Lecture (supplemental): Matrices and vectors

August 22, 2022

Summary: Summary of $\mathbb{R}^n$ vectors and matrices as linear transformations. Much of this material has been taken from "Linear algebra in a hurry for students of ODE" by Michael E. Taylor.

$\mathbb{R}^n$ Vectors

Basic operations. *The set $\mathbb{R}^n$ is a vector space over $\mathbb{R}$.* We'll work with vectors rather informally; however, just this once let's break down the meaning of our statement into the following three pieces. First, by $\mathbb{R}^n$ we mean the n–space comprised of all n–tuples,

$$\mathbf{v} = (v_1, v_2, \cdots, v_n), \tag{1}$$

where each *component* v_k of $\mathbf{v}$ is a real number itself, that is $\forall k, v_k \in \mathbb{R}$. Here, we've introduced the elements of $\mathbb{R}^n$ as row vectors (laid out horizontally), but could just as well have introduced them as column vectors (laid out vertically),

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{2}$$

Second, we view $\mathbb{R}^n$ as equipped with two operations: (i) vector addition and (ii) multiplication by scalars. Given two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, vector addition gives us a new element $\mathbf{v} + \mathbf{w} \in \mathbb{R}^n$, where

$$\mathbf{v} = (v_1, v_2, \cdots, v_n), \quad \mathbf{w} = (w_1, w_2, \cdots, w_n), \quad \mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2, \cdots, v_n + w_n). \tag{3}$$

Given $\mathbf{v} \in \mathbb{R}^n, a \in \mathbb{R}$ (the real number a is the "scalar") the operation of scalar multiplication determines a new vector $a\mathbf{v} \in \mathbb{R}^n$, where

$$\mathbf{v} = (v_1, v_2, \cdots, v_n), \quad a\mathbf{v} = (av_1, av_2, \cdots, av_n). \tag{4}$$

The "over $\mathbb{R}$" qualification in our opening statement refers to the operation of scalar multiplication and the restriction $a \in \mathbb{R}$. Third, with $a, b \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, the operations just defined satisfy the following properties.

(Vector addition laws)	Commutative law:	$\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{v}$,
	Associative law:	$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$,
	Zero vector:	$\exists \mathbf{0} \in \mathbb{R}^n, \mathbf{v} + \mathbf{0} = \mathbf{v}$,
	Negative:	$\exists -\mathbf{v} \in \mathbb{R}^n, \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$,
(Multiplication–by–scalar laws)	Associative law:	$a(b\mathbf{v}) = (ab)\mathbf{v}$,
	Unit:	$1 \cdot \mathbf{v} = \mathbf{v}$,
(Distributive laws)		$a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$
		$(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.

For our case, that of $\mathbb{R}^n$ over $\mathbb{R}$, it hardly seems worth it to mention these properties. Indeed, as a concrete example of commutative law, consider the following uninspiring identity:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}. \tag{5}$$

Nevertheless, the mathematical world is replete with vector spaces, ones for which the notion of vector may be rather different, and it's imperative to have a precise definition. All vector spaces satisfy essentially the same properties listed above. For example, we could define the complex vector space $\mathbb{C}^n$ over $\mathbb{C}$ by just replacing all $\mathbb{R}^n$'s and $\mathbb{R}$'s above with $\mathbb{C}^n$'s and $\mathbb{C}$'s. As another example, consider the set $\mathcal{P}_k$ of all polynomials (in a variable x and with real coefficients) of degree $\leq k$, which is a vector space over $\mathbb{R}$. If we replace $\mathbb{R}^n$ by $\mathcal{P}_k$ above, and trade $\mathbf{u}, \mathbf{v}, \mathbf{w}$ for polynomials $p(x), q(x), r(x)$, then the above conditions are satisfied. $\mathbb{R}^n$ and $\mathcal{P}_k$ are examples of finite–dimensional vector spaces, but there are also infinite–dimensional examples. To define the dimensionality of a vector space, we need the concept of a basis which we get to below.

Inner product and norm. As a vector space, $\mathbb{R}^n$ is special in that, beyond the standard properties above which make it a vector space, we may equip it will an auxiliary notion of an inner product (here the same as the dot product). This is the paring defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = (v_1, v_2, \cdots, v_n) \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \sum_{k=1}^{n} v_k u_k. \tag{6}$$

Notice that $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$. Furthermore, the paring is described as linear since $\forall a, b \in \mathbb{R}$

$$\langle \mathbf{u}, a\mathbf{v} + b\mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle + b\langle \mathbf{u}, \mathbf{w} \rangle. \tag{7}$$

Using the inner product we may define the *norm* (that is the length) of a vector via

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}. \tag{8}$$

Not all vector spaces have the notion of inner product and/or norm (but all inner product spaces are norm spaces). They are useful in that they allow for projection of one vector along the direction of another vector. For example,

$$\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|^2}\mathbf{u} = \langle \hat{\mathbf{u}}, \mathbf{v} \rangle \hat{\mathbf{u}} \tag{9}$$

is the projection of $\mathbf{v}$ along the unit direction $\hat{\mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$.

Linear subspaces. A subset U of a vector space V (say over $\mathbb{R}$) is a *linear subspace* provided

$$\mathbf{u}, \mathbf{w} \in U \text{ and } a, b \in \mathbb{R} \implies a\mathbf{u} + b\mathbf{w} \in U. \tag{10}$$

Notice that vectors in U are also vectors in V, and as a result U inherits the structure of a vector space. As an example consider $\mathbb{R}^3$, and the subset U of all vectors with the form

$$\mathbf{u} = (u_1, u_2, u_3) = (u_1, u_2, -u_1). \tag{11}$$

Then U is a linear subspace, and so a vector space which we may identify with $\mathbb{R}^2$, geometrically a plane through the origin. A line through the origin in $\mathbb{R}^3$ is also a subspace.

Linear transformations and matrices

Abstract definition and matrix example. Suppose V and W are vectors spaces (again, say over $\mathbb{R}$). The map

$$T : V \to W, \tag{12}$$

where $\mathbf{v} \in V$ and $\mathbf{w} = T\mathbf{v} \in W$, is a linear transformation provided,

$$T(a\mathbf{v} + b\mathbf{u}) = aT\mathbf{v} + bT\mathbf{u}, \qquad \forall a, b \in \mathbb{R} \text{ and } \forall \mathbf{v}, \mathbf{u} \in V. \tag{13}$$

You may think of T as a function $\mathbf{w} = \mathbf{f}(\mathbf{v}) = T\mathbf{v}$ taking a vector in V and returning one in W. Were we considering a linear scalar function $y = f(x) = cx$, the notation of linear transformations is such that we would just write $c : \mathbb{R} \to \mathbb{R}$, doing away with f altogether. Again, examples are manifold, such as

$$D : \mathcal{P}_{k+1} \to \mathcal{P}_k, \tag{14}$$

where D represents differentiation, that is $Dq(x) = q'(x)$.

However, we'll be interested in the important special case of an m–by–n matrix A, which defines the linear transformation

$$A : \mathbb{R}^n \to \mathbb{R}^m \tag{15}$$

via

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k} v_k \\ \vdots \\ \sum_{k=1}^n a_{mk} v_k \end{pmatrix}. \tag{16}$$

Whereas we speak of *components* v_k of a vector $\mathbf{v}$, we typically refer to the *entries* a_{jk} of the matrix A. But notice that $\mathbf{v}$ is itself a n–by–1 matrix, so this isn't an absolute terminology. Referring to a_{jk}, the row index is j and the column index is k. Multiplication of two matrices corresponds to composition of the associated linear maps. Although this is the case of interest, let's continue to develop some concepts which hold for more general linear transformations.

Kernel and range of a linear operator. The null space or kernel of a linear operator T is the set $\mathcal{N}(T) = \{\mathbf{v} \in V : T\mathbf{v} = \mathbf{0}\}$, where the zero vector here is $\mathbf{0} \in W$. The range of T is the set $\mathcal{R}(T) = \{\mathbf{w} \in W : w = T\mathbf{v}, \mathbf{v} \in V\} = \{T\mathbf{v} : \mathbf{v} \in V\}$. We note that $\mathcal{N}(T)$ is a linear subspace of V, whereas $\mathcal{R}(T)$ is a linear subspace of W. If $\mathcal{N}(T) = \{\mathbf{0}\}$ (that is just contains the zero vector), then T is one–to–one (injective), since due to the linearity of T,

$$T\mathbf{v} = T\mathbf{u} \implies \mathbf{v} = \mathbf{u}. \tag{17}$$

Were the last statement not to hold, then clearly $\mathbf{v} - \mathbf{u} \neq \mathbf{0}$ would be an element of $\mathcal{N}(T)$. If $\mathcal{R}(T) = W$, then T is onto (surjective), that is to say, any $\mathbf{w} \in W$ can be obtained by the action of T on some $\mathbf{v} \in V$. A one–to–one and onto linear transformation is called an *isomorphism*, and for such a transformation there is a well–defined inverse

$$T^{-1} : W \to V, \tag{18}$$

where T^{-1} is a linear transformation in its own right satisfying. $\mathbf{w} = T\mathbf{v} \iff \mathbf{v} = T^{-1}\mathbf{w}$. For case of matrices, isomorphisms will correspond to nonsingular (that is invertible) n–by–n (square) matrices. We'll reconsider this issue in the context of a linear system of n equations in n unknowns.

Consider the linear transformation $A : \mathbb{R}^4 \to \mathbb{R}^3$ defined by the 3–by–4 matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{19}$$

$\mathcal{N}(A)$ is the set of vectors with form

$$\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_4 \end{pmatrix}, \tag{20}$$

and $\mathcal{R}(A)$ the set with form

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}. \tag{21}$$

So the transformation defined by A is onto but not one–to–one (surjective but not injective).

Basis and dimension

Given a finite set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ in $\mathbb{R}^n$, the *span* of S is the set of all vectors in $\mathbb{R}^n$ expressible as

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k, \tag{22}$$

where $a_1, \cdots, a_k \in \mathbb{R}$ are arbitrary scalars (expansion coefficients). The set S is said to be *linearly dependent* if there exists scalars $a_1, \cdots, a_k \in \mathbb{R}$, not all zero, such that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}. \tag{23}$$

Otherwise the set S is said to be linearly independent, and we say S is a basis for $\text{Span}(S)$. If S is linearly dependent, there is a smaller set $S_0 \subset S$ with the same span as S.

The following canonical set of vectors is a basis for $\mathbb{R}^n$:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{24}$$

Here the component e_{jk} of $\mathbf{e}_k$ is 0 if $j \neq k$ and 1 if $j = k$. The collection above is a basis since it's linearly and any vector can be expanded as

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots v_n\mathbf{e}_n. \tag{25}$$

In fact, the above basis is orthonormal. By counting the number of basis elements we arrive at a definition of dimension, either for $\mathbb{R}^n$ itself or a linear subspace. Some useful facts are the following.

Fact 1: If $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{n+1}$ are vectors in $\mathbb{R}^n$, then they are linearly dependent.

Fact 2: If $U \in \mathbb{R}^n$ is a linear subspace, then U has a finite basis and $\text{dim}U \leq n$.

Fact 3: If $U \in \mathbb{R}^n$ is a linear subspace, with basis $\{u_1, u_2, \cdots, u_\ell\}$, then it's possible to pick as a basis of $\mathbb{R}^n$ the set $\{u_1, u_2, \cdots, u_\ell, q_1, q_2, \cdots, q_m\}$, where $\ell + m = n$.

Theorem 1 (Fundamental Theorem of Linear Algebra)**.** *Assume V and W are vectors spaces, with V finite dimensional, and $A : V \to W$ a linear map. Then*

$$\text{dim}\mathcal{N}(A) + \text{dim}\mathcal{R}(A) = \text{dim}V. \tag{26}$$

We don't give the proof, and concentrate on the case when A is an n–by–n matrix, that is when $A : \mathbb{R}^n \to \mathbb{R}^n$. Then the above result tells us that

$$A \text{ injective} \iff A \text{ surjective} \iff A \text{ isomorphism} \tag{27}$$

This will have ramifications for solving linear systems next time.