Imperial College London

Maths for Signals and Systems

Linear Algebra in Engineering

Lectures 13-14, Tuesday 10th November 2015

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING

IMPERIAL COLLEGE LONDON

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.
- For a random matrix with independent eigenvectors we have $A = S\Lambda S^{-1}$.
- For a symmetric matrix with orthonormal eigenvectors we have

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

Imperial College London

Real matrices

Problem:

Prove that the eigenvalues of a real 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

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 $\lambda = \lambda^*$.

Imperial College London

Complex matrices. Complex symmetric matrices.

- Let us find which complex matrices have real eigenvalues.
- 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 $\lambda = \lambda^*$.

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 $A^{*T} = A$. These are called Hermitian matrices.
- They have real eigenvalues and perpendicular unit 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:

$$(2-\lambda)(5-\lambda) - (3+i)(3-i) = 0$$

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

Imperial College London

Eigenvalue sign

- We proved that the eigenvalues of a symmetric matrix, either real or complex, are real.

- **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

Imperial College London

Positive definite matrices

- A symmetric or Hermitian matrix is positive definite if and only if (iff) all its eigenvalues are real and positive.
- Therefore, the pivots are positive and the determinant is positive.
- Positive determinant doesn’t guarantee positive definiteness.

Example: Consider the matrix

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

Eigenvalues are obtained from:

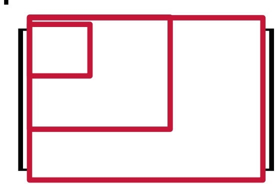
$$(5-\lambda)(3-\lambda)-4=0 \Rightarrow \lambda^2-8\lambda+11=0$$

$$\lambda_{1,2}=\frac{8\pm\sqrt{64-44}}{2}=\frac{8\pm\sqrt{20}}{2}=4\pm\sqrt{5}$$

The eigenvalues are positive and the matrix is symmetric, therefore, the matrix is positive definite.

Positive definite matrices cont.

- We are talking about symmetric matrices.
- We have various tests for positive definiteness. Consider the 2×2 case of a positive definite matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.
  - The eigenvalues are positive $\lambda_1 > 0, \lambda_2 > 0$.
  - The pivots are positive $a > 0$, $\frac{ac-b^2}{a} > 0$.
  - All determinates of leading ("north west") sub-matrices are positive $a > 0, ac - b^2 > 0$.



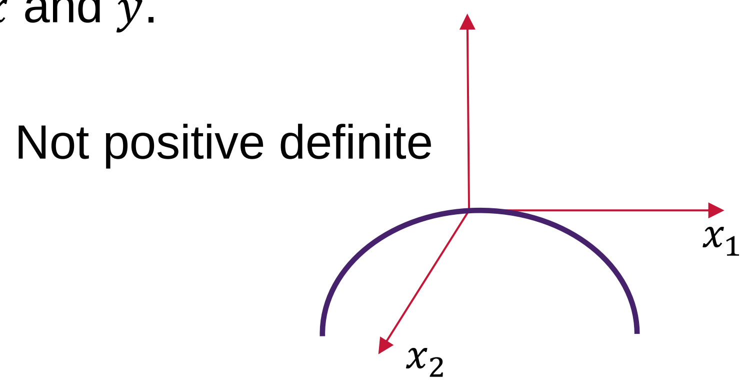
  - Quadratic form is positive $x^T A\, x > 0$, x is any vector.

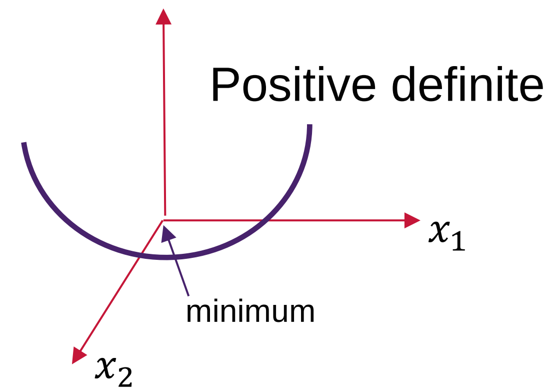
Positive semi-definite matrices

- **Example:** Consider the matrix $\begin{bmatrix} 2 & 6 \\ 6 & x \end{bmatrix}$
  - Which sufficiently large values of x makes the matrix positive definite? The answer is $x > 18$. In that case we obtain the matrix $\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$.
  - For $x = 18$ the matrix is **positive semi-definite**. The eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 20$. One of its eigenvalues is zero.
  - It has only one pivot since the matrix is singular. The pivots are 2 and 0.
  - Its quadratic form is $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 18x_2^2$.
  - This is equal to $ax_1^2 + 2bx_1x_2 + cx_2^2 \geq 0$. This formula is a so called **quadratic form**.
  - In that case the matrix marginally failed the test.

Graph of quadratic form

- In mathematics, a **quadratic form** is a homogeneous polynomial of degree two in a number of variables. For example, the condition for positive-definiteness of a 2×2 matrix, $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$, is a quadratic form in the variables x and y.





- For the positive definite case we have:
  - Obviously, first derivatives must be zero at the minimum. This condition is not enough.
  - Second derivatives' matrix is positive definite

$$\begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix}, f_{x_1x_1}f_{x_2x_2}-2f_{x_1x_2} > 0.$$

  - Positive for a number turns into positive definite for a matrix.

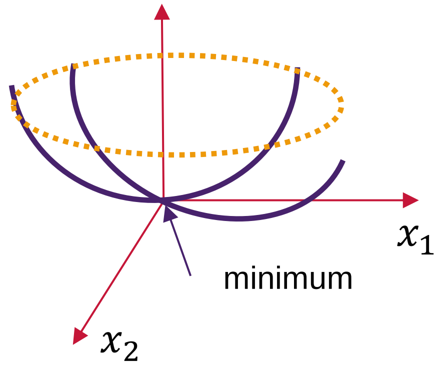
Imperial College London

Graph of quadratic form

- **Example:**

  $\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$, $\text{trace}(A) = 22 = \lambda_1 + \lambda_2$, $\det(A) = 4 = \lambda_1 \lambda_2$

  - $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 20x_2^2$
  - $f(x_1, x_1) = 2x_1^2 + 12x_1x_2 + 20x_2^2 = 2(x_1 + 3x_2)^2 + 2x_2^2.$



  - A horizontal intersection could be $f(x_1, x_1) = 1$. It is an ellipse.
  - Its quadratic form is $2(x_1 + 3x_2)^2 + 2x_2^2 = 1$.

Graph of quadratic form

- Example:

$\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$, $\text{trace}(A) = 22 = \lambda_1 + \lambda_2$, $\det(A) = 4 = \lambda_1 \lambda_2$

  - $f(x_1, x_1) = 2x_1^2 + 12x_1x_2 + 20x_2^2 = \mathbf{2}(x_1 + \mathbf{3}x_2)^2 + \mathbf{2}x_2^2$
  - Note that computing the square form is effectively elimination

$$A = \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \xrightarrow[(2)-3(1)]{} \begin{bmatrix} \mathbf{2} & 6 \\ 0 & \mathbf{2} \end{bmatrix} = u \text{ and } L = \begin{bmatrix} 1 & 0 \\ \mathbf{3} & 1 \end{bmatrix}$$

  - The pivots and the multipliers appear in the quadratic form when we compute the square.
  - Pivots are the square multipliers so positive pivots imply sum of squares and hence positive definiteness.

Graph of quadratic form

- **Example:** Consider the matrix $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$
  - The leading ("north west") determinants are 2,3,4.
  - The pivots are 2, 3/2, 4/3.
  - The quadratic form is $\boldsymbol{x}^T A\, \boldsymbol{x} = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3$.
  - The eigenvalues of A are $\lambda_1 = 2 - \sqrt{2}, \lambda_2 = 2, \lambda_3 = 2 + \sqrt{2}$
  - The matrix A is positive definite when $\boldsymbol{x}^T A\, \boldsymbol{x} > 0$.

The Discrete Fourier Transform (DFT) matrix

- The $n \times n$ Fourier matrix is defined as:

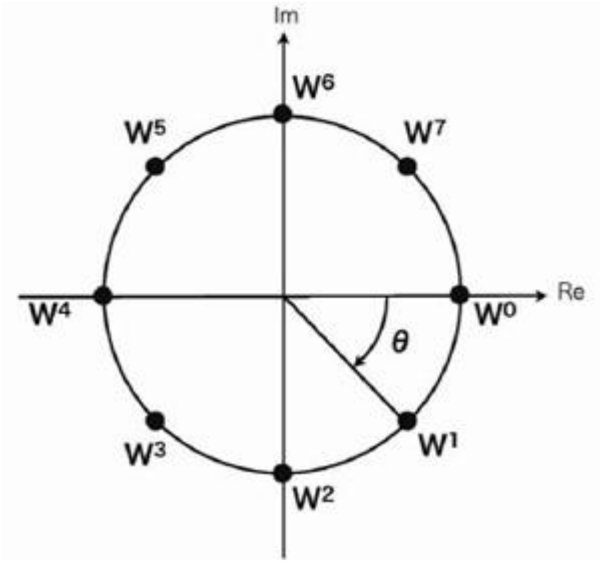
$$F_n = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & \dots & w^{(n-1)} \\ 1 & w^2 & w^4 & \dots & w^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{(n-1)} & w^{2(n-1)} & \dots & w^{(n-1)(n-1)} \end{bmatrix}$$

- In this matrix we will number the first row and column with 0.
- We define $w = e^{-i\frac{2\pi}{n}}$. For w is preferable to use polar representation.
- $F_n(i,j) = w^{ij}$.
- We must stress out that it is better to use the notation w_n instead of w.
- I have avoided this notation to make things look simpler.

Imperial College London

The Discrete Fourier Transform (DFT) matrix cont.

- The parameter $w = e^{-i\frac{2\pi}{n}}$ lies on the unit circle shown below. The case depicted below refers to $n = 8$ where the points $w^m = e^{-i\frac{2\pi m}{8}}, m = 0, \ldots, 7$ of the second row (row 1) of the Fourier matrix are shown.



- We must stress out that the Fourier matrix is totally constructed out of numbers of the form ${w_n}^k$.

The Discrete Fourier Transform (DFT) matrix for $n = 4$

- The parameter $w_4 = e^{-i\frac{2\pi}{4}} = e^{-i\frac{\pi}{2}} = \cos\left(\frac{\pi}{2}\right) - i\sin\left(\frac{\pi}{2}\right) = -i$.
- The quantities inside Fourier matrix are $1, i, i^2, i^3, i^4, i^6, i^9$.

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & i^2 & -i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & -i^3 & i^6 & -i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$$

- The columns of this matrix are orthogonal.
- Remember that the inner product of 2 complex vectors is $y^{*T}x = y^H x$.

The Discrete Fourier Transform (DFT) matrix for $n = 4$ cont.

- I can show that the columns are orthogonal but they are not orthonormal.
- I can fix this by dividing the Fourier matrix with the length of the rows (columns). In this case it is 2. Therefore, I can write:

$$F_4 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & i^2 & -i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & -i^3 & i^6 & -i^9 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$$

- We can easily show that ${F_4}^H F_4 = I$.

Imperial College London

The Fast Fourier Transform (FFT)

- It can be proven that there is a connection between F_{2n} and F_n.
- This is expected from the fact that ${w_{2n}}^2 = w_n$. It can be shown that:

diagonal matrix D_n

$$[F_{2n}] = \begin{bmatrix} I_n & D_n \\ I_n & -D_n \end{bmatrix} \begin{bmatrix} F_n & \mathbf{O}_n \\ \mathbf{O}_n & F_n \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

permutation matrix P_{2n}

- When $[F_{2n}]$ is multiplied by a column vector in order to obtain the Fourier Transform of the signal, we require $(2n)^2$ multiplications.
- When $[F_{2n}]$ is decomposed as above, P_{2n} does not contribute to multiplications, $\begin{bmatrix} F_n & \mathbf{O}_n \\ \mathbf{O}_n & F_n \end{bmatrix}$ requires $2 \times (n)^2$ multiplications and $\begin{bmatrix} I_n & D_n \\ I_n & -D_n \end{bmatrix}$ requires n multiplications.
- In total $2 \times (n)^2 + n < (2n)^2$.

The Fast Fourier Transform (FFT) cont.

- In the previous analysis the matrix D_n is defined as:

$$D_n = \begin{bmatrix} 1 & & & & \\ & w & & & \\ & & w^2 & & \\ & & & \ddots & \\ & & & & w^{n-1} \end{bmatrix}$$

- We start requiring $(2n)^2$ multiplications and manage to reduce them to $2 \times (n)^2 + n$ multiplications.

The Fast Fourier Transform (FFT) cont.

- The next step is to break the F_n down. We use the above idea recursively.

$$[F_{2n}] = \begin{bmatrix} I_n & D_n \\ I_n & -D_n \end{bmatrix} \begin{bmatrix} F_n & \mathbf{0}_n \\ \mathbf{0}_n & F_n \end{bmatrix} P_{2n} =$$

$$= \begin{bmatrix} I_n & D_n \\ I_n & -D_n \end{bmatrix} \begin{bmatrix} \begin{matrix} I_{n/2} & D_{n/2} \\ I_{n/2} & -D_{n/2} \end{matrix} & \mathbf{0}_n \\ \mathbf{0}_n & \begin{matrix} I_{n/2} & D_{n/2} \\ I_{n/2} & -D_{n/2} \end{matrix} \end{bmatrix} \begin{bmatrix} \begin{matrix} F_{n/2} & \mathbf{0}_{n/2} \\ \mathbf{0}_{n/2} & F_{n/2} \end{matrix} & \mathbf{0}_n \\ \mathbf{0}_n & \begin{matrix} F_{n/2} & \mathbf{0}_{n/2} \\ \mathbf{0}_{n/2} & F_{n/2} \end{matrix} \end{bmatrix} \begin{bmatrix} P_n & \mathbf{0}_n \\ \mathbf{0}_n & P_n \end{bmatrix} P_{2n}$$

- We started with $(2n)^2$ multiplications and manage to reduce them to $2 \times (n)^2 + n$ multiplication.
- Now the n^2 multiplications are reduced to $2 \times (n/2)^2 + n/2$ multiplications.

The Fast Fourier Transform (FFT) cont.

- We can carry on this recursive procedure until we reach 1×1 Fourier matrices.
- We will have a large number of matrices piling up.
- It can be proven that if we start with a matrix of size $n \times n$ the total number of multiplications is reduced to

$$\frac{1}{2} n \log_2(n)$$

- Consider $n = 1024 = 2^{10}$. In that case $n^2 > 1{,}000{,}000$.
- $\frac{1}{2} 1024 \log_2(1024) = 5 \times 1024$.
- We reduced the multiplications from 1024×1024 to 5×1024, i.e., by a factor of 200.
- **The Fast Fourier Transform is one of the most important algorithms in modern scientific computing.**