# IB Linear Algebra (from lecture 18)

Martin von Hodenberg (`mjv43@cam.ac.uk`)

December 27, 2021

Linear algebra description etc

This article constitutes my notes for the 'IB Linear Algebra' course, held in Michaelmas 2021 at Cambridge. The course was lectured by Prof. Pierre Raphael.

## Contents



## §1 Bilinear forms

**Lemma 1.1**

We have a bilinear form $\varphi : V \times V \to \mathbb{F}$, where $V$ is a finite vector space and $B, B'$ are bases of $V$. Let

$$\varphi = [\mathrm{Id}]_{B,B'}.$$

Then

$$[\varphi]_{B'} = P^T[\varphi]_B P.$$

*Proof.* This is just a special case of the general change of basis formula; see that proof. □

**Definition 1.2** (Congruent matrices)

Two square matrices $A, B$ are said to be **congruent** if there exists an invertible square matrix $P$ such that

$$A = P^T BP.$$

**Remark.** This defines an equivalence relation.

**Definition 1.3** (Symmetric bilinear form)
A bilinear form on $V$ is said to be **symmetric** if

$$\varphi(u, v) = \varphi(v, u) \ \forall u, v \in V.$$

**Remark.** We have encountered this definition before in IA Vectors and Matrices.

1. If $A$ is a square matrix, we say that $A$ is symmetric if $A^T = A$. Equivalently, $A_{ij} = A_{ji}$.
2. $\varphi$ is symmetric iff $[\varphi]_B$ is symmetric in *any* basis $B$.
3. To be able to represent $\varphi$ by a diagonal matrix in some basis $B$, it is necesssary that $\varphi$ is symmetric:

$$P^T AP = D = D^T = PA^T P^T \implies A = A^T \implies \varphi \text{ is symmetric.}$$

**Definition 1.4** (Quadratic form)
A map $Q : V \to F$ is said to be a **quadratic form** if there exists a bilinear form $\varphi : V \times V \to F$ such that

$$\forall u \in V,\ Q(u) = \varphi(u, u).$$

**Remark** (Computation in a basis)**.** Let $B = (e_i)_{1\leq i\leq n}$ be a basis of $V$, and let $A = [\varphi]_B$. Let $u = \sum_{i=1}^n u_i e_i$, then

$$Q(u) = \varphi(u, u) = \varphi\left(\sum_{i=1}^n u_i e_i, \sum_{j=1}^n u_j e_j\right) = \sum_{i,j=1}^n u_i u_j \varphi(e_i, e_j) = \sum_{i,j=1}^n a_{ij} u_i u_j.$$

(by bilinearity of $\varphi$) Therefore we essentially have

$$Q(u) = U^T AU, \text{ where } U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

**Remark.** We can note that

$$Q(u) = U^T AU = \sum_{i,j=1}^n a_{ij} u_i u_j = \sum_{i,j=1}^n \left(\frac{a_{ij} + a_{ji}}{2}\right) u_i u_j = U^T \left(\frac{A + A^T}{2}\right) U.$$

So this representation of $A$ is not necessarily unique.

**Proposition 1.5**
If $Q : V \to F$ is a quadratic form, then there exists a unique symmetric bilinear form $\varphi : V \times V \to F$ such that

$$Q(u) = \varphi(u, u) \ \forall u \in V.$$

*Proof.* (Polarisation identity)
**Proof of existence**
Let $\psi$ be a bilinear form on $V$ such that

$$\forall u \in V,\ Q(u) = \psi(u,u).$$

Let $\varphi(u,v) = \frac{1}{2}(\psi(u,v) + \psi(v,u))$. Thus we have that:

- $\varphi$ is a bilinear form
- $\varphi$ is symmetric
- $\varphi(u,u) = \psi(u,u) = Q(u)$.

This concludes the proof of existence.
**Proof of uniqueness**
Let $\varphi$ be a symmetric bilinear form such that

$$\forall u \in V,\ \varphi(u,u) = Q(u).$$

Then

$$\begin{aligned} Q(u+v) &= \varphi(u+v, u+v) \\ &= \varphi(u,u) + \varphi(u,v) + \varphi(v,u) + \varphi(v,v) \text{ by bilinearity} \\ &= Q(u) + 2\varphi(u,v) + Q(v) \text{ by symmetry} \end{aligned}$$

From this we get that

$$\varphi(u,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v)).$$

□

**Theorem 1.6** (Diagonalisation of symmetric bilinear forms)
Let $\varphi : V \times V \to F$ be a symmetric bilinear form. (dim $V$=n). Then there exists a basis $B$ of $V$ such that $[\varphi]_B$ is diagonal.

*Proof.* We proceed by induction on the dimension of $V$. For $n = 1$ it is trivially true. Suppose the theorem holds for all dimensions $< n$: then

- If $\varphi(u,u) = 0 \forall u \in V$, then $\varphi = 0$ by the polarisation identity ($\varphi$ is symmetric).
- If $\varphi \neq 0$, then there exists a $u \in V \backslash \{0\}$ such that $\varphi(u,u) \neq 0$. Let us call $u = e_1$.
- Let $U$ be the 'orthogonal' of $e_1$:

  $$\begin{aligned} U &= \{v \in V :\ \varphi(e_1, v) = 0\} \\ &= \ker\theta \text{ where } \theta : V \to F \text{ is given by } \theta(v) = \varphi(e_1, v). \end{aligned}$$

  Since it is a kernel of a linear map $V \to F$, $U$ is a vector subspace of $V$. By the Rank-Nullity theorem, we have

$$\dim V = n = r(\theta) + \operatorname{null}\theta = \dim U + 1.$$

We now claim that $U+ < e_1 >= U \oplus < e_1 >$. Indeed,

$$v =< e_1 > \cap U \implies v = \lambda e_1 \text{ and } \varphi(e_1, v) = 0.$$

$$\implies 0 = \varphi(e_1, v) = \varphi(e_1, \lambda e_1) = \lambda\varphi(e_1, e_1) \implies \lambda = 0 \implies v = 0.$$

$$\implies U+ < e_1 >= U \oplus < e_1 > .$$

Therefore $V = U \oplus < e_1 >$, and pick a basis $B' = (e_2, \ldots, e_n)$ such that $(e_1, e_2, \ldots, e_n)$ is a basis of $V$ (since the sum is direct). So

$$[\varphi]_B = (\varphi(e_i, e_j))_{1 \leq i,j \leq n} = \left(\begin{array}{c|c} \varphi(e_1, e_1) & 0 \\ \hline 0 & A' \end{array}\right).$$

Therefore $(A')^T = A'$, and $A' = [\varphi|_U]_{B'}$ where $\varphi|_U$ is the restriction of $\varphi$ onto $U$. Now we apply the induction hypothesis to find a basis $(e'_1, \ldots, e_n;)$ of $V$ such that $[\varphi|U]_B$ is diagonal. So

$$\hat{B} = (e_1, e'_2, \ldots, e'_n)$$

is a basis of $V$, and finally we have that $[\varphi]_{\hat{B}}$ is diagonal.

□

**Example 1.7**

Let $V = \mathbb{R}^3$, and

$$\begin{aligned} Q(x_1, x_2, x_3) &= x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3 \\ &= x^T A x \text{ where } A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix} \end{aligned}$$

There are two ways to diagonalise our quadratic form:

1. Diagonalise using the algorithm we developed in the previous proof.
2. Complete the square.

Let's complete the square:

$$\begin{aligned} Q(x_1, x_2, x_3) &= (x_1 + x_2 + x_3)^2 + x_3^2 - 4x_2x_3 \\ &= (x_1 + x_2 + x_3)^2 + (x_3 - 2x_1)^2 - (2x_2)^2 \\ &= (x'_1)^2 + (x'_2)^2 + (x'_3)^2. \end{aligned}$$

Therefore

$$P^T A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

To find $P$, note that

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

## §1.1 Sylvester's law and sesquilinear forms

Recall our theorem that a symmetric bilinear form has a diagonal basis.

**Corollary 1.8**

Let $F = \mathbb{C}$, and $\varphi$ be a symmetric bilinear form of rank $r$ on $V \times V$ where $\dim V = n$. Then there exists a basis $B$ of $V$ such that the matrix

$$[\varphi]_B = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

*Proof.* Pick $E = (e_1, \dots, e_n)$ such that

$$[\varphi]_E = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Order the $a_i$ such that

$$\begin{cases} a_i \neq 0 & 1 \leq i \leq r \\ a_i = 0 & i > r \end{cases}$$

For $i \leq r$, let $\sqrt{a_i}$ be a choice of complex root for $a_i$. Let

$$\begin{cases} v_i = \frac{e_i}{\sqrt{a_i}} & 1 \leq i \leq r \\ v_i = e_i & i > r \end{cases}$$

Then $B = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$ is a basis of $V$ and we can check that we have diagonalised and normalised $\varphi$:

$$[\varphi]_B = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

□

**Corollary 1.9**

Every symmetric matrix of $M_n(\mathbb{C})$ is congruent to a *unique* matrix of the form

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

*Proof.* Immediate. □

**Corollary 1.10**

Let $F = \mathbb{R}$. Let $\varphi$ be a symmetric bilinear form on $V \times V$ where $\dim V = n$. Then there exists a basis $B = (v_1, \ldots, v_n)$ basis of $V$ such that for some $p - q \geq 0$ with $p + q = r(\varphi)$,

$$[\varphi]_B = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

*Proof.* Pick $E = (e_1, \ldots, e_n)$ such that

$$[\varphi]_E = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Order the $a_i$ such that

$$\begin{cases} a_i > 0 & 1 \leq i \leq p \\ a_i < 0 & p+1 \leq i \leq p+q \\ a_i = 0 & i > p+q \end{cases}$$

Similarly to before we define

$$\begin{cases} v_i = \frac{e_i}{\sqrt{a_i}} & 1 \leq i \leq p \\ v_i = \frac{e_i}{\sqrt{-a_i}} & p+1 \leq i \leq p+q \\ v_i = e_i & i > p+q \end{cases}$$

Then this basis will do the job. □

**Definition 1.11** (Signature of a quadratic form)

For $F = \mathbb{R}$, we define the **signature** of $\varphi$

$$s(\varphi) = p - q.$$

This is also the signature of the associated quadratic form $Q$. We will soon see that this is well-defined.

**Definition 1.12** (Positive definite quadratic/bilinear form)

Let $\varphi$ be a symmetric bilinear form on a real vector space $V$. We say that

(i) $\varphi$ is positive definite if

$$\varphi(u, u) > 0 \; \forall u \in V \backslash \{0\}.$$

(ii) $\varphi$ is positive semidefinite if

$$\varphi(u, u) \geq 0 \; \forall u \in V \backslash \{0\}.$$

(iii) $\varphi$ is negative definite if

$$\varphi(u,u) < 0 \ \forall u \in V\backslash\{0\}.$$

(iv) $\varphi$ is negative semidefinite if

$$\varphi(u,u) \leq 0 \ \forall u \in V\backslash\{0\}.$$

**Example 1.13**

The matrix

$$\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right)$$

is

- positive definite for $p = n$
- positive semidefinite for $1 \leq p \leq n$.

**Theorem 1.14** (Sylvester's law of inertia)

The signature of a quadratic form is well defined.
Formally, if a real symmetric bilinear form is represented in two different ways by two identity blocks of size $p, q$ and $p', q'$ then we have

$$p = p', q = q'.$$

*Proof.* In order to prove uniqueness of $p$, it is enough to show that $p$ is the largest dimension of a subspace of $V$ on which $\varphi$ is positive definite. Say $B = (v_1, \dots, v_n)$ and

$$[\varphi]_B = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Let $X =< v_1, \dots, v_p$. Then $\varphi$ is positive definite on $X$. If we take a $u \in X$ with $u = \sum_{i=1}^p \lambda_i v_i$, then

$$Q(u) = \varphi(u,u) = \varphi(\sum_{i=1}^p \lambda_i v_i, \sum_{j=1}^p \lambda_j v_j) = \sum_{i,j=1}^p \lambda_i \lambda_j \varphi(v_i, v_j) = \sum_{i=1}^n \lambda_i^2.$$

This is ¿0 as long as $u \neq 0$. Suppose that $\varphi$ is positive definite on another subspace $X$. Let

$$X =< v_1, \dots, v_p >, Y =< v_{p+1}, \dots, v_n > .$$

Then arguing as above, looking at $[\varphi]_B$ we know that $\varphi$ is negative semidefinite on $Y$. This implies that

$$Y \cap X' = \{0\}.$$

Indeed, if $y \in Y \cap X$, then $Q(y) \leq 0$ since $y \in Y$, but this implies that $y = 0$ since

$y \in X$. Therefore

$$Y+X = Y \oplus X \implies n = \dim V \geq \dim(Y+X) = \dim Y + \dim X \implies n \geq n-p+\dim X \implies \text{dimX} \leq p.$$

Similarly, we show that $q$ is the largest dimension of a subspace on which $\varphi$ is negative definite. So we have a *geometric characterisation* of $p$ and $q$, which concludes the proof. □

**Definition 1.15** (Kernel of a bilinear form)
Define the kernel of the bilinear form

$$K = \{v \in V : \forall u \in V, \varphi(u, v) = 0\}.$$

**Remark.** $\dim K + r(\varphi) = n$.

One can show using the above notation that there is a subspace $T$ of dimension $m-(p+q)+\min p, q$ such that $\varphi_T = 0$. The subspace we want to take consists of all the basis vectors that are killed by the lower corner of $[\varphi]_B$, but it also includes the 'cancellations' between $I_p$ and $-I_q$.
Moreover, one can show that the dimension of $T$ is the largest possible dimension of such a subspace.

### §1.1.1 Sesquilinear forms

Recall that the standard inner product on $\mathbb{C}^n$ is

$$< x, y > = < \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i \overline{y_i}.$$

However note that this map is **not** a bilinear form, since the conjugation of the second coordinate prevents this. However, it is still bilinear 'in a way'. We now explore this further.

**Definition 1.16** (Sesquilinear form)
Let $V, W$ be vector spaces over $\mathbb{C}$. A sesquilinear form on $V \times W$ is a function

$$\varphi : V \times W \to \mathbb{C}.$$

such that

1. $\varphi(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \varphi(v_1, w) + \lambda_2 \varphi(v_2, w)$
2. $\varphi(v, \lambda_1 w_1 + \lambda_2 w_2) = \overline{\lambda_1} \varphi(v, w_1) + \overline{\lambda_2} \varphi(v, w_2)$

For $B$ a basis of $V$, C a basis of $W$, we define the matrix

$$[\varphi]_{B,C} = (\varphi(v_i, w_j))_{1 \leq i \leq m, 1 \leq j \leq n}.$$

**Lemma 1.17**

If $B$ basis for $V$, $C$ a basis for $W$,

$$\varphi(v,w) = [v]_B^T[\varphi]_{B,C}\overline{[w]_C}.$$

Further, if $B', C'$ are other bases for $V, W$ respectively, and $P = [Id]_{B',B}, Q = [Id]_{C',C}$ then

$$[\varphi]_{B',C'} = P^T[\varphi]_{B,C}\overline{Q}.$$

*Proof.* Same proofs as for bilinear forms, just slightly modified. □

## §1.2 Hermitian forms and skew symmetric forms

We will now generalise symmetric bilinear forms to sesquilinear forms.

**Definition 1.18** (Hermitian form)

Let $V$ be a finite dimensional vector space and $\varphi : V \times V \to C$ be a sesquilinear form. We call $\varphi$ **hermitian** if for all $(u,v) \in V \times V$ we have

$$\varphi(u,v) = \overline{\varphi(v,u)}.$$

**Remark.** If $\varphi$ is Hermitian, then $\varphi(u,u) = \overline{\varphi(u,u)} \implies \varphi(u,u) \in \mathbb{R}$. Moreover,

$$\varphi(\lambda u, \lambda u) = |\lambda|^2\varphi(u,u).$$

This allows us to talk about negative/ positive definite Hermitian form.

**Lemma 1.19**

A sesquilinear form $\varphi : V \times V \to \mathbb{C}$ is hermitian if and only if for every basis $B$ of $V$,

$$[\varphi]_B = \overline{[\varphi]_B^T}.$$

*Proof.* Let $A = [\varphi]_B$ have entries $a_{ij} = \varphi(e_i, e_j)$. If $\varphi$ is Hermitian, then

$$a_{ji} = \varphi(e_j, e_i) = \overline{\varphi(e_i, e_j)} = \overline{a_{ij}}.$$

Thus $[\varphi]_B = \overline{[\varphi]_B^T}$.
For the converse direction, if we have $[\varphi]_B = \overline{[\varphi]_B^T}$, and we write $u = \sum_{i=1}^n u_i e_i, v = \sum_{i=1}^n v_i e_i$, then

$$\varphi\left(u,v\right) = \varphi\left(\sum_{i=1}^n u_i e_i\right), \sum_{i=1}^n v_i e_i = \sum_{i,j=1}^n u_i \overline{v_j}\varphi\left(e_i, e_j\right) = \sum_{i,j=1}^n u_i \overline{v_j} a_i j.$$

We can similarly expand $\varphi(e_j, e_i)$ to get the same result and we are done. □

**Proposition 1.20** (Polarisation identity for sesquilinear forms)

A Hermitian form $\varphi$ on a complex vector space $V$ is entirely determined by $Q : V \to$

$R$ where $Q(v) = \varphi(v, v)$ via the formula

$$\varphi(u, v) = \frac{1}{4}\left(Q(u+v) - Q(u-v) + iQ(u+iv) - iQ(u-iv)\right).$$

*Proof.* Similar to the case of symmetric bilinear forms. □

**Theorem 1.21** (Hermitian formulation of Sylvester's law)

Let $V$ be an $n$-dimensional vector space over $\mathbb{C}$. Let $\varphi : V \times V \to C$ be a Hermitian form on $V$. Then there exists a basis $B = (v_1, \ldots, v_n)$ of $V$ so that

$$[\varphi]_B = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

*Proof.* We just provide a sketch proof as the proof is nearly identical to the case of real symmetric bilinear forms.

**Existence:**

$\varphi = 0$, done. Otherwise, using the polarisation identity, there exists $e_1 \neq 0$ such that $\varphi(e_1, e_1) \neq 0$. We then normalise $e_1$ to $v_1$ and consider the orthogonal of $< v_1 >$. Then we check (arguing like for real bilinear symmetric forms):

$$V = < v_1 > \oplus W,$$

where $\dim W = n - 1$. Finally, we argue by induction to diagonalise $\varphi|_W$.

**Uniqueness:**

We argue that $p$ is the maximal dimension of a subspace on which $\varphi$ is positive definite, and that $q$ is the maximal dimension of a subspace on which $\varphi$ is negative definite. Then by our geometric characterisation we are done. □

## §1.3 Skew symmetric bilinear forms

We now return to $F = \mathbb{R}$.

**Definition 1.22** (Skew symmetric bilinear forms)

Let $V$ be a vector space over $\mathbb{R}$, and $\dim V = n$. A bilinear form $\varphi : V \times V \to \mathbb{R}$ is **skew symmetric** if:

$$\forall (u, v) \in V \times V,\ \varphi(u, v) = -\varphi(v, u).$$

**Remark.** This leads to a few easy consequences:

1. $\varphi(u, u) = \varphi(u, u) \implies \varphi(u, u) = 0.$
2. For all bases $B$ of $V$, $[\varphi]_B = -[\varphi]_B^T$
3. For all $A \in M_n(\mathbb{R})$,
$$A = \frac{1}{2}\left(A + A^T\right) + \frac{1}{2}\left(A - A^T\right).$$

(we can decompose any matrix into symmetric and skew symmetric parts)

**Theorem 1.23** (Sylvester form of skew symmetric matrices)
Let $V$ be a finite $n$-dimensional vector space over $\mathbb{R}$. Let $\varphi$ be a skew symmetric bilinear form over $V$. Then there exists a basis $B$ of $V$ such that

$$B = (v_1, w_1, v_2, w_2, \dots v_m, w_m, v_{2m+1}, v_{2m+2}, \dots, v_n).$$

**Corollary 1.24**
Skew symmetric matrices have an even rank.

*Proof.* We give a sketch proof as this proof is again similar to past proofs. We proceed by induction on the dimension of $V$.

1. If $\varphi = 0$, then we are done.
2. Else there exists $(v_1, w_1)$ such that $\varphi(v_1, w_1) \neq 0$.
3. After scaling $v_1$, we can assume $\varphi(v_1, w_1) = 1$ and thus by skew symmetricity $\varphi(w_1, v_1) = -1$. Observe $v_1, w_1$ are linearly independent since

   $$\varphi(v_1, \lambda v_1) = \lambda\varphi(v_1, v_1) = 0.$$

   We now let W be the orthogonal of $U =< v_1, w_1 >$,

   $$W = \{v \in V : \ \varphi(v_1, v) = \varphi(v, w_1) = 0\}.$$

4. Finally we show $V = U \oplus W$ and apply our inductive hypothesis.

$\square$

## §1.4 Inner product spaces

We will find that positive definite bilinear forms will be useful for a form of inner product, which leads us to norms (notions of distance). Looking past this course, this extends to an infinite dimensional counterpart of Hilbert spaces in Part II Linear Analysis.

**Definition 1.25** (Inner product)
Let $V$ be a vector space over $\mathbb{R}$ (resp. $\mathbb{C}$). An **inner product** on $V$ is a positive definite symmetric (resp. Hermitian) bilinear form $\varphi$ on $V$. We use the notation

$$< u, v >= \varphi(u, v).$$

We call $V$ a real (resp. complex) inner product space.

**Example 1.26**
Here are several examples of inner product spaces.

1. $V = \mathbb{R}^n$, where we define

$$< x, y >=< \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i.$$

2. $V = \mathbb{C}^n$, $< x, y >= \sum_{i=1}^{n} x_i \overline{y_i}$.
3. $V = C^0([0,1], \mathbb{C})$. We define

$$< f, g >= \int_0^1 f(t)\,\overline{g(t)}\ dt.$$

4. We can fix a weight $\omega : [0,1] \to \mathbb{R}_+^*$ and define on $V = C^0([0,1], \mathbb{C})$

$$< f, g >= \int_0^1 f(t)\,\overline{g(t)}\omega(t)\ \ dt.$$

One can check that all these examples are inner products. The only non-trivial thing that needs to be checked in this case is the positive definite property:

$$< u, u >= 0 \implies u = 0.$$

**Remark.** The study of $L^2$ spaces is the heart of the definition of a new integral, the Lebesgue integral.

**Definition 1.27** (Norm/length)
We define the **norm** of $v$ by

$$||v|| = \sqrt{< v, v >}.$$

**Remark.** We say that the norm derives from a scalar product.

### §1.4.1 Gram-Schmidt

**Theorem 1.28** (Cauchy-Schwarz inequality)
In an inner product space,

$$| < u, v > | \leq ||u||||v||.$$

*Proof.* We have $F = \mathbb{R}$ or $\mathbb{C}$. Let $t \in F$, then

$$\begin{aligned} 0 \leq ||tu - v||^2 &=< tu - v, tu - v > \\ &= t\bar{t} < u, u > -t < u, v > -\bar{t} < v, u > +||v||^2 \\ &= |t|^2 ||u||^2 - 2\,\mathrm{Re}(t < u, v >) + ||v||^2. \end{aligned}$$

We now choose explicitly
$$t = \frac{\overline{< u, v >}}{||u||^2}.$$
This gives
$$\begin{aligned} 0 &\leq \frac{| < v, u > |^2}{||u||^4} ||u||^2 - 2 \operatorname{Re} \left( \frac{| < u, v > |^2}{||u||^2} \right) + ||v||^2 \\ &\implies 0 \leq ||v||^2 - \frac{| < v, u > |^2}{||u||^2} \\ &\implies | < v, u > |^2 \leq ||v||^2 ||u||^2. \end{aligned}$$
□

**Remark.** It is left as an exercise to show that equality holds iff the vectors $u, v$ are collinear.

**Corollary 1.29** (Triangle inequality)
$||u + v|| \leq ||u|| + ||v||$.

*Proof.*
$$\begin{aligned} ||u + v|| &= < u + v, u + v > \\ &= ||u||^2 + 2 \operatorname{Re} < v, u > + ||v||^2 \\ &\leq ||u||^2 + 2| < u, v > | + ||v||^2 \\ &\leq ||u||^2 + 2||u||||v|| + ||v||^2 \text{ by Cauchy-Scwarz} \\ &= (||u|| + ||v||)^2 . \end{aligned}$$
□

**Remark.** $||.||$ is a norm.

**Definition 1.30** (Orthogonal/orthonormal family)
A set $(e_1, \ldots, e_k)$ of vectors of $V$ is

(i) **orthogonal** if $< e_i, e_j >= 0 \ \forall i \neq j$

(ii) **orthonormal** if $< e_i, e_j >= \delta_{ij} \ \forall i, j$

**Lemma 1.31**
If $(e_1, \ldots, e_k)$ are orthogonal non-zero vectors, then

(i) they are linearly independent.

(ii) Moreover, writing $v = \sum_{j=1}^k \lambda_j e_j$, then
$$\lambda_j = \frac{< v, e_j >}{||e_j||^2}.$$

*Proof.* From the definitions:

(i)

$$\sum_{i=1}^{k} \lambda_i e_i = 0 \implies 0 =< \sum_{i=1}^{k} \lambda_i e_i, e_j >= \sum_{i=1}^{k} \lambda_i < e_i, e_j >= \lambda_j ||e_j||^2 \implies \lambda_j = 0 \forall 1 \leq j \leq k.$$

(ii) If $v = \sum_{i=1}^{k} \lambda_i e_i$, then

$$< v, e_j >=< \sum_{i=1}^{k} \lambda_i e_i, e_j >= \sum_{i=1}^{k} \lambda_i < e_i, e_j >= \lambda_j ||e_j||^2.$$

$$\implies \lambda_j = \frac{< v, e_j >}{||e_j||^2}.$$

□

**Lemma 1.32** (Parseval's identity)
If $V$ is a finite dimensional inner product space, and $(e_1, \ldots, e_n)$ is an *orthonormal* basis, then

$$< u, v >= \sum_{i=1}^{n} < u, e_i > \overline{< v, e_i >}.$$

In particular,

$$||u||^2 = \sum_{i=1}^{n} | < u, e_i > |^2.$$

*Proof.* By the previous lemma with $||e_i|| = 1$,

$$u = \sum_{i=1}^{n} < u, e_i > e_i, \ v = \sum_{i=1}^{n} < v, e_i > e_i.$$

Therefore

$$< u, v >=< \sum_{i=1}^{n} < u, e_i > e_i, \sum_{i=1}^{n} < v, e_i > e_i >= \sum_{i=1}^{n} < u, e_i > \overline{< v, e_i >}.$$

(by orthonormality, and since our form is sesquilinear). Setting $u = v$ recovers our other result. □

**Theorem 1.33** (Gram-Schmidt orthogonalisation process)
Let $V$ be an inner product space. Let $(v_i)_{i \in I}$ such that $I$ is countable, and the $v_i$ are non-zero and linearly independent. Then there exists a family $(r_i)_{i \in I}$ of *orthonormal* vectors such that

$$\forall k \geq 1, \text{span}\{v_1, \ldots, v_k\} = \text{span}\{e_1, \ldots, e_k\}.$$

*Proof.* The proof is an explicit construction of the family $(e_i)_{i\in I}$. We proceed by induction on $k$.
If $k = 1$, then $e_1 = \frac{v_1}{||v_1||}$ $(v_1 \neq 0)$. For the general case, say we have found $(e_1, \ldots, e_k)$ orthonormal with

$$\text{span}\{v_1, \ldots, v_k\} = \text{span}\{e_1, \ldots, e_k\}.$$

Let us compute $e_{k+1}$. We define

$$e_{k+1} = v_{k+1} - \sum_{i=1}^{k} < v_{k+1}, e_i > e_i.$$

Geometrically, we are projecting $v_{k+1}$ onto the plane spanned by $(e_1, \ldots, e_k)$.[drawing]
Note that $e_{k+1} \neq 0$, as otherwise

$$v_{k+1} \in \text{span}\{e_1, \ldots, e_k\} = \text{span}\{v_1, \ldots, v_k\},$$

a contradiction. Now take $j \in \{1, \ldots, k\}$, then

$$\begin{aligned} < e_{k+1}, e_j > &=< v_{k+1} - \sum_{i=1}^{k} < v_{k+1}, e_i > e_i, e_j > \\ &=< v_{k+1}, e_j > - \sum_{i=1}^{k} < v_{k+1}, e_j > \underbrace{< e_i, e_j >}_{=\delta_{ij}} \\ &=< v_{k+1}, e_j > - < v_{k+1}, e_j >= 0. \end{aligned}$$

So $e_{k+1}$ is orthogonal to $e_j$ for all $1 \leq j \leq k$. Thus

$$\text{span}\{v_1, \ldots, v_{k+1}\} = \text{span}\{e_1, \ldots, e_{k+1}\}.$$

Now define (since we know $e_{k+1} \neq 0$),

$$e'_{k+1} = \frac{e_{k+1}}{||e_{k+1}||}.$$

Then $\left(e_1, \ldots, e_k, e'_{k+1}\right)$ is orthonormal □

**Corollary 1.34**

If $V$ is a finite dimensional inner product space, then any orthonormal set of vectors can be extended to an orthonormal basis of $V$.

*Proof.* Pick $(e_1, \ldots, e_k)$ orthonormal. Then they are linearly independed, hence can extend to a basis $(e_1, \ldots, e_k, v_{k+1}, \ldots, v_n)$ of $V$. We apply the Gram-Schmidt algorithm to this basis to get an orthonormal set $S = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ with

$$\text{span}\, S = \text{span}\,(e_1, \ldots, e_k, v_{k+1}, \ldots, v_n) = V.$$

Hence this is an orthonormal basis of $V$. □

**Remark.** If we have a matrix $A \in M_n(\mathbb{R})$ (resp. $A \in M_n(\mathbb{C})$), then the column vectors of $A$ are orthonormal iff

$$A^T A = \mathrm{Id}(\mathbb{R}), A^T \overline{A} \mathrm{Id}(\mathbb{C}).$$

**Definition 1.35** (Orthogonal/unitary matrix)

We say $A \in M_n(\mathbb{R})$ is **orthogonal** if $A^T A = \mathrm{Id}$.
We say $A \in M_n(\mathbb{C})$ is **unitary** if $A^T \overline{A} = \mathrm{Id}$.

**Proposition 1.36**

Let $A \in M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) be non-singular. Then $A$ can be written as $A = RT$ where $T$ is upper triangular, and $R$ is orthogonal (resp. unitary).

*Proof.* Apply Gram-Schmidt to the column vectors of $A$ (the fact $T$ is upper triangular is because the Gram-Schmidt algorithm only uses the previous vectors in the set). □

## §1.5 Orthogonal complement and projection

**Definition 1.37** (Orthogonal direct sum)

Let $V$ be an inner product space, and $V_1, V_2 \leq V$. We say that $V$ is the **orthogonal direct sum** of $V_1$ and $V_2$ if

1. $V = v_1 \oplus V_2$.
2. $\forall (v_1, v_2) \in V_1 \times V_2 : < v_1, v_2 >= 0$.

**Remark.** The directness of the sum in 1. is redundant (it follows from 2.)

**Definition 1.38** (Orthogonal complement)

For $V$ an inner product space and $W \leq V$, the **orthogonal complement** of $W$ in $V$ is

$$W^{\perp} = \{v \in V : \ \forall w \in W, \ < v, w >= 0\}.$$