MATH2130-F17

Farid Aliniaeifard

CU BOULDER

Content

Theorem 1.1

Let $\mathcal{B}$ be a basis for a vector space V. Then for each x in V, there exists unique set of scalars $\{c_1, \ldots, c_n\}$ such that

$$x = c_1 b_1 + \ldots + c_n b_n.$$

Proof. Since $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis there are scalars $c_1, \ldots, c_n$ such that $x = c_1 b_1 + \ldots + c_n b_n$. Suppose also x has the representation

$$x = d_1 b_1 + \ldots + d_n b_n.$$

Then

$$0 = x - x = (c_1 - d_1) b_1 + \ldots + (c_n - d_n) b_n.$$

Note that $\{b_1, \ldots, b_n\}$ is linearly independent, so

$$c_1 - d_1 = 0, \ldots, c_n - d_n = 0 \Rightarrow c_1 = d_1, \ldots, c_n = d_n.$$

Definition 1.2

Suppose $\mathcal{B} = \{b_1, \ldots, b_n\}$ *is a basis for* V *and* x *is in* V. *Let*

$$x = c_1 b_1 + \ldots + c_n b_n.$$

The **coordinate vector for** x **relative to the basis** $\mathcal{B}$ *is*

$$[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Note that $[x]_{\mathcal{B}} \in \mathbb{R}^n$ *for any basis* $\mathcal{B}$ *of* V.

- **Coordinates in $\mathbb{R}^n$**

Example 1.3

Let $\mathcal{B} = \{b_1, b_2\}$ *be a basis for* $\mathbb{R}^2$ *where* $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $b_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *If* $[x]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. *Find* x.

Solution. $[x]_{\mathcal{B}} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 4 \end{bmatrix}$.

Example 1.4

Let $\mathcal{B}$ be the standard basis for $\mathbb{R}^2$, i.e., $\mathcal{B} = \{e_1, e_2\}$, where $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Let $x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ what is $[x]_{\mathcal{B}}$?

Solution. Since $\begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3e_1 + e_2$, we have $[x]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

- If $\mathcal{B}$ is the standard basis for $\mathbb{R}^n$, then $[x]_{\mathcal{B}} = x$.

Example 1.5

Let $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, *and* $x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, *and* $\mathcal{B} = \{b_1, b_2\}$. *find the coordinate vector* $[x]_{\mathcal{B}}$.

Solution. We have that $[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ where

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

i.e.,

$$\begin{bmatrix} 2c_1 - c_2 \\ c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

we can write it as

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Then you can solve this equation and find $c_1 = 3$ and $c_2 = 2$.

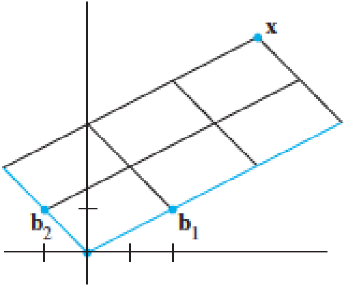


FIGURE 4

The $\mathcal{B}$-coordinate vector of $\mathbf{x}$ is $(3, 2)$.

In the above example the matrix

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

has a especial name.

Definition 1.6

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for $\mathbb{R}^n$. The matrix

$$P_{\mathcal{B}} = [b_1 | \ldots | b_n]$$

is called the **change-of-coordinates matrix from $\mathcal{B}$ to the standard basis of $\mathbb{R}^n$**. *Also when $x = c_1 b_1 + \ldots + c_n b_n$, we have*

$$x = P_{\mathcal{B}}[x]_{\mathcal{B}} = P_{\mathcal{B}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Remark.

1. The matrix $P_{\mathcal{B}}$ is an $n \times n$ matrix.
2. The columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, so $P_{\mathcal{B}}$ is invertible.
3. We can also write $P_{\mathcal{B}}^{-1}x = [x]_{\mathcal{B}}$.

- The coordinate mapping

Theorem 1.7

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping

$$\begin{aligned} T: \quad V &\to \mathbb{R}^n \\ x &\mapsto [x]_{\mathcal{B}} \end{aligned}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

Proof.

Let $u = c_1b_1 + \ldots + c_nb_n$ and $w = d_1b_1 + \ldots + d_nb_n$. Then

$$u + w = (c_1 + d_1)b_1 + \ldots + (c_n + d_n)b_n.$$

It follows that

$$[u + w]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [u]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

Now let $r \in \mathbb{R}$,

$$ru = r(c_1b_1 + \ldots + c_nd_n) = (rc_1)b_1 + \ldots + (rc_n)d_n.$$

Therefore,

$$[ru]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[u]_{\mathcal{B}}.$$

Definition 1.8

A linear transformation T from a vector space V to a vector space W is an isomorphism if T is one-to-one and onto. Moreover, we say V and W are **isomorphic**.

Week 9, Lecture 2, Oct.25, Linearly independent sets, basis, and dimension

Theorem 1.9

Let V and W be vector spaces, and $T : V \to W$ be a linear transformation. Then

1. *T is one-to-one if ker $(T) = \{v \in V : T(v) = 0\} = \{0\}$.*
2. *T is onto if range$(T) = \{T(v) : v \in V\} = W$.*

Definition 1.10

A linear transformation T from a vector space V to a vector space W is an **isomorphism** *if T is one-to-one and onto. Moreover, we say V and W are* **isomorphic**.

Theorem 1.11

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping

$$\begin{array}{rccc} T: & V & \to & \mathbb{R}^n \\ & x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

Solution. Previously we showed that T is a linear transformation. Now, we will show that it is one-to-one and onto.
one-to-one: $ker(T) = \{x \in V : [x]_{\mathcal{B}} = 0\}$. Note that if $[x]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, then $x = 0b_1 + \ldots + 0b_n = 0$. Therefore, $ker(T) = 0$ and so T is one-to-one.

onto: For any $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$, there is a vector $x = y_1 b_1 + \ldots + y_n b_n \in V$ such that $[x]_{\mathcal{B}} = y$.

Definition 1.12

Let $f(t) = a_0 + a_1 t + \ldots + a_n t^n = 0$ be a non-zero polynomial. A root for f is a number c such that

$$f(c) = a_0 + a_1 c + \ldots + a_n c^n = 0,$$

for example $f(t) = t^2 - 1$ has roots 1 and -1.

Theorem 1.13

Every polynomial in $\mathbb{P}_n$ has at most n roots.

Example 1.14

$S = \{1, t, t^2, \ldots, t^n\}$ is a basis for $\mathbb{P}_n$.

Solution. Any polynomial is of the form

$$f(t) = a_0 + a_1 t + \ldots + a_m t^m$$

where $m \leq n$ so $f(t) \in span\{1, t, \ldots, t^n\}$.
Now, we should show that $\{1, t, \ldots, t^n\}$ are linearly independent.
Let

$$c_0 + c_1 t + \ldots + c_n t^n = 0,$$

then it means the polynomial $c_0 + c_1 t + \ldots + c_n t^n$ has infinitely many roots which is not possible because every polynomial of degree at most n has at most n roots.

Example 1.15

Let $B = \{1, t, t^2, t^3\}$ be the standard basis for $\mathbb{P}_3$. Show that $\mathbb{P}_3$ is isomorphic to $\mathbb{R}^4$.

Solution. By Theorem 1.11 we have

$$T : \mathbb{P}_3 \longrightarrow \mathbb{R}^4$$

$$p = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \mapsto [p]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

is a isomorphism.

Example 1.16

Let

$$v_1 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix} \quad x = \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$$

and $B = \{v_1 v_2\}$*. Then* $\mathcal{B}$ *is a basis for* $H = span\{v_1, v_2\}$*. Determine if* x *is in* H*. Find* $[x]_{\mathcal{B}}$*.*

Solution. If the following system is consistent

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$$

Then $\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$ is in $span\{v_1, v_2\}$. The augmented matrix is

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 4 \\ 1 & -3 & -1 \end{bmatrix}$$

An echelon form is

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

so the system is consistent and if you solve it, you have $c_1 = 2$ and $c_2 = 1$. Therefore $[x]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Theorem 1.17

Let $T : V \longrightarrow W$ be an isomorphism. Then $v_1, \ldots, v_n$ are linearly independent (dependent) in V if and only if $T(v_1), \ldots, T(v_n)$ are linearly independent (dependent) in W.

Example 1.18

Verify that the polynomials $1+2t^2$, $4+t+5t^2$, and $3+2t$ are linearly independent.

Solution. Let $\mathcal{B}=\{1,t,t^2,t^3\}$ be the standard basis for $\mathbb{P}_3$. We have by Theorem 1.11 $T:\mathbb{P}_3 \longrightarrow \mathbb{R}^4$ where

$$p \mapsto [p]_B$$

is an isomorphism. Therefore by theorem above $1+2t^2$, $4+t+5t^2$ and $3+2t$ are linearly independent if and only if $\left[1+2t^2\right]_B$, $\left[4+t+5t^2\right]_B$, and $[3+2t]_B$ are linearly independent. So

$$\left[1+2t^2\right]_B=\begin{bmatrix}1\\0\\2\\0\end{bmatrix}, \left[4+t+5t^2\right]_B=\begin{bmatrix}4\\1\\5\\0\end{bmatrix}, [3+2t]_B=\begin{bmatrix}3\\2\\0\\0\end{bmatrix}$$

Therefore, we only need to show that

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

are linearly dependent. (Do it as an Exercise).

Week 9, Lecture 3, Oct.25, the dimension of vector space

Theorem 1.19

Let $T : V \longrightarrow W$ *be an isomorphism.*

1. $v_1, \ldots, v_n$ *are linearly independent (dependent) in* V *if and only if* $T(v_1), \ldots, T(v_n)$ *are linearly independent (dependent) in* W*.*
2. *A vector* x *is in* $span\{v_1, \ldots, v_n\}$ *if and only if* $T(x)$ *is in* $span\{T(v_1), \ldots, T(v_n)\}$*.*

Example 1.20

1. *Verify that the polynomials* $1+2t^2$, $4+t+5t^2$, *and* $3+2t$ *are linearly independent.*
2. *Is* $g(t) = t - 3t^2$ *in span*$\{1+2t^2, 4+t+5t^2, 3+2t\}$*?*

Proof. (1) Let $\mathcal{B} = \{1, t, t^2, t^3\}$ be the standard basis for $\mathbb{P}_3$. We have by Theorem 1.11 $T : \mathbb{P}_3 \longrightarrow \mathbb{R}^4$ where

$$p \mapsto [p]_B$$

is an isomorphism. Therefore by theorem above $1 + 2t^2$, $4 + t + 5t^2$ and $3 + 2t$ are linearly independent if and only if

$$\left[1 + 2t^2\right]_B, \left[4 + t + 5t^2\right]_B, [3 + 2t]_B$$

are linearly independent.

We have

$$[1+2t^2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, [4+t+5t^2]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, [3+2t]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, we only need to show that

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

are linearly independent. (Do it as an Exercise).

(2) By the above theorem we only need to show that

$$[g(t)]_{\mathcal{B}} \in span \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\},$$

i.e.,

$$\begin{bmatrix} 0 \\ 1 \\ -3 \\ 0 \end{bmatrix} \in span \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$\square$

•The dimension of a vector space

Theorem 1.21

If a vector space V has a basis $B = \{b_1, \ldots, b_n\}$ then any set containing more than n vectors must be linearly dependent.

Theorem 1.22

If V is a vector space and V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

Definition 1.23

1. *A vector space is said to be* **finite-dimensional** *if it is spanned by a finite set of vectors in* V
2. **Dimension** *of* V*, dim* V*, is the number of vectors in a basis of* V*. Also dimension of zero space* $\{0\}$ *is* 0*.*
3. *If* V *is not spanned by a finite set, then* V *is said to be* **infinite-dimensional**.

Example 1.24

Find dimension of the subspace

$$H = \left\{ \begin{bmatrix} a-3b+c \\ 2a+2d \\ b-3c-d \\ 2d-b \end{bmatrix} : a, b, c, d \text{ in } \mathbb{R} \right\}.$$

Solution. We have

$$\begin{bmatrix} a-3b+c \\ 2a+2d \\ b-3c-d \\ 2d-b \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \\ -1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ -3 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 2 \\ -1 \\ 2 \end{bmatrix}$$

Therefore,

$$H = span\left\{\begin{bmatrix}1\\2\\0\\0\end{bmatrix},\begin{bmatrix}-3\\0\\1\\-1\end{bmatrix},\begin{bmatrix}1\\0\\-3\\0\end{bmatrix},\begin{bmatrix}0\\2\\-1\\2\end{bmatrix}\right\}$$

Now, we want to find a basis for H, we had a process for finding the basis.(Do it as an exercise.)

Theorem 1.25

Let H be a subspace of a finite dimensional vector space V. Any linearly independent set in H can be expanded to a basis for H. Also

$$dim\ H \leq dim\ V$$

Theorem 1.26

(**The Basis Theorem**) *Let V be a p-dimensional vector space $p \geq 1$.*

1. *Any linearly independent set of exactly p elements in V is automatically a basis for V.*
2. *Any set of exactly p elements that spans V is automatically a basis for V.*

Remember: The dimension of $Nul\ A$ is the number of free variables in the equation $Ax = 0$, and the dimension of $Col\ A$ is the number of pivot columns in A, and the pivot columns of A gives a basis for column space of A.

Week 10, Lecture 1, Oct.30, change of basis

Example 1.27

Let $b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, c_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *Then* $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$ *are two basis for* $\mathbb{R}^2$. *Let* $x = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$. *Then*

$$x = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = b_1 + 2b_2$$

Therefore, $[x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. *Also*

$$x = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2c_1 + 0c_2 \text{ so } [x]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Then there is a matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ such that

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}} = [[b_1]_{\mathcal{C}} \quad [b_2]_{\mathcal{C}}][x]_{\mathcal{B}}.$$

Since

$$b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = (-1)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-1)c_1 + c_2$$

we have

$$[b_1]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Also

$$b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 3/2\begin{bmatrix} 0 \\ 1 \end{bmatrix} + (-1/2)\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 3/2c_1 - 1/2c_2$$

Therefore,

$$[x]_{\mathcal{C}} = \begin{bmatrix} -1 & 3/2 \\ 1 & -1/2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Theorem 1.28

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ *and* $\mathcal{C} = \{c_1, \ldots, c_n\}$ *be bases of a vector space* V. *Then there is a unique matrix* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *such that*

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}$$

The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *are the* $\mathcal{C}$*-coordinate vectors of the vectors in the basis* $\mathcal{B}$. *That is,*

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = [[b_1]_{\mathcal{C}} \quad [b_2]_{\mathcal{C}} \quad \ldots \quad [b_n]_{\mathcal{C}}].$$

Definition 1.29

The matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *in the above theorem is called* **change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$**.

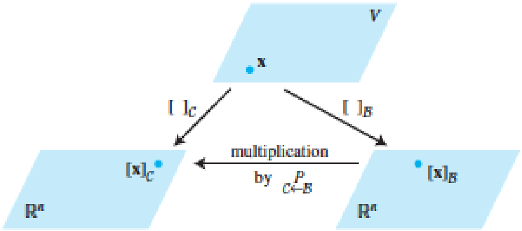


FIGURE 2 Two coordinate systems for V.

Remark. We have

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}$$

so

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1}[x]_{\mathcal{C}} = [x]_{\mathcal{B}}$$

Therefore,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$$

• Change of Basis in $\mathbb{R}^n$

Remark.

1. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ a basis for $\mathbb{R}^n$. Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. Then $P_{\mathcal{B}} = [b_1|\ldots|b_n]$ is the same as $\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$.
2. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{C} = \{c_1, \ldots, c_n\}$ be bases for $\mathbb{R}^n$. Then by row operation we can reduce the matrix

$$[c_1 \quad \ldots \quad c_n | b_1 \quad \ldots \quad b_n]$$

to

$$[I|\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}].$$

Example 1.30

Let $b_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $c_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, *and* $c_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, *and consider the bases for* $\mathbb{R}^2$ *given by* $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$. *Find the change-of-coordinate matrix from* $\mathcal{B}$ *to* $\mathcal{C}$.

Solution. We can reduce the matrix $[c_1 \; c_2 | b_1 \; b_2]$ to $[I | \underset{\mathcal{C} \leftarrow \mathcal{B}}{P}]$, and so we can find $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$. Therefore, we have

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \underset{\longleftrightarrow}{\text{Replace R2 by R2+4R1}}$$

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right] \underset{\longleftrightarrow}{\text{Scaling R2 by 1/7}}$$

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array}\right] \xleftrightarrow{\text{Replace R1 by R1}-3\text{R2}} \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right]$$

Therefore,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}.$$

Example 1.31

Let $b_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $b_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $c_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $c_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, *and consider the bases for* $\mathbb{R}^2$ *given by* $\mathcal{B} = \{b_1, b_2\}$ *and* $\mathcal{C} = \{c_1, c_2\}$.

1. *Find the change-of-coordinates matrix from* $\mathcal{C}$ *to* $\mathcal{B}$.
2. *Find the change-of-coordinates matrix from* $\mathcal{B}$ *to* $\mathcal{C}$.

Solution. (1) Note that we need to find $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$, so compute

$$[b_1 \ \ b_2 | c_1 \ \ c_2] = \left[\begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \leftrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right].$$

Therefore,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}.$$

(2) We now want to compute $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Note that

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}.$$

Remark. Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ and $\{c_1, \ldots, c_n\}$ be bases for $\mathbb{R}^n$. We have (see week 9, lecture 2)

$$P_{\mathcal{B}} = [b_1|b_2|\ldots|b_n] \qquad P_{\mathcal{C}} = [c_1|c_2|\ldots|c_n].$$

It was shown that

$$x = P_{\mathcal{B}}[x]_{\mathcal{B}} \qquad x = P_{\mathcal{C}}[x]_{\mathcal{C}}.$$

So we have

$$P_{\mathcal{C}}[x]_{\mathcal{C}} = P_{\mathcal{B}}[x]_{\mathcal{B}}.$$

Therefore,

$$[x]_{\mathcal{C}} = P_{\mathcal{C}}^{-1} P_{\mathcal{B}}[x]_{\mathcal{B}}.$$

We also have

$$[x]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[x]_{\mathcal{B}}.$$

So,

$$P_{\mathcal{C}}^{-1} P_{\mathcal{B}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$$

• Change of basis for polynomials

Example 1.32

Let $\mathcal{B} = \{1+t, 1+t^2, 1+t+t^2\}$ *and* $\mathcal{C} = \{2-t, -t^2, 1+t^2\}$ *be bases for* $\mathbb{P}_2$. *Find* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.

Solution. Solution. Let $\mathcal{E} = \{1, t, t^2\}$ be the standard basis for $\mathbb{P}_2$. Then

$$\begin{array}{rccc} T: & \mathbb{P}_2 & \to & \mathbb{R}^3 \\ & f & \mapsto & [f]_{\mathcal{E}} \end{array}$$

is an isomorphism.We have

$$[1+t]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, [1+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, [1+t+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$[2-t]_{\mathcal{E}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, [-t^2]_{\mathcal{E}} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, [1+t^2]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Now we have

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$\mathcal{C} = \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

be bases for $\mathbb{R}^3$. We are looking for the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$.

Week 10, Lecture 2, Nov. 1, Eigenvalues and eigenvectors

Example 1.33

Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $u = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *Then*

$$Au = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

$$Av = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Precisely we have $Av = 2v$.

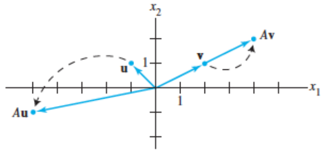


FIGURE 1 Effects of multiplication by A.

Definition 1.34

An **eigenvector** *of an* $n \times n$ *matrix* A *is a nonzero vector* x *such that* $Ax = \lambda x$ *for some scalar* λ. *A scalar* λ *is called an* **eigenvalue** *of* A *if there is a nonzero vector* x *such that* $Ax = \lambda x$; *such* x *is called an* **eigenvector corresponding to** λ.

Example 1.35

Let $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$, $v = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$, $u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

$$Av = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -12 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

so $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$ *is an eigenvector and* 3 *is an eigenvalue.* $Au = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ *for any* λ.

Example 1.36

Show that 7 is an eigenvalue of $A = \begin{bmatrix} 1 & 5 \\ 6 & 2 \end{bmatrix}$.

Solution. The number 7 is an eigenvalue. For some vector x we have

$$Ax = 7x$$

so

$$Ax - 7x = 0$$

we can write the above equation as

$$(A - 7I)x = 0$$

so if $(A - 7I)x = 0$ has a nonzero solution say x', then

$$(A - 7I)x' = 0 \Rightarrow Ax' - 7x' = 0$$

$$\Rightarrow Ax' = 7x'$$

and so 7 is an eigenvalue.

Therefore, we only need to solve

$$(A - 7I)x = 0, \quad i.e.,$$

$$\left(\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 7\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

when we solve the equation we have at least a nonzero solution $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Therefore 7 is an eigenvalue.

- **How to find all eigenvalues of a matrix A.**

λ is an eigenvalue for A if and only if

$$Ax = \lambda x \quad at\ least\ for\ a\ nonzero\ vector\ x.$$

So we can say λ is an eigenvalue of a matrix A if and only if

$$(A - \lambda I)x = 0 \quad at\ least\ for\ some\ nonzero\ x.$$

Which means the equation $(A - \lambda I)x = 0$ does not have only trivial solution if and only if

$$det(A - \lambda I) = 0.$$

Lemma 1.37

λ is an eigenvalue of A if and only if

$$det(A - \lambda I) = 0.$$

Definition 1.38

The equation $det(A - \lambda I) = 0$ *is called the* **characteristic equation**.

Definition 1.39

Let λ *be an eigenvalue of* $n \times n$ *matrix* A. *Then the* **eigenspace of** A **corresponding to** λ *is the solution set of*

$$(A - \lambda I)x = 0$$

Remark. Note that we already have the solution set of

$$(A - \lambda I)x = 0$$

is a subspace.

Example 1.40

let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

(a) Find all eigenvalues of A.

(b) For each eigenvalue λ *of* A*, find a basis for the eigenspace of* A *corresponding to* λ.

(a) To find all eigenvalues of A we must find all λ such that

$$det(A - \lambda I) = 0.$$

Note that

$$det(A - \lambda I) = det\left(\begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}\right) = 0$$

$$\Rightarrow det\left(\begin{bmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{bmatrix}\right) = 0$$

you already know how to compute the determinant. We have

$$det\left(\begin{bmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{bmatrix}\right) = -(\lambda - 9)(\lambda - 2)^2$$

so $\lambda = 9$ and $\lambda = 2$, are the eigenvalues of A.

(b) We first find the basis for eigenspace of A corresponding to $\lambda = 2$, which is the same as the finding the basis of the solution set of $(A-2I)x = 0$ which means we should find the basis for null space of $A - 2I$ (you know how to do it). The null space of $A - 2I$ contains all vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ such that $(A - 2I)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$. i.e.,

$$\begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

The augmented matrix is

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix}$$

and the reduced echelon form is

$$\begin{bmatrix} 1 & -1/2 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So x_1 is basic and x_2 and x_3 are free. We have $x_1 - 1/2x_2 + 3x_3 = 0$

$$\Rightarrow x_1 = 1/2x_2 - 3x_3$$

Let $x_2 = t$ and $x_3 = s$. Then

$$x_1 = 1/2t - 3s.$$

So

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1/2t - 3s \\ t \\ s \end{bmatrix} = t \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

so the eigenspace of A corresponding to 2 is

$$\left\{ t \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

and the basis for the eigenspace of A corresponding to 2 is

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Now you will find the eigenspace and the basis of it for $\lambda = 9$ (Do it as an exercise).

Week 10, Lecture 3, Nov. 3, Characteristic polynomial and diagonalization

Theorem 1.41

The eigenvalues of a triangular matrix are the entries on its main diagonal.

Example 1.42

Let $A = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$. *Then eigenvalues of* A *are* a, d, *and* f. *Why? because*

$$det(A - \lambda I) = det(\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}) =$$

$$det(\begin{bmatrix} a-\lambda & b & c \\ 0 & d-\lambda & e \\ 0 & 0 & f-\lambda \end{bmatrix}) = (a-\lambda)(d-\lambda)(f-\lambda)$$

Therefore, the eigenvalues are a, d and f, the entries on the main diagonal.

Theorem 1.43

If $v_1, \dots, v_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{v_1, \dots, v_r\}$ is linearly independent.

Example 1.44

let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. *Then* 2 *and* 9 *are eigenvalues of* A.

The eigenspace corresponding to 2 *has a basis*

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Also, the eigenspace corresponding to 9 has a basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Then

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad and \quad \left\{ \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

are linearly independent.

- **When 0 is an eigenvalue of an $n \times n$ matrix A:**

If 0 is an eigenvalue, then there is a nonzero vector x such that $Ax = 0x$

$$\Rightarrow \quad Ax = 0$$

which means that $Ax = 0$ has a nonzero solution, which also means A is not invertible and $det\ A = 0$.

Theorem 1.45

Let A be an $n \times n$ matrix. Then A is invertible if and only if one of the following holds:

1. *The number 0 is not eigenvalue of A.*
2. *The determinant of A is not zero.*

- **Similarity:**

Definition 1.46

Two $n \times n$ matrices A and B are said to be **similar** *if there exists an invertible matrix P such that $A = PBP^{-1}$.*

Definition 1.47

The expression $\det(A-\lambda I)$ is called the **characteristic polynomial**.

Let A and B are similar. Then there exists an invertible matrix P such that

$$A = PBP^{-1} \qquad \Leftrightarrow \qquad A - \lambda I = PBP^{-1} - \lambda I$$

Note that $PP^{-1} = I$, so

$$A - \lambda I = PBP^{-1} - \lambda PP^{-1} = P(B - \lambda I)P^{-1}$$

Now

$$\begin{aligned} det(A - \lambda I) &= det(P(B - \lambda I)P^{-1}) \\ &= det(P)det(B - \lambda I)det(P^{-1}) \\ &= det(P)det(P^{-1})det(B - \lambda I) = det(B - \lambda I) \end{aligned}$$

Therefore, A and B have the same characteristic polynomial and so they have the same eigenvalues.

Proposition 1.48

Similar matrices have the same characteristic polynomial and so they have the same eigenvalues.

• Diagonalization (Heads up)

Example 1.49

If $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, *Then*

$$D^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

$$D^3 = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

and for k *we have*

$$D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$$

Definition 1.50

A matrix D is a **diagonal matrix** *if it is of the form*

$$\begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}.$$

Definition 1.51

A matrix is called **diagonalizable** *if A is similar to a diagonal matrix, i.e., there is an invertible matrix P and a diagonal matrix D such that*

$$A = PDP^{-1}.$$

Theorem 1.52

An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

Example 1.53

- *How to diagonalize a matrix:*
  1. *First check that if the matrix has n linearly dependent eigenvectors, if so, the matrix is diagonalizable.*
  2. *Find a basis for the set of all eigenvectors, say $\{v_1, \ldots, v_n\}$.*
  3. *Let $P = [v_1|\ldots|v_n]$, then $D = P^{-1}AP$ is an diagonal matrix with eigenvalues on its diagonal.*

Example 1.54

Find if $A = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}$ *is diagonalizable, if so find an invertible matrix* P *and a diagonal matrix* D *such that* $D = P^{-1}AP$.

Solution. First we should find basis for eigenspaces. Note that $det(A-\lambda I) = (1-\lambda)(-3-\lambda)$. So, A has two eigenvalues 1 and -3. The eigenspace corresponding to 1 has the basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ and the eigenspace corresponding to -3 has the basis $\left\{ \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} \right\}$. Then we have $P = \begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix}$, and $D = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$. Check that $D = P^{-1}AP$.

Week 11, Lecture 1, Nov. 6, Diagonalization

Example 1.55

If $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, *Then*

$$D^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

$$D^3 = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

and for k *we have*

$$D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$$

Definition 1.56

A matrix D is a **diagonal matrix** *if it is of the form*

$$\begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}.$$

Definition 1.57

A matrix is called **diagonalizable** *if A is similar to a diagonal matrix, i.e., there is an invertible matrix P and a diagonal matrix D such that*

$$A = PDP^{-1}.$$

Example 1.58

Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. *Find a formula for* A^k, *given that* $A = PDP^{-1}$. *Where* $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ *and* $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$.

Solution. We can find the inverse of P which is

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Then

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} =$$

$$PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}^2 \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix} =$$

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

Again,

$$A^3 = AA^2 = (PDP^{-1})(PD^2P^{-1}) =$$

$$PD(P^{-1}P)D^2P^{-1} = PD^3P^{-1}.$$

In general, for $k >= 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 2.5^k - 3^k & 5^k - 3^k \\ 2.3^k - 2.5^k & 2.3^k - 5^k \end{bmatrix}.$$

Theorem 1.59

(The diagonal theorem) An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

Definition 1.60

An eigenvector basis of $\mathbb{R}^n$ corresponding to A is a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ such that $v_1, \ldots, v_n$ are eigenvectors of A.

- An $n \times n$ matrix A is diagonalizable if and only if there are eigenvectors $v_1, \ldots, v_n$ such that $\{v_1, \ldots, v_n\}$ are a basis for $\mathbb{R}^n$, i.e., $\{v_1, \ldots, v_n\}$ is an eigenvector basis for $\mathbb{R}^n$ corresponding to A.

Week 11, Lecture 2, Nov. 8, diagonalizable matrices, eigenvectors and linear transformations

How to diagonalize an $n \times n$ matrix A.

Step 1. First find the eigenvalues of A.
Step 2. Find a basis for each eigenspace. That is, if

$$det(A - \lambda I) = (x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2} \dots (x - \lambda_p)^{k_p},$$

we should find the basis of eigenspace corresponding to each λ_i.
Step 3. If the number of all vectors in bases in Step 2 is n, then A is diagonalizable, otherwise it is not and we stop.
Step 4. Let $v_1, v_2, \dots, v_n$ be all vectors in bases in Step 2, then

$$P = [v_1|v_2|\dots|v_n].$$

Step 5. Constructing D form eigenvalues. If the multiplicity of an eigenvalue λ_i is k_i, we repeat λ_i, k_i times, on the diagonal of D.

Example 1.61

Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Solution. Step 1. Find eigenvalues of A.

$$0 = det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2.$$

Therefore, $\lambda = 1$ and $\lambda = -2$ are the eigenvalues.

Step 2. Find a basis for each eigenspace. The eigenspace corresponding to $\lambda = 1$ is the solution set of

$$(A - I)x = 0.$$

A basis for this space is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

The eigenspace corresponding to $\lambda = -2$ is the solution set of

$$(A - (-2)I)x = 0.$$

A basis for this space is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Step 3. Since we find three vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

So A is diagonalizable.

Step 4.

$$P = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Step 5.

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

It is a good idea to check that P and D work, i.e.,

$$A = PDP^{-1} \quad \text{or} \quad AP = PD.$$

If we compute we have

$$AP = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \qquad PD = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

Example 1.62

Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution. First we find the eigenvalues, which are the roots of characteristic polynomial $det(A - \lambda I)$.

$$0 = det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

So $\lambda = 1$ and $\lambda = -2$ are eigenvalues.
A basis for eigenspace corresponding to $\lambda = 1$ is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

and a basis for eigenspace corresponding to $\lambda = -2$ is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Since we can not find 3 eigenvectors that are linearly independent, so A is not diagonalizable.

Theorem 1.63

An $n \times n$ matrix with n distinct eigenvalues i.e.,

$$det(A-\lambda I) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n) \quad \textit{with distinct } \lambda_i\textit{'s,}$$

is diagonalizable.

Theorem 1.64

Let characteristic polynomial of A is

$$(x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2}\dots(x-\lambda_p)^{k_p}.$$

1. *For each $1 \leq i \leq p$ The dimension of eigenspace corresponding to λ_i is at most k_i.*
2. *The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if*
    1. *the characteristic polynomial factors completely into linear factors and*
    2. *the dimension of the eigenspace for each λ_i equals the multiplicity of λ_i.*

If A is diagonalizable and $\mathcal{B}_i$ is a basis for the eigenspace corresponding to λ_i for each i, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Week 11, Lecture 3, Nov. 10, Eigenvectors and linear transformations

- **Eigenvectors and linear transformations**

When A is diagonalizable there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Our goal is to show that the following two linear transformations are essentially the same.

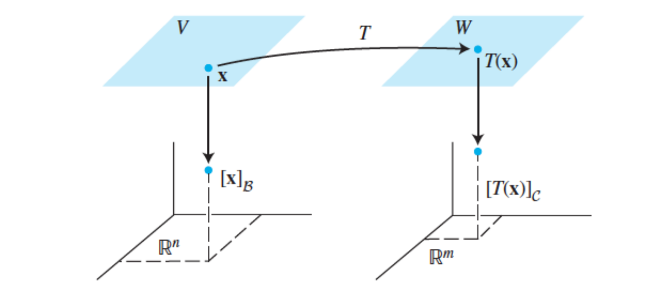
$$\begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ x & \mapsto & Ax \end{array} \qquad\qquad \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ u & \mapsto & Du \end{array}$$

Remark. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping

$$\begin{array}{rccc} T: & V & \to & \mathbb{R}^n \\ & x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

is a one-to-one linear transformation form V onto $\mathbb{R}^n$.

• The matrix of a linear transformation: Let V be an n-dimensional vector space and W be an m-dimensional vector space.



Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. The connection between $[x]_{\mathcal{B}}$ and $[T(x)]_{\mathcal{C}}$ is easy to find. Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be the basis of V. If $x = r_1b_1 + r_2b_2 + \ldots + r_nb_n$, then

$$x_{\mathcal{B}} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}.$$

Note that

$$T(x) = T(r_1b_1 + r_2b_2 + \ldots + r_nb_n) = r_1T(b_1) + r_2T(b_2) + \ldots + r_nT(b$$

Since the coordinate mapping from W to $\mathbb{R}^m$ is a linear transformation, we have

$$[T(x)]_{\mathcal{C}} = [r_1T(b_1) + r_2T(b_2) + \ldots + r_nT(b_n)]_{\mathcal{C}} =$$

$$r_1[T(b_1)]_{\mathcal{C}} + r_2[T(b_2)]_{\mathcal{C}} + \ldots + r_n[T(b_n)]_{\mathcal{C}} =$$

$$\begin{bmatrix} [T(b_1)]_{\mathcal{C}} & [T(b_2)]_{\mathcal{C}} & \dots & [T(b_n)]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} =$$

$$\begin{bmatrix} [T(b_1)]_{\mathcal{C}} & [T(b_2)]_{\mathcal{C}} & \dots & [T(b_n)]_{\mathcal{C}} \end{bmatrix} [x]_{\mathcal{B}}.$$

So

$$[T(x)]_{\mathcal{C}} = M[x]_{\mathcal{B}},$$

where

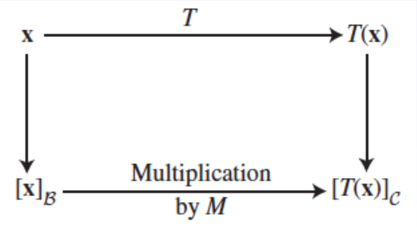
$$M = \begin{bmatrix} [T(b_1)]_{\mathcal{C}} & [T(b_2)]_{\mathcal{C}} & \dots & [T(b_n)]_{\mathcal{C}} \end{bmatrix}.$$

Theorem 1.65

Let V be an n-dimensional vector space with basis $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$, and let W be an m-dimensional vector space with basis $\mathcal{C}$. If T is a linear transformation form V to W, then

$$[T(x)]_{\mathcal{C}} = M[x]_{\mathcal{B}},$$

where $M = \begin{bmatrix} [T(b_1)]_{\mathcal{C}} & [T(b_2)]_{\mathcal{C}} & \ldots & [T(b_n)]_{\mathcal{C}} \end{bmatrix}$. M **is called matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$.**



Example 1.66

Let $\mathcal{B} = \{b_1, b_2\}$ be a basis for V and $\mathcal{C} = \{c_1, c_2, c_3\}$ be a basis for W. Let $T : V \to W$ be a linear transformation such that

$$T(b_1) = 3c_1 - 2c_2 + 5c_3 \qquad T(b_2) = 4c_1 + 7c_2 - c_3$$

Find matrix M for T relative to $\mathcal{B}$ and $\mathcal{C}$.

Solution. We have that

$$M = [[T(b_1)]_{\mathcal{C}} \quad [T(b_2)]_{\mathcal{C}}].$$

We have

$$[T(b_1)] = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \qquad [T(b_2)] = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}.$$

So

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}.$$

• **Linear transformation from V into V**

Now, we want to find the matrix M when V and W are the same, and the basis $\mathcal{C}$ is the same as $\mathcal{B}$. The matrix M in this case called **Matrix for T relative to $\mathcal{B}$**, or simply $\mathcal{B}$-matrix for T.

The $\mathcal{B}$-matrix for T satisfies

$$[T(x)]_{\mathcal{B}} = [T]_{\mathcal{B}}[x]_{\mathcal{B}} \quad \text{for all } x \text{ in } V.$$

So if $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$, then

$$[T]_{\mathcal{B}} = [[T(b_1)]_{\mathcal{B}} \quad [T(b_2)]_{\mathcal{B}} \quad \ldots [T(b_n)]_{\mathcal{B}}]$$

Example 1.67

The linear transformation $T : \mathbb{P}_2 \to \mathbb{P}_2$ *defined by*

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

is a linear transformation.

1. *Find the* $\mathcal{B}$*-matrix for* T*, when* $\mathcal{B}$ *is the basis* $\{1, t, t^2\}$.
2. *Verify that* $[T(p)]_{\mathcal{B}} = [T]_{\mathcal{B}}[p]_{\mathcal{B}}$ *for each* $p \in \mathbb{P}_2$.

Solution. (1) We have that

$$[T]_{\mathcal{B}} = [[T(1)]_{\mathcal{B}} \;\; [T(t)]_{\mathcal{B}} \;\; [T(t^2)]_{\mathcal{B}}].$$

Note that $T(1) = 0 \qquad T(t) = 1 \qquad T(t^2) = 2t$ Therefore,

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad [T(t^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

So

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

(2) Any polynomial $p(t) \in \mathbb{P}_2$ is of the form $p(t) = a_0 + a_1 t + a_2 t^2$ for some scalars a_0, a_1 and a_2. Thus,

$$[T(p)]_{\mathcal{B}} = [a_1 + 2a_2 t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

and

$$[T(p)]_{\mathcal{B}} = [T]_{\mathcal{B}}[p]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}.$$

- **Linear transformation on $\mathbb{R}^n$**

Theorem 1.68

(**Diagonal matrix representation**) *Suppose that $A = PDP^{-1}$ where P is an invertible matrix and D is a diagonal matrix. Assume that*

$$P = [v_1|v_2|\dots|v_n].$$

Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$. Let

$$\begin{aligned} T: \quad \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto Ax \end{aligned}$$

Then $D = [T]_{\mathcal{B}}$, i.e.,

$$[T(x)]_{\mathcal{B}} = D[x]_{\mathcal{B}}.$$

Example 1.69

Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $T(x) = Ax$, *where* $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. *Find a basis for* $\mathbb{R}^2$ *with the property that the* $\mathcal{B}$*-matrix for* T *is a diagonal matrix.*

Solution. By the previous Theorem if we find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$, then the columns of P produce the basis $\mathcal{B}$. We can find $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ such that $A = PDP^{-1}$.

So $\mathcal{B} = \{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\}$.

- **Similarity of matrix representations**

Theorem 1.70

Suppose that $A = PCP^{-1}$ *where* P *is an invertible matrix. Assume that*

$$P = [v_1|v_2|\dots|v_n].$$

Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$. *Let*

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & Ax \end{array}$$

Then $C = [T]_{\mathcal{B}}$, *i.e.,*

$$[T(x)]_{\mathcal{B}} = C[x]_{\mathcal{B}}.$$

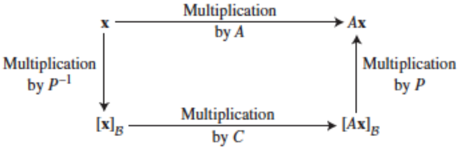


FIGURE 5 Similarity of two matrix representations: $A = PCP^{-1}$.

Week 12, Lecture 1, Nov. 13, Inner Product, length and orthogonality

Definition 2.1

A complex eigenvalue for a matrix A is a complex scalar λ such that there is a non-zero vector x in $\mathbb{C}^n$ s.t $Ax = \lambda x$. Moreover, x is called a complex eigenvector corresponding to λ.

Remark. The complex eigenvalues are the roots of $det(A - \lambda I)$. Also, the set of all eigenvectors corresponding to λ are the non-zero vectors $x \in \mathbb{C}^n$ such that

$$(A - \lambda I)x = 0.$$

Example 2.2

If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, *find eigenvalues.*

Solution. To find the eigenvalues, we should find the roots of $det(A - \lambda I)$.

$$det(A - \lambda I) = det \begin{bmatrix} 0 - \lambda & -1 \\ 1 & 0 - \lambda \end{bmatrix} = \lambda^2 + 1$$

The roots of $\lambda^2 + 1$ are i and $-i$. So eigenvalues are i and $-i$. And also we have

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}$$

So $\begin{bmatrix} 1 \\ i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ are eigenvectors corresponding to $-i$ and i respectively.

• The inner product

Let

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n,$$

then

$$u^T = [u_1 u_2 \ldots u_n].$$

The **inner product(or dot product)** of two vectors $u, v \in \mathbb{R}^n$ is the number $u^T v$, and often it is written as $u.v$.

Example 2.3

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

Solution.

$$u.v = u^T v = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} =$$

$$2 \times 3 + (-5) \times 2 + (-1) \times (-3) = -1$$

$$v.u = v^T u = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} =$$

$$3 \times 2 + 2 \times (-5) + (-3) \times (-1) = -1$$

Theorem 2.4

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

a. $u.v = v.u$

b. $(u+v).w = u.w + v.w$

c. $(cu).v = c(u.v) = u.(cv)$

d. $u.u \geq 0$ *and* $u.u = 0$ *if and only if* $u = 0$.

Combining (b) and (c) we have

$$(c_1u_1 + \ldots + c_pu_p).w = c_1(u_1.w) + \ldots + c_p(u_p.w).$$

- **The length of a vector:**

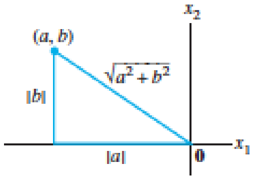


FIGURE 1
Interpretation of $\|\mathbf{v}\|$ as length.

Definition 2.5

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

$$\|v\| = \sqrt{v.v} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$$

and $\|v\|^2 = v.v.$

- For any scalar c, the length of cv is $|c|$ times the length of v, that is

$$\|cv\| = |c|\|v\|.$$

Definition 2.6

A vector v with $\|v\| = 1$ is called a **unit vector**.

Normalizing a vector: Let u be a vector, then $(1/\|u\|)u$ is a unit vector. The process of dividing a vector to its length is called **normalizing**. Moreover, u and $(1/\|u\|)u$ have the same direction.

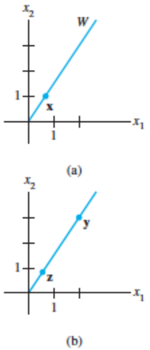


FIGURE 2
Normalizing a vector to produce a unit vector.

Example 2.7

Let $v = (1, -2, 2, 4)$. Find a unit vector u in the same direction as v.

Solution. First compute the length of v:

$$\|v\| = \sqrt{v.v} = \sqrt{1^2 + (-2)^2 + 2^2 + 4^2} = \sqrt{25} = 5$$

Then we multiply v by $1/\|v\|$ to obtain u.

$$u = (1/\|v\|)v = 1/5v = 1/5 \begin{bmatrix} 1 \\ -2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/5 \\ -2/5 \\ 2/5 \\ 4/5 \end{bmatrix}.$$

To check $\|u\| = 1$,

$$\|u\| = \sqrt{u.u} = \sqrt{(1/5)^2 + (-2/5)^2 + (2/5)^2 + (4/5)^2} =$$

$$\sqrt{1/25 + 4/25 + 4/25 + 16/25} = \sqrt{25/25} = 1$$

Example 2.8

Let W be a subspace of $\mathbb{R}^2$ spanned by $x = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}$. Find a unit vector z that is a basis for W.

Solution. Note that $W = \{c \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} : c \in \mathbb{R}\}$. We have that $1/||x|| \in \mathbb{R}$ so $(1/\|x\|)x$ is a vector in W, and spanning it. It is enough to compute $(1/\|x\|)x$.

$$\|x\| = \sqrt{x.x} = \sqrt{(3/2)^2 + 1^2} = \sqrt{9/4 + 1} = \sqrt{13/4} = \sqrt{13}/2$$

so $(1/\|x\|)x = \frac{1}{\sqrt{13}/2} \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} = 2/\sqrt{13} \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6/2\sqrt{13} \\ 2/\sqrt{13} \end{bmatrix}$.

Week 12, Lecture 2, Nov. 15, Distance in $\mathbb{R}^n$ and Orthogonality

- **Distance in $\mathbb{R}^n$**

Definition 2.9

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

Example 2.10

Compute the distance between the vectors $u = (7, 1)$ *and* $v = (3, 2)$.

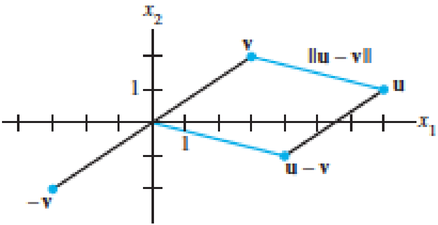


FIGURE 4 The distance between $\mathbf{u}$ and $\mathbf{v}$ is the length of $\mathbf{u} - \mathbf{v}$.

Solution.

$$u - v = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$||u - v|| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

Example 2.11

If $u = (u_1, u_2, u_3)$ *and* $v = (v_1, v_2, v_3)$*, then*

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

$$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$$

Definition 2.12

Two vectors u and v in $\mathbb{R}^n$ are **orthogonal** *to each other if $u.v = 0$.*

Theorem 2.13

(The pythagorean Theorem) Two vectors u and v are orthogonal if and only if

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Orthogonal Complement

Definition 2.14

- *If a vector z is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then z is said to be* **orthogonal to** W**.**
- *The set of all vectors z that are orthogonal to W is said* **orthogonal complement of** W *and is denoted by* $W^\perp$ *(W perp)*

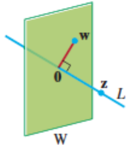


FIGURE 7
A plane and line through **0** as orthogonal complements.

Theorem 2.15

1. *A vector x is in $W^{\perp}$ if and only if x is orthogonal to every vector in a set that spans W.*
2. *$W^{\perp}$ is a subspace of $\mathbb{R}^n$.*

Definition 2.16

Let $A = [A_1|A_2|\ldots|A_n]$ *be an* $m \times n$ *matrix. Also* A *has* m *rows, denote them by* $A_1', \ldots, A_m'$.

$$Col\ A = span\{A_1, \cdots, A_n\} \qquad Row\ A = span\{A_1', \ldots, A_m'\}.$$

Theorem 2.17

Let A *be an* $m \times n$ *matrix.*

1. $(Row\ A)^{\perp} = Nul\ A$, *that is the orthogonal complement of the row space of* A *is the null space of* A.
2. $(Col\ A)^{\perp} = Nul\ A^T$, *that is the orthogonal complement of the column space of* A *is the null space of* A^T.

Angle between two vectors

- Let u and v be in $\mathbb{R}^2$ or $\mathbb{R}^3$, then

1. $$u.v = \|u\|\|v\|cos\theta,$$

   where θ is the angle between the two line segments from the origin to the points identified with u and v.

2. We also have

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|cos\theta$$

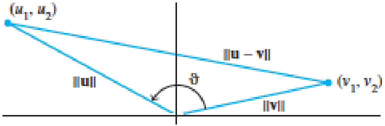


FIGURE 9 The angle between two vectors.

Example 2.18

Find the angle between $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *and* $v = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Solution. We have

$$u.v = \|u\|\|v\|cos\theta.$$

Note that $\|u\| = \sqrt{1^2 + 1^2} = \sqrt{2}$ and $\|v\| = \sqrt{(-1)^2 + 0^2} = 1$ and $u.v = u^T.v = -1$. So $-1 = \sqrt{2}.cos\theta$. Therefore, $\theta = \frac{3\pi}{4}$.

Orthogonal Sets:

Definition 2.19

A set of vectors $\{u_1, u_2, \ldots, u_p\}$ *in* $\mathbb{R}^n$ *is said to be* **orthogonal set** *if each pair of distinct vectors from the set are orthogonal, that is,* $u_i.u_j = 0$ *if* $i \neq j$.

Example 2.20

Show that $\{u_1, u_2, u_3\}$ *is an orthogonal set where*

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \text{ and } u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}.$$

Solution. We must show that $u_1.u_2 = 0$, $u_1.u_3 = 0$, and $u_2.u_3 = 0$.

$$u_1.u_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$u_1.u_3 = 3(-1/2) + 1(-2) + 1(7/2) = 0$$

$$u_2.u_3 = -1(-1/2) + 2(-2) + 1(7/2) = 0.$$

Theorem 2.21

If $S = \{u_1, u_2, \ldots, u_p\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

Definition 2.22

An **orthogonal basis** *for a subspace W of $\mathbb{R}^n$ is a basis for W that is also orthogonal set.*

Theorem 2.23

Let $\{u_1, \ldots, u_p\}$ *be an orthogonal basis for a subspace* W *of* $\mathbb{R}^n$*. For each* $y \in W$*, the weights in the linear combination*

$$y = c_1u_1 + \cdots + c_pu_p$$

are given by

$$c_j = \frac{y.u_j}{u_j.u_j} \quad (j = 1, 2, \ldots, p)$$

Example 2.24

The set $S = \{u_1, u_2, u_3\}$*, where*

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \text{ and } u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

is an orthogonal basis for $\mathbb{R}^3$*. Express the vector* $y = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ *as a linear combination of the vectors in* S*.*

Solution. If we write $y = c_1u_1 + c_2u_2 + c_3u_3$, then

$$c_1 = \frac{y.u_1}{u_1.u_1} = \frac{11}{11} = 1 \quad c_2 = \frac{y.u_2}{u_2.u_2} = \frac{-12}{6} = -2$$

$c_3 = \frac{y.u_3}{u_3.u_3} = \frac{-33}{33/2} = -2$. Therefore, $y = 1u_1 - 2u_2 - 2u_3$.

Week 12, Lecture 3, Nov. 17, Orthogonal projection and orthonormal sets

Assume that u is in $\mathbb{R}^n$. then $L = span\{u\} = \{cu : c \in \mathbb{R}\}$ is a line.

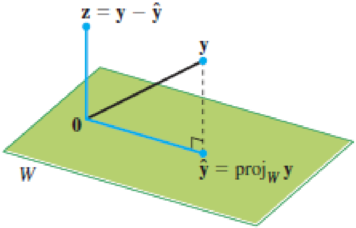


FIGURE 2
Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

We want to write a vector y as a sum of a vector in $L = span\{u\}$ and a vector orthogonal to u. Then $y = \hat{y} + (y - \hat{y})$, where

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

$\hat{y} = \mathbf{proj}_L y$ is called **orthogonal projection** of y onto L. Also $y - \hat{y}$ is called the **complement of** y **orthogonal to** u**.**

Example 2.25

Let $y = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$, *and* $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. *Find the orthogonal projection of* y *onto* u. *Then write* y *as the sum of two orthogonal vectors, one in span*$\{u\}$ *and one orthogonal to* u.

Solution.

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

$$u.u = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

$$\Rightarrow \hat{y} = \frac{y.u}{u.u}u = (40/20)u = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the complement of y orthogonal to u.

$$y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Visualizing Theorem 2.23

- It is easy to visualize the case in which $w = \mathbb{R}^2 = span\{u_1, u_2\}$ with u_1 and u_2 orthogonal. Any $y \in \mathbb{R}^2$ can be written in the form

$$y = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$

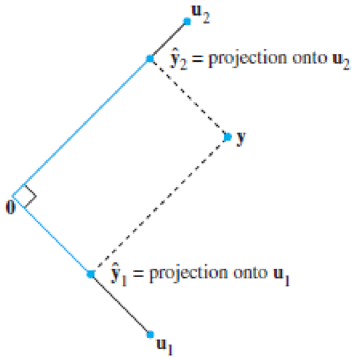


FIGURE 4 A vector decomposed into the sum of two projections.

Orthonormal sets

Definition 2.26

A set $\{u_1, \ldots, u_p\}$ is an **orthonormal set** *if it is an orthogonal of unit vectors.*

Example 2.27

Show that $\{v_1, v_2, v_3\}$ is an orthonormal basis of $\mathbb{R}^3$. Where

$$v_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, v_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \text{and } \; v_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Solution. Compute

$$v_1.v_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$v_1.v_3 = -3/\sqrt{726} + (-4)/\sqrt{726} + 7/\sqrt{726} = 0$$

$$v_2.v_3 = 1/\sqrt{396} + (-8)/\sqrt{396} + 7/\sqrt{396} = 0$$

so $\{v_1, v_2, v_3\}$ is an orthogonal set.
Now we show that v_1, v_2, v_3 are unit vector.

$$\|u_1\| = \sqrt{v_1.v_1} = \sqrt{9/11 + 1/11 + 1/11} = 1$$

$$\|u_2\| = \sqrt{v_2.v_2} = \sqrt{1/6 + 4/6 + 1/6} = 1$$

$$\|u_3\| = \sqrt{v_3.v_3} = \sqrt{1/66 + 16/66 + 49/66} = 1$$

So $\{v_1, v_2, v_3\}$ is orthonormal basis for $\mathbb{R}^3$.

Theorem 2.28

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

Theorem 2.29

Let U be an $m \times n$ matrix with orthonormal columns and let x and y be in $\mathbb{R}^n$. Then

1. $\|Ux\| = \|x\|$.
2. $(Ux).(Uy) = x.y$.
3. $(Ux).(Uy) = 0 \quad$ *if and only if* $\quad x.y = 0$

Example 2.30

Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ *and* $x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. *Notice that* U *has orthonormal columns and*

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

verify that $||Ux|| = ||x||$.

Solution.

$$Ux = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}.$$

$$\|Ux\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|Ux\| = \sqrt{2+9} = \sqrt{11}$$

Orthogonal matrix

Definition 2.31

An **orthonormal matrix** *is a square invertible matrix U such that*

$$U^{-1} = U^T.$$

Example 2.32

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 orthonormal matrix.

Week 14, Lecture 1, Nov. 27, Orthogonal Projection

Example 4.1

Let $\{u_1, \ldots, u_5\}$ *be an orthogonal basis for* $\mathbb{R}^5$ *and let*

$$y = c_1u_1 + \ldots + c_5u_5.$$

Consider the subspace $W = span\{u_1, u_2\}$, *and write* y *as the sum of a vector* z_1 *in* W *and a vector* z_2 *in* $W^\perp$.

Solution. Write

$$y = \underbrace{c_1u_1 + c_2u_2}_{z_1} + \underbrace{c_3u_3 + c_4u_4 + c_5u_5}_{z_2}$$

where $z_1 = c_1u_1 + c_2u_2$ is in $span\{u_1, u_2\} = W$ and $z_2 = c_3u_3 + c_4u_4 + c_5u_5$ is in $span\{u_3, u_4, u_5\}$.
To show that z_2 is in $W^\perp$ it is enough to show that $z_2.u_i = 0$, for $i = 1$ *and* $i = 2$.

$$z_2.u_1 = (c_3u_3 + c_4u_4 + c_5u_5).u_1$$
$$= c_3u_3.u_1 + c_4u_4.u_1 + c_5u_5.u_1 = 0$$

because $\{u_1, \dots, u_5\}$ is an orthogonal set.
Similarly $z_2.u_2 = 0$. Therefore $z_2 \in W^\perp$.

Theorem 4.2

(The Orthogonal Decomposition Theorem) Let W be a subspace of $\mathbb{R}^n$. Then each y in $\mathbb{R}^n$ can be written uniquely in the form

$$y = \widehat{y} + z \qquad (1)$$

where $\widehat{y}$ is in W and z in $W^\perp$. In fact if $\{u_1, \ldots, u_p\}$ is an orthogonal basis of W, then

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \ldots + \frac{y.u_p}{u_p.u_p}u_p$$

and $z = y - \widehat{y}$.

Definition 4.3

The vector $\widehat{y}$ in (1) is called **the orthogonal projection of y onto W**, *and it sometimes denoted by* $\mathbf{proj}_W y$.

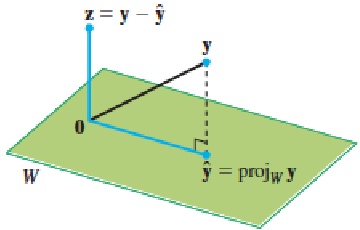


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

Example 4.4

Let $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, *and* $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. *Observe that* $\{u_1, u_2\}$ *is an orthogonal basis for* $W = span\{u_1, u_2\}$. *Write* y *as the sum of a vector in* W *and a vector orthogonal to* W.

Solution. The orthogonal projection of y onto W is

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$

$$= 9/30 \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + 3/6 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

Also

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

By previous theorem $y - \widehat{y}$ is in $W^{\perp}$. And

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

• A Geometric Interpretation of the Orthogonal Projection

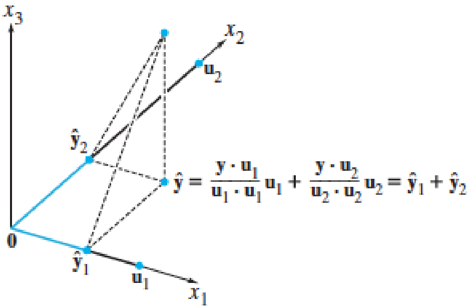


FIGURE 3 The orthogonal projection of **y** is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

- **Properties of Orthogonal Projections**

Proposition 4.5

If y *is in* $W = span\{u_1, \ldots, u_p\}$, *then* $\mathbf{proj}_W y = y$.

Theorem 4.6

(The Best Approximation Theorem) Let W *be a subspace of* $\mathbb{R}^n$, *let* y *be any vector in* $\mathbb{R}^n$, *and let* $\widehat{y}$ *be the orthogonal projection of* y *onto* W. *Then* $\widehat{y}$ *is the closest point in* W *to* y, *in the sense that*

$$\|y - \widehat{y}\| \leq \|y - v\|$$

for all v *in* W *distinct from* $\widehat{y}$.

Definition 4.7

The vector $\widehat{y}$ *is called* **the best approximation to** y **by elements of** W.

Definition 4.8

The vector $\widehat{y}$ *is called* **the best approximation to** y **by elements of** W**.**

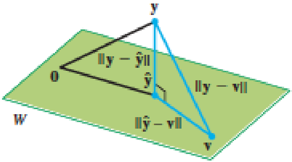


FIGURE 4 The orthogonal projection of **y** onto W is the closest point in W to **y**.

Example 4.9

If $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ *and* $W =$ *span*$\{u_1, u_2\}$. *Find the closest point in* W *to* y.

Solution. By the theorem the point is

$$\widehat{y} = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

(we already computed $\widehat{y}$ in one of the examples.)

Example 4.10

The **distance from a point** $y \in \mathbb{R}^n$ **to a subspace** W *is defined as the distance from* y *to the nearest point in* W*. Find the distance from* y *to* $W = span\{u_1, u_2\}$*, where*

$$y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Solution. By the best approximation theorem, the distance from y to W is $\|y - \widehat{y}\|$, where $\widehat{y} = \mathbf{proj}_W y$. Since $\{u_1, u_2\}$ is an orthogonal basis for W,

$$\widehat{y} = 15/30u_1 + (-21/6)u_2 = 1/2\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - 7/2\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$y - \widehat{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|y - \widehat{y}\| = \sqrt{3^2 + 6^2} = \sqrt{45}.$$

Therefore, the distance from y to W is $\sqrt{45} = 3\sqrt{5}$.

Theorem 4.11

If $\{u_1, \ldots, u_5\}$ *is an orthonormal basis for a subspace* W *of* $\mathbb{R}^n$*, then*

$$\mathbf{proj}_W y = (y.u_1)u_1 + (y.u_2)u_2 + \ldots + (y.u_p)u_p$$

if $U = [u_1 u_2 \ldots u_p]$*, then*

$$\mathbf{proj}_W y = UU^T y \quad \textit{for all } y \textit{ in } \mathbb{R}^n.$$

Week 14, Lecture 2, Nov. 29, The Gram-Schmidt process

Reminder from last lecture

Orthogonal Projection

Let $W = \{u_1, u_2, \ldots, u_p\}$ be an orthogonal subspace of $\mathbb{R}^n$. Let $y \in \mathbb{R}^n$. Then the orthogonal projection of y on W is

$$\widehat{y} = \mathbf{proj}_W y = \frac{u_1.y}{u_1.u_1}u_1 + \frac{u_2.y}{u_2.u_2}u_2 + \ldots + \frac{u_p.y}{u_p.u_p}u_p.$$

Also we can write

$$y = \widehat{y} + z,$$

where $\widehat{y} \in W$ and $z = y - \widehat{y} \in W^{\perp}$.

Example 4.12

Let $W = span\{x_1, x_2\}$, *where* $x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ *and* $x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

Construct an orthogonal basis $\{v_1, v_2\}$ *for* W.

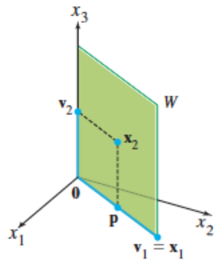


FIGURE 1
Construction of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

Solution. Let $v_1 = x_1$. Let p be orthogonal projection of x_2 onto x_1 , i.e.,

$$p = \frac{x_1.x_2}{x_1.x_1} x_1.$$

We have that

$$v_2 = x_2 - \frac{x_1.x_2}{x_1.x_1} x_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - 15/45 \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Then $\{v_1, v_2\}$ is an orthogonal set of non-zero vectors in W. Since $dim\ W = 2$, then set $\{v_1, v_2\}$ is a basis for W.

Example 4.13

Let $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, *and* $x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. *Then* $\{x_1, x_2, x_3\}$ *is clearly linearly independent and thus is a basis for* W. *Construct an orthogonal basis for* W.

Solution.

Step1. Let $v_1 = x_1$ and $W_1 = span\{x_1\} = span\{v_1\}$.
Step2. $v_2 = x_2 - \mathbf{proj}_{W_1} x_2$

$$= x_2 - \frac{x_2.v_1}{v_1.v_1} v_1$$

$$= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 3/4 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

Let $W_2 = span\{v_1, v_2\}$. Then $\{v_1, v_2\}$ is an orthogonal basis for $W_2 = span\{v_1, v_2\} = span\{x_1, x_2\}$.

Step3. $v_3 = x_3 - \mathbf{proj}_{W_2} x_3$

$$\mathbf{proj}_{w_2} x_3 = \frac{x_3.v_1}{v_1.v_1} v_1 + \frac{x_3.v_2}{v_2.v_2} v_2$$

$$= 1/2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 2/3 \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Then

$$v_3 = x_3 - \mathbf{proj}_{w_2} x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

So $\{v_1, v_2, v_3\}$ is an orthogonal basis for W.

Theorem 4.14

(The Gram-Schmidt process) Given a basis $\{x_1, \dots, x_p\}$ *for non-zero subspace* W *of* $\mathbb{R}^n$*, define*

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1} v_1$$

$$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1} v_1 - \frac{x_3.v_2}{v_2.v_2} v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p.v_1}{v_1.v_1} v_1 - \frac{x_p.v_2}{v_2.v_2} v_2 - \dots - \frac{x_p.v_{p-1}}{v_{p-1}.v_{p-1}} v_{p-1}$$

Then $\{v_1, \dots, v_p\}$ *is an orthogonal basis for* W*. In addition* $span\{v_1, \dots, v_k\} = span\{x_1, \dots, x_k\}$ *for* $1 \le k \le p$*.*

Theorem 4.15

(The QR factorization) 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 from an orthogonal basis for $Col\ A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Example 4.16

Let $W = span\{v_1, v_2, v_3\}$ *be a subspace of* $\mathbb{R}^4$*, where*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 3 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 4 \\ -4 \\ 5 \end{bmatrix}.$$

Find an orthogonal basis for W*.*

Week 14, Lecture 3, Dec. 1, Least squares problems

Sometimes $Ax = b$ does not have a solution. However, we can find the vector $\widehat{x}$ such that $A\widehat{x}$ is the best approximation to b.

Definition 4.17

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

$$\|b - A\widehat{x}\| \leq \|b - Ax\|$$

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

- Goal: Finding the set of least-squares solution of $Ax = b$.

Theorem 4.18

(Best Approximation Theorem): Let W be a subspace of $\mathbb{R}^n$, let y be any vector in $\mathbb{R}^n$, and let $\widehat{y}$ be the orthogonal projection of y onto W. Then $\widehat{y}$ is the closest point in W to y, in the sense that

$$\|y - \widehat{y}\| < \|y - v\|$$

for all v in W distinct from $\widehat{y}$.

• Solution of the general least-squares problem:
We apply the theorem above to find the set of least-squares solution of $Ax = b$.
Consider $Col\ A$. Let

$$\widehat{b} = \mathbf{proj}_{Col\ A} b$$

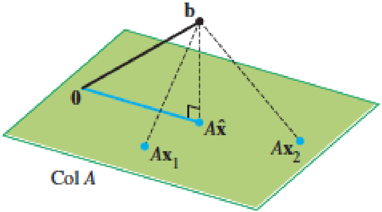


FIGURE 1 The vector **b** is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other **x**.

Since $\widehat{b} \in Col\ A$, there is $\widehat{x}$ such that

$$A\widehat{x} = \widehat{b} \qquad (1)$$

Note that $\widehat{b}$ is the closest point in $Col\ A$ to b. Therefore, a vector $\widehat{x}$ is a least-squares solution if and only if $\widehat{x}$ satisfies $A\widehat{x} = \widehat{b}$. We have by the Orthogonal Decomposition Theorem that $b - \widehat{b}$ is orthogonal to $Col\ A$. So $b - \widehat{b}$ is orthogonal to each column A_j of A. Therefore,

$$0 = A_j.(b - \widehat{b}) = A_j.(b - A\widehat{x})$$

$$= A_j^T(b - A\widehat{x}) = 0$$

$$\Rightarrow A^T(b - A\widehat{x}) = 0$$

$$\Rightarrow A^Tb = A^TA\widehat{x}.$$

So the set of least squares solutions of $Ax = b$ is the same as all $\widehat{x}$ such that $A^Tb = A^TA\widehat{x}$. So we have the following theorem.

Theorem 4.19

The set of least-squares solutions of $Ax = b$ coincides with the nonempty set of solution of the **normal equations** $A^T Ax = A^T b$.

Theorem 4.20

Let A be an $m \times n$ matrix. The following statements are logically equivalent:

(a) *The equation $Ax = b$ has a unique least-squares solution for each b in $\mathbb{R}^m$.*

(b) *The columns of A are linearly independent.*

(c) *The matrix $A^T A$ is invertible.*

When these statements are true, the least-squares solution $\widehat{x}$ is given by

$$\widehat{x} = (A^T A)^{-1} A^T b.$$

Example 4.21

Find a least-squares solution of the inconsistent system $Ax = b$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \textit{and} \quad b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

Solution. Example 1 page 364 of the textbook.

Example 4.22

Find a least-squares solution of $Ax = b$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad and \quad b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}.$$

Solution. Example 2 page 364 of the textbook.

Week 15, Lecture 1, Dec. 4, Inner product space

Definition 5.1

An **inner product** *on a vector space* V *is a function*

$$\langle ., . \rangle : V \times V \longrightarrow \mathbb{R}$$

satisfying the following axioms:

1. $\langle u, v \rangle = \langle v, u \rangle$

2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$

3. $\langle cu, v \rangle = c\langle u, v \rangle$

4. $\langle u, u \rangle \geq 0$ *and* $\langle u, u \rangle = 0$ *if and only if* $u = 0$.

A vector space with an inner product is called an **inner product space**.

Example 5.2

Show that $\mathbb{R}^2$ *with the following function*

$$\left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle = 4u_1v_1 + 5u_2v_2$$

is an inner product space.

Solution. We know that $\mathbb{R}^2$ is a vector space, so we only need to show that the function is an inner product, i.e., checking that the axioms are satisfied.

(1) $\left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle$

(2) Let $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ be another element in $\mathbb{R}^2$. Then

$$\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle = \langle \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle =$$

$$4(u_1+v_1)w_1+5(u_2+v_2)w_2 = 4u_1w_1+4v_1w_1+5u_2w_2+5v_2w_2$$

$$= (4u_1w_1 + 5u_2w_2) + (4v_1w_1 + 5v_2w_2)$$

$$= \langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle + \langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \rangle$$

(3) $\langle c \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle = \langle \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle$

$$= 4cu_1v_1 + 5cu_2v_2 = c(4u_1v_1 + 5u_2v_2) = c\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rangle.$$

(4) $\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \rangle = 4u_1^2 + 5u_2^2 \geq 0$

and also note that if $\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \rangle = 4u_1^2 + 5u_2^2 = 0$ then $u_1 = 0$ and $u_2 = 0$. Therefore, $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Example 5.3

Let $t_0, \ldots, t_n$ be distinct real numbers. For p and q in $\mathbb{P}_n$, define

$$\langle p, q \rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \ldots + p(t_n)q(t_n).$$

Solution. Axioms 1-3 are readily checked. For axiom 4,

$$\langle p, p \rangle = [p(t_0)]^2 + \ldots + [p(t_n)]^2 = 0.$$

So if $[p(t_0)]^2 + \ldots + [p(t_n)]^2 = 0$ we must have $p(t_0) = 0, \ldots, p(t_n) = 0$. It means $t_0, \ldots, t_n$ are roots for p. Therefore, p has $n+1$ roots, which is impossible if $p \neq 0$ since any non-zero polynomial of degree n has at most n roots.

Length, Distance, and Orthogonality

Definition 5.4

Let V be an inner product space and u and $v \in V$. Then we define

1. *the* **length or norm** *of a vector to be the scalar*

$$\|v\| = \sqrt{\langle v, v \rangle}$$

2. *A* **unit** *vector is one whose length is 1.*
3. *The* **distance** *between u and v is $\|u - v\| = \sqrt{\langle u - v, u - v \rangle}$.*
4. *Two vectors u and v are said to be* **orthogonal** *if and only if $\langle u, v \rangle = 0$.*

Example 5.5

Let $\mathbb{P}_2$ have the inner product

$$\langle p, q\rangle = p(0)q(0) + p(1/2)q(1/2) + p(1)q(1).$$

Compute the length of the following vectors $p(t) = 12t^2$ and $q(t) = 2t - 1$.

Solution. Note that $\|p\| = \sqrt{\langle p, p\rangle}$. We have

$$\langle p, p\rangle = [p(0)]^2 + [p(1/2)]^2 + [p(1)]^2 = 0 + 3^2 + 12^2 = 153.$$

Therefore, $\|p\| = \sqrt{153}$. Also, $\|q\| = \sqrt{2}$ (check it).

The Gram-Schmidt Process:

Theorem 5.6

(The Gram-Schmidt process) Given a basis $\{x_1, \ldots, x_p\}$ *for non-zero subspace* W *of* $\mathbb{R}^n$*, define*

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1}v_1$$

$$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1}v_1 - \frac{x_3.v_2}{v_2.v_2}v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p.v_1}{v_1.v_1}v_1 - \frac{x_p.v_2}{v_2.v_2}v_2 - \ldots - \frac{x_p.v_{p-1}}{v_{p-1}.v_{p-1}}v_{p-1}$$

Then $\{v_1, \ldots, v_p\}$ *is an orthogonal basis for* W*. In addition* $span\{v_1, \ldots, v_k\} = span\{x_1, \ldots, x_k\}$ *for* $1 \leq k \leq p$.

The Gram-Schmidt process for an inner product space

Theorem 5.7

(The Gram-Schmidt process for an inner product space) Given a basis $\{x_1, \ldots, x_p\}$ for non-zero subspace W of an inner product space V, define

$$v_1 = x_1$$

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

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

$$\vdots$$

$$v_p = x_p - \frac{\langle x_p, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x_p, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \ldots - \frac{\langle x_p, v_{p-1} \rangle}{\langle v_{p-1}, v_{p-1} \rangle} v_{p-1}$$

Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W. In addition $span\{v_1, \ldots, v_k\} = span\{x_1, \ldots, x_k\}$ for $1 \leq k \leq p$.

Example 5.8

Define the following inner product for $\mathbb{P}_4$*,*

$$\langle p, q\rangle = p(-2)q(-2)+p(-1)q(-1)+p(0)q(0)+p(1)q(1)+p(2)q(2).$$

Let $\mathbb{P}_2$ *be the subspace of* $\mathbb{P}_4$ *with the basis* $\{p_1, p_2, p_3\}$*, where* $p_1 = 1, p_2 = t, p_3 = t^2$*. Produce an orthogonal basis for* $\mathbb{P}_2$ *by applying the Gram-Schmidt Process.*

Solution.

$f_1 = p_1 = 1$

$f_2 = p_2 - \frac{\langle p_2, f_1\rangle}{\langle f_1, f_1\rangle} f_1$

$f_3 = p_3 - \frac{\langle p_3, f_1\rangle}{\langle f_1, f_1\rangle} f_1 - \frac{\langle p_3, f_2\rangle}{\langle f_2, f_2\rangle} f_2$

$$\langle t, 1\rangle = (-2)\times 1 + (-1)\times 1 + 0\times 1 + 1\times 1 + 2\times 1 = 0.$$

$$\langle f_1, f_1\rangle = \langle 1, 1\rangle = 1\times 1 + 1\times 1 + 1\times 1 + 1\times 1 + 1\times 1 = 5$$

Therefore, $f_2 = t - \frac{0}{5} = t$.

$$\langle p_3, f_1 \rangle = \langle t^2, 1 \rangle = (-2)^2 \times 1 + (-1)^2 \times 1 +$$

$$0^2 \times 1 + 1^2 \times 1 + 2^2 \times 1 = 10.$$

$$\langle p_3, f_2 \rangle = \langle t^2, t \rangle = (-2)^2 \times -2 + (-1)^2 \times (-1) +$$

$$0^2 \times 0 + 1^2 \times 1 + 2^2 \times 2 = 0.$$

$$\langle f_2, f_2 \rangle = \langle t, t \rangle = (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 = 10.$$

Therefore, $f_3 = t^2 - \frac{10}{5}1 - \frac{0}{10}t = t^2 - 2$. Therefore,

$$\{1, t, t^2 - 2\}$$

is an orthogonal basis for $\mathbb{P}_2$ (check orthogonality).