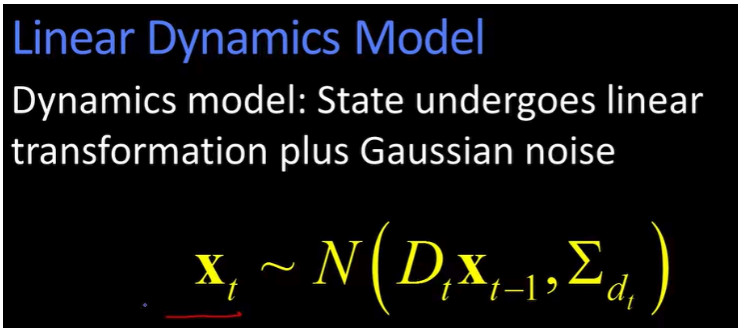
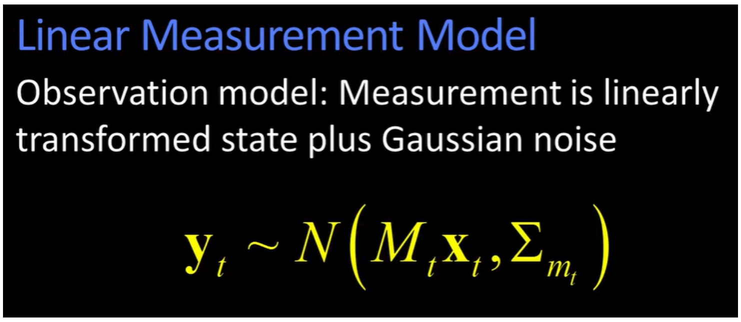
Kalman Filters

For the linear case, we have two models.

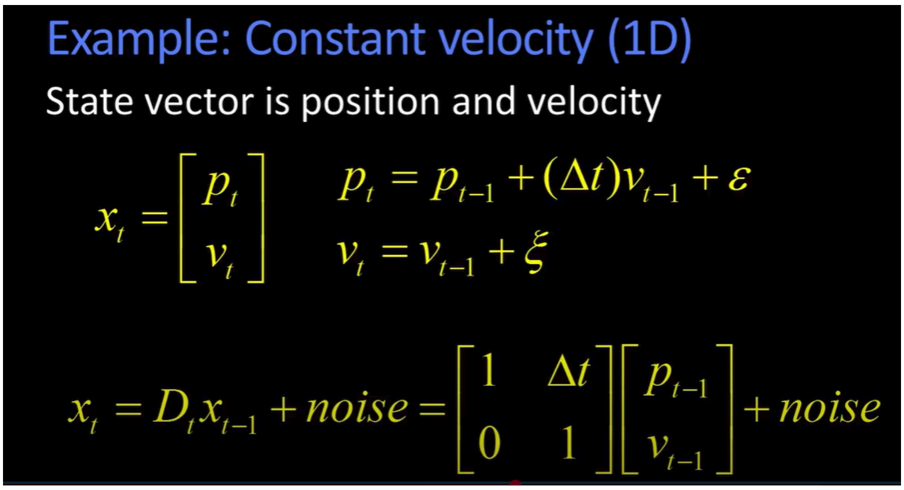
The linear dynamics model (how the state evolves over time)



The linear measurement mode



Constant Velocity (1D example)



Example: Constant velocity (1D)

Measurement is position only

$$y_t = Mx_t + noise = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} p_t \\ v_t \end{bmatrix} + noise$$

Constant Acceleration (1D example)

Example: Constant acceleration (1D)

State vector is position, velocity & acceleration

$$x_t = \begin{bmatrix} p_t \\ v_t \\ a_t \end{bmatrix} \qquad \begin{aligned} p_t &= p_{t-1} + (\Delta t)v_{t-1} + \varepsilon \\ v_t &= v_{t-1} + (\Delta t)a_{t-1} + \xi \\ a_t &= a_{t-1} + \zeta \end{aligned}$$

$$x_t = D_t x_{t-1} + noise = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ v_{t-1} \\ a_{t-1} \end{bmatrix} + noise$$

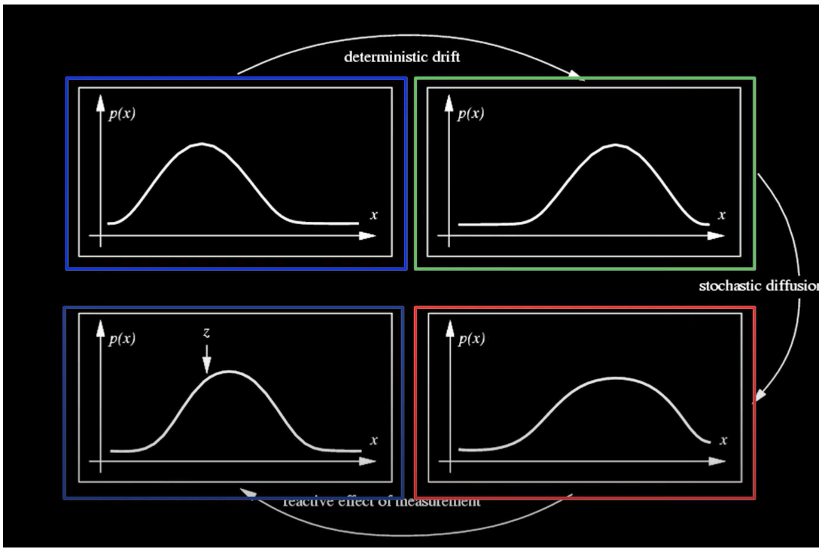
Example: Constant acceleration (1D)

Measurement is position only

$$y_t = Mx_t + noise = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_t \\ v_t \\ a_t \end{bmatrix} + noise$$

Kalman Filter

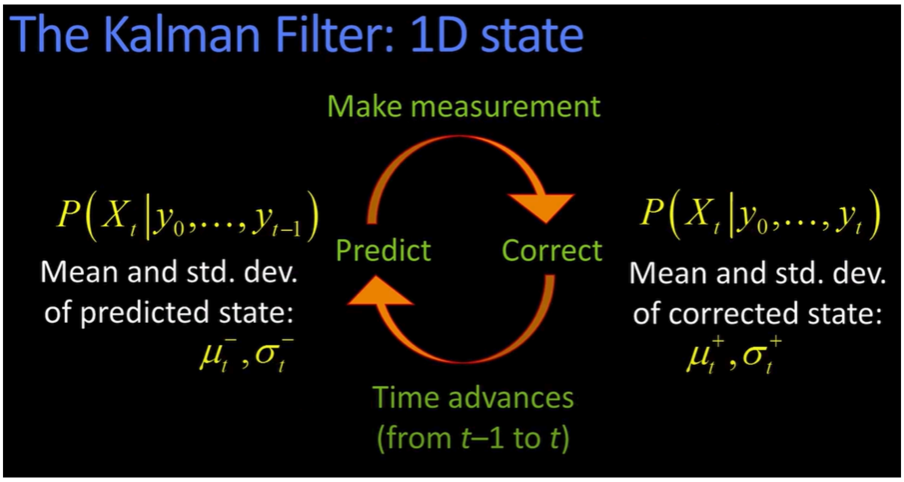
Method to track linear dynamical systems with GAUSSIAN NOISE. Both predicted and corrected states are both gaussian.



So you shift (scale mean) + expand (add noise) + measure + correct + shrink (because new information will only reduce uncertainty)

The first 3 images correspond to prediction.

Kalman Filter 1D State

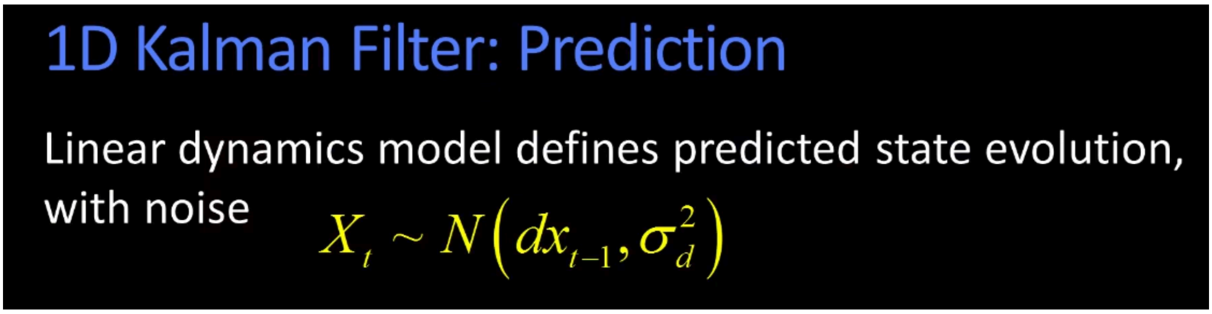


Note that prediction uses mu minus and sigma minus (mean and covariance BEFORE I take the measurement, this is based on mu plus and sigma plus at t -1)

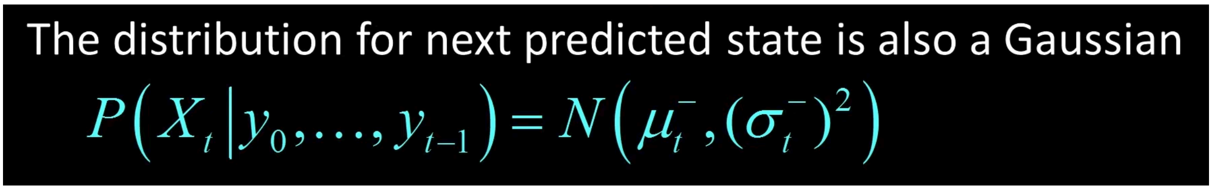
Then after you make measurement and correct mean to mu plus and sigma plus at t

1D Kalman Filter Prediction

In this linear dynamics model the next state is gotten by multiplying the previous state by a factor of d and adding some noise modeled by a gaussian.

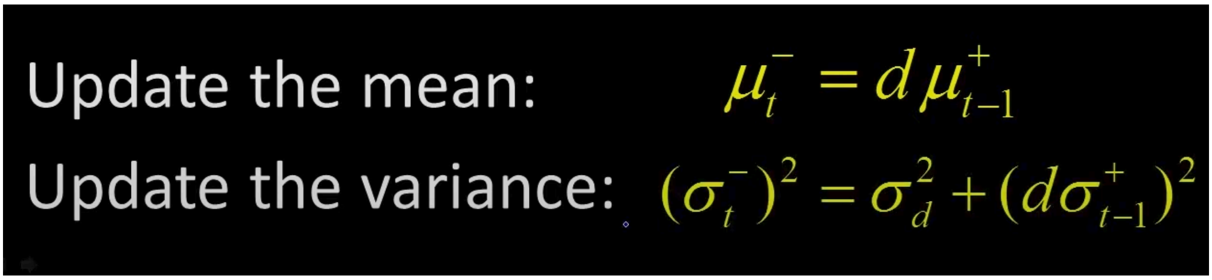


So, if that's the case, the distribution of the next predicted state is also gaussian:



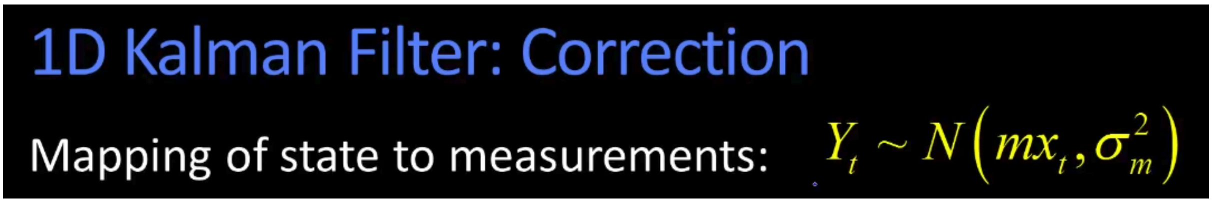
Crucially note that they are mu and sigma minus because this prediction was made before measurement

And remember that we need to update mu and sigma



1D Kalman Correction

In this linear dynamics model, the measurement is gotten by multiplying the previous state by a factor of m and adding some gaussian noise



Now, given our predicted state (that we got above), we need to correct it using the measurement!

Predicted state: $P\left(X_t \middle| y_0,\ldots,y_{t-1}\right) = N\left(\mu_t^-,(\sigma_t^-)^2\right)$

Want to estimate corrected distribution:

$$P\left(X_t \middle| y_0,\ldots,y_t\right) = \frac{P(y_t \mid X_t)P(X_t \mid y_0,\ldots,y_{t-1})}{\int P(y_t \mid X_t)P(X_t \mid y_0,\ldots,y_{t-1})dX_t}$$

Now, the corrected distribution is defined to be a new Gaussian with new mean mu plus and sigma plus.

Kalman: With linear, Gaussian dynamics and measurements, the corrected distribution to be:

$$P\left(X_t \middle| y_0,\ldots,y_t\right) \equiv N\left(\mu_t^+,(\sigma_t^+)^2\right)$$

Formula for mu plus and sigma plus

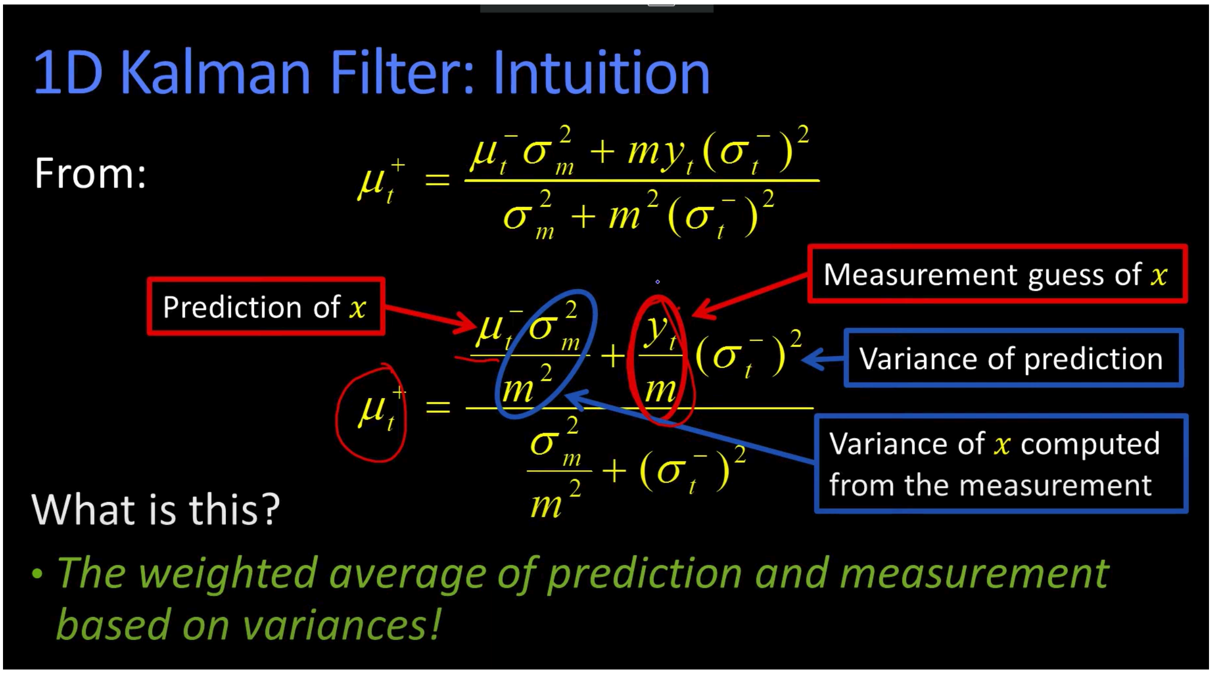
Update the mean:

$$\mu_t^+ = \frac{\mu_t^-\sigma_m^2 + m y_t(\sigma_t^-)^2}{\sigma_m^2 + m^2(\sigma_t^-)^2}$$

Update the variance:

$$(\sigma_t^+)^2 = \frac{\sigma_m^2(\sigma_t^-)^2}{\sigma_m^2 + m^2(\sigma_t^-)^2}$$

1D Kalman Intuition



Weights Based on Variances: The crucial insight is how the weights are determined. They are inversely proportional to the uncertainty (variance) of each estimate:

- The weight for the prediction (μ_t^-) is proportional to the variance of the measurement (σ_m^2).
- The weight for the measurement-derived estimate (y_t/m) is proportional to the variance of the prediction ($(\sigma_t^-)^2$).

This means:

- If the measurement is very certain (small σ_m^2), the weight for the prediction is smaller, and the weight for the measurement is larger. The filter trusts the measurement more.
- If the prediction is very certain (small $(\sigma_t^-)^2$), the weight for the measurement is smaller, and the weight for the prediction is larger. The filter trusts the prediction more.

Special Cases

Prediction vs. correction

$$\mu_t^+ = \frac{\mu_t^- \sigma_m^2 + m y_t (\sigma_t^-)^2}{\sigma_m^2 + m^2 (\sigma_t^-)^2} \qquad (\sigma_t^+)^2 = \frac{\sigma_m^2 (\sigma_t^-)^2}{\sigma_m^2 + m^2 (\sigma_t^-)^2}$$

What if there is no prediction uncertainty? $(\sigma_t^- = 0)$

$$\mu_t^+ = \mu_t^- \qquad (\sigma_t^+)^2 = 0$$

The measurement is ignored!

What if there is no measurement uncertainty? $(\sigma_m = 0)$

$$\mu_t^+ = \frac{y_t}{m} \qquad (\sigma_t^+)^2 = 0$$

Simplification of Mu+ and Sigma+

1D Kalman Filter: Intuition

Also: $$\mu_t^+ = \frac{\frac{\mu_t^- \sigma_m^2}{m^2} + \frac{y_t}{m}(\sigma_t^-)^2}{\frac{\sigma_m^2}{m^2} + (\sigma_t^-)^2}$$

$$\mu_t^+ = \frac{a\mu_t^- + b\frac{y_t}{m}}{a+b} = \frac{(a+b)\mu_t^- + b(\frac{y_t}{m} - \mu_t^-)}{a+b}$$

$$\mu_t^+ = \frac{a\mu_t^- + b\frac{y_t}{m}}{a+b} = \frac{(a+b)\mu_t^- + b(\frac{y_t}{m} - \mu_t^-)}{a+b}$$

$$\mu_t^+ = \mu_t^- + \frac{b(\frac{y_t}{m} - \mu_t^-)}{a+b} = \mu_t^- + k(\frac{y_t}{m} - \mu_t^-)$$

Mu t - is the prediction. The difference (red) ika the residual. The residual is the difference between yt/m (the measurements guess of x) and mu t -. Something like TD updates.

N-Dimensional Kalman Filters (the real deal)

PREDICT

$$x_t^- = D_t x_{t-1}^+$$

$$\Sigma_t^- = D_t \Sigma_{t-1}^+ D_t^T + \Sigma_{d_t}$$

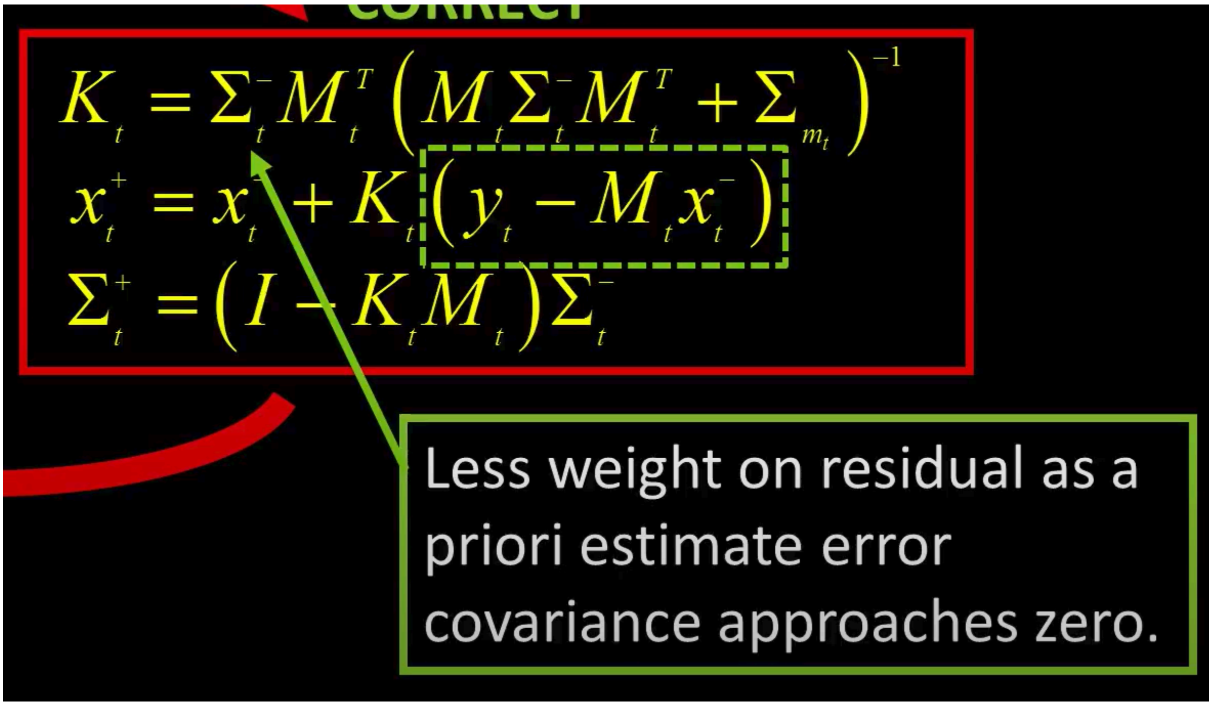
CORRECT

$$K_t = \Sigma_t^- M_t^T \left(M_t \Sigma_t^- M_t^T + \Sigma_{m_t} \right)^{-1}$$

$$x_t^+ = x_t^- + K_t \left(y_t - M_t x_t^- \right)$$

$$\Sigma_t^+ = \left(I - K_t M_t \right) \Sigma_t^-$$

Σt−: This is the **predicted state covariance matrix** at time t. It represents the uncertainty in the predicted state xt−. The covariance matrix describes the variances of each state variable (diagonal elements) and the covariances between them (off-diagonal elements).



K_t is Kalman Gain The green box is the residual

Prof Amir's slide