

Linear Algebra and Matrix Methods

Eigenvalues and Eigenvectors

Eigenvectors, Eigenvalues: “proper” directions, scalings

So far, we've seen ways to solve a system of linear equations,

$$A\vec{x} = b.$$

If $A\vec{x} = b$ has no solution, then the least squares (best fit) solution to

$$A^t A \hat{x} = A^t b$$

will suffice for “as close as we can get”.

If $A\vec{x} = b$ has a solution, then $b \in C(A)$.

We are now interested in the special case where b is a scaling of \vec{x} .

Eigenvectors, Eigenvalues: “proper” directions, scalings

Goal: Solve the $n \times n$ (square) system

$$A\vec{x} = \lambda\vec{x}$$

for \vec{x} and λ .

These are considered the “proper”^{*} directions of the matrix A , where transformation by the matrix A is equivalent to merely scaling that direction.

^{*}The “eigen” in these terms is German for “proper” or “characteristic”.

Eigenvectors, Eigenvalues: “proper” directions, scalings

If there are solutions to this **eigenproblem**[†]

$$A\vec{x} = \lambda\vec{x},$$

we will call the values in each pair (\vec{x}, λ) by the names

eigenvector for each $\vec{x} \in \mathbb{C}^n$

and

eigenvalue for each $\lambda \in \mathbb{C}$.

[†]Some of the examples in this chapter require you to understand arithmetic with complex numbers, i.e. from $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\}$.

Eigenvectors and eigenvalues via nullspaces

How can we solve this system?

Typically, if you want to solve an equation for a variable, you'll get all instances of that variable on the same side of the equals sign.

We'll do that here: seeing that scaling a vector by λ is the same as multiplying the vector by the scaled identity matrix λI ,

$$A\vec{x} = \lambda\vec{x} \implies A\vec{x} - \lambda\vec{x} = 0 \implies (A - \lambda I)\vec{x} = 0.$$

Thus, if \vec{x} is an eigenvector of A , then $\vec{x} \in N(A - \lambda I)$.

Eigenvectors and eigenvalues via determinants

$\vec{x} = 0$ is always a solution. However, the eigenvector 0 has *any* $\lambda \in \mathbb{C}$ as an eigenvalue; this tells us nothing about the matrix A .

If a nontrivial \vec{x} solves the eigenproblem, we see that the columns of $A - \lambda I$ are linearly dependent; that is,

$$\det(A - \lambda I) = 0.$$

Eigenvectors and eigenvalues via determinants

$$\det(A - \lambda I) = 0.$$

This is the **characteristic equation** we need to solve, for $\lambda \in \mathbb{C}$.

This also explains why we say $\lambda \in \mathbb{C}$ and $\vec{x} \in \mathbb{C}^n$:

$$\det(A - \lambda I) = 0$$

is an n th degree polynomial equation in λ , with n roots (with multiplicity) in \mathbb{C} .

Roots of the eigenequation: spectrum of the matrix

Unfortunately for the desire for a solution to the characteristic equation, we know[‡] that there is no general method of find roots of polynomials of degree $n \geq 5$.

That said, we can use various techniques to solve this problem.

[‡]thanks to Neils Abel (1802-1829)

Roots of the eigenequation: spectrum of the matrix

Define the **trace** of a square matrix A as the sum of its diagonal:

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

Then, if $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ are the n roots of the characteristic equation of A (i.e. the **spectrum** of A), we have

$$\det(A) = \prod_{i=1}^n \lambda_i \quad \text{and} \quad \text{tr}(A) = \sum_{i=1}^n \lambda_i.$$

Roots of the eigenequation: exponentiation

- ▶ Repeated multiplication by A on an eigenvector repeats scaling by λ : for any $k \in \mathbb{N}$,

$$A(A\vec{x}) = A(\lambda\vec{x}) \implies A^2\vec{x} = \lambda^2\vec{x} \implies A^k\vec{x} = \lambda^k\vec{x}.$$

- ▶ If A is invertible, then all of the eigenvalues of A are nonzero.

If \vec{x} is an eigenvector of A with eigenvalue λ ,
then \vec{x} is also an eigenvector of A^{-1} , with eigenvalue λ^{-1} :

$$\begin{aligned} A\vec{x} = \lambda\vec{x} &\implies A^{-1}A\vec{x} = \lambda A^{-1}\vec{x} = \vec{x} \\ &\implies A^{-1}\vec{x} = \lambda^{-1}\vec{x}. \end{aligned}$$

- ▶ If A is singular, then $\det(A) = 0$, and $\lambda = 0$ is one of its eigenvalues.

Roots of the eigenequation: triangular, projection

- ▶ If A is triangular, then A 's eigenvalues are the diagonal entries.
- ▶ If A is symmetric ($A^t = A$), then A 's eigenvalues are in \mathbb{R} .
- ▶ If $A = P$ is a projection matrix, then the idempotency of P determines λ : since $P^2 = P$, we get

$$\lambda^2 \vec{x} = P^2 \vec{x} = P \vec{x} = \lambda \vec{x} \implies \lambda \in \{0, 1\}.$$

The eigenvector(s) \vec{x} paired with

- ▶ $\lambda = 1$ have $P\vec{x}_1 = \vec{x}_1$, and so project to themselves under P ($\vec{x}_1 \in C(P) = C(P^t)$ since P is symmetric);
- ▶ $\lambda = 0$ have $P\vec{x}_0 = 0$, and so project to 0 under P ($\vec{x} \in N(P)$).

Roots of the eigenequation: orthogonal

- ▶ If $A = Q$ is orthogonal ($Q^{-1} = Q^t$), then

$$\begin{aligned} Q\vec{x} = \lambda\vec{x} &\implies (Q\vec{x})^t(Q\vec{x}) = (\lambda\vec{x})^t(\lambda\vec{x}) \\ &\implies \vec{x}^t Q^t Q \vec{x} = \vec{x}^t \vec{x} = \|\vec{x}\|^2 = \lambda^2 \|\vec{x}\|^2 \\ &\implies \lambda \in \{-1, 1\}. \end{aligned}$$

In particular, if $A = R$ is a reflection matrix, R is orthogonal and symmetric, implying R is an involution ($R^2 = I$). Thus,

$$R\vec{x} = \lambda\vec{x} \implies R^2\vec{x} = \lambda^2\vec{x} = \vec{x} \implies \lambda \in \{-1, 1\}.$$

- ▶ If A is **skew-symmetric** ($A^t = -A$), then A 's eigenvalues are “pure imaginary”: $\lambda = bi \in \mathbb{C}$ for $b \in \mathbb{R}$.

Example: 2x2 matrix

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.

$$\det(A) = 4 - 4 = 0$$

and A is symmetric, so both eigenvalues are real.

Thus, A is singular, and so one of A 's eigenvalues is 0.

Find the eigenvalues:

$$\det(A - \lambda I) = 0 \implies \det \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{pmatrix} = 0$$

$$\implies (1 - \lambda)(4 - \lambda) - 4 = 0$$

$$\implies \lambda^2 - 5\lambda = 0 \implies \lambda = 0, 5.$$

Example: 2x2 matrix

To find the eigenvectors, we need to solve

$$(A - \lambda I)\vec{x} = 0$$

for each λ .

The eigenvectors associated with $\lambda = 0$ are:

$$A\vec{x} = 0 \implies (\text{compute rref}(A)) \implies x_1 = -2x_2 \implies \vec{x}_n = x_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

The eigenvectors associated with $\lambda = 5$ are:

$$\begin{aligned} (A - 5I)\vec{x} = 0 &\implies \begin{pmatrix} 1-5 & 2 \\ 2 & 4-5 \end{pmatrix} \vec{x} = 0 \\ \implies (\text{compute rref}(A-5I)) &\implies x_1 = \frac{1}{2}x_2 \implies \vec{x}_n = x_2 \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}. \end{aligned}$$

Example: 3x3 matrix

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 1 & 5 & 4 \\ -6 & 3 & 7 \\ 0 & 0 & 2 \end{pmatrix}$.

Note: A is invertible:

$$\det(A) = 2[1(3) - 5(-6)] = 66 \neq 0.$$

Thus, A 's eigenvalues are all nonzero.

First, find the eigenvalues:

$$\begin{aligned} \det(A - \lambda I) = 0 &\implies \det \begin{pmatrix} 1 - \lambda & 5 & 4 \\ -6 & 3 - \lambda & 7 \\ 0 & 0 & 2 - \lambda \end{pmatrix} = 0 \\ &\implies (2 - \lambda)[(1 - \lambda)(3 - \lambda) + 30] = 0 \\ &\implies \lambda = 2 \text{ is an eigenvalue.} \end{aligned}$$

Example: 3x3 matrix, conjugate pair

There are two more: factoring out the $2 - \lambda$ term, we have

$$(1 - \lambda)(3 - \lambda) + 30 = 0 \implies 3 - 4\lambda + \lambda^2 + 30 = 0$$

$$\implies \lambda^2 - 4\lambda + 33 = 0$$

$$\implies \lambda = \frac{4 \pm \sqrt{16 - 4(1)(33)}}{2} = 2 \pm i\sqrt{29}.$$

These are the other two eigenvalues.

Note that they are a conjugate pair of complex numbers.

Example: 3x3 matrix, conjugate pair

Now for the eigenvectors: solving $(A - 2I)\vec{x} = 0$ for $\lambda = 2$ yields

$$\lambda = 2 : \quad \vec{x}_n = x_3 \begin{pmatrix} \frac{31}{29} \\ -\frac{17}{29} \\ 1 \end{pmatrix}, \text{ or } x_3 \begin{pmatrix} 31 \\ -17 \\ 29 \end{pmatrix}.$$

The conjugate pair of complex eigenvalues have complex conjugate pair eigenvectors:

$$\begin{aligned} \lambda = 2 + i\sqrt{29} : \quad \vec{x}_n &= x_2 \begin{pmatrix} \frac{1}{6}(1 - i\sqrt{29}) \\ 1 \\ 0 \end{pmatrix} \\ \lambda = 2 - i\sqrt{29} : \quad \vec{x}_n &= x_2 \begin{pmatrix} \frac{1}{6}(1 + i\sqrt{29}) \\ 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Example: eigenvalue multiplicity

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 5 & 2 \\ 0 & 5 \end{pmatrix}$.

This one is easy: A is triangular, so clearly the eigenvalues are $\lambda = 5$ with multiplicity 2.

What are the eigenvectors?

$$(A - 5I)\vec{x} = 0 \implies 2x_2 = 0 \implies \vec{x}_n = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Diagonalization of an invertible matrix

We know an $n \times n$ matrix A has, with multiplicity, n pairs of eigenvectors and eigenvalues. Label them

$$(\lambda_1, \vec{x}_1), (\lambda_2, \vec{x}_2), \dots, (\lambda_n, \vec{x}_n).$$

If the n eigenvectors of A are *linearly independent*, then we can construct an invertible **eigenvector matrix**

$$X = (\vec{x}_1 \ \vec{x}_2 \ \cdots \ \vec{x}_n)$$

of the n eigenvectors (which, naturally, are a basis of \mathbb{R}^n), and a corresponding diagonal **eigenvalue matrix**

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Diagonalization of an invertible matrix

Since $A\vec{x}_i = \lambda_i\vec{x}_i$ for all $i = 1, 2, \dots, n$, and X is invertible, we have

$$AX = X\Lambda \implies X^{-1}AX = \Lambda,$$

or

$$A = X\Lambda X^{-1},$$

called the **diagonalization** of A .

Note that the exponentiation property of eigenvalues is easy to see:

$$A^n = (X\Lambda X^{-1})^n = (X\Lambda X^{-1})(X\Lambda X^{-1}) \cdots (X\Lambda X^{-1}) = X\Lambda^n X^{-1},$$

confirming the notion that the eigenvalues of A^n are λ^n , with the same eigenvectors as A .

Finding eigenvalues is not a linear operation

Certain “shortcuts” to computing eigenvalues do not work:

If A and B are $n \times n$ matrices, with the sets of eigenvalues $\{\lambda_i\}_{i=1}^n$ for A and $\{\beta_i\}_{i=1}^n$ for B , then:

- ▶ $A + B$ does *not* have the eigenvalues $\lambda_i + \beta_i$ *unless* both A and B are diagonal.
- ▶ AB does *not* have the eigenvalues $\lambda_i \beta_i$.

Finding eigenvalues is not a linear operation

However,

Proposition

Assume A and B are diagonalizable, with diagonalizations

$$A = X\Lambda X^{-1}, \quad B = TKT^{-1}.$$

Then

$$X = T \iff AB = BA.$$

Invertibility vs Diagonalizability?

If any eigenvalue has multiplicity, then A is only diagonalizable if the **geometric multiplicity** of each eigenvalue's nullspace,

$$m_G(\lambda) = \dim(N(A - \lambda I)),$$

equals the **arithmetic multiplicity** of the eigenvalue,

$$m_A(\lambda) = \text{multiplicity of } \lambda \text{ in } \det(A - \lambda I) = 0.$$

Invertibility vs Diagonalizability?

Note that $m_A(\lambda) \geq m_G(\lambda)$ always holds.

If all n eigenvalues differ, then for each,

$$m_G(\lambda) = m_A(\lambda) = 1,$$

and A is diagonalizable (even if one of the eigenvalues is 0).

If the n eigenvectors of the matrix A are independent, then all eigenvalues differ, and A can be diagonalized.

Similar Matrices

If A is diagonalizable, in form

$$A = X\Lambda X^{-1},$$

there are multiple forms, depending on the scalings and orderings of the eigenvectors in X .

We will extend the conjugation paradigm to nondiagonalizable A .

Let M be an invertible matrix. Then we call

$$B = M^{-1}AM$$

similar to A .

If A is diagonalizable, then Λ is similar to A . In general,

Proposition

A and $M^{-1}AM$ have the same eigenvalues, and if \vec{x} is an eigenvector of A , then $M^{-1}\vec{x}$ is an eigenvector of $M^{-1}AM$.

Proof

$$\begin{aligned} A\vec{x} = \lambda\vec{x} &\implies (M^{-1}AM)(M^{-1}\vec{x}) = M^{-1}A\vec{x} \\ &= M^{-1}\lambda\vec{x} = \lambda(M^{-1}\vec{x}). \blacksquare \end{aligned}$$

Some similarity-invariant properties

What properties are *invariant* (do not change) under the transformation $A \mapsto M^{-1}AM$?

- ▶ eigenvalues
- ▶ trace
- ▶ determinant
- ▶ rank
- ▶ # of independent eigenvectors
- ▶ Jordan form

Some similarity-variant properties

What properties *do* change under this transformation?

- ▶ eigenvectors (from \vec{x} to $M^{-1}\vec{x}$)
- ▶ all four subspaces ($C(A)$ to $C(M^{-1}AM)$, etc.)
- ▶ singular values (in SVD)

Initial, long-term distributions

Once again examining powers of A , let $c \in \mathbb{R}^n$. Then the eigenvectors of A make a basis of \mathbb{R}^n , and so c can be represented via X in the form

$$c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \quad A = X\Lambda X^{-1} \implies Xc = \sum_{i=1}^n c_i \vec{x}_i.$$

Set $u_0 = Xc$ and define $u_k = A^k u_0$. We can quickly calculate u_k :

$$\begin{aligned} u_k &= A^k u_0 = A^k (c_1 \vec{x}_1 + c_2 \vec{x}_2 + \cdots + c_n \vec{x}_n) \\ &= c_1 \lambda_1^k \vec{x}_1 + c_2 \lambda_2^k \vec{x}_2 + \cdots + c_n \lambda_n^k \vec{x}_n. \end{aligned}$$

Initial, long-term distributions (Markov matrices)

Consider a particle Y_t at each time t moving along the states $\{1, 2, \dots, n\}$ with one-step probabilities

$$p_{ij} = P(Y_{t+1} = i \mid Y_t = j).$$

If A is a **Markov matrix**[§], whose entries are $a_{ij} = p_{ij}$, then A^k is the matrix of k -step probabilities

$$P(Y_{t+k} = i \mid Y_t = j).$$

[§]I am leaving out many important probability-based details here.

Initial, long-term distributions (Markov matrices)

If we define u_0 as the **initial distribution**[¶] of Y_0 , and define u_k iteratively, as

$$u_{k+1} = Au_k,$$

then the **long-term distribution** of Y is found by sending $k \rightarrow \infty$.

[¶]a probability mass function of the initial position Y_0 of the n possible states in the system

Initial, long-term distributions (Markov matrices)

A property of Markov matrices is that all of their eigenvalues have the property $|\lambda_i| \leq 1$, and, in particular, one of them *is* 1: we label this one $\lambda_1 = 1$.

Thus, the **long-term distribution** of Y is the eigenvector \vec{x}_1 , as we increase k in the k -step distribution u_k : recall, if $u_0 = Xc$, then

$$u_k = c_1 \lambda_1^k \vec{x}_1 + c_2 \lambda_2^k \vec{x}_2 + \cdots + c_n \lambda_n^k \vec{x}_n.$$

Initial, long-term distributions (Markov matrices)

As $k \rightarrow \infty$, $\lambda_i^k \rightarrow 0$ for $i > 1$ since $|\lambda_i| < 1$. But $\lambda_1 = 1$, so we get the limit

$$u_\infty = \lim_{k \rightarrow \infty} u_k = \lim_{k \rightarrow \infty} c_1 \lambda_1^k \vec{x}_1 + c_2 \lambda_2^k \vec{x}_2 + \cdots + c_n \lambda_n^k \vec{x}_n = c_1 \vec{x}_1.$$

The scaling of u_∞ that yields a probability vector is called π , the **long-term**, or **steady state**, distribution of Y .

The exponential of a square matrix

We can compute power series-like objects using exponentiation.

Recall, for any $r, t \in \mathbb{R}$, the natural exponential e^{rt} can be defined as a power series:

$$e^{rt} = \sum_{n=0}^{\infty} \frac{1}{n!} (rt)^n = 1 + rt + \frac{1}{2}(rt)^2 + \frac{1}{6}(rt)^3 + \dots$$

Likewise, if A is a square matrix and $t \in \mathbb{R}$, we will ignore the scalar convention of writing scalars on the left, and define the matrix exponential by

$$e^{At} = \sum_{n=0}^{\infty} \frac{1}{n!} (At)^n = I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \dots$$

The exponential of a diagonalizable matrix

If A is diagonalizable, then $A = X\Lambda X^{-1}$ for some eigenvalue matrix Λ and eigenvector matrix X .

Then the exponential is easy to compute:

$$\begin{aligned}e^{At} &= I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \dots \\&= I + X\Lambda tX^{-1} + \frac{1}{2}(X\Lambda tX^{-1})^2 + \frac{1}{6}(X\Lambda tX^{-1})^3 + \dots \\&= I + X(\Lambda t)X^{-1} + \frac{1}{2}X(\Lambda t)^2X^{-1} + \frac{1}{6}X(\Lambda t)^3X^{-1} + \dots \\&= Xe^{\Lambda t}X^{-1}.\end{aligned}$$

\therefore the eigenvalues of e^{At} are $e^{\Lambda t}$, with the same eigenvectors as A .

The complex exponential as a diagonalizable matrix

Consider the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Its powers have period 4, and e^{At} is orthogonal:

$$A^2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Thus,

$$e^{At} = \begin{pmatrix} 1 - \frac{1}{2}t^2 + \frac{1}{4}t^4 - \dots & t - \frac{1}{3}t^3 + \frac{1}{5}t^5 - \dots \\ -\left(t - \frac{1}{3}t^3 + \frac{1}{5}t^5 - \dots\right) & 1 - \frac{1}{2}t^2 + \frac{1}{4}t^4 - \dots \end{pmatrix}.$$

The complex exponential as a diagonalizable matrix

We can rewrite this form of e^{At} in power series terms:

$$e^{At} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \\ -\sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} \end{pmatrix} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix},$$

the two-dimensional rotation matrix.

The complex exponential as a diagonalizable matrix

The eigenvalue/eigenvector pairs of A are

$$(\lambda_1, \vec{x}_1) = \left(i, \begin{pmatrix} 1 \\ i \end{pmatrix} \right), \quad (\lambda_2, \vec{x}_2) = \left(-i, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right).$$

The eigenvalue/eigenvector pairs of e^{At} are

$$(e^{\lambda_1 t}, \vec{x}_1) = \left(e^{it}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right), \quad (e^{\lambda_2 t}, \vec{x}_2) = \left(e^{-it}, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right).$$

Note the connection to **Euler's formula**:

$$e^{it} = \cos(t) + i \sin(t).$$

Symmetric Matrices

Recall, a **symmetric matrix** A is a square matrix such that

$$A = A^t.$$

If A is symmetric and has independent columns, then A is diagonalizable, and so

$$A = X\Lambda X^{-1} \implies A^t = (X\Lambda X^{-1})^t = (X^{-1})^t \Lambda^t X^t = A.$$

Fact: $X^{-1} = X^t$, i.e. X is orthogonal.

Spectral Theorem

Theorem

Let A be a real-valued, symmetric matrix.

*Then its eigenvector matrix X is orthogonal,
i.e. $X^t X = I$, and can be chosen as unit length vectors.*

We will relabel this X as Q and write $A = Q \Lambda Q^t$.

Spectral Theorem: Supporting Propositions

Proposition

A real-valued, symmetric \implies A 's eigenvalues are real-valued.

Proof Suppose $A\vec{x} = \lambda\vec{x}$. Denote by $\bar{\lambda}$ the complex conjugate of λ : that is, if $\lambda = a + bi$, then $\bar{\lambda} = a - bi$.

We will prove that $\lambda = \bar{\lambda}$, which implies $b = 0$, and so $\lambda = a \in \mathbb{R}$.

Spectral Theorem: Supporting Propositions

Proof (continued)

Thus, $A = \bar{A}$ since A is real-valued, and so

$$A\vec{x} = \lambda\vec{x} \implies \overline{A\vec{x}} = \overline{\lambda\vec{x}} \implies A\vec{x} = \bar{\lambda}\vec{x}.$$

Transposing, we have

$$\begin{aligned}\vec{x}^t A^t &= \vec{x}^t \bar{\lambda} \implies \vec{x}^t A = \vec{x}^t \bar{\lambda} \\ &\implies \vec{x}^t A \vec{x} = \vec{x}^t \bar{\lambda} \vec{x} \\ &\implies \vec{x}^t \lambda \vec{x} = \vec{x}^t \bar{\lambda} \vec{x} \implies \lambda(\vec{x}^t \vec{x}) = \bar{\lambda}(\vec{x}^t \vec{x}).\end{aligned}$$

But $\vec{x}^t \vec{x}$ is a scalar. Thus, $\lambda = \bar{\lambda}$. ■

Spectral Theorem: Supporting Propositions

Proposition

Let A be a real-valued, symmetric matrix. Then, if \vec{x}_i and \vec{x}_j are eigenvectors corresponding to the eigenvalues λ_i and λ_j , $i \neq j$, and $\lambda_i \neq \lambda_j$, then $\vec{x}_i \perp \vec{x}_j$.

Proof We are given $A\vec{x}_i = \lambda_i\vec{x}_i$, $A\vec{x}_j = \lambda_j\vec{x}_j$, $\lambda_i \neq \lambda_j$, and $A = A^t$. Thus,

$$\begin{aligned}\lambda_i(\vec{x}_i^t \vec{x}_j) &= (\lambda_i \vec{x}_i^t) \vec{x}_j = (\lambda_i \vec{x}_i)^t \vec{x}_j \\ &= (A\vec{x}_i)^t \vec{x}_j = \vec{x}_i^t A^t \vec{x}_j = \vec{x}_i^t (A\vec{x}_j) \\ &= \vec{x}_i^t \lambda_j \vec{x}_j = \lambda_j(\vec{x}_i^t \vec{x}_j).\end{aligned}$$

Therefore, $\lambda_i = \lambda_j$ or $\vec{x}_i^t \vec{x}_j = 0$. As we are given $\lambda_i \neq \lambda_j$, it must then be that $\vec{x}_i^t \vec{x}_j = 0$, i.e. $\vec{x}_i \perp \vec{x}_j$. ■

Pivots vs Eigenvalues

We know that, if A is triangular, then the eigenvalues of A are the diagonal entries, i.e. the pivots, of A .

However, this is not true for a non-triangular matrix. For a matrix $A = LDU$, the pivots of A correspond to the diagonal entries of D , which are *not* the eigenvalues of A .

Pivots vs Eigenvalues

We have the following “compromise” result:

If d_1, d_2, \dots, d_n are the diagonal entries of D , i.e. the pivots of A , and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A , then

$$\det(A) = \prod_{i=1}^n \lambda_i = \prod_{i=1}^n d_i = \det(D).$$

Pivots vs Eigenvalues: Symmetric

If A is symmetric, then the signs of pivots and eigenvalues agree:

$$\# \text{ positive pivots of } A = A^t = \# \text{ positive eigenvalues of } A = A^t.$$

If all eigenvalues of A are positive (and hence all pivots are positive), then we call A is **positive definite** matrix.

(We will have another definition of this term shortly.)

Schur's Theorem

Theorem

If A is a square, complex-valued matrix, then A has a decomposition called **Schur form**,

$$A = QTQ^{-1},$$

where

- ▶ Q is a **unitary** matrix, i.e. $Q^{-1} = \overline{Q}^t$, its **conjugate transpose**, and
- ▶ T is upper triangular.

If A is a square, real-valued matrix, then the Schur form of A has Q an orthogonal matrix. If A is also symmetric, then T is diagonal.

Positive Definite Matrices

An $n \times n$ matrix A is called **symmetric positive definite (SPD)** if A is symmetric and satisfies all of the equivalent^{||} properties:

- ▶ $\vec{x}^t A \vec{x} > 0$ for any $\vec{x} \neq 0 \in \mathbb{R}^n$ (the “energy” definition)
- ▶ all n pivots $d_i > 0$
- ▶ all n eigenvalues $\lambda_i > 0$
- ▶ all n upper-left determinants (deleting successive bottom row/column pairs) > 0
- ▶ $A = R^t R$ for some R with independent columns.

If we relax the first (really, any) property to ≥ 0 , any A satisfying is called **symmetric positive semi-definite (SPSD)**.

^{||}satisfying one satisfies them all!

Positive Definite Matrices: Cholesky decomposition

If all the inequalities are flipped, we call A **symmetric negative (semi)definite**.

If A has positive and negative eigenvalues, we call A **indefinite**.

The last property,

► $A = R^t R$ for some R with independent columns,

is the **Cholesky decomposition** of A ; more to the point,

$$A = LDU \implies A = LDL^t = (L\sqrt{D})(L\sqrt{D})^t,$$

i.e. $R = (L\sqrt{D})^t$.

Positive Definite Matrices: two decompositions

If A is SPD, then there is a decomposition for pivots and another for eigenvalues:

$$A = LDL^t$$

gives the pivots in D (with triangular bookends), and

$$A = Q\Lambda Q^t$$

gives the eigenvalues in Λ (with orthogonal bookends).

Positive Definite Matrices: closed under addition, 2x2

For once, arithmetic intuition holds: if A and B are SPD, then so is $A + B$:

$$\vec{x}^t(A + B)\vec{x} = \vec{x}^t A \vec{x} + \vec{x}^t B \vec{x} > 0.$$

The converse, of course, is not necessarily true.

A 2x2 matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is SPD $\iff a > 0$ and $ac - b^2 > 0$.

Positive Semidefinite Matrices: Quadratic Forms, Ellipses

Note the energy definition of an SPSD matrix:

- ▶ $\vec{x}^t A \vec{x} \geq 0$ for any $\vec{x} \neq 0 \in \mathbb{R}^n$.

This kind of function, $f(\vec{x}) = \vec{x}^t A \vec{x}$, is called a **quadratic form**.

If A is SPSD, then the equation

$$\vec{x}^t A \vec{x} = 1$$

defines an n -dimensional ellipsoid in \mathbb{R}^n , centered at the origin.

In particular, the equation in \mathbb{R}^2 ,

$$\vec{x}^t A \vec{x} = 1,$$

$$\text{i.e. } \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2 = 1,$$

defines an ellipse.

Positive Semidefinite Matrices: principal axis theorem

This ellipse

$$\vec{x}^t A \vec{x} = ax^2 + 2bxy + cy^2 = 1$$

is tilted and stretched from the unit circle

$$\vec{x}^t I \vec{x} = x^2 + y^2 = 1$$

in that its eigenvectors point in the directions of the minor and major axes, and the eigenvalues determine the lengths of the axes: this is stated in the **Principal Axis Theorem**.

Positive Semidefinite Matrices: principal axis theorem

Factoring the matrix $A = Q\Lambda Q^t$ yields a way to “standardize” the ellipse equation in the following way:

$$\vec{x}^t A \vec{x} = (\vec{x}^t Q) \Lambda (Q^t \vec{x}) = 1.$$

Writing $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = Q^t \vec{x}$, the change into “ellipse coordinates”, we get

$$\vec{x}^t A \vec{x} = C^t \Lambda C = \lambda_1 c_1^2 + \lambda_2 c_2^2 = 1.$$

This implies that, if \vec{x}_1 and \vec{x}_2 the eigenvectors of A , the axes are in the directions of \vec{x}_1 and \vec{x}_2 , with axis half-lengths $\frac{1}{\sqrt{\lambda_1}}$ and $\frac{1}{\sqrt{\lambda_2}}$.