

卡尔曼滤波

最小二乘估计、状态传递与离散卡尔曼滤波

李丹

2026 年 6 月 1 日

卡尔曼滤波概述

卡尔曼滤波的历史与应用

最小二乘估计

状态传递

离散时间卡尔曼滤波

算例一：单自由度系统状态估计

算例二：MEMS 倾角传感与卡尔曼滤波

扩展：多模型估计 (MMAE)

小结

卡尔曼滤波概述

为什么需要卡尔曼滤波

卡尔曼滤波 (R. E. Kalman, 1960) 是动态系统状态估计的高效**递推**算法：在含噪声观测下，融合**系统动力学模型**与**实时观测**，给出状态的最优估计。

- 递推：只用上一步结果与当前观测，无需存储全部历史数据；
- 最优：在线性高斯假设下为最小均方误差估计；
- 给出估计的**不确定性** (协方差)，可量化置信度。

在结构健康监测中，卡尔曼滤波是时域模型修正、损伤诊断、状态/参数联合识别的基础工具。本讲路线：最小二乘 → 递归最小二乘 → 状态传递 → 离散卡尔曼滤波。

卡尔曼滤波的历史与应用

创始人与 1960 年的开山之作



Rudolf E. Kálmán
(1930–2016)
匈牙利裔美籍

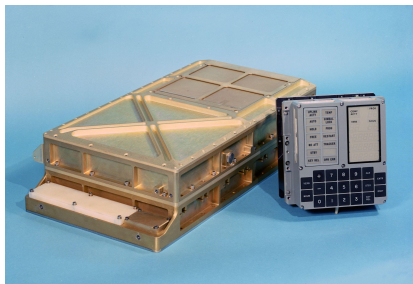
论文: *A New Approach to Linear Filtering and Prediction Problems*, Trans. ASME J. Basic Engineering, **82**(1): 35–45, 1960。

- 突破 Wiener 滤波器（频域、批处理、需平稳假设）的局限，首次给出**时域递推**的最优估计框架；
- 用**状态空间**描述动态系统，估计与不确定性**同步**递推；
- 投稿初期**未受重视**——审稿人质疑其工程价值；
- 1960 年 Kálmán 访问 NASA Ames Research Center, Stanley Schmidt 认定该方法对登月导航不可或缺，自此打开应用之门。

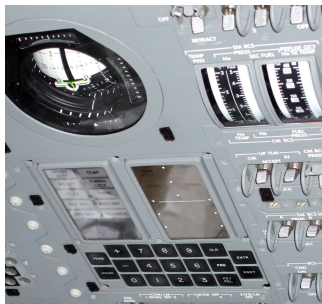
"I had to do it because nobody else would."

— R. E. Kálmán

Apollo 制导计算机

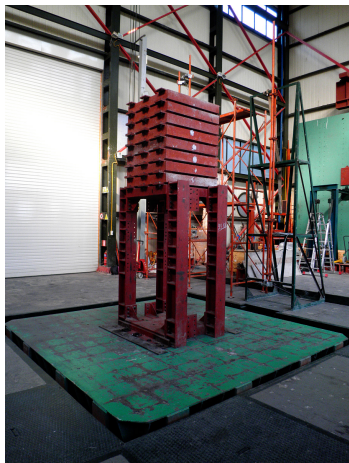


Apollo Guidance Computer (AGC) 与 DSKY 显示键盘 (NASA)



指令舱内的 DSKY 操作面板特写

- NASA Ames Research Center 团队 (Schmidt 等) 将 Kálmán 的线性递推方法推广到**非线性轨道动力学**, 发展出 **扩展卡尔曼滤波 (EKF)**;
- 1969 年 Apollo 11 登月: AGC 内存仅 **4 KB**, 主频 **2 MHz**, 却实时跑卡尔曼滤波融合 IMU、星敏、雷达数据;
- 自此卡尔曼滤波成为**所有现代航天器**制导导航的标配。



振动台试验 (NTUA, Athens)

当代应用版图：

- **航天与航空**：火箭/卫星轨道、姿态估计、组合导航；
- **GPS / 惯导融合**：手机、车载、无人机定位；
- **机器人与自动驾驶**：SLAM、传感器融合、多目标跟踪；
- **结构健康监测** (本课重点)：
 - 状态估计——由稀疏加速度反演位移与速度；
 - 参数识别——刚度/阻尼时变追踪；
 - 损伤诊断——多模型估计 (MMAE)、自适应 KF；
- **其他**：经济学 (动态因子模型)、生物医学 (脑机接口)、气象 (资料同化)。

共同特征：含噪观测 + 动态模型 + 实时性 $\Rightarrow$ KF 家族。

最小二乘估计

不考虑直接输入的线性观测系统

$$\mathbf{y} = \mathbf{C}_d \mathbf{x} + \mathbf{v}, \quad \mathbf{C}_d \in \mathbb{R}^{n_y \times n_x}, \quad \mathbb{E}[\mathbf{v}] = \mathbf{0}, \quad \mathbb{E}[v_i^2] = \sigma^2$$

不考虑直接输入的线性观测系统

$$\mathbf{y} = \mathbf{C}_d \mathbf{x} + \mathbf{v}, \quad \mathbf{C}_d \in \mathbb{R}^{n_y \times n_x}, \quad \mathbb{E}[\mathbf{v}] = \mathbf{0}, \quad \mathbb{E}[v_i^2] = \sigma^2$$

设状态估计为 $\hat{\mathbf{x}}$, 定义观测残差

$$\epsilon_y = \mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}}$$

不考虑直接输入的线性观测系统

$$\mathbf{y} = \mathbf{C}_d \mathbf{x} + \mathbf{v}, \quad \mathbf{C}_d \in \mathbb{R}^{n_y \times n_x}, \quad \mathbb{E}[\mathbf{v}] = \mathbf{0}, \quad \mathbb{E}[\mathbf{v}_i^2] = \sigma^2$$

设状态估计为 $\hat{\mathbf{x}}$, 定义观测残差

$$\boldsymbol{\epsilon}_y = \mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}}$$

最小二乘准则: 使残差的 $\mathcal{L}_2$ 范数平方最小

$$\hat{\mathbf{x}} = \arg \min_{\hat{\mathbf{x}}} \|\boldsymbol{\epsilon}_y\|_2^2 = \arg \min_{\hat{\mathbf{x}}} (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})^T (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})$$

正规方程的推导

展开目标函数（标量项 $\hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{y} = \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}}$ ）：

$$J = \mathbf{y}^T \mathbf{y} - 2 \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}} + \hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{C}_d \hat{\mathbf{x}}$$

正规方程的推导

展开目标函数（标量项 $\hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{y} = \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}}$ ）：

$$J = \mathbf{y}^T \mathbf{y} - 2 \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}} + \hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{C}_d \hat{\mathbf{x}}$$

对 $\hat{\mathbf{x}}$ 求梯度（用 $\partial(\mathbf{b}^T \hat{\mathbf{x}})/\partial \hat{\mathbf{x}} = \mathbf{b}$ 、 $\partial(\hat{\mathbf{x}}^T \mathbf{M} \hat{\mathbf{x}})/\partial \hat{\mathbf{x}} = 2\mathbf{M}\hat{\mathbf{x}}$ ）：

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = -2 \mathbf{C}_d^T \mathbf{y} + 2 \mathbf{C}_d^T \mathbf{C}_d \hat{\mathbf{x}} = \mathbf{0}$$

正规方程的推导

展开目标函数（标量项 $\hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{y} = \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}}$ ）：

$$J = \mathbf{y}^T \mathbf{y} - 2 \mathbf{y}^T \mathbf{C}_d \hat{\mathbf{x}} + \hat{\mathbf{x}}^T \mathbf{C}_d^T \mathbf{C}_d \hat{\mathbf{x}}$$

对 $\hat{\mathbf{x}}$ 求梯度（用 $\partial(\mathbf{b}^T \hat{\mathbf{x}})/\partial \hat{\mathbf{x}} = \mathbf{b}$ 、 $\partial(\hat{\mathbf{x}}^T \mathbf{M} \hat{\mathbf{x}})/\partial \hat{\mathbf{x}} = 2\mathbf{M}\hat{\mathbf{x}}$ ）：

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = -2 \mathbf{C}_d^T \mathbf{y} + 2 \mathbf{C}_d^T \mathbf{C}_d \hat{\mathbf{x}} = \mathbf{0}$$

当 $\mathbf{C}_d$ 列满秩（ $\text{rank } \mathbf{C}_d = n_x$ ）时 $\mathbf{C}_d^T \mathbf{C}_d$ 可逆，

$$\hat{\mathbf{x}} = (\mathbf{C}_d^T \mathbf{C}_d)^{-1} \mathbf{C}_d^T \mathbf{y} = \mathbf{C}_d^\dagger \mathbf{y}$$

$\mathbf{C}_d^\dagger$ 即 Moore–Penrose 伪逆，几何上是 $\mathbf{y}$ 在 $\mathbf{C}_d$ 列空间上的正交投影。

当各传感器精度不同时，观测噪声协方差为对角阵

$$\Sigma_v = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \text{diag}(\sigma_1^2, \dots, \sigma_{n_y}^2)$$

当各传感器精度不同时，观测噪声协方差为对角阵

$$\Sigma_v = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \text{diag}(\sigma_1^2, \dots, \sigma_{n_y}^2)$$

σ_i^2 越大，第 i 个观测越不可靠。以 Σ_v^{-1} 为权构造目标

$$J = (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})^T \Sigma_v^{-1} (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})$$

当各传感器精度不同时，观测噪声协方差为对角阵

$$\Sigma_v = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \text{diag}(\sigma_1^2, \dots, \sigma_{n_y}^2)$$

σ_i^2 越大，第 i 个观测越不可靠。以 Σ_v^{-1} 为权构造目标

$$J = (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})^T \Sigma_v^{-1} (\mathbf{y} - \mathbf{C}_d \hat{\mathbf{x}})$$

同样令 $\partial J / \partial \hat{\mathbf{x}} = \mathbf{0}$ ，得加权最小二乘解

$$\hat{\mathbf{x}} = (\mathbf{C}_d^T \Sigma_v^{-1} \mathbf{C}_d)^{-1} \mathbf{C}_d^T \Sigma_v^{-1} \mathbf{y}$$

可靠（方差小）的观测获得更大权重，主导估计结果。

递归最小二乘：动机与结构

工程中观测逐次到来。批处理需不断扩充观测矩阵，计算量随数据量急剧上升。**递归最小二乘 (RLS)**：已有 $\hat{\mathbf{x}}_{n-1}$ ，来了新观测 $\mathbf{y}_n$ ，用递推公式更新而不重算全部数据。

$$\mathbf{y}_n = \mathbf{C}_d \mathbf{x} + \mathbf{v}_n, \quad \hat{\mathbf{x}}_n = \hat{\mathbf{x}}_{n-1} + \mathbf{G}_n (\mathbf{y}_n - \mathbf{C}_d \hat{\mathbf{x}}_{n-1})$$

- $(\mathbf{y}_n - \mathbf{C}_d \hat{\mathbf{x}}_{n-1})$ ：**预测残差** (innovation)，新观测与基于历史数据的预测之差，代表观测中**未被预测捕获**的新信息；
- $\mathbf{G}_n$ ：待定**增益矩阵**，决定相信新观测的程度；
- 关键问题：如何选取 $\mathbf{G}_n$ 使估计最优。

记估计误差 $\epsilon_n^x = \mathbf{x} - \hat{\mathbf{x}}_n$ 。代入递推式：

$$\begin{aligned}\mathbb{E}[\epsilon_n^x] &= \mathbb{E}[\mathbf{x} - \hat{\mathbf{x}}_{n-1} - \mathbf{G}_n(\mathbf{C}_d\mathbf{x} + \mathbf{v}_n - \mathbf{C}_d\hat{\mathbf{x}}_{n-1})] \\ &= (\mathbf{I} - \mathbf{G}_n\mathbf{C}_d) \mathbb{E}[\epsilon_{n-1}^x] - \mathbf{G}_n \mathbb{E}[\mathbf{v}_n]\end{aligned}$$

若观测噪声零均值 $\mathbb{E}[\mathbf{v}_n] = \mathbf{0}$ ，且初值无偏 $\hat{\mathbf{x}}_0 = \mathbb{E}[\mathbf{x}]$ ，则对一切 n 有 $\mathbb{E}[\epsilon_n^x] = \mathbf{0}$ 。

即该递推式是**无偏估计**，与增益 $\mathbf{G}_n$ 的具体取值无关—— $\mathbf{G}_n$ 只影响估计的**方差**。

以误差方差之和为优化目标

$$J_n = \mathbb{E}[(\epsilon_n^x)^T \epsilon_n^x] = \text{Tr}(\Sigma_n^x), \quad \Sigma_n^x = \mathbb{E}[\epsilon_n^x (\epsilon_n^x)^T]$$

以误差方差之和为优化目标

$$J_n = \mathbb{E}[(\epsilon_n^x)^T \epsilon_n^x] = \text{Tr}(\Sigma_n^x), \quad \Sigma_n^x = \mathbb{E}[\epsilon_n^x (\epsilon_n^x)^T]$$

由 $\epsilon_n^x = (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \epsilon_{n-1}^x - \mathbf{G}_n \mathbf{v}_n$, 且 ϵ_{n-1}^x 与 $\mathbf{v}_n$ 独立 (交叉项期望为零), 展开得

$$\Sigma_n^x = (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n-1}^x (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d)^T + \mathbf{G}_n \Sigma_n^v \mathbf{G}_n^T$$

这是关于增益 $\mathbf{G}_n$ 的二次型, 下面对其求极小。

把协方差递推代入 $J_n = \text{Tr}(\Sigma_n^x)$, 对 $\mathbf{G}_n$ 求导 (用 $\partial \text{Tr}(\mathbf{G}\mathbf{M}\mathbf{G}^T)/\partial \mathbf{G} = 2\mathbf{G}\mathbf{M}$, $\mathbf{M}$ 对称):

$$\frac{\partial J_n}{\partial \mathbf{G}_n} = -2(\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n-1}^x \mathbf{C}_d^T + 2 \mathbf{G}_n \Sigma_n^v = \mathbf{0}$$

把协方差递推代入 $J_n = \text{Tr}(\Sigma_n^x)$, 对 $\mathbf{G}_n$ 求导 (用 $\partial \text{Tr}(\mathbf{G}\mathbf{M}\mathbf{G}^T)/\partial \mathbf{G} = 2\mathbf{G}\mathbf{M}$, $\mathbf{M}$ 对称):

$$\frac{\partial J_n}{\partial \mathbf{G}_n} = -2(\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n-1}^x \mathbf{C}_d^T + 2 \mathbf{G}_n \Sigma_n^v = \mathbf{0}$$

整理 ($\Sigma_{n-1}^x, \Sigma_n^v$ 正定):

$$\mathbf{G}_n (\mathbf{C}_d \Sigma_{n-1}^x \mathbf{C}_d^T + \Sigma_n^v) = \Sigma_{n-1}^x \mathbf{C}_d^T$$

把协方差递推代入 $J_n = \text{Tr}(\Sigma_n^x)$, 对 $\mathbf{G}_n$ 求导 (用 $\partial \text{Tr}(\mathbf{G}\mathbf{M}\mathbf{G}^T)/\partial \mathbf{G} = 2\mathbf{G}\mathbf{M}$, $\mathbf{M}$ 对称):

$$\frac{\partial J_n}{\partial \mathbf{G}_n} = -2(\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n-1}^x \mathbf{C}_d^T + 2 \mathbf{G}_n \Sigma_n^v = \mathbf{0}$$

整理 ($\Sigma_{n-1}^x, \Sigma_n^v$ 正定):

$$\mathbf{G}_n (\mathbf{C}_d \Sigma_{n-1}^x \mathbf{C}_d^T + \Sigma_n^v) = \Sigma_{n-1}^x \mathbf{C}_d^T$$

$$\mathbf{G}_n = \Sigma_{n-1}^x \mathbf{C}_d^T (\mathbf{C}_d \Sigma_{n-1}^x \mathbf{C}_d^T + \Sigma_n^v)^{-1}$$

这正是**卡尔曼增益**的形式 (仅差状态传递一步)。

状态传递

线性时不变系统

$$\mathbf{x}_n = \mathbf{A}_d \mathbf{x}_{n-1} + \mathbf{B}_d \mathbf{u}_{n-1} + \mathbf{w}_{n-1}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma_{n-1}^w)$$

若初始状态 $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0)$ 为高斯, 由于 **线性变换保持高斯性**, $\mathbf{x}_n$ 对任意 n 仍为高斯分布。

高斯分布由**均值**与**协方差**唯一确定, 故状态传递只需递推这两个量。

均值 (噪声零均值):

$$\mu_n^x = \mathbb{E}[\mathbf{x}_n] = \mathbf{A}_d \mu_{n-1}^x + \mathbf{B}_d \mathbf{u}_{n-1}$$

均值与协方差的递推

均值 (噪声零均值):

$$\mu_n^x = \mathbb{E}[\mathbf{x}_n] = \mathbf{A}_d \mu_{n-1}^x + \mathbf{B}_d \mathbf{u}_{n-1}$$

协方差: 由 $\mathbf{x}_n - \mu_n^x = \mathbf{A}_d(\mathbf{x}_{n-1} - \mu_{n-1}^x) + \mathbf{w}_{n-1}$, 且 $\mathbf{x}_{n-1}$ 与 $\mathbf{w}_{n-1}$ 独立, 展开外积期望 (交叉项为零):

$$\begin{aligned}\Sigma_n^x &= \mathbb{E}[(\mathbf{A}_d \delta \mathbf{x}_{n-1} + \mathbf{w}_{n-1})(\mathbf{A}_d \delta \mathbf{x}_{n-1} + \mathbf{w}_{n-1})^T] \\ &= \mathbf{A}_d \Sigma_{n-1}^x \mathbf{A}_d^T + \Sigma_{n-1}^w\end{aligned}$$

均值与协方差的递推

均值 (噪声零均值):

$$\mu_n^x = \mathbb{E}[\mathbf{x}_n] = \mathbf{A}_d \mu_{n-1}^x + \mathbf{B}_d \mathbf{u}_{n-1}$$

协方差: 由 $\mathbf{x}_n - \mu_n^x = \mathbf{A}_d(\mathbf{x}_{n-1} - \mu_{n-1}^x) + \mathbf{w}_{n-1}$, 且 $\mathbf{x}_{n-1}$ 与 $\mathbf{w}_{n-1}$ 独立, 展开外积期望 (交叉项为零):

$$\begin{aligned}\Sigma_n^x &= \mathbb{E}[(\mathbf{A}_d \delta \mathbf{x}_{n-1} + \mathbf{w}_{n-1})(\mathbf{A}_d \delta \mathbf{x}_{n-1} + \mathbf{w}_{n-1})^T] \\ &= \mathbf{A}_d \Sigma_{n-1}^x \mathbf{A}_d^T + \Sigma_{n-1}^w\end{aligned}$$

$$\Sigma_n^x = \mathbf{A}_d \Sigma_{n-1}^x \mathbf{A}_d^T + \Sigma_{n-1}^w$$

此即**离散时间 Lyapunov 方程**: 模型动力学传播不确定性, 过程噪声持续注入不确定性。

离散时间卡尔曼滤波

线性离散时间系统

$$\mathbf{x}_n = \mathbf{A}_d \mathbf{x}_{n-1} + \mathbf{B}_d \mathbf{u}_{n-1} + \mathbf{w}_{n-1}$$

$$\mathbf{y}_n = \mathbf{C}_d \mathbf{x}_n + \mathbf{v}_n$$

线性离散时间系统

$$\mathbf{x}_n = \mathbf{A}_d \mathbf{x}_{n-1} + \mathbf{B}_d \mathbf{u}_{n-1} + \mathbf{w}_{n-1}$$

$$\mathbf{y}_n = \mathbf{C}_d \mathbf{x}_n + \mathbf{v}_n$$

系统噪声与观测噪声为零均值、互不相关的高斯白噪声

$$\mathbf{w}_n \sim \mathcal{N}(\mathbf{0}, \Sigma_n^{\mathbf{w}}), \quad \mathbf{v}_n \sim \mathcal{N}(\mathbf{0}, \Sigma_n^{\mathbf{v}}), \quad \mathbb{E}[\mathbf{w}_m \mathbf{v}_n^T] = \mathbf{0}$$

目标：在每个时间步给出 $\mathbf{x}_n$ 的最优估计及其协方差。

用条件期望（最小均方误差意义下最优）定义两类估计：

$$\text{先验: } \hat{\mathbf{x}}_{n|n-1} = \mathbb{E}[\mathbf{x}_n | \mathbf{y}_{0:n-1}], \quad \text{后验: } \hat{\mathbf{x}}_{n|n} = \mathbb{E}[\mathbf{x}_n | \mathbf{y}_{0:n}]$$

对应误差协方差 $\Sigma_{n|n-1}^x$ 与 $\Sigma_{n|n}^x$ 。后验使用了更多信息（含 $\mathbf{y}_n$ ），精度高于先验。

卡尔曼滤波交替执行两步：

- **预测（时间更新）**：仅用模型，由 $n-1$ 推到 n ；
- **校正（测量更新）**：用新观测 $\mathbf{y}_n$ 修正先验。

区间 $[n-1, n)$ 内无新观测，状态预测完全依赖动力学模型。借用**状态传递**一节的结论：

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \mathbf{A}_d \hat{\mathbf{x}}_{n-1|n-1} + \mathbf{B}_d \mathbf{u}_{n-1} \\ \Sigma_{n|n-1}^x &= \mathbf{A}_d \Sigma_{n-1|n-1}^x \mathbf{A}_d^T + \Sigma_{n-1}^w\end{aligned}$$

模型把上一时刻的后验推进为本时刻的先验；过程噪声使先验协方差增大（不确定性上升）。

获得 y_n 后，借用**递归最小二乘**的最优增益（此处 $\Sigma_{n-1}^x \rightarrow \Sigma_{n|n-1}^x$ ）：

$$\mathbf{G}_n = \Sigma_{n|n-1}^x \mathbf{C}_d^T (\mathbf{C}_d \Sigma_{n|n-1}^x \mathbf{C}_d^T + \Sigma_n^v)^{-1}$$

获得 y_n 后，借用**递归最小二乘**的最优增益（此处 $\Sigma_{n-1}^x \rightarrow \Sigma_{n|n-1}^x$ ）：

$$\mathbf{G}_n = \Sigma_{n|n-1}^x \mathbf{C}_d^T (\mathbf{C}_d \Sigma_{n|n-1}^x \mathbf{C}_d^T + \Sigma_n^v)^{-1}$$

用预测残差修正先验，得后验估计与协方差：

$$\begin{aligned}\hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{G}_n (\mathbf{y}_n - \mathbf{C}_d \hat{\mathbf{x}}_{n|n-1}) \\ \Sigma_{n|n}^x &= (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n|n-1}^x (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d)^T + \mathbf{G}_n \Sigma_n^v \mathbf{G}_n^T\end{aligned}$$

协方差更新写成**Joseph 形式**，对舍入误差更稳健，且始终保持对称半正定。

初始化

$$\hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0], \quad \Sigma_{0|0}^{\mathbf{x}} = \mathbb{E}[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T]$$

初始化

$$\hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0], \quad \Sigma_{0|0}^{\mathbf{x}} = \mathbb{E}[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T]$$

迭代 ($n = 1, 2, \dots$)

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \mathbf{A}_d \hat{\mathbf{x}}_{n-1|n-1} + \mathbf{B}_d \mathbf{u}_{n-1} \\ \Sigma_{n|n-1}^{\mathbf{x}} &= \mathbf{A}_d \Sigma_{n-1|n-1}^{\mathbf{x}} \mathbf{A}_d^T + \Sigma_{n-1}^{\mathbf{w}} \\ \mathbf{G}_n &= \Sigma_{n|n-1}^{\mathbf{x}} \mathbf{C}_d^T (\mathbf{C}_d \Sigma_{n|n-1}^{\mathbf{x}} \mathbf{C}_d^T + \Sigma_n^{\mathbf{v}})^{-1} \\ \hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{G}_n (\mathbf{y}_n - \mathbf{C}_d \hat{\mathbf{x}}_{n|n-1}) \\ \Sigma_{n|n}^{\mathbf{x}} &= (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d) \Sigma_{n|n-1}^{\mathbf{x}} (\mathbf{I} - \mathbf{G}_n \mathbf{C}_d)^T + \mathbf{G}_n \Sigma_n^{\mathbf{v}} \mathbf{G}_n^T\end{aligned}$$

卡尔曼滤波与递归最小二乘的关系

离散卡尔曼滤波 = 引入状态传递机制的递归最小二乘。

若 $A_d = I$ 、无外部输入 $u_n = 0$ 、无系统噪声 $w_n = 0$ ，预测步退化为 $\hat{x}_{n|n-1} = \hat{x}_{n-1|n-1}$ 、 $\Sigma_{n|n-1}^x = \Sigma_{n-1|n-1}^x$ ，整个算法退化为对 **常量**的递归最小二乘。

反过来看：RLS 解决“静态参数的逐次估计”，卡尔曼滤波在其上叠加“状态随模型演化”，因而能跟踪**动态系统**。

噪声的影响：卡尔曼增益的极限

增益 $G_n = \Sigma_{n|n-1}^x C_d^T (C_d \Sigma_{n|n-1}^x C_d^T + \Sigma_n^v)^{-1}$ 。

- $\Sigma_n^v \rightarrow 0$ (观测很准): G_n 趋于伪逆形式, $G_n C_d \rightarrow I$, 滤波器高度信任观测;
- $\Sigma_n^v \rightarrow \infty$ (观测很差): $G_n \rightarrow 0$, 几乎不用观测, 依赖模型预测;
- Σ_{n-1}^w 增大 $\Rightarrow$ 先验协方差 $\Sigma_{n|n-1}^x$ 增大 $\Rightarrow$ 增益增大, 更信任观测。

卡尔曼滤波在**模型预测**与**观测**之间寻找最优平衡，由系统噪声与观测噪声的相对大小决定：

$$\frac{\Sigma_{n-1}^w}{\Sigma_n^v} \propto \frac{\text{对观测的信任}}{\text{对模型的信任}}$$

- $\Sigma^w \gg \Sigma^v$ ：更信任观测，跟踪快但易受噪声扰动；
- $\Sigma^w \ll \Sigma^v$ ：更信任模型，平滑但响应慢、可能滞后。

系统可观测时， $\Sigma_{n|n}^x$ 随时间收敛到稳态值（稳态卡尔曼滤波）。

连续系统-离散观测：混合卡尔曼滤波

实际结构多为连续时间系统、离散采样观测。两次观测间用连续模型预测，得到观测时刻执行离散测量更新。工程实现常将连续系统 **离散化**（零阶保持，ZOH）：

$$\mathbf{A}_d = e^{\mathbf{A}_c T}, \quad \mathbf{B}_d = \left(\int_0^T e^{\mathbf{A}_c \tau} d\tau \right) \mathbf{B}_c, \quad \Sigma_d^w = \int_0^T e^{\mathbf{A}_c \tau} \Sigma^w e^{\mathbf{A}_c^T \tau} d\tau$$

T 为采样间隔。离散化后即可套用前述离散卡尔曼滤波。下面的算例采用此方法。

MATLAB: ZOH 离散化 (expm, 免工具箱)

```
1 function [Ad, Bd, Qd] = c2d_zoh(Ac, Bc, Qc, T)
2 n = size(Ac,1); m = size(Bc,2);
3 % Ad, Bd from augmented matrix exponential
4 MM = [Ac, Bc; zeros(m, n+m)];
5 Phi = expm(MM*T);
6 Ad = Phi(1:n, 1:n);
7 Bd = Phi(1:n, n+1:n+m);
8 % discrete process-noise cov by Van Loan
9 F = [-Ac, Qc; zeros(n), Ac'] * T;
10 EG = expm(F);
11 Qd = Ad * EG(1:n, n+1:2*n);
12 Qd = 0.5*(Qd + Qd'); % symmetrize
13 end
```

MATLAB: 离散卡尔曼滤波 (核心)

```
1 x = x0; P = P0; I = eye(numel(x0));
2 for n = 1:N
3     % --- predict (time update) ---
4     x = Ad*x + Bd*U(:,n);
5     P = Ad*P*Ad' + Qd;
6     % --- correct (measurement update) ---
7     S = Cd*P*Cd' + Rv;           % innovation covariance
8     G = (P*Cd') / S;             % Kalman gain
9     x = x + G*(Y(:,n) - Cd*x);   % innovation correction
10    P = (I-G*Cd)*P*(I-G*Cd)' + G*Rv*G'; % Joseph form
11    xhat(:,n) = x; Pdiag(:,n) = diag(P);
12 end
```

算例一：单自由度系统状态估计

质量-弹簧-阻尼系统 $m\ddot{q} + c\dot{q} + kq = w(t)$, 取状态 $\mathbf{x} = (q, \dot{q})^T$:

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ \frac{1}{m} \end{pmatrix} w(t), \quad y_n = \begin{pmatrix} -\frac{k}{m} & -\frac{c}{m} \end{pmatrix} \mathbf{x}(t_n) + v_n$$

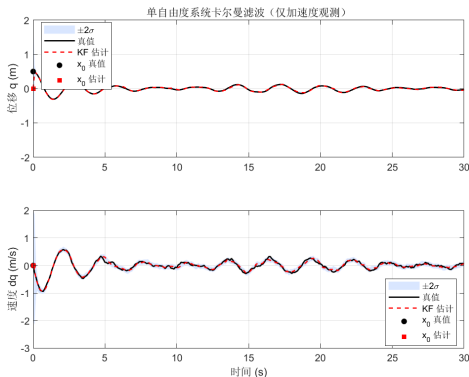
只测加速度, 不直接测位移与速度。参数取自教材例题: $m=1$ kg, $c=0.5$ N·s/m, $k=5$ N/m, $T=0.1$ s, $\sigma_w=0.1$, $\sigma_v=0.1$ 。真实初值 $\mathbf{x}_0=(0.5, 0)^T$, 估计初值 $\hat{\mathbf{x}}_{0|0}=(0, 0)^T$ (误差大)。

MATLAB: 仿真与滤波驱动 (节选)

```
1 par.m=1; par.c=0.5; par.k=5; par.T=0.1; par.N=300;
2 par.sigma_w=0.1; par.sigma_v=0.1; par.x0_true=[0.5;0];
3
4 [Xtrue, Y, tvec, sys] = simulate_sdof(par); % truth + noisy accel.
5
6 x0_est = [0;0]; P0 = diag([1,1]); U = zeros(1,par.N);
7 [xhat, Pdiag] = kalman_discrete(sys.Ad, sys.Bd, sys.Cd, ...
8                               sys.Qd, sys.Rv, x0_est, P0, Y, U);
9
10 rmse_q = sqrt(mean((xhat(1,:)-Xtrue(1,:)).^2));
```

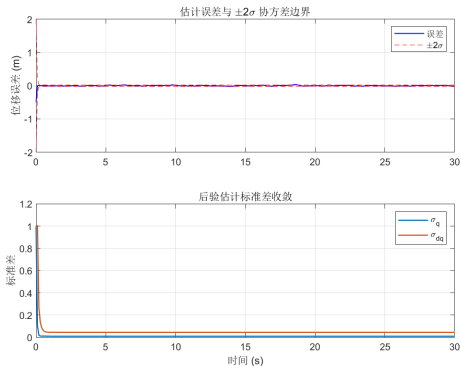
`simulate_sdof` 用 ZOH 离散模型生成真值与含噪加速度观测;
`kalman_discrete` 执行预测/校正循环并返回估计与协方差。

结果：状态估计



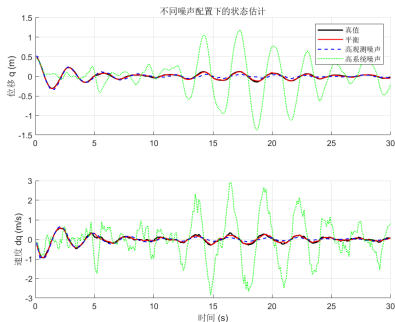
尽管初始状态误差大且**未直接观测**位移与速度，滤波器仍快速收敛并准确重构两者；阴影为 $\pm 2\sigma$ 置信带。RMSE: $q \approx 1.0 \times 10^{-2}$ m, $\dot{q} \approx 4.6 \times 10^{-2}$ m/s。

结果：误差与协方差收敛



估计误差始终落在 $\pm 2\sigma$ 协方差边界内；后验标准差迅速下降并趋于稳态，说明滤波器收敛、协方差预测可信。

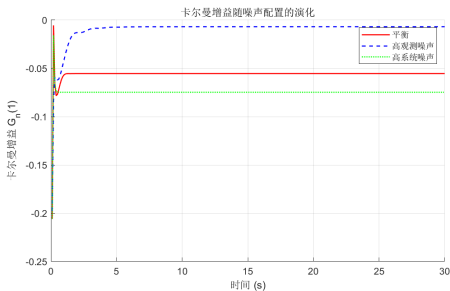
噪声配置对滤波性能的影响



配置	$RMSE_q$	$RMSE_{\dot{q}}$
平衡	$1.0e-2$	$4.6e-2$
高观测噪声	$3.7e-2$	$8.8e-2$
高系统噪声	$1.4e-2$	$1.5e-1$

平衡配置跟踪最好；高观测噪声时滤波器偏信模型、响应滞后；高系统噪声时偏信观测、波动加大。

卡尔曼增益随噪声配置的演化



高系统噪声 $\Rightarrow$ 先验协方差大 $\Rightarrow$ 增益大 (重观测); 高观测噪声 $\Rightarrow$ 增益小 (重模型)。与理论分析一致。

算例二：MEMS 倾角传感与卡尔曼滤波

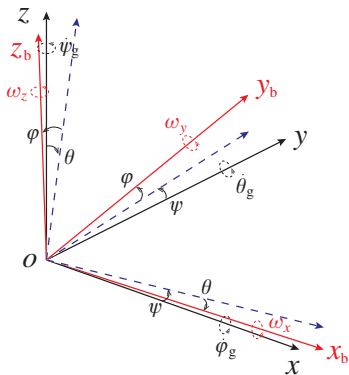
倾角是评估结构安全的关键指标（桥梁支座转角、塔架倾斜、滑坡监测等）。MEMS 传感器具有**低成本、小体积、低功耗**的优点，已成为 SHM 倾角测量的主流方案。

单一 MEMS 器件估角的局限

- **陀螺仪**：测角速度 ω ，积分得倾角——短时精度高，但偏置积累导致**长时漂移**；
- **加速度计**：测重力投影反推倾角——长时无偏，但动态环境下被线加速度污染，**高频噪声大**。

融合思路：用卡尔曼滤波综合二者优势，并**在线估计陀螺仪零偏**以抑制漂移。

陀螺仪估角：积分式



小倾角近似下，欧拉角率约等于体坐标角速度：

$$\dot{\theta}_g \approx \omega_y$$

实测含**零偏** δ_g 与噪声 e_g ：

$$\omega = \omega_{\text{true}} + \delta_g + e_g$$

直接积分 $\Rightarrow$ 零偏被累加 $\Rightarrow$ 倾角估计随时间
线性漂移。

加速度计估角：重力投影

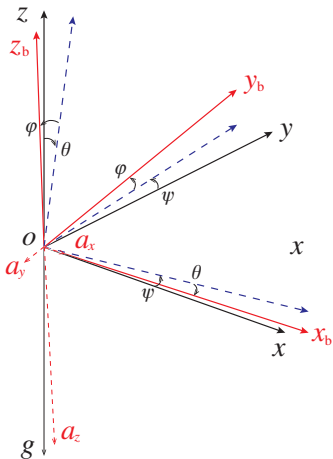
准静态下三轴加速度为重力在体坐标系的投影：

$$\begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \sin \phi \\ \cos \theta \cos \phi \end{pmatrix} g$$

由此反推倾角：

$$\theta_a = \arctan \left(\frac{-a_x}{\sqrt{a_y^2 + a_z^2}} \right)$$

动态工况下 $\mathbf{a}$ 含线加速度分量 $\Rightarrow \theta_a$ 被高频噪声污染，但长时无偏。



卡尔曼融合：状态空间建模

把**倾角**与**陀螺仪零偏**一起作为待估状态；陀螺仪测量作为已知输入，加速度计反推的倾角作为观测：

$$\mathbf{x} = \begin{pmatrix} \theta \\ \delta_{g,y} \end{pmatrix}, \quad u = \omega_y, \quad y = \theta_a$$

卡尔曼融合：状态空间建模

把**倾角**与**陀螺仪零偏**一起作为待估状态；陀螺仪测量作为已知输入，加速度计反推的倾角作为观测：

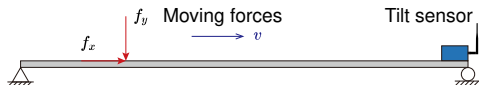
$$\mathbf{x} = \begin{pmatrix} \theta \\ \delta_{g,y} \end{pmatrix}, \quad u = \omega_y, \quad y = \theta_a$$

离散状态空间 (Δt 为采样周期)

$$\mathbf{x}_k = \begin{pmatrix} 1 & -\Delta t \\ 0 & 1 \end{pmatrix} \mathbf{x}_{k-1} + \begin{pmatrix} \Delta t \\ 0 \end{pmatrix} u_{k-1} + \mathbf{w}_{k-1}$$
$$y_k = \begin{pmatrix} 1 & 0 \end{pmatrix} \mathbf{x}_k + v_k$$

物理含义：状态方程为陀螺仪积分扣除零偏；观测方程将加速度计倾角作为对真倾角的含噪测量。 Σ_w 含陀螺仪噪声与零偏游走， Σ_v 取**偏大**以容纳动态线加速度造成的模型不确定性。

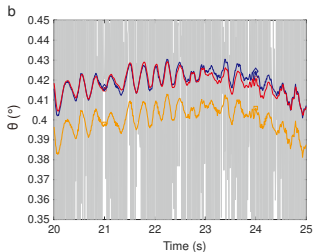
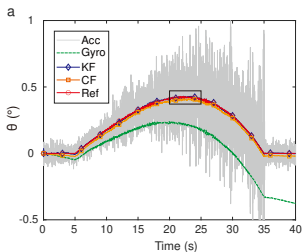
数值仿真：简支梁车致响应



参数	取值
跨度	6 m
截面	30×8 mm
弹模	6 GPa
密度	500 kg/m ³
车速	0.2 m/s
采样	$\Delta t = 0.01$ s

支座处虚拟 IMU 同步采集加速度与角速度，模拟动态倾角响应。

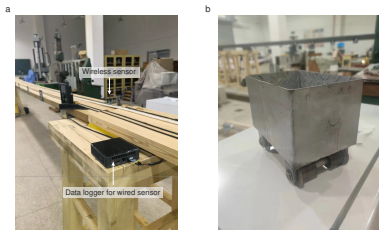
数值仿真：倾角估计对比



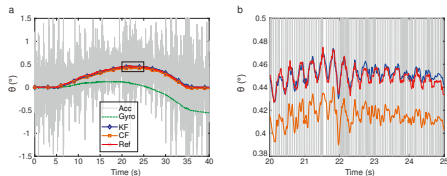
	加速度计 θ_a	陀螺仪 θ_g	卡尔曼 θ_{kf}
RMSE 比	35.50%	49.71%	0.83%

θ_a 高频抖动严重, θ_g 随时间漂移; 卡尔曼估计 θ_{kf} 与参考真值 θ_{ref} 几乎重合
——**在线估出陀螺仪零偏** 是关键。

实验验证：移动车致倾角



(a) 6 m 简支木梁 + 模型车 (5 kg)



(b) 支座倾角估计对比

无线 MEMS 倾角传感节点（陀螺仪 ADXRS290 + 加速度计 ADXL355，采样 100 Hz）与高精度有线倾角计同步采集。卡尔曼滤波估计与参考传感器吻合，RMSE 比 **2.50%**，验证了仿真结论。

扩展：多模型估计 (MMAE)

动机：模型参数不确定时怎么办

标准卡尔曼滤波假定系统矩阵 A_d, C_d 已知且唯一。**现实**中参数常存在不确定性：

- 结构刚度因损伤、温度、老化发生变化；
- 机动目标可能在多种运动模式间切换；
- 传感器特性随工况漂移。

多模型估计 (Multiple-Model Adaptive Estimation, MMAE)

预设 M 个候选模型 $\mathcal{M}_1, \dots, \mathcal{M}_M$ ，对每个模型 **并行运行一个卡尔曼滤波器**；用观测数据在线评估每个模型的可信度（后验概率），**概率会向最贴近真值的模型集中**。

设候选模型 i 的离散状态空间为

$$\mathbf{x}_k = \mathbf{A}_d^{(i)} \mathbf{x}_{k-1} + \mathbf{B}_d^{(i)} \mathbf{u}_{k-1} + \mathbf{w}_{k-1}^{(i)}, \quad \mathbf{y}_k = \mathbf{C}_d^{(i)} \mathbf{x}_k + \mathbf{v}_k^{(i)}$$

噪声 $\mathbf{w}^{(i)} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{w}}^{(i)})$, $\mathbf{v}^{(i)} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{v}}^{(i)})$ 。

对每个模型 i 同时跑一个 KF，得到

- 后验估计 $\hat{\mathbf{x}}_{k|k}^{(i)}$ 与协方差 $\Sigma_{\mathbf{x},k|k}^{(i)}$ ；
- **预测残差** $\mathbf{e}_k^{(i)} = \mathbf{y}_k - \mathbf{C}_d^{(i)} \hat{\mathbf{x}}_{k|k-1}^{(i)}$ ；
- 预测残差协方差 $\mathbf{s}_k^{(i)} = \mathbf{C}_d^{(i)} \Sigma_{\mathbf{x},k|k-1}^{(i)} (\mathbf{C}_d^{(i)})^T + \Sigma_{\mathbf{v}}^{(i)}$ 。

$\mathbf{e}^{(i)}$ 在模型 i 假设下应为零均值高斯——观测越偏离这一假设，说明模型 i 越不可信。

在模型 i 假设下，第 k 步观测的**似然**为预测残差的高斯密度：

$$\Lambda_k^{(i)} = p(\mathbf{y}_k \mid \mathcal{M}_i, \mathbf{y}_{1:k-1}) = \frac{1}{\sqrt{(2\pi)^p |\mathbf{S}_k^{(i)}|}} \exp\left(-\frac{1}{2}(\mathbf{e}_k^{(i)})^T (\mathbf{S}_k^{(i)})^{-1} \mathbf{e}_k^{(i)}\right)$$

在模型 i 假设下，第 k 步观测的**似然**为预测残差的高斯密度：

$$\Lambda_k^{(i)} = p(\mathbf{y}_k \mid \mathcal{M}_i, \mathbf{y}_{1:k-1}) = \frac{1}{\sqrt{(2\pi)^p |\mathbf{S}_k^{(i)}|}} \exp\left(-\frac{1}{2}(\mathbf{e}_k^{(i)})^T (\mathbf{S}_k^{(i)})^{-1} \mathbf{e}_k^{(i)}\right)$$

由贝叶斯公式递推模型后验概率 $p_k^{(i)} = \Pr(\mathcal{M}_i \mid \mathbf{y}_{1:k})$ ：

$$p_k^{(i)} = \frac{\Lambda_k^{(i)} p_{k-1}^{(i)}}{\sum_{j=1}^M \Lambda_k^{(j)} p_{k-1}^{(j)}} \quad p_0^{(i)} = \frac{1}{M}$$

MMAE 理论：贝叶斯模型更新

在模型 i 假设下，第 k 步观测的**似然**为预测残差的高斯密度：

$$\Lambda_k^{(i)} = p(\mathbf{y}_k \mid \mathcal{M}_i, \mathbf{y}_{1:k-1}) = \frac{1}{\sqrt{(2\pi)^p |\mathbf{S}_k^{(i)}|}} \exp\left(-\frac{1}{2}(\mathbf{e}_k^{(i)})^T (\mathbf{S}_k^{(i)})^{-1} \mathbf{e}_k^{(i)}\right)$$

由贝叶斯公式递推模型后验概率 $p_k^{(i)} = \Pr(\mathcal{M}_i \mid \mathbf{y}_{1:k})$ ：

$$p_k^{(i)} = \frac{\Lambda_k^{(i)} p_{k-1}^{(i)}}{\sum_{j=1}^M \Lambda_k^{(j)} p_{k-1}^{(j)}} \quad p_0^{(i)} = \frac{1}{M}$$

MMAE 加权状态估计

$$\hat{\mathbf{x}}_k = \sum_{i=1}^M p_k^{(i)} \hat{\mathbf{x}}_{k|k}^{(i)}$$

渐近性质：当真模型在候选集内时， $p_k^{(i^*)} \rightarrow 1$ ，其余 $\rightarrow 0$ ；多模型 KF 退化为最优单一 KF。

每个时间步 k

1. **并行 KF**: 对每个模型 $i = 1, \dots, M$ 执行预测/校正, 得到 $\hat{\mathbf{x}}_{k|k}^{(i)}, \Sigma_{\mathbf{x}, k|k}^{(i)}, \mathbf{e}_k^{(i)}, \mathbf{s}_k^{(i)}$;
2. **似然计算**: $\Lambda_k^{(i)} = \mathcal{N}(\mathbf{e}_k^{(i)}; \mathbf{0}, \mathbf{s}_k^{(i)})$;
3. **概率更新**: $p_k^{(i)} \propto \Lambda_k^{(i)} p_{k-1}^{(i)}$, 再归一化;
4. **加权融合**: $\hat{\mathbf{x}}_k = \sum_i p_k^{(i)} \hat{\mathbf{x}}_{k|k}^{(i)}$.

实现要点:

- 用**对数似然** $\ln \Lambda$ 并经 log-sum-exp 归一化, 防止数值下溢;
- 设定概率**下限** $p_{\min}$ (如 0.5% ~ 2%), 防止某模型概率早期跌至 0 后无法更新 (**锁定**问题);
- 观测矩阵 $\mathbf{C}_d^{(i)}$ 应**尽量与待识别参数无关**, 否则似然的归一化项会引入系统偏差。

算例三：4 自由度剪切结构

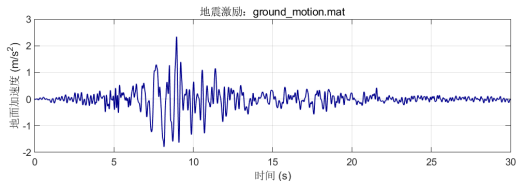
结构参数：

$$\mathbf{M} = 10\mathbf{I}_4 \text{ kg}, \quad \mathbf{K} = 1500 \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \text{ N/m}$$

Rayleigh 阻尼，前两阶模态阻尼比 2%。固有频率

$\omega \approx [4.3, 12.3, 18.8, 23.0] \text{ rad/s}$ 。

激励：实测地震加速度， $T = 0.01 \text{ s}$ ， $N = 3000$ 。观测：4 层位移 + 高斯噪声 $\sigma_v = 5 \text{ mm}$ 。



算例三：候选模型库

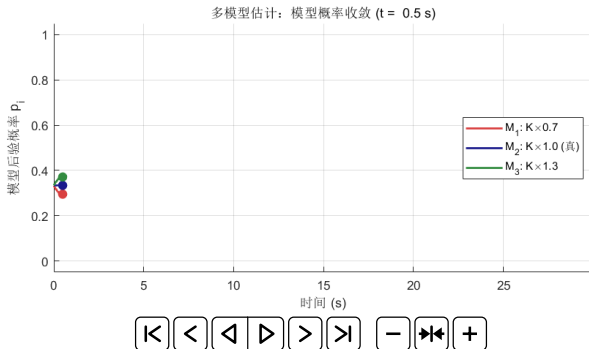
设结构刚度未知，假定 3 个候选模型（典型损伤/温度场景）：

$$\mathcal{M}_1 : \mathbf{K} \times 0.7; \quad \mathcal{M}_2 : \mathbf{K} \times 1.0 \text{ (真)}; \quad \mathcal{M}_3 : \mathbf{K} \times 1.3$$

每个模型构造对应的 $\mathbf{A}_d^{(i)}, \mathbf{B}_d^{(i)}$ (ZOH 离散化)，观测矩阵 $\mathbf{C}_d = [\mathbf{I}_4, \mathbf{0}]$ 与刚度无关，保证似然比较公平。

初始 $p_0^{(i)} = 1/3$ ，过程噪声协方差 $\Sigma_w = 10^{-4} \mathbf{I}$ ，概率下限 $p_{\min} = 0.5\%$ 。

算例三：模型概率收敛（动画）



- $t < 7 \text{ s}$: 地震激励小, 结构响应主要为噪声, 三模型概率**近似均匀** (信息不足);
- $t \approx 9 \text{ s}$: 地震主震到达, 结构进入大幅响应, 预测残差开始 **区分模型**;
- $t \in [10, 20] \text{ s}$: 自由振荡区, 真模型 M_2 概率升至 ≈ 0.95 , **正确识别**;
- $t > 25 \text{ s}$: 响应衰减回噪声水平, 模型再次趋于不可分。

小结

- **最小二乘**: 由正规方程得 $\hat{\mathbf{x}} = (\mathbf{C}_d^T \mathbf{C}_d)^{-1} \mathbf{C}_d^T \mathbf{y}$; 加权最小二乘以噪声协方差逆为权;
- **递归最小二乘**: 逐次更新, 无偏, 最优增益由误差协方差极小化得到;
- **状态传递**: 高斯分布经线性系统仍为高斯, 均值与协方差按系统状态方程递推;
- **离散卡尔曼滤波**: 递归最小二乘 + 状态传递, 交替预测与校正; 卡尔曼增益在模型与观测间最优加权;
- **多模型估计 (MMAE)**: 并行多组 KF, 通过贝叶斯模型概率更新, 自动识别最贴近真值的模型。

三个算例表明: 单自由度系统中即使初值误差大、仅测加速度, 滤波器仍能准确重构位移与速度; MEMS 倾角传感中将**陀螺仪零偏**纳入状态后可抑制漂移; 剪切结构中 MMAE 能在线识别刚度变化, 提供结构损伤诊断的基础工具。

谢谢

Q & A