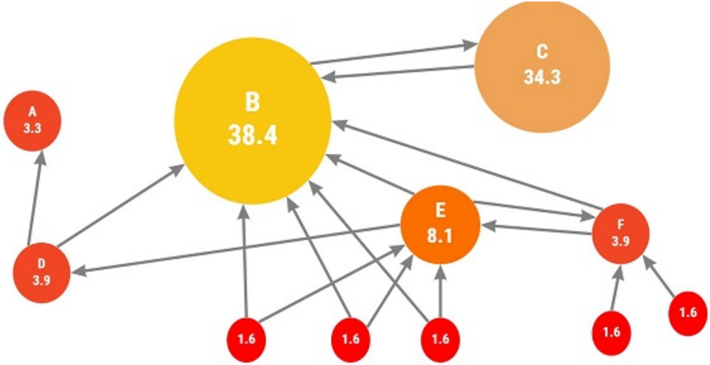
Course 9

Change of bases, eigenvectors and eigenvalues

Prof. dr. Septimiu Crivei

We describe a simplified version of *PageRank*.



Definition

Let V be a vector space over K, and let $B = (v_1, \ldots, v_n)$ and $B' = (v'_1, \ldots, v'_n)$ be bases of V. Then we can uniquely write

$$\begin{cases} v'_1 = t_{11}v_1 + t_{21}v_2 + \cdots + t_{n1}v_n \\ v'_2 = t_{12}v_1 + t_{22}v_2 + \cdots + t_{n2}v_n \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ v'_n = t_{1n}v_1 + t_{2n}v_2 + \cdots + t_{nn}v_n \end{cases}$$

for some $t_{ij} \in K$. Then the matrix $(t_{ij}) \in M_n(K)$, having as its columns the coordinates of the vectors of the basis B' in the basis B, is called the *change matrix* (or *transition matrix*) *from the basis B to the basis B'* and is denoted by $T_{BB'}$.

Properties of change matrices I

(1) B and B' are referred to as the "old" basis and the "new" basis respectively.

(2) The change matrix may be related to the matrix of a linear map as follows. For every $j \in \{1, \ldots, n\}$, the j^{th} column of $T_{BB'}$ consists of the coordinates of $v_j' = 1_V(v_j')$ in the basis B, hence $T_{BB'} = [1_V]_{B'B}$.

Theorem

Let V be a vector space over K, and let $B = (v_1, \ldots, v_n)$, $B' = (v_1', \ldots, v_n')$ and $B'' = (v_1'', \ldots, v_n'')$ be bases of V. Then

$$T_{BB''} = T_{BB'} \cdot T_{B'B''}.$$

Proof. Using a previous theorem, we have

$$T_{BB'} \cdot T_{B'B''} = [1_V]_{B'B} \cdot [1_V]_{B''B'} = [1_V \circ 1_V]_{B''B} = [1_V]_{B''B} = T_{BB''}.$$

Properties of change matrices II

Theorem

Let V be a vector space over K, and let B and B' be bases of V. Then the change matrix $T_{BB'}$ is invertible and its inverse is the change matrix $T_{B'B}$.

Proof. Using the above theorem for $B'' = B$, we have

$$T_{BB'}T_{B'B} = T_{BB} = I_n.$$

Also, changing the roles for B, B' and B'' by B', B and B' respectively, we have

$$T_{B'B}T_{BB'} = T_{B'B'} = I_n.$$

Hence $T_{BB'}$ is invertible and $T_{BB'}^{-1} = T_{B'B}$. $\square$

Properties of change matrices III

Theorem

Let V be a vector space over K, let $B = (v_1, \dots, v_n)$ and $B' = (v_1', \dots, v_n')$ be bases of V and let $v \in V$. Then

$$[v]_B = T_{BB'} \cdot [v]_{B'} \,.$$

Proof. Using a previous theorem, we have

$$T_{BB'} \cdot [v]_{B'} = [1_V]_{B'B} \cdot [v]_{B'} = [1_V(v)]_B = [v]_B. \qquad \square$$

Usually, we are interested in computing the coordinates of a vector v in the new basis B', knowing the coordinates of the same vector v in the old basis B and the change matrix from B to B'. Then:

$$[v]_{B'} = T_{BB'}^{-1} \cdot [v]_B = T_{B'B} \cdot [v]_B \,.$$

Example

Consider the bases $E = (e_1, e_2, e_3)$ and $B = (v_1, v_2, v_3)$ of the canonical real vector space $\mathbb{R}^3$, where E is the canonical basis and $v_1 = (0, 1, 1)$, $v_2 = (1, 1, 2)$, $v_3 = (1, 1, 1)$. Let us determine the change matrices from E to B and viceversa. After calculations [...]:

$$T_{EB} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \quad T_{BE} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

We must have $T_{BE} = T_{EB}^{-1}$, so that we could obtain T_{BE} by computing the inverse of T_{EB}.

Now consider the vector $u = (1, 2, 3)$. Clearly, its coordinates in the canonical basis E are 1, 2 and 3. It follows that

$$[u]_B = T_{BE} \cdot [u]_E = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Hence the coordinates of u in the basis B are 1, 1 and 0.

Theorem

Let $f \in \mathrm{Hom}_K(V, V')$, let B_1 and B_2 be bases of V and let B_1' and B_2' be bases of V'. Then

$$[f]_{B_2 B_2'} = T_{B_1' B_2'}^{-1} \cdot [f]_{B_1 B_1'} \cdot T_{B_1 B_2} \,.$$

Proof. We have

$$T_{B_1' B_2'}^{-1} \cdot [f]_{B_1 B_1'} \cdot T_{B_1 B_2} = T_{B_2' B_1'} \cdot [f]_{B_1 B_1'} \cdot T_{B_1 B_2}$$
$$= [1_V]_{B_1' B_2'} \cdot [f]_{B_1 B_1'} \cdot [1_V]_{B_2 B_1} = [1_V \circ f \circ 1_V]_{B_2 B_2'} = [f]_{B_2 B_2'} \qquad \square$$

Corollary

Let $f \in \mathrm{End}_K(V)$, and let B and B' be bases of V. Then

$$[f]_{B'} = T_{BB'}^{-1} \cdot [f]_B \cdot T_{BB'} \,.$$

Example

Consider the bases $E = (e_1, e_2, e_3)$ and $B = (v_1, v_2, v_3)$ of the canonical real vector space $\mathbb{R}^3$, where E is the canonical basis and $v_1 = (0,1,1)$, $v_2 = (1,1,2)$, $v_3 = (1,1,1)$. Let $f \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^3)$,

$$f(x,y,z) = (x+y, y-z, z+x), \ \forall (x,y,z) \in \mathbb{R}^3 .$$

We determine $[f]_E$ and $[f]_B$. We obtain [...]:

$$[f]_E = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} .$$

Using the computed change matrices T_{EB} and T_{BE}, we have

$$[f]_B = T_{EB}^{-1} \cdot [f]_E \cdot T_{EB} = T_{BE} \cdot [f]_E \cdot T_{EB} = \cdots = \begin{pmatrix} -1 & -3 & -2 \\ 1 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix} .$$

Alternatively, use the definition of the matrix of a linear map and express $f(v_1)$, $f(v_2)$, $f(v_3)$ as linear combinations of v_1, v_2, v_3.

Eigenvectors and eigenvalues

The study of endomorphisms of vector spaces also makes use of vectors whose images are just scalar multiples of themselves, in other words vectors that are "stretched" by an endomorphism. They are the subject of the present section.

Definition

Let $f \in \mathrm{End}_K(V)$. A non-zero vector $v \in V$ is called an *eigenvector of f* if there exists $\lambda \in K$ such that $f(v) = \lambda \cdot v$. Here λ is called an *eigenvalue of f*.

Clearly, each eigenvector has a unique corresponding eigenvalue. But different eigenvectors may have the same corresponding eigenvalue.

The characteristic subspace I

For $f \in \mathrm{End}_K(V)$, denote $V(\lambda) = \{v \in V \mid f(v) = \lambda v\}$, the set of the zero vector and the eigenvectors of f with eigenvalue λ.

Theorem

Then $V(\lambda)$ is a subspace of V.

Proof. Clearly, $0 \in V(\lambda)$, hence $V(\lambda) \neq \emptyset$. Now let $k_1, k_2 \in K$ and $v_1, v_2 \in V(\lambda)$. We have $f(v_1) = \lambda v_1$ and $f(v_2) = \lambda v_2$. Then

$$\begin{aligned} f(k_1 v_1 + k_2 v_2) &= k_1 f(v_1) + k_2 f(v_2) = k_1(\lambda v_1) + k_2(\lambda v_2) \\ &= (k_1\lambda)v_1 + (k_2\lambda)v_2 = \lambda(k_1 v_1 + k_2 v_2)\,. \end{aligned}$$

Hence, $k_1 v_1 + k_2 v_2 \in V(\lambda)$ and so, $V(\lambda)$ is a subspace of V. □

Definition

Then $V(\lambda)$ is called the *eigenspace* (or the *characteristic subspace*) of λ with respect to f.

Theorem

Let V be a vector space over K, B a basis of V and $f \in \mathrm{End}_K(V)$ with the matrix $[f]_B = A = (a_{ij}) \in M_n(K)$. Then $\lambda \in K$ is an eigenvalue of f if and only if

$$\det(A - \lambda \cdot I_n) = 0 \tag{1}$$

Proof. The element $\lambda \in K$ is an eigenvalue of f if and only if there exists a non-zero $v \in V$ such that $f(v) = \lambda v$. Consider $[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Then it follows that:

$$f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff (f - \lambda \cdot 1_V)(v) = 0$$
$$\iff [(f - \lambda \cdot 1_V)(v)]_B = [0]_B \iff [f - \lambda \cdot 1_V]_B \cdot [v]_B = [0]_B$$
$$\iff ([f]_B - \lambda \cdot [1_V]_B) \cdot [v]_B = [0]_B \iff (A - \lambda \cdot I_n) \cdot [v]_B = [0]_B$$

$$\Longleftrightarrow \begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Longleftrightarrow \begin{cases} (a_{11}-\lambda)x_1 + a_{12}x_2 \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n = 0. \end{cases} \qquad (S)$$

Then λ is an eigenvalue of $f \Leftrightarrow (S)$ has a non-zero solution $\Leftrightarrow$ $\det(A - \lambda \cdot I_n) = 0$. $\square$

(1) is called the *characteristic equation*,
(S) is called the *characteristic system*,
$p_A(\lambda) = \det(A - \lambda I_n)$ is called the *characteristic polynomial of A*.

If we take another basis B' of V and use the matrix $[f]_{B'}$, do we get the same eigenvalues and eigenvectors of f?

Theorem

Let V be a vector space over K, B and B' bases of V and $f \in End_K(V)$ with $[f]_B = A \in M_n(K)$ and $[f]_{B'} = A' \in M_n(K)$. Then $p_A(\lambda) = p_{A'}(\lambda)$.

Proof. We have $[f]_{B'} = T^{-1} \cdot [f]_B \cdot T$, where $T = T_{BB'}$. Hence we have $A' = T^{-1} \cdot A \cdot T$. Then

$$\begin{aligned} p_{A'}(\lambda) &= \det(A' - \lambda I_n) = \det(T^{-1}AT - \lambda I_n T^{-1}T) \\ &= \det(T^{-1}(A - \lambda I_n)T) \\ &= \det(T^{-1}) \cdot \det(A - \lambda I_n) \cdot \det(T) \\ &= \det(A - \lambda I_n) = p_A(\lambda). \end{aligned}$$

Remarks

(1) The eigenvalues and the eigenvectors *do not depend* on the basis chosen for writing the matrix of the endomorphism. Of course, the matrices might be different, but in the end we get the same characteristic polynomial. Consequently, we can say that the eigenvalues of an endomorphism (or simply, of a matrix) are just the roots in K of its unique characteristic polynomial.

(2) If V is a vector space over K with $\dim V = n$ and $f \in \mathrm{End}_K(V)$, then the degree of the characteristic polynomial of f is n, hence f may have at most n eigenvalues. If $K = \mathbb{C}$, then by the Fundamental Theorem of Algebra f has exactly n eigenvalues, not necessarily distinct.

(3) A non-zero vector $v \in K^n$ is an eigenvector of a matrix $A \in M_n(K)$ if and only if there exists $\lambda \in K$ such that $A[v]_E = \lambda[v]_E$, where E is the canonical basis of the canonical vector space K^n over K. In this case, λ is an eigenvalue of A.

Let $f \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^3)$ be defined by

$$f(x,y,z) = (2x, y+2z, -y+4z)\,,\ \forall (x,y,z) \in \mathbb{R}^3\,.$$

We write its matrix in the simplest basis, namely in the canonical basis E of $\mathbb{R}^3$. Then

$$[f]_E = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & 4 \end{pmatrix}\,.$$

The characteristic polynomial is $p(\lambda) = -(\lambda-2)^2(\lambda-3)$, so the eigenvalues are $\lambda_1 = \lambda_2 = 2$ and $\lambda_3 = 3$.
Let us take first $\lambda_1 = \lambda_2 = 2$. An eigenvector (x_1, x_2, x_3) is a non-zero solution of the characteristic system

$$\begin{pmatrix} 2-\lambda_1 & 0 & 0 \\ 0 & 1-\lambda_1 & 2 \\ 0 & -1 & 4-\lambda_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

that is,

$$\begin{cases} -x_2 + 2x_3 = 0 \\ -x_2 + 2x_3 = 0 \end{cases}.$$

Then $x_2 = 2x_3$ and $x_1, x_3 \in \mathbb{R}$, whence

$$V(2) = \{(x_1, 2x_3, x_3) \mid x_1, x_3 \in \mathbb{R}\} = \langle (1,0,0), (0,2,1) \rangle.$$

Any non-zero vector in $V(2)$ is an eigenvector of f with the associated eigenvalue $\lambda_1 = \lambda_2 = 2$.
Similarly, we have:

$$V(\lambda_3) = V(3) = \{(0, x_3, x_3) \mid x_3 \in \mathbb{R}\} = \langle (0,1,1) \rangle.$$

Any non-zero vector in $V(3)$ is an eigenvector of f with the associated eigenvalue $\lambda_3 = 3$.

Cayley-Hamilton Theorem

For $A \in M_n(K)$, $\mathrm{Tr}(A)$ is the *trace* of A, that is, the sum of the elements of the main diagonal of A.

Theorem

Let $A \in M_n(K)$ having eigenvalues $\lambda_1, \ldots, \lambda_n$. Then:
(i) $\lambda_1 + \cdots + \lambda_n = \mathrm{Tr}(A)$.
(ii) $\lambda_1 \cdots \lambda_n = \det(A)$.

Theorem (Cayley-Hamilton Theorem)

Every matrix $A \in M_n(K)$ is a root of its characteristic polynomial.

Corollary

Let $A \in M_2(K)$. Then:
(i) $p_A(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$.
(ii) $A^2 - \mathrm{Tr}(A) \cdot A + \det(A) \cdot I_2 = 0_2$.

Example I

Cayley-Hamilton Theorem may be used for computing the inverse or powers of a matrix. Let

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{R}).$$

Then $\det(A) = 2 \neq 0$, hence A is invertible. We have:

$$p_A(\lambda) = \det \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{pmatrix} = -\lambda^3 + 4\lambda^2 - 5\lambda + 2\,.$$

Then we have

$$A^3 - 4A^2 + 5A - 2I_3 = 0_3\,.$$

Example II

It follows that

$$A\left[\frac{1}{2}(A^2-4A+5I_3)\right]=\left[\frac{1}{2}(A^2-4A+5I_3)\right]A=I_3,$$

whence

$$A^{-1}=\frac{1}{2}(A^2-4A+5I_3)=\frac{1}{2}\begin{pmatrix}1&0&0\\0&2&0\\0&-2&2\end{pmatrix}.$$

For $k\geq 3$, the powers A^k can be computed using the recurrence relation given by the Cayley-Hamilton Theorem, namely

$$A^k=4A^{k-1}-5A^{k-2}+2A^{k-3}.$$

Diagonalization

Eigenvectors and eigenvalues are important for deciding whether an endomorphism is *diagonalizable* (i.e., its matrix has possibly non-zero entries only on its main diagonal), which is a much more useful computational form. As a sample result:

Theorem

Let V be a vector space over K with $\dim V = n$ and $f \in \mathrm{End}_K(V)$. Then f is diagonalizable if and only if it has n linearly independent eigenvectors. In particular, if f has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, then f is diagonalizable and

$$[f]_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

where B is the basis of the corresponding eigenvectors.

Extra: PageRank I

PageRank is a number assigned by Google to each web page. Pages with higher rank come higher in search results. We describe a simplified version, following [N. Strickland, *Linear Mathematics for Applications*].

- Consider pages $S_1, \dots, S_n$, with some links between them. A link from S_j to S_i is a vote by S_j that S_i is important.
- Links from important pages should count for more (because the probability of visiting S_i will clearly increase); links from pages with many links should count for less (because that will decrease the probability that we click the one that leads to S_i).
- We want rankings $r_1, \dots, r_n \geq 0$, normalized so that $\sum_{i=1}^{n} r_i = 1$.
- Say S_j links to N_j different pages, and assume $N_j > 0$. We use the rule: a link from S_j to S_i contributes $\frac{r_j}{N_j}$ to r_i.

- Thus, for every $i \in \{1, \ldots, n\}$, the following consistency condition should be satisfied:

$$r_i = \sum_{j \in J_i} \frac{r_j}{N_j},$$

where $J_i = \{j \in \{1, \ldots, n\} \mid \text{page } S_j \text{ links to page } S_i\}$.

- Define the matrix $P = (p_{ij}) \in M_n(\mathbb{R})$ by

$$p_{ij} = \begin{cases} \frac{1}{N_j} & \text{if there is a link from } S_j \text{ to } S_i \\ 0 & \text{otherwise.} \end{cases}$$

- Hence $\forall i \in \{1, \ldots, n\}$, the consistency condition becomes:

$$r_i = \sum_{j \in J_i} p_{ij} r_j.$$

- But this is equivalent to the matrix equation $Pr = r$, and thus r is an eigenvector of the matrix P with eigenvalue 1.