Linear Independence, Basis, and Dimensions

Making the Abstraction Concrete

A. Havens

Department of Mathematics
University of Massachusetts, Amherst

March 21, 2018

Outline

$\mathbb{F}$-Linear Combinations

Definition

Let V be an $\mathbb{F}$-vector space.

Given a finite collection of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$, and a collection of scalars (not necessarily distinct) $a_1, \ldots, a_k \in \mathbb{F}$, the expression

$$a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k = \sum_{i=1}^{k} a_i\mathbf{v}_i$$

is called an $\mathbb{F}$-linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ with scalar weights $a_1, \ldots a_k$.

It is called *nontrivial* if at least one $a_i \neq 0$, otherwise it is called trivial.

$\mathbb{F}$-Linear Spans

Definition

The $\mathbb{F}$-*linear span* of a finite collection $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ of vectors is the set of all linear combinations of those vectors:

$$\mathrm{Span}_{\mathbb{F}}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} := \left\{ \sum_{i=1}^{k} a_i \mathbf{v}_i \;\middle|\; a_i \in \mathbb{F}, i = 1, \ldots, k \right\}.$$

If $S \subset V$ is an infinite set of vectors, the span is defined to be the set of *finite linear combinations* made from finite collections of vectors in S.

Spans are Subspaces

Recall:

Proposition

Let V be an $\mathbb{F}$-vector space. Given a finite collection of vectors $S \subset V$, the span $\mathrm{Span}(S)$ is a vector subspace of V.

This was proven last time.

Definition of Linear Dependence/Independence

Definition

A collection $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ of vectors in an $\mathbb{F}$-vector space V are called *linearly independent* if and only if the only linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ equal to $\mathbf{0} \in V$ is the trivial linear combination:

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ linearly independent } \iff \left(\sum_{i=1}^{k} a_i \mathbf{v}_i = \mathbf{0} \implies a_1 = \ldots = a_k = 0\right).$$

Otherwise we say that $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is *linearly dependent*.

Criteria for Dependence

Proposition

An ordered collection of vectors $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *with* $\mathbf{v}_1 \neq \mathbf{0}$ *is linearly dependent if and only if there is some* $\mathbf{v}_i \in \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, $i > 1$ *which can be expressed as a linear combination of the preceding vectors vectors* $\mathbf{v}_j$ *for* $j < i$.

If any $\mathbf{v}_i = \mathbf{0}$ *in a collection of vectors, that set is linearly dependent.*

Example

Let $V = \mathbb{R}^n$, and suppose $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{R}^n$ is a collection of $k \leq n$ vectors. Then we have the following proposition:

Proposition

The set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent if and only if the matrix $\mathrm{A} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_k]$ has k pivot positions.

Proof.

Consider the system $\mathrm{A}\mathbf{x} = \mathbf{0}$. If $\mathrm{Nul}\, A \neq \{\mathbf{0}\}$, then there's some nonzero $\mathbf{x} \in \mathbb{R}^k$ such that $\sum_{i=1}^k x_i \mathbf{v}_i = \mathbf{0}$, which implies that $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly dependent. Thus, $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent if and only if $\mathrm{Nul}\, \mathrm{A}$ is trivial, which is true if and only if there k pivot positions. □

Towards Bases and Coordinates

How do we tell if, e.g., a collection of polynomials in $\mathscr{P}_n(\mathbb{R})$ are linearly independent? It would be nice to have an analogue of this result for general collections of vectors in any $\mathbb{F}$-vector space.

If a vector space V is spanned by a finite set, we will have an analogous result; but first we need to define a notion of *basis* and *linear coordinates*.

Using the idea of basis, we'll discover that a *finite dimensional vector* space V can be understood as being structurally equivalent in a precise way to some $\mathbb{F}^n$, which allows us to define coordinates.

With coordinates, we can turn a collection of vectors in V into some matrix whose columns are in $\mathbb{F}^n$. Such matrices can represent linear maps and linear systems, measure linear independence, et cetera.

Finite Dimensional versus Infinite Dimensional

Definition

A vector space V over $\mathbb{F}$ is called *finite dimensional* if and only if there exists a finite collection $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset V$ such that the $\mathbb{F}$-linear span of S is V. If no finite collection of vectors spans V, we say V is *infinite dimensional*.

A key result for our discussion is the following:

Proposition

Any nontrivial finite dimensional $\mathbb{F}$-vector space V contains a linearly independent set $\mathscr{B} \subset V$ such that $\operatorname{Span} \mathscr{B} = V$, and moreover, any other such set $\mathscr{B}' \subset V$ such that $\operatorname{Span} \mathscr{B}' = V$ has the same number of elements as $\mathscr{B}$.

Definition of Basis

We'll want to use the size of such a finite set $\mathscr{B}$ spanning V to define *dimension of V*. The preceding proposition, once proved, guarantees that dimension is well defined. Before proving the proposition, we need some terminology and a lemma.

Definition

Given a vector space V over $\mathbb{F}$, we say that a linearly independent set $\mathscr{B}$ such that $V = \mathrm{Span}_{\mathbb{F}}\mathscr{B}$ is a *basis* of V.
An *ordered basis* is a basis which has a specified order for the vectors, $\mathscr{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$.

Example

The standard basis $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ previously defined is an ordered basis of $\mathbb{R}^n$.

Standard basis of $\mathbb{F}^n$

Example

For any field $\mathbb{F}$, let $\mathbb{F}^n = \underbrace{\mathbb{F} \times \ldots \times \mathbb{F}}_{n \text{ times}}$ be the set of all n-tuples of elements of $\mathbb{F}$. This is naturally an $\mathbb{F}$-vector space with component-wise addition and scaling, analogous to $\mathbb{R}^n$.

The *standard basis* of $\mathbb{F}^n$ is the set $\mathscr{B}_S := (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ consisting of the vectors which are columns of I_n. In particular, for any $\mathbf{x} \in \mathbb{F}^n$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \sum_{i=1}^{n} x_i\mathbf{e}_i .$$

Clearly, the vectors of $\mathscr{B}_S$ are linearly independent since they are columns of the identity matrix.

Example: Bases for spaces of Polynomials

Example

Since any polynomial $p(x) \in \mathscr{P}_n(\mathbb{F})$ can by definition be written in the form

$$p(x) = \sum_{k=0}^{n} a_k x^k\,, \;\; a_0, \ldots, a_n \in \mathbb{F}\,,$$

we see that the monomials $1, x, \ldots, x^n$ span $\mathscr{P}_n(\mathbb{F})$.

But by definition, if a polynomial $p(x)$ is the zero polynomial, all of its coefficients $a_0, \ldots, a_n$ must be zero. This implies that $\{1, x \ldots, x^n\}$ is a linearly independent set, and thus a basis for $\mathscr{P}_n(\mathbb{F})$ as an $\mathbb{F}$ vector space.

Example: Basis of $\mathbb{C}$ as a Real Vector Space

Example

The complex numbers $\mathbb{C} = \{a(1) + b(i) \mid a, b \in \mathbb{R}, i = \sqrt{-1}\}$ can be regarded as a real vector space. Indeed, any complex number is a real linear combination of the real multiplicative unity 1 and the "imaginary unit" $i = \sqrt{-1}$.

Thus $\mathbb{C} = \mathrm{Span}_{\mathbb{R}}\{1, i\}$ is a real vector space. Clearly, 1 and i are independent, so $(1, i)$ gives an ordered basis of $\mathbb{C}$ as a real vector space.

Can you give another basis of $\mathbb{C}$ as a real vector space, e.g., one whose elements are both strictly complex? What's an example of a basis for $\mathbb{C}$ as a complex vector space?

A Lemma

A spanning set S such that $\mathrm{Span}_{\mathbb{F}} S = V$ need not be linearly independent. The key thing about a basis is that it is a spanning set which is linearly independent, and so in a sense has the minimum number of elements needed to build the space V with linear combinations.

The following lemma, which we will use in proving the proposition, captures this idea that a basis is more minimal than a general spanning set might be:

Lemma

If $S \subset V$ is a finite set and $B \subset \mathrm{Span}\, S$ is a linearly independent set, then $|B| \leq |S|$.

Proving the Lemma

Proof.

Let $S = \{\mathbf{v}_1 \ldots \mathbf{v}_m\}$ and suppose $B \subset \operatorname{Span} S$ is a linearly independent set. Choose some finite subset $E \subset B$. Since B is linearly independent, so is E.

Suppose $E = \{\mathbf{u}_1, \ldots \mathbf{u}_k\}$. Since $E \subset \operatorname{Span} S$, there's a linear relation $\mathbf{u}_k = a_1\mathbf{v}_1 + \ldots a_m\mathbf{v}_m$.

Since $\mathbf{u}_k \neq 0$ by linear independence of E, we deduce that at least one $a_j \neq 0$. W e may assume that $a_1 \neq 0$, whence we can write $\mathbf{v}_1$ as a linear combination of $\{\mathbf{u}_k, \mathbf{v}_2 \ldots \mathbf{v}_m\}$.

Observe that by construction $E \subset \operatorname{Span}_{\mathbb{F}}\{\mathbf{u}_k, \mathbf{v}_2 \ldots \mathbf{v}_m\}$. Thus $\mathbf{u}_{k-1} \in \{\mathbf{u}_k, \mathbf{v}_2 \ldots \mathbf{v}_m\}$, and repeating the argument above we have that $\mathbf{v}_2 \in \operatorname{Span}_{\mathbb{F}}\{\mathbf{u}_k, \mathbf{u}_{k-1}, \mathbf{v}_3 \ldots \mathbf{v}_m\}$ and $E \subset \operatorname{Span}_{\mathbb{F}}\{\mathbf{u}_k, \mathbf{u}_{k-1}, \mathbf{v}_3 \ldots \mathbf{v}_m\}$.

Proving the Lemma

Proof (continued).

We can repeat this procedure until either we've used up E, in which case $k \leq m$, or until we run out of elements of S.

If we were to run out of elements of S without running out of elements of E, then since E is in the span of each of the sets we are building, we'd be forced to conclude that there are elements of E which are linear combinations of other elements in E, contradicting the linear independence of E.

Thus, it must be the case that $|E| = k \leq m = |S|$. Finally, since any finite subset $E \subset B$ has no more elements than the finite set S, B is itself finite, and $|B| \leq |S|$, as desired. □

Proving the Proposition

With this lemma we can now prove the proposition.

Proof.

Let V be a nontrivial finite dimensional $\mathbb{F}$-vector space. Observe that because the V is finite dimensional, by definition there exists a subset $S \subset V$ such that $\mathrm{Span}\, S = V$.

If S is linearly independent then we merely have to show that no other linearly independent set has a different number of elements.

On the other hand, if S is linearly dependent, we can remove any vector which is a linear combination of the remaining vectors without altering the span: $\mathrm{Span}_{\mathbb{F}}(S) = \mathrm{Span}_{\mathbb{F}}(S - \{\mathbf{w}\})$ whenever $\mathbf{w} \in \mathrm{Span}_{\mathbb{F}}(S - \{\mathbf{w}\})$.

Proving the Proposition (continued)

Proof (continued).

Indeed, since $\mathrm{Span}_{\mathbb{F}} S$ is the set of $\mathbb{F}$-linear combinations of the vectors in S, if we throw a vector $\mathbf{w}$ out of S which is a linear combination of elements from $S - \{\mathbf{w}\}$, the set $S - \{\mathbf{w}\}$ still contains $\mathbf{w}$ in its span, and hence any other linear combination which potentially involved w can be constructed using only $S - \{\mathbf{w}\}$.

Then since S is finite, we can remove at most finitely many vectors in S without changing the span. Thus, after throwing out finitely many vectors, we have a set $\mathscr{B}$ which is linearly independent, such that $\mathrm{Span}_{\mathbb{F}} \mathscr{B} = \mathrm{Span}_{\mathbb{F}} S = V$.

Proof (continued).

It now remains to show that the size of any linearly independent set $\mathscr{B}'$ which also spans V is the same as that of $\mathscr{B}$.

Suppose $|\mathscr{B}| = n$ and $|\mathscr{B}'| = m$. Since $\mathrm{Span}\,\mathscr{B} = V \supset \mathscr{B}'$ and $\mathscr{B}'$ is linearly independent, we deduce that $m \leq n$ from the lemma.

We similarly conclude that since $\mathrm{Span}\,\mathscr{B}' = V \supset \mathscr{B}$ and $\mathscr{B}$ is linearly independent, $n \leq m$. Thus $m = n$ and we are done. □

Bases: Minimality and Maximality

The proposition we just proved says that any finite dimensional vector space has a basis, and the size of a basis is the minimum for a spanning set and gives a well defined *invariant* of V (meaning it doesn't depend on a choice of basis).

We can thus characterize a basis of a finite dimensional vector space V as follows:

- A basis for V is a spanning set of V of the *minimum size*: every vector in V is a linear combination of the basis elements, and no smaller set spans V.
- A basis for V is a linearly independent set of *maximum size*: if you try to construct any set larger than the basis, there will be a linear dependency relation among the elements.

Defining Dimension

Definition

Given a finite dimensional vector space V over $\mathbb{F}$, the *dimension of* V is the size of any $\mathbb{F}$-basis of V:

$$\dim_{\mathbb{F}} V := |\mathscr{B}|\,,$$

where $V = \operatorname{Span}_{\mathbb{F}}\mathscr{B}$ and $\mathscr{B}$ is linearly independent.

Remark

The subscript $\mathbb{F}$ is necessary at times, since a given set V may have different vector space structures over different fields, and consequently different dimensions. Specifying the field removes ambiguity. However, if the field is understood, the subscript may be dropped.

Example

Example

The complex numbers $\mathbb{C}$, regarded as a real vector space, have a basis with two elements: $\{1, i\}$ as described above, and thus $\dim_{\mathbb{R}} \mathbb{C} = 2$.

But as a $\mathbb{C}$-vector space, a basis choice for $\mathbb{C}$ could be any nonzero complex number, and in particular, $\{1\}$ is a basis of $\mathbb{C}$ as a vector space over $\mathbb{C}$, so $\dim_{\mathbb{C}} \mathbb{C} = 1$.

More generally, $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$ while $\dim_{\mathbb{C}} \mathbb{C}^n = n$.

Note that for any field, $\dim_{\mathbb{F}} \mathbb{F}^n = n$, which is established by considering at the standard basis.

Convention Going Forward

Many of the easiest and most informative examples tend to involve real numbers and concrete computations.

Thus, henceforth (in this and future slide shows) to make notation less cumbersome, we may write $\mathrm{Span}\, S := \mathrm{Span}_{\mathbb{R}} S$ and $\dim S = \dim_{\mathbb{R}} S$ whenever S is a real subset of a vector space V over $\mathbb{R}$. More generally, if S is a subset of an $\mathbb{F}$-vector space for some field $\mathbb{F}$, like $\mathbb{C}$ or $\mathbb{Q}$, $\dim S := \dim_{\mathbb{F}} S$ and $\mathrm{Span}\, S := \mathrm{Span}_{\mathbb{F}} S$ should be understood.

Theorems will still be stated as generally as possible (for vector spaces over an arbitrary field $\mathbb{F}$), but as always, you should try to relate the results to your intuition for vectors in spaces such as $\mathbb{R}^n$.

Example: Bases and Dimensions for Spaces of Matrices

Example

We can give an analogue of the standard basis in the case that our vector space is the space of real $m \times n$ matrices, $\mathbb{R}^{m\times n}$. Define $\mathscr{B}_S = \{\mathrm{E}_{ij} \,|\, 1 \le i \le m, 1 \le j \le n, i, j \in \mathbb{N}\}$ such that E_{ij} is the matrix containing a single 1 in the (i, j)-th entry, and zeros in all other entries.

Then $\mathbb{R}^{m\times n} = \mathrm{Span}_{\mathbb{R}}\mathscr{B}_S$. To show $\mathscr{B}_S$ is a basis we have to check that the matrices are linearly independent.

But if $\mathbf{0}_{m\times n} = \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathrm{E}_{ij}$, then clearly $a_{ij} = 0$ for all indices i and j. Thus, $\mathscr{B}_S$ is a basis of the space of $m \times n$ real matrices $\mathbb{R}^{m\times n}$, and $\dim \mathbb{R}^{m\times n} = mn$.

Example: Dimensions of Spaces Polynomials

Example

As established above, the set of monomials $\{1, x, \dots x^n\}$ are a basis for the space $\mathscr{P}_n(\mathbb{F})$ of polynomials of degree $\leq n$ with coefficients in $\mathbb{F}$.

Thus, $\dim \mathscr{P}_n(\mathbb{F}) = n + 1$.

Consider $\mathscr{P}(\mathbb{F})$, the space of all polynomials with coefficients in $\mathbb{F}$. Can we describe a basis? What is its dimension?

Example: Dimensions of Spaces Polynomials

Example

No finite basis can exist for $\mathscr{P}(\mathbb{F})$, since, given any finite list of polynomials, there is a maximal degree, while there are elements of $\mathscr{P}(\mathbb{F})$ of arbitrarily large degree.

Thus $\mathscr{P}(\mathbb{F})$ is not finite dimensional.

However, since *any given polynomial* is a finite linear combination of monomials of bounded degree, $\mathscr{P}(\mathbb{F}) = \mathrm{Span}_{\mathbb{F}}\{x^k \mid k \in \mathbb{Z}_{\geq 0}\}$. Since the only way to make the zero polynomial with monomials is to take all coefficients to be zero, $\{x^k \mid k \in \mathbb{Z}_{\geq 0}\}$ is a linearly independent set.

Since $\{x^k \mid k \in \mathbb{Z}_{\geq 0}\}$ is a linearly independent set spanning $\mathscr{P}(\mathbb{F})$, it provides a basis of $\mathscr{P}(\mathbb{F})$.

Basis of a Subspace

Since a subspace is itself a vector space, the proposition we proved implies we can find some subset $\mathscr{E} \subset W$ of any nontrivial subspace W of an $\mathbb{F}$-vector space such that $W = \mathrm{Span}_{\mathbb{F}}(\mathscr{E})$ and $\mathscr{E}$ is linearly independent.

One then says $\mathscr{E}$ is a basis of the subspace W.

Given a non-trivial subspace W of an $\mathbb{F}$-vector space with basis $\mathscr{E}$, one can always extend $\mathscr{E}$ to a basis of V.

Example: Dimensions of Subspaces in $\mathbb{R}^2$ and $\mathbb{R}^3$

Example

A line in $\mathbb{R}^2$ through $\mathbf{0}$ is a proper subspace of $\mathbb{R}^2$, and can be described as the real span of a single nonzero vector $\mathbf{v}$.

Thus lines in $\mathbb{R}^2$ are one dimensional subspaces.

To extend to a basis of $\mathbb{R}^2$, we need just one additional vector, not parallel to $\mathbf{v}$.

Example

Recall that a plane Π through $\mathbf{0}$ in $\mathbb{R}^3$ is a subset of $\mathbb{R}^3$ which is the solution space of a homogeneous equation of the form

$$ax_1 + bx_1 + cx_3 = 0\,.$$

We can express this as a span of two vectors in many ways.

Example

For example

$$\Pi := \mathrm{Span}_{\mathbb{R}} \left\{ \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}, \begin{bmatrix} c \\ 0 \\ -a \end{bmatrix} \right\} = \mathrm{Span}_{\mathbb{R}} \left\{ \begin{bmatrix} -c \\ 0 \\ a \end{bmatrix}, \begin{bmatrix} 0 \\ c \\ -b \end{bmatrix} \right\}.$$

Thus $\{b\mathbf{e}_1 - a\mathbf{e}_2, c\mathbf{e}_1 - a\mathbf{e}_3\}$ and $\{a\mathbf{e}_3 - c\mathbf{e}_1, c\mathbf{e}_2 - b\mathbf{e}_3\}$ are both bases of Π, and $\dim \Pi = 2$.

To extend to a basis of $\mathbb{R}^3$ it suffices to add any vector not in Π, since $\dim \mathbb{R}^3 = 3 = 1 + \dim \Pi$. For example, we can add in a normal vector to the plane, such as $\mathbf{n} = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3$.

Rank: The Dimension of the Column Space

For linear maps and for matrices, we have special names for the dimensions of associated subspaces.

Definition

The *rank of a linear map* $T : V \to W$ between finite dimensional $\mathbb{F}$-vector spaces V and W is the dimension of the image:

$$\mathrm{rank}\, T = \dim_{\mathbb{F}} T(V)\,.$$

Given a matrix $\mathrm{A} \in \mathbb{F}^{m\times n}$, the *rank of* A is the dimension of the column space of A:

$$\mathrm{rank}\,\mathrm{A} = \dim_{\mathbb{F}} \mathrm{Col}\,\mathrm{A}\,.$$

Nullity: The Dimension of the Null Space

Definition

The *nullity of a linear map* $T : V \to W$ between finite dimensional $\mathbb{F}$-vector spaces V and W is the dimension of the kernel:

$$\mathrm{null}\, T = \dim_{\mathbb{F}} \ker T\,.$$

Given a matrix $\mathrm{A} \in \mathbb{F}^{m\times n}$, the *nullity of* A is the dimension of the null space of A:

$$\mathrm{null}\, \mathrm{A} = \dim_{\mathbb{F}} \mathrm{Nul}\, \mathrm{A}\,.$$

Computing Rank and Nullity

Remark

One can compute the rank of a matrix A by determining a linearly independent subset of the columns of A which span $\mathrm{Col}\,A$. The pivot columns of A are precisely such a collection; they give a basis of $\mathrm{Col}\,A$! Thus the number of pivot positions of A is the rank.

One can compute the nullity of a matrix A by determining the number of non-pivot columns. The non-pivot columns of A determine the number of free variables for the homogeneous equation, and thus, the number of vectors needed to build a basis of $\mathrm{Nul}\,A$.

We will soon prove this carefully, and its generalization for any linear map between finite dimensional vector spaces.

Matrices Encoding $\mathbb{R}$-Linear Maps

Recall, to any linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ we can uniquely associate a matrix A such that $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$, and

$$\mathrm{A} = \left[\begin{array}{ccc} T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n) \end{array}\right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_n$ are the standard basis elements of the ordered basis $\mathscr{B}_S$.

Conversely, given a matrix $\mathrm{A} \in \mathbb{R}^{m\times n}$, there is a uniquely determined linear map T whose image is the span of the columns of A.

To generalize this, we might seek a method of describing linear maps by determining how they affect a basis. Then we can try to associate matrices to linear maps $T : V \to W$ given a basis of V.

An Analogue

Given a finite dimensional $\mathbb{F}$-vector space V with a basis $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, and a linear transformation $T : V \to W$ for W another $\mathbb{F}$-vector space, we wish to understand what the set $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ tells us about T.

The first observation is that the image $T(V)$ is spanned by these vectors. Indeed, if $\mathbf{w} \in T(V)$, then there is some $\mathbf{u} \in V$ such that $\mathbf{w} = T(\mathbf{u})$, and $\mathbf{u}$ can be written in terms of the basis $\mathscr{B}$: if $\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$ then
$\mathbf{w} = T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \ldots c_n T(\mathbf{v}_n)$.
This establishes that
$T(V) = \mathrm{Span}_{\mathbb{F}} T(\mathscr{B}) = \mathrm{Span}_{\mathbb{F}}\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$.
We can say better: given the basis $\mathscr{B}$ of V, the map $T : V \to W$ is uniquely determined by the images.

Linear Maps are determined by their effect on a basis

Theorem

Given an n-dimensional $\mathbb{F}$-vector space V with basis $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, and given a collection of n not necessarily distinct vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W$, there is a unique linear map $T : V \to W$ such that $\mathbf{w}_i = T(\mathbf{v}_i)$, $i = 1, \ldots, n$. Moreover, any linear map $T : V \to W$ can be built in this way.

Proof.

For any $\mathbf{u} \in V$ we can write $\mathbf{u}$ as a linear combination of the basis vectors: $\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$.

Suppose $T : V \to W$ is a linear map which satisfies the conditions $T(\mathbf{v}_1) = \mathbf{w}_1, \ldots, T(\mathbf{v}_n) = \mathbf{w}_n$. Then the claim is that the value $T(\mathbf{u})$ is determined uniquely.

Extending by linearity

Proof.

Indeed, since T is linear, one has

$$T(\mathbf{u}) = T\left(\sum_{i=1}^n c_i\mathbf{v}_i\right) = \sum_{i=1}^n c_i T(\mathbf{v}_i) = \sum_{i=1}^n c_i\mathbf{w}_i\,.$$

Moreover, we may construct a unique T from the data $T(\mathbf{v}_i) = \mathbf{w}_i$ by the above formula, and define this to be the *linear extension* of the map on the basis.

Finally, given an arbitrary map $\tilde{T} : V \to W$, we can restrict our attention to its affect on $\mathscr{B}$, defining $\mathbf{w}_i := \tilde{T}(\mathbf{v}_i)$. Then the image $\tilde{T}(\mathbf{u})$ of any $\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$ satisfies
$\tilde{T}(\mathbf{u}) = c_1\mathbf{w}_1 + \ldots + c_n\mathbf{w}_n$, which is just the linear extension of the map on basis elements $\mathbf{v}_i \mapsto \mathbf{w}_i = \tilde{T}(\mathbf{v}_i)$. □

Definition of a Linear Isomorphism

Definition

A *linear isomorphism* from an $\mathbb{F}$-vector space V to an $\mathbb{F}$-vector space W is a bijective linear map $\Phi : V \to W$. In particular, Φ is an invertible linear map from V to W.
If there is an isomorphism $\Phi : V \to W$, then $\Phi^{-1} : W \to V$ is also an isomorphism. Will will thus say things such as "V is isomorphic to W", "W is isomorphic to V", or "V and W are isomorphic," and write $V \cong W$.

Isomorphisms preserve dimension; they can be viewed as maps realizing the structural equivalence of a pair of vector spaces. Isomorphisms are an example of an equivalence relation.

We can apply the above theorem on determining linear maps to understand how isomorphisms arise from transformations of bases.

Theorem on Building Isomorphisms

Theorem

Let V and W be finite dimensional $\mathbb{F}$-vector spaces, and suppose $\mathscr{B}_V$ is a basis of V and $\mathscr{B}_W$ is a basis of W. Then

- *if V and W are isomorphic, then they have the same dimension, and so $|\mathscr{B}_V| = |\mathscr{B}_W|$, and*
- *if $|\mathscr{B}_V| = |\mathscr{B}_W|$, then there exists an isomorphism $\Phi : V \to W$ which is induced by linear extension of a bijective map $\hat{\Phi} : \mathscr{B}_V \to \mathscr{B}_W$.*

Proof.

This is left as an exercise (challenge problem). The main point is to use the definitions of dimension and isomorphism, and then apply the theorem on using linear extension to determine a linear map. □

Endomorphisms and Automorphisms

Definition

A *linear endomorphism* of an $\mathbb{F}$-vector space V is a linear map $T : V \to V$. An endomorphism which is an isomorphism is called a *vector space automorphism*. Thus a linear automorphism is just an invertible linear map $T : V \to V$.

Example: Linear Automorphisms of $\mathbb{R}^n$

If A is an invertible $n \times n$ matrix, then the map $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ is an automorphism.

Since the columns of A are the images of the standard basis vectors, by the theorem on constructing isomorphisms, we can conclude that the columns of A form a basis of $\mathbb{R}^n$.

We also could conclude this via the invertible matrix theorem: we know that for A to be invertible, the columns must span $\mathbb{R}^n$ and must be linearly independent.

Conversely, any linear automorphism $T : \mathbb{R}^n \to \mathbb{R}^n$ is associated to an invertible matrix $\mathrm{A} = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{array}\right]$, and can be viewed as a rising from a correspondence between elements of the standard basis and elements of a new basis $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$.

Examples: Polynomial Spaces and Space of Matrices

Example

Let $\mathbb{P}_n := \mathscr{P}_n(\mathbb{R}) = \mathrm{Span}_{\mathbb{R}}\{1, t, t^2, \dots t^n\}$ be the space of real polynomials of degree $\leq n$. There is a one-to-one correspondence between monomials $\{1, t, \dots, t^n\}$ and elements $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \mathbf{e}_{n+1}$ of the standard basis of $\mathbb{R}^{n+1}$. Observe that the correspondence is not unique.

Thus $\mathbb{P}_n$ is isomorphic to $\mathbb{R}^{n+1}$.

Exercise

Can you identify an isomorphism from $\mathbb{R}^{m\times n}$ to $\mathbb{R}^{mn}$?

$\mathscr{B}$-Coordinates

Definition

Let V be an n-dimensional $\mathbb{F}$-vector space, and let $\mathscr{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an ordered basis of V. Then a given $\mathbf{x}$ can be written as a linear combination $\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$. The vector

$$[x]_{\mathscr{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$$

is called the *$\mathscr{B}$-coordinate vector for* $\mathbf{x}$, and the numbers $c_1, \ldots c_n$ are called the *coordinates of* $\mathbf{x}$ *relative to the basis* $\mathscr{B}$.

Remark

The $\mathscr{B}$-coordinate vector $[\mathbf{x}]_{\mathscr{B}}$ of $\mathbf{x} \in V$ is the image of $\mathbf{x}$ under the isomorphism determined by mapping the ordered basis $\mathscr{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ of V onto the standard ordered basis $\mathscr{B}_S = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ of $\mathbb{F}^n$. This isomorphism sending $\mathbf{x} \mapsto [\mathbf{x}]_{\mathscr{B}}$ is called the *coordinate map for the basis* $\mathscr{B}$, or the *$\mathscr{B}$-coordinate map/isomorphism*.

We will see momentarily that we can compute coordinate vectors of real vectors $\mathbf{x}$ relative to a given basis by solving a system of equations. In this case, the coordinate isomorphism arises as an inverse to an automorphism determined by an invertible matrix describing a map of the standard basis onto a new basis.

Example: Change of basis of $\mathbb{R}^2$

Example

Let $\mathbf{b}_1 = \mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b}_2 = \mathbf{e}_1 + 3\mathbf{e}_2$. Then note that $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2)$ is a basis of $\mathbb{R}^2$. Find the $\mathscr{B}$-coordinate vector of $\mathbf{x} = 8\mathbf{e}_2 - 7\mathbf{e}_1$.

Example: Change of basis of $\mathbb{R}^2$

Example

Let $\mathbf{b}_1 = \mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b}_2 = \mathbf{e}_1 + 3\mathbf{e}_2$. Then note that $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2)$ is a basis of $\mathbb{R}^2$. Find the $\mathscr{B}$-coordinate vector of $\mathbf{x} = 8\mathbf{e}_2 - 7\mathbf{e}_1$.

Solution: The coordinate vector $[x]_{\mathscr{B}} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2$ has components that satisfy $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2$. We can thus write a system:

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} [x]_{\mathscr{B}} = \mathrm{P}_{\mathscr{B}}[x]_{\mathscr{B}} = \mathbf{x}\,.$$

$$\left[\begin{array}{cc|c} 1 & 1 & -7 \\ 2 & 3 & 8 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -29 \\ 0 & 1 & 22 \end{array}\right].$$

Thus the $\mathscr{B}$-coordinate vector of $\mathbf{x}$ is

$$[x]_{\mathscr{B}} = \begin{bmatrix} -29 \\ 22 \end{bmatrix}.$$

Coordinate Transformations

Let $\mathscr{B}$ be a basis for $\mathbb{F}^n$. The equation $\mathrm{P}_{\mathscr{B}}[x]_{\mathscr{B}} = \mathbf{x}$ implies that $[x]_{\mathscr{B}} = \mathrm{P}_{\mathscr{B}}^{-1}\mathbf{x}$. The matrix $\mathrm{P}_{\mathscr{B}}$ is the *change of basis matrix* for the basis change $\mathscr{B} \mapsto \mathscr{B}_S$ where $\mathscr{B}_S = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is the standard basis, and $\mathrm{P}_{\mathscr{B}}^{-1}$ is the matrix for the change of basis $\mathscr{B}_S \mapsto \mathscr{B}$.

Changing coordinates is one of the few situations where one might actually use the inverse of a matrix rather than reduction: if only a few vector's coordinates are needed, then it makes sense to just use row operations, but if one may need to routinely switch between coordinates for many vectors, it may be simpler to compute and store the change of basis matrices.

Coordinates for Subspaces

If $W \subseteq V$ is a subspace with basis $\mathscr{B}_W = (\mathbf{w}_1, \ldots, \mathbf{w}_k)$, then one can describe elements of W by coordinate vectors in $\mathbb{F}^k$.

Example

Let Π be the plane with equation $2x_1 - x_2 + 3x_3 = 0$. Find a basis $\mathscr{B}$ for Π and express the point $P(-4, 10, 6)$ as a coordinate vector $[\mathbf{p}]_{\mathscr{B}}$ relative to this basis.

Coordinates for Subspaces

If $W \subseteq V$ is a subspace with basis $\mathscr{B}_W = (\mathbf{w}_1, \ldots, \mathbf{w}_k)$, then one can describe elements of W by coordinate vectors in $\mathbb{F}^k$.

Example

Let Π be the plane with equation $2x_1 - x_2 + 3x_3 = 0$. Find a basis $\mathscr{B}$ for Π and express the point $P(-4, 10, 6)$ as a coordinate vector $[\mathbf{p}]_{\mathscr{B}}$ relative to this basis.

Solution: The plane Π is given by a single homogeneous equation in three variables, and consequently has two free variables. Writing $2s = x_2$ and $2t = x_3$ (so as to avoid fractions), and solving for x_1, a solution vector in Π can be written as

$$\mathbf{x} = (s - 3t)\mathbf{e}_1 + 2s\mathbf{e}_2 + 2t\mathbf{e}_3 = s(\mathbf{e}_1 + 2\mathbf{e}_2) + t(-3\mathbf{e}_1 + 2\mathbf{e}_3).$$

Thus, $\Pi = \mathrm{Span}\,\{\mathbf{e}_1 + 2\mathbf{e}_2, -3\mathbf{e}_1 + 2\mathbf{e}_3\}$.

Coordinates for Subspaces

Example

To find the coordinates of the given point, we solve the system

$$\begin{bmatrix} -4 \\ 10 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$ which is equivalent

to the augmented matrix $\left[\begin{array}{cc|c} 1 & -3 & -4 \\ 2 & 0 & 10 \\ 0 & 2 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right].$

Thus that the coordinate vector for the point $P(-4, 10, 6)$ relative to the basis $\mathscr{B} = (\mathbf{e}_1 + 2\mathbf{e}_2, -3\mathbf{e}_1 + 2\mathbf{e}_3)$ is

$$[\mathbf{p}]_{\mathscr{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Checking Linear Independence/Dependence

Proposition

Let $\mathscr{B} = \{\mathbf{v}_1, \dots \mathbf{v}_n\}$ be any basis of a vector space V, and suppose $\{\mathbf{u}_1, \dots \mathbf{u}_k\} \subset V$ are a collection of $k \leq n$ vectors. Then $\{\mathbf{u}_1, \dots \mathbf{u}_k\}$ is a linearly independent set if and only if the matrix

$$\begin{bmatrix} [\mathbf{u}_1]_{\mathscr{B}} & \dots & [\mathbf{u}_k]_{\mathscr{B}} \end{bmatrix}$$

has k pivot columns.

Proof.

We prove the contrapositive: that $\{\mathbf{u}_1, \dots \mathbf{u}_k\}$ is linearly dependent if and only if there are fewer than k pivot columns in the matrix $\begin{bmatrix} [\mathbf{u}_1]_{\mathscr{B}} & \dots & [\mathbf{u}_k]_{\mathscr{B}} \end{bmatrix}$.

Proof Using Contrapositive: Lifting Dependence Relations

Proof.

Consider a linear combination building the zero vector: $x_1[\mathbf{u}_1]_{\mathscr{B}} + \ldots + x_k[\mathbf{u}_k]_{\mathscr{B}} = \mathbf{0}$. Applying the inverse coordinate map $\Phi_{\mathscr{B}}^{-1} : \mathbb{F}^n \to V$ to the linear combination $x_1[\mathbf{u}_1]_{\mathscr{B}} + \ldots + x_k[\mathbf{u}_k]_{\mathscr{B}}$:

$$\Phi_{\mathscr{B}}^{-1}(x_1[\mathbf{u}_1]_{\mathscr{B}} + \ldots + x_k[\mathbf{u}_k]_{\mathscr{B}}) = x_1\mathbf{u}_1 + \ldots x_k\mathbf{u}_k = \Phi_{\mathscr{B}}^{-1}(\mathbf{0}) = \mathbf{0}\,.$$

The set $\{\mathbf{u}_1, \ldots \mathbf{u}_k\}$ is linearly dependent if and only if at least one x_i in the above combination is nonzero, which is true if and only if the original linear combination gives a nontrivial $\mathbf{x} = \sum_{i=1}^k x_i\mathbf{e}_i$ such that

$$\left[\ [\mathbf{u}_1]_{\mathscr{B}} \quad \ldots \quad [\mathbf{u}_k]_{\mathscr{B}}\ \right]\mathbf{x} = x_1[\mathbf{u}_1]_{\mathscr{B}} + \ldots + x_k[\mathbf{u}_k]_{\mathscr{B}} = \mathbf{0},$$

which is true if and only if the null space is nontrivial, which is true if and only if there are fewer than k pivot columns. □

Example Checking Linear Independence

Example

Let $\mathbb{P}_2 = \{a_0 + a_1 t + a_2 t^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}$. Check that the polynomials $1 + 2x$, $2 - 3x^2$, $-1 - x + x^2$ are linearly independent in $\mathbb{P}_2$.

Solution: Let $\mathscr{B} = \{1, t, t^2\}$ be the monomial basis for $\mathbb{P}_2$. Then there is a coordinate isomorphism $\varphi_2 : \mathbb{P}_2 \to \mathbb{R}^3$ sending a polynomial $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$ to the coordinate vector

$$[\mathbf{p}(t)]_{\mathscr{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}.$$

Example Checking Linear Independence

Example

Then to the set of polynomials given we can use φ_2 to associate a matrix

$$\left[\ [1+2t]_{\mathscr{B}}\quad [2-3t^2]_{\mathscr{B}}\quad [-1-t+t^2]_{\mathscr{B}}\ \right] = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & -1 \\ 0 & -3 & 1 \end{bmatrix}.$$

Then since $\mathrm{RREF}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & -1 \\ 0 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, we conclude that the polynomials $1+2t$, $2-3t^2$, $-1-t+t^2$ are linearly independent in $\mathbb{P}_2$

Had they been dependent, we could use the inverse coordinate map to lift a linear dependency from $\mathbb{R}^3$ to $\mathbb{P}_2$.

Expressing Linear Maps in Coordinates with Matrices

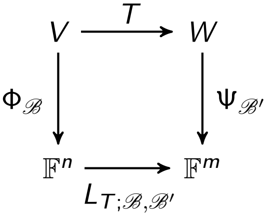
Coordinates allow us to express linear maps between general finite dimensional vector spaces using matrices.

Let V be be an n-dimensional vector space with basis $\mathscr{B}$ and $\mathscr{B}$-coordinate map $\Phi_{\mathscr{B}}$, and let W be an m-dimensional vector space with basis $\mathscr{B}'$ and $\mathscr{B}'$-coordinate map $\Psi_{\mathscr{B}'}$. Thus, $\Phi_{\mathscr{B}}$ and $\Psi_{\mathscr{B}'}$ give isomorphisms $V \cong \mathbb{F}^n$ and $W \cong \mathbb{F}^m$.

Suppose you wanted to express a linear map $T : V \to W$ using a matrix with coefficients in $\mathbb{F}$. That is, you want to understand how to go from a coordinate vector $[\mathbf{x}]_{\mathscr{B}} \in \mathbb{F}^n$ representing $\mathbf{x} \in V$ to the coordinate vector $[T(\mathbf{x})]_{\mathscr{B}'}$ in $\mathbb{F}^m$ representing its image $T(\mathbf{x})$ in W, using a matrix-vector product.

Expressing Linear Maps in Coordinates With Matrices

Let $L_{T;\mathscr{B},\mathscr{B}'} : \mathbb{F}^n \to \mathbb{F}^m$ be the linear transformation giving the coordinate map $[\mathbf{x}]_{\mathscr{B}} \mapsto [T(\mathbf{x})]_{\mathscr{B}'}$. The diagram below shows how the maps T, $\Phi_{\mathscr{B}}$, $\Psi_{\mathscr{B}'}$ and $L_{T;\mathscr{B},\mathscr{B}'}$ fit together into what is called a *commutative square diagram*.



Here, *commutative* means that the different ways of getting from the top left to the bottom right yield the same image: $\Psi_{\mathscr{B}'} \circ T = L_{T;\mathscr{B},\mathscr{B}'} \circ \Phi_{\mathscr{B}}$.

Expressing Linear Maps in Coordinates With Matrices

Since $\Psi_{\mathscr{B}'} \circ T = L_{T;\mathscr{B},\mathscr{B}'} \circ \Phi_{\mathscr{B}}$ and $\Phi_{\mathscr{B}}$ is an isomorphism, and thus invertible, we can solve for $L_{T;\mathscr{B},\mathscr{B}'}$ in terms of T, $\Psi_{\mathscr{B}'}$ and $\Phi_{\mathscr{B}}^{-1}$: $L_{T;\mathscr{B},\mathscr{B}'} = \Psi_{\mathscr{B}'} \circ T \circ \Phi_{\mathscr{B}}^{-1}$.

We thus seek a matrix $\mathrm{P}_{\mathscr{B},\mathscr{B}'} \in \mathbb{F}^{m\times n}$ representing the map $\Psi_{\mathscr{B}'} \circ T \circ \Phi_{\mathscr{B}}^{-1} : \mathbb{F}^n \to \mathbb{F}^m$, and thus representing the effect of T in coordinates.

The j-th column of $\mathrm{P}_{\mathscr{B},\mathscr{B}'}$ is $L_{T;\mathscr{B},\mathscr{B}'}(\mathbf{e}_j)$, which equivalently corresponds to computing the $\mathscr{B}'$-coordinate representations of the images $T(\mathbf{v}_i)$ of basis vectors $\mathbf{v}_i \in \mathscr{B}$.

The $(\mathscr{B}, \mathscr{B}')$-Matrix of a linear Map

Theorem

Let V be be an n-dimensional vector space with basis $\mathscr{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ and let W be an m-dimensional vector space with basis $\mathscr{B}' = (\mathbf{w}_1, \dots, \mathbf{w}_m)$. Given a linear map $T : V \to W$, the matrix $\mathrm{P}_{\mathscr{B},\mathscr{B}'}$ of the associated coordinate transformation $[\mathbf{x}]_{\mathscr{B}} \mapsto [T(\mathbf{x})]_{\mathscr{B}'}$ is uniquely determined and has the form

$$\mathrm{P}_{\mathscr{B},\mathscr{B}'} = \left[\ [T(\mathbf{v}_1)]_{\mathscr{B}'} \quad \dots \quad [T(\mathbf{v}_n)]_{\mathscr{B}'}\ \right].$$

Proof.

Let $L_{T;\mathscr{B},\mathscr{B}'} : \mathbb{F}^n \to \mathbb{F}^m$ be the map such that $L_{T;\mathscr{B},\mathscr{B}'}([\mathbf{x}]_{\mathscr{B}}) = [T(\mathbf{x})]_{\mathscr{B}'}$. It is easy to check that this is a linear map. Let $\mathrm{P}_{\mathscr{B},\mathscr{B}'}$ be the matrix representing it. It suffices to compute $L_{T;\mathscr{B},\mathscr{B}'}(\mathbf{e}_j)$ to obtain the columns of $\mathrm{P}_{\mathscr{B},\mathscr{B}'}$.

Let $\Phi_{\mathscr{B}} : V \to \mathbb{F}^n$ be the $\mathscr{B}$-coordinate map and $\Psi_{\mathscr{B}'} : W \to \mathbb{F}^n$ be the $\mathscr{B}'$-coordinate map. Then since $\Phi_{\mathscr{B}}(\mathbf{v}_j) = \mathbf{e}_j$,

$$\begin{aligned} L_{T;\mathscr{B},\mathscr{B}'}(\mathbf{e}_j) &= \Psi_{\mathscr{B}'} \circ T \circ \Phi_{\mathscr{B}}^{-1}(\mathbf{e}_j) = \Psi_{\mathscr{B}'} \circ T \circ \Phi_{\mathscr{B}}^{-1}\big(\Phi_{\mathscr{B}}(\mathbf{v}_j)\big) \\ &= \Psi_{\mathscr{B}'}\big(T(\mathbf{v}_j)\big) \\ &= [T(\mathbf{v}_j)]_{\mathscr{B}'} \end{aligned}$$

□

Example: The Derivative of a Quadratic Polynomial

We will illustrate this with a simple example: computing a matrix representing the derivative map acting on real degree two polynomials $\mathbb{P}_2$, giving real degree one polynomials in $\mathbb{P}_1$.

Example

Construct a matrix representing the linear map

$$\frac{\mathrm{d}}{\mathrm{d}t} : \mathbb{P}_2 \to \mathbb{P}_1 \,.$$

Solution: We can exploit the coordinate isomorphisms $\varphi_2 : \mathbb{P}_2 \to \mathbb{R}^3$, $\varphi_2(a_0 + a_1 t + a_2 t^2) = a_0\mathbf{e}_1 + a_1\mathbf{e}_2 + a_2\mathbf{e}_3 \in \mathbb{R}^3$, $\varphi_1 : \mathbb{P}_1 \to \mathbb{R}^2$, $\varphi_1(a_0 + a_1 t) = a_0\mathbf{e}_1 + a_1\mathbf{e}_2 \in \mathbb{R}^2$ arising from using the monomial bases of $\mathbb{P}_2$ and $\mathbb{P}_1$.

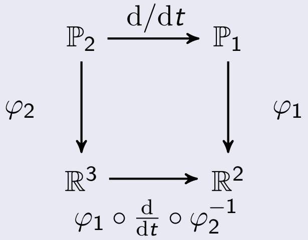
Example: The Derivative of a Quadratic Polynomial

Example

The matrix we desire will actually then be the standard matrix of the map

$$\varphi_1 \circ \frac{\mathrm{d}}{\mathrm{d}t} \circ \varphi_2^{-1} : \mathbb{R}^3 \to \mathbb{R}^2$$

which completes the diagram shown below:



Example: The Derivative of a Quadratic Polynomial

Example

Since $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}(t) = v\left(a_0 + a_1 t + a_2 t^2\right) = a_1 + 2a_2 t$, the bottom map in the diagram representing the coordinate transformation associated to the derivative is
$\left(\varphi_1 \circ \frac{\mathrm{d}}{\mathrm{d}t} \circ \varphi_2^{-1}\right)\left(a_0\mathbf{e}_1 + a_1\mathbf{e}_2 + a_2\mathbf{e}_3\right) = a_1\mathbf{e}_1 + 2a_2\mathbf{e}_2$.
In particular,

$$\left(\varphi_1 \circ \frac{\mathrm{d}}{\mathrm{d}t} \circ \varphi_2^{-1}\right)(\mathbf{e}_1) = \mathbf{0}$$
$$\left(\varphi_1 \circ \frac{\mathrm{d}}{\mathrm{d}t} \circ \varphi_2^{-1}\right)(\mathbf{e}_2) = \mathbf{e}_1$$
$$\left(\varphi_1 \circ \frac{\mathrm{d}}{\mathrm{d}t} \circ \varphi_2^{-1}\right)(\mathbf{e}_3) = 2\mathbf{e}_2$$

Example: The Derivative of a Quadratic Polynomial

Example

It follows that the matrix representing the derivative map $\frac{\mathrm{d}}{\mathrm{d}t} : \mathbb{P}_2 \to \mathbb{P}_1$ with respect to the standard monomial bases is

$$\mathrm{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Observe that the first column is a zero column, and this is entirely sensible since the derivative of a constant is 0.

Exercise

Expand on the above example and describe matrices representing the derivative of polynomials in $\mathbb{P}_n := \mathscr{P}_n(\mathbb{R})$, and do the same for the integral.

Preview: Eigen-Bases and Diagonalization

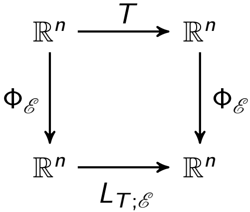
One can sometimes associate special coordinates on $\mathbb{F}^n$ to an endomorphism $T : \mathbb{F}^n \to \mathbb{F}^n$.

In particular, we will soon study the situation where a map $T : \mathbb{R}^n \to \mathbb{R}^n$ preserves or even fixes some subspaces $E_1, \ldots E_k \subset \mathbb{R}^n$.

Suppose T preserves k such subspaces $E_1, \ldots E_k$, such that $n = \dim E_1 + \dim E_2 + \ldots + \dim E_k$, and no E_i is contained in any E_j for $i \neq j$.
Then one can collect the bases of the individual subspaces into a basis $\mathscr{E}$ of $\mathbb{R}^n$ called an *eigenbasis*, and there is a special associated coordinate system called *eigen-coordinates*.

Eigen-Coordinates?

We have a commutative diagram:



What makes eigen-coordinates special is this: the transformation $[\mathbf{x}]_{\mathscr{E}} \mapsto [T(\mathbf{x})]_{\mathscr{E}} = [\mathrm{A}\mathbf{x}]_{\mathscr{E}}$ is given by a *diagonal matrix*! Indeed, $L_{T,\mathscr{E}}(\mathbf{y}) = \mathrm{D}\mathbf{y}$ for a diagonal matrix $\mathrm{D} = \mathrm{EAE}^{-1}$, where E is the matrix of the coordinate map $\Phi_{\mathscr{E}} : \mathbb{R}^n \to \mathbb{R}^n$.

Homework

- Please read sections 4.4, 4.5 and 4.6 of the textbook by Friday, 3/23.
- Homework in MyMathLab for section 4.2 on on Column and Null Spaces, Images/Ranges and Kernels is due Tuesday, 3/27.
- Homework in MyMathLab for section 4.3 on Basis and Linear Independence is due Thursday, 3/29.
- **Exam 2 will be held Tuesday, April 4/10/18, 7:00PM-9:00PM, in Hasbrouck Lab Addition room 124.** The syllabus for the second midterm is the following sections of the textbook: 2.2, 2.3, 3.1, 3.2, 3.3, 4.1, 4.2, 4.3, 4.4, 4.5.