SOME METHODS

•To test whether a set S of vectors is a subspace of a vector space:

Determine whether S is nonempty (check to see if $0 \in S$). Determine whether S is closed under scalar multiplication and vector addition.

•To test whether $u \in \mathrm{span}(v_1, \ldots, v_n)$:

Determine whether the equation $a_1 v_1 + \cdots + a_n v_n = u$ has any solution for the a_i's. (Such a vector equation, in an m-dimensional space, generally converts to a system of m linear equations in the n unknowns. This system can then be solved by the methods of Math. 33A.)

•To test whether $v_1, \ldots, v_n$ are linearly dependent:

See whether the equation $a_1 v_1 + \cdots + a_n v_n = 0$ has nontrivial solutions for the a_i's.

•To calculate the coordinate vector $[v]_\beta$, where β consists of $u_1, \ldots, u_n$ in that order:

Solve the equation $a_1 u_1 + \cdots + a_n u_n = v$ for the a_i's.

•To determine whether a function $T : V \to W$ is linear.

Compare $T(u + v)$ with $T(u) + T(v)$ and compare $T(av)$ with $aT(v)$. (For starters, check whether $T(0_V) = 0_W$.)

•To calculate the matrix $[T]_\beta^\gamma$ representing T.

Apply T to the basis vectors of the domain V, then take coordinate vectors relative to the basis for W. This gives the columns of the matrix.

•To solve the equation $T(v) = w$ for v.

Find the representing matrix $A = [T]_\beta^\gamma$ relative to some bases β and γ; solve the equation $Ax = [w]_\gamma$ for x; decoordinatize. (There may be more direct methods, depending on the situation.)

•To find a basis for $\ker T$.

Find a basis for the nullspace of a representing matrix, and decoordinatize it. (There may be more direct methods, depending on the situation.) If all you need is the nullity of T, then all you need is the dimension of the nullspace for the matrix.

•To find a basis for $\mathrm{ran}\, T$.

Find a basis for the column space of a representing matrix, and decoordinatize it. (There may be more direct methods, depending on the situation.) If all you need is the rank of T, then all you need is the rank of the matrix.

•To test whether a linear transformation T is an isomorphism.

Is it one-to-one (what is its nullity)? Is it onto (what is its rank)? Because

$$\text{nullity} + \text{rank} = \text{dimension of the domain},$$

the two questions are related.

•To test whether two vector spaces (over the same field) are isomorphic.

Do they have the same dimension?

•To find the coordinate vector with respect to a new basis.

Forget the old basis, and calculate the coordinate vector as before. Or else use the equation $[v]_\beta = Q^{-1}[v]_\sigma$, where $Q = [id]^\sigma_\beta$, so that column j of Q is $[u_j]_\sigma$, where u_j if the jth basis vector in β. (The latter method gives the connection between the old coordinate vector and the new one.)

•To find the matrix representation of a linear operator with respect to a new basis.

Forget the old basis, and calculate the representing matrix as before. Or else use the equation $[T]^\beta_\beta = Q^{-1}[T]^\sigma_\sigma Q$, where Q is as above. (The latter method gives the connection between the old matrix and the new one.)

Now suppose that we have a linear operator T on an n-dimensional vector space V. In order to understand T, we seek a diagonalization, if one exists.

• 1. Calculate the representing matrix $A = [T]^\sigma_\sigma$ for T, relative to some convenient basis σ, such as the standard basis, or the best basis we know of.

A is an $n \times n$ matrix. If T is L_A, matrix multiplication by A, then of course the representing matrix (relative to the standard basis for $\mathbf{F}^n$) is simply A.

• 2. Find the characteristic polynomial $p(\lambda)$ of T, which is $\det(A - \lambda I)$, and its roots $\lambda_1, \ldots, \lambda_k$. Let m_i be the multiplicity of the root λ_i. (For large n, this step might need to be altered.)

This polynomial does not depend on our choice of basis. It is a polynomial of degree n in the variable λ. The roots $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of T and of A. (A scalar λ is an eigenvalue if and only if the matrix $A - \lambda I$ is singular.)

Over $\mathbb{R}$, we have $p(t) = \pm(t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k} \cdot$(irreducible quadratics). And $m_1 + \cdots + m_k \le n$.

Over $\mathbf{C}$, we have simply $p(t) = \pm(t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$ and $m_1 + \cdots + m_k = n$.

• 3. For each eigenvalue λ_i, find a basis for the nullspace of $A - \lambda_i I$. By decoordinatizing, find a basis for E_{λ_i}.

E_{λ_i} is isomorphic to the nullspace of $A - \lambda_i I$ under the coordinate map $[\ ]_\sigma$. We know that

$$1 \le \dim E_{\lambda_i} \le m_i$$

that is, the "geometric multiplicity" $\dim E_{\lambda_i}$ does not exceed the "algebraic multiplicity" m_i.

Combine the bases for $E_{\lambda_1}, \ldots, E_{\lambda_k}$. This gives a maximal linearly independent set of eigenvectors for T. (It is linearly independent by the theorem on independence of eigenvectors for different eigenvalues.) And the number of independent eigenvectors is

$$(\star) \qquad \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} \le m_1 + \cdots + m_k \le n.$$

• 4. Now there are two cases.

4A. The good case: $\sum_i \dim E_{\lambda_i} = n$ (i.e., equality holds in ($\star$)).

This happens if and only if both (a) the characteristic polynomial splits completely into linear factors (no irreducible quadratics), and (b) for each i for which $m_i > 1$, the geometric multiplicity $\dim E_{\lambda_i}$ equals the algebraic multiplicity m_i. (For example, (a) always happens if the field is $\mathbf{C}$. And (b) always happens if $m_i = 1$ for each i, i.e., if every root of the characteristic polynomial is a simple root.)

Then we have a basis ε of n eigenvectors. The matrix $D = [T]^\varepsilon_\varepsilon$ is an $n \times n$ diagonal matrix, and the n entries on the diagonal are exactly the eigenvalues, each repeated according to its multiplicity (i.e., λ_i is repeated m_i times). The diagonal matrix D will equal $Q^{-1}AQ$, where $Q = [id]^\sigma_\varepsilon$, so that the columns of Q are the coordinate vectors (relative to σ) of the eigenvectors in ε (i.e., the column vectors from step 3).

4B. The bad case: $\sum_i \dim E_{\lambda_i} < n$.

This happens if and only if either (a) the characteristic polynomial does not split completely into linear factors, or (b) for some i, the eigenspace E_{λ_i} has dimension less than m_i. In this case, T is not diagonalizable. (The maximal linearly independent set from step 3 fails to span V.) Even in the bad case, we can form a basis with as many eigenvectors as possible. The resulting matrix may not be diagonal, but it might be better than the one we started with. Beyond that, there is something called *Jordan normal form*; see Math 115B.

Example. Consider the two matrices:

$$A = \begin{bmatrix} 5 & 3 \\ 0 & 5 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

A and B have the same characteristic polynomial, namely $p(\lambda) = (5 - \lambda)^2$. So in both cases there is a single eigenvalue, namely 5, with multiplicity 2. But the eigenspace of A for 5 is a line (the x-axis), while the eigenspace of B for 5 is the entire plane. Hence A is not diagonalizable (its eigenvectors do not span the plane), while B *is* a diagonal matrix.

This example shows that we cannot always tell from the characteristic polynomial alone whether or not a matrix is diagonalizable. If there are roots that are not simple (i.e., have multiplicity greater than 1), then we must investigate the dimension of the eigenspace before we can be sure if the matrix is diagonalizable.

By the way, the word "eigen" comes from the German. Sometimes instead of "eigenvalue," the English phrases "characteristic value" or "proper value" are used, but usually the German word is preferred. Literally, the German "Eigenwert" can be translated as "its own value." The idea is that the eigenvalues and eigenvectors of T are determined by T itself, independent of any choice of basis vectors or representations.

Eigenvalues and eigenvectors were introduced here to deal with the question of diagonalization. It should be noted that eigenvectors are useful in other situations as well, although we focus primarily on diagonalization.

- To calculate the (usual) inner product in $\mathbb{R}^n$: $\langle x, y\rangle = y^t x$.
  To calculate the (usual) inner product in $\mathbf{C}^n$: $\langle x, y\rangle = y^* x$.

- To find $S^\perp$: Solve the system $\{v \perp s \mid s \in S\}$ for v. Often the theorem that
$$(\text{column space of } A)^\perp = \text{nullspace of } A^*$$
  is useful.

- To find the coordinate vector $[v]_\beta$ relative to an orthonormal ordered basis $\beta = \{u_1, \ldots, u_n\}$: We can proceed as before, or we can use the Fourier coefficients:
$$[v]_\beta = \begin{bmatrix} \langle v, u_1\rangle \\ \vdots \\ \langle v, u_n\rangle \end{bmatrix}$$
  so that $v = \langle v, u_1\rangle u_1 + \cdots + \langle v, u_n\rangle u_n$.

- To find the representing matrix $[T]_\beta^\beta$ relative to an orthonormal ordered basis $\beta = \{u_1, \ldots, u_n\}$: We can proceed as in Chapter 5, or we can use
$$([T]_\beta^\beta)_{ij} = \langle T(u_j), u_i\rangle.$$

- To find the orthogonal projection p of a vector v onto an m-dimensional subspace W having basis $\{w_1, \ldots, w_m\}$:
$$p = \sum_{j=1}^m \frac{\langle v, w_j\rangle}{||w_j||^2} w_j \qquad \text{for an } \textit{orthogonal} \text{ basis}$$

$$p = \sum_{j=1}^{m} \langle v, w_j \rangle w_j \qquad \text{for an } \textit{orthonormal} \text{ basis}$$

•To replace linearly independent vectors $w_1, w_2, \ldots$ by nonzero orthogonal vectors $u_1, u_2, \ldots$ with the same span:

$$u_{k+1} = w_{k+1} - (\text{the projection of } w_{k+1} \text{ onto } \operatorname{span}\{u_1, \ldots, u_k\})$$

•To find an orthonormal basis for diagonalizing a self-adjoint linear operator (or a Hermitian matrix): Proceed as before to make a basis of eigenvectors, with the additional step of obtaining an orthonormal basis for each individual eigenspace E_λ.