Diagonalization

Guillaume Frèche

Version 1.0

We have introduced in the previous lecture the vector space $L^2_{T_0}(\mathbb{R}, \mathbb{K})$ of periodic signals with signal T_0 which are locally square-integrable. Before studying in detail the structure of this vector space, we review some notions of linear algebra, especially diagonalization, and apply these concepts to signal processing. We denote V a vector space over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

1 Linear algebra review

Definition 1.1 (Linear combination, generating set)
Let $v_1, \dots, v_n$ be vectors of V. Let $\text{Span}(v_1, \dots, v_n)$ be the set of **linear combinations** of $v_1, \dots, v_n$, i.e.

$$\text{Span}(v_1, \dots, v_n) = \left\{ \sum_{j=1}^{n} \lambda_j v_j, (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \right\}$$

Set $(v_1, \dots, v_n)$ is a **generating set** of V if $V = \text{Span}(v_1, \dots, v_n)$, i.e. any vector of V can be written as a linear combination of $v_1, \dots, v_n$.

Definition 1.2 (Linearly independent, linearly dependent)
A set $(v_1, \dots, v_n)$ is **linearly independent** if for any $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$, relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$ implies $\lambda_1 = \cdots = \lambda_n = 0$.
Otherwise, i.e. if there exists $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ such $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$, $(v_1, \dots, v_n)$ is **linearly dependent**.

Definition 1.3 (Basis, coordinates)
A set $(v_1, \dots, v_n)$ is a **basis** of V if it is a linearly independent generating set of V. In other words, any vector of V can be written uniquely as a linear combination of $v_1, \dots, v_n$. Scalars $\lambda_1, \dots, \lambda_n$ appearing in this linear combination are called the **coordinates** of the vector.

In the previous lectures, we have equipped the spaces $L^2(\mathbb{R}, \mathbb{K})$ and $L^2_{T_0}(\mathbb{R}, \mathbb{K})$ with a scalar product or a Hermitian product. We enrich these notions with some algebraic definitions.

Definition 1.4 (Orthogonal set, orthonormal set)
A set $(v_1, \dots, v_n)$ is **orthogonal** if for any $(j, k) \in [\![1, n]\!]^2$, with $j \neq k$, $\langle v_j, v_k \rangle = 0$. Moreover, if for any $j \in [\![1, n]\!]$, $\|v_j\| = 1$, this set is **orthonormal**.

Remark: Using these definitions, we can easily prove that any orthogonal or orthonormal set is linearly independent.

> **Definition 1.5 (Orthogonal basis, orthonormal basis)**
> An **orthogonal basis** (resp. **orthonormal basis**) of V is an orthogonal (resp. orthonormal) set which is a basis of V.

Remark: The interest in orthonormal bases is to easily express the coordinates of a vector from its scalar product with the vectors of the basis. Indeed, if $(v_1, \dots, v_n)$ is an orthonormal basis of V, then any vector x of V can be written

$$x = \sum_{j=1}^{n} \langle x, v_j \rangle v_j$$

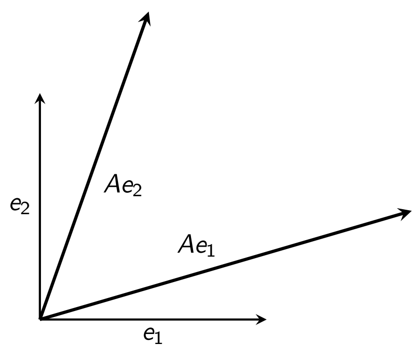
2 Diagonalization

A matrix provides the representation of a linear mapping within a vector space in a given basis, but this single data does not seem enough to "geometrically" interpret the behavior of this mapping. For instance, consider the following 2×2 matrix:

$$A = \begin{pmatrix} 1.64 & 0.48 \\ 0.48 & 1.36 \end{pmatrix}$$

We apply this matrix to vectors v_1 and v_2 of the standard basis of $\mathbb{R}^2$:

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad Av_1 = \begin{pmatrix} 1.64 \\ 0.48 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad Av_2 = \begin{pmatrix} 0.48 \\ 1.36 \end{pmatrix}$$



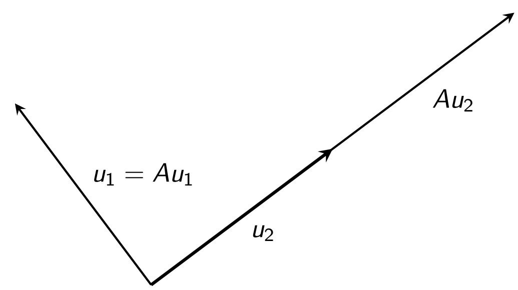
As shown in this figure, the transform rotates and scales these vectors, but it seems difficult to deduce a general behavior. The idea is to find cases for which the transform only scales the input vector. Thereby, λ is an **eigenvalue** of A if there exists a non-zero vector u such that $Au = \lambda u$. Vector u is called an **eigenvector** of A associated with eigenvalue λ. We possibly aim at determining a basis of eigenvectors, enabling the representation of the transform with a diagonal matrix only displaying the scalings in given directions. This process is called the **diagonalization** of A.
The first step consists in computing the eigenvalues, which are the roots of the characteristic polynomial:

$$\chi_A(\lambda) = \det(A - \lambda I_2) = \begin{vmatrix} 1.64 - \lambda & 0.48 \\ 0.48 & 1.36 - \lambda \end{vmatrix} = (\lambda - 1.64)(\lambda - 1.36) - 0.48^2 = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$$

Thus the eigenvalues of A are 2 et 1.
The second step consists in determining eigenvectors u_2 and u_1, respectively associated with eigenvalues 2 and 1. Moreover, we look for vectors with norm 1 to obtain, if possible, an orthonormal basis. To find eigenvector u_2, we write:

$$A \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \quad \Longrightarrow \quad \begin{cases} 1.64x + 0.48y = 2x \\ 0.48x + 1.36y = 2y \end{cases} \quad \Longrightarrow \quad x = \frac{4}{3} y$$

so that $u_2 = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}$ is an eigenvector of A associated with eigenvalue 2 and with norm 1. With the same reasoning, we find $u_1 = \begin{pmatrix} -0.6 \\ 0.8 \end{pmatrix}$ as eigenvector of A associated with eigenvalue 1 and with norm 1.



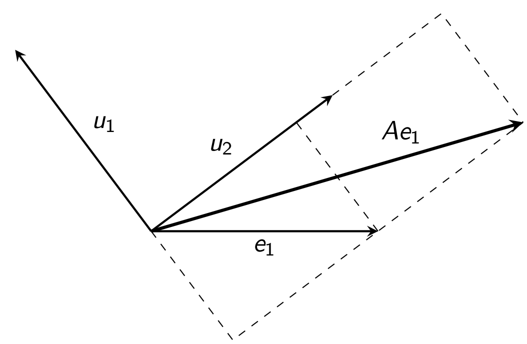
Computing $\langle u_2, u_1 \rangle = 0.8 \times (-0.6) + 0.6 \times 0.8 = 0$, we prove that (u_2, u_1) if an orthonormal basis of $\mathbb{R}^2$. We can write any vector $x \in \mathbb{R}^2$ as:

$$x = \langle x, u_2 \rangle u_2 + \langle x, u_1 \rangle u_1$$

along with its image by A:

$$Ax = \langle x, u_2 \rangle Au_2 + \langle x, u_1 \rangle Au_1 = 2\langle x, u_2 \rangle u_2 + \langle x, u_1 \rangle u_1$$

For instance e_1 can be written $e_1 = 0.8u_2 - 0.6u_1$, and its image $Ae_1 = 1.6u_2 - 0.6u_1$.



We can see on this figure that u_2 coordinate of vector e_1 is multiplied by 2, while the u_1 coordinate is unchanged (multiplied by 1).

3 Application to signal processing

We apply these ideas to our study of signals and systems by considering the signal vector space $\mathcal{F}(\mathbb{R}, \mathbb{K})$. In this case, LTI systems correspond to matrix A (since they are linear), and input signals are arguments of this mapping. In this context, we talk about **eigenfunctions** instead of eigenvectors to insist on the fact that $\mathcal{F}(\mathbb{R}, \mathbb{K})$ is a space containing functions. The following proposition provides a very important result about some eigenfunctions of an LTI system.

> **Proposition 3.1**
> For any $\omega \in \mathbb{R}$, complex exponential $e_\omega : \mathbb{R} \to \mathbb{C} \quad t \mapsto e^{i\omega t}$ is an eigenfunction of any LTI system. In other words, if L is an LTI system, then there exists $H(\omega) \in \mathbb{C}$ such that $L(e_\omega) = H(\omega) e_\omega$.

Proof : We set $\omega \in \mathbb{R}$ and L an LTI system of impulse response $h = L(\delta)$. Then for any $t \in \mathbb{R}$,

$$L(e_\omega)(t) = (e_\omega * h)(t) = \int_{-\infty}^{+\infty} h(u) e^{i\omega(t-u)} du = e^{i\omega t} \int_{-\infty}^{+\infty} h(u) e^{-i\omega u} du$$

thus $L(e_\omega) = H(\omega) e_\omega$ with $H(\omega) = \displaystyle\int_{-\infty}^{+\infty} h(u) e^{-i\omega u} du$, therefore e_ω is an eigenfunction of L. ■

Remarks:

- **WARNING**: the complex exponentials are some eigenfunctions but not necessarily all the eigenfunctions of a given LTI system. For instance, consider the complex-valued differential system $D : \mathcal{F}(\mathbb{R}, \mathbb{C}) \to \mathcal{F}(\mathbb{R}, \mathbb{C}) \quad x \mapsto x'$. Then any complex number $s = a + ib \in \mathbb{C}$ is an eigenvalue of D, whose associated eigenfunctions $Ke^{st} = Ke^{at} e^{ibt}$ are the solutions of the differential equation: $D(x) = x' = sx$. In this example, complex exponential e_ω corresponds to $a = 0$ and $b = \omega$.
- We will study in detail the expression of the eigenvalue $H(\omega)$ in the lecture about Fourier transform.
- We have seen in the previous lecture that if we input a sine or a cosine into the RC circuit, we obtain an output which is a linear combination of sine and cosine with the same fundamental impulse. This result is now explained with this proposition. We generalize this idea in the next lecture.