Vector Spaces and Linear Maps *by Dr. Dmitry Gokhman* 1996 .... *Vectors are carriers of disease.*

Vectors are things that can be added and scaled. [1]
Scalars come from a field, typically the field of real numbers $\mathbf{R}$ or the field of complex numbers $\mathbf{C}$.

Definition: A *vector space* [2] (a *linear space*) V is a set with addition $(+)$ and multiplication by scalars $(\cdot)$ from a field F.

The operations satisfy the following axioms: $(\forall u, v, w \in V; r, s \in F)$

(i) $u+(v+w)=(u+v)+w$ (iii) $r\cdot(u+v)=r\cdot u+r\cdot v$ (v) $r\cdot(s\cdot u)=rs\cdot u$

(ii) $u+v=v+u$ (iv) $(r+s)\cdot u=r\cdot u+s\cdot u$ (vi) $1\cdot u=u$, $0\cdot u=0$

The combined operations give *linear combinations*: $\sum_{k=1}^{m} r_k \cdot u_k = r_1\cdot u_1+...+r_m\cdot u_m$, where $u_k \in V, r_k \in F$.

Example 1: Real Euclidean space $\mathbf{E}^n$ (e.g. the plane $\mathbf{E}^2$) can be made into a real vector space by a choice of origin. In fact, we can pick any point p of $\mathbf{E}^n$ and obtain a vector space with p as the origin ($p=0$).

Given two points u and v in $\mathbf{E}^n$, draw segments connecting them to the origin p. Complete the parallelogram. The sum $u+v$ is defined to be the remaining vertex of the parallelogram. To scale a point u by a positive real number r draw a ray from the p through u and define $v=r\cdot u$ to be a point on the ray such that the ratio of the distances from p to v and p to u is r. To scale by -1 reflect u about p.

Example 2: Let $\mathbf{R}^n$ be the set of all sequences of real numbers of length n (n-tuples). Placed horizontally these sequences are known as *row vectors*. We shall often place them vertically as *column vectors*. Define addition and scaling coordinate-wise:

$$\text{(i)}\ \begin{pmatrix} a_1 \\ ... \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ ... \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ ... \\ a_n+b_n \end{pmatrix} \qquad \text{(ii)}\ c\begin{pmatrix} a_1 \\ ... \\ a_n \end{pmatrix} = \begin{pmatrix} ca_1 \\ ... \\ ca_n \end{pmatrix}$$

Example 3: Let $C(\mathbf{R})$ be the set of all continuous real-valued functions of $\mathbf{R}$. Define addition and scaling as follows:

(i) $(f_1+f_2)(x)=f_1(x)+f_2(x)$ (ii) $(cf)(x)=cf(x)$

Note that addition and scaling preserve continuity. Can we do the same with differentiable functions?

Exercise: Verify that the sets with operations as in the above examples satisfy the axioms of a vector space over $\mathbf{R}$.

Linear subspaces: A subset of a vector space is a linear subspace, if it is *closed* under the operations. In other words, any linear combination of vectors in the subspace should also be in the subspace.

Exercise: Show that the following are linear subspaces of $\mathbf{E}^2$ with an origin: the origin, the whole space, any line through the origin. Show that these are the only linear subspaces of $\mathbf{E}^2$.

Exercise: Show that the intersection of a collection of linear subspaces of a vector space is also a linear subspace. What happens in the case of union?

Bases: Given a set of vectors we would like consider the vector subspace that they generate (span). This is the smallest subspace containing them. It consists of all possible linear combinations of vector in the generating set.

Now consider the reverse problem: given a vector space V, find a set of generators. We could take all of V, but prefer to have as few generators as possible. If one of the generators is a linear combinations of others, i.e. the set of generators is linearly dependent, then we can throw that generator away and still have a set of generators, but smaller.

A set of vectors is *linearly independent* [3] when each of the vectors is outside the span of the rest of them. In other words, none of the elements of B can be written as a linear combination of the others. Equivalently we can phrase this as follows: the only way a linear combination of elements of B can equal 0 is by being trivial, i.e. by having all the coefficients be 0. [**Proof:** If $b_1=r_2\cdot b_2+...+r_m\cdot b_m$, move b_1 to the other side and obtain a nontrivial linear combination equal to 0. Conversely, given a nontrivial linear combination equal to 0, choose $r_k\neq 0$, divide the equation by r_k and move b_k to the other side of the equation.] This leads to the following definition.

Definition: Suppose V is a vector space and $B\subseteq U$. The set B is called *linearly independent* when for all $b_1,...b_m$ in B

$$\sum_{k=1}^{m} r_k\cdot b_k=0 \quad\Rightarrow\quad (\forall k)\ r_k=0.$$

If B generates V, then any vector in V can be written as a linear combination of elements of B. If B is linearly independent, then any such representation is *unique*. [**Proof:** Given two such representations, subtract them: $(r_1b_1+...+r_nb_n)-(s_1b_1+...+s_nb_n)=(r_1-s_1)\,b_1+...+(r_n-s_n)\,b_n=0$, so $r_1-s_1=...=r_n-s_n=0$.]

If B spans V and is linearly independent, then B is called a *basis* (or a set of coordinate vectors) for V. A basis can be thought of as a minimal set of generators or, equivalently, a maximal linearly independent set.

---

[1] Vectors were first used by Bernhard Placidus Johann Nepomuk Bolzano (1781–1848) in *Betrachtungen über einige Gegenstände der Elementargoemetrie*, 1804. They were popularized by Edwin Bidwell Wilson (1879–1964) in *Vector Analysis*, 1901 after the Yale lectures of Josiah Willard Gibbs (1839–1903) in 1899 .

[2] Linear spaces were introduced by Giuseppe Peano (1858–1932) in *Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva*, 1888.

[3] This idea and the notion of dimension was introduced by Hermann Grassmann (1808–1877) in *Die lineale Ausdehnundslehre, ein neuer Zweig der Mathematik*, 1844.

Example: The vector space $\mathbf{R}^n$ has a canonical basis $\{e_1, ...e_n\}$, where e_i consists of all zeros except a 1 in the i-th place.

Coordinate representation: If a basis B is finite, we may as well include *all* of the elements of B in our linear combinations, because if a given element of B does not appear in a particular combination, it can be added with coefficient 0. Furthermore, we can order $B = \{b_1, ...b_n\}$. In this case we may *represent* a vector u with the n-tuple $(r_1, ...r_n)$ of unique coefficients (known as the *coordinates* of u) in the representation of u as a linear combination of elements of B. In other words, a choice of a basis B with n elements is an identification of the vector space with $\mathbf{R}^n$, where b_i correspond to e_i.

Example: Consider $\mathbf{E}^2$ with an origin. Let $\{\widehat{\imath}, \widehat{\jmath}\}$ be an *orthonormal* set, meaning that the line segments connecting $\widehat{\imath}$ and $\widehat{\jmath}$ to the origin are $\perp$ (orthogonal) and have unit $(= 1)$ length. Then $\{\widehat{\imath}, \widehat{\jmath}\}$ form a basis known as *cartesian coordinates.* [4]

Each vector is a unique linear combination of $\widehat{\imath}$ and $\widehat{\jmath}$, which can be obtained by orthogonal projection to the *axes.* The x and y axes are the linear subspaces spanned by $\widehat{\imath}$ and $\widehat{\jmath}$ respectively. Orthogonal projection of a vector v to an axis, say the y-axis, is a point on the axis $v_y\widehat{\jmath}$. The scalar v_y is the y-coodinate of the vector. Similarly for the other axis. Then $v = v_x\widehat{\imath} + v_y\widehat{\jmath}$.

The Euclidean plane $\mathbf{E}^2$ with a choice of origin and cartesian coordinates is identified with $\mathbf{R}^2$, where the basis $\{\widehat{\imath}, \widehat{\jmath}\}$ corresponds to $\{e_1, e_2\}$. This is the source of two different common notations:

$$\begin{pmatrix} a_x \\ a_y \end{pmatrix} = a_x\widehat{\imath} + a_y\widehat{\jmath}.$$

Exercise: Show that vector operations on $\mathbf{E}^2$ correspond to the vector operations for $\mathbf{R}^2$.

A vector space may have many different bases. However if it has a finite basis, then the number of elements in various bases remains the same and is known as the *dimension* of the space.

Linear Maps: We shall consider functions (a.k.a. *linear maps* or *homomorphisms*) between vector spaces, that preserve the vector space structure, i.e. respect linear combinations. Given a function F between two vector spaces U and V, we require that for all vectors $u_1, ...u_m$ in U and scalars $r_1, ...r_m$ we have

$$F\left(\sum_{k=1}^{m} r_k u_k\right) = \sum_{k=1}^{m} r_k F(u_k).$$

Linear maps from U to V form a vector space of their own, designated $\mathrm{Hom}[U, V]$. Addition and scalar multiplication are pointwise. In other words, given scalars a, b and linear maps f, g we have $(af + bg)(u) = af(u) + bg(u)$.

Matrix representation: [5] To define a linear map F it is enough to specify the values of F at the basis elements of U.

Suppose U and V have finite bases C with m elements and D with n elements respectively. Let c_i be in C and uniquely represent (i.e. expand) $F(c_i)$ with respect to D. Also, given a vector u in U, expand it with respect to C. We get:

$$F(c_j) = \sum_{i=1}^{n} a_{ij} d_i. \qquad u = \sum_{j=1}^{m} r_j c_j. \qquad F(u) = \sum_{j=1}^{m} r_j F(c_j) = \sum_{j=1}^{m} r_j \sum_{i=1}^{n} a_{ij} d_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij} r_j\right) d_i.$$

The numbers a_{ij} have two indices, so it is natural to place them in rectangular grid with n rows and m columns (a *matrix*). We see that they uniquely determine the function F, once the two bases C and D are fixed.

The above formula shows how to calculate the coordinates of $F(u)$ from those of u. This process is known as the "multiplication" of a matrix by a column vector (i.e. $m \times 1$ matrix) of r_j (on the right) which produces an $n \times 1$ matrix (another column vector) of the coordinates of $F(u)$ (in parentheses).

Example 1: Suppose $F\colon \mathbf{R}^3 \to \mathbf{R}^2$ is a linear map determined by

$$F(e_1) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, F(e_2) = \begin{pmatrix} 0 \\ 5 \end{pmatrix}, F(e_3) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

We construct the matrix corresponding to F by stacking the images of the basis vectors e_i as columns. Then to find $F(1, -2, -1)$ we multiply the matrix corresponding to F by the column vector corresponding to the argument to obtain

$$F(1, -2, -1) = \begin{pmatrix} 3 & 0 & -1 \\ 2 & 5 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 + 0 \cdot (-2) + (-1) \cdot (-1) \\ 2 \cdot 1 + 5 \cdot (-2) + 0 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \end{pmatrix}.$$

Example 2: A *linear functional* (a.k.a. a 1-*form* or a *co-vector*) is a linear map from a vector space V to the field F. The set of all 1-forms on V is called the dual space $V^* = \mathrm{Hom}[V, F]$.

The double dual V^{**} is naturally identified with V by evaluation. An element v in V corresponds to the linear functional on V^* which takes a linear functional u on V to $u(v)$.

After a choice of basis $\{v_1, ...v_n\}$ for V a 1-form is represented by a $1 \times n$ matrix, i.e. a row vector. The basis for V^* dual to $\{v_1, ...v_n\}$ is given by 1-forms $\{v_1^*, ...v_n^*\}$ represented by row vectors e_i. In other words $v_i^*(v_j) = \delta_{ij}$. [6]

---

[4] After René Descartes (Renatus Cartesius) (1596 – 1650), whose book *Geometry*, published in 1637, did not mention coordinates. The term *coordinates* was introduced later by Gottfried Wilhelm Leibniz (1646 – 1716) .

[5] The term matrix is due to James Joseph Sylvester (1814–1897), 1848. Matrix algebra was developed by Arthur Cayley (1821–1895), 1855.

[6] δ_{ij} is known as the Kronecker delta and is equal to 0 unless $i = j$ in which case it equals 1. Named after Leopold Kronecker (1823–1891) .

Composition of linear maps, matrix multiplication: Consider that in the above situation we also have a linear map G from V into another vector space W with a basis E of p elements. Suppose G is represented by a $p \times n$ matrix b_{ki}. We would like to see how the composition $G \circ F$ transforms the coordinates of u.

$$G\left(d_i\right) = \sum_{k=1}^{n} b_{ki} e_k. \qquad G\left(F\left(u\right)\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} r_j \right) G\left(d_i\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} r_j \right) \sum_{k=1}^{p} b_{ki} e_k = \sum_{k=1}^{p} \left[\sum_{j=1}^{m} \left(\sum_{i=1}^{n} b_{ki} a_{ij} \right) r_j \right] e_k.$$

The coordinates of $G\left(F(u)\right)$ (in brackets) are of matrix-times-column-vector form again, where the $p \times m$ matrix (in parenthesis) representing $G \circ F$ is the "product" of the original matrices $b_{ki} \cdot a_{ij}$.

Example: Suppose $F\colon \mathbf{R}^3 \to \mathbf{R}^2$ and $G\colon \mathbf{R}^2 \to \mathbf{R}^4$ are linear map determined by

$$F(e_1) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, F(e_2) = \begin{pmatrix} 0 \\ 5 \end{pmatrix}, F(e_3) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \qquad G(e_1) = \begin{pmatrix} 4 \\ 1 \\ 0 \\ -2 \end{pmatrix}, G(e_2) = \begin{pmatrix} 0 \\ 1 \\ 5 \\ 0 \end{pmatrix}.$$

The matrix corresponding to $G \circ F$ is the product of the matrices corresponding to G and F:

$$\begin{pmatrix} 4 & 0 \\ 1 & 1 \\ 0 & 5 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & -1 \\ 2 & 5 & 0 \end{pmatrix} = \begin{pmatrix} 12 & 0 & -4 \\ 5 & 5 & -1 \\ 10 & 25 & 0 \\ -6 & 0 & 2 \end{pmatrix}.$$

Exercise: Verify that matrix multiplication (when defined) is associative. Show by example that matrix multiplication is not commutative.

Note that you can't always compose two maps, since you need to have the range of one be the domain of the other. Likewise, you can't always multiply matrices. You need the number of columns of one match the number of rows of the other.

Linear operators, square matrices: Linear operators are linear maps from a vector space V to itself. They are sometimes known as *self-maps* or *endomorphisms*. Given a choice of basis for V, they are represented by square matrices.

Since you can always compose self-maps, $\mathrm{End}\left[V\right] = \mathrm{Hom}\left[V, V\right]$ forms a ring with multiplication being composition. Likewise, since you can always multiply square matrices, they form a ring as well. A choice of basis $\{b_1, ... b_n\}$ for V gives a ring isomorphism between the set of all linear operators on V and the ring of $n \times n$ matrices over F.

Example 1: For any basis, the identity map is represented by the multiplicative identity of the matrix ring, i.e. the matrix

$$\begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{pmatrix}$$

Example 2: Let F be rotation about a point p by angle θ counter-clockwise in the plane E^2. If we choose the origin to be p and pick cartesian coordinates, then F is represented by the matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$