



Matrix Theory

22 Eigenvalues and Eigenvectors III (Symmetric matrices and the spectral theorem)

Dr. Konda Reddy Mopuri
Dept. of AI, IIT Hyderabad
July-Nov 2026

The best-behaved case



Take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which is symmetric ($A^T = A$).

The best-behaved case



Take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which is symmetric ($A^T = A$). $\text{tr } A = 6$, $\det A = 8$, so

$$\lambda^2 - 6\lambda + 8 = 0 \text{ and } \lambda = 4, 2.$$

The best-behaved case



Take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which is symmetric ($A^T = A$). $\text{tr } A = 6$, $\det A = 8$, so

$\lambda^2 - 6\lambda + 8 = 0$ and $\lambda = 4, 2$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (\lambda = 4), \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (\lambda = 2).$$

The best-behaved case



Take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which is symmetric ($A^T = A$). $\text{tr } A = 6$, $\det A = 8$, so

$\lambda^2 - 6\lambda + 8 = 0$ and $\lambda = 4, 2$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (\lambda = 4), \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (\lambda = 2).$$

Notice: $\mathbf{v}_1 \cdot \mathbf{v}_2 = 1 - 1 = 0$.

The best-behaved case



Take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which is symmetric ($A^T = A$). $\text{tr } A = 6$, $\det A = 8$, so

$\lambda^2 - 6\lambda + 8 = 0$ and $\lambda = 4, 2$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (\lambda = 4), \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (\lambda = 2).$$

Notice: $\mathbf{v}_1 \cdot \mathbf{v}_2 = 1 - 1 = 0$.

The eigenvectors are orthogonal. This is no accident.

Orthogonality is forced



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Theorem

If A is symmetric, eigenvectors for distinct eigenvalues are orthogonal.

Orthogonality is forced



Theorem

If A is symmetric, eigenvectors for distinct eigenvalues are orthogonal.

Proof

Let $A\mathbf{u} = \lambda\mathbf{u}$, $A\mathbf{v} = \mu\mathbf{v}$, $\lambda \neq \mu$. Then

$$\lambda(\mathbf{u} \cdot \mathbf{v}) = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T A\mathbf{v} = \mu(\mathbf{u} \cdot \mathbf{v}).$$

So $(\lambda - \mu)(\mathbf{u} \cdot \mathbf{v}) = 0$, and since $\lambda \neq \mu$ we get $\mathbf{u} \cdot \mathbf{v} = 0$.

Orthogonality is forced



Theorem

If A is symmetric, eigenvectors for distinct eigenvalues are orthogonal.

Proof

Let $A\mathbf{u} = \lambda\mathbf{u}$, $A\mathbf{v} = \mu\mathbf{v}$, $\lambda \neq \mu$. Then

$$\lambda(\mathbf{u} \cdot \mathbf{v}) = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T A\mathbf{v} = \mu(\mathbf{u} \cdot \mathbf{v}).$$

So $(\lambda - \mu)(\mathbf{u} \cdot \mathbf{v}) = 0$, and since $\lambda \neq \mu$ we get $\mathbf{u} \cdot \mathbf{v} = 0$.

Symmetry was used in exactly one place: $A^T = A$.



Spectral Theorem for symmetric matrices

Let A be a real symmetric $n \times n$ matrix. Then:

- 1 all n eigenvalues are **real**;



Spectral Theorem for symmetric matrices

Let A be a real symmetric $n \times n$ matrix. Then:

- ① all n eigenvalues are **real**;
- ② eigenspaces for distinct eigenvalues are mutually orthogonal;



Spectral Theorem for symmetric matrices

Let A be a real symmetric $n \times n$ matrix. Then:

- ① all n eigenvalues are **real**;
- ② eigenspaces for distinct eigenvalues are mutually orthogonal;
- ③ the geometric multiplicity equals the algebraic multiplicity for every eigenvalue, so A **is** diagonalizable;



Spectral Theorem for symmetric matrices

Let A be a real symmetric $n \times n$ matrix. Then:

- ① all n eigenvalues are **real**;
- ② eigenspaces for distinct eigenvalues are mutually orthogonal;
- ③ the geometric multiplicity equals the algebraic multiplicity for every eigenvalue, so A **is** diagonalizable;
- ④ in fact $A = QDQ^T$ with Q **orthogonal** ($Q^T Q = I$, so $Q^{-1} = Q^T$).

The spectral theorem



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Why this matters

No defective cases, no complex arithmetic, and the inverse of Q is free. Symmetric matrices are where the theory is cleanest — which is fortunate, since covariance matrices, Gram matrices, Hessians, and graph Laplacians are all symmetric.

Orthogonal diagonalization in practice



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

For $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, normalize the eigenvectors:

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Orthogonal diagonalization in practice



For $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, normalize the eigenvectors:

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \quad A = QDQ^T.$$

Orthogonal diagonalization in practice



For $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, normalize the eigenvectors:

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \quad A = QDQ^T.$$

Procedure

- ① Find each eigenspace.
- ② Apply Gram–Schmidt *within* each eigenspace of dimension ≥ 2 (across eigenspaces orthogonality is automatic).
- ③ Normalize every vector.

Reading $A = QDQ^T$ geometrically



Expand the factorization:

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T.$$

Reading $A = QDQ^T$ geometrically



Expand the factorization:

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T.$$

Each $\mathbf{q}_i \mathbf{q}_i^T$ is the orthogonal projection onto the line through $\mathbf{q}_i$. So:

Reading $A = QDQ^T$ geometrically



Expand the factorization:

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T.$$

Each $\mathbf{q}_i \mathbf{q}_i^T$ is the orthogonal projection onto the line through $\mathbf{q}_i$. So:

A symmetric matrix decomposes space into perpendicular axes and stretches each by its own factor.

Reading $A = QDQ^T$ geometrically



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Applying A to the unit circle gives an ellipse with axes along the $\mathbf{q}_i$ and semi-axis lengths $|\lambda_i|$. For our example: the unit circle becomes an ellipse stretched by 4 along $(1, 1)$ and by 2 along $(1, -1)$.

Reading $A = QDQ^T$ geometrically



Applying A to the unit circle gives an ellipse with axes along the $\mathbf{q}_i$ and semi-axis lengths $|\lambda_i|$. For our example: the unit circle becomes an ellipse stretched by 4 along $(1, 1)$ and by 2 along $(1, -1)$. This is the germ of principal component analysis and of the SVD.

Iterating a matrix



Many models have the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \quad \text{so} \quad \mathbf{x}_k = A^k \mathbf{x}_0.$$

Iterating a matrix



Many models have the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \quad \text{so} \quad \mathbf{x}_k = A^k \mathbf{x}_0.$$

If A has an eigenbasis $\mathbf{v}_1, \dots, \mathbf{v}_n$, write $\mathbf{x}_0 = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$. Then

$$\mathbf{x}_k = c_1 \lambda_1^k \mathbf{v}_1 + \dots + c_n \lambda_n^k \mathbf{v}_n$$

Iterating a matrix



Many models have the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \quad \text{so} \quad \mathbf{x}_k = A^k \mathbf{x}_0.$$

If A has an eigenbasis $\mathbf{v}_1, \dots, \mathbf{v}_n$, write $\mathbf{x}_0 = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$. Then

$$\mathbf{x}_k = c_1 \lambda_1^k \mathbf{v}_1 + \dots + c_n \lambda_n^k \mathbf{v}_n$$

The central idea of the hour The eigenvectors are the *modes* of the system; the eigenvalues are their *growth rates*. Long-run behaviour is decided by which $|\lambda_i|$ is largest.

Classification of long-run behaviour



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

Classification of long-run behaviour



Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

- All $|\lambda_i| < 1$: $\mathbf{x}_k \rightarrow \mathbf{0}$. The origin is an *attractor* (stable).
- All $|\lambda_i| > 1$: $\|\mathbf{x}_k\| \rightarrow \infty$. The origin is a *repeller*.

Classification of long-run behaviour



Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

- All $|\lambda_i| < 1$: $\mathbf{x}_k \rightarrow \mathbf{0}$. The origin is an *attractor* (stable).
- All $|\lambda_i| > 1$: $\|\mathbf{x}_k\| \rightarrow \infty$. The origin is a *repeller*.
- Mixed, e.g. $|\lambda_1| > 1 > |\lambda_2|$: a *saddle*. Trajectories approach the $\mathbf{v}_1$ -line and flee along it.

Classification of long-run behaviour



Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

- All $|\lambda_i| < 1$: $\mathbf{x}_k \rightarrow \mathbf{0}$. The origin is an *attractor* (stable).
- All $|\lambda_i| > 1$: $\|\mathbf{x}_k\| \rightarrow \infty$. The origin is a *repeller*.
- Mixed, e.g. $|\lambda_1| > 1 > |\lambda_2|$: a *saddle*. Trajectories approach the $\mathbf{v}_1$ -line and flee along it.
- $\lambda_1 = 1$ dominant: $\mathbf{x}_k \rightarrow c_1 \mathbf{v}_1$, a *steady state*.

Classification of long-run behaviour

Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

- All $|\lambda_i| < 1$: $\mathbf{x}_k \rightarrow \mathbf{0}$. The origin is an *attractor* (stable).
- All $|\lambda_i| > 1$: $\|\mathbf{x}_k\| \rightarrow \infty$. The origin is a *repeller*.
- Mixed, e.g. $|\lambda_1| > 1 > |\lambda_2|$: a *saddle*. Trajectories approach the $\mathbf{v}_1$ -line and flee along it.
- $\lambda_1 = 1$ dominant: $\mathbf{x}_k \rightarrow c_1 \mathbf{v}_1$, a *steady state*.
- Complex $\lambda = re^{i\varphi}$: spirals — inward if $r < 1$, outward if $r > 1$, a closed orbit if $r = 1$.

Classification of long-run behaviour

Assume a real eigenbasis, and order $|\lambda_1| > |\lambda_2| \geq \dots$.

- All $|\lambda_i| < 1$: $\mathbf{x}_k \rightarrow \mathbf{0}$. The origin is an *attractor* (stable).
- All $|\lambda_i| > 1$: $\|\mathbf{x}_k\| \rightarrow \infty$. The origin is a *repeller*.
- Mixed, e.g. $|\lambda_1| > 1 > |\lambda_2|$: a *saddle*. Trajectories approach the $\mathbf{v}_1$ -line and flee along it.
- $\lambda_1 = 1$ dominant: $\mathbf{x}_k \rightarrow c_1 \mathbf{v}_1$, a *steady state*.
- Complex $\lambda = re^{i\varphi}$: spirals — inward if $r < 1$, outward if $r > 1$, a closed orbit if $r = 1$.

Stability is entirely a statement about $\max_i |\lambda_i|$ — the *spectral radius*.

The dominant eigenvalue takes over

Factor out the largest:

$$\mathbf{x}_k = \lambda_1^k \left[c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \cdots \right].$$



The dominant eigenvalue takes over



Factor out the largest:

$$\mathbf{x}_k = \lambda_1^k \left[c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \cdots \right].$$

Since $|\lambda_j/\lambda_1| < 1$ for $j \geq 2$, every bracketed term after the first decays geometrically.

The dominant eigenvalue takes over



Factor out the largest:

$$\mathbf{x}_k = \lambda_1^k \left[c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \cdots \right].$$

Since $|\lambda_j/\lambda_1| < 1$ for $j \geq 2$, every bracketed term after the first decays geometrically.

$$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1 \quad \text{for large } k \quad (c_1 \neq 0).$$

The dominant eigenvalue takes over



Factor out the largest:

$$\mathbf{x}_k = \lambda_1^k \left[c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \cdots \right].$$

Since $|\lambda_j/\lambda_1| < 1$ for $j \geq 2$, every bracketed term after the first decays geometrically.

$$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1 \quad \text{for large } k \quad (c_1 \neq 0).$$

Interpretation

Whatever you start with, the system aligns itself with the dominant eigenvector. The *shape* stabilizes even when the *size* grows.

Example: a population model



Two age classes, juveniles j_k and adults a_k :

$$\begin{bmatrix} j_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1.5 \\ 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} j_k \\ a_k \end{bmatrix}.$$

Example: a population model



Two age classes, juveniles j_k and adults a_k :

$$\begin{bmatrix} j_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1.5 \\ 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} j_k \\ a_k \end{bmatrix}.$$

Reading the matrix: each adult produces 1.5 juveniles; 60% of juveniles survive to adulthood; 80% of adults survive.

Example: a population model

Two age classes, juveniles j_k and adults a_k :

$$\begin{bmatrix} j_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1.5 \\ 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} j_k \\ a_k \end{bmatrix}.$$

Reading the matrix: each adult produces 1.5 juveniles; 60% of juveniles survive to adulthood; 80% of adults survive. $\text{tr} = 0.8$, $\det = -0.9$, so $\lambda^2 - 0.8\lambda - 0.9 = 0$ and

$$\lambda = \frac{0.8 \pm \sqrt{0.64 + 3.6}}{2}, \quad \lambda_1 \approx 1.43, \quad \lambda_2 \approx -0.63.$$

Example: a population model

Two age classes, juveniles j_k and adults a_k :

$$\begin{bmatrix} j_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1.5 \\ 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} j_k \\ a_k \end{bmatrix}.$$

Reading the matrix: each adult produces 1.5 juveniles; 60% of juveniles survive to adulthood; 80% of adults survive. $\text{tr} = 0.8$, $\det = -0.9$, so $\lambda^2 - 0.8\lambda - 0.9 = 0$ and

$$\lambda = \frac{0.8 \pm \sqrt{0.64 + 3.6}}{2}, \quad \lambda_1 \approx 1.43, \quad \lambda_2 \approx -0.63.$$

$|\lambda_1| > 1$: the population **grows about 43% per period**, and the age distribution converges to the direction of $\mathbf{v}_1$ regardless of the starting mix.

Example: a Markov chain



Two cities, with yearly migration:

$$P = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}, \quad \mathbf{x}_{k+1} = P\mathbf{x}_k.$$

Example: a Markov chain



Two cities, with yearly migration:

$$P = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}, \quad \mathbf{x}_{k+1} = P\mathbf{x}_k.$$

Columns sum to 1, so P is *stochastic*, and $\lambda = 1$ is always an eigenvalue of such a matrix. (Why? P^T has row sums 1, so $P^T \mathbf{1} = \mathbf{1}$, and P, P^T share eigenvalues.)

Example: a Markov chain

Two cities, with yearly migration:

$$P = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}, \quad \mathbf{x}_{k+1} = P\mathbf{x}_k.$$

Columns sum to 1, so P is *stochastic*, and $\lambda = 1$ is always an eigenvalue of such a matrix. (Why? P^T has row sums 1, so $P^T \mathbf{1} = \mathbf{1}$, and P, P^T share eigenvalues.)

$$P - I = \begin{bmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{bmatrix} \implies \mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \xrightarrow{\text{normalize}} \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}.$$

Example: a Markov chain

Two cities, with yearly migration:

$$P = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}, \quad \mathbf{x}_{k+1} = P\mathbf{x}_k.$$

Columns sum to 1, so P is *stochastic*, and $\lambda = 1$ is always an eigenvalue of such a matrix. (Why? P^T has row sums 1, so $P^T \mathbf{1} = \mathbf{1}$, and P, P^T share eigenvalues.)

$$P - I = \begin{bmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{bmatrix} \implies \mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \xrightarrow{\text{normalize}} \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}.$$

The other eigenvalue is $\text{tr} - 1 = 0.92$, with $|0.92| < 1$. So $\mathbf{x}_k \rightarrow \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}$:

in the long run 37.5% live in city 1 — **whatever the initial distribution.**

Continuous time: $\mathbf{x}' = A\mathbf{x}$



The same decomposition solves linear ODE systems. If $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ satisfies $\mathbf{x}' = A\mathbf{x}$:

$$\mathbf{x}' = \lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}(A\mathbf{v}) = A\mathbf{x}.\checkmark$$

Continuous time: $\mathbf{x}' = A\mathbf{x}$

The same decomposition solves linear ODE systems. If $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ satisfies $\mathbf{x}' = A\mathbf{x}$:

$$\mathbf{x}' = \lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}(A\mathbf{v}) = A\mathbf{x}.$$

With an eigenbasis, the general solution is

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{v}_n.$$

Continuous time: $\mathbf{x}' = A\mathbf{x}$

The same decomposition solves linear ODE systems. If $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ satisfies $\mathbf{x}' = A\mathbf{x}$:

$$\mathbf{x}' = \lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}(A\mathbf{v}) = A\mathbf{x}.\checkmark$$

With an eigenbasis, the general solution is

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{v}_n.$$

Stability(continuous version)

Here the threshold is $\operatorname{Re} \lambda$, not $|\lambda|$:

- all $\operatorname{Re} \lambda_i < 0$: solutions decay (stable);
- some $\operatorname{Re} \lambda_i > 0$: solutions blow up;
- $\operatorname{Im} \lambda_i \neq 0$: oscillation at frequency $|\operatorname{Im} \lambda_i|$.

Continuous time: $\mathbf{x}' = A\mathbf{x}$

The same decomposition solves linear ODE systems. If $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ satisfies $\mathbf{x}' = A\mathbf{x}$:

$$\mathbf{x}' = \lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}(A\mathbf{v}) = A\mathbf{x}.\checkmark$$

With an eigenbasis, the general solution is

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{v}_n.$$

Stability(continuous version)

Here the threshold is $\operatorname{Re} \lambda$, not $|\lambda|$:

- all $\operatorname{Re} \lambda_i < 0$: solutions decay (stable);
- some $\operatorname{Re} \lambda_i > 0$: solutions blow up;
- $\operatorname{Im} \lambda_i \neq 0$: oscillation at frequency $|\operatorname{Im} \lambda_i|$.

Discrete: compare $|\lambda|$ to 1. Continuous: compare $\operatorname{Re} \lambda$ to 0.

Other places eigenvectors appear



- **PageRank.** The ranking vector is the dominant eigenvector of a stochastic link matrix. Google's original algorithm is a 10^9 -dimensional eigenvector computation.

Other places eigenvectors appear



- **PageRank.** The ranking vector is the dominant eigenvector of a stochastic link matrix. Google's original algorithm is a 10^9 -dimensional eigenvector computation.
- **Principal component analysis.** The principal axes of a data set are the eigenvectors of its (symmetric) covariance matrix; the eigenvalues are the variances along them.

Other places eigenvectors appear

- **PageRank.** The ranking vector is the dominant eigenvector of a stochastic link matrix. Google's original algorithm is a 10^9 -dimensional eigenvector computation.
- **Principal component analysis.** The principal axes of a data set are the eigenvectors of its (symmetric) covariance matrix; the eigenvalues are the variances along them.
- **Vibration and resonance.** Normal modes of a structure are eigenvectors; natural frequencies come from the eigenvalues. Resonance is an eigenvalue phenomenon.

Other places eigenvectors appear

- **PageRank.** The ranking vector is the dominant eigenvector of a stochastic link matrix. Google's original algorithm is a 10^9 -dimensional eigenvector computation.
- **Principal component analysis.** The principal axes of a data set are the eigenvectors of its (symmetric) covariance matrix; the eigenvalues are the variances along them.
- **Vibration and resonance.** Normal modes of a structure are eigenvectors; natural frequencies come from the eigenvalues. Resonance is an eigenvalue phenomenon.
- **Quantum mechanics.** Observables are Hermitian operators; measured values are their eigenvalues (real, by the spectral theorem), and states are their eigenvectors.

Other places eigenvectors appear



- **PageRank.** The ranking vector is the dominant eigenvector of a stochastic link matrix. Google's original algorithm is a 10^9 -dimensional eigenvector computation.
- **Principal component analysis.** The principal axes of a data set are the eigenvectors of its (symmetric) covariance matrix; the eigenvalues are the variances along them.
- **Vibration and resonance.** Normal modes of a structure are eigenvectors; natural frequencies come from the eigenvalues. Resonance is an eigenvalue phenomenon.
- **Quantum mechanics.** Observables are Hermitian operators; measured values are their eigenvalues (real, by the spectral theorem), and states are their eigenvectors.
- **Spectral graph theory.** Eigenvectors of the Laplacian detect clusters and connectivity in networks.

The power method



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

How does one actually compute a dominant eigenvector?

The power method



How does one actually compute a dominant eigenvector? Recall

$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1$. Just iterate, rescaling to prevent overflow:

The power method



How does one actually compute a dominant eigenvector? Recall

$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1$. Just iterate, rescaling to prevent overflow:

Algorithm Pick $\mathbf{x}_0$. For $k = 0, 1, 2, \dots$:

- ① $\mathbf{y} = A\mathbf{x}_k$;
- ② $\mu_k =$ the entry of $\mathbf{y}$ largest in absolute value;
- ③ $\mathbf{x}_{k+1} = \mathbf{y}/\mu_k$.

Then $\mathbf{x}_k \rightarrow$ (a multiple of) $\mathbf{v}_1$ and $\mu_k \rightarrow \lambda_1$.

The power method

How does one actually compute a dominant eigenvector? Recall

$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1$. Just iterate, rescaling to prevent overflow:

Algorithm Pick $\mathbf{x}_0$. For $k = 0, 1, 2, \dots$:

- ① $\mathbf{y} = A\mathbf{x}_k$;
- ② $\mu_k =$ the entry of $\mathbf{y}$ largest in absolute value;
- ③ $\mathbf{x}_{k+1} = \mathbf{y}/\mu_k$.

Then $\mathbf{x}_k \rightarrow$ (a multiple of) $\mathbf{v}_1$ and $\mu_k \rightarrow \lambda_1$.

Convergence rate is governed by $|\lambda_2/\lambda_1|$ — fast when the dominant eigenvalue is well separated, slow when it is not.

The power method

How does one actually compute a dominant eigenvector? Recall

$\mathbf{x}_k \approx c_1 \lambda_1^k \mathbf{v}_1$. Just iterate, rescaling to prevent overflow:

Algorithm Pick $\mathbf{x}_0$. For $k = 0, 1, 2, \dots$:

- ① $\mathbf{y} = A\mathbf{x}_k$;
- ② $\mu_k =$ the entry of $\mathbf{y}$ largest in absolute value;
- ③ $\mathbf{x}_{k+1} = \mathbf{y}/\mu_k$.

Then $\mathbf{x}_k \rightarrow$ (a multiple of) $\mathbf{v}_1$ and $\mu_k \rightarrow \lambda_1$.

Convergence rate is governed by $|\lambda_2/\lambda_1|$ — fast when the dominant eigenvalue is well separated, slow when it is not. Only matrix–vector products are needed, which is why it scales to the enormous sparse matrices of web search.

The three hours in one slide



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- ① **Definition.** $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: directions A merely scales. Find them via $\det(A - \lambda I) = 0$, then $\text{Nul}(A - \lambda I)$.

The three hours in one slide



- ① **Definition.** $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: directions A merely scales. Find them via $\det(A - \lambda I) = 0$, then $\text{Nul}(A - \lambda I)$.
- ② **Diagonalization.** An eigenbasis means $A = PDP^{-1}$, which turns A^k into D^k . It exists iff geometric and algebraic multiplicities agree everywhere. Symmetric matrices always qualify, and orthogonally so.

The three hours in one slide

- ① **Definition.** $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: directions A merely scales. Find them via $\det(A - \lambda I) = 0$, then $\text{Nul}(A - \lambda I)$.
- ② **Diagonalization.** An eigenbasis means $A = PDP^{-1}$, which turns A^k into D^k . It exists iff geometric and algebraic multiplicities agree everywhere. Symmetric matrices always qualify, and orthogonally so.
- ③ **Dynamics.** Eigenvectors are modes, eigenvalues are growth rates. $|\lambda|$ versus 1 (discrete) or $\text{Re } \lambda$ versus 0 (continuous) decides stability; the dominant eigenvector decides the long-run shape.

The three hours in one slide

- ① **Definition.** $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: directions A merely scales. Find them via $\det(A - \lambda I) = 0$, then $\text{Nul}(A - \lambda I)$.
- ② **Diagonalization.** An eigenbasis means $A = PDP^{-1}$, which turns A^k into D^k . It exists iff geometric and algebraic multiplicities agree everywhere. Symmetric matrices always qualify, and orthogonally so.
- ③ **Dynamics.** Eigenvectors are modes, eigenvalues are growth rates. $|\lambda|$ versus 1 (discrete) or $\text{Re } \lambda$ versus 0 (continuous) decides stability; the dominant eigenvector decides the long-run shape.

One idea: find the coordinates in which the problem is diagonal.

Common errors to avoid



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Writing $A - \lambda$ instead of $A - \lambda I$.

Common errors to avoid



- Writing $A - \lambda$ instead of $A - \lambda I$.
- Calling $\mathbf{0}$ an eigenvector (it never is), or refusing $\lambda = 0$ (it often is).

Common errors to avoid



- Writing $A - \lambda$ instead of $A - \lambda I$.
- Calling $\mathbf{0}$ an eigenvector (it never is), or refusing $\lambda = 0$ (it often is).
- Mismatching the order of the columns of P and the entries of D .

Common errors to avoid



- Writing $A - \lambda$ instead of $A - \lambda I$.
- Calling $\mathbf{0}$ an eigenvector (it never is), or refusing $\lambda = 0$ (it often is).
- Mismatching the order of the columns of P and the entries of D .
- Assuming a repeated eigenvalue means not diagonalizable. Check $\dim E_\lambda$.

Common errors to avoid



- Writing $A - \lambda$ instead of $A - \lambda I$.
- Calling $\mathbf{0}$ an eigenvector (it never is), or refusing $\lambda = 0$ (it often is).
- Mismatching the order of the columns of P and the entries of D .
- Assuming a repeated eigenvalue means not diagonalizable. Check $\dim E_\lambda$.
- Assuming the eigenvalues of $A + B$ are sums of eigenvalues, or that A and B share eigenvectors.

Common errors to avoid



- Writing $A - \lambda$ instead of $A - \lambda I$.
- Calling $\mathbf{0}$ an eigenvector (it never is), or refusing $\lambda = 0$ (it often is).
- Mismatching the order of the columns of P and the entries of D .
- Assuming a repeated eigenvalue means not diagonalizable. Check $\dim E_\lambda$.
- Assuming the eigenvalues of $A + B$ are sums of eigenvalues, or that A and B share eigenvectors.
- Forgetting that eigenvectors are determined only up to a nonzero scalar — your answer and the textbook's may differ legitimately.

Exercises



- ① Find the eigenvalues and eigenspaces of $\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ and verify the eigenvectors are orthogonal. Explain why that was predictable.
- ② Show that $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ is not diagonalizable.
- ③ Diagonalize $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$ and use it to compute A^{10} .
- ④ Prove: if λ is an eigenvalue of an invertible A , then λ^{-1} is an eigenvalue of A^{-1} , with the same eigenvector.
- ⑤ Find the steady-state vector of $\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$ and determine how fast convergence occurs.
- ⑥ Give a 3×3 matrix with a single eigenvalue of algebraic multiplicity 3 and geometric multiplicity 2.
- ⑦ For which c is $\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$ diagonalizable?

Rough Work



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Rough Work



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad