MA-110 Linear Algebra

Nazar Khan

PUCIT

12. Gram-Schmidt Process

Orthogonal & Orthonormal Sets

Content in this lecture applies to *sets* of 2 or more vectors.

- The solution of a problem can often be simplified by choosing a basis with orthogonal basis vectors.
- Further simplification is achieved if the orthogonal vectors are also unit vectors.

  A *set* of two or more vectors in a real inner product space is said to be *orthogonal* if all pairs of distinct vectors in the set are orthogonal.

  An orthogonal set in which each vector has norm 1 is said to be *orthonormal*.

- Roughly, orthonormal = orthogonal + normal.

Orthogonal & Orthonormal Basis

- Recall that any set of n linearly independent vectors constitutes a basis for $\mathbb{R}^n$.
- If a set of basis vectors is orthogonal as well, it is called an *orthogonal basis*.
- If an orthogonal basis is made from unit vectors, it is called an *orthonormal basis*.
- Any vector $\mathbf{v}$ can be normalized by dividing by its magnitude $\frac{1}{||\mathbf{v}||}\mathbf{v}$ which makes it a unit vector. Verify this.
- A familiar orthonormal basis is the standard basis for $\mathbb{R}^n$ with the Euclidean inner product:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \ldots \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Example

Verify that the following set is orthonormal with respect to the standard Euclidean inner product on $\mathbb{R}^3$.

$$\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{u}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \mathbf{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Why orthonormal basis?

- Assume $S = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is a *non-orthogonal* basis for $\mathbb{R}^n$.
- Any vector $\mathbf{u}$ can be represented in S as

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

which can be written as a linear system $V\mathbf{c} = \mathbf{u}$

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots \mathbf{v}_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{u}$$

- It can be solved for the coefficients $c_1, c_2, \ldots, c_n$ via matrix inversion $\mathbf{c} = V^{-1}\mathbf{u}$ which can be an expensive operation.

Why orthonormal basis?

- Had S been orthogonal (or orthonormal), finding the c_i would have been much easier.
  - *Orthogonal case:*

$$c_i = \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{||\mathbf{v}_i||^2}$$

  - *Orthonormal case:*

$$c_i = \langle \mathbf{u}, \mathbf{v}_i \rangle$$

- Notice the convenience – linear system versus simple inner products.

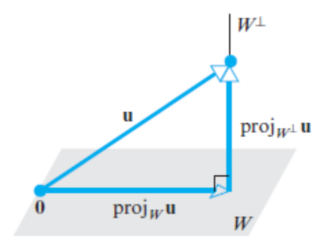
Why orthonormal basis?

Proof of orthogonal case

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{v}_i \rangle &= \langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \mathbf{v}_i \rangle \\
&= c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + c_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle \\
&= c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle \text{ since } S \text{ is an orthogonal basis} \\
\implies c_i &= \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{||\mathbf{v}_i||^2}
\end{aligned}$$

Orthonormal case can now be proven by observing that $||\mathbf{v}_i|| = 1$ if S is an orthonormal basis.

Orthogonal Projection



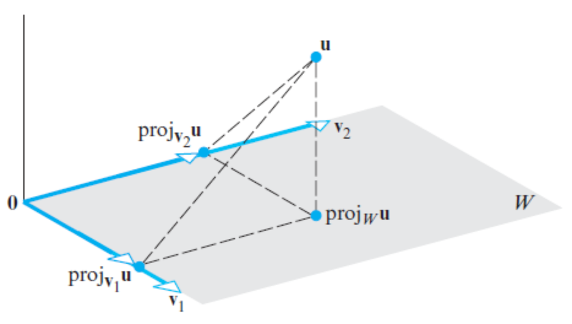
- $\mathbf{u}$ is a vector in inner product space V.
- W is a subspace of V.
- We can express

$$\mathbf{u} = \text{proj}_W \mathbf{u} + \text{proj}_{W^\perp} \mathbf{u}$$

or alternatively,

$$\text{proj}_{W^\perp} \mathbf{u} = \mathbf{u} - \text{proj}_W \mathbf{u}$$

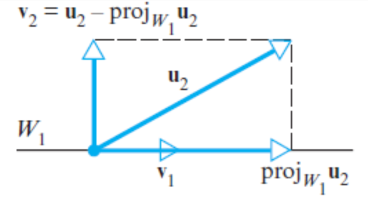
Orthogonal Projections

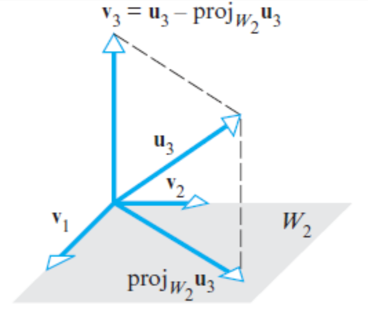


- Let W be an r-dimensional subspace of V and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ be an orthogonal basis for W.
- $\text{proj}_W \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2} \mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{||\mathbf{v}_2||^2} \mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_r \rangle}{||\mathbf{v}_r||^2} \mathbf{v}_r$.
- If basis is orthonormal, then $\text{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_r \rangle \mathbf{v}_r$.

Gram-Schmidt Process

- Every nonzero finite-dimensional inner product space has an orthonormal basis.
- A basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ can be converted into *an orthonormal basis* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ via the Gram-Schmidt process.
- Basic idea: $\mathbf{u} - \text{proj}_W\mathbf{u}$ is always orthogonal to W.





Gram-Schmidt Process

To convert a basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ into an *orthogonal* basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, perform the following computations:

1. $\mathbf{v}_1 = \mathbf{u}_1$
2. $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1$
3. $\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$
4. $\mathbf{v}_4 = \mathbf{u}_4 - \frac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 - \frac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2} \mathbf{v}_3$
5. Continue until $\mathbf{v}_r$.

Normalize the $\mathbf{v}_i$ vectors to obtain an *orthonormal* basis.

> Notice that *by construction* $\mathbf{v}_i$ is orthogonal to each of the vectors in $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}$.

Gram-Schmidt Process

Assume Euclidean inner product is defined on $\mathbb{R}^3$. Apply the Gram–Schmidt process to transform the basis vectors

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

into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

QR Decomposition

Every *invertible matrix* A can be written as the product $A = QR$ of an orthonormal matrix Q and an upper triangular matrix R.

- The column vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of A can be converted into an *orthonormal* set $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ using the Gram-Schmidt process.
- Each column vector $\mathbf{u}_i$ can be represented in the new basis as

$$\begin{aligned}\mathbf{u}_i &= \langle \mathbf{u}_i, \mathbf{q}_1\rangle \mathbf{q}_1 + \langle \mathbf{u}_i, \mathbf{q}_2\rangle \mathbf{q}_2 + \cdots + \langle \mathbf{u}_i, \mathbf{q}_n\rangle \mathbf{q}_n \\ &= \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n\end{bmatrix} \begin{bmatrix}\langle \mathbf{u}_i, \mathbf{q}_1\rangle \\ \langle \mathbf{u}_i, \mathbf{q}_2\rangle \\ \vdots \\ \langle \mathbf{u}_i, \mathbf{q}_n\rangle\end{bmatrix}\end{aligned}$$

QR Decomposition

- So the n columns of A can be represented as

$$\underbrace{\begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n\end{bmatrix}}_{A} = \underbrace{\begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n\end{bmatrix}}_{Q} \underbrace{\begin{bmatrix}\langle \mathbf{u}_1, \mathbf{q}_1\rangle & \langle \mathbf{u}_2, \mathbf{q}_1\rangle & \dots & \langle \mathbf{u}_n, \mathbf{q}_1\rangle \\ \langle \mathbf{u}_1, \mathbf{q}_2\rangle & \langle \mathbf{u}_2, \mathbf{q}_2\rangle & \dots & \langle \mathbf{u}_n, \mathbf{q}_2\rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{u}_1, \mathbf{q}_n\rangle & \langle \mathbf{u}_2, \mathbf{q}_n\rangle & \dots & \langle \mathbf{u}_n, \mathbf{q}_n\rangle\end{bmatrix}}_{R}$$

- The terms in the lower triangle (shown in red) are all 0 due to the Gram-Schmidt process. Hence R is an upper triangular matrix.
- Find a QR-decomposition of

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