Mappings of Matrices with Real Eigenvalues

CEE 629. System Identification
Duke University, Spring 2021

The matrix equation $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ defines a mapping of a vector $\boldsymbol{x}$ to another vector $\boldsymbol{y}$. For real symmetric matrices $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ and all vectors $\boldsymbol{x} \in \mathbb{R}^2$ s.t. $x_1^2 + x_2^2 = 1$, a visualization of this mapping helps in interpreting solutions to the standard eigenvalue problem $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}\boldsymbol{\lambda}$.

Matrices with real eigenvalues may be symmetric or non symmetric. All symmetric matrices have real-valued eigenvalues and eigenvectors. Some non-symmetric matrices have real-valued eigenvalues and eigenvectors. The eigenvectors of symmetric matrices with distinct eigenvalues are orthogonal and vice-versa. ($\boldsymbol{v}_i^{\mathsf{T}}\boldsymbol{v}_j = 0$ for $i \neq j \Leftrightarrow \boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}}$) So the eigenvectors of any non-symmetric matrix are not-orthonormal. Matrices with real eigenvalues are classified according to the signs of their eigenvalues. These classifications are illustrated in the following sections, in which,

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$$

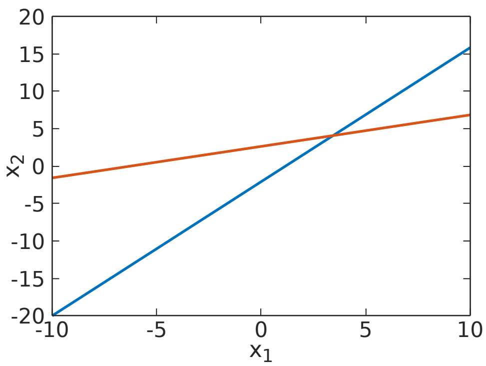
$$\boldsymbol{\Lambda} = \begin{bmatrix} \pm 0.5 \text{ or } 0 & \\ & \pm 1.5 \end{bmatrix}, \quad \boldsymbol{V}_{\rm s} = \left[\begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}, \begin{bmatrix} 0.6 \\ -0.8 \end{bmatrix} \right], \quad \boldsymbol{V}_{\rm ns} = \left[\begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right]$$

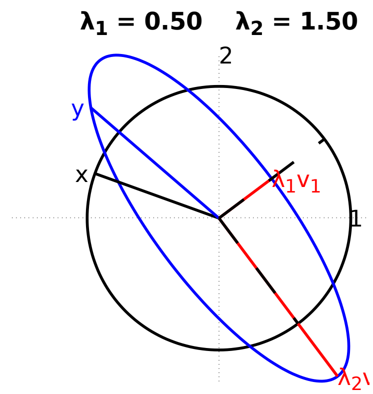
- the colums of $\boldsymbol{V}_{\rm s}$ are orthogonal ($\boldsymbol{V}_{\rm s}^{-1} = \boldsymbol{V}_{\rm s}^{\mathsf{T}}$), resulting in $\boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}}$
- the colums of $\boldsymbol{V}_{\rm ns}$ are not orthogonal, resulting in $\boldsymbol{A} \neq \boldsymbol{A}^{\mathsf{T}}$
- the black circle represents $\boldsymbol{x} \in \mathbb{R}^2$ s.t. $x_1^2 + x_2^2 = 1$, the unit circle
- the blue ellipse represents $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \ \forall \ \boldsymbol{x} : \ x_1^2 + x_2^2 = 1$
- the black solid line represents one of the elements of $\boldsymbol{x}$
- the blue solid line represents the corresponding $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$
- the black dashed lines represent the eigenvectors, $\boldsymbol{v}$, of $\boldsymbol{A}$
- the red lines represent the scaled eigenvectors, $\boldsymbol{v}\boldsymbol{\lambda}$, of $\boldsymbol{A}$
- the surfaces are graphs of the quadratics $z = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$

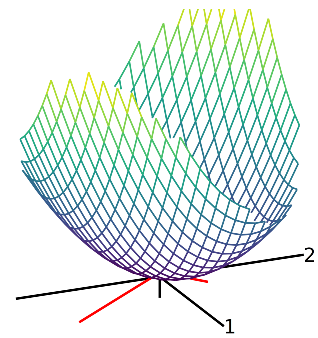
$\boldsymbol{A}$ is symmetric ($\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1} \quad \boldsymbol{V}^{-1} = \boldsymbol{V}^{\mathsf{T}}$)

positive eigenvalues (*positive definite*)

$$\begin{bmatrix} 0.86 & -0.48 \\ -0.48 & 1.14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

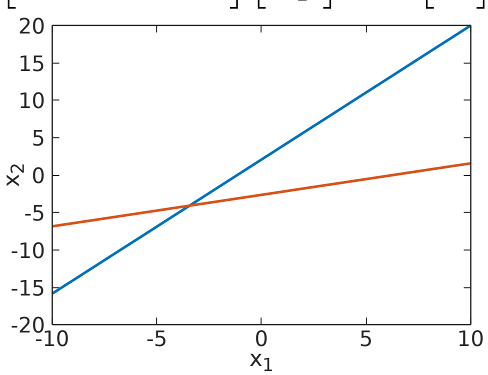


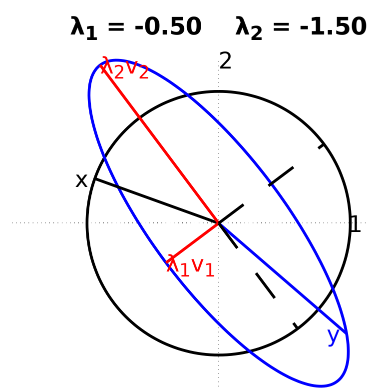


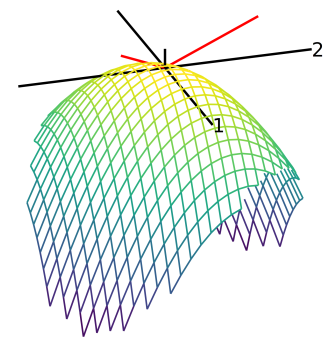


negative eigenvalues (*negative definite*)

$$\begin{bmatrix} -0.86 & 0.48 \\ 0.48 & -1.14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

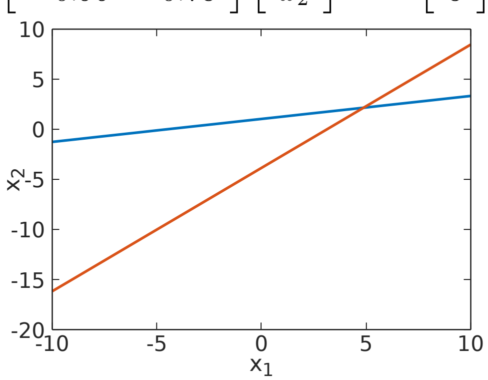


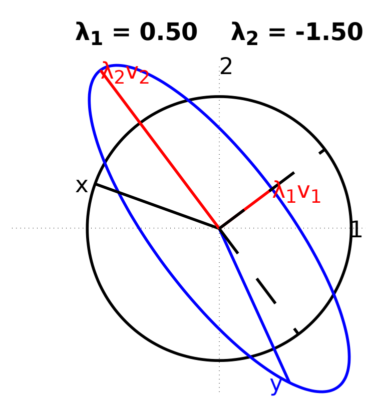


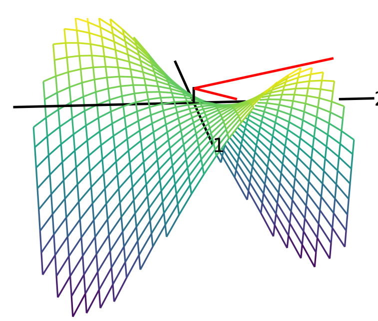


positive and negative eigenvalues (*indefinite*)

$$\begin{bmatrix} -0.22 & 0.96 \\ 0.96 & -0.78 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

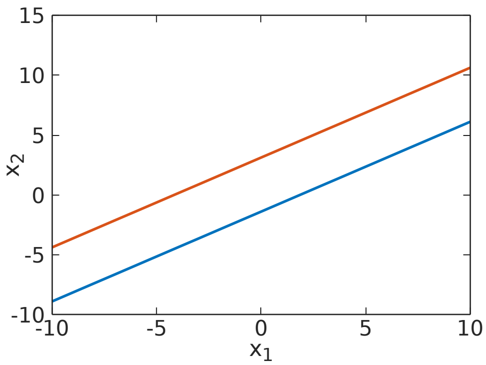


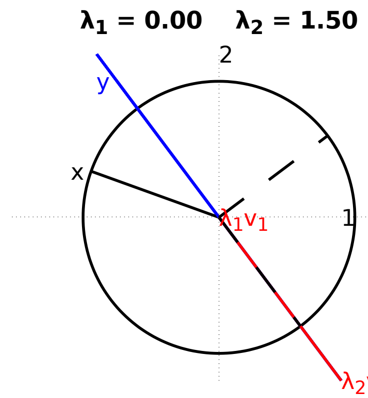


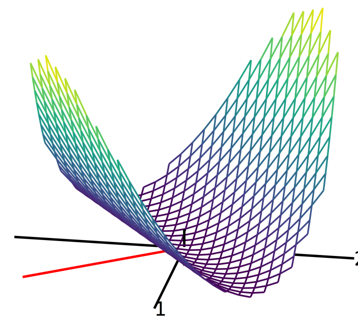


non-negative eigenvalues (*non-negative definite*)

$$\begin{bmatrix} 0.54 & -0.72 \\ -0.72 & 0.96 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



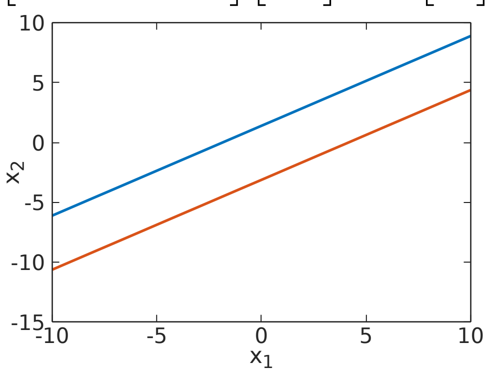


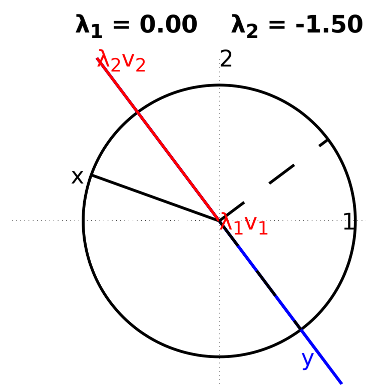


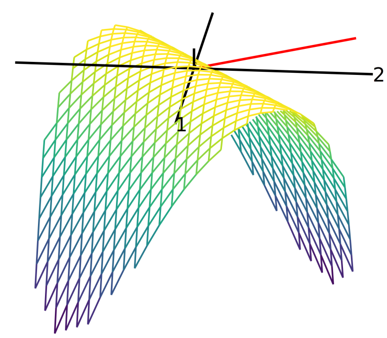
---

non-positive eigenvalues (*non-positive definite*)

$$\begin{bmatrix} -0.54 & 0.72 \\ 0.72 & -0.96 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



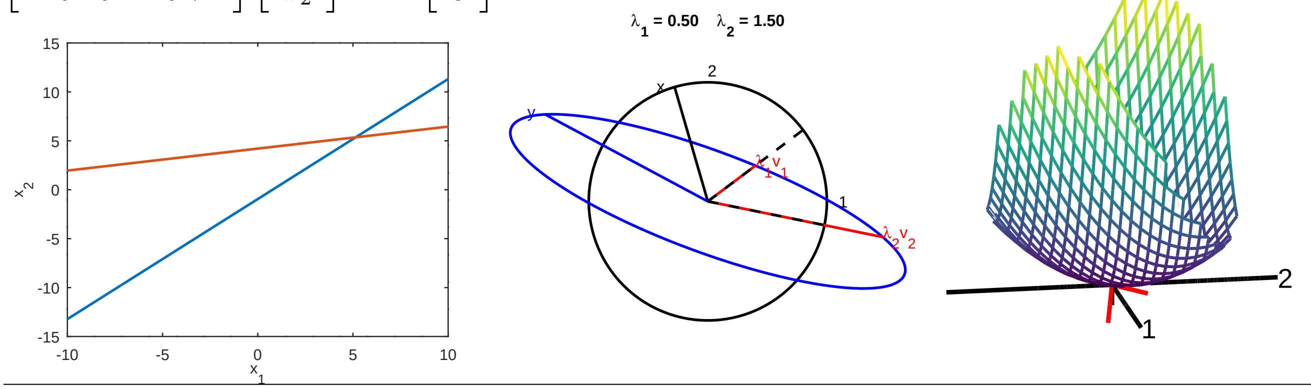




$\boldsymbol{A}$ is not symmetric ($\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1} \quad \boldsymbol{V}^{-1} \neq \boldsymbol{V}^{\mathsf{T}}$)

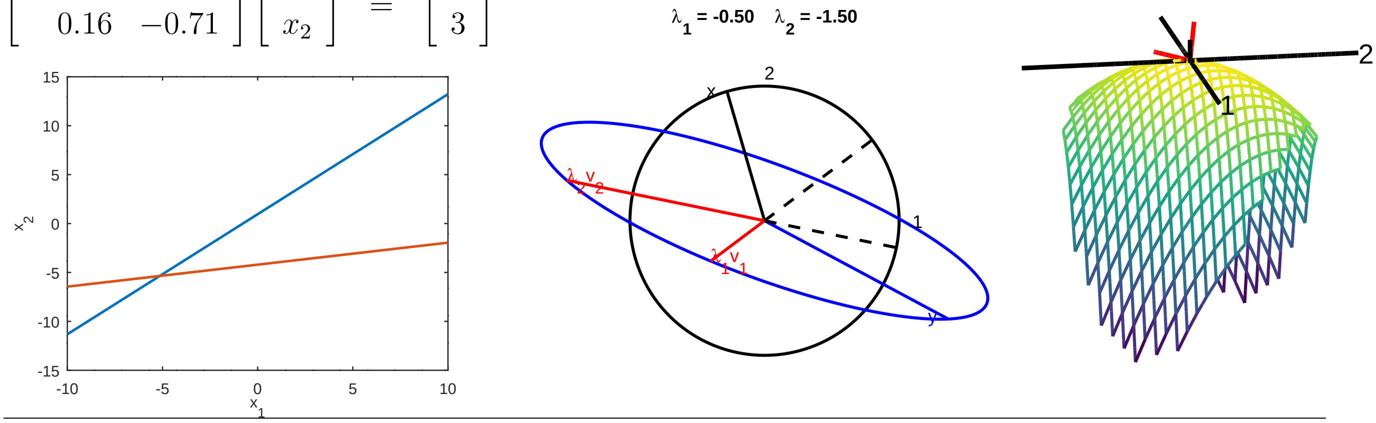
positive eigenvalues

$$\begin{bmatrix} 1.29 & -1.05 \\ -0.16 & 0.71 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



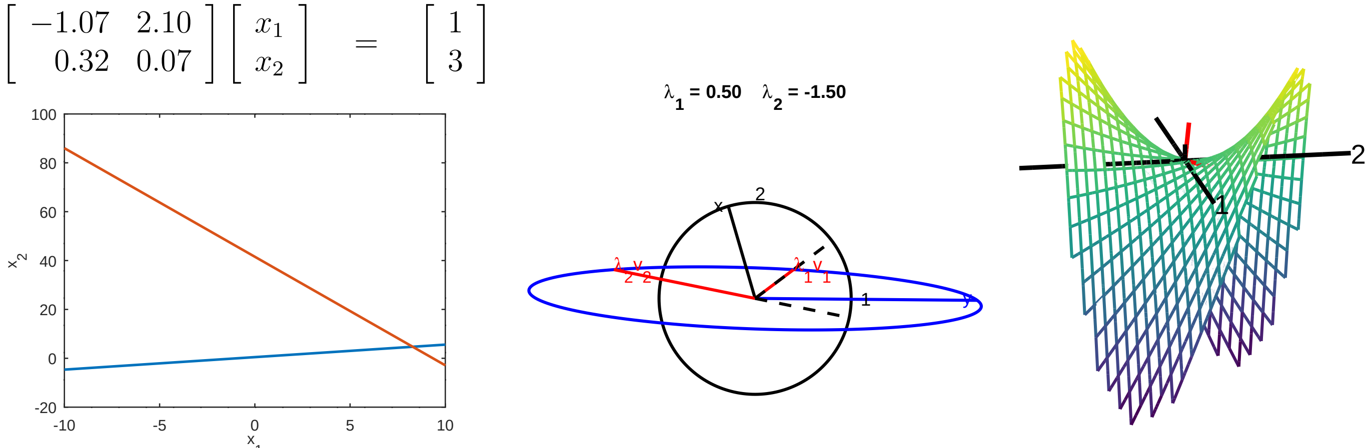
negative eigenvalues

$$\begin{bmatrix} -1.29 & 1.05 \\ 0.16 & -0.71 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



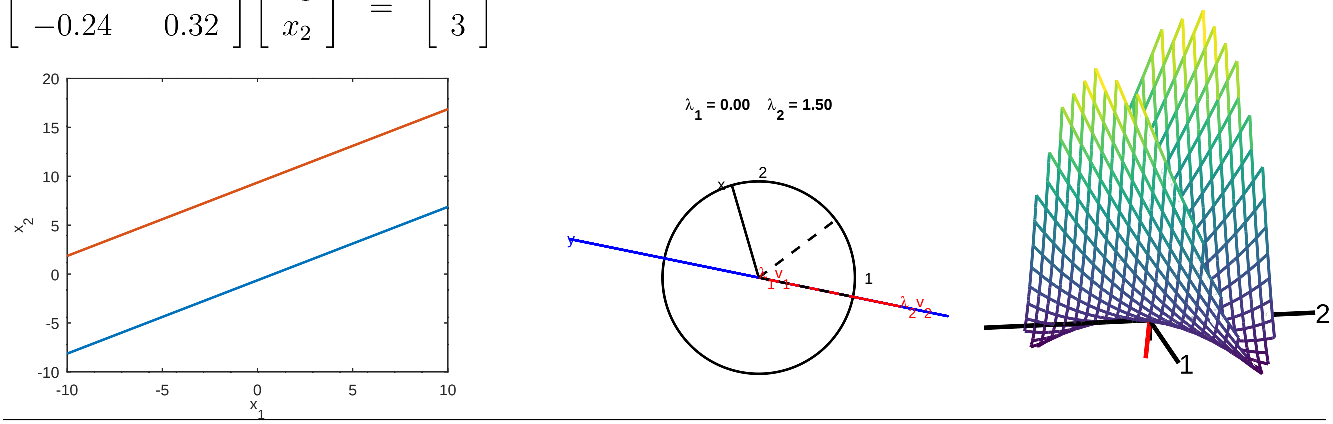
positive and negative eigenvalues

$$\begin{bmatrix} -1.07 & 2.10 \\ 0.32 & 0.07 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



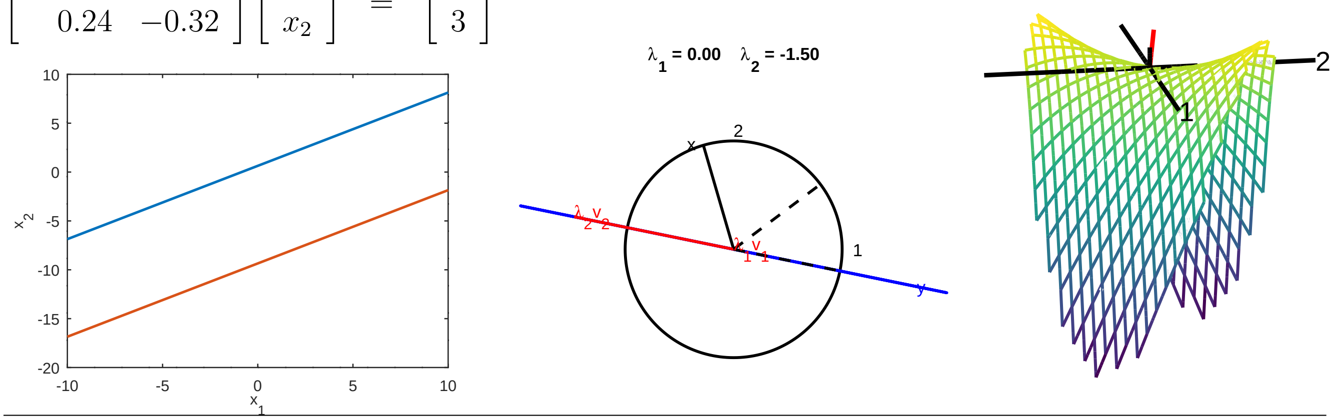
non-negative eigenvalues

$$\begin{bmatrix} 1.18 & -1.57 \\ -0.24 & 0.32 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



non-positive eigenvalues

$$\begin{bmatrix} -1.18 & 1.57 \\ 0.24 & -0.32 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



the transpose of a matrix product is the product of the transposes in reverse order

For two real-valued matrices $\boldsymbol{X} \in \mathbb{R}^{p\times q}$ and $\boldsymbol{Y} \in \mathbb{R}^{q\times r}$,

$$(\boldsymbol{XY})^{\mathsf{T}} = \boldsymbol{Y}^{\mathsf{T}}\boldsymbol{X}^{\mathsf{T}}$$

Consider the i, j and j, i elements of the matrix product $(\boldsymbol{XY})$

$$\begin{aligned}
(\boldsymbol{XY})_{ij} &= \sum_{k=1}^{q} X_{ik}Y_{kj} \\
(\boldsymbol{XY})_{ji} &= \sum_{k=1}^{q} X_{jk}Y_{ki} \\
&= \sum_{k=1}^{q} Y_{ki}X_{jk} \\
&= \sum_{k=1}^{q} Y^{\mathsf{T}}_{ik}X^{\mathsf{T}}_{kj} \\
&= (\boldsymbol{Y}^{\mathsf{T}}\boldsymbol{X}^{\mathsf{T}})_{ij}
\end{aligned}$$

And for two complex-valued matrices $\boldsymbol{X} \in \mathbb{C}^{p\times q}$ and $\boldsymbol{Y} \in \mathbb{C}^{q\times r}$,

$$(\boldsymbol{XY})^{*} = \boldsymbol{Y}^{*}\boldsymbol{X}^{*}$$

Consider the i, j element of $(\boldsymbol{XY})$ and j, i element of the complex conjugate of $(\boldsymbol{XY})$

$$\begin{aligned}
(\boldsymbol{XY})_{ij} &= \sum_{k=1}^{q} X_{ik}Y_{kj} \\
(\boldsymbol{XY})'_{ji} &= \sum_{k=1}^{q} X'_{jk}Y'_{ki} \\
&= \sum_{k=1}^{q} Y'_{ki}X'_{jk} \\
&= \sum_{k=1}^{q} Y^{*}_{ik}X^{*}_{kj} \\
&= (\boldsymbol{Y}^{*}\boldsymbol{X}^{*})_{ij}
\end{aligned}$$

where $\bullet'$ denotes the complex conjugate of a scalar and $\bullet^{*}$ denotes the complex conjugate transpose of a matrix.

the inner product of any matrix with itself is real and symmetric

For any real-valued matrix $\boldsymbol{X} \in \mathbb{R}^{p\times q}$, $(\boldsymbol{X}^\mathsf{T}\boldsymbol{X}) = (\boldsymbol{X}^\mathsf{T}\boldsymbol{X})^\mathsf{T}$
From the fact that $(\boldsymbol{X}^\mathsf{T})^\mathsf{T} = \boldsymbol{X}$ and the result $(\boldsymbol{X}\boldsymbol{Y})^\mathsf{T} = (\boldsymbol{Y}^\mathsf{T}\boldsymbol{X}^\mathsf{T})$, it is seen that $(\boldsymbol{X}^\mathsf{T}\boldsymbol{X})^\mathsf{T} = (\boldsymbol{X}^\mathsf{T}\boldsymbol{X})$.
So $(\boldsymbol{X}^\mathsf{T}\boldsymbol{X})$ is symmetric.

For any complex-valued matrix $\boldsymbol{X} \in \mathbb{C}^{p\times q}$, $(\boldsymbol{X}^*\boldsymbol{X}) = (\boldsymbol{X}^*\boldsymbol{X})^\mathsf{T}$
From the fact that $(\boldsymbol{X}^*)^* = \boldsymbol{X}$ and the result $(\boldsymbol{X}\boldsymbol{Y})^* = (\boldsymbol{Y}^*\boldsymbol{X}^*)$, it is seen that $(\boldsymbol{X}^*\boldsymbol{X})^* = (\boldsymbol{X}^*\boldsymbol{X})$.
Further, a complex value times its complex conjugate is real.
So $(\boldsymbol{X}^*\boldsymbol{X})$ is real and symmetric.

pairs eigenvectors of a real symmetric matrices corresponding to distinct eigenvalues are orthogonal

Consider two eigenvalues λ and μ of a real symmetric matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $(\boldsymbol{A} = \boldsymbol{A}^\mathsf{T})$, and their associated eigenvectors $\boldsymbol{v}$ and $\boldsymbol{u}$.

$$\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}\lambda \qquad\qquad \boldsymbol{A}\boldsymbol{u} = \boldsymbol{u}\mu$$

Pre-multiply the first by $\boldsymbol{u}^\mathsf{T}$ and the second by $\boldsymbol{v}^\mathsf{T}$

$$\boldsymbol{u}^\mathsf{T}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{u}^\mathsf{T}\boldsymbol{v}\lambda \qquad\qquad \boldsymbol{v}^\mathsf{T}\boldsymbol{A}\boldsymbol{u} = \boldsymbol{v}^\mathsf{T}\boldsymbol{u}\mu$$

Since $\boldsymbol{A}$ is symmetric, $\boldsymbol{u}^\mathsf{T}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{u} = \boldsymbol{v}^\mathsf{T}\boldsymbol{A}\boldsymbol{u}$ so subtracting the equations above gives

$$\boldsymbol{u}^\mathsf{T}\boldsymbol{A}\boldsymbol{v} - \boldsymbol{v}^\mathsf{T}\boldsymbol{A}\boldsymbol{u} = \boldsymbol{u}^\mathsf{T}\boldsymbol{v}\lambda - \boldsymbol{v}^\mathsf{T}\boldsymbol{u}\mu$$

$$0 = \boldsymbol{u}^\mathsf{T}\boldsymbol{v}(\lambda - \mu)$$

So, if the eigenvalues λ and μ are not equal to each other, then $\boldsymbol{u}^\mathsf{T}\boldsymbol{v}$ must be zero, or, in other words, the eigenvectors $\boldsymbol{u}$ and $\boldsymbol{v}$ must be orthogonal.

the eigenvalues of any real symmetric matrix are real

We can use the above method to show that the eigenvalues of any real symmetric matrix are real. Consider an eigenvalue problem

$$\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}\lambda$$

and its complex conjugate (not transpose)

$$(\boldsymbol{A}\boldsymbol{v})' = (\boldsymbol{v}\lambda)'$$

Distributing the complex-conjugates ...

$$\boldsymbol{A}'\boldsymbol{v}' = \boldsymbol{v}'\lambda'$$

Because $\boldsymbol{A}$ is presumed real $\boldsymbol{A}' = \boldsymbol{A}$, and

$$\boldsymbol{A}\boldsymbol{v}' = \boldsymbol{v}'\lambda'$$

Pre-multiplying the original eigenvalue probelm by v'^{T} and the last by v^{T} we get the two quadratic forms

$$\boldsymbol{v}'^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}'^{\mathsf{T}}\boldsymbol{v}\lambda$$

$$\boldsymbol{v}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v}' = \boldsymbol{v}^{\mathsf{T}}\boldsymbol{v}'\lambda'$$

Because $\boldsymbol{A}$ is symmetric $\boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}}$,

$$(\boldsymbol{v}'^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v})^{\mathsf{T}} = (\boldsymbol{v}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{v}') = (\boldsymbol{v}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v}')$$

Now, subtracting the two quadratic forms above,

$$\boldsymbol{v}'^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v} - \boldsymbol{v}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{v}' = \boldsymbol{v}'^{\mathsf{T}}\boldsymbol{v}\lambda - \boldsymbol{v}^{\mathsf{T}}\boldsymbol{v}'\lambda'$$

resulting in

$$0 = \boldsymbol{v}'^{\mathsf{T}}\boldsymbol{v}(\lambda - \lambda')$$

Since $\boldsymbol{v} \neq 0$, $(\lambda - \lambda')$ must be zero, or, in other words, the eigenvalue must equal its complex conjugate, or, in other words, the eigenvalue must be real.

matrix-vector multiplication via eigenvalue decomposition

As seen above, for real symmetric matrices $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, the eigenvalues are real and the eigenvectors are orthogonal. For normalized orthogonal eigenvectors arranged column-wise in a matrix

$$\boldsymbol{V} = \begin{bmatrix} | & | & & | \\ \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \\ | & | & & | \end{bmatrix}$$

the orthonormality implies $\boldsymbol{V}^{\mathsf{T}}\boldsymbol{V} = \boldsymbol{I}$, or, equivalently $\boldsymbol{V}^{\mathsf{T}} = \boldsymbol{V}^{-1}$. For matrices with orthogonal eigenvectors, i.e., real symmetric matrices, the eigenvalue problem may be written for all eigenvectors and eigenvalues as follows,

$$\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\Lambda$$

or, in general

$$\boldsymbol{A} = \boldsymbol{V}\Lambda\boldsymbol{V}^{-1}$$

or, specifically for real symmetric matrices $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{V}\Lambda\boldsymbol{V}^{\mathsf{T}}$$

$$
\begin{aligned}
\boldsymbol{y} &= \boldsymbol{A}\boldsymbol{x} \\
\boldsymbol{y} &= \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathsf{T}}\boldsymbol{x} \\
&= \begin{bmatrix} | & | & & | \\ \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \\ | & | & & | \end{bmatrix}
\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}
\begin{bmatrix} - & \boldsymbol{v}_1^{\mathsf{T}} & - \\ - & \boldsymbol{v}_2^{\mathsf{T}} & - \\ & \vdots & \\ - & \boldsymbol{v}_n^{\mathsf{T}} & - \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} | & | & & | \\ \lambda_1\boldsymbol{v}_1 & \lambda_2\boldsymbol{v}_2 & \cdots & \lambda_n\boldsymbol{v}_n \\ | & | & & | \end{bmatrix}
\begin{bmatrix} - & \boldsymbol{v}_1^{\mathsf{T}} & - \\ - & \boldsymbol{v}_2^{\mathsf{T}} & - \\ & \vdots & \\ - & \boldsymbol{v}_n^{\mathsf{T}} & - \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\end{aligned}
$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{v}_1\boldsymbol{v}_1^{\mathsf{T}} & \lambda_2\boldsymbol{v}_2\boldsymbol{v}_2^{\mathsf{T}} & \cdots & \lambda_n\boldsymbol{v}_n\boldsymbol{v}_n^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
$$

$$
\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathsf{T}}\right]_{n\times n}
$$

So a real symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ may be thought of as the wighted sum of rank-one matrices $[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathsf{T}}] \in \mathbb{R}^{n\times n}$ in which the weights are the eigenvalues λ_i.

positive definite

A *positive definite* is a symmetric matrix with postiive eigenvalues. Positive definite matrices are denoted $\boldsymbol{A} > 0$.
$\boldsymbol{A} > 0 \Leftrightarrow \forall \boldsymbol{x} \neq 0, \boldsymbol{x}^{\mathsf{T}} A \boldsymbol{x} > 0$.

negative definite

A *negative definite* is a symmetric matrix with negative eigenvalues. Negative definite matrices are denoted $\boldsymbol{A} < 0$.
$\boldsymbol{A} < 0 \Leftrightarrow \forall \boldsymbol{x} \neq 0, \boldsymbol{x}^{\mathsf{T}} A \boldsymbol{x} < 0$.

correlation of uncorrelated random variables

A *standard random variable*, Z, has an expected value of 0 and a variance of 1,

$$\mathsf{E}[Z] = 0 \quad , \quad \mathsf{E}[Z^2] = 1$$

A pair of *uncorrelated standard random variables*, Z_1 and Z_2, has a covariance of 0,

$$\mathsf{E}[Z_1] = 0 \quad , \quad \mathsf{E}[Z_2] = 0 \quad , \quad \mathsf{E}[Z_1^2] = 1 \quad , \quad \mathsf{E}[Z_2^2] = 1 \quad \mathsf{E}[Z_1 Z_2] = 0$$

The rows of a matrix of uncorrelated standard random variables $\boldsymbol{z} \in \mathbb{R}^{n \times N}$ can be interpreted as n uncorrelated time series of N points. The covariance of $\boldsymbol{z}$ is

$$\boldsymbol{C}_z = \mathsf{E}[\boldsymbol{z}\boldsymbol{z}^\mathsf{T}] = \lim_{N\to\infty} \frac{1}{N} \boldsymbol{z}\boldsymbol{z}^\mathsf{T} = \boldsymbol{I}_n$$

The rows of a matrix $\boldsymbol{a} = \boldsymbol{R}\boldsymbol{z}$, with $\boldsymbol{R} \in \mathbb{R}^{n \times n}$ can be interpreted as n correlated time series of N points. The covariance of $\boldsymbol{a}$ is

$$\boldsymbol{C}_a = \mathsf{E}[\boldsymbol{a}\boldsymbol{a}^\mathsf{T}] = \mathsf{E}[\boldsymbol{R}\boldsymbol{z}\boldsymbol{z}^\mathsf{T}\boldsymbol{R}^\mathsf{T}] = \boldsymbol{R}\,\mathsf{E}[\boldsymbol{z}\boldsymbol{z}^\mathsf{T}]\boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^\mathsf{T}$$

$$\boldsymbol{C}_a = \lim_{N\to\infty} \frac{1}{N} \boldsymbol{a}\boldsymbol{a}^\mathsf{T} = \lim_{N\to\infty} \frac{1}{N} \boldsymbol{R}\boldsymbol{z}\boldsymbol{z}^\mathsf{T}\boldsymbol{R}^\mathsf{T} = \lim_{N\to\infty} \frac{1}{N} N\boldsymbol{R}\boldsymbol{R}^\mathsf{T} = \boldsymbol{R}\boldsymbol{R}^\mathsf{T}$$

The eigenvalue decompostion $\boldsymbol{C}_a = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^\mathsf{T}$ may be used to determine the correlating matrix $\boldsymbol{R}$.

$$\boldsymbol{R} = \boldsymbol{V}\boldsymbol{\Lambda}^{1/2}$$

(Other matrix factorizations, e.g., Cholesky factorization or LDL$^\mathsf{T}$ factorization, may also be used to compute $\boldsymbol{R}$ from $\boldsymbol{C}_a$.)