Orthogonality and Orthonormal Bases

$x, y \in \mathbb{R}^n$ col vectors.

$x^T y = x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$
dot product

$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ length of x

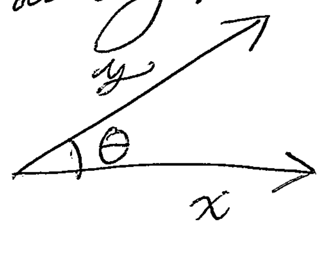
$\|x\| = \sqrt{x \cdot x}$

Fact: For all $x, y \in \mathbb{R}^n$

$$|x \cdot y| \leq \|x\| \, \|y\| \quad \text{(Cauchy–Schwarz inequality)}$$

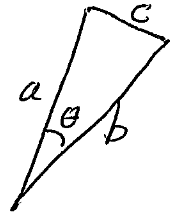
[This is equivalent to saying that $\left|\sum_{i=1}^{n} x_i y_i\right| \leq \left(\sum_{j=1}^{n} x_j^2\right)^{1/2} \left(\sum_{k=1}^{n} y_k^2\right)^{1/2}$]

We will prove this later.
But given this inequality, we can, for $\|x\| \neq 0$ and $\|y\| \neq 0$ define $\cos(\theta)$ for $\measuredangle$ between x and y.

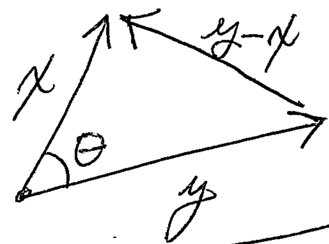


$$\boxed{\frac{x \cdot y}{\|x\| \, \|y\|} = \cos(\theta)}.$$

See this in 3-space via law of cosines:



$$c^2=a^2+b^2-2ab\cos(\theta)$$



$$\|y-x\|^2=\|x\|^2+\|y\|^2-2\|x\|\|y\|\cos(\theta)$$

$$\Rightarrow \|x\|\|y\|\cos(\theta)=\frac{1}{2}\left(\|x\|^2+\|y\|^2-\|y-x\|^2\right)$$

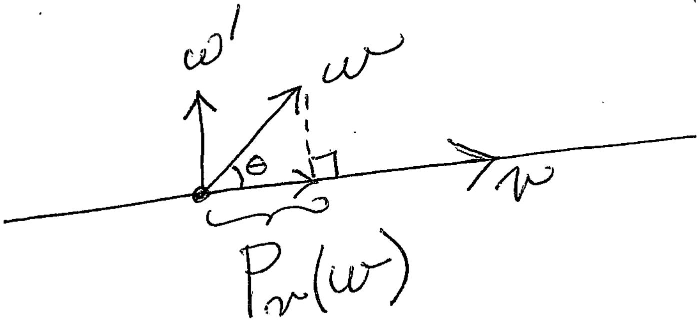
$$=\frac{1}{2}\left(x\cdot x+y\cdot y-(y-x)\cdot(y-x)\right)$$

$$=\frac{1}{2}\left(x\cdot x+y\cdot y-y\cdot y-x\cdot x+2x\cdot y\right)$$

$$=\frac{1}{2}(2x\cdot y) = x\cdot y.$$

Thus $x\cdot y=\|x\|\|y\|\cos\theta.$ //

Projection & Perpendicularity



$$w' = w - P_v(w)$$

$$P_v(w)=\left(\frac{v\cdot w}{\|v\|}\right)\frac{v}{\|v\|} = \left(\frac{v\cdot w}{v\cdot v}\right)v$$

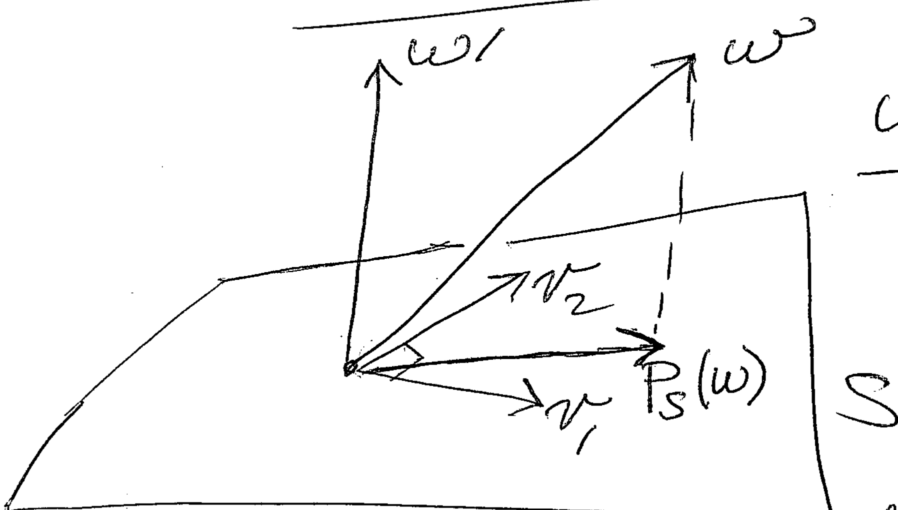
$$w' = w - \left(\frac{v\cdot w}{v\cdot v}\right)v$$

Note: $w'\cdot v = w\cdot v - \left(\frac{v\cdot w}{v\cdot v}\right)v\cdot v = 0$

So $\boxed{w' \perp v}$

Now suppose you have a subspace S with an <u>orthonormal basis</u> $\{v_1, v_2, \ldots, v_K\}$.

This means $v_i \perp v_j$ for $i \neq j$ & $\|v_i\| = 1$ all i.



$$w^{\perp} = w - P_S(w)$$

Suppose $w \notin S$ & define the projection $P_S(w) \in S$ by the equation

$$P_S(w) = \sum_{\ell=1}^{K} P_\ell(w) \quad , \quad P_\ell(w) = P_{v_\ell}(w)$$

<u>Note</u>: $P_\ell(w) = \left(\dfrac{w \cdot v_\ell}{v_\ell \cdot v_\ell}\right) v_\ell$

$\left(\text{since } \|v_\ell\| = 1\right)$ $\quad P_\ell(w) = (w \cdot v_\ell) v_\ell$

Thus
$$P_S(w) = \sum_{\ell=1}^{K} (w \cdot v_\ell) v_\ell .$$

<u>Claim</u> $w' = w - P_S(w) \perp S$.

<u>Proof</u>. It will suffice to show that $w' \cdot v_\ell = 0 \ \forall \ell = 1, 2, \ldots, K$.

$$w' \cdot v_\ell = \left(w - \sum_{i=1}^{K} (w \cdot v_i) v_i\right) \cdot v_\ell$$
$$= w \cdot v_\ell - (w \cdot v_\ell) v_\ell \cdot v_\ell \quad \left(\begin{array}{l}\text{since} \\ v_i \cdot v_\ell = 0 \\ \text{for } i \neq \ell\end{array}\right)$$
$$= w \cdot v_\ell - (w \cdot v_\ell) \quad (\text{sine } v_\ell \cdot v_\ell = 1)$$
$$= 0 \ \checkmark \ //$$

If we now replace w' by $w'' = w' / \|w'\|$, then $\{v_1, \ldots, v_K, w''\}$ is an orthonormal basis for S extended by w.

<u>Theorem</u>. Let $W \subset \mathbb{R}^n$ be any subspace of $\mathbb{R}^n$. Then W has an <u>orthonormal basis</u>.

Subspaces $X, Y \subseteq \mathbb{R}^n$ are said to be <u>orthogonal</u> $(X \perp Y)$ if $x \cdot y = 0$ for all $x \in X$ and $y \in Y$.

<u>Proposition</u>. X, Y subspaces of $\mathbb{R}^n$ & $X \perp Y$, then $X \cap Y = \{0\}$.

<u>Proof</u>. If $v \in X \cap Y$ then $v \in X$ <u>and</u> $v \in Y$. Hence $v \cdot v = 0$. But 0 is the only vector in $\mathbb{R}^n$ such that $v \cdot v = 0$. $\therefore X \cap Y = \{0\}$. //

Let $W \subseteq \mathbb{R}^n$ & define

$$W^{\perp} = \{x \in \mathbb{R}^n \mid x \cdot w = 0 \ \forall w \in W\}$$

$W^{\perp}$ is the <u>orthogonal subspace to W</u>. Note that $W \cap W^{\perp} = \{0\}$ since $W \perp W^{\perp}$. We claim that $v \in \mathbb{R}^n \Rightarrow v = w_1 + w_2$, $w_1 \in W$ & $w_2 \in W^{\perp}$.

To see this, let $v \in \mathbb{R}^n$

$$w_1 = P_W(v)$$

$$w_2 = v - P_W(v).$$

Then we know that $w_2 \perp W$ & hence $w_2 \in W^\perp$.

$$v = w_1 + w_2, \quad w_1 \in W, \ w_2 \in W^\perp.$$

This decomposition of v is <u>unique</u>. Suppose $v = w_1' + w_2'$, $w_1' \in W$, $w_2' \in W^\perp$. Then

$$w_1 + w_2 = w_1' + w_2'$$

$$\Rightarrow \underbrace{w_1 - w_1'}_{\in W} = \underbrace{w_2' - w_2}_{\in W^\perp}$$

$$\Rightarrow w_1 - w_1' \in W \cap W^\perp \text{ and } w_2' - w_2 \in W \cap W^\perp$$

$$\Rightarrow w_1 - w_1' = 0 \quad \text{and } w_2' - w_2 = 0$$

$$\Rightarrow w_1 = w_1' \quad \text{and } w_2' = w_2.$$

QED.

We write $\mathbb{R}^n = W \oplus W^\perp$.

Note that if $W \subseteq \mathbb{R}^n$ has orthonormal basis $\{w_1, \ldots, w_K\}$ & $v \in W$

then $v = a_1 w_1 + \cdots + a_K w_K$

and $v \cdot w_i = a_i (w_i \cdot w_i) = a_i$
(since $w_i \perp w_j$, $i \neq j$)

Thus
$$v = (v \cdot w_1) w_1 + (v \cdot w_2) w_2 + \cdots + (v \cdot w_K) w_K.$$

Will generalize to spaces of functions where we define
$$f \cdot g = K \int_a^b f(x) g(x)\, dx$$
for an approp $\int$, & K.

e.g. $\frac{1}{2\pi} \int_0^{2\pi} f(x) g(x)\, dx.$

$$b_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(nx)\, dx$$

We can choose orthonormal bases $\{w_1, \ldots, w_k\}$ for W & $\{w_1', \ldots, w_\ell'\}$ for $W^\perp$ so that $\{w_1, \ldots, w_k, w_1', \ldots, w_\ell'\} = \beta$ is an orthonormal basis for $\mathbb{R}^n$.

& $k + \ell = n$

Claim. $(W^\perp)^\perp = W$.

Proof. Suppose $v \in \mathbb{R}^n$ and $v \in (W^\perp)^\perp$.

Then $v = a_1 w_1 + \cdots + a_k w_k + b_1 w_1' + \cdots + b_\ell w_\ell'$

We are given that $v \perp W^\perp$.

This means $v \cdot w_i' = 0$ $i = 1, \ldots, \ell$.

But $v \cdot w_i' = b_i$ (easily seen)

$\therefore$ $b_i = 0$ $i = 1, \ldots, \ell$

& so $v = a_1 w_1 + \cdots + a_k w_k \in W$.

This shows that

$(W^\perp)^\perp \subseteq W$.

But clearly $W \subseteq (W^\perp)^\perp$.

$\therefore$ $W = (W^\perp)^\perp$ //

A an $m \times n$ matrix

$N(A) = \{v \in \mathbb{R}^n \mid Av = 0\}$ Let $A = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$ rows

$R(A^T) = \operatorname{Span}\{r_1^T, \ldots, r_m^T\}$ $A^T = [r_1^T \cdots r_m^T]$ cols

Note: $N(A) = \left\{ v \in \mathbb{R}^n \;\middle|\; \begin{array}{l} r_i \cdot v = 0 \\ i = 1, \ldots, m \end{array} \right\}$

$$N(A) = R(A^T)^{\perp}$$

or we can write $N(A)^{\perp} = R(A^T)$

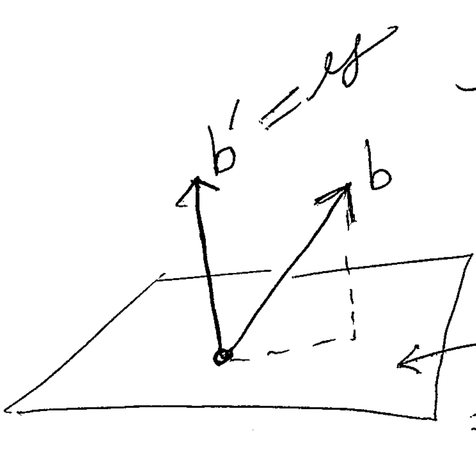
or $N(A^T)^{\perp} = R(A)$

Note: $Ax = b$ consistent iff $b \in R(A)$.

iff $b \in N(A^T)^{\perp}$.

Thus

Corollary 5.2.5 (p219)	A $m \times n$ matrix, $b \in \mathbb{R}^m$ then either there is a vector $x \in \mathbb{R}^n$ s.t. $Ax = b$, or there is a vector $y \in \mathbb{R}^m$ s.t. $A^T y = 0$ & $y \cdot b \neq 0$.



$\leftarrow R(A) = \{Av \mid v \in \mathbb{R}^n\} = \operatorname{Col}(A)$

$b \notin R(A)$
$\Rightarrow \exists\, b' \perp R(A)$
$\Rightarrow \exists\, b' \in N(A^T)$
$\Rightarrow A^T b' = 0.$

$$\boxed{P = A(A^TA)^{-1}A^T}$$

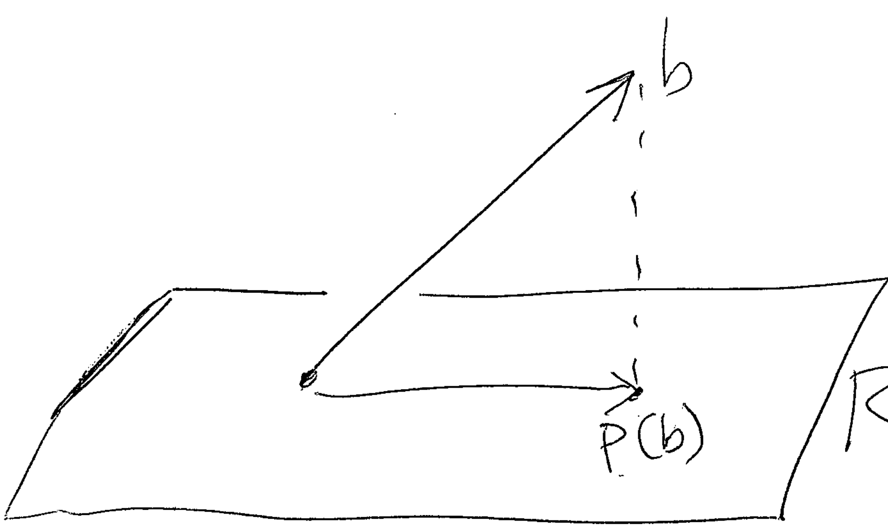
$$P(Av) = A\overbrace{(A^TA)^{-1}(A^TA)}v$$

$$= Av$$

$$v \perp R(A)$$

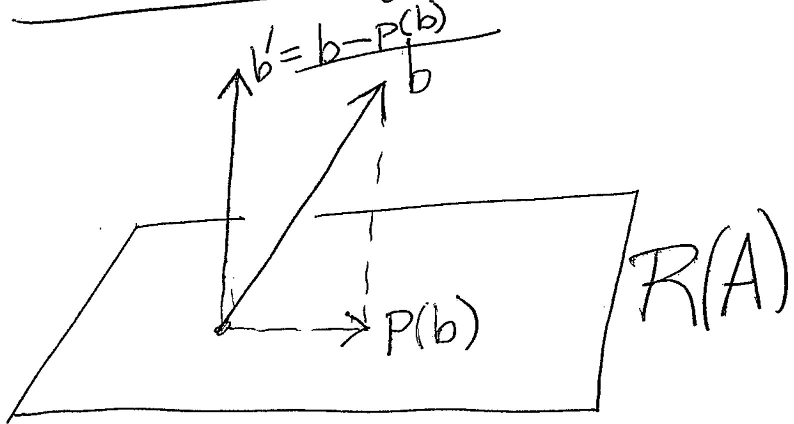
$$\Rightarrow Pv = 0$$

$$A\hat{x} = A(A^TA)^{-1}A^T\underline{b}$$



$R(A) = \{Av \mid v \in \mathbb{R}^n\}$

Least Squares Problem



$\boxed{Ax = b}$ unsolvable.

$A\cancel{x} = P(b)$ solvable

& solution are closest to solutions to $Ax=b$ if there were any.

$b' = b - P(b)$ is $\perp R(A)$

So $A^T b' = 0$

$0 = A^T b' = A^T b - A^T P(b)$

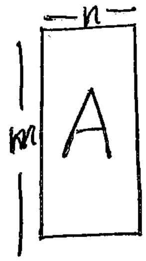
$A^T b = A^T P(b)$

iff $Ax = P(b)$

$\Longrightarrow \boxed{A^T b = A^T A x}$ The normal equations

Theorem. If A is $m \times n$ matrix of rank n, then the normal equations $A^T b = A^T A x$ have unique soln $\boxed{\hat{x} = (A^T A)^{-1} A^T b.}$

Proof. A $m \times n$ rank n means



Col(A) has dim n

Show: A^TA non-sing.

A^TA is $(n\times m)(m \times n) = n \times n$

Consider $A^TAx = 0$.

soln z $\boxed{A^T(Az) = 0}$

$Az \in N(A^T)$

$Az \in R(A) = N(A^T)^\perp$

$\Rightarrow Az = 0$

Now A rank $n \Rightarrow Ax = 0$ has only triv soln.

$\therefore$ $z = 0$

$\therefore$ A^TA non-sing.

So $A^TAx = A^Tb$

$\Rightarrow x = (A^TA)^{-1}A^Tb$ //

Example 1. $A=\begin{bmatrix}1&1\\-2&3\\2&-1\end{bmatrix}$, $\vec{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$, $\vec{b}=\begin{pmatrix}3\\1\\2\end{pmatrix}$

$$A\vec{x}=\vec{b}$$

First note that $\vec{b}\notin R(A)$:

$$\left[\begin{array}{cc|c}1&1&3\\-2&3&1\\2&-1&2\end{array}\right]\to\left[\begin{array}{cc|c}1&1&3\\0&5&7\\0&-3&-4\end{array}\right]\to\left[\begin{array}{cc|c}1&1&3\\0&1&7/5\\0&1&4/3\end{array}\right]\ \#.$$

$\text{Rank}(A)=2$ ✓

$$A^TA=\begin{bmatrix}1&-2&2\\1&3&-1\end{bmatrix}\begin{bmatrix}1&1\\-2&3\\2&-1\end{bmatrix}=\begin{bmatrix}1+4+4&1-6-2\\1-6-2&1+9+1\end{bmatrix}$$

$$A^TA=\begin{bmatrix}9&-7\\-7&11\end{bmatrix}\qquad \text{Det}(A^TA)=99-49=50$$

$$(A^TA)^{-1}=\frac{1}{50}\begin{bmatrix}11&7\\7&9\end{bmatrix}$$

least squares fit to soln.

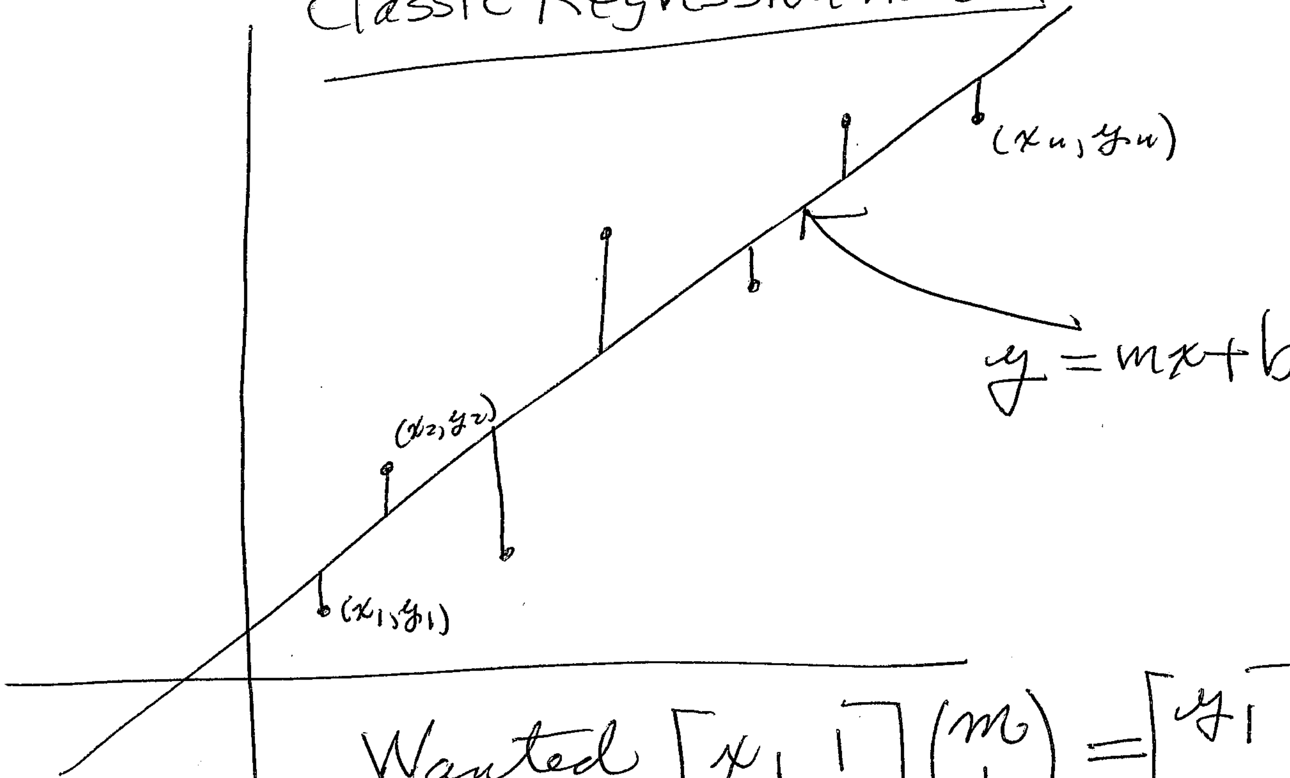
We solve $A^TA\vec{x}=A^Tb=\begin{bmatrix}1&-2&2\\1&3&-1\end{bmatrix}\begin{bmatrix}3\\1\\2\end{bmatrix}$

$$=\begin{bmatrix}3-2+4\\3+3-2\end{bmatrix}=\begin{bmatrix}5\\4\end{bmatrix}$$

$$\vec{x}=(A^TA)^{-1}\begin{bmatrix}5\\4\end{bmatrix}$$

$$=\frac{1}{50}\begin{bmatrix}11&7\\7&9\end{bmatrix}\begin{bmatrix}5\\4\end{bmatrix}=\frac{1}{50}\begin{bmatrix}83\\71\end{bmatrix}=\begin{bmatrix}83/50\\71/50\end{bmatrix}$$

Classic Regression Problem



Wanted
$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{bmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\boxed{\begin{aligned} mx_1 + b &= y_1 \\ mx_2 + b &= y_2 \\ &\cdots \\ mx_n + b &= y_n \end{aligned}}$$

$$A v = y$$

Assume $\vec{x}$ & $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ lin ind.

$$A^T A = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} x \cdot x & \sum_i x_i \\ \sum_i x_i & n \end{bmatrix}$$

$$A^TA = \begin{bmatrix} x\cdot x & \Sigma x \\ \Sigma x & n \end{bmatrix}$$

$$d = \mathrm{Det}(A^TA) = n(x\cdot x) - (\Sigma x)^2$$

$$(A^TA)^{-1} = \frac{1}{d}\begin{bmatrix} n & -\Sigma x \\ -\Sigma x & x\cdot x \end{bmatrix}$$

$$\hat{v} = \begin{pmatrix} m \\ b \end{pmatrix} = (A^TA)^{-1}A^T y$$

$$= (A^TA)^{-1}\begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$= (A^TA)^{-1}\begin{bmatrix} x\cdot y \\ \Sigma y \end{bmatrix}$$

$$= \frac{1}{d}\begin{bmatrix} n & -\Sigma x \\ -\Sigma x & x\cdot x \end{bmatrix}\begin{bmatrix} x\cdot y \\ \Sigma y \end{bmatrix}$$

$$= \frac{1}{d}\begin{bmatrix} n(x\cdot y) - (\Sigma x)(\Sigma y) \\ (x\cdot x)(\Sigma y) - (x\cdot y)(\Sigma x) \end{bmatrix}$$

$$m = \frac{n(x\cdot y) - (\Sigma x)(\Sigma y)}{n(x\cdot x) - (\Sigma x)^2}$$

$$b = \frac{(x\cdot x)(\Sigma y) - (x\cdot y)(\Sigma x)}{n(x\cdot x) - (\Sigma x)^2}$$

$$m = \frac{n(x \cdot y) - (\Sigma x)(\Sigma y)}{n(x \cdot x) - (\Sigma x)^2}$$

$$b = \frac{(x \cdot x)(\Sigma y) - (x \cdot y)(\Sigma x)}{n(x \cdot x) - (\Sigma x)^2}$$

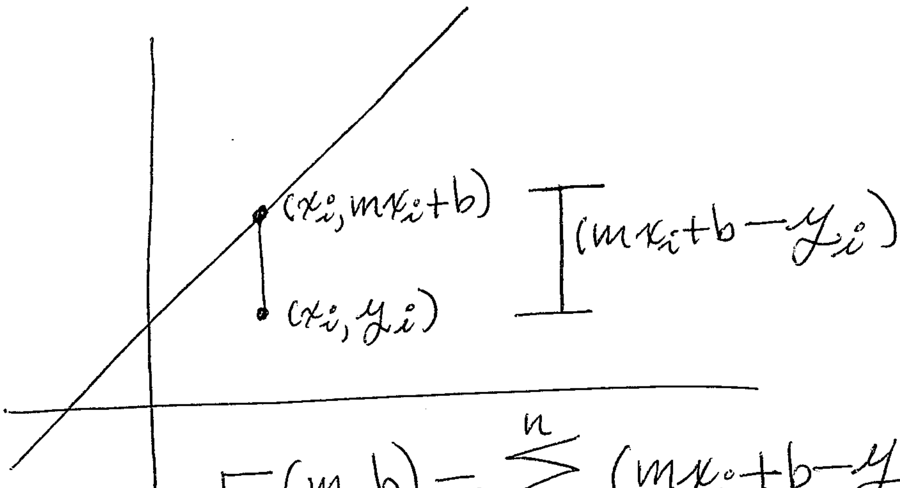
Let $\left.\begin{array}{l} \bar{x} = (\Sigma x)/n \\ \bar{y} = (\Sigma y)/n \end{array}\right\}$ average values of x_i + y_i resp.

Then $\Sigma x = n\bar{x}$, $\Sigma y = n\bar{y}$

and so

$$\boxed{\begin{aligned} m &= \frac{(x \cdot y) - n\bar{x}\bar{y}}{x \cdot x - n\bar{x}^2} \\ b &= \frac{(x \cdot x)\bar{y} - (x \cdot y)\bar{x}}{x \cdot x - n\bar{x}^2} \end{aligned}}$$

Compare Via Calculus



$$F(m,b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$$

$$\frac{\partial F}{\partial m} = \sum_{i=1}^{n} 2(mx_i + b - y_i)x_i$$

$$= 2m\sum_{i=1}^{n} x_i^2 + 2\left(\sum_{i=1}^{n} x_i\right)b - \sum_{i=1}^{n} x_i y_i$$

$$\frac{\partial F}{\partial m} = 2m\, x\cdot x + 2(\Sigma x)b - x\cdot y$$

So $0 = \dfrac{\partial F}{\partial m} \iff \boxed{(x\cdot x)m + (\Sigma x)b = x\cdot y}$

$$\frac{\partial F}{\partial b} = \sum_{i=1}^{n} 2(mx_i + b - y_i) = 2m(\Sigma x) + nb - \Sigma y$$

So $0 = \dfrac{\partial F}{\partial b} \iff \boxed{(\Sigma x)m + nb = (\Sigma y)}$

$$\therefore \begin{bmatrix} x\cdot x & \Sigma x \\ \Sigma x & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} x\cdot y \\ \Sigma y \end{bmatrix}$$

✓ This is exactly same as our linear algebraic equation.

Inner Product Spaces

An *inner product* on a vector space V is a generalization of the dot product $x \cdot y$ on $\mathbb{R}^n$. We will write $\langle x, y \rangle$ or $\langle x | y \rangle$ for this generalized ~~dot~~ product. $\langle x | y \rangle$ satisfies the following axioms: V a vector space over $\mathbb{R}$.

1. $\langle x | x \rangle \geq 0$ with $\langle x | x \rangle = 0$ only if $x = 0$.
2. $\langle x | y \rangle = \langle y | x \rangle$ for all $x, y \in V$.
3. $\langle \alpha x + \beta y | z \rangle = \alpha \langle x | z \rangle + \beta \langle y | z \rangle$ for all $x, y, z \in V$ and scalars $\alpha, \beta \in \mathbb{R}$.

Example 1. $V = \mathbb{R}^n$

$$\langle x | y \rangle = x^T y = x \cdot y.$$

(you can think $|y\rangle \equiv y$ + $\langle x| \equiv x^T$.)

Example 2. On $C[a,b] = V$.

$$\langle f|g\rangle = \int_a^b f(x)g(x)\,dx$$

If V has inner product $\langle\ |\ \rangle$,
define $\|v\| = \sqrt{\langle v|v\rangle}$ for $v \in V$.
Say $v \perp w$ (v orthogonal to w)
if $\langle v|w\rangle = 0$.

Note: $\|u+v\|^2 = \langle (u+v)|(u+v)\rangle$.
Suppose $u \perp v$. Then

$$\|u+v\|^2 = \langle u|u\rangle + \cancel{\langle v|u\rangle} + \cancel{\langle u|v\rangle} + \langle v|v\rangle$$

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 \checkmark$$

Example 3. $C[-\pi, \pi]$.

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

$$\Rightarrow \langle \cos x|\sin x\rangle = 0$$

$$\langle \cos x|\cos x\rangle = 1$$

$$\langle \sin x|\sin x\rangle = 1$$

Fourier Series

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos(m x) + b_m \sin(mx)$$

$$S(x+2\pi) = S(x) \qquad n, m = 1, 2, 3, \ldots$$

$$\int_{-\pi}^{\pi} \cos(mx)\cos(nx)\,dx = \begin{cases} 0 & m \neq n \\ \pi & m = n \end{cases}$$

$$\int_{-\pi}^{\pi} \cos(mx)\sin(nx)\,dx = 0 \quad \text{all } m, n$$

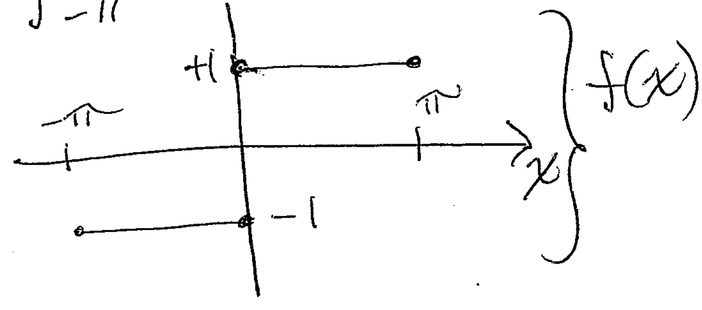
$$\int_{-\pi}^{\pi} \sin(mx)\sin(nx)\,dx = \begin{cases} 0 & m \neq n \\ \pi & m = n \end{cases}$$

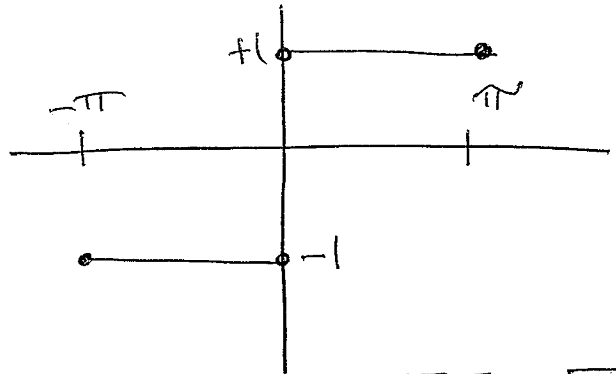
$$\int_{-\pi}^{\pi} f(x)\,dx = \pi a_0 \qquad \{1\} \cup \{\cos(nx)\} \cup \{\sin(mx)\}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx$$

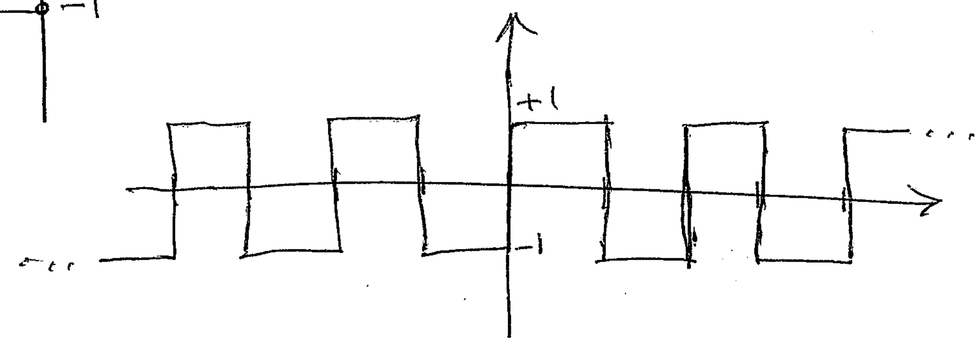
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$$

Exercise:





$y = f(x)$



$$b_n = \frac{2}{\pi}\int_0^{\pi} \sin(nx)\,dx = \frac{2}{\pi}\left(\frac{-1}{n}\cos(nx)\right)\Bigg|_0^{\pi}$$

$$= \frac{2}{\pi}\left(\frac{-1}{n}\cos(n\pi) + \frac{1}{n}\cos(0)\right)$$

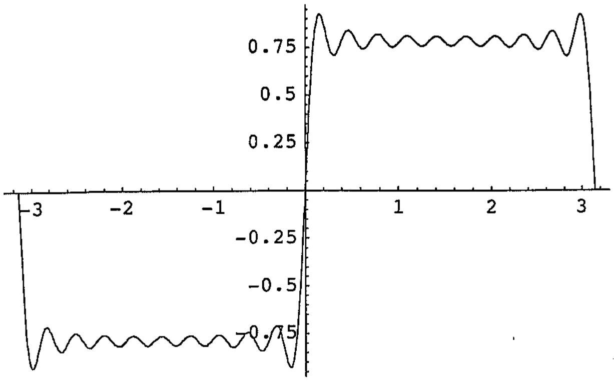
$$= \begin{cases} 0 & n \text{ even} \\ \frac{4}{\pi n} & n \text{ odd} \end{cases}$$

$$\Longrightarrow \quad f(x) = \sum_{m \text{ odd}} \frac{4}{\pi m} \sin(mx)$$

$$f(x) = \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{1}{2k+1} \sin\left((2k+1)x\right)$$

15.3

```
Plot[Sin[x] + Sin[3 x] / 3 + Sin[5 x] / 5 + Sin[7 x] / 7 + Sin[9 x] / 9 +
  Sin[11 x] / 11 + Sin[13 x] / 13 + Sin[15 x] / 15 + Sin[17 x] / 17 + Sin[19 x] / 19
 , { x, -Pi, Pi}]
```



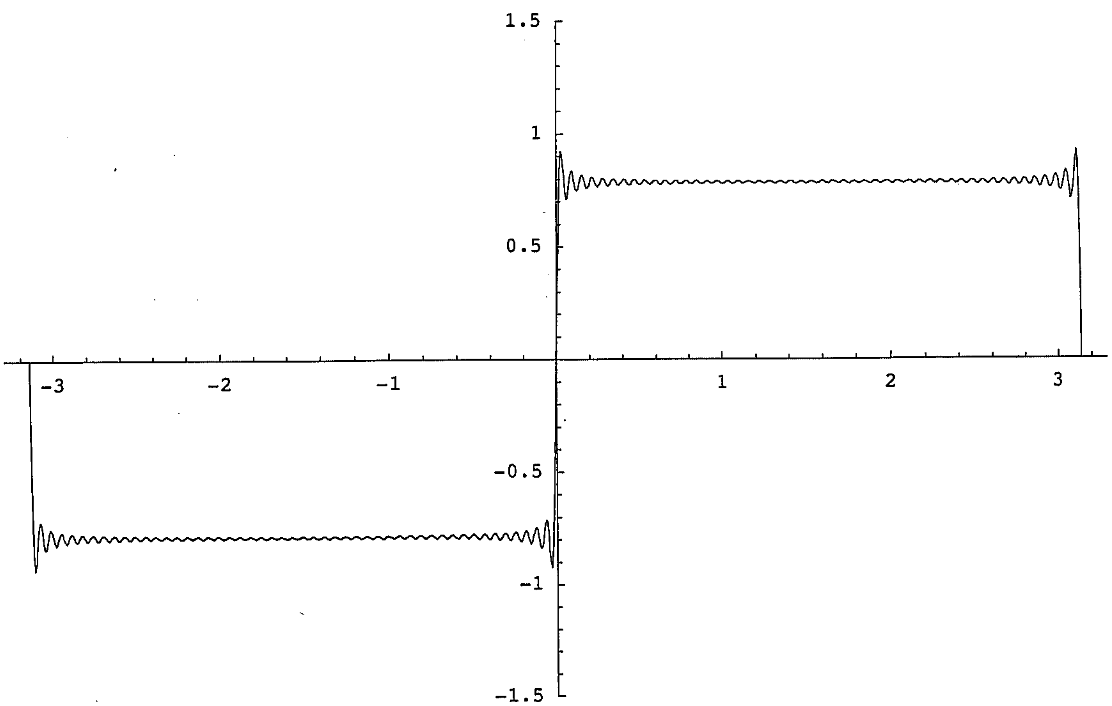
- Graphics -

```
$RecursionLimit = Infinity
```

∞

```
G[x_, 1] := Sin[x]
G[x_, n_] := G[x, n] = G[x, n - 1]  + Sin[ (2 n - 1) x] / (2 n - 1)
```

```
Plot[G[x, 50], {x, -Pi, Pi}, PlotRange → {-1.5, 1.5}]
```



- Graphics -

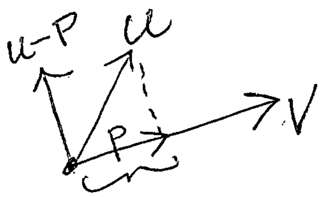
Theorem (Cauchy-Schwarz Inequality).

$$u, v \in V \Longrightarrow |\langle u|v\rangle| \leq \|u\|\,\|v\|$$

& $|\langle u|v\rangle| = \|u\|\,\|v\|$ iff u, v linearly dependent.

Proof. We assume $v \neq 0$

& P = vector proj of u to v.



$$P = \frac{\langle u|v\rangle}{\langle v|v\rangle} v, \quad P \cdot P = \frac{\langle u|v\rangle^2}{\langle v|v\rangle^2} \langle v|v\rangle$$

$P \perp (u - P)$ (check!)

& $\therefore \|p\|^2 + \|u-p\|^2 = \|u\|^2$

$$\therefore \frac{\langle u|v\rangle^2}{\|v\|^2} = \|p\|^2 = \|u\|^2 - \|u-p\|^2$$

$$\Longrightarrow \langle u|v\rangle^2 = \|u\|^2\|v\|^2 - \|u-p\|^2\|v\|^2 \leq \|u\|^2\|v\|^2$$

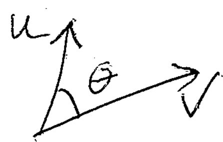
$$\therefore |\langle u|v\rangle| \leq \|u\|\,\|v\|.$$

& we see $|\langle u|v\rangle| = \|u\|\,\|v\|$ iff $u = P$, whence dep. //

So $-1 \leq \dfrac{\langle u|v\rangle}{\|u\|\,\|v\|} \leq 1$

& we can define

$$\cos(\theta) = \frac{\langle u|v\rangle}{\|u\|\,\|v\|}$$



Complex Vector Spaces

$$\mathbb{C}^n = \{ z = (z_1, z_2, \dots, z_n) \mid z \in \mathbb{C} \}$$

$$\mathbb{C} = \{ a+bi \mid a, b \text{ real} ;\ i^2 = -1 \}$$

$$z = a+bi, \quad w = c+di$$

$$zw = (ac-bd) + (ad+bc)i$$

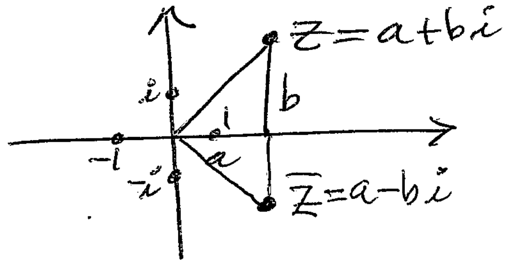
$$z+w = (a+c) + (b+d)i$$

$$\overline{z} \underset{def}{=} a-bi$$

$$z\overline{z} = (a+bi)(a-bi) = a^2+b^2.$$

$$z\overline{z} = |z|^2 \text{ where } |z| = \sqrt{a^2+b^2}$$

$$\boxed{\begin{array}{l} \dfrac{z+\overline{z}}{2} = Re(z) \\ \hline \dfrac{z-\overline{z}}{2i} = Im(z) \end{array}}$$



$\mathbb{C}$ The complex plane

Fact: $\overline{zw} = \overline{z}\,\overline{w}$.

Proof. $\overline{zw} = (ac-bd) - (ad+bc)i$

$$\overline{z}\,\overline{w} = (a-bi)(c-di)$$

$$= (ac-bd) + (-ad-bc)i \quad \checkmark \quad //$$

A Number-Theoretic Application of Complex Numbers

Thm. Let A, B, C, D be integers.

$$\Rightarrow (A^2+B^2)(C^2+D^2) = R^2+S^2$$

for some integers R, S.

---

e.g. $(1^2+2^2)(3^2+4^2)$

$= (5)(5^2) = 5^3 = 125$

$= 100 + 25 = 10^2+5^2$

$\boxed{A=1, B=2 \\ C=3, D=4}$

$(3-8)^2+(4+6)^2$

$= 5^2+10^2$

$z = A+iB \Rightarrow z\bar{z} = A^2+B^2$

$w = C+iD \Rightarrow w\bar{w} = C^2+D^2$

$(A^2+B^2)(C^2+D^2) = z\bar{z}w\bar{w}$

$= zw\,\bar{z}\bar{w}$

$= (zw)\overline{(zw)}$

$= (AC-BD)^2 + (AD+BC)^2$

$zw = AC-BD + (AD+BC)i$

Inner product on $\mathbb{C}^n$

$\overline{\bar{z} w} = \bar{\bar{z}}\,\bar{w} = z\bar{w} = \bar{w} z$

$$\langle \vec{z} | \vec{w} \rangle = \bar{z}_1 w_1 + \bar{z}_2 w_2 + \cdots + \bar{z}_n w_n$$

<u>Note</u>: $\langle \vec{z} | \vec{z} \rangle = \bar{z}_1 z_1 + \cdots + \bar{z}_n z_n$

$= |z_1|^2 + \cdots + |z_n|^2$

≥ 0

& $= 0$ only if $\vec{z} = 0$.

CX inner prod

$\langle \alpha \vec{z} | \vec{w} \rangle = \bar{\alpha} \langle \vec{z} | \vec{w} \rangle$

$\langle \vec{z} | \alpha \vec{w} \rangle = \alpha \langle \vec{z} | \vec{w} \rangle$

<u>Note</u>: $\langle \vec{w} | \vec{z} \rangle = \overline{\langle \vec{z} | \vec{w} \rangle}$!

$$\langle \vec{z} | \vec{w} \rangle = \vec{z}^{\,*} \vec{w}$$

where $x^* = \bar{x}^T$

$(\bar{x}_1, \ldots, \bar{x}_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x^* y = \langle x | y \rangle$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow x^* = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$$

$U : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is a <u>unitary</u> linear transformation (matrix) if $U^* = U^{-1}$.

So $U^* U = I$ & this means that $\mathrm{Col}(U)$ form an orthonormal basis in $\langle \, | \, \rangle$.

Note:

$$\langle \alpha \vec{z} \,|\, \alpha \vec{w} \rangle = \bar{\alpha}\,\alpha \langle \vec{z} \,|\, \vec{w} \rangle$$

$$\langle \alpha \vec{z} \,|\, \alpha \vec{z} \rangle = \alpha\,\bar{\alpha} \langle \vec{z} \,|\, \vec{z} \rangle$$

$$\begin{array}{c} \text{cols} \\ \begin{bmatrix} x & y \end{bmatrix} \end{array}$$

$$M = \begin{bmatrix} z & w \\ p & q \end{bmatrix} \quad z, w, p, q \in \mathbb{C}$$

$$M^* = \begin{bmatrix} \bar{z} & \bar{p} \\ \bar{w} & \bar{q} \end{bmatrix}$$

$$M^*M = \begin{bmatrix} \bar{z} & \bar{p} \\ \bar{w} & \bar{q} \end{bmatrix} \begin{bmatrix} z & w \\ p & q \end{bmatrix} = \begin{bmatrix} \langle x|x\rangle & \langle x|y\rangle \\ \langle y|x\rangle & \langle y|y\rangle \end{bmatrix}$$

$$M^*M = I \Rightarrow \langle x|x\rangle = 1$$
$$\langle y|y\rangle = 1$$
$$\langle x|y\rangle = \langle y|x\rangle = 0$$

$\{M\}$ called $\underline{SU(2)}$.

ex: $M = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$; $z, w \in \mathbb{C}$

and assume $\det(M) = 1$.

$$\det(M) = z\bar{z} + w\bar{w} \underset{\text{assume}}{=} 1.$$

$$M^* = \begin{pmatrix} \bar{z} & -\bar{\bar{w}} \\ \bar{w} & \bar{\bar{z}} \end{pmatrix} = \begin{pmatrix} \bar{z} & -w \\ \bar{w} & z \end{pmatrix}$$

$$M^*M = \begin{pmatrix} \bar{z} & -w \\ \bar{w} & z \end{pmatrix} \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$$

$$= \begin{pmatrix} \bar{z}z + w\bar{w} & 0 \\ 0 & \bar{z}z + w\bar{w} \end{pmatrix} = \begin{pmatrix} \det(M) & 0 \\ 0 & \det(M) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \checkmark$$

SU(2) is closed under Matrix Multiplication

$$\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}\begin{pmatrix} p & q \\ -\bar{q} & \bar{p} \end{pmatrix}$$

both have ~~det~~ = 1

$$\|$$

$$\begin{pmatrix} zp - w\bar{q} & zq + w\bar{p} \\ -\bar{w}p - \bar{z}\bar{q} & -\bar{w}q + \bar{z}\bar{p} \end{pmatrix}$$

$$\|$$

$$\begin{pmatrix} pz - \bar{q}w & qz + \bar{p}w \\ -\left(\overline{qz + qz}\right) & \overline{pz - \bar{q}w} \end{pmatrix}$$

det = 1

$$\|$$

$$\begin{pmatrix} A & B \\ -\bar{B} & \bar{A} \end{pmatrix}$$

$$\mathcal{M} = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}, \quad z = a + ib, \quad w = c + id \qquad \boxed{E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}$$

$$= \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$+ c\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

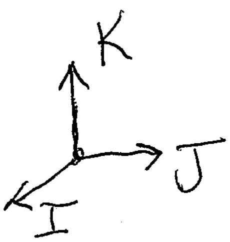
$$\left\{ \begin{aligned} &\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}^2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E \\ &\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = -E \\ &\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}^2 = -E \end{aligned} \right.$$

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

Let $I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $K = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

Then $I^2 = J^2 = K^2 = -E$



$$\begin{aligned} IJ &= K \\ JK &= I \\ KI &= J \end{aligned} \qquad \left\{ \begin{aligned} JI &= -K \\ KJ &= -I \\ IK &= -J \end{aligned} \right.$$

SU(2) gives a matrix representation of <u>Hamilton's Quaternions</u>.
(discovered by W.R. Hamilton in 18___.)

Note that

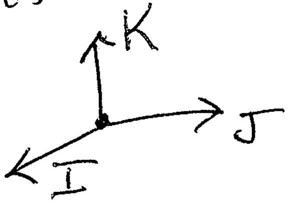
$$M = aE + bI + cJ + dK$$

and $a^2+b^2+c^2+d^2=1$

$$\iff \det(M) = 1.$$

M is a "generalized unit complex number".

Think of numbers $\alpha I + \beta J + \gamma K = v$. $(\alpha, \beta, \gamma \text{ real})$ as vectors in $\mathbb{R}^3$.



} orthonormal basis of $\mathbb{R}^3$.

a different b from above

Then $M \equiv a + bu$

where $u \in \mathbb{R}^3$ & $\|u\| = 1$.

Suppose $M = a + bu \quad a^2+b^2=1, \ \|u\|=1$

$N = c + dv \quad c^2+d^2=1, \ \|v\|=1$.

Then check that

$$uv = -\underbrace{u \cdot v}_{\text{scalar}} + \underbrace{u \times v}_{\text{in } \mathbb{R}^3}$$

(view $aE \equiv a$)

So $MN = (a+bu)(c+dv)$

$$MN = \underbrace{(ac - bd\, u \cdot v)}_{\text{scalar}} + \underbrace{(adv + bcu + bd(u \times v))}_{\mathbb{R}^3 \text{ part}}.$$

$\|u\|=1$, $u^2=-1$

$u \in \mathbb{R}^3$ Think of $a+bu$ with $a^2+b^2=1$

$e^{u\theta}=\cos(\theta)+u\sin(\theta)$

as $e^{u\theta}$, $a=\cos(\theta)$, $b=\sin(\theta)$.

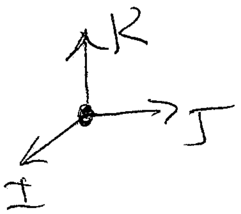
Note that $u^2 = -u\cdot u + u\times u$

$u^2 = -1$.

We have a 2-dim'l sphere's worth of square roots of -1.

Fact: If $g = e^{u\theta/2}$, $\bar{g} = e^{-u\theta/2}$ & $v \in \mathbb{R}^3$ then

$g v \bar{g}$ = Rotate v around the u-axis by the angle θ.



Exercise. Verify this method of producing rotations.

$\theta = 180^\circ$ rotation about $u = K$

$g = e^{\frac{\pi}{2}K} = \cos\left(\frac{\pi}{2}\right) + K\sin\left(\frac{\pi}{2}\right) = K$

$g x \bar{g} = -KxK$

$gK\bar{g} = -KKK = -(-1)K = K$ ✓

$-KIK = +KJ = -I$

$-KJK = -J$

$g \leftrightarrow 90^\circ$ around I

$h \leftrightarrow 90^\circ$ around J

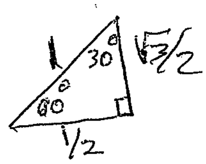
$$g = e^{\frac{\pi}{4}I} = \cos\left(\tfrac{\pi}{4}\right) + \sin\left(\tfrac{\pi}{4}\right) I = \tfrac{\sqrt{2}}{2} + \tfrac{\sqrt{2}}{2} I$$

$$h = e^{\frac{\pi}{4}J} = \tfrac{\sqrt{2}}{2} + \tfrac{\sqrt{2}}{2} J$$

$$hg = \tfrac{1}{2}(1+J)(1+I)$$

$$= \tfrac{1}{2}\left(1 + I + J + (-K)\right)$$

$$= \tfrac{1}{2} + \tfrac{1}{2}(I + J - K)$$



$$= \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}\left(\frac{I+J-K}{\sqrt{3}}\right)$$

$$= \cos(60^\circ) + \sin(60^\circ)\left(\frac{I+J-K}{\sqrt{3}}\right)$$

$$\Rightarrow hg \leftrightarrow \boxed{\text{Rotation of } 120^\circ \text{ around axis } \left(\frac{I+J-K}{\sqrt{3}}\right)}$$

Quantum Mechanics

1. The state of a physical system is represented by a vector $|\psi\rangle \in$ a Complex Vector Space $\mathcal{H}$ (We will use $\mathbb{C}^n$)
s.t. $\| |\psi\rangle \| = \langle\psi|\psi\rangle = 1.$
Thus if $\{|e_i\rangle\}$ is an orthonormal basis of $\mathcal{H}$, then
$|\psi\rangle = z_1|e_1\rangle + \cdots + z_n|e_n\rangle$
& $|z_1|^2 + \cdots + |z_n|^2 = 1.$

2. Physical processes are modeled by Unitary transformations
$U : \mathcal{H} \longrightarrow \mathcal{H}$ s.t.
$U|\psi\rangle$ represents the evolution of $|\psi\rangle$ after some time step.

3. <u>Measurement</u>

$\mathrm{Meas}\,|\psi\rangle \rightsquigarrow |e_i\rangle$ with probability $|z_i|^2$

when $|\psi\rangle = \sum_{i=1}^{n} z_i|e_i\rangle$
& $\langle\psi|\psi\rangle = 1.$

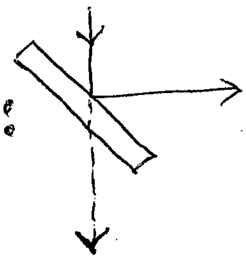
The quantum model comes from Schrodinger's Equation and we don't have time to go into the details.

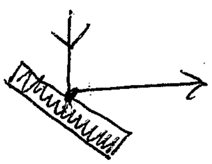
However, here is an idealized example.

1° Mach-Zender Interferometer

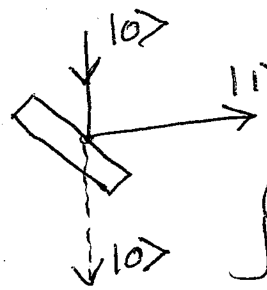
$\mathcal{H} = \mathbb{C}^2$ basis $\{|0\rangle, |1\rangle\}$ $\quad |0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$|1\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

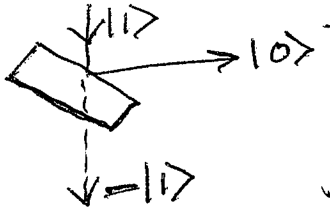
H:  Half Silvered Mirror

M:  Standard Mirror

$$\boxed{\begin{aligned} H|0\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ H|1\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} \end{aligned}}$$

Rules: 

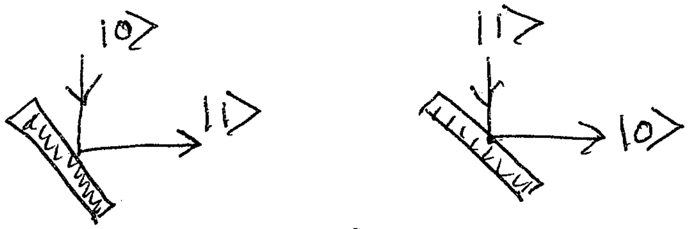
In superposition.



In superposition.

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

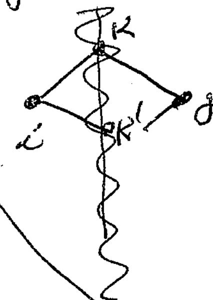
unitary matrix for half-silvered mirror.

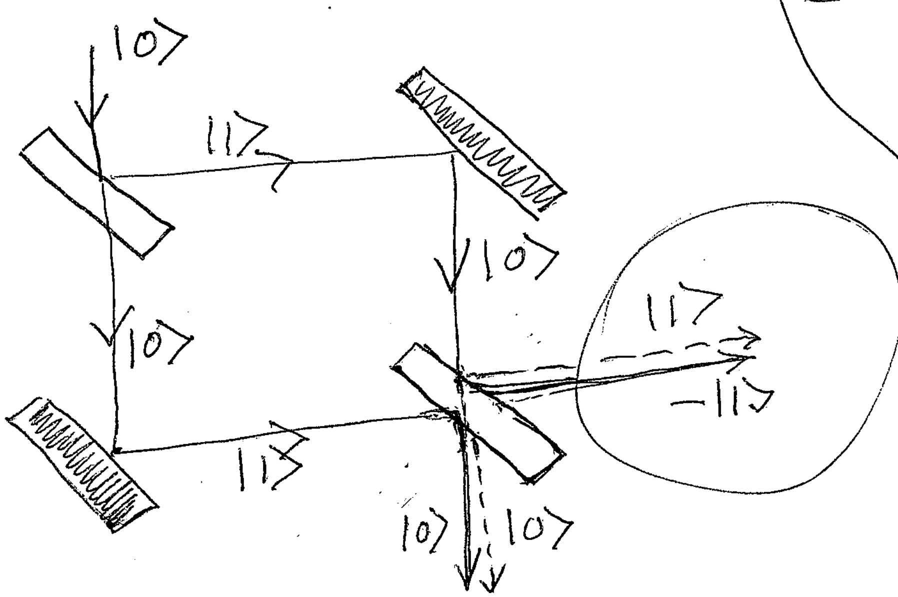
M: 

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ unitary.}$$

The Interferometer

$$(MN)_{ij} = \sum_k M_{ik} M_{kj}$$





The whole process is modeled by

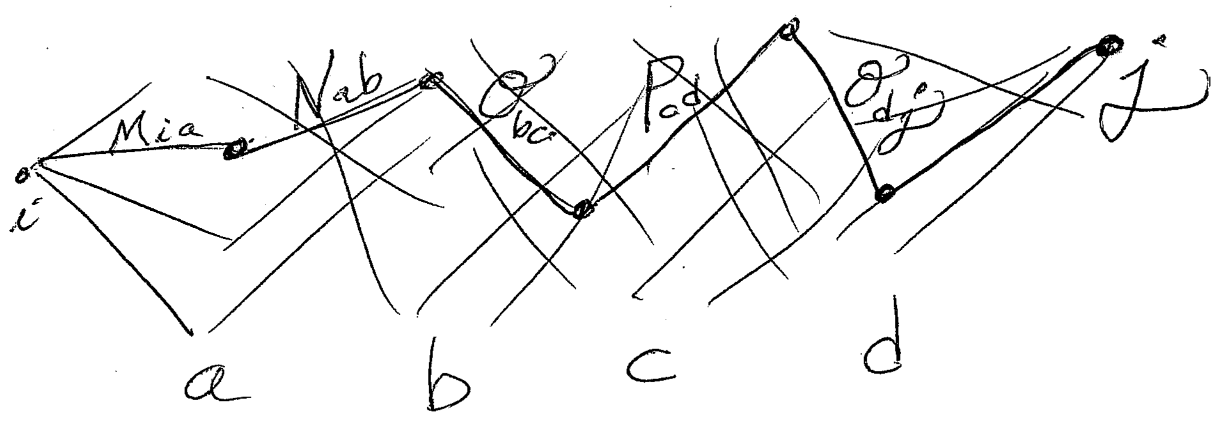
$$\mathcal{U} = HMH = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The result of $HMH = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the same as summing over all possible paths through the mirror system. This is because

Matrix Multiplication can be interpreted as a sum over paths.

$$(MNOPQ)_{ij} = \sum_{a,b,c,d} M_{ia} N_{ab} O_{bc} P_{cd} Q_{dj}$$



See: R. Feynman, QED
& R. Feynman, Lecture Notes on Physics

2° Entanglement

Notion of tensor product of vector spaces.

If V has basis $\{e_i\}_{i=1,\dots,n}$

W has basis $\{f_i\}_{i=1,\dots,m}$

then $V \otimes W$ has basis $\left\{ e_i \otimes f_j \,\middle|\, \begin{matrix} i=1,\dots,n \\ j=1,\dots,m \end{matrix} \right\}$

$\mathcal{H}$ as before

$\mathcal{H} \otimes \mathcal{H}$ has basis $|ij\rangle = |i\rangle \otimes |j\rangle \quad i,j \in \{0,1\}$.

So $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

$|ij\rangle = |i\rangle \otimes |j\rangle$ can represent the joint state of two particles that are in different locations.

$|i\rangle$ in San Francisco

$|j\rangle$ in New York

$$|\psi\rangle = \frac{|0\,\text{[redacted]}\rangle + |10\rangle}{\sqrt{2}}$$

$$|\psi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$

$$\Big\downarrow \text{ Measure in San Francisco}$$

$$|1\rangle \quad \text{or} \quad |0\rangle$$

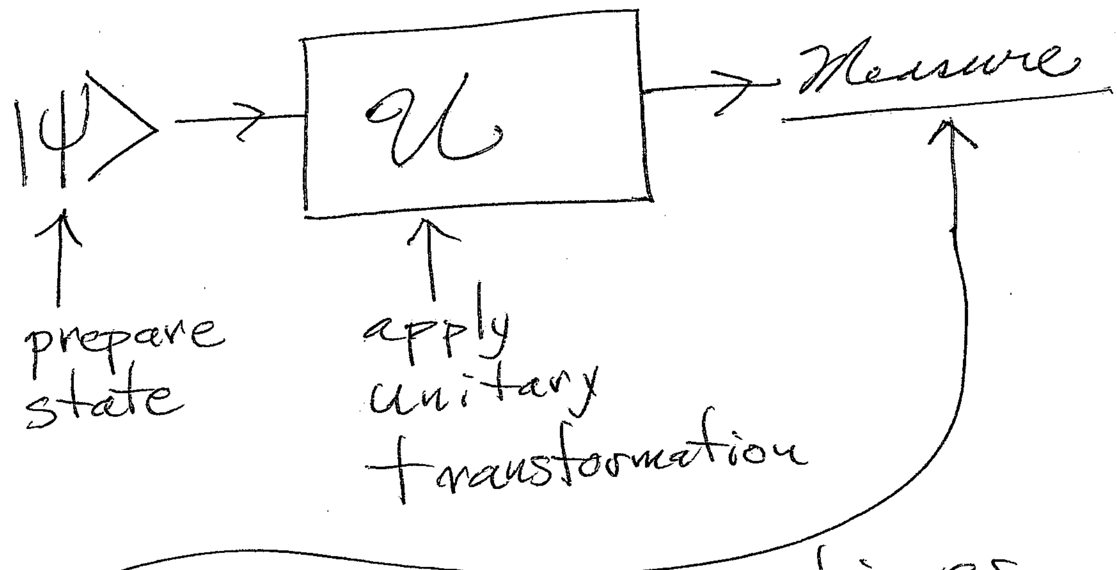
The New York "State"

with equal probability.

A measurement in San Francisco <u>determines</u> what <u>will be</u> measured in New York!

The state $|\psi\rangle$ represents particles that are <u>non-locally correlated</u>.

3.° Quantum Computers



repeat process many times to find frequency of outcomes.

<u>Design</u> algorithms so ~~that~~ one gets information from the frequency ~~of the~~ outcomes of the quantum computer.

<u>Peter Shor</u> (1996) designed such an algorithm to factor integers in principle faster than known classical algorithms.