arXiv:2303.02324v1 [eess.SP] 4 Mar 2023

Mo快速接近物体的建模与最快变化检测

Tim Brucks[a]、Taposh Banerjee[b]和Rahul Mishra[c]

[a]圣安东尼奥德州大学；[b]匹兹堡大学；[c]印度班加罗尔U. R. Rao卫星中心。

文章历史
编写于2023年3月7日

摘要
本文研究如何检测可能导致灾难的信号。决策者按时间顺序收集数据。在某个时刻（称为变化点），数据的分布会发生变化。这种分布变化可能由某个事件或敌方物体突然出现所引起。如果未能及时检测到，这种变化可能会酿成重大灾难。在太空和军事应用中，随着敌方物体逐渐接近目标，测量值可能会随时间随机增大。本文引入了一类称为爆炸过程的新型随机过程，用于对随机增长的数据进行建模。文中提出了一种算法，并证明当平均虚警时间趋于无穷大时，该算法渐近最优。

关键词
最快变化检测、非平稳过程、单调似然比、随机占优、极小极大最优性。

1. 引言

在最快变化检测（QCD）问题中，决策者会收集一系列测量值。该序列中的数值被视为某个随机过程的实现。假设该随机过程的分布律最初遵循一个被认为是正态分布的分布。在某个时刻（称为变化点），测量值的统计性质或过程的分布律会发生变化。QCD 问题的目标是在尽可能避免误报的同时，尽快检测到这一分布变化 [12, 15, 19]。最常见的变点模型是突变且持续的变化模型 [5, 9, 10, 14, 16, 17]。在该模型中，过程的分布律在变化点处突然改变，并且此后永久保持新的分布律。例如，在图 1（左）中，我们绘制了一列高斯随机变量的均值。在变化点（图中为 80）处，均值从 0 突然变为 4，并一直保持为 4。

在某些军事和太空应用中，变化后分布律可能具有爆炸性特征。

(1) 空间科学家关注对接近卫星并可能将其摧毁的碎片或其他危险物体的探测（见图2）。随着目标物体迅速接近卫星，测量值的均值将呈现

联系 Taposh Banerjee。电子邮箱：taposh.banerjee@pitt.edu

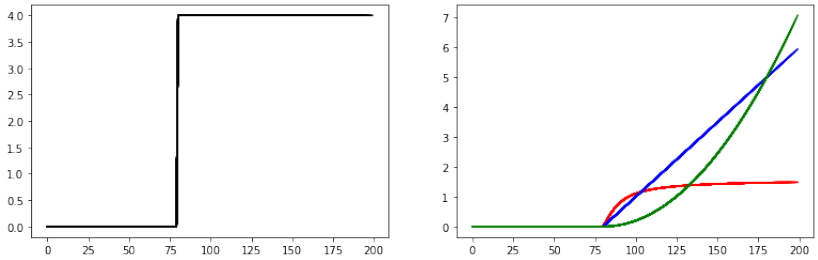

图 1. 左图：经典的突变点模型。图中展示了一列高斯随机变量的均值。变化前均值为 0，变化后为 4。右图：在某些空间或军事应用中，变化后的测量值可能会呈现爆发式增长（线性、超线性或次线性）。随着敌方物体逐渐靠近，测量值可能会随机增大。



图 2. 空间科学家关注碎片或其他危险物体接近卫星并将其摧毁的情况。导弹（或鱼雷）可以迅速接近目标（或潜艇）并将其摧毁。来源：https://images.google.com/。

如图1（右）所示的爆炸式变化。在第1.1节中，我们将对此应用进行详细讨论。

(2) 另一个经典示例是军事应用中的敌方物体检测。当导弹接近目标，或鱼雷接近潜艇或舰船时，测量值预计会随时间增大。

本文提出了一类新的随机过程，用以刻画过程随机增长的特性。随后，我们针对这类新过程，得到了一个用于检测分布变化的最优算法。

1.1. 卫星应用

本节简要讨论卫星安全这一问题，它是本文所研究的最快变化检测问题的主要研究动机。

对太空中接近物体的检测与跟踪有许多具体应用，如下所示：

(1) 在图 3（左侧和底部）中，我们给出了地球轨道上物体的尺寸和密度。这些物体可能会对执行舱外活动（EVA）的宇航员构成危险，也可能损害空间资源，例如在轨运行卫星和国际空间站。在图 3（右侧）中，我们可以看到碎片撞击国际空间站造成的影响。在 [8, 20], 中，作者绘制了碎片的分布图并对其进行检测

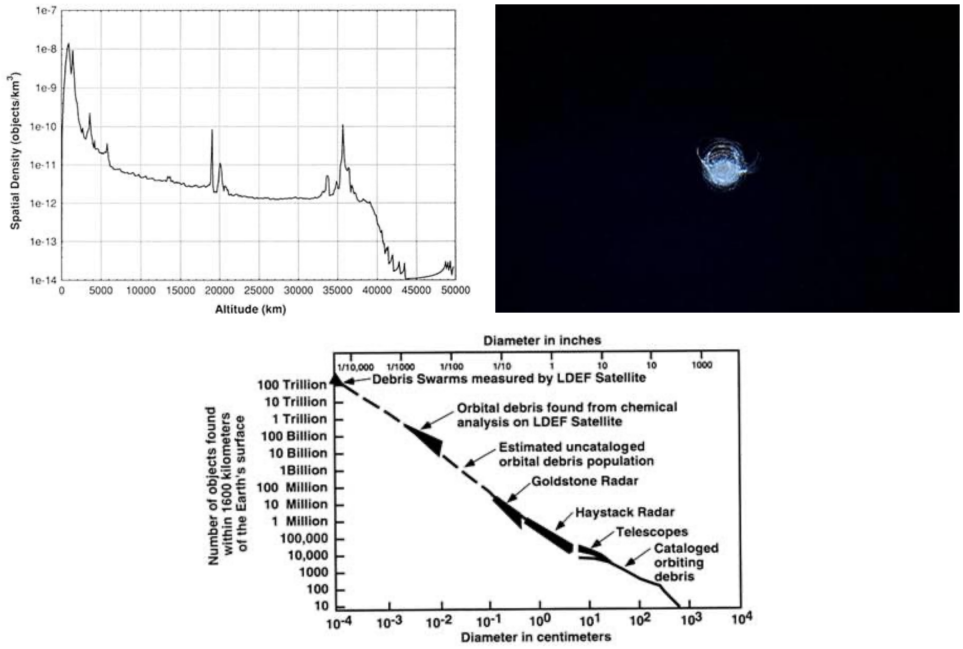

图3。左图：地球轨道中的碎片密度；右图：国际空间站遭碎片撞击；下图：碎片尺寸分布。

利用机载系统。

(2) 以下是一些由太空碰撞和空间碎片导致的重大空间事故：a) 太空垃圾损坏了国际空间站内实验舱的机械臂；机械臂的热防护层上发现了一处穿孔。b) 1996年，法国 Cerise 卫星与一块已编目的阿丽亚娜火箭碎片相撞，导致该卫星的稳定杆断裂。c) 铱星33（商业通信网络的一部分）与宇宙2251号相撞。这是地球轨道上两颗卫星因轨道拥堵且失效卫星未离轨而发生的首次重大卫星碰撞。此次碰撞产生了1000多颗碎片，可能进一步加剧碎片碰撞。d) 2021年3月，云海一号02星与天顶二号火箭的碎片发生碰撞，产生了20多块碎片。

（3）碎片：人造空间碎片与流星体统称为 MMOD（微流星体和轨道碎片）。这些物体可造成喷砂效应，尤其会影响难以通过任何防护罩保护的航天器附属部件和光学设备 [1]。空间碎片的一些类型如下：a）备用航天器部件，b）离轨卫星，以及 c）运载火箭部件。

（4）诸如在役卫星和失效卫星及其部件、运载火箭部件等轨道物体，以及大多数编目物体的状态，预计会在可利用本文所述统计方法充分建模的容差范围内确定性地演变。这些合理假设被用于得出有用结果，以实现未来卫星的自主能力。因此，本文结果可证明对自主流星体规避机制有所助益。

(5) 本研究旨在完善更贴近现实的假设，以涵盖凯斯勒综合征所导致的碎片以及智能追踪器等情形

系统，这些系统可能位于太空或地面，并可能对空间资源构成潜在威胁，造成灾难。

(6) 本工作的直接应用是自主航天器运行，这能为航天工业带来显著的运行或商业优势。空间运行的复杂性是卫星服务面临的主要开销之一，因此，控制这一部分对于在当前航天领域的竞争中取得优势至关重要。

(7) 最近的一个例子是，欧洲航天局的一颗卫星需要进行机动，以避免可能与 SPACEX 的星链卫星发生碰撞。

2. 模型与问题表述

2.1. 数据模型

我们引入一类新的随机过程，用于描述变化后过程的爆炸特性：

定义 2.1. 如果过程 $\{X_n\}_{n\geq 1}$ 的密度为 $\{f_n\}_{n\geq 0}$，则称其为爆炸过程，若

(1) $\{X_n\}$ 联合独立，(2) $X_n \sim f_{n-1}, \forall n$，(3) $\frac{f_{n+1}(x)}{f_n(x)}$ 关于 $x, \forall n$递增。我们用 $f_n \prec f_{n+1}$ 表示这种单调似然比（MLR）序。

如果对所有 n都有 $f_{n+1} = f_n$，那么我们得到一个独立同分布（i.i.d.）过程。请注意，MLR 占优意味着随机占优 [4]:

$$\int_x^\infty f_n(x)dx \leq \int_x^\infty f_{n+1}(x)dx, \quad \forall x, n.$$

2.2. 变点模型

我们假设变化发生之前，数据是密度为 $g \prec f_0$的独立同分布序列；变化发生之后，则是密度序列为 $\{f_n\}_{n\geq 0}$ 的爆炸过程。数学上，存在一个离散时间 ν，使得

$$X_n \sim \begin{cases} g, & \forall n < \nu, \\ f_{n-\nu} & \forall n \geq \nu. \end{cases}$$

这里我们使用记号 $X \sim f$ 表示随机变量 X 服从分布律f。请注意，观测值在变化发生后的密度取决于变化点 ν的位置。我们的目标是在满足误报率约束的前提下，尽可能快地检测到这一变化。

2.3. 问题表述

为了解决变化检测问题，我们关注最快变化检测的两种常用极小极大公式。为表述这些公式，我们定义：

对于 $1 \le \nu \le \infty$，将 E_ν 定义为变化发生在时间 ν 时的期望。我们考虑 Lorden [7]的问题公式：

$$\begin{aligned} \min_{\tau} \quad & \sup_{\nu \ge 1} \operatorname{ess\,sup} \mathsf{E}_\nu[(\tau - \nu + 1)^+ | X_1, \ldots, X_{\nu-1}], \\ \text{subj. to} \quad & \mathsf{E}_\infty[\tau] \ge \gamma, \end{aligned} \tag{1}$$

其中，γ 是对平均虚警时间的约束。我们还将考虑 Pollak [11]的表述：

$$\begin{aligned} \min_{\tau} \quad & \sup_{\nu \ge 1} \mathsf{E}_\nu[\tau - \nu | \tau \ge \nu], \\ \text{subj. to} \quad & \mathsf{E}_\infty[\tau] \ge \gamma, \end{aligned} \tag{2}$$

其中，γ 同样是对平均虚警时间的约束。

3. 候选算法：爆炸累积和算法

我们提出使用以下爆炸累积和（EX-CUSUM）统计量进行变化检测：

$$W_n = \max_{1 \le k \le n} \sum_{i=k}^{n} \log \frac{f_{i-k}(X_i)}{g(X_i)}. \tag{3}$$

项 $\sum_{i=k}^{n} \log \frac{f_{i-k}(X_i)}{g(X_i)}$ 是观测值在变化后分布与变化前分布之间的对数似然比，条件是变化发生在时间 k。由于我们不知道变化点，因此取时间 n时所有可能取值中的最大值，即 $1 \le k \le n$。

In 对于上述统计量，请注意，观测值 X_i 在时间 e i,

$$\frac{f_{i-k}(X_i)}{g(X_i)},$$

取决于时间 i 与变化点假设 k 之间的相对距离 $i - k$。正因如此，EX-CUSUM 统计量并非 Lai [5] 的广义CUSUM统计量的一个特例。

为了检测分布律从独立同分布 g （分布律为 g ）变为爆炸过程 $\{f_n\}$，当 EX-CUSUM 统计量首次超过阈值 A时，我们停止：

$$\tau_{ec} = \inf\{n \ge 1 : W_n > A\}. \tag{4}$$

我们选择阈值 A 来控制误报率：阈值越高，平均虚警时间 $\mathsf{E}_\infty[\tau_{ec}]$ 越长（这一点将在下文正式证明）。

本文旨在刻画以下条件：在这些条件下，EX-CUSUM 算法对于式（1）和式（2）中的 Lorden 问题与 Pollak 问题是渐近最优的。

4. 性能的渐近下界

4.1. Lai 渐近下界

在 [5], 中，Lai 提出了最快变化检测的一般极小极大理论。我们首先对其进行回顾，并讨论其局限性。

假设在 [5] 时，变化前密度为 $f_0(X_i|X_1, \ldots, X_{i-1})$，变化后密度为 $f_1(X_i|X_1, \ldots, X_{i-1})$，由此得到对数似然比

$$Z_i = \log \frac{f_1(X_i|X_1, \ldots, X_{i-1})}{f_0(X_i|X_1, \ldots, X_{i-1})}.$$

请注意，这些密度是一般的条件密度，允许存在数据依赖性，但并非变化点假设的函数（如论文中所定义）。Lai 证明，如果 Z_is 满足存在一个信息数 $I > 0$ s，使得

$$\lim_{n\to\infty} \sup_{\nu\geq 1} \operatorname{ess\,sup} \mathsf{P}_\nu \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_i \geq I(1+\delta)n \;\middle|\; X_1, \ldots, X_{\nu-1} \right) = 0, \tag{5}$$

那么我们有普适下界 $\gamma \to \infty$,

$$\begin{aligned} \min_{\tau} \sup_{\nu\geq 1} \operatorname{ess\,sup} \mathsf{E}_\nu[(\tau-\nu+1)^+|X_1, \ldots, X_{\nu-1}] \\ \geq \quad \min_{\tau} \sup_{\nu\geq 1} \mathsf{E}_\nu[\tau-\nu|\tau\geq\nu] \\ \geq \quad \frac{\log\gamma}{I}(1+o(1)). \end{aligned} \tag{6}$$

这里，对 τ 的最小值是在满足 $\mathsf{E}_\infty[\tau] \geq \gamma$ 的停止时刻范围内取的。Lai 进一步证明，在对 Z_is 施加某些额外条件时，广义 CUSUM 算法,

$$\tau_c = \min\left\{ n \geq 1 : \max_{1\leq k\leq n} \sum_{i=k}^{n} Z_i \geq \log(\gamma) \right\},$$

对于 Lorden 和 Pollak 的问题均为渐近最优。具体而言,

$$\mathsf{E}_\infty[\tau_c] \geq \gamma.$$

此外，如果 Z_is 满足

$$\lim_{n\to\infty} \sup_{k\geq\nu\geq 1} \operatorname{ess\,sup} \mathsf{P}_\nu \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_i \leq I-\delta \;\middle|\; X_1, \ldots, X_{k-1} \right) = 0, \tag{7}$$

于是，当 $\gamma \to \infty$、τ_c 达到下界时:

$$\sup_{\nu\geq 1} \operatorname{ess\,sup} \mathsf{E}_\nu[(\tau_c-\nu+1)^+|X_1, \ldots, X_{\nu-1}] \leq \frac{\log\gamma}{I}(1+o(1)), \quad \gamma\to\infty. \tag{8}$$

目前尚不清楚这些结果对于爆炸过程设定是否有效，因为在后一种设定中，似然比确实取决于变化点所在的位置。接下来我们将对此进行讨论。

4.2. 变化点依赖似然的充分条件

本节将 Lai 的结果推广到似然随变化点而变的情形。为陈述更一般的结果，我们继续允许数据存在时间依赖性。定义在时间 n 、变化发生于 ν 时的对数似然比为

$$Z_{n,\nu} = \log \frac{f_{n,\nu}(X_n|X_1,\ldots,X_{n-1})}{f_0(X_n|X_1,\ldots,X_{n-1})}.$$

定理 4.1。 *(1)* 设存在一个正数 I ，使得双变量对数似然比 $\{Z_{n,\nu}\}$ 满足以下条件:

$$\lim_{n\to\infty} \sup_{\nu\geq 1} \mathit{ess\ sup}\ \mathsf{P}_\nu \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_{i,\nu} \geq I(1+\delta)n \;\middle|\; X_1,\ldots,X_{\nu-1} \right) = 0. \tag{9}$$

则有如下普适下界: $\gamma \to \infty$,

$$\begin{aligned} \min_{\tau} \sup_{\nu\geq 1} \mathit{ess\ sup}\, \mathsf{E}_\nu[(\tau-\nu+1)^+|X_1,\ldots,X_{\nu-1}] \\ \geq \quad \min_{\tau} \sup_{\nu\geq 1} \mathsf{E}_\nu[\tau-\nu|\tau\geq\nu] \\ \geq \quad \frac{\log\gamma}{I}(1+o(1)). \end{aligned} \tag{10}$$

此处对 τ 取最小值，是在满足$\mathsf{E}_\infty[\tau] \geq \gamma$的停止时刻中取最小值。*(2)* 以下修正后的广义 *CUSUM* 算法，

$$\tau_{mc} = \min\left\{ n\geq 1 : \max_{1\leq k\leq n} \sum_{i=k}^{n} Z_{i,k} \geq \log(\gamma) \right\},$$

满足

$$\mathsf{E}_\infty[\tau_{mc}] \geq \gamma.$$

(3) 设二元对数似然比 $\{Z_{n,\nu}\}$ 还满足

$$\lim_{n\to\infty} \sup_{k\geq\nu\geq 1} \mathit{ess\ sup}\ \mathsf{P}_\nu \left(\frac{1}{n}\sum_{i=k}^{k+n} Z_{i,k} \leq I-\delta \;\middle|\; X_1,\ldots,X_{k-1} \right) = 0. \tag{11}$$

于是， $\gamma\to\infty$、 τ_{mc} 达到了下界:

$$\sup_{\nu\geq 1} \mathit{ess\ sup}\, \mathsf{E}_\nu[(\tau_{mc}-\nu+1)^+|X_1,\ldots,X_{\nu-1}] \leq \frac{\log\gamma}{I}(1+o(1)), \quad \gamma\to\infty. \tag{12}$$

证明。 将 [5] 中的 Z_is 替换为 $Z_{n,\ \nu}$s 后，下界结果仍然成立。这是因为证明依赖于测度变换论证来得到下界。由于此处的似然比是 $Z_{n,\ \nu}$ 的函数，相同的测度变换论证仍然适用。只要满足上述条件，并且在所有地方将 Z_i 替换为 $Z_{i,\ k}$，检测延迟的证明也仍然成立。作者不确定 Lai 关于平均虚警时间结果的证明能否推广到时变情形。不过，可以采用另一种基于 Shiryaev-Roberts 统计量和可选抽样定理的技术。由于后一种技术是经典方法[15], ，这里省略细节。

□

本文撰写期间，作者获悉了另一篇论文[6] ，该文确立了更一般的最优性充分条件（比 [5] 中的条件更一般）。与 [6], 相比，我们提出了一种新颖的方法（使用 MLR序）来刻画随机递增的非平稳变后过程。此外，我们用于验证这些最优性条件的证明技巧略有不同，也具有独立的研究价值。

5. EX-CUSUM 算法的最优性

本节首先简化定理 4.1 中针对定义 2.1 所定义爆炸过程的条件，然后对这些简化后的条件作进一步说明，以保证 EX-CUSUM 算法的最优性。

5.1. 简化爆炸过程的下界条件

下文的定理表明，对于爆炸过程，下界条件 (9) 可以简化。

定理 5.1. 满足

$$\lim_{n\to\infty} \sup_{\nu\geq 1} \ ess\ sup\ \mathsf{P}_\nu \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_{i,\nu} \geq I(1+\delta)n \ \middle|\ X_1, \ldots, X_{\nu-1} \right) = 0 \qquad (13)$$

对于某个 $0 < I < \infty$，只需满足

$$\frac{1}{n}\sum_{k=1}^{n} Z_{k,1} = \frac{1}{n}\sum_{k=1}^{n} \log \frac{f_{k-1}(X_k)}{g(X_k)} \ \to\ I, \quad a.s.\ under\ \mathsf{P}_1.$$

证明。 由于独立性，我们有

$$\begin{aligned} &\sup_{\nu\geq 1} \operatorname{ess\,sup}\ \mathsf{P}_\nu \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_{i,\nu} \geq I(1+\delta)n \ \middle|\ X_1, \ldots, X_{\nu-1} \right) \\ &\qquad = \sup_{\nu\geq 1} \mathsf{P}_\nu \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} \log \frac{f_{i-\nu}(X_i)}{g(X_i)} \geq I(1+\delta)n \right). \end{aligned}$$

由于似然比是相对于变化点 ν计算的，右侧的概率不是 ν 的函数。这意味着

$$\sup_{\nu\geq 1}\ \mathsf{P}_\nu\left(\max_{t\leq n}\sum_{i=\nu}^{\nu+t}\log\frac{f_{i-\nu}(X_i)}{g(X_i)}\geq I(1+\delta)n\right)=\mathsf{P}_1\left(\max_{t\leq n}\sum_{i=1}^{1+t}\log\frac{f_{i-1}(X_i)}{g(X_i)}\geq I(1+\delta)n\right)$$
$$=\mathsf{P}_1\left(\frac{1}{n}\max_{t\leq n}\sum_{i=1}^{1+t}\log\frac{f_{i-1}(X_i)}{g(X_i)}\geq I(1+\delta)\right).$$

请注意，若变化后过程对于不同的变化点以不同方式演化，则上述简化可能不成立。现在，若

$$\frac{1}{n}\sum_{i=1}^{n}\log\frac{f_{i-1}(X_i)}{g(X_i)}\ \to\ I,\quad \text{a.s. under } \mathsf{P}_1,$$

then

$$\frac{1}{n}\max_{t\leq n}\sum_{i=1}^{1+t}\log\frac{f_{i,1}(X_i)}{g(X_i)}\to I,\quad \text{a.s. under } \mathsf{P}_1.$$

关于上述事实的证明，请参见 [2] 中定理 5.1 的证明。这就证明了该定理，因为几乎必然收敛蕴含依概率收敛。 □

我们仍需刻画在何种条件下

$$\frac{1}{n}\sum_{i=1}^{n}\log\frac{f_{i-1}(X_i)}{g(X_i)}\ \to\ I,\quad \text{a.s. under } \mathsf{P}_1,$$

对于一个爆炸过程。这将在第 5.4 节中完成。

5.2. 控制虚警

本节中，我们证明，通过在 EX-CUSUM 算法中将阈值设为 $A=\log\gamma$，可以满足对平均虚警时间的约束。这正是下一个定理的内容。

请注意，

$$W_n=\max_{1\leq k\leq n}\ \sum_{i=k}^{n}\ \log\frac{f_{i-k}(X_i)}{g(X_i)},$$

$$\tau_{ec}=\inf\{n\geq 1: W_n>A\}.$$

定理 5.2. 设置 $A=$ 对数(γ) 可确保

$$\mathsf{E}_\infty[\tau_{ec}]\geq\gamma.$$

证明。 这种证明技巧很常见，例如在 [15] 中就用到了。由于对数函数是单调的，且正数中的最大值总是小于它们的和，因此我们有

$$\begin{aligned}\tau_{ec} &= \inf\left\{n \geq 1 : \max_{1\leq k\leq n} \sum_{i=k}^{n} \log \frac{f_{i-k}(X_i)}{g(X_i)} > A\right\} \\ &= \inf\left\{n \geq 1 : \max_{1\leq k\leq n} \prod_{i=k}^{n} \frac{f_{i-k}(X_i)}{g(X_i)} > e^A\right\} \\ &\geq \inf\left\{n \geq 1 : \sum_{1\leq k\leq n} \prod_{i=k}^{n} \frac{f_{i-k}(X_i)}{g(X_i)} > e^A\right\} := \tau_{esr}.\end{aligned}$$

该过程

$$R_n - n := \sum_{1\leq k\leq n} \prod_{i=k}^{n} \frac{f_{i-k}(X_i)}{g(X_i)} - n$$

是一个 P_∞-鞅。假设 $\mathsf{E}[\tau_{esr}] < \infty$（否则虚警约束显然成立），我们有

$$\begin{aligned}\mathsf{E}\left[|R_n - n|; \{\tau_{esr} > n\}\right] &\leq \mathsf{E}\left[e^A + \tau_{esr}; \{\tau_{esr} > n\}\right] \\ &\to 0, \quad n \to \infty.\end{aligned}$$

因此，根据 Doob 可选抽样定理 [3],

$$\mathsf{E}\left[R_{\tau_{esr}} - \tau_{esr}\right] = 0,$$

and

$$\mathsf{E}\left[\tau_{esr}\right] = \mathsf{E}\left[R_{\tau_{esr}}\right] \geq e^A.$$

现在，令 $A = \log\gamma$，即可完成证明。 □

5.3. 简化爆炸过程时延分析的上界条件

本节中，我们简化爆炸过程的条件 (11)。为完成这一步，我们需要一个中间结果。

引理 5.3。 设 $f(x_1, x_2, \ldots, x_n)$ 是一个在其他参数固定时关于每个自变量递增的连续函数。如果 $\{X_n\}$ 是根据爆炸过程生成的随机过程，那么对于所有 n、m、t,

$$\mathsf{P}\left(f(X_n, X_{n+1}, \ldots, X_{n+m}) \geq t\right) \leq \mathsf{P}\left(f(X_{n+1}, X_{n+2}, \ldots, X_{n+m+1}) \geq t\right). \tag{14}$$

证明。 证明与 [18]中给出的证明类似。我们的证明需要额外的随机化步骤。

□

与定理 5.1 中一样，我们将在下文证明，几乎必然条件足以满足式 (11) 中的条件。

定理 5.4. 要使对所有 $\delta > 0$,

$$\lim_{n\to\infty} \sup_{k\geq\nu\geq 1} \operatorname{ess\,sup} \mathsf{P}_\nu \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_{i,k} \leq I - \delta \;\middle|\; X_1, \ldots, X_{k-1} \right) = 0 \tag{15}$$

对于某个 $0 < I < \infty$，只需对数似然连续且

$$\frac{1}{n} \sum_{k=1}^{n} Z_{k,1} = \frac{1}{n} \sum_{k=1}^{n} \log \frac{f_{k-1}(X_k)}{g(X_k)} \;\to\; I, \quad a.s.\ under\ \mathsf{P}_1.$$

证明。 由于独立性以及爆炸过程的特性，我们有

$$\begin{aligned}
&\sup_{k\geq\nu\geq 1} \operatorname{ess\,sup} \mathsf{P}_\nu \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_{i,k} \leq I - \delta \;\middle|\; X_1, \ldots, X_{k-1} \right) \\
&\quad = \sup_{k\geq\nu\geq 1} \mathsf{P}_\nu \left(\frac{1}{n} \sum_{i=k}^{k+n} \log \frac{f_{i-k}(X_i)}{g(X_i)} \leq I - \delta \right) \\
&\quad = \sup_{k\geq 1} \mathsf{P}_1 \left(\frac{1}{n} \sum_{i=k}^{k+n} \log \frac{f_{i-k}(X_i)}{g(X_i)} \leq I - \delta \right).
\end{aligned}$$

根据引理 5.3，随机变量

$$\frac{1}{n} \sum_{i=k}^{k+n} \log \frac{f_{i-k}(X_i)}{g(X_i)}$$

随着 k 增大而在随机序意义上变大。因此， k 上的最大概率在 $k = 1$ 处取得。这给出

$$\sup_{k\geq 1} \mathsf{P}_1 \left(\frac{1}{n} \sum_{i=k}^{k+n} \log \frac{f_{i-k}(X_i)}{g(X_i)} \leq I - \delta \right) = \mathsf{P}_1 \left(\frac{1}{n} \sum_{i=1}^{1+n} \log \frac{f_{i-1}(X_i)}{g(X_i)} \leq I - \delta \right).$$

最后一项在以下条件下趋于零

$$\frac{1}{n} \sum_{k=1}^{n} \log \frac{f_{k-1}(X_k)}{g(X_k)} \;\to\; I, \quad \text{a.s. under } \mathsf{P}_1.$$

至此，证明完成。

□

因此，对于爆炸型 CUSUM 算法的最优性而言，只需找出能够满足前述定理中所述几乎必然收敛的条件即可。

5.4. 独立非同分布随机变量的大数定律

本节将给出爆炸过程满足以下条件的条件，以确保

$$\frac{1}{n}\sum_{k=1}^{n}\log\frac{f_{k-1}(X_k)}{g(X_k)} \;\to\; I, \quad \text{a.s. under } \mathsf{P}_1.$$

我们首先回顾 Cantelli 强大数定律。证明见[13]。

引理 5.5. 设 $Y_1, Y_2, \ldots$ 为具有有限四阶矩的独立随机变量，并且令

$$\mathsf{E}[|Y_n - \mathsf{E}[Y_n]|^4] \le C, \quad n \ge 1,$$

对于某个常数 C。那么，当 $n \to \infty$ 时，

$$\frac{S_n - \mathsf{E}[S_n]}{n} \to 0, \quad \textit{almost surely.}$$

这里 $S_n = Y_1 + Y_2 + \ldots Y_n$。

Cantelli 强大数定律为我们陈述条件提供了所需的工具。

定理5.6。 我们假设以下条件对每个 $k \ge 1$均成立:

$$\begin{aligned}
&\mathsf{E}\left[\left(\log\frac{f_{k-1}(X_k)}{g(X_k)}\right)^4\right] < \infty, \quad X_k \sim f_{k-1}, \\
&\mathsf{E}\left[\left(\log\frac{f_{k-1}(X_k)}{g(X_k)} - D(f_{k-1} \parallel g)\right)^4\right] \le C, \quad X_k \sim f_{k-1},
\end{aligned} \tag{16}$$

其中， C 为常数。此外，存在一个 $I > 0$ 使得

$$\frac{1}{n}\sum_{k=1}^{n} D(f_{k-1} \parallel g) \quad \to \quad I, \quad n \to \infty. \tag{17}$$

于是，根据 *Cantelli* 强大数定律（引理*5.5*），

$$\frac{1}{n}\sum_{k=1}^{n}\log\frac{f_{k-1}(X_k)}{g(X_k)} \;\to\; I, \quad \textit{a.s. under } \mathsf{P}_1.$$

5.5. EX-CUSUM 算法的渐近最优性

前述定理进一步简化了 EX-CUSUM 算法达到最优性所需的条件。我们将其表述为以下定理:

定理5.7。 在定理*5.6*所述的矩条件下，如果存在一个$I > 0$，使得

$$\frac{1}{n}\sum_{k=1}^{n} D(f_{k-1} \parallel g) \quad \to \quad I, \quad n \to \infty,$$

那么，对于 *Lorden* 和 *Pollak* 的极小极大公式，*EX-CUSUM* 算法都是渐近最优的。

5.6. 高斯示例

现在我们给出一个爆炸过程模型的示例，该模型满足本文中所述的所有条件。我们假设

$$g = \mathcal{N}(0, 1).$$

and

$$f_n = \mathcal{N}(\mu_n, 1)$$

with

$$0 \le \mu_n \uparrow \mu.$$

此示例中的所有似然比都是单调的，因为 $0 \le \mu_n \to \mu$。此外，对数似然比是连续的，因为它们是线性的。由于具有高斯性，对数似然比的四阶矩条件得以满足。

此外， f_n 与 g 之间的对数似然比为

$$\log \frac{f_n(X_{n+1})}{g(X_{n+1})} = \mu_n(X_{n+1} - \frac{\mu_n}{2}).$$

因此,

$$\begin{aligned} D(f_n \parallel g) &= \mathsf{E}_{X_{n+1}\sim f_n}\left[\log \frac{f_n(X_{n+1})}{g(X_{n+1})}\right] = \mu_n\left[\mathsf{E}_{X_{n+1}\sim f_n}[X_{n+1}] - \frac{\mu_n}{2}\right] \\ &= \mu_n\left(\mu_n - \frac{\mu_n}{2}\right) = \frac{\mu_n^2}{2}. \end{aligned} \tag{18}$$

这意味着

$$D(f_n \parallel g) = \frac{\mu_n^2}{2} \to \frac{\mu^2}{2}.$$

这进一步意味着

$$\frac{1}{n}\sum_{k=1}^{n} D(f_{k-1} \parallel g) \quad \to \quad \frac{\mu^2}{2}$$

通过 Cesàro 和极限。

最后，我们需要检查对数似然的四阶中心矩是否一致有界。令

$$\xi_n := \log \frac{f_n(X_{n+1})}{g(X_{n+1})} = \mu_n(X_{n+1} - \frac{\mu_n}{2}).$$

此外，令

$$\Xi_n := \mathsf{E}_{X_{n+1}\sim f_n}[\xi_n] = \frac{\mu_n^2}{2}.$$

于是，

$$\begin{aligned}
\mathsf{E}_{X_{n+1}\sim f_n}\left[(\xi_n - \Xi_n)^4\right] &= \mathsf{E}_{X_{n+1}\sim f_n}\left[\left(\mu_n(X_{n+1} - \frac{\mu_n}{2}) - \frac{\mu_n^2}{2}\right)^4\right] \\
&= \mathsf{E}_{X_{n+1}\sim f_n}\left[\left(\mu_n X_{n+1} - \mu_n^2\right)^4\right] \\
&= \mathsf{E}_{X_{n+1}\sim f_n}\left[\left(\mu_n(X_{n+1} - \mu_n)\right)^4\right] \\
&= \mu_n^4 \mathsf{E}_{X_{n+1}\sim f_n}\left[\left((X_{n+1} - \mu_n)\right)^4\right] \\
&\leq 3\mu^4.
\end{aligned} \tag{19}$$

最后一个不等式成立是因为 $\mu_n \uparrow \mu$，并且在 f_n 条件下，$X_{n+1} - \mu_n \sim \mathcal{N}(0, 1)$，而后者的四阶矩恰好等于 3。因此，四阶中心矩由 $C = 3\mu^4$ 一致有界。以上所有论证表明，对于该示例，EX-CUSUM 算法是渐近最优的。

6. 数值结果

本节将 EX-CUSUM 算法应用于第 5.6 节讨论的高斯数据示例。为生成图 4 中的数据，我们使用了

$$\mu_n = \arctan(n) \to \mu = \frac{\pi}{2}.$$

如图所示，在变化点 80 之前，EX-CUSUM 统计量始终接近零；变化点之后则逐渐增大至 ∞。通过精心设计的阈值可以检测到这一增长。

7. 结论

我们提出了一类新的随机过程，用于对某些变化点问题中变化发生后观测值不断增大的特性进行建模。这类观测在卫星和军事应用中很常见：逼近中的敌方物体会导致观测值随时间随机增大。我们证明，在 Kullback–Leibler 散度的增长满足温和条件时，随着平均虚警时间趋于无穷大，我们提出的 EX-CUSUM 算法是渐近最优的。在我们的

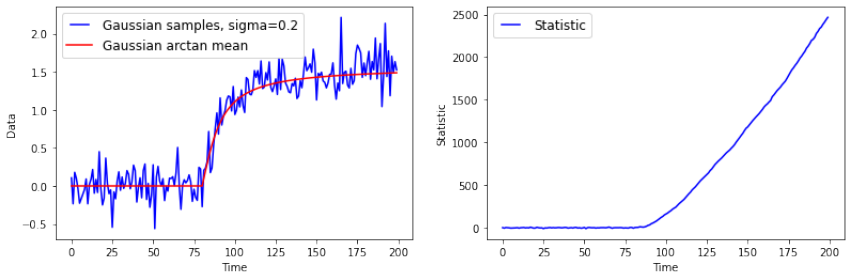

图4. 应用于高斯数据的 EX-CUSUM 算法。

未来工作中，我们将研究该算法的精确最优性，并提出计算高效的变化检测方法。

致谢

本研究的部分内容曾于2022年9月在伊利诺伊州蒙蒂塞洛举行的第58届阿勒顿通信、控制与计算年会上发表。美国国家科学基金会通过范德比尔特大学的分包项目，以第1917164号拨款部分资助了 Taposh Banerjee 的研究。我们撰写本文期间，TimBrucks 不幸离世。谨以本文纪念他。

参考资料

[1]轨道碎片的威胁及保护 *NASA* 航天资产免受卫星碰撞，`http://images.spaceref.com/news/2009/ODMediaBriefing28Apr09-1.pdf`。访问日期：2023-02-03。[2] T. Banerjee、P. Gurram 和 G.T. Whipps，统计周期随机过程变化检测的贝叶斯理论，IEEE信息论汇刊 67 (2021)，第 2562–2580 页。[3] Y.S. Chow、H. Robbins 和 D. Siegmund，远大期望：最优停止理论，Houghton Mifflin，1971。[4] V. Krishnamurthy，部分可观测马尔可夫决策过程，剑桥大学出版社，2016。[5] T.L. Lai，随机系统中参数变化的信息界与快速检测，IEEE信息论汇刊 44 (1998)，第 2917–2929 页。[6] Y. Liang、A.G. Tartakovsky 和 V.V. Veeravalli，变化后观测非平稳情况下的最快变化检测，即将在 IEEE信息论汇刊发表（2022）。[7] G. Lorden，对分布变化作出反应的程序，《统计年刊》42 (1971)，第 1897–1908 页。[8] S. Mohammed、M.J.S. Teja、P. Kora、D.E. Vunnava 和 M. Lokesh，航天器空间碎片检测装置，载于*2022*年计算机通信与信息学国际会议（*ICCCI*）。IEEE，2022，第 1–2 页。[9] G.V. Moustakides，检测分布变化的最优停止时刻，《统计年刊》14 (1986)，第 1379–1387 页。[10] S. Pergamenchtchikov 和 A.G. Tartakovsky，相关数据下渐近最优的逐点及极小极大最快变化点检测，《随机过程统计推断》21 (2018)，第 217–259 页。

[11] M. Pollak，分布变化的最优检测，《统计年刊》13 (1985)，第 206–227 页。[12] H.V. Poor 和 O. Hadjiliadis，最快变化检测，剑桥大学出版社，2009。[13] A.N. Shiryaev，概率论-*2*，第 95 卷，施普林格，2019。[14] A.N. Shiryayev，最优停止规则，施普林格出版社，纽约，1978。[15] A.G. Tartakovsky、I.V. Nikiforov 和 M. Basseville，序贯分析：假设检验与变点检测，统计学，CRC出版社，2014。[16] A.G. Tartakovsky 和 V.V. Veeravalli，最快变化检测的一般渐近贝叶斯理论，SIAM《概率论与应用》49 (2005)，第 458–497 页。[17] A.G. Tartakovsky，序贯变点检测中的渐近最优性：非独立同分布情形，IEEE信息论汇刊 63 (2017)，第 3433–3450 页。[18] J. Unnikrishnan、V.V. Veeravalli 和 S.P. Meyn，极小极大鲁棒最快变化检测，IEEE信息论汇刊 57 (2011)，第 1604–1614 页。[19] V.V. Veeravalli 和 T. Banerjee，最快变化检测，《学术出版社信号处理丛书：第 3 卷——阵列与统计信号处理》，2014。[20] Y. Xiang、J. Xi、M. Cong、Y. Yang、C. Ren 和 L. Han，采用快速网格学习的空间碎片检测，载于*2020*年*IEEE*第三届安全生产与信息化国际会议（*IICSPI*）。IEEE，2020，第 205–209 页。