Math 642:550 — Summer 1999
MTTh 6:15–8:45 PM   Hill 525
Prof. Bumby

Week 3: July 12, 13 & 15. Recall that each Tuesday class begins with a short exam based on homework from the previous week. The July 13 exam will contain problems that can be done based on the problems from sections 5.1 and 5.2.

This week will be devoted to finishing Chapter 5. We were not able to do Section 5.4 last week, so we will begin with that. In the remainder of Chapter 5, matrices will have complex entries and properties of the complex numbers will play an essential role in all work. When we move to chapter 6 next week, matrices will again be taken to have real entries, although the theory can be developed for complex matrices with only a few modifications. Results for real matrices will illustrate the theory well enough without demanding an extra level of care in the computations.

We have seen that real matrices do not always have real eigenvalues, and if an eigenvalue of real matrix isn't real, its corresponding eigenvectors must also fail to be real. The proof is easy: just look at $Ax = \lambda x$. We have also seen that some matrices, the simplest being

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

do not have enough eigenvectors to span the space. In Section 5.5, we show that symmetric matrices (matrices satisfying $A^{\mathrm{T}} = A$) are free from this annoyance. Indeed, not only is there a basis of eigenvectors, but this basis can be chosen to consist of orthogonal vectors. Indeed, since any multiple of eigenvector is also an eigenvector, the vectors in the basis can also be chosen to all have length 1. Putting such a basis into the matrix S that diagonalizes A, means that $S^{\mathrm{T}}S = I$, or $S^{\mathrm{T}} = S^{-1}$. Matrices S with this property are called **orthogonal matrices**, and can be though of as giving **rigid motions** in $\mathbb{R}^n$, since they preserve length and angle (although the matrices describing reflections are orthogonal, so orientation is not preserved unless one also demands $\det S = +1$).

To prove this result, it is helpful to begin by assuming that one has an eigenvector and then show that it must be real, rather than try to prove the *existence* of a real eigenvector directly. This proof includes the same result for **Hermitian matrices**, which are matrices A with complex entries such that A^{T} is the complex conjugate of A. The easiest way to work with Hermitian matrices is to define an operation $A \mapsto A^{\mathrm{H}}$ that takes a matrix to **the complex conjugate of its transpose**. When A is a real matrix, A^{H} is the same as A^{T}, and among complex matrices, Hermitian matrices are characterized by $A^{\mathrm{H}} = A$. The matrices playing the role of rigid motions are the **unitary matrices**, characterized by $U^{\mathrm{H}} = U^{-1}$. A useful class of matrices, containing both Hermitian and unitary matrices are **normal matrices**, characterized by $N^{\mathrm{H}}N = NN^{\mathrm{H}}$.

The main result of Section 5.5 is that every Hermitian matrix can be diagonalized by a unitary matrix. That is, it has a set of eigenvectors that form an orthonormal basis. A proof in the case where the matrix has distinct eigenvalues is given in Section 5.5, and the general case follows from Schur's triangularization theorem, which is result **5R** in Section 5.6. Supplementary notes on this have been prepared, since it is easy to overlook the treatment in the text.

Part of section 5.6 is devoted to a discussion of the claim that *change of basis* is the same as *similarity transformation*. The treatment is so condensed, while attempting to touch on many different points of view, that it may seem more complicated than it really is. The main point is that, while matrices are generally used to represent linear transformations from one vector space to another, if we limit attention to square matrices, we should think of them as representing linear transformations from a vector space to *itself*.

When m by n matrices are first introduced, they are used to describe linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$, i.e., as concrete things that will later be included in the general theory of vector spaces. When the generality is introduced, it is not always made clear that one works with it by producing very specific matrices that express things described abstractly. To see how this is done, let us look more closely at how matrices express linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

The entries in a column used to write an element of $\mathbb{R}^n$ are the coefficients in the unique expression of that element in terms of the standard basis. If we have a linear transformation into $\mathbb{R}^m$, the image of this vector is the same combination of the images of the standard basis. The images of these n vectors, like any other vector in $\mathbb{R}^m$ is written as a column of m numbers. The matrix representing the linear transformation is the matrix with these columns, and the definition of matrix multiplication models the way in which the image of a general vector is found from the image of the basis vectors.

If you want to do the same thing for abstract vector spaces, the first thing to do is to choose bases for the spaces. This allows us to use the coefficients of the representation of a vector in terms of the chosen basis in exactly the same way as the standard basis is used in $\mathbb{R}^m$.

Row operations introduce new bases on the codomain while retaining the same basis on the domain. This means that features of the domain, like the nullspace, continue to be described in terms of the same coordinates. However, row operations cause features of the codomain, like the actual image of the function, to have descriptions in different bases.

So far, so good. Unfortunately, at some point, you may want to think of $\mathbb{R}^m$ as an abstract vector space and choose a basis for it other than the standard basis. The intrinsic properties of a linear transformation should not depend on the basis that was chosen, so this leads to various equivalences between matrices, but there is some difficulty describing all this in words.

In an attempt to clarify matters, a new level of structure should be introduced: a *vector space with a basis* (VSWB). Making the basis part of the structure allows us to distinguish two ways of describing the same space. If the basis is different, then we have different VSWB's even if we feel that we have the same vector spaces. In particular, a change of basis amounts to a composition with the matrix representing the identity linear transformation from the space with one basis to the space with another.

If you have one basis of a space, then to obtain a matrix representing the identity map to that space with a second basis, it is necessary to express the vectors in the first basis in terms of the second. If this information is not directly available, it is still possible to characterize it as the solution of a system of linear equations. In particular, if you have a matrix giving the second basis in terms of the first, you need the inverse of that matrix. When you are setting up a matrix to describe a mapping from a space to itself, change of basis thus leads to a similarity.

This weeks homework is

section	page	problems
5.5	301	6, 7
5.6	315	8, 9, 16
Supp.	B	four exercises
5.R	319	1, 6, 12, 15

Notes. Problem 5.6.16 is a refinement of Problem 5.2.3. In order to diagonalize the matrix with an orthogonal matrix, you need an orthonormal basis of eigenvectors, so a diagonalizing matrix that worked for Problem 5.2.3 may not work here. However, it is still possible to get *essentially* different diagonalizing matrices. The matrix P that appears in part (b) of Problem 5.6.16 describes a projection into the 0-eigenspace of the matrix, so it *should be* independent of the basis. However, it is not immediately obvious when you calculate it in this form that the results will be the same.