Linear Algebra, revisited (beyond useless excercises)

This book is a pedagogical introduction to the coordinate-free approach in basic finite-dimensional linear algebra. Throughout this book, extensive use is made of the exterior (anticommutative, "wedge") product of vectors: $\mathbf{a} \wedge \mathbf{b}$. The coordinate-free formalism and the exterior product, while somewhat more abstract, provide a deeper understanding of the classical results in linear algebra. The reader should be already familiar with the elementary array-based formalism of vector and matrix calculations, in order to fully appreciate the approach based on exterior products. The standard properties of determinants, the ... m for multidimensional volumes, the Liouville formula, the Hamilton-Cayley theorem, ... some results concerning eigenspace projectors are derived using exterior products, ... e matrix calculations. For the benefit of students, every result is logically motivated and ... s with some hints are provided.

The tensor product is an abstract construction which is important in many applications. The motivation is that we would like to define a product of vectors, $\mathbf{u} \otimes \mathbf{v}$, which behaves as we expect a product to behave, e.g.

$$(\mathbf{a} + \lambda\mathbf{b}) \otimes \mathbf{c} = \mathbf{a} \otimes \mathbf{c} + \lambda\mathbf{b} \otimes \mathbf{c}, \quad \forall\lambda \in \mathbb{K}, \forall\mathbf{a}, \mathbf{b}, \mathbf{c} \in V,$$

and the same with respect to the second vector. This property is called **bilinearity**. A "trivial" product would be $\mathbf{a} \otimes \mathbf{b} = 0$ for all $\mathbf{a}, \mathbf{b}$; of course, this product has the bilinearity property but is useless. It turns out to be impossible to define a nontrivial product of vectors in a general vector space, such that the result is again a vector in the same space.[3] The solution is to define a product of vectors so that the resulting object $\mathbf{u} \otimes \mathbf{v}$ is not a vector from V but an element of *another space*. This space is constructed in the following definition.

[3]The impossibility of this is proved in abstract algebra but I do not know the proof.

elimination, a careful analysis of the proofs reveals that the Gaussian elimination and counting of the pivots do not disappear, they are just hidden in most of the proofs. So, instead of presenting very elegant (but not easy for a beginner to understand) "coordinate-free" proofs, which are typically presented in advanced linear algebra books, we use "row reduction" proofs, more common for the "calculus type" texts. The advantage here is that it is easy to see the common idea behind all the proofs, and such proofs are easier to understand and to remember for a reader who is not very mathematically

Chapter 4 is an introduction to spectral theory, and that is where the complex space $\mathbb{C}^n$ naturally appears. It was formally defined in the beginning of the book, and the definition of a complex vector space was also given there, but before Chapter 4 the main object was the real space $\mathbb{R}^n$. Now the appearance of complex eigenvalues shows that for spectral theory the most natural space is the complex space $\mathbb{C}^n$, even if we are initially dealing with real matrices (operators in real spaces). The main accent here is on the diagonalization, and the notion of a basis of eigesnspaces is also introduced.

Definition: Suppose V and W are two vector spaces over a field $\mathbb{K}$; then one defines a new vector space, which is called the **tensor product** of V and W and denoted by $V \otimes W$. This is the space of *expressions* of the form

$$\mathbf{v}_1 \otimes \mathbf{w}_1 + ... + \mathbf{v}_n \otimes \mathbf{w}_n, \tag{1.16}$$

where $\mathbf{v}_i \in V$, $\mathbf{w}_i \in W$. The plus sign behaves as usual (commutative and associative). The symbol $\otimes$ is a special separator symbol. Further, we postulate that the following combinations are equal,

$$\lambda(\mathbf{v} \otimes \mathbf{w}) = (\lambda\mathbf{v}) \otimes \mathbf{w} = \mathbf{v} \otimes (\lambda\mathbf{w}), \tag{1.17}$$

$$(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} = \mathbf{v}_1 \otimes \mathbf{w} + \mathbf{v}_2 \otimes \mathbf{w}, \tag{1.18}$$

$$\mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v} \otimes \mathbf{w}_1 + \mathbf{v} \otimes \mathbf{w}_2, \tag{1.19}$$

for any vectors $\mathbf{v}, \mathbf{w}, \mathbf{v}_{1,2}, \mathbf{w}_{1,2}$ and for any constant λ. (One could say that the symbol $\otimes$ "behaves as a noncommutative product sign".) The expression $\mathbf{v} \otimes \mathbf{w}$, which is by definition an element of $V \otimes W$, is called the **tensor product** of vectors $\mathbf{v}$ and $\mathbf{w}$. In the space $V \otimes W$, the operations of addition and multiplication by scalars are defined in the natural way. Elements of the tensor product space are called **tensors**.

Question: The logic behind the operation $\otimes$ is still unclear. How could we write the properties (1.17)–(1.19) if the operation $\otimes$ was not yet defined?

Answer: We actually *define* the operation $\otimes$ through these properties. In other words, the object $\mathbf{a} \otimes \mathbf{b}$ is defined as an expression with which one may perform certain manipulations. Here is a more formal definition of the tensor product space. We first consider the space of *all* formal linear combinations

$$\lambda_1\mathbf{v}_1 \otimes \mathbf{w}_1 + ... + \lambda_n\mathbf{v}_n \otimes \mathbf{w}_n,$$

which is a very large vector space. Then we introduce equivalence relations expressed by Eqs. (1.17)–(1.19). The space $V \otimes W$ is, by definition, the set of equivalence classes of linear combinations with respect to these relations. Representatives of these equivalence classes may be written in the form (1.16) and calculations can be performed using only the axioms (1.17)–(1.19). ■

This is simply because the acutal operation is dimension-dependent (?)

Determinants are introduced as a way to compute volumes. It is s... if we allow signed volumes, to make the determinant linear in ea... (and at that point students should be well aware that the linear ... lot, and that allowing negative volumes is a very small price to p... and assume some very natural properties, then we do not have any choice and arrive to the classical definition of the determinant. I would like to emphasize that initially I do not postulate antisymmetry of the determinant; I deduce it from other very natural properties of volume.

As we have seen, tensors from the space $V \otimes V$ are representable by linear combinations of the form $\mathbf{a} \otimes \mathbf{b} + \mathbf{c} \otimes \mathbf{d} + ...$, but not *uniquely* representable

[2]This is a theorem due to Grassmann (1862).

Add history of Cayley-Hamilton form

2015-September-7

Answer: It is true that the abstract definition of a linear map does not include a specification of a particular transformation, unlike the concrete definition in terms of a matrix. However, it does not mean that matrices are always needed. For a particular problem in linear algebra, a particular transformation is always specified either as a certain matrix in a given basis, or in a *geometric*, i.e. basis-free manner, e.g. "the transformation $\hat{B}$ multiplies a vector by 3/2 and then projects onto the plane orthogonal to the fixed vector **a**." In this book I concentrate on general properties of linear transformations, which are best formulated and studied in the geometric (coordinate-free) language rather than in the matrix language. Below we will see many coordinate-free calculations with linear maps. In Sec. 1.8 we will also see how to specify arbitrary linear transformations in a coordinate-free manner, although it will then be quite similar to the matrix notation.

because one can transform one such linear combination into another by using the axioms of the tensor product. Similarly, n-vectors are not uniquely representable by linear combinations of exterior products. For example,

$$\mathbf{a} \wedge \mathbf{b} + \mathbf{a} \wedge \mathbf{c} + \mathbf{b} \wedge \mathbf{c} = (\mathbf{a} + \mathbf{b}) \wedge (\mathbf{b} + \mathbf{c})$$

since $\mathbf{b} \wedge \mathbf{b} = 0$. In other words, the 2-vector $\omega \equiv \mathbf{a} \wedge \mathbf{b} + \mathbf{a} \wedge \mathbf{c} + \mathbf{b} \wedge \mathbf{c}$ has an alternative representation containing only a single-term exterior product, $\omega = \mathbf{r} \wedge \mathbf{s}$ where $\mathbf{r} = \mathbf{a} + \mathbf{b}$ and $\mathbf{s} = \mathbf{b} + \mathbf{c}$.

Exercise: Show that any 2-vector in a *three*-dimensional space is representable by a single-term exterior product, i.e. to a 2-vector of the form $\mathbf{a} \wedge \mathbf{b}$.

Hint: Choose a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and show that $\alpha\mathbf{e}_1 \wedge \mathbf{e}_2 + \beta\mathbf{e}_1 \wedge \mathbf{e}_3 + \gamma\mathbf{e}_2 \wedge \mathbf{e}_3$ is equal to a single-term product. ■

What about higher-dimensional spaces? We will show (see the Exercise at the end of Sec. 2.3.2) that n-vectors cannot be in general reduced to a single-term product. This is, however, always possible for $(N-1)$-vectors in an N-dimensional space. (You showed this for $N = 3$ in the exercise above.)

Statement: Any $(N-1)$-vector in an N-dimensional space can be written as a single-term exterior product of the form $\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_{N-1}$.

2.1 How to Multiply Vectors

of David's excellent summary 'A Unified Language for Mathematics and Physics' [4]. Anyone who is involved with Bayesian probability or MaxEnt is accustomed to the polemical style of writing, but his 6-page introduction on the deficiencies of our mathematics is strong stuff. In summary, David said that physicists had not learned properly how to multiply vectors and, as a result of attempts to overcome this, had evolved a variety of mathematical systems and notations that has come to resemble Babel. Four years on, having studied his work in more detail, we believe that he wrote no less than the truth and that, as a result of learning how to multiply vectors together, we can all gain a great increase in our mathematical facility and understanding.

The tensor product construction may appear an abstract plaything at this point, but in fact it is a universal tool to describe linear maps.

We have seen that the set of all linear operators $\hat{A} : V \to V$ is a vector space because one can naturally define the sum of two operators and the product of a number and an operator. This vector space is called the space of **endomorphisms** of V and denoted by End V.

In this section I will show that linear operators can be ... elements of the space $V \otimes V^*$. This gives a convenient way to represent a linear operator by a coordinate-free formula. Later we will see that the space Hom (V, W) of linear maps $V \to W$ is canonically isomorphic to $W \otimes V^*$.

Example 1: polynomials. Let V be the space of polynomi... ≤ 2 in the variable x, and let W be the space of polynomi... the variable y. We consider the tensor product of the ele... and $q(y) = y^2 - 2y$. Expanding the tensor product accor... we find

$$(1 + x) \otimes (y^2 - 2y) = 1 \otimes y^2 - 1 \otimes 2y + x \otimes y^2 - ...$$

Let us compare this with the formula we would obtain ... polynomials in the conventional way,

$$(1 + x)(y^2 - 2y) = y^2 - 2y + xy^2 - 2xy.$$

Note that $1 \otimes 2y = 2 \otimes y$ and $x \otimes 2y = 2x \otimes y$ according to the axioms of the tensor product. So we can see that the tensor product space $V \otimes W$ has a natural interpretation through the algebra of polynomials. The space $V \otimes W$ can be visualized as the space of polynomials in both x and y of degree at most 2 in each variable. To make this interpretation precise, we can construct a canonical isomorphism between the space $V \otimes W$ and the space of polyno...

Theorem 1: A set $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ of vectors from V is linearly independent if and only if $(\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_k) \neq 0$, i.e. it is a nonzero tensor from $\wedge^k V$.

$$\mathbf{x} \wedge \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N = x_1\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N = x_1\omega.$$

Therefore, exterior multiplication with $\mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N$ acts quite similarly to $\mathbf{v}_1^*$. To make the notation more concise, let us introduce a special **complement** operation[3] denoted by a star:

$$*(\mathbf{v}_1) \equiv \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N.$$

Then we can write $\mathbf{v}_1^*(\mathbf{x})\omega = \mathbf{x} \wedge *(\mathbf{v}_1)$. This equation can be used for computing $\mathbf{v}_1^*$: namely, for any $\mathbf{x} \in V$ the number $\mathbf{v}_1^*(\mathbf{x})$ is equal to the constant λ in the equation $\mathbf{x} \wedge *(\mathbf{v}_1) = \lambda\omega$. To make this kind of equation more convenient, let us write

$$\lambda \equiv \mathbf{v}_1^*(\mathbf{x}) = \frac{\mathbf{x} \wedge \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N}{\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge ... \wedge \mathbf{v}_N} = \frac{\mathbf{x} \wedge *(\mathbf{v}_1)}{\omega},$$

where the "division" of one tensor by another is to be understood as follows:

1.9 Index notation for tensors

So far we have used a purely coordinate-free formalism to define and describe tensors from spaces such as $V \otimes V^*$. However, in many calculations a basis in V is fixed, and one needs to compute the components of tensors in that basis. Also, the coordinate-free notation becomes cumbersome for computations in higher-rank tensor spaces such as $V \otimes V \otimes V^*$ because there is no direct means of referring to an individual component in the tensor product. The **index notation** makes such calculations easier.

Suppose a basis $\{\mathbf{e}_1, ..., \mathbf{e}_N\}$ in V is fixed; then the dual basis $\{\mathbf{e}_k^*\}$ is also fixed. Any vector $\mathbf{v} \in V$ is decomposed as $\mathbf{v} = \sum_k v_k \mathbf{e}_k$ and any covector as $\mathbf{f}^* = \sum_k f_k \mathbf{e}_k^*$. Any tensor from $V \otimes V$ is decomposed as

$$A = \sum_{j,k} A_{jk}\mathbf{e}_j \otimes \mathbf{e}_k \in V \otimes V$$

and so on. The action of a covector on a vector is $\mathbf{f}^*(\mathbf{v}) = \sum_k f_k v_k$, and action of an operator on a vector is $\sum_{j,k} A_{jk} v_k \mathbf{e}_k$. However, it is cum... some to keep writing these sums. In the index notation, one writes *only* components v_k or A_{jk} of vectors and tensors.

Remark: The preceding exercise serves to show that calculations in the coordinate-free approach are not always short! (I even specified some additional constraints on $\mathbf{u}, \mathbf{v}, \mathbf{f}^*, \mathbf{g}^*$ in order to make the solution shorter. Without these constraints, there are many more cases to be considered.) The coordinate-free approach does not necessarily provide a shorter way to find eigenvalues of matrices than the usual methods based on the evaluation of determinants. However, the coordinate-free method is efficient for the operator $\hat{A}$. The end result is that we are able to determine eigenvalues and eigenspaces of operators such as $\hat{A}$ and $\hat{B}$, regardless of the number of dimensions in the space, by using the special structure of these operators, which is specified in a purely geometric way.

Question: Can we define the complement operation for all $\mathbf{x} \in V$ by the equation $\mathbf{x} \wedge *(\mathbf{x}) = \omega$ where $\omega \in \wedge^N V$ is a fixed tensor? Does the complement really depend on the entire basis? Or perhaps a choice of ω is sufficient?

The main advantage of the index notation is that it makes computations with complicated tensors quicker. Consider, for example, the space $V \otimes V \otimes V^* \otimes V^*$ whose elements can be interpreted as operators from Hom $(V \otimes V, V \otimes V)$. The action of such an operator on a tensor $a^{jk} \in V \otimes V$ is expressed in the index notation as

$$b^{lm} = A^{lm}_{jk} a^{jk},$$

https://sites.google.com/site/winitzki/sergei-winitzkis-files

Classical mechanics: a minimal standard course

... is theoretical physics hard to learn no matter how much you already know?

https://sites.google.com/site/winitzki/top-index

Here I discuss, at some length, the motivation for introducing the exterior product. The motivation is geometrical and comes from considering the properties of areas and volumes in the framework of elementary Euclidean geometry. I will proceed with a formal definition of the exterior product in Sec. 2.2. In order to understand the definition explained there, it is not necessary to use this geometric motivation because the definition will be purely algebraic. Nevertheless, I feel that this motivation will be helpful for some readers.

Answer: No, yes, no. Firstly, $*(\mathbf{x})$ is not uniquely specified by that equation alone, since $\mathbf{x} \wedge A = \omega$ defines A only up to tensors of the form $\mathbf{x} \wedge ...$; secondly, the equation $\mathbf{x} \wedge *(\mathbf{x}) = \omega$ indicates that $*(\lambda\mathbf{x}) = \frac{1}{\lambda} *(\mathbf{x})$, so the complement map would not be linear if defined like that. It is important to keep in mind that the complement map requires an entire basis for its definition and depends not only on the choice of a tensor ω, but also on the choice of all the basis vectors. For example, in two dimensions we have $*(\mathbf{e}_1) = \mathbf{e}_2$; it is clear that $*(\mathbf{e}_1)$ depends on the choice of $\mathbf{e}_2$!

Remark: The situation is different when the vector space is equipped with a scalar product (see Sec. 5.4.2 below). In that case, one usually chooses an *orthonormal* basis to define the complement map; then the complement map is called the **Hodge star**. It turns out that the Hodge star is independent of the choice of the basis as long as the basis is orthonormal with respect to the given scalar product, and as long as the orientation of the basis is unchanged (i.e. as long as the tensor ω does not change sign). In other words, the Hodge star operation is invariant under orthogonal and orientation-preserving transformations of the basis; these transformations preserve the tensor ω. So the Hodge star operation depends not quite on the detailed choice of the basis, but rather on the choice of the scalar product and on the orientation of the basis (the sign of ω). However, right now we are working with a general space without a scalar product. In this case, the complement map depends on the entire basis.

Question: At this point it is still unclear why the antisymmetric definition is at all useful. Perhaps we could define something else, say the symmetric product, instead of the exterior product? We could try to define a product, say $\mathbf{a} \odot \mathbf{b}$, with some other property, such as

$$\mathbf{a} \odot \mathbf{b} = 2\mathbf{b} \odot \mathbf{a}.$$

Answer: This does not work because, for example, we would have

$$\mathbf{b} \odot \mathbf{a} = 2\mathbf{a} \odot \mathbf{b} = 4\mathbf{b} \odot \mathbf{a},$$

so all the "$\odot$" products would have to vanish.

We can define the *symmetric* tensor product, $\otimes_S$, with the property

$$\mathbf{a} \otimes_S \mathbf{b} = \mathbf{b} \otimes_S \mathbf{a},$$

but it is impossible to define anything else in a similar fashion.[2]

The antisymmetric tensor product is the eigenspace (within $V \otimes V$) of the exchange operator $\hat{T}$ with eigenvalue -1. That operator has only eigenvectors with eigenvalues ± 1, so the only other possibility is to consider the eigenspace with eigenvalue $+1$. This eigenspace is spanned by symmetric tensors of the form $\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}$, and can be considered as the space of symmetric tensor products. We could write

$$\mathbf{a} \otimes_S \mathbf{b} \equiv \mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a}$$

and develop the properties of this product. However, it turns out that the symmetric tensor product is much less useful for the purposes of linear algebra than the antisymmetric subspace. This book derives most of the results of linear algebra using the antisymmetric product as the main tool!

In this chapter I introduce one of the most useful constructions in basic linear algebra — the exterior product, denoted by $\mathbf{a} \wedge \mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ are vectors from a space V. The basic idea of the exterior product is that we would like to define an *antisymmetric* and bilinear product of vectors. In other words, we would like to have the properties $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$ and $\mathbf{a} \wedge (\mathbf{b} + \lambda\mathbf{c}) = \mathbf{a} \wedge \mathbf{b} + \lambda\mathbf{a} \wedge \mathbf{c}$.

Remark: Origin of the name "exterior." The construction of the exterior product is a modern formulation of the ideas dating back to H. Grassmann (1844). A 2-vector $\mathbf{a} \wedge \mathbf{b}$ is interpreted geometrically as the oriented area of the parallelogram spanned by the vectors $\mathbf{a}$ and $\mathbf{b}$. Similarly, a 3-vector $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c}$ represents the oriented 3-volume of a parallelepiped spanned by $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. Due to the antisymmetry of the exterior product, we have $(\mathbf{a} \wedge \mathbf{b}) \wedge (\mathbf{a} \wedge \mathbf{c}) = 0$, $(\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c}) \wedge (\mathbf{b} \wedge \mathbf{d}) = 0$, etc. We can interpret this geometrically by saying that the "product" of two volumes is zero if these volumes have a vector in common. This motivated Grassmann to call his antisymmetric product "exterior." In his reasoning, the product of two "extensive quantities" (such as lines, areas, or volumes) is nonzero only when each of the two quantities is geometrically "to the exterior" (outside) of the other.