An easy on Linear Transformations

Research by: Jose Portillo

$$\Big\langle \langle K, +_g, \cdot \rangle,\ \langle V, +_v \rangle,\ \circ \Big\rangle$$

The system above is called a *vector space* over the field K [2]. Elements of K are called *scalars*, and elements of V are called *vectors*. The operation $\circ : K \times V \to V$ is scalar multiplication, and $+_v$ is vector addition. In what follows, symbols $\alpha_1, \ldots, \alpha_k$ always denote scalars in K, and symbols $x_1, \ldots, x_k$ always denote vectors in V.

Example (Coordinate n-space, [3]).
Let Γ be a field. Consider the set

$$\Gamma^n = \Gamma \times \cdots \times \Gamma$$

of n-tuples $(\xi^1, \ldots, \xi^n)$, with $\xi^i \in \Gamma$. Define addition by

$$(\xi^1, \ldots, \xi^n) + (\eta^1, \ldots, \eta^n) = (\xi^1 + \eta^1, \ldots, \xi^n + \eta^n)$$

and scalar multiplication by

$$\lambda(\xi^1, \ldots, \xi^n) = (\lambda\xi^1, \ldots, \lambda\xi^n), \quad \lambda \in \Gamma.$$

With these operations, Γ^n is a vector space over Γ, called the *n-space over* Γ. In particular, Γ itself is a vector space over Γ, where scalar multiplication coincides with field multiplication.
This construction provides the standard coordinate model for finite-dimensional vector spaces.

A vector space like this is defined over a field K, where $\langle K, +_g, \cdot \rangle$ satisfies the usual *field axioms*: for all $\alpha, \beta, \gamma \in K$, addition $+_g$ is associative and commutative, has an identity $0 \in K$ and inverses $-\alpha$, multiplication $\cdot$ is associative and commutative, has an identity $1 \in K$, every nonzero element has a multiplicative inverse, and multiplication distributes over addition. Vector addition $+_v$ on V is associative and commutative, has an identity vector 0_V, and every $x \in V$ has an additive inverse $-x$. Scalar multiplication $\circ$ satisfies: $\alpha \circ (x +_v y) = \alpha \circ x +_v \alpha \circ y$, $(\alpha +_g \beta) \circ x = \alpha \circ x +_v \beta \circ x$, $(\alpha \cdot \beta) \circ x = \alpha \circ (\beta \circ x)$, and $1 \circ x = x$ for all $\alpha, \beta \in K$ and $x, y \in V$.

A finite set of vectors $\{b_1, \ldots, b_n\} \subset V$ is called a *basis* of the vector space V if every vector $x \in V$ can be written uniquely as a linear combination $x = \alpha_1 \circ_E b_1 + \alpha_2 \circ_E b_2 + \cdots + \alpha_n \circ_E b_n, \quad \alpha_1, \ldots, \alpha_n \in K$. A basis therefore provides a minimal and non-redundant set of building blocks for the entire space. Once a basis is fixed, every vector is completely determined by its coefficients relative to that basis.

Given vectors $x_1, \ldots, x_k \in V$ and scalars $\alpha_1, \ldots, \alpha_k \in K$, an expression of the form

$$x_t = \alpha_1 \circ_E x_1 + \alpha_2 \circ_E x_2 + \cdots + \alpha_k \circ_E x_k$$

is called a *linear combination* of the vectors $x_1, \ldots, x_k$.

By closure of vector addition and scalar multiplication, every linear combination is itself a vector in V. Linear combinations therefore describe the permitted way to construct new vectors from given ones inside the same vector space.

Let $\Big\langle \langle K, +_g, \cdot \rangle,\ \langle E, +_E \rangle,\ \circ_E \Big\rangle$ and $\Big\langle \langle K, +_g, \cdot \rangle,\ \langle G, +_G \rangle,\ \circ_G \Big\rangle$ be vector spaces over the same field K. A function $\psi : E \to G$ is called a *linear transformation* if it preserves linear combinations. That is, for all vectors $x_1, \ldots, x_k \in E$ and all scalars $\alpha_1, \ldots, \alpha_k \in K$, [1]

$$\psi(\alpha_1 \circ_E x_1 + \alpha_2 \circ_E x_2 + \cdots + \alpha_k \circ_E x_k) = \alpha_1 \circ_G \psi(x_1) + \alpha_2 \circ_G \psi(x_2) + \cdots + \alpha_k \circ_G \psi(x_k).$$

This equation expresses the defining property of linearity: vectors may be combined first and then transformed, or transformed first and then combined, with identical results. Linear transformations therefore preserve the algebraic structure of vector spaces independently of any choice of coordinates.

Let $\psi : E \to G$ be a linear transformation, and let $B_E = \{e_1, \ldots, e_n\} \subset E, B_G = \{g_1, \ldots, g_m\} \subset G$ be bases of the domain and codomain, respectively. For each basis vector e_i, the image $\psi(e_i)$ can be written uniquely as a linear combination of the basis vectors of G: $\psi(e_i) = a_{1i} \circ_G g_1 + a_{2i} \circ_G g_2 + \cdots + a_{mi} \circ_G g_m$. The scalars $a_{ji} \in K$ assemble into an $m \times n$ matrix

$$A = (a_{ji}),$$

called the *matrix of ψ relative to the bases B_E and B_G*. The matrix A represents the linear transformation ψ in coordinates. Different choices of bases generally lead to different matrices, but they all represent the same underlying linear transformation.

Matrices thus do not define linear transformations; rather, they are concrete representations of linear transformations once bases (coordinate systems) have been chosen.

This abstract formulation is independent of geometry, yet it underlies most quantitative models in science and engineering. In modern machine learning, for example, embeddings and weight matrices are linear transformations on high-dimensional vector spaces, composed with non-linear activations. Even where non-linearity dominates, linear transformations provide the structural backbone—the universality of a man-made linear structure, even though the universe itself is not linear.

Lets complement the definition of *linear transformation* with that of *dual space*. Starting with with our definition of vector space over the field K $\big\langle \langle K, +_g, \cdot \rangle,\ \langle E, +_E \rangle,\ \circ_E \big\rangle$

the set of all linear transformations from E into the base field K is denoted by $E^* = L(E, K)$ and is called the *dual space* of E. An element $x^* \in E^*$ is called a *linear functional*. Thus, a linear functional is a function $x^* : E \to K$ satisfying, for all $x_1, \ldots, x_k \in E$ and all $\alpha_1, \ldots, \alpha_k \in K$,

$$x^*(\alpha_1 \circ_E x_1 + \cdots + \alpha_k \circ_E x_k) = \alpha_1 x^*(x_1) + \cdots + \alpha_k x^*(x_k).$$

Linear functionals therefore preserve linear combinations, but unlike linear transformations between vector spaces, their values are scalars. They may be interpreted as algebraic measurements performed on vectors.

Assume now that E is finite dimensional, and let $B_E = \{e_1, \ldots, e_n\}$ be a basis of E. Any linear functional $x^* \in E^*$ is uniquely determined by its values on the basis vectors. Indeed, if

$$x = \alpha_1 \circ_E e_1 + \cdots + \alpha_n \circ_E e_n,$$

then linearity implies

$$x^*(x) = \alpha_1 x^*(e_1) + \cdots + \alpha_n x^*(e_n).$$

The scalars $x^*(e_1), \ldots, x^*(e_n)$ may therefore be assembled into a row vector

$$\begin{bmatrix} x^*(e_1) & \ldots & x^*(e_n) \end{bmatrix},$$

which represents the functional x^* relative to the basis B_E.

In particular, when $E = K^2$ with its canonical basis, the functional

$$x^*(x, y) = 3x + 2y$$

is represented by the row vector

$$[\,3 \;\; 2\,],$$

and its action on a vector $\begin{bmatrix} x \\ y \end{bmatrix}$ is given by matrix multiplication:

$$[\,3 \;\; 2\,] \begin{bmatrix} x \\ y \end{bmatrix} = 3x + 2y.$$

This representation shows that row vectors naturally encode linear functionals, whereas column vectors represent elements of the vector space itself. These are objects of different nature, even though they are closely related.

Let now

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}$$

be the matrix of a linear transformation $\psi : K^2 \to K^2$ relative to the canonical basis. The product

$$[\,3 \;\; 2\,]\, A = [\,19 \;\; 17\,]$$

represents the composition of linear maps

$$K^2 \xrightarrow{\psi} K^2 \xrightarrow{x^*} K,$$

which is again a linear functional on K^2. The result is therefore a row vector, representing a new element of the dual space.

This interpretation explains the rule governing matrix multiplication: products are defined precisely when the codomain of one linear map coincides with the domain of the next. Matrix dimensions encode this compatibility.

Finally, if E is equipped with an inner product, each vector $x \in E$ determines a linear functional via

$$x^*(y) = \langle x, y \rangle.$$

In coordinates, this correspondence is expressed using the transpose operation, which converts column vectors into row vectors. The transpose therefore provides a bridge between vectors and linear functionals, but it depends on additional structure and is not intrinsic to the vector space itself.

The dual space thus completes the algebraic picture: alongside linear transformations between vector spaces, linear functionals describe how vectors may be evaluated and compared through scalar quantities.

Note. An AI language model was used as a support tool for discussion, clarification of concepts, and stylistic refinement. All mathematical choices and final wording are the responsibility of the author.

References

[1] Iribarren Ignacio L., *Algebra Lineal*, Editorial Equinoccio, 2018.
[2] Cohen, Leon and Ehrlich, Gertrude, *The Structure Of The Real Number System*, D. Van Nostrand Company, 1963.
[3] Greub, Werner, *Linear Algebra*, Springer-Verlag, 1975.