Linear Transformations

Definition 1: *Linear Transformation*

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces. A **linear transformation** from $\mathbb{V}$ to $\mathbb{W}$ is a function $f : \mathbb{V} \longrightarrow \mathbb{W}$, such that:

1. For all $\vec{v}_1, \vec{v}_2 \in \mathbb{V}$, $f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)$.
2. For all $\vec{v} \in \mathbb{V}$ and $r \in \mathbb{R}$ $f(r.\vec{v}) = r.f(\vec{v}_1)$.

Example 2: *Linear Transformations from* $\mathbb{R}^3$ *to itself*

Heuristically, a linear transformation $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ can be thought of as a way of "warping" 3-dimensional space, in a manner such that many geometric properties are preserved. **Straight lines** are transformed into **straight lines**, and **flat planes** into **flat planes**. Two lines (or planes) which were **parallel** before the transformation will remain **parallel** afterwards. Finally, the **origin** point is unmoved by the transformation.

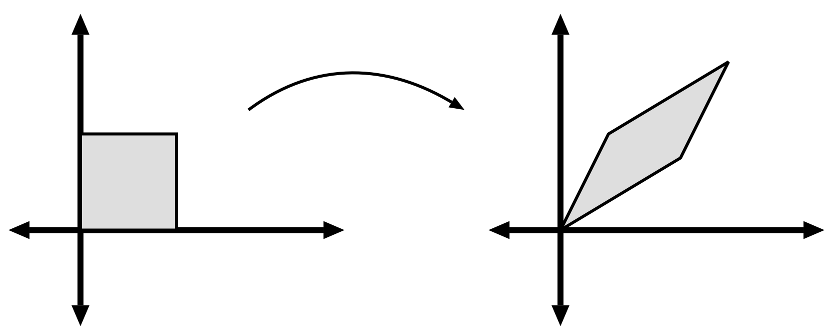

Figure 1: Linear functions send parallelograms to parallelograms, and send zero to itself

Thus, a rectilinear 3-dimensional "box" with one corner at the origin will get transformed, under the action of f, into a rather squashed looking box (with parallelograms for sides), still with one corner at the origin. ———

Example 3: *Matrix Multiplication*

Let $\boxed{\mathrm{A}} \in \mathcal{M}_{D\times M}$. We can define a linear transformation $f : \mathbb{R}^M \longrightarrow \mathbb{R}^D$ via **multiplication by** $\boxed{\mathbf{A}}$: for any $\vec{v} \in \mathbb{R}^M$, define:

$$f(\vec{v}) \; = \; \boxed{\mathrm{A}} \cdot \vec{v}$$

Explicitly, suppose that $\boxed{\mathrm{A}} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1M} \\ a_{21} & a_{22} & \ldots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{D1} & a_{D2} & \ldots & a_{DM} \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} v_1 & v_2 & \ldots & v_M \end{bmatrix}$.

Then

$$\begin{aligned} f(\vec{v}) \; &= \; \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1M} \\ a_{21} & a_{22} & \ldots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{D1} & a_{D2} & \ldots & a_{DM} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix} \\ &= \; \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \ldots + a_{1M}v_M \\ a_{21}v_1 + a_{22}v_2 + \ldots + a_{2M}v_M \\ \vdots \\ a_{D1}v_1 + a_{D,2}v_2 + \ldots + a_{D,M}v_M \end{bmatrix} \end{aligned}$$

Theorem 4: *A linear transformation is determined by its action on a Basis*

Let $\mathbb{V}$ be a **finite dimensional** vector space.

Let $\{\mathbf{a}_1, \ldots, \mathbf{a}_K\}$ be any **spanning set** for $\mathbb{V}$. Suppose that $f, g : \mathbb{V} \longrightarrow \mathbb{W}$ are two linear transformations. If $f(\mathbf{a}_k) = g(\mathbf{a}_k)$ for all of $\{\mathbf{a}_1, \ldots, \mathbf{a}_K\}$, then f and g are **equal everywhere**.

Proof: Let $\vec{v} \in \mathbb{V}$ be arbitrary. Since $\{\mathbf{a}_1, \ldots, \mathbf{a}_K\}$ is a spanning set, write:

$$\vec{v} = \sum_{k=1}^{K} v_k \mathbf{a}_k$$

for some numbers $v_1, \ldots, v_k \in \mathbb{R}$. Then

$$\begin{aligned} f(\vec{v}) &=_{(1)} f\left(\sum_{k=1}^{K} v_k \mathbf{a}_k\right) \\ &=_{(2)} \sum_{k=1}^{K} f(v_k \mathbf{a}_k) \\ &=_{(3)} \sum_{k=1}^{K} v_k f(\mathbf{a}_k) \\ &=_{(4)} \sum_{k=1}^{K} v_k g(\mathbf{a}_k) \\ &=_{(5)} \sum_{k=1}^{K} g(v_k \mathbf{a}_k) \\ &=_{(6)} g\left(\sum_{k=1}^{K} v_k \mathbf{a}_k\right) \\ &=_{(7)} g(\vec{v}) \end{aligned}$$

here, (1) and (7) are because $\vec{v} = \sum_{k=1}^{K} v_k \mathbf{a}_k$,

(2),(3) are because f is linear.

(4) is because f and g agree on $\{\mathbf{a}_1, \ldots, \mathbf{a}_K\}$.

(5),(6) are because g is linear.

□ [Theorem 4]

Theorem 5: *All Linear Transformations on $\mathbb{R}^N$ are Matrix Multiplications*

Let $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ be the "standard basis" for $\mathbb{R}^N$.

1. If $f : \mathbb{R}^N \longrightarrow \mathbb{R}^D$ is any linear transformation, then f is equivalent to matrix multiplication:

$$f(\vec{v}) \;=\; \boxed{\mathsf{A}} \cdot \vec{v}$$

where $\boxed{\mathsf{A}}$ is the $D \times N$ matrix whose *columns* are the *images* of $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$. Formally, let $\mathbf{a}_1 = f(\mathbf{e}_1), \mathbf{a}_2 = f(\mathbf{e}_2), \ldots, \mathbf{a}_N = f(\mathbf{e}_N)$. Then

$$\boxed{\mathsf{A}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_N \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$$

2. Let $\boxed{\mathsf{B}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_N \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$ be an arbitrary $D \times N$ matrix. Then there is a **unique** linear transformation $g : \mathbb{R}^N \longrightarrow \mathbb{R}^D$ so that

$$\mathbf{b}_1 = g(\mathbf{e}_1), \mathbf{b}_2 = g(\mathbf{e}_2), \ldots, \mathbf{b}_N = g(\mathbf{e}_N)$$

and this linear transformation is simply **multiplication by** $\boxed{\mathsf{B}}$:

$$g(\vec{v}) \;=\; \boxed{\mathsf{B}} \cdot \vec{v}$$

Proof:

Proof of Part 2: Suppose that $\boxed{\mathrm{B}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_N \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1N} \\ b_{21} & b_{22} & \ldots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{D1} & b_{D2} & \ldots & b_{DN} \end{bmatrix}$.

Then:

$$\boxed{\mathrm{B}} \cdot \vec{v} \;=\; \begin{bmatrix} b_{11}v_1 + b_{12}v_2 + \ldots + b_{1N}v_N \\ b_{21}v_1 + b_{22}v_2 + \ldots + b_{2N}v_N \\ \vdots \\ b_{D1}v_1 + b_{D,2}v_2 + \ldots + b_{D,N}v_N \end{bmatrix}$$

$$= v_1 \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{D1} \end{bmatrix} + v_2 \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{D2} \end{bmatrix} + \ldots + v_N \begin{bmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{Dn} \end{bmatrix}$$

$$= v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \ldots v_N\mathbf{b}_N$$

$$\text{In particular, if} \quad \mathbf{e}_n = \Bigg(\underbrace{0, \ldots, 0}_{n-1}, 1, 0, \ldots, 0\Bigg)$$

$$\text{then} \quad \boxed{\mathrm{B}} \cdot \mathbf{e}_n = \mathbf{b}_n$$

Proof of Part 1: This follows from **Part 2** and the previous theorem.

□ [Theorem 5]

Example 6: *Compression/expansion/reflection in the kth dimension*

Let $r \in \mathbb{R}$, and consider the transformation $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ that multiplies the kth dimension of $\mathbb{R}^N$ by a factor of r:

$$f(v_1, v_2, \ldots, v_k, \ldots, v_N) = (v_1, v_2, \ldots, r.v_k, \ldots, v_N)$$

$$\text{Notice: } \left\{ \begin{array}{c} f(\mathbf{e}_1) = \mathbf{e}_1 = (1, 0, 0, \ldots, 0, \ldots, 0) \\ f(\mathbf{e}_2) = \mathbf{e}_2 = (0, 1, 0, \ldots, 0, \ldots, 0) \\ \ldots \\ f(\mathbf{e}_k) = r.\mathbf{e}_k = (0, 0, 0, \ldots, r, \ldots, 0) \\ \ldots \\ f(\mathbf{e}_N) = \mathbf{e}_N = (0, 0, 0, \ldots, 0, \ldots, 1) \end{array} \right\}, \text{ so the matrix of } f \text{ is}$$

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Geometrically speaking...

- if $r = 0$, then f **annihilates** the kth dimension, thereby projecting $\mathbb{R}^N$ onto a $(N-1)$-dimensional subspace. (see **Part A** of Fig 2 on the facing page)
- if $0 < r < 1$, then f **compresses** the kth dimension. (see **Part B** of Fig 2 on the next page)
- if $r = 1$, then f is the **identity map**. (see **Part C** of Fig 2 on the facing page)
- if $r > 1$, then f **stretches** the kth dimension. (see **Part D** of Fig 2 on the next page)
- if $r = -1$, then f acts to **reflect** the kth dimension. (see **Part E** of Fig 2 on the facing page)
- What happens if $-1 < r < 0$? If $r < -1$?

Example 7: *Rotation in* $\mathbb{R}^2$
Let θ be an angle between 0 and 2π, and let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ "rotate" the plane about the origin by an angle of θ.

f doesn't change the lengths of vectors, so unit vectors get sent to unit vectors. Thus, $\mathbf{a} = f(\mathbf{e}_1)$ is a unit vector, and makes an angle of θ with the horizontal axis. In other words, the vector $\mathbf{a}$ is the **hypotenuse** of a right-angle triangle with sides of length a_1 and a_2, and an angle of θ; by trigonometry, we know that $a_1 = \cos(\theta)$, $a_2 = \sin(\theta)$.

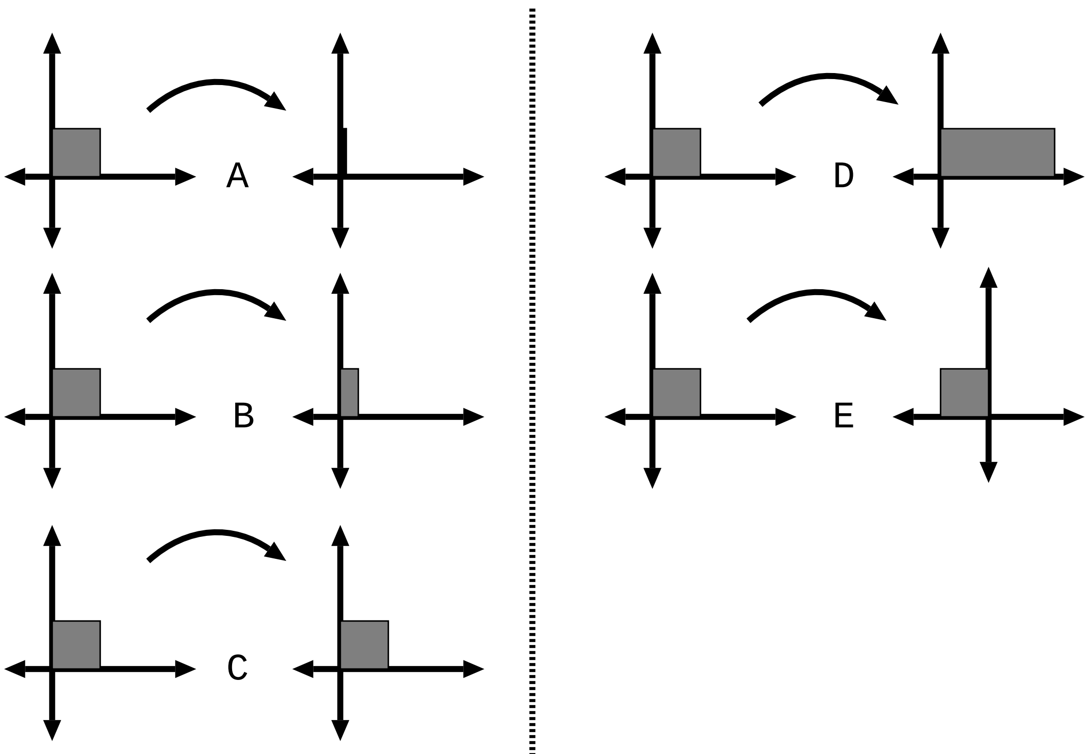

Figure 2: Stretching/Compressing/Reflecting a dimension

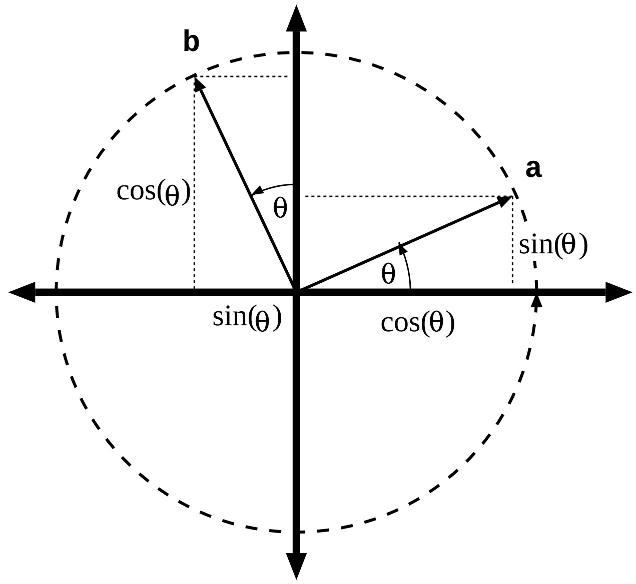

Figure 3: Rotation by θ

Similarly, $\mathbf{b} = f(\mathbf{e}_2)$ is a unit vector which makes an angle of θ with the *vertical* axis, so it is the **hypotenuse** of an upended right-angle triangle with sides of length b_1 and b_2, and an angle of θ; by trigonometry, we know that $b_1 = -\sin(\theta)$, $b_1 = \cos(\theta)$.

Hence, the matrix of f is:

$$\begin{bmatrix} \uparrow & \uparrow \\ f(\mathbf{e}_1) & f(\mathbf{e}_2) \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Example 8: *Orthogonal Projection*

Suppose $\mathbb{V}$ is a subspace of $\mathbb{R}^N$, with orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_D\}$. Then the **orthogonal projection** map $\mathbf{pr}_{\mathbb{V}} : \mathbb{R}^D \longrightarrow \mathbb{R}^D$, defined

$$\mathbf{pr}_{\mathbb{V}}(\vec{x}) = \sum_{d=1}^{D} \langle \vec{x}, \mathbf{v}_d \rangle \, \mathbf{v}_d$$

is a linear transformation. (See Figure 4 on the next page

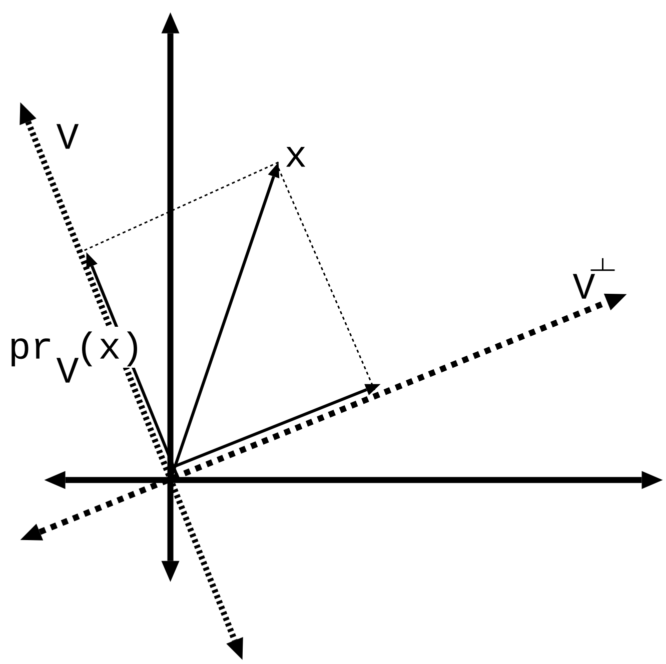

Figure 4: The projection of $\vec{x}$ onto $\mathbb{V}$ is $\mathbf{pr}_{\mathbb{V}}(\vec{x})$

Example 9: *Reflection across a Line*

Let $\mathbb{L}$ be a line in $\mathbb{R}^2$ passing through zero. Thus, $\mathbb{L}$ is a linear subspace. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by

$$f(\vec{x}) \;=\; 2 \cdot \mathbf{pr}_{\mathbb{L}}(\vec{x}) - \vec{x}.$$

This is equivalent to **reflecting** $\vec{x}$ across the line $\mathbb{L}$.

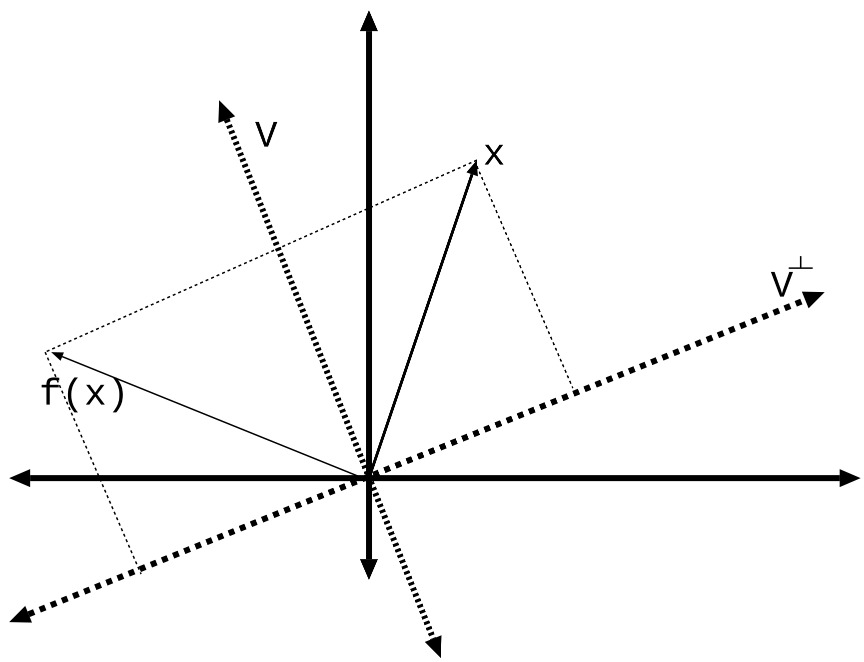

Figure 5: Reflection across a line

Example 10: *Reflection across a Plane*

Let $\mathbb{P}$ be a plane in $\mathbb{R}^3$ passing through zero. Thus, $\mathbb{P}$ is a linear subspace. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by

$$f(\vec{x}) \;=\; 2 \cdot \mathbf{pr}_{\mathbb{P}}(\vec{x}) - \vec{x}.$$

This is equivalent to **reflecting** $\vec{x}$ across the plane $\mathbb{P}$.

Example 11: *Reflection across a subspace*
Let $\mathbb{V}$ be a linear subspace in $\mathbb{R}^D$ passing through zero. Define $f : \mathbb{R}^D \longrightarrow \mathbb{R}^D$ by

$$f(\vec{x}) \;=\; 2 \cdot \mathbf{pr}_{\mathbb{V}}(\vec{x}) - \vec{x}.$$

This is "**reflecting**" $\vec{x}$ across $\mathbb{V}$.

Example 12: *Linear Actions on Matrices*
$\mathcal{M}_{N \times M}$ is also a vector space, and the following are linear transformations:

- **Matrix Multiplication:** If $\boxed{\text{A}} \in \mathcal{M}_{N \times M}$, define $f : \mathcal{M}_{M \times D} \longrightarrow \mathcal{M}_{N \times D}$ by $f\left(\boxed{\text{B}}\right) = \boxed{\text{A}} \cdot \boxed{\text{B}}$. Then f is a linear transformation.
- **Matrix Transposition:** Define $f : \mathcal{M}_{M \times D} \longrightarrow \mathcal{M}_{D \times M}$ by $f\left(\boxed{\text{B}}\right) = \boxed{\text{B}}^t$. Then f is linear.
- **Matrix Trace:** Define $f : \mathcal{M}_{N \times N} \longrightarrow \mathbb{R}$ by $f\left(\boxed{\text{B}}\right) = \mathsf{trace}\left[\boxed{\text{B}}\right] = b_{11} + b_{22} + \ldots + b_{NN}$. Then f is linear.

Nonexamples 13: The following are *not* linear transformations

- **Matrix Inversion:** Define $f : \mathcal{M}_{N \times N} \longrightarrow \mathcal{M}_{N \times N}$ by $f\left(\boxed{\text{B}}\right) = \boxed{\text{B}}^{-1}$.
  - f is not defined on all of $\mathcal{M}_{N \times N}$.
  - f is not linear, even where it is defined.
- **Polynomial Squaring:** Define $f : \mathcal{P}_N \longrightarrow \mathcal{P}_{2N}$ by $f(p(x)) = p^2(x)$. For example, $f(x+1) = x^2 + 2x + 1$. It is easy to find counterexamples to show this is not linear. (Try!)
- **Translation:** Let $\mathbf{v} \in \mathbb{R}^D$, and define $f : \mathbb{R}^D \longrightarrow \mathbb{R}^D$ by $f(\vec{x}) = \vec{x} + \mathbf{v}$. Then f is not linear because the **origin** does not remain fixed.
- **Rotation about a point other than zero**: If $\mathbf{x} \in \mathbb{R}^2$ is a point other than zero, and we rotate $\mathbb{R}^2$ about $\mathbf{x}$, then this function *moves the origin*, so it is not linear.
- **Reflection across a line not through zero:** If $\mathbb{L}$ is a line in $\mathbb{R}^2$ that does not pass through zero, then reflection across $\mathbb{L}$ *moves the origin*, so it is not linear.

Remark 14: *Affine Transformations*

The last three examples "would be" linear, except that they move zero. A map of this kind is called an **affine transformation**; it is a linear transformation, preceeded and/or followed by a *translation*.

Example 15: *Change of Basis*

Suppose $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ is some **basis** for $\mathbb{R}^N$, and $\vec{x} \in \mathbb{R}^N$. How can we find the **coordinates** for $\vec{x}$ in the basis $\mathcal{B}$?

Let $\boxed{\mathrm{B}} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$. $\mathcal{B}$ is a basis, so $\boxed{\mathrm{B}}$ is **invertible**.

Let $\boxed{\mathrm{A}} = \boxed{\mathrm{B}}^{-1}$.

Suppose $\vec{x} = (x_1, \ldots, x_N)$. Let $\vec{y} = \boxed{\mathrm{A}} \cdot \vec{x}$; with $\vec{y} = (y_1, \ldots, y_N)$. I claim that $(y_1, \ldots, y_N)$ are the coordinates of $\vec{x}$ with respect to $\mathcal{B}$. To see this:

$$\begin{aligned} \sum_{n=1}^{N} y_n \mathbf{b}_n &= y_1 \begin{bmatrix} \uparrow \\ \mathbf{b}_1 \\ \downarrow \end{bmatrix} + y_2 \begin{bmatrix} \uparrow \\ \mathbf{b}_2 \\ \downarrow \end{bmatrix} + \ldots + y_N \begin{bmatrix} \uparrow \\ \mathbf{b}_N \\ \downarrow \end{bmatrix} \\ &= \boxed{\mathrm{B}} \cdot \vec{y} \\ &= \boxed{\mathrm{B}} \cdot \boxed{\mathrm{B}}^{-1} \cdot \vec{x} \\ &= \vec{x} \end{aligned}$$

Example 16:

$$\begin{aligned} \text{Suppose } \mathbf{b}_1 &= \left(\frac{\sqrt{3}}{2}, \frac{1}{2} \right), \\ \mathbf{b}_1 &= \left(\frac{-1}{2}, \frac{\sqrt{3}}{2} \right), \\ \text{Then } \boxed{\mathrm{B}} &= \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}. \\ \text{Thus, } \boxed{\mathrm{A}} &= \boxed{\mathrm{B}}^{-1} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}. \end{aligned}$$

Thus, if $\vec{x} = (1, 2)$, then

$$\vec{y} \;=\; \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \;=\; \begin{bmatrix} \frac{\sqrt{3}}{2} + 1 \\ \frac{-1}{2} + \sqrt{3} \end{bmatrix}$$