Math 354
Homework

Charlie Cruz

Proof: We are told to assume the result of problem 3.D.4.
Now, since this is an if and only if, we need to prove both ways:

Proof of $\implies$:   We are given that there exists invertible operators $R \in \mathscr{L}(V), S \in \mathscr{L}(W)$ such that $T_1 = ST_2R$.
As S is invertible, we know that S^{-1} exists.
Thus, we can write:

$$S^{-1}T_1 = S^{-1}ST_2R \quad \text{Multiply both sides by } S^{-1}$$
$$= T_2R \quad \text{Identity property of linear maps (Axler 3.9)}$$

Thus, by problem 3.D.4, we know that $\ker T_1 = \ker T_2R$ as $S^{-1} \in \mathscr{L}(W)$
Now we want to show that $\ker T_2 = \ker T_2R$.
Pick an arbitrary $\vec{v} \in \ker T_2R$, then we know that $T_2R(\vec{v}) = \vec{0}$.
This implies that $T_2(R(\vec{v})) = \vec{0}$, which means that $R(\vec{v}) \in \ker T_2$.
In other words, for all $\vec{v} \in \ker T_2R$, R maps it to $\ker T_2$ as invertible maps are injective (Axler 3.59).
Since $\vec{v}$ was arbitrary, we know that $\ker T_2R \subseteq \ker T_2$.
Pick an arbitrary $\vec{w} \in \ker T_2$, then we know that $T_2(\vec{w}) = \vec{0}$.
Now since R is invertible, we know that R^{-1} exists.
Let $\vec{v} = R^{-1}(\vec{w})$, then we know that $R(\vec{v}) = \vec{w}$.
Now, consider $T_2R(\vec{v})$, we know that $T_2(R(\vec{v})) = T_2(\vec{w}) = \vec{0}$.
Thus, $R(\vec{v}) \in \ker T_2R$, which is the same as $\vec{w} \in \ker T_2R$.
Since $\vec{w}$ was arbitrary, we know that $\ker T_2 \subseteq \ker T_2R$.
Therefore, we have shown that $\ker T_2 = \ker T_2R$.
Meaning that $\ker T_1 = \ker T_2$.
Therefore, $\dim \ker T_1 = \dim \ker T_2$ as desired.

☺

Proof of $\impliedby$:   Suppose $\dim T_1 = \dim T_2 = m$.
Then we can construct a basis for $\ker T_1$ and $\ker T_2$: $(\vec{a}_1, \ldots, \vec{a}_m)$ and $(\vec{u}_1, \ldots, \vec{u}_m)$ respectively.
Now extend both to a basis of V respectively: $(\vec{a}_1, \ldots, \vec{a}_m, \vec{b}_1, \ldots, \vec{b}_n)$ and $(\vec{u}_1, \ldots, \vec{u}_m, \vec{v}_1, \ldots, \vec{v}_n)$.
By construction, we know that $(\vec{b}_1, \ldots, \vec{b}_n)$ and $(\vec{v}_1, \ldots, \vec{v}_n)$ are linearly independent lists in V.

Claim: Both $T_1(\vec{b}_1), \ldots, T_1(\vec{b}_n)$ and $T_2(\vec{v}_1), \ldots, T_2(\vec{v}_n)$ are linearly independent lists in W.

Proof: Suppose there exists scalars $\lambda_1, \ldots, \lambda_n$ such that:

$$\lambda_1 T_1(\vec{b}_1) + \ldots + \lambda_n T_1(\vec{b}_n) = \vec{0}$$

Then by linearity of T_1, we know that:

$$T_1(\lambda_1 \vec{b}_1 + \ldots + \lambda_n \vec{b}_n) = \vec{0}$$

Therefore, $\lambda_1 \vec{b}_1 + \ldots + \lambda_n \vec{b}_n \in \ker T_1$.
By our previous construction, we know that $(\vec{a}_1, \ldots, \vec{a}_m)$ is a basis for $\ker T_1$.
Thus for some scalars $\mu_1, \ldots, \mu_m$, we have:

$$\lambda_1 \vec{b}_1 + \ldots + \lambda_n \vec{b}_n = \mu_1 \vec{a}_1 + \ldots + \mu_m \vec{a}_m$$
$$\implies \lambda_1 \vec{b}_1 + \ldots + \lambda_n \vec{b}_n - \mu_1 \vec{a}_1 - \ldots - \mu_m \vec{a}_m = \vec{0}$$

Since $(\vec{a}_1, \ldots, \vec{a}_m, \vec{b}_1, \ldots, \vec{b}_n)$ is a linearly independent list in V, then this implies that:

$$\lambda_1 = \ldots = \lambda_n = \mu_1 = \ldots = \mu_m = 0$$

Thus, the only way to get a linear combination of $T_1(\vec{b}_1), \ldots, T_1(\vec{b}_n)$ to equal $\vec{0}$ is if all the scalars are 0.
Hence, $T_1(\vec{b}_1), \ldots, T_1(\vec{b}_n)$ is linearly independent.
We can show that $T_2(\vec{v}_1), \ldots, T_2(\vec{v}_n)$ is linearly independent in the same way with some minor modifications.
I'll leave it as an exercise to the reader. 😎😎😎😎😎😎😎😎😎😎
Thus, both lists are linearly independent in W as desired.

☺

1

Since both lists are linearly independent in W, we can extend them to a basis of V:
$(T_1(\vec{b_1}), \ldots, T_1(\vec{b_n}), \vec{c_1}, \ldots, \vec{c_p})$ and $(T_2(\vec{v_1}), \ldots, T_2(\vec{v_n}), \vec{w_1}, \ldots, \vec{w_p})$ respectively.
Now, let's define the following linear operators:

$$R : V \to V, R \in \mathcal{L}(V)$$

$$R = \begin{cases} \vec{a_i} \mapsto \vec{u_i} & \forall i \in \{1, \ldots, m\} \\ \vec{b_j} \mapsto \vec{v_j} & \forall j \in \{1, \ldots, n\} \end{cases}$$

$$S : W \to W, S \in \mathcal{L}(W)$$

$$S = \begin{cases} T_2(\vec{v_j}) \mapsto T_1(\vec{b_j}) & \forall j \in \{1, \ldots, n\} \\ \vec{w_k} \mapsto \vec{c_k} & \forall k \in \{1, \ldots, p\} \end{cases}$$

Notice that by construction that R maps a basis of V to a basis of V.
Therefore, the image of R is V, and by Axler 3.20, we know that R is surjective.
The same can be said for S, as it maps a basis of W to a basis of W, so S is surjective.
Thus, we have shown that $R \in \mathcal{L}(V)$ and $S \in \mathcal{L}(W)$ and are operators.
Thus, by Axler 3.69, we know that being surjective means that R and S are invertible.
Now, we want to show that $T_1 = ST_2R$. Let's do this directly.
Let $\vec{a}$ represent an arbitrary vector in the basis of $\ker T_1$, $\vec{a_1}, \ldots, \vec{a_m}$.
Thus, we know that $T_1(\vec{a}) = \vec{0}$.
We need to show that $ST_2R(\vec{a}) = \vec{0}$.

$$\begin{aligned} ST_2R(\vec{a}) &= ST_2(\vec{u}) \quad \text{where } \vec{u} \text{ is one of the } (\vec{u_1}, \ldots, \vec{u_m}) \\ &= S(\vec{0}) \quad \text{as } \vec{u} \in \ker T_1 \\ &= \vec{0} \quad \text{as } S \text{ is linear (Axler 3.11)} \end{aligned}$$

Furthermore, let $\vec{b}$ represent an arbitrary vector in the list of $(\vec{b_1}, \ldots, \vec{b_n})$.
Let's show that $ST_2R(\vec{b}) = T_1(\vec{b})$.

$$\begin{aligned} ST_2R(\vec{b}) &= ST_2(\vec{v}) \quad \text{where } \vec{v} \text{ is one of the } (\vec{v_1}, \ldots, \vec{v_n}) \\ &= S(T_2(\vec{v})) \quad \text{product of linear maps (Axler 3.8)} \\ &= T_1(\vec{b}) \quad \text{by definition of } S \end{aligned}$$

This completes the proof but, let's do one more thing.
Let $\vec{😎} \in V$ be an arbitrary vector.
That means that $\vec{😎}$ is in the span of $(\vec{a_1}, \ldots, \vec{a_m}, \vec{b_1}, \ldots, \vec{b_n})$.
Let's show that $T_1(\vec{😎}) = ST_2R(\vec{😎})$ to show my emojis.
Let's express $\vec{😎}$ as a linear combination of these vectors:

$$\vec{😎} = \sum_{i=1}^{m} \lambda_i \vec{a_i} + \sum_{j=1}^{n} \mu_j \vec{b_j}$$

Using the definition of R, we have:

$$R(\vec{😎}) = \sum_{i=1}^{m} \lambda_i \vec{u_i} + \sum_{j=1}^{n} \mu_j \vec{v_j}$$

Now, applying T_2:

2

$$T_2(R(\overset{\rightarrow}{😎})) = \sum_{i=1}^{m} \lambda_i T_2(\vec{u}_i) + \sum_{j=1}^{n} \mu_j T_2(\vec{v}_j)$$

Now, since $\vec{u}_i$ is in the kernel of T_1, $T_1(\vec{u}_i) = \vec{0}$.
So, the expression simplifies to:

$$T_2(R(\overset{\rightarrow}{😎})) = \sum_{j=1}^{n} \mu_j T_2(\vec{v}_j)$$

Finaly, applying S:

$$S(T_2(R(\overset{\rightarrow}{😎}))) = \sum_{j=1}^{n} \mu_j S(T_2(\vec{v}_j)) = \sum_{j=1}^{n} \mu_j T_1(\vec{b}_j)$$

Now, let's check:

$$T_1(\overset{\rightarrow}{😎}) = \sum_{i=1}^{m} \lambda_i T_1(\vec{a}_i) + \sum_{j=1}^{n} \mu_j T_1(\vec{b}_j) = \sum_{j=1}^{n} \mu_j T_1(\vec{b}_j)$$

Note that the first sum is zero as $\vec{a}_i$ is in the kernel of T_1.

Thus, we have shown that $T_1(\overset{\rightarrow}{😎}) = ST_2R(\overset{\rightarrow}{😎})$ for an arbitrary $\overset{\rightarrow}{😎} \in V$.
Thus, we have shown that $T_1 = ST_2R$.

☺

As we have shown both directions, the result follows.

☺

Lenma 0.0.1

Let $\vec{x}_s$ be a solution to $T(\vec{x}) = \vec{b}$.
Where T is a linear map that maps $\vec{x} \in \mathbb{R}^n$ to $\vec{b} \in \mathbb{R}^m$ by a matrix A: $T(\vec{x}) = A \cdot \vec{x}$.
Then, if there are other solutions, $\vec{x}_\star$, to $T(\vec{x}) = \vec{b}$,
Then there exists an $\vec{x}_k \in \ker T$ such that every other solution is given by:

$$\vec{x}_\star = \vec{x}_s + \vec{x}_k$$

Proof: We know $T(\vec{x}_s) = \vec{b}$ and $T(\vec{x}_\star) = \vec{b}$ as they are solutions to $T(\vec{x}) = \vec{b}$. Consider, $T(\vec{x}_\star - \vec{x}_s)$.
Then, by the fact that T is a linear map, the following is true:

$$T(\vec{x}_\star - \vec{x}_s) = T(\vec{x}_\star) - T(\vec{x}_s) = \vec{b} - \vec{b} = \vec{0}$$

.
Therefore, $\vec{x}_\star - \vec{x}_s \in \ker T$.
Thus, there exists $\vec{x}_k \in \ker T$ such that $\vec{x}_k = \vec{x}_\star - \vec{x}_s$.
By definition, $T(\vec{x}_k) = 0$, meaning it is a solution to $T(\vec{x}) = \vec{0}$.
Thus, every other solution to $T(\vec{x}) = \vec{b}$ is given by a solution, $\vec{x}_s$ plus a solution to $T(\vec{x}) = \vec{0}$, $\vec{x}_k$.   ☺

Question 1

Vandermonde Matrices.
(1) Prove that

$$\det \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = (a-b)(b-c)(c-a).$$

(b) Prove an analogous formula for $n \times n$ matrices, using appropriate row operations to clear out the first column.

3

(c) Use the determinant in part (b) to prove that there is a unique polynomial $p(x) \in \mathcal{P}_n(\mathbb{F})$ that takes arbitrary prescribed values at $n + 1$ points $x = x_0, x_1, \ldots, x_n$.

Proof of *1:* Not going to lie. This problem is so fun compare to the other ones. I love it.
   Let's begin by proving this directly.

$$
\det \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \cdot \det A_{i1}
$$

$$
= (-1)^{1+1} a_{1,1} \cdot \det A_{1,1} + (-1)^{2+1} a_{2,1} \cdot \det A_{2,1} + (-1)^{3+1} a_{3,1} \cdot \det A_{3,1}
$$

$$
= 1 \cdot \det \begin{pmatrix} b & c \\ b^2 & c^2 \end{pmatrix} + (-1) \cdot \det \begin{pmatrix} a & c \\ a^2 & c^2 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} a & b \\ a^2 & b^2 \end{pmatrix}
$$

$$
= 1 \cdot (bc^2 - b^2 c) + (-1) \cdot (ac^2 - a^2 c) + 1 \cdot (ab^2 - a^2 b)
$$

$$
= bc^2 - b^2 c - ac^2 + a^2 c + ab^2 - a^2 b
$$

$$
= bc^2 - b^2 c - ac^2 + a^2 c + ab^2 - a^2 b + (abc - abc) \text{ we can add a 0 value term}
$$

$$
= (abc - b^2 c - ac^2 + bc^2) - (a^2 b - ab^2 - a^2 c + abc)
$$

$$
= (ab - b^2 - ac + bc)(c - a)
$$

$$
= (a - b)(b - c)(c - a)
$$

☺

Proof of *2:* Given V, a Vandermonde matrix of size $n \times n$, then V is of the form:

$$
V = \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
a & b & c & \cdots & n \\
a^2 & b^2 & c^2 & \cdots & n^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a^n & b^n & c^n & \cdots & n^n
\end{pmatrix}
$$

We want to prove a formula for $\det V$.
We were given the hint to use row operations to clear out the first column.
But first we need to show how that would affect the determinant.
Let δ be the determinant function.
By theorem 4 of the handout, we know that if there is an elementary matrix of type (i), $E \in \mathbb{R}^{n,n}$,
then $\delta(E \cdot V) = \delta(V)$, where E adds a multiple of one row to another row.
That means that when we modify the first column of V through row operations,
we don't change the determinant.
Now, I'll show how to clear the first column.
Let's show two specific row operation that clears the first column's second to last entry and last entry.

$$
\begin{pmatrix}
1 & 1 & \cdots & 1_n \\
a & b & \vdots & n \\
a^2 & b^2 & \vdots & n^2 \\
\vdots & \vdots & \vdots & \vdots \\
a^n & b^n & \cdots & n^n
\end{pmatrix}
\xrightarrow{R_n - a R_{n-1} \mapsto R_n}
\begin{pmatrix}
1 & 1 & \cdots & 1_n \\
a & b & \vdots & n \\
a^2 & b^2 & \vdots & n^2 \\
\vdots & \vdots & \vdots & \vdots \\
a^{n-1} & b^{n-1} & \cdots & n^{n-1} \\
a^n - a \cdot a^{n-1} & b^n - a \cdot b^{n-1} & \cdots & n^n - a \cdot n^{n-1}
\end{pmatrix}
$$

4

$$= \begin{pmatrix} 1 & 1 & \cdots & 1_n \\ a & b & \vdots & n \\ a^2 & b^2 & \vdots & n^2 \\ \vdots & \vdots & \vdots & \vdots \\ a^{n-1} & b^{n-1} & \cdots & n^{n-1} \\ 0 & b^{n-1}(b-a) & \cdots & n^{n-1}(n-a) \end{pmatrix}$$

$$\xrightarrow{R_{n-1}-aR_{n-2}\mapsto R_{n-1}} \begin{pmatrix} 1 & 1 & \cdots & 1_n \\ a & b & \vdots & n \\ a^2 & b^2 & \vdots & n^2 \\ \vdots & \vdots & \vdots & \vdots \\ a^{n-2} & b^{n-2} & \cdots & n^{n-2} \\ 0 & b^{n-2}(b-a) & \cdots & n^{n-2}(n-a) \\ 0 & b^{n-1}(b-a) & \cdots & n^{n-1}(n-a) \end{pmatrix}$$

Let's now generalize this to clear the first column's ith entry for $i \in \{n, n-1, \cdots, 2\}$.
But before we do that, let's notate the original matrix as V_a, then apply the generalization:

$$\begin{pmatrix} 1 & 1 & \cdots & 1_n \\ a & b & \vdots & n \\ a^2 & b^2 & \vdots & n^2 \\ \vdots & \vdots & \vdots & \vdots \\ a^n & b^n & \cdots & n^n \end{pmatrix} \xrightarrow{R_i-aR_{i-1}\mapsto R_i} \begin{pmatrix} 1 & 1 & \cdots & 1_n \\ a & b & \cdots & n \\ \vdots & \vdots & \vdots & \vdots \\ a^{i-1} & b^{i-1} & \cdots & n^{i-1} \\ 0 & b^{i-1}(b-a) & \cdots & n^{i-1}(n-a) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & b^j(b-a) & \cdots & n^j(n-a) \end{pmatrix}, j > i$$

$$= \begin{pmatrix} 1 & 1 & \cdots & 1_n \\ 0 & b-a & \vdots & n-a \\ 0 & b(b-a) & \vdots & n(n-a) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & b^{i-2}(b-a) & \cdots & n^{i-2}(n-a) \\ 0 & b^{i-1}(b-a) & \cdots & n^{i-1}(n-a) \end{pmatrix}$$

Let's now introduce more notation:

$$V_{ai} = \begin{pmatrix} 1 & 1 & \cdots & 1_n \\ 0 & b-a & \vdots & n-a \\ 0 & b(b-a) & \vdots & n(n-a) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & b^{n-2}(b-a) & \cdots & n^{n-2}(n-a) \\ 0 & b^{n-1}(b-a) & \cdots & n^{n-1}(n-a) \end{pmatrix}$$

Where $V_{ai} = E_n^{ai} \cdots E_2^{ai} \cdot V_a$ where E_i^{ai} is the elementary matrix (EM) that clears the first column's ith entry.
Let's define $E_{V_{ai}}$ to be the list of all the EMs that clear the first column's entries i.e., $E_{V_{ai}} = (E_n^{ai}, \ldots, E_2^{ai})$.
Let's now apply the first column expansion δ function to V_{ai}:

$$\det V_{ai} = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \cdot \det(A_{i,1})$$

$$= (-1)^{1+1} a_{1,1} \cdot \det(A_{1,1}) + \cdots + (-1)^{n+1} a_{n,1} \cdot \det(A_{n,1})$$

$$= 1 \cdot \det(A_{1,1}) + 0 \cdot \det(A_{2,1}) + \cdots + 0 \cdot \det(A_{n,1})$$

$$= \det(A_{1,1}) \quad \textbf{Note}: \text{let } A_{1,1} = V_M^a$$

$$= \det \begin{pmatrix} b - a & \cdots & n - a \\ b(b-a) & \vdots & n(n-a) \\ \vdots & \vdots & \vdots \\ b^{n-2}(b-a) & \cdots & n^{n-2}(n-a) \\ b^{n-1}(b-a) & \cdots & n^{n-1}(n-a) \end{pmatrix}$$

Notice that V_M^a has factorable columns in the form $(\lambda - a)$ where $\lambda \in \{b, \ldots, n\}$.
Now, we will prove this in due time, but PSET 8 tell us that:

$$\det(V_M^a) = \det(V_{D^T}^a), \text{ where } V_{M^T}^a \text{ is the transpose of } V_M^a$$

$$V_{M^T}^a = \begin{pmatrix} b - a & b(b-a) & \cdots & b^{n-1}(b-a) \\ \vdots & \vdots & \vdots & \vdots \\ n - a & n(n-a) & \cdots & n^{n-1}(n-a) \end{pmatrix}$$

Again, we define transpose in PSET 8, but for now let's take this operation for granted.
There exists EMs of type (iii) such that we can factor out $(\lambda - a)$ from each row of $V_{M^T}^a$, where $\lambda \in \{b, \ldots, n\}$.
Denote this list of EMs as $E_{V_{af}}$.
Notice that there are $n - 1$ rows, so we have $n - 1$ EMs in this list.

> **Note:-**
>
> I will not show what $E_{V_{af}}$ is as it seems tedious and I don't want to do it.
> Trust me bro.

Remember that for all $E_i^{af} \in E_{V_{af}}$, we have $\delta(E_i^{af}) = (\lambda - a)$ by corollary 5 of the handout.
Let V_{bt} be the matrix where we factor out $(\lambda - a)$ from each row of $V_{M^T}^a$.
Then, we have:

$$V_{bt} = \begin{pmatrix} 1 & b & \cdots & b^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} \end{pmatrix}$$

In other words we get $V_{bt} = E_i^{af} \cdot V_{M^T}^a$, where $E_i^{af} \in E_{V_{af}}, i = n - 1, \ldots, 1$.
Combining this all together, we have:

$$\det(V) = \det(V_a) = \det(V_{M^T}^a) = \det(E_{n-1}^{af}) \cdot \cdots \cdot \det(E_1^{af}) \cdot \det(V_{bt})$$

Now, let's transpose V_{bt} and define the result as V_b.

$$V_b = \begin{pmatrix} 1 & \cdots & 1_{n-1} \\ b & \cdots & n \\ \vdots & \vdots & \vdots \\ b^{n-1} & \cdots & n^{n-1} \end{pmatrix}$$

Recall that $\det(V_b) = \det(V_{bt})$, so we have:

$$\det(V) = \det(V_a) = (b - a) \cdots (n - a) \cdot \det(V_b)$$

Now, the value of $\det(V) = \det(V_a)$ is expressed in terms of $\det(V_b)$.
However, have you noticed that V_b is a Vandermonde matrix but of size $(n-1) \times (n-1)$?
That means to find the determinant of V_b, we can:

1. Find elementary matrices that clear the first column of V_b, $E_{V_{bi}}$

2. Apply these EM to V_b to get V_{bi} i.e., $V_{bi} = E_{V_{bi}} \cdot V_b$

3. Find the determinant of V_{bi} using first column expansion and realize only one sub-matrix is non-zero.

4. Denote this sub-matrix as V_M^b.

5. Transpose V_M^b to get $V_{M^T}^b$.

6. Find elementary matrices that factor out $(\lambda - b)$ from each row of $V_{M^T}^b$, $E_{V_{bf}}$, for $\lambda \in \{c, \ldots, n\}$.

7. Apply these EM to $V_{M^T}^b$ to the resulting matrix V_{ct}, i.e., $V_{ct} = E_{V_{bf}} \cdot V_{M^T}^b$.

8. Transpose V_{ct} to get V_c.

9. Now the determinant of V_b is expressed in terms of $\det(V_c)$ i.e., $\det(V_b) = (c-b) \cdots (n-b) \cdot \det(V_c)$.

10. Notice that V_c is a Vandermonde matrix but of size $(n-2) \times (n-2)$.

In other words, we can repeat this process until we get to a Vandermonde matrix of size 2×2.
Let $\star$ represent the variable before the variable n.
Where we end up with $V_\star$ that looks like this:

$$V_\star = \begin{pmatrix} 1 & 1 \\ \star & n \end{pmatrix}$$

The determinant of $V_\star$ is:

$$\det(V_\star) = (1 \cdot n) - (1 \cdot \star) = (n - \star)$$

Note that V_n is $\begin{pmatrix} 1 \end{pmatrix}$, so $\det(V_n) = 1$.
Thus, we can express the determinant of a Vandermonde matrix of size $n \times n$, $n \in \mathbb{N}$ as:

$$
\begin{aligned}
\det(V) = \det(V_a) &= \det(E_{V_{af}}) \cdot \det(V_b) \\
&= (b-a) \cdots (n-a) \cdot \det(V_b) \\
&= (b-a) \cdots (n-a) \cdot (c-b) \cdots (n-b) \cdot \det(V_c) \\
&= \vdots \\
&= (b-a) \cdot \ldots \cdot \det(V_\star) \\
&= (b-a) \cdot \ldots \cdot \det(V_n) \\
&= (b-a) \cdot \ldots \cdot (c-b) \cdot \ldots \cdot (n-\star) \cdot (1) \\
&= \prod_{a \leqslant i < j \leqslant n} (x_j - x_i)
\end{aligned}
$$

Which is the formula we wanted to prove.
But I wrote a nice algorithm that shows all the steps of this process. See below.

Algorithm 1: Determinant of a Vandermonde Matrix, $DetVa$

Input: A Vandermonde matrix of size n, V
Output: The determinant of V

1 $det \leftarrow 1$;
2 $size \leftarrow n$;
3 if $n = 1$ **then**
4 | **return** 1;
5 else
6 | $\star \leftarrow V_{2,1}$;   /* the first entry of the second row */
7 | $E_{V_i} \leftarrow ()$;
8 | **foreach** $i \in size, size - 1, \ldots, 2$ **do**
9 | | $rowop \leftarrow V(i) = V(i) - \star \cdot V(i - 1)$;
10 | | $E_i \leftarrow$ an elementary matrix equivalent to $rowop$;
11 | | $E_{V_i} \leftarrow (E_i) + E_{V_i}$;
12 | **end**
13 | $V_i \leftarrow E_{V_i} \cdot V$;
14 | $V_M \leftarrow minor(V_i, 1, 1)$;
15 | $V_{M^T} \leftarrow transpose(V_M)$;
16 | $E_{V_f} \leftarrow ()$;
17 | **foreach** $i \in size - 1, \ldots, 1$ **do**
18 | | $term \leftarrow V_{M^T}(i, 1)$;   /* term of the form $(\lambda - \star)$ */
19 | | $det \leftarrow det \cdot term$;
20 | | $E_i \leftarrow$ an elementary matrix that factors out $(term)$ from each row of V_{M^T};
21 | | $E_{V_f} \leftarrow (E_i) + E_{V_f}$;
22 | **end**
23 | $V'_T \leftarrow E_{V_f} \cdot V_{M^T}$;
24 | $V' \leftarrow transpose(V'_T)$;
25 | **return** $det \cdot DetVa(n - 1, V')$;
26 end

☺

Proof of c: Let's remind ourselves what $\mathcal{P}_n(\mathbb{F})$ means:

$$\mathcal{P}_n(\mathbb{F}) = \left\{ p : \mathbb{F} \to \mathbb{F} \mid p(x) = \sum_{i=0}^{n} a_i x^i, a_i \in \mathbb{F} \right\}$$

We were given that there might be a polynomial $p(x) \in \mathcal{P}_n(\mathbb{F})$
that takes arbitrary prescribed values at $n + 1$ points $x = x_0, x_1, \ldots, x_n$.
Let's assume that the points are distinct (it wasn't clear).
Given $n + 1$ distinct points $x_0, x_1, \ldots, x_n$ with arbitrary values, the polynomial $p(x) \in \mathcal{P}_n(\mathbb{F})$ is of the form:

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

However, the coefficients $a_0, \ldots, a_n$ are not known. To determine them, we can set up a system of equations:

$$p(x_0) = a_0 + a_1 x_0 + \cdots + a_n x_0^n$$
$$p(x_1) = a_0 + a_1 x_1 + \cdots + a_n x_1^n$$
$$\vdots$$
$$p(x_n) = a_0 + a_1 x_n + \cdots + a_n x_n^n$$

This system can be expressed in matrix form as:

8

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_n) \end{pmatrix}$$

The matrix on the left is the transpose of a Vandermonde matrix.
As shown in part (b), the determinant of a Vandermonde matrix, provided the x_i values are distinct, is:

$$\det(V) = \prod_{i<j}(x_j - x_i)$$

Since the x_i values are distinct, $\det(V) \neq 0$ and that $\det(V) = \det(V^T)$,
then the matrix on the left is invertible.
Let's denote the matrix on the left by A, the vector of coefficients by $\vec{x}$,
and the vector of polynomial values by $\vec{p}$.
Now, because the matrix is invertible, we can do the following:

$$A \cdot \vec{x} = \vec{p}$$
$$A^{-1} \cdot A \cdot \vec{x} = A^{-1} \cdot \vec{p}$$
$$\vec{x} = A^{-1} \cdot \vec{p}$$

Given that the determinant of A is non-zero, this system of equations has a unique solution.
This is because the determinant A is non-zero,
which means that no row is a linear combination of the other rows.
Thus, the system of linear equation will have an answer as the mapping is unique.
This means there is a unique set of coefficients a_i, which in turn implies a unique polynomial.
Hence, there exists a unique polynomial $p(x) \in \mathcal{P}_n(\mathbb{F})$ with the stated property. ☺

Question 2: Column Operations

Let $E \in \mathbb{F}^{n,n}$ be an elementary matrix. Then the result of $A \cdot E$ is a column operation on A (you don't have to prove this):
(i) If E is of type (i), with a in the (ij)-th position, then $A \cdot E$ is the column operation "add $a \cdot$ (column i) of A to (column j) of A ";
(ii) If E is of type (ii), then $A \cdot E$ is the column operation "interchange (column i) with (column j)";
(iii) If E is of type (iii), then $A \cdot E$ is the column operation "multiply (column i) by the nonzero scalar c;
(a) Define a notion of "reduced column-echelon form", based on the notion of "reduced row-echelon form". Do so in such a way that if A is in reduced row-echelon form, then A^t is in reduced column-echelon form.
(b) Use induction on the number of columns of a matrix to show that every matrix has a reduced column-echelon form.

Theorem 0.0.1 Column Operations

Note:-
I did not read you did not have to prove this. I'm so tired.

1. **Type I (Scalar Addition to another column):**

   Given an elementary matrix E of type (i) with a in the (i,j)-th position, and all other off-diagonal elements being zero and diagonal entries being ones. This means that the j-th column of E has the value a at the i-th row and 1 at the j-th row.

   For matrix multiplication, each entry in the product matrix AE is obtained by taking the dot product of the rows of A with the columns of E.

   Let's consider the j-th column of AE.

For any row k of A, the dot product of the k-th row of A with the j-th column of E is:

$$A_{k1} \times 0 + \ldots + A_{ki} \times a + \ldots + A_{kj} \times 1 + \ldots + A_{kn} \times 0$$

This simplifies to:

$$A_{kj} + a \times A_{ki}$$

This operation adds a times the i-th entry of the k-th row to the j-th entry of the same row.

And, this holds true for every row k in matrix A.

Therefore, the j-th column of AE is the result of adding a times the i-th column of A to the j-th column of A.

Hence, $A \cdot E$ is the column operation "add $a\cdot$ (column i) of A to (column j) of A".

2. **Type II (Column Interchange):**

Given an elementary matrix E of type (ii), such that the i-th and j-th columns are swapped. This means that the i-th column of E has 1 at the j-th position and the j-th column has 1 at the i-th position, with all other entries being zero except the diagonal ones which are 1.

Consider the matrix product AE. The k-th column of AE for $k \neq i, j$ remains unchanged, since it will dot with columns of E filled with zeros except for the one at its own position. However, the i-th column of AE will be the j-th column of A and vice-versa.

Thus, post-multiplying A by E effectively interchanges the i-th and j-th columns of A. Therefore, $A \cdot E$ is the column operation "interchange (column i) with (column j)".

3. **Type III (Scalar Multiplication to a column):**

Given an elementary matrix E of type (iii) such that its i-th column has a scalar c at the diagonal position and zeros elsewhere. This means that the i-th column of E is $[0, \ldots, 0, c, 0, \ldots, 0]$.

For the product AE, every column other than the i-th column remains unchanged. However, the i-th column of AE is multiplied by c.

Thus, post-multiplying A by E effectively scales the i-th column of A by c. Therefore, $A \cdot E$ is the column operation "multiply (column i) by the nonzero scalar c".

Solution: Let's define a notion of "reduced column-echelon form".

Let's call my notation ACBAIDRCEDF:

"Acronyms Can Be Anything I Desired: REDUCED, Column Echelon DEADASS Form."

Trademarked, by the way.

Anyway, an $m \times n$ matrix M is in ACBAIDRCEDF if:

(i) If (column j) of M is entirely composed of zeros, then so is (column k) for all $k > j$.

For example, the matrix holds the following property:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(ii) If (column j) has a nonzero entry, then then its north most nonzero entry is 1. Let's call it a Cool Pivot.

For example, the matrix holds the following property:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(iii) If (column $j + 1$) has nonzero entries, the cool pivot in (column j) is strictly north-west of the cool pivot in (column $j + 1$).

That is to say, the cool pivot in (column j) is above and to the left of the cool pivot in (column $j + 1$).

For example, the matrix holds the following property:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(iv) Any row containing a cool pivot of a column has zeros everywhere else.

For example, the matrix holds the following property:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now given a matrix A that is in row reduced echelon form, we want to show that A^T is in ACBAIDRCEDF. If a matrix is in row reduced echelon form, then it looks something like this:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Thus, once we take the transpose of A, we get:

$$A^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Let's go down the list of properties and show that A^T satisfies them.

(i) No column here is entirely composed of zeros, so this property is vacuously true.

(ii) Every column has a cool pivot as the most north nonzero entry, so this property is true.

(iii) Every cool pivot is strictly north-west of the cool pivot in the column to its right, which checks out.

(iv) The entries left and right of the cool pivots are all zeros, so this property is true.

Thus, we have shown that A^T is in ACBAIDRCEDF.

Proof of b : We proceed by induction on the number of columns of a matrix $j \in \mathbb{N}$.

Base Case: Let $j = 1$

Then, the matrix A is of size $m \times 1$ for some $m \in \mathbb{N}$:

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

Let's consider the first nonzero entry a_λ. Without loss of generality, let's assume that it is the first entry. If not we can apply an elementary matrix of type (ii) to swap the first row with the row containing a_λ. We want a column operation that will make a_1 a cool pivot i.e., equal to 1. We can do this by multiplying the first column by $\frac{1}{a_1}$. i.e., $C_1 \leftarrow \frac{1}{a_1} \cdot C_1$. Let E_1 of type (iii) be the elementary matrix that represents this operation. Then, $A \cdot E_1$ looks something like this:

$$A \cdot E_1 = \begin{pmatrix} 1 \\ \frac{a_2}{a_1} \\ \vdots \\ \frac{a_m}{a_1} \end{pmatrix}$$

Notice, that by our notion of reduced column-echelon form, the matrix is now in ACBAIDRCEDF.

(i) The first column is not entirely composed of zeros, so we pass

(ii) The first column has a cool pivot as the most north nonzero entry

(iii) There is only one cool pivot, we can't compare it to anything, so by default it passes.

(iv) The entries left and right of the cool pivots don't exist, so this property is true.

Thus, every matrix with one column has a reduced column-echelon form.

Inductive Hypothesis: Assume the statement is true for every matrix with $k \in \mathbb{N}$ columns,
We want to show that matrix with $k + 1$ columns can be reduced to a reduced column-echelon form.
Let A be an arbitrary matrix with $k + 1$ columns of size $m \times (k + 1)$ for some $m \in \mathbb{N}$.
By our inductive hypothesis, we know that the first k columns of A,
can be reduced to a reduced column-echelon form with some elementary matrices $E_1, \ldots, E_\alpha$.
Where every E_i is an elementary matrices of type (i), (ii), or (iii)
and applied to the right of A. That is to say, each E_i apply column operations to A.
Let $B = A \cdot E_1 \cdots E_\alpha$.
Then, the first k columns of B are in reduced column-echelon form.
This means that due to our notion of reduced column-echelon form,
all k columns have a cool pivot as their north most nonzero entry.
Thus every row with a cool pivot has zeros in every other entry.
Now consider the $k + 1$-th column of B.
We have two cases to consider:

Case one: $m \leq k$***:*** If the number of rows is at most k, then notice that the following holds:
If $m = k$, there are k rows and there are k cool pivots.
By our notion, every row with a cool pivot has zeros in every other entry.
Thus, the $k + 1$-th column of B is entirely composed of zeros.
The same logic holds if $m < k$.
After the m column of B, the rest of the columns are entirely composed of zeros.
Thus, the matrix is in reduced column-echelon form for $k + 1$ if $m \leq k$.

Case Two: $m \geq k + 1$ ***:*** If there are at least as many rows as columns,
then the $k + 1$-th column of B has at least one nonzero entry.
Denote this entry as λ. It follows from our previous case that λ is the most north nonzero entry as $m > k$.
Now, we want to make λ a cool pivot.
Let E_λ be elementary matrix of type (iii) that scales the $k + 1$-th column of B by $\frac{1}{\lambda}$.
Then, $B \cdot E_\lambda$ looks something like this:

$$B \cdot E_\lambda = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & \cdots & \cdots & 1_{k,k} & 0_{k,k+1} \\ a_{k+1,1} & a_{k+1,2} & \cdots & \cdots & \cdots & a_{k+1,k} & 1_{k+1,k+1} \\ a_{k+2,1} & & & & & & \frac{a_{k+2,k+1}}{\lambda} \\ \vdots & & & & & & \\ a_m & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{a_{m,k+1}}{\lambda} \end{pmatrix}$$

Notice, that there may be elements below the $k + 1$-th row that are nonzero for any given column.
This is fine, and will not affect our proof as ACBAIDRCEDF only concerns itself with the first $k + 1$ rows.

Thus, we can safely ignore the rows below the $k + 1$-th row (though they will change as we see later on).
Let $C = B \cdot E_\lambda$.
Now, let $D = C \cdot E_\gamma \cdots E_{\gamma+k-1}$. Where $E_\gamma, \ldots, E_{\gamma+k-1}$ are type (i) elementary matrices that do the following:

$$C = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & 1_{k,k} & 0_{k,k+1} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} & 1_{k+1,k+1} \\ a_{k+2,1} & & & & \frac{a_{k+2,k+1}}{\lambda} \\ \vdots & & & & \\ a_{m,1} & \cdots & \cdots & \cdots & \frac{a_{m,k+1}}{\lambda} \end{pmatrix} \xrightarrow{C_k - a_{k+1,k}C_{k+1} = C_k} \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & 1_{k,k} - 0 & 0_{k,k+1} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} - a_{k+1,k} & 1_{k+1,k+1} \\ a_{k+2,1} & & & & \frac{a_{k+2,k+1}}{\lambda} \\ \vdots & & & & \\ a_m & \cdots & \cdots & \cdots - a_{k+1,k}\frac{a_{m,k+1}}{\lambda} & \frac{a_{m,k+1}}{\lambda} \end{pmatrix}$$

$$C \cdot E_\gamma = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & 1_{k,k} & 0_{k,k+1} \\ a_{k+1,1} & a_{k+1,2} & \cdots & 0 & 1_{k+1,k+1} \\ a_{k+2,1} & & & & \frac{a_{k+2,k+1}}{\lambda} \\ \vdots & & & & \\ a_{m,1} & \cdots & \cdots & \cdots - a_{k+1,k}\frac{a_{m,k+1}}{\lambda} & \frac{a_{m,k+1}}{\lambda} \end{pmatrix}$$

In other words, let γ be the element at the $k + 1$-th row and j-th column of C for some $1 \le j \le k$
Then, $E_{\gamma+i}, i \in \{0, \ldots, k-1\}$ is an elementary matrix type (i) that does the following column operation:

$$E_{\gamma+i} = C_j - \gamma C_{k+1} = C_j - a_{k+1,j}C_{k+1} \text{ for } j = k - i$$

Thus, after applying $E_{\gamma+i}$ for $i \in \{0, \ldots, k-1\}$ to C , we get the following matrix:

$$D = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & 1_{k,k} & 0_{k,k+1} \\ 0_{k+1,1} & 0_{k+1,2} & \cdots & 0_{k+1,k} & 1_{k+1,k+1} \\ a_{k+2,1} - \frac{a_{k+2,k+1}}{\lambda} & & & & \frac{a_{k+2,k+1}}{\lambda} \\ \vdots & & & & \\ a_{m,1} - \frac{a_{m,k+1}}{\lambda} & \cdots & \cdots & \cdots - a_{k+1,k}\frac{a_{m,k+1}}{\lambda} & \frac{a_{m,k+1}}{\lambda} \end{pmatrix}$$

Notice that the $k + 1$-th row of D is now in reduced column-echelon form.
This follows as it has a cool pivot as its most north nonzero entry with zeros elsewhere in its row.
Moreover, the k pivot is strictly north-west of the $k + 1$-th pivot (in fact a upward step).
Thus, with elementary matrices: $E_\lambda, E_\gamma, \ldots, E_{\gamma+k-1}$,
we have reduced the $k + 1$-th column of B to a reduced column-echelon form.
Hence, $D = A \cdot E_1 \cdots E_\alpha \cdot E_\lambda \cdot E_\gamma \cdots E_{\gamma+k-1}$ is in reduced column-echelon form.
Thus, this concludes our inductive step as we have shown that the $k + 1$-th column of B,
can be reduced to a reduced column-echelon form for $m \in \mathbb{N}$
Thus, through the principle of mathematical induction, we have proven that every matrix has a reduced column-echelon form for any number of columns $j \in \mathbb{N}$.

☺

Question 3

Determinants are column-linear. We say that a function $\delta : \mathbb{F}^{n,n} \to \mathbb{F}$ is column-linear if the following condition holds:
Let A, B and D be three $n \times n$ matrices, all of whose columns are equal, except for the k-th column, where we assume that $D_k = \lambda \cdot A_k + \mu \cdot B_k$. Then

$$\delta(D) = \lambda \cdot \delta(A) + \mu \cdot \delta(B).$$

Or, more visually:

$$\delta\left(\begin{array}{ccccc} | & & | & & | \\ A_1 & - & \lambda \cdot A_k + \mu \cdot B_k & - & A_m \\ | & & | & & | \end{array}\right) = \lambda \cdot \delta\left(\begin{array}{ccccc} | & & | & & | \\ A_1 & - & A_k & - & A_m \\ | & & | & & | \end{array}\right) + \mu \cdot \delta\left(\begin{array}{ccccc} | & & | & & | \\ A_1 & - & B_k & - & A_m \\ | & & | & & | \end{array}\right)$$

For example:

$$\delta\begin{pmatrix} 1 & 2\lambda + 10\mu & 3 \\ 4 & 5\lambda + 11\mu & 6 \\ 7 & 8\lambda + 12\mu & 9 \end{pmatrix} = \lambda \cdot \delta\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \mu \cdot \delta\begin{pmatrix} 1 & 10 & 3 \\ 4 & 11 & 6 \\ 7 & 12 & 9 \end{pmatrix}$$

Prove that the determinant function det: $\mathbb{F}^{n,n} \to \mathbb{F}$ is column-linear. [Hint: Problem 4(a).]

Proof: Let A, B and D be three $n \times n$ matrices, all of whose columns are equal, except for the k-th column, where we assume that $D_k = \lambda \cdot A_k + \mu \cdot B_k$.
We want to show that for every column j, the following equality holds:

$$d_{1,j} \cdot \det(D_{1,j}) = \lambda \cdot a_{1,j} \cdot \det(A_{1,j}) + \mu \cdot b_{1,j} \cdot \det(B_{1,j})$$

Before, we begin, notice that the determinant function is unique, which means that Theorem 3 of the handout holds for our δ.
That is to say that δ holds all the properties defined in theorem 3.
Therefore, theorem 4 and corollary 5, 6, 7 hold for δ.
Specifically we can use that if if E is an elementary matrix, then:

(a) If E is of type (i), then $\delta(E) = 1$

(b) If E is of type (ii), then $\delta(E) = -1$

(c) If E is of type (iii), then $\delta(E) = c$

Let's proceed by induction on $n \in \mathbb{N}$ to show that δ is column-linear.

Base Case: Remember that the determinant of a 1×1 matrix is just the entry of the matrix, thus:

$$\begin{aligned} \det(D) &= \lambda A_1 + \mu B_1 \\ &= \lambda \det(A) + \mu \det(B) \end{aligned}$$

Thus, our base case holds for $n = 1$.

Inductive Step: Assume that δ is column-linear for $\mathbb{F}^{n-1,n-1} \to \mathbb{F}$.
Now, we are given that D, A, B are the same except for the k-th column.
Which means that D_k is a linear combination of A_k and B_k: $D_k = \lambda A_k + \mu B_k$.
Notice that by our induction hypothesis, the determinant of each $(n-1) \times (n-1)$ matrix, $D'_{1,j}, A'_{1,j}, B'_{1,j}$ is:

(i) Column linear in the j-th column for all $j < k$

(ii) Column linear in the $j-1$-th column for all $j > k$

Thus, we can write:

$$\det(D) = \sum_{k=1}^{n} (-1)^{k+1} d_{1,k} \cdot \det(D'_{1,k})$$

$$= \sum_{k=1}^{j-1} (-1)^{k+1} d_{1,k} \cdot \det(D'_{1,k}) + (-1)^{j+1} d_{1,j} \cdot \det(D'_{1,j}) + \sum_{k=j+1}^{n} (-1)^{k+1} d_{1,k} \cdot \det(D'_{1,k})$$

$$= \sum_{k \neq j} (-1)^{k+1} d_{1,k} \cdot \det(D'_{1,k}) + (-1)^{j+1} d_{1,j} \cdot \det(D'_{1,j})$$

$$= \sum_{k \neq j} (-1)^{k+1} d_{1,k} \cdot \left(\lambda \det(A'_{1,k}) + \mu \det(B'_{1,k}) \right) + (-1)^{j+1} (\lambda A_k + \mu B_k) \cdot \det(D'_{1,j})$$

$$= \sum_{k \neq j} \ldots + (-1)^{j+1} \lambda a_{1,j} \cdot \det(D'_{1,j}) + (-1)^{j+1} \mu b_{1,j} \cdot \det(D'_{1,j})$$

$\det(D_{1,j})' = \det(A_{1,j})' = \det(B_{1,j})'$ by construction, as the (j) column was removed

$$= \sum_{k \neq j} \ldots + (-1)^{j+1} \lambda a_{1,j} \cdot \det(A'_{1,j}) + (-1)^{j+1} \mu b_{1,j} \cdot \det(B'_{1,j})$$

$$= \sum_{k=1}^{n} (-1)^{k+1} \lambda a_{1,k} \cdot \det(A'_{1,k}) + (-1)^{k+1} \mu b_{1,k} \cdot \det(B'_{1,k})$$

$$= \lambda \left(\sum_{k=1}^{n} (-1)^{k+1} a_{1,k} \cdot \det(A'_{1,k}) \right) + \mu \left(\sum_{k=1}^{n} (-1)^{k+1} b_{1,k} \cdot \det(B'_{1,k}) \right)$$

$$= \lambda \det(A) + \mu \det(B)$$

Thus $\det(D) = \lambda \det(A) + \mu \det(B)$ for n which completes our inductive step.
Thus, through the principle of mathematical induction, we have proven that δ is column-linear.

Note:-

This took me way too long, and I'm still not sure how I would use the hint.
I feel like there was an easier way, but I couldn't find it.

☺