Chapter 8
The Characteristic Polynomial and Eigenvalues of Matrices

We have already characterized matrices using their rank and their determinant. In this chapter we use the determinant map in order to assign to every square matrix a unique polynomial that is called the characteristic polynomial of the matrix. This polynomial contains important information about the matrix. For example, one can read off the determinant and thus see whether the matrix is invertible. Even more important are the roots of the characteristic polynomial, which are called the eigenvalues of the matrix.

8.1 The Characteristic Polynomial and the Cayley-Hamilton Theorem

Let R be a commutative ring with unit and let $R[t]$ be the corresponding ring of polynomials (cp. Example 3.17). For $A = [a_{ij}] \in R^{n,n}$ we set

$$tI_n - A := \begin{bmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & t - a_{nn} \end{bmatrix} \in (R[t])^{n,n}.$$

The entries of the matrix $tI_n - A$ are elements of the commutative ring with unit $R[t]$, where the diagonal entries are polynomials of degree 1, and the other entries are constant polynomials. Using Definition 7.4 we can form the determinant of the matrix $tI_n - A$, which is an element of $R[t]$.

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-24346-7_8

Definition 8.1 Let R be a commutative ring with unit and $A \in R^{n,n}$. Then

$$P_A := \det(tI_n - A) \;\in\; R[t]$$

is called the *characteristic polynomial* of A.

Example 8.2 If $n = 1$ and $A = [a_{11}]$, then

$$P_A = \det(tI_1 - A) = \det([t - a_{11}]) = t - a_{11}.$$

For $n = 2$ and

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

we obtain

$$P_A = \det\left(\begin{bmatrix} t - a_{11} & -a_{12} \\ -a_{21} & t - a_{22} \end{bmatrix}\right) = t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}).$$

Using Definition 7.4 we see that the general form of P_A for a matrix $A \in R^{n,n}$ is given by

$$P_A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i,\sigma(i)} t - a_{i,\sigma(i)}\right). \tag{8.1}$$

The following lemma presents basic properties of the characteristic polynomial.

Lemma 8.3 *For $A \in R^{n,n}$ we have $P_A = P_{A^T}$ and*

$$P_A = t^n - \alpha_{n-1} t^{n-1} + \ldots + (-1)^{n-1} \alpha_1 t + (-1)^n \alpha_0$$

with $\alpha_{n-1} = \sum_{i=1}^{n} a_{ii}$ and $\alpha_0 = \det(A)$.

Proof Using (5) in Lemma 7.10 we obtain

$$P_A = \det(tI_n - A) = \det((tI_n - A)^T) = \det(tI_n - A^T) = P_{A^T}.$$

Using P_A as in (8.1) we see that

$$P_A = \prod_{i=1}^{n} (t - a_{ii}) \;+ \sum_{\substack{\sigma \in S_n \\ \sigma \neq [1 \cdots n]}} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i,\sigma(i)} t - a_{i,\sigma(i)}\right).$$

The first term on the right hand side is of the form

$$t^n - \left(\sum_{i=1}^{n} a_{ii}\right) t^{n-1} + \text{ (polynomial of degree} \leq n-2),$$

and the second term is a polynomial of degree $\leq n - 2$. Thus, $\alpha_{n-1} = \sum_{i=1}^{n} a_{ii}$ as claimed. Moreover, Definition 8.1 yields

$$P_A(0) = \det(-A) = (-1)^n \det(A),$$

so that $\alpha_0 = \det(A)$. □

This lemma shows that the characteristic polynomial of $A \in R^{n,n}$ always is of degree n. The coefficient of t^n is $1 \in R$. Such a polynomial is called *monic*. The coefficient of t^{n-1} is given by the sum of the diagonal entries of A. This quantity is called the *trace* of A, i.e.,

$$\text{trace}(A) := \sum_{i=1}^{n} a_{ii}.$$

The following lemma shows that for every monic polynomial $p \in R[t]$ of degree $n \geq 1$ there exists a matrix $A \in R^{n,n}$ with $P_A = p$.

Lemma 8.4 *If $n \in \mathbb{N}$ and $p = t^n + \beta_{n-1}t^{n-1} + \ldots + \beta_0 \in R[t]$, then p is the characteristic polynomial of the matrix*

$$A = \begin{bmatrix} 0 & & & -\beta_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -\beta_{n-2} \\ & & 1 & -\beta_{n-1} \end{bmatrix} \in R^{n,n}.$$

(For $n = 1$ we have $A = [-\beta_0]$.) The matrix A is called the companion matrix *of p.*

Proof We prove the assertion by induction on n.

For $n = 1$ we have $p = t + \beta_0$, $A = [-\beta_0]$ and $P_A = \det([t + \beta_0]) = p$.

Let the assertion hold for some $n \geq 1$. We consider $p = t^{n+1} + \beta_n t^n + \ldots + \beta_0$ and

$$A = \begin{bmatrix} 0 & & & -\beta_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -\beta_{n-1} \\ & & 1 & -\beta_n \end{bmatrix} \in R^{n+1,n+1}.$$

Using the Laplace expansion with respect to the first row (cp. Corollary 7.22) and the induction hypothesis we get

$$\begin{aligned}
P_A &= \det(tI_{n+1} - A) \\
&= \det\left(\begin{bmatrix} t & & & \beta_0 \\ -1 & \ddots & & \vdots \\ & \ddots & t & \beta_{n-1} \\ & & -1 & t+\beta_n \end{bmatrix}\right) \\
&= t \cdot \det\left(\begin{bmatrix} t & & & \beta_1 \\ -1 & \ddots & & \vdots \\ & \ddots & t & \beta_{n-1} \\ & & -1 & t+\beta_n \end{bmatrix}\right) + (-1)^{n+2} \cdot \beta_0 \cdot \det\left(\begin{bmatrix} -1 & t & & \\ & \ddots & \ddots & \\ & & \ddots & t \\ & & & -1 \end{bmatrix}\right) \\
&= t \cdot (t^n + \beta_n t^{n-1} + \ldots + \beta_1) + (-1)^{2n+2}\beta_0 \\
&= t^{n+1} + \beta_n t^n + \ldots + \beta_1 t + \beta_0 \\
&= p.
\end{aligned}$$

□

Example 8.5 The polynomial $p = (t-1)^3 = t^3 - 3t^2 + 3t - 1 \in \mathbb{Z}[t]$ has the companion matrix

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix} \in \mathbb{Z}^{3,3}.$$

The identity matrix I_3 has the characteristic polynomial

$$P_{I_3} = \det(tI_3 - I_3) = (t-1)^3 = P_A.$$

Thus, different matrices may have the same characteristic polynomial.

In Example 3.17 we have seen how to evaluate a polynomial $p \in R[t]$ at a scalar $\lambda \in R$. Analogously, we can evaluate p at a matrix $M \in R^{m,m}$ (cp. Exercise 4.8). For

$$p = \beta_n t^n + \beta_{n-1} t^{n-1} + \ldots + \beta_0 \in R[t]$$

we define

$$p(M) := \beta_n M^n + \beta_{n-1} M^{n-1} + \ldots + \beta_0 I_m \in R^{m,m},$$

where the multiplication on the right hand side is the scalar multiplication of $\beta_j \in R$ and $M^j \in R^{m,m}$, $j = 0, 1, \ldots, n$. (Recall that $M^0 = I_m$.) Evaluating a given polynomial at matrices $M \in R^{m,m}$ therefore defines a map from $R^{m,m}$ to $R^{m,m}$.

In particular, using (8.1), the characteristic polynomial P_A of $A \in R^{n,n}$ satisfies

$$P_A(M) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i,\sigma(i)} M - a_{i,\sigma(i)} I_m\right) \quad \text{for all } M \in R^{m,m}.$$

Note that for $M \in R^{n,n}$ and $P_A = \det(tI_n - A)$ the "obvious" equation $P_A(M) = \det(M - A)$ is *wrong*. By definition, $P_A(M) \in R^{n,n}$ and $\det(M - A) \in R$, so that the two expressions cannot be the same, even for $n = 1$.

The following result is called the *Cayley-Hamilton theorem*.[1]

Theorem 8.6 *For every matrix $A \in R^{n,n}$ and its characteristic polynomial $P_A \in R[t]$ we have $P_A(A) = 0 \in R^{n,n}$.*

Proof For $n = 1$ we have $A = [a_{11}]$ and $P_A = t - a_{11}$, so that $P_A(A) = [a_{11}] - [a_{11}] = [0]$.

Let now $n \geq 2$ and let e_i be the ith column of the identity matrix $I_n \in R^{n,n}$. Then

$$Ae_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, \quad i = 1, \ldots, n,$$

which is equivalent to

$$(A - a_{ii}I_n)e_i + \sum_{\substack{j=1\\ j\neq i}}^{n} (-a_{ji}I_n)e_j = 0, \quad i = 1, \ldots, n.$$

The last n equations can be written as

$$\begin{bmatrix} A - a_{11}I_n & -a_{21}I_n & \cdots & -a_{n1}I_n \\ -a_{12}I_n & A - a_{22}I_n & \cdots & -a_{n2}I_n \\ \vdots & \vdots & & \vdots \\ -a_{1n}I_n & -a_{2n}I_n & \cdots & A - a_{nn}I_n \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{or} \quad B\varepsilon = \widehat{0}.$$

Hence $B \in (R[A])^{n,n}$ with $R[A] := \{p(A) \mid p \in R[t]\} \subset R^{n,n}$. The set $R[A]$ forms a commutative ring with unit given by the identity matrix I_n (cp. Exercise 4.8). Using Theorem 7.18 we obtain

$$\operatorname{adj}(B)B = \det(B)\widehat{I_n},$$

[1] Arthur Cayley (1821–1895) showed this theorem in 1858 for $n = 2$ and claimed that he had verified it for $n = 3$. He did not feel it necessary to give a proof for general n. Sir William Rowan Hamilton (1805–1865) proved the theorem for the case $n = 4$ in 1853 in the context of his investigations of quaternions. One of the first proofs for general n was given by Ferdinand Georg Frobenius (1849–1917) in 1878. James Joseph Sylvester (1814–1897) coined the name of the theorem in 1884 by calling it the "no-little-marvelous Hamilton-Cayley theorem".

where $\det(B) \in R[A]$ and $\widehat{I_n}$ is the identity matrix in $(R[A])^{n,n}$. (This matrix has n times the identity matrix I_n on its diagonal.) Multiplying this equation from the right by ε yields

$$\operatorname{adj}(B)B\varepsilon = \det(B)\widehat{I_n}\varepsilon,$$

which implies that $\det(B) = 0 \in R^{n,n}$. Finally, using Lemma 8.3 gives

$$\begin{aligned} 0 = \det(B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (\delta_{i,\sigma(i)} A - a_{\sigma(i),i} I_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (\delta_{\sigma(i),i} A - a_{\sigma(i),i} I_n) \\ &= P_{A^T}(A) \\ &= P_A(A), \end{aligned}$$

which completes the proof. □

8.2 Eigenvalues and Eigenvectors

In this section we present an introduction to the topic of eigenvalues and eigenvectors of square matrices over a field K. These concepts will be studied in more detail in later chapters.

Definition 8.7 Let $A \in K^{n,n}$. If $\lambda \in K$ and $v \in K^{n,1} \setminus \{0\}$ satisfy $Av = \lambda v$, then λ is called an *eigenvalue* of A and v is called an *eigenvector* of A corresponding to λ.

While by definition $v = 0$ can never be an eigenvector of a matrix, $\lambda = 0$ may be an eigenvalue. For example,

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

If v is an eigenvector corresponding to the eigenvalue λ of A and $\alpha \in K \setminus \{0\}$, then $\alpha v \neq 0$ and

$$A\,(\alpha v) = \alpha\,(Av) = \alpha\,(\lambda v) = \lambda\,(\alpha v).$$

Thus, also αv is an eigenvector of A corresponding to λ.

Theorem 8.8 *For $A \in K^{n,n}$ the following assertions hold:*

(1) λ is an eigenvalue of A if and only if λ is a root of the characteristic polynomial of A, i.e., $P_A(\lambda) = 0 \in K$.
(2) $\lambda = 0$ is an eigenvalue of A if and only if $\det(A) = 0$.
(3) λ is an eigenvalue of A if and only if λ is an eigenvalue of A^T.

Proof

(1) The equation $P_A(\lambda) = \det(\lambda I_n - A) = 0$ holds if and only if the matrix $\lambda I_n - A$ is not invertible (cp. (7.4)), and this is equivalent to $\mathscr{L}(\lambda I_n - A, 0) \neq \{0\}$. This, however, means that there exists a vector $\widehat{x} \neq 0$ with $(\lambda I_n - A)\widehat{x} = 0$, or $A\widehat{x} = \lambda\widehat{x}$.
(2) By (1), $\lambda = 0$ is an eigenvalue of A if and only if $P_A(0) = 0$. The assertion now follows from $P_A(0) = (-1)^n \det(A)$ (cp. Lemma 8.3).
(3) This follows from (1) and $P_A = P_{A^T}$ (cp. Lemma 8.3). □

Whether a matrix $A \in K^{n,n}$ has eigenvalues or not may depend on the field K over which A is considered.

Example 8.9 The matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{R}^{2,2}$$

has the characteristic polynomial $P_A = t^2 + 1 \in \mathbb{R}[t]$. This polynomial does not have roots, since the equation $t^2 + 1 = 0$ has no (real) solutions. If we consider A as an element of $\mathbb{C}^{2,2}$, then $P_A \in \mathbb{C}[t]$ has the roots $\mathbf{i}$ and $-\mathbf{i}$. Then these two complex numbers are the eigenvalues of A.

Item (3) in Theorem 8.8 shows that A and A^T have the same eigenvalues. An eigenvector of A, however, may not be an eigenvector of A^T.

Example 8.10 The matrix

$$A = \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2,2}$$

has the characteristic polynomial $P_A = t^2 - 4t = t \cdot (t - 4)$, and hence its eigenvalues are 0 and 4. We have

$$A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad A^T \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \neq \lambda \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

for all $\lambda \in \mathbb{R}$. Thus, $[1, \ -1]^T$ is an eigenvector of A corresponding to the eigenvalue 0, but it is not an eigenvector of A^T. On the other hand,

$$A^T \begin{bmatrix} 1 \\ -3 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -3 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} -6 \\ -2 \end{bmatrix} \neq \lambda \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

for all $\lambda \in \mathbb{R}$. Thus, $[1, \ -3]^T$ is an eigenvector of A^T corresponding to the eigenvalue 0, but it is not an eigenvector of A.

Theorem 8.8 implies further criteria for the invertibility of $A \in K^{n,n}$ (cp. (7.3)):

$$\begin{aligned} A \in GL_n(K) &\Leftrightarrow 0 \text{ is not an eigenvalue of } A \\ &\Leftrightarrow 0 \text{ is not a root of } P_A. \end{aligned}$$

Definition 8.11 Two matrices $A, B \in K^{n,n}$ are called *similar*, if there exists a matrix $Z \in GL_n(K)$ with $A = ZBZ^{-1}$.

One can easily show that this defines an equivalence relation on the set $K^{n,n}$ (cp. the proof following Definition 5.13).

Theorem 8.12 *If two matrices $A, B \in K^{n,n}$ are similar, then $P_A = P_B$.*

Proof If $A = ZBZ^{-1}$, then the multiplication theorem for determinants yields

$$\begin{aligned} P_A &= \det(tI_n - A) = \det(tI_n - ZBZ^{-1}) = \det(Z(tI_n - B)Z^{-1}) \\ &= \det(Z)\det(tI_n - B)\det(Z^{-1}) = \det(tI_n - B)\ \det(ZZ^{-1}) \\ &= P_B \end{aligned}$$

(cp. the remarks below Theorem 7.15). □

Theorem 8.12 and (1) in Theorem 8.8 show that two similar matrices have the same eigenvalues. The condition that A and B are similar is sufficient, but not necessary for $P_A = P_B$.

Example 8.13 Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Then $P_A = (t-1)^2 = P_B$, but for every matrix $Z \in GL_n(K)$ we have $ZBZ^{-1} = I_2 \neq A$. Thus, we have $P_A = P_B$ although A and B are not similar (cp. also Example 8.5).

MATLAB-Minute.
The roots of a polynomial $p = \alpha_n t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0$ can be computed (or approximated) in MATLAB using the command `roots(p)`, where `p` is a $1 \times (n+1)$ matrix with the entries `p(i)`$= \alpha_{n+1-i}$ for $i = 1, \ldots, n+1$. Compute `roots(p)` for the monic polynomial $p = t^3 - 3t^2 + 3t - 1 \in \mathbb{R}[t]$ and display the output using `format long`. What are the exact roots of p and how large is the numerical error in the computation of the roots using `roots(p)`?
Form the matrix `A=compan(p)` and compare its structure with the one of the companion matrix from Lemma 8.4. Can you transfer the proof of Lemma 8.4 to the structure of the matrix `A`?
Compute the eigenvalues of `A` with the command `eig(A)` and compare the output with the one of `roots(p)`. What do you observe?

8.3 Eigenvectors of Stochastic Matrices

We now consider the eigenvalue problem presented in Sect. 1.1 in the context of the PageRank algorithm. The mathematical modeling leads to the equations (1.1), which can be written in the form $Ax = x$. Here $A = [a_{ij}] \in \mathbb{R}^{n,n}$ (n is the number of documents) satisfies

$$a_{ij} \geq 0 \quad \text{and} \quad \sum_{i=1}^{n} a_{ij} = 1 \quad \text{for} \quad j = 1, \ldots, n.$$

Such a matrix A is called *column-stochastic*. Note that A is column-stochastic if and only if A^T is row-stochastic. Such matrices also occurred in the car insurance application considered in Sect. 1.2 and Example 4.7. We want to determine $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^{n,1} \setminus \{0\}$ with $Ax = x$, where the entry x_i describes the importance of document i. The importance values should be nonnegative, i.e., $x_i \geq 0$ for $i = 1, \ldots, n$. Thus, we want to determine an entrywise nonnegative eigenvector of A corresponding to the eigenvalue $\lambda = 1$.

We first check whether this problem has a solution, and then study whether the solution is unique. Our presentation is based on the article [BryL06].

Lemma 8.14 *A column-stochastic matrix $A \in \mathbb{R}^{n,n}$ has an eigenvector corresponding to the eigenvalue 1.*

Proof Since A is column-stochastic, we have $A^T[1, \ldots, 1]^T = [1, \ldots, 1]^T$, so that 1 is an eigenvalue of A^T. Now (3) in Theorem 8.8 shows that also A has the eigenvalue 1, and hence there exists a corresponding eigenvector. □

A matrix with real entries is called *positive*, if all its entries are positive.

Lemma 8.15 *If $A \in \mathbb{R}^{n,n}$ is positive and column-stochastic and if $x \in \mathbb{R}^{n,1}$ is an eigenvector of A corresponding to the eigenvalue 1, then either x or $-x$ is positive.*

Proof If $x = [x_1, \ldots, x_n]^T$ is an eigenvector of $A = [a_{ij}]$ corresponding to the eigenvalue 1, then

$$x_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, n.$$

Suppose that not all entries of x are positive or not all entries of x are negative. Then there exists at least one index k with

$$|x_k| = \Big| \sum_{j=1}^{n} a_{kj}\, x_j \Big| < \sum_{j=1}^{n} a_{kj}\, |x_j|,$$

which implies

$$\sum_{i=1}^{n} |x_i| < \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, |x_j| = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}\, |x_j| = \sum_{j=1}^{n} \Big(|x_j| \cdot \underbrace{\sum_{i=1}^{n} a_{ij}}_{=1} \Big) = \sum_{j=1}^{n} |x_j|.$$

This is impossible, so that indeed x or $-x$ must be positive. □

We can now prove the following uniqueness result.

Theorem 8.16 *If $A \in \mathbb{R}^{n,n}$ is positive and column-stochastic, then there exists a unique positive $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^{n,1}$ with $\sum_{i=1}^{n} x_i = 1$ and $Ax = x$.*

Proof By Lemma 8.15, A has a least one positive eigenvector corresponding to the eigenvalue 1. Suppose that $x^{(1)} = \big[x_1^{(1)}, \ldots, x_n^{(1)}\big]^T$ and $x^{(2)} = \big[x_1^{(2)}, \ldots, x_n^{(2)}\big]^T$ are two such eigenvectors. Suppose that these are normalized by $\sum_{i=1}^{n} x_i^{(j)} = 1$, $j = 1, 2$. This assumption can be made without loss of generality, since every nonzero multiple of an eigenvector is still an eigenvector.

We will show that $x^{(1)} = x^{(2)}$. For $\alpha \in \mathbb{R}$ we define $x(\alpha) := x^{(1)} + \alpha x^{(2)} \in \mathbb{R}^{n,1}$, then

$$Ax(\alpha) = Ax^{(1)} + \alpha A x^{(2)} = x^{(1)} + \alpha x^{(2)} = x(\alpha).$$

If $\widetilde{\alpha} := -x_1^{(1)}/x_1^{(2)}$, then the first entry of $x(\widetilde{\alpha})$ is equal to zero and thus, by Lemma 8.15, $x(\widetilde{\alpha})$ cannot be an eigenvector of A corresponding to the eigenvalue 1. Now $Ax(\widetilde{\alpha}) = x(\widetilde{\alpha})$ implies that $x(\widetilde{\alpha}) = 0$, and hence

$$x_i^{(1)} + \widetilde{\alpha} x_i^{(2)} = 0, \quad i = 1, \ldots, n. \tag{8.2}$$

Summing up these n equations yields

$$\underbrace{\sum_{i=1}^{n} x_i^{(1)}}_{=1} + \widetilde{\alpha} \underbrace{\sum_{i=1}^{n} x_i^{(2)}}_{=1} = 0,$$

so that $\widetilde{\alpha} = -1$. From (8.2) we get $x_i^{(1)} = x_i^{(2)}$ for $i = 1, \ldots, n$, and therefore $x^{(1)} = x^{(2)}$. □

The unique positive eigenvector x in Theorem 8.16 is called the *Perron eigenvector*[2] of the positive matrix A. The theory of eigenvalues and eigenvectors of positive (or more general nonnegative) matrices is an important area of Matrix Theory, since these matrices arise in many applications.

By construction, the matrix $A \in \mathbb{R}^{n,n}$ in the PageRank algorithm is column-stochastic but not positive, since there are (usually many) entries $a_{ij} = 0$. In order to obtain a uniquely solvable problem one can use the following trick:

Let $S = [s_{ij}] \in \mathbb{R}^{n,n}$ with $s_{ij} = 1/n$. Obviously, S is positive and column-stochastic. For a real number $\alpha \in (0, 1]$ we define the matrix

$$\widehat{A}(\alpha) := (1 - \alpha)A + \alpha S.$$

This matrix is positive and column-stochastic, and hence it has a unique positive eigenvector $\widehat{u}$ corresponding to the eigenvalue 1. We thus have

$$\widehat{u} = \widehat{A}(\alpha)\widehat{u} = (1 - \alpha)A\widehat{u} + \alpha S\widehat{u} = (1 - \alpha)A\widehat{u} + \frac{\alpha}{n}[1, \ldots, 1]^T.$$

For a very large number of documents (e.g. the entire internet) the number α/n is very small, so that $(1 - \alpha)A\widehat{u} \approx \widehat{u}$. Therefore a solution of the eigenvalue problem $\widehat{A}(\alpha)\widehat{u} = \widehat{u}$ for small α potentially gives a good approximation of a $u \in \mathbb{R}^{n,1}$ that satisfies $Au = u$. The practical solution of the eigenvalue problem with the matrix $\widehat{A}(\alpha)$ is a topic of the field of Numerical Linear Algebra.

The matrix S represents a link structure where all document are mutually linked and thus all documents are equally important. The matrix $\widehat{A}(\alpha) = (1 - \alpha)A + \alpha S$ therefore models the following internet "surfing behavior": A user follows a proposed link with the probability $1-\alpha$ and an arbitrary link with the probability α. Originally, Google Inc. used the value $\alpha = 0.15$.

[2]Oskar Perron (1880–1975).

Exercises

(In the following exercises K is an arbitrary field.)

8.1 Determine the characteristic polynomials of the following matrices over $\mathbb{Q}$:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 4 \\ -1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -4 & 0 & 2 \end{bmatrix}.$$

Verify the Cayley-Hamilton theorem in each case by direct computation. Are two of the matrices A, B, C similar?

8.2 Let R be a commutative ring with unit and $n \geq 2$.

(a) Show that for every $A \in GL_n(R)$ there exists a polynomial $p \in R[t]$ of degree at most $n-1$ with $\mathrm{adj}(A) = p(A)$. Conclude that $A^{-1} = q(A)$ holds for a polynomial $q \in R[t]$ of degree at most $n-1$.

(b) Let $A \in R^{n,n}$. Apply Theorem 7.18 to the matrix $tI_n - A \in (R[t])^{n,n}$ and derive an alternative proof of the Cayley-Hamilton theorem from the formula $\det(tI_n - A)\, I_n = (tI_n - A)\,\mathrm{adj}(tI_n - A)$.

8.3 Let $A \in K^{n,n}$ be a matrix with $A^k = 0$ for some $k \in \mathbb{N}$. (Such a matrix is called *nilpotent*.)

(a) Show that $\lambda = 0$ is the only eigenvalue of A.

(b) Determine P_A and show that $A^n = 0$.

(*Hint:* You may assume that P_A has the form $\prod_{i=1}^{n}(t-\lambda_i)$ for some $\lambda_1, \ldots, \lambda_n \in K$.)

(c) Show that $\mu I_n - A$ is invertible if and only if $\mu \in K \setminus \{0\}$.

(d) Show that $(I_n - A)^{-1} = I_n + A + A^2 + \ldots + A^{n-1}$.

8.4 Determine the eigenvalues and corresponding eigenvectors of the following matrices over $\mathbb{R}$:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 8 & 16 \\ 0 & 7 & 8 \\ 0 & -4 & -5 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

Is there any difference when you consider A, B, C as matrices over $\mathbb{C}$?

8.5 Let $n \geq 3$ and $\varepsilon \in \mathbb{R}$. Consider the matrix

$$A(\varepsilon) = \begin{bmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \varepsilon & & & 1 \end{bmatrix}$$

as an element of $\mathbb{C}^{n,n}$ and determine all eigenvalues in dependence of ε. How many pairwise distinct eigenvalues does $A(\varepsilon)$ have?

8.6 Determine the eigenvalues and corresponding eigenvectors of

$$A = \begin{bmatrix} 2 & 2-a & 2-a \\ 0 & 4-a & 2-a \\ 0 & -4+2a & -2+2a \end{bmatrix} \in \mathbb{R}^{3,3}, \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \in (\mathbb{Z}/2\mathbb{Z})^{3,3}.$$

(For simplicity, the elements of $\mathbb{Z}/2\mathbb{Z}$ are here denoted by k instead of $[k]$.)

8.7 Let $A \in K^{n,n}$, $B \in K^{m,m}$, $n \geq m$, and $C \in K^{n,m}$ with $\operatorname{rank}(C) = m$ and $AC = CB$. Show that then every eigenvalue of B is an eigenvalue of A.

8.8 Show the following assertions:

(a) $\operatorname{trace}(\lambda A + \mu B) = \lambda \operatorname{trace}(A) + \mu \operatorname{trace}(B)$ holds for all $\lambda, \mu \in K$ and $A, B \in K^{n,n}$.
(b) $\operatorname{trace}(AB) = \operatorname{trace}(BA)$ holds for all $A, B \in K^{n,n}$.
(c) If $A, B \in K^{n,n}$ are similar, then $\operatorname{trace}(A) = \operatorname{trace}(B)$.

8.9 Prove or disprove the following statements:

(a) There exist matrices $A, B \in K^{n,n}$ with $\operatorname{trace}(AB) \neq \operatorname{trace}(A)\operatorname{trace}(B)$.
(b) There exist matrices $A, B \in K^{n,n}$ with $AB - BA = I_n$.

8.10 Suppose that the matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$ has only real entries a_{ij}. Show that if $\lambda \in \mathbb{C}\backslash\mathbb{R}$ is an eigenvalue of A with corresponding eigenvector $v = [\nu_1, \ldots, \nu_n]^T \in \mathbb{C}^{n,1}$, then also $\overline{\lambda}$ is an eigenvalue of A with corresponding eigenvector $\overline{v} := [\overline{\nu}_1, \ldots, \overline{\nu}_n]^T$.