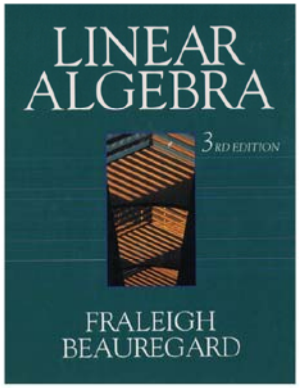
Linear Algebra

Chapter 3. Vector Spaces

Section 3.1. Vector Spaces—Proofs of Theorems



Example 3.1.2

Example 3.1.2. The set $\mathcal{P}$ of all polynomials in variable x with real coefficients is a vector space. Vector addition and scalar multiplication are the usual addition of polynomials and multiplication of a polynomial by a scalar.

Solution. Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, $q(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$, and $r(x) = c_\ell x^\ell + c_{\ell-1}x^{\ell-1} + \cdots + c_1 x + c_0$ be polynomials in $\mathcal{P}$ (where, say, $\ell \le n \le m$) and let s and t be real scalars. Then $p(x) + q(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0) \in \mathcal{P}$ and so $\mathcal{P}$ is closed under vector addition. Also, $sp(x) = (sa_n)x^n + (sa_{n-1})x^{n-1} + \cdots + (sa_1)x + (sa_0) \in \mathcal{P}$ and so $\mathcal{P}$ is closed under scalar multiplication. We take these computations as the definitions of vector addition and scalar multiplication in $\mathcal{P}$. We now check the 8 properties of Definition 3.1.

Example 3.1.2 (A2)

Solution (continued).
A2.

$$\begin{aligned} p(x) + q(x) &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n + b_n)x^n \\ &\quad + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0) \\ &\quad \text{(as above)} \\ &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (b_n + a_n)x^n \\ &\quad + (b_{n-1} + a_{n-1})x^{n-1} + \cdots + (b_1 + a_1)x + (b_0 + a_0) \\ &\quad \text{since addition is commutative in } \mathbb{R} \\ &= q(x) + p(x) \end{aligned}$$

Example 3.1.2 (A1)

Solution (continued).
A1. $(p(x) + q(x)) + r(x)$

$$\begin{aligned} &= ((a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) + (b_m x^m + b_{m-1}x^{m-1} \\ &\quad + \cdots + b_1 x + b_0)) + (c_\ell x^\ell + c_{\ell-1}x^{\ell-1} + \cdots + c_1 x + c_0) \\ &= (b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n + b_n)x^n \\ &\quad + (a_{n-1} + b_{n-1}x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0)) \\ &\quad + (c_\ell x^\ell + c_{\ell-1}x^{\ell-1} + \cdots + c_1 x + c_0) \\ &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n + b_n)x^n \\ &\quad + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_{\ell+1} + b_{\ell+1})x^{\ell+1} \\ &\quad + ((a_\ell + b_\ell) + c_\ell)x^\ell + ((a_{\ell-1} + b_{\ell-1}) + c_{\ell-1})x^{n-1} \\ &\quad + \cdots + ((a_1 + b_1) + c_1)x + ((a_0 + b_0) + c_0) \end{aligned}$$

Example 3.1.2 (A1) (continued)

Solution. A1. (continued) $(p(x)+q(x))+r(x)$

$$\begin{aligned} &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n+b_n)x^n \\ &\quad +(a_{n-1}+b_{n-1})x^{n-1} + \cdots + (a_{\ell+1}+b_{\ell+1})x^{\ell+1} \\ &\quad +(a_\ell + (b_\ell + c_\ell))x^\ell + (a_{\ell-1} + (b_{\ell-1}+c_{\ell-1}))x^{n-1} \\ &\quad + \cdots + (a_1 + (b_1+c_1))x + (a_0 + (b_0 + c_0)) \\ &\quad \text{since addition in } \mathbb{R} \text{ is associative} \\ &= (a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 + a_0) + (b_m x^m + b_{m-1}x^{m-1} + \cdots \\ &\quad + b_{\ell+1}x^{\ell+1} + (b_\ell + c_\ell)x^\ell + (b_{\ell-1}+c_{\ell-1})x^{\ell-1} + \cdots \\ &\quad +(b_1+c_1)x + (b_0+c_0)) \\ &= p(x) + (q(x)+r(x)). \end{aligned}$$

Notice that, by A2, we can permute $p(x)$, $q(x)$, and $r(x)$ and the associativity claim then holds in general (this is necessary to cover all cases of the relative degrees of the polynomials).

Example 3.1.2 (A3, A4)

Solution (continued).
A3. We take the zero vector as the polynomial with all coefficients 0: $0(x)=0$. Then

$$\begin{aligned} 0(x)+p(x) &= (0+a_n)x^n + (0+a_{n-1})x^{n-1} + \cdots + (0+a_1)x + (0+a_0) \\ &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \\ &\quad \text{since 0 is the additive identity in } \mathbb{R} \\ &= p(x). \end{aligned}$$

A4. For $p(x)$ as given, we define $-p(x) = (-a_n)x^n + (-a_{n-1})x^{n-1} + \cdots + (-a_1)x + (-a_0)$. Then

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

Example 3.1.2 (S1)

Solution (continued).
S1. We have:

$$\begin{aligned} s(p(x)+q(x)) &= s(b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (b_n + a_n)x^n \\ &\quad +(b_{n-1}+a_{n-1})x^{n-1} + \cdots + (b_1+a_1)x + (b_0+a_0)) \\ &= s(b_m)x^m + s(b_{m-1})x^{m-1} + \cdots + s(b_{n+1})x^{n+1} \\ &\quad +s(b_n+a_n)x^n + s(b_{n-1}+a_{n-1})x^{n-1} + \cdots \\ &\quad +s(b_1+a_1)x + s(b_0+a_0) \\ &= (sb_m)x^m + (sb_{m-1})x^{m-1} + \cdots + (sb_{n+1})x^{n+1} \\ &\quad +(sb_n + sa_n)x^n + (sb_{n-1}+sa_{n-1})x^{n-1} + \cdots \\ &\quad +(sb_1+sa_1)x + (sb_0+sa_0) \text{ since multiplication} \\ &\quad \text{distributes over addition in } \mathbb{R} \\ &= sp(x)+sq(x). \end{aligned}$$

Example 3.1.2 (S2)

Solution (continued).
S2. We have:

$$\begin{aligned} (s+t)p(x) &= (s+t)(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) \\ &= (s+t)a_n x^n + (s+t)a_{n-1}x^{n-1} + \cdots \\ &\quad +(s+t)a_1 x + (s+t)a_0 \\ &= (sa_n + ta_n)x^n + (sa_{n-1}+ta_{n-1})x^{n-1} + \cdots \\ &\quad +(sa_1+ta_1)x + (sa_0+ta_0) \\ &\quad \text{since multiplication distributes over addition in } \mathbb{R} \\ &= sp(x)+tp(x). \end{aligned}$$

Example 3.1.2 (S3)

Solution (continued).
S3. We have:

$$\begin{aligned} s(tp(x)) &= s(t(a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0)) \\ &= s(ta_nx^n + ta_{n-1}x^{n-1} + \cdots + ta_1x + ta_0) \\ &= s(ta_n)x^n + s(ta_{n-1})x^{n-1} + \cdots + s(ta_1)x + s(ta_0) \\ &= (st)a_nx^n + (st)a_{n-1}x^{n-1} + \cdots + (st)a_1x + (st)a_0 \\ &\quad \text{since multiplication is associative in } \mathbb{R} \\ &= (st)p(x). \end{aligned}$$

Example 3.1.2 (S4)

Solution (continued).
S4. We have:

$$\begin{aligned} 1p(x) &= 1(a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0) \\ &= (1a_n)x^n + (1a_{n-1})x^{n-1} + \cdots + (1a_1)x + (1a_0) \\ &= a_nx^n + a_{n-1}x^n + \cdots + a_1x + a_0 \\ &\quad \text{since 1 is the multiplicative identity in } \mathbb{R} \\ &= p(x). \end{aligned}$$

So all properties of Definition 3.1 are satisfied and $\mathcal{P}$ is a vector space. □

Page 189 Number 6

Page 189 Number 6. Consider the set $\mathcal{F}$ of all functions mapping $\mathbb{R}$ into $\mathbb{R}$, with scalar multiplication defined for scalar $r \in \mathbb{R}$ and $f \in \mathcal{F}$ as $(rf)(x) = rf(x)$, and vector addition $\tilde{+}$ defined as $(f \,\tilde{+}\, g)(x) = \max\{f(x), g(x)\}$. Is $\mathcal{F}$ a vector space?

Solution. The peculiar way of adding vectors yields some problems. For example there can be no additive identity and A3 does not hold. To see this, let $k \in \mathbb{R}$ be a constant and consider the constant function $f(x) = k$. If $e(x)$ is the additive identity in $\mathcal{F}$ then $e(x) \,\tilde{+}\, f(x) = \max\{e(x), f(x)\} = f(x) = k$ for all $x \in \mathbb{R}$. So it must be that $e(x) \leq k$ for all $x \in \mathbb{R}$. Since $k \in \mathbb{R}$ is arbitrary, we have that $e(x) \leq k$ for all $k \in \mathbb{R}$ and for all $x \in \mathbb{R}$. But then there is no value that can be assigned to $e(x)$ for any $x \in \mathbb{R}$ and so no identity vector exists. So **NO, $\mathcal{F}$ is not a vector space.** □

Theorem 3.1

Theorem 3.1. Elementary Properties of Vector Spaces.
Every vector space V satisfies:
1. the vector $\vec{0}$ is the unique additive identity in a vector space,
3. if $\vec{u} + \vec{v} = \vec{u} + \vec{w}$ then $\vec{v} = \vec{w}$,

Proof. 1. Suppose that there are two additive identities, $\vec{0}$ and $\vec{0}'$. Then consider:

$$\begin{aligned} \vec{0} &= \vec{0} + \vec{0}' \text{ (since } \vec{0}' \text{ is an additive identity)} \\ &= \vec{0}' \text{ (since } \vec{0} \text{ is an additive identity).} \end{aligned}$$

Therefore, $\vec{0} = \vec{0}'$ and the additive identity is unique.

Theorem 3.1 (continued)

Theorem 3.1. Elementary Properties of Vector Spaces.
Every vector space V satisfies:
1. the vector $\vec{0}$ is the unique additive identity in a vector space,
3. if $\vec{u}+\vec{v}=\vec{u}+\vec{w}$ then $\vec{v}=\vec{w}$,

Proof (continued).
3. Suppose $\vec{u}+\vec{v}=\vec{u}+\vec{w}$. Then we add $-\vec{u}$ to both sides of the equation and we get:

$$\begin{aligned}(\vec{u}+\vec{v})+(-\vec{u}) &= (\vec{u}+\vec{w})+(-\vec{u})\\ (\vec{v}+\vec{u})+(-\vec{u}) &= (\vec{w}+\vec{u})+(-\vec{u}) \text{ by commutivity, A2}\\ \vec{v}+(\vec{u}-\vec{u}) &= \vec{w}+(\vec{u}-\vec{u}) \text{ by associativity, A1}\\ \vec{v}+\vec{0} &= \vec{w}+\vec{0} \text{ by additive inverse, A4}\\ \vec{v} &= \vec{w} \text{ by additive identity, A3.}\end{aligned}$$

The conclusion holds. □

Page 190 Number 24

Page 190 Number 24. Let V be a vector space and let $\vec{v}$ and $\vec{w}$ be nonzero vectors in V. Prove that if $\vec{v}$ is not a scalar multiple of $\vec{w}$, then $\vec{v}$ is not a scalar multiple of $\vec{v}+\vec{w}$.

Proof. We consider the (logically equivalent) contrapositive of the claim: If $\vec{v}$ is a scalar multiple of $\vec{v}+\vec{w}$ then $\vec{v}$ is a scalar multiple of $\vec{w}$. We prove this and then the original claim follows.

Suppose $\vec{v}$ is a scalar multiple of $\vec{v}+\vec{w}$, say $\vec{v}=r(\vec{v}+\vec{w})$ where $r\in\mathbb{R}$ is a scalar. Then $\vec{v}=r\vec{v}+r\vec{w}$ by S1 and so $\vec{v}-r\vec{v}=(r\vec{v}+r\vec{w})-r\vec{v}$ or

$$\begin{aligned}(1-r)\vec{v} &= -r\vec{v}+(r\vec{v}+r\vec{w}) \text{ by S2 and A2}\\ &= (-r\vec{v}+r\vec{v})+r\vec{w} \text{ by A1}\\ &= \vec{0}+r\vec{w} \text{ by A4}\\ &= r\vec{w} \text{ by A3.} \qquad (*)\end{aligned}$$

If $r=1$ then $0\vec{v}=1\vec{w}$ or $\vec{0}=\vec{w}$ (by S4 and Theorem 3.1(4)), but $\vec{w}$ is a nonzero vector by hypothesis, so $r\neq 1$.

Page 190 Number 24 (continued)

Page 190 Number 24. Let V be a vector space and let $\vec{v}$ and $\vec{w}$ be nonzero vectors in V. Prove that if $\vec{v}$ is not a scalar multiple of $\vec{w}$, then $\vec{v}$ is not a scalar multiple of $\vec{v}+\vec{w}$.

Proof. Then $(*)$, $(1-r)\vec{v}=r\vec{w}$, implies that

$$\frac{1}{1-r}((1-r)\vec{v})=\frac{1}{1-r}(r\vec{w})$$

or, by S3, $1\vec{v}=\frac{r}{1-r}\vec{w}$ or, by S4, $\vec{v}=\frac{r}{1-r}\vec{w}$. That is, $\vec{v}$ is a scalar multiple of $\vec{w}$, as claimed. □

Page 190 Number 26

Page 190 Number 26. Use the universality of function spaces to explain how we can view the Euclidean vector space $\mathbb{R}^{mn}$ and the vector space $M_{m,n}$ of all $m\times n$ matrices as essentially the same vector space with just a different notation for the vectors.

Solution. We saw in the previous note that we can use set $S=\{(1,1),(1,2),\ldots,(1,n),(2,1),$ $(2,2),\ldots,(2,n),(3,1),(3,2),\ldots,(m-1,n),(m,1),(m,2),\ldots,(m,n)\}$ and function $f:S\to\mathbb{R}$ to represent an $m\times n$ matrix as

$$M_f=\begin{bmatrix} f((1,1)) & f((1,2)) & \cdots & f((1,n))\\ f((2,1)) & f((2,2)) & \cdots & f((2,n))\\ \vdots & \vdots & \ddots & \vdots\\ f((m,1)) & f((m,2)) & \cdots & f((m,n))\end{bmatrix}.$$

We can also use function $f:S\to\mathbb{R}$ to represent a vector in $\mathbb{R}^{mn}$ as $\vec{v}_f=[f((1,1)),f((2,1)),\ldots,f((m,1)),f((1,2)),f((2,2)),\ldots,f((m,n))]$.

Page 190 Number 26 (continued)

Solution (continued). For k with $1 \leq k \leq mn$, we can write k as $k = i + (j-1)m$ for some j with $1 \leq j \leq n$ and some i with $1 \leq i \leq m$ (this is the "Division Algorithm"). So the kth component of vector $\vec{v}_f$ equals the (i,j) entry of M (and conversely). When matrix M_f is multiplied by a scalar r, the (i,j) entry of M is $rf((i,j))$. When vector $\vec{v}_f$ is multiplied by a scalar r, the kth component of $r\vec{v}_f$ is $rf((i,j))$ where $k = i + (j-1)m$ as above. So scalar multiplication "behaves" in the same way on M_f and $\vec{v}_f$. If matrix M_g and vector $\vec{v}_g$ are similarly defined using function $g : S \to \mathbb{R}$ then the (i,j) entry of matrix $M_f + M_g$ is $f((i,j)) + g((i,j))$. The kth component of $\vec{v}_f + \vec{v}_g$ is $f((i,j)) + g((i,j))$ where $k = i + (j-1)m$ as above. So vector/matrix addition "behaves" the same way as well. The two basic properties of a vector space are scalar multiplication and vector addition. Since these are the same (or "behave" the same) then the vector spaces $\mathbb{R}^{mn}$ and $M_{m,n}$ are essentially the same.
☐ **Note.** We clarify this "essentially the same" idea in the Section 3.3, "Coordinatization of Vectors," when we define a vector space *isomorphism*.