MA313 · 2026–27 · ILIA PIRASHVILI

Linear Algebra

A revision guide

A revsion guide for the module, not a substitte for the lectures.
It gives an overview of the module only, and ommits many proofs and details.

ilia.pirashvili@universityofgalway.com

Contents

I Matrices 1

1 Matrices and vectors . . . 2
- 1.1 Shapes, sums and products . . . 2
- 1.2 The transpose . . . 4
- 1.3 Vectors are thin matrices, and the dot product . . . 4

2 Determinants . . . 7
- 2.1 What the determinant tells you . . . 7
- 2.2 Computing it . . . 7
- 2.3 Two rules, and what follows . . . 9

3 Inverses and row reduction . . . 11
- 3.1 What an inverse is . . . 11
- 3.2 Row operations and the Gauss–Jordan algorithm . . . 12
- 3.3 Echelon forms (Lecture 8) . . . 14

II Vector spaces 16

4 Vector spaces . . . 17
- 4.1 The definition . . . 17
- 4.2 The standard examples . . . 18
- 4.3 Deciding whether something is a vector space . . . 19

5 Subspaces . . . 21
- 5.1 The subspace test . . . 21
- 5.2 Recognising subspaces quickly . . . 22
- 5.3 Worked examples . . . 23
- 5.4 Simplify first, and combining subspaces . . . 24

III Building subspaces 27

6 Linear combinations and span . . . 28
- 6.1 Linear combinations . . . 28
- 6.2 Span . . . 30
- 6.3 Spanning sets (Lecture 7) . . . 31

7 The kernel of a matrix . . . 34
- 7.1 Definition and membership . . . 34
- 7.2 Computing a kernel . . . 35
- 7.3 Two ways of describing a subspace . . . 36

8 Column space and image . . . 38
- 8.1 Definition . . . 38
- 8.2 Is $\mathbf{b}$ in the column space? . . . 39

IV Linear maps 41

9 Linear transformations . . . 42

9.1 Linear maps . . . 42
9.2 Maps and matrices . . . 43
9.3 Kernel and image in general . . . 43

V Bases and dimension **45**

10 Linear independence and bases . . . **46**
10.1 Linear independence . . . 46
10.2 Bases and dimension . . . 47
10.3 Dimension and the Basis Theorem (Lecture 13) . . . 48

11 Bases of kernel and image . . . **50**
11.1 Basis of the kernel . . . 50
11.2 Basis of the column space . . . 51

12 Coordinates . . . **53**

VI Change of basis **55**

13 Isomorphism and change of basis . . . **56**
13.1 Injective, surjective, isomorphic . . . 56
13.2 The coordinate isomorphism and change of basis . . . 57

VII Rank and the fundamental subspaces **59**

14 Rank and nullity . . . **60**

15 The four fundamental subspaces . . . **63**

VIII Orthogonality **66**

16 Inner products and orthogonality . . . **67**
16.1 Inner product, length, angle . . . 67
16.2 Projection onto a vector . . . 68
16.3 Orthogonal complements . . . 68
16.4 Orthogonal sets . . . 69

17 Orthogonal projection and Gram–Schmidt . . . **70**
17.1 Projection onto a subspace . . . 70
17.2 Gram–Schmidt . . . 71
17.3 Orthogonal matrices . . . 71

18 Best approximation and least squares . . . **73**

IX Eigenvalues **76**

19 Eigenvalues and eigenvectors . . . **77**
19.1 Definitions . . . 77
19.2 Diagonalisation . . . 78

20 Powers and the long run . . . **81**

X Singular values **83**

21 The singular value decomposition . . . **84**
21.1 Symmetric matrices and $A^{\mathsf{T}}A$ . . . 84
21.2 Singular values . . . 84
21.3 The SVD . . . 85

22 Pseudoinverse and least squares . . . **88**

PART I

Matrices

CHAPTER 1

Matrices and vectors

LECTURE 1 · *shapes, products, transposes and the dot product*

THIS CHAPTER TEACHES HOW TO

- construct $A + B$ and AB if they are defined, and what shape the answer has;
- multiply two matrices, and transpose a product the right way round;
- explain, with an example, why $AB \neq BA$ in general;
- compute the length of a vector and the angle between two vectors, and test perpendicularity.

1.1 Shapes, sums and products

SHAPE, SUM, SCALAR MULTIPLE

An $m \times n$ **matrix** has m rows and n columns (*rows first*). Its entry in row i, column j is $a_{i,j}$; the set of all of them is $\mathrm{Mat}_{m\times n}(\mathbb{R})$. $A + B$ and λA are done entry by entry, and $A + B$ needs equal shapes.

THE PRODUCT

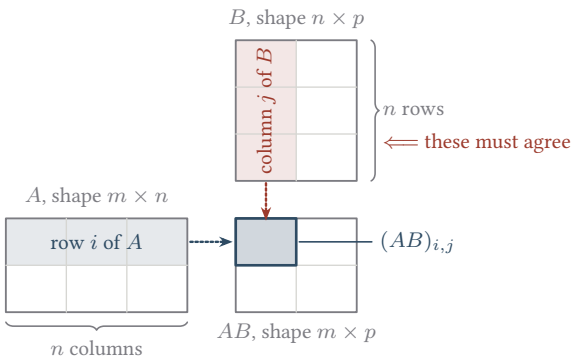
AB is defined *only* if the number of columns of A equals the number of rows of B. The two inner numbers cancel in their shapes:

$$\underset{m\times n}{A}\ \underset{n\times p}{B} = \underset{m\times p}{AB}, \qquad (AB)_{i,j} = \sum_{k=1}^{n} a_{i,k}\, b_{k,j} = a_{i,-} \cdot b_{-,j}.$$

In words: walk along row i of A and down column j of B, multiply the paired numbers, add them up. ($a_{i,-}$ is the i-th row of A and $b_{-,j}$ the j-th column of B; the dash means "let this index run".)

HOW TO MULTIPLY TWO MATRICES

1. Write the two shapes side by side. Are the inner numbers equal? If not, say "AB is not defined, because number of columns of A does not match number of rows of B".
2. The outer numbers are the shape of AB. Draw an empty matrix of that shape.
3. Fill entry (i, j) with $a_{i,-} \cdot b_{-,j}$.



EXAMPLE 1.1 · a product, entry by entry

Compute AB for $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix}$.

SOLUTION

Shapes: $(2 \times 3)(3 \times 2)$: the inner numbers agree, so AB is defined and is 2×2.

$$(1,1):\ 1\cdot 3 + 0\cdot 2 + 2\cdot 1 = 5 \qquad (1,2):\ 1\cdot 1 + 0\cdot 1 + 2\cdot 0 = 1$$
$$(2,1):\ (-1)\cdot 3 + 3\cdot 2 + 1\cdot 1 = 4 \qquad (2,2):\ (-1)\cdot 1 + 3\cdot 1 + 1\cdot 0 = 2$$

Hence $AB = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$. Note that BA is also defined, but it is 3×3: a different shape, so $AB \neq BA$ overtly.

LIVE: watch this product built entry by entry ▸.

ORDER MATTERS

$AB \neq BA$ in general. AB can be defined while BA is not, and even when both are defined they may have different shapes. And even for two square matrices of the same size,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}: \qquad AB = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \qquad BA = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}.$$

Think of a matrix as an instruction: AB means *do B first, then A*. A rotation followed by a reflection is not the same as the reflection followed by the rotation.

SHAPE CHECK

Before multiplying, predict the shape of the answer. If your final matrix has the wrong shape, you have mis-paired rows and columns. Each entry of AB is a sum of exactly n products, where n is the inner number.

1.1.1 The zero and the identity

ZERO AND IDENTITY

The **identity** $I = I_n$ is *square*: ones on the diagonal, zeros elsewhere. The **zero matrix** has only zeros, in any shape. There is one of each *per shape*, and whenever the shapes allow it

$$0 \times M \;=\; M \times 0 \;=\; 0, \qquad I \times M \;=\; M \times I \;=\; M.$$

1.2 The transpose

TRANSPOSE

$(A^\mathsf{T})_{i,j} = a_{j,i}$: rows and columns swap, so an $m \times n$ matrix becomes $n \times m$. Also $(A^\mathsf{T})^\mathsf{T} = A$ and, crucially,

$$(AB)^\mathsf{T} \;=\; B^\mathsf{T}A^\mathsf{T} \qquad \text{(the order reverses).}$$

EXAMPLE 1.2 · checking $(AB)^\mathsf{T} = B^\mathsf{T}A^\mathsf{T}$

With A, B from Example 1.1: $AB = \left[\begin{smallmatrix} 5 & 1 \\ 4 & 2 \end{smallmatrix}\right]$, so $(AB)^\mathsf{T} = \left[\begin{smallmatrix} 5 & 4 \\ 1 & 2 \end{smallmatrix}\right]$. And

$$B^\mathsf{T}A^\mathsf{T} \;=\; \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 3 \\ 2 & 1 \end{bmatrix} \;=\; \begin{bmatrix} 3+0+2 & -3+6+1 \\ 1+0+0 & -1+3+0 \end{bmatrix} \;=\; \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}.$$

USE SHAPES TO CATCH THE ORDER ERROR

$A^\mathsf{T}B^\mathsf{T}$ is $(3 \times 2)(2 \times 3) = 3 \times 3$, while $(AB)^\mathsf{T}$ is 2×2. They cannot be equal. Only $B^\mathsf{T}A^\mathsf{T}$, $(2 \times 3)(3 \times 2)$, has the right shape.

1.3 Vectors are thin matrices, and the dot product

VECTORS AS MATRICES

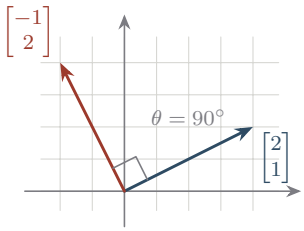
A **column vector** in $\mathbb{R}^n$ is an $n \times 1$ matrix; a **row vector** is a $1 \times n$ matrix, namely $\mathbf{u}^\mathsf{T}$. "Vector" means column vector unless said otherwise. Adding and scaling are the matrix operations, done one coordinate at a time. Two non-zero vectors are **parallel** when each is a multiple of the other; the multiples $\{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\}$ of one non-zero vector form a line through the origin.

DOT PRODUCT, LENGTH, ANGLE

For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the product $\mathbf{u}^\mathsf{T}\mathbf{v}$ is $(1 \times n)(n \times 1)$, a 1×1 matrix, i.e. a number:

$$\mathbf{u} \cdot \mathbf{v} \;=\; \mathbf{u}^\mathsf{T}\mathbf{v} \;=\; u_1v_1 + \cdots + u_nv_n, \qquad \|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}, \qquad \mathbf{u} \cdot \mathbf{v} \;=\; \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta.$$

So $\mathbf{u}$ and $\mathbf{v}$ are **perpendicular** exactly when $\mathbf{u} \cdot \mathbf{v} = 0$.



$\begin{bmatrix}2\\1\end{bmatrix} \cdot \begin{bmatrix}-1\\2\end{bmatrix} = -2 + 2 = 0$: perpendicular.

HOW TO FIND THE ANGLE BETWEEN TWO VECTORS

1. Compute the dot product $\mathbf{u} \cdot \mathbf{v}$.
2. Compute both lengths $\|\mathbf{u}\|$, $\|\mathbf{v}\|$.
3. $\cos\theta = \dfrac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$; then $\theta = \arccos(\cdot)$.

EXAMPLE 1.3 · an angle

Find the angle between $\mathbf{u} = (1, 2, 3)^\mathsf{T}$ and $\mathbf{v} = (4, -1, 0)^\mathsf{T}$.

SOLUTION

$\mathbf{u} \cdot \mathbf{v} = 4 - 2 + 0 = 2 \neq 0$, so they are not perpendicular. $\|\mathbf{u}\| = \sqrt{14}$, $\|\mathbf{v}\| = \sqrt{17}$, so

$$\cos\theta = \frac{2}{\sqrt{14}\sqrt{17}} = \frac{2}{\sqrt{238}} \approx 0.1296, \qquad \theta \approx 82.6°.$$

LIVE: see the angle between these two vectors ▸.

IS THE COSINE LEGAL?

$\cos\theta$ must lie in $[-1, 1]$. If $|\mathbf{u} \cdot \mathbf{v}| > \|\mathbf{u}\|\|\mathbf{v}\|$ you have made an arithmetic slip. A small positive dot product means an angle just under $90°$, a negative one means an obtuse angle.

$\mathbf{u}^\mathsf{T}\mathbf{v}$ IS NOT $\mathbf{u}\mathbf{v}^\mathsf{T}$

Swapping the transpose swaps the shapes. $\mathbf{u}^\mathsf{T}\mathbf{v}$ is $(1\times n)(n\times 1)$, a *number*; $\mathbf{u}\mathbf{v}^\mathsf{T}$ is $(n\times 1)(1\times n)$, an $n \times n$ *matrix*:

$$\begin{bmatrix}1\\2\end{bmatrix}^\mathsf{T}\begin{bmatrix}3\\4\end{bmatrix} = 11, \qquad \begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix}^\mathsf{T} = \begin{bmatrix}3&4\\6&8\end{bmatrix}.$$

Likewise two column vectors cannot be multiplied as they stand ($(n \times 1)(n \times 1)$ has no match): you must transpose one of them.

THE VECTOR PRODUCT

The *vector product* $\mathbf{u} \times \mathbf{v}$ exists only in $\mathbb{R}^3$ and is not used in this module. In this guide the dot product is the only product of two vectors.

THE CHAPTER AT A GLANCE

SHAPE	$m \times n$ is rows $\times$ columns. $A + B$ needs equal shapes; AB needs the inner numbers to agree, and $(m \times n)(n \times p) = (m \times p)$.

PRODUCT	$(AB)_{i,j} = a_{i,-} \cdot b_{-,j}$, a table of dot products. $AB \neq BA$: a matrix is an instruction, and order matters.
IDENTITY	$IM = MI = M$, $0M = M0 = 0$; one of each per shape.
TRANSPOSE	$(A^\mathsf{T})_{i,j} = a_{j,i}$, and $(AB)^\mathsf{T} = B^\mathsf{T}A^\mathsf{T}$.
VECTORS	$\mathbb{R}^n$ is the $n \times 1$ matrices; $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\mathsf{T}\mathbf{v}$ is a number; $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$; $\mathbf{u} \cdot \mathbf{v} = 0$ means perpendicular.

TRY IT YOURSELF

1. Let $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ -1 & 4 \end{bmatrix}$. Compute AB and BA, state their shapes, and check that $(AB)^\mathsf{T} = B^\mathsf{T}A^\mathsf{T}$.
2. C is 3×2, D is 2×4, E is 4×1. Which of CD, DC, CDE, $E^\mathsf{T}D^\mathsf{T}C^\mathsf{T}$ are defined? Give the shapes.
3. Show that $\mathbf{u} = (1, -2, 2)^\mathsf{T}$ and $\mathbf{v} = (2, 2, 1)^\mathsf{T}$ are perpendicular. Find t such that $(1, t, 2)^\mathsf{T}$ is perpendicular to $(3, 1, -1)^\mathsf{T}$.
4. Find the angle between $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$.

ANSWERS

1. AB is 2×2: $\begin{bmatrix} 5 & -2 \\ 1 & 11 \end{bmatrix}$. BA is 3×3: $\begin{bmatrix} 2 & 4 & -2 \\ 1 & 5 & 1 \\ -1 & 10 & 9 \end{bmatrix}$. $(AB)^\mathsf{T} = \begin{bmatrix} 5 & 1 \\ -2 & 11 \end{bmatrix} = B^\mathsf{T}A^\mathsf{T}$.
2. CD: 3×4. DC: not defined ($4 \neq 3$). CDE: 3×1. $E^\mathsf{T}D^\mathsf{T}C^\mathsf{T}$: $(1 \times 4)(4 \times 2)(2 \times 3) = 1 \times 3$, the transpose of CDE.
3. $\mathbf{u} \cdot \mathbf{v} = 2 - 4 + 2 = 0$. For the second: $3 + t - 2 = 0$, so $t = -1$.
4. $\cos\theta = \dfrac{1}{\sqrt{2} \cdot 1}$, so $\theta = 45°$.

CHAPTER 2

Determinants

LECTURE 1 · *one number that records scaling, orientation and collapse*

THIS CHAPTER TEACHES HOW TO

- compute a 2×2 determinant, and a 3×3 one by minors *and* by Sarrus;
- read off from $\operatorname{Det} A$ the area (or volume) factor, the orientation, and whether A collapses space;
- use $\operatorname{Det}(A^{\mathsf{T}}) = \operatorname{Det} A$ and $\operatorname{Det}(AB) = \operatorname{Det} A \operatorname{Det} B$;
- say when a determinant does not exist.

2.1 What the determinant tells you

WHAT $\operatorname{Det} A$ IS FOR

To a *square* matrix A we attach one number $\operatorname{Det} A$.

- $|\operatorname{Det} A|$ is the factor by which A scales area (in 2D) or volume (in 3D and up).
- The sign of $\operatorname{Det} A$ records whether A preserves $(+)$ or reverses $(-)$ orientation. Swapping two rows or two columns flips the sign.
- $\operatorname{Det} A = 0$ exactly when A collapses space onto something flatter (in 2D: the columns are parallel). Chapter 3 shows that this is exactly the case when A has no inverse.

NON-SQUARE MATRICES

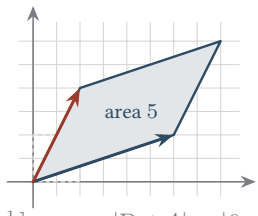
A matrix that is not square has *no* determinant at all. It is not zero, it is undefined. If a question asks for Det of a 2×3 matrix, the answer is "not defined".

2.2 Computing it

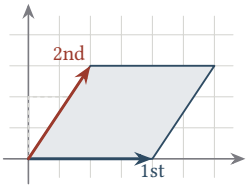
THE 2×2 DETERMINANT

$$\operatorname{Det} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

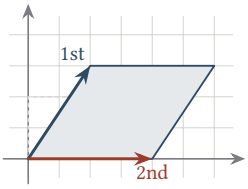
The columns $\left[\begin{smallmatrix} a \\ c \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} b \\ d \end{smallmatrix}\right]$ span a parallelogram of area $|ad - bc|$; the unit square has area 1, so this is the area factor.



$A = \left[\begin{smallmatrix} 3 & 1 \\ 1 & 2 \end{smallmatrix}\right]$: area $= |\operatorname{Det} A| = |6 - 1| = 5$.



Det $= +3$
1st to 2nd: counterclockwise



Det $= -3$
1st to 2nd: clockwise

THE 3×3 DETERMINANT, BY MINORS

Write A_{1j} for A with row 1 and column j deleted. Expanding along the first row, with alternating signs $+, -, +, \ldots$,

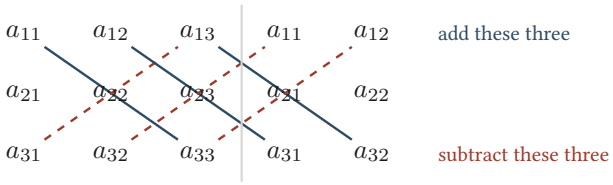
$$\operatorname{Det} A \;=\; a_{1,1} \operatorname{Det} A_{11} - a_{1,2} \operatorname{Det} A_{12} + a_{1,3} \operatorname{Det} A_{13} \quad (\text{in general } \textstyle\sum_j (-1)^{1+j} a_{1,j} \operatorname{Det} A_{1j}).$$

A large determinant is built from smaller ones; the 2×2 formula is where the recursion stops. The cost grows fast: a 4×4 needs four 3×3s.

THE 3×3 DETERMINANT, BY SARRUS

Copy the first two columns to the right of the matrix. Add the products down the three solid diagonals ($\searrow$) and subtract the products up the three dashed ones ($\nearrow$):

$$\operatorname{Det} A = \big(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}\big) - \big(a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}\big).$$



HOW TO COMPUTE A 3×3 DETERMINANT

1. Choose a method. Sarrus is quicker; minors is the one that generalises.
2. Write out every product before adding. Most errors are lost minus signs, so write the signs $+, -, +$ (minors) or *add down, subtract up* (Sarrus) *first.*
3. If time allows, compute it a second way. The two answers must agree.

SARRUS IS FOR 3×3 ONLY

The trick is false for 4×4 matrices: it would give six terms where the true determinant has twenty-four. For 4×4 and bigger, use minors (or the methods of later chapters).

EXAMPLE 2.1 · a 3×3 determinant, two ways

Compute $\operatorname{Det} M$ for $M = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$.

SOLUTION

Minors, first row:

$$\operatorname{Det} M = 2\begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 0 & 4 \end{vmatrix} + 0\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} = 2(12-1) - 1(4-0) + 0 = 22 - 4 = 18.$$

Sarrus: down: $2 \cdot 3 \cdot 4 + 1 \cdot 1 \cdot 0 + 0 \cdot 1 \cdot 1 = 24$; up: $0 \cdot 3 \cdot 0 + 2 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 4 = 6$. So $\operatorname{Det} M = 24 - 6 = 18$.

ANIMATION: expansion along the first row ▸; the rule of Sarrus ▸.

EXAMPLE 2.2 · reading a determinant

$\mathbf{u} = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$ are perpendicular, since $\mathbf{u} \cdot \mathbf{v} = -2 + 2 = 0$. What is the area of the parallelogram they span?

SOLUTION

$\left|\operatorname{Det}\left[\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right]\right| = |4 + 1| = 5$. As a check: the sides have length $\sqrt{5}$ and, being perpendicular, the parallelogram is a rectangle of area $\sqrt{5} \cdot \sqrt{5} = 5$. The determinant is positive, so $\mathbf{u}$ to $\mathbf{v}$ turns counterclockwise.

ANIMATION: the parallelogram and its area ▸.

AREA IS POSITIVE, DETERMINANT NEED NOT BE

The area is $|\operatorname{Det} A|$. A negative determinant does not mean negative area: it means orientation is reversed.

2.3 Two rules, and what follows

RULES FOR DETERMINANTS

For square matrices A, B of the same size:

$$\operatorname{Det}(A^{\mathsf{T}}) = \operatorname{Det} A, \qquad \operatorname{Det}(AB) = \operatorname{Det} A \operatorname{Det} B.$$

Consequently: anything true of rows is true of columns (transpose); areas multiply when maps are composed; and $\operatorname{Det}(AB) = \operatorname{Det}(BA)$ *even though* $AB \neq BA$ in general.

EXAMPLE 2.3 · $\operatorname{Det}(AB) = \operatorname{Det} A \operatorname{Det} B$

$A = \left[\begin{smallmatrix} 3 & 1 \\ 5 & 2 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right]$, $AB = \left[\begin{smallmatrix} 7 & 1 \\ 13 & 2 \end{smallmatrix}\right]$. Then $\operatorname{Det} A = 1$, $\operatorname{Det} B = 1$ and $\operatorname{Det}(AB) = 14 - 13 = 1 = 1 \cdot 1$.

QUICK CHECKS

- If one column is a multiple of another, expect $\operatorname{Det} = 0$ (parallel columns).
- Swapping two columns must only flip the sign of your answer.
- $\operatorname{Det}(A + B)$ is *not* $\operatorname{Det} A + \operatorname{Det} B$; only products are well behaved.

THE CHAPTER AT A GLANCE

EXISTENCE	Square matrices only. Non-square: undefined, not zero.
MEANING	$\lvert\operatorname{Det} A\rvert$ is the area/volume factor; the sign is orientation; $\operatorname{Det} A = 0$ means collapse.
2×2	$\operatorname{Det}\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = ad - bc$.
3×3	Minors along row 1 with signs $+, -, +$; or Sarrus (3×3 *only*).
RULES	$\operatorname{Det}(A^{\mathsf{T}}) = \operatorname{Det} A$ and $\operatorname{Det}(AB) = \operatorname{Det} A \operatorname{Det} B$.

TRY IT YOURSELF

1. Compute $\operatorname{Det}\left[\begin{smallmatrix} 4 & -1 \\ 2 & 3 \end{smallmatrix}\right]$ and $\operatorname{Det}\left[\begin{smallmatrix} 6 & 3 \\ 4 & 2 \end{smallmatrix}\right]$. Which matrix collapses the plane onto a line?
2. Compute $\operatorname{Det} N$ for $N = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ 2 & 4 & 0 \end{bmatrix}$ by minors and by Sarrus.
3. A and B are 3×3 with $\operatorname{Det} A = 2$ and $\operatorname{Det} B = -3$. Find $\operatorname{Det}(AB)$, $\operatorname{Det}(BA^{\mathsf{T}})$ and $\operatorname{Det}(A^{\mathsf{T}}B^{\mathsf{T}})$. Is AB a collapse?
4. Find the area of the parallelogram spanned by $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$, and say whether the first-to-second turn is clockwise or counterclockwise.
5. Can you compute $\operatorname{Det}\left[\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{smallmatrix}\right]$?

ANSWERS

1. $12 + 2 = 14$ and $12 - 12 = 0$; the second matrix collapses the plane (its columns $\left[\begin{smallmatrix} 6 \\ 4 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 3 \\ 2 \end{smallmatrix}\right]$ are parallel).
2. Minors: $1(0 + 4) - 0 + 2(12 - 2) = 24$. Sarrus: down $0 + 0 + 24 = 24$, up $4 - 4 + 0 = 0$, so 24.
3. $\operatorname{Det}(AB) = -6$; $\operatorname{Det}(BA^{\mathsf{T}}) = (-3)(2) = -6$; $\operatorname{Det}(A^{\mathsf{T}}B^{\mathsf{T}}) = -6$. Not a collapse ($\operatorname{Det} \neq 0$); orientation is reversed (negative sign).
4. $\operatorname{Det}\left[\begin{smallmatrix} 3 & -1 \\ 1 & 2 \end{smallmatrix}\right] = 6 + 1 = 7$: area 7, positive, so counterclockwise.
5. No: the matrix is 2×3, not square, so the determinant is not defined.

CHAPTER 3

Inverses and row reduction

LECTURES 2 AND 8 · *inverting a matrix, and the algorithm that does it*

THIS CHAPTER TEACHES HOW TO

- write down the inverse of a 2×2 matrix, or say why none exists;
- invert a 3×3 matrix by Gauss–Jordan elimination, and recognise from the working when the matrix is not invertible;
- understand $(AB)^{-1} = B^{-1}A^{-1}$, $(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}$ and $\operatorname{Det}(A^{-1}) = 1/\operatorname{Det} A$;
- recognise A invertible $\iff \operatorname{Det} A \neq 0$, and understand the three row operations.

3.1 What an inverse is

INVERSE

A square matrix A is **invertible** if there is a matrix A^{-1} with

$$AA^{-1} = A^{-1}A = I.$$

The inverse, when it exists, is *unique*, so we say *the* inverse. Only square matrices can have one.

THE 2×2 FORMULA

If $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ has $\operatorname{Det} A = ad - bc \neq 0$, then

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

Swap the diagonal, negate the off-diagonal, divide by the determinant. If $ad - bc = 0$ there is no inverse.

INVERSION CRITERION

A square matrix A is invertible $\iff \operatorname{Det} A \neq 0$. In that case

$$\operatorname{Det}(A^{-1}) \;=\; \frac{1}{\operatorname{Det} A}.$$

(Reason in one line: $\operatorname{Det} A \operatorname{Det} A^{-1} = \operatorname{Det} I = 1$.) This also explains the geometry: if A multiplies area by $\operatorname{Det} A$, then A^{-1} must divide it back.

RULES OF INVERSION

If A and B are invertible $n \times n$ matrices, then so are AB and A^{T}, and

$$(AB)^{-1} = B^{-1}A^{-1}, \qquad (A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}.$$

"Socks and shoes": to undo *socks, then shoes*, take off the shoes first.

FOUR CLASSIC MISTAKES

- $(AB)^{-1} = A^{-1}B^{-1}$ is wrong: the order reverses.
- $(A+B)^{-1}$ is *not* $A^{-1} + B^{-1}$. (Take $A = I$, $B = -I$: $A + B = 0$ has no inverse at all, so the sum of two invertible matrices need not be invertible.)
- Never write $\dfrac{B}{A}$. Because $A^{-1}B$ and BA^{-1} are different matrices, there is no "division"; write $A^{-1}B$ or BA^{-1}.
- Only *square* matrices are invertible, and $\operatorname{Det} A \neq 0$ is the test, not "no zero entries".

EXAMPLE 3.1 · 2×2 inverses

Invert $A = \left[\begin{smallmatrix} 3 & 1 \\ 5 & 2 \end{smallmatrix}\right]$ and $C = \left[\begin{smallmatrix} 4 & 2 \\ 3 & 1 \end{smallmatrix}\right]$.

SOLUTION

$\operatorname{Det} A = 6 - 5 = 1$, so $A^{-1} = \left[\begin{smallmatrix} 2 & -1 \\ -5 & 3 \end{smallmatrix}\right]$. Check: $AA^{-1} = \left[\begin{smallmatrix} 6-5 & -3+3 \\ 10-10 & -5+6 \end{smallmatrix}\right] = I$. $\operatorname{Det} C = 4 - 6 = -2 \neq 0$, so $C^{-1} = \dfrac{1}{-2}\left[\begin{smallmatrix} 1 & -2 \\ -3 & 4 \end{smallmatrix}\right] = \begin{bmatrix} -\frac{1}{2} & 1 \\ \frac{3}{2} & -2 \end{bmatrix}$.

EXAMPLE 3.2 · using the rules

With $A = \left[\begin{smallmatrix} 3 & 1 \\ 5 & 2 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right]$ we have $B^{-1} = \left[\begin{smallmatrix} 1 & 0 \\ -4 & 1 \end{smallmatrix}\right]$, and

$$(AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -13 & 7 \end{bmatrix}.$$

Directly: $AB = \left[\begin{smallmatrix} 7 & 1 \\ 13 & 2 \end{smallmatrix}\right]$ has determinant $14 - 13 = 1$, so its inverse is $\left[\begin{smallmatrix} 2 & -1 \\ -13 & 7 \end{smallmatrix}\right]$.

3.2 Row operations and the Gauss–Jordan algorithm

THE THREE ELEMENTARY ROW OPERATIONS

1. **Swap** two rows: $R_i \leftrightarrow R_j$.
2. **Scale** a row by a non-zero number: $R_i \leftarrow \lambda R_i$, $\lambda \neq 0$.
3. **Add** a multiple of one row to another: $R_i \leftarrow R_i + \lambda R_j$, $i \neq j$.

Each is reversible, so none destroys information; each is multiplication by an invertible matrix. Scaling by $\lambda = 0$ is not allowed: it cannot be undone.

GAUSS–JORDAN

Write A and I side by side and apply the *same* row operations to both until the left block is I:

$$[\,A \mid I\,] \xrightarrow{\text{row operations}} [\,I \mid A^{-1}\,].$$

If the left block cannot be brought to I then A is not invertible. This happens exactly when $\operatorname{Det} A = 0$.

HOW TO INVERT A MATRIX BY GAUSS–JORDAN

1. Write the augmented matrix $[A \mid I]$.
2. Work column by column, left to right. Get a non-zero entry into the pivot position (swap rows if it holds a zero), and clear the entries below it.
3. Scale each pivot row so that its pivot is 1, then clear the entries *above* each pivot, working from the last column back to the first.
4. **Stop** if a row of the *left* block becomes entirely zero: that is a definite proof that A has no inverse. (A zero in a pivot position is not yet proof: try a row swap first.)
5. Otherwise read A^{-1} off the right block, and check $AA^{-1} = I$.

EXAMPLE 3.3 · Gauss–Jordan on a 3×3 matrix

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

SOLUTION

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow \frac{1}{2} R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 - R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

$$\xrightarrow{R_1 \leftarrow R_1 - R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right] \qquad A^{-1} = \tfrac{1}{2} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Check: the first row of A times A^{-1} gives $\frac{1}{2}(1+1,\ -1+1,\ 1-1) = (1, 0, 0)$. Also $\operatorname{Det} A = 1(1-0) - 1(0-1) + 0 = 2$, and $\operatorname{Det} A^{-1} = \frac{1}{8} \cdot 4 = \frac{1}{2} = 1/\operatorname{Det} A$.

ANIMATION: watch Gauss–Jordan invert this matrix ▸.

EXAMPLE 3.4 · a matrix with no inverse

Show that $C = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ is not invertible.

SOLUTION

The first pivot is 2; swap $R_1 \leftrightarrow R_3$ to get a leading 1 at once.

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 2 & 0 & 1 & 1 & 0 & 0 \end{array}\right] \xrightarrow[R_3 \leftarrow R_3 - 2R_1]{R_2 \leftarrow R_2 - R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 1 & -1 \\ 0 & -2 & -1 & 1 & 0 & -2 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 & -3 \end{array}\right]$$

The third row of the left block is entirely zero, so the left block can never become I: C is not invertible. Consistently, $\operatorname{Det} C = 2(3-2) - 0 + 1(1-3) = 0$.

ANIMATION: watch the reduction fail ▸.

CHECKS BEFORE YOU HAND IN

- Multiply: AA^{-1} should be I. Checking one row or one entry catches most slips.
- If A has integer entries and $\operatorname{Det} A = \pm 1$, then A^{-1} has integer entries (as for $\left[\begin{smallmatrix} 3 & 1 \\ 5 & 2 \end{smallmatrix}\right]$); fractions with denominator d usually point to $\operatorname{Det} A = \pm d$.
- If you have a nonzero determinant but your row reduction produced a zero row (or vice versa), one of them contains an arithmetic error.

3.3 Echelon forms (Lecture 8)

REF AND RREF

The first non-zero entry of a non-zero row is its **pivot**. A matrix is in

- **row echelon form** (REF) if every zero row is at the bottom and each pivot lies strictly to the right of the pivot in the row above (a staircase, with zeros below it);
- **reduced row echelon form** (RREF) if moreover every pivot is 1 and is the *only* non-zero entry in its column (zeros above the pivots too).

Columns containing a pivot are **pivot columns**; the others are **non-pivot** columns, and these may contain anything. Every matrix can be brought to RREF by elementary row operations, and the RREF is unique: $\operatorname{RREF}(A)$. A REF is *not* unique (rescale any row). Gauss–Jordan on $[A \mid I]$ was exactly this reduction.

EXAMPLE 3.5 · REF, RREF or neither?

$\left[\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 4 & 1 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 & 1 & 0 & 3 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right]$.

SOLUTION

First: neither (a zero row sits above a non-zero row, and the last pivot is left of the one above). Second: REF, not RREF (the pivots 2, 5 are not 1, and there is a 1 above the second pivot). Third: RREF (pivots 1, alone in their columns).

WHAT ROW OPERATIONS PRESERVE

Row operations rewrite the equations without changing the solutions, so they preserve $\operatorname{Ker} A$, i.e. $\operatorname{Ker} A = \operatorname{Ker} \operatorname{RREF}(A)$, and which columns are redundant. They do not preserve the column space: $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ and its reduction $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ span different lines. Use row reduction on $\operatorname{Ker} A$ directly, and on $\operatorname{Col} A$ only to ask whether $A\mathbf{x} = \mathbf{b}$ is solvable (Chapter 8).

THE CHAPTER AT A GLANCE

INVERSE	$AA^{-1} = A^{-1}A = I$; unique; square matrices only; no division, so write $A^{-1}B$ or BA^{-1}.
2×2	$A^{-1} = \dfrac{1}{ad - bc}\left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$.
CRITERION	A invertible $\iff \operatorname{Det} A \neq 0$; then $\operatorname{Det}(A^{-1}) = 1/\operatorname{Det} A$.
RULES	$(AB)^{-1} = B^{-1}A^{-1}$; $(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}$.
ROW OPS	Swap; scale by $\lambda \neq 0$; add a multiple of one row to another. All reversible.
ALGORITHM	$[A \mid I] \to [I \mid A^{-1}]$; a zero row on the left means $\operatorname{Det} A = 0$ and no inverse.
ECHELON	REF: staircase of pivots, zero rows last. RREF: pivots 1, alone in their column; unique. Row operations keep $\operatorname{Ker} A$ but not $\operatorname{Col} A$.

TRY IT YOURSELF

1. Invert $\left[\begin{smallmatrix}5 & 2\\ 7 & 3\end{smallmatrix}\right]$, or explain why no inverse exists. Do the same for $\left[\begin{smallmatrix}2 & 4\\ 1 & 2\end{smallmatrix}\right]$.
2. Invert $B = \begin{bmatrix}2 & 1 & 0\\ 0 & 1 & 1\\ 1 & 0 & 1\end{bmatrix}$ by Gauss–Jordan, and check your answer.
3. Use row operations to show that $D = \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix}$ is not invertible. Confirm with the determinant.
4. A, B are 3×3 with $\operatorname{Det} A = 5$ and $\operatorname{Det} B = -2$. Find $\operatorname{Det}(A^{-1})$, $\operatorname{Det}\big((AB)^{-1}\big)$ and $\operatorname{Det}(A^{\mathsf{T}}B^{-1})$.
5. True or false (justify): *if A and B are invertible, so is $A + B$.*

ANSWERS

1. $\operatorname{Det} = 15 - 14 = 1$: $\left[\begin{smallmatrix}3 & -2\\ -7 & 5\end{smallmatrix}\right]$. The second has $\operatorname{Det} = 4 - 4 = 0$: no inverse.
2. $\operatorname{Det} B = 3 \neq 0$, and $B^{-1} = \frac{1}{3}\begin{bmatrix}1 & -1 & 1\\ 1 & 2 & -2\\ -1 & 1 & 2\end{bmatrix}$.
3. $R_2 \leftarrow R_2 - 4R_1$, $R_3 \leftarrow R_3 - 7R_1$ give rows $(0, -3, -6)$ and $(0, -6, -12)$; then $R_3 \leftarrow R_3 - 2R_2$ gives a zero row. $\operatorname{Det} D = 0$ (e.g. by Sarrus: $45 + 84 + 96 - 105 - 48 - 72 = 0$).
4. $\frac{1}{5}$; $\dfrac{1}{5 \cdot (-2)} = -\frac{1}{10}$; $\operatorname{Det}(A^{\mathsf{T}}B^{-1}) = 5 \cdot (-\frac{1}{2}) = -\frac{5}{2}$.
5. False. $A = I$ and $B = -I$ are invertible but $A + B = 0$ is not.

PART II

Vector spaces

CHAPTER 4

Vector spaces

LECTURES 2–3 · *the nine axioms, the standard examples, and how to spot a non-example*

THIS CHAPTER TEACHES HOW TO

- list the nine axioms and say what each one demands;
- name the standard vector spaces and their operations ($\mathbb{R}^n$, matrices, polynomials, functions, $\mathbb{C}$ over $\mathbb{R}$);
- decide whether a set with given operations is a vector space, and if not *name the axiom that fails* with an explicit counterexample;
- avoid the usual near-misses: degree *exactly* n, lines missing the origin, redefined scalings.

4.1 The definition

VECTOR SPACE OVER $\mathbb{R}$

A **vector space** is a set V with an addition and a scaling $\mathbb{R}\times V\to V$ such that, for all $\mathbf{u},\mathbf{v},\mathbf{w}\in V$ and $\lambda,\mu\in\mathbb{R}$:

(A1)	$\mathbf{u}+\mathbf{v}\in V$	closure under addition
(A2)	$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$	associativity
(A3)	there is $\mathbf{0}\in V$ with $\mathbf{v}+\mathbf{0}=\mathbf{v}$	neutral element
(A4)	each $\mathbf{v}$ has $-\mathbf{v}$ with $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$	inverses
(A5)	$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$	commutativity
(S0)	$\lambda\mathbf{u}\in V$	closure under scaling
(S1)	$\lambda(\mathbf{u}+\mathbf{v})=\lambda\mathbf{u}+\lambda\mathbf{v}$	distributivity in V
(S2)	$(\lambda+\mu)\mathbf{v}=\lambda\mathbf{v}+\mu\mathbf{v}$	distributivity in $\mathbb{R}$
(S3)	$(\lambda\mu)\mathbf{v}=\lambda(\mu\mathbf{v})$	compatibility
(S4)	$1\,\mathbf{v}=\mathbf{v}$	the unit acts trivially

(A1)–(A5) say $(V,+)$ is an **abelian group**; (S0)–(S4) govern the scaling.

CONSEQUENCES OF THE AXIOMS

In every vector space

$$0\,\mathbf{v}=\mathbf{0},\qquad \lambda\mathbf{0}=\mathbf{0},\qquad (-1)\mathbf{v}=-\mathbf{v},$$

and the zero vector is unique. So if a candidate fails any one of these, it is not a vector space.

WHAT THE AXIOMS DO *NOT* MENTION

There is no product of two vectors, no division, no length and no angle. The dot product returns a number, not a vector; the vector product exists only in $\mathbb{R}^3$. A vector space remembers only $+$ and scaling by real numbers, which is why one theorem can serve $\mathbb{R}^n$, matrices, polynomials and functions at once.

4.2 The standard examples

Space	Elements	Operations
$\mathbb{R}^n$	columns $(v_1, \dots, v_n)^{\mathsf{T}}$	entrywise
$\mathrm{Mat}_{m\times n}(\mathbb{R})$	$m \times n$ matrices	entrywise (no matrix product needed)
P_n	polynomials of degree *at most* n	coefficientwise
$\mathbb{R}[x]$	all polynomials	coefficientwise
$\mathcal{F}(\mathbb{R}, \mathbb{R})$	all functions $\mathbb{R} \to \mathbb{R}$	pointwise: $(f+g)(x) = f(x)+g(x)$
$\mathcal{C}(\mathbb{R})$, $\mathcal{C}^1(\mathbb{R})$, $\mathcal{C}^\infty(\mathbb{R})$	continuous / once / infinitely differentiable	pointwise
$\mathbb{R}[x, x^{-1}]$	Laurent polynomials	coefficientwise
$\mathbb{C}$ over $\mathbb{R}$	$a + bi$	usual $+$; scale by *real* numbers only

ZERO VECTORS AND NEGATIVES IN THE STANDARD EXAMPLES

In $\mathbb{R}^n$ the zero vector is $(0, \dots, 0)^{\mathsf{T}}$; in $\mathrm{Mat}_{m\times n}(\mathbb{R})$ it is the zero matrix; in P_n and $\mathbb{R}[x]$ it is the polynomial 0; in $\mathcal{F}(\mathbb{R}, \mathbb{R})$ it is the constant function 0 ("the x-axis"). The negative of $\mathbf{v}$ is $(-1)\mathbf{v}$.

WHY THE STANDARD EXAMPLES WORK

Everything is checked *coordinatewise* (or *pointwise* for functions), and each axiom then reduces to the same axiom for real numbers. For instance (S1) in $\mathbb{R}^n$: coordinate i of $\lambda(\mathbf{u} + \mathbf{v})$ is $\lambda(u_i + v_i) = \lambda u_i + \lambda v_i$, which is coordinate i of $\lambda\mathbf{u} + \lambda\mathbf{v}$.

$\mathbb{C}$ AS A REAL VECTOR SPACE

Every element of $\mathbb{C}$ is $a \cdot 1 + b \cdot i$, and $\mathbb{C}$ behaves as a copy of $\mathbb{R}^2$ (addition is the parallelogram rule). The complex *product*, e.g. $(2 + 3i)(1 - i) = 5 + i$, is extra structure that the axioms never use.

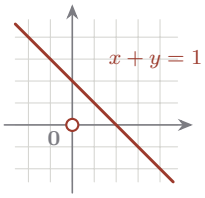
4.3 Deciding whether something is a vector space

HOW TO DECIDE: "IS V A VECTOR SPACE?"

1. **Is V a subset of a vector space you know, with the same $+$ and scaling?** Then do not check nine axioms. Use the three-condition subspace test of Chapter 5.
2. **Otherwise** (new operations, or a new kind of object), go hunting for a failure, in this order: is there a zero vector? is the set closed under $+$ and scaling? does $1\,\mathbf{v} = \mathbf{v}$? do the two distributive laws hold?
3. **To say "no":** name *one* axiom and give *one* explicit numerical counterexample. That is a complete answer.
4. **To say "yes":** verify every axiom for *general* $\mathbf{u}, \mathbf{v}, \lambda, \mu$ (examples prove nothing), or reduce them coordinatewise to the real numbers.

NEAR MISSES THAT COME UP EVERY YEAR

- **Degree *exactly* n.** $\{a + bx \mid b \neq 0\}$ is not a vector space: 0 has not degree 1 (A3 fails) and $(1 + x) + (1 - x) = 2$ leaves the set (A1 fails). Degree *at most* n is fine.
- **A line missing the origin.** $\{(x, y) \mid x + y = 1\}$: no zero vector, and $(1, 0) + (0, 1) = (1, 1)$ leaves the set.
- **The non-negative reals** $\mathbb{R}_{\geq 0}$ (or the first quadrant): closed under $+$, but $(-1) \cdot 2 = -2$ escapes (S0 and A4 fail).
- **Redefined scaling.** Everything else may be fine and one axiom, often (S4) or (S2), fails. Always test $1 \cdot \mathbf{v}$ and $(\lambda + \mu)\mathbf{v}$.



A line that misses the origin is not a vector space.

EXAMPLE 4.1 · redefining the scaling on $\mathbb{R}^2$

Keep the usual addition on $\mathbb{R}^2$. For each scaling $\odot$, is $\mathbb{R}^2$ a vector space? (a) $\lambda \odot (x, y) = (\lambda x, 0)$; (b) $\lambda \odot (x, y) = (\lambda x, y)$; (c) $\lambda \odot (x, y) = (\lambda x, \lambda y + 1)$.

SOLUTION

(a) No. (S4) fails: $1 \odot (1, 1) = (1, 0) \neq (1, 1)$.

(b) No. (S2) fails: for $\lambda = \mu = 1$ and $\mathbf{v} = (0, 1)$, $(\lambda + \mu) \odot \mathbf{v} = 2 \odot (0, 1) = (0, 1)$, but $\lambda \odot \mathbf{v} + \mu \odot \mathbf{v} = (0, 1) + (0, 1) = (0, 2)$.

(c) No. (S4) fails: $1 \odot (x, y) = (x, y + 1) \neq (x, y)$.

EXAMPLE 4.2 · degree at most 2 versus exactly 2

Is $D = \{a + bx + cx^2 \mid c \neq 0\}$ a vector space?

SOLUTION

No. (A1) fails: $(x^2 + x + 1) + (-x^2) = x + 1$, which has degree 1 and is not in D. (A3) fails too: the zero polynomial is not in D. The space P_2 of polynomials of degree *at most* 2 *is* a vector space.

THREE INSTANT TESTS

Before anything else ask: (i) is $\mathbf{0}$ in the set? (ii) is $-\mathbf{v} = (-1)\mathbf{v}$ still in the set? (iii) does $0 \cdot \mathbf{v}$ give the zero vector, and $1 \cdot \mathbf{v}$ give $\mathbf{v}$? A "no" to any of these ends the question.

THE CHAPTER AT A GLANCE

DEFINITION	Nine axioms: (A1)–(A5) make $(V, +)$ an abelian group, (S0)–(S4) govern scaling by $\mathbb{R}$.
CONSEQUENCES	$0\mathbf{v} = \mathbf{0}$, $\lambda\mathbf{0} = \mathbf{0}$, $(-1)\mathbf{v} = -\mathbf{v}$; the zero vector is unique.
EXAMPLES	$\mathbb{R}^n$, $\mathrm{Mat}_{m\times n}(\mathbb{R})$, P_n, $\mathbb{R}[x]$, $\mathcal{F}(\mathbb{R},\mathbb{R})$, $\mathcal{C}(\mathbb{R})$, $\mathcal{C}^1(\mathbb{R})$, $\mathcal{C}^\infty(\mathbb{R})$, $\mathbb{R}[x, x^{-1}]$, and $\mathbb{C}$ over $\mathbb{R}$.
METHOD	Subset of a known space? Use the subspace test. New operations? Hunt for the failing axiom. "No" needs one counterexample; "yes" needs a general argument.
NEAR MISSES	Degree *exactly* n; a line missing the origin; $\mathbb{R}_{\geq 0}$; a scaling with $1 \odot \mathbf{v} \neq \mathbf{v}$.

TRY IT YOURSELF

1. Keep the usual addition on $\mathbb{R}^2$ and put $\lambda \odot (x, y) = (\lambda^2 x, \lambda^2 y)$. Is this a vector space? Justify.
2. With the usual operations, is the set of invertible 2×2 matrices a vector space?
3. With the usual operations, decide each of the following (justify): (a) P_3; (b) polynomials of degree exactly 3; (c) functions $f : \mathbb{R} \to \mathbb{R}$ with $f(x) \geq 0$ for all x; (d) $\{(x, y) \in \mathbb{R}^2 \mid y = 2x\}$.

ANSWERS

1. No. (S2) fails: $(1 + 1) \odot (1, 0) = (4, 0)$, but $1 \odot (1, 0) + 1 \odot (1, 0) = (2, 0)$. (S4 does hold, so one must test further axioms to find the failure.)
2. No. (A3) fails: the zero matrix is not invertible. Also (A1): $I + (-I) = 0$.
3. (a) Yes (P_3 is the polynomials of degree *at most* 3). (b) No: $0 \notin$, and $x^3 + (-x^3) = 0$ leaves the set. (c) No: $f(x) = x^2$ is in the set, but $(-1)f$ takes negative values (S0 fails). (d) Yes: a line through the origin (see Chapter 5).

CHAPTER 5

Subspaces

LECTURES 3–4 · *three checks instead of nine, and how to apply them quickly*

THIS CHAPTER TEACHES HOW TO

- state the subspace test and write a complete, justified answer to "is U a subspace of V?";
- refute a non-subspace with one explicit counterexample, naming the failing condition;
- recognise at a glance if a set is defined by homogeneous linear conditions, or not;
- use the chain $\mathbb{R}[x] \subset \mathcal{C}^\infty(\mathbb{R}) \subset \mathcal{C}^1(\mathbb{R}) \subset \mathcal{C}(\mathbb{R}) \subset \mathcal{F}(\mathbb{R},\mathbb{R})$ and the two trivial subspaces.

5.1 The subspace test

SUBSPACE TEST

Let V be a vector space. A subset $U \subseteq V$ is a **subspace** if

(U1) $\mathbf{0} \in U$ — contains the zero vector (so is non-empty)

(U2) $\mathbf{u}, \mathbf{w} \in U \;\Rightarrow\; \mathbf{u}+\mathbf{w} \in U$ — closed under addition

(U3) $\lambda \in \mathbb{R},\ \mathbf{u} \in U \;\Rightarrow\; \lambda\mathbf{u} \in U$ — closed under scaling

Then U is itself a vector space (with the operations of V).

WHY THREE CONDITIONS ARE ENOUGH

The axioms (A2), (A5), (S1) – (S4) are *equations*: they hold in V, hence in any subset. Only the "leaving the set" axioms can fail: (A1), (A3), (A4), (S0). (U1) gives (A3), (U2) is (A1), (U3) is (S0), and (A4) follows from (U3) with $\lambda = -1$.

ALWAYS SUBSPACES

$\{\mathbf{0}\}$ and V itself are subspaces of every vector space V (the **trivial** ones). The standard function spaces form a chain of subspaces:

$$\mathbb{R}[x] \;\subset\; \mathcal{C}^\infty(\mathbb{R}) \;\subset\; \mathcal{C}^1(\mathbb{R}) \;\subset\; \mathcal{C}(\mathbb{R}) \;\subset\; \mathcal{F}(\mathbb{R},\mathbb{R}),$$

and $P_n \subseteq \mathbb{R}[x]$.

HOW TO ANSWER "IS U A SUBSPACE OF V?"

1. Name the set: $U = \{\text{elements of } V \mid \text{condition}\}$.
2. **(U1)** Does $\mathbf{0}$ satisfy the condition? If not, stop: not a subspace.
3. **(U2)** Take *arbitrary* $\mathbf{u}, \mathbf{w} \in U$, so the condition holds for each. Show that it holds for $\mathbf{u} + \mathbf{w}$.
4. **(U3)** Take arbitrary $\lambda \in \mathbb{R}$ and $\mathbf{u} \in U$. Show that it holds for $\lambda\mathbf{u}$.
5. **Conclude** in a sentence: "all three hold, so U is a subspace (and hence a vector space)".
6. **To refute,** do *not* write a general argument: give *one explicit* $\mathbf{u}, \mathbf{w}$ (or $\lambda, \mathbf{u}$) that are in U and show $\mathbf{u} + \mathbf{w}$ (or $\lambda\mathbf{u}$) is not. One failure is enough.

COMMON WAYS TO LOSE MARKS

- **Examples do not prove "yes".** Checking that two particular vectors add to something in U shows nothing; (U2) and (U3) are for *all* elements.
- **Counterexamples must lie in U.** Check that your $\mathbf{u}, \mathbf{w}$ really satisfy the defining condition before showing the sum does not.
- **(U1) alone is not enough.** $\{(x, y) \mid xy = 0\}$ contains $\mathbf{0}$ but fails (U2).
- **Stay inside V.** A subspace of $\mathbb{R}^3$ must consist of vectors of $\mathbb{R}^3$, with their three coordinates.

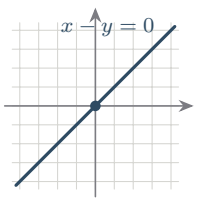
5.2 Recognising subspaces quickly

RULE OF THUMB FOR SUBSETS OF $\mathbb{R}^n$

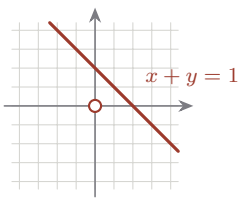
A subset of $\mathbb{R}^n$ defined by one or more *homogeneous linear equations* (expression where every term is exactly linear and equls $= 0$) is a subspace. For example, $x + y = 1 \iff x + y - 1 = 0$ is not homogenious since -1 has degree 0 and is thus not linear. Neither are $(x^2 = y^2, xy = z)$, as they contain degree two terms (note, xy has degree two). The inequality $(x \geq 0)$ might be homogenious and linear, but is not an equation.

PICTURES

In $\mathbb{R}^2$ the subspaces are just $\{\mathbf{0}\}$, lines *through the origin*, and $\mathbb{R}^2$; in $\mathbb{R}^3$, $\{\mathbf{0}\}$, lines and planes through the origin, and $\mathbb{R}^3$. A line or plane that misses the origin, or anything curved, will fail.



Yes. through $\mathbf{0}$



No. misses $\mathbf{0}$

5.3 Worked examples

EXAMPLE 5.1 · a full answer: a plane in $\mathbb{R}^3$

Is $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 0\}$ a subspace of $\mathbb{R}^3$?

SOLUTION

(U1) $0 + 2 \cdot 0 - 0 = 0$, so $\mathbf{0} \in W$.

(U2) Let $\mathbf{u} = (x_1, y_1, z_1)$ and $\mathbf{w} = (x_2, y_2, z_2)$ be in W, so $x_1 + 2y_1 - z_1 = 0$ and $x_2 + 2y_2 - z_2 = 0$. Then $(x_1 + x_2) + 2(y_1 + y_2) - (z_1 + z_2) = 0 + 0 = 0$, so $\mathbf{u} + \mathbf{w} \in W$.

(U3) For $\lambda \in \mathbb{R}$: $\lambda x_1 + 2\lambda y_1 - \lambda z_1 = \lambda \cdot 0 = 0$, so $\lambda\mathbf{u} \in W$.

All three hold, so W is a subspace of $\mathbb{R}^3$ (and hence a vector space).

ANIMATION: the three tests on this plane ▸.

EXAMPLE 5.2 · three ways to fail

Show that none of these is a subspace of the given space. (a) $\{(x, y, z) \mid x + y + z = 1\} \subseteq \mathbb{R}^3$; (b) $\{(x, y) \mid xy = 0\} \subseteq \mathbb{R}^2$; (c) $\{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \mid f(x) \geq 0 \text{ for all } x\}$.

SOLUTION

(a) $(0, 0, 0)$ gives $0 \neq 1$, so (U1) fails. (For the record, $(1, 0, 0)$ and $(0, 1, 0)$ lie in the set, but their sum $(1, 1, 0)$ has coordinate sum 2: (U2) fails as well.)

(b) $(1, 0)$ and $(0, 1)$ are in the set ($1 \cdot 0 = 0$) but their sum $(1, 1)$ has $1 \cdot 1 = 1 \neq 0$. (U2) fails. ($\mathbf{0}$ is in the set and (U3) holds: the set is the union of the two axes.)

(c) $f(x) = x^2$ is in the set, but $(-1)f$ takes the value -1 at $x = 1$: (U3) fails.

ANIMATION: part (a) failing the first test ▸.

EXAMPLE 5.3 · polynomials and functions

Which are subspaces? (a) $\{p \in P_2 \mid p(1) = 0\}$; (b) $\{p \in P_2 \mid p(0) = 1\}$; (c) $\{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \mid f(-x) = f(x) \text{ for all } x\}$.

SOLUTION

(a) Yes. The zero polynomial has $p(1) = 0$. If $p(1) = q(1) = 0$ then $(p+q)(1) = p(1)+q(1) = 0$ and $(\lambda p)(1) = \lambda p(1) = 0$. (Equivalently $a + b + c = 0$ for $p = a + bx + cx^2$: a homogeneous linear condition.)

(b) No. The zero polynomial has $0(0) = 0 \neq 1$: (U1) fails. (Also $(p + q)(0) = 2$.)

(c) Yes (the even functions). $0(-x) = 0 = 0(x)$; if f, g are even then $(f + g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f + g)(x)$, and $(\lambda f)(-x) = \lambda f(x)$.

EXAMPLE 5.4 · matrices

Which are subspaces of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$? (a) the symmetric matrices $A^\mathsf{T} = A$; (b) $\{A \mid a_{1,1} = 1\}$; (c) $\{A \mid \mathrm{Det}\, A = 0\}$.

SOLUTION

(a) Yes. $0^\mathsf{T} = 0$. If $A^\mathsf{T} = A$ and $B^\mathsf{T} = B$ then $(A+B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T} = A + B$ and $(\lambda A)^\mathsf{T} = \lambda A^\mathsf{T} = \lambda A$.

(b) No. The zero matrix has $a_{1,1} = 0$: (U1) fails.

(c) No. $\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$ both have determinant 0, but their sum is I, with $\mathrm{Det}\, I = 1$: (U2) fails. (Det is not linear, so a condition on Det is not a "linear condition".)

5.4 Simplify first, and combining subspaces

SIMPLIFY THE CONDITION *BEFORE* TESTING IT

A quadratic-looking description need not describe a curved set.

- $x^2 + y^2 + 1 = 1$ says $x^2 + y^2 = 0$, so $x = y = 0$: the set is $\{\mathbf{0}\}$. **Yes.**
- $x^2 + y^2 + z^2 \geq 0$ holds for every vector: the set is all of $\mathbb{R}^3$. **Yes.**
- $x^2 + y^2 \geq 0$ and $z = 0$: the first condition says nothing, leaving the plane $z = 0$, a subspace (a copy of $\mathbb{R}^2$). **Yes.**
- $x^2 + y^2 = 1$ (a circle) misses $\mathbf{0}$, so (U1) fails. **No.**
- $x + y \geq 0$, $z = 0$ (a half-plane): contains $\mathbf{0}$ and is closed under $+$, but $\mathbf{u} = (1,0,0)^\mathsf{T}$ is in it and $-\mathbf{u}$ is not, so (U3) fails. A genuine inequality always breaks at $\lambda < 0$. **No.**

INTERSECTIONS, UNIONS, SUMS

Let U, W be subspaces of V.

- $U \cap W$ is *always* a subspace. ($\mathbf{0}$ lies in both; sums and multiples stay in each.) Hence a *system* of homogeneous linear equations always defines a subspace: it is the intersection of the subspaces cut out by the single equations.
- $U \cup W$ is *not* a subspace unless $U \subseteq W$ or $W \subseteq U$. Example: the two axes of $\mathbb{R}^2$; $(1,0)^\mathsf{T} + (0,1)^\mathsf{T} = (1,1)^\mathsf{T}$ is on neither.
- The **sum** $U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U,\ \mathbf{w} \in W\}$ *is* a subspace: the smallest one containing both. For the two axes of $\mathbb{R}^2$, $U + W = \mathbb{R}^2$.

WHAT SUBSPACES LOOK LIKE

The subspaces of $\mathbb{R}^2$ are $\{\mathbf{0}\}$, the lines through the origin, and $\mathbb{R}^2$. The subspaces of $\mathbb{R}^3$ are $\{\mathbf{0}\}$, the lines through the origin, the planes through the origin, and $\mathbb{R}^3$. They are always given by linear equations (that is why the subject is called *linear* algebra). They are always flat, never curve and never leave the origin.

EXAMPLE 5.5 · an "or" is a union

Is $H = \{(x, y, z) \in \mathbb{R}^3 \mid x = y \text{ or } y = z\}$ a subspace?

SOLUTION

No. H is the union of two planes. $(1, 1, 0)$ satisfies $x = y$ and $(0, 1, 1)$ satisfies $y = z$, so both lie in H; their sum $(1, 2, 1)$ has $x \neq y$ and $y \neq z$. (U2) fails.

THE CHAPTER AT A GLANCE

TEST	$U \subseteq V$ is a subspace iff (U1) $\mathbf{0} \in U$, (U2) closed under $+$, (U3) closed under scaling.
WHY	A subspace is a vector space in its own right; the other axioms are inherited.
TRIVIAL	$\{\mathbf{0}\}$ and V are always subspaces. $\mathbb{R}[x] \subset \mathcal{C}^\infty(\mathbb{R}) \subset \mathcal{C}^1(\mathbb{R}) \subset \mathcal{C}(\mathbb{R}) \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$.
PATTERN	Homogeneous linear conditions give subspaces; right-hand side $\neq 0$, squares, products, inequalities and $\mathrm{Det} = 0$ do not.
COMBINING	$U \cap W$ always a subspace; $U \cup W$ almost never; $U + W$ always, and it is the smallest subspace containing both.
SIMPLIFY	Simplify the condition first: $x^2 + y^2 + 1 = 1$ is $\{\mathbf{0}\}$; $x^2 + y^2 + z^2 \geq 0$ is everything; inequalities and "or" fail.
WRITING	Yes: three general checks. No: one explicit counterexample, naming (U1), (U2) or (U3).

TRY IT YOURSELF

1. Which of these are subspaces of $\mathbb{R}^3$? Justify each. (a) $\{(x, y, z) \mid x = y\}$; (b) $\{(x, y, z) \mid x + y + z = 0 \text{ and } x - z = 1\}$; (c) $\{(x, y, z) \mid x \geq 0\}$; (d) $\{(x, y, z) \mid xy = z\}$.
2. Which are subspaces of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$? (a) $\{A \mid a_{1,2} = 0\}$; (b) $\{A \mid A^\mathsf{T} = -A\}$; (c) $\{A \mid a_{1,1}a_{2,2} = 0\}$.
3. Which are subspaces of $\mathcal{F}(\mathbb{R}, \mathbb{R})$? (a) $\{f \mid f(2) = 0\}$; (b) $\{f \mid f(0) = 1\}$; (c) $\{f \mid f(2) = f(3)\}$.
4. Which are subspaces of $\mathbb{R}^3$? (a) $\{x^2 + z^2 = 0\}$; (b) $\{x = y \text{ or } z = 0\}$; (c) $\{x \geq y\}$; (d) $\{x + y + z = 0 \text{ and } x - y = 0\}$.
5. Let U and W be the x-axis and y-axis of $\mathbb{R}^3$. Is $U \cup W$ a subspace? What is $U + W$?

ANSWERS

1. (a) Yes: homogeneous linear ($x - y = 0$); (U1)–(U3) as in the worked example. (b) No: $\mathbf{0}$ gives $0 - 0 = 0 \neq 1$, (U1) fails. (c) No: $(1, 0, 0)$ is in it, $(-1)(1, 0, 0) = (-1, 0, 0)$ is not (U3). (d) No: $(1, 1, 1)$ is in it, $2(1, 1, 1) = (2, 2, 2)$ has $xy = 4 \neq 2 = z$ (U3).
2. (a) Yes (0 has $a_{1,2} = 0$; sums and multiples keep $a_{1,2} = 0$). (b) Yes (skew-symmetric: $0^\mathsf{T} = -0$; $(A + B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T} = -A - B = -(A + B)$; $(\lambda A)^\mathsf{T} = -\lambda A$). (c) No: $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ are in it, their sum I is not.
3. (a) Yes. (b) No: the zero function has $0(0) = 0 \neq 1$. (c) Yes: $(f + g)(2) = f(2) + g(2) = f(3) + g(3) = (f + g)(3)$, likewise for λf, and 0 works.
4. (a) Yes: forces $x = z = 0$, the y-axis. (b) No: $(1, 1, 1)$ and $(1, 0, 0)$ are in it, their sum $(2, 1, 1)$

is not. (c) No: $(1,0,0)$ is in it, $(-1,0,0)$ is not (U3). (d) Yes: homogeneous linear equations (an intersection of two planes: the line through $(1,1,-2)^{\mathsf{T}}$).

5. No: $(1,0,0)+(0,1,0)=(1,1,0)$ is on neither axis. $U+W$ is the plane $z=0$.

PART III

Building subspaces

CHAPTER 6

Linear combinations and span

LECTURES 4–7 · *building subspaces outwards from a list of vectors*

THIS CHAPTER TEACHES HOW TO

- decide whether a vector is a linear combination of given vectors (in $\mathbb{R}^n$, in P_n, in $\mathrm{Mat}_{m\times n}(\mathbb{R})$), and find the coefficients;
- read $A\mathbf{x}$ as a combination of the columns of A;
- write down $\mathrm{Span}(S)$, say when a vector is redundant, and recognise a span in a parametrised set;
- find an equation for a plane given as a span.

6.1 Linear combinations

LINEAR COMBINATION

A **linear combination** of $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ is any vector

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_k\mathbf{v}_k, \qquad \lambda_i \in \mathbb{R},$$

the λ_i being the **coefficients**. Only $+$ and scaling occur, so this is the most general thing built from the list. Any subspace containing the $\mathbf{v}_i$ must contain every linear combination of them.

MATRICES PRODUCE COMBINATIONS OF COLUMNS

$A\mathbf{x}$ is the linear combination of the *columns* of A whose coefficients are the entries of $\mathbf{x}$:

$$\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

So "is $\mathbf{w}$ a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$?" is exactly the question "does $A\mathbf{x} = \mathbf{w}$ have a solution?" for $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$. If A is square with $\mathrm{Det}\, A \neq 0$, the answer is yes for *every* $\mathbf{w}$, with unique coefficients $\mathbf{x} = A^{-1}\mathbf{w}$.

HOW TO DECIDE WHETHER $\mathbf{w}$ IS A COMBINATION OF $\mathbf{v}_1, \dots, \mathbf{v}_k$

1. Write $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{w}$ with unknown coefficients.
2. Compare components: coordinates in $\mathbb{R}^n$, coefficients of $1, x, x^2, \dots$ in P_n (two polynomials are equal exactly when all coefficients agree), entries in $\mathrm{Mat}_{m\times n}(\mathbb{R})$. This gives a system of linear equations.
3. Solve. A solution: yes, and say the coefficients. A contradiction between equations: no, and say which equations clash.
4. Check by substituting the coefficients back.
5. If the $\mathbf{v}_i$ are n vectors in $\mathbb{R}^n$, compute $\mathrm{Det}[\mathbf{v}_1 \cdots \mathbf{v}_n]$ first: non-zero means yes for every $\mathbf{w}$.

COEFFICIENTS NEED NOT BE UNIQUE

If the answer is yes there may be infinitely many choices of coefficients (when one of the $\mathbf{v}_i$ is already a combination of the others). One valid choice answers the question; do not claim it is the only one unless the determinant (or the equations) show it.

EXAMPLE 6.1 · in $\mathbb{R}^3$

Let $\mathbf{v}_1 = (1,0,1)^\mathsf{T}$, $\mathbf{v}_2 = (1,1,0)^\mathsf{T}$, $\mathbf{v}_3 = (0,1,-1)^\mathsf{T}$. Is (a) $\mathbf{w} = (3,2,1)^\mathsf{T}$, (b) $\mathbf{z} = (1,2,3)^\mathsf{T}$ a linear combination of them?

SOLUTION

$a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 = (a+b,\ b+c,\ a-c)^\mathsf{T}$.

(a) $a+b=3$, $b+c=2$, $a-c=1$. The first two give $a = 3-b$, $c = 2-b$, and then $a-c=1$ holds automatically. **Yes**: e.g. $b=0$, $a=3$, $c=2$: $3\mathbf{v}_1 + 2\mathbf{v}_3 = (3,0,3) + (0,2,-2) = (3,2,1)$. Other choices of b also work.

(b) $a+b=1$, $b+c=2$, $a-c=3$: the first two give $a-c = (1-b)-(2-b) = -1 \neq 3$. **No.** (Reason: $\mathbf{v}_3 = \mathbf{v}_2 - \mathbf{v}_1$, so all three lie in the plane $x-y-z=0$; $\mathbf{w}$ satisfies it, $\mathbf{z}$ does not.)

ANIMATION: part (a) ▸; part (b) ▸.

EXAMPLE 6.2 · in P_2

Let $p_1 = 1 - x^2$, $p_2 = x + x^2$. Is $q = 2 + 3x + x^2$ a combination? Is $r = 1 + x + x^2$?

SOLUTION

$ap_1 + bp_2 = a + bx + (b-a)x^2$. For q: $a=2$, $b=3$, $b-a=1$: consistent, so $q = 2p_1 + 3p_2$ (yes). For r: $a=1$, $b=1$, but $b-a = 0 \neq 1$: no. (Two polynomials cannot produce all of P_2, which needs three coefficients.)

ANIMATION: is q in the span? ▸; is r in the span? ▸.

EXAMPLE 6.3 · reading $A\mathbf{x}$ by columns

$\begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$, agreeing with the row-by-column computation.

6.2 Span

SPAN

For a finite set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq V$,

$$\mathrm{Span}(S) = \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k \mid \lambda_i \in \mathbb{R}\},$$

the set of *all* linear combinations; $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$. Then

- $\mathrm{Span}(S)$ is always a *subspace*, and it is the *smallest* subspace containing S (every $\mathbf{v}_i$ is in it, and any subspace containing S contains all of it). So every subspace is a span.
- The generating set is not unique: $\mathrm{Span}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\} = \mathrm{Span}\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\} = \mathrm{Span}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}\} = \mathbb{R}^2$.

REDUNDANT VECTORS

If $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$ then $\mathrm{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \mathrm{Span}\{\mathbf{u}, \mathbf{v}\}$: a vector already reachable from the others adds nothing. Such a $\mathbf{w}$ is called **linearly dependent** on the others (made precise later). *Counting vectors is not enough; what matters is whether any is redundant.*

SPANS IN SMALL CASES

- One non-zero vector spans a *line* through the origin; $\mathbf{0}$ spans $\{\mathbf{0}\}$.
- Two vectors in $\mathbb{R}^2$ not on a common line span $\mathbb{R}^2$: check $\mathrm{Det}[\mathbf{u}\ \mathbf{v}] \neq 0$. n vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$ when $\mathrm{Det} \neq 0$, and then each target has a unique set of coefficients.
- Three vectors in $\mathbb{R}^3$ with $\mathbf{w} = \mathbf{u} + \mathbf{v}$ span only a plane.
- In $\mathbb{R}^2$ no number of vectors spans more than $\mathbb{R}^2$.

HOW TO SEE THAT A PARAMETRISED SET IS A SUBSPACE

1. Write the general element with its parameters, e.g. $(2s - t,\ s,\ 3t)^\mathsf{T}$.
2. Factor the parameters out, one at a time: $s(2, 1, 0)^\mathsf{T} + t(-1, 0, 3)^\mathsf{T}$.
3. If only *fixed vectors times parameters* remain, the set is their span, hence a subspace (one parameter: a line; two independent: a plane).
4. If a leftover *constant* vector remains (e.g. $(s + 1, 2s, s)^\mathsf{T} = s(1, 2, 1)^\mathsf{T} + (1, 0, 0)^\mathsf{T}$), the set is *affine*: it misses $\mathbf{0}$, so (U1) fails.

HOW TO FIND AN EQUATION OF A PLANE $\mathrm{Span}\{\mathbf{u}, \mathbf{v}\} \subset \mathbb{R}^3$

1. Look for $\alpha x + \beta y + \gamma z = 0$ satisfied by both $\mathbf{u}$ and $\mathbf{v}$: two homogeneous equations in α, β, γ.
2. Solve (one free choice, e.g. set $\gamma = 1$).
3. Check that every spanning vector satisfies the equation.

EXAMPLE 6.4 · from span to equation

Find an equation for $\mathrm{Span}\{(1,0,2)^\mathsf{T}, (0,1,-1)^\mathsf{T}\}$.

SOLUTION

$\alpha + 2\gamma = 0$ and $\beta - \gamma = 0$. Take $\gamma = 1$: $\alpha = -2$, $\beta = 1$. The plane is $-2x + y + z = 0$. Check: $-2 + 0 + 2 = 0$ and $0 + 1 - 1 = 0$.

ANIMATION: the normal vector and equation of this plane ▸.

EXAMPLE 6.5 · spans in a parametrisation

Is $H = \{(2s - t,\ s,\ 3t)^\mathsf{T}\}$ a subspace? Is $K = \{(s + 1,\ 2s,\ s)^\mathsf{T}\}$?

SOLUTION

H: factoring gives $s(2,1,0)^\mathsf{T} + t(-1,0,3)^\mathsf{T}$, so $H = \mathrm{Span}\{(2,1,0)^\mathsf{T}, (-1,0,3)^\mathsf{T}\}$, a plane through $\mathbf{0}$: **yes**. K: $s(1,2,1)^\mathsf{T} + (1,0,0)^\mathsf{T}$ has a leftover constant; $\mathbf{0} \in K$ would need $s = -1$ and $2s = 0$ at once. **No** (affine line).

EXAMPLE 6.6 · a determinant settles it

Do $(1,1,0)^\mathsf{T}, (0,1,1)^\mathsf{T}, (1,0,1)^\mathsf{T}$ span $\mathbb{R}^3$? With these as columns, $\mathrm{Det}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = 1(1) - 0 + 1(1) = 2 \neq 0$. Yes, and every vector of $\mathbb{R}^3$ has *unique* coefficients $A^{-1}\mathbf{w}$.

QUICK CHECKS

- The zero vector must be in any span; if your answer set misses $\mathbf{0}$, it is not a span.
- To confirm $\mathbf{w} \in \mathrm{Span}$, substitute the coefficients back and add.
- To rule out membership in a plane, test the plane's equation.

6.3 Spanning sets (Lecture 7)

"S SPANS V"

S **spans** V (V is *generated* by S; S is a **spanning set** of V) when $\mathrm{Span}(S) = V$. The inclusion $\mathrm{Span}(S) \subseteq V$ is automatic; the content is that *every* vector of V is a combination of S. Spanning is always spanning *of something*: adding vectors never leaves the ambient space, so no set of vectors of $\mathbb{R}^2$ spans $\mathbb{R}^3$. Say "the set generates"; whether each vector is needed is a separate question (linear independence, still to come).

STANDARD SPANNING SETS

- $\mathbb{R}^n$ is spanned by $\mathbf{e}_1, \ldots, \mathbf{e}_n$ (a 1 in position i): $\mathbf{x} = \sum x_i \mathbf{e}_i$.
- P_n is spanned by $1, x, \ldots, x^n$.
- $\mathrm{Mat}_{m\times n}(\mathbb{R})$ is spanned by the mn **matrix units** (a single 1, all else 0).
- $\mathbb{R}[x]$ has *no* finite spanning set (degrees cannot exceed the largest one in the set).

Spanning sets are not unique, and not even of fixed size.

HOW TO SHOW A SET SPANS

Recover the *standard* spanning set from it: write each of $\mathbf{e}_i$ (or $1, x, x^2$; or the matrix units) as a combination of your vectors. Everything else follows. In $\mathbb{R}^n$ with n vectors, $\mathrm{Det} \neq 0$ is quicker. If one vector is a combination of the others, drop it. To show a set does *not* span, exhibit a vector it cannot reach (e.g. by an equation of the span, as above).

EXAMPLE 6.7 · spanning sets

(a) $\{1,\ 1+x,\ 1+x+x^2\}$ spans P_2: $1 = 1$, $x = (1+x) - 1$, $x^2 = (1+x+x^2) - (1+x)$.
(b) $\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}3\\1\end{bmatrix}\}$ spans $\mathbb{R}^2$: the first two have $\mathrm{Det} = -2 \neq 0$ and the third is redundant. $\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}\}$ does not: it spans only the line $y = 2x$.
(c) $\{1+x,\ x+x^2,\ 1+x^2\}$ spans P_2: $p_1 - p_2 + p_3 = 2$, so $1 = \frac{1}{2}(p_1 - p_2 + p_3)$, $x = p_1 - 1$, $x^2 = p_3 - 1$.

THE CHAPTER AT A GLANCE

COMBINATION	$\lambda_1\mathbf{v}_1 + \cdots + \lambda_k\mathbf{v}_k$; asking whether $\mathbf{w}$ is one is a system of linear equations.
COLUMNS	$A\mathbf{x}$ is the combination of the columns of A with coefficients x_i. Square with $\mathrm{Det} \neq 0$: every $\mathbf{w}$ is reached, uniquely.
SPAN	All linear combinations; always a subspace, the smallest containing S. Every subspace is a span; the generating set is not unique.
REDUNDANT	$\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$ changes nothing: $\mathrm{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \mathrm{Span}\{\mathbf{u}, \mathbf{v}\}$.
PARAMETRISED	Factor out the parameters: fixed vectors left means a span; a leftover constant means affine, not a subspace.
SPANS V	$\mathrm{Span}(S) = V$. Standard sets: $\mathbf{e}_i$ for $\mathbb{R}^n$, $1, \ldots, x^n$ for P_n, matrix units for $\mathrm{Mat}_{m\times n}(\mathbb{R})$. Show a set spans by recovering the standard one.

TRY IT YOURSELF

1. Is $(5, 2)^\mathsf{T}$ a linear combination of $(1, 1)^\mathsf{T}$ and $(1, -1)^\mathsf{T}$? Find the coefficients.
2. Is $(1, 2, 3)^\mathsf{T}$ a combination of $(1, 0, 1)^\mathsf{T}$ and $(0, 1, 1)^\mathsf{T}$? What about $(1, 2, 4)^\mathsf{T}$?
3. In P_2, is $q = 2 + x + x^2$ a combination of $1 + x$ and $x + x^2$?
4. Show that $H = \{(s + t,\ 2s,\ s - t)^\mathsf{T}\}$ is a subspace, write it as a span, and find an equation for it.
5. Is $\mathrm{Span}\{(1, 0, 1)^\mathsf{T}, (0, 1, 1)^\mathsf{T}, (2, 1, 3)^\mathsf{T}\} = \mathbb{R}^3$? Describe it.
6. Show that $\{1 + x,\ x + x^2,\ 1 + x^2\}$ spans P_2. Does $\{(1, 1)^\mathsf{T}, (1, -1)^\mathsf{T}, (3, 1)^\mathsf{T}\}$ span $\mathbb{R}^2$?

ANSWERS

1. $a + b = 5$, $a - b = 2$: $a = \frac{7}{2}$, $b = \frac{3}{2}$. Yes ($\mathrm{Det} = -2 \neq 0$, so unique).
2. $a = 1$, $b = 2$, then $a + b = 3$: yes. For $(1, 2, 4)$: $a + b = 3 \neq 4$: no.
3. $a(1 + x) + b(x + x^2) = a + (a + b)x + bx^2$: $a = 2$, $b = 1$, $a + b = 3 \neq 1$: no.
4. $s(1, 2, 1)^\mathsf{T} + t(1, 0, -1)^\mathsf{T}$, so $H = \mathrm{Span}\{(1, 2, 1)^\mathsf{T}, (1, 0, -1)^\mathsf{T}\}$. Equation: $\alpha + 2\beta + \gamma = 0$, $\alpha - \gamma = 0$; $\alpha = 1$, $\beta = -1$, $\gamma = 1$: $x - y + z = 0$.
5. No: $(2, 1, 3)^\mathsf{T} = 2(1, 0, 1)^\mathsf{T} + (0, 1, 1)^\mathsf{T}$ is redundant; the span is the plane $x + y - z = 0$.
6. Recover $1, x, x^2$: $p_1 - p_2 + p_3 = 2$ gives 1, then $x = p_1 - 1$, $x^2 = p_3 - 1$. Yes. Second: yes,

since $\mathrm{Det}\left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right] = -2 \neq 0$ and the third vector is redundant.

CHAPTER 7

The kernel of a matrix

LECTURES 6 AND 8 · *cutting a subspace out of $\mathbb{R}^n$ with equations*

THIS CHAPTER TEACHES HOW TO

- test whether a given vector lies in $\operatorname{Ker} A$;
- compute $\operatorname{Ker} A$ and write it as a span;
- explain why $\operatorname{Ker} A$ is a subspace of the *domain*;
- go back and forth between a span and a matrix whose kernel it is.

7.1 Definition and membership

MATRICES AS MAPS, AND THE KERNEL

An $m \times n$ matrix A is a map $A\colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto A\mathbf{x}$. Its **kernel** is the set of vectors sent to zero,

$$\operatorname{Ker} A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

$\operatorname{Ker} A$ is a subspace of $\mathbb{R}^n$, the *domain* (source), not of $\mathbb{R}^m$. (Reason: $A\mathbf{0} = \mathbf{0}$, and $A(\mathbf{u}+\mathbf{w}) = A\mathbf{u} + A\mathbf{w}$, $A(\lambda\mathbf{u}) = \lambda A\mathbf{u}$.)

SQUARE MATRICES

For a square matrix, $\operatorname{Ker} A \neq \{\mathbf{0}\}$ exactly when $\operatorname{Det} A = 0$: the collapse of Chapter 2 seen from the other side (a whole line or plane is crushed onto $\mathbf{0}$).

EXAMPLE 7.1 · membership is one multiplication

Let $A = \left[\begin{smallmatrix} 1 & -1 & 2 \\ 2 & -1 & 3 \end{smallmatrix}\right]$. Is $(-2,2,2)^{\mathsf{T}} \in \operatorname{Ker} A$? Is $(1,1,0)^{\mathsf{T}}$?

SOLUTION

$A(-2,2,2)^{\mathsf{T}} = (-2-2+4,\ -4-2+6)^{\mathsf{T}} = (0,0)^{\mathsf{T}}$: yes. $A(1,1,0)^{\mathsf{T}} = (0,1)^{\mathsf{T}} \neq \mathbf{0}$: no.

WHERE THE KERNEL LIVES

$\operatorname{Ker} A$ consists of *inputs* (vectors with n entries, n= number of columns), not outputs. For $A = \left[\begin{smallmatrix} 1 & -1 & 2 \\ 2 & -1 & 3 \end{smallmatrix}\right]$, $\operatorname{Ker} A \subseteq \mathbb{R}^3$, even though $A\mathbf{x} \in \mathbb{R}^2$.

7.2 Computing a kernel

HOW TO FIND $\operatorname{Ker} A$

1. Row-reduce A to $\operatorname{RREF}(A)$ (Chapter 3); this does not change $\operatorname{Ker} A$.
2. Variables of pivot columns are **basic**, the others **free**. Give each free variable its own parameter $(s, t, \ldots)$.
3. Read the pivot rows to express each basic variable in terms of the parameters.
4. Write $\mathbf{x} = s\mathbf{k}_1 + t\mathbf{k}_2 + \cdots$ and conclude $\operatorname{Ker} A = \operatorname{Span}\{\mathbf{k}_1, \mathbf{k}_2, \ldots\}$. State the answer as a span.
5. Check: $A\mathbf{k}_i = \mathbf{0}$ for every spanning vector.

One spanning vector per free column, so the number of vectors is $n - r$ for r pivots. If $r = n$, nothing is free and $\operatorname{Ker} A = \{\mathbf{0}\}$.

EXAMPLE 7.2 · one parameter

Find $\operatorname{Ker} A$ for $A = \left[\begin{smallmatrix} 1 & -1 & 2 \\ 2 & -1 & 3 \end{smallmatrix}\right]$.

SOLUTION

$a - b + 2c = 0$ and $2a - b + 3c = 0$. Subtracting: $a + c = 0$, so $a = -c$; then $b = a + 2c = c$. With $c = t$: $\mathbf{x} = t(-1, 1, 1)^{\mathsf{T}}$, so $\operatorname{Ker} A = \operatorname{Span}\{(-1, 1, 1)^{\mathsf{T}}\}$, a line. Check: $-1 - 1 + 2 = 0$ and $-2 - 1 + 3 = 0$.

EXAMPLE 7.3 · two parameters: redundant rows

Find $\operatorname{Ker} A$ for $A = \left[\begin{smallmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \end{smallmatrix}\right]$.

SOLUTION

Row 2 is $-2\times$row 1, so the system is the single equation $x - 2y + z = 0$. Two unknowns are free: $y = s$, $z = t$ give $x = 2s - t$, so

$$\mathbf{x} = s\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \operatorname{Ker} A = \operatorname{Span}\{(2, 1, 0)^{\mathsf{T}}, (-1, 0, 1)^{\mathsf{T}}\},$$

a plane through the origin.

EXAMPLE 7.4 · a square matrix

$A = \left[\begin{smallmatrix} 1 & 2 \\ 2 & 4 \end{smallmatrix}\right]$ has $\operatorname{Det} A = 0$. $x + 2y = 0$ (the second equation is twice the first), so $y = t$, $x = -2t$: $\operatorname{Ker} A = \operatorname{Span}\{(-2, 1)^{\mathsf{T}}\}$. By contrast $\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ has $\operatorname{Det} = -2 \neq 0$, so its kernel is $\{\mathbf{0}\}$.

EXAMPLE 7.5 · with row reduction

Find $\operatorname{Ker} A$ for $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 3 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ (a 4×3 matrix, so $\operatorname{Ker} A \subseteq \mathbb{R}^3$).

SOLUTION

$R_2 \leftarrow R_2 - 2R_1$, $R_3 \leftarrow R_3 - R_1$, swap rows 2 and 3, $R_4 \leftarrow R_4 - R_2$, $R_1 \leftarrow R_1 - 2R_2$ give $\operatorname{RREF}(A) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Pivots in columns $1, 2$; $z = t$ is free; $x = t$, $y = -2t$. So $\operatorname{Ker} A = \operatorname{Span}\{(1, -2, 1)^{\mathsf{T}}\}$, a line. Check: $A(1, -2, 1)^{\mathsf{T}} = (1{-}4{+}3,\ 2{-}8{+}6,\ 1{-}6{+}5,\ 0{-}2{+}2)^{\mathsf{T}} = \mathbf{0}$.

ANIMATION: watch the kernel being found ▸.

EXAMPLE 7.6 · two free columns

$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{bmatrix}$ reduces to $\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Pivots in columns $1, 3$; $y = s$, $w = t$ are free, $x = -s + t$, $z = -2t$: $\operatorname{Ker} B = \operatorname{Span}\{(-1, 1, 0, 0)^{\mathsf{T}}, (1, 0, -2, 1)^{\mathsf{T}}\}$, a plane inside $\mathbb{R}^4$.

ANIMATION: two free columns, two parameters ▸.

CHECKS

- Every spanning vector must satisfy $A\mathbf{k} = \mathbf{0}$; multiply it out.
- The number of free parameters is the number of unknowns minus the number of *independent* equations; here 1 parameter gave a line and 2 gave a plane.
- If A is square with $\operatorname{Det} A \neq 0$ the only solution is $\mathbf{x} = \mathbf{0}$.

7.3 Two ways of describing a subspace

SPAN VERSUS KERNEL

Span builds a subspace *outwards* from a list of vectors; **kernel** cuts one *out* of $\mathbb{R}^n$ with equations. Every kernel we compute ends up written as a span, and every span in $\mathbb{R}^3$ that is a plane is a kernel: if $\operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$ satisfies $\alpha x + \beta y + \gamma z = 0$ (Chapter 6), then it is $\operatorname{Ker} \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix}$.

EXAMPLE 7.7 · from span to a matrix

$H = \operatorname{Span}\{(1, 0, 2)^{\mathsf{T}}, (0, 1, -1)^{\mathsf{T}}\}$ satisfies $-2x + y + z = 0$, so $H = \operatorname{Ker} \begin{bmatrix} -2 & 1 & 1 \end{bmatrix}$.

THE CHAPTER AT A GLANCE

KERNEL	$\operatorname{Ker} A = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\}$, a subspace of the domain $\mathbb{R}^n$.
MEMBERSHIP	One multiplication: is $A\mathbf{x} = \mathbf{0}$?
COMPUTING	Reduce to RREF; pivot columns basic, others free; parametrise and read off a span; $n - r$ vectors; check by multiplying.
SQUARE	$\operatorname{Ker} A \neq \{\mathbf{0}\}$ iff $\operatorname{Det} A = 0$.
BOTH WAYS	Spans are built outwards, kernels cut inwards; a plane given as a span is the kernel of a one-row matrix.

TRY IT YOURSELF

1. Find $\operatorname{Ker}\begin{bmatrix}1&2&-1\\2&1&1\end{bmatrix}$ as a span.
2. Find $\operatorname{Ker}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\end{bmatrix}$ as a span.
3. Find $\operatorname{Ker}\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$. (Careful: what is the domain?)
4. Find a matrix A with $\operatorname{Ker} A = \operatorname{Span}\{(1,1,0)^{\mathsf{T}}, (2,-1,3)^{\mathsf{T}}\}$.
5. Find $\operatorname{RREF}(A)$ and a spanning set of $\operatorname{Ker} A$ for $A = \begin{bmatrix}1&3&-2\\2&6&-4\\-1&-3&2\end{bmatrix}$.
6. An $m \times n$ matrix in RREF has r pivots. How many spanning vectors does the recipe give? What if $r = n$?

ANSWERS

1. $a + 2b - c = 0$, $2a + b + c = 0$: adding gives $3a + 3b = 0$, so $b = -a$, $c = a + 2b = -a$. $\operatorname{Ker} = \operatorname{Span}\{(1,-1,-1)^{\mathsf{T}}\}$. Check: $1 - 2 + 1 = 0$, $2 - 1 - 1 = 0$.
2. Adding the equations: $2x + 2z = 0$; subtracting: $2y + 2w = 0$. So $x = -z$, $y = -w$: $\operatorname{Ker} = \operatorname{Span}\{(-1,0,1,0)^{\mathsf{T}}, (0,-1,0,1)^{\mathsf{T}}\}$.
3. The domain is $\mathbb{R}^2$. The columns $(1,3,5)^{\mathsf{T}}$, $(2,4,6)^{\mathsf{T}}$ are not parallel, so only $\mathbf{x} = \mathbf{0}$ solves it: $\operatorname{Ker} = \{\mathbf{0}\}$.
4. $\alpha + \beta = 0$, $2\alpha - \beta + 3\gamma = 0$; $\alpha = 1$: $\beta = -1$, $\gamma = -1$. $A = \begin{bmatrix}1&-1&-1\end{bmatrix}$. Check: $1 - 1 - 0 = 0$, $2 + 1 - 3 = 0$.
5. $R_2 - 2R_1$, $R_3 + R_1$ give $\operatorname{RREF}(A) = \begin{bmatrix}1&3&-2\\0&0&0\\0&0&0\end{bmatrix}$. One pivot; $y = s$, $z = t$ free, $x = -3s + 2t$: $\operatorname{Ker} A = \operatorname{Span}\{(-3,1,0)^{\mathsf{T}}, (2,0,1)^{\mathsf{T}}\}$, a plane.
6. $n - r$ vectors (one per free column); $r = n$ gives $\operatorname{Ker} A = \{\mathbf{0}\}$.

CHAPTER 8

Column space and image

LECTURES 7–8 · *the subspace hiding in the columns, and membership by row reduction*

THIS CHAPTER TEACHES HOW TO

- write down a spanning set of $\operatorname{Col} A = \operatorname{Im} A$ at once;
- decide whether $\mathbf{b} \in \operatorname{Col} A$ by reducing $[A \mid \mathbf{b}]$;
- say which of kernel and image is easy at what, and why row operations may not be used to read off a column space.

8.1 Definition

COLUMN SPACE = IMAGE

For an $m \times n$ matrix A with columns $a_{-,1}, \dots, a_{-,n} \in \mathbb{R}^m$,

$$\begin{aligned} \operatorname{Col} A = \operatorname{Span}\{a_{-,1}, \dots, a_{-,n}\} &= \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} \\ &= \{\mathbf{b} \in \mathbb{R}^m \mid A\mathbf{x} = \mathbf{b} \text{ is solvable}\} = \operatorname{Im} A. \end{aligned}$$

It is a subspace of the *target* $\mathbb{R}^m$ (while $\operatorname{Ker} A$ lives in the source $\mathbb{R}^n$). Being a span, it is automatically a subspace: nothing to check.

KERNEL AGAINST IMAGE

	$\operatorname{Ker} A \subseteq \mathbb{R}^n$	$\operatorname{Im} A \subseteq \mathbb{R}^m$
given by	an equation $A\mathbf{x} = \mathbf{0}$	a list: the columns
membership	easy: one multiplication	work: solve $A\mathbf{x} = \mathbf{b}$
spanning set	work: solve the system	free: the columns

EXAMPLE 8.1 · spanning sets by inspection

I_n has $\operatorname{Col} I_n = \operatorname{Span}\{\mathbf{e}_1, \dots, \mathbf{e}_n\} = \mathbb{R}^n$. $A = \left[\begin{smallmatrix} 1 & 2 & -1 \\ -1 & 0 & 4 \end{smallmatrix}\right]$ has three columns in $\mathbb{R}^2$; the first two are not parallel ($\operatorname{Det} \left[\begin{smallmatrix} 1 & 2 \\ -1 & 0 \end{smallmatrix}\right] = 2$), so $\operatorname{Col} A = \mathbb{R}^2$ and the third is redundant. $B = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ has two non-parallel columns in $\mathbb{R}^3$: $\operatorname{Col} B$ is a plane through $\mathbf{0}$, never all of $\mathbb{R}^3$.

ROW OPERATIONS CHANGE THE COLUMN SPACE

$\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \to \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ by $R_2 - R_1$, but $\mathrm{Span}\{(1,1)^\mathsf{T}\} \neq \mathrm{Span}\{(1,0)^\mathsf{T}\}$. Never read a column space off $\mathrm{RREF}(A)$; read it from the *original* columns.

8.2 Is $\mathbf{b}$ in the column space?

HOW TO DECIDE $\mathbf{b} \in \mathrm{Im}\, A$

1. $\mathbf{b} \in \mathrm{Im}\, A \iff A\mathbf{x} = \mathbf{b}$ has a solution (consistent system). You need existence, not necessarily the solution.
2. Row-reduce the augmented matrix $[A \mid \mathbf{b}]$, applying every operation to $\mathbf{b}$ too.
3. If a row $\begin{bmatrix}0 & \cdots & 0 \mid c\end{bmatrix}$ with $c \neq 0$ appears (a pivot in the augmented column), the system says $0 = c$: no, $\mathbf{b} \notin \mathrm{Im}\, A$. Otherwise yes.
4. For a witness, set the free variables to 0 and solve; check $A\mathbf{x} = \mathbf{b}$.

EXAMPLE 8.2 · the same A, two targets

$A = \left[\begin{smallmatrix}1&2&0&1\\2&4&1&3\\1&2&1&2\end{smallmatrix}\right]$. Is $\mathbf{b} = (1,3,2)^\mathsf{T}$ in $\mathrm{Col}\, A$? Is $\mathbf{c} = (1,3,3)^\mathsf{T}$?

SOLUTION

$R_2 - 2R_1$, $R_3 - R_1$, $R_3 - R_2$ give $\left[\begin{smallmatrix}1&2&0&1&1\\0&0&1&1&1\\0&0&0&0&0\end{smallmatrix}\right]$ for $\mathbf{b}$: the last row is $0 = 0$, so **yes**. With $y = w = 0$: $z = 1$, $x = 1$, and $1\left[\begin{smallmatrix}1\\2\\1\end{smallmatrix}\right] + 1\left[\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}1\\3\\2\end{smallmatrix}\right]$.

For $\mathbf{c}$ the same operations end in $\begin{bmatrix}0 & 0 & 0 & 0 \mid 1\end{bmatrix}$, i.e. $0 = 1$: **no**.

Seen from outside: $a_{-,2} = 2a_{-,1}$ and $a_{-,4} = a_{-,1} + a_{-,3}$, so $\mathrm{Col}\, A = \mathrm{Span}\{(1,2,1)^\mathsf{T}, (0,1,1)^\mathsf{T}\}$, the plane $x - y + z = 0$. $1 - 3 + 2 = 0$ but $1 - 3 + 3 = 1 \neq 0$.

ANIMATION: solving $A\mathbf{x} = \mathbf{b}$ ▸; solving $A\mathbf{x} = \mathbf{c}$ ▸.

CROSS-CHECK

A column space that is a plane in $\mathbb{R}^3$ has one equation; test $\mathbf{b}$ in it. For a square A with $\mathrm{Det}\, A \neq 0$, every $\mathbf{b}$ is in $\mathrm{Col}\, A$ (Chapter 6).

THE CHAPTER AT A GLANCE

COLUMN SPACE	$\mathrm{Col}\, A = \mathrm{Im}\, A = \mathrm{Span}(\text{columns}) = \{A\mathbf{x}\} \subseteq \mathbb{R}^m$; spanning set free.
MEMBERSHIP	Reduce $[A \mid \mathbf{b}]$. A row $[0 \cdots 0 \mid c]$, $c \neq 0$, means no; otherwise yes.
WARNING	Row operations keep $\mathrm{Ker}\, A$, not $\mathrm{Col}\, A$. Take columns from the original A.

TRY IT YOURSELF

1. Let $A = \left[\begin{smallmatrix}1&0&2\\0&1&1\\1&1&3\end{smallmatrix}\right]$. Is $(1,2,3)^\mathsf{T} \in \mathrm{Col}\, A$? Is $(1,2,4)^\mathsf{T}$?
2. Let $A = \left[\begin{smallmatrix}1&2&1\\0&1&1\\1&1&0\end{smallmatrix}\right]$. Write down a spanning set of $\mathrm{Col}\, A$ with no computation, then decide

whether $(1,1,1)^\mathsf{T} \in \operatorname{Col} A$.

3. Is $(1,2,3)^\mathsf{T}$ in $\operatorname{Col} \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{smallmatrix}\right]$? And $(1,2,4)^\mathsf{T}$?

ANSWERS

1. $R_3 - R_1 - R_2$ gives $\left[\begin{smallmatrix} 0 & 0 & 0 \mid 0 \end{smallmatrix}\right]$ for $(1,2,3)^\mathsf{T}$: yes. For $(1,2,4)^\mathsf{T}$ it gives $\left[\begin{smallmatrix} 0 & 0 & 0 \mid 1 \end{smallmatrix}\right]$ $(4-1-2)$: no.
2. The three columns span it. $\operatorname{Det} A = 0$ (column 3 is column 2 minus column 1), so $\operatorname{Col} A = \operatorname{Span}\{(1,0,1)^\mathsf{T}, (2,1,1)^\mathsf{T}\}$. $a(1,0,1)^\mathsf{T} + b(2,1,1)^\mathsf{T} = (1,1,1)^\mathsf{T}$ forces $b = 1$, $a = -1$, but the third row gives $0 \neq 1$: no.
3. Vectors of the column space are $(a, b, a+b)^\mathsf{T}$: third entry is the sum. $(1,2,3)^\mathsf{T}$ yes ($a = 1, b = 2$); $(1,2,4)^\mathsf{T}$ no.

PART IV

Linear maps

CHAPTER 9

Linear transformations

LECTURE 9 · *structure-preserving maps, which are matrices in disguise*

THIS CHAPTER TEACHES HOW TO

- decide, with justification, whether a map is linear;
- find the standard matrix of a linear map $\mathbb{R}^n \to \mathbb{R}^m$;
- find $\operatorname{Ker} T$ and $\operatorname{Im} T$, including for maps on polynomials and matrices;
- present a subspace as an image and as a kernel.

9.1 Linear maps

LINEAR TRANSFORMATION

$T\colon V \to W$ is **linear** (a vector space homomorphism) if for all $\mathbf{u}, \mathbf{v} \in V$, $\lambda \in \mathbb{R}$:

$$T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}), \qquad T(\lambda \mathbf{v}) = \lambda T(\mathbf{v}).$$

Consequences: $T(\mathbf{0}) = \mathbf{0}$ always, and $T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$, so T is determined by its values on a spanning set.

HOW TO DECIDE WHETHER T IS LINEAR

1. Quick test: $T(\mathbf{0}) \neq \mathbf{0}$ means not linear (affine maps such as $(x, y) \mapsto (x+1, y)$ fail).
2. If T is given by formulas that are sums of constant multiples of the coordinates (no products, squares, constants), it is T_A for a matrix A: linear.
3. Otherwise check both rules in general, or disprove one with a counterexample, such as $T(2\mathbf{v}) \neq 2T(\mathbf{v})$.

EXAMPLE 9.1 · linear or not?

(a) $T(x, y) = (2x - y, 0)$: yes, T_A with $A = \left[\begin{smallmatrix} 2 & -1 \\ 0 & 0 \end{smallmatrix}\right]$. (b) $T(x, y) = (xy, x)$: no, $T(2\mathbf{v}) = (4xy, 2x) \neq 2T(\mathbf{v}) = (2xy, 2x)$. (c) $T(x, y) = (x+1, y-1)$: no, $T(\mathbf{0}) = (1, -1) \neq \mathbf{0}$. (d) $T\colon P_2 \to P_2$, $T(p) = p'' + p$: yes, $(p+q)'' + (p+q) = (p'' + p) + (q'' + q)$ and likewise for scalars.

SQUARES AND SHIFTS

Any square, product or added constant breaks linearity (that is why it is *linear* algebra). Keeping $T(\mathbf{0}) = \mathbf{0}$ is necessary, not sufficient: $(x, y) \mapsto (x^2, y)$ fixes $\mathbf{0}$ but is not linear.

EXAMPLE 9.2 · beyond $\mathbb{R}^n$

Linear: differentiation $D\colon P_2 \to P_1$, $D(p) = p'$; the trace $\mathrm{Tr}\colon \mathrm{Mat}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$; evaluation $E_3\colon P_2 \to \mathbb{R}$, $p \mapsto p(3)$.

9.2 Maps and matrices

STANDARD MATRIX

Every $m \times n$ matrix gives a linear map $T_A(\mathbf{x}) = A\mathbf{x}\colon \mathbb{R}^n \to \mathbb{R}^m$, and *every* linear $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is T_A for exactly one matrix, the **standard matrix**

$$[T] = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}:$$

the columns are the images of the standard vectors.

EXAMPLE 9.3 · finding $[T]$

$T(x_1, x_2) = (x_1 - x_2,\ x_1 + x_2)$: $T(\mathbf{e}_1) = (1, 1)^{\mathsf{T}}$, $T(\mathbf{e}_2) = (-1, 1)^{\mathsf{T}}$, so $[T] = \left[\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right]$, with $\mathrm{Det} = 2$: it rotates by $45°$ and stretches by $\sqrt{2}$ (areas double). $T(x, y, z) = (x - z,\ 2y + z)$ has $[T] = \left[\begin{smallmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{smallmatrix}\right]$.

9.3 Kernel and image in general

KERNEL AND IMAGE OF A LINEAR MAP

$$\mathrm{Ker}\, T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\} \subseteq V, \qquad \mathrm{Im}\, T = \{T(\mathbf{v}) \mid \mathbf{v} \in V\} \subseteq W,$$

both always subspaces (of source and target respectively). For T_A: $\mathrm{Ker}\, A$ and $\mathrm{Col}\, A$. Examples: for $D\colon P_2 \to P_1$, $\mathrm{Ker}\, D$ is the constants and $\mathrm{Im}\, D = P_1$; for E_3, $\mathrm{Ker}\, E_3$ is the polynomials vanishing at 3, and $\mathrm{Im}\, E_3 = \mathbb{R}$.

EXAMPLE 9.4 · from the formula

$T(x, y, z) = (x - z,\ 2y + z)$, $A = [T] = \left[\begin{smallmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{smallmatrix}\right]$. $x = z$, $y = -z/2$; put $z = 2t$: $\mathrm{Ker}\, T = \mathrm{Span}\{(2, -1, 2)^{\mathsf{T}}\}$ (check $2 - 2 = 0$, $-2 + 2 = 0$). The first two columns are not parallel, so $\mathrm{Im}\, T = \mathrm{Col}\, A = \mathbb{R}^2$.

EVERY SUBSPACE IS AN IMAGE AND A KERNEL

For every subspace H of a vector space there are linear maps S and T with $\mathrm{Im}\, S = H = \mathrm{Ker}\, T$. So the two descriptions of a subspace (built from vectors, or cut out by equations) are *always* both available. For $H \subseteq \mathbb{R}^3$: put the spanning vectors as columns of S; put equations satisfied by the spanning vectors as rows of T (Chapter 6).

EXAMPLE 9.5 · both descriptions

$H = \mathrm{Span}\{(1,0,1)^{\mathsf{T}}\}$: $S = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ has $\mathrm{Im}\, S = H$; $T = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$ has $T\mathbf{x} = \mathbf{0} \iff y = 0,\ x = z$, so $\mathrm{Ker}\, T = H$. For the plane $H = \mathrm{Span}\{(1,1,0)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\}$: $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $T = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$ (check $1 - 1 + 0 = 0$, $0 - 1 + 1 = 0$). A plane needs one equation; a line in $\mathbb{R}^3$ needs two.

THE CHAPTER AT A GLANCE

LINEAR	$T(\mathbf{u} + \mathbf{v}) = T\mathbf{u} + T\mathbf{v}$, $T(\lambda\mathbf{v}) = \lambda T\mathbf{v}$. Forces $T(\mathbf{0}) = \mathbf{0}$. Squares, products, shifts are out.
MATRIX	$[T] = [T(\mathbf{e}_1) \cdots T(\mathbf{e}_n)]$; every linear $\mathbb{R}^n \to \mathbb{R}^m$ is T_A.
Ker, Im	Subspaces of source and target; for T_A they are $\mathrm{Ker}\, A$ and $\mathrm{Col}\, A$.
BOTH WAYS	Every subspace is an image and a kernel: columns span it, rows cut it out.

TRY IT YOURSELF

1. Which are linear? (a) $T(x, y) = (2x - y, 0)$; (b) $T(x, y) = (xy, x)$; (c) $T(x, y) = (x + 1, y - 1)$; (d) $T(p) = p'' + p$ on P_2.
2. Find $[T]$ for $T(x, y, z) = (x - z,\ 2y + z)$; describe $\mathrm{Ker}\, T$ as a span, and say what $\mathrm{Im}\, T$ is.
3. Let $H = \mathrm{Span}\{(1,1,0)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\}$. Find S with $\mathrm{Im}\, S = H$ and T with $\mathrm{Ker}\, T = H$.

ANSWERS

1. (a) yes; (b) no, $T(2\mathbf{v}) = (4xy, 2x) \neq 2T(\mathbf{v})$; (c) no, $T(\mathbf{0}) \neq \mathbf{0}$; (d) yes.
2. $[T] = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix}$; $\mathrm{Ker}\, T = \mathrm{Span}\{(2, -1, 2)^{\mathsf{T}}\}$; $\mathrm{Im}\, T = \mathbb{R}^2$.
3. $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$; $\alpha + \beta = 0$, $\beta + \gamma = 0$ gives $T = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$.

PART V

Bases and dimension

CHAPTER 10

Linear independence and bases

LECTURES 10 AND 13 · *minimal spanning sets, and how many vectors a space needs*

THIS CHAPTER TEACHES HOW TO

- decide whether vectors are independent or dependent, and justify it;
- recognise a basis, and cast redundant vectors out of a spanning set;
- state the dimension of the standard spaces.

10.1 Linear independence

INDEPENDENT / DEPENDENT

$\mathbf{v}_1, \dots, \mathbf{v}_p$ are **linearly independent** if

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \implies c_1 = \cdots = c_p = 0.$$

Otherwise they are **dependent**: some combination with *not all coefficients zero* gives $\mathbf{0}$. Equivalent (for $p \geq 2$): dependent $\iff$ one of the vectors is a linear combination of the others. Independence is a property of the *whole list*.

QUICK CHECKS

- $\{\mathbf{v}\}$ is independent iff $\mathbf{v} \neq \mathbf{0}$; any list containing $\mathbf{0}$ is dependent.
- Two non-zero vectors are independent iff neither is a multiple of the other.
- n vectors in $\mathbb{R}^n$: independent $\iff$ $\mathrm{Det}[\mathbf{v}_1 \cdots \mathbf{v}_n] \neq 0$.
- More than n vectors in $\mathbb{R}^n$ are always dependent (two on a line, three in a plane, four in space).
- Independence survives removing vectors; spanning survives adding them.

HOW TO DECIDE INDEPENDENCE

1. Write $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ and turn it into equations for the c_i (coordinates; coefficients of $1, x, x^2$; entries).
2. Solve. Only $c = \mathbf{0}$: independent. A non-zero solution: dependent, and that solution *is* the relation (exhibit it).
3. In $\mathbb{R}^n$ with n vectors, use $\mathrm{Det} \neq 0$ instead. Alternatively spot a vector that is a combination of the others.

EXAMPLE 10.1 · independent or dependent?

(a) $(1,2,3)^{\mathsf{T}}, (2,4,6)^{\mathsf{T}}$: dependent, $2\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$.
(b) $(1,0,1)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}, (1,1,0)^{\mathsf{T}}$: $\det = 1(0-1) - 0 + 1(0-1) = -2 \neq 0$: independent.
(c) $1+x,\ 1-x,\ x^2$ in P_2: $a(1+x)+b(1-x)+cx^2 = 0$ gives $a+b=0$, $a-b=0$, $c=0$, so all zero: independent.
(d) $(1,0)^{\mathsf{T}}, (0,1)^{\mathsf{T}}, (1,1)^{\mathsf{T}}$: no two parallel, yet $(1,0)^{\mathsf{T}} + (0,1)^{\mathsf{T}} - (1,1)^{\mathsf{T}} = \mathbf{0}$: dependent.
(e) $p_1 = 1+x$, $p_2 = x+x^2$, $p_3 = 1+2x+x^2$: $p_3 = p_1 + p_2$, dependent. In $\mathrm{Mat}_{2\times 2}(\mathbb{R})$: $\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right], I$ are dependent since $I = A + B$.

10.2 Bases and dimension

BASIS

A **basis** of V is a list that *spans* V and is *linearly independent*: exactly the right number of vectors. Bases are never unique (except $\{\mathbf{0}\}$). Standard bases: $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for $\mathbb{R}^n$; $1, x, \ldots, x^n$ for P_n; the mn matrix units for $\mathrm{Mat}_{m\times n}(\mathbb{R})$. Other bases: $\{(1,1)^{\mathsf{T}}, (1,-1)^{\mathsf{T}}\}$ of $\mathbb{R}^2$, $\{1, 1+x, 1+x+x^2\}$ of P_2.

CASTING OUT, AND DIMENSION

If a list is dependent, one vector that is a combination of the others can be removed without changing the span. Repeating this turns *any finite spanning set into a basis.* All bases of V have the same size, the **dimension** $\dim V$:

$$\dim \mathbb{R}^n = n, \quad \dim P_n = n+1, \quad \dim \mathrm{Mat}_{m\times n}(\mathbb{R}) = mn, \quad \dim\{\mathbf{0}\} = 0.$$

Consequently, in an n-dimensional space, n independent vectors form a basis, and so do n spanning vectors.

CAST OUT ONE AT A TIME

If $\mathbf{v}_i$ and $\mathbf{v}_j$ are each a combination of the others, you may remove *either*, not automatically both. In $\{(1,0)^{\mathsf{T}}, (0,1)^{\mathsf{T}}, (1,1)^{\mathsf{T}}\}$, $(1,0)^{\mathsf{T}} = (1,1)^{\mathsf{T}} - (0,1)^{\mathsf{T}}$ and $(0,1)^{\mathsf{T}} = (1,1)^{\mathsf{T}} - (1,0)^{\mathsf{T}}$, but removing both leaves $(1,1)^{\mathsf{T}}$ alone, which no longer spans.

EXAMPLE 10.2 · basis of $\mathbb{R}^3$ and casting out

$\{(1,1,0)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}, (1,0,1)^{\mathsf{T}}\}$ has $\det = 2 \neq 0$, so it is independent, hence a basis of $\mathbb{R}^3$ (dimension 3). Adding $(2,2,0)^{\mathsf{T}} = 2(1,1,0)^{\mathsf{T}}$ gives a dependent spanning set; you may cast out $(1,1,0)^{\mathsf{T}}$ (it is a multiple of the new vector), but not one of the other two.

10.3 Dimension and the Basis Theorem (Lecture 13)

WHY DIMENSION IS WELL DEFINED

If V has a basis of n vectors, then any $n+1$ vectors in V are linearly dependent (in coordinates there are $n+1$ columns but at most n pivots, so some column is free). Hence all bases of V have the same size, and $\dim V$ makes sense. In particular *more than* $\dim V$ *vectors are always dependent*: any three polynomials in P_1, any four vectors in $\mathbb{R}^3$, any five matrices in $\mathrm{Mat}_{2\times 2}(\mathbb{R})$.

SUBSPACES AND EXTENSION

Let $\dim V = n$ and $H \subseteq V$ a subspace.

- H is finitely generated and $\dim H \leq n$; and $\dim H = n \iff H = V$.
- Any linearly independent set in V can be extended to a basis of V (add vectors outside the span).
- $\dim \mathbb{C} = 2$ over $\mathbb{R}$ (basis $1, i$); $\mathbb{R}[x]$ is not finitely generated.

BASIS THEOREM

Let $\dim V = n$ and take exactly n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$. The following are equivalent: (1) they form a basis; (2) they are independent; (3) they span V. So count the vectors, then check only one condition, usually independence ($\mathrm{Det} \neq 0$ in coordinates).

IS IT A BASIS? FAST ROUTE

1. Count: the number of vectors must equal $\dim V$ (otherwise: not a basis; fewer cannot span, more are dependent).
2. Write coordinate vectors with respect to a known basis and compute $\mathrm{Det} \neq 0$.
3. Conclude from the Basis Theorem; no need to check spanning.

EXAMPLE 10.3 · Basis Theorem in use

(a) $\left(\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\right)$: two vectors in $\mathbb{R}^2$, $\mathrm{Det} = -2 \neq 0$: a basis. (b) $1+x$, $x+x^2$, $1+x^2$ in P_2: three vectors, $\dim = 3$, coordinates have $\mathrm{Det}\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&1&1\end{smallmatrix}\right] = 2 \neq 0$: a basis. (c) $\{(1,2,1)^\mathsf{T}, (0,1,1)^\mathsf{T}, (2,1,-1)^\mathsf{T}\}$: $\mathrm{Det} = 1(-1-1) - 0 + 2(2-1) = 0$, so not a basis; the dependence is $2\mathbf{v}_1 - 3\mathbf{v}_2 = \mathbf{v}_3$ (the span is a plane). (d) Extend $\{(1,1,0)^\mathsf{T}\}$ to a basis of $\mathbb{R}^3$: add $(0,1,0)^\mathsf{T}, (0,0,1)^\mathsf{T}$; $\mathrm{Det}\left[\begin{smallmatrix}1&0&0\\1&1&0\\0&0&1\end{smallmatrix}\right] = 1$, so it is a basis. Many other completions work.

ANIMATION: part (c): which vector is redundant, and how it is built from the others ▸.

EXAMPLE 10.4 · dimension of a subspace

$H = \{p \in P_2 \mid p(1) = 0\}$ is a subspace (kernel of evaluation at 1). $x-1$ and x^2-1 vanish at 1 and are independent; if $p = a+bx+cx^2$ has $p(1) = 0$ then $a = -b-c$, so $p = b(x-1)+c(x^2-1)$. Basis $\{x-1, x^2-1\}$, $\dim H = 2$. One linear condition in a 3-dimensional space loses one dimension.

THE CHAPTER AT A GLANCE

INDEPENDENT	$\sum c_i \mathbf{v}_i = \mathbf{0} \Rightarrow$ all $c_i = 0$. Dependent iff one is a combination of the others.
CHECKS	$\mathbf{v} \neq \mathbf{0}$; contains $\mathbf{0}$ is dependent; two vectors: not multiples; n in $\mathbb{R}^n$: $\mathrm{Det} \neq 0$; more than n in $\mathbb{R}^n$: dependent.
BASIS	Spanning and independent. Not unique. Standard bases of $\mathbb{R}^n$, P_n, $\mathrm{Mat}_{m\times n}(\mathbb{R})$.
DIMENSION	Number of vectors in any basis: n, $n+1$, mn. Casting out turns a finite spanning set into a basis.
DIMENSION	$n+1$ vectors in an n-dimensional space are dependent, so $\dim$ is well defined. $H \subseteq V$: $\dim H \leq \dim V$, equality iff $H = V$. Independent sets extend to bases.
BASIS THEOREM	With exactly $\dim V$ vectors: basis $\iff$ independent $\iff$ spanning. Count, then test one.

TRY IT YOURSELF

1. Independent or dependent? (a) $(1,2,3)^\mathsf{T}, (2,4,6)^\mathsf{T}$; (b) $(1,0,1)^\mathsf{T}, (0,1,1)^\mathsf{T}, (1,1,0)^\mathsf{T}$; (c) $1+x,\ 1-x,\ x^2$ in P_2.
2. Show that $\{(1,1,0)^\mathsf{T}, (0,1,1)^\mathsf{T}, (1,0,1)^\mathsf{T}\}$ is a basis of $\mathbb{R}^3$. Adding $(2,2,0)^\mathsf{T}$ still spans. Which old vector may be cast out, and why?
3. Are $1,\ 1+x,\ 1+x+x^2$ a basis of P_2? What is $\dim \mathrm{Mat}_{2\times 3}(\mathbb{R})$?
4. Is $\{(1,2,1)^\mathsf{T}, (0,1,1)^\mathsf{T}, (2,1,-1)^\mathsf{T}\}$ a basis of $\mathbb{R}^3$? If not, exhibit a dependence.
5. Extend $\{(1,1,0)^\mathsf{T}\}$ to a basis of $\mathbb{R}^3$ and justify. Show that $H = \{p \in P_2 \mid p(1) = 0\}$ is a subspace and find $\dim H$.

ANSWERS

1. (a) dependent ($2\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$); (b) independent ($\mathrm{Det} = -2$); (c) independent ($a+b=0,\ a-b=0,\ c=0$).
2. $\mathrm{Det} = 2 \neq 0$: three independent vectors in $\mathbb{R}^3$, a basis. $(2,2,0)^\mathsf{T} = 2(1,1,0)^\mathsf{T}$, so $(1,1,0)^\mathsf{T}$ may be cast out (not the others).
3. Yes: three vectors, each of $1,\ x = (1+x) - 1,\ x^2 = (1+x+x^2) - (1+x)$ is reachable, so they span a 3-dimensional space; $\dim \mathrm{Mat}_{2\times 3}(\mathbb{R}) = 6$.
4. Three vectors, $\dim \mathbb{R}^3 = 3$; $\mathrm{Det} = 0$, so no. $2(1,2,1)^\mathsf{T} - 3(0,1,1)^\mathsf{T} = (2,1,-1)^\mathsf{T}$.
5. Add $(0,1,0)^\mathsf{T}, (0,0,1)^\mathsf{T}$ ($\mathrm{Det} = 1$, Basis Theorem). $H = \mathrm{Ker}(p \mapsto p(1))$, basis $\{x-1, x^2-1\}$, $\dim H = 2$.

CHAPTER 11

Bases of kernel and image

LECTURES 10–11 · *how to produce a basis on demand, and the count* $\dim \operatorname{Col} A + \dim \operatorname{Ker} A = n$

THIS CHAPTER TEACHES HOW TO

- write down a basis of $\operatorname{Ker} A$ and state $\dim \operatorname{Ker} A$;
- write down a basis of $\operatorname{Col} A$ and state $\dim \operatorname{Col} A$;
- read the relations among the columns of A from $\operatorname{RREF}(A)$;
- check the count $\dim \operatorname{Col} A + \dim \operatorname{Ker} A = n$.

11.1 Basis of the kernel

THE RECIPE OF CHAPTER 7 ALREADY GIVES A BASIS

The spanning vectors from $\operatorname{RREF}(A)$ (one per free variable) are independent: each carries a 1 in its own free slot and 0 in the other free slots. So they form a *basis* of $\operatorname{Ker} A$ and $\dim \operatorname{Ker} A =$ number of free columns $= n - r$.

EXAMPLE 11.1 · basis of a kernel

$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 7 \\ 1 & 2 & 1 & 4 \end{bmatrix} \to \operatorname{RREF}(A) = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Pivots in columns $1, 3$; $x_2 = s$, $x_4 = t$ free; $x_1 = -2s - 3t$, $x_3 = -t$:

$$\operatorname{Ker} A = \operatorname{Span}\{(-2, 1, 0, 0)^{\mathsf{T}},\ (-3, 0, -1, 1)^{\mathsf{T}}\}, \qquad \dim \operatorname{Ker} A = 2.$$

Free slots (positions 2 and 4) show $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ against $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$: independent. Likewise $\operatorname{Ker} \left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 2 & 6 & 1 & 5 \end{smallmatrix}\right]$ has basis $\{(-3, 1, 0, 0)^{\mathsf{T}}, (-2, 0, -1, 1)^{\mathsf{T}}\}$, dimension 2.

ANIMATION: the kernel, step by step ▸.

11.2 Basis of the column space

RELATIONS AMONG COLUMNS SURVIVE ROW OPERATIONS

A relation $c_1 a_{-,1} + \cdots + c_n a_{-,n} = \mathbf{0}$ says $\mathbf{c} \in \operatorname{Ker} A$, and $\operatorname{Ker} A = \operatorname{Ker} \operatorname{RREF}(A)$. So A and $\operatorname{RREF}(A)$ satisfy *exactly the same* relations among their columns. In $\operatorname{RREF}(A)$, the pivot columns are $\mathbf{e}_1, \mathbf{e}_2, \ldots$ (independent), and each non-pivot column is a combination of the pivot columns to its left, with its own entries as coefficients.

BASIS OF $\operatorname{Col} A$

The columns of the original A in the positions of the pivot columns of $\operatorname{RREF}(A)$ form a basis of $\operatorname{Col} A$. Do *not* take the pivot columns of $\operatorname{RREF}(A)$ themselves (they say which positions, not which vectors). Hence $\dim \operatorname{Col} A = r$, the number of pivots (the **rank**).

BASES FOR $\operatorname{Col} A$ AND $\operatorname{Ker} A$ FROM ONE REDUCTION

1. Reduce A to $\operatorname{RREF}(A)$; mark the pivot columns.
2. $\operatorname{Col} A$: take those columns *of* A.
3. $\operatorname{Ker} A$: parametrise the free variables as in Chapter 7.
4. Check $r + (n - r) = n$.

A discarded column is a combination of the kept ones; the coefficients are read from the reduced form.

EXAMPLE 11.2 · both bases

For the A above, pivots in columns $1, 3$, so $\operatorname{Col} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\}$, dimension 2 (a plane). The other columns: $\begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (coefficients 2 and 3, 1 from $\operatorname{RREF}(A)$). Count: $2 + 2 = 4$ columns.

$A = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 2 & 6 & 1 & 5 \\ 1 & 3 & 1 & 3 \end{bmatrix}$ reduces to $\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$: $\operatorname{Col} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\}$, $\operatorname{Ker} A = \operatorname{Span}\{(-3,1,0,0)^{\mathsf{T}}, (-2,0,-1,1)^{\mathsf{T}}\}$, and $2 + 2 = 4$.

ANIMATION: the image: pivot columns and the relations among the others ▸.

PIVOT COLUMNS OF WHICH MATRIX?

Row operations change the column space. The basis of $\operatorname{Col} A$ consists of columns of A, never of $\operatorname{RREF}(A)$.

THE CHAPTER AT A GLANCE

$\operatorname{Ker} A$	Free-variable vectors form a basis; $\dim = n - r$.
$\operatorname{Col} A$	Columns of A at the pivot positions form a basis; $\dim = r$.
COUNT	$\dim \operatorname{Col} A + \dim \operatorname{Ker} A = n$: each column is a pivot or a free column.
RELATIONS	A and $\operatorname{RREF}(A)$ have the same column relations; non-pivot entries are the coefficients.

TRY IT YOURSELF

1. Find a basis of $\operatorname{Ker} A$ for $A = \left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 2 & 6 & 1 & 5 \end{smallmatrix}\right]$, and state $\dim \operatorname{Ker} A$.
2. For $A = \left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 2 & 6 & 1 & 5 \\ 1 & 3 & 1 & 3 \end{smallmatrix}\right]$ find $\operatorname{RREF}(A)$, a basis of $\operatorname{Col} A$ and a basis of $\operatorname{Ker} A$. Check the count.
3. Are $(1, 2, 1)^{\mathsf{T}}, (0, 1, 1)^{\mathsf{T}}, (1, 1, 0)^{\mathsf{T}}$ independent? What is the dimension of their span?

ANSWERS

1. $R_2 - 2R_1$ gives $\left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{smallmatrix}\right]$. Pivots $1, 3$; $x_2 = s$, $x_4 = t$, $x_1 = -3s - 2t$, $x_3 = -t$: basis $\{(-3, 1, 0, 0)^{\mathsf{T}}, (-2, 0, -1, 1)^{\mathsf{T}}\}$, $\dim = 2$.
2. $\operatorname{RREF}(A) = \left[\begin{smallmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right]$; $\operatorname{Col} A$ basis $\{(1, 2, 1)^{\mathsf{T}}, (0, 1, 1)^{\mathsf{T}}\}$; $\operatorname{Ker} A$ basis as in 1; $2 + 2 = 4$.
3. $\operatorname{Det} = 1(0-1) - 0 + 1(2-1) = 0$: dependent, since $(1, 2, 1)^{\mathsf{T}} - (0, 1, 1)^{\mathsf{T}} = (1, 1, 0)^{\mathsf{T}}$. Cast out the third: two non-parallel vectors, so the span is a plane, dimension 2.

CHAPTER 12

Coordinates

LECTURE 11 · *a basis is a coordinate system*

THIS CHAPTER TEACHES HOW TO

- find the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of a vector in $\mathbb{R}^n$, P_n, $\mathrm{Mat}_{m\times n}(\mathbb{R})$;
- explain why coordinates exist and are unique.

UNIQUE REPRESENTATION

If $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ is a basis of V, every $\mathbf{x} \in V$ can be written as $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$ in exactly one way: spanning gives existence, independence gives uniqueness. The **coordinate vector** is

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n.$$

The basis is an *ordered list*, since order matters. Coordinates belong to the pair (vector, basis), not to the vector alone.

HOW TO FIND $[\mathbf{x}]_{\mathcal{B}}$

1. Solve $c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{x}$ for the c_i: a linear system whose matrix has the $\mathbf{b}_i$ as *columns*.
2. In P_n compare coefficients (work from the top power down if the basis is triangular); in $\mathrm{Mat}_{m\times n}(\mathbb{R})$ compare entries.
3. Check by recombining $\sum c_i\mathbf{b}_i$.

For the standard basis nothing changes: $[\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$ in $\mathbb{R}^n$; $\left[a + bx + cx^2\right] = (a, b, c)^{\mathsf{T}}$ for $(1, x, x^2)$.

EXAMPLE 12.1 · coordinates

(a) $\mathcal{B} = \left(\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\right)$, $\mathbf{x} = \left[\begin{smallmatrix}5\\1\end{smallmatrix}\right]$: $a + b = 5$, $a - b = 1$, so $[\mathbf{x}]_{\mathcal{B}} = \left[\begin{smallmatrix}3\\2\end{smallmatrix}\right]$. With $\mathbf{x} = \left[\begin{smallmatrix}4\\2\end{smallmatrix}\right]$: $a = 3$, $b = 1$.
(b) $\mathcal{C} = (1,\ 1 + x,\ 1 + x + x^2)$, $p = 2 + 3x + 4x^2$: $c_3 = 4$, $c_2 + c_3 = 3 \Rightarrow c_2 = -1$, $c_1 + c_2 + c_3 = 2 \Rightarrow c_1 = -1$: $[p]_{\mathcal{C}} = (-1, -1, 4)^{\mathsf{T}}$. The same polynomial has coordinates $(2, 3, 4)^{\mathsf{T}}$ in $(1, x, x^2)$.
(c) $\mathcal{B} = \left(\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right], \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right]\right)$: $[\mathbf{e}_1]_{\mathcal{B}} = \left[\begin{smallmatrix}-5\\2\end{smallmatrix}\right]$ and $[\mathbf{e}_2]_{\mathcal{B}} = \left[\begin{smallmatrix}3\\-1\end{smallmatrix}\right]$ (check $3\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] - \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$). And for $3 - 2x + 5x^2$ in $\mathcal{C}$: $c_3 = 5$, $c_2 = -7$, $c_1 = 5$, i.e. $(5, -7, 5)^{\mathsf{T}}$.

ANIMATION: part (a): the point $(5, 1)$ read in this basis ▸.

CHECK

Recombine: $5(1) - 7(1+x) + 5(1+x+x^2) = 3 - 2x + 5x^2$. Always multiply the coordinates back onto the basis.

THE CHAPTER AT A GLANCE

UNIQUENESS	A basis gives each vector exactly one list of coefficients (existence: spans; uniqueness: independent).
COORDINATES	$[\mathbf{x}]_{\mathcal{B}}$ is that list; solve $[\mathbf{b}_1 \cdots \mathbf{b}_n]\mathbf{c} = \mathbf{x}$. Order of the basis matters.
STANDARD	$[\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$; $(1, x, x^2)$-coordinates of $a + bx + cx^2$ are $(a, b, c)^{\mathsf{T}}$.

TRY IT YOURSELF

1. $\mathcal{B} = \left(\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\5\end{bmatrix}\right)$ is a basis of $\mathbb{R}^2$. Find $[\mathbf{e}_1]_{\mathcal{B}}$ and $[\mathbf{e}_2]_{\mathcal{B}}$.
2. In P_2 with $\mathcal{C} = (1,\ 1+x,\ 1+x+x^2)$, find $\left[3 - 2x + 5x^2\right]_{\mathcal{C}}$.
3. $\mathcal{B} = \left(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right)$. Find $[\mathbf{x}]_{\mathcal{B}}$ for $\mathbf{x} = (4, 2)^{\mathsf{T}}$.

ANSWERS

1. $a + 3b = 1$, $2a + 5b = 0$: $b = 2$, $a = -5$, so $(-5, 2)^{\mathsf{T}}$. For $\mathbf{e}_2$: $b = -1$, $a = 3$, so $(3, -1)^{\mathsf{T}}$.
2. $c_3 = 5$, $c_2 + c_3 = -2 \Rightarrow c_2 = -7$, $c_1 + c_2 + c_3 = 3 \Rightarrow c_1 = 5$: $(5, -7, 5)^{\mathsf{T}}$.
3. $a + b = 4$, $a - b = 2$: $(3, 1)^{\mathsf{T}}$.

PART VI

Change of basis

CHAPTER 13

Isomorphism and change of basis

LECTURE 12 · *translating between coordinate systems, and when two spaces are "the same"*

THIS CHAPTER TEACHES HOW TO

- decide whether a linear map is injective, surjective, an isomorphism;
- convert coordinates between two bases with $P_{\mathcal{B}}$ and $P_{\mathcal{C}}$;
- compose linear maps via their matrices;
- use isomorphisms to move a problem into $\mathbb{R}^n$.

13.1 Injective, surjective, isomorphic

FOR MAPS OF SETS

$T\colon X \to Y$ is **injective** if $T(x) = T(x') \Rightarrow x = x'$; **surjective** if $\operatorname{Im} T = Y$; **bijective** if both, equivalently **invertible** (an inverse T^{-1} with $T^{-1}T = \mathrm{Id}$, $TT^{-1} = \mathrm{Id}$ exists, and is unique).

ISOMORPHISMS

An **isomorphism** is a bijective *linear* map $T\colon V \to W$; then $V \cong W$. Facts:

- the inverse of an isomorphism is automatically linear (an isomorphism);
- T is injective $\iff \operatorname{Ker} T = \{\mathbf{0}\}$. So T is an isomorphism iff $\operatorname{Ker} T = \{\mathbf{0}\}$ and $\operatorname{Im} T = W$;
- composites of linear maps are linear, and $[T \circ S] = [T]\,[S]$ ("do S, then T"). Composites of isomorphisms are isomorphisms, $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$.
- isomorphic spaces are "the same" as far as vector space notions go: subspaces, spans, independence, bases, dimension, kernels, images all cross over. Extra structure (length, angle) does not.

HOW TO SHOW $T\colon V \to W$ IS AN ISOMORPHISM

1. Check T is linear.
2. Compute $\operatorname{Ker} T$: only $\mathbf{0}$? (injective). Compute $\operatorname{Im} T$: is it all of W? (surjective).
3. Or: take bases, write the matrix of T column by column, and check $\operatorname{Det} \neq 0$ (square, same dimension).

EXAMPLE 13.1 · an isomorphism

$T\colon P_2 \to \mathbb{R}^3, T(p) = (p(0), p(1), p(2))^{\mathsf{T}}$ is linear (evaluation is). On the basis $(1, x, x^2)$: $T(1) = (1,1,1)^{\mathsf{T}}$, $T(x) = (0,1,2)^{\mathsf{T}}$, $T(x^2) = (0,1,4)^{\mathsf{T}}$, with $\operatorname{Det}\left[\begin{smallmatrix}1&0&0\\1&1&1\\1&2&4\end{smallmatrix}\right] = 4 - 2 = 2 \neq 0$. So T is an isomorphism (a quadratic is determined by its values at three points).

13.2 The coordinate isomorphism and change of basis

COORDINATE MAP

For a basis $\mathcal{B}$ of an n-dimensional V, $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is an isomorphism $V \to \mathbb{R}^n$. So $P_2 \cong \mathbb{R}^3$, $\mathrm{Mat}_{2\times 2}(\mathbb{R}) \cong \mathbb{R}^4$, $\mathbb{C} \cong \mathbb{R}^2$ (over $\mathbb{R}$): choose a basis and the space becomes $\mathbb{R}^n$.

CHANGE OF BASIS IN $\mathbb{R}^n$

With $P_{\mathcal{B}} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$ (the basis vectors as columns),

$$\mathbf{x} = P_{\mathcal{B}}\,[\mathbf{x}]_{\mathcal{B}}\,, \qquad [\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}\mathbf{x}.$$

$P_{\mathcal{B}}$ turns $\mathcal{B}$-coordinates into standard ones; its inverse goes the other way. To convert $\mathcal{B}$-coordinates into $\mathcal{C}$-coordinates, go *through* the standard basis:

$$P_{\mathcal{B}}^{\mathcal{C}} = P_{\mathcal{C}}^{-1}P_{\mathcal{B}}, \qquad [\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{B}}^{\mathcal{C}}\,[\mathbf{x}]_{\mathcal{B}}\,, \qquad P_{\mathcal{C}}^{\mathcal{B}} = (P_{\mathcal{B}}^{\mathcal{C}})^{-1}.$$

ORDER AND DIRECTION

$P_{\mathcal{B}}^{\mathcal{C}}$ takes $\mathcal{B}$-coordinates *to* $\mathcal{C}$-coordinates: first undo $\mathcal{B}$ ($P_{\mathcal{B}}$), then translate to $\mathcal{C}$ ($P_{\mathcal{C}}^{-1}$). The product is written in that reverse order. When unsure, test with a vector you know.

EXAMPLE 13.2 · change of basis

$\mathcal{B} = \left(\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\right)$, $\mathcal{C} = \left(\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right)$. $P_{\mathcal{B}} = \left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right]$, $P_{\mathcal{C}} = \left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$, $P_{\mathcal{C}}^{-1} = \left[\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right]$, so

$$P_{\mathcal{B}}^{\mathcal{C}} = \left[\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right]\left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right] = \left[\begin{smallmatrix}0&2\\1&-1\end{smallmatrix}\right].$$

Check: $[\mathbf{x}]_{\mathcal{B}} = (1,0)^{\mathsf{T}}$ means $\mathbf{x} = (1,1)^{\mathsf{T}} = 0\cdot(1,0)^{\mathsf{T}} + 1\cdot(1,1)^{\mathsf{T}}$, and $P_{\mathcal{B}}^{\mathcal{C}}(1,0)^{\mathsf{T}} = (0,1)^{\mathsf{T}}$.
Another: $\mathcal{B}$ as above, $\mathcal{C} = \left(\left[\begin{smallmatrix}2\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right]\right)$: $P_{\mathcal{C}}^{-1} = \frac{1}{6}\left[\begin{smallmatrix}3&-1\\0&2\end{smallmatrix}\right]$ and $P_{\mathcal{B}}^{\mathcal{C}} = \frac{1}{3}\left[\begin{smallmatrix}1&2\\1&-1\end{smallmatrix}\right]$. For $[\mathbf{x}]_{\mathcal{B}} = (2,1)^{\mathsf{T}}$: $[\mathbf{x}]_{\mathcal{C}} = \frac{1}{3}(4,1)^{\mathsf{T}}$. Direct check: $\mathbf{x} = 2(1,1)^{\mathsf{T}} + (1,-1)^{\mathsf{T}} = (3,1)^{\mathsf{T}} = \frac{4}{3}(2,0)^{\mathsf{T}} + \frac{1}{3}(1,3)^{\mathsf{T}}$.

LENGTHS ARE NOT PRESERVED

Changing basis does not (in general) preserve lengths, angles, areas: $\mathbf{b}_1 = (2,1)^{\mathsf{T}}$ has length $\sqrt{5}$, but $[\mathbf{b}_1]_{\mathcal{B}} = (1,0)^{\mathsf{T}}$ has length 1. Length is not a vector space notion. Isomorphisms preserve only what is expressed with $+$ and scaling.

EXAMPLE 13.3 · move a problem to $\mathbb{R}^n$

Are $1+x$, $1-x$, x^2 independent in P_2? Their coordinates in $(1, x, x^2)$ are $(1,1,0)^\mathsf{T}, (1,-1,0)^\mathsf{T}, (0,0,1)^\mathsf{T}$, with $\mathrm{Det} = -2 \neq 0$: independent, no polynomial manipulation needed.

EXAMPLE 13.4 · composition

$S = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ (rotation), $T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ (shear). $[T \circ S] = TS = \left[\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, $[S \circ T] = ST = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right]$: they differ. $\mathrm{Det}(TS) = 1$, so $(TS)^{-1} = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right] = S^{-1}T^{-1}$.

ANIMATION: rotate then shear, and shear then rotate ▸.

THE CHAPTER AT A GLANCE

ISOMORPHISM	Bijective linear map; inverse is linear. Linear T is injective iff $\mathrm{Ker}\, T = \{\mathbf{0}\}$; isomorphism iff also $\mathrm{Im}\, T = W$.
COORDINATES	$\mathbf{x} \mapsto [\mathbf{x}]_\mathcal{B}$ is an isomorphism $V \to \mathbb{R}^n$. Isomorphic spaces share all vector space properties, not length.
CHANGE	$P_\mathcal{B} = [\mathbf{b}_1 \cdots \mathbf{b}_n]$; $P_\mathcal{B}^\mathcal{C} = P_\mathcal{C}^{-1} P_\mathcal{B}$; $P_\mathcal{C}^\mathcal{B} = (P_\mathcal{B}^\mathcal{C})^{-1}$.
COMPOSITION	$[T \circ S] = [T]\,[S]$; order matters; $(TS)^{-1} = S^{-1}T^{-1}$.

TRY IT YOURSELF

1. With $\mathcal{B} = \left(\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]\right)$ and $\mathcal{C} = \left(\left[\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right]\right)$, find $P_\mathcal{B}^\mathcal{C}$ and use it to convert $[\mathbf{x}]_\mathcal{B} = (2,1)^\mathsf{T}$ to $\mathcal{C}$-coordinates. Check directly.
2. Show that $T\colon P_2 \to \mathbb{R}^3$, $T(p) = (p(0), p(1), p(2))^\mathsf{T}$, is an isomorphism.
3. Compute $[T \circ S]$ and $[S \circ T]$ for $S = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, $T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, confirm they differ, and find $[T \circ S]^{-1}$.

ANSWERS

1. $P_\mathcal{B}^\mathcal{C} = \frac{1}{6}\left[\begin{smallmatrix} 3 & -1 \\ 0 & 2 \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right] = \frac{1}{3}\left[\begin{smallmatrix} 1 & 2 \\ 1 & -1 \end{smallmatrix}\right]$; $P_\mathcal{B}^\mathcal{C}(2,1)^\mathsf{T} = \frac{1}{3}(4,1)^\mathsf{T}$. Direct: $\mathbf{x} = (3,1)^\mathsf{T} = \frac{4}{3}(2,0)^\mathsf{T} + \frac{1}{3}(1,3)^\mathsf{T}$.
2. Linear (evaluation is). Matrix on $(1, x, x^2)$: columns $(1,1,1)^\mathsf{T}, (0,1,2)^\mathsf{T}, (0,1,4)^\mathsf{T}$, $\mathrm{Det} = 2 \neq 0$: invertible, so isomorphism.
3. $TS = \left[\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, $ST = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right]$: different. $(TS)^{-1} = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right]$ ($\mathrm{Det} = 1$).

PART VII

Rank and the fundamental subspaces

CHAPTER 14

Rank and nullity

LECTURE 14 · *one equation that ties the module together*

THIS CHAPTER TEACHES HOW TO

- compute $\operatorname{Rank} A$ and $\operatorname{Nullity} A$ and use $\operatorname{Rank} A + \operatorname{Nullity} A = n$;
- apply rank–nullity to linear maps (polynomials, matrices);
- use the invertibility equivalences for square matrices;
- describe the solution set of $A\mathbf{x} = \mathbf{b}$ and bound possible ranks.

RANK AND NULLITY

For an $m \times n$ matrix A:

$$\operatorname{Rank} A = \dim(\operatorname{Im} A) = \#\text{pivots}, \qquad \operatorname{Nullity} A = \dim(\operatorname{Ker} A) = \#\text{free columns},$$

$$\boxed{\operatorname{Rank} A + \operatorname{Nullity} A = n} \qquad (n = \text{number of columns}).$$

Every column is a pivot column or a free column. Abstractly, for linear $T\colon V \to W$ with $\dim V < \infty$: $\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V$ (the *source*).

HOW TO FIND RANK AND NULLITY

1. Row-reduce to $\operatorname{RREF}(A)$; count pivots $= r = \operatorname{Rank} A$.
2. $\operatorname{Nullity} A = n - r$. Bases: pivot columns of A for $\operatorname{Im} A$; free-variable vectors for $\operatorname{Ker} A$ (Chapter 11).
3. Check $r + (n - r) = n$.

EXAMPLE 14.1 · computing

$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 10 \\ 1 & 2 & 4 & 6 \end{bmatrix}$ reduces to $\begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$: pivots in columns $1, 3$, so $\operatorname{Rank} A = 2$, $\operatorname{Im} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (3,7,4)^{\mathsf{T}}\}$. Free $x_2 = s$, $x_4 = t$: $x_1 = -2s + 2t$, $x_3 = -2t$, $\operatorname{Ker} A = \operatorname{Span}\{(-2,1,0,0)^{\mathsf{T}}, (2,0,-2,1)^{\mathsf{T}}\}$, $\operatorname{Nullity} A = 2$, and $2 + 2 = 4$.

Abstract example: $D\colon P_2 \to P_1$, $p \mapsto p'$: $\operatorname{Ker} D$ = constants (dim 1), $\operatorname{Im} D = P_1$ (dim 2), $1 + 2 = 3 = \dim P_2$.

ANIMATION: rank–nullity for this matrix ▸.

INVERTIBILITY, EVERYTHING AT ONCE

For an $n \times n$ matrix A the following are equivalent: A invertible $\iff$ $\operatorname{Det} A \neq 0$ $\iff$ $\operatorname{RREF}(A) = I$ $\iff$ $\operatorname{Rank} A = n$ $\iff$ $\operatorname{Nullity} A = 0$ $\iff$ the columns form a basis of $\mathbb{R}^n$ $\iff$ $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $\mathbf{b}$.

SOLVING $A\mathbf{x} = \mathbf{b}$ THROUGH THE TWO SUBSPACES

- A solution exists iff $\mathbf{b} \in \operatorname{Im} A$. If $\operatorname{Rank} A = m$ then $\operatorname{Im} A = \mathbb{R}^m$ and a solution always exists.
- If one solution $\mathbf{x}_0$ exists, the full solution set is $\mathbf{x}_0 + \operatorname{Ker} A$. The solution is unique iff $\operatorname{Nullity} A = 0$.
- Reducing $[A \mid \mathbf{b}]$ shows the rank (pivots), the nullity (free columns) and consistency (is there a pivot in the $\mathbf{b}$ column?).

TRANSPOSE AND SIZE OF THE RANK

$\operatorname{Rank} A^{\mathsf{T}} = \operatorname{Rank} A$ (row rank equals column rank), so

$$\operatorname{Rank} A \leq \min(m, n).$$

In particular, if $n > m$ then $\operatorname{Nullity} A \geq n - m > 0$: a wide matrix always has a non-trivial kernel.

EXAMPLE 14.2 · sizes

Largest rank of a 4×7 matrix: $\min(4, 7) = 4$; of a 7×4: also 4. If $\operatorname{Ker} A$ of a 4×7 matrix has dimension 3, then $\operatorname{Rank} A = 7 - 3 = 4$ (maximal). A 4×7 matrix cannot have nullity 2 (that would force rank $5 > 4$).

RANK COUNTS *COLUMNS* OF THE SAME MATRIX

$\operatorname{Rank} + \operatorname{Nullity} = n$, the number of *columns* (dimension of the source), not m. The kernel lives in $\mathbb{R}^n$, the image in $\mathbb{R}^m$.

THE CHAPTER AT A GLANCE

DEFINITIONS	$\operatorname{Rank} A = \dim \operatorname{Im} A$ =#pivots; $\operatorname{Nullity} A = \dim \operatorname{Ker} A$ =#free columns; $\operatorname{Rank} + \operatorname{Nullity} = n$. For T: $\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V$.
SQUARE	Invertible $\iff$ $\operatorname{Det} \neq 0$ $\iff$ $\operatorname{Rank} = n$ $\iff$ $\operatorname{Nullity} = 0$ $\iff$ columns a basis $\iff$ unique solution for all $\mathbf{b}$.
SYSTEMS	Solvable iff $\mathbf{b} \in \operatorname{Im} A$; solution set $\mathbf{x}_0 + \operatorname{Ker} A$; unique iff $\operatorname{Nullity} A = 0$.
BOUND	$\operatorname{Rank} A = \operatorname{Rank} A^{\mathsf{T}} \leq \min(m, n)$; wide matrices have non-trivial kernel.

TRY IT YOURSELF

1. For $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 10 \\ 1 & 2 & 4 & 6 \end{bmatrix}$ find $\operatorname{Rank} A$, $\operatorname{Nullity} A$, a basis of $\operatorname{Col} A$ and of $\operatorname{Ker} A$, and verify rank–nullity.
2. (a) Can a 3×5 matrix have nullity 1? (b) Can a 5×3 matrix have rank 4? (c) A 6×4 matrix has $\operatorname{Nullity} A = 0$: what is $\operatorname{Rank} A$?

3. A is 5×5 with $\operatorname{Rank} A = 5$. Explain, using $\operatorname{Col} A$ and $\operatorname{Ker} A$ separately, why $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $\mathbf{b}$.

ANSWERS

1. Computed above: $\operatorname{Rank} = 2$, $\operatorname{Nullity} = 2$; $\operatorname{Col} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (3,7,4)^{\mathsf{T}}\}$; $\operatorname{Ker} A = \operatorname{Span}\{(-2,1,0,0)^{\mathsf{T}}, (2,0,-2,1)^{\mathsf{T}}\}$.
2. (a) No: $\operatorname{Rank} \leq 3$ so $\operatorname{Nullity} \geq 2$. (b) No: $\operatorname{Rank} \leq \min(5,3) = 3$. (c) $\operatorname{Rank} = 4 - 0 = 4$: columns independent, but they cannot span $\mathbb{R}^6$.
3. $\dim \operatorname{Im} A = 5 = \dim \mathbb{R}^5$, so $\operatorname{Im} A = \mathbb{R}^5$: every $\mathbf{b}$ is reached (existence). $\operatorname{Nullity} A = 5 - 5 = 0$, so $\operatorname{Ker} A = \{\mathbf{0}\}$; if $A\mathbf{x} = A\mathbf{y}$ then $A(\mathbf{x} - \mathbf{y}) = \mathbf{0}$, so $\mathbf{x} = \mathbf{y}$ (uniqueness).

CHAPTER 15

The four fundamental subspaces

LECTURE 15 · *row space, cokernel, and the full dimension count*

THIS CHAPTER TEACHES HOW TO

- find a basis of the row space, and of the cokernel;
- state the dimensions of all four subspaces and the spaces they live in;
- compute all four for a given matrix.

ROW SPACE

$\operatorname{Row} A = \operatorname{Span}\{\text{rows of } A\} \subseteq \mathbb{R}^n$ (rows written as columns, it is $\operatorname{Im} A^{\mathsf{T}}$). Row operations preserve the row space (the opposite of the column space), so $\operatorname{Row} A = \operatorname{Row} \operatorname{RREF}(A)$, and

$$\text{a basis of } \operatorname{Row} A : \text{ the non-zero rows of } \operatorname{RREF}(A).$$

Hence $\dim \operatorname{Row} A = \#\text{pivots} = \operatorname{Rank} A = \dim \operatorname{Col} A$: row rank equals column rank, and $\operatorname{Rank} A^{\mathsf{T}} = \operatorname{Rank} A$.

COKERNEL

$\operatorname{Coker} A = \{\mathbf{y} \mid \mathbf{y}A = \mathbf{0}\} \subseteq \mathbb{R}^m$ (row vectors, equivalently $\mathbf{y}^{\mathsf{T}} \in \operatorname{Ker} A^{\mathsf{T}}$). Compute it as $\operatorname{Ker} A^{\mathsf{T}}$: transpose, reduce, parametrise. It measures what A misses (more abstractly $\operatorname{Coker} A = \mathbb{R}^m / \operatorname{Im} A$).

THE FOUR SUBSPACES

For an $m \times n$ matrix of rank r:

subspace	lives in	dimension
$\operatorname{Row} A$	$\mathbb{R}^n$ (source)	r
$\operatorname{Ker} A$	$\mathbb{R}^n$ (source)	$n - r$
$\operatorname{Im} A = \operatorname{Col} A$	$\mathbb{R}^m$ (target)	r
$\operatorname{Coker} A$	$\mathbb{R}^m$ (target)	$m - r$

In each space the two dimensions add up to that space's dimension ($r+(n-r) = n$, $r+(m-r) = m$). A carries $\operatorname{Row} A$ onto $\operatorname{Im} A$ and sends $\operatorname{Ker} A$ to $\mathbf{0}$.

ORTHOGONALITY OF THE FOUR SUBSPACES (CHAPTER 16)

In the source, $(\operatorname{Row} A)^{\perp} = \operatorname{Ker} A$; in the target, $(\operatorname{Im} A)^{\perp} = \operatorname{Coker} A$. So $\mathbb{R}^n = \operatorname{Row} A \oplus \operatorname{Ker} A$ and $\mathbb{R}^m = \operatorname{Im} A \oplus \operatorname{Coker} A$ as orthogonal pieces; e.g. $\operatorname{Coker} A$ is the set of vectors perpendicular to every column of A.

ALL FOUR FROM TWO REDUCTIONS

1. Reduce A: pivots give r. $\operatorname{Im} A$: pivot columns *of* A. $\operatorname{Row} A$: non-zero rows of $\operatorname{RREF}(A)$. $\operatorname{Ker} A$: free variables.
2. Reduce A^{T} and read off the kernel: that is $\operatorname{Coker} A$.
3. Check the four dimensions: $r,\ r,\ n-r,\ m-r$.

EXAMPLE 15.1 · all four

$A = \left[\begin{smallmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 1 & 3 \\ 1 & 2 & 1 & 2 \end{smallmatrix}\right]$ $(m = 3, n = 4)$ has $\operatorname{RREF}(A) = \left[\begin{smallmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right]$, so $r = 2$.

- $\operatorname{Im} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\}$, $\dim 2$.
- $\operatorname{Row} A = \operatorname{Span}\{(1,2,0,1), (0,0,1,1)\}$, $\dim 2$.
- $\operatorname{Ker} A = \operatorname{Span}\{(-2,1,0,0)^{\mathsf{T}}, (-1,0,-1,1)^{\mathsf{T}}\}$, $\dim 2 = n - r$.
- $\operatorname{Coker} A$: $A^{\mathsf{T}} = \left[\begin{smallmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 2 \end{smallmatrix}\right] \to \left[\begin{smallmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$, so $z = t$, $x = t$, $y = -t$: $\operatorname{Coker} A = \operatorname{Span}\{(1,-1,1)\}$, $\dim 1 = m - r$. Check: $(1,-1,1)A = (1-2+1,\ 2-4+2,\ 0-1+1,\ 1-3+2) = \mathbf{0}$.

CHECKS

Dimensions must be $r, r, n-r, m-r$. Every vector of $\operatorname{Ker} A$ satisfies $A\mathbf{x} = \mathbf{0}$; every $\mathbf{y} \in \operatorname{Coker} A$ satisfies $\mathbf{y}A = \mathbf{0}$ (a combination of the rows of A that vanishes). $\dim \operatorname{Coker} A = 0$ means $\operatorname{Rank} A = m$, so $A\mathbf{x} = \mathbf{b}$ is solvable for every $\mathbf{b}$.

EXAMPLE 15.2 · a matrix of a map

$T\colon P_2 \to P_2$, $T(p) = p' + p$: $T(1) = 1$, $T(x) = 1 + x$, $T(x^2) = 2x + x^2$; in the basis $(1, x, x^2)$ the matrix is $\left[\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{smallmatrix}\right]$, triangular with $\operatorname{Det} = 1 \neq 0$: invertible, so T is an isomorphism.

THE CHAPTER AT A GLANCE

ROW SPACE	$\operatorname{Row} A = \operatorname{Im} A^{\mathsf{T}}$; preserved by row operations; basis: non-zero rows of $\operatorname{RREF}(A)$; $\dim = r$, so $\operatorname{Rank} A^{\mathsf{T}} = \operatorname{Rank} A$.
COKERNEL	$\operatorname{Coker} A = \{\mathbf{y} \mid \mathbf{y}A = \mathbf{0}\} = \operatorname{Ker} A^{\mathsf{T}}$; $\dim = m - r$.
FOUR	In $\mathbb{R}^n$: $\operatorname{Row} A$ (r), $\operatorname{Ker} A$ $(n-r)$. In $\mathbb{R}^m$: $\operatorname{Im} A$ (r), $\operatorname{Coker} A$ $(m-r)$.

TRY IT YOURSELF

1. For $A = \left[\begin{smallmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 1 & 0 & 1 \end{smallmatrix}\right]$ find $\operatorname{Rank} A$ and a basis of each of the four fundamental subspaces. Check the four dimensions.
2. An $m \times n$ matrix has $\dim \operatorname{Coker} A = 0$. What does that say about $\operatorname{Rank} A$ and about solvability of $A\mathbf{x} = \mathbf{b}$?

3. $T\colon P_2 \to P_2$, $T(p) = p' + p$: write its matrix in the basis $(1, x, x^2)$ and decide whether T is an isomorphism.

ANSWERS

1. $\mathrm{RREF}(A) = \left[\begin{smallmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right]$, $r = 2$. $\operatorname{Im} A = \operatorname{Span}\{(1,2,1)^{\mathsf{T}}, (1,2,0)^{\mathsf{T}}\}$; $\operatorname{Row} A = \operatorname{Span}\{(1,0,1), (0,1,1)\}$; $\operatorname{Ker} A = \operatorname{Span}\{(-1,-1,1)^{\mathsf{T}}\}$; $\operatorname{Coker} A = \operatorname{Ker} A^{\mathsf{T}} = \operatorname{Span}\{(2,-1,0)\}$. Dimensions $2, 2, 1, 1$ ($= r, r, n-r, m-r$).
2. $m - r = 0$, so $\operatorname{Rank} A = m$; $\operatorname{Im} A$ has full dimension, hence $\operatorname{Im} A = \mathbb{R}^m$, and $A\mathbf{x} = \mathbf{b}$ is solvable for *every* $\mathbf{b}$.
3. $\left[\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{smallmatrix}\right]$, $\det = 1 \neq 0$: invertible, so T is an isomorphism.

PART VIII

Orthogonality

CHAPTER 16

Inner products and orthogonality

LECTURE 16 · *putting length and angle back into the space*

THIS CHAPTER TEACHES HOW TO

- compute inner products, lengths, unit vectors, angles and distances in $\mathbb{R}^n$;
- test orthogonality and use Pythagoras;
- project one vector onto another;
- compute the orthogonal complement $H^\perp$ of a subspace;
- use an orthogonal basis to find coordinates.

16.1 Inner product, length, angle

INNER PRODUCT AND LENGTH

$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + \cdots + u_nv_n = \mathbf{u}^{\mathsf{T}}\mathbf{v}$. It is symmetric, linear in each entry, and $\mathbf{u}\cdot\mathbf{u} \geq 0$ with equality only for $\mathbf{u} = \mathbf{0}$. The **length** is $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$; $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$. A **unit vector** has length 1; the **normalisation** of $\mathbf{v} \neq \mathbf{0}$ is $\hat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$. The **distance** is $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$.

ANGLE, ORTHOGONALITY

$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$, so $\cos\theta = \hat{\mathbf{u}}\cdot\hat{\mathbf{v}}$: positive for an acute angle, 0 for a right angle, negative for an obtuse one. $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal**, $\mathbf{u}\perp\mathbf{v}$, if $\mathbf{u}\cdot\mathbf{v} = 0$; $\mathbf{0}$ is orthogonal to everything and is the only vector orthogonal to itself. **Pythagoras:** $\mathbf{u}\perp\mathbf{v} \iff \|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$. **Cauchy–Schwarz:** $|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$; **triangle inequality:** $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

EXAMPLE 16.1 · computations

$\mathbf{u} = (1,2,2)^{\mathsf{T}}$, $\mathbf{v} = (2,-1,0)^{\mathsf{T}}$: $\mathbf{u}\cdot\mathbf{v} = 2-2+0 = 0$, so $\mathbf{u}\perp\mathbf{v}$. $\|\mathbf{u}\| = 3$, $\|\mathbf{v}\| = \sqrt{5}$, $\hat{\mathbf{u}} = \frac{1}{3}(1,2,2)^{\mathsf{T}}$, $\mathbf{u}-\mathbf{v} = (-1,3,2)^{\mathsf{T}}$ so $d(\mathbf{u},\mathbf{v}) = \sqrt{14}$, and Pythagoras checks: $9+5 = 14$. Also $(3,4)^{\mathsf{T}}$ has length 5, normalisation $(\frac{3}{5},\frac{4}{5})^{\mathsf{T}}$.

ANIMATION: the distance between these two vectors ▸.

16.2 Projection onto a vector

PROJECTION ONTO A LINE

For $\mathbf{u} \neq \mathbf{0}$, the **projection** of $\mathbf{v}$ on $\mathbf{u}$ is

$$\mathrm{Proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\,\mathbf{u}.$$

Then $\mathbf{v} = \mathbf{p} + \mathbf{w}$ with $\mathbf{p} = \mathrm{Proj}_{\mathbf{u}}(\mathbf{v})$ along $\mathbf{u}$ and $\mathbf{w} = \mathbf{v} - \mathbf{p} \perp \mathbf{u}$; also $\mathbf{p} \perp \mathbf{w}$, so $\|\mathbf{v}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{w}\|^2$ and $\|\mathbf{p}\| \leq \|\mathbf{v}\|$.

EXAMPLE 16.2 · projection

$\mathbf{v} = (3,4)^{\mathsf{T}}$, $\mathbf{u} = (1,0)^{\mathsf{T}}$: $\mathbf{p} = 3\mathbf{u} = (3,0)^{\mathsf{T}}$, $\mathbf{w} = (0,4)^{\mathsf{T}}$, and $9 + 16 = 25$. For $\mathbf{u} = (3,0)^{\mathsf{T}}$, $\mathbf{v} = (2,5)^{\mathsf{T}}$: $\mathbf{p} = \frac{6}{9}(3,0)^{\mathsf{T}} = (2,0)^{\mathsf{T}}$, $\mathbf{w} = (0,5)^{\mathsf{T}}$.

ANIMATION: the second projection ▸.

16.3 Orthogonal complements

ORTHOGONAL COMPLEMENT

For a subspace $H \subseteq \mathbb{R}^n$, $H^{\perp} = \{\mathbf{v} \mid \mathbf{v}\cdot\mathbf{h} = 0 \text{ for all } \mathbf{h} \in H\}$ is a subspace. To test $\mathbf{v} \perp H$, it suffices to test against a *spanning set* of H. In particular, if the rows of A span H,

$$(\mathrm{Row}\,A)^{\perp} = \mathrm{Ker}\,A.$$

HOW TO COMPUTE $H^{\perp}$

1. Write a spanning set $\mathbf{h}_1, \ldots, \mathbf{h}_k$ of H as the *rows* of a matrix A.
2. $H^{\perp} = \mathrm{Ker}\,A$: solve $A\mathbf{x} = \mathbf{0}$ (Chapters 7, 11); write as a span.
3. Check $\dim H + \dim H^{\perp} = n$ and that each spanning vector is orthogonal to each $\mathbf{h}_i$.

EXAMPLE 16.3 · complements

$H = \mathrm{Span}\{(1,1,0)^{\mathsf{T}}, (0,1,1)^{\mathsf{T}}\} \subseteq \mathbb{R}^3$: $x + y = 0$, $y + z = 0$, so $H^{\perp} = \mathrm{Span}\{(-1,1,-1)^{\mathsf{T}}\}$, $2 + 1 = 3$.

$H = \mathrm{Span}\{(1,0,1,0)^{\mathsf{T}}, (0,1,0,1)^{\mathsf{T}}\} \subseteq \mathbb{R}^4$: $x_1 + x_3 = 0$, $x_2 + x_4 = 0$, so $H^{\perp} = \mathrm{Span}\{(-1,0,1,0)^{\mathsf{T}}, (0,-1,0,1)^{\mathsf{T}}\}$, $2 + 2 = 4$.

ANIMATION: the first complement: the vector perpendicular to both ▸.

16.4 Orthogonal sets

ORTHOGONAL AND ORTHONORMAL SETS

A set of non-zero vectors is **orthogonal** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for $i \neq j$, **orthonormal** if moreover all $\|\mathbf{u}_i\| = 1$ (normalise each vector). An orthogonal set is automatically linearly independent; n of them in $\mathbb{R}^n$ form a basis. In an *orthogonal* basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ the coordinates are one inner product each:

$$\mathbf{v} = \sum c_i \mathbf{u}_i, \qquad c_i = \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}.$$

EXAMPLE 16.4 · an orthogonal basis

$\{(1,1)^\mathsf{T}, (1,-1)^\mathsf{T}\}$ is orthogonal, not orthonormal; divide each by $\sqrt{2}$. $\mathbf{v} = (4,2)^\mathsf{T}$: $c_1 = \frac{6}{2} = 3$, $c_2 = \frac{2}{2} = 1$, i.e. $[\mathbf{v}]_\mathcal{B} = (3,1)^\mathsf{T}$, without solving a system.

ANIMATION: coordinates as dot products ▸.

THE CHAPTER AT A GLANCE

BASICS	$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\mathsf{T}\mathbf{v}$; $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$; $\cos\theta = \mathbf{u} \cdot \mathbf{v}/(\|\mathbf{u}\|\|\mathbf{v}\|)$; $\mathbf{u} \perp \mathbf{v} \iff \mathbf{u} \cdot \mathbf{v} = 0$.
PYTHAGORAS	$\mathbf{u} \perp \mathbf{v} \iff \|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.
PROJECTION	$\mathrm{Proj}_\mathbf{u}\,\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$; remainder is orthogonal to $\mathbf{u}$.
COMPLEMENT	$H^\perp$ is a subspace; test on spanning vectors; $(\mathrm{Row}\,A)^\perp = \mathrm{Ker}\,A$; $\dim H + \dim H^\perp = n$.
ORTHOGONAL SETS	Independent. In an orthogonal basis $c_i = \mathbf{v} \cdot \mathbf{u}_i / \mathbf{u}_i \cdot \mathbf{u}_i$.

TRY IT YOURSELF

1. For $\mathbf{u} = (1,2,2)^\mathsf{T}$, $\mathbf{v} = (2,-1,0)^\mathsf{T}$: compute $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, $d(\mathbf{u},\mathbf{v})$, $\hat{\mathbf{u}}$. Are they orthogonal?
2. Write $\mathbf{v} = (3,4)^\mathsf{T}$ as $\mathbf{p} + \mathbf{w}$ with $\mathbf{p}$ a multiple of $\mathbf{u} = (1,0)^\mathsf{T}$ and $\mathbf{w} \perp \mathbf{u}$. Verify Pythagoras.
3. Find $H^\perp$ for $H = \mathrm{Span}\{(1,0,1,0)^\mathsf{T}, (0,1,0,1)^\mathsf{T}\} \subseteq \mathbb{R}^4$, and check $\dim H + \dim H^\perp = 4$.

ANSWERS

1. $2 - 2 + 0 = 0$: orthogonal. $\|\mathbf{u}\| = 3$, $\|\mathbf{v}\| = \sqrt{5}$, $d = \sqrt{1+9+4} = \sqrt{14}$, $\hat{\mathbf{u}} = \frac{1}{3}(1,2,2)^\mathsf{T}$. $(9 + 5 = 14.)$
2. $\mathbf{p} = (3,0)^\mathsf{T}$, $\mathbf{w} = (0,4)^\mathsf{T}$; $9 + 16 = 25$.
3. $H^\perp = \mathrm{Span}\{(-1,0,1,0)^\mathsf{T}, (0,-1,0,1)^\mathsf{T}\}$; $2 + 2 = 4$.

CHAPTER 17

Orthogonal projection and Gram–Schmidt

LECTURE 17 · *splitting a vector into a part in W and a part perpendicular to it*

THIS CHAPTER TEACHES HOW TO

- compute $\mathrm{Proj}_W(\mathbf{v})$ and the perpendicular part $\mathbf{z} \in W^\perp$;
- run Gram–Schmidt on a basis and normalise;
- recognise and use orthogonal matrices.

17.1 Projection onto a subspace

ORTHOGONAL PROJECTION AND DECOMPOSITION

If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $W \subseteq \mathbb{R}^n$, then

$$\mathrm{Proj}_W(\mathbf{v}) = \sum_{i=1}^{k} \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}\, \mathbf{u}_i.$$

Orthogonal Decomposition Theorem: every $\mathbf{v} \in \mathbb{R}^n$ is *uniquely* $\mathbf{v} = \hat{\mathbf{v}} + \mathbf{z}$ with $\hat{\mathbf{v}} = \mathrm{Proj}_W(\mathbf{v}) \in W$ and $\mathbf{z} \in W^\perp$. Also Proj_W is linear, with image W and kernel $W^\perp$.

ORTHOGONAL BASIS ONLY

The formula is wrong for a non-orthogonal basis (the terms interfere). Run Gram–Schmidt first.

HOW TO PROJECT ONTO W

1. Get an orthogonal basis of W (check pairwise inner products; otherwise Gram–Schmidt).
2. Compute each coefficient $\mathbf{v} \cdot \mathbf{u}_i / \mathbf{u}_i \cdot \mathbf{u}_i$ and sum.
3. $\mathbf{z} = \mathbf{v} - \hat{\mathbf{v}}$. Check $\mathbf{z} \perp \mathbf{u}_i$ for each i, and Pythagoras $\|\mathbf{v}\|^2 = \|\hat{\mathbf{v}}\|^2 + \|\mathbf{z}\|^2$.

EXAMPLE 17.1 · projection onto a plane

$W = \mathrm{Span}\{(1,1,0)^\mathsf{T}, (-1,1,0)^\mathsf{T}\}$ (orthogonal: $-1+1=0$), $\mathbf{v} = (2,4,3)^\mathsf{T}$: coefficients $\frac{6}{2} = 3$ and $\frac{2}{2} = 1$, so $\mathrm{Proj}_W \mathbf{v} = 3(1,1,0)^\mathsf{T} + (-1,1,0)^\mathsf{T} = (2,4,0)^\mathsf{T}$, $\mathbf{z} = (0,0,3)^\mathsf{T}$.
$W = \mathrm{Span}\{(1,0,1)^\mathsf{T}, (0,1,0)^\mathsf{T}\}$, $\mathbf{v} = (3,2,1)^\mathsf{T}$: coefficients $\frac{4}{2} = 2$, $\frac{2}{1} = 2$: $\mathrm{Proj}_W \mathbf{v} = (2,2,2)^\mathsf{T}$, $\mathbf{z} = (1,0,-1)^\mathsf{T}$; $12 + 2 = 14 = 9 + 4 + 1$.

17.2 Gram–Schmidt

GRAM–SCHMIDT

Given a basis $\mathbf{x}_1, \ldots, \mathbf{x}_k$ of W:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{x}_1, \\ \mathbf{u}_2 &= \mathbf{x}_2 - \tfrac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1, \\ \mathbf{u}_3 &= \mathbf{x}_3 - \tfrac{\mathbf{x}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \tfrac{\mathbf{x}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2, \ \ldots \end{aligned}$$

Each $\mathbf{u}_j$ is $\mathbf{x}_j$ minus its projection onto the span of the previous ones. Then $\{\mathbf{u}_j\}$ is an orthogonal basis of W with $\mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_j\} = \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_j\}$. Normalise to get an orthonormal basis. You may rescale a $\mathbf{u}_j$ (clear fractions) at any time. Check the first few inner products to catch slips early.

EXAMPLE 17.2 · Gram–Schmidt

$\mathbf{x}_1 = (1,1,0)^\mathsf{T}$, $\mathbf{x}_2 = (2,0,1)^\mathsf{T}$, $\mathbf{x}_3 = (0,0,3)^\mathsf{T}$. $\mathbf{u}_1 = \mathbf{x}_1$ ($\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$). $\mathbf{x}_2 \cdot \mathbf{u}_1 = 2$, so $\mathbf{u}_2 = (2,0,1)^\mathsf{T} - (1,1,0)^\mathsf{T} = (1,-1,1)^\mathsf{T}$ (check $1-1+0=0$). $\mathbf{x}_3 \cdot \mathbf{u}_1 = 0$, $\mathbf{x}_3 \cdot \mathbf{u}_2 = 3 = \mathbf{u}_2 \cdot \mathbf{u}_2$, so $\mathbf{u}_3 = (0,0,3)^\mathsf{T} - (1,-1,1)^\mathsf{T} = (-1,1,2)^\mathsf{T}$ (checks 0 and $-1-1+2=0$). Orthonormal basis: $\frac{1}{\sqrt{2}}(1,1,0)^\mathsf{T}$, $\frac{1}{\sqrt{3}}(1,-1,1)^\mathsf{T}$, $\frac{1}{\sqrt{6}}(-1,1,2)^\mathsf{T}$.
$\mathbf{x} = (1,1,1)^\mathsf{T}, (1,1,0)^\mathsf{T}, (1,0,0)^\mathsf{T}$: $\mathbf{u}_1 = (1,1,1)^\mathsf{T}$; $\mathbf{u}_2 = (1,1,0)^\mathsf{T} - \frac{2}{3}(1,1,1)^\mathsf{T} \to (1,1,-2)^\mathsf{T}$ (rescaled); $\mathbf{u}_3 = (1,0,0)^\mathsf{T} - \frac{1}{3}\mathbf{u}_1 - \frac{1}{6}(1,1,-2)^\mathsf{T} = \frac{1}{2}(1,-1,0)^\mathsf{T} \to (1,-1,0)^\mathsf{T}$. Normalised: $\frac{1}{\sqrt{3}}(1,1,1)^\mathsf{T}$, $\frac{1}{\sqrt{6}}(1,1,-2)^\mathsf{T}$, $\frac{1}{\sqrt{2}}(1,-1,0)^\mathsf{T}$.
ANIMATION: the first example ▸; the second example ▸.

CONSEQUENCES

For a subspace $W \subseteq \mathbb{R}^n$: $W \cap W^\perp = \{\mathbf{0}\}$; $\dim W + \dim W^\perp = n$; $(W^\perp)^\perp = W$. (With $(\mathrm{Row}\, A)^\perp = \mathrm{Ker}\, A$ this is rank–nullity from the geometric side.)

17.3 Orthogonal matrices

ORTHOGONAL MATRICES

For a square matrix A, these are equivalent: (1) the columns are an orthonormal basis of $\mathbb{R}^n$; (2) $A^\mathsf{T}A = I = AA^\mathsf{T}$, i.e. $A^{-1} = A^\mathsf{T}$; (3) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$; (4) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$. Such A are **orthogonal**: rigid motions (rotations, reflections), with $\mathrm{Det}\, A = \pm 1$ ($+1$ rotation, -1 reflection). Orthonormal bases of $\mathbb{R}^n$ give them as columns. (Why (2) gives (4): $\|A\mathbf{x}\|^2 = (A\mathbf{x})^\mathsf{T}(A\mathbf{x}) = \mathbf{x}^\mathsf{T}A^\mathsf{T}A\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{x} = \|\mathbf{x}\|^2$.)

IS A ORTHOGONAL?

Compute $A^{\mathsf{T}}A$ (entry (i,j) is the inner product of columns i and j). Orthogonal iff $A^{\mathsf{T}}A = I$: columns of length 1, pairwise orthogonal. Orthogonal columns alone are not enough: $\mathrm{Diag}(1,2)$ has orthogonal columns but is not orthogonal.

EXAMPLE 17.3 · orthogonal matrices

Rotation $\left[\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right]$ (Det 1); reflection $\left[\begin{smallmatrix}1 & 0 \\ 0 & -1\end{smallmatrix}\right]$ (Det -1); the matrix with the three orthonormal vectors above as columns. $\left[\begin{smallmatrix}1 & 0 \\ 0 & 2\end{smallmatrix}\right]$: no (second column has length 2).

THE CHAPTER AT A GLANCE

PROJECTION	$\mathrm{Proj}_W \mathbf{v} = \sum \frac{\mathbf{v}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}\mathbf{u}_i$ for an orthogonal basis. $\mathbf{v} = \hat{\mathbf{v}} + \mathbf{z}$, $\hat{\mathbf{v}} \in W$, $\mathbf{z} \in W^{\perp}$, uniquely.
GRAM–SCHMIDT	Subtract from each $\mathbf{x}_j$ its projection onto the earlier $\mathbf{u}$'s; normalise at the end; rescale freely.
DIMENSION	$\dim W + \dim W^{\perp} = n$, $(W^{\perp})^{\perp} = W$.
ORTHOGONAL MATRIX	Orthonormal columns $\iff A^{\mathsf{T}}A = I$; preserves lengths and angles; $\mathrm{Det} = \pm 1$.

TRY IT YOURSELF

1. $W = \mathrm{Span}\{(1,0,1)^{\mathsf{T}}, (0,1,0)^{\mathsf{T}}\}$. Verify the basis is orthogonal, find $\mathrm{Proj}_W(3,2,1)^{\mathsf{T}}$ and the component in $W^{\perp}$. Check Pythagoras.
2. Apply Gram–Schmidt to $(1,1,1)^{\mathsf{T}}, (1,1,0)^{\mathsf{T}}, (1,0,0)^{\mathsf{T}}$, then normalise.
3. Is $\frac{1}{5}\left[\begin{smallmatrix}3 & -4 \\ 4 & 3\end{smallmatrix}\right]$ orthogonal? What is its determinant? Is $\left[\begin{smallmatrix}1 & 1 \\ 1 & -1\end{smallmatrix}\right]$ orthogonal?

ANSWERS

1. $(1,0,1)\cdot(0,1,0) = 0$. Coefficients $\frac{4}{2} = 2$, $\frac{2}{1} = 2$: $\mathrm{Proj} = (2,2,2)^{\mathsf{T}}$, $\mathbf{z} = (1,0,-1)^{\mathsf{T}}$; $12 + 2 = 14$.
2. $(1,1,1)^{\mathsf{T}}, (1,1,-2)^{\mathsf{T}}, (1,-1,0)^{\mathsf{T}}$; normalised by $\sqrt{3}, \sqrt{6}, \sqrt{2}$.
3. $A^{\mathsf{T}}A = \frac{1}{25}\left[\begin{smallmatrix}25 & 0 \\ 0 & 25\end{smallmatrix}\right] = I$: yes, $\mathrm{Det} = \frac{9+16}{25} = 1$ (a rotation). Second: columns orthogonal but of length $\sqrt{2}$, so $A^{\mathsf{T}}A = 2I$: no.

CHAPTER 18

Best approximation and least squares

LECTURE 18 · *what to do when a system has no solution*

THIS CHAPTER TEACHES HOW TO

- find the point of a subspace closest to a given vector, and the distance;
- solve $A^{\mathsf{T}}A\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b}$ for an inconsistent system;
- fit a line, or any curve linear in its unknown parameters, to data.

BEST APPROXIMATION THEOREM

Let $W \subseteq \mathbb{R}^n$ be a subspace and $\mathbf{v} \in \mathbb{R}^n$. Then $\hat{\mathbf{v}} = \mathrm{Proj}_W(\mathbf{v})$ is the unique closest point of W to $\mathbf{v}$: $\|\mathbf{v} - \hat{\mathbf{v}}\| < \|\mathbf{v} - \mathbf{w}\|$ for all $\mathbf{w} \in W$, $\mathbf{w} \neq \hat{\mathbf{v}}$. The distance from $\mathbf{v}$ to W is $\|\mathbf{v} - \hat{\mathbf{v}}\| = \|\mathbf{z}\|$.

EXAMPLE 18.1 · distance to a line

$W = \mathrm{Span}\{(1,1,0)^{\mathsf{T}}\}$, $\mathbf{v} = (3,1,4)^{\mathsf{T}}$: $\hat{\mathbf{v}} = \frac{4}{2}(1,1,0)^{\mathsf{T}} = (2,2,0)^{\mathsf{T}}$, $\mathbf{v} - \hat{\mathbf{v}} = (1,-1,4)^{\mathsf{T}}$, distance $\sqrt{18} = 3\sqrt{2}$.

LEAST SQUARES AND THE NORMAL EQUATIONS

If $A\mathbf{x} = \mathbf{b}$ is inconsistent, a **least-squares solution** $\hat{\mathbf{x}}$ minimises $\|\mathbf{b} - A\mathbf{x}\|$, equivalently $A\hat{\mathbf{x}} = \mathrm{Proj}_{\mathrm{Im}\,A}(\mathbf{b})$ (consistent by construction). The residual $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to $\mathrm{Im}\,A$, hence to every column, which gives the **normal equations**

$$\boxed{A^{\mathsf{T}}A\,\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b}}.$$

If the columns of A are independent, $A^{\mathsf{T}}A$ is invertible and $\hat{\mathbf{x}} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$ is unique. If $\mathrm{Ker}\,A \neq \{\mathbf{0}\}$, $\hat{\mathbf{x}}$ is not unique, but the projection $A\hat{\mathbf{x}}$ always is.

LEAST-SQUARES SOLUTION

1. Form $A^{\mathsf{T}}A$ and $A^{\mathsf{T}}\mathbf{b}$ (a small square system, whatever the number of equations).
2. Solve $A^{\mathsf{T}}A\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b}$.
3. Check: the residual $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ satisfies $A^{\mathsf{T}}\mathbf{r} = \mathbf{0}$; quote $\|\mathbf{r}\|^2$ (sum of squared residuals) if asked.

EXAMPLE 18.2 · an inconsistent system

$A = \left[\begin{smallmatrix}1&0\\0&1\\1&1\end{smallmatrix}\right]$, $\mathbf{b} = (1,1,0)^\mathsf{T}$: $A^\mathsf{T}A = \left[\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right]$, $A^\mathsf{T}\mathbf{b} = (1,1)^\mathsf{T}$, so $\hat{\mathbf{x}} = (\frac{1}{3},\frac{1}{3})^\mathsf{T}$, $A\hat{\mathbf{x}} = (\frac{1}{3},\frac{1}{3},\frac{2}{3})^\mathsf{T}$, residual $(\frac{2}{3},\frac{2}{3},-\frac{2}{3})^\mathsf{T}$, orthogonal to both columns.

$A = \left[\begin{smallmatrix}1&1\\1&2\\1&3\end{smallmatrix}\right]$, $\mathbf{b} = (2,1,3)^\mathsf{T}$: $A^\mathsf{T}A = \left[\begin{smallmatrix}3&6\\6&14\end{smallmatrix}\right]$, $A^\mathsf{T}\mathbf{b} = (6,13)^\mathsf{T}$, so $\hat{\mathbf{x}} = (1,\frac{1}{2})^\mathsf{T}$, residual $(0.5,-1,0.5)^\mathsf{T}$ with inner products 0 with $(1,1,1)^\mathsf{T}$ and $0.5-2+1.5=0$ with $(1,2,3)^\mathsf{T}$.

Non-unique: $A = \left[\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right]$, $\mathbf{b} = (1,3)^\mathsf{T}$: $A^\mathsf{T}A = \left[\begin{smallmatrix}2&4\\4&8\end{smallmatrix}\right]$ is singular; $2x_1+4x_2=4$ has infinitely many solutions (e.g. $(2,0)^\mathsf{T}$ and $(0,1)^\mathsf{T}$), all giving $A\hat{\mathbf{x}} = (2,2)^\mathsf{T}$.

ANIMATION: the first system: projecting $\mathbf{b}$ onto the column space ▸.

LINE OF BEST FIT $y = c + dx$

One row $(1, x_i)$ per data point, right-hand side y_i. Then $A^\mathsf{T}A = \left[\begin{smallmatrix}m&\sum x_i\\\sum x_i&\sum x_i^2\end{smallmatrix}\right]$, $A^\mathsf{T}\mathbf{b} = \left[\begin{smallmatrix}\sum y_i\\\sum x_i y_i\end{smallmatrix}\right]$. Solve for c, d; the residuals are $y_i - (c + dx_i)$ and they sum to 0.

EXAMPLE 18.3 · line of best fit

Points $(1,2),(2,2),(3,3),(4,5)$: $\sum x = 10$, $\sum x^2 = 30$, $\sum y = 12$, $\sum xy = 35$; $4c+10d=12$, $10c+30d=35$ give $d=1$, $c=0.5$: $y = 0.5 + x$, residuals $0.5, -0.5, -0.5, 0.5$, sum of squares 1.

Points $(0,1),(1,1),(2,4),(3,4)$: $\sum x = 6$, $\sum x^2 = 14$, $\sum y = 10$, $\sum xy = 21$; $4c+6d=10$, $6c+14d=21$ give $d=1.2$, $c=0.7$; residuals $0.3, -0.9, 0.9, -0.3$, sum of squares 1.8.

ANIMATION: the first data set ▸; the second data set ▸.

OTHER CURVES

Any model whose *unknown parameters appear linearly* works: one row per data point, one column per parameter (the thing multiplying it, evaluated at the data point). For $y = c_0 + c_1x + c_2x^2$ rows are $(1, x_i, x_i^2)$; for $h = ct + dt^2$ rows are (t_i, t_i^2); for $y = a + b\cdot 2^x$ rows are $(1, 2^{x_i})$. The curve may be non-linear in x; it must be linear in the unknowns.

ANIMATION: a curve fitted this way, step by step ▸.

EXAMPLE 18.4 · curve fits

Ball: $(t,h) = (1,7.3),(2,11.7),(3,15.1)$ with $h = ct + dt^2$: $A = \left[\begin{smallmatrix}1&1\\2&4\\3&9\end{smallmatrix}\right]$, $A^\mathsf{T}A = \left[\begin{smallmatrix}14&36\\36&98\end{smallmatrix}\right]$, $A^\mathsf{T}\mathbf{b} = (76,190)^\mathsf{T}$; $7c+18d=38$, $18c+49d=95$ give $d=-1$, $c=8$: $h = 8t - t^2$, predicting $7, 12, 15$, residual $(0.3,-0.3,0.1)^\mathsf{T}$, orthogonal to both columns.

$y = a + b\cdot 2^x$ through $(0,2),(1,6.5),(2,8.5)$: $A = \left[\begin{smallmatrix}1&1\\1&2\\1&4\end{smallmatrix}\right]$, $A^\mathsf{T}A = \left[\begin{smallmatrix}3&7\\7&21\end{smallmatrix}\right]$, $A^\mathsf{T}\mathbf{b} = (17,49)^\mathsf{T}$; $3a+7b=17$, $a+3b=7$ give $b=2$, $a=1$: $y = 1 + 2\cdot 2^x$ (predicts $3,5,9$, residual $(-1,1.5,-0.5)^\mathsf{T}$).

COMMON SLIPS

The columns of A hold the *coefficients of the unknowns*, not the data values of y. Form $A^\mathsf{T}A$ and $A^\mathsf{T}\mathbf{b}$ with the *same* A. A model like $y = c_0c_1 + c_1x$ (unknowns multiplied together) is not linear in the unknowns and does not fit this method.

THE CHAPTER AT A GLANCE

BEST APPROX.	$\text{Proj}_W \mathbf{v}$ is the unique closest point of W; distance $= \|\mathbf{v} - \text{Proj}_W \mathbf{v}\|$.
LEAST SQUARES	Inconsistent $A\mathbf{x} = \mathbf{b}$: minimise $\|\mathbf{b} - A\mathbf{x}\|$, i.e. $A\hat{\mathbf{x}} = \text{Proj}_{\text{Im}\,A}\,\mathbf{b}$; normal equations $A^\mathsf{T}A\hat{\mathbf{x}} = A^\mathsf{T}\mathbf{b}$.
UNIQUENESS	Independent columns $\Rightarrow$ $\hat{\mathbf{x}}$ unique; otherwise $A\hat{\mathbf{x}}$ is still unique.
FITTING	One row per data point, one column per unknown; line: rows $(1, x_i)$. Unknowns must appear linearly.

TRY IT YOURSELF

1. Find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ for $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, $\mathbf{b} = (2, 1, 3)^\mathsf{T}$, and verify the residual is orthogonal to both columns.
2. Find the line of best fit through $(0, 1), (1, 1), (2, 4), (3, 4)$ and the sum of squared residuals.
3. $W = \text{Span}\{(1, 1, 0)^\mathsf{T}\}$, $\mathbf{v} = (3, 1, 4)^\mathsf{T}$. Find the point of W closest to $\mathbf{v}$ and the distance.
4. Find the best fit of the form $y = a + b \cdot 2^x$ to $(0, 2), (1, 6.5), (2, 8.5)$.

ANSWERS

1. $3c + 6d = 6$, $6c + 14d = 13$: $d = \frac{1}{2}$, $c = 1$; $A\hat{\mathbf{x}} = (1.5, 2, 2.5)^\mathsf{T}$, residual $(0.5, -1, 0.5)^\mathsf{T}$, orthogonal to $(1, 1, 1)^\mathsf{T}$ and $(1, 2, 3)^\mathsf{T}$.
2. $y = 0.7 + 1.2x$; residuals $0.3, -0.9, 0.9, -0.3$; sum of squares 1.8.
3. $(2, 2, 0)^\mathsf{T}$; distance $\sqrt{18} = 3\sqrt{2}$.
4. $y = 1 + 2 \cdot 2^x$ ($a = 1$, $b = 2$).

PART IX

Eigenvalues

CHAPTER 19

Eigenvalues and eigenvectors

LECTURE 19 · *the directions a matrix does not turn*

THIS CHAPTER TEACHES HOW TO

- find eigenvalues from the characteristic polynomial and check them with trace and determinant;
- find a basis of each eigenspace;
- decide whether a matrix is diagonalisable and, if so, write $A = EDE^{-1}$.

19.1 Definitions

EIGENVECTOR, EIGENVALUE, EIGENSPACE

For $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, a non-zero vector $\mathbf{v}$ with $A\mathbf{v} = \lambda\mathbf{v}$ is an **eigenvector**, and λ is its **eigenvalue**. The number 0 *is* allowed as an eigenvalue (it means $A\mathbf{v} = \mathbf{0}$); the vector $\mathbf{0}$ is not allowed as an eigenvector. The **eigenspace** of λ is $\mathrm{Ker}(A - \lambda I)$: all eigenvectors for λ together with $\mathbf{0}$. It is a subspace, and any non-zero multiple of an eigenvector is again an eigenvector, so there is no single correct answer.

THE CHARACTERISTIC POLYNOMIAL

λ is an eigenvalue $\iff$ $(A - \lambda I)\mathbf{v} = \mathbf{0}$ has a non-zero solution $\iff$ $\mathrm{Det}(A - \lambda I) = 0$. The **characteristic polynomial** $\chi_A(\lambda) = \mathrm{Det}(A - \lambda I)$ has degree n; its roots are the eigenvalues.

SUBTRACT ON THE DIAGONAL ONLY

$A - \lambda I$ has λ subtracted from the diagonal entries only. If $A - \lambda I$ row reduces with no free variable, then λ is not an eigenvalue and there is a mistake further up.

EIGENVALUES AND EIGENSPACES

1. Compute $\mathrm{Det}(A - \lambda I)$ (expand along the row or column with most zeros) and factorise it. Pull out common factors rather than multiplying out; otherwise try small integers dividing the constant term. The roots are the eigenvalues.
2. For each eigenvalue λ, row reduce $A - \lambda I$ and read off a basis of its kernel (Chapter 7).
3. Check: $A\mathbf{v} = \lambda\mathbf{v}$ for each basis vector; $\sum \lambda_i = \mathrm{Tr}\, A$, $\prod \lambda_i = \mathrm{Det}\, A$ (eigenvalues listed with repetition).

A triangular matrix has its eigenvalues on the diagonal. In this module every eigenvalue we compute is real (a $90°$ rotation has $\lambda^2 + 1$ and no real eigenvalue).

EXAMPLE 19.1 · 2×2

$A = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$: $\mathrm{Det}(A - \lambda I) = (2-\lambda)^2 - 1 = (\lambda - 1)(\lambda - 3)$. $\lambda = 3$: $A - 3I = \left[\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, eigenspace $\mathrm{Span}\{(1,1)^\mathsf{T}\}$. $\lambda = 1$: $A - I = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, eigenspace $\mathrm{Span}\{(1,-1)^\mathsf{T}\}$. Trace $4 = 3 + 1$, $\mathrm{Det}\, 3 = 3 \cdot 1$.

ANIMATION: the computation ▸; the picture ▸.

EXAMPLE 19.2 · 3×3, repeated eigenvalue

$A = \left[\begin{smallmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{smallmatrix}\right]$. Expanding along the first row, every term has a factor $\lambda - 2$: $\mathrm{Det}(A - \lambda I) = (\lambda - 2)\big[(3-\lambda)(\lambda - 5) + 3\big] = -(\lambda - 2)^2(\lambda - 6)$.

$\lambda = 6$: $A - 6I$ reduces to $\left[\begin{smallmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{smallmatrix}\right]$, eigenspace $\mathrm{Span}\{(1,2,1)^\mathsf{T}\}$.

$\lambda = 2$: $A - 2I$ reduces to $\left[\begin{smallmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$, two free variables, basis $(-1,1,0)^\mathsf{T}, (-1,0,1)^\mathsf{T}$.

Check: $A(1,2,1)^\mathsf{T} = (6,12,6)^\mathsf{T}$; $6 + 2 + 2 = 10 = 3 + 4 + 3$.

ANIMATION: the computation ▸.

EXAMPLE 19.3 · Arguing from $A\mathbf{v} = \lambda\mathbf{v}$

Sometimes you are asked what an eigenvalue *can* be, from a property of A. Start from $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, apply the property, and cancel $\mathbf{v}$. If $P^2 = P$: $\lambda\mathbf{v} = P\mathbf{v} = P^2\mathbf{v} = P(\lambda\mathbf{v}) = \lambda^2\mathbf{v}$, so $(\lambda^2 - \lambda)\mathbf{v} = \mathbf{0}$ and, as $\mathbf{v} \neq \mathbf{0}$, $\lambda \in \{0, 1\}$. Say explicitly where $\mathbf{v} \neq \mathbf{0}$ is used. If $\|A\mathbf{x}\| \leq c\|\mathbf{x}\|$ for all $\mathbf{x}$, take lengths: $|\lambda|\|\mathbf{v}\| = \|\lambda\mathbf{v}\| = \|A\mathbf{v}\| \leq c\|\mathbf{v}\|$; divide by $\|\mathbf{v}\| > 0$ to get $|\lambda| \leq c$.

19.2 Diagonalisation

DIAGONALISABLE

A is **diagonalisable** if $A = EDE^{-1}$ with E invertible and D diagonal. This happens exactly when A has n independent eigenvectors, i.e. when the dimensions of the eigenspaces add up to n. Then the columns of E are the eigenvectors and D has the eigenvalues on its diagonal, in the same order. Eigenvectors from different eigenspaces are automatically independent; n distinct eigenvalues always suffice. (In the basis of eigenvectors, A just scales each coordinate: this is a change of basis, Chapter 13.)

DIAGONALISE A

1. Find all eigenvalues and a basis of each eigenspace.
2. Add up the eigenspace dimensions. If the sum is n, put the basis vectors as columns of E and the matching eigenvalues in D. If less than n, say A is not diagonalisable.
3. Check $AE = ED$ (this avoids computing E^{-1}).

REPEATED IS NOT THE SAME AS DEFECTIVE

A repeated eigenvalue is not by itself a failure (the 3×3 example above has eigenspaces of dimensions $1 + 2 = 3$). What counts is the dimension of the eigenspace, not the multiplicity of the root.

EXAMPLE 19.4 · diagonalisable or not

$\left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right] = EDE^{-1}$ with $E = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, $D = \mathrm{Diag}(3, 1)$, $E^{-1} = \frac{1}{2}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$.
The 3×3 example: $E = \left[\begin{smallmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right]$, $D = \mathrm{Diag}(6, 2, 2)$.
$\left[\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right]$: $\lambda = 2$ only; $A - 2I = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ has a one-dimensional kernel $\mathrm{Span}\{(1, 0)^{\mathsf{T}}\}$, so $1 < 2$ and it is not diagonalisable.

ANIMATION: the first matrix ▸; the 3×3 matrix ▸; the matrix that fails ▸.

THE CHAPTER AT A GLANCE

EIGEN PAIR	$A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$. Eigenspace $= \mathrm{Ker}(A - \lambda I)$.
EIGENVALUES	Roots of $\mathrm{Det}(A - \lambda I)$; check $\sum \lambda_i = \mathrm{Tr}\, A$, $\prod \lambda_i = \mathrm{Det}\, A$; triangular: diagonal entries.
DIAGONALISE	$A = EDE^{-1}$ iff eigenspace dimensions sum to n; columns of E match the order of D; check $AE = ED$.
FAILURE	Dimensions sum to less than n, e.g. $\left[\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right]$.

TRY IT YOURSELF

1. Find the eigenvalues and eigenvectors of $A = \left[\begin{smallmatrix} 4 & 1 \\ 2 & 3 \end{smallmatrix}\right]$ and write $A = EDE^{-1}$.
2. $B = \left[\begin{smallmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{smallmatrix}\right]$: find the characteristic polynomial, a basis of each eigenspace, and diagonalise B if possible.
3. Is $\left[\begin{smallmatrix} 3 & 1 \\ 0 & 3 \end{smallmatrix}\right]$ diagonalisable? Explain.
4. Show that if $A^2 = I$ then every eigenvalue of A is 1 or -1.

ANSWERS

1. $(4 - \lambda)(3 - \lambda) - 2 = (\lambda - 5)(\lambda - 2)$. $\lambda = 5$: $(1, 1)^{\mathsf{T}}$; $\lambda = 2$: $(1, -2)^{\mathsf{T}}$. $E = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -2 \end{smallmatrix}\right]$, $D = \mathrm{Diag}(5, 2)$, $E^{-1} = \frac{1}{3}\left[\begin{smallmatrix} 2 & 1 \\ 1 & -1 \end{smallmatrix}\right]$.
2. Expanding along row 2: $-(\lambda - 1)^2(\lambda - 3)$. $\lambda = 3$: $(1, 0, 1)^{\mathsf{T}}$. $\lambda = 1$: $(-1, 1, 0)^{\mathsf{T}}, (-1, 0, 1)^{\mathsf{T}}$. $E = \left[\begin{smallmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right]$, $D = \mathrm{Diag}(3, 1, 1)$; trace $5 = 3 + 1 + 1$.
3. $(3 - \lambda)^2$ so $\lambda = 3$ only; $A - 3I = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ has a one-dimensional kernel, so only one independent eigenvector: not diagonalisable.

4. $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: $\mathbf{v} = A^2\mathbf{v} = \lambda^2\mathbf{v}$, so $(\lambda^2 - 1)\mathbf{v} = \mathbf{0}$ and $\lambda^2 = 1$.

CHAPTER 20

Powers and the long run

LECTURE 19, CONTINUED · *what diagonalisation is for*

THIS CHAPTER TEACHES HOW TO

- compute $A^k = ED^kE^{-1}$;
- solve $\mathbf{x}_{k+1} = A\mathbf{x}_k$ and describe its limit;
- read the long-term behaviour from the eigenvalues.

POWERS

If $A = EDE^{-1}$ then $A^k = ED^kE^{-1}$ with $D^k = \mathrm{Diag}(\lambda_1^k, \dots, \lambda_n^k)$, because the middle $E^{-1}E$ cancel. For a system that moves in steps, $\mathbf{x}_{k+1} = A\mathbf{x}_k$ has solution $\mathbf{x}_k = A^k\mathbf{x}_0 = ED^kE^{-1}\mathbf{x}_0$.

SOLVE $\mathbf{x}_{k+1} = A\mathbf{x}_k$

1. Diagonalise $A = EDE^{-1}$.
2. Work from the right so that only vectors appear: $\mathbf{c} = E^{-1}\mathbf{x}_0$, then $D^k\mathbf{c}$, then $E(D^k\mathbf{c})$. Equivalently write $\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$; then $\mathbf{x}_k = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n$.
3. Check $k = 0$ and $k = 1$.

THE LONG RUN

$|\lambda| < 1$: that term dies out. $\lambda = 1$: it stays. $|\lambda| > 1$: it grows, and the largest $|\lambda|$ wins. $\lambda < 0$: the sign alternates from step to step. The eigenvalues give the long-term behaviour without iterating.

EXAMPLE 20.1 · powers

$A = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$: $A^k = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right] \left[\begin{smallmatrix} 3^k & 0 \\ 0 & 1 \end{smallmatrix}\right] \frac{1}{2} \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right] = \frac{1}{2} \begin{bmatrix} 3^k+1 & 3^k-1 \\ 3^k-1 & 3^k+1 \end{bmatrix}$; $k = 1$ gives A.

ANIMATION: $A^k = ED^kE^{-1}$ for this matrix ▸.

EXAMPLE 20.2 · machine working or broken

$M = \left[\begin{smallmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{smallmatrix}\right]$, $\mathbf{x}_0 = (1, 0)^{\mathsf{T}}$ (columns sum to 1). $\mathrm{Det}(M - \lambda I) = \lambda^2 - 1.4\lambda + 0.4 = (\lambda - 1)(\lambda - 0.4)$. $\lambda = 1$: $(5, 1)^{\mathsf{T}}$; $\lambda = 0.4$: $(1, -1)^{\mathsf{T}}$. $E = \left[\begin{smallmatrix} 5 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, $E^{-1} = \frac{1}{6} \left[\begin{smallmatrix} 1 & 1 \\ 1 & -5 \end{smallmatrix}\right]$, $E^{-1}\mathbf{x}_0 = \frac{1}{6}(1, 1)^{\mathsf{T}}$, so

$$\mathbf{x}_k = \frac{1}{6} \begin{bmatrix} 5 + 0.4^k \\ 1 - 0.4^k \end{bmatrix} \to \frac{1}{6} \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

Check $k = 1$: $\frac{1}{6}(5.4, 0.6)^{\mathsf{T}} = (0.9, 0.1)^{\mathsf{T}} = M\mathbf{x}_0$. The limit is a multiple of the $\lambda = 1$ eigenvector:

the machine works 5 days in 6.

EXAMPLE 20.3 · averaging

$A = \left[\begin{smallmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{smallmatrix}\right]$, $\mathbf{x}_0 = (4,0)^{\mathsf{T}}$: eigenvalues $1, 0.5$ with $(1,1)^{\mathsf{T}}, (1,-1)^{\mathsf{T}}$; $E^{-1}\mathbf{x}_0 = (2,2)^{\mathsf{T}}$; $\mathbf{x}_k = \big(2 + 2(0.5)^k,\ 2 - 2(0.5)^k\big)^{\mathsf{T}} \to (2,2)^{\mathsf{T}}$. Check $k = 1$: $(3,1)^{\mathsf{T}}$.

EXAMPLE 20.4 · growth

$\mathbf{x}_{k+1} = \left[\begin{smallmatrix} 0 & 2 \\ 1 & 1 \end{smallmatrix}\right]\mathbf{x}_k$: $(\lambda - 2)(\lambda + 1)$, eigenvectors $(1,1)^{\mathsf{T}}$ (for 2) and $(2,-1)^{\mathsf{T}}$ (for -1). From $\mathbf{x}_0 = (3,0)^{\mathsf{T}} = (1,1)^{\mathsf{T}} + (2,-1)^{\mathsf{T}}$: $\mathbf{x}_k = 2^k(1,1)^{\mathsf{T}} + (-1)^k(2,-1)^{\mathsf{T}}$. The 2^k term dominates, so the ratio settles to $1:1$.

THE CHAPTER AT A GLANCE

POWERS	$A^k = ED^kE^{-1}$; $D^k = \mathrm{Diag}(\lambda_i^k)$.
DYNAMICS	$\mathbf{x}_k = A^k\mathbf{x}_0 = \sum c_i\lambda_i^k\mathbf{v}_i$ where $\mathbf{x}_0 = \sum c_i\mathbf{v}_i$.
LONG RUN	$\lvert\lambda\rvert < 1$ dies; $\lambda = 1$ stays; $\lvert\lambda\rvert > 1$ grows (largest wins); negative λ alternates.

TRY IT YOURSELF

1. Use $A = \left[\begin{smallmatrix} 4 & 1 \\ 2 & 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -2 \end{smallmatrix}\right] \mathrm{Diag}(5,2) \frac{1}{3}\left[\begin{smallmatrix} 2 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ to compute A^k and check $k = 1$.
2. $\mathbf{x}_{k+1} = A\mathbf{x}_k$ with $A = \left[\begin{smallmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{smallmatrix}\right]$, $\mathbf{x}_0 = (4,0)^{\mathsf{T}}$: find $\mathbf{x}_k$ and its limit.

ANSWERS

1. $A^k = \frac{1}{3}\left[\begin{smallmatrix} 2\cdot 5^k + 2^k & 5^k - 2^k \\ 2\cdot 5^k - 2^{k+1} & 5^k + 2\cdot 2^k \end{smallmatrix}\right]$; $k = 1$: $\frac{1}{3}\left[\begin{smallmatrix} 12 & 3 \\ 6 & 9 \end{smallmatrix}\right] = A$.
2. $\mathbf{x}_k = (2 + 2(0.5)^k,\ 2 - 2(0.5)^k)^{\mathsf{T}} \to (2,2)^{\mathsf{T}}$: the total 4 is shared equally.

PART X

Singular values

CHAPTER 21

The singular value decomposition

LECTURE 20 · *every matrix is a rotation, a stretch and another rotation*

THIS CHAPTER TEACHES HOW TO

- compute the singular values of any matrix;
- compute a full or reduced SVD $A = U\Sigma V^{\mathsf{T}}$;
- read the rank and bases of the fundamental subspaces off an SVD.

21.1 Symmetric matrices and $A^{\mathsf{T}}A$

SYMMETRIC MATRICES

If $A^{\mathsf{T}} = A$ then (1) every eigenvalue is real; (2) eigenvectors for distinct eigenvalues are orthogonal; (3) $\mathbb{R}^n$ has an orthonormal basis of eigenvectors (the spectral theorem). For repeated eigenvalues, apply Gram–Schmidt *within* the eigenspace (Chapter 17).

$A^{\mathsf{T}}A$

For $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $A^{\mathsf{T}}A$ is $n \times n$ and symmetric, and every eigenvalue satisfies $\lambda = \|A\mathbf{v}\|^2/\|\mathbf{v}\|^2 \geq 0$ (for an eigenvector $\mathbf{v}$).

21.2 Singular values

SINGULAR VALUES

Let $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ be the eigenvalues of $A^{\mathsf{T}}A$ (with multiplicity, decreasing). The **singular values** are $s_i = \sqrt{\lambda_i}$, so $s_1 \geq \cdots \geq s_n \geq 0$. There are always n of them (one per *column*). $s_i = \|A\mathbf{v}_i\|$ for a unit eigenvector $\mathbf{v}_i$. Geometrically, A maps the unit circle (sphere) to an ellipse (ellipsoid) with semi-axes s_i; s_1 is the greatest stretch of any unit vector and s_n the least. A zero singular value signals a collapsed direction; the number of non-zero ones is $\mathrm{Rank}\, A$.

EXAMPLE 21.1 · singular values

$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$: $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$, $\lambda(\lambda - 4)$, so $s_1 = 2$, $s_2 = 0$, rank 1.
$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$: $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, eigenvalues $3, 1$, so $s_1 = \sqrt{3}$, $s_2 = 1$ (only two, although A is 3×2).

21.3 The SVD

SINGULAR VALUE DECOMPOSITION

Every $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ factors as $A = U\Sigma V^{\mathsf{T}}$ with $U \in \mathrm{Mat}_{m\times m}(\mathbb{R})$ and $V \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ orthogonal (Chapter 17) and $\Sigma \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ having $s_1, s_2, \ldots$ down its leading diagonal and zeros elsewhere (same shape as A, so usually not square: e.g. $\begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \end{bmatrix}$ for $m > n$). The columns satisfy $A^{\mathsf{T}}A\mathbf{v}_i = s_i^2\mathbf{v}_i$ and $A\mathbf{v}_i = s_i\mathbf{u}_i$. Read from the right: V^{T} rotates, Σ stretches each axis by s_i, U rotates (an orthogonal matrix may include a reflection).

HOW TO COMPUTE AN SVD

1. Form $A^{\mathsf{T}}A$ ($n \times n$, symmetric).
2. Find its eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$; set $s_i = \sqrt{\lambda_i}$.
3. For each λ_i find an *orthonormal* basis of $\mathrm{Ker}(A^{\mathsf{T}}A - \lambda_i I)$; these are the columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V, in the same order as the s_i.
4. Build Σ, the same shape as A, with the s_i on the diagonal.
5. For each i with $s_i > 0$ set $\mathbf{u}_i = \frac{1}{s_i}A\mathbf{v}_i$.
6. If fewer than m columns $\mathbf{u}_i$ exist, extend to an orthonormal basis of $\mathbb{R}^m$ (solve for a vector orthogonal to those found, normalise; Gram–Schmidt if several needed).
7. Check $V^{\mathsf{T}}V = I$, $U^{\mathsf{T}}U = I$, and multiply out $U\Sigma V^{\mathsf{T}}$ against A.

ORDER AND LENGTH

The $\mathbf{v}_i$ must be *unit* vectors, listed in the same order as the s_i. A repeated λ needs Gram–Schmidt inside its eigenspace. In the answer, give the size of U that was asked for (full or reduced).

EXAMPLE 21.2 · a 3×2 matrix, full and reduced

$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $s_1 = \sqrt{3}$, $s_2 = 1$. $\lambda = 3$: $(1,1)^{\mathsf{T}}$; $\lambda = 1$: $(1,-1)^{\mathsf{T}}$; so $V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.
$\mathbf{u}_1 = \frac{1}{\sqrt{3}}A\mathbf{v}_1 = \frac{1}{\sqrt{6}}(1,1,2)^{\mathsf{T}}$, $\mathbf{u}_2 = A\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1,0)^{\mathsf{T}}$ (orthonormal: $1 - 1 + 0 = 0$).
Reduced SVD ($m > n$, Rank $= n$): U is 3×2 with columns $\mathbf{u}_1, \mathbf{u}_2$, $\Sigma = \mathrm{Diag}(\sqrt{3}, 1)$.
Full SVD: need $\mathbf{u}_3 \perp \mathbf{u}_1, \mathbf{u}_2$: $w_1 + w_2 + 2w_3 = 0$, $w_1 - w_2 = 0$ give $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1,1,-1)^{\mathsf{T}}$, and
$U = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{bmatrix}$.
ANIMATION: the SVD of this matrix ▸.

EXAMPLE 21.3 · a zero singular value

$A = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$: $s_1 = 2$, $s_2 = 0$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1)^\mathsf{T}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1)^\mathsf{T}$. Only $s_1 > 0$, so $\mathbf{u}_1 = \frac{1}{2}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1)^\mathsf{T}$; complete with $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,-1)^\mathsf{T}$. $A = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right]\left[\begin{smallmatrix}2&0\\0&0\end{smallmatrix}\right]\frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right]$. Check: $\Sigma V^\mathsf{T} = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}2&2\\0&0\end{smallmatrix}\right]$, and U times it is $\left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$.

ANIMATION: the SVD of this matrix ▸.

EXAMPLE 21.4 · more

$A = \left[\begin{smallmatrix}2&0\\0&3\end{smallmatrix}\right]$: $A^\mathsf{T}A = \operatorname{Diag}(4,9)$, $s_1 = 3$, $s_2 = 2$ with $\mathbf{v}_1 = (0,1)^\mathsf{T}$, $\mathbf{v}_2 = (1,0)^\mathsf{T}$, $\mathbf{u}_1 = (0,1)^\mathsf{T}$, $\mathbf{u}_2 = (1,0)^\mathsf{T}$: $A = \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]\operatorname{Diag}(3,2)\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$. The order of the s_i is decreasing, so the diagonal entries are swapped.

$A = \left[\begin{smallmatrix}1&1\\0&0\end{smallmatrix}\right]$: $A^\mathsf{T}A = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$, $s_1 = \sqrt{2}$, $s_2 = 0$, $\mathbf{u}_1 = (1,0)^\mathsf{T}$, $\mathbf{u}_2 = (0,1)^\mathsf{T}$, $U = I$.

$A = \begin{bmatrix}1&2\\2&1\\2&-1\end{bmatrix}$: $A^\mathsf{T}A = \left[\begin{smallmatrix}9&2\\2&6\end{smallmatrix}\right]$, $(\lambda-10)(\lambda-5)$, $s_1 = \sqrt{10}$, $s_2 = \sqrt{5}$; $\mathbf{v}_1 = \frac{1}{\sqrt{5}}(2,1)^\mathsf{T}$, $\mathbf{v}_2 = \frac{1}{\sqrt{5}}(1,-2)^\mathsf{T}$; $\mathbf{u}_1 = \frac{1}{\sqrt{50}}(4,5,3)^\mathsf{T}$, $\mathbf{u}_2 = \frac{1}{5}(-3,0,4)^\mathsf{T}$ (reduced U, 3×2).

ANIMATION: the 3×2 matrix ▸.

WHAT THE PIECES MEAN

Let r be the number of non-zero singular values. Then $\operatorname{Rank} A = r$; $\operatorname{Im} A = \operatorname{Span}\{\mathbf{u}_1,\dots,\mathbf{u}_r\}$; $\operatorname{Ker} A = \operatorname{Span}\{\mathbf{v}_{r+1},\dots,\mathbf{v}_n\}$; $\operatorname{Coker} A = \operatorname{Span}\{\mathbf{u}_{r+1},\dots,\mathbf{u}_m\}$ (and $\operatorname{Row} A = \operatorname{Span}\{\mathbf{v}_1,\dots,\mathbf{v}_r\}$). All are orthonormal bases of the four fundamental subspaces at once. E.g. for $\left[\begin{smallmatrix}1&1\\0&0\end{smallmatrix}\right]$: rank 1, $\operatorname{Im} A = \operatorname{Span}\{(1,0)^\mathsf{T}\}$, $\operatorname{Ker} A = \operatorname{Span}\{(1,-1)^\mathsf{T}\}$.

THE CHAPTER AT A GLANCE

SYMMETRIC	Real eigenvalues, orthonormal eigenbasis. $A^\mathsf{T}A$ is symmetric with eigenvalues ≥ 0.
SINGULAR VALUES	$s_i = \sqrt{\lambda_i(A^\mathsf{T}A)}$, decreasing; n of them; $r = \#$ non-zero $= \operatorname{Rank} A$.
SVD	$A = U\Sigma V^\mathsf{T}$, U, V orthogonal, Σ same shape as A. Recipe: $A^\mathsf{T}A \to s_i \to \mathbf{v}_i \to \mathbf{u}_i = A\mathbf{v}_i/s_i \to$ extend U.
REDUCED	Drop the zero rows of Σ and matching columns of U ($m > n$, full rank).
READS OFF	$\operatorname{Im} A$: $\mathbf{u}_{1..r}$; $\operatorname{Ker} A$: $\mathbf{v}_{r+1..n}$; $\operatorname{Coker} A$: $\mathbf{u}_{r+1..m}$.

TRY IT YOURSELF

1. Find an SVD of $A = \left[\begin{smallmatrix}2&0\\0&3\end{smallmatrix}\right]$ and of $B = \left[\begin{smallmatrix}0&2\\3&0\end{smallmatrix}\right]$.
2. Find an SVD of $A = \left[\begin{smallmatrix}1&1\\0&0\end{smallmatrix}\right]$ and read off $\operatorname{Rank} A$, a basis of $\operatorname{Im} A$ and of $\operatorname{Ker} A$.
3. Find a reduced SVD of $A = \begin{bmatrix}1&2\\2&1\\2&-1\end{bmatrix}$.

ANSWERS

1. A: $s_=3, 2$; $A = \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]\operatorname{Diag}(3,2)\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$. B: $B^\mathsf{T}B = \operatorname{Diag}(9,4)$, $\mathbf{v}_1 = (1,0)^\mathsf{T}$, $\mathbf{v}_2 = (0,1)^\mathsf{T}$, $\mathbf{u}_1 = (0,1)^\mathsf{T}$, $\mathbf{u}_2 = (1,0)^\mathsf{T}$: $B = \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]\operatorname{Diag}(3,2)I$.
2. $s_=\sqrt{2}, 0$; $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1)^\mathsf{T}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1)^\mathsf{T}$, $\mathbf{u}_1 = (1,0)^\mathsf{T}$, $\mathbf{u}_2 = (0,1)^\mathsf{T}$, $U = I$. Rank 1, $\operatorname{Im} A = \operatorname{Span}\{(1,0)^\mathsf{T}\}$, $\operatorname{Ker} A = \operatorname{Span}\{(1,-1)^\mathsf{T}\}$.

3. $s_{=}\sqrt{10},\sqrt{5}$; $A = \begin{bmatrix} 4/\sqrt{50} & -3/5 \\ 5/\sqrt{50} & 0 \\ 3/\sqrt{50} & 4/5 \end{bmatrix} \operatorname{Diag}(\sqrt{10},\sqrt{5}) \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$.

CHAPTER 22

Pseudoinverse and least squares

LECTURE 20, CONTINUED · *least squares without inverting $A^{\mathsf{T}}A$*

THIS CHAPTER TEACHES HOW TO

- build Σ^{+} and $A^{+} = V\Sigma^{+}U^{\mathsf{T}}$;
- solve least-squares problems with $\hat{\mathbf{x}} = A^{+}\mathbf{b}$.

PSEUDOINVERSE

Let $A = U\Sigma V^{\mathsf{T}}$ with r non-zero singular values. $\Sigma^{+} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ is obtained by transposing Σ and replacing each non-zero s_i by $1/s_i$ (zeros stay zero: never divide by zero). The **pseudoinverse** is

$$A^{+} = V\Sigma^{+}U^{\mathsf{T}}.$$

$\hat{\mathbf{x}} = A^{+}\mathbf{b}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ (Chapter 18): *the* solution if the columns of A are independent, and otherwise the one of smallest length. Unlike $(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$, it is defined even for dependent columns.

LEAST SQUARES BY SVD

1. Find the SVD (full U) of A.
2. Compute $\mathbf{c} = U^{\mathsf{T}}\mathbf{b}$, then $\Sigma^{+}\mathbf{c}$ (divide the first r entries by s_i, drop the rest).
3. $\hat{\mathbf{x}} = V(\Sigma^{+}\mathbf{c})$.

The entries c_i with $i > r$ are the part of $\mathbf{b}$ that cannot be reached; they form the residual.

EXAMPLE 22.1 · pseudoinverse

$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \Sigma^{+} = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.
$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{b} = (1,1,0)^{\mathsf{T}}$, with the SVD of Chapter 21: $U^{\mathsf{T}}\mathbf{b} = (2/\sqrt{6},\ 0,\ 2/\sqrt{3})^{\mathsf{T}}$, $\Sigma^{+}U^{\mathsf{T}}\mathbf{b} = (2/\sqrt{18},\ 0)^{\mathsf{T}}$, $\hat{\mathbf{x}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}(2/\sqrt{18}, 0)^{\mathsf{T}} = (\frac{1}{3}, \frac{1}{3})^{\mathsf{T}}$, the same as from the normal equations in Chapter 18.

EXAMPLE 22.2 · dependent columns

$A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] = U\,\mathrm{Diag}(2,0)V^{\mathsf{T}}$ (Chapter 21), $\Sigma^{+} = \mathrm{Diag}(\frac{1}{2},0)$, so $A^{+} = \frac{1}{4}\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, and $\hat{\mathbf{x}} = A^{+}(1,3)^{\mathsf{T}} = (1,1)^{\mathsf{T}}$: of all solutions of $x_1 + x_2 = 2$, the one of least length.

THE CHAPTER AT A GLANCE

A^{+}	$V\Sigma^{+}U^{\mathsf{T}}$; invert the non-zero s_i, transpose, leave zeros.
LEAST SQUARES	$\hat{\mathbf{x}} = A^{+}\mathbf{b}$; shortest solution when columns dependent; no $A^{\mathsf{T}}A$ inverse needed.

TRY IT YOURSELF

1. Write Σ^{+} for $\Sigma = \left[\begin{smallmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{smallmatrix}\right]$ and for $\Sigma = \left[\begin{smallmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$.
2. Use $A = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right]$ to find A^{+} and the solution of $A\mathbf{x} = (4,3)^{\mathsf{T}}$.

ANSWERS

1. $\left[\begin{smallmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \end{smallmatrix}\right]$; $\left[\begin{smallmatrix} 1/2 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.
2. $A^{+} = A^{-1} = \mathrm{Diag}(\frac{1}{2},\frac{1}{3})$; $\hat{\mathbf{x}} = (2,1)^{\mathsf{T}}$.