8.1. Inner Product Spaces

Inner product

Linear functional

Adjoint

- Assume F is a subfield of R or C.
- Let V be a v.s. over F.
- An inner product on V is a function VxV -> F i.e., a,b in V -> (a|b) in F s.t.
  - (a) (a+b|r)=(a|r)+(b|r)
  - (b)(ca|r)=c(a|r)
  - (c) (b|a)=(a|b)$^{-}$
  - (d) (a|a) >0 if a $\neq$0.
  - Bilinear (nondegenerate) positive.
  - A ways to measure angles and lengths.

- Examples:
- F^n has a standard inner product.
  - $((x_1,..,x_n)|(y_1,\ldots,y_n)) = \sum_i x_i \overline{y_i}$
  - If F is a subfield of R, then $= x_1y_1+\ldots+x_ny_n$.
- A,B in F^{nxn}.
  - $(A|B) = tr(AB^*)=tr(B^*A)$
    - Bilinear property: easy to see.
    - $tr(AB^*)= \sum_j (AB^*)_j = \sum_j \sum_k A_{jk} B_{kj}{}^* = \sum_j \sum_k A_{jk} \overline{B}_{jk}$

$$tr(AA^*) = \sum \sum \left| A_{jk} \right|^2$$

- $(X|Y)=Y^*Q^*QX$, where X,Y in F^{nx1}, Q nxn invertible matrix.
  - Bilinearity follows easily
  - $(X|X)=X^*Q^*QX=(QX|QX)_{std} \geq 0$.
  - In fact almost all inner products are of this form.
- Linear T:V->W and W has an inner product, then V has "induced" inner product.
  - $p_T(a|b):= (Ta|Tb)$.

- (a) For any basis $B=\{a_1,\ldots,a_n\}$, there is an inner-product s.t. $(a_i|a_j)=\delta_{ij}$.
  - Define $T:V\to F^n$ s.t. $a_i \to e_i$.
  - Then $p_T(a_i|a_j)=(e_i|e_j)=\delta_{ij}$.
- (b) $V=\{f:[0,1]\to C \mid f \text{ is continuous}\}$.
  - $(f|g)=\int_0^1 f\bar{g}\,dt$ for f,g in V is an inner product.
  - $T:V\to V$ defined by $f(t) \to tf(t)$ is linear.
  - $p_T(f,g)= \int_0^1 tf\overline{tg}\,dt=\int_0^1 t^2 f\bar{g}\,dt$ is an inner product.

- Polarization identity: Let F be an imaginary field in C.
- $(a|b)=\mathrm{Re}(a|b)+i\mathrm{Re}(a|ib)$ (*):
  - $(a|b)=\mathrm{Re}(a|b)+i\mathrm{Im}(a|b)$.
  - Use the identity $\mathrm{Im}(z)=\mathrm{Re}(-iz)$.
  - $\mathrm{Im}(a|b)=\mathrm{Re}(-i(a|b))=\mathrm{Re}(a|ib)$
- Define $||a|| := (a|a)^{1/2}$ norm
- $||a\pm b||^2=||a||^2\pm 2\mathrm{Re}(a|b)+||b||^2$ (**).
- $(a|b)=||a+b||^2/4-||a-b||^2/4+i||a+ib||^2/4 -i||a-ib||^2/4$. (proof by (*) and (**).)
- $(a|b)= ||a+b||^2/4-||a-b||^2/4$ if F is a real field.

- When V is finite-dimensional, inner products can be classified.
- Given a basis $B=\{a_1,\ldots,a_n\}$ and any inner product (|):
  $(a|b) = Y^*GX$ for $X=[a]_B$, $Y=[b]_B$
  - G is an nxn-matrix and $G=G^*$, $X^*GX>0$ for any X, $X\neq 0$.
- Proof: (->) Let $G_{jk}=(a_k|a_j)$.

$$a=\sum_k x_k a_k, b=\sum_j y_j a_j$$

$$(a\,|\,b)=(\sum_k x_k a_k, b)=\sum_k x_k \sum_j \overline{y}_j (a_k\,|\,a_j)=Y^* GX$$

– $G=G^*$: $(a_j|a_k)=(a_k|a_j)^-$. $G_{kj}=G_{jk}^-$.
– $X^*GX =(a|a) > 0$ if $X\neq 0$.
– (G is invertible. $GX\neq 0$ by above for $X\neq 0$.)
– (<-) X^*GY is an inner-product on F^{nx1}.
  - (a|b) is an induced inner product by a linear transformation T sending a_i to e_i.

– Recall Cholesky decomposition: Hermitian positive definite matrix $A = L L^*$. L lower triangular with real positive diagonal. (all these are useful in appl. Math.)

8.2. Inner product spaces

- Definition: An inner product space (V, (|))
- F⊂R -> Euclidean space
- F⊂C -> Unitary space.
- Theorem 1. V, (|). Inner product space.

1. $||ca||=|c|||a||$.
2. $||a|| > 0$ for $a\neq 0$.
3. $|(a|b)| \leq ||a||||b||$ (Cauchy-Schwarz)
4. $||a+b|| \leq ||a||+||b||$

- Proof (ii)

$$a=0 \text{ trivial}$$

$$a\neq 0,\quad r=b-\frac{(b\mid a)}{\|a\|^2}a,\ (r\mid a)=0$$

$$0\leq\|r\|^2=\left(b-\frac{(b\mid a)}{\|a\|^2}a \,\middle|\, b-\frac{(b\mid a)}{\|a\|^2}a\right)=(b\mid b)-\frac{(b\mid a)(b\mid a)}{\|a\|^2}$$

$$=\|b\|^2-\frac{|(a\mid b)|^2}{\|a\|^2}$$

- Proof (iii)

$$\|a+b\|^2=\|a\|^2+(a\mid b)+(b\mid a)+\|b\|^2$$

$$=\|a\|^2+2\operatorname{Re}(a\mid b)+\|b\|^2$$

$$\leq\|a\|^2+2\|a\|\|b\|+\|b\|^2=(\|a\|^2+\|b\|^2)^2$$

- In fact many inequalities follows from Cauchy-Schwarz inequality.
- The triangle inequality also follows.
- See Example 7.
- Example 7 (d) is useful in defining Hilbert spaces. Similar inequalities are used much in analysis, PDE, and so on.
- Note Example 7, no computations are involved in proving these.

- On inner product spaces one can use the inner product to simplify many things occurring in vector spaces.
  - Basis -> orthogonal basis.
  - Projections -> orthogonal projections
  - Complement -> orthogonal complement.
  - Linear functions have adjoints
  - Linear functionals become vector
  - Operators -> orthogonal operators and self adjoint operators (we restrict to)

Orthogonal basis

- Definition:
  - a,b in V, $a \perp b$ if $(a|b)=0$.
  - The zero vector is orthogonal to every vector.
  - An orthogonal set S is a set s.t. all pairs of distinct vectors are orthogonal.
  - An orthonormal set S is an orthogonal set of unit vectors.

- Theorem 2. An orthogonal set of nonzero-vectors is linearly independent.
- Proof: Let $a_1,\ldots,a_m$ be the set.
  - Let $0=b=c_1a_1+\ldots+c_ma_m.$
  - $0=(b,a_k)=(c_1a_1+\ldots+c_ma_m,\ a_k\ )=c_k(a_k\ |a_k\ )$
  - $c_k=0.$
- Corollary. If b is a linear combination of orthogonal set $a_1,\ldots,a_m$ of nonzero vectors, then $b=\sum_{k=1}^{m}((b|a_k)/||a_k||^2)\ a_k$
- Proof: See above equations for $b\neq 0$.

- Gram-Schmidt orthogonalization:
- Theorem 3. $b_1,\ldots,b_n$ in V independent. Then one may construct orthogonal basis $a_1,\ldots,a_n$ s.t. $\{a_1,\ldots,a_k\}$ is a basis for $\langle b_1,\ldots,b_k\rangle$ for each $k=1,..,n$.
- Proof: $a_1 := b_1$. $a_2=b_2-((b_2|a_1)/||a_1||^2)a_1,\ldots,$
  - Induction: $\{a_1,..,a_m\}$ constructed and is a basis for $\langle b_1,\ldots,b_m\rangle$.
  - Define

$$a_{m+1} = b_{m+1} - \sum_{k=1}^{m} \frac{(b_{m+1} \mid a_k)}{\| a_k \|^2} a_k$$

– Then

$$(a_{m+1} \mid a_j) = (b_{m+1} \mid a_j) - \sum_{k=1}^{m} \frac{(b_{m+1} \mid a_k)}{\| a_k \|^2} (a_k \mid a_j)$$

$$= (b_{m+1} \mid a_j) - (b_{m+1} \mid a_j) = 0$$

– Use Theorem 2 to show that the result $\{a_1, \ldots, a_{m+1}\}$ is independent and hence is a basis of $<b_1,\ldots,b_{m+1}>$.

- See p.281, equation (8-10) for some examples.
- See examples 12 and 13.

Best approximation, Orthogonal complement, Orthogonal projections

- This is often used in applied mathematics needing approximations in many cases.
- Definition: W a subspace of V. b in W. Then the best approximation of b by a vector in W is *a* in W s.t. $||b-a|| \leq ||b-c||$ for all c in W.
- Existence and Uniqueness. (finite-dimensional case)

- Theorem 4: W a subspace of V. b in V.
  - (i). a is a best appr to b <-> b-a $\perp$ c for all c in W.
  - (ii). A best appr is unique (if it exists)
  - (iii). W finite dimensional.
    $\{a_1,..,a_k\}$ any orthonormal basis.

$$a = \sum_k \frac{(b \mid a_k)}{\| a_k \|^2} a_k$$

is the best approx. to b by vectors in W.

- Proof: (i)
  - Fact: Let c in W. b-c =(b-a)+(a-c). $||b-c||^2=||b-a||^2+2Re(b-a|a-c)+||a-c||^2$(*)
  - (<-) b-a $\perp$W. If c $\neq$a, then $||b-c||^2=||b-a||+||a-c||^2 > ||b-a||^2$. Hence a is the best appr.
  - (->) $||b-c||\geq||b-a||$ for every c in W.
    - By (*) $2Re(b-a|a-c)+||a-c||^2 \geq 0$
    - <-> $2Re(b-a|t)+||t||^2 \geq 0$ for every t in W.
    - If a$\neq$c, take t = $-\frac{(b-a|a-c)}{||a-c||^2}(a-c)$

$$2\operatorname{Re}\left(b-a \,\middle|\, -\frac{(b-a|a-c)}{\|a-c\|^2}(a-c)\right)+\frac{|(b-a|a-c)|^2}{\|a-c\|^4}\|a-c\|^2 \geq 0$$

$$-2\operatorname{Re}\left(\overline{\frac{(b-a|a-c)}{\|a-c\|^2}}(b-a|a-c)\right)+\frac{|(b-a|a-c)|^2}{\|a-c\|^2} \geq 0$$

$$-2\operatorname{Re}\frac{|(b-a|a-c)|^2}{\|a-c\|^2}+\frac{|(b-a|a-c)|^2}{\|a-c\|^2} \geq 0$$

$$\frac{|(b-a|a-c)|^2}{\|a-c\|^2} \leq 0$$

$$-2\frac{|(b-a|a-c)|^2}{||a-c||^2}+\frac{|(b-a|a-c)|^2}{||a-c||^2}\geq 0$$

- This holds <-> (b-a|a-c)=0 for any c in W.
- Thus, b-a is ⊥ every vector in W.

– (ii) a,a' best appr. to b in W.

- b-a ⊥ every v in W. b-a' ⊥ every v in W.
- If a≠a', then by (*)
  $||b-a'||^2=||b-a||^2+2\text{Re}(b-a|a-a')+||a-a'||^2$.
  Hence, ||b-a'||>||b-a||.
- Conversely, ||b-a||>||b-a'||.
- This is a contradiction and a=a'.

– (iii) Take inner product of a_k with

$$b - a = b - \sum_k \frac{(b \mid a_k)}{\| a_k \|^2} a_k$$

– This is zero. Thus b-a $\perp$ every vector in W.

Orthogonal projection

- Orthogonal complement. S a set in V.
- $S^{\perp}$:={v in V| $v \perp w$ for all w in S}.
- $S^{\perp}$ is a subspace. $V^{\perp}=\{0\}$.
- If S is a subspace, then $V=S \oplus S^{\perp}$ and $(S^{\perp})^{\perp}=S$.
- Proof: Use Gram-Schmidt orthogonalization to a basis $\{a_1,\ldots,a_r,a_{r+1},\ldots,a_n\}$ of V where $\{a_1,\ldots,a_r\}$ is a basis of V.

- Orthogonal projection: E_W:V->W.
  a in V -> b the best approximation in W.
- By Theorem 4, this is well-defined for any subspace W.
- E_W is linear by Theorem 5.
- E_W is a projection since $E_W \circ E_W(v) = E_W(v)$.

- Theorem 5: W subspace in V. E orthogonal projection V->W. Then E is an projection and $W^{\perp}$=nullE and $V=W\oplus W^{\perp}$.
- Proof:
  - Linearity:
    - a,b in V, c in F. a-Ea, b-Eb $\perp$ all v in W.
    - c(a-Ea)+(b-Eb)=(ca+b)-(cE(a)+E(b)) $\perp$ all v in W.
    - Thus by uniqueness E(ca+b)=cEa+Eb.
  - null E $\subset W^{\perp}$: If b is in nullE, then b=b-Eb is in $W^{\perp}$.
  - $W^{\perp} \subset$ null E: If b is in $W^{\perp}$, then b-0 is in $W^{\perp}$ and 0 is the best appr to b by Theorem 4(i) and so Eb=0.
  - Since V=ImE$\oplus$nullE, we are done.

- Corollary: b-> b-E_wb is an orthogonal projection to $W^\perp$. I-E_w is an idempotent linear transformation; i.e., projection.
- Proof: b-> b-E_wb is in $W^\perp$ by Theorem 4 (i).
  - Let c be in $W^\perp$. b-c=Eb+(b-Eb-c).
  - Eb in W, (b-Eb-c) in $W^\perp$.
  - $||b-c||^2=||Eb||^2+||b-Eb-c||^2 \geq ||b-(b-Eb)||^2$ and $>$ if $c \neq b-Eb$.
  - Thus, b-Eb is the best appr to b in $W^\perp$.

Bessel’s inequality

- $\{a_1,\ldots,a_n\}$ orthogonal set of nonzero vectors. Then

$$\sum_{k=1}^{n} \frac{|(b \mid a_k)|^2}{\|a_k\|^2} \leq \|b\|^2$$

- **= <->** $$b = \sum_{k=1}^{n} \frac{(b \mid a_k)}{\|a_k\|^2} a_k$$