Gram-Schmidt orthogonality process

Recall: Dot Product (inner product)

- Think of the dot product as a matrix multiplication

$$\vec{x} \cdot \vec{y} =$$

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_N \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_N y_N$$

$$= \sum_{i=1}^{N} x_i y_i$$

1 X N N X 1 1 X 1

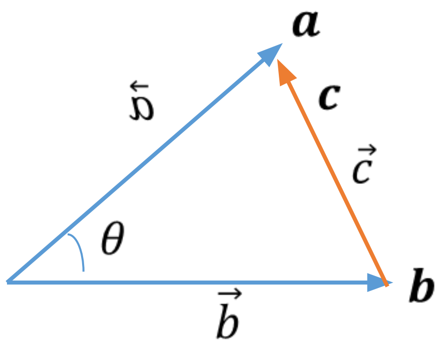
- MATLAB: 'inner matrix dimensions must agree'

Outer dimensions give size of resulting matrix

Recall: Dot Product (inner product)

$$\boldsymbol{a} \cdot \boldsymbol{b} \; = ||a|| \; ||b|| \cos\theta$$

The dot product is also related to the angle between the two vectors



Law of cosines

$$||\boldsymbol{c}||^2 = ||\boldsymbol{a}||^2 + ||\boldsymbol{b}||^2 - 2||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||\boldsymbol{cos\theta} \;\; (\mathbf{1})$$

$|| \cdot ||$ The magnitude/length of a vector. For example, let $\boldsymbol{a} = (a_1, a_2)$ be a two-dimensional vector, then the formula for its magnitude is $\sqrt{\mathrm{a}_1^2 + a_2^2}$

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$$
$$\boldsymbol{c} \cdot \boldsymbol{c} = (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b})$$
$$||\boldsymbol{c}||^2 = \boldsymbol{a} \cdot \boldsymbol{a} - \boldsymbol{a} \cdot \boldsymbol{b} - \boldsymbol{b} \cdot \boldsymbol{a} + \boldsymbol{b} \cdot \boldsymbol{b}$$
$$||\boldsymbol{c}||^2 = ||\boldsymbol{a}||^2 + ||\boldsymbol{b}||^2 - 2\boldsymbol{a} \cdot \boldsymbol{b} \;\; (\mathbf{2})$$

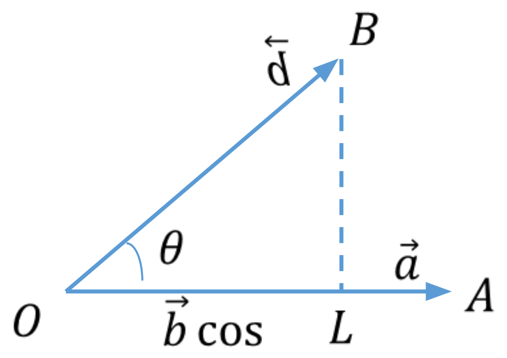
Recall: Dot Product Geometrical Interpretation

$$\boldsymbol{a} \cdot \boldsymbol{b} = ||a||\ ||b||\cos\theta$$

The dot product is also related to the angle between the two vectors

- In ΔOLB

$$\cos\theta = \frac{OL}{OB}$$



$$OL = OB\cos\theta$$
$$OL = ||\boldsymbol{b}||\cos\theta$$

$$A.B = ||\boldsymbol{a}||OL$$

Magnitude of vector A

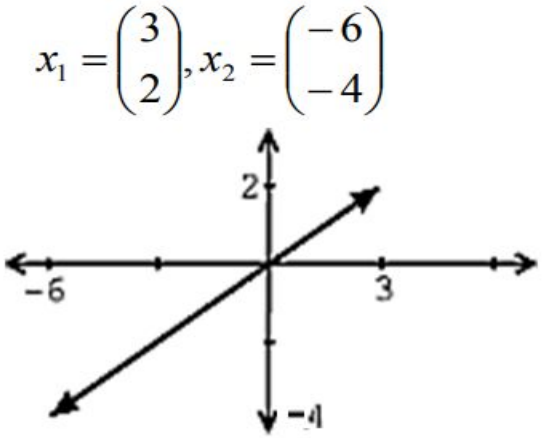
Projection of vector B on A

$$A.B = 0 \implies A \perp B$$

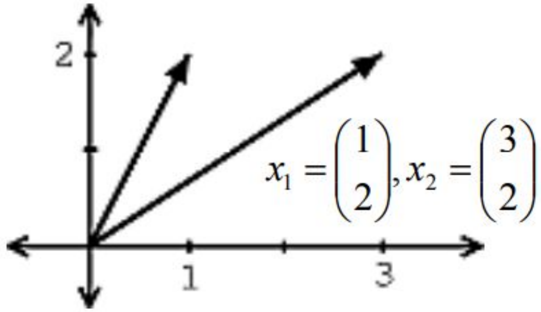
Recall: Linear Independence

- A set of d-dimensional vectors $x_i \in \mathbb{R}^d$, are said to be linearly independent if none of them can be written as a linear combination of the others.
- In other words,

$$c_1 x_1 + c_2 x_2 + \cdots + c_k x_k = 0,$$
$$iff\ c_1 = c_2 = c_3 \ldots = c_n = 0$$



Not linearly independent vectors



Linearly independent vectors

Recall: Span

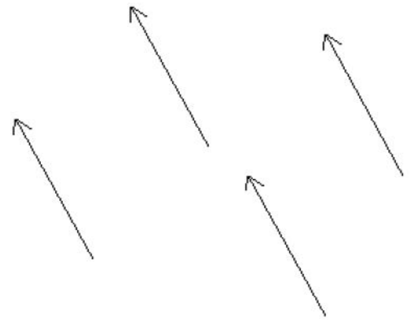
- **A span of a set of vectors $x_1, x_2, \dots, x_k$ is the set of vectors that can be written as a linear combination of $x_1, x_2, \dots, x_k$.**

$$span(x_1, x_2, \dots, x_k) = \{c_1 x_1 + c_2 x_2 + \cdots + c_k x_k \mid c_1, c_2, \dots, c_k \in \mathbb{R}\}$$

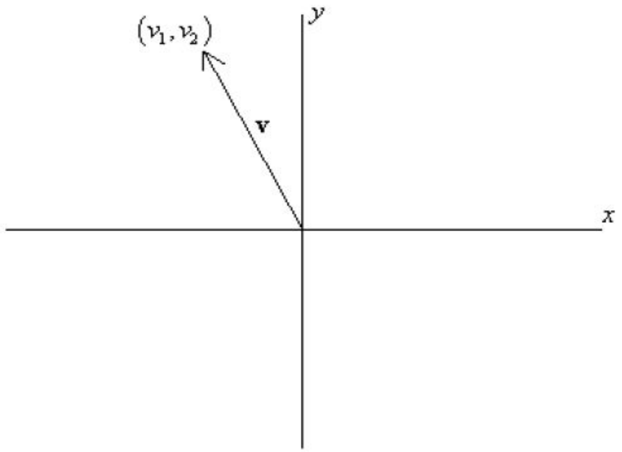
Recall: Bases & Orthonormal Bases

- A basis for $\mathbb{R}^d$ is a set of vectors which:
  - Spans $\mathbb{R}^d$, i.e. any vector in this d-dimensional space can be written as linear combination of these basis vectors.
  - Are linearly independent
- Clearly, any set of d-linearly independent vectors form basis vectors for $\mathbb{R}^d$
- Ortho-Normal: orthogonal + normal
  - Orthogonal: dot product is zero
  - Normal: magnitude is one

$$x = [1 \quad 0 \quad 0]^T \qquad x \cdot y = 0$$
$$y = [0 \quad 1 \quad 0]^T \qquad x \cdot z = 0$$
$$z = [0 \quad 0 \quad 1]^T \qquad y \cdot z = 0$$



VS



Inner product space

- In linear algebra, an inner product space is a vector space with an additional structure called an inner product.
- This additional structure associates each pair of vectors in the space with a scalar quantity known as the inner product of the vectors.
- Let u, v, and w be vectors in $\mathbb{R}^n$, and let c be a scalar. Then
  1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
  2. $(\mathbf{u} + \mathbf{v})\mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
  3. $(\mathrm{c}\,\mathbf{u}) \cdot \mathbf{v} = \mathrm{c}\,(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\mathrm{c}\mathbf{v})$
  4. $\mathbf{u} \cdot \mathbf{u} \geq 0$; and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathrm{u} = 0$.

The Length of Vector

-
- The length (or norm) of $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is the nonnegative scalar ||v|| defined by

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \ \ and \ \ \|v\|^2 = v \cdot v.$$

- $||cv|| = |c| \cdot ||v||, \quad \forall v \in \mathbb{R}^n, c \in \mathbb{R}$

Normalizing v

- A vector whose length is 1 is called a unit vector
- Normalizing v
  - Let, $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be a non-zero vector, then $u = \frac{v}{||v||}$ unit vector in the same direction as v

Distance between two vectors

- For u and v in $\mathbb{R}^n$, the distance between u and v, written as dist(u, v), is the length of the vector $u - v$. That is,

$$\text{dist}(u, v)=||u-v||$$

Orthogonal sets

- Let V be a vector space with an inner product.
- **Definition**. Nonzero vectors $v_1, v_2, \ldots, v_k \in V$ form an orthogonal set if they are orthogonal to each other: $< v_i , v_j > = 0$ for $i \neq j$.
- If, in addition, all vectors are of unit norm, $||v_i|| = 1$, then $v_1, v_2, \ldots, v_k$ is called an orthonormal set.
- Any orthogonal set is linearly independent. [proof: next slide]
- If $S = \{u_1, u_2, \ldots, u_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

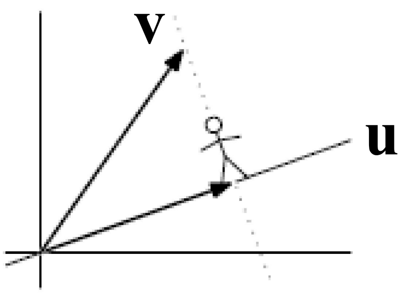
Orthogonal sets are linearly independent

- **If $S = \{ v_1, v_2, \dots v_n\}$is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.**
- Proof:
  - S is an orthogonal set of nonzero vectors, i.e., $\langle v_i, v_j \rangle = \mathbf{0}, \boldsymbol{i \neq j} \; and. \langle \boldsymbol{v_i}, \boldsymbol{v_i} \rangle > \mathbf{0}.$
  - To be independent $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = 0, \boldsymbol{iff} \; \boldsymbol{c_1} = \dots = \boldsymbol{c_n} = \mathbf{0}$
    $\Rightarrow \langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \mathbf{v}_i \rangle = \langle 0, \mathbf{v}_i \rangle = 0 \;\; \forall i$
    $\Rightarrow c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle + \cdots + c_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle$
    $\Rightarrow c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle \Rightarrow c_i = 0, \forall i$ **, and hence S is linearly independent**

Orthogonal projection

- Let u and v be two vectors in an inner product space V, such that $u \neq v$. Then the orthogonal projection of v onto u is given by

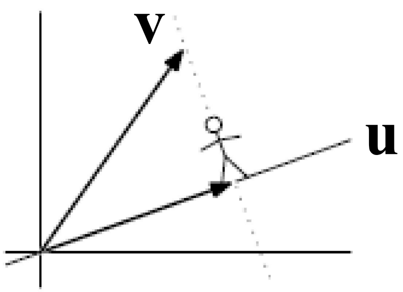
$$proj_u v = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$$



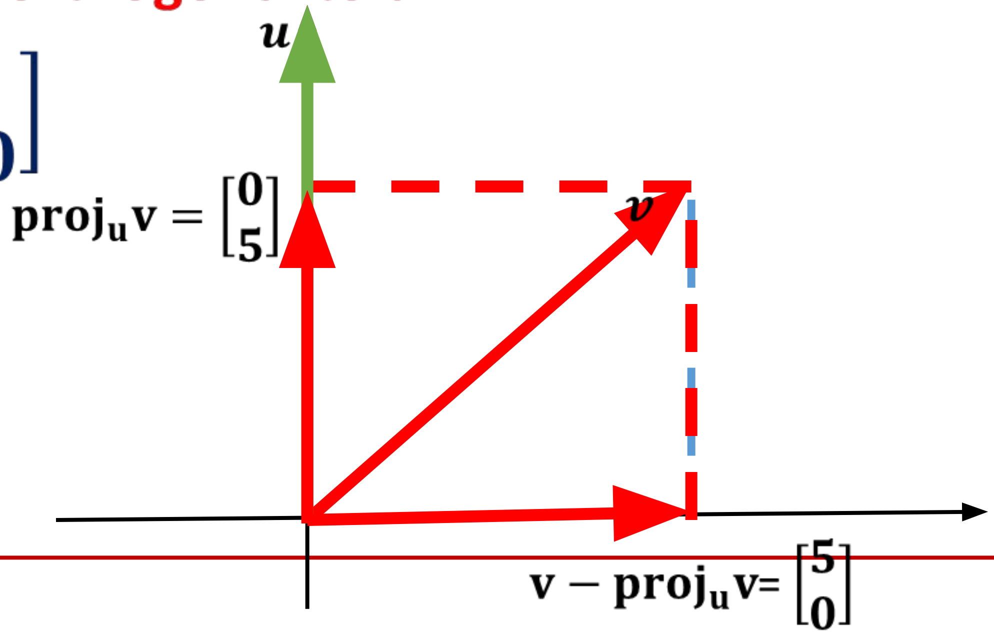
- Example: Use the Euclidean inner product in R^3 to find the orthogonal projection of v=(6, 2, 4) onto u=(1, 2, 0).
  - $< v, u > = (6)(1) + 2(2) + 4(0) = 10$
  - $< u, u > = 1^2 + 2^2 + 0^2 = 5$
  - $proj_u v = \frac{<v,u>}{<u,u>} u = \frac{10}{5}(1, 2, 0) = (2, 4, 0)$

Orthogonal projection

- Given two vectors u and v, we can decompose v into a sum of two vectors, one a multiple of u and the other orthogonal to u.
  1. $\mathbf{proj_u v} = \frac{\langle \mathbf{v},\mathbf{u}\rangle}{\langle \mathbf{u},\mathbf{u}\rangle}\mathbf{u}$ is the orthogonal projection of v onto u
  2. $\mathbf{v} - \mathbf{proj_u v}$ is the component of v orthogonal to u



- Example: Let $v = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$ and $u = \begin{bmatrix} 0 \\ 10 \end{bmatrix}$
- $\mathbf{proj_u v} = \frac{50}{100} u = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$
- $\mathbf{v} - \mathbf{proj_u v} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$



Gram-Schmidt Orthogonalization

Any basis		Orthogonal basis
$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$	$\longrightarrow$	$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

- **Let V be a vector space with an inner product. Suppose $x_1, x_2, \ldots, x_n$ is a basis for V . Let**

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{\langle x_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$$

$$v_3 = x_3 - \frac{\langle x_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2$$

$$\vdots$$

$$v_n = x_n - \frac{\langle x_n, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x_n, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \cdots - - \frac{\langle x_n, v_{n-1} \rangle}{\langle v_{n-1}, v_{n-1} \rangle} v_n$$

Then, $v_1, v_2, \ldots, v_n$ is an orthogonal basis for V

Properties of Gram-Schmidt Process

Any basis	$\longrightarrow$	Orthogonal basis
$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$		$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

- $v_k = x_k - (\alpha_1 x_1 + \cdots + \alpha_{k-1} x_{k-1}), 1 \leq k \leq n$
  - **The span of $v_1, v_2, \ldots, v_k$ is the same as the span of $x_1, \ldots, x_k$**
  - v_k **is orthogonal to $x_1, \ldots, x_{k-1}$**
  - $v_k = x_k - p_k$, **where p_k is the orthogonal projection of the vector x_k on the subspace spanned by $x_1, \ldots, x_{k-1}$**
  - $||v_k||$ **is the distance from x_k to the subspace spanned by $x_1, \ldots, x_{k-1}$**

Normalization

- Let V be a vector space with an inner product. Suppose $v_1, v_2, ..., v_n$ is an orthogonal basis for V
- Let $\boldsymbol{w_1} = \frac{\boldsymbol{v_1}}{||\boldsymbol{v_1}||}, \boldsymbol{w_2} = \frac{\boldsymbol{v_2}}{||\boldsymbol{v_2}||}, \ldots, \boldsymbol{w_n} = \frac{\boldsymbol{v_n}}{||\boldsymbol{v_n}||}$
- Then $w_1, w_2, ..., w_n$ is an orthonormal basis for V
- An alternative form of the Gram-Schmidt process combines orthogonalization with normalization

$$\boldsymbol{v_1} = \boldsymbol{x_1}, \boldsymbol{w_1} = \frac{\boldsymbol{v_1}}{||\boldsymbol{v_1}||}$$

$$\boldsymbol{v_2} = \boldsymbol{x_2} - \frac{\langle \boldsymbol{x_2}, \boldsymbol{w_1} \rangle}{\langle \boldsymbol{w_1}, \boldsymbol{w_1} \rangle} \boldsymbol{w_1}, \boldsymbol{w_2} = \frac{\boldsymbol{v_2}}{||\boldsymbol{v_2}||}$$

$$\vdots$$

Example: Gram-Schmidt Orthogonalization

- Apply the Gram-Schmidt process to the following basis.

$$B = \{\overset{\mathbf{u}_1}{(1,1,0)},\ \overset{\mathbf{u}_2}{(1,2,0)},\ \overset{\mathbf{u}_3}{(0,1,2)}\}$$

Sol:

$$\mathbf{v}_1 = \mathbf{u}_1 = (1,1,0)$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = (1,2,0) - \frac{3}{2}(1,1,0) = (-\frac{1}{2},\frac{1}{2},0)$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

$$= (0,1,2) - \frac{1}{2}(1,1,0) - \frac{1/2}{1/2}(-\frac{1}{2},\frac{1}{2},0) = (0,0,2)$$

Example: Gram-Schmidt Orthogonalization

$$\Rightarrow B' = \{\mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3\} = \{(1,1,0), (\frac{-1}{2}, \frac{1}{2}, 0), (0,0,2)\}$$

Orthonormal basis

$$\Rightarrow B'' = \{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}\} = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0,0,1)\}$$

The QR Factorization: Basic Idea

- If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthogonal basis for $Col\ A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its main diagonal

- Applications
  - linear equations
  - least squares problems
  - constrained least squares problems

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 3 \end{bmatrix}.$$

Using the Gram Schmidt process, we can find orthonormal basis for col A

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- $A = QR$
- $\boldsymbol{Q^T A = Q^T Q R = R}$

$$R = Q^T A = \begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 0 & 3 \end{bmatrix}.$$

Source: http://www2.math.uconn.edu/~troby/math2210f16/LT/sec6_4.pdf

Topics for the next class

- Linear Mappings
- Kernel and Image space of a linear map
- Matrix associated with linear map
- Eigenvectors & Eigenvalues

Reference

- Introduction to Linear Algebra, Mark Goldman, Emily Mackevicius
- Lecture Notes on Linear Algebra, Arbind K Lal, Sukant Pati
- Lecture notes on linear algebra, David Lerner
- 6502 : Mathematics for Engineers 2, James Burnett, Department of Mathematics, University College London
- ENGG2013, Advanced Engineering Math. Kenneth Shum
- Matrices and Determinants – 2, Carrer launcher
- http://www.shsu.edu/ldg005/data/mth199/chapter5.ppt
- Lecture 12, Inner Product Space & Linear Transformation