Section 5.4: Similarity transformations, Hermitian and real symmetric matrices.

This chapter is about constructing matrix factorizations based on eigenvalues and eigenvectors, and seeing how adding properties to a matrix (like semisimple or symmetric) adds properties to the factorizations.

We need to talk about complex matrices; mainly because, unlike all of the other properties of real matrices and all the other factorizations that we have thus far studied, eigenvalues of a real matrix (and the corresponding eigenvectors) need not be real. But eigenvalues and eigenvectors all exist over $\mathbb{C}$ and, respectively, $\mathbb{C}^n$, and so our “generic” matrices will also be in $\mathbb{C}^{n \times n}$.

But first, let us examine one tool that will explain all the others: the similarity transformation.

**Definition** A similarity transformation is easy enough to state: given square $n \times n$ matrices $A, B$ with entries that are either real or complex, and an invertible $n \times n$ matrix $S$, with the property that

$$B = S^{-1}AS$$

then we say that $A$ and $B$ are similar, via the similarity transformation $B = S^{-1}AS$.

**Remark 1.** Note that if $A$ is similar to $B$ via $S$, then $B$ is similar to $A$ via $S^{-1}$: $B = S^{-1}AS$ can easily be rewritten as $A = SBS^{-1}$.

The deeper meaning of similarity transformations is this: one can see each square matrix $A$ as the expression, in the standard basis, of a linear transformation of the space. Recall that linear transformations of $\mathbb{R}^n$ into itself (or $\mathbb{C}^n$ into itself) are functions with the property that $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$, for all $\alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$) and $x, y \in \mathbb{R}^n$ (or $\mathbb{C}^n$).

Suppose that one would like to change basis, from a different basis to the standard one. Given the change-of-basis matrix $S$, in the other basis, the linear transformation has matrix $B = S^{-1}AS$. But the underlying linear transformation is the same, therefore the eigenvalues of $A$ should be the same as the eigenvalues of $B$. We can see this also using determinants.

**Lemma 1.** $A$ and $S^{-1}AS$ have the same eigenvalues with the same multiplicities.

**Proof.** We know that the eigenvalues of $A$ are the roots of the characteristic polynomial (this takes care of multiplicities, too). But

$$\det(\lambda I - S^{-1}AS) = \det(S^{-1}(\lambda S - AS)) = \det(S^{-1}(\lambda I - A)S) = \det S^{-1} \cdot \det(\lambda I - A) \cdot \det(S) = \det(\lambda I - A),$$

since $\det S^{-1} \cdot \det S = 1$. The conclusion follows.

Moreover, there is a simple, trivial mapping of an eigenvector of $A$ into one of $B = S^{-1}AS$.

**Lemma 2.** If $(\lambda, v)$ is an eigenpair for $A$, then $(\lambda, S^{-1}v)$ is an eigenpair for $B = S^{-1}AS$. 

Proof. Let \( w = S^{-1}v \); then
\[
Bw = BS^{-1}v = S^{-1}Av = S^{-1}\lambda v = \lambda S^{-1}v = \lambda w ,
\]
where we have used that \( B = S^{-1}AS \) means that \( BS^{-1} = S^{-1}A \) (by simple multiplication of the equation with \( S^{-1} \) to the right.) The conclusion follows.

Similarity is a very useful concept; it allows us to see that most matrices are diagonalizable (recall that we argued that most matrices are semisimple).

**Lemma 3.** Let \( A \) be an \( n \times n \) matrix. Then \( A \) is semisimple if there exist \( V \) an \( n \times n \) invertible, and \( D \) an \( n \times n \) diagonal matrix, such that \( A = VD^{-1} \).

Proof. \( \Rightarrow \) Suppose \( A \) is semisimple, then it has a set of \( n \) linearly independent vectors \( v_1, \ldots, v_n \), with corresponding eigenvalues \( \lambda_1, \ldots, \lambda_n \). Let \( V = [v_1 \ldots v_n] \), the matrix whose columns are the \( v_i \)'s. As the \( v_i \)'s are independent, \( V \) is invertible. Let \( D \) be the diagonal matrix with diagonal entries \( \lambda_1, \ldots, \lambda_n \).

Since \( Av_i = \lambda_i v_i \) for all \( i \), putting all these together we get that \( AV = VD \), which leads to \( A = VDV^{-1} \).

\( \Leftarrow \) If \( A = VDV^{-1} \), let \( V = [v_1, \ldots, v_n] \) and label the diagonal entries of \( D \) by \( \lambda_1, \ldots, \lambda_n \). The equality shows that \( AV = VD \) and hence \( Av_i = \lambda_i v_i \), thus \( (\lambda_i, v_i) \) is an eigenpair, and since \( V \) is invertible, \( v_1, \ldots, v_n \) are linearly independent, so \( A \) is semisimple.

Hence, any semisimple matrix is similar to the diagonal matrix of its eigenvalues (with multiplicities).

This is an important result; it shows that since most matrices are semisimple, most matrices are similar to the diagonal matrix of their eigenvalues, and the similarity transformation simply involves finding the basis of their eigenvectors.

But not all matrices are semisimple. For matrices who are not, might it be possible to find some other kind of structured matrix with easily “read” eigenvalues which is similar to the original matrix? It turns out the answer is yes.

**Theorem 1.** (due to Issai Schur). For any matrix \( A \in \mathbb{C}^{n \times n} \), there exist an \( n \times n \) unitary matrix \( U \) and an \( n \times n \) upper triangular matrix \( T \) such that \( A = UTU^* \).

**Remark 2.** When \( A \) is real, the theorem above has an all-real formulation. One may pick \( Q \) real and orthogonal, but instead of the all-real \( T \) being upper triangular, it is almost upper triangular: the diagonal of \( T \) consists of \( 1 \times 1 \) and \( 2 \times 2 \) blocks (with all entries below being 0); and \( A = QTQ^T \).

The Schur Theorem has a couple of very consequential corollaries.
**Corollary 1.** Let $A$ be an $n \times n$ Hermitian matrix (for which $A = A^*$). Then $A$ has all real eigenvalues, and a complete set of orthogonal eigenvectors.

**Proof.** From Schur’s theorem we see $A = UTU^*$ for a unitary $U$ and a upper triangular $T$. Now, taking the $*$ operation, we obtain

$$A^* = (UTU^*)^* = (U^*)^* T^* U^* = UT^* U^* .$$

On the other hand, $A$ being Hermitian, $A = A^*$ implies $UTU^* = UT^* U^*$, and as $U$ is unitary and invertible, one must have $T = T^*$. But $T$ is upper triangular, while $T^*$ is lower triangular. Hence $T$ is diagonal, and its diagonal entries satisfy $\lambda = \lambda^*$, or in other words, $\lambda$ is real. As the first lemma of this section shows, the diagonal entries of $T$ are the eigenvalues of $A$, so the eigenvalues of $A$ are real. Moreover, $U$ contains the set of eigenvectors, and since $U$ is orthogonal, the eigenvectors are all orthogonal. \[\Box\]

**Remark 3.** Note that if a matrix is real symmetric (and thus also complex Hermitian), its eigenvalues will be real, but—as an added bonus—so will be the eigenvectors.

**Normal matrices.** A matrix $A \in \mathbb{C}^{n \times n}$ is called **normal** if $AA^* = A^* A$. Normal matrices are diagonalizable via a unitary similarity transformation, i.e. there exist $U$ unitary and $D$ diagonal such that $A = UDU^*$. What is the difference between normal matrices and Hermitian ones? For normal matrices, the diagonal of $D$, that is, the eigenvalues of $A$, need not be real as for Hermitian matrices.