Fundamentals of linear algebra

- Revision: linear dependence and independence
- Linearly independent sets
- Basis of a vector space
- Spanning set theorem
- Coordinate systems
- Linear Transformations

Wednesday, 1 May 2019

Revision: Linear independence

- A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_m\}$ is *linearly independent* if

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$$

  has only the trivial solution (with all $c_i = 0$).

Theorem:

- A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_m\}$ of two or more vectors in $\mathbb{R}^n$ is linearly dependent if and only if at least one of the vectors in S is a linear combination of the other vectors in S.
- If S is a linearly dependent set, then some $\mathbf{v}_j$ is a linear combination of the preceding vectors $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_{j-1}$.

Theorem:

- Any set $\{\mathbf{v}_1, \dots \mathbf{v}_p\} \in \mathbb{R}^n$ is linearly dependent if $p > n$.

Revision: Vector spaces

A *vector space* in $\mathbb{R}^n$ is a non-empty set V of vectors, on which two operations — addition and multiplication by real scalars — are defined, subject to these axioms:

(i) $(\mathbf{u}+\mathbf{v}) \in V$

(ii) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$

(iii) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$

(iv) $\exists\, \mathbf{0}$ such that $\mathbf{u}+\mathbf{0} = \mathbf{0}+\mathbf{u} = \mathbf{u}$

(v) $\forall\, \mathbf{u}\ \exists\,(-\mathbf{u})$ such that $\mathbf{u}+(-\mathbf{u}) = (-\mathbf{u})+\mathbf{u} = \mathbf{0}$

(vi) $c\,\mathbf{u} \in V$

(vii) $c\,(\mathbf{u}+\mathbf{v}) = c\,\mathbf{u}+c\,\mathbf{v}$

(viii) $(c+d)\mathbf{u} = c\,\mathbf{u}+d\,\mathbf{u}$

(ix) $c\,(d\,\mathbf{u}) = (c\,d)\,\mathbf{u}$

(x) $1\mathbf{u} = \mathbf{u}, \quad (-\mathbf{u}) = (-1)\mathbf{u}$

These rules must hold for all $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in V$ and for any c, $d \in \mathbb{R}$.

Revision: Subspaces

Definition: A **subspace** H of a vector space V is a subset of vectors with the following properties:

- H includes the zero vector of V
- H is closed under vector addition: $\forall\,(\mathbf{u},\mathbf{v}) \in H,\ \ \mathbf{u}+\mathbf{v} \in H$
- H is closed under multiplication by scalars:
  $\forall\,\mathbf{u} \in H$ and $\forall\,c \in \mathbb{R}$, $c\mathbf{v} \in H$

Every subspace is a vector space and satisfies the ten axioms.

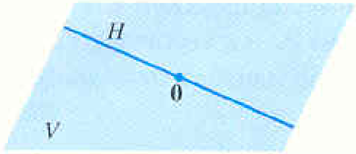


FIGURE 6

A subspace of V.

Subspaces spanned by a set

Theorem:

For $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_p \in V$, $\operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_p\}$ is a subspace of V.

This is called a subspace spanned (generated) by $\{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_p\}$.

Example 1: Let H be a set of all vectors of the form $\big[(a-3b);\ (b-a);\ a;\ b\big]$ where a, b are arbitrary scalars:

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \equiv a\mathbf{v}_1 + b\mathbf{v}_2.$$

This rearrangement demonstrates that $H = \operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Therefore, H is a subspace of $\mathbb{R}^4$ generated by $\mathbf{v}_1$ and $\mathbf{v}_2$.

Subspaces spanned by a set

Example 2: Find h such that $\mathbf{y} \in \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$, if

$$\mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}; \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}.$$

Solution: vector $\mathbf{y} \in \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if $\mathbf{y} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$:

$$x_1 \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}.$$

This vector equation corresponds to the augmented matrix

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}$$

which is only consistent if $h = 5$, so then $\mathbf{y} = [-4; 3; 5]$.

Subspaces spanned by a set

Continue reduction towards REF, taking into account $h = 5$:

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 7 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

which is now equivalent to the system

$$\begin{cases} x_1 + 7x_3 = 1 \\ x_2 - 2x_3 = -1 \end{cases} \quad \Bigg| \quad \begin{aligned} x_1 &= 1 - 7x_3 \\ x_2 &= -1 + 2x_3 \end{aligned}$$

Setting the free variable $x_3 = 0$, we get $x_1 = 1$ and $x_2 = -1$.

Thus $\mathbf{y} = 1\mathbf{v}_1 - 1\mathbf{v}_2 + 0\mathbf{v}_3 = \mathbf{v}_1 - \mathbf{v}_2$, which is easily checked:

$$\begin{bmatrix} -4 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} - \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}.$$

Linearly independent sets

To identify subsets that span a vector space V or its subspace H, we use linear independence defined in the same way as in $\mathbb{R}^n$.

Definition: An indexed set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_p\} \in V$ is *linearly independent* if the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution ($c_1 = 0, \ldots c_p = 0$).

Conversely, the set is *linearly dependent* if there there are $c_1, \ldots c_p$ not all equal to zero, such that the above equation holds.

Theorem:

An indexed set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_p\}$ of $p \geqslant 2$ vectors with $\mathbf{v}_1 \neq \mathbf{0}$ is linearly dependent if and only if some $\mathbf{v}_j$ with $j > 1$ is a linear combination of the preceding vectors $\mathbf{v}_1, \ldots \mathbf{v}_{j-1}$.

Linearly independent sets: Examples

- **Example 1:** Given $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$, and $\mathbf{p}_3(t) = 4 - t$, this polynomial set $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is linearly dependent in $\mathcal{P}$, because $\mathbf{p}_3 = 4\mathbf{p}_1 - \mathbf{p}_2$, that is, $4\mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3 = 0$ ("vectors" are linearly dependent with weights 4, -1, -1).

- **Example 2:** The set $\{\sin t, \cos t\}$ is linearly independent in space $\mathcal{C}_{[0,\, 2\pi]}$ (continuous functions on $0 \leqslant t \leqslant 2\pi$ interval).

  The relation $c_1 \sin t + c_2 \cos t = 0$ only has the trivial solution:

  There is no scalar c such that $\cos t = c \sin t \quad \forall\, t \in [0,\, 2\pi]$, so functions $\sin t$ and $\cos t$ are not multiples of each other.

- **Example 3:** The set $\{\sin t \cos t,\, \sin 2t\}$ is linearly dependent in $\mathcal{C}_{[0,\, \pi]}$, because $\sin 2t = 2 \sin t \cos t \;\; \forall\, t \in [0,\, \pi]$, and the functions are linearly dependent: $2 \sin t \cos t - \sin 2t = 0$.

  The weights in this linear dependence are 2 and -1.

Basis

Definition: Let H be a subspace of a vector space V.

An indexed set $\mathcal{B} = \{\mathbf{b}_1, \dots \mathbf{b}_p\}$ in V is a **basis** in H if

(i) $\mathcal{B}$ is a linearly independent set, and

(ii) $H = \mathrm{Span}\{\mathbf{b}_1, \dots \mathbf{b}_p\}$

Notes:

- The definition of a basis also applies for $H = V$, because any vector space is a subspace of itself.
- So, a basis of V is a linearly independent set that spans V.
- When $H \neq V$ condition (ii) includes the requirement that each of the vectors $\mathbf{b}_1, \dots \mathbf{b}_p \in H$ because $\mathrm{Span}\,\mathcal{B}$ contains all these vectors.

Basis: examples

Example 1: Determine if the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$ if

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

Solution: Check that the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0} \quad \Rightarrow \quad c_i = 0 \ \forall\, i$$

and spans $\mathbb{R}^3$, so $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{b}$ is consistent $\forall\, \mathbf{b} \in \mathbb{R}^3$.

We therefore need to check if there are pivots in every column of

$$\begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 6 & -2 & 5 \end{bmatrix} \to \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 0 & 6 & 9 \end{bmatrix} \to \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

There are pivots in every column, therefore the homogeneous system only has the trivial solution, and the inhomogeneous system has a unique solution for every $\mathbf{b}$. Thus, $\{\mathbf{v}_i\}$ is a basis for $\mathbb{R}^3$.

Basis: examples

Example 2: Set $\mathcal{S} = \{1, t, t^2, \ldots\; t^n\}$ is a basis for $\mathcal{P}_n$.

This basis is called the *standard basis* for polynomial space $\mathcal{P}_n$.

Proof: It is obvious that any polynomial of degree at most n can be written as a combination of the members of $\mathcal{S}$.

Suppose that coefficients $c_0, \ldots c_n$ satisfy

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \ldots + c_n t^n = 0 \qquad \forall\, t$$

However a polynomial of degree n has at most n zeros.

Therefore the above relation can only be satisfied if $c_i = 0\ \forall\, i$, which means that the set $\mathcal{S}$ is linearly independent.

Basis: examples

Example 3: Consider $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}$$

It is easy to see that $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, therefore $\forall\, \mathbf{u} \in H$

$$\begin{aligned} \mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(5\mathbf{v}_1 + 3\mathbf{v}_2) \\ &= (c_1 + 5c_3)\mathbf{v}_1 + (c_2 + 3c_3)\mathbf{v}_2. \end{aligned}$$

Thus $\mathbf{u} \in \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and so H is identical to $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

In other words, it turns out $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Clearly, every vector in $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ belongs to H because

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3.$$

Note that $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent ($\mathbf{v}_1 \neq c\mathbf{v}_2$).

Therefore, we can conclude that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for H.

Spanning set theorem

- 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.

Spanning set theorem:

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

(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.

Spanning set theorem (proof)

(a): Suppose $\mathbf{v}_p = a_1\mathbf{v}_1 + \ldots + a_{p-1}\mathbf{v}_{p-1}$. Then $\forall\, \mathbf{x} \in H$,

$$\begin{aligned}
\mathbf{x} &= c_1\mathbf{v}_1 + \ldots + c_{p-1}\mathbf{v}_{p-1} + c_p\mathbf{v}_p \\
&= c_1\mathbf{v}_1 + \ldots + c_{p-1}\mathbf{v}_{p-1} + c_p(a_1\mathbf{v}_1 + \ldots + a_{p-1}\mathbf{v}_{p-1}) \\
&= (c_1 + c_p a_1)\mathbf{v}_1 + \ldots + (c_{p-1} + c_p a_{p-1})\mathbf{v}_{p-1}
\end{aligned}$$

Thus $\{\mathbf{v}_1, \ldots\, \mathbf{v}_{p-1}\}$ spans H, because this holds $\forall\, \mathbf{x} \in H$.

(b): If $\mathcal{S}$ is linearly independent, then it is already a basis for H.

- If this is not the case then one of the vectors can be dropped (part a). We can continue to drop vectors until the remaining set is linearly independent and hence is a basis for H.
- If the set is eventually reduced to one vector, that vector will be non-zero because $H \neq \{\mathbf{0}\}$. Since a single non-zero vector $\mathbf{v}$ is linearly independent, it will be a basis for H.

Spanning set theorem (notes)

Notes:

- 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.

A linearly independent set can be enlarged to form a basis, but further enlargement destroys the linear independence.

A larger spanning set can be reduced to a basis, but further shrinking destroys the spanning property.

Spanning set theorem (example)

A linearly independent set which does **not** span $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}.$$

A basis for $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\}.$$

A set that spans $\mathbb{R}^3$ but is linearly dependent:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}.$$

(Any) 3 linearly independent vectors form a basis for $\mathbb{R}^3$.

2 (or less) are not sufficient, while 4 (or more) are too many.

Coordinate systems

- An important reason for specifying a basis $\mathcal{B}$ for a vector space V is to impose a coordinate system on V.
- If $\mathcal{B}$ for V contains n vectors, then a coordinate system will make V behave like $\mathbb{R}^n$.

Theorem: (the *unique representation theorem*)

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

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

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

Coordinate systems

Proof: As $V = \operatorname{Span}\mathcal{B}$ there exists a set of scalars $\{c_i\}^n$ such that

$$\mathbf{x} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n.$$

Suppose another set $\{d_i\}^n$ also satisfies $\mathbf{x} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n$.

Then we can write

$$\begin{aligned}\mathbf{0} &= \mathbf{x} - \mathbf{x} \\ &= c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n - d_1\mathbf{b}_1 - \ldots - d_n\mathbf{b}_n \\ &= (c_1 - d_1)\mathbf{b}_1 + \ldots + (c_n - d_n)\mathbf{b}_n.\end{aligned}$$

However, because $\mathcal{B}$ is a linearly independent set, the $(c_i - d_i)$ coefficients must be zero $\forall\, i$:

$$c_i = d_i \qquad 1 \leqslant i \leqslant n.$$

Therefore, representation $\mathbf{x} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$ is unique.

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 $\mathbb{R}^n$

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

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 $\mathcal{B}$.

Coordinate systems: Examples

Example 1a: Consider a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for $\mathbb{R}^2$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Let $\mathbf{x} \in \mathbb{R}^2$ have the following $\mathcal{B}$-coordinate vector:

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

The $\mathcal{B}$-coordinates of $\mathbf{x}$ directly produce $\mathbf{x}$ from the vectors of $\mathcal{B}$:

$$\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2 = -2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}.$$

Coordinate systems: Examples

Example 1b: The same basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, and $\mathbf{x} \in \mathbb{R}^2$:

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \qquad \mathbf{x} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

To find the $\mathcal{B}$-coordinates for a vector $\mathbf{x}$, we need to solve

$$\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 1 & 6 \\ 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 4 \\ x_2 = 2 \end{cases} \quad \text{thus} \quad [\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

This can be easily verified: $\mathbf{x} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$

Coordinate systems: Examples

Example 2: Coordinates of $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ in the basis

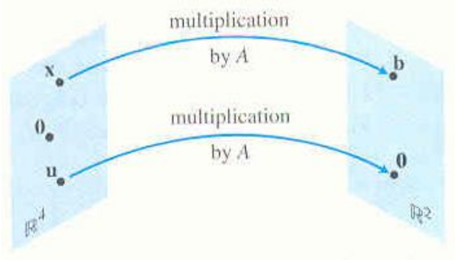
$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

are obvious, but can be formally found from

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 6 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 4 \\ x_2 = 5 \\ x_3 = 6 \end{cases} \quad \text{so} \quad [\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

Linear transformations

- Equation $\mathbf{Ax} = \mathbf{b}$ can be regarded as a mapping (function).
- Matrix $\mathbf{A}$ acts on $\mathbf{x}$, producing a new vector $\mathbf{b}$, which is analogous to an action of a function $y = f(x)$.
- The resulting correspondence between $\mathbf{x}$ and $\mathbf{b}$ is then a mapping from one set of vectors to another.
- Left-multiplication by $\mathbf{A}$ *transforms* $\mathbf{x}$ into $\mathbf{b}$.



Linear transformations

Example:

Solving an equation $\mathbf{A}\mathbf{x} = \mathbf{b}$, such as

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

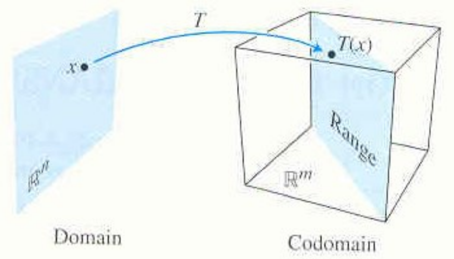
$$\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

amounts to finding all $\mathbf{x} \in \mathbb{R}^4$ that are transformed into $\mathbf{b} \in \mathbb{R}^2$.

Linear 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 $\mathcal{T}$.



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))

Linear Transformation

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

Proof: given that $\mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$

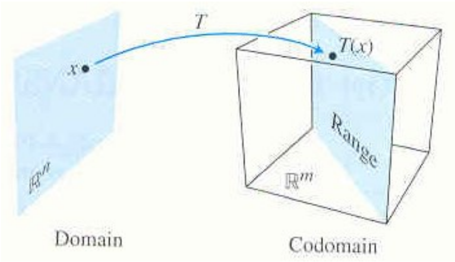
$$\begin{aligned}\mathcal{T}(\mathbf{x}+\mathbf{y}) &= (x_1+y_1)\mathbf{a}_1 + \ldots + (x_n+y_n)\mathbf{a}_n \\ &= x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n + y_1\mathbf{a}_1 + \ldots + y_n\mathbf{a}_n = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})\end{aligned}$$

$$\begin{aligned}\mathcal{T}(c\,\mathbf{x}) &= (cx_1)\mathbf{a}_1 + \ldots + (cx_n)\mathbf{a}_n \\ &= c\,(x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n) = c\,T(\mathbf{x})\end{aligned}$$

Thus $\mathcal{T}(\mathbf{x}+\mathbf{y}) = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})$ and $\mathcal{T}(c\,\mathbf{x}) = c\mathcal{T}(\mathbf{x})$
so $\mathcal{T} = \mathbf{A}\mathbf{x}$ is a linear transformation, by definition.

Linear transformations

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



Linear transformations

Example: Find the image of $\mathbf{u}$ obtained with $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

The transformation is given by $\mathcal{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ so

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

$$\mathcal{T}(\mathbf{u}) = \mathbf{A}\mathbf{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

This is a $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^3$ transformation.

Linear transformations

Example: Let $\mathcal{T}(x_1, x_2) = \big((2x_1 - 3x_2);\ (x_1 + 4);\ (5x_2)\big)$.

We can show that the above $\mathcal{T}$ is *not a linear* transformation:

$$\mathcal{T}(x_1, x_2) = \begin{bmatrix} 2x_1 - 3x_2 \\ x_1 + 4 \\ 5x_2 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}.$$

Thus $\mathcal{T}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{q}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 5 \end{bmatrix}, \quad \text{and} \quad \mathbf{q} = \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}.$$

If $\mathcal{T}$ were a linear transformation, then $\mathcal{T}(\mathbf{0}) = \mathbf{0}$.

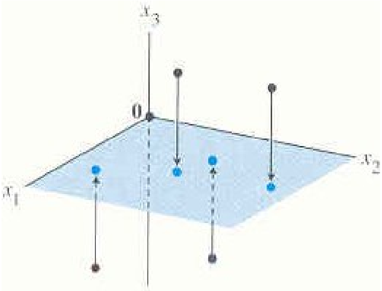
However, here $\mathcal{T}(\mathbf{0}) = \mathbf{q} \neq \mathbf{0}$.

Linear transformations

Example:

Consider the following matrix:

$$\mathbf{O_{12}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



Action of this matrix on $\mathbf{x} \in \mathbb{R}^3$ is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

This transformation projects points of $\mathbb{R}^3$ into $\mathbb{R}^2$.

This is called a *projection* transformation.

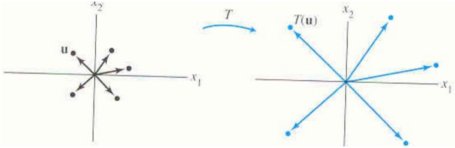
Linear transformations

Example: Let $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$, $\mathcal{T}(\mathbf{x}) = r\,\mathbf{x}$, where $r \in \mathbb{R}$.

This is clearly a linear transformation:

$$\begin{aligned}\mathcal{T}(c\,\mathbf{u} + d\,\mathbf{v}) &= r(c\,\mathbf{u} + d\,\mathbf{v}) \\ &= c(r\,\mathbf{u}) + d(r\,\mathbf{v}) = c\,\mathcal{T}(\mathbf{u}) + d\,\mathcal{T}(\mathbf{v}).\end{aligned}$$

$\mathcal{T}$ is called a *dilation* (or scaling) transformation; matrix $\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$.

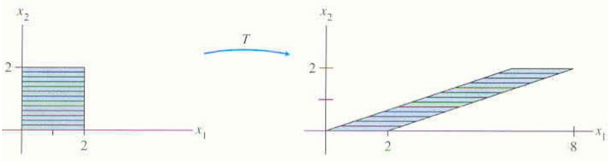


Linear transformations

Example: Let $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be achieved with matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

This is called a *shear* transformation (a square to a parallelogram):



$$\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}.$$

Next lecture

See you next Wednesday

8 May 2019

Assignment 4 is due this week (on 1–3 May)

Written assignment (essay) is due Week 9 (on 15–17 May)