ESTIMATION & LEAST SQUARES:

PLAN:

- THIS WEEK: LEAST SQUARES
- LEAST SQUARES & PROBABILITY
- KALMAN FILTER:

NOTATION:

state/parameters

x : true state (unknown)
(parameters)

$\tilde{x}$: measured state (known)

$\hat{x}$: estimated state (compute)

noise terms

v : measurement noise ~ distribution
"noise in the sensor"

w : process noise ~ distribution
"noise in the model"
- inaccurate modeling
- dynamic noise

e : residual error

$e = \tilde{x} - \hat{x}$ state meas. vs. current estimate

LEAST SQUARES:

MODEL: $y(t) = \sum_{i=1}^{n} x_i h_i(t)$

$y(t)$: output

x_i: parameters

h_i: basis function

$h_i(t)$: something we can compute.

t: measurements vary w/ t

- take a sequence of measurements $(y(t))$
- use to fit the parameters x_i

Real World:

$$\begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_m) \end{bmatrix} = \begin{bmatrix} h_1(t_1) & \cdots & h_n(t_1) \\ h_1(t_2) & \cdots & h_n(t_2) \\ \vdots & & \vdots \\ h_1(t_m) & \cdots & h_n(t_m) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$\longleftarrow x \longrightarrow$

t

Measure:

$$\underbrace{\begin{bmatrix} \tilde{y}(t_1) \\ \vdots \\ \tilde{y}(t_m) \end{bmatrix}}_{\tilde{y}} = \underbrace{\begin{bmatrix} h_1(t_1) & \cdots & h_n(t_1) \\ & \vdots & \\ h_1(t_m) & \cdots & h_n(t_m) \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} v(t_1) \\ \vdots \\ v(t_m) \end{bmatrix}}_{\tilde{v}}$$

$\tilde{y}$: outputs we measure
H: data compute
v: noise
x: true parameters to estimate

---

linear in x but not necessarily lin in t

Ex: $y(t) = \boxed{a_1}\underline{\cos(t)} + \boxed{a_2}\underline{\sin(t)}$ ←

BASIS FNCS

$$x = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad H = \begin{bmatrix} \cos(t_1) & \sin(t_1) \\ \cos(t_2) & \sin(t_2) \\ & \vdots \end{bmatrix}$$

$h_1(t) = \cos(t)$
$h_2(t) = \sin(t)$

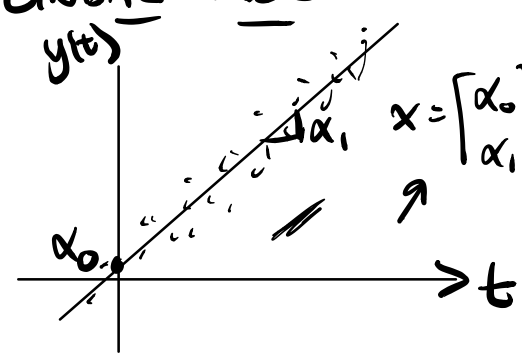
Ex. $y(t) = \boxed{\alpha_n}\underline{t^n} + \boxed{\alpha_{n-1}}\underline{t^{n-1}} + \cdots + \boxed{\alpha_1}\underline{t} + \boxed{\alpha_0}$

$$x = \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{bmatrix} \qquad H = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^n \\ 1 & t_2 & t_2^2 & \cdots & t_2^n \\ & & \vdots & & \end{bmatrix}$$

BASIS
$h_1(t) = 1$
$h_2(t) = t$
$h_3(t) = t^2$
$\vdots$

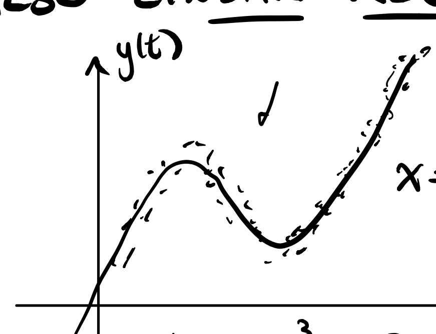
curve fitting: different basis functions

LINEAR REG.



$x = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}$

ALSO LINEAR REG

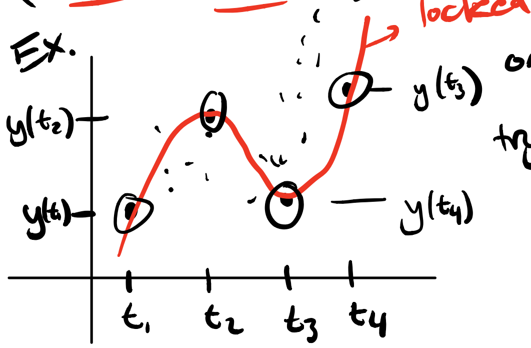


$x = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$

$y(t) = \alpha_3 t^3 + \alpha_2 t^2 + \alpha_1 t + \alpha_0$

more parameters → requires more ↗
(more coeffs.) data

Ex.



only 4 data points

trying to fit

$$y(t) = \alpha_3 t^3 + \alpha_2 t^2 + \alpha_1 t + \alpha_0$$

$$H = \begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ 1 & \vdots & & \vdots \\ 1 & & & \\ 1 & t_4 & t_4^2 & t_4^3 \end{bmatrix}$$

→ $y = \underset{\sim}{H} x$ ←

⇒ $x = H^{-1} y$. BAD

any noise in an individual meas. → big impact

NOTE: x is constant in time. ←

values of basis functions change in "time"

if x is changing w/ time:

→ need measurements to come much faster than changes in x...

→ treating x as constant over a short time interval.

OR. try to predict how x is changing

model the dynamics x. ⇐

compare $\underline{y}$ w/ what we expect x to be...

KALMAN FILTER
(dynamics: Linear)

SOLVING:

MODEL: $\underset{\mathbb{R}^m}{\underline{y}} = \underset{\mathbb{R}^{m \times n}}{H}\, \underset{\mathbb{R}^n}{x}$ ⇐

Measure: $\tilde{y} = Hx + \underline{v} \Rightarrow \tilde{y} \notin \mathcal{R}(H)$

if we try to solve $\tilde{y} = Hx \Longrightarrow$ no solution

no noise... → $\tilde{y} = Hx$ H tall
(no solution)

$\tilde{y}$ most likely not in range of H

$$\begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \vdots \\ \tilde{y}_m \end{bmatrix} = m \begin{bmatrix} \overset{n}{\overline{\quad\quad}} \\ \\ H \\ \\ \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

System of Eqns:

$y = Hx$	$y = Hx$	$y = Hx$
H tall	H square	H fat.
↓	↓	↓
no solns	unique solution $x = H^{-1}y$	subspace of solns.

try to find x. that assumes noise has as little impact as possible

find x.

$$\min_x \left|\tilde{y} - y\right|^2 = \left|\underbrace{\tilde{y}}_{\text{measure}} - \underbrace{Hx}_{\text{what we expect true to be}}\right|^2$$

solving:

$$\min_x \; (\tilde{y} - Hx)^T(\tilde{y} - Hx) = J(x)$$

$$\frac{\partial J}{\partial x} = 0 : \quad \frac{\partial}{\partial x}\left(\tilde{y}^T\tilde{y} - 2\tilde{y}^T Hx + x^T H^T H x\right)$$

$$-2\tilde{y}^T H + 2x^T H^T H = 0$$

or use chain rule . . .

$$J(x) = z^T z \quad \text{where} \quad z = \tilde{y} - Hx$$

$$\frac{\partial J}{\partial x} = \left.\frac{\partial J}{\partial z}\right|_{z = \tilde{y} - Hx} \frac{\partial z}{\partial x} = \underbrace{2z^T}_{\frac{\partial J}{\partial z}} \underbrace{(-H)}_{\frac{\partial z}{\partial x}}$$

$$= 2(\tilde{y} - Hx)^T(-H)$$

$$= -2\tilde{y}^T H + 2x^T H^T H = 0$$

$$x^T = \tilde{y}^T H \left(\underline{H^T H} \right)^{-1}$$

need $(\underline{H^T H})^{-1}$ to exist $\Updownarrow$ H is full col rank.

$$\boxed{x = (H^T H)^{-1} H^T \tilde{y}} \Leftarrow \text{least squares solution}$$

Ex. $\underline{\tilde{y}} = H \underline{x}$ $\quad H \in \mathbb{R}^{3 \times 2}$

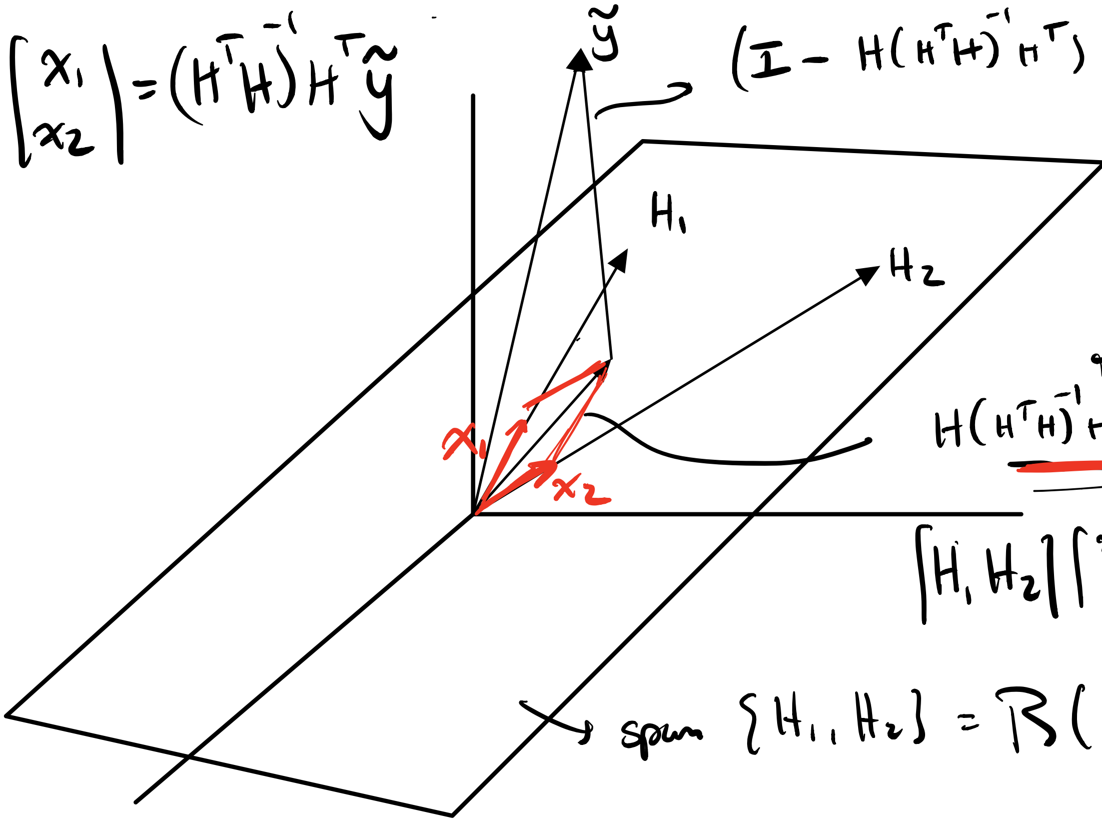
$H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}$

$\alpha_n s^n + \cdots \alpha_1 s + \alpha_0$

$\alpha_n' s^n \qquad + \alpha_0'$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (H^T H)^{-1} H^T \tilde{y}$$

$(I - H(H^T H)^{-1} H^T)\tilde{y}$



$H(\underline{H^T H)^{-1} H^T \tilde{y}}$

$\begin{bmatrix} H_1 & H_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

span $\{H_1, H_2\} = \mathcal{R}(H)$

if x is the least squares solution:

Hx is projection of $\tilde{y}$ onto the range of H