Lecture 10: Observers and Kalman Filters

CS 344R: Robotics

Benjamin Kuipers

Stochastic Models of an Uncertain World

$$\begin{aligned}\dot{\mathbf{x}} &= F(\mathbf{x},\mathbf{u}) \\ \mathbf{y} &= G(\mathbf{x})\end{aligned} \quad \Rightarrow \quad \begin{aligned}\dot{\mathbf{x}} &= F(\mathbf{x},\mathbf{u},\varepsilon_1) \\ \mathbf{y} &= G(\mathbf{x},\varepsilon_2)\end{aligned}$$

- Actions are uncertain.
- Observations are uncertain.
- $\varepsilon_i \sim N(0,\sigma_i)$ are random variables

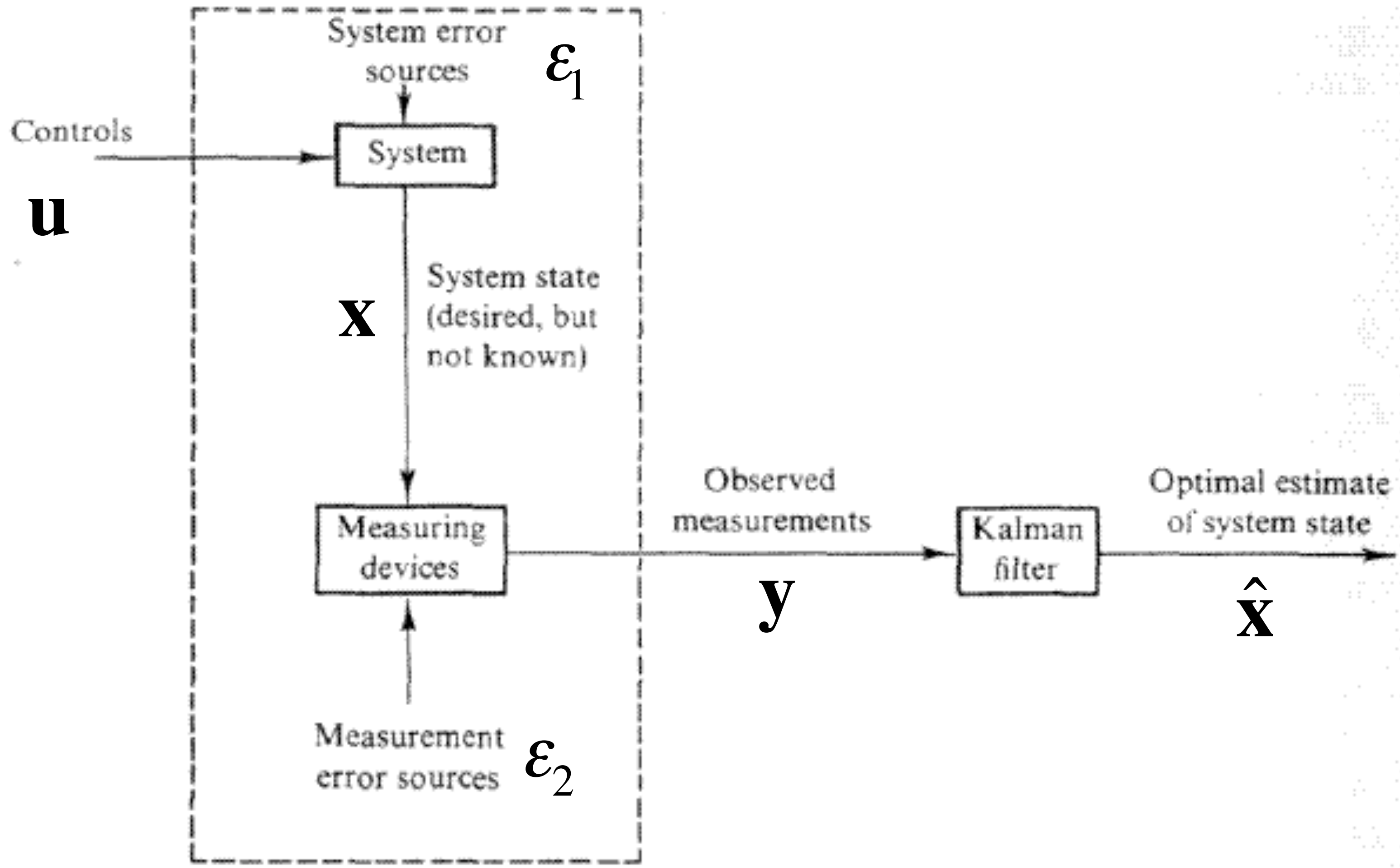
Observers

$$\begin{aligned} \dot{\mathbf{x}} &= F(\mathbf{x}, \mathbf{u}, \varepsilon_1) \\ \mathbf{y} &= G(\mathbf{x}, \varepsilon_2) \end{aligned}$$

- The state $\mathbf{x}$ is unobservable.
- The sense vector $\mathbf{y}$ provides noisy information about $\mathbf{x}$.
- An *observer* $\hat{\mathbf{x}} = Obs(\mathbf{y})$ is a process that uses sensory history to estimate $\mathbf{x}$.
- Then a control law can be written

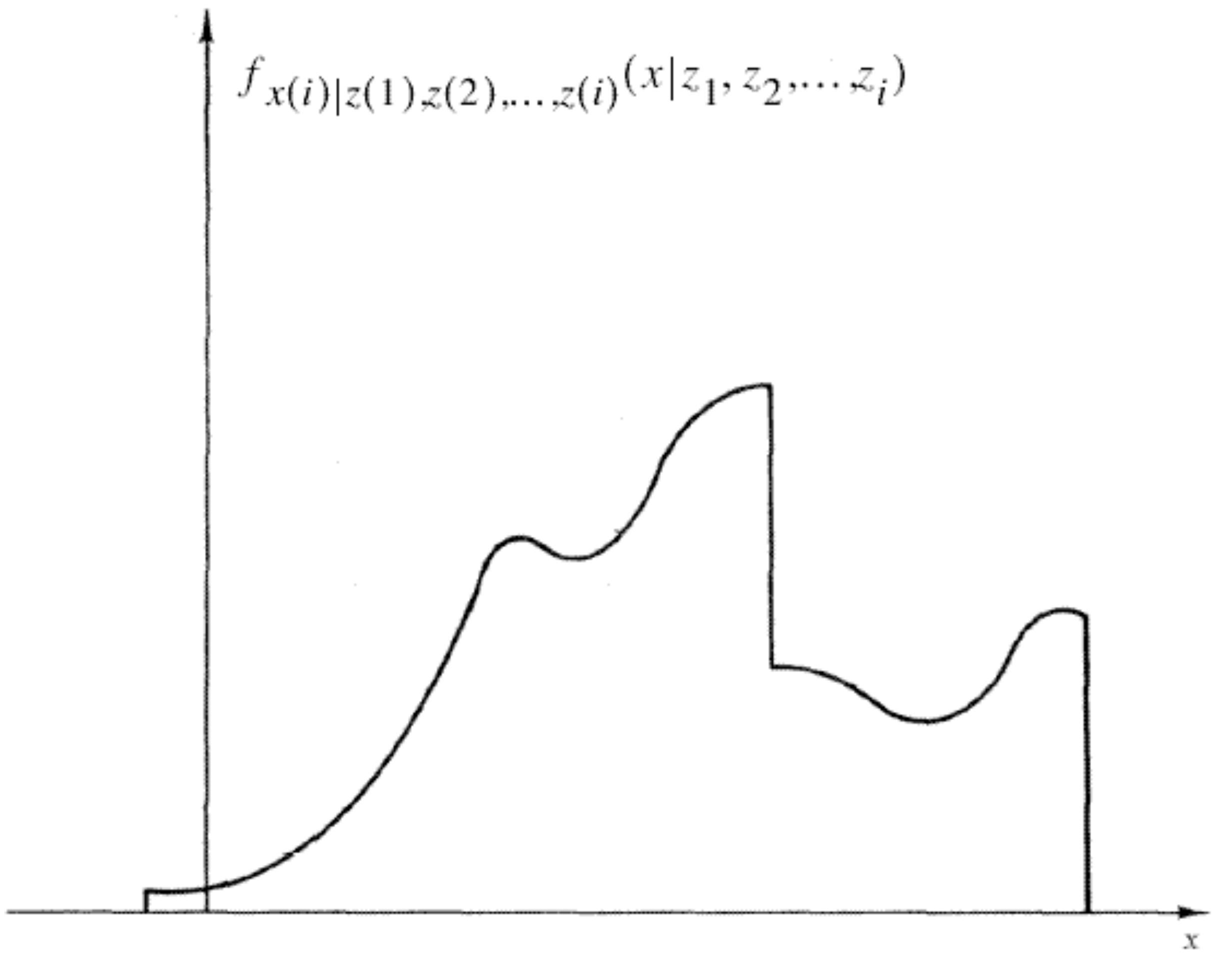
$$\mathbf{u} = H_i(\hat{\mathbf{x}})$$

Kalman Filter: Optimal Observer



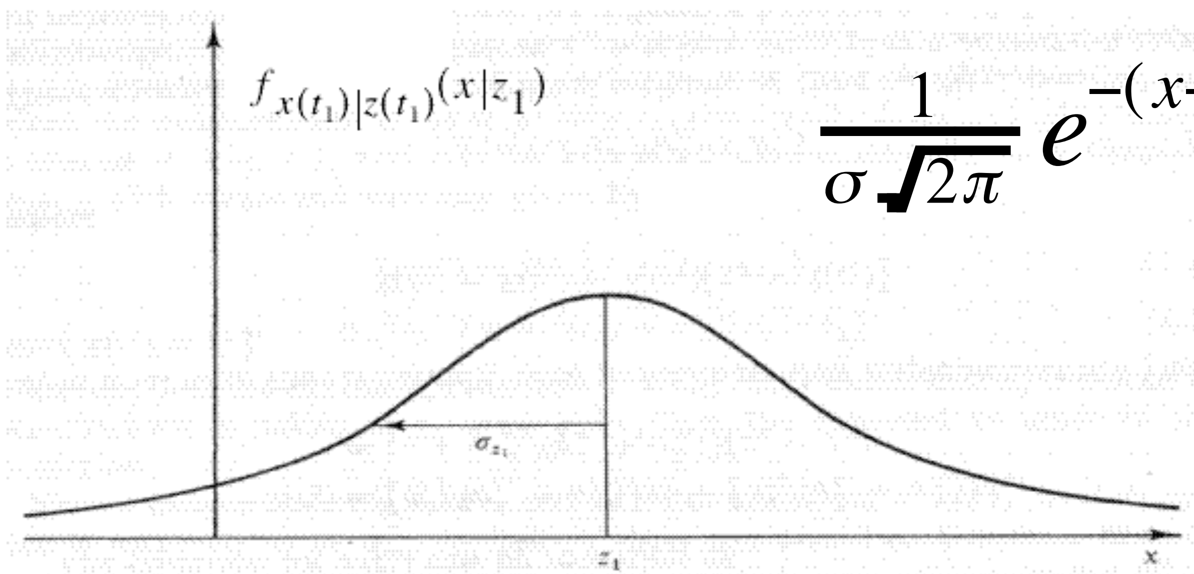
Estimates and Uncertainty

- Conditional probability density function



Gaussian (Normal) Distribution

- Completely described by $N(\mu,\sigma)$
  - Mean μ
  - Standard deviation σ, variance σ^2



$$\frac{1}{\sigma\sqrt{2\pi}}\,e^{-(x-\mu)^2/2\sigma^2}$$

The Central Limit Theorem

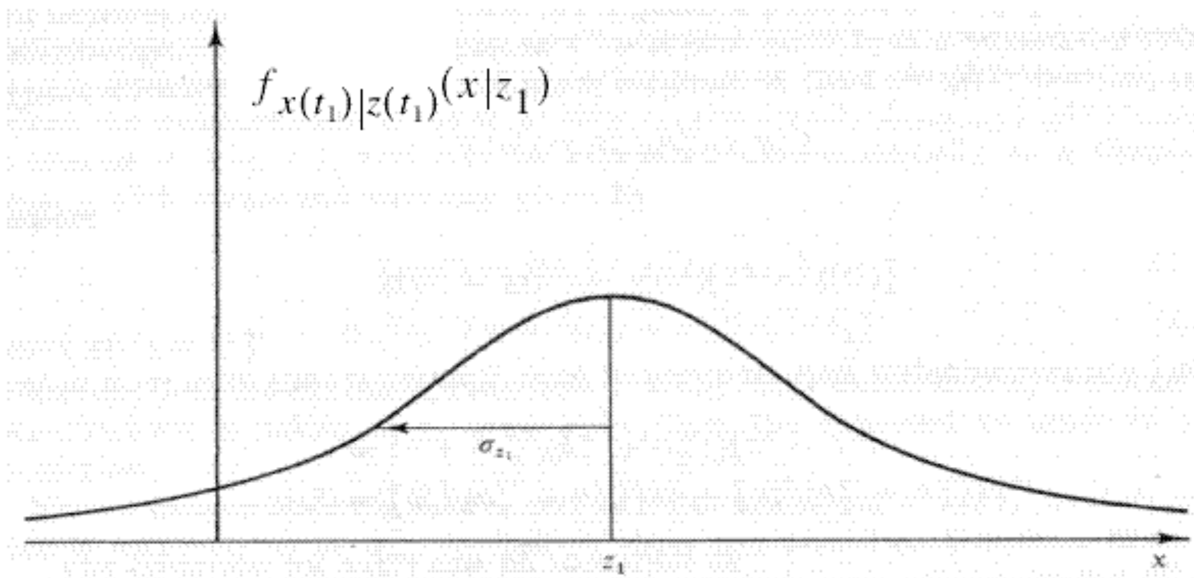
- The sum of many random variables
  - with the same mean, but
  - with arbitrary conditional density functions,

  converges to a Gaussian density function.

- If a model omits many small unmodeled effects, then the resulting error should converge to a Gaussian density function.

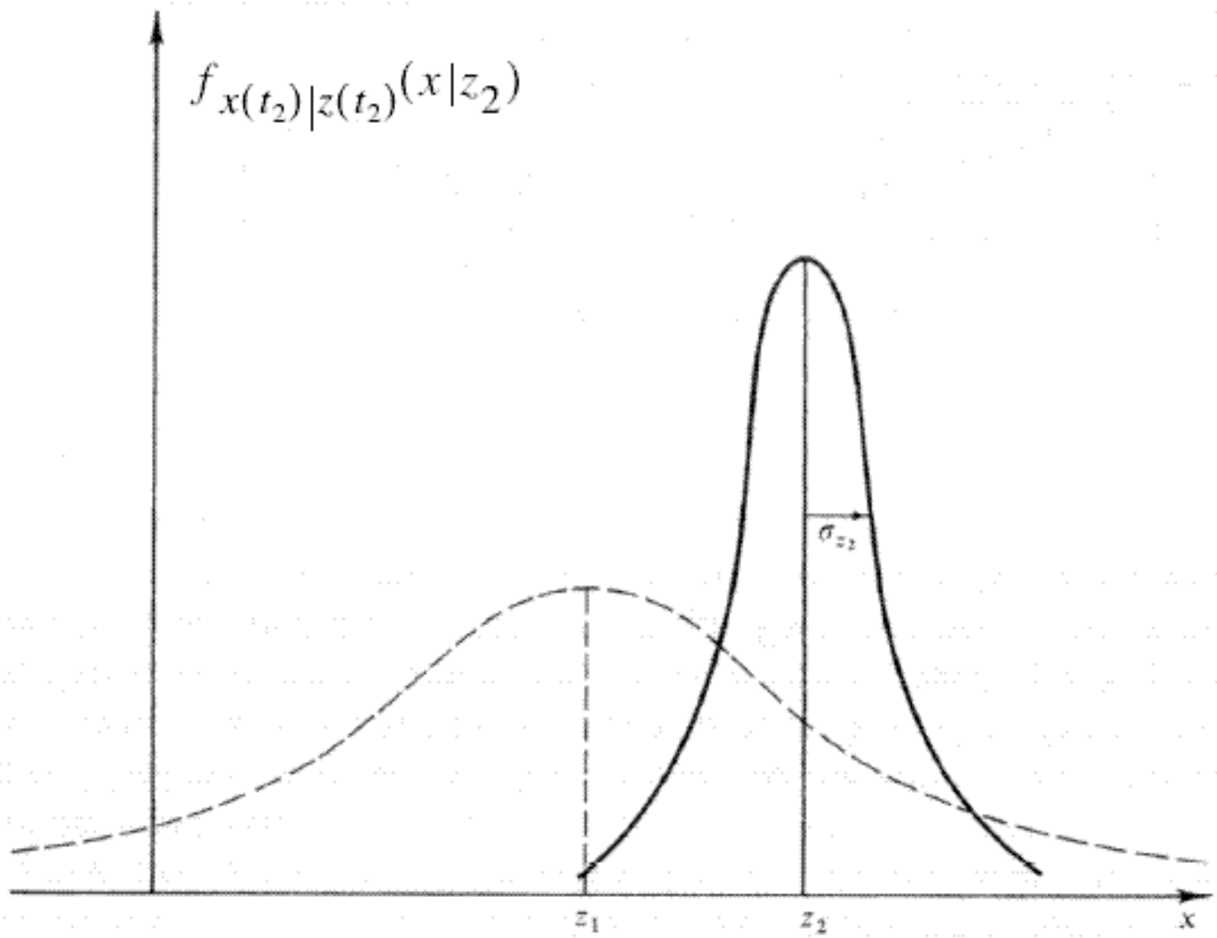
Estimating a Value

- Suppose there is a *constant* value x.
  - Distance to wall; angle to wall; etc.
- At time t_1, observe value z_1 with variance σ_1^2
- The optimal estimate is $\hat{x}(t_1) = z_1$ with variance σ_1^2

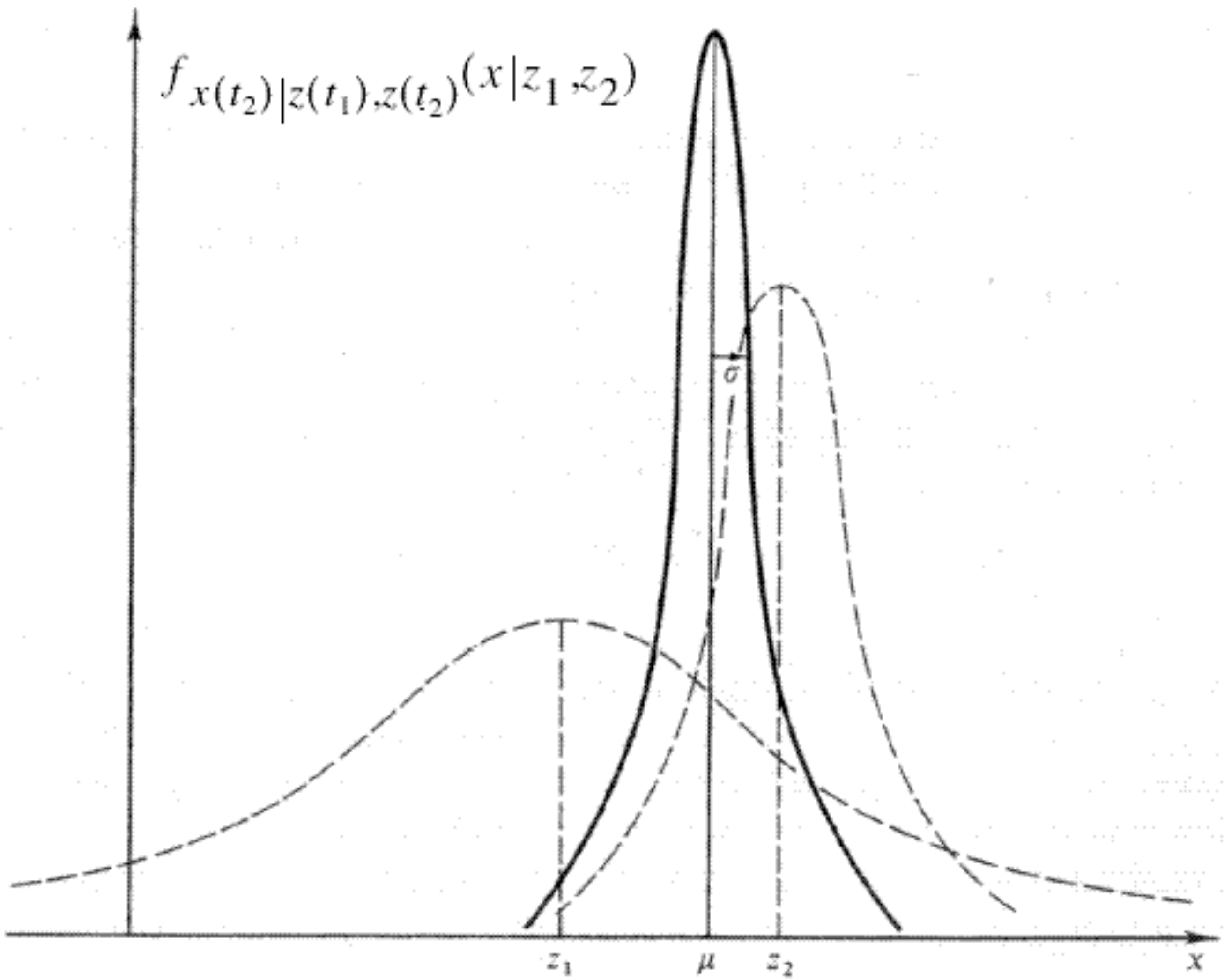


A Second Observation

- At time t_2, observe value z_2 with variance σ_2^2



Merged Evidence



Update Mean and Variance

- Weighted average of estimates.

$$\hat{x}(t_2) = Az_1 + Bz_2 \qquad\qquad A + B = 1$$

- The weights come from the variances.
  - Smaller variance = more certainty

$$\hat{x}(t_2) = \left[\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right] z_1 + \left[\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right] z_2$$

$$\frac{1}{\sigma^2(t_2)} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

From Weighted Average to Predictor-Corrector

- Weighted average:

$$\hat{x}(t_2) = Az_1 + Bz_2 = (1 - K)z_1 + Kz_2$$

- Predictor-corrector:

$$\begin{aligned}\hat{x}(t_2) &= z_1 + K(z_2 - z_1)\\ &= \hat{x}(t_1) + K(z_2 - \hat{x}(t_1))\end{aligned}$$

- This version can be applied "recursively".

Predictor-Corrector

- Update best estimate given new data

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)(z_2 - \hat{x}(t_1))$$

$$K(t_2) = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

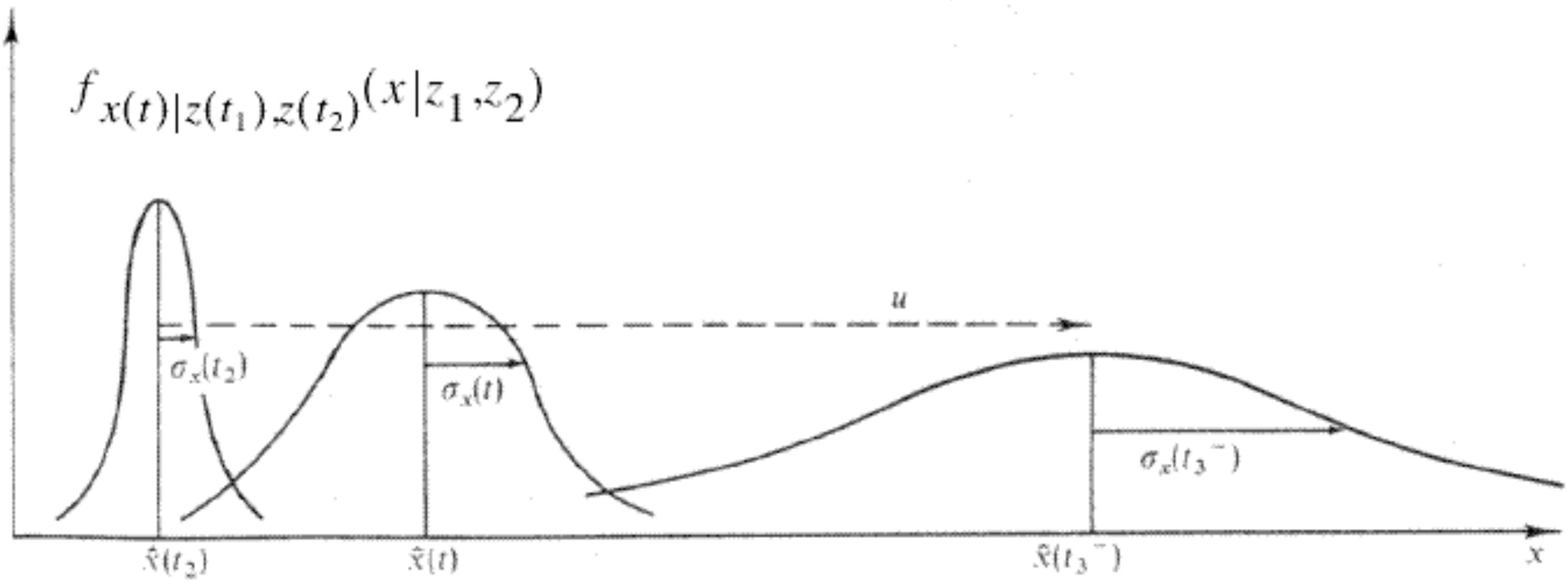
- Update variance:

$$\sigma^2(t_2) = \sigma^2(t_1) - K(t_2)\sigma^2(t_1)$$

$$= (1 - K(t_2))\,\sigma^2(t_1)$$

Static to Dynamic

- Now suppose *x* changes according to
  $\dot{x} = F(x, u, \varepsilon) = u + \varepsilon \qquad (N(0, \sigma_\varepsilon))$



Dynamic Prediction

- At t_2 we know $\hat{x}(t_2)$ $\sigma^2(t_2)$
- At t_3 after the change, before an observation.

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma^2(t_3^-) = \sigma^2(t_2) + \sigma_\varepsilon^2[t_3 - t_2]$$

- Next, we correct this prediction with the observation at time t_3.

Dynamic Correction

- At time t_3 we observe z_3 with variance σ_3^2
- Combine prediction with observation.

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)(z_3 - \hat{x}(t_3^-))$$

$$\sigma^2(t_3) = (1 - K(t_3))\sigma^2(t_3^-)$$

$$K(t_3) = \frac{\sigma^2(t_3^-)}{\sigma^2(t_3^-) + \sigma_3^2}$$

Qualitative Properties

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)(z_3 - \hat{x}(t_3^-))$$

$$K(t_3) = \frac{\sigma^2(t_3^-)}{\sigma^2(t_3^-) + \sigma_3^2}$$

- Suppose measurement noise σ_3^2 is large.
  - Then $K(t_3)$ approaches 0, and the measurement will be mostly ignored.
- Suppose prediction noise $\sigma^2(t_3^-)$ is large.
  - Then $K(t_3)$ approaches 1, and the measurement will dominate the estimate.

Kalman Filter

- Takes a stream of observations, and a dynamical model.
- At each step, a weighted average between
  - prediction from the dynamical model
  - correction from the observation.
- The Kalman gain $K(t)$ is the weighting,
  - based on the variances $\sigma^2(t)$ and σ_ε^2
- With time, $K(t)$ and $\sigma^2(t)$ tend to stabilize.

Simplifications

- We have only discussed a one-dimensional system.
  – Most applications are higher dimensional.
- We have assumed the state variable is observable.
  – In general, sense data give indirect evidence.

$$\dot{x} = F(x, u, \varepsilon_1) = u + \varepsilon_1$$
$$z = G(x, \varepsilon_2) = x + \varepsilon_2$$

- We will discuss the more complex case next.