w/ S&B PAGE #'s

Vector Spaces

Definition: The **usual addition and scalar multiplication** of n-tuples $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ (also called **vectors**) are the addition and scalar multiplication operations defined component-wise:

$$\mathbf{x} + \mathbf{y} := (x_1 + y_1, \ldots, x_n + y_n) \text{ and } \lambda\mathbf{x} := (\lambda x_1, \ldots, \lambda x_n).$$

Remark: The usual addition and scalar multiplication in $\mathbb{R}^n$ are functions:

$$\text{Addition} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$
$$\text{Scalar multiplication} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n.$$

P.751

Remark: The usual addition and scalar multiplication in $\mathbb{R}^n$ have the following properties:

(VS1) $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n : \mathbf{x} + \mathbf{y} \in \mathbb{R}^n$. ($\mathbb{R}^n$ is **closed under vector addition**.)

(VS2) $\forall \lambda \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n : \lambda\mathbf{x} \in \mathbb{R}^n$. ($\mathbb{R}^n$ is **closed under scalar multiplication**.)

(VS3) $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n : \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$. (Vector addition is **commutative**.)

(VS4) $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n : (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$,
and $\forall \lambda, \mu \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n : \lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$. (Both operations are **associative**.)

(VS5) $\forall \lambda \in \mathbb{R}, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n : \lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$,
and $\forall \lambda, \mu \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n : (\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$. (The operations are **distributive**.)

(VS6) $\exists \widehat{\mathbf{x}} \in \mathbb{R}^n : \forall \mathbf{x} \in \mathbb{R}^n : \widehat{\mathbf{x}} + \mathbf{x} = \mathbf{x}$. (Note that $\widehat{\mathbf{x}}$ is the origin of $\mathbb{R}^n$, *viz.* $\mathbf{0}$. It's called the **additive identity**.)

(VS7) $\forall \mathbf{x} \in \mathbb{R}^n : \exists \mathbf{x}' \in \mathbb{R}^n : \mathbf{x} + \mathbf{x}' = \mathbf{0}$. (Note that for each $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}'$ is $-\mathbf{x}$. It's called the **additive inverse** of $\mathbf{x}$.)

(VS8) $\forall \mathbf{x} \in \mathbb{R}^n : \lambda\mathbf{x} = \mathbf{x}$ for the scalar $\lambda = 1$.

This particular algebraic structure — operations that satisfy (VS1) - (VS8) — is not unique to $\mathbb{R}^n$. In fact, it's pervasive in mathematics (and in economics and statistics). So we generalize in the following definition and say that any set V with operations that behave in this way — *i.e.*, that satisfy (VS1) - (VS8) — is a **vector space**. And in order to highlight definitions, theorems, etc., that are about general vector spaces (and not just about $\mathbb{R}^n$), I'll indicate these general propositions by this symbol: ♦.

P.771

Definition:♦ A **vector space** is a set V together with operations

$$\text{Addition} : V \times V \to V$$
$$\text{Scalar multiplication} : \mathbb{R} \times V \to V$$

that satisfy the conditions (VS1) - (VS8) if $\mathbb{R}^n$ is replaced throughout with V.

Notation:♦ In any vector space V, we denote the additive identity by $\mathbf{0}$ and the additive inverse of any $\mathbf{x} \in V$ by $-\mathbf{x}$. We'll use boldface for vectors and regular font for scalars and other numbers.

Some examples of vector spaces are:

(1) $\mathbb{M}_{m,n}$, the set of all $m \times n$ matrices, with component-wise addition and scalar multiplication.

(2) $\mathbb{R}^\infty$, the set of all sequences $\{x_k\}$ of real numbers, with operations defined component-wise.

(3) The set $\mathcal{F}$ of all real functions $f : \mathbb{R} \to \mathbb{R}$, with $f + g$ and λf defined by

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x),\ \forall x \in \mathbb{R} \\ (\lambda f)(x) &:= \lambda f(x),\ \forall x \in \mathbb{R}. \end{aligned}$$

(4) For any set X, the set of all real-valued functions on X, with operations defined as in (3).

(5) The set S of all solutions $(x_1, \ldots, x_n)$ of the $m \times n$ system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots& \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0, \end{aligned}$$

i.e., the set of all solutions of the matrix equation $A\mathbf{x} = \mathbf{0}$, where a_{ij} is the i, j^{th} element of A.

Notice that the set S in (5) consists of n-tuples of real numbers, so S is a subset of $\mathbb{R}^n$. Assuming that S retains the component-wise definitions of the operations from $\mathbb{R}^n$, we can show that S is a vector space in its own right — a *vector subspace* of $\mathbb{R}^n$. In order to verify that S is a vector space we have to verify that S satisfies the conditions (VS1) - (VS8). The conditions (VS3) - (VS5) are satisfied automatically, since S retains the operations from $\mathbb{R}^n$, where we know (VS3) - (VS5) are satisfied. (VS6) is satisfied, because $\mathbf{0} \in S$ — *i.e.*, $A\mathbf{0} = \mathbf{0}$. (VS7) is satisfied, because if $\mathbf{x}$ is a solution of the equation system, then so is $-\mathbf{x}$ — *i.e.*, $A\mathbf{x} = \mathbf{0} \Rightarrow A(-\mathbf{x}) = \mathbf{0}$. (VS8) is trivially satisfied. What about (VS1) and (VS2)? We verify that S is closed under vector addition: $\mathbf{x}, \mathbf{y} \in S \Rightarrow \mathbf{x} + \mathbf{y} \in S$ — *i.e.*, $[A\mathbf{x} = \mathbf{0} \,\&\, A\mathbf{y} = \mathbf{0}] \Rightarrow A(\mathbf{x} + \mathbf{y}) = \mathbf{0}$. And we verify that S is closed under scalar multiplication: $[\lambda \in \mathbb{R}, \mathbf{x} \in S] \Rightarrow \lambda\mathbf{x} \in S$ — *i.e.*, $A\mathbf{x} = \mathbf{0} \Rightarrow A\lambda\mathbf{x} = \mathbf{0}$. We've verified that S satisfies all eight of the conditions that define a vector space.

To a large extent, the subject of vector spaces is about situations like the one in the preceding paragraph, where a subset S of a vector space V turns out to be itself a vector space — a **vector subspace** of V (also called a **linear subspace** of V, or just a **subspace** of V).

P.752 **Definition:**♦ A **subspace** of a vector space V is a subset of V which is itself a vector space, under the same operations as in V.

The following remark tells us that in order to determine whether a nonempty subset of a vector space V is actually a subspace of V, it's sufficient to merely check whether it satisfies conditions (VS1) and (VS2) — whether it's closed under vector addition and scalar multiplication. If it is, then its operations, under which V is a vector space, must necessarily satisfy (VS3) - (VS8).

P.752
(BUT THEY OMITTED "NONEMPTY")

Remark:♦ If V is a vector space and $S \subseteq V$, then S is a subspace of V if and only if it is nonempty and satisfies (VS1) and (VS2) — *i.e.*, if and only if it is nonempty and is closed under vector addition and scalar multiplication.

Proof: Obviously, if S is itself a vector space, then it satisfies (VS1) and (VS2). Conversely, we must show that if S satisfies (VS1) and (VS2), then it satisfies the remaining conditions (VS3) - (VS8) as well. Conditions (VS3) - (VS5) are immediate from the fact that S has the same operations as V and that V satisfies (VS3) - (VS5). You should be able to establish on your own that S satisfies (VS6), (VS7), and (VS8), given that it is nonempty and satisfies (VS1) - (VS5). ||

Question:♦ Why does the Remark include the restriction that S be nonempty, while the definition of a subspace doesn't include that restriction? Could the empty set satisfy (VS1) and (VS2)? Could the empty set satisfy (VS3) - (VS8)?

Exercise:♦ Verify that if a nonempty subset S of a vector space V satisfies (VS1) - (VS5), then it satisfies (VS6), (VS7), and (VS8).

Some additional examples of subspaces:

(6) The set $\mathcal{Q}$ of all quadratic forms on $\mathbb{R}^n$ is a subspace of the vector space V of all functions $f : \mathbb{R}^n \to \mathbb{R}$. (Recall that there is a one-to-one correspondence between quadratic forms and symmetric matrices.)

(7) A sequence $\{x_k\}$ of real numbers is said to be **bounded** if there is a number M such that $|x_k| \leqq M$ for every $k = 1, 2, \ldots$. The set B of bounded sequences is a subspace of the space $\mathbb{R}^\infty$ of all real sequences. It may help here to write a sequence as $\mathbf{x} = (x_1, x_2, x_3, \ldots) \in \mathbb{R}^\infty$.

(8) The set C of all continuous real functions, with operations defined as in (3), is a subspace of the vector space $\mathcal{F}$ in (3). In order to verify (VS1), we have to show that the sum of any two continuous real functions is continuous (*i.e.*, C is closed under vector addition). In order to verify (VS2), we have to show that any multiple λf of a continuous function is continuous (*i.e.*, C is closed under scalar multiplication). What is the additive identity in this vector space?

Exercise: Verify that the sets in (6) and (7) are subspaces of the respective vector spaces.

Here are some examples of subsets $S \subseteq \mathbb{R}^2$ that are *not* subspaces of $\mathbb{R}^2$:

(9) S is any singleton other than $\mathbf{0}$ — for example, $S = \{(1,1)\}$.

(10) $S = \{(0,0), (1,1)\}$.

(11) $S = \{(0,0), (1,1), (-1,-1)\}$.

(12) $S = \{(x_1, x_2) \mid x_1 + x_2 = 1\}$.

(13) $S = \{\mathbf{x} \in \mathbb{R}^2 \mid |x_2| = |x_1|\}$.

Exercise: In each of the examples (9) - (13) determine which of the conditions (VS1), (VS2), (VS6), (VS7), and (VS8) are satisfied and which are violated. You'll probably find it helpful to draw diagrams of the sets S.

P.753, P.238

Definition:♦ In a vector space V, let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a nonempty finite set of vectors. A vector $\mathbf{w} \in V$ is a **linear combination** of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ if there are scalars $a_1, \ldots, a_m$ such that $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m$.

P.237, P.244

Definition:♦ In a vector space V, let $S \subseteq V$ be a (not necessarily finite) nonempty subset of V. The **span** of S, denoted $\mathcal{L}[S]$, is the set of all linear combinations of members of S.

Note that while the set S in this definition can be infinite, every linear combination of members of S must be a linear combination of only a finite number of members of S, according to the definition of linear combination. We don't define linear combinations of an infinite number of vectors.

P.753 (EXAMPLE 27.2)

Theorem:♦ Let V be a vector space and let S be a nonempty subset of V. The span of S, $\mathcal{L}[S]$, is a subspace of V.

Proof: To show that $\mathcal{L}[S]$ is closed under scalar multiplication, let $\mathbf{x} \in \mathcal{L}[S]$ and let $\lambda \in \mathbb{R}$. Then $\mathbf{x} = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m$ for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in S$ and some numbers $a_1, \ldots, a_m \in \mathbb{R}$. Then $\lambda\mathbf{x} = \lambda a_1\mathbf{v}_1 + \cdots + \lambda a_m\mathbf{v}_m$ — *i.e.*, $\lambda\mathbf{x}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and is therefore in $\mathcal{L}[S]$. To show that $\mathcal{L}[S]$ is closed under vector addition, let $\mathbf{x}$ and $\mathbf{y}$ be members of $\mathcal{L}[S]$, and we show that $\mathbf{x} + \mathbf{y} \in \mathcal{L}[S]$. We have $\mathbf{x} = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m$ and $\mathbf{y} = b_1\mathbf{w}_1 + \cdots + b_K\mathbf{w}_K$ for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_1, \ldots, \mathbf{w}_K \in S$ and some $a_1, \ldots, a_m, b_1, \ldots, b_K \in \mathbb{R}$. Therefore $\mathbf{x} + \mathbf{y} = \sum_{i=1}^m a_i\mathbf{v}_i + \sum_{k=1}^K b_k\mathbf{w}_k = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m + b_1\mathbf{w}_1 + \cdots + b_K\mathbf{w}_K$, which is a linear combination of members of S, and is therefore in $\mathcal{L}[S]$. (Note that this is still correct even if some of the vectors $\mathbf{v}_i$ and $\mathbf{w}_k$ coincide.) ||

Because $\mathcal{L}[S]$ is always a vector space, it is sometimes referred to as the vector space **spanned by** the set S, or the vector space **generated by** the set S.

P.240

Definition:♦ In a vector space V, a set S of vectors is **linearly dependent** if some vector in S is a linear combination of other vectors in S. A set S is **linearly independent** if it is not linearly dependent — *i.e.*, if no vector in S is a linear combination of other vectors in S.

Remark:♦ A set S of vectors is linearly dependent if and only if the vector $\mathbf{0}$ is a linear combination of some of the members of S. A set S is linearly independent if and only if

$$[\mathbf{v}_1, \ldots, \mathbf{v}_m \in S \;\;\&\;\; a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}] \Rightarrow a_1 = \cdots = a_m = 0.$$

P. 247

Definition:♦ A **basis** of a vector space V is a linearly independent subset of V that spans V.

P.755 (THM 27.2)

Theorem:♦ Let $\mathcal{B}$ be a finite basis of a vector space V. Any set $S \subseteq V$ containing more vectors than $\mathcal{B}$ is linearly dependent.

Proof: Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ and let $\mathbf{w}_1, \ldots, \mathbf{w}_n \in S$, with $n > m$. We will show that the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ are linearly dependent. Since $\mathcal{B}$ is a basis of V, each vector $\mathbf{w}_j$ can be written as

$$\mathbf{w}_j = a_{1j}\mathbf{v}_1 + \cdots + a_{mj}\mathbf{v}_m \quad \text{for some numbers } a_{1j}, \ldots, a_{mj}.$$

For any linear combination $c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$ of the $\mathbf{w}_j$ vectors, we have

$$\begin{aligned} c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n &= c_1(a_{11}\mathbf{v}_1 + \cdots + a_{m1}\mathbf{v}_m) + \cdots + c_n(a_{1n}\mathbf{v}_1 + \cdots + a_{mn}\mathbf{v}_m) \\ &= \left(\sum_{j=1}^{n} a_{1j}c_j\right)\mathbf{v}_1 + \cdots + \left(\sum_{j=1}^{n} a_{mj}c_j\right)\mathbf{v}_m \\ &= (\mathbf{a}_1 \cdot \mathbf{c})\mathbf{v}_1 + \cdots + (\mathbf{a}_m \cdot \mathbf{c})\mathbf{v}_m, \end{aligned}$$

where $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$, $i = 1, \ldots, m$, and $\mathbf{c} = (c_1, \ldots, c_n)$.

Now assume that $c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n = \mathbf{0} \in V$, which yields $(\mathbf{a}_1 \cdot \mathbf{c})\mathbf{v}_1 + \cdots + (\mathbf{a}_m \cdot \mathbf{c})\mathbf{v}_m = \mathbf{0}$. Since the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent ($\mathcal{B}$ is a basis), we must have $\mathbf{a}_i \cdot \mathbf{c} = 0$ for all $i = 1, \ldots, m$. Because $n > m$, this $m \times n$ equation system has solutions $\mathbf{c} \neq \mathbf{0}$. Equivalently, let A be the $m \times n$ matrix whose rows are the n-tuples $\mathbf{a}_i$; then, because $n > m$, the equation $A\mathbf{c} = \mathbf{0}$ has solutions $\mathbf{c} \neq (0, \ldots, 0)$. Therefore the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ are linearly dependent. ||

P.756 (THM 27.3) P.248

Corollary:♦ If a vector space V has a finite basis, then every basis of V has the same number of vectors.

This corollary ensures that the following notion of the **dimension** of a (finite-dimensional) vector space is well-defined:

P.756

Definition:♦ The **dimension** of a vector space that has a finite basis $\mathcal{B}$ is the number of vectors in $\mathcal{B}$. We write $\dim V$ for the dimension of V.

Remark:♦ If $\dim V = n$ and the n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V.

Proof: Exercise.

Definition:♦ If a vector space has a finite basis, it is said to be **finite-dimensional**; otherwise it is said to be **infinite-dimensional**.

Examples:

P.248 (EXAMPLE 11.8)

(14) For the vector space $\mathbb{R}^n$ the n **unit vectors**

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ \mathbf{e}_n = (0, 0, \ldots, 0, 1)$$

constitute a basis. Therefore (fortunately!) $\dim \mathbb{R}^n = n$.

(15) Let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and let $S = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} = 0\}$, the set of solutions of the linear equation $a_1x_1 + \cdots + a_nx_n = 0$. We've already seen that S is a subspace of $\mathbb{R}^n$. Now we also have $\dim S = n - 1$: S is an $(n-1)$-dimensional hyperplane in $\mathbb{R}^n$ that contains $\mathbf{0}$.

Exercise: Verify in Example (14) that the unit vectors form a basis of $\mathbb{R}^n$, and verify in Example (15) that the dimension of the set S is $n-1$.

Note that we've defined the dimension of a finite-dimensional vector space, but we haven't defined the dimension of an infinite-dimensional vector space. This is not an oversight. Infinite-dimensional vector spaces are important — in fact, we've already seen four important ones: the set $\mathcal{F}$ of all real functions; the set C of all continuous real functions; the set $\mathbb{R}^\infty$ of all sequences of real numbers; and the set B of all bounded sequences of real numbers. And it's a fact (but a fact not so easily proved!) that every vector space does have a basis. But for an infinite-dimensional space V, it's typically not clear how to identify, or enumerate, any of its bases, like we did above for $\mathbb{R}^n$. So while we'll be working with some important infinite-dimensional vector spaces, we won't be using the concept of a basis for them.

Exercise: Show that in $\mathbb{R}^\infty$ the unit vectors

$$\mathbf{e}_1 = (1,0,0,\ldots),\ \mathbf{e}_2 = (0,1,0,0,\ldots),\ \mathbf{e}_3 = (0,0,1,0,0,\ldots),\ \ldots$$

do not constitute a basis, unlike the situation in $\mathbb{R}^n$. (**Hint:** Recall that linear combinations are defined only for finite sets of vectors.) Also note that here (as we'll often do) we're writing a sequence as $(x_1, x_2, x_3, \ldots)$ instead of $\{x_k\}$.

Linear Functions

P.287

Definition:♦ Let V and W be vector spaces. A function $f : V \to W$ is a **linear function** if

$$(*) \qquad \forall \mathbf{x}, \mathbf{y} \in V : f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \qquad \text{and} \qquad \forall \mathbf{x} \in V, \forall \lambda \in \mathbb{R} : f(\lambda \mathbf{x}) = \lambda f(\mathbf{x}).$$

A linear function from a vector space into itself — *i.e.*, $f : V \to V$ — is sometimes called a **linear transformation**.

Theorem:♦ If V and W are vector spaces and $f : V \to W$ is a linear function, then for every subspace S of V, $f(S)$ is a subspace of W.

Proof: Exercise.

For linear functions on $\mathbb{R}^n$, the following theorem is the fundamental characterization in terms of matrices:

P.288 (THM 13.2)

Theorem: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if there is a unique $m \times n$ matrix A such that $\forall \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = A\mathbf{x}$.

Proof: If $\forall \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = A\mathbf{x}$, it's trivial that $(*)$ is satisfied, and f is therefore linear. Conversely, suppose f is linear; we show that there is a matrix A such that $f(\mathbf{x}) = A\mathbf{x}$. For each $j = 1, \ldots, n$ let $\mathbf{e}_j$ be the j-th unit vector in $\mathbb{R}^n$, and let $\mathbf{a}_j = f(\mathbf{e}_j)$; note that $\mathbf{a}_j \in \mathbb{R}^m$. Define A to be the $m \times n$ matrix in which the column vector $\mathbf{a}_j$ is the j-th column. Then we have

$$\begin{aligned} f(\mathbf{x}) &= f(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) \\ &= f(x_1\mathbf{e}_1) + \cdots + f(x_n\mathbf{e}_n), \quad \text{because } f \text{ is linear} \\ &= x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n), \quad \text{because } f \text{ is linear} \\ &= x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n \\ &= A\mathbf{x}. \end{aligned}$$

To see that A is unique, note that if $f(\mathbf{x}) = A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$, then the columns $\mathbf{a}_j$ and $\mathbf{b}_j$ of A and B satisfy $\mathbf{a}_j = f(\mathbf{e}_j) = \mathbf{b}_j$. $\|$

Corollary: A real-valued linear function $f : \mathbb{R}^n \to \mathbb{R}$ has the form $f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} = \sum_{j=1}^n a_j x_j$.

I like the following excerpt from Simon & Blume (p. 288): "[The above characterization theorem] underlines the one-to-one correspondence between linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $m \times n$ matrices. Each linear function f is [defined by] a unique $m \times n$ matrix. This fact will play an important role in the rest of this book. Matrices are not simply rectangular arrays of numbers .... Matrices are representations of linear functions. When we use calculus to do its main task, namely to approximate a nonlinear function F at a given point by a linear function, [namely] the derivative of F, we will write that linear function as a matrix. In other words, derivatives of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ are $m \times n$ matrices."