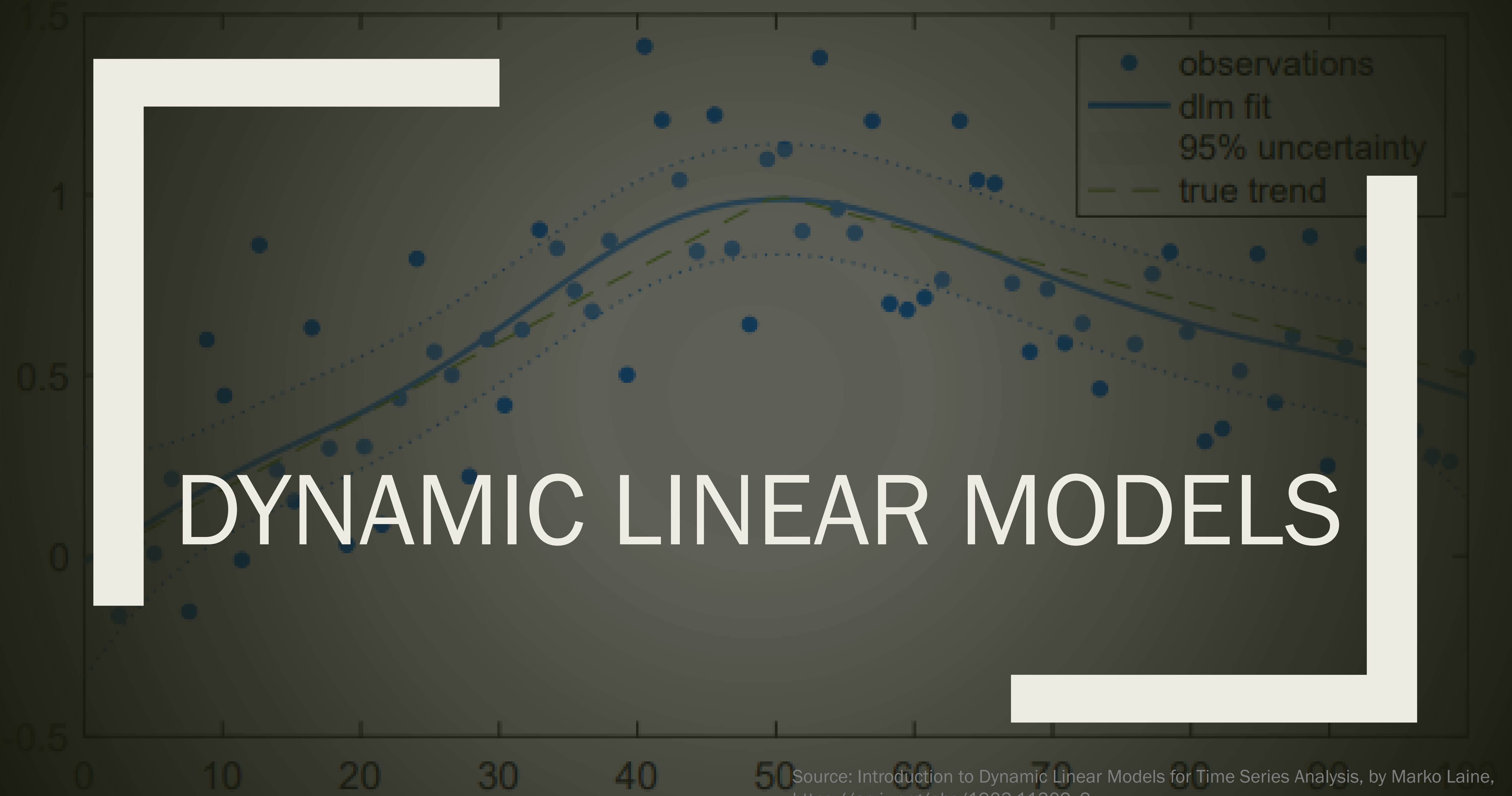
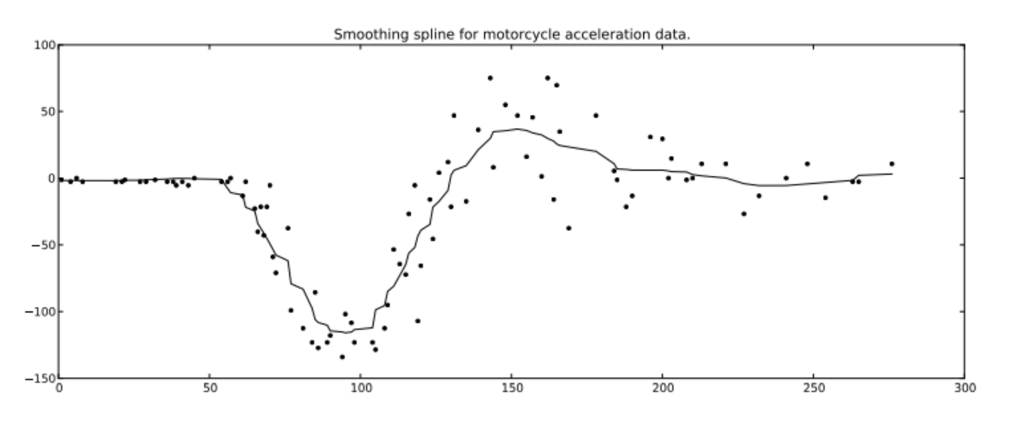
DYNAMIC LINEAR MODELS



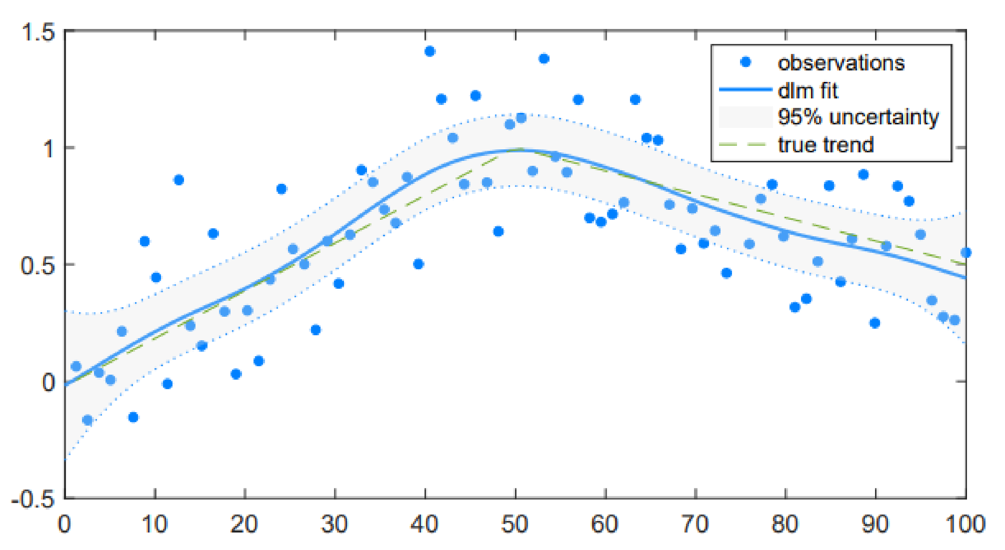
Source: Introduction to Dynamic Linear Models for Time Series Analysis, by Marko Laine, https://arxiv.org/abs/1903.11309v2

Introduction to DLM

- It is difficult to deal with the non-stationary time series
- We need to assume that some distributional properties of the process that generate the observations do not change with time
- Dynamic regression avoids this by explicitly allowing temporal variability in the regression coefficients and by letting some of the system properties to change in time
- Many classical time series models can be formulated as DLMs, including ARMA, ARIMA, GARCH models
- The main goals are short-term forecasting, intervention analysis and monitoring



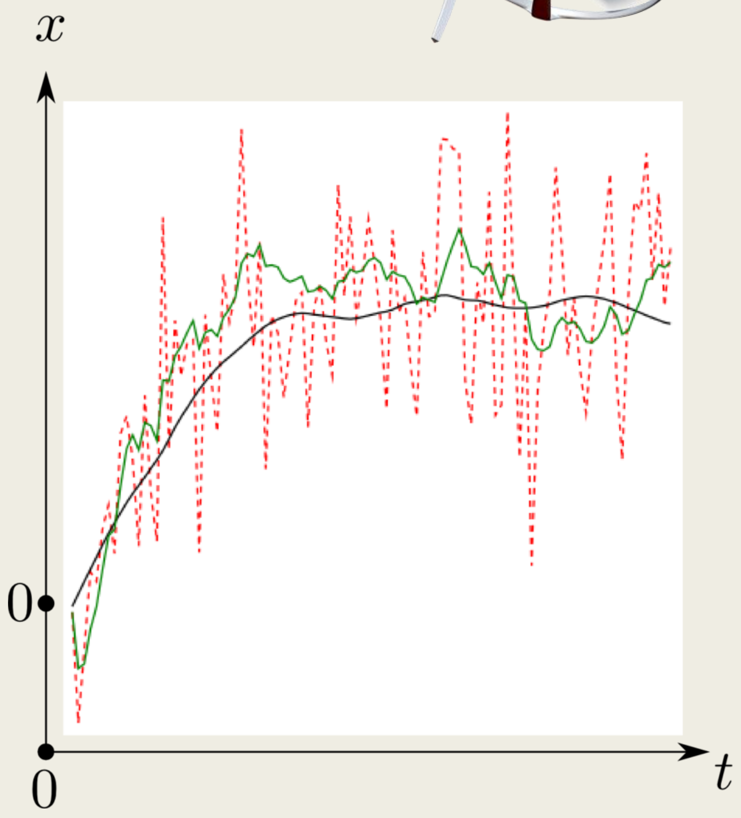
Source: PySSM: Bayesian Inference of Linear Gaussian State Space Models in Python, by C.M.Strickland et al.



Source: Introduction to Dynamic Linear Models for Time Series Analysis, by Marko Laine, https://arxiv.org/abs/1903.11309v2

State space modeling (SSM)

- We have a set of states that evolve in time, but our observations of these states contain statistical noise, and hence we are unable to ever directly observe the "true" states.
- The goal of the state space model is to infer information about the states, given the observations, as new information arrives.
- A famous algorithm for carrying out this procedure is the Kalman Filter.
- Applications: engineering control problems, including guidance & navigation, spacecraft trajectory analysis, etc., quantitative finance and macroeconomics
- Example: a radar tracks a target, i.e. determine its location, speed, and acceleration, while the estimates of its location were received gradually and with noise. We need to construct a model of optimal, continuously updated estimates of the position and speed of an object based on the results of a time series of inaccurate measurements of its location.



Truth; filtered process; observations.
Source: https://en.wikipedia.org/wiki/Kalman_filter

State space models

- The model:
  - *Discrete time* $t_1, t_2, \dots, t_T$,
  - *Model for states of the system,*
  - *Model for observations of the system.*
- The state of the system is described by a vector of finite dimension (the state vector).
- At each time cycle, the linear operator converts the state vector into another state vector (deterministic state change), adds some normal noise vector (random factors) and, probably, a control vector that simulates the effect of the control system.

Source: http://www.machinelearning.ru/wiki/images/8/83/SMAIS2009-10.pdf

Dynamic linear models

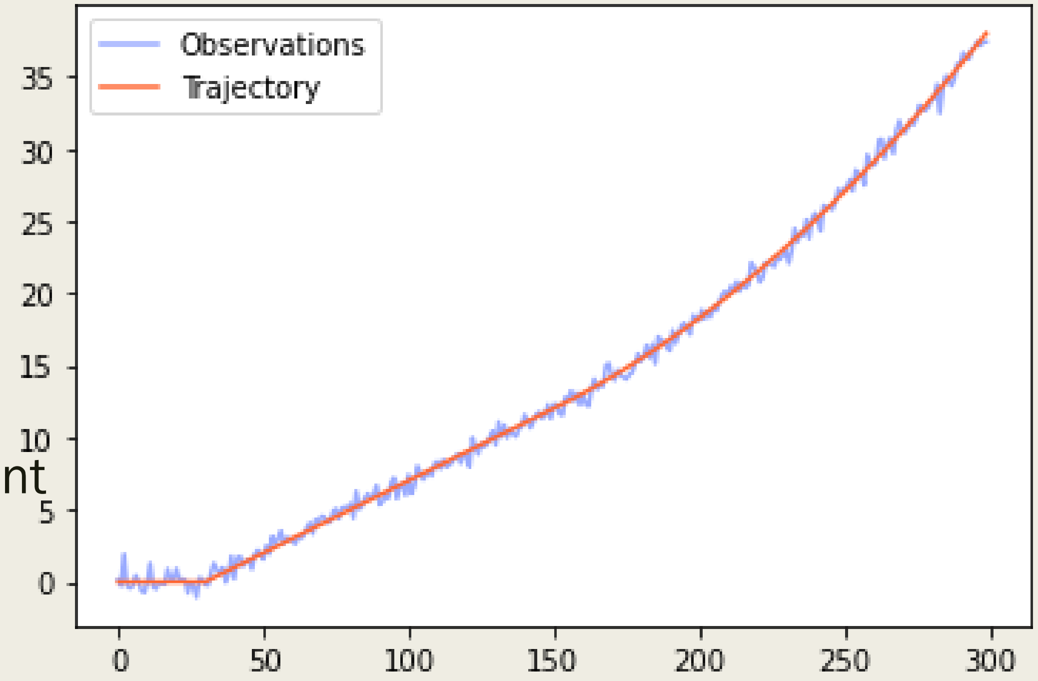
- The model:
  - *Discrete time* $t_1, t_2, \dots, t_T$,
  - *Model for states of the system:* $x_t = Fx_{t-1} + Bu_t + w_t$

    Evolution of the system

    Control over system
    - x_t is a current state of the system,
    - F is a state transition matrix,
    - u_t is a control vector,
    - B is a control matrix,
    - w_t is an error of the model, $w_t \sim N(0, Q)$, Q is a covariance matrix
  - *Model for observations of the system.*

Source: http://www.machinelearning.ru/wiki/images/8/83/SMAIS2009-10.pdf

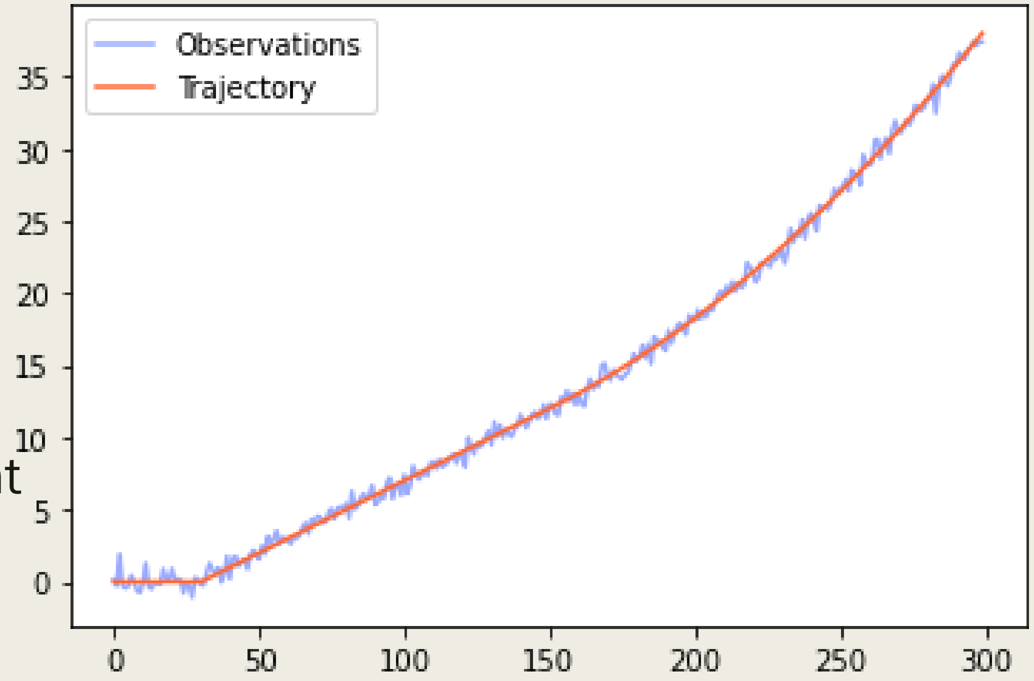
Dynamic linear models



- Simple example: robot, one-dimensional movement
- $x_t = Fx_{t-1} + Bu_t + w_t$
  - x_t *is a current state of the system,*
  - F is a state transition matrix,
  - u_t *is a control vector,*
  - B is a control matrix,
  - w_t *is an error of the model,* $w_t \sim N(0, Q), Q$ *is a covariance matrix*

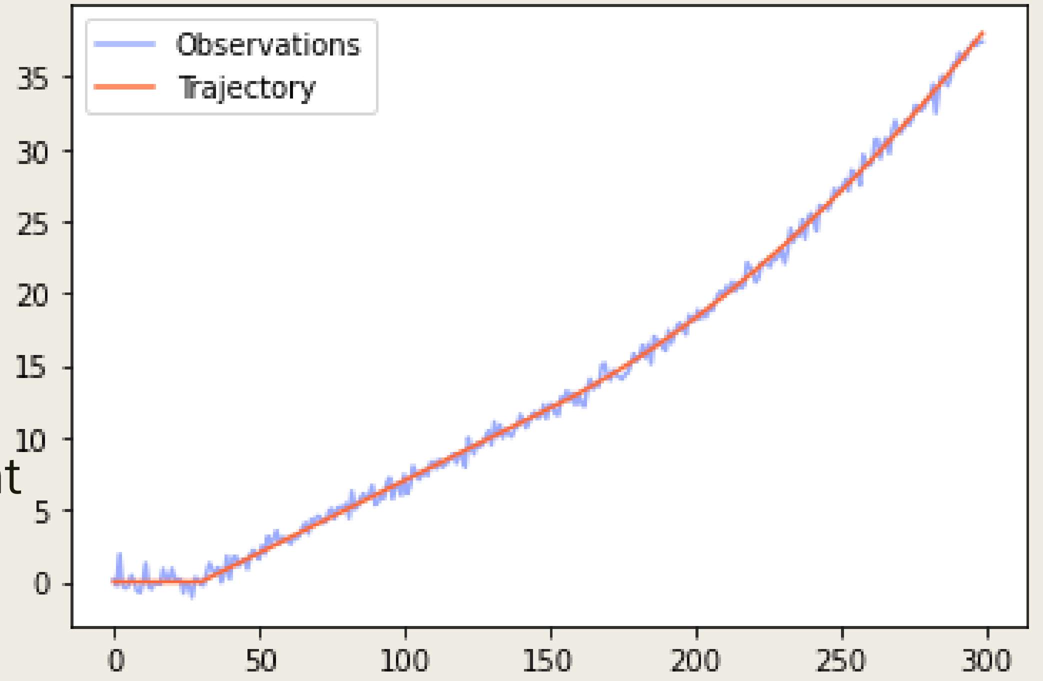
We cannot control the robot (just watch),
so this part is 0

Dynamic linear models



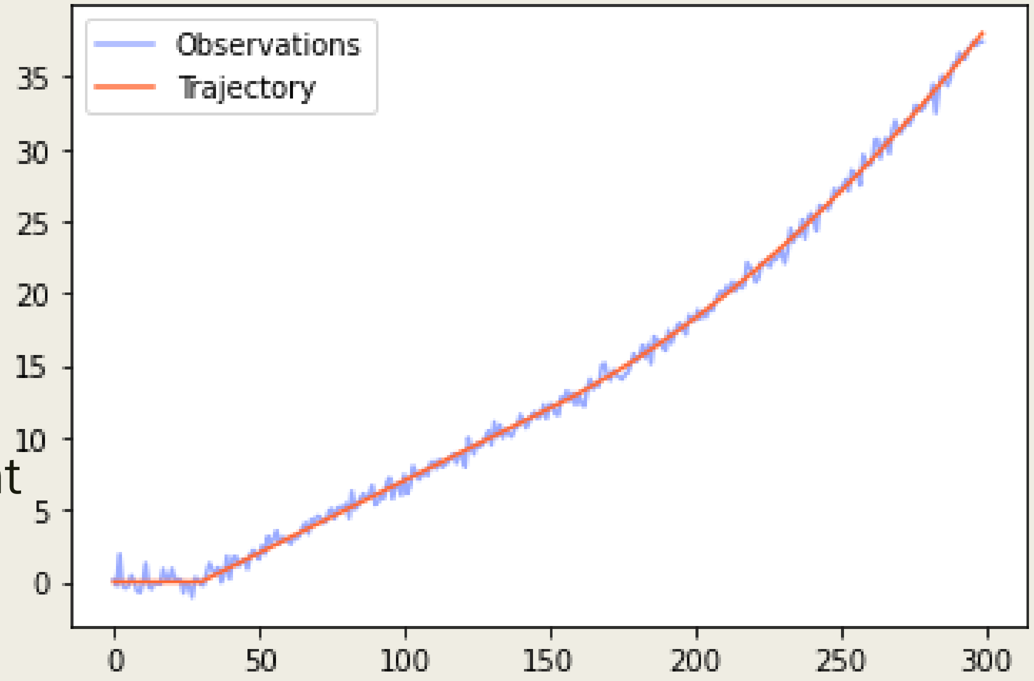
- Simple example: robot, one-dimensional movement
- $x_t = F x_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
  - *w_t is an error of the model, $w_t \sim N(0, Q)$, Q is a covariance matrix*

Dynamic linear models



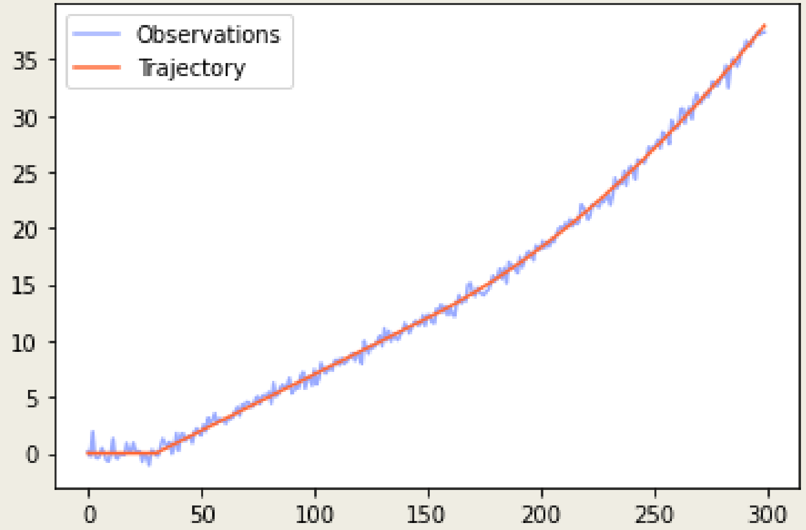
- Simple example: robot, one-dimensional movement
- $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model, $w_t \sim N(0, Q)$, Q is a covariance matrix*

Dynamic linear models



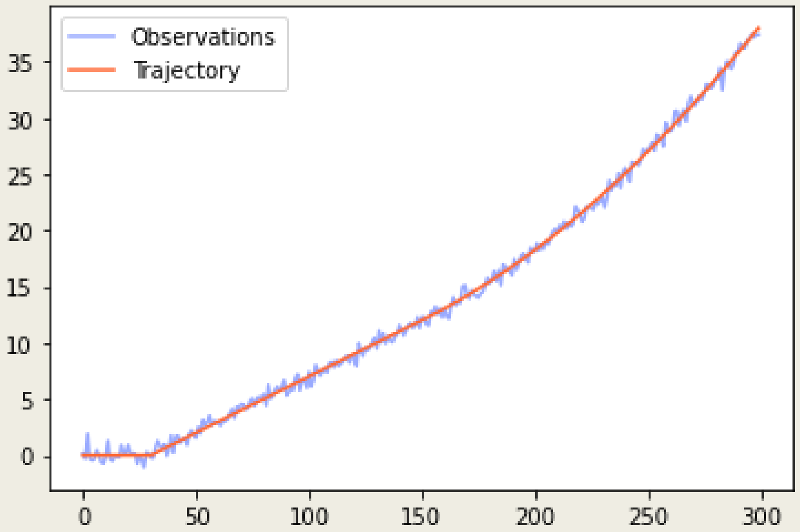
- Simple example: robot, one-dimensional movement
- $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*

Dynamic linear models



- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

Dynamic linear models

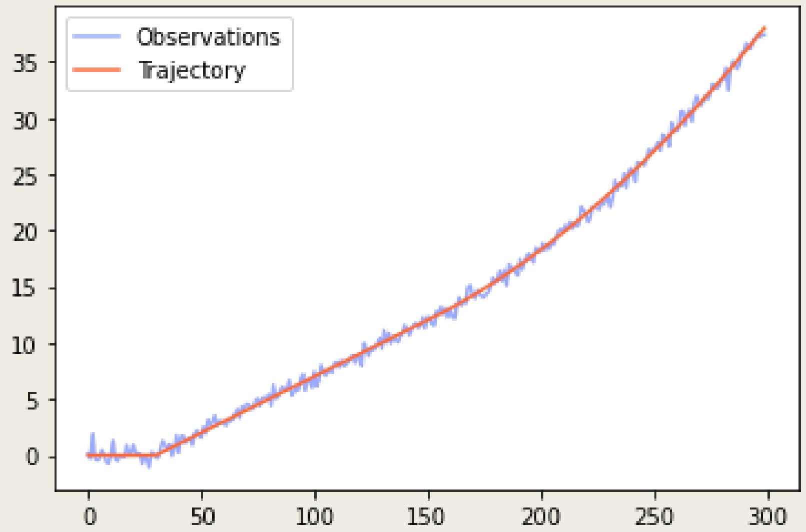


- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - x_t *is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - F *is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - w_t *is an error of the model (bumps in the road, wind, noise, etc.),* $w_t \sim N(0, Q)$, Q *is a covariance matrix*

In general $\dim(z_t) \neq \dim(x_t)$!

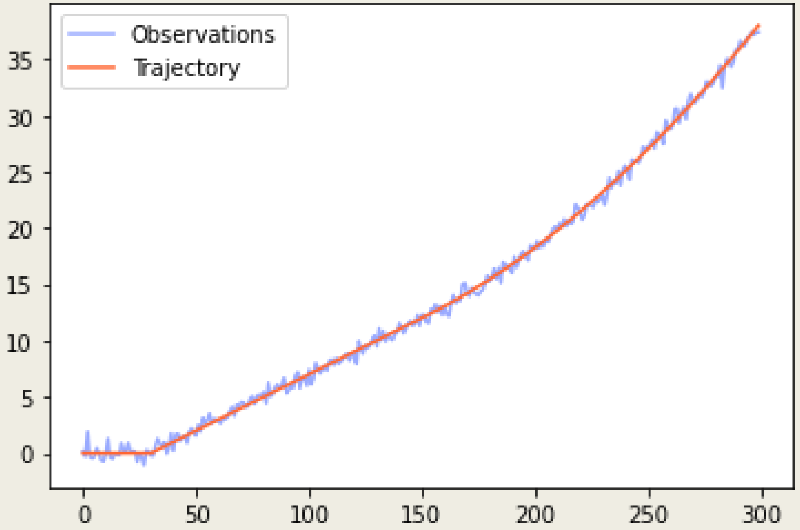
- Observation model: $z_t = Hx_t + v_t$,
  - z_t *is measurements made by sensors,*
  - v_t *is a vector of measurement errors,* $v_t \sim N(0, R)$, R *is a covariance matrix*

Dynamic linear models



- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,

  If we can measure only position, then $H_t = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

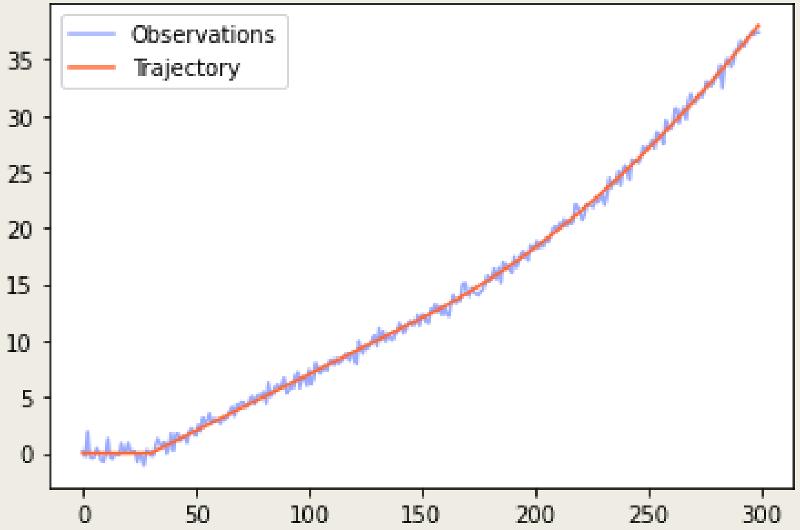


Dynamic linear models

- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,

  If we can measure position and speed, then $H_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

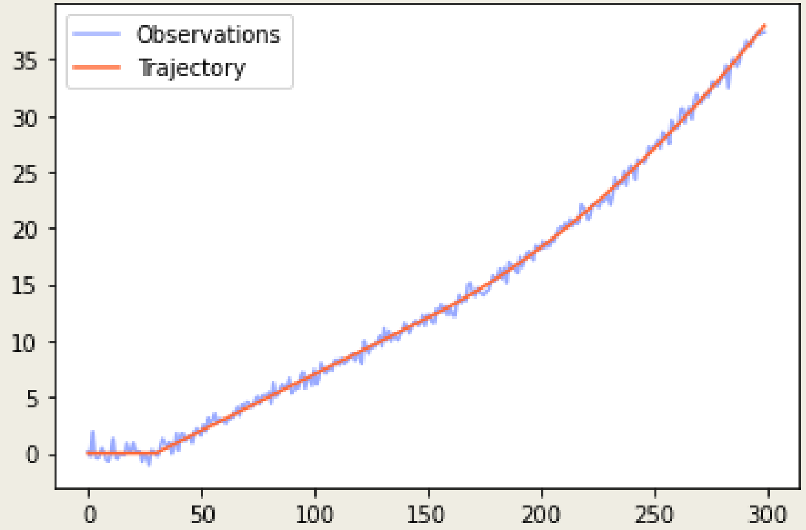
Dynamic linear models



- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,

  If we can measure speed from two sensors, then $H_t = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

Dynamic linear models



- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

If we can measure speed from two sensors, one in km/h (as in model) and the second one uses mph, then $H_t = \begin{bmatrix} 0 & 0 \\ 1 & 1.609 \\ 0 & 0 \end{bmatrix}$

More examples

- Second-order Difference Equation: $y_{t+1} = \phi_0 + \phi_1 y_t + \phi_2 y_{t-1}$

- Model for states of the system $x_t = F x_{t-1}$
  - $x_t = (1, y_t, y_{t-1})^T$ *is a state vector;*
  - $F = \begin{bmatrix} 1 & 0 & 0 \\ \phi_0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix}$ *is a state transition matrix;*

- Observation model: $z_t = H x_t$,
  - z_t *is measurements made by sensors and* $H = [0 \quad 1 \quad 0]$

More examples

- Autoregressive Process: $y_{t+1} = \phi_1 y_t + \phi_2 y_{t-1} + \phi_3 y_{t-2} + \phi_4 y_{t-3} + w_{t+1}$

- Model for states of the system $x_t = F x_{t-1} + w_t$
  - $x_t = (y_t, y_{t-1}, y_{t-2}, y_{t-3})^T$ is *a state vector;*
  - $F = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ *is a state transition matrix;*
  - w_t *is an error of the model (bumps in the road, wind, noise, etc.),* $w_t \sim N(0, Q)$, Q *is a covariance matrix*

- Observation model: $z_t = H x_t$,
  - z_t *is measurements made by sensors and* $H = [1 \quad 0 \quad 0 \quad 0]$

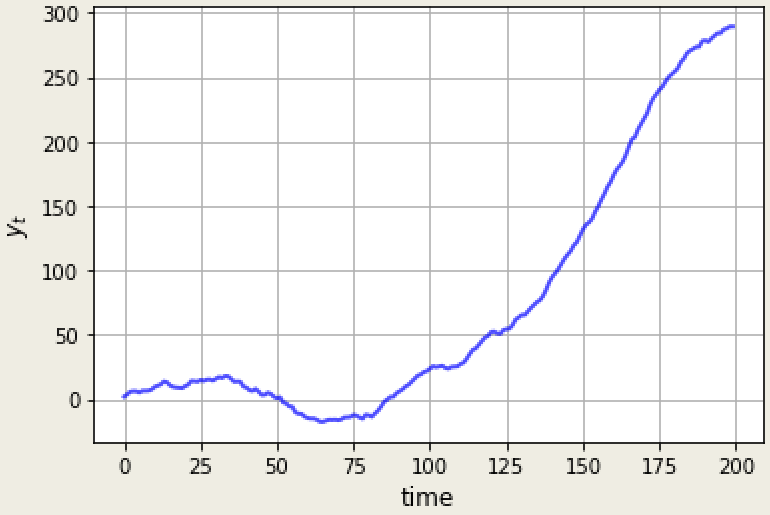
More examples

- Linear Time Trend: $y_t = at + b$
- Model for states of the system $x_t = Fx_{t-1}$
  - $x_t = (1,1)^T$ is *a state vector;*
  - $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ *is a state transition matrix;*
- Observation model: $z_t = Hx_t$,
  - z_t *is measurements made by sensors and* $H = [a \quad b]$

-

More examples

- Local level model
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - $x_t = \mu_t$, is a *state vector;*
  - $F = [1]$ *is a state transition matrix;*
- Observation model: $z_t = Hx_t + v_t$,
  - z_t *is measurements made by sensors and* $H = [1]$

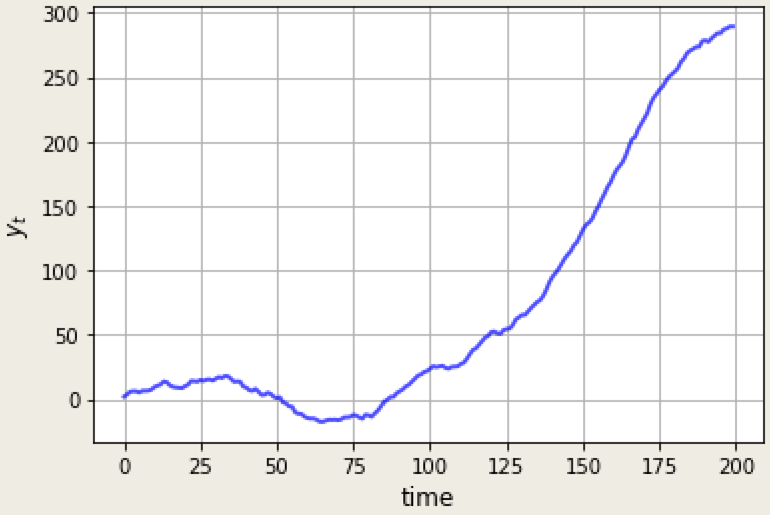


$z_t = \mu_t + v_t$

mean level process $\mu_t = \mu_{t-1} + \epsilon_1$

More examples

- Local trend model
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - $x_t = (\mu_t, \alpha_t)^T$ is *a state vector;*
  - $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$*is a state transition matrix;*
- Observation model: $z_t = Hx_t + v_t$,
  - z_t *is measurements made by sensors and* $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$



$$z_t = \mu_t + v_t$$

mean level process $\mu_t = \mu_{t-1} + \alpha_{t-1} + \epsilon_1$
change in the mean $\mu_t - \mu_{t-1}$ is controlled by the trend process $\alpha_t = \alpha_{t-1} + \epsilon_2$

Kalman filter

- Kalman filtering is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe.
- The algorithm works in a two-step process:
  - *prediction step, where the Kalman filter produces estimates of the current state variables, along with their uncertainties (because next measurement will be necessarily corrupted with some amount of error, including random noise).*
  - *correction (or update) step, where these estimates are updated using a weighted average, with more weight being given to estimates with higher certainty.*
- The algorithm is recursive. It can run in real time,; no additional past information is required. using only the present input measurements and the previously calculated state and its uncertainty matrix

Kalman filter

- Simplest example first (from [KALMAN FILTERS (archive.org)]())
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$

Kalman filter

- Simplest example first (from [KALMAN FILTERS (archive.org)]())
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$
- The initial estimate (a priori estimate) of x_0 is 1000, the variance of error is P=40000
- So, our estimation of x_0 will be called x_e and x_e is 1000. The variance of the error in this estimate is defined by $P = E[(x_0 - x_e)^2]$

Kalman filter

- Simplest example first (from KALMAN FILTERS (archive.org))
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$
- The initial estimate (a priori estimate) of x_0 is 1000, the variance of error is P=40000
- So, our estimation of x_0 will be called x_e and x_e is 1000. The variance of the error in this estimate is defined by $P = E[(x_0 - x_e)^2]$
- Prediction step. Now we would like to estimate x_1. Using the dynamic equation, we get

$$x_1 = Fx_0 + w_0$$

- The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)

Kalman filter

- Simplest example first (from [KALMAN FILTERS (archive.org)])
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$
- The initial estimate (a priori estimate) of x_0 is 1000, the variance of error is P=40000
- So, our estimation of x_0 will be called x_e and x_e is 1000. The variance of the error in this estimate is defined by $P = E[(x_0 - x_e)^2]$
- Prediction step. Now we would like to estimate x_1. Using the dynamic equation, we get
$$x_1 = Fx_0 + w_0$$
- The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$

Kalman filter

- Simplest example first (from KALMAN FILTERS (archive.org))
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$
- The initial estimate (a priori estimate) of x_0 is 1000, the variance of error is P=40000
- So, our estimation of x_0 will be called x_e and x_e is 1000. The variance of the error in this estimate is defined by $P = E[(x_0 - x_e)^2]$
- Prediction step. Now we would like to estimate x_1. Using the dynamic equation, we get
$$x_1 = Fx_0 + w_0$$
- The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$

Kalman filter

- Simplest example first (from KALMAN FILTERS (archive.org))
- One dimensional random variable x_t satisfies a linear dynamic equation $x_t = Fx_{t-1} + w_t$. F is a known number, assume $F = 0.9$. White noise $w_t \sim N(0,100)$
- The initial estimate (a priori estimate) of x_0 is 1000, the variance of error is P=40000
- So, our estimation of x_0 will be called x_e and x_e is 1000. The variance of the error in this estimate is defined by $P = E[(x_0 - x_e)^2]$
- Prediction step. Now we would like to estimate x_1. Using the dynamic equation, we get
$$x_1 = Fx_0 + w_0$$
- The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2]$

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2]$

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2] = E[F^2(x_0 - x_e)^2] + E[w_0^2] + 2E[F(x_0 - x_e)w_0]$, but w_0 is uncorrelated with x_0 and x_e, so
$$P^{new} = F^2E[(x_0 - x_e)^2] + E[w_0^2]$$

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2] = E[F^2(x_0 - x_e)^2] + E[w_0^2] + 2E[F(x_0 - x_e)w_0]$, but w_0 is uncorrelated with x_0 and x_e, so
$$P^{new} = F^2E[(x_0 - x_e)^2] + E[w_0^2]$$
- As $E[(x_0 - x_e)^2] = P$ and $E[w_0^2] = Q$, then $P^{new} = F^2P + Q$ (2)

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2] = E[F^2(x_0 - x_e)^2] + E[w_0^2] + 2E[F(x_0 - x_e)w_0]$, but w_0 is uncorrelated with x_0 and x_e, so
$$P^{new} = F^2E[(x_0 - x_e)^2] + E[w_0^2]$$
- As $E[(x_0 - x_e)^2] = P$ and $E[w_0^2] = Q$, then $P^{new} = F^2P + Q$ (2)
- For our example $P^{new} = F^2P + Q = 0.81 * 40000 + 100 = 32500$

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2] = E[F^2(x_0 - x_e)^2] + E[w_0^2] + 2E[F(x_0 - x_e)w_0]$, but w_0 is uncorrelated with x_0 and x_e, so
$$P^{new} = F^2E[(x_0 - x_e)^2] + E[w_0^2]$$
- As $E[(x_0 - x_e)^2] = P$ and $E[w_0^2] = Q$, then $P^{new} = F^2P + Q$ (2)
- For our example $P^{new} = F^2P + Q = 0.81 * 40000 + 100 = 32500$
- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$

Kalman filter

- Prediction step. The new best estimate of x_1 should be $x_e^{new} = Fx_e$ (1)
- This is because mean value of w_t is zero. Here $x_e^{new} = Fx_e = 0.9 * 1000 = 900$
- The variance of the error of this estimate is $P^{new} = E[(x_1 - x_e^{new})^2]$. As $x_1 = Fx_0 + w_0$ and $x_e^{new} = Fx_e$, the variance of error is $P^{new} = E[(x_1 - x_e^{new})^2] = E[(Fx_0 + w_0 - Fx_e)^2] = E[F^2(x_0 - x_e)^2] + E[w_0^2] + 2E[F(x_0 - x_e)w_0]$, but w_0 is uncorrelated with x_0 and x_e, so
$$P^{new} = F^2E[(x_0 - x_e)^2] + E[w_0^2]$$
- As $E[(x_0 - x_e)^2] = P$ and $E[w_0^2] = Q$, then $P^{new} = F^2P + Q$ (2)
- For our example $P^{new} = F^2P + Q = 0.81 * 40000 + 100 = 32500$
- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

The noisy estimations from our sensor

The our prediction of what the measurement should be according to our model

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- In our example $(z_1 - z_e) = 1200 - 900 = 300$. Let's find optimal K using optimization technique

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = E[(x_1 - x_e^{newer})^2] =$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = E[(x_1 - x_e^{newer})^2] =$

$$=E\left[\left(x_1 - x_e^{new} - K(z_1 - z_e)\right)^2\right] = E\left[\left(x_1 - x_e^{new} - K(Hx_1 + v_1 - H * x_e^{new})\right)^2\right] =$$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1-K)z_e$

$$x_e^{newer} = Kz_1 + (1-K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = E[(x_1 - x_e^{newer})^2] =$

$$=E\left[\left(x_1 - x_e^{new} - K(z_1 - z_e)\right)^2\right] = E\left[\left(x_1 - x_e^{new} - K(Hx_1 + v_1 - H * x_e^{new})\right)^2\right] =$$

$$=E[(1-KH)(x_1 - x_e^{new}) - Kv_1)^2] = (1-KH)^2 E[(x_1 - x_e^{new})^2] + K^2 * E[v_1^2] =$$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = E[(x_1 - x_e^{newer})^2] =$

$$=E\left[\left(x_1 - x_e^{new} - K(z_1 - z_e)\right)^2\right] = E\left[\left(x_1 - x_e^{new} - K(Hx_1 + v_1 - H * x_e^{new})\right)^2\right] =$$

$$=E[(1 - KH)(x_1 - x_e^{new}) - Kv_1)^2] = (1 - KH)^2 E[(x_1 - x_e^{new})^2] + K^2 * E[v_1^2] =$$

$$=(1 - KH)^2 P^{new} + K * R$$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = (1 - KH)^2 P^{new} + K^2 * R$ (5)

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1-K)z_e$
$$x_e^{newer} = Kz_1 + (1-K)Hx_e^{new}$$
$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$
- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = (1-KH)^2 P^{new} + K^2 * R$ (5)
- Let's find optimal K using optimization technique. We want to minimize the estimation error:
$$P^{newer} = (1-KH)^2 P^{new} + K^2 * R \to \min_K$$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1-K)z_e$

$$x_e^{newer} = Kz_1 + (1-K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = (1-KH)^2 P^{new} + K^2 * R$ (5)
- Let's find optimal K using optimization technique. We want to minimize the estimation error:

$$P^{newer} = (1-KH)^2 P^{new} + K^2 * R \to \min_K$$

FOC: $\frac{d}{dk} P^{newer} = 0 \Leftrightarrow -2HP^{new}(1-KH) + 2RK = 0 \Leftrightarrow K = \frac{HP^{new}}{H^2P^{new}+R}$ (4)

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1-K)z_e$

$$x_e^{newer} = Kz_1 + (1-K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- K is a number called the Kalman gain.
- Let us compute the variance of the resulting error: $P^{newer} = (1-KH)^2 P^{new} + K^2 * R$ (5)
- Let's find optimal K using optimization technique. We want to minimize the estimation error:

$$P^{newer} = (1-KH)^2 P^{new} + K^2 * R \to \min_K$$

- In our example $K = \frac{HP^{new}}{H^2P^{new}+R} = \frac{1*32500}{1^2*32500+10000} = 0.7647$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1-K)z_e$

$$x_e^{newer} = Kz_1 + (1-K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- In our example Kalman gain $K = 0.7647$
- So $x_e^{newer} = x_e^{new} + K(z_1 - z_e) = 1129$ and $P^{newer} = (1 - KH)^2 P^{new} + K^2 * R = 7647$

Kalman filter

- Correction step. We make a noisy measurement y of x according to the linear equation $z_t = Hx_t + v_t$, here v_t is a white noise $N(0, R), R = 10000$, H =1, so $z_1 = Hx_1 + v_1$.
- Suppose $z_1 = 1200$, but if we wanted to estimate z_1 before we look at the measured value we would use $z_e = H * x_e^{new} = 1 * 900 = 900$
- The new best estimate of z_1 is given by $x_e^{newer} = Kz_1 + (1 - K)z_e$

$$x_e^{newer} = Kz_1 + (1 - K)Hx_e^{new}$$

$$x_e^{newer} = x_e^{new} + K(z_1 - z_e) \quad (3)$$

- In our example Kalman gain $K = 0.7647$
- So $x_e^{newer} = x_e^{new} + K(z_1 - z_e) = 1129$ and $P^{newer} = (1 - KH)^2 P^{new} + K^2 * R = 7647$
- These are the five equations of the Kalman filter.
- At time t=2, we start again using x_e^{newer} as the value of x_e^1 in $x_e^{new,2} = Fx_e^1$ (1) and P^{newer} as the value of P in $P^{new} = F^2P + Q$ (2). Then we calculate K from equation 4 and use that along with the new measurement z_2 in equation 3 to get another estimate of x and we use equation 5 to get the corresponding P. And repeat these steps….

Kalman filter in simplest example

- Prediction step:
  - *Predicted (a priori) state estimate* $x_e^{new,k} = F x_e^{newer,k-1}$ *(1)*
  - Predicted (*a priori*) estimate covariance $P^{new,k} = F^2 P^{newer,k-1} + Q$ *(2)*
- Correction step:
  - Updated (*a posteriori*) state estimate $x_e^{newer,k} = x_e^{new,k} + K\left(z_k - H x_e^{new,k}\right)$ (3)
  - *Optimal* Kalman gain $K = \frac{HP^{new,k}}{H^2 P^{new,k} + R}$ *(4)*
  - Updated (*a posteriori*) estimate covariance $P^{newer,k} = (1 - KH)^2 P^{new,k} + K^2 * R$ *(5)*

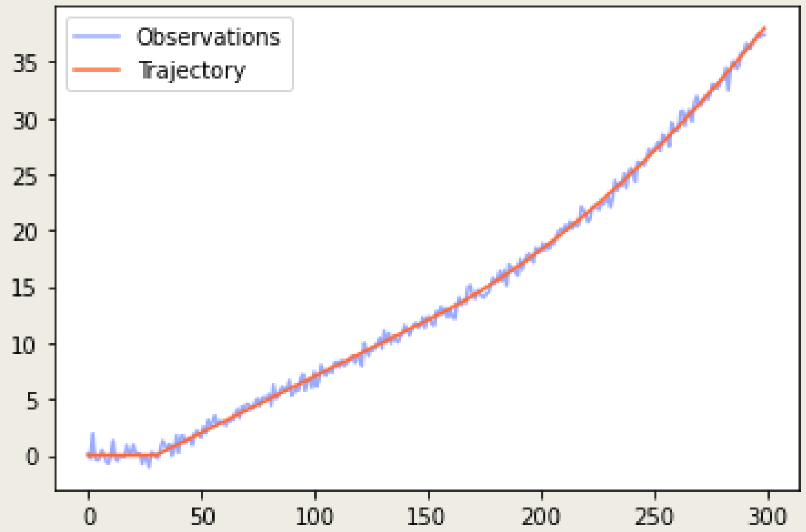
Kalman filter in simplest example

- Prediction step:
  - *Predicted (a priori) state estimate* $x_e^{new,k} = Fx_e^{newer,k-1}$ *(1)*
  - Predicted (*a priori*) estimate covariance $P^{new,k} = F^2 P^{newer,k-1} + Q$ *(2)*
- Correction step:
  - Updated (*a posteriori*) state estimate $x_e^{newer,k} = x_e^{new,k} + K\left(z_k - Hx_e^{new,k}\right)$ (3)
  - *Optimal* Kalman gain $K = \frac{HP^{new,k}}{H^2P^{new,k}+R}$ *(4)*
  - Updated (*a posteriori*) estimate covariance $P^{newer,k} = (1-KH)^2 P^{new,k} + K^2 * R$ *(5)*
  - *In case of optimal Kalman gain* $P^{newer,k}$ *can be simplified:*
  - $P^{newer,k} = (1-KH)^2 P + K^2 * R = \left(1 - \frac{H^2P}{H^2P+R}\right)^2 P + \frac{H^2P^2}{(H^2P+R)^2} * R =$
  - $= \frac{R^2}{(H^2P+R)^2} P + \frac{H^2P^2}{(H^2P+R)^2} * R = \frac{PR(R+H^2P)}{(H^2P+R)^2} = \frac{PR}{R+H^2P} = (1-KH)P$

Kalman filter in general case

- Prediction step:
  - *Predicted (a priori) state estimate* $x_e^{new,k} = Fx_e^{newer,k-1}$ *(1)*
  - Predicted (*a priori*) estimate covariance $P^{new,k} = FP^{newer,k-1}F^T + Q$ *(2)*
- Correction step:
  - Updated (*a posteriori*) state estimate $x_e^{newer,k} = x_e^{new,k} + K\left(z_k - Hx_e^{new,k}\right)$ (3)
  - *Optimal* Kalman gain $K = P^{new,k}H^T\left(HP^{new,k}H^T + R\right)^{-1}$ *(4)*
  - Updated (*a posteriori*) estimate covariance $P^{newer} = (I - KH)P^{new}$ *(5)*

Dynamic linear models



- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system:*
    - The state of the system is complex: the position of the drone, its velocity and acceleration, so $x_t = (s_t, v_t, a_t)^T$;
  - *F is a state transition matrix:*
    - In case of movement with constant acceleration $x_t = x_0 + v_0 t + \frac{at^2}{2}$, so $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*
- Observation model: $z_t = Hx_t + v_t$,
  - *z_t is measurements made by sensors,*
  - *v_t is a vector of measurement errors, $v_t \sim N(0, R)$, R is a covariance matrix*

Dynamic linear models

- Observation model: $z_t = Hx_t + v_t$,
  - z_t *is measurements made by sensors,*
  - v_t *is a vector of measurement errors,* $v_t \sim N(0, R)$, R *is a covariance matrix*

- In many cases we can assume that the measurements do not correlate with each other.

- Then the matrix $R = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_m \end{pmatrix}$.

- It is sufficient to set the variance values for each measured parameter. Sometimes this data can be found in the documentation for the sensors used. However, if there is no reference information, you can estimate the variance by measuring a pre-known reference value with the sensor, or use the 3σ −rule.

Dynamic linear models

- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
- w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q), Q$ is a covariance matrix
- What about Q? Remember the $P^{new,k} = FP^{newer,k-1}F^T + Q$ (2)
- If you set Q very small, the uncertainty of the estimate $P^{new,k}$ will slightly increase on the prediction stage. This means that we believe that our model accurately describes the process.
- If you set Q large, the uncertainty of the estimate $P^{new,k}$ will greatly increase on the prediction stage. In this way, we show that the model may contain inaccuracies or factors that we didn't describe.
- This matrix indicates which state variables will be primarily affected by model errors or factors that we didn't describe.

Dynamic linear models

- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
- w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q), Q$ is a covariance matrix
- This matrix indicates which state variables will be primarily affected by model errors or factors that we didn't describe.
- The state is the position of the robot, its velocity and acceleration $x_t = (s_t, v_t, a_t)^T$;
- When the robot passes a bump, the sensor estimations and model predictions will start to diverge. The matrix structure will determine how the filter should responds to this discrepancy.
- We can make various assumptions about the nature of the noise. For our example, we can assume that undescribed factors (road bumps) primarily affect acceleration. The matrix is chosen so that the highest value corresponds to the highest order of the derivative a_t.

Dynamic linear models

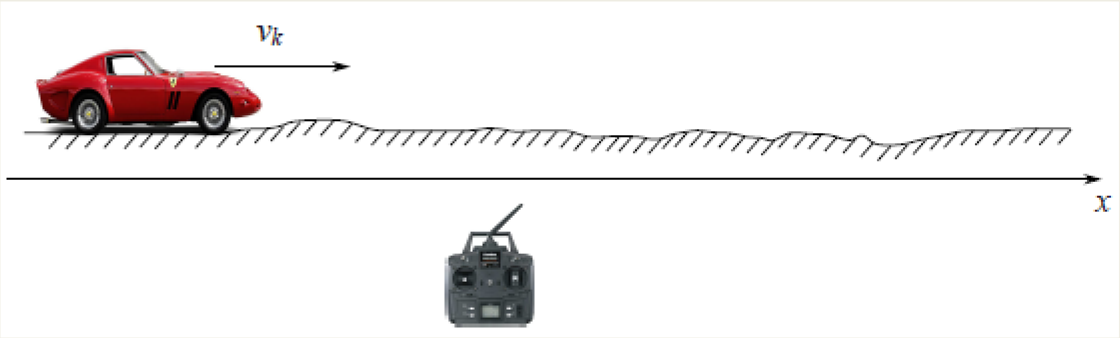
- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
- w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q), Q$ is a covariance matrix
- This matrix indicates which state variables will be primarily affected by model errors or factors that we didn't describe.
- The state is the position of the robot, its velocity and acceleration $x_t = (s_t, v_t, a_t)^T$;
- When the robot passes a bump, the sensor estimations and model predictions will start to diverge. The matrix structure will determine how the filter should responds to this discrepancy.
- We can make various assumptions about the nature of the noise. For our example, we can assume that undescribed factors (road bumps) primarily affect acceleration. The matrix is chosen so that the highest value corresponds to the highest order of the derivative a_t.
- For the simplest case we may take even $\mathrm{Q} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}$

Dynamic linear models

- Simple example: robot, one-dimensional movement
- Model for states of the system $x_t = Fx_{t-1} + w_t$
- w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix
- This matrix indicates which state variables will be primarily affected by model errors or factors that we didn't describe.
- The state is the position of the robot, its velocity and acceleration $x_t = (s_t, v_t, a_t)^T$;
- When the robot passes a bump, the sensor estimations and model predictions will start to diverge. The matrix structure will determine how the filter should responds to this discrepancy.
- In Discrete Constant White Noise Model matrix Q should be defined by intensity of noise Γ, that is the column for the highest derivative: if $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$, then $\Gamma = \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \\ 1 \end{bmatrix}$ and $Q = \Gamma \sigma_v^2 \Gamma^{\mathrm{T}}$

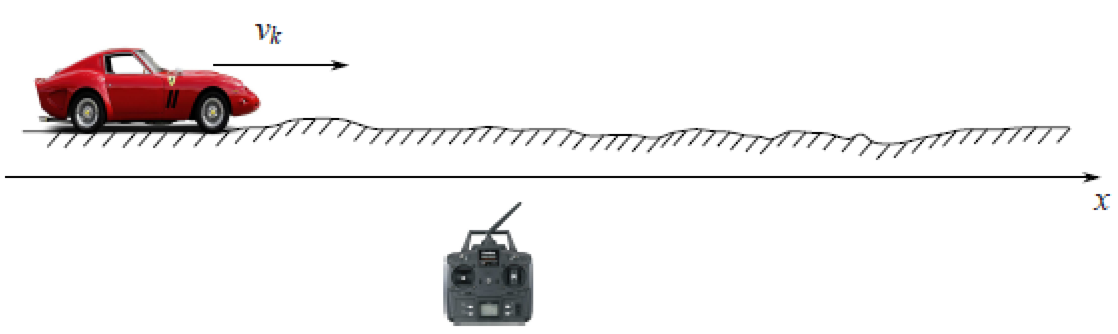
Example in Python

- Let the robot stand still for the first 20% of the time, then move with constant speed, and then start moving with constant acceleration.



Source: https://habr.com/ru/post/166693/

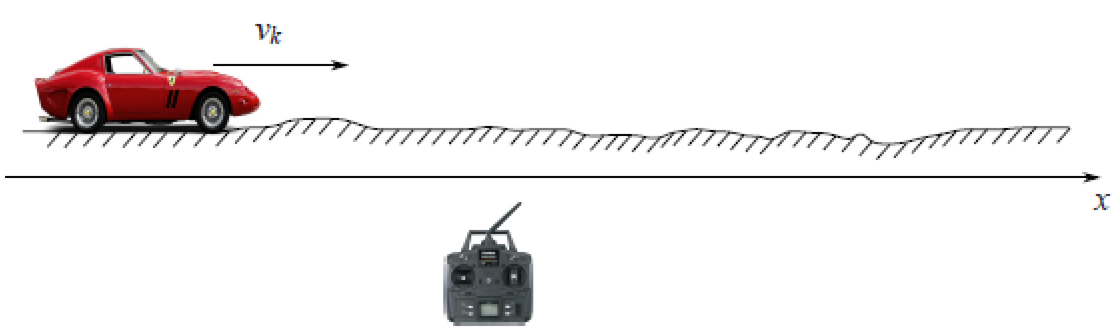
Example in Python



- First model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system, i.e. the position of the car, its velocity and acceleration, so $x_t = s_t$;*
  - *F is a state transition matrix and $F = 1$*
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*

- Observation model: $z_t = Hx_t + v_t$, where $H = [1 \quad 0 \quad 0]$, we can only measure the position by GPS
- Initial position $x_0 = [0 \quad 0 \quad 0]$ with uncertainty measure $P = 10$

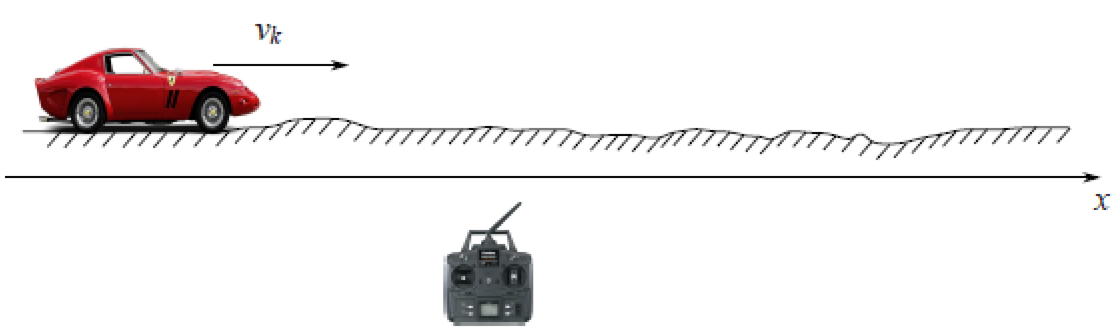
Example in Python



- Second model for states of the system $x_t = Fx_{t-1} + w_t$
  - x_t *is a current state of the system, i.e. the position of the car and its velocity, so* $x_t = (s_t, v_t)^T$;
  - F *is a state transition matrix and* $F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}$
  - w_t *is an error of the model (bumps in the road, wind, noise, etc.),* $w_t \sim N(0, Q)$, Q *is a covariance matrix,* $Q = \Gamma \sigma_v^2 \Gamma^\mathrm{T} = \begin{bmatrix} \Delta t \\ 1 \end{bmatrix} [\Delta t \quad 1] \sigma_v^2 = \begin{bmatrix} \Delta t^2 & \Delta t \\ \Delta t & 1 \end{bmatrix} \sigma_v^2$

- Observation model: $z_t = Hx_t + v_t$, where $H = [1 \quad 0]$, we can only measure the position by GPS
- Initial position $x_0 = [0 \quad 0]$ with uncertainty measure $P = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$

Example in Python



- Third model for states of the system $x_t = Fx_{t-1} + w_t$
  - *x_t is a current state of the system, i.e. the position of the car, its velocity and acceleration, $x_t = (s_t, v_t, a_t)^T$;*
  - *F is a state transition matrix and* $F = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}$
  - *w_t is an error of the model (bumps in the road, wind, noise, etc.), $w_t \sim N(0, Q)$, Q is a covariance matrix*

- Observation model: $z_t = Hx_t + v_t$, where $H = [1 \quad 0 \quad 0]$, we can only measure the position by GPS

- Initial position $x_0 = [0 \quad 0 \quad 0]$ with uncertainty measure $P = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}$

Extended Kalman filter

- Here the state transition and observation models need not be linear functions of the state but may instead be nonlinear functions of differentiable type:
  - $x_t = Fx_{t-1} + w_t$ *converts to* $x_t = f(x_{t-1}) + w_t$
  - $z_t = Hx_t + v_t$ *converts to* $z_t = h(x_t) + v_t$
- The functions f and h cannot be applied to the covariance directly. Instead a matrix of partial derivatives (the Jacobian) is computed.
- At each timestep the Jacobian is evaluated with current predicted states: $F_t = \left.\frac{\partial f}{\partial x}\right|_{x_e^{new,k}}, H_t = \left.\frac{\partial h}{\partial x}\right|_{x_e^{new,k}},$
- These matrices can be used in the Kalman filter equations. This process essentially linearizes the nonlinear function around the current estimate.

THANK YOU!

Source: Introduction to Dynamic Linear Models for Time Series Analysis, by Marko Laine, https://arxiv.org/abs/1903.11309v2