鲁棒 H_∞ 滤波

Prof. Yuan-Li Cai

Spring 2025

0. Outline

当系统模型及噪声统计特性存在大的不确定性或者完全未知时，标准的卡尔曼滤波算法将无法保证状态估计在最小均方误差意义下的最优，甚至会出现发散现象。为此，近年人们基于 H_∞ 控制理论，发展起来了一种称为 H_∞ 滤波的方法，在理论界和工程领域引起了极大关注。

本节将在讨论动态约束优化的基础上，介绍 H_∞ 滤波器的基本原理，并讨论、分析与卡尔曼滤波器之间的区别和联系。

1. 动态约束优化

考虑如下动态系统

$$\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k \quad (k = 0, \cdots, N-1) \tag{1}$$

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

$$J = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k(\boldsymbol{x}_k, \boldsymbol{w}_k) \tag{2}$$

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

(1) 和 (2) 便构成了一个带有约束的动态优化问题，(1) 称为动态约束，(2) 称为优化目标函数或性能指标。和经典的最优控制问题略有不同，主要差异在 $\psi(\boldsymbol{x}_0)$ 项。和所有约束优化问题一样，我们通过引入拉格朗日乘子来解决上述动态约束优化问题。

设拉格朗日乘子为 $\boldsymbol{\lambda}_{k+1}$(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

$$J_a = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \left[\mathcal{L}_k(\boldsymbol{x}_k, \boldsymbol{w}_k) + \boldsymbol{\lambda}_{k+1}^{\mathrm{T}} (F_k \boldsymbol{x}_k + \boldsymbol{w}_k - \boldsymbol{x}_{k+1})\right] \tag{3}$$

上式可以改写整理为

$$\begin{aligned} J_a &= \psi(\boldsymbol{x_0}) + \sum_{k=0}^{N-1} [\mathcal{L}_k + \boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(F_k\boldsymbol{x}_k + \boldsymbol{w}_k)] - \sum_{k=0}^{N-1} \boldsymbol{\lambda}_{k+1}^{\mathrm{T}}\boldsymbol{x}_{k+1} \\ &= \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} [\mathcal{L}_k + \boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(F_k\boldsymbol{x}_k + \boldsymbol{w}_k)] - \sum_{k=0}^{N} \boldsymbol{\lambda}_{k}^{\mathrm{T}}\boldsymbol{x}_{k} + \boldsymbol{\lambda}_{0}^{\mathrm{T}}\boldsymbol{x}_{0} \end{aligned}$$

式中，$\boldsymbol{\lambda}_0$ 为拉格朗日乘子序列的附加项，它并不在原始的增广目标函数中。后面我们将会看到当约束优化问题得到解决后，它的取值也随之可以确定下来。

定义如下哈密尔顿函数

$$\mathcal{H}_k(\boldsymbol{x}_k, \boldsymbol{w}_k, \boldsymbol{\lambda}_{k+1}) = \mathcal{L}_k(\boldsymbol{x}_k, \boldsymbol{w}_k) + \boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(F_k\boldsymbol{x}_k + \boldsymbol{w}_k) \tag{4}$$

增广目标函数可以重写为

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

显然，最优解的必要条件（驻点）为

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

进一步地可以写为

$$\frac{\partial J_a}{\partial \boldsymbol{x}_0} = 0$$
$$\frac{\partial J_a}{\partial \boldsymbol{x}_N} = 0$$
$$\frac{\partial J_a}{\partial \boldsymbol{x}_k} = 0 \quad (k = 1, \cdots, N-1)$$
$$\frac{\partial J_a}{\partial \boldsymbol{w}_k} = 0 \quad (k = 0, \cdots, N-1)$$
$$\frac{\partial J_a}{\partial \boldsymbol{\lambda}_k} = 0 \quad (k = 0, \cdots, N)$$

即

$$\boldsymbol{\lambda}_0 + \frac{\partial \psi_0}{\partial \boldsymbol{x}_0} = 0 \tag{5}$$

$$\boldsymbol{\lambda}_N = 0 \tag{6}$$

$$\boldsymbol{\lambda}_k = \frac{\partial \mathcal{H}_k}{\partial \boldsymbol{x}_k} \quad (k = 0, 1, \cdots, N-1) \tag{7}$$

$$\frac{\partial \mathcal{H}_k}{\partial \boldsymbol{w}_k} = 0 \quad (k = 0, 1, \cdots, N-1) \tag{8}$$

$$\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k \quad (k = 0, 1, \cdots, N-1) \tag{9}$$

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 $\boldsymbol{\lambda}_k$ 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

$$\{\boldsymbol{x}_k^* | k = 0, 1, 2, \cdots, N\}$$

$$\{\boldsymbol{w}_k^* | k = 0, 1, 2, \cdots, N-1\}$$

$$\{\boldsymbol{\lambda}_k | k = 0, 1, 2, \cdots, N\}$$

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。

2. H_∞ 滤波算法

考虑如下离散时间系统

$$\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k \tag{10}$$

$$\boldsymbol{y}_k = H_k \boldsymbol{x}_k + \boldsymbol{v}_k \tag{11}$$

式中，$\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 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

$$\boldsymbol{z}_k = L_k \boldsymbol{x}_k \tag{12}$$

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 $\boldsymbol{x}_k$，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。

采用惯例符号，用 $\hat{\boldsymbol{z}}_k$ 表示 $\boldsymbol{z}_k$ 的估计量。那么，我们的目的就是利用所有可能的信息，获得使 $\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{S_k}^2$ 尽可能小的估计 $\hat{\boldsymbol{z}}_k$，建立类似卡尔曼滤波那样的递推算法。其中，S_k 是对称正定的权重矩阵。

显然，代表不确定性的环境（以后称为对方）不会配合我们，而会使得我们希望的估计变差。在 H_∞ 滤波中，假设对方可以调动所有资源，即最优地选择 $\boldsymbol{w}_k$、$\boldsymbol{v}_k$ 和初始状态 $\boldsymbol{x}_0$，使得 $\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{S_k}^2$ 尽可能的大。

如果不加限制，对方可以取 $\boldsymbol{w}_k$、$\boldsymbol{v}_k$ 和初始状态 $\boldsymbol{x}_0$ 为无穷大，这种情况明显没有实际意义。因此，我们可以利用博弈理论的思想，定义如下目标函数

$$J_1 = \frac{\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{S_k}^2}{\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1} \left(\|\boldsymbol{w}_k\|_{Q_k^{-1}}^2 + \|\boldsymbol{v}_k\|_{R_k^{-1}}^2\right)} \tag{13}$$

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

$$J_1^* = \min_{\hat{\boldsymbol{z}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{v}_k, \boldsymbol{x}_0} J_1 \tag{14}$$

由此可见 H_∞ 滤波和卡尔曼滤波本质上的区别。在卡尔曼滤波中，我们对对方是漠不关心的，同时认为噪声的统计特性是已知的，我们可以利

用先验知识得到统计意义下的最优状态估计，可以认为是一种单边最优控制，对方无法改变噪声的统计特性来影响我们的状态估计效果。然而在 H_∞ 滤波中，认为对手会想尽办法降低我们的估计效果。可以认为是双方智能博弈。

在卡尔曼滤波中，没有考虑矩阵 S_k。如果在卡尔曼滤波中，最小化 S_k 矩阵加权的估计误差方差，结果不受影响。而在 H_∞ 滤波中，后面会发现，S_k 矩阵的取值会影响滤波增益。

直接求解上述 min-max 问题 (14) 非常困难，下面通过设定一个性能边界来获取满足边界条件的估计策略，即保证一定性能的鲁棒策略。

数学上，我们把问题转化为寻找估计 $\hat{z}_k$，同时使得

$$J_1 < \epsilon \tag{15}$$

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

$$J = -\epsilon\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{P_0^{-1}} + \sum_{k=0}^{N-1}[\|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{S_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{Q_k^{-1}} + \|\boldsymbol{v}_k\|^2_{R_k^{-1}})] < 1 \quad (16)$$

于是，我们的问题转化为如下 min-max 问题

$$J^* = \min_{\hat{\boldsymbol{z}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{v}_k, \boldsymbol{x}_0} J \quad (17)$$

由于 $\boldsymbol{z}_k = L_k\boldsymbol{x}_k$，很自然地可以选择 $\hat{\boldsymbol{z}}_k = L_k\hat{\boldsymbol{x}}_k$，从而只需要设法获得最小化 J 的状态估计 $\hat{\boldsymbol{x}}_k$。因此，上述 min-max 问题可以改写为

$$J^* = \min_{\hat{\boldsymbol{x}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{v}_k, \boldsymbol{x}_0} J \quad (18)$$

在博弈过程中，对方将选择 $\boldsymbol{x}_0$、$\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 去最大化 J。在给定 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 条件下，$\boldsymbol{v}_k$ 完全决定了 $\boldsymbol{y}_k$，所以我们可以将上述 min-max 问题进一步化为

$$J^* = \min_{\hat{\boldsymbol{x}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{y}_k, \boldsymbol{x}_0} J \tag{19}$$

由 $\boldsymbol{y}_k = H_k \boldsymbol{x}_k + \boldsymbol{v}_k$，可知

$$\|\boldsymbol{v}_k\|^2_{R_k^{-1}} = \|\boldsymbol{y}_k - H_k \boldsymbol{x}_k\|^2_{R_k^{-1}} \tag{20}$$

另外，$\boldsymbol{z}_k = L_k \boldsymbol{x}_k$，$\hat{\boldsymbol{z}}_k = L_k \hat{\boldsymbol{x}}_k$，所以

$$
\begin{aligned}
\|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{S_k}^2 &= (\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^{\mathrm{T}} S_k (\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k) \\
&= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} L_k^{\mathrm{T}} S_k L_k (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k) \\
&= \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|_{\bar{S}_k}^2
\end{aligned} \tag{21}
$$

式中，$\bar{S}_k$ 定义为

$$
\bar{S}_k = L_k^{\mathrm{T}} S_k L_k \tag{22}
$$

综合上述讨论，我们可导出

$$
\begin{aligned}
J &= -\epsilon \|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{P_0^{-1}} + \sum_{k=0}^{N-1} [\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|^2_{\bar{S}_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{Q_k^{-1}} + \|\boldsymbol{y}_k - H_k \boldsymbol{x}_k\|^2_{R_k^{-1}})] \\
&= \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k
\end{aligned} \tag{23}
$$

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

为了解决上述 min-max 问题，我们首先寻找目标函数 J 关于 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 的驻点，然后求解 J 关于 $\hat{\boldsymbol{x}}_k$ 和 $\boldsymbol{y}_k$ 的驻点。

2.1 关于 x_0 和 w_k 的驻点

本小节的问题是寻找 $J = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$（满足约束条件 $\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k$）关于 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 的驻点。显然，这是我们前面讨论过的动态约束优化问题。

现在的哈密尔顿函数定义为

$$\mathcal{H}_k = \mathcal{L}_k + 2\epsilon \boldsymbol{\lambda}_{k+1}^{\mathrm{T}} (F_k \boldsymbol{x}_k + \boldsymbol{w}_k) \tag{24}$$

式中，$2\epsilon\boldsymbol{\lambda}_{k+1} (k = 0, \cdots, N-1)$ 为时变的拉格朗日乘子。需要注意的是此时拉格朗日乘子由 $\boldsymbol{\lambda}_{k+1}$ 变为了 $2\epsilon\boldsymbol{\lambda}_{k+1}$。但这并不改变问题的解，而只是

利用一个常数对拉格朗日乘子进行了缩放，从而使得问题的数学描述更加简洁。

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

$$2\epsilon\boldsymbol{\lambda}_0 + \frac{\partial\psi_0}{\partial\boldsymbol{x}_0} = 0 \tag{25}$$

$$2\epsilon\boldsymbol{\lambda}_N = 0 \tag{26}$$

$$\frac{\partial\mathcal{H}_k}{\partial\boldsymbol{w}_k} = 0 \tag{27}$$

$$2\epsilon\boldsymbol{\lambda}_k = \frac{\partial\mathcal{H}_k}{\partial\boldsymbol{x}_k} \tag{28}$$

由 (25) 式可得

$$\boldsymbol{x}_0 = \hat{\boldsymbol{x}}_0 + P_0 \boldsymbol{\lambda}_0 \tag{29}$$

由 (26) 式可得

$$\boldsymbol{\lambda}_N = 0 \tag{30}$$

由 (27) 式可得

$$\boldsymbol{w}_k = Q_k \boldsymbol{\lambda}_{k+1} \tag{31}$$

代入到状态方程 (10)，则有

$$\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + Q_k \boldsymbol{\lambda}_{k+1} \tag{32}$$

令

$$\theta = \frac{1}{\epsilon} \tag{33}$$

由 (28) 式可导出

$$\boldsymbol{\lambda}_k = F_k^{\mathrm{T}} \boldsymbol{\lambda}_{k+1} + \theta \bar{S}_k (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k \boldsymbol{x}_k) \tag{34}$$

考虑到 $\boldsymbol{x}_0 = \hat{\boldsymbol{x}}_0 + P_0 \boldsymbol{\lambda}_0$，我们假定

$$\boldsymbol{x}_k = \boldsymbol{\mu}_k + P_k \boldsymbol{\lambda}_k \tag{35}$$

式中，$\boldsymbol{\mu}_k$ 和 P_k 待定，但它们的初值分别由初始估计 $\boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0$ 和权重矩阵 P_0 决定。将 (35) 代入 (32)，可以得到

$$\boldsymbol{\mu}_{k+1} + P_{k+1}\boldsymbol{\lambda}_{k+1} = F_k\boldsymbol{\mu}_k + F_kP_k\boldsymbol{\lambda}_k + Q_k\boldsymbol{\lambda}_{k+1} \tag{36}$$

将 (35) 代入 (34)，可得

$$\boldsymbol{\lambda}_k = F_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1} + \theta\bar{S}_k(\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}}R_k^{-1}[\boldsymbol{y}_k - H_k(\boldsymbol{\mu}_k + P_k\lambda_k)] \tag{37}$$

上式可以进一步整理为

$$\begin{aligned}\boldsymbol{\lambda}_k =& [I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k]^{-1} \\ & \times [F_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1} + \theta\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)]\end{aligned}$$

将上式代入 (36)，可导出

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

由上式，可导出

$$\begin{gathered}\boldsymbol{\mu}_{k+1} - F_k\boldsymbol{\mu}_k - F_kP_k[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k]^{-1}\times \\ \times[\theta\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)] = \\ [-P_{k+1} + F_kP_k[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k]^{-1}F_k^{\mathrm{T}} + Q_k]\boldsymbol{\lambda}_{k+1} \qquad (38)\end{gathered}$$

在假设的解式 (35) 中，我们未对 $\boldsymbol{\mu}_k$ 和 P_k 的取值引入任何限制。式 (38) 等式两边都等于 0，(38) 仍然成立。

令 (38) 等式左边为 0，可导出

$$\boldsymbol{\mu}_{k+1} = F_k\boldsymbol{\mu}_k + F_kP_k[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k]^{-1}\times \tag{39}$$

$$\times\,[\theta\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)] \tag{40}$$

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

$$\boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0 \tag{41}$$

令 (38) 等式右边为 0, 可导出

$$P_{k+1} = F_kP_k[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k]^{-1}F_k^{\mathrm{T}} + Q_k = F_k\tilde{P}_kF_k^{\mathrm{T}} + Q_k \tag{42}$$

这是关于 P_k 的演化方程，式中 $\tilde{P}_k$ 的定义为

$$\tilde{P}_k = P_k[I - \theta\bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k]^{-1} = [P_k^{-1} - \theta\bar{S}_k + H_k^{\mathrm{T}} R_k^{-1} H_k]^{-1} \quad (43)$$

公式 (43) 表明，如果 P_k、S_k 和 R_k 均为对称矩阵，那么 $\tilde{P}_k$ 也是对称矩阵。从 (42) 可以看出，如果 Q_k 也是对称矩阵，则 P_{k+1} 对称。所以如果 P_0、Q_k、R_k 和 S_k 均为对称矩阵，那么 $\tilde{P}_k$ 和 P_k 也是对称的。

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

$$\boldsymbol{x}_0 = \hat{\boldsymbol{x}}_0 + P_0\boldsymbol{\lambda}_0 \tag{44}$$

$$\boldsymbol{w}_k = Q_k\boldsymbol{\lambda}_{k+1} \tag{45}$$

$$\begin{aligned}\boldsymbol{\lambda}_k =& [I - \theta\bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k]^{-1} \times \\ & \times [F_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1} + \theta\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}} R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)], \quad \boldsymbol{\lambda}_N = 0\end{aligned} \tag{46}$$

$$P_{k+1} = F_k P_k [I - \theta\bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k]^{-1} F_k^{\mathrm{T}} + Q_k, \quad P_0 = P_0 \tag{47}$$

$$\begin{aligned}\boldsymbol{\mu}_{k+1} =& F_k\boldsymbol{\mu}_k + F_k P_k [I - \theta\bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k]^{-1} \times \\ & \times [\theta\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + H_k^{\mathrm{T}} R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)], \quad \boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0\end{aligned} \tag{48}$$

2.2 关于 $\hat{x}_k$ 和 y_k 的驻点

假定 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 已经取得最大值的前提下，本小节的问题是寻找 $J = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$（满足约束条件 $\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k$）关于 $\hat{\boldsymbol{x}}_k$ 和 $\boldsymbol{y}_k$ 的驻点。根据 (35) 和初始条件 (41)，可以得到

$$\boldsymbol{\lambda}_k = P_k^{-1}(\boldsymbol{x}_k - \boldsymbol{\mu}_k) \tag{49}$$

$$\boldsymbol{\lambda}_0 = P_0^{-1}(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0) \tag{50}$$

由此可知

$$\begin{aligned}
\|\boldsymbol{\lambda}_0\|_{P_0}^2 &= \boldsymbol{\lambda}_0^{\mathrm{T}} P_0 \boldsymbol{\lambda}_0 \\
&= (\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}} P_0^{-T} P_0 P_0^{-1} (\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0) \\
&= (\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}} P_0^{-1} (\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0) \\
&= \|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|_{P_0^{-1}}^2
\end{aligned}$$

因此，性能指标 (23) 可以改写为

$$J = -\epsilon \|\boldsymbol{\lambda}_0\|_{P_0}^2 + \sum_{k=0}^{N-1} \left[\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|_{\bar{S}_k}^2 - \epsilon \left(\|\boldsymbol{w}_k\|_{Q_k^{-1}}^2 + \|\boldsymbol{y}_k - H_k \boldsymbol{x}_k\|_{R_k^{-1}}^2 \right) \right] \quad (51)$$

将 (35) 代入上式，有

$$J = -\epsilon\|\boldsymbol{\lambda}_0\|^2_{P_0} + \tag{52}$$

$$+\sum_{k=0}^{N-1}\left[\|\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|^2_{\bar{S}_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{Q_k^{-1}} + \|\boldsymbol{y}_k - H_k(\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k)\|^2_{R_k^{-1}})\right] \tag{53}$$

将 (31) 式代入 $\|\boldsymbol{w}_k\|^2_{Q_k^{-1}}$，得到

$$\|\boldsymbol{w}_k\|^2_{Q_k^{-1}} = \boldsymbol{\lambda}^{\mathrm{T}}_{k+1}Q_k^{\mathrm{T}}Q_k^{-1}Q_k\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}^{\mathrm{T}}_{k+1}Q_k\boldsymbol{\lambda}_{k+1} \tag{54}$$

于是，式 (53) 可以重写为

$$
\begin{aligned}
J =& -\epsilon\|\boldsymbol{\lambda}_0\|_{P_0}^2 - \epsilon\sum_{k=0}^{N-1}\|\boldsymbol{\lambda}_{k+1}\|_{Q_k}^2 + \\
&+ \sum_{k=0}^{N-1}\left[\|\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|_{\bar{S}_k}^2 - \epsilon\|\boldsymbol{y}_k - H_k(\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k)\|_{R_k^{-1}}^2\right]
\end{aligned}
\tag{55}
$$

由于 $\boldsymbol{\lambda}_N = 0$，可以得到如下等式

$$
\sum_{k=0}^{N}\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\boldsymbol{\lambda}_k - \sum_{k=0}^{N-1}\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\boldsymbol{\lambda}_k = 0 \tag{56}
$$

上式可以改写为

$$
\begin{aligned}
0 &= \boldsymbol{\lambda}_0^{\mathrm{T}} P_0 \boldsymbol{\lambda}_0 + \sum_{k=1}^{N} \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \boldsymbol{\lambda}_k - \sum_{k=0}^{N-1} \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \boldsymbol{\lambda}_k \\
&= \boldsymbol{\lambda}_0^{\mathrm{T}} P_0 \boldsymbol{\lambda}_0 + \sum_{k=0}^{N-1} \boldsymbol{\lambda}_{k+1}^{\mathrm{T}} P_{k+1} \boldsymbol{\lambda}_{k+1} - \sum_{k=0}^{N-1} \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \boldsymbol{\lambda}_k \\
&= -\epsilon \|\boldsymbol{\lambda}_0\|_{P_0}^2 - \epsilon \sum_{k=0}^{N-1} (\boldsymbol{\lambda}_{k+1}^{\mathrm{T}} P_{k+1} \boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \boldsymbol{\lambda}_k)
\end{aligned}
$$

性能指标 (55) 左右端分别减去上式左右端，有

$$
\begin{aligned}
J =& \sum_{k=0}^{N-1}[\|\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|^2_{\bar{S}_k} - \epsilon\|\boldsymbol{y}_k - H_k(\boldsymbol{\mu}_k + P_k\boldsymbol{\lambda}_k)\|^2_{R_k^{-1}}] \\
& - \epsilon\sum_{k=0}^{N-1}\|\boldsymbol{\lambda}_{k+1}\|^2_{Q_k} + \epsilon\sum_{k=0}^{N-1}(\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}P_{k+1}\boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k^{\mathrm{T}}P_k\boldsymbol{\lambda}_k) \\
=& \sum_{k=0}^{N-1}[(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + 2(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_kP_k\boldsymbol{\lambda}_k + \boldsymbol{\lambda}_k^{\mathrm{T}}P_k\bar{S}_kP_k\boldsymbol{\lambda}_k \\
& + \epsilon\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1} - Q_k)\boldsymbol{\lambda}_{k+1} - \epsilon\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\boldsymbol{\lambda}_k - \epsilon(\boldsymbol{\lambda}_k - H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k) \\
& + 2\epsilon(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_kP_k\boldsymbol{\lambda}_k - \epsilon\boldsymbol{\lambda}_k^{\mathrm{T}}P_kH_k^{\mathrm{T}}R_k^{-1}H_kP_k\boldsymbol{\lambda}_k]
\end{aligned}
\tag{57}
$$

考虑上式中的 $\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1} - Q_k)\boldsymbol{\lambda}_{k+1}$ 项，将 P_{k+1} 表达式 (42) 代入其中，可

得

$$\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1}-Q_k)\boldsymbol{\lambda}_{k+1}=\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(Q_k+F_k\tilde{P}_kF_k^{\mathrm{T}}-Q_k)\boldsymbol{\lambda}_{k+1}=\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}F_k\tilde{P}_kF_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1}$$

根据 (37) 式，有

$$F_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1}=\boldsymbol{\lambda}_k-\theta\bar{S}_k(\boldsymbol{\mu}_k+P_k\boldsymbol{\lambda}_k-\hat{\boldsymbol{x}}_k)+H_k^{\mathrm{T}}R_k^{-1}[\boldsymbol{y}_k-H_k(\boldsymbol{\mu}_k+P_k\boldsymbol{\lambda}_k)] \tag{58}$$

因此

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

因为 $(I-\theta P_k\bar{S}_k+P_kH_k^{\mathrm{T}}R_k^{-1}H_k)=P_k\tilde{P}_k^{-1}$，代入上式可得

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1}-Q_k)\boldsymbol{\lambda}_{k+1}\\
&=\{\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\tilde{P}_k^{-1}-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k-(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_k\}\\
&\quad\times\tilde{P}_k\{\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\tilde{P}_k^{-1}-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k-(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_k\}^{\mathrm{T}}\\
&=\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\tilde{P}_k^{-1}P_k\boldsymbol{\lambda}_k-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_kP_k\boldsymbol{\lambda}_k-(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_kP_k\boldsymbol{\lambda}_k\\
&\quad-\theta\boldsymbol{\lambda}_kP_k\bar{S}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\theta^2(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k\tilde{P}_k\bar{S}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)\\
&\quad+\theta(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_k\tilde{P}_k\bar{S}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)-\boldsymbol{\lambda}_k^{\mathrm{T}}P_kH_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)\\
&\quad+\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k\tilde{P}_kH_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)\\
&\quad+(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_k\tilde{P}_kH_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)
\end{aligned}
$$

这里需要注意的是，上式的结果是一个标量，即等式右边的每一项均为标量。由于是标量，所以有

$$\theta(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k P_k \boldsymbol{\lambda}_k = \theta \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)$$

因此，上面等式可以重写为

$$\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1} - Q_k)\boldsymbol{\lambda}_{k+1} \\
&= \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \tilde{P}_k^{-1} P_k \boldsymbol{\lambda}_k - 2\theta(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k P_k \boldsymbol{\lambda}_k \\
&\quad - 2(\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)^{\mathrm{T}} R_k^{-1} H_k P_k \boldsymbol{\lambda}_k + \theta^2(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k \tilde{P}_k \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) \\
&\quad + 2\theta(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k) \\
&\quad + (\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k)^{\mathrm{T}} R_k^{-1} H_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k\boldsymbol{\mu}_k) \qquad (59)
\end{aligned}$$

从 (43) 式可以得到

$$\tilde{P}_k^{-1} = [I - \theta\bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k] P_k^{-1} = P_k^{-1}[I - \theta P_k \bar{S}_k + P_k H_k^{\mathrm{T}} R_k^{-1} H_k]$$

因此

$$\begin{aligned}\boldsymbol{\lambda}_k^{\mathrm{T}} P_k \tilde{P}_k^{-1} P_k \boldsymbol{\lambda}_k =& \boldsymbol{\lambda}_k^{\mathrm{T}}[I - \theta P_k \bar{S}_k + P_k H_k^{\mathrm{T}} R_k^{-1} H_k] P_k \boldsymbol{\lambda}_k \\ =& \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \boldsymbol{\lambda}_k - \theta \boldsymbol{\lambda}_k^{\mathrm{T}} P_k \bar{S}_k P_k \boldsymbol{\lambda}_k + \boldsymbol{\lambda}_k^{\mathrm{T}} P_k H_k^{\mathrm{T}} R_k^{-1} H_k P_k \boldsymbol{\lambda}_k\end{aligned}$$

将上式代入 (59) 式中，得到

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^{\mathrm{T}}(P_{k+1}-Q_k)\boldsymbol{\lambda}_{k+1}\\
&\quad=\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\boldsymbol{\lambda}_k-\theta\boldsymbol{\lambda}_k^{\mathrm{T}}P_k\bar{S}_kP_k\boldsymbol{\lambda}_k+\boldsymbol{\lambda}_k^{\mathrm{T}}P_kH_k^{\mathrm{T}}R_k^{-1}H_kP_k\boldsymbol{\lambda}_k-2\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_kP_k\boldsymbol{\lambda}_k\\
&\qquad-2(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_kP_k\boldsymbol{\lambda}_k+\theta^2(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k\tilde{P}_k\bar{S}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)\\
&\qquad+2\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^{\mathrm{T}}\bar{S}_k\tilde{P}_kH_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)\\
&\qquad+(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)^{\mathrm{T}}R_k^{-1}H_k\tilde{P}_kH_k^{\mathrm{T}}R_k^{-1}(\boldsymbol{y}_k-H_k\boldsymbol{\mu}_k)
\end{aligned}
$$

将上式代入 (57) 式，有

$$
\begin{aligned}
J = \sum_{k=0}^{N-1} \Big[& (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) - \epsilon (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k)^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) \\
& + \theta (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k \tilde{P}_k \bar{S}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + 2 (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) \\
& + \epsilon (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k)^{\mathrm{T}} R_k^{-1} H_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) \Big] \\
= \sum_{k=0}^{N-1} \Big[& (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} (\bar{S}_k + \theta \bar{S}_k \tilde{P}_k \bar{S}_k) (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) \\
& + 2 (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}} \bar{S}_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) \\
& + \epsilon (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k)^{\mathrm{T}} (R_k^{-1} H_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} - R_k^{-1}) (\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) \Big]
\end{aligned}
$$

现在我们来回顾本小节的目标：找出 J 关于 $\hat{\boldsymbol{x}}_k$ 和 $\boldsymbol{y}_k$ 的驻点。利用上

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

$$\frac{\partial J}{\partial \hat{\boldsymbol{x}}_k} = 2(\bar{S}_k + \theta \bar{S}_k \tilde{P}_k \bar{S}_k)(\hat{\boldsymbol{x}}_k - \boldsymbol{\mu}_k) + 2\bar{S}_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1}(H_k \boldsymbol{\mu}_k - \boldsymbol{y}_k) = 0 \tag{60}$$

$$\frac{\partial J}{\partial \boldsymbol{y}_k} = 2\epsilon(R_k^{-1} H_k \tilde{P}_k H_k^{\mathrm{T}} R_k^{-1} - R_k^{-1})(\boldsymbol{y}_k - H_k \boldsymbol{\mu}_k) + 2R_k^{-1} H_k \tilde{P}_k \bar{S}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) = 0 \tag{61}$$

当 $\hat{\boldsymbol{x}}_k = \boldsymbol{\mu}_k$ 以及 $\boldsymbol{y}_k = H_k \boldsymbol{\mu}_k$ 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

$$\frac{\partial^2 J}{\partial \hat{\boldsymbol{x}}_k^2} = 2(\bar{S}_k + \theta \bar{S}_k \tilde{P}_k \bar{S}_k) \tag{62}$$

如果 $(\bar{S}_k+\theta\bar{S}_k\tilde{P}_k\bar{S}_k)>0$，那么 $\hat{\boldsymbol{x}}_k$ 为极小值点。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 (43) 可知，$\hat{\boldsymbol{x}}_k$ 为极小值点的条件为

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

等价于

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

其中，P_k^{-1}、$\theta\bar{S}_k$ 和 $H_k^{\mathrm{T}}R_k^{-1}H_k$ 均正定。因此，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，可采用如下 3 种方式：

1）$\theta = 1/\epsilon$ 变小可以使 $\theta\bar{S}_k$ 变小，这意味着 (15) 或 (16) 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。这个结论是根据式中 $\bar{S}_k$ 和 L_k 的关系得出的。目标函数的分子为

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

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。

2.3 H_∞ 鲁棒滤波算法小结

总结以上两小节的结果，我们即可建立所谓的 H_∞ 鲁棒滤波算法。

系统模型

考虑如下动态系统及输出方程

$$\begin{cases} \boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k \\ \boldsymbol{y}_k = H_k \boldsymbol{x}_k + \boldsymbol{v}_k \\ \boldsymbol{z}_k = L_k \boldsymbol{x}_k \end{cases} \tag{63}$$

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

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

$$J_1 = \frac{\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{S_k}^2}{\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1} \left(\|\boldsymbol{w}_k\|_{Q_k^{-1}}^2 + \|\boldsymbol{v}_k\|_{R_k^{-1}}^2\right)} \tag{64}$$

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

将 $\hat{\boldsymbol{x}}_k = \boldsymbol{\mu}_k$ 代入 (48)，并令

$$K_k = P_k\left[I - \theta \bar{S}_k P_k + H_k^{\mathrm{T}} R_k^{-1} H_k P_k\right]^{-1} H_k^{\mathrm{T}} R_k^{-1}$$

可得

$$\hat{\boldsymbol{x}}_{k+1} = F_k \hat{\boldsymbol{x}}_k + F_k K_k (\boldsymbol{y}_k - H_k \hat{\boldsymbol{x}}_k), \quad \hat{\boldsymbol{x}}_0 = \hat{\boldsymbol{x}}_0$$

考虑到 P_k 的计算公式 (47) 以及 $\bar{\boldsymbol{S}}_k$ 的计算公式 (22)，如下称为 H_∞ 滤波的递推滤波算法使目标函数 J_1 小于 ϵ，其中 $\theta = 1/\epsilon$。

$$\begin{cases} \hat{\boldsymbol{x}}_{k+1} = F_k\hat{\boldsymbol{x}}_k + F_kK_k(\boldsymbol{y}_k - H_k\hat{\boldsymbol{x}}_k), \quad \hat{\boldsymbol{x}}_0 = \hat{\boldsymbol{x}}_0 \\ P_{k+1} = F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k\right]^{-1}F_k^{\mathrm{T}} + Q_k, \quad P_0 = P_0 \end{cases} \tag{65}$$

式中

$$\begin{cases} K_k = P_k\left[I - \theta\bar{S}_kP_k + H_k^{\mathrm{T}}R_k^{-1}H_kP_k\right]^{-1}H_k^{\mathrm{T}}R_k^{-1} \\ \bar{S}_k = L_k^{\mathrm{T}}S_kL_k \end{cases} \tag{66}$$

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 $P_k^{-1} - \theta\bar{S}_k + H_k^{\mathrm{T}}R_k^{-1}H_k > 0$ 成立。

3. H_∞ 滤波和卡尔曼滤波的区别和联系

为了比较 H_∞ 滤波与卡尔曼滤波，这里我们简要回顾卡尔曼滤波器的另外一种形式——基于一步预测的卡尔曼滤波算法。仍考虑 (10) 和 (11) 描述的线性动态系统，假设 $\boldsymbol{w}_k$ 和 $\boldsymbol{v}_{k+1}$ 为不相关的零均值高斯分布白噪声，协方差分别为 Q_k 和 R_{k+1}。

根据卡尔曼滤波理论，我们有

$$\hat{\boldsymbol{x}}_{k+1|k} = F_k \hat{\boldsymbol{x}}_{k|k} \tag{67}$$

$$P_{k+1|k} = F_k P_{k|k} F_k^{\mathrm{T}} + Q_k \tag{68}$$

$$
\begin{aligned}
\hat{\boldsymbol{x}}_{k+1|k+1} &= \hat{\boldsymbol{x}}_{k+1|k} + K_{k+1}(\boldsymbol{y}_{k+1} - H_{k+1}\hat{\boldsymbol{x}}_{k+1|k}) \qquad (69) \\
&= P_{k+1k+1}[P_{k+1|k}^{-1}\hat{\boldsymbol{x}}_{k+1|k} + H_{k+1}^{\mathrm{T}} R_{k+1}^{-1}\boldsymbol{y}_{k+1}]
\end{aligned}
$$

其中

$$
P_{k+1k+1} = [P_{k+1|k}^{-1} + H_{k+1}^{\mathrm{T}} R_{k+1}^{-1} H_{k+1}]^{-1} \qquad (70)
$$

$$
\begin{aligned}
K_{k+1} &= [P_{k+1|k}^{-1} + H_{k+1}^{\mathrm{T}} R_{k+1}^{-1} H_{k+1}]^{-1} H_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \\
&= P_{k+1|k}[I + H_{k+1}^{\mathrm{T}} R_{k+1}^{-1} H_{k+1} P_{k+1|k}]^{-1} H_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \qquad (71)
\end{aligned}
$$

将 (70) 代入 (68)，可知

$$\begin{aligned} P_{k+1|k} &= F_k P_{k|k} F_k^{\mathrm{T}} + Q_k \\ &= F_k [P_{k|k-1}^{-1} + H_k^{\mathrm{T}} R_k^{-1} H_k]^{-1} F_k^{\mathrm{T}} + Q_k \\ &= F_k P_{k|k-1} [I + H_k^{\mathrm{T}} R_k^{-1} H_k P_{k|k-1}]^{-1} F_k^{\mathrm{T}} + Q_k \end{aligned} \tag{72}$$

由 (67) 和 (69) 有

$$\hat{\boldsymbol{x}}_{k+1|k} = F_k \hat{\boldsymbol{x}}_{k|k} = F_k \hat{\boldsymbol{x}}_{k|k-1} + F_k K_k (\boldsymbol{y}_k - H_k \hat{\boldsymbol{x}}_{k|k-1}) \tag{73}$$

综上，一步预测形式的卡尔曼滤波方程可表示为

$$\hat{\boldsymbol{x}}_{k+1|k} = F_k \hat{\boldsymbol{x}}_{k|k-1} + F_k K_k (\boldsymbol{y}_k - H_k \hat{\boldsymbol{x}}_{k|k-1}) \tag{74}$$

$$K_k = P_{k|k-1} \left(I + H_k^{\mathrm{T}} R_k^{-1} H_k P_{k|k-1}\right)^{-1} H_k^{\mathrm{T}} R_k^{-1} \tag{75}$$

$$P_{k+1|k} = F_{k-1} P_{k|k-1} \left(I + H_k^{\mathrm{T}} R_k^{-1} H_k P_{k|k-1}\right)^{-1} F_k^{\mathrm{T}} + Q_k \tag{76}$$

我们知道，卡尔曼滤波器除了需要知道知状态转移矩阵 F_k 和量测矩阵 H_k 外，需要知道过程噪声 $\boldsymbol{w}_k$ 和量测噪声 $\boldsymbol{v}_k$ 在每一时刻的均值（标准算法中假设为 0）以及协方差矩阵 Q_k 和 R_k。只有当噪声服从高斯分布时，卡尔曼滤波器是最小方差估计器；当噪声非高斯分布时，卡尔曼滤波器是线性最小方差估计器。

对比分析 (65)、(66) 和 (74)~(76)，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：

1. 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 $\boldsymbol{w}_k$、量测扰动 $\boldsymbol{v}_k$ 和初始估计误差 $(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)$ 幅度的先验信息事先设置；而在卡尔曼滤波算法中，认为 $\boldsymbol{w}_k$、$\boldsymbol{v}_k$ 和 $(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)$ 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。

2. 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 $\epsilon = +\infty(\theta = 0)$，$H_\infty$ 滤波变为卡尔曼滤波。由此可以重新定义卡尔曼滤波，即卡尔曼滤波为式 (64) 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。

3. 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。也就是说，如果我们利用卡尔曼滤波估计 $L_k\boldsymbol{x}_k$，那么结果与我们选择的 L_k 矩阵无关。然而，当我们使用 H_∞ 滤波时，结果与 L_k 和我们想要估计的状态组合有很大关系。

4. 观察 (65) 和 (66) 式，当去掉 K_k 和 P_{k+1} 方程中的 $\theta\bar{S}_kP_k$ 项后，H_∞ 滤波算法形式上变为卡尔曼滤波完全一样。

5. 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。

6. 在 H_∞ 滤波中，从 (65) 式可以看出，P_{k+1} 等式右边 $(-\theta\bar{S}_kP_k)$ 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 3.1 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1，均方根误差 *(RSME)* 结果比较见图2。不同 θ 取值下 H_∞ 滤波平均均方根误差结果比较如表1所示，表中也给出了标准卡尔曼滤波的平均均方根误差。

Table 1: 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

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

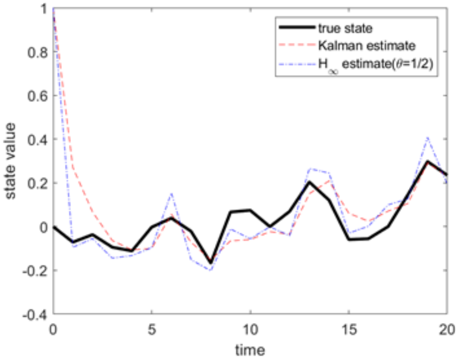


Figure 1: 标称条件滤波结果比较

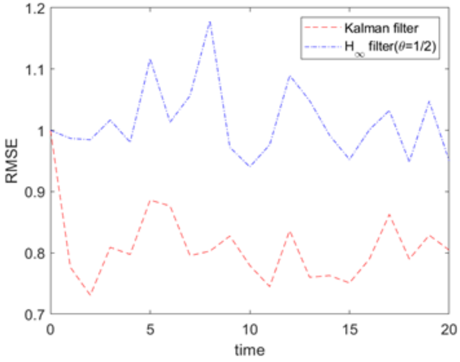


Figure 2: 标称条件均方根误差比较

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

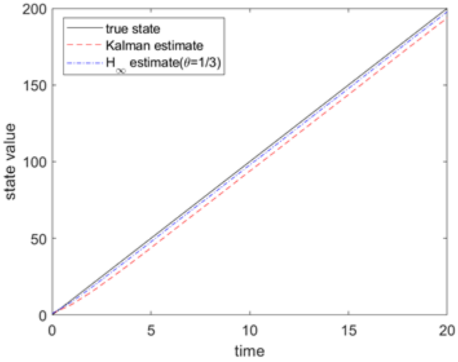
图3，均方根误差 (RSME) 结果比较见图4。不同 θ 取值下 H_∞ 滤波平均均方根误差结果比较如表2所示，表中也给出了标准卡尔曼滤波的平均均方根误差。

Table 2: 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

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；



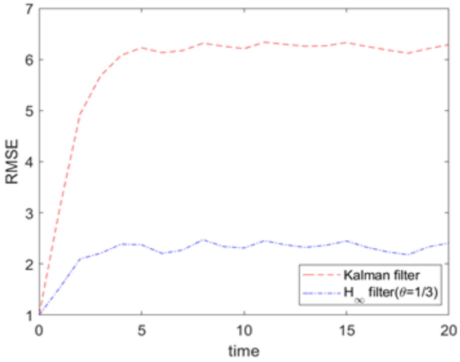


Figure 3: 过程噪声统计特性未知滤波结果比较

Figure 4: 过程噪声统计特性未知均方根误差比较

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。

$\square$

References (参考文献)

[1] Simon D. Optimal state estimation: Kalman, Hinfinity, and nonlinear approaches[M]. USA: John Wiley & Sons,2006.

[2] Lewis F L, Xie L, Popa D. Optimal and robust estimation: with an introduction to stochastic control theory[M]. USA: CRC press, 2017.

[3] 韩崇昭, 朱洪艳, 段战胜, 等. 多源信息融合 [M]. 北京: 清华大学出版社, 2010.

[4] Särkkä S. 贝叶斯滤波与平滑 [J]. 北京: 国防工业出版社， 2015.

[5] Anderson B D O, Moore J B. Optimal filtering[M]. USA: Courier Corporation, 2012.

[6] Kalman R E. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering, 1960, 82(1): 35–45.

[7] Reif K, Gunther S, Yaz E, et al. Stochastic stability of the discrete-time extended Kalman filter[J]. IEEE Transactions on Automatic Control, 1999, 44(4): 714-728.

[8] Psiaki M L. Backward-smoothing extended Kalman filter[J]. Journal of Guidance, Control, and Dynamics, 2005, 28(5): 885-894.

[9] 王小旭, 潘泉, 黄鹤, 等. 非线性系统确定采样型滤波算法综述 [J]. 控制与决策, 2012, 27(6): 801-812.

[10] Julier S, Uhlmann J, Durrant-Whyte H F. A new method for the nonlinear transformation of means and covariances in filters and estimators[J]. IEEE

Transactions on Automatic Control, 2000, 45(3):477-482.

[11] Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation[J]. Proceedings of the IEEE, 2004, 92(3):401-422.

[12] Kandepu R, Foss B, Imsland L. Applying the unscented Kalman filter for nonlinear state estimation[J]. Journal of Process Control, 2008, 18(7-8):753-768.

[13] Zhan R, Wan J. Iterated unscented Kalman filter for passive target tracking[J]. IEEE Transactions on Aerospace and Electronic Systems, 2007, 43(3): 1155-1163.

[14] 廖瑛, 刘光明, 文援兰, 等. 空间非合作目标被动跟踪技术与应用 [M]. 北京: 国防工业出版社, 2015.

[15] Arasaratnam I, Haykin S. Cubature Kalman filters[J]. IEEE Transactions on Automatic Control, 2009, 54(6): 1254-1269.

[16] Arasaratnam I, Haykin S. Cubature Kalman smoothers[J]. Automatica, 2011, 47(10): 2245-2250.

[17] Jia B, Xin M, Cheng Y. High-degree cubature Kalman filter[J]. Automatica, 2013, 49(2): 510-518.

[18] Arasaratnam I, Haykin S, Hurd T R. Cubature Kalman filtering for continuous-discrete systems:theory and simulations[J]. IEEE Transactions on Signal Processing,2010, 58(10): 4977-4993.

[19] Arulampalam M S, Maskell S, Gordon N, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions

on Signal Processing, 2002, 50(2): 174-188.

[20] 胡士强, 敬忠良. 粒子滤波算法综述 [J]. 控制与决策, 2005, 20(4): 361-365.

[21] Carpenter J, Clifford P. Improved particle filter for nonlinear problems[J] . IEE Proceedings-Radar, Sonar and Navigation, 1999, 146(1): 2-7.

[22] Handschin J E. Monte Carlo techniques for prediction and filtering of non-linear stochastic processes[J]. Automatica, 1970, 6(4): 555-563.

[23] Dunik J, Straka O, Simandl M, et al. Random-point-based filters: Analysis and comparison in target tracking[J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51(2): 1403-1421.

[24] Alspach D, Sorenson H. Nonlinear Bayesian estimation using Gaussian sum

approximations[J]. IEEE Transactions on Automatic Control, 1972, 17(4): 439-448.

[25] Kotecha J H, Djuric P M. Gaussian sum particle filtering[J]. IEEE Transactions on Signal Processing, 2003, 51(10): 2602-2612.

[26] Gandhi M A, Mili L. Robust Kalman filter based on a generalized maximum-likelihood-type estimator[J]. IEEE Transactions on Signal Processing, 2009, 58(5): 2509-2520.

[27] Xie L, Soh Y C, De Souza C E. Robust Kalman filtering for uncertain discrete-time systems[J]. IEEE Transactions on Automatic Control, 1994, 39(6): 1310-1314.

[28] Chang L, Hu B, Chang G, et al. Huber-based novel robust unscented

Kalman filter[J]. IET Science, Measurement & Technology, 2012, 6(6): 502-509.

[29] Aidala V, Hammel S. Utilization of modified polar coordinates for bearings-only tracking[J]. IEEE Transactions on Automatic Control, 1983, 28(3): 283-294.

[30] Peach N. Bearings-only tracking using a set of range-parameterised extended Kalman filters[J]. IEE Proceedings-Control Theory and Applications, 1995, 142(1): 73-80.

[31] Li X R, Jilkov V P. Survey of maneuvering target tracking. Part I. Dynamic models[J]. IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(4): 1333-1364.

[32] Li X R, Jilkov V P. Survey of maneuvering target tracking. Part II: Motion models of ballistic and space targets[J]. IEEE Transactions on Aerospace and Electronic Systems, 2010, 46(1): 96-119.

[33] Li X R, Jilkov V P. Survey of maneuvering target tracking. Part V. Multiple-model methods[J]. IEEE Transactions on Aerospace and Electronic Systems, 2005, 41(4): 1255-1321.

$\square$

Questions?