Rational Minimax Filtering

Arthur J. Krener **Wei Kang**

ajkrener@nps.edu wkang@nps.edu

Research supported in part by AFOSR and NSF

Dedicated to our Esteemed Colleague

Eduardo Sontag

on the occasion of his 60^{th} birthday

Kalman Filtering

We assume that the dynamics and measurement processes can be modeled by a linear system

$$\begin{aligned} \dot{x} &= Ax \\ y &= Cx \end{aligned}$$

The state is $x \in \mathbb{R}^n$, the measurement is $y \in \mathbb{R}^p$ and $p \leq n$.

Kalman Filtering

We assume that the dynamics and measurement processes can be modeled by a linear system

$$\dot{x} = Ax$$
$$y = Cx$$

The state is $x \in \mathbb{R}^n$, the measurement is $y \in \mathbb{R}^p$ and $p \leq n$.

The model is said to be observable (more precisely, reconstructable) if the past measurements $y(s), \ s \leq t$ uniquely determine the current state $x(t)$.

Kalman Filtering

We assume that the dynamics and measurement processes can be modeled by a linear system

$$\begin{aligned} \dot{x} &= Ax \\ y &= Cx \end{aligned}$$

The state is $x \in \mathbb{R}^n$, the measurement is $y \in \mathbb{R}^p$ and $p \leq n$.

The model is said to be observable (more precisely, reconstructable) if the past measurements $y(s),\ s \leq t$ uniquely determine the current state $x(t)$.

One might try to reconstruct the state by differentiating the measurements

$$\begin{aligned} y(t) &= Cx(t) \\ \dot{y}(t) &= CAx(t) \\ \ddot{y}(t) &= CA^2x(t) \\ &\vdots \end{aligned}$$

Kalman Filtering

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

Kalman Filtering

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

But the model is probably not completely accurate.

- The process is not linear.

Kalman Filtering

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

But the model is probably not completely accurate.

- **The process is not linear.**
- **There are unmodeled dynamics.**

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

But the model is probably not completely accurate.

- **The process is not linear.**
- **There are unmodeled dynamics.**
- **There are unknown exogenous inputs affecting the dynamics.**

Kalman Filtering

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

But the model is probably not completely accurate.

- **The process is not linear.**
- **There are unmodeled dynamics.**
- **There are unknown exogenous inputs affecting the dynamics.**
- **There are unknown exogenous noises affecting the measurements.**

Kalman Filtering

If the matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is of full column rank n then the system is observable.

But the model is probably not completely accurate.

- **The process is not linear.**
- **There are unmodeled dynamics.**
- **There are unknown exogenous inputs affecting the dynamics.**
- **There are unknown exogenous noises affecting the measurements.**

Kalman Filtering

To cope with these inaccuracies Kalman added driving and observation noises to the model.

$$
\begin{aligned}
\dot{x} &= Ax + Bv \\
y &= Cx + Dw
\end{aligned}
$$

Kalman Filtering

To cope with these inaccuracies Kalman added driving and observation noises to the model.

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

He assumed that v, w are standard white Gaussian noises (WGN) of dimensions m, p .

Kalman Filtering

To cope with these inaccuracies Kalman added driving and observation noises to the model.

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

He assumed that $v,\ w$ are standard white Gaussian noises (WGN) of dimensions $m,\ p$.

What is standard white Gaussian noise?

Kalman Filtering

To cope with these inaccuracies Kalman added driving and observation noises to the model.

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

He assumed that v, w are standard white Gaussian noises (WGN) of dimensions m, p .

What is standard white Gaussian noise?
It is the formal derivative of a standard Weiner process and is mathematically characterized by the following properties.

Kalman Filtering

To cope with these inaccuracies Kalman added driving and observation noises to the model.

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

He assumed that $v,\ w$ are standard white Gaussian noises (WGN) of dimensions $m,\ p$.

What is standard white Gaussian noise?
It is the formal derivative of a standard Weiner process and is mathematically characterized by the following properties.

If $f(t) \in L^2([t_1, t_2], \mathbb{R}^m)$ then the random variable

$$X = \int_{t_1}^{t_2} f'(t)w(t)\ dt$$

is Gaussian with zero mean and variance

$$\mathbf{E}(X^2) = \int_{t_1}^{t_2} \|f(t)\|^2\ dt$$

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**
- **This requires that there is noise in every measurement so we assume that D is invertible.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**
- **This requires that there is noise in every measurement so we assume that D is invertible.**
- **Because it has a constant power spectrum density at all frequencies.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**
- **This requires that there is noise in every measurement so we assume that D is invertible.**
- **Because it has a constant power spectrum density at all frequencies.**
- **Unfortunately this means that it has infinite power.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**
- **This requires that there is noise in every measurement so we assume that D is invertible.**
- **Because it has a constant power spectrum density at all frequencies.**
- **Unfortunately this means that it has infinite power.**
- **Since we don't know the errors in the dynamics and measurements, modeling them as white is appropriate.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- **Because it is "real".**
- **To keep us from doing something dumb like differentiating the output to reconstruct the state.**
- **This requires that there is noise in every measurement so we assume that D is invertible.**
- **Because it has a constant power spectrum density at all frequencies.**
- **Unfortunately this means that it has infinite power.**
- **Since we don't know the errors in the dynamics and measurements, modeling them as white is appropriate.**
- **This overlooks the fact that we have to choose B, D which fixes the covariances of the errors.**

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- Because it is "real".
- To keep us from doing something dumb like differentiating the output to reconstruct the state.
- This requires that there is noise in every measurement so we assume that D is invertible.
- Because it has a constant power spectrum density at all frequencies.
- Unfortunately this means that it has infinite power.
- Since we don't know the errors in the dynamics and measurements, modeling them as white is appropriate.
- This overlooks the fact that we have to choose B, D which fixes the covariances of the errors.
- Because standard white Gaussian noise is relatively easy to handle mathematically in a linear setting.

Kalman Filtering

Why white Gaussian noise? There are several possible answers.

- Because it is "real".
- To keep us from doing something dumb like differentiating the output to reconstruct the state.
- This requires that there is noise in every measurement so we assume that D is invertible.
- Because it has a constant power spectrum density at all frequencies.
- Unfortunately this means that it has infinite power.
- Since we don't know the errors in the dynamics and measurements, modeling them as white is appropriate.
- This overlooks the fact that we have to choose $B, \ D$ which fixes the covariances of the errors.
- Because standard white Gaussian noise is relatively easy to handle mathematically in a linear setting.

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \le s \le t$

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathbf{E}v(t) \neq 0, \mathbf{E}w(t) \neq 0$

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \le s \le t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**
- **An additional known input.**

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \le s \le t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**
- **An additional known input.**
- **Extended Kalman filters for nonlinear systems.**

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**
- **An additional known input.**
- **Extended Kalman filters for nonlinear systems.**
- **Unscented Kalman filters for nonlinear systems.**

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**
- **An additional known input.**
- **Extended Kalman filters for nonlinear systems.**
- **Unscented Kalman filters for nonlinear systems.**
- **Particle filters for nonlinear systems.**

Kalman Filtering

For simplicity of exposition we are restricting the discussion to continuous time Kalman filtering of a time invariant linear system where the measurements are available over the infinite past.

There are generalizations and extensions to handle the following.

- **Discrete time systems,** $x(t+1) = Ax(t), \ldots$
- **Time varying linear systems,** $A = A(t), C = C(t), \ldots$
- **Finite interval of measurements** $y(s),\ t_0 \leq s \leq t$
- **Partial knowledge of the initial state** $\hat{x}(t_0) \approx N(\hat{x}^0, P^0)$
- **Known bias in the noises,** $\mathrm{E}v(t) \neq 0, \mathrm{E}w(t) \neq 0$
- **Correlation between the noises.**
- **An additional known input.**
- **Extended Kalman filters for nonlinear systems.**
- **Unscented Kalman filters for nonlinear systems.**
- **Particle filters for nonlinear systems.**

Derivation of the Kalman Filter

$$\begin{aligned}\dot{\boldsymbol{x}} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{v}\\ \boldsymbol{y} &= \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{w}\end{aligned}$$

Derivation of the Kalman Filter

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

We assume that the filter for $x_i(t)$ is a weighted sum of the past observations. The estimate is

$$\hat{x}_i(t) = \int_0^\infty k(s) y(t-s)\, ds$$

Derivation of the Kalman Filter

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

We assume that the filter for $x_i(t)$ is a weighted sum of the past observations. The estimate is

$$\hat{x}_i(t) = \int_0^\infty k(s)y(t-s)\, ds$$

We wish to choose the weighing pattern $k(s) \in \mathbb{R}^{1\times p}$ to minimize $\mathbf{E}(\tilde{x}_i(t))^2$ where $\tilde{x}_i(t) = x_i(t) - \hat{x}_i(t)$.

Derivation of the Kalman Filter

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx + Dw \end{aligned}$$

We assume that the filter for $x_i(t)$ is a weighted sum of the past observations. The estimate is

$$\hat{x}_i(t) = \int_0^\infty k(s) y(t-s)\, ds$$

We wish to choose the weighing pattern $k(s) \in I\!R^{1\times p}$ to minimize $\mathrm{E}(\tilde{x}_i(t))^2$ where $\tilde{x}_i(t) = x_i(t) - \hat{x}_i(t)$.
Given a $k(s)$ define $h(s) \in I\!R^{1\times n}$ by

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -e^i \end{aligned}$$

where e^i is the i^{th} unit row vector.

Derivation of the Kalman Filter

$$\begin{aligned}
\hat{x}_i(t) &= \int_0^\infty k(s)y(t-s)\ ds \\
&= \int_0^\infty k(s)Cx(t-s) + k(s)Dw(t-s)\ ds \\
&= \int_0^\infty \left(\dot{h}(s) - h(s)A\right)x(t-s) + k(s)Dw(t-s)\ ds \\
&= [h(s)x(t-s)]_0^\infty + \int_0^\infty h(s)Bv(t-s) + k(s)Dw(t-s)\ d
\end{aligned}$$

We assume that $h(\infty) = 0$ so

$$\begin{aligned}
\tilde{x}_i(t) &= -\int_0^\infty h(s)Bv(t-s) + k(s)Dw(t-s)\ ds \\
\mathbf{E}(\tilde{x}_i(t))^2 &= \int_0^\infty h(s)BB'h'(s) + k(s)DD'k'(s)\ ds
\end{aligned}$$

Linear Quadratic Regulator

So we have the optimal control problem of minimizing by choice of $k(s)$

$$\int_0^\infty h(s)BB'h'(s) + k(s)DD'k'(s)\ ds$$

subject to

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= h^0 \end{aligned}$$

Linear Quadratic Regulator

So we have the optimal control problem of minimizing by choice of $k(s)$

$$\int_0^\infty h(s)BB'h'(s) + k(s)DD'k'(s)\ ds$$

subject to

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= h^0 \end{aligned}$$

We assume that the minimum is a quadratic form in h^0

$$h^0 P (h^0)' = \min_k \int_0^\infty h(s)BB'h'(s) + k(s)DD'k'(s)\ ds$$

Completing the Square

$$h^0 P (h^0)' = \min_k \int_0^\infty hBB'h' + kDD'k' \, ds$$

Completing the Square

$$h^0 P (h^0)' = \min_k \int_0^\infty hBB'h' + kDD'k' \; ds$$

$$\begin{aligned} \left[h(s) P h'(s) \right]_0^\infty &= \int_0^\infty \frac{d}{ds} h(s) P h'(s) \; ds \\ h^0 P (h^0)' &= -\int_0^\infty (hA + kC) \, Ph' + hP \, (hA + kC)' \; ds \end{aligned}$$

Completing the Square

$$h^0 P(h^0)' = \min_k \int_0^\infty hBB'h' + kDD'k' \, ds$$

$$\left[h(s)Ph'(s)\right]_0^\infty = \int_0^\infty \frac{d}{ds} h(s)Ph'(s) \, ds$$

$$h^0 P(h^0)' = -\int_0^\infty (hA + kC)\, Ph' + hP\,(hA + kC)' \, ds$$

Subtracting

$$0 = \min_k \int_0^\infty [h, k] \begin{bmatrix} AP + PA' + BB' & PC' \\ CP & DD' \end{bmatrix} [h, k]' \, ds$$

Completing the Square

If

$$\begin{aligned} 0 &= AP + PA' + BB' - PC'(DD')^{-1}CP \\ G &= PC'(DD')^{-1} \end{aligned}$$

then the above reduces to a perfect square

$$0 = \min_k \int_0^\infty (k + hG)DD'(k + hG)' \, ds$$

so the optimal $k = -hG$ **.**

Kalman Filtering

To filter all states at once we let $H(s) \in \mathbb{R}^{n \times n}$ satisfy

$$\begin{aligned} \dot{H} &= H(A - GC) \\ H(0) &= -I \end{aligned}$$

and $K(s) = H(s)G$ then

$$\dot{H} = (A - GC)H$$

Kalman Filtering

To filter all states at once we let $H(s) \in \mathbb{R}^{n\times n}$ satisfy

$$\begin{aligned} \dot{H} &= H(A-GC) \\ H(0) &= -I \end{aligned}$$

and $K(s) = H(s)G$ then

$$\dot{H} = (A-GC)H$$

$$\begin{aligned} \hat{x}(t) &= \int_0^\infty K(s)y(t-s)\ ds \\ &= -\int_{-\infty}^t H(t-s)Gy(s)\ ds \end{aligned}$$

Kalman Filtering

To filter all states at once we let $H(s) \in \mathbb{R}^{n\times n}$ satisfy

$$\begin{aligned} \dot{H} &= H(A - GC) \\ H(0) &= -I \end{aligned}$$

and $K(s) = H(s)G$ then

$$\dot{H} = (A - GC)H$$

$$\begin{aligned} \hat{x}(t) &= \int_0^\infty K(s)y(t-s)\ ds \\ &= -\int_{-\infty}^t H(t-s)Gy(s)\ ds \end{aligned}$$

$$\frac{d}{dt}\hat{x}(t) = (A - GC)\hat{x}(t) + Gy(t)$$

Kalman Filtering

Kalman Filter

$$\frac{d}{dt}\hat{x}(t) = (A - GC)\hat{x}(t) + Gy(t)$$

Kalman Filtering

Kalman Filter

$$\frac{d}{dt}\hat{x}(t) \quad = \quad (A - GC)\hat{x}(t) + Gy(t)$$

Riccati equation

$$0 \quad = \quad AP + PA' + BB' - PC'(DD')^{-1}CP$$

Kalman Filtering

Kalman Filter

$$\frac{d}{dt}\hat{x}(t) = (A - GC)\hat{x}(t) + Gy(t)$$

Riccati equation

$$0 = AP + PA' + BB' - PC'(DD')^{-1}CP$$

Filter Gain

$$G = PC'(DD')^{-1}$$

Kalman Filtering

Kalman Filter

$$\frac{d}{dt}\hat{x}(t) \;=\; (A - GC)\hat{x}(t) + Gy(t)$$

Riccati equation

$$0 \;=\; AP + PA' + BB' - PC'(DD')^{-1}CP$$

Filter Gain

$$G \;=\; PC'(DD')^{-1}$$

This derivation is easily extended to discrete time, time varying and/or finite horizon linear systems.

Johansen and Berkovitz-Pollard Problem

Independently Johansen (1966) and Berkovitz-Pollard (1967) considered the following filtering problem.

$$
\begin{aligned}
\ddot{x} &= u, && |u| \leq 1 \\
y &= x + w, && w \ \text{WGN}
\end{aligned}
$$

Johansen and Berkovitz-Pollard Problem

Independently Johansen (1966) and Berkovitz-Pollard (1967) considered the following filtering problem.

$$\begin{aligned} \ddot{x} &= u, & |u| \leq 1 \\ y &= x + w, & w \ \text{WGN} \end{aligned}$$

They assumed a linear filter

$$\hat{x}(t) = \int_0^\infty k(s) y(t-s)\, ds$$

Johansen and Berkovitz-Pollard Problem

Independently Johansen (1966) and Berkovitz-Pollard (1967) considered the following filtering problem.

$$\begin{aligned} \ddot{x} &= u, \qquad |u| \le 1 \\ y &= x + w, \qquad w \ \textbf{WGN} \end{aligned}$$

They assumed a linear filter

$$\hat{x}(t) = \int_0^\infty k(s)y(t-s)\, ds$$

where the weighing pattern $k(s)$ is chosen to

$$\min_k \max_{|u|\le 1} E_w(\tilde{x}(t))^2$$

Johansen and Berkovitz-Pollard Problem

Given a $k(s)$ define $h(s)$ by

$$\begin{aligned} \ddot{h} &= k \\ h(0) &= -1 \\ \dot{h}(0) &= 0 \end{aligned}$$

Johansen and Berkovitz-Pollard Problem

Given a $k(s)$ define $h(s)$ by

$$\begin{aligned}
\ddot{h} &= k \\
h(0) &= -1 \\
\dot{h}(0) &= 0
\end{aligned}$$

$$\begin{aligned}
\hat{x}(t) &= \int_0^\infty k(s)y(t-s)\ ds \\
&= \int_0^\infty \ddot{h}(s)x(t-s) + k(s)w(t-s)\ ds \\
&= x(t) + \int_0^\infty h(s)u(t-s) + k(s)w(t-s)\ ds \\
\tilde{x}(t) &= -\int_0^\infty h(s)u(t-s) + k(s)w(t-s)\ ds
\end{aligned}$$

Johansen and Berkovitz-Pollard Problem

Then

$$E_w(\tilde{x}(t))^2 = \left(\int_0^\infty h(s)u(s)\ ds\right)^2 + \int_0^\infty (k(s))^2\ ds$$

and we have a differential game.

Johansen and Berkovitz-Pollard Problem

Then

$$E_w(\tilde{x}(t))^2 \;=\; \left(\int_0^\infty h(s)u(s)\ ds\right)^2 + \int_0^\infty (k(s))^2\ ds$$

and we have a differential game.
Our adversary wishes to choose $u(s)$ to maximize this quantity subject to $|u(s)| \leq 1$.

Johansen and Berkovitz-Pollard Problem

Then

$$E_w(\tilde{x}(t))^2 = \left(\int_0^\infty h(s)u(s)\, ds\right)^2 + \int_0^\infty (k(s))^2\, ds$$

and we have a differential game.
Our adversary wishes to choose $u(s)$ to maximize this quantity subject to $|u(s)| \le 1$.
We wish to choose $k(s), h(s)$ to minimize this maximum subject to

$$\begin{aligned} \ddot{h} &= k \\ h(0) &= -1 \\ \dot{h}(0) &= 0 \end{aligned}$$

Johansen and Berkovitz-Pollard Problem

Then

$$E_w(\tilde{x}(t))^2 = \left(\int_0^\infty h(s)u(s)\,ds\right)^2 + \int_0^\infty (k(s))^2\,ds$$

and we have a differential game.
Our adversary wishes to choose $u(s)$ to maximize this quantity subject to $|u(s)| \leq 1$.
We wish to choose $k(s), h(s)$ to minimize this maximum subject to

$$\begin{aligned}\ddot{h} &= k \\ h(0) &= -1 \\ \dot{h}(0) &= 0\end{aligned}$$

Clearly for a given $k(s), h(s)$, the maximizing $u(s)$ are

$$u(s) = \pm\,\mathbf{sign}(h(s))$$

Johansen and Berkovitz-Pollard Problem

So

$$\max_{|u|\leq 1} E_w(\tilde{x}(t))^2 \;=\; \left(\int_0^\infty |h(s)|\; ds\right)^2 + \int_0^\infty (k(s))^2\; ds$$

Johansen and Berkovitz-Pollard Problem

So

$$\max_{|u|\leq 1} E_w(\tilde{x}(t))^2 = \left(\int_0^\infty |h(s)|\ ds\right)^2 + \int_0^\infty (k(s))^2\ ds$$

The differential game reduces to a non standard optimal control of choosing $k(s), h(s)$ to minimize this quantity subject to

$$\begin{aligned} \ddot{h} &= k \\ h(0) &= -1 \\ \dot{h}(0) &= 0 \end{aligned}$$

Johansen and Berkovitz-Pollard Problem

So

$$\max_{|u|\leq 1} E_w(\tilde{x}(t))^2 = \left(\int_0^\infty |h(s)|\, ds\right)^2 + \int_0^\infty (k(s))^2\, ds$$

The differential game reduces to a non standard optimal control of choosing $k(s), h(s)$ to minimize this quantity subject to

$$\begin{aligned} \ddot{h} &= k \\ h(0) &= -1 \\ \dot{h}(0) &= 0 \end{aligned}$$

The Euler Lagrange equation for this problem is

$$h^{(4)} = -\gamma\, \mathsf{sign}(h)$$

where

$$\gamma = \int_0^\infty |h(s)|\, ds$$

Johansen and Berkovitz-Pollard Problem

Consider the related differential equation

$$\phi^{(4)} = -\,\mathbf{sign}(\phi)$$

Johansen and Berkovitz-Pollard Problem

Consider the related differential equation

$$\phi^{(4)} = -\operatorname{sign}(\phi)$$

Two one parameter groups act on the space of solutions of this equation.

$$\begin{aligned} \phi(s) &\to \phi(s+\sigma), & \sigma &\in \mathbb{R} \\ \phi(s) &\to \alpha^4 \phi(s/\alpha), & \alpha &\in \mathbb{R}_{>0} \end{aligned}$$

Johansen and Berkovitz-Pollard Problem

Consider the related differential equation

$$\phi^{(4)} = -\,\mathrm{sign}(\phi)$$

Two one parameter groups act on the space of solutions of this equation.

$$\phi(s) \to \phi(s+\sigma), \qquad \sigma \in \mathbb{R}$$
$$\phi(s) \to \alpha^4 \phi(s/\alpha), \qquad \alpha \in \mathbb{R}_{>0}$$

We look for a self similar solution that has consecutive simple zeros at $s = 0,\ s = 1$ and satisfies for $s \in [0, \alpha]$

$$\phi(s+1) = -\alpha^4 \phi(s/\alpha)$$

Johansen and Berkovitz-Pollard Problem

Consider the related differential equation

$$\phi^{(4)} = -\,\mathrm{sign}(\phi)$$

Two one parameter groups act on the space of solutions of this equation.

$$\phi(s) \rightarrow \phi(s+\sigma), \qquad \sigma \in \mathbb{R}$$
$$\phi(s) \rightarrow \alpha^4 \phi(s/\alpha), \qquad \alpha \in \mathbb{R}_{>0}$$

We look for a self similar solution that has consecutive simple zeros at $s = 0,\ s = 1$ and satisfies for $s \in [0, \alpha]$

$$\phi(s+1) = -\alpha^4 \phi(s/\alpha)$$

On $s \in [0, 1]$

$$\phi(s) = c_1 s + c_2 s^2/2 + c_3 s^3/6 + c_4 s^4/24$$

where $c_4 = -\,\mathrm{sign}(c_1) \neq 0$

Johansen and Berkovitz-Pollard Problem

Matching $\phi(s)$ and its first three derivatives at $s = 1^{\pm}$ we obtain

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1/2! & 1/3! & 1/4! \\ 1+\alpha^3 & 1 & 1/2! & 1/3! \\ 0 & 1+\alpha^2 & 1 & 1/2! \\ 0 & 0 & 1+\alpha & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

so the determinant of this matrix must be zero.

Johansen and Berkovitz-Pollard Problem

Matching $\phi(s)$ and its first three derivatives at $s = 1^{\pm}$ we obtain

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1/2! & 1/3! & 1/4! \\ 1+\alpha^3 & 1 & 1/2! & 1/3! \\ 0 & 1+\alpha^2 & 1 & 1/2! \\ 0 & 0 & 1+\alpha & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

so the determinant of this matrix must be zero.
The determinant is

$$p(s) = (-\alpha^6 + 3\alpha^5 + 5\alpha^4 - 5\alpha^2 - 3\alpha + 1)/24$$

and it has three positive roots

$$\alpha = \begin{cases} 0.2421 \\ 1 \\ 1/0.2421 \end{cases}$$

Johansen and Berkovitz-Pollard Problem

Matching $\phi(s)$ and its first three derivatives at $s = 1^{\pm}$ we obtain

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1/2! & 1/3! & 1/4! \\ 1+\alpha^3 & 1 & 1/2! & 1/3! \\ 0 & 1+\alpha^2 & 1 & 1/2! \\ 0 & 0 & 1+\alpha & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

so the determinant of this matrix must be zero.
The determinant is

$$p(s) = (-\alpha^6 + 3\alpha^5 + 5\alpha^4 - 5\alpha^2 - 3\alpha + 1)/24$$

and it has three positive roots

$$\alpha = \begin{cases} 0.2421 \\ 1 \\ 1/0.2421 \end{cases}$$

The first and third roots yield self similar solutions to $\phi^{(4)} = -\,\mathrm{sign}(\phi)$ while the second root yields a periodic solution to $\phi^{(4)} = \mathrm{sign}(\phi)$.

Johansen and Berkovitz-Pollard Problem

We choose the first root because that solution chatters to zero at $s = 1/(1-\alpha) = 1.3194$.

Johansen and Berkovitz-Pollard Problem

We choose the first root because that solution chatters to zero at $s = 1/(1-\alpha) = 1.3194$.

Then

$$h(s) \;=\; \gamma\beta^4\phi(s/\beta)$$

where β is chosen so that

$$1 \;=\; \int_0^\infty |\beta^4\phi(s/\beta)|\, ds$$

Then γ is chosen so that

$$h(0) \;=\; -1$$

Johansen and Berkovitz-Pollard Problem

For $s \in [0, \beta]$

$$\begin{aligned} h(s) &= -s + 0.872575492926169s^2 - 0.253795996951782s^3 \\ &\quad +0.024616157365051s^4 \\ k(s) &= 1.745150985852338 - 1.522775981710693s \\ &\quad +0.295393888380611s^2 \end{aligned}$$

and it chatters to zero at $\beta/(1-\alpha) = 4.2244$ **.**

Johansen and Berkovitz-Pollard Problem

For $s \in [0, \beta]$

$$
\begin{aligned}
h(s) &= -s + 0.872575492926169s^2 - 0.253795996951782s^3 \\
&\quad +0.024616157365051s^4 \\
k(s) &= 1.745150985852338 - 1.522775981710693s \\
&\quad +0.295393888380611s^2
\end{aligned}
$$

and it chatters to zero at $\beta/(1-\alpha) = 4.2244$ **.**
Integration by parts yields the minmax expected error variance

$$
\ddot{h}(0) = k(0) = 1.745150985852338
$$

Johansen and Berkovitz-Pollard Problem

For $s \in [0, \beta]$

$$
\begin{aligned}
h(s) &= -s + 0.872575492926169s^2 - 0.253795996951782s^3 \\
&\quad +0.024616157365051s^4 \\
k(s) &= 1.745150985852338 - 1.522775981710693s \\
&\quad +0.295393888380611s^2
\end{aligned}
$$

and it chatters to zero at $\beta/(1-\alpha) = 4.2244$.
Integration by parts yields the minmax expected error variance

$$
\ddot{h}(0) = k(0) = 1.745150985852338
$$

The problem is that the resulting filter is infinite dimensional as it requires storing the values of $y(t-s)$ for $s \in [0, 4.2244]$.

Johansen and Berkovitz-Pollard Problem

For $s \in [0, \beta]$

$$\begin{aligned} h(s) &= -s + 0.872575492926169s^2 - 0.253795996951782s^3 \\ &\quad +0.024616157365051s^4 \\ k(s) &= 1.745150985852338 - 1.522775981710693s \\ &\quad +0.295393888380611s^2 \end{aligned}$$

and it chatters to zero at $\beta/(1-\alpha) = 4.2244$.
Integration by parts yields the minmax expected error variance

$$\ddot{h}(0) = k(0) = 1.745150985852338$$

The problem is that the resulting filter is infinite dimensional as it requires storing the values of $y(t-s)$ for $s \in [0, 4.2244]$.

And what about a general linear system?

Linear Time Invariant Minimax Filtering

Plant:

$$\begin{aligned} \dot{x} &= Ax + Bu, & \|u\|_\infty &\le 1 \\ y &= Cx + Dw, & w &\ \textbf{WGN} \\ z &= Lx, & z &\in \mathbb{R} \end{aligned}$$

Linear Time Invariant Minimax Filtering

Plant:

$$\dot{x} = Ax + Bu, \qquad \|u\|_\infty \leq 1$$
$$y = Cx + Dw, \qquad w \text{ WGN}$$
$$z = Lx, \qquad z \in \mathbb{R}$$

Linear Filter:

$$\hat{z} \; = \; \int_0^\infty k(s) y(t-s) \; ds$$

Goal:

$$\min_k \max_{\|u\|_\infty \leq 1} \mathsf{E}_w (\tilde{x}_i)^2$$

Linear Time Invariant Minimax Filtering

Given a $k(s)$ define $h(s)$ as before

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -L \end{aligned}$$

Linear Time Invariant Minimax Filtering

Given a $k(s)$ define $h(s)$ as before

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -L \end{aligned}$$

After integration by parts

$$\begin{aligned} \tilde{z}(t) &= \int_0^\infty h(s)Bu(t-s) + k(s)Dw(t-s)\ ds \\ \mathsf{E}_w(\tilde{z}(t))^2 &= \left(\int_0^\infty h(s)Bu(t-s)\ ds\right)^2 \\ &\quad + \int_0^\infty k(s)DD'k'(s)\ ds \\ \max_{\|u\|_\infty \le 1} \mathsf{E}_w(\tilde{z}(t))^2 &= \left(\int_0^\infty \|h(s)B\|_1\ ds\right)^2 \\ &\quad + \int_0^\infty k(s)DD'k'(s)\ ds \end{aligned}$$

Non Standard Optimal Control Problem

Minimize

$$\left(\int_0^\infty \|h(s)B\|_1 \; ds\right)^2 + \int_0^\infty k(s)DD'k'(s) \; ds$$

subject to

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -L \end{aligned}$$

- **State** $h(s) \in \mathbb{R}^{1\times n}$ **,** **Control** $k(s) \in \mathbb{R}^{1\times p}$

Non Standard Optimal Control Problem

Minimize

$$\left(\int_0^\infty \|h(s)B\|_1 \; ds \right)^2 + \int_0^\infty k(s)DD'k'(s) \; ds$$

subject to

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -L \end{aligned}$$

State $h(s) \in I\!R^{1\times n}$ **,** **Control** $k(s) \in I\!R^{1\times p}$

This optimization problem is too complicated for the Euler-Lagrange approach so we apply the Pontryagin Maximum Principle instead.

Pontryagin Maximum Principle

Add an extra state coordinate

$$\dot{h}_{n+1} = \|hB\|_1$$

Pontryagin Maximum Principle

Add an extra state coordinate

$$\dot{h}_{n+1} \;=\; \|hB\|_1$$

Adjoint variables $\xi \in \mathbb{R}^{n\times 1}$ **,** $\zeta \in \mathbb{R}$ **.**

Pontryagin Maximum Principle

Add an extra state coordinate

$$\dot{h}_{n+1} \;=\; \|hB\|_1$$

Adjoint variables $\xi \in I\!R^{n\times 1}$ **,** $\zeta \in I\!R$ **.**

Control Hamiltonian

$$\mathcal{H} \;=\; hA\xi + kC\xi + \|hB\|_1\zeta + kDD'k$$

Adjoint Dynamics

$$\begin{aligned}\dot{\xi} \;&=\; -\left(\frac{\partial \mathcal{H}}{\partial h}\right)' = -A\xi - B\,(\ \mathrm{sign}(hB))'\,\zeta \\ \dot{\zeta} \;&=\; -\left(\frac{\partial \mathcal{H}}{\partial h_{n+1}}\right) = 0\end{aligned}$$

Pontryagin Maximum Principle

Add an extra state coordinate

$$\dot{h}_{n+1} = \|hB\|_1$$

Adjoint variables $\xi \in \mathbb{R}^{n\times 1}$ **,** $\zeta \in \mathbb{R}$ **.**

Control Hamiltonian

$$\mathcal{H} = hA\xi + kC\xi + \|hB\|_1\zeta + kDD'k$$

Adjoint Dynamics

$$\dot{\xi} = -\left(\frac{\partial \mathcal{H}}{\partial h}\right)' = -A\xi - B\left(\ \mathbf{sign}(hB)\right)'\zeta$$

$$\dot{\zeta} = -\left(\frac{\partial \mathcal{H}}{\partial h_{n+1}}\right) = 0$$

Pontryagin Maximum Principle

Maximize the Hamiltonian with respect to the control

$$
\begin{aligned}
0 &= \frac{\partial \mathcal{H}}{\partial k} = C\xi + 2DD'k' \\
k &= -\frac{\xi' C' (DD')^{-1}}{2}
\end{aligned}
$$

and plug into the dynamics.

Pontryagin Maximum Principle

Hamiltonian Dynamics and Transversality Conditions

$$
\begin{aligned}
\dot{h} &= hA - \frac{\xi' C' (DD')^{-1} C}{2} \\
\dot{h}_{n+1} &= \|hB\|_1 \\
\dot{\xi} &= -A\xi - B\,(\,\mathsf{sign}(hB))'\,\zeta \\
\dot{\zeta} &= -2\|hB\|_1
\end{aligned}
$$

$$
\begin{aligned}
h(0) &= -L \\
h_{n+1}(0) &= 0 \\
\xi(\infty) &= 0 \\
\zeta(\infty) &= 0
\end{aligned}
$$

Pontryagin Maximum Principle

Hamiltonian Dynamics and Transversality Conditions

$$
\begin{aligned}
\dot{h} &= hA - \frac{\xi' C' (DD')^{-1} C}{2} \\
\dot{h}_{n+1} &= \|hB\|_1 \\
\dot{\xi} &= -A\xi - B\,(\,\mathrm{sign}(hB))'\,\zeta \\
\dot{\zeta} &= -2\|hB\|_1
\end{aligned}
$$

$$
\begin{aligned}
h(0) &= -L \\
h_{n+1}(0) &= 0 \\
\xi(\infty) &= 0 \\
\zeta(\infty) &= 0
\end{aligned}
$$

This is usually too complicated to solve explicitly and even if we could the resulting filter would probably be infinite dimensional.

Rational Minimax Filtering

Therefore we restrict the optimization to weighing patterns $k(s)$ that are the impulse responses of finite dimensional linear systems.

Rational Minimax Filtering

Therefore we restrict the optimization to weighing patterns $k(s)$ that are the impulse responses of finite dimensional linear systems.

In other words we restrict to $k(s)$ whose Laplace transforms are rational.

$$k(s) = \sum_{i=1}^{N} \gamma_i e^{\lambda_i s}$$

Rational Minimax Filtering

Therefore we restrict the optimization to weighing patterns $k(s)$ that are the impulse responses of finite dimensional linear systems.

In other words we restrict to $k(s)$ whose Laplace transforms are rational.

$$k(s) \;=\; \sum_{i=1}^{N} \gamma_i e^{\lambda_i s}$$

This guarantees that the resulting filter is finite dimensional, it can be realized by a finite dimensional time invariant linear system.

Rational Minimax Filtering

$$\hat{z}(t) = \int_0^\infty k(s)y(t-s)\,ds$$
$$k(s) = \sum_{i=1}^{N} \gamma_i e^{\lambda_i s}$$

is realized by

$$\dot{\xi} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix} \xi + \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} y$$
$$\hat{z}(t) = \begin{bmatrix} \gamma_1 & \dots & \gamma_N \end{bmatrix} \xi$$

Rational Minimax Filtering

If we look for a filter the same size as the original system $N = n$, A, B is a controllable pair and all the eigenvalues of A are in the closed right half plane then the filter takes the form

$$k(s) = -h(s)G$$

for some G .

Rational Minimax Filtering

If we look for a filter the same size as the original system $N = n$, A, B is a controllable pair and all the eigenvalues of A are in the closed right half plane then the filter takes the form

$$k(s) = -h(s)G$$

for some G .

In other words we are finding the linear feedback that

$$\min_G \left(\int_0^\infty \|h(s)B\|_1 \; ds \right)^2 + \int_0^\infty k(s)DD'k'(s) \; ds$$

subject to

$$\begin{aligned} \dot{h} &= hA + kC \\ h(0) &= -L \\ k(s) &= -h(s)G \end{aligned}$$

Rational Minimax Filtering

One virtue of this approach is that the resulting filter is realized by the linear system

$$\begin{aligned} \dot{\xi} &= (A - GC)\xi + Gy = A\xi + G(y - C\xi) \\ \hat{z} &= L\xi \end{aligned}$$

and it looks like a Kalman filter or linear observer.

Rational Minimax Filtering

One virtue of this approach is that the resulting filter is realized by the linear system

$$\begin{aligned} \dot{\xi} &= (A - GC)\xi + Gy = A\xi + G(y - C\xi) \\ \hat{z} &= L\xi \end{aligned}$$

and it looks like a Kalman filter or linear observer.

Notice that there may be a different gain G and different filter for each linear functional of the state $z = Lx$.

One virtue of this approach is that the resulting filter is realized by the linear system

$$\begin{aligned} \dot{\xi} &= (A - GC)\xi + Gy = A\xi + G(y - C\xi) \\ \hat{z} &= L\xi \end{aligned}$$

and it looks like a Kalman filter or linear observer.

Notice that there may be a different gain G and different filter for each linear functional of the state $z = Lx$.

This suggests the following approach. Use numerical routines to minimize the optimal control problem with and without the restriction that $k(s) = h(s)G$. If the optimal cost of the former is close enough to that of the latter, accept the filter. If not expand the class of rational filters that are considered.

Single Integrator

We tried this approach on some model problems.

Single Integrator

We tried this approach on some model problems.

$$\begin{array}{ll} A = 0 & B = 1 \\ C = 1 & D = 1 \end{array}$$

$$z = x$$

Optimal Cost	Suboptimal Rational Cost	Ratio
1.1006	1.1906	1.0818

We were able to compute the optimal infinite dimensional filter explicitly.
The suboptimal filter was computed using a numerical optimization routine.

Double Integrator

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 \end{bmatrix}$$

(JBP Problem) $$z = x_1$$

Optimal Cost	Suboptimal Rational Cost	Ratio
1.7452	1.7880	1.0245

Double Integrator

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 \end{bmatrix}$$

(JBP Problem)

$$z = x_1$$

Optimal Cost	Suboptimal Rational Cost	Ratio
1.7452	1.7880	1.0245

$$z = x_2$$

Optimal Cost	Suboptimal Rational Cost	Ratio
2.1269	2.2733	1.0688

Again we were able to compute the optimal infinite dimensional filters explicitly.
The suboptimal filters were computed using a numerical optimization routine.

Triple Integrator

Estimate x_1

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 \end{bmatrix}$$

$$z = x_1$$

Approx. Optimal Cost	Suboptimal Rational Cost	Ratio
2.4074	2.4282	1.009

We computed the optimal filter and the suboptimal filter using numerical optimization routines.

Quadruple Integrator

Estimate x_1

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 \end{bmatrix}$$

$$z = x_1$$

Approx. Optimal Cost	Suboptimal Rational Cost	Ratio
3.0722	3.0901	1.006

We computed the optimal filter and the suboptimal filter using numerical optimization routines.

Harmonic Oscillator

Estimate x_1

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 \end{bmatrix}$$

$$z = x_1$$

Approx. Optimal Cost	Suboptimal Cost	Ratio
1.26	1.3536	1.07

We computed the optimal filter and the suboptimal filter using numerical optimization routines.

Remarks

For the single integrator

- **The best first order filter that we found was 8.2% above optimal.**

Remarks

For the single integrator

- **The best first order filter that we found was 8.2% above optimal.**
- **It is the Kalman filter we would have constructed if we had assumed that the driving noise covariance was 2.5198.**

Remarks

For the single integrator

- **The best first order filter that we found was 8.2% above optimal.**
- **It is the Kalman filter we would have constructed if we had assumed that the driving noise covariance was 2.5198.**
- **The Kalman filter with driving noise covariance 1 was 36% above optimal.**

Remarks

For the single integrator

- **The best first order filter that we found was 8.2% above optimal.**
- **It is the Kalman filter we would have constructed if we had assumed that the driving noise covariance was 2.5198.**
- **The Kalman filter with driving noise covariance 1 was 36% above optimal.**

Remarks

For the double integrator

- The best second order filter for estimating x_1 that we found was 2.4% above optimal.

Remarks

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal.**
- **The best Kalman filter for estimating x_1 that we found was 2.6% above optimal. The driving noise covariance was 3.4.**

Remarks

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal.**
- **The best Kalman filter for estimating x_1 that we found was 2.6% above optimal. The driving noise covariance was 3.4.**
- **The Kalman filter for estimating x_1 with driving noise covariance 1 was 6.4% above optimal.**

Remarks

For the double integrator

- The best second order filter for estimating x_1 that we found was 2.4% above optimal.
- The best Kalman filter for estimating x_1 that we found was 2.6% above optimal. The driving noise covariance was 3.4.
- The Kalman filter for estimating x_1 with driving noise covariance 1 was 6.4% above optimal.
- The best gain for estimating x_2 was 7% above optimal.

Remarks

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal.**
- **The best Kalman filter for estimating x_1 that we found was 2.6% above optimal. The driving noise covariance was 3.4.**
- **The Kalman filter for estimating x_1 with driving noise covariance 1 was 6.4% above optimal.**
- **The best gain for estimating x_2 was 7% above optimal.**
- **If we used the best gain for estimating x_1 to estimate x_2 the performance was 9% above optimal.**

Remarks

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal.**
- **The best Kalman filter for estimating x_1 that we found was 2.6% above optimal. The driving noise covariance was 3.4.**
- **The Kalman filter for estimating x_1 with driving noise covariance 1 was 6.4% above optimal.**
- **The best gain for estimating x_2 was 7% above optimal.**
- **If we used the best gain for estimating x_1 to estimate x_2 the performance was 9% above optimal.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**
- **The best second order filter that we found was 1.4% above optimal. The poles were complex at $-1.9572 \pm 1.0372i$.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**
- **The best second order filter that we found was 1.4% above optimal. The poles were complex at $-1.9572 \pm 1.0372i$.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**
- **The best second order filter that we found was 1.4% above optimal. The poles were complex at $-1.9572 \pm 1.0372i$.**

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal. The best Kalman filter was 2.6% above optimal.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**
- **The best second order filter that we found was 1.4% above optimal. The poles were complex at $-1.9572 \pm 1.0372i$.**

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal. The best Kalman filter was 2.6% above optimal.**
- **The best fourth order filter for estimating x_1 that we found was 0.7% above optimal. The poles were at $-1.4442 \pm 0.9460i$ and $-1.7142 \pm 1.8055i$.**

Remarks

From this we might conclude that a Kalman filter can be a nearly optimal rational filter provided that we choose the driving noise covariance correctly.

This conclusion is wrong!

For the single integrator

- **The best first order filter that we found was 8.2% above optimal. It is the best Kalman filter.**
- **The best second order filter that we found was 1.4% above optimal. The poles were complex at $-1.9572 \pm 1.0372i$.**

For the double integrator

- **The best second order filter for estimating x_1 that we found was 2.4% above optimal. The best Kalman filter was 2.6% above optimal.**
- **The best fourth order filter for estimating x_1 that we found was 0.7% above optimal. The poles were at $-1.4442 \pm 0.9460i$ and $-1.7142 \pm 1.8055i$.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**
- **Rational minimax filtering is a computationally feasible alternative to Kalman filtering for low dimensional systems.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**
- **Rational minimax filtering is a computationally feasible alternative to Kalman filtering for low dimensional systems.**
- **It is possible to compute how close to optimal is a rational filter.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**
- **Rational minimax filtering is a computationally feasible alternative to Kalman filtering for low dimensional systems.**
- **It is possible to compute how close to optimal is a rational filter.**
- **Increasing the dimension of the filter over that of the plant can significantly improve performance.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**
- **Rational minimax filtering is a computationally feasible alternative to Kalman filtering for low dimensional systems.**
- **It is possible to compute how close to optimal is a rational filter.**
- **Increasing the dimension of the filter over that of the plant can significantly improve performance.**
- **More research is needed to understand how to choose a good suboptimal rational filter particularly when the dimension of the filter is greater than that of the original system.**

Conclusions

- **Minimax filters focus on worse case rather than average case performance.**
- **Minimax filters do not require knowledge of the driving noise covariance, instead, a bound on its magnitude.**
- **Rational minimax filtering is a computationally feasible alternative to Kalman filtering for low dimensional systems.**
- **It is possible to compute how close to optimal is a rational filter.**
- **Increasing the dimension of the filter over that of the plant can significantly improve performance.**
- **More research is needed to understand how to choose a good suboptimal rational filter particularly when the dimension of the filter is greater than that of the original system.**
- **Happy Birthday Eduardo**