Tutorial 6 Inner Product Spaces and their Operators

Def. 6.1. An <u>inner product</u> is a function $\langle\ ,\ \rangle : V\times V \to \mathbb{F}$, where $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, s.t.

(IP1). Sesquilinear $\langle ax+y, z\rangle = a\langle x,z\rangle + \langle y,z\rangle$. $\forall x,y,z\in V,\ a\in\mathbb{F}$.

Linear in first and conjugate linear in second $\langle x, ay+z\rangle = \bar{a}\langle x,y\rangle + \langle x,z\rangle$ conjugate : $\overline{a+bi} = a-bi$

(IP2). Conjugate symmetric $\langle x,y\rangle = \overline{\langle y,x\rangle}$ $\forall x,y\in V$.

(IP3). positive definite $\langle x,x\rangle \geq 0$ with $=$ iff $x=0$. $\forall x\in V$.

Remark. (i) (IP2) implies $\langle x,x\rangle\in\mathbb{R}$, so (IP3) make sense as $\mathbb{R}$ has a natural order $\geq$.

(ii) When $F=\mathbb{R}$, a function $\langle\ ,\ \rangle$ satisfying (IP1) is called a <u>bilinear form</u>.

a function satisfying (IP1)+(IP2) is called a <u>symmetric bilinear form</u>.

So an inner product is a positive definite symmetric bilinear form.

(a function $\langle\ ,\ \rangle : V\times V\to\mathbb{F}$ is called a <u>form</u>)

They are all very important objects in linear algebra, and appear everywhere in geometry.

When $\mathbb{F}=\mathbb{C}$, we also study <u>Hermitian form</u> : A function $H(\ ,\ ): V\times V\to\mathbb{F}$ satisfying (IP1)+(IP2). They are heavily studied in analysis, especially functional analysis.

(iii). The dictionary :

$\mathbb{F}=\mathbb{R}$ Symmetric bilinear form $\langle Tv,w\rangle = \langle v,Tw\rangle \longleftrightarrow$ symmetry matrix $A^T=A \longleftarrow A=[T]_\beta$.

$\mathbb{F}=\mathbb{C}$ Hermitian form $\langle Tv,w\rangle = \langle v,Tw\rangle \longleftrightarrow$ Hermitian matrix $A^H=A$.

orthonormal basis

(iv) For every symmetry bilinear form B we can associate a <u>quadratic form</u> $q: V\to\mathbb{F}$ defined by $q(x) = B(x,x)$. A <u>quadratic form</u> is a function $q: V\to\mathbb{F}$ satisfying

(QF1). $q(ax) = a^2 q(x)$ $\forall x\in V,\ a\in\mathbb{F}$.

(QF2). $q(x+y)-q(x)-q(y)$ is bilinear $\forall x,y\in V$.

Eg. $q(x) = x^TQx$, where Q is a matrix corresponding to q if a basis is chosen.

Eg. $q(x,y) = 4x^2+2xy-3y^2 = (x\ \ y)\begin{pmatrix}4 & 1\\ 1 & -3\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ so we may just study $Q=\begin{pmatrix}4 & 1\\ 1 & -3\end{pmatrix}$.

Classification of quadratic form gives classification of quadrics in $\mathbb{R}^n$.

When $n=3$, we may classify all <u>conic sections</u> (圓錐曲線) in high school!

Eg. $q(x,y) = x^2+4xy+y^2$

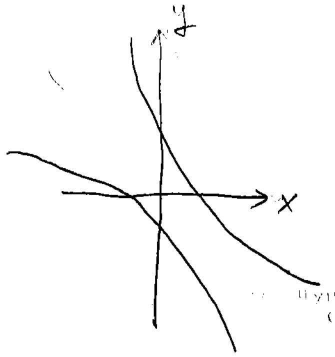
$$= (x\ \ y)\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

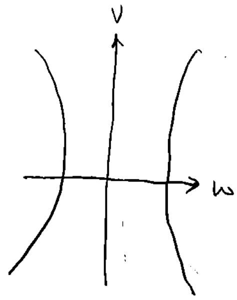
change of basis.

$$= (x\ \ y)\, P\begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}P^{-1}\begin{pmatrix}x\\ y\end{pmatrix} \qquad P = \frac{1}{2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} = \text{rotation by } \frac{\pi}{4}$$

$$= (w\ \ v)\begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}w\\ v\end{pmatrix} \qquad \text{where } (w,v) = (x,y)\,P.$$

$$= 3w^2 - v^2 \qquad\qquad = \left(\tfrac{1}{\sqrt{2}}(x+y),\ \tfrac{1}{\sqrt{2}}(y-x)\right)$$





(v) A bilinear form $B: V\times V \to \mathbb{F}$ is equivalent to a linear map $\tilde{B}: V\otimes V \to \mathbb{F}$.

tensor product.

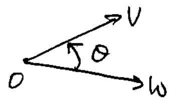
In this language, we can study multilinear maps: $f: V^{\otimes n} \to \mathbb{F}$.

This is another generalization of this topic.

* all remarks above are non-examinable except (i).

Def. 6.2. An <u>inner product space</u> is a vector space with a inner product on it.

Remark.(i) Recall $\cos\theta = \dfrac{\langle v,w\rangle}{\sqrt{\langle v,v\rangle\langle w,w\rangle}}$



, $\langle v,w\rangle = v^T w$ here.

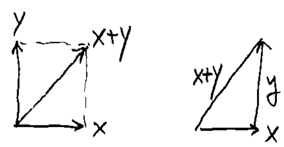
$\|v\| = \sqrt{\langle v,v\rangle}$ length of v.

So on inner product space we may talk above "angle" and "length".

(ii) Pythagorean theorem (勾股定理)

$\langle x,y\rangle = 0$

$\|x\|^2 + \|y\|^2 = \|x+y\|^2$ where $x \perp y$



This works for arbitrary dimension!

$\sum_{i=1}^{n}\|x_i\|^2 = \|\sum_{i=1}^{n} x_i\|^2$ if $\langle x_i, x_j\rangle = 0$ for all $i\neq j$.

(iii). Cauchy inequality (柯西不等式)

$$(x_1^2+x_2^2)(y_1^2+y_2^2)\geqslant(x_1y_1+x_2y_2)^2 \iff \|x\|\|y\|\geqslant\|\langle x,y\rangle\| \qquad x=(x_1,x_2)\quad y=(y_1,y_2).$$

Again this works for arbitrary dimension! <u>Cauchy-Schwarz inequality</u>.

$$\|x\|\,\|y\|\geqslant\|\langle x,y\rangle\| \qquad x=(x_1,x_2,\cdots,x_n)\quad y=(y_1,\cdots,y_n)$$

$$\Updownarrow$$

$$\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)=\left(\sum_{i=1}^n x_iy_i\right)^2$$

(iv). matrix transpose Let $T: V\to V$ be a linear operator.

When we define $\langle v,w\rangle=v^*w$ the <u>standard inner product</u>.

$$\langle Tv,w\rangle=\langle v,T^*w\rangle \iff (Tv)^*w=v^*T^*w$$

where $*$ is the transpose when $\mathbb{F}=\mathbb{R}$

hermitian where $\mathbb{F}=\mathbb{C}$.

$\quad$ $\underset{\text{conjugate + transpose}}{\parallel}$

T^* is called the <u>adjoint operator</u> of T.

So we have notion of matrix transpose and matrix Hermitian in our language.

$*$ not all remarks above are examinable, depends on the progress of lecture.

Eg. Frobenius inner product: $\langle A,B\rangle=tr(AB^H)$. gives a inner product structure on $M_n(\mathbb{C})$.

<u>Euclidean space</u>: $\mathbb{R}^n$ with an inner product, $\mathbb{F}=\mathbb{R}$.

Inner product gives additional structure to a vector space, and such restrictions give many astonishing consequences.

Thm 6.3 Let V be an inner product space, $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Suppose $\dim V<\infty$. Then

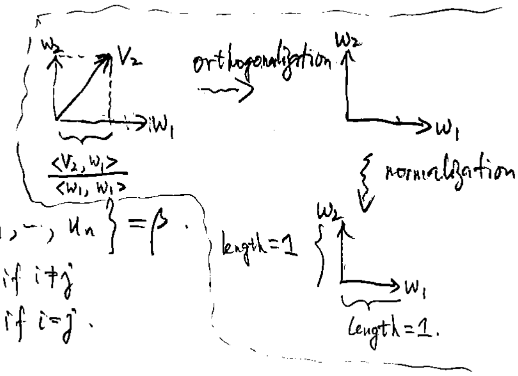
(i). There exist an orthonormal basis, via <u>Gram-Schmidt orthonormalisation</u> process.

$$w_1=v_1$$
$$w_2=v_2-\frac{\langle w_1,v_2\rangle}{\langle w_1,w_1\rangle}w_1$$
$$\vdots$$
$$w_k=v_k-\sum_{i=1}^{k-1}\frac{\langle w_i,v_k\rangle}{\langle w_i,w_i\rangle}w_i$$

orthogonalization

$$u_i=\frac{w_i}{\|w_i\|},\quad \|w_i\|=\sqrt{\langle w_i,w_i\rangle}.$$

normalization



Any basis $\{v_1,\cdots,v_n\}\xrightarrow{\text{orthonormalization}}$ <u>Orthonormal</u> basis $\{u_1,\cdots,u_n\}=\beta$.

$$\langle u_i,u_j\rangle=\delta_{ij}=\begin{cases}0 & \text{if } i\neq j\\ 1 & \text{if } i=j.\end{cases}$$

orthonormal basis

$$v=\sum_{i=1}^n\underbrace{\langle v,u_i\rangle}_{\text{scalar}}u_i \quad \text{is the linear combination for any } v\in V.$$

$$\Rightarrow\left([T]_\beta\right)_{ij}=\langle u_i,T(u_j)\rangle.$$

(ii). Every subspace $U \subseteq V$ has an orthogonal complement

$$U^\perp := \{ v \in V \mid \langle v, u \rangle = 0 \text{ for all } u \in U \}.$$

So that $V = U \oplus U^\perp$

This is direct result from (i): the existence of orthonormal basis.

Moreover, ~~one may~~ extend an orthonormal basis in U to a orthonormal basis in V. This is done by an arbitrary extension of basis, and then do Gram-Schmidt. As the basis in U is already orthonormal, this will do nothing to them.

(iii) The adjoint T^* always exists and linear. (T^* always unique if exist)

<u>Spectral Theorem.</u> (eigenspace) Let $T: V \to V$ be a linear operator on a <u>finite dimensional</u> v.s. V.

$$\begin{array}{ccc} \text{Diagonalizable} & \xrightarrow[???]{\text{Inner product}} & \text{orthonormally diagonalizable ?} \\ \Updownarrow & & \Updownarrow \\ \exists \text{ eigenbasis} & \xrightarrow{???} & \exists \text{ orthonormal eigenbasis} \end{array}$$

The spectral theorem gives the condition when $T: V \to V$ is orthonormally diagonalizable.

(spectral) Thm 6.4 (i) $\mathbb{F} = \mathbb{R}$. $\underbrace{T = T^*}_{\text{self adjoint}} \iff \exists$ orthonormal eigenbasis $\iff T$ is orthonormally diagonalizable.

(ii). $\mathbb{F} = \mathbb{C}$. $\underbrace{TT^* = T^*T}_{\text{normal}} \iff \exists$ orthonormal eigenbasis $\iff T$ is unitarily diagonalizable.

Thm 6.5 (i). $T = T^* \iff \langle Tv, v \rangle \in \mathbb{R} \ \forall v \in V$. $\mathbb{F} = \mathbb{C}$.

(ii) $T = T^* \Rightarrow$ all eigenvalues are real. $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

(iii) $TT^* = T^*T$ and all eigenvalues are real $\Rightarrow T = T^*$ $\mathbb{F} = \mathbb{C}$.

(iv) $TT^* = T^*T \iff \|Tv\| = \|T^*v\| \iff \exists$ orthonormal eigenbasis $\mathbb{F} = \mathbb{C}$.

(v) $TT^* = T^*T \iff \exists$ orthonormal basis β s.t. $[T]_\beta$ is block diagonal with each block is a scalar or $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ where $b > 0$. $\mathbb{F} = \mathbb{R}$.

Isometry Let $T: V \to V$ be a linear operator on a finite dimensional vector space V.

We want to study distance preserving symmetries of space V, this is the notion of isometry. Examples include rotations and reflection. [cf. Artin Algebra].

Def. 6.6. T is an isometry if $\|Tv\| = \|v\|$ for all $v \in V$.

Lemma 6.7. $\|Tv\| = \|v\| \Leftrightarrow \langle Tv, Tw \rangle = \langle v, w \rangle \Leftrightarrow T^*T = Id \Leftrightarrow \begin{cases} A^tA = I & \text{orthogonal } F = \mathbb{R}. \\ A^HA = I & \text{unitrary } F = \mathbb{C}. \end{cases}$
$ = TT^*$

$\Rightarrow$ all eigenvalues $|\lambda_i| = 1$. where $A = [T]_\beta$ ← orthonormal basis

Thm. 6.8. (i) $F = \mathbb{R}$. $T^*T = Id \Leftrightarrow \exists$ orthonormal basis β s.t. $[T]_\beta = \begin{pmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_\ell} \end{pmatrix}$ $\theta_i \neq 0, \pi$.

where $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. ← corresponding to $\lambda, \bar{\lambda}$ in $\mathbb{C}$.

(ii) $F = \mathbb{C}$. $T^*T = id \Leftrightarrow \exists$ orthonormal eigenbasis s.t. $[T]_\beta = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ with $|\lambda_i| = 1$. (eigenvalues)

$\Leftrightarrow$ T is unitarily diagonalizable with eigenvalues $|\lambda_i| = 1$. ↖ unitrary.

(note that $T^*T = Id$ implies $T^*T = TT^*$ so spectral thm is applicable).

Q. (Unitary matrix v.s. unitarily diagonalizable:

Unitary matrix means $A^HA = I = AA^H$. Unitarily diagonalizable means $\exists$ unitary P s.t. $A = P^{-1}DP$ for diagonal D. By 6.8 (ii), unitary matrix is untarily diagonalizable with eigenvalues $|\lambda_i| = 1$.

Q. normal v.s. unitary. Over $\mathbb{C}$.

A is normal if $AA^* = A^*A$. By 6.4 (ii) and 6.8 (ii) we see that

A is normal iff A is unitary with all eigenvalues $|\lambda_i| = 1$.

[Ref. Axler Linear Algebra Done Right.]

Fourier Transform and Hilbert space.

Eg. Consider $L^2([-1,1]) = \{ f : [-1,1] \to \mathbb{R} \mid \int_{-1}^{1} |f(x)|^2 dx < \infty \}$ ← L^2-integrable

It is a vector space with $(f_1 + f_2)(x) = f_1(x) + f_2(x) \quad x \in [-1,1]$

$(af_1)(x) = a \cdot f_1(x) \quad a \in \mathbb{R}$.

Define an inner product on $L^2([-1,1])$ by

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x)\, dx$$

This makes $(L^2(-1,1), \langle , \rangle)$ an inner product space.

$\sin\alpha\cos\beta = \frac{1}{2}[\sin(\alpha-\beta) + \sin(\alpha+\beta)]$

$\sin\alpha\sin\beta = \frac{1}{2}[\cos(\alpha-\beta) - \cos(\alpha+\beta)]$

$\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha-\beta) + \cos(\alpha+\beta)]$

For example, for $n, m \in \mathbb{Z}$,

$\langle \sin m\pi x, \cos n\pi x \rangle = \int_{-1}^{1} \sin m\pi x \cdot \cos n\pi x \, dx = 0 \quad \forall m, n \in \mathbb{Z}$

$\langle \sin m\pi x, \sin n\pi x \rangle = \int_{-1}^{1} \sin m\pi x \sin n\pi x \, dx = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n \end{cases}$

$\langle \cos m\pi x, \cos n\pi x \rangle = \int_{-1}^{1} \cos m\pi x \cos n\pi x \, dx = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n \end{cases}$

Hence $\{\sin n\pi x, \cos m\pi x\}_{m,n \in \mathbb{Z}}$ are orthonormal elements in $(L^2([-1,1]), \langle , \rangle)$.

This is useful for later when we define fourier transform.

Eg. Let $\{e_i\}_{i=1}^n$ be an orthonormal basis of an inner product space $(V, \langle , \rangle)$. Let $S = \{v_i\}_{i=1}^n$ be a subset of V s.t.

$$\|v_i - e_i\| < \frac{1}{\sqrt{n}}$$

Then S is a basis.

pf. If $\{v_i\}_{i=1}^n$ are linearly dependent, $0 = \sum a_i v_i = \sum a_i (v_i - e_i) + \sum a_i e_i$.

Hence $\left\| \sum a_i (v_i - e_i) \right\| = \left\| \sum a_i e_i \right\|$.

But $\|\sum a_i (v_i - e_i)\| \underset{\text{triangle}}{\leq} \sum |a_i| \|v_i - e_i\| < \frac{1}{\sqrt{n}} \sum |a_i|$.

$\|\sum a_i e_i\| = (\sum |a_i|^2)^{\frac{1}{2}}$ since $\{e_i\}$ are orthonormal.

As $(\sum |a_i|^2)(\sum 1^2) \underset{\text{Cauchy Swarz}}{\geq} (\sum |a_i|)^2$, this gives a contradiction unless $a_i = 0$ for all i. □

Q1. Let $A \in M_n(\mathbb{R})$. Show that

$\exists B \in M_n(\mathbb{R})$ s.t. $A = B^t B \Leftrightarrow A$ is symmetric and all eigenvalues are non-negative.

Pf. $(\Rightarrow)$. $A^t = (B^t B)^t = B^t B = A$. Hence A is symmetric.

Suppose λ is any eigenvalue with eigenvector v.

$$\begin{array}{l} \langle Av, v \rangle = \lambda \langle v, v \rangle. \\ \quad \| \\ \langle B^t B v, v \rangle = \langle Bv, Bv \rangle. \end{array}$$

If $Bv = 0$, then $Av = 0$ so $\lambda = 0$

If $Bv \neq 0$, $\lambda = \dfrac{\langle Bv, Bv \rangle}{\langle v, v \rangle} > 0$.

$(\Leftarrow)$. By spectral theorem, $A = PDP^t$, where P is an orthonormal matrix, D is diagonal with eigenvalues at the diagonal. As eigenvalues are non-negative, $\sqrt{D}$ (all entries are square root of D) exist. Then let $B = P\sqrt{D}$ gives the result as $(\sqrt{D})^t = \sqrt{D}$ and $(\sqrt{D})^2 = D$. $\square$.

Q2. Let $T \in \mathcal{L}(V)$ be a self-adjoint operator on inner product space V over $\mathbb{F}$.

Suppose $\exists \lambda \in \mathbb{F}$, $\|v\| = 1$, $\varepsilon > 0$ s.t.

$$\|Tv - \lambda v\| < \varepsilon.$$

Show that T has an eigenvalue $\lambda' \in \mathbb{F}$ s.t.

$$|\lambda' - \lambda| < \varepsilon.$$

Pf. It suffices to show for matrix version. Suppose A is self-adjoint and $A = PDP^t$, for P orthogonal. In particular, $\|Pw\| = \|w\|$ for any $w \in V$.

Now $\|Tv - \lambda v\| = \|P(D - \lambda I)P^{-1}v\| = \|(D - \lambda I)P^{-1}v\|$.

Write $D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$. $P^{-1}v = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. As $\|P^{-1}v\| = \|v\| = 1$, $\left\| \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right\| = 1$.

Suppose for contradiction that $|\lambda_i - \lambda| \geq \varepsilon$ for all i. Then.

$$\varepsilon \geq \|Tv - \lambda v\| = \left\|(D - \lambda I)\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}\right\| = \left\| \begin{pmatrix} (\lambda_1 - \lambda)c_1 \\ \vdots \\ (\lambda_n - \lambda)c_n \end{pmatrix} \right\| > \varepsilon \left\| \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right\| = \varepsilon \quad \text{contradiction.}$$

Hence $\exists \lambda_i$ s.t. $|\lambda_i - \lambda| < \varepsilon$. $\square$.