Imperial College London

Maths for Signals and Systems

Linear Algebra in Engineering

Lectures 10-12, Tuesday 1st and Friday 4th November2016

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING

IMPERIAL COLLEGE LONDON

Imperial College London

In this set of lectures we will talk about:

- Eigenvectors and eigenvalues
- Matrix diagonalization
- Applications of matrix diagonalization
- Stochastic matrices

Eigenvectors and eigenvalues

- Consider a matrix A and a vector x.
- The operation Ax produces a vector y at some direction.
- I am interested in vectors y which lie in the same direction as x.
- In that case I have $Ax = \lambda x$ with λ being a scalar.
- When the above relationship holds, x is called an **eigenvector** and λ is called an **eigenvalue** of matrix A.
- If A is singular then $\lambda = 0$ is an eigenvalue.
- **Problem:** How do we find the eigenvectors and eigenvalues of a matrix?

Imperial College London

Eigenvectors and eigenvalues of a projection matrix

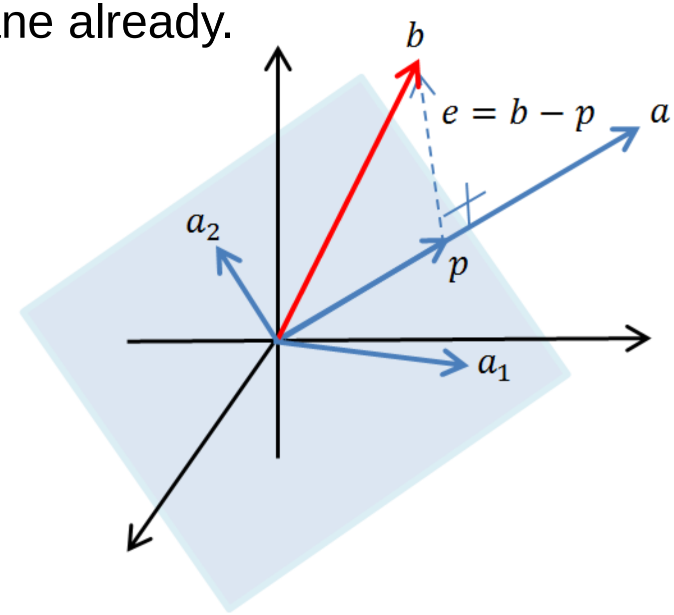
- **Problem:** What are the eigenvectors x' and eigenvalues λ' of a projection matrix P? In the figure, consider the matrix P which projects vector b onto vector p.

Question: Is b an eigenvector of P?

Answer: No, because b and Pb lie in different directions.

Question: What vectors are eigenvectors of P?

Answer: Vectors x which lie on the projection plane already.
In that case $Px = x$ and therefore
x is an eigenvector with eigenvalue 1.



Eigenvectors and eigenvalues of a projection matrix (cont.)

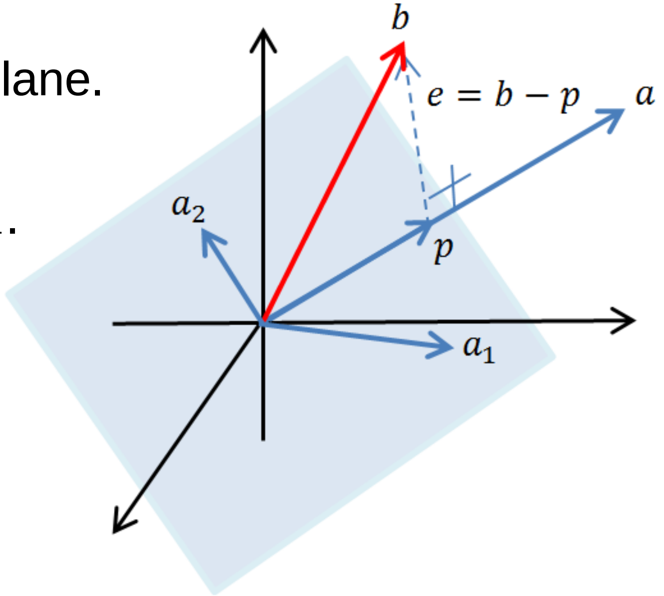
- The eigenvectors of P are vectors x which lie on the projection plane already. In that case $Px = x$ and therefore x is an eigenvector with eigenvalue 1.
- We can find 2 independent eigenvectors of P which lie on the projection plane both associated with an eigenvalue of 1.

Problem: In the 3D space we can find 3 independent vectors. Can you find a third eigenvector of P that is perpendicular to the eigenvectors of P that lie on the projection plane?

Answer: YES! Any vector e perpendicular to the plane.

In that case $Pe = \mathbf{0} = 0e$. Therefore,

the eigenvalues of P are $\lambda = 0$ and $\lambda = 1$.



Eigenvectors and eigenvalues of a permutation matrix

- Consider the permutation matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

  Problem: Can you give an eigenvector of the above matrix? Or can you think of a vector that if permuted is still a multiple of itself?

  Answer: YES. It is the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the corresponding eigenvalue is $\lambda = 1$.

  And furthermore, the vector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda = -1$.

- $n \times n$ matrices will have n eigenvalues.
- It is not easy to find them.
- **The sum of the eigenvalues, called the trace of a matrix, equals the sum of the diagonal elements of the matrix.**
- **The product of the eigenvalues equals the determinant of the matrix.**
- Therefore, in the previous example, once I found an eigenvalue $\lambda = 1$, I should know that there is another eigenvalue $\lambda = -1$.

Problem: Solve $Ax = \lambda x$

- Consider an eigenvector x of matrix A. In that case $Ax = \lambda x \Rightarrow Ax - \lambda x = \mathbf{0}$ ($\mathbf{0}$ is the zero vector). Therefore, $(A - \lambda I)x = \mathbf{0}$.

  In order for the above set of equations to have a non-zero solution, the nullspace of $(A - \lambda I)$ must be non-zero, i.e., the matrix $(A - \lambda I)$ must be singular. Therefore, $\det(A - \lambda I) = 0$.

- I now have an equation for λ. It is called the **characteristic equation**, or the **eigenvalue equation**. From the roots of this equation we can find the eigenvalues.

- I might have repeated λs. **This might cause problems but I will deal with it later.**

- After I find λ, I can find x from $(A - \lambda I)x = \mathbf{0}$. Basically, I will be looking for the nullspace of $(A - \lambda I)$.
- The eigenvalues of A^T are obtained through the equation $\det(A^T - \lambda I) = 0$. But: $\det(A^T - \lambda I) = \det(A^T - \lambda I^T) = \det[(A - \lambda I)^T] = \det(A - \lambda I)$.
- Therefore, the eigenvalues of A^T are the same as the eigenvalues of A.

Imperial College London

Solve $Ax = \lambda x$. An example.

- Consider the matrix $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. **A symmetric matrix always has real eigenvalues.**
- Furthermore, **the eigenvectors of a symmetric matrix can be chosen to be orthogonal.**
- $\det(A - \lambda I) = (3-\lambda)^2 - 1 = 0 \Rightarrow 3 - \lambda = \pm 1 \Rightarrow \lambda = 3 \pm 1 \Rightarrow \lambda_1 = 4,\ \lambda_2 = 2$.
  Or $\det(A - \lambda I) = \lambda^2 - 6\lambda + 8 = 0$. Note that $6 = \lambda_1 + \lambda_2$ and $8 = \det(A) = \lambda_1\lambda_2$.
- Find the eigenvector for $\lambda_1 = 4$.
  $A - 4I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x = y$
- Find the eigenvector for $\lambda_2 = 2$.
  $A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x = -y$
- Notice that there are **families of eigenvectors**, not single eigenvectors.

Compare the two matrices given previously

- Consider the matrix $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. As shown it has eigenvectors $\begin{bmatrix} x \\ x \end{bmatrix}$ and $\begin{bmatrix} -x \\ x \end{bmatrix}$ with eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$.

- Consider the matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, also with eigenvectors $\begin{bmatrix} x \\ x \end{bmatrix}$ and $\begin{bmatrix} -x \\ x \end{bmatrix}$ and eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$.

- **We observe that $A = B + 3I$. The eigenvalues of A are obtained from the eigenvalues of B if we increase them by 3.**

- **The eigenvectors of A and B are the same.**

Generalization of the above observation

- Consider the matrix $A = B + cI$.
- Consider an eigenvector x of B with eigenvalue λ. Then $Bx = \lambda x$ and therefore,
  $Ax = (B + cI)x = Bx + cIx = Bx + cx = \lambda x + cx = (\lambda + c)x$
  A has the same eigenvectors with B with eigenvalues $\lambda + c$.
- **There aren't any properties that enable us to find the eigenvalues of $\boldsymbol{A + B}$ and $\boldsymbol{AB}$.**

Example

- Take a matrix that rotates every vector by 90^{o}.
- This is $Q = \begin{bmatrix} \cos(90) & -\sin(90) \\ \sin(90) & \cos(90) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
- $\lambda_1 + \lambda_2 = 0$ and $\det(Q) = \lambda_1 \lambda_2 = 1$.
- **What vector can be parallel to itself after rotation?**
- $\det(Q - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1 = 0 \Rightarrow \lambda = \pm i$.
- In that case we have a **skew symmetric** (or **anti-symmetric**) matrix with
$$Q^T = Q^{-1} = -Q.$$
- We observe that the eigenvalues are complex.
- Complex eigenvalues always come in complex conjugate pairs if the associated matrix is real.

Example

- Consider $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$.
- $\lambda_1 + \lambda_2 = 6$ and $\det(\lambda_1 \lambda_2) = 9$.
- $\det(A - \lambda I) = \det \begin{bmatrix} 3 - \lambda & 1 \\ 0 & 3 - \lambda \end{bmatrix} = (3 - \lambda)^2 = 0 \Rightarrow \lambda_{1,2} = 3$
- **The eigenvalues of a triangular matrix are the values of the diagonal.**
- For that particular case we have

$\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix} \Rightarrow y = 0$ and x can be any number.

Matrix diagonalization
The case of independent eigenvectors

- Suppose we have n independent eigenvectors of a matrix A. We call them x_i.
- The associated eigenvalues are λ_i.
- We put them in the columns of a matrix S.
- We form the matrix:

$$AS = A[x_1 \quad x_2 \quad \cdots \quad x_n] = [\lambda_1 x_1 \quad \lambda_2 x_2 \quad \ldots \quad \lambda_n x_n] =$$

$$[x_1 \quad x_2 \quad \cdots \quad x_n]\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = S\Lambda \Rightarrow AS = S\Lambda$$

$S^{-1}AS = \Lambda$ or $A = S\Lambda S^{-1}$

- The above formulation of A is very important in Mathematics and Engineering and is called **matrix diagonalization**.

Matrix diagonalization: Eigenvalues of A^k

- If $Ax = \lambda x \Rightarrow A^2x = \lambda Ax \Rightarrow A^2x = \lambda^2 x$.
- Therefore, the eigenvalues of A^2 are λ^2.
- The eigenvectors of A^2 remain the same since $A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}$
- In general $A^k = S\Lambda^k S^{-1}$
- $\lim(A^k) = 0$ if the eigenvalues of A have the property $|\lambda_i| < 1$.
- A matrix has n independent eigenvectors and therefore is diagonalizable if all the eigenvalues are different.
- If there are repeated eigenvalues a matrix may, or may not have independent eigenvectors. As an example consider the identity matrix. Its eigenvectors are the row (or column) vectors. They are all independent. However, the eigenvalues are all equal to 1.
- Find the eigenvalues of $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, a \neq 0$. **This is a standard textbook matrix which cannot be diagonalized.**

  (**Answer:** The eigenvectors are of the form $[x \quad 0]^T$)

Application of matrix diagonalization cont. A first order system which evolves with time

- Consider a system that follows an equation of the form $u_{k+1} = Au_k$.
- u_k is the vector which consists of the system parameters which evolve with time.
- The eigenvalues of A characterize fully the behavior of the system.
- I start with a given vector u_0.

  $u_1 = Au_0$, $u_2 = A^2u_0$ and in general $u_k = A^k u_0$
- In order to solve this system I assume that I can write $u_0 = c_1x_1 + c_2 x_2 + \cdots + c_nx_n$ where x_i are the eigenvectors of matrix A. This is a standard approach to the solution of this type of problem.
- $Au_0 = c_1Ax_1 + c_2 Ax_2 + \cdots + c_nAx_n = c_1\lambda_1x_1 + c_2 \lambda_2x_2 + \cdots + c_n\lambda_nx_n$

  $A^{100}u_0 = c_1{\lambda_1}^{100}x_1 + c_2 {\lambda_2}^{100}x_2 + \cdots + c_n{\lambda_n}^{100}x_n = S\Lambda^{100}c$

  c is a column vector that contains the coefficients c_i.

  S, Λ are the matrices defined previously.

Application of matrix diagonalization cont. Fibonacci example: Convert a second order scalar problem into a first order system

- I will take two numbers which I denote with $F_0 = 0$ and $F_1 = 1$.
- The Fibonacci sequence of numbers is given by the two initial numbers given above and the relationship $F_k = F_{k-1} + F_{k-2}$.
- The sequence is 0,1,1,2,3,5,8,13 and so on.
- How can I get a closed form formula for the 100th Fibonacci number or any Fibonacci number?
- I will present a standard approach to this problem.

  I define a vector $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$ and an extra equation $F_{k+1} = F_{k+1}$. By merging the two equations I obtain the following matrix form:

$$u_{k+1} = \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = Au_k$$

- I managed to convert the second order scalar problem into a first order matrix problem.

Application of matrix diagonalization cont. Fibonacci example cont. Convert a second order scalar problem into a first order system

- The eigenvalues of A are obtained from
$$\det\begin{bmatrix}1-\lambda & 1\\ 1 & -\lambda\end{bmatrix} = -(1-\lambda)\lambda - 1 = 0 \Rightarrow \lambda^2-\lambda-1=0$$
- Observe the analogy between $\lambda^2-\lambda-1=0$ and $F_k - F_{k-1} - F_{k-2} = 0$.
- $\lambda_{1,2} = \frac{1\pm\sqrt{5}}{2}$. Eigenvalues add up to 1. The matrix A is diagonalizable.
- It can be shown that the eigenvectors are $x_1 = \begin{bmatrix}\lambda_1\\ 1\end{bmatrix}$ and $x_2 = \begin{bmatrix}\lambda_2\\ 1\end{bmatrix}$.
- How can I get a formula for the 100th Fibonacci number?
- $u_{k+1} = \begin{bmatrix}F_{k+2}\\ F_{k+1}\end{bmatrix} = Au_k = A^2u_{k-1}$ etc. Therefore, $u_{100} = A^{100}u_0 = S\Lambda^{100}S^{-1}u_0$
- $\Lambda^{100} = \begin{bmatrix}{\lambda_1}^{100} & 0\\ 0 & {\lambda_2}^{100}\end{bmatrix} \cong \begin{bmatrix}{\lambda_1}^{100} & 0\\ 0 & 0\end{bmatrix}$, $S = \begin{bmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{bmatrix}$, $S^{-1} = \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{bmatrix}$
- $u_{100} \cong \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{bmatrix}\begin{bmatrix}{\lambda_1}^{100} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{bmatrix}\begin{bmatrix}{\lambda_1}^{100} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix} =$
$\frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{bmatrix}\begin{bmatrix}{\lambda_1}^{100}\\ 0\end{bmatrix} = \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}{\lambda_1}^{101}\\ {\lambda_1}^{100}\end{bmatrix} \Rightarrow F_{100} = \frac{1}{\lambda_1-\lambda_2}{\lambda_1}^{100}$

Applications of matrix diagonalization cont.

First order differential equations $\frac{du}{dt} = Au$

Problem:

Solve the system of differential equations:

$$\frac{du_1(t)}{dt} = -u_1(t) + 2u_2(t), \qquad \frac{du_2(t)}{dt} = u_1(t) - 2u_2(t)$$

Solution:

- We convert the system of equations into a matrix form $\begin{bmatrix} \frac{du_1(t)}{dt} \\ \frac{du_2(t)}{dt} \end{bmatrix} = A \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$.
- We set $u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$. We impose the **initial conditions** $u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- The system's matrix is $A = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$. The matrix A is singular. One of the eigenvalues is zero. Therefore, from the trace we conclude that the eigenvalues are $\lambda_1 = 0, \ \lambda_2 = -3$.
- The solution of the above system depends exclusively on the eigenvalues of A.
- We can easily show that the eigenvectors are $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Applications of matrix diagonalization cont.
First order differential equations $\frac{du}{dt} = Au$

- The solution of the above system of equations is of the form

$$u(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2$$

Problem:

Verify the above by plugging-in $e^{\lambda_i t} x_i$ to the equation $\frac{du}{dt} = Au$.

- Let's find $u(t) = c_1 e^{0t} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- c_1, c_2 comes from the initial conditions.

  $c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow S \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = u(0) \Rightarrow c_1 = \frac{1}{3}, c_2 = \frac{1}{3}$.
- **The Steady State** of the system is defined as $u(\infty) = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$.
- **Stability** is achieved if the real part of the eigenvalues is negative.
- Note that if complex eigenvalues exists, they appear in conjugate pairs.

Imperial College London

Applications of matrix diagonalization cont.
First order differential equations $\frac{du}{dt} = Au$

- **Stability** is achieved if the real part of the eigenvalues is negative.

- We **do** have a **steady state** if at least one eigenvalue is 0 and the rest of the eigenvalues have negative real part.

- The system is unstable if at least one eigenvalue has a positive real part.

- For stability the trace of the system's matrix must be negative.

- The reverse is not true. Obviously a negative trace does not guarantee stability.

Applications of matrix diagonalization cont. First order differential equations $\frac{du}{dt} = Au$

- Consider a first-order differential equation $\frac{du}{dt} = Au$.
- Recall the matrix S which contains the eigenvectors of A defined previously.
- I set $u(t) = Sv(t)$ and therefore the differential equation becomes:
- $S\frac{dv(t)}{dt} = ASv(t) \Rightarrow \frac{dv(t)}{dt} = S^{-1}ASv(t) = \Lambda v(t)$
- This is an interesting result which can be found in various engineering applications.
- I start from a system of equations $\frac{du(t)}{dt} = Au(t)$ which are **coupled** (or **dependent** or **correlated**) and I end up with a set of equations which are **decoupled** and easier to solve.
- From the relationship $\frac{dv(t)}{dt} = \Lambda v(t)$ we get $v(t) = e^{\Lambda t}v(0)$ and $u(t) = Se^{\Lambda t}S^{-1}u(0)$ with $e^{At} = Se^{\Lambda t}S^{-1}$

Question: What is the exponential of a matrix? See next slide.

Applications of matrix diagonalization cont. Second order homogeneous differential equations

- How do I change the second order homogeneous differential equation

$$y''(t) + by'(t) + ky(t) = 0$$

  to two first order ones?

- I use the equation above and the additional equation $y'(t) = y'(t)$.

- I define $u(t) = \begin{bmatrix} y'(t) \\ y(t) \end{bmatrix}$ and from the two available equations I form the matrix representation shown below.

$$u'(t) = \begin{bmatrix} y''(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y'(t) \\ y(t) \end{bmatrix}$$

$$u'(t) = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} u(t)$$

Imperial College London

Applications of matrix diagonalization cont. Higher order homogeneous differential equations

- How do I change the n^{th} order homogeneous differential equation

$$y^{(n)}(t) + b_1 y^{(n-1)}(t) + \cdots + b_{n-1} y(t) = 0$$

to n first order ones?

- I define $u(t) = \begin{bmatrix} y^{(n-1)}(t) \\ \vdots \\ y'(t) \\ y(t) \end{bmatrix}$ and n additional equations $y^{(i)}(t) = y^{(i)}(t)$, $i = 1, \ldots, (n-1)$ and therefore, I obtain the matrix representation below:

$$u'(t) = \begin{bmatrix} y^{(n)}(t) \\ y^{(n-1)}(t) \\ \vdots \\ y''(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} -b_1 & -b_2 & \ldots & -b_{n-2} & -b_{n-1} \\ 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix} u(t)$$

- As previously analysed the solution of this system is explicitly affiliated to the eigenvalues of the system's matrix.

Diagonal matrix exponentials

- The exponential $e^{\Lambda t}$ of a diagonal matrix is given by:

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix}$$

  with λ_i being the elements of the diagonal.
- Furthermore,

$$\Lambda^t = \begin{bmatrix} {\lambda_1}^t & 0 & \dots & 0 \\ 0 & {\lambda_2}^t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & {\lambda_n}^t \end{bmatrix}$$

- As we already showed

$$\lim_{t\to\infty} e^{\Lambda t} = 0 \text{ if } \mathrm{Re}(\lambda_i) < 0, \forall i$$

$$\lim_{t\to\infty} \Lambda^t = 0 \text{ if } |\lambda_i| < 1, \forall i$$

Matrix exponentials e^{At}

- Consider the exponential e^{At}.
  - The Taylor series expansion is $e^{x} = \sum_0^\infty \frac{x^n}{n!}$.
  - Similarly, $e^{At} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots + \frac{(At)^n}{n!} + \cdots$
  - $e^{At} = I + S\Lambda S^{-1}t + \frac{S\Lambda^2 S^{-1}t^2}{2} + \frac{S\Lambda^3 S^{-1}t^3}{6} + \cdots + \frac{S\Lambda^n S^{-1}t^n}{n!} + \cdots = Se^{\Lambda t}S^{-1}$
  - The assumption built-in to the above formula is that A must be diagonalizable.

- Furthermore, note that $\frac{1}{1-x} = \sum_0^\infty x^n$.
  - For matrices we have $(I - At)^{-1} = I + At + (At)^2 + (At)^3 + \cdots$
  - This sum converges if the magnitudes of the eigenvalues of matrix At are of magnitude less than 1, i.e., $|\lambda(At)| < 1$. This statement is proven using diagonalization.

Stochastic matrices

- Consider a matrix A with the following properties:
  - It is square
  - All entries are positive and real.
  - The elements of each column or each row or both each column and each row add up to 1.
  - Based on the above, a matrix that exhibits the above properties will have all entries ≤ 1.
- This is called a **stochastic matrix**.
- Stochastic matrices are also called **Markov, probability, transition, or substitution matrices**.
- The entries of a stochastic matrix usually represent a probability.
- Stochastic matrices are widely used in probability theory, statistics, mathematical finance and linear algebra, as well as computer science.

Stochastic matrices. Types.

- There are several types of stochastic matrices:
  - A **right stochastic matrix** is a matrix of nonnegative real entries, with each row's elements summing to 1.
  - A **left stochastic matrix** is a matrix of nonnegative real entries, with each column's elements summing to 1.
  - A **doubly stochastic matrix** is a matrix of nonnegative real entries with each row's and each column's elements summing to 1.
- A stochastic matrix often describes a so called **Markov chain** X_t over a finite state space S.
- Generally, an $n \times n$ stochastic matrix is related to n "states".
- If the probability of moving from state i to state j is $P_r(j/i) = p_{ij}$, the stochastic matrix P is given by using p_{ij} as the i^{th} row and j^{th} column element:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

- Depending on the particular problem, the above matrix can be formulated in such a way so that it is either right or left stochastic.

Imperial College London

Products of stochastic matrices. Stochastic vectors.

- An example of a left stochastic matrix is the following:

$$A = \begin{bmatrix} 0.1 & 0.01 & 0.3 \\ 0.2 & 0.99 & 0.3 \\ 0.7 & 0 & 04 \end{bmatrix}$$

- You can prove that if A and B are stochastic matrices of any type, then AB is also a stochastic matrix of the same type.
- Consider two left stochastic matrices A and B with elements a_{ij} and b_{ij} respectively, and $C = AB$ with elements c_{ij}.
- Let us find the sum of the elements of the j^{th} column of C:

$\sum_{i=1}^{n} c_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{n} a_{ik} = 1 \cdot 1 = 1$

- Based on the above any power of a stochastic matrix is a stochastic matrix.
- Furthermore, a vector with real, nonnegative entries p_k, for which all the p_k add up to 1, is called a **stochastic vector**. For a stochastic matrix, every column or row or both is a stochastic vector.
- I am interested in the eigenvalues and eigenvectors of a stochastic matrix.

Stochastic matrices and their eigenvalues

- I would like to prove that $\lambda = 1$ is always an eigenvalue of a stochastic matrix.
- Consider again a left stochastic matrix A.
- Since the elements of each column of A add up to 1, the elements of each row of A^T should add up to 1. Therefore,

$$A^T \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

- Therefore, 1 is an eigenvalue of A^T .
- As already shown, the eigenvalues of A and A^T are the same, which implies that 1 is also an eigenvalue of A.
- Since $\det(A - I) = 0$, the matrix $A - I$ is singular, which means that there is a vector x for which

$$(A - I)x = \mathbf{0} \Rightarrow Ax = x$$

- A vector of the null space of $A - I$ is the eigenvector of A that corresponds to eigenvalue $\lambda = 1$.

Stochastic matrices and their eigenvalues cont.

- I would like now to prove that the eigenvalues of a stochastic matrix have magnitude smaller or equal to 1.
- Assume that $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is an eigenvector of a **right** stochastic matrix A with an associated eigenvalue $|\lambda| > 1$. Then $Av = \lambda v$ implies $\sum_{j=1}^{n} A_{ij} v_j = \lambda v_i$.
- We can see that $\sum_{j=1}^{n} A_{ij} v_j \leq v_{\max} \sum_{j=1}^{n} A_{ij} = v_{\max}$.
- All the elements of the vector Av are smaller or equal to $v_{\max}$.
- If $|\lambda| > 1$ at least one element of the vector λv (which is equal to Av) will be greater than $v_{\max}$ $(\lambda v_{\max})$.
- Based on the above, the assumption of an eigenvalue being larger than 1 can not be valid.
- For left stochastic matrices we use A^T in the proof above, and the fact that the eigenvalues of A and A^T are the same.

Imperial College London

An application of stochastic matrices: First order systems

- Consider again the first order system described by an equation of the form $u_k = A^k u_0$, where A is now a stochastic matrix.
- Previously, we managed to write $u_k = A^k u_0 = c_1 {\lambda_1}^k x_1 + c_2 {\lambda_2}^k x_2 + \cdots$ where λ_i and x_i are the eigenvalues and eigenvectors of matrix A, respectively.
- Note that the above relationship requires a complete set of eigenvectors.
- If $\lambda_1 = 1$ and $|\lambda_i| < 1, i > 1$ then the steady state of the system is $c_1 x_1$ (which is part of the initial condition u_0).
- I will use an example where A is a 2×2 matrix. Generally, an $n \times n$ stochastic matrix is related to n "states". Assume that the 2 "states" are 2 UK cities.
- I take London and Oxford. I am interested in the population of the two cities and how it evolves.
- I assume that people who inhabit these two cities move between them only.

$$\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k+1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k}$$

- It is now obvious that the column elements are positive and also add up to 1 because they represent probabilities.

Application of stochastic matrices (cont.)

- I assume that $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$. Then $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{k=1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = \begin{bmatrix} 200 \\ 800 \end{bmatrix}$.
- What is the population of the two cities after a long time?
- Consider the matrix $\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0.7$.

  (Notice that the second eigenvalue is found by the trace of the matrix.)
- The eigenvectors of this matrix are $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

$$\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_k = c_1 {\lambda_1}^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 {\lambda_2}^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 0.7^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- I find $c_1, \ c_2$ from the initial condition $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{k=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$

  $\begin{bmatrix} 0 \\ 1000 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and therefore, $c_1 = \frac{1000}{3}$ and $c_2 = \frac{2000}{3}$.

- $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{k\to\infty} = \frac{1000}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{2000}{3} 0.7^{k\to\infty} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2000}{3} \\ \frac{1000}{3} \end{bmatrix}$

Application of stochastic matrices (cont.)

- Stochastic models facilitate the modeling of various real life engineering applications.
- An example is the modeling of the movement of people without gain or loss: total number of people is conserved.

Symmetric matrices

- In this lecture we will be interested in symmetric matrices.
- In case of real matrices, symmetry is defined as $A = A^T$.
- In case of complex matrices, symmetry is defined as $A^* = A^T$ or $A^{*T} = A$. A matrix which possesses this property is called **Hermitian**.
- We can also use the symbol $A^H = A^{*T}$.
- We will prove that the eigenvalues of a symmetric matrix are real.
- The eigenvectors of a symmetric matrix can be chosen to be orthogonal. If we also choose them to have a magnitude of 1, then the eigenvectors can be chosen to form an orthonormal set of vectors.
- However, the eigenvectors of a symmetric matrix that correspond to different eigenvalues **are** orthogonal (prove is given in subsequent slide).
- For a random matrix with independent eigenvectors we have $A = S\Lambda S^{-1}$.
- For a symmetric matrix with orthonormal eigenvectors we have

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

Real symmetric matrices

Problem:

Prove that the eigenvalues of a symmetric matrix occur in complex conjugate pairs.

Solution:

Consider $Ax = \lambda x$.

If we take complex conjugate in both sides we get

$$(Ax)^* = (\lambda x)^* \Rightarrow A^*\, x^* = \lambda^*\, x^*$$

If A is real then $Ax^* = \lambda^*\, x^*$. Therefore, if λ is an eigenvalue of A with corresponding eigenvector x then λ^* is an eigenvalue of A with corresponding eigenvector x^*.

Real symmetric matrices cont.

Problem:

Prove that the eigenvalues of a symmetric matrix are real.

Solution:

We proved that if A is real then $Ax^* = \lambda^* x^*$.

If we take transpose in both sides we get

$$x^{*T}A^T = \lambda^* x^{*T} \Rightarrow x^{*T}A = \lambda^* x^{*T}$$

We now multiply both sides from the right with x and we get

$$x^{*T}Ax = \lambda^* x^{*T}x$$

We take now $Ax = \lambda x$. We now multiply both sides from the left with x^{*T} and we get

$$x^{*T}Ax = \lambda x^{*T}x.$$

From the above we see that $\lambda x^{*T}x = \lambda^* x^{*T}x$ and since $x^{*T}x \neq 0$, we see that $\lambda = \lambda^*$.

Real symmetric matrices cont.

Problem:

Prove that the eigenvectors of a symmetric matrix **which correspond to different eigenvalues** are always perpendicular.

Solution:

Suppose that $Ax = \lambda_1 x$ and $Ay = \lambda_2 y$ with $\lambda_1 \neq \lambda_2$.

$$(\lambda_1 x)^T y = x^T \lambda_1 y = (Ax)^T y = x^T A y = x^T \lambda_2 y$$

The conditions $x^T \lambda_1 y = x^T \lambda_2 y$ and $\lambda_1 \neq \lambda_2$ give $x^T y = 0$.

The eigenvectors x and y are perpendicular.

Complex matrices. Complex symmetric matrices.

- Let us find which complex matrices have real eigenvalues and orthogonal eigenvectors.
- Consider $Ax = \lambda x$ with A possibly complex.
- If we take complex conjugate in both sides we get

$$(Ax)^* = (\lambda x)^* \Rightarrow A^* \, x^* = \lambda^* \, x^*$$

- If we take transpose in both sides we get

$$x^{*T} A^{*T} = \lambda^* x^{*T}$$

- We now multiply both sides from the right with x we get

$$x^{*T} A^{*T} x = \lambda^* x^{*T} x$$

- We take now $Ax = \lambda x$. We now multiply both sides from the left with x^{*T} and we get

$$x^{*T} A x = \lambda x^{*T} x.$$

- From the above we see that if $A^{*T} = A$ then $\lambda x^{*T} x = \lambda^* x^{*T} x$ and since $x^{*T} x \neq 0$, we see that $\lambda = \lambda^*$.
- If $A^{*T} = A$ the matrix is called **Hermitian**.

Complex vectors and matrices

- Consider a complex column vector $z = [z_1 \quad z_2 \quad \cdots \quad z_n]^T$.
- Its length is $z^{*T} z = \sum_{i=1}^{n} |z_i|^2$.
- As already mentioned, when we both transpose and conjugate we can use the symbol $z^H = z^{*T}$ (Hermitian).
- The inner product of two complex vectors is $y^{*T} x = y^H x$.
- For complex matrices the symmetry is defined as $A^{*T} = A$. As already mentioned, these are called Hermitian matrices.
- They have real eigenvalues and perpendicular eigenvectors. If these are complex we check their length using ${q_i}^{*T} q_i$ and also $Q^{*T} Q = I$.

  Example: Consider the matrix

$$A = \begin{bmatrix} 2 & 3+i \\ 3-i & 5 \end{bmatrix}$$

  Eigenvalues are found from:

$$(2-\lambda)(5-\lambda) - (3+i)(3-i) = 0$$
$$\Rightarrow \lambda^2 - 7\lambda + 10 - (9 - 3i + 3i - i^2) = 0 \Rightarrow \lambda(\lambda - 7) = 0$$

Eigenvalue sign

- We proved that:
  - The eigenvalues of a symmetric matrix, either real or complex, are real.
  - The eigenvectors of a symmetric matrix can be chosen to be orthogonal.
  - The eigenvectors of a symmetric matrix that correspond to different eigenvalues are orthogonal.
- **Do not forget the definition of symmetry for complex matrices.**
- It can be proven that the signs of the pivots are the same as the signs of the eigenvalues.
- Just to remind you:

  Product of pivots=Product of eigenvalues=Determinant