Autovalores y autovectores

Credit: This notes are 100% from chapter 4 of the book entitled *Linear Algebra. A Modern Introduction* by David Poole. Thomson. Australia. 2006.

Introduction

For a square matrix A, we ask whether there exist nonzero vectors $\bar{x}$ such that $A\bar{x}$ is just a scalar multiple of $\bar{x}$. This is the eigenvalue problem, and it is one of the most central problems in linear algebra. It has applications thorough mathematics and in many other fields as well.

Eigenvalue: Let A be an $n \times n$ matrix. A scalar λ is called an eigenvalue of A if there is a nonzero vector $\bar{x}$ such that $A\bar{x} = \lambda\bar{x}$. Such a vector $\bar{x}$ is called an eigenvector of A corresponding to λ.

Geometric interpretation: In $\mathbb{R}^2$, the eigenvalue equation $A\bar{x} = \lambda\bar{x}$ says that the vectors $A\bar{x}$ and $\bar{x}$ are parallel. Thus, $\bar{x}$ is an eigenvector of A if and only if A transforms $\bar{x}$ into a parallel vector, i.e. $T_A(\bar{x})$ is parallel to $\bar{x}$, where T_A is the matrix transformation (MT) corresponding to A.

Example: Let us calculate the eigenvalue of $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ associate to the eigenvector $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
We compute $A\bar{x}$,

$$A\bar{x} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\bar{x} \tag{1}$$

then $\lambda = 4$.

Example 4.2: Let us calculate the eigenvectors of $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ associate to the eigenvalue $\lambda = 5$.
From $A\bar{x} = \lambda\bar{x}$ we have

$$(A - \lambda I)\bar{x} = (A - 5I)\bar{x} = 0 \tag{2}$$

This implies we have to compute the null space of the matrix $A - 5I$,

$$\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix} \rightarrow \left[\begin{array}{rr|r} -4 & 2 & 0 \\ 4 & -2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \tag{3}$$

then

$$\begin{aligned} x_1 - \frac{1}{2}x_2 &= 0 \qquad (4)\\ 0x_1 + 0x_2 &= 0 \qquad (5)\end{aligned}$$

Then, any vector of the form $\begin{bmatrix} t \\ 2t \end{bmatrix}$ is an eigenvector of A with eigenvalue $\lambda = 5$ with t in $\mathbb{R}$.

Eigenspace: Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. The collection of all eigenvectors corresponding to λ, together with the zero vector, is called the eigenspace of λ and is denoted by E_λ.

Example 1: In the previous example 4.2 the eigenspace is $E_5 = t\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Example 2 Let us verifies that $\lambda = 6$ is an eigenvalue of $A = \begin{bmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{bmatrix}$ and find a basis for its eigenspace.

First we calculate the null space of $A - 6I$

$$\begin{aligned} A - 6I &= \begin{bmatrix} 1 & 1 & -2 \\ -3 & -3 & 6 \\ 2 & 2 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad (6)\\ &\implies x_1 + x_2 - 2x_3 = 0 \qquad (7)\end{aligned}$$

Then,

$$\begin{aligned} E_6 &= \left\{ \begin{bmatrix} -x_2 + 2x_3 \\ x_2 \\ x_3 \end{bmatrix} \right\} = \left\{ x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\} \qquad (8)\\ &= span\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right) \qquad (9)\end{aligned}$$

Calculation of eigenvalue for 2×2 matrices: How do we find the eigenvalue and eigenvector of a matrix if we do not know any of them from the very beginning?. The key is the observation that λ is an eigenvalue of A if and only if the null space of $A - \lambda I$ is nontrivial. For a two-by-two matrix we know from section 3.3 that A is invertible if and only if $det(A)$ is nonzero. We also know from the fundamental theorem (FT) of invertible matrices that a matrix has a nontrivial null space if and only if it is noninvertible, hence, if and only if its determinant is zero. Putting these facts together, we see that (for two by two matrices at least) λ is an eigenvalue of A if and only if

$$det(A - \lambda I) = 0 \qquad (10)$$

Application: Find all of the eigenvalues and the corresponding eigenvectos of the matrix $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

- First we find the zeros of $det(A - \lambda I)$. They are $\lambda_1 = 4, \lambda_2 = 2$.
- Next we find the null space of the matrix $A - \lambda I$ separately for λ_1 and λ_2. They are $E_4 = span([1\ 1]^T)$ and $E_2 = span([-1\ 1]^T)$, respectively.

Exercise for the student in class: Find all of the eigenvalues and the corresponding eigenvectos of the matrix $A = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}$
Solution: Pending...

Determinants

Tema visto en Álgebra II .
For any square matrix $A = [a_{ij}]$ of order n, the determinant is the scalar

$$det\, A = |A| = \sum_{j=1}^{n} (-1)^{j+1}\, a_{1j}\, det\, A_{1j} \tag{11}$$

where A_{ij}, called the (i, j)-minor of A, is the matrix obtained by deleting the row i and column j from A.

Example 4.8: By computing the determinant of $A = \begin{bmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$ we get: det$A$=5.

About the column expansion: The election of the first row is arbitrary. Any column can be use to expand the determinant. But is an even row is used then the sing must change consistently. For example, using the second column the definition change to

$$det\, A = |A| = \sum_{j=1}^{n} (-1)^{j}\, a_{2j}\, det\, A_{2j} \tag{12}$$

Exercise for the student in class: Apply the above definition to the example 4.8.

Determinants of $n \times n$ Matrices

Cofactor: It is convenient to combine a minor with its plus or minus sign. To this end, we define the (i, j)-cofactor of A to be

$$C_{ij} = (-)^{i+j}\, det\, A_{ij} \tag{13}$$

Theorem 4.1: The Laplace expansion Theorem. The previous definition of the determinant were done in terms of row expansion. They are also valid if we used instead column expansion: The determinant of an $n \times n$ matrix $A = [a_{ij}]$, where $n \geq 2$, can be computed as (row expansion)

$$det\, A = |A| = \sum_{j=1}^{n} a_{ij}\, C_{ij} \tag{14}$$

and also as (column expansion)

$$det\, A = |A| = \sum_{i=1}^{n} a_{ij}\, C_{ij} \tag{15}$$

Proof: See libro, pag. 279 (see before the Lemmas from pag. 276.)

Theorem 4.2: The determinant of a triangular matrix is the product of the entries on its main diagonal: $det\, A = \sum_{i=1}^{n} a_{ii}$.

Example 4.12: By computing the determinant of $A = \begin{bmatrix} 2 & -3 & 1 & 0 & 4 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$ using columns expansion (for example) we get: det$A$=-30. While the product of the diagonal give $detA = 2 \times 3 \times 1 \times 5 \times (-1)$=-30.

Properties of Determinants

Theorem 4.3: The most efficient way to compute determinants is to use row reduction. This theorem summarizes the main properties needed in order to used row reduction effectively. Let $A = [a_{ij}]$ be square matrix.

a. If A has a zero row (column), then $det(A) = 0$.

b. If B is obtained by interchanging two rows (columns) of A, then $det(B) = -det(A)$.

c. If A has two identical rows (columns), then $det(A) = 0$.

d. If B is obtained by multiplying a row (column) of A by k, then $det(B) = k\, det(A)$.

e. If A, B, and C are identical except that the ith row (column) of C is the sum of the ith rows (columns) of A and B, then $det(C) = det(A) + det(B)$.

f. If B is obtained by adding a multiple of one row (column) of A to another row (column), then $det(B) = det(A)$.

Proof: See book, page 269.

Example 4.13: Let us calculate the determinant of the matrices A and B, with

$$A \;=\; \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 3 \\ -4 & -6 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 2 & -4 & 5 \\ 3 & 0 & -3 & 6 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{bmatrix} \tag{16}$$

We should get: $det(A) = 0$ and $det(B) = 585$.

Determinants of Elementary Matrices

Recalled that an elementary matrix results from performing an elementary row operation on an identity matrix. Setting $A = I_n$ in Theorem 4.3 yields the following theorem.

Theorem 4.4: Let E be an $n \times n$ elementary matrix.

a. If E results from interchanging two rows of I_n, then $det(E) = -1$.

b. If E results from multiplying one row of I_n by k, then $det(E) = k$.

c. If E results from adding a multiple of row of I_n to another row, then $det(E) = 1$.

Proof: Since $det(I_n) = 1$, applying (b), (d), and (f) of Theorem 4.3 immediately gives (a), (b), and (c), respectively.

Lemma 4.5: Since that multiplying a matrix B by an elementary matrix on the left performs the corresponding elementary row operation of B. We can rephrase (b), (d), and (f) of Theorem 4.3 by $\det(EB)=(\det E)(\det B)$.

Theorem 4.6: A square matrix A is invertible if and only if $\det A \neq 0$.
Proof: Let A be an $n \times n$ matrix and let R be the reduced row echelon form of A, i.e. $E_r \cdots E_1 A = R$. By taking determinant and applying Lemma 4.5, we get $(det\, E_r) \cdots (det\, E_1)(det\, A) = (det\, R)$. By Theorem 4.4, the determinants $det\, E_i$ are nonzero for all $i = 1, \cdots, r$. Then, $\det A \neq 0$ if and only if $\det R \neq 0$.
Now suppose that A is invertible. Then, by the FT of IM, $R = I_n$, so $\det R = 1 \neq 0$. Hence, $\det A \neq 0$ also.
Conversely, if $\det A \neq 0$, then $\det R \neq 0$, then R cannot contain a zero row by Theorem 4.3.a. It follows that R must be I_n, so A is invertible, by the FT.

Determinants and Matrix Operations

This section is about to determine relationships between determinants and some of the basic matrix operations.

Theorem 4.7: If A is an $n \times n$ matrix, then

$$det(kA) = k^n\, detA \tag{17}$$

Proof: Exercise 44.

Theorem 4.8: If A and B are $n \times n$ matrices, then

$$det(AB) = (detA)(detB) \tag{18}$$

Proof: See book, pag. 272.

Theorem 4.9: If A is invertible, then

$$det(A^{-1}) = \frac{1}{detA} \tag{19}$$

Proof: See book, pag. 273.

Theorem 4.10: For any square matrix A,

$$det(A) = detA^T \tag{20}$$

Proof: Since the rows of A^T are the columns of A evaluating $\det A^T$ by expanding along the first row is identical to evaluating $\det A$ by expanding along its first column.

Cramer's Rule and the Adjoint

The Cramer's Rule gives a formula for describing the solution of certain systems of n linear equations in n variables entirely in terms of determinants.

Theorem 4.11: Cramer's Rule. Let A be an invertible $n \times n$ matrix and let $\bar{b}$ be a vector in $\mathbb{R}^n$. Then the unique solution $\bar{x}$ of the system $A\bar{x} = \bar{b}$ is given by

$$x_i = \frac{det(A_i(\bar{b}))}{detA} \tag{21}$$

for $i = 1, \cdots, n$ and $A_i(\bar{b})$ the matrix obtained by replacing the ith column of A by $\bar{b}$, i.e. $A_i(\bar{b}) = [\bar{a}_1 \ \cdots \ \bar{a}_{i-1} \ \bar{b} \ \bar{a}_{i+1} \ \cdots \ \bar{a}_n]$.
Proof: See book, pag. 274.

Example 4.16: *Para trabajar en clase:* Let us check that the solution of the following LSE

$$x_1 + 2x_2 = 2 \tag{22}$$
$$-x_1 + 4x_2 = 1 \tag{23}$$

is $x_1 = 1$, $x_2 = 0.5$, by using the Cramer's Rule.

Theorem 4.12: Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{detA} adjA \tag{24}$$

where $adjA$ the so called adjoint matrix of A, defined as $[C_{ji}] = [C_{ij}]^T$, where $[C_{ij}]$ is a matriz with elements C_{ij}, with C_{ij} the co-factors defined in Eq. (13).
Proof: See book, pag. 275.

Example 4.17 Calculates the inverse of the matrix A using the adjoint matrix, with $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$
Solution: the determinant is 2, the cofactors are

$$C_{11} = -18 \tag{25}$$
$$C_{12} = 10 \tag{26}$$
$$C_{13} = 4 \tag{27}$$
$$C_{21} = 10 \tag{28}$$
$$C_{22} = -2 \tag{29}$$
$$C_{23} = -1 \tag{30}$$
$$C_{31} = 10 \tag{31}$$
$$C_{32} = -6 \tag{32}$$
$$C_{33} = -2 \tag{33}$$

then

$$A^{-1} = \frac{1}{detA} adjA = \frac{1}{2} \begin{bmatrix} -18 & 10 & 4 \\ 3 & -2 & -1 \\ 10 & -6 & -2 \end{bmatrix}^T = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix} \tag{34}$$

Compare this result with example 3.30

Lemma 4.13: Let A be an $n \times n$ matrix. Then

$$\sum_{i=1}^{n} a_{1i}C_{1i} = detA = \sum_{i=1}^{n} a_{i1}C_{i1} \tag{35}$$

Lemma 4.14: Let A be an $n \times n$ matrix and let B be obtained by interchanging any two rows (columns) of A. Then

$$detB = -detA \tag{36}$$

Eigenvalues and Eigenvectors of $n \times n$ Matrices

The eigenvalues of a square matrix A are precisely the solutions λ of the equation

$$det(A - \lambda I) = 0 \tag{37}$$

Characteristic polynomial/equation: When we expand $det(A - \lambda I)$, we get a polynomial in λ, called the characteristic polynomial of A. The equation $det(A - \lambda I) = 0$ is called the characteristic equation of A.

- If A is $n \times n$, its characteristic polynomial will be of degree n.
- According to the Fundamental Theorem of Algebra (Theorem D.4 of the Appendix D of the book, pag. 668), a polynomial of degree n with real or complex coefficients has at most n distinct roots.
- Then an $n \times n$ matrix with real or complex entries has at most n distinct eigenvalues.

Example 4.18: Find the eigenvalues and the corresponding eigenspaces of the matrix A, with $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$.

First we calculates the eigenvalues from $det(A - \lambda I)$. We should get

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0 \tag{38}$$

Using the Rational Roots Theorem (Theorem D.3 in page 665 in theAppendix D of the book) we get $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$.

Some details about the Rational Roots Theorem: Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients and a/b be a rational number. If a/b is a zero of $f(x)$, then a_0 is a multiple of a and a_n is a multiple of b. In our case we would have

$$2 = ka \tag{39}$$

$$-1 = k'b \tag{40}$$

the values for k and k' which give a and b integers are $k = 1, 2$ and $k' = 1$. Then the possible values of a which are multiples of a_0 are $a = \pm 2, \pm 1$. While for $b = \pm 1$. Then, the possible roots of (38) of the form a/b are

$$\frac{a}{b} = \pm 2, \pm 1 \tag{41}$$

The next step is just try each one of the a/b possibility.

Second we find the eigenvectors by finding the null space of the matrix $A - \lambda_i I$ for each eigenvalue.

Using row reduction for $\lambda = 1$ we get

$$[A - I|\bar{0}] = \left[\begin{array}{rrr|r} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & -5 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{42}$$

Then,

$$\begin{aligned} x_1 - x_3 &= 0 \\ x_2 - x_3 &= 0 \end{aligned} \tag{43}$$

Then, the eigenvectors $\bar{u}$ for $\lambda = 1$ are

$$\bar{u} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} \tag{44}$$

with t in $\mathbb{R}$.

The eigenspase E_1 is

$$E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} = span\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right) \tag{45}$$

Similarly for $\lambda = 2$ we get

$$[A - 2I|\bar{0}] = \left[\begin{array}{rrr|r} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 2 & -5 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{46}$$

Then,

$$\begin{aligned} x_1 - \tfrac{1}{4}x_3 &= 0 \\ x_2 - \tfrac{1}{2}x_3 &= 0 \end{aligned} \tag{47}$$

Then, the eigenvectors $\bar{v}$ for $\lambda = 2$ are

$$\bar{v} = \begin{bmatrix} \frac{1}{4}t \\ \frac{1}{2}t \\ t \end{bmatrix} \tag{48}$$

with t in $\mathbb{R}$.

The eigenspase E_3 is

$$E_3 = \left\{ t \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ 1 \end{bmatrix} \right\} = span\left(\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right) \tag{49}$$

Algebraic multiplicity: The algebraic multiplicity of an eigenvalue is the multiplicity as a root of the characteristic equation. Thus, $\lambda = 1$ has algebraic multiplicity 2 and $\lambda = 2$ has algebraic multiplicity 1: $(\lambda - 1)^2(\lambda - 2)$.

Geometric multiplicity: The geometric multiplicity of an eigenvalue λ is the dimension of its eigenspace, i.e. $dim(E_\lambda)$.

Exercise for the student in class (Example 4.19): Find all of the eigenvalues and the corresponding eigenvectors of the matrix $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$.

Solution:

- $\lambda_1 = \lambda_2 = 0$, *span* $\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right)$

  Notice that any linear combination of eigenvectors $\{\bar{v}_1, \cdots, \bar{v}_k\}$ of a given eigenspace E_λ is also an eigenvector of the matrix A with the same eigenvalue λ, since

$$\begin{aligned} A\bar{x} &= A(c_1\bar{v}_1 + \cdots + c_k\bar{v}_k) && (50) \\ &= c_1 A\bar{v}_1 + \cdots + c_k A\bar{v}_k && (51) \\ &= c_1 \lambda\bar{v}_1 + \cdots + c_k \lambda\bar{v}_k && (52) \\ &= \lambda(c_1\bar{v}_1 + \cdots + c_k\bar{v}_k) && (53) \\ A\bar{x} &= \lambda\bar{x} && (54) \end{aligned}$$

- $\lambda_3 = -2$, *span* $\left(\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right)$

In terms of the multiplicity we have

- The algebraic multiplicity of $\lambda = 0$ is 2 and its geometric multiplicity is also 2.
- The algebraic multiplicity of $\lambda = -2$ is 1 and its geometric multiplicity is also 1.

Theorem 4.15: The eigenvalues of a triangular matrix are the entries on its main diagonal.

Theorem 4.16: A square matrix A is invertible if and only if 0 is not an eigenvalue of A.
Proof: Let A be a square matrix. By Theorem 4.6, A is invertible if and only if $\det A \neq 0$. But $\det A \neq 0$ is equivalent to $\det(A - 0I) \neq 0$, which says that 0 is not a root of the characteristic equation of A, i.e. 0 is not and eigenvalue of A.

Theorem 4.17: Fundamental Theorem (FT) of Invertible Matrices. Version 3 of 5
Let A be an $n \times n$ matrix. The following statements are equivalent:

From Version 1

a. A is invertible.

b. $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b}$ in $\mathbb{R}^n$.

c. $A\bar{x} = 0$ has only the trivial solution.

d. The reduced row echelon form of A is I_n.

e. A is a product of elementary matrices.

From Version 2

f. rank(A)=n

g. nullity(A)=0

h. The column vectors of A are LI

i. The column vectors of A span $\mathbb{R}^n$

j. The column vectors of A form a basis for $\mathbb{R}^n$

k. The row vectors of A are LI

l. The row vectors of A span $\mathbb{R}^n$

m. The row vectors of A form a basis for $\mathbb{R}^n$

New statements

n. $detA \neq 0$

o. 0 is not an eigenvalue of A

Proof: The equivalence (a)⇔(n) is Theorem 4.6. Theorem 4.16 proves (a)⇔(o).

Theorem 4.18: Let A be a square matrix with eigenvalue λ and corresponding eigenvector $\bar{x}$.

a. For any positive integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector $\bar{x}$.

b. If A is invertible, then $1/\lambda$ is an eigenvalue of A^{-1} with corresponding eigenvector $\bar{x}$.

c. For any integer n (la diferencia con el punto (a) es que pedía que n sea positivo), λ^n is an eigenvalue of A^n with corresponding eigenvector $\bar{x}$.

Proof: See book, pag. 293.

Application of the theorem 4.18: Calculate the action of A^3 in the vector $\bar{u}$ with, $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ and $\bar{u} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

The strategy consist in expanding the vector $\bar{u}$ in the basis (if it exist) generated by the eigenspaces of the eigenvalues of A.

Then, we first calculate the eigenvalues and get $\lambda_1 = -1$ and $\lambda_2 = 2$, with eigenvectors $\bar{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\bar{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, respectively.

Next, we express the given vector $\bar{u}$ as linear combination of the eigenvectors $\bar{v}_1$ and $\bar{v}_2$:

$$\bar{u} = c_1\bar{v}_1 + c_2\bar{v}_2 \tag{55}$$

which gives for the coefficients: $c_1 = 3$ and $c_2 = 2$.

Finally, we apply the matrix A^3 and use the fact that it is a linear operator and that $\bar{v}_i$ are eigenvectors of A,

$$\begin{aligned}
A^3\bar{u} &= A^3(3\bar{v}_1) + A^3(2\bar{v}_2) && (56)\\
&= 3\lambda_1^3\bar{v}_1 + 2\lambda_1^3\bar{v}_2 && (57)\\
&= 3(-1)^3\bar{v}_1 + 2(2)^3\bar{v}_2 && (58)\\
&= -3\bar{v}_1 + 16\bar{v}_2 && (59)\\
&= -3\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 16\begin{bmatrix} 1 \\ 2 \end{bmatrix} && (60)\\
&= \begin{bmatrix} 13 \\ 35 \end{bmatrix} && (61)
\end{aligned}$$

This result should be the same that the one obtained by multiplying three times the matrix A with it self and the resulting matrix multiplied by the vector $\bar{u}$. **Check it!**. You will need the explicit matrix A^3 below.

Theorem 4.19: Suppose the $n \times n$ matrix A has eigenvectors $\bar{v}_1, \cdots, \bar{v}_m$ with corresponding eigenvalues $\lambda_1, \cdots, \lambda_m$. If $\bar{x}$ is a vector in $\mathbb{R}^n$ that can be expressed as a linear combination of these eigenvectors,

$$\bar{x} = c_1\bar{v}_1 + \cdots + c_m\bar{v}_m \tag{62}$$

then, for any integer k,

$$A^k\bar{x} = c_1\lambda_1^k\bar{v}_1 + \cdots + c_m\lambda_m^k\bar{v}_m \tag{63}$$

Theorem 4.20: Let A be an $n \times n$ matrix and let $\lambda_1, \cdots, \lambda_m$ be distinct eigenvalues of A with corresponding eigenvectors $\bar{v}_1, \cdots, \bar{v}_m$. Then $\bar{v}_1, \cdots, \bar{v}_m$ are LI.
Proof: See book, pag. 295.

Similarity and Diagonalization

Triangular and diagonal matrices expose their eigenvalues explicitly. We wish to relate a given matrix with its triangular or diagonal form. Since the Gauss elimination does not preserve the eigenvalue of the matrix another procedure is called for. This is the goal of this section.

Similar Matrices

Let A and B be $n \times n$ matrices. We say that A is similar to B if there is an invertible $n \times n$ matrix P such that

$$P^{-1}AP = B \tag{64}$$

and write

$$A \sim B \tag{65}$$

Remarks:

- If $A \sim B$, we can write, equivalently, that $A = PBP^{-1}$ or $AP = PB$.
- Similarity is a relation on square matrices.
- There is a direction (or order) implicit in the definition of similarity. We should not assume that $A \sim B$ implies $B \sim A$ (even when it is true) since it does not follow immediately from the definition.
- The matrix P depends on A and B.
- The matrix P is not unique for a given pair of similar matrices A and B. For example, if $A = B = I$ then $I \sim I$ for any invertible matrix P, because $P^{-1}IP = P^{-1}P = I$.

Theorem 4.21: Equivalence relation. Any relation satisfying the following three properties is called an equivalence relation. Let A, B and C be $n \times n$ matrices,

a. $A \sim A$

b. If $A \sim B$, then $B \sim A$.

c. If $A \sim B$ and $B \sim C$, then $A \sim C$.

Proof:

a. $I^{-1}AI = A$

b. If $A \sim B$, then $P^{-1}AP = B \Rightarrow A = PBP^{-1}$. Let us renamed $Q = P^{-1}$, then $Q^{-1}BQ = A$, so $B \sim A$.

c. Exercise 30.

Theorem 4.22: Let A and B be $n \times n$ matrices with $A \sim B$. Then

a. $detA = detB$

b. A is invertible if and only if B is invertible.

c. A and B have the same rank

d. A and B have the same characteristic polynomial.

e. A and B have the same eigenvalues.

Proof: See book, pag. 299.

Remarks:

1. Two matrices may have properties (a) through (e) (and more) in common and yet still not be similar.

2. Theorem 4.22 is most useful in showing that two matrices are not similar, since A and B cannot be similar if any of properties (a) through (e) fails.

Example of the remark 1: Be $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$:

a. $detA = detB = 1$

b. A and B both are invertible

c. $rankA = rankB = 2$

d. The characteristic polynomial of A and B is the same: $(1-\lambda)^2$

e. The eigenvalues of A and B are the same: $\lambda_1 = \lambda_2 = 1$

But, $P^{-1}AP = P^{-1}P$ (since $A = I$), then $P^{-1}AP = I \neq B$ for any invertible matrix P.

Example of the remark 2: The following two matrices have the same determinant and rank but they have not the same characteristic polynomial. $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$:

- Characteristic polynomial of A: $\lambda^2 - 3\lambda - 4$.

- Characteristic polynomial of B: $\lambda^2 - 4$.

Check it!

Diagonalization

The best possible situation is when a square matrix is similar to a diagonal matrix. Whether a matrix is diagonalizable is closely related to the eigenvalues and eigenvectors of the matrix. This is the topic of this section.

An $n \times n$ matrix A is diagonalizable if there is a diagonal matrix D such that $A \sim D$, i.e. if there is an invertible $n \times n$ matrix P such that $P^{-1}AP = D$. The entries of the matrices D and P are related to the eigenvalues and eigenvectors, respectively of A.

Theorem 4.23: Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if A has n linearly independent eigenvectors.
More precisely, there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ is and only if the columns of P are n LI eigenvectors of A and the diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors in P in the same order.
Proof: See book, pag. 301.

Example 1 of the Theorem 4.23: The matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$ has only two LI eigenvectors (it was calculate previously, see Example 4.18). Therefore, A is not diagonalizable

Example 2 of the Theorem 4.23: The matrix $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$ has three LI (check it!) eigenvectors (it was calculate previously, see example 4.19) with eigenvalues $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = -2$:

$$\bar{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \bar{p}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \bar{p}_3 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \tag{66}$$

If we take

$$P = [\bar{p}_1\ \bar{p}_2\ \bar{p}_3] = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{bmatrix} \tag{67}$$

then P is invertible (find the inverse!). Furthermore,

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = D \tag{68}$$

About the order in P and D: The eigenvectors can be placed into the columns of P in any order. However, the eigenvalues will come up on the diagonal of D in the same order as their corresponding eigenvectors in P.

Theorem 4.24: Let A be an $n \times n$ matrix and let $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues of A. If $\mathcal{B}_i$ is a basis for the eigenspace E_{λ_i}, then $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is LI.
Proff: See book, pag. 303.

Example of Theorem 4.24: In the 'Example 2 of the Theorem 4.23' above we were asked to demonstrate the independence of the three eigenvectors $\bar{p}_1, \bar{p}_2, \bar{p}_3$. It was unnecessarily since,

- $\bar{p}_1, \bar{p}_2$ is a basis for the eigenspace $E_{\lambda=0}$, then they are LI
- $\bar{p}_3$ is a basis for the eigenspace $E_{\lambda=-2}$
- Then, from Theorem 4.24 the vectors which result from the union of $E_{\lambda=0}$ and $E_{\lambda=-2}$ are LI.

Theorem 4.25: If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.
Proof: Let $\bar{v}_1, \cdots, \bar{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of A. By Theorem 4.20, $\bar{v}_1, \cdots, \bar{v}_n$ are LI, so, by Theorem 4.23, A is diagonalizable.

Exercise for the student in class: Find the matrix P which transform the matrix A to its diagonal form, with $A = \begin{bmatrix} 2 & -3 & 7 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{bmatrix}$

Solution: pending...

Lemma 4.26: If A is an $n \times n$ matrix, then the geometric multiplicity of each eigenvalue is less than or equal to its algebraic multiplicity.
Proof: See book, pag. 304.

Theorem 4.27: The Diagonalization Theorem. Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \cdots, \lambda_k$. The following statements are equivalent:

a. A is diagonalizable

b. The union $\mathcal{B}$ of the bases of the eigenspaces of A contains n vectors.

c. The algebraic multiplicity of each eigenvalue equals its geometric multiplicity.

Proof: See book, pag. 304.

Example 1 to the Diagonalization Theorem: The matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$ has two distinct eigenvalues (see Example 4.18) $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$. Since the algebraic multiplicity of $\lambda = 1$ is 2 while the geometric multiplicity 1 (the eigenspace contain only one vector), the matrix A is not diagonalizable.

Example 2 to the Diagonalization Theorem: The matrix $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$ has two distinct eigenvalues (see Example 4.19) $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = -2$.Since the geometric multiplicity of $\lambda = 0$ is also 2 (as its algebraic multiplicity), the eigenvalue $\lambda = -2$ has algebraic and geometric multiplicity 1, the matrix A is diagonalizable.

Application of the diagonalization $P^{-1}AP = D$: Let us assume we want a power of a given matrix A of dimension n, let us say A^k. Let us write

$$P^{-1}AP = D \Rightarrow A = PDP^{-1} \tag{69}$$

with D diagonal

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \tag{70}$$

with λ_i the eigenvalues of A. Then

$$A^2 = (PDP^{-1})^2 = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1} \tag{71}$$

in general

$$A^k = PD^kP^{-1} \tag{72}$$

with

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix} \tag{73}$$

Example: Calculates A^3 where $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

From some previous calculation we know the eigenvalues and eigenvectors:

- $\lambda_1 = -1$, $\bar{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- $\lambda_2 = 2$, $\bar{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

then

$$P = [\bar{v}_1 \ \bar{v}_2] = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \tag{74}$$

and (using the Theorem for the inverse of a 2 by 2 matrix)

$$P^{-1} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \tag{75}$$

and

$$D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \tag{76}$$

The application of $A^3 = PD^3P^{-1}$ gives,

$$\begin{aligned} A^3 &= \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} (-1)^3 & 0 \\ 0 & (2)^3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} && (77) \\ &= \textit{(pending)} && (78) \end{aligned}$$

Compare the result with the previous calculation of A^3.

Application: exponential of a matrix. By analogy with the power series expansion of $e^x = 1 + x + \frac{x^2}{2!} + \cdots$, let us define the exponential of a matrix A in terms of powers of A,

$$e^A = I + A + \frac{A^2}{2!} + \cdots \tag{79}$$

with A an square matrix. It can be shown (it is not demostrated in the book) that the series converges for any real matrix A.

For practical calculation, let us assume that the matrix A is diagonalizable,

$$P^{-1}AP = D \tag{80}$$

with D a diagonal matrix, and $A = PDP^{-1}$, then

$$\begin{aligned}
e^A &= I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots && (81)\\
&= I + (PDP^{-1}) + \frac{(PDP^{-1})^2}{2!} + \frac{(PDP^{-1})^3}{3!} + \cdots && (82)\\
&= PIP^{-1} + PDP^{-1} + \frac{PD^2P^{-1}}{2!} + \frac{PD^3P^{-1}}{3!} + \cdots && (83)\\
&= P\left(I + D + \frac{D^2}{2!} + \frac{D^3}{3!} + \cdots\right)P^{-1} && (84)
\end{aligned}$$

where

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} \tag{85}$$

then,

$$\begin{aligned}
e^A &= P\left(I + D + \frac{D^2}{2!} + \frac{D^3}{3!} + \cdots\right)P^{-1} && (86)\\
&= P\left(\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\right. \\
&\quad \left. + \frac{1}{2!}\begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n^3 \end{bmatrix} + \cdots \right)P^{-1}
\end{aligned} \tag{87}$$

$$e^A = P\begin{bmatrix} 1 + \lambda_1 + \frac{1}{2!}\lambda_1^2 + \cdots & 0 & \cdots & 0 \\ 0 & 1 + \lambda_2 + \frac{1}{2!}\lambda_2^2 + \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 + \lambda_n + \frac{1}{2!}\lambda_n^2 + \cdots \end{bmatrix}P^{-1} \tag{88}$$

$$e^A = P \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & e^{\lambda_n} \end{bmatrix} P^{-1} \tag{89}$$

where λ_i are the eigenvalues of A and P is the matrix which has as columns the eigenvectors of A in the same order as the eigenvalues is the matrix D.