Strang LAAIA
ch 3

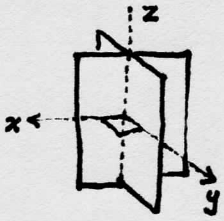
Linear Algebra With OrthoGonality

1/28/2012
1/28/2020

Let $x, y \in \mathbb{R}^n$ Then $x^T y = \sum^n x_i y_i = \langle x, y \rangle$ inner prod

$x^T y = 0 \Leftrightarrow x \perp y$ and $\|x\|_2^2 = x^T x$

Def O.G. Subspaces: V, W are OG if $v^T w = 0$ for all $v \in V$, $w \in W$



These planes do NOT qualify as OG because they share a common line (z axis)

Consider
$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \arrow[l, bend left, "A^T"] \end{tikzcd}$$

Really $(\mathbb{R}^n)^* \xleftarrow{A^T} (\mathbb{R}^m)^*$ dual spaces
but this reduces to $A^T: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ since natural iso
$\mathbb{R}^k \cong (\mathbb{R}^k)^*$ See Schuams LA p.$\overset{\text{ch 11}}{252}$, $\overset{\text{ch 13}}{284}$

$V^\perp := \{ \text{all } x \in \mathbb{R}^n \mid \langle x, v \rangle = 0 \ \forall v \in V \}$

Lemma $A: \mathbb{R}^n \longrightarrow \mathbb{R}^m \Longrightarrow \ker(A) = [\mathrm{Image}(A^T)]^\perp$ a.k.a. $\ker(A) = [\mathrm{Col}(A^T)]^\perp$

Pf show $\ker(A) \subseteq [\mathrm{Im}(A^T)]^\perp$

choose any $x \in \ker(A) \Rightarrow Ax = 0$

any $v \in \mathrm{Im}(A^T) \Rightarrow \exists\ z_v \ni A^T z_v = v$

then $v^T x = (A^T z_v)^T x = z_v^T A x = z_v^T 0 = 0$ so $x \perp$ any $v \in \mathrm{Im}(A^T)$

▷ Now show $\supseteq$

Let $y \in [\mathrm{Im}(A^T)]^\perp$ Then $y^T v = 0 \ \forall\ v \in \mathrm{Im}(A)$ by def

$\Rightarrow y^T A^T u = 0 \ \forall\ u \in \mathbb{R}^m$

$\Rightarrow (Ay)^T u = 0$

$\Rightarrow Ay = 0$ since u arb

$\Rightarrow y \in \ker(A)$

□

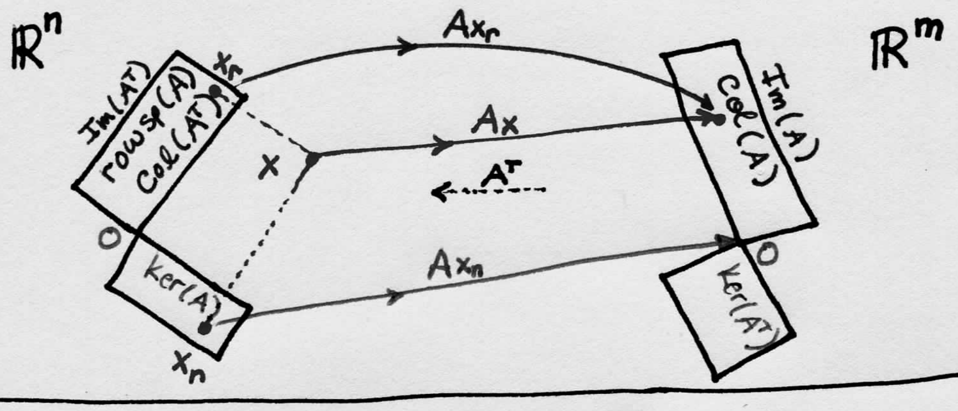
COR $\ker(A^T) = [\mathrm{Im}(A)]^\perp$

A^T is just another linear map; apply Lemma.

Strang LAAIA
ch 3.1

1/29/2020

4 Fundamental Subspaces (of Strang?)



Strang's Fund Thm of Lin Alg, Part II

$$ker(A) = [A^T(\mathbb{R}^m)]^\perp$$
$$ker(A^T) = [A(\mathbb{R}^n)]^\perp$$

Thus

$$\mathbb{R}^n = ker(A) \oplus \underbrace{Im(A^T)}_{ker(A)^\perp}$$
$$\mathbb{R}^m = ker(A^T) \oplus Im(A)$$

Part I was
$rank(A) + nullity(A) = n$

This result follows directly from the prev Lemma and it corollary, but to really understand it, you need the SVD discussion from Appendix A, which I have written up.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \\
Im(A^T) \oplus ker(A) \arrow[u, "Q_2"] \arrow[r, "\Sigma"] & Im(A) \oplus ker(A^T) \arrow[u, "Q_1"']
\end{tikzcd}$$

The O.N matrices Q_1, Q_2 change the co-ord representation of vectors (see Append A write up)

Here are some consequences and remarks:

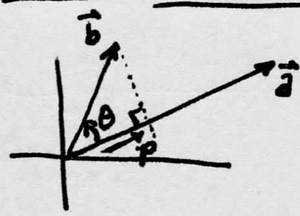
Thm: $Ax = b$ is solveable $\Longleftrightarrow$ $[A^T y = 0 \Rightarrow b^T y = 0]$

Fredholm Alternative: Either $b \in Im(A)$ or b has non-zero O.G proj into $ker(A^T)$

Every matrix A maps its row space $[Im(A^T)]$ to its col space $[Im(A)]$ one-to-one onto

$A\big|_{Im(A^T)}$ is then invertible but $A^T \neq A^{-1}$ in general, only for ON Q is $Q^T = Q^{-1}$.
See SVD write up

ch 3.2 Inner Products and Projection onto lines



Projection of $\vec{b}$ onto $\vec{a}$: $\vec{p} = \alpha\vec{a} = \frac{a^Tb\vec{a}}{a^Ta}$

$$= \vec{a}\left(\frac{a^T}{a^Ta}\right)b$$

$$= P_a(b) \text{ where matrix } P_a = \frac{1}{a^Ta}aa^T$$

$$= \frac{1}{\|a\|^2}\left[\ \right]\left[\ \ \ \right]$$

Note: $P = P^T$ symm
$P^2 = P$

$$\boxed{\underline{\text{Cauchy-Schwartz}} \quad |a^Tb| \leq \|a\|\,\|b\|}$$

Pf. $0 \leq \|b-p\|^2 = \|b-\alpha a\|^2 = \langle b-\alpha a, b-\alpha a\rangle = \langle b,b\rangle - 2\alpha\langle a,b\rangle + \alpha^2\langle a,a\rangle$

$$= b^Tb - \frac{2a^Tb}{a^Ta}a^Tb + \frac{(a^Tb)^2}{(a^Ta)^{\cancel{2}}}\cancel{a^Ta}$$

$0 \leq (b^Tb)(a^Ta) - (a^Tb)^2$ and it follows from here □

Transpose of A is really defined by $\langle Ax, y\rangle = \langle x, A^Ty\rangle$

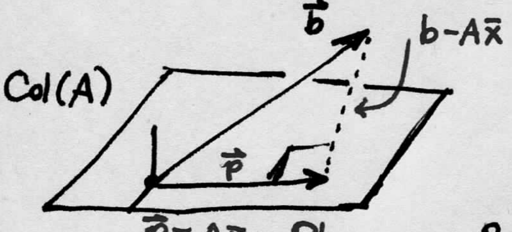
ch 3.3 Projections and Least Squares

$A: \mathbb{R}^n \longrightarrow \mathbb{R}^M$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} \\ A \\ \\ \end{bmatrix} \end{array}\begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

we want to solve $Ax = b$ but the system is overdetermined, inconsistent

Idea: minimize the error $e = \|Ax-b\|$
we will call the minimizer $\bar{x}$



$\vec{p} = A\bar{x} = Pb$

Derive the Normal Eqs for $\bar{x}$

Proj b into Col(A) [which is a subsp of $\mathbb{R}^M$, hyperplane thru 0]
then $\exists$ some vector $\bar{x}$ that maps to this projection image: $p = A\bar{x}$
then $\vec{e} := A\bar{x}-b$ is $\perp$ to Col(A)
thus $\|e\|$ is the minimum dist (Euclidean) from b to Col(A)

By <u>Fund Thm of LA</u> $e \in \ker(A^T)$ because $\mathbb{R}^M = \underbrace{\text{Im}(A)}_{\text{Col}(A)} \oplus \underbrace{\ker(A^T)}_{\text{Col}(A)^{\perp}}$ Prop $\perp$

so $A^T(A\bar{x}-b) = A^T(e) = 0$

$$\Rightarrow \boxed{A^TA\bar{x} = A^Tb \quad \text{Normal eqs}}$$

▷ We can also derive this from calculus:

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \|Ax-b\|_2^2 = \underbrace{(Ax-b)^T(Ax-b)}_{\substack{\text{bilinear form} \\ B(v,v)}} = (x^TA^T-b^T)(Ax-b) = x^TA^TAx - x^TA^Tb \\ &\qquad - b^TAx + b^Tb \end{aligned}$$

For more details of differentiating bilinear form directly like this, see my writeup of M&T Ch 4.2 or Avez DC p. 4, 31

$$Df_x(h) = 2x^T(A^TA)h - \underbrace{h^TA^Tb}_{b^TAh} - b^TAh \overset{!}{=} 0 \text{ for extrema}$$

$$\cancel{2}x^T(A^TA)h = \cancel{2}b^TAh \text{ for any } h$$

$$\Rightarrow x^T(A^TA) = b^TA \xrightarrow[\text{transpose}]{} (A^TA)^Tx = A^Tb \Rightarrow \boxed{A^TAx = A^Tb}$$

Strang LAAIA
ch 3.3 cont'd

12/13/2019

2a

So we have established $A^TA\bar{x} = A^Tb$

If A^TA would be invertible, we could solve: $\bar{x} = (A^TA)^{-1}A^Tb$

and the projection $\vec{p} = A\bar{x} = \underbrace{A((A^TA)^{-1}A^T}\, b)$

This is P the proj matrix. Note how it is still the same form $p = a\left(\frac{1}{a^Ta}\right)a^T$ as for projecting to a line.

Observe this P is a true projection:

$P^2 = P$ because $P^2 = A(A^TA)^{-1}\underbrace{\boxed{A^T \cdot A(A^TA)^{-1}}}_{I}A^T = A(A^TA)^{-1}A^T = P$

$P^T = P$ because $P^T = [A(A^TA)^{-1}A^T]^T$

$= A^{TT}((A^TA)^{-1})^TA^T$ ←

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

If $S = S^T$ does $S^{-1} = (S^{-1})^T$?

Yes, because $SS^{-1} = I$

$(SS^{-1})^T = I^T = I$

$(S^{-1})^TS^T = I$

$\Rightarrow \underset{\overset{\parallel}{S}}{(S^T)^{-1}} = (S^{-1})^T$

$S^{-1} = (S^{-1})^T$ □

$(I-P)$ projects to $Col(A)^{\perp}$

Thm A has LI cols $\Longrightarrow$ A^TA is invertible

[Cols of A not LI leads us to SVD]

(we already know it is square and symm; In ch 6.2 we show it is $\boxed{\text{pos def}}$ since all EWs $\lambda > 0$

Pf If we show A^TA is One-to-One, then we know it is invertible since we are in fin dim.
[see sheet ⑥ of 'Lin Alg w/out OG']

We know for any linear B, $\ker(B) = \{0\} \Leftrightarrow$ B one-to-one.

A has LI cols $\underset{\text{sheet ③ op cit}}{\Longrightarrow}$ A one-to-one $\underset{\text{sheet ⑥ op cit}}{\Longrightarrow}$ $\ker(A) = \{0\}$

So now we are done if we can show $\ker(A^TA) = \ker(A)$:

$(\supseteq)$ $Ax = 0 \Rightarrow A^TAx = 0 \Rightarrow x \in \ker(A^TA)$

$(\subseteq)$ $A^TAx = 0 \Rightarrow x^TA^TAx = 0 \Rightarrow \langle Ax, Ax\rangle = 0 \Rightarrow \|Ax\|_2^2 = 0 \Rightarrow Ax = 0 \Rightarrow x \in \ker(A)$

$\boxed{\text{QED}}$

▷ Now a discussion of Linear Least Squares as a LC of basis fcns to curve fit data pts [following Numerical Recipes ch 15 p. 671-673]

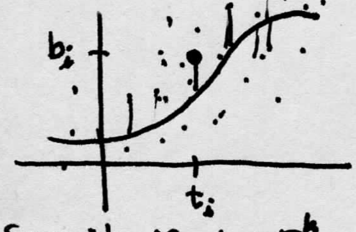
Given data pts (t_i, b_i) $i = 1, \dots, M$ and a set of basis fcns $\{\varphi_j\}$ (for example, $\varphi_1(t) = 1$, $\varphi_2(t) = t$, $\varphi_3(t) = t^2$) we want to find the best curve

$p(t) = \sum_{j=1}^{n} x_j\varphi_j(t)$ that approximates the data pts.

↑ coeff

Define total error $\chi^2 := \sum_{i=1}^{M} (\underbrace{b_i - p(t_i)}_{\text{'residual' } r_i})^2$ Vertical distance

Connection to statistics



Same idea if $t_i \in \mathbb{R}^k$

cont'd →

$$\chi^2 = \sum_{\text{data pts} \rightarrow i=1}^{M} \left(b_i - p(t_i) \right)^2 = \sum_{i=1}^{M} \Big(b_i - \sum_{j=1}^{n} x_j \underbrace{\varphi_j(t_i)}_{a_{ij} \text{ for matrix}} \Big)^2$$

$$= \sum^{M} \left[b^{\textcircled{i}} - (Ax)^{\textcircled{i}} \right]^2$$

$$= (Ax-b)^T (Ax-b)$$

For example

$$A = \text{M data pts} \left\{ \overbrace{\begin{bmatrix} \varphi_1(t_1) & \varphi_2(t_1) & \varphi_3(t_1) \\ \varphi_1(t_2) & \varphi_2(t_2) & \varphi_3(t_2) \\ \varphi_1(t_3) & \varphi_2(t_3) & \varphi_3(t_3) \\ \varphi_1(t_4) & \varphi_2(t_4) & \varphi_3(t_4) \\ \varphi_1(t_5) & \varphi_2(t_5) & \varphi_3(t_5) \end{bmatrix}}^{n \text{ basis fcns}} \right.$$

Let's minimize χ^2 wrt the coeffs x_j:

$$\frac{\partial(\chi^2)}{\partial x_k} = \sum_{i=1}^{M} 2 \Big(b_i - \sum^{n} x_j a_{ij} \Big)(-a_{ik}) \overset{!}{=} 0$$

$$= \sum_i b_i a_{ik} - \sum_i \sum_j a_{ik} a_{ij} x_j = 0$$

Observe that

$$(A^T A)_{ij} = \underset{\text{row } i}{\begin{bmatrix} — a_{pi} — \end{bmatrix}} \overset{\text{col } j}{\begin{bmatrix} | \\ a_{pj} \\ | \end{bmatrix}} = \sum_{p=1}^{M} a_{pi} a_{pj} \longrightarrow$$

$$= (A^T b)^{\textcircled{k}} - (A^T A x)^{\textcircled{k}} = 0$$

This holds for each component k

$$\boxed{A^T A x = A^T b \quad \text{Normal eqs}}$$

Then $(A^T A x)^{\textcircled{k}} = \sum_j \sum_p a_{pk} a_{pj} x_j$

Let's work a specific example from Strang p. 160-161 : Fitting the best straight line.

$$\boxed{\begin{aligned} q(t) &= c\,\varphi_1(t) + d\,\varphi_2(t) \\ &= c \cdot 1 + d \cdot t \end{aligned}}$$

The set of eqs is

$$\begin{array}{ccccc} c & + & d\,t_1 & = & b_1 \\ c & + & d\,t_2 & = & b_2 \\ \vdots & & \vdots & & \vdots \\ c & & d\,t_M & = & b_M \\ \hline t & & b & & \end{array} \Rightarrow \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_M \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}$$

$$A^T A \bar{x} = A^T b \Rightarrow \begin{bmatrix} M & \Sigma t_i \\ \Sigma t_i & \Sigma t_i^2 \end{bmatrix} \begin{bmatrix} \bar{c} \\ \bar{d} \end{bmatrix} = \begin{bmatrix} \Sigma b_i \\ \Sigma t_i b_i \end{bmatrix}$$

Say we have data pts $(-1, 1)$
$(1, 1)$
$(2, 3)$

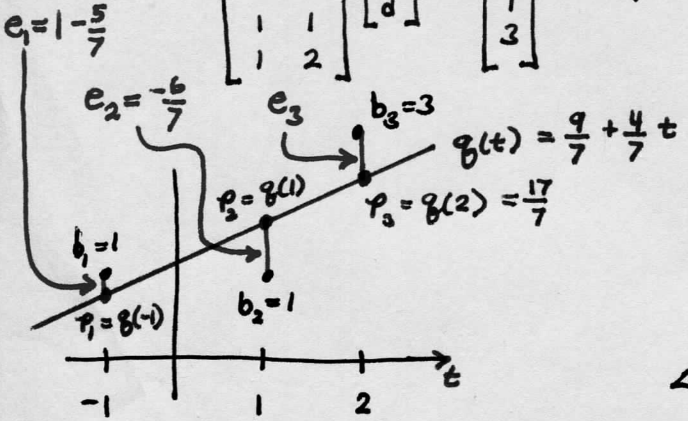
$$\begin{array}{cc} \varphi_1 & \varphi_2 \\ \hline 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{array} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \Rightarrow \begin{bmatrix} \bar{c} \\ \bar{d} \end{bmatrix} = \begin{bmatrix} 9/7 \\ 4/7 \end{bmatrix}$$

$$q(t) = \frac{9}{7} \cdot 1 + \frac{4}{7} t$$

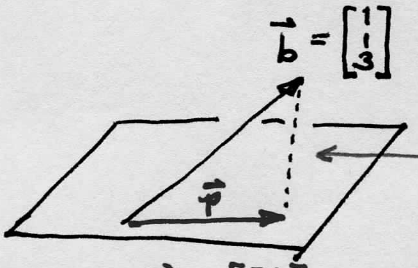
$$q(-1) = \frac{9}{7} - \frac{4}{7} = \frac{5}{7}$$

$$q(1) = 13/7$$

$$q(2) = 17/7$$



$$\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$



$$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} 3/7 \\ -6/7 \\ 4/7 \end{bmatrix}$$

$$\vec{p} = \begin{bmatrix} 5/7 \\ 13/7 \\ 17/7 \end{bmatrix} = A\bar{x} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 9/7 \\ 4/7 \end{bmatrix}$$

Unable to solve $Ax = b$, we solve $A\bar{x} = \vec{p}$

- There is non-linear least squares – see NR
- Weighted L Sq and relation to Statistics
- More later in this chapter on QR and SVD

Some material from Atkinson AITNA ch 7

p. 412 We know $\|x\|_2^2 = x^T x$

Now for $p \in [1, \infty)$ define $\|x\|_p := \left(\sum^n |x_i|^p \right)^{1/p}$

$\|x\|_\infty := \max\limits_{1 \le j \le n} |x_j|$

Thm $\lim\limits_{p \nearrow \infty} \|x\|_p = \|x\|_\infty$

Thm In $\mathbb{C}^n$ (fin dim), any 2 norms N, M are equivalent:
i.e. $\exists\ c_1, c_2 > 0 \ni \ c_1 M(x) \le N(x) \le c_2 M(x) \ \forall x \in \mathbb{C}^n$

Def of Norm $\|\cdot\| : \mathbb{R}^n \longrightarrow \mathbb{R}^{(+)}$

1. $\|x\| \ge 0$, $\|x\| = 0 \Leftrightarrow x = 0$
2. $|\alpha x\| = |\alpha| \, \|x\|$
3. $\|x+y\| \le \|x\| + \|y\|$

different

Def of metric Rudin POMA p.30

1. $d(x,y) > 0 \ \forall x \ne y$, $d(x,x) = 0$
2. $d(x,y) = d(y,x)$
3. $d(x,y) \le d(x,z) + d(z,y)$

For matrix norms, add the requirements
4. $\|AB\| \le \|A\| \, \|B\|$
5. $\|Ax\|_v \le \|A\| \, \|x\|_v$ where $\|\cdot\|_v$ is vector norm.

Frobenius norm $\|A\|_2 \overset{:=}{?} \left(\sum \sum |a_{ij}|^2 \right)^{1/2}$ Rudin POMA p.211

$\|A\|_1 := \max\limits_{1 \le j \le n} \sum^n |a_{ij}|$ $\boxed{\text{Column Norm}}$

$\|A\|_\infty := \max\limits_{1 \le i \le n} \sum^n |a_{ij}|$ $\boxed{\text{Row Norm}}$

Sup or Operator norm $\|A\|_{op} := \sup\limits_{\|x\|_v = 1} \|Ax\|_v$

Spectral radius $=$ max EW $= r_\sigma(A) := \max |\lambda_i|$

p. 421 **Thm** Let A be arb square matrix $\Rightarrow$ $r_\sigma(A) \le \|A\|$ op norm
Given $\varepsilon > 0 \ \exists$ "$\|\cdot\|_\varepsilon$" $\ni$ $\|A\|_\varepsilon \le r_\sigma(A) + \varepsilon$

COR $r_\sigma(A) < 1 \Leftrightarrow \|A\| < 1$ for some op norm.

Thm $A^m \longrightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow r_\sigma(A) < 1$

Thm $r_\sigma(A) < 1 \Leftrightarrow \exists\ (I-A)^{-1}$ and $(I-A)^{-1} = \sum\limits_{k=0}^{\infty} A^k$

$\boxed{\text{Geometric Series}}$
Cheney notes p. 19 calls it Neuman's Thm
Cheney has a lot of material on this in ch 6

Thm $\|A\| < 1$ in some op norm $\Rightarrow$ Same result as above
Also $\|(I-A)^{-1}\| \le \dfrac{1}{1 - \|A\|}$

Thm $\left.\begin{array}{l} A, B \ n{\times}n \text{ matrices} \\ A \text{ is invertible} \\ \|A - B\| < \dfrac{1}{\|A^{-1}\|} \end{array}\right\} \Rightarrow$
- B invertible
- $\|B^{-1}\| \le \dfrac{\|A^{-1}\|}{1 - \|A^{-1}\| \, \|A-B\|}$
- $\|A^{-1} - B^{-1}\| \le \dfrac{\|A^{-1}\|^2 \, \|A-B\|}{1 - \|A^{-1}\| \, \|A-B\|}$

$GL(n)$
The set of invertible matrices is open in $\mathbb{R}^{2n}$ $(\mathbb{C}^{2n})$.

Strang LAAIA ch 3.4 | O.G Bases, O.G. Matrices and Gram-Schmidt

It is customary to say "O.G" when they really mean "O.N." (I will say O.N.)

Let $Q: \mathbb{R}^n \to \mathbb{R}^n$ be an O.N. matrix (all cols are O.N.)

Remark Q preserves lengths $\|Qx\|^2 = (Qx)^TQx = x^TQ^TQx = x^Tx = \|x\|^2$

Then a priori $Q^TQ = I$.

Thm $QQ^T = I$ also.

[Thus the <u>cols</u> being O.N. forces the <u>rows</u> to be O.N.! Does <u>NOT</u> hold if cols are only O.G!! see below

pf. sheets ④-⑥ of my '<u>Linear Algebra without O.G.</u>' were building techniques to solve this problem!

First some lemmas:

Lemma: A O.N. set of vectors is LI.

<u>Lemma</u> Let $\{q_1 \dots q_n\}$ be an O.N. basis $\Rightarrow$ It is easy to find the components of any vector b:

$$b = \langle b, q_1\rangle q_1 + \dots + \langle b, q_n\rangle q_n$$

pf. a priori $b = \beta_1 q_1 + \beta_2 q_2 + \dots + \beta_n q_n$

Mult both sides by q_i^T: $q_i^Tb = 0 + \dots + \beta_i \underbrace{q_i^Tq_i}_{=1 \text{ since O.N.}} + \dots 0$

$\Rightarrow$ coeff $\beta_i = q_i^Tb = \langle b, q_i\rangle$ $\square$

<u>NOTE</u> This only works if $\{q_i\}$ is O.N. so $q_i^Tq_j = \delta_{ij}$

Now show $\underset{g\ f}{Q^TQ} = I \Rightarrow QQ^T = I$

straight forward application of 'Discussion of Inverses' p. 4 of 'L.A. w/out O.G.'

$f: X \to Y$ $\left.\begin{array}{l} g \text{ is left inv of } f \\ f \text{ maps Onto } Y\end{array}\right\} \Rightarrow$ g is also a right inv of f $f \circ g = Id_Y$

So I just need to show that $f = Q$ maps onto $\mathbb{R}^n$ $[Q: \mathbb{R}^n \to \mathbb{R}^n]$

The cols of Q are O.N. so they are n LI vectors in $\mathbb{R}^n$

By dimension counting, cols $\{q_1 \dots q_n\}$ are a basis for $\mathbb{R}^n$, even this one

Is Q onto? choose any $z \in \mathbb{R}^n$ and show $\exists\ x \ni Qx = z$

Since cols are basis, $z = \xi^1 q_1 + \xi^2 q_2 + \dots + \xi^n q_n$

$\Rightarrow x = \begin{bmatrix}\xi^1 \\ \vdots \\ \xi^n\end{bmatrix}$ and we are done! $\square$

▷ Counter-example showing it is not good enough just for cols of Q to be O.G.

[From my sheets on Symon 'Mechanics' p. 426-428]

Take the matrix $A = \begin{bmatrix}11 & -3\sqrt{2} \\ -3\sqrt{2} & 8\end{bmatrix}$ It is Real & Symmetric, so it has O.G. EVs (Discussed in Strang ch 5)

The EVs I calculate are $\begin{bmatrix}\frac{1}{\sqrt{2}} \\ 1\end{bmatrix}$ and $\begin{bmatrix}-\sqrt{2} \\ 1\end{bmatrix}$ but they are <u>not</u> normalized.

Let $S = \begin{bmatrix}\frac{1}{\sqrt{2}} & \sqrt{2} \\ 1 & -1\end{bmatrix}$ then $S^TS = \begin{bmatrix}\frac{1}{\sqrt{2}} & 1 \\ \sqrt{2} & -1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \sqrt{2} \\ 1 & -1\end{bmatrix} = \begin{bmatrix}\frac{3}{2} & 0 \\ 0 & 3\end{bmatrix}$ Not I, but at least diag.

but $SS^T = \begin{bmatrix}\frac{1}{\sqrt{2}} & \sqrt{2} \\ 1 & -1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & 1 \\ \sqrt{2} & -1\end{bmatrix} = \begin{bmatrix}2\frac{1}{2} & (\frac{1}{\sqrt{2}} - \sqrt{2}) \\ (\frac{1}{\sqrt{2}} - \sqrt{2}) & 2\end{bmatrix}$ Not even diag! $\square$

Strang ch 3.4

Rectangular Matrices with O.N. Cols

How might we solve an overdetermined system if the cols were ON : $Qx = b$

$$\begin{array}{cc} & M \\ n & \begin{bmatrix} - q_i - \end{bmatrix} \end{array} \begin{bmatrix} | \\ q_j \\ | \end{bmatrix} = \begin{array}{cc} n & \\ \begin{bmatrix} I \end{bmatrix} & n \end{array}$$

Here Q^T is only a <u>left</u> inverse of Q

$$Q^TQ = I$$

[O.N. matrices are crucial for numerical Linear Alg as they don't change lengths and thus keep round off error under control]

To solve $Qx=b$ least squares becomes simpler :

$$\underbrace{Q^TQ}_{I}\bar{x} = Q^Tb$$
$$I\ \bar{x} = Q^Tb$$

$$\vec{p} = Q\bar{x} = QQ^Tb = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix}\begin{bmatrix} - q_1^T - \\ - q_n^T - \end{bmatrix}\begin{bmatrix} \\ b \\ \\ \end{bmatrix} = \sum^n \hat{q}_i\, q_i^T b \neq b$$ because here $\{q_i\}$ is not a basis for $\mathbb{R}^M$.

Proj Matrix

$$P = A(A^TA)^{-1}A^T$$
$$= Q(Q^TQ)^{-1}Q^T$$
$$= Q\ I\ Q^T$$
$$= QQ^T \quad \text{simple form for } P.$$

example 2 Let's revisit fitting data pts to a straight line.

$c + dt_1 = b_1$
$c + dt_2 = b_2$
$c + dt_3 = b_3$

$$\Rightarrow \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$
$$\quad \vec{a}_1 \quad \vec{a}_2$$

$\vec{a}_1 \cdot \vec{a}_2 = \sum t_i$

If $\vec{a}_1 \cdot \vec{a}_2 = 0$, the cols are O.G.

so if $\sum t_i = 0$, we could solve by $(\Delta\bar{x} = A^Tb)$

Following Strang, take $\vec{a}_2 = \begin{bmatrix} -3 \\ 0 \\ 3 \end{bmatrix}$

If the cols of A are O.G. byt not O.N. then $A^TA\bar{x} = A^Tb$ becomes

$$\begin{bmatrix} a_1 a_1 & 0 \\ 0 & a_2 a_2 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a_1 \cdot b \\ a_2 \cdot b \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 3 \end{bmatrix}$$

so

$$\bar{c} = \frac{a_1^Tb}{a_1^Ta_1} = \frac{\sum b_i}{3}$$ average, but is this just accidental?

$$\bar{d} = \frac{a_2^Tb}{a_2^Ta_2} = \frac{-3b_1 + 3b_3}{3^2 + 3^2}$$

Now Strang considers if $\sum t_i \neq 0$ but I don't understand

Gram-Schmidt OrthoNormalization Process

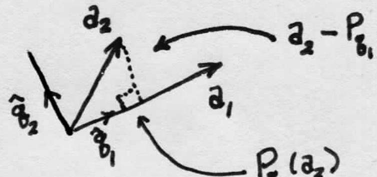
Take an arb set of vectors $\{\vec{a}_1, \ldots \vec{a}_n\}$ and produce an O.N. set $\{\hat{q}_1 \ldots \hat{q}_n\}$ with same span.
[Actually, if $\{a_i\}$ is not LI, we end up with fewer than n vectors $\{\hat{q}_i\}$.]

$$a_1' := a_1 \qquad q_1 := \frac{1}{\|a_1'\|} a_1'$$

$$a_2' := a_2 - \langle q_1, a_2 \rangle q_1 \qquad q_2 = \frac{1}{\|a_2'\|} a_2'$$

$$a_3' := a_3 - \langle q_1 a_3 \rangle q_1 - \langle q_2 a_3 \rangle q_2 \qquad q_3 = \frac{1}{\|a_3'\|} a_3'$$

$$\vdots$$



Take the next vector a_j. O.N. Project it into the subsp spanned by prev $\{q_1 \ldots q_{j-1}\}$
Subtract that off from a_j and get a new vector a_j' which is OG to that subsp.
Normalize a_j' to become q_j.

$$a_j' := a_j - \sum_{k=1}^{j-1} \langle q_k a_j \rangle q_k \quad \text{and} \quad \underbrace{q_j = \frac{1}{\|a_j'\|} a_j'}$$

$$\Rightarrow a_j' = \|a_j'\| q_j$$

$$\cancel{\Rightarrow q_j \cdot a_j' = \|a_j'\| 1}$$

$$|a_j'| q_j = a_j - \sum^{j-1} \langle q_k a_j \rangle q_k$$

$$\Rightarrow |a_j'| q_j \cdot q_j = a_j \cdot q_j - \sum 0$$

$$\Rightarrow \boxed{\|a_j'\| = q_j \cdot a_j}$$

rearranging

$$a_j = a_j' + \sum \langle \ \ \rangle q_k$$

$$= \|a_j'\| q_j + \sum \langle \ \ \rangle q_k$$

$$= \langle q_j a_j \rangle q_j + \sum \langle \ \ \rangle q_k$$

$$\Rightarrow \boxed{a_j = \langle q_1 a_j \rangle q_1 + \langle q_2 a_j \rangle q_2 + \ldots + \langle q_j a_j \rangle q_j}$$

thus

$$a_1 = \langle q_1 a_1 \rangle q_1$$

$$a_2 = \langle q_1 a_2 \rangle q_1 + \langle q_2 a_2 \rangle q_2$$

$$a_3 = \langle q_1 a_3 \rangle q_1 + \langle q_2 a_3 \rangle q_2 + \langle q_3 a_3 \rangle q_3$$

$$\Rightarrow \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \hat{q}_1 & \hat{q}_2 & \hat{q}_3 \\ | & | & | \end{bmatrix} \underbrace{\begin{bmatrix} q_1^T a_1 & q_1^T a_2 & q_1^T a_3 \\ 0 & q_2^T a_2 & q_2^T a_3 \\ 0 & 0 & q_3^T a_3 \end{bmatrix}}_{\begin{bmatrix} |a_1'| & * & * \\ 0 & |a_2'| & * \\ 0 & 0 & |a_3'| \end{bmatrix}}$$

upper triangular
Square, not rectangular despite letter "R"

$$\boxed{A = QR}$$

main diag elts always pos
$\Rightarrow \det(R) > 0$
$\Rightarrow$ R is invertible.

This ~~[illegible]~~ factorization is useful for linear L Sq, and calculating EWs (ch 7)

We can factor A = QR by Gram-Schmidt, also by Householder or Givens rotations.

Lets do an example $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ with columns $\vec{a}\ \vec{b}\ \vec{c}$

we modify the GS algo slightly to make hand calculations easier

we don't have to repeatedly calc square roots.

we have

show for $k=3$ case

$$a_3' = a_3 - \underbrace{\langle q_1, a_3\rangle q_1}_{\frac{\langle a_1', a_3\rangle a_1'}{\langle a_1', a_1'\rangle}} - \underbrace{\langle q_2\, a_3\rangle q_2}_{\frac{\langle a_2', a_3\rangle a_2'}{\langle a_2', a_2'\rangle}}$$

since $q_j = \frac{a_j'}{\|a_j'\|}$

Lets drop subscripts and call cols a, b, c

$$a' = a = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$b' = b - \frac{\langle a'\, b\rangle}{\langle a', a'\rangle} a' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$$

$$c' = c - \frac{\langle a'\, c\rangle}{\langle a'\, a'\rangle} a' - \frac{\langle b'\, c\rangle}{\langle b'\, b'\rangle} b' = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{1/2}\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \qquad q_2 = \frac{1}{1/\sqrt{2}}\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} \qquad q_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Here Q and R also happen to be square $n = 3 = M$

$$\underset{A}{\begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}} = \underset{Q}{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}} \underset{R}{\begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}}$$

$q_1 \cdot a = \frac{2}{\sqrt{2}} = \sqrt{2}$

$q_1 \cdot b = 1/\sqrt{2}$

$q_1 \cdot c = 2 \cdot \frac{1}{\sqrt{2}} = \sqrt{2}$

$q_2 \cdot b = \frac{1}{\sqrt{2}}$

$q_2 \cdot c = \sqrt{2}$

$q_3 \cdot c = 1$

▷ We want to solve the linear Least Squares problem for $Ax = b$ when $A = QR$

$$A^T A \bar{x} = A^T b$$

$$R^T \underbrace{Q^T Q}_{I} R \bar{x} = R^T Q^T b$$

$$R^T R \bar{x} = R^T Q^T b$$

we know R is invertible because diag elts are all pos $\Rightarrow \det(R) > 0$

Then R^T is also invertible and $(R^T)^{-1} = (R^{-1})^T$ because $RR^{-1} = I \Rightarrow (R^{-1})^T R^T = I^T = I$

$$\Rightarrow \boxed{R\bar{x} = Q^T b}$$

(P.172)

Now lets revisit the line fitting problem again, using Gram-Schmidt and A=QR ⑧

(P. 172)

$$\begin{bmatrix}1 & t_1\\ 1 & t_2\\ 1 & t_3\end{bmatrix}\begin{bmatrix}c\\ d\end{bmatrix} = \begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}$$

In particular, from p.160 (sheet (2b))

$$\begin{array}{c}\begin{bmatrix}1 & -1\\ 1 & 1\\ 1 & 2\end{bmatrix}\\ a_1 \quad a_2\end{array}\begin{bmatrix}c\\ d\end{bmatrix} = \begin{bmatrix}1\\ 1\\ 3\end{bmatrix}$$

$a_1 \cdot a_2 = 2 \neq 0$ not O.G.

$$a_1' = a_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$a_2 \text{[redacted]}' = \text{[redacted]}\, a_2 - \frac{\langle a' \text{[redacted]} a_2 \rangle}{\langle a' a' \rangle} a' = \begin{bmatrix}-1\\ 1\\ 2\end{bmatrix} - \frac{2}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-\frac{3}{3}-\frac{2}{3}\\ \\ \end{bmatrix} = \begin{bmatrix}-5/3\\ 1/3\\ 4/3\end{bmatrix}$$

$$q_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} \qquad q_2 = \text{[redacted]} \frac{\sqrt{3}}{\sqrt{14}}\begin{bmatrix}-5/3\\ 1/3\\ 4/3\end{bmatrix}$$

$$\sqrt{\left(\frac{-5}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{4}{3}\right)^2} = \sqrt{\frac{25}{9} + \frac{1}{9} + \frac{16}{9}} = \sqrt{\frac{42}{9}} = \sqrt{\ }$$

$6 \cdot 7 = 42$

$$\frac{2 \cdot \cancel{3} \cdot 7}{3 \cdot \cancel{3}}$$

$$\sqrt{\frac{14}{3}} \qquad \longrightarrow \frac{\sqrt{42}}{3}$$

$$\boxed{\underset{A}{\begin{bmatrix}1 & -1\\ 1 & 1\\ 1 & 2\end{bmatrix}} = \underset{Q}{\begin{bmatrix}\frac{1}{\sqrt{3}} & \frac{-5\sqrt{3}}{3\sqrt{14}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}\sqrt{14}}\\ \frac{1}{\sqrt{3}} & \frac{4}{\sqrt{3}\sqrt{14}}\end{bmatrix}} \underset{R}{\begin{bmatrix}\sqrt{3} & 2/\sqrt{3}\\ 0 & \frac{14}{\sqrt{42}}\end{bmatrix}}}$$

$a_1 \cdot q_1 = \frac{3}{\sqrt{3}} = \sqrt{3}$

$a_2 \cdot q_1 = 2/\sqrt{3}$

$a_2 \cdot q_2 =$

$$\begin{bmatrix}-1 & 1 & 2\end{bmatrix}\begin{bmatrix}-5/3\\ 1/3\\ 4/3\end{bmatrix}\left(\frac{\sqrt{3}}{\sqrt{14}}\right) = \left(\frac{5}{3} + \frac{1}{3} + \frac{8}{3}\right)\left(\frac{\text{[redacted]}3}{\sqrt{42}}\right) = \frac{14}{\sqrt{42}}$$

$$\frac{14}{\sqrt{42}} = \frac{\cancel{14}\sqrt{14}}{\sqrt{3}\cancel{\sqrt{14}}}$$

then $R\bar{x} = Q^T b$

$$= \text{[redacted]} \begin{bmatrix}1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3}\\ \frac{-5}{\sqrt{42}} & \frac{1}{\sqrt{42}} & \frac{4}{\sqrt{42}}\end{bmatrix}\begin{bmatrix}1\\ 1\\ 3\end{bmatrix} = \begin{bmatrix}5/\sqrt{3}\\ \\ \frac{8}{\sqrt{42}}\end{bmatrix}$$

$$\bar{x} = R^{-1}\begin{bmatrix}5/\sqrt{3}\\ 8/\sqrt{42}\end{bmatrix} = \frac{1}{\sqrt{3} \cdot \frac{14}{\sqrt{42}}}\begin{bmatrix}\cancel{14/\sqrt{42}} & -\frac{2}{\sqrt{3}}\\ 0 & \sqrt{3}\end{bmatrix}\begin{bmatrix}\ \\ \ \end{bmatrix} = \frac{1}{\sqrt{14}}\begin{bmatrix}\frac{14}{\sqrt{42}} & \frac{2}{\sqrt{3}}\\ 0 & \sqrt{3}\end{bmatrix}\begin{bmatrix}5/\sqrt{3}\\ 8/\sqrt{42}\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1} \rightsquigarrow \frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$$

$$= \frac{1}{\sqrt{14}}\begin{bmatrix}\frac{5\sqrt{14}}{3} + \frac{-8}{\sqrt{3}\sqrt{42}}\\ \frac{\cancel{\sqrt{3}}8}{\cancel{\sqrt{3}}\sqrt{14}}\end{bmatrix}$$

$$= \begin{bmatrix}\frac{5}{3} - \frac{8}{\sqrt{3}\sqrt{3}\sqrt{14}\sqrt{14}}\\ \frac{8}{14} = \boxed{\frac{4}{7}}\end{bmatrix}$$

wrong! →

correct!

$$= \Big[$$

$$\frac{7 \cdot 5}{7 \cdot 3} - \frac{\cancel{8}4}{3 \cdot \cancel{14}7} = \frac{\text{[redacted]}35 - 4}{7 \cdot 3} = \frac{31}{7 \cdot 3}$$ this is wrong should be a/1

The goal was to try to understand Strang's shift $y = c + d(t - \bar{t})$ where $\bar{t} = \frac{\sum t_i}{m}$ mean.

$$\text{col } a_2 \longmapsto a_2 - \bar{t} = \begin{bmatrix}t_1 - \bar{t}\\ t_2 - \bar{t}\\ t_3 - \bar{t}\end{bmatrix}$$

COV

Function Spaces and Fourier Series

① Instead of $\mathbb{R}^n$, lets imagine $\mathbb{R}^\infty$ with vectors $x = [x^{\textcircled{1}}, x^{\textcircled{2}}, x^{\textcircled{3}}, ...]$

Before $\langle x,y\rangle = \sum^n x_i y_i$ Now $\langle x,y\rangle = \sum_{i=1}^{\infty} x_i y_i$

$\|x\| = \sqrt{\langle x,x\rangle}$ and we omit all sequences x where $\|x\| = \infty$

thus $(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...)$ would be allowed $\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1}{1-\frac{1}{2}} = 2 < \infty$

This is the Hilbert space "$\mathbb{R}^\infty$" or $\ell^2(\mathbb{R})$

② $L^2[0,2\pi]$ we can regard a fcn $f: [0,2\pi] \longrightarrow \mathbb{R}$ as a vector with infinitely many "components" because

$$\langle f, g\rangle := \int_0^{2\pi} f g \, dx \quad \text{analogous to} \quad \sum f_i g_i$$

$$\|f\|_2^2 = \langle f, f\rangle = \int_0^{2\pi} f(t)^2 \, dt \quad \text{and we require} \quad \|f\|_2 < \infty$$

③ Fourier Series

$$y(x) = a_0 + \sum_{k=1}^{\infty} \left(a_k \cos(kx) + b_k \sin(kx)\right)$$

$$\langle y, \sin x\rangle = \int_0^{2\pi} y(x)\sin(x)\,dx = a_0 \int_0 \sin x \, dx + a_1 \int_0^{2\pi} \cos x \sin x \, dx + b_1 \int_0^{2\pi} \sin^2 x \, dx$$
$$+ a_2 \int_0^{2\pi} \cos(2x) \sin x \, dx + ...$$

$\Rightarrow b_1 = \frac{\langle y, \sin\rangle}{\langle \sin, \sin\rangle}$ This is just like projection $\bar{x} = \frac{b^T a}{a^T a}$

$\{\sin kx, \cos kx\}$ is a basis

④ O.G. Polys, Legendre Polys

Consider $\varphi_1(x) = 1$ on $[0,1]$
$\varphi_2(x) = x$
$\varphi_3(x) = x^2$

$$A^TA = \begin{bmatrix} \langle 1,1\rangle & \langle 1,x\rangle & \langle 1,x^2\rangle \\ \langle x,1\rangle & \langle x,x\rangle & \langle x,x^2\rangle \\ \langle x^2,1\rangle & \langle x^2,x\rangle & \langle x^2,x^2\rangle \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

ill conditioned Hilbert matrix

Use Gram-Schmidt to orthogonalize: [we switch to $[-1,1]$]

$v_1 = 1$
$v_2 = x$
$v_3 = x^2 - \frac{\langle 1,x^2\rangle}{\langle 1,1\rangle} 1 - \frac{\langle x,x^2\rangle}{\langle x,x\rangle} x = x^2 - \frac{1}{3}$

⑤ Lets approx the fcn $y=x^5$ by a straight line $y = c+dx$ on $[0,1]$.

$\boxed{\text{Method 1}}$ solve $Ax=b$ by LLSq

$$[1\ \ x]\begin{bmatrix}c\\d\end{bmatrix} = [x^5] \qquad A^TA\bar{x} = A^Tb$$

$$\langle 1,1\rangle = \int_0^1 1\,dx = x\Big]_0^1 = \textcircled{1}$$

$$\langle 1,x\rangle = \int_0^1 x\,dx = \tfrac{1}{2}x^2\Big]_0^1 = \textcircled{\tfrac{1}{2}}$$

$$\langle x,x\rangle = \int_0^1 x^2\,dx = \tfrac{1}{3}x^3\Big]_0^1 = \textcircled{\tfrac{1}{3}}$$

$$\langle 1,x^5\rangle = \int x^5 = \frac{1}{6}$$

$$\langle x,x^5\rangle = \int x^6 = \frac{1}{7}$$

$$\begin{bmatrix}1\\x\end{bmatrix}[1\ \ x]\begin{bmatrix}c\\d\end{bmatrix} = \begin{bmatrix}1\\x\end{bmatrix}[x^5]$$

$$\begin{bmatrix}\langle 1,1\rangle & \langle 1,x\rangle\\ \langle x,1\rangle & \langle x,x\rangle\end{bmatrix}\begin{bmatrix}c\\d\end{bmatrix} = \begin{bmatrix}\langle 1,x^5\rangle\\ \langle x,x^5\rangle\end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix}1 & 1/2\\ 1/2 & 1/3\end{bmatrix}\begin{bmatrix}c\\d\end{bmatrix} = \begin{bmatrix}1/6\\ 1/7\end{bmatrix}$$

This has sol'n $c = -8/42$, $d = 5/7$

$\boxed{\text{Method 2}}$

$$E^2 = \|Ax-b\|^2 = \int_0^1 [(c+dx)-x^5]^2\,dx$$

$$0 \overset{!}{=} \frac{\partial E^2}{\partial c} = \int_0^1 \cancel{2}(c+dx-x^5)\,dx = \left[cx + \tfrac{1}{2}dx^2 - \tfrac{1}{6}x^6\right]_0^1 = c + \tfrac{1}{2}d - \tfrac{1}{6} \overset{!}{=} 0$$

$$[1\ \ 1/2]\begin{bmatrix}c\\d\end{bmatrix} = 1/6$$

$$0 \overset{!}{=} \frac{\partial E^2}{\partial d} = \int_0^1 \cancel{2}(c+dx-x^5)x\,dx = \int (cx + dx^2 - x^6)\,dx = \tfrac{1}{2}c + \tfrac{1}{3}d - \tfrac{1}{7}$$

$$\Longrightarrow \begin{bmatrix}1 & 1/2\\ 1/2 & 1/3\end{bmatrix}\begin{bmatrix}c\\d\end{bmatrix} = \begin{bmatrix}1/6\\ 1/7\end{bmatrix} \quad \underline{\text{Same}}$$

$\boxed{\text{Method 3}}$ Gram-Schmidt

We need to solve $R\bar{x} = Q^Tb$

$A = [1\ \ x]$

so $a_1 = \textcircled{1} = a_1' = q_1$

$a_2 = x$

$$a_2' = x - \frac{\langle a_1', a_2\rangle}{\langle a_1', a_1'\rangle}a_1' = x - \tfrac{1}{2}\cdot 1 = \textcircled{x - 1/2}$$

$$q_2 = \frac{1}{\|a_2'\|}a_2'$$

$$\langle a_2', a_2'\rangle = \int_0^1 (x-1/2)^2\,dx = \tfrac{1}{3}(x-1/2)^3\Big]_0^1 = \frac{1}{12}$$

$$\longrightarrow \frac{1}{1/\sqrt{12}}(x-1/2) = \sqrt{12}(x-1/2)$$

$$Q^Tb = \begin{bmatrix}-\ q_1\ -\\ -\ q_2\ -\end{bmatrix}[x^5] = \begin{bmatrix}1\\ \sqrt{12}(x-1/2)\end{bmatrix}[x^5] = \begin{bmatrix}\langle 1,x^5\rangle\\ \langle\sqrt{12}(x-1/2),x^5\rangle\end{bmatrix}$$

$$\langle 1,x^5\rangle = 1/6$$

$$\langle\sqrt{12}(x-1/2),x^5\rangle = \sqrt{12}\int_0^1 (x-\tfrac{1}{2})x^5\,dx = \sqrt{12}\left(\frac{1}{7}-\frac{1}{12}\right) = \frac{5}{7\sqrt{12}}$$

$$R = \begin{bmatrix}q_1\cdot a_1 & q_1\cdot a_2\\ 0 & q_2\cdot a_2\end{bmatrix} = \begin{bmatrix}1 & 1/2\\ 0 & \frac{1}{\sqrt{12}}\end{bmatrix}$$

$$q_1\cdot a_1 = \int_0^1 1\,dx = 1$$

$$q_1\cdot a_2 = \int_0^1 1\cdot x\,dx = 1/2$$

$$q_2\cdot a_2 = \sqrt{12}\int_0^1 (x-\tfrac{1}{2})x\,dx = \frac{1}{\sqrt{12}}$$

$$\begin{bmatrix}c\\d\end{bmatrix} = \underbrace{\sqrt{12}\begin{bmatrix}\frac{1}{\sqrt{12}} & -\frac{1}{2}\\ 0 & 1\end{bmatrix}}_{R^{-1}}\underbrace{\begin{bmatrix}1/6\\ \frac{5}{7\sqrt{12}}\end{bmatrix}}_{Q^Tb} = \begin{bmatrix}-8/42\\ 5/7\end{bmatrix}$$

$\boxed{\text{QED}}$

Strang has $c + dx = \frac{1}{6} + \frac{5}{7}(x - \frac{1}{2})$ which has same values for c, d

Atkinson
AITNA

Ch 4 Approx of Fcns

7/5/2011

Thm 4.1 **Weirstrauss** $f \in C[a,b]$, $\varepsilon > 0$ $\Rightarrow \exists$ poly $p \ni |f(x) - p(x)| < \varepsilon$ on $[a,b]$

$$\|f\|_\infty = \max_{[a,b]} |f(x)| \text{ on } C[a,b]$$

Minimax approx:

$$\rho_n(f) = \inf_{\substack{\text{all polys} \\ g \text{ w/ } \deg(g) \le n}} \|f - g\|_\infty \text{ on } [a,b]$$

4.3 Least Sq Approx

$$\|g\|_2 = \left[\int_{[a,b]} |g|^2 \right]^{1/2} \text{ std norm.}$$

$$M_n(f) := \inf_{\text{poly } r} \|f - r\|_2$$

O.G. polys

w: weight fcn $\qquad |\int f| < M$

1. $\int |x|^n w(x)\, dx$ is integrable $\forall\, n \ge 0$
2. $\int g w \, dx = 0 \Rightarrow g \equiv 0$

$$\langle g, w \rangle = 0$$

$$F(a_0 \dots a_n) := \int_a^b w(x) \left[f(x) - \sum_{j=0}^{n} a_j x^j \right]^2 dx$$

Gram-Schmidt $\exists$ seq of polys (φ_n) where $\deg(\varphi_n) = n$

$$\langle \varphi_n, \varphi_m \rangle = \delta_{nm} \begin{cases} \langle \varphi_n, \varphi_m \rangle = 0 & \forall\, n \ne m \\ \langle \varphi_n, \varphi_n \rangle = 1 \end{cases}$$

- Coeff of x^n in φ_n is pos.

.175

[redacted]

Legendre polys: $w(x) = 1$ on $[-1,1]$

Chebyshev $w(x) = \frac{1}{\sqrt{1-x^2}}$ $[-1,1]$

Laguerre $w(x) = e^{-x}$ on $[0,\infty)$

Thm 4.3 f poly of degree m $\Rightarrow$ $f(x) = \sum_{n=0}^{m} \frac{\langle f,\varphi_n\rangle}{\langle \varphi_n,\varphi_n\rangle}\varphi_n(x)$

(φ_n) O.G family w/ w

Create O.G. polys with Gram-Schmidt.

4.5 Least Sq Approx of fcn

(φ_n) O.G. family of polys w/ weight fcn w, $\deg(\varphi_k) = k$.

poly $r = \sum_{i=0}^{n} b_i \varphi_i(x)$ constants

$f \in C[a,b]$

$$\boxed{\underbrace{\|f-r\|_2^2}_{\text{Error}} = \underbrace{\int_{[a,b]} \left[f(x) - \sum_{j=0}^{n} b_j \varphi_j(x)\right]^2 \omega(x)\,dx}_{G(b_0,\dots,b_n)}}$$

MINIMIZE THIS

we could set $\frac{\partial G}{\partial b_i} = 0$

Another way: $0 \le G(\vec{b}) = \left\langle f - \sum b_i\varphi_i, f - \sum b_j\varphi_j \right\rangle$

$$= \langle f,f\rangle - 2\sum_{j=0}^{n} b_j \langle f,\varphi_j\rangle + \sum_i\sum_j b_i b_j \langle \varphi_i,\varphi_j\rangle$$

$$= \|f\|_2^2 - 2\sum b_j \langle f,\varphi_j\rangle + \sum b_j^2$$

$$= \|f\|_2^2 - \sum \langle f,\varphi_j\rangle^2 + \underbrace{\sum_{j=0}^{n} \left[\langle f,\varphi_j\rangle - b_j\right]^2}$$

always pos

The min would be attained if this term were 0.

0 iff $b_j = \langle f,\varphi_j\rangle$

Then LSq approx exists and is $r_n^* = \sum^{n} \langle f,\varphi_j\rangle \varphi_j$

More over

$$\|f-r_n^*\|_2 = \left[\|f\|_2^2 - \sum \langle f, \varphi_j\rangle^2\right]^{1/2}$$

$$= \sqrt{\|f\|_2^2 - \|r_n^*\|_2^2}$$

$$\|f\|_2^2 = \|r_n^*\|_2^2 + \|f-r_n^*\|_2^2$$

Note

$$r_{n+1}^*(x) = r_n^*(x) + \langle f, \varphi_{n+1}\rangle \varphi_{n+1}(x)$$