13 大名鼎鼎的 Kalman Filter

1. Bayes Filer

1. Prediction

$$\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\, bel(x_{t-1})\, dx_{t-1}$$

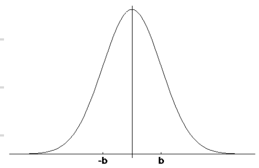
2. Correction

$$bel(x_t) = \eta\, p(z_t \mid x)\, \overline{bel}(x_t)$$

2. Gaussians

1. $P(x) \sim N(\mu, \sigma^2)$:

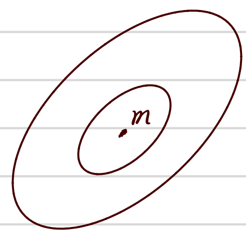
Univariate $$P(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



2. $P(x) \sim N(\mu, \Sigma)$:

Multivariate 多元 $$P(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^t \Sigma^{-1} (x-\mu)}$$

3. Properties of Gaussian



$$\left.\begin{array}{l} X \sim N(\mu, \sigma^2) \\ Y = aX + b \end{array}\right\} \Rightarrow Y \sim N(a\mu + b, a^2\sigma^2)$$

$$\left.\begin{array}{l} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \end{array}\right\} \Rightarrow P(X_1)\cdot P(X_2) \sim N\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1 + \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\mu_2, \frac{1}{\sigma_1^{-2}+\sigma_2^{-2}}\right)$$

4. Multivariate Gaussians

$$\left.\begin{array}{l} X \sim N(\mu, \Sigma) \\ Y = AX + B \end{array}\right\} \Rightarrow Y \sim N(A\mu + B, A\Sigma A^T)$$

$$\left.\begin{array}{l} X_1 \sim N(\mu_1, \Sigma_1) \\ X_2 \sim N(\mu_2, \Sigma_2) \end{array}\right\} \Rightarrow P(X_1)\cdot P(X_2) \sim N\left(\frac{\Sigma_2}{\Sigma_1+\Sigma_2}\mu_1 + \frac{\Sigma_1}{\Sigma_1+\Sigma_2}\mu_2, \frac{1}{\Sigma_1^{-1}+\Sigma_2^{-1}}\right)$$

方差 $$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 p(x)\, dx = \overset{\text{mean}}{\overset{\downarrow}{E}}\left[(x - E(x))^2\right]$$

协方差 $$cov(x, y) = E\left[(X - E(X))\cdot(Y - E(Y))\right]$$ $=0$, 相互独立；$\neq 0$, 相关

相关系数 $$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x, \sigma_y}$$

5. Discrete Kalman Filter

① Estimate the state x of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$x_t = A x_{t-1} + B_t u_t + \varepsilon_t$$ Transition equation

$\varepsilon_t \sim N(0, R_t)$

with a measurement

$$z_t = C_t x_t + \delta_t \qquad \delta_t \sim N(0, Q_t)$$

② Components of a Kalman Filter

$n \times n$ A_t 状态转移矩阵，由t-1时间的状态转移到t时间状态

$n \times 1$ B_t: 描述系统输入如何改变状态

$k \times n$ C_t: 描述状态如何侧偪证实测量

ε_t:, δ_t: 随机误差

5. Kalman Filter Updates in 1D

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - \bar{\mu}_t) \\ \sigma_t^2 = (1 - K_t)\bar{\sigma}_t^2 \end{cases} \quad \text{with } K_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \bar{\sigma}_{obs,t}^2}$$

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - C_t\bar{\mu}_t) \\ \Sigma_t = (I - K_tC_t)\bar{\Sigma}_t \end{cases} \quad \text{with } K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = a_t \mu_{t-1} + b_t u_t \\ \bar{\sigma}_t^2 = a_t^2 \sigma_t^2 + \sigma_{act,t}^2 \end{cases}$$

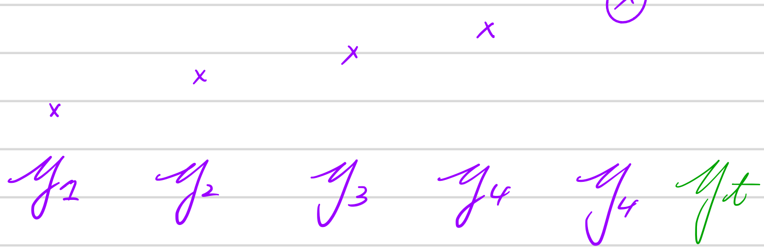
$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t \end{cases}$$

Dynamic Model $\longleftrightarrow$ State Space Model

Series based on time and space cannot reverse( ) or change order

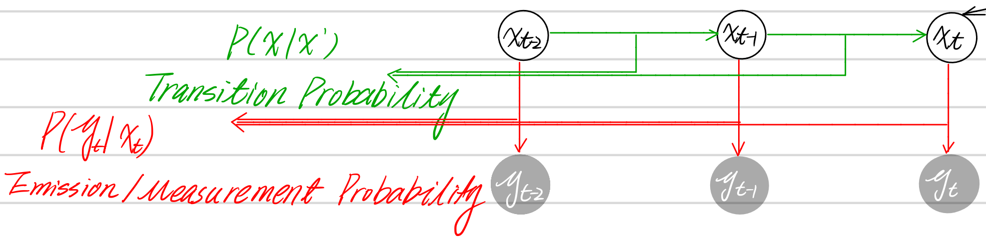
Assume we have a series of measurements

(x) → have high probability here



Write the series of data/mesurements into graph model in grey

The connection between measurements represented by the hidden state



$P(x_1)$ initial prob

$\therefore P(x_1), P(x|x'), P(y_t|x_t)$ defines the state space model

Main Idea

Measurements are dependent since they're in a series. But, assume we've known the hidden state of measurements, then measurements are independent.

$\Downarrow$

The relationship between measurements are decided by their hidden state.

	$P(x_t \mid x_{t-1})$ Transition Prob	Measurement Prob $P(y_t \mid x_t)$	$P(x_1)$
HMM Discrete State DM All hidden states are discrete 1 2	A_{x_{t-1}, x_t} $A = \begin{bmatrix} 0.3 & 0.6 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \end{bmatrix}$ $P(x_t=2 \mid x_{t-1}=1) = 0.6$ $P(x_t=1 \mid x_{t-1}=2) = 0.2$	Any	π $\lambda = \{\underline{A}, B, \pi\}$ ← Parameter matrix ↓ Case 1 (Special form (Matrix) Mesurement are discrete) Case 2 (Continuous measurement, voice process ⇒ audio ⇒ high dimen vector
Kalman Filter 4 Linear Gaussian DM	$N(Ax_{t-1}+B, Q)$	$N(Hx_t+C, R)$	$N(\mu_0, \Sigma_0)$
Non-Linear, Non-Gaussian DM Particle Filter 4	$f(x_{t-1})$ any function	$g(y_t)$ any	$f_0(x_1)$

Dynamic Model can do?

1. $P(y_1, \cdots, y_t)$ —— Evaluation

2. $\underset{\theta}{\text{argmax}} \log P(y_1 \cdots y_t \mid \theta)$ —— Parameter Learning

3. $P(x_1 \cdots x_t \mid y_1 \cdots y_t)$ —— State Decoding

4. $P(x_t \mid y_1 \cdots y_t)$ —— Measurements we have from time $t_0 - t_t$, now we need to get the probability of the state at time t = filtering

不一定Matrix

1. Kalman Filter $HMM=\{A, B, \pi\}$ $LDM=\{A,B,Q,H,C,R\}$

Assume

Transition $\begin{cases} P(x_t|x_{t-1}) = N(Ax_{t-1}+B,\ Q) \\ x_t = Ax_{t-1} + B + \underline{w};\ \text{noise } w \sim N(0,Q) \end{cases}$

Measurement $\begin{cases} P(y_t|x_t) = N(Hx_t + C,\ R) \\ y_t = Hx_t + C + \underline{v};\ \text{noise } v \sim N(0,R) \end{cases}$

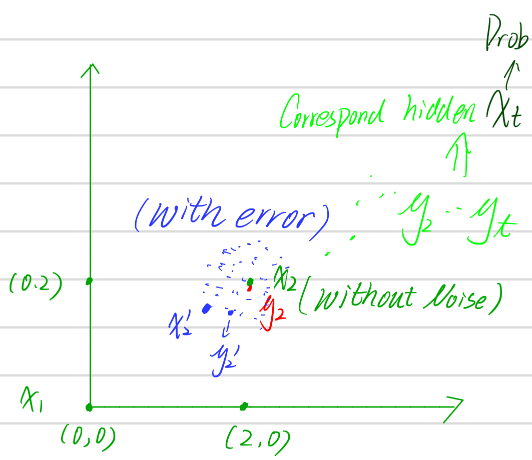
2. How to understand this model?

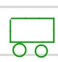
Example: $LDM = \{A, B, Q, H, C, R\}$

Assume $A=I$, $B=\begin{bmatrix}2\\2\end{bmatrix}$, $R=Q=I=H$

$x_t = Ix_{t-1} + \begin{bmatrix}2\\2\end{bmatrix} + w \Leftrightarrow w \sim N(0, I)$

$y_t = I y_{t-1} + v \Leftrightarrow v \sim N(0,1)$



Example 2: 

$y_1 \quad \underbrace{\qquad\qquad}_{\Delta t} \quad y_2 \qquad y_3$

$\ddot{x} = a \sim \{0, v\}$

$\bar{x}_t = \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix}$ current position, current speed $\begin{cases} x_t = x_{t-1} + \dot{x}_t\Delta t + \frac{1}{2}a(\Delta t)^2 \\ \dot{x}_t = \dot{x}_{t-1} + a\Delta t \end{cases}$

$\Updownarrow$

$$\underbrace{\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix}}_{\bar{x}_t} = \underbrace{\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_{t-1} \\ \dot{x}_{t-1} \end{bmatrix}}_{\bar{x}_{t-1}} + \begin{bmatrix} \frac{1}{2}a\Delta t^2 \\ a\Delta t \end{bmatrix}$$

协方差 Covariance $Var(\bar{x}_t) = E[(\bar{x}_t - \mu)\cdot(\bar{x}_t - \mu)^T]$ 两个变量总体的误差

$$= E\left[\begin{bmatrix} \frac{1}{2}a(\Delta t)^2 \\ a\Delta t \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2}a\Delta t^2 & a\Delta t \end{bmatrix}\right]$$

$$= E\left\{\begin{bmatrix} \frac{1}{4}a^2(\Delta t)^2 & \frac{1}{2}a^2(\Delta t)^2 \\ \frac{1}{2}a^2(\Delta t)^2 & a^2(\Delta t)^2 \end{bmatrix}\right\} = \underbrace{E[a^2]}_{v^2} \underbrace{\begin{bmatrix} \frac{1}{4}\Delta t^2 & \frac{1}{2}\Delta t^2 \\ \frac{1}{2}\Delta t^2 & \Delta t^2 \end{bmatrix}}_{Q}$$

Measurement $y_t = [1\ 0]\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + v,\ v \sim N(0, R)$

3. KF $P(x_t|y_1,\cdots,y_t)$ Filtering

$$\underbrace{\overbrace{P(x_t|y_1\cdots,y_t)}^{N(\hat{\mu},\hat{\Sigma})}}_{\text{Update}} \propto P(x_t,y_1,\cdots,y_t)$$

常数 => 可去掉

$$\propto \underbrace{P(y_t|x_t,y_1,\cdots,y_{t-1})}_{\downarrow \text{Simplified}} \cdot \underbrace{P(x_t|y_1,\cdots,y_{t-1})}_{\text{prediction}} \cdot \boxed{P(y_1,\cdots,y_{t-1})}$$

$$\propto P(y_t|x_t)\cdot \underline{P(x_t|y_1,\cdots,y_{t-1})}$$

$\Downarrow$

$$\underbrace{P(x_t|y_1,\cdots,y_{t-1})}_{N(\bar{\mu}_t,\bar{\Sigma}_t)} = \int_{x_{t-1}} P(x_t,x_{t-1}|y_1,\cdots,y_{t-1})\,dx_{t-1}$$

$$= \int_{x_{t-1}} P(x_t|x_{t-1},y_1,\cdots,y_{t-1})\,P(x_{t-1}|y_1,\cdots,y_{t-1})\,dx_{t-1}$$

递归

$$= \int_{x_{t-1}} \underbrace{P(x_t|x_{t-1})}_{\text{Gaussian}} \cdot \underbrace{P(x_{t-1}|y_1,\cdots,y_{t-1})}_{N(\hat{\mu}_{t-1},\hat{\Sigma}_{t-1})\ \text{Gaussian}}\,dx_{t-1}$$

均值

Update $\longrightarrow$ $t=1: P(x_1|y_1)\sim N(\hat{\mu}_1,\hat{\Sigma}_1)$

Prediction $\longrightarrow$ $t=2: P(x_2|y_1)\sim N(\bar{\mu}_2,\bar{\Sigma}_2)$ $\rightarrow$ prediction

$P(x_2|y_1,y_2)\sim N(\hat{\mu}_2,\hat{\Sigma}_2)$

$\vdots$

$t=t: P(x_t|y_1,\cdots,y_{t-1})\sim N(\bar{\mu}_t,\bar{\Sigma}_t)$

$P(x_t|y_1,\cdots,y_t)\sim N(\hat{\mu}_t,\hat{\Sigma}_t)$

In a conclusion, what we gonna do is take parameters into these formular, and then compute $\underbrace{\bar{\mu},\hat{\mu},\bar{\Sigma}_t,\hat{\Sigma}_t}_{\text{auxiliary variable}}$ recursively

FK is about the way that comput the auxiliary variable from t-1 state to t state.

$$\underbrace{P(x_t \mid y_1, \cdots, y_{t-1})}_{N(\bar{\mu}_t, \bar{\Sigma}_t)\,:\ \text{Prediction}} = \int P(x_t \mid x_{t-1})\, P(x_{t-1} \mid y_1, \cdots, y_{t-1})$$

$$\underbrace{P(x_t \mid y_1, \cdots, y_t)}_{N(\hat{\mu}_t, \hat{\Sigma}_t)\,:\ \text{Update}} = P(y_t \mid x_t) \cdot P(x_t \mid y_1, \cdots, y_{t-1})$$

$$\begin{cases} x_t = Ax_{t-1} + w, & w \sim N(0, Q) \\ y_t = Hx_t + v, & v \sim N(0, R) \end{cases}$$

Linear Gaussian ($B=C=0$)

v, w has no relationship with x_{t-1} ⇓ $\Rightarrow$

w, v has no connection, independently

$Cov(x_{t-1}, w) = 0$ $\quad$ $Cov(x_{t-1}, v) = 0$ $\quad$ $Cov(w, v) = 0$

Gaussian Property $\Rightarrow$

$$\begin{pmatrix} u \\ v \end{pmatrix} \sim N\left(\overset{\text{mean}}{\begin{bmatrix} \mu_u \\ \mu_v \end{bmatrix}} \cdot \underbrace{\begin{bmatrix} \Sigma_u & \Sigma_{u,v}^T \\ \Sigma_{u,v} & \Sigma_v \end{bmatrix}}_{\text{Covariance Matrix}} \right)$$

If: u is 17-dimen, v is 6 dimen
$\Rightarrow$ The mean of union distribution (u, v) is 23 dimen

Conditional Gaussian

$$P(u \mid v) \sim N\Big(\underbrace{\mu_u + \Sigma_{u,v} \Sigma_v^{-1} (v - \mu_v)}_{\text{mean}},\ \underbrace{\Sigma_u - \Sigma_{uv} \Sigma_v^{-1} \Sigma_{uv}^T}_{\sigma^2} \Big)$$

u, v: is a vector and random variable

$$P(x_{t-1} | y_1 \cdots y_{t-1}) = N(\hat{\mu}_{t-1}, \hat{\Sigma}_{t-1})$$

$$x_{t-1} | y_1 \cdots y_{t-1} = E[x_{t-1}] + \Delta x_{t-1} \quad , \Delta x_{t-1} \sim N(0, \hat{\Sigma}_{t-1})$$

类似的，

$$\begin{aligned} x_t | y_1 \cdots y_{t-1} &= A(x_{t-1}) + w \\ &= A(E[x_{t-1}] + \Delta x_{t-1}) + w \\ &= \underbrace{AE[x_{t-1}]}_{E[x_t]} + \underbrace{A\Delta x_{t-1} + w}_{\Delta x_t} \end{aligned}$$

$$\begin{aligned} y_t | y_1, \cdots, y_{t-1} &= Hx_t + v \\ &= H(AE[x_{t-1}] + A\Delta x_{t-1} + w) + v \\ &= \underbrace{HAE[x_{t-1}]}_{E[y_t]} + \underbrace{HA\Delta x_{t-1} + Hw + v}_{\Delta y_t} \end{aligned}$$

$$x_t | y_1 \cdots y_{t-1} = \underbrace{AE[x_{t-1}]}_{E[x_t]} + \underbrace{A\Delta x_{t-1} + w}_{\Delta x_t}$$

$$y_t | y_1, \cdots, y_{t-1} = \underbrace{HAE[x_{t-1}]}_{E[x_t]} + \underbrace{HA\Delta x_{t-1} + Hw + v}_{\Delta y_t}$$

边缘分布

$$P(x_t | y_1, \cdots, y_{t-1}) = N(\overbrace{AE[x_{t-1}]}^{\bar{\mu}_t}, \overbrace{E[(\Delta x)(\Delta x)^T]}^{\bar{\Sigma}_t})$$

随机变量的方差 x 为 0 均值 $\Rightarrow Var(x) = E[xx^T]$

$$P(y_t | y_1, \cdots, y_{t-1}) = N(HAE[x_{t-1}], E[(\Delta y)(\Delta y)^T])$$

联合分布

$$P(x_t, y_t | y_1, \cdots, y_{t-1}) = N\left(\begin{bmatrix} \underline{AE[x_{t-1}]} \\ \hat{\mu}_{t-1} \\ \downarrow \\ \overline{HAE[x_{t-1}]} \end{bmatrix} \begin{bmatrix} \overset{A\hat{\Sigma}_{t-1}A^T + Q}{E[(\Delta x)(\Delta x)^T]} & \overset{\bar{\Sigma}_t^T H}{E(\Delta x)(\Delta y)^T} \\ \overset{H\bar{\Sigma}_t}{E[(\Delta y)(\Delta x)^T]} & E[(\Delta y)(\Delta y)^T] \end{bmatrix} \right)$$

$$\Updownarrow$$

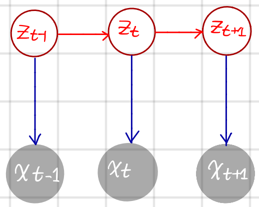
$$P(x_t | y_t, y_1, \cdots, y_{t-1})$$

$$\begin{aligned} E[(\Delta x)(\Delta x)^T] &= E[(A\Delta x_{t-1} + w)(A\Delta x_{t-1} + w)^T] \\ &= E[A(\Delta x_{t-1})(\Delta x_{t-1})^T A^T \quad + ww^T] \\ &= AE[\Delta x_{t-1} \cdot \Delta x_{t-1}^T]A^T + E(w \cdot w^T)] \\ &= A\hat{\Sigma}_{t-1}A^T + Q \end{aligned}$$

$COV(x_{t-1}, w) = 0$

$w \cdot \Delta x_{t-1} = 0$

Dynamic Model (State Space Model)
- HMM — State is discrete
- Linear Dynamic System (Kalman Filter) — Linear Gaussian
- Particle Filter — Non-Linear, Non Gaussian



- Learning
- Inference
  - Decoding: $P(z_1, z_2, \cdots, z_t \mid x_1, x_2, \cdots, x_t)$
  - Prob of evidence: $P(X \mid \theta) = P(x_1, \cdots, x_t \mid \theta)$
  - Filtering: $P(z_t \mid x_1, \cdots, x_t)$ —— online
  - Smoothing: $P(z_t \mid x_1, \cdots, x_t)$ —— offline
  - Prediction: $P(z_{t+1}, z_{t+2} \mid x_1, \cdots, x_t)$
    - $P(x_{t+1}, x_{t+2} \mid x_1, \cdots, x_t)$

Transition Probability: $P(z_{t+1} \mid z_t)$

Emission Probability: $P(x_{t-1} \mid z_{t-1})$

Linear Gaussian

1. Linear Property:

$z_t = A z_{t-1} + B + \varepsilon$

线性关系

$x_t = C z_t + D + \delta$

Gaussian Distribution ⇒ Noise

$\varepsilon \sim N(0, Q)$

$\delta \sim N(0, R)$

2. Parameter: HMM: $\lambda = (\pi, A, B)$

$A = [a_{ij}]$ $\quad a_{ij} = P(i_{t+1} = q_j \mid i_t = q_i)$ 状态矩阵

represent state

$B = [b_j(k)]$ $\quad b_j(k) = P(O_t = v_t \mid i_t = q_j)$

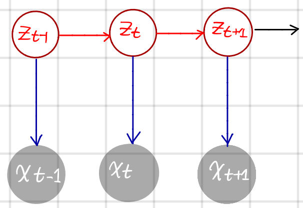
Kalman: $P(z_t \mid z_{t-1}) \sim N(A z_{t-1} + B, Q)$

$P(x_t \mid z_t) \sim N(C z_t + D, R)$

Conditional Prob ~ Gaussian

Initial $\quad z_1 \sim N(\mu_1, \Sigma_1)$

$\theta = (A, B, C, D, Q, R, \mu, \Sigma)$



$$\begin{cases} P(z_t \mid z_{t-1}) = N(A \cdot z_{t-1} + B, Q) \\ P(x_t \mid z_t) = N(C \cdot z_t + D, R) \\ P(z_1) = N(\mu, \Sigma) \end{cases}$$

$\theta = (A, B, C, D, Q, R, \mu, \Sigma)$

$z_t = A z_{t-1} + B + \varepsilon$, $\varepsilon \sim N(0, Q)$

$x_t = C \cdot z_t + D + \delta$, $\delta \sim N(0, R)$

Learning

Inference: $P(Z \mid X)$
后验概率分布

1. decoding → HMM
2. Prob of evidence
3. Filtering $P(z_t \mid x_1, \cdots, x_t)$
   Marginal posterior
4. Smoothing $P(z_t \mid x_1, \cdots, x_T)$
5. Prediction $P(z_t \mid x_1, \cdots, x_{t-1})$

1. Linear System:

先验 预测 理论预测

$$x_k = A x_{k-1} + B u_{k-1} + w_{k-1}$$

x_k: n×1 state vector 当前状态

A: n×n transition matrix

B: Transition from input to state 输入→状态

u_{k-1}: k×1 System Input 输入

w_{k-1}: System noise 噪声

后验 测量 测量反馈

$$z_k = H x_k + v_k$$

v_k: measurement noise 系统噪声

System Noise $P(w) \sim N(0, Q)$

Measurement Noise $P(v) \sim N(0, R)$ } 相互独立

理论预测与测量值，通过加权获得最优

2. $\hat{x}_k'$ 为预测值；$\hat{x}_k$ 估计值；$\hat{z}_k$ 测量值的预测

$$\hat{x}_k = \hat{x}_k' + K_k (z_k - \hat{z}_k)$$

$$= \hat{x}_k' + K_k (z_k - H\hat{x}_k')$$

残差：预测值与真实值之间的距离 if $(z_k - H\hat{x}_k') = 0$ 则完全吻合

$$= \hat{x}_k' + K_k (H x_k + v_k - H\hat{x}_k')$$

$$= \hat{x}_k' + K_k H x_k + K_k v_k - H K_k \hat{x}_k'$$

3. 估计值和真实值间误差的协方差矩阵

$$P_k = E[e_k e_k^T] = E[(x - \hat{x}_k)(x_k - \hat{x}_k)^T]$$

e_k 为向量，各个系统状态变量的误差组成

匀速运动 $e_k \Rightarrow$ 速度误差，位移误差

(1) $$P_k = \begin{bmatrix} E(S_{err} S_{err}^T) & E(S_{err} V_{err}^T) \\ E(V_{err} S_{err}^T) & E(V_{err} V_{err}^T) \end{bmatrix}$$

$$= [[(I - K_k H)(x_k - x_k') - K_k v_k] [(I - K_k H)(x_k - x_k') - K_k v_k]^T]$$

4. 预测值与真实值之间误差的协方差矩阵

$$P_k' = E[e_k' e_k'^T] = E[(x_k - \hat{x}_k')(x_k - \hat{x}_k')^T]$$

系统状态与噪声之间相互独立

5. 展开(1)式

$$P_k = (I - K_k H) E[(x_k - \hat{x}_k')(x_k - \hat{x}_k')^T](I - K_k H) + K_k E[v_k v_k^T] K_k^T$$

$$P_k = (I - K_k H) P_k' (I - K_k H)^T + K_k R K_k^T$$

$$P_k = P_k' - K_k H P_k' - P_k' H^T K_k^T + K_k (H P_k' H^T + R) K_k^T$$

最小均方差，协方差对角线为方差，T 代表对角线元素 Sum

$$T[P_k] = T[P_k'] - 2T[K_k H P_k'] + T[K_k (H P_k' H^T + R) K_k^T]$$

使之最小，求导等于 0

$$\frac{d\,T[P_k]}{d K_k} = -2 (H P_k')^T + 2 K_k (H P_k' H^T + R)$$

Assume $H = 1$, $P_k' \neq 0 \Rightarrow K_k = 1/(1 + R/P_k')$ $P_k' \uparrow$ $K \uparrow$，权重越重视反馈

Take K into P_k

$P_k = (I - K_k H) P_k'$ 预测值与真实值之间误差的协方差矩阵

$$\hat{x}_{k+1}' = A\hat{x}_k + B u_k$$

$$P_{k+1}' = E[e_{k+1}' e_{k+1}'^T] = E[(x_{k+1} - \hat{x}_{k+1}')(x_{k+1} - \hat{x}_{k+1}')^T]$$

$$= E[(A(x_k - \hat{x}_k) + w_k)(A(x_k - \hat{x}_k) + w_k)^T]$$

State & noise 独立

$$P_{k+1}' = E[(Ae_k)(Ae_k)^T] + E[w_k w_k^T]$$

$$= A P_k A^T + Q$$

只要设置初始 P_k 就能不断递推

Kalman Filter 思路：

① Compute (1) Predicate (2) Covariance Matrix of error of predicate and real value。

预测 (1) $\hat{x}_k' = A\hat{x}_{k-1} + Bu_{k-1}$

预与真的关系 (2) $P_k' = AP_{k-1}A^T + Q$

② Compute Kalman Gain K, then compute estimation
(1)

KG (1) $K_k = P_k' H^T (HP_k' H + R)^{-1}$

估计值 (2) $\hat{x}_k = \hat{x}_k' + K_k(z_k - H\hat{x}_k')$

③ Compute corvariance matrix of estimate and real value

估与真的关系 $P_k = (I - K_k H)P_k'$