
卡尔曼滤波简介

薛文超

中国科学院数学与系统科学研究院

2020年12月

概要

- 研究背景和问题描述
- 卡尔曼滤波算法及其主要特性
- 扩展卡尔曼滤波算法

研究背景

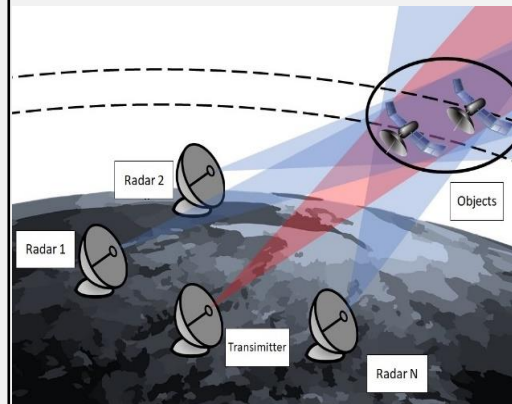
滤波：最优的状态估计

实际系统

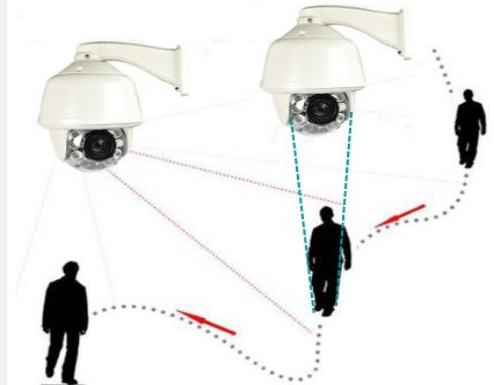
Volatility Estimation



Orbit Determination



Camera Tracking



Observation data	Disturbed stock price	Radar measurements	Measurements of sensors in network
System state	Volatility	Position and velocity	Position and velocity

准备知识：随机变量

- 随机变量：随机试验各种结果的实值单值函数

X：将随机试验各种结果映射到实数

(扔硬币出现正面记为1，扔硬币出现反面记为0)

- 随机变量**X**最基本的特性：概率分布函数

(probability distribution function, PDF) $F_X(x) = P(X \leq x)$

- PDF满足的特性：

$$F_X(x) \in [0, 1]$$

$$F_X(-\infty) = 0$$

$$F_X(\infty) = 1$$

$$F_X(a) \leq F_X(b) \quad \text{if } a \leq b$$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

准备知识： 概率分布函数(PDF)和概率密度函数(pdf)

- 概率密度函数(Probability density function, pdf): PDF的导数

$$f_X(x) = \frac{dF_X(x)}{dx}$$

- pdf满足的特性:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_X(z) dz \\ f_X(x) &\geq 0 \\ \int_{-\infty}^{\infty} f_X(x) dx &= 1 \\ P(a < x \leq b) &= \int_a^b f_X(x) dx \end{aligned}$$

准备知识：条件概率分布函数和条件概率密度函数

➤ 给定事件A条件下的概率分布和概率密度

$$\begin{aligned} F_X(x|A) &= P(X \leq x|A) \\ &= \frac{P(X \leq x, A)}{P(A)} \\ f_X(x|A) &= \frac{dF_X(x|A)}{dx} \end{aligned}$$

$$\begin{aligned} f_{X_1|X_2}(x_1|x_2) &= P[(X_1 \leq x_1)|(X_2 = x_2)] \\ &= \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_2}(x_2)} \end{aligned}$$

准备知识：期望和方差

➤ $g(X)$ 为随机变量的一个函数（将随机变量映射为随机变量），

则 $g(X)$ 的期望为 $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

➤ X 的期望为： $\bar{x} \triangleq E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$

➤ X 的方差：
$$\begin{aligned}\sigma_X^2 &= E[(X - \bar{x})^2] \\ &= \int_{-\infty}^{\infty} (x - \bar{x})^2 f_X(x) dx\end{aligned}$$

➤ 常用标记： $X \sim (\bar{x}, \sigma^2)$

➤ 条件期望： $E(X|A) = \int_{-\infty}^{\infty} x f_X(x|A) dx$

准备知识：高维随机变量

$$F_X(x) = P(X \leq x)$$

$$F_Y(y) = P(Y \leq y)$$

➤ **联合分布函数：** $F_{XY}(x, y) = P(X \leq x, Y \leq y)$

(扔两个硬币，出现正面+1，出现反面+0)

$$F(x, y) \in [0, 1]$$

$$F(x, -\infty) = F(-\infty, y) = 0$$

$$F(\infty, \infty) = 1$$

$$F(a, c) \leq F(b, d) \quad \text{if } a \leq b \text{ and } c \leq d$$

$$P(a < x \leq b, c < y \leq d) = F(b, d) + F(a, c) - F(a, d) - F(b, c)$$

$$F(x, \infty) = F(x)$$

$$F(\infty, y) = F(y)$$

准备知识：高维密度函数

➤ 联合密度函数： $f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}$

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(z_1, z_2) dz_1 dz_2$$

$$f(x, y) \geq 0$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

$$P(a < x \leq b, c < y \leq d) = \int_c^d \int_a^b f(x, y) dx dy$$

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

准备知识：独立性

➤ **随机变量X和Y独立：** $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for all x, y

➤ **随机变量X和Y独立：**

$$\begin{aligned} F_{XY}(x, y) &= F_X(x)F_Y(y) \\ f_{XY}(x, y) &= f_X(x)f_Y(y) \end{aligned}$$

准备知识：高维随机变量的相关函数

➤ 互相关函数矩阵: $R_{XY} = E(XY^T)$

$$= \begin{bmatrix} E(X_1Y_1) & \cdots & E(X_1Y_m) \\ \vdots & & \vdots \\ E(X_nY_1) & \cdots & E(X_nY_m) \end{bmatrix}$$

➤ 互协方差矩阵: $C_{XY} = E[(X - \bar{X})(Y - \bar{Y})^T]$

$$= E(XY^T) - \bar{X}\bar{Y}^T$$

准备知识：高维随机变量的协方差阵

➤ 自相关函数矩阵 $R_X = E[XX^T]$

$$= \begin{bmatrix} E[X_1^2] & \cdots & E[X_1 X_n] \\ \vdots & & \vdots \\ E[X_n X_1] & \cdots & E[X_n^2] \end{bmatrix}$$

➤ 自协方差矩阵 $C_X = E[(X - \bar{X})(X - \bar{X})^T]$

$$= \begin{bmatrix} E[(X_1 - \bar{X}_1)^2] & \cdots & E[(X_1 - \bar{X}_1)(X_n - \bar{X}_n)] \\ \vdots & & \vdots \\ E[(X_n - \bar{X}_n)(X_1 - \bar{X}_1)] & \cdots & E[(X_n - \bar{X}_n)^2] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix}$$

自协方差矩阵为半正定矩阵

Note that $\sigma_{ij} = \sigma_{ji}$ so $C_X = C_X^T$.

$$\begin{aligned} z^T C_X z &= z^T E[(X - \bar{X})(X - \bar{X})^T] z \\ &= E[z^T (X - \bar{X})(X - \bar{X})^T z] \\ &= E[(z^T (X - \bar{X}))^2] \\ &\geq 0 \end{aligned}$$

准备知识：随机过程

随机过程：一系列随机变量 $X(t)$ ， t 为时间

$$F_X(x, t) = P(X(t) \leq x)$$

$$F_X(x, t) = P[X_1(t) \leq x_1 \text{ and } \cdots X_n(t) \leq x_n(t)]$$

$$f_X(x, t) = \frac{d^n F_X(x, t)}{dx_1 \cdots dx_n}$$

$$\bar{x}(t) = \int_{-\infty}^{\infty} x f(x, t) dx$$

$$\begin{aligned} C_X(t) &= E\{[X(t) - \bar{x}(t)][X(t) - \bar{x}(t)]^T\} \\ &= \int_{-\infty}^{\infty} [x - \bar{x}(t)][x - \bar{x}(t)]^T f(x, t) dx \end{aligned}$$

$$R_X(t_1, t_2) = E[X(t_1)X^T(t_2)]$$

$$C_X(t_1, t_2) = E\{[X(t_1) - \bar{X}(t_1)][X(t_2) - \bar{X}(t_2)]^T\}$$

均为时变的函数：

离散时变线性系统的滤波问题

$$\begin{cases} X_{k+1} = A_k X_k + B_k u_k + w_k \\ Y_{k+1} = C_{k+1} X_{k+1} + n_{k+1} \end{cases}$$
$$k = 0, 1, 2, \dots$$

$\{n_k\}$ 和 $\{w_k\}$ 为随机过程

$$E(n_k) = 0, \forall k > 0$$

$$E(w_k) = 0, \forall k > 0$$

$$C_n(k, j) = \begin{cases} 0, \forall k \neq j \\ R_k, \forall k = j \end{cases}$$

$$C_w(k, j) = \begin{cases} 0, \forall k \neq j \\ Q_k, \forall k = j \end{cases}$$

离散时变线性系统的滤波问题

$$\begin{cases} X_{k+1} = A_k X_k + B_k u_k + w_k \\ Y_{k+1} = C_{k+1} X_{k+1} + n_{k+1} \end{cases}$$
$$k = 0, 1, 2, \dots$$

$\{n_k\}$ 和 $\{w_k\}$ 为随机过程

$$E(n_k) = 0, \forall k > 0$$

$$E(w_k) = 0, \forall k > 0$$

$$C_n(k, j) = \begin{cases} 0, \forall k \neq j \\ R_k, \forall k = j \end{cases}$$

$$C_w(k, j) = \begin{cases} 0, \forall k \neq j \\ Q_k, \forall k = j \end{cases}$$

目标：利用量测获得 X_k 的线性最小方差估计

$$\hat{X}_k = \arg \min_{Z_k \in \{\text{linear function of } Y_i, 1 \leq i \leq k\}} E \left\{ (Z_k - X_k)(Z_k - X_k)^T \right\}$$

概要

- 研究背景和问题描述
- 卡尔曼滤波算法及其主要特性
- 扩展卡尔曼滤波算法

卡尔曼滤波算法及参数设计



离散线性系统:
$$\begin{cases} X_{k+1} = A_k X_k + B_k \mathbf{u}_k + \mathbf{w}_k \\ Y_{k+1} = C_{k+1} X_{k+1} + \mathbf{n}_{k+1} \end{cases}$$
$$k = 0, 1, 2, \dots$$

卡尔曼滤波:

$$\bar{X}_{k+1} = A_k \hat{X}_k + B_k \mathbf{u}_k: \text{预测过程}$$

$$\bar{P}_{k+1} = A_k P_k A_k^T + Q_k: \text{预测误差的协方差阵}$$

$$\hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1} (Y_{k+1} - C_{k+1} \bar{X}_{k+1}): \text{更新过程}$$

$$K_{k+1} = (\bar{P}_{k+1} C_{k+1}^T) (C_{k+1} \bar{P}_{k+1} C_{k+1}^T + R_{k+1})^{-1}: \text{更新增益}$$

$$P_{k+1} = \bar{P}_{k+1} - K_{k+1} C_{k+1} \bar{P}_{k+1}: \text{估计 (滤波) 误差的协方差阵}$$

1960s, Kalman

考虑线性定常系统:

$$\begin{cases} X_{k+1} = AX_k + Bu_k + w_k \\ Y_{k+1} = CX_{k+1} + n_{k+1} \end{cases}$$

$$k = 0, 1, 2, \dots$$

$$\begin{cases} \hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1}(Y_{k+1} - C\bar{X}_{k+1}) \\ \bar{X}_{k+1} = A\hat{X}_k + Bu_k \\ K_{k+1} = (\bar{P}_{k+1}C^T)(C\bar{P}_{k+1}C^T + R)^{-1} \\ P_{k+1} = \bar{P}_{k+1} - K_{k+1}C\bar{P}_{k+1} \\ \bar{P}_{k+1} = AP_kA^T + Q \end{cases}$$

定理：卡尔曼滤波满足如下特性：

1. 线性最小方差估计

$$2. P_k = \mathbb{E} \left\{ (X_k - \hat{X}_k)(X_k - \hat{X}_k)^T \right\}$$

3. 当(A,C)可观时，滤波误差均方有界

引理1. 设 x 和 y 分别是具有前二阶矩的随机向量,
则利用 y 对 x 的线性无偏最小方差估计为

$$\hat{x}_y = \mathbb{E}\{x\} + R_{xy}R_y^{-1}(y - \mathbb{E}\{y\}),$$

其中

$$R_{xy} \triangleq \mathbb{E}\{(x - \mathbb{E}\{x\})(y - \mathbb{E}\{y\})^T\},$$

$$R_y \triangleq \mathbb{E}\{(y - \mathbb{E}\{y\})(y - \mathbb{E}\{y\})^T\}.$$

引理1的证明： 由估计的线性无偏性可知 \hat{x}_y 具有如下形式：

$$\hat{x}_y = \mathbb{E}\{x\} + K(y - \mathbb{E}\{y\}).$$

由此易得

$$\begin{aligned} & \mathbb{E}\left\{(\hat{x}_y - x)(\hat{x}_y - x)^T\right\} \\ &= \mathbb{E}\{(x - \mathbb{E}\{x\} - K(y - \mathbb{E}\{y\}))(x - \mathbb{E}\{x\} - K(y - \mathbb{E}\{y\}))^T\} \\ &= R_x - KR_{yx} - R_{xy}K^T + KR_yK^T \\ &= (K - R_{xy}R_y^{-1})R_y(K - R_{xy}R_y^{-1})^T + R_x - R_{xy}R_y^{-1}R_{yx} \end{aligned}$$

其中最后一个等式的第1项为非负定项，故欲令估计误差的协方差阵最小，当且仅当 $K = R_{xy}R_y^{-1}$.

引理2. 设 u 为已知向量, x 和 y 是具有前二阶矩的随机向量, 记 y 对 x 的线性无偏最小方差估计为 \hat{x}_y , 则利用 y 对 $z \triangleq Ax + Bu + Cy$ 的线性无偏最小方差估计为 $\hat{z}_y = A\hat{x}_y + Bu + Cy$.

证明过程

引理3. 设 x , y 和 y 是具有前二阶矩的随机向量, 记 $w \triangleq [y^T \quad y^T]^T$, 若利用 y 对 x 和 y 的估计分别为 \hat{x}_y 和 \hat{y}_y , 则利用 w 对 x 的估计满足

$$\hat{x}_w = \hat{x}_y + K(y - \hat{y}_y),$$

其中

$$K = R_{xy|y} R_{y|y}^{-1},$$

$$R_{xy|y} \triangleq \mathbb{E}\{(x - \hat{x}_y)(y - \hat{y}_y)^T\},$$

$$R_{y|y} \triangleq \mathbb{E}\{(y - \hat{y}_y)(y - \hat{y}_y)^T\}.$$

引理3的证明: 记 $e = x - \hat{x}_y - K(y - \hat{y}_y)$, 则由引理2 知 $\hat{e}_w = \hat{x}_w - \hat{x}_y - K(y - \hat{y}_y)$, 从而只需验证 $\hat{e}_w = 0$. 根据引理1易得

$$\mathbb{E} \left\{ (x - \hat{x}_y)(y - \mathbb{E}\{y\})^T \right\} = \mathbf{0},$$

$$\mathbb{E} \left\{ (y - \hat{y}_y)(y - \mathbb{E}\{y\})^T \right\} = \mathbf{0}.$$

从而 $\mathbb{E}\{e(y - \mathbb{E}\{y\})^T\} = \mathbf{0}$. 进而有

$$\mathbb{E}\{e(y - \mathbb{E}\{y\})^T\} = \mathbb{E} \left\{ e(y - \hat{y}_y)^T \right\} = \mathbf{0}.$$

注意到 $\mathbb{E}\{e\} = \mathbf{0}$, 从而 $R_{ew} = \mathbb{E}\{e(w - \mathbb{E}\{w\})^T\} = \mathbf{0}$. 则根据引理1知 $\hat{e}_w = 0$.

定理的证明

记 $\mathbf{y}_k \triangleq [Y_0^T \ Y_1^T \ \cdots \ Y_k^T]^T$ 对 X_k, X_{k+1} 和 Y_{k+1} 的线性无偏最小方差估计分别为 \hat{X}_k, \bar{X}_{k+1} 和 \bar{Y}_{k+1} , 则根据引理2可得 $\bar{X}_{k+1} = A\hat{X}_k + Bu_k, \bar{Y}_{k+1} = C\bar{X}_{k+1}$. 同时记 \hat{X}_k 和 \bar{X}_{k+1} 的估计误差协方差阵为 P_k 和 \bar{P}_{k+1} , 则 $\bar{P}_{k+1} = AP_kA^T + Q$. 进一步地, 将引理3中的 x, \mathbf{y} 和 y 分别替换成 X_{k+1}, \mathbf{y}_k 和 Y_{k+1} , 可得 $\hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1}(Y_{k+1} - C\bar{X}_{k+1})$, 且

$$\begin{aligned} K_{k+1} &= \mathbb{E}\{(X_{k+1} - \bar{X}_{k+1})(Y_{k+1} - \bar{Y}_{k+1})^T\}(\mathbb{E}\{(Y_{k+1} - \bar{Y}_{k+1})(Y_{k+1} - \bar{Y}_{k+1})^T\})^{-1} \\ &= \bar{P}_{k+1} C^T (C\bar{P}_{k+1}C^T + R)^{-1}, \end{aligned}$$

$$P_{k+1} = (\mathbf{I} - K_{k+1}C)\bar{P}_{k+1}(\mathbf{I} - K_{k+1}C)^T + K_{k+1}RK_{k+1} = \bar{P}_{k+1} - K_{k+1}C\bar{P}_{k+1}.$$

记 $Y_{k(n)} \triangleq [Y_{k-n+1}^T \quad Y_{k-n}^T \quad \cdots \quad Y_k^T]^T$, $M \triangleq [C^T \quad (CA)^T \quad \cdots \quad (CA^{n-1})^T]^T$, 由于 (A, C) 能观, 亦即 M 是列满秩的, 从而可以构造如下估计

$$\tilde{X}_k = \begin{cases} A^k \mathbb{E}\{x_0\}, & k < n \\ A^{n-1}(M^T M)^{-1} M^T Y_{k(n)}, & k \geq n \end{cases}$$

易得 \tilde{X}_k 的估计误差协方差阵 \tilde{P}_k 是有界的。由于 \tilde{P}_k 是最小方差的线性估计, 从而 $P_k \leq \tilde{P}_k$ 也是有界的。

概要

- 研究背景和问题描述
- 卡尔曼滤波算法及其主要特性
- **扩展卡尔曼滤波算法**

非线性系统的滤波问题:

$$\begin{cases} X_{k+1} = f(X_k) + w_k \\ Y_{k+1} = g(X_{k+1}) + v_{k+1} \end{cases}$$

$k = 0, 1, 2, \dots$

X_k : state, Y_k : measured output,

$(f(\cdot), g(\cdot))$: known C^1 functions

$(\{n_k\}, \{w_k\})$: uncorrelated zero-mean stochastic noise

Objective: linear minimum variance estimation, i.e.,

$$\hat{X}_k = \arg \min_{Z_k \in \{\text{linear function of } Y_i, 1 \leq i \leq k\}} E \left\{ (Z_k - X_k)(Z_k - X_k)^T \right\}$$

对付非线性的主要方法：线性化

线性化
linearization

+

Kalman Filter, Kalman.
H-infinity Filter, Banavar, etc.
Set-valued Filter, Zhu, etc.

...

线性系统的卡尔曼滤波

$$\begin{cases} X_{k+1} = AX_k + BF(X_k, k) + w_k \\ Y_{k+1} = CX_{k+1} + n_{k+1} \end{cases}$$

$$k = 0, 1, 2, \dots$$

X_k : state, Y_k : measured output

(A, B, C) : known matrices

$F(X_k, k)$: nonlinear dynamics

卡尔曼
滤波算法

$$\begin{cases} \hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1}(Y_{k+1} - C\bar{X}_{k+1}) \\ \bar{X}_{k+1} = A\hat{X}_k + Bu_k \\ K_{k+1} = (\bar{P}_{k+1}C^T)(C\bar{P}_{k+1}C^T + R)^{-1} \\ P_{k+1} = \bar{P}_{k+1} - K_{k+1}C\bar{P}_{k+1} \\ \bar{P}_{k+1} = AP_kA^T + Q \end{cases}$$

Parameters design:

$$\begin{cases} Q_k = E(w_k w_k^T) \\ R_{k+1} = E(n_{k+1} n_{k+1}^T) > 0 \\ P_0 = E(X_0 - EX_0)(X_0 - EX_0)^T \\ \bar{X}_0 = EX_0 \end{cases}$$

linear model

Require the exact information of : noise's model

initial value's model

扩展卡尔曼滤波

$$\begin{cases} X_{k+1} = f(X_k) + w_k \\ Y_{k+1} = g(X_{k+1}) + v_{k+1} \end{cases}$$

$k = 0, 1, 2, \dots$

X_k : state, Y_k : measured output,

$(f(\cdot), g(\cdot))$: known C^1 functions

$(\{v_k\}, \{w_k\})$: uncorrelated zero-mean

$$A_k = \left. \frac{\partial f}{\partial X} \right|_{\hat{X}_k},$$

第k+1步:
在滤波值点的线性部分;

$$C_{k+1} = \left. \frac{\partial g}{\partial X} \right|_{\bar{X}_{k+1}}$$

第k+1步:
在预测值点的线性部分

扩展卡尔曼滤波

$$\begin{cases} X_{k+1} = f(X_k) + w_k \\ Y_{k+1} = g(X_{k+1}) + v_{k+1} \end{cases}$$

$$k = 0, 1, 2, \dots$$

X_k : state, Y_k : measured output,
 $(f(\cdot), g(\cdot))$: known C^1 functions
 $(\{v_k\}, \{w_k\})$: uncorrelated zero-mean

$$\begin{cases} \hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1} (Y_{k+1} - g(\bar{X}_{k+1})) \\ \bar{X}_{k+1} = F(\hat{X}_k, k) \\ K_{k+1} = (\bar{P}_{k+1} \bar{C}_{k+1}^T) (\bar{C}_{k+1} \bar{P}_{k+1} \bar{C}_{k+1}^T + R_{k+1})^{-1} \\ P_{k+1} = \bar{P}_{k+1} - K_{k+1} \bar{C}_{k+1} \bar{P}_{k+1} \\ \bar{P}_{k+1} = \bar{A}_k P_k \bar{A}_k^T + Q_k \end{cases}$$

$$A_k = \left. \frac{\partial f}{\partial X} \right|_{\hat{X}_k}, C_{k+1} = \left. \frac{\partial g}{\partial X} \right|_{\bar{X}_{k+1}}$$

Parameters design:

$$\begin{cases} Q_k = E(w_k w_k^T) \\ R_{k+1} = E(n_{k+1} n_{k+1}^T) > 0 \\ P_0 = E(X_0 - EX_0)(X_0 - EX_0)^T \\ \bar{X}_0 = EX_0 \end{cases}$$

Utilizing the linear part at the point of estimation value

$$\begin{cases} X_{k+1} = f(X_k) + w_k \\ Y_{k+1} = g(X_{k+1}) + v_{k+1} \end{cases}$$

$$k = 0, 1, 2, \dots$$

Reif K., et al., Stochastic stability of the discrete-time extended Kalman filter, IEEE TAC, 1999

$$\begin{cases} \hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1} (Y_{k+1} - g(\bar{X}_{k+1})) \\ \bar{X}_{k+1} = F(\hat{X}_k, k) \\ K_{k+1} = (\bar{P}_{k+1} \bar{C}_{k+1}^T) (\bar{C}_{k+1} \bar{P}_{k+1} \bar{C}_{k+1}^T + R_{k+1})^{-1} \\ P_{k+1} = \bar{P}_{k+1} - K_{k+1} \bar{C}_{k+1} \bar{P}_{k+1} \\ \bar{P}_{k+1} = \bar{A}_k P_k \bar{A}_k^T + Q_k, \quad A_k = \left. \frac{\partial f}{\partial X} \right|_{\hat{X}_k}, C_{k+1} = \left. \frac{\partial g}{\partial X} \right|_{\bar{X}_{k+1}} \end{cases}$$

EKF算法的主要理论结果

- Conditions:**
- Exact information of nonlinear model
 - Sufficiently small noise and initial error
 - Uniformly observable condition

稳定性: $\sup_k \left\| E \left\{ \left(\hat{X}_k - X_k \right) \left(\hat{X}_k - X_k \right)^T \right\} \right\| < \infty$

最优性?

一致性?

谢谢！