Extra: More details on the spectral theorem

Let us add $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$ to our notations for the dot product: $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{v} \cdot \boldsymbol{w}$.

- In our story of orthogonality, the important player has been the dot product. However, one could argue that the fundamental quantity is actually the norm:
  $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \frac{1}{4}(\|\boldsymbol{v} + \boldsymbol{w}\|^2 - \|\boldsymbol{v} - \boldsymbol{w}\|^2)$. See Example 28.
- Accepting the dot product as immensely important, we see that symmetric matrices (i.e. matrices A such that $A = A^T$) are of interest.
  For every matrix A, $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle$.
  It follows that, a matrix A is symmetric if and only if $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.
- Similarly, let Q be an orthogonal matrix (i.e. Q is a square matrix with $Q^T Q = I$).
  Then, $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
  In fact, a matrix A is orthogonal if and only if $\langle A\boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.
  Comment. We observed in Example 150 that orthogonal matrices Q correspond to rotations ($\det Q = 1$) or reflections ($\det Q = -1$) [or products thereof]. The equality $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ encodes the fact that these types (and only these!) of geometric transformations preserve angles and lengths.

(spectral theorem)

A $n \times n$ matrix A is symmetric if and only if it can be decomposed as $A = PDP^T$, where

- D is a diagonal matrix, $(n \times n)$

The diagonal entries λ_i are the **eigenvalues** of A.

- P is orthogonal. $(n \times n)$

The columns of P are **eigenvectors** of A.

Note that, in particular, A is always diagonalizable, the eigenvalues (and hence, the eigenvectors) are all real, and, most importantly, the eigenspaces of A are orthogonal.
The "only if" part says that, if A is symmetric, then we get a diagonalization $A = PDP^T$. The "if" part says that, if $A = PDP^T$, then A is symmetric (which follows from $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$).

Let us prove the following important parts of the spectral theorem.

We already proved the first part in Theorem 94 using the same argument and only slightly different notation.

Theorem 168.

(a) If A is symmetric, then the eigenspaces of A are orthogonal.

(b) If A is real and symmetric, then the eigenvalues of A are real.

Armin Straub
straub@southalabama.edu

Proof.

(a) We need to show that, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are eigenvectors of A with different eigenvalues, then $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{0}$.
Suppose that $A\boldsymbol{v} = \lambda \boldsymbol{v}$ and $A\boldsymbol{w} = \mu \boldsymbol{w}$ with $\lambda \neq \mu$.
Then, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \lambda \boldsymbol{v}, \boldsymbol{w} \rangle = \langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \mu \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
However, since $\lambda \neq \mu$, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ is only possible if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$.

(b) Suppose λ is a nonreal eigenvalue with nonzero eigenvector $\boldsymbol{v}$. Then, $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector and, since $\lambda \neq \bar{\lambda}$, we have two eigenvectors with different eigenvalues. By the first part, these two eigenvectors must be orthogonal in the sense that $\bar{\boldsymbol{v}}^T \boldsymbol{v} = 0$. But $\bar{\boldsymbol{v}}^T \boldsymbol{v} = \boldsymbol{v}^* \boldsymbol{v} = \|\boldsymbol{v}\|^2 \neq 0$. This shows that it is impossible to have a nonzero eigenvector for a nonreal eigenvalue.

Alternative proof. Note that a complex number λ is real if and only if $\bar{\lambda} = \lambda$. Suppose that λ is an eigenvalue with nonzero eigenvector $\boldsymbol{v}$ so that $A\boldsymbol{v} = \lambda \boldsymbol{v}$. We now observe that $\lambda \boldsymbol{v}^* \boldsymbol{v} = \boldsymbol{v}^* (\lambda \boldsymbol{v}) = \boldsymbol{v}^* A \boldsymbol{v} = \boldsymbol{v}^* A^* \boldsymbol{v} = (A\boldsymbol{v})^* \boldsymbol{v} = (\lambda \boldsymbol{v})^* \boldsymbol{v} = \bar{\lambda} \boldsymbol{v}^* \boldsymbol{v}$. Dividing by $\|\boldsymbol{v}\|^2 = \boldsymbol{v}^* \boldsymbol{v}$ (which is not zero!) we find $\lambda = \bar{\lambda}$ from which we conclude that λ is real. □

Advanced comment. Note that the alternative proof of the second part shows that any Hermitian matrix A (that is, a complex matrix A such that $A^* = A$) has only real eigenvalues. If A is Hermitian, what can we conclude about the eigenspaces if we follow the argument in the first part?

Let us highlight the following point we used in our proof:

> Let A be a real matrix. If $\boldsymbol{v}$ is a λ-eigenvector, then $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector.

See, for instance, Example 89. This is just a consequence of the basic fact that we cannot algebraically distinguish between $+i$ and $-i$.

Singular value decomposition

(Singular value decomposition)

Every $m \times n$ matrix A can be decomposed as $A = U \Sigma V^T$, where

- Σ is a (rectangular) diagonal matrix with nonnegative entries, $(m \times n)$

The diagonal entries σ_i are called the **singular values** of A.

- U is orthogonal, $(m \times m)$
- V is orthogonal. $(n \times n)$

Comment. If A is symmetric, then the singular value decomposition is already provided by the spectral theorem (the diagonalization of A). Moreover, in that case, $V = U$.

Important observations. If $A = U \Sigma V^T$, then $A^T A = V \Sigma^T \Sigma V^T$.

- Note that $\Sigma^T \Sigma$ is an $n \times n$ diagonal matrix. Its entries are σ_i^2 (the squares of the entries in Σ).
- $A^T A$ is a symmetric matrix! (Why?!) Hence, by the spectral theorem, we are able to find V and $\Sigma^T \Sigma$.

In other words, V is obtained from the (orthonormally chosen) eigenvectors of $A^T A$. Likewise, the entries of $\Sigma^T \Sigma$ are the eigenvalues of $A^T A$; their square roots are the entries of Σ, the singular values.
Finally, the equation $AV = U\Sigma$ allows us to determine U. How?! (Hint: $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$)

This results in the following **recipe** to determine the SVD $A = U\Sigma V^T$ for any matrix A.

Find an orthonormal basis of eigenvectors $\boldsymbol{v}_i$ of $A^T A$. Let λ_i be the eigenvalue of $\boldsymbol{v}_i$.

- V is the matrix with columns $\boldsymbol{v}_i$.
- Σ is the diagonal matrix with entries $\sigma_i = \sqrt{\lambda_i}$.
- U is the matrix with columns $\boldsymbol{u}_i = \frac{1}{\sigma_i} A\boldsymbol{v}_i$. If needed, fill in additional columns to make U orthogonal.

Example 169. Determine the SVD of $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

Solution. $A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ has 8-eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and 2-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Since $A^T A = V\Sigma^2 V^T$ (here, $\Sigma^T\Sigma = \Sigma^2$), we conclude that $V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{8} & \\ & \sqrt{2} \end{bmatrix}$.

From $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$, we find $\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1 = \frac{1}{\sqrt{8}}\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Likewise, $\boldsymbol{u}_2 = \frac{1}{\sigma_2} A\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Hence, $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Check that, indeed, $A = U\Sigma V^T$!

Comment. For applications, it is common to arrange the singular values in decreasing order like we did.

Comment. If we had chosen $V = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$ instead, then $U = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{8} & \\ & \sqrt{2} \end{bmatrix}$.

As with diagonalization, there are choices! (A lot fewer choices though.) This is another perfectly fine SVD. In fact, it's what Sage computes below.

Sage. Let's have Sage do the work for us. In Sage, the SVD is currently only implemented for floating point numbers. (RDF is the real numbers as floating point numbers with double precision)

```
Sage] A = matrix(RDF, [[2,2],[-1,1]])
Sage] U,S,V = A.SVD()
Sage] U
```

$$\begin{bmatrix} -1.0 & 1.11022302463 \times 10^{-16} \\ 8.64109131471 \times 10^{-17} & 1.0 \end{bmatrix}$$

```
Sage] S
```

$$\begin{bmatrix} 2.82842712475 & 0.0 \\ 0.0 & 1.41421356237 \end{bmatrix}$$

```
Sage] V
```

$$\begin{bmatrix} -0.707106781187 & -0.707106781187 \\ -0.707106781187 & 0.707106781187 \end{bmatrix}$$