More examples of change of basis

We will do a few more examples of change of basis and say a bit more on dividing 2×2 matrices into equivalences classes by similarity. Here is a table of content for the examples. Important phrases are highlighted in red.

1. Geometric point of view.
2. Change of basis on a subspace (non-examinable).
3. Diagonalization and Fibonacci numbers.
4. Equivalence classes of 2×2 matrices + many more examples.

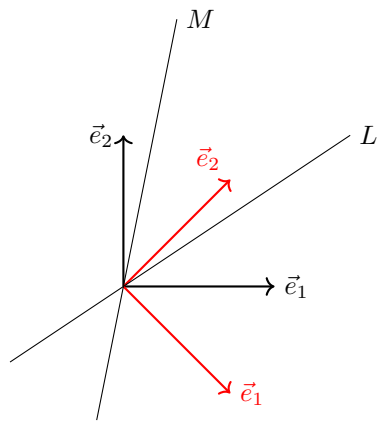
Example 1 (Quiz 1, #3)**.** Let $\mathfrak{B} = \{\vec{v}_1, \vec{v}_2\}$, where

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(these are in the standard basis). The change of basis matrix is

$$S = [\vec{v}_1, \vec{v}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

The two sets of bases are related by a rotation by $\frac{\pi}{4}$.



Let L be the line $2x = 3y$, M be the line L rotated counter-clockwise by $\frac{\pi}{4}$. The equation of M is $y = 5x$. Let P_L and P_M be the two projection transformations. The matrix of P_L (in the standard basis) is

$$[P_L] = [P_L]_{\text{std}} = \frac{1}{13} \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}$$

So the matrix of P_L in the basis $\mathfrak{B}$ is

$$[T]_{\mathfrak{B}} = S^{-1}[T]S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{13} \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{26} \begin{bmatrix} 1 & 5 \\ 5 & 25 \end{bmatrix}$$

which is the projection onto the line with equation $y = 5x$. This is not a coincidence: $y = 5x$ is the equation of L in the basis $\mathfrak{B}$.

Geometrically, this is what is happening: to project onto M, we can first rotate everything clockwise by $\frac{\pi}{4}$ so M becomes L, then project on to L, and then undo the first operation by rotating counter-clockwise by $\frac{\pi}{4}$. In matrix languages,

$$[P_M] = [\text{Rot}_{\frac{\pi}{4}}] \cdot [P_L] \cdot [\text{Rot}_{-\frac{\pi}{4}}] = [\text{Rot}_{-\frac{\pi}{4}}]^{-1} \cdot [P_L] \cdot [\text{Rot}_{-\frac{\pi}{4}}]$$

This is the view that change of basis is doing the same thing in a different frame of reference. It also follows that $[R_M] = S^{-1}[R_L]S$ if R denotes a reflection. This answers part (d).

Example 2. This example does change of basis in a subspace without reference to the standard basis. This is beyond the course, but it includes some more example computations with bases, which are useful to see and are in the course.

Let $V \subseteq \mathbb{R}^3$ be the plane $x + y + z = 0$. It has two bases $\mathfrak{B} = \{\vec{v}_1, \vec{v}_2\}$, $\mathfrak{C} = \{\vec{w}_1, \vec{w}_2\}$, where

$$\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \vec{w}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix},$$

If we interpret V as the kernel of $A = [1\ 1\ 1]$, then the first basis is what the algorithm in class would write down, and the second is choosing a different way of writing down all solutions.

Let T be the linear transformation which changes the y and z coordinates, leaving the x-coordinate fixed (this is the reflection across the plane $y = z$ in $\mathbb{R}^3$). In the basis $\mathfrak{B}$, the matrix of T is

$$[T]_{\mathfrak{B}} = \begin{bmatrix} [T(\vec{v}_1)]_{\mathfrak{B}} & [T(\vec{v}_2)]_{\mathfrak{B}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}_{\mathfrak{B}} & \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}_{\mathfrak{B}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

In other words, $T(\vec{v}_1) = \vec{v}_2$, and $T(\vec{v}_2) = \vec{v}_1$. Similarly, observe that $T(\vec{w}_1) = \vec{w}_1 + \vec{w}_2$ and $T(\vec{w}_2) = -\vec{w}_2$, so

$$[T]_{\mathfrak{C}} = \begin{bmatrix} [T(\vec{w}_1)]_{\mathfrak{C}} & [T(\vec{w}_2)]_{\mathfrak{C}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$$

Finally, we introduce the base change matrix from $\mathfrak{B}$ to $\mathfrak{C}$

$$S_{\mathfrak{B},\mathfrak{C}} = \begin{bmatrix} [\vec{w}_1]_{\mathfrak{B}} & [\vec{w}_2]_{\mathfrak{B}} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$$

so we can check that

$$S_{\mathfrak{B},\mathfrak{C}}^{-1}[T]_{\mathfrak{B}} S_{\mathfrak{B},\mathfrak{C}} = \begin{bmatrix} -1 & -1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = [T]_{\mathfrak{C}}$$

The whole process can be explained by the commutative diagram

$$\begin{tikzcd}
 & V \arrow[dr, "T"] & & \\
\mathbb{R}^2 \arrow[ur, "\mathfrak{B}"] \arrow[dr, "{[T]_{\mathfrak{B}}}"'] & & \mathbb{R}^2 \arrow[ul, "\mathfrak{C}"'] \arrow[ll, "S_{\mathfrak{B},\mathfrak{C}}"'] \arrow[dr, "{[T]_{\mathfrak{C}}}"] & V \\
 & \mathbb{R}^2 \arrow[ur, "\mathfrak{B}"] & & \mathbb{R}^2 \arrow[ll, "S_{\mathfrak{B},\mathfrak{C}}"'] \arrow[ul, "\mathfrak{C}"']
\end{tikzcd}$$

Example 3. Consider the matrix

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

Say for some reason we want to compute A^{1000}. The practical way is to use exponentiation by squaring, but you are not a computer, so you need a closed formula. We will find one for A^n. To do this, "observe" that

$$A = SDS^{-1}, \quad D = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}, \ S = \begin{bmatrix} \frac{\sqrt{5}-1}{2} & \frac{-\sqrt{5}-1}{2} \\ 1 & 1 \end{bmatrix}$$

We will see how to actually find them in Chapter 7, but in the mean time, we have

$$A^2 = (SDS^{-1})(SDS^{-1}) = SD(S^{-1}S)DS^{-1} = SD^2S^{-1}$$

This pattern continues, so $A^n = SD^nS^{-1}$ (we are *not* switching the order of matrices, just re-grouping them). From the more abstract point of view, A and D represent the same linear transformation in different bases, so A^n should also be similar to D^n using the same change of basis.

But D is diagonal, so

$$D^n = \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}$$

because D just scales each basis vectors independently. Therefore, we can write down an explicit formula for $A^n = SD^nS^{-1}$, and in particular compute A^{1000}, which have really large entries, because $\frac{1+\sqrt{5}}{2} > 1$.

This is related to the Fibonacci numbers because

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_n + F_{n+1} \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_{n+2} \end{bmatrix}$$

so if we let $F_0 = 0$, $F_1 = 1$, then

$$A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}$$

So the above formula for A^n gives the famous closed form formula for the Fibonacci numbers

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right)$$

Example 4. In this example, we group all 2×2 matrices into *equivalence classes*, so all matrices in the same class are similar to each other, and matrices from different classes are not similar. This is possible because similarity is an equivalence relation.

There are three types of classes. In the next list, we write down one element from each class, and the class contains all 2×2 matrices similar to that element.

1. Diagonal: given real numbers $a \leq b$, we have a class $\mathcal{H}_{a,b}$ containing the diagonal matrix $\left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$.
2. Scaled shear: given a real number λ, we have a class $\mathcal{U}_\lambda$ containing the matrix $\left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right]$.
3. Scaled rotation: given an angle $\theta \in (0, \pi)$ and a scale $r > 0$, we have a class $\mathcal{E}_{r,\theta}$ containing the scaled rotation matrix $r\left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right]$.

They are also called hyperbolic, unipotent/parabolic, and elliptic classes, respectively. Their diagonalizability properties are as follows:

1. Diagonal classes (hyperbolic): always diagonalizable, by definition.
2. Scaled shear (parabolic): never diagonalizable.
3. Scaled rotation (elliptic): not diagonalizable over $\mathbb{R}$, but diagonalizable over $\mathbb{C}$.

We will see in Chapter 7 that they are exactly the equivalence classes. A similar result holds in higher dimensions, and it is called the Jordan normal form. For now, we will just look at some examples.

- Rotation by θ clockwise is similar to rotation by θ counter-clockwise:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

  Geometrically, flipping the y-axis changes clockwise to counter-clockwise. Note that we could have used $\left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, so the S-matrix in the definition of similarity is not unique.

- Rotation by π has matrix $-I_2$, so it is covered by case (1).

– The matrix $\left[\begin{smallmatrix} 4 & -15 \\ 2 & 7 \end{smallmatrix}\right]$ is similar to $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$:

$$\begin{bmatrix} 4 & -15 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

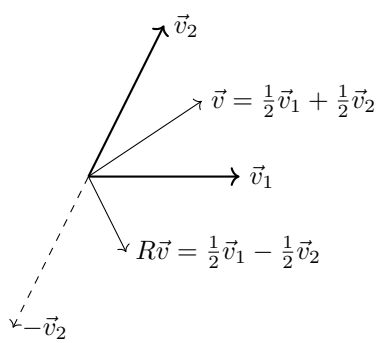
so it is diagonalizable and belongs to $\mathcal{H}_{1,2}$. Observe that the basis vectors do not need to be orthogonal.

– By changing the order of the basis, $\left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$ is similar to $\left[\begin{smallmatrix} b & 0 \\ 0 & a \end{smallmatrix}\right]$, so we have to eliminate over-counting by asking $a \leq b$ in case 1.

– The horizontal shear is similar to the vertical shear, again by changing the order of the axes. In matrix language, this is represented by

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

– Projections are in the class $\mathcal{H}_{0,1}$, and reflections are in the class $\mathcal{H}_{-1,1}$, but the converse is false. The matrix $\frac{1}{7}\left[\begin{smallmatrix} 1 & 3 \\ 2 & 6 \end{smallmatrix}\right]$ is similar to $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, but it does not represent an orthogonal projection. In this case, we need the new basis vectors to be orthogonal to each other for the result to represent an orthogonal projection, because $\mathbb{R}^n$ carries an additional structure of length and angle that is needed to define an orthogonal projection or reflection. Here is a picture of an oblique reflection.



– The matrices $\left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right]$ and $\lambda I_2 = \left[\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right]$ are not similar, since for any invertible $n \times n$ matrix S, we have

$$S^{-1}(\lambda I_2)S = \lambda S^{-1}S = \lambda I_2$$

In particular, the class $\mathcal{H}_{\lambda,\lambda}$ contains only one matrix.

All matrices in the same equivalence class have the same trace and determinant. The pair $(\operatorname{Tr} A, \det A)$ almost completely determines the class of A. The only exception is that they cannot distinguish between $\mathcal{H}_{\lambda,\lambda}$ and $\mathcal{U}_\lambda$, which are different classes by the above example.