dc3451
David Chen

Homework 6

Apostol p.20 no.3

Claim.

$$(x, y) = 0 \iff ||x + y|| = ||x - y||$$

Proof.

$$\begin{aligned}
||x + y|| &= \sqrt{(x + y, x + y)} \\
&= \sqrt{(x, x + y) + (y, x + y)} \\
&= \sqrt{(x, x) + (x, y) + (y, x) + (y, y)} \\
&= \sqrt{(x, x) + (y, y) + 2(x, y)} \\
||x - y|| &= \sqrt{(x, x) + (-y, -y) + 2(x, -y)} \\
&= \sqrt{(x, x) + (y, y) - 2(x, y)}
\end{aligned}$$

($\implies$) From above, $||x+y|| = \sqrt{(x, x) + (y, y) + 2(0)}, ||x-y|| = \sqrt{(x, x) + (y, y) - 2(0)} \implies ||x + y|| = ||x - y||$.

($\impliedby$) From $\sqrt{(x, x) + (y, y) + 2(x, y)} = \sqrt{(x, x) + (y, y) - 2(x, y)} \implies (x, x) + (y, y) + 2(x, y) = (x, x)+(y, y)-2(x, y) \implies 2(x, y) = -2(x, y) \implies 4(x, y) = 0 \implies (x, y) = 0$. ☐

Apostol p.20 no.4

Claim.

$$(x, y) = 0 \iff ||x + y||^2 = ||x||^2 + ||y||^2$$

Proof. ($\implies$) From above, we have that $||x+y||^2 = (x, x)+(y, y)+2(x, y) = (x, x)+(y, y) = ||x||^2 + ||y||^2$.

($\impliedby$) From above, we have that $||x + y||^2 = ||x||^2 + ||y||^2 \implies (x, x) + (y, y) + 2(x, y) = (x, x) + (y, y) \implies 2(x, y) = 0 \implies (x, y) = 0$. ☐

Apostol p.20 no.5

Claim.

$$(x, y) = 0 \iff \forall c \in \mathbb{R}, ||x + cy|| \geq ||x||$$

Proof. ($\implies$) From above, $||x + cy|| = \sqrt{(x, x) + (cy, cy) + 2(x, cy)} = \sqrt{(x, x) + c^2(y, y)}$. Since we have from the axioms that $(y, y) \geq 0$, $(x, x)+c^2(y, y) \geq 0 \implies \sqrt{(x, x) + c^2(y, y)} \geq \sqrt{(x, x)} \implies ||x + cy|| \geq ||x||$.

($\impliedby$) From above, $\sqrt{(x, x) + c^2(y, y) + 2(x, cy)} \geq \sqrt{(x, x)} \implies (x, x)+c^2(y, y)+2(x, cy) \geq (x, x) \implies c^2(y, y) + 2c(x, y) \geq 0$.

If $y = 0$, then $(x, y) = (x, 0) = 0(x, 0) = 0$. Otherwise, $(y, y) > 0$. Take $c = -\frac{(x,y)}{(y,y)}$, such that $\frac{(x,y)^2}{(y,y)} - 2\frac{(x,y)^2}{(y,y)} = -\frac{(x,y)^2}{(y,y)} \geq 0$. Since $(y, y) > 0$, we must have that $-(x, y)^2 \geq 0 \implies (x, y)^2 \leq 0 \implies (x, y) = 0$. $\square$

Problem 1

a

Claim. Let $A \in M_{n\times n}(F)$, A_{ij} the i, j element of A, and A^{ij} be the submatrix of A with the i row and j column removed. Then,

$$p_A(\lambda) = \prod_{i=1}^{n}(\lambda - A_{ii}) + g(\lambda)$$

where the degree of g is at most $n - 2$.

Proof. Induct on n.
For $n = 1$, we have that $\lambda I_1 - A = \begin{bmatrix}\lambda - a_{11}\end{bmatrix}$, and that $\det(\lambda I_1 - A) = \lambda - a_{11}$.
Now assume the hypothesis for $n = k$. Since we have that $p_A(\lambda) = \det(\lambda I_{k+1} - A)$, cofactor expansion gives that

$$p_A(\lambda) = \det(\lambda I_{k+1} - A) = (\lambda - A_{11})\det((\lambda I_{k+1} - A)^{11}) + \sum_{i=2}^{k+1}(-1)^{i-1}A_{1i}\det((\lambda I_{k+1} - A)^{1i})$$

Now, we have that the $1, 1$ minor of $I_n \in M_{n\times n}$ is $I_{n-1} \in M_{(n-1)\times(n-1)}$. To see this, consider that $(I_n^{11})_{ij} = (I_n)_{(i+1)(j+1)}$ (as for the i, j cofactor matrix, we have that $A^{kl}_{ij} = A_{(i+1)(j+1)}$ for $i \geq k, j \geq l$). Then, since we have that $(I_n)_{(i+1)(j+1)} = \delta_{(i+1)(j+1)} = \delta_{ij}$, then $I_n^{11} = I_{n-1}$.
Now, $(\lambda I_{k+1} - A)^{11} = \lambda I_{k+1}^{11} - A^{11} = \lambda I_k - A^{11}$. Then, by the inductive hypothesis, $p_{A^{11}}(\lambda) = \det((\lambda i_{k+1} - A)^{11}) = \prod_{i=2}^{k+1}(\lambda - a_{ii}) + g_{11}(\lambda)$, where g_{11} is of degree at most $k-2$ by assumption. Then,

$$\begin{aligned}
p_A(\lambda) &= (\lambda - A_{11})(\prod_{i=2}^{k+1}(\lambda - A_{ii}) + g_{11}(\lambda)) + \sum_{i=2}^{k+1}(-1)^{i-1}A_{1i}\det((\lambda I_{k+1} - A)^{1i}) \\
&= \prod_{i=1}^{k+1}(\lambda - A_{ii}) + (\lambda - A_{11})g_{11}(\lambda) + \sum_{i=2}^{k+1}(-1)^{i-1}A_{1i}\det((\lambda I_{k+1} - A)^{1i})
\end{aligned}$$

Now, for any $A_{1i}\det((\lambda I_{k+1} - A)^{1i})$, consider the permutation formula for the determinant; since $\lambda I_{k+1} - A$ has only $k + 1$ elements with the form $\lambda - x$, i.e. the diagonal elements (as these are the only nonzero elements in the identity matrix). Then, we have that since removing the 1 row and j column removes $(\lambda I_{k+1} - A)_{11}$ and $(\lambda I_{k+1} - A)_{jj}$, the resulting cofactor matrix can only have at most $k-1$ elements with λ (i.e. elements that have the form $\lambda - A_{ij}$).

Then, since the permutation formula has that

$$\det((\lambda I_{k+1} - A)^{1j}) = \sum_{\sigma \in S_k} \prod_{i=1}^{k} (\lambda I_{k+1} - A)^{1j}_{i\sigma(i)} \operatorname{sgn}(\sigma)$$

any individual $\prod_{i=1}^{k} (\lambda I_{k+1} - A)^{1j}_{i\sigma(i)} \operatorname{sgn}(\sigma)$ is a polynomial of λ of at most degree $k-1$, as it is a product of k distinct elements of $(\lambda I_{k+1} - A)^{1j}$, which has at most $k-1$ terms which are binomials of degree 1 in λ. Call $\det((\lambda I_{k+1} - A)^{1j}) = g_{1j}(\lambda)$.

Now,

$$\begin{aligned} p_A(\lambda) &= \prod_{i=1}^{k+1} (\lambda - A_{ii}) + (\lambda - A_{11}) g_{11}(\lambda) + \sum_{i=2}^{k+1} (-1)^{i-1} A_{1i} \det((\lambda I_{k+1} - A)^{1i}) \\ &= \prod_{i=1}^{k+1} (\lambda - A_{ii}) + (\lambda - A_{11}) g_{11}(\lambda) + \sum_{i=2}^{k+1} (-1)^{i-1} A_{1i} g_{1j}(\lambda) \end{aligned}$$

Since we have that $g_{11}(\lambda)$ has degree at most $k-2$, and A_{1i} for $i \neq 1$ has no λ term, and that g_{1j} has degree at most $k-1$, we have that $(\lambda - A_{11}) g_{11}(\lambda) + \sum_{i=2}^{k+1} (-1)^{i-1} A_{1i} g_{1j}(\lambda)$ is a polynomial of at most degree $k-1$. Call that $g(\lambda)$.

Finally,

$$p_A(\lambda) = \prod_{i=1}^{k+1} (\lambda - A_{ii}) + g(\lambda)$$

and the induction is finished. $\square$

b

Claim. $p_A(\lambda)$ has leading term 1, second term $-\operatorname{tr}(A)$, and constant term $(-1)^n \det(A)$.

Proof. Since we have that

$$p_A(\lambda) = \det(\lambda I - A) = \prod_{i=1}^{n} (\lambda - A_{ii}) + g(\lambda)$$

taking $p_A(0) = \det(0I - A) = \det(-A) = (-1)^n \det(A)$ (the $(-1)^n$ comes from multiplying n rows by -1 to get from $-A$ to A). Further, since p_A is a polynomial, $p_A(0)$ is only the constant term. Thus, the constant term of p_A is $(-1)^n \det(A)$.

We will now prove the other two via induction on n; specifically, if

$$p_n(x) = \sum_{i=1}^{n}(x - a_i)$$

then the leading and second terms will be 1 and $-\sum_{i=1}^{n} a_i$ respectively.
For the base case of $n = 1$, the polynomial is $x - a_1$, and the criteria are met.
Assuming the hypothesis for $n = k$, we have that

$$\begin{aligned}
p_{k+1}(x) &= \sum_{i=1}^{k+1}(x - a_i) \\
&= (x - a_{k+1})\sum_{i=1}^{k}(x - a_i) \\
&= (x - a_{k+1})(x^k - (\sum_{i=1}^{k} a_i)x^{k-1} + \sum_{i=1}^{k-2} c_i x^{k-2}) \\
&= x^{k+1} - (\sum_{i=1}^{k} a_i)x^k + \sum_{i=1}^{k-2} c_i x^{k-1} - a_{k+1}x^k + (\sum_{i=1}^{k} a_i)a_{k+1}x^{k-1} + \sum_{i=1}^{k-2} a_{k+1}c_i x^{k-2} \\
&= x^{k+1} - (\sum_{i=1}^{k+1} a_i)x^k + \sum_{i=1}^{k-1} c'_i x^i
\end{aligned}$$

Now, notice that $g(\lambda)$ has at most degree $n-2$, while $\prod_{i=1}^{n}(\lambda - A_{ii})$ is of degree n, so g contributes nothing to the first or second terms. Now, from above we have that the leading coefficient is 1 and the second coefficient is $-\sum_{i=1}^{n} A_{ii} = -\mathrm{tr}(A)$. □

Problem 2

a

Claim. If $D = \mathrm{diag}(d_1, \ldots, d_n)$, then $D^k = \mathrm{diag}(d_1^k, \ldots, d_n^k)$.

Proof. Induct on k. For $k = 1$, it is immediate.
Now assume the claim for $k = l$. Then,

$$D^{l+1} = (D^l)D$$

such that

$$D_{ij}^{l+1} = \sum_{m=1}^{n} D_{im}^{l} D_{mj}$$

Note that if $i \neq j$, then either $i \neq m$ or $m \neq j$, so $D^l_{im} = 0$ or $D_{mj} = 0$, so the only nonzero entries can be the diagonals. Then,

$$D^{l+1}_{ii} = \sum_{m=1}^{n} D^l_{im} D_{mi} = D^l_{ii} D_{ii} = d^l_i d_i = d^{l+1}_i$$

And so $D^{l+1} = \operatorname{diag}(d^{l+1}_1, \dots, d^{l+1}_n)$.

□

b

Claim. Let $A = B^{-1}DB$. Then,

$$A^k = B^{-1}D^k B$$

Proof. Induct again on k. For $k = 1$, it is immediate.
Assuming the claim for $k = l$, then

$$A^{l+1} = A^l A = (B^{-1}D^l B)(B^{-1}DB) = (B^{-1}D^l)(BB^{-1})(DB) = B^{-1}(D^l D)B = B^{-1}D^{l+1}B$$

□

c

We first diagonalize A.

$$\begin{aligned}
\det(\lambda I - A) &= \det\left(\begin{bmatrix} \lambda - 3 & -2 \\ -3 & \lambda - 4 \end{bmatrix}\right) \\
&= (\lambda - 3)(\lambda - 4) - 6 \\
&= \lambda^2 - 7\lambda + 6 \\
&= (\lambda - 1)(\lambda - 6)
\end{aligned}$$

Taking $\lambda = 1$,

$$(I - A) = \begin{bmatrix} -2 & -2 \\ -3 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Thus, the eigenspace belonging to $\lambda = 1$ is spanned by $(1, -1)$. Taking $\lambda = 6$,

$$(6I - A) = \begin{bmatrix} 3 & -2 \\ -3 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

Thus, the eigenspace belonging to $\lambda = 6$ is spanned by $(2, 3)$. Finally, inverting $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ we have:

$$A = \begin{bmatrix} \frac{3}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$$

Problem 3

We first compute:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}(a_{11} + a_{22}) \\ a_{21}(a_{11} + a_{22}) & a_{22}^2 + a_{12}a_{21} \end{bmatrix}$$

a

Claim. The real matrix $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ has exactly two square roots.

Proof. We see that $a_{22}^2 - a_{11}^2 = 1 \implies (a_{22} - a_{11})(a_{22} + a_{11}) = 1$. This means that $a_{11} + a_{22} \neq 0$, and as $a_{12}(a_{11} + a_{22})$ and $a_{21}(a_{11} + a_{22})$ are both zero, $a_{12} = a_{21} = 0$.
Then, we have that $a_{11}^2 = 0 \implies a_{11} = 0$, and $a_{22}^2 = 1 \implies a_{22} = \pm 1$. Thus, there are only two square roots and they are $\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. □

b

Claim. The real matrix $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ has exactly four square roots.

Proof. Similarly to above, we have that $a_{11}^2 - a_{22}^2 = 3 \implies (a_{11} - a_{22})(a_{11} + a_{22}) = 3$, such that now $a_{12} = a_{21} = 0$. Further, we have now that $a_{11} = \pm 2$, $a_{22} = \pm 1$, so the four square roots are $\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$. □

c

Claim. The real matrix $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has no square roots.

Proof. Suppose that $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a square root.

We have that $a_{11}^2 + a_{12}a_{21} = a_{22}^2 + a_{12}a_{21} \implies a_{11}^2 = a_{22}^2$. This means that $a_{11} = \pm a_{22}$; however, if $a_{11} = -a_{22}$, $a_{12}(a_{11} + a_{22}) = 0$, $\Rightarrow\Leftarrow$.

However, if $a_{11} = a_{22}$, then as $a_{21}(a_{11} + a_{22}) = 0$, we need $a_{21} = 0$. Then, as $a_{11}^2 + a_{12}a_{21} = a_{11}^2 = 0, a_{22}^2 + a_{12}a_{21} = a_{22}^2 = 0$, we need that $a_{11} = a_{22} = 0$. Then, $a_{12}(a_{11} + a_{22}) = 0$, $\Rightarrow\Leftarrow$.

Therefore no such square root exists. □

d

Claim. The 2×2 zero matrix has infinitely many square roots.

Proof. We must have the following:

$$\begin{cases} a_{11}^2 + a_{12}a_{21} = 0 \\ a_{21}(a_{11} + a_{22}) = 0 \\ a_{12}(a_{11} + a_{22}) = 0 \\ a_{22}^2 + a_{12}a_{21} = 0 \end{cases}$$

Thus, $a_{21}(a_{11} + a_{22}) = a_{12}(a_{11} + a_{22}) \implies (a_{12} - a_{21})(a_{11} + a_{22}) = 0$.

Since we know in any field, $ab = 0 \iff a = 0$ or $b = 0$, we have two cases.

First, let $a_{12} = a_{21}$. Then, $a_{11}^2 + a_{12}^2 = 0$. Thus, $a_{11} = a_{12} = 0$. Similarly, $a_{22} = 0$.

Second, let $a_{11} = -a_{22}$. Then, $a_{11}^2 + a_{12}a_{21} = 0 \implies -a_{12}a_{21} = a_{11}^2$. All such roots are then parameterized by

$$\begin{bmatrix} x & y \\ -\frac{x^2}{y} & -x \end{bmatrix}, \begin{bmatrix} x & -\frac{x^2}{y} \\ y & -x \end{bmatrix}$$

for any $x, y \in \mathbb{R}, y \neq 0$.

Thus, we have that all square roots are either the above form, or the zero matrix. □

Problem 5

a

The equivalency between the characteristic polynomial and $\det(\lambda I - A)$ holds only if λ is a scalar in the base field of A, since this property only holds when $(A - \lambda I)x = 0$ where x is an eigenvector belonging to $\lambda \in F$. Taking $\lambda = A$ is very obviously not a scalar in the base field of A, so the proof fails.

b

Claim. Cayley-Hamilton holds for 2×2 matrices.

Proof.

$$\begin{aligned}
p_A(\lambda) =& \begin{bmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{bmatrix} \\
=& (\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21} \\
=& \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) \\
p_A(A) =& A^2 - (a_{11} + a_{22})A + (a_{11}a_{22} - a_{12}a_{21})I \\
=& \begin{bmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}(a_{11} + a_{22}) \\ a_{21}(a_{11} + a_{22}) & a_{22}^2 + a_{12}a_{21} \end{bmatrix} - \begin{bmatrix} (a_{11} + a_{22})a_{11} & (a_{11} + a_{22})a_{12} \\ (a_{11} + a_{22})a_{21} & (a_{11} + a_{22})a_{22} \end{bmatrix} \\
& + \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} \\
=& \begin{bmatrix} a_{11}^2 + a_{12}a_{21} - (a_{11} + a_{22})a_{11} + a_{11}a_{22} - a_{12}a_{21} & a_{12}(a_{11} + a_{22}) - (a_{11} + a_{22})a_{12} \\ a_{21}(a_{11} + a_{22}) - (a_{11} + a_{22})a_{21} & a_{22}^2 + a_{12}a_{21} - (a_{11} + a_{22})a_{22} + a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} \\
=& \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

□

Problem 6

a

Claim. Cayley-Hamilton holds for diagonal matrices.

Proof. The determinant of a diagonal matrix is just the product of the diagonal elements. To see this, consider that

$$\det(D) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} D_{i\sigma(1)} \mathrm{sgn}(\sigma)$$

There is only one permutation that takes all $A_{i\sigma(i)}$ to a nonzero entry, and that is when $\forall i \in [1, n], i = \sigma(i)$. This is the identity permutation, and thus $\det(D) = \prod_{i=1}^{n} D_{ii}$.

Now take any diagonal matrix D. For any $i, j, i \neq j$, $(\lambda I - D)_{ij} = \lambda I_{ij} - D_{ij} = 0 - 0 = 0$. Thus, $\lambda I - D$ is also diagonal. Then,

$$p_D(\lambda) = \det(\lambda I - D) = \prod_{i=1}^{n} (\lambda - D_{ii}) \implies p_D(D) = \prod_{i=1}^{n} (D - D_{ii}I)$$

We will show that

$$\prod_{i=1}^{m} \operatorname{diag}(d_{i1}, \ldots, d_{in}) = \operatorname{diag}(\prod_{i=1}^{m} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in})$$

.

Induct on m. For $m = 1$, this is immediate.

Assume the above for $m = k$. Then,

$$\begin{aligned}
A = \prod_{i=1}^{k+1} \operatorname{diag}(d_{i1}, \ldots, d_{in}) &= (\prod_{i=1}^{k} \operatorname{diag}(d_{i1}, \ldots, d_{in}))\operatorname{diag}(d_{(k+1)1}, \ldots, d_{(k+1)n}) \\
&= \operatorname{diag}(\prod_{i=1}^{k} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in})\operatorname{diag}(d_{(k+1)1}, \ldots, d_{(k+1)n}) \\
A_{uv} &= \sum_{l=1}^{n} (\operatorname{diag}(\prod_{i=1}^{k} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in}))_{ul}(\operatorname{diag}(d_{(k+1)1}, \ldots, d_{(k+1)n}))_{lv}
\end{aligned}$$

When $u \neq v$, we have that both factors on the right are zero. Otherwise,

$$A_{uu} = \sum_{l=1}^{n} (\operatorname{diag}(\prod_{i=1}^{k} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in}))_{ul}(\operatorname{diag}(d_{(k+1)1}, \ldots, d_{(k+1)n}))_{lu}$$

When $u \neq l$, both factors are again zero.

$$\begin{aligned}
&= (\operatorname{diag}(\prod_{i=1}^{k} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in}))_{uu}(\operatorname{diag}(d_{(k+1)1}, \ldots, d_{(k+1)n}))_{uu} \\
&= (\prod_{i=1}^{k} d_{iu})d_{(k+1)u} \\
&= \prod_{i=1}^{k+1} d_{iu}
\end{aligned}$$

Thus,

$$\prod_{i=1}^{m} \operatorname{diag}(d_{i1}, \ldots, d_{in}) = \operatorname{diag}(\prod_{i=1}^{m} d_{i1}, \ldots, \prod_{i=1}^{n} d_{in})$$

.

Finally, we arrive at that

$$\prod_{i=1}^{n} (D - D_{ii}I) = \operatorname{diag}(\prod_{i=1}^{n} (D - D_{ii}I)_{11}, \prod_{i=1}^{n} (D - D_{ii}I)_{22}, \ldots, \prod_{i=1}^{n} (D - D_{ii}I)_{nn})$$

However, note that for any of these diagonal entries $\prod_{i=1}^{n}(D-D_{ii}I)_{jj}$, we have that $\prod_{i=1}^{n}(D-D_{ii}I)_{jj} = (\prod_{i=1,i\neq j}^{n}(D-D_{ii}I)_{jj})(D-D_{jj}I)_{jj} = (\prod_{i=1,i\neq j}^{n}(D-D_{ii}I)_{jj})(0) = 0$.
Thus, we have that $p_D(D) = \prod_{i=1}^{n}(D-D_{ii}I) = 0$. □

b

Claim. Cayley-Hamilton holds for diagonalizable matrices.

Proof. Let $f(A) = \sum_{i=0}^{n} c_i A^i$. Then, if $A = P^{-1}BP$, we will show that $f(A) = P^{-1}f(B)P$.

$$\begin{aligned}
P^{-1}f(B)P &= P^{-1}(\sum_{i=0}^{n} c_i B^i)P \\
&= (\sum_{i=0}^{n} P^{-1}(c_i B^i))P \\
&= \sum_{i=0}^{n} P^{-1}(c_i B^i)P \\
&= \sum_{i=0}^{n} c_i P^{-1} B^i P \\
&= \sum_{i=0}^{n} c_i A^i = f(A)
\end{aligned}$$

The last line comes from problem 2, part b, which doesn't actually rely on D being a diagonal matrix.

The first three lines come from the left/right distributivity of matrix multiplication.

Now, let $A = P^{-1}DP$, where D is diagonal. Then, since Apostol shows all similar matrices have the same characterizing polynomial, we put $p(\lambda)$ for the characterizing polynomials of both A and D. Then, $p(A) = P^{-1}p(D)P = P^{-1}0P = 0$. □

Problem 7

a

$$(1+i)^2 = 1^2 + 2i + i^2 = 2i$$

b

$$\frac{1}{i} = \frac{i}{i^2} = -i$$

c

$$\frac{1+i}{1-2i} = \frac{(1+i)(1+2i)}{(1-2i)(1+2i)} = \frac{1+3i+2i^2}{1-4i^2} = -\frac{1}{5} + \frac{3i}{5}$$

d

$$i^5 + i^{16} = (i^4)i + (i^4)^4 = i + 1$$

Problem 8

Let p be a polynomial over $\mathbb{R}$.

a

Claim.

$$\overline{f(z)} = f(\overline{z})$$

Proof. We will show that conjugation is a automorphism on $\mathbb{C}$.
First, we show that $\overline{zw} = \overline{z} \cdot \overline{w}$. Let $z = z_1 + z_2 i, w = w_1 + w_2 i$. Then,

$$\overline{zw} = \overline{z_1 w_1 - z_2 w_2 + (z_1 w_2 + z_2 w_1) i} = z_1 w_1 - z_2 w_2 - (z_1 w_2 + z_2 w_1) i$$

and

$$\overline{z} \cdot \overline{w} = (z_1 - z_2 i)(w_1 - w_2 i) = z_1 w_1 - z_2 w_2 - (z_1 w_2 + z_2 w_1) i$$

so we have that $\overline{zw} = \overline{z} \cdot \overline{w}$.
Further,

$$\overline{z + w} = z_1 + w_1 - (z_2 + w_2) i = z_1 - z_2 i + w_1 - w_2 i = \overline{z} + \overline{w}$$

Now, induct on the degree n of f. For $n = 1$, we have that $f = c_o$, which has the desired property immediately.
Now assume the hypothesis for $n = k$. Then,

$$\overline{f(z)} = \overline{\sum_{i=0}^{k+1} c_i z^i} = \overline{\sum_{i=0}^{k} c_i z^i + c_{k+1} z^{k+1}} = \overline{\sum_{i=0}^{k} c_i z^i} + \overline{c_{k+1} z^{k+1}} = \sum_{i=0}^{k} c_i \overline{z}^i + \overline{c} \overline{z^{k+1}} = \sum_{i=0}^{k+1} c_i \overline{z}^i = f(\overline{z})$$

$\square$

b

Claim. Any nonreal zeros of f must occur in conjugate pairs.

Proof. Suppose that $f(z) = 0, z \in \mathbb{C}$. Then, $f(\overline{z}) = \overline{0} = 0$. Thus, any nonreal zero of f must also have that its conjugate is a zero of f. In particular, if $z \notin \mathbb{R}$, $z \neq \overline{z}$ and so complex roots come in conjugate pairs. □

c

Claim. For $A \in M_{n\times n}(\mathbb{R})$, any nonreal eigenvectors of A occur in complex conjugate pairs.

Proof. Note that the characteristic polynomial of A was shown to be a real valued polynomial in problem 1. Then, as $p_A(\lambda) = \det(\lambda I - A)$, the eigenvalues of A are exactly the roots of p_A. Thus, from part b, any nonreal eigenvalues must come in conjugate pairs. □

Problem 9

Claim.

$$(x, y) = \frac{1}{4}(||x+y||^2 - ||x-y||^2)$$

Proof.

$$\begin{aligned}
\frac{1}{4}(||x+y||^2 - ||x-y||^2) &= \frac{1}{4}((x+y, x+y) - (x-y, x-y)) \\
&= \frac{1}{4}((x+y, x) + (x+y, y) - ((x-y, x) - (x-y, y))) \\
&= \frac{1}{4}((x, x) + (y, x) + (x, y) + (y, y) - ((x, x) - (y, x) - ((x, y) - (y, y)))) \\
&= \frac{1}{4}((x, x) + (x, y) + (x, y) + (y, y) - (x, x) + (x, y) + (x, y) - (y, y)) \\
&= \frac{1}{4}(4(x, y)) \\
&= (x, y)
\end{aligned}$$

□