Notes on Eigenvalues and Eigenvectors

by Arunas Rudvalis

Definition 1: Given a linear transformation $\mathbf{T : R^n \to R^n}$ a **non-zero** vector $\mathbf{v}$ in $\mathbf{R^n}$ is called an **eigenvector** of $\mathbf{T}$ if $\mathbf{Tv = \lambda v}$ for some **real number** λ. The number λ is called the **eigenvalue** of $\mathbf{T}$ corresponding to $\mathbf{v}$. Given an $n \times n$ matrix $\mathbf{A}$ we know that there is a linear transformation $\mathbf{T = T_A : R^n \to R^n}$ defined by $\mathbf{T(v) = Av}$ for every vector $\mathbf{v}$ in $\mathbf{R^n}$. Consequently, **eigenvectors** and **eigenvalues** of the matrix $\mathbf{A}$ are precisely those of the linear transformation $\mathbf{T = T_A}$.

We have already seen that the main problem of linear algebra is solving systems of linear equations. The second most important problem in linear algebra is finding the eigenvectors and eigenvalues of a linear transformation $\mathbf{T : R^n \to R^n}$ or equivalently of an $n \times n$ matrix $\mathbf{A}$. As mentioned in lecture, the technique for doing this comes from consideration of determinants and kernels and we spell out the details below. We find it particularly convenient to work with the matrix $\mathbf{A}$ whose columns are the vectors $\mathbf{a}_j = \mathbf{T(e}_j\mathbf{)}$ for $j = 1, 2, 3, \dots, n$. The equation $\mathbf{Av = \lambda v}$ has a **non-zero** solution $\mathbf{v}$ $\Leftrightarrow$ the equation $\mathbf{Av = (\lambda I)v}$ has a **non-zero** solution $\mathbf{v}$ $\Leftrightarrow$ the equation $\mathbf{Av - (\lambda I)v = 0}$ has a **non-zero** solution $\mathbf{v}$ $\Leftrightarrow$ the equation $\mathbf{[A - (\lambda I)]v}$ has a **non-zero** solution $\mathbf{v}$ $\Leftrightarrow$ $\mathbf{v}$ is a **non-zero** vector in $\mathbf{Kernel([A - (\lambda I)])}$. In order for $\mathbf{Kernel([A - (\lambda I)])}$ to have **non-zero** vectors, the matrix $\mathbf{[A - (\lambda I)]}$ must be **singular** (i.e **not invertible)** and we know from our work on determinants that this happens if and only if the $\mathbf{det[A - (\lambda I)] = 0}$.

Definition 2 Given an $n \times n$ matrix $\mathbf{A}$ and a variable $\mathbf{t}$ the determinant of the matrix $\mathbf{[A - (tI)]}$ is is a polynomial of degree exactly n in the variable $\mathbf{t}$. This polynomial is called the **characteristic polynomial** of the matrix $\mathbf{A}$ and is denoted $\chi_{\mathbf{A}}\mathbf{(t)}$.

The roots (i.e. zeros) of this polynomial are the **characteristic values** (an older and now obsolete term for **eigenvalues)** of **A.**

Since we can compute $\chi_A(t)$ by our previous work on determinants, the problem of finding the **eigenvalues** of **A** becomes the problem of finding the roots of a polynomial, namely $\chi_A(t)$ of degree n. In order to get a better understanding of this situation we consider the special case n = 2, i.e. finding the **eigenvalues** and **eigenvectors** for a 2 × 2 matrix $\mathbf{A} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Since $\mathbf{A - tI} = \begin{bmatrix} a-t & c \\ b & d-t \end{bmatrix}$ it follows that $\chi_A(t) = \det[A - (tI)] = (a-t)(d-t) - bc = ad - (a+d)t + t^2 - bc = t^2 - (a+d)t + (ad - bc)$. Now **(a+d),** the negative of the coefficient of t, is the sum of the diagonal entries of **A** which is called **trace(A),** while **(ad – bc)** is the determinant of the 2 × 2 matrix **A.** In particular, for 2 × 2 matrices finding the **eigenvalues** is equivalent to finding the roots of $\chi_A(t) = \det[A - (tI)] = t^2 - (a+d) + (ad - bc)$ $= t^2 - \text{trace}(A) + \det(A)$ a polynomial of degree 2. Before looking at examples we mention the following theorem about the characteristic polynomial which we prove only in the 2 × 2 case:

Theorem 1 (Cayley-Hamilton Theorem). $\chi_A(A) = 0$.

Before giving the proof (only in the 2 × 2 case) we take a moment to explain what the above is saying: First compute $\chi_A(t)$ which we have already observed is a polynomial of degree n in the variable **t,** and then replace every occurrence of **t** by the matrix **A**, in particular replacie all powers t^k with the matrix powers A^k. The Cayley-Hamilton theorem says the n × n **matrix** $\chi_A(A)$ obtained by such substitutions must be the zero matrix. The proof in the 2 × 2 case is a simple computation where you compute $A^2 = \begin{bmatrix} (a^2 + bc) & (a+d)c \\ (a+d)b & (bc + d^2) \end{bmatrix}$, subtract $(a+d)\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ and add $(ad - bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and see the result of these computations is the 2 × 2 zero matrix. The proof in the n × n case is much more complicated as purely computational proofs such as the one above are not practical for n > 3.

Next we have some examples of **computing** eigenvalues and eigenvectors for 2 × 2 matrices:

Example 1. Suppose $\mathbf{A} = \begin{bmatrix} 6 & 8 \\ -1 & 0 \end{bmatrix}$ so $\mathbf{trace(A) = 6 + 0 = 6}$ and $\mathbf{det(A) = 0 - (-1)(8) = 8}$ so $\chi_{\mathbf{A}}\mathbf{(t)} =$

$\mathbf{t^2 - 6t + 8 = (t - 2)(t - 4)},$ so it follows the eigenvalues are $\boldsymbol{\lambda_1 = 2}$ and $\boldsymbol{\lambda_2 = 4}$. Next we compute the corresponding eigenvectors: For $\boldsymbol{\lambda_1 = 2}$ this means we must solve the system of linear equations:

$$\begin{bmatrix} (6-2) & 8 \\ -1 & (0-2) \end{bmatrix} \begin{matrix} x \\ y \end{matrix} = \begin{matrix} 0 \\ 0 \end{matrix} \Rightarrow \begin{matrix} 4x + 8y = 0 \\ -1x + -2y = 0 \end{matrix} \quad \text{so} \quad \begin{matrix} x = -2y \\ y = \ \ \ y \end{matrix} = \begin{matrix} -2 \\ 1 \end{matrix} \ y, \ \text{so non-zero multiples of } \mathbf{v}_1 = \begin{matrix} -2 \\ 1 \end{matrix}$$

are **all** the eigenvectors for $\boldsymbol{\lambda_1 = 2}$. For $\boldsymbol{\lambda_2 = 4}$, the system of linear equations we must solve is:

$$\begin{bmatrix} (6-4) & 8 \\ -1 & (0-4) \end{bmatrix} \begin{matrix} x \\ y \end{matrix} = \begin{matrix} 0 \\ 0 \end{matrix} \Rightarrow \begin{matrix} 2x + 8y = 0 \\ -1x + -4y = 0 \end{matrix} \quad \text{so} \quad \begin{matrix} x = -4y \\ y = \ \ \ y \end{matrix} = \begin{matrix} -4 \\ 1 \end{matrix} \ y, \ \text{so non-zero multiples of } \mathbf{v}_1 = \begin{matrix} -4 \\ 1 \end{matrix}$$

are all the eigenvectors for $\boldsymbol{\lambda_2 = 4}$. In particular there are two linearly independent eigenvectors for **A.**

Example 2. Suppose $\mathbf{B} = \begin{bmatrix} 6 & 9 \\ -1 & 0 \end{bmatrix}$ so $\mathbf{trace(B) = 6 + 0 = 6}$ and $\mathbf{det(B) = 0 - (-1)(9) = 9}$ so $\chi_{\mathbf{B}}\mathbf{(t)} =$ $\mathbf{t^2 - 6t + 9 = (t - 3)(t - 3)},$ so it follows the eigenvalues are $\boldsymbol{\lambda_1 = 3}$ and $\boldsymbol{\lambda_2 = 3}$. Next we compute the corresponding eigenvectors: For $\boldsymbol{\lambda_1 = 3}$ this means we must solve the system of linear equations:

$$\begin{bmatrix} (6-3) & 9 \\ -1 & (0-3) \end{bmatrix} \begin{matrix} x \\ y \end{matrix} = \begin{matrix} 0 \\ 0 \end{matrix} \Rightarrow \begin{matrix} 3x + 9y = 0 \\ -1x + -3y = 0 \end{matrix} \quad \text{so} \quad \begin{matrix} x = -3y \\ y = \ \ \ y \end{matrix} = \begin{matrix} -3 \\ 1 \end{matrix} \ y, \ \text{so non-zero multiples of } \begin{matrix} -3 \\ 1 \end{matrix}$$

are all the eigenvectors for $\boldsymbol{\lambda_1 = 3}$. Since the second eigenvalue $\boldsymbol{\lambda_2 = 3}$ is the same as the first we do not get any more eigenvalues from $\boldsymbol{\lambda_2}$. In particular, even though we are working in $\mathbf{R^2}$ we **do not** find two linearly independent eigenvalues for the matrix **B.**

The following is a more extreme example where there are no real number eigenvlaues nor eigenvectors:

Example 3. Suppose $\mathbf{C} = \begin{bmatrix} 6 & 10 \\ -1 & 0 \end{bmatrix}$ so $\mathbf{trace(C) = 6 + 0 = 6}$ and $\mathbf{det(C) = 0 - (-1)(10) = 10}$ so $\chi_{\mathbf{C}}\mathbf{(t) = t^2 - 6t + 10 = (t - 3)^2 + 1},$ so it follows the eigenvalues are complex conjuagates $\boldsymbol{\lambda_1 = 3 + i}$ and $\boldsymbol{\lambda_2 = 3 - i}$. Since these are complex numbers, this matrix has no real eigenvalues and thus no real eigenvectors. It does have complex eigenvectors, but that is a story for another course in linear algebra for which complex numbers are a prerequisite unlike our course here.

These three examples illustrate all the possibilities for the eigenvalues and eigenvectors for 2×2 matrices as the roots of a polynomial of degree 2 can only be of one of three types:

1) Two distinct real roots, as in example 1 above. 2) One (repeated) real root, as in example 2 above.

3) Two complex conjugate complex roots, as in example 3 above.

Next we consider the problem of recovering (i.e. reconstructing/computing) the standard basis matrix $\mathbf{A}$ of a linear transformation $\mathbf{T : R^n \to R^n}$ from a basis of $\mathbf{R}^n$ consisting of eigenvectors of $\mathbf{T}$ and their corresponding eigenvalues. Although I am not claiming that it is **obvious** that such a reconstruction is possible, I will first try to persuade you that such a reconstruction is **plausible** in the case $n = 2$, then we work out a specific example and finally I will generalize this to give a procedure for actually reconstructing the matrix $\mathbf{A}$ whatever the value of n.

Suppose $\mathbf{T : R^2 \to R^2}$ is a linear transformation and that $\mathbf{s}_1$ and $\mathbf{s}_2$ are two linearly independent eigenvectors of $\mathbf{T}$ with respective eigenvalues $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ and we want to use these eigenvectors and eigenvalues to compute the matrix $\mathbf{A} = \begin{bmatrix} \mathbf{a} & \mathbf{c} \\ \mathbf{b} & \mathbf{d} \end{bmatrix}$ of $\mathbf{T}$ with respect to the standard basis of $\mathbf{R}^2$.

What we **know** is $\mathbf{As}_1 = \boldsymbol{\lambda}_1\mathbf{s}_1$ and $\mathbf{As}_2 = \boldsymbol{\lambda}_2\mathbf{s}_2$ where the entries **a, b, c,** and **d** of $\mathbf{A}$ are **unknowns** while the vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ and the scalars $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are all **given**. The first equation is a linear system of two equations in four unknowns while the second equation is also a linear system of two equations in four unknowns so we have a total of four linear equations in four unknowns so it is **plausible** that we can solve these four equations in four unknowns in order to recover the matrix **A.** Next we do an example:

Example 4. We try example 1 above where we computed the eigenvalues and eigenvectors from the matrix $\mathbf{A}$ to see if we can recover (i.e. reconstruct) $\mathbf{A}$ from the eigenvectors and corresponding eigenvalues. We found $\mathbf{s}_1 = \begin{matrix} \mathbf{-2} \\ \mathbf{1} \end{matrix}$ with $\boldsymbol{\lambda}_1 = \mathbf{2}$ and $\mathbf{s}_2 = \begin{matrix} \mathbf{-4} \\ \mathbf{1} \end{matrix}$ with $\boldsymbol{\lambda}_2 = \mathbf{4}$. This gives the pairs of equations

$$\begin{matrix} \mathbf{-2a + c = -4} \\ \mathbf{-2b + d = \ \ 2} \end{matrix} \quad \text{and} \quad \begin{matrix} \mathbf{-4a + c = -16} \\ \mathbf{-4b + d = \ \ 4} \end{matrix}$$

where the two equations on the left are $\mathbf{As}_1 = \boldsymbol{\lambda}_1\mathbf{s}_1$ while the two equations on the right are $\mathbf{As}_2 = \lambda_2\mathbf{s}_2$. Subtracting the top equation on the right from the top equation on the left gives $\mathbf{2a = 12}$ so $\mathbf{a = 6}$ and thus $\mathbf{c = 8}$ while subtracting the bottom equation on the right from the bottom equation on the left gives $\mathbf{2b = -2}$ so $\mathbf{b = -1}$ and thus $\mathbf{d = 0}$, so we have recovered $\mathbf{A}$.

Next we generalize this example to arbitrary n. To this end suppose $\mathbf{T : R^n \rightarrow R^n}$ is a linear transformation and that $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots, \mathbf{s}_n$ is a set of n linearly independent eigenvectors of $\mathbf{T}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$. As before, $\mathbf{A}$ is the $n \times n$ matrix of $\mathbf{T}$ with respect to the standard basis of $\mathbf{R}^n$, so each of the vector equations $\mathbf{As}_j = \lambda_j \mathbf{s}_j$ $j = 1, 2, 3, \dots, n$ provides us with a system of n equations in n^2 unknowns (namely the coefficients of the unknown matrix $\mathbf{A}$). It turns out to be convenient to organize these vector equations into a matrix equation as follows:

Let $\mathbf{S}$ be the $n \times n$ matrix whose columns are the eigenvectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots, \mathbf{s}_n$ so that the above systems of vector equations can be combined into one matrix equation: $\mathbf{A\,S} = [\,\lambda_1\mathbf{s}_1, \lambda_2\mathbf{s}_2, \dots, \lambda_n\mathbf{s}_n\,]$, where $\lambda_j\mathbf{s}_j$ is the column vector obtained by multiplying the eigenvector $\mathbf{s}_j$ by the eigenvalue λ_j. It is easy to see (for example calculation in the 2×2 case) that the matrix $[\,\lambda_1\mathbf{s}_1, \lambda_2\mathbf{s}_2, \dots, \lambda_n\mathbf{s}_n\,]$ can be written as $\mathbf{S\,\Lambda}$ where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal entries are the corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ (appearing in the same order as the eigenvectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots, \mathbf{s}_n$ appear in $\mathbf{S}$).

The advantage of writing the system of n^2 equations in n^2 unknowns as $\mathbf{A\,S = S\,\Lambda}$ becomes obvious if we recall that any matrix (in particular $\mathbf{S}$) with linearly independent columns has an inverse so multiplying this equation (on the right) by $\mathbf{S}^{-1}$ gives $\mathbf{A\,S\,S^{-1} = A = S\,\Lambda\,S^{-1}}$ giving us an explicit formula for $\mathbf{A}$ in terms of the eigenvectors of $\mathbf{T}$ (which are the columns of $\mathbf{S}$) and the corresponding eigenvalues (which are the diagonal entries of the matrix $\mathbf{\Lambda}$). Thus we can recover the matrix $\mathbf{A}$ of $\mathbf{T}$ whenever we are given a basis of $\mathbf{R}^n$ consisting of n linearly independent eigenvectors of $\mathbf{T}$ together with the corresponding eigenvalues.

If we rearrange the equation $\mathbf{A = S\,\Lambda\,S^{-1}}$ as $\mathbf{S^{-1}A\,S = \Lambda}$, we see that this means that **whenever** there is a basis of $\mathbf{R}^n$ consisting of eigenvectors of $\mathbf{A}$, then $\mathbf{A}$ is **similar** to a diagonal matrix (specifically $\mathbf{\Lambda}$) with the matrix $\mathbf{S}$ of eigenvectors of $\mathbf{A}$ providing the similarity.

Remark: Similarity of matrices [see p. 145 of the textbook] is much stronger than the similarity of triangles you learned in geometry, where two triangles are similar if they have the same shape even if they have different sizes. Similarity of matrices is much more like congruence of triangles where they must have the same shape and the same size, as similar matrices can be thought of as matrices of the same linear transformation with respect to different bases of $\mathbf{R}^n$ as is carefully explained on pages 143-146 of the textbook. In particular we have proved the following result:

Theorem 2 Diagonalization Theorem Suppose $\mathbf{T} : \mathbf{R}^n \to \mathbf{R}^n$ is a linear transformation for which $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots, \mathbf{s}_n$ is a set of n linearly independent eigenvectors for $\mathbf{T}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$. If we let $\mathbf{S}$ be the $n \times n$ matrix whose columns are the above eigenvectors and let $\mathbf{\Lambda}$ be the diagonal matrix whose diagonal entries are the corresponding eigenvalues, then the matrix $\mathbf{A}$ of $\mathbf{T}$ with respect to the standard basis of $\mathbf{R}^n$ is given by $\mathbf{A} = \mathbf{S}\,\mathbf{\Lambda}\,\mathbf{S}^{-1}$ where $\mathbf{S}^{-1}$ is the inverse of $\mathbf{S}$. Furthermore, any square matrix $\mathbf{A}$ is diagonalizable (i.e.similar to a diagonal matrix) if and only if there is a basis of $\mathbf{R}^n$ consisting of eigenvectors of $\mathbf{A}$.

For diagonalizable matrices it is easy to compute integer powers as well as more complicated functions (as we will see later) and we describe this now. In general, given a square matrix $\mathbf{A}$ it is difficult to compute powers $\mathbf{A}^k$ of $\mathbf{A}$, (not conceptually, but simply because it involves a lot of calculation). For example to compute $\mathbf{A}^k$ requires us to perform k matrix multiplications so if k is large this is a lot of computation. The number of computations can be reduced drastically if we observe that $\mathbf{A}^k = (\mathbf{S}\,\mathbf{\Lambda}\,\mathbf{S}^{-1})^k = \mathbf{S}\,\mathbf{\Lambda}^k\,\mathbf{S}^{-1}$ so the required computations are first computie the kth powers of the diagonal entries of $\mathbf{\Lambda}$, and then only two matrix multiplications, one on the left by $\mathbf{S}$ and one on the right by $\mathbf{S}^{-1}$ (although to be completely honest we also have to compute $\mathbf{S}^{-1}$ which is about as much work as a matrix multiplication) so computing $\mathbf{A}^k$ is no more expensive than doing three matrix multiplications (no matter how large the exponent k). For large values of k this is a huge savings and also we have an explicit formula for $\mathbf{A}^k$ which is useful in theoretical work as well as computations.

We just explained how we can practically compute large powers of any matrix that can be diagonalized so next we look at cases where **A** cannot be diagonalized. Although we will explain (in principle) how to do this in general later, we now concentrate on cases where our matrices are 2×2 as in such cases we give explicit formulas/procedures (like that above for diagonalizable matrices) for doing such calculations and along the way we will have useful comments on performing such calculations for $n \times n$ matrices that cannot be diagonalized.

Matrices cannot be diagonalized if there are **at most** k linearly independent eigenvectors where $k < n$. In the 2×2 case we saw in examples 2 and 3 that this can happen for only two reasons: In example 2 we had a repeated real eigenvalue and there was only one eigenvector (up to scalar multiples). In example 3 we had no real number eigenvalues and therefore no real number eigenvectors. A more precise statement of what happened in example 3 is that the eigenvalues and thus the eigenvectors were complex conjugates so there are no **real number** eigenvalues and no **real number** eigenvectors. Although the formula $\mathbf{A} = \mathbf{S}\,\mathbf{\Lambda}\,\mathbf{S}^{-1}$ above still works if we allow the eigenvalues and eigenvectors to be complex (and therefore provides us with a procedure for computing powers of **A**) it seems somehow inappropriate to have to use complex numbers, complex eigenvectors and complex matrices to perform computations of powers of real number matrices. Rather surprisingly, there is a procedure for practically computing powers (and even more complicated functions) of 2×2 real matrices whose eigenvalues happen to be complex conjugates and this procedure does not require any knowledge of complex numbers. The heart of this procedure is to "separate" such matrices into their "real" and "imaginary" parts with both parts being **real number** matrices and we illustrate this procedure with the following example which is very much like example 3 above but a little bit more complicated:

Example 5. Suppose $\mathbf{C} = \begin{bmatrix} 6 & 13 \\ -1 & 0 \end{bmatrix}$ so $\mathbf{trace(C) = 6 + 0 = 6}$ and $\mathbf{det(C) = 0 - (-1)(13) = 13}$ so

$\chi_C(t) = t^2 - 6t + 10 = (t-3)^2 + 4$, and it follows the eigenvalues are complex conjuagates $\lambda_1 = 3 + 2i$

and $\lambda_2 = 3 - 2i$. If we subtract $3I$ from the matrix C, we obtain the matrix $K = \begin{bmatrix} 3 & 13 \\ -1 & -3 \end{bmatrix}$ whose

eigenvalues are $2i$ and $-2i$, so the matrix $J = (1/2)K$ has eigenvalues i and $-i$. If we compute J^2

$$(1/2)\begin{bmatrix} 3 & 13 \\ -1 & -3 \end{bmatrix} (1/2)\begin{bmatrix} 3 & 13 \\ -1 & -3 \end{bmatrix} = (1/4)\begin{bmatrix} (9-13) & (39-39) \\ (-3+3) & (-13+9) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I,$$ we see that the matrix J is

like a **real number** matrix version of the complex number i. In particular, the original matrix $C = 3I + 2J$ has been "separated" into its "real" part $3I$ and its "imaginary" part $2J$.

To compute powers of C we simply compute powers of $[3I + 2J]$ using the binomial theorem in conjunction with the fact that all powers of I are equal to I while the powers of J are: $J^2 = -I$, $J^3 = -J$, and $J^4 = I$ with all higher powers of J computed from repetitions of this sequence of powers of J. For example, $C^2 = [3I + 2J]^2 = (3I)^2 + 2(3I)(2J) + (2J)^2 = 9I + 12J + 4(-I) = 5I + 12J$ and a similar calculation gives $C^3 = -9I + 46J$. In particular, we can compute large powers of C without having to do any matrix multiplications at all as all bookkeeping is done using the binomial theorem in conjunction with our above formulas for the powers of the matrix J.

This approach can also be used to compute more complicated functions of the matrix C. For instance, if we let t be a variable then the matrix e^{Ct} can be computed as $e^{[3I + 2J]t} = e^{3It + 2Jt} = e^{3It} e^{2Jt} = (e^{3t})I\, e^{2Jt}$ $= (e^{3t})e^{2Jt}$, where e^{3t} is the ordinary real function of t and the **matrix** e^{2Jt} can be computed using the power series expansion of $e^x = 1 + x + x^2/2! + x^3/3! + \ldots + x^k/k! + \ldots$ We replace the variable x by the matrix $(2t)J$ and make use of the formulas for the powers of J to see that $e^{2Jt} = \cos(2t)I + \sin(2t)J$ which you can think of as a **real number matrix** version of Euler's formula $e^{ix} = \cos(x) + i\sin(x)$ with $x = 2t$ and the complex number i replaced by the **real** matrix J. Thus $e^{Ct} = (e^{3t})\,[\cos(2t)I + \sin(2t)J]$.

The details of this computation are as follows:

$\mathbf{e^{2Jt}} = \mathbf{I + (2tJ) + (2tJ)^2/2! + (2tJ)^3/3!} + \ldots + \mathbf{(2tJ)^k/k!} + \ldots$

$= \mathbf{I + (2t)J + [(2t)^2/2!]J^2} + \mathbf{[(2t)^3/3!]J^3} + \ldots + \mathbf{[(2t)^k/k!]J^k} + \ldots$ separate into even and odd powers

$= \mathbf{I} + \mathbf{[(2t)^2/2!]J^2} + \mathbf{[(2t)^4/4!]J^4} + \ldots + \mathbf{[(2t)^{2m}/(2m)!]J^{2m}} + \ldots$ use $\mathbf{J^2 = -I}$ and $\mathbf{J^4 = I}$ here

$+ \mathbf{(2t)J + [(2t)^3/3!]J^3} + \ldots + \mathbf{[(2t)^{2m+1}/(2m+1)!]J^{2m+1}} + \ldots$ use $\mathbf{J^3 = -J}$ here to find

$= \mathbf{I - [(2t)^2/2!]\,I + [(2t)^4/4!]\,I - [(2t)^6/6!]\,I + \ldots + (-1)^k\,[(2t)^{2k}/(2k)!]\,I + \ldots}$ factor out $\mathbf{I}$

$\mathbf{(2t)J - [(2t)^3/3!]J + [(2t)^5/5!]J - \ldots + (-1)^k[(2t)^{2k+1}/(2k+1)!]J + \ldots}$ factor out $\mathbf{J}$

$= \mathbf{(1 - [(2t)^2/2!] + [(2t)^4/4!] - [(2t)^6/6!] + \ldots + (-1)^k\,[(2t)^{2k}/(2k)!] + \ldots)\,I}$ series for $\mathbf{\cos(2t)\,I}$

$\mathbf{((2t) - [(2t)^3/3!] + [(2t)^5/5!] - \ldots + (-1)^k[(2t)^{2k+1}/(2k+1)!] + \ldots)\,J}$ series for $\mathbf{\sin(2t)\,J}$

$= \mathbf{\cos(2t)I + \sin(2t)J}$ as claimed**.**

This example generalizes to any 2×2 matrix **A** whose eigenvalues are complex conjugates $\mathbf{a + bi}$ and $\mathbf{a - bi}$ as we write **A** in the form $\mathbf{A = aI + bJ}$ where $\mathbf{J = [A - aI]/b}$ is a **real number** matrix such that $\mathbf{J^2 = -I}$ and the above argument shows that $\mathbf{e^{At} = (e^{at})\,[\cos(bt)I + \sin(bt)J]}$ where $\mathbf{J = [A - aI]/b}$.

We will return to matrices of the form $\mathbf{e^{At}}$ later when we discuss one of the main applications of eigenvalues and eigenvectors, namely the solution of systems of first order linear differential equations with constant coefficients.

Next we look at cases that are generalizations of Example 2 and for this we need to introduce:

Definition 3. A linear transformation $\mathbf{T : R^n \to R^n}$ is **nilpotent** if some positive integer power of **T** is the zero transformation **0** (i.e. the transformation that sends every vector in $\mathbf{R^n}$ to the zero vector). The smallest positive exponent k such that $\mathbf{T^k = 0}$ is called the **index of nilpotence** of **T.** The same terminology is used if the linear transformation **T** is replaced by an n × n matrix **A.**

For example, if **B** is the matrix in example 3 above, then

$$\mathbf{N = B - 3I = \begin{bmatrix} (6-3) & 9 \\ -1 & (0-3) \end{bmatrix} = \begin{bmatrix} 3 & 9 \\ -1 & -3 \end{bmatrix}}$$

is nilpotent with index of nilpotence 2 because $\mathbf{N^2 = 0.}$ If we write $\mathbf{B = 3I + N}$ then we have separated **B** into its "real" part **3I** and its "nilpotent" part $\mathbf{N} = \mathbf{B} - \mathbf{3I}$ and much the same ideas that were used above for matrices with complex conjugate eigenvalues can be used to compute powers of **B.** Specifically, $\mathbf{B}^k = [\mathbf{3I} + \mathbf{N}]^k = [\mathbf{3I}]^{\mathbf{k}} + k[\mathbf{3I}]^{k-1}\mathbf{N} + \ldots$ where all remaining terms from the binomial theorem involve $\mathbf{N}^k$ for $k \geq 2$ and since $\mathbf{N}^2 = \mathbf{0}$ all such terms are **0,** so $\mathbf{B}^k = \mathbf{3}^{k-1}[\mathbf{3I} + k\mathbf{N}]$.

The computation of $\mathbf{e}^{\mathbf{Bt}}$ is also easy in this case: $\mathbf{e}^{\mathbf{Bt}} = \mathbf{e}^{[\mathbf{3I+N}]\mathbf{t}} = \mathbf{e}^{\mathbf{3It}}\mathbf{e}^{\mathbf{Nt}}$ and if we compute $\mathbf{e}^{\mathbf{Nt}}$ using the power series for the exponential function $\mathbf{e}^{\mathbf{Nt}} = \mathbf{I} + (\mathbf{Nt}) + (\mathbf{Nt})^2/2! + \ldots$ all the terms after the first two in the power series are **0** because $\mathbf{N}^2 = \mathbf{0}$, so $\mathbf{e}^{\mathbf{Nt}} = \mathbf{I} + \mathbf{Nt}$ therefore $\mathbf{e}^{\mathbf{Bt}} = \mathbf{e}^{\mathbf{3t}}[\mathbf{I} + \mathbf{Nt}]$.

All of this generalizes to any 2×2 matrix **B** having a repeated eigenvalue λ and only one eigenvector (up to scalar multiples) as $\mathbf{B}^k = (\lambda)^{k-1}[\lambda\mathbf{I} + k\mathbf{N}]$ where $\mathbf{N} = \mathbf{B} - \lambda\mathbf{I}$ is the "nilpotent" part of **B** and similarly $\mathbf{e}^{\mathbf{Bt}} = \exp(\lambda t)\,[\mathbf{I} + \mathbf{Nt}]$.

Before making some comments about $n \times n$ matrices that **cannot** be diagonalized we summarize our results from all the 2×2 cases:

Case 1 **A** has two linearly independent eigenvectors $\mathbf{s}_1$ and $\mathbf{s}_2$ with corresponding eigenvalues λ_1 and λ_2. **Remark:** It is possible that $\lambda_1 = \lambda_2 = \lambda$, but if so then $\mathbf{A} = \lambda\mathbf{I}$ is a scalar multiple of **I.** Let **S** be the matrix whose columns are the eigenvectors $\mathbf{s}_1$ and $\mathbf{s}_2$ and let Λ be the diagonal matrix whose diagonal entries are the corresponding eigenvalues λ_1 and λ_2. Then $\mathbf{A}^k = \mathbf{S}\,\Lambda^k\,\mathbf{S}^{-1}$ where Λ^k is the diagonal matrix whose diagonal entries are the k-th powers of the diagonal entries of Λ and $\mathbf{S}^{-1}$ is the inverse of S. Furthermore, $\mathbf{e}^{\mathbf{At}} = \mathbf{S}\,\mathbf{e}^{\Lambda\mathbf{t}}\,\mathbf{S}^{-1}$ where $\mathbf{e}^{\Lambda\mathbf{t}}$ is the diagonal matrix whose diagonal entries are the exponential functions $\exp(\lambda_1 t)$ and $\exp(\lambda_2 t)$.

Case 2 **A** has repeated eigenvalues λ and λ and only one eigenvector (up to scalar multiples). We split **A** into its "real" and "nilpotent" parts $\mathbf{A} = \lambda\mathbf{I} + \mathbf{N}$, where $\mathbf{N} = \mathbf{A} - \lambda\mathbf{I}$. Then $\mathbf{A}^k = (\lambda)^{k-1}[\lambda\mathbf{I} + k\mathbf{N}]$. Furthermore, $\mathbf{e}^{\mathbf{At}} = \exp(\lambda t)\,[\mathbf{I} + \mathbf{Nt}]$.

Case 3 **A** has a pair of complex conjugate eigenvalues $\mathbf{a + bi}$ and $\mathbf{a - bi}$. We split **A** into its "real" and "imaginary" parts $\mathbf{A = aI + bJ}$, where $\mathbf{J = [A - aI]/b}$ is a matrix such that $\mathbf{J^2 = -I}$. Then $\mathbf{A^k} = \mathbf{[aI + bJ]^k}$ is computed using the binomial formula and we do not have to compute any matrix products or powers because $\mathbf{I^j = I}$ for every exponent j while the powers of **J** are computed using $\mathbf{J}^2 = -I$ so $\mathbf{J}^3 = -J$ and $\mathbf{J}^4 = I$, etc. Furthermore, $\mathbf{e^{At}} = \mathbf{(e^{at})\,[\cos(bt)I + \sin(bt)J]}$.

The case of $n \times n$ matrices that cannot be diagonalized is considerably more complicated as there can be several eigenvalues that are repeated and others that come in complex conjugate pairs in the same problem so I limit the discussion here to the cases where the matrix **A** has only one eigenvalue $\boldsymbol{\lambda}$ that is repeated n times as this is a very important special case that needs to be looked at if you want to look at arbitrary matrices that cannot be diagonalized. As before we separate **A** into its "real" part $\boldsymbol{\lambda}\mathbf{I}$ and its "nilpotent" part $\mathbf{N = A - \lambda I}$. In the case $n = 2$ the index of nilpotence for any non-zero matrix must be 2, but in the $n \times n$ case this index is only guaranteed to be at most n. If the index is n then there will be only one eigenvector (up to scalar multiples) but if the index is smaller there will be several independent eigenvectors and as many as (n-1) of them if the index is only 2. The details are quite messy and we will not go into them here. In the extreme case where the index of nilpotence is n, in order to compute powers of $\mathbf{A = \lambda I + N}$ we must first compute all the non-zero powers of **N** i.e. up to $\mathbf{N}^{n-1}$ because all these powers will be non-zero and will appear in the expansion of $\mathbf{[\lambda I + N]}^k$ by the binomial theorem. Similarly, if you want to compute $\mathbf{e^{Nt}} = \mathbf{I + (Nt)/1! + (Nt)}^2/2! + \ldots + \mathbf{(Nt)}^k/k! + \ldots$ then the first n terms (through $k = n-1$) will be non-zero. For example if $n = 4$, then $\mathbf{e^{Nt}} = \mathbf{I + Nt + N^2t^2}/2! + \mathbf{N^3t^3}/3!$ rather than the much simpler formula $\mathbf{e^{Nt} = I + Nt}$ when $n = 2$.