# MATH211: Linear Methods I

Matthew Burke

Thursday 29th November, 2018

# Lecture on Thursday 29th November, 2018

Diagonalisation

Similar matrices

Powers

Dynamical systems

Markov chains

# Last time

- Spectral theory
- Finding eigenvalues
- Finding eigenspaces of eigenvectors

# Diagonalisation

# Motivation

If $x \in \mathbb{R}^n$ is a linear combination of eigenvectors of $A$ then:

$$A(x) = A\left(\sum_{i=1}^{n} a_i v_{\lambda_i}\right) = \sum_{i=1}^{n} a_i A(v_{\lambda_i}) = \sum_{i=1}^{n} a_i \lambda_i v_{\lambda_i}$$

So if *every* vector in $\mathbb{R}^n$ can be written as a linear combination of eigenvectors of $A$ the the *entire* matrix action simplifies.

## Definition

An $n \times n$ matrix $A$ is *diagonalisable* iff every vector in $\mathbb{R}^n$ can be written as a linear combination of eigenvectors.

# Example

## Example

$$\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \text{ has eigenvectors } \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and so is diagonalisable.

## Example

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \text{ has eigenvectors } v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } v_4 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and so is diagonalisable because all vectors in $\mathbb{R}^2$ can be written as a linear combination of $v_2$ and $v_4$.

# Motivation

Suppose we want to find $Av$.

$$\begin{tikzcd}
v = \sum_{i=1}^n v_i e_i \arrow[d, mapsto, "P^{-1}"] & Av = \sum_{i=1}^n (Av)_i e_i \\
\sum_{i=1}^n a_i v_{\lambda_i} \arrow[r, mapsto, "\text{Apply } A"] & \sum_{i=1}^n a_i \lambda_i v_{\lambda_i} \arrow[u, mapsto, "P"']
\end{tikzcd}$$

- where $P$ writes a linear combination of eigenvectors as a linear combination of standard basis vectors
- we say that the bottom horizontal action is *diagonal*

# Diagonalisation

## Definition

The matrix $P$ *diagonalises* $A$ iff there is a diagonal matrix $D$ such that

$$A = PDP^{-1}$$

**Question:** How do we find the matrix $P$ for a matrix $A$?

**Answer:** If $v_{\lambda_1}, v_{\lambda_2} \ldots v_{\lambda_n}$ are eigenvectors such that every $x \in \mathbb{R}^n$ is a linear combination of the $v_{\lambda_i}$ then the matrix

$$P = [v_{\lambda_1} v_{\lambda_2} \ldots v_{\lambda_n}]$$

with eigenvectors as the columns is a diagonalising matrix for $A$.

# When is a matrix diagonalisable?

### Theorem

*The following are equivalent for a matrix $A$:*

1. *there is a $P$ such that $P^{-1}AP$ is diagonal*
2. *there are $n$ eigenvectors such that any $x \in \mathbb{R}^n$ is a linear combination of these eigenvectors*
3. *there are eigenvectors $v_1 \dots v_n$ such that $[v_1 v_2 \dots v_n]$ is invertible*
4. *for every eigenvalue $\lambda$ the geometric multiplicity is equal to the algebraic multiplicity*

### Lemma

*If the characteristic polynomial of $A$ has $n$ distinct eigenvalues then $A$ is diagonalisable by (4).*

# Examples

**Example**

If possible diagonalise

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

**Example**

If possible diagonalise

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

# Examples

## Example

If possible diagonalise

$$\begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$$

## Example

If possible diagonalise

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

# Summary

1. Find the eigenvalues by solving $|A - \lambda \cdot I|$.
   - The no. of $(\lambda - \lambda_i)$ factors is the *algebraic multiplicity* $alg(\lambda_i)$.

2. Find the eigenspaces $E_{\lambda_i}$ by solving $(A - \lambda_i \cdot I) = 0$.
   - Only need to do this for the eigenvalues $\lambda_i$ found in (1).
   - The no. of parameters is the *geometric multiplicity* $geom(\lambda_i)$.

3. - If for all $\lambda_i$ found in (1) we have $geom(\lambda_i) = alg(\lambda_i)$ then the diagonalising matrix is $P = [v_{\lambda_1} v_{\lambda_2} \dots v_{\lambda_n}]$ where the $v_{\lambda_i}$ are the basic eigenvectors.
   - If for any of the $\lambda_i$ found in (1) has $geom(\lambda_i) < alg(\lambda_i)$ then the matrix is not diagonalisable.

# Similar matrices

# Similar matrices

**Definition**

Two matrices $A$ and $B$ are *similar* iff there exists an invertible $P$ such that

$$A = PBP^{-1}$$

and we write $A \sim B$.

# Trace of a matrix

**Definition**

The trace of a matrix is the sum of its diagonal elements:

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

**Theorem**

*If $A \sim B$ then $tr(A) = tr(B)$.*

# Powers

# Taking powers of a diagonalisable matrix

Suppose that there exists a diagonal matrix $D$ such that

$$A = PDP^{-1}$$

for some invertible matrix $P$. Then

$$A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$

and indeed

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

# Examples

### Example

Find $A^9$ if

$$A = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

### Example

Find $A^{50}$ if

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

# Dynamical systems

# Markov chains

### Definition

A *dynamical system* consists of a function $\alpha(t)$ that prescribes how the state of the system changes over time.

### Definition

A *discrete linear dynamical system* consists of a sequence of vectors

$$x_0, x_1, x_2, \dots, x_k, \dots$$

such that $x_{k+1} = Ax_k$ for some matrix $A$.

# Long term behaviour using eigenvectors

If $x_0$ is a linear combination of the eigenvectors $v_{\lambda_i}$ of $A$ then

$$x_k = A^k x_0 = A^k \left( \sum_{i=1}^{n} b_i v_{\lambda_i} \right) = \sum_{i=1}^{n} b_i A^k \left( v_{\lambda_i} \right) = \sum_{i=1}^{n} b_i (\lambda_i)^k v_{\lambda_i}$$

and so the long-term behaviour is determined by the limits:

$$\lim_{k \to \infty} (\lambda_i)^k$$

# Dominant eigenvalue

### Definition

If a is a square matrix then a *dominant eigenvalue* $\lambda_{max}$ is one for which $|\lambda_{max}| > |\lambda_i|$ for all other eigenvalues $\lambda_i$.

$$x_k = \sum_{i=1}^{n} b_i(\lambda_i)^k v_{\lambda_i} \approx b_i(\lambda_{max})^k v_{\lambda_{max}}$$

and we can read off the long term behaviour. E.g.

- if $|\lambda_{max}| < 1$ then the system converges to 0
- if $|\lambda_{max}| = 1$ then the system converges to $b_i v_{\lambda_{max}}$
- if $|\lambda_{max}| = -1$ then the system oscillates between $\pm b_i v_{\lambda_{max}}$
- if $|\lambda_{max} > 1|$ then the system diverges

# Examples

### Example

Find a formula for $x_k$ if $x_{k+1} = Ax_k$,

$$x_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}$$

### Example

Estimate the long term behaviour of the dynamical system with

$$x_0 = \begin{bmatrix} 100 \\ 40 \end{bmatrix} \text{ and } A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{bmatrix}$$

# Markov chains

# Markov chains

A *Markov chain* consists of:-

- a finite set of *states* $x_1, x_2, \ldots, x_n$
- a repeated *transition interval* at the end of which the system transitions between states
- a *non-deterministic rule* for predicting the probability that the system will transition into a certain state
  - **This probability only depends on the current state.**
  - (Not the entire history of the chain.)

# Markov transition matrices

This means that a Markov chain is described by a matrix $A$ such that

$$\begin{aligned} A_{ij} &= \mathbb{P}(X_1 = j | X_0 = i) \\ &= \text{ the probability that the next state will be } j \\ &\quad\ \text{ if the current state is } i \end{aligned}$$

Therefore:-

- all of the entries are between 0 and 1
  - (I.e. they are probabilities.)
- in any column the sum of the entries is 1
  - (The system must transition into one of the states.)

# Example

## Example

Find the probability that $x_3$ is in state 1 if

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ and } A = \begin{bmatrix} 0.4 & 0.25 & 0.2 \\ 0.4 & 0.35 & 0.5 \\ 0.2 & 0.4 & 0.3 \end{bmatrix}$$