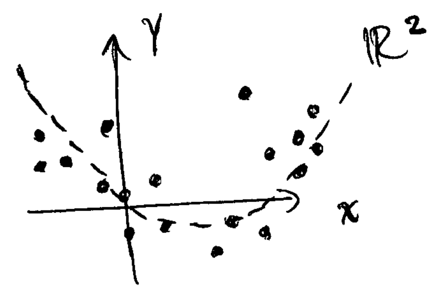
Lecture 27
03/30/12

$A\underline{x} = \underline{b}$ has sol'n $\underline{x}$ $\iff$ $\underline{b} \in \operatorname{Ran}(A)$ (column space of A).

Q: What if $\underline{b} \notin \operatorname{Ran}(A)$? Can we find $\underline{x}$ that "almost" solves $A\underline{x} = \underline{b}$? What do we mean by "almost"?

Motivation: Fitting curves to data.



Have data $\{(x_i, y_i)\}_{i=1}^{m}$.

Suppose we expect the observed data to fit the model $y = cx + dx^2$.

c, d: unknown.

What choice for c, d? According to model,

$$\begin{aligned} cx_1 + dx_1^2 &= y_1 \\ cx_2 + dx_2^2 &= y_2 \\ &\vdots \\ cx_m + dx_m^2 &= y_m \end{aligned} \iff \underbrace{\begin{pmatrix} x_1 & x_1^2 \\ \vdots & \vdots \\ x_m & x_m^2 \end{pmatrix}}_{A \in M_{m,2}(\mathbb{R})} \overbrace{\begin{pmatrix} c \\ d \end{pmatrix}}^{\underline{x} \in \mathbb{R}^2} = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}}_{\underline{b} \in \mathbb{R}^m}.$$

LEAST SQUARES (6.7):

DEFINE ERROR $E(\underline{x}) = \| \underbrace{A}_{\text{GIVEN}} \underline{x} - \underbrace{\underline{b}}_{\text{GIVEN}} \|^2$.

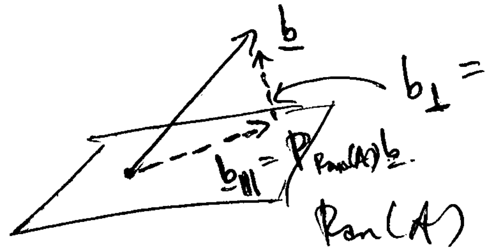
NOTE THAT $E(\underline{x}) = 0$ IFF $\underline{x}$ EXACTLY SOLVES $A\underline{x} = \underline{b}$!

DEF. $\underline{x}$ IS A LEAST SQUARES SOL'N TO $A\underline{x} = \underline{b}$

IF $E(\underline{x})$ IS MINIMIZED AT $\underline{x}$.

Q: GIVEN $A, \underline{b}$, HOW TO FIND $\underline{x}$? IS IT UNIQUE?

A:



$$\Rightarrow \quad \underline{b} = \underline{b}_\perp + \underline{b}_\parallel$$

BY PYTHAGOREAN THM.,

$$E(\underline{x}) = \| \underline{b} - A\underline{x} \|^2 = \| \underbrace{\underline{b}_\perp}_{\in (Ran(A))^\perp} + \overbrace{(\underline{b}_\parallel - A\underline{x})}^{\in Ran(A)} \|^2$$

$$= \| \underline{b}_\perp \|^2 + \underbrace{\| \underline{b}_\parallel - A\underline{x} \|^2}_{\text{WANT TO MINIMIZE THIS.}}$$

SINCE $\underline{b}_\parallel \in Ran(A)$, THERE IS AT LEAST ONE SOL'N. $\underline{x}$ TO $A\underline{x} = \underline{b}_\parallel$, SO $\min_{\underline{x}} E(\underline{x}) = \| \underline{b}_\perp \|^2$.

SO, $\underline{x}$ IS A LEAST SQUARES SOL'N IF AND ONLY IF

$$A\underline{x} = b_{\parallel} = P_{\mathrm{Ran}(A)}\,\underline{b}.$$

GIVEN

<u>NOTE:</u> LEAST SQUARES SOLN. IS <u>UNIQUE</u> IF AND ONLY IF A HAS FULL RANK (I.E., RANK OF A EQUALS NUMBER OF COLUMNS).

<u>REMARK:</u> IT IS USUALLY NOT STRAIGHTFORWARD TO COMPUTE $P_{\mathrm{Ran}(A)}$. INSTEAD, WE USE ANOTHER APPROACH.

FIRST, NOTE THAT

$$\underline{b}_{\perp} \perp \mathrm{Ran}(A) \iff \underline{b}_{\perp} \perp \text{ columns of } A$$

$$\iff A^{*}\underline{b}_{\perp} = \underline{0}.$$

THEN, FOR ANY LEAST SQUARES SOL'N $\underline{x}$,

$$A^{*}(A\underline{x}) = A^{*}(\underline{b}_{\parallel}) = A^{*}(\underline{b} - \underline{b}_{\perp}) = A^{*}\underline{b}$$

<u>DEF.</u> (NORMAL EQN.) $(A^{*}A)\,\underline{x} = A^{*}\underline{b}$, $A, \underline{b}$ GIVEN.

<u>THM.</u>
- $\underline{x}$ SOL'N. TO NORMAL EQN. $\iff$ $\underline{x}$ LEAST SQUARES SOL'N.
- $\underline{x}$ UNIQUE $\iff$ A HAS ~~RANK~~ $n = \#$ OF COLUMNS.

EX. FIND BEST LINE THROUGH POINTS

$$(0,-1),\ (1,1),\ (2,4),\ (3,9).$$

THAT IS, WE SEEK CONSTANTS c, d FOR $y = c + dx$.

WRITING THE EQN'S $y_i = c + dx_i$ FOR $i = 1, .., 4$.

IN MATRIX FORM,

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} c \\ d \end{pmatrix}}_{\underline{x}} = \underbrace{\begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \end{pmatrix}}_{\underline{b}}$$

$$A^*A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}$$

$$A^*\underline{b} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \end{pmatrix} = \begin{pmatrix} 13 \\ 36 \end{pmatrix}$$

NORMAL EQN: $$\underbrace{\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}}_{A^*A} \underbrace{\begin{pmatrix} c \\ d \end{pmatrix}}_{\underline{x}} = \underbrace{\begin{pmatrix} 13 \\ 36 \end{pmatrix}}_{\underline{b}}$$

$$\Rightarrow \begin{pmatrix} c \\ d \end{pmatrix} = \frac{1}{10}\begin{pmatrix} -17 \\ 33 \end{pmatrix}.$$ (UNIQUE SOLN. SINCE A HAS FULL RANK.)

$\Rightarrow y = -1.7 + 3.3x$ LINE OF BEST FIT.
(REGRESSION)

NOTE: NEED LINEAR EQN. FOR UNKNOWN PARAMETERS $c_0, c_1, \ldots, c_n$. MODEL FOR DATA DOESN'T NEED TO BE LINEAR!

EX. FIT DATA TO $\underbrace{y = c_0 + c_1 x}_{\text{LINEAR IN } c_i\text{'s.}}$. ✓

EX. FIT DATA TO $\underbrace{c_0 y^2 - c_1 x^2 = c_2}_{\text{LINEAR IN } c_i\text{'s.}}$ ✓

EX. FIT DATA TO $\underbrace{y = c_0 e^{c_1 x}}_{\text{NOT LINEAR IN } c_i\text{'s}}$ ✗.

LECTURE 28
04/02/12

SUPPOSE V IS AN INNER PRODUCT SPACE AND $L : V \to V$ IS A LINEAR OPERATOR WHICH IS DIAGONALIZABLE.

$\Rightarrow$ V HAS BASIS $\mathcal{B}$ OF EIGENVECTORS OF L.

Q: IS $\mathcal{B}$ AN ORTHOGONAL BASIS?

A: GENERALLY NOT.

EX. $V = \mathbb{R}^2$ w/ STD. INNER PRODUCT.

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \in M_{2,2}(\mathbb{R}) \quad (\text{i.e., } A : V \to V).$$

E-VALUES OF A ARE $\lambda = 0$ AND 3, WITH E-VECTORS $\underline{\xi}_1 = (2,-1)^T$ AND $\underline{\xi}_2 = (1,1)^T$.

HOWEVER, $\langle \underline{\xi}_1 | \underline{\xi}_2 \rangle = (2,-1)(1,1)^T = 1 \neq 0$,

SO $\underline{\xi}_1 \not\perp \underline{\xi}_2$.

Q: GIVEN AN INNER PRODUCT SPACE, WHICH CLASS OF OPERATORS GENERATE AN ORTHOGONAL BASIS OF E-VECTORS? WHAT IS GEOMETRIC STRUCTURE OF THESE OPERATORS?

ADJOINTS (7.1):

DEF. V INNER PROD. SPACE, $L : V \to V$.

THE ADJOINT OPERATOR L^* OF L IS THE UNIQUE OPERATOR THAT SATISFIES

$$\langle L^* \underline{x} \mid \underline{y} \rangle = \langle \underline{x} \mid L \underline{y} \rangle \quad \text{FOR ALL } \underline{x}, \underline{y} \in V.$$

- IF $V = \mathbb{C}^n$ w/ STD. INNER PRODUCT, THE ADJOINT OF $A \in M_{n,n}(\mathbb{C})$ IS $A^* = \overline{A}^T$.
- FOR A GENERAL VECTOR SPACE w/ ORTHONORMAL BASIS $\mathcal{E} = \{\underline{e}_i\}_{i=1}^n$:

MATRIX REPRESENTATION OF $L : V \to V$ IS

$$([L]_\mathcal{E})_{ij} = \langle \underline{e}_i \mid L \underline{e}_j \rangle$$

(EASY TO CHECK THIS, AND THAT $L = \sum_{i=1}^n \sum_{j=1}^n ([L]_\mathcal{E})_{ij} \, |\underline{e}_i\rangle\langle \underline{e}_j|$).

THEN,

$$\begin{aligned}
([L^*]_\mathcal{E})_{ij} = \langle \underline{e}_i \mid L^* \underline{e}_j \rangle &= \overline{\langle L^* \underline{e}_j \mid \underline{e}_i \rangle} \\
&= \overline{\langle \underline{e}_j \mid L \underline{e}_i \rangle} \\
&= \overline{([L]_\mathcal{E})_{ji}}
\end{aligned}$$

$$\Rightarrow \quad [L^*]_{\mathcal{E}} = [L]_{\mathcal{E}}^* .$$

EX. $V = \mathbb{C}^3$, w/ STD. INNER PROD.

LET $L\underline{x} = (3x_1 + ix_2,\ ix_2 - 2x_3,\ (1+i)x_1 + 5x_3)^T$

WHERE $\underline{x} = (x_1, x_2, x_3)^T \in \mathbb{C}^3$.

Q: WHAT IS $L^*\underline{x}$?

A: $L = \begin{pmatrix} 3 & i & 0 \\ 0 & i & -2 \\ 1+i & 0 & 5 \end{pmatrix} \Rightarrow L^* = \overline{L}^T = \begin{pmatrix} 3 & 0 & 1-i \\ -i & -i & 0 \\ 0 & -2 & 5 \end{pmatrix}$

$$\Rightarrow L^*\underline{x} = (3x_1 + (1-i)x_3,\ -ix_1 - ix_2,\ -2x_2 + 5x_3)^T.$$

EX. $V = L^2(\mathbb{R})$ (INFINITE DIMENSIONAL SPACE OF SQUARE INTEGRABLE FCN'S)

I.E, $f \in L^2(\mathbb{R}) \Leftrightarrow \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$.

LET $\langle f | g \rangle = \int_{-\infty}^{\infty} \overline{f(x)}\, g(x)\, dx$ BE THE INNER PRODUCT (THIS IS WELL DEFINED SINCE

$$|\langle f|g\rangle| \leq \|f\|\,\|g\| = \left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right)^{1/2} \left(\int_{-\infty}^{\infty} |g(x)|^2 dx\right)^{1/2} < \infty .)$$

IT IS EASY TO CHECK THAT $L = \frac{d}{dx}$ IS A LINEAR OPERATOR ON V (ASSUMING THAT ~~ALL~~ FCN'S IN V ARE DIFFERENTIABLE).

__Q:__ WHAT IS L^*?

__A:__ $$\langle f | L g \rangle = \int_{-\infty}^{\infty} f(x)\left(\frac{d}{dx} g(x)\right) dx$$

INTEGRATION BY PARTS

$$= \underbrace{\Big[f(x)\, g(x)\Big]_{-\infty}^{\infty}}_{=0} - \int_{-\infty}^{\infty} \left(\frac{d}{dx} f(x)\right) g(x)\, dx$$

$= 0$ SINCE SQUARE INTEGRABILITY $\Rightarrow f(x), g(x) \to 0$ AS $x \to \pm\infty$

$$= \int_{-\infty}^{\infty} \left(-\frac{d}{dx} f(x)\right) g(x)\, dx$$

$$= \langle L^* f | g \rangle$$

SO, $L^* = -\frac{d}{dx}$.

- SIMILARLY, IF $\tilde{L}$ WERE $\frac{d^2}{dx^2}$, WE WOULD FIND THAT $\tilde{L}^* = \frac{d^2}{dx^2}$ BY INTEGRATING BY PARTS TWICE. IN THIS CASE, $\tilde{L}^* = \tilde{L}$ AND $\tilde{L}$ WOULD BE CALLED __SELF ADJOINT__.

PROPERTIES OF ADJOINTS :-

$$\text{(i)}\quad (A+B)^* = A^* + B^*$$

$$\text{(ii)}\quad (AB)^* = B^* A^*$$

$$\text{(iii)}\quad (cA)^* = \bar{c} A^*$$

$$\text{(iv)}\quad (A^*)^* = A.$$

THM. (i) $(\text{Ker } A^*) = (\text{Ran } A)^{\perp}$

(ii) $(\text{Ker } A) = (\text{Ran}(A^*))^{\perp}$

(iii) $(\text{Ker } A^*)^{\perp} = \text{Ran } A$

(iv) $(\text{Ker } A)^{\perp} = \text{Ran } A^*$.

PF. (i) $\underline{x} \in (\text{Ran } A)^{\perp} \iff \langle \underline{x} \mid A \underline{y} \rangle = 0$ for all $\underline{y}$

$\iff \langle A^* \underline{x} \mid \underline{y} \rangle = 0$ for all $\underline{y}$

$\iff A^* \underline{x} = 0$

$\iff \underline{x} \in \text{Ker } A^*$.

Then, (i) $\iff$ (ii) $\iff$ (iv) $\iff$ (iii) and we are done.

LECTURE 29
04/04/12

DEF. AN OPERATOR $L: V \to V$ ON AN INNER PRODUCT SPACE V IS SELF-ADJOINT (OR HERMITIAN) IF $L^* = L$.

- IF V IS A REAL INNER PRODUCT SPACE, WE HAVE THAT $L^T = L$ AND L IS CALLED SYMMETRIC.

THM. EVERY EIGENVALUE OF A SELF-ADJOINT OPERATOR IS REAL.

PF. LET λ BE AN E-VALUE OF L WITH E-VECTOR $\underline{v}$. THEN,

$$\lambda \|\underline{v}\|^2 = \langle \underline{v} | \lambda \underline{v} \rangle = \langle \underline{v} | L\underline{v} \rangle = \langle L^* \underline{v} | \underline{v} \rangle$$
$$= \langle L\underline{v} | \underline{v} \rangle = \langle \lambda \underline{v} | \underline{v} \rangle = \bar{\lambda} \|\underline{v}\|^2.$$

$$\Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}.$$

THM. EIGENSPACES OF DISTINCT EIGENVALUES ARE ORTHOGONAL.

PF. LET λ, μ BE DISTINCT E-VALUES WITH CORRESPONDING EIGENSPACES E_λ AND E_μ. LET $\underline{v} \in E_\lambda$, $\underline{w} \in E_\mu$. THEN,

$$\langle L\underline{v} | \underline{w} \rangle = \bar{\lambda} \langle \underline{v} | \underline{w} \rangle = \lambda \langle \underline{v} | \underline{w} \rangle$$

SINCE $\lambda \in \mathbb{R}$.

$$\|$$

$$\langle \underline{v} | L\underline{w} \rangle = \mu \langle \underline{v} | \underline{w} \rangle$$

SO $\overbrace{(\lambda - \mu)}^{\neq 0} \langle \underline{v} | \underline{w} \rangle = 0$, AND $\langle \underline{v} | \underline{w} \rangle = 0 \Rightarrow \underline{v} \perp \underline{w}$.

NOTE: EIGENVECTORS CORRESPONDING TO SAME EIGENVALUE DON'T HAVE TO BE ORTHOGONAL !

EX.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow \lambda = 1, -1$$

$$\Rightarrow \underbrace{\underline{\xi}_1 = (1,0,0)^T, \ \underline{\xi}_2 = (1,1,1)^T}_{\text{CORRESPONDING TO } \lambda = 1}, \ \underbrace{\underline{\xi}_3 = (0,1,-1)^T}_{\text{CORRESPONDING TO } \lambda = -1.}$$

EASY TO SEE $\underline{\xi}_3 \perp \text{SPAN}\{\underline{\xi}_1, \underline{\xi}_2\}$. ✓

BUT $\underline{\xi}_1 \not\perp \underline{\xi}_2$! HOWEVER, $\underline{\xi}_1$ AND $\underline{\xi}_2$ FORM A <u>BASIS</u> OF E_1, SO WE CAN ORTHOGONALIZE BY GRAM-SCHMIDT.

$$\Rightarrow \underline{b}_1 = \underline{\xi}_1 = (1,0,0)^T$$

$$\underline{b}_2 = (I - P_{\underline{b}_1})\underline{\xi}_2 = \underline{\xi}_2 - P_{\underline{b}_1}\underline{\xi}_2$$

$$= \underline{\xi}_2 - \frac{\langle \underline{b}_1 | \underline{\xi}_2 \rangle}{\|\underline{b}_1\|^2}\underline{b}_1$$

$$= (1,1,1)^T - \frac{1}{1}(1,0,0)^T$$

$$= (0,1,1)^T$$ ← STILL AN EIGENVECTOR CORRESPONDING TO $\lambda = 1$!

$$\left.\begin{aligned} \Rightarrow \underline{e}_1 &= \frac{\underline{b}_1}{\|\underline{b}_1\|} = (1,0,0)^T \\ \underline{e}_2 &= \frac{\underline{b}_2}{\|\underline{b}_2\|} = \frac{1}{\sqrt{2}}(0,1,1)^T. \\ \underline{e}_3 &= \frac{\underline{\xi}_3}{\|\underline{\xi}_3\|} = \frac{1}{\sqrt{2}}(0,1,-1)^T \end{aligned}\right\} \text{ORTHONORMAL BASIS OF EIGENVECTORS OF } A.$$

THM. LET L BE SELF-ADJOINT. ASSUME L IS DIAGONALIZABLE. THEN THERE IS AN ORTHONORMAL BASIS OF V CONSISTING OF EIGENVECTORS OF L.

PF. L HAS DISTINCT E-VALUES $\lambda_1, \ldots, \lambda_r$, $r \leq n$, WITH EIGENSPACES $E_1, \ldots, E_r$. THEN ALL THE E-SPACES ARE ORTHOGONAL. NOW USE GRAM-SCHMIDT TO FIND ORTHONORMAL BASIS $\mathcal{E}_i$ FOR EACH E_i. SINCE L IS DIAGONALIZABLE, $\mathcal{E} = \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_r$ IS A BASIS OF ORTHONORMAL E-VECTORS FOR V.

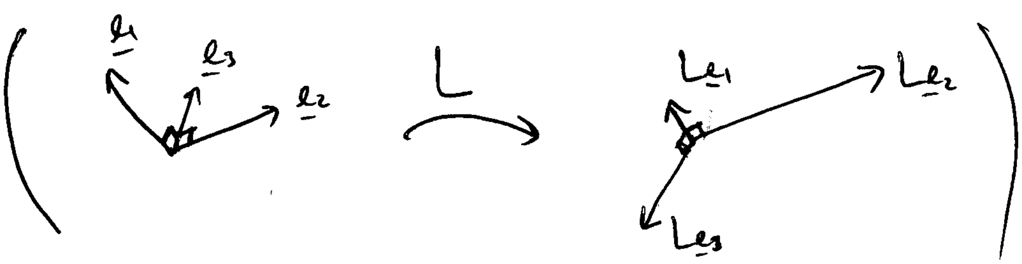
NOTE: WE WILL SHOW LATER THAT L SELF-ADJOINT IMPLIES L IS DIAGONALIZABLE, SO THE ASSUMPTION ABOVE CAN BE DROPPED.

THEREFORE,

THM. L SELF-ADJOINT $\Longleftrightarrow$ $L = UDU^{-1}$

COLUMNS ARE ORTHONORMAL EVECTORS OF L.

DIAGONAL MATRIX OF REAL EIGENVALUES OF L



PF. "$\Rightarrow$" ALREADY DONE.

"$\Leftarrow$" $L = UDU^{-1}$ WITH D REAL DIAGONAL

$$U = \underbrace{(e_1, \ldots, e_n)}_{\text{ORTHONORMAL BASIS } \mathcal{E}}.$$

SO, $[L]_{\mathcal{E}} = D = D^* = [L]_{\mathcal{E}}^* = [L^*]_{\mathcal{E}}$,

AND $L^* = L$.

__COROLLARY__: (REAL SPECTRAL THM.)

IF V __REAL__ INNER PRODUCT SPACE,

S SYMMETRIC $\iff$ $S = ODO^{-1}$

ORTHONORMAL E-VECTORS OF S.

DIAGONAL MATRIX OF REAL EIGENVALUES

__Q__: IS EVERY OPERATOR WHICH HAS AN ORTHONORMAL BASIS OF E-VECTORS SELF-ADJOINT?

__A__: NO. FOR EXAMPLE, ANY $L = UDU^{-1}$.

ORTHONORMAL E-VECTORS

DIAGONAL MATRIX w/ __COMPLEX__ ENTRIES.

__EX__ $A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$

$\Rightarrow \lambda_1 = 1 + 2i,\ \lambda_2 = 1 - 2i$

A __NOT__ SELF-ADJOINT SINCE $\lambda \notin \mathbb{R}$

$\underline{e}_1 = \frac{1}{\sqrt{2}}(1, i)^T, \quad \underline{e}_2 = \frac{1}{\sqrt{2}}(1, -i)^T$

$$\Rightarrow \langle \underline{e}_1 | \underline{e}_2 \rangle = 0 \Rightarrow \mathbf{A} = \underbrace{(\underline{e}_1\ \underline{e}_2)}_{U} \overbrace{\begin{pmatrix} 1+2i & \\ & 1-2i \end{pmatrix}}^{D} \underbrace{(\underline{e}_1\ \underline{e}_2)^{-1}}_{U^{-1}}$$