Eigenvalues and Matrix Decompositions

Last modified: 2026-11-24 09:19:45 (PST)

Matrix Decompositions

Definition 0.1 (Eigenvalue)**.** Let $\mathbf{A}$ be a $p \times p$ matrix. A real number λ is an **eigenvalue** of $\mathbf{A}$ if some real vector $\underset{\sim}{v} \neq \underset{\sim}{0}$ of length p satisfies

$$\underset{p\times p}{\underset{\sim}{\mathbf{A}}} \; \underset{p\times 1}{\underset{\sim}{v}} = \lambda \; \underset{p\times 1}{\underset{\sim}{v}}$$

Definition 0.2 (Eigenvector)**.** In Definition 0.1, any such $\underset{\sim}{v}$ is an **eigenvector** of $\mathbf{A}$ for λ.

Example 0.1 (Eigenvectors of a 2×2 matrix)**.** Let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

$(1, 1)$ is an eigenvector for the eigenvalue 3:

$$\begin{aligned} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 2+1 \\ 1+2 \end{bmatrix} && \text{(definition of matrix-vector multiplication)} \\ &= 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} && \text{(factor out 3)} \end{aligned}$$

$(1, -1)$ is an eigenvector for the eigenvalue 1:

$$\begin{aligned} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= \begin{bmatrix} 2-1 \\ 1-2 \end{bmatrix} && \text{(definition of matrix-vector multiplication)} \\ &= 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} && \text{(factor out 1)} \end{aligned}$$

Remark (Multiples of an eigenvector)

Remark 0.1 (Multiples of an eigenvector). Multiplying by $\mathbf{A}$ rescales an eigenvector by λ and does not change the line it lies on. Any nonzero multiple $c\,\underset{\sim}{v}$ of an eigenvector is also an eigenvector for the same λ:

$$\begin{aligned}
\mathbf{A}(c\,\underset{\sim}{v}) &= c\,\mathbf{A}\underset{\sim}{v} && \text{(move the scalar } c \text{ to the front)} \\
&= c\,\lambda\underset{\sim}{v} && (\underset{\sim}{v} \text{ is an eigenvector for } \lambda) \\
&= \lambda\,(c\,\underset{\sim}{v}) && \text{(multiplication of numbers is commutative)}
\end{aligned}$$

For example, in Example 0.1, $2\,(1,1) = (2,2)$ is also an eigenvector for the eigenvalue 3:

$$\begin{aligned}
\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} &= \begin{bmatrix} 2 \cdot 2 + 1 \cdot 2 \\ 1 \cdot 2 + 2 \cdot 2 \end{bmatrix} && \text{(definition of matrix-vector multiplication)} \\
&= \begin{bmatrix} 4 + 2 \\ 2 + 4 \end{bmatrix} && \text{(multiply)} \\
&= \begin{bmatrix} 6 \\ 6 \end{bmatrix} && \text{(add)} \\
&= 3 \begin{bmatrix} 2 \\ 2 \end{bmatrix} && \text{(factor out 3)}
\end{aligned}$$

Example 0.2 (A vector that is not an eigenvector)**.** For the same $\mathbf{A}$, $(1,0)$ is not an eigenvector:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

and no number λ gives $(2,1) = \lambda\,(1,0)$, because the second entries would need $1 = \lambda \cdot 0$.

Remark (Real and complex eigenvalues)

Remark 0.2 (Real and complex eigenvalues). These notes take λ and $\underset{\sim}{v}$ to be real. Banerjee and Roy [1, Ch. 11], Definition 11.1, pp. 312-313 also allows complex eigenvalues and eigenvectors, because some real matrices, such as a rotation by 90 degrees, have no real eigenvalues [1, Ch. 11], Example 11.2, p. 312.

For example, $\mathbf{R} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotates each vector in the plane counterclockwise by 90 degrees: $\mathbf{R}\,(v_1, v_2) = (-v_2, v_1)$. Suppose $\mathbf{R}\underset{\sim}{v} = \lambda\underset{\sim}{v}$ for a real number λ. Matching entries gives $-v_2 = \lambda v_1$ and $v_1 = \lambda v_2$. Substituting the second equation into the first gives $-v_2 = \lambda^2 v_2$, so $(1 + \lambda^2)\,v_2 = 0$. Since $1 + \lambda^2 > 0$, $v_2 = 0$, and then

$$
\begin{aligned}
v_1 &= \lambda v_2 \\
&= 0.
\end{aligned}
$$

So the only solution is $\underset{\sim}{v} = \underset{\sim}{0}$, and $\mathbf{R}$ has no real eigenvalue.

Video

Chapter 14 of the *Essence of linear algebra* series by 3Blue1Brown shows the geometric meaning of eigenvectors and eigenvalues:

- Eigenvectors and eigenvalues shows how certain vectors maintain their direction during a linear transformation and are merely stretched or squished by an eigenvalue factor (Definition 0.1).

Theorem 0.1 (Spectral theorem for symmetric matrices)**.** *If* $\mathbf{A}$ *is a* $p \times p$ *symmetric matrix (Definition 3 in Special Matrices, Quadratic Forms and Trace) with real entries, then there are a* $p \times p$ *orthogonal matrix* $\mathbf{Q}$ *(Definition 10 in Special Matrices, Quadratic Forms and Trace) and a* $p \times p$ *diagonal matrix* $\mathbf{\Lambda}$ *(Definition 4 in Special Matrices, Quadratic Forms and Trace) with real diagonal entries* $\lambda_1, \ldots, \lambda_p$ *such that*

$$
\underset{p\times p}{\mathbf{A}} = \underset{p\times p}{\mathbf{Q}}\ \underset{p\times p}{\mathbf{\Lambda}}\ \underset{p\times p}{\mathbf{Q}^\top}
$$

Each λ_i *is an eigenvalue of* $\mathbf{A}$ *(Definition 0.1), and column* i *of* $\mathbf{Q}$ *is an eigenvector of* $\mathbf{A}$ *for* λ_i*.*

Remark (An equivalent form of the spectral theorem)

Remark 0.3 (An equivalent form of the spectral theorem). The proof, by induction on p, is outside the scope of these notes; see Banerjee and Roy [1, Ch. 11], Theorem 11.27, p. 349, which states the result in the equivalent form $\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \mathbf{\Lambda}$. The two forms are equivalent because

$$
\begin{aligned}
\mathbf{Q}^\top \mathbf{Q} &= \mathbf{Q}\mathbf{Q}^\top \\
&= \mathbf{I}_p
\end{aligned}
$$

(Remark 6 in Special Matrices, Quadratic Forms and Trace). Multiplying $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$ by $\mathbf{Q}^\top$ on the left and by $\mathbf{Q}$ on the right gives

$$
\begin{aligned}
\mathbf{Q}^\top \mathbf{A}\mathbf{Q} &= \mathbf{Q}^\top(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top)\mathbf{Q} && \text{(substitute)} \\
&= (\mathbf{Q}^\top\mathbf{Q})\mathbf{\Lambda}(\mathbf{Q}^\top\mathbf{Q}) && \text{(regroup; matrix multiplication is associative)} \\
&= \mathbf{I}_p\mathbf{\Lambda}\mathbf{I}_p && (\mathbf{Q} \text{ is orthogonal}) \\
&= \mathbf{\Lambda} && \text{(multiplying by } \mathbf{I}_p \text{ changes nothing)}
\end{aligned}
$$

Multiplying $\mathbf{Q}^\top\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ by $\mathbf{Q}$ on the left and by $\mathbf{Q}^\top$ on the right gives back the first form:

$$
\begin{aligned}
\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top &= \mathbf{Q}(\mathbf{Q}^\top\mathbf{A}\mathbf{Q})\mathbf{Q}^\top && \text{(substitute)} \\
&= (\mathbf{Q}\mathbf{Q}^\top)\mathbf{A}(\mathbf{Q}\mathbf{Q}^\top) && \text{(regroup; matrix multiplication is associative)} \\
&= \mathbf{I}_p\mathbf{A}\mathbf{I}_p && (\mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_p) \\
&= \mathbf{A} && \text{(multiplying by } \mathbf{I}_p \text{ changes nothing)}
\end{aligned}
$$

For example, let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, which is symmetric, so $\mathbf{Q}^\top = \mathbf{Q}$. Then

$$
\begin{aligned}
\mathbf{Q}^\top\mathbf{A}\mathbf{Q} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(substitute)} \\
&= \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(move the scalars to the front)} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(multiply the scalars)} \\
&= \frac{1}{2}\begin{bmatrix} 1\cdot 2 + 1\cdot 1 & 1\cdot 1 + 1\cdot 2 \\ 1\cdot 2 + (-1)\cdot 1 & 1\cdot 1 + (-1)\cdot 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(definition of matrix multiplication, first two matric} \\
&= \frac{1}{2}\begin{bmatrix} 2+1 & 1+2 \\ 2-1 & 1-2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{2}\begin{bmatrix} 3 & 3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(add)} \\
&= \frac{1}{2}\begin{bmatrix} 3\cdot 1 + 3\cdot 1 & 3\cdot 1 + 3\cdot(-1) \\ 1\cdot 1 + (-1)\cdot 1 & 1\cdot 1 + (-1)\cdot(-1) \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \frac{1}{2}\begin{bmatrix} 3+3 & 3-3 \\ 1-1 & 1+1 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{2}\begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} && \text{(add)} \\
&= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} && \text{(multiply by } \tfrac{1}{2}\text{)}
\end{aligned}
$$

a diagonal matrix with the eigenvalues 3 and 1 of $\mathbf{A}$ from Example 0.1 on its diagonal.

Theorem 0.2 (Eigenvectors of a symmetric matrix for distinct eigenvalues are orthogonal)**.** *Let* $\mathbf{A}$ *be a symmetric* $p \times p$ *matrix (Definition 3 in Special Matrices, Quadratic Forms and Trace), and let* $\underset{\sim}{u}$ *and* $\underset{\sim}{v}$ *be eigenvectors of* $\mathbf{A}$ *(Definition 0.2) for eigenvalues* $\lambda_1 \neq \lambda_2$*. Then* $\underset{\sim}{u} \perp \underset{\sim}{v}$ *(Definition 22 in Vectors).*

Proof

Proof.

$$\begin{aligned}
\lambda_1 (\underset{\sim}{u} \cdot \underset{\sim}{v}) &= (\lambda_1 \underset{\sim}{u}) \cdot \underset{\sim}{v} && \text{(Definition 10 in Vectors)} \\
&= (\mathbf{A}\underset{\sim}{u}) \cdot \underset{\sim}{v} && (\underset{\sim}{u} \text{ is an eigenvector for } \lambda_1) \\
&= (\mathbf{A}\underset{\sim}{u})^\top \underset{\sim}{v} && \text{(Definition 10 in Vectors)} \\
&= \underset{\sim}{u}^\top \mathbf{A}^\top \underset{\sim}{v} && \text{(Theorem 6 in Matrix Multiplication)} \\
&= \underset{\sim}{u}^\top \mathbf{A}\, \underset{\sim}{v} && (\mathbf{A} \text{ is symmetric}) \\
&= \underset{\sim}{u}^\top (\lambda_2 \underset{\sim}{v}) && (\underset{\sim}{v} \text{ is an eigenvector for } \lambda_2) \\
&= \lambda_2 (\underset{\sim}{u} \cdot \underset{\sim}{v}). && \text{(Definition 10 in Vectors)}
\end{aligned}$$

So $(\lambda_1 - \lambda_2)(\underset{\sim}{u} \cdot \underset{\sim}{v}) = 0$, and $\lambda_1 - \lambda_2 \neq 0$, so $\underset{\sim}{u} \cdot \underset{\sim}{v} = 0$. □

Example 0.3 (Eigenvectors of a symmetric 2×2 matrix)**.** $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is symmetric. Its eigenvectors $(1, 1)$ for $\lambda_1 = 3$ and $(1, -1)$ for $\lambda_2 = 1$ (Example 0.1) satisfy

$$\begin{aligned}
(1, 1) \cdot (1, -1) &= 1 \cdot 1 + 1 \cdot (-1) \\
&= 0.
\end{aligned}$$

Definition 0.3 (Eigendecomposition, spectral decomposition)**.** Let $\mathbf{A}$ be a $p \times p$ symmetric matrix with real entries. An **eigendecomposition** [2], or **spectral decomposition**, of $\mathbf{A}$ is a factorization

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$$

with $\mathbf{Q}$ a $p \times p$ orthogonal matrix (Definition 10 in Special Matrices, Quadratic Forms and Trace) and $\mathbf{\Lambda}$ a $p \times p$ diagonal matrix (Definition 4 in Special Matrices, Quadratic Forms and Trace).

Example 0.4 (An eigendecomposition of a 2×2 symmetric matrix)**.** For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, Example 0.1 found the eigenvectors $(1, 1)$ for 3 and $(1, -1)$ for 1. They are orthogonal (Definition 22 in Vectors):

$$
\begin{aligned}
(1,1)\cdot(1,-1) &= 1\cdot 1 + 1\cdot(-1) && \text{(definition of the dot product)} \\
&= 1 - 1 && \text{(multiply)} \\
&= 0 && \text{(add)}
\end{aligned}
$$

Each has norm $\sqrt{1^2+1^2} = \sqrt{2}$ (Definition 23 in Vectors), so dividing each by $\sqrt{2}$ gives two orthonormal eigenvectors (Definition 28 in Vectors). Put them in the columns of $\mathbf{Q}$, and the matching eigenvalues on the diagonal of $\mathbf{\Lambda}$:

$$
\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \mathbf{\Lambda} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}
$$

This $\mathbf{Q}$ is symmetric, so $\mathbf{Q}^\top = \mathbf{Q}$. Then

$$
\begin{aligned}
\mathbf{Q\Lambda Q}^\top &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(substitute)} \\
&= \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(move the scalars to the front)} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(multiply the scalars)} \\
&= \frac{1}{2}\begin{bmatrix} 1\cdot 3 + 1\cdot 0 & 1\cdot 0 + 1\cdot 1 \\ 1\cdot 3 + (-1)\cdot 0 & 1\cdot 0 + (-1)\cdot 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(definition of matrix multiplication, first two matrice} \\
&= \frac{1}{2}\begin{bmatrix} 3+0 & 0+1 \\ 3+0 & 0-1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{2}\begin{bmatrix} 3 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(add)} \\
&= \frac{1}{2}\begin{bmatrix} 3\cdot 1 + 1\cdot 1 & 3\cdot 1 + 1\cdot(-1) \\ 3\cdot 1 + (-1)\cdot 1 & 3\cdot 1 + (-1)\cdot(-1) \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \frac{1}{2}\begin{bmatrix} 3+1 & 3-1 \\ 3-1 & 3+1 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{2}\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} && \text{(add)} \\
&= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} && \text{(multiply by } \tfrac{1}{2}\text{)}
\end{aligned}
$$

which is $\mathbf{A}$.

Remark (An eigendecomposition is not unique)

Remark 0.4 (An eigendecomposition is not unique). A symmetric matrix can have more than one eigendecomposition. Reordering the eigenvalues on the diagonal of $\mathbf{\Lambda}$, and the columns of $\mathbf{Q}$ with them, gives another one, and so does multiplying a column of $\mathbf{Q}$ by -1.

For example, take $\mathbf{A}$ from Example 0.4. Swapping the two eigenvalues and the two columns gives

$$\mathbf{Q}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad \mathbf{\Lambda}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

and multiplying the second column of $\mathbf{Q}$ by -1 gives

$$\begin{aligned} \mathbf{Q}_2 &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ \mathbf{\Lambda}_2 &= \mathbf{\Lambda} \\ &= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

The columns of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the columns of $\mathbf{Q}$, reordered or multiplied by -1, so they are still orthonormal, and $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are orthogonal. Both products give back $\mathbf{A}$:

$$
\begin{aligned}
\mathbf{Q}_1\mathbf{\Lambda}_1{\mathbf{Q}_1}^\top &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\right)^\top && \text{(substitute)}\\
&= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(definition of the transpose)}\\
&= \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(move the scalars to the front)}\\
&= \frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(multiply the scalars)}\\
&= \frac{1}{2}\begin{bmatrix}1\cdot 1+1\cdot 0 & 1\cdot 0+1\cdot 3\\ -1\cdot 1+1\cdot 0 & -1\cdot 0+1\cdot 3\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(definition of matrix multiplication, first two matrice}\\
&= \frac{1}{2}\begin{bmatrix}1+0 & 0+3\\ -1+0 & 0+3\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(multiply)}\\
&= \frac{1}{2}\begin{bmatrix}1 & 3\\ -1 & 3\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} && \text{(add)}\\
&= \frac{1}{2}\begin{bmatrix}1\cdot 1+3\cdot 1 & 1\cdot(-1)+3\cdot 1\\ -1\cdot 1+3\cdot 1 & -1\cdot(-1)+3\cdot 1\end{bmatrix} && \text{(definition of matrix multiplication)}\\
&= \frac{1}{2}\begin{bmatrix}1+3 & -1+3\\ -1+3 & 1+3\end{bmatrix} && \text{(multiply)}\\
&= \frac{1}{2}\begin{bmatrix}4 & 2\\ 2 & 4\end{bmatrix} && \text{(add)}\\
&= \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix} && \text{(multiply by } \tfrac{1}{2}\text{)}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Q}_2\mathbf{\Lambda}_2{\mathbf{Q}_2}^\top &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\right)^\top && \text{(substitute)}\\
&= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(definition of the transpose)}\\
&= \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(move the scalars to the front)}\\
&= \frac{1}{2}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(multiply the scalars)}\\
&= \frac{1}{2}\begin{bmatrix}1\cdot 3+(-1)\cdot 0 & 1\cdot 0+(-1)\cdot 1\\ 1\cdot 3+1\cdot 0 & 1\cdot 0+1\cdot 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(definition of matrix multiplication, first two mat}\\
&= \frac{1}{2}\begin{bmatrix}3+0 & 0-1\\ 3+0 & 0+1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(multiply)}\\
&= \frac{1}{2}\begin{bmatrix}3 & -1\\ 3 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} && \text{(add)}\\
&= \frac{1}{2}\begin{bmatrix}3\cdot 1+(-1)\cdot(-1) & 3\cdot 1+(-1)\cdot 1\\ 3\cdot 1+1\cdot(-1) & 3\cdot 1+1\cdot 1\end{bmatrix} && \text{(definition of matrix multiplication)}\\
&= \frac{1}{2}\begin{bmatrix}3+1 & 3-1\\ 3-1 & 3+1\end{bmatrix} && \text{(multiply)}\\
&= \frac{1}{2}\begin{bmatrix}4 & 2\\ 2 & 4\end{bmatrix} && \text{(add)}\\
&= \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix} && \text{(multiply by } \tfrac{1}{2}\text{)}
\end{aligned}
$$

Theorem 0.3 (Singular value decomposition)**.** *Let* $\mathbf{A}$ *be an* $n \times p$ *matrix with real entries and* $\text{rank}(\mathbf{A}) = r$ *(Definition 2 in Subspaces and Rank). Then there are an* $n \times n$ *orthogonal matrix* $\mathbf{U}$*, a* $p \times p$ *orthogonal matrix* $\mathbf{V}$ *(Definition 10 in Special Matrices, Quadratic Forms and Trace), and numbers* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ *such that*

$$\underset{n\times p}{\mathbf{A}} = \underset{n\times n}{\mathbf{U}}\ \underset{n\times p}{\mathbf{D}}\ \underset{p\times p}{\mathbf{V}^\top}$$

where $\mathbf{D}$ *has entries* $d_{ii} = \sigma_i$ *for* $i = 1, \ldots, r$ *and every other entry* 0*.*

Remark (The SVD applies to every real matrix)

Remark 0.5 (The SVD applies to every real matrix). The proof is outside the scope of these notes; see Banerjee and Roy [1, Ch. 12], Theorem 12.1, p. 373. Unlike the spectral theorem (Theorem 0.1), Theorem 0.3 applies to every real matrix, including one that is not square or not symmetric.
For example, the 1×2 matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ is not square, so Theorem 0.1 does not apply to it. It has rank 1, and Theorem 0.3 holds with $\mathbf{U} = \begin{bmatrix} 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} \sqrt{2} & 0 \end{bmatrix}$, and $\mathbf{V}$ the orthogonal matrix $\mathbf{Q}$ from Example 0.4, which is symmetric, so $\mathbf{V}^\top = \mathbf{V}$:

$$
\begin{aligned}
\mathbf{U}\mathbf{D}\mathbf{V}^\top &= \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(substitute)} \\
&= \begin{bmatrix} \sqrt{2} & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(multiplying by the } 1 \times 1 \text{ identity changes nothing)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(move the scalar to the front)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} \cdot 1 + 0 \cdot 1 & \sqrt{2} \cdot 1 + 0 \cdot (-1) \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} + 0 & \sqrt{2} + 0 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} & \sqrt{2} \end{bmatrix} && \text{(add)} \\
&= \begin{bmatrix} 1 & 1 \end{bmatrix} && \text{(divide by } \sqrt{2}\text{)}
\end{aligned}
$$

$\mathbf{U}$ is orthogonal because

$$
\begin{aligned}
\mathbf{U}^\top \mathbf{U} &= \begin{bmatrix} 1 \end{bmatrix} \\
&= \mathbf{I}_1,
\end{aligned}
$$

and the single singular value is $\sigma_1 = \sqrt{2}$.

Definition 0.4 (Singular value decomposition)**.** Let $\mathbf{A}$ be an $n \times p$ matrix with real entries and $\operatorname{rank}(\mathbf{A}) = r$. A **singular value decomposition** [3] (SVD) of $\mathbf{A}$ is a factorization $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ with $\mathbf{U}$, $\mathbf{D}$ and $\mathbf{V}$ as in Theorem 0.3.

Definition 0.5 (Singular values)**.** The numbers $\sigma_1 \geq \cdots \geq \sigma_r > 0$ on the diagonal of $\mathbf{D}$ in a singular value decomposition of $\mathbf{A}$ (Definition 0.4) are the **singular values** of $\mathbf{A}$.

Example 0.5 (An SVD of a 3×2 matrix)**.** Let

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \qquad \mathbf{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$\mathbf{V}$ is the orthogonal matrix $\mathbf{Q}$ from Example 0.4. $\mathbf{U}$ is orthogonal too. Entry (i, j) of $\mathbf{U}^\top\mathbf{U}$ is the dot product of columns i and j of $\mathbf{U}$, and those columns are $\underset{\sim}{u}_1 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $\underset{\sim}{u}_2 = (0, 1, 0)$ and $\underset{\sim}{u}_3 = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$:

$$\begin{aligned} \underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1 &= \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}} + 0 \cdot 0 + \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}} && \text{(definition of the dot product)} \\ &= \tfrac{1}{2} + 0 + \tfrac{1}{2} && \text{(multiply)} \\ &= 1 && \text{(add)} \end{aligned}$$

$$\begin{aligned} \underset{\sim}{u}_2 \cdot \underset{\sim}{u}_2 &= 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 && \text{(definition of the dot product)} \\ &= 0 + 1 + 0 && \text{(multiply)} \\ &= 1 && \text{(add)} \end{aligned}$$

$$\begin{aligned} \underset{\sim}{u}_3 \cdot \underset{\sim}{u}_3 &= \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}} + 0 \cdot 0 + \left(-\tfrac{1}{\sqrt{2}}\right) \cdot \left(-\tfrac{1}{\sqrt{2}}\right) && \text{(definition of the dot product)} \\ &= \tfrac{1}{2} + 0 + \tfrac{1}{2} && \text{(multiply)} \\ &= 1 && \text{(add)} \end{aligned}$$

$$\begin{aligned} \underset{\sim}{u}_1 \cdot \underset{\sim}{u}_2 &= \tfrac{1}{\sqrt{2}} \cdot 0 + 0 \cdot 1 + \tfrac{1}{\sqrt{2}} \cdot 0 && \text{(definition of the dot product)} \\ &= 0 + 0 + 0 && \text{(multiply)} \\ &= 0 && \text{(add)} \end{aligned}$$

$$\begin{aligned} \underset{\sim}{u}_1 \cdot \underset{\sim}{u}_3 &= \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}} + 0 \cdot 0 + \tfrac{1}{\sqrt{2}} \cdot \left(-\tfrac{1}{\sqrt{2}}\right) && \text{(definition of the dot product)} \\ &= \tfrac{1}{2} + 0 - \tfrac{1}{2} && \text{(multiply)} \\ &= 0 && \text{(add)} \end{aligned}$$

$$\begin{aligned} \underset{\sim}{u}_2 \cdot \underset{\sim}{u}_3 &= 0 \cdot \tfrac{1}{\sqrt{2}} + 1 \cdot 0 + 0 \cdot \left(-\tfrac{1}{\sqrt{2}}\right) && \text{(definition of the dot product)} \\ &= 0 + 0 + 0 && \text{(multiply)} \\ &= 0 && \text{(add)} \end{aligned}$$

The dot product is symmetric, so these six values fill in all nine entries, and $\mathbf{U}^\top\mathbf{U} = \mathbf{I}_3$. $\mathbf{V}$ is symmetric, so $\mathbf{V}^\top = \mathbf{V}$. Then

$$
\begin{aligned}
\mathbf{DV}^\top &= \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(substitute)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(move the scalar to the front)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \cdot 1 + 0 \cdot 1 & 2 \cdot 1 + 0 \cdot (-1) \\ 0 \cdot 1 + \sqrt{2} \cdot 1 & 0 \cdot 1 + \sqrt{2} \cdot (-1) \\ 0 \cdot 1 + 0 \cdot 1 & 0 \cdot 1 + 0 \cdot (-1) \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 + 0 & 2 + 0 \\ 0 + \sqrt{2} & 0 - \sqrt{2} \\ 0 + 0 & 0 + 0 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ \sqrt{2} & -\sqrt{2} \\ 0 & 0 \end{bmatrix} && \text{(add)} \\
&= \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{bmatrix} && \text{(divide by } \sqrt{2}\text{)}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{UDV}^\top &= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{bmatrix} && \text{(substitute)} \\
&= \begin{bmatrix} \frac{1}{\sqrt{2}} \cdot \sqrt{2} + 0 \cdot 1 + \frac{1}{\sqrt{2}} \cdot 0 & \frac{1}{\sqrt{2}} \cdot \sqrt{2} + 0 \cdot (-1) + \frac{1}{\sqrt{2}} \cdot 0 \\ 0 \cdot \sqrt{2} + 1 \cdot 1 + 0 \cdot 0 & 0 \cdot \sqrt{2} + 1 \cdot (-1) + 0 \cdot 0 \\ \frac{1}{\sqrt{2}} \cdot \sqrt{2} + 0 \cdot 1 - \frac{1}{\sqrt{2}} \cdot 0 & \frac{1}{\sqrt{2}} \cdot \sqrt{2} + 0 \cdot (-1) - \frac{1}{\sqrt{2}} \cdot 0 \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \begin{bmatrix} 1 + 0 + 0 & 1 + 0 + 0 \\ 0 + 1 + 0 & 0 - 1 + 0 \\ 1 + 0 - 0 & 1 + 0 - 0 \end{bmatrix} && \text{(multiply)} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} && \text{(add)}
\end{aligned}
$$

which is $\mathbf{A}$. $\mathbf{A}$ has rank 2, because its two columns are linearly independent, and its singular values are $\sigma_1 = 2$ and $\sigma_2 = \sqrt{2}$.

Remark (An SVD is not unique)

Remark 0.6 (An SVD is not unique). An SVD is not unique [1, Ch. 12], Examples 12.2 and 12.3, p. 378. For any $i \leq r$, multiplying column i of both $\mathbf{U}$ and $\mathbf{V}$ by -1 gives another one. The singular values do not depend on which SVD is chosen [1, Ch. 12], pp. 371 and 378.
For example, in Example 0.5, multiplying the first columns of $\mathbf{U}$ and $\mathbf{V}$ by -1 gives

$$\mathbf{U}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad \mathbf{V}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$$

Their columns are the columns of $\mathbf{U}$ and $\mathbf{V}$, some multiplied by -1, so they are still orthonormal. With the same $\mathbf{D}$:

$$\begin{aligned}
\mathbf{D}{\mathbf{V}_1}^\top &= \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \right)^\top && \text{(substitute)} \\
&= \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} && \text{(definition of the transpose)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} && \text{(move the scalar to the front)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \cdot (-1) + 0 \cdot 1 & 2 \cdot (-1) + 0 \cdot (-1) \\ 0 \cdot (-1) + \sqrt{2} \cdot 1 & 0 \cdot (-1) + \sqrt{2} \cdot (-1) \\ 0 \cdot (-1) + 0 \cdot 1 & 0 \cdot (-1) + 0 \cdot (-1) \end{bmatrix} && \text{(definition of matrix multiplication)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} -2 + 0 & -2 + 0 \\ 0 + \sqrt{2} & 0 - \sqrt{2} \\ 0 + 0 & 0 + 0 \end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} -2 & -2 \\ \sqrt{2} & -\sqrt{2} \\ 0 & 0 \end{bmatrix} && \text{(add)} \\
&= \begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{bmatrix} && \text{(divide by } \sqrt{2}\text{)}
\end{aligned}$$

and

$$
\begin{aligned}
\mathbf{U}_1\mathbf{D}\mathbf{V}_1{}^\top &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{bmatrix} && \text{(substitute)} \\
&= \begin{bmatrix} -\frac{1}{\sqrt{2}} \cdot \left(-\sqrt{2}\right) + 0 \cdot 1 + \frac{1}{\sqrt{2}} \cdot 0 & -\frac{1}{\sqrt{2}} \cdot \left(-\sqrt{2}\right) + 0 \cdot (-1) + \frac{1}{\sqrt{2}} \cdot 0 \\ 0 \cdot \left(-\sqrt{2}\right) + 1 \cdot 1 + 0 \cdot 0 & 0 \cdot \left(-\sqrt{2}\right) + 1 \cdot (-1) + 0 \cdot 0 \\ -\frac{1}{\sqrt{2}} \cdot \left(-\sqrt{2}\right) + 0 \cdot 1 - \frac{1}{\sqrt{2}} \cdot 0 & -\frac{1}{\sqrt{2}} \cdot \left(-\sqrt{2}\right) + 0 \cdot (-1) - \frac{1}{\sqrt{2}} \cdot 0 \end{bmatrix} && \text{(definition of matrix multipl} \\
&= \begin{bmatrix} 1+0+0 & 1+0+0 \\ 0+1+0 & 0-1+0 \\ 1+0-0 & 1+0-0 \end{bmatrix} && \text{(multiply)} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} && \text{(add)}
\end{aligned}
$$

which is $\mathbf{A}$ again. Both SVDs have the singular values $\sigma_1 = 2$ and $\sigma_2 = \sqrt{2}$.

i Remark (Counting zero singular values)

Remark 0.7 (Counting zero singular values). Some texts, and R's `svd()`, also count $\min(n, p) - r$ singular values equal to 0, so that every $n \times p$ matrix has $\min(n, p)$ singular values.

For example, $\mathbf{B} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$ has rank $r = 1$. Taking

$$
\begin{aligned}
\mathbf{U} &= \mathbf{V} \\
&= \mathbf{I}_2
\end{aligned}
$$

and $\mathbf{D} = \mathbf{B}$ gives an SVD $\mathbf{B} = \mathbf{I}_2\mathbf{B}\mathbf{I}_2{}^\top$, so by Definition 0.5, $\mathbf{B}$ has one singular value, $\sigma_1 = 3$. R's `svd()` reports $\min(2, 2) = 2$ singular values for $\mathbf{B}$: 3 and 0.

Theorem 0.4 (An SVD gives an eigendecomposition of $\mathbf{A}^\top\mathbf{A}$). *Let $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ be a singular value decomposition (Definition 0.4) of an $n \times p$ matrix $\mathbf{A}$ with singular values $\sigma_1, \ldots, \sigma_r$. Then $\mathbf{A}^\top\mathbf{A}$ is symmetric (Definition 3 in Special Matrices, Quadratic Forms and Trace), and*

$$
\underbrace{\mathbf{A}^\top\mathbf{A}}_{p\times p} = \underbrace{\mathbf{V}}_{p\times p}\ \underbrace{\mathbf{\Lambda}}_{p\times p}\ \underbrace{\mathbf{V}^\top}_{p\times p}
$$

where $\mathbf{\Lambda} \overset{def}{=} \mathbf{D}^\top\mathbf{D}$ is the $p \times p$ diagonal matrix whose i-th diagonal entry λ_i is σ_i^2 for $i \leq r$ and 0 for $i > r$. So $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ is an eigendecomposition (Definition 0.3) of $\mathbf{A}^\top\mathbf{A}$: column i

of $\mathbf{V}$ *is an eigenvector of* $\mathbf{A}^\top\mathbf{A}$ *for the eigenvalue* λ_i.

Proof

Proof. **Symmetry.**

$$\begin{aligned}(\mathbf{A}^\top\mathbf{A})^\top &= \mathbf{A}^\top(\mathbf{A}^\top)^\top && \text{(transpose of a product)}\\ &= \mathbf{A}^\top\mathbf{A} && \text{(transposing twice changes nothing)}\end{aligned}$$

The first step is Theorem 6 in Matrix Multiplication, and the second follows from Definition 6 in Matrices, which swaps rows and columns, so swapping them again restores $\mathbf{A}$.

The factorization.

$$\begin{aligned}\mathbf{A}^\top\mathbf{A} &= (\mathbf{U}\mathbf{D}\mathbf{V}^\top)^\top\mathbf{U}\mathbf{D}\mathbf{V}^\top && \text{(substitute the SVD)}\\ &= (\mathbf{D}\mathbf{V}^\top)^\top\mathbf{U}^\top\mathbf{U}\mathbf{D}\mathbf{V}^\top && \text{(transpose of the product of } \mathbf{U} \text{ and } \mathbf{D}\mathbf{V}^\top)\\ &= (\mathbf{V}^\top)^\top\mathbf{D}^\top\mathbf{U}^\top\mathbf{U}\mathbf{D}\mathbf{V}^\top && \text{(transpose of the product of } \mathbf{D} \text{ and } \mathbf{V}^\top)\\ &= \mathbf{V}\mathbf{D}^\top\mathbf{U}^\top\mathbf{U}\mathbf{D}\mathbf{V}^\top && \text{(transposing twice changes nothing)}\\ &= \mathbf{V}\mathbf{D}^\top(\mathbf{U}^\top\mathbf{U})\mathbf{D}\mathbf{V}^\top && \text{(regroup; matrix multiplication is associative)}\\ &= \mathbf{V}\mathbf{D}^\top\mathbf{I}_n\mathbf{D}\mathbf{V}^\top && (\mathbf{U} \text{ is orthogonal})\\ &= \mathbf{V}(\mathbf{D}^\top\mathbf{D})\mathbf{V}^\top && \text{(identity matrix)}\\ &= \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top && \text{(definition of } \mathbf{\Lambda})\end{aligned}$$

The entries of $\mathbf{\Lambda}$. Write λ_{ij} for entry (i,j) of $\mathbf{\Lambda}$, so $\lambda_i = \lambda_{ii}$. Only the entries d_{kk} with $k \le r$ of $\mathbf{D}$ can be nonzero.

$$\begin{aligned}\lambda_{ij} &= \sum_{k=1}^{n}(\mathbf{D}^\top)_{ik}\,d_{kj} && \text{(definition of matrix multiplication)}\\ &= \sum_{k=1}^{n} d_{ki}\,d_{kj} && \text{(definition of the transpose)}\end{aligned}$$

The term for k is nonzero only if $k = i \le r$ and $k = j$. So $\lambda_{ij} = 0$ when $i \ne j$. On the diagonal, only the term $k = i$ can be nonzero, and for $i \le r$

$$\begin{aligned}\lambda_i &= \lambda_{ii} && \text{(notation)}\\ &= d_{ii}\,d_{ii} && \text{(the only term that can be nonzero)}\\ &= \sigma_i \cdot \sigma_i && \text{(definition of } \mathbf{D})\\ &= \sigma_i^2 && \text{(notation for a square)}\end{aligned}$$

For $i > r$, every term is 0, so $\lambda_i = 0$.

The eigenvectors. Write $\underset{\sim}{v}_i$ for column i of $\mathbf{V}$, and $\underset{\sim}{e}_i$ for the indicator vector (Definition 21 in Vectors) of length p. Column i of $\mathbf{V}^\top\mathbf{V} = \mathbf{I}_p$ says $\mathbf{V}^\top\underset{\sim}{v}_i = \underset{\sim}{e}_i$. Then

$$\begin{aligned}
\mathbf{A}^\top\mathbf{A}\,\underset{\sim}{v}_i &= \mathbf{V}\mathbf{\Lambda}\,\mathbf{V}^\top\underset{\sim}{v}_i && \text{(the factorization)}\\
&= \mathbf{V}\mathbf{\Lambda}\,\underset{\sim}{e}_i && (\mathbf{V}^\top\underset{\sim}{v}_i = \underset{\sim}{e}_i)\\
&= \mathbf{V}\,(\lambda_i\,\underset{\sim}{e}_i) && \text{(column } i \text{ of the diagonal matrix } \mathbf{\Lambda})\\
&= \lambda_i\,\mathbf{V}\underset{\sim}{e}_i && \text{(move the scalar } \lambda_i \text{ to the front)}\\
&= \lambda_i\,\underset{\sim}{v}_i && (\mathbf{V}\underset{\sim}{e}_i \text{ is column } i \text{ of } \mathbf{V})
\end{aligned}$$

and $\underset{\sim}{v}_i \neq \underset{\sim}{0}$ because it has norm 1. □

Example 0.6 ($\mathbf{A}^\top\mathbf{A}$ for the matrix of the SVD example)**.** For $\mathbf{A}$ in Example 0.5:

$$\begin{aligned}
\mathbf{A}^\top\mathbf{A} &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} && \text{(definition of the transpose)}\\
&= \begin{bmatrix} 1+1+1 & 1-1+1 \\ 1-1+1 & 1+1+1 \end{bmatrix} && \text{(definition of matrix multiplication)}\\
&= \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} && \text{(add)}
\end{aligned}$$

Theorem 0.4 says its eigenvalues are $\sigma_1^2 = 4$ and $\sigma_2^2 = 2$, with the columns of $\mathbf{V}$ as eigenvectors. Checking the first column, up to its factor $\frac{1}{\sqrt{2}}$:

$$\begin{aligned}
\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 3\cdot 1 + 1\cdot 1 \\ 1\cdot 1 + 3\cdot 1 \end{bmatrix} && \text{(definition of matrix-vector multiplication)}\\
&= \begin{bmatrix} 3+1 \\ 1+3 \end{bmatrix} && \text{(multiply)}\\
&= \begin{bmatrix} 4 \\ 4 \end{bmatrix} && \text{(add)}\\
&= 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} && \text{(factor out 4)}
\end{aligned}$$

and the second:

$$
\begin{aligned}
\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= \begin{bmatrix} 3 \cdot 1 + 1 \cdot (-1) \\ 1 \cdot 1 + 3 \cdot (-1) \end{bmatrix} && \text{(definition of matrix-vector multiplication)} \\
&= \begin{bmatrix} 3 - 1 \\ 1 - 3 \end{bmatrix} && \text{(multiply)} \\
&= \begin{bmatrix} 2 \\ -2 \end{bmatrix} && \text{(add)} \\
&= 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} && \text{(factor out 2)}
\end{aligned}
$$

Remark (Building an SVD from an eigendecomposition)

Remark 0.8 (Building an SVD from an eigendecomposition). Theorem 0.4 matches Banerjee and Roy [1, Ch. 12], pp. 372-373, which works in the other direction: it constructs an SVD from a spectral decomposition of $\mathbf{A}^\top \mathbf{A}$ and sets $\sigma_i = \sqrt{\lambda_i}$.
For example, in Example 0.6 the eigenvalues of $\mathbf{A}^\top \mathbf{A}$ are 4 and 2, and $\sqrt{4} = 2$ and $\sqrt{2}$ are the singular values σ_1 and σ_2 of $\mathbf{A}$ from Example 0.5.

[1] S. Banerjee and A. Roy, *Linear algebra and matrix analysis for statistics*, vol. 181. Crc Press Boca Raton, 2014. Available: https://www.routledge.com/Linear-Algebra-and-Matrix-Analysis-for-Statistics/Banerjee-Roy/p/book/9781420095388

[2] Wikipedia contributors, "Eigendecomposition of a matrix — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

[3] Wikipedia contributors, "Singular value decomposition — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Singular_value_decomposition