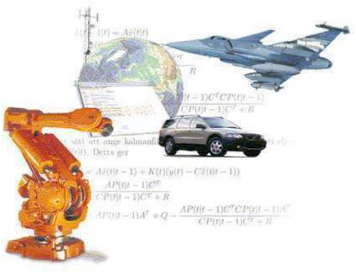
ML Estimation of Process Noise Variance in Dynamic Systems



Patrik Axelsson, Umut Orguner, Fredrik Gustafsson and Mikael Norrlöf

{axelsson,umut,fredrik,mino}@isy.liu.se

Division of Automatic Control
Department of Electrical Engineering
Linköping University, Sweden

Outline

1. Problem Formulation
2. Estimation of process noise covariance
3. Alternative Methods
4. Simulation Results

Outline

1. Problem Formulation
2. Estimation of process noise covariance
3. Alternative Methods
4. Simulation Results

Background

- The elasticity (in joints, arms, . . .) for an industrial robot has increased over the last years.
- The motion control must be improved for these new robots.

ABB IRB4600 (`www.abb.se`)

Background

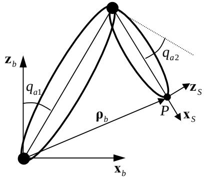
- The elasticity (in joints, arms, . . .) for an industrial robot has increased over the last years.
- The motion control must be improved for these new robots.
- One solution can be to estimate the tool position and use that estimate in the feedback loop.
- The tool position can be estimated with e.g. a EKF which uses knowledge about the process noise.

ABB IRB4600 (www.abb.se)

- The elasticity (in joints, arms, ...) for an industrial robot has increased over the last years.
- The motion control must be improved for these new robots.
- One solution can be to estimate the tool position and use that estimate in the feedback loop.
- The tool position can be estimated with e.g. a EKF which uses knowledge about the process noise.
- The accuracy is very important, < 1 mm.

ABB IRB4600 (www.abb.se)

- Serial robot with two degrees of freedom.
- Nonlinear stiffness.
- Nonlinear friction.

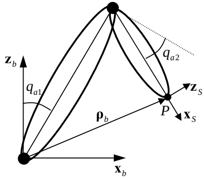


$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) + T(q) + D\dot{q} + F(\dot{q}) = u$$

- The model structure is common for mechanical systems derived by Newton's second law of motion or Lagrange's equation.

Problem Formulation

- Serial robot with two degrees of freedom.
- Nonlinear stiffness.
- Nonlinear friction.



$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) + T(q) + D\dot{q} + F(\dot{q}) = u$$

- The model structure is common for mechanical systems derived by Newton's second law of motion or Lagrange's equation.
- Accelerometer

$$\ddot{\rho}_s(q_a,\dot{q}_a,\ddot{q}_a) = R_s^b(q_a)\Big(\ddot{\rho}_b(q_a,\dot{q}_a,\ddot{q}_a) + G_b\Big) + \delta_s$$

- States

$$x = \begin{pmatrix} q_a^T & q_m^T & \dot{q}_a^T & \dot{q}_m^T \end{pmatrix}^T \in \mathbb{R}^8$$

- Discrete state space model (cont. $\rightarrow$ disc. using Euler forward)

$$x_{k+1} = F_1(x_k, u_k) + F_2(x_k)v_k, \quad v_k \sim \mathcal{N}(0, Q)$$
$$y_k = \begin{pmatrix} q_{m,k} \\ \ddot{\rho}_{s,k} \end{pmatrix} + e_k = h(x_k, u_k) + e_k, \quad e_k \sim \mathcal{N}(0, R)$$

where

$$F_2(x_k) = \begin{pmatrix} 0 \\ \widetilde{F}_2(x_k) \end{pmatrix}.$$

- All model parameters known except for the covariance matrix Q.

- Maximum Likelihood (ML) method

$$\hat{\theta}^{ML} = \arg\max_{\theta \in \Theta} \log p_\theta(y_{1:N}).$$

- Maximum Likelihood (ML) method

$$\hat{\theta}^{ML} = \arg\max_{\theta \in \Theta} \log p_\theta(y_{1:N}).$$

- Solution hard to find ⇒ Expectation Maximisation (EM) algorithm.

- Maximum Likelihood (ML) method

$$\hat{\theta}^{ML} = \arg\max_{\theta \in \Theta} \log p_\theta(y_{1:N}).$$

- Solution hard to find $\Rightarrow$ Expectation Maximisation (EM) algorithm.
  1. Select an initial value θ_0 and set $l = 0$.
  2. Expectation Step (E-step): Calculate

$$\Gamma(\theta; \theta_l) = E_{\theta_l}\left[\log p_\theta(y_{1:N}, x_{1:N}) | y_{1:N}\right].$$

  3. Maximisation Step (M-step): Compute

$$\theta_{l+1} = \arg\max_{\theta \in \Theta} \Gamma(\theta; \theta_l).$$

  4. If converged, stop. If not, set $l = l + 1$ and go to step 2.

Outline

- Estimate Q with the Expectation Maximisation (EM) algorithm.

- Estimate Q with the Expectation Maximisation (EM) algorithm.
- Main ideas
  - Calculate $\Gamma(Q; Q_l)$ using linearisation.
  - Use the smoothed states (EKS) as approximation when necessary. Possible since the smoothed densities are peaky when the SNR is high which is the case here.

- Conditional densities

$$x_{k+1} \sim p(x_{k+1}|x_k) = \mathcal{N}\left(x_{k+1}; F_{1,k}, F_{2,k} Q F_{2,k}^T\right)$$
$$y_k \sim p(y_k|x_k) = \mathcal{N}\left(y_k; h_k, R\right)$$

- Conditional densities

$$
\begin{aligned}
x_{k+1} &\sim p(x_{k+1}|x_k) = \mathcal{N}\left(x_{k+1}; F_{1,k}, F_{2,k} Q F_{2,k}^T\right) \\
y_k &\sim p(y_k|x_k) = \mathcal{N}\left(y_k; h_k, R\right)
\end{aligned}
$$

- The joint log likelihood function

$$
\begin{aligned}
L_Q(y_{1:N}, x_{1:N}) = \log p_Q(y_{1:N}, x_{1:N}) = \widetilde{L} \\
- \frac{1}{2}\sum_{i=2}^{N} (x_i - F_{1,i-1})^T \left(F_{2,i-1} Q F_{2,i-1}^T\right)^{\dagger} (x_i - F_{1,i-1}) \\
+ \frac{1}{2}\sum_{i=2}^{N} \log\left(\prod_{\lambda_j \neq 0} \lambda_j\left(\left(F_{2,i-1} Q F_{2,i-1}^T\right)^{\dagger}\right)\right)
\end{aligned}
$$

Estimation of process noise covariance

- Conditional densities

$$x_{k+1} \sim p(x_{k+1}|x_k) = \mathcal{N}\left(x_{k+1}; F_{1,k}, F_{2,k} Q F_{2,k}^T\right)$$
$$y_k \sim p(y_k|x_k) = \mathcal{N}\left(y_k; h_k, R\right)$$

- The joint log likelihood function

$$\begin{aligned} L_Q(y_{1:N}, x_{1:N}) &= \log p_Q(y_{1:N}, x_{1:N}) = \widetilde{L} \\ &- \frac{1}{2} \sum_{i=2}^{N} (x_i - F_{1,i-1})^T \left(F_{2,i-1} Q F_{2,i-1}^T\right)^{\dagger} (x_i - F_{1,i-1}) \\ &+ \frac{1}{2} \sum_{i=2}^{N} \log\left(\left|\left(\widetilde{F}_{2,i-1} Q \widetilde{F}_{2,i-1}^T\right)^{\dagger}\right|\right) \end{aligned}$$

- Expectation of the joint log likelihood function

$$
\begin{aligned}
\Gamma(Q;Q_l) &= E_{Q_l}\left[L_Q(y_{1:N}, x_{1:N}) | y_{1:N}\right] = \bar{L} \\
&+ \frac{1}{2}\sum_{i=2}^{N} E_{Q_l}\left[\log\left(\left|\left(\widetilde{F}_{2,i-1} Q \widetilde{F}_{2,i-1}^T\right)^{\dagger}\right|\right) \middle| y_{1:N}\right] \\
&- \frac{1}{2}\operatorname{tr}\sum_{i=2}^{N} E_{Q_l}\left[\left(F_{2,i-1} Q F_{2,i-1}^T\right)^{\dagger} \left(x_i - F_{1,i-1}\right)\left(x_i - F_{1,i-1}\right)^T \middle| y_{1:N}\right]
\end{aligned}
$$

- Calculation of the first expectation

$$E_{Q_l}\left[\log\left(\left|\left(\widetilde{F}_{2,i-1}Q\widetilde{F}_{2,i-1}^T\right)^\dagger\right|\right)\middle| y_{1:N}\right]$$
$$= \int \log\left(\left|\left(\widetilde{F}_2(x_{i-1})Q\widetilde{F}_2^T(x_{i-1})\right)^\dagger\right|\right)p_{Q_l}(x_{i-1}|y_{1:N})\,\mathrm{d}x_{i-1}$$

Estimation of process noise covariance
cont'd

- Calculation of the first expectation

$$E_{Q_l}\left[\log\left(\left|\left(\widetilde{F}_{2,i-1}Q\widetilde{F}^T_{2,i-1}\right)^{\dagger}\right|\right)\middle| y_{1:N}\right]$$
$$= \int \log\left(\left|\left(\widetilde{F}_2(x_{i-1})Q\widetilde{F}^T_2(x_{i-1})\right)^{\dagger}\right|\right)p_{Q_l}(x_{i-1}|y_{1:N})\,\mathrm{d}x_{i-1}$$

- Cannot be solved $\Rightarrow$ Approximation: Use the smoothed states.

$$E_{Q_l}\left[\log\left(\left|\left(\widetilde{F}_{2,i-1}Q\widetilde{F}^T_{2,i-1}\right)^{\dagger}\right|\right)\middle| y_{1:N}\right]$$
$$\approx \log\left(\left|\left(\widetilde{F}_2(\hat{x}^s_{i-1|N})Q\widetilde{F}^T_2(\hat{x}^s_{i-1|N})\right)^{\dagger}\right|\right)$$

- Calculation of the second expectation

$$E_{Q_l}\left[\left(F_{2,i-1}QF_{2,i-1}^T\right)^{\dagger}(x_i - F_{1,i-1})(x_i - F_{1,i-1})^T \middle| y_{1:N}\right]$$

$$= \int \left(F_2(x_{i-1})QF_2^T(x_{i-1})\right)^{\dagger}(x_i - F_{1,i-1})(x_i - F_{1,i-1})^T$$

$$\times\, p_{Q_l}(x_i, x_{i-1}|y_{1:N})\,\mathrm{d}x_i\,\mathrm{d}x_{i-1}$$

- Calculation of the second expectation

$$E_{Q_l}\left[\left(F_{2,i-1}QF_{2,i-1}^T\right)^{\dagger}(x_i - F_{1,i-1})(x_i - F_{1,i-1})^T \middle| y_{1:N}\right]$$
$$= \int \left(F_2(x_{i-1})QF_2^T(x_{i-1})\right)^{\dagger}(x_i - F_{1,i-1})(x_i - F_{1,i-1})^T$$
$$\times\, p_{Q_l}(x_i, x_{i-1}|y_{1:N})\, \mathrm{d}x_i\, \mathrm{d}x_{i-1}$$

- Use the smoothed states again

$$E_{Q_l}\left[\left(F_{2,i-1}QF_{2,i-1}^T\right)^{\dagger}(x_i - F_{1,i-1})(x_i - F_{1,i-1})^T \middle| y_{1:N}\right]$$
$$\approx \left(F_2(\hat{x}^s_{i-1|N})QF_2^T(\hat{x}^s_{i-1|N})\right)^{\dagger}$$
$$\times \int (x_i - F_{1,i-1})(x_i - F_{1,i-1})^T p_{Q_l}(x_i, x_{i-1}|y_{1:N})\, \mathrm{d}x_i\, \mathrm{d}x_{i-1}$$

- A first order Taylor approximation gives

$$
\begin{aligned}
M &\triangleq \int (x_i - F_{1,i-1}) (x_i - F_{1,i-1})^T p_{Q_l}(x_i, x_{i-1} | y_{1:N}) \, \mathrm{d}x_i \, \mathrm{d}x_{i-1} \\
&= \begin{pmatrix} -J_1 & I \end{pmatrix} P^{\xi,s}_{i|N} \begin{pmatrix} -J_1 & I \end{pmatrix}^T \\
&\quad + \left(\hat{x}^s_{i|N} - F_1(\hat{x}^s_{i-1|N})\right) \left(\hat{x}^s_{i|N} - F_1(\hat{x}^s_{i-1|N})\right)^T \\
J_1 &= \left. \frac{\partial F_1(x)}{\partial x} \right|_{x=\hat{x}^s_{i-1|N}}
\end{aligned}
$$

- $P^{\xi,s}_{i|N}$ is the smoothed state covariance for the augmented state vector $\xi = \begin{pmatrix} x^T_{i-1} & x^T_i \end{pmatrix}^T$

- Finally

$$\Gamma(Q;Q_l) = \bar{\bar{\bar{L}}} + \frac{N-1}{2}\log\left|Q^{-1}\right| - \frac{1}{2}\operatorname{tr} Q^{-1}W$$

where

$$W = \sum_{i=2}^{N} F_2^{\dagger}(\hat{x}^s_{i-1|N}) M \left(F_2^{\dagger}(\hat{x}^s_{i-1|N})\right)^T$$

- Finally

$$\Gamma(Q;Q_l) = \bar{\bar{L}} + \frac{N-1}{2}\log\left|Q^{-1}\right| - \frac{1}{2}\operatorname{tr} Q^{-1}W$$

where

$$W = \sum_{i=2}^{N} F_2^{\dagger}(\hat{x}^s_{i-1|N}) M \left(F_2^{\dagger}(\hat{x}^s_{i-1|N})\right)^T$$

- Maximisation of $\Gamma(Q;Q_l)$ w.r.t. Q gives

$$Q_{l+1} = \frac{1}{N-1}\sum_{i=2}^{N} F_2^{\dagger}(\hat{x}^s_{i-1|N}) M \left(F_2^{\dagger}(\hat{x}^s_{i-1|N})\right)^T$$

Outline

1. Problem Formulation
2. Estimation of process noise covariance
3. Alternative Methods
4. Simulation Results

Alt. 1 Minimisation of the estimation error, given by e.g. EKF, w.r.t. Q, where Q is parametrised as a diagonal matrix.

Alt. 2 Calculate v_k from the state space model.

Alt. 1 Minimisation of the estimation error, given by e.g. EKF, w.r.t. Q, where Q is parametrised as a diagonal matrix.

Alt. 2 Calculate v_k from the state space model.

1. Select an initial value Q_0 and set $l = 0$.
2. Use the EKS with Q_l.
3. Calculate the noise according to

$$v_k = F_2^{\dagger}\left(\hat{x}^s_{k|N}\right)\left(\hat{x}^s_{k+1|N} - F_1\left(\hat{x}^s_{k|N}, u_k\right)\right).$$

4. Let Q_{l+1} be the covariance matrix for v_k according to

$$Q_{l+1} = \frac{1}{N}\sum_{k=1}^{N} v_k^T v_k.$$

5. If converged, stop, If not, set $l = l + 1$ and go to step 2.

Outline

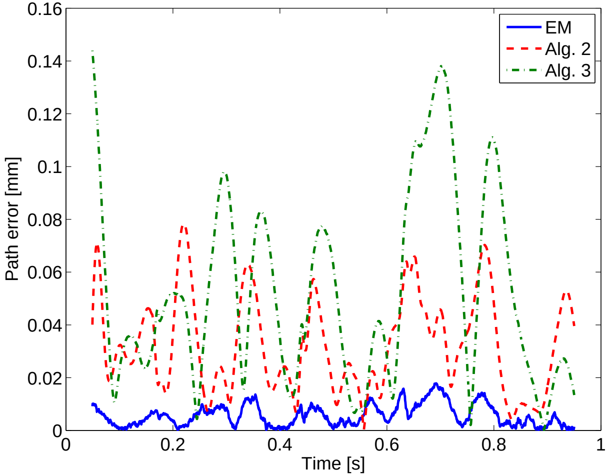
1. Problem Formulation
2. Estimation of process noise covariance
3. Alternative Methods
4. Simulation Results

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

- Monte Carlo simulations on one simulated path with different initial values.
- No true values for $Q \Rightarrow$ Use Q in an EKF and calculate the estimation error (RMSE) for the path.

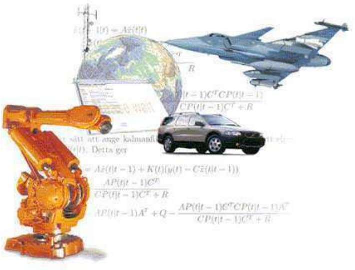
- Monte Carlo simulations on one simulated path with different initial values.
- No true values for $Q \Rightarrow$ Use Q in an EKF and calculate the estimation error (RMSE) for the path.
- The EM algorithm converges to the same RMSE for all initial values. Same thing happens for Alt. 2. Alt. 1 ends up in different errors.
- The EM algorithm gives the lowest RMSE (2-norm of the RMSE).

	Max	Min
EM	0.2999	0.2996
Alt. 1	3.3769	1.5867
Alt. 2	2.6814	2.6814



Path error [mm]
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
Time [s]
0
0.2
0.4
0.6
0.8
1
EM
Alg. 2
Alg. 3

ML Estimation of Process Noise Variance in Dynamic Systems



Patrik Axelsson, Umut Orguner, Fredrik Gustafsson and Mikael Norrlöf
{axelsson,umut,fredrik,mino}
@isy.liu.se

Division of Automatic Control
Department of Electrical Engineering
Linköping University, Sweden