Time Series Analysis

ECE 6563, Electrical and Computer Engineering
University of New Brunswick

Recap

In the last section, we expanded our models to include additional variables (multivariate)

- Regression
  Multivariate regression, exogenous vs endogenous
- ARIMAX
  Regression + (S)ARIMA model, ARIMA what is left from regression
- Vector Autoregression (VAR)
  Interrelated variables
- Causality vs. Correlation
  Granger Causality, Cross Correlation Function (CCF), Cointegration, Vector Error Correction

In this section, we'll explore state space

- State Space, Dynamic Linear Models
- Kalman Filters

State Space Models

So far, we've treated data almost empirically, without thinking too much about where the data come from.

- We've just modeled the autocorrelations using a data-driven approach
- We didn't really question the underlying processes that generated the data
- We assumed that the statistical properties of the underlying process remained the same (and/or did our best to make them)

State space models are a more general formulation of time series models that are characterized by two main principles

State Space Models

So far, we've treated data almost empirically, without thinking too much about where the data come from.

- We've just modeled the autocorrelations using a data-driven approach
- We didn't really question the underlying processes that generated the data
- We assumed that the statistical properties of the underlying process remained the same (and/or did our best to make them)

State space models are a more general formulation of time series models that are characterized by two main principles

1. There is a *hidden* or *latent* process, x_t, called the state process.

- The state process is assumed to be a Markov process, meaning that future values dependent only on the current value

State Space Models

So far, we've treated data almost empirically, without thinking too much about where the data come from.

- We've just modeled the autocorrelations using a data-driven approach
- We didn't really question the underlying processes that generated the data
- We assumed that the statistical properties of the underlying process remained the same (and/or did our best to make them)

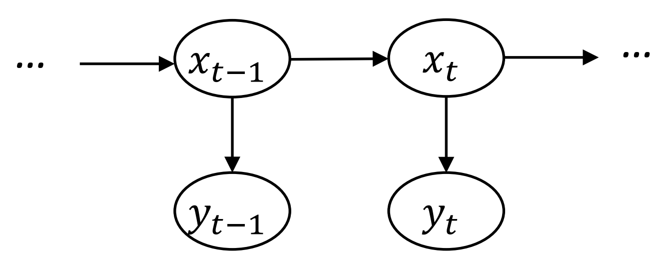
State space models are a more general formulation of time series models that are characterized by two main principles

1. There is a *hidden* or *latent* process, x_t, called the state process.

- The state process is assumed to be a Markov process, meaning that future values dependent only on the current value

2. The observations, y_t, are independent given the states x_t

- Any relationship between observations is dictated by the originating states



Linear State Space Models

The linear Gaussian state space model, or dynamic linear model (DLM), in its basic form, uses a first order autoregression as the *state equation:*

$x_t = Ax_{t-1} + w_t$ where w_t is IID with zero mean, normal distribution with finite variance

We also assume that the process starts somewhere with a normally distributed x_0

We can't see the state vector x_t directly, though, and instead observe a linearly transformed version of it with added noise, written as the *observation equation:*

$y_t = Cx_t + v_t$ where v_t is again IID with zero mean, normal distribution with finite variance

We also assume for simplicity that x_0, w_t, and v_t are uncorrelated.

Linear State Space Models

As with ARIMAX models, exogenous variables (other inputs) can be incorporated, either to the states themselves, or to the observations:

$$x_t = Ax_{t-1} + Bu_t + w_t \qquad\qquad y_t = Cx_t + Du_t + v_t$$

The variables can all be multivariate vectors, yielding matrix forms of the equations. For example:

$$\begin{pmatrix} x_{t,1} \\ x_{t,2} \\ x_{t,3} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{13} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix} \begin{pmatrix} x_{t-1,1} \\ x_{t-1,2} \\ x_{t-1,3} \end{pmatrix} + \begin{pmatrix} w_{t,1} \\ w_{t,2} \\ w_{t,3} \end{pmatrix}$$

The corresponding observation equation would be $y_t = C_t x_t + v_t$

where $y_t = \begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \end{pmatrix}$, and C_t would be a 3x3 observation matrix.

Linear State Space Models

As with ARIMAX models, exogenous variables (other inputs) can be incorporated, either to the states themselves, or to the observations:

$$x_t = Ax_{t-1} + Bu_t + w_t \qquad\qquad y_t = Cx_t + Du_t + v_t$$

Although simple, this state format is quite generalizable.

- For example, for an AR(2) model, $x_t = A_1 x_{t-1} + A_2 x_{t-2} + \varepsilon_t$
- We could write the following state and observation equations:

$$\begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x_{t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix} \qquad\qquad y_t = (C, 0) \begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix}$$

Similarly, with sufficient interpretation, the whole family of ARIMA models can be rewritten as state space models

Dynamic Linear Models

The ability to define special forms of the weighting and transitions matrices enables a variety of different models to be defined

For example, for regular regression, we can rewrite the formula in state space form:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad \Rightarrow \quad y = [1 \;\; x] \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \varepsilon$$

or $$y = X'\theta + \varepsilon$$

where $$X' = [1 \;\; x] \qquad \theta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

Dynamic Linear Models

The ability to define special forms of the weighting and transitions matrices enables a variety of different models to be defined

For example, for regular regression, we can rewrite the formula in state space form:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad \Rightarrow \quad y = \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \varepsilon$$

or $$y = X'\theta + \varepsilon$$

where $X' = \begin{bmatrix} 1 & x \end{bmatrix}$ $\quad \theta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$

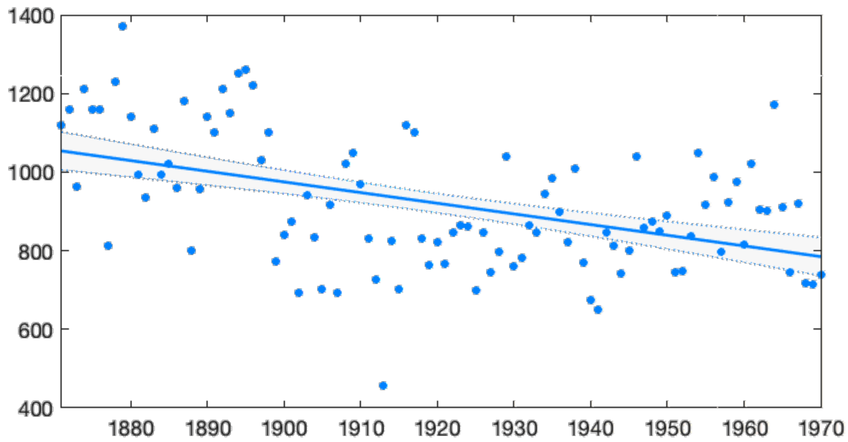
But, by explicitly incorporating time, and allowing the regression coefficients to change over time, we can define a dynamic linear model

$$y_t = X_t'\theta_t + \varepsilon_t$$

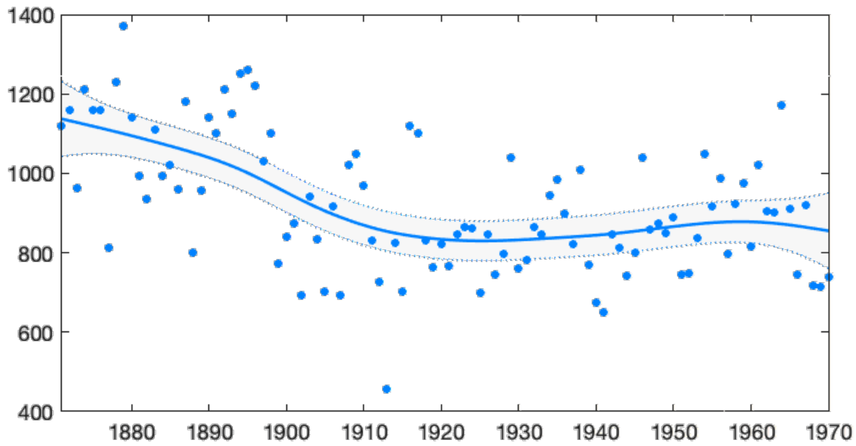
Dynamic Linear Models

By explicitly incorporating time, and allowing the regression coefficients to change over time, we can define a dynamic linear model

$$y_t = X_t'\theta_t + \varepsilon_t$$



DLM without dynamical evolution
(regular linear regression)



DLM with dynamical evolution

https://mjlaine.github.io/dlm/dlmtut.html

Dynamic Linear Models

This extension emphasizes two important points:

1. The addition of the subscript t explicitly acknowledges that there is information in the time ordering of the data in y (as we've seen)

$$y_t = X_t' \theta_t + \varepsilon_t$$

2. The relationship between X and y is potentially different, for different values of t

This could be a powerful extension, but poses a problem for parameter estimation

- We only have one sample (data point) per t – we can't estimate statistical properties/distributions
- For this reason, it is common to constrain the behaviour of the parameters to be dependent from t to $t + 1$ (we'll come back to this)

https://mjlaine.github.io/dlm/dlmtut.html

State Space Models

In our original definition of the state space model, we included an additive noise term in the observation equation

- We left this term out in showing how we could write an AR(2) process in this format

Let's now model a univariate AR(1) process, but accommodate the fact that it may be measured using a noisy sensor

- The state equation is, as before:

$$x_t = Ax_{t-1} + w_t$$

- The observation equation is also simple (here, $C = 1$):

$$y_t = x_t + v_t$$

- So, when working with this signal, we must now consider the additional v_t term.

$$y_t = x_{t-1} + w_t + v_t$$

State Space Models

So, when working with this signal, we must now consider the additional v_t term.

$$y_t = Ax_{t-1} + w_t + v_t$$

Then, if we were to look at the ACF of this new noisy version of y_t:

$$\gamma_{y(0)} = var(y_t) = var(x_t + v_t) = \frac{\sigma_w^2}{1+\phi^2} + \sigma_v^2$$

$$\gamma_{y(h \neq 0)} = cov(y_t, y_{t-h}) = cov(x_t + v_t, x_{t-h} + v_{t-h})$$

Unless $\sigma_v^2 = 0$, (i.e. no noise), this no longer looks like an AR(1) process!

State Space Models

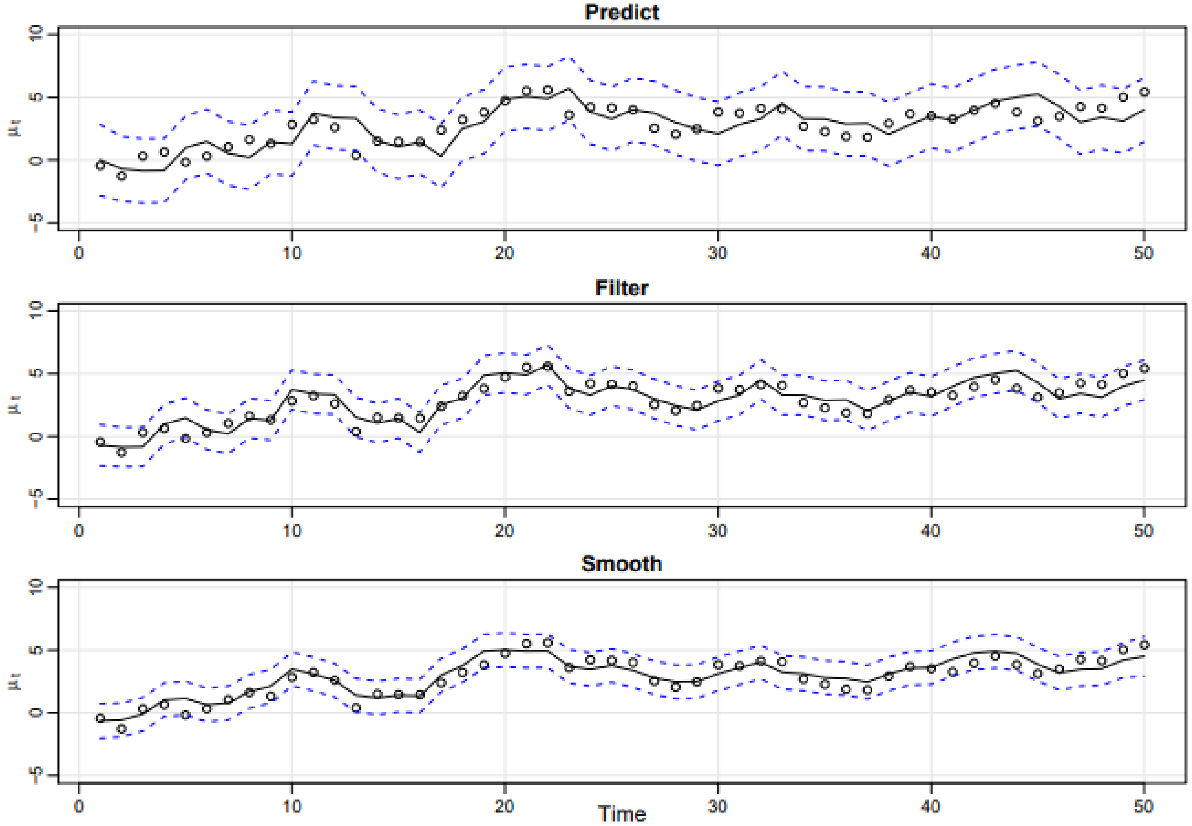
A primary and common aim of using state space models is therefor to produce estimations of an underlying state given the presence of noise (or uncertainty) in the observation.

- Given a time series of observations, $\{y_1 \dots y_s\}$, we want to understand what is happening with the underlying, generating state, x_t
- If for our time series $s < t$, then the goal of determining x_t is *forecasting*
- If $s = t$ (we have a measurement for the current time, but want to know what the corresponding state is), then the model is *filtering*
- If $s > t$ (we have an existing 'dataset' beyond t), then the model is used for *smoothing*

State Space Models

The confidence intervals for each of these cases varies, intuitively.

- Predicting into the future results in wider confidence intervals
- Smoothing past values, with the benefit of hindsight, yields tighter confidence intervals



Shumway & Stoffer (2016). Time Series Analysis and Its Applications

Kalman Filters

So far, in our state space models, we've said that

- Our observations are a measurement of a true underlying value/state, but with some amount of error
  - Think of a scale that measures your weight, but has +/- 2kg of error
- We may also have multiple predicting variables that each have a certain amount of error (or degree of certainty/confidence)
  - Think about the positioning of a robot based on multiple sensors (camera, lidar, IMU, GPS, etc…)
- The evolution of the parameters must be somehow constrained to have dependence from t to $t+1$

But, we need to somehow estimate the corresponding parameters:

- Transition matrix, A,
- Observation matrix, C,
- Noise variances σ_v^2 and σ_w^2

$$x_t = Ax_{t-1} + w_t \qquad\qquad y_t = Cx_t + v_t$$

Kalman Filters

So far, in our state space models, we've said that

- Our observations are a measurement of a true underlying value/state, but with some amount of error
  - Think of a scale that measures your weight, but has +/- 2kg of error
- We may also have multiple predicting variables that each have a certain amount of error (or degree of certainty/confidence)
  - Think about the positioning of a robot based on multiple sensors (camera, lidar, IMU, GPS, etc…)
- The evolution of the parameters must be somehow constrained to have dependence from t to $t+1$

The Kalman filter provides an optimal estimation of the states of a system when we have indirect and/or uncertain measurements of those states.

It is used when:

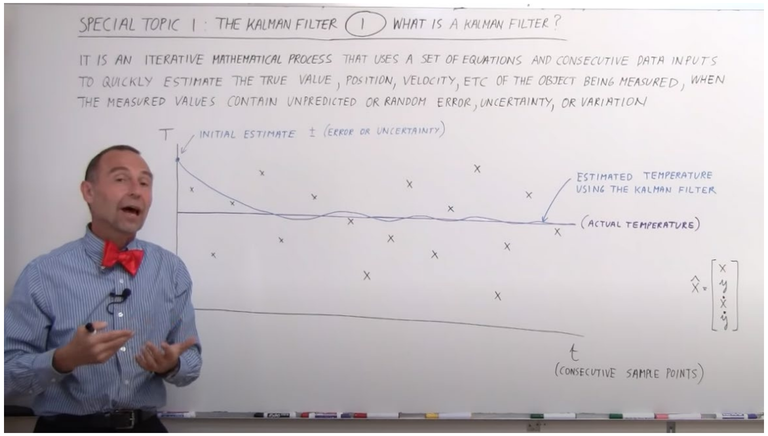
1. There is some level of uncertainty about a dynamic system (e.g. measurements have noise)
2. You can make an educated guess about what the system will do next (you know the dynamics)

Kalman Filters

Disclaimer: We are only going to glance over Kalman filters, even though they deserve far more time in engineering applications

- There are other courses at UNB that can expose you to Kalman Filters
- Great online courses and material (e.g. check out Michel van Biezen on YouTube)
- Heavily used in position tracking (robotics, autonomous vehicles, etc.)

- To “shift” the perspective in many tutorials to be move aligned with our time series, simply consider a plot of position, velocity, direction, etc., over time.

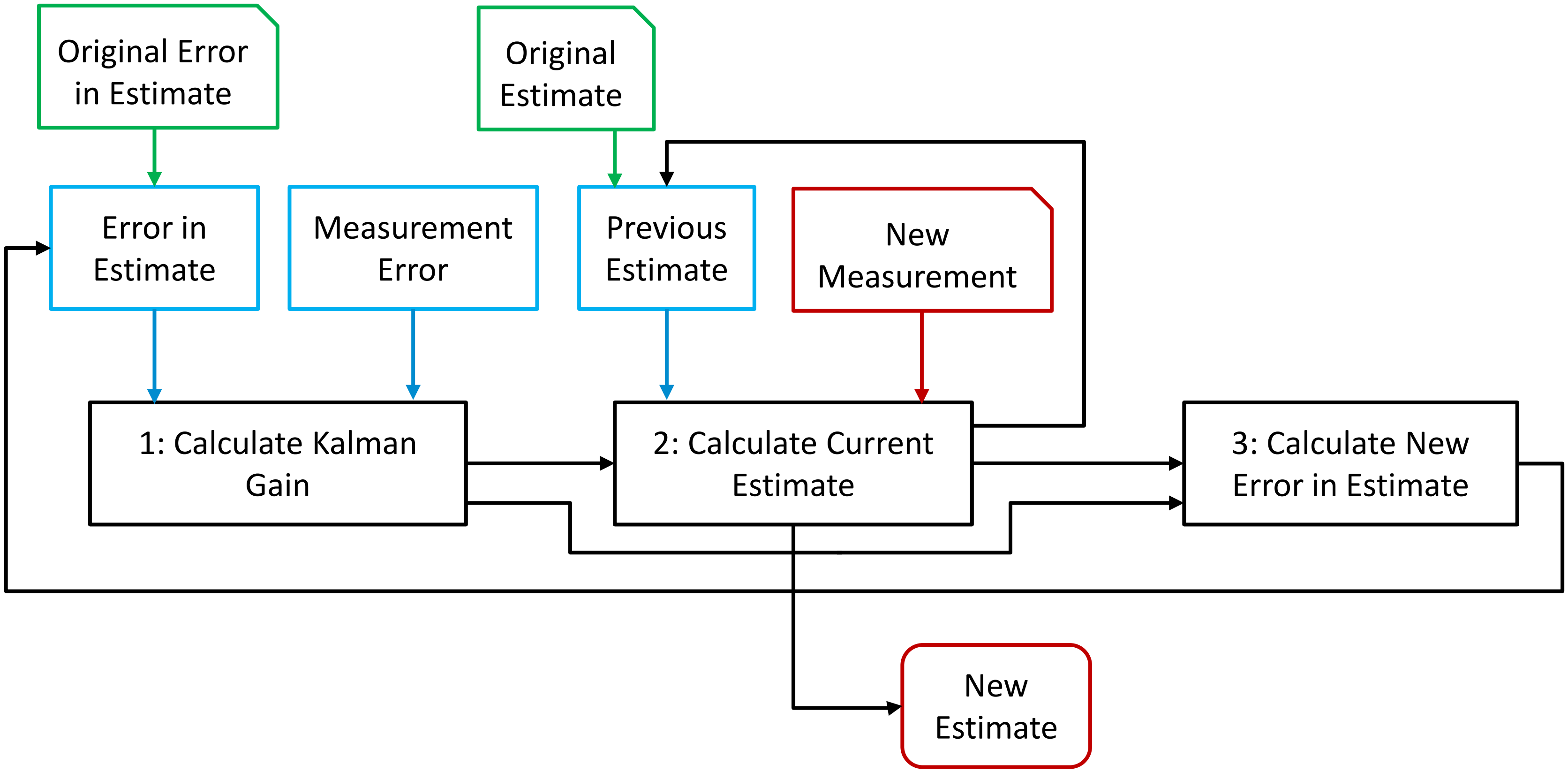


https://www.youtube.com/watch?v=-cD7WkbAIL0&ab_channel=MichelvanBiezen

Kalman Filters

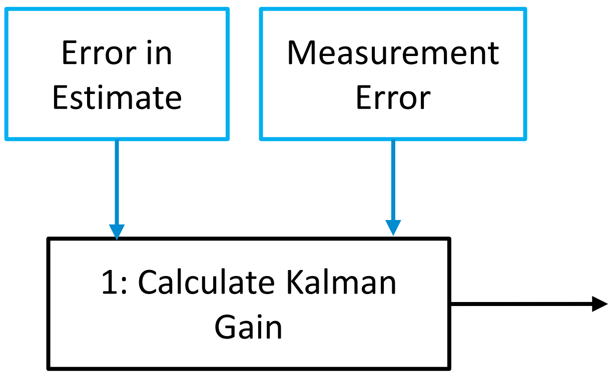
First, let's look at what's happening, before we go back to the equations

- There are three main operations that drive the iterations of the Kalman Filter



Note: Here, error and *uncertainty* used synonymously

Kalman Filters



Kalman Gain (K)
Measurement ($Meas$)
Measurement Error (Err_{meas})

Current Estimate (Est_t)
Previous Estimate (Est_{t-1})
Error in Estimate (Err_{est})

$$K = \frac{Err_{est}}{Err_{est} + Err_{meas}} \qquad 0 \leq K \leq 1$$

If $Err_{est} >> Err_{meas}$, then K is close to 1; we want to trust the measurement

$$Est_t = Est_{t-1} + \underline{K(Meas - Est_{t-1})}$$

If $Err_{meas} >> Err_{est}$, then K is close to 0; we want to trust the estimate

$$Est_t = \underline{Est_{t-1}} + K(Meas - Est_{t-1})$$

Over time, K tends to decrease, as the model improves

Kalman Filters

1: Calculate Kalman Gain → 2: Calculate Current Estimate → 3: Calculate New Error in Estimate

Kalman Gain (K)
Measurement ($Meas$)
Measurement Error (Err_{meas})

Current Estimate (Est_t)
Previous Estimate (Est_{t-1})
Error in Estimate (Err_{est})

1. $$K = \frac{Err_{est}}{Err_{est} + Err_{meas}} \qquad 0 \leq K \leq 1$$

2. $$Est_t = Est_{t-1} + K(Meas - Est_{t-1})$$

3. $$Err_{est(t)} = \frac{Err_{meas}\, Err_{est(t-1)}}{Err_{meas} + Err_{est(t-1)}}$$ or $$Err_{est(t)} = [1 - K]Err_{est(t-1)}$$

Kalman Filters

This has been a simplified example of what is often a multidimensional problem

- We'll leave it to independent learning to explore this extension

$$X_t = AX_{t-1} + Bu_t + w_t \qquad\qquad Y_t = CX_t + v_t$$

An important consideration for Kalman filters, though, is that the model is often based on physical considerations.

- We said that they are good when we can make an educated guess about what the system will do next
- This leads to dynamical models that dictate the state matrices
- e.g. a system may model the position and velocity of an object, with updates being constrained by the corresponding physics

Kalman Filters

An important consideration for Kalman filters, though, is that the model is often based on physical considerations.

- We said that they are good when we can make an educated guess about what the system will do next
- This leads to dynamical models that dictate the state matrix
- For example, if a system is tracking an object along the x-dimension in space, it may have a state matrix:

$$X = \begin{bmatrix} x \\ \dot{x} \end{bmatrix} \begin{array}{l} \longleftarrow \text{Position} \\ \longleftarrow \text{Velocity} \end{array}$$

and be governed by the laws of physics $$x = x_0 + \dot{x}t + \frac{1}{2}\ddot{x}t^2$$

Then, the state equation might be

$$X_t = AX_{t-1} + Bu_t + w_t$$

Updates state based on past position & velocity

The Control variable updates the state based on acceleration

Kalman Filters

For example, if a system is tracking an object along the x-dimension in space, it may have a state matrix:

$$X = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$$

Position

Velocity

and be governed by the kinematics laws of physics $x = x_0 + \dot{x}t + \frac{1}{2}\ddot{x}t^2$

$$X_t = AX_{t-1} + Bu_t + w_t$$

Updates state based on past position & velocity

The Control variable, in this case, updates the state based on acceleration

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \qquad AX = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} x + \Delta t\dot{x} \\ 0 + \quad \dot{x} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} \qquad u = [\ddot{x}] \qquad Bu = \begin{bmatrix} \frac{1}{2}\ddot{x}\Delta t^2 \\ \ddot{x}\Delta t \end{bmatrix} \qquad X_t = \begin{bmatrix} x + \Delta t\dot{x} \\ 0 + \quad \dot{x} \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\ddot{x}\Delta t^2 \\ \ddot{x}\Delta t \end{bmatrix} + w_t$$

any acceleration

Kalman Filters

If instead, we want to use a Kalman filter to work with stock market predictions, these same considerations must be made

- Not as intuitive or as well defined; what are Newton's equations for stocks?
  (actually, it turns out that a similar model is sometimes used)
- This is sometimes a drawback of Kalman filters; although powerful, they can require more work and domain knowledge to set up

We have also only touched on a few parts of Kalman filters.

- For example, we haven't covered how to update the estimates of errors in our estimates.
- If the source of potential errors aren't fully modelled, it can lead to ill-conditions models that don't perform well.

So, while fascinating and powerful, we must move on…

- It is left to you to explore further if interested

Q&A