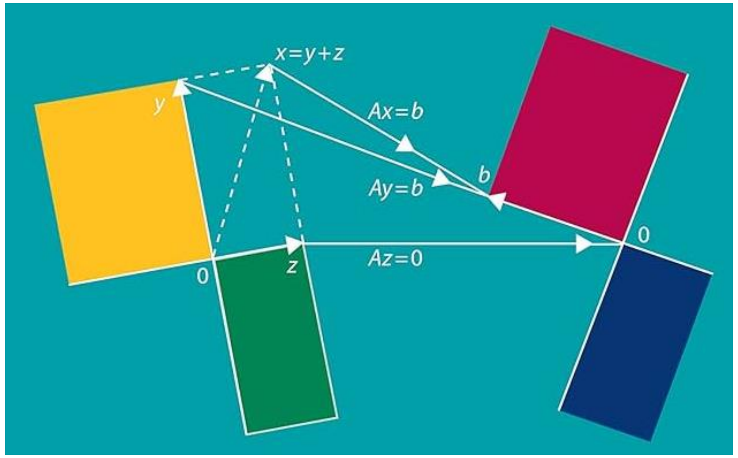
Eigenvalues and Eigenvectors

(Linear Algebra)

Randil Pushpananda, PhD
rpn@ucsc.cmb.ac.lk



When a square matrix A acts upon a vector $\mathbf{x}$, it generally outputs a new vector $A\mathbf{x}$. Usually this new vector will be stretched and rotated. For example, if we take the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and apply it to the vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ we get:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

If we apply A to the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ we get:

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

if we apply A to the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ we get:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The matrix A leaves the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ unchanged! In particular, it does not rotate the vector. When a matrix A acts upon a vector and does not rotate it, we have $A\mathbf{x} = \lambda\mathbf{x}$, where λ is a scaling factor. In our example $\lambda = 1$. We call such a vector an *eigenvector* for the matrix A, and the associated scaling factor λ an *eigenvalue*.

Definition

Let A be an n × n matrix. If there exist a real value λ and a non-zero n × 1 vector x satisfying,

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}$$

then

we refer to λ as an eigenvalue of A, and x as an eigenvector of A corresponding to λ.

The basic equation for eigenvectors and eigenvalues is

$$A\mathbf{x} = \lambda\mathbf{x}$$

where A is a matrix and λ is a number. One property we see right away is

$$A^2\mathbf{x} = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

and in general

$$A^n\mathbf{x} = \lambda^n\mathbf{x}.$$

So, the eigenvectors of A^n are the same as the eigenvectors of A, while the eigenvalues for A^n are the eigenvalues for A raised to the nth power.

Purpose of Learning in Computer Science

- Data Compression and Dimensionality Reduction:
  - reduce the dimensionality of large datasets, increasing interpretability while minimizing information loss.
- Machine Learning:
  - Feature Extraction and Selection
- Computer Vision and Image Processing:
  - Image Compression and Face Recognition
- Natural Language Processing (NLP):
  - Word Embedding

How to calculate?

$$A\mathbf{x} = \lambda\mathbf{x}$$

then

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

So, the matrix $(A - \lambda I)$ has a nontrivial nullspace, and therefore must be singular. So,

$$det(A - \lambda I) = 0.$$

So, if λ is an eigenvalue of A then $det(A - \lambda I) = 0$. It turns out that this equation can be used to calculate *every* eigenvalue. The equation $det(A - \lambda I) = 0$ will be a polynomial equation in λ, and its roots will give us all the eigenvalues. We call this polynomial equation the *characteristic equation* of A.

Eigenvalues first. If $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nonzero solution, $A - \lambda I$ is not invertible. ***The determinant of A − λI must be zero.*** This is how to recognize an eigenvalue λ:

Eigenvalues The number λ is an eigenvalue of A if and only if $A - \lambda I$ is singular.

Equation for the n eigenvalues of A $$\boxed{\det(A - \lambda I) = 0.} \tag{5}$$

This "*characteristic polynomial*" $\det(A - \lambda I)$ involves only λ, not $\boldsymbol{x}$. Since λ appears all along the main diagonal of $A - \lambda I$, the determinant in (5) includes $(-\lambda)^n$. Then equation (5) has n solutions λ_1 to λ_n and **A has n eigenvalues.**

An n by n matrix has n eigenvalues (repeated λ's are possible !) Each λ leads to $\boldsymbol{x}$:

For each eigenvalue λ solve $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ or $A\boldsymbol{x} = \lambda\boldsymbol{x}$ to find an eigenvector $\boldsymbol{x}$.

Example 4 $A = \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{2} & \mathbf{4} \end{bmatrix}$ is already singular (zero determinant). Find its λ's and $\boldsymbol{x}$'s.

When A is singular, $\lambda = 0$ is one of the eigenvalues. The equation $A\boldsymbol{x} = 0\boldsymbol{x}$ has solutions. They are the eigenvectors for $\lambda = 0$. But $\det(A - \lambda I) = 0$ is the way to find *all* λ's and $\boldsymbol{x}$'s. Always subtract λI from A:

$$\textit{\textbf{Subtract } } \boldsymbol{\lambda} \textit{\textbf{ along the diagonal to find}} \quad A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}. \tag{6}$$

Take the determinant "$ad - bc$" of this 2 by 2 matrix. From $1 - \lambda$ times $4 - \lambda$, the "ad" part is $\lambda^2 - 5\lambda + 4$. The "bc" part, not containing λ, is 2 times 2.

$$\det \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} = (1-\lambda)(4-\lambda) - (2)(2) = \lambda^2 - 5\lambda. \tag{7}$$

Set this determinant $\boldsymbol{\lambda^2 - 5\lambda}$ to zero. One solution is $\lambda = 0$ (as expected, since A is singular). Factoring into λ times $\lambda - 5$, the other root is $\lambda = 5$:

$$\boxed{\det(A - \lambda I) = \lambda^2 - 5\lambda = 0 \quad \text{yields the eigenvalues} \quad \boldsymbol{\lambda_1 = 0} \quad \text{and} \quad \boldsymbol{\lambda_2 = 5}.}$$

Now find the eigenvectors. Solve $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ separately for $\lambda_1 = 0$ and $\lambda_2 = 5$:

$$(A - 0I)\boldsymbol{x} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \mathbf{2} \\ \mathbf{-1} \end{bmatrix} \text{ for } \boldsymbol{\lambda_1} = \mathbf{0}$$

$$(A - 5I)\boldsymbol{x} = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \end{bmatrix} \text{ for } \boldsymbol{\lambda_2} = \mathbf{5}$$

Note that these eigenvectors are *not* unique. In fact, any non-zero multiple $c\mathbf{x}$ ($c \neq 0$) of an eigenvector is another eigenvector.

Oh, and we should probably mention right now that if $\mathbf{x} = 0$ then $A\mathbf{x} = \lambda\mathbf{x}$ for *any* λ. So, we have to put in a qualifier that *the zero vector is never an eigenvector*. Please keep in mind that the number 0 can certainly be an eigenvalue.

Some Facts About Eigenvectors

First, some bad news. We cannot use elimination to calculate eigenvalues. Sorry. If we use elimination to convert a a matrix A into an upper triangular matrix U, the eigenvalues of U could be different than the eigenvalues of A. However, the two will not be completely unrelated.

What relates them is the amazing fact that the determinant of a matrix is equal to the product of its eigenvalues. That is to say, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of a matrix A, then

$$det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

So, because A and U have the same determinant, the product of their eigenvalues will be the same.

Some Facts About Eigenvectors

Finally, we note that the trace of a matrix is defined as being the sum of the diagonal elements.

$$trace(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

The trace of a matrix will be equal to the sum of the eigenvalues of the matrix

$$trace(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Diagonalizing a Matrix

Diagonalization Suppose the n by n matrix A has n linearly independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Put those $\boldsymbol{x}_i$ into the columns of an invertible ***eigenvector matrix*** $\boldsymbol{X}$. Then $X^{-1}AX$ is the diagonal ***eigenvalue matrix*** $\boldsymbol{\Lambda}$:

Eigenvector matrix X
Eigenvalue matrix Λ

$$\boldsymbol{X^{-1}AX = \Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \qquad (1)$$

The matrix A is "diagonalized." We use capital lambda for the eigenvalue matrix, because the small λ's (the eigenvalues) are on its diagonal.

Diagonalizing a Matrix

Example 1 This A is triangular so its eigenvalues are on the diagonal: $\lambda = 2$ and $\lambda = 6$.

$$\boxed{\textbf{Eigenvectors go into } \boldsymbol{X} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \qquad \underset{X^{-1}}{\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} 2 & 4 \\ 0 & 6 \end{bmatrix}} \underset{X}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}} \underset{=}{=} \underset{\Lambda}{\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}}$$

In other words $A = X\Lambda X^{-1}$. Then watch $A^2 = X\Lambda X^{-1}X\Lambda X^{-1}$. So A^2 is $X\Lambda^2X^{-1}$.

A^2 has the same eigenvectors in X. It has squared eigenvalues 4 and 36 in Λ^2.

Diagonalizing a Matrix

we want to express the matrix as a product of three matrices in the form:

$$A = S\Lambda S^{-1}$$

where Λ is a diagonal matrix. In particular, the diagonal entries of Λ will be the eigenvalues of A, and the columns of S will be the corresponding eigenvectors. Having A in this form can greatly simplify many calculations, particularly calculations involving powers of A.

Diagonalizing a Matrix

$$\boxed{AX = X\Lambda \quad \text{is} \quad \boldsymbol{X^{-1}AX = \Lambda} \quad \text{or} \quad \boldsymbol{A = X\Lambda X^{-1}}.} \tag{2}$$

The matrix X has an inverse, because its columns (the eigenvectors of A) were assumed to be linearly independent. *Without n independent eigenvectors, we can't diagonalize A.*

A and Λ have the same eigenvalues $\lambda_1, \ldots, \lambda_n$. The eigenvectors are different. The job of the original eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ was to diagonalize A. Those eigenvectors in X produce $A = X\Lambda X^{-1}$. You will soon see their simplicity and importance and meaning. The kth power will be $A^k = \boldsymbol{X\Lambda^k X^{-1}}$ which is easy to compute using $\boldsymbol{X^{-1}X = I}$:

$$\boldsymbol{A^k = (X\Lambda X^{-1})(X\Lambda X^{-1})\ldots(X\Lambda X^{-1}) = X\Lambda^k X^{-1}}.$$

■

Things to Remember

Remark 1 Suppose the eigenvalues are n different numbers (like 2 and 6). Then it is automatic that the n eigenvectors will be independent. The eigenvector matrix X will be *invertible*. ***Any matrix that has no repeated eigenvalues can be diagonalized.***

Remark 2 *We can multiply eigenvectors by any nonzero constants.* $A(c\boldsymbol{x}) = \lambda(c\boldsymbol{x})$ is still true. In Example 1, we can divide $\boldsymbol{x} = (1, 1)$ by $\sqrt{2}$ to produce a *unit vector*.

MATLAB and virtually all other codes produce eigenvectors of length $||\boldsymbol{x}|| = 1$.

Remark 3 The eigenvalues in Λ come in the same order as the eigenvectors in X. To reverse the order of 2 and 6 in Λ, put the eigenvector $(1, 1)$ before $(1, 0)$ in X:

$$X^{-1}AX = \Lambda_{\text{new}} \qquad \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 0 \\ 0 & \mathbf{2} \end{bmatrix} = \Lambda_{\text{new}}$$

Things to Remember

Remark 4 (repeated warning for repeated eigenvalues) Some matrices have too few eigenvectors. *Those matrices cannot be diagonalized*. Here are two examples:

$$\textbf{Not diagonalizable} \qquad A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Their eigenvalues happen to be 0 and 0. Nothing is special about $\lambda = 0$, the problem is the repetition of λ. All eigenvectors of the first matrix are multiples of $(1, 1)$:

$$\begin{array}{l}\textbf{Only one line}\\ \textbf{of eigenvectors}\end{array} \qquad A\boldsymbol{x} = 0\boldsymbol{x} \quad \text{means} \quad \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{x} = c \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

There is no second eigenvector, so this unusual matrix A cannot be diagonalized.

Those matrices are the best examples to test any statement about eigenvectors. In many true-false questions, non-diagonalizable matrices lead to *false*.

Remember that there is no connection between invertibility and diagonalizability:

Things to Remember

Remark 1 Suppose the eigenvalues are n different numbers (like 2 and 6). Then it is automatic that the n eigenvectors will be independent. The eigenvector matrix X will be *invertible*. ***Any matrix that has no repeated eigenvalues can be diagonalized.***

Remark 2 *We can multiply eigenvectors by any nonzero constants.* $A(c\boldsymbol{x}) = \lambda(c\boldsymbol{x})$ is still true. In Example 1, we can divide $\boldsymbol{x} = (1, 1)$ by $\sqrt{2}$ to produce a *unit vector*.
MATLAB and virtually all other codes produce eigenvectors of length $||\boldsymbol{x}|| = 1$.

Remark 3 The eigenvalues in Λ come in the same order as the eigenvectors in X. To reverse the order of 2 and 6 in Λ, put the eigenvector $(1, 1)$ before $(1, 0)$ in X:

$$X^{-1}AX = \Lambda_{\text{new}} \qquad \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 0 \\ 0 & \mathbf{2} \end{bmatrix} = \Lambda_{\text{new}}$$