Contents

1 Introduction .................................................. 1

2 Geometry in two dimensions ................................ 3
   2.1 Vector operations ........................................... 3
   2.2 Lines ........................................................ 4
   2.3 Triangles .................................................... 6
   2.4 Distances .................................................... 7
   2.5 Dot product .................................................. 7
   2.6 Equation of a line .......................................... 8
   2.7 Perpendicular distance .................................... 9
   2.8 Two famous inequalities ................................. 10

3 Linear inequalities ........................................... 13

4 Conics .......................................................... 15
   4.1 Transforming conics ....................................... 18
      4.1.1 Translation ............................................. 18
      4.1.2 Rotation ................................................ 19
   4.2 The Theory .................................................. 20
   4.3 Geometrical definition of conics ........................ 22

5 Matrices and linear equations .............................. 25
   5.1 Method – Gaussian elimination .......................... 27

6 Some applications of matrix algebra ....................... 37
   6.1 Linear equations .......................................... 37
   6.2 Population Distribution ................................... 38
   6.3 Matrices and Geometry .................................... 41
      6.3.1 Rotation ............................................... 41
      6.3.2 Reflection ............................................. 41
      6.3.3 Combining Transformations .......................... 41

7 Inverses ........................................................ 43
   7.1 Relation to linear equations ............................. 43
   7.2 Finding inverses ............................................ 44
   7.3 Finding inverses in general .............................. 45
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Determinants</td>
<td>47</td>
</tr>
<tr>
<td>8.1</td>
<td>Properties</td>
<td>48</td>
</tr>
<tr>
<td>8.2</td>
<td>Effects of row operations on determinant</td>
<td>49</td>
</tr>
<tr>
<td>8.3</td>
<td>$n \times n$ determinants</td>
<td>51</td>
</tr>
<tr>
<td>9</td>
<td>Eigenvalues and eigenvectors</td>
<td>53</td>
</tr>
<tr>
<td>9.1</td>
<td>Eigenvectors</td>
<td>53</td>
</tr>
<tr>
<td>9.2</td>
<td>How to find eigenvectors and eigenvalues</td>
<td>54</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Back to Fibonacci</td>
<td>55</td>
</tr>
<tr>
<td>9.3</td>
<td>Diagonalization</td>
<td>56</td>
</tr>
<tr>
<td>9.3.1</td>
<td>Repeated eigenvalues</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>Conics (again) and Quadrics</td>
<td>63</td>
</tr>
<tr>
<td>10.1</td>
<td>Reduction of conics</td>
<td>65</td>
</tr>
<tr>
<td>10.1.1</td>
<td>Orthogonal matrices</td>
<td>67</td>
</tr>
<tr>
<td>10.2</td>
<td>Quadric surfaces</td>
<td>67</td>
</tr>
<tr>
<td>10.3</td>
<td>Linear Geometry in 3 dimensions</td>
<td>70</td>
</tr>
<tr>
<td>10.4</td>
<td>Geometry of planes</td>
<td>70</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Some Greek highlights
Chapter 2

Geometry in two dimensions

This means geometry in the plane. Based on the real numbers. Think of real numbers as points on the real line.

Definition 2.1. $\mathbb{R}^2$ is the set of all ordered pairs $(x_1, x_2)$ with $x_1, x_2$ real numbers.
Note: $(1, 3) \neq (3, 1)$

Call $(x_1, x_2)$ a vector. Geometrically, vector $(x_1, x_2)$ will be used to represent several things:

- Point $p$ in the plane with coordinates $(x_1, x_2)$
- Position vector $\overrightarrow{OP}$, line in direction $\overrightarrow{OP}$, length $OP$
- Any vector $\overrightarrow{AB}$ of the same direction as $\overrightarrow{OP}$

Usually we write $x = (x_1, x_2)$.
The origin or zero vector is $\mathbf{O} = (0, 0)$.

2.1 Vector operations

Definition 2.2.

1) Addition
define the sum of two vectors by

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

2) Scalar multiplication
A scalar is a real number. If $x = (x_1, x_2)$ is a vector and $\lambda \in \mathbb{R}$ is a scalar

$$\lambda x = \lambda (x_1, x_2) = (\lambda x_1, \lambda x_2)$$

- Subtraction

$$x - y = x + (-y) = (x_1 - y_1, x_2 - y_2)$$
2.2 LINES

**Note.** If $\lambda, \nu \in \mathbb{R}$ and $x, y \in \mathbb{R}^2$

1. $\lambda(x + y) = \lambda x + \lambda y$
2. $(\lambda + \nu)x = \lambda x + \nu x$

**Proof.**

(i) 

\[
\lambda(x + y) = \lambda(x_1 + y_1, x_2 + y_2) \\
= (\lambda(x_1 + y_1), \lambda(x_2 + y_2)) \\
= (\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2) \\
= (\lambda x_1, \lambda y_1) + (\lambda y_1, \lambda y_2) \\
= \lambda x + \lambda y
\]

(ii) exercise

\[\blacksquare\]

### 2.2 Lines

In Euclid, points and lines are undefined. But we need to define points and lines. We have already defined points.

**Definition 2.3.** Let $a, u$ be vectors in $\mathbb{R}^2$ with $u \neq \mathbf{0}$. The set of vectors \( \{a + \lambda u \mid \lambda \in \mathbb{R}\} \) is called a line.

**Definition 2.4.** Two lines \( \{a + \lambda u \mid \lambda \in \mathbb{R}\} \) and \( \{b + \lambda v \mid \lambda \in \mathbb{R}\} \) are parallel iff
\[
v = \alpha u \text{ for some } \alpha \in \mathbb{R}.
\]

Next we will prove the three famous propositions from Euclid.

**Proposition 2.1.** If $L, L'$ are parallel lines which have a point in common, then $L = L'$.

**Proposition 2.2.** Through any two points, there is exactly one line.

**Proposition 2.3.** Two non parallel lines meet in exactly one point.

**Proof.** Proposition 2.1.

Let 

\[
L = \{a + \lambda u \mid \lambda \in \mathbb{R}\} \\
L' = \{b + \lambda v \mid \lambda \in \mathbb{R}\}
\]

As $L, L'$ are parallel, $v = \alpha u$ for some scale $\alpha \in \mathbb{R}$, so 

\[
L' = \{b + \lambda \alpha u \mid \lambda, \alpha \in \mathbb{R}\} = \{b + \nu u \mid \nu \in \mathbb{R}\}
\]
Given that $L$ and $L'$ have $x$ in common, we have

$$ x = a + \lambda_1 u = b + \lambda_2 u \quad \lambda_1, \lambda_2 \in \mathbb{R} $$

$$ b = a + \lambda_1 u - \lambda_2 u $$

$$ b = a + \lambda_3 u \quad (\lambda_3 = (\lambda_1 - \lambda_2) \rightarrow \lambda_3 \in \mathbb{R}) $$

Finally, we get

$$ L' = \{ b + \lambda u \mid \lambda \in \mathbb{R} \} $$

$$ = \{ a + \lambda_3 u + \lambda u \mid \lambda \in \mathbb{R} \} $$

$$ = \{ a + (\lambda_3 + \lambda) u \mid \lambda \in \mathbb{R} \} $$

$$ = \{ a + \lambda u \mid \lambda \in \mathbb{R} \} $$

$$ = L $$

**Proof.** Proposition 2.2.

Let $u, v \in \mathbb{R}^2$ to be our two points. Let us take the line

$$ L = \{ u + \lambda (v - u) \mid \lambda \in \mathbb{R} \} $$

Suppose now that $L'$ is some line through $u$ and $v$

$$ L' = \{ a + \lambda w \mid \lambda \in \mathbb{R} \} $$

Since points $u$ and $v$ line on $L'$, we get

$$ u = a + \lambda_1 w $$

$$ v = a + \lambda_2 w $$

$$ v - u = (\lambda_2 - \lambda_1) w $$

Therefore, lines $L$ and $L'$ are parallel. As they also have a common point, they are (by proposition 1) identical.

**Example 2.1.** Find common point of

$$ L_1 = \{ (0, 1) + \lambda(1, 1) \mid \lambda \in \mathbb{R} \} $$

$$ L_2 = \{ (4, 0) + \lambda(2, 1) \mid \lambda \in \mathbb{R} \} $$

By Proposition 3, there exists a unique common point, call it $x$.

$$ x = (0, 1) + \lambda_1 (1, 1) $$

$$ x = (4, 0) + \lambda_2 (2, 1) $$

Writing equations for individual coordinates, we get

$$ 0 + \lambda_1 = 4 + 2 \lambda_2 $$

$$ 1 + \lambda_1 = 0 + \lambda_2 $$

$$ -1 = 4 + \lambda_2 $$

$$ \lambda_2 = -5 $$

$$ \lambda_1 = 4 + 2(-5) $$

$$ \lambda_1 = -6 $$
Therefore

\[ x = (0, 1) - 6(1, 1) = (-6, -5) \]
\[ = (4, 0) - 5(2, 1) = (-6, -5) \]

2.3 Triangles

**Definition 2.5.** A *triangle* is a set of 3 non-colinear points \( \{a, b, c \mid a, b, c \in \mathbb{R}^2\} \).

*Edges* of a triangle are the line segments \( ab, ac, bc \), where

\[ ab = \{a + \lambda(b - a) \mid 0 \leq \lambda \leq 1\} \]

*Midpoint* of \( ab \) is the point

\[ a + \frac{1}{2}(b - 1) = \frac{1}{2}(a + b) \]

A *median* of a triangle is a line joining one of \( a, b, c \) with the midpoint of the opposite side (line segment).

**Proposition 2.4.** The 3 medians of a triangle meet in a common point.

*Proof.* Let the three medians be \( M_a, M_b \) and \( M_c \)

\[ M_a = \{a + \lambda(\frac{1}{2}(b + c) - a) \mid \lambda \in \mathbb{R}\} \]
\[ M_b = \{b + \lambda(\frac{1}{2}(a + c) - b) \mid \lambda \in \mathbb{R}\} \]
\[ M_c = \{c + \lambda(\frac{1}{2}(a + b) - c) \mid \lambda \in \mathbb{R}\} \]

We just show, that for \( \lambda = \frac{2}{3} \), medians meet at the same point

\[ M_a = a + \frac{1}{3}b + \frac{1}{3}c - \frac{2}{3}a = \frac{1}{3}(a + b + c) \]
\[ M_b = b + \frac{1}{3}a + \frac{1}{3}c - \frac{2}{3}b = \frac{1}{3}(a + b + c) \]
\[ M_c = c + \frac{1}{3}a + \frac{1}{3}b - \frac{2}{3}c = \frac{1}{3}(a + b + c) \]

Therefore, all the medians contain the point (sometimes called *centroid*) \( \frac{1}{3}(a + b + c) \).

*Note.* Other interesting properties of triangles are

- 3 altitudes meet at a point (sometimes called *orthocentre*)
- 3 perpendicular bisectors meet at a point (sometimes called *circumcentre*)
- the 3 centres (orthocentre, circumcentre, centroid) are colinear, (they lie on *Euler line*)
2.4 Distances

Definition 2.6. The length of a vector \( x = (x_1, x_2) \) (denoted \( \|x\| \)) is a real number

\[
\|x\| = \sqrt{x_1^2 + x_2^2}
\]

The distance between \( x \) and \( y \) \((x, y \in \mathbb{R}^2)\) is the length of \((x - y)\) called \( \text{dist}(x, y) \)

\[
\text{dist}(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}
\]

2.5 Dot product

Definition 2.7. For \( x, y \in \mathbb{R}^2 \), dot product (scalar product) of \( x \) and \( y \) is a real number

\[
x \cdot y = x_1y_1 + x_2y_2
\]

Example 2.2. \( x = (1, 2), y = (-1, 0) \)

\[
\begin{align*}
\|x\| &= \sqrt{1^2 + 2^2} = \sqrt{5} \\
\text{dist}(x, y) &= \sqrt{2^2 + 2^2} = 2\sqrt{2} \\
x \cdot y &= -1 + 0 = -1
\end{align*}
\]

Proposition 2.5.  
(i) \( \|x\|^2 + \|y\|^2 - \|x - y\|^2 = 2x \cdot y \)

(ii) If \( \theta \) is the angle between \( x \) and \( y \) then \( x \cdot y = \|x\| \|y\| \cos \theta \)

(iii) The position vectors of \( x \) and \( y \) are at right angles iff \( x \cdot y = 0 \)

Proof.

(i)

\[
\begin{align*}
\text{LHS} &= x_1^2 + x_2^2 + y_1^2 + y_2^2 - (y_1 - x_1)^2 - (y_2 - x_2)^2 \\
&= -(2x_1y_1 - 2x_2y_2) \\
&= 2x \cdot y
\end{align*}
\]

(ii) From cosine rule

\[
\frac{\|x\|^2 + \|y\|^2 - \|y - x\|^2}{2\|x\|\|y\|} = \cos \theta
\]

Using part (i), we get

\[
2x \cdot y = 2\|x\|\|y\| \cos \theta \\
x \cdot y = \|x\|\|y\| \cos \theta
\]

\[\blacksquare\]
2.6 Equation of a line

Consider a line \( L = \{ a + \lambda u \mid \lambda \in \mathbb{R} \} \). If \( u = (u_1, u_2) \), define \( n = (-u_2, u_1) \), so that \( n.u = 0 \).

For any \( x \in L \), \( x = a + \lambda u \)

\[
x.n = (a + \lambda u).n = a.n + \lambda u.n = a.n
\]

Proposition 2.6. Let \( L = \{ a + \lambda y \mid \lambda \in \mathbb{R} \} \), \( n = (-u_2, u_1) \). Then

(i) Every \( x \) in \( L \) satisfies \( x.n = a.n \)

(ii) Every solution of \( x.n = a.n \) lies on \( L \)

Proof.

(ii) Suppose \( x.n = a.n \) (we need to show that \( x = a + \lambda u \), i.e. \( x - a = \lambda u \)). Then

\[
(x - a).n = 0.
\]

Let \( y = x - a \), and say \( y = (y_1, y_2) \). As \( y.n = 0 \)

\[
-y_1 u_2 + y_2 u_1 = 0
\]

\[
y_2 u_1 = y_1 u_2
\]

So, \( y = \frac{u_1}{u_2}(u_1, u_2) \) if \( u_1 \neq 0 \), or \( y = \frac{u_2}{u_1}(u_1, u_2) \) if \( u_2 \neq 0 \).

Therefore

\[
y = \lambda u = x - a
\]

\[
x = a + \lambda u \in L
\]

Definition 2.8. For a line \( L = \{ a + \lambda u \mid \lambda \in \mathbb{R} \} \), the vector \( n = (-u_2, u_1) \) (or any scalar multiple of it) is called a normal to \( L \). The equation \( x.n = a.n \) is called the equation of \( L \).

Proposition 2.7. The linear equation \( px_1 + qx_2 + r = 0 \) (1) is an equation of a line with normal \((p, q)\) and with direction \((-q, p)\).

Proof.

Suppose \( q \neq 0 \). Then the solution of (1)

\[
\frac{p}{q} x_1 + x_2 + \frac{r}{q} = 0
\]

Let \( x_1 = \lambda \), \( x_2 = \frac{-r}{q} - \frac{p\lambda}{q} \), so

\[
(x_1, x_2) = \left( \lambda, \frac{-r}{q} - \frac{p\lambda}{q} \right)
\]

\[
= \left( 0, \frac{-r}{q} \right) + \lambda \left( 1, \frac{-p}{q} \right)
\]

which is a line with direction vector \( (1, \frac{-p}{q}) = \frac{1}{q}(-q, p) \). Normal is clearly \((p, q)\).
When \( q = 0 \). Equation (1) is \( px + r = 0 \), with solution

\[
(x_1, x_2) = \left( \frac{-r}{p}, 0 \right) + \lambda(0, 1)
\]

Which is clearly a line with equation \( \frac{1}{p}(0, p) = \frac{1}{p}(-q, p) \).

**Definition 2.9.** Lines \( L_1 = \{a + \lambda u \mid \lambda \in \mathbb{R}\} \) and \( L_2 = \{b + \lambda v \mid \lambda \in \mathbb{R}\} \) are perpendicular iff \( u \cdot v = 0 \).

**Proposition 2.8.** Let lines \( L_1, L_2 \) have equations

\[
L_1 : p_1x_1 + q_1x_2 + r_1 = 0
\]

\[
L_2 : p_2x_1 + q_2x_2 + r_2 = 0
\]

Then

(i) \( L_1, L_2 \) are parallel iff \( (p_1, q_1) = \alpha(p_2, q_2) \) for some scalar \( \lambda \)

(ii) \( L_1, L_2 \) are perpendicular iff \( (p_1, q_1) \cdot (p_2, q_2) = 0 \)

**Proof.**

(i) By Proposition 2.7

\[
L_1 = \{a + \lambda(-q_1, p_1) \mid \lambda \in \mathbb{R}\}
\]

\[
L_2 = \{b + \lambda(-q_2, p_2) \mid \lambda \in \mathbb{R}\}
\]

These are parallel iff \( \exists \alpha \in \mathbb{R} \) such that \( (-q_2, p_2) = \alpha(-q_1, p_1) \), i.e. \( \alpha(p_1, q_1) = (p_2, q_2) \).

(ii) excercise

\[\blacksquare\]

### 2.7 Perpendicular distance

**Definition 2.10.** For a line \( L \) and point \( p, p \notin L \), \( \text{dist}(p, L) \) is the length of the perpendicular from \( p \) to \( L \).

**Proposition 2.9.** If \( a \) is a point on the line and \( n \) is a normal to the line, then

\[
\text{dist}(p, L) = \left| \frac{(p - a) \cdot n}{\|n\|} \right|
\]

**Note.** A unit vector is a vector of length 1. If \( w \) is any vector, then \( \hat{w} = \frac{w}{\|w\|} \) is a unit vector. Therefore we can write

\[
\text{dist}(p, l) = |(p - a) \cdot \hat{n}|
\]
2.8. TWO FAMOUS INEQUALITIES

**Proposition 2.10.** (Cauchy - Schwartz inequality)
Let $x, y \in \mathbb{R}^2$.

(i) $|x \cdot y| \leq \|x\|\|y\|

(ii) If $|x \cdot y| = \|x\|\|y\|$ then one of $x, y$ is a scalar multiple of the other

**Proof.**

(i) $|x \cdot y| \leq \|x\|\|y\|

$\iff (x_1y_1 + x_2y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2)

\iff x_1^2y_2^2 + 2x_1y_1x_2y_2 + x_2^2y_2^2 \leq x_1^2y_1^2 + x_2^2y_1^2 + x_2^2y_2^2

\iff 0 \leq x_1^2y_2^2 + x_2^2y_1^2 - 2x_1y_1x_2y_2

\iff (x_1y_2 - x_2y_1)^2 \geq 0$

(ii) If $|x \cdot y| = \|x\|\|y\|$, above proof shows

$0 = x_1y_2 - x_2y_1

\frac{x_1}{x_2} = \frac{y_1}{y_2}$

Therefore one of $x$ and $y$ is a scalar multiple of the other.

Note that $x \cdot x = (x_1^2 + x_2^2) = \|x\|^2$.

**Proposition 2.11.** (Triangle inequality) If $x, y \in \mathbb{R}^2$, then

$\|x + y\| \leq \|x\| + \|y\|$
2.8. TWO FAMOUS INEQUALITIES

Proof.

\[ \|x + y\|^2 = (x + y)(x + y) = x \cdot x + y \cdot y + 2x \cdot y \]
\[ = \|x\|^2 + \|y\|^2 + 2x \cdot y \]
\[ \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \quad \text{by Cauchy - Schwartz} \]
\[ \leq (\|x\| + \|y\|)^2 \]

Therefore \( \|x + y\| \leq \|x\| + \|y\| \).

\[ \blacksquare \]
2.8. TWO FAMOUS INEQUALITIES
Chapter 3

Linear inequalities
Chapter 4

Conics

In $\mathbb{R}^2$, a linear equation $px_1 + qx_2 + r = 0$ defines a line. Now we define a curve in $\mathbb{R}^2$ by quadratic equation.

Example 4.1. Circle of center $c$, radius $r$. It contains all the points $x$ such that

$$\|x - c\| = r$$
$$\|x - c\|^2 = r^2$$
$$(x_1 - c_1)^2 + (x_2 - c_2)^2 = r^2$$
$$x_1^2 + x_2^2 - 2c_2x_2 + c_1^2 + c_2^2 - r^2 = 0$$

Definition 4.1. A conic section in $\mathbb{R}^2$ is the set of points $x \in \mathbb{R}^2$, $x = (x_1, x_2)$, satisfying a quadratic equation

$$ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0 \quad (4.1)$$

where not all of $a$, $b$ and $c$ are 0.

Here are some basic examples

(1) \[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1
\]

is an ellipse

(2) \[
\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1
\]

is a hyperbola
(3) \[ x_2 = ax_1^2 + b \]

(a \neq 0) is a parabola.

(4) \[ \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 0 \]

is a pair of lines.

(5) \[ \frac{x_1^2}{a^2} = 1 \]

is a pair of lines.
We call the conics (6), (7) and (8) *degenerate* conics. We’ll see how to reduce an arbitrary conic to one of these basic examples.
Note. Why conic sections?

4.1 Transforming conics

Aim – start with an arbitrary conic in the form (4.1), do some geometrical changes to coordinates so that the equation becomes one of our standard equations (1) – (8). Two kinds of changes are allowed.

4.1.1 Translation

To change coordinates from \((x_1, x_2)\) to \((y_1, y_2)\) with new origin \((\alpha, \beta)\)

\[
\begin{align*}
y_1 &= x_1 - \alpha \\
y_2 &= x_2 - \beta
\end{align*}
\]

This is called a translation.

Example 4.2. What type is the following conic?

\[
x_1^2 + 2x_2^2 - 2x_1 + 4x_2 - 10 = 0
\]

We complete the square

\[
(x_1 - 1)^2 + 2(x_2 + 1)^2 - 13 = 0
\]

So the new coordinates are

\[
\begin{align*}
y_1 &= x_1 - 1 \\
y_2 &= x_2 + 1
\end{align*}
\]

Now, we can easily see that the our conic is an ellipse.

Remark – If (4.1) has no \(x_1x_2\) term \((c = 0)\), then we can find a translation which reduces the equation to one of the standard ones (examples (1) – (6)).
4.1.2 Rotation

Rotate axes anticlockwise through \( \theta \). What happens to the coordinates of a general point \( P \)?

Say \( P \) has old coordinates \((x_1, x_2)\) and new coordinates \((y_1, y_2)\).

Now

\[
\begin{align*}
y_1 &= r \cos \alpha \\
y_2 &= r \sin \alpha
\end{align*}
\]

and

\[
\begin{align*}
x_1 &= r \cos(\alpha + \theta) \\
x_2 &= r \sin(\alpha + \theta)
\end{align*}
\]

Hence

\[
\begin{align*}
x_1 &= r \cos \theta \cos \alpha - r \sin \alpha \sin \theta \\
    &= y_1 \cos \theta - y_2 \sin \theta
\end{align*}
\]

and

\[
\begin{align*}
x_2 &= r \sin \alpha \cos \theta + \cos \alpha \sin \theta \\
    &= y_2 \cos \theta + y_1 \sin \theta
\end{align*}
\]

Summarizing – change of coordinates when we do a rotation through angle \( \theta \) is

\[
\begin{align*}
x_1 &= y_1 \cos \theta - y_2 \sin \theta \\
x_2 &= y_1 \sin \theta + y_2 \cos \theta
\end{align*}
\]
**Example 4.3.** Rotation through $\frac{\pi}{4}$.

\[
x_1 = \frac{1}{\sqrt{2}}(y_1 - y_2) \\
x_2 = \frac{1}{\sqrt{2}}(y_1 + y_2)
\]

**Example 4.4.** What is the following conic (find the rotation and translation to change coordinates and get standard equation)

\[
x_1^2 + x_2^2 + 4x_1x_2 = 1
\]

(From the hat method) Let’s rotate through $\frac{\pi}{4}$. By magic (see later). Then our equation becomes

\[
\frac{1}{2}(y_1 - y_2)^2 + \frac{1}{2}(y_1 + y_2)^2 + 4 \frac{1}{2}(y_1 - y_2)(y_1 + y_2) = 1 \Rightarrow 3y_1^2 - y_2^2 = 1
\]

This is a hyperbola.

### 4.2 The Theory

Start with general conic

\[
a x_1^2 + b x_2^2 + c x_1 x_2 + d x_1 + e x_2 + f = 0
\]

Aim

(i) find the rotation which gets rid of $x_1 x_2$

(ii) complete the square to find the translation which changes the equation to one of our six standard equations

Part (i).

If $c = 0$, we don’t need to rotate. So assume that $c \neq 0$. When we do general rotation through $\theta$ we change coordinates to $y_1, y_2$

\[
x_1 = y_1 \cos \theta - y_2 \sin \theta \\
x_2 = y_1 \sin \theta + y_2 \cos \theta
\]

We aim to find $\theta$ so that the new equation has no $y_1 y_2$ term.

\[
a x_1^2 = a(y_1 \cos \theta - y_2 \sin \theta)^2 \\
b x_2^2 = b(y_1 \sin \theta + y_2 \cos \theta)^2 \\
c x_1 x_2 = c(y_1 \sin \theta + y_2 \cos \theta)(y_1 \cos \theta - y_2 \sin \theta)
\]

So the $y_1 y_2$ term when we change coordinates will be

\[
-2a \sin \theta \cos \theta + 2b \sin \theta \cos \theta + c(\cos^2 \theta - \sin^2 \theta)
\]

\[
= (b - a) \sin 2\theta + c \cos 2\theta
\]
4.2. THE THEORY

So we want to choose \( \theta \) to make this expression zero.

\[
(a - b) \sin 2\theta = c \cos 2\theta
\]

If \( a \neq b \) then

\[
\tan 2\theta = \frac{c}{a - b}
\]

If \( a = b \), we want to \( \cos 2\theta = 0 \), so take \( \theta = \frac{\pi}{4} \).

**Summary**

- **Step 1 – Rotation.** If \( c \neq 0 \) in (4.1), we rotate through \( \theta \), where

  \[
  \theta = \frac{\pi}{4}
  \]

  when \( a = b \) or

  \[
  \tan 2\theta = \frac{c}{a - b}
  \]

  when \( a \neq b \).

- **Step 2 – Translation.** After Step 1, equation becomes

  \[
  a'y_1^2 + b'y_2^2 + d'y_1 + e'y_2 + f' = 0
  \]

  Now we complete the square to find a translation which changes equation to one of the standard ones.

We’ve proved:

**Theorem 4.1.** Every conic in the form (4.1) can be changed by rotation and translation of the axes to one of the standard equations (1) – (8). Thus every conic is either an ellipse, hyperbola, parabola, or one of the degenerate conics.

**Example 4.5.** Reduce conic

\[
2x_1^2 + 2\sqrt{3}x_1x_2 + \left(3\sqrt{3} - 1\right)x_1 + \left(3 + \sqrt{3}x_2 = 0\right)
\]

to standard form by rotation and translation.

- **Step 1 – Rotation.** Here \( a = 2, b = 0, c = 2\sqrt{3} \). So rotate through \( \theta \), where

  \[
  \tan 2\theta = \sqrt{3}
  \]

  Therefore \( 2\theta = \frac{\pi}{3}, \) so \( \theta = \frac{\pi}{6} \). So, our new coordinates are

  \[
  x_1 = y_1 \cos \theta - y_2 \sin \theta = \frac{\sqrt{3}}{2}y_1 - \frac{1}{2}y_2
  \]

  \[
  x_2 = y_1 \sin \theta + y_2 \cos \theta = \frac{1}{2}y_1 + \frac{\sqrt{3}}{2}y_2
  \]

  So, the equation becomes

  \[
  2\frac{1}{4} \left(\sqrt{3}y_1 - y_2\right)^2 + 2\sqrt{3} \frac{1}{4} \left(\sqrt{3}y_1 - y_2\right) (y_1 + \sqrt{3}y_2) \\
  + (3\sqrt{3} - 1) \frac{1}{2} (\sqrt{3}y_1 - y_2) (3 + \sqrt{3}) \frac{1}{2} (y_1 + \sqrt{3}y_2) = 0
  \]

  Which is

  \[
  3y_1^2 - y_2^2 + 6y_1 + 2y_2 = 0
  \]
Step 2 – Translation – complete the square

\[3(y_1 + 1)^2 - (y_2 - y_1)^2 = 2\]

So put \(z_1 = y_1 + 1\), \(z_2 = y_2 - 1\), and equation is now standard equation

\[3z_1^2 - z_2^2 = 2\]

This is a hyperbola. Sketch – new origin has \(y_1, y_2\) coordinates \((-1, 1)\), so has \(x_1, x_2\) coordinates \((-\frac{1}{2}(\sqrt{3}+1), \frac{1}{2}(-1 + \sqrt{3}))\)

**Note.** Usually \(\tan 2\theta = \frac{2\tan \theta}{1-\tan^2 \theta}\) is not so convinient (can’t write what \(2\theta\) is). In general, use formula

\[\tan 2\theta = \frac{2\tan \theta}{1-\tan^2 \theta}\]

\[
\text{to work out } \tan \theta, \text{ hence } \sin \theta \text{ and } \cos \theta.
\]

**Note.** Standard parabola is \(x_2 = ax_1^2 + b\). What about parabola \(x_1 = ax_2^2 + b\)? We can rotate this through \(\frac{\pi}{2}\) to the standard equation.

### 4.3 Geometrical definition of conics

**Ingredients of this definition**

- a line \(L\)
- a point \(p\) not on \(L\)
- a real number \(e > 0\)

**Definition 4.2.** Curve \(C\) is a set of all points \(x \in \mathbb{R}^2\), such that

\[\|x - p\| = e \cdot \text{dist}(x, L)\]

i.e.

\[C = \{x \in \mathbb{R}^2 \mid \|x - p\| = e \cdot \text{dist}(p, L)\}\]

**Example 4.6.** Let \(e = 1\).

Let \(e = \frac{1}{2}\)

**Theorem 4.2.** Curve \(C\) (from previous definition) is a conic. It is

- a parabola if \(e = 1\)
- an ellipse if \(e < 1\)
- a hyperbola if \(e > 1\)

**Proof.** Do a rotation and translation to make \(p\) the origin and \(L\) vertical line \(x_1 = s\). Then the equation defining \(C\) is

\[\|x - p\| = e \cdot \text{dist}(x, L)\]
4.3. GEOMETRICAL DEFINITION OF CONICS

i.e. \[ \|x\| = e \cdot |x_1 - s| \]
i.e. \[ \|x\|^2 = e^2(x_1 - s)^2 \]
i.e. \[ x_1^2 + x_2^2 = e^2(x_1^2 - 2sx_1 + s^2) \]
i.e. \[ x_1^2(1-e^2) + x_2^2 + 2se^2x_1 - s^2e^2 = 0 \]

If \( e = 1 \), the \( x_1^2 \) term vanishes – this is a parabola.

Now suppose \( e \neq 1 \).

Complete square

\[ (1-e^2) \left( x_1 + \frac{se^2}{1-e^2} \right)^2 + x_2^2 - s^2e^2 - \frac{s^2e^4}{1-e^2} = 0 \]

So put \( y_1 = x_1 + \frac{se^2}{1-e^2} \) and \( y_2 = x_2 \). Then get standard equation

\[ (1-e^2)y_1^2 + y_2^2 = s^2e^2 \left( 1 + \frac{e^2}{1-e^2} \right) \]
i.e. \[ y_1^2 + \frac{y_2^2}{1-e^2} = \frac{s^2e^2}{(1-e^2)^2} \] (4.2)

If \( e < 1 \), this is an ellipse. If \( e > 1 \), this is hyperbola.

\[ \text{Definition 4.3.} \quad \text{The conic has focus } p, \text{ directrix } L, \text{ excentricity } e \text{ from the previous proof.} \]

\[ \text{Example 4.7.} \quad \text{Find } e, p \text{ and } L \text{ for the ellipse} \]

\[ \frac{x_1^2}{2} + x_2^2 = 1 \]

This is the standard equation. We compare it with (4.2)

\[ y_1^2 + \frac{y_2^2}{1-e^2} = \frac{s^2e^2}{(1-e^2)^2} \] (4.3)

From the \( y \) coordinate picture, the focus is \( \left( \frac{se^2}{1-e^2}, 0 \right) \), directrix is \( x_1 = s + \frac{se^2}{1-e^2} \).

Compare equations

\[ y_1^2 + \frac{y_2^2}{1-e^2} = \frac{s^2e^2}{(1-e^2)^2} \]

\[ x_1^2 + 2x_2^2 = 2 \]

So,

\[ \frac{1}{1-e^2} = 2 \]

\[ \frac{s^2e^2}{(1-e^2)^2} = 2 \]

So \( e = \frac{1}{\sqrt{2}} \), \( s = 1 \). So the excentricity is \( \frac{1}{\sqrt{2}} \), focus \( (1,0) \) and the directrix is \( x_1 = 2 \).

\[ \text{Note.} \quad \text{Ellipse has in fact two foci } \pm p, \text{ and two directrices } L, L'. \]
Chapter 5

Matrices and linear equations

Definition 5.1. The space \( \mathbb{R}^n \) – define

\[
\begin{align*}
\mathbb{R}^1 &= \mathbb{R} \\
\mathbb{R}^2 &= \text{set of ordered pairs} \\
\mathbb{R}^3 &= \text{set of all triples} \\
\mathbb{R}^4 &= \text{set of all quadruples}
\end{align*}
\]

In general, \( \mathbb{R}^n \) is the set of all \( n \)-tuples \((x_1, x_2, \ldots, x_n)\) with \( x_i \in \mathbb{R} \). Call these \( n \)-tuples vectors.

Note that \( \mathbb{R}^i \nsubseteq \mathbb{R}^{i+1} \).

The \( \mathbb{R}^n \) has interesting structure

- geometric – points, lines, curves
- algebraic

Definition 5.2. Addition of vectors

\[
(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)
\]

Scalar multiplication

\[
\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)
\]

Definition 5.3. A linear equation in \( x_1, x_2, \ldots, x_n \) is

\[
a_1x_1 + a_2x_2 + \cdots + a_nx_n = b
\]

where coefficients \( a_i, b \in \mathbb{R} \).

Definition 5.4. A solution to a linear equation is a vector \((k_1, \ldots, k_n)\) such that the equation is satisfied when we put \( x_i = k_i \).

Example 5.1. \((1, -2, 3)\) is a solution to the linear equation \( x_1 + x_2 + x_3 = 2 \).

Definition 5.5. A system of linear equations in \( x_1, \ldots, x_n \) is a collection of one or more linear equations in these variables.

Example 5.2.
1) 
\[
3x_1 + 4x_2 = 5 \\
8x_1 - x_2 = -2
\]

is a system of 2 linear equations in \(x_1, x_2\).

2) 
\[
x_1 + x_2 + x_3 + x_4 = 0 \\
2x_1 - x_2 + 5x_4 = -2 \\
x_1 + x_2 - x_4 = 3
\]

is a system of 3 equations in \(x_1, x_2, x_3, x_4\).

**Definition 5.6.** General system
\[
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\]

where \(x_1, \ldots, x_n\) are unknowns and \(a_{ij}, b_i\) are constants. A solution to this system is a vector \((k_1, \ldots, k_n)\) which satisfies all the equations.

Aim is to find the method to find all solutions of any system of the form from previous definition.

**Example 5.3.** System
\[
x_1 + x_1 = 1 \\
-2x_1 + 3x_2 = 5
\]

Eliminate \(x_1\) – take twice first, add to second
\[
5x_2 = 7
\]

One solution \((x_1, x_2) = (-2/5, 7/5)\).

**Example 5.4.**
\[
2x_1 + x_2 = 2 \\
-6x_1 - 3x_2 = 1
\]

These are just two parallel (nonidentical) lines – the system has no solution.

**Example 5.5.** System
\[
x_1 + x_2 = 0 \\
x_2 - x_3 = 0 \\
x_1 - x_2 + 2x_3 = 0
\]

Notice (cleverly), that third equation is equal to first minus two times second. So any solution of system of first and second equation will automatically satisfy third.

So general solution is
\[
(x_1, x_2, x_3) = (-a, a, a)
\]

for any \(a \in \mathbb{R}\). This system has infinitely many solutions. Soon we will see that every system has either no solution, one solution, or unlimited number of solution.
5.1 Method – Gaussian elimination

Example 5.6. System

\[
\begin{align*}
    x_1 + 2x_2 + 3x_3 &= 9 \\
    4x_1 + 5x_2 + 6x_3 &= 24 \\
    3x_1 + x_2 - 2x_3 &= 4 \\
\end{align*}
\]

Step 1 – Eliminate \(x_1\) from second and third, using first. We get equivalent system

\[
\begin{align*}
    x_1 + 2x_2 + 3x_3 &= 9 \\
    -3x_2 - 6x_3 &= -12 \quad (2) - 4(1) \\
    -5x_2 - 11x_2 &= -23 \quad (3) - 3(1) \\
\end{align*}
\]

This system has the same solutions as the original one as new equations are combinations of original ones, and vice versa.

Step 2 – Eliminate \(x_2\) from the third equation using only second equation.

\[
\begin{align*}
    x_1 + 2x_2 + 3x_3 &= 9 \\
    -3x_2 - 6x_3 &= -12 \\
    -x_3 &= -3 \quad (3) - \frac{5}{3}(2) \\
\end{align*}
\]

Step 3 – Solve!

By third, \(x_3 = 3\). By the second, \(3x_2 = 12 - 18\), so \(x_2 = -2\). By the first, \(x_1 = 4\). So the system has one solution \((4, -2, 3)\).

For bigger systems, we need better notation. This is provided by matrices.

Definition 5.7. A matrix is a rectangular array of numbers. Eg

\[
\begin{pmatrix}
    1 & 2 & 3 \\
    4 & 5 & 6
\end{pmatrix}
\]

This is a \(2 \times 3\) matrix.

\[
\begin{pmatrix}
    1 \\
    5 \\
    -\pi
\end{pmatrix}
\]

This is a \(3 \times 1\) matrix.

Call matrix \(m \times n\) if it has \(m\) rows, \(n\) columns. We use matrices to encapsulate systems of linear equations. System from previous example has the coefficients matrix

\[
\begin{pmatrix}
    1 & 2 & 3 \\
    4 & 5 & 6 \\
    3 & 1 & -2
\end{pmatrix}
\]

and augmented matrix

\[
\begin{pmatrix}
    1 & 2 & 3 & 9 \\
    4 & 5 & 6 & 24 \\
    3 & 1 & -2 & 4
\end{pmatrix}
\]
For general system

coeff matrix = \[
\begin{pmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
\vdots & & & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn}
\end{pmatrix}
\]

augmented matrix = \[
\begin{pmatrix}
a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\
\vdots & & & \vdots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn} & b_m
\end{pmatrix}
\]

A row of augmented matrix corresponds to an equation in the system. The number in the \(i\)-th row and \(j\)-th column of a matrix is called the \(ij\)-entry.

Operation on the equations in a system are operations on the rows of the augmented matrix.

\[
\begin{pmatrix}
1 & 2 & 3 & 9 \\
4 & 5 & 6 & 24 \\
3 & 1 & -2 & 4
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 3 & 9 \\
0 & -3 & -6 & -12 \\
0 & -5 & -11 & -23
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 3 & 9 \\
0 & -3 & -6 & -12 \\
0 & 0 & -1 & -3
\end{pmatrix}
\]

**Definition 5.8.** Elementary row operations are the following operations on the rows of an (augmented) matrix

1) Add a scalar multiple of one row to another

\[r_1 \rightarrow r_i + \lambda r_j, \lambda \in \mathbb{R}\]

2) Swap two rows

\[r_i \leftrightarrow r_j\]

3) Multiply any row by a nonzero scalar

\[r_i \rightarrow \lambda r_i, \lambda \neq 0\]

Doing these elementary row operations to an augmented matrix does not change the solution of the system.

**Idea of Gaussian elimination**

We start with a “hard” system. Then we do a couple of row operations and get an “easy” system. What makes it easy are zeros under the main diagonal.

**Definition 5.9.** An \(m \times n\) matrix is in echelon form if

(i) The first non-zero number in each row occurs to the right of the first non-zero number in any higher row

(ii) All rows \((0, 0, \ldots, 0)\) appear at the bottom.
Example 5.7. The following matrices
\[
\begin{pmatrix}
1 & 2 & 3 & 9 \\
0 & -3 & -6 & -12 \\
0 & 0 & -3 & -9 \\
\end{pmatrix}
\begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
\end{pmatrix}
\]
are in echelon form. This one is not
\[
\begin{pmatrix}
0 & 1 & 3 \\
1 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
\]

The point If a system has its augmented matrix in echelon form, the system is easy
to solve.

Example 5.8. Solve the system of equations with augmentated matrix
\[
\begin{pmatrix}
2 & -1 & 3 & 0 \\
0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]
This system is
\[
2x_1 - x_2 + 3x_3 = 0 \\
x_2 + x_3 = 1
\]
Solve from bottom up. Equation 3 tells us nothing. Let \(x_3 = a\) (any \(a \in \mathbb{R}\). Then eq
2 implies that \(x_2 = 1 - a\).
Equation 1 implies \(2x_1 = x_2 - 3x_2\), therefore \(x_1 = \frac{1-4a}{2}\).
Therefore the solutions are \((x_1, x_2, x_3) = (\frac{1}{2}(1 - 4a), 1 - a, a)\). E.g when \(a = 0\), we
get solution \((\frac{1}{2}, 1, 0)\).

Example 5.9. Solve the system with augmentated matrix
\[
\begin{pmatrix}
2 & -1 & 3 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 2 \\
\end{pmatrix}
\]
System thus is
\[
2x_1 - x_2 + 3x_3 = 0 \\
x_2 + x_3 = 1 \\
0 = 2
\]
Third equation instantly implies no solution at all.
Example 5.10.

$$\begin{pmatrix} 1 & 1 & 1 & 3 & 2 & 4 & 0 \\ 0 & 0 & 0 & 2 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

The system is

$$x_1 + x_2 + x_3 + 3x_4 + 2x_5 + 4x_6 = 0$$
$$2x_4 - x_5 = 3$$
$$x_5 + x_6 = 0$$

Equation 3 sets $x_6 = a$ (any $a$). Then $x_5 = -a$. From 2 we get $x_4 = \frac{1}{2}(x_5 + 3) = \frac{1}{2}(3 - a)$. From 1 we get $x_3 = b$ and $x_2 = c$. Then $x_1 = -c - b - \frac{3}{2}(3 - a) + 2a - 4a = \frac{-9}{2} - \frac{1}{2}a - b - c$. The general solution is

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (-\frac{9}{2} - \frac{1}{2}a - b - c, c, b, \frac{1}{2}(3 - a), a, a)$$

for any $a, b, c \in \mathbb{R}$.

In general, if augmented matrix is in echelon form, solve the system by solving the last equation, then the next last, and so on.

This method is called **back substitution**.

The variables we can put free equal to $a, d, c$ etc are **free variables**. E.g. in previous example, the free variables are $x_6, x_3, x_2$.

**Theorem 5.1.** (Gaussian elimination theorem)

Any matrix can be reduced by elementary row operations to a matrix which is in echelon form.

**Method 5.1.** (Gaussian elimination method)

System of linear equations (augmented matrix). We put the augmented to echelon form using elementary row operations. Solve the new system by back substitution.

Example 5.11. Solve the system with augmentation matrix

$$A = \begin{pmatrix} 1 & 0 & 3 & 2 & 5 & 1 \\ 1 & 0 & 3 & 5 & 8 & 1 \\ 1 & 1 & 4 & 5 & 7 & 0 \\ 1 & -1 & 2 & -1 & 3 & 3 \end{pmatrix}$$

Answer.

- Step 1 – clear first column using top left hand entry (i.e. equations 2, 3, 4 take away equation 1)

$$A \rightarrow \begin{pmatrix} 1 & 0 & 3 & 2 & 5 & 1 \\ 0 & 0 & 0 & 3 & 3 & 0 \\ 0 & 1 & 1 & 3 & 2 & -1 \\ 0 & -1 & -1 & -3 & -2 & 1 \end{pmatrix}$$
• Step 2 – swap rows 2 and 4

\[
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 1 & 1 & 3 & 2 \\
0 & 0 & 0 & 3 & 3 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 1 & 1 & 3 & 2 \\
0 & 0 & 0 & 3 & 3 \\
\end{pmatrix}
\]

• Step 3 – clear second column using the row 2

\[
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 3 & 3 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

• Step 4

\[
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 0 & 3 & 2 & 5 \\
0 & -1 & -1 & -3 & -2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

• Step 5 – solve by back substitution Let \( x_5 = a \). Then \( x_4 = -a \), \( x_3 = b \), \( x_2 = -1 - b + a \), \( x_1 = 1 - 3b - 3a \).

**Matrix Algebra**

**Definition 5.10.** Matrix multiplication

Dot product in \( \mathbb{R}^n \)

\[
x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}
\]

\[
x \cdot y = x_1y_1 + \cdots + x_ny_n
\]

It is convenient given a row vector \( z = (z_1, \ldots, z_n) \) to also define \( z \cdot y = z_1y_1 + \cdots + z_ny_n \).

Then one can define matrix multiplication by

\[
A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad B = (b_1 \ldots b_n)
\]

\[
AB = \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mp} \end{pmatrix}
\]

where \( c_{ij} = a_i b_j \)

**Note.** If \( A \) is \( 1 \times n \ \begin{pmatrix} a_1 \ldots a_n \end{pmatrix} = a \), \( B \) is \( n \times 1 \ \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = b \), \( AB \) is a \( 1 \times 1 \) matrix \( (a \cdot b) \).
Key properties of matrix multiplication

1) One can multiply an \( m \times n \) matrix \( A \) with an \( r \times s \) matrix \( B \) iff \( r = n \). Then, \( AB \) will be an \( m \times s \) matrix.

2) Commutativity fails, i.e. in general \( AB \neq BA \).
   In order for both \( AB \) and \( BA \) to be defined, they must be square matrices of same size, i.e. \( n \times n \) for some \( n \).

Example 5.12.

\[
A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}
\]

Then

\[
AB = \begin{pmatrix} -2 & 5 \\ -4 & 11 \end{pmatrix}, \quad BA = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}
\]

3) Matrix multiplication, like for \( \times \) on \( \mathbb{R} \), is associative.

Theorem 5.2. Given any matrices \( A m \times n \), \( B n \times p \), \( C p \times q \)

\[(AB)C = A(BC)\]

Example 5.13.

\[
A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}
\]

Then

\[
AB = \begin{pmatrix} -2 & 5 \\ -4 & 11 \end{pmatrix}, \quad (AB)C = \begin{pmatrix} 13 & -4 \\ 29 & -8 \end{pmatrix}
\]

\[
BC = \begin{pmatrix} 3 & 0 \\ 5 & -2 \end{pmatrix}, \quad A(BC) = \begin{pmatrix} 13 & -4 \\ 29 & -8 \end{pmatrix}
\]

Proof. First, we do special case, when \( C \) is \( p \times 1 \) matrix, i.e. column vector. Let \( A \) is \( m \times n \), \( B \) is \( n \times p \), \( x \) is \( p \times 1 \). We show that

\[(AB)x = A(Bx)\]

Let \( y = Bx \) be a vector and \( z = Ay \). So \( z \) is the right hand side. Claim is \( z = (AB)x \).

\[
y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_{11} & \ldots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \ldots & b_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}
\]

\[
(n \times 1) (n \times p) (p \times 1)
\]
5.1. METHOD – GAUSSIAN ELIMINATION

\[
z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}
\]

For \( i \leq n \) \( z_i = a_{i1}y_1 + \cdots + a_{in}y_n \), \( y_j = b_{ji}x_1 + \cdots + b_{jp}x_p \).

Then

\[
z_i = a_{i1}(b_1x_1 + \cdots + b_{1p}) + a_{i2}(b_2x_1 + \cdots + b_{2p}) + \\
\vdots \\
+ a_{in}(b_mx_1 + \cdots + b_{np})
\]

\[
= (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})x_1 + (a_{i2}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2})x_2 \\
+ \cdots + (a_{i2}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np})x_p
\]

Thus coefficient of \( x_i \) in \( z_i \) is

\[
(a_{i1}b_{ij} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj})
\]

which is \( a_i.b_j \), where

\[
a_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}, \quad b_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}
\]

which by definition is the \((i,j)\) entry of \( AB \).

Claim

\[(AB)C = A(BC)\]

Write \( C = \begin{pmatrix} \vdots \\ c_1 \\ \vdots \\ c_q \end{pmatrix} \)

Then

\[
BC = (Bc_1 \ldots Bc_q)
\]

Therefore

\[
A(BC) = (A(Bc_1) \ldots A(Bc_q))
\]

4) Again, as for + and × on \( \mathbb{R} \), distributivity holds

**Proposition 5.3.** Given matrices \( A_{m \times n}, B, C \), both \( n \times p \)

\[
A(B + C) = AB + AC
\]
Powers of matrices

**Definition 5.11.** A square matrix is one which is $n \times n$ for some $n$.

If $A$ is $n \times n$, define

\[ A^2 = AA \]

\[ A^3 = (AA)A = A(AA) \]

\[ A^4 = A^3A \]

\[ \ldots \]

\[ A^n = A^{n-1}A \]

**Example 5.14.**

\[
\begin{pmatrix}
1 & 2 \\
0 & 1
\end{pmatrix}^2 = \begin{pmatrix} 1 & 2 \\
0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\
0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\
0 & 4 \end{pmatrix}
\]

**Definition 5.12.** The identity matrix is

\[
2 \times 2 I_2 = \begin{pmatrix} 1 & 0 \\
0 & 1 \end{pmatrix}
\]

\[
3 \times 3 I_3 = \begin{pmatrix} 1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \end{pmatrix}
\]

\[
n \times n I_n = \begin{pmatrix} 1 & 0 & \cdots \\
0 & 1 & \cdots \\
\vdots & \ddots & \ddots \end{pmatrix}
\]

**Proposition 5.4.** If $A$ is $m \times n$ and $B$ is $n \times p$ then

\[ AI_n = A \]

\[ I_nB = B \]

**Proposition 5.5.**

(i) If $A$ is $m \times n$ and $B$, $C$ are $n \times p$, then

\[ A(B + C) = AB + AC \]

(ii) If $D$, $E$ are $m \times n$ and $F$ is $n \times p$, then

\[ (D + E)F = DF + EF \]

**Proof.**

(i) Let $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. The $ij$-entry of $AB$ is

\[
(a_{j1} \ldots a_{jn}) \begin{pmatrix} b_{ij} \\
\vdots \\
b_{nj} \end{pmatrix} = a_{i1}b_{ij} + \cdots + a_{in}b_{nj}
\]
So $ij$-entry of $A(B + C)$ is

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

$$= a_{i1}b_{1j} + a_{i1}c_{1j} + \cdots + a_{in}b_{nj} + a_{in}c_{nj}$$

$$= (a_{i1}b_{1j} + \cdots + a_{in}b_{nj}) + (a_{i1}c_{1j} + \cdots + a_{in}c_{nj})$$

This is the $ij$ entry of $AB + AC$.

(ii) Left to enthusiastic reader. □
5.1. METHOD – GAUSSIAN ELIMINATION
Chapter 6

Some applications of matrix algebra

6.1 Linear equations

We know that a general system of linear equations
\[ a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\
\vdots \\
a_{m1}x_1 + \cdots + a_{mn}x_n = b_n \]
can be expressed as a matrix product
\[ Ax = b \]
where \( A = (a_{ij}) \), \( x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \) and \( b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \).

Proposition 6.1. The number of solutions of a system \( Ax = b \) is 0, 1 or \( \infty \).

Proof. Suppose the number of solutions is not 0 or 1, i.e. there are at least two solutions. We prove that this implies that there are infinitely many solutions. Let \( p \) and \( q \) be two different solutions (\( p,q \in \mathbb{R}^n \)), so
\[ Ap = b \\
Aq = b \]
and \( p \neq q \). Then
\[ A(p - q) = Ap - Aq = 0 \]
For any scalar \( \lambda \in \mathbb{R} \)
\[ A\lambda(p - q) = \lambda Ap - \lambda Aq = \lambda 0 \]
So
\[ A(p + \lambda(p - q)) = Ap + A\lambda(p - q) = Ap + 0 = b \]
So \( p + \lambda(p - q) \) is a solution of the original system.
Since \( p \neq q \), \( p - q \neq 0 \) and so each different scalar \( \lambda \) gives a different solution. So there are \( \infty \) solutions. \( \blacksquare \)
Structure of solutions

**Proposition 6.2.**

(i) System $Ax = 0$ either has one solution (which must be $x = \mathbf{0}$) or it has an infinite number of solutions.

(ii) Suppose $p$ is a solution of a system $Ax = b$ (i.e. $p \in \mathbb{R}^n$, $Ap = b$). Then all solutions of $Ax = b$ take the form

$$p + h$$

where $h$ is a solution of the system $Ax = 0$.

**Example 6.1.** System

$$
\begin{pmatrix}
1 & 0 & 3 \\
0 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
\begin{pmatrix} x_1 \\
x_2 \\
x_3 \end{pmatrix}
\end{pmatrix} =
\begin{pmatrix}
2 \\
1
\end{pmatrix}.
$$

The general solution is

$$x = (2 - 3a, 1 - a, a)$$

Particular solution $p = (2, 1, 0)$. So general solution is

$$(2 - 3a, 1 - a, a) = (2, 1, 0) + (-3a, -a, a) = p + h$$

Where $h = (-3a, -a, a)$ is the general solution of

$$
\begin{pmatrix}
1 & 0 & 3 \\
0 & 1 & 1
\end{pmatrix}
\begin{pmatrix} x_1 \\
x_2 \\
x_3 \end{pmatrix} =
\begin{pmatrix} 0 \\
0 \end{pmatrix}.
$$

**Proof.**

(i) clear from previous proposition

(ii) Let $q$ be a solution of $A$. So $Ap = b$, $Aq = b$. So $A(q - p) = 0$. Put $q - p = h$, a solution to $Ax = 0$. Then $q = p + h$.

6.2 Population Distribution

**Example 6.2.** Population classified in 3 income states

(1) Poor (P)

(2) Middle incom (M)

(3) Rich(R)

Over one generation (20 year period)

$P \rightarrow 20\%M, 5\%R \text{ rest stay } P$

$M \rightarrow 25\%P, 10\%R$

$R \rightarrow 5\%P, 30\%M$
Summarize this information in a matrix. It’s 3 × 3 matrix \( T, T = (t_{ij}), \) where

\[
t_{ij} = \text{proportion moving from state } j \rightarrow \text{state } i
\]

So

\[
T = \begin{pmatrix}
0.75 & 0.25 & 0.5 \\
0.20 & 0.65 & 0.30 \\
0.05 & 0.10 & 0.65
\end{pmatrix}
\]

This is called the \textit{transition} matrix. All entries are \( \geq 0, \) and column sums are 1.

Say we start with proportions \( p_1, p_2, p_3 \) in states 1, 2, 3.

\[
p^{(0)} = \begin{pmatrix}
p_1 \\
p_2 \\
p_3
\end{pmatrix}
\]

is the initial \textit{population vector}.

After 1 generation

- proportion in state 1 \( = t_{11}p_1 + t_{12}p_2 + t_{13}p_3 \)
- proportion in state 2 \( = t_{21}p_1 + t_{22}p_2 + t_{23}p_3 \)
- proportion in state 3 \( = t_{31}p_1 + t_{32}p_2 + t_{33}p_3 \)

So population vector after 1 generation is

\[
p^{(1)} = Tp^{(0)}
\]

Similarly, after 2 generations

\[
p^{(2)} = Tp^{(1)} = T^2p^{(0)}
\]

Continuing, see that after \( n \) generations, population vector is

\[
p^{(n)} = T^np^{(0)}
\]

This is an example of a \textit{Markov chain} – a population is divided into states 1, \ldots, \( n, \) and we’re given the proportion \( t_{ij} \) moving from state \( j \) to state \( i \) over a generation. The \textit{transition matrix} \( T = (t_{ij}) \) is \( n \times n \) with properties

1) \( t_{ij} \geq 0 \)

2) all column sums are 1

Markov chain is \textit{regular} if some power of \( T \) has no zero entries.

**Example 6.3.** The above example is regular.

\[
T = \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
\]

is not regular.
Basic Fact
In a regular Markov chain, as \( n \) grows, the vector \( T^n p^{(0)} \) gets closer and closer to a steady state vector \( s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \) where

1) \( s_1 + \cdots + s_n = 1 \)

2) \( Ts = s \)

This is true, whatever the initial population vector \( p^{(0)} \). Proof is not hard.

Example 6.4. On a desert island, a veg-prone community dies according to

(1) no-one eats meat 2 days in a row

(2) if someone doesn’t eat meat one day, they toss a coin: heads eat meat next day, tails don’t

What proportion can be expected to eat meat on a given day?

Answer – Markov chain.

- State 1: meat
- State 2: no meat
- “generation”: 1 day

Transition matrix

\[
T = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix}
\]

Notice \( T^2 = \begin{pmatrix} \frac{1}{4} & \frac{1}{3} \\ \frac{3}{4} & \frac{2}{3} \end{pmatrix} \). So it is regular. By our Basic Fact, we have a steady state vector \( s = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \), where \( s_1 + s_2 = 1 \) and \( Ts = s \).

So

\[
\frac{1}{2}s_2 = s_1 \\
\frac{1}{2}s_2 + s_1 = s_2 \\
s_1 + s_2 = 1
\]

So in long run, \( \frac{1}{3} \) of population will be eating meat on a given day.
6.3 Matrices and Geometry

6.3.1 Rotation

Consider a rotation about the origin through angle $\theta$. Then

\[
y_1 = r \cos(\theta + \alpha) \\
= r(\cos \theta \cos \alpha - \sin \theta \sin \alpha) \\
= x_1 \cos \theta - x_2 \sin \theta \\
y_2 = r \sin(\theta + \alpha) \\
= r(\sin \theta \cos \alpha + \cos \theta \sin \alpha) \\
= x_1 \sin \theta + x_2 \cos \theta
\]

So

\[
\begin{pmatrix}
y_1 \\
y_2
\end{pmatrix} =
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\]

Call $R_\theta = \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}$, a rotation matrix.

6.3.2 Reflection

Let $s$ be the reflection in the $x_1$ axis sending

\[
x = \begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} \rightarrow \begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\]

So matrix $S = \begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}$ represents the reflection $s$.

6.3.3 Combining Transformations

Two rotations, say through $\theta$ then $\gamma$

\[
x \rightarrow r_\gamma(r_\theta(x)) \\
x \rightarrow R_\gamma(R_\theta(x)) \\
= (R_\gamma R_\theta)x = \begin{pmatrix}
\cos \gamma & -\sin \gamma \\
\sin \gamma & \cos \gamma
\end{pmatrix}
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\]

Reflection+Rotation

Now consider $sr_\theta$, sending $x \rightarrow s(r_\theta(x))$. This sends

\[
x \rightarrow S(R_\theta x) \\
= (SR_\theta)x \\
= \begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & -\cos \theta
\end{pmatrix}
\begin{pmatrix}
x
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\cos & -\sin \\
-\sin & -\cos
\end{pmatrix}
\begin{pmatrix}
x
\end{pmatrix}
\]
Chapter 7

Inverses

**Definition 7.1.** Let $A$ be a square matrix $(n \times n)$. Say another $n \times n$ matrix $B$ is an inverse of $A$ iff $AB = BA = I_n$.
Inverse of $A$ is denoted $A^{-1}$. If $A$ has an inverse, $A$ is invertible.

**Example 7.1.** $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then

\[
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]

\[
\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]

**Example 7.2.**

\[
\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + 2c & b + 2d \\ 0 & 0 \end{pmatrix}
\]

which cannot equal to $I$.

**Proposition 7.1.** If $A$ is invertible then its inverse is unique.

**Proof.** Suppose $B$, $C$ are inverses of $A$. Then

\[
AB = BA = I
\]

\[
AC = CA = I
\]

\[
B = BI = B(AC) = (BA)C = IC = C
\]

\[\blacksquare\]

7.1 Relation to linear equations

**Proposition 7.2.** Suppose $A$ is invertible. Then any system $Ax = b$ has a unique solution

\[
x = A^{-1}b
\]
44 7.2. FINDING INVERSES

Proof.

\[ Ax = b \]
\[ \iff \]
\[ A^{-1}(Ax) = A^{-1}b \]
\[ (A^{-1}A)x = A^{-1}b \]
\[ x = A^{-1}b \]

\[ \Box \]

Example 7.3. System

\[ x_1 + x_2 = 2 \]
\[ x_2 = 3 \]

is

\[ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \]

By previous proposition

\[ x = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \]

7.2 Finding inverses

2 \times 2 matrices

Let \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \). Observe

\[ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = (ad - bc)I \]

Using this, we can prove the following

Proposition 7.3.

1) If \( ad - bc \neq 0 \) then \( A \) is invertible and

\[ A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \]

2) if \( ad - bc = 0 \) then \( A \) is not invertible.

Proof.

2) Suppose \( ad - bc = 0 \). Then \( AB = 0 \) (\( B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \)). Assume \( A \) is invertible.

\[ A^{-1}(AB) = A^{-1}0 = 0 \]
\[ A^{-1}(AB) = (A^{-1}A)B = IB \]

Therefore \( IB = 0 \) and thus \( A = B = 0 \). But zero matrix is not invertible.

\[ \Box \]
7.3 Finding inverses in general

Let $A = (a_{ij})$ is $n \times n$ matrix. We want $n \times n$ matrix $X = x_{ij}$ such that $AX = I$. To solve, we write one giant matrix

\[
\begin{pmatrix}
a_{11} & \ldots & a_{1n} & 1 & \ldots & 0 \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
a_{n1} & \ldots & a_{nn} & 0 & \ldots & 1
\end{pmatrix} = (A|I)
\]

Now use Gaussian elimination to reduce to Gaussian form. There are two possibilities

1) There is a row with zeros in the "left side" and non zero elements in the "right side". There is no inverse in this case.

2) We don’t get into the first situation. Therefore we can put the left matrix into Echelon form. Then, we can get rid of non-zero elements above the main diagonal of the left matrix and get a matrix

\[(I|E)\]

So $AE = I$.
Start with $(E|I)$ and reverse everything, we end up with $(I|A)$. So solving $EX = I$ gives $X = A$, completing proof that $E = A^{-1}$.

**Proposition 7.4.**

1) If we can reduce $(A|I)$ to $(I|E)$ using elementary row operations, then $E = A^{-1}$.

2) If we can reduce $(A|I)$ to matrix with zeros in the left side of the last line and a non-zero element in the right side of the last line, then $A$ is not invertible.

**Example 7.4.** Find inverse of

\[
\begin{pmatrix}
1 & 3 & -2 \\
2 & 5 & -3 \\
-3 & 2 & -4
\end{pmatrix}
\]
Augmented matrix is
\[
\begin{pmatrix}
1 & 3 & -2 & 1 & 0 & 0 \\
2 & 5 & -3 & 0 & 1 & 0 \\
-3 & 2 & -4 & 0 & 0 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 3 & -2 & 1 & 0 & 0 \\
0 & -1 & 1 & -2 & 1 & 0 \\
0 & 11 & -10 & 3 & 0 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 3 & -2 & 1 & 0 & 0 \\
0 & -1 & 1 & -2 & 1 & 0 \\
0 & 0 & 1 & -19 & 11 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 3 & -2 & 1 & 0 & 0 \\
0 & 1 & -1 & 2 & -11 & 0 \\
0 & 0 & 1 & -19 & 11 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 3 & 0 & -37 & 22 & 2 \\
0 & 1 & 0 & -17 & 10 & 1 \\
0 & 0 & 1 & -19 & 11 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 0 & 0 & 14 & -8 & -1 \\
0 & 1 & 0 & -17 & 10 & 1 \\
0 & 0 & 1 & -19 & 11 & 1 \\
\end{pmatrix}
\]

Thus the inverse
\[
\begin{pmatrix}
14 & -8 & -1 \\
-17 & 10 & 1 \\
-19 & 11 & 1 \\
\end{pmatrix}
\]

A result linking inverses, linear equations and echelon forms

**Proposition 7.5.** Let \( A \) be a square matrix \( n \times n \). The following four statements are equivalent

1. \( A \) is invertible
2. Any system \( Ax = b \) has a unique solution
3. The system \( Ax = 0 \) has the unique solution \( x = 0 \)
4. \( A \) can be reduced to the identity \( I_n \) using ERO.

**Proof.** We prove (1) \( \rightarrow \) (2), (2) \( \rightarrow \) (3), (3) \( \rightarrow \) (4), (4) \( \rightarrow \) (1).

(1) \( \rightarrow \) (2) Is done, \( x = A^{-1}b \)

(2) \( \rightarrow \) (3) Is obvious

(3) \( \rightarrow \) (4) Needs to be proved.

(4) \( \rightarrow \) (1) Is proved earlier.

\[\blacksquare\]
Chapter 8

Determinants

Recall that matrix \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) is invertible iff \( ad - bc \neq 0 \).

**Definition 8.1.** The determinant of a 2 × 2 matrix is
\[
\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc
\]
Also write as
\[
\begin{vmatrix} a & b \\ c & d \end{vmatrix}
\]
or \( |A| \).

**Example 8.1.**
\[
\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2
\]
\[
|R_\theta| = \begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix} = 1
\]
Recall

**Proposition 8.1.** For a 2 × 2 matrix \( A \)
- \( |A| \neq 0 \) if \( A \) is invertible if \( Ax = b \) has unique solution.

Aim is to define \( \det(A) \) for 3 × 3 and larger matrices in such a way that this result is still true.

**Definition 8.2.** For a 3 × 3 matrix
\[
A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}
\]
define the determinant of \( A \) to be
\[
\det(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}
\]
\[
= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})
\]
Example 8.2.

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
1 & 1 & 2
\end{vmatrix} = \begin{vmatrix}
5 & 6 \\
1 & 2
\end{vmatrix} - 2 \begin{vmatrix}
4 & 6 \\
1 & 2
\end{vmatrix} + 3 \begin{vmatrix}
4 & 5 \\
1 & 1
\end{vmatrix} = -3
\]

Example 8.3.

\[
\begin{vmatrix}
a & x & y \\
0 & b & z \\
0 & 0 & c
\end{vmatrix} = abc
\]

8.1 Properties

Definition 8.3. If \( A \) is a 3 \( \times \) 3 matrix, the \( ij \)-minor of \( A \) is the 2 \( \times \) 2 matrix \( A_{ij} \) obtained by deleting \( i \)th row and \( j \)th column of \( A \).

Then

\[
|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}|
\]

This is called the expansion of \( |A| \) by the first row.

Example 8.4.

\[
A = \begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
1 & 1 & 2
\end{pmatrix}, \quad A_{11} = \begin{pmatrix} 5 & 6 \end{pmatrix}
\]

We don’t have to expand by the first row.

Proposition 8.2. Expansion by the second row

\[
|A| = -a_{21}|A_{21}| + a_{22}|A_{22}| - a_{23}|A_{23}|
\]

third row

\[
|A| = a_{31}|A_{31}| - a_{32}|A_{32}| + a_{33}|A_{33}|
\]

Proof. The second row.

\[
RHS = -a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{23}(a_{11}a_{32} - a_{12}a_{31}) = |A|
\]

Check third row.

Proposition 8.3. If \( A \) (3 \( \times \) 3) has two equal rows, then \( |A| = 0 \).

Proof. WLOG\(^1\) say that

\[
A = \begin{pmatrix} a_1 & a_2 & a_3 \\
a_1 & a_2 & a_3 \\
c_1 & c_2 & c_3 \end{pmatrix}
\]

Expand by the third row to get \( |A| = 0 \).

\(^1\)Without Loss Of Generality
8.2 Effects of row operations on determinant

**Proposition 8.4.** Let $A = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$ be $3 \times 3$.

1) Row operation $r_i \rightarrow r_i + \lambda r_j$ ($i \neq j$) does not change $|A|$.

2) Swapping two rows changes $|A|$ to $-|A|$.

3) Row operation $r_i \rightarrow \lambda r_i$ changes $|A|$ to $\lambda |A|$.

**Proof.**

1) Say $i = 2$, $j = 1$, so the row op. sends

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \rightarrow A' = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 + \lambda a_1 & b_2 + \lambda a_2 & b_3 + \lambda a_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Expand by the second row

$$|A'| = -(b_1 + \lambda a_1) \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + (b_2 + \lambda a_2) \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - (b_3 + \lambda a_3) \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}$$

$$= |A| + \lambda \begin{vmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= |A|$$

2) Say we swap rows 1 and 2 to get

$$\begin{vmatrix} b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} + \ldots$$

3) It is obvious when we do the expansion by the row we multiply.

**Example 8.5.**

$$\begin{vmatrix} 1 & 2 & -1 \\ -2 & 3 & 4 \\ 5 & 9 & -4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 7 & 2 \\ 0 & -1 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 7 & 2 \\ -1 & 1 \end{vmatrix} + 0 + 0$$

$$= 9$$
Example 8.6.
\[
\begin{vmatrix}
1 & x & x^2 \\
1 & y & y^2 \\
1 & z & z^2 \\
\end{vmatrix} = \begin{vmatrix}
1 & x & x^2 \\
0 & y - x & y^2 - x^2 \\
0 & z - x & z^2 - x^2 \\
\end{vmatrix}
\]
\[
= (y - x)(z - x) \begin{vmatrix}
1 & x & x^2 \\
0 & 1 & y + x \\
0 & 1 & z + x \\
\end{vmatrix}
\]
\[
= (y - x)(z - x) \begin{vmatrix}
1 & x & x^2 \\
0 & 1 & y + x \\
0 & 0 & z - y \\
\end{vmatrix}
\]
\[
= (y - x)(z - x)(z - y)
\]
This is the $3 \times 3$ Vandermonde determinant.

Proposition 8.5. Let $A$ be $3 \times 3$ matrix, and let $A'$ be obtained from $A$ by el. row. ops. Then

\[
|A| = 0 \iff |A'| = 0
\]

Proof. Doing a row op. changes $|A|$ to $|A|$, $-|A|$ or $\lambda |A|$ (where $\lambda \neq 0$).

Main result

Theorem 8.6. Let $A$ be $3 \times 3$ matrix. Then

\[
|A| \neq 0 \iff A \text{ is invertible}
\]

(or the system $Ax = 0$ has unique solution, or $A \rightarrow I_3$ by row ops.)

Proof.

$\Rightarrow$ Suppose $|A| \neq 0$. Reduce $A$ to echelon form $A'$ by row operations. By 8.5,

$|A'| \neq 0$. If $A'$ has a zero row, $|A| = 0$. Hence $A' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ so can be reduced to $I$. So $A$ is invertible.

$\Leftarrow$ Suppose $A$ is invertible. $A$ can be reduced to $I$ by row ops. Since $|I| = 1$, so $|A| \neq 0$.

Example 8.7. For which values of $a$ is $A$ invertible?

\[
A = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 3 & 7 \\ 1 & 4 & a \end{pmatrix}
\]

\[
|A| = \begin{vmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ 0 & 2 & a - 5 \end{vmatrix} = a - 9
\]

So by 8.6, $A$ is invertible iff $a \neq 9$. 
Important consequence

**Corollary 8.7.** Let $A$ be $3 \times 3$ matrix. Suppose the system $Ax = 0$ has a non-zero solution $x \neq \mathbf{0}$. Then $|A| = 0$.

**Example 8.8.** Let $A = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 3 & 7 \\ 1 & 4 & a \end{pmatrix}$. System $Ax = 0$ has augmented matrix

$\begin{pmatrix} 1 & 2 & 5 & 0 \\ 1 & 3 & 7 & 0 \\ 1 & 4 & a & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & a - 5 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & a - 9 & 0 \end{pmatrix}$

So if $a \neq 9$, only solution is $x = 0$. If $a = 9$, system has $\infty$ solutions. We saw that $|A| = 0 \leftrightarrow a = 9$.

### 8.3 $n \times n$ Determinants

It is possible to define $\det(A)$ for any $n \times n$ matrix $A$, and to prove that all the results of this section are true for $n \times n$ matrices. Wait for proofs until next year.

**Definition 8.4.** $4 \times 4$ determinant. If $A = (a_{ij})$ is $4 \times 4$,

$$|A| = a_{11} |A_{11}| - a_{12} |A_{12}| + a_{13} |A_{13}| - a_{14} |A_{14}|$$

Similarly, define $n \times n$ determinant in terms of $(n-1) \times (n-1)$ determinants (recursive definition).

**Example 8.9.**

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 0 & 1 & 5 \\ 2 & 1 & 0 & 2 \\ 0 & -1 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 5 \\ 1 & 0 & 2 \\ -1 & 1 & 3 \end{vmatrix} - 2 \begin{vmatrix} -1 & 1 & 5 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{vmatrix} + 3 + \ldots$$
8.3. $N \times N$ DETERMINANTS
Chapter 9

Eigenvalues and eigenvectors

Fibonacci’s Rabbits
Rabbits, born in pairs. Newly born pair produce no offspring in the first month, but then one pair in each following month. No deaths. Start with one pair. How many pairs after n months?
Answer. Let $F_n =$ number of pairs. Then

\[ F_n = F_{n-1} + F_{n-2} \]

So $F_0, F_1, F_2, \ldots = 0, 1, 1, 2, 3, 5, 8, 13, \ldots$ is the Fibonacci sequence. In matrix form

\[ \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix} \]

Now \( \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \), so \( \begin{pmatrix} F_2 \\ F_1 \end{pmatrix} = A \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} \) and so on.

\[ \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^n \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} \]

9.1 Eigenvectors

Definition 9.1. Let $A$ be a $n \times n$ matrix. Vector $\nu = \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix} \in \mathbb{R}$ is an eigenvector of $A$ if

1. $\nu \neq 0$
2. $A\nu = \lambda \nu$ ($\lambda \in \mathbb{R}$)

The scalar $\lambda$ is an eigenvalue of $A$.

Example 9.1.

\[ \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = -1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} \]

So \( \begin{pmatrix} 1 \\ -2 \end{pmatrix} \) is an eigenvector of $A = \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}$ and $-1$ is its eigenvalue.
Example 9.2.

\[
\begin{pmatrix}
3 & 2 \\
2 & 0
\end{pmatrix}
\begin{pmatrix}
1 \\
1
\end{pmatrix} = \begin{pmatrix}
5 \\
2
\end{pmatrix} \neq \lambda \begin{pmatrix}
1 \\
1
\end{pmatrix}
\]

So \(\begin{pmatrix}
1 \\
1
\end{pmatrix}\) is not an eigenvector of \(\begin{pmatrix}
3 & 2 \\
2 & 0
\end{pmatrix}\).

### 9.2 How to find eigenvectors and eigenvalues

Let \(A\) be a \(n \times n\) matrix. Vector \(x\) is a non-zero solution of the system

\[Ax = \lambda x\]

Example 9.3.

\[
\begin{pmatrix}
3 & 2 \\
2 & 0
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} = \lambda \begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\]

We can get

\[
(3 - \lambda)x_1 + 2x_2 = 0 \\
2x_1 - \lambda x_2 = 0
\]

So the equation is

\[
\begin{vmatrix}
3 - \lambda & 2 \\
2 & -\lambda
\end{vmatrix} = 0
\]

\[
\lambda^2 - 3\lambda - 4 = 0
\]

Eigenvalues are \(-1\) and \(4\).

For \(\lambda = -1\), eigenvectors are non-zero solution of

\[
A + Ix = 0 \\
\begin{pmatrix}
4 & 2 \\
2 & 1
\end{pmatrix} x = 0
\]

So eigenvectors are \(\begin{pmatrix}
a \\
-2a
\end{pmatrix}\) \((a \in \mathbb{R}, a \neq 0)\).

For \(\lambda = 4\)

\[
\begin{pmatrix}
-1 & 2 \\
2 & -4
\end{pmatrix} x = 0
\]

So eigenvectors are \(\begin{pmatrix}
2b \\
b
\end{pmatrix}\) \((b \in \mathbb{R}, b \neq 0)\).

When does have the following equation a non-zero solution?

\[
Ax - \lambda x = 0 \\
(A - \lambda I)x = 0
\]

Precisely when \(|A - \lambda I| = 0\) (by 8.6).
**Proposition 9.1.**

(1) If $A$ is a $3 \times 3$ or $2 \times 2$, then the eigenvalues of $A$ are the solutions $\lambda$ of $|A - \lambda I| = 0$.

(2) If $\lambda$ is an eigenvalue, then the corresponding eigenvectors are the non-zero solutions of

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

**Definition 9.2.** The equation $|A - \lambda I| = 0$ is the characteristic equation of $A$, and $|A - \lambda I|$ is the characteristic polynomial of $A$.

### 9.2.1 Back to Fibonacci

We had $\left( \begin{array}{l} F_{n+1} \\ F_n \end{array} \right) = A^n \left( \begin{array}{l} 1 \\ 0 \end{array} \right)$ where $A = \left( \begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right)$.

**Strategy**

(1) Find eigenvalues $\lambda_1, \lambda_2$ and eigenvectors $v_1$ and $v_2$ of $A$. Observe

$$Av_1 = \lambda_1 v_1$$
$$Av_2 = \lambda_2 v_2$$

Then

$$A^2 v_1 = A(Av_1)$$
$$= A(\lambda_1 v_1)$$
$$= \lambda_1 Av_1$$
$$= \lambda_1^2 v_1$$

Similarly

$$A^n v_1 = \lambda_1^n v_1$$
$$A^n v_2 = \lambda_2^n v_2$$

(2) Express $\left( \begin{array}{l} 1 \\ 0 \end{array} \right)$ as a combination of $v_1$ and $v_2$.

$$\left( \begin{array}{l} 1 \\ 0 \end{array} \right) = \alpha v_1 + \beta v_2 \quad (\alpha, \beta \in \mathbb{R})$$

Then

$$A^n \left( \begin{array}{l} 1 \\ 0 \end{array} \right) = A^n(\alpha v_1 + \beta v_2)$$
$$= \alpha A^n v_1 + \beta A^n v_2$$
$$\left( \begin{array}{l} 1 \\ 0 \end{array} \right) = \alpha \lambda_1^n v_1 + \beta \lambda_2^n v_2 \quad (9.1)$$
Calculations

(1) Characteristic equation is \(|A - \lambda I| = 0\), i.e.

\[
\begin{vmatrix}
1 - \lambda & 1 \\
1 & -\lambda
\end{vmatrix} = 0
\]

i.e.

\[
\lambda^2 - \lambda - 1 = 0
\]

\[
\lambda_1 = \frac{1}{2}(1 + \sqrt{5})
\]

\[
\lambda_2 = \frac{1}{2}(1 - \sqrt{5})
\]

Eigenvectors for \(\lambda_1\)

\[
\begin{pmatrix}
1 - \lambda \\
1
\end{pmatrix} x = 0
\]

So \(v_1 = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}\) is an eigenvector.

Similarly \(v_2 = \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}\).

(2) We now find \(\alpha\) and \(\beta\) such that

\[
\begin{pmatrix}
1 \\
0
\end{pmatrix} = \alpha \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}
\]

\[
\alpha = \frac{1}{\lambda_1 - \lambda_2}
\]

\[
\beta = \frac{1}{\lambda_2 - \lambda_1}
\]

Putting all this into (9.1)

\[
\begin{pmatrix}
F_{n+1} \\
F_n
\end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\]

\[
= \alpha \lambda_1^n v_1 + \beta \lambda_2^n v_2
\]

\[
= \frac{1}{\sqrt{5}} \left( \frac{1}{2} \left( 1 - \sqrt{5} \right) \right)^n \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{5}} \left( \frac{1}{2} \left( 1 - \sqrt{5} \right) \right)^n \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}
\]

To get formula

\[
F_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right)
\]

9.3 Diagonalization

Want to investigate functions of matrices, e.g. \(A^n\), \(A^{\frac{1}{n}}\), \(f(A)\).
Definition 9.3. An $n \times n$ matrix $D$ is diagonal matrix if

$$D = \begin{pmatrix} \alpha_1 & 0 \\ \vdots & \ddots \\ 0 & \alpha_n \end{pmatrix}$$

Example 9.4.

$$\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$$

It’s easy to find powers of diagonal matrices.

Proposition 9.2. Let

$$D = \begin{pmatrix} \alpha_1 & 0 \\ \vdots & \ddots \\ 0 & \alpha_n \end{pmatrix}$$

$$E = \begin{pmatrix} \beta_1 & 0 \\ \vdots & \ddots \\ 0 & \beta_n \end{pmatrix}$$

Then

$$DE = \begin{pmatrix} \alpha_1\beta_1 & 0 \\ \vdots & \ddots \\ 0 & \alpha_n\beta_n \end{pmatrix}$$

and

$$D^k = \begin{pmatrix} \alpha_1^k & 0 \\ \vdots & \ddots \\ 0 & \alpha_n^k \end{pmatrix}$$

Proof. $DE$ given by definition of matrix multiplication. Take $E = D$ to get $D^2$ and repeat to get $D^k$.

Aim - To relate an arbitrary square matrix $A$ to a diagonal matrix, and exploit this to find $A^n$, etc.

2 × 2

Let $A$ be $2 \times 2$ and suppose $A$ has eigenvalues $\alpha_1, \alpha_2$ with eigenvectors $v_1, v_2$. Assume $\alpha_1 \neq \alpha_2$. We get

$$Av_1 = \alpha_1 v_1$$
$$Av_2 = \alpha_2 v_2$$

Cleverly define $2 \times 2$ matrix $P$ ($v_1, v_2$ are column vectors)

$$P = (v_1 \ v_2)$$
Then
\[ AP = A(v_1 v_2) \]
\[ = (Av_1 \quad Av_2) \]
\[ = (\alpha_1 v_1 \quad \alpha_2 v_2) \]
\[ = (v_1 \quad v_2) \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \]

So if we write \( D = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \), we’ve shown
\[ AP = PD \quad (9.2) \]

We claim that \( P \) is invertible. For if not, then \( |P| = 0 \), which means that \( v_1 = \lambda v_2 \), which is false as
\[ Av_1 = \alpha_1 v_1 = \alpha_1 \lambda v_2 \]
\[ A\lambda v_2 = \lambda \alpha_2 v_2 \]
and these are not equal as \( \alpha_1 \neq \alpha_2 \).
So from (9.2)
\[ P^{-1}AP = P^{-1}PD = D \]

Summary

**Proposition 9.3.** Let \( A \) be \( 2 \times 2 \) with distinct eigenvalues \( \alpha_1, \alpha_2 \), eigenvectors \( v_1, v_2 \).

Let
\[ P = (v_1 \quad v_2) \]

Then \( P \) is invertible and
\[ P^{-1}AP = D = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \]

**Note.** Also true for \( 3 \times 3, \ldots \) matrices.

**Example 9.5.** Let \( A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \).

1) Find \( P \) such that \( P^{-1}AP \) is diagonal.

2) Find the formula \( A^n \).

3) Find a 5th root of \( A \), i.e. find \( B \) such that
\[ B^5 = A \]

**Answer.**
1) Characteristic equation of $A$ is 

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ -2 & 3 - \lambda \end{vmatrix} = 0$$

i.e.

$$\lambda^2 - 3\lambda + 2 = 0$$

$$(\lambda - 1)(\lambda - 2) = 0$$

So eigenvectors are

$\lambda = 1 \ a \begin{pmatrix} 1 \\ 1 \end{pmatrix}, (a \neq 0).$

$\lambda = 2 \ b \begin{pmatrix} 1 \\ 2 \end{pmatrix}, (b \neq 0)$

Let $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Then by Proposition 9.3,

$$P^{-1}AP = D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Note that many other $P$’s work, e.g. $\begin{pmatrix} 2 & -1 \\ 2 & -2 \end{pmatrix}$. Or if $Q = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, then

$$Q^{-1}AQ = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

2) Find $A^n$. Know $P^{-1}AP = D$. By the 9.2, we know that

$$D^n = \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix}$$

So

$$\begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} = D^n = (P^{-1}AP)^n$$

$$= \underbrace{P^{-1}APP^{-1}AP \cdots P^{-1}APP^{-1}AP}_{n}$$

$$= P^{-1}A^nP$$

So

$$P^{-1}A^nP = D^n$$

$$PP^{-1}A^nPP^{-1} = PD^nP^{-1}$$

$$A^n = PD^nP^{-1}$$

So

$$A^n = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2^n \\ 1 & 2^{n+1} \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$
Find $B$ such that $B^5 = A$.

Well, if

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 2^{1/5} \end{pmatrix}$$

Then $C^5 = D$. So

$$(PCP^{-1})^5 = PCP^{-1} \cdots PCP^{-1} = PC^5P^{-1} = PDP^{-1} = A$$

So take

$$B = PCP^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^{1/5} \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 - 2^{1/5} & 2^{1/5} - 1 \\ 2 - 2^{2/5} & 2^{2/5} - 1 \end{pmatrix}$$

**Note.** Usually a matrix has many square roots, fifth roots, etc. Eg, $I$ has infinitely many.

**Note.** Similarly can calculate polynomial functions

$$p(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I$$

**Summary**

If a square matrix $A$ has distinct eigenvalues, then it can be diagonalized, i.e. there exists an invertible $P$ such that $P^{-1}AP$ is diagonal.

**9.3.1 Repeated eigenvalues**

If the characteristic polynomial of $A$ has a repeated root $\lambda$, we call $\lambda$ a repeated eigenvalue of $A$.

Some $A$'s with the repeated eigenvalue can be diagonalised and some can’t.

**Example 9.6.** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then the characteristic polynomial is

$$\begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2$$

so 1 is a repeated eigenvalue. **Claim** $A$ cannot be diagonalized (i.e. no invertible $P$ exists such that $P^{-1}AP$ is diagonal).

**Proof.** Assume there exists an invertible $P$ such that

$$P^{-1}AP = D = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$
Then

$$AP = PD$$

Writing

$$P = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$$

$$A\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}D = \begin{pmatrix} \alpha_1 v_1 \\ \alpha_2 v_2 \end{pmatrix}$$

Hence $Av_1 = \alpha_1 v_1$, $Av_2 = \alpha_2 v_2$.
So $v_1$, $v_2$ are eigenvectors of $A$.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

eigenvectors are $a\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $a \neq 0$. Hence

$$P = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

But this is not invertible. Contradiction.

Point $P$ must have evectors of $A$ as columns, but $A$ does not have enough “independent” evectors to make invertible $P$.

Example 9.7. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}$. Then the characteristic polynomial is

$$\begin{vmatrix} 1 - \lambda & 0 & 0 \\ -1 & 2 - \lambda & 0 \\ 1 & -1 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 (2 - \lambda)$$

So 1 is a repeated evalue.
Can $A$ be diagonalized?
Let’s try.
Evecors for $\lambda = 2$

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 0 \end{pmatrix}$$

eigenvectors $a\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $a \neq 0$.

For $\lambda = 1$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$
62 9.3. DIAGONALIZATION

eigenvectors \( \begin{bmatrix} b \\ b \\ c \end{bmatrix} \).

So for columns of \( P \) choose eigenvectors.

\[
\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\]

So take

\[
P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}
\]

Then \(|P| = -1\), so \( P \) is invertible, and from 9.2

\[
P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

Summary If \( A \) has enough independent vectors for the repeated eigenvalue, \( A \) can be diagonalized. If not, it can’t.
Chapter 10

Conics(again) and Quadrics

Recall equation of conic
\[ ax_1^2 + bx_1 x_2 + cx_2^2 + dx_1 + ex_2 + f = 0 \]
We’ll write this in matrix form
\[
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
\begin{pmatrix}
  a & b/2 & c/2 \\
  b/2 & b/2 & b/2 \times & x_2 & c/2
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= a x_1^2 + bx_1 x_2 + cx_2^2
\]
So in matrix form, the equation of conic is
\[ x^T A x + (d \quad e) x + f = 0 \]
where \( x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \), \( x^T = (x_1 \quad x_2) \), \( A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \)

Digression – Transposes

Definition 10.1. If \( A = (a_{ij}) \) is \( m \times n \), the transpose of \( A \) is the \( m \times n \) matrix \( A^T = (a_{ji}) \).

Example 10.1. \( A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \), \( A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \)
\[ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ x^T = (x_1 \quad x_2), \ A = \begin{pmatrix} a & b \\ b/2 & c \end{pmatrix} \]

Note.
\[ (A^T)^T = A \]
\[ x.y = (x_1 \quad x_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x^T y \]

Definition 10.2. A square matrix \( A \) is symmetric if \( A = A^T \).

Example 10.2. \( \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \) is symmetric.

Proposition 10.1. Let \( A \) be \( m \times n \), \( B \) be \( n \times p \). Then
\[ (AB)^T = B^T A^T \]
Proof. The \( ij \)-th entry of \((AB)^T\) is the \( ji \)-th entry of \(AB\), which is the \( j \)-th row of \(A\) multiplied by \( i \)-th column of \(B\). And \( ij \)-th entry of \(B^TA^T\) is \( i \)-th row of \(B^T\) multiplied by \( j \)-th column of \(A^T\), i.e. \( i \)-th column of \(B\) multiplied by \( j \)-th row of \(A\). These are equal.

\[ \square \]

**Back to conics**

We formulatied the equation of a conic as

\[ x^T Ax + (d \ e) x + f = 0 \quad (10.1) \]

where \( x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \), \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \).

Notice

\[ x^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \lambda_1 x_1^2 + \lambda_2 x_2^2 \]

So, aim is to find a rotation matrix \( P \), such that the change of coordinates to \( y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \), where \( x = Py \) changes the equation (10.1) to

\[ (Py)^T A(Py) + (d \ e) Py + f = 0 \]

\[ y^T P^T A P y + (d' \ e') y + f = 0 \]

and \( P^T A P \) is diagonal.
We can do this.

**Theorem 10.2.** Let \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \) be symetrix, with \( b \neq 0 \). Then

1) \( A \) has two real eigenvalues \( \lambda_1, \lambda_2 \), with \( \lambda_1 \neq \lambda_2 \).

2) If \( v_1, v_2 \) are eigenvectors of \( A \) corresponding to \( \lambda_1 \) and \( \lambda_2 \), then

\[ v_1 \cdot v_2 = 0 \]

3) We can choose unit eigenvectors \( v_1, v_2 \), such that

\[ P = \begin{pmatrix} v_1 & v_2 \end{pmatrix} \]

has determinant 1 and \( P \) is a rotation matrix. Moreover, \( P^{-1} = P^T \), and \( P^T A P \) is diagonal.

**Proof.**

1) Characteristic polynomial of \( A \) is

\[ \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = \lambda^2 - (a + c)\lambda + ac - \frac{b^2}{4} \]

Roots are

\[ \frac{1}{2} \left( a + c \pm \sqrt{(a + c)^2 - 4(ac + b^2)} \right) = \frac{1}{2} \left( a + c \pm \sqrt{(a - c)^2 + b^2} \right) \]
Since $b \neq 0$, $(a - c)^2 + b^2 > 0$, so the roots are real and distinct. Call them $\lambda_1$ and $\lambda_2$.

2) Let

$$Av_1 = \lambda_1 v_1$$
$$Av_2 = \lambda_2 v_2$$

Consider

$$v_1^T A v_2$$

(a) This is $v_1^T (\lambda_2 v_2) = \lambda_2 v_1^T v_2$.

(b) It is also (because $A$ is symmetric, i.e. $A = A^T$)

$$(A^T v_1)^T v_2 = (Av_1)^T v_2 = \lambda_1 v_1^T v_2$$

So

$$\lambda_2 v_1^T v_2 = \lambda_1 v_1^T v_2$$

As $\lambda_1 \neq \lambda_2$, this forces $v_1^T v_2 = 0$, i.e. $v_1 \cdot v_2 = 0$.

3) Choose unit $v_1$, $v_2$ (see picture). Then

$$v_1 = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$
$$v_2 = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$$

So

$$P = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$$
$$= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$
$$= R_\theta$$

Finally, by 9.3,

$$P^T A P = P^{-1} A P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

10.1 Reduction of conics

Start with equation (10.1).

1) Find the evalves and evectors of $A$

2) Find unit eigenvectors $v_1$, $v_2$, such that $P = (v_1 v_2)$ has determinant 1, so is a rotation matrix.
3) Change coordinates to \( y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \) when \( x = Py \), so equation (10.1) becomes

\[
y^T (P^T AP) y + (d' \ e') y + f = 0
\]
i.e.
\[
y^T \begin{pmatrix} y_1 & 0 \\ 0 & y_2 \end{pmatrix} y + (d' \ e') y + f = 0
\]
i.e.
\[
\lambda_1 y_1^2 + \lambda_2 y_2^2 + d'y_1 + e'y_2 + f = 0
\]

This is an ellipse if \( \lambda_1 \lambda_2 > 0 \), hyperbola if \( \lambda_1 \lambda_2 < 0 \), parabola if \( \lambda_1 \lambda_2 = 0 \) (or possible degenerate cases).

**Example 10.3.** Reduce conic

\[
5x_1^2 + 4x_1x_2 + 2x_2^2 = 1
\]
to standard form. This is

\[
x^T \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} x = 1
\]

1) Characteristic polynomial

\[
\begin{vmatrix} 5 - \lambda & 2 \\ 2 & 2 - \lambda \end{vmatrix} = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda + 6)
\]

For \( \lambda = 1 \), eigenvectors are \( a \begin{pmatrix} 1 \\ -2 \end{pmatrix} \).

For \( \lambda = 6 \), eigenvectors are \( b \begin{pmatrix} 2 \\ 1 \end{pmatrix} \).

2) Unit vectors are \( \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \), \( \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \).

Take

\[
P = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}
\]

Then \( |P| = 1 \), and \( P = R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \), where \( \cos \theta = \frac{1}{\sqrt{5}} \), \( \sin \theta = \frac{2}{\sqrt{5}} \).

3) Let \( x = Py \) (i.e. \( x_1 = \frac{1}{\sqrt{5}}(y_1 + 2y_2) \), \( x_2 = \frac{1}{\sqrt{5}}(-2y_1 + y_2) \)). Then equation becomes

\[
y_1^2 + 6y_2^2 = 1
\]

which is an ellipse.
10.2. QUADRIC SURFACES

10.1.1 Orthogonal matrices
Recall, rotation matrix satisfies

\[ R_\theta^T = R_\theta \]

So does a reflection matrix

\[
\begin{pmatrix}
\cos \theta & \sin \theta \\
\sin \theta & -\cos \theta
\end{pmatrix}
\]

**Definition 10.3.** A square matrix \( P \) is an **orthogonal** matrix if \( P^T = P^{-1} \), i.e. \( PP^T = I \).

The key property:

**Proposition 10.3.** Orthogonal matrices preserve lengths, i.e.

\[ \| Pv \| = \| v \| \]

**Proof:**

\[
\| Pv \|^2 = (Pv)(Pv) = (Pv)^T(Pv) = v^T P^T P v = v^T I v = \| v \|^2
\]

■

10.2 Quadric surfaces

**Definition 10.4.** A **quadric surface** is a surface in \( \mathbb{R}^3 \) defined by a quadratic equation

\[ ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3 + gx_1 + hx_2 + jx_3 + k = 0 \quad (10.2) \]

**Standard examples**

1) **Ellipsoid**

\[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1
\]

2) **Hyperboloid** of 1 sheet

\[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 1
\]

3) **Hyperboloid** of 2 sheets

\[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = -1
\]

4) **Elliptic cone**

\[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 0
\]
5) Elliptic paraboloid

\[ x_3 = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \]

6) Hyperbolic paraboloid

\[ x_3 = \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} \]

Some degenerate cases

7) Elliptic cylinder

\[ \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \]

8) Parabolic cylinder

\[ x_1^2 - ax_2 = 0 \]

Aim is to find a rotation and translation, which reduces (10.2) to one of the standard examples. Here is the procedure.

1. Write (10.2) in matrix form

\[ x^T A x + (g \ h \ j) x + k = 0 \]

where

\[ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \]

\[ A = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} \]

Notice that A is symmetric.

2. Find the eigenvalues \(\lambda_1, \lambda_2, \lambda_3\) of A (theory – these are real). Find corresponding unit eigenvectors \(v_1, v_2, v_3\) which are perpendicular to each other (theory – this can be done). Choose a directions for the \(v_i\) so that the matrix

\[ P = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \]

has determinant 1.

Then, \(P\) is a rotation matrix, and \(P^{-1} = P^T\). Then

\[ P^{-1} A P = P^T A P \]

\[ = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \]
3. Change coordinates to \( y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \), where \( x = Py \). Then (??) reduces to
\[
y^T (P^T AP)y + (g \ h \ j) Py + k = 0
\]
i. e.
\[
\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 + g' y_1 + h' y_2 + j' y_3 + k = 0
\]
Finally, complete the square to find translation reducing to a standard equation.

**Note.** All the assumed bits of theory will be covered in Algebra II next year.

**Example 10.4.** Reduce the quadric
\[
2x_1 x_2 + 2x_1 x_3 - x_2 - 1 = 0
\]
to standard form.

**Answer.**

1. Equation is
\[
x^T \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} x = x_2 + 1
\]

2. Characteristic polynomial is
\[
\begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - 2)
\]
We get \( \lambda \) equal to 0, \( \sqrt{2} \), \( -\sqrt{2} \).

**Eigenvectors.**

For \( \lambda = 0 \) a \( \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \), unit eigenvectors \( \pm \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \)

For \( \lambda = \sqrt{2} \) \( \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix} \), unit eigenvectors \( \pm \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \frac{\sqrt{2}}{2} \\ \frac{1}{2} \end{pmatrix} \)

For \( \lambda = -\sqrt{2} \) \( \begin{pmatrix} -\sqrt{2} \\ 1 \\ 1 \end{pmatrix} \), unit eigenvectors \( \pm \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ \frac{\sqrt{2}}{2} \\ \frac{1}{2} \end{pmatrix} \)

Let
\[
P = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{2} & \frac{1}{2} \end{pmatrix}
\]
Then \(|P| = 1\) and \(P\) is a rotation.
Change of coordinates \(x = Py\) changes equation to
\[
\sqrt{2}y_2^2 - \sqrt{2}y_3^2 = -\frac{1}{\sqrt{2}}y_1 + \frac{1}{2}y_2 + \frac{1}{2}y_3 + 1
\]
Finally, complete the square
\[
\sqrt{2}(y_2 - \alpha)^2 - \sqrt{2}(y_3 - \beta)^2 = -\frac{1}{\sqrt{2}}(y_1 - \gamma)
\]
This is a hyperbolic paraboloid.

10.3 Linear Geometry in 3 dimensions

Definition 10.5. For two vectors \(x = (x_1, x_2, x_3), y = (y_1, y_2, y_3)\) in \(\mathbb{R}^3\), define
- \(x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3)\)
- \(\lambda x = (\lambda x_1, \lambda x_2, \lambda x_3), \lambda \in \mathbb{R}\)
- length of \(x\), denoted \(\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}\)
- dot product \(x.y = x_1y_1 + x_2y_2 + x_3y_3\) (note that \(x.x = \|x\|^2\))

**Proposition 10.4.** For \(x, y \in \mathbb{R}^3\)
\[
x.y = \|x\|\|y\| \cos \theta
\]

**Proof.**
\[
\|y - x\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos \theta
\]
\[
2\|x\|\|y\| \cos \theta = x.x + y.y - (y - x)(y - x)
\]
\[
= 2x.y
\]

10.4 Geometry of planes

Fact – there is a unique plane through any 3 points in \(\mathbb{R}^3\) provided they aren’t colinear. We want to describe planes in terms of vectors and equations.
A plane \(\Pi\) in \(\mathbb{R}^3\) is specified by
1) a point \(A\) on the plane
2) a vector \(n\) normal to the plane
Then \(x \in \Pi \iff (x - a).n = n \iff x.n = a.n.\)

**Proposition 10.5.** If there are points \(A, B, C \in \mathbb{R}^3\) that are not collinear, there exists a unique plane \(\Pi\) through \(A, B, C\).
Proof. Because both $A - B$ and $B - C$ are both in the plane

\[ n.(a - b) = 0 \]
\[ n.(c - a) = 0 \]

i.e.

\[ r_1n_1 + r_2n_2 + r_3n_3 = 0 \]
\[ s_1n_1 + s_2n_2 + s_3n_3 = 0 \]

where $(r_1, r_2, r_3) = a - b$, $(s_1, s_2, s_3) = c - a$.

We want to show that all solutions are $\{\lambda n \mid \lambda \in \mathbb{R}\}$. Because $A, B, C$ are not colinear, $r$ and $s$ are not parallel.

How to find $\text{dist}(P, \Pi)$?
Find foot of perpendicular, $Q$ and arbitrary $A$ in the plane.

\[ \text{dist}(P, Q) = \|a - p\| \cos \theta \]

But

\[ (a - p).n = \|a - p\|\|n\| \cos \theta \]
\[ \text{dist}(P, Q) = \frac{(a - p).n}{\|n\|} \]