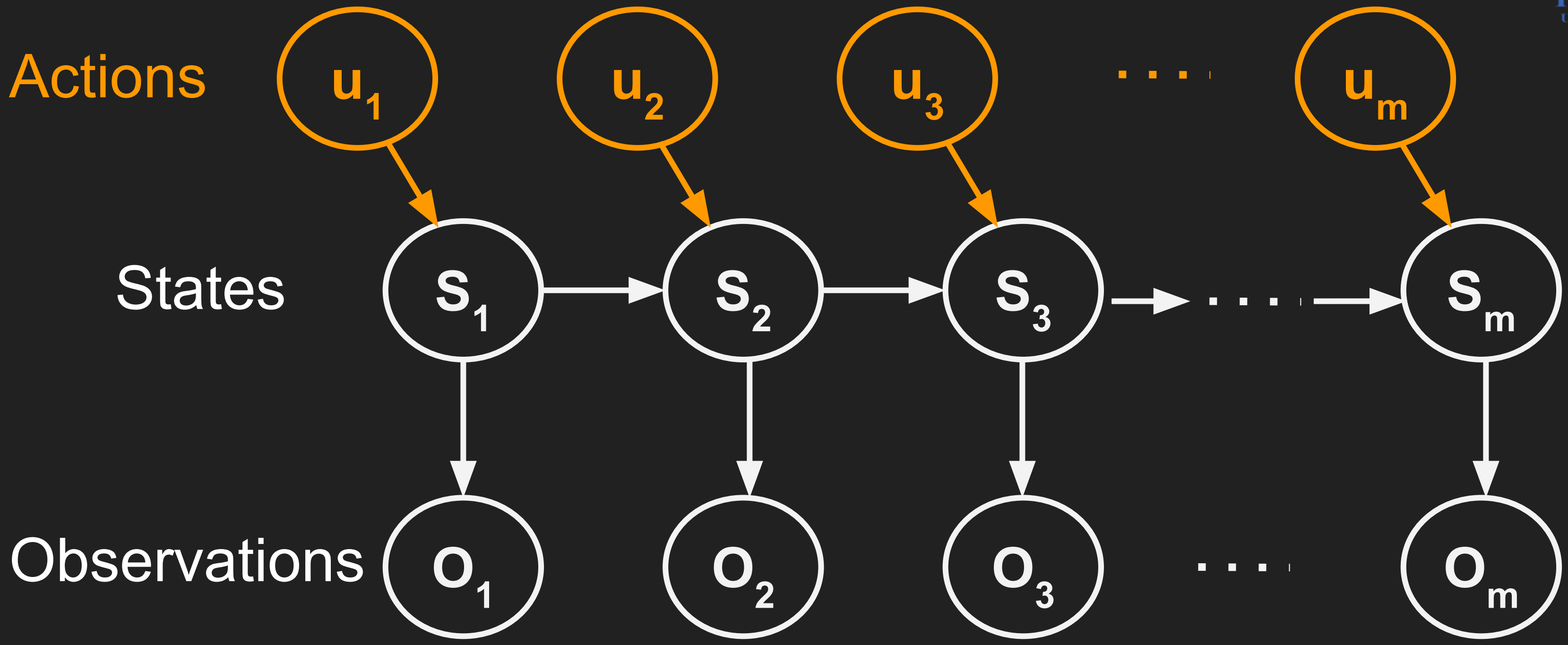
Probabilistic Graphical Models

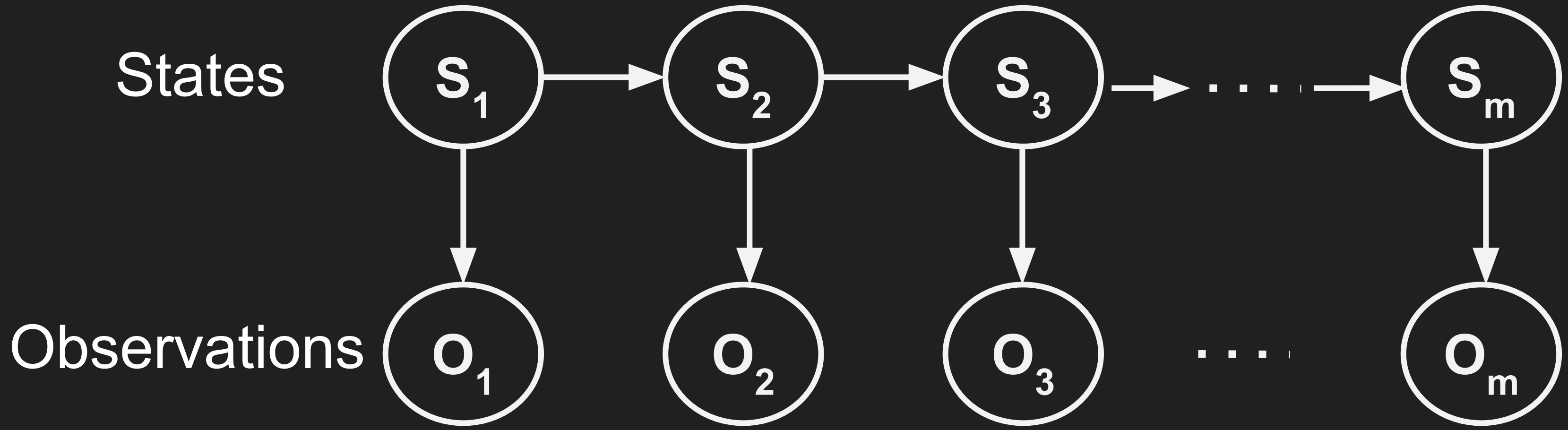
Lectures 18

Kalman Filters
Extended Kalman Filter
UKF, Particle Filter

Remember: Temporal Models



Remember: Temporal Models



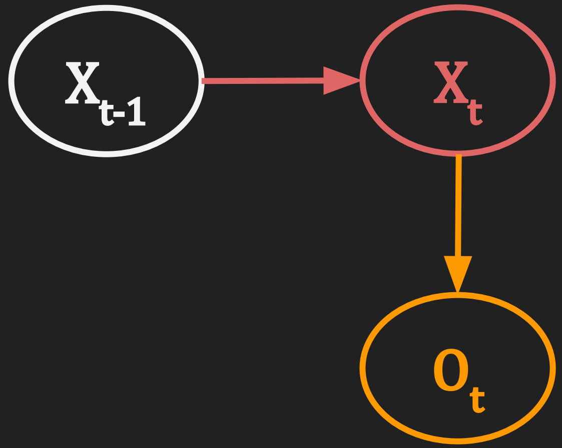
Remember: Transition and Observation Models

K. N. Toosi
University of Technology

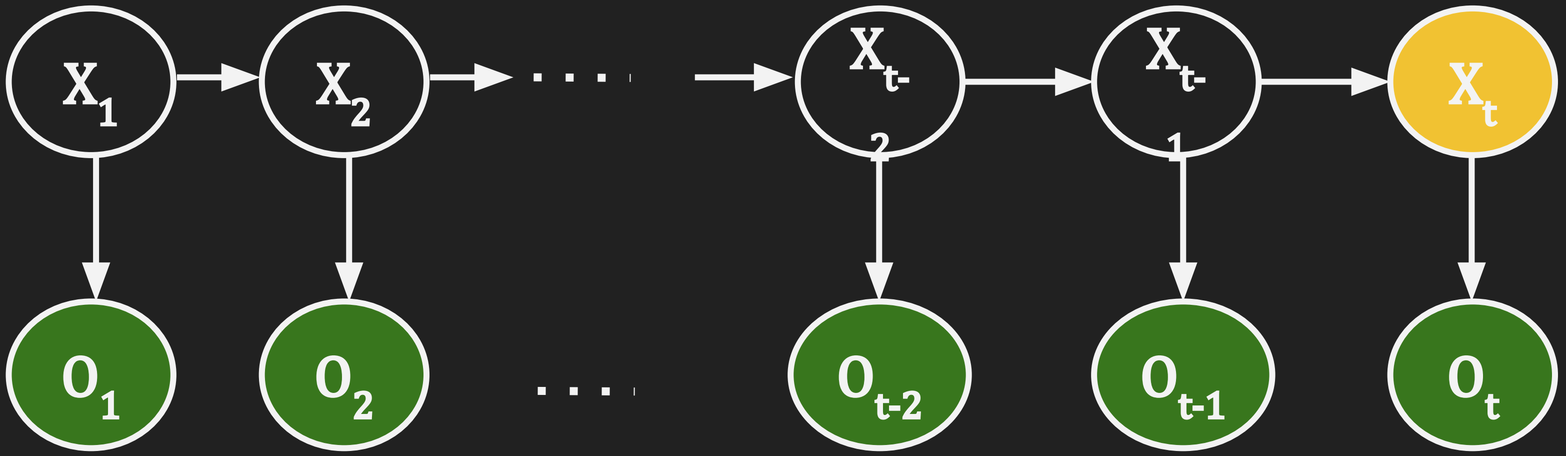
(No input)

$P(X_t \mid X_{t-1})$ Transition Model

$P(O_t \mid X_t)$ Observation Model



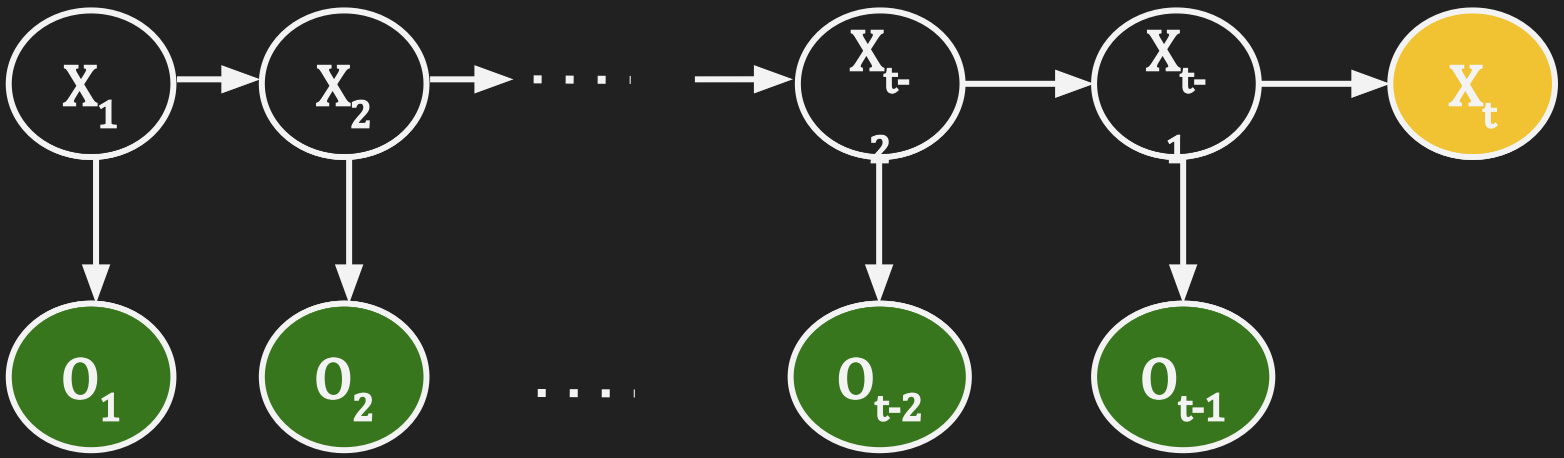
Inference: State Estimation



Given $O_{1,}$ $O_{2},$ … , O_t what is X_t?

$P(X_t \mid O_{1,} O_{2}, \ldots, O_t) = ?$

Inference: Prediction



Given $O_1, O_2, \ldots, O_{t-1}$ what is X_t?

$P(X_t \mid O_1, O_2, \ldots, O_{t-1}) = ?$

Inference Problems

K. N. Toosi University of Technology

$\text{pred}(X_t) = P(X_t \mid O_{1,} O_{2}, \dots, O_{t-1})$ Prediction

$\text{corr}(X_t) = P(X_t \mid O_{1,} O_{2}, \dots, O_{t})$ Correction (update)

1. Variable Elimination
2. Recursive Bayesian Filtering

$$\cdots \rightarrow \text{pred}(X_t) \rightarrow \text{corr}(X_t) \rightarrow \text{pred}(X_{t+1}) \rightarrow \text{corr}(X_{t+1}) \rightarrow \cdots$$

Recursive Bayesian Filtering

K. N. Toosi
University of Technology

$$\mathrm{pred}(\mathsf{X}_t) = \sum_{\mathsf{X}_{t-1}} P(\mathsf{X}_t \mid \mathsf{X}_{t-1})\ \mathrm{corr}(\mathsf{X}_{t-1})$$

$$\mathrm{corr}(\mathsf{X}_t) = \frac{P(\mathsf{O}_t \mid \mathsf{X}_t)\ \mathrm{pred}(\mathsf{X}_t)}{\sum_{\mathsf{X}_t} P(\mathsf{O}_t \mid \mathsf{X}_t)\ \mathrm{pred}(\mathsf{X}_t)}$$

Recursive Bayesian Filtering

K. N. Toosi
University of Technology

$$\mathrm{pred}(\mathsf{X}_t) = \int P(\mathsf{X}_t \mid \mathsf{X}_{t-1})\ \mathrm{corr}(\mathsf{X}_{t-1})\ d\mathsf{X}_{t-1}$$

$$\mathrm{corr}(\mathsf{X}_t) = \frac{P(\mathsf{O}_t \mid \mathsf{X}_t)\ \mathrm{pred}(\mathsf{X}_t)}{\int P(\mathsf{O}_t \mid \mathsf{X}_t)\ \mathrm{pred}(\mathsf{X}_t)\ d\mathsf{X}_t}$$

Kalman Filter

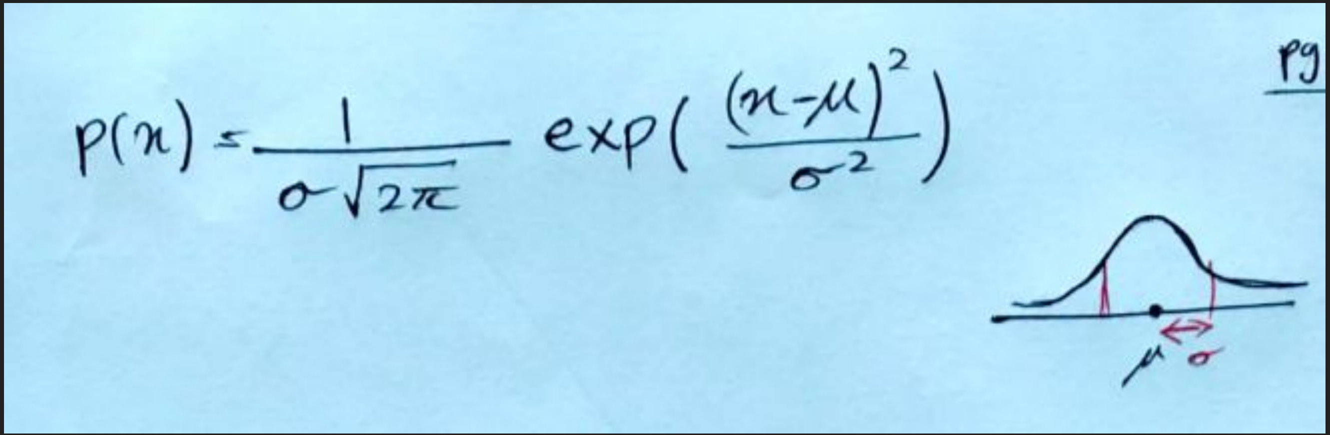
K. N. Toosi University of Technology

- Linear Transition and Observation Models
  - $X_t = \mathbf{A} X_{t-1} + \varepsilon_x$, $O_t = \mathbf{B} X_t + \varepsilon_o$
- Additive Gaussian Noise
  - $\varepsilon_x \sim N(0, \Sigma_x)$, $\varepsilon_o \sim N(0, \Sigma_o)$

$$P(X_t \mid X_{t-1}) = P_{\varepsilon x}(X_t - \mathbf{A} X_{t-1}) = \mathrm{Normal}(X_t - \mathbf{A} X_{t-1}; 0, \Sigma_x)$$

$$P(O_t \mid X_t) = P_{\varepsilon o}(O_t - \mathbf{B} X_t) = \mathrm{Normal}(O_t - \mathbf{B} X_t; 0, \Sigma_o)$$

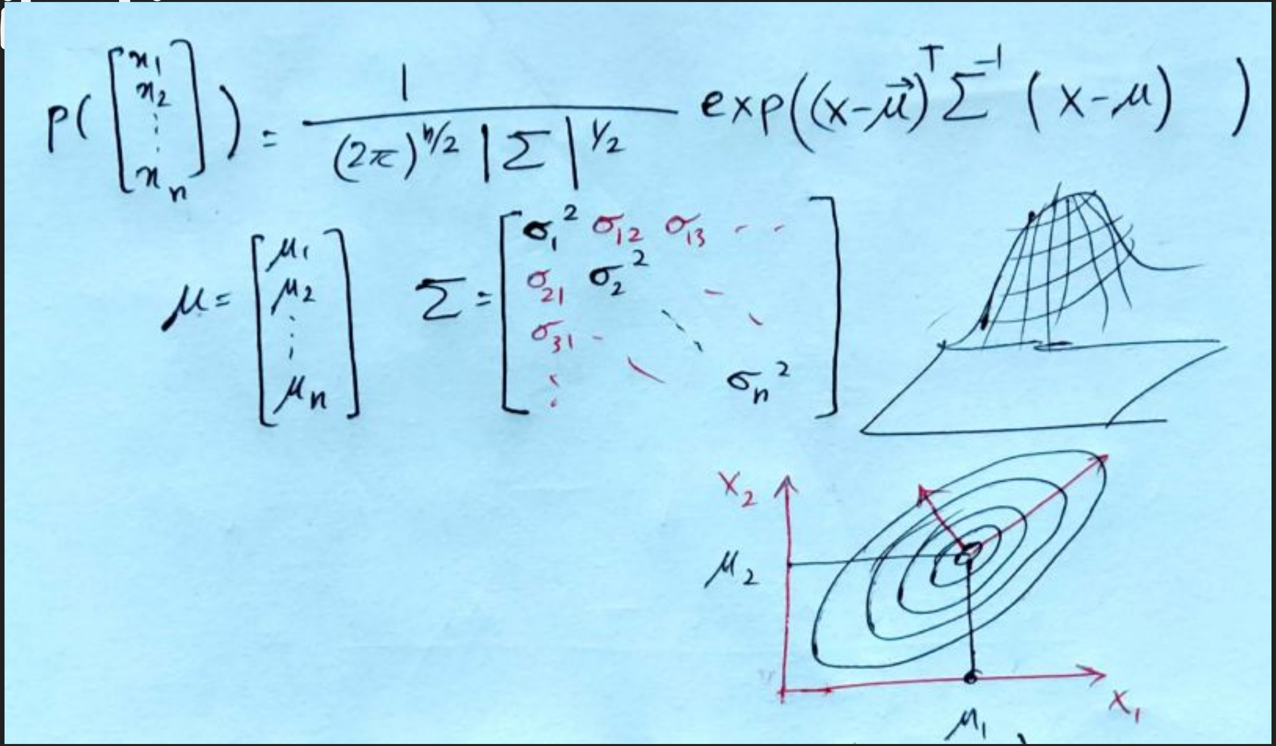
Remember: Gaussian Distribution



$p(x)$ can be represented with μ, σ

Remember: Multivariate Gaussian Distri

K. N. Toosi
University of Technology



$p(\boldsymbol{x})$ can be represented with $\boldsymbol{\mu}, \boldsymbol{\Sigma}$

Kalman Filters

Gaussian   Gaussian

$$\text{pred}(X_t) = \int P(X_t \mid X_{t-1})\, \text{corr}(X_{t-1})\, dX_{t-1}$$

Gaussian

Gaussian   Gaussian

$$\text{corr}(X_t) = \frac{P(O_t \mid X_t)\, \text{pred}(X_t)}{\int P(O_t \mid X_t)\, \text{pred}(X_t)\, dX_t}$$

Gaussian

K. N. Toosi University of Technology

Kalman Filters

K. N. Toosi
University of Technology

$$\mathrm{pred}(X_t) = \int P(X_t \mid X_{t-1})\ \mathrm{corr}(X_{t-1})\ dX_{t-1}$$

$\mathcal{N}(\mu'_t, \Sigma'_t)$

$$\mathrm{corr}(X_t) = \frac{P(O_t \mid X_t)\ \mathrm{pred}(X_t)}{\int P(O_t \mid X_t)\ \mathrm{pred}(X_t)\ dX_t}$$

$\mathcal{N}(\mu_t, \Sigma_t)$

Kalman Filters

$\mathcal{N}(\mu_{t-1}, \Sigma_{t-1})$

$$\text{pred}(X_t) = \int P(X_t \mid X_{t-1})\, \text{corr}(X_{t-1})\, dX_{t-1}$$

$\mathcal{N}(\mu'_t, \Sigma'_t)$

$\mathcal{N}(\mu'_t, \Sigma'_t)$

$$\text{corr}(X_t) = \frac{P(O_t \mid X_t)\, \text{pred}(X_t)}{\int P(O_t \mid X_t)\, \text{pred}(X_t)\, dX_t}$$

$\mathcal{N}(\mu_t, \Sigma_t)$

Kalman Filters

$$\underbrace{\text{pred}(X_t)}_{\mathcal{N}(\mu'_t\,,\,\Sigma'_t)} = \int \overbrace{P(X_t \mid X_{t-1})}^{\mathcal{N}(A\,X_{t-1}\,,\,\Sigma_x)}\ \overbrace{\text{corr}(X_{t-1})}^{\mathcal{N}(\mu_{t-1}\,,\,\Sigma_{t-1})}\ dX_{t-1}$$

$$\underbrace{\text{corr}(X_t)}_{\mathcal{N}(\mu_t\,,\,\Sigma_t)} = \frac{\overbrace{P(O_t \mid X_t)}^{\mathcal{N}(B\,X_t\,,\,\Sigma_o)}\ \overbrace{\text{pred}(X_t)}^{\mathcal{N}(\mu'_t\,,\,\Sigma'_t)}}{\int P(O_t \mid X_t)\ \text{pred}(X_t)\ dX_t}$$

Kalman Filter

$$pred(X_t) = P(X_t \mid O_1, O_2, \ldots, O_{t-1}) = \mathcal{N}(X_t;\ \mu'_t, \Sigma'_t)$$

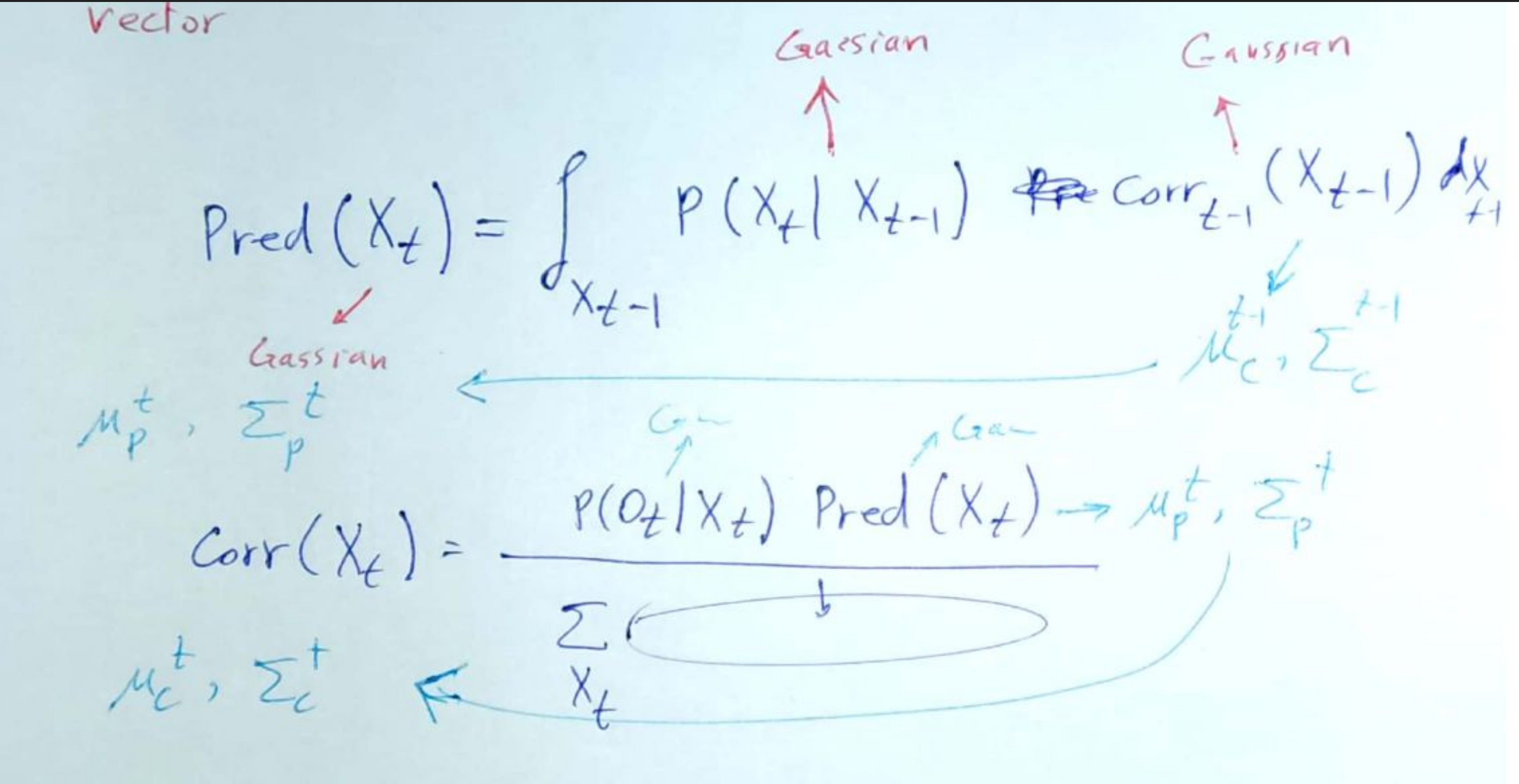
$$corr(X_t) = P(X_t \mid O_1, O_2, \ldots, O_t) = \mathcal{N}(X_t;\ \mu_t, \Sigma_t)$$

$$P(X_t \mid X_{t-1}) = Normal(X_t - \mathbf{A} X_{t-1}; 0, \Sigma_x) = \mathcal{N}(X_t; \mathbf{A} X_{t-1}, \Sigma_x)$$

$$P(O_t \mid X_t) = Normal(O_t - \mathbf{B} X_t; 0, \Sigma_o) = \mathcal{N}(O_t; \mathbf{B} X_t, \Sigma_o)$$

Kalman Filters

K. N. Toosi
University of Technology



Kalman Filter

$$p(X_t \mid X_{t-1}) = \text{Normal}\left(X_t - AX_{t-1} ; \emptyset , \Sigma_x\right)$$

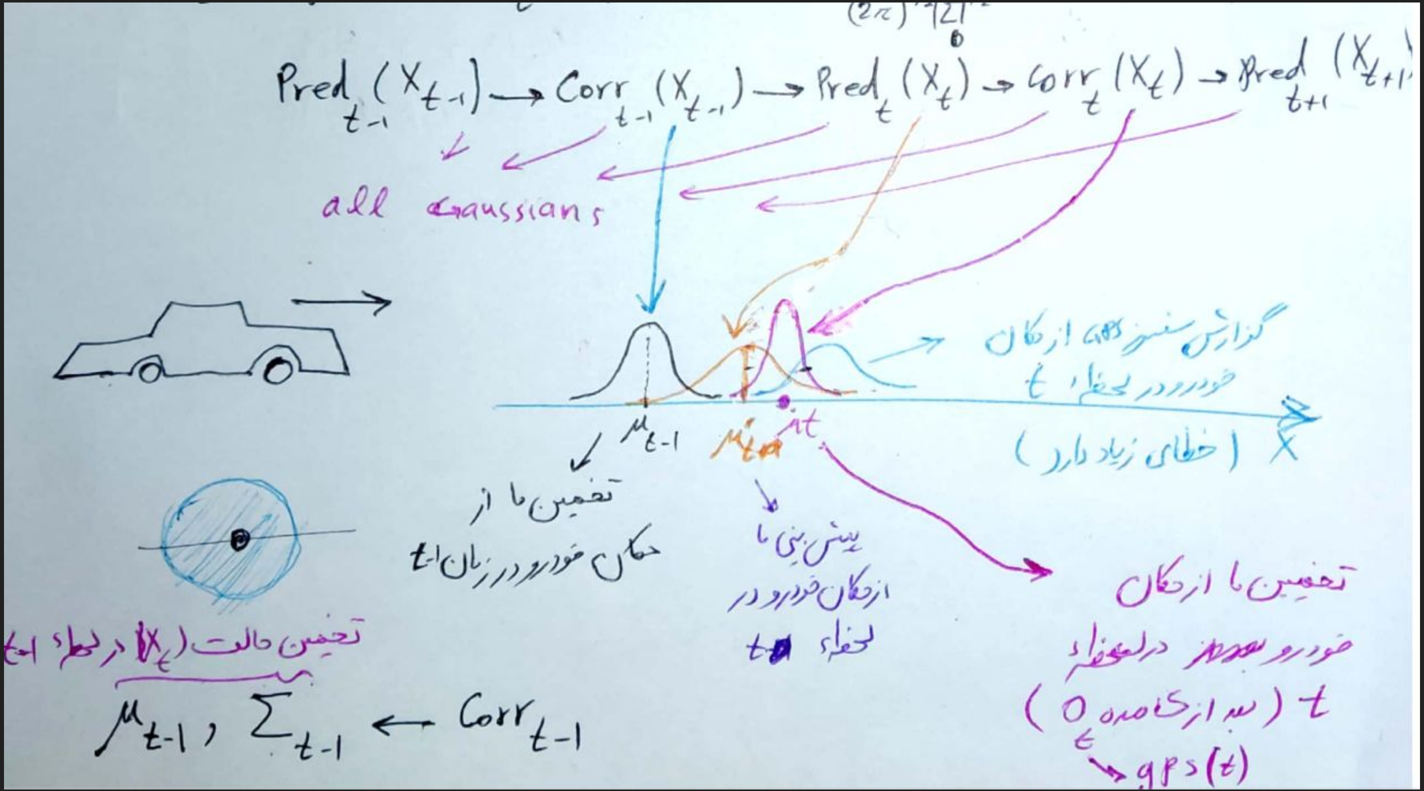
$$= \frac{1}{(2\pi)^{n/2} \, |\Sigma_x|^{1/2}} \exp\left((X_t - AX_{t-1})^T \Sigma_x^{-1} (X_t - AX_{t-1}) \right)$$

observation model

$$p(O_t \mid X_t) = \text{Normal}\left(O_t - BX_t , \emptyset , \Sigma_o\right)$$

$$= \frac{1}{(2\pi)^{n/2} \, |\Sigma_o|^{1/2}} \exp\left((O_t - BX_t)^T \Sigma_o^{-1} (O_t - BX_t) \right)$$

Kalman Filters



Kalman Filters - Prediction Phase

$\mathcal{N}(X_t ; \mu'_t , \Sigma'_t)$ $\quad$ $\mathcal{N}(X_t ; A X_{t-1} , \Sigma_x)$ $\quad$ $\mathcal{N}(X_{t-1} ; \mu_{t-1} , \Sigma_{t-1})$

$$\text{pred}(\text{X}_t) = \int P(\text{X}_t \mid \text{X}_{t-1}) \, \text{corr}(\text{X}_{t-1}) \, d\text{X}_{t-1}$$

$$\mu'_t = A \ \mu_{t-1}$$

$$\Sigma'_t = \Sigma_x + A \ \Sigma_{t-1} \ A^T$$

Kalman Filters - Prediction Phase

K. N. Toosi University of Technology

$$X_t = \mathbf{A}\, X_{t-1} + \varepsilon_x$$

$$\mu'_t = A\ \mu_{t-1}$$

$$\Sigma'_t = \Sigma_x + A\ \Sigma_{t-1}\ A^T$$

Kalman Filters - Correction Phase

$\mathcal{N}(X_t ; \mu_t , \Sigma_t)$ $\quad$ $\mathcal{N}(O_t ; B X_t , \Sigma_o)$ $\quad$ $\mathcal{N}(X_t ; \mu'_t , \Sigma'_t)$

$$\mathrm{corr}(X_t) = \frac{P(O_t \mid X_t)\ \mathrm{pred}(X_t)}{\int P(O_t \mid X_t)\ \mathrm{pred}(X_t)\ dX_t}$$

$$K = \Sigma'_t B^T (\Sigma_o + B \Sigma'_t B^T)^{-1}$$

$$\mu_t = \mu'_t + K (O_t - B \mu'_t)$$

$$\Sigma_t = \Sigma'_t - K B \Sigma'_t$$

Kalman Filters - Correction Phase

K. N. Toosi
University of Technology

$$O_t = B X_t + \varepsilon_o$$

$$K = \Sigma'_t B^T (\Sigma_o + B \Sigma'_t B^T)^{-1}$$

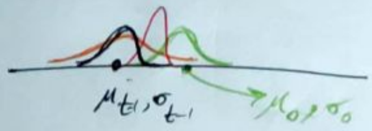
$$\mu_t = \mu'_t + K (O_t - B \mu'_t)$$

$$\Sigma_t = \Sigma'_t - K B \Sigma'_t = (I - K B) \Sigma'_t$$

Example - 1D case, no velocity

$$x_{t+1} = x_t + \varepsilon_x \rightarrow \mu_x, \sigma_x$$

$$o_t = x_t + \varepsilon_o \rightarrow \mu_o, \sigma_o$$



$$\mu'_t = \mu_{t-1}$$

$$\sigma'^2_t = \sigma_t^2 + \sigma_x^2$$

$$K = \frac{\sigma_t^2 + \sigma_x^2}{\sigma_o^2 + \sigma_t^2 + \sigma_x^2}$$

$$K = \frac{\sigma'^2}{\sigma_o^2 + \sigma'^2}$$

$$\mu_t = \mu_{t-1} + \frac{\sigma'^2}{\sigma_o^2 + \sigma'^2}(\mu_o - \mu_{t-1})$$

$$\mu_t = \frac{\sigma_o^2}{\sigma_o^2 + \sigma'^2}\mu_{t-1} + \frac{\sigma'^2}{\sigma_o^2 + \sigma'^2}\mu_o$$

$$\sigma_t^2 = \sigma'^2_t - \frac{\sigma'^2}{\sigma'^2 + \sigma_o^2}\sigma'^2_t$$

$$\sigma_t^2 = \frac{\sigma_o^2}{\sigma'^2 + \sigma_o^2}\sigma'^2_t$$

Kalman Filter - Limitations

K. N. Toosi
University of Technology

1. What if the Transition and/or Observation Models are not linear?
   - $X_t = \mathbf{A} X_{t-1} + \varepsilon_x \quad \rightarrow \quad X_t = f(X_{t-1}) + \varepsilon_x$
   - $O_t = \mathbf{B} X_{t-1} + \varepsilon_o \quad \rightarrow \quad O_t = g(X_t) + \varepsilon_o$
2. What if the noise is not Additive?
   - $X_t = f(X_{t-1}, \varepsilon_x)$
   - $O_t = g(X_t, \varepsilon_o)$

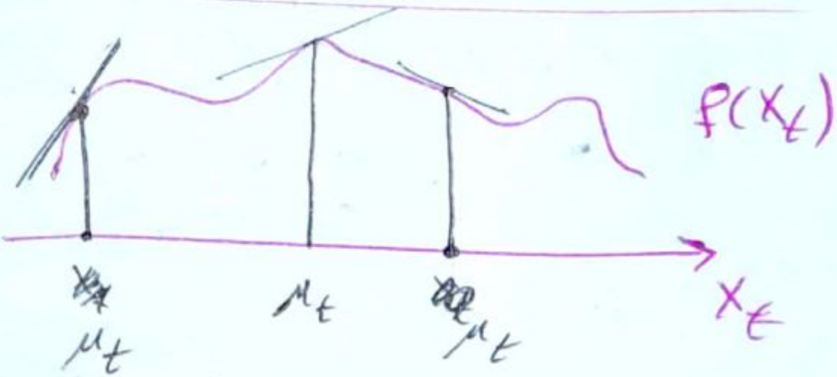
Solution 1: Linearization

Solution 1: Linearization

1D $x \in \mathbb{R}$

$f(x)$ near μ

$x \in (\mu-\sigma, \mu+\sigma)$ $f(x) \simeq f(\mu) + f'(\mu)(x-\mu)$



n-dimentional $x, \mu \in \mathbb{R}^n$ $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$

Extended Kalman Filter (EKF)

$X_{t+1} = f(X_t) + \varepsilon_x$ $\qquad \mu_t \rightarrow$ تخمین میانگین در حال حاضر

$$X_{t+1} \simeq J_f \Big|_{x=\mu_t} (X_t - \mu_t) + f(\mu_t) + \varepsilon_x$$

$f(\mu_t)$ ← ثابت

$$\simeq J_f X_t + \left(f(\mu_t) - J_f \mu_t \right)$$

J_f varies dung the algorithm execution

$O_t = g(X_t) + \varepsilon_o \qquad O_t \simeq g(\mu_t) + J_g \Big|_{x=\mu_t} (X_t - \mu_t)$

$g(\mu_t)$ ← بهترین تخمینی که از X_t داریم

Extended Kalman Filter (EKF)

prediction

$$\mu'_t = f(\mu_{t-1})$$

$$\Sigma'_t = J_f \Sigma_{t-1} J_f^T + \Sigma_x$$

Jacobian at μ_{t-1}

correction

$$K = \Sigma'_t J_g (J_g \Sigma' J_g^T + \Sigma_o)^{-1}$$

Jacobian of g at μ'_t

$$\mu_t = \mu'_t + K(o_t - g(\mu'_t))$$

$$\Sigma_t = \Sigma'_t - K J_g \Sigma'_t$$

Kalman Filters - Prediction Phase

K. N. Toosi University of Technology

$$X_t = f(X_{t-1}) + \varepsilon_x$$

$$\mu'_t = f(\mu_{t-1})$$

$$\Sigma'_t = \Sigma_x + J_f\, \Sigma_{t-1}\, J_f^T$$

Kalman Filters - Correction Phase

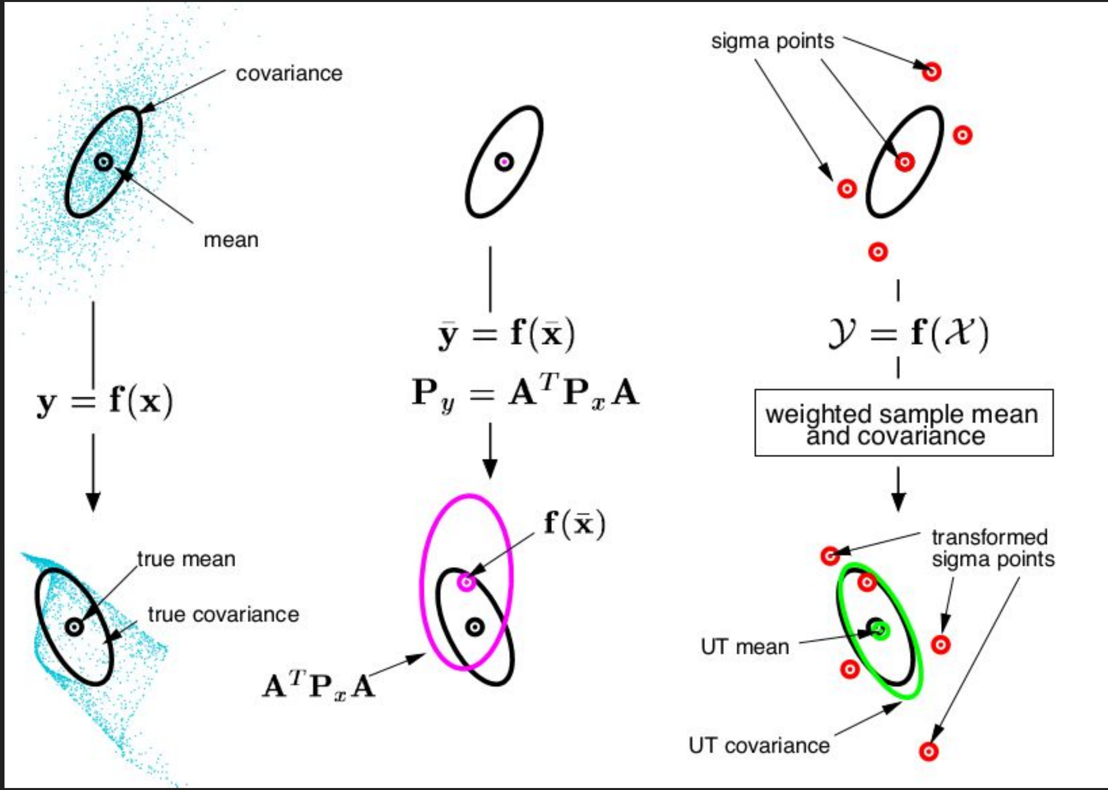
$$O_t = g(X_t) + \varepsilon_o$$

$$K = \Sigma'_t J_g^T (\Sigma_o + J_g \Sigma'_t J_g^T)^{-1}$$

$$\mu_t = \mu'_t + K (O_t - g(\mu'_t))$$

$$\Sigma_t = \Sigma'_t - K J_g \Sigma'_t = (I - K J_g) \Sigma'_t$$

Unscented Kalman Filter (UKF)



Wan et al. "The unscented Kalman filter for nonlinear estimation." 2000.

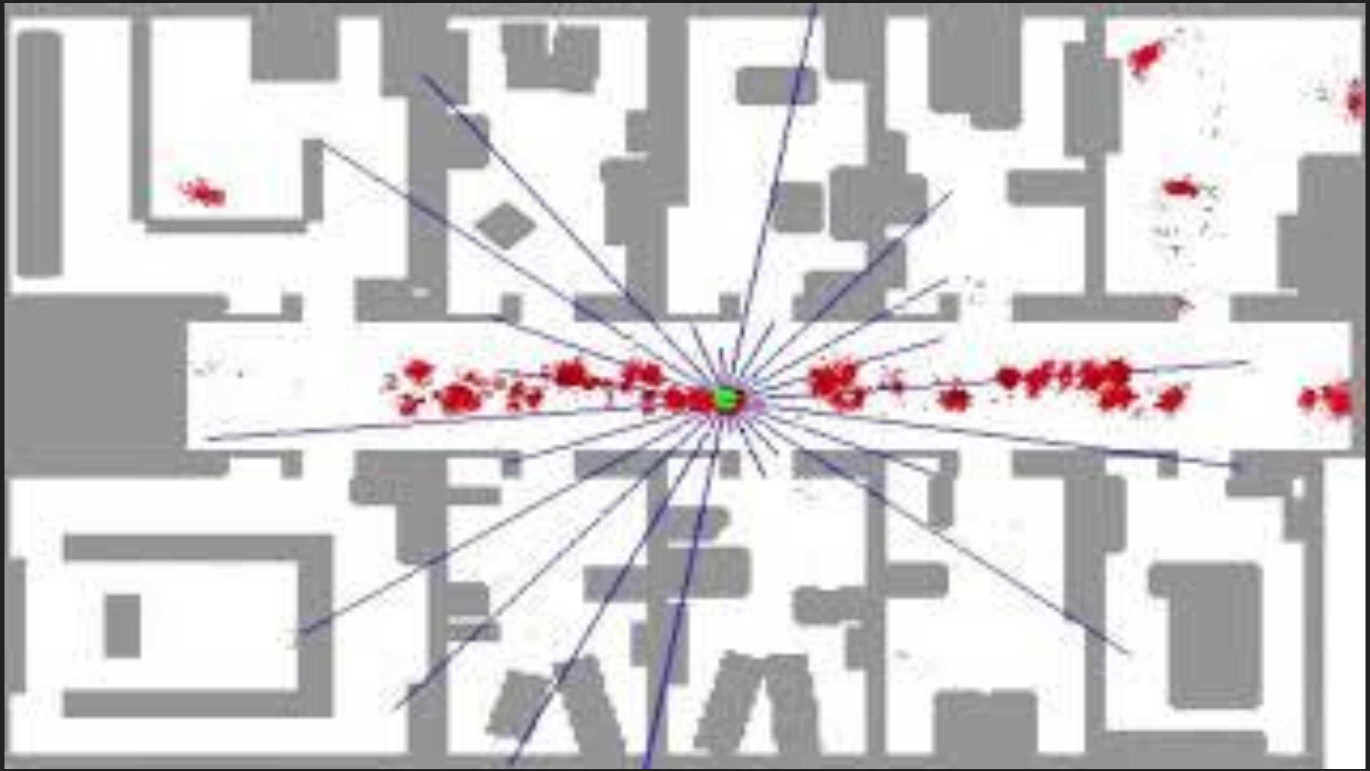
Particle Filter

K. N. Toosi
University of Technology

- What if the transition and/or observation models are far from parametric?
- What if the noise model is far from Gaussian?

=> Represent the (distribution of the) state with a set of particles.

Particle



https://www.youtube.com/watch?v=ZyVWLw0dPN0

Particle