Notes on Linear Algebra

Will Troiani

October 7, 2021

1 Trace

Notice that since $\text{Trace}(ABC) = \text{Trace}(BCA)$ for any triple of complex, square matrices A, B, C, the following is a well defined definition:

Definition 1.0.1. The **trace** of an operator $T : \mathbb{H} \longrightarrow \mathbb{H}$ is the trace of any matrix representation of T.

Let $|v_1\rangle, ..., |v_n\rangle$ be an orthogonal basis for $\mathbb{H}$, and let A be an operator on $\mathbb{H}$ and write $A|v_j\rangle = a_{1j}|v_1\rangle + ... + a_{nj}|v_n\rangle$, then we have

$$\langle v_i| A |v_j\rangle = a_{ij} \tag{1}$$

and so

$$\text{Trace}\, A = \sum_{i=1}^{n} \langle v_i| A |v_i\rangle \tag{2}$$

We can push this slightly further to find the trace of A precomposed with a projector. Let $|\psi\rangle$ be a unit vector in $\mathbb{H}$ and let $\{|\psi\rangle, |v_2\rangle, ..., |v_n\rangle\}$ be an orthogonal basis for $\mathbb{H}$. Then

$$\text{Trace}(A |\psi\rangle \langle\psi|) = \langle\psi| A |\psi\rangle \langle\psi|\psi\rangle + \sum_{i=2}^{n} \langle v_i| A |\psi\rangle \langle\psi|v_i\rangle$$

$$= \langle\psi| A |\psi\rangle$$

In a statement:

Lemma 1.0.2. *Let A be an operator on a Hilbert space $\mathbb{H}$ and let $|\psi\rangle \in \mathbb{H}$ be a unit vector in $\mathbb{H}$. We have the following formula:*

$$\text{Trace}(A |\psi\rangle \langle\psi|) = \langle\psi| A |\psi\rangle \tag{3}$$

Proof. Apply the previous discussion and the Gram-Schmidt process (Lemma 1.0.3). $\square$

Lemma 1.0.3 (Gram-Schmidt process)**.** *Every Hilbert space $\mathbb{H}$ admits an orthonormal basis. Moreover, there is an algorithm which converts a (not necessarily orthonormal) basis for $\mathbb{H}$ into an orthonormal basis.*

Proof. Every Hilbert space admits a basis so it suffices to prove the second claim.

Let $|v_1\rangle, ..., |v_n\rangle$ be a basis for $\mathbb{H}$ and inductively define the following set of vectors:

$$|u_1\rangle := |v_i\rangle / || |v_i\rangle ||, \qquad |u_i\rangle := \frac{|v_i\rangle - \sum_{k=1}^{i-1} \langle u_k|v_i\rangle |u_k\rangle}{|| |v_i\rangle - \sum_{k=1}^{i-1} \langle u_k|v_i\rangle |u_k\rangle ||} \tag{4}$$

These are clearly normal vectors which span all of $\mathbb{H}$, we prove orthogonality by induction on j, the key calculation is the following: let $i < j$

$$|| |u_j\rangle |||| |u_i\rangle || \langle u_j|u_i\rangle = (\langle v_i| - \sum_{k=1}^{i-1} \overline{\langle u_k|v_i\rangle} \langle u_k|) |u_i\rangle$$

$$= \langle v_i|u_i\rangle - \sum_{k=1}^{i-1} \overline{\langle u_k|v_i\rangle} \langle u_k|u_i\rangle$$

$$= \langle v_i|u_i\rangle - \overline{\langle u_i|v_i\rangle}$$

$$= 0$$

$\square$

2 Matrix algebra

Observe the following calculation, where all entries are elements of some ring R:

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \tag{5}$$

it follows that the image of a linear transformation T is given by the span of the columns of the matrix representing T. This is where the importance of the *column space* comes from.

Here is another important calculation: let a matrix M have columns $c_1, ..., c_n$ and assume these are linearly dependent, write

$$\alpha_1 c_1 + \ldots + \alpha_n c_n = 0 \tag{6}$$

where not all α_i are zero. If we multiply row i of M by a scalar λ then

$$\lambda \alpha_1 c_{1i} + \ldots + \lambda \alpha_n c_{ni} = \lambda(\alpha_1 c_{1i} + \ldots + \alpha_n c_{ni}) = 0 \tag{7}$$

and so linear dependence is preserved. Also, if we interchange two rows $i \neq j$ then linear dependence is still preserved. Finally, if we replace row i with $\lambda R_j + R_i$, where R_k denotes the kth row of M, we have

$$\alpha_1(\lambda c_{1j} + c_{1i}) + \ldots + \alpha_n(\lambda c_{nj} + c_{ni}) = \lambda(\alpha_1 c_{1j} + \ldots + \alpha_n c_{nj}) + \alpha_1 c_{1i} + \ldots + \alpha_n c_{ni} = 0 \tag{8}$$

We have proved the first half of:

Proposition 2.0.1. *Elementary row operations preserve linear dependence and independence of columns.*

Proof. Easy. $\square$

Proposition 2.0.2. *Let M be a matrix, the largest number of linearly independent columns is equal to the largest number of linearly independent rows.*

Proof. Let M' be a matrix is row echelon form which is equivalent under row reduction to M. Then the number of non-zero rows of M' is equal to the number of columns with leading 1s which in turn is equal to the number of linearly independent columns of M'. Clearly, the number of linearly independent rows of M and the number of linearly independent rows of M' agree, and that the same claim holds for the columns of M and M' follows from Proposition 2.0.1. $\square$

Definition 2.0.3. By Proposition 2.0.2, the greatest integer r such that a matrix M has r linearly independent rows/columns is a well defined notion, which we call the **rank of** M, denoted $\text{rank}\,M$.

Example 2.0.4. Let $Y \subseteq \mathbb{A}^n$ be an affine variety, $P \in Y$ a point, and $f_1, ..., f_m$ generators for the ideal IY. The linear transformation

$$\theta : IY \longrightarrow k^n$$
$$f \longmapsto \left(\frac{\partial f}{\partial x_1}(P), ..., \frac{\partial f}{\partial x_n}(P)\right)$$

is such that $\dim \operatorname{im} \theta = \operatorname{rank} ||(\partial f_i / \partial x_j)(P)||$. This is because

$$\operatorname{im} \theta = \operatorname{Span}(\theta f_1, ..., \theta f_n) \tag{9}$$

and consider the matrix M whose i^{th} row is θf_i (ie, $M = ||(\partial f_i / \partial x_j)(P)||$). We can reduce M to row echelon form (as k is a field) which results in a matrix M' whose number d of non-zero rows is the dimension of $\operatorname{im} \theta$. The last observation to make is that d is exactly $\operatorname{rank} M$.

Definition 2.0.5. Given an $n \times m$ matrix M and $r \leq \min\{n, m\}$, an r-**minor** of M is an $r \times r$ submatrix of M given by deleting rows and/or columns.

Corollary 2.0.6 (Corollary of Proposition 2.0.2)**.** *The rank of a matrix M is equal to the greatest integer r for which there exists an r-minor N of M with $\det N \neq 0$.*

Corollary 2.0.6 also uses the following observation:

Observation 2.0.7. Let M be a matrix whose columns are linearly dependent, then $\det M = 0$.

Proof. If M' is equivalent to M under row reduction then there is at least one zero row of M' and so $\det M' = 0$ which implies $\det M = 0$. $\square$

2.1 Determinant

We have:

Proposition 2.1.1. *If M is a square $n \times n$ matrix with entries in some field k then*

- *for any $l \in k$:*
$$\det(lM) = l^n \det M \tag{10}$$

- *If N is another such square matrix then:*
$$\det(MN) = \det M \det N \tag{11}$$

Proof. The first claim is obvious.

First say M is non-invertible. Then MN is also non-invertible and so both sides of (11) are zero.

Now say that M is invertible. We use the fact that M can be written as a product of elementary matrices, this follows easily from the fact that the reduced row echelon form of an invertible matrix is the identity. It now suffices to show that for any elementary matrix E,

$$\det(EM) = \det E \det M \tag{12}$$

The result then follows from a case analysis on E. $\square$

3 Diagonalisation

Throughout, we assume k is a field.

Definition 3.0.1. A square matrix A with coefficients in k is **diagonalisable** if there exists an invertible matrix P such that PAP^{-1} is diagonal.

Let V be some k-vector space. A linear transformation $T : V \longrightarrow V$ is **diagonalisable** if there exists a basis with respect to which the matrix representation of T is diagonal.

Remark 3.0.2. A linear transformation $T : V \longrightarrow V$ being diagonalisable is equivalent to the statement that the matrix representation of T (with respect to any basis) is diagonalisable. This is clumsy though because we are then saying T with respect to any basis can be written diagonally if we change the basis...

Clearly, we have:

Lemma 3.0.3. *Let V be some k-vector space. A linear transformation $T : V \longrightarrow V$ is diagonalisable if there exists a basis for V consisting of eigenvectors for T. The diagonal entries will then be the eigenvalues.*

We wish to prove the following:

Proposition 3.0.4. *Let $T : V \longrightarrow V$ be a linear transformation and $v_1, ..., v_k$ a set of eigenvectors with eigenvalues $\lambda_1, ..., \lambda_k$. If all of $\lambda_1, ..., \lambda_k$ are distinct then the set $v_1, ..., v_k$ is linearly independent.*

Proof. Assume $v_1, ..., v_k$ is linearly dependent, let j be such that $v_1, ..., v_j$ is linearly dependent but $v_1, ..., v_{j-1}$ is linearly independent (reordering if necessary). Write

$$\alpha_1 v_1 + \ldots + \alpha_j v_j = 0 \tag{13}$$

Applying T to both sides of (13) we obtain:

$$\alpha_1 \lambda_1 v_1 + \ldots + \alpha_j \lambda_j v_j = 0 \tag{14}$$

and multiplying (13) by λ_j we obtain

$$\alpha_1 \lambda_k v_1 + \ldots + \alpha_j \lambda_j v_j = 0 \tag{15}$$

Subtracting (15) from (14) we obtain

$$\alpha_1 (\lambda_1 - \lambda_k) v_1 + \ldots + \alpha_{j-1} (\lambda_{j-1} - \lambda_j) v_{j-1} = 0 \tag{16}$$

which by linear independence implies for all $i = 1, ..., j-1$ that

$$\alpha_i (\lambda_1 - \lambda_i) = 0 \tag{17}$$

That all eigenvalues are distinct then implies $\alpha_i = 0$ for $i = 1, ..., j-1$ which then implies $\alpha_j = 0$, a contradiction. $\square$

Remark 3.0.5. The proof of Proposition 3.0.4 is probably more naturally written as an inductive argument.

3.1 Jordan normal form

Wikipedia described a process for constructing a linearly independent set of *generalised eigenvectors* with respect to a linear transformation $T : V \longrightarrow V$. With respect to this basis, the matrix representation of T consists of *eigenblocks*. Corollaries of this includes that the product of eigenvalues is the determinant of the matrix, and the sum is the trace. We also arrive at the fact that if the *minimal polynomial* of T splits into linear factors, then T is diagonalisable. In particular, this implies that any *involution* (a linear transformation $T : V \longrightarrow V$ such that $T^2 = \text{id}_V$) is diagonalisable, as T satisfies the polynomial $x^2 - 1 = (x+1)(x-1)$. Hence, any vector space V which admits an involution T decomposes $V \cong V_0 \oplus V_1$ where $V_0 := \{v \in V \mid Tv = v\}$ and $V_1 := \{v \in V \mid Tv = -v\}$.

4 Quadratic forms

Throughout, V is a finite dimensional vector space over a field k.

In Linear Algebra, we study morphisms between vector spaces which (given a fixed basis) are given by matrix multiplication:

$$T : V \longrightarrow V, \qquad T(x_1, ..., x_n) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{18}$$

A **bilinear form** is a bilinear map $T : V \times V \longrightarrow k$ (where k is the bass field of V), which (again, given a fixed basis) can always be written in matrix form:

$$T : V \times V \longrightarrow k, \qquad T\big((x_1, ..., x_n), (y_1, ..., y_n)\big) = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{19}$$

First we prove that such a representation always exists.

Let V be finite dimensional and let $v_1, ..., v_n$ a basis for V. Let $a_{ij} = T(v_i, v_j)$. Then for any pair of vectors $v = \sum_{i=1}^n a_i v_i$, $u = \sum_{i=1}^n b_i v_i$ we have

$$\begin{aligned} T(v, u) &= T(\sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_i v_i) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j T(v_i, v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j a_{ij} \end{aligned}$$

and

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^n a_1 j b_j \\ \vdots \\ \sum_{j=1}^n a_{nj} b_j \end{bmatrix}$$
$$= \big[\sum_{i=1}^n \sum_{j=1}^n a_i a_{ij} b_j \big]$$

The result follows by commutativity.

Definition 4.0.1. A bilinear form $B : V \times V \longrightarrow k$ is **symmetric** if for all $v, u \in V$ we have $B(v, u) = B(u, v)$.

Definition 4.0.2. A **quadratic form** is a function $Q : V \longrightarrow k$ satisfying the following properties:

- for all $a \in k$ and $v \in V$, we have $Q(av) = a^2 Q(v)$,

- the function $B : V \times V \longrightarrow k$ given by $B(v, u) = Q(v + u) - Q(v) - Q(u)$ is linear.

Example 4.0.3. The prototypical example is given using the dot product: let $V = \mathbb{R}^n$ and define $Q(x_1, ..., x_n) = \sum_{i=1}^n x_i^2$.

Bilinear forms and quadratic forms are closely related, as is shown by the following Proposition:

Proposition 4.0.4. *Let $B : V \times V \longrightarrow k$ be a symmetric bilinear form and k a field of characteristic not equal to 2. Then the function $Q_B : V \longrightarrow k$ given by $Q(v) = B(v, v)$ is a bilinear form.*

Also, given a quadratic form $Q : V \longrightarrow k$, the function $B_Q : V \times V \longrightarrow k$ given by $B_Q(v, u) = \frac{1}{2}\big(Q(v + u) - Q(v) - Q(u)\big)$ is a bilinear form.

Furthermore, the maps $B \longrightarrow Q_B$ and $Q \longrightarrow B_Q$ are mutually inverse bijections.

Proof. We have

$$B_{Q_B}(v, u) = \frac{1}{2}\big(Q_B(v+u) - Q_B(v) - Q_B(u)\big)$$
$$= \frac{1}{2}\big(B(v+u, v+u) - B(v,v) - B(u,u)\big)$$
$$= \frac{1}{2}\big(B(v,u) + B(v,u)\big)$$
$$= B(v,u)$$

where we use symmetry in the final line.

Also,

$$Q_{B_Q}(v) = B_Q(v,v)$$
$$= \frac{1}{2}\big(Q(v+v) - Q(v) - Q(v)\big)$$
$$= Q(v)$$

The rest follows in a similar, simple fashion. □

Example 4.0.5. The assumption that k have characteristic not equal to 2 is crucial: let $\mathbb{F}_2$ denote the field with two elements and consider the following quadratic form:

$$Q : \mathbb{F}_2^2 \longrightarrow \mathbb{F}_2$$
$$(x, y) \longmapsto xy$$

we claim that this cannot arise as Q_B for any symmetric bilinear form $B : \mathbb{F}_2^2 \times \mathbb{F}_2^2 \longrightarrow \mathbb{F}_2$. To see this, notice that

$$Q_B(x, y) = B\big((x,y),(x,y)\big)$$
$$= x^2 B\big((1,0),(1,0)\big) + yx B\big((0,1),(1,0)\big) + xy B\big((1,0),(0,1)\big) + y^2 B\big((0,1),(0,1)\big)$$
$$= x^2 B\big((1,0),(1,0)\big) + 2xy B\big((0,1),(1,0)\big) + y^2 B\big((0,1),(0,1)\big)$$
$$= x^2 B\big((1,0),(1,0)\big) + y^2 B\big((0,1),(0,1)\big)$$
$$= x^2 Q_B(1,0) + y^2 Q_B(0,1)$$

which does not fit the form of Q. Thus in the characteristic 2 case, there are quadratic forms which do not arise from symmetric bilinear forms.

4.1 Clifford algebras

4.2 Graded, differential algebras

We generalise the concept of polynomial algebras to *differential, graded algebras.*

Definition 4.2.1. Let G be a totally ordered, abelian group, a G-**graded algebra** is an algebra A along with a decomposition

$$G \cong \bigoplus_{g \in G} A_g \tag{20}$$

satisfying $A_g A_h \subseteq A_{g+h}$.

A nonzero element of A_g is a **homogeneous element of degree** g and we write $\deg a = g$ for $a \in A_g$.

Definition 4.2.2. A **differential $\mathbb{Z}$-graded algebra** is a $\mathbb{Z}$-graded algebra $A := \bigoplus_{i \in \mathbb{Z}} A_i$ along with a **differential** $d : A \longrightarrow A$ satisfying

$$\deg d(a) = \deg a - 1 \tag{21}$$

The following Definitions are justified by Remark ??????

Definition 4.2.3. Let $A \cong \bigoplus_{i \in \mathbb{Z}} A_i, B \cong \bigoplus_{i \in \mathbb{Z}} B_i$ be graded k-algebras (where k is some ring). The **graded tensor product** $A \otimes B$ is the normal tensor product of k-modules but with multiplication defined on pure tensors as:

$$(a_1 \otimes_k b_1)(a_2 \otimes_k b_2) = (-1)^{\deg a_2 \deg b_1}(a_1 a_2 \otimes_k b_1 b_2) \tag{22}$$

The reason for this definition is given in Remark ?????
The k-algebra $\operatorname{Hom}_k(A, B)$ is also naturally a graded algebra. For $n \in \mathbb{Z}$ we define

$$\operatorname{Hom}_k(A, B)_n := \{ f \in \operatorname{Hom}_k(A, B) \mid \forall m \in \mathbb{Z}, \forall a \in A_m, \deg f(a) = n + \deg a \} \tag{23}$$

In fact this is a graded algebra with grading

$$A \otimes B \cong \bigoplus_{\substack{i \in \mathbb{Z} \\ n+m=i}} A_n \oplus B_m \tag{24}$$

Furthermore, say A, B were *differential*, graded algerbas with respective differentials d_A, d_B. Then $A \otimes B$ is also differential graded with differential given on pure tensors by

$$d_{A \otimes B}(a \otimes b) = d_A(a) \otimes b + (-1)^{\deg a \deg b} a \otimes d_B(b) \tag{25}$$