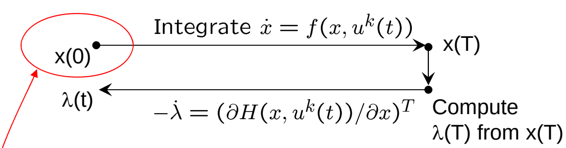
Estimation

D. MacMynowski
Nov 10, 2009

Optimization…

- For a control input u^k(t), compute the sensitivity:



$$\frac{\delta \tilde{J}}{\delta u} = \frac{\partial L}{\partial u} + \lambda^T \frac{\partial f}{\partial u}$$

- For linear system:

$$-\dot{P} = PA + A^T P + Q - PBR^{-1}B^T P \qquad P(T) = 0$$
$$u = -R^{-1}B^T P x$$

- Need model of the system (that is sufficiently good)
- Need to know current state!

Goal

- Implementing optimal control requires estimate of the entire state, not just the measured variables
  - How do we reconstruct this estimate?
  - How do we quantify how "good" the estimate is?
- We can build an estimator if the system is observable

A system is *observable* if for any $T > 0$, the state of the system can be determined uniquely from $y(t),\ t \in [0, T]$.

- Two approaches:
  - Understand covariance propagation & minimize
  - Pose as an optimization problem
- Basic approach: follow dynamics, with linear correction

$$\dot{x} = f(x,u) \qquad y = h(x,u)$$

$$\dot{\hat{x}} = \underbrace{f(\hat{x},u)}_{\text{prediction (copy of dynamics)}} + \underbrace{L(y - h(\hat{x},u))}_{\text{correction (based on output error)}}$$

D. G. MacMynowski
Ae240

3

Covariance Propagation

- System: $\dot{x} = Ax + Bu + Fv \qquad E\{v(t)v(t+\tau)^T\} = \Sigma_v \delta(\tau)$

  $y = Cx + n \qquad E\{n(t)n(t+\tau)^T\} = \Sigma_n \delta(\tau)$
  - where process noise *v* and measurement noise *n* are both white
    - Any spectrum of *v*, *n* can be captured in system dynamics
  - Assume *v* and *n* are uncorrelated so $E\{vn^T\}=0$, and also $E\{x_0 v(\tau)^T\}=0$
  - Ignore control input, since can always subtract off
- Covariance of state is $P(t) = E\{x(t)x^T(t)\}$
- Substitute

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Fv(\tau)$$

- Then

$$P(t) = e^{At}P_0 e^{A^T t} + \int_0^t e^{A(t-\tau)} F \Sigma_v F^T e^{A^T(t-\tau)} d\tau$$

- Hence:

$$\boxed{\dot{P} = AP + PA^T + F\Sigma_v F}$$

Change due to state evolution ($AP + PA^T$)

Increment from disturbance ($F\Sigma_v F$)

D. G. MacMynowski
Ae240

4

Brief Aside…

$$P_e(t) = e^{At}P_0 e^{A^T t} + \int_0^t e^{A(t-\tau)} F \Sigma_v F^T e^{A^T(t-\tau)} d\tau$$

$$\dot{P}_e = AP_e + P_e A^T + F\Sigma_v F$$

- With zero initial condition, the state covariance is the finite-time controllability Gramian for the system with disturbance input (weighted by disturbance covariance)
- Can derive LQR by writing the cost for the system $\dot{x} = Ax$ as

$$\begin{aligned} J(t) &= \int_t^T x^T Q x \, \mathrm{d}\tau \\ &= \int_t^T x(t)^T e^{A^T(\tau-t)} Q e^{A(\tau-t)} x(t) \mathrm{d}\tau \\ &= x(t)^T \left(\int_t^T e^{A^T(\tau-t)} Q e^{A(\tau-t)} \mathrm{d}\tau \right) x(t) \\ &\triangleq x(t)^T P_c(t) x(t) \end{aligned}$$

$$-\dot{P}_c = A^T P_c + P_c A + Q$$

$$\dot{x} = Ax + Bu, \quad u = -Kx$$

$$\Rightarrow \dot{x} = (A - BK)x$$

Substitute closed-loop dynamics & minimize wrt K

- In linear-quadratic optimal control problem, the matrix P_c (the remaining cost-to-go), is the observability Gramian for the system

Kalman Filter

- Minimize $E\{(x-\hat{x})^T(x-\hat{x})\}$ or $\hat{x} = E\{x(t)|y(\tau), \ \tau \le t\}$

  Minimum variance — Conditional mean

- The optimal estimate is a linear observer:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$$

- Error dynamics satisfy:

$$\dot{e} = (A - LC)e + \xi \qquad \xi = Fv - Ln \qquad \Sigma_\xi = F\Sigma_v F^T + L\Sigma_n L^T$$

- Substitute into covariance evolution and minimize:

$$\dot{P} = (A - LC)P + P(A - LC)^T + (F\Sigma_v F + L\Sigma_n L^T)$$

$$\Rightarrow \quad \begin{aligned} \dot{P} &= AP + PA^T + F\Sigma_v F^T - PC^T\Sigma_n^{-1}CP \\ L &= PC^T\Sigma_n^{-1} \end{aligned}$$

D. G. MacMynowski
Ae240

6

Notes

- Credit generally given to Kalman (1960) in discrete-time, and Kalman-Bucy (1961) in continuous time
- Gives both the estimate and the covariance $\Rightarrow$ see how well you are converging
- Straightforward to include cross-covariance
- With optimal estimate, residual process $r = y - C\hat{x}$ is white
  - No remaining information
- If noise is stationary (covariances independent of time) then both the estimator gain and the error covariance converge to constant values
- Filtering is causal (estimate current state with past measurements).
  - Can use same framework for *smoothing* (a posteriori improvement of prior state estimate based on current measurements)

Discrete-Time

- System: $x_{k+1} = Ax_k + Fv_k \qquad E\{v_k v_j^T\} = \Sigma_v \delta(k-j)$

  $y_k = Cx_k \qquad E\{n_k n_j^T\} = \Sigma_n \delta(k-j)$

- Covariance:

$$P_{k+1} = \{x_{k+1} x_{k+1}^T\}$$
$$= AP_k A^T + F\Sigma_v F^T$$

- Note that given y_0, y_1,…y_k, the problem could be cast as a giant minimum variance (weighted least-squares) problem
  - Estimate initial condition and all disturbances given measurements is of the form Y=MX, hence $X=(M^TM)^{-1}M^TY$
  - Recursive solution is much simpler!

Discrete-Time Kalman Filter

- Initialize: $\hat{x}_{0|-1} = E\{x_0\}$

  $P_{0|-1} = E\left\{x_0 x_0^T\right\}$

Best estimate with measurement

Best estimate before measurement

- Correction: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(y - C\hat{x}_{k|k-1})$

  Prediction: $\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k$

- Correction: $P_{k|k} = P_{k|k-1} - P_{k|k-1}C^T\left(CP_{k|k-1}C^T + \Sigma_n\right)^{-1}CP_{k|k-1}$

  Prediction: $P_{k+1|k} = AP_{k|k}A^T + F\Sigma_v F^T$

  where

$$L = P_{k|k-1}C^T\left(CP_{k|k-1}C^T + \Sigma_n\right)^{-1} = P_{k|k}C^T\Sigma_n^{-1}$$

In "information" form:

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + C^T\Sigma_n^{-1}C$$

The covariance inverse is the "information"

Ae240

Nonlinear?

- Linearize system about current state ("extended Kalman filter")
  - This works because the estimation error is hopefully small and so the error dynamics are approximately linear
- EKF works for parameter estimation too

$$\dot{x} = f(x,u,p) \qquad \Rightarrow \qquad \frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} f(x,u,p) \\ 0 \end{bmatrix}$$

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} \simeq \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial p} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} \frac{\partial f}{\partial u} \\ 0 \end{bmatrix} u$$

- Strong non-linearity means non-Gaussian statistics
  - Basic idea is particle filter to track higher order moments
  - Unscented Kalman filter: choose specific particles ("sigma points") at each time step & propagate them to remove bias in estimating mean & covariance

D. G. MacMynowski
Ae240 10

Estimation is Optimization

- Goal is to estimate unknown initial condition x_0, unknown disturbances *d*, and maybe unknown parameters *p*.
- Define cost functional:

$$J = \underbrace{\int_0^t \left\| y(\tau) - C\hat{x}(\tau) \right\| d\tau}_{\text{Residual}} + \underbrace{\left\| \hat{x}_0 - \hat{x}_{0|0} \right\|}_{\text{Initial condition estimate (minus prior)}} + \underbrace{\left\| \hat{p} - \hat{p}_{|0} \right\|}_{\text{parameter estimate (minus prior)}} + \underbrace{\left\| \hat{d} \right\|}_{\text{Disturbance estimate}}$$

  - And choose appropriate inner product on each norm
  - This cost minimizes some weighting between the estimation error and the “size” of the disturbance (etc) error
  - ***What is the minimum norm explanation for the observed sensor time history?***
- Now include dynamics with a Lagrange multiplier
  - Same machinery as before

Separation Principle / Theorem

1. If we have
   a) a system that is stable under linear state feedback *u* = -K*x*
   b) and an estimator that is stable, $\lim_{t\to\infty} \hat{x} - x = 0$, then
   c) the system with feedback of the estimate $u = -K\hat{x}$ is stable

$$\left.\begin{aligned} \dot{x} &= Ax + Bu + Fv \\ y &= Cx + n \\ \dot{\hat{x}} &= A\hat{x} + Bu + L(y - C\hat{x}) \\ u &= -K\hat{x} \end{aligned}\right\} \Rightarrow \frac{d}{dt}\begin{bmatrix} x \\ e \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}\begin{bmatrix} x \\ e \end{bmatrix} + \cdots$$

2. If the state feedback is optimal for the deterministic problem and the estimator is optimal, then the combination is *optimal* for the stochastic problem!

$$J = E\left\{ \lim_{T\to\infty} \frac{1}{T} \int_0^T \left(x^T Q x + u^T R u \right) dt \right\}$$

LQG = Linear Quadratic Gaussian

- Linear plant
- Quadratic cost
- Gaussian statistics
- Then optimal control is

$$
\begin{aligned}
u &= -K\hat{x} \\
\dot{\hat{x}} &= A\hat{x} + Bu + L(y - C\hat{x}) \\
K &= -R^{-1}B^T P_c \\
L &= P_e C \Sigma_n^{-1} \\
\dot{P}_c &= PA_c + A^T P_c + Q - P_c B R^{-1} B^T P_c \\
\dot{P}_e &= PA_e^T + AP_e + F\Sigma_v F^T - P_e C^T \Sigma_n^{-1} C P_e
\end{aligned}
$$

Note that estimation and control are "dual" problems: control of (A,B,Q,R) is the same as estimation for (A^T,C^T,$F\Sigma_v F^T$,Σ_n)

Examples

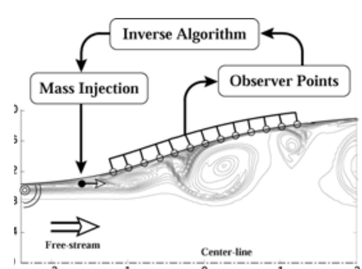
- Could estimate entire flow state using a few measurements
  - Clearly not practical in real-time
- Could estimate coefficients of a few POD modes in real-time
  - E.g. Glauser *et al.*
- Takao Suzuki: estimate location of a vortex from wall measurements
- Estimate frequency of a nearly-periodic shedding
  - E.g. Tadmor

Estimation techniques in a turbulent flow field

Julie Ausseur, Jeremy Pinier and Mark Glauser

Syracuse University 149 Link Hall Syracuse Ny 13244, USA jausseur@syr.edu

Part of our continuous efforts to implement an effective closed-loop feedback control of the flow over a NACA 4412 airfoil is the obtention of an accurate estimate of the actual flow state. An elaborate controller combines both prediction and measurement techniques to obtain a precise estimation of the control variable. In this paper we focus on the measurement process. We first present the different candidates for the control variable, and describe the low-dimensional techniques employed. We then focus on the estimation methods that will be incorporated in the controller in order to access the state variable from the pressure real-time measurements. The investigation of the dynamics in the flow field and the correlations between the variables at stake reveal the benefits that a spectral estimation approach will bring.



Frequency Tracking

- System model: $y(t) = c_0 + c_1 \sin\theta(t) + n, \qquad \dot{\theta}(t) = \omega(t)$
  - Mean (c_0), Periodic component ($c_1 \sin\theta(t)$), Noise (n)
- Write in state-space:
  - "Disturbance" reflects fact that these are slowly-varying parameters

$$\begin{bmatrix}\omega\\ \theta\\ c_0\\ c_1\end{bmatrix}_{k+1} = \begin{bmatrix}1 & 0 & 0 & 0\\ \Delta t & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\omega\\ \theta\\ c_0\\ c_1\end{bmatrix}_k + w_k$$

  - Measurement equation: take local linearization about current estimate:

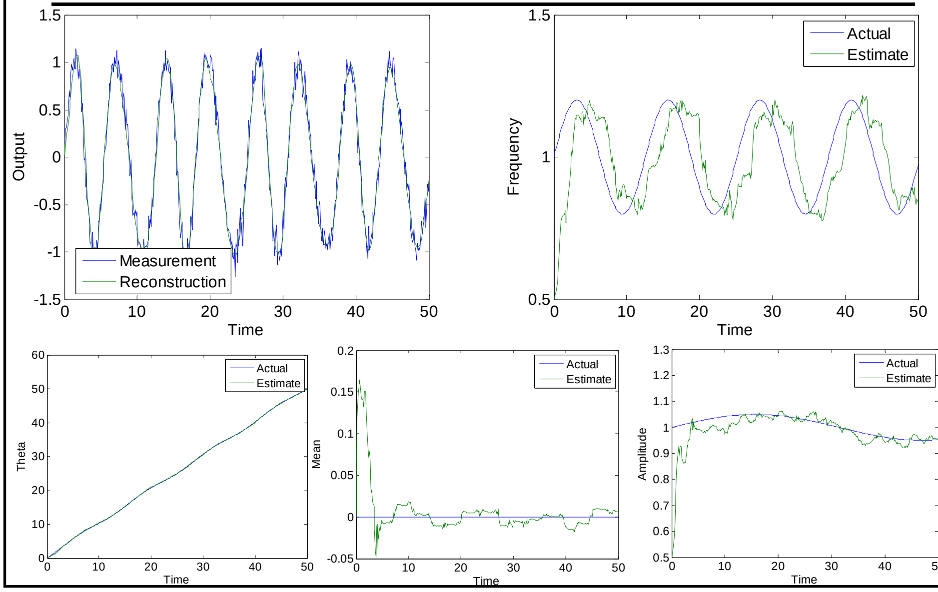
$$y = \begin{bmatrix}0 & (\hat{c}_1)_k \cos\hat{\theta}_k & 1 & \sin\hat{\theta}\end{bmatrix}\begin{bmatrix}\omega\\ \theta\\ c_0\\ c_1\end{bmatrix}_k + n_k$$

Matlab code:

```
% do a simple KF example:
%
% DGM 11/10/09
%%
%Some coefficients:
noiseamp=.1;
%% Set up data:
t=linspace(0,50,500);
omvec=1+0.2*sin(t/2);
thvec=cumtrapz(t,omvec);
plot(t,sin(thvec))
avec=1+0.05*sin(t/10);
yvec=avec.*sin(thvec)+
     noiseamp*randn(size(t));
nt=length(t);
dt=t(2)-t(1);
%
%% Model:
A=eye(4);A(2,1)=dt;
Q=diag([1 0.00 0.01 0.1]);
R=50;  % no clue, off hand
```

```
%% Kalman filter:
% initialize:
xh=zeros(4,nt);xh(1,1)=.5;xh(4,1)=.5;
r=zeros(nt,1);
Pdiag=zeros(4,nt);
P=1e2*eye(4);
% now loop:
for k=2:nt,
    % prediction step:
    xh(:,k)=A*xh(:,k-1);
    P=A*P*A'+Q;
    % get linearized output equation:
    C=[0 xh(4,k)*cos(xh(2,k)) 1 sin(xh(2,k))];
    % get residual:
    r(k)=yvec(k)-(xh(3,k)+xh(4,k)*sin(xh(2,k)));
    % measurement update:
    L=P*C'/(C*P*C'+R);
    P=(eye(4)-L*C)*P;
    xh(:,k)=xh(:,k)+L*r(k);
    % save covariance info:
    Pdiag(:,k)=diag(P);
end
```

Results:



Summary

- As long as the system is *observable*, the system state can be estimated from measurements
  - Relies on system model
- Optimal trade-off exists in balancing convergence speed with responding to sensor noise
  - What is the minimum-norm (minimum covariance) explanation for the observation?
- Residual (or "innovations") process is white
  - No remaining information
- Applies to nonlinear systems as well
  - Uses local linearization (assuming estimation error sufficiently small)
- Can use for parameter estimation as well