February 16, 2022

I.3 MATRIX DIAGONALIZATION

1. Diagonalization of matrices

Definition 1.1. Two matrices A and B of order n are similar if there exists a matrix P such that

$$B = P^{-1}AP.$$

Definition 1.2. A matrix A is diagonalizable if it is similar to a diagonal matrix D, that is, there exists D diagonal and P invertible such that $D = P^{-1}AP$.

Of course, D diagonal means that every element out of the diagonal is null

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Proposition 1.3. *If A is diagonalizable, then for all $m \geq 1$*

$$A^m = PD^mP^{-1}, \tag{1.1}$$

where

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^m \end{pmatrix}.$$

Proof. Since A is diagonalizable

$$\begin{aligned} A^m &= (PDP^{-1})(PDP^{-1}) \overset{m}{\cdots} (PDP^{-1}) \\ &= PD(P^{-1}P)D \cdots D(P^{-1}P)DP^{-1} \\ &= PDI_nD \cdots DI_nDP^{-1} = PD^mP^{-1}. \end{aligned}$$

The expression for D^m is readily obtained by induction on m. □

Definition 1.4. Let A be a matrix of order n.

- We say that $\overline{u} \in \mathbb{R}^n$, $\overline{u} \neq \overline{0}$, is an eigenvector or proper vector of A if and only if there is $\lambda \in \mathbb{R}$ such that $A\overline{u} = \lambda\overline{u}$.

- We say that $\lambda \in \mathbb{R}$ is an eigenvalue or proper value of A if and only if there is $\overline{u} \in \mathbb{R}^n$, $\overline{u} \neq \overline{0}$, such that $A\overline{u} = \lambda\overline{u}$.

Remark 1.5. If $\overline{u} \in \mathbb{R}^n$, $\overline{u} \neq \overline{0}$ and $A\overline{u} = \lambda\overline{u}$, we say that $\overline{u}$ is an eigenvector of A associated to the eigenvalue λ.

Definition 1.6. For a matrix $A \in \mathcal{M}_{n\times n}$ and a real number λ, let the set

$$\mathrm{V}(\lambda) = \{\overline{u} \in \mathbb{R}^n \,:\, A\overline{u} = \lambda\overline{u}\}.$$

Note that $\overline{0} \in \mathrm{V}(\lambda)$ for all $\lambda \in \mathbb{R}$.

Remark 1.7.

(1) $\mathrm{V}(\lambda)$ is the set of solutions of the linear homogenous system

$$(A - \lambda I_n)\overline{u} = \overline{0}.$$

(2) When λ is an eigenvalue of A, $\mathrm{V}(\lambda)$ is the set of all the eigenvectors associated to λ, together with the null vector $\overline{0}$, and it is called the proper subspace associated to λ.

(3) If λ is not a proper value of A, then $\mathrm{V}(\lambda) = \{\overline{0}\}$.

The following result shows that an eigenvector can only be associated to a unique eigenvalue.

Proposition 1.8. *Let $A \in \mathcal{M}_{n\times n}$ and let $\lambda, \mu \in \mathbb{R}$ two eigenvalues of A.*

(1) For all $\overline{u} \in \mathrm{V}(\lambda)$, $A\overline{u} \in \mathrm{V}(\lambda)$.

(2) If $\lambda \neq \mu$, then $\mathrm{V}(\lambda) \cap \mathrm{V}(\mu) = \emptyset$.

Proof.

(1) For $\overline{u} \in \mathrm{V}(\lambda)$, $A(A\overline{u}) = A(\lambda\overline{u}) = \lambda A\overline{u}$, thus $A\overline{u} \in \mathrm{V}(\lambda)$.

(2) Suppose $\overline{0} \neq \overline{u} \in \mathrm{V}(\lambda) \cap \mathrm{V}(\mu)$. Then

$$A\overline{u} = \lambda\overline{u}$$
$$A\overline{u} = \mu\overline{u}.$$

Subtracting both equations we obtain $\overline{0} = (\lambda - \mu)\overline{u}$ and, since $\overline{0} \neq \overline{u}$, we must have $\lambda = \mu$.

□

Recall that for an arbitrary matrix A, *the rank of the matrix* is the number of linearly independent columns or rows (both numbers necessarily coincide). It is also given by the order of the largest non null minor of A.

Theorem 1.9. *The real number λ is an eigenvalue of A if and only if*

$$|A - \lambda I_n| = 0.$$

Moreover, $\mathrm{V}(\lambda)$ *is the set of solutions (including the null vector) of the linear homogeneous system*

$$(A - \lambda I_n)\overline{u} = \overline{0},$$

and hence it is a vector subspace, which dimension is

$$\dim \mathrm{V}(\lambda) = n - \mathrm{rank}(A - \lambda I_n).$$

Proof. Suppose that $\lambda \in \mathbb{R}$ is an eigenvalue of A. Then the system $(A - \lambda I_n)\overline{u} = \overline{0}$ admits some non–trivial solution $\overline{u}$. Since the system is homogeneous, this implies that the determinant of the system is zero, $|A - \lambda I_n| = 0$. The second part about $\mathrm{V}(\lambda)$ follows also from the definition of eigenvector, and the fact that the set of solutions of a linear homogenous system is a subspace (the sum of two solutions is again a solution, as well as it is the product of a real number by a solution). Finally, the dimension of the space of solutions is given by the Theorem of Rouche–Frobenius. $\square$

Definition 1.10. The characteristic polynomial of A is the polynomial of order n given by

$$p_A(\lambda) = |A - \lambda I_n|.$$

Notice that the eigenvalues of A are the real roots of p_A. This polynomial is of degree n. The Fundamental Theorem of Algebra estates that a polynomial of degree n has n complex roots (not necessarily different, some of the roots may have multiplicity grater than one). It could be the case that some of the roots of p_A were not real numbers. For us, a root of $p_A(\lambda)$ which is not real is not an eigenvalue of A.

Example 1.11. Find the eigenvalues and the proper subspaces of

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

ANSWER:

$$A-\lambda I=\begin{pmatrix}-\lambda & -1 & 0\\ 1 & -\lambda & 0\\ 0 & 0 & 1-\lambda\end{pmatrix};\quad p(\lambda)=(1-\lambda)\begin{vmatrix}-\lambda & -1\\ 1 & -\lambda\end{vmatrix}=(1-\lambda)(\lambda^2+1).$$

The characteristic polynomial has only one real root, hence the spectrum of A is $\sigma(A)=\{1\}$. The proper subspace V(1) is the set of solutions of the homogeneous linear system $(A-I_3)\overline{u}=\mathbf{0}$, that is, the set of solutions of

$$(A-I_3)\overline{u}=\begin{pmatrix}-1 & -1 & 0\\ 1 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

Solving the above system we obtain

$$\mathrm{V}(1)=\{(0,0,z)\,:\,z\in\mathbb{R}\}=<(0,0,1)>\ \text{(the subspace generated by }(0,0,1)).$$

Notice that $p_A(\lambda)$ has other roots that are not reals. They are the *complex* numbers $\pm i$, that are not (real) eigenvalues of A. If we would admit complex numbers, then they would be eigenvalues of A in this extended sense.

Example 1.12. Find the eigenvalues and the proper subspaces of

$$B=\begin{pmatrix}2 & 1 & 0\\ 0 & 1 & -1\\ 0 & 2 & 4\end{pmatrix}.$$

ANSWER: The eigenvalues are obtained solving

$$\begin{vmatrix}2-\lambda & 1 & 0\\ 0 & 1-\lambda & -1\\ 0 & 2 & 4-\lambda\end{vmatrix}=0.$$

The solutions are $\lambda=3$ (simple root) and $\lambda=2$ (double root). To find $\mathrm{V}(3)=\{\overline{u}\in\mathbb{R}^3 : (B-3I_3)\overline{u}=\overline{0}\}$ we compute the solutions to

$$(B-3I_3)\overline{u}=\begin{pmatrix}-1 & 1 & 0\\ 0 & -2 & -1\\ 0 & 2 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

which are $x = y$ and $z = -2y$, and hence V(3) $=< (1, 1, -2) >$. To find V(2) we solve the system

$$(B - 2I_3)\overline{u} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

from which $x = y = 0$ and hence V(2) $=< (1, 0, 0) >$.

Example 1.13. Find the eigenvalues and the proper subspaces of

$$C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 3 \end{pmatrix}$$

Answer: To compute the eigenvalues we solve the characteristic equation

$$0 = |C - \lambda I_3| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ 1 & 1 & 0-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 2-\lambda \end{vmatrix} \begin{vmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)(3-\lambda)$$

So, the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$. We now compute the eigenvectors. The eigenspace V(1) is the solution of the homogeneous linear system whose associated matrix is $C - \lambda I_3$ with $\lambda = 1$. That is, V(1) is the solution of the following homogeneous linear system

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

Solving the above system we find that

$$\mathrm{V}(1) = \{(-2z, 0, z) : z \in \mathbb{R}\} =< (-2, 0, 1) >$$

On the other hand, V(2) is the set of solutions of the homogeneous linear system whose associated matrix is $C - \lambda I_3$ with $\lambda = 2$. That is, V(2) is the solution of the following

$$\begin{pmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

So,

$$\mathrm{V}(2) = \{(2y, y, -3y) : y \in \mathbb{R}\} =< (2, 1, -3) >$$

Finally, V(3) is the set of solutions of the homogeneous linear system whose associated matrix is $A - \lambda I_3$ with $\lambda = 3$. That is, V(3) is the solution of the following

$$\begin{pmatrix} -2 & 2 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and we obtain

$$\mathrm{V}(3) = \{(0,0,z) : z \in \mathbb{R}\} =< (0,0,1) >$$

We now start describing the *procedure to diagonalize a matrix.* Fix a square matrix A. Let

$$\lambda_1, \lambda_2, \dots, \lambda_k$$

be distinct *real* roots of the characteristic polynomial $p_A(\lambda)$ an let m_k be the multiplicity of each λ_k (Hence $m_k = 1$ if λ_k is a simple root, $m_k = 2$ if it is double, etc.). Note that $m_1 + m_2 + \cdots + m_k = n$.

The following result estates that the number of independent vectors in the subspace $\mathrm{V}(\lambda)$ can never be bigger than the multiplicity of λ.

Proposition 1.14. *For each $j = 1, \dots, k$*

$$1 \leq \dim \mathrm{V}(\lambda_j) \leq m_j.$$

The following theorem gives necessary and sufficient conditions for a matrix A to be diagonalizable.

Theorem 1.15. *A matrix A is diagonalizable if and only if the two following conditions hold.*

(1) Every root, $\lambda_1, \lambda_2, \dots, \lambda_k$ of the charateristic polynomial $p_A(\lambda)$ is real.

(2) For each $j = 1, \dots, k$

$$\dim \mathrm{V}(\lambda_j) = m_j.$$

Corollary 1.16. *If the matrix A has n distinct real eigenvalues, then it is diagonalizable.*

Theorem 1.17. *If A is diagonalizable, then the diagonal matrix D is formed by the eigenvalues of A in its main diagonal, with each λ_j repeated n_j times. Moreover, a matrix P such that $D = P^{-1}AP$ has as columns independent eigenvectors selected from each proper subspace* $\mathrm{V}(\lambda_j)$, $j = 1, \dots, k$.

Comments on the examples above.

- Matrix A of Example 1.11 is not diagonalizable, since p_A has complex roots.
- Although all roots of p_B are real, B of Example 1.12 is not diagonalizable, because $\dim \mathrm{V}(2) = 1$, which is smaller than the multiplicity of the eigenvalue 2.
- Matrix C of Example 1.13 is diagonalizable, since p_C has 3 different real roots. In this case

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \qquad P = \begin{pmatrix} -2 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}.$$

Example 1.18. The manager of a certain NGO observes the following behavior in its affiliates: 80% of affiliates who made donations in a certain year (x_t), also contribute to the following year, while 30% of affiliates who did not contribute in that year (y_t), do will do a contribution the following year. Calculate the percentage of affiliates who make donations to the NGO in the long term, if it is known that at the present time this percentage is 50%.

Actually, we have the system

$$\begin{cases} x_{t+1} = 0.8x_t + 0.3y_t \\ y_{t+1} = 0.2x_t + 0.7y_t \end{cases}$$

The matrix associated to the system of difference equations is the matrix A

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}$$

The eigenvalues of A are the solutions of $p_A(\lambda) = 0$, where

$$p_A(\lambda) = |A - \lambda I| = \begin{vmatrix} 0.8 - \lambda & 0.3 \\ 0.2 & 0.7 - \lambda \end{vmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = 0.$$

Then $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{2}$.

For $V(\lambda_1 = 1)$ we have the following homogeneous system of linear equations

$$\begin{pmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

Then

$$V(\lambda_1 = 1) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \cdot t, \text{ where } t \in \mathbb{R}.$$

For $V(\lambda_2 = \frac{1}{2})$ we have the following homogeneous system of linear equations

$$\begin{pmatrix} 0.3 & 0.3 \\ 0.2 & 0.2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Then

$$V(\lambda_2 = \frac{1}{2}) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot t, \text{ where } t \in \mathbb{R}.$$

Then

$$D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, P = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}, P^{-1} = -\frac{1}{5} \begin{pmatrix} -1 & -1 \\ -2 & 3 \end{pmatrix}.$$

$$\lim_{t\to\infty} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & (\frac{1}{2})^t \end{pmatrix} \cdot \left(-\frac{1}{5}\right) \begin{pmatrix} -1 & -1 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \frac{3}{5}(x_0 + y_0) \\ \frac{2}{5}(x_0 + y_0) \end{pmatrix} = \begin{pmatrix} \frac{3}{5} \\ \frac{2}{5} \end{pmatrix}.$$

Thus, (x_t, y_t) converges to the stationary distribution $(\frac{3}{5}, \frac{2}{5})$ independently of the values of x_0 and y_0, which means that 60% of affiliates to the NGO make donations in the long run.

Proposition 1.19. *Let $A \in \mathcal{M}_{n\times n}$ be symmetric. Then all the roots of the characteristic polynomial $P_A(\lambda)$ are real, A is diagonalizable, and there are matrices $P, D \in \mathcal{M}_{n\times n}$, with D diagonal and P orthogonal ($P^{-1} = P^t$), such that $A = P^t DP$.*