Five.II Similarity

Linear Algebra

Jim Hefferon

http://joshua.smcvt.edu/linearalgebra

Definition and Examples

We've defined two matrices H and $\hat{H}$ to be matrix equivalent if there are nonsingular P and Q such that $\hat{H} = PHQ$. We were motivated by this diagram showing H and $\hat{H}$ both representing a map h, but with respect to different pairs of bases, B, D and $\hat{B}, \hat{D}$.

$$\begin{tikzcd}
V_{wrt\ B} \arrow[r, "h", "H"'] \arrow[d, "id"'] & W_{wrt\ D} \arrow[d, "id"'] \\
V_{wrt\ \hat{B}} \arrow[r, "h", "\hat{H}"'] & W_{wrt\ \hat{D}}
\end{tikzcd}$$

We've defined two matrices H and $\hat{H}$ to be matrix equivalent if there are nonsingular P and Q such that $\hat{H} = PHQ$. We were motivated by this diagram showing H and $\hat{H}$ both representing a map h, but with respect to different pairs of bases, B, D and $\hat{B}, \hat{D}$.

$$\begin{array}{ccc} V_{wrt\ B} & \xrightarrow[H]{h} & W_{wrt\ D} \\ \mathrm{id}\big\downarrow & & \mathrm{id}\big\downarrow \\ V_{wrt\ \hat{B}} & \xrightarrow[\hat{H}]{h} & W_{wrt\ \hat{D}} \end{array}$$

We now consider the special case of transformations, where the codomain equals the domain, and we add the requirement that the codomain's basis equals the domain's basis. So, we are considering representations with respect to B, B and D, D.

$$\begin{array}{ccc} V_{wrt\ B} & \xrightarrow[T]{t} & V_{wrt\ B} \\ \mathrm{id}\big\downarrow & & \mathrm{id}\big\downarrow \\ V_{wrt\ D} & \xrightarrow[\hat{T}]{t} & V_{wrt\ D} \end{array}$$

In matrix terms, $\mathrm{Rep}_{D,D}(t) = \mathrm{Rep}_{B,D}(\mathrm{id})\ \mathrm{Rep}_{B,B}(t)\ \big(\mathrm{Rep}_{B,D}(\mathrm{id})\big)^{-1}$.

Similar matrices

1.2 *Definition* The matrices T and $\hat{T}$ are *similar* if there is a nonsingular P such that $\hat{T} = PTP^{-1}$.

Example Consider the derivative map $d/dx\colon \mathcal{P}_2 \to \mathcal{P}_2$. Fix the basis $B = \langle 1, x, x^2 \rangle$ and the basis $D = \langle 1, 1+x, 1+x+x^2 \rangle$. In this arrow diagram we will first get T, and then calculate $\hat{T}$ from it.

$$\begin{tikzcd}
V_{\mathit{wrt}\ B} \arrow[r, "t", "T"'] \arrow[d, "\mathrm{id}"'] & V_{\mathit{wrt}\ B} \arrow[d, "\mathrm{id}"'] \\
V_{\mathit{wrt}\ D} \arrow[r, "t", "\hat{T}"'] & V_{\mathit{wrt}\ D}
\end{tikzcd}$$

Similar matrices

1.2 *Definition* The matrices T and $\hat{T}$ are *similar* if there is a nonsingular P such that $\hat{T} = PTP^{-1}$.

Example Consider the derivative map $d/dx\colon \mathcal{P}_2 \to \mathcal{P}_2$. Fix the basis $B = \langle 1, x, x^2 \rangle$ and the basis $D = \langle 1, 1+x, 1+x+x^2 \rangle$. In this arrow diagram we will first get T, and then calculate $\hat{T}$ from it.

$$\begin{tikzcd}
V_{\mathit{wrt}\ B} \arrow[r, "t", "T"'] \arrow[d, "\mathrm{id}"'] & V_{\mathit{wrt}\ B} \arrow[d, "\mathrm{id}"'] \\
V_{\mathit{wrt}\ D} \arrow[r, "t", "\hat{T}"'] & V_{\mathit{wrt}\ D}
\end{tikzcd}$$

The action of d/dx on the elements of the basis B is $1 \mapsto 0$, $x \mapsto 1$, and $x^2 \mapsto 2x$.

$$\mathrm{Rep}_B(\frac{d}{dx}(1)) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_B(\frac{d}{dx}(x)) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_B(\frac{d}{dx}(x^2)) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

So we have this matrix representation of the map.

$$T = \mathrm{Rep}_{B,B}(\frac{d}{dx}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

So we have this matrix representation of the map.

$$T = \mathrm{Rep}_{B,B}(\frac{d}{dx}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The matrix changing bases from B to D is $\mathrm{Rep}_{B,D}(\mathrm{id})$. We find these by eye

$$\mathrm{Rep}_D(\mathrm{id}(1)) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(\mathrm{id}(x)) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(\mathrm{id}(x^2)) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

to get this.

$$P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, by following the arrow diagram we have $\hat{T} = PTP^{-1}$.

$$\hat{T} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

To check that, and to underline what the arrow diagram says

$$\begin{array}{ccc} V_{\mathit{wrt}\ B} & \xrightarrow[T]{t} & V_{\mathit{wrt}\ B} \\ {\scriptstyle \mathrm{id}}\big\downarrow & & {\scriptstyle \mathrm{id}}\big\downarrow \\ V_{\mathit{wrt}\ D} & \xrightarrow[\hat{T}]{t} & V_{\mathit{wrt}\ D} \end{array}$$

we calculate $\hat{T}$ directly. The effect of the map on the basis elements is $d/dx(1) = 0$, $d/dx(1+x) = 1$, and $d/dx(1+x+x^2) = 1+2x$. Representing of those with respect to D

$$\mathrm{Rep}_D(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(1+2x) = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

gives the same matrix $\hat{T} = \mathrm{Rep}_{D,D}(d/dx)$ as above.

The definition doesn't require that we consider the underlying maps. We can just multiply matrices.

Example Where

$$T = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 4 & 5 & 2 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

(note that P is nonsingular) we can compute this $\hat{T} = PTP^{-1}$.

$$\hat{T} = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 4/3 \\ 27/2 & 3/2 & 2 \end{pmatrix}$$

The definition doesn't require that we consider the underlying maps. We can just multiply matrices.

Example Where

$$T = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 4 & 5 & 2 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

(note that P is nonsingular) we can compute this $\hat{T} = PTP^{-1}$.

$$\hat{T} = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 4/3 \\ 27/2 & 3/2 & 2 \end{pmatrix}$$

1.4 *Example* The only matrix similar to the zero matrix is itself: $PZP^{-1} = PZ = Z$. The identity matrix has the same property: $PIP^{-1} = PP^{-1} = I$.

Similarity is an equivalence

Exercise 15 checks that similarity is an equivalence relation.

Similarity is an equivalence

Exercise 15 checks that similarity is an equivalence relation.

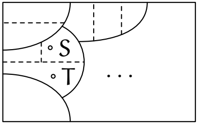
Since matrix similarity is a special case of matrix equivalence, if two matrices are similar then they are matrix equivalent. What about the converse: must any two matrix equivalent square matrices be similar? No; the matrix equivalence class of an identity consists of all nonsingular matrices of that size while the prior example shows that an identity matrix is alone in its similarity class.

Similarity is an equivalence

Exercise 15 checks that similarity is an equivalence relation.

Since matrix similarity is a special case of matrix equivalence, if two matrices are similar then they are matrix equivalent. What about the converse: must any two matrix equivalent square matrices be similar? No; the matrix equivalence class of an identity consists of all nonsingular matrices of that size while the prior example shows that an identity matrix is alone in its similarity class.

So some matrix equivalence classes split into two or more similarity classes — similarity gives a finer partition than does equivalence. This pictures some matrix equivalence classes subdivided into similarity classes.

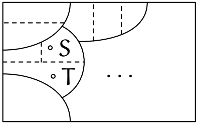


Similarity is an equivalence

Exercise 15 checks that similarity is an equivalence relation.

Since matrix similarity is a special case of matrix equivalence, if two matrices are similar then they are matrix equivalent. What about the converse: must any two matrix equivalent square matrices be similar? No; the matrix equivalence class of an identity consists of all nonsingular matrices of that size while the prior example shows that an identity matrix is alone in its similarity class.

So some matrix equivalence classes split into two or more similarity classes — similarity gives a finer partition than does equivalence. This pictures some matrix equivalence classes subdivided into similarity classes.



We naturally want a canonical form to represent the similarity classes. Some classes, but not all, are represented by a diagonal form.

Diagonalizability

2.1 *Definition* A transformation is *diagonalizable* if it has a diagonal representation with respect to the same basis for the codomain as for the domain. A *diagonalizable matrix* is one that is similar to a diagonal matrix: T is diagonalizable if there is a nonsingular P such that PTP^{-1} is diagonal.

Example This matrix

$$\begin{pmatrix} 6 & -1 & -1 \\ 2 & 11 & -1 \\ -6 & -5 & 7 \end{pmatrix}$$

is diagonalizable by using this

$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ -1/2 & 1/4 & 1/4 \\ -1/2 & -3/4 & 1/4 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{pmatrix}$$

to get this $D = PSP^{-1}$.

$$D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{pmatrix}$$

Example This matrix is not diagonalizable.

$$N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

The fact that N is not the zero matrix means that it cannot be similar to the zero matrix, because the zero matrix is similar only to itself. Thus if N were to be similar to a diagonal matrix D then D would have at least one nonzero entry on its diagonal.

The crucial point is that a power of N is the zero matrix, specifically N^2 is the zero matrix. This implies that for any map n represented by N with respect to some B, B, the composition $n \circ n$ is the zero map. This in turn implies that any matrix representing n with respect to some $\hat{B}, \hat{B}$ has a square that is the zero matrix. But for any nonzero diagonal matrix D^2, the entries of D^2 are the squares of the entries of D, so D^2 cannot be the zero matrix. Thus N is not diagonalizable.

2.4 *Lemma* A transformation t is diagonalizable if and only if there is a basis $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ and scalars $\lambda_1, \ldots, \lambda_n$ such that $t(\vec{\beta}_i) = \lambda_i \vec{\beta}_i$ for each i.

2.4 *Lemma* A transformation t is diagonalizable if and only if there is a basis $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ and scalars $\lambda_1, \ldots, \lambda_n$ such that $t(\vec{\beta}_i) = \lambda_i \vec{\beta}_i$ for each i.

Proof Consider a diagonal representation matrix.

$$\mathrm{Rep}_{B,B}(t) = \begin{pmatrix} \vdots & & \vdots \\ \mathrm{Rep}_B(t(\vec{\beta}_1)) & \cdots & \mathrm{Rep}_B(t(\vec{\beta}_n)) \\ \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix}$$

Consider the representation of a member of this basis with respect to the basis $\mathrm{Rep}_B(\vec{\beta}_i)$. The product of the diagonal matrix and the representation vector

$$\mathrm{Rep}_B(t(\vec{\beta}_i)) = \begin{pmatrix} \lambda_1 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{pmatrix}$$

has the stated action. QED

Example This matrix is not diagonal

$$T = \begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix}$$

but we can find a diagonal matrix similar to it, by finding an appropriate basis.

Example This matrix is not diagonal

$$T = \begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix}$$

but we can find a diagonal matrix similar to it, by finding an appropriate basis.

Suppose that $T = \mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(t)$ for $t\colon \mathbb{R}^2 \to \mathbb{R}^2$. We will find a basis $B = \langle \vec{\beta}_1, \vec{\beta}_2 \rangle$ giving a diagonal representation.

$$D = \mathrm{Rep}_{B,B}(t) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Here is the arrow diagram.

$$\begin{array}{ccc}
V_{\mathit{wrt}\ \mathcal{E}_2} & \xrightarrow[T]{t} & V_{\mathit{wrt}\ \mathcal{E}_2} \\
{\scriptstyle \mathrm{id}}\big\downarrow & & {\scriptstyle \mathrm{id}}\big\downarrow \\
V_{\mathit{wrt}\ B} & \xrightarrow[D]{t} & V_{\mathit{wrt}\ B}
\end{array}$$

We want λ_1 and λ_2 making these true.

$$\begin{pmatrix}4 & 1\\0 & -1\end{pmatrix}\vec{\beta}_1 = \lambda_1 \cdot \vec{\beta}_1 \qquad \begin{pmatrix}4 & 1\\0 & -1\end{pmatrix}\vec{\beta}_2 = \lambda_2 \cdot \vec{\beta}_2$$

More precisely, we want all scalars $x \in \mathbb{C}$ such that this system

$$\begin{pmatrix}4 & 1\\0 & -1\end{pmatrix}\begin{pmatrix}b_1\\b_2\end{pmatrix} = x \cdot \begin{pmatrix}b_1\\b_2\end{pmatrix}$$

has solutions $b_1, b_2 \in \mathbb{C}$ that are not both zero (the zero vector is not an element of any basis).

We want λ_1 and λ_2 making these true.

$$\begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix} \vec{\beta}_1 = \lambda_1 \cdot \vec{\beta}_1 \qquad \begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix} \vec{\beta}_2 = \lambda_2 \cdot \vec{\beta}_2$$

More precisely, we want all scalars $x \in \mathbb{C}$ such that this system

$$\begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

has solutions $b_1, b_2 \in \mathbb{C}$ that are not both zero (the zero vector is not an element of any basis).

Rewrite that as a linear system.

$$\begin{aligned} (4 - x) \cdot b_1 + \qquad\qquad\quad b_2 &= 0 \\ (-1 - x) \cdot b_2 &= 0 \end{aligned}$$

One solution is $\lambda_1 = -1$, associated with those (b_1, b_2) such that $b_1 = (-1/5)b_2$. The other solution is $\lambda_2 = 4$, associated with the (b_1, b_2) such that $b_2 = 0$.

Thus the original matrix

$$T = \begin{pmatrix} 4 & 1 \\ 0 & -1 \end{pmatrix}$$

is diagonalizable to

$$D = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$$

where this is a basis.

$$B = \langle \begin{pmatrix} -1 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$$

Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors

3.1 *Definition* A transformation $t \colon V \to V$ has a scalar *eigenvalue* λ if there is a nonzero *eigenvector* $\vec{\zeta} \in V$ such that $t(\vec{\zeta}) = \lambda \cdot \vec{\zeta}$.

Eigenvalues and eigenvectors

3.1 *Definition* A transformation $t\colon V \to V$ has a scalar *eigenvalue* λ if there is a nonzero *eigenvector* $\vec{\zeta} \in V$ such that $t(\vec{\zeta}) = \lambda \cdot \vec{\zeta}$.

3.5 *Definition* A square matrix T has a scalar *eigenvalue* λ associated with the nonzero *eigenvector* $\vec{\zeta}$ if $T\vec{\zeta} = \lambda \cdot \vec{\zeta}$.

Example The matrix

$$D = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

has an eigenvalue $\lambda_1 = 4$ and a second eigenvalue $\lambda_2 = 2$. The first is true because an associated eigenvector is $\vec{e_1}$

$$\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 4 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and similarly for the second an associated eigenvector is e_2.

$$\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Thinking of the matrix as representing a transformation of the plane, the transformation acts on those vectors in a particularly simple way, by rescaling.

Not every vector is simply rescaled.

$$\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \neq x \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Computing eigenvalues and eigenvectors

Example We will find the eigenvalues and associated eigenvectors of this matrix.

$$T = \begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix}$$

We want to find scalars x such that $T\vec{\zeta} = x\vec{\zeta}$ for some nonzero $\vec{\zeta}$. Bring the terms to the left side.

$$\begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} - x \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and factor out the vector.

$$\begin{pmatrix} 0-x & 5 & 7 \\ -2 & 7-x & 7 \\ -1 & 1 & 4-x \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (*)$$

This homogeneous system has nonzero solutions if and only if the matrix is singular, that is, has a determinant of zero.

Some computation gives the determinant and its factors.

$$\begin{aligned} 0 &= \begin{vmatrix} 0-x & 5 & 7 \\ -2 & 7-x & 7 \\ -1 & 1 & 4-x \end{vmatrix} \\ &= x^3 - 11x^2 + 38x - 40 = (x-5)(x-4)(x-2) \end{aligned}$$

So the eigenvalues are $\lambda_1 = 5$, $\lambda_2 = 4$, and $\lambda_3 = 2$.

Some computation gives the determinant and its factors.

$$0=\begin{vmatrix}0-x&5&7\\-2&7-x&7\\-1&1&4-x\end{vmatrix}$$

$$=x^3-11x^2+38x-40=(x-5)(x-4)(x-2)$$

So the eigenvalues are $\lambda_1=5$, $\lambda_2=4$, and $\lambda_3=2$.

To find the eigenvectors associated with the eigenvalue of 5 specialize equation $(*)$ for $x=5$.

$$\begin{pmatrix}-5&5&7\\-2&2&7\\-1&1&-1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Gauss's Method gives this solution set; its nonzero elements are the eigenvectors.

$$V_5=\{\begin{pmatrix}1\\1\\0\end{pmatrix}z_2 \mid z_2\in\mathbb{C}\}$$

Similarly, to find the eigenvectors associated with the eigenvalue of 4 specialize equation $(*)$ for $x = 4$.

$$\begin{pmatrix} -4 & 5 & 7 \\ -2 & 3 & 7 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gauss's Method gives this.

$$V_4 = \{ \begin{pmatrix} -7 \\ -7 \\ 1 \end{pmatrix} z_3 \mid z_3 \in \mathbb{C} \}$$

Similarly, to find the eigenvectors associated with the eigenvalue of 4 specialize equation $(*)$ for $x = 4$.

$$\begin{pmatrix} -4 & 5 & 7 \\ -2 & 3 & 7 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gauss's Method gives this.

$$V_4 = \{ \begin{pmatrix} -7 \\ -7 \\ 1 \end{pmatrix} z_3 \mid z_3 \in \mathbb{C} \}$$

Specializing $(*)$ for $x = 2$

$$\begin{pmatrix} -2 & 5 & 7 \\ -2 & 5 & 7 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gives this.

$$V_2 = \{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} z_3 \mid z_3 \in \mathbb{C} \}$$

Example To find the eigenvalues and associated eigenvectors for the matrix

$$T = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

start with this equation.

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} 3-x & 1 \\ 1 & 3-x \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (*)$$

Example To find the eigenvalues and associated eigenvectors for the matrix

$$T = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

start with this equation.

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} 3-x & 1 \\ 1 & 3-x \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (*)$$

That system has a nontrivial solution if this determinant is nonzero.

$$\begin{vmatrix} 3-x & 1 \\ 1 & 3-x \end{vmatrix} = x^2 - 6x + 8 = (x-2)(x-4)$$

Example To find the eigenvalues and associated eigenvectors for the matrix

$$T = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

start with this equation.

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} 3-x & 1 \\ 1 & 3-x \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (*)$$

That system has a nontrivial solution if this determinant is nonzero.

$$\begin{vmatrix} 3-x & 1 \\ 1 & 3-x \end{vmatrix} = x^2 - 6x + 8 = (x-2)(x-4)$$

First take the $x = 2$ version of $(*)$.

$$\begin{aligned} 1 \cdot b_1 + \quad b_2 &= 0 \\ b_1 + 1 \cdot b_2 &= 0 \end{aligned} \quad \Longrightarrow \quad V_2 = \{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \mid b_1 = -b_2 \text{ where } b_2 \in \mathbb{C} \}$$

Solving the second system is just as easy.

$$\begin{aligned} -1 \cdot b_1 + \quad b_2 &= 0 \\ b_1 - 1 \cdot b_2 &= 0 \end{aligned} \quad \Longrightarrow \quad V_4 = \{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \mid b_1 = b_2 \text{ where } b_2 \in \mathbb{C} \}$$

Example If the matrix is upper diagonal or lower diagonal

$$T = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

then the polynomial is easy to factor.

$$0 = \begin{vmatrix} 2-x & 1 & 0 \\ 0 & 3-x & 1 \\ 0 & 0 & 2-x \end{vmatrix} = (3-x)(2-x)^2$$

Example If the matrix is upper diagonal or lower diagonal

$$T = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

then the polynomial is easy to factor.

$$0 = \begin{vmatrix} 2-x & 1 & 0 \\ 0 & 3-x & 1 \\ 0 & 0 & 2-x \end{vmatrix} = (3-x)(2-x)^2$$

These are the solutions for $\lambda_1 = 3$.

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad V_3 = \{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} z_2 \mid z_2 \in \mathbb{C} \}$$

These are for $\lambda_2 = 2$.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad V_2 = \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} z_1 \mid z_1 \in \mathbb{C} \}$$

Matrices that are similar have the same eigenvalues, but needn't have the same eigenvectors.

Example These two are similar

$$T = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{pmatrix} \qquad S = \begin{pmatrix} 6 & -1 & -1 \\ 2 & 11 & -1 \\ -6 & -5 & 7 \end{pmatrix}$$

since $S = PTP^{-1}$ for this P.

$$P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ -1/2 & 1/4 & 1/4 \\ -1/2 & -3/4 & 1/4 \end{pmatrix}$$

For the first matrix

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

is an eigenvector associated with the eigenvalue 4 but that does not hold for the second matrix.

Characteristic polynomial

3.9 *Definition* The *characteristic polynomial of a square matrix* T is the determinant $|T - xI|$ where x is a variable. The *characteristic equation* is $|T - xI| = 0$. The *characteristic polynomial of a transformation* t is the characteristic polynomial of any matrix representation $\mathrm{Rep}_{B,B}(t)$.

Note The characteristic polynomial of an $n \times n$ matrix, or of a transformation $t\colon \mathbb{C}^n \to \mathbb{C}^n$, is of degree n. Exercise 35 checks that the characteristic polynomial of a transformation is well-defined, that is, that the characteristic polynomial is the same no matter which basis we use for the representation.

Characteristic polynomial

3.9 *Definition* The *characteristic polynomial of a square matrix* T is the determinant $|T - xI|$ where x is a variable. The *characteristic equation* is $|T - xI| = 0$. The *characteristic polynomial of a transformation* t is the characteristic polynomial of any matrix representation $\mathrm{Rep}_{B,B}(t)$.

Note The characteristic polynomial of an $n \times n$ matrix, or of a transformation $t\colon \mathbb{C}^n \to \mathbb{C}^n$, is of degree n. Exercise 35 checks that the characteristic polynomial of a transformation is well-defined, that is, that the characteristic polynomial is the same no matter which basis we use for the representation.

3.10 *Lemma* A linear transformation on a nontrivial vector space has at least one eigenvalue.

Characteristic polynomial

3.9 *Definition* The *characteristic polynomial of a square matrix* T is the determinant $|T - xI|$ where x is a variable. The *characteristic equation* is $|T - xI| = 0$. The *characteristic polynomial of a transformation* t is the characteristic polynomial of any matrix representation $\mathrm{Rep}_{B,B}(t)$.

Note The characteristic polynomial of an $n \times n$ matrix, or of a transformation $t\colon \mathbb{C}^n \to \mathbb{C}^n$, is of degree n. Exercise 35 checks that the characteristic polynomial of a transformation is well-defined, that is, that the characteristic polynomial is the same no matter which basis we use for the representation.

3.10 *Lemma* A linear transformation on a nontrivial vector space has at least one eigenvalue.

Proof Any root of the characteristic polynomial is an eigenvalue. Over the complex numbers, any polynomial of degree one or greater has a root.

QED

Characteristic polynomial

3.9 *Definition* The *characteristic polynomial of a square matrix* T is the determinant $|T - xI|$ where x is a variable. The *characteristic equation* is $|T - xI| = 0$. The *characteristic polynomial of a transformation* t is the characteristic polynomial of any matrix representation $\mathrm{Rep}_{B,B}(t)$.

Note The characteristic polynomial of an $n \times n$ matrix, or of a transformation $t\colon \mathbb{C}^n \to \mathbb{C}^n$, is of degree n. Exercise 35 checks that the characteristic polynomial of a transformation is well-defined, that is, that the characteristic polynomial is the same no matter which basis we use for the representation.

3.10 *Lemma* A linear transformation on a nontrivial vector space has at least one eigenvalue.

Proof Any root of the characteristic polynomial is an eigenvalue. Over the complex numbers, any polynomial of degree one or greater has a root.

QED

Remark This result is why we switched in this chapter from working with real number scalars to complex number scalars.

Eigenspace

3.12 *Definition* The *eigenspace of a transformation* t *associated with the eigenvalue* λ is $V_\lambda = \{\vec{\zeta} \mid t(\vec{\zeta}) = \lambda\vec{\zeta}\}$. The eigenspace of a matrix is analogous.

Example Recall that this matrix has three eigenvalues, 5, 4, and 2.

$$T = \begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix}$$

Earlier, we found that these are the eigenspaces.

$$V_5 = \{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} c \mid c \in \mathbb{C}\} \quad V_4 = \{\begin{pmatrix} -7 \\ -7 \\ 1 \end{pmatrix} c \mid c \in \mathbb{C}\} \quad V_2 = \{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} c \mid c \in \mathbb{C}\}$$

3.13 *Lemma* An eigenspace is a subspace. It is a nontrivial subspace.

3.13 *Lemma* An eigenspace is a subspace. It is a nontrivial subspace.

Proof Notice first that V_λ is not empty; it contains the zero vector since $t(\vec{0}) = \vec{0}$, which equals $\lambda \cdot \vec{0}$. To show that an eigenspace is a subspace, what remains is to check closure of this set under linear combinations. Take $\vec{\zeta}_1, \ldots, \vec{\zeta}_n \in V_\lambda$ and then

$$\begin{aligned} t(c_1\vec{\zeta}_1 + c_2\vec{\zeta}_2 + \cdots + c_n\vec{\zeta}_n) &= c_1 t(\vec{\zeta}_1) + \cdots + c_n t(\vec{\zeta}_n) \\ &= c_1\lambda\vec{\zeta}_1 + \cdots + c_n\lambda\vec{\zeta}_n \\ &= \lambda(c_1\vec{\zeta}_1 + \cdots + c_n\vec{\zeta}_n) \end{aligned}$$

that the combination is also an element of V_λ.

The space V_λ contains more than just the zero vector because by definition λ is an eigenvalue only if $t(\vec{\zeta}) = \lambda\vec{\zeta}$ has solutions for $\vec{\zeta}$ other than $\vec{0}$. QED

3.18 *Theorem* For any set of distinct eigenvalues of a map or matrix, a set of associated eigenvectors, one per eigenvalue, is linearly independent.

3.18 *Theorem* For any set of distinct eigenvalues of a map or matrix, a set of associated eigenvectors, one per eigenvalue, is linearly independent.

Proof We will use induction on the number of eigenvalues. The base step is that there are zero eigenvalues. Then the set of associated vectors is empty and so is linearly independent.

3.18 *Theorem* For any set of distinct eigenvalues of a map or matrix, a set of associated eigenvectors, one per eigenvalue, is linearly independent.

Proof We will use induction on the number of eigenvalues. The base step is that there are zero eigenvalues. Then the set of associated vectors is empty and so is linearly independent.

For the inductive step assume that the statement is true for any set of $k \geqslant 0$ distinct eigenvalues. Consider distinct eigenvalues $\lambda_1, \ldots, \lambda_{k+1}$ and let $\vec{v}_1, \ldots, \vec{v}_{k+1}$ be associated eigenvectors. Suppose that $\vec{0} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1}$. Derive two equations from that, the first by multiplying by λ_{k+1} on both sides $\vec{0} = c_1\lambda_{k+1}\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1}$ and the second by applying the map to both sides $\vec{0} = c_1 t(\vec{v}_1) + \cdots + c_{k+1} t(\vec{v}_{k+1}) = c_1\lambda_1\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1}$ (applying the matrix gives the same result). Subtract the second from the first.

$$\vec{0} = c_1(\lambda_{k+1} - \lambda_1)\vec{v}_1 + \cdots + c_k(\lambda_{k+1} - \lambda_k)\vec{v}_k + c_{k+1}(\lambda_{k+1} - \lambda_{k+1})\vec{v}_{k+1}$$

The $\vec{v}_{k+1}$ term vanishes. Then the induction hypothesis gives that $c_1(\lambda_{k+1} - \lambda_1) = 0, \ldots, c_k(\lambda_{k+1} - \lambda_k) = 0$. The eigenvalues are distinct so the coefficients $c_1, \ldots, c_k$ are all 0. With that we are left with the equation $\vec{0} = c_{k+1}\vec{v}_{k+1}$ so c_{k+1} is also 0. QED

Example This matrix from above has three eigenvalues, 5, 4, and 2.

$$T = \begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix}$$

Picking a nonzero vector from each eigenspace we get this linearly independent set (which is a basis because it has three elements).

$$\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -14 \\ -14 \\ 2 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} \}$$

Example This matrix from above has three eigenvalues, 5, 4, and 2.

$$T = \begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix}$$

Picking a nonzero vector from each eigenspace we get this linearly independent set (which is a basis because it has three elements).

$$\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -14 \\ -14 \\ 2 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} \}$$

Example This upper-triangular matrix has the eigenvalues 2 and 3

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Picking a vector from each of V_3 and V_2 gives this linearly independent set.

$$\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \}$$

A criteria for diagonalizability

3.20 *Corollary* An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof Form a basis of eigenvectors. Apply Lemma 2.4 . QED

Geometry of eigenvectors

Lines go to lines

Consider a real space transformation $t\colon \mathbb{R}^n \to \mathbb{R}^n$. A defining property of linear maps is that $t(r \cdot \vec{v}) = r \cdot t(\vec{v})$.

In a real space $\mathbb{R}^n$ a line through the origin is a set $\{r \cdot \vec{v} \mid r \in \mathbb{R}\}$. So t's action

$$r \cdot \vec{v} \xmapsto{t} r \cdot t(\vec{v})$$

is to send members of the line $\{r \cdot \vec{v} \mid r \in \mathbb{R}\}$ in the domain to members of the line $\{s \cdot t(\vec{v}) \mid s \in \mathbb{R}\}$ in the codomain.

Thus, lines through the origin transform to lines through the origin. Further, the action of t is determined by its effect $t(\vec{v})$ on any nonzero vector element of the domain line.

Example Consider the line $y = 2x$ in the plane

$$\{r \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid r \in \mathbb{R}\}$$

and this transformation $t \colon \mathbb{R}^2 \to \mathbb{R}^2$ of the plane.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + 3y \\ 2x + 4y \end{pmatrix}$$

The map's effect on any vector in the line is easy to compute.

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} 7 \\ 10 \end{pmatrix}$$

The scalar multiplication property in the definition of linear map $t(r \cdot \vec{v}) = r \cdot t(\vec{v})$ imposes a uniformity on t's action: it has twice the effect on $2\vec{v}$, three times the effect on $3\vec{v}$, etc.

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} 14 \\ 20 \end{pmatrix} \qquad \begin{pmatrix} -3 \\ -6 \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} -21 \\ -30 \end{pmatrix} \qquad \begin{pmatrix} r \\ 2r \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} 7r \\ 10r \end{pmatrix}$$

In short: the action of t on any nonzero $\vec{v}$ determines its action on any other vector $r\vec{v}$ in the line $[\vec{v}]$.

Pick one, any one

Every plane vector is in some line through the origin so to understand what $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ does to plane elements it suffices to understand what it does to lines through the origin. By the prior slide, to understand what t does to a line through the origin it suffices to understand what it does to a single nonzero vector in that line.

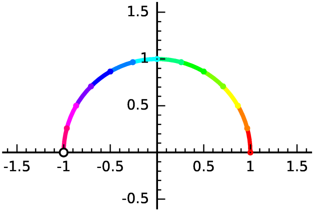
Pick one, any one

Every plane vector is in some line through the origin so to understand what $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ does to plane elements it suffices to understand what it does to lines through the origin. By the prior slide, to understand what t does to a line through the origin it suffices to understand what it does to a single nonzero vector in that line.

So one way to understand a transformation's action is to take a set containing one nonzero vector from each line through the origin, and describe where the transformation maps the elements of that set.

A natural set with one nonzero element from each line through the origin is the upper half unit circle (we will explain the colors below).

$$\{\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \mid 0 \leqslant t < \pi\}$$

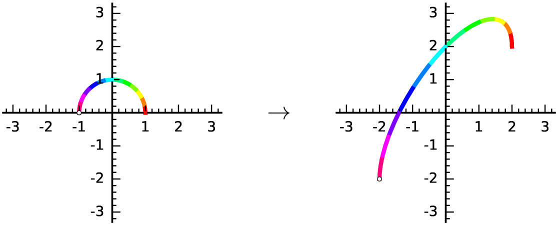


Angles

Example This plane transformation.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 2x + 2y \end{pmatrix}$$

is a skew.

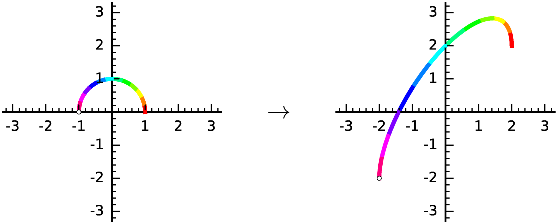


Angles

Example This plane transformation.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 2x + 2y \end{pmatrix}$$

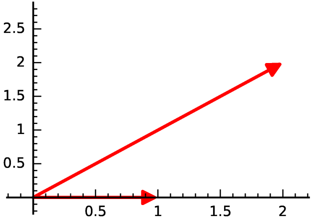
is a skew.



As we move through the unit half circle on the left, the transformation has varying effects on the vectors. The dilation vary, that is, different vectors get their length multiplied by different factors, and they are turned through varying angles. The next slide gives examples.

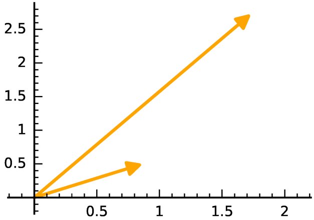
The prior slide's vector from the left shown in red is dilated by a factor of $2\sqrt{2}$ and rotated counterclockwise by $\pi/4 \approx 0.78$ radians.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$



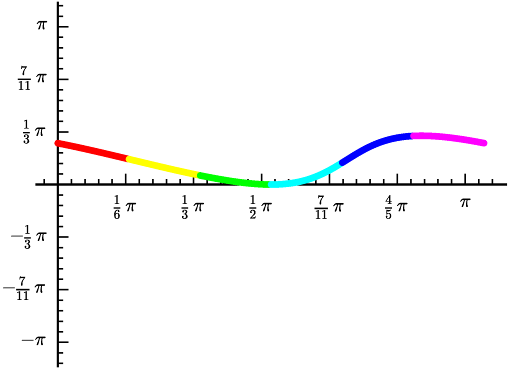
The orange vector is dilated by a factor of $2\sqrt{\cos^2(\pi/6)+1} = \sqrt{7}$ and rotated by about 0.48 radians.

$$\begin{pmatrix} \cos(\pi/6) \\ \sin(\pi/6) \end{pmatrix} \mapsto \begin{pmatrix} 2\cos(\pi/6) \\ 2\cos(\pi/6) + 2\sin(\pi/6) \end{pmatrix}$$



On the graph below the horizontal axis is the angle of a vectors from the upper half unit circle, while the vertical axis is the angle through which that vector is rotated.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 2x + 2y \end{pmatrix}$$



The rotation angle of interest is 0 radians, here achieved by some green vector.

Definition

A vector that is rotated through an angle of 0 radians or of π radians, while being dialated by a nonzero factor, is an eigenvector. The factor by which it is dilated is the eigenvalue.

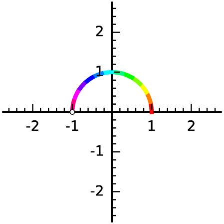
Example The plane transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ 2y \end{pmatrix}$$

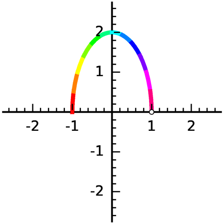
represented with respect to the standard bases by a diagonal matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

has this simple action on the upper half unit circle.

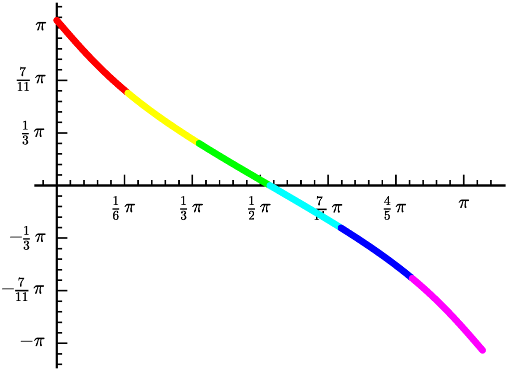


$\longrightarrow$



This plots the angle of each vector in the upper half unit circle against the angle through which it is rotated.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ 2y \end{pmatrix}$$



One vector gets zero rotation, the vector with $x = 0$.

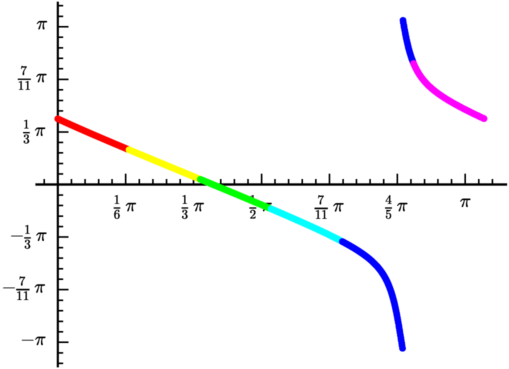
Example This generic plane transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+2y \\ 3x+4y \end{pmatrix}$$

has this action on the upper half unit circle.

Plotting the angle of each vector in the upper half unit circle against the angle through which it is rotated

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix}$$



gives that one vector gets a rotation of 0 radians, while another gets a rotation of π radians.