Fundamentals of linear algebra

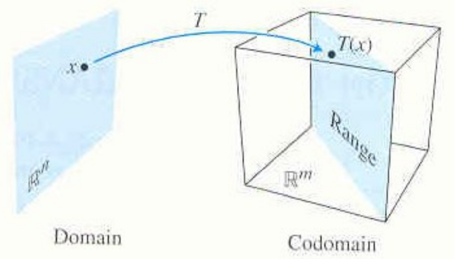
- Linear transformations
- Basis and coordinates
- Coordinate mapping
- Change of basis
- Dimensions of vectors spaces

Wednesday, 8 May 2019

Revision: Transformations

Definition: Transformation (mapping, function) $\mathcal{T}$ from $\mathbb{R}^n$ into $\mathbb{R}^m$ is a rule assigning a vector $\mathcal{T}(\mathbf{x}) = \mathbf{y} \in \mathbb{R}^m$ to each vector $\mathbf{x} \in \mathbb{R}^n$.

- $\mathbb{R}^n$ is called the **domain**, and $\mathbb{R}^m$ the **codomain** of $\mathcal{T}$.
- The usual notation is: $\mathcal{T} : \mathbb{R}^n \mapsto \mathbb{R}^m$.
- Vector $\mathbf{y} = \mathcal{T}(\mathbf{x})$ is called the **image** of $\mathbf{x}$.
- The set of all images $\{\mathcal{T}(\mathbf{x})\}$ is called the **range** of $\mathcal{T}$.



Revision: Linear transformations

Definition:

Transformation $\mathcal{T}$ is **linear** if: $\forall\, \mathbf{u}, \mathbf{v}$ in the domain of $\mathcal{T}$

(i) $\mathcal{T}(\mathbf{u}+\mathbf{v}) = \mathcal{T}(\mathbf{u}) + \mathcal{T}(\mathbf{v})$

(ii) $\mathcal{T}(c\,\mathbf{u}) = c\,\mathcal{T}(\mathbf{u}) \quad \forall\, c \in \mathbb{R}$

Notes:

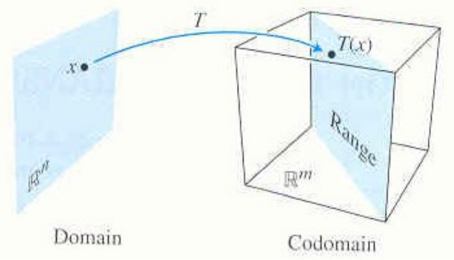
If $\mathcal{T}$ is a linear transformation, then

- $\mathcal{T}(\mathbf{0}) = \mathbf{0}$ (proof: $\mathcal{T}(\mathbf{0}) = \mathcal{T}(0\cdot\mathbf{0}) = 0\cdot\mathcal{T}(\mathbf{0}) = \mathbf{0}$)
- $\mathcal{T}(c\,\mathbf{u} + d\,\mathbf{v}) = c\,\mathcal{T}(\mathbf{u}) + d\,\mathcal{T}(\mathbf{v})$ (follows from (i) and (ii))

If $\mathbf{T}$ is an $m \times n$ matrix, $\mathcal{T}(\mathbf{x}) = \mathbf{T}\mathbf{x}$ is a linear transformation.

Revision: Linear transformations

- For $\mathbf{x} \in \mathbb{R}^n$, a linear transformation $\mathcal{T}(\mathbf{x})$ into $\mathbb{R}^m$ can be computed as $\mathbf{Tx}$ where $\mathbf{T}$ is an $m \times n$ matrix.
- If the domain of $\mathcal{T}$ is $\mathbb{R}^n$ then $\mathbf{T}$ has n columns, and if the codomain of $\mathcal{T}$ is $\mathbb{R}^m$ then $\mathbf{T}$ has m rows.
- The range of $\mathcal{T}$ is $\mathrm{Span}\{\mathbf{a}_1 \ldots \mathbf{a}_n\}$ and image of $\mathbf{x}$ is $\mathbf{Tx}$.



Standard vectors and standard matrix

In $\mathbb{R}^2$: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

The vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are called the *standard unit vectors in* $\mathbb{R}^2$, and $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$.

Similarly, in $\mathbb{R}^3$

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$.

Generally in $\mathbb{R}^n$ $\quad \mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$ $\quad$ with $\quad \mathbf{e}_n : \begin{cases} e_n^{(i=n)} = 1 \\ e_n^{(i\neq n)} = 0 \end{cases}$

Standard vectors and standard matrix

For any linear transformation $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ we can write

$$\mathcal{T}(\mathbf{x}) = \mathcal{T}(x_1\mathbf{e}_1+x_2\mathbf{e}_2) = \mathcal{T}(x_1\mathbf{e}_1)+\mathcal{T}(x_2\mathbf{e}_2) = x_1\mathcal{T}(\mathbf{e}_1)+x_2\mathcal{T}(\mathbf{e}_2)$$

This can be written as $\mathcal{T}(\mathbf{x}) = \left[\begin{array}{c|c} \mathcal{T}(\mathbf{e}_1) & \mathcal{T}(\mathbf{e}_2) \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] \equiv \mathbf{Tx}$

$\mathbf{T} = \left[\begin{array}{c|c} \mathcal{T}(\mathbf{e}_1) & \mathcal{T}(\mathbf{e}_2) \end{array}\right]$ is called the *standard matrix* of $\mathcal{T}$.

Similarly, for any linear transformation $\mathcal{T} : \mathbb{R}^n \mapsto \mathbb{R}^n$ we can write

$$\begin{aligned} \mathcal{T}(\mathbf{x}) = \mathcal{T}(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) &= \mathcal{T}(x_1\mathbf{e}_1) + \ldots + \mathcal{T}(x_n\mathbf{e}_n) \\ &= x_1\mathcal{T}(\mathbf{e}_1) + \ldots + x_n\mathcal{T}(\mathbf{e}_n) = \mathbf{Tx} \end{aligned}$$

with the standard matrix $\mathbf{T} = \left[\begin{array}{c|c|c} \mathcal{T}(\mathbf{e}_1) & \ldots & \mathcal{T}(\mathbf{e}_n) \end{array}\right]$.

Linear transformations and standard matrix

Example in 3D:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{y}_1 = \begin{bmatrix} 3 \\ 5 \\ -7 \end{bmatrix}, \qquad \mathbf{y}_2 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}, \qquad \mathbf{y}_3 = \begin{bmatrix} -1 \\ 3 \\ 5 \end{bmatrix}.$$

Let $\mathcal{T} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ be a linear transformation that maps $\mathbf{e}_i$ to $\mathbf{y}_i$.

Let us obtain the standard matrix $\mathbf{T}$ for transformation $\mathcal{T}$,

and then find the image of $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$.

Linear transformations and standard matrix

The transformation is defined so that $\mathcal{T}(\mathbf{e}_i) = \mathbf{y}_i$.

The standard matrix $\mathbf{T}$ then has columns directly given by $\mathbf{y}_k$:

$$\mathbf{T} = \left[\mathcal{T}(\mathbf{e}_1)\ \mathcal{T}(\mathbf{e}_2)\ \mathcal{T}(\mathbf{e}_3)\right] = [\mathbf{y}_1\ \mathbf{y}_2\ \mathbf{y}_3] = \begin{bmatrix} 3 & 2 & -1 \\ 5 & 0 & 3 \\ -7 & 3 & 5 \end{bmatrix}$$

The image of $\mathbf{u}$ is

$$\mathcal{T}(\mathbf{u}) = \mathbf{T}\mathbf{u} = \begin{bmatrix} 3 & 2 & -1 \\ 5 & 0 & 3 \\ -7 & 3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 14 \\ 12 \\ -20 \end{bmatrix}$$

Linear transformations and standard matrix

Example: Find the standard matrix for dilation $\mathcal{T}(\mathbf{x}) = 3\mathbf{x}$.

We first check the action of dilation on standard vectors:

$$\mathcal{T}(\mathbf{e}_1) = 3\mathbf{e}_1 = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

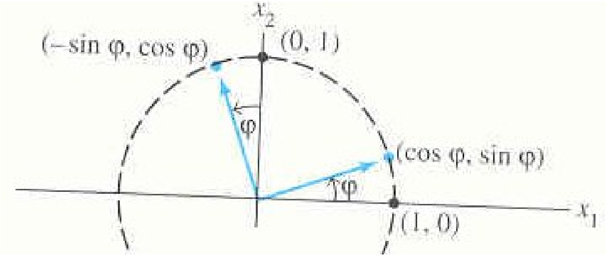
$$\mathcal{T}(\mathbf{e}_2) = 3\mathbf{e}_2 = 3\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Then we can form the required standard matrix:

$$\mathbf{D}_3 = \left[\mathcal{T}(\mathbf{e}_1)\ \mathcal{T}(\mathbf{e}_2)\right] = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Linear transformations and standard matrix

Example: Find the standard matrix of a transformation that rotates any vector in $\mathbb{R}^2$ around the origin by an angle φ radians.



$$\mathcal{T}(\mathbf{e}_1) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \quad \text{and} \quad \mathcal{T}(\mathbf{e}_2) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$$

Therefore

$$\mathbf{R}_\varphi = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Linear transformations and standard matrix

Example: Consider a linear transformation $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ by

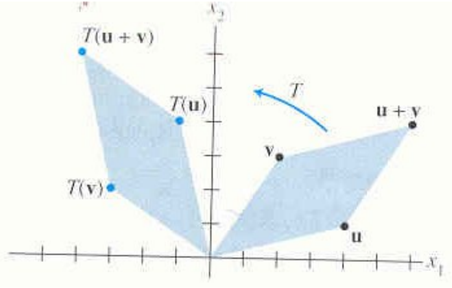
$$\mathbf{T} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{so} \quad \mathcal{T}(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Find the images of: $\mathbf{u} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{u} + \mathbf{v} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

$$\mathcal{T}(\mathbf{u}) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

$$\mathcal{T}(\mathbf{v}) = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

$$\mathcal{T}(\mathbf{u} + \mathbf{v}) = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$



$\mathcal{T}$ rotates vectors by 90^0 counterclockwise about the origin.

Linear transformations, examples

Example: $\mathbf{T} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}, \quad \mathcal{T}(\mathbf{x}) = \mathbf{T}\mathbf{x}.$

To find $\mathbf{x}$ such that $\mathcal{T}(\mathbf{x}) = \mathbf{b}$ we need to solve the system:

$$\left[\begin{array}{cc|c} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{array}\right] \quad \rightarrow \quad \left[\begin{array}{cc|c} 1 & 0 & 3/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{array}\right]$$

So $x_1 = \dfrac{3}{2}$, $x_2 = -\dfrac{1}{2}$ and thus $\mathbf{b}$ is the image of $\mathbf{x} = \begin{bmatrix} 1.5 \\ -0.5 \end{bmatrix}$

Any vector producing image $\mathbf{b}$ under $\mathcal{T}$ must satisfy this system.

A unique solution implies only one $\mathbf{x}$ with the image $\mathbf{b}$.

Linear transformations, examples

Example: $\mathbf{T} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}, \quad \mathcal{T}(\mathbf{x}) = \mathbf{T}\mathbf{x}.$

Determine whether or not $\mathbf{c}$ is in the range of transformation $\mathcal{T}$.

Vector $\mathbf{c}$ is in the range of $\mathcal{T}(\mathbf{x})$ if $\mathbf{c}$ is an image of some $\mathbf{x} \in \mathbb{R}^2$.

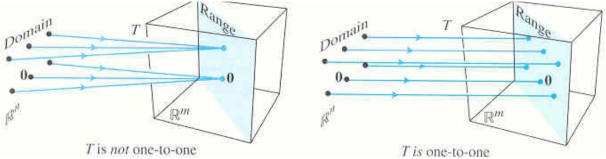
Solving the system $\mathbf{T}\mathbf{x} = \mathbf{c}$ gives

$$\left[\begin{array}{cc|c} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{array}\right] \quad \to \quad \left[\begin{array}{cc|c} 1 & -3 & 3 \\ 0 & 1 & 1/2 \\ 0 & 0 & 10 \end{array}\right]$$

The system is inconsistent and thus $\mathbf{c}$ is **not** in the range of $\mathcal{T}$.
There are no vectors in $\mathbb{R}^2$ with an image $\mathbf{c}$ under $\mathcal{T}$.

Linear transformations

Definition: A mapping $\mathcal{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be *one-to-one* mapping if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$.



- Equivalently, $\mathcal{T}$ is a one-to-one transformation if $\forall\, \mathbf{b} \in \mathbb{R}^m$ the equation $\mathcal{T}(\mathbf{x}) = \mathbf{b}$ has either a unique or no solutions.
- A one-to-one transformation is a unique transformation.

Linear transformations

Theorem: A linear transformation $\mathcal{T} : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a one-to-one transformation if and only if the equation $\mathcal{T}(\mathbf{x}) = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

Proof: Since $\mathcal{T}$ is linear then $\mathcal{T}(\mathbf{0}) = \mathbf{0}$. If $\mathcal{T}$ is one-to-one, then the equation $\mathcal{T}(\mathbf{x}) = \mathbf{0}$ has at most one solution, $\mathbf{x} = \mathbf{0}$.

Suppose $\mathcal{T}$ is *not* one-to-one. Then $\exists\, \mathbf{b} \in \mathbb{R}^m$ which is the image of at least 2 different vectors $\mathbf{u} \neq \mathbf{v}$ in $\mathbb{R}^n$: $\mathcal{T}(\mathbf{u}) = \mathbf{b}$, $\mathcal{T}(\mathbf{v}) = \mathbf{b}$.

But since $\mathcal{T}$ is linear, $\mathcal{T}(\mathbf{u} - \mathbf{v}) = \mathcal{T}(\mathbf{u}) - \mathcal{T}(\mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$.

However $\mathbf{u} - \mathbf{v} \neq \mathbf{0}$ and then the equation $\mathcal{T}(\mathbf{x}) = \mathbf{0}$ is going to have more than one solution: $\mathbf{0}$ and $\mathbf{u} - \mathbf{v}$.

Therefore, the assumption that $\mathcal{T}$ is not one-to-one must be false, which proves that $\mathcal{T}$ is one-to-one.

Linear transformations

Theorem: Let $\mathcal{T} : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation with the standard matrix $\mathbf{T}$. Then:

- $\mathcal{T}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of $\mathbf{T}$ span $\mathbb{R}^m$.
- $\mathcal{T}$ is one-to-one if and only if the columns of $\mathbf{T}$ are linearly independent.

Proof:

(a) The columns of $\mathbf{T}$ span $\mathbb{R}^m$ if and only if $\forall\, \mathbf{b} \in \mathbf{R}^m$ the equation $\mathbf{Tx} = \mathbf{b}$ is consistent, so it has a solution for every $\mathbf{b}$. This is true only when $\mathcal{T}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

(b) From the previous theorem, if $\mathcal{T}$ is one-to-one, the equation $\mathbf{Tx} = \mathbf{0}$ has only the trivial solution. This can only happen if the columns of $\mathbf{T}$ are linearly independent.

Linear transformations, examples

Example: Consider a linear transformation with standard matrix

$$\mathbf{T} = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Check if this transformation $\mathcal{T}$ is a one-to-one linear transformation, and find the domain, codomain, and range of $\mathcal{T}$.

- $\mathcal{T}$ can be presented as $\mathbf{Tx}$ so it is a linear transformation.
- The domain of $\mathcal{T}$ is $\mathbb{R}^4$ and the codomain is $\mathbb{R}^3$.
- There are three pivots in $\mathbf{T}$ so the columns of $\mathbf{T}$ span $\mathbb{R}^3$, thus the range of $\mathcal{T}$ is the entire $\mathbb{R}^3$.
- The augmented matrix $[\mathbf{T} \,|\, \mathbf{b}]$ has a solution for any $\mathbf{b}$. However there is one free variable so each $\mathbf{b}$ is an image for more then one $\mathbf{x}$, therefore this mapping is **not** one-to-one.

Linear transformations, examples

Example: Check if the transformation $\mathcal{T}(x_1, x_2) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix}$ is a one-to-one linear transformation, and find its range.

$$\mathcal{T}(\mathbf{x}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{T}\mathbf{x}$$

- The columns of $\mathbf{T}$ are linearly independent (not multiples of each other), therefore $\mathcal{T}$ is a one-to-one linear transformation.
- The domain of $\mathcal{T}$ is $\mathbb{R}^2$ and the codomain is $\mathbb{R}^3$.
- There are only two columns in $\mathbf{T}$, so they cannot span $\mathbb{R}^3$; the columns of $\mathbf{T}$ span a plane within $\mathbb{R}^3$;
- That plane is the range of $\mathcal{T}$ — any image produced by $\mathcal{T}$ is a linear combination of the columns of $\mathbf{T}$.

Revision: Basis

An indexed set $\mathcal{B} = \{\mathbf{b}_1, \dots \mathbf{b}_p\}$ is a **basis** in V if:
$\mathcal{B}$ is a linearly independent set, and $V = \mathrm{Span}\{\mathbf{b}_1, \dots \mathbf{b}_p\}$

Theorem: (the *unique representation theorem*)

Let $\mathcal{B} = \{\mathbf{b}_1, \dots \mathbf{b}_n\}$ be a basis for a vector space V.

Then $\forall\, \mathbf{x} \in V$ there exists a *unique* set of scalars $x_1, \dots x_n$ such that

$$\mathbf{x} = x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n.$$

- A basis is an "efficient" spanning set
  that contains only "necessary" vectors.
- A basis can be constructed from a spanning set
  by discarding unnecessary vectors.

Revision: Spanning set theorem

Spanning set theorem:

Let $\mathcal{S} = \{\mathbf{v}_1, \dots \mathbf{v}_p\}$ be a set in V, and $H = \mathrm{Span}\{\mathbf{v}_1, \dots \mathbf{v}_p\}$.

(a) If one of the vectors in $\mathcal{S}$, say $\mathbf{v}_i$, is a linear combination of the other vectors in $\mathcal{S}$, then the set formed from $\mathcal{S}$ by removing $\mathbf{v}_i$ still spans H.

(b) If $H \neq \{\mathbf{0}\}$, some subset of $\mathcal{S}$ is a basis for H.

- A basis is a smallest possible spanning set.
- A basis is also a largest possible linearly independent set.
- If $\mathcal{S}$ is a basis for V and is enlarged by one vector $\mathbf{w} \in V$ then the enlarged set cannot be linearly independent, because $\mathcal{S}$ spans V so $\mathbf{w}$ is a linear combination of the vectors of $\mathcal{S}$.
- If $\mathcal{S}$ is a basis for V and if $\mathcal{S}$ is made smaller by one vector $\mathbf{u} \in V$ then the reduced set cannot serve as a basis, because it will not span V anymore.

Revision: Coordinate systems

Definition: Let $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ be a basis for V, and $\mathbf{x} \in V$. The *coordinates* of $\mathbf{x}$ *relative to basis* $\mathcal{B}$ (or $\mathcal{B}$-coordinates of $\mathbf{x}$) are the coefficients $x_1, \ldots x_n$ such that

$$\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n.$$

If $x_1, \ldots x_n$ are the $\mathcal{B}$-coordinates of $\mathbf{x}$, then the vector in $\mathbb{R}^n$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

is the *coordinate vector* of $\mathbf{x}$ (relative to $\mathcal{B}$), or the $\mathcal{B}$*-coordinate vector* of $\mathbf{x}$.

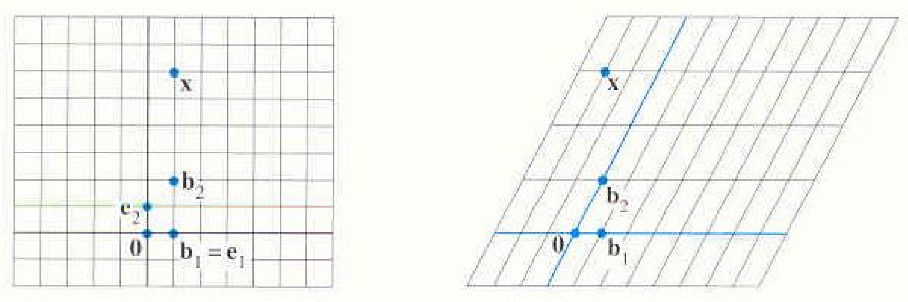
Mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is the coordinate mapping defined by $\mathcal{B}$.

Revision: Coordinate systems

Example: Standard basis $\{\mathbf{e}_1; \mathbf{e}_2\}$, and basis $\mathcal{B} = \{\mathbf{b}_1; \mathbf{b}_2\}$ with $\mathbf{b}_1 = \mathbf{e}_1$ and $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, for vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$.

Coordinates $\begin{bmatrix} 1 \\ 6 \end{bmatrix}$ locate $\mathbf{x}$ relative to the standard basis.

$\mathcal{B}$-coordinates $[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$ locate $\mathbf{x}$ relative to $\mathcal{B}$.



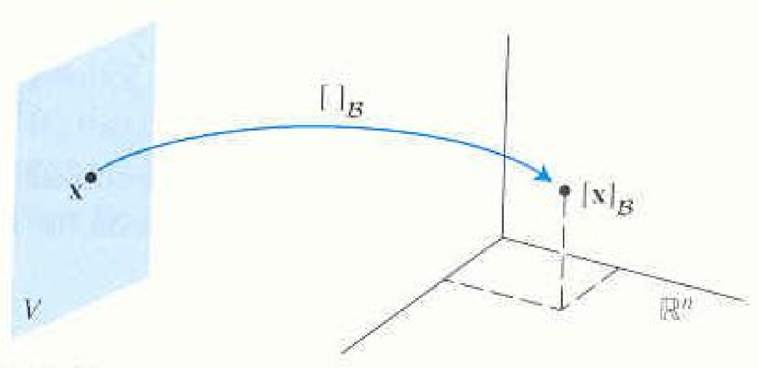
Coordinate mapping

By choosing a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ in a space V we introduce a coordinate system in V.

Coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ maps V to the space $\mathbb{R}^n$.

Points in V can then be identified by equivalent points in $\mathbb{R}^n$.

Coordinate change is a one-to-one linear transformation.



Coordinate mapping

Theorem: Let $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ be a basis for a vector space V. Then coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from V onto $\mathbb{R}^n$.

Proof: Consider two arbitrary vectors in V,

$$\mathbf{u} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n,$$
$$\mathbf{v} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n.$$

Then $\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{b}_1 + \ldots + (c_n + d_n)\mathbf{b}_n$ and

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

Thus coordinate mapping preserves addition.

Coordinate mapping

Proof (continuation): If r is an arbitrary scalar, then

$$\begin{aligned} r\mathbf{u} &= r(c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n) = \\ &= (rc_1)\mathbf{b}_1 + \ldots + (rc_n)\mathbf{b}_n, \end{aligned}$$

so

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

Thus coordinate mapping also preserves scalar multiplication.

Therefore, coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a linear mapping.

Since $\mathcal{B}$ is a basis, the columns of $[\mathbf{b}_1 \ldots \mathbf{b}_n]$ are linearly independent, therefore the transformation is one-to-one.

Coordinate mapping

- The linearity of the coordinate mapping extends to linear combinations: for $\mathbf{u}_1, \dots \mathbf{u}_n \in V$ and scalars $c_1, \dots c_n$

  $$[c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n]_{\mathcal{B}} = c_1 [\mathbf{u}_1]_{\mathcal{B}} + \ldots + c_n [\mathbf{u}_n]_{\mathcal{B}}$$

  (the $\mathcal{B}$-coordinates of a linear combination of vectors $\mathbf{u}_1, \dots \mathbf{u}_n$, are the coefficients in the linear combination of their coordinate vectors).

- A one-to-one linear transformation from vector space V onto vector space W is called an *isomorphism* from V onto W.

  Notation and terminology may be different in V and W, however as vector spaces they have identical structures.

- Coordinate mapping is an isomorphism between V and $\mathbb{R}^n$.

Coordinate mapping: Example

Example:

$\mathcal{P} = \{1, t, t^2, t^3\}$ is the standard basis in the space $\mathbb{P}^3$ (polynomials of a degree up to 3).

A typical element $\mathbf{p}$ of $\mathbb{P}^3$ has the form

$$\mathbf{p} = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

This is a linear combination of the standard basis vectors, so

$$[\mathbf{p}]_{\mathcal{P}} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Coordinate mapping $\mathbf{p} \mapsto [\mathbf{p}]_{\mathcal{P}}$ is an isomorphism of $\mathbb{P}^3$ onto $\mathbb{R}^4$.

All vector operations in $\mathbb{P}^n$ correspond to operations in $\mathbb{R}^{(n+1)}$.

Coordinate mapping: Example

Example: Check if these polynomials are linearly dependent:

$$\mathbf{p}_1 = 1 + 2t^2, \quad \mathbf{p}_2 = 4 + t + 5t^2, \quad \mathbf{p}_3 = 3 + 2t$$

Solution: These $\mathbf{p}_i \in \mathbb{P}^2$ so their coordinate vectors are in $\mathbb{R}^3$:

$$[\mathbf{p}_1]_{\mathcal{P}} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad [\mathbf{p}_2]_{\mathcal{P}} = \begin{pmatrix} 4 \\ 1 \\ 5 \end{pmatrix}, \quad [\mathbf{p}_3]_{\mathcal{P}} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}.$$

We then write these vectors as columns of a matrix, and check the linear dependence $x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + x_3\mathbf{p}_3 = \mathbf{0}$ by row reduction:

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 2 \\ 2 & 5 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 2 \\ 0 & -3 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

No pivot in the third column so the columns are linearly dependent.

Coordinate mapping: Example

So

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

Here, x_3 is a free variable, and then $x_2 = -2x_3$ and $x_1 = 5x_3$.

Choosing $x_3 = 1$, we have $x_1 = 5$ and $x_2 = -2$:

$$5\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} = \mathbf{0}.$$

The corresponding relation $5\mathbf{p}_1 - 2\mathbf{p}_2 + \mathbf{p}_3 = \mathbf{0}$ then reads:

$$5(1 + 2t^2) - 2(4 + t + 5t^2) + (3 + 2t) = 0 \quad \forall\, t$$

Coordinate mapping: Example

Example: Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ be a basis for $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$

$$\text{with} \quad \mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}; \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

Determine if $\mathbf{x} \in H$ and, if so, find $[\mathbf{x}]_{\mathcal{B}}$.

Solution: If $\mathbf{x} \in H$, then the following equation is consistent:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}.$$

The scalars c_1 and c_2, if they exist, are the $\mathcal{B}$ coordinates of $\mathbf{x}$. Row reduction yields

$$\left[\begin{array}{cc|c} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 3 & -1 & 3 \\ 0 & 2 & 6 \\ 0 & 1 & 3 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right].$$

Coordinate mapping: Example

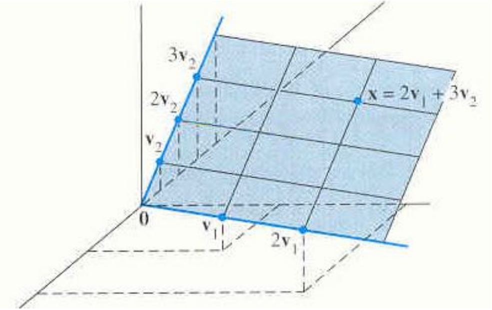
Thus the system is consistent; a unique solution is $c_1 = 2,\ c_2 = 3$.

The coordinate vector of $\mathbf{x}$ relative to $\mathcal{B}$ is $[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix};$$

and

$$\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix} = 2\mathbf{v}_1 + 3\mathbf{v}_2$$



Change of basis in $\mathbb{R}^2$

For a given basis $\mathcal{B}$ in $\mathbb{R}^2$, $\mathcal{B}$-coordinates are easily found.

Example: $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$

To find the coordinate vector $[\mathbf{x}]_{\mathcal{B}} = (c_1; c_2)$ of $\mathbf{x}$ with respect to basis $\mathcal{B}$, we need to solve the equation $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = \mathbf{x}$:

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

which is

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

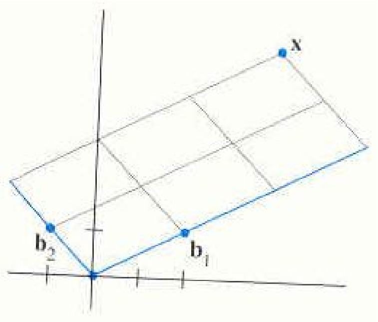
This equation can be solved by row-reducing the augmented matrix; the solution is $c_1 = 3$ and $c_2 = 2$.

Change of basis in $\mathbb{R}^2$

Thus $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$

and we can write

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$



Note: The matrix $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$

transforms $\mathcal{B}$-coordinates of vector $\mathbf{x}$

into standard coordinates for $\mathbf{x}$.

Change of basis in $\mathbb{R}^n$

Generally, let $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ be a basis in $\mathbb{R}^n$.

Construct the following matrix: $\mathrm{P}_{\mathcal{B}} = [\mathbf{b}_1, \ldots \mathbf{b}_n]$.

Then the vector equation $\mathbf{x} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$ is equivalent to

$$\mathbf{x} = \mathrm{P}_{\mathcal{B}} \, [\mathbf{x}]_{\mathcal{B}} \qquad \text{where} \qquad [\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$\mathrm{P}_{\mathcal{B}}$ is the *change of coordinates matrix* from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$: $\{\mathbf{e}_1, \ldots \mathbf{e}_n\}$.

The columns of $\mathrm{P}_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, therefore $\mathrm{P}_{\mathcal{B}}$ is invertible.

Left-multiplication by $\mathrm{P}_{\mathcal{B}}^{-1}$ converts $\mathbf{x}$ into $\mathcal{B}$-coordinate vector

$$\mathrm{P}_{\mathcal{B}}^{-1}\mathbf{x} = [\mathbf{x}]_{\mathcal{B}}$$

Change of basis, a special example

Example: Consider two bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ such that $\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2$ and $\mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2$, and some $\mathbf{x}$, for which we only know that $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, that is,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Even though we do not know the vectors $\mathbf{b}_i$ and $\mathbf{c}_i$, a connection between the two coordinate systems can be established because we know how $\mathbf{b}_1$ and $\mathbf{b}_2$ are related to $\mathbf{c}_1$ and $\mathbf{c}_2$.

Apply $\mathcal{C}$ coordinates to $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$. Since the mapping is linear,

$$[\mathbf{x}]_{\mathcal{C}} = [3\mathbf{b}_1 + \mathbf{b}_2]_{\mathcal{C}} = 3\,[\mathbf{b}_1]_{\mathcal{C}} + [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

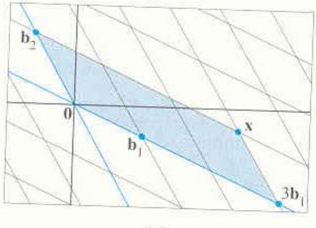
Change of basis, a special example

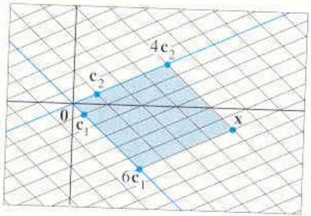
Given that $\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2$ and $\mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2$, the columns are:

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{pmatrix} -6 \\ 1 \end{pmatrix} \qquad \Rightarrow \qquad \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

So the coordinates are connected by the above matrix, and

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$





On the left, $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$. On the right, the same $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$.

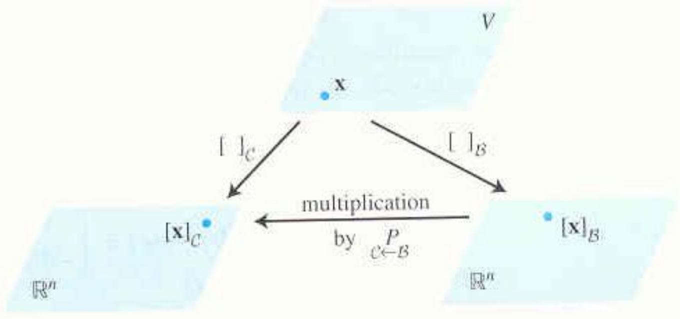
Change of basis

Let $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots \mathbf{c}_n\}$ be bases of space V. Then there is a unique $n \times n$ matrix $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$ such that

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

The columns of $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$ are the $\mathcal{C}$-coordinate vectors of the vectors in basis $\mathcal{B}$. That is, $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}} = \left[\ [\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}} \quad \ldots \quad [\mathbf{b}_n]_{\mathcal{C}} \ \right]$.

$\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$ is called the *change of coordinate matrix* from $\mathcal{B}$ to $\mathcal{C}$.



Change of basis

The columns of $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ are linearly independent because they are the coordinate vectors of the linearly independent set $\mathcal{B}$.
$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ is a square matrix, and coordinate mapping is one-to-one, therefore $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible. By multiplying both the sides of

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

by $(\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$ we obtain

$$(\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}\;[\mathbf{x}]_{\mathcal{C}} = [\mathbf{x}]_{\mathcal{B}}.$$

Therefore the matrix $(\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$ converts $\mathcal{C}$ coordinates into $\mathcal{B}$ coordinates:

$$(\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Note: For the standard basis $\mathcal{E} = \{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ in $\mathbb{R}^n$, each $[\mathbf{b}_i]_{\mathcal{E}} = \mathbf{b}_i$. Then $\mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}}$ (sometimes denoted as $\mathbf{P}_{\mathcal{B}}$) is:

$$\mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}} \equiv \mathbf{P}_{\mathcal{B}} = \left[\mathbf{b}_1, \ldots \mathbf{b}_n\right] = \Big[[\mathbf{b}_1]_{\mathcal{E}}, \ldots [\mathbf{b}_n]_{\mathcal{E}}\Big].$$

Change of basis

Example: Find $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ from $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ to $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, where

$$\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}.$$

Solution: The columns of $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ are $\mathcal{C}$-coordinates of $\mathbf{b}_1$ and $\mathbf{b}_2$.

Let $[\mathbf{b}_1]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $[\mathbf{b}_2]_{\mathcal{C}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Then $\begin{cases} \mathbf{b}_1 = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 \\ \mathbf{b}_2 = y_1\mathbf{c}_1 + y_2\mathbf{c}_2 \end{cases}$

or, in matrix-vector form,

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{b}_1, \qquad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{b}_2.$$

To find x_1, x_2, y_1, y_2 at once, we form the doubly augmented matrix: $\left[\begin{array}{cc|cc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{array}\right]$.

Change of basis

$$\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}.$$

$$\left[\mathbf{c}_1\ \mathbf{c}_2 \mid \mathbf{b}_1\ \mathbf{b}_2\right] = \left[\begin{array}{rr|rr} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rr|rr} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right] \rightarrow \left[\begin{array}{rr|rr} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right]$$

Thus

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ -5 \end{pmatrix}, \qquad [\mathbf{b}_2]_{\mathcal{C}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$$

so the change of coordinate matrix is

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}.$$

Change of basis

Note:

$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ appeared in the right-hand side upon row reduction:

$$\left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right]$$

The first column of $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ results from row reducing $[\,\mathbf{c}_1\ \mathbf{c}_2 \mid \mathbf{b}_1\,]$ to $\left[\,\mathbf{I} \mid [\mathbf{b}_1]_{\mathcal{C}}\right]$ and likewise for the second column of $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$, so

$$\left[\begin{array}{cc|cc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{array}\right] \to \left[\begin{array}{c|c} \mathbf{I} & \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} \end{array}\right]$$

This procedure is used for changing between any two bases in $\mathbb{R}^n$.

Change of basis

Another way to change coordinates from $\mathcal{B}$ to $\mathcal{C}$ is to combine the transitions into standard coordinates, realised by $\mathbf{P}_{\mathcal{B}}$ and $\mathbf{P}_{\mathcal{C}}$.

$$\forall\, \mathbf{x} \in \mathbb{R}^n \qquad \mathbf{P}_{\mathcal{B}}\, [\mathbf{x}]_{\mathcal{B}} = \mathbf{x} \qquad \mathbf{P}_{\mathcal{C}}\, [\mathbf{x}]_{\mathcal{C}} = \mathbf{x}$$

From the last relation we obtain

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}}^{-1}\, \mathbf{x}$$

and subsequently using the first relation we obtain

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

Therefore

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{P}_{\mathcal{C}}^{-1}\, \mathbf{P}_{\mathcal{B}} \qquad \text{and} \qquad \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} = \mathbf{P}_{\mathcal{B}}^{-1}\, \mathbf{P}_{\mathcal{C}}$$

This approach is however slower than the direct transformation.

Change of basis

Example: Find $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ by using $\mathbf{P}_{\mathcal{C}}$ and $\mathbf{P}_{\mathcal{B}}$ for the two bases:

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \; \mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}; \quad \mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}, \; \mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}.$$

We first construct $\mathbf{P}_{\mathcal{C}}$ and $\mathbf{P}_{\mathcal{B}}$ and then use $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = (\mathbf{P}_{\mathcal{C}})^{-1}\mathbf{P}_{\mathcal{B}}$.

$$\mathbf{P}_{\mathcal{B}} = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}, \qquad \mathbf{P}_{\mathcal{C}} = \begin{bmatrix} -7 & -5 \\ 9 & 7 \end{bmatrix}.$$

Then

$$\mathbf{P}_{\mathcal{C}}^{-1} = \begin{bmatrix} -7/4 & -5/4 \\ 9/4 & 7/4 \end{bmatrix}.$$

and so

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

Dimension of a vector space

We have seen that a vector space V with the basis containing n vectors is isomorphic to $\mathbb{R}^n$. This number n is an intrinsic property of V (called the *dimension* of V) which does not depend on a particular choice of basis vectors.

Definition: If V is spanned by a finite set, then V is called a *finite-dimensional* space, and the dimension of V, written as $\dim V = n$, is the number n of vectors in a basis for V.

> **Note:** the dimension is a number of vectors in a basis, **not** a number of entries in each vector.

The dimension of the $\{\mathbf{0}\}$ vector space is defined to be zero.

If V is not spanned by a finite set then V is *infinite-dimensional*.

Dimension of a vector space

Example: The basis for $\mathbb{R}^n$ contains n vectors, so $\dim \mathbb{R}^n = n$.

Example: For standard polynomial basis $\{1, t, t^2\}$, $\dim \mathbb{P}^2 = 3$.

In general, $\dim \mathbb{P}^n = n + 1$.

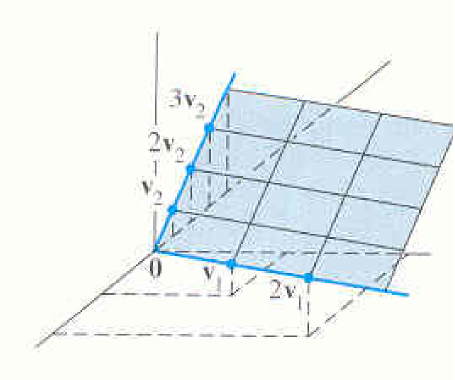
The space $\mathbb{P}$ of all polynomials is infinite-dimensional.

Example:

Let $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, with

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for H, since $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. So $\dim H = 2$.



Dimension of a subspace

Theorem: Let H be a subspace of a finite-dimensional space V. Then any linearly independent set in H can be expanded to a basis for H, and $\dim H \leqslant \dim V$.

Proof: If $H = \{\mathbf{0}\}$, then certainly $\dim H = 0 \leqslant \dim V$.

Otherwise let $S = \{\mathbf{u}_1 \ldots \mathbf{u}_k\}$ be a linearly independent set in H.

If S spans H then S is a basis for H.

Otherwise there is a some $\mathbf{u}_{k+1}$ in H that is not in span of S.

Then the set $S' = \{\mathbf{u}_1, \ldots \mathbf{u}_k, \mathbf{u}_{k+1}\}$ is linearly independent.

As long as the new set S does not span subspace H we can continue to add linearly independent vectors to S, expanding it to a larger linearly independent set in H.

But the number of vectors can never exceed $\dim V$, which is a number of linearly independent vectors in the entire space V.

Eventually the expanded set S will span H and $\dim H \leqslant \dim V$.

Dimension of a subspace

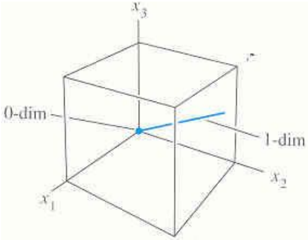
Example: Various subspaces of $\mathbb{R}^3$:

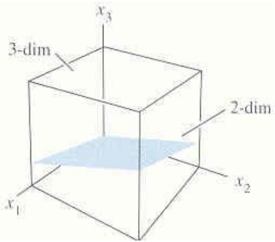
(a) *0*-dimensional subspace: Only the zero subspace.

(b) *1*-dimensional subspaces: Any subspace spanned by a single nonzero vector. Such subspaces are lines through the origin.

(c) *2*-dimensional subspaces: Any subspace spanned by two linearly independent vectors. Planes through the origin.

(d) *3*-dimensional subspace: Only $\mathbb{R}^3$ itself. Any three linearly independent vectors in $\mathbb{R}^3$ span the entire $\mathbb{R}^3$.





Dimension of a subspace

Example: Find the dimension of a subspace defined as

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} \right\} \qquad a, b, c, d \in \mathbb{R}.$$

Solution: Decomposing these vectors, we have

$$\begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} = a \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Therefore H is the set of all linear combination of vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}.$$

Dimension of a subspace

By analysing these vectors, we notice the following

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}.$$

- $\mathbf{v}_1 \neq \mathbf{0}$
- $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$.
- $\mathbf{v}_3$ is a multiple of $\mathbf{v}_2$ (since $\mathbf{v}_3 = -2\mathbf{v}_2$).
  Using the spanning set theorem, we may discard $\mathbf{v}_3$ and have the remaining set still spanning H.
- $\mathbf{v}_4$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is a linearly independent set and a basis for H.

Thus $\dim H = 3$.

Supplementary theorems

Theorem: If a vector space V has a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ with n vectors, then any set in V containing more than n vectors must be linearly dependent.

Note:

This theorem implies that if the vector space V has a basis with n vectors, then any linearly independent set in V must have no more than n vectors.

Proof strategy:

- Create isomorphism to $\mathbb{R}^n$.
- Prove in $\mathbb{R}^n$ using matrix techniques.
- Apply a reverse isomorphism.

Supplementary theorems

Proof: Let $\{\mathbf{u}_1, \dots \mathbf{u}_p\}$ be a set in V with $p > n$ vectors. The coordinate vectors $[\mathbf{u}_1]_{\mathcal{B}}, \dots [\mathbf{u}_p]_{\mathcal{B}}$ are linearly dependent in $\mathbb{R}^n$ because there are more vectors (p) than entries (n). Therefore $\exists\ c_1, \dots c_p$ not all zero such that

$$c_1 [\mathbf{u}_1]_{\mathcal{B}} + \dots + c_p [\mathbf{u}_p]_{\mathcal{B}} = \mathbf{0}.$$

The coordinate mapping is a linear transformation $V \mapsto \mathbb{R}^n$, so we therefore have

$$[c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p]_{\mathcal{B}} = \mathbf{0} = [\mathbf{0}]_{\mathcal{B}} \qquad \text{in} \quad \mathbb{R}^n.$$

Since the transformation is one-to-one, we must have

$$c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p = \mathbf{0} \qquad \text{in} \quad V$$

Since the c_i are not all zero, $\mathbf{u}_1, \dots \mathbf{u}_p$ are linearly dependent.

Supplementary theorems

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

Proof strategy: Double inequality: $a \leqslant b$ and $a \geqslant b \implies a = b$ (show that upper and lower bounds coincide)

Proof: Let $\mathcal{B}_1$ be a basis in V with n vectors, and $\mathcal{B}_2$ be any other basis in V with m vectors.

Since $\mathcal{B}_1$ is a basis and $\mathcal{B}_2$ is a linearly independent set, from the previous theorem $\mathcal{B}_2$ has no more than n vectors, so $m \leqslant n$.

Vice versa, since $\mathcal{B}_2$ is a basis and $\mathcal{B}_1$ is a linearly independent set, $\mathcal{B}_1$ has to have no more than m vectors, so $n \leqslant m$.

Thus $m = n$ and $\mathcal{B}_2$ consists of exactly n vectors.

Note: If a non-zero space V is spanned by a finite set S then there is a subset of S that is a basis in V by spanning set theorem.

Supplementary theorems

"The basis theorem":

Let V be a p-dimensional vector space, $p \geqslant 1$. Then:

- any linearly independent set of p elements in V is a basis for V;
- any set of p elements that spans V is a basis for V.

Proof:

Any linearly independent set S of p elements can be extended to a basis for V. But this basis should contain exactly p elements since $\dim V = p$. So S is a basis for V.

Suppose S has p elements and spans V. Since V is nonzero, the spanning set theorem implies that there exists a subset S' of S which is a basis of V. Since $\dim V = p$, S' must contain p vectors. Hence $S = S'$.

Next lecture

See you next Wednesday

15 May 2019

Assignment 5 is due this week (on 8–10 May)

Written assignment (essay) is due next week (on 15–17 May)