Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

September 27, 2021 (unfinished!)

Contents

1. Schur triangularization (cont'd) **1**
1.1. Schur triangularization 1

Math 504 Lecture 4

1. Schur triangularization (cont'd)

1.1. Schur triangularization

We are now ready for one more matrix decomposition, the so-called **Schur triangularization** (aka **Schur decomposition**):

Theorem 1.1.1 (Schur triangularization theorem)**.** Let $A \in \mathbb{C}^{n\times n}$. Then, there exists a unitary matrix $U \in \mathrm{U}_n(\mathbb{C})$ and an upper-triangular matrix $T \in \mathbb{C}^{n\times n}$ such that $A = UTU^*$.

In other words, A is unitary similar to some upper-triangular matrix.

Example 1.1.2. Let $A = \begin{pmatrix} 1 & 3 \\ -3 & 7 \end{pmatrix}$. Then, a Schur triangularization of A is $A = UTU^*$, where

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad \text{and} \qquad T = \begin{pmatrix} 4 & 6 \\ 0 & 4 \end{pmatrix}.$$

*Drexel University, Korman Center, 15 S 33rd Street, Philadelphia PA, 19104, USA

Proof. Induction on n.

Base case ($n = 0$) is trivial.

Induction step: (From $n - 1$ to n:)

Suppose that we have proved the theorem for $(n-1) \times (n-1)$-matrices.

Let us now prove it for $n \times n$-matrices.

So let $A \in \mathbb{C}^{n \times n}$. Since $n > 0$, this matrix A has at least one eigenvalue (by the FTA).

Fix some eigenvalue λ of A, and let $x \neq 0$ be an eigenvector for this eigenvalue.

Let $u_1 = \dfrac{1}{||x||} x$. Choose vectors $u_2, u_3, \ldots, u_n$ such that $(u_1, u_2, \ldots, u_n)$ is an orthonormal basis of $\mathbb{C}^n$. (We can do this, by a lemma from the lectures before.)

Let U be the matrix with columns $u_1, u_2, \ldots, u_n$. Then, U is unitary (by Theorem 1.5.3 $\mathcal{E} \Longrightarrow \mathcal{A}$). Hence, U^* is unitary.

Now, by standard properties of matrix multiplication, we have

$$(AU)_{\bullet,1} = A \underbrace{U_{\bullet,1}}_{=u_1} = Au_1 = \lambda u_1$$

(since u_1, being a scalar multiple of x, is an eigenvector of A for eigenvalue λ). Thus,

$$(U^* AU)_{\bullet,1} = U^* (AU)_{\bullet,1} = U^* \lambda u_1 = \lambda U^* \underbrace{u_1}_{=U_{\bullet,1}}$$

$$= \lambda \underbrace{U^* U_{\bullet,1}}_{=(U^* U)_{\bullet,1}} = \lambda \left(\underbrace{U^* U}_{=I_n} \right)_{\bullet,1} = \lambda \left(I_n\right)_{\bullet,1}.$$

In other words,

$$U^* AU = \begin{pmatrix} \lambda & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \cdots & * \end{pmatrix},$$

where the asterisk ($*$) means an entry that you don't know or don't care about.

Let us write this in block-matrix notation:

$$U^* AU = \begin{pmatrix} \lambda & p \\ 0 & B \end{pmatrix},$$

where $p \in \mathbb{C}^{1 \times (n-1)}$ is a row vector and $B \in \mathbb{C}^{(n-1) \times (n-1)}$ is a matrix. (The "0" here is actually the zero vector in $\mathbb{C}^{n-1}$.)

Now, by the induction hypothesis, B is unitary similar to an upper-triangular matrix. In other words, there is a unitary matrix $V \in \mathbb{C}^{(n-1) \times (n-1)}$ and an upper-triangular S such that $VBV^* = S$. Consider these V and S.

Now, let $W = \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$. This is a block-diagonal matrix, with $\begin{pmatrix} 1 \end{pmatrix}$ and V being its diagonal blocks. Hence, $W^* = \begin{pmatrix} 1 & 0 \\ 0 & V^* \end{pmatrix}$. Moreover, W is a unitary matrix (since it is a block-diagonal matrix whose diagonal blocks are unitary). Thus, WU^* is unitary (being the product of the two unitary matrices W and U^*).

Now,

$$\begin{aligned}
&\underbrace{W}_{=\begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}} \underbrace{U^*AU}_{=\begin{pmatrix} \lambda & p \\ 0 & B \end{pmatrix}} \underbrace{W^*}_{=\begin{pmatrix} 1 & 0 \\ 0 & V^* \end{pmatrix}} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \lambda & p \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & V^* \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot \lambda \cdot 1 & 1 \cdot p \cdot V^* \\ V \cdot 0 \cdot 1 & V \cdot B \cdot V^* \end{pmatrix} = \begin{pmatrix} \lambda & pV^* \\ 0 & VBV^* \end{pmatrix} = \begin{pmatrix} \lambda & pV^* \\ 0 & S \end{pmatrix}.
\end{aligned}$$

This matrix is upper-triangular (since S is upper-triangular). However, WU^*AUW^* is unitary similar to A, because

$$WU^*A \underbrace{UW^*}_{\substack{=U^{**}W^* \\ =(WU^*)^*}} = \underbrace{WU^*}_{\text{unitary}} A \left(WU^*\right)^*.$$

Thus, we have found an upper-triangular matrix (namely, WU^*AUW^*) that is unitary similar to A. Qed. $\square$

So much for making a single matrix triangular.

Can we make a whole bunch of matrices triangular simultaneously (using the same unitary U)?

Recall that two square matrices A and B are said to *commute* if $AB = BA$. For example, any two powers of a single matrix commute (because if A is a square matrix, then $A^k A^\ell = A^\ell A^k$).

Lemma 1.1.3. Let $n > 0$. Let $\mathcal{F}$ be a subset of $\mathbb{C}^{n \times n}$. Assume that any two matrices in $\mathcal{F}$ commute (i.e., for any $A, B \in \mathcal{F}$, we have $AB = BA$).

Then, there exists a nonzero $x \in \mathbb{C}^n$ such that x is an eigenvector of each $A \in \mathcal{F}$.

In short: A family of pairwise commuting matrices always has a common eigenvector.

Proof. An $\mathcal{F}$**-invariant subspace** of $\mathbb{C}^n$ shall mean a vector subspace V of $\mathbb{C}^n$ such that

$$AV \subseteq V \qquad \text{for each } A \in \mathcal{F}.$$

Here,

$$AV := \{Av \mid v \in V\}.$$

For example, $\{0\}$ and $\mathbb{C}^n$ itself are two $\mathcal{F}$-invariant subspaces. There might be further $\mathcal{F}$-invariant subspaces between these two.

[**Example:** If $n = 2$ and $\mathcal{F} = \left\{ \begin{pmatrix} 1 & a \\ 0 & 2 \end{pmatrix} \mid a \in \mathbb{R} \right\}$, then $\operatorname{span}(e_1)$ is an $\mathcal{F}$-invariant subspace.]

Let W be an $\mathcal{F}$-invariant subspace of $\mathbb{C}^n$ that has lowest possible positive dimension.

We will show that each $x \in W$ is an eigenvector of each $A \in \mathcal{F}$.

Indeed, fix any $A \in \mathcal{F}$. Then, $AW \subseteq W$ (since W is $\mathcal{F}$-invariant). Thus, A restricts to a $\mathbb{C}$-linear map $A\mid_W : W \to W$. Since $\dim W > 0$, this $\mathbb{C}$-linear map $A\mid_W$ has an eigenvalue λ and a corresponding nonzero eigenvector $w \neq 0$. So $w \in W$ is a nonzero vector satisfying $Aw = \lambda w$.

Let

$$W_{A,\lambda} := \{ y \in W \mid Ay = \lambda y \}.$$

This $W_{A,\lambda}$ is a vector subspace of W. It has positive dimension, because it contains the nonzero vector w. Furthermore, I claim that this subspace $W_{A,\lambda}$ is $\mathcal{F}$-invariant.

[*Proof:* Let $B \in \mathcal{F}$ be arbitrary. We must prove that $BW_{A,\lambda} \subseteq W_{A,\lambda}$. In other words, we must prove that $Bz \in W_{A,\lambda}$ for each $z \in W_{A,\lambda}$.

Indeed, let $z \in W_{A,\lambda}$. Then, we must show that $Bz \in W_{A,\lambda}$. It is clear that $Bz \in W$, since $z \in W_{A,\lambda} \subseteq W$ and because W is $\mathcal{F}$-invariant. Furthermore, we have $ABz = \lambda Bz$ because

$$\underbrace{AB}_{=BA} z = B \underbrace{Az}_{\substack{=\lambda z \\ (\text{since } z \in W_{A,\lambda})}} = B\lambda z = \lambda Bz.$$

So we conclude that $Bz \in W_{A,\lambda}$. This shows that $W_{A,\lambda}$ is $\mathcal{F}$-invariant.]

So $W_{A,\lambda}$ is an $\mathcal{F}$-invariant subspace of $\mathbb{C}^n$ of positive dimension that is a subspace of W. However, W is an $\mathcal{F}$-invariant subspace of smallest positive dimension. Thus, $W_{A,\lambda}$ must have the same dimension as W. Hence, $W_{A,\lambda} = W$ (since a subspace of W having the same dimension as W must just be W itself). This shows that any vector in W is an eigenvector of A (since it belongs to $W_{A,\lambda} = \{ y \in W \mid Ay = \lambda y \}$).

Forget that we fixed A. We thus have shown that any vector in W is an eigenvector of each $A \in \mathcal{F}$. Since W has positive dimension, there exists some nonzero vector in W. Thus, this vector is an eigenvector of each $A \in \mathcal{F}$. This proves the lemma. □

Theorem 1.1.4 (Schur triangularization for many commuting matrices)**.** Let $\mathcal{F}$ be a subset of $\mathbb{C}^{n \times n}$. Assume that any two matrices in $\mathcal{F}$ commute (i.e., for any $A, B \in \mathcal{F}$, we have $AB = BA$).

Then, there exists a unitary $n \times n$-matrix U such that

$$UAU^* \text{ is upper-triangular for all } A \in \mathcal{F}.$$

Proof. Same induction as for the previous theorem. But now, instead of picking an eigenvector of a single matrix A, we pick a common eigenvector for all matrices in $\mathcal{F}$. The existence of such an eigenvector is guaranteed by the preceding lemma. $\square$

Note that the theorem has no converse. Indeed, it is well possible that a set $\mathcal{F}$ of matrices can be simultaneously triangularized by one and the same unitary matrix U, but these matrices do not pairwise commute. For example, $\mathcal{F} = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right\}$.