Moving Horizon Estimation and Minimum Energy Estimation

Arthur J. Krener

ajkrener@nps.edu

Supported by AFOSR and DARPA

Discrete Time Filtering

There are three variations on filtering problems:

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**
- **Continuous time dynamics, discrete time observations**

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**
- **Continuous time dynamics, discrete time observations**
- **Discrete time dynamics, discrete time observations**

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**
- **Continuous time dynamics, discrete time observations**
- **Discrete time dynamics, discrete time observations**

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**
- **Continuous time dynamics, discrete time observations**
- **Discrete time dynamics, discrete time observations**

We shall focus on the last. Here is the model of the plant.

Discrete Time Filtering

There are three variations on filtering problems:

- **Continuous time dynamics, continuous time observations**
- **Continuous time dynamics, discrete time observations**
- **Discrete time dynamics, discrete time observations**

We shall focus on the last. Here is the model of the plant.

$$
\begin{aligned}
x^+ &= f(t, x, u) \\
y &= h(t, x, u) \\
x(0) &= x^0
\end{aligned}
$$

where the state $x(t) \in \mathbb{R}^{n\times 1}$, the control $u(t) \in \mathbb{R}^{m\times 1}$, the measurement $y(t) \in \mathbb{R}^{p\times 1}$ and $x^+(t) = x(t+1)$.

Discrete Time Filtering

There are three variations on filtering problems:

- Continuous time dynamics, continuous time observations
- Continuous time dynamics, discrete time observations
- Discrete time dynamics, discrete time observations

We shall focus on the last. Here is the model of the plant.

$$\begin{aligned} x^+ &= f(t,x,u) \\ y &= h(t,x,u) \\ x(0) &= x^0 \end{aligned}$$

where the state $x(t) \in I\!R^{n\times 1}$, the control $u(t) \in I\!R^{m\times 1}$, the measurement $y(t) \in I\!R^{p\times 1}$ and $x^+(t) = x(t+1)$.

The filtering problem is to estimate $x(t)$ from $u(s), 0 \le s \le t-1$, from $y(s), 1 \le s \le t-1$ and from some inexact knowledge about x^0 . We denote this estimate by $\hat{x}(t|t-1)$. We also consider estimating $x(t)$ using the additional measurement $y(t)$, we denote this estimate $\hat{x}(t|t)$.

Three Noises

The standard approach is add three noises to the model, a driving noise $v(t) \in \mathbb{R}^{k\times 1}$, an observation noise $w(t) \in \mathbb{R}^{p\times 1}$ and an initial condition noise $\tilde{x}^0 \in \mathbb{R}^{n\times 1}$ to obtain

$$\begin{aligned} x^+ &= f(t,x,u) + g(t,x,u)v \\ y &= h(t,x,u) + w \\ x(0) &= \hat{x}^0 + \tilde{x}^0 \end{aligned}$$

where $\hat{x}^0$ is our estimate of $x(0)$ based on prior information.

Three Noises

The standard approach is add three noises to the model, a driving noise $v(t) \in \mathbb{R}^{k\times 1}$, an observation noise $w(t) \in \mathbb{R}^{p\times 1}$ and an initial condition noise $\tilde{x}^0 \in \mathbb{R}^{n\times 1}$ to obtain

$$\begin{aligned} x^+ &= f(t,x,u) + g(t,x,u)v \\ y &= h(t,x,u) + w \\ x(0) &= \hat{x}^0 + \tilde{x}^0 \end{aligned}$$

where $\hat{x}^0$ is our estimate of $x(0)$ based on prior information.

Stochastic Filtering

The stochastic version of this approach is to assume that $v(t)$ and $w(t)$ are independent, standard white Gaussian noise processes and $\tilde{x}^0$ is an independent Gaussian random vector of mean zero and known covariance.

Stochastic Filtering

The stochastic version of this approach is to assume that $v(t)$ and $w(t)$ are independent, standard white Gaussian noise processes and $\tilde{x}^0$ is an independent Gaussian random vector of mean zero and known covariance.

Then the conditional density $\rho(x, t|t-1)$ of the state given the past controls $u(s),\ 0 \le s \le t-1$ and past measurements $y(s),\ 1 \le s \le t-1$ satisfies an integral-differential-difference equation that is very difficult to solve.

Stochastic Filtering

The stochastic version of this approach is to assume that $v(t)$ and $w(t)$ are independent, standard white Gaussian noise processes and $\tilde{x}^0$ is an independent Gaussian random vector of mean zero and known covariance.

Then the conditional density $\rho(x, t|t-1)$ of the state given the past controls $u(s),\ 0 \leq s \leq t-1$ and past measurements $y(s),\ 1 \leq s \leq t-1$ satisfies an integral-differential-difference equation that is very difficult to solve.

About the only time it can be solved is when the system is linear for then the conditional density is Gaussian. Then its mean and covariance can be computed by a Kalman filter.

Deterministic Filtering

The minimum energy version of this approach is to assume the noises are deteministic but unknown. One finds the noise triple that minimizes a so-called "energy" like

$$\frac{1}{2}\left(\|\tilde{x}^0\|^2 + \sum_{s=0}^{t-1}\|v(s)\|^2 + \sum_{s=1}^{t}\|w(s)\|^2\right)$$

Deterministic Filtering

The minimum energy version of this approach is to assume the noises are deteministic but unknown. One finds the noise triple that minimizes a so-called "energy" like

$$\frac{1}{2}\left(\|\tilde{x}^0\|^2 + \sum_{s=0}^{t-1}\|v(s)\|^2 + \sum_{s=1}^{t}\|w(s)\|^2\right)$$

subject to

$$\begin{aligned} z^+ &= f(s,z,u) + g(s,z,u)v \\ y &= h(s,z,u) + w \\ z(0) &= \hat{x}^0 + \tilde{x}^0 \end{aligned}$$

where $u(s), 0 \leq s \leq t-1$ **and** $y(s), 1 \leq s \leq t-1$ **are the actual control and measurement sequences.**

Deterministic Filtering

The minimum energy version of this approach is to assume the noises are deteministic but unknown. One finds the noise triple that minimizes a so-called "energy" like

$$\frac{1}{2}\left(\|\tilde{x}^0\|^2 + \sum_{s=0}^{t-1}\|v(s)\|^2 + \sum_{s=1}^{t}\|w(s)\|^2\right)$$

subject to

$$\begin{aligned} z^+ &= f(s,z,u) + g(s,z,u)v \\ y &= h(s,z,u) + w \\ z(0) &= \hat{x}^0 + \tilde{x}^0 \end{aligned}$$

where $u(s), 0 \le s \le t-1$ **and** $y(s), 1 \le s \le t-1$ **are the actual control and measurement sequences.**

The MME is the endpoint of the minimizing state trajectory

$$\hat{x}(t|t) = z(t)$$

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

This idea goes back to Gauss' method of least squares to determine of the orbit of the planetoid Ceres.

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

This idea goes back to Gauss' method of least squares to determine of the orbit of the planetoid Ceres.

Its application to filtering of control systems goes back at least to Mortenson (1968) who called it Maximum Likelyhood Estimation.

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

This idea goes back to Gauss' method of least squares to determine of the orbit of the planetoid Ceres.

Its application to filtering of control systems goes back at least to Mortenson (1968) who called it Maximum Likelyhood Estimation.

Hijab (1980) called it Minimum Energy Estimation (MEE).

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

This idea goes back to Gauss' method of least squares to determine of the orbit of the planetoid Ceres.

Its application to filtering of control systems goes back at least to Mortenson (1968) who called it Maximum Likelyhood Estimation.

Hijab (1980) called it Minimum Energy Estimation (MEE).

Both Mortenson and Hijab dealt with systems in continuous time, but their methods readily extend to discrete time systems.

Minimum Energy Estimation

The heuristic behind MME is that the noises can be thought of as "slack variables" to make the system equations hold. The Minimum Energy Estimate requires the least slack.

This idea goes back to Gauss' method of least squares to determine of the orbit of the planetoid Ceres.

Its application to filtering of control systems goes back at least to Mortenson (1968) who called it Maximum Likelyhood Estimation.

Hijab (1980) called it Minimum Energy Estimation (MEE).

Both Mortenson and Hijab dealt with systems in continuous time, but their methods readily extend to discrete time systems.

If the system is linear then MME filtering is identical to Kalman filtering provided that the weights of the norms in the deterministic version are the inverses of the covariances of the noises in the stochastic version.

Minimum Horizon Estimation

Minimum Horizon Estimation (MHE) is a moving window version of MEE. It has become very popular recently as it is the mathematical dual of Model Predictive Control (MPC).

Minimum Horizon Estimation

Minimum Horizon Estimation (MHE) is a moving window version of MEE. It has become very popular recently as it is the mathematical dual of Model Predictive Control (MPC).

Again the system model is

$$x^+ = f(t, x, u), \qquad y = h(t, x, u)$$

Minimum Horizon Estimation

Minimum Horizon Estimation (MHE) is a moving window version of MEE. It has become very popular recently as it is the mathematical dual of Model Predictive Control (MPC).

Again the system model is

$$x^+ = f(t,x,u), \qquad y = h(t,x,u)$$

One chooses a past window of length T . At time t minimize

$$\min_{z^{t-T},v,w} \left\{ \|z^{t-T} - \hat{x}(t-T|t-T)\|^2_{P(t-T)} + \sum_{s=t-T}^{t-1} \|v(s)\|^2_{Q(s)} + \sum_{s=t-T+1}^{t} \|w(s)\|^2_{R(s)} \right\}$$

subject to

$$\begin{aligned} z^+ &= f(s,z,u) + g(s,z,u)v \\ y &= h(s,z,u) + w \end{aligned}$$

where $u(s)$ is the actual input and $y(s)$ is the actual output.

Minimum Horizon Estimation

The MHE estimate is

$$\hat{x}(t|t) \;\; = \;\; z(t)$$

where $z(s)$ is the minimizing state trajectory.

Minimum Horizon Estimation

The MHE estimate is

$$\hat{x}(t|t) \;\; = \;\; z(t)$$

where $z(s)$ is the minimizing state trajectory.

This minimization problem is solved in real time by a nonlinear program solver. The complexity of the MHE nonlinear program is dictated by the nonlinearities of f, g, h and the length of the past observation window T .

Minimum Horizon Estimation

The MHE estimate is

$$\hat{x}(t|t) = z(t)$$

where $z(s)$ is the minimizing state trajectory.

This minimization problem is solved in real time by a nonlinear program solver. The complexity of the MHE nonlinear program is dictated by the nonlinearities of f, g, h and the length of the past observation window T .

There is not much we can do about f, h as they are given by the plant and we can take g constant as we did before. So the question is can we shorten T and still get reasonable estimates.

Minimum Horizon Estimation

The MHE estimate is

$$\hat{x}(t|t) \;=\; z(t)$$

where $z(s)$ is the minimizing state trajectory.

This minimization problem is solved in real time by a nonlinear program solver. The complexity of the MHE nonlinear program is dictated by the nonlinearities of f, g, h and the length of the past observation window T .

There is not much we can do about f, h as they are given by the plant and we can take g constant as we did before. So the question is can we shorten T and still get reasonable estimates.

We believe that the power series techniques described below will allow us to take T as short as 1 time step. Instead of taking a fixed past window of length T we use the whole past but with a forgetting factor $0 < \alpha \leq 1$.

MEE

Since we know $u(t)$ we can simplify notation by redefining

$$\begin{aligned} f(t,x) &= f(t,x,u(t)) \\ g(t,x) &= g(t,x,u(t)) \\ h(t,x) &= h(t,x,u(t)) \end{aligned}$$

MEE

Since we know $u(t)$ we can simplify notation by redefining

$$\begin{aligned} f(t,x) &= f(t,x,u(t)) \\ g(t,x) &= g(t,x,u(t)) \\ h(t,x) &= h(t,x,u(t)) \end{aligned}$$

Then the noisy model is

$$\begin{aligned} z(t+1) &= f(t,z(t)) + g(t,z(t))v(t) \\ y &= h(t,z(t)) + w(t) \\ z(0) &= \hat{x}^0 + \tilde{z}^0 \end{aligned}$$

MEE

Since we know $u(t)$ we can simplify notation by redefining

$$\begin{aligned} f(t,x) &= f(t,x,u(t)) \\ g(t,x) &= g(t,x,u(t)) \\ h(t,x) &= h(t,x,u(t)) \end{aligned}$$

Then the noisy model is

$$\begin{aligned} z(t+1) &= f(t,z(t)) + g(t,z(t))v(t) \\ y &= h(t,z(t)) + w(t) \\ z(0) &= \hat{x}^0 + \tilde{z}^0 \end{aligned}$$

The simplest definition of "energy' to time t is

$$\frac{\alpha^t}{2}\|z^0\|^2 + \frac{1}{2}\sum_{s=0}^{t-1}\alpha^{t-s}\|v(s)\|^2 + \frac{1}{2}\sum_{s=1}^{t}\alpha^{t-s}\|w(s)\|^2$$

with a discount factor $0 < \alpha \leq 1$.

MEE

We can extend MME by considering a more general form of "energy"

$$\alpha^t \pi^0(z^0) + \frac{1}{2}\sum_{s=0}^{t-1} \alpha^{t-s}\|v(s)\|^2_{Q(s)} + \frac{1}{2}\sum_{s=1}^{\tau} \alpha^{t-s}\|w(s)\|^2_{R(s)}$$

where $Q(s) > 0,\ R(s) > 0$ **, an initial energy** $\pi^0(x) \geq 0$ **and**

$$\|v(s)\|^2_{Q(s)} = v'(s)Q(s)v(s), \qquad \|w(s)\|^2_{R(s)} = w'(s)R(s)w(s)$$

MEE

We can extend MME by considering a more general form of "energy"

$$\alpha^t \pi^0(z^0) + \frac{1}{2} \sum_{s=0}^{t-1} \alpha^{t-s} \|v(s)\|^2_{Q(s)} + \frac{1}{2} \sum_{s=1}^{\tau} \alpha^{t-s} \|w(s)\|^2_{R(s)}$$

where $Q(s) > 0, \ R(s) > 0$, an initial energy $\pi^0(x) \geq 0$ and

$$\|v(s)\|^2_{Q(s)} = v'(s)Q(s)v(s), \qquad \|w(s)\|^2_{R(s)} = w'(s)R(s)w(s)$$

Notice if we only consider measurements up to time $\tau \leq t$ then causality implies that the minimizing $v(s) = 0$ for $s > \tau$.

MEE

We can extend MME by considering a more general form of "energy"

$$\alpha^t \pi^0(z^0) + \frac{1}{2} \sum_{s=0}^{t-1} \alpha^{t-s} \|v(s)\|^2_{Q(s)} + \frac{1}{2} \sum_{s=1}^{\tau} \alpha^{t-s} \|w(s)\|^2_{R(s)}$$

where $Q(s) > 0, \; R(s) > 0$, an initial energy $\pi^0(x) \geq 0$ and

$$\|v(s)\|^2_{Q(s)} = v'(s)Q(s)v(s), \qquad \|w(s)\|^2_{R(s)} = w'(s)R(s)w(s)$$

Notice if we only consider measurements up to time $\tau \leq t$ then causality implies that the minimizing $v(s) = 0$ for $s > \tau$.

Or even more generally

$$\alpha^t \pi^0(z^0) + \frac{1}{2} \sum_{s=0}^{t-1} \alpha^{t-s} l(s, z(s), v(s), w(s))$$

subject to constraints on the Lagrangian $l(s, z, v, w)$. For simplicity of exposition we shall stick with quadratic Lagrangians.

MME Function

The MEE function $\pi(x, t|\tau)$ is the minimum of the generalized energy subject to

$$\begin{aligned} z^{+} &= f(s, z(s)) + g(s, z(s))v(s) \\ y(s) &= h(s, z) + w(s) \\ z(0) &= \hat{x}^{0} + \tilde{x}^{0} \\ z(t) &= x \end{aligned}$$

where $y(s)$ are the actual observations.

MME Function

The MEE function $\pi(x, t|\tau)$ is the minimum of the generalized energy subject to

$$\begin{aligned} z^{+} &= f(s, z(s)) + g(s, z(s))v(s) \\ y(s) &= h(s, z) + w(s) \\ z(0) &= \hat{x}^{0} + \tilde{x}^{0} \\ z(t) &= x \end{aligned}$$

where $y(s)$ are the actual observations.

Notice the additional terminal condition, $z(t) = x$.

MME Function

The MEE function $\pi(x, t|\tau)$ is the minimum of the generalized energy subject to

$$\begin{aligned} z^{+} &= f(s, z(s)) + g(s, z(s))v(s) \\ y(s) &= h(s, z) + w(s) \\ z(0) &= \hat{x}^{0} + \tilde{x}^{0} \\ z(t) &= x \end{aligned}$$

where $y(s)$ are the actual observations.

Notice the additional terminal condition, $z(t) = x$.

The minimum energy estimate is then

$$\hat{x}(t|\tau) = \operatorname{argmin}_x \pi(x, t|\tau)$$

MME Function

The MEE function $\pi(x, t|\tau)$ is the minimum of the generalized energy subject to

$$\begin{aligned} z^+ &= f(s, z(s)) + g(s, z(s))v(s) \\ y(s) &= h(s, z) + w(s) \\ z(0) &= \hat{x}^0 + \tilde{x}^0 \\ z(t) &= x \end{aligned}$$

where $y(s)$ are the actual observations.

Notice the additional terminal condition, $z(t) = x$.

The minimum energy estimate is then

$$\hat{x}(t|\tau) = \text{argmin}_x \pi(x, t|\tau)$$

We are primarily interested in the cases where $\tau = t - 1$ or $\tau = t$.

MEE

Notice that the MEE function $\pi(x, t|\tau)$ contains more information than its minimum, the simple point estimate $\hat{x}(t|\tau) = \textbf{argmin}_x \pi(x, t|\tau)$.

MEE

Notice that the MEE function $\pi(x, t|\tau)$ contains more information than its minimum, the simple point estimate $\hat{x}(t|\tau) = \text{argmin}_x \pi(x, t|\tau)$.

For starters its minimum value $\pi(\hat{x}(t|\tau), t|\tau)$ is a residual which measures the "reliability" of the point estimate. The smaller it is the more confident we are of the estimate.

MEE

Notice that the MEE function $\pi(x, t|\tau)$ contains more information than its minimum, the simple point estimate $\hat{x}(t|\tau) = \textbf{argmin}_x \pi(x, t|\tau)$.

For starters its minimum value $\pi(\hat{x}(t|\tau), t|\tau)$ is a residual which measures the "reliability" of the point estimate. The smaller it is the more confident we are of the estimate.

The eigenvalues of the Hessian of $\pi(x, t|\tau)$ at $\hat{x}(t|\tau)$ measure how "tight" is the estimate. The bigger they are, the more precise is the point estimate.

MEE

Notice that the MEE function $\pi(x, t|\tau)$ contains more information than its minimum, the simple point estimate $\hat{x}(t|\tau) = \textbf{argmin}_x \pi(x, t|\tau)$.

For starters its minimum value $\pi(\hat{x}(t|\tau), t|\tau)$ is a residual which measures the "reliability" of the point estimate. The smaller it is the more confident we are of the estimate.

The eigenvalues of the Hessian of $\pi(x, t|\tau)$ at $\hat{x}(t|\tau)$ measure how "tight" is the estimate. The bigger they are, the more precise is the point estimate.

Finally $\pi(x, t|\tau)$ might have multiple local minima. As these change with t the global minimum might jump from one to another.

MEE

Notice that the MEE function $\pi(x, t|\tau)$ contains more information than its minimum, the simple point estimate $\hat{x}(t|\tau) = \text{argmin}_x \pi(x, t|\tau)$.

For starters its minimum value $\pi(\hat{x}(t|\tau), t|\tau)$ is a residual which measures the "reliability" of the point estimate. The smaller it is the more confident we are of the estimate.

The eigenvalues of the Hessian of $\pi(x, t|\tau)$ at $\hat{x}(t|\tau)$ measure how "tight" is the estimate. The bigger they are, the more precise is the point estimate.

Finally $\pi(x, t|\tau)$ might have multiple local minima. As these change with t the global minimum might jump from one to another.

Except for the residual, this additional information is not available with MHE.

Prediction Step

Filtering of a discrete time system is usually done in two alternating steps.

Prediction Step

Filtering of a discrete time system is usually done in two alternating steps.

In MEE the prediction step starts with $\hat{x}(t|t)$ and $\pi(x,t|t)$ and computes $\hat{x}(t+1|t)$ and $\pi(x,t+1|t)$ by a solving functional equation for π of dynamic programming type,

$$\pi(x,t+1|t) \;\; = \;\; \min_{z,v} \left\{ \alpha\pi(z,t|t) + \frac{1}{2}\|v(t)\|^2_{Q(t)} \right\}$$

subject to the constraint

$$x \;\; = \;\; f(t,z) + g(t,z)v$$

Prediction Step

Filtering of a discrete time system is usually done in two alternating steps.

In MEE the prediction step starts with $\hat{x}(t|t)$ and $\pi(x,t|t)$ and computes $\hat{x}(t+1|t)$ and $\pi(x,t+1|t)$ by a solving functional equation for π of dynamic programming type,

$$\pi(x,t+1|t) \;=\; \min_{z,v}\left\{\alpha\pi(z,t|t)+\frac{1}{2}\|v(t)\|^2_{Q(t)}\right\}$$

subject to the constraint

$$x \;=\; f(t,z)+g(t,z)v$$

The predicted estimate $\hat{x}(t+1|t)$ is given by

$$\hat{x}(t+1|t) \;=\; \text{argmin}_x \pi(x,t+1|t)$$

Assimilation Step

When $y(t)$ becomes known the assimilation step is given by the functional equation,

$$\pi(x,t|t) \;=\; \pi(x,t|t-1) + \frac{1}{2}\|y(t) - h(t,x)\|^2_{R(t)}$$

Assimilation Step

When $y(t)$ becomes known the assimilation step is given by the functional equation,

$$\pi(x,t|t) \;\; = \;\; \pi(x,t|t-1) + \frac{1}{2}\|y(t) - h(t,x)\|^2_{R(t)}$$

The assimilated estimate $\hat{x}(t|t)$ is

$$\hat{x}(t|t) \;\; = \;\; \text{argmin}_x \pi(x,t|t)$$

Constrained Optimiztion

The prediction step requires the solution of a family of constrained optimization problems indexed by the variables $t,\ x$ of the terminal constraint $z(t) = x$.

Constrained Optimiztion

The prediction step requires the solution of a family of constrained optimization problems indexed by the variables t, x of the terminal constraint $z(t) = x$.

To get unconstrained problems we add the constraint to the criterion with a Lagrange multiplier $\lambda(t,x) \in \mathbb{R}^{n\times 1}$ to get a family of unconstrained minimization problems indexed by t, x ,

$$\min_{z,v,\lambda} \left\{ \alpha\pi(z,t|t) + \frac{1}{2}\|v\|^2_{Q(t)} + \lambda'(t,x)\left(x - f(t,z) - g(t,z)v\right) \right\}$$

Constrained Optimiztion

The prediction step requires the solution of a family of constrained optimization problems indexed by the variables $t,\ x$ of the terminal constraint $z(t) = x$.

To get unconstrained problems we add the constraint to the criterion with a Lagrange multiplier $\lambda(t,x) \in I\!R^{n\times 1}$ to get a family of unconstrained minimization problems indexed by $t,\ x$,

$$\min_{z,v,\lambda} \left\{ \alpha\pi(z,t|t) + \frac{1}{2}\|v\|^2_{Q(t)} + \lambda'(t,x)\left(x - f(t,z) - g(t,z)v\right) \right\}$$

To get the first order necessary conditions we set to zero the partials of this quantity with respect to z, v, λ .

Constrained Optimiztion

The prediction step requires the solution of a family of constrained optimization problems indexed by the variables t, x of the terminal constraint $z(t) = x$.

To get unconstrained problems we add the constraint to the criterion with a Lagrange multiplier $\lambda(t,x) \in \mathbb{R}^{n \times 1}$ to get a family of unconstrained minimization problems indexed by t, x ,

$$\min_{z,v,\lambda} \left\{ \alpha\pi(z,t|t) + \frac{1}{2}\|v\|^2_{Q(t)} + \lambda'(t,x)\left(x - f(t,z) - g(t,z)v\right) \right\}$$

To get the first order necessary conditions we set to zero the partials of this quantity with respect to z, v, λ .

To simplify the presentation we shall assume $g(t,x) = G(t)x$. The general case follows in a similar fashion.

Necessary Conditions

Setting the partial of the augmented criterion with respect to $\boldsymbol{\lambda}$ to zero yields the constraint

$$\boldsymbol{x} = \boldsymbol{f}(t, \boldsymbol{z}) + \boldsymbol{G}(t)\boldsymbol{v}$$

Necessary Conditions

Setting the partial of the augmented criterion with respect to $\boldsymbol{\lambda}$ to zero yields the constraint

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{t}, \boldsymbol{z}) + \boldsymbol{G}(\boldsymbol{t})\boldsymbol{v}$$

Setting the partial of the augmented criterion with respect to $\boldsymbol{v}$ to zero yields

$$\boldsymbol{0} \;=\; \boldsymbol{v}'\boldsymbol{Q}(\boldsymbol{t}) - \boldsymbol{\lambda}'(\boldsymbol{t}, \boldsymbol{x})\boldsymbol{G}(\boldsymbol{t})$$

Necessary Conditions

Setting the partial of the augmented criterion with respect to λ to zero yields the constraint

$$x = f(t,z) + G(t)v$$

Setting the partial of the augmented criterion with respect to v to zero yields

$$0 = v'Q(t) - \lambda'(t,x)G(t)$$

Because $Q(t)$ is assumed to be invertible this can be solved for $v(t,x)$ as a function of $\lambda(t,x)$

$$v(t,x) = Q^{-1}(t)G'(t)\lambda(t,x)$$

Necessary Conditions

Setting the partial of the augmented criterion with respect to λ to zero yields the constraint

$$x = f(t, z) + G(t)v$$

Setting the partial of the augmented criterion with respect to v to zero yields

$$0 = v'Q(t) - \lambda'(t, x)G(t)$$

Because $Q(t)$ is assumed to be invertible this can be solved for $v(t, x)$ as a function of $\lambda(t, x)$

$$v(t, x) = Q^{-1}(t)G'(t)\lambda(t, x)$$

Then the constraint becomes

$$x = f(t, z) + \bar{Q}(t)\lambda(t, x)$$

where

$$\bar{Q}(t) = G(t)Q^{-1}(t)G'(t)$$

Necessary Conditions

Setting the partial of of the augmented criterion with respect to z to zero yields

$$0 = \alpha \frac{\partial \pi}{\partial z}(z(t,x), t|t) - \lambda'(t,x) \frac{\partial f}{\partial z}(t, z(t,x))$$

Necessary Conditions

Setting the partial of of the augmented criterion with respect to z to zero yields

$$0 = \alpha \frac{\partial \pi}{\partial z}(z(t,x), t|t) - \lambda'(t,x)\frac{\partial f}{\partial z}(t, z(t,x))$$

So at each t, x we must solve this equation and the constraint equation

$$x = f(t, z(t,x)) + \bar{Q}(t)\lambda(t,x)$$

in the two unknowns $z(t,x)$, $\lambda(t,x)$.

Necessary Conditions

Setting the partial of of the augmented criterion with respect to z to zero yields

$$0 = \alpha \frac{\partial \pi}{\partial z}(z(t,x), t|t) - \lambda'(t,x)\frac{\partial f}{\partial z}(t, z(t,x))$$

So at each t, x we must solve this equation and the constraint equation

$$x = f(t, z(t,x)) + \bar{Q}(t)\lambda(t,x)$$

in the two unknowns $z(t,x)$, $\lambda(t,x)$.

Solving these two nonlinear equations is a daunting task so we turn to a series approach.

Series Approach

Expand f and π in series in $\tilde{z} = z - \hat{x}(t|t)$,

$$\begin{aligned} f(t,z) &= f^{[0]}(t) + F(t)\tilde{z} + f^{[2]}(t,\tilde{z}) + \dots \\ \pi(z,t|t) &= \pi^{[0]}(t|t) + \frac{1}{2}\tilde{z}' P(t|t)\tilde{z} + \pi^{[3]}(\tilde{z},t|t) + \dots \end{aligned}$$

where $^{[d]}$ denotes a homogeneous polynomial term of degree d in $\tilde{z}$ with time dependent coefficients.

Series Approach

Expand f and π in series in $\tilde{z} = z - \hat{x}(t|t)$,

$$\begin{aligned} f(t,z) &= f^{[0]}(t) + F(t)\tilde{z} + f^{[2]}(t,\tilde{z}) + \dots \\ \pi(z,t|t) &= \pi^{[0]}(t|t) + \frac{1}{2}\tilde{z}' P(t|t)\tilde{z} + \pi^{[3]}(\tilde{z},t|t) + \dots \end{aligned}$$

where $^{[d]}$ denotes a homogeneous polynomial term of degree d in $\tilde{z}$ with time dependent coefficients.

It is not hard to see that

$$\begin{aligned} \pi^{[0]}(t|t) &= \min_x \pi(x,t|t) \\ \hat{x}(t|t) &= \operatorname{argmin}_x \pi(x,t|t) \\ & \min_x \pi(x,t+1|t) = \alpha \pi^{[0]}(t|t) \\ \hat{x}(t+1|t) &= \operatorname{argmin}_x \pi(x,t+1|t) = f^{[0]}(t) \end{aligned}$$

Series Approach

Expand $\tilde{z}(t,x)$ and $\lambda(t,x)$ in series in $\tilde{x} = x - \hat{x}(t+1|t)$,

$$\begin{aligned}
\tilde{z}(t,x) &= Z(t)\tilde{x} + z^{[2]}(t,\tilde{x}) + z^{[3]}(t,\tilde{x}) + \dots \\
\lambda(t,x) &= \Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x}) + \dots
\end{aligned}$$

The two nonlinear equations become

$$0 = \alpha \frac{\partial \pi}{\partial z}(z(t,x), t|t) - \lambda'(t,x)\frac{\partial f}{\partial z}(t, z(t,x))$$

$$\tilde{x} = f(t, z(t,x)) - f^{[0]}(t) + \bar{Q}(t)\lambda(t,x)$$

Series Approach

Expand $\tilde{z}(t,x)$ and $\lambda(t,x)$ in series in $\tilde{x} = x - \hat{x}(t+1|t)$,

$$\begin{aligned} \tilde{z}(t,x) &= Z(t)\tilde{x} + z^{[2]}(t,\tilde{x}) + z^{[3]}(t,\tilde{x}) + \dots \\ \lambda(t,x) &= \Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x}) + \dots \end{aligned}$$

The two nonlinear equations become

$$0 = \alpha \frac{\partial \pi}{\partial z}(z(t,x), t|t) - \lambda'(t,x)\frac{\partial f}{\partial z}(t, z(t,x))$$

$$\tilde{x} = f(t, z(t,x)) - f^{[0]}(t) + \bar{Q}(t)\lambda(t,x)$$

We collect from these equations the terms linear in $\tilde{x}$ to obtain

$$\begin{aligned} \tilde{x} &= F(t)Z(t)\tilde{x} + \bar{Q}(t)\Lambda(t)\tilde{x} \\ 0 &= \alpha P(t|t)Z(t)\tilde{x} - F'(t)\Lambda(t)\tilde{x} \end{aligned}$$

First Degree Terms

These equations must hold for any $\tilde{x}$ so $Z(t), \Lambda(t)$ must satisfy

$$\begin{bmatrix} I \\ 0 \end{bmatrix} = \mathcal{H}(t) \begin{bmatrix} Z(t) \\ \Lambda(t) \end{bmatrix}$$

where

$$\mathcal{H}(t) = \begin{bmatrix} F(t) & \bar{Q}(t) \\ \alpha P(t|t) & -F'(t) \end{bmatrix}$$

The solvability of these equations depends on the invertibility of $\mathcal{H}(t)$ which we will discuss in a moment.

Second Degree Terms

The terms of second degree in $\tilde{x}$ are

$$\mathcal{H}(t) \begin{bmatrix} z^{[2]}(t,\tilde{x}) \\ \lambda^{[2]}(t,\tilde{x}) \end{bmatrix} = -\begin{bmatrix} k^{[2]}(t,\tilde{x})) \\ l^{[2]}(t,\tilde{x}) \end{bmatrix}$$

Second Degree Terms

The terms of second degree in $\tilde{x}$ are

$$\mathcal{H}(t)\begin{bmatrix} z^{[2]}(t,\tilde{x}) \\ \lambda^{[2]}(t,\tilde{x}) \end{bmatrix} = -\begin{bmatrix} k^{[2]}(t,\tilde{x})) \\ l^{[2]}(t,\tilde{x}) \end{bmatrix}$$

where

$$\begin{aligned} k^{[2]}(t,\tilde{x}) &= f^{[2]}(t,Z(t)\tilde{x}) \\ l^{[2]}(t,\tilde{x}) &= \alpha\left(\frac{\partial \pi^{[3]}}{\partial z}(t,Z(t)\tilde{x})\right)' \\ &\quad -\left(\frac{\partial f^{[2]}}{\partial z}(t,Z(t)\tilde{x})\right)'\Lambda(t)\tilde{x} \end{aligned}$$

Again the solvability of these equations depends on the invertibility of $\mathcal{H}(t)$.

Third Degree Terms

The degree three terms are

$$\mathcal{H}(t)\begin{bmatrix} z^{[3]}(t,\tilde{x}) \\ \lambda^{[3]}(t,\tilde{x}) \end{bmatrix} = -\begin{bmatrix} k^{[3]}(t,\tilde{x}) \\ l^{[3]}(t,\tilde{x}) \end{bmatrix}$$

Third Degree Terms

where

$$k^{[3]}(t,\tilde{x}) = f^{[3]}(t, Z(t)\tilde{x}) + \left(f^{[2]}(t, Z(t)\tilde{x}) + z^{[2]}(t,\tilde{x})\right)^{[2]}$$

$$l^{[3]}(t,\tilde{x}) = \alpha\left(\left(\frac{\partial \pi^{[3]}}{\partial z}(t, Z(t)\tilde{x} + z^{[2]}(t,\tilde{x})\right)^{[3]}\right)'$$

$$+ \quad \alpha\left(\frac{\partial \pi^{[4]}}{\partial z}(t, Z(t)\tilde{x})\right)' - \left(\frac{\partial f^{[2]}}{\partial z}(t, Z(t)\tilde{x})\right)' \lambda^{[2]}(t,\tilde{x})$$

$$- \quad \left(\frac{\partial f^{[2]}}{\partial z}(t, z^{[2]}(t,\tilde{x})\right)' \Lambda(t)\tilde{x} - \left(\frac{\partial f^{[3]}}{\partial z}(t, Z(t)\tilde{x})\right)' \Lambda(t)\tilde{x}$$

and $(\cdot)^{[d]}$ indicates the degree d terms of the enclosed expression.

Third Degree Terms

where

$$\begin{aligned}
k^{[3]}(t,\tilde{x}) &= f^{[3]}(t,Z(t)\tilde{x}) + \left(f^{[2]}(t,Z(t)\tilde{x}) + z^{[2]}(t,\tilde{x})\right)^{[2]} \\
l^{[3]}(t,\tilde{x}) &= \alpha\left(\left(\frac{\partial \pi^{[3]}}{\partial z}(t,Z(t)\tilde{x} + z^{[2]}(t,\tilde{x})\right)^{[3]}\right)' \\
+ \quad & \alpha\left(\frac{\partial \pi^{[4]}}{\partial z}(t,Z(t)\tilde{x})\right)' - \left(\frac{\partial f^{[2]}}{\partial z}(t,Z(t)\tilde{x})\right)' \lambda^{[2]}(t,\tilde{x}) \\
- \quad & \left(\frac{\partial f^{[2]}}{\partial z}(t,z^{[2]}(t,\tilde{x})\right)' \Lambda(t)\tilde{x} - \left(\frac{\partial f^{[3]}}{\partial z}(t,Z(t)\tilde{x})\right)' \Lambda(t)\tilde{x}
\end{aligned}$$

and $(\cdot)^{[d]}$ indicates the degree d terms of the enclosed expression.

The higher degree terms are found in a similar fashion.

Prediction Step

Suppose we have the expansions

$$\begin{aligned}\tilde{z}(t,x) &\approx Z(t)\tilde{x} + z^{[2]}(t,\tilde{x}) + z^{[3]}(t,\tilde{x}) \\ \lambda(t,x) &\approx \Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x})\end{aligned}$$

then

$$\begin{aligned}v(t,x) &= Q^{-1}(t)G'(t)\lambda(t,x) \\ &\approx Q^{-1}(t)G'(t)\left(\Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x})\right)\end{aligned}$$

and

$$\pi(x,t+1|t) \approx \alpha\pi(z(t,x),t|t) + \frac{1}{2}\|v(t,z(t,x),x)\|^2_{Q(t)}$$

$$\hat{x}(t+1|t) \approx \textsf{argmin}_x \pi(x,t+1|t)$$

Prediction Step

Suppose we have the expansions

$$\begin{aligned} \tilde{z}(t,x) &\approx Z(t)\tilde{x} + z^{[2]}(t,\tilde{x}) + z^{[3]}(t,\tilde{x}) \\ \lambda(t,x) &\approx \Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x}) \end{aligned}$$

then

$$\begin{aligned} v(t,x) &= Q^{-1}(t)G'(t)\lambda(t,x) \\ &\approx Q^{-1}(t)G'(t)\left(\Lambda(t)\tilde{x} + \lambda^{[2]}(t,\tilde{x}) + \lambda^{[3]}(t,\tilde{x})\right) \end{aligned}$$

and

$$\pi(x,t+1|t) \approx \alpha\pi(z(t,x),t|t) + \frac{1}{2}\|v(t,z(t,x),x)\|^2_{Q(t)}$$

$$\hat{x}(t+1|t) \approx \text{argmin}_x \pi(x,t+1|t)$$

This completes the prediction step.

Relation to EKF

If $d = 1$ the prediction step from $\hat{x}(t|t)$ to $\hat{x}(t+1|t)$ is the same as that of the Extended Kalman Filter (EKF)

Relation to EKF

If $d = 1$ the prediction step from $\hat{x}(t|t)$ to $\hat{x}(t+1|t)$ is the same as that of the Extended Kalman Filter (EKF)

But the prediction step from $P(t|t)$ to $P(t+1|t)$ is not the same as that of the EKF.

Relation to EKF

If $d = 1$ the prediction step from $\hat{x}(t|t)$ to $\hat{x}(t+1|t)$ is the same as that of the Extended Kalman Filter (EKF)

But the prediction step from $P(t|t)$ to $P(t+1|t)$ is not the same as that of the EKF.

In the EKF, the matrices $P(t|\tau)$ are covariances and the EKF prediction step is

$$P(t+1|t) = F(t)P(t|t)F'(t) + G(t)G'(t)$$

assuming that the driving noise is standard white Gaussian.

Relation to EKF

If $d = 1$ the prediction step from $\hat{x}(t|t)$ to $\hat{x}(t+1|t)$ is the same as that of the Extended Kalman Filter (EKF)

But the prediction step from $P(t|t)$ to $P(t+1|t)$ is not the same as that of the EKF.

In the EKF, the matrices $P(t|\tau)$ are covariances and the EKF prediction step is

$$P(t+1|t) \;=\; F(t)P(t|t)F'(t) + G(t)G'(t)$$

assuming that the driving noise is standard white Gaussian.

In the MEE, the matrices $P(t|\tau)$ are Hessians and loosely speaking the inverses of the covariances. The MEE prediction step is

$$P(t+1|t) \;=\; \alpha Z'(t)P(t|t)Z(t) + \Lambda'(t)\bar{Q}\Lambda(t)$$

Assimilation Step

The assimilation step starts with $y(t)$, $\hat{x}(t|t-1)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t-1)$ expanded in $\tilde{x} = x - \hat{x}(t|t-1)$ and computes $\hat{x}(t|t)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t)$.

Assimilation Step

The assimilation step starts with $y(t)$, $\hat{x}(t|t-1)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t-1)$ expanded in $\tilde{x} = x - \hat{x}(t|t-1)$ and computes $\hat{x}(t|t)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t)$.

To compute $\hat{x}(t|t)$ we assume that the polynomial $\pi(x,t|t)$ is strictly convex in x in a neighborhood around $\hat{x}(t|t-1)$ and that $\hat{x}(t|t)$ is in this neighborhood.

Assimilation Step

The assimilation step starts with $y(t)$, $\hat{x}(t|t-1)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t-1)$ expanded in $\tilde{x} = x - \hat{x}(t|t-1)$ and computes $\hat{x}(t|t)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t)$.

To compute $\hat{x}(t|t)$ we assume that the polynomial $\pi(x,t|t)$ is strictly convex in x in a neighborhood around $\hat{x}(t|t-1)$ and that $\hat{x}(t|t)$ is in this neighborhood.

So we can use Newton's method to minimize

$$\pi(x,t|t) \;\; = \;\; \pi(x,t|t-1) + \frac{1}{2}\|y(t) - h(t,x)\|^2_{R(t)}$$

with initial guess $\hat{x}(t|t-1)$. The argminimum is $\hat{x}(t|t)$.

Assimilation Step

The assimilation step starts with $y(t)$, $\hat{x}(t|t-1)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t-1)$ expanded in $\tilde{x} = x - \hat{x}(t|t-1)$ and computes $\hat{x}(t|t)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t)$.

To compute $\hat{x}(t|t)$ we assume that the polynomial $\pi(x,t|t)$ is strictly convex in x in a neighborhood around $\hat{x}(t|t-1)$ and that $\hat{x}(t|t)$ is in this neighborhood.

So we can use Newton's method to minimize

$$\pi(x,t|t) \;=\; \pi(x,t|t-1) + \frac{1}{2}\|y(t) - h(t,x)\|^2_{R(t)}$$

with initial guess $\hat{x}(t|t-1)$. The argminimum is $\hat{x}(t|t)$.

Then $\pi(x,t|t)$ is transformed into a polynomial of degree $d+1$ in the new $\tilde{x} = x - \hat{x}(t|t)$.

Assimilation Step

The assimilation step starts with $y(t)$, $\hat{x}(t|t-1)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t-1)$ expanded in $\tilde{x} = x - \hat{x}(t|t-1)$ and computes $\hat{x}(t|t)$ and a degree $d+1$ polynomial approximation to $\pi(x,t|t)$.

To compute $\hat{x}(t|t)$ we assume that the polynomial $\pi(x,t|t)$ is strictly convex in x in a neighborhood around $\hat{x}(t|t-1)$ and that $\hat{x}(t|t)$ is in this neighborhood.

So we can use Newton's method to minimize

$$\pi(x,t|t) \;\; = \;\; \pi(x,t|t-1) + \frac{1}{2}\|y(t) - h(t,x)\|^2_{R(t)}$$

with initial guess $\hat{x}(t|t-1)$. The argminimum is $\hat{x}(t|t)$.

Then $\pi(x,t|t)$ is transformed into a polynomial of degree $d+1$ in the new $\tilde{x} = x - \hat{x}(t|t)$.

This completes the assimilation step.

Open Question

But suppose $\pi(x,t|t)$ has several local minima that are close to $\hat{x}(t|t-1)$, what can happen?

Theorem

Suppose that $P(0|0)$ is positive definite and the pair $F(t), G(t)$ is stabilizable at each $t \geq 0$ then $\mathcal{H}(t)$ is invertible for each $t \geq 0$.

Theorem

Suppose that $P(0|0)$ is positive definite and the pair $F(t), G(t)$ is stabilizable at each $t \geq 0$ then $\mathcal{H}(t)$ is invertible for each $t \geq 0$.

Stabilizable is in the continuous time sense. Why?

Example

Here is the Euler discretion of the Van der Pol Oscillator with time step 0.1 second,

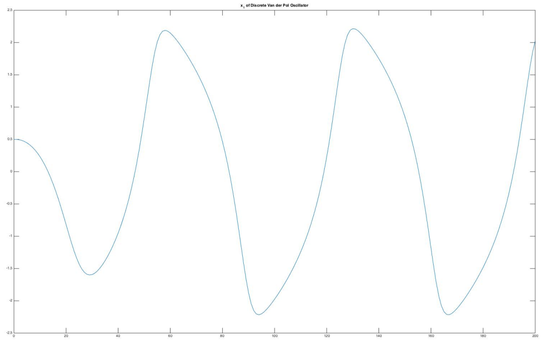
$$
\begin{aligned}
x_1^+ &= x_1 + 0.1x_2 \\
x_2^+ &= -0.1x_1 + x_2 + 0.1(1 - x_1^2)x_2
\end{aligned}
$$

Example

Here is the Euler discretion of the Van der Pol Oscillator with time step 0.1 second,

$$
\begin{aligned}
x_1^+ &= x_1 + 0.1x_2 \\
x_2^+ &= -0.1x_1 + x_2 + 0.1(1 - x_1^2)x_2
\end{aligned}
$$

A typical trajectory of x_1 is below Notice that the period of the limit cycle is approximately seventy time steps.

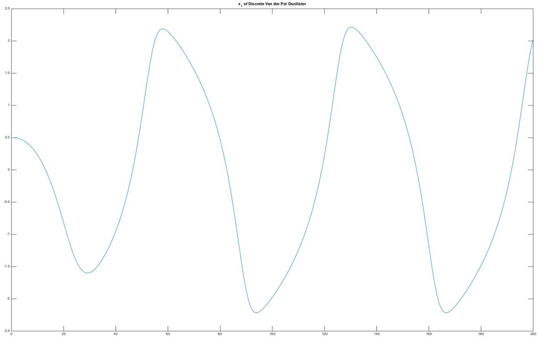


Example

Here is the Euler discretion of the Van der Pol Oscillator with time step 0.1 second,

$$\begin{aligned} x_1^+ &= x_1 + 0.1x_2 \\ x_2^+ &= -0.1x_1 + x_2 + 0.1(1 - x_1^2)x_2 \end{aligned}$$

A typical trajectory of x_1 is below Notice that the period of the limit cycle is approximately seventy time steps.



Example

The corresponding phase portrait is below.

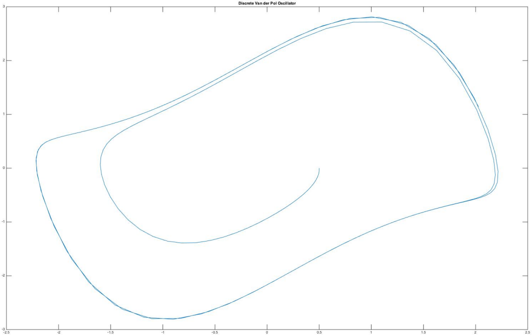


Figure: Typical Phase Portrait

Example

We constructed a MEE filter with

$$
\begin{aligned}
G &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
Q &= 0.1 \\
R &= 0.01 \\
\alpha &= 0.5 \\
x(0) &= \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \\
\hat{x}(0|0) &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\pi(x, 0|-1)) &= \|x - \hat{x}(0|0)\|^2
\end{aligned}
$$

Example

The error $\tilde{x}_1(t|t)$ is shown below and is essentially zero after about twenty time steps.

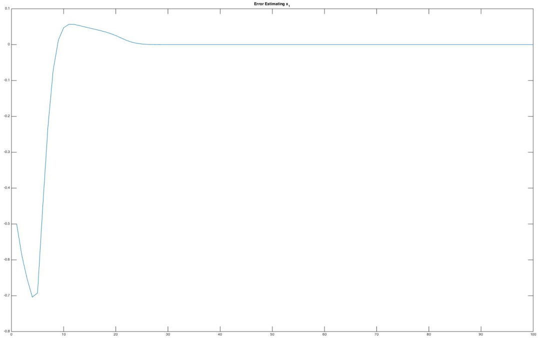


Figure: $\tilde{x}_1(t|t)$

Example

The phase portrait of $x(t)$ is in blue and that of $\hat{x}(t|t)$ is in red. Notice that $\hat{x}(t|t)$ converges to $x(t)$ faster than $x(t)$ converges to the limit cycle.

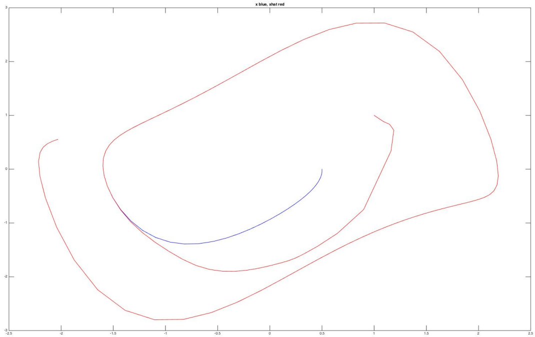


Figure: $\boldsymbol{x(t)}$ blue, $\boldsymbol{\hat{x}(t|t)}$ red

Conclusion

We have extended Minimum Energy Estimation of Mortenson and Hijab to discrete time systems.

Conclusion

We have extended Minimum Energy Estimation of Mortenson and Hijab to discrete time systems.

We showed how a power series method can be employed to ease the computational burden.

Conclusion

We have extended Minimum Energy Estimation of Mortenson and Hijab to discrete time systems.

We showed how a power series method can be employed to ease the computational burden.

We believe that such techniques complement standard Moving Horizon Estimation techniques for real time filtering of faster processes as they essentially reduce the horizon length to one time step.

Conclusion

We have extended Minimum Energy Estimation of Mortenson and Hijab to discrete time systems.

We showed how a power series method can be employed to ease the computational burden.

We believe that such techniques complement standard Moving Horizon Estimation techniques for real time filtering of faster processes as they essentially reduce the horizon length to one time step.

Thank You.

Conclusion

We have extended Minimum Energy Estimation of Mortenson and Hijab to discrete time systems.

We showed how a power series method can be employed to ease the computational burden.

We believe that such techniques complement standard Moving Horizon Estimation techniques for real time filtering of faster processes as they essentially reduce the horizon length to one time step.

Thank You.

Questions?