CDS 270-2: Lecture 4-2
Moving Horizon Estimation

Henrik Sandberg
19 April 2006

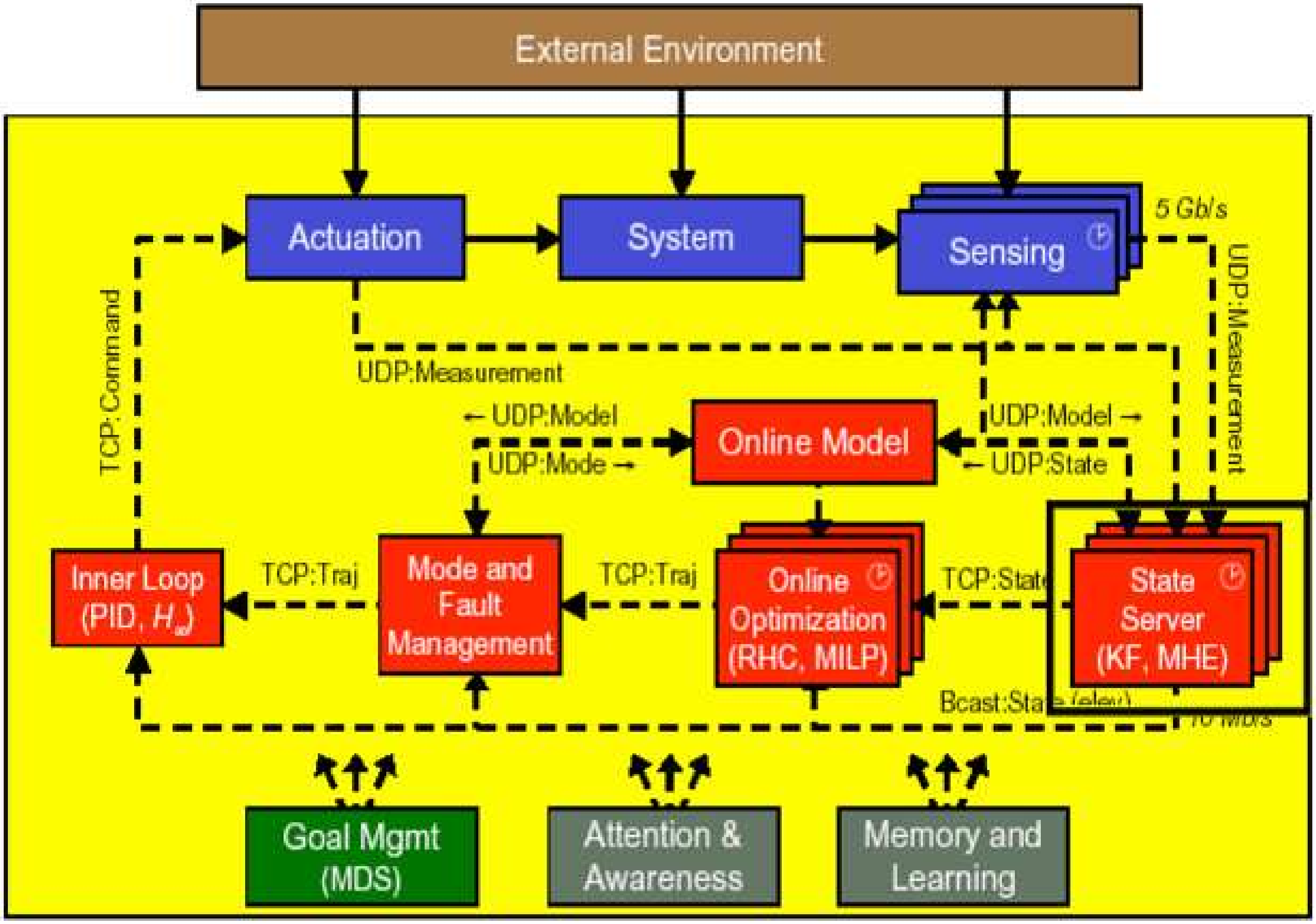
Goals:

- To learn how extended Kalman filters (EKF) and moving horizon estimation (MHE) works.
- To learn sufficient conditions for stability of MHE.

Reading:

- E.L. Haseltine and J.B. Rawlings: Critical evaluation of extended Kalman filtering and moving horizon estimation, Ind. Eng. Chem. Res., vol. 44, no.8, 2005.
- C.V. Rao, J.B. Rawlings, and D.Q. Mayne: Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations, IEEE TAC, vol.48, no.2, 2003.

Networked Control Systems



Today: The state server with extended Kalman filter and moving horizon estimation.

Problem Formulation

The model:

$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ y_k &= h_k(x_k) + v_k \end{aligned}$$

(include possible inputs u_k in f_k) with constraints

$$x_k \in \mathbb{X}_k, \quad w_k \in \mathbb{W}_k, \quad v_k \in \mathbb{V}_k.$$

Possibly nonlinear f_k and h_k!

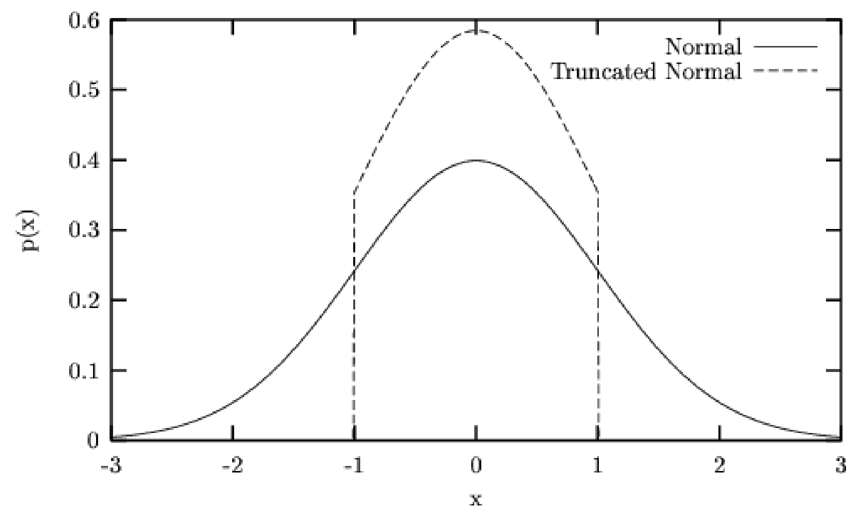
The problem: *Given* the data

$$Y_k = \{y_i : 0 \leq i \leq k\},$$

find the "best" (to be defined) estimate $\hat{x}_{k+m}$ of x_{k+m}.
($m = 0$ filtering, $m > 0$ prediction, and $m < 0$ smoothing.)

Constraints

- The noise w_k and v_k can typically have truncated Gaussian distributions:



- Be careful with bounds on measurement noise due to possibility of outliers.
- Constraints on state x_k more problematic to explain. May lead to acausality. A detailed enough model f_k should not need them. Nevertheless, state constraints are useful together with simplified models.

First Approach: The Extended Kalman Filter

Forget about constraints and linearize along estimated trajectory:

$$f_k(x_k, w_k) \approx f_k(\hat{x}_{k|k}, 0) + A_k(x_k - \hat{x}_{k|k}) + N_k w_k$$
$$h_k(x_k) \approx h_k(\hat{x}_{k|k-1}) + C_k(x_k - \hat{x}_{k|k-1})$$

where

$$A_k = \frac{\partial}{\partial x} f_k(x,0)\Big|_{x=\hat{x}_{k|k}}, N_k = \frac{\partial}{\partial w} f_k(\hat{x}_{k|k}, w)\Big|_{w=0} \quad C_k = \frac{\partial}{\partial x} h_k(x,0)\Big|_{x=\hat{x}_{k|k-1}}$$

Assume

$$\mathbb{E}\left\{ \begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_l^T & v_l^T \end{bmatrix} \right\} = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix} \delta_{kl} = \Sigma_k \delta_{kl}$$

and

$$\mathbb{E}\{x_0\} = \bar{x}_0 \quad \text{and} \quad \mathbb{E}\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} = P_0.$$

Extended Kalman Filter (EKF)

0. Initialization:

$$\hat{x}_{0|-1} = \bar{x}_0$$
$$P_{0|-1} = P_0$$

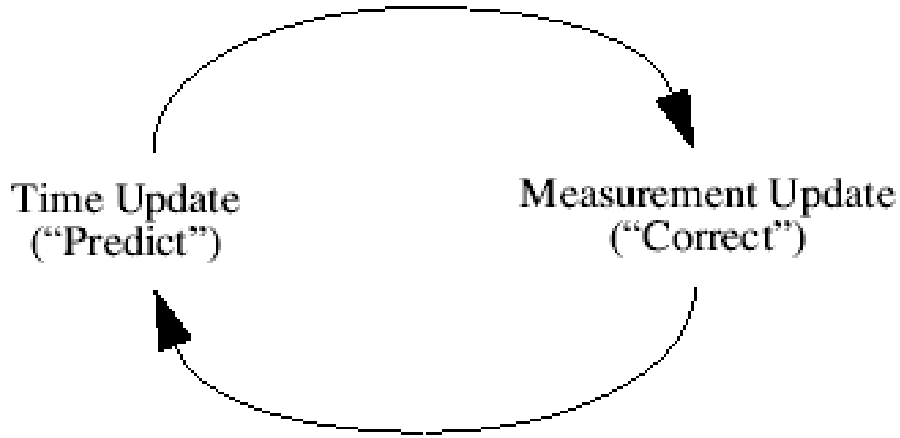
1. Corrector:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - h_k(\hat{x}_{k|k-1}))$$
$$K_k = P_{k|k-1}C_k^T(C_kP_{k|k-1}C_k^T + R_k)^{-1}$$
$$P_{k|k} = P_{k|k-1} - K_kC_kP_{k|k-1}$$

2. One-step predictor:

$$\hat{x}_{k+1|k} = f_k(\hat{x}_{k|k}, 0)$$
$$P_{k+1|k} = A_kP_{k|k}A_k^T + N_kQ_kN_k^T$$

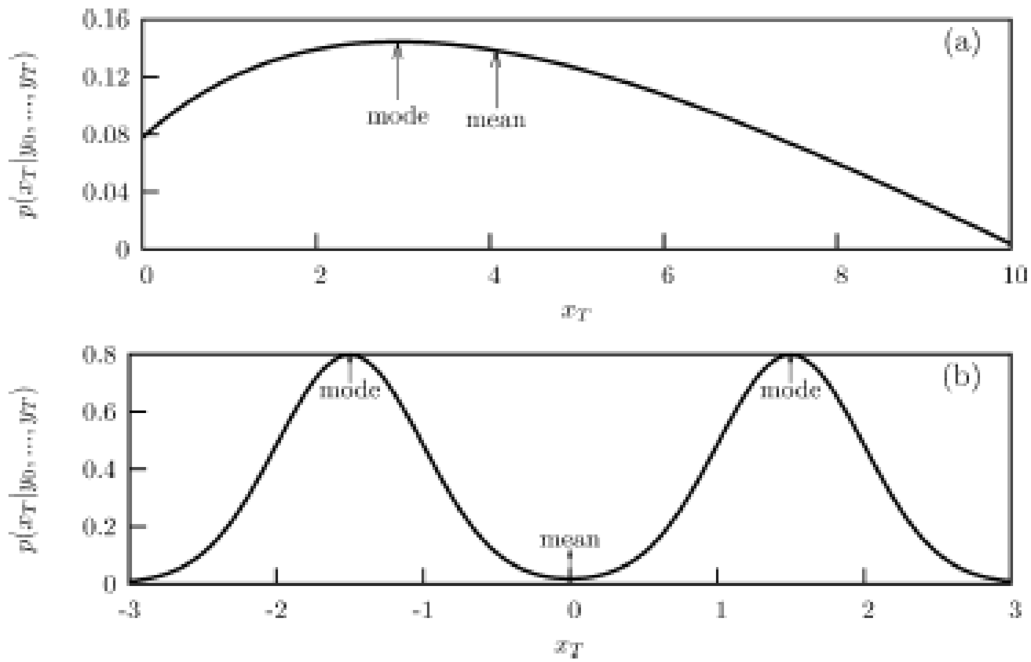
Comments on EKF



- EKF is simple and works often well (if linear approximation is valid, the noise is "small", and constraints not important).
- Hard to prove anything.
- $P_{k|k}$ and $P_{k|k-1}$ are no longer exact estimation error covariances. They give a rough estimate of the uncertainty, though.

Problems with Expectations and non-Gaussian Distributions

Ind. Eng. Chem. Res., Vol. 44, No. 8, 2005



The expectation (“mean”) may be a very unlikely process state!

Another Approach: MAP Estimates

- Idea: "Compute most likely values of the states $x_0, \ldots, x_T$, given measurements $y_0, \ldots, y_{T-1}$."
- Use Bayesian *maximum a posteriori* (MAP) criteria:

$$\{\hat{x}_0, \ldots, \hat{x}_T\} = \arg \max_{\{x_0, \ldots, x_T\}} p(x_0, \ldots, x_T | Y_{T-1})$$

- Assuming Gaussian white noise and $x_{k+1} = f_k(x_k) + w_k$:

$$\begin{aligned}
&\arg \max_{\{x_0, \ldots, x_T\}} p(x_0, \ldots, x_T | y_0, \ldots, y_{T-1}) \\
&= \arg \max_{\{x_0, \ldots, x_T\}} p_{x_0}(x_0) \prod_{k=0}^{T-1} p_{v_k}(y_k - h(x_k)) p(x_{k+1}|x_k) \\
&= \arg \max_{\{x_0, \ldots, x_T\}} \sum_{k=0}^{T-1} \log p_{v_k}(y_k - h_k(x_k)) + \log p(x_{k+1}|x_k) + \log p_{x_0}(x_0) \\
&= \arg \min_{\{x_0, \ldots, x_T\}} \sum_{k=0}^{T-1} \|y_k - h_k(x_k)\|^2_{R_k^{-1}} + \|x_{k+1} - f(x_k)\|^2_{Q_k^{-1}} + \|x_0 - \bar{x}_0\|^2_{P_0^{-1}}.
\end{aligned}$$

- Often the variables x_0 and $w_0, \ldots, w_{T-1}$ are used instead.

Interpretation

$$\min_{x_0,\{w_0,\ldots,w_{T-1}\}} \sum_{k=0}^{T-1} \|y_k - h_k(x_k)\|^2_{R_k^{-1}} + \|w_k\|^2_{Q_k^{-1}} + \|x_0 - \bar{x}_0\|^2_{P_0^{-1}}$$

Find initial state and process noise sequence such that

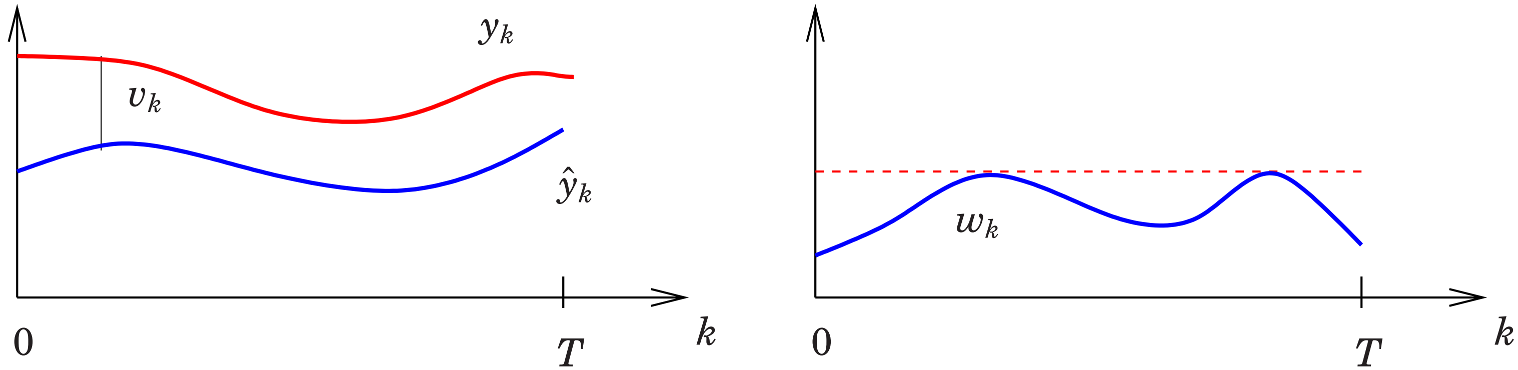
1. measurement data is matched (first term small);
2. process noise not larger than expected (second term small);
3. initial state not too far away from initial guess (third term small).

More generally we will consider the problem

$$\min_{x_0,\{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0),$$

subject to dynamics and constraints, for positive-definite functions L_k and Γ.

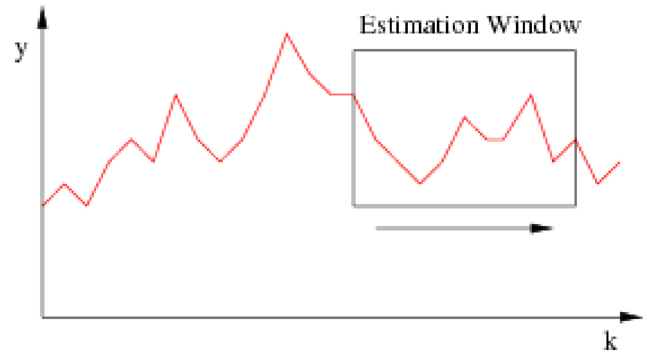
Possibilities and Problems



We note that

- we can try to enforce the constraints $x_k \in \mathbb{X}_k$, $w_k \in \mathbb{W}_k$, $v_k \in \mathbb{V}_k$ in the optimization;
- the optimization problem may have multiple local minima;
- the problem complexity grows at least linearly with the horizon T;
- for linear systems without constraints and quadratic cost, the problem can be solved recursively and leads to the Kalman filter.

Moving Horizon Estimation (MHE)



To reduce the computation cost with increasing T, consider a moving horizon backwards in time of length N (compare with RHC):

$$\Phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \left(\sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \sum_{k=0}^{T-N-1} L_k(w_k, v_k) + \Gamma(x_0) \right)$$

$$= \min_{z \in \mathcal{R}_{T-N}, \{w_k\}_{k=T-N}^{T-1}} \left(\sum_{k=T-N}^{T-1} L_k(w_k, v_k) + Z_{T-N}(z) \right).$$

The Arrival Cost

The *arrival cost* is defined as

$$\mathcal{Z}_{T-N}(z) = \min_{x_0, \{w_k\}_{k=0}^{T-N-1}} \sum_{k=0}^{T-N-1} L_k(w_k, v_k) + \Gamma(x_0)$$

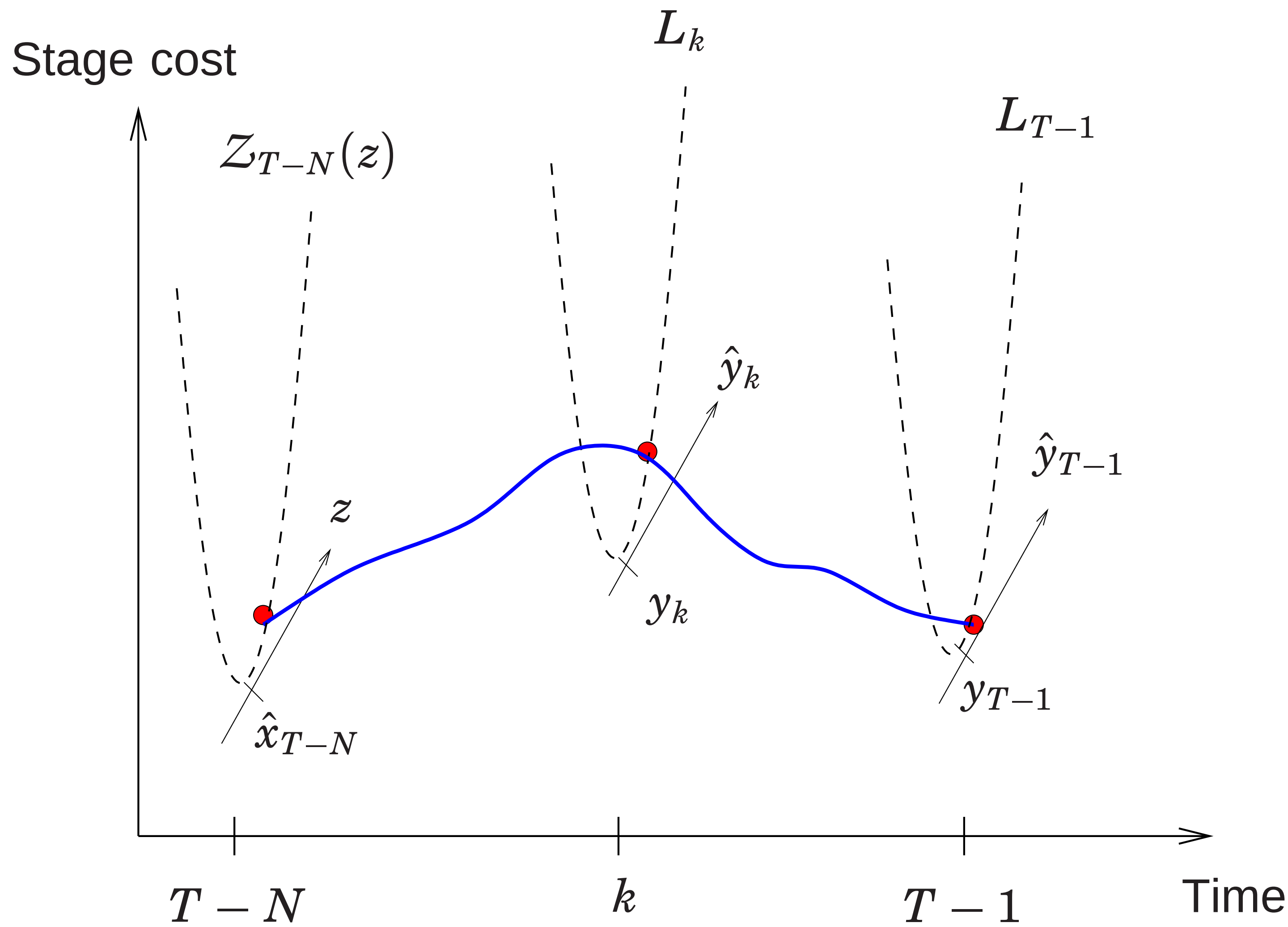
subject to constraints, dynamics, and $x_{T-N} = z$.

- Compare the arrival cost to the *terminal cost* in RHC.
- The arrival cost is hard to compute. Approximations $\hat{\mathcal{Z}}_{T-N}$ often necessary. Examples of this later.
- For linear systems without constraints and quadratic cost, the *exact* expression is:

$$\mathcal{Z}_{T-N}(z) = (z - \hat{x}_{T-N})^T P_{T-N}^{-1} (z - \hat{x}_{T-N}) + \Phi_{T-N}^*,$$

Interpretation? Just minimizing arrival cost should give back $\hat{x}_{T-N|T-N-1}$ (our previous best estimate)!

Minimize the Total Cost Over the Horizon N



The state z is the optimal updated estimate of the state at time $T-N$, after taking $y_{T-N}, \ldots, y_{T-1}$ into account.

“Implementation”

Sketch of algorithm

0. Fix horizon N, and find L_k and $\mathcal{Z}_k$(hard!).
1. Solve

$$\{z^*, \{\hat{w}_k\}_{T-N}^{T-1}\} = \arg \min_{z \in \mathcal{R}_{T-N}, w_k} \left(\sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \mathcal{Z}_{T-N}(z) \right)$$

   subject to dynamics and $x_k \in \mathbb{X}_k$, $w_k \in \mathbb{W}_k$, $v_k \in \mathbb{V}_k$.

   (Use (S)QP, NPSOL, NTG,. . .)
2. Compute estimates $\hat{x}_{k|T-1}$, $k \in [T-N, T]$:

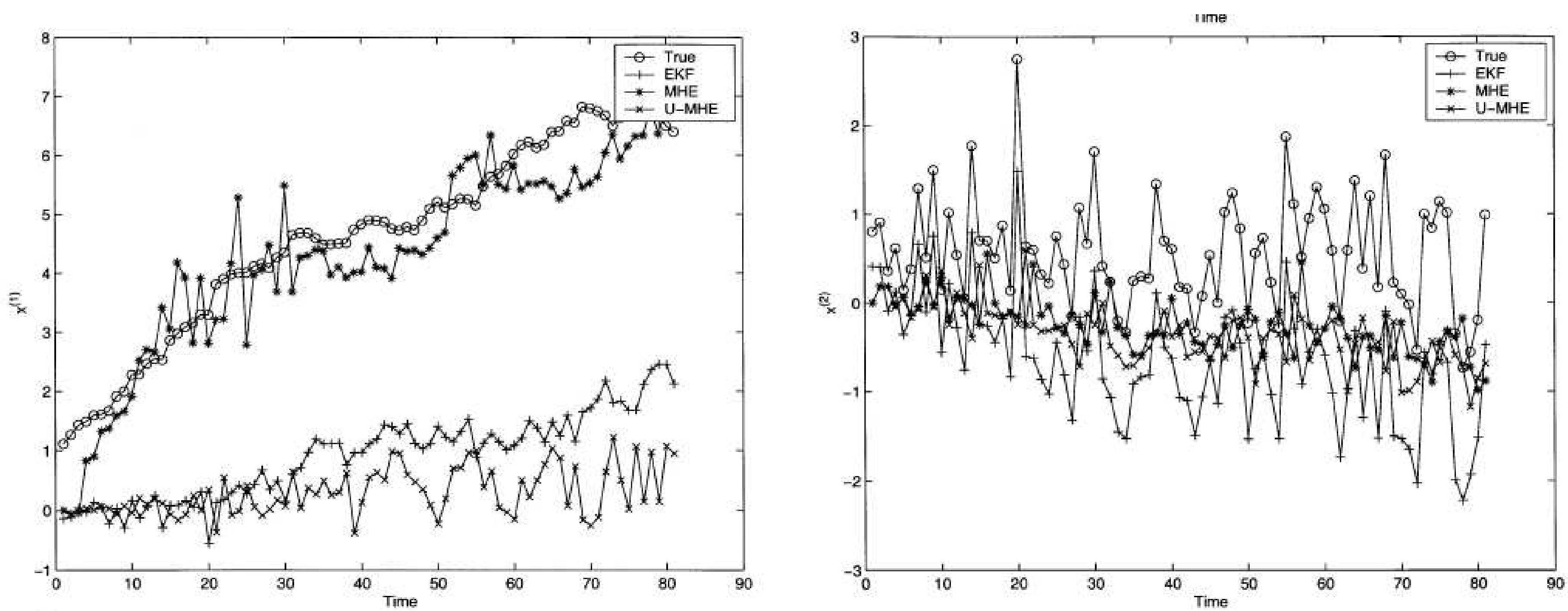
$$\hat{x}_{T-N|T-1} := z^*, \quad \dots, \quad \hat{x}_{T|T-1} := f_{T-1}(\hat{x}_{T-1|T-1}, \hat{w}_{T-1})$$

3. Increase T, update $\mathcal{Z}_{T-N}$, and goto 1.

Example 1: EKF vs. MHE (Rao et al. 2003)

$$x_{1,k+1} = 0.99x_{1,k} + 0.2x_{2,k}$$
$$x_{2,k+1} = -0.1x_{1,k} + \frac{0.5x_{2,k}}{1 + x_{2,k}^2} + w_k$$
$$y_k = x_{1,k} - 3x_{2,k} + v_k$$

v_k Gaussian, w_k truncated Gaussian such that $w_k \geq 0$. Compare EKF and MHE ($N = 10$, with/without constraints).



Example 1: EKF vs. MHE (Rao et al. 2003)

- Constraints are important! Make all the difference in this example.
- Computation time (interpreted code, 500 MHz processor):
  - MHE: $\sim$3 s per time step
  - EKF: negligible
- GNU Octave toolbox NMPC is available at

  `http://jbrwww.che.wisc.edu/home/tenny/nmpc/`

  (Use CVS version!)

Approximate MHE

$$\hat{\Phi}_T = \min_{z \in \mathcal{R}_{T-N}, \{w_k\}_{k=T-N}^{T-1}} \left(\sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \right)$$

Two common approximate arrival costs $\hat{\mathcal{Z}}_{T-N}$:

(A) Forget initial guess (independent of z):

$$\hat{\mathcal{Z}}_{T-N}(z) = \hat{\Phi}_{T-N}.$$

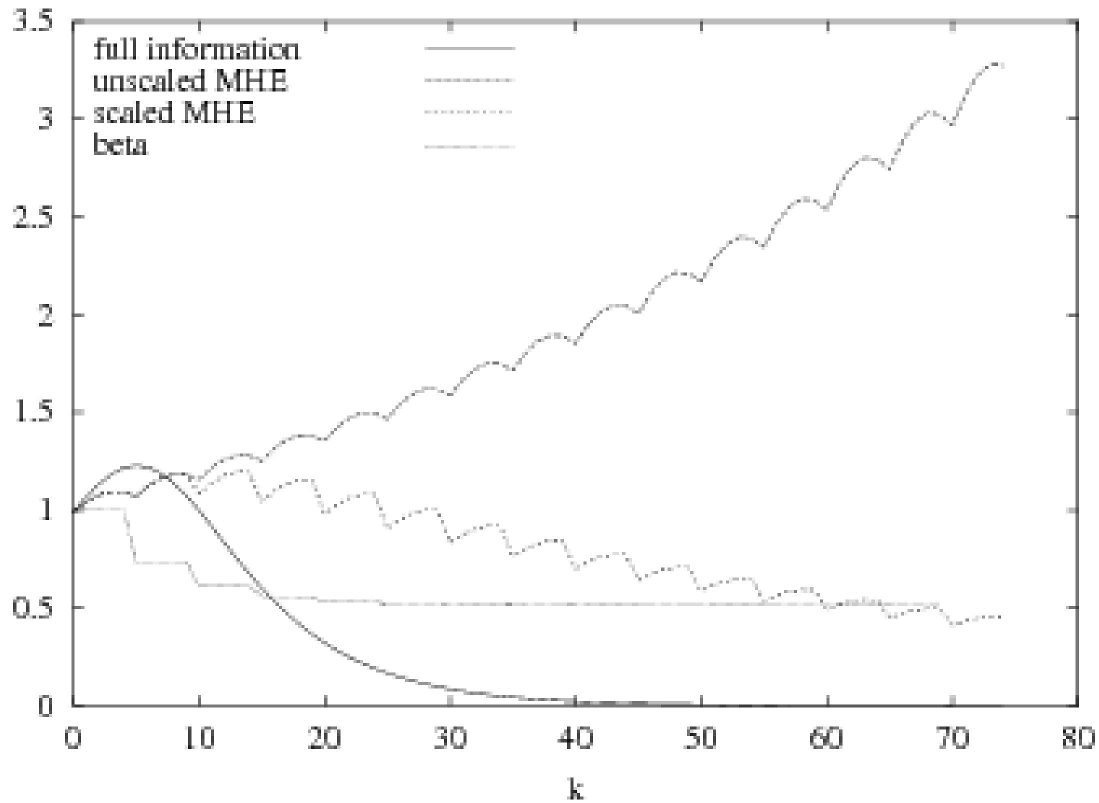
(B) Use EKF around estimated trajectory, $\{\hat{x}_k\}_0^{T-N}$, beyond the moving horizon:

$$\hat{\mathcal{Z}}_{T-N}(z) = (z - \hat{x}_{T-N})^T P_{T-N}^{-1} (z - \hat{x}_{T-N}) + \hat{\Phi}_{T-N}.$$

P_{T-N} computed with a second-order Taylor expansion of L_k and Γ, and the one-step predictor formulas.

Example 2: Approximation Problems (Rao 2000)

Estimation error plotted:



A bad approximation $\hat{\mathcal{Z}}_{T-N}$ may result in unstable estimations!

Next: We give some sufficient conditions for asymptotic stability of (approximate) MHE.

Asymptotic Stability of Estimation

- Asymptotic stability means that when there is no noise in the system, i.e., $w_k = v_k = 0$, then

$$\|x_k - \hat{x}_k\| \to 0, \quad k \to \infty,$$

  if the initial estimate is good enough. That is, the estimate converges to the true state.

- When bounded noise is present, we require that $\|x_k - \hat{x}_k\|$ is uniformly bounded for all k.

Three Conditions for Asymptotic Stability

C1. Uniform observability: With initial conditions x^1 and x^2

$$0 \leq \varphi(\|x^1 - x^2\|) \leq \sum_{j=k}^{k+N-1} \|y_k^1 - y_k^2\|, \quad \text{for all} \quad k.$$

C2. A cost to change initial estimate:

$$0 \leq \hat{\mathcal{Z}}_{T-N}(z) - \hat{\Phi}_{T-N} \leq \gamma(\|z - \hat{x}_{T-N}\|).$$

C3. Add no information: Some forgetting of information outside the moving horizon

$$\hat{\mathcal{Z}}_T(p) \leq \min_{z \in \mathcal{R}_{T-N}, \{w_k\}_{k=T-N}^{T-1}} \left(\sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \hat{\mathcal{Z}}_{T-N}(z) \right)$$

subject to $x_T = p \in \mathcal{R}_T$.

Discussion

- $\mathbf{C1}$ is system dependent. A long horizon N is good.
- Condition $\mathbf{C3}$ is often hard to prove. If it is violated, an estimate p is considered worse than there are reason for.
- The approximation $\mathbf{(A)}$ satisfies $\mathbf{C2}$-$\mathbf{C3}$. May have bad performance, though, since it has complete forgetting.
- The approximation $\mathbf{(B)}$ (using EKF) does not necessarily satisfy $\mathbf{C3}$. Exception: linear dynamics, quadratic cost and convex constraints.

> If $\mathbf{C1}$-$\mathbf{C3}$ and some technical assumptions on f_k, h_k, L_k, Γ are true (there are solutions etc.), then *MHE is asymptotically stable.*

Proof. Lyapunov-type argument. See Rao et al. 2003. Compare with RHC.

Summary

- The extended Kalman filter (EKF) is a simple and often good estimator for nonlinear systems. No constraints are taken into account. Not many proofs.
- Moving horizon estimation (MHE) is more powerful (and complicated): nonlinear models, constraints, and asymmetric noise distributions. Dual to RHC.
- Constraints come from physical insight ("positive flows") or approximations.
- The arrival cost $\mathcal{Z}_T$ often needs to be approximated. (Arrival cost $\sim$ Terminal cost in RHC.)
- MHE performance relies on good optimization performance and good approximations.
- Asymptotic stability follows from some natural conditions: observability, cost to update, some forgetting,...