

Introduction to Kalman Filter



Rudolf Emil Kalman Journal of Basic Engineering, 1960

Speaker: Shih-Shinh Huang

January 30, 2022

Rudolf Emil Kalman, "A New Approach to Linear Filtering and Prediction Problems", Transactions of the ASME, Journal of Basic Engineering, 1960



Outline

- Introduction
 - About Kalman Filter
 - Idea of Kalman Filter
- Gaussian Distribution
 - Standard Gaussian
 - Multivariant Gaussian
 - Product of Two Gaussians

- Kalman Filter Algorithm
 - Problem Statement
 - Algorithm Overview
 - Prediction Stage
 - Update Stage





- About Kalman Filter
 - Kalman filter is used to estimate <u>states</u> from the <u>measurements</u> in a dynamic system.
 - Kalman filter has demonstrated its usefulness in various applications.
 - visual object tracking
 - robot/vehicle navigation
 - data prediction task





- About Kalman Filter
 - estimate the states in a recursive manner
 - very fast: without reprocessing all data at each time
 - light on memory: without storing all previous data
 - provide **good** estimates from the measurements with uncertainty
 - linear system: optimal estimate
 - non-linear system: qualified estimate



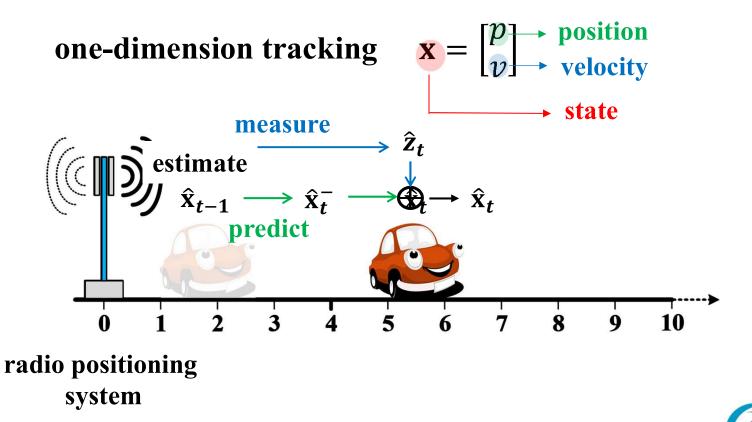


- Idea of Kalman Filter
 - fuse the estimates from two sources that are with uncertainty (inaccuracy).
 - prediction: evolve from the estimate at previous time
 - measurement: perceive system state by sensors.
 - provide the estimate with less uncertainty than both from two sources.





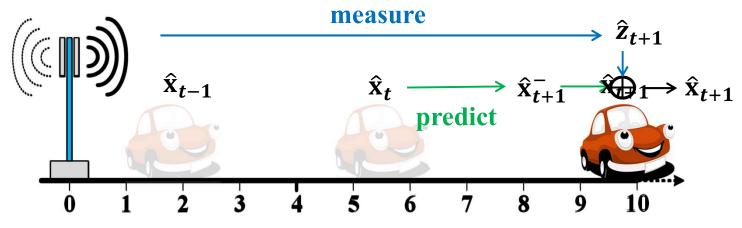
• Idea of Kalman Filter





• Idea of Kalman Filter

one-dimension tracking
$$\mathbf{x} = \begin{bmatrix} p \\ v \end{bmatrix}$$
 position velocity



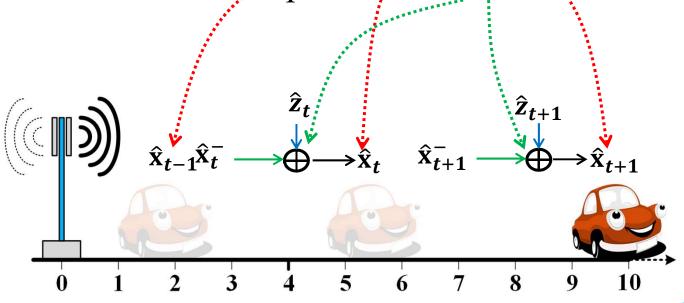
radio positioning system





- Idea of Kalman Filter
 - estimate modeling: Gaussian distribution

• estimate fusion: product of two Gaussians.





- Standard Gaussian $\eta(\mu, \sigma^2)$
 - A <u>random variable</u> x is Gaussian $(x \sim \eta(\mu, \sigma^2))$

$$Pr(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2$$

three samples of *x*

$$x_{1} = 3.0 \mu = \frac{1}{3} \times (3.0 + 4.0 + 2.0) = 3.0$$

$$x_{2} = 4.0 \sigma^{2} = \frac{1}{3} \times \begin{pmatrix} (3.0 - \mu 3.0)^{2} \\ (4.0 - \mu 3.0)^{2} \\ (2.0 - \mu 3.0)^{2} \end{pmatrix} \approx 0.67$$

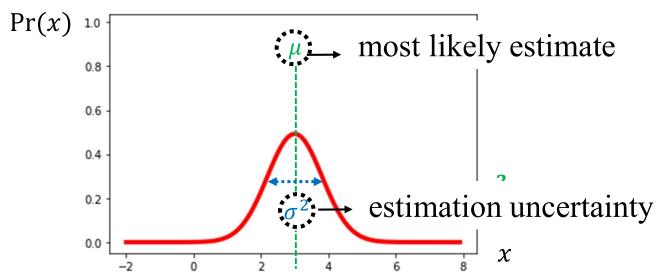
$$x_{3} = 2.0$$





• Standard Gaussian $\eta(\mu, \sigma^2)$

$$\Pr(x) = \frac{1}{\sqrt{2\pi \times 0.67}} \times \exp\left(-\frac{(x - 3.0)^2}{2 \times 0.67}\right)$$



 (μ, σ^2) models an estimate of a random variable x





- Multivariant Gaussian $\eta(\mu, \Sigma)$
 - A *n*-D <u>random vector</u> x is Gaussian $(x \sim \eta(\mu, \Sigma))$

$$\Pr(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \times |\mathbf{\Sigma}|^{1/2}} \times \exp\left(-\frac{1}{2} \times (\mathbf{x} - \mathbf{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mathbf{\mu})\right)$$
mean violatore matrix

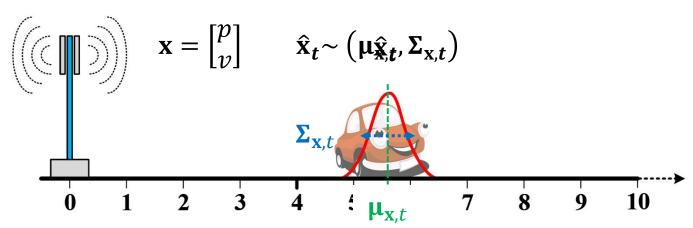
three samples of 2-D \mathbf{x}

$$x_{1} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \qquad \mu = \frac{1}{3} \times \left\{ \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} \right\} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$$

$$x_{2} = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix} \qquad \mathbf{\Sigma} = \begin{bmatrix} 1 + \frac{2}{3} & \left[(\mathbf{X}_{1} \frac{2}{2}) \mu (\mathbf{X}_{1} - \mu)^{T} + \mathbf{X}_{1} - \mathbf{X}_{1} \mu (\mathbf{X}_{1} - \mu)^{T} + \mathbf{X}_{2} \mu (\mathbf{X}_{2} \frac{2}{2}) \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{1} \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{2} \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{3} \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{3} \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{4} \mu (\mathbf{X}_{3} - \mu)^{T} \mu (\mathbf{X}_{3} - \mu)^{T} + \mathbf{X}_{4} \mu (\mathbf{X}_{3} - \mu)^{T} \mu (\mathbf{X}_{3} - \mu)^{T$$



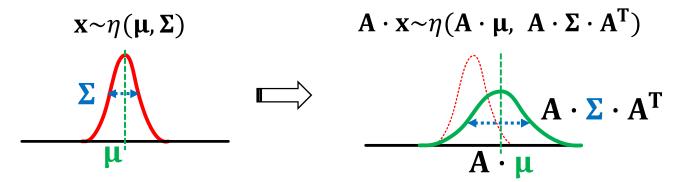
- Multivariant Gaussian $\eta(\mu, \Sigma)$
 - (μ, Σ) models an estimate $\hat{\mathbf{x}}$ of a random vector \mathbf{x}
 - μ : most likely estimate of x
 - Σ : uncertainty of \mathbf{x}



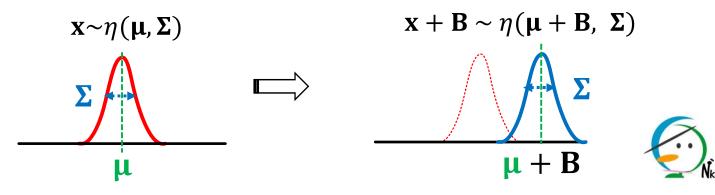




- Multivariant Gaussian $\eta(\mu, \Sigma)$
 - multiplication formula: multiplying x by A



• addition formula: adding **B** to **x**





- Product of Two Gaussians
 - Let x and y be two Gaussian random variables.

$$x \sim \eta(\mu_x, \sigma_x^2)$$
 $y \sim \eta(\mu_y, \sigma_y^2)$

• The product of x and y is a scaled Gaussian.

$$xy \sim \lambda \times \eta(\mu_{xy}, \sigma_{xy}^2)$$

$$\mu_{xy} = \frac{\mu_x \sigma_y^2 + \mu_y \sigma_x^2}{\sigma_x^2 + \sigma_y^2}$$
 $\sigma_{xy}^2 = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$





- Product of Two Gaussians
 - Property: multiplying two Gaussians results in a **smaller** variance σ_{xy}^2 .

$$\begin{aligned} &\textit{proof: } \sigma_{xy}^2 \leq \sigma_x^2 & \textit{proof: } \sigma_{xy}^2 \leq \sigma_y^2 \\ &\sigma_{xy}^2 - \sigma_x^2 = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2} - \sigma_x^2 & \sigma_{xy}^2 - \sigma_y^2 = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2} - \sigma_y^2 \\ &= \frac{\sigma_x^2 \sigma_y^2 - (\sigma_x^4 + \sigma_x^2 \sigma_y^2)}{\sigma_x^2 + \sigma_y^2} = \frac{-\sigma_x^4}{\sigma_x^2 + \sigma_y^2} \leq 0 & = \frac{-\sigma_y^4}{\sigma_x^2 + \sigma_y^2} \leq 0 \end{aligned}$$





$$\mu_{xy} = \frac{\mu_x \sigma_y^2 + \mu_y \sigma_x^2}{\sigma_x^2 + \sigma_y^2} = \frac{\{\mu_x (\sigma_x^2 + \sigma_y^2) + \sigma_x^2 (\mu_y - \mu_x)\}}{\sigma_x^2 + \sigma_y^2}$$

$$= \mu_x + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} \times (\mu_y - \mu_x) \Longrightarrow \mu_{xy} = \mu_x + \mathbf{k} \times (\mu_y - \mu_x)$$

$$\sigma_{xy}^{2} = \frac{\sigma_{x}^{2}\sigma_{y}^{2}}{\sigma_{x}^{2} + \sigma_{y}^{2}} = \frac{\left\{\sigma_{x}^{2}\left(\sigma_{x}^{2} + \sigma_{y}^{2}\right) - \sigma_{x}^{2}\sigma_{x}^{2}\right\}}{\sigma_{x}^{2} + \sigma_{y}^{2}}$$

$$= \sigma_{x}^{2} - \frac{\sigma_{x}^{2}}{\sigma_{x}^{2} + \sigma_{y}^{2}} = k$$

$$\times \sigma_{x}^{2} \implies \sigma_{xy}^{2} = \sigma_{x}^{2} - k \times \sigma_{x}^{2}$$





Product of Two Gaussians

$$\mathbf{x} \sim \eta(\mu_{\mathbf{x}}, \mathbf{\Sigma}_{\mathbf{x}})$$
 $\mathbf{y} \sim \eta(\mu_{\mathbf{y}}, \mathbf{\Sigma}_{\mathbf{y}})$ $\implies \mathbf{x} \mathbf{y} \sim \lambda \times \eta(\mu_{\mathbf{x}\mathbf{y}}, \mathbf{\Sigma}_{\mathbf{x}\mathbf{y}})$





- Problem Statement
 - assumption: the system is **linearly dynamic**
 - linear prediction process: evolving $\hat{\mathbf{x}}_t$ from $\hat{\mathbf{x}}_{t-1}$

$$\hat{\mathbf{x}}_t = \mathbf{F}_t \cdot \hat{\mathbf{x}}_{t-1} + \mathbf{B}_t \cdot \mathbf{u}_t$$
 state transition control input vector

• <u>linear measure process:</u> operating on $\hat{\mathbf{x}}_t$ to produce sensor reading $\hat{\mathbf{z}}_t$

$$\hat{\mathbf{z}}_t = \mathbf{H}_t \cdot \hat{\mathbf{x}}_t$$
 measurement matrix

 \mathbf{F}_t , \mathbf{B}_t , \mathbf{u}_t and \mathbf{H}_t : known system parameters





- Problem Statement
 - assumption: the system is with uncertainty
 - noisy control input

$$\hat{\mathbf{x}}_t = \mathbf{F}_t \cdot \hat{\mathbf{x}}_{t-1} + \mathbf{B}_t \cdot \mathbf{u}_t + \mathbf{w}_t \sim \eta(0, \mathbf{Q})$$

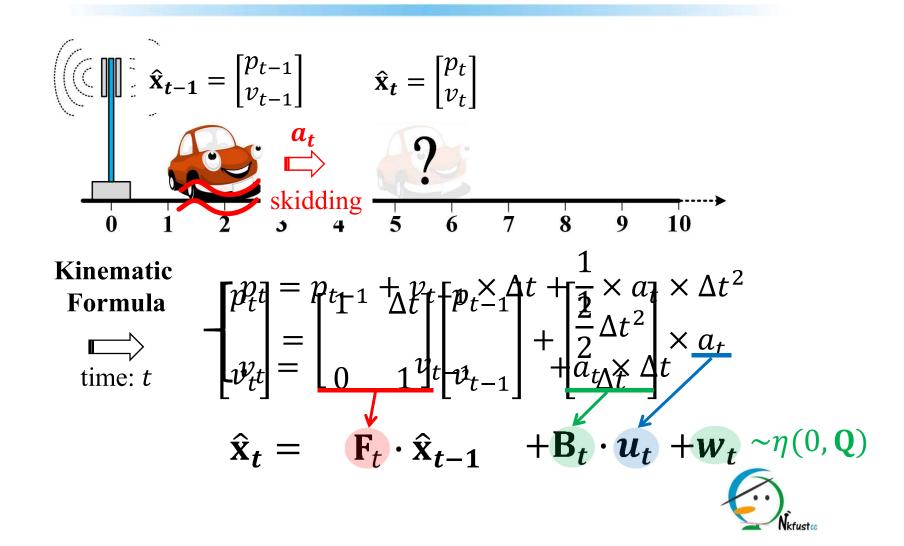
noisy sensing

$$\hat{\mathbf{z}}_t = \mathbf{H}_t \cdot \hat{\mathbf{x}}_t + \mathbf{v}_t \sim \eta(0, \mathbf{R})$$
 zero-mean Gaussian

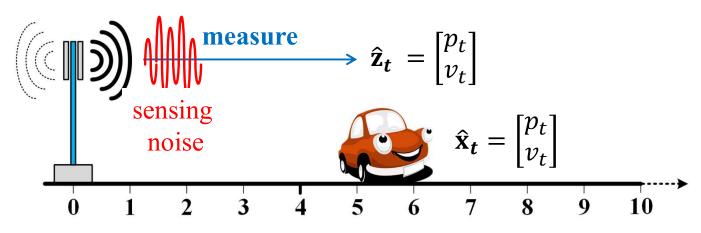
Q and **R** are the **tuning** parameters











radio positioning system

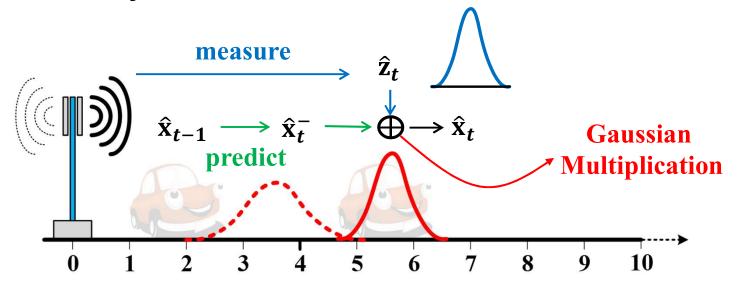
$$\hat{\mathbf{z}}_{t} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \hat{\mathbf{x}}_{t}$$

$$\hat{\mathbf{z}}_{t} = \mathbf{H}_{t} \cdot \hat{\mathbf{x}}_{t} + \mathbf{v}_{t} \sim \eta(0, \mathbf{R})$$





- Problem Statement
 - objective: obtain the estimate $\hat{\mathbf{x}}_t$ by fusing $\hat{\mathbf{x}}_t^-$ and $\hat{\mathbf{z}}_t$ Gaussian estimates







- Algorithm Overview
 - Input
 - $\hat{\mathbf{x}}_0 \sim (\boldsymbol{\mu}_{\mathbf{x},0}, \boldsymbol{\Sigma}_{\mathbf{x},0})$: initial estimate
 - Q and R: tuning parameters
 - $\{z_1, z_2, ..., z_t\}$: a sequence of measurements
 - Output
 - $\hat{\mathbf{x}}_t \sim (\mu_{\mathbf{x},t}, \mathbf{\Sigma}_{\mathbf{x},t})$: an estimate of the state \mathbf{x} at time t

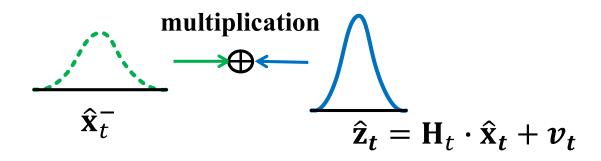




- Algorithm Overview
 - prediction stage: predict $\hat{\mathbf{x}}_t^-$ at time t from $\hat{\mathbf{x}}_{t-1}$

$$\hat{\mathbf{x}}_t^- = \mathbf{F}_t \cdot \hat{\mathbf{x}}_{t-1} + \mathbf{B}_t \cdot \mathbf{u}_t + \mathbf{w}_t$$

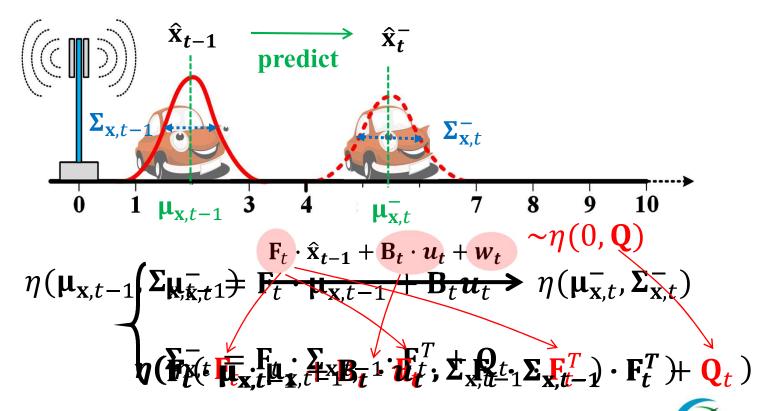
• update stage: fuse $\hat{\mathbf{x}}_t^-$ and $\hat{\mathbf{z}}_t$ by Gaussian product





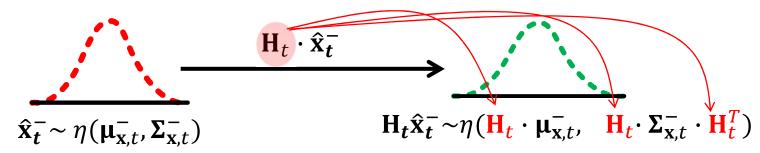


• Prediction Stage: $\hat{\mathbf{x}}_{t-1} \to \hat{\mathbf{x}}_t^-$





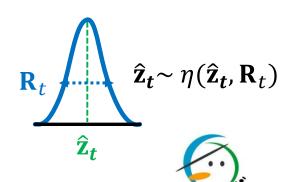
- Update Stage
 - transform $\hat{\mathbf{x}}_t^-$ to sensor-reading **z** space.



• model the measurement as a Gaussian distribution

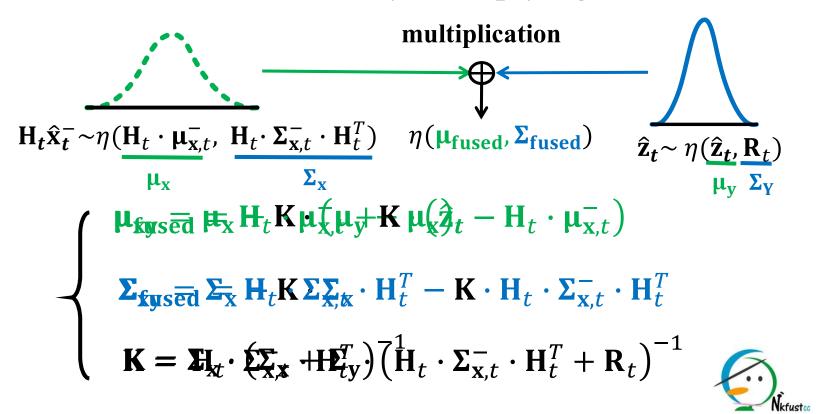
$$\hat{\mathbf{z}}_t = \mathbf{H}_t \cdot \hat{\mathbf{x}}_t + \boldsymbol{v}_t \sim \eta(0, \mathbf{R}_t)$$

- $\hat{\mathbf{z}}_t$: most likely estimate
- \mathbf{R}_t : measuring uncertainty



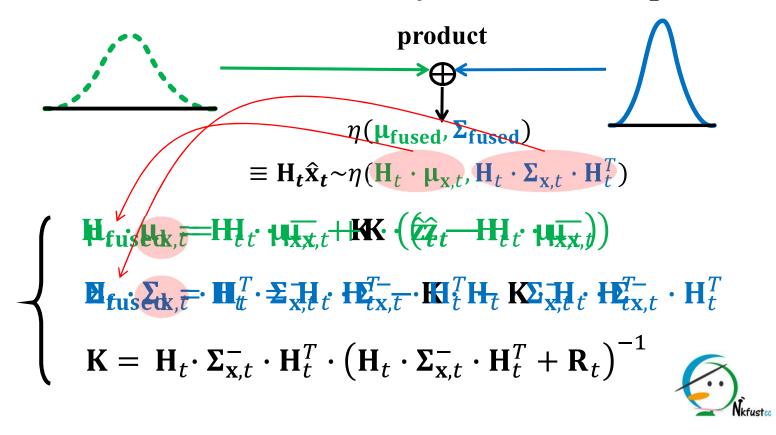


- Update Stage
 - refine the estimate by multiplying two Gaussians





- Update Stage
 - determine the estimate $\hat{\mathbf{x}}_t$ in the state \mathbf{x} space



$$\mathbf{H}_{t}^{-1}$$

$$\mathbf{H}_{t}^{-1}$$

$$= \mathbf{I}$$

$$\Rightarrow \mu_{\mathbf{x},t} = \mathbf{H}_{t}^{-1} \cdot \mathbf{H}_{t} \cdot \mu_{\mathbf{x},t}^{-} + \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot (\hat{\mathbf{z}}_{t} - \mathbf{H}_{t} \cdot \mu_{\mathbf{x},t}^{-})$$

$$\Rightarrow \mu_{\mathbf{x},t} = \mu_{\mathbf{x},t}^{-1} + \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot (\hat{\mathbf{z}}_{t} - \mathbf{H}_{t} \cdot \mu_{\mathbf{x},t}^{-})$$

$$\Rightarrow \mu_{\mathbf{x},t} = \mu_{\mathbf{x},t}^{-} + \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot (\hat{\mathbf{z}}_{t} - \mathbf{H}_{t} \cdot \mu_{\mathbf{x},t}^{-})$$

$$\mathbf{H}_{t}^{-1}$$

$$\mathbf{H}_{t}^{-1}$$

$$\Rightarrow \Sigma_{\mathbf{x},t} \cdot \mathbf{H}_{t}^{T} = \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot \mathbf{H}_{t}^{T} - \mathbf{K} \cdot \mathbf{H}_{t} \cdot \mathbf{\Sigma}_{\mathbf{x},t}^{-} \cdot \mathbf{H}_{t}^{T}$$

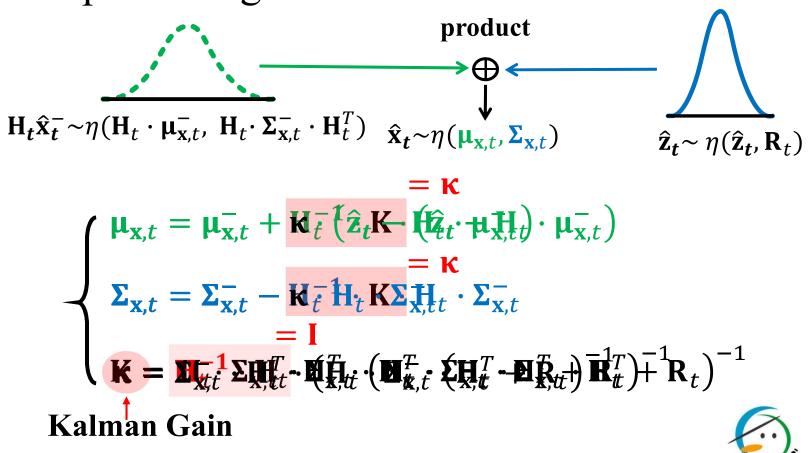
$$\Rightarrow \Sigma_{\mathbf{x},t} \cdot \mathbf{H}_{t}^{T} = \Sigma_{\mathbf{x},t}^{-} \cdot \mathbf{H}_{t}^{T} - \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot \mathbf{H}_{t} \cdot \mathbf{\Sigma}_{\mathbf{x},t}^{-} \cdot \mathbf{H}_{t}^{T}$$

$$\Rightarrow \Sigma_{\mathbf{x},t} = (\Sigma_{\mathbf{x},t}^{-} \cdot \mathbf{H}_{t}^{T} - \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot \mathbf{H}_{t} \cdot \mathbf{\Sigma}_{\mathbf{x},t}^{-} \cdot \mathbf{H}_{t}^{T}) (\mathbf{H}_{t}^{T})^{-1}$$

$$\Rightarrow \Sigma_{\mathbf{x},t} = \Sigma_{\mathbf{x},t}^{-} - \mathbf{H}_{t}^{-1} \cdot \mathbf{K} \cdot \mathbf{H}_{t} \cdot \mathbf{\Sigma}_{\mathbf{x},t}^{-}$$



Update Stage





$$t=1 \qquad \begin{cases} \mu_{x,1}^- = F_1 \cdot \mu_{x,0} + B_1 u_1 \\ \Sigma_{x,1}^- = F_1 \cdot \Sigma_{x,0} \cdot F_1^T + Q_1 \end{cases}$$

$$\hat{x}_0 \sim (\mu_{x,0}, \Sigma_{x,0})$$

$$\hat{x}_1^- \sim (\mu_{x,1}^-, \Sigma_{x,1}^-)$$

$$update \\ stage$$

$$\hat{z}_1 \sim \eta(\hat{z}_1, R_1)$$

$$\xi_{x,1} = \mu_{x,1}^- + \kappa \cdot (\hat{z}_1 - H_1 \cdot \mu_{x,1}^-)$$

$$\Sigma_{x,1} = \Sigma_{x,1}^- - \kappa \cdot H_1 \cdot \Sigma_{x,1}^-$$

$$\kappa = \Sigma_{x,1}^- \cdot H_1^T \cdot (H_1 \cdot \Sigma_{x,1}^- \cdot H_1^T + R_1)^{-1}$$



$$\begin{aligned} \textbf{t=2} & \begin{cases} \mu_{x,2}^{-} = F_{2} \cdot \mu_{x,1} + B_{2}u_{2} \\ \Sigma_{x,2}^{-} = F_{2} \cdot \Sigma_{x,1} \cdot F_{2}^{T} + Q_{2} \end{cases} \\ & \\ \hat{\Sigma}_{2}^{-} \sim (\mu_{x,2}, \Sigma_{x,2}) & \hat{\Sigma}_{2}^{-} \sim (\mu_{x,2}^{-}, \Sigma_{x,2}^{-}) \\ & \\ \frac{update}{stage} & \hat{z}_{2} \sim \eta(\hat{z}_{2}, R_{2}) \end{cases} \\ & \\ \begin{cases} \mu_{x,2} = \mu_{x,2}^{-} + \kappa \cdot (\hat{z}_{2} - H_{2} \cdot \mu_{x,2}^{-}) \\ \Sigma_{x,2} = \Sigma_{x,2}^{-} - \kappa \cdot H_{2} \cdot \Sigma_{x,2}^{-} \\ \kappa = \Sigma_{x,2}^{-} \cdot H_{2}^{T} \cdot (H_{2} \cdot \Sigma_{x,2}^{-} \cdot H_{2}^{T} + R_{2})^{-1} \end{cases} \end{aligned}$$

