GES-207-4 — Mathematical Methods of Physics

National Institute for Space Research (INPE)
Graduate Program in Space Geophysics (PPGGES)

Lecture Notes 04

Eigenvalues, Diagonalization, and Matrix Functions

Thursday, October 1, 2026 · Prof. Dr. José Paulo Marchezi · DICEP · 3rd Term 2026

Learning goals and study scope. This lecture explains why eigenvectors are the natural directions of a linear operator, how eigenvalues are found, when a matrix can be diagonalized, and how diagonalization turns matrix powers, exponentials, and other functions into scalar calculations. We emphasize derivations, diagnostic checks, exceptional cases, and applications to coupled dynamical systems, normal modes, stability, and quantum observables.

Contents

1 Eigenvalues as Natural Responses of a Linear Operator **2**
1.1 Eigenspaces and multiplicities 3

2 Diagonalization **3**
2.1 Repeated eigenvalues: diagonalizable and defective cases 4

3 Symmetric and Hermitian Matrices **5**

4 Matrix Powers and Functions **6**
4.1 Spectral projectors 7

5 The Matrix Exponential and Linear Dynamical Systems **7**
5.1 Rotation from complex eigenvalues 8

6 Normal Modes and Physics Connections **8**
6.1 Coupled oscillators 8
6.2 Quantum-mechanical observables 8
6.3 Diffusion and stability 9

7 Cayley–Hamilton and Alternatives to Diagonalization **9**

8 A Reliable Workflow and Common Pitfalls **9**

9 Lecture Summary **10**

1. Eigenvalues as Natural Responses of a Linear Operator

For a general vector $\mathbf{v}$, applying a linear operator A can change both magnitude and direction. Certain nonzero directions are special: the operator changes only their scale (and possibly their sign or complex phase).

Eigenvalue and eigenvector

A scalar λ is an *eigenvalue* of A if there exists a nonzero vector $\mathbf{v}$ such that

$$\boxed{A\mathbf{v} = \lambda\mathbf{v}.}$$

The vector $\mathbf{v}$ is an eigenvector associated with λ. The zero vector is excluded because it would satisfy the equation for every scalar and would therefore identify no distinguished direction.

Geometrically, a real positive eigenvalue stretches an eigenvector, a negative one reverses its direction as well as scaling it, and zero collapses it to the origin. Complex eigenvalues naturally encode simultaneous growth/decay and rotation. Their appearance for a real matrix does not mean the calculation has failed; complex eigenvalues occur in conjugate pairs.

Rearranging the eigenvalue equation gives

$$(A - \lambda I)\mathbf{v} = \mathbf{0}.$$

A nonzero solution exists precisely when $A - \lambda I$ is singular. Therefore

$$\boxed{p_A(\lambda) = \det(A - \lambda I) = 0.} \tag{1}$$

The polynomial p_A is the characteristic polynomial. Some authors define it as $\det(\lambda I - A)$; the two conventions differ by $(-1)^n$ but have the same roots.

Example 4.1: a complete 2×2 calculation

Let

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$

Then

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 10.$$

Thus $(\lambda - 5)(\lambda - 2) = 0$, so $\lambda_1 = 5$ and $\lambda_2 = 2$. For $\lambda_1 = 5$,

$$(A - 5I)\mathbf{v} = 0, \qquad \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

so $y = x$ and one choice is $\mathbf{v}_1 = (1, 1)^T$. For $\lambda_2 = 2$,

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

so $y = -2x$ and one choice is $\mathbf{v}_2 = (1, -2)^T$. Always check:

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad A\begin{pmatrix} 1 \\ -2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Any nonzero scalar multiple of an eigenvector represents the same eigendirection.

1.1. Eigenspaces and multiplicities

For a fixed eigenvalue,

$$E_\lambda = \ker(A - \lambda I)$$

is its eigenspace. It contains the zero vector and all eigenvectors associated with λ. The *algebraic multiplicity* is the multiplicity of λ as a root of p_A; the *geometric multiplicity* is $\dim E_\lambda$. They satisfy

$$1 \leq \dim E_\lambda \leq \text{algebraic multiplicity of } \lambda.$$

Failure of equality is exactly what can prevent diagonalization.

Two fast checks follow from the characteristic polynomial:

$$\sum_{j=1}^{n} \lambda_j = \operatorname{tr} A, qquad \prod_{j=1}^{n} \lambda_j = \det A,$$

with algebraic multiplicities and over $\mathbb{C}$. Trace and determinant do not by themselves determine all eigenvalues for $n > 2$, but they catch many arithmetic errors.

2. Diagonalization

Suppose A has n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Assemble them as columns of

$$S = [\mathbf{v}_1 | \cdots | \mathbf{v}_n], qquad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

The column equations $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$ combine into

$$AS = SD.$$

Since the columns are independent, S is invertible, and

$$\boxed{S^{-1}AS = D \qquad \text{or equivalently} \qquad A = SDS^{-1}.} \tag{2}$$

Diagonalization criterion

An $n \times n$ matrix is diagonalizable over a field if and only if it has a basis of n eigenvectors over that field. Equivalently, for every eigenvalue, geometric and algebraic multiplicities agree and the eigenspace dimensions sum to n. Distinct eigenvalues always have linearly independent eigenvectors; hence n distinct eigenvalues are sufficient, but not necessary.

Diagonalization is a change of basis. In the eigenvector basis, the same operator acts independently on coordinates:

$$D\mathbf{c} = (\lambda_1 c_1, \ldots, \lambda_n c_n)^T.$$

The matrix S translates eigenbasis coordinates into the original basis; S^{-1} translates back. The order of these matrices matters.

Example 4.2: diagonalizing Example 4.1

Using the eigenvectors already found,

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, qquad D = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}, qquad S^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}.$$

One can verify directly that $AS = SD$ and

$$SDS^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$

The columns of S may be rescaled without changing the final matrix, provided the inverse is recomputed consistently.

2.1. Repeated eigenvalues: diagonalizable and defective cases

Repeated roots require inspecting eigenspaces; the characteristic polynomial alone is not enough.

Example 4.3: repetition does not imply failure

The identity I_2 has the repeated eigenvalue 1 with algebraic multiplicity two, but $E_1 = \mathbb{R}^2$ has dimension two. Every nonzero vector is an eigenvector, so the matrix is already diagonal and is certainly diagonalizable.

Example 4.4: a defective matrix

For

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \qquad p_J(\lambda) = (2-\lambda)^2.$$

Yet

$$J - 2I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

so $y = 0$ and $E_2 = \mathrm{span}\{(1,0)^T\}$ has dimension one. There is no basis of eigenvectors; J is not diagonalizable. It is already a Jordan block. This case will matter when defining matrix functions: the superdiagonal produces derivatives of the scalar function.

3. Symmetric and Hermitian Matrices

The best-behaved operators are self-adjoint. A real matrix is symmetric if $A^T = A$; a complex matrix is Hermitian if $A^H = A$.

Spectral theorem

Every real symmetric matrix admits an orthogonal diagonalization

$$\boxed{A = QDQ^T, \qquad Q^TQ = I,}$$

and every complex Hermitian matrix admits a unitary diagonalization $A = UDU^H$. All eigenvalues are real, and eigenvectors belonging to distinct eigenvalues are orthogonal. Within a repeated eigenspace, Gram–Schmidt can be used to choose an orthonormal basis.

To see the orthogonality mechanism in the real case, suppose $A\mathbf{u} = \lambda\mathbf{u}$ and $A\mathbf{v} = \mu\mathbf{v}$. Then

$$\lambda\langle \mathbf{u}, \mathbf{v}\rangle = \langle A\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, A\mathbf{v}\rangle = \mu\langle \mathbf{u}, \mathbf{v}\rangle.$$

If $\lambda \neq \mu$, the inner product must vanish. This is the bridge from Lecture 3: orthonormal eigenvectors make $Q^{-1} = Q^T$, improving both geometry and numerical conditioning.

Example 4.5: principal directions of a quadratic form

Let

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Its eigenvalues are 3 and 1, with orthonormal eigenvectors $\mathbf{q}_1 = (1,1)^T/\sqrt{2}$ and $\mathbf{q}_2 = (1,-1)^T/\sqrt{2}$. Thus

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad D = \mathrm{diag}(3,1), \quad A = QDQ^T.$$

For the quadratic form $\mathbf{x}^TA\mathbf{x}$, set $\mathbf{y} = Q^T\mathbf{x}$. Then

$$\mathbf{x}^TA\mathbf{x} = \mathbf{y}^TD\mathbf{y} = 3y_1^2 + y_2^2.$$

The eigenvectors are the principal axes and the eigenvalues measure the strength along those axes. The same idea underlies principal stresses, moments of inertia, covariance analysis, and normal modes.

4. Matrix Powers and Functions

If $A = SDS^{-1}$, repeated multiplication telescopes:

$$A^2 = SDS^{-1}SDS^{-1} = SD^2S^{-1}, qquad \boxed{A^k = SD^kS^{-1}.}$$

Since $D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$, a difficult matrix power becomes scalar exponentiation.

Example 4.6: a closed form for A^k

For the matrix of Example 4.1,

$$A^k = S\,\operatorname{diag}(5^k, 2^k)S^{-1} = \frac{1}{3}\begin{pmatrix} 2\cdot 5^k + 2^k & 5^k - 2^k \\ 2\cdot 5^k - 2\cdot 2^k & 5^k + 2\cdot 2^k \end{pmatrix}.$$

At $k = 0$ the formula gives I, and at $k = 1$ it recovers A; these are useful checks. For large k, the term 5^k dominates unless the initial vector has zero component along its eigendirection.

For a polynomial $f(t) = \sum_{m=0}^{M} a_m t^m$, define $f(A) = \sum_{m=0}^{M} a_m A^m$. Diagonalization gives

$$f(A) = S\left(\sum_{m=0}^{M} a_m D^m\right)S^{-1} = S\,\operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n))S^{-1}.$$

The same principle extends to convergent power series and suitable scalar functions.

Matrix function for a diagonalizable matrix

If $A = SDS^{-1}$ and f is defined on the spectrum of A, set

$$\boxed{f(A) = Sf(D)S^{-1}, \qquad f(D) = \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n)).}$$

This definition is independent of the scaling and ordering of eigenvectors. For non-diagonalizable matrices, values at eigenvalues alone are insufficient; derivatives of f enter through Jordan blocks.

Important examples are

$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!}, \qquad \sin A = \sum_{m=0}^{\infty} \frac{(-1)^m A^{2m+1}}{(2m+1)!}, \qquad \cos A = \sum_{m=0}^{\infty} \frac{(-1)^m A^{2m}}{(2m)!}.$$

If A is diagonalizable, simply apply the scalar function to each eigenvalue.

4.1. Spectral projectors

For distinct eigenvalues, write $S = [\mathbf{v}_1|\cdots|\mathbf{v}_n]$ and let $\mathbf{w}_j^T$ be the jth row of S^{-1}. Then

$$P_j = \mathbf{v}_j \mathbf{w}_j^T, qquad P_i P_j = \delta_{ij} P_j, qquad \sum_j P_j = I,$$

and

$$\boxed{A = \sum_j \lambda_j P_j, qquad f(A) = \sum_j f(\lambda_j) P_j.}$$

For a symmetric matrix, left and right eigenvectors coincide and $P_j = \mathbf{q}_j \mathbf{q}_j^T$ is the orthogonal projector from Lecture 3.

5. The Matrix Exponential and Linear Dynamical Systems

Consider the initial-value problem

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}, qquad \mathbf{x}(0) = \mathbf{x}_0.$$

Its solution is

$$\boxed{\mathbf{x}(t) = e^{At}\mathbf{x}_0.}$$

Indeed, term-by-term differentiation of the power series gives $\frac{d}{dt}e^{At} = Ae^{At}$ and $e^{A0} = I$. If $A = SDS^{-1}$,

$$e^{At} = S\,\mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})S^{-1}.$$

Writing $\mathbf{x}_0 = \sum_j c_j \mathbf{v}_j$ reveals the physics directly:

$$\mathbf{x}(t) = \sum_j c_j e^{\lambda_j t}\mathbf{v}_j.$$

Each eigenvector is an independent mode and its eigenvalue controls temporal behavior.

For continuous-time systems, $\mathrm{Re}\,\lambda < 0$ gives exponential decay, $\mathrm{Re}\,\lambda > 0$ gives growth, and a nonzero imaginary part gives oscillation. For discrete iterations $\mathbf{x}_{k+1} = A\mathbf{x}_k$, the relevant criterion is $|\lambda| < 1$ for decay. Do not interchange the continuous- and discrete-time criteria.

Example 4.7: solving a coupled system

Let A be the matrix in Example 4.1 and $\mathbf{x}_0 = (3, 0)^T$. Since

$$\mathbf{x}_0 = c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\-2\end{pmatrix},$$

the equations $c_1 + c_2 = 3$ and $c_1 - 2c_2 = 0$ give $c_1 = 2$, $c_2 = 1$. Therefore

$$\boxed{\mathbf{x}(t) = 2e^{5t}\begin{pmatrix}1\\1\end{pmatrix} + e^{2t}\begin{pmatrix}1\\-2\end{pmatrix}.}$$

At $t = 0$ this returns $(3, 0)^T$. Differentiating verifies $\mathbf{x}' = A\mathbf{x}$. The faster e^{5t} mode dominates forward time.

5.1. Rotation from complex eigenvalues

For

$$A = \begin{pmatrix}0 & -\omega\\ \omega & 0\end{pmatrix}, \qquad A^2 = -\omega^2 I,$$

the power series separates into even and odd powers:

$$e^{At} = I\cos(\omega t) + \frac{A}{\omega}\sin(\omega t) = \begin{pmatrix}\cos\omega t & -\sin\omega t\\ \sin\omega t & \cos\omega t\end{pmatrix}.$$

The eigenvalues are $\pm i\omega$. The complex eigenmodes are a compact description of a real rotation; real initial data remain real because the conjugate contributions combine.

6. Normal Modes and Physics Connections

6.1. Coupled oscillators

For a conservative small-oscillation system

$$M\ddot{\mathbf{q}} + K\mathbf{q} = 0,$$

seek $\mathbf{q}(t) = \mathbf{a}e^{i\omega t}$. Substitution yields the generalized eigenvalue problem

$$K\mathbf{a} = \omega^2 M\mathbf{a}.$$

The eigenvectors are normal-mode shapes; ω_j are natural frequencies. When M is positive definite, the problem can be transformed into a symmetric one, giving real ω_j^2 and an appropriate orthogonality relation.

6.2. Quantum-mechanical observables

In finite-dimensional quantum mechanics, an observable is represented by a Hermitian operator. The spectral theorem ensures real measurement values and an orthonormal eigenbasis. If $A = \sum_j \lambda_j P_j$, then a state decomposes into spectral components $|\psi\rangle = \sum_j P_j|\psi\rangle$. The same projector language used for least squares now organizes measurement subspaces.

6.3. Diffusion and stability

After spatial discretization, a diffusion equation often has the form $\dot{\mathbf{u}} = L\mathbf{u}$, where L approximates a Laplacian. Its eigenvectors are spatial modes and its typically nonpositive eigenvalues produce decay factors $e^{\lambda_j t}$. High-frequency modes often have more negative eigenvalues and decay faster, explaining the smoothing action of diffusion.

7. Cayley–Hamilton and Alternatives to Diagonalization

Cayley–Hamilton theorem

Every square matrix satisfies its own characteristic equation:

$$\boxed{p_A(A) = 0.}$$

Consequently, powers $A^n, A^{n+1}, \ldots$ can be reduced to combinations of $I, A, \ldots, A^{n-1}$. Matrix functions can therefore often be represented by a polynomial of degree at most $n-1$, with extra derivative conditions when eigenvalues are defective.

For Example 4.1, $p_A(\lambda) = \lambda^2 - 7\lambda + 10$, so

$$A^2 - 7A + 10I = 0, qquad A^2 = 7A - 10I.$$

This produces recurrences for powers without explicitly multiplying large matrices. It also explains why, for a 2×2 matrix with distinct eigenvalues, any analytic function can be written as $f(A) = \alpha I + \beta A$, where $\alpha + \beta\lambda_j = f(\lambda_j)$.

For the Jordan block $J = 2I + N$ of Example 4.4, $N^2 = 0$. Hence

$$f(J) = f(2)I + f'(2)N, \qquad e^{Jt} = e^{2t}(I + tN) = e^{2t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

The factor t is the signature of the missing eigenvector. This is why the naive rule "apply f only to the eigenvalues" is valid for diagonalizable matrices but incomplete for defective ones.

8. A Reliable Workflow and Common Pitfalls

1. Compute $p_A(\lambda)$ and use trace/determinant as checks.
2. For each eigenvalue, solve $(A - \lambda I)\mathbf{v} = 0$; do not row-reduce A before finding its eigenvalues, because ordinary row operations generally change them.
3. Count independent eigenvectors. Repeated eigenvalues require explicit eigenspace dimensions.
4. Build S from eigenvectors and place the corresponding eigenvalues in the same column order in D.

5. Verify $AS = SD$ before computing S^{-1}; this catches mismatched columns.

6. Use $A = SDS^{-1}$, not $S^{-1}DS$. For symmetric/Hermitian matrices use the stronger form $A = QDQ^T$ or UDU^H.

7. Distinguish exact theory from numerical practice: eigenvalue algorithms do not normally form the characteristic polynomial for large matrices, and a nearly defective matrix can be highly sensitive to perturbations.

9. Lecture Summary

- Eigenvectors are invariant directions; eigenvalues give the associated scaling or modal rate.
- A matrix is diagonalizable exactly when its eigenvectors form a basis. Distinct eigenvalues guarantee this, while repeated ones require care.
- Symmetric and Hermitian matrices have real spectra and orthonormal eigenbases.
- Diagonalization turns powers and functions of matrices into scalar operations on eigenvalues.
- The matrix exponential solves linear systems, and the spectrum controls growth, decay, oscillation, and stability.
- Spectral projectors connect diagonalization directly to the orthogonal projections of Lecture 3.

Thursday Assignment — Prepare and Present in the Next Class

Instructions. Solve both problems in full. Your solution should show the intermediate calculations, justify each conclusion, and include the requested verification. Be prepared to present and explain your reasoning in the next class.

1. **Orthogonal diagonalization and matrix powers.** Consider

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

   (a) Find the eigenvalues and an orthonormal basis of eigenvectors.

   (b) Construct Q and D such that $A = QDQ^T$, and verify the factorization by direct multiplication.

   (c) Use the diagonalization to obtain a closed expression for A^n and compute A^{10}.

   (d) Explain why orthogonal diagonalization is available in this problem and identify the dominant eigendirection as $n \to \infty$.

2. **A defective matrix and its exponential.** Consider

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

(a) Find the characteristic polynomial, eigenvalue, algebraic multiplicity, eigenspace, and geometric multiplicity.

(b) Explain precisely why B cannot be diagonalized.

(c) Write $B = I + N$, show that $N^2 = 0$, and use the power series to compute e^{Bt}.

(d) Solve $\dot{\mathbf{x}} = B\mathbf{x}$ for $\mathbf{x}(0) = (x_0, y_0)^T$, and verify the solution by differentiation.

(e) Interpret the factor proportional to te^t as the dynamical consequence of the missing eigenvector.