Eigenvalues & Eigenvectors

Mathematical Methods for Economics (771)

Stellenbosch University

Contents:

Readings

- Simon & Blume, Chapter 23.1-23.3, 23.6-23.7 (recommended self-study 23.8-23.9)
- Additional: 3Blue1Brown

Linear combinations and span

The set of all scalar multiplies of a nonzero vector $\mathbf{v}$ is a straight line through the origin.

Formally, we denote this set by:

$$\mathcal{L}[\mathbf{v}] \equiv \{r.\mathbf{v} : r \in \mathbf{R}\} \ , \tag{1}$$

and call it the line *generated* or *spanned* by $\mathbf{v}$.

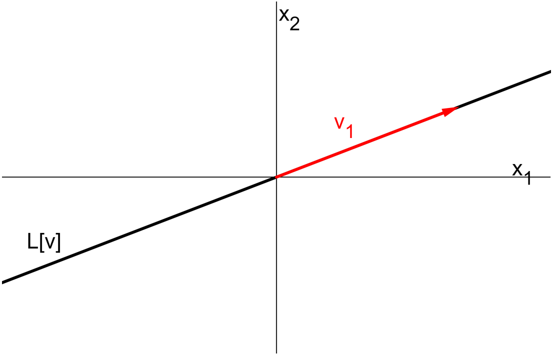

Figure: The line $\mathcal{L}$ spanned by vector $\mathbf{v}$

Linear combinations and span

The set spanned by two nonzero vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ is given by:

$$\mathcal{L}[\mathbf{v_1}, \mathbf{v_2}] \equiv \{r_1.\mathbf{v_1} + r_2.\mathbf{v_2} : r_1, r_2 \in \mathbf{R}\} \ , \tag{2}$$

If $\mathbf{v_1}$ is a multiple of $\mathbf{v_2}$, or vice versa, we say $\mathbf{v_1}$ and $\mathbf{v_2}$ are **linearly dependent**. Otherwise, we say that $\mathbf{v_1}$ and $\mathbf{v_2}$ are **linearly *in*dependent**.

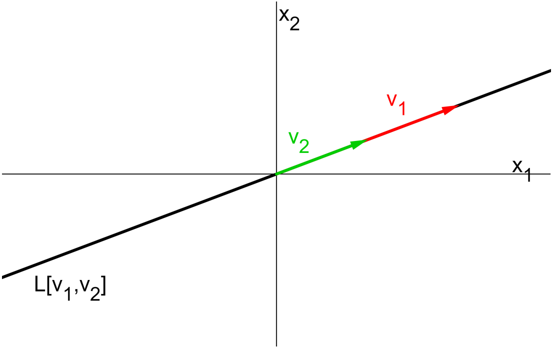

Figure: If $\mathbf{v_1}$ is a multiple of $\mathbf{v_2}$, then $\mathcal{L}[\mathbf{v_1}, \mathbf{v_2}] = \mathcal{L}[\mathbf{v_2}]$ is simply a line spanned by $\mathbf{v_2}$.

Linear combinations and span

That is, given the linear combination of two nonzero vectors $\mathbf{v_1}$ and $\mathbf{v_2}$, $c_1\mathbf{v_1} + c_2\mathbf{v_2} = 0$, we say $\mathbf{v_1}$ and $\mathbf{v_2}$ are:

- **linearly dependent** if c_1 or c_2 nonzero ($\mathcal{L}[\cdot]$ is a line);
- **linearly *in*dependent** if $c_1 = c_2 = 0$ ($\mathcal{L}[\cdot]$ is a plane).

Definition

Vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbf{R^n}$ are **linearly dependent** if and only if there exists scalars $c_1, c_2, \ldots, c_k$, *not all zero*, such that

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_k\mathbf{v_k} = 0$$

Vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbf{R^n}$ are **linearly independent** if and only if $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_k\mathbf{v_k} = 0$ for scalars $c_1, \ldots, c_k$ implies that $c_1 = \cdots = c_k = 0$.

Theorem (11.1)

*Vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbf{R^n}$ are **linearly dependent** if and only if the linear system*

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = 0$$

has a nonzero solution $(c_1, c_2, \ldots, c_k)$, where A is the $n \times k$ matrix whose columns are the vectors $\mathbf{v_1}, \ldots, \mathbf{v_k}$ under study:

$$A = \begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_k} \end{pmatrix} .$$

Theorem (11.2, case $k = n$)

*A set of n vectors $\mathbf{v_1}, \ldots, \mathbf{v_n}$ in $\mathbf{R^n}$ are **linearly independent** if and only if*

$$\det A = \det \begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_n} \end{pmatrix} \neq 0 .$$

Theorem (11.3, case $k > n$)

*If $k > n$, any set A set of k vectors in $\mathbf{R^n}$ is **linearly dependent**.*

Spanning sets

Recall that the set spanned by $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ can be written as

$$\mathcal{L}[\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}] \equiv \{c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_k\mathbf{v_k} : c_1, c_2, \ldots, c_k \in \mathbf{R}\} \; .$$

Suppose we are given a subset V of $\mathbf{R^n}$. Then there exists $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbf{R^n}$ such that every vector in V can be written as a linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$:

$$V = \mathcal{L}[\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}] \; .$$

That is, $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ **span** V.

Spanning sets

Example (11.4)

The x_1x_2-plane in $\mathbf{R^3}$ is the span of the unit vectors $e_1 = (1, 0, 0)$ and $e_2 = (0, 1, 0)$, because any vector $(a, b, 0)$ in this plane can be written as a linear combination of e_1 and e_2:

$$\begin{pmatrix} a \\ b \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Theorem (11.6)

A set of k vectors that spans $\mathbf{R^n}$ must contain at least n vectors ($k \geq n$).

Basis and dimension in $\mathbf{R^n}$

It is clear from Theorem 11.6 that we can find a “more efficient” spanning set:

Definition

Let $\mathbf{v_1}, \dots, \mathbf{v_k}$ be a fixed set of k vectors in $\mathbf{R^n}$. Let V be the set $\mathcal{L}[\mathbf{v_1}, \dots, \mathbf{v_k}]$ spanned by $\mathbf{v_1}, \dots, \mathbf{v_k}$. Then, $\mathbf{v_1}, \dots, \mathbf{v_k}$ forms a **basis** of V if:

(a) $\mathbf{v_1}, \dots, \mathbf{v_k}$ span V, and

(b) $\mathbf{v_1}, \dots, \mathbf{v_k}$ are linearly independent.

Example (11.8)

The unit vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

form a basis of $\mathbf{R^n}$. Since this is such a natural basis, it is called the **canonical basis**.

Basis and dimension in $\mathbf{R^n}$

Theorem (11.7)

Every basis of $\mathbf{R^n}$ *contains* n *vectors.*

Theorem (11.8)

Let $\mathbf{v_1}, \ldots, \mathbf{v_n}$ *be a collection of* n *vectors in* $\mathbf{R^n}$*. Form the* $n \times n$ *matrix* A *whose columns are these* $\mathbf{v}_j$*'s:* $A = \begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_n} \end{pmatrix}$*.*
Then, the following statements are equivalent:

(a) $\mathbf{v_1}, \ldots, \mathbf{v_n}$ *are linearly independent,*

(b) $\mathbf{v_1}, \ldots, \mathbf{v_n}$ *span* $\mathbf{R^n}$*,*

(c) $\mathbf{v_1}, \ldots, \mathbf{v_n}$ *form a basis of* $\mathbf{R^n}$*, and*

(d) *the determinant of* $A_{n \times n}$ *is nonzero*

The **dimension** of a vector space V is the # of vectors in any basis of V.

Since every basis of $\mathbf{R^n}$ contains exactly n vectors, there are n independent directions in $\mathbf{R^n}$, and $\mathbf{R^n}$ is n-dimensional.

Chapter 23. Eigenvalues and Eigenvectors

Eigenvalues (“characteristic values”) and eigenvectors of a square matrix summarize the essential properties of linear and nonlinear systems of equations:

- They are the components of explicit solutions to *linear* dynamic models.
- The signs of eigenvalues determine the stability of equilibria in *nonlinear* dynamic models, and the definiteness of a symmetric matrix.
- Important for economic optimization problems.

Eigenvalues and Eigenvectors

Example (See 23.1, 23.2, 23.3, & 23.4., pp. 580-1)

Definition

Let A be a square matrix. An **eigenvalue** of A is a number r which when subtracted from each of the diagonal entries of A converts A into a singular matrix. Therefore, r is an eigenvalue of A if and only if $A - rI$ is a singular matrix.

Theorem (23.1)

The diagonal entries of a diagonal matrix D are eigenvalues of D.

Theorem (23.2)

A square matrix A is singular if and only if 0 is an eigenvalue of A.

A matrix is singular if and only if its determinant is zero. That is, $A - rI$ is a singular matrix, if and only if $\det(A - rI) = 0$.

Finding eigenvalues and eigenvectors

For an $n \times n$ matrix A, $\det(A - rI)$ is an nth order polynomial in the variable r, called the **characteristic polynomial**.

Example

For a general 2×2 matrix, the characteristic polynomial is

$$\begin{aligned} \det(A - rI) &= \det \begin{pmatrix} a_{11} - r & a_{12} \\ a_{21} & a_{22} - r \end{pmatrix} \\ &= r^2 - (a_{11} + a_{22})r + (a_{11}a_{22} - a_{12}a_{21}) \ , \end{aligned} \tag{3}$$

a second-order polynomial.

Therefore,

- The eigenvalues r of A are the roots of the characteristic polynomial;
- a 2×2 matrix has at most two eigenvalues (we can use the quadratic formula); a $n \times n$ matrix has at most n eigenvalues.

Finding eigenvalues and eigenvectors

When r is an eigenvalue of A, a *nonzero* vector $\mathbf{v}$ such that

$$(A - rI)\mathbf{v} = \mathbf{0} \tag{4}$$

is called an **eigenvector** of A *corresponding* to eigenvalue r. Multiplying out (4) yields

$$\begin{aligned} A\mathbf{v} - rI\mathbf{v} &= \mathbf{0} \\ A\mathbf{v} - r\mathbf{v} &= \mathbf{0} \\ A\mathbf{v} &= r\mathbf{v} \end{aligned}$$

If r is an eigenvalue and $\mathbf{v}$ is a corresponding eigenvector, then $A\mathbf{v} = r\mathbf{v}$.

Finding eigenvalues and eigenvectors

We can summarize the above as follows:

Theorem (23.3)

Let A be an $n \times n$ matrix and let r be a scalar. Then, the following statements are equivalent:

(a) *Subtracting r from each diagonal entry of A transforms A into a singular matrix.*

(b) *$A - rI$ is a singular matrix.*

(c) $\det(A - rI) = 0$

(d) $(A - rI)\mathbf{v} = \mathbf{0}$ *for some nonzero vector $\mathbf{v}$.*

(e) $A\mathbf{v} = r\mathbf{v}$ *for some nonzero vector $\mathbf{v}$.*

Finding eigenvalues and eigenvectors

Example

Examples 23.5 and 23.6 (pp. 583-4)

In general, one chooses the "simplest" of the nonzero candidates.

The **eigenspace** of A is the span of the set of all eigenvectors, including $\mathbf{v} = 0$. i.e., the set of all solutions to (4).

In some problems, one will need to use Guassian elimination to solve the linear system $(A - rI)\mathbf{v} = \mathbf{0}$ for an eigenvector $\mathbf{v}$.

Solving linear difference equations

One-dimensional equations

To solve a one-dimensional equation $y_{n+1} = ay_n$,

$$
\begin{aligned}
y_1 &= ay_0 \\
y_2 &= ay_1 = a(ay_0) = a^2 y_0 \\
y_3 &= ay_2 = a(a^2 y_0) = a^3 y_0 \\
\vdots &= \vdots \\
y_n &= a^n y_0
\end{aligned}
\tag{5}
$$

In what settings could the solution (5) arise?

Solving linear difference equations

Two-dimensional systems

Now, consider a system of two linear difference equations:

$$x_{n+1} = ax_n + by_n \tag{6}$$
$$y_{n+1} = cx_n + dy_n \tag{7}$$

In matrix form, the system of difference equations becomes:

$$\mathbf{z}_{n+1} = \begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} \equiv A\mathbf{z}_n \tag{8}$$

- If $b = c = 0$, (6) and (7) are uncoupled, and can be solved as two seperate one-dimensional problems (5).
- If $b \neq 0$ or $c \neq 0$, we need to transform the coefficient matrix A into a diagonal matrix so that (6) and (7) become uncoupled and therefore more easily solved; and
- To transform matrix A into a diagonal matrix D we use the technique of a change-of-coordinates (or change-of-bases): $P^{-1}AP = D$

Visual example: conic section

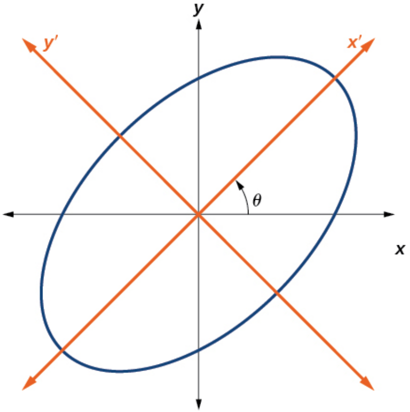

Figure: A conic section in a coordinate system adapted for it.

$$Ax^2 + Bxy + Cy^2 - D = 0 \qquad , B \neq 0$$

$\therefore$ find $\quad x = \alpha x' + \beta y' \;, \quad y = \gamma x' + \delta y'$

to remove xy-term

Solving linear difference equations

Two-dimensional systems

Consider the abstract system of difference equations:

$$\mathbf{z}_{n+1} = A\mathbf{z}_n \ .$$

We want to choose a transformation P and P^{-1} such that:*

$$\mathbf{z} = P\mathbf{Z}, \text{ or } \mathbf{Z} = P^{-1}\mathbf{z} \ .$$

$$\begin{aligned}
\mathbf{Z}_{n+1} &= P^{-1}\mathbf{z}_{n+1} \\
&= P^{-1}(A\mathbf{z}_n) \\
&= (P^{-1}A)\mathbf{z}_n \\
&= (P^{-1}A)(P\mathbf{Z}_n) \\
&= (P^{-1}AP)\mathbf{Z}_n \\
\mathbf{Z}_{n+1} &= D\mathbf{Z}_n \qquad (9)
\end{aligned}$$

*Trace these steps in the Leslie population model example.

It turns out ... since P is an invertible matrix then $AP = PD$. In the two-dimensional system we can write this as:

$$A[\mathbf{v}_1 \;\; \mathbf{v}_2] = [\mathbf{v}_1 \;\; \mathbf{v}_2] \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} \,,$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the two column vectors of the 2×2 matrix P. Therefore:

$$\begin{array}{ccc} \left[A\mathbf{v}_1 \quad A\mathbf{v}_2\right] & = & \left[r_1\mathbf{v}_1 \quad r_2\mathbf{v}_2\right] \\ A\mathbf{v}_1 = r_1\mathbf{v}_1 \,, & \text{and} & A\mathbf{v}_2 = r_2\mathbf{v}_2 \end{array}$$

... r_1 and r_2 must be eigenvalues of A, and $\mathbf{v}_1$ and $\mathbf{v}_2$ are the corresponding eigenvectors!! (see Theorem 23.3)

This holds for k-Dimensional systems ... which gives us

Theorem (23.4)

Let A be a $k \times k$ matrix. Let $r_1, r_2, \ldots, r_k$ be eigenvalues of A, and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ the corresponding eigenvectors. Form the matrix

$$P = [\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_k]$$

whose columns are these k eigenvectors. If P is invertible, then

$$P^{-1}AP = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix} \tag{10}$$

Conversely, if $P^{-1}AP$ is a diagonal matrix D, the columns of P must be eigenvectors of A and the diagonal entries of D must be eigenvalues of A.

Theorem (23.6)

Let A be a $k \times k$ matrix with k distinct real eigenvalues $r_1, \ldots, r_k$ and corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. The general solution of the system of difference equations $\mathbf{z}_{n+1} = A\mathbf{z}_n$ is

$$\mathbf{z}_n = c_1 r_1^n \mathbf{v}_1 + c_2 r_2^n \mathbf{v}_2 + \cdots + c_k r_k^n \mathbf{v}_k \ . \tag{11}$$

Note: We need to know the initial vector $\mathbf{z}_0$ to solve the numerical formula $\mathbf{z}_n$ (recall (5)). To see this, set $n = 0$ in (11):

$$\begin{aligned} \mathbf{z}_0 &= c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \\ &= [\mathbf{v}_1 \cdots \mathbf{v}_k] \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = P \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \ . \end{aligned} \tag{12}$$

Therefore, for *any* specific initial vector $\mathbf{z}_0$, the proper choice of $c_1, \ldots, c_k$ gives the solution for (11).

An alternative approach: the powers of a matrix

Theorem (23.7)

Let A be a $k \times k$ matrix. Suppose that there is a nonsingular matrix P such that

$$P^{-1}AP = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k \end{pmatrix}, \quad \textit{[(10)]}$$

a diagonal matrix. Then,

$$A^n = P \begin{pmatrix} r_1^n & 0 & \cdots & 0 \\ 0 & r_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k^n \end{pmatrix} P^{-1} .$$

The solution of the corresponding system of difference equations $\mathbf{z}_{n+1} = A\mathbf{z}_n$ with initial vector $\mathbf{z}_0$ is

$$\mathbf{z}_n = P \begin{pmatrix} r_1^n & 0 & \cdots & 0 \\ 0 & r_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_k^n \end{pmatrix} P^{-1}\mathbf{z}_0 .$$

Stability of equilibria

It follows:

- For $\mathbf{z}_0 = \mathbf{0}$, we have $\mathbf{z}_n = \mathbf{0}$. Such a solution is called a **steady state**, **equilibrium**, or **stationary solution**.
- The solution is **asymptotically stable** if every solution of $\mathbf{z}_{n+1} = A\mathbf{z}_n$ tends to the steady state $\mathbf{z} = \mathbf{0}$ as n tends to infinity.
- For every solution to have a steady state, the absolute value of all the eigenvalues of A must be less than 1 $(|r_i| < 1)$ such that $r_i^n \to 0$ as $n \to \infty$.

Theorem (23.8)

If the $k \times k$ matrix A has k distinct real eigenvalues, then every solution of the general system of linear difference equations $\mathbf{z}_{n+1} = A\mathbf{z}_n$ tends to $\mathbf{0}$ if and only if all the eigenvalues of A have absolute value less than 1.

Markov processes

Definition

A **stochastic process** is a rule which gives the probability that the system (or the individual in this system) will be in state i at time $n+1$, given the probabilities of its being in the various states in previous periods.

When the probability that the system is in any state i at time $n+1$ depends only on what state the system was in at time n, the stochastic process is called a **Markov process**.

That is, only the *immediate* past matters.

The key elements of a Markov process are:

(1) the probability $x^i(n)$ that state i occurs at time period n, or alternatively, the fraction of the population under study that is in state i at time period n; and

(2) the transition probabilities m_{ij}, where m_{ij} is the probability that the process will be in state i at time $n+1$ if it is in state j at time n.

Markov processes

$$\begin{pmatrix} x^1(n+1) \\ \vdots \\ x^k(n+1) \end{pmatrix} = \underbrace{\begin{pmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & \ddots & \vdots \\ m_{k1} & \cdots & m_{kk} \end{pmatrix}}_{\textbf{Markov}\text{ (transition) }\textbf{matrix}} \begin{pmatrix} x^1(n) \\ \vdots \\ x^k(n) \end{pmatrix} ; \qquad (13)$$

that is, $\mathbf{x}(n+1) = M\mathbf{x}(n)$,

where M is any nonnegative matrix whose column sums $\sum_i m_{ij}$ all equal 1.

For example, each element m_{ij} in first column of M gives the (conditional) probability that the system will be in state $i = 1, \ldots, k$ next period, given that it was in state $j = 1$ today. We therefore have $\sum_{i=1}^{k} m_{i1} = 1$.

Example (23.20, p. 617)

$$\begin{pmatrix} x^{em}(n+1) \\ x^{un}(n+1) \end{pmatrix} = \begin{pmatrix} 0.9 & 0.4 \\ 0.1 & 0.6 \end{pmatrix} \begin{pmatrix} x^{em}(n) \\ x^{un}(n) \end{pmatrix}$$

Markov processes

Some general principles that example 23.20 illustrates

Theorem (23.15)

Let M be a positive Markov matrix. Then,

(a) $r_1 = 1$ *is an eigenvalue of every* M*;*

(b) *every other eigenvalue* r *of* M *satisfies* $|r| < 1$

(c) *eigenvalue* $r = 1$ *has an eigenvector* $\mathbf{w}_1$ *with strictly positive components; and*

(d) *if we write* $\mathbf{v}_1$ *for* $\mathbf{w}_1$ *divided by its components, then* $\mathbf{v}_1$ *is a probability vector and each solution* $\mathbf{x}(n)$ *of* $\mathbf{x}(n+1) = M\mathbf{x}(n)$ *tends to* $\mathbf{v}_1$ *as* $n \to \infty$*.*

Symmetric matrices

Most matrices that arise in optimization and econometrics are symmetric matrices (e.g., Hessians). If A is a symmetric matrix. Then,

Definition

- $A = A^T$;
- A is a $k \times k$ square matrix;
- its “counter-diagonals” have the same entries: $a_{ij} = a_{ji}$;
- it is orthogonally diagonalizable, meaning that we can find an orthogonal matrix P ($P^{-1} = P^T$) which diagonalizes A: $D = P^{-1}AP$.

See also Section 23.8: Definiteness