Real-valued symmetric matrices always diagonalize

Vera Sacristán

Definition 1 *The sum of two subspaces* W_1 *and* W_2 *of a vector space* V *is defined as*

$$W = W_1 + W_2 = \{w \in V \mid w = w_1 + w_2, w_1 \in W_1, w_2 \in W_2\}.$$

If $W_1 \cap W_2 = \{0\}$*, the sum is called direct, and we write* $W_1 \oplus W_2$*.*

Lemma 1 *The sum* $W_1 + W_2$ *of two subspaces of a vector space* V *is a subspace of* V*.*

Proof: Immediate. □

Definition 2 *Let* W *be a subspace of an Euclidean vector space* V*. The subset of* V *orthogonal to* W *is defined as* $W^\perp = \{v \in V \mid v \perp w \ \forall\, w \in W\}$*.*

Lemma 2 *If* W *is a subspace of* V*, then the set* $W^\perp$ *is a subspace of* V*.*

Proof: Immediate. □

Theorem 3 *If* W *is a subspace of an Euclidean vector space* V*, then* $V = W \oplus W^\perp$*.*

Proof: The fact that $W \cap W^\perp = \{0\}$ is easy to prove: if $u \in W \cap W^\perp$ then $u \in W$ and $u \cdot w = 0$ forall $w \in W$. In particular, then, $u \cdot u = 0$ and we get $u = 0$. Therefore, the sum of W and $W^\perp$ is direct. Trivially, $W \oplus W^\perp \subseteq V$. In order to prove that $V \subseteq W \oplus W^\perp$, consider $w_1, \ldots, w_r$ an orthonormal basis of W which can be obtained using Gram-Schmidt method, for example. Let $w_{r+1}, \ldots, w_n$ be its completion to an orthonormal basis of V, which can also be obtained using Gram-Schmidt method. It is immediate to prove that $w_{r+1}, \ldots, w_n$ is a basis of $W^\perp$ due to the orthonormality of the basis $w_1, \ldots, w_n$. Therefore, every vector $x \in V$ can be written as

$$x = \sum_{i=1}^{n} x_i w_i = \sum_{i=1}^{r} x_i w_i + \sum_{i=r+1}^{n} x_i w_i \in W \oplus W^\perp.$$

□

Lemma 4 *Let* f *be an endomorphism in an Euclidean vector space* V *whose associated matrix in some orthonormal basis is symmetric. Then* $f(u) \cdot v = u \cdot f(v)\ \forall u, v \in V$*.*

Proof: Let $e_1, \ldots, e_n$ be the orthonormal basis in which f is represented by the symmetric matrix $A = (a_i^j)$. Due to the bilinearity of the dot product, we only need to prove that $f(e_i) \cdot e_j = e_i \cdot f(e_j)\ \forall i, j \in \{1, \ldots, n\}$:

$$f(e_i) \cdot e_j = \left(\sum_{k=1}^{n} a_k^i w_k \right) \cdot w_j = \sum_{k=1}^{n} a_k^i \delta_{k,j} = a_j^i$$

$$e_i \cdot f(e_j) = e_i \cdot \left(\sum_{k=1}^{n} a_k^j w_k \right) = \sum_{k=1}^{n} a_k^j \delta_{i,k} = a_i^j$$

Since A is symmetric, we obtain $f(e_i) \cdot e_j = a_j^i = a_i^j = e_i \cdot f(e_j)\ \forall i, j \in \{1, \ldots, n\}$. □

Proposition 5 *Let A be a symmetric matrix of size n with real coefficients, and let $\lambda_1, \ldots, \lambda_r$ be its eigenvalues, with multiplicities $k_1, \ldots, k_r$, respectively. Then $\lambda_i \in \mathbb{R}$ $\forall i \in \{1, \ldots, r\}$ and $\sum_{i=1}^{r} k_i = n$.*

Proof: Let us consider $A \in M_n(\mathbb{R}) \subset M_n(\mathbb{C})$. The Fundamental Theorem of Algebra guarantees that the characteristic polynomial of A has complex roots $\lambda_1, \ldots, \lambda_r$, with multiplicities $k_1, \ldots, k_r$ respectively, and $\sum_{i=1}^{r} k_i = n$. We will prove that $\lambda_i \in \mathbb{R}$ for all $i \in \{1, \ldots, r\}$. Indeed, if $z = (z_1, \ldots, z_n)$ is a (complex) eigenvector of eigenvalue λ, we have $Az = \lambda z$. Since the coefficients of A are real numbers, when we conjugate the previous equality we obtain that $\overline{\lambda}\overline{z} = \overline{A}\overline{z} = A\overline{z}$. In other words, $\overline{z}$ is an eigenvector of A with eigenvalue $\overline{\lambda}$. Then,

$$\lambda|z| = \lambda\overline{z}^T z = \overline{z}^T \lambda z = \overline{z}^T A z = \overline{z}^T A^T z = (z^T A \overline{z})^T \stackrel{*}{=} z^T A \overline{z} = z^T \overline{\lambda}\overline{z} = \overline{\lambda} z^T \overline{z} = \overline{\lambda}|z|,$$

where the starred equality holds because $\lambda|z| \in \mathbb{R}$. Hence, $\lambda = \overline{\lambda}$, and $\lambda \in \mathbb{R}$. □

Theorem 6 *Let V be an Euclidean real vector space V, and let A be a symmetric matrix. Then V admits a basis of orthonormal eigenvectors of A.*

Proof: By induction over the dimension of V, denoted n. The base case corresponds to $n = 1$ and is immediate: each non null vector is an eigenvector of A and can be normalized. The induction step is proved as follows: let $\lambda \in \mathbb{R}$ be an eigenvalue of A, and let v be a unit eigenvector for λ. We know that $< v > \oplus < v >^{\perp} = V$. We will prove that $< v >^{\perp}$ is invariant under the endomorphism f associated to A. If $u \in < v >^{\perp}$, then $u \cdot v = 0$. As a consequence, $f(u) \cdot v = u \cdot f(v) = u \cdot \lambda v = \lambda u \cdot v = 0$. This proves that $f(u) \in < v >^{\perp}$. By inductive hypothesis, $< v >^{\perp}$ has an orthonormal basis made of eigenvectors of f restricted to $< v >^{\perp}$. Adding v to this basis we obtain an orthonormal basis made of eigenvectors of f. □

Corollary 7 *Let V be an Euclidean real vector space V, and let A be a symmetric matrix. Then A diagonalizes in orthonormal basis.*

Proof: Immediate from Theorem 6. □