Math 4326
Fall 2018

Matrices related to linear transformations

We have encountered several ways in which matrices relate to linear transformations. In this note, I summarize the important facts and formulas we have encountered.

The matrix of a linear transformation from R^n to R^m.

Theorem 1 *Given a linear transformation $T : R^n \to R^m$, there is an $m \times n$ matrix A for which $T(v) = Av$ for all v in R^n. This matrix is*

$$A = (T(e_1) \,|\, T(e_2) \,|\, \cdots \,|\, T(e_n)), \tag{1}$$

where $\{e_1, e_2, \ldots, e_n\}$ is the standard basis for R^n.

Thus, the only transformations from R^n to R^m are matrix transformations, and we can calculate the range of T by finding the column space of A and the kernel of T by finding the null space of A.

Transition matrices

Given two bases $B = \{v_1, v_2, \ldots, v_n\}$ and $C = \{w_1, w_2, \ldots, w_n\}$ for the same vector space V, we can change coordinates with respect to one basis to coordinates with respect to the other via matrix multiplication by an appropriate matrix. That is, $[v]_C = P[v]_B$ for some matrix P. This matrix, P, is called the transition matrix from B to C. Our book uses the notation $\underset{C \leftarrow B}{P}$ to emphasize that we are changing from B to C coordinates using P. The formula for P is

$$\underset{C \leftarrow B}{P} = ([v_1]_C \,|\, [v_2]_C \,|\, \cdots \,|\, [v_n]_C). \tag{2}$$

In words, to go from B to C, P is obtained by forming the C-coordinates of each B-basis vector. Also, you should note the similarity between this formula and the one above. To make it look even more similar, define $T : V \to R^n$ by $T(v) = [v]_C$. That is, let T be the coordinate map with respect to C-coordinates. As mentioned in class, the coordinate map is linear. We can rewrite our matrix in the form:

$$\underset{C \leftarrow B}{P} = (T(v_1) \,|\, T(v_2) \,|\, \cdots \,|\, T(v_n)). \tag{3}$$

There are two ways to get the transition matrix from C to B. If we call this matrix Q, then $Q = \underset{B \leftarrow C}{P}$. The direct approach is that

$$Q = \underset{B \leftarrow C}{P} = ([w_1]_B \,|\, [w_2]_B \,|\, \cdots \,|\, [w_n]_B).$$

The indirect approach is that $Q = P^{-1}$. That is,

$$\underset{B \leftarrow C}{P} = (\underset{C \leftarrow B}{P})^{-1}. \tag{4}$$

The matrix of a linear transformation from V to W, with respect to bases B for V and C for W

If we don't have $T : R^n \to R^m$ then then Theorem 1 does not apply. In fact, there are matrices that represent linear transformations, **but only if we use coordinate systems**.

Theorem 2 *Let $T : V \to W$ be a linear transformation. Suppose $B = \{v_1, v_2, \dots, v_n\}$ is a basis for V, and C is a basis for W. Then there is a matrix A, for which $[T(v)]_C = A[v]_B$. In fact, we have a formula for A:*

$$A = ([T(v_1)]_C \,|\, [T(v_2)]_C \,|\, \cdots \,|\, [T(v_n)]_C). \tag{5}$$

We have notation for this matrix. It is usually denoted $[T]_{B,C}$. Again, you should note how similar (5) is to (3), and even to (1).

Proof of theorem: Let v be a vector in V. Then for some scalars $c_1, c_2, \dots, c_n$, we have $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$. Given this,

$$\begin{aligned} T(v) &= T(c_1v_1 + c_2v_2 + \cdots + c_nv_n) \\ &= c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n). \end{aligned}$$

Now

$$\begin{aligned} [T(v)]_C &= [c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n)]_C \\ &= c_1[T(v_1)]_C + c_2[T(v_2)]_C + \cdots + c_n[T(v_n)]_C \\ &= ([T(v_1)]_C \,|\, [T(v_2)]_C \,|\, \cdots \,|\, [T(v_n)]_C) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\ &= A[v]_B, \end{aligned}$$

as desired.

A natural question to ask is how the range and kernel of T relate to A. It is NOT the case that the range of T is the column space of A, or that $\mathrm{Ker}(T) = \mathrm{Null}(A)$. It is important that you understand why they are not equal, so take some time to think about this. What IS true is the following: The column space of A gives the **coordinates** for the vectors in $\mathrm{Range}(T)$ and $\mathrm{Null}(A)$ gives the **coordinates** for the vectors in $\mathrm{Ker}(T)$. To find a basis for the range of T, find a basis for the column space of A and these will be the coordinates for the basis vectors for the range. Similarly, to find a basis for the kernel of T, first find a basis for $\mathrm{Null}(A)$. These vectors are the coordinates for the basis for $\mathrm{Ker}(T)$.

As an example, consider a problem from Homework 8: $T : P_2 \to M_{2\times 2}$ defined by $T(p(t)) = \begin{pmatrix} p(1) + p(2) & p(2) \\ p(1) & p(1) - p(2) \end{pmatrix}$. We will calculate $[T]_{B,C}$, where B is the standard basis for P_2 and C is the standard basis for $M_{2\times 2}$. We have

$$T(1) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \to [T(1)]_C = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

$$T(t) = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix} \to [T(t)]_C = \begin{pmatrix} 3 \\ 2 \\ 1 \\ -1 \end{pmatrix},$$

$$T(t^2) = \begin{pmatrix} 5 & 4 \\ 1 & -3 \end{pmatrix} \to [T(t^2)]_C = \begin{pmatrix} 5 \\ 4 \\ 1 \\ -3 \end{pmatrix}.$$

We now have the columns for $A = [T]_{B,C}$, giving

$$A = [T]_{B,C} = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 0 & -1 & -3 \end{pmatrix}.$$

Suppose we wish to calculate $T(2t^2 + t + 1)$ by matrix multiplication using the formula $[T(v)]_C = A[v]_B$ instead of just using the formula for T. Here is what we would do: First find the coordinates for $2t^2+t+1$: $[2t^2+t+1]_B = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$. Next, multiply by A:
$\begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 0 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 15 \\ 11 \\ 4 \\ -7 \end{pmatrix}$. This is NOT the answer. Instead, what we have is
$[T(v)]_C = \begin{pmatrix} 15 \\ 11 \\ 4 \\ -7 \end{pmatrix}$. Finally, converting from coordinates to matrices, $T(v) = \begin{pmatrix} 15 & 11 \\ 4 & -7 \end{pmatrix}$.

To find the kernel and range of T, we may first find the null space and column space of A. This will give us coordinates of the respective bases. We reduce A:

$$\begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 0 & -1 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

A basis for the column space of A consists of the first two columns of A. These are the coordinates for the basis vectors for Range(T) so this basis is $\left\{ \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix} \right\}$. A basis for Null($A$) is $\left\{\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}\right\}$. Converting to P_2, a basis for the kernel of T is $\{t^2 - 3t + 2\}$.

Mixing bases

You can safely skip this section.

Suppose that $T : V \to W$ and we have several bases: B_1, B_2 for V and C_1, C_2 for W. Then we have our choice of bases to use to find the matrix of the transformation. How are the matrices related to each other? They are related by transition matrices among the bases.

Theorem 3 *Let $T : V \to W$ be a linear transformation and suppose B_1 and B_2 are bases for V and C_1 and C_2 are bases for W. Then*

$$[T]_{B_2,C_2} = \underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} \underset{B_1 \leftarrow B_2}{P}.$$

Proof. In general, if you have two matrices, A and B, and if $Av = Bv$ for all v, then $A = B$. (Can you prove this?) We use this result on matrices to show that the two matrices

$$[T]_{B_2,C_2} \text{ and } \underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} \underset{B_1 \leftarrow B_2}{P}$$

are the same. Given a vector v in V, by the defining property of the matrix of a transformation,

$$[T]_{B_2,C_2}[v]_{B_2} = [T(v)]_{C_2}.$$

On the other hand,

$$\begin{aligned} \left(\underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} \underset{B_1 \leftarrow B_2}{P}\right)[v]_{B_2} &= \underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} \left(\underset{B_1 \leftarrow B_2}{P} [v]_{B_2}\right) \\ &= \underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} [v]_{B_1} \\ &= \underset{C_2 \leftarrow C_1}{P} [T(v)]_{C_1} \\ &= [T(v)]_{C_2}. \end{aligned}$$

That is, for all vectors v,

$$[T]_{B_2,C_2}[v]_{B_2} = \left(\underset{C_2 \leftarrow C_1}{P} [T]_{B_1,C_1} \underset{B_1 \leftarrow B_2}{P}\right)[v]_{B_2},$$

so the two matrices must be the same.

I think of this as follows: $[T]_{B_1,C_1}$ expects to see B_1 coordinates, and $[T]_{B_2,C_2}$ expects to see B_2 coordinates. If you want to feed a matrix B_2 coordinates and get C_2-coordinate answers, the direct approach is to use $[T]_{B_2,C_2}$. However, there is the following indirect route:

use the transition matrix, $\underset{B_1 \leftarrow B_2}{P}$, to convert B_2 coordinates to B_1 coordinates, then multiply by $[T]_{B_1,C_1}$, giving an answer in C_1 coordinates. Finally, multiply by another transition matrix, $\underset{C_2 \leftarrow C_1}{P}$, to get the desired C_2-coordinate answer.

Continuing our previous example, suppose that $B_1 = \{1, t-1, t^2-3t+2\}$, $C_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ and we wish to calculate $[T]_{B_1,C_1}$ for the transformation on Page 3. We could do this in two ways. The direct approach is to use:

$$\begin{aligned} [T]_{B_1,C_1} &= ([T(1)]_{C_1} \mid [T(t-1)]_{C_1} \mid [T(t^2-3t+2)]_{C_1}) \\ &= \left\{ \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}_{C_1} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}_{C_1} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}_{C_1} \right\} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Alternatively, we may use the formula from Theorem 3:

$$[T]_{B_1,C_1} = \underset{C_1 \leftarrow C}{P} [T]_{B,C} \underset{B \leftarrow B_1}{P}$$

where B is the standard basis for P_2 and C is the standard basis for $M_{2\times 2}$. The transition matrices are:

$$\underset{B \leftarrow B_1}{P} = ([1]_B \mid [t-1]_B \mid [t^2-3t+2]_B) = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\underset{C_1 \leftarrow C}{P} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_{C_1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_{C_1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_{C_1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_{C_1} \right\} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix}.$$

By the theorem,

$$\begin{aligned} [T]_{B_1,C_1} &= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 0 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

This second matrix also contains all the information about T. For example, if we were asked to find the kernel of T, we could use this second matrix. The null space of the matrix is spanned by $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. This is different from before, but remember, this is not the kernel, it is a coordinate vector for a kernel vector. Using the basis B_1, the kernel is spanned by the vector $0 \cdot 1 + 0(t-1) + 1(t^2 - 3t + 2) = t^2 - 3t + 2$. That is, once we convert from coordinates to a vector in P_2, we get exactly the same answer.

Linear operators, eigenvectors, and diagonalization
This section is important, at least for the examples.

A linear operator is a linear transformation from a space to itself. That is, rather than $T : V \to W$, we have $T : V \to V$. That is, V and W are the same space. In this case, we only need one basis, B of V. If $B = \{v_1, v_2, \ldots, v_n\}$, then we write

$$[T(v)]_B = A[v]_B,$$

where

$$A = ([T(v_1)]_B \,|\, [T(v_2)]_B \,|\, \cdots \,|\, [T(v_n)]_B).$$

We use the notation $A = [T]_B$ rather than $[T]_{B,B}$. In the case of a linear operator, with two bases, B and C for V, Theorem 3 becomes

$$[T]_C = \underset{C \leftarrow B}{P}[T]_B \underset{B \leftarrow C}{P}.$$

Since $\underset{C \leftarrow B}{P} = \left(\underset{B \leftarrow C}{P}\right)^{-1}$, we can write this

$$[T]_C = P^{-1}[T]_B P,$$

where $P = \underset{B \leftarrow C}{P}$. That is, all matrices that represent T are **similar** to each other. In fact, this is what **similarity** actually means! That is, if A and B are similar to each other, then there is a linear transformation T for which A and B are the matrices representing T with respect to some bases.

Here are three examples of these formulas. First, consider $T : P_3 \to P_3$ defined by $T(p(t)) = (t+2)p'(t)$. Let S be the standard basis for P_3, $\{1, t, t^2, t^3\}$ and let $C = \{1, t+2, (t+2)^2, (t+2)^3\}$ be a second basis. We have

$$A = [T]_S = ([T(1)]_S \,|\, [T(t)]_S \,|\, [T(t^2)]_S \,|\, [T(t^3)]_S)$$

$$= ([0]_S \,|\, [t+2]_S \,|\, [2t^2 + 4t]_S \,|\, [3t^3 + 6t^2]_S)$$

$$= \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Letting $B = [T]_C$, we could calculate B directly:

$$\begin{aligned} B = [T]_C &= ([T(1)]_C \mid [T(t+2)]_C \mid [T((t+2)^2)]_C \mid [T((t+2)^3)]_C) \\ &= ([0]_C \mid [t+2]_C \mid [2(t+2)^2]_C \mid [3(t+2)^3]_C) \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

To calculate B using the formula, we have $B = P^{-1}AP$, where P is the transition matrix from C to S. First we get the transition matrix:

$$P = \underset{S \leftarrow C}{P} = ([1]_S \mid [t+2]_S \mid [(t+2)^2]_S \mid [(t+2)^3]_S) = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

I will let you check that $P^{-1} = \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Given this, we should have

$$\begin{aligned} B &= \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 8 & 24 \\ 0 & 1 & 8 & 36 \\ 0 & 0 & 2 & 18 \\ 0 & 0 & 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

As you should notice from this example, the columns of P are eigenvectors of A, and P is a diagonalizing matrix for A. One of the aspects of linear operators is that they have eigenvectors and eigenvalues, just like matrices. We say c is an eigenvalue for a linear operator T if there is a nonzero vector v in V with $T(v) = cv$. Similarly, we call v an eigenvector for T.

How does one go about finding eigenvectors and eigenvalues for a linear operator? In differential equations, you studied "differential operators," which were linear operators on vector spaces of functions, with differentiation being part of the definition of the operator. Part of what happened in that class involved finding eigenvectors and eigenvalues for differential operators. In this class, we relate everything back to matrices. So for the example above, we calculated the matrix of the transformation. It turns out that the eigenvalues of a transformation are the same as the eigenvalues for any matrix representing that transformation. Thus, T had 0, 1, 2, 3 as eigenvalues. What about eigenvectors? The are NOT the columns of the diagonalizing matrix P. However, the columns of P are the coordinates for the eigenvectors. So, for this example, $\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ is an eigenvector for the matrix representing T, and it gives the coordinates for an eigenvector for T. Going from coordinates to the vector, $\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \to 2 \cdot 1 + 1 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 = t + 2.$

We define a linear operator T to be diagonalizable if there is some basis B for which $[T]_B$ is a diagonal matrix. So our example above, $T(p(t)) = (t+2)p'(t)$ is diagonalizable since $[T]_B$ is a diagonal matrix with the basis $\{1, t+2, (t+2)^2, (t+2)^3\}$. Eigenvalue/eigenvector/diagonalizability properties for linear operators are very similar to those for matrices. For example, here is a good Final Exam question: Prove that if c is an eigenvalue for T, then the set of all v for which $T(v) = cv$ is a subspace of V, the domain space. That is, prove that the eigenspace is a subspace. The following is the main theorem on diagonalizability of linear transformations.

Theorem 4 *A linear operator T on a finite dimensional vector space V is diagonalizable if and only if V has a basis consisting of eigenvectors of T.*

Here is half of the proof: Suppose that $B = \{v_1, v_2, \ldots, v_n\}$ is a basis for V and each v_i is an eigenvector for T. That is, suppose that $T(v_i) = c_i v_i$ for each i. The matrix of T with respect to B is $([T(v_1)]_B \,|\, [T(v_2)]_B \,|\, \cdots \,|\, [T(v_n)]_B)$. That is, we apply T to each basis vector, get the coordinates of the answer, and make that the appropriate column of the matrix. Since each vector is an eigenvector, we have $([T(v_1)]_B \,|\, [T(v_2)]_B \,|\, \cdots \,|\, [T(v_n)]_B) = ([c_1 v_1]_B \,|\, [c_2 v_2]_B \,|\, \cdots \,|\, [c_n v_n]_B)$. But the B-coordinates of v_i is e_i, the i'th standard basis vector. That is, $[T]_B = (c_1 e_1 \,|\, c_2 e_2 \,|\, \cdots \,|\, c_n e_n)$, but this is just the diagonal matrix with c's on the diagonal.

For a second example, let's use matrix spaces: Define $T : M_{2\times 2} \to M_{2\times 2}$ by

$$T(A) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} A \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Problem: Find the matrix of T with respect to the standard basis for $M_{2\times 2}$, find all eigenvalues of T, a basis for each eigenspace, and finally, find the matrix of T with respect to a basis of eigenvectors.

Solution: As an aid to the calculations, let's first calculate a formula for $T(A)$ in terms of the entries of the matrix A. We have

$$\begin{aligned} T\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2a+b & a+2b \\ 2c+d & c+2d \end{pmatrix} \\ &= \begin{pmatrix} 2a+b+2c+d & a+2b+c+2d \\ 2a+b+2c+d & a+2b+c+2d \end{pmatrix}. \end{aligned}$$

We now apply T to the standard basis vectors:

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

Getting the standard coordinates of each vector gives

$$[T]_S = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}.$$

Next, the eigenvalues of T are the same as those for $[T]_S$. We get the characteristic polynomial for this matrix:

$$\det(xI - [T]_S) = \begin{vmatrix} x-2 & -1 & -2 & -1 \\ -1 & x-2 & -1 & -2 \\ -2 & -1 & x-2 & -1 \\ -1 & -2 & -1 & x-2 \end{vmatrix} = \begin{vmatrix} x & 0 & -x & 0 \\ 0 & x & 0 & -x \\ -2 & -1 & x-2 & -1 \\ -1 & -2 & -1 & x-2 \end{vmatrix}$$

(I subtracted the third row from the first and the fourth from the second)

$$\begin{aligned} &= x^2\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -2 & -1 & x-2 & -1 \\ -1 & -2 & -1 & x-2 \end{vmatrix} = x^2\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & x-4 & -2 \\ 0 & 0 & -2 & x-4 \end{vmatrix} \\ &= x^2\begin{vmatrix} x-4 & -2 \\ -2 & x-4 \end{vmatrix} = x^2[(x-4)^2 - 4] = x^2(x-2)(x-6). \end{aligned}$$

Thus, the eigenvalues are 0, 2, 6. Next, we find bases for the eigenspaces of $[T]_S$. These will NOT be bases for the eigenspaces for T. I will let you go through the details, but the matrix $[T]_S$ has eigenspace bases:

$$0 : \left\{\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}\right\}, \quad 2 : \left\{\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}\right\}, \quad 6 : \left\{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\right\}.$$

These give COORDINATES for the bases for the actual eigenspaces of T. That is, if we convert from coordinate vectors to matrices, we get the actual eigenvectors. These give us

$$0 : \left\{ \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} \right\}, \quad 2 : \left\{ \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \right\}, \quad 6 : \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}.$$

A basis for $M_{2\times 2}$ consisting of eigenvectors of T is

$$B = \left\{ \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}.$$

With respect to this basis, the matrix of T is

$$[T]_B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

You should check that $[T]_B = P^{-1}[T]_S P$ for an appropriate P. As a check on the calculations above, I will use this second matrix for T to calculate $T\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}$. To do this, we need the formula $[T(v)]_B = [T]_B[v]_B$. We need to write $\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}$ as a linear combination of vectors in B. Calling these vectors b_1, b_2, b_3, b_4, we have $\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} = b_1 + 2b_2 - b_3 + b_4$, so

$$\left[T\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} \right]_B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 6 \end{pmatrix}.$$

This tells us the coordinates of the answer, we need -2(third basis vector) + 6(fourth basis vector). Thus,

$$T\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} = -2\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} + 6\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 4 \\ 8 & 4 \end{pmatrix}.$$

The direct calculation would have been

$$T\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -3 \\ 8 & 7 \end{pmatrix} = \begin{pmatrix} 8 & 4 \\ 8 & 4 \end{pmatrix}.$$

One last example. Let $V = \{(x, y, z) \,|\, x + y + z = 0\}$. That is, let V be a plane in R^3. A basis for V is $B = \{u, v\} = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$. Let $T : V \to V$ be defined by

$$T(x, y, z) = (x + 2y, x + 2z, 0).$$

I will leave it to you to show that T really is a linear operator on V. (What this means is that if w is any vector in V, then $T(w)$ should also be in V. For example, (5, -3, -2) is in V, and $T(5,-3,-2) = (-1,1,0)$ is also in V.)

The problem: Find the matrix of T with respect to B, find all eigenvalues and a basis for each eigenspace, and find the matrix of T with respect to a basis of eigenvectors. Implicit in this problem is that T IS diagonalizable.

Note that even though T is defined in terms of vectors in R^3, the matrix of T will be a 2×2 matrix. That is, it is the size of the basis (the dimension of V) that determines the size of the matrix. To get the matrix of T, we apply T to each basis vector and write than answer as a combination of basis vectors. We have

$$T(u) = (-1,1,0) = u - v, \qquad T(v) = (-2,2,0) = 2u - 2v.$$

Getting the coordinates for $T(u)$ and $T(v)$, we put these in as columns in the matrix:

$$[T]_B = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} = A.$$

The eigenvalues of T are the same as the eigenvalues of $[T]_B$, so

$$\det(xI - [T]_B) = \begin{vmatrix} x-1 & -2 \\ 1 & x+2 \end{vmatrix} = (x-1)(x+2) + 2 = x^2 + x.$$

The eigenvalues of T are 0 and -1. For 0, we have $\begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, giving a basis $\left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$. For -1, $A + I = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, giving a basis $\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$. These are not eigenvectors for T, but they are the coordinate vectors for the eignevectors. The corresponding eigenvectors for T are $-2u + v = (2,-1,-1)$ and $-u + v = (1,-1,0)$. Since V is 2-dimensional, and these come from different eigenspaces, they must be a basis for V. So we have found our basis of eigenvectors: $C = \left\{ \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\} = \{w_1, w_2\}$. Finally, we calculate the matrix of T with respect to C: $T(w_1) = 0$, $T(w_2) = -w_2$, so $[T]_C = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$. Based on all we've done, the matrix $\begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$ should be similar to $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$. It is, of course, using $P = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}$, which has inverse $P^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}$. That is, $P^{-1}AP = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$, which we could write

$$[T]_C = P^{-1}[T]_B P,$$

and interpret either as (1): P is the matrix with eigenvectors of A as its columns, or (2): P is the transition matrix $\underset{B \leftarrow C}{P}$.

One final comment. Not all linear operators are diagonalizable. For example, the linear operator $T : P_2 \to P_2$ defined by $T(p(t)) = p(t+1)$ is not diagonalizable. This T is sometimes called a shift operator. As an example, $T(t^2 + 2t + 3) = (t+1)^2 + 2(t+1) + 3 = t^2 + 2t + 1 + 2t + 2 + 3 = t^2 + 4t + 6$. You should be able to verify the following: The matrix of T with respect to the standard basis of P_2 is $[T]_S = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, and this matrix has only one eigenvalue ($\lambda = 1$), and only a 1-dimensional eigenspace. Since the matrix is not diagonalizable, neither is T.