Vectors

In what follows n will always be a positive integer.

A **scalar** is a real number. An n**-vector** is an ordered n-tuple of scalars. When it is clear from the context what n is we will frequently say "vector" instead of "n-vector". If $(x^1, \ldots, x^n)$ is an n-vector then, for each $i = 1, \ldots, n$, the scalar x^i is called the i**-th component** of the n-vector. Let

$$\mathbf{R}^n$$

denote the set of vectors. For each $i = 1, \ldots, n$ we let $\mathbf{e}^i$ be the scalar valued function on $\mathbf{R}^n$ which assigns to each n-vector its i-th component.

Given a vector $\mathbf{x}$ and a scalar c we define

$$c\mathbf{x}$$

to be the vector whose i-th component is c times the i-th component of $\mathbf{x}$, $i = 1, \ldots, n$; $c\mathbf{x}$ is called **scalar multiplication of x by** c. Given a vector $\mathbf{x}$ and a vector $\mathbf{y}$ we define

$$\mathbf{x} + \mathbf{y}$$

to be the vector whose i-th component is the sum of the i-th component of $\mathbf{x}$ and the i-th component of $\mathbf{y}$, $i = 1, \ldots, n$; $\mathbf{x} + \mathbf{y}$ is called the **vector sum of x and y**.

We let $\mathbf{0} \in \mathbf{R}^n$ be the n-vector all of whose components are zero; we call this vector the **zero vector**. Given an n-vector $\mathbf{x}$ we set $-\mathbf{x} = (-1)\mathbf{x}$.

Make sure you understand the geometric and physical interpretation of these operations! Check out the book for lots of pictures. Better still, draw some of your own.

We urge the reader to verify the following properties of the vector operations we have just introduced:

Suppose c, d are scalars and $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are n-vectors. Then

(v1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;

(v2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$;

(v3) $\mathbf{x} + \mathbf{0} = \mathbf{x}$;

(v4) $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;

(v5) $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$;

(v6) $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$;

(v7) $c(d\mathbf{x}) = (cd)\mathbf{x}$;

(v8) $1\mathbf{x} = \mathbf{x}$.

For each $i = 1, \ldots, n$ we let

$$\mathbf{e}_i$$

be the vector whose i-th component is 1 and whose other components are 0 and we let

$$\mathbf{e}^i$$

be the function with domain $\mathbf{R}^n$ which assigns to a n-vector its i-th component. Note that

$$\mathbf{x} \;=\; \mathbf{e}^1(\mathbf{x})\mathbf{e}_1 + \cdots + \mathbf{e}^n(\mathbf{x})\mathbf{e}_n \;=\; \sum_{i=1}^{n} \mathbf{e}^i(\mathbf{x})\mathbf{e}_i \qquad \text{whenever } \mathbf{x} \in \mathbf{R}^n.$$

The dot product. Given n-vectors $\mathbf{x} = (x^1 \ldots, x^n)$ and $\mathbf{y} = (y^1, \ldots, y^n)$ we set

$$\mathbf{x} \bullet \mathbf{y} = \sum_{j=1}^{n} x^j y^j$$

and call this scalar the **dot product of x and y**. As you will see, the dot product allows us to deal with lengths and angles. We urge the reader to verify the following properties of the dot product:

Suppose c, d are scalars and $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are n-vectors. Then

(d1) $(\mathbf{x} + \mathbf{y}) \bullet \mathbf{z} = \mathbf{x} \bullet \mathbf{z} + \mathbf{y} \bullet \mathbf{z}$;

(d2) $(c\mathbf{x}) \bullet \mathbf{y} = c(\mathbf{x} \bullet \mathbf{y})$;

(d3) $\mathbf{x} \bullet \mathbf{y} = \mathbf{y} \bullet \mathbf{x}$;

(d4) $\mathbf{x} \bullet \mathbf{x} \geq 0$ with equality only if $\mathbf{x} = \mathbf{0}$.

Note that (1) and (3) imply that

(d1') $\mathbf{x} \bullet (\mathbf{y} + \mathbf{z}) = \mathbf{x} \bullet \mathbf{y} + \mathbf{x} \bullet \mathbf{z}$

and that (2) and (3) imply that

(d2') $\mathbf{x} \bullet (c\mathbf{y}) = c(\mathbf{x} \bullet \mathbf{y})$.

Keeping in mind (4),for each $\mathbf{x} \in \mathbf{R}^n$ we set

$$|\mathbf{x}| = \sqrt{\mathbf{x} \bullet \mathbf{x}}$$

and call this nonegative real number the **norm** or **length of x**. Evidently,

$$|c\mathbf{x}| = |c||\mathbf{x}| \qquad \text{whenever } c \in \mathbf{R} \text{ and } \mathbf{x} \in \mathbf{R}^n.$$

The use of this terminology is justified by the Pythagorean Theorem from Euclidean geometry, the relevance of which to the real world has been established by millenia of experience. All good things about the dot product, and there are many, follow from the

Cauchy-Schwarz Inequality. Suppose $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. Then

$$|\mathbf{x} \bullet \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|$$

with equality only if there is a scalar c such that either $\mathbf{y} = c\mathbf{x}$ or $\mathbf{x} = c\mathbf{y}$.

Proof. We may assume that $\mathbf{y}$ is nonzero since otherwise the assertion holds trivially. For any $t \in \mathbf{R}$ we have

$$(1) \qquad 0 \leq (\mathbf{x} + t\mathbf{y}) \bullet (\mathbf{x} + t\mathbf{y}) = |\mathbf{x}|^2 + 2t(\mathbf{x} \bullet \mathbf{y}) + t^2|\mathbf{y}|^2;$$

We leave as **Exercise 1** for the reader to prove this using (d1)-(d4) above. Substitute

$$t = -\frac{\mathbf{x} \bullet \mathbf{y}}{|\mathbf{y}|^2}$$

and transpose a bit to get the inequality. The remaining assertion follows by observing that if equality holds in (1) then $(\mathbf{x} + t\mathbf{y}) \bullet (\mathbf{x} + t\mathbf{y})$ must be zero which forces $\mathbf{x} + t\mathbf{y} = \mathbf{0}$ by (d4). $\square$

Important Remark. Note that the proof depended only on (d1)-(d4) and *not* the definition of the dot product. Thus the Cauchy-Schwarz Inequality holds in any context where (d1)-(d4) (and, implicitly, (v1)-(v8)) hold. For example, given continuous $f, g : [a, b] \to \mathbf{R}$ we could set

$$f \bullet g = \int_a^b f(x)g(x)\,dx$$

and conclude that

$$\Big|\int_a^b f(x)g(x)\,dx\Big| \le \Big(\int_a^b f(x)^2\,dx\Big)^{1/2}\Big(\int_a^b g(x)^2\,dx\Big)^{1/2}$$

with equality only $f = cg$ or $g = cf$ for some constant c. The utility of the above dot product on functions is amazing; modern communications depend on the mathematics derived from it.

The triangle inequality. Suppose $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. Then

$$|\mathbf{x} + \mathbf{y}| \le |\mathbf{x}| + |\mathbf{y}|$$

with equality only if there is a scalar c such that either $\mathbf{y} = c\mathbf{x}$ or $\mathbf{x} = c\mathbf{y}$.

Proof. Square both sides and use the Schwarz Inequality. Give the details in **Exercise 2**. $\square$

Orthogonal projection onto a line through 0 and a nonzero vector. Suppose $\mathbf{x}$ is a nonzero n-vector. Let

$$\mathbf{u} = \frac{1}{|\mathbf{x}|}\mathbf{x}.$$

$\mathbf{u}$ is of unit length and is called the **normalization of x**. We let

$$\mathbf{proj}_{\mathbf{x}}(\mathbf{y}) = \mathbf{x} \bullet \mathbf{u}\,\mathbf{u} = \frac{\mathbf{x} \bullet \mathbf{y}}{|\mathbf{x}|^2}\,\mathbf{x} \qquad \text{for each } n\text{-vector } \mathbf{y}$$

and call this vector the **orthogonal projection of y on the line passing through 0 and x**. Let us explain. Let

$$L = \{t\mathbf{x} : t \in \mathbf{R}\}.$$

By taking $t = 0$ and $t = 1$ we find that $\mathbf{0} \in L$ and $\mathbf{x} \in L$, respectively. Moreover, L is a line; in fact, it is the unique line containing $\mathbf{0}$ and $\mathbf{x}$. (We will define what a **line** is shortly; maybe you can do it now.) We claim that $\mathbf{proj}_{\mathbf{x}}(\mathbf{y})$ is that point on L which is closest to $\mathbf{y}$. To see this we set

$$S(t) = |\mathbf{y} - t\mathbf{x}|^2 = |\mathbf{y}|^2 - 2\mathbf{x} \bullet \mathbf{y} + t^2|\mathbf{x}|^2 \qquad \text{for } t \in \mathbf{R}$$

and note that S has a unique minimum point when $t = \frac{\mathbf{x} \bullet \mathbf{y}}{|\mathbf{x}|^2}$.

The angle between two vectors. Given nonzero n-vectors $\mathbf{x}, \mathbf{y}$ we keep in mind the Cauchy-Schwarz Inequality and define the **angle** between them to be

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos\Big(\frac{\mathbf{x} \bullet \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\Big) \in [0, \pi].$$

Evidently,

$$\mathbf{x} \bullet \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos\angle(\mathbf{x}, \mathbf{y}).$$

This terminology will be justified if we can show that it is consistent with Euclidean geometry. We do this by considering the parallelogram

$$P = \{u\mathbf{x} + v\mathbf{y} : u, v \in [0, 1]\}$$

whose vertices are $\mathbf{0}, \mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y}$. The area of this parallelogram is

$$|\mathbf{x}||\mathbf{y} - \mathbf{proj}_{\mathbf{x}}(\mathbf{y})| = \sqrt{|\mathbf{x}|^2|\mathbf{y}|^2 - (\mathbf{x} \bullet \mathbf{y})^2};$$

we leave it as **Exercise 3** for the reader to show that the square of both sides is the same. But the right hand side of this equation is

$$|\mathbf{x}||\mathbf{y}| \sin \angle(\mathbf{x}, \mathbf{y})$$

which is as it should be. In particular, we find that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular in the sense of Euclidean geometry if and only if there dot product is zero which is the case if and only if the angle between them is $\frac{\pi}{2}$.

Volume. Given n n-vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ we set

$$[\mathbf{x}_1, \ldots, \mathbf{x}_n] = \mathbf{det}\, X$$

where X is the $n \times n$ matrix in whose i-th row and j-th column is the i-th component of $\mathbf{x}_j$. We call this scalar the **volume** of the sequence of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$. The justification of this terminology will be furnished shortly.

Some special features of $\mathbf{R}^2$. We set

$$\mathbf{i} = \mathbf{e}_1 \qquad \text{and} \qquad \mathbf{j} = \mathbf{e}_2.$$

Suppose $\mathbf{x} = (x, y) \in \mathbf{R}^2$. We set

$$\mathbf{x}^{\perp} = (-y, x).$$

Note that

$$\mathbf{i}^{\perp} = \mathbf{j} \qquad \text{and} \qquad \mathbf{j}^{\perp} = -\mathbf{i}.$$

Draw some pictures to verify that $\mathbf{x}^{\perp}$ is obtained by rotating $\mathbf{x}$ counterclockwise through an angle of $\frac{\pi}{2}$.

Proposition. Suppose $\mathbf{x}, \mathbf{y} \in \mathbf{R}^2$ and c is a scalar. Then

(1) $(\mathbf{x}+\mathbf{y})^{\perp} = \mathbf{x}^{\perp} + \mathbf{y}^{\perp}$;

(2) $(c\mathbf{x})^{\perp} = c(\mathbf{x}^{\perp})$;

(3) $\mathbf{x}^{\perp} \bullet \mathbf{x} = 0$;

(4) $\mathbf{x}^{\perp\perp} = -\mathbf{x}$;

(5) $\mathbf{x}^{\perp} \bullet \mathbf{y}^{\perp} = \mathbf{x} \bullet \mathbf{y}$;

(6) $|\mathbf{x}^{\perp}| = |\mathbf{x}|$;

(7) $[\mathbf{x}, \mathbf{y}] = \mathbf{x} \bullet \mathbf{y}^{\perp}$.

Proof. Exercise 4 for the reader. □

Some special features of $\mathbf{R}^3$. We set

$$\mathbf{i} = \mathbf{e}_1, \qquad \mathbf{j} = \mathbf{e}_2, \qquad \text{and} \qquad \mathbf{k} = \mathbf{e}_3.$$

Proposition. Suppose $\mathbf{x}, \mathbf{y} \in \mathbf{R}^3$. Then there is one and only one vector

$$\mathbf{x} \times \mathbf{y},$$

called the **cross product of x and y**, with the property that

$$(\mathbf{x} \times \mathbf{y}) \bullet \mathbf{z} = [\mathbf{x}, \mathbf{y}, \mathbf{z}] \qquad \text{for any } \mathbf{z} \in \mathbf{R}^3.$$

Moreover, if $\mathbf{x} = (x^1, x^2, x^3)$ and $\mathbf{y} = (y^1, y^2, y^3)$ then

$$\mathbf{x} \times \mathbf{y} = (x^2 y^3 - x^3 y^2, x^3 y^1 - x^1 y^3, x^1 y^2 - x^2 y^1).$$

Proof. Supposing the cross product of $\mathbf{x}$ and $\mathbf{y}$ exists as in the defining property, we set $\mathbf{z}$ equal to $\mathbf{i}$,$\mathbf{j}$ and $\mathbf{k}$, respectively, to deduce the given formula for the cross product and its uniqueness. Stuffing the formula into the defining property, we see that the defining property holds. □

Check out the book for all the properties of the cross product. Perhaps its most important property is that $\mathbf{x} \times \mathbf{y}$ is perpendicular to both $\mathbf{x}$ and $\mathbf{y}$ and that it is $\mathbf{0}$ if and only if there is a scalar c such that either $\mathbf{x} = c\mathbf{y}$ or $\mathbf{y} = c\mathbf{x}$. We leave for the reader as **Exercise 5** to verify that

$$|\mathbf{x} \times \mathbf{y}| = \sqrt{|\mathbf{x}|^2 |\mathbf{y}|^2 - (\mathbf{x} \bullet \mathbf{y})^2};$$

if neither $\mathbf{x}$ nor $\mathbf{y}$ is zero, this is

$$|\mathbf{x}||\mathbf{y} \sin \angle(\mathbf{x}, \mathbf{y})$$

which is the area of the parallelogram

$$P = \{u\mathbf{x} + v\mathbf{y} : (u, v) \in [0, 1] \times [0, 1]\}.$$

Using all of the above we find that

$$(\mathbf{x} \times \mathbf{y}) \bullet \mathbf{z} = [\mathbf{x}, \mathbf{y}, \mathbf{z}]$$

is the volume of the solid

$$S = \{u\mathbf{x} + v\mathbf{y} + w\mathbf{z} : (u, v, w) \in [0, 1] \times [0, 1] \times [0, 1]\}.$$

Operations on vector valued functions. Suppose $\mathbf{f}$ is an n-vector valued function. For each $i = 1, \ldots, n$ its i-th component is the function $\mathbf{e}^i \circ \mathbf{f}$.

Suppose c is a scalar and $\mathbf{f}$ is an n-vector valued function. Then

$$c\mathbf{f}$$

is the n-vector valued function whose domain is the domain of $\mathbf{f}$ and whose value at $\mathbf{x}$ in the domain of $\mathbf{f}$ is $c\mathbf{f}(\mathbf{x})$.

Suppose f is a scalar valued function and $\mathbf{v}$ is an n-vector. Then

$$f\mathbf{v}$$

is the n-vector valued function whose domain is the domain of f and whose value at $\mathbf{x}$ in the domain of f is $f(\mathbf{x})\mathbf{v}$.

Suppose $\mathbf{f}$ and $\mathbf{g}$ are n-vector valued functions.

$$\mathbf{f} + \mathbf{g}$$

is the n-vector valued function whose domain is the intersection of the domains of $\mathbf{f}$ and $\mathbf{g}$ and which whose value at $\mathbf{x}$ in this intersection is $\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})$.

Suppose f is a scalar valued function and $\mathbf{g}$ is an n-vector valued function. Then

$$f\mathbf{g}$$

is the n-vector valued function whose domain is the intersection of the domains of f and $\mathbf{g}$ and which whose value at $\mathbf{x}$ in this intersection is $f(\mathbf{x})\mathbf{g}(\mathbf{x})$.

Covectors. We say α is an n-**covector** if

(c1); $$\alpha : \mathbf{R}^n \to \mathbf{R}$$

(c2) $\alpha(\mathbf{x} + \mathbf{y}) = \alpha(\mathbf{x}) + \alpha(\mathbf{y})$ whenever $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$;

(c3) $\alpha(c\mathbf{x}) = c\alpha(\mathbf{x})$ whenever $c \in \mathbf{R}$ and $\mathbf{x} \in \mathbf{R}^n$.

We let

$$\mathbf{R}_n$$

be the set of n-covectors. It is a simple matter to verify that a scalar multiple of a covector is a covector and that the sum of two covectors is a covector. Note that

$$\mathbf{e}^1, \ldots, \mathbf{e}^n$$

are n-covectors.

Proposition. Suppose $\alpha : \mathbf{R}^n \to \mathbf{R}$. Then α is a covector if and only if there is a vector $\mathbf{a}$ such that

$$\alpha(\mathbf{x}) = \mathbf{x} \bullet \mathbf{a} \qquad \text{whenver } \mathbf{x} \in \mathbf{R}^n.$$

Proof. The sufficiency follows from properties (d1) and (d2) of the dot product.

Suppose α is a covector. Let $\mathbf{a}$ be the vector whose i-th component equals $\alpha(\mathbf{e}_i)$, $i = 1, \ldots, n$. Then for any vector $\mathbf{x} = (x^1, \ldots, x^n)$ we have

$$\alpha(\mathbf{x}) = \alpha(\sum_{i=1}^n x^j \mathbf{e}_j) = \sum_{i=1}^n x^j \alpha(\mathbf{e}_j) = \mathbf{x} \bullet \mathbf{a}.$$

$\square$

Linear functions. Suppose

($l1$) $$\mathbf{l} : \mathbf{R}^n \to \mathbf{R}^m.$$

We say $\mathbf{l}$ is **linear** if

(l1) $\mathbf{l}(c\mathbf{x}) = c\mathbf{l}(\mathbf{x})$ whenever $c \in \mathbf{R}$ and $\mathbf{x} \in \mathbf{R}^n$ and

(l2) $\mathbf{l}(\mathbf{x} + \mathbf{y}) = \mathbf{l}(\mathbf{x} + \mathbf{l}(\mathbf{y})$ whenever $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$.

We let

$$\mathbf{R}_n^m$$

be the set of linear functions from $\mathbf{R}^n$ to $\mathbf{R}^m$. It is a simple matter to verify that the sum of two members of $\mathbf{R}^m_n$ is in $\mathbf{R}^m_n$ and that a scalar multiple of a member of $\mathbf{R}^m_n$ is in $\mathbf{R}^m_n$. Proceed as in **Exercise 6**. We let

$$\mathbf{R}^m_n$$

be the set of real valued linear functions from $\mathbf{R}^n$ to $\mathbf{R}$. Note that $\mathbf{R}^1_n$ is the set of n-covectors.

Proposition. Suppose $l \in \mathbf{R}_n$ and $\mathbf{v} \in \mathbf{R}^m$. Then

$$l\mathbf{v}$$

is linear.

Proof. Exercise 7. $\square$

Proposition. $\mathbf{e}^j$ is linear for each $j = 1, \ldots, n$.

Proof. This is a direct consequence of the definitions. $\square$

Proposition. Suppose $\mathbf{l} \in \mathbf{R}^m_n$ and $\mathbf{k} \in \mathbf{R}^m_l$. Then $\mathbf{l} \circ \mathbf{k} \in \mathbf{R}^m_l$.

Proof. Exercise 7. $\square$

Proposition. Suppose

$$\mathbf{l} : \mathbf{R}^n \to \mathbf{R}^m.$$

The $\mathbf{l}$ is linear if and only if each its components is linear.

Proof. Suppose $\mathbf{l}$ is linear. Then, by the preceding Proposition, $l^j = \mathbf{e}^j \circ \mathbf{l}$ is linear for each $j = 1, \ldots, n$.

Suppose each component of $\mathbf{l}$ is linear. Then for any $\mathbf{x} \in \mathbf{R}^n$ we have

$$\mathbf{l}(\mathbf{x}) = \sum_i^n l^i(\mathbf{x})\mathbf{e}_i = (\sum_i^n l^i \mathbf{e}_i)(\mathbf{x})$$

which is to say that

$$\mathbf{l} = \sum_i^n l^i \mathbf{e}_i.$$

But we have already noted that scalar multiples and sums of linear functions are linear. $\square$

Proposition. Suppose $\mathbf{l} \in \mathbf{R}^m_n$. Let

$$a^i_j = \mathbf{e}^i(\mathbf{l}(\mathbf{e}_j)) \qquad i = 1 \ldots, m,\ j = 1, \ldots, n.$$

Then

$$\mathbf{l}(\mathbf{x}) = \sum_{i=1}^m (\sum_{j=1}^n a^i_j x^j)\mathbf{e}_i \qquad \text{for } \mathbf{x} \in \mathbf{R}^n.$$

Proof. Suppose $\mathbf{x} \in \mathbf{R}^n$. Then

$$\begin{aligned}
\mathbf{l}(\mathbf{x}) &= \mathbf{l}(\sum_{j=1}^{n} x^j \mathbf{e}_j) \\
&= \sum_{j=1}^{n} \mathbf{l}(x^j \mathbf{e}_j) \\
&= \sum_{j=1}^{n} x^j \mathbf{l}(\mathbf{e}_j) \\
&= \sum_{j=1}^{n} x^j (\sum_{i=1}^{m} \mathbf{e}^i(\mathbf{l}(\mathbf{e}_j))\mathbf{e}_i) \\
&= \sum_{j=1}^{n} x^j (\sum_{i=1}^{m} a_i^j \mathbf{e}_i) \\
&= \sum_{i=1}^{m} (\sum_{j=1}^{n} a_i^j x^j)\mathbf{e}_i).
\end{aligned}$$

As an exercise, give reasons justifying each of these steps. □

Discussion. Thus a linear function from $\mathbf{R}^n$ to $\mathbf{R}^m$ is determined by its **matrix** which is, by definition, the rectangular array with m rows and n columns which has the scalar

$$a_j^i = \mathbf{e}^i(\mathbf{l}(\mathbf{e}_j))$$

in its i-th row and j-th column $i = 1 \ldots, m,\ j = 1, \ldots, n$. This array is usually depicted as follows:

$$\begin{bmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \ldots & a_n^m \end{bmatrix}.$$

Suppose $\mathbf{l} \in \mathbf{R}_n^m$ and A is its matrix. The if $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{y} = \mathbf{l}(\mathbf{x}) \in \mathbf{R}^m$ then the preceding Proposition says that

$$y^i = \sum_{j=1}^{n} a_j^i x^j, \qquad i = 1, \ldots, m$$

which amounts to

$$\begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{bmatrix} = A \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}.$$

The interested reader can verify that all the definitions and properties of matrix operations follow from corresponding but usually more simply established definitions and properties of linear functions.