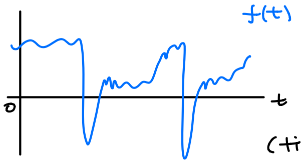
1/25/21

Applied Harmonic Analysis

Background :



(time space)

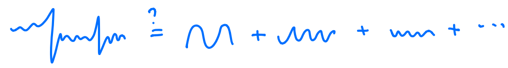
signal { voice, sound
images (2D, 3D)
radio waves (FM)

Fourier analysis :

Fourier series (periodic f)

$f(t) \overset{?}{=} a_0/2 + \sum_{n=1}^{\infty} a_n \cos(w_n t) + b_n \sin(w_n t)$

amplitude — frequency



Fourier transform (nonperiodic f)

$\mathcal{F}[f] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t) e^{-ikt}\, dt$ where $i = \sqrt{-1}$

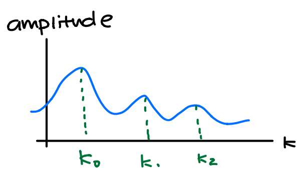
Function $f \in V \leftarrow$ some vector space (say $L^2(\mathbb{R})$ or $\ell^2(\mathbb{R})$)

↓

Decompose f (into say sines/cosines or wavelets)

↓

In frequency space, can manipulate f : compression, denoising, apply other filters ..

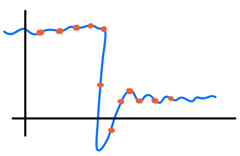


(frequency)

ex: low-pass filter

$f(t) \approx \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(w_n t) + b_n \sin(w_n t)$

In practice, need to discretize ! - DFT (Discrete Fourier Transform)
- DWT (Discrete Wavelet Transform)



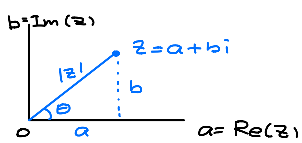
need to sample points of f.

Number Systems.

countable {
$\mathbb{N} = \{0, 1, 2, 3 \ldots\}$ naturals
$\mathbb{Z} = \{\ldots -1, 0, 1, 2 \ldots\}$ integers
$\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z},\ q \neq 0\}$ rationals.

uncountable {
$\mathbb{R} = \{$ real numbers $\}$ ex: $\pi, -\frac{1}{2}, 0, e, \ldots$
$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$ complex.

Complex numbers



$|z| = z\bar{z} = \sqrt{a^2 + b^2}$ (modulus)

conjugate $\bar{z} = a - bi$

polar form : $z = re^{i\theta}$ where $r = |z|$, $\theta = \tan^{-1}(\frac{b}{a})$

$e^{i\theta} = \cos\theta + i\sin\theta$ (Euler's Formula)

Complex Fourier series :

$f(t) = \sum_{n=-\infty}^{\infty} a_n e^{int}$ where $n =$ frequency

Roots of complex $z \in \mathbb{C}$

if $w^n = z$, then w = n^{th} root of z

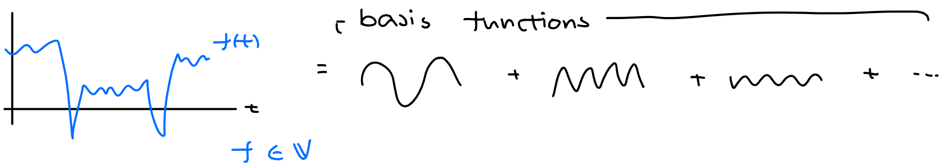
$w = \rho e^{i\alpha}$, $z = re^{i\theta}$

$w^n = \rho^n e^{in\alpha}$

$\Rightarrow$ $w = \sqrt[n]{r}\ \exp\left(i\,\frac{\theta + 2k\pi}{n}\right) \quad (k = 0, 1, 2 \ldots n-1)$

1/27/21

Vector (linear) spaces.



Examples of vector spaces : $\mathbb{R}^n$ finite dimensional

$\mathbb{C}^n$ 2n dimensional

$\mathbb{P}_n = \{n^{th}$ degree polynomials$\}$ = poly

infinite dimensional $\begin{cases} C[0,1] = \{u : [0,1] \to \mathbb{R} \mid u \text{ is continuous}\} \\ L^2(a,b) = \{u : (a,b) \to \mathbb{R} \mid \int_a^b |u(t)|^2\, dt < \infty\} \end{cases}$

Vector Space Axioms :

let $\underline{u, v, w}$ $\in \mathbb{V}$ = some set and let $\alpha, \beta \in \underline{\mathbb{R}}$ or $\underline{\mathbb{C}}$

linearity — vectors — fields

1) $u + (v + w) = (u + v) + w$

2) $u + v = v + u$

3) $\exists\, 0 \in \mathbb{V}$ s.t $u + 0 = u$

4) $\exists (-u) \in \mathbb{V}$ s.t $u + (-u) = 0$

5) $\alpha(\beta u) = (\alpha\beta) u$

6) $(\alpha + \beta) u = \alpha u + \beta u$

7) $\alpha(u + v) = \alpha u + \alpha v$

Defn : If $\{v_1, \ldots, v_n\} \subset \mathbb{V}$ and $\{a_1, \ldots, a_n\} \subset \mathbb{R}$ (or $\mathbb{C}$)

then $\sum_{i=1}^n a_i v_i = a_1 v_1 + \cdots + a_n v_n$ is a linear combination of the $\{v_i\}_{i=1}^n$

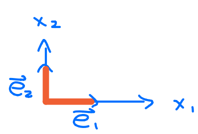
note : n has to be finite

ex : $\sum_{k=1}^{10} a_k \cos(kt)$ is a lin. comb. of $\{\cos(kt)\}_{k=1}^{10}$

Defn : The span of $\{v_1, \ldots, v_n\} \subset \mathbb{V}$ is the set

$\text{span}\{v_1, \ldots, v_n\} = \{\sum_{i=1}^n a_i v_i \mid v_i \in \mathbb{V},\ a_i \in \mathbb{K}\}$ $\quad \mathbb{K} = \mathbb{R}$ or $\mathbb{C}$

ex : $\mathbb{R}^2$, $\vec{e}_1 = \binom{1}{0}$, $\vec{e}_2 = \binom{0}{1}$



$\text{span}\{\vec{e}_1\}$ = entire x_1 axis $= \{a\vec{e}_1 \mid a \in \mathbb{R}\}$

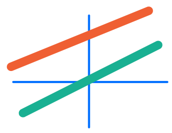
$\text{span}\{\vec{e}_1, \vec{e}_2\} = \mathbb{R}^2 = \{a_1\vec{e}_1 + a_2\vec{e}_2 \mid a_1, a_2 \in \mathbb{R}\} = \binom{a_1}{a_2}$

Subspaces :

ex : span $\{v_i\}_{i=1}^n$ is a subspace of $\mathbb{V}$.

A subspaces of $\mathbb{V}$ is a set $\mathbb{X} \subset \mathbb{V}$ s.t $u + cv \in \mathbb{V}$ $\forall c \in \mathbb{R}$ or $\mathbb{C}$, $u, v \in \mathbb{X}$

- we say $\mathbb{X} \subset \mathbb{V}$ is closed under scalar multiplication and addition.
- A subspace is itself a vector space.

ex: 

← not a subspace - doesn't go through 0.

← Subspace of $\mathbb{R}^2$

ex: $\mathbb{P}_2$ is a subspace of $C[0,1]$

Defn: A set $\{v_1, \ldots, v_n\} \subset \mathbb{V}$ is linearly independent if $\sum_{i=1}^{n} a_i v_i = 0$, then $a_i = 0$ for $i = 1, \ldots, n$. note: n is finite.

ex: $\binom{1}{0}, \binom{2}{0}, \binom{4}{0}$ are linearly independent.
since $\binom{1}{0} = \frac{1}{2}\binom{2}{0} + 0\binom{4}{0}$

Defn: A basis B of $\mathbb{V}$ is a linearly independent set of vectors that span $\mathbb{V}$.

ex: $B = \{1, t, t^2\}$ form a basis of $\mathbb{P}_2$.

Defn: The dimension of $\mathbb{V}$ is the number of vectors in any basis.

ex: $\dim(\mathbb{R}^2) = 2$, $\dim(\mathbb{P}_2) = 3$, $\dim(C[0,1]) = \infty$
note: need to be careful in $C[0,1]$
since $\sum_{i=1}^{\infty} a_i v_i$ not allowed in vector space axiom.

1/29/21

Basis in ∞-dim spaces.

Hamel Basis (ignore) - uses finite # of vectors.

Schauder Basis:

B is a schauder basis if

i) Any finite set of vectors of B is linearly independent.

ii) $\forall \vec{v} \in \mathbb{V}$, $\vec{v} = \sum_{i=1}^{\infty} c_i \vec{v}_i$ where $B = \{\vec{v}_i\}_{i=1}^{\infty}$ (completeness)

interpreted as $\lim_{n \to \infty} \| \vec{v} - \sum_{i=1}^{\infty} c_i \vec{v}_i \|_{\mathbb{V}} = 0$

iii) The coefficients c_i are unique.

example of Schauder basis: Fourier series in $L^2(0, 2\pi) = \mathbb{V}$

$f(t) = \sum_{n=-\infty}^{\infty} a_n e^{int}$

↳ form a Schauder basis $\{1, e^{it}, e^{-it}, e^{2it}, e^{-2it} \ldots\}$

Inner-Product Spaces: $\mathbb{K}^n = \mathbb{R}^n$ or $\mathbb{C}^n$

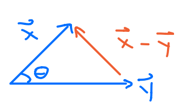
Define inner-product $\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i$ $\quad x, y \in \mathbb{K}^n$ $\quad$ note $z = x + iy$, $\bar{z} = x - iy$

i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

ii) $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ $\quad a, b \in \mathbb{K}$

iii) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ iff $x = 0$

note: $\langle x, ay \rangle = \overline{\langle ay, x \rangle} = \overline{a\langle y, x \rangle} = \bar{a}\,\overline{\langle y, x \rangle} = \bar{a}\langle x, y \rangle$

ex: $\mathbb{R}^2$ 

$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \|\vec{y}\| \cos\theta$

pf: law of cosines $\|\vec{x} - \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\|\vec{x}\|\|\vec{y}\|\cos\theta$

$\langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\|\vec{x}\|\|\vec{y}\|\cos\theta$

$\cancel{\|\vec{x}\|^2} - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, -\vec{x} \rangle + \cancel{\|\vec{y}\|^2} = \cancel{\|\vec{x}\|^2} + \cancel{\|\vec{y}\|^2} - 2\|\vec{x}\|\|\vec{y}\|\cos\theta$

Normed spaces

$(V, \|\cdot\|)$ is a normed space where $\|\cdot\| : V \to \mathbb{R}$ and satisfies

i) $\|u+v\| \le \|u\| + \|v\| \quad \forall u, v \in V$. (triangle inequality)

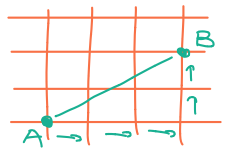
ii) $\|au\| = |a|\,\|u\| \quad a \in \mathbb{K}$

iii) $\|u\| \ge 0$ and $\|u\| = 0$ iff $u = 0$

All inner-products generate a norm: $\|x\| = \sqrt{\langle x, x\rangle}$

Other norms on $\mathbb{R}^n$: let $p \ge 1$, and set $\|\vec{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$

$\|\vec{x}\|_1 = \sum_{i=1}^n |x_i|$ (1-norm, taxicab norm)



Dist btw. A.B in $\|\vec{x}\|_1$ is 5

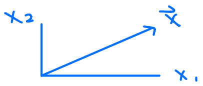
$\|\vec{x}\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ (2-norm)

generated by $\langle x, y\rangle = \sum_{i=1}^n x_i \bar{y}_i$

$\|\vec{x}\|_\infty = \max\limits_{i=1,\dots,n} |x_i|$ (∞-norm, max norm)

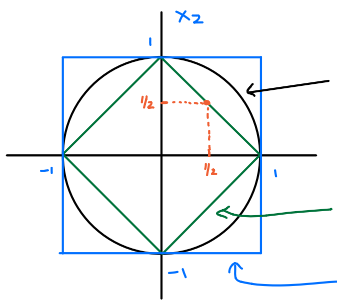
note $\|\vec{x}\|_\infty = \lim\limits_{p\to\infty} \|\vec{x}\|_p$

Ex:



$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2} = \left(\sum_{i=1}^2 x_i^2\right)^{1/2}$ (Euclidian norm)

Unit spheres in various norms $\mathbb{R}^2$



$\|\vec{x}\|_2 = 1 \Rightarrow \|x\|_2^2 = 1$

$x_1^2 + x_2^2 = 1$

$\|\vec{x}\|_1 = \sum_{i=1}^2 |x_i| = |x_1| + |x_2| = 1$

$\|\vec{x}\|_\infty = \max\limits_{i=1,2} |x_i| = 1$

2/1/21

Equivalent Norms

we say any two norms $\|x\|_\alpha$, $\|x\|_\beta$ are equivalent if $\exists A, B > 0$ s.t $A\|x\|_\alpha \le \|x\|_\beta \le B\|x\|_\alpha$ (norm are comparable)

Theorem: In finite dimensional spaces, ALL norms are equivalent.

Cauchy-Schwarz Inequality

For any $x, y \in V$ (V inner-product space)

we have $|\langle x, y\rangle| \le \|x\|\|y\|$ recall $\|x\| = \sqrt{\langle x, x\rangle}$

pf: If $x = 0$ or $y = 0$, done.

so, suppose $x, y \ne 0$. suppose $V = \mathbb{R}$

$\underbrace{0 \le \|x - ty\|^2}_{\text{constant}} = \langle x - ty, x - ty\rangle = \underbrace{\|x\|^2}_{= c} - \underbrace{2t\langle x, y\rangle}_{= 2tb} + \underbrace{t^2\|y\|^2}_{= t^2 a}$

$= c - 2tb + at^2$ (quadratic in t)

$t = \frac{2b \pm \sqrt{(-2b)^2 - 4ac}}{2a}$ want $(-2b)^2 - 4ac \le 0$

$b^2 \le ac$ ($b^2 = |\langle x, y\rangle|^2$, $a = \|y\|^2$, $c = \|x\|^2$) $\Rightarrow |\langle x, y\rangle| \le \|x\|\|y\|$

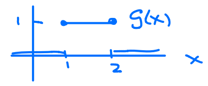
The vector space $L^2[a,b]$

$L^2[a,b] := \{ f : [a,b] \to \mathbb{C} \mid \int_a^b |f(x)|^2 \, dx < \infty \}$

this is an inner-product space with $\langle f, g \rangle = \langle f, g \rangle_{L^2[a,b]} := \int_a^b f(x) \overline{g(x)} \, dx$

an norm $\|f\|_{L^2[a,b]} = \|f\|_{L^2} = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$

ex: $g \in L^2(\mathbb{R})$



$\|g\|_{L^2(\mathbb{R})} = \left(\int_{\mathbb{R}} |g(x)|^2 dx \right)^{1/2} = \left(\int_1^2 1^2 \, dx \right)^{1/2} = 1$

Cauchy – Schwarz in L^2:

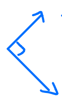
$|\langle f, g \rangle| \le \|f\| \|g\|$

$$\left| \int_a^b f(x) \overline{g(x)} \, dx \right| \le \left(\int_a^b |f|^2 dx \right)^{1/2} \left(\int_a^b |g|^2 dx \right)^{1/2}$$

Orthogonality

let $u, v \in V$, we say u, v are orthogonal if $\langle u, v \rangle = 0$

ex:



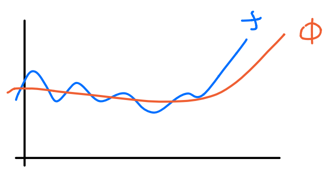
$\vec{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\vec{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\langle \vec{x}, \vec{y} \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = 1 - 1 = 0$

note $\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\|_2 \|\vec{y}\|_2 \cos\theta = 0$ $\quad \theta = \frac{\pi}{2}$

plus if $\|u\|, \|v\| = 1$, then u, v are orthonormal.



- $\|f - \phi\|_2$ is the average distance btw f, ϕ.
- $\|f - \phi\|_\infty$ is a distance that tries to minimize the maximum dist.

note $L^3[a,b] = \{ f : [a,b] \to \mathbb{R} \mid \int_a^b |f(x)|^3 dx < \infty \}$

$\|f\|_{L^3} = \left(\int_a^b |f(x)|^3 dx \right)^{1/3}$

while $\|\vec{x}\|_3 = \left(\sum_{i=1}^n |x_i|^3 \right)^{1/3}$

Orthonormal Bases

Suppose $B = \{e_1, e_2, \dots, e_n\} \subset V$ is an orthonormal basis of V.

Have $\langle e_i, e_j \rangle = 0$ unless $i = j$ in which case $\langle e_i, e_j \rangle = \|e_i\|^2 = 1$

Kronecker Delta: $\langle e_i, e_j \rangle = \delta_{ij} = \delta(i-j) = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$

and $\forall u \in V$, $u = \sum_{i=1}^n c_i e_i$

use orthogonality: $\langle u, e_j \rangle = \langle \sum_{i=1}^n c_i e_i, e_j \rangle = \sum_{i=1}^n c_i \underbrace{\langle e_i, e_j \rangle}_{= \delta_{ij}} = c_j$

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

2/3/21

Recall:

Orthonormal basis $\{e_k\}_{k=1}^n \subset V$.

$u \in V$, $u = \sum_{k=1}^n \langle u, e_k \rangle e_k$

$\langle u, u \rangle = \langle \sum_{k=1}^n \langle u, e_k \rangle e_k, u \rangle$

$\|u\|^2 = \sum_{k=1}^n \langle u, e_k \rangle \underbrace{\langle e_k, u \rangle}_{= \overline{\langle u, e_k \rangle}} = \sum_{k=1}^n |\langle u, e_k \rangle|^2 = \|u\|^2$

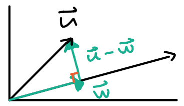
Each $\langle u, e_k \rangle e_k$ is an orthogonal projection of u onto e_k

$P_{e_k} u = \mathrm{Proj}_{e_k} u = \langle u, e_k \rangle e_k$ $\quad$ - $\{e_k\}_{k=1}^n$ is orthonormal here

If $M = \text{span}\{e_1, \dots, e_m\}$ then $P_M u = \text{Proj}_M u = \sum_{k=1}^{m} \langle u, e_k \rangle e_k$

└ set   └ $m \in \mathbb{Z}^+ = \mathbb{N}$ (index)

In $\mathbb{R}^2$



$\vec{w} = \text{proj}_{\vec{v}} \vec{u}$ have $\vec{w} = \lambda \vec{v}$ for some $\lambda \in \mathbb{R}$

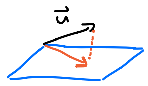
$0 = \langle \vec{u} - \vec{w}, \vec{v} \rangle = \langle \vec{u} - \lambda \vec{v}, \vec{v} \rangle = \langle u, \vec{v} \rangle - \underbrace{\langle \lambda \vec{v}, \vec{v} \rangle}_{\lambda \|\vec{v}\|^2}$

$\lambda = \langle \vec{u}, \vec{v} \rangle / \|\vec{v}\|^2$

$\vec{w} = \lambda \vec{v} = \dfrac{\langle \vec{u}, \vec{v} \rangle \vec{v}}{\|\vec{v}\|^2}$ when $\|\vec{v}\| = 1$, $\vec{w} = \langle \vec{u}, \vec{v} \rangle \vec{v}$

If $\{e_k\}_{k=1}^{n}$ is on orthogonal basis then $u = \sum_{k=1}^{n} \dfrac{\langle u, e_k \rangle}{\|e_k\|^2} e_k$

P_m is a projection if $P_m(P_m) = P_m^2 = P_m$



let's check for $P_M u = \sum_{k=1}^{m} \langle u, e_k \rangle e_k$

$P_M(P_M u) = \sum_{k=1}^{m} \langle P_m u, e_k \rangle e_k$

$= \sum_{k=1}^{m} \langle \underbrace{\textstyle\sum_{j=1}^{m} \langle u, e_j \rangle e_j, e_k}_{} \rangle e_k$

$= \sum_{j=1}^{m} \langle u, e_j \rangle \underbrace{\langle e_j, e_k \rangle}_{= \delta_{jk}} = \langle u, e_k \rangle$

$= \sum_{k=1}^{m} \langle u, e_k \rangle e_k$

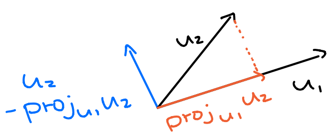
$= P_M u$

Consider $\{1, t\} \subset L^2[0,1] = \mathbb{V}$

$\langle 1, t \rangle = \int_0^1 1 \cdot t \, dt = 1/2 \neq 0 \implies 1, t$ not orthogonal

How to orthogonalize ?

$\{u_1, u_2, \dots, u_n\}$ non-orthogonal want to turn into $\{v_1, v_2 \dots v_n\}$ orthogonal.



Now $\{\underbrace{u_1}_{=v_1}, \underbrace{u_2 - \text{proj}_{u_1} u_2}_{= v_2}\}$ is an orthogonal set spanning the same set as $\{u_1, u_2\}$

Gram - Schmidt Process

To turn $\{u_i\}_{i=1}^{n}$ non-orthogonal set into $\{v_i\}_{i=1}^{n}$ orthogonal.

$v_1 = u_1$

$v_2 = u_2 - \text{proj}_{v_1} u_2$

$v_3 = u_3 - \text{proj}_{\text{span}(v_1, v_2)} u_3 = u_3 - \text{proj}_{v_1} u_3 - \text{proj}_{v_2} u_3$

$\vdots$

$v_n = u_n - \text{proj}_{\text{span}(v_1 \dots v_{n})} u_n$

we have $\text{span}\{u_1, \dots, u_n\} = \text{span}\{v_1 \dots v_n\}$, $\{v_1 \dots v_n\}$ is orthogonal,

we can normalize by $\left\{ \dfrac{v_1}{\|v_1\|}, \dfrac{v_2}{\|v_2\|}, \dots, \dfrac{v_n}{\|v_n\|} \right\}$

2/5/21

Linear Operators (map , function)

$T: \mathbb{V} \to \mathbb{W}$ is a linear map if $\forall c \in \mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) and $\forall u, v \in \mathbb{V}$.

i) $T(u+v) = T(u) + T(v)$

ii) $T(cu) = cT(u)$

Ex: $T: \mathbb{R}^3 \to \mathbb{R}^2$

$T(\vec{x}) = A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ A = linear operator (Matrix)

Ex: $T: L^2(0,1) \to \mathbb{R}$

$T(u) = \int_0^1 u(t) e^t \, dt$

$T(au+bv) = \int_0^1 (au+bv)(t) e^t \, dt = a T(u) + b T(v)$

Ex: $T: \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, $x \in \ell^2(\mathbb{N})$, $x = \{x_1, x_2 \dots\}$

$T(x) = \{ \frac{1}{2} x_1, \frac{1}{2^2} x_2, \frac{1}{2^3} x_3, \dots, \frac{1}{2^n} x_n, \dots \}$

$T(ax+by) = \{ \frac{1}{2}(ax_1+by_1), \frac{1}{2^2}(ax_2+by_2) \dots \}$

$= a\{ \frac{1}{2}x_1, \frac{1}{2^2}x_2, \dots \} + b\{ \frac{1}{2}y_1, \frac{1}{2^2}y_2, \dots \}$

$= a T(x) + b T(y)$

In finite dimensions, any linear map $T: V \to W$ can be represented as matrix multiplication. where V, W are vector spaces.

Let $\{V_1 \dots V_n\}$ be a basis of V and let $\{w_1 \dots w_m\}$ be a basis of W.

Note $T(V_j) \in W \quad \forall j = 1, \dots, n$

$\Rightarrow T(V_j) = \sum_{i=1}^m a_{ij} w_i$

let $v = \sum_{j=1}^n x_j V_j \in V$

Then $T(v) = T\left(\sum_{j=1}^n x_j v_j\right) = \sum_{j=1}^n x_j T(v_j) = \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij} w_i \right) = \sum_{i=1}^m \overbrace{\sum_{j=1}^n a_{ij} x_j}^{= c_i} w_i = \sum_{i=1}^m c_i w_i$

Here $c_i = \sum_{j=1}^n a_{ij} x_j$ is precisely matrix-vector multiplication : $\vec{c} = A\vec{x}$

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Think of $T(\vec{v}) = A\vec{v}$

The column of A store the coefficients of $T(V_j)$ $j = 1 \dots n$

where $\{V_j\}_{j=1}^n$ is a basis of V.

Ex: Find the matrix representation of $T: V \to W$ where $V = \mathbb{R}^2$, $W = \text{span}\{\cos t, \sin t\}$. and $T(\vec{v}) = (v_1 + 2v_2)\cos t$, $v_2 \sin t$

basis for V let's use $\{\vec{e}_1, \vec{e}_2\} = \{\binom{1}{0}, \binom{0}{1}\}$

and for W let's use $\{w_1, w_2\} = \{\cos t, \sin t\}$

Have $T(\vec{e}_1) = T\binom{1}{0} = 1 \cdot \cos t + 0 \cdot \sin t$

$T(\vec{e}_2) = T\binom{0}{1} = 2 \cdot \cos t + 1 \cdot \sin t$ $\Rightarrow A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

$\Rightarrow \vec{v} = x_1 \binom{1}{0} + x_2 \binom{0}{1} \in V = \mathbb{R}^2$, $\vec{c} = \binom{c_1}{c_2}$, $\vec{c} = A\vec{x}$.

basis of $\mathbb{R}^2$

So $T(\vec{v}) = c_1 w_1 + c_2 w_2 = c_1 \cos t + c_2 \sin t$

And if used basis $\{\binom{1}{0}, \binom{1}{1}\}$ for $V = \mathbb{R}^2$.

$T\binom{1}{0} = 1 \cdot \cos t + 0 \cdot \sin t$

$T\binom{1}{1} = 3 \cdot \cos t + 1 \cdot \sin t$ $\Rightarrow A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$

Ex: if basis $\{\binom{1}{0}, \binom{0}{1}\}$ of $V = \mathbb{R}^2$, basis $\{w_1, w_2\} = \{\cos t + \sin t, \sin t\}$ of W.

then $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$ $\vec{v} = x_1 \binom{1}{0} + x_2 \binom{0}{1}$, $T(\vec{v}) = c_1(\cos t + \sin t) + c_2 \sin t$

2/8/21

Adjoint Operators

$T: V \to W$ linear operator (map)

inner-product spaces with $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle_W$

Can we write $\langle Tv, w \rangle_W$ as $\langle v, T^* w \rangle_V$ for some new operator T^* the adjoint of T.

Tv or $T(v)$ denotes action of T on v

i.e. $T^*: W \to V$ s.t $\langle Tv, w \rangle_W = \langle v, T^* w \rangle_V$

Goal: Try to write $\langle Tv, w \rangle_W = \langle v, z \rangle_V$ for some $z \in V$ and define $T^* w = z$.

Theorem: If V, W are finite-dimensional and $T: V \to W$ is linear, then $T^*: W \to V$ exists and is unique

If $T = T^*$ we call T self-adjoint, $\langle Tv, w \rangle = \langle v, T^* w \rangle = \langle v, Tw \rangle$

Ex: The Fourier transform $F[f] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{ix^3} dx$ is self-adjoint with the $L^2(\mathbb{R})$ inner-product $\langle f, g \rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} f(x) g(x)\, dx$

so $\langle Ff, g \rangle_{L^2} = \langle f, Fg \rangle_{L^2}$

Ex: $T = I$ $\quad I: V \to V$ (identity map) $\quad Iv = v$

$\langle Iv, w \rangle = \langle v, Iw \rangle$

so $I^* w = Iw \Rightarrow I^* = I$ by definition

Ex: Let $a(t)$ be bounded. $\langle f, g \rangle_{L^2} = \int_a^b f(x) \overline{g(x)}\, dx$

Define $T: L^2(a,b) \to L^2(a,b)$ by $(Tf)(t) = a(t) f(t)$

Find T^*

Have $\langle Tf, g \rangle = \langle af, g \rangle = \int_a^b a(t) f(t) \overline{g(t)}\, dt = \int_a^b f(t) \cdot \underline{a(t) \overline{g(t)}}\, dt$

$= \int_a^b f(t) \overline{\overline{a(t)} g(t)}\, dt$ $\qquad = \overline{\overline{a(t)}\, g(t)}$

$= \langle f, \bar{a} g \rangle$ $\qquad$ so $(T^* g)(t) = \overline{a(t)}\, g(t)$

If $a(t)$ real-valued, then $T^* = T$ and T self-adjoint.

Integration by Parts:

$$\int_a^b f(t) g'(t)\, dt = f(t) g(t) \Big|_{t=a}^{t=b} - \int_a^b f'(t) g(t)\, dt$$

Ex: $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = A\vec{x} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

Fact: $A^* = \bar{A}^T$ so $A^* = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}$ $\qquad A^*: \mathbb{R}^2 \to \mathbb{R}^3$

$\langle A \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \rangle_2 = \langle \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \rangle_2 = \langle \begin{pmatrix} x+y \\ x+2z \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \rangle_2 = (x+y)a + (x+2z)b$

$= x(a+b) + ya + z(2b) = \langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} a+b \\ a \\ 2b \end{pmatrix} \rangle_3$

we define $A^* \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ a \\ 2b \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$ $\qquad A: \mathbb{R}^3 \to \mathbb{R}^2$, $A^* = \mathbb{R}^2 \to \mathbb{R}^3$

$\langle A\vec{x}, \vec{y} \rangle_2 = \langle \vec{x}, A^* \vec{y} \rangle_3$

2/10/21

Intro to Frames

V with orthonormal basis $\{v_1, v_2, v_3, \dots\}$ (ONB)

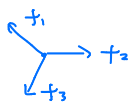
- can be restrictive in practice
- easy to compute c_i s.t $x = \sum c_i v_i$, $c_i = \langle x_i, v_i \rangle$

$\{f_1, f_2, f_3 \dots\}$ is a frame of V if the $\{f_i\}_{i \in I}$ span V and satisfy some other conditions.

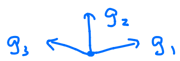
- frame can be linearly dependent.

ex: 

a basis of $\mathbb{R}^2$ (orthonormal)

ex: 

a frame for $\mathbb{R}^2$ (contains redundancies)

ex: 

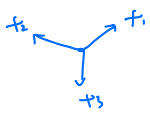
another frame

ex: Reconstruction , orthonormal basis $\{e_1, e_2\}$

person A , $\{1, 2\}$, $x = 1 \cdot e_1 + 2 \cdot e_2$.

↓ " erasure" " information loss" , lose all info. in the e_2 dire.

person B , $\{1, ?\}$, $x = 1 \cdot e_1 + ?$



person A , $\{1, 2, 3\}$, $x = 1 \cdot f_1 + 2 \cdot f_2 + 3 \cdot f_3$

↓ retains more info. than with ONB and sometimes can get " perfect reconstruction "

person B , $\{1, 2, ?\}$, $x = 1 \cdot f_1 + 2 \cdot f_2 + ?$

Primer on Hilbert Spaces

A Hilbert space $\mathbb{H}$ is a complete inner-product space.

- also normed $\|x\|_{\mathbb{H}} = \sqrt{\langle x, x \rangle_H}$

Defn: $\{x_n\}_{n=1}^{\infty} \subset \mathbb{H}$ is a Cauchy-sequence if
$\forall \varepsilon > 0 \ \exists N \in \mathbb{N}$ s.t $\forall n, m \geq N$, $\|x_n - x_m\|_H < \varepsilon$

Defn: $\mathbb{H}$ is complete if every Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ converges to a point $x \in \mathbb{H}$.

i.e. $\lim_{n \to \infty} x_n = x \in \mathbb{H}$, $\|x_n - x\|_H \to 0$ as $n \to \infty$

Hilbert spaces : $\mathbb{R}, \mathbb{R}^n, \mathbb{C}^n, \ell^2(\mathbb{Z}), L^2(a,b), L^2(\mathbb{R})$
any finite dimensional inner-product space i.e. P_n

Non-Hilbert spaces : $S = (0,1) \subseteq \mathbb{R}$, $x_n = 1/n$ for $n \geq 2$ converges to $0 \notin S$

Defn: H is separable if it contains a countable basis.

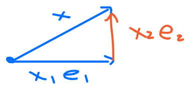
- All examples above are separable.

For ONB, of a Hilbert space H, $x = \sum_{i \in I} \langle x, e_i \rangle e_i \quad \forall x \in H$

Dotting with x we get

$\|x\|^2 = \langle x, \sum_{i\in I} \langle x, e_i\rangle e_i \rangle = \sum_{i\in I} \overline{\langle x, e_i\rangle} \langle x, e_i\rangle = \sum_{i\in I} |\langle x, e_i\rangle|^2$

└ parseval equality

ex : 

$\mathbb{R}^2$-case : $\|x\|^2 = x_1^2 + x_2^2 = |\langle x, e_1\rangle|^2 + |\langle x, e_2\rangle|^2 = \sum_{i=1}^{2} |\langle x, e_i\rangle|^2$

pythegorean theorem

Frames generzlize Parseval equality :

Defn : $\{f_i\}_{i\in I}$ is a frame for $\mathbb{H}$ if I is countable and if $\exists\, 0 < A \le B < \infty$

s.t $\forall x \in \mathbb{H}$ $\quad A\|x\|^2 \le \sum_{i\in I} |\langle x, f_i\rangle|^2 \le B\|x\|^2$

frame bounds.

- largest A and smallest B that work are called optimal frame bounds.
- For $A = B = 1$, equality holds and we call $\{f_i\}_{i\in I}$ a Parseval frame

  i.e. an ONB is parseval
- $\{f_i\}_{i\in I}$ is a tight frame if $A = B$

2/12/21

Recall : $\{e_i\}_{i\in I}$ ONB : $\quad x = \sum_{i\in I} \langle x, e_i\rangle e_i$

Parseval equality : $\|x\|^2 = \sum_{i\in I} |\langle x, e_i\rangle|^2$

frames $\{f_i\}_{i\in I}$ satisfy $\quad A\|x\|^2 \le \sum_{i\in I} |\langle x, f_i\rangle|^2 \le B\|x\|^2$

note : An OBE is a frame , Let $A = B = 1$

union of N orthonormal bases is a frame

call these $\{e_i^{(j)}\}_{i\in I}$ $\quad j = 1, \dots, N$ $\quad$ every j is a different ONB

 bases of $\mathbb{R}^2$, union is a frame of $\mathbb{R}^2$

note by Perseval , $\forall j = 1, \dots, n$. $\|x\|^2 = \sum_{i\in I} |\langle x, e_i^{(j)}\rangle|^2$

$\Rightarrow \sum_{j=1}^{N} \|x\|^2 = \sum_{j=1}^{N} \sum_{i\in I} |\langle x, e_i^{(j)}\rangle|^2$

$\Rightarrow N\|x\|^2 = \sum_{j=1}^{N} \sum_{i\in I} |\langle x, e_i^{(j)}\rangle|^2$

$A = B = N$ and $\{e_i^{(j)}\}$ is a tight frame

If we have just $\{f_1\} = \{\binom{1}{0}\}$ for $\mathbb{R}^2$,

then no A exists s.t $A\|x\|^2 \le \sum_{i=1}^{1} |\langle x, f_i\rangle|^2 = |\langle x, f_1\rangle|^2$

because $|\langle x, f_1\rangle|^2 = x_1^2 \overset{?}{\ge} A\|x\|^2 = A(x_1^2 + x_2^2)$

If such an A existed , let $x_2 \to \infty$ to get contradiction

Lemma : Young's Inequality :

if $a \ge 0$, $b \ge 0$, $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$

then $ab \le \frac{a^p}{p} + \frac{b^q}{q}$

If $p = q = 2$, then $ab \le \frac{a^2}{2} + \frac{b^2}{2} \Rightarrow 2ab \le a^2 + b^2$

i.e. $(x_1 \pm x_2)^2 \ge 0 \Rightarrow x_1^2 + x_2^2 \pm 2x_1x_2 \ge 0$

$\Rightarrow \|x\|^2 \pm 2x_1x_2 \ge 0$

$\|x\|^2 \ge 2x_1x_2 \quad \& \quad 2x_1x_2 \ge -\|x\|^2$

Ex: $\{f_i\}_{i=1}^3 = \{f_1 = \binom{1}{0},\ f_2 = \binom{0}{1},\ f_3 = \binom{1}{1}\} \subset \mathbb{R}^2$

Goal: $\exists A, B$. s.t $A\|x\|^2 \le \sum_{i=1}^3 |\langle x, f_i\rangle|^2 \le B\|x\|^2$, $\forall x \in \mathbb{R}^2$

$$\sum_{i=1}^3 |\langle x, f_i\rangle|^2 = x_1^2 + x_2^2 + (x_1+x_2)^2 = 2(x_1^2+x_2^2) + 2x_1x_2$$
$$= 2\|x\|^2 + \underbrace{2x_1x_2}_{\le \|x\|^2} \le 3\|x\|^2$$

$$\sum_{i=1}^3 |\langle x, f_i\rangle|^2 = 2\|x\|^2 + \underbrace{2x_1x_2}_{\ge -\|x\|^2} \ge \|x\|^2$$

thus $\|x\|^2 \le \sum_{i=1}^3 |\langle x, f_i\rangle|^2 \le 3\|x\|^2$

Bounds $A=1$, $B=3$ are optimal,

to get equality, choose $x = \binom{x_1}{x_2}$ s.t $\{f_i\}_{i=1}^3$ is not tight

Now will develop theory about frames:

ONB: $x = \sum_{i\in I} \langle x, e_i\rangle e_i$ $\quad \{e_i\}_{i\in I} \subset H$

define $\langle x, e_i\rangle := T(x)$ takes x and produces $\langle x, e_i\rangle \in \mathbb{C}$, $T: \mathbb{H} \to \ell^2(I)$

$\sum_{i\in I} \langle x, e_i\rangle e_i := S(\langle x, e_i\rangle) = S(c_i)$ if $c_i := \langle x, e_i\rangle$, $S: \ell^2(I) \to \mathbb{H}$

we call $T(x) = \{\langle x, e_i\rangle\}_{i\in I}$ the analysis operator, $T: \mathbb{H} \to \ell^2(I)$

we call $S(c) = T^*(c) = \sum_{i\in I} c_i e_i$ the synthesis operator

$c = \{c_i\}_{i\in I}$ and $T^*: \ell^2(I) \to \mathbb{H}$

Same definition holds for frames $\{f_i\}_{i\in I}$

2/15/21

Check $T: \mathbb{H} \to \ell^2(I)$ $\qquad T(x) = \{\langle x, f_i\rangle\}_{i\in I}$

Have $\|\{\langle x, f_i\rangle\}_{i\in I}\|^2_{\ell^2(I)} := \sum_{i\in I} |\langle x, f_i\rangle|^2 \le B\|x\|^2_{\mathbb{H}} < \infty$

since $\{f_i\}$ are frame

$\Rightarrow T(x) \in \ell^2(I)$

Defn: A linear operator $T: X \to Y$ is bounded if $\exists c > 0$

s.t $\|T(x)\|_Y \le c\|x\|_X$ $\quad \forall x \in X$.

thus the analysis operator is bounded $\|T(x)\|_{\ell^2(I)} \le \sqrt{B}\|x\|_{\mathbb{H}}$

Adjoint of Analysis operator T

$$\langle \underbrace{T(x)}_{=\{\langle x, f_i\rangle\}_{i\in I}}, c\rangle_{\ell^2(I)} = \sum_{i\in I} \langle x, f_i\rangle_{\mathbb{H}} \bar{c}_i = \langle x, \underbrace{\sum_{i\in I} f_i c_i}_{=:\, T^*c}\rangle_{\mathbb{H}}$$

$T^*c = \sum_{i\in I} f_i c_i$ synthesis operator

Frame Operator: $S := T^*T: \mathbb{H} \to \mathbb{H}$, $\quad T^*: \ell^2 \to \mathbb{H}$, $T: \mathbb{H} \to \ell^2$

$S(x) = \sum_{i\in I} \langle x, f_i\rangle_{\mathbb{H}} f_i$

Now $\langle S(x), x\rangle_{\mathbb{H}} = \sum_{i\in I} \langle x, f_i\rangle \underbrace{\langle f_i, x\rangle}_{=\overline{\langle x, f_i\rangle}} = \sum_{i\in I} |\langle x, f_i\rangle|^2$

$A\|x\|^2 \le \langle S(x), x\rangle \le B\|x\|^2$

Remark: $S = T^*T$ is self-adjoint since

$S^* = (T^*T)^* = T^*(T^*)^* = T^*T = S$

Ex: 2.1.4 in Czaja's notes

finite frame $\{f_i\}_{i=1,\dots,n} \subset \mathbb{R}^d$ $\quad f_1 = \binom{a}{b}, \dots f_n$

Analysis operator $T(x) = \{\langle x, f_i\rangle\}_{i=1}^n$ where $x \in \mathbb{R}^d$

we have $T: \mathbb{R}^d \to \mathbb{R}^n$ and T has an $n \times d$ matrix representation $T(x) = Ax$ $\forall x \in \mathbb{R}^d$.

Let's find A:

suppose x, f_i are all column vectors.

T takes $x \in \mathbb{R}^d$ and maps it to $\{\underbrace{\langle x, f_i\rangle}\}_{i=1}^n \in \mathbb{R}^n$

note $\langle x, f_i\rangle = f_i^* x$

$*$ denotes conjugate transpose $A^* := \bar{A}^T$ $\quad$ recall $\vec{a}\cdot\vec{b} = \vec{a}^T\vec{b}$

$\underbrace{\binom{1}{0}\cdot\binom{2}{3}}_{2} = \underbrace{(1 \ \ 0)\binom{2}{3}}_{2}$

So $T(x) = \begin{pmatrix} f_1^* x \\ f_2^* x \\ \vdots \\ f_n^* x \end{pmatrix} = \underbrace{\begin{pmatrix} f_1^* \\ f_2^* \\ \vdots \\ f_n^* \end{pmatrix}}_{:= A_{n\times d}} x = Ax$

and $T = \begin{pmatrix} f_1^* \\ \vdots \\ f_n^* \end{pmatrix}$ analysis operator

Ex: $\{f_1, f_2, f_3\} = \{\binom{1}{0}, \binom{1}{1}, \binom{4}{2}\} \subset \mathbb{R}^2$ $\quad$ Have $T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 4 & 2 \end{pmatrix}$

$\forall x \in \mathbb{R}^2$, $T(x) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 4 & 2 \end{pmatrix}\binom{x_1}{x_2} = \begin{pmatrix} (1\ 0)\binom{x_1}{x_2} \\ (1\ 1)\binom{x_1}{x_2} \\ (4\ 1)\binom{x_1}{x_2} \end{pmatrix} = \begin{pmatrix} \langle x, f_1\rangle \\ \langle x, f_2\rangle \\ \langle x, f_3\rangle \end{pmatrix}$

Frame operator: $S = T^* T$

$= (f_1, f_2 \dots f_n)\begin{pmatrix} f_1^* \\ \vdots \\ f_n^* \end{pmatrix}$

$= f_1 f_1^* + f_2 f_2^* + \cdots + f_n f_n^*$

$= \sum_{k=1}^n f_k f_k^*$

componentwise, $S_{ij} = \sum_{k=1}^n f_k(i)\, \overline{f_k(j)}$

2/17/21

Ex. $\{f_1, f_2\} = \{\binom{1}{0}, \binom{1}{1}\} \subset \mathbb{R}^2$

Find A, B s.t $A\|x\|^2 \le \underbrace{\sum_{i\in I} |\langle x, f_i\rangle|^2}_{\langle S(x), x\rangle_H} \le B\|x\|^2$

Let $\{\lambda_i\}_{i=1}^2$ be the eigenvalue of $S = T^*T$

Then $A = \min_i \lambda_i$ $\quad B = \max_i \lambda_i$

$T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $T^* = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad S = T^*T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ frame operator

Recall the eigenvalue of S is a number λ_i s.t $Se_i = \lambda_i e_i$,

and eigenvector e_i to

In this question $\lambda_1, \lambda_2 = \frac{3 \pm \sqrt{5}}{2}$ $\quad A = \frac{3-\sqrt{5}}{2}$, $B = \frac{3+\sqrt{5}}{2}$

note: Eigenvalues always positive since S is a positive operator meaning

$S \geq A$ or $\langle S(x), x\rangle \geq A\langle x, x\rangle = A\|x\|^2 > 0$ $\quad \forall x \in H$

Let $\{(\lambda_i, e_i)\}_{i=1}^n$ be eigen pairs of S. let $\lambda_{min} \leq \lambda_i \leq \lambda_{max}$ $\forall i$

Since S is self-adjoint, we can choose $\{e_i\}$ to be orthonormal.

$\forall x \in \mathbb{R}^n$, $x = \sum_{i=1}^n \langle x, e_i \rangle e_i$

$$S(x) = \sum_{i=1}^n \langle x, e_i \rangle \underbrace{S(e_i)}_{\lambda_i e_i}$$ since S is linear. since e_i are eigenvectors

$$S(x) = \sum_{i=1}^n \langle x, e_i \rangle \lambda_i e_i$$

$$\langle S(x), x \rangle = \sum_{i=1}^n \langle x, e_i \rangle \lambda_i \langle e_i, x \rangle$$
$$= \underbrace{\sum_{i=1}^n |\langle x, e_i \rangle|^2}_{\|x\|^2} \lambda_i$$ parsevals equality

$$\lambda_{min} \|x\|^2 \leq \langle S(x), x \rangle \leq \lambda_{max} \|x\|^2$$

Theorem: (Tight Frames) $A = B$

let $\{f_i\}_{i\in I} \subset \mathbb{H}$ be a frame for $\mathbb{H}$.

then $\underset{\text{frame operator}}{S} = A \cdot \underset{\text{identity operator}}{Id}$ iff $\{f_i\}_{i\in I}$ is a tight frame w/ A frame bound.

Lemma: let $\|A\|^2 \leq \underbrace{\langle S(x), x \rangle}_{\sum_{i\in I} |\langle x, f_i \rangle|^2} \leq B\|x\|^2$

then S^{-1} satisfies $B^{-1} \cdot Id \leq S^{-1} \leq A^{-1} \cdot Id$

meaning $\langle B^{-1} \cdot Id, x \rangle \leq \langle S^{-1}x, x \rangle \leq \langle A^{-1} \cdot Id, x \rangle$

or $B^{-1} \|x\|^2 \leq \langle S^{-1}x, x \rangle \leq A^{-1} \|x\|^2$

so S^{-1} generates a frame

idea: Have $A \cdot Id \leq S \leq B \cdot Id$

$$S^{-1} \cdot A \cdot Id \leq S^{-1} \cdot S \leq S^{-1} \cdot B \cdot Id$$

$S^{-1} \leq A \cdot Id$ $\quad$ $S^{-1} \geq B^{-1} \cdot Id$

Reconstruction + Dual Frames.

Given $\{\langle x, f_i \rangle\}_{i\in I} \subset \ell^2(I)$, reconstruction of x is possible.

Is $x = \sum_{i\in I} \langle x, f_i \rangle \underset{\text{dual of } f_i}{f_i^*}$

2/19/21

Dual Frames + Reconstruction

Given $\{\langle x, f_i \rangle\}_{i\in I}$, how to reconstruct $x \in \mathbb{H}$?

If $\{e_i\}_{i\in I}$ is an ONB, then $\underset{\text{reconstructed } x}{x} = \sum_{i\in I} \underset{\text{out of these coefficients.}}{\langle x, e_i \rangle} e_i$

let $\{f_i\}_{i\in I} \subset \mathbb{H}$ be a frame of $\mathbb{H}$.

A dual frame of $\{f_i\}_{i\in I}$ is a set of vectors $\{f_i^*\}_{i\in I} \subset \mathbb{H}$

s.t $\forall x \in \mathbb{H}$, $x = \sum_{i\in I} \langle x, f_i \rangle f_i^*$ reconstruction formula.

note, f_i^* is not conjugate transpose.

If $\{f_i\}$ is an ONB, then $f_i^* = f_i$ $\forall i \in I$.

To find f_i^*, using S^{-1} (inverse of the frame operator)

we call $\{S^{-1} f_i\}_{i\in I}$ the canonical dual frame of $\{f_i\}_{i\in I}$

Have $S(x) = (T^*T)(x) = \sum_{i\in I} \langle x, f_i\rangle f_i$

let $y = S(x)$ then $x = S^{-1}(y)$

then $y = \sum_{i\in I} \langle S^{-1}(y), f_i\rangle f_i = \sum_{i\in I} \langle y, S^{-1}(f_i)\rangle f_i$

$\downarrow$

Because S^{-1} is self-adjoint, let $u = Sx$, $v = Sy$

then $\langle S^{-1}u, v\rangle = \langle S^{-1}(Sx), Sy\rangle = \langle x, Sy\rangle = \langle Sx, y\rangle$

$= \langle S(S^{-1}u), S^{-1}v\rangle = \langle u, S^{-1}v\rangle$

then $S(x) = \sum_{i\in I} \langle x, f_i\rangle f_i$

$x = S^{-1}\left(\sum_{i\in I} \langle x, f_i\rangle f_i\right) = \sum_{i\in I} \langle x, f_i\rangle \underbrace{S^{-1}(f_i)}_{= f_i^*}$

thus $y = \sum_{i\in I} \langle y, S^{-1}f_i\rangle f_i$ solves the representation problem.

$x = \sum_{i\in I} \langle x, f_i\rangle S^{-1}f_i$ solves the reconstruction problem.

For Parseval frames $(A = B = 1)$ we have $x = \sum_{i\in I} \langle x, f_i\rangle f_i \quad \forall x \in \mathbb{H}$.

so $f_i^* = f_i$ in this case.

This is a tight frame, so $S = A \cdot \mathrm{Id} = 1 \cdot \mathrm{Id}$.

so $S^{-1} = \mathrm{Id}$, $S^{-1}f_i = (\mathrm{Id}) \cdot f_i = f_i$

we say Parseval frames are self-dual.

Theorem: If $\mathbb{H}$ is a Hilbert space. $\dim(\mathbb{H}) = n$,

and $\{e_i\}_{i\in I}^n \subseteq \mathbb{H}$, with $\|e_i\|_{\mathbb{H}} = 1$.

Then $\{e_i\}_{i\in I}^n$ is a Parseval frame iff $\{e_i\}_{i=1}^n$ is an ONB

Ex: Consider $\{f_1, f_2, f_3\} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}4\\2\end{pmatrix}\right\} \subset \mathbb{R}^2$

To compute $\{S^{-1}f_i\}_{i=1}^3$

$S = T^*T = \begin{pmatrix}18 & 9\\ 9 & 5\end{pmatrix}$ frame operator $\qquad S^{-1} = \begin{pmatrix}5/9 & -1\\ -1 & 2\end{pmatrix}$

$T(x) = \begin{pmatrix}1 & 0\\ 1 & 1\\ 4 & 2\end{pmatrix}$, $T^* = \begin{pmatrix}1 & 1 & 4\\ 0 & 1 & 2\end{pmatrix}$

$f_1^* = S^{-1}f_1 = \begin{pmatrix}5/9 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}5/9\\ -1\end{pmatrix}$

$f_2^* = S^{-1}f_2 = \begin{pmatrix}5/9 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}-4/9\\ 1\end{pmatrix}$

$f_3^* = S^{-1}f_3 = \begin{pmatrix}2/9\\ 0\end{pmatrix}$

$x = \sum_{i=1}^3 \langle x, f_i\rangle f^*$

Ex: Suppose $\langle x, f_3\rangle$ is lost, use dual vectors $\{f_1, f_2\}$ for reconstruction.

$f_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $f_2 = \begin{pmatrix}1\\1\end{pmatrix}$

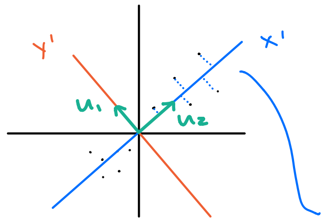
$T = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$, $T^* = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, $S = T^*T = \begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}$, $S^{-1} = \begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}$

So $f_1^* = S^{-1}f_1 = \begin{pmatrix}1\\-1\end{pmatrix}$ and $f_2^* = S^{-1}f_2 = \begin{pmatrix}0\\1\end{pmatrix}$

Get $x = \langle x, f_1\rangle f_1^* + \langle x, f_2\rangle f_2^*$

2/22/21

Principal Component Analysis (PCA)



data can be points, vectors, images, sounds...

Compression, dimensionality reduction

best 1-d subspace (or k-dimensional in general) representing the points.

data can be compressed by projecting onto the x' axis. $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto x'$

- u_1, u_2 are the principal components, ("left-singular vectors" from the SVD)

PCA { covariance matrix
SVD

Data + Mean Centering

data is mean-centered

let $x_i \in \mathbb{R}^m$ $(i=1,\dots,n)$ be some data points.
Often $m \leq n$. Store in   assume x_i column vectors.

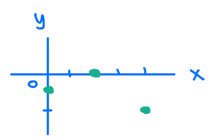
$$X = [x_1 \mid x_2 \mid \cdots \mid x_n] \in \mathbb{R}^{m\times n}$$

We suppose the $\{x_i\}$ is mean-centered.

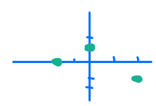
Ex: $X = \begin{bmatrix} 2 & 0 & 4 \\ 0 & -1 & -2 \end{bmatrix}$

let $\bar{x}$ = mean of x-values $= \frac{1}{3}[2+0+4] = 2$

$\bar{y}$ = mean of y-values $= \frac{1}{3}[0-1-2] = -1$

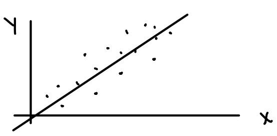


Replace $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$, $X_{centered} = \begin{bmatrix} 0 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}$

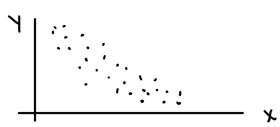


In matrix form, $X_{centered} = X - \underbrace{\left(\frac{1}{n} X \mathbb{1}_n\right)}_{=:\,\bar{x}} \mathbb{1}_n^T = X - \bar{x}\,\mathbb{1}_n^T$ $\mathbb{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n\times 1}$

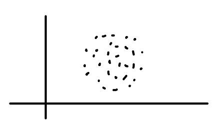
Data Covariance = data spread



- large positive covariance
- larger x → larger y



- large negative covariance
- larger x → smaller y



- low covariance (≈ 0)
- no correlation

Defn: The (sample) covariance matrix is $C_{xx} := \frac{1}{n-1} XX^T \in \mathbb{R}^{m\times m}$
$X \in \mathbb{R}^{m\times n}$

This definition holds if X is mean-centered.

properties: ① C_{xx} is symmetric.
② C_{xx} is positive semi-definite

Ex: $X = \begin{bmatrix} 0 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}$ ← centered $\qquad C_{xx} = \frac{1}{3-1} XX^T = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$

Interpretation of $C_{xx} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$

Suppose we have $X = [x_1 \quad x_2]$. let $x_i^{(k)}$ denote the k^{th} component of x_i

then $C_{xx} = \frac{1}{n-1} XX^T$

$$= \frac{1}{n-1} [x_1 \; x_2] \begin{bmatrix} x_1^T \\ x_2^T \end{bmatrix} = \frac{1}{n-1} [x_1 x_1^T + x_2 x_2^T]$$

$$= \frac{1}{n-1} \begin{pmatrix} {x_1^1}^2 + {x_2^1}^2 & x_1^2 x_1^1 + x_2^2 x_2^1 \\ x_1^1 x_1^2 + x_2^1 x_2^2 & {x_1^2}^2 + {x_2^2}^2 \end{pmatrix}$$

the diagonal terms are variances of the k^{th} components of x_i.

Defn: the (sample) variance of the k^{th} components of x_i is

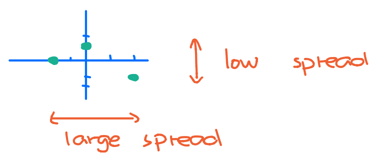
$$\mathrm{Var}(\{x_i^k\}) := \frac{1}{n-1} \sum_{i=1}^n (x_i^k)^2$$

If $\mathrm{Var}(\{x_i^k\})$ is large, then the k^{th} components of x_i are spread out more.

For $C_{xx} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$ know spread (variance) of x_i^1,
is greater than the spread (variance) of x_i^2.

2/24/21

Ex: $X = \begin{bmatrix} 0 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}$ $\quad C_{xx} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$



For the off-diagonal terms in C_{xx},
these measure the correlation between x_i^1 and x_i^2
(weak correlation in this case)
The closer to 0 that the off-diagonal terms are, the less correlation
there is between those components.

Recall $\mathrm{Var}\{x_i^k\} := \frac{1}{n-1} \sum_{i=1}^n (x_i^k)^2$

Defn: $\mathrm{Var}(\{x_i\})$ = total variance of the points $\{x_i\}$

$$:= \frac{1}{n-1} \sum_{i=1}^n \|x_i\|_2^2 = \sum_{i=1}^n \mathrm{Var}\{x_i^k\}$$

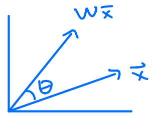
Defn: $\mathrm{tr}(A) = \sum_{i=1}^n (A)_{ii}$ $\qquad$ Ex: $\mathrm{tr}\begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} = 5$

So $\mathrm{Var}(\{x_i\}) = \mathrm{tr}(C_{xx})$ = total variance

Goal: We seek a new representation Y of X with the least correlation
amongst the data components.
we will do this by an orthogonal linear transformation W,
so that $X = WY$

Defn: A square matrix $W \in \mathbb{R}^{n \times n}$ is orthogonal if $W^{-1} = W^T$ (so $W^T W = Id$)
- Equivalently, the columns (and rows) of W form an orthonormal basis of $\mathbb{R}^n$.

Ex:



$W = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is orthogonal, $\|Wx\|_2 = \|x\|$ $\forall x \in \mathbb{R}^n$

$C_{xx} = \frac{1}{n-1} \underset{WY}{X} \underset{Y^T W^T}{X^T} = \boxed{\frac{1}{n-1}} W \boxed{YY^T} W^T = W C_{YY} W^T$

$\underbrace{\frac{1}{n-1} \quad YY^T}_{= C_{YY}}$

$C_{YY} = W^T C_{xx} W$

want a Y that s.t points in Y have the least corr. possible.

Goal now is to choose a W that causes components of points in Y to have the least correlation. Best possible case would be to make C_{YY} diagonal.

Theorem: (Diagonalization)

let $A \in \mathbb{R}^{n \times n}$, TFAE

(i) A is symmetric

(ii) A has an orthonormal set of e-vectors

(iii) A is orthogonally diagonalizable.

i.e. $\exists \Lambda$ diagonal matrix and an orthogonal matrix V s.t. $V^T A V = \Lambda$

note: $V = [v_1 \mid v_2 \mid \dots \mid v_n]$ where v_i are unit e-vectors of A,

and $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_n \end{pmatrix}$ where $\{\lambda_i\}$ are the corresponding e-values.

Now recall C_{xx} is symmetric, so we can orthogonally diagonalize C_{xx}.

$\exists$ V orthogonal matrix and Λ diagonal matrix

s.t $V^T C_{xx} V = \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_n \end{pmatrix}$

suppose $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ are the e-values of C_{xx}

and $V = [v_1 \mid v_2 \mid \dots \mid v_n]$

└ corresponding unit e-vectors of C_{xx}.

these are the principal components of X

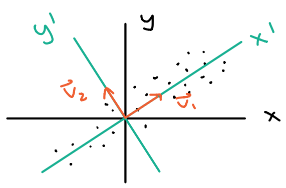
For $C_{YY} = W^T C_{xx} W$, if we let $W = V$

then $C_{YY} = V^T C_{xx} V = \Lambda$ = diagonal matrix

└ covariance matrix, where the correlation amongst point components is minimal.

The change of variables. $X = WY$ is now $X = VY$ or $Y = V^T X$

now Y containing points that represent the initial data from X in a better coordinate system.



X contains $\binom{x}{y}$ points

Y contains $\binom{x'}{y'}$ points.

$\vec{v}_1, \vec{v}_2$ are the principal components.

2/26/21

Recall $X = [x_1 \vdots \cdots \vdots x_n]$ and letting $V = [v_1 \vdots \cdots \vdots v_n]$ be the matrix consisting of e-vector of C_{xx}, then $X = VY$ (or $Y = V^T X$) gives an orthogonal linear transformation that better represents points in the new coordinates in Y.

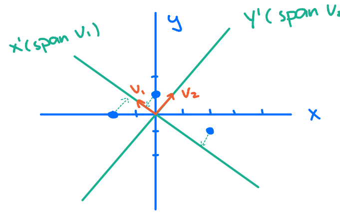
Remarks:

① $C_{YY} = \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, and the eigenvalues $\{\lambda_k\}$ of C_{YY} represent the variances of $\{y_i^k\}$ (the k^{th} components of the data points $\{y_i\}$)

$tr(C_{xx}) = tr(C_{YY})$ total variance is preserved.

② The first principal component v_1 corresponds to the largest e-value of Λ (i.e λ_1) i.e the direction v_1 captures the most variance of the data points.

Ex: $X = \begin{bmatrix} 0 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}$ with $C_{xx} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$



e-pairs of C_{xx} are $(\lambda_1, v_1) = \left(4.3, \begin{pmatrix} -3.3 \\ 1 \end{pmatrix}\right)$

$(\lambda_2, v_2) = \left(0.7, \begin{pmatrix} 0.3 \\ 1 \end{pmatrix}\right)$,

total variance $= 4.3 + 0.7 = 5$.

$4.3/5 = 0.86$, 86% of data is captured by the 1st principal component v_1 direction.

Recall $Y = V^T X$, $V = [v_1 \;\; v_2] = \begin{bmatrix} -0.957 & 0.29 \\ 0.29 & 0.957 \end{bmatrix}$, $C_{YY} = \Lambda = \begin{pmatrix} 4.3 & 0 \\ 0 & 0.7 \end{pmatrix}$

normalize

note $tr(C_{xx}) = tr(C_{YY})$

so $Y = \begin{bmatrix} -0.957 & 0.29 \\ 0.29 & 0.957 \end{bmatrix} \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0.29 & 1.914 & -2.2 \\ 0.957 & -0.58 & -0.38 \end{bmatrix}$

Compression:

- let's project the data points to the line span(v_1)
- we will do this by zeroing out v_2 in $V = [v_1 \;\; \cancel{v_2}]$ 0

Define $\tilde{V} = [v_1 \;\; 0] = \begin{bmatrix} -0.957 & 0 \\ 0.29 & 0 \end{bmatrix}$

$\tilde{Y} = \tilde{V}^T X = \begin{bmatrix} -0.957 & 0.29 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0.29 & 1.914 & -2.204 \\ 0 & 0 & 0 \end{bmatrix}$

To convert the projected points from $\tilde{Y}$ back to the original X coordinate system.

use $V\tilde{Y} = V\tilde{V}^T X =: \underline{\tilde{X}}$ contains points projected to span(v_1) with coordinates in the original X coord. system

note $Y = V^T X$, $X = VY$

$\tilde{X} = V\tilde{Y} = \begin{bmatrix} -0.957 & 0.29 \\ 0.29 & 0.957 \end{bmatrix} \begin{bmatrix} 0.29 & 1.914 & -2.204 \\ 0 & 0 & 0 \end{bmatrix}$

$= \begin{bmatrix} -0.27 & -1.8 & 2.1 \\ 0.08 & 0.55 & -0.6 \end{bmatrix}$