[MATH-07] Eigenvalues and eigenvectors

Miguel-Angel Canela
Associate Professor, IESE Business School

Eigenvalues and eigenvectors

Let $\mathbf{A}$ be an $(n \times n)$–matrix, $\mathbf{v} \neq \mathbf{0}$ an n-vector, and λ a scalar, such that $\mathbf{A}\mathbf{v} = \lambda\,\mathbf{v}$. We say then that $\mathbf{v}$ is an **eigenvector** of $\mathbf{A}$, and that λ is the associated **eigenvalue**. In particular, any non-zero vector of the null space (if it exists) is an eigenvector, with eigenvalue zero.

Just a few numbers can be eigenvalues for a particular matrix. To see why, note that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \Longleftrightarrow \big(\mathbf{A} - \lambda\mathbf{I}\big)\mathbf{v} = \mathbf{0}.$$

Therefore, λ is an eigenvalue when the null space of $\mathbf{A} - \lambda\,\mathbf{I}$ contains nonzero vectors, i.e. when λ is a solution of the **characteristic equation**

$$\det\big(\mathbf{A} - \lambda\,\mathbf{I}\big) = 0.$$

Example 1. Consider the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$. Here,

$$\det\big(\mathbf{A} - \lambda\,\mathbf{I}\big) = \begin{vmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 2 = 0$$

has two solutions, $\lambda_1 = -1$ and $\lambda_2 = 2$. These are the eigenvalues of $\mathbf{A}$. The eigenvectors associated to $\lambda_1 = -1$ are those vectors satisfying

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow x_1 + x_2 = 0.$$

The set of these eigenvectors is a linear subspace of dimension 1 (the null space of $\mathbf{A} + \mathbf{I}$). A basis of this subspace is given by the vector

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The eigenvectors associated to $\lambda_2 = 2$ are the solutions of

$$\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow x_1 - 2x_2 = 0.$$

These vectors form a linear subspace of dimension 1 (the null space of $\mathbf{A} - 2\,\mathbf{I}$), with basis

$$\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \ \square$$

Example 2. Let me consider now

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

$\mathbf{A}$ has two eigenvalues, $\lambda_1 = 0$ and $\lambda_2 = 3$, which are the solutions of the characteristic equation

$$\begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = \lambda^2(3-\lambda) = 0.$$

The eigenvectors associated to $\lambda_1 = 0$ satisfy

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longleftrightarrow x_1 + x_2 + x_3 = 0,$$

which defines a linear subspace of dimension 2 (the null space of $\mathbf{A}$). A basis is given by

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

The eigenvectors associated to $\lambda_2 = 3$ satisfy

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longleftrightarrow x_1 = x_2 = x_3,$$

which defines a linear subspace of dimension 1 (the null space of $\mathbf{A} - 3\,\mathbf{I}$). A basis is given by

$$\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \ \square$$

Diagonalization

Diagonalization is the key of important multivariate statistical methods, such as principal component analysis. A square matrix $\mathbf{A}$ is said to be diagonalizable if it admits a factorization $\mathbf{A} = \mathbf{B}\boldsymbol{\Lambda}\mathbf{B}^{-1}$, with $\mathbf{B}$ nonsingular and $\boldsymbol{\Lambda}$ diagonal. This is equivalent to the existence of a non-singular matrix $\mathbf{B}$ such that $\boldsymbol{\Lambda} = \mathbf{B}^{-1}\mathbf{A}\mathbf{B}$ is diagonal. This is the origin of the name.

It can be proved that $\mathbf{A}$ is diagonalizable if and only if there exists a basis of eigenvectors of $\mathbf{A}$. Then, the vectors of that basis coincide with the columns of $\mathbf{B}$ and the eigenvalues with the diagonal terms in $\boldsymbol{\Lambda}$. When is this possible? Suppose that the dimension of the matrix is n.

- If the number of eigenvalues coincides with n (the maximum possible), we pick an eigenvector for each eigenvalue, forming a basis, so that the matrix is diagonalizable.
- When the number of eigenvalues is less than n, it can fail to be diagonalizable. If the characteristic equation has complex non-real solutions, it is not diagonalizable. If all the solutions are real, but some are multiple, we associate to each multiple solution a subspace of eigenvectors. If the dimension equals the multiplicity, diagonalization is possible, $\boldsymbol{\Lambda}$ having the multiple eigenvalue repeated according to its multiplicity.

Example 3. The matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

is not diagonalizable, since the characteristic equation $\lambda^2 + 1 = 0$ has no real solution. $\square$

For a diagonalizable matrix, there are two formulas which are sometimes useful, when searching the eigenvalues (or their signs). They relate the sum and the product of the eigenvalues to the trace and the determinant:

- $\operatorname{tr}(\mathbf{A}) = \lambda_1 + \cdots + \lambda_n$.
- $\det \mathbf{A} = \lambda_1 \cdots \lambda_n$.

In any of these formulas, a multiple eigenvalue must be repeated according to its multiplicity.

Example 1 (continuation). In Example 1,

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}, \qquad \mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.$$

The trace is 1 and the determinant -2. Note that we can find the eigenvalues solving the system

$$\lambda_1 + \lambda_2 = 1, \qquad \lambda_1\lambda_2 = -2.$$

Of course, this is not possible if the dimension is higher than two. □

Example 2 (continuation). In Example 2,

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \qquad \mathbf{\Lambda} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

The trace equals 3 and the determinant equals 0. □

Orthogonal matrices

Orthogonal matrices are special nonsingular matrices that appear in many applications, in particular in multivariate statistics. $\mathbf{A}$ is said to be orthogonal when its inverse coincides with its transpose, i.e. when $\mathbf{A}^{-1} = \mathbf{A}^{\mathsf{T}}$. This is equivalent to $\mathbf{A}^{\mathsf{T}}\mathbf{A} = \mathbf{I}$. It is can be easily proved that a matrix is orthogonal when its columns (or its rows) are pairwise orthogonal unit vectors. This is the easy way of checking the orthogonality of a matrix.

Example 4. The matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is symmetric and orthogonal. Hence, $\mathbf{A}^{-1} = \mathbf{A}$. □

Linear operators defined by orthogonal matrices are sometimes (improperly) called **rotations**. Characterizing orthogonal two-dimensional matrices helps to understand how orthogonal matrices are related to rotations, as we seen next.

A unit vector $\mathbf{u}$ in E_2 can be placed as an arrow in the plane, starting at the origin and ending at a point of the unit circle. This allows for a representation of this vector as

$$\mathbf{u} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix},$$

where θ denotes the angle from the positive x half-axis to the arrow $\mathbf{u}$ (counterclockwise). Now, I call $\mathbf{u}_1$ and $\mathbf{u}_2$ the columns of an orthogonal 2–dimensional matrix $\mathbf{A}$. Since these vectors are orthogonal, $\theta_2 = \theta_1 \pm \pi/2$ (use a picture if you do not see this immediately). If $\theta_2 = \theta_1 + \pi/2$,

$$\cos\theta_2 = -\sin\theta_1, \qquad \sin\theta_2 = \cos\theta_1,$$

so $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Here, the columns correspond to the images of the points $(1,0)$ and $(0,1)$ under the transformation defined by $\mathbf{A}$. It is not hard to see, in a picture, that we have a rotation of angle θ around the origin. A similar argument proves that, when $\theta_2 = \theta_1 - \pi/2$, $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

What is this? It follows from the identity

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

that $\mathbf{A}$ defines a rotation on top of a change of sign in the second coordinate. Such change-of-sign transformation is called an **inversion**. Therefore, orthogonal matrices of dimension 2 define either a rotation or an inversion plus a rotation.

Diagonalization of symmetric matrices

Not every matrix is diagonalizable, as shown by Example 3. But an important theorem of matrix algebra ensures that every symmetric matrix is diagonalizable. Since, in most applications, the matrices are symmetric (e.g. covariance matrices in statistics, Hessian matrices in differential calculus), this theorem will probably cover your needs.

Theorem. If $\mathbf{A}$ is a symmetric (n,n)–matrix, there is a basis of E_n formed by orthogonal eigenvectors.

Then, for a symmetric matrix, an orthogonal matrix of (unit) eigenvectors can be found. The diagonalization formula becomes $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, called the **spectral decomposition**.

Example 2 (continuation). In Example 2, $\mathbf{B}$ is not orthogonal, but replacing $\mathbf{v}_2$ by

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix},$$

we get a basis of orthogonal eigenvectors. Now I divide these vectors by the respective norms, obtaining a basis of orthogonal unit eigenvectors,

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ -2/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}. \ \square$$

The spectral decomposition is then

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & -2/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 0 & & \\ & 0 & \\ & & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & -2/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}. \ \square$$

Homework

A. Find the eigenvalues and eigenvectors of $\mathbf{A} = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}$.

B. Diagonalize, if possible, $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$.

C. Check that the operator defined by $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is a rotation of the plane.

D. Find the spectral decomposition of

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}.$$

E. Check that $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ does not rotate all the vectors in the same way.

F. Find the eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$