Nonstationary processes

Definition: $x(t)$ is a Markov stochastic process if

$$f_X\big(x(t) \mid x(t_1), x(t_2) \dots x(t_k)\big) = f_X\big(x(t) \mid x(t_k)\big),$$

$$\forall\, t > t_k > t_{k-1} > \dots > t_2 > t_1 \quad , \quad \forall\, k \in \mathbb{N}_+$$

A (large) class of Markov processes

$$x(t+1) = f\big(x(t), u(t)\big) + g\big(x(t), w(t)\big)$$
$$y(t) = h\big(x(t), u(t)\big) + v(t)$$

- $w(t)$ and $v(t)$ are white processes such that

$$E\left[\begin{pmatrix} w(t) \\ v(t) \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$E\left[\begin{pmatrix} w(t) \\ v(t) \end{pmatrix}\begin{pmatrix} w(t) \\ v(t) \end{pmatrix}^T\right] = \begin{pmatrix} Q & \emptyset \\ \emptyset & R \end{pmatrix}$$

- $w(t)$ and $v(t)$ are independent
- $x(0)$ is a random vector, independent from $w(t)$ and $v(t)$, such that

$$E[x(0)] = m_0 \quad E\left[(x(0)-m_0)(x(0)-m_0)^T\right] = P_0$$

- $u(t)$ is a deterministic signal.

Then, $x(t)$ is a <u>Markov process</u>.

<u>LTI case</u> (same assumptions on $w(t)$, $v(t)$, $x(0)$)

$$x(t+1) = Ax(t) + Bu(t) + Gw(t)$$
$$y(t) = Cx(t) + Du(t) + v(t)$$

$x \in \mathbb{R}^n \quad u \in \mathbb{R}^m \quad y \in \mathbb{R}^p \quad w \in \mathbb{R}^d$

$w(t)$: disturbance process
$v(t)$: measurement noise

$\rightarrow$ $x(t)$ is a Markov process

Result $\left(\begin{smallmatrix} B=0 \\ D=0 \end{smallmatrix}\right)$

Let $m_x(t) = E[x(t)]$, $m_y(t) = E[y(t)]$,

$R_x(t,s) = E\left[\left(x(t)-m_x(t)\right)\left(x(s)-m_x(s)\right)^T\right]$.

Then,

* $m_x(t) = A^t m_x(0) = A^t m_0$
* $R_x(t+\tau, t) = A^\tau R_x(t,t) \triangleq A^\tau P(t)$

  where $\tau > 0$, and

  $P(t+1) = A\,P(t)\,A^T + G Q G^T$
* $m_y(t) = C\, m_x(t)$
* $R_y(t+\tau, t) = C A^\tau P(t)\, C^T \quad \tau > 0$

  and

  $R_y(t,t) = C P(t)\, C^T + R$

Key fact: the result holds also if matrices A, G, C, Q, R are time-varying!

Result (LTI case)

If all the eigenvalues of A, λ_i $i = 1 \ldots n$, satisfy $|\lambda_i| < 1$, and Q is constant,

$$\lim_{t \to +\infty} P(t) = \bar{P} \qquad \forall P(0) > 0$$

where $\bar{P} = \bar{P}^T \geq 0$ is the unique solution of $\bar{P} = A \bar{P} A^T + G Q G^T$

and $x(t)$ and $y(t)$ are asymptotically stationary processes.

THE STATE ESTIMATION PROBLEM

Consider the linear stochastic system

$$x(t+1) = Ax(t) + Bu(t) + Gw(t)$$
$$y(t) = Cx(t) + v(t)$$

where $u(t)$ is a known deterministic signal,

$$\begin{pmatrix} w(t) \\ v(t) \end{pmatrix} \sim WN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix} \right),$$

$x(0)$ is a random vector with mean m_0, covariance P_0, independent on $w(t)$ and $v(t)$.

Find an estimate of $x(t)$ based on the knowledge of $u(k)$, $\forall k$ and

$$Y^t = \{ y(t), y(t-1), \dots\dots, y(0) \}.$$

The matrices $A = A(t)$, $B = B(t)$, $G = G(t)$, $C = C(t)$, $Q = Q(t)$, $R = R(t)$ are known at every time t (but not necessarily constant)

Notation

LMSE: linear mean square estimate

$\hat{x}(t|t)$: LMSE of $x(t)$ based on Y^t

$\hat{x}(t+1|t)$: LMSE of $x(t+1)$ based on Y^t
(one-step-ahead prediction)

$$P(t|t) = E\left[\left(x(t) - \hat{x}(t|t)\right)\left(x(t) - \hat{x}(t|t)\right)^T\right] :$$

covariance of estimation error at time t

$$P(t+1|t) = E\left[\left(x(t+1) - \hat{x}(t+1|t)\right)\left(x(t+1) - \hat{x}(t+1|t)\right)^T\right]$$

covariance of prediction error at time t

A two-step recursive solution

* Step 1: prediction
  Given $\hat{x}(t|t)$, $P(t|t)$, compute
  $\hat{x}(t+1|t)$, $P(t+1|t)$
* Step 2: Correction
  Given $\hat{x}(t+1|t)$, $P(t+1|t)$ and $y(t+1)$, compute
  $\hat{x}(t+1|t+1)$, $P(t+1|t+1)$
* iterate

1st step: prediction

$$\hat{x}(t+1|t) = A\,\hat{x}(t|t) + Bu(t)$$

$$P(t+1|t) = A\,P(t|t)\,A^T + GQG^T$$

2nd step: correction

$$\hat{x}(t+1|t+1) = \hat{x}(t+1|t) +$$

$$+ \underbrace{P(t+1|t)\,C^T\left[C\,P(t+1|t)\,C^T + R\right]^{-1}}_{K(t+1)\text{ : the Kalman gain}}\left(y(t+1) - C\,\hat{x}(t+1|t)\right)$$

$$P(t+1|t+1) = P(t+1|t)\left[I - C^T K^T(t+1)\right] =$$

$$= P(t+1|t) - P(t+1|t)\,C^T\left[C\,P(t+1|t)\,C^T + R\right]^{-1} C\,P(t+1|t)$$

Result [information]

$$P^{-1}(t+1|t+1) = P^{-1}(t+1|t) + \underbrace{C^T R^{-1} C}$$

P^{-1} : information matrix

new information from $y(t+1)$

Initialization: how to choose the initial condition $\hat{x}(0|-1)$, $P(0|-1)$

Theoretically: $\hat{x}(0|-1) = m_0 = E[x(0)]$
$P(0|-1) = P_0 = Cov(x(0))$

From a practical point of view:

* choose $\hat{x}(0|-1)$ from the available information on $x(0)$
* pick $P(0|-1) = \lambda \cdot I$ where λ is big enough to model uncertainty associated to $\hat{x}(0|-1)$