[5] The Basis

Coordinate systems

In terms of vectors (and generalized beyond two dimensions),

- *coordinate system* for a vector space $\mathcal{V}$ is specified by generators $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of $\mathcal{V}$
- Every vector $\mathbf{v}$ in $\mathcal{V}$ can be written as a linear combination

$$\mathbf{v} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n$$

- We represent vector $\mathbf{v}$ by the vector $[\alpha_1, \ldots, \alpha_n]$ of coefficients. called the *coordinate representation* of $\mathbf{v}$ in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

But assigning coordinates to points is not enough. In order to avoid confusion, we must ensure that each point is assigned coordinates in exactly one way. How?

We will discuss unique representation later.

Coordinate representation

Definition: The *coordinate representation* of $\mathbf{v}$ in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the vector $[\alpha_1, \ldots, \alpha_n]$ such that

$$\mathbf{v} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n$$

In this context, the coefficients are called the *coordinates*.

Example: The vector $\mathbf{v} = [1, 3, 5, 3]$ is equal to

$$1\,[1, 1, 0, 0] + 2\,[0, 1, 1, 0] + 3\,[0, 0, 1, 1]$$

so the coordinate representation of $\mathbf{v}$ in terms of the vectors $[1, 1, 0, 0], [0, 1, 1, 0], [0, 0, 1, 1]$ is $[1, 2, 3]$.

Example: What is the coordinate representation of the vector $[6, 3, 2, 5]$ in terms of the vectors $[2, 2, 2, 3], [1, 0, -1, 0], [0, 1, 0, 1]$?

Since

$$[6, 3, 2, 5] = 2\,[2, 2, 2, 3] + 2\,[1, 0, -1, 0] - 1\,[0, 1, 0, 1],$$

the coordinate representation is $[2, 2, -1]$.

Coordinate representation

Definition: The *coordinate representation* of $\mathbf{v}$ in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the vector $[\alpha_1, \ldots, \alpha_n]$ such that

$$\mathbf{v} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n$$

In this context, the coefficients are called the *coordinates*.

Now we do an example with vectors over *GF* (2).

Example: What is the coordinate representation of the vector [0,0,0,1] in terms of the vectors [1,1,0,1], [0,1,0,1], and [1,1,0,0]?

Since

$$[0, 0, 0, 1] = 1\,[1, 1, 0, 1] + 0\,[0, 1, 0, 1] + 1\,[1, 1, 0, 0]$$

the coordinate representation of [0, 0, 0, 1] is [1, 0, 1].

Coordinate representation

Definition: The *coordinate representation* of $\mathbf{v}$ in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the vector $[\alpha_1, \ldots, \alpha_n]$ such that

$$\mathbf{v} = \alpha_1\,\mathbf{a}_1 + \cdots + \alpha_n\,\mathbf{a}_n$$

In this context, the coefficients are called the *coordinates*.

Why put the coordinates in a vector?

Makes sense in view of linear-combinations definitions of matrix-vector multiplication.

Let $A = \left[\begin{array}{c|c|c} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{array}\right]$.

- "$\mathbf{u}$ is the coordinate representation of $\mathbf{v}$ in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_n$" can be written as matrix-vector equation $A\mathbf{u} = \mathbf{v}$
- To go from a coordinate representation $\mathbf{u}$ to the vector being represented, we multiply A times $\mathbf{u}$.
- To go from a vector $\mathbf{v}$ to its coordinate representation, we can solve the matrix-vector equation $A\mathbf{x} = \mathbf{v}$.
  (Because the columns of A are generators for $\mathcal{V}$ and $\mathbf{v}$ belongs to $\mathcal{V}$, the equation must have at least one solution.)

Linear Combinations: Lossy compression

Say you need to store or transmit many 2-megapixel images:

How do we represent the image compactly?

- *Obvious method:* 2 million pixels $\Longrightarrow$ 2 million numbers

- *Strategy 1:* Use sparsity! Find the "nearest" k-sparse vector. Later we'll see this consists of suppressing all but the largest k entries.

- *More sophisticated strategy?*

Linear Combinations: Lossy compression

Strategy 2: Represent image vector by its coordinate representation:

- Before compressing any images, select vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
- Replace each image vector with its coordinate representation in terms of $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

For this strategy to work, we need to ensure that *every* image vector can be represented as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

> Given some D-vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ over $\mathbb{F}$, how can we tell whether *every* vector in $\mathbb{F}^D$ can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$?

We also need the number of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ to be much smaller than the number of pixels.

> Given D, what is minimum number of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that every vector in $\mathbb{F}^D$ can be written as a linear combination?

Linear Combinations: Lossy compression

Strategy 3: A hybrid approach

Step 1: Select vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Step 2: For each image to compress, find its coordinate representation $\mathbf{u}$ in terms of $\mathbf{v}_1, \ldots, \mathbf{v}_n$

Step 3: Replace $\mathbf{u}$ with the closest k-sparse vector $\tilde{\mathbf{u}}$, and store $\tilde{\mathbf{u}}$.

Step 4: To recover an image from $\tilde{\mathbf{u}}$, calculate the corresponding linear combination of $\mathbf{v}_1, \ldots \mathbf{v}_n$.

Greedy algorithms for finding a set of generators

Question: *For a given vector space* $\mathcal{V}$*, what is the minimum number of vectors whose span equals* $\mathcal{V}$*?*

How can we obtain a minimum number of vectors?

Two natural approaches come to mind, the *Grow* algorithm and the *Shrink* algorithm.

Grow algorithm

```
def Grow(V)
   S = ∅
   repeat while possible:
        find a vector v in V that is not in Span S, and put it in S.
```

The algorithm stops when there is no vector to add, at which time S spans all of $\mathcal{V}$.
Thus, if the algorithm stops, it will have found a generating set.

But is it bigger than necessary?

Shrink Algorithm

def $\mathrm{Shrink}(\mathcal{V})$
  S = some finite set of vectors that spans $\mathcal{V}$
  repeat while possible:
    find a vector $\mathbf{v}$ in S such that Span $(S - \{v\}) = \mathcal{V}$, and remove $\mathbf{v}$ from S.

The algorithm stops when there is no vector whose removal would leave a spanning set.
At every point during the algorithm, S spans $\mathcal{V}$, so it spans $\mathcal{V}$ at the end.
Thus, if the algorithm stops, the algorithm will have found a generating set.

The question is, again: is it bigger than necessary?

Linear Dependence: The Superfluous-Vector Lemma

Grow and Shrink algorithms both test whether a vector is superfluous in spanning a vector space $\mathcal{V}$. Need a criterion for superfluity.

Superfluous-Vector Lemma: For any set S and any vector $\mathbf{v} \in S$, if $\mathbf{v}$ can be written as a linear combination of the other vectors in S then Span $(S - \{\mathbf{v}\})$ = Span S

Proof: Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Suppose $\mathbf{v}_n = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{n-1} \mathbf{v}_{n-1}$

To show: every vector in Span S is also in Span $(S - \{\mathbf{v}_n\})$.

Every vector $\mathbf{v}$ in Span S can be written as $\mathbf{v} = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \cdots \beta_n \mathbf{v}_n$

Substituting for $\mathbf{v}_n$, we obtain

$$\begin{aligned} \mathbf{v} &= \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \cdots + \beta_n (\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{n-1} \mathbf{v}_{n-1}) \\ &= (\beta_1 + \beta_n \alpha_1)\mathbf{v}_1 + (\beta_2 + \beta_n \alpha_2)\mathbf{v}_2 + \cdots + (\beta_{n-1} + \beta_n \alpha_{n-1})\mathbf{v}_{n-1} \end{aligned}$$

which shows that an arbitrary vector in Span S can be written as a linear combination of vectors in $S - \{\mathbf{v}_n\}$ and is therefore in Span $(S - \{\mathbf{v}_n\})$.

QED

Defining linear dependence

Definition: Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are *linearly dependent* if the zero vector can be written as a **nontrivial** linear combination of the vectors:

$$\mathbf{0} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

In this case, we refer to the linear combination as a *linear dependency* in $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

On the other hand, if the *only* linear combination that equals the zero vector is the trivial linear combination, we say $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly *in*dependent.

Example: The vectors [1, 0, 0], [0, 2, 0], and [2, 4, 0] are linearly dependent, as shown by the following equation:

$$2\,[1, 0, 0] + 2\,[0, 2, 0] - 1\,[2, 4, 0] = [0, 0, 0]$$

Therefore:

$2\,[1, 0, 0] + 2\,[0, 2, 0] - 1\,[2, 4, 0]$ is a linear dependency in [1, 0, 0], [0, 2, 0], [2, 4, 0].

Linear dependence

Example: The vectors [1, 0, 0], [0, 2, 0], and [0, 0, 4] are linearly independent.
How do we know?
Easy since each vector has a nonzero entry where the others have zeroes.
Consider any linear combination

$$\alpha_1 [1, 0, 0] + \alpha_2 [0, 2, 0] + \alpha_3 [0, 0, 4]$$

This equals $[\alpha_1, 2\alpha_2, 4\alpha_3]$

If this is the zero vector, it must be that $\alpha_1 = \alpha_2 = \alpha_3 = 0$
That is, the linear combination is trivial.

We have shown the only linear combination that equals the zero vector is the trivial linear combination.

Properties of linear independence: hereditary

Lemma: If a finite set S of vectors is linearly dependent and S is a subset of T then T is linearly dependent.

In graphs, if a set S of edges includes a cycle then a superset of S also includes a cycle.

Properties of linear independence: hereditary

Lemma: If a finite set S of vectors is linearly dependent and S is a subset of T then T is linearly dependent.
Proof: If the zero vector can be written as a nontrivial linear combination of some vectors, it can be so written even if we allow some extra vectors to be in the linear combination because we can use zero coefficients on the extra vectors.

More formal proof: Write $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ and $T = \{\mathbf{s}_1, \ldots, \mathbf{s}_n, \mathbf{t}_1, \ldots, \mathbf{t}_k\}$. Suppose S is linearly dependent. Then there are coefficients $a_1, \ldots, a_n$, not all zero, such that

$$\mathbf{0} = a_1\mathbf{s}_1 + \cdots + a_n\mathbf{s}_n$$

Therefore

$$\mathbf{0} = a_1\mathbf{s}_1 + \cdots + a_n\mathbf{s}_n + 0\,\mathbf{t}_1 + \cdots 0\,\mathbf{t}_k$$

which shows that the zero vector can be written as a nontrivial linear combination of the vectors of T ,i.e. that T is linearly dependent.
QED

Properties of linear (in)dependence

Linear-Dependence Lemma Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors.

A vector $\mathbf{v}_i$ is in the span of the other vectors

if and only if

the zero vector can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$

in which the coefficient of $\mathbf{v}_i$ is nonzero.

In graphs, the Linear-Dependence Lemma states that an edge e is in the span of other edges if there is a cycle consisting of e and a subset of the other edges.

Properties of linear (in)dependence

Linear-Dependence Lemma Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors.

A vector $\mathbf{v}_i$ is in the span of the other vectors
if and only if
the zero vector can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$
in which the coefficient of $\mathbf{v}_i$ is nonzero.

Proof: *First direction*: Suppose $\mathbf{v}_i$ is in the span of the other vectors. That is, there exist coefficients $\alpha_1, \ldots, \alpha_{n-1}$ such that

$$\mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_{i-1} \mathbf{v}_{i-1} + \alpha_{i+1} \mathbf{v}_{i+1} + \cdots \alpha_n \mathbf{v}_n$$

Moving $\mathbf{v}_i$ to the other side, we can write

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \cdots + (-1)\mathbf{v}_i + \cdots + \alpha_n \mathbf{v}_n$$

which shows that the all-zero vector can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in which the coefficient of $\mathbf{v}_i$ is nonzero.

Properties of linear (in)dependence

Linear-Dependence Lemma Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors.

A vector $\mathbf{v}_i$ is in the span of the other vectors
if and only if
the zero vector can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$
in which the coefficient of $\mathbf{v}_i$ is nonzero.

Proof: *Now for the other direction*. Suppose there are coefficients $\alpha_1, \ldots, \alpha_n$ such that

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_i \mathbf{v}_i + \cdots + \alpha_n \mathbf{v}_n$$

and such that $\alpha_i \neq 0$.

Dividing both sides by α_i yields

$$\mathbf{0} = (\alpha_1/\alpha_i)\,\mathbf{v}_1 + (\alpha_2/\alpha_i)\,\mathbf{v}_2 + \cdots + \mathbf{v}_i + \cdots + (\alpha_n/\alpha_i)\,\mathbf{v}_n$$

Moving every term from right to left except $\mathbf{v}_i$ yields

$$-(\alpha_1/\alpha_i)\,\mathbf{v}_1 - (\alpha_2/\alpha_i)\,\mathbf{v}_2 - \cdots - (\alpha_n/\alpha_i)\,\mathbf{v}_n = \mathbf{v}_i$$

QED

Properties of linear (in)dependence

Linear-Dependence Lemma Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors. A vector $\mathbf{v}_i$ is in the span of the other vectors if and only if the zero vector can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in which the coefficient of $\mathbf{v}_i$ is nonzero.

Contrapositive:

$\mathbf{v}_i$ is *not* in the space of the other vectors
if and only if
for any linear combination equaling the zero vector

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_i \mathbf{v}_i + \cdots + \alpha_n \mathbf{v}_n$$

it must be that the coefficient α_i *is* zero.

Analyzing the Grow algorithm

```
def Grow(V)
    S = ∅
    repeat while possible:
        find a vector v in V that is not in Span S, and put it in S.
```

Grow-Algorithm Corollary: The vectors obtained by the Grow algorithm are linearly independent.

In graphs, this means that the solution obtained by the Grow algorithm has no cycles (is a forest).

Analyzing the Grow algorithm

Grow-Algorithm Corollary: The vectors obtained by the Grow algorithm are linearly independent.

Proof: For $n = 1, 2, \ldots,$ let $\mathbf{v}_n$ be the vector added to S in the n^{th} iteration of the Grow algorithm. We show by induction that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent.

For $n = 0$, there are no vectors, so the claim is trivially true.
Assume the claim is true for $n = k - 1$. We prove it for $n = k$.
The vector $\mathbf{v}_k$ added to S in the k^{th} iteration is not in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$.
Therefore, by the Linear-Dependence Lemma, for any coefficients $\alpha_1, \ldots, \alpha_k$ such that

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_{k-1} \mathbf{v}_{k-1} + \alpha_k \mathbf{v}_k$$

it must be that α_k equals zero. We may therefore write

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_{k-1} \mathbf{v}_{k-1}$$

By claim for $n = k - 1$, $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ are linearly independent, so $\alpha_1 = \cdots = \alpha_{k-1} = 0$

The linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is *trivial*. We have proved that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent. This proves the claim for $n = k$. QED

Analyzing the Shrink algorithm

def $\mathrm{Shrink}(\mathcal{V})$
  S = some finite set of vectors that spans $\mathcal{V}$
  repeat while possible:
    find a vector **v** in S such that Span $(S - \{v\}) = \mathcal{V}$, and remove **v** from S.

Shrink-Algorithm Corollary: The vectors obtained by the Shrink algorithm are linearly independent.

In graphs, this means that the Shrink algorithm outputs a solution that is a forest.

Recall:

Superfluous-Vector Lemma For any set S and any vector $\mathbf{v} \in S$, if **v** can be written as a linear combination of the other vectors in S then Span $(S - \{\mathbf{v}\})$ = Span S

Analyzing the Shrink algorithm

Shrink-Algorithm Corollary: The vectors obtained by the Shrink algorithm are linearly independent.

Proof: Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be the set of vectors obtained by the Shrink algorithm. Assume for a contradiction that the vectors are linearly dependent.

Then $\mathbf{0}$ can be written as a nontrivial linear combination

$$\mathbf{0} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

where at least one of the coefficients is nonzero.

Let α_i be one of the nonzero coefficients.

By the Linear-Dependence Lemma, $\mathbf{v}_i$ can be written as a linear combination of the other vectors.

Hence by the Superfluous-Vector Lemma, Span $(S - \{\mathbf{v}_i\}) = \text{Span } S$,

so the Shrink algorithm should have removed $\mathbf{v}_i$. QED

Basis

If they successfully finish, the Grow algorithm and the Shrink algorithm each find a set of vectors spanning the vector space $\mathcal{V}$. In each case, the set of vectors found is linearly independent.

Definition: Let $\mathcal{V}$ be a vector space. A *basis* for $\mathcal{V}$ is a linearly independent set of generators for $\mathcal{V}$.

Thus a set S of vectors of $\mathcal{V}$ is a *basis* for $\mathcal{V}$ if S satisfies two properties:

Property B1 (*Spanning*) Span $S = \mathcal{V}$, and

Property B2 (*Independent*) S is linearly independent.

Most important definition in linear algebra.

Basis: Examples

A set S of vectors of $\mathcal{V}$ is a *basis* for $\mathcal{V}$ if S satisfies two properties:

Property B1 (*Spanning*) Span $S = \mathcal{V}$, and

Property B2 (*Independent*) S is linearly independent.

Example: Let $\mathcal{V}$ = Span $\{[1, 0, 2, 0], [0, -1, 0, -2], [2, 2, 4, 4]\}$.
Is $\{[1, 0, 2, 0], [0, -1, 0, -2], [2, 2, 4, 4]\}$ a basis for $\mathcal{V}$?

The set *is* spanning but is *not* independent

$$1\,[1, 0, 2, 0] - 1\,[0, -1, 0, -2] - \frac{1}{2}[2, 2, 4, 4] = \mathbf{0}$$

so not a basis

However, $\{[1, 0, 2, 0], [0, -1, 0, -2]\}$ *is* a basis:

- Obvious that these vectors are independent because each has a nonzero entry where the other has a zero.
- To show
  Span $\{[1, 0, 2, 0], [0, -1, 0, -2]\}$ = Span $\{[1, 0, 2, 0], [0, -1, 0, -2], [2, 2, 4, 4]\}$,
  can use Superfluous-Vector Lemma:

$$[2, 2, 4, 4] = 2\,[1, 0, 2, 0] - 2\,[0, -1, 0, -2]$$

Basis: Examples

Example: A simple basis for R^3: the standard generators
$\mathbf{e}_1 = [1, 0, 0]$, $\mathbf{e}_2 = [0, 1, 0]$, $\mathbf{e}_3 = [0, 0, 1]$.

- *Spanning*: For any vector $[x, y, z] \in R^3$,

$$[x, y, z] = x\,[1, 0, 0] + y\,[0, 1, 0] + z\,[0, 0, 1]$$

- *Independent*: Suppose

$$\mathbf{0} = \alpha_1\,[1, 0, 0] + \alpha_2\,[0, 1, 0] + \alpha_3\,[0, 0, 1] = [\alpha_1, \alpha_2, \alpha_3]$$

Then $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Instead of "standard generators", we call them *standard basis vectors*.
We refer to $\{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ as *standard basis* for R^3.

In general the standard generators are usually called *standard basis vectors*.

Basis: Examples

Example: Another basis for R^3: [1, 1, 1], [1, 1, 0], [0, 1, 1]

- *Spanning*: Can write standard generators in terms of these vectors:

$$
\begin{aligned}
[1, 0, 0] &= [1, 1, 1] - [0, 1, 1] \\
[0, 1, 0] &= [1, 1, 0] + [0, 1, 1] - [1, 1, 1] \\
[0, 0, 1] &= [1, 1, 1] - [1, 1, 0]
\end{aligned}
$$

  Since $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ can be written in terms of these new vectors, every vector in Span $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is in span of new vectors. Thus R^3 equals span of new vectors.

- *Linearly independent*: Write zero vector as linear combination:

$$\mathbf{0} = x\,[1, 1, 1] + y\,[1, 1, 0] + z\,[0, 1, 1] = [x + y, x + y + z, x + z]$$

Looking at each entry, we get

$$
\begin{aligned}
0 &= x + y \\
0 &= x + y + z \\
0 &= x + z
\end{aligned}
$$

Plug $x + y = 0$ into second equation to get $0 = z$.
Plug $z = 0$ into third equation to get $x = 0$.
Plug $x = 0$ into first equation to get $y = 0$.
Thus the linear combination is trivial.

Uniqueness of representation in terms of a basis

Unique-Representation Lemma: Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be a basis for V. For any vector $\mathbf{v} \in V$, there is exactly one representation of $\mathbf{v}$ in terms of the basis vectors.

Proof: Let $\mathbf{v}$ be any vector in V.

The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ span V, so there is at least one representation of $\mathbf{v}$ in terms of the basis vectors.

Suppose there are two such representations:

$$\mathbf{v} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n = \beta_1 \mathbf{a}_1 + \cdots + \beta_n \mathbf{a}_n$$

We get the zero vector by subtracting one from the other:

$$\begin{aligned} \mathbf{0} &= \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n - (\beta_1 \mathbf{a}_1 + \cdots + \beta_n \mathbf{a}_n) \\ &= (\alpha_1 - \beta_1)\mathbf{a}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{a}_n \end{aligned}$$

Since the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent, the coefficients $\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n$ must all be zero, so the two representations are really the same.

QED

Change of basis

Proposition: If $a_1, \dots a_n$ and $c_1, \dots c_k$ are bases for F^m then multiplication by the matrix

$$B = [a_1| \dots |a_n]^{-1}[c_1| \dots |c_k]$$

maps

- from the representation of a vector with respect to $c_1, \dots c_k$
- to the representation of that vector with respect to $a_1, \dots a_n$.

Conclusion: Given two bases of F^m, there is a matrix B such that multiplication by B converts from one coordinate representation to the other.

Remark: Converting between vector itself and its coordinate representation is a special case:

- Think of the vector itself as coordinate representation with respect to standard basis.

Change of basis: simple example

Example: To map from coordinate representation with respect to $[1,2,3], [2,1,0], [0,1,4]$ to coordinate representation with respect to $[2,0,1], [0,1,-1], [1,2,0]$ multiply by the matrix

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

Which is

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

Which is

$$\begin{bmatrix} -1 & 1 & -5/3 \\ -4 & 1 & -17/3 \\ 3 & 9 & 10/3 \end{bmatrix}$$

Simplified Exchange Lemma

We need a tool to iteratively transform one set of generators into another.

- You have a set S of vectors.
- You have a vector z you want to inject into S.
- You want to maintain same size so must eject a vector from S.
- You want the span to not change.

Exchange Lemma tells you how to choose vector to eject.

Simplified Exchange Lemma:

- Suppose S is a set of vectors.
- Suppose $\mathbf{z}$ is a nonzero vector in Span S.
- Then there is a vector $\mathbf{w}$ in S such that

$$\text{Span}\,(S \cup \{\mathbf{z}\} - \{\mathbf{w}\}) = \text{Span}\,S$$

Simplified Exchange Lemma proof

Simplified Exchange Lemma: Suppose S is a set of vectors, and $\mathbf{z}$ is a nonzero vector in Span S. Then there is a vector $\mathbf{w}$ in S such that Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}\}) =$ Span S.

Proof: Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Since $\mathbf{z}$ is in Span S, can write

$$\mathbf{z} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$$

By Superfluous-Vector Lemma, Span $(S \cup \{z\}) =$ Span S.
Since $\mathbf{z}$ is nonzero, at least one of the coefficients is nonzero, say α_i.
Rewrite as

$$\mathbf{z} - \alpha_1 \mathbf{v}_1 - \cdots - \alpha_{i-1} \mathbf{v}_{i-1} - \alpha_{i+1} \mathbf{v}_{i+1} - \cdots - \alpha_n \mathbf{v}_n = \alpha_i \mathbf{v}_i$$

Divide through by α_i:

$$(1/\alpha_i)\mathbf{z} - (\alpha_1/\alpha_i)\mathbf{v}_1 - \cdots - (\alpha_{i-1}/\alpha_i)\mathbf{v}_{i-1} - (\alpha_{i+1}/\alpha_i)\mathbf{v}_{i+1} - \cdots - (\alpha_n/\alpha_i)\mathbf{v}_n = \mathbf{v}_i$$

By Superfluous-Vector Lemma, Span $(S \cup \{\mathbf{z}\}) =$ Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}\})$. QED

Exchange Lemma

Simplified Exchange Lemma: Suppose S is a set of vectors, and $\mathbf{z}$ is a nonzero vector in Span S. Then there is a vector $\mathbf{w}$ in S such that Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}\})$ = Span S.

Need to enhance this lemma. Set of *protected* elements is A:

Exchange Lemma:

- Suppose S is a set of vectors and A is a subset of S.
- Suppose $\mathbf{z}$ is a vector in Span S such that $A \cup \{\mathbf{z}\}$ is linearly independent.
- Then there is a vector $\mathbf{w} \in S - A$ such that Span S = Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}\})$

Now, not enough that $\mathbf{z}$ be nonzero—need A to be linearly independent.

Exchange Lemma proof

Exchange Lemma: Suppose S is a set of vectors and A is a subset of S. Suppose $\mathbf{z}$ is a vector in Span S such that $A \cup \{\mathbf{z}\}$ is linearly independent.
Then there is a vector $\mathbf{w} \in S - A$ such that Span S = Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}\})$

Proof: Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}_1, \ldots, \mathbf{w}\}$ and $A = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.
Since $\mathbf{z}$ is in Span S, can write

$$\mathbf{z} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k + \beta_1 \mathbf{w}_1 + \cdots + \beta \mathbf{w}$$

By Superfluous-Vector Lemma, Span $(S \cup \{z\})$ = Span S.
If coefficients $\beta_1, \ldots, \beta$ were all zero then we would have $\mathbf{z} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k$, contradicting the linear independence of $A \cup \{\mathbf{z}\}$.
Thus one of the coefficients $\beta_1, \ldots, \beta$ must be nonzero... say β_1. Rewrite as

$$\mathbf{z} - \alpha_1 \mathbf{v}_1 - \cdots - \alpha_k \mathbf{v}_k - \beta_2 \mathbf{w}_2 - \cdots - \beta \mathbf{w} = \beta_1 \mathbf{w}_1$$

Divide through by β_1:

$$(1/\beta_1)\mathbf{z} - (\alpha_1/\beta_1)\mathbf{v}_1 - \cdots - (\alpha_k/\beta_1)\mathbf{v}_k - (\beta_2/\beta_1)\mathbf{w}_2 - \cdots - (\beta\ /\beta_1)\mathbf{w} = \mathbf{w}_1$$

By Superfluous-Vector Lemma, Span $(S \cup \{\mathbf{z}\})$ = Span $(S \cup \{\mathbf{z}\} - \{\mathbf{w}_1\})$. QED