



次优及鲁棒滤波算法

蔡远利 教授
西安交通大学自动化学院

0. Outline

- 1 衰减记忆滤波 / 3
- 2 限定记忆滤波 / 14
- 3 协方差平方根滤波 / 23
- 4 自适应滤波 / 33
- 5 常值增益次优滤波 / 65
- 6 鲁棒 H_∞ 滤波 / 69

- ♠ 标准卡尔曼滤波算法只有在系统数学模型和噪声统计描述与实际情况一致的情况下，同时计算精度足够时，最优性才有保证。否则，可能出现**滤波发散**现象。
- ♠ 对于一些复杂的高阶系统，卡尔曼滤波算法每一步所需的计算时间可能太长，从而超过容许的采样时间，将无法实现有效的**实时控制**。
- ♠ 为此，我们需要改良卡尔曼滤波算法，希望解决上述问题的同时，使滤波性能损失不太大，即增加滤波算法的鲁棒性。

1. 衰减记忆滤波

在递推估计系统状态的过程中，如果逐渐减少历史量测数据的影响、相对增加新量测数据的影响，可以抑制舍入误差的积累和传播，从而达到克服滤波发散的目的。

1.1 指数衰减记忆滤波

考虑离散时间随机系统

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + G_k \mathbf{w}_k \quad (1)$$

$$\mathbf{y}_{k+1} = H_{k+1} \mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (2)$$

其中, $\mathbf{w}_k \sim (0, Q_k)$ 与 $\mathbf{v}_k \sim (0, R_k)$ 相互独立, 它们与初始状态 $\mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)$ 不相关。另外, 设 $Q_k \geq 0, R_k > 0$ 。

最优滤波可以表达为

$$\hat{\mathbf{x}}_{k+1|k+1} = [I - K_{k+1}H_{k+1}]F_k\hat{\mathbf{x}}_{k|k} + K_{k+1}\mathbf{y}_{k+1} \quad (3)$$

可见:

- (1) 任意时刻的滤波值 $\hat{\mathbf{x}}_{k+1|k+1}$ 是量测数据 $\mathbf{y}_1 \sim \mathbf{y}_{k+1}$ 及初始估计 $\hat{\mathbf{x}}_{0|0}$ 的线性组合。
- (2) 所有的量测噪声的协方差矩阵 $R_1 \sim R_{k+1}$ 和初始估计误差协方差矩阵 $P_{0|0}$ 都进入了 $\hat{\mathbf{x}}_{k+1|k+1}$ 的计算。
- (3) 设当前时刻为 N , 如果要降低 \mathbf{y}_k 及 $\hat{\mathbf{x}}_{0|0}$ 对 $\hat{\mathbf{x}}_{N|N}$ 的影响, 可以通过增大 R_k 及 $P_{0|0}$ 的值来实现。

最优滤波也可以表达为

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + K_{k+1}\tilde{\mathbf{y}}_{k+1} \quad (4)$$

表明:

- ♠ 一步预测和量测信息都直接与系统状态模型 (1) 有关;
- ♠ 为了抑制远离当前时刻 N 模型不确定性对 $\hat{\mathbf{x}}_{N|N}$ 的影响, 可以通过增大 Q_k 的值来实现。

注意到最优增益计算公式为

$$K_{k+1} = P_{k+1|k} H_{k+1}^T [H_{k+1} P_{k+1|k} H_{k+1}^T + R_{k+1}]^{-1} \quad (5)$$

或

$$K_{k+1} = P_{k+1|k+1} H_{k+1}^T R_{k+1}^{-1} \quad (6)$$

如果将时刻 $k+1$ 的量测噪声协方差增加为

$$R_{k+1} \exp\left[\sum_{i=k+1}^N c_i\right], c_i \geq 0 \quad (7)$$

上式表明根据距离当前时刻 N 的远近加不同的权重，即离当前时刻 N 越远加越大的权重。记

$$P_{k+1|k+1}^* = P_{k+1|k+1} \exp\left[-\sum_{i=k+1}^N c_i\right] \quad (8)$$

$$P_{k+1|k}^* = P_{k+1|k} \exp\left[-\sum_{i=k+1}^N c_i\right] \quad (9)$$

那么 (5)、(6) 变为

$$\begin{aligned} K_{k+1}^* &= P_{k+1|k}^* H_{k+1}^T [H_{k+1} P_{k+1|k}^* H_{k+1}^T + R_{k+1}]^{-1} \\ &= P_{k+1|k+1}^* H_{k+1}^T R_{k+1}^{-1} \end{aligned}$$

考虑一步预测的误差协方差矩阵

$$P_{k+1|k} = F_k P_{k|k} F_k^T + G_k Q_k G_k^T \quad (10)$$

如果将过程噪声协方差增加为

$$Q_k \exp\left[\sum_{i=k}^N c_i\right], c_i \geq 0 \quad (11)$$

那么可得

$$P_{k+1|k}^* = [F_k P_{k|k}^* F_k^T + G_k Q_k G_k^T] \exp(c_k) \quad (12)$$

如果将 $P_{0|0}$ 增加为

$$P_{0|0} \exp\left[\sum_{i=0}^N c_i\right], c_i \geq 0 \quad (13)$$

于是有

$$P_{0|0}^* = P_{0|0} \quad (14)$$

这样我们便建立了统计特性修改后的滤波算法，如表1所示。上标“*”代表次优滤波参数。

由表1不难发现，指数衰减记忆滤波与标准卡尔曼滤波的区别仅在于一步预测误差的协方差矩阵，即式 (12)。

Table 1: 指数衰减记忆滤波算法

状态方程与量测方程

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + G_k \mathbf{w}_k$$

$$\mathbf{y}_{k+1} = H_{k+1} \mathbf{x}_{k+1} + \mathbf{v}_{k+1}$$

滤波初值

$$\hat{\mathbf{x}}_{0|0}^* = E\mathbf{x}_0 = \bar{\mathbf{x}}_0, \quad P_{0|0}^* = \text{var}[\mathbf{x}_0] = P_0$$

一步预测

$$\hat{\mathbf{x}}_{k+1|k}^* = F_k \hat{\mathbf{x}}_{k|k}^*$$

$$P_{k+1|k}^* = [F_k P_{k|k}^* F_k^T + G_k Q_k G_k^T] \exp(c_k)$$

滤波增益

$$\begin{aligned} K_{k+1}^* &= P_{k+1|k}^* H_{k+1}^T [H_{k+1} P_{k+1|k}^* H_{k+1}^T + R_{k+1}]^{-1} \\ &= P_{k+1|k+1}^* H_{k+1}^T R_{k+1}^{-1} \end{aligned}$$

滤波计算

$$\hat{\mathbf{x}}_{k+1|k+1}^* = \hat{\mathbf{x}}_{k+1|k}^* + K_{k+1}^* [\mathbf{y}_{k+1} - H_{k+1} \hat{\mathbf{x}}_{k+1|k}^*]$$

$$P_{k+1|k+1}^* = [I - K_{k+1}^* H_{k+1}] P_{k+1|k}^*$$

1.2 几何级数衰减记忆滤波

在指数衰减记忆滤波算法中，如果取 $c_0 = c_1 = \cdots = c_N = c \geq 0$ ，并记 $s = \exp(c)$ ，我们便得到几何级数衰减滤波算法。一步预测误差的协方差矩阵计算公式变为

$$P_{k+1|k}^* = [F_k P_{k|k}^* F_k^T + G_k Q_k G_k^T] s \quad (15)$$

表1中其他公式不变。

在任意的时刻 N ，几何级数衰减滤波相当于将量测噪声协方差矩阵调整为

$$s^{N-k} R_k, k = 1, 2, \cdots, N \quad (16)$$

将过程噪声协方差矩阵调整为

$$s^{N-k}Q_k, k = 0, 1, \dots, N - 1 \quad (17)$$

而初始估计的误差协方差调整为

$$s^N P_{0|0} \quad (18)$$

显然，几何级数衰减滤波算法是指数衰减记忆滤波算法的一种特例。关于衰减记忆滤波详细论述见 Sorenson 等经典文献。

2. 限定记忆滤波

- ♠ 标准卡尔曼滤波器对量测数据的记忆是无限增长的，即计算 $\hat{\mathbf{x}}_{k|k}$ 时用到了所有过去的量测值。
- ♠ 限定记忆滤波计算 $\hat{\mathbf{x}}_{k|k}$ 时，只用到离时刻 k 最近的 N 个量测值 $\{\mathbf{y}_{k-N+1}, \mathbf{y}_{k-N+2}, \dots, \mathbf{y}_k\}$ ，完全截断了时刻 $k - N + 1$ 以前量测数据对滤波值的影响。
- ♠ 和衰减记忆滤波一样，限定记忆滤波也是非常实用的一种技术，有时也称为加窗、滑动窗口滤波技术。

2.1 量测数据分组

$$\mathbf{y}^k = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_d, \underbrace{\mathbf{y}_{d+1}, \dots, \mathbf{y}_{k-1}, \mathbf{y}_k}_{\mathbf{y}_d^k}\} \quad (19)$$

$$\mathbf{y}_d^{k-1} = \{\mathbf{y}_d, \dots, \mathbf{y}_{k-2}, \mathbf{y}_{k-1}\} \quad (20)$$

注意, \mathbf{y}^k 表示所有到时刻 k 的量测值; \mathbf{y}_{d+1}^k 表示时刻 $d+1$ 到时刻 k 的 N 个量测值; \mathbf{y}_d^{k-1} 表示时刻 d 到时刻 $k-1$ 的 N 个量测值; \mathbf{y}_d^k 表示时刻 d 到时刻 k 的 $N+1$ 个量测值。

我们的目标是获得 \mathbf{x}_k 基于 \mathbf{y}_{d+1}^k 的最优估计 (线性最小方差估计), 并建立递推算法。

2.2 递推算法

为描述方便，研究如下不含过程噪声的随机动态系统

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k \quad (21)$$

$$\mathbf{y}_{k+1} = H_{k+1} \mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (22)$$

其中， $\mathbf{v}_k \sim (0, R_k)$ 是与初始状态 $\mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)$ 不相关的噪声， $R_k > 0$ 。另外，我们也会用到 $F_{k+1,k}$ 来表示 F_k 。

设基于 \mathbf{y}_d^{k-1} 对 \mathbf{x}_k 的线性最小方差估计为 $\hat{\mathbf{x}}_{k|k-1}^N$ ，对 \mathbf{x}_{k-1} 的线性最小方差估计为 $\hat{\mathbf{x}}_{k-1|k-1}^N$ 。那么

$$\hat{\mathbf{x}}_{k|k-1}^N = F_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^N \quad (23)$$

基于 \mathbf{y}_d^k 对 \mathbf{x}_k 的线性最小方差估计为

$$\hat{\mathbf{x}}_{k|k}^{N+1} = \hat{\mathbf{x}}_{k|k-1}^N + J_k [\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}^N] \quad (24)$$

其中

$$J_k = P_{k|k-1}^N H_k^T [H_k P_{k|k-1}^N H_k^T + R_k]^{-1} \quad (25)$$

$$P_{k|k-1}^N = F_{k-1} P_{k-1|k-1}^N F_{k-1}^T \quad (26)$$

$$P_{k|k}^{N+1} = [I - J_k H_k] P_{k|k-1}^N = [(P_{k|k-1}^N)^{-1} + H_k^T R_k^{-1} H_k]^{-1} \quad (27)$$

式 (24)~ 式 (27) 实际上完全套用了标准卡尔曼滤波公式。

又设基于 \mathbf{y}_{d+1}^k 对 \mathbf{x}_k 的线性最小方差估计为 $\hat{\mathbf{x}}_{k|k}^N$, 基于 \mathbf{y}_d^k 对 \mathbf{x}_k 的线性最小方差估计为 $\hat{\mathbf{x}}_{k|k}^{N+1}$ 。显然, $\hat{\mathbf{x}}_{k|k}^{N+1}$ 比 $\hat{\mathbf{x}}_{k|k}^N$ 多用了—个量测数据 \mathbf{y}_d 。

由 (21) 和 (22) 可知

$$\mathbf{y}_d = H_d \mathbf{x}_d + \mathbf{v}_d = H_d F_{d,k} \mathbf{x}_k + \mathbf{v}_d \quad (28)$$

其中 $F_{d,k}$ 参照 $F_{k+1,k} = F_k$ 进行定义, 因此

$$\hat{\mathbf{x}}_{k|k}^{N+1} = \hat{\mathbf{x}}_{k|k}^N + \bar{J}_k [\mathbf{y}_d - H_d F_{d,k} \hat{\mathbf{x}}_{k|k}^N] \quad (29)$$

其中

$$\bar{J}_k = P_{k|k}^N F_{d,k}^T H_d^T [F_{d,k} H_d P_{k|k}^N H_d^T F_{d,k}^T + R_d]^{-1}$$

$$P_{k|k}^N = E \tilde{\mathbf{x}}_{k|k}^N (\tilde{\mathbf{x}}_{k|k}^N)^T$$

$$P_{k|k}^{N+1} = [I - \bar{J}_k H_d F_{d,k}] P_{k|k}^N = [(P_{k|k}^N)^{-1} + F_{d,k}^T H_d^T R_d^{-1} H_d F_{d,k}]^{-1}$$

由 (24)、(29) 可知

$$\begin{aligned}\hat{\mathbf{x}}_{k|k}^N - \hat{\mathbf{x}}_{k|k-1}^N &= J_k[\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}^N] - \bar{J}_k[\mathbf{y}_d - H_d F_{d,k} \hat{\mathbf{x}}_{k|k}^N] \\ &= J_k[\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}^N] - \bar{J}_k[\mathbf{y}_d - H_d F_{d,k} \hat{\mathbf{x}}_{k|k-1}^N] \\ &\quad + \bar{J}_k H_d F_{d,k} [\hat{\mathbf{x}}_{k|k}^N - \hat{\mathbf{x}}_{k|k-1}^N]\end{aligned}$$

即

$$\begin{aligned}\hat{\mathbf{x}}_{k|k}^N - \hat{\mathbf{x}}_{k|k-1}^N &= [I - \bar{J}_k H_d F_{d,k}]^{-1} J_k [\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}^N] \\ &\quad - [I - \bar{J}_k H_d F_{d,k}]^{-1} \bar{J}_k [\mathbf{y}_d - H_d F_{d,k} \hat{\mathbf{x}}_{k|k-1}^N]\end{aligned}$$

令

$$K_k = [I - \bar{J}_k H_d F_{d,k}]^{-1} J_k \quad (30)$$

$$\bar{K}_k = [I - \bar{J}_k H_d F_{d,k}]^{-1} \bar{J}_k \quad (31)$$

那么

$$\hat{\mathbf{x}}_{k|k}^N - \hat{\mathbf{x}}_{k|k-1}^N = K_k[\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}^N] - \bar{K}_k[\mathbf{y}_d - H_d F_{d,k} \hat{\mathbf{x}}_{k|k-1}^N]$$

注意到 (23), 最后可得

$$\begin{aligned} \hat{\mathbf{x}}_{k|k}^N &= F_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^N + K_k[\mathbf{y}_k - H_k F_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^N] \\ &\quad - \bar{K}_k[\mathbf{y}_d - H_d F_{d,k} F_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^N] \end{aligned} \quad (32)$$

此外, 整理可得

$$K_k = P_{k|k}^N H_k^T R_k^{-1}, \quad \bar{K}_k = P_{k|k}^N F_{d,k}^T H_d^T R_d^{-1} \quad (33)$$

还有

$$(P_{k|k}^N)^{-1} = F_{k-1}^T (P_{k-1|k-1}^N)^{-1} F_{k-1} + H_k^T R_k^{-1} H_k - F_{d,k}^T H_d^T R_d^{-1} H_d F_{d,k} \quad (34)$$

公式 (32)~ (34) 构成了一套限定记忆滤波递推计算算法, 适用于 $k > N$ 。但滤波初值还需要进一步研究。

2.3 初始条件

- 当 $k \leq N$ 时, 量测数据长度小于记忆长度 N , 尚不能进行限定记忆滤波计算, 只能采用常规卡尔曼滤波算法。
- 从 $k = N + 1$ 以后, 不能将 $\hat{\mathbf{x}}_{N|N}$ 和 $P_{N|N}$ 作为限定记忆滤波的初值 $\hat{\mathbf{x}}_{N|N}^N$ 和 $P_{N|N}^N$, 否则 $\hat{\mathbf{x}}_{0|0} = \bar{\mathbf{x}}_0, P_{0|0} = \text{var}[\mathbf{x}_0]$ 将一直影响到后续的滤波计算。
- 对于时刻 N 的状态 \mathbf{x}_N , 除了从常规卡尔曼滤波算法得到 $\hat{\mathbf{x}}_{N|N}$ 和 $P_{N|N}$ 外, 还可以建立两种估计:

- 用 $\{\mathbf{y}_1, \dots, \mathbf{y}_N\}$ 得到的 $\hat{\mathbf{x}}_{N|N}^N$ 和 $P_{N|N}^N$;
- 基于 \mathbf{x}_0 先验信息的估计, 实际上是 N 步预测。

基于融合估计原理, 我们有

$$\hat{\mathbf{x}}_{N|N} = P_{N|N}[(P_{N|N}^N)^{-1}\hat{\mathbf{x}}_{N|N}^N + F_{0,N}^T P_{0|0}^{-1}\hat{\mathbf{x}}_{0|0}]$$

$$P_{N|N}^{-1} = (P_{N|N}^N)^{-1} + F_{0,N}^T P_{0|0}^{-1} F_{0,N}$$

由此我们便可建立限定记忆滤波递推计算初值计算公式如下

$$\hat{\mathbf{x}}_{N|N}^N = P_{N|N}^N [P_{N|N}^{-1}\hat{\mathbf{x}}_{N|N} - F_{0,N}^T P_{0|0}^{-1}\hat{\mathbf{x}}_{0|0}] \quad (35)$$

$$P_{N|N}^N = [P_{N|N}^{-1} - F_{0,N}^T P_{0|0}^{-1} F_{0,N}]^{-1} \quad (36)$$

限定记忆滤波技术的灵活应用, 可以较好地解决许多实际问题, 是一种非常经典的实用技术。

3. 协方差平方根滤波

- ♠ 除了模型不准确会引起滤波发散外，计算舍入误差的积累也会引发滤波发散。
- ♠ 计算精度产生发散的主要原因是 $P_{k|k}$ 、 $P_{k|k-1}$ 在计算过程中丧失了应该具有的对称性、非负定性。
- ♠ 通过引入矩阵平方根的概念，可以避免递推计算直接处理 $P_{k|k}$ 、 $P_{k|k-1}$ ，从而增加算法的鲁棒性。

3.1 矩阵的下三角分解

对于任意的对称非负定矩阵 P ，均可分解为

$$P = SS^T \quad (37)$$

其中 S 是一个下三角矩阵，称为 P 矩阵的平方根矩阵。如果 P 矩阵是正定的，那么 S 矩阵还将是非奇异的。

MATLAB 内嵌函数 chol 可以完成矩阵的下三角分解, 例如 $S = \text{chol}(P)'$ 。三阶矩阵平方根分解算法为

$$\begin{aligned} \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{21} & p_{22} & p_{32} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} &= \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix}^T \\ &= \begin{bmatrix} c_{11}^2 & c_{11}c_{21} & c_{11}c_{31} \\ c_{11}c_{21} & c_{21}^2 + c_{22}^2 & c_{21}c_{31} + c_{22}c_{32} \\ c_{11}c_{31} & c_{21}c_{31} + c_{22}c_{32} & c_{31}^2 + c_{32}^2 + c_{33}^2 \end{bmatrix} \end{aligned}$$

一般地, 设

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}, \quad S = \begin{bmatrix} s_{11} & 0 & \cdots & 0 \\ \vdots & s_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ s_{n1} & \cdots & \cdots & s_{nn} \end{bmatrix}$$

由 (37) 可得

$$s_{ii} = \sqrt{p_{ii} - \sum_{j=1}^{i-1} s_{ij}^2} \quad (38)$$

$$s_{ij} = \begin{cases} 0, & i < j \\ \frac{1}{s_{jj}}(p_{ij} - \sum_{k=1}^{j-1} s_{ik}s_{jk}), & i > j \end{cases} \quad (39)$$

3.2 平方根滤波基本思想

考虑无过程噪声、标量量测情况下的系统

$$\mathbf{x}_k = F_{k-1}\mathbf{x}_{k-1} \quad (40)$$

$$y_k = H_k\mathbf{x}_k + v_k \quad (41)$$

其中, $v_k \sim (0, R_k)$ 是与初始状态 $\mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)$ 不相关的噪声, $R_k > 0$ 。该系统的卡尔曼滤波基本方程为

$$\hat{\mathbf{x}}_{k|k} = F_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + K_k[y_k - H_kF_{k-1}\hat{\mathbf{x}}_{k-1|k-1}] \quad (42)$$

$$K_k = P_{k|k-1}H_k^T[H_kP_{k|k-1}H_k^T + R_k]^{-1} \quad (43)$$

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T \quad (44)$$

$$P_{k|k} = [I - K_kH_k]P_{k|k-1} \quad (45)$$

设

$$P_{k|k} = S_{k|k} S_{k|k}^T \quad (46)$$

由 (44) 可得

$$P_{k|k-1} = S_{k|k-1} S_{k|k-1}^T \quad (47)$$

其中 $S_{k|k-1} = F_{k-1} S_{k-1|k-1}$ 。而由 (43)、(45) 可得

$$P_{k|k} = S_{k|k-1} \{I - \mathcal{F}_k [\mathcal{F}_k^T \mathcal{F}_k + R_k]^{-1} \mathcal{F}_k^T\} S_{k|k-1}^T \quad (48)$$

其中 $\mathcal{F}_k = S_{k|k-1}^T H_k^T$ 。

考虑到

$$\alpha_k = [\mathcal{F}_k^T \mathcal{F}_k + R_k]^{-1} \quad (49)$$

是一标量, (48) 即为

$$P_{k|k} = S_{k|k}[I - \alpha_k \mathcal{F}_k \mathcal{F}_k^T] S_{k|k-1}^T \quad (50)$$

如果令

$$I - \alpha_k \mathcal{F}_k \mathcal{F}_k^T = [I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T][I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T]^T \quad (51)$$

那么

$$2r_k - r_k^2 \alpha_k \mathcal{F}_k \mathcal{F}_k^T = 1 \quad (52)$$

计及 (49), 上式即为

$$2r_k - r_k^2[1 - \alpha_k R_k] = 1 \quad (53)$$

由此可解出

$$r_k = \frac{1}{1 \pm \sqrt{\alpha_k R_k}} \quad (54)$$

(48) 于是可化为

$$P_{k|k} = S_{k|k-1} [I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T] [I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T]^T S_{k|k-1}^T \quad (55)$$

考虑到 (46), 我们得到

$$S_{k|k} = S_{k|k-1} [I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T] \quad (56)$$

另外, 注意到

$$[H_k P_{k|k-1} H_k^T + R_k]^{-1} = [\mathcal{F}_k^T \mathcal{F}_k + R_k]^{-1} = \alpha_k \quad (57)$$

Table 2: 协方差平方根滤波算法

状态方程与量测方程

$$\begin{aligned}\mathbf{x}_k &= F_{k-1}\mathbf{x}_{k-1}, \quad v_k \sim (0, R_k) \\ y_k &= H_k\mathbf{x}_k + v_k, \quad \mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)\end{aligned}$$

滤波方程

$$\hat{\mathbf{x}}_{k|k} = F_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + K_k[y_k - H_k F_{k-1}\hat{\mathbf{x}}_{k-1|k-1}]$$

滤波增益

$$K_k = \alpha_k S_{k|k-1} \mathcal{F}_k$$

协方差平方根

$$\begin{aligned}S_{k|k} &= S_{k|k-1} [I - r_k \alpha_k \mathcal{F}_k \mathcal{F}_k^T] \\ S_{k|k-1} &= F_{k-1} S_{k-1|k-1}\end{aligned}$$

其他方程

$$\alpha_k = [\mathcal{F}_k^T \mathcal{F}_k + R_k]^{-1}, \mathcal{F}_k = S_{k|k-1}^T H_k^T, r_k = \frac{1}{1 \pm \sqrt{\alpha_k R_k}}$$

滤波增益 (43) 即为

$$K_k = \alpha_k P_{k|k-1} H_k^T = \alpha_k S_{k|k-1} \mathcal{F}_k \quad (58)$$

无过程噪声、标量量测情况下的协方差平方根滤波算法汇总于表2中。关于平方根协方差滤波的进一步分析和讨论，参加 Kaminski、Carlo 及 Chandra 等人工作。

4. 自适应滤波

- 卡尔曼滤波器的设计模型或噪声统计特性参数与实际情况不一致时，滤波性能将大打折扣，甚至出现滤波发散现象。
- 我们希望利用量测数据进行滤波的同时，能够基于设计的滤波器输出，不断地估计和修正模型中不精确的参数和噪声统计特性，这就是自适应滤波的基本思想。
- 自适应滤波是状态估计领域的热门主题。

4.1 噪声自适应滤波

我们首先讨论噪声统计特性参数不准确的情况，简要说明自适应滤波的原理。这里仅研究线性定常系统，即

$$\mathbf{x}_{k+1} = F\mathbf{x}_k + G\mathbf{w}_k \quad (59)$$

$$\mathbf{y}_{k+1} = H\mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (60)$$

其中， $\mathbf{w}_k \sim (0, Q)$ 与 $\mathbf{v}_k \sim (0, R)$ 是互不相关的白噪声，它们与初始状态 $\mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)$ 不相关。另外，设 $Q \geq 0, R > 0$ 。

该系统的卡尔曼滤波器为

$$\hat{\mathbf{x}}_{k+1|k+1} = F\hat{\mathbf{x}}_{k|k} + K_{k+1}\tilde{\mathbf{y}}_{k+1} \quad (61)$$

$$\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_{k+1|k} = \mathbf{y}_{k+1} - HF\hat{\mathbf{x}}_{k|k} \quad (62)$$

$$K_{k+1} = P_{k+1|k}H^T[HP_{k+1|k}H^T + R]^{-1} \quad (63)$$

$$P_{k+1|k} = FP_{k|k}F^T + GQG^T \quad (64)$$

$$P_{k+1|k+1} = [I - KH]P_{k+1|k} \quad (65)$$

根据最优估计无偏性和随机正交原理，有

$$E\tilde{\mathbf{y}}_k = 0, \quad \forall k > 0$$

$$E\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_j^T = 0, \quad \forall k \neq j, k, j > 0$$

以上两式表明, 新息序列 $\{\tilde{\mathbf{y}}_k, k > 0\}$ 是一均值为零的白噪声序列, 其协方差阵为

$$P_{\tilde{\mathbf{y}}_{k+1}} = E\tilde{\mathbf{y}}_{k+1}\tilde{\mathbf{y}}_{k+1}^T = HP_{k+1,k}H^T + R \quad (66)$$

如果 Q 与 R 不精确时, 按 (61)~(65) 计算出的 $\hat{\mathbf{x}}_{k|k}$ 将不是状态 \mathbf{x}_k 的最优估计, 新息序列 $\{\tilde{\mathbf{y}}_k, k > 0\}$ 也将不是白色的。后者是自适应滤波算法的基本出发点。

4.1.1 量测噪声 R 不精确情况

由于 R 阵不确定，需要不断进行估计，记时刻 k 的估计为 R_k 。如果时刻 k 及以前的 $\hat{\mathbf{x}}_{k|k}$ 接近最优，那么 $\{\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_k\}$ 将接近白噪声序列，其协方差矩阵的估计为

$$E\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_k^T = \frac{1}{k} \sum_{i=1}^k \tilde{\mathbf{y}}_i\tilde{\mathbf{y}}_i^T \quad (67)$$

结合到 (66)，我们有

$$R_k = \frac{1}{k} \sum_{i=1}^k \tilde{\mathbf{y}}_i\tilde{\mathbf{y}}_i^T - HP_{k+1|k}H^T \quad (68)$$

以此代入前面的卡尔曼滤波公式 (61)~(65), 便得到了此种情况下的自适应滤波算法

$$\hat{\mathbf{x}}_{k+1|k+1} = F\hat{\mathbf{x}}_{k|k} + K_{k+1}\tilde{\mathbf{y}}_{k+1} \quad (69)$$

$$\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_{k+1|k} = \mathbf{y}_{k+1} - HF\hat{\mathbf{x}}_{k|k} \quad (70)$$

$$K_{k+1} = P_{k+1|k}H^T(P_{\tilde{\mathbf{y}}_{k+1}})^{-1} \quad (71)$$

$$P_{k+1|k} = FP_{k|k}F^T + GQG^T \quad (72)$$

$$P_{\tilde{\mathbf{y}}_{k+1}} = \frac{k}{k+1}P_{\tilde{\mathbf{y}}_k} + \frac{1}{k+1}\tilde{\mathbf{y}}_{k+1}\tilde{\mathbf{y}}_{k+1}^T \quad (73)$$

$$P_{k+1|k+1} = [I - KH]P_{k+1|k} \quad (74)$$

上述算法与卡尔曼滤波基本公式的主要差异在于 (73), 其滤波初值为

$$\hat{\boldsymbol{x}}_{0|0} = E\boldsymbol{x}_0 = \bar{\boldsymbol{x}}_0 \quad (75)$$

$$P_{0|0} = \text{var}[\boldsymbol{x}_0] = P_0 \quad (76)$$

$P_{\tilde{\boldsymbol{y}}_0} = 0$ 或取定 $P_{\tilde{\boldsymbol{y}}_1}$ 。

4.1.2 过程噪声 Q 不精确情况

记过程噪声协方差阵的估计值为 Q_k ，类似地可以建立如下自适应滤波算法

$$\hat{\mathbf{x}}_{k+1|k+1} = F\hat{\mathbf{x}}_{k|k} + K_{k+1}\tilde{\mathbf{y}}_{k+1} \quad (77)$$

$$\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_{k+1|k} = \mathbf{y}_{k+1} - HF\hat{\mathbf{x}}_{k|k} \quad (78)$$

$$K_{k+1} = P_{k+1|k}H^T[HP_{k+1|k}H^T + R]^{-1} \quad (79)$$

$$P_{k+1|k} = FP_{k|k}F^T + GQ_kG^T \quad (80)$$

$$Q_k = G_A P_{\tilde{\mathbf{y}}_{k+1}} G_A^T - R^* - G_B P_{k|k} G_B^T \quad (81)$$

$$P_{\tilde{\mathbf{y}}_{k+1}} = \frac{k}{k+1} P_{\tilde{\mathbf{y}}_k} + \frac{1}{k+1} \tilde{\mathbf{y}}_{k+1} \tilde{\mathbf{y}}_{k+1}^T \quad (82)$$

$$P_{k+1|k+1} = [I - KH] P_{k+1|k} \quad (83)$$

其中 $G_A = [(HG)^T(HG)]^{-1}(HG)^T$, $G_B = G_A H F$, $R^* = G_A R G_A^T$ 。

滤波初值为 $\hat{\mathbf{x}}_{0|0} = E\mathbf{x}_0 = \bar{\mathbf{x}}_0$, $P_{0|0} = \text{var}[\mathbf{x}_0] = P_0$, $P_{\tilde{\mathbf{y}}}(0) = 0$ 。

4.1.3 Q 与 R 同时不精确情况

当过程噪声和量测噪声都不精确时，可以先选取一个适当的量测噪声协方差阵 R （通常这是比较容易做到的），并把它固定下来，然后按过程噪声都不精确的情况来设计自适应滤波器。

关于自适应滤波早期的思想及发展，见 William 等人文献。

4.2 多模型自适应滤波

对一个实际系统，假设我们并行地运行一族滤波器，每个滤波器的设计参数均不同，我们可以根据不同滤波器的残差来判断哪个滤波器更加适应当前的物理系统。进一步地，可以根据每个滤波器的输出获得系统当前最好的状态估计。这就是多模型滤波算法的基本思想或出发点，其工作原理如图1所示。

多模型方法已经获得了非常广泛的应用，包括目标跟踪、组合导航、故障诊断与识别、飞行器制导与控制、过程控制等等。

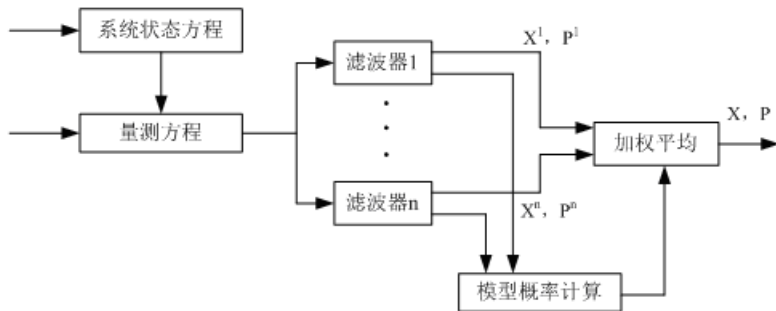


Figure 1: 多模型滤波原理示意图

4.2.1 多模型组合滤波公式

假设共有 r 个滤波器 $\{m_i\}_{i=1}^r$, 记 k 时刻第 i 个滤波器的输出为 $\hat{\mathbf{x}}_{k,m_i}$ 和 P_{k,m_i} , 那么系统当前时刻的组合估计为

$$\hat{\mathbf{x}}_k = \sum_{i=1}^r p_k(m_i) \hat{\mathbf{x}}_{k,m_i} \quad (84)$$

对应的估计误差协方差为

$$P_k = E(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T = \sum_{i=1}^r p_k(m_i) [P_{k,m_i} + (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k,m_i})(*)^T] \quad (85)$$

以上两式中, $p_k(m_i)$ 表示 k 时刻模型 (滤波器) m_i 的概率, 即

$$p_k(m_i) = \Pr(m = m_i | \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k) \quad (86)$$

4.2.2 静态多模型滤波

假设对象的真实模型是 r 个模型 $\{m_i\}_{i=1}^r$ 中的一个，不随时间变化，但事先并不知道具体是哪个。由此建立的多模型方法称为**静态多模型滤波方法**，这里的“静态”是指对象真实的模型是固定的，当然对象的状态仍然随时间变化，是动态的。

模型概率递推计算公式

设 k 时刻的量测为 \mathbf{y}_k ，那么模型 m_i 的似然函数为

$$\Lambda(\mathbf{y}_k | m_i) = \frac{1}{\sqrt{2\pi |C_{k,m_i}|}} e^{-\frac{1}{2}[\mathbf{y}_k - H_{k,m_i} \hat{\mathbf{x}}_{k,m_i}^-] C_{k,m_i}^{-1} [*]^T} \quad (87)$$

式中

$$C_{k,m_i} = H_{k,m_i} P_{k,m_i}^- H_{k,m_i}^T + R_{k,m_i} \quad (88)$$

根据全概率公式，可导出模型（滤波器） m_i 的概率递推公式如下：

$$p_k(m_i) = \frac{\Lambda(\mathbf{y}_k|m_i)p_{k-1}(m_i)}{\sum_{j=1}^r \Lambda(\mathbf{y}_k|m_j)p_{k-1}(m_j)} \quad (89)$$

注意到 $\Lambda(\mathbf{y}_k|m_i)$ 是一个标量，此外每一时刻有

$$\sum_{i=1}^r p_k(m_i) = 1 \quad (90)$$

初始 $k = 0$ 时，一般取

$$p_0(m_i) = \frac{1}{r}, \quad i = 1, 2, \dots, r \quad (91)$$

算法步骤

1. 根据当前时刻的量测数据 \mathbf{y}_k ，按 (87) 式计算每个模型 m_i 的似然函数；
2. 根据 (89) 式递推计算每个模型 m_i 的概率 $p_k(m_i)$ ；
3. 计算每个模型的后验估计 $\hat{\mathbf{x}}_{k,m_i}$ 及误差协方差 P_{k,m_i} ；
4. 根据 (84) 和 (85) 式计算系统状态的组合滤波 $\hat{\mathbf{x}}_k$ 和误差协方差 P_k 。

Remark 4.1 以上多模型滤波算法同样适用于非线性系统，对应地，将针对每个模型的滤波算法改为合适的非线性滤波算法即可，例如下一章将要介绍的 *EKF*、*UKF*、*CKF* 等。

模型估计的收敛性

上述算法在进行状态估计的同时，对系统的模型也同步进行了估计。已有研究表明，如果真实系统模型与模型族 $\{m_i\}_{i=1}^r$ 中某模型一致，那么上述多模型滤波算法将收敛到对象的真实模型。一般情况下，将收敛到一个加权平均表示的定常模型。

如果真实系统的模型在运行过程中是不断变化的，即可能从一个模型跳变到另外一个模型。此时，可以采用直观的处理方法避免收敛到单一模型。一些可能的途径包括：

- 人为地限制每个模型概率的下界；
- 在模型似然函数计算中采用衰减记忆技术；
- 在模型似然函数计算中采用有限记忆（滑动窗）技术。

Remark 4.2 采用这些直观处理方法可能带来一些负效应，例如不正确模型的估计误差可能越来越大，甚至发散。

Remark 4.3 必要的时候需要将所有的滤波器重新初始化，例如采用当前时刻最可能的滤波值及其协方差重启所有的滤波器。

4.2.3 动态多模型滤波

如果对象的模型不是一成不变的，每个时刻都可能从一个模型跳变到另外一个模型，此时的多模型方法研究起来要困难许多。一种简单的方法是按似然函数 (87) 计算任意时刻 k 模型 m_i 的概率，即

$$p_k(m_i) = \frac{\Lambda(\mathbf{y}_k|m_i)}{\sum_{j=1}^r \Lambda(\mathbf{y}_k|m_j)} \quad (92)$$

其他计算与上述静态多模型方法相同。当然，严格的处理方法需要引入一些新的假设和数学工具。

马尔可夫模型假设 假设模型序列 $\{M(t_0), M(t_1), \dots, M(t_k)\}$ 构成一个马尔可夫链，模型之间的转移概率可表示为

$$\pi_{ij}(t_k, t_{k-1}) = \Pr[M(t_k) = M_j | M(t_{k-1}) = M_i] \quad (93)$$

根据马尔可夫链的性质，有 $\sum_{i=1}^r \pi_{ij}(t_k, t_{k-1}) = 1$ （假设共有 r 个可能模型）。注意，如果 $\pi_{ij} = \delta_{ij}$ ，退化为静态多模型方法。

Remark 4.4 如果严格按照马尔可夫链模型，最优估计需要的计算时间和计算机内存将随时间成指数增加，没有任何实用价值。

Remark 4.5 已经发展起来了若干次优算法，例如广义伪贝叶斯方法(GPB)、交互多模型方法(IMM)等，使得每步都可以在有限的时间内获得次优的状态估计及模型估计。

交互多模型滤波算法

交互多模型算法是一种高效的多模型滤波算法，广泛应用于目标跟踪问题。IMM 滤波器结构见图2。

马尔可夫跳变系统

许多对象可以用线性马尔可夫跳变系统来描述，对应的状态方程和量测方程为

$$\mathbf{x}(k+1) = F_j \mathbf{x}(k) + G_j \mathbf{w}_j(k) \quad (94)$$

$$\mathbf{y}(k) = H_j \mathbf{x}(k) + \mathbf{v}_j(k) \quad (95)$$

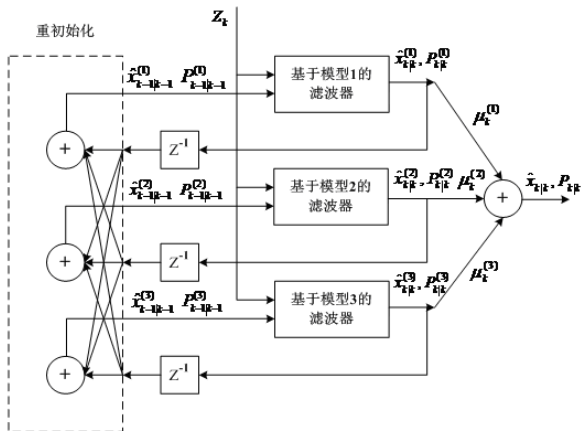


Figure 2: IMM 滤波器结构图

式中, $\mathbf{x}(0) \sim \mathcal{N}(\bar{x}_0, P_0)$; $\mathbf{w}_j(k)$ 与 $\mathbf{v}_j(k)$ 表示模型相关的过程噪声和量测噪声, 假设两者相互独立 (与 $\mathbf{x}(0)$ 也不相关), 都是零均值的高斯白噪声。

$$E\mathbf{w}_j(k) = 0, \quad E\mathbf{w}_j(k)\mathbf{w}_j(k)^T = Q_j \quad (96)$$

$$E\mathbf{v}_j(k) = 0, \quad E\mathbf{v}_j(k)\mathbf{v}_j(k)^T = R_j \quad (97)$$

记 $M_j(k) = \{F_j, G_j, H_j, Q_j, R_j\}$, 表示 k 时刻系统的模型, 假设为有限状态马尔可夫链。模型状态的转移概率设为

$$\pi_{ij} \triangleq \Pr[M_j(k) | M_i(k-1)] \geq 0 \quad (98)$$

$$\sum_{i=1}^r \pi_{ij} = 1, \quad i, j \in N \triangleq \{1, 2, \dots, r\} \quad (99)$$

模型状态马尔可夫链的初始分布为 $\{\mu_1(0), \mu_2(0), \dots, \mu_r(0)\}$, 满足

$$0 \leq \mu_j(0) \leq 1, \quad \sum_{j=1}^r \mu_j(0) = 1, \quad j \in N \quad (100)$$

IMM 算法步骤

设已知: $\hat{\mathbf{x}}_j(k-1), P_j(k-1), \mu_j(k-1), \pi_{ij}$ (设计参数)。下面给出一个采样周期内的算法步骤。

1. 计算模型交互概率:

$$\begin{aligned}
 \mu_{i|j}(k-1) &\triangleq \Pr[M_i(k-1)|M_j(k), Y^{k-1}] \\
 &= \frac{\Pr[M_j(k)|M_i(k-1), Y^{k-1}]\Pr[M_i(k-1)|Y^{k-1}]}{\Pr[M_j(k)|Y^{k-1}]} \\
 &= \frac{\pi_{ij}\mu_i(k-1)}{\bar{c}_j}, \quad i, j \in N
 \end{aligned} \tag{101}$$

2. 计算交互滤波初值 ($j \in N$):

$$\hat{\mathbf{x}}_{0j}(k-1) = \sum_{i=1}^r \mu_{i|j}(k-1)\hat{\mathbf{x}}_i(k-1) \tag{102}$$

$$P_{0j}(k-1) = \sum_{i=1}^r \mu_{i|j}(k-1)\{P_i(k-1) + [\hat{\mathbf{x}}_i((k-1) - \hat{\mathbf{x}}_{0j}(k-1))][*]^T\} \tag{103}$$

3. 滤波及模型概率修正 ($j \in N$):

$$[\hat{\mathbf{x}}_j(k), P_j(k), \mathbf{v}_j(k), S_j(k)] = \text{KF}(\hat{\mathbf{x}}_{0j}(k-1), P_{0j}(k-1), \mathbf{y}_k) \quad (104)$$

其中, $\mathbf{v}_j(k)$ 、 $S_j(k)$ 分别表示新息及其协方差, \mathbf{y}_k 为 k 时刻系统的量测 (矢量)。KF(\cdot, \cdot, \cdot) 表示标准卡尔曼滤波推算。

k 时刻模型 $M_j(k)$ 的似然函数为

$$\Lambda_j(k) = \mathcal{N}(\mathbf{v}_j(k); 0, S_j(k)) \quad (105)$$

模型概率为

$$\begin{aligned}\mu_j(k) &\triangleq \Pr[M_j(k)|Y^k] \\ &= \frac{1}{c} p[\mathbf{y}_k | M_j(k), Y^{k-1}] \Pr[M_j(k) | Y^{k-1}] \\ &= \frac{1}{c} \Lambda_j(k) \sum_{i=1}^r \Pr[M_j(k) | M_i(k-1), Y^{k-1}] \Pr[M_i(k-1) | Y^{k-1}] \\ &= \frac{1}{c} \Lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) = \frac{1}{c} \Lambda_j(k) \bar{c}_j\end{aligned}\quad (106)$$

4. 组合滤波输出:

$$\hat{\mathbf{x}}(k) = \sum_{j=1}^r \mu_j(k) \hat{\mathbf{x}}_j(k) \quad (107)$$

$$P(k) = \sum_{j=1}^r \mu_j(k) \{P_j(k) + [\hat{\mathbf{x}}(k) - \hat{\mathbf{x}}_j(k)][*]^T\} \quad (108)$$

上面的 \bar{c}_j 和 c 为归一化常数, 分别为

$$\bar{c}_j = \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad (109)$$

$$c = \sum_{j=1}^r \Lambda_j(k) \bar{c}_j \quad (110)$$

数值算例——机动目标跟踪

- (1) 假设目标初始位置为 $(1, 1, 1)$ (单位: km), 先以速度 $v = (50, 30, 10)$ (单位: m/s) 做匀速直线运动, 之后依次做匀速圆周运动、匀加速运动、匀速直线运动、匀速圆周运动和匀加速运动;
- (2) 第三阶段和第六阶段目标运动加速度分别为 $a = (0.5, -0.3, 0.5)$ 与 $a = (0.5, 0.5, -1)$ (单位: m/s^2);
- (3) 第二阶段和第五阶段目标匀速圆周运动角速度为 $\omega = \frac{\pi}{270} \text{rad/s}$ 。

采用 CV 模型（匀速模型）、CA 模型（匀加速模型）与 CS 模型（“当前”统计模型）三个模型进行滤波器设计。

模型状态转移概率矩阵为

$$[\pi_{ij}] = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.4 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

模型初始概率为

$$[\mu(0)] = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

N 次蒙特卡洛仿真计算的均方根误差定义为

$$RMSE_k = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\hat{\mathbf{x}}^{(i)}(k) - \mathbf{x}^{(i)}(k)\|^2} \quad (111)$$

仿真结果如图3与图4所示。结果表明，IMM 滤波器能够较好地实现机动目标的跟踪，比任何单一模型滤波器的效果都要好。

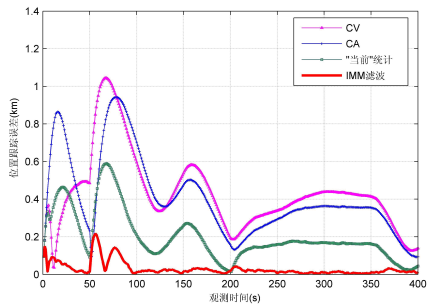


Figure 3: 机动目标的位置跟踪均方根误差

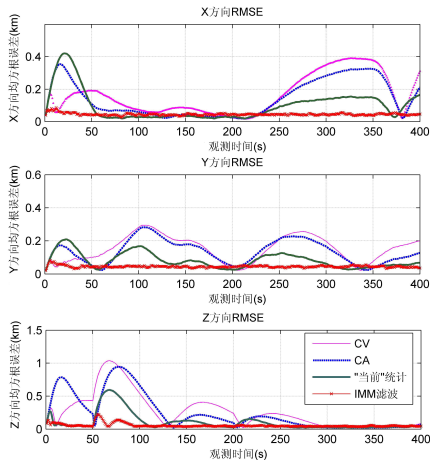


Figure 4: 机动目标三方向位置估计均方根误差

5. 常值增益次优滤波

- ♠ 在卡尔曼滤波算法中，最主要的计算量来自卡尔曼增益矩阵的计算。
- ♠ 为了减少卡尔曼滤波算法的计算量，可以合理地简化卡尔曼增益矩阵的计算。
- ♠ 由于滤波的稳态对工程实际具有特别的重要性，如果卡尔曼增益矩阵的稳态值为 K_∞ ，取

$$K_{k+1} = K_\infty \quad (112)$$

从而形成常值增益次优滤波。

对于定常离散时间系统

$$\mathbf{x}_{k+1} = F\mathbf{x}_k + G\mathbf{w}_k \quad (113)$$

$$\mathbf{y}_{k+1} = H\mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (114)$$

其中, $\mathbf{w}_k \sim (0, Q)$ 与 $\mathbf{v}_{k+1} \sim (0, R)$ 是互不相关的白噪声, 它们与初始状态 $\mathbf{x}_0 \sim (\bar{\mathbf{x}}_0, P_0)$ 不相关。另外, 设 $Q \geq 0, R > 0$ 。

如果系统是完全能控和完全能观的, 则有

$$K_\infty = \lim_{k \rightarrow \infty} K_k \quad (115)$$

$$P_\infty = \lim_{k \rightarrow \infty} P_{k|k} \quad (116)$$

$$\bar{P} = \lim_{k \rightarrow \infty} P_{k|k-1} \quad (117)$$

进一步地

$$K_{\infty} = \bar{P}H^T(H\bar{P}H^T + R)^{-1} \quad (118)$$

$$P_{\infty} = (I - K_{\infty}H)\bar{P} \quad (119)$$

$$\bar{P} = FP_{\infty}F^T + GQG^T \quad (120)$$

关于常值增益次优滤波，我们有如下结论。

Theorem 5.1 如果所研究的系统完全能控和完全能观，则存在常值增益滤波器

$$\begin{aligned}\hat{\mathbf{x}}_{k+1|k+1}^* &= \bar{F}\hat{\mathbf{x}}_{k|k}^* + K_\infty \mathbf{y}_{k+1}, & \hat{\mathbf{x}}_{0|0}^* &= \bar{\mathbf{x}}_0 \\ P_{k+1|k+1}^* &= \bar{F}P_{k|k}^* \bar{F}^T + \bar{Q}, & P_{0|0}^* &= P_0\end{aligned}$$

它在滤波初始阶段是次优的，而在稳态则是最优的。其中

$$\begin{aligned}\bar{F} &= (I - K_\infty H)F \\ \bar{Q} &= (I - K_\infty H)GQG^T(I - K_\infty H)^T + K_\infty RK_\infty^T\end{aligned}$$

稳态增益 K_∞ 由 (118) 唯一给出。

关于常增益卡尔曼滤波的进一步讨论及上述结论的证明见 Kleinman、Silva 及 Guosl 等人论文。

6. 鲁棒 H_∞ 滤波

6.1 动态约束优化

考虑如下动态系统

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k \quad (k = 0, \dots, N-1) \quad (121)$$

式中, \mathbf{x}_k 为 n 维状态向量。我们的目标是最小化如下标量函数

$$J = \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k(\mathbf{x}_k, \mathbf{w}_k) \quad (122)$$

式中, $\psi(\mathbf{x}_0)$ 是关于 \mathbf{x}_0 的已知函数, $\mathcal{L}_k(\mathbf{x}_k, \mathbf{w}_k)$ 是关于 \mathbf{x}_k 和 \mathbf{w}_k 的已知函数。通常要求 $\psi(\mathbf{x}_0)$ 和 $\mathcal{L}_k(\mathbf{x}_k, \mathbf{w}_k)$ 关于相关变量是光滑可导的。

设拉格朗日乘子为 λ_{k+1} (对应动态约束方程, 共有 N 个 n 维向量), 我们可以获得如下增广目标函数

$$J_a = \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} [\mathcal{L}_k(\mathbf{x}_k, \mathbf{w}_k) + \lambda_{k+1}^T (F_k \mathbf{x}_k + \mathbf{w}_k - \mathbf{x}_{k+1})] \quad (123)$$

上式可以改写整理为

$$\begin{aligned} J_a &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} [\mathcal{L}_k + \lambda_{k+1}^T (F_k \mathbf{x}_k + \mathbf{w}_k)] - \sum_{k=0}^{N-1} \lambda_{k+1}^T \mathbf{x}_{k+1} \\ &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} [\mathcal{L}_k + \lambda_{k+1}^T (F_k \mathbf{x}_k + \mathbf{w}_k)] - \sum_{k=0}^N \lambda_k^T \mathbf{x}_k + \lambda_0^T \mathbf{x}_0 \end{aligned}$$

定义如下哈密顿函数

$$\mathcal{H}_k(\mathbf{x}_k, \mathbf{w}_k, \boldsymbol{\lambda}_{k+1}) = \mathcal{L}_k(\mathbf{x}_k, \mathbf{w}_k) + \boldsymbol{\lambda}_{k+1}^T (F_k \mathbf{x}_k + \mathbf{w}_k) \quad (124)$$

增广目标函数可以重写为

$$\begin{aligned} J_a &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} \mathcal{H}_k - \sum_{k=0}^N \boldsymbol{\lambda}_k^T \mathbf{x}_k + \boldsymbol{\lambda}_0^T \mathbf{x}_0 \\ &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} \mathcal{H}_k - \sum_{k=0}^{N-1} \boldsymbol{\lambda}_k^T \mathbf{x}_k - \boldsymbol{\lambda}_N^T \mathbf{x}_N + \boldsymbol{\lambda}_0^T \mathbf{x}_0 \\ &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} (\mathcal{H}_k - \boldsymbol{\lambda}_k^T \mathbf{x}_k) - \boldsymbol{\lambda}_N^T \mathbf{x}_N + \boldsymbol{\lambda}_0^T \mathbf{x}_0 \end{aligned}$$

显然，最优解的必要条件（驻点）为

$$\begin{aligned}\frac{\partial J_a}{\partial \mathbf{x}_k} &= 0 \quad (k = 0, \dots, N) \\ \frac{\partial J_a}{\partial \mathbf{w}_k} &= 0 \quad (k = 0, \dots, N - 1) \\ \frac{\partial J_a}{\partial \lambda_k} &= 0 \quad (k = 0, \dots, N)\end{aligned}$$

进一步地可以写为

$$\frac{\partial J_a}{\partial \mathbf{x}_0} = 0$$

$$\frac{\partial J_a}{\partial \mathbf{x}_N} = 0$$

$$\frac{\partial J_a}{\partial \mathbf{x}_k} = 0 \quad (k = 1, \dots, N - 1)$$

$$\frac{\partial J_a}{\partial \mathbf{w}_k} = 0 \quad (k = 0, \dots, N - 1)$$

$$\frac{\partial J_a}{\partial \boldsymbol{\lambda}_k} = 0 \quad (k = 0, \dots, N)$$

即

$$\boldsymbol{\lambda}_0 + \frac{\partial \psi_0}{\partial \mathbf{x}_0} = 0 \quad (125)$$

$$\boldsymbol{\lambda}_N = 0 \quad (126)$$

$$\boldsymbol{\lambda}_k = \frac{\partial \mathcal{H}_k}{\partial \mathbf{x}_k} \quad (k = 0, 1, \dots, N-1) \quad (127)$$

$$\frac{\partial \mathcal{H}_k}{\partial \mathbf{w}_k} = 0 \quad (k = 0, 1, \dots, N-1) \quad (128)$$

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k \quad (k = 0, 1, \dots, N-1) \quad (129)$$

通过求解上述方程组，即可获得动态约束优化问题的最优解（如果存在）。

按惯例，我们用上标“*”表示最优解，将 λ_k 称为协态变量，对应这里的动态约束优化问题，求得的最优解可表示为

$$\begin{aligned} & \{\mathbf{x}_k^* | k = 0, 1, 2, \dots, N\} \\ & \{\mathbf{w}_k^* | k = 0, 1, 2, \dots, N - 1\} \\ & \{\lambda_k | k = 0, 1, 2, \dots, N\} \end{aligned}$$

不难发现，未上述知变量的个数与约束方程个数相等，理论上有解。以上关于动态约束优化的方法及结果可以用来解决下面的 H_∞ 滤波问题。

6.2 H_∞ 滤波算法

考虑如下离散时间系统

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k \quad (130)$$

$$\mathbf{y}_k = H_k \mathbf{x}_k + \mathbf{v}_k \quad (131)$$

式中, \mathbf{w}_k 和 \mathbf{v}_k 为噪声项, 分别表示模型及数据的不确定性。这些噪声可能是随机的 (统计特性未知) 也可能是确定的, 它们的均值可能不为 0。

我们的目标是对状态的线性组合进行估计，可以表示为

$$\mathbf{z}_k = L_k \mathbf{x}_k \quad (132)$$

其中， L_k 是自定义的矩阵（假设 L_k 满秩）。如果我们想要直接估计 \mathbf{x}_k ，设置 $L_k = I$ 。但通常情况下我们可能只对状态量的某种线性组合感兴趣。

利用博弈理论的思想，定义如下目标函数

$$J_1 = \frac{\sum_{k=0}^{N-1} \|\mathbf{z}_k - \hat{\mathbf{z}}_k\|_{S_k}^2}{\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1} (\|\mathbf{w}_k\|_{Q_k^{-1}}^2 + \|\mathbf{v}_k\|_{R_k^{-1}}^2)} \quad (133)$$

式中， P_0 、 Q_k 、 R_k 和 S_k 均为对称正定权重矩阵，是设计参数。

于是，我们得到如下 min-max 问题

$$J_1^* = \min_{\hat{z}_k} \max_{\mathbf{w}_k, \mathbf{v}_k, \mathbf{x}_0} J_1 \quad (134)$$

由此可见 H_∞ 滤波和卡尔曼滤波本质上的区别。数学上，我们把问题转化为寻找估计 \hat{z}_k ，同时使得

$$J_1 < \epsilon \quad (135)$$

式中， $\epsilon > 0$ 为自定义的性能边界参数。结合 J_1 的表达式，上式可以重写为

$$J = -\epsilon \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{P_0}^2 + \sum_{k=0}^{N-1} [\|\mathbf{z}_k - \hat{\mathbf{z}}_k\|_{S_k}^2 - \epsilon(\|\mathbf{w}_k\|_{Q_k}^2 + \|\mathbf{v}_k\|_{R_k}^2)] < 0 \quad (136)$$

于是，我们的问题转化为如下 min-max 问题

$$J^* = \min_{\hat{z}_k} \max_{\mathbf{w}_k, \mathbf{v}_k, \mathbf{x}_0} J \quad (137)$$

上述 min-max 问题可以改写为

$$J^* = \min_{\hat{\mathbf{x}}_k} \max_{\mathbf{w}_k, \mathbf{v}_k, \mathbf{x}_0} J \quad (138)$$

可以进一步化为

$$J^* = \min_{\hat{\mathbf{x}}_k} \max_{\mathbf{w}_k, \mathbf{y}_k, \mathbf{x}_0} J \quad (139)$$

定义

$$\bar{S}_k = L_k^T S_k L_k \quad (140)$$

可导出

$$\begin{aligned} J &= -\epsilon \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1} [\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_{\bar{S}_k}^2 - \epsilon(\|\mathbf{w}_k\|_{Q_k^{-1}}^2 + \|\mathbf{y}_k - H_k \mathbf{x}_k\|_{R_k^{-1}}^2)] \\ &= \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k \end{aligned} \quad (141)$$

上式给出了 $\psi(\mathbf{x}_0)$ 和 \mathcal{L}_k 的定义。

6.2.1 关于 x_0 和 w_k 的驻点

定义哈密尔顿函数

$$\mathcal{H}_k = \mathcal{L}_k + 2\epsilon \boldsymbol{\lambda}_{k+1}^T (F_k \mathbf{x}_k + \mathbf{w}_k) \quad (142)$$

式中, $2\epsilon \boldsymbol{\lambda}_{k+1} (k = 0, \dots, N-1)$ 为时变的拉格朗日乘子。

根据上小节介绍的动态约束优化, 不难发现, J 关于 \mathbf{x}_0 和 \mathbf{w}_k 的驻点方程为

$$2\epsilon\boldsymbol{\lambda}_0 + \frac{\partial\psi_0}{\partial\mathbf{x}_0} = 0 \quad (143)$$

$$2\epsilon\boldsymbol{\lambda}_N = 0 \quad (144)$$

$$\frac{\partial\mathcal{H}_k}{\partial\mathbf{w}_k} = 0 \quad (145)$$

$$2\epsilon\boldsymbol{\lambda}_k = \frac{\partial\mathcal{H}_k}{\partial\mathbf{x}_k} \quad (146)$$

由此可导出

$$\begin{aligned}
 \boldsymbol{\mu}_{k+1} - F_k \boldsymbol{\mu}_k - F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} \\
 \times [\theta \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^T R_k^{-1} (\mathbf{y}_k - H_k \boldsymbol{\mu}_k)] \\
 = [-P_{k+1} + F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} F_k^T + Q_k] \boldsymbol{\lambda}_{k+1}
 \end{aligned} \tag{147}$$

令 (147) 左端为 0, 可导出

$$\begin{aligned}
 \boldsymbol{\mu}_{k+1} = F_k \boldsymbol{\mu}_k + F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} \\
 \times [\theta \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^T R_k^{-1} (\mathbf{y}_k - H_k \boldsymbol{\mu}_k)]
 \end{aligned} \tag{148}$$

这是关于 $\boldsymbol{\mu}_k$ 的演化方程, 其初始条件为 $\boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0$ 。

令 (147) 右端为 0, 可导出

$$\begin{aligned} P_{k+1} &= F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} F_k^T + Q_k \\ &= F_k \tilde{P}_k F_k^T + Q_k \end{aligned} \quad (149)$$

这是关于 P_k 的演化方程, 式中 \tilde{P}_k 的定义为

$$\begin{aligned} \tilde{P}_k &= P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} \\ &= [P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k]^{-1} \end{aligned} \quad (150)$$

优化目标函数 J 的最优解 \mathbf{x}_0 和 \mathbf{w}_k 的求解过程可以总结如下:

$$\mathbf{x}_0 = \hat{\mathbf{x}}_0 + P_0 \boldsymbol{\lambda}_0 \quad (151)$$

$$\mathbf{w}_k = Q_k \boldsymbol{\lambda}_{k+1} \quad (152)$$

$$\begin{aligned} \boldsymbol{\lambda}_k &= [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} \\ &\quad \times [F_k^T \boldsymbol{\lambda}_{k+1} + \theta \bar{S}_k (\boldsymbol{\mu}_k - \hat{\mathbf{x}}_k) + H_k^T R_k^{-1} (\mathbf{y}_k - H_k \boldsymbol{\mu}_k)], \quad \boldsymbol{\lambda}_N = 0 \end{aligned} \quad (153)$$

$$P_{k+1} = F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} F_k^T + Q_k, \quad P_0 = P_0 \quad (154)$$

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &= F_k \boldsymbol{\mu}_k + F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} \\ &\quad \times [\theta \bar{S}_k (\boldsymbol{\mu}_k - \hat{\mathbf{x}}_k) + H_k^T R_k^{-1} (\mathbf{y}_k - H_k \boldsymbol{\mu}_k)], \quad \boldsymbol{\mu}_0 = \hat{\mathbf{x}}_0 \end{aligned} \quad (155)$$

6.2.2 关于 $\hat{\mathbf{x}}_k$ 和 \mathbf{y}_k 的驻点

假定 \mathbf{x}_0 和 \mathbf{w}_k 已经取得最大值的前提下, 本小节的问题是寻找 $J = \psi(\mathbf{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$ (满足约束条件 $\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k$) 关于 $\hat{\mathbf{x}}_k$ 和 \mathbf{y}_k 的驻点。

性能指标 (141) 可以改写为

$$J = -\epsilon \|\boldsymbol{\lambda}_0\|_{P_0}^2 + \sum_{k=0}^{N-1} [\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_{\bar{S}_k}^2 - \epsilon(\|\mathbf{w}_k\|_{Q_k^{-1}}^2 + \|\mathbf{y}_k - H_k \mathbf{x}_k\|_{R_k^{-1}}^2)] \quad (156)$$

利用上式, 求 J 关于 $\hat{\mathbf{x}}_k$ 和 \mathbf{y}_k 的偏导数并令其等于零, 可以得到如下方程

$$\begin{aligned} \frac{\partial J}{\partial \hat{\mathbf{x}}_k} &= 2(\bar{S}_k + \theta \bar{S}_k \tilde{P}_k \bar{S}_k)(\hat{\mathbf{x}}_k - \boldsymbol{\mu}_k) \\ &\quad + 2\bar{S}_k \tilde{P}_k H_k^T R_k^{-1}(H_k \boldsymbol{\mu}_k - \mathbf{y}_k) = 0 \end{aligned} \quad (157)$$

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{y}_k} &= 2\epsilon(R_k^{-1} H_k \tilde{P}_k H_k^T R_k^{-1} - R_k^{-1})(\mathbf{y}_k - H_k \boldsymbol{\mu}_k) \\ &\quad + 2R_k^{-1} H_k \tilde{P}_k \bar{S}_k(\boldsymbol{\mu}_k - \hat{\mathbf{x}}_k) = 0 \end{aligned} \quad (158)$$

当 $\hat{\mathbf{x}}_k = \boldsymbol{\mu}_k$ 以及 $\mathbf{y}_k = H_k \boldsymbol{\mu}_k$ 时, 上述方程组显然成立。

上述驻点条件是最优的必要条件，为了验证 $\hat{\boldsymbol{x}}_k = \boldsymbol{\mu}_k$ 是我们所求的最优解，我们可以导出 J 关于 $\hat{\boldsymbol{x}}_k$ 的二阶导数

$$\frac{\partial^2 J}{\partial \hat{\boldsymbol{x}}_k^2} = 2(\bar{S}_k + \theta \bar{S}_k \tilde{P}_k \bar{S}_k) \quad (159)$$

通常情况下， \bar{S}_k 的正定性可以保证，如果 \tilde{P}_k 正定则 $\hat{\boldsymbol{x}}_k$ 为极小值点。

根据式中 \tilde{P}_k 的定义 (150) 可知， $\hat{\boldsymbol{x}}_k$ 为极小值点的条件为

$$(P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k)^{-1} > 0$$

等价于

$$(P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k) > 0$$

为了保证正定性， $\theta \bar{S}_k$ 应尽可能小。为了使 $\theta \bar{S}_k$ 变小，有如下 3 种途径：

1) $\theta = 1/\epsilon$ 变小可以使 $\theta\bar{S}_k$ 变小, 这意味着 (135) 或 (136) 式中定义的性能要求不能太严格。如果性能要求不是很严格, 可以求解; 反之, 则无法求解。

2) L_k 变小可以使 $\theta\bar{S}_k$ 变小。这个结论是根据式 (140) 中 \bar{S}_k 和 L_k 的关系得出的。目标函数的分子为

$$(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T L_k^T S_k L_k (\mathbf{x}_k - \hat{\mathbf{x}}_k)$$

L_k 变小可以使目标函数的分子变小, 最小化目标函数将变得容易; L_k 过大可能会导致问题无解。

3) S_k 变小可以使 $\theta\bar{S}_k$ 变小。与上面的结论类似, S_k 变小可以使目标函数的分子变小, 最小化目标函数将变得容易; S_k 过大可能会导致问题无解。

6.2.3 H_∞ 鲁棒滤波算法小结

系统模型

动态系统及输出方程:

$$\begin{cases} \mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k = H_k \mathbf{x}_k + \mathbf{v}_k \\ \mathbf{z}_k = L_k \mathbf{x}_k \end{cases} \quad (160)$$

式中, \mathbf{w}_k 和 \mathbf{v}_k 分别表示过程噪声和量测噪声, 目标是估计状态 \mathbf{x}_k 。

目标函数

选取合适的初始估计及权重矩阵，构造如下目标函数：

$$J_1 = \frac{\sum_{k=0}^{N-1} \|\mathbf{z}_k - \hat{\mathbf{z}}_k\|_{S_k}^2}{\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1} (\|\mathbf{w}_k\|_{Q_k^{-1}}^2 + \|\mathbf{v}_k\|_{R_k^{-1}}^2)} \quad (161)$$

式中，矩阵 P_0 ， Q_k ， R_k 和 S_k 均为正定对称矩阵。

滤波算法

如下 H_∞ 递推滤波算法使目标函数 J_1 小于 ϵ ，其中 $\theta = 1/\epsilon$ 。

$$\begin{cases} \hat{\mathbf{x}}_{k+1} = F_k \hat{\mathbf{x}}_k + F_k K_k (\mathbf{y}_k - H_k \hat{\mathbf{x}}_k), & \hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_0 \\ P_{k+1} = F_k P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} F_k^T + Q_k, & P_0 = P_0 \end{cases} \quad (162)$$

式中

$$\begin{cases} K_k = P_k [I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k]^{-1} H_k^T R_k^{-1} \\ \bar{S}_k = L_k^T S_k L_k \end{cases} \quad (163)$$

需要注意， H_∞ 滤波能够求解问题的前提条件是：在每个采样时刻 k ，不等式 $P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k > 0$ 成立。

6.3 H_∞ 滤波和卡尔曼滤波的区别和联系

一步预测形式的卡尔曼滤波方程可表示为

$$\hat{\mathbf{x}}_{k+1|k} = F_k \hat{\mathbf{x}}_{k|k-1} + F_k K_k (\mathbf{y}_k - H_k \hat{\mathbf{x}}_{k|k-1}) \quad (164)$$

$$K_k = P_{k|k-1} (I + H_k^T R_k^{-1} H_k P_{k|k-1})^{-1} H_k^T R_k^{-1} \quad (165)$$

$$P_{k+1|k} = F_{k-1} P_{k|k-1} (I + H_k^T R_k^{-1} H_k P_{k|k-1})^{-1} F_k^T + Q_k \quad (166)$$

对比分析 (162)、(163) 和 (164)~(166)，可以发现 H_∞ 滤波算法与卡尔曼滤波算法之间的区别和联系。

- ♠ 在 H_∞ 滤波算法中, Q_k 、 R_k 和 P_0 等权重矩阵是设计参数, 需要根据过程扰动 w_k 、量测扰动 v_k 和初始估计误差 $(x_0 - \hat{x}_0)$ 幅度的先验信息事先设置;
- ♠ 在卡尔曼滤波算法中, 认为 w_k 、 v_k 和 $(x_0 - \bar{x}_0)$ 服从均值为零的高斯分布, Q_k 、 R_k 和 P_0 分别是它们各自已知的协方差矩阵。
- ◇ 在 H_∞ 滤波中取 $L_k = S_k = I$, 对应于对状态变量全体进行估计, 同时同等对待所有的估计误差。此时, 当取 $\epsilon = +\infty$ ($\theta = 0$), H_∞ 滤波变为卡尔曼滤波。
- ◇ 由上一点可以重新定义卡尔曼滤波, 即卡尔曼滤波为式 (161) 表示的性能指标上界为 ∞ 时的 min-max 滤波。

- ◇ 尽管卡尔曼滤波能够最小化估计误差方差，但是并不能限制最差情况下的估计误差。因此，卡尔曼滤波不能保证目标函数的界限。
- 卡尔曼滤波和 H_∞ 滤波算法有一个很有趣的区别。如果想利用卡尔曼滤波估计状态的线性组合，与不考虑线性组合情况下是一致的。
- 当使用 H_∞ 滤波时，结果与 L_k 和想要估计的状态组合有很大关系。
- ♣ 观察 (162) 和 (163) 式，当去掉 K_k 和 P_{k+1} 方程中的 $\theta \bar{S}_k P_k$ 项后， H_∞ 滤波算法形式上变为卡尔曼滤波完全一样。

- 在卡尔曼滤波中，对于未建模的动态系统和噪声，可以通过增大 Q_k 以改善卡尔曼滤波的鲁棒性。这样做会使得协方差矩阵 $P_{k+1|k}$ 增大，同时使得卡尔曼滤波增益矩阵 K_k 变大。
- 在 H_∞ 滤波中，从 (162) 式可以看出， P_{k+1} 等式右边 $(-\theta \bar{S}_k P_k)$ 项相当于变大 P_{k+1} 。类似地，也会使得矩阵 K_k 变大。

以上分析表明， H_∞ 滤波器可以认为是一种鲁棒的卡尔曼滤波器。但 P_{k+1} 不再具备估计误差协方差的含义！

为了直观上比较卡尔曼滤波算法和 H_∞ 滤波算法，下面给出一个简单算例。

Example 6.1 考虑标量系统

$$\begin{cases} x_{k+1} = x_k + w_k \\ y_k = x_k + v_k \\ z_k = x_k \end{cases}$$

其中, $w_k \sim \mathcal{N}(0, 1)$, $v_k \sim \mathcal{N}(0, 1)$ 。对照上面介绍的算法, 令 $F_k = H_k = L_k = 1$, $Q_k = R_k = S_k = 1$ 。

标称条件仿真 假设初始状态 $x_0 = 0$ ，在系统模型参数和噪声统计特性均已知的条件下，分别采用标准卡尔曼滤波和 H_∞ 滤波算法进行 100 次蒙特卡洛仿真实验，状态估计仿真结果对比见图 5，均方根误差 (RSME) 结果比较见图 6。不同 θ 取值下 H_∞ 滤波平均均方根误差结果比较如表 3 所示。

噪声统计特性未知情况仿真 同样假设初始状态 $x_0 = 0$ ，在过程噪声统计特性未知条件下（真实过程噪声均值取 10），分别采用标准卡尔曼滤波和 H_∞ 滤波算法进行 100 次蒙特卡洛仿真实验，状态估计仿真结果对比见图 7，均方根误差 (RSME) 结果比较见图 8。不同 θ 取值下 H_∞ 滤波平均均方根误差结果比较如表 4 所示。

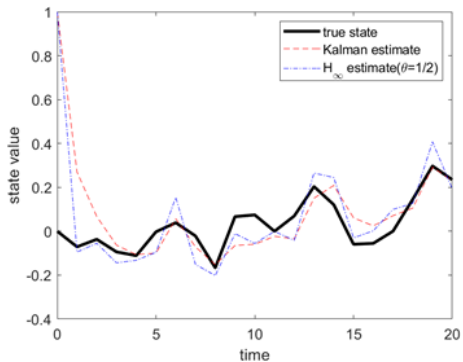


Figure 5: 标称条件滤波结果比较

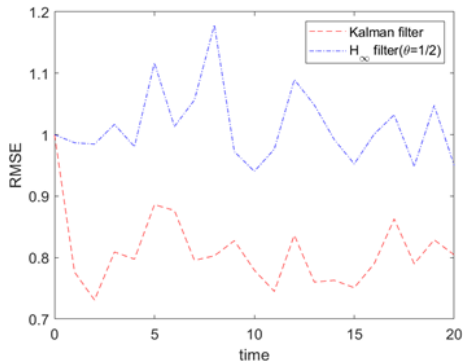


Figure 6: 标称条件均方根误差比较

Table 3: 标称条件平均均方根误差结果比较

方法	卡尔曼滤波	H_∞ 滤波 (不同 θ 值)				
		0	1/10	1/3	7/16	1/2
平均均方根误差	0.7954	0.7954	0.8026	0.8580	0.9193	0.9576



Table 4: 过程噪声统计特性未知平均均方根误差结果比较

方法	卡尔曼滤波	H_∞ 滤波 (不同 θ 值)				
		0	1/10	1/3	7/16	1/2
平均均方根误差	5.7441	5.7441	4.6687	2.2215	1.1936	0.9995

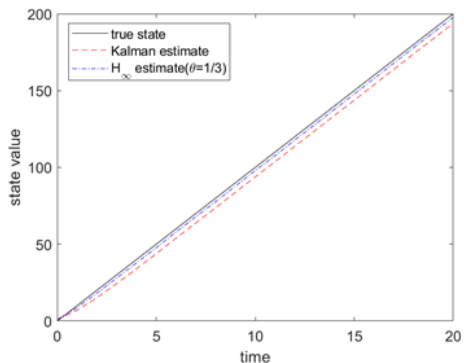


Figure 7: 过程噪声统计特性未知滤波结果比较

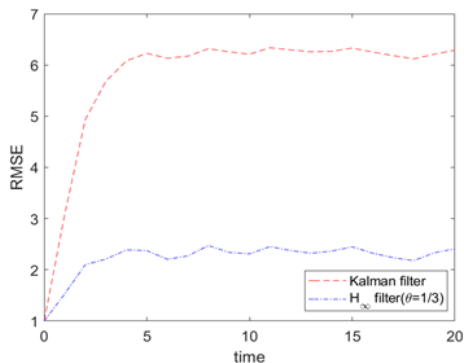


Figure 8: 过程噪声统计特性未知均方根误差比较

分析上述两组仿真实验结果，不难发现：

1. 在标称条件下，即已知系统参数及噪声统计特性时，卡尔曼滤波的估计精度高于 H_∞ 滤波；
2. 在滤波器设计参数明显不同于实际参数时，卡尔曼滤波性能下降很多；而合适选择 θ 值， H_∞ 滤波可以获得接近最优的估计精度；
3. 当取 $\theta = 0$ 时，即 $\epsilon = +\infty$ ， H_∞ 滤波等同于卡尔曼滤波；
4. 当系统参数及噪声统计特性存在不确定性时， H_∞ 滤波的估计精度随 θ 取值的增大而提高；但太大的 θ 取值， H_∞ 滤波不存在可行解。

□

6. 本章小结

本章探讨了标准卡尔曼滤波理论的若干扩展主题，分别介绍了衰减记忆滤波、限定记忆滤波、协方差平方根滤波、自适应滤波、常增益滤波以及 H_∞ 滤波等算法。

主要目的是提高卡尔曼滤波算法在实际应用中的鲁棒性，许多算法以牺牲一定的最优性为代价，所以也称为次优滤波算法。

Questions?