MATH 1300 – Diagonalisation

Julien Arino

November 29, 2018

Purpose

This is a version of Section 5.2 in the Anton book that does not rely on material not covered in MATH 1300.

1 Matrix diagonalisation problem

Recall that a linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ can be represented using its associated *standard matrix* A, where for $\mathbf{x} \in \mathbb{R}^n$,

$$T(\mathbf{x}) = A\mathbf{x}.$$

If T is a linear transformation and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are so-called *standard basis vectors*, then the vectors

$$\mathbf{f}_1 = T(\mathbf{e}_1), \ldots, \mathbf{f}_n = T(\mathbf{e}_n)$$

can be used to understand the effect of the transformation. We have also seen that if the standard matrix A associated to T is invertible, then A^{-1} can be used to find the *pre-image* of points: if $\mathbf{y} \in \mathbb{R}^n$ is a point/vector, then $\mathbf{x} \in \mathbb{R}^n$ defined by $\mathbf{x} = A^{-1}\mathbf{y}$ is the point/vector such that $T(\mathbf{x}) = \mathbf{y}$.

Example 1:
The standard matrix of

$$T(x, y, z) = (2x, x + y, 4x + y + 3z),$$

is

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

We have

$$T(\mathbf{e}_1) = T(1,0,0) = (2,1,4) = 2\mathbf{e}_1 + \mathbf{e}_2 + 4\mathbf{e}_3,$$
$$T(\mathbf{e}_2) = T(0,1,0) = (0,1,1) = \mathbf{e}_2 + \mathbf{e}_2$$

and

$$T(\mathbf{e}_3) = T(0,0,1) = (0,0,3) = 3\mathbf{e}_3.$$

The matrix diagonalisation problem can be stated as follows in terms of linear transformations.

Diagonalisation problem (1^{st} version). Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, is it possible to find a change of coordinates and a linear transformation $R : \mathbb{R}^n \to \mathbb{R}^n$ having the same effect as T but such that

$$R(x_1, \ldots, x_n) = (k_1 x_1, \ldots, k_n x_n),$$

with $k_1, \ldots, k_n$ scalars (numbers)?

If we do find such a linear transformation R, then computations using the transformation will be much easier. We will see later that in Example 1, with a suitable change of coordinates, we have a linear transformation R such that, in terms of the new coordinates $(\tilde{x}, \tilde{y}, \tilde{z})$,

$$R(\tilde{x}, \tilde{y}, \tilde{z}) = (\tilde{x}, 2\tilde{y}, 3\tilde{z}).$$

2 Similarity

Definition 1. *If A and B are square matrices, then we say that B **is similar to** A if there is an invertible matrix P such that $B = P^{-1}AP$.*

We say that B is a **similarity transformation** of A. Note that if B is similar to A, then it is also true that A is similar to B, since we can express B as $B = Q^{-1}AQ$ by taking $Q = P^{-1}$. This being the case, we will usually say that A and B are **similar matrices** if either is similar to the other.

Similarity is a very convenient property. Suppose we have a given $n \times n$ matrix A. Then all $n \times n$ matrices X that are similar to A enjoy a certain number of the same properties as A. For instance, suppose X and A are similar, i.e., $X = P^{-1}AP$ for some invertible matrix P, then it follows that

$$\det(X) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \det(A)\frac{\det(P)}{\det(P)} = \det(A),$$

so all matrices similar to A have the same determinant as A. We say that the determinant is **invariant under similarity**. There are a number of such similarity invariants. Some are summarized in Table 2. You should try to prove some of these as we have done for the determinant.

Using similarity, the diagonalisation problem can be stated as follows.

Table 1: Similarity invariants.

Property	Description
Determinant	A and $P^{-1}AP$ have same determinant
Invertibility	A invertible if and only if $P^{-1}AP$ invertible
Characteristic polynomial	A and $P^{-1}AP$ have same characteristic polynomial
Eigenvalues	A and $P^{-1}AP$ have same eigenvalues

Diagonalisation problem (2$^{\text{nd}}$ version). Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with associated standard matrix A, is it possible to find a linear transformation $R : \mathbb{R}^n \to \mathbb{R}^n$ with a diagonal standard matrix D, with A and D similar, i.e., such that

$$D = P^{-1}AP$$

for some invertible matrix P.

3 The "nice" case

We will only consider a case that works very well. More elaborate cases require more advanced material and are beyond the scope of this course. The following result tells us that, under some assumptions about the eigenvalues of A, the answer to the question in the diagonalisation problem is positive.

Theorem 2. *Let A be an $n \times n$ matrix. If A has n distinct eigenvalues, then there exists an invertible $n \times n$ matrix P such that*

$$D = P^{-1}AP,$$

where D is a diagonal $n \times n$ matrix with the eigenvalues of A along the diagonal, i.e., if the distinct eigenvalues of A are $\lambda_1, \dots, \lambda_n$, then

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

We also write $D = \mathsf{diag}(\lambda_1, \dots, \lambda_n)$. We skip the proof of this result for now and focus on how to use it. For this, we give the following procedure, in which A is $n \times n$.

Procedure for diagonalising a matrix.

1. Compute all eigenvalues $\lambda_1, \dots, \lambda_n$ of A and check that the conditions of Theorem 2 hold, i.e., that they are all distinct.

2. Compute all eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ associated to these eigenvalues, i.e., for $j = 1, \ldots, n$, find the $\mathbf{v}_j \neq \mathbf{0}$ such that

$$(A - \lambda_j \mathbb{I})\mathbf{v}_j = \mathbf{0},$$

where $\mathbb{I}$ is the $n \times n$ identity matrix.

3. Form the matrix P with the eigenvectors as columns, i.e.,

$$P = [\mathbf{v}_1 \cdots \mathbf{v}_n].$$

4. The matrix D with eigenvalues on the diagonal is then given by $D = P^{-1}AP$.

Be careful to list the eigenvectors as columns in P and eigenvalues as diagonal entries in D in the same order.

Example 2:

We return to Example 1, where the matrix was

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

Let us compute the eigenvalues. As A is a lower triangular matrix, its eigenvalues are the diagonal entries, i.e, 2, 1 and 3. If you to not know the result, remark that

$$\begin{aligned} 0 &= \det(A - \lambda\mathbb{I}) && (1) \\ &= \begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 1-\lambda & 0 \\ 4 & 1 & 3-\lambda \end{vmatrix} && (2) \\ &= (2-\lambda)\begin{vmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} && (3) \\ &= (2-\lambda)(1-\lambda)(3-\lambda), && (4) \end{aligned}$$

from which the eigenvalues 2, 1 and 3 follow directly. As all eigenvalues are distinct, the conditions of Theorem 2 hold and we can use the procedure for diagonalising the matrix. So we now seek eigenvectors of A corresponding to the eigenvalues we found.

Let us start with $\lambda_1 = 1$. We seek $\mathbf{v}_1 \neq 0$ such that $(A - \mathbb{I})\mathbf{v}_1 = \mathbf{0}$, i.e., denoting $\mathbf{v}_1 = (x, y, z)$,

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 4 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This gives the augmented matrix (we omit the column of zeros to the right as it will remain unchanged by elementary row operations),

$$\begin{pmatrix}1&0&0\\1&0&0\\4&1&2\end{pmatrix} \xrightarrow{R_2\leftrightarrow R_3} \begin{pmatrix}1&0&0\\4&1&2\\1&0&0\end{pmatrix} \xrightarrow{R_3\leftarrow R_3-R_1} \begin{pmatrix}1&0&0\\4&1&2\\0&0&0\end{pmatrix} \xrightarrow{R_2\leftarrow R_2-4R_1} \begin{pmatrix}1&0&0\\0&1&2\\0&0&0\end{pmatrix}.$$

Thus we find

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\-2t\\t\end{pmatrix} = \begin{pmatrix}0\\-2\\1\end{pmatrix} t, \quad t \in \mathbb{R},$$

i.e., an eigenvector associated to $\lambda_1 = 1$ is

$$\mathbf{v}_1 = \begin{pmatrix}0\\-2\\1\end{pmatrix}.$$

Similarly, we find an eigenvector associated to $\lambda_2 = 2$ to be $\mathbf{v}_2 = (1, 1, -5)^T$ and one associated to $\lambda_3 = 3$ as $\mathbf{v}_3 = (0, 0, 1)^T$. We therefore form the matrix P using $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ as columns, giving

$$P = \begin{pmatrix}0&1&0\\-2&1&0\\1&-5&1\end{pmatrix}.$$

It is then easy to check that

$$P^{-1}AP = \begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}.$$

Remarks.

1. In Example 3, if we had used the eigenvectors in a different order to form P, then the eigenvalues would have appeared on the diagonal of D in a different order.

2. Recall that all nonzero scalar multiples of an eigenvector are eigenvectors corresponding to the same eigenvalue (if you do not recall, this is an easy exercise that you should try). So, in Example 2, depending on how we performed the computation, we could have for instance found $\mathbf{v}_1 = (0, 1, -\frac{1}{2})$, i.e., $-1/2$ times the eigenvector we found. It can seem strange at first that using such a different vector would yield the same result. However, recall that $\mathbf{v}_1$ plays a role not only in P, but also in P^{-1}, so that whatever change happens because of the change in P is "counter-balanced" by whatever change this induces in P^{-1}.

Powers of a matrix.

An important application of diagonalisation is when computing powers of a matrix, which is something one is required to do for a variety of reasons.

Suppose we must compute the kth power of a matrix A, i.e, the product of A with itself k times. In general, computing A^k can be quite costly. (See how long it takes you to multiply two matrices together. Suppose you have to do this, say, 1,000 times.) Assume, however, that A is diagonalisable. Then there is an invertible matrix P such that

$$P^{-1}AP = D,$$

with D a diagonal matrix. Take the kth power of both sides of this equation:

$$\left(P^{-1}AP\right)^k = D^k. \tag{5}$$

Consider the left hand side of (5):

$$\begin{aligned}
\left(P^{-1}AP\right)^k &= \left(P^{-1}AP\right)\left(P^{-1}AP\right)\cdots\left(P^{-1}AP\right) \qquad [k \text{ times}] \\
&= P^{-1}APP^{-1}AP\cdots P^{-1}AP \\
&= P^{-1}A(PP^{-1})AP\cdots P^{-1}AP \\
&= P^{-1}A\cdots AP \\
&= P^{-1}A^kP.
\end{aligned}$$

The fourth equality comes from the fact that "within" the product, between any two A matrices, there is a term PP^{-1} that is present, which simplifies to $\mathbb{I}$.

Now consider D^k in (5). We know that $D = \mathsf{diag}(\lambda_1, \dots, \lambda_n)$ and it is easy to check (do it as an exercise) that it follows that

$$D^k = \mathsf{diag}(\lambda_1^k, \dots, \lambda_n^k),$$

i.e., D^k is the diagonal matrix with the kth power of the eigenvalues on the diagonal. As a consequence, (5) can be rewritten as

$$P^{-1}A^kP = \mathsf{diag}(\lambda_1^k, \dots, \lambda_n^k),$$

or, in terms of A,

$$A^k = P \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} P^{-1}.$$

Thus, instead of computing k matrix products, one just needs to compute the kth power of n numbers and multiply by two matrices.

Eigenvalues of powers of a matrix.

We can use the same type of reasoning to show that if λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k. Indeed, suppose that λ is an eigenvalue of A corresponding to the eigenvector $\mathbf{v}$, i.e., $\mathbf{v} \neq \mathbf{0}$ and $A\mathbf{v} = \lambda\mathbf{v}$. Multiply both sides by A (to the left):

$$AA\mathbf{v} = A\lambda\mathbf{v}.$$

The left hand side of this equality is $A^2\mathbf{v}$, while the right hand side takes the form

$$A\lambda\mathbf{v} = \lambda A\mathbf{v} = \lambda(A\mathbf{v}) = \lambda(\lambda\mathbf{v}) = \lambda^2\mathbf{v},$$

so that

$$A^2\mathbf{v} = \lambda^2\mathbf{v}, \tag{6}$$

in turn implying that λ^2 is an eigenvalue of A^2 associated to the same eigenvector as λ is for A. Continue the reasoning: multiplying both sides in (6) by A gives λ^3 as an eigenvalue of A^3 and so on and so forth. Thus we have proved the following result.

Theorem 3. *Let A be an $n \times n$ matrix and assume that λ is an eigenvalue of A. Then λ^k is an eigenvalue of A^k.*

4 Exercises

The following are exercises in Section 5.2 of Anton 10th edition that can be attempted with the material in this note.

1. 1–4.
2. 6.a.
3. In exercises 7–21, do the following: Determine if the matrix is diagonalisable using Theorem 2 and if so, find a matrix P that diagonalises A and compute $P^{-1}AP$.
4. 22–23, 25–27.
5. True-False exercises: a–f.