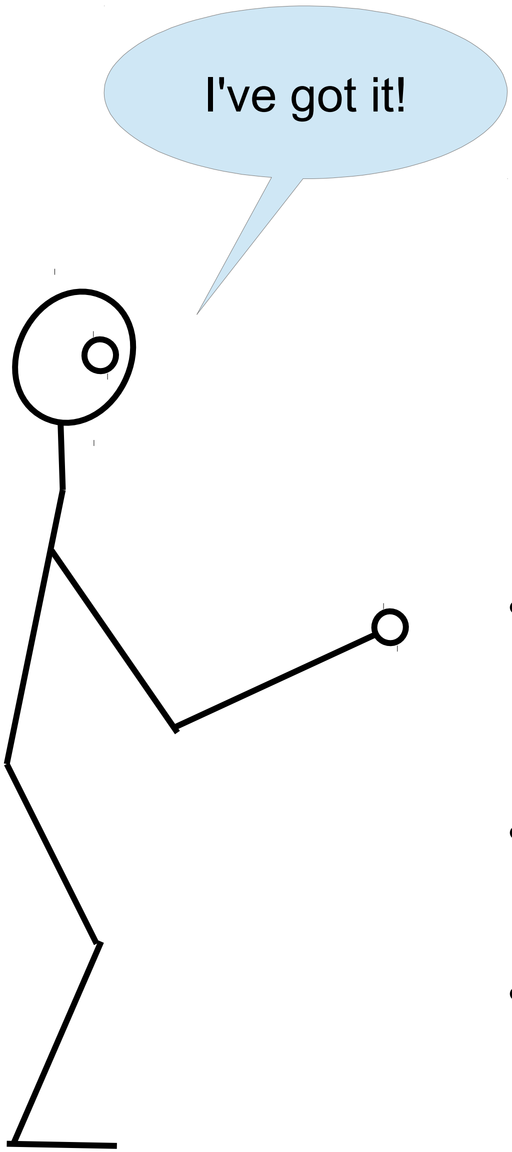
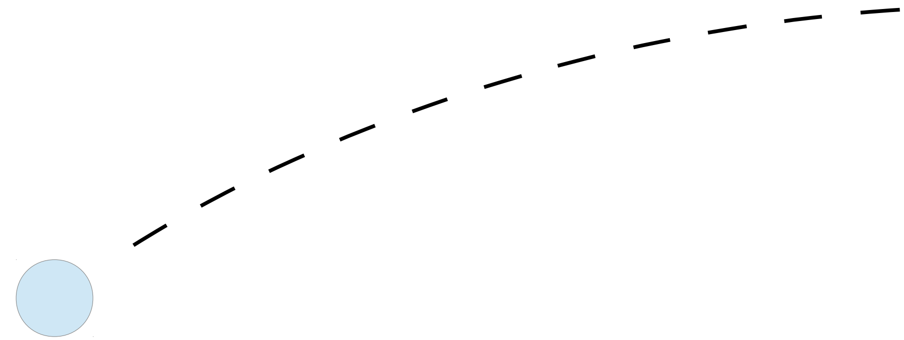
Kalman Filter





- Sequential Bayes Filtering is a general approach to state estimation that gets used all over the place.
- But, implementations like histogram filters or Kalman filters are computationally complex.
- Is it always this way? Is Bayes filtering ever simple?

Why Kalman filtering?

In general: state estimation

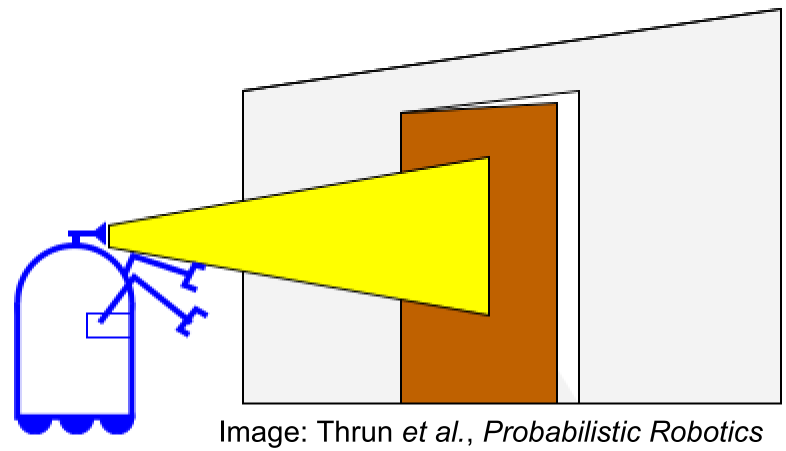


Image: Thrun *et al.*, *Probabilistic Robotics*

Kalman filters are particularly well suited for tracking moving objects

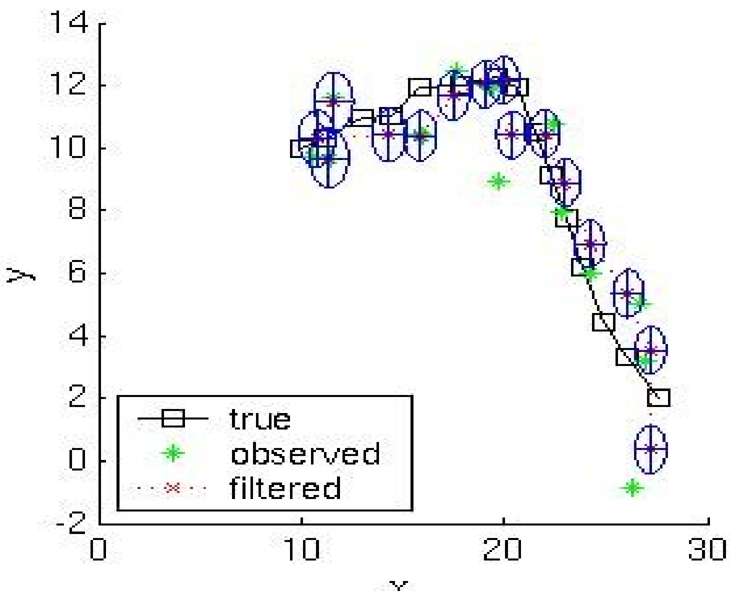
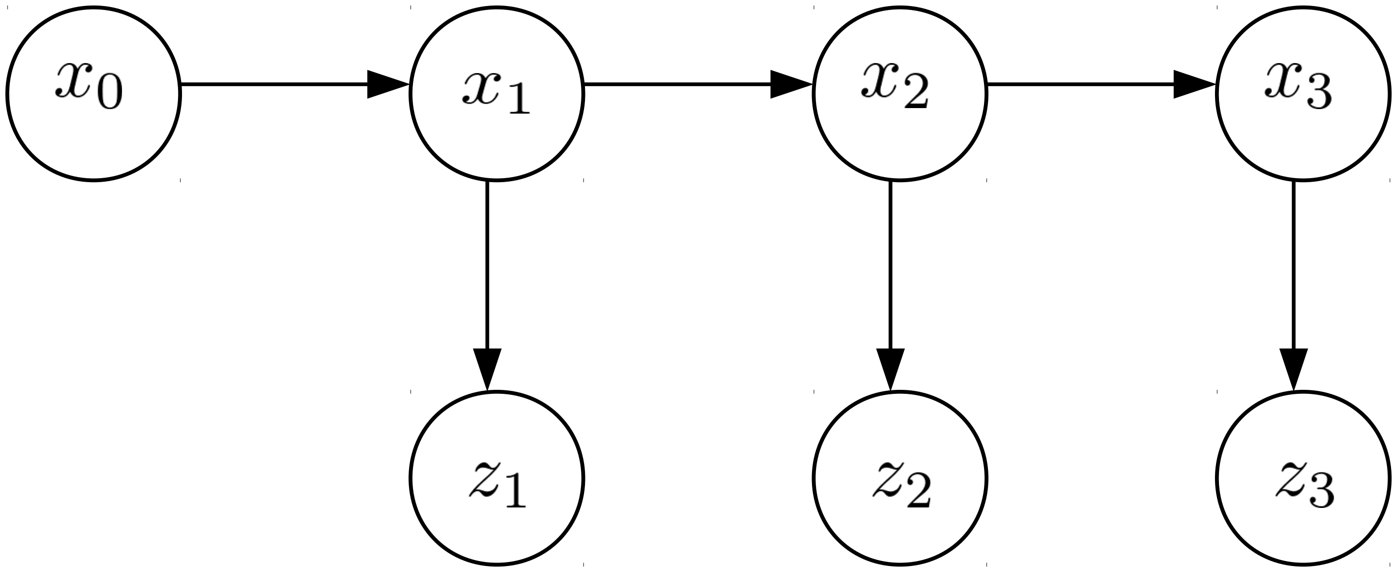


Image: UBC, Kevin Murphy Matlab toolbox

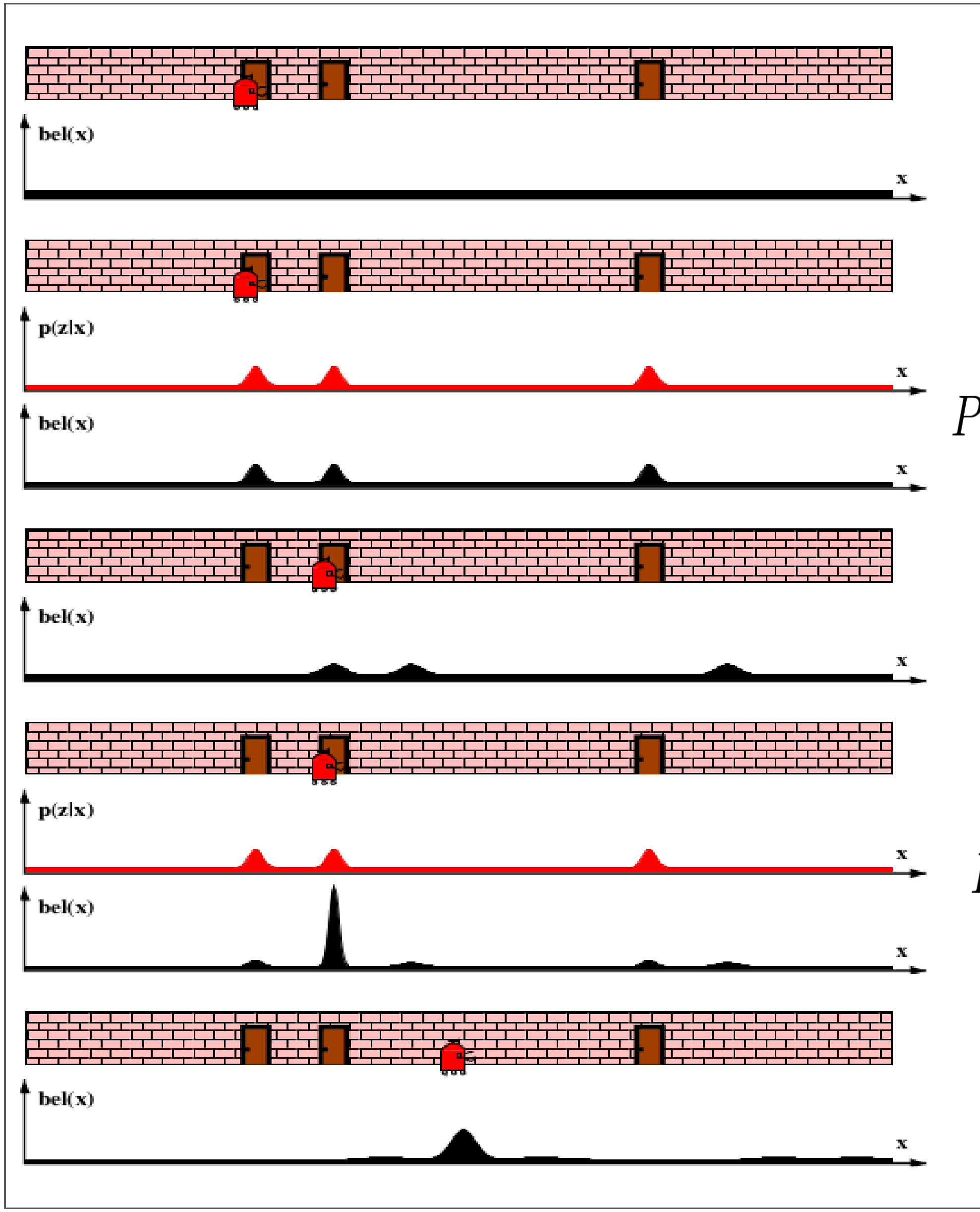
Review of sequential Bayes filtering



Process update: $$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

Measurement update: $$P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1})P(x_{t+1}|z_{0:t})$$

Review of sequential Bayes filtering



$$P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1}) P(x_{t+1}|z_{0:t})$$

$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t, z_{0:t}) P(x_t|z_{0:t})$$

$$P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1}) P(x_{t+1}|z_{0:t})$$

$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t, z_{0:t}) P(x_t|z_{0:t})$$

Image: Thrun *et al.*, *Probabilistic Robotics*

Transition function

Recall the state transition function:

$$P(x_{t+1}|x_t)$$

Next state Current state

– this probability can be expressed as a table:

$P(x_{t+1}\|x_t)$	x_t

Can also be expressed as a function:

$$x_{t+1} = f(x_t) + w_t \qquad\qquad w_t \sim N(w_t|0, Q)$$

or: $$P(x_{t+1}|x_t) = N(x_{t+1}|f(x_t), Q)$$

Linear system

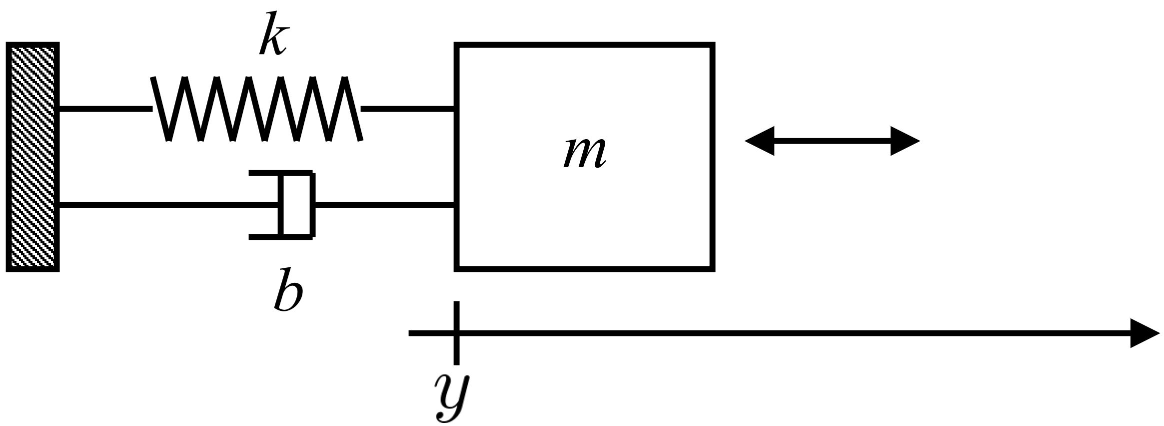
A linear system is any system where:

$$P(x_{t+1}|x_t) = N(x_{t+1}|Ax_t, Q)$$

(technically, this is a linear Gaussian system)

Also written: $$x_{t+1} = Ax_t + w_t \qquad w_t \sim N(w_t|0, Q)$$

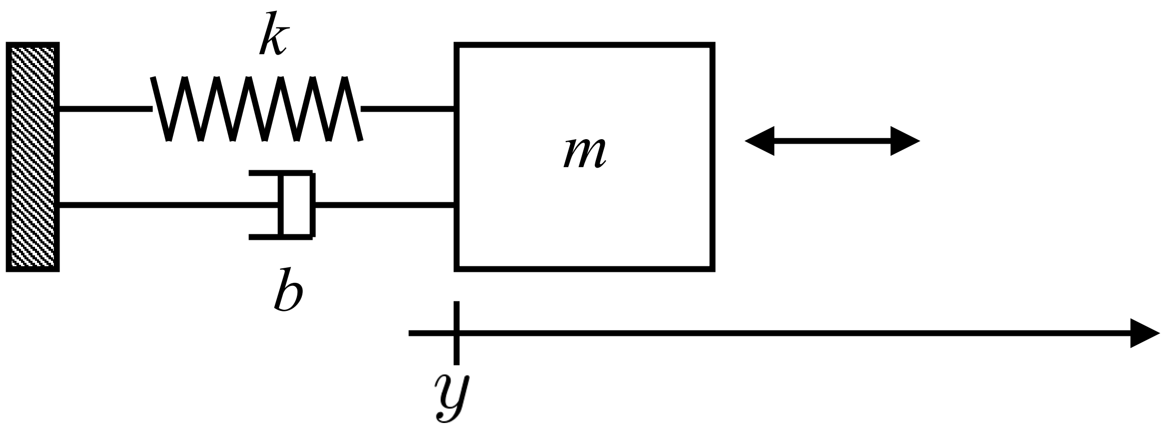
Linear system: Example



Equation of motion: $0 = m\ddot{y} + b\dot{y} + ky$

How get it into this form? $x_{t+1} = Ax_t + w_t$

Linear system: Example



Equation of motion: $0 = m\ddot{y} + b\dot{y} + ky$

How get it into this form? $x_{t+1} = Ax_t + w_t$

Integrate forward one timestep:

$$y_{t+1} = y_t + \dot{y}dt$$

$$\dot{y}_{t+1} = \dot{y}_t + \ddot{y}dt$$

$$x = \begin{pmatrix} y \\ \dot{y} \end{pmatrix}$$

$$\begin{pmatrix} y_{t+1} \\ \dot{y_{t+1}} \end{pmatrix} = \begin{pmatrix} 1 & dt \\ -\frac{k}{m}dt & 1 - \frac{b}{m}dt \end{pmatrix} \begin{pmatrix} y_t \\ \dot{y}_t \end{pmatrix} + w_t$$

Linear system

A linear system is any system where:

$$P(x_{t+1}|x_t) = N(x_{t+1}|Ax_t, Q)$$

(technically, this is a linear Gaussian system)

Also written: $x_{t+1} = Ax_t + w_t$ $\qquad w_t \sim N(w_t|0, Q)$

Also, assume that the observation function is linear Gaussian:

$$P(z_{t+1}|x_{t+1}) = N(z_{t+1}|Cx_{t+1}, R)$$

$$z_{t+1} = Cx_{t+1} + r_{t+1} \qquad r_t \sim N(r_t|0, R)$$

Kalman Idea

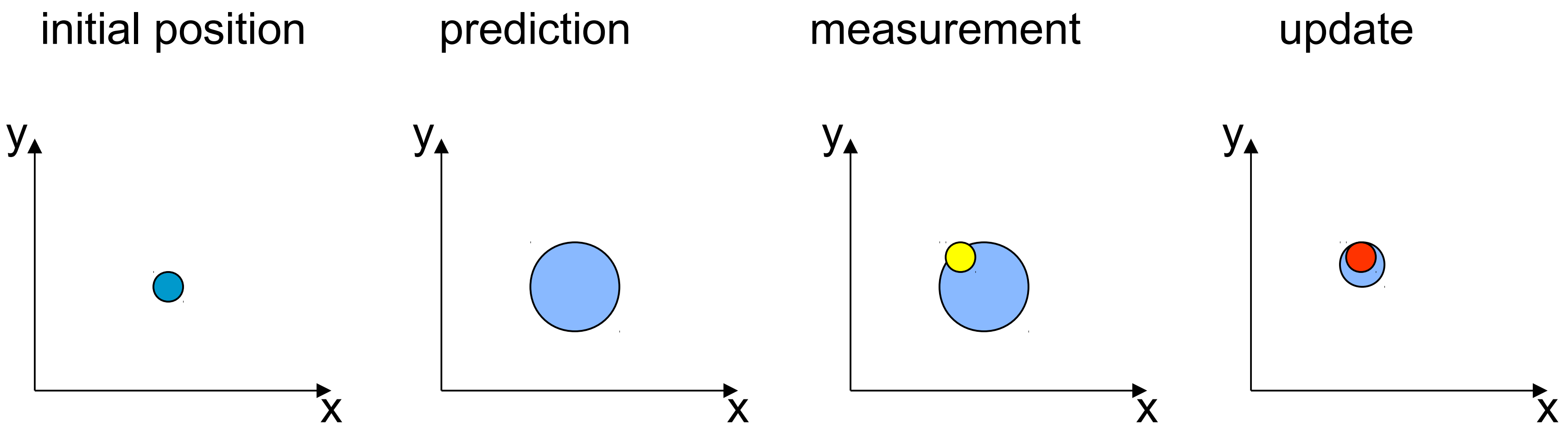


Image: Thrun *et al.*, *CS233B course notes*

Kalman Idea

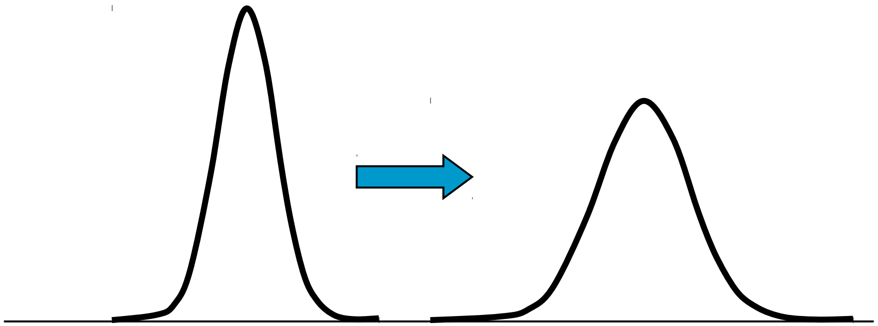


Image: Thrun *et al.*, *CS233B course notes*

$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

$$P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1})P(x_{t+1}|z_{0:t})$$

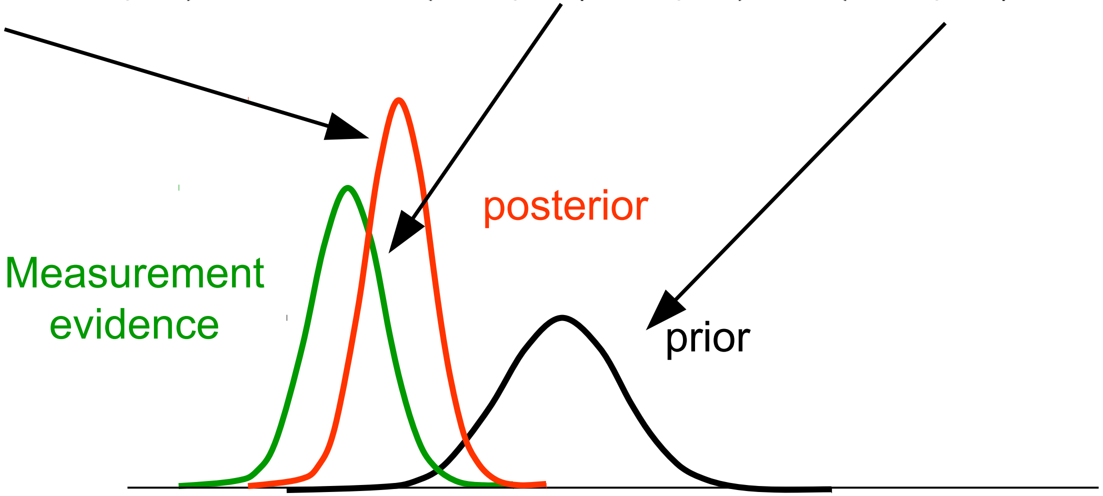


Image: Thrun *et al.*, *CS233B course notes*

Gaussians

Univariate Gaussian: $$P(x) = \eta e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

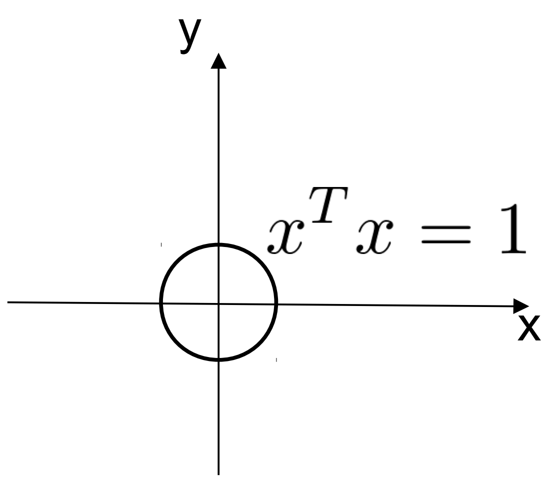
Multivariate Gaussian: $$P(x) = \eta e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

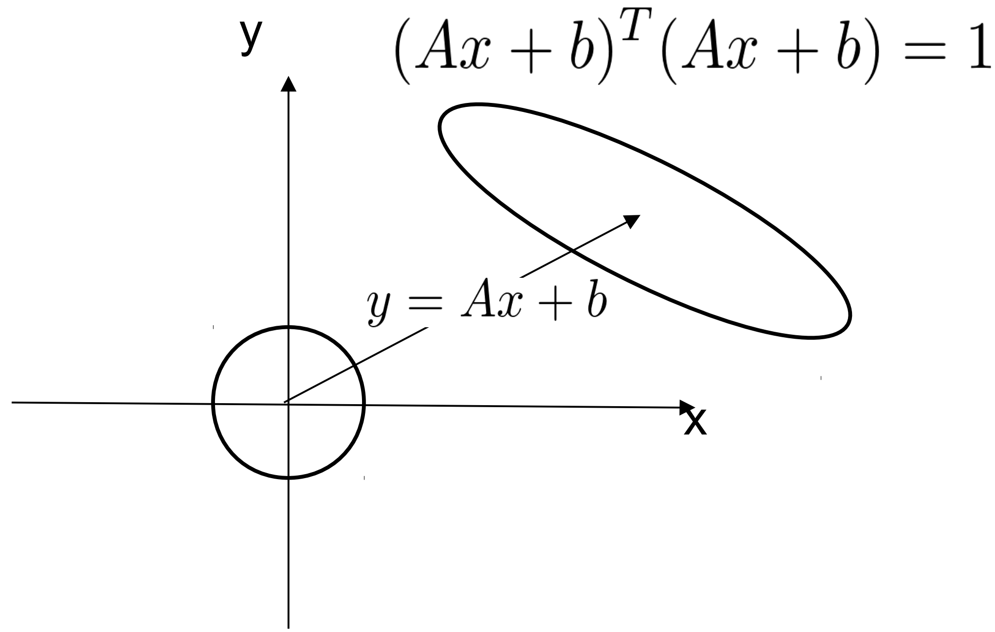
$$P(x) = N(x; \mu, \Sigma)$$

Playing w/ Gaussians

Suppose:
$$P(x) = N(x; \mu, \Sigma)$$
$$y = Ax + b$$

Calculate:
$$P(y) = ?$$
$$P(y) = N(y; Ax + b, A\Sigma A^T)$$





In fact

Suppose:
$$P(x) = N(x; \mu, \Sigma)$$
$$y = Ax + b$$

Then:

$$P\begin{pmatrix} x \\ y \end{pmatrix} = N\left[\begin{array}{c} x \\ y \end{array} : \begin{array}{c} \mu \\ A\mu + b \end{array}, \begin{pmatrix} \Sigma & \Sigma A^T \\ A\Sigma & A\Sigma A^T \end{pmatrix}\right]$$

Illustration

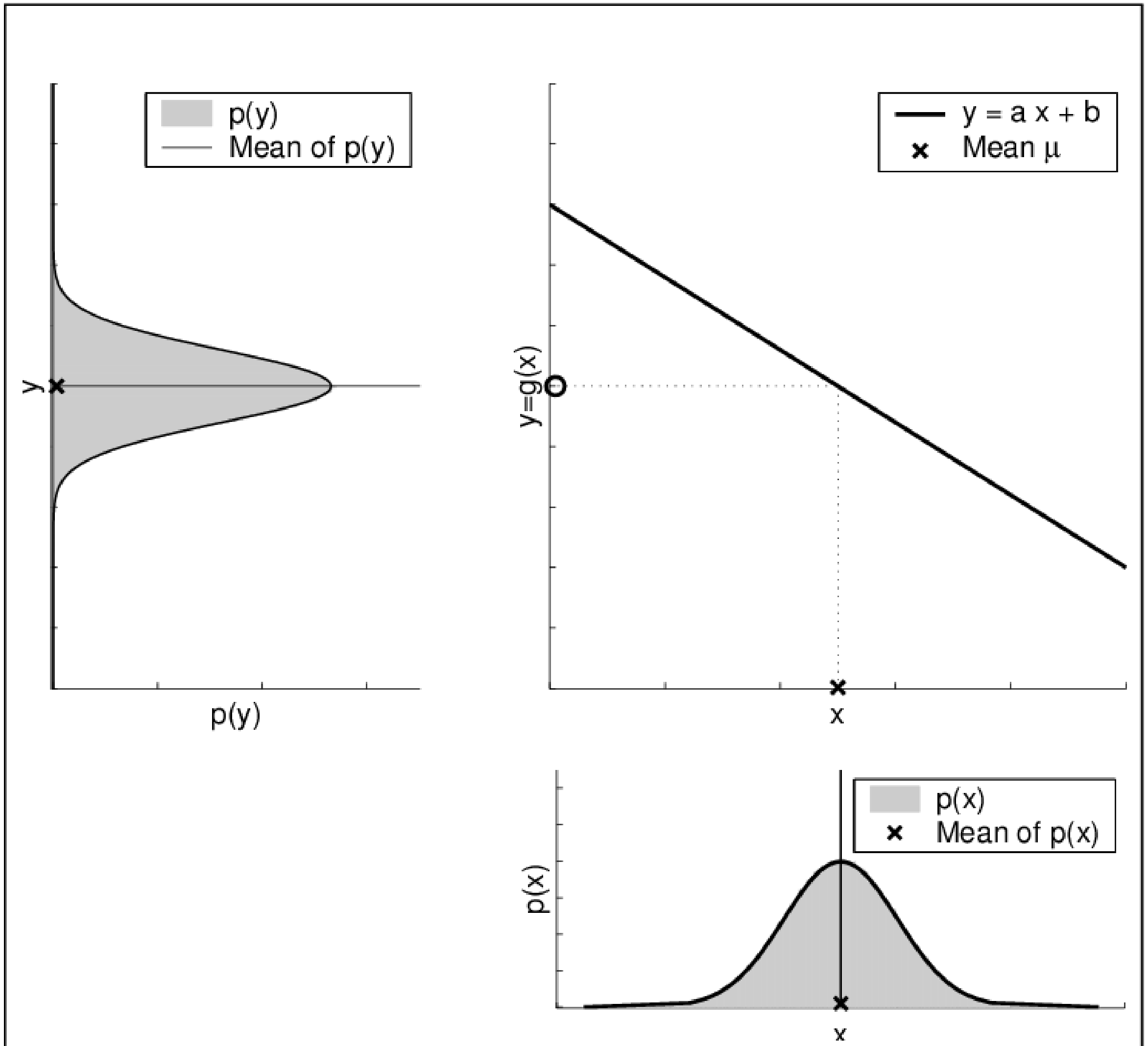


Image: Thrun *et al.*, *CS233B course notes*

And

Suppose: $$P(x) = N(x; \mu, \Sigma)$$

$$P(y|x) = N(y; Ax + b, R)$$

Then:

$$P\begin{pmatrix} x \\ y \end{pmatrix} = N\left[\begin{matrix} x \\ y \end{matrix} : \begin{matrix} \mu \\ A\mu + b \end{matrix}, \begin{pmatrix} \Sigma & \Sigma A^T \\ A\Sigma & A\Sigma A^T + R \end{pmatrix}\right]$$

$$P(y) = N(y; A\mu + b, A\Sigma A^T + R)$$

Marginal distribution

Does this remind us of anything?

Does this remind us of anything?

Process update (discrete):

$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

Process update (continuous):

$$P(x_{t+1}|z_{0:t}) = \int_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

Does this remind us of anything?

Process update (discrete):
$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

Process update (continuous):
$$P(x_{t+1}|z_{0:t}) = \int_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

$$N(x_{t+1}|Ax_t, Q)$$

Transition dynamics

$$N(x_t|\mu_t, \Sigma_t)$$

prior

Does this remind us of anything?

Process update (discrete):
$$P(x_{t+1}|z_{0:t}) = \sum_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

Process update (continuous):
$$P(x_{t+1}|z_{0:t}) = \int_{x_t} P(x_{t+1}|x_t)P(x_t|z_{0:t})$$

$$N(x_{t+1}|Ax_t, Q)$$

Transition dynamics

$$N(x_t|\mu_t, \Sigma_t)$$

prior

$$P(x_{t+1}|z_{0:t}) = \int_{x_t} N(x_{t+1}|Ax_t, Q)N(x_t; \mu_t, \Sigma_t)$$

$$P(x_{t+1}|z_{0:t}) = N(x_{t+1}|A\mu_t, A\Sigma_t A^T + Q)$$

Observation update

Observation update: $P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1}) P(x_{t+1}|z_{0:t})$

$$N(z_{t+1}|Cx_{t+1}, R) \qquad N(x_t|\mu'_t, \Sigma'_t)$$

Where: $\mu'_t = A\mu_t$

$$\Sigma'_t = A\Sigma_t A^T + Q$$

Observation update

Observation update: $P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1}) P(x_{t+1}|z_{0:t})$

$N(z_{t+1}|Cx_{t+1}, R)$ $\quad$ $N(x_t|\mu'_t, \Sigma'_t)$

Where: $\mu'_t = A\mu_t$

$\Sigma'_t = A\Sigma_t A^T + Q$

$$P(z_{t+1}, x_{t+1}|z_{0:t}) = \eta N(z_{t+1}|Cx_t, R) N(x_t; \mu'_t, \Sigma'_t)$$

Observation update

Observation update: $P(x_{t+1}|z_{0:t+1}) = \eta P(z_{t+1}|x_{t+1}) P(x_{t+1}|z_{0:t})$

$$N(z_{t+1}|Cx_{t+1}, R) \qquad N(x_t|\mu'_t, \Sigma'_t)$$

Where: $\mu'_t = A\mu_t$

$$\Sigma'_t = A\Sigma_t A^T + Q$$

$$P(z_{t+1}, x_{t+1}|z_{0:t}) = \eta N(z_{t+1}|Cx_t, R) N(x_t; \mu'_t, \Sigma'_t)$$

$$P(z_{t+1}, x_{t+1}|z_{0:t}) = N\left[\begin{array}{ccc} x_{t+1} & : & \mu'_t \\ z_{t+1} & & C\mu'_t \end{array}, \begin{pmatrix} \Sigma'_t & \Sigma'_t C^T \\ C\Sigma'_t & C\Sigma'_t A^T + R \end{pmatrix}\right]$$

Observation update

$$P(z_{t+1}, x_{t+1} | z_{0:t}) = N \left[\begin{array}{ccc} x_{t+1} & : & \mu'_t \\ z_{t+1} & & C\mu'_t \end{array}, \left(\begin{array}{cc} \Sigma'_t & \Sigma'_t C^T \\ C\Sigma'_t & C\Sigma'_t A^T + R \end{array} \right) \right]$$

But we need: $P(x_{t+1} | z_{0:t+t}) = ?$

Another Gaussian identity...

Suppose: $$N\left[\begin{array}{ccc} x & : & a \\ y & & b \end{array}, \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}\right]$$

Calculate: $$P(y|x) = ?$$

$$P(y|x) = N(y|b + C^T A^{-1}(x-a), B - C^T A^{-1} C)$$

Observation update

$$P(z_{t+1}, x_{t+1} | z_{0:t}) = N\left[\begin{array}{ccc} x_{t+1} & : & \mu'_t \\ z_{t+1} & & C\mu'_t \end{array}, \begin{pmatrix} \Sigma & \Sigma C^T \\ C\Sigma & C\Sigma A^T + R \end{pmatrix}\right]$$

But we need: $P(x_{t+1} | z_{0:t+1}) = ?$

$$P(x_{t+1} | z_{0:t+1}) = N(x_{t+1}; \mu_{t+1}, \Sigma_{t+1})$$

$$\mu_{t+1} = \mu'_t + \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1} (z_{t+1} - C\mu'_t)$$
$$\Sigma_{t+1} = \Sigma'_t - \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1} C\Sigma'_t$$

To summarize the Kalman filter

System:

$$P(x_{t+1}|x_t) = N(x_{t+1}|Ax_t, Q)$$
$$P(z_{t+1}|x_{t+1}) = N(z_{t+1}|Cx_{t+1}, R)$$

Prior: μ_t

Σ_t

Process update:

$$\mu'_t = A\mu_t$$
$$\Sigma'_t = A\Sigma_t A^T + Q$$

Measurement update:

$$\mu_{t+1} = \mu'_t + \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1}(z_{t+1} - C\mu'_t)$$
$$\Sigma_{t+1} = \Sigma'_t - \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1} C\Sigma'_t$$

Suppose there is an action term...

System:
$$P(x_{t+1}|x_t) = N(x_{t+1}|Ax_t + u_t, Q)$$
$$P(z_{t+1}|x_{t+1}) = N(z_{t+1}|Cx_{t+1}, R)$$

Prior: μ_t

Σ_t

Process update:
$$\mu_t' = A\mu_t + u_t$$
$$\Sigma_t' = A\Sigma_t A^T + Q$$

Measurement update:
$$\mu_{t+1} = \mu_t' + \Sigma_t' C^T (R + C\Sigma_t' C^T)^{-1}(z_{t+1} - C\mu_t')$$
$$\Sigma_{t+1} = \Sigma_t' - \Sigma_t' C^T (R + C\Sigma_t' C^T)^{-1} C\Sigma_t'$$

To summarize the Kalman filter

Prior: μ_t

Σ_t

Process update: $\mu'_t = A\mu_t$

$\Sigma'_t = A\Sigma_t A^T + Q$

Measurement update: $\mu_{t+1} = \mu'_t + \boxed{\Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1}} (z_{t+1} - C\mu'_t)$

This factor is often called the "Kalman gain"

$$\Sigma_{t+1} = \Sigma'_t - \boxed{\Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1}} C\Sigma'_t$$

Things to note about the Kalman filter

Process update: $$\mu'_t = A\mu_t$$

$$\Sigma'_t = A\Sigma_t A^T + Q$$

Measurement update: $$\mu_{t+1} = \mu'_t + \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1}(z_{t+1} - C\mu'_t)$$
$$\Sigma_{t+1} = \Sigma'_t - \Sigma'_t C^T (R + C\Sigma'_t C^T)^{-1} C\Sigma'_t$$

– covariance update is independent of observation

– Kalman is only optimal for linear-Gaussian systems

– the distribution “stays” Gaussian through this update

– the error term can be thought of as the different between the observation and the prediction

Kalman in 1D

System:

$$P(x_{t+1}|x_t) = N(x_{t+1} : x_t + u_t, q)$$
$$P(z_{t+1}|x_{t+1}) = N(z_{t+1}|2x_{t+1}, r)$$

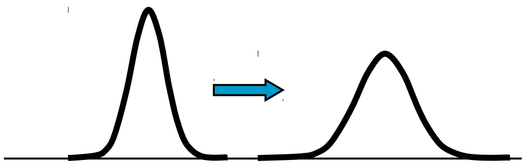


Image: Thrun *et al.*, *CS233B course notes*

Process update:

$$\bar{\mu}_t = \mu_t + u_t$$
$$\bar{\sigma}_t^2 = \sigma_t^2 + q$$

Measurement update:

$$\mu_{t+1} = \bar{\mu}_t + \frac{2\bar{\sigma}_t^2}{r + 4\bar{\sigma}_t^2}(z_{t+1} - \bar{\mu}_t)$$
$$\sigma_{t+1} = \bar{\sigma}_t^2 - \frac{4(\bar{\sigma}_t^2)^2}{r + 4\bar{\sigma}_t^2}$$

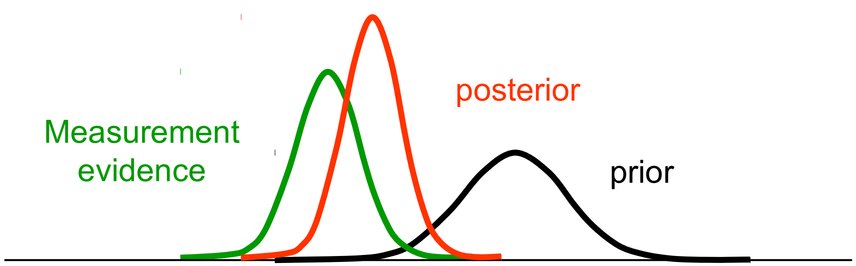


Image: Thrun *et al.*, *CS233B course notes*

Kalman Idea

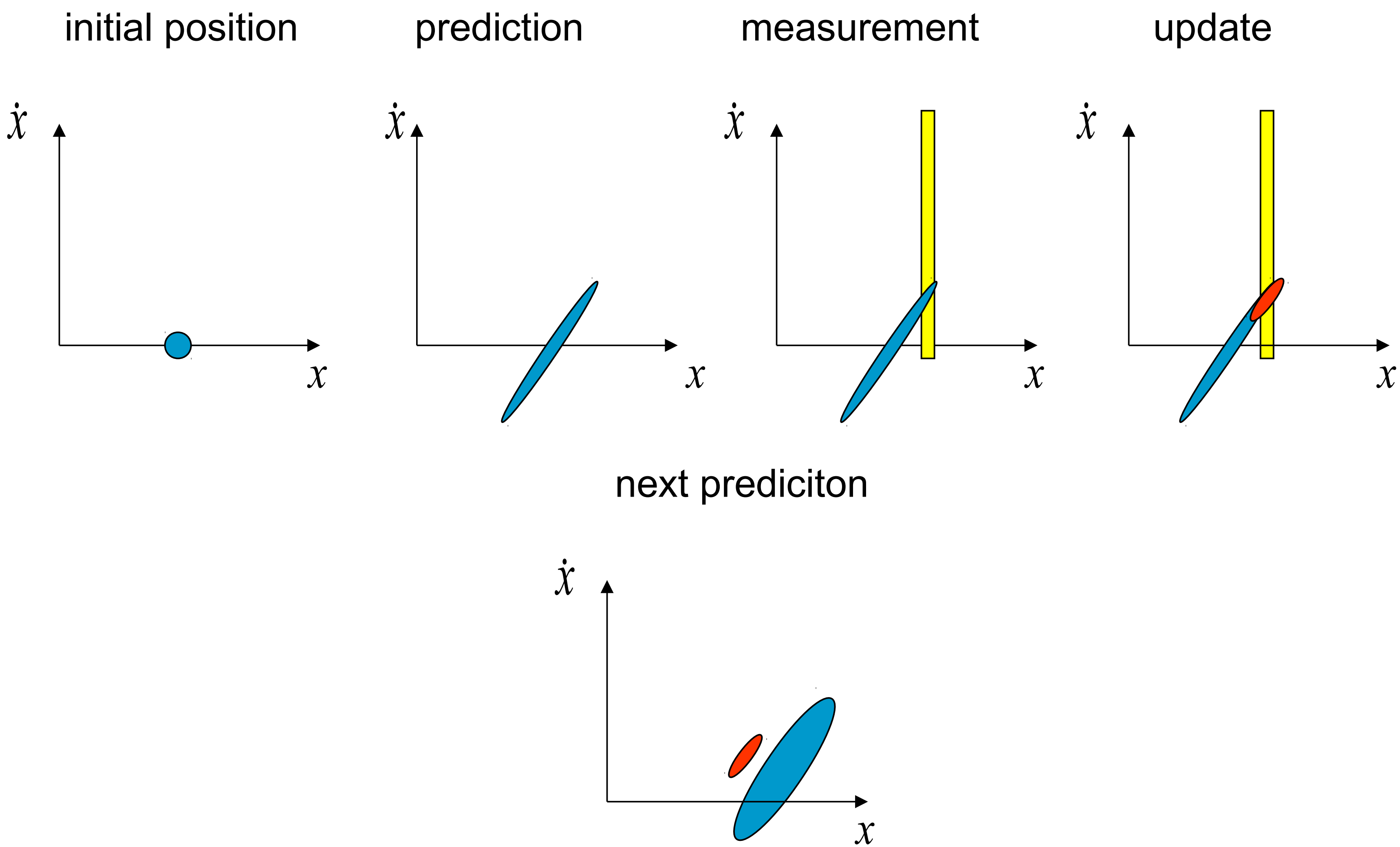


Image: Thrun *et al.*, *CS233B course notes*

Example: estimate velocity

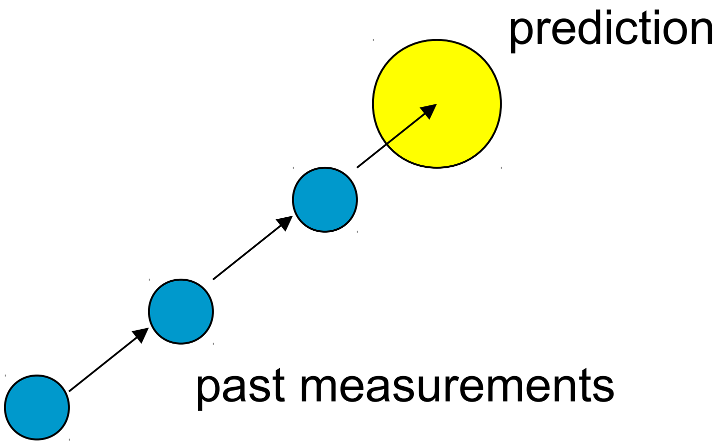


Image: Thrun *et al.*, *CS233B course notes*

Example: filling a tank

$$x = \begin{pmatrix} l \\ f \end{pmatrix} \begin{matrix} \longleftarrow \text{Level of tank} \\ \longleftarrow \text{Fill rate} \end{matrix}$$

$$l_{t+1} = l_t + f\,dt$$

Process: $$x_{t+1} = \begin{pmatrix} 1 & dt \\ 0 & 1 \end{pmatrix} x_t + q$$

Observation: $$z_{t+1} = \begin{pmatrix} 1 & 0 \end{pmatrix} x_{t+1} + r$$

Example: estimate velocity

$$x_{t+1} = Ax_t + w_t$$

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \\ \dot{x}_{t+1} \\ \dot{y}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \end{pmatrix} + \mathrm{w}_t$$

$$z_{t+1} = Cx_{t+1} + r_{t+1}$$

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \\ \dot{x}_{t+1} \\ \dot{y}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{t+1} \\ y_{t+1} \\ \dot{x}_{t+1} \\ \dot{y}_{t+1} \end{pmatrix} + \mathrm{r}_{t+1}$$

But, my system is NON-LINEAR!

$$
\begin{aligned}
x_{t+1} &= f(x_t, u_t) \\
&\neq Ax_t + Bu_t
\end{aligned}
$$

What should I do?

But, my system is NON-LINEAR!

$$
\begin{aligned}
x_{t+1} &= f(x_t, u_t) \\
&\neq Ax_t + Bu_t
\end{aligned}
$$

What should I do?

Well, there are some options...

But, my system is NON-LINEAR!

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ &\neq Ax_t + Bu_t \end{aligned}$$

What should I do?

Well, there are some options...

But none of them are great.

But, my system is NON-LINEAR!

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ &\neq Ax_t + Bu_t \end{aligned}$$

What should I do?

Well, there are some options...

But none of them are great.

Here's one: the Extended Kalman Filter

Extended Kalman filter

Take a Taylor expansion:

$$x_{t+1} = f(x_t, u_t)$$

$$\approx f(\mu_t, u_t) + A_t(x_t - \mu_t)$$

Where: $A_t = \frac{\partial f}{\partial x}(\mu_t, u_t)$

$$z_{t+1} = h(x_t)$$

$$\approx h(\mu_t) + C_t(x_t - \mu_t)$$

Where: $C_t = \frac{\partial h}{\partial x}(\mu_t)$

Extended Kalman filter

Take a Taylor expansion:

$$x_{t+1} = f(x_t, u_t)$$
$$\approx f(\mu_t, u_t) + A_t(x_t - \mu_t)$$

Where: $$A_t = \frac{\partial f}{\partial x}(\mu_t, u_t)$$

$$z_{t+1} = h(x_t)$$
$$\approx h(\mu_t) + C_t(x_t - \mu_t)$$

Where: $$C_t = \frac{\partial h}{\partial x}(\mu_t)$$

Then use the same equations...

To summarize the EKF

Prior: μ_t

Σ_t

Process update: $\mu_t' = f(\mu_t, u_t)$

$\Sigma_t' = A_t \Sigma_t A_t^T + Q$

Measurement update: $\mu_{t+1} = \mu_t' + \Sigma_t' C^T (R + C\Sigma_t' C^T)^{-1}(z_{t+1} - h(\mu_t'))$

$\Sigma_{t+1} = \Sigma_t' - \Sigma_t' C^T (R + C\Sigma_t' C^T)^{-1} C\Sigma_t'$

Extended Kalman filter

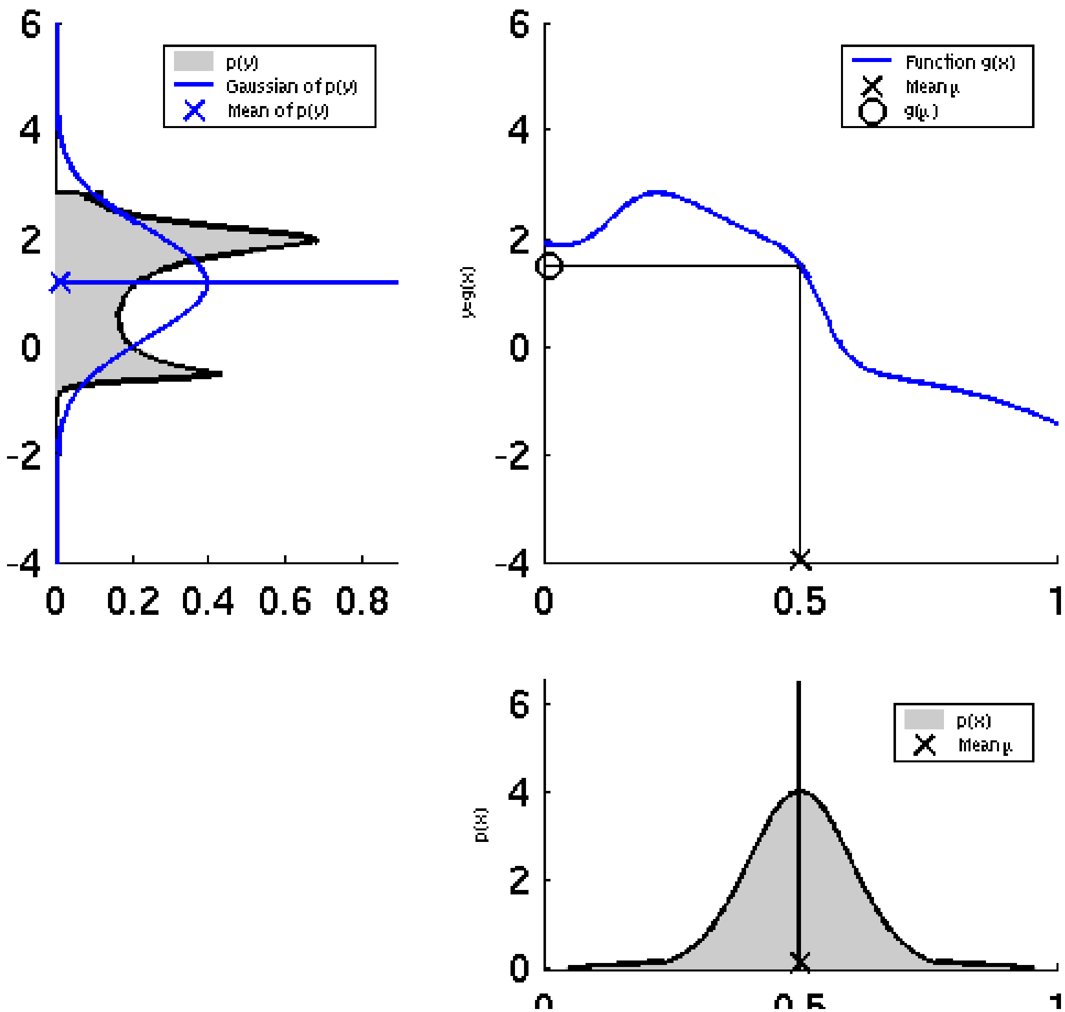


Image: Thrun *et al.*, *CS233B course notes*

Extended Kalman filter

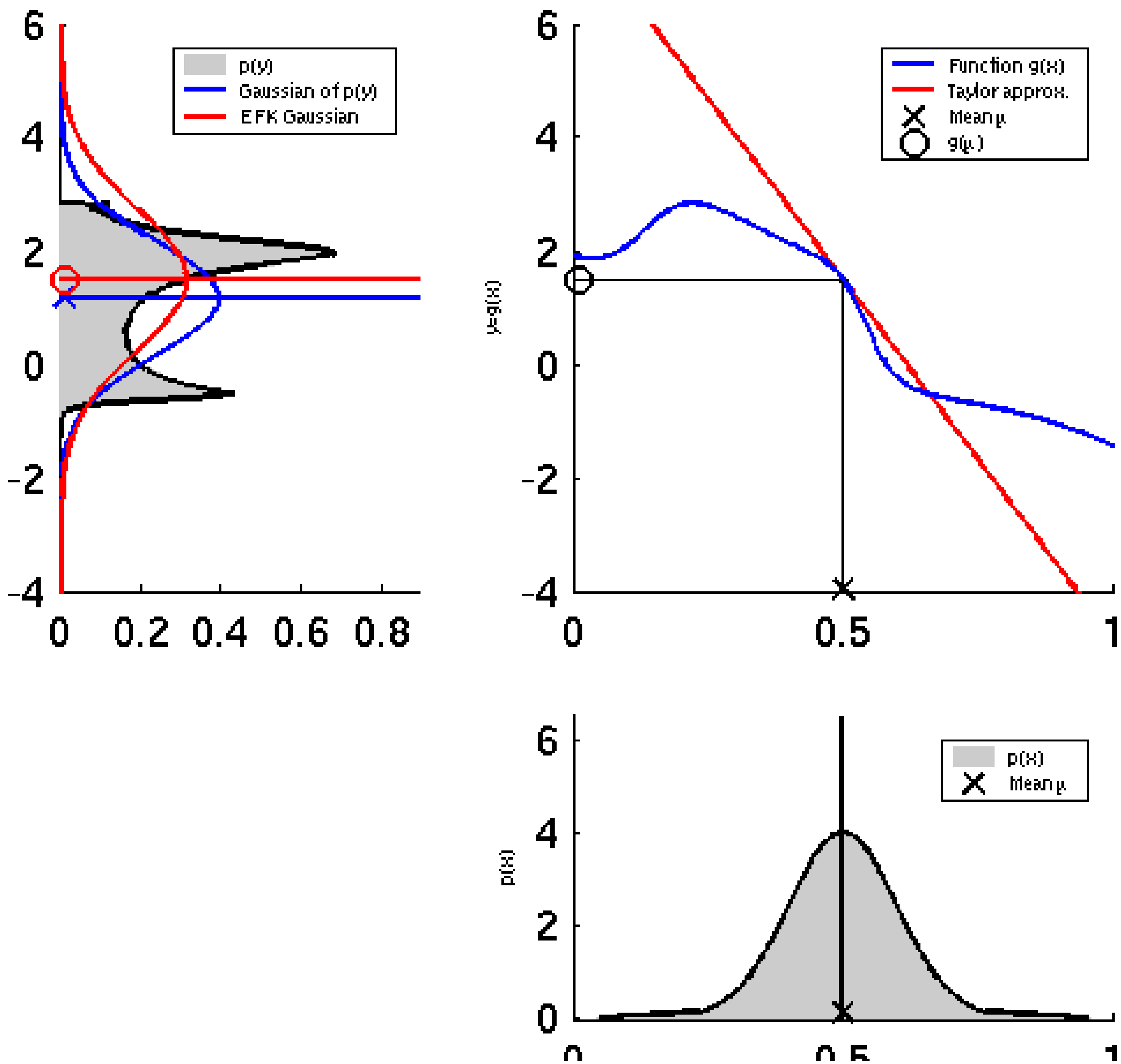


Image: Thrun *et al.*, *CS233B course notes*