CLAUSSEN

Eigenvectors & Eigenvalues (Def.)

Let F be a field and F^n a n-dim vector space with linear mappings (in the vector space) represented by matrices $M \in F^{n \times n}$.

Then $\lambda_i \in F$ is called an eigenvalue of M to the eigenvector $\vec{x}_i \in F^n$, $\vec{x}_i \neq \vec{0}$ if $M\vec{x}_i = \lambda_i \vec{x}_i$

We cannot expect more than n $(\lambda_i, \vec{x}_i)$ pairs (solutions) in a n-dim. vector space.

Let's inspect some basic examples!

$$M = \mathbb{1}_3. \text{ Then } \det(M - \lambda_i \mathbb{1}) = \det\left(\begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{pmatrix} - \begin{pmatrix} \lambda_i & & 0 \\ & \lambda_i & \\ 0 & & \lambda_i \end{pmatrix}\right)$$

$$= \det\begin{pmatrix} (1-\lambda_i) & & 0 \\ & (1-\lambda_i) & \\ 0 & & (1-\lambda_i) \end{pmatrix}$$

$$= (1-\lambda_i)(1-\lambda_i)(1-\lambda_i) \overset{!}{=} 0$$

which has the unique (triple-degenerate!) solution $\lambda_i = 1$.

As the solution space for the $\vec{x}_i$ is degenerate, we would have some freedom to choose them, but

$(\lambda_1 = 1, \vec{x}_1 = \vec{e}_1), (\lambda_2 = 1, \vec{x}_2 = \vec{e}_2), (\lambda_3 = 1, \vec{x}_3 = \vec{e}_3)$ would do.

All eigenvalues of a unit matrix are equal to 1.

More specifically, for any diagonal matrix

$\begin{pmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & & \ddots \; d_n \end{pmatrix}$ we have the eigenvalues $\lambda_i = d_i$ (with eigenvectors $\vec{e}_i$)

VERY important to remember & recognize

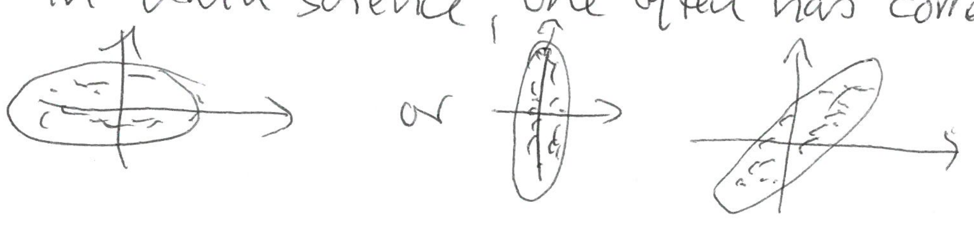
Corrolary: All eigenvalues of a zero matrix are zero ☺

CLAUSSEN

Eigenvalues, Eigenvectors (and Diagonalization)

A matrix M (linear mapping) provides some form of coordinate transformation from the basis $(\vec{e}_1, \vec{e}_2, \dots \vec{e}_n)$ to the span of $M\vec{e}_1, M\vec{e}_2, \dots M\vec{e}_n$ ((which might not always be a basis, but we come to that later))

- For instance, $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ "rotates" the unit vector basis (along the $(1,1,1)$ axis) such that $x \to y \to z \to x$ (cyclic permutation of vector components)
- For instance $\begin{pmatrix} 2.54 & 0 & 0 \\ 0 & 2.54 & 0 \\ 0 & 0 & 2.54 \end{pmatrix}$ rescales axes between cm and inch, hence volume is "increased in number" by factor $(2.54)^3$ (of course, the volume stays the same but $(1\text{ in})^3 = (2.54\text{ cm})^3$)
- One might wish to scale units differently for the coordinates, like km and miles in x and y, but m or yard for z (land maps: mountain height vs horizontal distances)
  $\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}$ ← we get a matrix like this
- In data science, one often has correlated data to approximate:



but we might get the same data in different coordinate systems.

Can we "extract" some properties from a given matrix that describes "what it does"?

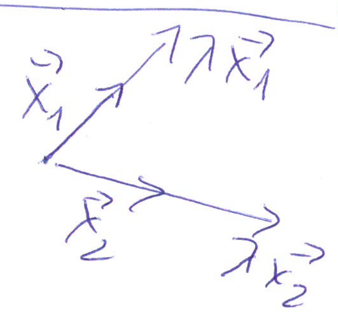
CLAUSSEN

Eigenvalues (& vectors) – motivation, cont'd

If for some vectors $\vec{x}_i$ (from basis, span, ... tbc) the matrix M only stretches the vector $\vec{x}_i$ by some factor λ_i, this means that

$$\boxed{M\vec{x}_i = \lambda_i \vec{x}_i \quad \text{Eigenvalue \& vector equation}}$$

We need to find out what that means.



(First, we should exclude the zero vector $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \end{pmatrix}$ from this discussion, as $M\vec{x} = \lambda\vec{x}$ would be always valid for $\vec{x} = \vec{0}$, whatever M and λ.)

We can use $M\vec{x}_i = \lambda_i \vec{x}_i = \lambda_i \mathbb{1} \vec{x}_i \Leftrightarrow \underline{(M - \lambda_i \mathbb{1})\vec{x}_i = \vec{0}}$.

(Excluding cases which have no solution, resulting in $0=1$ after Gaussian elimination)

Recalling from linear systems of equations

If $A\vec{x} = \vec{0}$ "homogeneous system", we have 2 cases:
- an unique solution $\vec{x} = \vec{0}$
- if $\vec{x} \neq \vec{0}$ and $A\vec{x} = \vec{0}$, we have for any $\lambda \in F$ $A(\lambda\vec{x}) = \lambda(A\vec{x}) = \vec{0}$ hence $\forall_{\lambda \in F}$ $(\lambda\vec{x})$ is also a solution. If $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, this would be an infinite solution space (geometrically, a line)

The homogeneous system has nontrivial (other than zero) solutions $\Leftrightarrow \det(A) = 0$

Hence, $\vec{x}_i$ is a nontrivial ($\vec{x}_i \neq \vec{0}$) solution if and only if

$\det(M - \lambda_i \mathbb{1}) = 0$ for some $\lambda_i \in F$.

Through this equivalence, we can use this for definition of eigenvectors + eigenvalues, and to calculate the λ_i.

CLAUSSEN

Some more examples (Eigenvalues & Eigenvectors)

Let's inspect a general symmetric 2x2 matrix

$$M = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \text{ with } a, b, d \in \mathbb{R}.$$

We will rarely refer to $\mathbb{C}$, for a while...

Then our eigenvalue equation reads

$$0 \overset{!}{=} \det(M - \lambda \mathbb{1}) = \det\left(\begin{pmatrix} a & b \\ b & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right)$$

$$= \det \begin{pmatrix} a-\lambda & b \\ b & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - b^2$$

$$= \lambda^2 - (a+d)\lambda + (ad - b^2)$$

This is a <u>quadratic equation</u> <u>for λ</u> with the general

COOL! You can solve that!

Solution:

$$\lambda_{1,2} = \frac{a+d}{2} \pm \sqrt{\left(\frac{a+d}{2}\right)^2 - (ad - b^2)}$$

$$= \frac{a+d}{2} \pm \frac{1}{2}\sqrt{a^2 + 2ad + d^2 - 4ad + 4b^2}$$

$$= \frac{a+d}{2} \pm \frac{1}{2}\sqrt{a^2 - 2ad + d^2 + 4b^2}$$

$$= \frac{a+d}{2} \pm \frac{1}{2}\sqrt{\underbrace{\underbrace{(a-d)^2}_{\geq 0} + \underbrace{4b^2}_{\geq 0}}_{\geq 0}}$$

$\Rightarrow$ <u>EVERY</u> symmetric real 2x2 matrix has two real-valued eigenvalues ((they may be zero)).

<u>In general</u>, any <u>symmetric</u> real $n \times n$ matrix has n <u>real-valued</u> eigenvalues (some may be identical $\equiv$ "degenerate")

CLAUSSEN 2

Eigenvalues of a non-symmetric matrix?

in 2d, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $b \neq c$. OK, inspect $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ...

$$0 \stackrel{!}{=} \det(M - \lambda_i \mathbb{1}) = \det\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right)$$

$$= \det\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$$

Hence, we have to find solutions of $\boxed{\lambda^2 = -1}$

- Knowing the square of any $\lambda \in \mathbb{R}$ is nonnegative, $\lambda^2 \geq 0$ there are no $\lambda \in \mathbb{R}$ solving $\lambda^2 = -1$.
- This $\boxed{\text{SHOULD}}$ remind (!) us of the equation $\boxed{\lambda^2 = 2}$ which has no rational ($\in \mathbb{Q}$) solution, but allowing for irrational numbers $\lambda_1 = +\sqrt{2}$ and $\lambda_2 = -\sqrt{2}$, both $\in \mathbb{R}$, solve this polynomial equation (while the coefficients are $\in \mathbb{Q}$)
- Similarily, if we extend the field $\mathbb{R}$ by imaginary ((non $\equiv$ real, literally!)) numbers iy with $y \in \mathbb{R}$,

We can solve $\boxed{\text{ANY}}$ quadratic equation by assuming complex numbers $z = x + iy$ with $x \in \mathbb{R}$, $y \in \mathbb{R}$ and $i^2 = -1$, $i = \sqrt{-1}$

Oh! I may have to look up the complex numbers 1-pager? LINK →

(After recap or explore into $\mathbb{C}$)

the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ has eigenvalues $\lambda_1 = i = \sqrt{-1}$, $\lambda_2 = -i = -\sqrt{-1}$

<u>Exercise</u>: Find the two eigenvectors to λ_1 and λ_2, you can treat i just as a scalar $\in \mathbb{C}$.

Complex numbers in a nutshell

CLAUSSEN

For any complex number $z \in \mathbb{C}$, there are $x, y \in \mathbb{R}$ such $\boxed{z = x + iy}$

Here, $i (\notin \mathbb{R})$ is the imaginary unit and solution of $z^2 = -1$.

↳ hence, $\boxed{i = \sqrt{-1}}$

For any $z \in \mathbb{C}$, $(z = x+iy)$ we can also define the complex conjugate $\underbrace{z^* = \bar{z}}_{\text{both notations common}} = x - iy$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \quad \underset{\substack{x = (z+z^*)/2 \\ y = (z-z^*)/2i}}{\overset{z = x+iy}{\rightleftarrows}} \quad z \in \mathbb{C}$$

Hence, $\mathbb{C}$ inherits the vector space properties of addition and multiplication with a scalar ($\in \mathbb{R}$).

We also want to multiply by complex numbers!

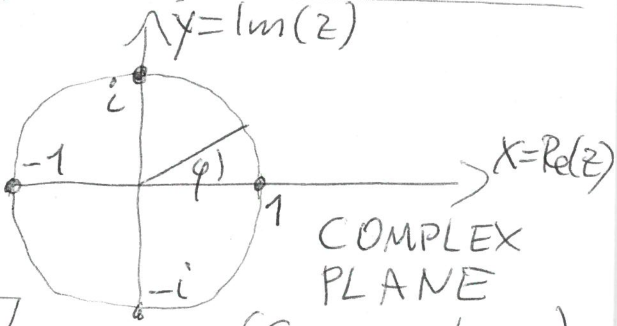
$$(x+iy)(a+ib) = xa + iya + ixb + \underbrace{(i)^2}_{=-1} yb$$
$$= (xa - yb) + i(ya + xb)$$

Using polar coordinates →

$$x = r \cos\varphi$$
$$y = r \sin\varphi$$

$$\rightarrow x + iy = r(\cos\varphi + i\sin\varphi) =: re^{i\varphi}$$

$\boxed{\text{EULER relation}: \exp(i\varphi) = \cos\varphi + i\sin\varphi}$



COMPLEX PLANE (Gauss plane)

We need to define the norm! Require $\|z\|^2 = \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 = x^2 + y^2$

$$= x^2 - (iy)^2$$
$$= (x - iy)(x + iy) = z^* z$$

Hence, norm in $\mathbb{C}$ (absolute value): $|z| := \sqrt{z^* z}$

This forces us (in $\mathbb{C}$) to "replace" transposition by ("transpose and conjugate complex" =: adjoint) for the scalar (inner) product: $\langle \vec{z}_1, \vec{z}_2 \rangle = (\vec{z}_1^{\,T})^* \vec{z}_2 = \vec{z}_1^{\,\dagger} \vec{z}_2$

Also, orthogonality $R^T R = \mathbb{1}$ (in $\mathbb{R}$) becomes unitarity $U^\dagger U = \mathbb{1}$ (in $\mathbb{C}$).

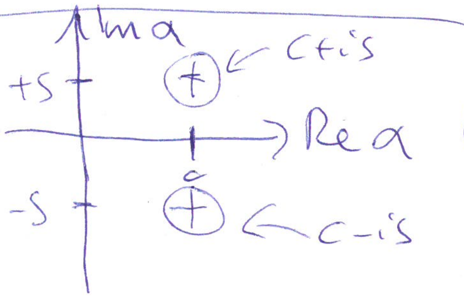
Eigenvalues (& vectors) of a rotation matrix?

$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$ with $c^2+s^2=1$ $(\cos^2\varphi+\sin^2\varphi=1)$

$c=\cos\varphi$ $s=\sin\varphi$, $\varphi\in\mathbb{R}$

$0 \overset{!}{=} \det\begin{vmatrix} c-\alpha & -s \\ s & c-\alpha \end{vmatrix} = (c-\alpha)^2+s^2 = (\alpha-c)^2-(i\cdot s)^2$ $\boxed{i=\sqrt{-1}\in\mathbb{C}}$

hence $\alpha - c = \pm is \Leftrightarrow \alpha_{1,2} = c \pm is$



or $\alpha_{1,2} = \cos\varphi \pm i\sin\varphi \overset{\text{Euler formula}}{=} \exp(\pm i\varphi)$

we have a pair of conjugate complex eigenvalues, $c+is$ and $c-is$.

Eigenvectors? To solve: $(R-\alpha\mathbb{1})\vec{x}=\vec{0}$

$\alpha_1 = c+is$ $\quad \begin{pmatrix} c-(c+is) & -s \\ s & c-(c+is) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -is & -s \\ s & -is \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

or $-xis - sy = 0$

$xs - isy = 0$ (lin. dep.)

$x=1 \to y=-ix$ EV: $\begin{pmatrix} 1 \\ -i \end{pmatrix}$

$\alpha_2 = c-is$ $\quad \begin{pmatrix} c-(c-is) & -s \\ s & c-(c-is) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

$= \begin{pmatrix} +is & -s \\ s & +is \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xis - ys \\ xs + isy \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ or $y = ix$ EV: $\begin{pmatrix} 1 \\ i \end{pmatrix}$

Observations:

- Eigenvalues of a real, but not-symmetric matrix are complex ($\in\mathbb{C}$), but come in complex conjugate pairs (like $c\pm is$)
- Rotation matrices have complex eigenvalues (rotation angle via Euler formula ☺)
- Eigenvectors: Rotation means that $Mx = \lambda x$ may not have real-valued solution.

Here, EV/EV pairs are $\left(c+is, \begin{pmatrix} 1 \\ -i \end{pmatrix}\right)$ and $\left((c-is), \begin{pmatrix} 1 \\ i \end{pmatrix}\right)$.

Eigenvectors & coordinate transform

CLAUSSEN

Consider again our rotation matrix, $R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$

with eigenvalues & vectors

$\alpha_1 = c + is \quad v_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$

$\alpha_2 = c - is \quad v_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$

$c := \cos\varphi$
$s := \sin\varphi, \ \varphi \in \mathbb{R}$

using: Euler relation

$$\boxed{\exp(i\varphi) = e^{i\varphi} = \cos\varphi + i\sin\varphi}$$

$\downarrow$

$e^{i\varphi} + e^{-i\varphi} = 2\cos\varphi$

$e^{i\varphi} + e^{-i\varphi} = 2i\sin\varphi$

Define $D := \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$, $T = \begin{pmatrix} | & | \\ \vec{v}_1 & \vec{v}_2 \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$

Then $T^{-1} = \frac{1}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$ (Inverse)
(Verify $TT^{-1} = \mathbb{1}$!)

Let us now calculate

$$\underline{TDT^{-1}} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \cdot \frac{1}{2}$$

$$= \frac{1}{2}\begin{pmatrix} e^{i\varphi} & e^{-i\varphi} \\ -ie^{i\varphi} & ie^{-i\varphi} \end{pmatrix}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} e^{i\varphi} + e^{-i\varphi} & i(e^{i\varphi} - e^{-i\varphi}) \\ -i(e^{i\varphi} - e^{-i\varphi}) & e^{i\varphi} + e^{-i\varphi} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2\cos\varphi & -2\sin\varphi \\ 2\sin\varphi & 2\cos\varphi \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \underline{R}$$

That means, we have expressed (or decomposed) the matrix R by a diagonal matrix (containing the eigenvalues), the matrix formed by (column) eigenvectors, and the inverse of this matrix.

A matrix having such a decomposition is called <u>diagonalizable</u>.

(obviously the inverse of T must exist! Eigenvectors need to be linearly independent, hence form a basis)

The matrix T provides the coordinate transform from the standard basis to the "eigensystem" (the basis comprised by the basis vectors.

CLAUSSEN

Diagonalization (cont'd)

Why does this work? Let $\alpha_1 \dots \alpha_n$ be EV to $\vec{t}_1 \dots \vec{t}_n$, $A\vec{t}_i = \lambda_i \vec{t}_i$

Then,

$$\underline{AT} = \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix}\begin{pmatrix} | & | & & | \\ \vec{t}_1 & \vec{t}_2 & \cdots & \vec{t}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ (A\vec{t}_1) & (A\vec{t}_2) & \cdots & (A\vec{t}_n) \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & | & & | \\ (\alpha_1 \vec{t}_1) & (\alpha_2 \vec{t}_2) & \cdots & (\alpha_n \vec{t}_n) \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & | & & | \\ \vec{t}_1 & \vec{t}_2 & \cdots & \vec{t}_n \\ | & | & & | \end{pmatrix}\begin{pmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_n \end{pmatrix} = \underline{TD}$$

this is just writing all n eigenvector equations in one matrix equation, $AT = TD$.

If we multiply by the inverse of T (provided it exists)

$$T^{-1}AT = \underbrace{(T^{-1}T)}_{=\mathbb{1}} D = D \qquad A = A\underbrace{TT^{-1}}_{=\mathbb{1}} = TDT^{-1}$$

(note that the $\vec{t}_i$ need not to be normalized, but if we do so calculating T^{-1} becomes easier)

This has many useful applications!!

$$A^2 = (TDT^{-1})(TDT^{-1}) = TD(T^{-1}T)DT^{-1} = TD^2T^{-1}$$

$$= T\begin{pmatrix} \alpha_1^2 & & & \\ & \alpha_2^2 & & \\ & & \ddots & \\ 0 & & & \alpha_n^2 \end{pmatrix}T^{-1}$$

$$A^k = T\begin{pmatrix} \alpha_1^k & & & 0 \\ & \alpha_2^k & & \\ & & \ddots & \\ 0 & & & \alpha_n^k \end{pmatrix}T^{-1}$$

even

$$\exp(A) = T\begin{pmatrix} \exp(\alpha_1) & & & 0 \\ & \exp(\alpha_2) & & \\ & & \ddots & \\ 0 & & & \exp(\alpha_n) \end{pmatrix}T^{-1}$$

((for the proof, you may use the Taylor expansion of $\exp()$:

$$\exp(x) = \lim_{n\to\infty} \sum_{\ell=0}^{n} \frac{x^\ell}{\ell!}$$
))