Symmetric matrices and SVD

San Jose State University

Prof. Guangliang Chen

Fall 2022

Outline

Symmetric matrices

Def 0.1. A *square* matrix $\mathbf{A}$ is said to be **symmetric** if $\mathbf{A}^T = \mathbf{A}$, i.e., $a_{ij} = a_{ji}$ for all i, j.

Example 0.1. The following matrices are all symmetric:

$$\begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Remark. Symmetric matrices have many nice properties.

For example, they have no complex eigenvalues, and eigenvectors corresponding to distinct eigenvalues must be orthogonal to each other.

Additionally, they are always diagonalizable via orthogonal matrices.

We present these results in a theorem, divided into two parts.

The Spectral Theorem (part 1)

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be any symmetric matrix. Then

- $\mathbf{A}$ has n real eigenvalues, counting multiplicities (there are no complex eigenvalues)
- For each distinct eigenvalue λ_i, the geometric multiplicity must coincide with the algebraic multiplicity, i.e., $a_i = g_i$. This implies that $\mathbf{A}$ must be diagonalizable.

Example 0.2. Diagonalize the following symmetric matrix:

$$\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$$

Answer: From

$$0 = \det(\mathbf{A} - \lambda \mathbf{I}) = (7-\lambda)(4-\lambda) - 2^2 = (\lambda - 8)(\lambda - 3)$$

we obtain two distinct eigenvalues $\lambda_1 = 8, \lambda_2 = 3$.

The corresponding eigenvectors can be found by directly solving $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{v}_i = \mathbf{0}$:

$$A - \lambda_1 \mathbf{I} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \quad \longrightarrow \quad \mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$A - \lambda_2 \mathbf{I} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \quad \longrightarrow \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

It follows that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, where

$$\mathbf{P} = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 8 & \\ & 3 \end{bmatrix}$$

Example 0.3. Diagonalize the following symmetric matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Answer: We start by computing the characteristic equation of $\mathbf{A}$:

$$0 = \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = (3-\lambda)\lambda^2$$

This shows that the matrix has two distinct eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 0$ with corresponding algebraic multiplicities $a_1 = 1$ and $a_2 = 2$.

For the eigenvalue $\lambda_1 = 3$ (we must have $g_1 = a_1 = 1$), there is only one linearly independent eigenvector, found as follows:

$$[\mathbf{A} - 3\mathbf{I} \mid \mathbf{0}] = \begin{bmatrix} -2 & 1 & 1 & | & 0 \\ 1 & -2 & 1 & | & 0 \\ 1 & 1 & -2 & | & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \longrightarrow \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For the eigenvalue $\lambda_2 = 0$ (we must have $g_2 = a_2 = 2$ due to the symmetry of $\mathbf{A}$), there are two linearly independent eigenvectors, found as follows:

$$[\mathbf{A} - 0\mathbf{I} \mid \mathbf{0}] = \begin{bmatrix} 1 & 1 & 1 & | & 0 \\ 1 & 1 & 1 & | & 0 \\ 1 & 1 & 1 & | & 0 \end{bmatrix} \longrightarrow \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Thus, if we let

$$\mathbf{P} = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_2) = \begin{bmatrix} 3 & & \\ & 0 & \\ & & 0 \end{bmatrix}$$

then we have

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

i.e.,

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} 3 & & \\ & 0 & \\ & & 0 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1}}_{\mathbf{P}^{-1}}.$$

Eigenspaces of a symmetric matrix corresponding to different eigenvalues must be orthogonal to each other

Theorem 0.1. If $\mathbf{A}$ is symmetric, then any two eigenvectors corresponding to different eigenvalues are orthogonal.

Proof. Let $\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$ such that $\lambda_1 \neq \lambda_2$. We need to show that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ (i.e., they are orthogonal to each other) based on the assumption that $\mathbf{A}$ is symmetric (i.e., $\mathbf{A}^T = \mathbf{A}$).

Write

$$\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = (\mathbf{A}\mathbf{v}_1) \cdot \mathbf{v}_2 = (\mathbf{A}\mathbf{v}_1)^T \mathbf{v}_2 = \mathbf{v}_1^T \mathbf{A}\mathbf{v}_2 = \mathbf{v}_1^T(\lambda_2 \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$$

Since $\lambda_2 \neq \lambda_1$, we must have $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. □

The Spectral Theorem (part 2)

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be any symmetric matrix. Then $\mathbf{A}$ is orthogonally diagonalizable, i.e., there exists an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ (i.e., $\mathbf{Q}^{-1} = \mathbf{Q}^T$) such that

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T.$$

Remark. This is also a spectral decomposition of $\mathbf{A}$:

- $\mathbf{Q}$ consists of (orthonormal) eigenvectors of $\mathbf{A}$: $\mathbf{Q} = [\mathbf{v}_1 \ldots \mathbf{v}_n]$
- $\mathbf{D}$ contains the corresponding eigenvalues of $\mathbf{A}$ along its diagonal: $\mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$

Example 0.4. Orthogonally diagonalize the following symmetric matrix:

$$\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$$

Example 0.5. Orthogonally diagonalize the following symmetric matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Answer: We have previously obtained that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, where

$$\mathbf{P} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_2) = \begin{bmatrix} 3 & & \\ & 0 & \\ & & 0 \end{bmatrix}$$

Since $\{\mathbf{v}_1\}$ and $\{\mathbf{v}_2, \mathbf{v}_3\}$ correspond to different eigenvalues ($\lambda_1 = 3,\ \lambda_2 = 0$) of the symmetric matrix $\mathbf{A}$, we must have that $\mathbf{v}_1$ is orthogonal to both of $\mathbf{v}_2, \mathbf{v}_3$

(but $\mathbf{v}_2, \mathbf{v}_3$ are not necessarily orthogonal because they come from the same eigenvalue $\lambda_2 = 0$).

To obtain an orthogonal matrix for diagonalizing $\mathbf{A}$, we just need to

(1) normalize $\mathbf{v}_1$ and

(2) apply the Gram-Schmidt process to convert $\{\mathbf{v}_2, \mathbf{v}_3\}$ to an orthonormal basis for the eigenspace associated to $\lambda_2 = 0$.

Specifically,

(1) $\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(2) $\mathbf{u}_2 = \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and

$$\mathbf{u}_3 = \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_2}\mathbf{v}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$$

$$= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{6/4}} \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

Therefore, the orthogonal matrix $\mathbf{Q}$ that is needed for diagonalizing $\mathbf{A}$ (i.e., $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ for the same diagonal matrix $\mathbf{D}$) is

$$\mathbf{Q} = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}$$

Quadratic forms

Symmetric matrices can be used to define the so-called quadratic forms.

Def 0.2. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. A **quadratic form** based on $\mathbf{A}$ is a function $Q : \mathbb{R}^n \mapsto \mathbb{R}$ with

$$Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}, \qquad \text{for all} \quad \mathbf{x} \in \mathbb{R}^n.$$

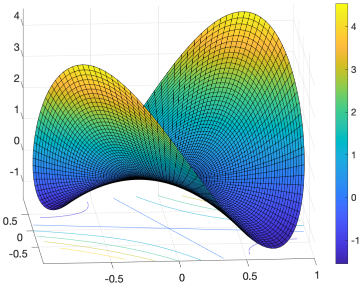
Remark. A quadratic form is a second-order polynomial in the components of $\mathbf{x}$ without linear or constant terms:

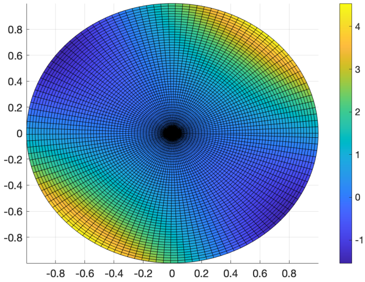
$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j = \sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$$

For all symmetric matrices $\mathbf{A}$, we have $Q(\mathbf{0}) = \mathbf{0}^T\mathbf{A}\mathbf{0} = 0$

For example, if $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = x_1^2 + 2x_2^2 + 6x_1x_2$$





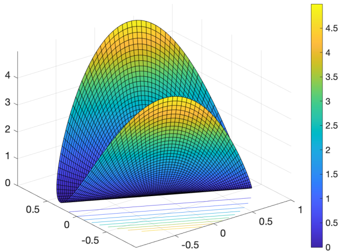
Question: Which symmetric matrix corresponds to

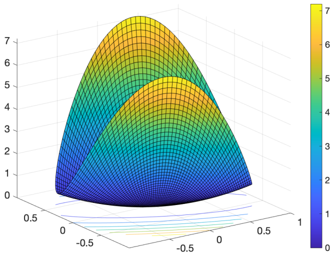
$$Q(\mathbf{x}) = x_1^2 + 2x_2^2 + 3x_3^2 + 6x_1x_2 - 4x_1x_3 + 10x_2x_3$$

Positive (semi)definite matrices

Consider the following symmetric matrices, and their associated quadratic forms

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$





For some symmetric matrices (like those two on the preceding slide), their quadratic forms are never negative.

A *symmetric* matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is said to be **positive definite** if $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n$.

A *symmetric* matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is said to be **positive semidefinite** if $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

Positive definite matrices must be positive semidefinite; the converse is not true.

It turns out that one can check the eigenvalues of a symmetric matrix to determine its positive definiteness.

Theorem 0.2. A symmetric matrix is positive definite (positive semidefinite) if and only if all of its eigenvalues are strictly positive (nonnegative).

Example 0.6. Determine the positive definiteness of each of the following symmetric matrices by finding their eigenvalues:

$$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

Theorem 0.3. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be any matrix. Then $\mathbf{A}^T\mathbf{A}$ is square, symmetric and positive definite.

Proof.

- square: $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n \times n}$.
- symmetric: $(\mathbf{A}^T\mathbf{A})^T = (\mathbf{A})^T(\mathbf{A}^T)^T = \mathbf{A}^T\mathbf{A}$.
- positive semidefinite: For any $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^T(\mathbf{A}^T\mathbf{A})\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) = \|\mathbf{A}\mathbf{x}\|^2 \geq 0.$$

This result will be used later when deriving the singular value decomposition.

Further learning

Math 250: Mathematical Data Visualization[1]

Refer to the following lectures:

- **Basic matrix algebra**: Diagonalization of idempotent matrices, matrix square roots, the generalized eigenvalue problem
- **Matrix computing in Matlab**
- **Singular value decomposition (SVD)**

Math 251: Statistical and Machine Learning Classification[2]

[1] https://www.sjsu.edu/faculty/guangliang.chen/Math250.html
[2] https://www.sjsu.edu/faculty/guangliang.chen/Math251.html