LUNDS UNIVERSITET

Multiple Model Minimax Estimation

Olle Kjellqvist

Outline

1. Multiple Model
   Adaptive Estimation
2. Bounded operator norm
3. Multiple Model Minimax Filters
4. Future Research
5. Summary

Multiple Model Adaptive Estimation

- Estimate properties of a system using a finite set of feasible models.
- Model complex engineering systems as sets of linear systems
- State estimation
- Fault Detection

Example: F-16 — Maybeck

- Design with sensor and actuator redundancy
- One model for each sensor/actuator fault
- One model for healthy operation
- MMAE to detect sensor fault and compensate estimate
- MMAE to detect actuator fault and reroute control
- Fault detection and compensation < 1 second

1763	Bayesian Estimation
1795	Least Square Estimation
1906	Markov Process
1908	Fisher Information Theory
1929	Lattice Filter
1930	Monte Carlo method
1945	Cramer-Rao Bound
1949	Wiener Kolmogorov Filter
1960	Kalman Filter
1965	**Multiple Models Filter**
1981	$\mathcal{H}_\infty$ filter
1984	Interacting Multiple Model
1993	Particle Filter
1997	Unscented Kalman Filter
2000	Gauss-Hermite Filter
2003	Gaussian Particle Filter
2007	Smooth Variable Structure Filter
2009	Cubature Kalman Filter

Gausian noise — Magill, Lainiotis

Given matrices $F_1, \ldots, F_K$, $H_1, \ldots, H_K$, covariance matrices $Q_1, \ldots Q_K$ and $R_1, \ldots R_K$. Assume the dynamics can be described by

$$x_{t+1} = F_i x_t + w_t$$
$$y_t = H_i x_t + v_t$$
$$w_t \sim \mathcal{N}(0, Q_i), \quad v_t \sim \mathcal{N}(0, R_i).$$

The probability of i being active is

$$p(i|y_t, \ldots, y_0) = \frac{\overbrace{p(y_t|i, y_{t-1}, \ldots y_0)}^{\text{computable}} \overbrace{p(i|y_{t-1} \ldots y_0)}^{\text{recursion}}}{\sum_{j=1}^{K} \left(p(y_t|j, y_{t-1}, \ldots y_0) p(j|y_{t-1} \ldots y_0)\right)}.$$

Gausian noise — Magill, Lainiotis

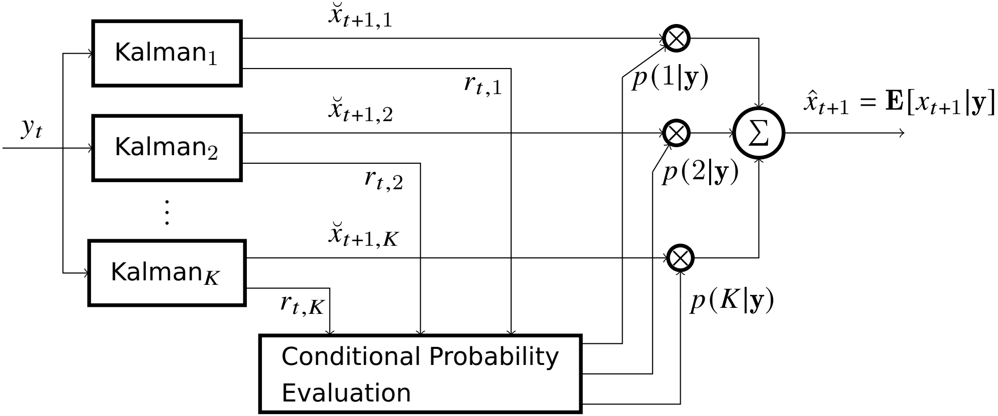
Given matrices $F_1, \ldots, F_K$, $H_1, \ldots, H_K$, covariance matrices $Q_1, \ldots Q_K$ and $R_1, \ldots R_K$. Assume the dynamics can be described by

$$
\begin{aligned}
x_{t+1} &= F_i x_t + w_t \\
y_t &= H_i x_t + v_t \\
w_t \sim \mathcal{N}(0, Q_i), &\quad v_t \sim \mathcal{N}(0, R_i).
\end{aligned}
$$

The probability of i being active is

$$
p(i|y_t, \ldots, y_0) = \frac{\overbrace{p(y_t|i, y_{t-1}, \ldots y_0)}^{\text{computable}} \overbrace{p(i|y_{t-1} \ldots y_0)}^{\text{recursion}}}{\sum_{j=1}^{K} \left(p(y_t|j, y_{t-1}, \ldots y_0) p(j|y_{t-1} \ldots y_0)\right)}.
$$

Block Diagram



Jump-linear systems — Bar-Shalom

The discrete variable can jump!

$$
\begin{aligned}
x_{t+1} &= F(\delta_t)x_t + w_t \\
y_t &= H(\delta_t)x_t + v_t \\
p(\delta_{t+1}) &= Ap(\delta_t) \leftarrow \text{Markov Chain} \\
w_t \sim \mathcal{N}(0, Q_i), &\quad v_t \sim \mathcal{N}(0, R_i)
\end{aligned}
$$

- Enumerating each feasible trajectory $\neq$ tractable
- Interacting Multiple Model algorithm most used
- Mode probabilities $\longrightarrow$ Filtering update $\longrightarrow$ mixing probabilities $\longrightarrow$ overall prediction

Weaknesses

- **Weakness:** Strong assumptions on gaussian white noise
- **Solution:** Use Interacting multiple models with particle filter
- **New weakness:** Complex and computationally heavy

1763	Bayesian Estimation
1795	Least Square Estimation
1906	Markov Process
1908	Fisher Information Theory
1929	Lattice Filter
1930	Monte Carlo method
1945	Cramer-Rao Bound
1949	Wiener Kolmogorov Filter
1960	Kalman Filter
1965	Multiple Models Filter
1981	$\mathcal{H}_\infty$ **filter**
1984	Interacting Multiple Model
1993	Particle Filter
1997	Unscented Kalman Filter
2000	Gauss-Hermite Filter
2003	Gaussian Particle Filter
2007	Smooth Variable Structure Filter
2009	Cubature Kalman Filter

Hardy space $\mathcal{H}^\infty$

$\mathcal{H}^\infty$ is a closed space with functions that are analytical and bounded inside the unit disc (open right half-plane). Normed space

$$\|G\|_\infty := \sup_{|z|<1} \bar{\sigma}[G(z)],$$

and

$$\|Gf\|_2 \leq \|G\|_\infty \|f\|_2.$$

Let $G : (w, v) \mapsto (\hat{x} - x)$ — map from disturbances to estimation error.

Idea: Find G with smallest $\|G\|_\infty$.

Godfrey Harold Hardy

1877 – 1944

- British mathematician
- number theory and analysis
- 1940 essay "A Mathematician’s Apology"

“It is never worth a first-class person’s time to express a majority opinion. By definition, there are plenty of others to do that.” — *paraphrased*

- Fredric Riesz
- 1880 – 1956
- Coined "Hardy Space"
- Teaching with style

Marcel Riesz

- Fredric Riesz
- 1880 – 1956
- Coined "Hardy Space"
- Teaching with style

- Marcel Riesz
- 1886 – 1969
- Math — Stockholm then Lund

- Fredric Riesz
- 1880 – 1956
- Coined "Hardy Space"
- Teaching with style

MARCEL RIESZ

- Marcel Riesz
- 1886 – 1969
- Math — Stockholm then Lund

- Harald Cramér
- 1893 - 1985
- Ph. D — M. Riesz
- Cramér-Rao lower bound...

$\mathcal{H}^\infty$-filter

- Find filter with smallest ℓ^2 (L^2) gain from disturbance to error.
- Introduced by Mike Grimble 1987 — Fundamentals by Zames 1981.

$$\min_{\hat{x}} \max_{w,v,x_0} \frac{\sum |\hat{x}_t - x_t|^2}{|\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}})}$$

- Different characterizations — different results

Linear-Quadratic Game — Shen, Deng

- Algorithm for solving

$$\min_{\hat{x}} \max_{w,v,x_0} \left\{ \sum |\hat{x}_t - x_t|^2 - \gamma^2 \left(\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum (|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}}) \right) \right\}$$

- Guarantees $\|G\|_\infty < \gamma$. Bisect over γ.
- Multiple Model Adaptive Estimation?

Minimax criterion

$$x_{t+1} = F_i x_t + w_t$$
$$y_t = H_i x_t + v_t$$
$$\min_{\hat{x}_N} \max_{w,v,x_0,i} \left\{ |\hat{x}_N - x_N|^2 - \gamma^2 \left(|\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}} \right) \right) \right\}$$

Forward dynamic programming

- $w_t = x_{t+1} - F_i x_t$
- $v_t = y_t - H_i x_t$
- $\mathbf{y}^t = \{y_{t-1}, \ldots, y_0\}$

$$\min_{w,v,x_0} |\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} (|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}})$$

$$= \min_{x_{N-1},\ldots,x_0} |\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} (|x_{t+1} - F_i x_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}})$$

$$=: V_N(x_N, \mathbf{y}^N)$$

Solution 1

- Solve using forward dynamic programming, i.e.

$$V_{t+1}(x_{t+1}, \mathbf{y}^{t+1}) = \min_{x_t} \left\{ |x_{t+1} - F_i x_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}} + V_t(x_t, \mathbf{y}^t) \right\}$$

$$V_t(x_t, \mathbf{y}^t) = |\breve{x}_{t,i} - x_t|^2_{P_{t,i}^{-1}} + \underbrace{\sum_{k=0}^{t-1} |y_t - H_i \breve{x}_{t,i}|^2_{(R+H_i P_{t,i} H_i^\top)^{-1}}}_{c_{t,i}}$$

- $\breve{x}_{t,i}$ is a Kalman Filter prediction for the ith model
- $P_{t,i}$ is the error covariance of a kalman filter for the ith model
- We have compressed the measurement history.

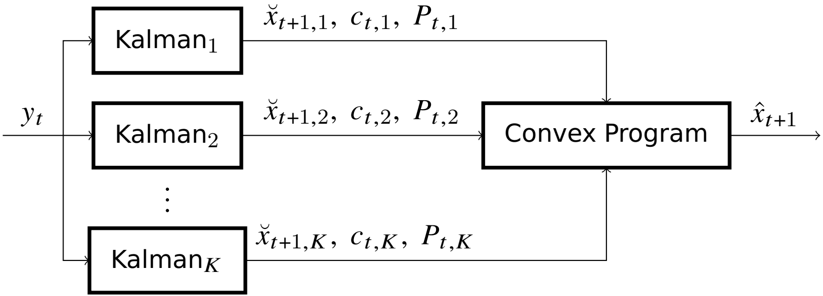
Solution 2

$$\min_{\hat{x}_N} \max_{w,v,x_0,i} \left\{ |\hat{x}_N - x_N|^2 - \gamma^2 \left(|\hat{x}_0 - x_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}} \right) \right) \right\}$$

$$= \min_{\hat{x}_N} \max_{x_N, i} \left\{ |\hat{x}_N - x_N|^2 - \gamma^2 |\breve{x}_{N,i} - x_N|^2_{P_{t,i}^{-1}} - c_{N,i} \right\}$$

$$= \min_{\hat{x}_N} \max_{i} |\hat{x}_N - \breve{x}_{N,i}|^2_{(I - \gamma^{-2} P_{N,i})^{-1}} - \gamma^2 c_{N,i}$$

- Convex program
- Solve using CVX, or JuMP with Mosek or Hypatia.jl


$\breve{x}_{t+1,1},\ c_{t,1},\ P_{t,1}$
Kalman$_1$
y_t
$\breve{x}_{t+1,2},\ c_{t,2},\ P_{t,2}$
Kalman$_2$
Convex Program
$\hat{x}_{t+1}$
$\breve{x}_{t+1,K},\ c_{t,K},\ P_{t,K}$
Kalman$_K$

Future Research

- Multiple Model Adaptive Estimation with a bound on ℓ^2 gain.
- Uncountable parameter space.
- Output feedback minimax adaptive control. $\leftarrow$ connect with Anders' research.

Summary

1. Multiple Model
   Adaptive Estimation
2. Bounded operator norm
3. Multiple Model Minimax Filters
4. Future Research
5. Summary

Resources

- Minimax Adaptive Estimation for Finite Sets of Linear Systems (Olle and Anders, 2021): https://arxiv.org/abs/2103.02479
- Julia package for recreating results and playing around: https://github.com/kjellqvist/MinimaxEstimation.jl
- Great review article — Gaussian Filters for Parameter and State Estimation (Afshari, et al., 2017): A General Review of Theory and Recent Trends: https://www.sciencedirect.com/science/article/pii/S0165168417300014
- Standard reference for all things related to linear filtering: Optimal Filtering by Anderson & Moore
- Maybecks case study (1999): https://doi.org/10.1002/(SICI)1099-1239(19991215)9:14<1051::AID-RNC452>3.0.CO;2-0