H5: inner products and orthogonality

5.1 prologue: bilinear and sesquilinear forms

def: bilinear map	For three vector spaces V_1, V_2, V_3 over a common field $\mathbb{F}$ > a bilinear map is a map B: $V_1 \times V_2 \rightarrow V_3$ that is linear in both first and second argument: $\forall a, b \in \mathbb{F}, v_1, w_1 \in V_1$ and $v_2, w_2 \in V_2$, it satisfies $$B(av_1 + bw_1, v_2) = aB(v_1, v_2) + bB(w_1, v_2),$$ $$B(v_1, av_2 + bw_2) = aB(v_1, v_2) + bB(v_1, w_2).$$
def: bilinear form	bilinear form on a vector space V over a field $\mathbb{F}$ = a bilinear map B: V x V→$\mathbb{F}$
def: symmetric bilinear map	For a bilinear form B on a vector space V For all v,w∈V > B is symmetric if B(v,w) = B(w,v)
def: antisymmetric bilinear map	For a bilinear form B on a vector space V For all v,w∈V > B is symmetric if B(v,w) = -B(w,v)
vector space basis	We can choose a basis $B_V = \{e_1,\ldots,e_n\}$ for a finite-dimensional vector space such that: and completely characterise the bilinear form as $$B(v, w) = B(v^i e_i, w^j e_j) = v^i B(e_i, e_j) w^j = v^i B_{ij} w^j = \boldsymbol{v}^{\mathsf{T}} \mathsf{B} \boldsymbol{w} \qquad (5.1)$$ in terms of the coefficients $B_{ij} = B(e_i, e_j) \in \mathbb{F}$, which we have organised in a a matrix $\mathsf{B} \in \mathbb{F}^{n \times n}$.
bilinear decomposition	Any bilinear form B can be decomposed in a symm. B_S and antisymm. B_A: $$B(v,w) = \underbrace{\frac{1}{2}\left(B(v,w) + B(w,v)\right)}_{B_S(v,w)} + \underbrace{\frac{1}{2}\left(B(v,w) - B(w,v)\right)}_{B_A(v,w)}$$
def: quadratic form	A quadratic form on a real vector space V is a map q:V→$\mathbb{R}$ such that: $$B(v,w) = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right)$$ is a bilinear form > it implies that $q(av) = a^2 q(v)$
def: associated bilinear form	= the bilinear form B in the definition of the quadratic form > by definition symmetric²
def: sesquilinear form C	For a vector space V over $\mathbb{F}$ For any a,b∈$\mathbb{F}$ and u,v,w∈V > the sesquilinear form C is a map B: V x V →$\mathbb{F}$ that is: - linear in second argument - antilinear in first argument • $C(au + bv, w) = \bar{a}C(u, w) + \bar{b}C(v, w)$ • $C(u, av + bw) = aC(u, v) + bC(u, w)$
def: Hermitian	= a sesquilinear form C on the vector space V for which for all v,w∈V: $$C(v,w) = \overline{C(w,v)}$$
def: anti-Hermitian	= a sesquilinear form C on the vector space V for which for all v,w∈V: $$C(v,w) = -\overline{C(w,v)}$$

def: matrix congruence	For V an n-dimensional vector space over $\mathbb{F}$, with a basis and sesquilinear form C > the matrix congruence is the transformation behaviour of the matrix C with respect to a basis transform T∈GL(n,$\mathbb{F}$) ie: $\tilde{C} = T^{-H}CT^{-1}$
def: degenerate form	For C a sesquilinear form on a vector space V For all w∈V > C is degenerate if there exists a nonzero v∈V such that C(w,v)=0
def: positive definite positive semidefinite indefinite	For C a Hermitian sesquilinear form on the vector space V For all nonzero vectors v∈V = C(v,v)>0 = C(v,v)≥0 = C(v,v) can be both pos. or neg.
prop: degenerate form	If the sesquilinear form C is positive semidefinite but not positive definite it is degenerate
5.2 inner product spaces	
def: inner product	inner product on a real or complex vector space V = positive Hermitian sesquilinear VxV→$\mathbb{F}$ not: $\langle v, w\rangle$
properties of inner product	• Linearity in the second argument: $\langle u, av + bw\rangle = a\langle u, v\rangle + b\langle u, w\rangle$, $\forall u, v, w \in V$ and $\forall a, b \in \mathbb{F}$ • Hermiticity: $\langle v, w\rangle = \overline{\langle w, v\rangle}$, $\forall v, w \in V$ • Antilinearity in the first argument (this is actually implied by combining the previous two properties) • Positive definiteness: $\langle v, v\rangle > 0$ for all $v \neq o$.
inner product space	= vector space with an inner product $(V, \langle\cdot,\cdot\rangle)$
Euclidean space	= inner product space for which $\mathbb{F}=\mathbb{R}$
Gram matrix	For a chosen basis $B_V = \{e_1,\ldots,e_n\}$ > the matrix representation of the inner product is denoted as g with: $g_{ij} = \overline{g_{ji}} = \langle e_i, e_j\rangle$
5.2.1 Cauchy-Schwarz inequality and its consequences	
Theorem: Cauchy-Schwarz	$\lvert\langle v, w\rangle\rvert^2 \leq \langle v, v\rangle\langle w, w\rangle$ > inequality becomes equal if and only if v and w are linearly dependent

5.3 orthogonality and unitarity	
5.3.1 orthogonality and orthonormality	
def: orthogonal	For two vectors v,w ∈V > these two are orthogonal if: v⊥w $$\langle v, w\rangle = 0.$$ two subsets A,B can be orthogonal if: $$\langle v, w\rangle = 0 \text{ for any } v \in A \text{ and } w \in B.$$
def: normalized	A vector v is normalized if ‖v‖ = 1
def: orthogonal set	For a set of vectors $\{v_i; i∈I\}$ with I finite, countable infinite or uncountable infinite > this set is orthogonal if: - $v_i \neq 0$ - $\langle v_i, v_j\rangle = 0$ for all $i \neq j \in I$.
def: orthonormal set	= orthogonal set where for all v_i: $$\|v_i\| = 1 \text{ for all } i \in I$$
prop: orth. vectors lin. ind.	A set of orthogonal vectors is linearly independent
> prop: amount of orthogonal vect.	In a finite dimensional V, we can construct at most n=dim(V) orthogonal vectors
th: Pythagoras	Given a finite set of orthogonal vectors $\{v_i; i=1,\ldots,n\} \subseteq V$, then: $$\left\|\sum_{i=1}^{n} v_i\right\|^2 = \sum_{i=1}^{n} \|v_i\|^2$$
5.3.2 orthogonal complements and orthogonal projections	
def: orthogonal complement	For a subset S⊆V of an inner product space $(V, \langle\ ,\ \rangle)$ > define $S^\perp$ as: $$S^\perp = \{v \in V \mid \langle w, v\rangle = 0 \text{ for all } w \in S\}.$$
prop: orth. compl. closed subspace	The orthogonal complement $S^\perp$ is a closed subspace of V
prop: S and 𝔽S	For a subset S⊆V, it holds: $$S^\perp = (\mathbb{F}S)^\perp = (\overline{\mathbb{F}S})^\perp.$$
prop: intersection of $S^\perp$	For a subset S⊆V, it holds that: $S \cap S^\perp$ *is* $\{o\}$ or empty
th: projection	For a closed subspace W≼V of a Hilbert space V > then: • *For any vector $v \in V$, there exists a unique closest vector $w \in W$ such that* $$\|v - w\| = \inf_{w' \in W} \|v - w'\|.$$ • *The vector $w \in W$ closest to v is the unique element in W satisfying $v - w \in W^\perp$.* *The vector $w \in W$ is known as the* ***orthogonal projection*** *of v onto W.*
prop: orthogonal direct sum	For W a closed subspace of a Hilbert space V > then V = W⊕$W^\perp$ and this orthogonal direct sum decomposition is unique

def: orthogonal projector	For a Hilbert space V > this is a linear operator ^P∈End(V) which satisfies: $\hat{P}^2 = \hat{P}, \qquad \langle \hat{P}v, w \rangle = \langle v, \hat{P}w \rangle, \forall v, w \in V.$
lemma: norm of ^P	**Lemma 5.15.** *Any nonzero orthogonal projector $\hat{P}$ on a Hilbert space V has operator norm $\|P\| = 1$. It is thus bounded and continuous.*
theory: im and ker of ^P 5.16	For a Hilbert space V For W a closed subspace such that V=W⊕W⊥ > then $\hat{P}_W$ the orthogonal projector
> theory: vice versa of 5.16	For an orthogonal projector $\hat{P}_w$ > this gives rise to an orthogonal direct sum decomposition: $V = \mathrm{im}(\hat{P}) \oplus \ker(\hat{P})$, *with thus* $\mathrm{im}(\hat{P}) = \ker(\hat{P})^\perp$ *and both are closed subspaces of* V.
prop: $S^{\perp\perp}$	For any subset S of a Hilbert space V, it holds that V = $S^\perp \oplus S^{\perp\perp}$ where $S^{\perp\perp} = (S^\perp)^\perp$
prop: $W^{\perp\perp}$	*For any closed subspace W of a Hilbert space V, $W^{\perp\perp} = W$.*
prop: $S^{\perp\perp} = \overline{\mathbb{F}S}$.	*For any subset S of a Hilbert space V, it holds that $S^{\perp\perp} = \overline{\mathbb{F}S}$.*
prop: $S^\perp = \{0\}$	for S a complete set of a Hilbert space V > span of S is dense in V > then: *thus* $\overline{\mathbb{F}S} = V$, *then* $S^\perp = \{o\}$.
5.3.3 orthonormal basis for Hilbert spaces	
lemma: orthogonal proj. of v	For an infinite-dimensional Hilbert space V For an orthonormal sequence $(e_i)_{i\in\mathbb{N}_0}$. For a vector v ∈ V > the orthogonal projection of v onto subspace $W_n = \mathbb{F}\{e_i;\ i = 1,\ldots,n\}$ is given by: $v_n = \sum_{i=1}^{n} e_i \langle e_i, v \rangle.$
lemma: Bessel's inequality	For an infinite-dimensional Hilbert space V For an orthonormal sequence $(e_i)_{i\in\mathbb{N}_0}$. For a vector v ∈ V and any $n \in \mathbb{N}_0$ > $\sum_{i=1}^{n} \lvert\langle e_i, v \rangle\rvert^2 \leq \|v\|^2.$ *so that in particular the series* $\sum_{i=1}^{+\infty} \lvert\langle e_i, v\rangle\rvert^2$ *converges to value upper bounded by* $\|v\|^2$.
th: expansion theorem	For V a Hilbert space for S a complete orthonormal sequence $S = (e_i)_{i\in\mathbb{N}_0}$ > then $(e_i)_{i\in\mathbb{N}_0}$ *is a Schauder basis and any vector* $v \in V$ *can be expanded* $v = \sum_{i=1}^{+\infty} \langle e_i, v \rangle e_i.$
prop: Plancherel's identity	Bessel's inequality becomes an equality: $\|v\|^2 = \langle v, v \rangle = \sum_{i=1}^{+\infty} \lvert\langle e_i, v \rangle\rvert^2 = \sum_{i=1}^{+\infty} \langle v, e_i \rangle \langle e_i, v \rangle$
prop: Perseval's identity	more general, Plancheral's identity becomes: $\langle v, w \rangle = \sum_{i=1}^{+\infty} \langle v, e_i \rangle \langle e_i, w \rangle$

5.3.4 Gram-Schmidt orthonormalization	
def: Gram-Schmidt process	For a countable set of linearly independent vectors S={v_1,v_2,…} For these vectors in an inner product space $(V, \langle\ ,\ \rangle)$. > the Gram-Schmidt process is a strategy to construct an orthonormal sequence (q_1,q_2,…) that has the same linear span as S. $w_1 = v_1 \qquad q_1 = w_1 / \|w_1\|$ $w_2 = v_2 - \langle q_1, v_2 \rangle q_1 \qquad q_2 = w_2 / \|w_2\|$ $\dots$ $w_k = v_k - \sum_{j=1}^{k-1} \langle q_j, v_k \rangle q_j \qquad q_k = w_k / \|w_k\|$ $\dots$
def: QR decomposition	= the decomposition of a matrix $V \in \mathbb{F}^{m \times n}$ with m≥n in the form V=QR with $Q \in \mathbb{F}^{m \times n}$ satisfying $Q^H Q = I_n$ and $R \in \mathbb{F}^{m \times n}$ an upper-triangle matrix > find the coefficients via: $R^j{}_k = \left\langle \boldsymbol{q}_j, \boldsymbol{v}_k \right\rangle = \boldsymbol{q}_j^{\mathsf{H}} \boldsymbol{v}_k, \qquad R^k{}_k = \left\| \boldsymbol{v}_k - \sum_{j=1}^{k-1} R^j{}_k \boldsymbol{q}_j \right\|,$
5.4 linear functionals and the duality of Hilbert space	
prop: inner prod. in Hilbert space	On a Hilbert space V, the inner product structure defines a canonical injective antilinear map $V \to V^* : v \mapsto \chi_v$ defined via the action: $\chi_v(w) = \langle v, w \rangle, \qquad \forall w \in V.$ *Furthermore, the resulting linear functional χ_v is continuous and thus bounded.*
5.4.2 Riesz representation theorem	
th: Riesz representation theorem	every bounded linear functional ς on Hilbert space V is uniquely associated with a vector $v_ς \in V$ such that: $\xi[w] = \langle v_\xi, w \rangle$ *for all* $w \in W$.
5.5 bounded linear maps in Hilbert spaces	
5.5.1 preliminaries	
Induced norm	For $\hat{A}$ the space of bounded linear maps between Hilbert spaces: $(V, \langle\ ,\ \rangle_V = \langle\ ,\ \rangle)$ and $(W, \langle\ ,\ \rangle_W)$, > We can reformulate the induced norm: $\|\hat{A}\| = \sup_{v \in V, v \neq o} \frac{\|\hat{A}v\|_W}{\|v\|_V} = \sup_{v \in V, \|v\|_V = 1} \|\hat{A}v\|_W$ $= \sup_{v \in V, w \in W, v \neq o, w \neq o} \frac{\lvert\langle w, \hat{A}v \rangle_W\rvert}{\|w\|_W \|v\|_V} = \sup_{v \in V, w \in V, \|v\|_V = 1, \|w\|_W = 1} \lvert\langle w, \hat{A}v \rangle_W\rvert.$

5.5.2 adjoint of a linear map	
prop: adjoint/dagger	For every map Â∈𝔅(V,W) > we can construct a map Â† such that for any v∈V and w∈W: $$\langle w, \hat{A}v\rangle_W = \left\langle \hat{A}^\dagger w, v\right\rangle_V.$$ *It furthermore holds that* $\|\hat{A}\| = \|\hat{A}^\dagger\|$.
def: dagger	= map † from Hom(V,W) to Hom(W,V) > properties: • anti-linearity: $(a\hat{A} + b\hat{B})^\dagger = \bar{a}\hat{A}^\dagger + \bar{b}\hat{B}^\dagger$; • anti-homomorphism for composition: $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$; • involution: $\hat{A}^{\dagger\dagger} = (\hat{A}^\dagger)^\dagger = \hat{A}$. furthermore: $\hat{1}_V^\dagger = \hat{1}_V$ and that, if $\hat{A}$ is invertible, $(\hat{A}^{-1})^\dagger = (\hat{A}^\dagger)^{-1}$.
Hermitian conjugate and dagger	For finite dimensional space V and W For corresponding basis B_V and B_W Rewrite the defining equation as: $$\overline{w}^i (g_W)_{ij} A^j{}_k v^k = \overline{(A^\dagger)^l{}_i w^i} (g_V)_{lk} v^k.$$ From this, we infer that $$(A^\dagger)^l{}_i = \overline{(g_W)_{ij} A^j{}_k (g_V^{-1})^{kl}}.$$ In the common case where both B_V and B_W are orthonormal bases, this becomes $$(A^\dagger)^l{}_i = \overline{A^i{}_l} = (\mathsf{A}^\mathsf{H})^l{}_i \qquad \Longleftrightarrow \qquad \Phi_{B_V,B_W}(\hat{A}^\dagger) = \mathsf{A}^\mathsf{H} = \Phi_{B_W,B_V}(\hat{A})^\mathsf{H}.$$ > adjoint coincides with Hermitian conjugate > $A^H = A^\dagger$
prop: norm of †	For Â a bounded linear map between Hilbert spaces V and W: $$\left\|\hat{A}^\dagger\hat{A}\right\| = \left\|\hat{A}\hat{A}^\dagger\right\| = \|\hat{A}\|^2.$$
prop: ker and im with †	For Â a bounded linear map between Hilbert spaces V and W > we already know $\ker(\hat{A}^\dagger) = \operatorname{im}(\hat{A})^\perp$ *and thus* $$W = \ker(\hat{A}^\dagger) \oplus \overline{\operatorname{im}(\hat{A})}, \qquad V = \ker(\hat{A}) \oplus \overline{\operatorname{im}(\hat{A}^\dagger)}$$
5.5.3 self-adjoint operators	
def: self adjoint operator	= a bounded linear operator Â on the Hilbert space V for which Â = Â†
† and definitions of operators	a linear operator Â∈V is called • skew-adjoint or anti-Hermitian or skew-Hermitian if $\hat{A}^\dagger = -\hat{A}$; • positive semidefinite if it is self-adjoint and $\langle v, \hat{A}v\rangle \geq 0$ for all $v \in V$, and positive definite if $\langle v, \hat{A}v\rangle > 0$ for all $v \neq o$.
5.5.4 isometric and unitary maps	
prop: isometry	For a linear map ^Q:V→W between Hilbert spaces V and W > ^Q is isometric with respect to $d_W(w', w) = \|w' - w\|_W$ *and* $d_V(v', v) = \|v - v'\|_V$ *if and only if* $$\hat{Q}^\dagger\hat{Q} = \hat{1}_V$$ *Such a map is said to be a (linear)* ***isometry***.

prop: unitary linear map	= An isometry Û:V→W that is surjective, and thus invertible, and has Û† as its inverse $$\hat{U}^\dagger \hat{U} = \hat{1}_V, \qquad \hat{U}\hat{U}^\dagger = \hat{1}_W.$$ *Such a linear map is said to be **unitary** or, when $\mathbb{F} = \mathbb{R}$, **orthogonal**.*
isomorphic Hilbert spaces	= Hilbert spaces V and W such that there exists a unitary map Û:V→W
5.5.5 antiunitary transformation	
prop: adjoint for every map	For every antilinear map Â:V→W between Hilbert spaces V and W > there exists an antilinear map A† such that: $$\langle w, \hat{A}v\rangle = \overline{\langle \hat{A}^\dagger w, v\rangle} = \left\langle v, \hat{A}^\dagger w\right\rangle$$
def: antiunitary map	= isometric antilinear map Â that is invertible > isometric: $\hat{A}^\dagger \hat{A} = \hat{1}_V{}^6$, which implies that $\langle \hat{A}v_1, \hat{A}v_2\rangle = \langle v_2, v_1\rangle$ for all $v_1, v_2 \in V$.
5.5.6 normal operators	
def: normal operator	A linear operator Â on a Hilbert space V is normal if: $$\hat{A}^\dagger \hat{A} = \hat{A}\hat{A}^\dagger \iff \left[\hat{A}^\dagger, \hat{A}\right] = \hat{0}.$$
decomposition of an operator	every operator can be split into two self-adjoint parts: $$\hat{A} = \frac{\hat{A} + \hat{A}^\dagger}{2} + \mathrm{i}\frac{\hat{A} - \hat{A}^\dagger}{2\mathrm{i}} = \hat{A}_\mathrm{r} + \mathrm{i}\hat{A}_\mathrm{i}$$
prop: condition for normal op.	An operator Â on a Hilbert space V is normal if and only if for all v∈V: $$\lVert\hat{A}v\rVert = \lVert\hat{A}^\dagger v\rVert$$
prop: eigenspace of Â†	For Â a normal operator on V For v an eigenvector with eigenvalue λ > it holds: $\hat{A}^\dagger v = \overline{\lambda} v$ ie: Â and Â† share eigenspaces up to the complex conjugate
prop: orthogonality of eigenvectors	For a normal operator Â, eigenvectors v_λ and v_ϕ are orthogonal if λ≠ϕ
prop: norm on normal operator	Using the operator norm, a normal operator Â on a Hilbert space V satisfies: $$\lVert\hat{A}^n\rVert = \lVert\hat{A}\rVert^n$$
prop: norm and spectral radius	For a normal operator Â: $\lVert\hat{A}\rVert = \rho_{\hat{A}}$ ie: the operator norm equals the spectral radius
prop: nilpotent normal operator	A normal operator Â CANNOT be nilpotent
5.5.7 Hilbert-Schmidt inner product	
def: Hilbert-Schmidt inner product	Between bounded linear maps $\hat{A}, \hat{B} \in \mathcal{B}(V, W)$ is defined as $$\begin{aligned}\langle \hat{A}, \hat{B}\rangle_{\mathrm{HS}} &= \mathrm{tr}_V(\hat{A}^\dagger \hat{B}) \\ &= \sum_n \left\langle e_n, \hat{A}^\dagger \hat{B} e_n\right\rangle_V = \sum_n \langle \hat{A}e_n, \hat{B}e_n\rangle_W = \sum_{m,n} \langle \hat{A}e_n, f_m\rangle_W \langle f_m, \hat{B}e_n\rangle_W \\ &= \sum_{m,n} \left\langle e_n, \hat{A}^\dagger f_m\right\rangle_V \left\langle \hat{B}^\dagger f_m, e_n\right\rangle_W = \sum_m \left\langle \hat{B}^\dagger f_m, \hat{A}^\dagger f_m\right\rangle = \mathrm{tr}_W(\hat{B}\hat{A}^\dagger).\end{aligned}$$ with orthonormal bases $B_V = \{e_n; n \in I \subseteq \mathbb{N}_0\}$ and $B_W = \{f_m; m \in J \subseteq \mathbb{N}_0\}$
Hilbert-Schmidt norm	$\lVert\hat{A}\rVert_{\mathrm{HS}} = \mathrm{tr}(\hat{A}^\dagger \hat{A})^{1/2}.$

5.6 application: least squares solutions	
def: least squares solution	For a matrix $A \in \mathbb{F}^{m \times n}$ with $m \geq n$ that has full rank > the vector **x*** is the least squares solution of the overdetermined linear system $A\mathbf{x} = \mathbf{y}$: $$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \mathbb{F}^n} \|A\boldsymbol{x} - \boldsymbol{y}\|_2 .$$
def: Moore-Penrose pseudoinverse	For a full rank matrix $A \in \mathbb{F}^{m \times n}$ with $m \geq n$ this is: $$A^+ = (A^H A)^{-1} A^H.$$ properties: - $AA^+A = A$, - $A^+AA^+ = A^+$.