Distributed Predictive Control and Estimation

–File 4–

João Miranda Lemos
jlml@inesc-id.pt

Instituto Superior Técnico

2022

7
State estimation

Need for state estimation

There are situations in which the state is not available for direct measurement:

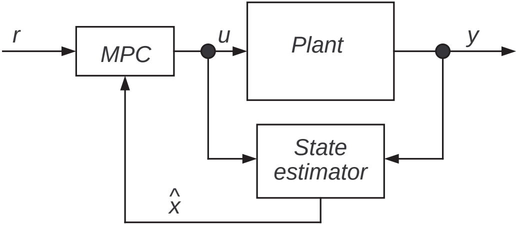
- Physical reasons (too harsh environments to place a sensor)
- The state has no physical meaning (only a mathematical meaning)
- The state is corrupted by noise and one wants to improve the measure

The state of a glass furnace consists of the temperature in a grid of points in a pool of molten glass.
The temperature of the glass is of the order of $1600^{\circ} C$, well above the molten point of the metal used as a case to thermocouple sensors (well below 1000°C).

One may only measure the temperature at the glass pool surface, at the top of the glass chamber or bellow it, and estimate the state from these measurements using models.

TÉCNICO LISBOA

MPC and state estimation



When the state is not available for direct estimation, it must be estimated to be used in the MPC algorithm.

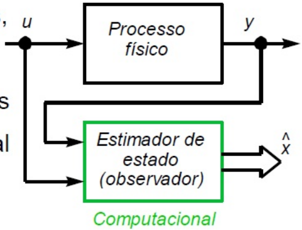
Linear state estimation: predictive observer

Estimates the state using the observation up to the previous time.

Asymptotic estimate of the state based on observations of the input and output:

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K_o\left(y(k) - C\hat{x}(k|k-1)\right)$$

This observer is a replica of the plant dynamics, driven by the same input, and corrected by a term that is the difference between what is expected to be the output at time k and the actual observation of the output.



Estimation error

$$\tilde{x} = x - \hat{x}$$

$$x(k+1) = \Phi x(k) + \Gamma u(k)$$

$$\hat{x}(k+1 \mid k) = \Phi \hat{x}(k \mid k-1) + \Gamma u(k) + K_o\big(y(k) - C\hat{x}(k \mid k-1)\big)$$

Subtracting these equations yields the error equation

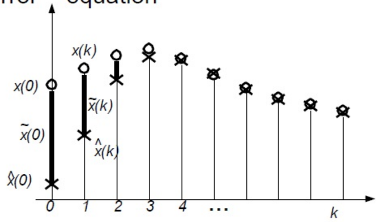
$$\tilde{x}(k+1) = (\Phi - K_o C)\tilde{x}(k)$$

Solution of the error equation

$$\tilde{x}(k) = (\Phi - K_o C)^k \tilde{x}(0)$$

The error will tend to zero iff

$$|\lambda(\Phi - K_o C)| < 1$$



Polynomial observer: $A_o(z) = \det(zI - \Phi + K_o C)$

Example: state estimation in a double integrator

Consider the double integrator

$$x(k+1) = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} h^2/2 \\ h \end{bmatrix} u(k)$$

$$y(k) = [1 \quad 0] x(k)$$

Design a predictive observer such that the roots of the polynomial observer are placed at $0.4 \pm j0.4$.

Example (cont.)

Desired characteristic polynomial

$$\alpha_o(z) = z^2 - 0.8z + 0.32$$

$$\det(zI - \Phi + K_o C) = \det\left(z\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - \begin{bmatrix}1 & h\\ 0 & 1\end{bmatrix} + \begin{bmatrix}K_{o1}\\ K_{o2}\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix}\right)$$

$$= z^2 + (K_{o1} - 2)z + hK_{o2} + 1 - K_{o1}$$

Comparing the coefficients

$$K_{o1} - 2 = -0.8$$

$$hK_{o2} + 1 - K_{o1} = 0.32$$

Solution:

$$K_{o1} = 1.2 \qquad K_{o2} = \frac{0.52}{h}$$

Linear state estimation: current observer

Estimates the state using the most recent observation.

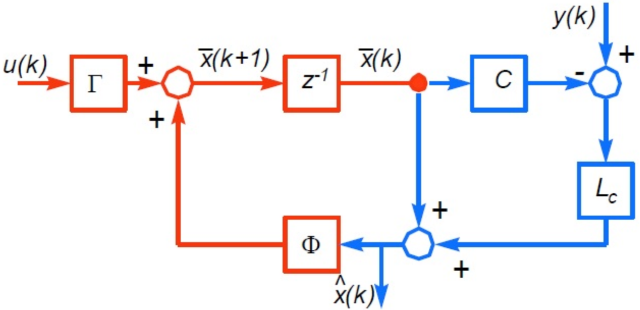
$$\hat{x}(k|k) = \Phi\hat{x}(k-1|k-1) + \Gamma u(k-1)$$
$$+ K_o[y(k) - C(\Phi\hat{x}(k-1|k-1) + \Gamma u(k-1))]$$

The estimate of the state at time k depends on the observations of the output up to time k (current time).

Prediction/Correction Equations

Prediction: $\bar{x}(k+1) = \Phi\hat{x}(k) + \Gamma u(k)$

Correction: $\hat{x}(k) = \bar{x}(k) + L_c(y(k) - C\bar{x}(k))$



The observer as a filter

Up to now, nothing has been said about sensor noise or stochastic disturbances.

The observer is a filter that has a frequency response.

If the bandwidth is to small, part of the state dynamics will be cancelled (undesirable).

If the bandwidth is too large, the state dynamics has already passed through the filter, and we are just aiding high frequency noise.

Problem: optimize the observer bandwidth to get the best compromise between getting the state dynamics and cancelling high-frequency noise.

This is done by the Kalman-Bucy filter (also known as the Kalman filter).

Optimizing the observer: the Kalman-Bucy filter

Process model with sensor noise and stochastic disturbances

$$x(k+1) = \Phi x(k) + \Gamma u(k) + v_1(k),$$

$$y(k) = Hx(k) + v_2(k).$$

Noise characterization

$$E\left[v_1(k) v_1^T(k)\right] = Q_E,$$

$$E\left[v_2^2(k)\right] = R_E.$$

Unbiased, optimal estimator

Use an unbiased estimator ⇒ Use the observer structure

Select the gains by minimizing the steady state power (variance) of the estimation error

$$J_o = E\left[\sum_{k=0}^{\infty} \|x(k) - \hat{x}(k)\|^2\right]$$

Use the MATLAB function *dlqe* to compute the optimal observer gain (Kalman gain).

Select Q_E and R_E as tuning knobs (and not as true noise variances).

The Kalman filter is a current observer with the gain (Kalman gain) optimized.

Optimal nonlinear filtering - model

Dado o modelo

$$x(t+1) = f(x(t)) + e(t)$$

$$y(t) = h(x(t)) + v(t)$$

em que $\{e(t)\}$ $\{v(t)\}$ são sequências de v.a. independentes e identicamente

distribuídas, pretende-se propagar recursivamente a f. d. p.

$$p(x(t)|Y^t)$$

em que

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

σ-algebra

$$Y^t = \sigma\{y(t),\ y(t-1),\ \ldots,\ y(0)\}$$

Significa a σ –álgebra gerada pelas observações até à amostra t.

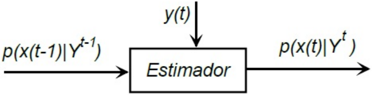
Conhecida a f. d. p. do estado dadas as observações, pode obter-se uma estimativa do estado por

$$\hat{x}(t) = \int_{-\infty}^{\infty} x(t)p(x(t)|Y^t)dx(t)$$

Optimal nonlinear filtering: solution structure

No instante t admitimos disponível uma densidade $\boldsymbol{p}(\boldsymbol{x}(\boldsymbol{t}-\mathbf{1})|\boldsymbol{Y}^{\boldsymbol{t}-\mathbf{1}})$. É feita

uma observação $\boldsymbol{y}(\boldsymbol{t})$, a partir da qual se calcula $p(x(t)|Y^t)$:



Explora-se o facto de $\{x(t)\}$ ser um processo de Markov.

Optimal nonlinear filtering: Markov processes

The state of a deterministic system is a vector of variables such that, if we know the control inputs and the plant model, allow us to compute its future evolution, starting from an initial condition.

Markov processes represent the notion of state for stochastic systems.

Sejam t_i e t inteiros tal que $t_1 < t_2 < \cdots < t_n < t$. Um processo estocástico $\{x(t)\}$ diz-se um processo de Markov se

$$P\{x(t) \leq \xi | x(t_1), \ \ldots, \ x(t_n)\} = P\{x(t) \leq \xi | x(t_n)\}$$

em que $P\{\cdot \, | x(t_1), \ \ldots, \ x(t_n)\}$ representa a probabilidade condicional dados $\{x(t_1), \ \ldots, \ x(t_n)\}$.

Some probability formulas

A and B random variables.

$$p(A, B) = p(A|B)p(B)$$

$$p(A) = \int p(A, B)dB$$

Optimal nonlinear filtering: steps of the estimate

A estimação é feita em dois passos:

- Predição
- Filtragem

Na Predição calcula-se $p(x(t)|Y^{t-1})$ a partir de $p(x(t-1)|Y^{t-1})$. Esta operação apenas envolve a equação dinâmica do estado x.

Na filtragem calcula-se $p(x(t)|Y^{t})$ a partir de $p(x(t)|Y^{t-1})$ e da última observação $y(t)$. Esta operação apenas envolve a equação de saída (modelo das observações).

TÉCNICO LISBOA

Optimal nonlinear filtering: prediction

Admita-se $p(x(t-1)|Y^{t-1})$ conhecida. Tem-se

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t), x(t-1)|Y^{t-1})dx(t-1)$$

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t)|x(t-1), Y^{t-1})p(x(t-1)|Y^{t-1})dx(t-1)$$

Esta equação é um caso particular da equação de Chapman-Kolmogorov.

Optimal nonlinear filtering: prediction (cont.)

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t)|x(t-1), Y^{t-1}) p(x(t-1)|Y^{t-1}) dx(t-1)$$

Sendo $\{x(t)\}$ um processo de Markov:

$$p(x(t)|x(t-1), Y^{t-1}) = p(x(t)|x(t-1))$$

pelo que

$$\boxed{p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t)|x(t-1)) p(x(t-1)|Y^{t-1}) dx(t-1)}$$

que define a *predição*.

TÉCNICO LISBOA

Optimal nonlinear filtering: filtering

Pela lei de Bayes

$$p(x(t)|Y^t) = \frac{p(x(t), Y^t)}{p(Y^t)} = \frac{p(y(t)|x(t), Y^{t-1})p(x(t), Y^{t-1})}{p(Y^t)}$$

$$p(x(t)|Y^t) = p(y(t)|x(t), Y^{t-1})p(x(t)|Y^{t-1})\frac{p(Y^{t-1})}{p(Y^t)}$$

Optimal nonlinear filtering: filtering (cont.)

Sendo $\{x(t)\}$ um processo de Markov:

$$p(y(t)|x(t), Y^{t-1}) = p(y(t)|x(t))$$

Por outro lado, pela lei de Bayes:

$$\frac{p(Y^{t-1})}{p(Y^t)} = \frac{1}{p(y(t)|Y^{t-1})} = K(t)$$

que é uma constante de normalização independente de $x(t)$.

TÉCNICO LISBOA

Optimal nonlinear filtering: filtering (cont.)

Let's now put all together:

$$p(x(t)|Y^t) = p(y(t)|x(t), Y^{t-1})p(x(t)|Y^{t-1})\frac{p(Y^{t-1})}{p(Y^t)}$$

$$p(y(t)|x(t), Y^{t-1}) = p(y(t)|x(t)) \qquad \frac{p(Y^{t-1})}{p(Y^t)} = \frac{1}{p(y(t)|Y^{t-1})} = K(t)$$

A *filtragem* é assim calculada por

$$\boxed{p(x(t)|Y^t) = K(t)p(y(t)|x(t))p(x(t)|Y^{t-1})}$$

Depende apenas do modelo das observações.

Optimal nonlinear filtering: filter equations

A f. d. p. do estado dadas as observações, $p(x(t)|Y^t)$ é propagada no tempo pelas equações:

Predição:

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t)|x(t-1))p(x(t-1)|Y^{t-1})dx(t-1)$$

Filtragem:

$$p(x(t)|Y^t) = K(t)p(y(t)|x(t))p(x(t)|Y^{t-1})$$

The Kalman filter revisited

Objective: derive the equations of the Kalman filter as a special case of the optimal nonlinear filter.

These procedure will give us a new of the Kalman filter, based on probability.

Product of 2 gaussian pdf

Mostre que o produto de duas gaussianas

$$p_i(\alpha) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{(\alpha-\mu_i)^2}{2\sigma_i^2}\right) \qquad i = 1,\ 2$$

é uma gaussiana não normalizada, dada por

$$p_1(\alpha)p_2(\alpha) = K\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(\alpha-\mu)^2}{2\sigma^2}\right)$$

de média μ e variância σ^2 dadas por $\frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$ $\qquad \mu = \sigma^2\left(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right)$

Sugestão: Complete o quadrado no expoente do produto.

Product of 2 gaussian pdf (cont.)

Solução

$$p_1(\alpha)p_2(\alpha) = \frac{1}{2\pi\sigma_1\sigma_2} exp\left(-\frac{(\alpha-\mu_1)^2}{2\sigma_1^2} - \frac{(\alpha-\mu_2)^2}{2\sigma_2^2}\right)$$

Desenvolva-se o simétrico do dobro do expoente

$$\frac{1}{\sigma_1^2}(\alpha-\mu_1)^2 + \frac{1}{\sigma_2^2}(\alpha-\mu_2)^2 =$$

$$= \frac{1}{\sigma_1^2}(\alpha^2 - 2\alpha\mu_1 + \mu_1^2) + \frac{1}{\sigma_2^2}(\alpha^2 - 2\alpha\mu_2 + \mu_2^2) =$$

$$= \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\alpha^2 - 2\alpha\left(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right) + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}$$

Product of 2 gaussian pdf (cont.)

$$= \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\alpha^2 - 2\alpha\left(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right) + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} =$$

$$\frac{1}{\sigma^2} \qquad\qquad \frac{\mu}{\sigma^2}$$

$$= \left[\frac{\alpha^2}{\sigma^2} - 2\frac{\alpha\mu}{\sigma^2} + \frac{\mu^2}{\sigma^2}\right] - \frac{\mu^2}{\sigma^2} + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} =$$

$$= \frac{1}{\sigma^2}(\alpha - \mu)^2 - \frac{\mu^2}{\sigma^2} + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}$$

Product of 2 gaussian pdf (cont.)

Compare-se com

$$p_1(\alpha)p_2(\alpha) = K\frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{(\alpha-\mu)^2}{2\sigma^2}\right)$$

A constante de normalização é

$$K = \frac{\sigma}{\sigma_1\sigma_2\sqrt{2\pi}}\cdot exp\left\{\frac{1}{2}\left(\frac{\mu^2}{\sigma^2}-\frac{\mu_1^2}{\sigma_1^2}-\frac{\mu_2^2}{\sigma_2^2}\right)\right\}$$

$\square$

Convolution of 2 gaussian pdf

A convolução de duas Gaussianas

$$p_i(\alpha) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{(\alpha-\mu_i)^2}{2\sigma_i^2}\right) \qquad i = 1,\ 2$$

é ainda uma gaussiana. Determine a sua média e a sua variância.

Convolution of 2 gaussian pdf (cont.)

Sugestão:

A função característica de uma v. a. é definida como a transformada de Fourier da sua f. d. p.:

$$\Phi_i(j\omega) = \int_{-\infty}^{\infty} p_i(\alpha) e^{j\omega\alpha} d\alpha$$

No caso da Gaussiana, a função característica é

$$\Phi_i(j\omega) = exp\left(j\omega\mu_i - \frac{1}{2}\omega^2\sigma_i^2\right)$$

Demonstra-se que a função característica da convolução de duas densidades é o produto das respetivas funções características.

Obtenha a função característica da convolução das Gaussianas.

Convolution of 2 gaussian pdf (cont.)

Solução

$$\Phi_{1*2}(j\omega) = \Phi_1(j\omega) \cdot \Phi_2(j\omega) =$$

$$= exp\left(j\omega\mu_1 - \frac{1}{2}\omega^2\sigma_1^2\right) \cdot exp\left(j\omega\mu_2 - \frac{1}{2}\omega^2\sigma_2^2\right) =$$

$$= exp\left(j\omega(\mu_1 + \mu_2) - \frac{1}{2}\omega^2(\sigma_1^2 + \sigma_2^2)\right)$$

Conclusão:

$$\mu_{1*2} = \mu_1 + \mu_2 \qquad \sigma_{1*2}^2 = \sigma_1^2 + \sigma_2^2$$

□

Conclusion: product of 2 gaussian pdf

Produto de duas gaussianas

O produto de duas gaussianas

$$p_i(\alpha) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{(\alpha-\mu_i)^2}{2\sigma_i^2}\right) \qquad i = 1,\ 2$$

é uma gaussiana não normalizada dada por

$$p_1(\alpha)p_2(\alpha) = K\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(\alpha-\mu)^2}{2\sigma^2}\right)$$

de média μ e variância σ^2 dadas por

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \qquad \mu = \sigma^2\left(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right)$$

TÉCNICO LISBOA

Conclusion: convolution of 2 gaussian pdf

A convolução de duas Gaussianas

$$p_i(\alpha) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{(\alpha-\mu_i)^2}{2\sigma_i^2}\right) \qquad i = 1,\ 2$$

é ainda uma gaussiana, com média e variância dadas por:

$$\mu_{1*2} = \mu_1 + \mu_2 \qquad \sigma_{1*2}^2 = \sigma_1^2 + \sigma_2^2$$

□

TÉCNICO LISBOA

Coordinate transform in a pdf (scalar)

Seja z uma v. a. com f. d. p. $p_z(\alpha)$.

Considere-se a transformação de variável

$$x = T(z) \qquad T \text{ monótona}$$

Então a f. d. p. da v. a. transformada x é

$$p_x(\beta) = p_z(T^{-1}(\beta)) \cdot \left|\frac{d}{d\beta} T^{-1}(\beta)\right|$$

Linear coordinate transform in a Gaussian pdf (scalar)

$$p_z(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(\alpha-\mu)^2}{2\sigma^2}\right)$$

$$x = a\ z$$

$$p_x(\beta) = \frac{1}{\sqrt{2\pi\sigma^2 a^2}} exp\left(-\frac{(\alpha-\mu a)^2}{2(\sigma a)^2}\right)$$

Conclusão

$$z \approx N(\mu,\ \sigma^2) \quad \rightarrow \quad x \approx N(\mu a,\ (\sigma a)^2)$$

Example: linear scalar model

$$x(t) = ax(t-1) + e(t)$$

$$y(t) = hx(t) + w(t)$$

$e, \; w$ sequências brancas e independentes de variâncias q e r

Example: linear scalar prediction

A) Predição

Da equação de estado

$$x(t) = ax(t-1) + e(t)$$

conclui-se que a densidade de transição é

$$p(x(t)|x(t-1)) = \frac{1}{\sqrt{2\pi q}} exp\left\{-\frac{1}{2} \cdot \frac{(x(t) - ax(t-1))^2}{q}\right\}$$

Example: linear scalar prediction (cont.)

$$p(x(t)|x(t-1)) = \frac{1}{\sqrt{2\pi q}} exp\left\{-\frac{1}{2}\cdot\frac{(x(t)-ax(t-1))^2}{q}\right\}$$

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} p(x(t)|x(t-1))p(x(t-1)|Y^{t-1})dx(t-1)$$

A equação de predição toma a forma

$$p(x(t)|Y^{t-1}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi q}} exp\left\{-\frac{(x(t)-ax(t-1))^2}{2q}\right\} p(x(t-1)|Y^{t-1})dx(t-1)$$

A densidade predita tem a forma da convolução de duas gaussianas, pelo que é uma Gaussiana cuja média é a soma das médias e cuja variância é a soma das variâncias.

Example: linear scalar prediction (cont.)

Seja

$$p(x(t)|Y^t) = N(\mu_t^F, \sigma_t^F) \qquad \text{“f. d. p. filtrada”}$$

$$p(x(t)|Y^{t-1}) = N(\mu_t^P, \sigma_t^P) \qquad \text{“f. d. p. predita”}$$

O resultado anterior permite concluir as expressões para a predição

$$\mu_t^P = a\mu_{t-1}^F$$

$$\sigma_t^P = a^2\sigma_{t-1}^F + q$$

TÉCNICO LISBOA

Example: linear scalar filtering

Tendo em conta o modelo das observações:

$$y(t) = hx(t) + w(t)$$

é

$$p(y(t)|x(t)) = \frac{1}{\sqrt{2\pi r}} \cdot exp\left\{-\frac{(y(t) - hx(t))^2}{2r}\right\}$$

Example: linear scalar filtering

$$p(y(t)|x(t)) = \frac{1}{\sqrt{2\pi r}} \cdot exp\left\{-\frac{(y(t)-hx(t))^2}{2r}\right\}$$

A densidade filtrada

$$p(x(t)|Y^t) = K(t)p(y(t)|x(t))p(x(t)|Y^{t-1})$$

é o produto de duas Gaussianas

$$\mu_t^F = \mu_{t-1}^P + K(t)[y(t) - h\mu_{t-1}^P] \qquad \sigma_t^F = [1 - K(t)h]^2\sigma_{t-1}^P + K^2(t)r$$

em que $K(t) = \frac{\sigma_t^P h}{r+h^2\sigma_t^P}$ é o “ganho de Kalman.

Example: linear scalar filter

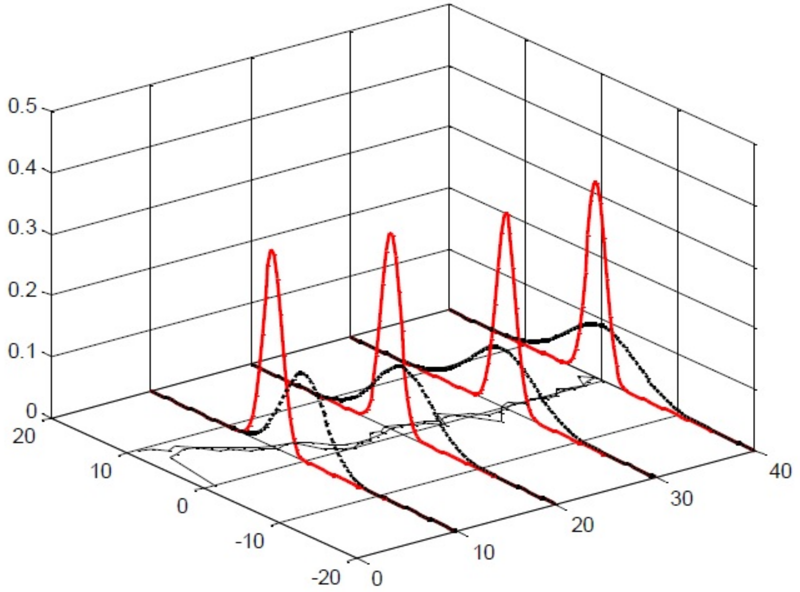
Conjugando os resultados de predição e filtragem, obtém-se um resultado que tem uma interpretação simples:

$\hat{x}(t) := \mu_t^F$ "estimativa do estado"

$$\hat{x}(t) = a\hat{x}(t-1) + K(t)[y(t) - ha\hat{x}(t-1)]$$

A estimativa é gerada por um sistema dinâmico (filtro de Kalman) que corresponde a uma réplica do sistema original "guiada" pela estimativa das inovações.

Example: numerical results



General Kalman filter: model

State dynamics:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$$

Observation model:

$$z_k = Hx_k + v_k$$

Process noise:

$$p(w)\ \mathcal{N}(0, Q), \quad Q \succeq 0$$

Sensor noise:

$$p(v)\ \mathcal{N}(0, R), \quad R \succ 0$$

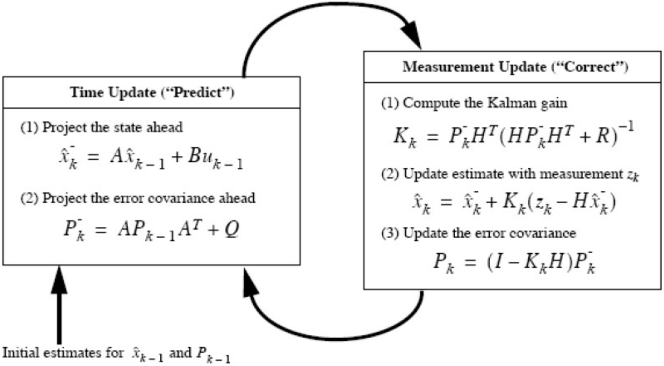
TÉCNICO LISBOA

General Kalman filter

The Kalman filter propagates in time the pdf of the state given the observations.

In the linear case, this pdf is a gaussian. Hence, only the mean and the covariance need to be propagated in time.

Kalman Filter: Prediction/correction



Extended Kalman Filter (EKF)

Nonlinear dynamics and/or observations:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$$

$$z_k = h(x_k, v_k)$$

Make a linearization about the last state estimate:

$$x_k \approx \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + W w_{k-1},$$

$$z_k \approx \tilde{z}_k + H(x_k - \tilde{x}_k) + V v_k.$$

$$\tilde{x}_k = f(\hat{x}_{k-1}, u_{k-1}, 0) \qquad \tilde{z}_k = h(\tilde{x}_k, 0)$$

Then, apply the equations of the Kalman Filter.

Extended Kalman Filter: jacobian approximations

- A is the Jacobian matrix of partial derivatives of f with respect to x, that is

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0),$$

- W is the Jacobian matrix of partial derivatives of f with respect to w,

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0),$$

- H is the Jacobian matrix of partial derivatives of h with respect to x,

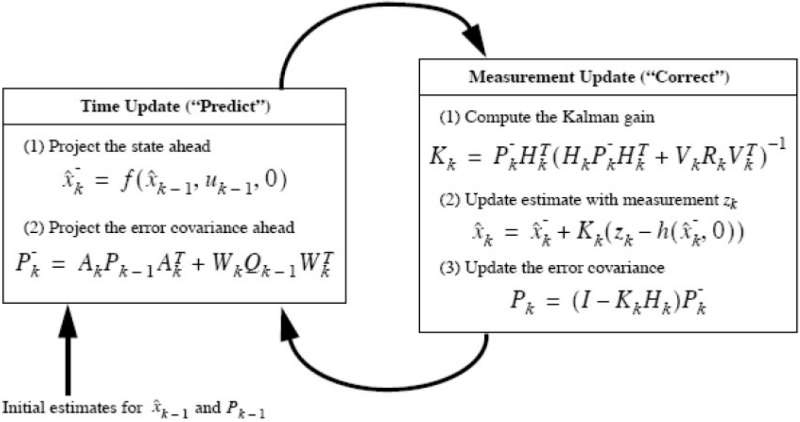
$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_k, 0),$$

- V is the Jacobian matrix of partial derivatives of h with respect to v,

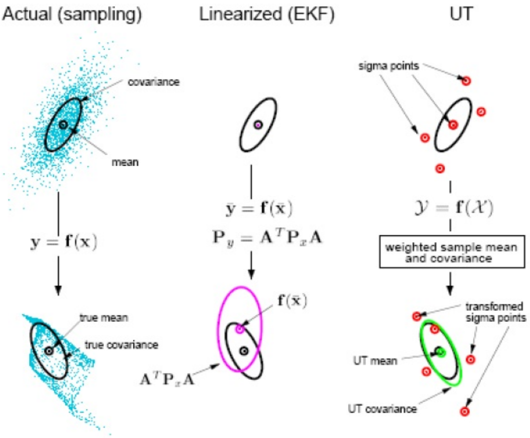
$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{x}_k, 0).$$

TÉCNICO LISBOA

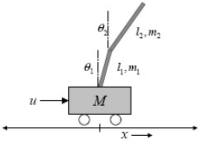
Extended Kalman Filter: prediction/correction



EKF and Unscented Kalman Filter: gaussian approximation



Comparison example: car with a double inverted pendulum



$$(M + m_1 + m_2)\ddot{x} - (m_1 + 2m_2)l_1\theta_1 \cos\theta_1 - m_2 l_2 \theta_2 \cos\theta_2$$
$$= u + (m_1 + 2m_2)l_1\dot{\theta_1}^2 \sin\theta_1 + m_2 l_2 \dot{\theta_2}^2 \sin\theta_2$$

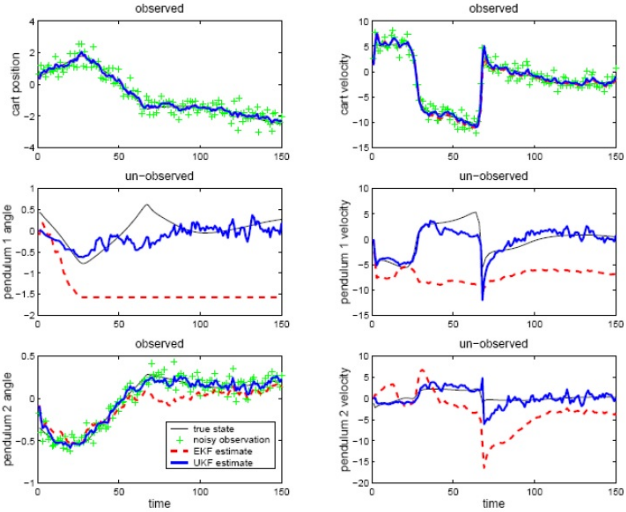
$$-(m_1 + 2m_2)l_1\ddot{x}\cos\theta_1 + 4(\frac{m_1}{3} + m_2)l_1{}^2\ddot{\theta_1} + 2m_2 l_1 l_2 \ddot{\theta_2}\cos(\theta_2 - \theta_1)$$
$$= (m1 + 2m_2)g l_1 \sin\theta_1 + 2m_2 l_1 l_2 \dot{\theta_2}^2 \sin(\theta_2 - \theta_1)$$

$$-m_2\ddot{x}l_2\cos\theta_2 + 2m_2 l_1 l_2 \ddot{\theta_1}\cos(\theta_2 - \theta_1) + \frac{4}{3}m_2 l_2{}^2\ddot{\theta_2}$$
$$= m_2 g l_2 \sin\theta_2 - 2m_2 l_1 l_2 \dot{\theta_1}^2 \sin(\theta_2 - \theta_1)$$

$$\mathbf{x} = [x, \dot{x}, \theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2]$$

TÉCNICO LISBOA

Comparison example: car with a double inverted pendulum



Joint state and parameter estimation

State model depends on a vector of parameters w_k:

$$x_{k+1} = F(x_k,\, u_k,\, w_k) + Bv_k$$

Model for the enlarged state (original state + parameters):

$$\begin{bmatrix} x_{k+1} \\ w_{k+1} \end{bmatrix} = \begin{bmatrix} F(x_k,\, u_k,\, w_k) \\ w_k \end{bmatrix} + \begin{bmatrix} Bv_k \\ r_k \end{bmatrix}$$

$$y_k = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} x_k \\ w_k \end{bmatrix} + n_k$$

If not all the state components are directly measured, the output equation is changed.

Joint state and parameter estimation (cont.)