CHAPTER 27

Lecture No. 27

27.1 The Inner Product

If **u** and **v** are vectors in $\mathbb{R}^n$, then we regard **u** and **u** as $n \times 1$ matrices.

The transpose $\mathbf{u}^T$ is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, which we write as a single real number (a scalar) without brackets.

The number $\mathbf{u}^T\mathbf{v}$ is called the inner product of **u** and **v**, and often it is written as **u.v**.

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix}$$

then the inner product of **u** and **v** is

$$\mathbf{u}.\mathbf{v} = u_1v_1 + u_2v_2 + ... + u_nv_n.$$

Example: Compute **u.v** and **v.u** for **u** and **v** for $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

Theorem: Let **u**, **v**, and **w** be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

1. $\mathbf{v}.\mathbf{u} = \mathbf{u}.\mathbf{v}$
2. $(\mathbf{v} + \mathbf{u}).\mathbf{w} = \mathbf{v}.\mathbf{w} + \mathbf{u}.\mathbf{w}$
3. $(c\mathbf{u}).\mathbf{v} = \mathbf{u}.(c\mathbf{v}) = c(\mathbf{v}.\mathbf{u})$
4. $\mathbf{u}.\mathbf{u} \geq 0$, and $\mathbf{u}.\mathbf{u} = 0$ if and only if $\mathbf{u} = 0$

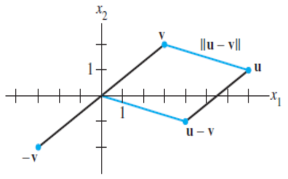
The Length of a Vector: The length (or norm) of **v** is the nonnegative scalar **v** defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}.\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + ... + v_n^2} = \mathbf{v}.\mathbf{v}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}.\mathbf{v}$$

Example: Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector **u** in the direction of **v**.

Solution: the unit vector is $\frac{\mathbf{v}}{\sqrt{\mathbf{v}.\mathbf{v}}}$.
Example: Let W be the subspace of $\mathbb{R}^2$ spanned by $\mathbf{x} = (2/3, 1)$. Find a unit vector $\mathbf{z}$ that is a basis for W.
Distance in $\mathbb{R}^n$: For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the distance between $\mathbf{u}$ and $\mathbf{v}$,written as dist($\mathbf{u}$,$\mathbf{v}$), is the length of the vector $\mathbf{u} - \mathbf{v}$. That is

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



Example: Compute the distance between the vectors $\mathbf{u} = (2, -1, 2, 1)$, $\mathbf{v} = (1, -2, 2, 0)$.
Example: Compute the distance between the vectors $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$.

27.2 Orthogonal Vectors

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (to each other) if $\mathbf{u}.\mathbf{v} = 0$.

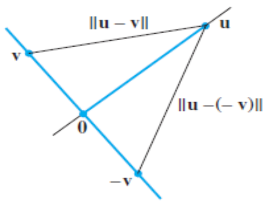


Figure 27.1: Orthogonal vectors

Theorem: Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

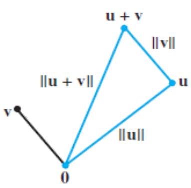
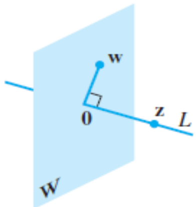


Figure 27.2: Orthogonal vectors

Orthogonal Complements: If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be orthogonal to W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the orthogonal complement of W and is denoted by $W^{\perp}$.

Example: Let W be a plane through the origin in $\mathbb{R}^3$, and let L be the line through the origin and perpendicular to W. If $\mathbf{z}$ and $\mathbf{w}$ are nonzero, $\mathbf{z}$ is on L, and $\mathbf{w}$ is in W then the line segment from $\mathbf{0}$ to $\mathbf{z}$ is perpendicular to the line segment from $\mathbf{0}$ to $\mathbf{w}$; that is, $\mathbf{z}.\mathbf{w} = \mathbf{0}$.



So each vector on L is orthogonal to every w in W.

In fact, L consists of all vectors that are orthogonal to the $w's$ in W, and W consists of all vectors orthogonal to the $z's$ in L. That is,

$$L = W\perp \quad \text{and} \quad W = L^{\perp}.$$

Orthogonal Complements:

1. A vector $\mathbf{x}$ is in $W^{\perp}$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.

2. W is a subspace of $\mathbb{R}^n$.

Theorem: Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:

$$(\text{Row} \quad A)^{\perp} = \text{Nul} \quad A \quad \text{and} \quad (\text{Col} \quad A)^{\perp} = \text{Nul} \quad A^T.$$

Orthogonal Sets: A set of vectors $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an orthogonal set if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i.\mathbf{u}_j = 0$ whenever $i \neq j$.

Example: Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Theorem: If $S = \{\mathbf{u}_1, ..., \mathbf{u}_p\}$ 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 Basis: An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

Theorem: If $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$, For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + ... + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}.\mathbf{u}_j}{\mathbf{u}_j.\mathbf{u}_j}, \quad (j = 1, ..., p).$$

Example: The set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbb{R}^3$ where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}.$$

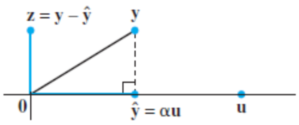
Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.

An Orthogonal Projection: Given a nonzero vector $\mathbf{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\mathbf{u}$ and the other orthogonal to $\mathbf{u}$. We wish to write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z}$ is some vector orthogonal to $\mathbf{u}$.

$$\hat{\mathbf{y}} = \frac{\mathbf{y}.\mathbf{u}}{\mathbf{u}.\mathbf{u}}\mathbf{u}.$$



Example: Let $\mathbf{y} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$. Then write $\mathbf{y}$ as the sum of two orthogonal vectors, one in Span $\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.

Solution: Compute

$$\mathbf{y}.\mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} . \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40, \quad \mathbf{u}.\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} . \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20.$$

The orthogonal projection is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}.\mathbf{u}}{\mathbf{u}.\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}.$$

27.3 Orthonormal Sets

A set of vectors $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an orthogonal set if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i.\mathbf{u}_j = 0$ whenever $i \neq j$ and $\mathbf{u}_i.\mathbf{u}_i = 1$.

Example: Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis in $\mathbb{R}^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}.$$

Theorem: An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.
Theorem: Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

1. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
2. $(U\mathbf{x}).(U\mathbf{y}) = \mathbf{x}.\mathbf{y}$
3. $(U\mathbf{x}).(U\mathbf{y}) = 0$ if and only if $\mathbf{x}.\mathbf{y} = 0$.

Example: Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Show that U has orthonormal columns and verify that $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix because it is square and because its columns are orthonormal. Verify that the rows are orthonormal, too.

27.4 Some Practice Problems

Question: Determine which set of vectors are orthogonal

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}.$$

Question: Compute orthogonal projection of $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ onto the line passing through $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the origin.

Question: Suppose W is a subspace of $\mathbb{R}^n$ spanned by n nonzero orthogonal vectors. Explain why $W = \mathbb{R}^n$.

Question: Let U be a square matrix with orthonormal columns. Explain why U is invertible.
Question: Let U be an $n \times n$ orthogonal matrix. how that the rows of U form an orthonormal basis of $\mathbb{R}^n$.

CHAPTER 28

Lecture No. 28

Question: Suppose W is a subspace of $\mathbb{R}^n$ spanned by n nonzero orthogonal vectors. Explain why $W = \mathbb{R}^n$.
Solution: We know that an orthogonal set of vectors is linearly independent. Hence form a basis for the space W. Consequently, dimension of W is n and $W = \mathbb{R}^n$.

28.1 Orthogonal Projections

The orthogonal projection of a point in $\mathbb{R}^2$ onto a line through the origin has an important analogue in $\mathbb{R}^n$.

Given a vector $\mathbf{y}$ and a subspace W in $\mathbb{R}^n$, there is a vector $\hat{\mathbf{y}}$ in W such that
(1) $\hat{\mathbf{y}}$ is the unique vector in W for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W, and
(2) $\hat{\mathbf{y}}$ is the unique vector in W closest to $\mathbf{y}$.

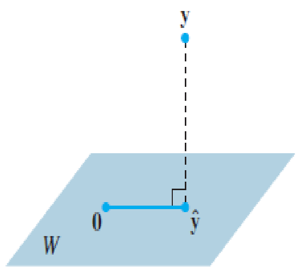


Figure 28.1: Shortest distance

These two properties of $\hat{\mathbf{y}}$ provide the key to finding least-squares solutions of linear systems
Example: Let $\{\mathbf{u}_1, ..., \mathbf{u}_5\}$ be an orthonormal basis for $\mathbb{R}^5$ and let

$$\mathbf{y} = c_1\mathbf{u}_1 + ... + c_5\mathbf{u}_5.$$

Consider the subspace $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, and write $\mathbf{y}$ as the sum of a vector $\mathbf{z}_1$ in W and a vector $\mathbf{z}_2$ in $W^\perp$.
Solution: Write $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_5\mathbf{u}_5$

where $\mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ is in $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ and $\mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5$ is in $\text{Span}\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.
Theorem: The Orthogonal Decomposition Theorem

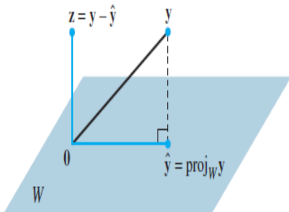
Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

In fact, if $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}.\mathbf{u}_1}{\mathbf{u}_1.\mathbf{u}_1}\mathbf{u}_1 + ... + \frac{\mathbf{y}.\mathbf{u}_p}{\mathbf{u}_p.\mathbf{u}_p}\mathbf{u}_p.$$



The vector $\hat{\mathbf{y}}$ is called the orthogonal projection of $\mathbf{y}$ onto W and is often is written $proj_W$.

Example: Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

Solution: The orthogonal projection of $\mathbf{y}$ onto W is

$$\begin{aligned} \hat{\mathbf{y}} &= \frac{\mathbf{y}.\mathbf{u}_1}{\mathbf{u}_1.\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}.\mathbf{u}_2}{\mathbf{u}_2.\mathbf{u}_2}\mathbf{u}_2 \\ &= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}. \end{aligned}$$

Also

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}.$$

28.2 A Geometric Interpretation of the Orthogonal Projection

When W is a subspace of $\mathbb{R}^3$ spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Here $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$ denote the projections of $\mathbf{y}$ onto the lines spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively.

The orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto W is the sum of the projections of $\mathbf{y}$ onto one-dimensional subspaces that are orthogonal to each other.

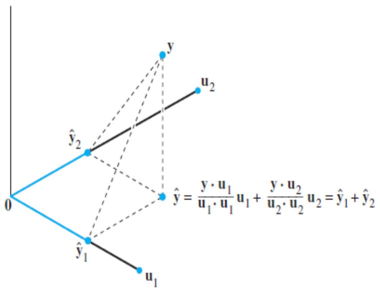


Figure 28.2: Geometric interpretation of orthogonal projection

Properties of Orthogonal Projections:

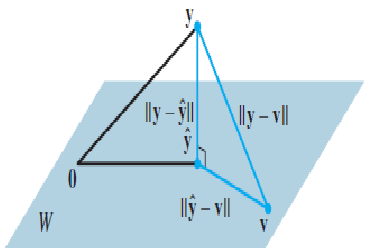
$$\text{If } \mathbf{y} \text{ is in } W = Span\{\mathbf{u}_1, ..., \mathbf{u}_p\}, \text{ then } proj_W\, \mathbf{y} = \mathbf{y}.$$

28.3 The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{b}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{b}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{b}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{b}}$.



Remark: The vector $\hat{\mathbf{b}}$ is called the best approximation to $\mathbf{y}$ by elements of W.

Example: If $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$ then the closest point in W to $\mathbf{y}$ is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}.\mathbf{u}_1}{\mathbf{u}_1.\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}.\mathbf{u}_2}{\mathbf{u}_2.\mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}.$$

Example: The distance from a point $\mathbf{y}$ in $\mathbb{R}^n$ to a subspace W is defined as the distance from $\mathbf{y}$ to the nearest point in W. Find the distance from $\mathbf{y}$ to $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$, and $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.

Solution: By the Best Approximation Theorem, the distance from $\mathbf{y}$ to W is $\|\mathbf{y}-\hat{\mathbf{y}}\|$, where $\hat{\mathbf{y}} = proj_W\mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W,

$$\begin{aligned}\hat{\mathbf{y}} &= \frac{\mathbf{y}.\mathbf{u}_1}{\mathbf{u}_1.\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}.\mathbf{u}_2}{\mathbf{u}_2.\mathbf{u}_2}\mathbf{u}_2 = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 \\ &= \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}.\end{aligned}$$

$$\|\mathbf{y}-\hat{\mathbf{y}}\|^2 = 0^2 + 3^2 + -^2 = 45.$$

Theorem: If $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$proj_W\ \mathbf{y} = (\mathbf{y}.\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}.\mathbf{u}_2)\mathbf{u}_2 + ... + (\mathbf{y}.\mathbf{u}_p)\mathbf{u}_p.$$

If $U = [\mathbf{u}_1\ \ \mathbf{u}_1\ ...\ \mathbf{u}_p]$ then

$$proj_W\ \mathbf{y} = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \in \mathbb{R}^n.$$

Example: Let $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, and $\mathbf{u}_2 = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$. Use this fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $proj_W\ \mathbf{y}$.

28.4 Some Practice Problems

Question: Let $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ -3 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$. Write $\mathbf{x}$ as the sum of two vectors, one in Span$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and the other in Span$\{\mathbf{u}_4\}$.

Question: Find the closed point to $\mathbf{y}$ in the subspace W spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$ $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$.

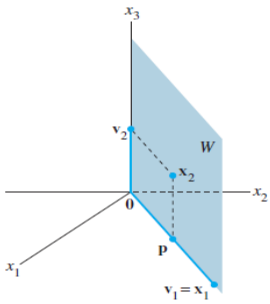
Question: Find find the best approximation to $\mathbf{x}$ by vectors of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ as the vectors in previous example.

CHAPTER 29

Lecture No. 29

29.1 The Gram-Schmidt process

The Gram-Schmidt process is a simple algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of $\mathbb{R}^n$.



Example: Let $W = Span\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for W.

Solution: Let $\mathbf{v}_1 = \mathbf{x}_1$ and

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2.\mathbf{x}_1}{\mathbf{x}_1.\mathbf{x}_1}\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Example: Let $W = Span\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, where $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for W which is subspace of $\mathbb{R}^4$.

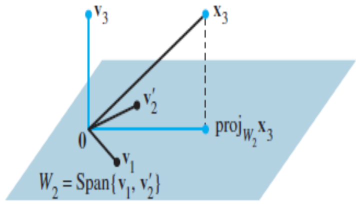
Solution: **Step I**: Let $\mathbf{v}_1 = \mathbf{x}_1$ and $W_1 = Span\{\mathbf{x}_1\} = Span\{\mathbf{v}_1\}$.

Step II: Let $\mathbf{v}_2$ be the vector produced by subtracting from $\mathbf{x}_2$ its projection onto the subspace W_1. That is, let

$$\mathbf{v}_2 = \mathbf{x}_2 - proj_{W_1}\mathbf{x}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2.\mathbf{v}_1}{\mathbf{v}_1.\mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

$\mathbf{v}_2$ is the component of $\mathbf{x}_2$ orthogonal to $\mathbf{x}_1$ and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for the subspace W_2 Spanned by $\mathbf{x}_1$ and $\mathbf{x}_2$.

Step II' Optional: Scale $\mathbf{v}_2$, we get $\mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.



Step III: Let $\mathbf{v}_3$ be the vector produced by subtracting from $\mathbf{x}_3$ its projection onto the subspace W_2. Use the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2'\}$ to compute this projection onto W_2:

$$proj_{W_2}\mathbf{x}_3 = \frac{\mathbf{x}_3.\mathbf{v}_1}{\mathbf{v}_1.\mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{x}_3.\mathbf{v}_2'}{\mathbf{v}_2'.\mathbf{v}_2'}\mathbf{v}_2' = \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Then $\mathbf{v}_3$ is the component of $\mathbf{x}_3$ orthogonal to W_2, namely,

$$\mathbf{v}_3 = \mathbf{x}_3 - proj_{W_2}\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

Theorem: The Gram-Schmidt Process

Given a basis $\{\mathbf{x}_1, ..., \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2.\mathbf{v}_1}{\mathbf{v}_1.\mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3.\mathbf{v}_1}{\mathbf{v}_1.\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3.\mathbf{v}_2}{\mathbf{v}_2.\mathbf{v}_2}\mathbf{v}_2 \\
&. \\
&. \\
&. \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p.\mathbf{v}_1}{\mathbf{v}_1.\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p.\mathbf{v}_2}{\mathbf{v}_2.\mathbf{v}_2}\mathbf{v}_2 - ... - \frac{\mathbf{x}_p.\mathbf{v}_{p-1}}{\mathbf{v}_{p-1}.\mathbf{v}_{p-1}}\mathbf{v}_{p-1}.
\end{aligned}$$

Then $\{\mathbf{v}_1, ..., \mathbf{v}_p\}$ is an orthogonal basis for W. In addition, we have

$$Span\{\mathbf{v}_1, ..., \mathbf{v}_k\} = Span\{\mathbf{x}_1, ..., \mathbf{x}_k\} \qquad \text{for } 1 \leq k \leq p.$$

Example: Let $W = Span\{\mathbf{x}_1, \mathbf{x}_2\}$, where

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}.$$

Construct an orthonormal basis for W.
Orthonormal Basis: Construct an orthonormal basis for the subspace spanned by the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Example: Find an orthogonal basis for the column space and null space of the matrix

$$\begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}.$$

29.2 Some Practice Problems

Question: Find an orthogonal basis for the column space and null space of each matrix

$$\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & 5 & 0 \\ -1 & -3 & 1 & 1 \\ 0 & 2 & 3 & 1 \\ 1 & 5 & 2 & 2 \end{bmatrix}.$$

Question: Mark each statement as true or false

1. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W , then multiplying $\mathbf{v}_3$ by a scalar c gives a new orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

2. If $W = Span\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

3. If $\mathbf{x}$ is not in a subspace W, then $\mathbf{x} - proj_W\mathbf{x}$ is not zero.

Question: Construct an orthogonal basis using Gram-Schmidt process for $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}.$

CHAPTER 30

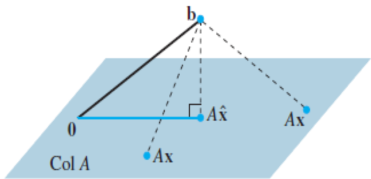
Lecture No. 30

30.1 Least Squares Solutions

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a least-squares solution of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.



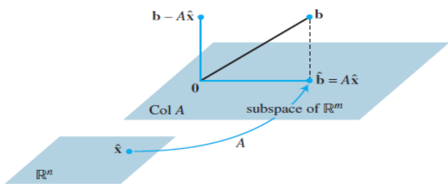
Remark: The most important aspect of the least-squares problem is that no matter what $\mathbf{x}$ we select, the vector $A\mathbf{x}$ will necessarily be in the column space, Col A. So we seek an $\mathbf{x}$ that makes $A\mathbf{x}$ the closest point in Col A to $\mathbf{b}$.

Solution of the General Least-Squares Problem: Given A and $\mathbf{b}$, apply the Best Approximation Theorem to the subspace Col A. Let

$$\hat{\mathbf{b}} = proj_{Col\ A}\mathbf{b}$$

Because $\hat{\mathbf{b}}$ is in the column space of A, the equation $A\mathbf{x} = \hat{\mathbf{b}}$ is consistent, and there is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}}.$$



Since $\hat{\mathbf{x}}$ is the closest point in Col A to $\mathbf{b}$, a vector $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Suppose $\hat{\mathbf{x}}$ satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. By the Orthogonal Decomposition Theorem the projection $\hat{\mathbf{b}}$ has the property that $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to Col A, so $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to each column of A.