0 Preliminaries

0.9 Problems

P.0.1 Let $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ be a function. Show that the following are equivalent: (a) f is one to one. (b) f is onto. (c) f is a permutation of $1, 2, \ldots, n$.

Solution. If S is a finite set, let $\#S$ denote the number of elements in S. The following arguments rely on the fact that if R and T are subsets of a finite set, then $\#(R \cup T) \leq \#R + \#T$, with equality if and only if R and T are disjoint.

(a) $\Rightarrow$ (b) If $n = 1$ there is nothing to prove. Proceed by induction. For each $k = 1, 2, \ldots, n$ let S_k be the statement $\#\{f(1), f(2), \ldots, f(k)\} = k$. Then $S_1 = \{f(1)\}$ contains one element, so S_1 is true. Assume that $1 \leq k \leq n-1$ and that S_k is true. Observe that

$$\{f(1), f(2), \ldots, f(k+1)\} = \{f(1), f(2), \ldots, f(k)\} \cup \{f(k+1)\}.$$

Because f is one to one, $\{f(k+1)\}$ is disjoint from $\{f(1), f(2), \ldots, f(k)\}$. Therefore, the induction hypothesis ensures that

$$\#\{f(1), f(2), \ldots, f(k+1)\} = \#\{f(1), f(2), \ldots, f(k)\} + \#\{f(k+1)\} = k + 1,$$

which shows that S_{k+1} is true. The principle of mathematical induction ensures that S_n is true, so $\#\{f(1), f(2), \ldots, f(n)\} = n$. Since

$$\{f(1), f(2), \ldots, f(n)\} \subseteq \{1, 2, \ldots, n\}$$

and both sets contain n elements, they are identical. This means that f is onto.

(b) $\Rightarrow$ (a) If $n = 1$ there is nothing to prove, so assume that $n \geq 2$. Let $k \in \{1, 2, \ldots, n\}$ be given and let $F_k = \{f(1), f(2), \ldots, f(n)\} \setminus \{f(k)\}$ denote the set obtained by omitting the element $f(k)$ from $\{f(1), f(2), \ldots, f(n)\}$. Since f is onto,

$$\{1, 2, \ldots, n\} = \{f(1), f(2), \ldots, f(n)\} = F_k \cup \{f(k)\}.$$

Therefore,

$$n = \#\{1, 2, \ldots, n\} = \#(F_k \cup \{f(k)\}) \leq \#F_k + \#\{f(k)\}, \tag{0.9.1}$$

with equality if and only if F_k and $\{f(k)\}$ are disjoint. Since $\#F_k \leq n-1$ and $\#\{f(k)\} = 1$, the inequality in (0.9.1) is an equality; we conclude that F_k and $\{f(k)\}$ are disjoint. Therefore, $f(k) \neq f(i)$ for all $i \in \{1, 2, \ldots, n\}$ such that $i \neq k$. Since $k \in \{1, 2, \ldots, n\}$ is arbitrary, it follows that f is one to one.

(a) $\Leftrightarrow$ (c) This is a definition.

P.0.2 Show that (a) the diagonal entries of a Hermitian matrix are real; (b) the diagonal entries of a skew-Hermitian matrix are purely imaginary; (c) the diagonal entries of a skew-symmetric matrix are zero.

Solution. (a) If A is Hermitian, then $A^* = [\overline{a}_{ji}] = [a_{ij}] = A$, so $\overline{a}_{jj} = a_{jj}$ (that is, each a_{jj} is real) for all $j = 1, 2, \ldots, n$.

(b) If A is skew-Hermitian, then $A^* = [\overline{a}_{ji}] = [-a_{ij}] = -A$, so $\overline{a}_{jj} = -a_{jj}$ (that is, each a_{jj} is pure imaginary) for all $j = 1, 2, \ldots, n$.

(c) If A is skew-symmetric, then $A^{\mathsf{T}} = [a_{ji}] = [-a_{ij}] = -A^{\mathsf{T}}$, so $a_{jj} = -a_{jj}$ (that is, each $a_{jj} = 0$) for all $j = 1, 2, \ldots, n$.

P.0.3 Use mathematical induction to prove that $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$ for $n = 1, 2, \ldots$.

Solution. Let $n \geq 1$ and let S_n be the statement that

$$\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)/6.$$

Then S_1 is the assertion that

$$1 = \frac{1(1+1)(2+1)}{6},$$

which is true. Let $n \geq 1$ and assume that S_n is true. Then

$$\begin{aligned}\sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^{n} k^2 + (n+1)^2 \\ &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{(n+1)(2n^2+7n+6)}{6} \\ &= \frac{(n+1)(n+2)(2n+3)}{6},\end{aligned}$$

which shows that S_{n+1} is true. The principle of mathematical induction ensures that S_n is true for all $n = 1, 2, \ldots$.

P.0.4 Use mathematical induction to prove that $1^3 + 2^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2$ for $n = 1, 2, \ldots$.

Solution. Let $n \geq 1$ and let S_n be the statement that $\sum_{k=1}^{n} k^3 = n^2(n+1)^2/4$.

The S_1 is the assertion that

$$1 = \frac{1^2 2^2}{4} = 1,$$

which is true. Let $n \geq 1$ and assume that S_n is true. Then

$$\begin{aligned}\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3 &= \frac{n^2(n+1)^2}{4} + (n+1)^3 \\ &= (n+1)^2 \frac{n^2 + 4(n+1)}{4} = \frac{(n+1)^2(n+2)^2}{4},\end{aligned}$$

which shows that S_{n+1} is true. The principle of mathematical induction ensures that S_n is true for all $n = 1, 2, \ldots$.

P.0.5 Let $A \in \mathsf{M}_n$ be invertible. Use mathematical induction to prove that $(A^{-1})^k = (A^k)^{-1}$ for all integers k.

Solution. For each $k \in \mathbb{Z}$ we must prove that $A^k(A^{-1})^k = (A^{-1})^k A^k = I$; denote this statement by S_k. Then S_0 is true because $A^0 = I$ by definition, $A^0(A^{-1})^0 = II = I$, and $(A^{-1})^0 A^0 = II = I$. If $k \in \mathbb{Z}$ is negative, then by definition $A^k(A^{-1})^k = (A^{-1})^{-k}((A^{-1})^{-1})^{-k} = (A^{-1})^{-k} A^{-k}$. Therefore, it suffices to prove that S_k is true for each positive integer k. The statement S^1 is $A(A^{-1}) = (A^{-1})A = I$, which is true. Assume that $k \geq 1$ and S^k is true. Then S_{k+1} is the statement

$$A^{k+1}(A^{-1})^{k+1} = (A^{-1})^{k+1} A^{k+1} = I.$$

The induction hypothesis ensures that

$$A^{k+1}(A^{-1})^{k+1} = A(A^k(A^{-1})^k)A^{-1} = AIA^{-1} = I$$

and

$$(A^{-1})^{k+1} A^{k+1} = A^{-1}((A^{-1})^k A^k)A = A^{-1}IA = I,$$

so S_{k+1} is true. The principle of mathematical induction ensures that S_k is true for all $k = 1, 2, \ldots$.

P.0.6 Let $A \in \mathsf{M}_n$. Use mathematical induction to prove that $A^{j+k} = A^j A^k$ for all integers j, k.

Solution. Since A is not assumed to be invertible, we prove this assertion for all $j, k \in \mathbb{N}$. Let $j \in \mathbb{N}$ be given and let S_k be the statement that $A^{j+k} = A^j A^k$. Then $A^{j+0} = A^j = A^j I = A^j A^0$, so S_0 is true. Suppose that S_k is true for some $k \geq 0$. Then $A^{j+k+1} = (A^{j+k})A = (A^j A^k)A = A^j A^k A = A^j A^{k+1}$, so S_{k+1} is true. The principle of mathematical induction ensures that S_k is true for all $k \in \mathbb{N}$.

An alternative approach is to invoke the associativity of matrix multiplication:

If $j, k \geq 1$, then

$$A^{j+k} = \underbrace{A \cdots A}_{j+k \text{ factors}} = \underbrace{A \cdots A}_{j \text{ factors}} \underbrace{A \cdots A}_{k \text{ factors}} = (\underbrace{A \cdots A}_{j \text{ factors}})(\underbrace{A \cdots A}_{k \text{ factors}}) = A^j A^k.$$

If $j = 0$, then $A^{0+k} = A^k = IA^k = A^0 A^k$. If $k = 0$, then $A^{j+0} = A^j = A^j I = A^j A^0$.

P.0.7 Use mathematical induction to prove Binet's formula (9.5.5) for the Fibonacci numbers.

Solution. Define f_k by $f_1 = f_2 = 1$ and $f_{k+1} = f_k + f_{k-1}$ for $k = 2, 3, \ldots$. Let $\phi = (1 + \sqrt{5})/2$ and $\tau = (1 - \sqrt{5})/2$. Compute $(\phi - \tau)/\sqrt{5} = 1 = f_1$ and $(\phi^2 - \tau^2)/\sqrt{5} = 1 = f_2$. We must show that $(\phi^k - \tau^k)/\sqrt{5} = f_k$ for all $k \geq 2$. Suppose that $z \in \mathbb{C}$ and $z^2 - z - 1 = 0$, that is, $z^2 = z + 1$; check that ϕ and τ satisfy this equation. Notice that $z^2 = f_2 z + f_1$ and

$$\begin{aligned} z^3 &= z(f_2 z + f_1) = f_2 z^2 + f_1 z \\ &= f_2(z+1) + f_1 z = (f_2 + f_1)z + f_2 \\ &= f_3 z + f_2. \end{aligned}$$

Let $k \geq 2$ and let S_k be the statement that $t^k = f_k z + f_{k-1}$. We have shown that S_1 and S_2 are true. If $k \geq 2$ and S_k is true, then

$$\begin{aligned} z^{k+1} &= z(f_k z + f_{k-1}) = f_k z^2 + f_{k-1} z \\ &= f_k(z+1) + f_{k-1} z = (f_k + f_{k-1})z + f_k \\ &= f_{k+1} z + f_k, \end{aligned}$$

so S_{k+1} is true. The principle of mathematical induction ensures that S_k is true for all $k = 1, 2, \ldots$.

Since ϕ and τ satisfy the equation $z^2 - z - 1 = 0$, we have

$$\phi^k = f_k \phi + f_{k-1}$$

and

$$\tau^k = f_k \tau + f_{k-1}$$

for all $k = 1, 2, \ldots$. Therefore, $\phi^k - \tau^k = f_k(\phi - \tau) = f_k \sqrt{5}$ and hence

$$f_k = \frac{\phi^k - \tau^k}{\sqrt{5}}$$

for all $k = 1, 2, \ldots$.

P.0.8 Use mathematical induction to prove that $1 + z + z^2 + \cdots + z^{n-1} = \frac{1-z^n}{1-z}$ for complex $z \neq 1$ and all positive integers n.

Solution. Let S_n be the statement that $(1-z)(1 + z + \cdots + z^{n-1}) = 1 - z^n$. Then

S_1 is the statement that $(1-z)(1) = 1-z$, which is true. Suppose that $n \geq 2$ and S_n is true. Then

$$\begin{aligned}(1-z)(1+z+\cdots+z^{n-1}+z^n) &= (1-z)(1+z+\cdots+z^{n-1}) + (1-z)z^n \\ &= (1-z^n) + z^n - z^{n+1} \\ &= 1 - z^{n+1},\end{aligned}$$

which shows that S_{n+1} is true. The principle of mathematical induction ensures that S_n is true for all $n = 1, 2, \ldots$. If $z \neq 1$, it follows that

$$1 + z + \cdots + z^{n-1} + z^n = \frac{1-z^{n+1}}{1-z}$$

for all $n = 1, 2, \ldots$.

P.0.9 (a) Compute the determinants of the matrices

$$V_2 = \begin{bmatrix} 1 & z_1 \\ 1 & z_2 \end{bmatrix}, \qquad V_3 = \begin{bmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{bmatrix}, \qquad V_4 = \begin{bmatrix} 1 & z_1 & z_1^2 & z_1^3 \\ 1 & z_2 & z_2^2 & z_2^3 \\ 1 & z_3 & z_3^2 & z_3^3 \\ 1 & z_4 & z_4^2 & z_4^3 \end{bmatrix},$$

and simplify your answers as much as possible. (b) Use mathematical induction to evaluate the determinant of the $n \times n$ *Vandermonde matrix*

$$V_n = \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_n & z_n^2 & \cdots & z_n^{n-1} \end{bmatrix}. \tag{0.9.2}$$

(c) Find conditions on $z_1, z_2, \ldots, z_n$ that are necessary and sufficient for V_n to be invertible.

Solution. (a) To compute $\det V_2$, subtract z_1 times the first column from the second column:

$$\det V_2 = \det \begin{bmatrix} 1 & 0 \\ 1 & z_2 - z_1 \end{bmatrix} = z_2 - z_1.$$

To compute $\det V_3$, subtract z_1 times the third column from the fourth column, subtract z_1 times the first column from the second column, expand by minors across the first row, factor each row, pull out the factors, and use the 2×2 case:

$$\det V_3 = \det \begin{bmatrix} 1 & z_1 & 0 \\ 1 & z_2 & z_2^2 - z_2 z_1 \\ 1 & z_3 & z_3^2 - z_3 z_1 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ 1 & z_2 - z_1 & z_2^2 - z_2 z_1 \\ 1 & z_3 - z_1 & z_3^2 - z_3 z_1 \end{bmatrix}$$

$$= \det \begin{bmatrix} z_2 - z_1 & z_2^2 - z_2 z_1 \\ z_3 - z_1 & z_3^2 - z_3 z_1 \end{bmatrix} = (z_2 - z_1)(z_3 - z_1) \det \begin{bmatrix} 1 & z_2 \\ 1 & z_3 \end{bmatrix}$$

$$= (z_3 - z_2)(z_3 - z_1)(z_2 - z_1)$$

$$= \prod_{\substack{i,j=1,2,3 \\ i>j}} \prod_{\substack{i,j=1,2,3 \\ i>j}} (z_i - z_j).$$

To compute $\det V_4$, proceed as in the 3×3 case to create zero entries in the first row. Subtract a suitable multiple of a column from the column to its right, starting at the right. Expand by minors along the first row, remove a factor from each row, and use the result for the 3×3 case:

$$\det V_4 = \det \begin{bmatrix} 1 & z_1 & z_1^2 & z_1^3 \\ 1 & z_2 & z_2^2 & z_2^3 \\ 1 & z_3 & z_3^2 & z_3^3 \\ 1 & z_4 & z_4^2 & z_4^3 \end{bmatrix} = \det \begin{bmatrix} 1 & z_1 & z_1^2 & 0 \\ 1 & z_2 & z_2^2 & z_2^3 - z_2^2 z_1 \\ 1 & z_3 & z_3^2 & z_3^3 - z_3^2 z_1 \\ 1 & z_4 & z_4^2 & z_4^3 - z_4^2 z_1 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & z_1 & 0 & 0 \\ 1 & z_2 & z_2^2 - z_2 z_1 & z_2^3 - z_2^2 z_1 \\ 1 & z_3 & z_3^2 - z_3 z_1 & z_3^3 - z_3^2 z_1 \\ 1 & z_4 & z_4^2 - z_{4 z_1} & z_4^3 - z_4^2 z_1 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & z_2 - z_1 & z_2^2 - z_2 z_1 & z_2^3 - z_2^2 z_1 \\ 1 & z_3 - z_1 & z_3^2 - z_3 z_1 & z_3^3 - z_3^2 z_1 \\ 1 & z_4 - z_4 & z_4^2 - z_{4 z_1} & z_4^3 - z_4^2 z_1 \end{bmatrix}$$

$$= \det \begin{bmatrix} z_2 - z_1 & z_2^2 - z_2 z_1 & z_2^3 - z_2^2 z_1 \\ z_3 - z_1 & z_3^2 - z_3 z_1 & z_3^3 - z_3^2 z_1 \\ z_4 - z_1 & z_4^2 - z_1 & z_4^3 - z_4^2 z_1 \end{bmatrix}$$

$$= (z_4 - z_1)(z_3 - z_1)(z_2 - z_1) \det \begin{bmatrix} 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \\ 1 & z_4 & z_4^2 \end{bmatrix}$$

$$= (z_4 - z_1)(z_3 - z_1)(z_2 - z_1) \prod_{\substack{i,j=2,3,4 \\ i>j}} \prod_{\substack{i,j=2,3,4 \\ i>j}} (z_i - z_j)$$

$$= \prod_{\substack{i,j=1,2,3,4 \\ i>j}} \prod_{\substack{i,j=1,2,3,4 \\ i>j}} (z_i - z_j).$$

(b) Let $n \geq 2$ and let S_n be the statement that

$$\det V_n = \prod_{\substack{i,j=1,2,\ldots,n \\ i>j}} \prod_{\substack{i,j=1,2,\ldots,n \\ i>j}} (z_i - z_j).$$

We have shown that S_n is true for $n = 2, 3, 4$. Suppose that $n \geq 4$ and S_n is true. Use the column-wise elimination process demonstrated in the preceding cases and the induction hypothesis to obtain

$$\begin{aligned}
\det V_{n+1} &= \det \begin{bmatrix} 1 & z_1 & \cdots & z_1^{n-1} & z_1^n \\ 1 & z_2 & \cdots & z_2^{n-1} & z_2^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & z_{n+1} & \cdots & z_{n+1}^{n-1} & z_{n+1}^n \end{bmatrix} \\
&= \det \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & z_2 - z_1 & \cdots & z_2^{n-1} - z_2^{n-2} z_1 & z_2^n - z_2^n z_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & z_{n+1} - z_1 & \cdots & z_{n+1}^{n-1} - z_{n+1}^{n-2} z_1 & z_{n+1}^n - z_{n+1}^n z_1 \end{bmatrix} \\
&= (z_{n+1} - z_1)(z_n - z_1) \cdots (z_2 - z_1) \det \begin{bmatrix} 1 & z_2 & \cdots & z_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{n+1} & \cdots & z_{n+1}^{n-1} \end{bmatrix} \\
&= (z_{n+1} - z_1)(z_n - z_1) \cdots (z_2 - z_1) \prod_{\substack{i,j=2,3,\ldots,n+1 \\ i>j}} \prod_{\substack{i,j=2,3,\ldots,n+1 \\ i>j}} (z_i - z_j) \\
&= \prod_{\substack{i,j=1,2,\ldots,n+1 \\ i>j}} \prod_{\substack{i,j=1,2,\ldots,n+1 \\ i>j}} (z_i - z_j).
\end{aligned}$$

This shows that S_{n+1} is true. The principle of mathematical induction ensures that S_n is true for all $n = 1, 2, \ldots$.

(c) The formula for $\det S_n$ shows that V_n is invertible if and only if $z_i \neq z_j$ for all $i, j = 1, 2, \ldots, n$ such that $i \neq j$.

P.0.10 Consider the polynomial $p(z) = c_k z^k + c_{k-1} z^{k-1} + \cdots + c_1 z + c_0$, in which $k \geq 1$, each coefficient c_i is a nonnegative integer, and $c_k \geq 1$. Prove the following statements: (a) $p(t+2) = c_k t^k + d_{k-1} t^{k-1} + \cdots + d_1 t + d_0$, in which each d_i is a nonnegative integer and $d_0 \geq 2^k$. (b) $p(nd_0 + 2)$ is divisible by d_0 for each $n = 1, 2, \ldots$. (c) $p(n)$ is not a prime for infinitely many positive integers n. This was proved by C. Goldbach in 1752.

Solution. (a) Compute

$$\begin{aligned}
p(t+2) &= c_k (t+2)^k + c_{k-1} (t+2)^{k-1} + \cdots + c_1 (t+2) + c_0 \\
&= c_k (t^k + \cdots + 2^k) + c_{k-1} (t^{k-1} + \cdots + 2^{k-1}) + \cdots c_1 (t+2) + c_0 \\
&= c_k t^k + d_{k-1} t^{k-1} + \cdots + d_1 t + d_0,
\end{aligned}$$

in which each d_j is a nonnegative integer because it is a sum of nonnegative integer multiples of the integers $c_0, c_1, \ldots, c_k$. Since each $c_i \geq 0$, $c_k \geq 1$, and $k \geq 1$, we have

$$d_0 = c_k 2^k + c_{k-1}2^{k-1} + \cdots + 2c_1 + c_0 \geq c_k 2^k \geq 2^k > 1.$$

(b) For each positive integer n, $p(nd_0 + 2)$ is a sum

$$p(nd_0 + 2) = c_k(nd_0)^k + d_{k-1}(nd_0)^{k-1} + \cdots + d_1(nd_0) + d_0,$$

in which each summand is either zero or a positive integer divisible by d_0. Therefore, $p(nd_0 + 2)$ is divisible by the positive integer $d_0 > 1$.

(c) In (b) we have exhibited infinitely many positive integers m (namely, $m = nd_0+2$ for $n = 1, 2, \ldots$) such that $p(m)$ is not prime.

P.0.11 If p is a real polynomial, show that $p(\lambda) = 0$ if and only if $p(\overline{\lambda}) = 0$.

Solution. Let $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, in which $a_0, a_1, \ldots, a_n$ are real. If $p(\lambda) = 0$, then

$$\begin{aligned} 0 = \overline{0} = \overline{p(\lambda)} &= \overline{a_n}\,\overline{\lambda}^n + \overline{a_{n-1}}\,\overline{\lambda}^{n-1} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} \\ &= a_n\overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \cdots + a_1\overline{\lambda} + a_0 \\ &= p(\overline{\lambda}). \end{aligned}$$

If λ is a non-real root of p, could λ have multiplicity 2 while $\overline{\lambda}$ has multiplicity 3? This problem provides no information about the answer to this question, but the following problem shows why λ and $\overline{\lambda}$ have the same multiplicities as zeros of p.

P.0.12 Show that a real polynomial can be factored into real linear factors and real quadratic factors that have no real zeros.

Solution. Let p be a real polynomial of degree $n \geq 1$. If $n = 1$ then $p(z) = c_1 z + c_0$, $c_1 \neq 0$, and the real number $-c_0/c_1$ is the only zero of p. Now suppose that $n \geq 2$. If a non-real complex number μ_1 is a zero of p, the preceding problem ensures that $\overline{\mu_1}$ is also a zero of p. Therefore, p is divisible by $(z - \mu_1)$, by $(z - \overline{\mu_1})$, and therefore by their product, which is the real quadratic polynomial

$$g(z, \mu_1) = (z - \mu_1)(z - \overline{\mu_1}) = z^2 - 2(\operatorname{Re}\mu_1)z + |\mu_1|^2,$$

that is, $p(z) = g(z, \mu_1)q_{n-2}(z)$, in which the quotient q_{n-2} is a real polynomial of degree $n - 2$. If q_{n-2} has any non-real zeros, let μ_2 be one of them. The preceding argument shows that $q_{n-2}(z) = g(z, \mu_2)q_{n-4}(z)$, in which the quotient q_{n-4} is a real polynomial of degree $n - 4$ and $p(z) = g(z, \mu_1)g(z, \mu_2)q_{n-4}(z)$. Continue this process until the quotient has no non-real zeroes, that is,

$$p(z) = g(z, \mu_1)g(z, \mu_2)\cdots g(z, \mu_k)q_{n-2k}(z),$$

in which q_{n-2k} is a real polynomial of degree $n-2k$ that has no non-real zeros. If $n =$

$2k$, then $q_{n-2k}(z) = c$ is a nonzero scalar and $p(z) = cg(z, \mu_1)g(z, \mu_2)\cdots g(z, \mu_k)$. If $n > 2k$, then $q_{n-2k}(z)$ has only real zeros $\lambda_1, \lambda_2, \ldots, \lambda_{n-2k}$ and $q_{n-2k}(z) = c(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_{n-2k})$ for some nonzero scalar c. In this case,

$$p(z) = cg(z, \mu_1)g(z, \mu_2)\cdots g(z, \mu_k)(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_{n-2k}).$$

This argument shows that each non-real zero of p has the same multiplicity as its complex conjugate.

P.0.13 Show that every real polynomial of odd degree has a real zero. *Hint*: Use the Intermediate Value Theorem.

Solution. Let $p(z) = c_n z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$, in which $n \geq 1$ is odd, $c_n \neq 0$, and all the coefficients are real. Let $t \in \mathbb{R}$ be nonzero and define

$$g(t) = c^{n-1}t^{-1} + c_{n-2}t^{-2} + \cdots + c_1 t^{-n+1} + c_0 t^{-n}.$$

Then $p(t) = t^n(c_n + g(t))$. Since $\lim_{t\to\pm\infty} g(t) = 0$, for sufficiently large positive or negative M, the value $p(M)$ has the same sign as $M^n c_n$ and $p(-M)$ has the same sign as $(-M)^n c_n$. Since n is odd, M^n and $(-M)^n$ (and therefore also $p(M)$ and $p(-M)$) have opposite signs. Since p is a continuous real valued function, the intermediate value theorem ensures that $p(t) = 0$ for some $t \in [-M, M]$.

P.0.14 Let $h(z)$ be a polynomial and suppose that $z(z-1)h(z) = 0$ for all $z \in [0, 1]$. Prove that h is the zero polynomial.

Solution. Since $z(z-1)h(z) = 0$ for all $z \in [0, 1]$ and $z(z-1) \neq 0$ for all $z \in (0, 1)$, it follows that $h(z) = 0$ for all $z \in (0, 1)$. A polynomial has infinitely many zeros if and only if it is the zero polynomial, so we conclude that h is the zero polynomial.

P.0.15 (a) Prove that the $n \times n$ Vandermonde matrix (0.9.2) is invertible if and only if the n complex numbers $z_1, z_2, \ldots, z_n$ are distinct. *Hint*: Consider the system $V_n\mathbf{c} = \mathbf{0}$, in which $\mathbf{c} = [c_0\ c_1\ \ldots\ c_{n-1}]^\mathsf{T}$, and the polynomial $p(z) = c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$. (b) Use (a) to prove the Lagrange Interpolation Theorem (Theorem 0.7.6).

Solution. (a) The assertion has already been proved in P.0.9, but the hint directs us to give a different proof. Let $n \geq 2$, let $\mathbf{c} = [c_0\ c_1\ \ldots\ c_{n-1}]^\mathsf{T} \in \mathbb{C}^n$, and let $p(z) = c_{n-1}z^{n-1} + c_{n-2}z^{n-2} + \cdots + c_1 z + c_0$. Observe that $V_n\mathbf{c} = [p(z_1)\ p(z_2)\ \ldots\ p(z_n)]^\mathsf{T}$.

Suppose that $z_1, z_2, \ldots, z_n$ are distinct. If V_n is not invertible then there is a nonzero vector $\mathbf{c}$ such that $V_n\mathbf{c} = \mathbf{0}$, and hence $p(z_1) = p(z_2) = \cdots = p(z_n) = 0$. But p is a polynomial of degree at most $n-1$, so it has more than $n-1$ distinct zeros if and only if it is the zero polynomial, that is, if and only if $c_0 = c_1 = \cdots = c_{n-1} = 0$, which is not possible since $\mathbf{c} \neq \mathbf{0}$. This contradiction shows that V_n must be invertible.

Conversely, suppose that $z_1, z_2, \ldots, z_n$ are not distinct. Then two rows of V_n are identical, so $\det V_n = 0$ and V_n is not invertible. This shows that $z_1, z_2, \ldots, z_n$ are distinct if and only if V_n is invertible.

(b) Using the notation of (a), the Lagrange Interpolation Theorem says that if $z_1, z_2, \ldots, z_n$ are distinct, then the linear system $V_n\mathbf{c} = [p(z_1)\ p(z_2)\ \ldots\ p(z_n)]^\mathsf{T} = \mathbf{w}$ has a unique solution $\mathbf{c}$ for any given $\mathbf{w}$. A linear system has a unique solution for any given right-hand side if and only if its coefficient matrix is invertible, and part (a) ensures that V_n is invertible if $z_1, z_2, \ldots, z_n$ are distinct. If $\mathbf{w}$ and the distinct values $z_1, z_2, \ldots, z_n$ are real, then $\mathbf{c} = V_n^{-1}\mathbf{w}$ is real, so the interpolating polynomial p has real coefficients.

P.0.16 If c is a nonzero scalar and p, q are nonzero polynomials, show that (a) $\deg(cp) = \deg p$, (b) $\deg(p + q) \leq \max\{\deg p, \deg q\}$, and (c) $\deg(pq) = \deg p + \deg q$. What happens if p is the zero polynomial?

Solution. Suppose that m, n are nonnegative integers, $p(z) = a_m z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$, $q(z) = b_n z^n + b_{n-1}z^{-1} + \cdots + b_1 z + b_0$, $a_m b_n \neq 0$, and $c \neq 0$. Then $\deg p = m$ and $\deg q = n$. (a) $cp(z) = ca_m z^m + \cdots$ and $ca_m \neq 0$, so $\deg(cp) = m = \deg p$. (b) $p(z) + q(z) = a_m z^m + \cdots + b_n z^n + \cdots$. If $m \neq n$, the highest order nonzero term in $p+q$ is either $a_m z^m$ or $b_n z^n$, so $\deg(p+q) = \max\{m, n\} = \max\{\deg p, \deg q\}$. If $m = n$, the highest order term in $p + q$ is $(a_n + b_n)z^n$ if $a_n + b_n \neq 0$, in which case $\deg(p + q) = n = \max\{\deg p, \deg q\}$. If $m = n$, $a_n + b_n = 0$, and $p + q \neq 0$, the highest order term in $p + q$ has nonnegative degree less than n, so $\deg(p+q) < n = \max\{\deg p, \deg q\}$. If $m = n$ and $p+q = 0$, then $-\infty = \deg(p+q) < \deg p + \deg q$. Therefore, in all cases we have $\deg(p + q) \leq \max\{\deg p, \deg q\}$. (c) $p(z)q(z) = a_m b_n z^{m+n} + \cdots$, so $\deg(pq) = m + n = \deg p + \deg q$.

If p is the zero polynomial, calculations in the extended real number system show that (a) cp is the zero polynomial, so $\deg(cp) = -\infty = \deg(p)$; (b) $p + q = q$, so $\deg(p + q) = \deg q = \max\{-\infty, \deg q\} = \max\{\deg p, \deg q\}$; (c) pq is the zero polynomial, so $\deg(pq) = -\infty = -\infty + \deg q = \deg p + \deg q$.

The problem does not ask, "What happens if p and q are both zero?", but if they are in (a) we have both p and cp zero polynomials, so $-\infty = \deg(cp) = \deg p$; in (b) we have p, q, and $p+q$ all zero polynomials, so $-\infty = \deg(p+q) = \max\{-\infty, -\infty\} = \max\{\deg p, \deg q\}$; in (c) we have p, q, and pq all zero polynomials, so $-\infty = \deg(pq) = -\infty + (-\infty) = \deg p + \deg q$.

P.0.17 Prove the uniqueness assertion of the division algorithm. That is, if f and g are polynomials such that $1 \leq \deg g \leq \deg f$ and if q_1, q_2, r_1 and r_2 are polynomials such that $\deg r_1 < \deg g$, $\deg r_2 < \deg g$, and $f = gq_1 + r_1 = gq_2 + r_2$, then $q_1 = q_2$ and $r_1 = r_2$.

Solution. If $f = gq_1 + r_1 = gq_2 + r_2$, then $0 = g(q_1 - q_2) + (r_1 - r_2)$ so that $g(q_1 - q_2) = r_2 - r_1$. If $r_2 - r_1 = 0$, then the assumption that $g \neq 0$ ($\deg g \geq 1$)

ensures that $q_1 - q_2 = 0$ and we have uniqueness. Now suppose that $r_2 - r_1 \neq 0$, so $\deg(r_2 - r_1) \geq 0$, which implies that $q_1 - q_2 \neq 0$. We are given that $\deg r_1 < \deg g$ and $\deg r_2 < \deg g$, so part (c) of the preceding problem ensures that

$$\deg g > \deg(r_2 - r_1) = \deg(g(q_1 - q_2)) = \deg g + \deg(q_1 - q_2) \geq \deg g,$$

that is, $\deg g > \deg g$. This contradiction ensures that $r_2 - r_1 \neq 0$ is impossible, so $r_2 - r_1 = 0$ is the only possibility and we have uniqueness.

P.0.18 Give an example of a nonconstant function $f : \mathbb{R} \to \mathbb{R}$ such that $f(t) = 0$ for infinitely many distinct values of t. Is f a polynomial?

Solution. $f(t) = \sin t$ is a real-valued function that has infinitely many real zeros. It is not a polynomial.

P.0.19 Let $A = \operatorname{diag}(1, 2)$ and $B = \operatorname{diag}(3, 4)$. If $X \in \mathsf{M}_2$ intertwines A and B, what can you say about X? For a generalization, see Theorem 10.4.1.

Solution. Let

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The intertwining relation $AX - BX = 0$ in this case is

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which is

$$\begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix} - \begin{bmatrix} 3a & 4b \\ 3c & 4d \end{bmatrix} = -\begin{bmatrix} 2a & 3b \\ c & 2d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore, $a = b = c = d = 0$ and $X = 0$.

P.0.20 Verify the identity (0.5.2) for a 2×2 matrix, and show that the identity (0.3.4) is (0.5.3).

Solution. If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$\operatorname{adj} A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Compute

$$A \operatorname{adj} A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (\det A)I$$

and

$$(\operatorname{adj} A)A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (\det A)I.$$

If $\det A \neq 0$, then $A((\det A)^{-1} \operatorname{adj} A) = ((\det A)^{-1} \operatorname{adj} A)A = I$, so $A^{-1} = (\det A)^{-1} \operatorname{adj} A$.

P.0.21 Deduce (0.5.3) from the identity (0.5.2).

Solution. Suppose that $\det A \neq 0$ and let $B = (\det A)^{-1} \operatorname{adj} A$. Since $AB = BA = I$, B is, by definition, the inverse of A.

P.0.22 Deduce the second assertion in Theorem 0.8.1 from the first.

Solution. Let $X = B = A$. Then $AA = AA$, so the first assertion becomes $p(A)A = Ap(A)$ in this case. This is the second assertion.

P.0.23 Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$. Show that $AB = AC$ even though $B \neq C$.

Solution. Compute

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} = AC,$$

in which $B \neq C$. We cannot cancel A in the equation $AB = AC$ (that is, we cannot multiply both sides by A^{-1}) because A is not invertible.

P.0.24 Let $A \in \mathsf{M}_n$. Show that A is idempotent if and only if $I - A$ is idempotent.

Solution. We have

$$(I - A)^2 = I - 2A + A^2 = (I - A) + (A^2 - A).$$

Therefore, $(I - A)^2 = (I - A)$ if and only if $A^2 - A = 0$. That is, $I - A$ is idempotent if and only if A is idempotent.

P.0.25 Let $A \in \mathsf{M}_n$ be idempotent. Show that A is invertible if and only if $A = I$.

Solution. If $A^2 = A$ and A is invertible, then $A = A^{-1}A^2 = A^{-1}A = I$. If $A = I$ then A is an invertible idempotent matrix.

P.0.26 Let $A, B \in \mathsf{M}_n$ be idempotent. Show that $\operatorname{tr}((A - B)^3) = \operatorname{tr}(A - B)$.

Solution. Compute

$$\begin{aligned}(A - B)^3 &= (A - B)(A^2 - AB - BA + B^3) = (A - B)(A - AB - BA + B) \\ &= A^2 - A^2B - ABA + AB - BA + BAB + B^2A - B^2 \\ &= A - B + BAB - ABA.\end{aligned}$$

Therefore,

$$\begin{aligned}\operatorname{tr}(A-B)^3 &= \operatorname{tr}(A-B)+\operatorname{tr}(BAB-ABA)\\ &= \operatorname{tr}(A-B)+\operatorname{tr}(AB^2-A^2B)\\ &= \operatorname{tr}(A-B)+\operatorname{tr}(AB-AB)\\ &= \operatorname{tr}(A-B).\end{aligned}$$

1 Vector Spaces

1.7 Problems

P.1.1 In the spirit of the examples in Section 1.2, explain how $\mathcal{V} = \mathbb{C}^n$ can be thought of as a vector space over $\mathbb{R}$. Is $\mathcal{V} = \mathbb{R}^n$ a vector space over $\mathbb{C}$?

Solution. Vector addition and scalar multiplication are defined entrywise as addition and scalar multiplication of the real and imaginary parts of each entry. That is, if $\mathbf{v} = [a_1 + b_1 i \;\; a_2 + b_2 i \;\; \ldots \;\; a_n + b_n i]^\mathsf{T}$ and $\mathbf{w} = [c_1 + d_1 i \;\; c_2 + d_2 i \;\; \ldots \;\; c_n + d_n i]^\mathsf{T}$ are in $\mathcal{V}$ and $c \in \mathbb{R}$, then

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} a_1 + c_1 + (b_1 + d_1)i \\ a_2 + c_2 + (b_2 + d_2)i \\ \vdots \\ a_n + c_n + (b_n + d_n)i \end{bmatrix} \text{ and } c\mathbf{v} = \begin{bmatrix} ca_1 + cb_1 i \\ ca_2 + cb_2 i \\ \vdots \\ ca_n + cb_n i \end{bmatrix}.$$

The zero vector is $[0 \; 0 \; \ldots \; 0]^\mathsf{T}$.

$\mathcal{V} = \mathbb{R}^n$ is *not* a vector space over $\mathbb{C}$. For example, $\mathbf{u} = [1 \; 1 \; \ldots \; 1]^\mathsf{T} \in \mathcal{V}$ and $i \in \mathbb{C}$ but $i\mathbf{u} = [i \; i \; \ldots \; i]^\mathsf{T} \notin \mathcal{V}$.

P.1.2 Let $\mathcal{V}$ be the set of real 2×2 matrices of the form $\mathbf{v} = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix}$. Define $\mathbf{v} + \mathbf{w} = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & w \\ 0 & 1 \end{bmatrix}$ (ordinary matrix multiplication) and $c\mathbf{v} = \begin{bmatrix} 1 & cv \\ 0 & 1 \end{bmatrix}$. Show that $\mathcal{V}$ together with these two operations is a real vector space. What is the zero vector in $\mathcal{V}$?

Solution. We show that the eight axioms hold.

(i) We have

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v + w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \mathbf{v}$$

if and only if $w = 0$, that is, if and only if $\mathbf{w}$ is the identity matrix. Since $I_2 \in \mathcal{V}$, we see that $\mathcal{V}$ has a zero vector, namely, I_2.

(ii) We have

$$\mathbf{v}+\mathbf{w}=\begin{bmatrix}1 & v\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & w\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & v+w\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & w+v\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & w\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & v\\ 0 & 1\end{bmatrix}=\mathbf{w}+\mathbf{v}$$

so vector addition is commutative.

(iii) Matrix multiplication is associative so vector addition in $\mathcal{V}$ is also associative.

(iv) We have

$$\mathbf{v}+\mathbf{w}=\begin{bmatrix}1 & v\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & w\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & v+w\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

if and only if $w=-v$, that is, if and only if

$$\mathbf{w}=\begin{bmatrix}1 & -v\\ 0 & 1\end{bmatrix}.$$

Thus, additive inverses exist and are unique.

(v) We have

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

(vi) We have

$$a(b\mathbf{v})=a\left(\begin{bmatrix}1 & bv\\ 0 & 1\end{bmatrix}\right)=\begin{bmatrix}1 & abv\\ 0 & 1\end{bmatrix}=(ab)\mathbf{v}.$$

(vii) We have

$$\begin{aligned}c(\mathbf{v}+\mathbf{w})&=c\left(\begin{bmatrix}1 & v\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & w\\ 0 & 1\end{bmatrix}\right)\\ &=c\begin{bmatrix}1 & v+w\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & c(v+w)\\ 0 & 1\end{bmatrix}\\ &=\begin{bmatrix}1 & cv+cw\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & cv\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & cw\\ 0 & 1\end{bmatrix}=c\mathbf{v}+c\mathbf{w}.\end{aligned}$$

(viii) We have

$$(a+b)\mathbf{v}=\begin{bmatrix}1 & (a+b)v\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & av+bv\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & av\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & bv\\ 0 & 1\end{bmatrix}=a\mathbf{v}+b\mathbf{v}.$$

P.1.3 Show that the intersection of any (possibly infinite) collection of subspaces of an $\mathbb{F}$-vector space is a subspace.

Solution. Let $\mathcal{V}$ be a $\mathbb{F}$-vector space and let $\{\mathcal{U}_\alpha : \alpha \in I\}$ be a collection of subspaces of $\mathcal{V}$; I is some index set. Let

$$\mathcal{W}=\bigcap_{\alpha\in I}\mathcal{U}_\alpha$$

Theorem 1.3.3 ensures that it is sufficient to show that $c\mathbf{u}+\mathbf{v}\in\mathcal{W}$ whenever $\mathbf{u},\mathbf{v}\in\mathcal{W}$ and $c\in\mathbb{F}$. Let $\mathbf{u},\mathbf{v}\in\mathcal{W}$ and $c\in\mathbb{F}$. Then for all $\alpha\in I$, $\mathbf{u},\mathbf{v}\in\mathcal{U}_\alpha$.