鲁棒 H_∞ 滤波

Prof. Yuan-Li Cai

Spring 2024

0. Outline

当系统模型及噪声统计特性存在大的不确定性或未知时，标准的 Kalman 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_∞ 滤波的基本原理，并讨论、分析与 Kalman 滤波之间的区别和联系。

1. 动态约束优化

考虑如下动态系统:

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

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:

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

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：

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

上式可以改写整理为

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

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 Hamilton 函数：

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

增广目标函数可以重写为

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

显然，最优解的必要条件（驻点）如下：

$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{x}_k} = 0 \ (k = 0, \cdots, N)$$
$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{w}_k} = 0 \ (k = 0, \cdots, N-1)$$
$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{\lambda}_k} = 0 \ (k = 0, \cdots, N)$$

进一步地可以写为

$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{x}_0} = 0 \tag{5}$$

$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{x}_N} = 0 \tag{6}$$

$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{x}_k} = 0 \ (k = 1, \cdots, N-1) \tag{7}$$

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

$$\frac{\partial \boldsymbol{J}_a}{\partial \boldsymbol{\lambda}_k} = 0 \ (k = 0, \cdots, N) \tag{9}$$

即

$$\boldsymbol{\lambda}_0 + \frac{\partial \psi_0}{\partial \boldsymbol{x}_0} = 0 \tag{10}$$

$$\boldsymbol{\lambda}_N = 0 \tag{11}$$

$$\boldsymbol{\lambda}_k = \frac{\partial \mathcal{H}_k}{\partial \boldsymbol{x}_k} \quad (k = 0, 1, \cdots, N-1) \tag{12}$$

$$\frac{\partial \mathcal{H}_k}{\partial \boldsymbol{w}_k} = 0 \quad (k = 0, 1, \cdots, N-1) \tag{13}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k \quad (k = 0, 1, \cdots, N-1) \tag{14}$$

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

$$\{\boldsymbol{x}_k^*|k=0,1,2,\cdots,N\} \tag{15}$$

$$\{\boldsymbol{w}_k^*|k=0,1,2,\cdots,N-1\} \tag{16}$$

$$\{\boldsymbol{\lambda}_k|k=0,1,2,\cdots,N\} \tag{17}$$

不能发现，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:

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k \tag{18}$$

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k \tag{19}$$

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

$$\boldsymbol{z}_k = \boldsymbol{L}_k \boldsymbol{x}_k \tag{20}$$

其中，$\boldsymbol{L}_k$ 是自定义的矩阵（假设 $\boldsymbol{L}_k$ 满秩）。如果我们想要直接估计 $\boldsymbol{x}_k$，设置 $\boldsymbol{L}_k = \boldsymbol{I}$。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。

设 $\boldsymbol{z_k}$ 的估计量表示为 $\hat{\boldsymbol{z}}_k$，我们的目的是利用所有可能的信息获得使 $\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{\boldsymbol{S}_k}^2$ 尽可能小的估计 $\hat{\boldsymbol{z}}_k$，建立类似 Kalman 滤波那样的递推算法。其中，$\boldsymbol{S}_k$ 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 $\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{\boldsymbol{S}_k}^2$ 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:

$$\boldsymbol{J}_1 = \frac{\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{\boldsymbol{S}_k}}{\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{k=0}^{N-1} \left(\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}}\right)} \tag{21}$$

式中，$\boldsymbol{P}_0$、$\boldsymbol{Q}_k$、$\boldsymbol{R}_k$ 和 $\boldsymbol{S}_k$ 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：

$$\boldsymbol{J}_1^* = \min_{\hat{\boldsymbol{z}}_k} \max_{\boldsymbol{w}_k,\boldsymbol{v}_k,\boldsymbol{x}_0} \boldsymbol{J}_1 \qquad (*)$$

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 $\boldsymbol{S}_k$。如果在卡尔曼滤波中，最小化 $\boldsymbol{S}_k$ 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，$\boldsymbol{S}_k$ 矩阵的取值会影响滤波增益。

直接求解上述 min-max 问题 (*) 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 $\hat{\boldsymbol{z}}_k$，同时使得

$$\boldsymbol{J}_1 < \epsilon \tag{22}$$

式中，$\epsilon(> 0)$ 为自定义的性能边界参数。结合 $\boldsymbol{J}_1$ 的表达式，上式可以重写为

$$\boldsymbol{J} = -\epsilon\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{k=0}^{N-1}\left[\|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{\boldsymbol{S}_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}})\right] < 1 \tag{23}$$

于是，我们的问题转化为如下 min-max 问题：

$$\boldsymbol{J}^* = \min_{\hat{\boldsymbol{z}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{v}_k, \boldsymbol{x}_0} \boldsymbol{J} \tag{24}$$

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

$$\boldsymbol{J}^* = \min_{\hat{\boldsymbol{x}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{v}_k, \boldsymbol{x}_0} \boldsymbol{J} \tag{25}$$

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 $\boldsymbol{J}$。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

$$\boldsymbol{J}^* = \min_{\hat{\boldsymbol{x}}_k} \max_{\boldsymbol{w}_k, \boldsymbol{y}_k, \boldsymbol{x}_0} \boldsymbol{J} \tag{26}$$

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

$$\|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}} = \|\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k\|^2_{\boldsymbol{R}_k^{-1}} \tag{27}$$

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

$$\begin{aligned}\|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{\boldsymbol{S}_k} &= (\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^{\mathsf{T}}\boldsymbol{S}_k(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)\\ &= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathsf{T}}\boldsymbol{L}_k^{\mathsf{T}}\boldsymbol{S}_k\boldsymbol{L}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)\\ &= \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|^2_{\overline{\boldsymbol{S}}_k}\end{aligned} \tag{28}$$

式中，$\overline{\boldsymbol{S}}_k$ 定义为

$$\overline{\boldsymbol{S}}_k = \boldsymbol{L}_k^{\mathsf{T}}\boldsymbol{S}_k\boldsymbol{L}_k \tag{29}$$

综合上述讨论，我们可导出

$$\begin{aligned}\boldsymbol{J} &= -\epsilon \left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{k=0}^{N-1}\left[\left\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right\|^2_{\boldsymbol{S}_k} - \epsilon\left(\left\|\boldsymbol{w}_k\right\|^2_{\boldsymbol{Q}_k^{-1}} + \left\|\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k\right\|^2_{\boldsymbol{R}_k^{-1}}\right)\right] \\ &= \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1}\mathcal{L}_k \qquad (30)\end{aligned}$$

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 $\boldsymbol{J}$ 关于 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 的驻点，然后求解 $\boldsymbol{J}$ 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 $\boldsymbol{J} = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$（满足约束条件 $\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k$）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 Hamilton 函数定义为

$$\mathcal{H}_k = \mathcal{L}_k + 2\epsilon \boldsymbol{\lambda}_{k+1}^T (\boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k) \tag{31}$$

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，$\boldsymbol{J}$ 关于 $\boldsymbol{x}_0$ 和 $\boldsymbol{w}_k$ 的驻点方程为

$$2\epsilon\boldsymbol{\lambda}_0 + \frac{\partial \psi_0}{\partial \boldsymbol{x}_0} = 0 \tag{32}$$

$$2\epsilon\boldsymbol{\lambda}_N = 0 \tag{33}$$

$$\frac{\partial \mathcal{H}_k}{\partial \boldsymbol{w}_k} = 0 \tag{34}$$

$$2\epsilon\boldsymbol{\lambda}_k = \frac{\partial \mathcal{H}_k}{\partial \boldsymbol{x}_k} \tag{35}$$

由 (32) 式可得

$$\boldsymbol{x}_0 = \hat{\boldsymbol{x}}_0 + \boldsymbol{P}_0\boldsymbol{\lambda}_0 \tag{36}$$

由 (33) 式可得

$$\boldsymbol{\lambda}_N = 0 \tag{37}$$

由 (34) 式可得

$$\boldsymbol{w}_k = \boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1} \tag{38}$$

代入到状态方程 (18)，则有

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k\boldsymbol{x}_k + \boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1} \tag{39}$$

令

$$\theta = \frac{1}{\epsilon} \tag{40}$$

由 (35) 式可导出

$$\boldsymbol{\lambda}_k = \boldsymbol{F}_k^T \boldsymbol{\lambda}_{k+1} + \theta \overline{\boldsymbol{S}}_k (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{x}_k) \tag{41}$$

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

$$\boldsymbol{x}_k = \boldsymbol{\mu}_k + \boldsymbol{P}_k \boldsymbol{\lambda}_k \tag{42}$$

式中，$\boldsymbol{\mu}_k$ 和 $\boldsymbol{P}_k$ 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 $\boldsymbol{P}_0$ 决定。将 (42) 代入 (39)，可以得到

$$\boldsymbol{\mu}_{k+1} + \boldsymbol{P}_{k+1}\boldsymbol{\lambda}_{k+1} = \boldsymbol{F}_k\boldsymbol{\mu}_k + \boldsymbol{F}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k + \boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1} \tag{43}$$

将 (42) 代入 (41)，可得

$$\boldsymbol{\lambda}_k = \boldsymbol{F}_k^T\boldsymbol{\lambda}_{k+1} + \theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k + \boldsymbol{P}_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}[\boldsymbol{y}_k - \boldsymbol{H}_k(\boldsymbol{\mu}_k + \boldsymbol{P}_k\lambda_k)] \tag{44}$$

上式可以进一步整理为

$$\boldsymbol{\lambda}_k = [\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\times$$
$$[\boldsymbol{F}_k^T\boldsymbol{\lambda}_{k+1} + \theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)]$$

将上式代入 (43)，可导出

$$\begin{aligned}\boldsymbol{\mu}_{k+1}+\boldsymbol{P}_{k+1}\boldsymbol{\lambda}_{k+1}&=\boldsymbol{F}_k\boldsymbol{\mu}_k+\boldsymbol{F}_k\boldsymbol{P}_k\left[\boldsymbol{I}-\theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\right]^{-1}\times\\&\left[\boldsymbol{F}_k^T\boldsymbol{\lambda}_{k+1}+\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)\right]+\boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1}\end{aligned}$$

由上式，可导出

$$\begin{aligned}&\boldsymbol{\mu}_{k+1}-\boldsymbol{F}_k\boldsymbol{\mu}_k-\boldsymbol{F}_k\boldsymbol{P}_k\left[\boldsymbol{I}-\theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\right]^{-1}\times\\&\left[\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)\right]=\\&\left[-\boldsymbol{P}_{k+1}+\boldsymbol{F}_k\boldsymbol{P}_k\left[\boldsymbol{I}-\theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\right]^{-1}\boldsymbol{F}_k^T+\boldsymbol{Q}_k\right]\boldsymbol{\lambda}_{k+1}\end{aligned}\tag{45}$$

在假设的解式 (42) 中，我们未对 $\boldsymbol{\mu}_k$ 和 $\boldsymbol{P}_k$ 的取值引入任何限制。式 (45) 两边都等于 0，(45) 仍然成立。

令 (45) 左边为 0，可导出

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{F}_k\boldsymbol{\mu}_k+\boldsymbol{F}_k\boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\times [\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)] \tag{46}$$

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

$$\boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0 \tag{47}$$

令 (45) 右边为 0, 可导出

$$\begin{aligned}\boldsymbol{P}_{k+1} &= \boldsymbol{F}_k\boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\boldsymbol{F}_k^T + \boldsymbol{Q}_k \\ &= \boldsymbol{F}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{F}_k^T + \boldsymbol{Q}_k\end{aligned} \tag{48}$$

这是关于 $\boldsymbol{P}_k$ 的演化方程，式中 $\widetilde{\boldsymbol{P}}_k$ 的定义为

$$\begin{aligned}\widetilde{\boldsymbol{P}}_k &= \boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1} \\ &= [\boldsymbol{P}_k^{-1} - \theta\overline{\boldsymbol{S}}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k]^{-1}\end{aligned} \tag{49}$$

公式 (49) 表明，如果 $\boldsymbol{P}_k$、$\boldsymbol{S}_k$ 和 $\boldsymbol{R}_k$ 均为对称矩阵，那么 $\widetilde{\boldsymbol{P}}_k$ 也是对称矩阵。从 (48) 可以看出，如果 $\boldsymbol{Q}_k$ 也是对称矩阵，则 $\boldsymbol{P}_{k+1}$ 对称。所以如果 $\boldsymbol{P}_0$、$\boldsymbol{Q}_k$、$\boldsymbol{R}_k$ 和 $\boldsymbol{S}_k$ 均为对称矩阵，那么 $\widetilde{\boldsymbol{P}}_k$ 和 $\boldsymbol{P}_k$ 也是对称的。

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

$$\boldsymbol{x}_0 = \hat{\boldsymbol{x}}_0 + \boldsymbol{P}_0 \boldsymbol{\lambda}_0 \tag{50}$$

$$\boldsymbol{w}_k = \boldsymbol{Q}_k \boldsymbol{\lambda}_{k+1} \tag{51}$$

$$\begin{aligned}\boldsymbol{\lambda}_k =& [\boldsymbol{I} - \theta \overline{\boldsymbol{S}}_k \boldsymbol{P}_k + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{P}_k]^{-1} \times \\ & [\boldsymbol{F}_k^T \boldsymbol{\lambda}_{k+1} + \theta \overline{\boldsymbol{S}}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{\mu}_k)], \quad \boldsymbol{\lambda}_N = 0\end{aligned} \tag{52}$$

$$\boldsymbol{P}_{k+1} = \boldsymbol{F}_k \boldsymbol{P}_k [\boldsymbol{I} - \theta \overline{\boldsymbol{S}}_k \boldsymbol{P}_k + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{P}_k]^{-1} \boldsymbol{F}_k^T + \boldsymbol{Q}_k, \quad \boldsymbol{P}_0 = \boldsymbol{P}_0 \tag{53}$$

$$\begin{aligned}\boldsymbol{\mu}_{k+1} =& \boldsymbol{F}_k \boldsymbol{\mu}_k + \boldsymbol{F}_k \boldsymbol{P}_k [\boldsymbol{I} - \theta \overline{\boldsymbol{S}}_k \boldsymbol{P}_k + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{P}_k]^{-1} \times \\ & [\theta \overline{\boldsymbol{S}}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{\mu}_k)], \quad \boldsymbol{\mu}_0 = \hat{\boldsymbol{x}}_0\end{aligned} \tag{54}$$

2.2 关于 $\hat{\mathrm{x}}_k$ 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 $\boldsymbol{J} = \psi(\boldsymbol{x}_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$（满足约束条件 $\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k$）关于 $\hat{\boldsymbol{x}}_k$ 和 $\boldsymbol{y}_k$ 的驻点。根据 (42) 和初始条件 (47)，可以得到

$$\boldsymbol{\lambda}_k = \boldsymbol{P}_k^{-1}(\boldsymbol{x}_k - \boldsymbol{\mu}_k) \tag{55}$$

$$\boldsymbol{\lambda}_0 = \boldsymbol{P}_0^{-1}(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0) \tag{56}$$

由此可知

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

因此，性能指标 (30) 可以改写为

$$
\boldsymbol{J} = -\epsilon\|\boldsymbol{\lambda}_0\|^2_{\boldsymbol{P}_0} + \sum_{k=0}^{N-1} [\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|^2_{\boldsymbol{S}_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \|\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{x}_k\|^2_{\boldsymbol{R}_k^{-1}})] \tag{57}
$$

将代 (42) 入上式，有

$$
\begin{aligned}
\boldsymbol{J} &= -\epsilon\|\boldsymbol{\lambda}_0\|^2_{\boldsymbol{P}_0} + \\
&\sum_{k=0}^{N-1}[\|\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|^2_{\boldsymbol{S}_k} - \epsilon(\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \|\boldsymbol{y}_k - \boldsymbol{H}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k)\|^2_{\boldsymbol{R}_k^{-1}})]
\end{aligned}
\tag{58}
$$

将 (38) 式代入到 $\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}}$，得到

$$
\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} = \boldsymbol{\lambda}_{k+1}^T\boldsymbol{Q}_k^T\boldsymbol{Q}_k^{-1}\boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_{k+1}^T\boldsymbol{Q}_k\boldsymbol{\lambda}_{k+1} \tag{59}
$$

于是，式 (58) 可以重写为

$$\boldsymbol{J} = -\epsilon\|\boldsymbol{\lambda}_0\|^2_{\boldsymbol{P}_0} - \epsilon\sum_{k=0}^{N-1}\|\boldsymbol{\lambda}_{k+1}\|^2_{\boldsymbol{Q}_k} + \sum_{k=0}^{N-1}[\|\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|^2_{\boldsymbol{S}_k} - \epsilon\|\boldsymbol{y}_k - \boldsymbol{H}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k)\|^2_{\boldsymbol{R}_k^{-1}}] \tag{60}$$

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

$$\sum_{k=0}^{N}\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{\lambda}_k - \sum_{k=0}^{N-1}\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{\lambda}_k = 0 \tag{61}$$

上式可以改写为

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

性能指标 (60) 减去上式的零项，有

$$\begin{aligned}
\boldsymbol{J} = \sum_{k=0}^{N-1} [\|\boldsymbol{\mu}_k + \boldsymbol{P}_k\boldsymbol{\lambda}_k - \hat{\boldsymbol{x}}_k\|^2_{\boldsymbol{S}_k} - \epsilon\|\boldsymbol{y}_k - \boldsymbol{H}_k(\boldsymbol{\mu}_k + \boldsymbol{P}_k\boldsymbol{\lambda}_k)\|^2_{\boldsymbol{R}_k^{-1}}] - \\
\epsilon \sum_{k=0}^{N-1} \|\boldsymbol{\lambda}_{k+1}\|^2_{\boldsymbol{Q}_k} + \epsilon \sum_{k=0}^{N-1} (\boldsymbol{\lambda}_{k+1}^T \boldsymbol{P}_{k+1}\boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k^T \boldsymbol{P}_k \boldsymbol{\lambda}_k) \\
= \sum_{k=0}^{N-1} [(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T \overline{\boldsymbol{S}}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + 2\,(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T \overline{\boldsymbol{S}}_k \boldsymbol{P}_k \boldsymbol{\lambda}_k + \\
\boldsymbol{\lambda}_k^T \boldsymbol{P}_k \overline{\boldsymbol{S}}_k \boldsymbol{P}_k \boldsymbol{\lambda}_k + \epsilon \boldsymbol{\lambda}_{k+1}^T (\boldsymbol{P}_{k+1} - \boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1} - \epsilon \boldsymbol{\lambda}_k^T \boldsymbol{P}_k \boldsymbol{\lambda}_k - \\
\epsilon(\boldsymbol{\lambda}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k) + \\
2\epsilon(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{\lambda}_k - \epsilon \boldsymbol{\lambda}_k^T \boldsymbol{P}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{\lambda}_k]
\end{aligned} \tag{62}$$

考虑上式中的 $\boldsymbol{\lambda}_{k+1}^T(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1}$ 项，将 $\boldsymbol{P}_{k+1}$ 表达式 (48) 代入其中，可得

$$\begin{aligned}\boldsymbol{\lambda}_{k+1}^T\left(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k\right)\boldsymbol{\lambda}_{k+1} &= \boldsymbol{\lambda}_{k+1}^T(\boldsymbol{Q}_k+\boldsymbol{F}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{F}_k^T-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1} \\ &= \boldsymbol{\lambda}_{k+1}^T\boldsymbol{F}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{F}_k^T\boldsymbol{\lambda}_{k+1}\end{aligned}$$

根据 (44) 式，有

$$\boldsymbol{F}_k^T\boldsymbol{\lambda}_{k+1}=\boldsymbol{\lambda}_k-\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k-\hat{\boldsymbol{x}}_k)+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}[\boldsymbol{y}_k-\boldsymbol{H}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k)] \tag{63}$$

因此

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^T(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1}\\
&=\{\boldsymbol{\lambda}_k-\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k-\hat{\boldsymbol{x}}_k)+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}[\boldsymbol{y}_k-\boldsymbol{H}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k)]\}^T\times\\
&\widetilde{\boldsymbol{P}}_k\{\boldsymbol{\lambda}_k-\theta\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k-\hat{\boldsymbol{x}}_k)+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}[\boldsymbol{y}_k-\boldsymbol{H}_k(\boldsymbol{\mu}_k+\boldsymbol{P}_k\boldsymbol{\lambda}_k)]\}\\
&=\{\boldsymbol{\lambda}_k^T(\boldsymbol{I}-\theta\boldsymbol{P}_k\overline{\boldsymbol{S}}_k+\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k-\\
&(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\}\widetilde{\boldsymbol{P}}_k\{\boldsymbol{\lambda}_k^T(\boldsymbol{I}-\theta\boldsymbol{P}_k\overline{\boldsymbol{S}}_k+\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)-\\
&\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k-(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\}^T
\end{aligned}
$$

因为 $(\boldsymbol{I}-\theta\boldsymbol{P}_k\overline{\boldsymbol{S}}_k+\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)=\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}$，代入上式中可得

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^T(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1}\\
&\quad=\{\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k-(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\}\\
&\quad\widetilde{\boldsymbol{P}}_k\{\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k-(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\}^T\\
&\quad=\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}\boldsymbol{P}_k\boldsymbol{\lambda}_k-\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k-(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k-\\
&\quad\ \theta\boldsymbol{\lambda}_k\boldsymbol{P}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\theta^2(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\\
&\quad\ \theta(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)-\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)+\\
&\quad\ \theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)+\\
&\quad\ (\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{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 $\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k=\theta\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)$。因此，上式可以重写为

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^T(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1}\\
&\quad=\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}\boldsymbol{P}_k\boldsymbol{\lambda}_k-2\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k-\\
&\quad\ 2(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k+\theta^2(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+\\
&\quad\ 2\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)+\\
&\quad\ (\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)
\end{aligned}\tag{64}
$$

从 (49) 式可以得到

$$
\begin{aligned}
\widetilde{\boldsymbol{P}}_k^{-1} &= [\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]\boldsymbol{P}_k^{-1} \\
&= \boldsymbol{P}_k^{-1}[\boldsymbol{I} - \theta\boldsymbol{P}_k\overline{\boldsymbol{S}}_k + \boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k]
\end{aligned}
$$

因此

$$
\begin{aligned}
\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\widetilde{\boldsymbol{P}}_k^{-1}\boldsymbol{P}_k\boldsymbol{\lambda}_k &= \boldsymbol{\lambda}_k^T[\boldsymbol{I} - \theta\boldsymbol{P}_k\overline{\boldsymbol{S}}_k + \boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k]\boldsymbol{P}_k\boldsymbol{\lambda}_k \\
&= \boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{\lambda}_k - \theta\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k + \boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k
\end{aligned}
$$

将上式代入 (64) 式中，得到

$$
\begin{aligned}
&\boldsymbol{\lambda}_{k+1}^T(\boldsymbol{P}_{k+1}-\boldsymbol{Q}_k)\boldsymbol{\lambda}_{k+1}\\
&=\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{\lambda}_k-\theta\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k+\boldsymbol{\lambda}_k^T\boldsymbol{P}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k-\\
&2\theta(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\overline{\boldsymbol{S}}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k-2(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{\lambda}_k+\theta^2(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\times\\
&\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)+2\theta\,(\boldsymbol{\mu}_k-\hat{\boldsymbol{x}}_k)^T\,\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)+\\
&(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\widetilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}(\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{\mu}_k)
\end{aligned}
\tag{65}
$$

将上式代入 (62) 式，有

$$
\begin{aligned}
\boldsymbol{J} = \sum_{k=0}^{N-1} [& (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T \overline{\boldsymbol{S}}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) - \epsilon(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k) + \\
& \theta\,(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T\, \overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \overline{\boldsymbol{S}}_k\,(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + 2\,(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T\, \overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1}\,(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k) + \\
& \epsilon(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)] \\
= \sum_{k=0}^{N-1} [& (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T (\overline{\boldsymbol{S}}_k + \theta \overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \overline{\boldsymbol{S}}_k)(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) + \\
& 2(\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k)^T \overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k) + \\
& \epsilon(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)^T (\boldsymbol{R}_k^{-1} \boldsymbol{H}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} - \boldsymbol{R}_k^{-1})(\boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{\mu}_k)]
\end{aligned}
$$

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

式，求 $\boldsymbol{J}$ 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：

$$\frac{\partial \boldsymbol{J}}{\partial \hat{\boldsymbol{x}}_k} = 2(\overline{\boldsymbol{S}}_k + \theta \overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \overline{\boldsymbol{S}}_k)(\hat{\boldsymbol{x}}_k - \boldsymbol{\mu}_k) + 2\overline{\boldsymbol{S}}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1}(\boldsymbol{H}_k \boldsymbol{\mu}_k - \boldsymbol{y}_k) = 0 \tag{66}$$

$$\begin{aligned} \frac{\partial \boldsymbol{J}}{\partial \boldsymbol{y}_k} = {} & 2\epsilon\left(\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{\mu}_k\right)^T \left(\boldsymbol{R}_k^{-1} \boldsymbol{H}_k \widetilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} - \boldsymbol{R}_k^{-1}\right)(\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{\mu}_k) + \\ & 2\boldsymbol{R}_k^{-1} \boldsymbol{H}_k \widetilde{\boldsymbol{P}}_k \overline{\boldsymbol{S}}_k (\boldsymbol{\mu}_k - \hat{\boldsymbol{x}}_k) = 0 \end{aligned} \tag{67}$$

当 $\hat{\boldsymbol{x}}_k = \boldsymbol{\mu}_k$ 以及 $\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{\mu}_k$ 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 $\boldsymbol{J}$ 关于 $\hat{\boldsymbol{x}}_k$ 的二阶导数：

$$\frac{\partial^2\boldsymbol{J}}{\partial\hat{\boldsymbol{x}}_k^2}=2(\overline{\boldsymbol{S}}_k+\theta\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k) \tag{68}$$

如果 $(\overline{\boldsymbol{S}}_k+\theta\overline{\boldsymbol{S}}_k\widetilde{\boldsymbol{P}}_k\overline{\boldsymbol{S}}_k)>0$，那么 $\hat{\boldsymbol{x}}_k$ 为极小值点。通常情况下，$\overline{\boldsymbol{S}}_k$ 的正定性可以保证，如果 $\widetilde{\boldsymbol{P}}_k$ 正定则 $\hat{\boldsymbol{x}}_k$ 为极小值点。

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

$$(\boldsymbol{P}_k^{-1}-\theta\overline{\boldsymbol{S}}_k+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)^{-1}>0$$

等价于

$$(\boldsymbol{P}_k^{-1}-\theta\overline{\boldsymbol{S}}_k+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)>0$$

其中，$\boldsymbol{P}_k^{-1}$、$\theta\overline{\boldsymbol{S}}_k$ 和 $\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k$ 均正定。因此，为了保证正定性，$\theta\overline{\boldsymbol{S}}_k$ 应尽可能小。为了使 $\theta\overline{\boldsymbol{S}}_k$ 变小，可采用如下 3 种方式：

1）$\theta = 1/\epsilon$ 变小可以使 $\theta\overline{\boldsymbol{S}}_k$ 变小，这意味着 (22) 或 (23) 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）$\boldsymbol{L}_k$ 变小可以使 $\theta\overline{\boldsymbol{S}}_k$ 变小。这个结论是根据式中 $\overline{\boldsymbol{S}}_k$ 和 $\boldsymbol{L}_k$ 的关系得出的。目标函数的分子为 $(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^T\boldsymbol{L}_k^T\boldsymbol{S}_k\boldsymbol{L}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)$。$\boldsymbol{L}_k$ 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；$\boldsymbol{L}_k$ 过大可能会导致问题无解。

3）$\boldsymbol{S}_k$ 变小可以使 $\theta\overline{\boldsymbol{S}}_k$ 变小。与上面的结论类似。$\boldsymbol{S}_k$ 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；$\boldsymbol{S}_k$ 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。

2.3.1 系统模型

考虑如下动态系统及输出方程：

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k \\ \boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k \\ \boldsymbol{z}_k = \boldsymbol{L}_k \boldsymbol{x}_k \end{cases} \tag{69}$$

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 $\boldsymbol{x}_k$。

2.3.2 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：

$$\boldsymbol{J}_1 = \frac{\sum_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|_{\boldsymbol{S}_k}^2}{\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|_{\boldsymbol{P}_0^{-1}}^2 + \sum_{k=0}^{N-1} \left(\|\boldsymbol{w}_k\|_{\boldsymbol{Q}_k^{-1}}^2 + \|\boldsymbol{v}_k\|_{\boldsymbol{R}_k^{-1}}^2\right)} \tag{70}$$

式中，矩阵 $\boldsymbol{P}_0$，$\boldsymbol{Q}_k$，$\boldsymbol{R}_k$ 和 $\boldsymbol{S}_k$ 均为正定对称矩阵。

2.3.3 滤波算法

将 $\hat{\boldsymbol{x}}_k = \boldsymbol{\mu}_k$ 代入 (54)，并令

$$\boldsymbol{K}_k = \boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}$$

可得

$$\hat{\boldsymbol{x}}_{k+1} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_k + \boldsymbol{F}_k\boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k), \quad \hat{\boldsymbol{x}}_0 = \hat{\boldsymbol{x}}_0$$

考虑到 $\boldsymbol{P}_k$ 的计算公式 (53) 以及 $\bar{\boldsymbol{S}}_k$ 的计算公式 (29)，如下称为 H_∞ 滤波的递推滤波算法使目标函数 $\boldsymbol{J}_1$ 小于 ϵ，其中 $\theta = 1/\epsilon$.

$$\begin{cases} \hat{\boldsymbol{x}}_{k+1} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_k + \boldsymbol{F}_k\boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k), \quad \hat{\boldsymbol{x}}_0 = \hat{\boldsymbol{x}}_0 \\ \boldsymbol{P}_{k+1} = \boldsymbol{F}_k\boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\boldsymbol{F}_k^T + \boldsymbol{Q}_k, \quad \boldsymbol{P}_0 = \boldsymbol{P}_0 \end{cases} \tag{71}$$

其中:

$$\begin{cases} \boldsymbol{K}_k = \boldsymbol{P}_k[\boldsymbol{I} - \theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k]^{-1}\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1} \\ \overline{\boldsymbol{S}}_k = \boldsymbol{L}_k^\mathsf{T}\boldsymbol{S}_k\boldsymbol{L}_k \end{cases} \tag{72}$$

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 $\boldsymbol{P}_k^{-1} - \theta\overline{\boldsymbol{S}}_k + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k > 0$ 成立。

3. 卡尔曼滤波的另一种形式

为了比较 H_∞ 滤波与 Kalman 滤波，这里给出卡尔曼滤波的另外一种形式——基于一步预测的 Kalman 滤波算法。

仍然考虑如下线性动态系统：

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{w}_k \\ \boldsymbol{y}_{k+1} = \boldsymbol{H}_{k+1} \boldsymbol{x}_{k+1} + \boldsymbol{v}_{k+} \end{cases} \tag{73}$$

式中：$\boldsymbol{w}_k$ 和 $\boldsymbol{v}_{k+1}$ 均为零均值高斯分布白噪声，协方差分别为 $\boldsymbol{Q}_k$ 和 $\boldsymbol{R}_{k+1}$。

根据 Kalman 滤波理论，我们有

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{F}_k \hat{\boldsymbol{x}}_{k|k} \tag{74}$$

$$\boldsymbol{P}_{k+1|k} = \boldsymbol{F}_k \boldsymbol{P}_{k|k} \boldsymbol{F}_k^T + \boldsymbol{Q}_k \tag{75}$$

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k+1} &= \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1}(\boldsymbol{y}_{k+1} - \boldsymbol{H}_{k+1}\hat{\boldsymbol{x}}_{k+1|k}) \\ &= \boldsymbol{P}_{k|+1k+1}[\boldsymbol{P}_{k+1|k}^{-1}\hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{y}_{k+1}] \end{aligned} \tag{76}$$

其中：

$$\boldsymbol{P}_{k|+1k+1} = [\boldsymbol{P}_{k+1|k}^{-1} + \boldsymbol{H}_{k+1}^T \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1}]^{-1} \tag{77}$$

$$\begin{aligned}\boldsymbol{K}_{k+1} &= [\boldsymbol{P}_{k+1|k}^{-1} + \boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1}\boldsymbol{H}_{k+1}]^{-1}\boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1} \\ &= \boldsymbol{P}_{k+1|k}[\boldsymbol{I} + \boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1}\boldsymbol{H}_{k+1}\boldsymbol{P}_{k+1|k}]^{-1}\boldsymbol{H}_{k+1}^T\boldsymbol{R}_{k+1}^{-1} \qquad (78)\end{aligned}$$

将 (77) 代入 (75)，可知

$$\begin{aligned}\boldsymbol{P}_{k+1|k} &= \boldsymbol{F}_k\boldsymbol{P}_{k|k}\boldsymbol{F}_k^T + \boldsymbol{Q}_k \\ &= \boldsymbol{F}_k[\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k]^{-1}\boldsymbol{F}_k^T + \boldsymbol{Q}_k \\ &= \boldsymbol{F}_k\boldsymbol{P}_{k|k-1}[\boldsymbol{I} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}]^{-1}\boldsymbol{F}_k^T + \boldsymbol{Q}_k \qquad (79)\end{aligned}$$

由 (74) 和 (76) 有

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_{k|k} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{F}_k\boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}) \qquad (80)$$

综上，一步预测形式的卡尔曼滤波方程可表示为：

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{F}_k\boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}) \tag{81}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\left(\boldsymbol{I} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\right)^{-1}\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1} \tag{82}$$

$$\boldsymbol{P}_{k+1|k} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k|k-1}\left(\boldsymbol{I} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\right)^{-1}\boldsymbol{F}_k^T + \boldsymbol{Q}_k \tag{83}$$

我们知道，卡尔曼滤波算法除了需要知道知状态转移矩阵 $\boldsymbol{F}_k$ 和量测矩阵 $\boldsymbol{H}_k$ 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 $\boldsymbol{Q}_k$ 和 $\boldsymbol{R}_k$。只有当噪声服从高斯分布时，卡尔曼滤波是最小方差估计器；当噪声非高斯分布时，卡尔曼滤波是线性最小方差估计器。

4. H_∞ 滤波和卡尔曼滤波的区别和联系

对比分析 (71)、(72) 和 (81)~(83) 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，$\boldsymbol{Q}_k$、$\boldsymbol{R}_k$ 和 $\boldsymbol{P}_0$ 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，$\boldsymbol{Q}_k$、$\boldsymbol{R}_k$ 和 $\boldsymbol{P}_0$ 分别是它们各自已知的协方差矩阵。

- 在 H_∞ 滤波中取 $\boldsymbol{L}_k = \boldsymbol{S}_k = \boldsymbol{I}$，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 (70) 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。也就是，如果我们利用卡尔曼滤波估计 $\boldsymbol{L}_k\boldsymbol{x}_k$，那么结果与我们选择的 $\boldsymbol{L}_k$ 矩阵无关。然而，当我们使用 H_∞ 滤波时，结果与 $\boldsymbol{L}_k$ 和我们想要估计的状态组合有很大关系。

- 观察 (71) 和 (72) 式，当去掉 $\boldsymbol{K}_k$ 和 $\boldsymbol{P}_{k+1}$ 方程中的 $\theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k$ 项后，H_∞ 滤波算法形式上变为卡尔曼滤波完全一样。

- 在 Kalman 滤波中，对于未建模的动态系统和噪声，可以通过增大 $\boldsymbol{Q}_k$ 以改善卡尔曼滤波的鲁棒性。这样做会使得协方差矩阵 $\boldsymbol{P}_{k+1|k}$ 增大，同时使得卡尔曼滤波增益矩阵 $\boldsymbol{K}_k$ 变大。

- 在 H_∞ 滤波中，从 (71) 式可以看出，$\boldsymbol{P}_{k+1}$ 等式右边 $(-\theta\overline{\boldsymbol{S}}_k\boldsymbol{P}_k)$ 项相当于变大 $\boldsymbol{P}_{k+1}$。类似地，也会使得矩阵 $\boldsymbol{K}_k$ 变大。

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

5. 小结

H_∞ 滤波算法能够最小化最差情况下的估计误差，可以认为是一种鲁棒的卡尔曼滤波算法。与卡尔曼滤波相比，H_∞ 滤波的优势在于更适合处理系统模型存在不确定性的情况。不难验证，H_∞ 滤波比 Kalman 滤波对于参数的选取要敏感些。

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?