4.6 Special Types of Matrices

Session 21
April 27, 2020

For some types of matrices we know even more very useful facts about eigenvectors.
Let us first introduce a special type of matrix, and then discuss why it is useful.
Here again, it will be more convenient to allow for matrices with complex entries; we will specialize to real matrices later.

Let us introduce the following notation:

Definition: For an $n\times n$ matrix A, we define the Hermitian conjugate — also called adjoint — as $A^\dagger := \bar{A}^T$, i.e., the transposed and complex conjugate matrix.

not to be confused with the "classical adjoint" we introduced earlier

Note: • $A^\dagger$ is pronounced as "A dagger"

meaning the imaginary unit i is replaced by $-i$

• In other words: $(A^\dagger)_{ij} = \bar{A}_{ji}$ ← i and j interchanged, and each entry is complex conjugated

• By the definition, we have $(A^\dagger)^\dagger = A$

• Example: For $A = \begin{pmatrix} 3 & 5 & 4 \\ 3 & i & 2i \\ 1 & 2 & 2+i \end{pmatrix}$, we have $A^\dagger = \begin{pmatrix} 3 & 3 & 1 \\ 5 & -i & 2 \\ 4 & -2i & 2-i \end{pmatrix}$.

Where does this come from? If we allow for vectors and matrices with complex entries, we need to write the $\underbrace{\text{scalar product}}_{\text{also called "inner product"}}$ of $\vec{x}$ and $\vec{y}$ as $\bar{\vec{x}}\cdot\vec{y} = \sum_{i=1}^{n} \bar{x}_i y_i$, such that $\bar{\vec{x}}\cdot\vec{x} = \underbrace{|\vec{x}|^2}_{\text{always real}} \geq 0$.

Then $\bar{\vec{x}}\cdot(A\vec{y}) = \sum_{i=1}^{n} \bar{x}_i (Ay)_i = \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i A_{ij} y_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{(A^\dagger)_{ji} x_i}\, y_j = \sum_{j=1}^{n} \overline{(A^\dagger \vec{x})_j}\, y_j = \overline{(A^\dagger\vec{x})}\cdot\vec{y}$,

$A_{ij} = A^T_{ji} = \overline{(A^\dagger)_{ji}}$

so it is a useful notation.

A much nicer notation is to define $\boxed{\underbrace{\langle\vec{x},\vec{y}\rangle}_{} := \sum_{i=1}^{n} \overline{x_i}\, y_i}$. Then the equation above becomes

recall the Chapter "The inner product" from Calc. Lin. Alg. I.

$$\underbrace{\langle\vec{x}, A\vec{y}\rangle = \langle A^{+}\vec{x}, \vec{y}\rangle}.$$

"We can move a matrix to the other side of the scalar product if we dagger it."

Now for the discussion of eigenvectors an interesting type of matrices is the following:

$$\boxed{\text{\underline{Definition:} An } n\times n \text{ matrix } A \text{ is called } \boxed{\text{normal}} \text{ if } AA^{+} = A^{+}A.}$$

Examples:

- $A = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix}$, then $A^{+} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}$, and we have:

  $\hookrightarrow AA^{+} = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix}\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} = \begin{pmatrix} -i^2+1 & -i-i \\ i+i & 1-i^2 \end{pmatrix} \overset{i^2=-1}{=} \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix}$

  $\hookrightarrow A^{+}A = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} = \begin{pmatrix} -i^2+1 & -i-i \\ i+i & 1-i^2 \end{pmatrix} = \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix}$

  $\Rightarrow$ A is a normal matrix

- $A = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}$, then $A^{+} = \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix}$, and we have:

  $\hookrightarrow AA^{+} = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}\begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix} = \begin{pmatrix} -i^2+1 & -i \\ i & -i^2 \end{pmatrix} = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix}$

  $\hookrightarrow A^{+}A = \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix}\begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix} = \begin{pmatrix} -i^2 & -i \\ i & 1-i^2 \end{pmatrix} = \begin{pmatrix} 1 & -i \\ i & 2 \end{pmatrix}$

  $\Rightarrow$ A is <u>not</u> a normal matrix

Two remarks:

- If A is normal and invertible, then A^{-1} is also normal.

  (Why? $(A^{-1})(A^{-1})^{+} = A^{-1}(A^{+})^{-1} = (A^{+}A)^{-1} = (AA^{+})^{-1} = (A^{+})^{-1}(A^{-1}) = (A^{-1})^{+}(A^{-1})$)

• More importantly: If A is normal, then so is $A-\lambda I$ for any $\lambda \in \mathbb{C}$.

$$\text{Why? } (A-\lambda I)^{+}(A-\lambda I) = (A^{+}-\bar{\lambda} I)(A-\lambda I) = \underbrace{A^{+}A}_{=AA^{+} \text{ if } A \text{ is normal}} - \bar{\lambda} A - \lambda A^{+} + \bar{\lambda}\lambda$$

$$= AA^{+} - \bar{\lambda} A - \lambda A^{+} + \bar{\lambda}\lambda = (A-\lambda I)(A-\lambda I)^{+}$$

From the second remark we see the following:

If λ is an eigenvalue, and $\vec{x}$ a corresponding eigenvector, then $(A-\lambda I)\vec{x} = \vec{0}$.

If A is normal, we have

$$0 = \langle \vec{x}, (A-\lambda I)^{+}\underbrace{(A-\lambda I)\vec{x}}_{=\vec{0}} \rangle \overset{(A-\lambda I) \text{ normal}}{=} \langle \vec{x}, (A-\lambda I)(A-\lambda I)^{+}\vec{x} \rangle \overset{\text{as noted above}}{=} \langle (A-\lambda I)^{+}\vec{x}, (A-\lambda I)^{+}\vec{x} \rangle,$$

so $(A-\lambda I)^{+}\vec{x} = \vec{0}$, i.e., $A^{+}\vec{x} = \bar{\lambda}\vec{x}$, so $\bar{\lambda}$ is an eigenvalue of A^{+}.

To summarize: **If a normal matrix A has an eigenvalue λ, then A^{+} has an eigenvalue $\bar{\lambda}$.**

Now suppose we have two eigenvectors $\vec{x}_i$ and $\vec{x}_j$ corresponding to two distinct eigenvalues λ_i and λ_j, i.e., $A\vec{x}_i = \lambda_i \vec{x}_i$ and $A\vec{x}_j = \lambda_j \vec{x}_j$ with $\lambda_i \neq \lambda_j$.

Then $\langle \vec{x}_i, A\vec{x}_j \rangle = \langle \vec{x}_i, \lambda_j \vec{x}_j \rangle$, but also $\langle \vec{x}_i, A\vec{x}_j \rangle = \langle A^{+}\vec{x}_i, \vec{x}_j \rangle = \langle \bar{\lambda}_i \vec{x}_i, \vec{x}_j \rangle$

$$= \lambda_i \langle \vec{x}_i, \vec{x}_j \rangle.$$

So $0 = \lambda_j \langle \vec{x}_i, \vec{x}_j \rangle - \lambda_i \langle \vec{x}_i, \vec{x}_j \rangle = \underbrace{(\lambda_j - \lambda_i)}_{\neq 0 \text{ by assumption}} \langle \vec{x}_i, \vec{x}_j \rangle$, so $\langle \vec{x}_i, \vec{x}_j \rangle$ must be zero, i.e.,

$\vec{x}_i$ and $\vec{x}_j$ are orthogonal to each other!

Since eigenvectors are still eigenvectors when we multiply them with a number, we can choose $\vec{x}_i$ and $\vec{x}_j$ to be normalized, i.e., $\vec{x}_i$ and $\vec{x}_j$ are $\underbrace{\text{orthonormal}}_{\text{meaning they are orthogonal (scalar product = 0) and normalized (norm = 1)}}$.

Now one can show (but we omit the details here) that this even works when some eigenvalues have algebraic multiplicities bigger than 1!

In fact, the following is true:

Theorem: An $n \times n$ matrix A is normal if and only if it is diagonalizable with orthonormal eigenvectors, i.e., $\underbrace{\Lambda}_{\text{matrix with eigenvalues on diagonal and 0 otherwise}} = V^{-1} A V$ where V has orthonormal columns.

We have proven this above for the special case when all eigenvalues are distinct, but proving the general case is more difficult.

Examples:

- Let us connect this to the simple example of $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ from the beginning.

  Is A normal? Here, $A^{+} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A$, so surely $A^{+}A = AA = AA^{+}$, i.e., A is normal.

  We already showed that $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $-\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are the two normalized eigenvectors.

  Indeed, they are orthonormal, since $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{2}\left(1 \cdot 1 + 1 \cdot (-1)\right) = \frac{1}{2}(1-1) = 0.$

- Last time we also discussed that the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{pmatrix}$ is diagonalizable.

  Is it normal? We check: $\cdot$ $AA^{+} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 6 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 37 & 13 \\ 1 & 13 & 5 \end{pmatrix}$

$$\cdot\ A^{\dagger}A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 6 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 8 \\ 0 & 8 & 40 \end{pmatrix}$$

So A is not normal. So the eigenvector cannot all be orthonormal. We found

$$E_{\lambda_0=0} = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\},\ E_{\lambda_+=4} = \text{span}\left\{\begin{pmatrix} \frac{1}{2} \\ 2 \\ 1 \end{pmatrix}\right\},\ E_{\lambda_-=-1} = \text{span}\left\{\begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix}\right\}.$$

These are clearly not orthogonal to each other.