Lesson 6:
Diagonalization

Algebra

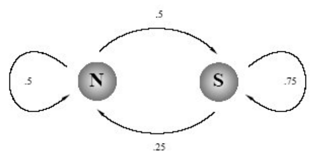
Computer Science Engineering Degree

May 4, 2016

Index

Problem

Suppose that the population moves from two geographical regions, say the North and South, as follows: each year 50% of the North population migrates to the south, while 25% of the South population migrates to the North.



Problema

If this migration pattern continues, will the North population decrease until all population will be concentrated in the South, or is there a stabilization tendency?

Statement or the problem using matrices

Let us denote n_0 and s_0 the proportions of initial populations in the North and the South, respectively. In the same way, let us denote n_k and s_k thr proportions at the end of the kth year. Observe that $n_k + s_k = 1$.
The migration pattern shows that

$$\left.\begin{array}{l} n_{k+1} = 0,5\, n_k + 0,25\, s_k \\ s_{k+1} = 0,5\, n_k + 0,75\, s_k \end{array}\right\}$$

Using matrices:

$$\begin{bmatrix} n_{k+1} \\ s_{k+1} \end{bmatrix} = \begin{bmatrix} 0,5 & 0,25 \\ 0,5 & 0,75 \end{bmatrix} \begin{bmatrix} n_k \\ s_k \end{bmatrix}$$

Statement of the problem using matrices

Let

$$\vec{x}_k = \begin{bmatrix} n_k \\ s_k \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 0,5 & 0,25 \\ 0,5 & 0,75 \end{bmatrix}$$

Then: $\vec{x}_{k+1} = AX_k$ para todo $k = 0, 1, 2, \ldots$, es decir,

$$\vec{x}_1 = A\vec{x}_0, \quad \vec{x}_2 = A\vec{x}_1 = A^2\vec{x}_0, \quad \vec{x}_3 = A\vec{x}_2 = A^3\vec{x}_0, \quad \cdots$$

That is:

$$\vec{x}_k = A^k\vec{x}_0 \text{ for all } k \geq 1$$

Therefore, the succesive powers of A determine the process.

Task

To give an expression of A^k for all k and, using it, to study the behaviour of $\vec{x}_k$ when k is big.

Idea

The theory we shall see in this lesson will allow us to obtain an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$. In fact, we can take:

$$P = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 1/4 \end{bmatrix}$$

Then $A = PDP^{-1}$ and, therefore:

$$A^k = \underbrace{PDP^{-1}}\underbrace{PDP^{-1}}\cdots\underbrace{PDP^{-1}} =$$

$$P \cdot D^k \cdot P^{-1} = P \begin{bmatrix} 1^k & 0 \\ 0 & (1/4)^k \end{bmatrix} P^{-1} = P \begin{bmatrix} 1 & 0 \\ 0 & 1/4^k \end{bmatrix} P^{-1}$$

Solution

Since

$$\vec{x}_k = A^k\vec{x}_0 = P \begin{bmatrix} 1 & 0 \\ 0 & 1/4^k \end{bmatrix} P^{-1} \begin{bmatrix} n_0 \\ s_0 \end{bmatrix}$$

taking limints when $k \to \infty$ we have:

$$\lim_{k\to\infty} \vec{x}_k = P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1} \begin{bmatrix} n_0 \\ s_0 \end{bmatrix} =$$

$$\begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} n_0 \\ s_0 \end{bmatrix} = \begin{bmatrix} (1/3)n_0 + (1/3)s_0 \\ (2/3)n_0 + (2/3)s_0 \end{bmatrix}$$

$$= \begin{bmatrix} (1/3)(n_0 + s_0) \\ (2/3)(n_0 + s_0) \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}$$

since $n_0 + s_0 = 1$. Then, the long-term tendency is a stabilization of $1/3$ of the population in the North and $2/3$ in the South, independently of the initial distribution (n_0, s_0).

Similarity of matrices and diagonalizable matrices

Definition

Two square matrices A and B are *similar* if there exists an invertible matrix P such that $\mathrm{A} = \mathrm{PBP}^{-1}$.

This is an equivalence relation in the set of square matrices.

Definition

A square matrix A is *diagonalizable* if it is similar to a diagonal matrix.

Problem to analize: Given a square matrix A is it diagonalizable? In the affirmative case, how can we compute a diagonal matrix D and an invertible matrix P such that $\mathrm{A} = \mathrm{PDP}^{-1}$?

Eigenvalues and eigenvectors

Definition

Let A be a square matrix.

- We shall say that a scalar λ is an eigenvalue of A if there exists a non-zero vector $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$. In this case, each vector $\vec{x} \neq \vec{0}$ satisfying this condition is called eigenvector associated with λ.
- Given an eigenvalue λ of A, we call eigenspace associated with λ to the vector subspace V_λ whose elements are $\vec{0}$ and all the eigenvectors associated to λ, that is, $V_\lambda := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda\vec{x}\} = Ker(A - \lambda I)$.

Computation of eigenvalues

If A is a square matrix:

λ is an eigenvalue of A $\Longleftrightarrow$

There exists $\vec{x} \neq 0$ such that $A\vec{x} = \lambda\vec{x}$ $\Longleftrightarrow$

There exists $\vec{x} \neq 0$ such that $(A - \lambda I)\vec{x} = \vec{0}$ $\Longleftrightarrow$

The homogeneous system $(A - \lambda I)\vec{x} = \vec{0}$ has non-zero solutions $\Longleftrightarrow$

The homogeneous system $(A - \lambda I)\vec{x} = \vec{0}$ has infinitely many solutions $\Longleftrightarrow$

The matrix $A - \lambda I$ is not invertible $\Longleftrightarrow$

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

Computation of eigenvalues

Definition

Given a square matrix A, its *characteristic polynomial* is

$$p_A(\lambda) := \det(A - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

From the previous slide:

Theorem

The eigenvalues of A are the roots of the characteristic polynomial.

Some immediate properties

(1) $p_A(\lambda)$ is a polynomial of degree n, the order of A.

(2) There eixist, at most, n different eigenvalues.

(3) Two similar matrices have the same characteristic polynomial.

Exercise 1: Prove these properties.

EXAMPLE: Consider the matrix $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$. Its characteristic polynomial is:

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

Then A has 2 different eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 3$.

Eamples

EXAMPLE: Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Its characteristic polynomial is:

$$p_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1\,.$$

Since it has no real roots, A has not eigenvalues

EXAMPLE: If $A = \begin{bmatrix} -2 & 4 & 5 \\ -3 & 5 & 5 \\ 0 & 0 & 1 \end{bmatrix}$ then

$$p_A(\lambda) = \begin{vmatrix} -2-\lambda & 4 & 5 \\ -3 & 5-\lambda & 5 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (-2-\lambda)(5-\lambda)(1-\lambda) + 12(1-\lambda) =$$

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda-1)^2(\lambda-2)$$

Therefore A has 2 different eigenvalues: $\lambda_1 = 1$ and $\lambda_2 = 2$.

Algebraic and geometric multiplicity

Definition

Let A a square matrix of order n and λ_i an eigenvalue.

- The *algebraic multiplicity* of λ_i is its multiplicity as a root of the characteristic polynomial of A, that is, the greatest exponent α_i for which the factor $(\lambda - \lambda_i)^{\alpha_i}$ appears in the decomposition of $p_A(\lambda)$.
- The *geometric multiplicity* of λ_i is the dimension d_i of its associated eigenspace, that is, $d_i = \dim\ V_{\lambda_i}$.

Since $V_{\lambda_i} = Ker(A - \lambda_i I)$, by a formula of Lesson 4 we have:

Proposición

$$d_i = n - \mathrm{rank}(A - \lambda_i I),$$

Algebraic and geometric multiplicities

EXAMPLE: Consider the matrix $A = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$.

Te characteristic polynomial is:

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

A has a unique eigenvalue $\lambda_1 = 2$ whose algebraic multiplicity is $\alpha_1 = 3$. The geometric multiplicity is:

$$d_1 = 3 - \operatorname{rank}(A - 2\,I) = 3 - \operatorname{rank} \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = 3 - 2 = 1 .$$

Properties of the eigenspaces

Let A a square matrix of order n and let $\lambda_1, \lambda_2, \dots, \lambda_r$ be its eigenvalues.

(a) $d_i = \dim V_{\lambda_i} \geq 1$ for all i.

(b) $V_{\lambda_i} \cap (V_{\lambda_i} + \cdots + V_{\lambda_{i-1}} + V_{\lambda_{i+1}} + \cdots + V_{\lambda_i}) = \{\vec{0}\}$ for all i. In particular $V_{\lambda_i} \cap V_{\lambda_j} = \{\vec{0}\}$ si $i \neq j$.

(c) Eigenvectors associated to different eigenvalues are linearly independent.

(d) If $\mathcal{B}_i$ is a basis of V_{λ_i} (for all $i = 1, \dots, r$) then $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ is a basis of $V_{\lambda_1} + \cdots + V_{\lambda_r}$.

(f) The following inequality is satisfied (very useful!!!): $1 \leq d_i \leq \alpha_i$ for all $i = 1, 2, \dots, r$.

Diagonalization criteria

Theorem (first criterion of diagonalization)

A square matrix A of order n is diagonalizable if and only if $V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_r} = \mathbb{R}^n$, where $\lambda_1, \lambda_2, \dots, \lambda_r$ are all the distinct eigenvalues of A (using (d) this is equivalent to say that there exists a basis of $\mathbb{R}^n$ whose elements are eigenvectors of A).

Theorem (second criterion of diagonalization) (extremely important!!!)

If A and $\lambda_1, \dots, \lambda_r$ are as in the previous theorem, A is diagonalizable if and only if these two properties are satisfied:

(1) $\alpha_1 + \alpha_2 + \cdots + \alpha_r = n$ (that is, all the roots of $p_A(\lambda)$ are real),

(2) $d_i = \alpha_i$ for all $i = 1, 2, \dots, r$.

As a consequence of this criterion and Property (f) of the previous slide:

Corollary (very useful!!!)

If a square matrix of order n has n distinct (real) eigenvalues then it is diagonalizable.

Diagonalization algorithm

Steps to determine if a square matrix A of order n is diagonalizable:

(1) Compute the characteristic polynomial $p_A(\lambda)$.

(2) Decompose into factors the characteristic polyonomial, obtaining the eigenvalues $\lambda_1, \dots, \lambda_r$ and their algebraic multiplicities $\alpha_1, \dots, \alpha_r$. If $\alpha_1 + \cdots + \alpha_r < n$ therefore A is not diagonalizable (by the second criterion). Otherwise go to the next step

(3) Compute the geometric multiplicities: $d_i = n - \operatorname{rank}(A - \lambda_i I)$. If, for some i, i, $d_i \neq \alpha_i$ then the matrix is not diagonalizable (by the second criterion). Otherwise A is diagonalizable.

Diagonalization algorithm

Steps to compute the matrices P and D such that D is diagonal and $P^{-1}AP = D$:

(4) The diagonal elements of D are $\overbrace{\lambda_1, \lambda_1, \ldots, \lambda_1}^{\alpha_1}, \overbrace{\lambda_2, \lambda_2, \ldots, \lambda_2}^{\alpha_2}, \ldots, \overbrace{\lambda_r, \lambda_r, \ldots, \lambda_r}^{\alpha_r}$ (each eigenvalue appears repeated according with its algebraic multiplicity).

(5) Obtain bases $\mathcal{B}_i$ of the eigenspaces V_{λ_i}.

(6) The matrix P is the one whose columns are the components of the vectors of the basis of $\mathbb{R}^n$ given by $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ (ordered accordingly with the order of the eigenvalues in the diagonal of D).

Example

We shall study if the matrix $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ is diagonalizable and we shall compute, in the affirmative case, the matrices D and P. We begin computing the characteristic polynomial of A:

$$p_A(\lambda) = \begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 24\lambda + 20\,.$$

We decompose $p_A(\lambda)$ using Ruffini:

$$p_A(\lambda) = (2-\lambda)^2(5-\lambda)\,.$$

Then the eigenvalues of A and their algebraic multiplicities are:

$$\begin{array}{ll} \lambda_1 = 2\,, & \alpha_1 = 2 \\ \lambda_2 = 5\,, & \alpha_1 = 1 \end{array}$$

Example

The next step consists of computing the geometric multiplicities:

$$d_1 = 3 - \operatorname{rank}(A - 2\,I) = 3 - \operatorname{rank}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = 3 - 1 = 2\,.$$

For d_2, applying that $1 \le d_2 \le \alpha_2 = 1$, we obtain directly $d_2 = 1$. We have, then, that $d_1 = \alpha_1$ and $d_2 = \alpha_2$. Applying the second criterion, the matrix A is diagonalizable and its diagonal form is

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Ejemplo

To compute the matrix P we need the bases of the eigenspaces V_{λ_1} and V_{λ_2}. We compute a basis of V_{λ_1}:

Since $A - 2\,I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ it holds that

$$(x, y, z) \in V_{\lambda_1} \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \{\; x + y + z = 0$$

Therefore, a basis of V_{λ_1} is $\{(-1, 1, 0), (-1, 0, 1)\}$.

Example

In the same way: $A - 5\,I = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$ and therefore

$$(x, y, z) \in V_{\lambda_2} \Leftrightarrow \begin{bmatrix} 2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} -2x + y + z = 0 \\ x - 2y + z = 0 \\ x + y - 2z = 0 \end{cases}$$

Solving the system one obtains that a basis of V_{λ_2} is $\{(1, 1, 1)\}$.
As a consequence, a basis of $\mathbb{R}^3$ given by eigenvectors is $\{(-1, 1, 0), (-1, 0, 1), (1, 1, 1)\}$ The matrix P is

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

Some interesting properties

(1) If A and B are similar matrices then $|A| = |B|$ and, moreover, they have the same rank.
(2) If A is a *symmetric matrix* then it is diagonalizable.
(3) If A is triangular, its eigenvalues are the diagonal elements.
(4) If λ is an eigenvalue of A then λ^k is an eigenvalue of A^k for all $k \in \mathbb{N}$.