MATH 291 WORKSHOP 4

MARK H. KIM

Contents

1. Vector-Space Dimensions

Informally, the *dimension* of a space is the minimum number of quantities required to specify the configuration of its elements. For example, to describe an arbitrary point in the Cartesian plane, two coordinates are just enough—it is not enough to give one coordinate, and three coordinates are superfluous. In fact, the Euclidean n-space

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R} \text{ for } 1 \leq i \leq n\}$$

should be of dimension n under any reasonable definition of dimension. In determining the dimension of a space X, it therefore makes sense to compare X to Euclidean spaces and try to find an $n \in \mathbb{N}$ such that X is, in a suitable sense, similar to $\mathbb{R}^n$.

We first review the notion of dimension discussed in §1.3 of [1]. It is natural to imagine that the Cartesian coordinate system for $\mathbb{R}^n$ entails n axes perpendicular to one another. Since every collection of orthonormal vectors establishes such a coordinate system, we expect the span of k orthonormal vectors to be a k-dimensional space. Conversely, every subset V of $\mathbb{R}^n$ that is closed under vector addition and scalar multiplication—called a *linear subspace* of $\mathbb{R}^n$—admits an *orthonormal basis*. The number of vectors in each orthonormal basis of V is a fixed constant k, whence it makes sense to set k to be the dimension of V.

Defining the dimension in terms of orthonormal bases is intuitive and pictorally satisfying, but the orthonormality is superfluous in determining the dimension of a subspace of $\mathbb{R}^n$. Recall from §4.5 of [1] that a set $\{v_1, \ldots, v_k\} \subseteq \mathbb{R}^n$ is *linearly independent* if the identity

$$\lambda_1 v_1 + \cdots + \lambda_k v_n = 0$$

forces each scalar λ_i to be zero. We shall see that k should be the dimension of the span V of $\{v_1, \dots, v_k\}$.

To this end, we first recall the Gram-Schmidt process (see Theorem 53 of [1] for a proof):

Theorem 1 (Gram-Schmidt)**.** *If $\{v_1 \dots, v_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, then there is a set $\{u_1, \dots, u_k\}$ of orthonormal vectors in $\mathbb{R}^n$ such that each u_i is a linear combination of $\{v_1, \dots, v_n\}$ and that each v_i is a linear combination of $\{u_1, \dots, u_n\}$. In fact, it suffices to set*

$$\begin{aligned} u_1 &= \frac{v_1}{\|v_1\|} \\ u_2 &= \frac{v_2 - (v_2 \cdot u_1)u_1}{\|v_2 - (v_2 \cdot u_1)u_1\|} \\ &\vdots \\ u_k &= \frac{v_k - (v_k \cdot u_1)u_1 - \cdots - (v_k \cdot u_{k-1})u_{k-1}}{\|v_k - (v_k \cdot u_1)u_1 - \cdots - (v_k \cdot u_{k-1})u_{k-1}\|}. \end{aligned}$$

Since each u_i is a linear combination of $\{v_1, \dots, v_k\}$ the subspace V contains the orthonormal set $\{u_1, \dots, u_k\}$. Therefore, the dimension of V should be at least k. To show that the dimension of V should not exceed k, we establish the following theorem, whose proof is taken from pp.48-49 of [2]:

Theorem 2. *If $\{w_1, \dots, w_p\}$ is a linearly independent subset of the span of $s_1, \dots, s_q$, then $p \leq q$.*

Proof. For notational convenience, we let W denote the span of $s_1, \dots, s_q$. Suppose for a contradiction that $p > q$. We list the two sets of vectors—the spanning set, and then the linearly independent set:

$$s_1, \dots, s_q || w_1, \dots, w_p.$$

Since $s_1, \dots, s_q$ span V, we can write w_1 as

$$w_1 = \lambda_1 s_1 + \cdots + \lambda_q s_q$$

for some scalars $\lambda_1, \dots, \lambda_q$, at least one of which is nonzero. Reindexing if necessary, we can assume that $\lambda_1 \neq 0$. This, in particular, implies that

$$s_1 = \frac{1}{\lambda_1}(w_1 + (-\lambda_2)s_2 + \cdots + (-\lambda_q)s_q)$$

is in the span of $w_1, s_2, \dots, s_q$, and so $w_1, s_2, \dots, s_q$ spans W. Since every subset of a linearly independent set is linearly independent, we see that the list

$$w_1, s_2, \dots, s_q || w_2, \dots, w_p.$$

preserves the pattern exhibited above—a spanning set, and then a linearly independent set. Since $p > q$, we can move w_2 through w_q from right to left and cancel out s_2 through s_q by similar reasoning. We then end up with the list

$$w_1, \dots, w_q || w_{q+1}, \dots, w_p,$$

which implies that $w_1, \dots, w_q$ spans W. This is evidently absurd: by linear independence of the set $\{w_1, \dots, w_p\}$, the span of $w_1, \dots, w_q$ cannot contain, for example, w_{q+1}, whence the span cannot be all of W. It follows that $p \leq q$. □

Since V was span of $v_1, \ldots, v_k$, the above theorem implies that V cannot contain a linearly independent subset of size larger than k. Recalling that orthonormal sets are linearly independent, we see that V cannot contain an orthonormal subset of size larger than k. It therefore suffices to employ linearly independent sets to specify the dimension of linear subspaces of $\mathbb{R}^n$:

Definition 3. A *basis* of a linear subspace V of $\mathbb{R}^n$ is a linearly independent subset of V whose span is V. The *vector-space dimension* of V is the number of vectors in a basis of V.

The definition makes sense, as Theorem 2 tells us that all bases of V contain the same number of vectors.

Remark. If A is an m-by-n matrix $\{v_1, \ldots, v_k\}$ a set of linearly independent vectors in $\mathbb{R}^n$, then $\{Av_1, \ldots, Av_k\}$ is a set of linearly independent vectors in $\mathbb{R}^m$—why is this true? This, in particular, implies that linear transformations map a basis of its domain into a basis of its range: see §2 for terminology.

Let us now consider a few examples of linear subspaces.

Example 1. Consider the solution set H_1 of the equation

$$x_1 + 2x_2 + 3x_3 + 4x_4 = 0$$

in $\mathbb{R}^4$. We first note that H_1 is a linear subspace of H_1. Since

$$\left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right\}$$

is a linearly independent subset of H_1, the linear subspace H_1 is of dimension at least 3. Now,

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is not in the solution set H_1, and so H_1 is of dimension at most 3. It follows that $\dim H_1 = 3$.

Example 2. Consider now the linear subspace H_2 of $\mathbb{R}^5$ spanned by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 4 \\ 1 \end{pmatrix},$$

which is of dimension 4. We claim that H_2 is the solution set of the linear equation

$$12x_1 + 6x_2 + 4x_3 + 3x_4 - 12x_1 = 0.$$

Indeed, the solution set is a linear subspace containing the four spanning vectors of H_2, hence it must contain H_2. Since

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is not in the solution set, hence the solution set must be of dimension 4. Two linear subspaces of the same dimension with the same spanning set must be equal, whence our claim is proved.

Example 3 (Hyperplanes)**.** A *hyperplane* in $\mathbb{R}^n$ is a linear subspace of dimension $n-1$. As illustrated above, every hyperplane in $\mathbb{R}^n$ can be described as the solution set of a linear equation in n variables, and, conversely, the solution set of each linear equation in n variables is a hyperplane in $\mathbb{R}^n$. This property will be illustrated further in the problem set.

2. Rank and Nullity of Matrices and Linear Transformations

Given an m-by-n matrix A, we define the *range* of A as

$$\mathrm{Ran}(A) = \{y \in \mathbb{R}^m : \text{ There exists } x \in \mathbb{R}^n \text{ such that } Ax = y\}$$

and the *null space* of A as

$$\mathrm{Nullspace}(A) = \{x \in \mathbb{R}^m : Ax = 0\}.$$

It is easy to see that the range and nullspace of A are linear subspaces, whence it makes sense to talk of their (vector-space) dimensions.

Definition 4. Let A be an m-by-n matrix. The *rank* of A is the dimension of the range of A and is denoted by $\mathrm{rank}(A)$. The *nullity* of A is the dimension of the null space of A and is denoted by $\mathrm{null}(A)$.

Since every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ admits a unique m-by-n matrix A such that $T(x) = Ax$, we can extend the above definitions to linear transformations. It therefore makes sense to talk about the range, null space, rank, and nullity of linear transformations.

We now state the main theorem of the section (see Theorem 52 in [1]):

Theorem 5. *If A is an m-by-n matrix, then*

$$\mathrm{rank}(A) + \mathrm{null}(A) = n.$$

Similarly, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$\mathrm{rank}(T) + \mathrm{null}(T) = n.$$

Let us now return to the problem of determining the dimension of linear subspaces. Let V be the linear subspace spanned by vectors $v_1, \ldots, v_k$ in $\mathbb{R}^n$. The elements of V are precisely the vectors

$$\left(\begin{bmatrix} \\ v_1 \\ \\ \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \\ v_k \\ \\ \end{bmatrix} \right) \begin{pmatrix} x_1 \\ \ddots \\ x_k \end{pmatrix} = x_1 v_1 + \cdots + x_k v_k$$

for arbitrary scalars $x_1, \ldots, x_k$, and so

$$V = \operatorname{Ran}\left(\begin{bmatrix} \\ v_1 \\ \\ \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \\ v_k \\ \\ \end{bmatrix} \right).$$

Therefore, finding the dimension of V reduces to finding the rank of the above matrix, which we shall investigate in the problem set. We conclude the section by quoting a useful computational principle, which can be found as Theorem 55 in [1]:

Theorem 6. *Let A be an m-by-n matrix. If $\{r_1, \ldots, r_m\}$ are the row vectors of A, then*

$$\operatorname{rank}(A) = \dim(\operatorname{span}\{r_1, \ldots, r_m\}).$$

3. Suggested Solutions to the Challenge Problems

3.1. **Problem 1.** Let

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ -2 & 0 \end{pmatrix}.$$

Since the column vectors of A are linearly independent[1], we can apply the Gram-Schmidt process to the column vectors to deduce that

$$\left\{ \begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix}, \begin{pmatrix} -10/3\sqrt{17} \\ 7/3\sqrt{17} \\ 2/3\sqrt{17} \end{pmatrix} \right\}$$

is an orthonormal basis of $\operatorname{Ran}(A)$. The rank of A is therefore 2, which is the dimension of $\operatorname{Ran}(A)$.

Since the range of A is a two-dimensional linear subspace of $\mathbb{R}^3$, it is a hyperplane in $\mathbb{R}^3$. We therefore expect to be able to find a single linear equation in variables x_1, x_2, and x_3 whose solution set is $\operatorname{Ran}(A)$.

$$\left\{ \begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix}, \begin{pmatrix} -10/3\sqrt{17} \\ 7/3\sqrt{17} \\ 2/3\sqrt{17} \end{pmatrix} \right\}$$

constitute a basis of $\operatorname{Ran}(A)$, whence the solution set of

$$2x_1 + 2x_2 - 3x_3 = 0 \tag{1}$$

is $\operatorname{Ran}(A)$. To see if the equation

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ -2 & 0 \end{pmatrix} x = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \tag{2}$$

has a solution, we simply check if $(2, 1, 1)$ satisfies equation (1). Since

$$2 \cdot 2 + 2 \cdot 1 - 3 \cdot 1 = 3 \neq 0,$$

we conclude that equation (2) does not have a solution. □

[1] How can we tell whether a set of vectors is linearly independent? The problem sheet suggests a method via the Gram-Schmidt process—this certainly works. I, however, prefer not to apply Gram-Schmidt before I have a set of linearly independent vectors, as the hypothesis calls for such a set. Here I would bring in the machinery of linear algebra—specifically, the Gauss-Jordan elimination—to test the linear independence: See Theorem 56 in [1]

3.2. **Problem 2.** We now let

$$A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 2 & 0 & -2 & -2 \\ -2 & 1 & 0 & 1 \end{pmatrix}.$$

Since the set of column vectors

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$$

contains four vectors in $\mathbb{R}^3$, it cannot be linearly independent. We therefore must find a basis of $\operatorname{Ran}(A)$ before we can apply the Gram-Schmidt process. Noting that $\operatorname{rank}(A) \leq 3$, we simply throw away the last vector:

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\}$$

Is this set linearly independent? Unfortunately,

$$\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

and so the set is not linearly independent. Since no vector is a scalar multiple of another, we are free to choose any one vector to throw away and obtain a linearly independent set[2]

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

We now apply the Gram-Schmidt process to obtain an orthonormal basis

$$\left\{ \begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix}, \begin{pmatrix} -2/3 \\ 2/3 \\ 1/3 \end{pmatrix} \right\}$$

of $\operatorname{Ran}(A)$, and conclude that $\operatorname{rank}(A) = 2$.

Since the range of A is a two-dimensional linear subspace of $\mathbb{R}^3$, it is a hyperplane in $\mathbb{R}^3$. We therefore expect to be able to find a single linear equation in variables x_1, x_2, and x_3 whose solution set is $\operatorname{Ran}(A)$.

$$\left\{ \begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix}, \begin{pmatrix} -2/3 \\ 2/3 \\ 1/3 \end{pmatrix} \right\}$$

constitute a basis of $\operatorname{Ran}(A)$, whence the solution set of

$$2x_1 + x_2 + 2x_3 = 0. \tag{3}$$

---

[2]Note that the set $\{v_1, v_2\}$ of two vectors are not linearly independent if and only if we can find scalars λ_1 and λ_2, at least one of which is nonzero, such that $\lambda_1 v_1 + \lambda_2 v_2 = 0$. Relabeling if necessary, we can assume that $\lambda_1 \neq 0$, which implies $v_2 = (\lambda_2/\lambda_1)v_1$. Therefore, $\{v_1, v_2\}$ is not linearly independent precisely in case one vector is a scalar multiple of another.

is Ran(A). To see if the equation

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 2 & 0 & -2 & -2 \\ -2 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix} \tag{4}$$

has a solution, we simply check if $(2, 1, 1)$ satisfies equation (3). Since

$$2 \cdot 1 - 4 + 2 \cdot 1 = 0,$$

we conclude that equation (4) has at least one solution.

Let

$$b = \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}.$$

What is the full solution set of $Ax = b$? Let x_p be a solution of the equation, whose existence is guaranteed by the argument above. If x_h is a solution of the equation $Ax = 0$, then

$$A(x_p + x_h) = Ax_p + Ax_h = b + 0 = b.$$

Therefore, the set

$$S = \{x_p + x_h : Ax_h = 0\}$$

is contained in the solution set of $Ax = b$. On the other hand, if x is *not* of the form $x_p + x_h$, then we can find x_0 such that $x = x_p + x_0$ and $Ax_0 \neq 0$. It then follows that

$$Ax = A(x_p + x_0) = Ax_p + Ax_0 \neq b,$$

whence S is precisely the solution set of $Ax = b$.[3] □

3.3. **Problem 3.** Let $a_1 = (x_1, y_1, z_1)$ and $a_2 = (x_2, y_2, y_z)$ be three-dimensional row vectors and define a 2-by-3 matrix

$$A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}.$$

We claim that rank(A) = 2 if and only if $a_1 \times a_2 \neq 0$. By Theorem 6, rank(A) = 2 if and only if dim(span$\{a_1, a_2\}$) = 2. This is true precisely in case $\{a_1, a_2\}$ are linearly independent, or, equivalently, in case a_1 is not a constant multiple of a_2. It now suffices to recall that $a_1 \times a_2 \neq 0$ if and only if a_1 is not a constant multiple of a_2. □

[3]The general principle illustrated here is as follows: Given a nonzero vector b, the solution set of an *inhomogeneous equation* $Ax = b$ is precisely the solution set of the corresponding *homogeneous equation* $Ax = 0$ translated by a particular solution x_p of $Ax = b$, viz., the set

$$S = x_p + H = \{x_p + x_h : x_h \in H\}, \text{ where } H = \{x_h : Ax_h = 0\}.$$

4. Concluding Remark

Linear algebra provides a rich theoretical background for the dimension theory of linear subspaces, the category of linear subspaces is rather limited. After all, it only encompasses flat subspaces that go through the origin. One remedy is to consider *affine subspaces*, which are flat subspaces that do not necessarily go through the origin. In particular, an *affine hyperplane* in $\mathbb{R}^n$ is the solution set of an inhomogeneous linear equation

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = b$$

in n variables, where $b \neq 0$ and at least one of the scalars $\lambda_1, \ldots, \lambda_n$ is nonzero—compare this to the corresponding hyperplane, which is the solution set of

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0.$$

Affine hyperplanes appear, for example, in differentiation theory: the gradient ∇f of a differentiable scalar field f defines an affine hyperplane, referred to as the *tangent hyperplane*, by specifying the normal vector of the hyperplane. Typically, the theory of planes (affine hyperplanes in $\mathbb{R}^3$) and tangent planes (tangent hyperplanes in $\mathbb{R}^3$) is discussed.

Can we consider "curved" subspaces of $\mathbb{R}^n$? Before we proceed, we remark that every linear subspace "looks like" some Euclidean space. What do we mean by this? If V is a k-dimensional linear subspace of $\mathbb{R}^n$ with a basis $\{v_1, \ldots, v_k\}$, then the map $T : V \to \mathbb{R}^k$ defined by setting

$$T(\lambda_1 v_1 + \cdots + \lambda_k v_k) = (\lambda_1, \ldots, \lambda_k)$$

is a bijective linear map whose inverse is also linear—such a map is called a *linear isomorphism*. Geometrically, the existence of a linear isomorphism implies that the two linear subspaces ($\mathbb{R}^k$ is a linear subspace of $\mathbb{R}^k$, after all!) are "the same", as far as the structure of linear subspaces are concerned. By a similar reasoning, we can see that all affine hyperplanes in $\mathbb{R}^n$ are "affine-isomorphic" to $\mathbb{R}^{n-1}$.

Extending this idea, we now consider a category of "locally flat" subspaces, the manifolds. Given two open sets U and V in $\mathbb{R}^n$, we say that a map $h : U \to V$ is a *homeomorphism* if h is a bijective continuous map whose inverse is also continuous. Now, a *k-submanifold* of $\mathbb{R}^n$ ($k \leq n$) is a subset $M \subseteq \mathbb{R}^n$ such that, for each $x \in M$, there is an open set U containing x, another open set V, and a homeomorphism $h : U \to V$ such that

$$h(U \cap M) = V \cap (\mathbb{R}^k \times \{0\}) = \{y \in V : y^{k+1} = \cdots = y^n = 0\}.$$

In other words, at each point of M, there is a "neighborhood" of it that "looks like" $\mathbb{R}^k$, the "flat" Euclidean space. In fact, the neighborhood *is* $\mathbb{R}^k$, as far as the structure preserved by continuous maps—this is called a *topology*—is concerned.

Analogously, we define a map $h : U \to V$ between two open sets U and V in $\mathbb{R}^n$ to be a *diffeomorphism* if h is bijective, partial derivatives of h of all orders exist and are continuous—"infinitely differentiable", in other words—and the inverse of h is also infinitely differentiable. Then, a *differentiable k-submanifold* is obtained by replacing "homeomorphism" with "diffeomorphism". But just why do we care about differentiable manifolds? We conclude this section with a brief advertisement from our sponsor—differential geometry. As it turns out, the study of these subspaces is greatly facilitated by the ability to do calculus on them. Now, homeomorphisms do not bring enough structure onto the subspace it defines, so we cannot do calculus on garden-variety manifolds. We need to carry over the *differential structure* as well, which diffeomorphisms are well-equipped to do. With this structure in hand, we can generalize the notion of tangent hyperplanes ("tangent spaces") to develop a theory of differentiation. Furthermore, the theory of vector calculus and its glorious Green-Gauss-Stokes trifecta (which you will learn in Chapter 9) lend itself to a natural generalization on differentiable manifolds.

References

[1] Carlen, Eric. *Multivariable Calculus, Linear Algebra and Differential Equations*
[2] Roman, Steven. *Advanced Linear Algebra* (2e)
[3] Spivak, Michael. *Calculus on Manifolds*
[4] Spivak, Michael. *A Comprehensive Introduction to Differential Geometry* (3e)