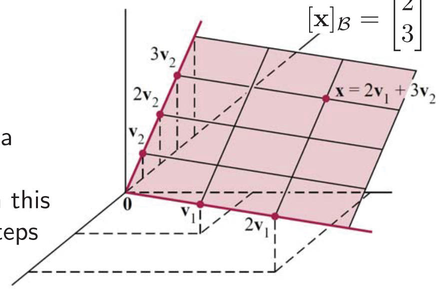
§4.4, 4.7, 5.4: Change of Basis

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

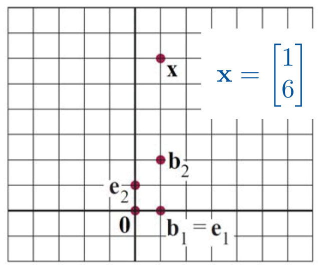
- The $\mathcal{B}$-coordinate vector of $\mathbf{x}$ is $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ where $\mathbf{x} = c_1\mathbf{b}_1 + \dots + c_n\mathbf{b}_n$.
- The matrix for a linear transformation $T: V \to V$ relative to $\mathcal{B}$ is

$$[T]_\mathcal{B} = \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_\mathcal{B} & \dots & [T(\mathbf{b}_n)]_\mathcal{B} \\ | & & | \end{bmatrix}.$$

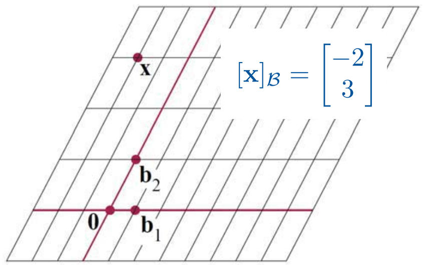
A basis for this plane in $\mathbb{R}^3$ allows us to draw a coordinate grid on the plane. The coordinate vectors then describe the location of points on this plane relative to this coordinate grid (e.g. 2 steps in $\mathbf{v}_1$ direction, 3 steps in $\mathbf{v}_2$ direction.)



Although we already have the standard coordinate grid on $\mathbb{R}^n$, some computations are much faster and more accurate in a different basis i.e. using a different coordinate grid (later, p17-19).



standard coordinate grid



$\mathcal{B}$-coordinate grid

Important questions:

i how are $\mathbf{x}$ and $[\mathbf{x}]_\mathcal{B}$ related (p3-6, §4.4 in textbook);
ii how are $[\mathbf{x}]_\mathcal{B}$ and $[\mathbf{x}]_\mathcal{F}$ related for two bases $\mathcal{B}$ and $\mathcal{F}$ (p7-10, §4.7);
iii how are the standard matrix of T and the matrix $[T]_\mathcal{B}$ related (p11-14, §5.4).

Changing from any basis to the standard basis of $\mathbb{R}^n$

EXAMPLE: (see the picture on p3) Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be a basis of $\mathbb{R}^2$.

a. If $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, then what is $\mathbf{x}$?

b. If $[\mathbf{v}]_\mathcal{B} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then what is $\mathbf{v}$?

Solution: (a) Use the definition of coordinates:

$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ means that $\mathbf{x} =$ ______ $\mathbf{b}_1 +$ ______ $\mathbf{b}_2 =$

(b) Use the definition of coordinates:

$[\mathbf{v}]_\mathcal{B} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ means that $\mathbf{v} =$ ______ $\mathbf{b}_1 +$ ______ $\mathbf{b}_2 =$

In general, if $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for $\mathbb{R}^n$, and $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, then

$\mathbf{x} =$ ______ $\mathbf{b}_1 +$ ______ $\mathbf{b}_2 + \dots +$ ______ $\mathbf{b}_n = \begin{bmatrix} \quad & \quad & \quad \end{bmatrix} [\mathbf{x}]_\mathcal{B}$.

This is the **change-of-coordinates matrix from $\mathcal{B}$ to the standard basis** ($\mathcal{P}_\mathcal{B}$ in textbook).

In the opposite direction
Changing from the standard basis to any other basis of $\mathbb{R}^n$

EXAMPLE: (see the picture on p3) Let $\mathbf{b}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix}1\\2\end{bmatrix}$ and let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be a basis of $\mathbb{R}^2$.

a. If $\mathbf{x} = \begin{bmatrix}1\\6\end{bmatrix}$, then what are its $\mathcal{B}$-coordinates $[\mathbf{x}]_\mathcal{B}$?

b. If $\mathbf{v} = \begin{bmatrix}v_1\\v_2\end{bmatrix}$, then what are its $\mathcal{B}$-coordinates $[\mathbf{v}]_\mathcal{B}$?

Solution: (a) Suppose $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}c_1\\c_2\end{bmatrix}$. This means that

$$\begin{bmatrix}1\\6\end{bmatrix} = \mathbf{x} =$$

So (c_1, c_2) is the solution to the linear system $\left[\begin{array}{cc|c}1 & 1 & 1\\0 & 2 & 6\end{array}\right]$.

Row reduction: $\left[\begin{array}{cc|c}1 & 0 & -2\\0 & 1 & 3\end{array}\right]$

So $[\mathbf{x}]_\mathcal{B} =$

(b) The $\mathcal{B}$-coordinate vector $\begin{bmatrix}c_1\\c_2\end{bmatrix}$ of $\mathbf{v}$ satisfies $\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = \begin{bmatrix} & \\ & \end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}$.

So $[\mathbf{v}]_\mathcal{B}$ is the solution to

In general, if $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for $\mathbb{R}^n$, and $\mathbf{v}$ is any vector in $\mathbb{R}^n$, then

$[\mathbf{v}]_\mathcal{B}$ is a solution to $\underbrace{\begin{bmatrix}| & & |\\ \mathbf{b}_1 & \dots & \mathbf{b}_n\\ | & & |\end{bmatrix}}_{\mathcal{P}_\mathcal{B}} \mathbf{x} = \mathbf{v}$.

Because $\mathcal{B}$ is a basis, the columns of $\mathcal{P}_\mathcal{B}$ are linearly independent, so by the Invertible Matrix Theorem, $\mathcal{P}_\mathcal{B}$ is invertible, and the unique solution to $\mathcal{P}_\mathcal{B}\mathbf{x} = \mathbf{v}$ is

$$[\mathbf{v}]_\mathcal{B} = \begin{bmatrix}| & & |\\ \mathbf{b}_1 & \dots & \mathbf{b}_n\\ | & & |\end{bmatrix}^{-1} \mathbf{v}.$$

In other words, the change-of-coordinates matrix from the standard basis to $\mathcal{B}$ is $\mathcal{P}_\mathcal{B}^{-1}$.

Indeed, in the previous example, $\mathcal{P}_\mathcal{B}^{-1}\mathbf{x} = \begin{bmatrix}1 & 1\\0 & 2\end{bmatrix}^{-1}\begin{bmatrix}1\\6\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2 & -1\\-0 & 1\end{bmatrix}\begin{bmatrix}1\\6\end{bmatrix} = \begin{bmatrix}-2\\3\end{bmatrix}$.

A very common mistake is to get the direction wrong:

Does multiplication by $\mathcal{P}_\mathcal{B}$ change from standard coordinates to $\mathcal{B}$-coordinates, or from $\mathcal{B}$-coordinates to standard coordinates?

Don't memorise the formulas. Instead, remember the definition of coordinates:

$$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix} \text{ means } \quad \mathbf{x} = c_1\mathbf{b}_1 + \dots + c_n\mathbf{b}_n = \begin{bmatrix}| & & |\\ \mathbf{b}_1 & \dots & \mathbf{b}_n\\ | & & |\end{bmatrix}[\mathbf{x}]_\mathcal{B}$$

and you won't go wrong.

ii: Changing between two non-standard bases:

Example: As before, $\mathbf{b}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix}1\\2\end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Another basis: $\mathbf{f}_1 = \begin{bmatrix}1\\1\end{bmatrix}, \mathbf{f}_2 = \begin{bmatrix}0\\1\end{bmatrix}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$.

If $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}-2\\3\end{bmatrix}$, then what are its $\mathcal{F}$-coordinates $[\mathbf{x}]_\mathcal{F}$?

Answer 1: $\mathcal{B}$ to standard to $\mathcal{F}$ - works only in $\mathbb{R}^n$, in general easiest to calculate.

$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}-2\\3\end{bmatrix}$ means $\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2 = -2\begin{bmatrix}1\\0\end{bmatrix} + 3\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\6\end{bmatrix}$.

So if $[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}d_1\\d_2\end{bmatrix}$, then $d_1\begin{bmatrix}1\\1\end{bmatrix} + d_2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\6\end{bmatrix}$.

Row-reducing $\left[\begin{array}{cc|c}1 & 0 & 1\\1 & 1 & 6\end{array}\right]$ shows $d_1 = 1, d_2 = 5$ so $[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}1\\5\end{bmatrix}$.

In other words, $\mathbf{x} = \mathcal{P}_\mathcal{B}[\mathbf{x}]_\mathcal{B}$ and $[\mathbf{x}]_\mathcal{F} = \mathcal{P}_\mathcal{F}^{-1}\mathbf{x}$, so $[\mathbf{x}]_\mathcal{F} = \mathcal{P}_\mathcal{F}^{-1}\mathcal{P}_\mathcal{B}[\mathbf{x}]_\mathcal{B}$.

Answer 2: A different view that works for abstract vector spaces (without reference to a standard basis) - important theoretically, but may be hard to calculate for general examples in $\mathbb{R}^n$.

$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}-2\\3\end{bmatrix}$ means $\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2$.

So $[\mathbf{x}]_\mathcal{F} = [-2\mathbf{b}_1 + 3\mathbf{b}_2]_\mathcal{F} = -2[\mathbf{b}_1]_\mathcal{F} + 3[\mathbf{b}_2]_\mathcal{F} = \begin{bmatrix}| & |\\ [\mathbf{b}_1]_\mathcal{F} & [\mathbf{b}_2]_\mathcal{F}\\ | & |\end{bmatrix}\begin{bmatrix}-2\\3\end{bmatrix}$.

because $\mathbf{x} \mapsto [\mathbf{x}]_\mathcal{F}$ is an isomorphism, so every vector space calculation is accurately reproduced using coordinates.

$\mathbf{b}_1 = \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix} = \mathbf{f}_1 - \mathbf{f}_2$ so $[\mathbf{b}_1]_\mathcal{F} = \begin{bmatrix}1\\-1\end{bmatrix}$.

$\mathbf{b}_2 = \begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} = \mathbf{f}_1 + \mathbf{f}_2$ so $[\mathbf{b}_2]_\mathcal{F} = \begin{bmatrix}1\\1\end{bmatrix}$.

This step can be hard to calculate if the $\mathbf{b}_i$ are not "easy" linear combinations of the $\mathbf{f}_i$. But if you need to change bases in a practical application, the bases are probably "nicely" related.

So $[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}\begin{bmatrix}-2\\3\end{bmatrix} = \begin{bmatrix}1\\5\end{bmatrix}$.

Theorem 15: Change of Basis: Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be two bases of a vector space V. Then, for all $\mathbf{x}$ in V,

$$[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}| & & |\\ [\mathbf{b}_1]_\mathcal{F} & \dots & [\mathbf{b}_n]_\mathcal{F}\\ | & & |\end{bmatrix}[\mathbf{x}]_\mathcal{B}.$$

Notation: write $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}}$ for the matrix $\begin{bmatrix}| & & |\\ [\mathbf{b}_1]_\mathcal{F} & \dots & [\mathbf{b}_n]_\mathcal{F}\\ | & & |\end{bmatrix}$, the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{F}$.

A tip to get the direction correct:

$$[\mathbf{x}]_\mathcal{F} = \underbrace{\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}}[\mathbf{x}]_\mathcal{B}}$$

A $\mathcal{F}$-coordinate vector

a linear combination of columns of $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}}$, so these columns should be $\mathcal{F}$-coordinate vectors

Theorem 15: Change of Basis: Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be two bases of a vector space V. Then, for all $\mathbf{x}$ in V,

$$[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}| & & |\\ [\mathbf{b}_1]_\mathcal{F} & \dots & [\mathbf{b}_n]_\mathcal{F}\\ | & & |\end{bmatrix}[\mathbf{x}]_\mathcal{B}.$$

Properties of the change-of-coordinates matrix $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}} = \begin{bmatrix}| & & |\\ [\mathbf{b}_1]_\mathcal{F} & \dots & [\mathbf{b}_n]_\mathcal{F}\\ | & & |\end{bmatrix}$:

- $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}} = \underset{\mathcal{B}\leftarrow\mathcal{F}}{\mathcal{P}}^{-1}$.
- If V is $\mathbb{R}^n$ and $\mathcal{E}$ is the standard basis $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$, then $\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_\mathcal{B} = \begin{bmatrix}| & & |\\ \mathbf{b}_1 & \dots & \mathbf{b}_n\\ | & & |\end{bmatrix}$, because $[\mathbf{b}_i]_\mathcal{E} = \mathbf{b}_i$. Also $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \mathcal{P}_\mathcal{B}^{-1}$.
- If V is $\mathbb{R}^n$, then $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_\mathcal{F}^{-1}\mathcal{P}_\mathcal{B}$ (see p8).

iii: Change of coordinates and linear transformations:

Remember that the matrix for a linear transformation $T : V \to V$ relative to $\mathcal{B}$ is

$$[T]_\mathcal{B} = \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_\mathcal{B} & \dots & [T(\mathbf{b}_n)]_\mathcal{B} \\ | & & | \end{bmatrix}$$, and this matrix is useful because

$$[T(\mathbf{x})]_\mathcal{B} = [T]_\mathcal{B}[\mathbf{x}]_\mathcal{B} \qquad (*)$$

(i.e. if you're working in $\mathcal{B}$-coordinates, then applying the function T is the same as multiplying by the matrix $[T]_\mathcal{B}$).

Often, it is easier to find the matrix for T relative to one basis than to another (later, p14). So it's important to know how to find $[T(\mathbf{x})]_\mathcal{F}$ if we know $[T(\mathbf{x})]_\mathcal{B}$. In $\mathbb{R}^n$, the following is true:

$$T(\mathbf{x}) = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T(\mathbf{x})]_\mathcal{B} \overset{(*)}{=} \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}[\mathbf{x}]_\mathcal{B} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}.$$

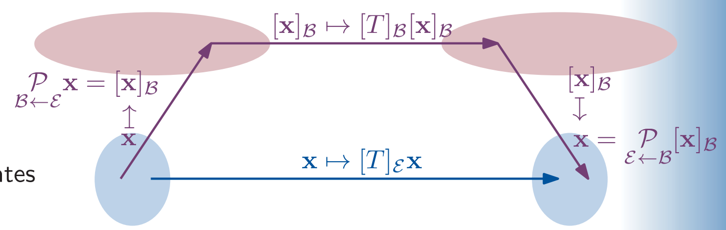
Let $[T]_\mathcal{E}$ be the standard matrix of T. Then the equation above shows that $[T]_\mathcal{E}\mathbf{x} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}$ for all $\mathbf{x}$. So

$$\boxed{[T]_\mathcal{E} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}.}$$

A picture to illustrate $[T]_\mathcal{E} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}$:

$\mathcal{B}$-coordinates

Standard coordinates



Because $\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_\mathcal{B}$ and $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \mathcal{P}_\mathcal{B}^{-1}$:

$$[T]_\mathcal{E} = \mathcal{P}_\mathcal{B}[T]_\mathcal{B}\mathcal{P}_\mathcal{B}^{-1}.$$

Multiply both sides by $\mathcal{P}_\mathcal{B}^{-1}$ on the left and by $\mathcal{P}_\mathcal{B}$ on the right:

$$\mathcal{P}_\mathcal{B}^{-1}[T]_\mathcal{E}\mathcal{P}_\mathcal{B} = [T]_\mathcal{B}$$

These two equations are hard to remember ("where does the inverse go?"). Instead, remember $\boxed{[T]_\mathcal{E} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_\mathcal{B}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}}$ (which works for all vector spaces, not just $\mathbb{R}^n$).

EXAMPLE: Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be a basis of $\mathbb{R}^2$.

Suppose T is a linear transformation satisfying $T(\mathbf{b}_1) = \mathbf{b}_1$ and $T(\mathbf{b}_2) = -\mathbf{b}_2$. Find $[T]_\mathcal{E}$, the standard matrix of T.

Solution: From the given information, it is easy to find $[T]_\mathcal{B}$, the matrix for T relative to $\mathcal{B}$:

$$[T]_\mathcal{B} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} =$$

Now use change of coordinates:

$$[T]_\mathcal{E} = \qquad\qquad = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

To find the change-of-coordinate matrices, use the definition of coordinates:

$$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \text{ means } \mathbf{x} = \quad \mathbf{b}_1 + \quad \mathbf{b}_2 = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}[\mathbf{x}]_\mathcal{B}.$$

So $\quad [T]_\mathcal{E} =$

$$= \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

$$= \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

Check that our answer satisfies the conditions given in the question:

$[T]_\mathcal{E}\mathbf{b}_1 =$

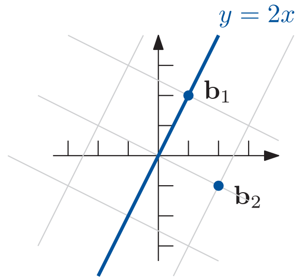
$[T]_\mathcal{E}\mathbf{b}_2 =$

Information of the type $T(\mathbf{b}_1) = \mathbf{b}_1$, $T(\mathbf{b}_2) = -\mathbf{b}_2$ arise naturally when considering geometric linear transformations. Indeed, the linear transformation on the previous page is reflection through the line $y = 2x$, because:

$\begin{bmatrix}1\\2\end{bmatrix}$ is on the line $y = 2x$, so it is unchanged by the reflection: $T\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}1\\2\end{bmatrix}$.

$\begin{bmatrix}2\\-1\end{bmatrix}$ is perpendicular to $y = 2x$, so its image is its negative: $T\left(\begin{bmatrix}2\\-1\end{bmatrix}\right) = \begin{bmatrix}-2\\1\end{bmatrix}$.

In this case, $[T]_{\mathcal{B}} = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ is more useful than $[T]_{\mathcal{E}} = \begin{bmatrix}-3/5 & 4/5\\4/5 & 3/5\end{bmatrix}$ - for example, it is clear from the diagonal matrix $[T]_{\mathcal{B}}$ that T^2 is the identity transformation.



Remember

$$[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}.$$

This motivates the following definition:

Definition: Two square matrices A and D are *similar* if there is an invertible matrix P such that $A = PDP^{-1}$.

Similar matrices represent the same linear transformation in different bases.

Similar matrices have the same determinant and the same rank, because the signed volume scaling factor and the dimension of the image are coordinate-independent properties of the linear transformation. (Exercise: prove that $\det D = \det(PDP^{-1})$ using the multiplicative property of determinants.)

Why is change of basis important?

Example: If x, y are the prices of two stocks on a particular day, then their prices the next day are respectively $\frac{1}{2}y$ and $-x+\frac{3}{2}y$. How are the prices after many days related to the prices today?

Answer: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the function representing the changes in stock prices from one day to the next, i.e. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}y \\ -x+\frac{3}{2}y \end{bmatrix}$. We are interested in T^k for large k. (You will NOT be required to do this step.)

T is a linear transformation; its standard matrix is $[T]_{\mathcal{E}} = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}$. Calculating

$\begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}^k$ by direct matrix multiplication will take a long time.

Answer: (continued) Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

$$T(\mathbf{b}_1) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \tfrac{1}{2}\mathbf{b}_1, \quad T(\mathbf{b}_2) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{b}_2,$$

so $[T]_{\mathcal{B}} = \begin{bmatrix} | & | \\ [T(\mathbf{b}_1)]_{\mathcal{B}} & [T(\mathbf{b}_2)]_{\mathcal{B}} \\ | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$. Use $[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}$:

$$\begin{aligned}
[T]_{\mathcal{E}}^k &= \left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}\right)^k \\
&= \left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}}\right)\left(\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}}[T]_{\mathcal{B}}\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}}\right)\cdots\cdots\left(\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}}[T]_{\mathcal{B}}\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}\right) \\
&= \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}^k\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}^k\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{2}^k & 0 \\ 0 & 1^k \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1+\frac{1}{2}^{k-1} & 1-\frac{1}{2}^k \\ -2+\frac{1}{2}^{k-1} & 2-\frac{1}{2}^k \end{bmatrix}.
\end{aligned}$$

So $[T]_{\mathcal{E}}^k = \begin{bmatrix} -1+\frac{1}{2}^{k-1} & 1-\frac{1}{2}^k \\ -2+\frac{1}{2}^{k-1} & 2-\frac{1}{2}^k \end{bmatrix}$. When k is very large, this is very close to $\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$.

So essentially the stock prices after many days is $-x+y$ and $-2x+2y$, where x, y are the prices today. (In particular, the prices stabilise, which was not clear from $[T]_{\mathcal{E}}$.)

The important points in this example:

- We have a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ and we want to find T^k for large k.
- We find a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ where $T(\mathbf{b}_1) = \lambda_1\mathbf{b}_1$ and $T(\mathbf{b}_2) = \lambda_2\mathbf{b}_2$ for some scalars λ_1, λ_2. (In the example, $\lambda_1 = \frac{1}{2}, \lambda_2 = 1$.)
- Relative to the basis $\mathcal{B}$, the matrix for T is a diagonal matrix $[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$.
- It is easy to compute with $[T]_{\mathcal{B}}$, and we can then use change of coordinates to transfer the result to the standard matrix $[T]_{\mathcal{E}}$.

Next week (§5): does a "magic" basis like this always exist, and how to find it?
(Don't worry: you can do many of the computations in §5 without fully understanding change of coordinates.)