MATHEMATICS FOR SCIENCE STUDENTS

An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

$$a^b = e^{b\log(a)}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$T\left(\alpha\vec{u} + \beta\vec{v}\right) = \alpha T\left(\vec{u}\right) + \beta T\left(\vec{v}\right)$$

$$A = Q\Lambda Q^{-1}$$

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t$$

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

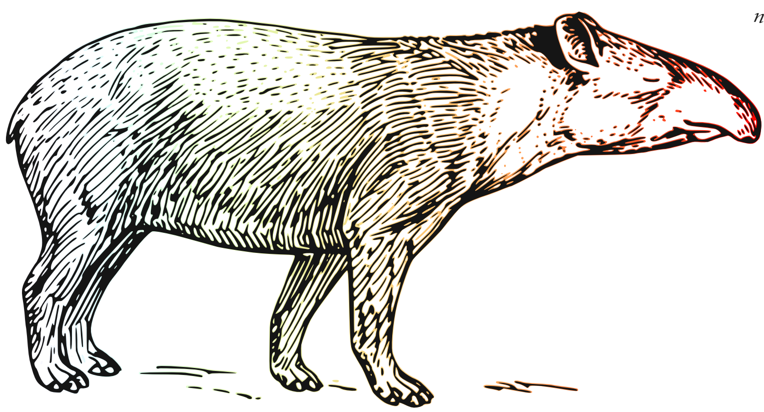
$$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$$

$$\vec{v} = \sum_{i=1}^{n} \alpha_i \hat{e}_i$$

$$e^{\pi i} + 1 = 0$$

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$



PUBLISHED IN THE WILD

HERE BE TABLE

CHAPTER 0

INTRODUCTION

In this chapter we introduce key concepts that will be used in later chapters. For this reason, unlike other chapters it contains many statements, sometimes given without thorough explanations or reasoning. While all of these statements are grounded in deep ideas and can be formulated in a rigorous manner, it is advised to first get an intuitive understanding of the ideas before diving into their more formal construction.

Note 0.1 In case you are already familiar with the topics

It is recommended for readers who are familiar with the topics to at least gloss over this chapter and make sure they know and understand all the concepts presented here.

!

0.1 EIGENVECTORS AND EIGENVALUES

0.1.1 Definition

Some linear transformations have special directions which only scale by the application of the transformation and are not mapped to different directions. Take for example the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, which scales space by 2 in the y-direction. All vectors pointing in the y-direction get scaled by T (namely by a factor of 2) and still point in the y-direction after the application of T. All vectors pointing in the x-direction do not change at all (i.e. they are ”scaled” by a factor of 1), and of course still point in the x-direction after the application of T. Any other vector - i.e. those that have both components different than zero - change their direction after the application of T (see Figure 0.1).

We call such vectors the **eigenvectors** of the transformation. The amount by which the are scaled is then their respective **eigenvalues**.

Note 0.2 Pronounciation

The word *eigen* is a German word meaning ”own” (as in ”own rules”), or ”self” (as in self-made). We will see how this meaning fits the concept later in the section. The *ei* part it pronounced the same as the English word ”eye”, and the *g* is pronounced like the *g* in the English word dog (i.e. unlike the *g* in *generation*).

!

Example 0.1 Eigenvectors and eigenvalues

Text here

In matrix form, a vector $\vec{v}$ is an eigenvector of a transformation represented by the matrix A, if

$$A\vec{v} = \lambda\vec{v}, \tag{0.1.1}$$

where $\lambda \in \mathbb{R}$, $\lambda \neq 0$. This kind of equation is typically called an **eigenvector equation**. When there are several eigenvectors for a transformation, each with its distinct eigenvalue, we simply add indeces to all relevant parts:

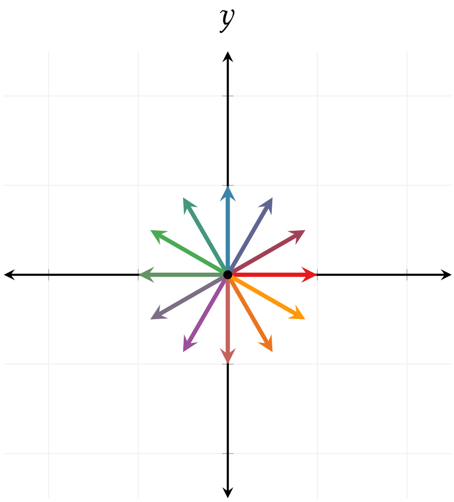
$$A\vec{v}_i = \lambda_i\vec{v}_i, \tag{0.1.2}$$

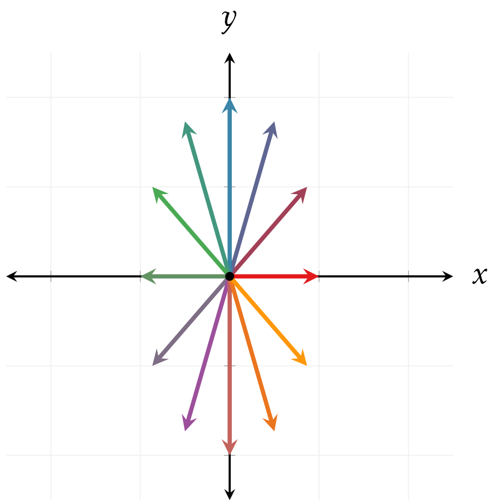
where again $\lambda_i \in \mathbb{R}$ and $\lambda_i \neq 0$.

Note 0.3 The zero-vector

Although technically the zero vector is indeed ”scaled” by any linear transformation - by infinitely many scalars - it is **not** considered an eigenvector, exactly because of the fact it has no unique eigenvalue.

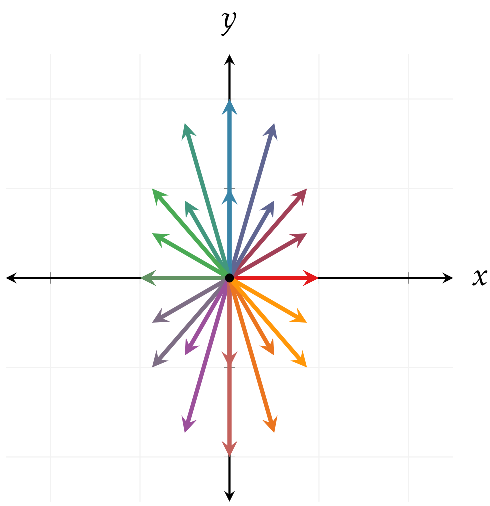
!





(a) Some vectors.

(b) Same vectors after the application of T.



(c) The vectors before and after the application of T layered on top of eachother.

Figure 0.1 Some vectors before and after application of the y-scaling transformation T. Note how only the vectors pointing in the direction of the x- and y-axes stay in the same direction, while all the other vectors change their directions.

Before continuing to explore some more examples of eigenvectors, there are two properties[1] of eigenvectors that are important to mention. Given a linear transformation T,

- A scale of any eigenvector $\vec{v}$ of T is also an eigenvector of T, with the same eigenvalue.

Proof 0.1 Eigenvector scale

Let $T:\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation represented by the square matrix A, with eigenvector $\vec{v}$ and its respective eigenvalue λ. Then

$$A\vec{v} = \lambda\vec{v}.$$

Replacing $\vec{v}$ with a scale of itself, i.e. $\vec{u} = \alpha\vec{v}$, then applying A to $\vec{u}$ gives us

$$A\vec{u} \overset{(1)}{=} A\left(\alpha\vec{v}\right) \overset{(2)}{=} \alpha A\vec{v} \overset{(3)}{=} \alpha\lambda\vec{v} \overset{(4)}{=} \lambda\alpha\vec{v} \overset{(5)}{=} \lambda\vec{u}.$$

where

(1) Substitution of $\vec{u}$ by its definition $\vec{u} = \alpha\vec{v}$.

(2) Due to the linearity of A we can bring α out of the product.

(3) Resulting due to $\vec{v}$ being an eigenvector of A.

(4) The product of real numbers is commutative.

(5) Substituting back $\alpha\vec{v} = \vec{u}$.

Therefore, $\vec{u}$ is also an eigenvector of A (and thus T) with the same eigenvalue λ as $\vec{v}$.

QED

Since a linear transformation never has just a single eigenvector but infinitely many (i.e. its entire span), we will refer from now on the **families** of eigenvectors, all pointing in the same direction, represented by a single vector (usually a unit vector, but not necesseraly).

- The linear combination of two eigenvectors of T is **not neccessarily an eigenvector of T**! For example, consider the above transformation which scales all vectors by 2 in the y-direction: as we saw, any vector in the x-direction is an eigenvector of the transformation, and so does any vector in the y-direction. Specifically, the vectors $\vec{a} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{b} = \begin{bmatrix}0\\1\end{bmatrix}$ are two separate eigenvectors of the transformation (with eigenvalues 1 and 2, respectively), however the vector

$$\vec{c} = \vec{a} + \vec{b} = \begin{bmatrix}1\\1\end{bmatrix}$$

is **NOT** an eigenvector of the transformation, since

$$\begin{bmatrix}1\\1\end{bmatrix} \overset{T}{\mapsto} \begin{bmatrix}1\\2\end{bmatrix} \neq 2\begin{bmatrix}1\\1\end{bmatrix}.$$

[1] actually one property and one non-property

0.1.2 Some examples

Example 0.2 Eigenvectors and eigenvalues of common linear transformations

Let's find the eigenvectors and their respective eigenvalues for the some basic linear transformations in 2-dimensions.

- **Skew (shear) in the x-direction**: any vector pointing in the x-direction is unchanged (i.e. scaled by a factor of 1), while any other vector changes its direction. Thus, the transformation has only a single family of eigenvectors, which we will represent by the vector $\hat{x}$, and their respective eigenvalue $\lambda = 1$.

- **Rotation around the origin**: no vector (except the zero-vector) is scaled by rotation, and therefore rotations have no eigenvectors. However, in 3-dimensions any rotation transformation does have an eigenvector, with eigenvalue $\lambda = 1$: it is of course the family of vectors pointing in the direction of the axis of rotation. We will discuss this further later in the section.

- **Reflection across the x-axis**: in this case, any vector pointing in the x-axis is not being changed - an eigenvector with eigenvalue $\lambda = 1$, and any vector pointing in the y-direction is reflected verically, i.e.

$$\begin{bmatrix} 0 \\ \beta \end{bmatrix} \overset{T}{\mapsto} \begin{bmatrix} 0 \\ -\beta \end{bmatrix} = -\begin{bmatrix} 0 \\ \beta \end{bmatrix},$$

  and is therefore an eigenvector with eigenvalue $\lambda = -1$. The transformation has no other eigenvectors.

- **Reflection across a line going through the origin**: much like the previous case, any vector lying on the reflection line will not change ($\lambda = 1$), and any vector pointing in an orthogonal direction to the line will flip ($\lambda = -1$). No other eigenvectors exist for this transformation.

0.1.3 Calulating eigenvectors

Calculating the eigenvectors of a given transformation, and their respective eigenvalues, is a rather easy procedure to perform once one has the transformation in matrix form. However, in order to understand *why* the procedure works it is useful to derive it first, which is what we'll do now.

We take the eigenvector equation (Equation 0.1.1) and rearrange it slightly:

$$A\vec{v} - \lambda\vec{v} = \vec{0}. \tag{0.1.3}$$

We can then group together all parts which include $\vec{v}$, but we must be careful: A is a matrix while λ is a scalar. This means that the term $A - \lambda$ has no meaning, since we haven't defined how to add or subtract matrices and real numbers. We therefore change Equation 0.1.3 a bit without changing its validity, by replacing $\lambda\vec{v}$ with $\lambda I\vec{v}$: i.e. instead of scaling $\vec{v}$ by a scalar, we scale it using the matrix λI, which yields the same result:

$$A\vec{v} - \lambda I\vec{v} = \vec{0}. \tag{0.1.4}$$

Note 0.4 That one weird trick

If you are not convinced the above trick works, consider the following:

$$\lambda = 3, \vec{v} = \begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix}.$$

Then the direct scaling of $\vec{v}$ by λ is

$$\lambda\vec{v} = \begin{bmatrix} 3 \\ 15 \\ -21 \end{bmatrix},$$

and scaling it using λI yields

$$\lambda I\vec{v} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix} = \begin{bmatrix} 3 \\ 15 \\ -7 \end{bmatrix},$$

i.e. we get exactly the same result.

!

Grouping together the parts with $\vec{v}$ gives:

$$(A - \lambda I)\vec{v} = \vec{0}. \tag{0.1.5}$$

Note that $A - \lambda I$ is a matrix, with the following form:

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}. \end{aligned} \tag{0.1.6}$$

Equation 0.1.5 tells us that $A-\lambda I$ sends the vector $\vec{v}$ to $\vec{0}$, and therefore $\vec{v}$ is in the kernel of $(A-\lambda I)$. Sinve by the definition of an eigenvector $\vec{v} \neq \vec{0}$, this means that the kernel of $A-\lambda I$ has more then just the zero vector, and thus

$$|A-\lambda I| = 0. \tag{0.1.7}$$

(this is derived from **??** and **??**)

Therefore, if we solve Equation 0.1.7 for λ, we will get all the values of λ for which the eigenvector equation holds, and in turn we get all the eigenvalues λ_i of the linear transformation represented by A. We can then substitute each λ_i into the eigenvector equation and find its respective eigenvector family.

Example 0.3 Eigenvectors and eigenvalues of a 2×2 matrix

The matrix representing the y-scaling transformation discussed in the beginning of this section is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Therefore, in order to find the eigenvectors of A we solve the equation

$$0 = |A-\lambda I| = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda).$$

In this case there are two solutions: $\lambda_1 = 1$ and $\lambda_2 = 2$. To find their corresponding eigenvectors, we subtitute them into the eigenvectors equation. First λ_1:

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

which corresponds to the following system of equations:

$$\begin{cases} 1\cdot x + 0\cdot y = x, \\ 0\cdot x + 2\cdot y = y, \end{cases}$$

for which the solution is $x \in \mathbb{R}$ and $y = 0$ - i.e. any vector pointing in the x-direction. Now for $\lambda_2 = 2$:

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

i.e. the following system of equations:

$$\begin{cases} 1\cdot x + 0\cdot y = 2x, \\ 0\cdot x + 2\cdot y = 2y. \end{cases}$$

The solution is of course $x = 0$ and $y \in \mathbb{R}$, meaning any vector pointing in the y-direction.

Example 0.4 Eigenvectors and eigenvalues of a 3×3 matrix

Let us now calculate the eigenvectors and their respective eigenvalues for the following 3×3 matrix:

$$A = \begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}.$$

We start by calculating the determinant $|A-\lambda I|$:

$$\begin{aligned} |A-\lambda I| &= \begin{vmatrix} 5-\lambda & -6 & -3 \\ 0 & -1-\lambda & 0 \\ 0 & 2 & 2-\lambda \end{vmatrix} \\ &= (5-\lambda)\begin{vmatrix} -1-\lambda & 0 \\ 2 & 2-\lambda \end{vmatrix} - (-6)\cancel{\begin{vmatrix} 0 & 0 \\ 0 & 2-\lambda \end{vmatrix}} + 3(-3)\cancel{\begin{vmatrix} 0 & -1-\lambda \\ 0 & 2 \end{vmatrix}} \\ &= (5-\lambda)(-1-\lambda)(2-\lambda) \end{aligned}$$

The solutions of $|A-\lambda I| = 0$ are therefore

$$\lambda_1 = 5,\ \lambda_2 = -1,\ \lambda_3 = 2.$$

- $\boldsymbol{\lambda_1 = 5}$: solving $A\vec{v} = 5\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 5\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5x \\ 5y \\ 5z \end{bmatrix},$$

  i.e.

$$\begin{cases} 5x - 6y - 3z & = 5x, \\ 0x - 1y + 0z & = 5y, \\ 0x + 2y + 2z & = 5z, \end{cases}$$

  for which the solution is

$$x \in \mathbb{R},\ y = 0,\ z = 0.$$

  The first family of eigenvectors are the vectors pointing in the x-direction, and their respective eigenvalue is $\lambda = 5$. We can verify this:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x - 6\cdot 0 - 3\cdot 0 \\ 0x - 1\cdot 0 + 0\cdot 0 \\ 0x + 2\cdot 0 + 2\cdot 0 \end{bmatrix} = \begin{bmatrix} 5x \\ 0 \\ 0 \end{bmatrix} = 5\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}.$$

- $\boldsymbol{\lambda_2 = -1}$: solving $A\vec{v} = -\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = -\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix},$$

i.e.

$$\begin{cases} 5x - 6y - 3z &= -x, \\ 0x - 1y + 0z &= -y, \\ 0x + 2y + 2z &= -z, \end{cases}$$

for which the solution is

$$x = \frac{2}{3}y,\ y \in \mathbb{R},\ z = -x.$$

Verifying the solution using the representative vector $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}$:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \cdot 2 - 6 \cdot 3 + 3 \cdot 2 \\ 0 \cdot 2 - 1 \cdot 3 + 0 \cdot (-2) \\ 0 \cdot 2 + 2 \cdot 3 + 2 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 10 - 18 + 6 \\ -3 \\ 6 - 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \\ 2 \end{bmatrix} = -\begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}.$$

- $\lambda_3 = 2$: solving $A\vec{v} = 2\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix},$$

i.e.

$$\begin{cases} 5x - 6y - 3z &= 2x, \\ 0x - 1y + 0z &= 2y, \\ 0x + 2y + 2z &= 2z, \end{cases}$$

for which the solution is

$$x = z,\ y = 0.$$

Verifying the solution using the representative vector $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 - 6 \cdot 0 - 3 \cdot 1 \\ 0 \cdot 1 - 1 \cdot 0 + 0 \cdot 0 \\ 0 \cdot 1 + 2 \cdot 0 + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

To summarize: the linear transformation represented by the matrix

$$A = \begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

has three families of eigenvectors:

Eigenvalue	Eigenvector
$\lambda_1 = 5$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
$\lambda_2 = -1$	$\begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}$
$\lambda_3 = 2$	$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

0.1.4 Characteristic polynomial

Did you notice that in both above examples the expression $|A - \lambda I|$ is a polynomial in λ? This is not a coincidence - any expression of this form is a polynomial in λ, its degree depending on the form of A. We call this polynomial the **characteristic polynomial** of A and its roots are of course the eigenvecors of A.