SF 1684 Algebra and Geometry
Lecture 15

Patrick Meisner

KTH Royal Institute of Technology

Topics for Today

1. Least Squares Process
2. Change of Basis
3. Gram-Schmidt Process

Approximate Solutions to Matrix Equations

For a given $m \times n$ matrix A, and a vector $\vec{b}$ in $\mathbb{R}^m$ we are interested in finding solutions to $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{b}$.

Approximate Solutions to Matrix Equations

For a given $m \times n$ matrix A, and a vector $\vec{b}$ in $\mathbb{R}^m$ we are interested in finding solutions to $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{b}$. As we have seen, there is not always a solution. Hence, we sometimes have to settle for a *best approximate solution*.

Approximate Solutions to Matrix Equations

For a given $m \times n$ matrix A, and a vector $\vec{b}$ in $\mathbb{R}^m$ we are interested in finding solutions to $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{b}$. As we have seen, there is not always a solution. Hence, we sometimes have to settle for a *best approximate solution*.

Definition

If A is an $m \times n$ matrix and $\vec{b}$ is a vector in $\mathbb{R}^m$, then a vector $\vec{x}$ in $\mathbb{R}^n$ is called a **best approximate solution** or a **least squares solution** to $A\vec{x} = \vec{b}$ if

$$\underbrace{\left\| \vec{b} - A\vec{x} \right\|}_{\text{distance between } A\vec{x} \text{ \& } \vec{b}} \leq \underbrace{\left\| \vec{b} - A\vec{y} \right\|}_{\text{distance between } A\vec{y} \text{ \& } \vec{b}}$$

for all $\vec{y}$ in $\mathbb{R}^n$

Approximate Solutions to Matrix Equations

For a given $m \times n$ matrix A, and a vector $\vec{b}$ in $\mathbb{R}^m$ we are interested in finding solutions to $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{b}$. As we have seen, there is not always a solution. Hence, we sometimes have to settle for a *best approximate solution*.

Definition

If A is an $m \times n$ matrix and $\vec{b}$ is a vector in $\mathbb{R}^m$, then a vector $\vec{x}$ in $\mathbb{R}^n$ is called a **best approximate solution** or a **least squares solution** to $A\vec{x} = \vec{b}$ if

$$\left\|\vec{b} - A\vec{x}\right\| \leq \left\|\vec{b} - A\vec{y}\right\|$$

for all $\vec{y}$ in $\mathbb{R}^n$. The vector $\vec{b} - A\vec{x}$ is called the **least squares error vector**

For a given $m \times n$ matrix A, and a vector $\vec{b}$ in $\mathbb{R}^m$ we are interested in finding solutions to $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{b}$. As we have seen, there is not always a solution. Hence, we sometimes have to settle for a *best approximate solution*.

Definition

If A is an $m \times n$ matrix and $\vec{b}$ is a vector in $\mathbb{R}^m$, then a vector $\vec{x}$ in $\mathbb{R}^n$ is called a **best approximate solution** or a **least squares solution** to $A\vec{x} = \vec{b}$ if

$$\left\|\vec{b} - A\vec{x}\right\| \leq \left\|\vec{b} - A\vec{y}\right\|$$

for all $\vec{y}$ in $\mathbb{R}^n$. The vector $\vec{b} - A\vec{x}$ is called the **least squares error vector**, and the scalar $\left\|\vec{b} - A\vec{x}\right\|$ is called the **least squares error**.

Remark: the least square error is 0 iff there is a solution to $Ax = b$.

Why "least squares"?

Note that if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

Why "least squares"?

Note that if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

then we get

$$\left\| \vec{b} - A\vec{x} \right\| = \sqrt{c_1^2 + c_2^2 + \cdots + c_m^2}$$

Why "least squares"?

Note that if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

then we get

$$\left\| \vec{b} - A\vec{x} \right\| = \sqrt{c_1^2 + c_2^2 + \cdots + c_m^2}$$

And we are wishing to *minimize* this value.

Why "least squares"?

Note that if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

then we get

$$\left\| \vec{b} - A\vec{x} \right\| = \sqrt{c_1^2 + c_2^2 + \cdots + c_m^2}$$

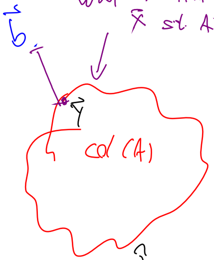
And we are wishing to *minimize* this value.

Note that the set $\{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ is the column space of A.

Why "least squares"?

If $A = [c_1 \dots c_n]$ then $\mathrm{col}(A) = \mathrm{Span}(c_1, \dots, c_n)$

want to find this $\vec{x}$ s.t. $A\vec{x} = \vec{y}$



all the $\vec{y}$ that have a solution to $A\vec{x} = \vec{y}$

Note that if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

then we get

$$\left\| \vec{b} - A\vec{x} \right\| = \sqrt{c_1^2 + c_2^2 + \cdots + c_m^2}$$

And we are wishing to *minimize* this value.

Note that the set $\{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ is the column space of A. Hence, we are really looking for the "minimal distance between the vector $\cancel{\vec{x}}$ $\vec{b}$ and the subspace col(A)"

Distance to a Subspace

Question (Minimal Distance to a Subspace)

Given a subspace W of $\mathbb{R}^n$ and a vector $\vec{b} \in \mathbb{R}^n$, can we find a vector $\vec{w}$ in W that is closest to $\vec{b}$ in the sense that

$$\left\|\vec{w} - \vec{b}\right\| \leq \left\|\vec{v} - \vec{b}\right\|$$

distance between $\vec{w}$ & $\vec{b}$

distance between $\vec{v}$ & $\vec{b}$

for all $\vec{v}$ in W?

Distance to a Subspace

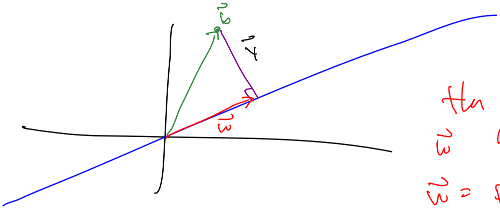
Question (Minimal Distance to a Subspace)

Given a subspace W of $\mathbb{R}^n$ and a vector $\vec{b} \in \mathbb{R}^n$, can we find a vector $\vec{w}$ in W that is closest to $\vec{b}$ in the sense that

$$\left\|\vec{w} - \vec{b}\right\| \leq \left\|\vec{v} - \vec{b}\right\|$$

for all $\vec{v}$ in W? Such a vector $\vec{w}$ is called a **best approximation to $\vec{b}$ from W**.

$n=2$ $W =$ space of $\mathbb{R}^2$ $\rightarrow$ $W =$ line $= \text{span}\{\vec{v}\}$



$\|\vec{y}\|$ is the shortest distance from $\vec{b}$ to line

then $\|\vec{w} - \vec{b}\| = \|\vec{y}\|$ shortest

$\vec{w}$ best app to $\vec{b}$

$\vec{w} = \text{proj}_{\text{Line}}\, \vec{b}$

Distance to a Subspace

Question (Minimal Distance to a Subspace)

Given a subspace W of $\mathbb{R}^n$ and a vector $\vec{b} \in \mathbb{R}^n$, can we find a vector $\vec{w}$ in W that is closest to $\vec{b}$ in the sense that

$$\left\|\vec{w} - \vec{b}\right\| \leq \left\|\vec{v} - \vec{b}\right\|$$

for all $\vec{v}$ in W? Such a vector $\vec{w}$ is called a **best approximation to $\vec{b}$ from W**.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\vec{b}$ is a vector in $\mathbb{R}^n$, then there is a unique best approximation to $\vec{b}$ from W

Distance to a Subspace

Question (Minimal Distance to a Subspace)

Given a subspace W of $\mathbb{R}^n$ and a vector $\vec{b} \in \mathbb{R}^n$, can we find a vector $\vec{w}$ in W that is closest to $\vec{b}$ in the sense that

$$\left\|\vec{w} - \vec{b}\right\| \leq \left\|\vec{v} - \vec{b}\right\|$$

for all $\vec{v}$ in W? Such a vector $\vec{w}$ is called a **best approximation to $\vec{b}$ from W**.

Theorem (Best Approximation Theorem)

If W is a subspace of $\mathbb{R}^n$ and $\vec{b}$ is a vector in $\mathbb{R}^n$, then there is a unique best approximation to $\vec{b}$ from W, namely $\vec{w} = proj_W \vec{b}$.

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$.

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$.

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$. That is:

$$A\vec{x} = \vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$. That is:

$$\vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Hence, to solve our original problem of finding $\vec{x}$, it remains to solve

$$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$$

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$. That is:

$$\vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Hence, to solve our original problem of finding $\vec{x}$, it remains to solve

$$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$$

CAUTION!!!!!

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$. That is:

$$\vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Hence, to solve our original problem of finding $\vec{x}$, it remains to solve

$M(M^T M)^{-1} M^T$

$$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$$

CAUTION!!!!!

It is NOT the case that $\text{proj}_{\text{col}(A)}\vec{b} = A(A^T A)^{-1}A^T\vec{b}$.

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from col(A). That is:

$$\vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Hence, to solve our original problem of finding $\vec{x}$, it remains to solve

$M(M^T M)^{-1} M^T$

$$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$$

CAUTION!!!!!

It is NOT the case that $\text{proj}_{\text{col}(A)}\vec{b} = A(A^TA)^{-1}A^T\vec{b}$. Recall that the matrix we use to describe the projection onto W must be one whose columns form a *basis* for W.

Solving Least Squares

So, given an $m \times n$ matrix A and a vector $\vec{b}$ in $\mathbb{R}^m$, we want to find a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x}$ is "closest" to $\vec{b}$. Setting $\vec{y} = A\vec{x}$, we see that $\vec{y} \in \text{col}(A)$ and so $\vec{y}$ would be the best approximation to $\vec{b}$ from $\text{col}(A)$. That is:

$$\vec{y} = \text{proj}_{\text{col}(A)}\vec{b}$$

Hence, to solve our original problem of finding $\vec{x}$, it remains to solve

$$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$$

CAUTION!!!!!

It is NOT the case that $\text{proj}_{\text{col}(A)}\vec{b} = A(A^TA)^{-1}A^T\vec{b}$. Recall that the matrix we use to describe the projection onto W must be one whose columns form a *basis* for W. While the columns of A do form a spanning set for $\text{col}(A)$, they may not be linearly independent and so would not form a basis!

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ are the exact solutions to the equation

$$A^T A\vec{x} = A^T \vec{b}$$

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ *are the exact solutions to the equation*

$$A^T A\vec{x} = A^T \vec{b}$$

Proof.

We have seen that it is enough to solve $A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$.

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ are the exact solutions to the equation

$$A^T A\vec{x} = A^T \vec{b}$$

Proof.

We have seen that it is enough to solve $A\vec{x} = \mathrm{proj}_{\mathrm{col}(A)}\vec{b}$. Moreover, we know that we can write

$$\vec{b} = \mathrm{proj}_{\mathrm{col}(A)}\vec{b} + \mathrm{proj}_{\mathrm{col}(A)^\perp}\vec{b}$$

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ are the exact solutions to the equation

$$A^T A\vec{x} = A^T\vec{b}$$

Proof.

We have seen that it is enough to solve $A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$. Moreover, we know that we can write

$$\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{col}(A)^\perp}\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{null}(A^T)}\vec{b}$$

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ are the exact solutions to the equation

$$A^T A\vec{x} = A^T \vec{b}$$

Proof.

We have seen that it is enough to solve$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$. Moreover, we know that we can write

$$\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{col}(A)^\perp}\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{null}(A^T)}\vec{b}$$

Hence, multiplying the first equation on both sides by A^T, we find

$$A^T A\vec{x} = A^T \text{proj}_{\text{col}(A)}\vec{b}$$

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ are the exact solutions to the equation

$$A^T A\vec{x} = A^T\vec{b}$$

Proof.

We have seen that it is enough to solve$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$. Moreover, we know that we can write

$$\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{col}(A)^\perp}\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{null}(A^T)}\vec{b}$$

in null(A^T)

Hence, multiplying the first equation on both sides by A^T, we find

T

$$A^T A\vec{x} = A^T\text{proj}_{\text{col}(A)}\vec{b} = A^T(\vec{b} - \text{proj}_{\text{null}(A^T)}\vec{b})$$

$= A^T\vec{b} - A^T(\text{proj}_{\text{null}\,A^T} b)$

$= A^T\vec{b} - 0$

$= A^T\vec{b}$

Least Squares Theorem

Theorem

The least squares solutions of a linear system $A\vec{x} = \vec{b}$ *are the exact solutions to the equation*

$$A^T A\vec{x} = A^T \vec{b}$$

Proof.

We have seen that it is enough to solve$A\vec{x} = \text{proj}_{\text{col}(A)}\vec{b}$. Moreover, we know that we can write

$$\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{col}(A)^\perp}\vec{b} = \text{proj}_{\text{col}(A)}\vec{b} + \text{proj}_{\text{null}(A^T)}\vec{b}$$

Hence, multiplying the first equation on both sides by A^T, we find

$$A^T A\vec{x} = A^T \text{proj}_{\text{col}(A)}\vec{b} = A^T(\vec{b} - \text{proj}_{\text{null}(A^T)}\vec{b}) = A^T\vec{b}$$

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

So, we set up A,$\vec{b}$,$A^T A$ and $A^T \overset{\vec{b}}{\cancel{A}}$:

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

So, we set up A,$\vec{b}$,$A^T A$ and $A^T A$:

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix}$$

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

So, we set up A,$\vec{b}$,A^TA and A^TA:

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \quad \vec{b} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

So, we set up A,$\vec{b}$,$A^T A$ and $A^T A$:

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \quad \vec{b} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \quad A^T = \begin{pmatrix} 3 & 1 & 1 \\ 2 & -4 & 10 \\ -1 & 3 & -7 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{pmatrix}$$

Example

Find the least squares solution and least squares error for the linear system

$$3x_1 + 2x_2 - x_3 = 2$$

$$x_1 - 4x_2 + 3x_3 = -2$$

$$x_1 + 10x_2 - 7x_3 = 1$$

So, we set up A,$\vec{b}$,A^TA and A^TA:

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \quad \vec{b} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \quad A^T = \begin{pmatrix} 3 & 1 & 1 \\ 2 & -4 & 10 \\ -1 & 3 & -7 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{pmatrix} \qquad A^T\vec{b} = \begin{bmatrix} 5 \\ 22 \\ -15 \end{bmatrix}$$

Example continued

Hence we need to solve $A^T A\tilde{\vec{x}} = A^T\vec{b}$ and so putting it in an augmented matrix we get

$$(A^T A | A^T\vec{b})$$

Example continued

Hence we need to solve $A^T A = A^T \vec{b}$ and so putting it in an augmented matrix we get

$$(A^T A | A^T \vec{b}) = \left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right)$$

Example continued

Hence we need to solve $A^T A = A^T \vec{b}$ and so putting it in an augmented matrix we get

$$(A^T A | A^T \vec{b}) = \left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right) \implies \left(\begin{array}{ccc|c} 1 & 0 & 1/7 & 2/7 \\ 0 & 1 & -57 & 13/84 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Example continued

Hence we need to solve $A^TA = A^T\vec{b}$ and so putting it in an augmented matrix we get

$$(A^TA|A^T\vec{b}) = \left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right) \implies \left(\begin{array}{ccc|c} 1 & 0 & 1/7 & 2/7 \\ 0 & 1 & -57 & 13/84 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

And we see that

$$\vec{x} = \begin{bmatrix} 2/7 - t/7 \\ 13/84 + 5t/7 \\ t \end{bmatrix}$$

Example continued

Hence we need to solve $A^TA = A^T\vec{b}$ and so putting it in an augmented matrix we get

$$(A^TA|A^T\vec{b}) = \left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right) \implies \left(\begin{array}{ccc|c} 1 & 0 & 1/7 & 2/7 \\ 0 & 1 & -5/7 & 13/84 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

And we see that

$$\vec{x} = \begin{bmatrix} 2/7 - t/7 \\ 13/84 + 5t/7 \\ t \end{bmatrix} = \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/7 \\ 5/7 \\ 1 \end{bmatrix} t$$

is a least squares solution for any t.

Check whether $A\begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} = \vec{b}$

If so have a solution. If not have an app solution.

Example continued

Hence we need to solve $A^T A = A^T \vec{b}$ and so putting it in an augmented matrix we get

$$(A^T A | A^T \vec{b}) = \left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right) \implies \left(\begin{array}{ccc|c} 1 & 0 & 1/7 & 2/7 \\ 0 & 1 & -57 & 13/84 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

And we see that

$$\vec{x} = \begin{bmatrix} 2/7 - t/7 \\ 13/84 + 5t/7 \\ t \end{bmatrix} = \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/7 \\ 5/7 \\ 1 \end{bmatrix} t$$

is a least squares solution for any t.

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?)

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?) So, setting $t = 0$, we get

$$\left\|\vec{b} - A\vec{x}\right\|$$

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?) So, setting $t = 0$, we get

$$\left\|\vec{b} - A\vec{x}\right\| = \left\| \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} \right\|$$

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?) So, setting $t = 0$, we get

$$\left\|\vec{b} - A\vec{x}\right\| = \left\| \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 5/6 \\ -5/3 \\ -5/6 \end{bmatrix} \right\|$$

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?) So, setting $t = 0$, we get

$$\left\|\vec{b} - A\vec{x}\right\| = \left\| \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 5/6 \\ -5/3 \\ -5/6 \end{bmatrix} \right\| = \sqrt{(5/6)^2 + (-5/3)^2 + (-5/6)^2} = \frac{5}{6}\sqrt{6}$$

Example Continued

To find the least squares error, it is enough now to find $\left\|\vec{b} - A\vec{x}\right\|$ for *any* of the $\vec{x}$ we found above. (Exercise: Why does it not matter which $\vec{x}$ we take?) So, setting $t = 0$, we get

$$\left\|\vec{b} - A\vec{x}\right\| = \left\| \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \begin{bmatrix} 2/7 \\ 13/84 \\ 0 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 5/6 \\ -5/3 \\ -5/6 \end{bmatrix} \right\| = \sqrt{(5/6)^2 + (-5/3)^2 + (-5/6)^2} = \frac{5}{6}\sqrt{6}$$

Remark

The least squares error of a linear system will be 0 if and only if there is a solution to $A\vec{x} = \vec{b}$

Coordinates of Vectors in Other Basis

$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $e_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e}_1, \vec{e}_2, \vec{e}_3$.

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e}_1, \vec{e}_2, \vec{e}_3$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$.

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e}_1, \vec{e}_2, \vec{e}_3$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$. We commonly condense this notation to just write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e}_1, \vec{e}_2, \vec{e}_3$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$. We commonly condense this notation to just write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

However, we know that there are many other bases for $\mathbb{R}^3$.

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e}_1, \vec{e}_2, \vec{e}_3$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$. We commonly condense this notation to just write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

However, we know that there are many other bases for $\mathbb{R}^3$. In particular

$$\text{if } B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e_1}, \vec{e_2}, \vec{e_3}$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3}$. We commonly condense this notation to just write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

However, we know that there are many other bases for $\mathbb{R}^3$. In particular

$$\text{if } B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\} \text{ then } \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Coordinates of Vectors in Other Basis

We know that a basis for $\mathbb{R}^3$ is given by $\vec{e_1}, \vec{e_2}, \vec{e_3}$.And so any $\vec{x} \in \mathbb{R}^3$ can be written as $\vec{x} = x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3}$. We commonly condense this notation to just write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

However, we know that there are many other bases for $\mathbb{R}^3$. In particular

$$\text{if } B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\} \text{ then } \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

So, what to do with these new numbers $2, 2, -1$?

Coordinates with Respect to a Basis

Definition

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an ordered basis for a subspace W of $\mathbb{R}^n$

Coordinates with Respect to a Basis

Definition

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an ordered basis for a subspace W of $\mathbb{R}^n$ and if

$$\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

is the expression for a vector $\vec{w}$ in W,

Coordinates with Respect to a Basis

Definition

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an ordered basis for a subspace W of $\mathbb{R}^n$ and if

$$\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

is the expression for a vector $\vec{w}$ in W,then we call $a_1, a_2, \ldots, a_k$ the **coordinates of $\vec{w}$ with respect to B**.

Coordinates with Respect to a Basis

Definition

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an ordered basis for a subspace W of $\mathbb{R}^n$ and if

$$\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

is the expression for a vector $\vec{w}$ in W,then we call $a_1, a_2, \ldots, a_k$ the **coordinates of $\vec{w}$ with respect to** B.More specifically, we call a_j the $\vec{v}_j$-**coordinate of** $\vec{w}$.

Coordinates with Respect to a Basis

Definition

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an ordered basis for a subspace W of $\mathbb{R}^n$ and if

$$\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

is the expression for a vector $\vec{w}$ in W,then we call $a_1, a_2, \ldots, a_k$ the **coordinates of $\vec{w}$ with respect to** B.More specifically, we call a_j the $\vec{v}_j$**-coordinate of** $\vec{w}$.We denote this as either

$$(\vec{w})_B = (a_1, a_2, \ldots, a_k) \qquad \text{or} \qquad [\vec{w}]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix}$$

$$[w]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \Longleftrightarrow W = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$$

Example

If we define the two bases we had above as

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \qquad \text{and} \qquad B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

Example

If we define the two bases we had above as

$$S = \left\{ \overset{e_1}{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}, \overset{e_2}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}, \overset{e_3}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} \right\} \qquad \text{and} \qquad B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

Then we can write our vector $\vec{v} = (1, 0, 3)$ as

$$[\vec{v}]_S = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \text{ since } \vec{v} = 1 \underset{e_1}{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} + 0 \underset{e_2}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} + 3 \underset{e_3}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}$$

Example

If we define the two bases we had above as

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad B = \left\{ \overset{v_1}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}, \overset{v_2}{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}}, \overset{v_3}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} \right\}$$

Then we can write our vector $\vec{v} = (1, 0, 3)$ as

$$[\vec{v}]_S = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \text{ since } \vec{v} = 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

but

$$[\vec{v}]_B = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \text{ since } \vec{v} = 2 \underset{v_1}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}} + 2 \underset{v_2}{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}} + (-1) \underset{v_3}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}$$

Example

If we define the two bases we had above as

$$S = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\} \qquad \text{and} \qquad B = \left\{ \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix} \right\}$$

Then we can write our vector $\vec{v} = (1, 0, 3)$ as

$$\begin{bmatrix}1\\0\\1\end{bmatrix}_B = \begin{pmatrix}1\\0\\0\end{pmatrix}$$

$$[\vec{v}]_S = \begin{bmatrix}1\\0\\3\end{bmatrix} \text{ since } \vec{v} = 1\begin{bmatrix}1\\0\\0\end{bmatrix} + 0\begin{bmatrix}0\\1\\0\end{bmatrix} + 3\begin{bmatrix}0\\0\\1\end{bmatrix}$$

because $1 \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} + 0 \begin{pmatrix}0\\1\\2\end{pmatrix} + 0 \cdot \begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{pmatrix}1\\0\\1\end{pmatrix}$

but

$$[\vec{v}]_B = \begin{bmatrix}2\\2\\-1\end{bmatrix} \text{ since } \vec{v} = 2\begin{bmatrix}1\\0\\1\end{bmatrix} + 2\begin{bmatrix}0\\1\\2\end{bmatrix} + (-1)\begin{bmatrix}1\\2\\3\end{bmatrix}$$

Remarks

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$, then

$$[\vec{w}]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \iff \vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

Remarks

If $B = \{\vec{v}_1, \dots, \vec{v}_k\}$, then

$$[\vec{w}]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \iff \vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

Hence, for any of the $\vec{v}_i$, we get

i^{th} coordinate

$$[\vec{v}_i]_B = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{e}_i \text{ since } \vec{v}_i = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 1\vec{v}_i + \cdots + 0\vec{v}_k$$

Remarks

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$, then

$$[\vec{w}]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \iff \vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$$

Hence, for any of the $\vec{v}_i$, we get

$$[\vec{v}_i]_B = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{e}_i \text{ since } \vec{v}_i = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 1\vec{v}_i + \cdots + 0\vec{v}_k$$

That is, looking at vectors with respect to a certain basis can simplify matters.

Change of Basis Problem

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]_{B'}$ related?

Change of Basis Problem

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]'_B$ related?

In the case that $n = 2$, we would have $B = \{\vec{v}_1, \vec{v}_2\}$ and $B' = \{\vec{v}'_1, \vec{v}'_2\}$.

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]'_B$ related?

In the case that $n = 2$, we would have $B = \{\vec{v}_1, \vec{v}_2\}$ and $B' = \{\vec{v}'_1, \vec{v}'_2\}$. Now, by definition

$$[\vec{v}_1]_{B'} = \begin{bmatrix} a \\ b \end{bmatrix} \overset{\text{def}}{\iff} \vec{v}_1 = a\vec{v}'_1 + b\vec{v}'_2$$

Change of Basis Problem

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]'_B$ related?

In the case that $n = 2$, we would have $B = \{\vec{v}_1, \vec{v}_2\}$ and $B' = \{\vec{v}'_1, \vec{v}'_2\}$. Now, by definition

$$[\vec{v}_1]_{B'} = \begin{bmatrix} a \\ b \end{bmatrix} \overset{\text{def}}{\iff} \vec{v}_1 = a\vec{v}'_1 + b\vec{v}'_2$$

$$[\vec{v}_2]_{B'} = \begin{bmatrix} c \\ d \end{bmatrix} \overset{\text{def}}{\iff} \vec{v}_2 = c\vec{v}'_1 + d\vec{v}'_2$$

Change of Basis Problem

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]'_B$ related?

In the case that $n = 2$, we would have $B = \{\vec{v}_1, \vec{v}_2\}$ and $B' = \{\vec{v}_1', \vec{v}_2'\}$. Now, by definition

$$[\vec{v}_1]_{B'} = \begin{bmatrix} a \\ b \end{bmatrix} \overset{\text{def}}{\iff} \vec{v}_1 = a\vec{v}_1' + b\vec{v}_2'$$

$$[\vec{v}_2]_{B'} = \begin{bmatrix} c \\ d \end{bmatrix} \overset{\text{def}}{\iff} \vec{v}_2 = c\vec{v}_1' + d\vec{v}_2'$$

Now, let $\vec{w}$ be any vector, then we have

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \overset{\text{def}}{\iff} \vec{w} = k_1\vec{v}_1 + k_2\vec{v}_2$$

Question

If $\vec{w}$ is a vector in $\mathbb{R}^n$, and if we change the basis for $\mathbb{R}^n$ from a basis B to a basis B', how are $[\vec{w}]_B$ and $[\vec{w}]'_B$ related?

In the case that $n = 2$, we would have $B = \{\vec{v}_1, \vec{v}_2\}$ and $B' = \{\vec{v}'_1, \vec{v}'_2\}$. Now, by definition

$$[\vec{v}_1]_{B'} = \begin{bmatrix} a \\ b \end{bmatrix} \iff \vec{v}_1 = a\vec{v}'_1 + b\vec{v}'_2$$

$$[\vec{v}_2]_{B'} = \begin{bmatrix} c \\ d \end{bmatrix} \iff \vec{v}_2 = c\vec{v}'_1 + d\vec{v}'_2$$

Now, let $\vec{w}$ be any vector, then we have

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = k_1\vec{v}_1 + k_2\vec{v}_2 = k_1(a\vec{v}'_1 + b\vec{v}'_2) + k_2(c\vec{v}'_1 + d\vec{v}'_2)$$

Change of Basis Problem 2

Expanding and collecting like terms we see that

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = (ak_1 + ck_2)\vec{v}_1' + (bk_1 + dk_2)\vec{v}_2'$$

Change of Basis Problem 2

Expanding and collecting like terms we see that

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = (ak_1 + ck_2)\vec{v}_1' + (bk_1 + dk_2)\vec{v}_2'$$

$$\overset{\text{def}}{\iff} [\vec{w}]_{B'} = \begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$$

Change of Basis Problem 2

Expanding and collecting like terms we see that

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = (ak_1 + ck_2)\vec{v}_1' + (bk_1 + dk_2)\vec{v}_2'$$

$$\iff [\vec{w}]_{B'} = \begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$$

So the question becomes: how are $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ and $\begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$ related?

Change of Basis Problem 2

Expanding and collecting like terms we see that

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = (ak_1 + ck_2)\vec{v}_1' + (bk_1 + dk_2)\vec{v}_2'$$

$$\iff [\vec{w}]_{B'} = \begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$$

So the question becomes: how are $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ and $\begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$ related?

$$\underbrace{\begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}}_{[w]_{B'}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \underbrace{\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}}_{(w)_B}$$

Change of Basis Problem 2

Expanding and collecting like terms we see that

$$[\vec{w}]_B = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \iff \vec{w} = (ak_1 + ck_2)\vec{v}_1' + (bk_1 + dk_2)\vec{v}_2'$$

$$\iff [\vec{w}]_{B'} = \begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$$

So the question becomes: how are $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ and $\begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix}$ related?

$$\begin{bmatrix} ak_1 + ck_2 \\ bk_1 + dk_2 \end{bmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

And we see that the columns of our matrices are exactly $[\vec{v}_1]_{B'}$ and $[\vec{v}_2]_{B'}$

$[v_1]_{B'}$ $[v_2]_{B'}$

Change of Basis Theorem

Theorem (Change of Basis Theorem)

If $\vec{w}$ is a vector in $\mathbb{R}^n$ and if $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $B' = \{\vec{v}'_1, \ldots, \vec{v}'_n\}$ are bases for $\mathbb{R}^n$, then

$$[\vec{w}]_{B'} = P_{B\to B'}[\vec{w}]_B$$

where $P_{B\to B'}$ is a matrix whose columns are the vectors of B in the bassi B':

$$P_{B\to B'} = \begin{pmatrix} [\vec{v}_1]_{B'} & [\vec{v}_2]_{B'} & \ldots & [\vec{v}_n]_{B'} \end{pmatrix}$$

Change of Basis Theorem

Theorem (Change of Basis Theorem)

If $\vec{w}$ is a vector in $\mathbb{R}^n$ and if $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $B' = \{\vec{v}'_1, \ldots, \vec{v}'_n\}$ are bases for $\mathbb{R}^n$, then

$$[\vec{w}]_{B'} = P_{B\to B'}[\vec{w}]_B$$

where $P_{B\to B'}$ is a matrix whose columns are the vectors of B in the bassi B':

$$P_{B\to B'} = \begin{pmatrix} [\vec{v}_1]_{B'} & [\vec{v}_2]_{B'} & \ldots & [\vec{v}_n]_{B'} \end{pmatrix}$$

The matrix $P_{B\to B'}$ is called the **transition matrix** *(or the* **change of coordinates matrix***) from B to B'.*

Example

Find the change of coordinate matrix $P_{B \to B'}$ and $P_{B' \to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Example

Find the change of coordinate matrix $P_{B\to B'}$ and $P_{B'\to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Now, we know that

$$P_{B\to B'} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} \right)$$

Example

Find the change of coordinate matrix $P_{B \to B'}$ and $P_{B' \to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Now, we know that

$$P_{B \to B'} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} \right)$$

Further, we see that

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

Example

Find the change of coordinate matrix $P_{B \to B'}$ and $P_{B' \to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Now, we know that

$$P_{B \to B'} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} \right)$$

Further, we see that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Example

Find the change of coordinate matrix $P_{B\to B'}$ and $P_{B'\to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Now, we know that

$$P_{B\to B'} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} \right)$$

Further, we see that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

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

Example

Find the change of coordinate matrix $P_{B \to B'}$ and $P_{B' \to B}$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ to } B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Now, we know that

$$P_{B \to B'} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} \right) = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Further, we see that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{B'} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Example continued

And so, we conclude that

$$P_{B \to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Example continued

And so, we conclude that

$$P_{B \to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Similarly, we see that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Example continued

And so, we conclude that

$$P_{B\to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Similarly, we see that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Example continued

And so, we conclude that

$$P_{B \to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Similarly, we see that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 2 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Example continued

And so, we conclude that

$$P_{B\to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Same here is not surprising since $B = \{e_1, e_2\}$ the standard basis.

Similarly, we see that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 2 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and so

$$P_{B\to B'} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

Example continued

And so, we conclude that

$$P_{B\to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Similarly, we see that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 1 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 2 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and so

$$P_{B\to B'} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

Invertible Change of Basis

Theorem

If B and B' are two basis, then the change of basis matrices $P_{B \to B'}$ and $P_{B' \to B}$ are invertible and each other's inverse.

Invertible Change of Basis

Theorem

If B and B' are two basis, then the change of basis matrices $P_{B\to B'}$ and $P_{B'\to B}$ are invertible and each other's inverse. That is:

$$P^{-1}_{B\to B'} = P_{B'\to B} \text{ and } P^{-1}_{B'\to B} = P_{B\to B'}$$

$P_{B\to B'}$ goes from B to B'

$P_{B'\to B}$ goes from B' to B

So $P_{B'\to B} \cdot P_{B\to B'}$ goes from B to B' to B

Hence $P_{B'\to B} \cdot P_{B\to B'}$ is the transition from $B \to B$ which is identity.

Invertible Change of Basis

Theorem

If B and B' are two basis, then the change of basis matrices $P_{B\to B'}$ and $P_{B'\to B}$ are invertible and each other's inverse. That is:

$$P_{B\to B'}^{-1} = P_{B'\to B} \text{ and } P_{B'\to B}^{-1} = P_{B\to B'}$$

Exercise

Show that the two matrices we found from the previous example

$$\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

are inverses of each other.

Algorithm for Computing $P_{B \to B'}$

Let B and B' be two bases.

Algorithm for Computing $P_{B \to B'}$

$B = (v_1 \dots v_k)$ $B' = \{v_1', \dots v_k'\}$

$(B|B') = \left(v_1 \dots v_k \,\middle|\, v_1' \dots v_k' \right)$

Let B and B' be two bases.

1. Form the matrix $(B|B')$ where the columns of B are the vectors in basis B and the columns of B' are the vectors in B'

Algorithm for Computing $P_{B \to B'}$

Let B and B' be two bases.

1. Form the matrix $(B|B')$ where the columns of B are the vectors in basis B and the columns of B' are the vectors in B'
2. Use elementary row operations to reduce B to the identity matrix

Algorithm for Computing $P_{B \to B'}$

Let B and B' be two bases.

1. Form the matrix $(B|B')$ where the columns of B are the vectors in basis B and the columns of B' are the vectors in B'
2. Use elementary row operations to reduce B to the identity matrix
3. The resulting matrix will be $(I|P_{B \to B'})$

Orthogonal and Orthonormal Basis

As we have seen, working with some basis gives us an advantage.

Orthogonal and Orthonormal Basis

As we have seen, working with some basis gives us an advantage.

Definition

We say a basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is **orthogonal** if

$$\vec{v}_i \cdot \vec{v}_j = 0 \text{ for all } i \neq j.$$

all vectors in the basis are orthogonal to each other (perpindicular)

$V_1 \cdot V_2 = 0$

$V_1 \cdot V_3 = 0$

$V_3 \cdot V_5 = 0$

$\vdots$

Orthogonal and Orthonormal Basis

As we have seen, working with some basis gives us an advantage.

Definition

We say a basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is **orthogonal** if

$$\vec{v}_i \cdot \vec{v}_j = 0 \text{ for all } i \neq j.$$

We say the basis is **orthonormal** if it is orthogonal plus

$$\|v_i\| \qquad \|\vec{v}_i\| = 1 \text{ for all } i.$$

all vectors in the basis are normal (unit length).

Orthogonal and Orthonormal Basis

As we have seen, working with some basis gives us an advantage.

Definition

We say a basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is **orthogonal** if

$$\vec{v}_i \cdot \vec{v}_j = 0 \text{ for all } i \neq j.$$

We say the basis is **orthonormal** if it is orthogonal plus

$$\|\vec{v}_i, \vec{v}_i\| = 1 \text{ for all } i.$$

Properties of Orthogonal and Orthonormal Basis

Theorem

1. *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is an orthonormal basis for a subspace* W *and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

Properties of Orthogonal and Orthonormal Basis

Theorem

1. *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is an orthonormal basis for a subspace* W *and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

$$\vec{w} = proj_W \vec{w}$$

Properties of Orthogonal and Orthonormal Basis

Theorem

1. *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is an orthonormal basis for a subspace* W *and* $\vec{w} \in W$ *then* $= \mathcal{B}$

$$proj_W \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

$$\vec{w} = proj_W \vec{w} = (\vec{w} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{w} \cdot \vec{v}_k)\vec{v}_k$$

$$[w]_{\mathcal{B}} = \begin{pmatrix} w \cdot v_1 \\ \vdots \\ w \cdot v_k \end{pmatrix}$$

v_1-coordinate of $\vec{w}$ in $\mathcal{B}$

v_k-coordinate of $\vec{w}$ in $\mathcal{B}$

v_i-coordinate of $\bar{w}$ in $\mathcal{B}$ is $(w \cdot v_i)$

Properties of Orthogonal and Orthonormal Basis

Theorem

1. *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is an orthonormal basis for a subspace* W *and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

$$\vec{w} = proj_W \vec{w} = (\vec{w} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{w} \cdot \vec{v}_k)\vec{v}_k$$

2. *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is an orthogonal basis for a subspace* W*, and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = \frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2}\vec{v}_k$$

If $\mathcal{B}$ is orthonormal then $\|\vec{v}_i\|^2 = 1$

Properties of Orthogonal and Orthonormal Basis

Theorem

1. *If* $\{\vec{v}_1, \dots, \vec{v}_k\}$ *is an orthonormal basis for a subspace* W *and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

$$\vec{w} = proj_W \vec{w} = (\vec{w} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{w} \cdot \vec{v}_k)\vec{v}_k$$

2. *If* $\{\vec{v}_1, \dots, \vec{v}_k\}$ *is an orthogonal basis for a subspace* W*, and* $\vec{w} \in W$ *then*

$$proj_W \vec{x} = \frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2}\vec{v}_k$$

$$\vec{w} = proj_W \vec{w} = \frac{\vec{w} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots + \frac{\vec{w} \cdot \vec{v}_k}{\|\vec{v}_k\|^2}\vec{v}_k$$

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B$$

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B = \begin{bmatrix} \vec{w} \cdot \vec{v}_1 \\ \vec{w} \cdot \vec{v}_2 \\ \vdots \\ \vec{w} \cdot \vec{v}_n \end{bmatrix}$$

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B = \begin{bmatrix} \vec{w} \cdot \vec{v}_1 \\ \vec{w} \cdot \vec{v}_2 \\ \vdots \\ \vec{w} \cdot \vec{v}_n \end{bmatrix}$$

and hence if $B' = \{\vec{v}'_1, \ldots, \vec{v}'_n\}$ is another basis, then finding the transition matrix $P_{B' \to B}$ is easy:

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B = \begin{bmatrix} \vec{w} \cdot \vec{v}_1 \\ \vec{w} \cdot \vec{v}_2 \\ \vdots \\ \vec{w} \cdot \vec{v}_n \end{bmatrix}$$

and hence if $B' = \{\vec{v}_1', \dots, \vec{v}_n'\}$ is another basis, then finding the transition matrix $P_{B' \to B}$ is easy:

$$P_{B' \to B} = \begin{pmatrix} [\vec{v}_1']_B & \dots & [\vec{v}_n']_B \end{pmatrix}$$

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B = \begin{bmatrix} \vec{w} \cdot \vec{v}_1 \\ \vec{w} \cdot \vec{v}_2 \\ \vdots \\ \vec{w} \cdot \vec{v}_n \end{bmatrix}$$

and hence if $B' = \{\vec{v}_1', \dots, \vec{v}_n'\}$ is another basis, then finding the transition matrix $P_{B' \to B}$ is easy:

$$P_{B' \to B} = \begin{pmatrix} [\vec{v}_1']_B & \dots & [\vec{v}_n']_B \end{pmatrix} = \begin{pmatrix} \vec{v}_1' \cdot \vec{v}_1 & \dots & \vec{v}_n' \cdot \vec{v}_1 \\ \vdots & \ddots & \vdots \\ \vec{v}_1' \cdot \vec{v}_n & \dots & \vec{v}_n' \cdot \vec{v}_n \end{pmatrix}$$

$$P_{B \to B'} = (P_{B' \to B})^{-1}$$

Niceness of Orthonormal Basis

In particular, this theorem states that if $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\vec{w} \in \mathbb{R}^n$, then,

$$[\vec{w}]_B = \begin{bmatrix} \vec{w} \cdot \vec{v}_1 \\ \vec{w} \cdot \vec{v}_2 \\ \vdots \\ \vec{w} \cdot \vec{v}_n \end{bmatrix}$$

and hence if $B' = \{\vec{v}_1', \ldots, \vec{v}_n'\}$ is another basis, then finding the transition matrix $P_{B' \to B}$ is easy:

$$P_{B' \to B} = \begin{pmatrix} [\vec{v}_1']_B & \ldots & [\vec{v}_n']_B \end{pmatrix} = \begin{pmatrix} \vec{v}_1' \cdot \vec{v}_1 & \ldots & \vec{v}_n' \cdot \vec{v}_1 \\ \vdots & \ddots & \vdots \\ \vec{v}_1' \cdot \vec{v}_n & \ldots & \vec{v}_n' \cdot \vec{v}_n \end{pmatrix}$$

NOTE: it was imperative that we took B to be an *orthonormal* basis. This does NOT hold in general!

Gram-Schmidt Process

So we see that orthonormal bases are quite nice. Thus we want to work them as much as possible.

Gram-Schmidt Process

So we see that orthonormal bases are quite nice. Thus we want to work them as much as possible. Luckily, there is a process that will take any basis and create an orthonormal basis out of it. This is called the **Gram-Schmidt** process.

Gram-Schmidt Process

So we see that orthonormal bases are quite nice. Thus we want to work them as much as possible. Luckily, there is a process that will take any basis and create an orthonormal basis out of it. This is called the **Gram-Schmidt** process.

Suppose we have a basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ of a subspace W of R^n the algorithm on the next slide creates a new set of vectors $\{\vec{w}_1, \ldots, \vec{w}_k\}$ that is an *orthogonal* basis for W

Gram-Schmidt Process

So we see that orthonormal bases are quite nice. Thus we want to work them as much as possible. Luckily, there is a process that will take any basis and create an orthonormal basis out of it. This is called the **Gram-Schmidt** process.

Suppose we have a basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ of a subspace W of R^n the algorithm on the next slide creates a new set of vectors $\{\vec{w}_1, \ldots, \vec{w}_k\}$ that is an *orthogonal* basis for Was well as a set of vectors $\{\vec{u}_1, \ldots, \vec{u}_k\}$ that is an *orthonormal* basis for W

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2$$

Want w_2 to be orthogonal to w_1

$$V_2 = \text{proj}_{w_1} v_2 + \underline{\text{proj}_{w_1^\perp} V_2}$$

$V_2 - \text{proj}_{w_1} V_2$ is orthogonal to w_1

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2 - \mathrm{proj}_{\mathrm{span}\{\vec{w}_1\}} \vec{v}_2$$

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2 - \mathrm{proj}_{\mathrm{span}\{\vec{w}_1\}} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\left\| \vec{w}_1 \right\|^2} \vec{w}_1$$

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2} \vec{w}_1$$

3. Set

$$\vec{w}_3 = \vec{v}_3$$

$V_3 = \text{proj}_{\text{span}(w_1, w_2)} V_3 + \text{proj}_{\text{span}(w_1, w_2)^\perp} V_3$

$V_3 = \text{proj}_{\text{span}(v_1, w_2)} V_3$ is orthogonal to both

w_1 & w_2

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\left\| \vec{w}_1 \right\|^2} \vec{w}_1$$

3. Set

$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\text{span}\{\vec{w}_1, \vec{w}_2\}} \vec{v}_3$$

Gram-Schmidt Algorithm

1. Set
$$\vec{w}_1 = \vec{v}_1$$

2. Set
$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\left\|\vec{w}_1\right\|^2}\vec{w}_1$$

3. Set
$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\text{span}\{\vec{w}_1,\vec{w}_2\}}\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\left\|\vec{w}_1\right\|^2}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\left\|\vec{w}_2\right\|^2}\vec{w}_2$$

Gram-Schmidt Algorithm

1. Set

$$\vec{w}_1 = \vec{v}_1$$

2. Set

$$\vec{w}_2 = \vec{v}_2 - \mathrm{proj}_{\mathrm{span}\{\vec{w}_1\}}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1$$

3. Set

$$\vec{w}_3 = \vec{v}_3 - \mathrm{proj}_{\mathrm{span}\{\vec{w}_1,\vec{w}_2\}}\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2$$

4. Continue the process to get $\vec{w}_1, \ldots, \vec{w}_k$.

$w_j = v_j - \mathrm{proj}_{\mathrm{span}\{w_1 \ldots w_{j-1}\}} v_j \in \mathrm{span}(w_1 \ldots w_{j-1})^{\perp}$

will be orthogonal to $w_1, \ldots w_j$

Gram-Schmidt Algorithm

1. Set
$$\vec{w}_1 = \vec{v}_1$$

2. Set
$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1$$

3. Set
$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\text{span}\{\vec{w}_1,\vec{w}_2\}}\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2$$

4. Continue the process to get $\vec{w}_1, \ldots, \vec{w}_k$.This will be an orthogonal basis.

Gram-Schmidt Algorithm

1. Set
$$\vec{w}_1 = \vec{v}_1$$
2. Set
$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1$$
3. Set
$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\text{span}\{\vec{w}_1,\vec{w}_2\}}\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2$$
4. Continue the process to get $\vec{w}_1, \dots, \vec{w}_k$.This will be an orthogonal basis.
5. Set
$$\vec{u}_i = \frac{1}{\|\vec{w}_i\|}\vec{w}_i$$

Gram-Schmidt Algorithm

1. Set
$$\vec{w}_1 = \vec{v}_1$$

2. Set
$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\text{span}\{\vec{w}_1\}}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1$$

3. Set
$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\text{span}\{\vec{w}_1,\vec{w}_2\}}\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2$$

4. Continue the process to get $\vec{w}_1, \ldots, \vec{w}_k$.This will be an orthogonal basis.
5. Set
$$\vec{u}_i = \frac{1}{\|\vec{w}_i\|}\vec{w}_i$$
then $\{\vec{u}_1, \ldots, \vec{u}_k\}$ will be an orthonormal basis.