Comparison of Prediction-Error-Modelling Criteria

John Bagterp Jørgensen[1] Sten Bay Jørgensen[2]

[1]Informatics and Mathematical Modelling
Technical University of Denmark
jbj@imm.dtu.dk

[2]CAPEC
Department of Chemical Engineering
Technical University of Denmark

American Control Conference 2007
New York, NY, USA, 11-13 July 2007

Outline

Model Predictive Control

The MPC regulator objective function

$$\phi_k = \frac{1}{2}\sum_{j=1}^{N_p}(\hat{y}_{k+j|k} - r_{k+j|k})'Q(\hat{y}_{k+j|k} - r_{k+j|k})$$
$$+ \frac{1}{2}\sum_{j=0}^{N_c}\Delta\hat{u}'_{k+j|k}S\Delta\hat{u}_{k+j|k}$$

requires a multi-step N_p-step-ahead prediction at each time point.
Idea:

1. Use a system identification criterion that is consistent with this objective.
2. Parameterize the predictor using a continuous-discrete stochastic model.

Standard Regression Problem

$$\mathbf{y}_k = \hat{y}_k(\theta) + \mathbf{e}_k,\ \mathbf{e}_k \sim N(0, R_k(\theta)),\ k = 0, 1, \ldots, N-1$$

$$\hat{\theta} = \arg\min_{\theta \in \Theta} V(\theta)$$

$$e_k(\theta) = y_k - \hat{y}_k(\theta)$$

$$V_{LS}(\theta) = \frac{1}{2} \sum_{k=0}^{N-1} \| e_k(\theta) \|_2^2 \tag{1}$$

$$\begin{aligned} V_{ML}(\theta) = \frac{N n_y}{2} \ln(2\pi) + \frac{1}{2} \sum_{k=0}^{N-1} \ln\left(\det R_k(\theta)\right) \\ + \frac{1}{2} \sum_{k=0}^{N-1} e_k(\theta)' R_k(\theta)^{-1} e_k(\theta) \end{aligned} \tag{2}$$

$$\begin{aligned} V_{MAP}(\theta) = V_{ML}(\theta) + \frac{n_\theta}{2} \ln(2\pi) + \frac{1}{2} \ln\left(\det P_{\theta_0}\right) \\ + \frac{1}{2} (\theta - \theta_0)' P_{\theta_0}^{-1} (\theta - \theta_0) \end{aligned} \tag{3}$$

Model Parameterization

Continuous-Discrete Stochastic Transfer Function Model

$$\mathbf{Z}(s) = G(s;\theta)U(s) + H(s;\theta)\mathbf{E}(s)$$
$$\mathbf{y}(t_k) = \mathbf{z}(t_k) + \mathbf{v}(t_k)$$

$U(s)$ is the process input vector

$\mathbf{E}(s)$ is a vector with white noise components

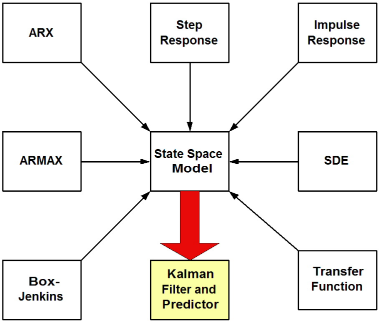
$\mathbf{Z}(s)$ is the process output vector

$\mathbf{v}(t_k) \sim N(0, R_{vv}(\theta))$ is the measurement noise vector

$\mathbf{y}(t_k)$ is the measured process output vector at time t_k.

$$g_{ij}(s) = \frac{b_{ij}(s;\theta)}{a_{ij}(s;\theta)} \exp(-\tau_{ij}(\theta)s)$$
$$h_{ij}(s) = \frac{d_{ij}(s;\theta)}{c_{ij}(s;\theta)} \exp(-\lambda_{ij}(\theta)s)$$

State Space Model Realization



1. Realize the chosen model parameterization as a stochastic discrete-time state space model.
2. The optimal filter and predictor for the state space model is the Kalman filter and predictor.

$$\mathbf{x}_{k+1} = A(\theta)\mathbf{x}_k + B(\theta)u_k + \mathbf{w}_k$$
$$\mathbf{y}_k = C(\theta)\mathbf{x}_k + \mathbf{v}_k$$

with

$$\begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \sim N_{iid}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_{ww}(\theta) & R_{wv}(\theta) \\ R_{wv}(\theta)' & R_{vv}(\theta) \end{bmatrix} \right)$$
$$\mathbf{x}_0 \sim N(\hat{x}_{0|-1}(\theta), P_{0|-1}(\theta))$$

Kalman Filter

Dynamic Kalman Filter

Innovation and gains

$$\begin{aligned}
\hat{y}_{k|k-1} &= C\hat{x}_{k|k-1} \\
e_k &= y_k - \hat{y}_{k|k-1} \\
R_{e,k} &= CP_{k|k-1}C' + R_{vv} \\
K_{fx,k} &= P_{k|k-1}C'R_{e,k}^{-1} \\
K_{fw,k} &= R_{wv}R_{e,k}^{-1}
\end{aligned}$$

Filtered estimates

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_{fx,k}e_k \\
\hat{w}_{k|k} &= K_{fw,k}e_k
\end{aligned}$$

and covariances

$$\begin{aligned}
P_{k|k} &= P_{k|k-1} - K_{fx,k}R_{e,k}K'_{fx,k} \\
Q_{k|k} &= R_{ww} - K_{fw,k}R_{e,k}K'_{fw,k}
\end{aligned}$$

Static Kalman Filter

Riccati equation ($P = \lim_{k\to\infty} P_{k|k-1}$)

$$\begin{aligned}
P = {} & APA' + R_{ww} \\
& - (APC' + R_{wv})(R_{vv} + CPC')^{-1}(APC' + R_{wv})'
\end{aligned}$$

Gains

$$\begin{aligned}
R_e &= CPC' + R_{vv} \\
K_{fx} &= PC'R_e^{-1} \\
K_{fw} &= R_{wv}R_e^{-1}
\end{aligned}$$

Innovation

$$\begin{aligned}
\hat{y}_{k|k-1} &= C\hat{x}_{k|k-1} \\
e_k &= y_k - \hat{y}_{k|k-1}
\end{aligned}$$

Filtered estimates

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_{fx}e_k \\
\hat{w}_{k|k} &= K_{fw}e_k
\end{aligned}$$

Kalman Predictors

One-Step Predictor

Estimated predictions

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + B\hat{u}_{k|k} + \hat{w}_{k|k}$$
$$\hat{y}_{k+1|k} = C\hat{x}_{k+1|k}$$

and covariances

$$\begin{aligned} P_{k+1|k} &= AP_{k|k}A' + Q_{k|k} \\ &\quad - AK_{fx,k}R'_{wv} - R_{wv}K'_{fx,k}A' \\ R_{k+1|k} &= CP_{k+1|k}C' + R_{vv} \end{aligned}$$

j-Step Predictor

Estimated predictions for $j \geq 2$

$$\hat{x}_{k+j|k} = A\hat{x}_{k+j-1|k} + B\hat{u}_{k+j-1|k}$$
$$\hat{y}_{k+j|k} = C\hat{x}_{k+j|k}$$

and covariances

$$P_{k+j|k} = AP_{k+j-1|k}A' + R_{ww}$$
$$R_{k+j|k} = CP_{k+j|k}C' + R_{vv}$$

One-Step Prediction Error Estimation

The innovations (prediction errors) are normally distributed

$$\mathbf{e}_k(\theta) \sim N(0, R_{e,k}(\theta))$$

and computed from the Kalman filter recursions

$$e_k(\theta) = y_k - \hat{y}_{k|k-1}(\theta)$$
$$R_{e,k}(\theta) = R_{vv}(\theta) + C(\theta)P_{k|k-1}(\theta)C(\theta)'$$

Estimation problem

$$\hat{\theta} = \arg\min_{\theta \in \Theta} V(\theta)$$

Estimation criteria

$$\text{LS}: \quad V(\theta) = \frac{1}{2}\sum_{k=1}^{N} \|e_k(\theta)\|_2^2$$

$$\text{ML}: \quad V(\theta) = \frac{1}{2}\sum_{k=1}^{N} \left[\ln\left(\det R_{e,k}(\theta)\right) + e_k(\theta)' R_{e,k}^{-1}(\theta) e_k(\theta)\right]$$

Single j-Step Prediction-Error Estimation

The innovations (prediction errors) are normally distributed

$$\mathbf{e}_{k+j|k}(\theta) \sim N(0, R_{e,k+j|k}(\theta))$$

and computed from the Kalman filter recursions

$$e_{k+j|k}(\theta) = y_{k+j} - \hat{y}_{k+j|k}(\theta)$$
$$R_{e,k+j|k}(\theta) = R_{vv}(\theta) + C(\theta)P_{k+j|k}(\theta)C(\theta)'$$

Estimation problem

$$\hat{\theta} = \arg\min_{\theta \in \Theta} V(\theta)$$

Estimation criteria

$$\text{LS}: \quad V(\theta) = \frac{1}{2}\sum_{k=-1}^{N-1-j} \left\| e_{k+j|k}(\theta) \right\|_2^2$$

$$\text{ML}: \quad V(\theta) = \frac{1}{2}\sum_{k=-1}^{N-1-j} \left[\ln\left(\det R_{e,k+j|k}(\theta)\right) + e_{k+j|k}(\theta)' R_{e,k+j|k}^{-1}(\theta) e_{k+j|k}(\theta) \right]$$

Multi-Step Prediction Error Maximum-Likelihood Estimation

$$\hat{\theta} = \arg\min_{\theta \in \Theta} V_{ML}(\theta)$$

in which the likelihood function is

$$V_{ML}(\theta) = \frac{n_y f}{2} \ln(2\pi) + \frac{1}{2} \sum_{k=-1}^{N-2} \ln\left(\det R_k\right) + \epsilon_k R_k^{-1} \epsilon_k$$

$f = N_p[N - \frac{1}{2}(N_p - 1)]$, $\boldsymbol{\epsilon}_k = \mathbf{Y}_k - \hat{Y}_k(\theta)$, $R_k = \langle \boldsymbol{\epsilon}_k, \boldsymbol{\epsilon}_k \rangle$, and

$$\mathbf{Y}_k = \begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \vdots \\ \mathbf{y}_{k+N_p} \end{bmatrix} \quad \hat{Y}_k(\theta) = \begin{bmatrix} \hat{y}_{k+1|k}(\theta) \\ \hat{y}_{k+2|k}(\theta) \\ \vdots \\ \hat{y}_{k+N_p|k}(\theta) \end{bmatrix}$$

Data used for SYSID: $\{(y_k, u_k)\}_{k=0}^{N-1}$

System

$$\mathbf{Z}(s) = g(s)U(s) + h(s)\mathbf{E}(s)$$
$$\mathbf{y}(t_k) = \mathbf{z}(t_k) + \mathbf{v}(t_k)$$

$\mathbf{E}(s)$ standard white noise

$\mathbf{v}(t_k) \sim N_{iid}(0, r^2)$

Model

$$g(s) = \frac{1.0}{(1.0s+1)(3.0s+1)}\, e^{-5.2s}$$

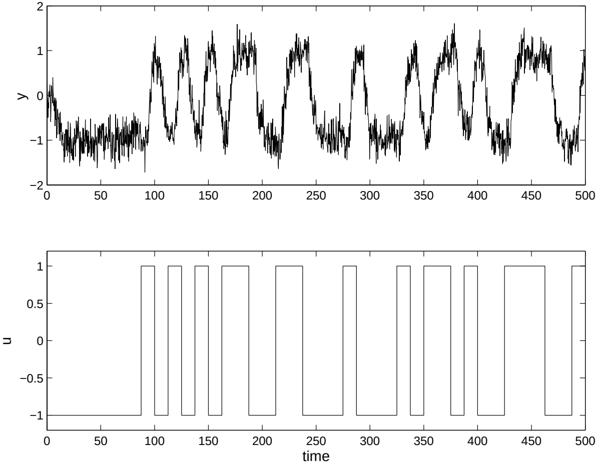
Process noise model

$$h(s) = \frac{0.2}{s+1}$$

Measurement noise

$$r = 0.2$$

IO-Data



The inputs, $\{u(t)\}$, are PRBS with bandwidth [0 0.02] and levels $[-1\ 1]$.

Sampling time: $T_s = 0.25$

System

$$\mathbf{Z}(s) = g(s)U(s) + h(s)\mathbf{E}(s)$$

$$\mathcal{S} = \{g(s), h(s)\}$$

$$g(s) = \frac{K}{(\alpha_1 s + 1)(\alpha_2 s + 1)} e^{-\tau s}$$

$$h(s) = \frac{\sigma}{\gamma s + 1}$$

$$\mathbf{y}(t_k) = \mathbf{z}(t_k) + \mathbf{v}(t_k) \qquad \mathbf{v}(t_k) \sim N(0, r^2)$$

K	α_1	α_2	τ	σ	γ	r	σ/r
1.0	1.0	3.0	5.2	0.2	1.0	0.2	1.0

Identical Model and System Structure

Let the model structure $\mathcal{M} = \left\{ \hat{g}(s), \hat{h}(s) \right\}$ be defined by

$$\hat{g}(s) = \frac{\hat{K}}{(\hat{\alpha}_1 s + 1)(\hat{\alpha}_2 s + 1)} e^{-\hat{\tau} s}$$

$$\hat{h}(s) = \frac{\hat{\sigma}}{\hat{\gamma} s + 1}$$

$\mathcal{S} \in \mathcal{M}$. Hence we expect unbiased estimates.

Identical Model and System Structure.

Single-Step LS Estimation

j	K	α_1	α_2	τ	σ	γ	r	σ/r	V	CPU sec.
1	0.9797	0.5641	3.4216	5.3171	0.8705	1.3757	0.8198	1.0618	109.9	113
4	0.9792	0.5798	3.4053	5.3119	0.4802	1.1548	0.3756	1.2785	116.9	154
8	0.9790	0.7239	3.3496	5.2037	0.8301	1.1885	0.5636	1.4730	120.0	227
20	0.9832	0.7086	3.3815	5.2019	1.9202	9.7771	2.0543	0.9347	119.2	351
40	0.9786	0.8639	3.2871	5.1016	0.1824	0.9056	0.1776	1.0268	122.5	325
80	0.9719	0.7612	3.3374	5.1578	0.1800	0.9000	0.1800	1.0000	129.5	394
100	1.0087	0.9820	3.3471	4.8656	0.1800	0.9000	0.1800	1.0000	200.3	376
200	0.9428	1.1445	2.8801	5.0634	0.1800	0.9000	0.1800	1.0000	130.4	532
	1.0	1.0	3.0	5.2	0.2	1.0	0.2	1.0		

Multi-Step LS Estimation

N_p	K	α_1	α_2	τ	σ	γ	r	σ/r	V	CPU sec.
1	0.9797	0.5632	3.4219	5.3179	0.3377	1.3754	0.3180	1.0620	110.0	87
4	0.9796	0.5657	3.4194	5.3170	0.0080	1.3861	0.0075	1.0603	449.1	421
8	0.9794	0.6136	3.3981	5.2827	0.3805	1.3641	0.3595	1.0585	924.2	251
20	0.9792	0.7116	3.3563	5.2107	0.3053	1.4301	0.2897	1.0539	2370	393
40	0.9823	0.7394	3.3641	5.1836	0.7805	8.0826	0.8318	0.9383	4763	644
80	0.9804	0.7597	3.3357	5.1767	0.4734	7.1230	0.4975	0.9514	9481	1101
100	0.9796	0.7586	3.3305	5.1782	0.6825	6.5664	0.7271	0.9386	11804	1426
200	0.9760	0.7739	3.2966	5.1802	0.5728	6.1969	0.6151	0.9314	23023	2382
	1.0	1.0	3.0	5.2	0.2	1.0	0.2	1.0		

Identical Model and System Structure.

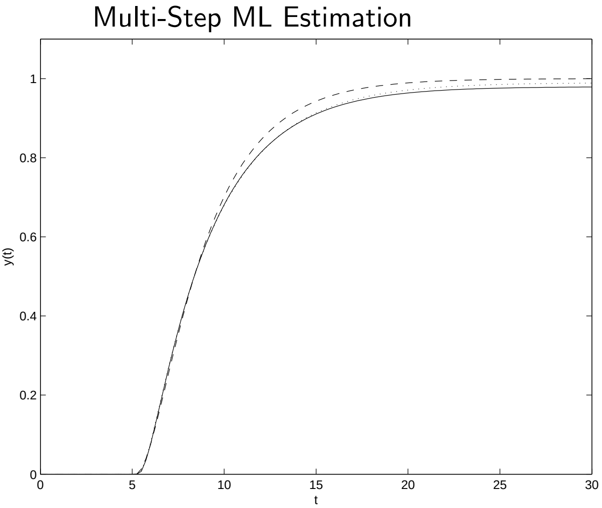
Single-Step ML estimation

j	K	α_1	α_2	τ	σ	γ	r	σ/r	V	CPU sec.
1	0.9797	0.5651	3.4211	5.3164	0.2204	1.3762	0.2077	1.0613	-63.27	115
4	0.9793	0.5798	3.4048	5.3120	0.2432	1.1321	0.1868	1.3022	-1.987	371
8	0.9789	0.7184	3.3500	5.2088	0.2853	1.0881	0.1526	1.8694	23.57	725
20	0.9832	0.7081	3.3829	5.2007	0.2243	9.0000	0.2391	0.9384	17.62	2038
40	0.9786	0.8639	3.2871	5.1060	0.0002	0.1380	0.2475	0.0007	45.37	3297
80	0.9719	0.7608	3.3376	5.1580	0.0002	0.1985	0.2545	0.0007	100.7	6082
100	1.0088	0.9817	3.3474	4.8656	0.0002	0.2748	0.3165	0.0006	536.9	6696
200	0.9426	1.1232	2.8931	5.0718	0.0002	0.0291	0.2553	0.0007	107.4	13227
	1.0	1.0	3.0	5.2	0.2	1.0	0.2	1.0		

Multi-Step ML Estimation

j	K	α_1	α_2	τ	σ	γ	r	σ/r	V	CPU sec.
1	0.9797	0.5651	3.4211	5.3164	0.2204	1.3762	0.2077	1.0613	-63.27	111
4	0.9798	0.5607	3.4251	5.3186	0.2182	1.4745	0.2102	1.0383	-248.7	160
8	0.9798	0.5307	3.4369	5.3425	0.2051	1.5453	0.2140	0.9581	-468.5	237
20	0.9801	0.5262	3.4700	5.3392	0.1860	1.5200	0.2189	0.8498	-829.0	490
40	0.9835	0.5302	3.5242	5.3124	0.1807	1.3614	0.2189	0.8254	-1204	878
80	0.9872	0.5294	3.5495	5.3043	0.1840	1.3603	0.2190	0.8403	-1977	1616
100	0.9879	0.5295	3.5535	5.3027	0.1843	1.3535	0.2189	0.8421	-2397	2508
200	0.9898	0.5298	3.5640	5.2987	0.1840	1.3310	0.2185	0.8420	-4715	6730
	1.0	1.0	3.0	5.2	0.2	1.0	0.2	1.0		

Identical Model and System Structure. Step Response



Solid: $N_p = 1$, Dotted: $N_p = 200$, Dashed: True model

Simplified Model with Output Integrator

Let the model structure $\mathcal{M} = \left\{\hat{g}(s), \hat{h}(s)\right\}$ be defined by

$$\hat{g}(s) = \frac{\hat{K}}{\hat{\alpha} s + 1} \, e^{-\hat{\tau} s}$$

$$\hat{h}(s) = \frac{\hat{\sigma}}{s}$$

Simplified Model with Output Integrator

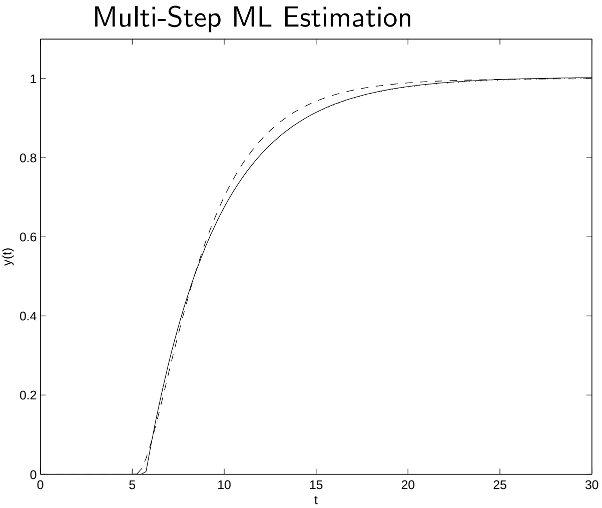
Single-Step ML Estimation.

j	K	α	τ	σ	r	σ/r	V	CPU sec.
1	1.0043	3.8386	5.7243	0.0658	0.2226	0.2959	-19.78	123
4	0.9911	3.6390	5.7547	0.0124	0.2424	0.0511	30.24	399
8	0.9811	3.5585	5.7792	0.0006	0.2490	0.0025	34.02	857
20	0.9812	3.5568	5.7802	0.0004	0.2455	0.0018	29.20	2245
40	0.9822	3.5750	5.7697	0.0002	0.2479	0.0009	48.26	4192
80	0.9747	3.5664	5.7618	0.0002	0.2547	0.0008	102.8	7448
100	1.0107	3.6458	5.6487	0.0002	0.3169	0.0006	539.5	8212
200	0.9465	3.3713	5.8331	0.0006	0.2556	0.0023	110.5	17885

Multi-Step ML Estimation.

j	K	α	τ	σ	r	σ/r	V	CPU sec.
1	1.0043	3.8386	5.7243	0.0658	0.2226	0.2959	-19.78	120
4	1.0043	3.8387	5.7244	0.0659	0.2424	0.2962	-79.72	160
8	1.0043	3.8386	5.7243	0.0658	0.2490	0.2956	-161.4	257
20	1.0043	3.8389	5.7244	0.0660	0.2455	0.2968	-398.4	382
40	1.0044	3.8394	5.7245	0.0666	0.2479	0.2995	-780.4	722
80	1.0039	3.8319	5.7256	0.0669	0.2547	0.3011	-1541	1550
100	1.0033	3.8277	5.7261	0.0670	0.3169	0.3018	-1954	2082
200	1.0024	3.8209	5.7269	0.0672	0.2556	0.3027	-4234	4268

Simplified Model with Output Integrator. Step Response



Solid: $N_p = 1$, Dotted: $N_p = 200$, Dashed: True model

Conclusions

1. Method for estimation of parameters in SISO (MIMO) continuous-discrete-time stochastic systems described by transfer functions with time delays.
2. Multi-Step Maximum Likelihood Prediction-Error-Method compatible with MPC objective function.
3. Computational feasibility of the method is demonstrated on two parameterizations of a SISO system.

Future work:

Comparison of closed loop MPC performance based on predictive models obtained by

1. One-step Prediction-Error-Method
2. Multi-Step Maximum Likelihood Prediction-Error-Method
3. ARX parameterizations
4. Subspace identification (ARX parameterization)