

1. 问题建模 (Problem Formulation)

在原论文中，误差动力学被建模为 $e_{t+1} = A(e_t)e_t + \xi'_t$ ，其中 ξ'_t 是均值为 0 的无偏噪声。但在现实场景中（例如使用了正则化的 MLE、贝叶斯估计或有偏差的预训练模型），估计器往往带有系统性偏差（**Systematic Bias**）。

我们重新定义非线性随机差分系统如下：

$$\mathbf{e}_{t+1} = A(\mathbf{e}_t)\mathbf{e}_t + \mathbf{b}(\mathbf{e}_t) + \xi_t$$

其中：

- $\mathbf{e}_t = \theta_t - \theta^*$: 第 t 步的参数估计误差。
- $A(\mathbf{e}_t)$: 由 Filter 诱导出的状态依赖收缩算子（State-dependent Contraction Operator）。
- $\mathbf{b}(\mathbf{e}_t)$: 系统性偏差项。这是一个确定性但未知的向量，代表估计器固有的偏移。
- ξ_t : 纯随机估计噪声，满足 $\mathbb{E}[\xi_t | \mathcal{F}_t] = 0$ 且 $\mathbb{E}[\|\xi_t\|^2] \leq \sigma^2$ 。

2. 假设条件 (Assumptions)

为了进行理论分析，我们需要对偏差和收缩算子做出以下假设：

假设 2.3 (Revised): 强化收缩条件 (Strong Contraction)

存在正定矩阵 $P \succ 0$ 和连续函数 $c: \mathbb{R}^p \rightarrow [0, 1]$ ，使得对于任意 \mathbf{e} :

$$A(\mathbf{e})^\top P A(\mathbf{e}) \preceq (1 - c(\mathbf{e}))P$$

且存在凸函数 $f(\cdot)$ 使得 $c(\mathbf{e})V(\mathbf{e}) \geq f(V(\mathbf{e}))$ ，其中 $V(\mathbf{e}) = \mathbf{e}^\top P \mathbf{e}$ 。

假设 2.14 (New): 有界偏差 (Bounded Bias)

系统性偏差 $\mathbf{b}(\mathbf{e}_t)$ 在加权范数下是有界的：

$$\sup_{\mathbf{e}} \|\mathbf{b}(\mathbf{e})\|_P \leq \beta$$

其中 $\|\mathbf{x}\|_P = \sqrt{\mathbf{x}^\top P \mathbf{x}}$ 。同时，噪声方差有界 $\mathbb{E}[\xi_t^\top P \xi_t] \leq \sigma^2$ 。

为了简化符号，我们将总干扰项记为 $\epsilon = \beta^2 + \sigma^2$ （偏差能量 + 噪声能量）。

3. 理论推导

步骤 1：能量函数的条件期望展开

考察下一时刻误差能量的条件期望 $\mathbb{E}[V(\mathbf{e}_{t+1})|\mathcal{F}_t]$ 。

将动力学方程 $\mathbf{e}_{t+1} = A(\mathbf{e}_t) + \mathbf{b}(\mathbf{e}_t) + \xi_t$ 代入：

$$\begin{aligned}\mathbb{E}[V_{t+1}|\mathbf{e}_t] &= \mathbb{E}[(A\mathbf{e}_t + \mathbf{b} + \xi_t)^\top P(A\mathbf{e}_t + \mathbf{b} + \xi_t) | \mathbf{e}_t] \\ &= \underbrace{(A\mathbf{e}_t)^\top P(A\mathbf{e}_t)}_{\text{基础收缩项}} + \underbrace{\mathbf{b}^\top P\mathbf{b} + \mathbb{E}[\xi_t^\top P\xi_t]}_{\text{常数干扰项}} + \underbrace{2(A\mathbf{e}_t)^\top P\mathbf{b}}_{\text{交叉项}}\end{aligned}$$

这里利用了 $\mathbb{E}[\xi_t] = 0$ 消去了包含 ξ_t 的一次项。

根据假设 2.14 (有界偏差与噪声)，令 $\epsilon_{total} = \beta^2 + \sigma^2$ ，则常数干扰项 $\leq \epsilon_{total}$ 。

步骤 2：交叉项的线性化处理 (Young's Inequality)

交叉项 $2(A\mathbf{e}_t)^\top P\mathbf{b}$ 是导致分析困难的主要原因。为了避免引入非线性的平方根项，我们使用带参数 $\eta > 0$ 的 **Young's Inequality** 进行放缩：

$$2x^\top Py \leq \eta x^\top Px + \frac{1}{\eta} y^\top Py$$

将其应用于交叉项 (令 $x = A\mathbf{e}_t, y = \mathbf{b}$)：

$$\begin{aligned}2(A\mathbf{e}_t)^\top P\mathbf{b} &\leq \eta(A\mathbf{e}_t)^\top P(A\mathbf{e}_t) + \frac{1}{\eta} \mathbf{b}^\top P\mathbf{b} \\ &\leq \eta(A\mathbf{e}_t)^\top P(A\mathbf{e}_t) + \frac{1}{\eta} \beta^2\end{aligned}$$

将此结果代回主不等式，合并 $(A\mathbf{e}_t)^\top P(A\mathbf{e}_t)$ 项：

$$\begin{aligned}\mathbb{E}[V_{t+1}|\mathbf{e}_t] &\leq (1 + \eta) \underbrace{(A\mathbf{e}_t)^\top P(A\mathbf{e}_t)}_{\text{应用收缩假设 2.3}} + \underbrace{\left(1 + \frac{1}{\eta}\right)\beta^2 + \sigma^2}_{\text{定义为常数 } C_\eta} \\ &\leq (1 + \eta)(V_t - f(V_t)) + C_\eta\end{aligned}$$

此时，我们得到了关于 V_t 的条件递推不等式。

步骤 3：从条件期望到全期望 (Jensen's Inequality)

为了分析系统的全局收敛性，我们需要对不等式两边取无条件期望 (**Unconditional Expectation**)。

令 $x_t = \mathbb{E}[V(\mathbf{e}_t)]$ 表示第 t 步的平均误差能量。

对不等式两边取 $\mathbb{E}[\cdot]$ ：

$$x_{t+1} \leq (1 + \eta)(x_t - \mathbb{E}[f(V_t)]) + C_\eta$$

由于 $f(\cdot)$ 是定义在凸集上的**凸函数 (Convex Function)** (假设 2.5)，根据 **Jensen 不等式**：

$$\mathbb{E}[f(V_t)] \geq f(\mathbb{E}[V_t]) = f(x_t)$$

因此有 $-\mathbb{E}[f(V_t)] \leq -f(x_t)$ 。代入上式得到确定性递推关系：

$$x_{t+1} \leq (1 + \eta)(x_t - f(x_t)) + C_\eta$$

步骤 4：一致最终有界性分析 (UUB Analysis)

系统收敛意味着能量不再增长，即 $x_{t+1} < x_t$ 。我们需要找到满足该条件的区域。

考察能量差分：

$$x_{t+1} - x_t \leq \eta x_t - (1 + \eta)f(x_t) + C_\eta$$

要使误差下降 ($x_{t+1} - x_t < 0$)，必须满足：

$$(1 + \eta)f(x_t) > \eta x_t + C_\eta$$

即收缩力必须大于扩张力与常数干扰之和：

$$f(x_t) > \frac{\eta}{1 + \eta}x_t + \frac{C_\eta}{1 + \eta}$$

定理结论 (Theorem):

在有偏估计下，系统的期望误差能量是一致最终有界的 (Uniformly Ultimately Bounded)。误差序列 $\{x_t\}$ 最终将收敛至集合：

$$\mathcal{S} = \{x \in \mathbb{R}_{\geq 0} \mid x \leq R^*\}$$

其中 R^* 是方程 $f(x) = \frac{\eta}{1 + \eta}x + \frac{C_\eta}{1 + \eta}$ 的最大实根。

物理意义解释：

- 如果不等式左边 $f(x)$ 的增长速度快于右边的线性项（例如 $f(x)$ 是超线性的，或者即使是线性但斜率足够大），则必然存在一个交点 R^* 。
- 当误差 $x_t > R^*$ 时，Filter 提供的强收缩力 (Strong Contraction) 占据主导，将误差拉回。
- 当误差 $x_t \leq R^*$ 时，偏差和噪声占据主导，误差无法进一步缩小至 0，而是在 R^* 范围内波动。这证明了在有偏情况下，模型虽然无法完美复原真值，但可以避免崩溃。