Advanced Linear Algebra – Week 7

Schur Triangularization and the Spectral Decomposition(s)

This week we will learn about:

- Schur triangularization,
- The Cayley–Hamilton theorem,
- Normal matrices, and
- The real and complex spectral decompositions.

Extra reading and watching:

- Section 2.1 in the textbook
- Lecture videos 25, 26, 27, 28, and 29 on YouTube
- Schur decomposition at Wikipedia
- Normal matrix at Wikipedia
- Spectral theorem at Wikipedia

Extra textbook problems:

⋆ 2.1.1, 2.1.2, 2.1.5

⋆⋆ 2.1.3, 2.1.4, 2.1.6, 2.1.7, 2.1.9, 2.1.17, 2.1.19

⋆⋆⋆ 2.1.8, 2.1.11, 2.1.12, 2.1.18, 2.1.21

☠ 2.1.22, 2.1.26

We're now going to start looking at **matrix decompositions**, which are ways of writing down a matrix as a product of (hopefully simpler!) matrices. For example, we learned about diagonalization at the end of introductory linear algebra, which said that...

While diagonalization let us do great things with certain matrices, it also raises some new questions:

Over the next few weeks, we will thoroughly investigate these types of questions, starting with this one:

Schur Triangularization

We know that we cannot hope in general to get a diagonal matrix via unitary similarity (since not every matrix is diagonalizable via *any* similarity). However, the following theorem says that we can get partway there and always get an upper triangular matrix.

Theorem 7.1 — Schur Triangularization

Suppose $A \in \mathcal{M}_n(\mathbb{C})$. Then there exists a unitary matrix $U \in \mathcal{M}_n(\mathbb{C})$ and an upper triangular matrix $T \in \mathcal{M}_n(\mathbb{C})$ such that

Proof. We prove the result by induction on n (the size of A). For the base case, we simply notice that the result is trivial if $n = 1$: every 1×1 matrix is upper triangular.

■

Let's make some notes about Schur triangularizations before proceeding...

- The diagonal entries of T are the eigenvalues of A. To see why, recall that the eigenvalues of a triangular matrix are its diagonal entries (theorem from previous course), and...

- The other pieces of Schur triangularization are

- To compute a Schur decomposition, follow the method given in the proof of the theorem:

The beauty of Schur triangularization is that it applies to *every* square matrix (unlike diagonalization), which makes it very useful when trying to prove theorems. For example...

Theorem 7.2 — Trace and Determinant in Terms of Eigenvalues

Suppose $A \in \mathcal{M}_n(\mathbb{C})$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then

Proof. Use Schur triangularization to write $A = UTU^*$ with U unitary and T upper triangular. Then...

∎

As another application of Schur triangularization, we prove an important result called the Cayley–Hamilton theorem, which says that every matrix satisfies its own characteristic polynomial.

Theorem 7.3 — Cayley–Hamilton

Suppose $A \in \mathcal{M}_n(\mathbb{C})$ has characteristic polynomial $p(\lambda) = \det(A - \lambda I)$. Then $p(A) = O$.

For example...

Proof of Theorem 7.3. Because we are working over $\mathbb{C}$, the Fundamental Theorem of Algebra says that we can factor the characteristic polynomial as a product of linear terms:

Well, let's Schur triangularize A:

■

One useful feature of the Cayley–Hamilton theorem is that if $A \in \mathcal{M}_n(\mathbb{C})$ then it lets us write every power of A as a linear combination of $I, A, A^2, \ldots, A^{n-1}$. In particular,

Example. *Use the Cayley–Hamilton theorem to come up with a formula for A^4 as a linear combination of A and I, where*

$$A =$$

Example. *Use the Cayley–Hamilton theorem to find the inverse of the same matrix.*

Normal Matrices and the Spectral Decomposition

We now start looking at when Schur triangularization actually results in a *diagonal* matrix, rather than just an upper triangular one. We first need to introduce another new family of matrices:

Definition 7.1 — Normal Matrix

A matrix $A \in \mathcal{M}_n(\mathbb{C})$ is called **normal** if $A^*A = AA^*$.

Many of the important families of matrices that we are already familiar with are normal. For example...

However, there are also other matrices that are normal:

Example. *Show that the matrix* $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ *is normal.*

Our primary interest in normal matrices comes from the following theorem, which says that normal matrices are exactly those that can be diagonalized by a unitary matrix:

Theorem 7.4 — Complex Spectral Decomposition

Suppose $A \in \mathcal{M}_n(\mathbb{C})$. Then there exists a unitary matrix $U \in \mathcal{M}_n(\mathbb{C})$ and diagonal matrix $D \in \mathcal{M}_n(\mathbb{C})$ such that

if and only if A is normal (i.e., $A^*A = AA^*$).

In other words, normal matrices are the ones with a diagonal Schur triangularization.

Proof. To see the "only if" direction, we just compute

∎

While we proved the spectral decomposition via Schur triangularization, that is not how it is computed in practice. Instead, we notice that the spectral decomposition is a special case of diagonalization where the invertible matrix that does the diagonalization is unitary, so we compute it via eigenvalues and eigenvectors (like we did for diagonalization last semester). Just be careful to choose the eigenvectors to have length 1 and be mutually orthogonal.

Example. *Find a spectral decomposition of the matrix...*

Example. *Find a spectral decomposition of the matrix*

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

Sometimes, we can just "eyeball" an orthonormal set of eigenvectors, but if we can't, we can instead apply the Gram–Schmidt process to any basis of the eigenspace.

The Real Spectral Decomposition

In the previous example, the spectral decomposition ended up making use only of real matrices. We now note that this happened because the original matrix was symmetric:

Theorem 7.5 — Real Spectral Decomposition

Suppose $A \in \mathcal{M}_n(\mathbb{R})$. Then there exists a unitary matrix $U \in \mathcal{M}_n(\mathbb{R})$ and diagonal matrix $D \in \mathcal{M}_n(\mathbb{R})$ such that

if and only if A is symmetric (i.e., $A^T = A$).

To give you a rough idea of why this is true, we note that every Hermitian (and thus every symmetric) matrix has real eigenvalues:

It follows that if A is Hermitian then we can choose the "D" piece of the spectral decomposition to be real. Also, it should not be too surprising, that if A is *real* and Hermitian (i.e., symmetric) that we can choose the "U" piece to be real as well.

We thus get the following 3 types of spectral decompositions for different types of matrices:

Geometrically, the real spectral decomposition says that real symmetric matrices are exactly those that act as follows: