linear transformation if

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for any vectors $x, y \in V$ and scalars α, β.

Example: $T: x \to Ax$ where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$

Example: Differentiation on $C^1([-1,1])$:

$$\frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha \frac{df}{dx} + \beta \frac{dg}{dx}$$

Example: Cumulative Integration on $C([-1,1])$:

$$\int_{-1}^{x}(\alpha f(y) + \beta g(y))dy = \alpha \int_{-1}^{x} f(y)dy + \beta \int_{-1}^{x} g(y)dy$$

If $T: V \to W$ is a linear transformation and $\{x_1, \dots, x_n\} \subset V$, $\{y_1, \dots, y_n\} \subset W$ are bases for V, W, we can <u>always</u> write down a matrix representation of T!