Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-11-04

Darij Grinberg

December 7, 2019

1. Determinants

1.1. Laplace expansion (recall)

Recall from last time:

Theorem 1.1.1 (Laplace expansion along the p-th row)**.** Let A be an $n \times n$-matrix. For each $p, q \in [n]$, we let $M_{p,q}$ be the $(n-1) \times (n-1)$-matrix obtained from A by removing row p and column q. Then, for each $p \in [n]$, we have

$$\det A = \sum_{q=1}^{n} (-1)^{p+q} A_{p,q} \det \left(M_{p,q} \right).$$

Recall also a basic fact from 2019-10-23:

Theorem 1.1.2. If an $n \times n$-matrix A has two equal rows, then $\det A = 0$.

1.2. Laplace expansion in a column

Here is the analogue of Theorem 1.1.1 for columns instead of rows:

Theorem 1.2.1 (Laplace expansion along the q-th column)**.** Let A be an $n \times n$-matrix. For each $p, q \in [n]$, we let $M_{p,q}$ be the $(n-1) \times (n-1)$-matrix obtained from A by removing row p and column q. Then, for each $q \in [n]$, we have

$$\det A = \sum_{p=1}^{n} (-1)^{p+q} A_{p,q} \det \left(M_{p,q} \right).$$

Example 1.2.2. Let's compute $\det A$ for $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 4 \end{pmatrix}$. Theorem 1.2.1 (for $q = 2$) yields

$$\begin{aligned}
&\det A \\
&= -A_{1,2} \det \begin{pmatrix} A_{2,1} & A_{2,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} + A_{2,2} \det \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} - A_{3,2} \det \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{2,1} & A_{2,3} \end{pmatrix} \\
&= \underbrace{-0 \det \begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix}}_{=0} + 1 \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \underbrace{- 0 \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}}_{=0} \\
&= \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2.
\end{aligned}$$

Theorem 1.2.1 is [Strickland, Proposition B.25].

1.3. The adjugate matrix

Let me define a rather weird matrix:

Definition 1.3.1. Let A be an $n \times n$-matrix. Let $M_{p,q}$ be as in Theorem 1.1.1. Then, we define the **adjugate matrix** $\operatorname{adj} A$ of A to be the $n \times n$-matrix

$$\left((-1)^{p+q} \det \left(M_{q,p} \right) \right)_{1 \leq p \leq n,\ 1 \leq q \leq n}.$$

(Sic! This is really saying $M_{q,p}$, not $M_{p,q}$.)

In other words,

$$\operatorname{adj} A = \begin{pmatrix} \det(M_{1,1}) & -\det(M_{2,1}) & \cdots & (-1)^{n+1} \det(M_{n,1}) \\ -\det(M_{1,2}) & \det(M_{2,2}) & \cdots & (-1)^{n+2} \det(M_{n,2}) \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{1+n} \det(M_{1,n}) & (-1)^{2+n} \det(M_{2,n}) & \cdots & \det(M_{n,n}) \end{pmatrix}.$$

The signs in this matrix follow the "chessboard pattern": The top-left cell has a $+$ sign; any two adjacent cells (with a common edge, not just a common corner) always have opposite signs. (Of course, we are talking about the $(-1)^{p+q}$ signs here; the $\det \left(M_{q,p} \right)$ factors may include their own signs.)

Example 1.3.2. Let $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then,

$$\operatorname{adj} A = \begin{pmatrix} \det(M_{1,1}) & -\det(M_{2,1}) \\ -\det(M_{1,2}) & \det(M_{2,2}) \end{pmatrix} = \begin{pmatrix} \det(\ d\ ) & -\det(\ b\ ) \\ -\det(\ c\ ) & \det(\ a\ ) \end{pmatrix}$$
$$= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(since the determinant of a 1×1-matrix is just its unique entry).

Example 1.3.3. Let $n = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Then,

$$\operatorname{adj} A = \begin{pmatrix} \det(M_{1,1}) & -\det(M_{2,1}) & \det(M_{3,1}) \\ -\det(M_{1,2}) & \det(M_{2,2}) & -\det(M_{3,2}) \\ \det(M_{1,3}) & -\det(M_{2,3}) & \det(M_{3,3}) \end{pmatrix}$$
$$= \begin{pmatrix} \det\begin{pmatrix} e & f \\ h & i \end{pmatrix} & -\det\begin{pmatrix} b & c \\ h & i \end{pmatrix} & \det\begin{pmatrix} b & c \\ e & f \end{pmatrix} \\ -\det\begin{pmatrix} d & f \\ g & i \end{pmatrix} & \det\begin{pmatrix} a & c \\ g & i \end{pmatrix} & -\det\begin{pmatrix} a & c \\ d & f \end{pmatrix} \\ \det\begin{pmatrix} d & e \\ g & h \end{pmatrix} & -\det\begin{pmatrix} a & b \\ g & h \end{pmatrix} & \det\begin{pmatrix} a & b \\ d & e \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} ie - fh & ch - ib & bf - ce \\ fg - id & ia - cg & cd - af \\ dh - ge & bg - ah & ae - bd \end{pmatrix}.$$

Any entry of the matrix $\operatorname{adj} A$ is a degree-$(n-1)$ polynomial in the entries of A.

Theorem 1.3.4. Let A be an $n \times n$-matrix. Then,

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

Example 1.3.5. Let $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then, $\operatorname{adj} A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Thus,

$$A \cdot \operatorname{adj} A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$
$$= \underbrace{(ad - bc)}_{= \det A} I_2 = \det A \cdot I_2$$

and

$$\operatorname{adj} A \cdot A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}$$
$$= \underbrace{(ad-bc)}_{=\det A} I_2 = \det A \cdot I_2.$$

Example 1.3.6. Let $n = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, so

$$\operatorname{adj} A = \begin{pmatrix} ie-fh & ch-ib & bf-ce \\ fg-id & ia-cg & cd-af \\ dh-ge & bg-ah & ae-bd \end{pmatrix}.$$

Then,

$$A \cdot \operatorname{adj} A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \cdot \begin{pmatrix} ie-fh & ch-ib & bf-ce \\ fg-id & ia-cg & cd-af \\ dh-ge & bg-ah & ae-bd \end{pmatrix}.$$

Let us compute this product by hand. The $(1,1)$-entry is

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} ie-fh \\ fg-id \\ dh-ge \end{pmatrix}$$
$$= a(ie-fh) + b(fg-id) + c(dh-ge)$$
$$= \sum_{q=1}^{n} (-1)^{1+q} A_{1,q} \det\left(M_{1,q}\right) = \det A \qquad \text{(by Laplace expansion along 1-st row)}.$$

Similarly, all other diagonal entries of $A \cdot \operatorname{adj} A$ are $\det A$.

The $(1,2)$-entry of $A \cdot \operatorname{adj} A$ is

$$\begin{pmatrix} a & b & c \end{pmatrix} \cdot \begin{pmatrix} ch-ib \\ ia-cg \\ bg-ah \end{pmatrix}$$
$$= a(ch-ib) + b(ia-cg) + c(bg-ah)$$
$$= \underbrace{\det \begin{pmatrix} a & b & c \\ a & b & c \\ g & h & i \end{pmatrix}}_{\substack{=0 \\ \text{(since this matrix has two equal rows)}}}$$
$$\left(\begin{array}{c} \text{because if we compute } \det \begin{pmatrix} a & b & c \\ a & b & c \\ g & h & i \end{pmatrix} \\ \text{by Laplace expansion along the 2-nd row,} \\ \text{then we get precisely } a(ch-ib) + b(ia-cg) + c(bg-ah) \end{array} \right)$$

$= 0.$

Similarly, all off-diagonal entries of $A \cdot \operatorname{adj} A$ are 0.

Thus, $A \cdot \operatorname{adj} A = \det A \cdot I_3$.

A similar argument using columns instead of rows shows that $\operatorname{adj} A \cdot A = \det A \cdot I_3$. This uses Laplace expansion along a column (i.e., Theorem 1.2.1). Thus, $A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_3$.

In the last example, we have seen why Theorem 1.3.4 holds for $n = 3$. The general case can be proved along the same lines. See [Strickland, Proposition B.28] for the general proof written out rigorously.

Corollary 1.3.7. If an $n \times n$-matrix A is invertible, then $A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A$.

Example 1.3.8. For $n = 2$, this leads to

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \frac{d}{ad - bc} & \frac{-b}{ad - bc} \\ \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{pmatrix}.$$

We've seen this before.

Proof of Corollary 1.3.7. Assume that A is invertible. Then, $\det A \neq 0$ (by Theorem 1.4.1 in the classwork from 2019-10-30). But Theorem 1.3.4 yields $A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n$. We can divide all sides of this equality by $\det A$ (since $\det A \neq 0$), and thus obtain

$$\frac{1}{\det A} \cdot A \cdot \operatorname{adj} A = \frac{1}{\det A} \cdot \operatorname{adj} A \cdot A = I_n.$$

This rewrites as

$$A \cdot \left(\frac{1}{\det A} \operatorname{adj} A \right) = \left(\frac{1}{\det A} \operatorname{adj} A \right) \cdot A = I_n$$

(since $\frac{1}{\det A}$ is just a number, and thus can be moved around the products freely). But this shows that $\frac{1}{\det A} \operatorname{adj} A$ is an inverse of A. Hence, $A^{-1} = \frac{1}{\det A} \operatorname{adj} A$. This proves Corollary 1.3.7. □

Remark 1.3.9. The adjugate of a matrix is also called its **classical adjoint** or sometimes just **adjoint**. But beware of the latter word, as it can mean many different things. I suggest you use "adjugate" simply because it only has one meaning.

Corollary 1.3.7 is known as **Cramer's rule**. However, there is something else also called Cramer's rule (see, e.g., the eponymous Wikipedia page), so beware of confusion.

In order to compute $\operatorname{adj} A$ according to the definition, you have to define n^2 determinants of $(n-1) \times (n-1)$-matrices. Thus, I do not recommend using Corollary 1.3.7 for finding A^{-1} (unless n is small), let alone for solving systems of linear equations (as it only applies when A is invertible).

2. Eigenvalues and eigenvectors ("eigenstuff")

We shall now study **eigenvalues** and **eigenvectors** of (square) matrices. This subject, at first, will look like an intellectual game; but soon, we will see what it is good for, and later will (if time allows...) even see what it "really means".

We follow [Strickland, §13] at first.

2.1. Definition and examples

Definition 2.1.1. Let A be an $n \times n$-matrix. Let λ be a scalar (i.e., a real number).

(a) A **λ-eigenvector of** A means a nonzero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$.

(b) We say that λ is an **eigenvalue** of A if and only if there exists a λ-eigenvector of A.

Caution: Many authors drop the "nonzero" in the definition of "λ-eigenvector of A". Then, of course, 0 (the zero vector) is **always** a λ-eigenvector of A, since $A \cdot 0 = 0 = \lambda \cdot 0$. Thus, if you do so, you need to modify the definition of "eigenvalue" so that it requires the existence of a **nonzero** λ-eigenvector.

Remark: This all only makes sense for square matrices A. If A is not square, then Av and λv will have different sizes, thus never equal.

Example 2.1.2. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Can we find the eigenvalues of A ? Can we find the corresponding eigenvectors?

Let λ be a scalar. We want to find the λ-eigenvectors of A. These are the nonzero vectors $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ such that $A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ (by definition). Now, let us rewrite this condition:

$$\begin{aligned} & \left(A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \right) \\ \Longleftrightarrow \ & \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \right) \\ \Longleftrightarrow \ & \left(\begin{pmatrix} x+y \\ x+y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \right) \end{aligned}$$

$$\iff (x + y = \lambda x \text{ and } x + y = \lambda y).$$

If we regard λ as fixed, this is a system of linear equations in x, y, and its augmented matrix is $\begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \end{pmatrix}$. We can solve it by Gaussian elimination, for example:

$$\begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \end{pmatrix} \overset{\text{swap row 1 with row 2}}{\longrightarrow} \begin{pmatrix} 1 & 1-\lambda & 0 \\ 1-\lambda & 1 & 0 \end{pmatrix}$$
$$\overset{\text{add } (\lambda-1)\cdot\text{row 1 to row 2}}{\longrightarrow} \begin{pmatrix} 1 & 1-\lambda & 0 \\ 0 & 2\lambda-\lambda^2 & 0 \end{pmatrix}.$$

Thus, we see that:

- If $2\lambda - \lambda^2 \neq 0$, then there is a pivot in each of the first two columns; thus, the linear system has a unique solution, which of course must be the zero vector 0 (because 0 is surely a solution), and this means that A has no λ-eigenvector (since a λ-eigenvector must be nonzero).
- If $2\lambda - \lambda^2 = 0$, then the second column has no pivot; thus, the linear system has a free variable, and therefore there exists a nonzero solution; this means that A has a λ-eigenvector.

Thus, A has a λ-eigenvector if and only if $2\lambda - \lambda^2 = 0$. But $2\lambda - \lambda^2 = 0$ is equivalent to $\lambda \in \{0, 2\}$. Therefore, 0-eigenvectors and 2-eigenvectors exist (for our A), and no other eigenvectors do (i.e., if $\lambda \notin \{0, 2\}$, then λ-eigenvectors don't exist). In other words, the eigenvalues of A are 0 and 2.

How do we find the eigenvectors? By solving the respective systems of linear equations:

- The 0-eigenvectors of A are the nonzero solutions of $A\begin{pmatrix} x \\ y \end{pmatrix} = 0\begin{pmatrix} x \\ y \end{pmatrix}$; they are all of the form $\begin{pmatrix} -y \\ y \end{pmatrix}$. Picking y nonzero gives a 0-eigenvector.
- The 2-eigenvectors of A are the nonzero solutions of $A\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix}$; they are all of the form $\begin{pmatrix} y \\ y \end{pmatrix}$. Picking y nonzero gives a 2-eigenvector.

This was rather haphazard: We cannot hope that Gaussian elimination will always proceed this nicely. If A is more complicated, then the choice of pivot might eventually depend on the value of λ. Thus, we have the following problem: How do we compute the eigenvalues and the corresponding eigenvectors of an $n \times n$-

matrix in general?

We will see this soon enough; for now, let us recall the Inverse Matrix Theorem (Theorem 1.2.1 in the classwork from 2019-10-16), but smuggle two more equivalent statements into it:

Theorem 2.1.3. (The Inverse Matrix Theorem, updated.)

Let A be an $n \times n$-matrix.

Then, the following statements are equivalent (i.e., if any of them holds, then so do all the others):

- **(a)** The matrix A can be row-reduced to I_n.
- **(b)** The columns of A are linearly independent.
- **(c)** The columns of A span $\mathbb{R}^n$.
- **(d)** The columns of A form a basis of $\mathbb{R}^n$.
- **(e)** The matrix A^T can be row-reduced to I_n.
- **(f)** The columns of A^T are linearly independent.
- **(g)** The columns of A^T span $\mathbb{R}^n$.
- **(h)** The columns of A^T form a basis of $\mathbb{R}^n$.
- **(i)** The matrix A has a left inverse.
- **(j)** The matrix A has a right inverse.
- **(k)** The matrix A has an inverse (i.e., A is invertible).
- **(l)** We have $\det A \neq 0$.
- **(m)** The only column vector $v \in \mathbb{R}^n$ satisfying $Av = 0$ is the zero vector 0.

Proof of Theorem 2.1.3. The Inverse Matrix Theorem (Theorem 1.2.1 in the classwork from 2019-10-16) shows that the 11 statements **(a)**, **(b)**, ..., **(k)** are equivalent. Thus, we only need to link **(l)** and **(m)** to these 11 statements.

The equivalence **(k)** $\Longleftrightarrow$ **(l)** is precisely Theorem 1.4.1 in the classwork from 2019-10-30. Thus, statement **(l)** is added to our list of 11 equivalent statements. It remains to link **(m)**.

It is easy to see that **(k)** $\Longrightarrow$ **(m)**: Indeed, if **(k)** holds, then the matrix A has an inverse A^{-1}, and therefore every column vector $v \in \mathbb{R}^n$ satisfying $Av = 0$ must be the zero vector 0 (because comparing $\underbrace{A^{-1}A}_{=I_n} v = I_n v = v$ with $A^{-1} \underbrace{Av}_{=0} = A^{-1}0 = 0$, we find $v = 0$); but this means that **(m)** holds.

Let us now prove that **(m)** $\Longrightarrow$ **(k)**. Indeed, assume that **(m)** holds. We need to prove that **(k)** holds. Equivalently, we need to prove that **(b)** holds (since we already know that **(b)** is equivalent to **(k)**). Let $v_1, v_2, \ldots, v_n$ be the columns of A (so that $A = [v_1 \mid v_2 \mid \cdots \mid v_n]$). Let $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$ be any relation between these columns. We shall prove that this relation is trivial. Indeed, let v denote the column vector $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$. Thus, from $A = [v_1 \mid v_2 \mid \cdots \mid v_n]$ and $v = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$, we obtain

$$\begin{aligned} Av &= [v_1 \mid v_2 \mid \cdots \mid v_n] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \qquad \text{(by Lemma 1.1.6 in the classwork from 2019-10-07)} \\ &= 0. \end{aligned}$$

Hence, according to statement **(k)**, the vector v must be the zero vector 0. Thus, its entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all 0. Hence, the relation $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$ must be trivial. Thus, we have proved that every relation between the columns of A is trivial. In other words, the columns of A are linearly independent. In other words, statement **(b)** holds. As explained above, this proves that **(m)** $\Longrightarrow$ **(k)**.

Combining **(k)** $\Longrightarrow$ **(m)** with **(m)** $\Longrightarrow$ **(k)**, we obtain the equivalence **(k)** $\Longleftrightarrow$ **(m)**. Thus, statement **(m)** is added to our list of equivalent statements. Theorem 2.1.3 is thus proved. □

Let us state the contrapositive of Theorem 2.1.3:

Theorem 2.1.4. (The Non-Inverse Matrix Theorem, updated.)
Let A be an $n \times n$-matrix.
Then, the following statements are equivalent (i.e., if any of them holds, then so do all the others):

- **(a′)** The matrix A cannot be row-reduced to I_n.
- **(b′)** The columns of A are linearly dependent.
- **(c′)** The columns of A do not span $\mathbb{R}^n$.
- **(d′)** The columns of A do not form a basis of $\mathbb{R}^n$.
- **(e′)** The matrix A^T cannot be row-reduced to I_n.
- **(f′)** The columns of A^T are linearly dependent.

- **(g′)** The columns of A^T do not span $\mathbb{R}^n$.
- **(h′)** The columns of A^T do not form a basis of $\mathbb{R}^n$.
- **(i′)** The matrix A has no left inverse.
- **(j′)** The matrix A has no right inverse.
- **(k′)** The matrix A has no inverse (i.e., A is not invertible).
- **(l′)** We have $\det A = 0$.
- **(m′)** There exists a nonzero column vector $v \in \mathbb{R}^n$ satisfying $Av = 0$.

Proof of Theorem 2.1.4. Each of the statements in Theorem 2.1.4 is the negation of the corresponding statement in Theorem 2.1.3. (For example, statement **(a′)** in Theorem 2.1.4 is the negation of statement **(a)** in Theorem 2.1.3.) Thus, Theorem 2.1.4 follows from Theorem 2.1.3 (because if two statements are equivalent, then so are their negations). □

We now revisit Example 2.1.2:

Example 2.1.5. Let us return to Example 2.1.2, so $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, but now let us try a different approach (instead of Gaussian elimination):

For any scalar λ, we have the following equivalence:

$$\begin{aligned}
&(\lambda \text{ is an eigenvalue of } A) \\
\Longleftrightarrow\ & (\text{there exists a } \lambda\text{-eigenvector of } A) \\
& \quad (\text{by the definition of “eigenvalue”}) \\
\Longleftrightarrow\ & (\text{there exists a nonzero column vector } v \in \mathbb{R}^n \text{ satisfying } Av = \lambda v) \\
& \quad \left(\begin{array}{c} \text{since a } \lambda\text{-eigenvector of } A \text{ is defined to be} \\ \text{a nonzero column vector } v \in \mathbb{R}^n \text{ satisfying } Av = \lambda v \end{array} \right) \\
\Longleftrightarrow\ & (\text{there exists a nonzero column vector } v \in \mathbb{R}^n \text{ satisfying } (A - \lambda I_n) v = 0) \\
& \quad \left(\begin{array}{c} \text{since } Av = \lambda v \text{ is equivalent to } (A - \lambda I_n) v = 0 \\ (\text{because } (A - \lambda I_n) v = Av - \lambda \underbrace{I_n v}_{=v} = Av - \lambda v) \end{array} \right) \\
\Longleftrightarrow\ & (\det (A - \lambda I_n) = 0) \\
& \quad \left(\begin{array}{c} \text{by the equivalence } \textbf{(m′)} \Longleftrightarrow \textbf{(l′)} \text{ in Theorem 2.1.4,} \\ \text{applied to the matrix } A - \lambda I_n \text{ instead of } A \end{array} \right).
\end{aligned}$$

Thus, in order to find the eigenvalues of A, you just compute $\det (A - \lambda I_n)$ for general λ, and find out what values of λ make it 0.

In our case, $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $n = 2$, so

$$A - \lambda I_n = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix}.$$

Thus,

$$\det(A - \lambda I_n) = \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)(1-\lambda) - 1 \cdot 1 = -2\lambda + \lambda^2.$$

So $\det(A - \lambda I_n) = 0$ if and only if $-2\lambda + \lambda^2 = 0$, which is equivalent to $\lambda \in \{0, 2\}$. Thus, the eigenvalues of A are 0 and 2.

The method we just showed works in general: For any $n \times n$-matrix A, the eigenvalues of A are the scalars λ for which $\det(A - \lambda I_n) = 0$. Let's give this determinant a name:

Definition 2.1.6. Let A be an $n \times n$-matrix. We define

$$\chi_A(t) = \det(A - tI_n).$$

This is a polynomial in t, and is called the **characteristic polynomial** of A.

Proposition 2.1.7. Let A be an $n \times n$-matrix. Then, the eigenvalues of A are the roots of the characteristic polynomial $\chi_A(t)$.

Proof. For any scalar λ, we have the equivalence

$$\begin{aligned}
&(\lambda \text{ is an eigenvalue of } A) \\
\Longleftrightarrow\ & (\det(A - \lambda I_n) = 0) \qquad \text{(as we have seen in Example 2.1.5)} \\
\Longleftrightarrow\ & (\det(A - tI_n) \text{ becomes 0 if we substitute } \lambda \text{ for } t) \\
\Longleftrightarrow\ & (\lambda \text{ is a root of the polynomial } \det(A - tI_n)) \\
\Longleftrightarrow\ & (\lambda \text{ is a root of the polynomial } \chi_A(t)) \qquad (\text{since } \det(A - tI_n) = \chi_A(t)).
\end{aligned}$$

Thus, the eigenvalues of A are the roots of the characteristic polynomial $\chi_A(t)$. □

Example 2.1.8. Let $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then, the characteristic polynomial $\chi_A(t)$ of A is

$$\begin{aligned}
\chi_A(t) = \det(A - tI_n) &= \det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} - t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\
&= \det \begin{pmatrix} a-t & b \\ c & d-t \end{pmatrix} = (a-t)(d-t) - bc
\end{aligned}$$

$$= t^2 - \underbrace{(a+d)}_{\substack{=\operatorname{Tr} A \\ \text{(the trace of } A\text{)}}} t + \underbrace{(ad-bc)}_{=\det A}.$$

Thus, the eigenvalues of A are the roots of this polynomial; they are

$$\frac{a+d \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2} = \frac{a+d \pm \sqrt{(a-d)^2 + 4bc}}{2}.$$

For example, if $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, then the eigenvalues of A are

$$\frac{1+4 \pm \sqrt{(1-4)^2 + 4 \cdot 2 \cdot 3}}{2} = \frac{5 \pm \sqrt{33}}{2}.$$

Example 2.1.9. We can try to play the same game with $n = 3$: Let $n = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Then,

$$\chi_A(t) = \det(A - tI_n) = \det \begin{pmatrix} a-t & b & c \\ d & e-t & f \\ g & h & i-t \end{pmatrix}$$
$$= -t^3 + \underbrace{(a+e+i)}_{=\operatorname{Tr} A} t^2 - \underbrace{(ia+ie+ae-bd-cg-fh)}_{=\det(M_{1,1})+\det(M_{2,2})+\det(M_{3,3})} t$$
$$+ \underbrace{(iae - ibd - cge - afh + bfg + cdh)}_{=\det A}.$$

Unfortunately, this is a polynomial of degree 3, and there is no good way to compute the roots of such a polynomial explicitly.

However, if A is "nice" (e.g., strategically chosen for a homework problem), then the polynomial may have neat rational or integer roots.

Example 2.1.10. Let $n = 3$ and $A = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}$ (an arbitrary lower-triangular matrix). Then,

$$\chi_A(t) = \det \begin{pmatrix} a-t & 0 & 0 \\ b & c-t & 0 \\ d & e & f-t \end{pmatrix} = (a-t)(c-t)(f-t).$$

The roots of $\chi_A(t)$ clearly are a, c, f, which are of course the diagonal entries of

A.

This holds more generally:

Proposition 2.1.11. If a matrix A is triangular, then its eigenvalues are its diagonal entries.

Proof. Let A be triangular. Then, the matrix $A - tI_n$ is triangular as well, and thus its determinant $\det(A - tI_n)$ equals the product of its diagonal entries $A_{1,1} - t, A_{2,2} - t, \ldots, A_{n,n} - t$. Hence,

$$\chi_A(t) = \det(A - tI_n) = (A_{1,1} - t)(A_{2,2} - t)\cdots(A_{n,n} - t).$$

Thus, the roots of $\chi_A(t)$ are $A_{1,1}, A_{2,2}, \ldots, A_{n,n}$. But these roots are the eigenvalues of A (by Proposition 2.1.7). Hence, the eigenvalues of A are $A_{1,1}, A_{2,2}, \ldots, A_{n,n}$. That is, they are the diagonal entries of A. This proves the proposition. □

Now, assuming you have found all eigenvalues of a matrix A, how do you find the eigenvectors?

Method for finding eigenvalues and eigenvectors of a matrix:

Given an $n \times n$-matrix A, we can find all eigenvalues and eigenvectors of A as follows:

- Calculate the characteristic polynomial $\chi_A(t) = \det(A - tI_n)$ of A.
- Find all roots $\lambda_1, \lambda_2, \ldots, \lambda_k$ of $\chi_A(t)$. These are the eigenvalues of A.
- For each eigenvalue λ_i, compute the nonzero solutions v to $(A - \lambda_i I_n)v = 0$ (for example, using Gaussian elimination). These are the λ_i-eigenvectors of A.

Example 2.1.12. Let

$$A = \begin{pmatrix} 16 & 2 & 1 & 1 \\ 2 & 16 & 1 & 1 \\ 1 & 1 & 16 & 2 \\ 1 & 1 & 2 & 16 \end{pmatrix}.$$

What are the eigenvalues and eigenvectors of A ?

The characteristic polynomial of A is

$$\chi_A(t) = \det(A - tI_n) = \det\begin{pmatrix} 16-t & 2 & 1 & 1 \\ 2 & 16-t & 1 & 1 \\ 1 & 1 & 16-t & 2 \\ 1 & 1 & 2 & 16-t \end{pmatrix}$$

$$= \det\begin{pmatrix} 16-t & 2 & 1 & 1 \\ -14+t & 14-t & 0 & 0 \\ 1 & 1 & 16-t & 2 \\ 1 & 1 & 2 & 16-t \end{pmatrix}$$

$$= \det \begin{pmatrix} 16-t & 2 & 1 & 1 \\ -14+t & 14-t & 0 & 0 \\ 1 & 1 & 16-t & 2 \\ 0 & 0 & -14+t & 14-t \end{pmatrix}$$
$$= \det \begin{pmatrix} 16-t & 2 & 2 & 1 \\ -14+t & 14-t & 0 & 0 \\ 1 & 1 & 18-t & 2 \\ 0 & 0 & 0 & 14-t \end{pmatrix}$$
$$= \det \begin{pmatrix} 18-t & 2 & 2 & 1 \\ 0 & 14-t & 0 & 0 \\ 2 & 1 & 18-t & 2 \\ 0 & 0 & 0 & 14-t \end{pmatrix}$$
$$= (14-t) \det \begin{pmatrix} 18-t & 2 & 2 \\ 0 & 14-t & 0 \\ 2 & 1 & 18-t \end{pmatrix}$$
$$= \underbrace{(14-t)(14-t)}_{=(14-t)^2} \underbrace{\det \begin{pmatrix} 18-t & 2 \\ 2 & 18-t \end{pmatrix}}_{\substack{=(18-t)(18-t)-2\cdot 2 \\ =t^2-36t+320 \\ =(t-16)(t-20)}}$$
$$= (14-t)^2 (t-16)(t-20).$$

So the roots of $\chi_A(t)$ (thus the eigenvalues of A) are 14, 16 and 20.

What about the eigenvectors?

The 14-eigenvectors of A are the nonzero vectors v such that $(A - 14I_4) v = 0$. Since

$$A - 14I_4 = \begin{pmatrix} 16-14 & 2 & 1 & 1 \\ 2 & 16-14 & 1 & 1 \\ 1 & 1 & 16-14 & 2 \\ 1 & 1 & 2 & 16-14 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix},$$

this boils down to solving

$$\begin{pmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

You can do this by Gaussian elimination, and obtain

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} -y \\ y \\ -w \\ w \end{pmatrix}.$$

So the 14-eigenvectors of A are the nonzero vectors of the form $\begin{pmatrix} -y \\ y \\ -w \\ w \end{pmatrix}$.

We can compute the 16-eigenvectors and the 20-eigenvectors as well; we get

$$\begin{aligned} A - 16I_4 &= \begin{pmatrix} 16-16 & 2 & 1 & 1 \\ 2 & 16-16 & 1 & 1 \\ 1 & 1 & 16-16 & 2 \\ 1 & 1 & 2 & 16-16 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{(RREF)}, \end{aligned}$$

so the 16-eigenvectors are

$$\begin{pmatrix} -w \\ -w \\ w \\ w \end{pmatrix} \qquad \text{with } w \neq 0.$$

You can find the 20-eigenvectors in the same way.

Let me comment on what can go "wrong" in the process:

Example 2.1.13. Let

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

Does it have eigenvalues? Its eigenvalues are the roots of $\chi_A(t)$. But

$$\chi_A(t) = \det \begin{pmatrix} 0-t & 1 \\ -1 & 0-t \end{pmatrix} = (0-t)(0-t) - 1(-1) = t^2 + 1.$$

What are the roots of this polynomial?

If we stick to real numbers, then this polynomial has no roots, since $x^2 + 1 > 0$ for every real x. Thus, if we stick to real numbers, then A has no eigenvalues and thus no eigenvectors. However, there is a way to force A to have eigenvalues:

Just extend the "field of scalars" to the complex numbers, in which case i and $-i$ become eigenvalues. Linear algebra using complex numbers as scalars is analogous to linear algebra using real numbers as scalars, but with the nice feature that each non-constant polynomial has at least one root (and, in fact, can be factored into linear factors), so every $n \times n$-matrix A (with $n > 0$) has at least one eigenvalue.

References

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
http://neil-strickland.staff.shef.ac.uk/courses/MAS201/