14. Diagonalization

14.1. Change of basis

Let $\mathcal{B} = (b_1, b_2)$ be an ordered basis for $\mathbb{R}^2$ and let $B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$ (the matrix with $b_1$ and $b_2$ as columns). If $x$ is a vector in $\mathbb{R}^2$, then its coordinate vector $[x]_{\mathcal{B}}$ relative to $\mathcal{B}$ satisfies the formula

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

We can see this by writing $[x]_{\mathcal{B}} = [\alpha_1, \alpha_2]^T$ to get

$$B[x]_{\mathcal{B}} = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \alpha_1 b_1 + \alpha_2 b_2 = x.$$

The columns of $B$ are linearly independent (since they form a basis) so $B$ has full rank and is therefore invertible. This allows us to solve the equation above for the coordinate vector:

$$[x]_{\mathcal{B}} = B^{-1}x$$

14.1.1 Example  Let $\mathcal{B} = (b_1, b_2)$ be the ordered basis of $\mathbb{R}^2$ with $b_1 = [1, 2]^T$ and $b_2 = [2, 3]^T$. Use the formula above to find the coordinate vector $[x]_{\mathcal{B}}$ where $x = [1, -1]^T$. 


Solution  We have

\[ B = \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \]

so

\[ [x]_B = B^{-1}x = \frac{1}{-1} \begin{bmatrix} 3 \\ -2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = - \begin{bmatrix} 5 \\ -3 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}. \]

(Check: \((-5)b_1 + (3)b_2 = -5 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = x).\]

Let \( C = (c_1, c_2) \) be another ordered basis for \( \mathbb{R}^2 \) and put \( C = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \). For any vector \( x \) in \( \mathbb{R}^2 \), the coordinate vectors of \( x \) relative to \( B \) and \( C \) satisfy the equation

\[ B[x]_B = C[x]_C \]

This equation holds since both sides equal \( x \) by the first equation of the section. We can solve for \( [x]_B \) to get

\[ [x]_B = B^{-1}C[x]_C \]

The matrix \( P = B^{-1}C \) is called the “transformation matrix from \( C \) to \( B \).”

14.1.2 Example  Let \( B = (b_1, b_2) \) and \( C = (c_1, c_2) \) be the ordered bases of \( \mathbb{R}^2 \) with

\[ b_1 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad c_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}. \]
and let \( \mathbf{x} \) and \( \mathbf{y} \) be vectors in \( \mathbb{R}^2 \).

(a) Find the transformation matrix \( \mathbf{P} \) from \( \mathcal{C} \) to \( \mathcal{B} \).

(b) Given that \( \mathbf{x}_\mathcal{C} = [-3, 6]^T \), find \( \mathbf{x}_\mathcal{B} \).

(c) Given that \( \mathbf{y}_\mathcal{B} = [7, 2]^T \), find \( \mathbf{y}_\mathcal{C} \).

**Solution**

(a) We have

\[
\mathbf{P} = \mathbf{B}^{-1} \mathbf{C} = \begin{bmatrix} 2 & 1 \\ 6 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ -1 & 5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 4 & -1 \\ -6 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 4 \end{bmatrix}.
\]

(b) We have

\[
\mathbf{x}_\mathcal{B} = \mathbf{P} \mathbf{x}_\mathcal{C} = \begin{bmatrix} 1/2 & -1/2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ 6 \end{bmatrix} = \begin{bmatrix} -9/2 \\ 15 \end{bmatrix}.
\]

(c) Multiplying both sides of the change of basis formula by \( \mathbf{P}^{-1} \) we get

\[
\mathbf{y}_\mathcal{C} = \mathbf{P}^{-1} \mathbf{y}_\mathcal{B} = \frac{1}{2} \begin{bmatrix} 2 & 1/2 \\ 1 & 1/2 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 15 \\ 8 \end{bmatrix} = \begin{bmatrix} 30 \\ 16 \end{bmatrix}.
\]
Here is the general formulation for a change of basis:

**CHANGE OF BASIS.**

Let $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ and $\mathcal{C} = (c_1, c_2, \ldots, c_n)$ be two ordered bases for $\mathbb{R}^n$, and put $B = [b_1 \; b_2 \; \cdots \; b_n]$ and $C = [c_1 \; c_2 \; \cdots \; c_n]$. For any vector $x$ in $\mathbb{R}^n$, we have

$$B[x]_B = C[x]_C$$

so that

$$[x]_B = B^{-1}C[x]_C.$$ 

The matrix $P = B^{-1}C$ is called the **transformation matrix from $\mathcal{C}$ to $\mathcal{B}$**.
14.2. Linear function and basis change

**Theorem.**

Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear function, let $\mathcal{B}$ and $\mathcal{C}$ be ordered bases for $\mathbb{R}^n$, let $P$ be the transformation matrix from $\mathcal{C}$ to $\mathcal{B}$, and let $A$ be the matrix of $L$ relative to $\mathcal{B}$. The matrix of $L$ relative to $\mathcal{C}$ is $P^{-1}AP$, that is,

$$[L(x)]_C = P^{-1}AP[x]_C$$

for all $x \in \mathbb{R}^n$.

**Proof.** For all $x \in \mathbb{R}^n$ we have

$$P^{-1}AP[x]_C = P^{-1}A[x]_B$$

definition of transformation matrix

$$= P^{-1}[L(x)]_B$$

$A$ is matrix of $L$ relative to $\mathcal{B}$

$$= [L(x)]_C$$

$P^{-1}$ is transformation matrix from $\mathcal{B}$ to $\mathcal{C}$.

14.2.1 Example Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear function given by

$$L(x) = \begin{bmatrix} \frac{1}{2}x_1 + \frac{3}{2}x_2 \\ \frac{3}{2}x_1 + \frac{1}{2}x_2 \end{bmatrix}.$$ 

(a) Find the matrix $A$ of $L$ (relative to the standard ordered basis $\mathcal{E} = (e_1, e_2)$).

(b) Find the transformation matrix $P$ from $\mathcal{C} = ([1, 1]^T, [-1, 1]^T)$ to $\mathcal{E}$. 

(c) Use the theorem to find the matrix of $L$ relative to $C$.

Solution

(a) We have

$$A = [L(e_1) \quad L(e_2)] = \begin{bmatrix} \frac{1}{3} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{3} \end{bmatrix}.$$

(b) Writing $E = [e_1 \quad e_2]$ we see that $E = I$, so that

$$P = E^{-1}C = I^{-1}C = C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

(c) According to the theorem, the matrix of $L$ relative to $C$ is

$$P^{-1}AP = C^{-1}AC = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

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

\[\square\]
In this example, the matrix of \( L \) relative to the basis \( C \) turned out to be much simpler than the matrix relative to the standard basis. The reason is that \( C \) is a basis for \( \mathbb{R}^2 \) consisting of eigenvectors of \( L \). We investigate such bases in the next section.

### 14.3. Method for diagonalization

The matrix
\[
\begin{bmatrix}
2 & 0 & 0 \\
0 & -3 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]
is an example of a “diagonal matrix.” In general, a **diagonal matrix** is a matrix having the property that every entry not on the main diagonal is 0.

An \( n \times n \) matrix \( A \) is **diagonalizable** if there exists an invertible \( n \times n \) matrix \( P \) such that \( P^{-1}AP = D \), where \( D \) is a diagonal matrix.

**Theorem.**

Let \( A \) be an \( n \times n \) matrix. The matrix \( A \) is diagonalizable if and only if there exists a basis for \( \mathbb{R}^n \) consisting of eigenvectors of \( A \). In this case, if \( P \) is the matrix with the eigenvectors as columns, then
\[
P^{-1}AP = D
\]
with \( D \) diagonal.

**Proof.** We will prove only the case \( n = 2 \). Assume that there exists a basis \( \{b_1, b_2\} \) for \( \mathbb{R}^2 \)
consisting of eigenvectors of \( A \). Let \( \lambda_1 \) and \( \lambda_2 \) be the corresponding eigenvalues and write 
\[
D = \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}.
\]
We have 
\[
AP = A \begin{bmatrix} b_1 & b_2 \end{bmatrix} = [Ab_1 \ Ab_2] = [\lambda_1 b_1 \ \lambda_2 b_2] = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix} = PD
\]
Therefore, \( P^{-1}AP = D \).

Conversely, if \( P \) diagonalizes \( A \), then the equation above shows that the columns of \( P \) must be eigenvectors of \( A \) and these columns are linearly independent since \( P \) is invertible. \( \square \)

14.3.1 Example Let \( A = \begin{bmatrix} 1 & 1 \\
-2 & 4
\end{bmatrix} \). If possible, find an invertible \( 2 \times 2 \) matrix \( P \) such that \( P^{-1}AP = D \), where \( D \) is a diagonal matrix.

Solution According to the theorem, such a matrix \( P \) exists if and only if there exists a basis for \( \mathbb{R}^2 \) consisting of eigenvectors of \( A \). The characteristic polynomial of \( A \) is 
\[
\det(A - \lambda I) = \begin{vmatrix}
1 - \lambda & 1 \\
-2 & 4 - \lambda
\end{vmatrix} = (1 - \lambda)(4 - \lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).
\]
The eigenvalues of \( A \) are the zeros of this polynomial, namely, \( \lambda = 2, 3 \).

Next, the \( \lambda \)-eigenspace of \( A \) is the solution set of the equation \( (A - \lambda I)x = 0 \):
\[
(\lambda = 2)
\]
\[
[A - 2I | 0] = \begin{bmatrix}
-1 & 1 & 0 \\
-2 & 2 & 0
\end{bmatrix} \sim \begin{bmatrix}
1 & -1 & 0 \\
0 & 0 & 0
\end{bmatrix},
\]
so the 2-eigenspace of \( A \) is \( \{ [t, t]^T \mid t \in \mathbb{R} \} \). Letting \( t = 1 \), we get a 2-eigenvector \( [1, 1]^T \).
(λ = 3)

\[
\begin{bmatrix}
    A - 3I & 0 \\
\end{bmatrix} = 
\begin{bmatrix}
    -2 & 1 & 0 \\
    -2 & 1 & 0
\end{bmatrix} \sim 
\begin{bmatrix}
    1 & -\frac{1}{2} & 0 \\
    0 & 0 & 0
\end{bmatrix},
\]

so the 3-eigenspace of \( A \) is \( \{[\tfrac{1}{2}, t]^T | t \in \mathbb{R} \} \). Letting \( t = 2 \) (to avoid fractions), we get a 3-eigenvector \( [1, 2]^T \).

The eigenvectors \([1, 1]^T \) and \([1, 2]^T \) of \( A \) form a basis for \( \mathbb{R}^2 \) (neither is a multiple of the other so they are linearly independent; since \( \text{dim} \mathbb{R}^2 = 2 \), they form a basis). According to the theorem, the matrix \( P \) with these vectors as columns should have the indicated property:

\[
P = \begin{bmatrix}
    1 & 1 \\
    1 & 2
\end{bmatrix}.
\]

Computing we get

\[
P^{-1}AP = \begin{bmatrix}
    1 & 1 \\
    1 & 2
\end{bmatrix}^{-1} \begin{bmatrix}
    1 & 1 \\
    1 & 2
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
    2 & -1 \\
    -1 & 1
\end{bmatrix} \begin{bmatrix}
    -1 & 1 \\
    1 & 2
\end{bmatrix} = \begin{bmatrix}
    4 & -2 \\
    -3 & 3
\end{bmatrix} \begin{bmatrix}
    1 & 1 \\
    1 & 2
\end{bmatrix} = \begin{bmatrix}
    2 & 0 \\
    0 & 3
\end{bmatrix} = D.
\]

(Note that the eigenvalues of \( A \) appear along the main diagonal of \( D \).)

14.3.2 Example  Is the matrix \( A = \begin{bmatrix}
    0 & -1 \\
    1 & 0
\end{bmatrix} \) diagonalizable? Explain.

Solution  The characteristic polynomial of \( A \) is

\[
\det(A - \lambda I) = \left| \begin{array}{cc}
    -\lambda & -1 \\
    1 & -\lambda
\end{array} \right| = \lambda^2 + 1.
\]
Since this polynomial has no zeros (in \( \mathbb{R} \)), the matrix \( A \) has no eigenvalues. Therefore, there is no basis for \( \mathbb{R}^2 \) consisting of eigenvectors of \( A \) and \( A \) is not diagonalizable according to the theorem.

(Here’s another way to see that \( A \) has no eigenvalues. \( A \) is the matrix of the linear function “90° clockwise rotation.” Since this function sends no nonzero vector to a multiple of itself, it has no eigenvalues, and therefore \( A \) has no eigenvalues either.)

### 14.4. Power of matrix

It is often necessary to compute high powers of a square matrix \( A \), such as \( A^{10} \). Just multiplying the matrix \( A \) by itself over and over again can be quite tedious. However, if \( A \) is diagonalizable, then there is an observation that greatly reduces the number of computations:

Let \( A \) be a diagonalizable matrix, so that \( P^{-1}AP = D \) with \( D \) diagonal. Solving this equation for \( A \), we get \( A = PDP^{-1} \). Note that

\[
A^2 = AA = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} \\
= PDIDP^{-1} = PDDP^{-1} = PD^2P^{-1},
\]

and in less detail

\[
A^3 = (PDP^{-1})(PDP^{-1})(PDP^{-1}) = PD^3P^{-1}.
\]

In general,

\[
P^{-1}AP = D \quad \Rightarrow \quad A^n = PD^nP^{-1}
\]
Note that since $D$ is diagonal, the power $D^n$ is obtained by raising each diagonal entry to the $n$th power.

### 14.4.1 Example

Compute $A^{10}$, where

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}.$$ 

**Solution** In Example 14.3.1 we found a matrix $P$ such that $P^{-1}AP = D$ with $D$ diagonal. Using the results of that example we get

$$A^{10} = PD^{10}P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2^{10} & 0 \\ 0 & 3^{10} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2^{10} & 3^{10} \\ 2^{10} & 2 \cdot 3^{10} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2^{11} - 3^{10} & -2^{10} + 3^{10} \\ 2^{11} - 2 \cdot 3^{10} & -2^{10} + 2 \cdot 3^{10} \end{bmatrix}$$

$$= \begin{bmatrix} -57001 & 58025 \\ -116050 & 117074 \end{bmatrix}$$
14 – Exercises

14–1 Let \( \mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2) \) and \( \mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2) \) be the ordered bases of \( \mathbb{R}^2 \) with

\[
\mathbf{b}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}.
\]

and let \( \mathbf{x} \) be a vector in \( \mathbb{R}^2 \).

(a) Find the transformation matrix \( \mathbf{P} \) from \( \mathcal{C} \) to \( \mathcal{B} \).

(b) Given that \( [\mathbf{x}]_\mathcal{C} = [1, 2]^T \), use part (a) to find \( [\mathbf{x}]_\mathcal{B} \) and check your answer by showing that both coordinate vectors yield the same vector \( \mathbf{x} \).

14–2 Let \( \mathcal{L} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) be “reflection across the line \( x_2 = x_1/\sqrt{3} \)” (the line through the origin making an angle of \( 30^\circ \) with the \( x_1 \)-axis).

(a) Find the matrix \( \mathbf{A} \) of \( \mathcal{L} \) (relative to the standard ordered basis \( \mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2) \)).

(b) Find the transformation matrix \( \mathbf{P} \) from \( \mathcal{C} = ([\sqrt{3}, 1]^T, [-1, \sqrt{3}]^T) \) to \( \mathcal{E} \).

(c) Use the theorem in Section 14.2 to find the matrix of \( \mathcal{L} \) relative to \( \mathcal{C} \).

HINT: Recall the 30-60-90° triangle with legs 1, 2, and \( \sqrt{3} \).
Let \( A = \begin{bmatrix} 5 & 6 \\ -2 & -2 \end{bmatrix} \).

(a) If possible, find an invertible \( 2 \times 2 \) matrix \( P \) such that \( P^{-1}AP = D \), where \( D \) is a diagonal matrix.

(b) Find \( A^{10} \). (Hint: Use the formula in Section 14.4.)