IB Linear Algebra (from lecture 18)

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

February 2, 2022

These are my notes for the IB course Linear Algebra, which was lectured in Lent 2022 at Cambridge by Prof. P. Raphael. These notes are written in LaTeX for my own revision purposes. Any suggestions or feedback is welcome.

Contents

§1 Bilinear forms revisited

Lemma 1.0.1

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
$$P = [\text{Id}]_{B,B'}.$$
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. $\qquad\square$

Definition (Congruent matrices)

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

Remark. This defines an equivalence relation.

Definition (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. φ is symmetric iff $[\varphi]_B$ is symmetric in *any* basis B.

3. To be able to represent φ by a diagonal matrix in some basis B, it is necessary that φ is symmetric:
$$P^T A P = D = D^T = P A^T P^T \implies A = A^T \implies \varphi \text{ is symmetric.}$$

Definition (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 $B = (e_i)_{1 \le i \le n}$ 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 φ) Therefore we essentially have

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

Remark. We can note that

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

So this representation of A is not necessarily unique.

Proposition 1.0.2

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 ψ 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:

- φ is a bilinear form

- φ is symmetric

- $\varphi(u, u) = \psi(u, u) = Q(u)$.

This concludes the proof of existence.
Proof of uniqueness
Let φ 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

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

$\square$

Theorem 1.0.3 (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 (φ 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 $\langle e_1 \rangle$:

$$U = \langle e_1 \rangle^{\perp}$$
$$= \{v \in V : \varphi(e_1, v) = 0\}$$
$$= \ker \theta \text{ where } \theta : V \to F \text{ is given by } \theta(v) = \varphi(e_1, v).$$

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

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

We now claim that $U + \langle e_1 \rangle = U \oplus \langle e_1 \rangle$. Indeed,

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

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

$$\implies U + \langle e_1 \rangle = U \oplus \langle e_1 \rangle.$$

Therefore $V = U \oplus \langle e_1 \rangle$, and pick a basis $B' = (e_2, \dots, e_n)$ such that $(e_1, e_2, \dots, e_n)$ 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 φ onto U. Now we apply the induction hypothesis to find a basis $(e_1', \dots, e_n';)$ of V such that $[\varphi|U]_B$ is diagonal. So

$$\hat{B} = (e_1, e_2', \dots, e_n')$$

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

$\square$

Example

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

$$Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3$$

$$= x^T A x \text{ where } A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

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

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

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

Corollary 1.0.4

Let $F = \mathbb{C}$, and φ 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 $[\varphi]_B$ can be expressed in the form

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

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

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

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

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

$\square$

Corollary 1.0.5

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. $\square$

Corollary 1.0.6

Let $F = \mathbb{R}$. Let φ 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. $\square$

Definition (Signature of a quadratic form)

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

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

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

Definition (Positive definite quadratic/bilinear form)

Let φ be a symmetric bilinear form on a real vector space V. We say that

  (i) φ is positive definite if

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

  (ii) φ is positive semidefinite if

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

  (iii) φ is negative definite if

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

  (iv) φ is negative semidefinite if

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

Example

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 $1 \leq p \leq n$.

§1.1 Sylvester's law

Theorem 1.1.1 (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 φ is positive definite. Say $B = (v_1, \ldots, v_n)$ 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 $X = < v_1, \ldots, v_p$. Then φ 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 φ is positive definite on another subspace X. Let

$$X = \langle v_1, \ldots, v_p \rangle, Y = \langle v_{p+1}, \ldots, v_n \rangle.$$

Then arguing as above, looking at $[\varphi]_B$ we know that φ 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

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

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

Definition (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.

§1.2 Sesquilinear forms

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

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

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** (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 1.2.1**
>
> 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.   □

§1.3 Hermitian forms

We will now generalise symmetric bilinear forms to sesquilinear forms.

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

Remark. If φ 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 1.3.1

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 φ 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

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

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

Proposition 1.3.2 (Polarisation identity for sesquilinear forms)

A Hermitian form φ 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. $\square$

Theorem 1.3.3 (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 φ is positive definite, and that q is the maximal dimension of a subspace on which φ is negative definite. Then by our geometric characterisation we are done. $\qquad\qquad\square$

§1.4 Skew symmetric forms

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

> **Definition** (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 1.4.1** (Sylvester form of skew symmetric matrices)
>
> Let V be a finite n-dimensional vector space over $\mathbb{R}$. Let φ be a skew symmetric bilinear form over V. Then there exists a basis B of V such that
>
> $$B = (v_1, w_1, v_2, w_2, \ldots v_m, w_m, v_{2m+1}, v_{2m+2}, \ldots, v_n).$$

> **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.

□

> **Corollary 1.4.2**
>
> Skew symmetric matrices have an even rank.

§2 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** (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 φ on V. We use the notation
>
> $$\langle u, v \rangle = \varphi(u, v).$$
>
> We call V a real (resp. complex) inner product space.

> **Example**
>
> Here are several examples of inner product spaces.
>
> 1. $V = \mathbb{R}^n$, where we define
>
> $$\langle x, y \rangle = \langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{i=1}^{n} x_i y_i.$$
>
> 2. $V = \mathbb{C}^n$, $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$.
> 3. $V = C^0([0,1], \mathbb{C})$. We define
>
> $$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)} \, dt.$$
>
> 4. We can fix a weight $\omega : [0,1] \to \mathbb{R}_+^*$ and define on $V = C^0([0,1], \mathbb{C})$
>
> $$\langle f, g \rangle = \int_0^1 \omega(t) f(t)\overline{g(t)} \, dt.$$
>
> 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:
>
> $$\langle u, u \rangle = 0 \implies u = 0.$$

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

Definition (Norm/length)

We define the **norm** of v by
$$||v|| = \sqrt{\langle v, v \rangle}.$$

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

Theorem 2.0.1 (Cauchy-Schwarz inequality)

In an inner product space,
$$|\langle u, v \rangle| \leq ||u|| ||v||.$$

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

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

We now choose explicitly
$$t = \frac{\overline{\langle u, v \rangle}}{||u||^2}.$$

This gives

$$0 \leq \frac{|\langle v, u \rangle|^2}{||u||^4}||u||^2 - 2\operatorname{Re}\left(\frac{|\langle u, v \rangle|^2}{||u||^2}\right) + ||v||^2$$

$$\implies 0 \leq ||v||^2 - \frac{|\langle v, u \rangle|^2}{||u||^2}$$

$$\implies |\langle v, u \rangle|^2 \leq ||v||^2 ||u||^2.$$

$\square$

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

Corollary 2.0.2 (Triangle inequality)

$||u + v|| \leq ||u|| + ||v||.$

Proof.

$$\begin{aligned}
||u + v|| &= \langle u + v, u + v \rangle \\
&= ||u||^2 + 2\operatorname{Re}\langle v, u \rangle + ||v||^2 \\
&\leq ||u||^2 + 2|\langle u, v \rangle| + ||v||^2 \\
&\leq ||u||^2 + 2||u|| ||v|| + ||v||^2 \text{ by Cauchy-Scwarz} \\
&= (||u|| + ||v||)^2.
\end{aligned}$$

$\square$

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

Definition (Orthogonal/orthonormal family)

A set $(e_1, \ldots, e_k)$ of vectors of V is

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

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

Lemma 2.0.3

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

$$\lambda_j = \frac{\langle v, e_j \rangle}{||e_j||^2}.$$

Proof. From the definitions:

(i) Let us take $\sum_{i=1}^{k} \lambda_i e_i = 0$. Then

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

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

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

$$\implies \lambda_j = \frac{\langle v, e_j \rangle}{||e_j||^2}.$$

$\square$

Lemma 2.0.4 (Parseval's identity)

If V is a finite dimensional inner product space, and $(e_1, \ldots, e_n)$ is an *orthonormal* basis, then

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

In particular,

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

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

$$u = \sum_{i=1}^{n} \langle u, e_i \rangle e_i, \ \ v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i.$$

Therefore

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

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

Having proved these facts, we now turn our attention to an important result which will lead to us understanding bases of inner product spaces.

§2.1 Gram-Schmidt orthogonalisation

Theorem 2.1.1 (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 $(e_i)_{i \in I}$ of *orthonormal* vectors such that

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

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

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

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

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

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

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

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

$$\langle e_{k+1}, e_j \rangle = \langle \left(v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i \right), e_j \rangle$$

$$= \langle v_{k+1}, e_j \rangle - \sum_{i=1}^{k} \langle v_{k+1}, e_j \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}}$$

$$= \langle v_{k+1}, e_j \rangle - \langle v_{k+1}, e_j \rangle = 0.$$

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 $\qquad\square$

Corollary 2.1.2

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. $\qquad\square$

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

Definition (Orthogonal/unitary matrix)

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

Proposition 2.1.3

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). $\qquad\square$

§2.2 Orthogonal complement and projection

Definition (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. $\forall (v_1, v_2) \in V_1 \times V_2 : \langle v_1, v_2 \rangle = 0$.

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

Definition (Orthogonal complement)

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

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

Lemma 2.2.1

Let V be a finite dimensional inner product space, and $W \leq V$. Then

$$V = W \oplus W^\perp.$$

Proof. If $w \in W, w \in W^\perp$ then

$$||w||^2 = \langle w, w \rangle = 0 \implies w = 0.$$

We need to show $V = W + W^\perp$. Let $(e_1, \ldots, e_k)$ be an orthonormal basis of W. Extend it to an orthonormal basis of V: $(e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$. Observe that $(e_{k+1}, \ldots, e_n) \in W^\perp$. Therefore $V = W + W^\perp$. $\qquad\square$

Remark. To make sense of this in infinite dimensions we would have to introduce a notion of *topology*.

Definition (Projection map)

Suppose we can decompose a vector space V as $V = U \oplus W$. (U is a **complement** of W in V). Now define

$$\pi : V \to W, \pi(v) = \pi(u + w) = w.$$

We have that π is well-defined, linear, and $\pi^2 = \pi$. We say that π is the **projection operator** onto W.

Remark. $\mathrm{Id} - \pi$ is the projection onto U. We can make the project map very explicit when $U = W^\perp$, and V is an inner product space.

Lemma 2.2.2

Let V be an inner product space. Let $W \leq V$ be finite dimensional. Let $(e_1, \ldots, e_k)$ be an orthonormal basis of W. Then

  (i) $\pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i \ \forall v \in V$

  (ii) For all $v \in V$, for all $w \in W$, we have

$$||v - \pi(v)|| \leq ||v - w||,$$

  with equality iff $w = \pi(v)$. ($\pi(v)$ is the point in W which is closest to v.)

Proof. We prove both statements.

  (i) We define $\pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i$. Since $W = \operatorname{span}\{e_1, \ldots, e_k\}$, $\pi(v) \in W$. We write
$$v = \underbrace{v - \pi(v)}_{\in W^{\perp}} + \underbrace{\pi(v)}_{\in W}.$$

  Since we have claimed that $v - \pi(v) \in W^{\perp}$, we need to show for all $w \in W$, $\langle v - \pi(v), w \rangle = 0$. This happens iff for all $1 \leq j \leq k$,

$$\langle v - \pi(v), w \rangle = 0.$$

  But

$$\langle v - \pi(v), e_j \rangle = \langle v, e_j \rangle - \langle \sum_{i=1}^{k} \langle v, e_i \rangle e_i, e_j \rangle$$
$$= \langle v, e_j \rangle - \sum_{i=1}^{k} \langle v, e_i \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}}$$
$$= \langle v, e_j \rangle - \langle v, e_j \rangle$$
$$= 0$$

  So our claim is true and hence $V = W + W^{\perp}$. Since $W \cap W^{\perp} = \{0\}$, we have $V = W \oplus W^{\perp}$.

  (ii) Let $v \in V$, $w \in W$, and let us compute

$$||v - w||^2 = ||\underbrace{v - \pi(v)}_{\in W^{\perp}} + \underbrace{\pi(v) - w}_{\in W}||^2$$
$$= ||v - \pi(v)||^2 + ||\pi(v) - w||^2$$
$$\geq ||v - \pi(w)||^2$$

  with equality iff $w = \pi(v)$.

  $\square$

§2.3 The adjoint map

> **Definition** (Adjoint map)
>
> Let V, W be finite dimensional inner product spaces and let $\alpha \in L(V, W)$. Then there is a unique linear map $\alpha^* : W \to V$ such that for all $(v, w) \in V \times W$,
>
> $$\langle \alpha(v), w \rangle = \langle v, \alpha^*(w) \rangle.$$
>
> Moreover, if B is an orthonormal basis of V and C is an orthonormal basis of W, then
> $$[\alpha^*]_{C,B} = (\overline{[\alpha]_{B,C}})^T.$$

> **Proof.** We just do a brute force computation. Let $B = \{v_1, \ldots, v_n\}$ and $C = \{\omega_1, \ldots, \omega_m\}$. Let $A = [\alpha]_{B,C} = (a_{ij})$. Let C be the matrix be defined by $c_{ij} = \overline{a_{ij}}$. We compute
>
> $$\langle \alpha(\sum \lambda_i v_i), \sum \mu_i \omega_i \rangle = \langle \sum_{i,k} a_{ki} \omega_k, \sum \mu_j \omega_j \rangle$$
>
> $$= \sum_{i,j} \lambda_i a_{ji} \overline{\mu_j} \text{ by orthonormality of } \omega_i.$$
>
> Then we also have
>
> $$\langle \sum_i \lambda_i v_i, \alpha^* \left(\sum_j \mu_J \omega_j \right) \rangle = \langle \sum_i \lambda_i v_i, \sum_{j,k} \mu_j c_{kj} v_k \rangle$$
>
> $$= \sum_{i,j} \lambda_i \overline{c_{ij} \mu_j}$$
>
> Therefore $\overline{c_{ij}} = a_{ij}$ so we have existence, and uniqueness also follows as our two expressions are equal iff $\omega_j = a_{ji}$. $\qquad \square$

Remark. We denote $\overline{A}^T = A^\dagger$. We are using the same notation α^* for both the adjoint and the dual map. We can check that if V, W are real product inner spaces and $\alpha \in L(V, W)$, then we can define

$$\psi_{R,V} : V \to V^*, v \to \langle ., v \rangle$$

and

$$\psi_{R,W} : W \to W^*, w \to \langle ., w \rangle.$$

we can then find that the adjoint of α is given by:

$$W \underset{\psi_{R,W}}{\to} W^* \underset{\text{dual of } \alpha}{\to} V^* \underset{\psi_{R,V}^{-1}}{\to} V.$$

§2.4 Self-adjoint maps and isometries

> **Proposition 2.4.1**
>
> Let V be an inner product space, and $\alpha \in L(V)$. Let $\alpha^* \in L(V)$ be the adjoint map. Then the following are equivalent:
>
> 1. For all $(v, w) \in V \times V$, $\langle \alpha v, w \rangle = \langle (v), \alpha(w) \rangle$.

2. $\alpha = \alpha^*$

We call the map α **self-adjoint**. In $\mathbb{R}$ we call it symmetric, and in $\mathbb{C}$ we call it Hermitian.

Proposition 2.4.2

Let V be an inner product space, and $\alpha \in L(V)$. Let $\alpha^* \in L(V)$ be the adjoint map. Then the following are equivalent:

1. For all $(v, w) \in V \times W$, $\langle \alpha(v), \alpha(w) \rangle = \langle v, w \rangle$.

2. $\alpha = \alpha^{*-1}$.

We call the map an **isometry**. In $\mathbb{R}$ we call it orthogonal, and in $\mathbb{C}$ we call it unitary.

Proof. Prove both directions.
1 $\implies$ 2:
$\langle \alpha(v), \alpha(w) \rangle = \langle v, w \rangle$ Now take $v = w$. $\implies ||\alpha(v)||^2 = ||v||^2$. Therefore $\ker \alpha = \{0\}$, and hence α is bijective (since we are in finite dimension). So α^{-1} is well defined. Then for all $(v, w) \in V \times V$,

$$\langle v, \alpha^*(w) \rangle = \langle \alpha(v), w \rangle = \langle \alpha(v), \alpha(\alpha^{-1}(w)) \rangle = \langle v, \alpha^{-1}(w)w \rangle.$$

We have show that for all v and w,

$$\langle v, (\alpha^* - \alpha^{-1})w \rangle = 0.$$

Thus if we choose $v = (\alpha^* - \alpha^{-1})w$, $||(\alpha^* - \alpha^{-1})w||^2 = 0$. So for all w, $(\alpha^* - \alpha^{-1})w = 0$. So we can conclude that $\alpha^* = \alpha^{-1}$.
2 $\implies$ 1:
If we have that $\alpha = \alpha^{*-1}$, then

$$\langle \alpha(v), \alpha(w) \rangle = \langle v, \alpha(\alpha^*(w)) \rangle = \langle v, \alpha(\alpha^{-1}(w)) \rangle = \langle v, w \rangle.$$

$\square$

Remark. Using the polarisation identity, we can show that if α is an isometry if and only if for all $v \in V$, $||\alpha(v)|| = ||v||$ (it preserves the norm).

Lemma 2.4.3

Let V be a finite dimensional real (complex) inner product space. Then given an $\alpha \in L(V)$, α:

(i) is self-adjoint iff for *any* orthonormal basis B of V, $[\alpha]_B$ is symmetric (Hermitian).

(ii) is an isomety iff for *any* orthonormal basis of V, $[\alpha]_B$ is orthogonal (unitary).

Proof. If B is an orthonormal basis of V, we have that $[\alpha^*]_B = (\overline{[\alpha]_B})^T$.
In (i) we have that $(\overline{[\alpha]_B})^T = [\alpha]_B$, and in (ii) we have that $(\overline{[\alpha]_B})^T = [\alpha]_B^{-1}$. $\square$

Definition (Orthogonal/unitary group)

Let V be a finite dimensional inner product space.
In $\mathbb{R}$ we define the **orthogonal group** of V by

$$O(V) = \{\alpha \in L(V) : \ \alpha \text{ is an isometry}\}.$$

In $\mathbb{C}$ we define the **unitary group** of V by

$$U(V) = \{\alpha \in L(V) : \ \alpha \text{ is an isometry}\}.$$

§2.5 Spectral theory for self-adjoint maps and isometries

Lemma 2.5.1

Let V be a finite dimensional inner product space. Let $\alpha \in L(V)$ be self-adjoint. Then

1. α has real eigenvalues.

2. The eigenvalues of α wrt. different eigenvalues are orthogonal.

Proof.   1. Let $\lambda \in \mathbb{C}$, $v \in V \setminus \{0\}$. Suppose $\alpha(v) = \lambda v$. Then

$$\begin{aligned}
\lambda ||v||^2 &= \langle \lambda v, v \rangle \\
&= \langle \alpha(v), v \rangle \\
&= \langle v, \alpha(v) \rangle = \langle v, \lambda v \rangle \\
&= \overline{\lambda} ||v^2|| \\
&\implies (\lambda - \overline{\lambda}) ||v||^2 = 0 \\
&\implies \lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}.
\end{aligned}$$

2. Let us consider two eigenvectors v, w for different eigenvalues λ, μ respectively. Then

$$\begin{aligned}
\lambda \langle v, w \rangle = \langle \lambda v, w \rangle &= \langle \alpha(v), w \rangle \\
&= \langle v, \alpha(w) \rangle = \langle v, \mu w \rangle \\
&= \overline{\mu} \langle v, w \rangle = \mu \langle v, w \rangle.
\end{aligned}$$

Since $\lambda \neq \mu$, we have $\langle v, w \rangle = 0$.

$\square$

Theorem 2.5.2 (Spectral theorem for self-adjoint operators)

Let V be a finite dimensional inner product space. Let $\alpha \in L(V)$ be self-adjoint. Then V has an orthonormal basis of eigenvectors of α. This implies that α is diagonalisable.

Proof. Let $F = \mathbb{R}$ or $\mathbb{C}$. We argue on induction on the dimension of V. The $n = 1$

case is trivial. Suppose it holds for $n-1$ dimensions. Say $A = [\alpha]_B$ with respect to the fundamental basis B. By the Fundamental Theorem of Algebra, we know that the characteristic polynomial $\chi_A(\lambda)$ has a root. This root is an eigenvalue of α, and so it is real. Let us call this real eigenvalue $\lambda \in \mathbb{R}$. Pick an eigenvector $v_1 \in V \setminus \{0\}$ such that

$$\alpha(v_1) = \lambda v_1, \|v_1\| = 1.$$

Let $U = \langle v_1 \rangle^\perp \leq V/$ The key observation is that $\alpha(U) \leq U$. Indeed, let $u \in U$, then

$$\langle \alpha(u), v_1 \rangle = \langle u, \alpha(v_1) \rangle = \langle u, \lambda v_1 \rangle = \lambda \langle u, v_1 \rangle = 0.$$

Therefore $\alpha(u) \perp v_1 \implies \alpha(u) \in U$. Hence we may consider $\alpha|_U \in L(U)$ which is also self-adjoint, and

$$n = \dim V = \dim U + 1.$$

Therefore $\dim U \leq n - 1$, so we apply our induction hypothesis: there exist an orthonormal basis $(v_2, \ldots v_n)$ of eigenvectors of $\alpha|_U$. Since $V = \langle v_1 \rangle \oplus U$, we have that $(v_1, v_2, \ldots, v_n)$ is an orthonormal basis of V. $\qquad\square$

Corollary 2.5.3

If V is a finite dimensional inner product space and $\alpha \in L(V)$ is self-adjoint, then V is the orthogonal direct sum of all the eigenspaces of α.

Proof. Immediate. $\qquad\square$

Lemma 2.5.4

Let V be a *complex* inner product space. Let $\alpha \in L(V)$ be unitary ($\alpha = \alpha^{*-1}$). Then

  (i) All eigenvalues of α lie on the unit circle.

 (ii) Eigenvectors corresponding to different eigenvectors are orthogonal.

Proof.   (i) Let $\lambda \in \mathbb{C}$ and $v \in V \setminus \{0\}$ such that $\alpha(v) = \lambda v$. Note that $\lambda \neq 0$, since we showed that unitary maps are invertible ($\ker \alpha = 0$). We compute

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle \alpha(v), v \rangle = \langle v, \alpha^*(v) \rangle = \langle v, \alpha^{-1}(v) \rangle.$$

Applying α^{-1}, we have $v = \lambda \alpha^{-1}(v)$. Thus

$$\lambda \langle v, v \rangle = \langle v, \alpha^{-1}(v) \rangle = \langle v, \frac{1}{\lambda} v \rangle = \frac{1}{\overline{\lambda}} \langle v, v \rangle.$$

So

$$(\lambda \overline{\lambda} - 1) \|v\|^2 = 0 \implies |\lambda| = 1.$$

 (ii) Let us take two evecs v, w with distinct eigenvalues λ, μ respectively. Then

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle \alpha(v), w \rangle = \langle v, \alpha^*(w) \rangle = \langle v, \alpha^{-1}(w) \rangle = \langle v, \frac{1}{\mu} w \rangle = \frac{1}{\overline{\mu}} \langle v, w \rangle.$$

Thus $(\lambda - \mu)\langle v, w \rangle = 0$. Since $\lambda \neq \mu$, $\langle v, w \rangle = 0$.

$\square$

Theorem 2.5.5 (Spectral theorem for unitary maps)

Let V be a finite n-dimensional *complex* inner product space. Let $\alpha \in L(V)$ be unitary. Then V has an orthonormal basis consisting of eigenvectors of α.

Proof. This is very similar to the self-adjoint case. Take $A = [\alpha]_B$, where B is an orthonormal basis. Let $\chi_A(\lambda)$ be our characteristic polynomial. It has a complex root, say $\lambda \in \mathbb{C}$. Fix $v_1 \in V \setminus \{0\}$ with $\alpha(v_1) = \lambda v_1$ and $||v_1|| = 1$. Letting $U = \langle v_1 \rangle^\perp$ as in the previous proof, we can show that $\alpha(U) \leq U$. Hence $\alpha|_U \in L(U)$, which is unitary and $\dim U = n - 1$. By the inductive hypothesis, we can find an orthonormal basis of U composed of evecs of $\alpha|_U$ and then combine it with v_1 to get an orthonormal basis of V. $\square$

Remark. We used the complex structure to assert that there is a complex eigenvalue. In general, a real valued orthonormal matrix A with $AA^T = \text{Id}$ CANNOT be diagonalised over $\mathbb{R}$.

Example (Rotation in $\mathbb{R}^2$)

Consider the matrix
$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

This is clearly an orthogonal matrix. However
$$\chi_A(\lambda) = (\cos(\theta) - \lambda)^2 + \sin^2(\theta).$$

The roots of this are $\lambda = e^{\pm i\theta}$, so the roots are not generally real.

§2.6 Applications to bilinear forms

Corollary 2.6.1

Let $A \in M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) be a symmetric (resp. Hermitian) matrix. Then there is an orthogonal (resp. unitary) matrix B such that $P^T A P$ (resp. $P^\dagger A P$) is diagonal with real-valued entries.

Proof. Let $F = \mathbb{R}$ WLOG. Let $\langle \, , \, \rangle$ be the standard inner product over $\mathbb{R}^n$. Then $A \in L(F^n)$ is self-adjoint, hence we can find an orthonormal (for the standard inner product) basis $\mathcal{B}$ of F^n such that A is diagonal in that basis $\mathcal{B} = (v_1, \ldots, v_n)$. Let

$$P = (v_1 | \ldots | v_n) \iff P \text{ is unitary} \iff P^T P = \text{Id}.$$

We know that $P^{-1} A P = D$, where D is diagonal. Then

$$P^T A P = P^{-1} A P = D.$$

Its entries are real since the elements of P on the diagonal, $(\lambda_i)_{1 \leq i \leq n}$ are real

(eigenvalues of a symmetric operator). $\qquad\qquad\qquad\qquad\qquad\qquad$ □

Corollary 2.6.2

Let V be a finite n-dimensional real (resp. complex) inner product space. Let $\varphi : V \times V \to F$ be a symmetric (resp. Hermitian) form. Then there exists an orthonormal basis of V such that φ in this basis is represented by a diagonal matrix.

Proof. Let $B = (v_1, \ldots, v_n)$ be an orthonormal basis of V. Let $A = [\varphi]_B$. We know $A = A^T$ (resp. $A^\dagger$), and hence there is an orthogonal (unitary) matrix P such that $P^T A P$ ($P^\dagger A P$) is diagonal, say D. Let (v_i) be the ith row of P^T ($P^\dagger$), then $(v_1, \ldots, v_n)$ is an orthonormal basis B of V and

$$[\varphi]_B = D.$$

This is using the change of basis formula for bilinear forms. $\qquad\qquad\qquad$ □

Remark. In this process, the diagonal entries of $P^T A P$ ($P^\dagger A P$) are the eigenvalues of A. Moreover, the signature $s(\varphi) = $ (no. positive evals of A)-(no. negative evals of A).

Corollary 2.6.3 (Simultaneous diagonalisation)

Let V be a finite dimensional real (resp. complex) vector space. Let $\varphi, \psi : V \times V \to F$ be symmetric (resp. Hermitian) bilinear forms. Assume that φ is positive definite. Then there exists a basis $(v_1, \ldots, v_n)$ of V with respect to which *both* forms are represented by a diagonal matrix.

Proof. The key point is that φ is positive definite, which means that V equipped with φ is a finite dimensional inner product space. We define

$$\langle u, v \rangle = \varphi(u, v).$$

Hence there exists an orthonormal (for the φ-induced scalar product) basis of V in which ψ is represented by a diagonal matrix (since it's symmetric). Observe that φ in this basis is just the identity matrix (since $\varphi(v_i, v_j) = \langle v_i, v_j \rangle = \delta_{ij}$). So we are done. $\qquad\qquad$ □

Corollary 2.6.4 (Simultaneous diagonalisation for matrices)

Let $A, B \in M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) that are symmetric (resp. Hermitian). Assume for all $x \neq 0$,

$$\overline{x}^T A x > 0 \qquad\qquad\qquad\qquad\qquad\qquad (*)$$

Then there exists $Q \in M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) such that both matrices

$$Q^T A Q, Q^T B Q$$

are diagonal.

Proof. $(*)$ just expresses the fact that A induces a positive definite symmetric bilinear form by $\bar{x}^T A x = \varphi(x, x)$. Then the proof follows from the previous result. $\quad\square$

Remark. End of notes!