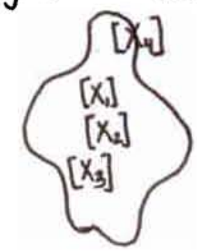
6.581 / 20.482

LECTURE 17: PARAMETER FITTING & ESTIMATION

THURSDAY
13 APRIL 2006

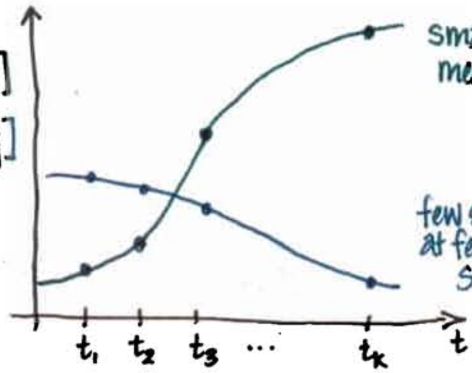
Dynamical Net Cell



known:
$[u_1]$
$[u_2]$

$$\dot{\vec{X}} = A^{(1)}\vec{X} + A^{(2)}(\vec{X}\otimes\vec{X}) + B^{(1)}\vec{u} + \underbrace{B^{(2)}(\vec{u}\otimes\vec{X}) + B^{(3)}(\vec{u}\otimes\vec{u})}_{\text{ignored for simplicity}}$$

$$\frac{d}{dt}\begin{bmatrix}[X_1]\\ \vdots \\ [X_n]\end{bmatrix}$$



what was $\vec{X}(0)$? (Initial Condition)

ESTIMATING INITIAL CONDITIONS

1.) START by looking at the linear case

$$\dot{\vec{X}} = A^{(1)}\vec{X} + B^{(1)}\vec{u}$$

2.) NONLINEAR case

$$\dot{\vec{X}} = A^{(1)}\vec{X} + A^{(2)}(\vec{X}\otimes\vec{X}) + B^{(1)}\vec{u}$$

LINEAR CASE

$$\dot{\vec{X}} = A\vec{X} + B\vec{u}$$

$$\vec{X}(t) = \underbrace{e^{At}}_{\text{MATRIX}}X(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

Scalar case:

$$\dot{x} = ax + \cancel{bu}$$

$$x(t) = e^{at}x(0)$$

$$\dot{x} = ax + bu$$

$$x(t) = e^{at}x(0) + \int_0^t e^{a(t-\tau)}bu(\tau)d\tau$$

$$[X_i(t_k)] = 1\updownarrow \underbrace{\begin{array}{c} \overset{n}{\longleftrightarrow} \\ \begin{bmatrix} 0 \dots 0 & 1 & 0 \dots 0 \end{bmatrix} \\ \uparrow_{i\text{th column}} \end{array}}_{C} \begin{bmatrix} X_1(t_k) \\ \vdots \\ X_n(t_k) \end{bmatrix}$$

$$\begin{bmatrix}[X_i(t_k)]\\ [X_j(t_k)]\end{bmatrix} = 2\updownarrow \underbrace{\begin{array}{c} \overset{n}{\longleftrightarrow} \\ \begin{bmatrix} 0 \cdots 0 1 0 \cdots 0 0 0 \cdots 0 \\ 0 \cdots 0 0 0 \cdots 0 1 0 \cdots 0 \end{bmatrix} \\ \uparrow_{i\text{th}} \qquad \uparrow_{j\text{th}} \end{array}}_{C} [\vec{X}(t_k)]$$

$$\underbrace{\begin{bmatrix}\vec{X}_M(t_1)\\ \vdots \\ \vec{X}_M(t_k)\end{bmatrix}}_{\text{measurements}} = \underbrace{\begin{bmatrix} Ce^{At_1} \\ \vdots \\ Ce^{At_k}\end{bmatrix}}_{Mk,\ M = \#\text{ of measurements}} [\vec{X}(0)] + \underbrace{\begin{bmatrix}\int_0^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau \\ \vdots \\ \int_0^{t_k} e^{A(t_k-\tau)}Bu(\tau)d\tau\end{bmatrix}}_{\text{inputs}}$$

COMPARE MEASURED DATA TO THE SIMULATION VIEW

$$\begin{bmatrix} C\vec{X}(t_1) \\ \vdots \\ C\vec{X}(t_k)\end{bmatrix} \qquad \begin{bmatrix} \vec{b}(t_1) \\ \vdots \\ \vec{b}(t_k)\end{bmatrix}$$

$$\min_{\vec{X}_0} \text{measure}\left\{ \vec{Y} - L\vec{X}_0 \right\}$$

$$\vec{Y} = \begin{bmatrix} X_M(t_1) \\ \vdots \\ X_M(t_k)\end{bmatrix} - \begin{bmatrix} \vec{b}(t_1) \\ \vdots \\ \vec{b}(t_k)\end{bmatrix} \qquad L = \begin{bmatrix} Ce^{At_1} \\ \vdots \\ Ce^{At_k}\end{bmatrix}$$

$$\min_{\vec{X}_0} \max_i |(\vec{Y} - L\vec{X}_0)_i|$$

$$\underbrace{\min_{\vec{X}_0} \sum_i (\vec{Y} - L\vec{X}_0)_i^2}_{\text{LEAST SQUARES}}$$

$$\min_{\vec{X}_0} \underbrace{\underbrace{(\vec{Y} - L\vec{X}_0)^T}_{\vec{e}^T} \underbrace{(\vec{Y} - L\vec{X}_0)}_{\vec{e}} = \sum \vec{e}_i^2}_{f(\vec{X}_0)}$$

LINEAR LEAST SQUARES

- Dynamical system is linear
  $\dot{\vec{X}} = A^{(1)}\vec{X} + B\vec{u} \dots$
- Unknowns are initial conditions
- Measured quantities were linear functions of $\vec{X}$
  $C\vec{X}(t)$

ABSTRACT PROBLEM

$\underbrace{f(\vec{x})}_{\text{cost function}}$ $\quad f: \mathbb{R}^n \to \mathbb{R} \quad f$: vector $\to$ scalar

$\min_{\vec{x}} f(\vec{x}) \leftarrow$ optimization problem

SCALAR CASE

$f(x^*) = \min_x f(x)$

$\frac{\partial f(x^*)}{\partial x} = 0$ (local min or max)

$$\text{VECTOR CASE}\quad \frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 0 \quad \left(\begin{matrix}\text{local}\\ \text{min or}\\ \text{max}\end{matrix}\right)$$

vector function or a vector

For some functions f

$\exists\, x^*$ s.t. $\frac{\partial f(x^*)}{\partial x} = 0$

means x^* is the global optimum

$$\min_{\vec{X}_0} (\vec{Y} - L\vec{X}_0)^T(\vec{Y} - L\vec{X}_0) = f(\vec{X}_0)$$

$$\frac{\partial f}{\partial \vec{X}_0} = \frac{\partial}{\partial \vec{X}_0}\left[\vec{Y}^T\vec{Y} - \underbrace{(L\vec{X}_0)^T\vec{Y} - \vec{Y}^T(L\vec{X}_0)}_{\text{equal}} + (L\vec{X}_0)^T(L\vec{X}_0)\right]$$

$$= \underbrace{\frac{\partial}{\partial \vec{X}_0}\left[\vec{Y}^T\vec{Y}\right]}_{0} - \frac{\partial}{\partial \vec{X}_0}\left[2\vec{Y}^T(L\vec{X}_0)\right] + \frac{\partial}{\partial \vec{X}_0}\left[(L\vec{X}_0)^T(L\vec{X}_0)\right]$$

$$= -2L^T\vec{Y} + 2L^TL\vec{X}_0 = 0$$

$$\Rightarrow \underbrace{L^TL\vec{X}_0 = L^T\vec{Y}}_{\text{NORMAL EQNS}}$$

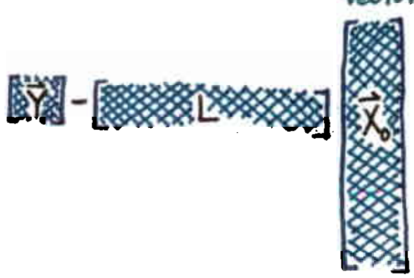
lots of data

$$\begin{array}{c} \\ mk \end{array}\overset{\longleftarrow n \longrightarrow}{\begin{bmatrix} Ce^{At_1} \\ Ce^{At_2} \\ \vdots \\ Ce^{At_k} \end{bmatrix}}\begin{bmatrix}\vec{X}_0\end{bmatrix}$$

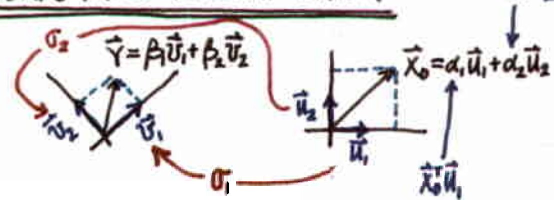
limited data → typical

$$\begin{bmatrix} Ce^{At_1} \\ \vdots \\ Ce^{At_k} \end{bmatrix}\begin{bmatrix} \\ \vec{X}_0 \\ \\ \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} L^T \end{bmatrix}\begin{bmatrix} \overset{\longleftarrow n \longrightarrow}{mk \; L} \end{bmatrix}}_{n\times n}\begin{bmatrix}\vec{X}_0\end{bmatrix} = \underbrace{\begin{bmatrix} L^T \end{bmatrix}\begin{bmatrix}\vec{Y}\end{bmatrix}}_{n \text{ length vector}}$$



SINGULAR VALUE DECOMPOSITION



$$\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_{mk}\end{bmatrix} = \begin{bmatrix} \longleftarrow u_1^T \longrightarrow \\ \vdots \\ \longleftarrow u_{mk}^T \longrightarrow \end{bmatrix}\begin{bmatrix} \\ \vec{X}_0 \\ \\ \end{bmatrix}$$

$$\begin{bmatrix} \uparrow & & \uparrow \\ \vec{V}_1 & \cdots & \vec{V}_{mk} \\ \downarrow & & \downarrow \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{mk}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_{mk}\end{bmatrix}$$

$$\vec{Y} = \overbrace{\underbrace{\begin{bmatrix} \uparrow & & \uparrow \\ \vec{V}_1 & \cdots & \vec{V}_{mk} \\ \downarrow & & \downarrow \end{bmatrix}}_{\substack{\vec{V}_i^T\vec{V}_j = 0 \\ \text{for } i \neq j \\ \text{(orthogonal coordinate system)}}}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{mk}\end{bmatrix}\underbrace{\begin{bmatrix} \longleftarrow \vec{u}_1 \longrightarrow \\ \vdots \\ \longleftarrow \vec{u}_{mk} \longrightarrow \end{bmatrix}}_{\substack{\vec{u}_i^T\vec{u}_j = 0 \\ \text{for } i \neq j}}}^{\text{SVD } L}\begin{bmatrix} \\ \vec{X}_0 \\ \\ \end{bmatrix}$$

$$\begin{bmatrix} \uparrow & & \uparrow \\ \vec{u}_1 & \cdots & \vec{u}_{mk} \\ \downarrow & & \downarrow \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_{mk}}\end{bmatrix}\begin{bmatrix} \longleftarrow \vec{V}_1^T \longrightarrow \\ \vdots \\ \longleftarrow \vec{V}_{mk}^T \longrightarrow \end{bmatrix}\vec{Y} = \vec{X}_0$$