122 Solution Set 4

1 4.3.9

(a) Let $\{1, x, x^2, ..., x^n\}$ be a basis for P. With respect to this basis,

$$D := \begin{bmatrix} 0 & 1 & 0 & & & \cdots & 0 \\ 0 & 0 & 2 & 0 & & \cdots & 0 \\ 0 & 0 & 0 & 3 & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & & \cdots & & \cdots & & n \\ 0 & & \cdots & & \cdots & & 0 \end{bmatrix}.$$

The characteristic polynomial of D is easily calculated to be x^n. By the Cayley-Hamilton theorem (that is, the last problem on this set), we have $D^n = 0$, which implies that D is nilpotent (quick proof by Jason Adaska).

(b). Define $P_i := span\{1, x, ..., x^j\}$ for $1 \leq j \leq n$. These P_i are clearly invariant subspaces, as almost everyone noted. However, most people didn't show (or didn't show correctly) that these were the only invariant subspaces. Here we go:

Suppose W is a nontrivial invariant subspace of P. Pick an element of maximal degree in W, say $p = a_k x^k + ... + a_0$, where necessarily $a_k \neq 0$. I claim $W = P_k$. We proceed by induction. $D^{k-1}p \in \mathbb{R}$, so $< 1 >$ is contained in W. Suppose we've shown $P_{n-1} \leq W$. Then $D^{k-n-2}p - c_{k-1}x^{k-1} + ... + c_0 = d_n x^n$ for some $c_i, d_0 \in \mathbb{R}$, which implies $x^n \in W$. Hence $P_n \leq W$. Continuing inductively, we have that $W = P_k$.

2 4.3.10

If $a \neq d$, B clearly has the distinct eigenvalues a, d (note that the eigenvectors of an upper triangular matrix are the entries along the diagonal). Thus, by Proposition (6.3) and Theorem (6.4), we may write B, via a change of basis, as the diagonal matrix A, where the new basis is the basis of eigenvectors for B guaranteed to exist by Proposition (6.3). Saying that two matrices are the same up to a change of basis is the same as saying that they are similar.

Conversely, suppose $a = d$. Then, for all matrices $P \in GL_n(\mathbb{R})$, $PAP^{-1} = B$ implies that $APP^{-}1 = B$ (because A is diagonal) which implies $A = B$, a contradiction, for b is nonzero.

3 4.4.2

(a) Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a symmetric matrix in $GL_2(\mathbb{R})$. The characteristic polynomial of A has discriminant $(a-c)^2+4b^2$ which is always nonnegative, thus the characteristic polynomial has real roots. The roots of the characteristic polynomial are the

(b) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL_2(\mathbb{R})$ such that $b, c > 0$. The characteristic polynomial of A has discriminant $(a-d)^2 + 4bc \geq 0$ because $b, c > 0$, thus the characteristic polynomial has real roots.

4 4.4.4

A real 3×3 matrix has a characteristic polynomial of degree three. A degree three polynomial has at least one real root (from elementary analysis), hence the 3×3 matrix has at least one real eigenvalue.

5 4.5.4

(a) Suppose $A, B \in O_n$.

Closure: $AB(AB)^T = ABB^TA^T = AIA^T = AA^T = I$.

Inverses: $AA^T = A^TA = I$ implies $A^{-1} = A^T$.

Identity: $II^T = I$.

Note $SO_n = SL_n \cap O_n$. The intersection of two subgroups is also a subgroup, so SO_n is a subgroup.

Note that $det : O_n \to \{-1, 1\}$ is a homomorphism. Therefore $[O_n : SO_n] = |O_n/SO_n| = 2$ by Lagrange.

(b) O_2 is NOT isomorphic to $SO_2 \times \{I, -I\}$. The only elements of order two in $SO_2 \times \{I, -I\}$ are $\rho_\pi \times I$, $\rho_\pi \times -I$, $I \times I$, $I \times -I$. That is, there are four of them; any other rotation ρ in SO_2 has order greater than 2, which forces the order of $\rho \times I$, $\rho \times -I$ to be greater than two. In contrast, O_2 has infinitely many elements of order two; any reflection through the origin is such an element. Because an isomorphism preserves order, the two groups cannot be isomorphic.

O_3 is isomorphic to $SO_3 \times \{I, -I\}$. $-I \notin SO_3$, so $SO_3 \cap \{I, -I\} = I$ Considering $\{I, -I\}$ as a subgroup of O_3, it is easy to see that it is normal because both of these elements commute with any matrix. We've already commented that SO_n is normal in O_n above. Because $-I \notin SO_3$ and $[O_n : SO_n] = | = 2$, we have that $SO_3\{I, -I\} = O_3$. Thus, by Proposition 2.8.6.c, O_3 is isomorphic to $SO_3 \times \{I, -I\}$.

6 4.5.7

The product of the two eigenvalues of A is equal to the determinant of A. Further, we proved in class that if λ is an eigenvalue of A, then so is λ^{-1}. The only possible eigenvalues satisfying these two stipulations are $1, -1$. Let the associated eigenvectors be v_1, v_2, respectively. Because they are eigenvectors corresponding to different eigenvalues of A, they are linearly independent. I claim that they are orthogonal. We have

$$< v_1, v_2 >=< Av_1, Av_2 >=< v_1, -v_2 >= - < v_1, v_2 >$$

which implies $< v_1, v_2 >= 0$. Noting that $Av_1 = v_1$ and $Av_2 = -v_2$, we have that A is a reflection of the plane fixing the span of the vector v_1.

7 4.5.8

We proved in class that any element A of SO_3 fixes a vector v_1 (where we may assume v_1 is a unit vector) and rotates the plane normal to that vector. Extend v_1 to an orthonormal basis for $\mathbb{R}^3$, say $\{v_1, v_2, v_3\}$. With respect to this basis A can be represented in matrix form as

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

The lower right hand corner of this matrix is just a rotation of θ degrees the plane normal to the vector v_1. We have already chosen an orthonormal basis; and hence the bottom right corner takes on the form below:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

(see 4.5.2). The trace of this matrix clearly satisfies the given relation. Because trace is invariant under similarity transformations, we $\cos(\theta)) = \frac{1}{2}(trA - 1)$

8 4.6.1

(a) The characteristic polynomial of this matrix is $(2-x)^2 - -1 = (x-3)(x-1)$, so the eigenvalues are $3, 1$. The corresponding eigenvectors are $(1,1)^T, (1,-1)^T$, which you can readily check (this is just high school algebra).

(b) Because the basis for the diagonal matrix will be $(1,1)^T, (1,-1)^T$, we know that these two vectors form the columns of the matrix P, the change of basis matrix from the new basis to the standard basis. Thus we have:

$$PAP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 3 & 3 \\ 1 & -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

(c)

$$A = P^{-1} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} P$$

which implies

$$A^{30} = (P^{-1} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} P)^{30}$$

$$= P^{-1} (\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix})^{30} P$$

$$= P^{-1} \begin{bmatrix} 3^{30} & 0 \\ 0 & 1 \end{bmatrix} P$$

$$= \frac{1}{2} \begin{bmatrix} 3^{30}+1 & 3^{30}-1 \\ 3^{30}-1 & 3^{30}+1 \end{bmatrix}.$$

9 4.8.1

(a)

$$\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \begin{bmatrix} & 1 \\ -1 & \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -1 & \\ & -1 \end{bmatrix} + \frac{1}{3!} \begin{bmatrix} & -1 \\ 1 & \end{bmatrix} + \frac{1}{4!} \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + ... =$$

$$\begin{bmatrix} \sum_{i=0}^{\infty} \frac{(-1)^n}{(2n)!} & \sum_{i=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \\ -\sum_{i=0}^{\infty} \frac{(-1)^n}{(2n+1)!} & \sum_{i=0}^{\infty} \frac{(-1)^n}{(2n)!} \end{bmatrix} =$$

$$\begin{bmatrix} \cos(1) & \sin(1) \\ -\sin(1) & \cos(1) \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \begin{bmatrix} a & b \\ & \end{bmatrix} + \frac{1}{2} \begin{bmatrix} a^2 & ab \\ & \end{bmatrix} + \frac{1}{3!} \begin{bmatrix} a^3 & a^2 b \\ & \end{bmatrix} + \frac{1}{4!} \begin{bmatrix} a^4 & a^3 b \\ & \end{bmatrix} + ... =$$

$$\begin{bmatrix} \sum_{i=0}^{\infty} \frac{a^i}{i!} & \frac{b}{a} \sum_{i=1}^{\infty} \frac{a^i}{i!} \\ & 1 \end{bmatrix} =$$

$$\begin{bmatrix} e^a & \frac{b}{a}(e^a - 1) \\ 0 & 1 \end{bmatrix}.$$

10 4.MISC.15

(a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then charpoly$(A) = (a-x)(d-x)-bd = x^2-(a+d)x+ad-bc$. We have

$$A^2 - (a+d)A + (ad-bc)I =$$

$$\begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} + \begin{pmatrix} -a^2-ad & -ab-bd \\ -ac-cd & -ad-d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(b) Let $A = \begin{pmatrix} \lambda_1 & 0 & 0 & & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ where the λ_i are not necessarily unique. Further, let $p(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$ be the characteristic polynomial of A. Then we have

$$A^n + a_{n-1}x^{n-1} + ... + a_0 I =$$

$$\begin{pmatrix} \lambda_1^n & 0 & 0 & & 0 \\ 0 & \lambda_2^n & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3^n & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda_n^n \end{pmatrix} + \begin{pmatrix} a_{n-1}\lambda_1^{n-1} & 0 & 0 & & 0 \\ 0 & a_{n-1}\lambda_2^{n-1} & 0 & \cdots & 0 \\ 0 & 0 & a_{n-1}\lambda_3^{n-1} & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & a_{n-1}\lambda_n^{n-1} \end{pmatrix} + ...$$

$$... + \begin{pmatrix} a_0 & 0 & 0 & & 0 \\ 0 & a_0 & 0 & \cdots & 0 \\ 0 & 0 & a_0 & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & a_0 \end{pmatrix} =$$

$$\begin{pmatrix} p(\lambda_1) & 0 & 0 & & 0 \\ 0 & p(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & p(\lambda_3) & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & p(\lambda_n) \end{pmatrix}$$

But the λ_i, being the diagonal entries of a diagonal matrix, are the eigenvalues of A. Hence $p(\lambda_i) = 0$ for all i, and we have our result.

(c) By hypothesis, we can find some $n \times n$ matrix P such that PAP^{-1} is diagonal. Again, let $p(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$ be the characteristic polynomial of A. We have, from (b),

$$p(PAP^{-1}) = (PAP^{-1})^n + a_{n-1}(PAP^{-1})^{n-1} + ... + a_0 I =$$

$$PA^nP^{-1} + a_{n-1}PA^{n-1}P^{-1} + ... + a_0 I =$$

$$P(A^n + a_{n-1}A^{n-1} + ... + a_0 I)P^{-1} = 0.$$

This directly implies, by multiplying on the left by P^{-1} and on the right by P, that $p(A) = 0$.

(d) For this problem, we will choose the operator norm $|| \cdot ||$ on $n \times n$ matrices.

Let A be an arbitrary $n \times n$ matrix. By Proposition 6.1, there is an $n \times n$ matrix P such that $C := PAP^{-1}$ is upper triangular. The diagonal entries of C are the eigenvalues of C, and hence also A. Call these eigenvalues $\lambda_1, ..., \lambda_n$. Pick some $\epsilon > 0$, and then choose $\epsilon_1, ..., \epsilon_n$ such that $\lambda_1 + \epsilon_1, ..., \lambda_n + \epsilon_n$ are all distinct and

$\epsilon_i < (\frac{\epsilon}{2n})^{1/2}$ for all i. Let $E = \begin{pmatrix} \epsilon_1 & 0 & 0 & & 0 \\ 0 & \epsilon_2 & 0 & \cdots & 0 \\ 0 & 0 & \epsilon_3 & & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & \cdots & \epsilon_n \end{pmatrix}$. Then $C+E$ is a diagonalizable matrix, for it has distinct eigenvalues. This implies that $D := P^{-1}(C+E)P$ is also a diagonalizable matrix. Now, we have

$$||D - A|| = ||P^{-1}(C+E)P - A|| = ||P^{-1}CP + P^{-1}EP - A||$$

$$\leq ||P^{-1}CP - A|| + ||P^{-1}EP|| = 0 + ||P^{-1}||||E||||P||$$

$$= ||E|| \leq n((\frac{\epsilon}{2n})^{1/2})^2 < \epsilon.$$

Thus, we have a diagonalizable matrix D arbitrarily close to our original matrix A. Let $p(x)$ be the characteristic polynomial of A. From continuity of polynomials, we have that $p(D) = 0$ (from c above) implies $\lim_{\epsilon \to 0} p(D) = p(A) = 0$.