The Shur Decomposition

Recall that the Jordan canonical form of a matrix A allows to write it similar to a direct sum of Jordan blocks (and thus similar to an upper triangular). The Schur decomposition is a unitarily equivalent form of the matrix, thus preserving norms (although not as in the specified form of Jordan blocks).

<u>Definition</u>. A matrix $A \in M_n(\mathbb{C})$ is called:

- unitary, if A is invertible with $A^{-1} = A^*$.
- normal, if $A^*A = AA^*$.
- selfadjoint, if $A = A^*$.

<u>Theorem</u>. Let $A \in M_n(\mathbb{C})$. Then there is a unitary $U \in M_n(\mathbb{C})$ and an upper triangular $T \in M_n(\mathbb{C})$ such that: $T = U^*AU$.

<u>Proof</u>. Let λ_1 be an eigenvalue of A and let x_1 be one of the associated eigenvectors. Hence $\exists\, i \in \{1, .., n\}$ such that $x_1(i) \neq 0$. (i-th coordinate)

Let P_1 be a change of basis such that $e_i \to e_1$

Then P_1 is unitary and λ_1 is an eigenvalue

of $P_1^* A P_1$ with eigenvector $P_1 x_1 =: v_1$.
Note now that $v_1(1) = x_1(i) \neq 0$. Wlog we normalise so that $\|v_1\|_2 = \|x_1\|_2 = 1$.
Consider the matrix:

$$R_1 = I - 2 v_1 v_1^* \in M_n(\mathbb{C})$$

Using that $v_1^* v_1 = \|v_1\|_2^2 = 1$ we can see that R_1 is a unitary with $R_1^* = R_1^{-1} = R_1$ and moreover $R_1 v_1 = \lambda_1 e_1$ $(\Rightarrow \lambda_1 R_1 e_1 = v_1)$

Set $Q_1 = P_1 R_1$ which is a unitary and by construction we have:

$$Q_1^* A Q_1 = R_1 P_1^* A P_1 R_1 = \left[\begin{array}{c|ccc} \lambda_1 & * & & * \\ \hline 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{array}\right]$$

Apply now the above for A_2 with eigenvalue λ_2 to find $Q_2' \in M_{n-1}(\mathbb{C})$ unitary s.t.

$$(Q_2')^* A_2 Q_2' = \left[\begin{array}{c|ccc} \lambda_2 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A_3 & \\ 0 & & & \end{array}\right].$$

Set $Q_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q_2' & \\ 0 & & & \end{bmatrix} \in M_n(\mathbb{C})$ which is a unitary and by construction:

$$(Q_1 Q_2)^* A Q_1 Q_2 = Q_2^* Q_1^* A Q_1 Q_2 =$$

$$= Q_2^* \begin{bmatrix} \lambda_1 & * \\ * & A_2 \end{bmatrix} Q_2 = \begin{bmatrix} \lambda_1 & * \\ 0 & (Q_2')^* A_2 Q_2' \end{bmatrix}$$

$$= \left[\begin{array}{cc|cc} \lambda_1 & * & * \cdots & * \\ 0 & \lambda_2 & * \cdots & * \\ \hline 0 & & & \\ \vdots & 0 & A_3 & \\ 0 & & & \end{array}\right]$$

Continuing inductively we find $Q_1 \dots Q_n$ unitaries such that:

$$(Q_1 \cdots Q_n)^* A (Q_1 \cdots Q_n) = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ \vdots & & \ddots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

and $Q_1 \cdots Q_n$ is unitary. $\boxtimes$

Note here that if $Q^* A Q = T$ for Q unitary and T upper triangular then the diagonal entries of T are its eigenvalues and they are also the eigenvalues of A.

This appears also in the construction above.

<u>Corollary</u>. Let $A \in M_n(\mathbb{C})$. If $A = A^*$ then there is a unitary $U \in M_n(\mathbb{C})$ such that U^*AU is diagonal.

<u>Proof</u>. By Theorem there is unitary U and an upper triangular T such that $U^*AU = T$.

Then: $T^* = (U^*AU)^* = U^*A^*U = U^*AU = T$.

Hence T is upper and lower triangular, thus it is diagonal. $\square$

<u>Corollary</u>. Let $A \in M_n(\mathbb{C})$. The following are equivalent: (1) A is normal

(2) there is a unitary U such that U^*AU is diagonal.

<u>Proof</u>. (1) $\Rightarrow$ (2): Let U be a unitary and T be upper triangular so that $U^*AU = T$.

Since A is normal then T is normal:

$$\left.\begin{aligned} T^*T &= U^*A^*UU^*AU = U^*A^*AU \\ TT^* &= U^*AUU^*A^*U = U^*AA^*U \end{aligned}\right\} \underline{A^*A = AA^*}$$

$$T^*T = TT^*.$$

The proof then follows by the next claim.

<u>Claim</u> If $T \in M_n(\mathbb{C})$ is upper triangular and normal, then T is diagonal.

<u>proof</u>: For $n=2$ take $T = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ Then we have:

$$\begin{pmatrix} |a|^2 & \bar{a}b \\ \bar{b}a & |d|^2+|b|^2 \end{pmatrix} = T^*T = TT^* = \begin{pmatrix} |a|^2+|b|^2 & b\bar{d} \\ d\bar{b} & |d|^2 \end{pmatrix}$$

By equating the $(1,1)$ entries we have $|b|^2 = 0$ and thus $b=0$.

Suppose the statement holds for any $k \le n$ and we will show it holds for $n+1$. We can write

$$T = \begin{bmatrix} a & x \\ 0_{n,1} & S \end{bmatrix} \quad \text{where: } x \in M_{1,n}(\mathbb{C}),\ a \in \mathbb{C}$$
$$S \in M_n(\mathbb{C}).$$

Then we have:

$$\begin{pmatrix} |a|^2 & \bar{a}x \\ x^* a & S^*S + x^*x \end{pmatrix} = T^*T = TT^* = \begin{pmatrix} |a|^2 + xx^* & xS^* \\ Sx^* & SS^* \end{pmatrix}$$

By equating the $(1,1)$ entries we get:

$$\sum_{i=1}^{n} |x_i|^2 = xx^* = 0 \Rightarrow x_i = 0\ \forall i \Rightarrow x = 0.$$

By equating the $(2,2)$ entries and using $x=0$ we get that S is upper triangular and normal in $M_n(\mathbb{C})$. Hence it is diagonal by the inductive hypothesis and we are done ☒

We can say more for <u>real</u> matrices.

<u>Theorem</u> Let $A \in M_n(\mathbb{R})$. The following are equivalent:

(1) $A = A^t$ (where t is for transpose)

(2) there exists $P \in M_n(\mathbb{R})$ with $P^{-1} = P^t$ and $D \in M_n(\mathbb{R})$ diagonal such that $P^t A P = D$.

<u>proof</u>: (1) $\Rightarrow$ (2): Follows with similar ideas to the spectral theorem for complex normal matrices but with some care. The key idea is the following:

See A in $M_n(\mathbb{C})$ and let λ be an eigenvalue with eigenvector v. Notice that $A = A^t$ implies $A = A^*$ in $M_n(\mathbb{C})$ as A has real entries. Then:

$$\left.\begin{array}{l} v^* A v = v^*(\lambda v) = \lambda \|v\|_2^2 \\ v^* A^* v = (Av)^* v = (\lambda v)^* v = \bar{\lambda} \|v\|_2^2 \end{array}\right\} \xRightarrow{v \neq 0} \lambda = \bar{\lambda}.$$

Hence $\lambda \in \mathbb{R}$ and this imposes $v \in M_{n,1}(\mathbb{R})$.

Consider the space $\{v\}^\perp = \{x \in \mathbb{R}^n : x^t v = 0\}$ and find an orthonormal basis $\{v_2, \dots, v_n\}$.

Then $\{v, v_2, \dots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$ which implies that the matrix

$$P = [v; v_2; \dots; v_n]$$ is orthogonal, i.e. $P^{-1} = P^t$.

We now see by construction of P that:

$$P^t A P e_1 = P^t A v = \lambda P^t v = \lambda P^{-1} v = \lambda e_1$$

and thus:

$$P^t A P = \left[\begin{array}{c|ccc} \lambda & a_2 & \cdots & a_n \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$$

Since $(P^t A P)^t = P^t A^t (P^t)^t = P^t A P$ we have that $a_2 = \cdots = a_n = 0$. Hence we have:

$$P^t A P = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}$$

Moreover $(P^t A P)^t = P^t A P$ implies $B = B^t$. Then the proof finishes by induction. ☒

<u>Remark</u>. Note that if $A \in M_n(\underline{\mathbb{C}})$ and $A = \underline{A^t}$ then A may <u>not</u> be <u>orthogonally</u> diagonalisable i.e., having $P \in M_n(\mathbb{R})$ with $P^{-1} = P^t$ and D diagonal with $P^t A P = D$. If that were true then P would consist of the eigenvectors of A. Take $A = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ for a counterexample.

Remark It is not true the following for $A \in M_n(\mathbb{R})$:

$A = A^t \Leftrightarrow A$ is diagonalisable $\Leftrightarrow A$ is normal

where is the error? (one direction is ok) $AA^t = A^t A$.

For a counterexample take $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Then $A \neq A^t$ but $AA^t = A^t A$.

However if we assume that $A \in M_n(\mathbb{R})$ and it has real eigenvalues then the above equivalence works.

For the counterexample $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ note that it has $a \pm ib$ as eigenvalues.