ML Estimation of Process Noise Variance in Dynamic Systems

Patrik Axelsson, Umut Orguner, Fredrik Gustafsson and Mikael Norrlöf

Contribution

This work presents a method to estimate the *process noise variance* for a *non-linear dynamic system* with *high state dimension*. The proposed method makes use of the *expectation maximization* algorithm, where the E-step is solved by *linearisation*.

Introduction

The performance of a non-linear filter hinges in the end on the accuracy of the assumed non-linear model of the process. In particular, the process noise covariance Q. For non-linear models, there is on-going research on using the expectation maximization (EM) algorithm with a particle smoother to estimate the parameters. However, the particle smoother is not applicable for models with high state dimension. The idea here is to:

- ➢ Linearise the non-linear model.
- ➢ Use an extended Kalman smoother (EKS).

Let the model be given by

$$x_{k+1} = F_1(x_k, u_k) + F_2(x_k)v_k \quad (1a)$$
$$y_k = h(x_k, u_k) + e_k \quad (1b)$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$, $v_k \sim \mathcal{N}(0, Q)$ and $e_k \sim \mathcal{N}(0, R)$. All model parameters are assumed to be known except for $Q \in S_+^p$. Assume also that $F_2(x_k)$ has the following structure

$$F_2(x_k) = \begin{pmatrix} 0 \\ \tilde{F}_2(x_k) \end{pmatrix}. \quad (2)$$

This type of model structure is common for mechanical systems derived by Newton's law or Lagrange's equation.

The EM Algorithm (Alg. 1)

1. Select an initial value Q_0 and set $l = 0$.
2. Expectation Step (E-step): Calculate
$$\Gamma(Q; Q_l) = E_{Q_l}\left[\log p_Q(y_{1:N}, x_{1:N}) | y_{1:N}\right].$$
3. Maximisation Step (M-step): Compute $Q_{l+1} = \arg\max_{Q \in \mathcal{Q}} \Gamma(Q; Q_l)$.
4. If converged[a], stop. If not, set $l = l + 1$ and go to step 2.

[a] Here an estimate of the log-likelihood function is used. The algorithm stops if no increase in the log-likelihood function can be observed.

The E-step

The expectation of the log-likelihood function $\log p_Q(y_{1:N}, x_{1:N})$ is calculated using the EKS and it can be expressed as

$$\Gamma(Q; Q_l) = \bar{L} - \frac{1}{2}\mathrm{Tr}Q^{-1}\sum_{i=2}^{N} F_2^\dagger(\hat{x}_{i-1|N}^s)M\left(F_2^\dagger(\hat{x}_{i-1|N}^s)\right)^T + \frac{1}{2}\sum_{i=2}^{N}\left[\log\left|Q^{-1}\right| + \log\left|\tilde{F}_2^\dagger(\hat{x}_{i-1|N}^s)\right| + \log\left|\left(\tilde{F}_2^\dagger(\hat{x}_{i-1|N}^s)\right)^T\right|\right],$$

where $\bar{L}$ is a function independent of Q,

$$M = \left(-J_1\ I\right)P_{i|N}^{\xi,s}\left(-J_1\ I\right)^T + \left(\hat{x}_{i|N}^s - F_1(\hat{x}_{i-1|N}^s)\right)\left(\hat{x}_{i|N}^s - F_1(\hat{x}_{i-1|N}^s)\right)^T$$

and J_1 is the Jacobian of $F_1(x, u)$ evaluated at $x = \hat{x}_{i-1|N}^s$. The variables $\hat{x}_{i-1|N}^s$, $\hat{x}_{i|N}^s$ and $P_{i|N}^{\xi,s}$ are obtained if the augmented state vector $\xi_i = \left(x_{i-1}^T\ x_i^T\right)^T$ with the new model $\xi_{k+1} = \begin{pmatrix} x_k \\ F_1(x_k, u_k)\end{pmatrix}$ is used in the EKS. That is, the EKS calculates

$$\hat{\xi}_{i|N}^s = \begin{pmatrix} \hat{x}_{i-1|N}^s \\ \hat{x}_{i|N}^s \end{pmatrix} \quad \text{and} \quad P_{i|N}^{\xi,s} = \begin{pmatrix} P_{i-1|N}^s & P_{i-1,i|N}^s \\ \left(P_{i-1,i|N}^s\right)^T & P_{i|N}^s \end{pmatrix}$$

where $\hat{x}_{i-1|N}^s$, $\hat{x}_{i|N}^s$, $P_{i-1|N}^s$ and $P_{i|N}^s$ are the first and second order moments of the smoothed $\hat{x}_{i-1}$ and $\hat{x}_i$ respectively.

The M-step

Take the derivative of $\Gamma(Q; Q_l)$ with respect to Q^{-1} and let the result be equal to zero to get the solution in the maximisation step according to

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

Two Alternative Algorithms

Two alternative methods are compared to the EM algorithm.

Alg. 2: Minimisation of the path error

1. Select diagonal $Q_0 \in \mathbb{R}^{4\times 4}$.
2. Minimise $\sqrt{\sum_{k=1}^{N} |\mathsf{e}_k|^2}$ subject to $\lambda_j > 0$, $j = 1, \ldots, 4$, $Q = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)Q_0$ and $(\hat{\mathsf{x}}, \hat{\mathsf{y}}) = \mathrm{EKF}(Q)$.
3. $Q = \mathrm{diag}(\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*)Q_0$, where λ_j^* is the optimal value from step 2.

Alg. 3: Iterative covariance estimation with EKS

1. Select $Q_0 \in \mathbb{R}^{4\times 4}$ and set $l = 0$.
2. Use the EKS with Q_l.
3. Calculate the noise v_k from (1a).
4. Let Q_{l+1} be the covariance matrix for v_k.
5. If converged, stop, if not, set $l = l + 1$ and go to step 2.

Application to Industrial Robots

Consider the non-linear joint flexible two axes robot model

$$\dot{x} = \begin{pmatrix} x_3 \\ x_4 \\ M_a^{-1}(x_1)\left(-C(x_1, x_3) - G(x_1) - A(x) + v_a\right) \\ M_m^{-1}\left(A(x) + \kappa(x_4) + u + v_m\right) \end{pmatrix},$$

where $x = \left(x_1^T\ x_2^T\ x_3^T\ x_4^T\right)^T = \left(q_a^T\ q_m^T\ \dot{q}_a^T\ \dot{q}_m^T\right)^T$ and $A(x) = D(x_3 - x_4) + \tau_s(x_1, x_2)$. The model structure (1a) and (2) is obtained when an Euler forward approximation is used to discretise the robot model.

Notation

$M_a(x_1)$	Inertia matrix for the arms
M_m	Inertia matrix for the motors
$C(x_1, x_3)$	Coriolis- and centrifugal terms
$G(x_1)$	Gravitaion torque
$\tau_s(x_1, x_2)$	Nonlinear stiffness torque
$D(x_3 - x_4)$	Damping torque
$\kappa(x_4)$	Nonlinear friction torque
$v = \left(v_a\ \ v_m\right)^T$	Process noise

I. The model was simulated to get the control signal u_k and the measurements, i.e., the motor angles q_m and the acceleration of the tool.

II. The true tool position, used in Alg. 2 was also calculated.

III. The three algorithms were applied to the data to get Q.

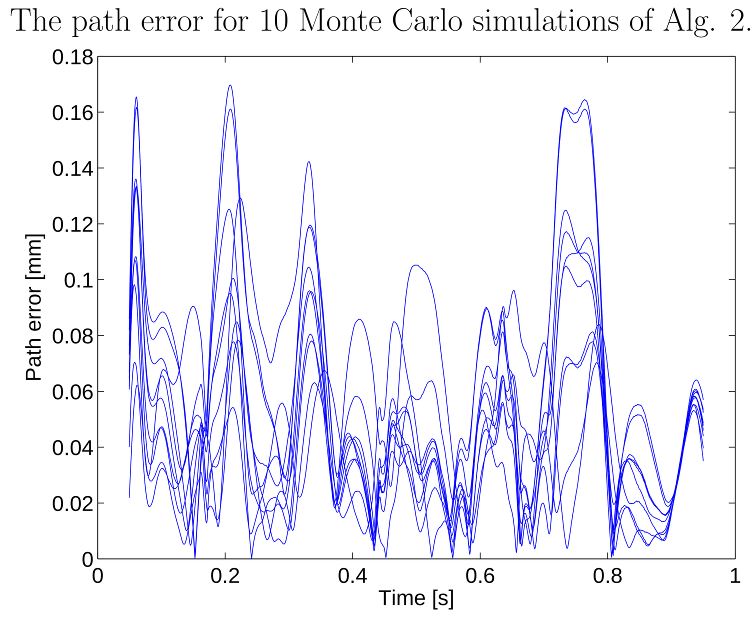
IV. The three Q-matrices were used in an extended Kalman filter to obtain an estimate of the tool position.

V. The path errors $\mathsf{e}_k = \sqrt{|\mathsf{x}_k - \hat{\mathsf{x}}_k|^2 + |\mathsf{y}_k - \hat{\mathsf{y}}_k|^2}$ for these estimates were used to compare the three algorithms.

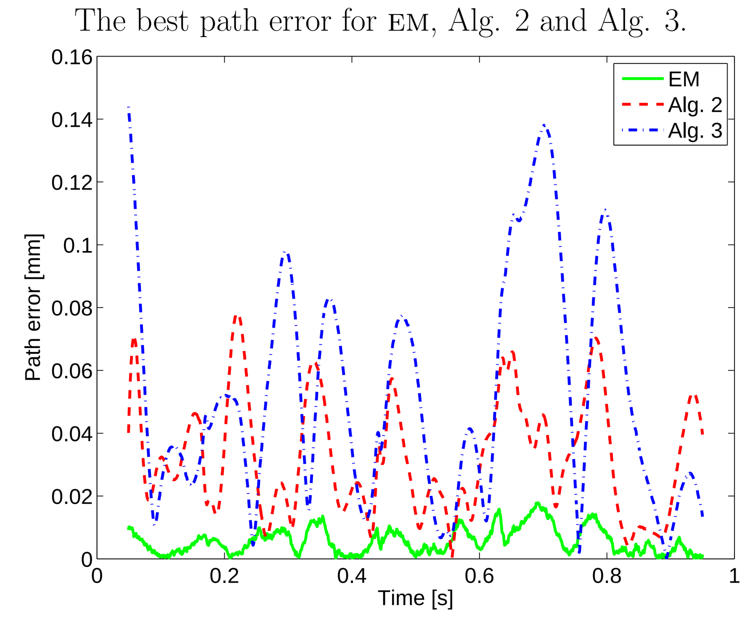
Result

- Alg. 2 gives different solutions for different initial values.
- EM and Alg. 3 give consistent solutions for different initial values.
- The path error for EM is much lower than Alg. 2 and Alg. 3.
- EM converges in around 50 iterations.

The path error for 10 Monte Carlo simulations of Alg. 2.



The best path error for EM, Alg. 2 and Alg. 3.



Max and min of the 2-norm of the path error for different initial values.

	EM	Alg. 2	Alg. 3
Max	0.2999	3.3769	2.6814
Min	0.2996	1.5867	2.6814