UNSW Mathematics Society Presents...

MATH2501/2601 Seminar

Presented by Jay Liang, Bruce Chen

Overview I

1. Group Theory (MATH2601 Only)
2. Group Theory (MATH2601 Only)
   - Groups
   - Fields
   - Subgroups/Subfields
   - Morphisms
3. Vector Spaces
   - Subspaces
   - Span
   - Linear Independence/Dependence
   - Basis
   - Coordinates
   - Sum of Subspaces (MATH2601 Only)
4. Linear Transformations
   - Kernel and Image

Overview II

Groups

From first-year linear algebra, you have gone through some core concepts of vector spaces. Before we dive back into that, we'll look at essentially a simpler variation as to build up some intuition towards later topics.

Groups

Definition 1: Group

A **group** G is an non-empty set with an operation $(*)$ defined onto it. They satisfy the four conditions:

- **Closure**: Suppose $a, b \in G$ then $a * b \in G$
- **Associativity**: $(a * b) * c = a * (b * c)$ for any $a, b, c \in G$.
- **Existence of identity**: There exists some $e \in G$ such that $a * e = e * a$ for all $a \in G$.
- **Existence of inverses**: For any $a \in G$, there exists $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

If the following condition is also met, we call it an **Abelian** group.

- **Commutativity**: $a * b = b * a$ for any $a, b \in G$.

This is typically denoted as $(G, *)$.

Group Theory Example

MATH2601 2017 Q4 (a)

Let a be a fixed integer, and let G be the set of all bijections $f : \mathbb{Z} \to \mathbb{Z}$ having the property

$$f(n + a) = f(n) + a \text{ for all } n \in \mathbb{Z}.$$

Prove that G is a group under composition of functions.

Group Theory Example

MATH2601 2017 Q4 (a)

Let a be a fixed integer, and let G be the set of all bijections $f : \mathbb{Z} \to \mathbb{Z}$ having the property

$$f(n + a) = f(n) + a \text{ for all } n \in \mathbb{Z}.$$

Prove that G is a group under composition of functions.

i) Closure under composition: Let g and $h \in G$.
This implies that $g(n + a) = g(n) + a$ and $h(n + a) = h(n) + a$.

$$g(h(n + a)) = g(h(n) + a) = g(h(n)) + a \in G$$

Thus, G is closed under composition of functions.

Group Theory Example

ii) Associativity: It has been stated in the course that composition of functions are always associative.

Group Theory Example

ii) Associativity: It has been stated in the course that composition of functions are always associative.

iii) Existence of identity: I claim that the identity function id is the identity element of G.

$$id(f(n+a)) = id(f(n)+a) = id(f(n))+a = f(n)+a = f(n+a)$$
$$f(id(n+a)) = f(n+a) = f(n)+a = f(n+a)$$
$$\text{So}, id \circ f = f = f \circ id$$

Thus, it has been verified that the identity function is indeed a valid identity for G.

Group Theory Example

4. Existence of inverses: As f is bijective, f^{-1} exists and is also a bijective function from $\mathbb{Z} \to \mathbb{Z}$. We are left with proving that f^{-1} is also in G.

4. Existence of inverses: As f is bijective, f^{-1} exists and is also a bijective function from $\mathbb{Z} \to \mathbb{Z}$. We are left with proving that f^{-1} is also in G.

   Let $m = f^{-1}(n)$

   Then,

$$n + a = f(m) + a = f(m + a)$$
$$f^{-1}(n + a) = f^{-1}(f(m + a)) = m + a = f^{-1}(n) + a \in G$$

   So, f^{-1} also lies in G.

Thus, G is indeed a group under composition of functions.

Properties of Groups

From the base definition of groups, we can prove some basic properties that we have become accustomed to assuming.

- Uniqueness of identity and inverses.
- $(a^{-1})^{-1} = a$.
- $(a * b)^{-1} = b^{-1} * a^{-1}$.
- If $a * b = a * c$, then $b = c$, where $a, b, c \in (G, *)$.

Fields

Extending the definition of a group by including an additional operation and a few more conditions gives us a **field** $\mathbb{F}$.

Definition 2: Field

A **field** $(\mathbb{F}, +, \times)$ is the set $\mathbb{F}$ with two operations defined on it, such that:

- $(\mathbb{F}, +)$ is **abelian**;
- $(\mathbb{F} \setminus \{0\}, \times)$ is **abelian**;
- **Multiplicative Distributivity**: $a \times (b + c) = a \times b + a \times c$, for any $a, b, c \in \mathbb{F}$.

Looks familiar?!

These rules are very reminiscent of the **10** axioms of vector spaces, although they aren't exactly the same.

Subgroups and Subfields

From first year, you have dealt with the idea of **subspaces**. The idea here is fairly similar to it, as we only have to prove a portion of the properties are satisfied.

Theorem 1: Subgroup Theorem

Consider a (non-empty) set $A \subset G$, where $(G, *)$ is a group. Then $(A, *)$ is a **subgroup** of $(G, *)$ iff all elements of A satisfy:

- **Closure under operation**: Suppose $a, b \in A$ then $a * b \in A$.
- **Existence of inverse**: For any $a \in A$, $a^{-1} \in A$.

Subgroup Example

Subgroup Theorem Question

Consider $b \in G$, where G is a group under the operation $\circ$.
Prove that $H_b \coloneqq \{b \circ a \circ b^{-1} : a \in G\}$ is a subgroup of G.

Subgroup Example

Subgroup Theorem Question

Consider $b \in G$, where G is a group under the operation $\circ$.
Prove that $H_b := \{b \circ a \circ b^{-1} : a \in G\}$ is a subgroup of G.

We must firstly check that H_b is non-empty.

We can simply check whether or not the identity element, e is in H_b.

If we choose $a = e$ then,

$$b \circ e \circ b^{-1} = b \circ b^{-1} = e$$

Thus, it is verified that $e \in H_b$ and so is non-empty.

Subgroup Example

Closure: Consider $x, y \in H_b$, such that $x = b \circ a \circ b^{-1}$ and $y = b \circ c \circ b^{-1}$, for some $a, c \in G$. We can see that:

$$\begin{aligned} x \circ y &= (b \circ a \circ b^{-1}) \circ (b \circ c \circ b^{-1}) \\ &= (b \circ a) \circ (b^{-1} \circ b) \circ (c \circ b^{-1}) && \text{(Associativity)} \\ &= (b \circ a) \circ (c \circ b^{-1}) && \text{(Identity)} \\ &= b \circ (a \circ c) \circ b^{-1} && \text{(Associativity)} \end{aligned}$$

As $a, c \in G$ and G is a group under $\circ$ then $a \circ c \in G$, i.e. $x \circ y \in H_b$.

Subgroup Example

Inverse Consider $z = b \circ a^{-1} \circ b^{-1}$ and x from before. $z \in H_b$ as $a^{-1} \in G$.
Following the working out before, we just replace c with a^{-1}.

$$
\begin{aligned}
x \circ z &= \dots \\
&= b \circ (a \circ a^{-1}) \circ b^{-1} \\
&= b \circ b^{-1} && \text{(Identity)} \\
&= e.
\end{aligned}
$$

Subgroup Example

Inverse Consider $z = b \circ a^{-1} \circ b^{-1}$ and x from before. $z \in H_b$ as $a^{-1} \in G$.
Following the working out before, we just replace c with a^{-1}.

$$\begin{aligned} x \circ z &= \dots \\ &= b \circ (a \circ a^{-1}) \circ b^{-1} \\ &= b \circ b^{-1} && \text{(Identity)} \\ &= e. \end{aligned}$$

A similar argument applies for $z \circ x = e$, and so we have that every $x \in H_b$ has an inverse.

Hence, by the Subgroup theorem, H_b is also a group under $\circ$.
Notation-wise, we say that $H_b \leq G$ under the operation, $\circ$.

Morphisms

Just like how we can define a mapping between any two sets, like from $[0, 1)$ to $\mathbb{R}$, we can also define something similar between two groups.

Definition 2: Morphism

Consider two groups $(G, *)$ and $(H, \circ)$. The mapping $f : G \to H$, is defined as a **homomorphism** from G to H if it satisfies the following:

$$f(a * b) = f(a) \circ f(b)$$

for any $a, b \in G$.

If this mapping is a bijection, we call say that both groups are **isomorphic** to each other.

Properties of Homomorphisms

From the definition in the previous slide, some neat properties that hold are:

- **Inverse maps to inverse**: $f(a^{-1}) = f^{-1}(a)$.
- **Identity maps to identity**: $f(e) = e'$, where $e,\ e'$ are the identity elements of G and H, respectively.

Homomorphism Example

MATH2601 2008 Q3c)(ii)

Suppose that $G = \left\{ \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} : t \in \mathbb{R} \right\}$ is a group under matrix multiplication. Is $(G, \times)$ isomorphic to $(\mathbb{R}, +)$?

Homomorphism Example

Consider $f : (\mathbb{R}, +) \to (G, \times)$ where $f(t)$ gives the matrix above, and $s, t \in \mathbb{R}$. We'll firstly show that this is a homomorphism.

$$\begin{aligned} f(s) \times f(t) &= \begin{pmatrix} 1 & s & \frac{s^2}{2} \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & s+t & \frac{t^2}{2} + st + \frac{s^2}{2} \\ 0 & 1 & s+t \\ 0 & 0 & 1 \end{pmatrix} = f(s+t) \end{aligned}$$

From this, $(G, \times)$ is homomorphic to $(\mathbb{R}, +)$ and as f is a bijection as well (it is both injective and surjective). Hence these two groups are also **isomorphic** to each other.

Kernel and Image

Consider a homomorphism $f : G \to H$. We define two special types of sets that come from this mapping, namely the **kernel** and **image**.

Definition 3: Kernel

The **kernel** is the 'roots' of f:

$$ker(f) := \{a \in G : f(a) = e'\}$$

where e' is the identity element of H.

Definition 4: Image

The **image** of f is the 'projections' of f:

$$im(f) := \{y \in H : f(x) = y, \quad \text{for some } x \in G\}.$$

Vector Spaces

The formal definition is the following:

Definition 5: Vector Spaces

A vector space over field $\mathbb{F}$ is valid if

a) $(V, +)$ is an Abelian group
b) Scalar multiplication by any $\alpha \in \mathbb{F}$ is valid
c) Associativity of scalar multiplication, i.e, $\alpha(\beta v) = (\alpha\beta)v$
d) $1v = v$
e) $\alpha(u + v) = \alpha u + \alpha v$
f) $(\alpha + \beta)u = \alpha u + \beta u$

*u, v are vectors in V and α, β is in $\mathbb{F}$

Vector Subspaces

Theorem 2: Subspace Theorem

Suppose U is a non-empty subset of the vector space, $(V, \mathbb{F})$. Then $(U, \mathbb{F})$ is a **subspace** of $(V, \mathbb{F})$ if the following is condition is met:

$$\lambda \mathbf{x} + \mathbf{y} \in U$$

for any $\lambda \in \mathbb{F}$ and $\mathbf{x}, \mathbf{y} \in U$.

Why does this work?

If this condition is met, then we can show that:

- **Closure under scalar multiplication**: Set $\mathbf{y} = \mathbf{0}$.
- **Closure under vector addition**: Set $\lambda = 1$.
- **Existence of vector inverse**: Set $\lambda = -1$ and $\mathbf{y} = \mathbf{x}$.

Subspace Example

MATH2601 2017 Q2a (i)

Consider the set:

$$W_1 = \left\{ \begin{pmatrix} x_1 & -x_2 \\ x_1 + x_2 & 3x_1 - x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R} \right\}.$$

Prove that this is a subspace of $M_{2,2}(\mathbb{R})$.

Subspace Example

Clearly W_1 is a (non-empty) subset of $M_{2,2}(\mathbb{R})$. Consider $X, Y \in G$ and $\lambda \in \mathbb{R}$.

$$\begin{aligned} X + \lambda Y &= \begin{pmatrix} x_1 & -x_2 \\ x_1 + x_2 & 3x_1 - x_2 \end{pmatrix} + \lambda \begin{pmatrix} y_1 & -y_2 \\ y_1 + y_2 & 3y_1 - y_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + \lambda y_1 & -(x_2 + \lambda y_2) \\ x_1 + \lambda y_1 + x_2 + \lambda y_2 & 3(x_1 + \lambda y_1) - (x_2 + \lambda y_2) \end{pmatrix} \end{aligned}$$

Subspace Example

Clearly W_1 is a (non-empty) subset of $M_{2,2}(\mathbb{R})$. Consider $X, Y \in G$ and $\lambda \in \mathbb{R}$.

$$\begin{aligned} X + \lambda Y &= \begin{pmatrix} x_1 & -x_2 \\ x_1 + x_2 & 3x_1 - x_2 \end{pmatrix} + \lambda \begin{pmatrix} y_1 & -y_2 \\ y_1 + y_2 & 3y_1 - y_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + \lambda y_1 & -(x_2 + \lambda y_2) \\ x_1 + \lambda y_1 + x_2 + \lambda y_2 & 3(x_1 + \lambda y_1) - (x_2 + \lambda y_2) \end{pmatrix} \end{aligned}$$

As $X + \lambda Y$ is in the same form of W_1, where $u_1 = x_1 + \lambda y_1$ and $u_2 = x_2 + \lambda y_2$, then $X + \lambda Y \in W_1$.

Thus, by the Subspace Theorem, we have shown that $W_1 \leq M_{2,2}(\mathbb{R})$.

Span

Definition 6: Spans

Suppose we have a set of vectors $S := \{v_1, v_2, \cdots, v_n\} \subset V$. We define the span of W as the set of all possible linear combinations of the elements of W, i.e.

$$\text{span}(S) = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n$$

where $\lambda_i \in \mathbb{F}\ \forall i = 1, 2, \ldots, n$.

It is quite easy to see that $\text{span}(S) \leq V$, and in the case when $\text{span}(S) = V$, we call S a **spanning set**.

Linear Independent and Dependent

Definition 7: Linear Independence

We call a set **linearly independent** if only the trivial linear combination maps to the zero vector, i.e. if we have

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

then $\lambda_i = 0$ for all $i = 1, 2, \cdots, n$.

We call the set **linearly dependent** otherwise.

Linearly 'independent'

Linear independence means that we are unable to express any of the vectors in the set as a linear combination of all the other vectors.

Linear Independence/Dependence Example

MATH2601 2017 Q1d

Show that the set $B = \{cos^3(t), cos^2(t)sin(t), cos(t)sin^2(t), sin^3(t)\}$ is linearly independent.

Linear Independence/Dependence Example

MATH2601 2017 Q1d

Show that the set $B = \{cos^3(t), cos^2(t)sin(t), cos(t)sin^2(t), sin^3(t)\}$ is linearly independent.

Consider

$$\lambda_1 cos^3(t) + \lambda_2 cos^2(t)sin(t) + \lambda_3 cos(t)sin^2(t) + \lambda_4 sin^3(t) = 0$$

We are intending to show that every $\lambda_i = 0$. We are going to do this by using the fact that these λ's have to work for all values of t.

If we set $t = 0$ then,

$$\lambda_1 cos^3(t) + \lambda_2 cos^2(t)sin(t) + \lambda_3 cos(t)sin^2(t) + \lambda_4 sin^3(t) = \lambda_1$$

And so, $\lambda_1 = 0$

Linear Independence/Dependence Example

Similarly, using the same trick, by substituting in $t = \frac{\pi}{2}$, we would get that,

$$\lambda_1 cos^3(t) + \lambda_2 cos^2(t) sin(t) + \lambda_3 cos(t) sin^2(t) + \lambda_4 sin^3(t) = \lambda_4$$

And so, $\lambda_4 = 0$.

Linear Independence/Dependence Example

Similarly, using the same trick, by substituting in $t = \frac{\pi}{2}$, we would get that,

$$\lambda_1 cos^3(t) + \lambda_2 cos^2(t)sin(t) + \lambda_3 cos(t)sin^2(t) + \lambda_4 sin^3(t) = \lambda_4$$

And so, $\lambda_4 = 0$.

Now, as λ_1 and $\lambda_4 = 0$, then our equation simplifies to,

$$\lambda_2 cos^2(t)sin(t) + \lambda_3 cos(t)sin^2(t) = 0$$

Setting $t = \frac{\pi}{4}$ and $-\frac{\pi}{4}$, we would end up with λ_2 and λ_3 also equal to 0.

Thus, the set B is linearly independent.

Basis

Naturally we'll be interested in ways we can **minimally** represent a vector space, such as through **spanning sets**.

Definition 8: Basis

The most basic sets we can choose are called **basis sets**, which obtain the following two properties:

- $\text{span}(S) = V$ i.e. spanning set of V
- S is linearly independent

Why these two properties only?

The first one is simple as we want to describe the whole set. By obtaining l.i., we are saying that each vector is 'pulling their own weight'.

Notable qualities of Basis sets

- **Uniqueness of representation**: Each $\mathbf{x} \in V$ is uniquely represented as a linear combination of the basis vectors.
- **Dependent on type of field**: Valid choices of basis also depend on the field accompanying the vector space, e.g. $(\mathbb{C}, +, \times, \mathbb{R})$ vs $(\mathbb{C}, +, \times, \mathbb{C})$.

Standard Basis

The most simplistic basis we generally call 'standard' basis, some of which you should be very familiar with:

- $\mathbb{R}^n : \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$
- $M_{n,m}(\mathbb{R}) : \{E_{1,1}, E_{1,2}, \cdots, E_{n,m}\}$

Dimensions

Not only are there multiple basis for each vector space, but they all have the **same number of elements**. As a result, we define the following useful concept of the **dimension** of a vector space.

Definition 9: Dimension

Consider a vector space, V, which has a **finite spanning basis**, S. Then, we define the size of S to be the dimension of V, $\dim(V) = |S|$.

Lemma: Equivalent ways of proving basis

The following three statements are equivalent (where $\dim(V) = n$):

- S is a basis of V
- S is linearly independent and $|S| = n$
- S is a spanning set and $|S| = n$

Basis Sets

Basis Vectors Example

Consider the degree-2 polynomial vector space, $\mathcal{P}_2(\mathbb{R})$. Show that the following set is a basis of this vector space,
$S = \{2 + 3x, 4x - x^2, 1 + x^2\}$.

Basis Sets

Basis Vectors Example

Consider the degree-2 polynomial vector space, $\mathcal{P}_2(\mathbb{R})$. Show that the following set is a basis of this vector space,
$S = \{2 + 3x, 4x - x^2, 1 + x^2\}$.

We know that $\dim(\mathcal{P}_2(\mathbb{R})) = 3$ and so we only need to show that S is linearly independent. To do this, we'll again begin by considering the usual linear independence equation,

$$\lambda_1(2 + 3x) + \lambda_2(4x - x^2) + \lambda_3(1 + x^2) = 0$$
$$(2\lambda_1 + \lambda_3) + x(3\lambda_1 + 4\lambda_2) + x^2(-\lambda_2 + \lambda_3) = 0$$

Basis Sets Example

And so, comparing the coefficients on both sides we can form the augmented matrix,

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right)$$

Basis Sets Example

And so, comparing the coefficients on both sides we can form the augmented matrix,

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right)$$

When reduced, it becomes

$$\left(\begin{array}{ccc|c} 3 & 4 & 0 & 0 \\ 0 & -\frac{8}{3} & 1 & 0 \\ 0 & 0 & \frac{5}{8} & 0 \end{array}\right)$$

So, clearly every column is leading and thus every column is linearly independent.
Thus, by the previous dimension theorem, we can say that S is a basis of $\mathcal{P}_2(\mathbb{R})$.

Coordinates

When looking back at the Cartesian plane, we described all of the points **uniquely** as 'coordinates'. This was possible as the x-y directional vectors were **basis vectors of** $\mathbb{R}^2$. We'll be extending this idea to apply to more general basis'.

Definition 10: Coordinate Vector

Suppose we have a basis $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ and $\mathbf{x} \in V$ such that:

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.$$

Then, we define the vector $[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ to be the **coordinate vector** of $\mathbf{x}$ w.r.t $\mathcal{B}$.

How to find Coordinate Vectors?

The most basic approach for finding these coordinate vectors is solving the matrix equation: $V\boldsymbol{\alpha} = \mathbf{x}$, where the columns of V are the basis vectors.

Other approaches exist, but are mostly circumstantial, e.g. for polynomials we can sub in n different values and find the coefficients.

Coordinate Vector Example

Coordinate Vector Example

Let V be the vector space of all 2×2 real symmetric matrices, and $A = \begin{pmatrix} 9 & 5 \\ 5 & -4 \end{pmatrix}$. Find the corresponding coordinate vector w.r.t the following basis:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ -1 & -5 \end{pmatrix} \right\}.$$

Coordinate Vector Example

Coordinate Vector Example

Let V be the vector space of all 2×2 real symmetric matrices, and $A = \begin{pmatrix} 9 & 5 \\ 5 & -4 \end{pmatrix}$. Find the corresponding coordinate vector w.r.t the following basis:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ -1 & -5 \end{pmatrix} \right\}.$$

Aside: Despite a 2 x 2 matrix containing 4 entries, when restricted to the real symmetric matrices, the basis required is only of dimension 3. This is because, the standard basis for the symmetric matrices would be, $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$.

Coordinate Vector Example

We want to find a vector α such that:

$$A = \sum_{i=1}^{3} \alpha_i B_i.$$

Rearranging this, so that we focus on each individual component of the matrices yields:

$$\left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ -2 & 1 & -1 & 5 \\ 1 & 3 & -5 & -4 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ 0 & 5 & 7 & 23 \\ 0 & -1 & 9 & 13 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ 0 & 5 & 7 & 23 \\ 0 & 0 & 52 & 88 \end{array}\right)$$

Coordinate Vector Example

We want to find a vector α such that:

$$A = \sum_{i=1}^{3} \alpha_i B_i.$$

Rearranging this, so that we focus on each individual component of the matrices yields:

$$\left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ -2 & 1 & -1 & 5 \\ 1 & 3 & -5 & -4 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ 0 & 5 & 7 & 23 \\ 0 & -1 & 9 & 13 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 4 & 9 \\ 0 & 5 & 7 & 23 \\ 0 & 0 & 52 & 88 \end{array}\right)$$

Solving this augmented matrix leads to:

$$\alpha_3 = \frac{22}{13},\ \alpha_2 = \frac{29}{13},\ \alpha_1 = -\frac{29}{13}$$

Hence, the coordinate vector of A w.r.t. $\mathcal{B}$ is:

$$[A]_{\mathcal{B}} = \frac{1}{13}(-29, 29, 22)^T$$

Sums of Vector Spaces (MATH2601)

Suppose that we have two subspaces, $U, W \leq V$. We define their **sum** as:

Definition 11: Sum of Subspaces

$U + W = \{\mathbf{y} \in V : \mathbf{y} = \mathbf{u} + \mathbf{w}, \text{for some } \mathbf{u} \in U, \mathbf{w} \in W\}$.

In the case that these intersection of these two subspaces only contains the $\mathbf{0}$ vector, then we call this a **direct sum** (denoted by $U \oplus W$).

Sums of Vector Spaces (MATH2601)

The concept of summing vector spaces actually leads to a very familiar result about sets, i.e. the cardinality relationship.

Theorem: Dimensions of Sum of Subspaces

Suppose that U, W are finite subspaces of V, then we have

$$\dim(U) + \dim(W) = \dim(U + W) + \dim(U \cap W).$$

This can be used to help determine whether the sum is direct or not.

Sum of Subspaces Example (MATH2601)

MATH2601 2019 1b

Let $V = M_{2,3}(\mathbb{C})$, the complex space of all 2×3 complex matrices. Also let,

$$S = \left\{ A \in V : A \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

and

$$T = \{A \in V : (1\ 2)A = (0\ 0\ 0)\}.$$

i) Find bases for $S \cap T, S$ and T.

ii) Does $S + T = V$? Give reasons for your answer.

Sum of Subspaces Example (MATH2601)

i) Let $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$.

Then for S,

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} a + 2b - c \\ d + 2e - f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Sum of Subspaces Example (MATH2601)

i) Let $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$.

Then for S,

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} a + 2b - c \\ d + 2e - f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

And so from the above matrices, we can observe that,

$$c = a + 2b$$
$$f = d + 2e$$

Sum of Subspaces Example (MATH2601)

i) Let $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$.

Then for S,

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} a + 2b - c \\ d + 2e - f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

And so from the above matrices, we can observe that,

$$c = a + 2b$$
$$f = d + 2e$$

Thus, if we set a, b, d and e as our free parameters, then the matrix A in the set S will look like,

$$\begin{pmatrix} a & b & a + 2b \\ d & e & d + 2e \end{pmatrix}$$

Sum of Subspaces Example (MATH2601)

So a suitable basis for S would be,

$$\left\{ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix} \right\}.$$

Similarly, applying the exact same method to T we would obtain that the restrictions on T are,

$$\begin{aligned} a + 2d &= 0 \\ b + 2e &= 0 \\ c + 2f &= 0 \end{aligned}$$

And so, a suitable basis for T would be,

$$\left\{ \begin{pmatrix} -2 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix} \right\}.$$

Sum of Subspaces Example (MATH2601)

Now, for $S \cap T$. We observe that this would require all 6 restrictions shared amongst S and T.

From the restrictions on S: c and f have become fixed values.

Sum of Subspaces Example (MATH2601)

Now, for $S \cap T$. We observe that this would require all 6 restrictions shared amongst S and T.

From the restrictions on S: c and f have become fixed values.

And from the restrictions on T: any 2 from a, b, d or e have become fixed.

And so, we would expect $S \cap T$ to be dimension 2, as 4 of the 6 entries are now fixed.

So rather than calculating $S \cap T$, we can just try cook something up from the bases for S and T.

Sum of Subspaces Example (MATH2601)

We see that -2 $\times$ the first element added to the third element of the basis for S is equal to the first element added to the third element of the basis for T.

Sum of Subspaces Example (MATH2601)

We see that -2 $\times$ the first element added to the third element of the basis for S is equal to the first element added to the third element of the basis for T.

Similarly, we observe that -2 $\times$ the second element added to the fourth element of the basis for S is equal to the second element added to 2 $\times$ the third element of the basis for T.

Sum of Subspaces Example (MATH2601)

We see that -2 $\times$ the first element added to the third element of the basis for S is equal to the first element added to the third element of the basis for T.

Similarly, we observe that -2 $\times$ the second element added to the fourth element of the basis for S is equal to the second element added to 2 $\times$ the third element of the basis for T.

And so, we have found 2 unique matrices that lie in $S \cap T$, thus our basis can be,

$$\left\{ \begin{pmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -2 & -4 \\ 0 & 1 & 2 \end{pmatrix} \right\}.$$

Sum of Subspaces Example (MATH2601)

ii) Does $S + T = V$? Give reasons for your answer.

By the Theorem on the Dimensions on the Sum of Subspaces,

$$\begin{aligned} dim(S+T) &= dim(S) + dim(T) - dim(S \cap T) \\ &= 4 + 3 - 2 = 5 \end{aligned}$$

The dimensions of V is however 6, and so **no**, $S + T \neq V$.

Linear Transformations

Definition: Linear Transformation

The mapping $T : V \to W$ (between vector spaces over the same field $\mathbb{F}$) is called a linear transformation iff

- $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$
- $T(\lambda \mathbf{x}) = \lambda T(\mathbf{x})$

for all $\mathbf{x}, \mathbf{y} \in V, \lambda \in \mathbb{F}$.

Luckily, we can prove these two conditions just in a single line!

Lemma

Consider a mapping $T : V \to W$, in which V, W are both vector spaces. Then T is a linear map iff

$$T(\mathbf{x} + \lambda \mathbf{y}) = T(\mathbf{x}) + \lambda T(\mathbf{y})$$

for any $\mathbf{x}, \mathbf{y} \in V, \lambda \in \mathbb{F}$.

Linear Transformation Example

Example

The function $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$ is defined by:

$$T(p) = (p(1), p(3), p'(2)).$$

Prove that this is linear.

Linear Transformation Example

Example

The function $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$ is defined by:

$$T(p) = (p(1), p(3), p'(2)).$$

Prove that this is linear.

Consider $p, q \in \mathbb{P}_2(\mathbb{R})$ and $\lambda \in \mathbb{R}$. Now

$$T(p + \lambda q) = ((p + \lambda q)(1), (p + \lambda q)(3), (p + \lambda q)'(2)).$$

Linear Transformation Example

Example

The function $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$ is defined by:

$$T(p) = (p(1), p(3), p'(2)).$$

Prove that this is linear.

Consider $p, q \in \mathbb{P}_2(\mathbb{R})$ and $\lambda \in \mathbb{R}$. Now

$$T(p + \lambda q) = ((p + \lambda q)(1), (p + \lambda q)(3), (p + \lambda q)'(2)).$$

Since differentiation is itself linear, we have

$$\begin{aligned} T(p + \lambda q) &= (p(1) + \lambda q(1), p(3) + \lambda q(3), p'(2) + \lambda q'(2)) \\ &= (p(1), p(3), p'(2)) + \lambda (q(1), q(3), q'(2)) \\ &= T(p) + \lambda T(q). \end{aligned}$$

So T is a linear transformation.

Kernel and Image of T

Consider the linear transformation $T : V \to W$.

Definition: Kernel

The **kernel** is the set of roots of T, that is,

$$\ker(T) \coloneqq \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$$

Kernel and Image of T

Consider the linear transformation $T : V \to W$.

Definition: Kernel

The **kernel** is the set of roots of T, that is,

$$\ker(T) \coloneqq \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$$

Definition: Image

The **image** is the set of possible values that $T(\mathbf{x})$ can take, that is,

$$\operatorname{im}(T) \coloneqq \{\mathbf{y} \in W : T(\mathbf{x}) = \mathbf{y}, \text{for some } \mathbf{x} \in V\}$$

These sets are both **vector spaces** under the field $\mathbb{F}$ as well.

Rank and Nullity

The dimensions of the **kernel** and the **image** are given special names: the **nullity** and the **rank** respectively.

Rank-Nullity Theorem

Consider a linear transformation $T : V \to W$, where V is finite-dimensional vector space. We have:

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Rank and Nullity

The dimensions of the **kernel** and the **image** are given special names: the **nullity** and the **rank** respectively.

Rank-Nullity Theorem

Consider a linear transformation $T : V \to W$, where V is finite-dimensional vector space. We have:

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(V).$$

Now suppose that $\dim(V) = \dim(W) = n$. Then the following statements are **equivalent**:

Properties of the Rank & Nullity

- $\operatorname{nullity}(T) = 0$.
- $\operatorname{rank}(T) = \dim(V) = n$.
- T is an invertible mapping.

Rank and Nullity Example

Example

Consider the linear transformation $T : \mathbb{R}^5 \to \mathbb{R}^4$ given by

$$T(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 + 2x_4 + x_5 \\ -x_3 - 2x_4 + 2x_5 \\ -x_1 - x_2 + x_3 + 4x_4 - x_5 \\ x_1 + x_1 + x_4 + 2x_5 \end{pmatrix}.$$

Find the rank and nullity of T. What is a basis for image(T)?

Rank and Nullity Example

Example

Consider the linear transformation $T : \mathbb{R}^5 \to \mathbb{R}^4$ given by

$$T(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 + 2x_4 + x_5 \\ -x_3 - 2x_4 + 2x_5 \\ -x_1 - x_2 + x_3 + 4x_4 - x_5 \\ x_1 + x_1 + x_4 + 2x_5 \end{pmatrix}.$$

Find the rank and nullity of T. What is a basis for image(T)?

Remember the equivalence between matrices and linear maps between $\mathbb{R}^n$ and $\mathbb{R}^m$. In this case, $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{pmatrix} \\ \\ \\ \end{pmatrix}.$$

Rank and Nullity Example

Example

Consider the linear transformation $T : \mathbb{R}^5 \to \mathbb{R}^4$ given by

$$T(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 + 2x_4 + x_5 \\ -x_3 - 2x_4 + 2x_5 \\ -x_1 - x_2 + x_3 + 4x_4 - x_5 \\ x_1 + x_1 + x_4 + 2x_5 \end{pmatrix}.$$

Find the rank and nullity of T. What is a basis for image(T)?

Remember the equivalence between matrices and linear maps between $\mathbb{R}^n$ and $\mathbb{R}^m$. In this case, $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 & 2 \\ -1 & -1 & 1 & 4 & -1 \\ 1 & 1 & 0 & 4 & 2 \end{pmatrix}.$$

Rank and Nullity Example

Once we've gotten our matrix, we can do our regular row-reductions until we reach row-echelon form.

$$\begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 & 2 \\ -1 & -1 & 1 & 4 & -1 \\ 1 & 1 & 0 & 4 & 2 \end{pmatrix} \begin{array}{c} \xrightarrow{R_1+R_3} \\ \xrightarrow{R_3+R_4} \end{array} \begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

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

Rank and Nullity Example

Once we've gotten our matrix, we can do our regular row-reductions until we reach row-echelon form.

$$\begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 & 2 \\ -1 & -1 & 1 & 4 & -1 \\ 1 & 1 & 0 & 4 & 2 \end{pmatrix} \begin{matrix} \xrightarrow{R_1+R_3} \\ \xrightarrow{R_3+R_4} \end{matrix} \begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \color{red}{1} & 1 & 0 & 2 & 1 \\ 0 & 0 & \color{red}{-1} & -2 & 2 \\ 0 & 0 & 0 & \color{red}{4} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Because this has three leading columns, $\text{rank}(A) = 3$. ($\text{nullity}(A) = 2$, the number of non-leading columns.)

Rank and Nullity Example

To find a basis for $\mathrm{im}(A)$ we simply select the columns in the original matrix corresponding to the leading columns, that is,

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 4 \\ 4 \end{pmatrix} \right\}.$$

Matrix Representation of Linear Maps

It turns out that **every** finite-dimensional linear map can be represented by matrix multiplication.

Matrix Representation of Linear Maps

It turns out that **every** finite-dimensional linear map can be represented by matrix multiplication.

Theorem: Matrix Representation of Linear Maps

Consider the linear transformation $T : V \to W$, with basis $\mathcal{B}$ for the domain and $\mathcal{C}$ for the codomain. We can find an unique **matrix representation** A such that:

$$[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$$

for any $\mathbf{v} \in V$.

So if we know the coordinate vector of $\mathbf{v}$ with respect to a given basis $\mathcal{B}$, we can use this matrix to determine $T(\mathbf{v})$. We will use $A = [T]^{\mathcal{B}}_{\mathcal{C}}$ as our notation for this **matrix representation**.

Procedure: Finding the Matrix Representation

For the linear map $T : V \to W$, suppose that $\dim(V) = p$ and $\dim(W) = q$.

1. Choose ordered bases for the domain and co-domain, $\mathcal{B}$ and $\mathcal{C}$.
2. Take the first vector of $\mathcal{B}$, say, $\mathbf{v}_1$. Calculate $T(\mathbf{v}_1)$ and determine its corresponding coordinate vector with respect to $\mathcal{C}$.
3. The vector from 2, that is, $[T(\mathbf{v}_1)]_{\mathcal{C}}$, will be the first column of A.
4. Repeat steps 2 and 3 for the other $p - 1$ basis vectors, completing the next $p - 1$ columns of A.

Procedure: Finding the Matrix Representation

For the linear map $T : V \to W$, suppose that $\dim(V) = p$ and $\dim(W) = q$.

1. Choose ordered bases for the domain and co-domain, $\mathcal{B}$ and $\mathcal{C}$.
2. Take the first vector of $\mathcal{B}$, say, $\mathbf{v}_1$. Calculate $T(\mathbf{v}_1)$ and determine its corresponding coordinate vector with respect to $\mathcal{C}$.
3. The vector from 2, that is, $[T(\mathbf{v}_1)]_{\mathcal{C}}$, will be the first column of A.
4. Repeat steps 2 and 3 for the other $p-1$ basis vectors, completing the next $p-1$ columns of A.

When $\mathcal{B}$ and $\mathcal{C}$ are standard bases, this will generally be quite easy. However, sometimes our bases will be substantially more complicated, in which case it will be quite difficult to compute coordinate vectors. In these cases, we use the **indirect approach**.

Indirect Approach to Finding $[T]^{\mathcal{B}}_{\mathcal{C}}$

The Change of Basis Matrix

The change of basis matrix between matrices is the matrix representation of the identity matrix with respect to different bases $\mathcal{B}$ and $\mathcal{C}$ in the domain and the co-domain respectively, that is, $[\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}$.

$$[\mathbf{v}]_{\mathcal{C}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

Indirect Approach to Finding $[T]^{\mathcal{B}}_{\mathcal{C}}$

The Change of Basis Matrix

The change of basis matrix between matrices is the matrix representation of the identity matrix with respect to different bases $\mathcal{B}$ and $\mathcal{C}$ in the domain and the co-domain respectively, that is, $[\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}$.

$$[\mathbf{v}]_{\mathcal{C}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

Theorem: Matrix Representation of Linear Composition

Consider linear transformations $S : U \to V$ and $T : V \to W$, with bases $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ for vector spaces U, V and W. Then

$$[T \circ S]^{\mathcal{A}}_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}.$$

Indirect Approach to Finding $[T]^{\mathcal{B}}_{\mathcal{C}}$

The Change of Basis Matrix

The change of basis matrix between matrices is the matrix representation of the identity matrix with respect to different bases $\mathcal{B}$ and $\mathcal{C}$ in the domain and the co-domain respectively, that is, $[\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}$.

$$[\mathbf{v}]_{\mathcal{C}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

Theorem: Matrix Representation of Linear Composition

Consider linear transformations $S : U \to V$ and $T : V \to W$, with bases $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ for vector spaces U, V and W. Then

$$[T \circ S]^{\mathcal{A}}_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}.$$

Corollary

For invertible T, $[T^{-1}]^{\mathcal{C}}_{\mathcal{B}}[T]^{\mathcal{B}}_{\mathcal{C}} = I$. Hence, $[\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}}[\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}} = I$.

Matrix Representation Example

2016 FE

Consider the linear transformation from earlier: $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$ and the following two respective bases: $\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\}, \mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$. Find the representation matrix of T with respect to these two bases.

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\}, \ \mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

These bases are a tad annoying, so let's break this up.

$$[T]^{\mathcal{B}}_{\mathcal{C}} = [\mathrm{id}_W]^{\mathcal{S}'}_{\mathcal{C}} \times [T]^{\mathcal{S}}_{\mathcal{S}'} \times [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{S}}.$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1+t, 2-t^2, 3t-t^2\}$, $\mathcal{C} = \{(1,0,0),(0,1,0),(-1,2,1)\}$

These bases are a tad annoying, so let's break this up.

$$[T]^{\mathcal{B}}_{\mathcal{C}} = [\mathrm{id}_W]^{\mathcal{S}'}_{\mathcal{C}} \times [T]^{\mathcal{S}}_{\mathcal{S}'} \times [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{S}}.$$

Here,

1. $[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{S}}$: changes the basis from $\mathcal{B}$ to the standard basis of $\mathbb{P}_2(\mathbb{R})$
2. $[T]^{\mathcal{S}}_{\mathcal{S}'}$: is our matrix representation with respect to the standard basis (the one that isn't too hard to calculate!)
3. $[\mathrm{id}_W]^{\mathcal{S}}_{\mathcal{C}}$: changes the basis from the standard basis of $\mathbb{R}^3$ back to $\mathcal{C}$.

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\}$, $\mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

1. Convert from our annoying basis $\mathcal{B}$ to the standard basis of degree-2 polynomials:

$$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\} \xrightarrow{\text{id}} \{1, t, t^2\}.$$

$$[\text{id}_V]^{\mathcal{B}}_{\mathcal{S}} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & -1 & -1 \end{pmatrix}.$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1+t, 2-t^2, 3t-t^2\}$, $\mathcal{C} = \{(1,0,0), (0,1,0), (-1,2,1)\}$

2. Find the matrix representation with respect to the standard bases:

$$\{1, t, t^2\} \xrightarrow{T} \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

$$\begin{aligned} [T(1)]_{\mathcal{S}'} &= [(1,1,0)]_{\mathcal{S}'} \\ &= (1,1,0)^T \end{aligned}$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\}$, $\mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

2. Find the matrix representation with respect to the standard bases:

$$\{1, t, t^2\} \xrightarrow{T} \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

$$\begin{aligned} [T(1)]_{\mathcal{S}'} &= [(1, 1, 0)]_{\mathcal{S}'} \\ &= (1, 1, 0)^T \\ [T(t)]_{\mathcal{S}'} &= [(1, 2, 1)]_{\mathcal{S}'} \\ &= (1, 2, 1)^T \end{aligned}$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\}$, $\mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

2. Find the matrix representation with respect to the standard bases:

$$\{1, t, t^2\} \xrightarrow{T} \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

$$\begin{aligned}
[T(1)]_{\mathcal{S}'} &= [(1, 1, 0)]_{\mathcal{S}'} \\
&= (1, 1, 0)^T \\
[T(t)]_{\mathcal{S}'} &= [(1, 2, 1)]_{\mathcal{S}'} \\
&= (1, 2, 1)^T \\
[T(t^2)]_{\mathcal{S}'} &= [(1, 4, 4)]_{\mathcal{S}'} \\
&= (1, 4, 4)^T.
\end{aligned}$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1+t, 2-t^2, 3t-t^2\}$, $\mathcal{C} = \{(1,0,0), (0,1,0), (-1,2,1)\}$

2. Matrix representation with standard basis:

$$\{1, t, t^2\} \xrightarrow{T} \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

So the matrix representation for the standard basis is:

$$[T]^{\mathcal{S}}_{\mathcal{S}'} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 1 & 4 \end{pmatrix}.$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\},\ \mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

3. Convert from the standard basis back to $\mathcal{C}$, our complicated basis:

$$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \xrightarrow{\text{id}} \mathcal{C}.$$

$$[\text{id}_W]_{\mathcal{C}}^{\mathcal{S}} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

Matrix Representation Example

2016 FE

$T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$, $T(p) = (p(1), p(2), p'(2))$
$\mathcal{B} = \{1 + t, 2 - t^2, 3t - t^2\},\ \mathcal{C} = \{(1, 0, 0), (0, 1, 0), (-1, 2, 1)\}$

From this, we have the full representation matrix

$$[T]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & -1 & -1 \end{pmatrix}$$

with respect to bases $\mathcal{B}$ and $\mathcal{C}$.

Invariant Transformations (MATH2601)

Definition: Invariant Subspace

Consider a linear mapping $T : V \to V$, with $X \leq V$. If $T(X)$ is a subspace of X then we call X **an invariant subspace under** T.

In other words, X maps back to a subset of itself.

Invariant Transforms Example (MATH2601)

2018 FE

Let V be a vector space, and S, T be linear transforms from V to V. Define $W = \ker(S - T)$. Show that if $ST = TS$ then W is invariant under T.

Invariant Transforms Example (MATH2601)

2018 FE

Let V be a vector space, and S, T be linear transforms from V to V. Define $W = \ker(S - T)$. Show that if $ST = TS$ then W is invariant under T.

Required to prove: $T(W) \subseteq W$.
Consider any $\mathbf{w} \in W$, and now consider $T(\mathbf{w})$. Remember, we want to show that $T(\mathbf{w}) \in W = \ker(S - T)$.

Invariant Transforms Example (MATH2601)

2018 FE

Let V be a vector space, and S, T be linear transforms from V to V. Define $W = \ker(S - T)$. Show that if $ST = TS$ then W is invariant under T.

Required to prove: $T(W) \subseteq W$.
Consider any $\mathbf{w} \in W$, and now consider $T(\mathbf{w})$. Remember, we want to show that $T(\mathbf{w}) \in W = \ker(S - T)$.

$$\begin{aligned}(S - T)(T(\mathbf{w})) &= ST(\mathbf{w}) - T^2(\mathbf{w}) \\ &= TS(\mathbf{w}) - T^2(\mathbf{w}) \qquad \text{(assumption of } TS = ST\text{)} \\ &= T((S - T)(\mathbf{w}))\end{aligned}$$

Invariant Transforms Example (MATH2601)

2018 FE

Let V be a vector space, and S, T be linear transforms from V to V. Define $W = \ker(S - T)$. Show that if $ST = TS$ then W is invariant under T.

Required to prove: $T(W) \subseteq W$.

Consider any $\mathbf{w} \in W$, and now consider $T(\mathbf{w})$. Remember, we want to show that $T(\mathbf{w}) \in W = \ker(S - T)$.

$$\begin{aligned}
(S - T)(T(\mathbf{w})) &= ST(\mathbf{w}) - T^2(\mathbf{w}) \\
&= TS(\mathbf{w}) - T^2(\mathbf{w}) && \text{(assumption of } TS = ST\text{)} \\
&= T((S - T)(\mathbf{w})) \\
&= T(\mathbf{0}) && (\mathbf{w} \in W) \\
&= \mathbf{0}.
\end{aligned}$$

So $T(\mathbf{w}) \in W$, and hence W is invariant under T.

Similarity

Definition: Similarity

Suppose we have two matrices $A, B \in M_{p,p}(\mathbb{F})$ and an invertible $P \in \mathrm{GL}(\mathbb{F}, p)$ such that

$$B = P^{-1}AP.$$

We say that A is **similar to** B.

The idea here is that these matrices **represent the same linear transformation** (just with different bases!). As such, a few properties remain intact. These properties are called similarity invariants.

Similarity

Definition: Similarity

Suppose we have two matrices $A, B \in M_{p,p}(\mathbb{F})$ and an invertible $P \in \mathrm{GL}(\mathbb{F}, p)$ such that

$$B = P^{-1}AP.$$

We say that A is **similar to** B.

The idea here is that these matrices **represent the same linear transformation** (just with different bases!). As such, a few properties remain intact. These properties are called similarity invariants.

- $\mathrm{rank}(A) = \mathrm{rank}(B)$;
- $\mathrm{nullity}(A) = \mathrm{nullity}(B)$;
- $\mathrm{tr}(A) = \mathrm{tr}(B)$;
- $|A| = |B|$

Note, however, that all of these properties being the same does not imply that the matrices are similar.

Similarity Example

2008 FE

Show that

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 5 \\ 3 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -4 & -2 & 0 \\ 2 & 3 & 0 \\ 1 & -5 & 1 \end{pmatrix}$$

are not similar.

Similarity Example

2008 FE

Show that

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 5 \\ 3 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -4 & -2 & 0 \\ 2 & 3 & 0 \\ 1 & -5 & 1 \end{pmatrix}$$

are not similar.

Looking at the similarity invariants, we try to find a mismatch:

- **Trace:** $\operatorname{tr}(A) = 1 + -2 + 1 = 0$ and $\operatorname{tr}(B) = -4 + 3 + 1 = 0$

Similarity Example

2008 FE

Show that

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 5 \\ 3 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -4 & -2 & 0 \\ 2 & 3 & 0 \\ 1 & -5 & 1 \end{pmatrix}$$

are not similar.

Looking at the similarity invariants, we try to find a mismatch:

- **Trace:** $\mathrm{tr}(A) = 1 + -2 + 1 = 0$ and $\mathrm{tr}(B) = -4 + 3 + 1 = 0$
- **Determinant:** $|A| = 1(-2(1) - 5(1)) - 0 + 0 = -7$ and $|B| = 1(-4(3) - -2(2)) = -8$.

Similarity Example

2008 FE

Show that

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 5 \\ 3 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -4 & -2 & 0 \\ 2 & 3 & 0 \\ 1 & -5 & 1 \end{pmatrix}$$

are not similar.

Looking at the similarity invariants, we try to find a mismatch:

- **Trace:** $\operatorname{tr}(A) = 1 + -2 + 1 = 0$ and $\operatorname{tr}(B) = -4 + 3 + 1 = 0$
- **Determinant:** $|A| = 1(-2(1) - 5(1)) - 0 + 0 = -7$ and $|B| = 1(-4(3) - -2(2)) = -8$.

We stop here as the determinants aren't the same, so the matrices **can't be similar to each other**.

Dot Product

Dot Product

Consider any two $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. The dot product is defined as:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Complex Dot Product

Consider any two $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. The complex dot product is defined as:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} \bar{x}_i y_i.$$

If the dot product is 0, we say that $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** to each other.

Inner Product

Definition: Inner Product

Consider a vector space, V alongside its respective field, $\mathbb{F}$. An inner product is a function $\langle \rangle : V \times V \to \mathbb{F}$ with the following properties:

- $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$
- $\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$
- $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ with equality iff $\mathbf{x} = \mathbf{0}$.

Inner Products

The inner product is a generalisation of the dot product. All of the properties of the inner product apply to the dot product.

Orthogonality

Definition: Orthogonal Vectors

Any two vectors $\mathbf{x}, \mathbf{y} \in V$ with $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, are said to be **orthogonal to each other**. This can be denoted as $\mathbf{x} \perp \mathbf{y}$.

Definition: Orthonormal Vectors

Suppose that you have a set of vectors, $S = \{\mathbf{v}_i\}_{i=1}^n$. This set is said to be **orthonormal** if for $i, j = 1, 2, \ldots, n$:

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

We can simplify this condition as $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$. δ_{ij} is the Kronecker delta symbol, equal to 1 if $i = j$ and equal to 0 otherwise.

Projection

Definition: Projection

Suppose that $\mathbf{y}, \mathbf{x} \in V$. We define the **projection** of $\mathbf{x}$ onto $\mathbf{y}$ as:

$$\text{proj}_{\mathbf{y}}(\mathbf{x}) = \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}\mathbf{y}.$$

This formula is defined so that $\mathbf{x} - \text{proj}_{\mathbf{y}}(\mathbf{x}) \perp \mathbf{y}$.

Note that the order of the numerator is actually important for complex inner products! The vector that you are projecting onto the other one must come second. That is, in $\text{proj}_{\mathbf{y}}(\mathbf{x})$, $\mathbf{x}$ must come second.

Gram-Schmidt Process

Theorem: Gram-Schmidt Process

Any finite-dimensional inner product space has an orthonormal basis.

Gram-Schmidt Process

Theorem: Gram-Schmidt Process

Any finite-dimensional inner product space has an orthonormal basis.

A finite-dimensional inner product space must have a finite basis. And, given any finite basis, it turns out that we can apply a set process (the Gram-Schmidt process) to transform it into an orthonormal basis.

Procedure:

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_2) \\
\mathbf{w}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_3) \\
&\vdots \\
\mathbf{w}_n &= \mathbf{v}_n - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_n) - \ldots - \mathrm{proj}_{\mathbf{w}_{n-1}}(\mathbf{v}_n)
\end{aligned}$$

After this, we can just normalise each vector.

Gram-Schmidt Example

2011 FE

Let W be a subspace of $\mathbb{R}^4$ that is spanned by the following vectors: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -3 \\ -1 \\ 4 \\ 1 \end{pmatrix}$. Find an orthonormal basis for W (under the regular dot product).

Gram-Schmidt Example

2011 FE

Let W be a subspace of $\mathbb{R}^4$ that is spanned by the following vectors: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -3 \\ -1 \\ 4 \\ 1 \end{pmatrix}$. Find an orthonormal basis for W (under the regular dot product).

We use the Gram-Schmidt process on $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 = (2, -1, 0, 2)^T \\
\mathbf{w}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_2) \\
&= (3, 0, -2, 0)^T - \frac{\langle (2, -1, 0, 2)^T, (3, 0, -2, 0)^T \rangle}{\|(2, -1, 0, 2)^T\|^2}(2, -1, 0, 2)^T
\end{aligned}$$

Gram-Schmidt Example

2011 FE

Let W be a subspace of $\mathbb{R}^4$ that is spanned by the following vectors: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -3 \\ -1 \\ 4 \\ 1 \end{pmatrix}$. Find an orthonormal basis for W (under the regular dot product).

$$\begin{aligned}
\mathbf{w}_2 &= (3,0,-2,0)^T - \frac{\langle (2,-1,0,2)^T, (3,0,-2,0)^T \rangle}{\|(2,-1,0,2)^T\|^2}(2,-1,0,2)^T \\
&= (3,0,-2,0)^T - \frac{6+0+0+0}{4+1+0+4}(2,-1,0,2)^T \\
&= \frac{1}{3}(5,2,-6,-4)^T.
\end{aligned}$$

Gram-Schmidt Example

2011 FE

Let W be a subspace of $\mathbb{R}^4$ that is spanned by the following vectors: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -3 \\ -1 \\ 4 \\ 1 \end{pmatrix}$. Find an orthonormal basis for W (under the regular dot product).

$$\begin{aligned}
\mathbf{w}_3 &= \mathbf{v}_3 - \operatorname{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \operatorname{proj}_{\mathbf{w}_2}(\mathbf{v}_3) \\
&= (-3, -1, 4, 1)^T - \frac{-3}{9}(2, -1, 0, 2)^T - \frac{-15}{9}(5, 2, -6, -4)^T \\
&= \frac{1}{3}\left[(-9, -3, 12, 3)^T + (2, -1, 0, 2)^T + (25, 10, -30, -20)^T\right] \\
&= (6, 2, -6, -5)^T.
\end{aligned}$$

Gram-Schmidt Example

So an orthogonal basis of W is:

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \ \frac{1}{3}\begin{pmatrix} 5 \\ 2 \\ -6 \\ -4 \end{pmatrix}, \ \begin{pmatrix} 6 \\ 2 \\ -6 \\ -5 \end{pmatrix} \right\}.$$

Gram-Schmidt Example

Our **orthonormal** basis is:

$$\mathcal{B}_3 = \left\{ \frac{1}{3}\begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix},\ \frac{1}{9}\begin{pmatrix} 5 \\ 2 \\ -6 \\ -4 \end{pmatrix},\ \frac{1}{\sqrt{101}}\begin{pmatrix} 6 \\ 2 \\ -6 \\ -5 \end{pmatrix} \right\}.$$

Projections onto Subspaces

Definition: The Orthogonal Complement

$W^\perp$, the orthogonal complement of W, is the space of all vectors $\mathbf{x} \in V$ such that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in W$.

Projections onto Subspaces

Definition: The Orthogonal Complement

$W^\perp$, the orthogonal complement of W, is the space of all vectors $\mathbf{x} \in V$ such that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in W$.

Theorem

For any finite-dimensional vector space V and $W \leq V$, $W \oplus W^\perp = V$.

This means that any vector $\mathbf{x}$ in a finite-dimensional vector space V can be expressed uniquely as $\mathbf{x} = \mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in W$ and $\mathbf{b} \in W^\perp$.

Projections onto Subspaces

Definition: The Orthogonal Complement

$W^\perp$, the orthogonal complement of W, is the space of all vectors $\mathbf{x} \in V$ such that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in W$.

Theorem

For any finite-dimensional vector space V and $W \leq V$, $W \oplus W^\perp = V$.

This means that any vector $\mathbf{x}$ in a finite-dimensional vector space V can be expressed uniquely as $\mathbf{x} = \mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in W$ and $\mathbf{b} \in W^\perp$.

Definitions: Orthogonal Projections

If $\mathbf{x} = \mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in W$ and $\mathbf{b} \in W^\perp$, then $\mathbf{a} = \text{proj}_W \mathbf{x}$, and $\mathbf{b} = \text{proj}_{W^\perp} \mathbf{x}$.

Projections onto Subspaces

Definition: The Orthogonal Complement

$W^\perp$, the orthogonal complement of W, is the space of all vectors $\mathbf{x} \in V$ such that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in W$.

Theorem

For any finite-dimensional vector space V and $W \leq V$, $W \oplus W^\perp = V$.

This means that any vector $\mathbf{x}$ in a finite-dimensional vector space V can be expressed uniquely as $\mathbf{x} = \mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in W$ and $\mathbf{b} \in W^\perp$.

Definitions: Orthogonal Projections

If $\mathbf{x} = \mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in W$ and $\mathbf{b} \in W^\perp$, then $\mathbf{a} = \operatorname{proj}_W \mathbf{x}$, and $\mathbf{b} = \operatorname{proj}_{W^\perp} \mathbf{x}$.

We have the incredibly useful equation $\mathbf{x} - \operatorname{proj}_W \mathbf{x} \perp \mathbf{y}$ for all $\mathbf{y}$ in W.

Projection Example

2017 FE

For the degree-2 polynomial vector space, we consider

$$f(t) = 3 + t^2 \quad g_1(t) = 1 - t \quad g_2(t) = 2 + t - t^2.$$

Find the projection of f onto $W = \text{span}\{g_1, g_2\}$, under the following inner product:

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

Projection Example

2017 FE

For the degree-2 polynomial vector space, we consider

$$f(t) = 3 + t^2 \quad g_1(t) = 1 - t \quad g_2(t) = 2 + t - t^2.$$

Find the projection of f onto $W = \text{span}\{g_1, g_2\}$, under the following inner product:

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

Essentially, we want the vector h in W that satisfies the all-important statement $f - h \in W^{\perp}$. What this means is that $f - h$ will need to be orthogonal to the basis vectors g_1 and g_2. So let's first set $h = \alpha_1 g_1 + \alpha_2 g_2$, and see what we'll need to do to solve.

Projection Example

2017 FE

For the degree-2 polynomial vector space, we consider

$$f(t) = 3 + t^2 \quad g_1(t) = 1 - t \quad g_2(t) = 2 + t - t^2.$$

Find the projection of f onto $W = \text{span}\{g_1, g_2\}$, under the following inner product:

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

So $f - \alpha_1 g_1 - \alpha_2 g_2 \perp g_1$, and $f - \alpha_1 g_1 - \alpha_2 g_2 \perp g_2$. Now all we have to do is convert these two equations to inner products, and that should give us two equations for α_1 and α_2.

$$\langle f - \alpha_1 g_1 - \alpha_2 g_2, g_1 \rangle = 0, \text{ and } \langle f - \alpha_1 g_1 - \alpha_2 g_2, g_2 \rangle = 0.$$

Projection Example

2017 FE

For the degree-2 polynomial vector space, we consider

$$f(t) = 3 + t^2 \quad g_1(t) = 1 - t \quad g_2(t) = 2 + t - t^2.$$

Find the projection of f onto $W = \operatorname{span}\{g_1, g_2\}$, under the following inner product:

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

At this point, we want to expand using linearity of the real inner product, so that we can get the following linear equations:

$$\langle f, g_1 \rangle - \alpha_1 \langle g_1, g_1 \rangle - \alpha_2 \langle g_2, g_1 \rangle = 0, \text{ and } \langle f, g_2 \rangle - \alpha_1 \langle g_1, g_2 \rangle - \alpha_2 \langle g_2, g_2 \rangle = 0.$$

Projection Example

Finding the coefficients:

$$\begin{aligned}
\langle f, g_1 \rangle &= 3(1) + 4(0) + 7(-1) \\
&= -4, \\
\langle g_1, g_1 \rangle &= 1(1) + 0(0) + -1(-1) \\
&= 2, \\
\langle g_2, g_1 \rangle &= 2(1) + 2(0) + 0(-1) \\
&= 2.
\end{aligned}$$

$$\begin{aligned}
\langle f, g_2 \rangle &= 3(2) + 4(2) + 7(0) \\
&= 14, \\
\langle g_1, g_2 \rangle &= \overline{\langle g_2, g_1 \rangle} \\
&= \bar{2} = 2, \\
\langle g_2, g_2 \rangle &= 2(2) + 2(2) + 0(0) \\
&= 8.
\end{aligned}$$

Projection Example

2017 FE

For the degree-2 polynomial vector space, we consider

$$f(t) = 3 + t^2 \quad g_1(t) = 1 - t \quad g_2(t) = 2 + t - t^2.$$

Find the projection of f onto $W = \text{span}\{g_1, g_2\}$, under the following inner product:

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

Subbing all of this in leads to:

$$2\alpha_1 + 2\alpha_2 = -4 \qquad\qquad 2\alpha_1 + 8\alpha_2 = 14.$$

We see that $\alpha_1 = -5, \alpha_2 = 3$. Hence, the projection of f onto W is

$$\text{proj}_W(f) = -5g_1 + 3g_2 = 1 + 8t - 3t^2.$$

QR Factorisation

Theorem: QR Factorisation

Suppose that we have a **full-rank** $p \times q$ matrix A (linearly independent columns). We can represent A as $A = QR$, where Q is a $p \times q$ orthogonal matrix and R is an invertible $q \times q$ upper triangular matrix.

$$A = \begin{pmatrix} \mathbf{q}_1 & | & \mathbf{q}_2 & | & \dots & | & \mathbf{q}_q \end{pmatrix} \begin{pmatrix} \|\mathbf{w}_1\| & \langle \mathbf{q}_1, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{q}_1, \mathbf{a}_q \rangle \\ & \|\mathbf{w}_2\| & \dots & \langle \mathbf{q}_2, \mathbf{a}_q \rangle \\ & & \ddots & \vdots \\ & & & \|\mathbf{w}_q\| \end{pmatrix}$$

where

- $\mathbf{w}_k$ are the vectors found directly by using the Gram-Schmidt process on the columns of A
- $\mathbf{q}_k$ is the corresponding normalised vector
- $\mathbf{a}_k$ are the column vectors of A.

QR Factorisation

Theorem: QR Factorisation

Suppose that we have a **full-rank** $p \times q$ matrix A (linearly independent columns). We can represent A as $A = QR$, where Q is a $p \times q$ orthogonal matrix and R is an invertible $q \times q$ upper triangular matrix.

Making Q square

If A is non-square, then Q will be too. If we want to **make Q a square matrix**, simply extend the Gram-Schmidt process to $\mathbb{F}^p$ and make the constructed vectors the column vectors of Q. For R, simply add $p - q$ extra rows full of zeros.

QR Factorisation Example

2017 FE

Find a QR factorisation of $B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$.

QR Factorisation Example

2017 FE

Find a QR factorisation of $B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$.

As our goal is to find a QR factorisation, we'll start by using the Gram-Schmidt approach on the column vectors.

But instead of normalising all the vectors at the end, we'll normalise our vectors as we go. We'll see that this will make our final step of determining the matrix R a bit easier.

QR Factorisation Example

2017 FE

Find a QR factorisation of $B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$.

$$\begin{aligned} \mathbf{w}_1 &= (1,2,0,2)^T \\ \|\mathbf{w}_1\| &= 3, \text{ so} \\ \mathbf{q}_1 &= \frac{1}{3}(1,2,0,2)^T \end{aligned}$$

QR Factorisation Example

$$
\begin{aligned}
\mathbf{w}_1 &= (1, 2, 0, 2)^T \\
\|\mathbf{w}_1\| &= 3, \text{ so} \\
\mathbf{q}_1 &= \frac{1}{3}(1, 2, 0, 2)^T \\
\mathbf{w}_2 &= (2, 1, 1, 1)^T - \text{proj}_{\mathbf{q}_1}((2, 1, 1, 1)^T) \\
&= (2, 1, 1, 1)^T - \langle \frac{1}{3}(1, 2, 0, 2)^T, (2, 1, 1, 1)^T \rangle \frac{1}{3}(1, 2, 0, 2)^T \\
&= (2, 1, 1, 1)^T - 2 \times \frac{1}{3}(1, 2, 0, 2)^T \\
&= \frac{1}{3}(4, -1, 3, -1)^T
\end{aligned}
$$

QR Factorisation Example

$\|\mathbf{w}_2\| = \sqrt{3}$, so

$$\mathbf{q}_2 = \frac{1}{3\sqrt{3}}(4, -1, 3, -1)^T$$

$$\mathbf{w}_3 = (1, 0, 1, 1)^T - \langle \mathbf{q}_1, (1, 0, 1, 1)^T \rangle \mathbf{q}_1 - \langle \mathbf{q}_2, (1, 0, 1, 1)^T \rangle \mathbf{q}_2$$

QR Factorisation Example

$\|\mathbf{w}_2\| = \sqrt{3}$, so

$$\mathbf{q}_2 = \frac{1}{3\sqrt{3}}(4, -1, 3, -1)^T$$

$$\begin{aligned}\mathbf{w}_3 &= (1,0,1,1)^T - \langle \mathbf{q}_1, (1,0,1,1)^T \rangle \mathbf{q}_1 - \langle \mathbf{q}_2, (1,0,1,1)^T \rangle \mathbf{q}_2 \\ &= (1,0,1,1)^T - 1 \times \frac{1}{3}(1,2,0,2)^T - \frac{2}{\sqrt{3}} \times \frac{1}{3\sqrt{3}}(4,-1,3,-1)^T \\ &= \frac{1}{9}(-2,-4,3,5)^T\end{aligned}$$

$\|\mathbf{w}_3\| = \dfrac{\sqrt{2}}{\sqrt{3}}$, so

$$\mathbf{q}_3 = \frac{1}{3\sqrt{6}}(-2, -4, 3, 5)^T$$

QR Factorisation Example

First, the vectors $\mathbf{q}_1, \mathbf{q}_2$, and $\mathbf{q}_3$ make up the columns of Q.

QR Factorisation Example

First, the vectors $\mathbf{q}_1, \mathbf{q}_2$, and $\mathbf{q}_3$ make up the columns of Q.

$$\mathbf{q}_1 = \frac{1}{3\sqrt{6}}(\sqrt{6}, 2\sqrt{6}, 0, 2\sqrt{6})^T$$
$$\mathbf{q}_2 = \frac{1}{3\sqrt{6}}(4\sqrt{2}, -\sqrt{2}, 3\sqrt{2}, -\sqrt{2})^T$$
$$\mathbf{q}_3 = \frac{1}{3\sqrt{6}}(-2, -4, 3, 5)^T$$

So

$$Q = \frac{1}{3\sqrt{6}}\begin{pmatrix} \sqrt{6} & 4\sqrt{2} & -2 \\ 2\sqrt{6} & -\sqrt{2} & -4 \\ 0 & 3\sqrt{2} & 3 \\ 2\sqrt{6} & -\sqrt{2} & 5 \end{pmatrix}$$

QR Factorisation Example

Now, since we normalised as we went, we'll actually find that we have all the components of R already.

We found $\|\mathbf{w}_1\|$, $\|\mathbf{w}_2\|$, and $\|\mathbf{w}_3\|$ when normalising our vectors.

And we found the inner products with $\mathbf{q}_1$ in the Gram-Schmidt process.

QR Factorisation Example

Now, since we normalised as we went, we'll actually find that we have all the components of R already.

We found $\|\mathbf{w}_1\|$, $\|\mathbf{w}_2\|$, and $\|\mathbf{w}_3\|$ when normalising our vectors.

And we found the inner products with $\mathbf{q}_1$ in the Gram-Schmidt process. For instance, we had

$$\begin{aligned} \mathbf{w}_3 &= (1,0,1,1)^T - \langle \mathbf{q}_1, (1,0,1,1)^T \rangle \mathbf{q}_1 - \langle \mathbf{q}_2, (1,0,1,1)^T \rangle \mathbf{q}_2 \\ &= (1,0,1,1)^T - 1 \times \frac{1}{3}(1,2,0,2)^T - \frac{2}{\sqrt{3}} \times \frac{1}{3\sqrt{3}}(4,-1,3,-1)^T, \end{aligned}$$

giving us

$$R = \begin{pmatrix} 3 & 2 & 1 \\ & \sqrt{3} & \frac{2}{\sqrt{3}} \\ & & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}.$$

QR Factorisation Example

Hence, a QR factorisation of B is

$$B = \frac{1}{3\sqrt{6}} \begin{pmatrix} \sqrt{6} & 4\sqrt{2} & -2 \\ 2\sqrt{6} & -\sqrt{2} & -4 \\ 0 & 3\sqrt{2} & 3 \\ 2\sqrt{6} & -\sqrt{2} & 5 \end{pmatrix} \times \begin{pmatrix} 3 & 2 & 1 \\ & \sqrt{3} & \frac{2}{\sqrt{3}} \\ & & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}.$$

Adjoint (MATH2601)

Theorem: Adjoint Linear Maps

Consider the linear mapping, $T : V \to W$. There exists a unique linear mapping, $T^* : W \to V$ such that:

$$\langle \mathbf{w}, T(\mathbf{v}) \rangle = \langle T^*(\mathbf{w}), \mathbf{v} \rangle$$

for all $\mathbf{v} \in V, \mathbf{w} \in W$. We call any such T^* as the **adjoint of** T.

Some properties of adjoints include:

- $(S + T)^* = S^* + T^*$
- $(\alpha T)^* = \bar{\alpha} T^*$ for any scalar α.
- $(T^*)^* = T$
- If $U : W \to V$ is a linear map, then we have: $(U \circ T)^* = T^* \circ U^*$

Adjoint Example (MATH2601)

2016 FE

Consider the following linear mapping, $T : \mathbb{R}^2 \to \mathbb{P}_1(\mathbb{R})$, defined by:

$$T(x_1, x_2) = (x_1 + x_2) - (2x_2)t.$$

Use the **standard inner product for** $\mathbb{R}^2$, and the following inner product for $\mathbb{P}_1(\mathbb{R})$:

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

Find the adjoint of T with respect to these inner products.

Adjoint Example (MATH2601)

2016 FE

Consider the following linear mapping, $T : \mathbb{R}^2 \to \mathbb{P}_1(\mathbb{R})$, defined by:

$$T(x_1, x_2) = (x_1 + x_2) - (2x_2)t.$$

Use the **standard inner product for** $\mathbb{R}^2$, and the following inner product for $\mathbb{P}_1(\mathbb{R})$:

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

Find the adjoint of T with respect to these inner products.

Consider $\mathbf{x} = (x_1, x_2)$ and $p(t) = p_0 + p_1 t$. We'll first find the inner product of $T(\mathbf{x})$ and p. Remember, T maps to $\mathbb{P}_1(\mathbb{R})$, so we'll use the inner product on $\mathbb{P}_1(\mathbb{R})$.

$$\langle p, T(\mathbf{x}) \rangle = (x_1 + x_2)p_0 + (x_1 - x_2)(p_0 + p_1)$$

Adjoint Example (MATH2601)

$$\langle p, T(\mathbf{x}) \rangle = (x_1 + x_2)p_0 + (x_1 - x_2)(p_0 + p_1)$$

The question we're really answering is: what do we have to do to p such that the dot product of the result with $\mathbf{x}$ is going to be this expression?

Adjoint Example (MATH2601)

$$\langle p, T(\mathbf{x})\rangle = (x_1 + x_2)p_0 + (x_1 - x_2)(p_0 + p_1)$$

The question we're really answering is: what do we have to do to p such that the dot product of the result with $\mathbf{x}$ is going to be this expression?

$$\begin{aligned} &= x_1(p_0 + p_0 + p_1) + x_2(p_0 - p_0 - p_1) \\ &= x_1(2p_0 + p_1) + x_2(-p_1). \end{aligned}$$

So $x_1(2p_0 + p_1) + x_2(-p_1) = \langle T^*(p), \mathbf{x}\rangle$.

Adjoint Example (MATH2601)

$$\langle p, T(\mathbf{x})\rangle = (x_1 + x_2)p_0 + (x_1 - x_2)(p_0 + p_1)$$

The question we're really answering is: what do we have to do to p such that the dot product of the result with $\mathbf{x}$ is going to be this expression?

$$\begin{aligned} &= x_1(p_0 + p_0 + p_1) + x_2(p_0 - p_0 - p_1) \\ &= x_1(2p_0 + p_1) + x_2(-p_1). \end{aligned}$$

So $x_1(2p_0 + p_1) + x_2(-p_1) = \langle T^*(p), \mathbf{x}\rangle$.

Hence, $T^*(p) = (2p_0 + p_1, -p_1)^T$, where $p(t) = p_0 + p_1 t$.

Types of Maps (MATH2601)

Definitions: Unitary and Hermitian Linear Maps

Consider the linear transformation $T : V \to V$. T is

- **unitary** if $T^* = T^{-1}$.
- **Hermitian** if $T^* = T$.

Definition: Isometry

A linear transformation is an isometry (distance-preserving map) if $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v} \in V$.

Types of Maps (MATH2601)

Definitions: Unitary and Hermitian Linear Maps

Consider the linear transformation $T : V \to V$. T is

- **unitary** if $T^* = T^{-1}$.
- **Hermitian** if $T^* = T$.

Definition: Isometry

A linear transformation is an isometry (distance-preserving map) if $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v} \in V$.

Theorem

A linear map is an isometry if and only if it is unitary.

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

We just discussed how to find the adjoint in the general case. However, for linear transformations $T : \mathbb{C}^n \to \mathbb{C}^m$ with the standard inner product, we have a much simpler way of finding the adjoint.

Adjoint of a Linear Transformation $T : \mathbb{C}^n \to \mathbb{C}^m$

Recall that any linear transformation $T : \mathbb{C}^n \to \mathbb{C}^m$ can be expressed as $T(\mathbf{x}) = A\mathbf{x}$ for an $m \times n$ matrix A. Now the adjoint is simply given by $T^* : \mathbb{C}^n \to \mathbb{C}^m$ where $T^*(\mathbf{x}) = \overline{A^T}\mathbf{x}$, $\overline{A^T}$ the conjugate-transpose of A. Because of this relationship, we often just write the matrix $\overline{A^T}$ as A^*.

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

We could calculate the adjoint from first principles, but that would be quite an involved task.

Let's instead express T as the matrix product $T(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x}$.

So the matrix of our transformation is $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

The adjoint matrix is therefore $A^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Now we can move on to determining which properties hold.

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

Unitary? Recall that for a unitary matrix we need $A^* = A^{-1}$. So let's calculate $AA^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. So our linear transformation is indeed **unitary**.

Types of Maps (MATH2601)

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

Unitary? Recall that for a unitary matrix we need $A^* = A^{-1}$. So let's calculate $AA^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. So our linear transformation is indeed **unitary**.

Isometric? From the previous theorem, we know that all unitary transformations are also isometries, so T **is also isometric**.

Types of Maps (MATH2601)

2007 FE

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(w, z) = (-z, w)$, defined with the standard inner product. Is T unitary? Isometric? Hermitian?

Unitary? Recall that for a unitary matrix we need $A^* = A^{-1}$. So let's calculate $AA^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. So our linear transformation is indeed **unitary**.

Isometric? From the previous theorem, we know that all unitary transformations are also isometries, so T **is also isometric**.

Hermitian? For a Hermitian transformation, we would need $A = A^*$. But we saw above that A and A^* are different. So the transformation is **not Hermitian**.

Method of Least Squares

Sometimes we want to solve a system of linear equations $A\mathbf{x} = \mathbf{b}$, but then find that there are fewer unknowns than relations (i.e. fewer columns than rows). Generally, there **won't be an exact solution**. We can only find the **best solution in the least squares sense**.

Theorem: Method of Least Squares

The least squares solution to $A\mathbf{x} = \mathbf{b}$ is a solution to the **normal equations**

$$A^*A\mathbf{x} = A^*\mathbf{b}.$$

For MATH2501, the normal equations are simply

$$A^TA\mathbf{x} = A^T\mathbf{b}.$$

Method of Least Squares Example

Example

Find the parabola that best fits in the least squares sense to the points:

$$(-1, 6), \quad (1, 2), \quad (2, -1), \quad (2, 7).$$

Method of Least Squares Example

Example

Find the parabola that best fits in the least squares sense to the points:

$$(-1, 6), \quad (1, 2), \quad (2, -1), \quad (2, 7).$$

First, write down the system of equations that we're trying to solve. We want the parabola $y = a + bx + cx^2$ that forms the best fit possible to the points given.

So we're trying to solve

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ -1 \\ 7 \end{pmatrix}.$$

Method of Least Squares Example

Example

Find the parabola that best fits in the least squares sense to the points:

$$(-1, 6), \quad (1, 2), \quad (2, -1), \quad (2, 7).$$

Now we can multiply both sides by A^* on the left.

$$A^*A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 2 & 2 \\ 1 & 1 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 2 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 4 & 10 \\ 4 & 10 & 16 \\ 10 & 16 & 34 \end{pmatrix}$$

Method of Least Squares Example

Example

Find the parabola that best fits in the least squares sense to the points:

$$(-1,6), \quad (1,2), \quad (2,-1), \quad (2,7).$$

Now we can multiply both sides by A^* on the left.

$$A^*A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 2 & 2 \\ 1 & 1 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 2 & 4 \end{pmatrix} \qquad A^*\mathbf{b} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 2 & 2 \\ 1 & 1 & 4 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \\ -1 \\ 7 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 4 & 10 \\ 4 & 10 & 16 \\ 10 & 16 & 34 \end{pmatrix} \qquad = \begin{pmatrix} 14 \\ 8 \\ 32 \end{pmatrix}$$

Method of Least Squares Example

Example

Find the parabola that best fits in the least squares sense to the points:

$$(-1, 6), \quad (1, 2), \quad (2, -1), \quad (2, 7).$$

Solving the system of equations yields

$$\left(\begin{array}{ccc|c} 4 & 4 & 10 & 14 \\ 4 & 10 & 16 & 8 \\ 10 & 16 & 34 & 32 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 4 & 4 & 10 & 14 \\ 0 & 6 & 6 & -6 \\ 0 & 6 & 9 & -3 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 4 & 4 & 10 & 14 \\ 0 & 6 & 6 & -6 \\ 0 & 0 & 3 & 3 \end{array}\right).$$

Thus, the least squares solution is

$$y = x^2 - 2x + 3.$$