Solutions to Simultaneous Diagonalization Worksheet

Math 110 Final Prep by Dan Sparks

I hope you find these problems interesting. I did! Two of them I borrowed from other GSI's (Mike Hartglass and Mohammad Safdari). These two problems, as well as one other (Problem 4), have already appeared in the worksheets. The rest are new and build upon the old ones.

Problem 1 (Mike Hartglass): Let T be an operator on a finite dimensional vector space. Show that T is diagonalizable if and only $T = a_1P_1 + \cdots + a_kP_k$ where the $a_i \in \mathbf{F}$ are scalars and where the P_i are projections that commute with each other: i.e., $P_iP_j = P_jP_i$. $\square$

Solution: First, suppose that T is diagonalizable. Let $E_{i,j}$ denote the matrix with 0's everywhere except in the (i, j) spot, where it has a 1. Then

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \sum_{i=1}^{n} \lambda_i E_{i,i}$$

now the $E_{i,i}$'s are projections onto one dimensional subspaces and they clearly commute as $E_{i,i}E_{j,j} = 0$ for $i \neq j$. [Remark: The more conceptual way is to write $V = \bigoplus_\lambda \text{Null}(T - \lambda I)$ and consider the projections onto the eigenspaces. What this amounts to is grouping by eigenvalue and factoring out the common eigenvalue in $\sum \lambda_i E_{i,i}$ to get projections onto subspaces of larger dimension (namely the eigenspaces).]

Conversely, suppose that $T = a_1P_1 + \cdots + a_nP_n$. We prove by induction that n commuting projections are simultaneously diagonalizable. For careful detail, see Problem 7, which is a generalization of this. The base case is clear, as a single projection is diagonalizable. To complete the induction decompose $V = V_0 \oplus V_1$, for P_1's $\lambda = 0$ and $\lambda = 1$ eigenspaces respectively. The V_0 and V_1's are P_j invariant for $j \geq 2$, so we restrict the $n - 1$ operators $P_2, \cdots, P_n$ each to V_0 and V_1, simultaneously diagonalize them there, and concatenate bases.

Problem 2 (Mohammad Safdari): Let S, T be self-adjoint operators on a finite dimensional $\mathbf{R}$-inner product space [or let S, T be normal operators on a finite dimensional $\mathbf{C}$-inner product space]. Suppose also that $ST = TS$. Show that S, T are simultaneously orthogonally diagonalizable. That is, show there is an orthonormal basis consisting of vectors which are eigenvectors for both S and T. $\square$

Solution: See the more general case below.

These were the original two problems which sparked my interest. First, let me point out a generalization of Problem 2 which has an important application in the theory of modular forms. For the curious, it is the space of newforms with the Petersson inner product, and the operators are the Hecke operators.

Problem 3 (New): Let S_i, for i in some index set I, be a collection of normal operators on a finite dimensional complex inner product space (or self-adjoint on a real inner product space). Suppose, for any $i, j \in I$, that $S_iS_j = S_jS_i$. Show that the S_i are simultaneously orthogonally diagonalizable. That is, show that there exists an orthonormal basis consisting of vectors which are eigenvectors for *every* operator S_i. [Suggestion: It would be entirely sufficient for this worksheet to suppose that I is a finite indexing set and use induction.] $\square$

Solution: Induction on the number of operators. The case of one operator is exactly the content of the real and complex spectral theorems. Suppose the result is true for all integers from 1 to $n - 1$. Let $S_1, \cdots, S_n$ be self adjoint or normal operators which commute with each other. By the spectral theorems applied to S_1 we have an orthogonal decomposition

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$$

where $\lambda_1, \cdots, \lambda_m$ are the distinct eigenvalues of S_1, and $V_{\lambda_i} = \mathrm{Null}(S_1 - \lambda_i I)$ is the corresponding eigenspace. Observe that $S_j(V_{\lambda_i}) \subseteq V_{\lambda_i}$ because $S_j S_1 = S_1 S_j$. In detail: let $(S_1 - \lambda_i I)x = 0$. Then

$$\begin{aligned}(S_1 - \lambda_i I)Sx &= (S_1 S - \lambda_i S)x \\ &= (SS_1 - S(\lambda_i I))x \\ &= S(S_1 - \lambda_i I)x \\ &= S(0) \\ &= 0\end{aligned}$$

That means we can restrict the $n-1$ operators $S_2, \cdots, S_n$ to each V_{λ_i}. Using the fact that the adjoint of a restricted operator is the restriction of the adjoint (see lemma below) we see that the restriction of a self adjoint or normal operator is self adjoint or normal, respectively. From here we are able to apply the inductive hypothesis to $S_2|_{V_{\lambda_i}}, \cdots, S_n|_{V_{\lambda_i}}$ to produce an orthonormal simultaneous eigenbasis $\{b_{i,1}, \cdots, b_{i,e_i}\}$ for each V_{λ_i}. While the inductive hypothesis only gives us that the $b_{i,j}$ are eigenvectors for $S_2, \cdots, S_n$, notice that $b_{i,j}$ already lives inside of V_{λ_i}, making it an eigenvector also for S_1. Therefore the concatenated basis $\{b_{1,1}, \cdots, b_{1,e_1}, \cdots, b_{m,1}, \cdots, b_{m,e_m}\}$ consists of eigenvectors for each $S_1, \cdots, S_n$. Finally, the basis is normalized because it was built up of unit length vectors, and it is orthogonal because $b_{i,j} \perp b_{l,k}$ (for $i = l$ this is by construction, for $i \neq l$ this is because $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$ is an orthogonal direct sum).

Lemma: Let $U \subseteq V$ be a subspace of an inner product space and let T be an operator for which $T(U) \subseteq U$. Then $(T|_U)^* = (T^*)|_U$. [Here we give U the inner product induced by V.]

Proof: We check, for each $u, w \in U$

$$\begin{aligned}\langle (T|_U)(u), w\rangle &= \langle T(u), w\rangle \\ &= \langle u, T^*(w)\rangle \\ &= \langle u, (T^*)|_U(w)\rangle\end{aligned}$$

which implies that $(T|_U)^*(w) = (T^*)|_U(w)$.

Recall a problem that I gave on an earlier worksheet.

Problem 4 (Dan Sparks): Let P be a projection on a finite dimensional inner product space. Prove that P is self-adjoint if and only if it is an orthogonal projection. □

Solution: This was discussed in lecture. To summarize, a projection is diagonalizable and has real eigenvalues (namely 0 and 1). So a projection is normal if and only if it is self-adjoint (normal operators with real eigenvalues are self-adjoint, self-adjoint operators are automatically normal). A diagonalizable operator is normal if and only if its eigenspaces are orthogonal to each other by the spectral theorem. Therefore, since projections are always diagonalizable, P is self adjoint if and only if $V_0 \perp V_1$ where $V_0 = \mathrm{Null}(P)$ and $V_1 = \mathrm{Range}(P)$.

Problem 5 (New): Prove the following orthogonal version of Problem 1. Let T be an operator on a finite dimensional inner product space. Then T is orthogonally diagonalizable (i.e., has an orthonormal eigenbasis) if and only if $T = a_1 P_1 + \cdots + a_k P_k$ where the a_i are scalars and the P_i are *orthogonal* projections such that $P_i P_j = P_j P_i$ for all i, j. [Suggestion: There's an easy proof using the preview two exercises.] □

Solution: If T has an orthonormal eigenbasis, then there is an orthogonal direct sum $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$ into eigenspaces. Let P_i be the projection onto V_{λ_i}. Then $T = \sum_i \lambda_i P_i$. These are commuting orthogonal projections because the direct sum is orthogonal. (See also the discussion in the solution of Problem 1.)

For the harder direction, suppose that $T = a_1 P_1 + \cdots + a_n P_n$. where the P_i's are commuting orthogonal projections. By Problem 4, the P_i's are self-adjoint. Since they commute, by Problem 3, they are simultaneously orthogonally diagonalizable. Take $\{b_1, \cdots, b_k\}$ a simultaneous orthogonal eigenbasis, and notice that with respect to this basis $\mathcal{M}(T)$ is a linear combination of diagonal matrices and hence, diagonal.

Problem 6 (New): Let T be a diagonalizable operator on a finite dimensional vector space V. Suppose that U is a T-invariant subspace. Show that $T\big|_U$ is diagonalizable. □

Solution: Perhaps this is easier than I think, but here is a solution anyway. Consider any basis $\{u_1, \cdots, u_k\}$ for U. Let $v_1, \cdots, v_n$ be a basis for V consisting of eigenvectors of T (since T is diagonalizable). If $\text{Span}\{u_1, \cdots, u_k\} = V$, then $\{u_1, \cdots, u_k\}$ is a basis for V. Otherwise, at least one eigenvector v_i must fail to be in $\text{Span}\{u_1, \cdots, u_k\}$. We choose such a v_i and call it w_1 and add it to our list $\{u_1, \cdots, u_k, w_1\}$ and repeat. In this way we can extend the given basis to a basis

$$\beta = \{u_1, \cdots, u_k, w_1, \cdots, w_l\}$$

where w_i is an eigenvector and $u_j \in U$. Consider the corresponding direct sum $U \oplus W$, and the corresponding projection P onto U along W.

We observe that T commutes with P. It suffices to check this for each vector in β. $TP(u_i) = Tu_i$ and $PT(u_i) = T(u_i)$ because $T(U) \subseteq U$ and P is a projection onto U. Also $TP(w_i) = T(0) = 0$ and $PT(w_i) = P(\alpha w_i) = 0$. Hence the operators T and P commute. We now forget the basis β and the decomposition $V = U \oplus W$.

Now consider the decomposition of V into eigenspaces for T, i.e. $V = \bigoplus_\lambda V_\lambda = \bigoplus_\lambda \text{Null}(T - \lambda I)$. Notice that $P(V_\lambda) = V_\lambda$ because P and T commute. (See the solution to Problem 3.) So, we may restrict P to V_λ. The observation is that $(P|_{V_\lambda})^2 = P|_{V_\lambda}$ so that $P|_{V_\lambda}$ is a projection. In this way we reduce the statement of the problem to the case of a projection (rather than a general diagonalizable operator). This means we can diagonalize $P|_{V_\lambda}$, finding bases for each V_λ such that P is diagonal, and we can concatenate these into one basis $\gamma = (z_1, \cdots, z_n)$ such that P and T are both diagonal. We form a basis consisting of those z_i such that P has a nonzero entry at that place on its main diagonal. More precisely, let I be the subset of $\{1, \cdots, n\}$ consisting of i such that $P(z_i) = z_i$. (In other words, $i \notin I$ if and only if $P(z_i) = 0$.)

We claim that $\{z_i : i \in I\}$ is a basis for U. Linear independence is clear. Now let $u \in U$, and write $u = a_1 z_1 + \cdots + a_n z_n$. But then $u = Pu = a_1 P(z_1) + \cdots + a_n P(z_n) = \sum_{i \in I} a_i z_i$, so that this set spans U.

The matrix for T with respect to $\{z_i : i \in I\}$ is diagonal, since each z_i is an eigenvector of T.

[Remark: My apologies if there is a much easier way to do this problem. I would be surprised if there wasn't an easier way, and I would not be surprised if there was a *much* easier way. However, after thinking about the other problems, this seemed natural though a bit clumsy.]

Problem 7 (New): Let S_i, for i in some index set I, be a collection of diagonalizable operators on a finite dimensional vector space V. Suppose, for any $i, j \in I$, that $S_i S_j = S_j S_i$. Show that the S_i are simultaneously diagonalizable. [Suggestion: If necessary, first do the case of just two operators S, T. Again, suppose that I is finite, and try induction.] □

Solution: Induction. The base case is a tautology: a single diagonalizable operator is diagonalizable. Suppose the result is true for any collection of commuting diagonalizable operator of size at most $n-1$. Suppose that $S_1, \cdots, S_n$ are diagonalizable, commuting operators.

Decompose $V = \bigoplus V_{\lambda_i}$ into eigenspaces for S_1. The V_{λ_i} are S_j invariant, so we restrict the $n-1$ operators $S_2, \cdots, S_n$ to each V_{λ_i} where they are a collection of $n-1$ commuting diagonalizable (by Problem 6) operators. By inductive hypothesis, we can simultaneously diagonalize $S_2, \cdots, S_n$ on each V_{λ_i} obtaining bases $\{b_{i,1}, \cdots, b_{i,e_i}$ for V_{λ_i}. The inductive hypothesis only tells us that $b_{i,j}$ is an eigenvector for $S_2, \cdots, S_n$, but since $b_{i,j} \in V_{\lambda_i}$ we see that it is an eigenvector for S_1 as well. Concatenating the bases into $\{b_{1,1}, \cdots, b_{1,e_1}, \cdots, b_{m,1}, \cdots, b_{m,e_m}\}$ which is a simultaneous basis of eigenvectors.

Problem 8 (New): Give an easy proof of Problem 1 using Problem 7. □

Solution: A diagonalizable operator is clearly a sum of commuting projections, namely the projections onto the eigenspaces (see the first paragraph of the solution to Problem 1).

For the harder direction, since projections are diagonalizable we can use Problem 7 on the collection $P_1, \cdots, P_n$. Once they are simultaneously diagonalized, it is clear that a linear combination of diagonal matrices is diagonal, so that T is diagonalizable.

Note that Problem 3 and Problem 7 are analogous to each other, in the same way that Problem 1 and Problem 5 are analogous to each other. Notice especially that neither the plain version nor the orthogonal version is more general than the other. The non-orthogonal versions apply in more contexts, but the orthogonal versions give sharper results. On the other hand, I do think that Problem 3 and Problem 7 are more general/powerful than Problem 1 and Problem 5, since you can give easy proofs of 1 and 5 using 7 and 3 respectively.