Eigenvalues and Eigenvectors—Why are they important?

R (Chandra) Chandrasekhar

2015-12-13 | 2020-12-01

Stimulating interest in an arcane topic

A university academic friend of mine recently remarked that it was not easy to motivate students to study eigenvalues and eigenvectors, let alone appreciate their importance: the subject itself was abstract, and the applications tended to be domain-specific and somewhat arcane.

A cursory Web search turned up results that confirmed his assertions and concerns. There have been pro and con views about motivating students to learn about eigenvalues and eigenvectors, and especially to convey intuitively their importance.

I then asked, "Can I explain to *myself* what eigenvalues and eigenvectors are, and why they are important?". It also occurred to me that the harried and hurried students of today might derive some benefit from my efforts; hence this blog. It is a brief, largely qualitative, and mathematically non-rigorous article on eigenvalues and eigenvectors that aims to provide meaning and motivation for their study. Corrections and suggestions for improvement are most welcome. ☺

Eigenvalues and eigenvectors

As a general rule, the more powerful an idea, the more prevalent it becomes. Think about words and numbers, and you will see what I mean.

Eigenvalues and eigenvectors are one such powerful idea. It is no surprise that they appear in different guises in different contexts: in oscillating electronic circuits, in dynamical systems, in computer games, in the spectra of atoms, and in Google searches, to name just a few.

The word *eigen* is German in origin and means "inherent, characteristic, natural, own, or peculiar (to)". So the prefix "eigen" captures the natural essence of the noun it qualifies. Perhaps the word "idiosyncratic" comes closest to conveying its import.

Matrices

Eigenvalues and eigenvectors are associated traditionally with *matrices*. If numbers are like tea-leaves, matrices are like tea-bags. They are rectangular arrays of numbers, whether real or complex, that have been hijacked by mathematicians to serve as a shorthand in a variety of contexts. What they mean depends on context and level of abstraction. They can represent geometric transformations in Euclidean space, or systems of linear equations, or systems of linear differential equations with constant coefficients, or linear transformations in vector spaces. Note the recurrence of the word *linear* here.

Invariance and identity elements

Invariance is a central concept in mathematics and physics. Adding zero to a number leaves it unchanged. Multiplying a number by one again leaves it unchanged. And zero and one are important numbers, usually called the *additive* and *multiplicative identity elements* respectively. Consider now the matrix equivalent of multiplying by 1, an example of which is:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 1 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{1}$$

The 2×2 matrix on the extreme left of Equation (1) is the *identity matrix* of dimension 2, analogous to the multiplicative identity. We could write this equation more succinctly as:

$$I\boldsymbol{v} = 1\boldsymbol{v} \tag{2}$$

I, on the left, is the *identity matrix*, the number 1 on the right is called an *eigenvalue* and the vector $\boldsymbol{v}$ is called an *eigenvector*. Note that there are no strictures on $\boldsymbol{v}$. So, in this particular case, *all* vectors $\boldsymbol{v}$ are eigenvectors but there is only *one* eigenvalue, namely 1. This example, however, is both unusual and contrived, because the identity matrix is a *special* type of *square matrix* with ones on its principal diagonal and zeros elsewhere.

Equation (2) is a particular case of the general equation for eigenvalues and eigenvectors, which is written:

$$M\boldsymbol{v} = \lambda\boldsymbol{v} \tag{3}$$

where *M* is a *general square matrix*, λ is a *real or complex scalar* called an *eigenvalue*, and $\boldsymbol{v}$ is a *non-zero vector* called an *eigenvector*. The matrix *M* is assigned meaning according to context. Geometrically, the *orientation* of the vector $\boldsymbol{v}$ is unchanged by the transformation *M*, although if λ is negative, the direction is reversed. Specifically, the eigenvector corresponding to an eigenvalue of 1 remains unchanged by the transformation—an example of invariance.

Calculus

The operation of taking a derivative may be denoted by the *differential operator, D*. We know that

$$\frac{d}{dt}e^t = D(e^t) = e^t$$

and further that

$$\frac{d}{dt}e^{st} = D(e^{st}) = se^{st}$$

where *s* is a scalar and *t* is the independent variable, usually time. Although *D* is not a matrix here, it is nevertheless a linear transformation that operates on *functions*. And in place of a vector, we have a function, e^{st}, which is therefore called an *eigenfunction* of the differential operator, with eigenvalue *s*. The importance of the complex exponentials in signal and system analysis cannot be over-emphasized: just recall the Laplace and Fourier transforms.

Differential Equations

Linear homogeneous *differential equations* with constant coefficients may be written using the *D* notation already introduced. A second order homogeneous equation with independent variable *t* and dependent variable *y* may be written as $a_2D^2(y) + a_1D(y) + a_0(y) = 0$. Plugging in a solution of the form $y = e^{st}$, we get $(a_2s^2 + a_1s + a_0)e^{st} = 0$. Since e^{st} can never be zero, we may divide by it to get the *characteristic polynomial*

$$a_2s^2 + a_1s + a_0 = 0. \tag{4}$$

The *roots* of this characteristic polynomial give us the eigenvalues of the system. Perhaps, the prefix "eigen" came to be used because of the adjective "characteristic".

These ideas masquerade under different terminology in linear system and control theory where transfer function, poles and zeros, natural frequency and resonance, and stability are encountered.

Characteristic polynomial of a square matrix

The characteristic polynomial of a square matrix is obtained likewise. The equation $M\boldsymbol{v} = \lambda\boldsymbol{v}$ may be re-written as $(M - \lambda I)\boldsymbol{v} = \boldsymbol{0}$, where the right hand side is the *zero vector*. Since the eigenvector $\boldsymbol{v}$ is non-zero, this implies that the matrix $(M - \lambda I)$ is singular or non-invertible, which in turn implies that its determinant is zero. So, the characteristic polynomial is the equation

$$\det(M - \lambda I) = 0 \tag{5}$$

and its roots are the eigenvalues of *M*. The determinant of a square matrix is a *number* associated with it, obtained by adding and subtracting products of its elements in a specific order.

Linear transformations and vector spaces

A vector space is a powerful mathematical abstraction that allows us to unify many disparate branches of mathematics under a uniform taxonomy. Linear transformations are a particular type of mapping between two vector spaces over a scalar field, satisfying:

$$T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$$
$$T(\alpha\boldsymbol{u}) = \alpha T(\boldsymbol{u})$$

(6)

where T is the transformation, $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors, and α is a scalar.

T may be represented by a matrix, and under certain conditions, its eigenvalues and eigenvectors can characterize the transformation *completely*. This happens when (a) all the eigenvectors are *linearly independent,* i.e., no two eigenvectors are parallel, and (b) when they *span* the vector space, i.e., any vector within the space can be constructed from a linear combination of the eigenvectors. The eigenvectors are then said to form a *basis* for the space.

As a case in point, let us say T is a 3×3 matrix whose eigenvectors $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, and $\boldsymbol{e_3}$ are linearly independent and form a basis. Then, if $\boldsymbol{v} = \alpha_1\boldsymbol{e_1} + \alpha_2\boldsymbol{e_2} + \alpha_3\boldsymbol{e_3}$, where the α_is are scalars, by virtue of the fact that T is a linear transformation, we have

$$\begin{aligned} T(\boldsymbol{v}) &= T(\alpha_1\boldsymbol{e_1} + \alpha_2\boldsymbol{e_2} + \alpha_3\boldsymbol{e_3}) \\ &= T(\alpha_1\boldsymbol{e_1}) + T(\alpha_2\boldsymbol{e_2}) + T(\alpha_3\boldsymbol{e_3}) \\ &= \alpha_1 T(\boldsymbol{e_1}) + \alpha_2 T(\boldsymbol{e_2}) + \alpha_3 T(\boldsymbol{e_3}) \\ &= \alpha_1\lambda_1\boldsymbol{e_1} + \alpha_2\lambda_2\boldsymbol{e_2} + \alpha_3\lambda_3\boldsymbol{e_3} \end{aligned}$$

Notice how the right hand side is now expressed *purely as a sum of scaled eigenvectors.* This is the essence of why eigenvalues and eigenvectors are so important: they are *sufficient* to describe what is taking place. *Eigenvalues and eigenvectors encode the transformation succinctly, just as DNA encodes biological information.*

If, in addition, T is *symmetric*, its eigenvectors form an *orthonormal basis*. Such basis vectors confer *parsimony* (think low bandwidth) when images or audio signals need to be deconstructed for transmission and reconstructed on reception. Eigenvectors are also useful in techniques like *principal component analysis* which is used in statistical pattern recognition.

The applications of eigenvalues and eigenvectors in linear algebra run far and deep. Suffice it here to merely mention that an extension, fortuitously called *spectral theory,* even explains the observed spectra of atoms in quantum theory!

A property of eigenvectors

I will here belabour a point that might seem blindingly obvious to some but frustratingly obscure to others. Let $\boldsymbol{v}$ be an eigenvector associated with a distinct eigenvalue λ as in Equation (3), and let k be a non-zero scalar. Then, using the second of the equation-pair Equation (6), we have

$$M(k\boldsymbol{v}) = k(M\boldsymbol{v}) = k(\lambda\boldsymbol{v}) = \lambda(k\boldsymbol{v}),$$

(7)

which means that if $\boldsymbol{v}$ is an eigenvector, any non-zero scalar multiple of $\boldsymbol{v}$ is also an eigenvector for that same eigenvalue. So, strictly speaking, we really should be referring to *an* eigenvector—rather than *the* eigenvctor—corresponding to any given eigenvalue.

Worked example

A worked example would normally have made its way here at this point in the article. But because the example is long and might not interest everyone, I have relegated it to the end of the article. Stay tuned if you are enthused.

Resources

I hope that this article has not been so brief as to be cryptic and off-putting. To those in search of greater rigour or a more formal exposition, I would recommend a good linear algebra textbook. The venerable

tome that I used at university went by the acronym "KKOP" after the initials of the surnames of the four authors.[1] Fortunately, it is freely available online in a variety of formats, and its actual title is *An Introduction to Linear Analysis* [1].

For something more contemporary, I would recommend the textbooks and lectures of Professor Gilbert Strang of MIT. They are attuned to those who *apply* mathematics, like engineers and scientists. There is an archived video of his lecture on eigenvalues and eigenvectors. There are also links to his MIT Open Course Ware (OCW) page for Course 18.06 of Spring 2010, his linear algebra textbook home page [2], and his academic home page.

Many academics make their lecture notes freely available online: google for them. You Tube videos of lectures are another source of information and knowledge, which offer the immediacy of a classroom lecture with the convenience of instant rewind in case you need to catch something you missed.

Online forums offer a slightly more interactive learning experience but again, their depth and quality varies. The Mathematics StackExchange and Quora are two sites that you might explore.

Examples of all the above types of resources have been tucked away within the various links in this article: try them out to get a flavour of what is available.

Importance and applications

If, after all this, you are still unconvinced about the utility of eigenvalues and eigenvectors, think of this analogy. Crystals have natural *cleavage planes* that allow them to be fractured easily along specific directions. This exploits the *symmetry* in the crystals. Likewise, eigenvalues and eigenvectors exploit the naturally occurring symmetries of mathematical structures and transformations to allow us to view them more simply and insightfully. Without eigenvalues and eigenvectors, we would have neither radios nor lasers.

To get an idea of the broad sweep of eigenvalues and their applicability, I strongly recommend that you should read a charming article entitled "Favourite Eigenvalue Problems". Another article that takes a breezy look at the subject of this writeup is "What the Heck are Eigenvalues and Eigenvectors?". It has a disputed explanation (see comments on the article) of how a bridge collapsed—so take that *cum grano salis*. It also contains a link to a PDF paper interestingly entitled "The 25,000,000,000.00 Dollar Eigenvector: The Linear Algebra Behind Google", which, in good faith, I think is not a spoof! Indeed, the citation to the original Stanford InfoLab technical report and the actual report are both available online.

Worked example: modelling weather with a transition matrix

Now for the promised example of eigenvalues at work—in a simplified real-life situation, modelling the weather. Let us assume that yesterday's weather influences the *probability* of today's weather, and today's weather influences the *probability* of tomorrow's weather. Each day's weather depends only on the previous day's weather, i.e., the weather has a "memory" of one day.

To keep it simple, let us have only three weather states: sunny, cloudy, and rainy, with the stipulation that each day can only be *one* of these three. Further, in our matrix, let the ordering be sunny, cloudy, and rainy, both left to right, and top to bottom. Then, the column headings represent *today's weather* and the row headings represent *tomorrow's weather*. We then have the *state-transition matrix* or *Markov matrix M* given in Equation (8):

$$M = \begin{bmatrix} 0.65 & 0.30 & 0.10 \\ 0.25 & 0.45 & 0.40 \\ 0.10 & 0.25 & 0.50 \end{bmatrix} \tag{8}$$

Note that each *column* of M represents the probabilities of *mutually exclusive* events, which must therefore sum to one. The matrix element m_{ij} is the probability that today's weather is in column j and that tomorrow's weather is in row i. For example, The probability of today being rainy and tomorrow cloudy is given by $m_{23} = 0.40$.

---

[1]Kreider, Kuller, Ostberg, and Perkins.

Let the column-vector $\boldsymbol{w}_k$ represent the probabilities for a particular day's weather and the column-vector $\boldsymbol{w}_{k+1}$, the next day's weather. The two are then related by:

$$\boldsymbol{w}_{k+1} = M\boldsymbol{w}_k \tag{9}$$

Equation (9)) is called a *recurrence relation* or *difference equation*, and in our case, it represents the evolution of a dynamical system in time, namely the weather. Just for completeness, let the initial condition be given by:

$$\boldsymbol{w}_0 = \begin{bmatrix} 0.55 \\ 0.34 \\ 0.11 \end{bmatrix} \tag{10}$$

We want to know whether, for this model, there will be an equilibrium or steady-state in the weather, represented by a probability vector with values that remain steady with temporal evolution. The question is how do we find that out?

One obvious way is to compute the downstream weather one day at a time: think of forging a chain one link at a time because the weather has a memory of only one day. From Equation (9) we can compute the following:

$$\boldsymbol{w}_1 = M\boldsymbol{w}_0$$
$$= \begin{bmatrix} 0.47050 \\ 0.33450 \\ 0.19500 \end{bmatrix}$$
$$\boldsymbol{w}_2 = M\boldsymbol{w}_1$$
$$= M(M\boldsymbol{w}_0)$$
$$= M^2\boldsymbol{w}_0$$
$$= \begin{bmatrix} 0.42568 \\ 0.34615 \\ 0.228180 \end{bmatrix}$$

By induction, the weather vector n days downstream is given by

$$\boldsymbol{w}_n = M^n\boldsymbol{w}_0. \tag{11}$$

In this manner, we can trace the time evolution of the weather and, if desired, draw a three-dimensional parametric plot of the successive weather vectors in $\mathbb{R}^3$ with time as parameter. This could be insightful, but it is a laborious and time-consuming way to find out the steady-state weather vector if there is one. Could we do better?

A rough and ready method would be to evaluate Equation (11) with n set to large numbers, say 50 and 100, and check if the resulting weather vectors, $\boldsymbol{w}_{50}$ and $\boldsymbol{w}_{100}$ were equal. If they were, we might hazard a guess that this unchanging value is the steady-state weather vector.

But computing the fiftieth or one-hunderdth power of a matrix is tedious and error-prone if done by hand, and computationally expensive if done by machine, especially if the matrix in question is large.

To devise a better solution, we need to digress briefly to examine diagonal matrices and the diagonalization of square matrices.

Diagonal matrix raised to a power

Suppose that D is a 3×3 diagonal matrix with non-zero entries on its principal diagonal and zeros elsewhere:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Observe that:

$$D^n = \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix}. \tag{12}$$

If we could somehow decompose M into a matrix product where a diagonal matrix was featured, we might be able to circumvent the matrix-raised-to-a large-power problem.

Matrix diagonalization or eigen decomposition

We need to *diagonalize* the transition matrix—a procedure called *eigen decomposition.* A square matrix with non-repeating eigenvalues and therefore, linearly independent eigenvectors, can be diagonalized. We demonstrate how this is done for the 3×3 case below.

Let the three eigenvectors be so denoted:

$$\boldsymbol{e}_k = \begin{bmatrix} e_{k1} \\ e_{k2} \\ e_{k3} \end{bmatrix}; k = 1, 2, 3.$$

The matrix P whose *columns* are the eigenvectors is therefore

$$P = \begin{bmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{bmatrix}$$

Now,

$$\begin{aligned} MP &= M \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{bmatrix} \\ &= \begin{bmatrix} M\boldsymbol{e}_1 & M\boldsymbol{e}_2 & M\boldsymbol{e}_3 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \lambda_3\boldsymbol{e}_3 \end{bmatrix} \\ &= \begin{bmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \\ &= PD \end{aligned} \tag{13}$$

To get only M on the left-hand side of Equation (13), we post-multiply both sides by P^{-1}, the inverse of P:

$$MPP^{-1} = MI = M = PDP^{-1} \tag{14}$$

If we now square M, we get

$$\begin{aligned} M^2 &= (PDP^{-1})(PDP^{-1}) \\ &= PD(P^{-1}P)DP^{-1} \\ &= P(DID)P^{-1} \\ &= PD^2P^{-1} \end{aligned} \tag{15}$$

By induction,

$$M^n = PD^nP^{-1} \tag{16}$$

The role of eigenvalues and eigenvectors in the plot of raising a square matrix to a power is now fully revealed: recall that in Equation (16), P is the matrix whose columns are the eigenvectors, and D is the diagonal matrix whose non-zero elements are the corresponding eigenvalues, and P^{-1} is the inverse of P.

Software Implementation

To get numerical results, I initially tried implementing the above steps with the free open-source mathematics software system SageMath, but found it less than convenient for my purpose.

I then experimented with GNU Octave, which is a free, platform-neutral, open source, high-level interpreted language, primarily intended for numerical computations. It was better suited to the task at hand, and I easily obtained the results discussed below.

The self-explanatory file, `weather.m`, may be downloaded and executed on the command line in the Octave command window. The discussion below will make better sense after you have thus executed the file `weather.m`. Instructions on how to download and set up Octave are given here.

Discussion of results from `weather.m`

The roots of the characteristic polynomial of M are first evaluated, and compared to the eigenvalues and eigenvectors obtained from an Octave function designed for that explicit purpose.

There are *three* distinct eigenvalues for the transition matrix, M. Moreover, it may be proved that 1 is an eigenvalue of any transition matrix. The eigenvalues we get are:

$$\lambda_1 = 1.00000$$
$$\lambda_2 = 0.48028$$
$$\lambda_3 = 0.11972$$

From Equation (16) we may surmise that the contributions from λ_2 and λ_3, being both less than one, will diminish progressively as n increases, and in the limit, only the eigenvalue 1 will dominate the steady-state behaviour.

The eigenvectors associated respectively with these eigenvalues, as spewed out by Octave, are:

$$\boldsymbol{e}_1 = \begin{bmatrix} -0.65375 \\ -0.61639 \\ -0.43895 \end{bmatrix} \quad \boldsymbol{e}_2 = \begin{bmatrix} -0.80430 \\ 0.28039 \\ 0.52391 \end{bmatrix} \quad \boldsymbol{e}_3 = \begin{bmatrix} 0.37921 \\ -0.81582 \\ 0.43662 \end{bmatrix}$$

None of the column sums of these eigenvectors sums to one. Indeed, the column sums of $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are close to zero, whereas the column sum of $\boldsymbol{e}_1$ is negative. We will return to the eigenvector $\boldsymbol{e}_1$, to wrest meaning out of it, a little later. In any case, we confirm that $M\boldsymbol{e}_1 = \boldsymbol{e}_1$.

Assembling the matrices P and P^{-1} from the eigenvectors is trivial, as is putting together D as the diagonal matrix of the corresponding eigenvalues. With these numerical values, the truth of Equation (14) is also easily demonstrated.

The time evolution of the initial weather vector is then tracked with 1, 10, 20, 50, and 100 iterations of Equation (9). In this case, the weather vector stabilizes after about twenty iterations to a steady-state vector, $\boldsymbol{w}_\infty$, given by

$$\boldsymbol{w}_\infty = \begin{bmatrix} 0.38251 \\ 0.36066 \\ 0.25683 \end{bmatrix} \tag{17}$$

When we track the same temporal evolution for eigenvector $\boldsymbol{e}_1$, the result after each iteration is $\boldsymbol{e}_1$ itself. This is the expected behaviour for an eigenvector associated with an eigenvalue of 1.

What may be disconcerting, though, is that we now seem to have *two* steady-state vectors, $\boldsymbol{w}_\infty$ and $\boldsymbol{e}_1$.

Observe, however, that $\boldsymbol{e}_1$ is not a probability vector whose columns sum to one. To convert it to a probability vector, we *normalize $\boldsymbol{e}_1$* by dividing it by its column sum, to get the *normalized eigenvector:*

$$\boldsymbol{n}_1 = \begin{bmatrix} 0.38251 \\ 0.36066 \\ 0.25683 \end{bmatrix} \tag{18}$$

Lo and behold! $\boldsymbol{n}_1$ and $\boldsymbol{w}_\infty$ are one and the same, and all is well. Recalling from Equation (7) that non-zero scalar multiples of eigenvectors are also eigenvectors themselves, this result, even if a little magical, really should not surprise us.

We do not bother normalizing the eigenvectors associated with λ_2 and λ_3 because their column sums almost vanish, and moreover, their contribution to the steady state decreases with increasing number of iterations.

To round things off, we substitute a random initial weather vector in place of $\boldsymbol{w}_0$, and view its evolution over time for twenty iterations, and find that it too converges to the steady-state weather vector after about 15 iterations.

This means that regardless of what initial weather vector we start with, in about two weeks we will end up with a vector that represents the steady-state.

Observations like these suggest that our inferences are only as good as our assumptions and models. Oversimplification could lead to absurd results, and weather prediction over time is a seriously non-trivial problem.

One general hypothesis that we could examine is whether it is generally true that the normalized eigenvector associated with an eigenvalue of 1 does indeed represent the steady state of the system represented by any transition matrix. If so, we would have a theorem on our hands.

Feedback Please email me your comments and corrections.

A PDF version of this article is available for download here.

References

[1]   D. L. Kreider, R. G. Kuller, D. R. Ostberg, and F. W. Perkins, *An Introduction to Linear Analysis*. Addison-Wesley Publishing Company, 1966.

[2]   G. Strang, *Introduction to Linear Algebra*, 5th ed. Wellesley-Cambridge Press, 2016.