arXiv:2602.04322v1 [stat.ME] 4 Feb 2026

基于最小有效分割的精确多重变点检测

VincentRunge[1*], AnicaKostic[2], Alexandre Combeau[1], Gaetano Romano[3]

[1*]埃夫里数学与建模实验室（Laboratoire de Math´ematiques etMod´elisation d'Evry），巴黎-萨克雷大学（Universit´e Paris-Saclay），法国国家科学研究中心，埃夫里大学，埃夫里-库尔库罗讷，91037，法国。[2]伦敦政治经济学院统计学系，哥伦比亚大厦，霍顿街，伦敦，WC2A 2AE，英国。[3]兰卡斯特大学数学科学学院，菲尔德学院，兰卡斯特，LA1 4YF，英国。

*通讯作者。电子邮箱：vincent.runge@univ-evry.fr；其他作者：kostic.ani@gmail.com;alexandre.combeau.77@gmail.com; g.romano@lancaster.ac.uk;

摘要

我们提出最小有效分割（SVP），这是一种用于时间序列多变点检测的分段方法。SVP依赖于分段有效性的局部概念：仅当候选分段通过用户选定的有效性检验（例如单变点检验）时，才予以保留。基于有效分段的集合，我们提出一种连贯的聚合程序，用以构建作为某一优化问题精确解的全局分段。我们的主要贡献在于对该优化问题采用优先考虑简约性的字典序。我们分析了所得程序的计算复杂度，其时间复杂度依所选代价与有效性函数、数据情形以及检测到的变点数量，介于线性与立方级之间。最后，我们通过与标准最优分割算法的比较来评估SVP的质量，结果表明SVP在显式强制分段有效性的同时，能够得到具有竞争力的分段。SVP的灵活性使其适用于广泛的问题类别；作为示例，我们通过将稳健性编码到有效性准则中，展示了稳健变点检测。

关键词：多变点检测、动态规划、剪枝、字典序

1 引言

在时间序列分析中，分段平稳时间序列十分常见。这类时间序列的行为会发生突变，并被划分为一系列连续分段，各分段中的数据呈现相似的性质。在这种情况下，我们可能希望根据观测数据识别其潜在的变化序列。对这一问题的研究历史悠久，自二十世纪中叶以来取得了长足进展，尤其是在序贯分析 [1–3] 和变点检测 [4–7]领域。这里可以识别出两个不同的挑战。第一是单变点检测，即在在线数据流中识别一次变化。第二是多变点检测，旨在将固定长度的时间序列划分为一个或多个连续分段。在这项工作中，我们提出了一种新策略，通过按字典序最小化定义明确的目标函数，将单变点检测结果汇总为全局分段。为简单起见，我们聚焦于一种典型情形：检测数据生成分布的位置（集中趋势）参数发生的突变，例如均值或中位数。如果底层单次检验程序未在某个分段中检测到变化，我们就称该分段为有效分段（尽管这一概念可以推广）。通过汇总单变点检验来构建全局分段，是离线变点检测中的标准范式 [8, 9]；我们的贡献是将这种汇总过程形式化为一个精确的优化框架。

处理多变点情形有多种算法策略。特别是，动态规划解在过去二十年中日益受到关注，因为它们通过精确最小化全局代价函数来获得最优分段——该函数通常表述为惩罚似然。在这一领域，两种开创性方法是最优分割（OP）算法 [10]和分段邻域（SN）方法 [11]。近年来，这些方法得到了大幅改进，以降低其二次时间复杂度 [12, 13] ，并处理更复杂的数据结构，例如对相邻分段均值的约束[14] ，或连续性约束下的斜率变化 [15]。模型复杂度通常由惩罚项控制，用于选择合理数量的变点。该惩罚项既可以直接纳入分段算法[16]，也可以在之后通过模型选择步骤应用，利用不同长度的多个候选分段 [17, 18] ，这些候选分段通常由 SN 算法生成。在所有这些方法中，惩罚项的校准仍是一项重大挑战；尽管它能对整体分段进行全局控制，但每个分段内部的局部拟合质量仅仅是校准方式所带来的结果。

在序贯分析或在线检测中，可以使用多种方法识别单个变点，例如累积和 [19, 20] 及其近似方法 [21, 22]。人们提出了各种策略，将单变点方法扩展到处理多个变点。通过在不同的数据分段上反复进行单变点检测，可以利用二分分段（BS）算法及其变体 [23–25] 重建多变点结构。尽管该算法的运行时间近似线性，但所得分段并不保证是

全局优化问题的精确极小解。一种称为 SMUCE 的多尺度变点检测程序 [9] 实现了与 SVP 类似的聚合方式，将单次检验步骤的有效性与聚合步骤分离（不使用字典序）。在他们的研究中，有效性检验是一个精确的单变点程序，它会检验所有子分段；其解与 SVP 的约束最大似然解相同，但这仅在某些条件下成立（参见其中的引理 3.1。）[9]。

借助我们提出的新方法，我们希望在全局分段过程中发挥单变点检测的精细控制能力。最小有效分割（SVP）算法通过有效性检验筛选候选分段，剔除可能包含变点的过大分段，从而应对这一挑战。我们为多变点检测提出了一种新颖的目标函数，并将这一有效性准则明确纳入其中。SVP 算法通过汇总局部单分段检验构建全局分段，从而精确地最小化该目标函数。为限制分段数量并避免过度分段，我们采用字典序最小化策略，在偏好较小分段（由有效性检验保证）与尽可能减少分段数量的目标之间取得平衡。据我们所知，这是首个利用字典序在统一且有序的框架中系统整合单变点检测与多变点检测的方法。因此，SVP 为成熟的 OP 和 SN 算法提供了一种互补的算法方法。

本文组织如下。第2 节介绍优化问题，以及分段代价和分段有效性的概念。在第3节中，我们介绍称为 SVP 算法的精确动态规划解、它与 OP方法的联系，以及其时间复杂度分析。第4 节详细说明我们针对一些精确的单变点有效性检验所采用的方法（特别是使用 CUSUM FOCuS [20]）。在第5节中，我们在均值模式发生变化的模拟数据上评估我们的方法，并在各种有效性检验函数下将其性能与 OP 进行比较。我们将 SVP 与中位数变化有效性检验（威尔科克森和 Mood）相结合，结果表明，与稳健的最优划分（OP）基线相比，该方法展现出良好的性能。最后，在第6节中，对测井数据的应用证明了 SVP 生成的分段具有现实性。

2 问题描述

2.1 分段代价和分段有效性

给定长度为 n 的单变量时间序列 $(y_t)_{t\geq 1}$，我们将分段 $y_{a..b}$ 定义为连续的数据点 $y_{a+1}, \ldots, y_b$，其中 $a < b$为整数，其对应的分段代价由数值 $\mathcal{C}(y_{a..b})$ 给出。索引集合 $\{1, \ldots, n\}$ 的一种分段记为 $\tau = \{0 = \tau_0, \tau_1, \ldots, \tau_{K-1}, \tau_K = n\}$，其中 τ_k 为递增排列的整数， K 为分段数。分段的总体代价为 $Q_n(\tau;\ y) = \sum_{k=0}^{K-1} \mathcal{C}(y_{\tau_k..\tau_{k+1}})$。

代价函数通常基于数据模型的负对数似然而构建。如果数据点 y_t 取自密度为 $p(y_t, \cdot)$ 的分布，我们得到$\mathcal{C}(y_{a..b}) = \min_\theta\{-\sum_{t=a+1}^{b} \log\ p(y_t|\theta)\}$。例如，对于分段均值 $\overline{y}_{a..b} = \frac{1}{b-a}\sum_{i=a+1}^{b} y_i$，有：

$$\mathcal{C}^{\text{Gauss}}(y_{a..b}) = \frac{1}{2}\sum_{i=a+1}^{b}(y_i - \overline{y}_{a..b})^2\,, \quad \mathcal{C}^{\text{Poisson}}(y_{a..b}) = (b-a)\overline{y}_{a..b}(1-\log(\overline{y}_{a..b}))\,, \tag{1}$$

它们分别是高斯模型和泊松模型的分段代价。当底层模型属于指数族时，只要预先计算好累积统计向量，即得到 $\mathtt{cs}(y)_s = \sum_{t=1}^{s} y_t$, $s = 1, \ldots, n$，无论分段长度如何，都可以在常数时间内计算分段代价。也可以使用其他代价度量来量化分段围绕集中趋势度量的变异性。中位数绝对偏差反映了围绕分段中位数的离散程度，

$$\mathcal{C}^{\text{MAD}}(y_{a..b}) = \sum_{i=a+1}^{b}\left|y_i - \text{median}(y_{a..b})\right|,$$

而基于分位数的代价

$$\mathcal{C}^{\text{quant}}(y_{a..b}; x) = q_{1-x}(y_{a..b}) - q_x(y_{a..b}),$$

是一种稳健的段内离散程度度量，其中 $2x$ 表示在两端忽略的数据所占的比例。令 $x = 0$ 即可得到极差代价。数据趋势的变化也可以纳入变点框架。在这种情况下，基于线性拟合进行分段是自然的做法，合适的代价是线性回归的残差平方和 [15]。通常，目标函数也倾向于使分段代价 $\mathcal{C}(y_{\tau_k..\tau_{k+1}})$尽可能小。

在本文中，选择有效分段 $\tau^\star$ 时，只考虑通过有效性检验的分段。为形式化这一概念，我们定义一个有效集合 $\mathcal{V}(\gamma) \subset \mathbb{R}^v$，其大小由用户定义参数 $\gamma \in \mathbb{R}^v$控制。我们还引入一个函数 $F_l : \mathbb{R}^l \to \mathbb{R}^v$，并要求返回的分段中的每个分段 $y_{a..b}$ 都满足约束 $F_{b-a}(y_{a..b}) \in \mathcal{V}(\gamma)$。函数 F_l也由用户定义。在我们的研究情境中，这自然是一个单变点检测检验。一旦检测到变点，该分段就会被拒绝，并被视为无效。我们在第 4 节提出了此类函数的若干实例，并将在第 5 节的模拟研究中评估它们对分段质量的影响。为简便起见，我们取 $v = 1$，现在稍微滥用一下符号，将有效分段写作 $f(y_{a..b}) \leq \gamma$。有效性检验 f 的选择将强烈影响所检测到的变点类型。

2.2 优化问题

通常采用的方法如下。如果分段数量 K已知，变点检测方法会从所有 $\tau^\star$ 种可能的候选方案中，选择使 $Q_n(\tau;\, y)$ 最小的分段向量 $\binom{n-1}{K}$，其中包含 K 个分段。如果 K 是自由参数，则候选方案有 2^{n-1} 种。随后使用模型选择准则，选择一个具有一定统计保证 [16–18]的良好划分。为使 $Q_n(\tau;\, y)$ 最小，前一种方法可以通过一种称为 Segment Neighborhood（SN）的动态规划方法精确求解；后一种方法则采用另一种称为最优分割（OP）的动态规划方法。

我们通过在$\mathbb{N}^\star \times \mathbb{R}$中定义双点 $R_n = (K_n(\tau)\ Q_n(\tau;\, y))$，提出了一种新方法，其中 $Q_n(\tau;\, y)$ 仍表示总体代价，而 $K_t(\tau)$ 表示分段 τ中的分段数量。

给定 $\mathbb{N}^\star \times \mathbb{R}$中的元素有限集 $\mathcal{R}$，我们定义两类极小值。第一个极小值记为 $\min_K \mathcal{R}$，返回 $\mathcal{R}$ 中分段数量最少的元素子集（第一分量）。类似地，第二个极小值记为 $\min_Q \mathcal{R}$，返回全局代价最小的元素子集（第二分量）。使用这些记号，寻找最佳分段——SmallestValidPartition——的优化问题可表述如下：

$$R_n = \min_Q \min_K \left\{ \left(K,\, \sum_{k=0}^{K-1} \mathcal{C}(y_{\tau_k..\tau_{k+1}}) \right),\ f(y_{\tau_k..\tau_{k+1}}) \le \gamma,\ k = 0, \ldots, K-1 \right\}. \quad (2)$$

可用解的集合 $\mathcal{R}$（其中 $R_n = \min_Q \min_K \mathcal{R}$）包含具有$K_{\min}$个分段、 $K_{\min}+1$个分段、 $K_{\min}+2$个分段等的元素，其中 $K_{\min}$表示可能的最小分段数量。直观而言，由于有效性约束会排除较长的分段， $K_{\min}$ 通常会大于 1；但在不存在变化的序列以及 γ取值非常大时例外。算子 $\min_K$ 选出具有 $K_{\min}$ 个分段的所有解，而 $\min_Q$ 随后从中选出函数代价 $\mathcal{C}$ 下拟合优度最佳的解。约束$f(y_{\tau_k..\tau_{k+1}}) \le \gamma$ 施加于划分的所有分段，并确保所选分段均不包含变化。若减小参数 γ，有效性条件会变得更加严格，从而增加分段数量。因此， γ充当权衡参数，发挥上述标准模型选择步骤的作用。

对这些元组（双点）分别取两个极小值，自然会引入元组的字典序。因此，接下来将用 $\min_\preceq$替代记号 $\min_Q \min_K$ ，其定义如下：

$$(K, Q) \preceq (K', Q') \iff \Big\{ K < K' \quad \text{or} \quad (K = K' \text{ and } Q \le Q') \Big\},$$

其中 (K, Q) 和 (K', Q') 属于 $\mathbb{N}^\star \times \mathbb{R}$。空集的最小值始终为$+\infty$ ($\min_\preceq \emptyset = +\infty$)。

3 自适应动态规划

3.1 序贯优化

我们的优化问题（2）可以通过动态规划（DP）技术求解。所得估计量具有自适应性，因为分段数量由分区算法推断得出，而非预先固定，也不是在计算出划分后再通过外部模型选择步骤确定。

在变点检测中，动态规划递归会通过确定最后一段的起始索引， t对每个 $y_{0..t}$ ，求出前缀数据 t的最优分段方式。对于每个 $\mathbb{N}^\star \times \mathbb{R}$，保存目标函数的最小值（中的一个元素）及其对应的极小点，即可通过回溯重建（2）的最优划分。以下命题概述了这一策略。

命题 1 *For a given integer t and $1 \leq r < t$ we consider the best solution, (,) $R_r =$ K_r Q_r with $Q_r = Q_r$ τ $= \sum_{k=0}^{K_r-1} \mathcal{C}$ $y_{\tau_k..\tau_{k+1}}$ for segmenting data $y_{0..r}$ of the opti-() () mization problem (2) as known. Based on these values, we can easily obtain the next solution R_t by solving:*

$$R_t = \min_{\substack{\preceq \\ 0 \leq s < t}} \left\{ R_s + (1, \mathcal{C}(y_{s..t})) \, , \, f(y_{s..t}) \leq \gamma \right\}, \tag{3}$$

with the initial condition $R_0 =$ (0,0) and where the minimum is considered in lexicographic order ($\preceq$) with t elements.

证明见附录A。回溯步骤是标准做法，并在文献中有广泛记载。

最小值可以通过利用字典序的性质高效计算。具体而言，我们按分段数量递增的顺序检查（3）中的元素，一旦找到第一个有效元素，便忽略所有分段数量更大的元素。形式上，我们写作：

$$\mathcal{R}_k = \left\{ (K_s, Q_s) + (1, \mathcal{C}(y_{s..t})) \, , \, 0 \leq s < t \, , \, K_s = k \, , \, f(y_{s..t}) \leq \gamma \right\},$$

并得到

$$R_t = \min_{\preceq} \bigcup_{k \geq 0} \mathcal{R}_k \, .$$

注意，按照约定， $\mathcal{R}_0$ 是空集。值 R_t 是第一个非空 $\mathcal{R}_k$ 集合的最小值：

$$R_t = \min_{\preceq} \mathcal{R}_K \, , \quad \text{where} \quad \mathcal{R}_0 = \mathcal{R}_1 = \ldots = \mathcal{R}_{K-1} = \emptyset \, , \, \mathcal{R}_K \neq \emptyset \, . \tag{4}$$

我们定义以下将在后文使用的概念。

定义1 如果有效性函数 f 具有以下性质，我们就称其为 γ-稳定：

$$f(y_{s..t}) > \gamma \implies f(y_{s..u}) > \gamma \, , \forall u > t \, .$$

范围代价函数就是一个简单的 γ-稳定函数示例。这一性质可直接用于剪枝索引。要证明 PELT 等经典剪枝规则，则需要一个类似的定义。

定义 2 如果有效性函数 f 具有以下性质，我们称其为 γ^{-1}-稳定：

$$f(y_{t..u}) > \gamma \implies f(y_{s..u}) > \gamma\,, \forall s < t\,.$$

可以通过在有效性检验中进行设计来实现 γ-稳定性。为此，我们将该检验写作 $f_\gamma(y_{s..t}) \leq \gamma$，其含义是 $F_{t-s}(y_{s..t}) = (f(y_{s..s+1}), f(y_{s..s+2}), \ldots, f(y_{s..t})) \in (-\infty, \gamma]^{t-s}$。$\gamma$稳定性的一个显著性质是其剪枝能力。任何满足 $f_\gamma(y_{s..t}) > \gamma$ 的 s 都可以从最小化中移除，即 $R_s + (1, \mathcal{C}(y_{s..t}))$ 不再是式 (3) 中需要考虑的元素。这种剪枝与最小化策略 (4) 相结合，可得到高效的算法（见第 3.3 节）。当分段数 K_t 是递增函数时，我们可以利用 (4) 更容易地找到 R_t。在这种情况下，集合 $\mathcal{R}_k$ 按索引 s 递增的顺序包含元素 $(K_s, Q_s) + (1, \mathcal{C}(y_{s..t}))$。我们有以下重要结果。

引理 1. 函数 $t \mapsto K_t$ 递增，当函数 f 是 γ-稳定的。

Proof 考虑值 $R_t = (K_t, Q_t)$ 中的一个。与之关联的分段为 $(\tau_0, \ldots, \tau_{K_t})$，并且对于任何小于 t 的索引 s，都存在 $k^\star$ 使得 $\tau_{k^\star - 1} \leq s < \tau_{k^\star}$。双点

$$\left(k^\star - 1, \sum_{k=1}^{k^\star - 1} \mathcal{C}(y_{\tau_{k-1}..\tau_k})\right) + (1, \mathcal{C}(y_{\tau_{k^\star-1}..s}))$$

有 $k^\star \leq K_t$ 个分段。如果数据分段 $y_{\tau_{k^\star-1}..s}$ 有效，那么这个双点就是字典序最小值中的一个可用值，因此 $K_s \leq k^\star \leq K_t$。我们能否有$y_{\tau_{k^\star-1}..s}$ 无效的情况？如果可以，那么我们有 $f(y_{\tau_{k^\star-1}..s}) > \gamma$，根据 γ-稳定性，也有 $f(y_{\tau_{k^\star-1}..\tau_{k^\star}}) > \gamma$。然而，后一分段是有效的，因为它是解 R_t 中的一个有效分段。这证明了分段 $y_{\tau_{k^\star-1}..s}$ 的有效性，从而证明了引理。□

3.2 与 OP方法的联系

一种自然的有效性检验基于代价函数：

$$f_{OP}(y_{a..b}) = \max_{a<u<b} \{\mathcal{C}(y_{a..b}) - (\mathcal{C}(y_{a..u}) + \mathcal{C}(y_{u..b}))\}\,,$$

，它直接源自广义似然比检验（参见下文其与公式（5）的联系）。当 $f_{OP}(y_{a..b}) > \gamma$ 存在 $u^\star$ 使得 $\mathcal{C}(y_{a..b}) > \mathcal{C}(y_{a..u^\star}) + \mathcal{C}(y_{u^\star..b}) + \gamma$。在 OP 框架中，我们优化量 $\sum_{k=0}^{K-1} \{\mathcal{C}(y_{\tau_k..\tau_{k+1}}) + \gamma\}$；满足 $f_{OP}(y_{\tau_k..\tau_{k+1}}) > \gamma$ 的分段是次优的，因此用 $(\mathcal{C}(y_{\tau_k..u^\star}) + \mathcal{C}(y_{u^\star..\tau_{k+1}})) + 2\gamma$ 替换 $\mathcal{C}(y_{\tau_k..\tau_{k+1}}) + \gamma$ 更可取。

命题 2 *If we choose the GLR validity test (5) then the obtained segmentation is a penalty value γ. We write $K_n^{SVP} \leq K_n^{OP}$.*

Proof 我们考虑通过最优分割算法获得的分段，以及其中一个记作 $y_{a..b}$ 的分段。如果其代价满足存在 u， $a < u < b$使得$\mathcal{C}(y_{a..b}) > \mathcal{C}(y_{a..u}) + \mathcal{C}(y_{u..b}) + \gamma$，则存在一个更优的分段，这与 OP 结果的最优性相矛盾。这意味着 OP 结果中的所有分段都是有效分段。因此，OP 划分是优化问题（2）的一个元素。SVP 算法返回分段数量最少的那个，这意味着 $K_n^{SVP} \leq K_n^{OP}$。□

3.3 SVP 算法

小型有效分割算法 1 与其他基于动态规划的变点检测算法具有相同的结构。第 5 行的条件用于筛选与有效分段相连接的可用索引。算法按照字典序寻找最小值，而不是考虑最小的全局代价（SN）或惩罚全局代价（OP）。为完整起见，我们在第 13 行加入了通用的剪枝步骤。该算法返回一个由元素组成的矩阵；在第 t 行中，元素（$R(t)$ $S(t)$）是一个三元组，其中包含值（Q_t， K_t， s_t），其中 s_t 是最后一个变点的索引。使用标准的回溯过程，可以恢复 SVP 优化问题（2）下的最优分段。

算法 1 SVP算法

```
1:  R(0) = (0,0), S(0) = 0, τ ← (0)                                  ▷ 初始化
2:  对 t = 1 至 n 执行                                                ▷ 按时间循环
3:      R* ← (+∞, +∞)
4:      对于 s ∈ τ 执行                                               ▷ 遍历可访问的最后变点索引
5:          如果 f(y_{s..t}) ≤ γ 则 R = R(s) + (1, C(y_{s..t}))        ▷ 有效性检验
6:              如果 R ⪯ R* 则 R* ← R, s* ← s                          ▷ 字典序
7:              结束如果
8:          结束如果
9:      结束循环
10:     R(t) ← R*, S(t) ← s*                                          ▷ 保存最小值和最佳索引
11:     τ ← (τ, t)
12:     对于 s ∈ τ 执行                                               ▷ 剪枝
13:         如果 剪枝(s) = 真 则 τ ← τ \ {s}
14:         结束如果
15:     结束循环
16: 结束循环
17: 返回 (R, S)
```

该算法的时间复杂度为立方级，因为计算 $f(y_{s..t})$ 和 $\mathcal{C}(y_{s..t})$ 相对于片段长度而言是线性的（即 $\mathcal{O}(t-s)$）。然而，

对于从指数族分布中推导出的所有代价，代价值都可以在常数时间内获得（参见（1））。此外，从值 $f(y_{s..t})$ 更新到值 $f(y_{s..(t+1)})$ 通常可以在常数时间或对数时间内完成。当有效性函数同样具有 γ-稳定性时，我们可以得到更优的界，其时间复杂度从线性到二次不等，并带有对数项。以下命题给出了此类界限，并假设 K_t 序列是已知的。

命题 3 *We consider that the vector $(K_t)_t$ contains n_k times the value k for k between 1 and K with $n_K > 0$. We have $\sum_{k=1}^{K} n_k = n$ and define $n_0 = 1$. This means that the obtained segmentation contains K segments. The time for updating the validity test from index a to b is expressed as $T_f\ b-a$. For a γ-stable validity function with a constant per iteration update (to compute the value $f(y_{s..(t+1)})$ from $f(y_{s..t})$ so that $T_f(b-a) = \mathcal{O}(b-a)$), the time complexity $T(n)$ of the SVP algorithm is bounded by the relations:*

$$T(n) \leq \mathcal{O}\Big(\sum_{k=1}^{K}\{n_{k-1}n_k + T_f(n_{k-1}+n_k)n_{k-1}\}\Big) \leq \mathcal{O}(n^2)\,.$$

Considering that the segmentation is regular, $n_k \approx n/K$, with a number of changesproportional to the data length, $K \approx \alpha n$, $(0 < \alpha < 1)$ we obtain a linear complexity, $T(n) = \mathcal{O}\binom{3n}{\alpha}$.

Proof 结合 γ 稳定性（引理 1 的结论）与字典序 (4) 的组织方式，对于给定的时间 t，我们只有两种可用的分段数，分别为 $k-1$ 和 k。对于具有 $k-1$ 个分段的 n_{k-1} 索引（在 $(K_t)_{t\geq 1}$ 中），只要有效性检验允许，我们就计算其最小值；也就是说，计算具有 k 个分段的所有 n_k 值。这样需要计算 $n_{k-1}n_k$ 个值。由于更新时不考虑具有 n_k 个分段的值，因此随着时间增加，这些值也不会针对有效性检验进行更新。再往前看一步，我们还必须更新具有 $k-1$ 个分段的解的有效性检验。最多，我们遗漏了一个大小为 $n_{k-1}+n_k$、更新时间为 $T_f(n_{k-1}+n_k)$ 的分段，而这项更新会重复 n_{k-1} 次。□

请注意，γ存在可在常数时间内更新的稳定函数；极差代价就是一个显而易见的例子。重要的 CUSUM 检验更新在期望意义下为对数时间（见第 4 节）。我们可以调整 PELT [26] 和 FPOP [12] 的剪枝规则，舍弃 (3) 中的一些索引。这些规则仅适用于分段数相同的元素，且有效性函数满足 γ^{-1} 稳定性。我们在附录B中介绍 PELT 规则；实际上，即使存在变化，它的剪枝能力也很低，因为我们比较的是分段数相同的划分。

4 单变化有效性函数

我们介绍两类有效性函数，以展示 SVP 方法的灵活性。基于代价的有效性检验（例如累积和）是构建单变点程序并将其整合为全局分段时的自然选择。然而，SVP 也支持解耦策略，此时有效性检验无需与

分割代价函数相匹配。为了增强对离群值的稳健性，我们因此考虑基于秩的统计量，即 Wilcoxon检验和 Mood 中位数检验，并评估它们在 SVP 方法中的表现。

4.1 指数族模型的基于代价的 FOCuS 检验

有效性函数的一个自然参数化选择是用于检测单变点的广义似然比（GLR）检验。设 $y_{s..t}$ 为一个分段，并假设观测值服从自然形式下密度为

$$p(y_t \mid \theta) = \exp\{\langle T(y_t), \theta\rangle - A(\theta) + B(y_t)\}\,,$$

的指数族，其中 $T(y_t)$ 是充分统计量，θ 是自然参数，$A(\theta)$ 是从 $\mathbb{R}^p$ 到 $\mathbb{R}$ 的凸函数，而 $B(y_t)$ 是数据的函数。对于这类模型，检验无变点与在 $y_{s..t}$ 中的 τ 处存在一个变点的广义似然比统计量为

$$f_{\mathrm{LR}}(y_{s..t}) = \max_{\tau\in\{s,\ldots,t-1\}} \left\{\ell(\hat{\theta}_{s..\tau}; y_{s..\tau}) + \ell(\hat{\theta}_{\tau..t}; y_{\tau..t})\right\} - \ell(\hat{\theta}_{s..t}; y_{s..t}), \tag{5}$$

其中 $\ell(\theta;\ y_{u..v}) = \sum_{i=u+1}^{v} \log\, p(y_i \mid \theta)$，而 $\hat{}\,\theta_{u..v}$ 是分段 $y_{u..v}$ 上的最大似然估计。

在已知方差的单变量高斯均值偏移情形下，$y_t \sim \mathcal{N}(\mu,\ 1)$，该统计量可化简为经典的累积和统计量 [27]。将变点前后均值记为 $(\mu_0,\ \mu_1)$，则广义似然比为

$$f_{\mathrm{LR}}(y_{s..t}) = \max_{\substack{\tau\in\{s,\ldots,t-1\}\\ \mu_0,\mu_1\in\mathbb{R}}} \left\{-\frac{1}{2}\sum_{i=s+1}^{\tau}(y_i-\mu_0)^2 - \frac{1}{2}\sum_{i=\tau+1}^{t}(y_i-\mu_1)^2\right\}$$

$$-\max_{\mu\in\mathbb{R}}\left\{-\frac{1}{2}\sum_{i=s+1}^{t}(y_i-\mu)^2\right\}.$$

计算该统计量需要同时对分割点 τ 和各段均值进行最大化。虽然对 $(\mu_0、\ \mu_1)$ 的优化可以通过闭式解求得，但必须在所有位置 $s,\ \ldots,\ t-1$ 上搜索 τ。因此，直接计算 $f_{\mathrm{LR}}(y_{s..t})$ 需要 $\mathcal{O}(t-s)$ 次运算，而在 SVP 中使用该方法会使总复杂度达到 $\mathcal{O}(n^3)$。

FOCuS 算法 [28] 通过对广义似然比进行精确的序贯更新，解决了这一瓶颈：$y_{s..t}$ 上的统计量可直接根据其在 $y_{s..(t-1)}$ 上的取值以及新观测 y_t 得到。这使得每次更新的期望摊销代价降至 $\mathcal{O}(\log(t-s))$，并使 SVP 的期望复杂度降至 $\mathcal{O}(n^2\log n)$。此外，同样的加速方法也适用于一大类指数族代价[29] 和低维多变量情形 [30]，其复杂度为 $\log^p(n)$，而非 $\log(n)$，且不影响统计最优性。

4.2 非参数稳健有效性检验

我们使用中心化 Wilcoxon 扫描定义一种基于秩的代价 [31–33]。给定一个分段$y_{s..t}$，对于任意分割点 $u \in \{s+1, \dots, t-1\}$，中心化 Wilcoxon 统计量为

$$W_u(s,t) \;=\; \sum_{i=s+1}^{u} \sum_{j=u+1}^{t} \left(I\{y_i \le y_j\} - 1/2\right) \qquad \text{and} \quad f_W(y_{s..t}) = \max_{u=s+1,\dots,t-1} |W_u(s,t)|.$$

将 $f_W(y_{s..t})$ 更新为 $f_W(y_{s..(t+1)})$ 时，所有秩都可能发生变化；因此，通常需要检查所有分割点 u，所需时间至少为 $\mathcal{O}(t-s)$。该有效性函数主要用于检测分段内的分布变化，但对位置偏移也很敏感，这表明它对离群值具有稳健性。

在模拟中，我们设定 $\gamma = 1.5\sqrt{\ell^3/12}$，其中 ℓ 是典型的片段长度（例如，对于 K 个变点，预言机平均值为 $\ell = n/K$ ）。这一选择与中心化 Wilcoxon 扫描的零假设尺度相符，因为 $Var(W_u) = \mathcal{O}(\ell^3/12)$，因此 $\mathrm{sd}(W_u)$ 的量级为 $\ell^{3/2}$；因子 1.5 是对分割点取最大值时的一种实用校准。

我们还考虑使用 Mood 中位数检验 [34] 的基于中位数的扫描。给定一个长度为 $\ell = t - s$ 的分段 $y_{s..t}$，令 $\tilde{y}$ 为 $\{y_{s+1}, \dots, y_t\}$ 的合并中位数。对于任意分割点 $u \in \{s+1, \dots, t-1\}$，定义 2×2 表，用于统计两个子分段中低于/高于中位数的观测值数量，

$$N_1^-(u) = \#\{i \le u : y_i \le \tilde{y}\}, \quad N_1^+(u) = \#\{i \le u : y_i > \tilde{y}\},$$

$$N_2^-(u) = \#\{i > u : y_i \le \tilde{y}\}, \quad N_2^+(u) = \#\{i > u : y_i > \tilde{y}\}.$$

分割点 u 处的 Mood 中位数检验统计量是常用的独立性 Pearson 卡方统计量，

$$M_u \;=\; \sum_{a \in \{1,2\}} \sum_{b \in \{-,+\}} \frac{\left(N_a^b(u) - E_a^b(u)\right)^2}{E_a^b(u)},$$

其中，$E_a^b(u)$ 是零假设下的期望计数：两个子分段具有相同分布（相对于合并中位数）。接着，我们将基于扫描的有效性（或代价）定义为 $f_M(y_{s..t}) = \max_{u=s+1,\dots,t-1} M_u$。将分段从 t 延伸至 $t+1$ 时，合并中位数 $\tilde{y}$（因而所有计数 $N_a^b(u)$）都可能发生变化，因此更新 f_M 通常需要考虑所有分割点 u，这至少需要 $\mathcal{O}(\ell) = \mathcal{O}(t-s)$ 的运行时间。该有效性函数对离群值尤其稳健，因为它只取决于与中位数的比较，而非数值大小。为进行校准，我们对 $t-s-1$ 个候选分割点采用 Dunn–Šidák 校正 [35]，分段级别设为 $\alpha = 0.01$。

5 模拟研究

本节中，我们研究 SVP 的性能和实测运行时间。首先，我们评估采用基于代价的有效性检验时，高斯均值变点的定位和检验功效。为展示我们方法的灵活性，随后我们使用稳健检验，在存在离群值的情况下进行变点检测。最后，我们研究该方法在有变化和无变化情景下的实测运行时间。

5.1 均值变点场景

现在，我们通过模拟研究评估 SVP 在四种均值变点场景下的性能，如图 1 所示。这些是用于评估检测性能的标准场景，正如 [20] 中此前所报告的那样。对于每次实验，我们生成长度为 $n = 1000$ 的数据，并加入方差为$\sigma^2 = 1$的独立同分布高斯噪声。根据每种情景的结构，我们在从 0.1 到 2 的数值网格范围内引入具有不同跳跃幅度（变化幅度）的变化。对于每种配置，我们进行 100 次重复实验，其中仅改变噪声。

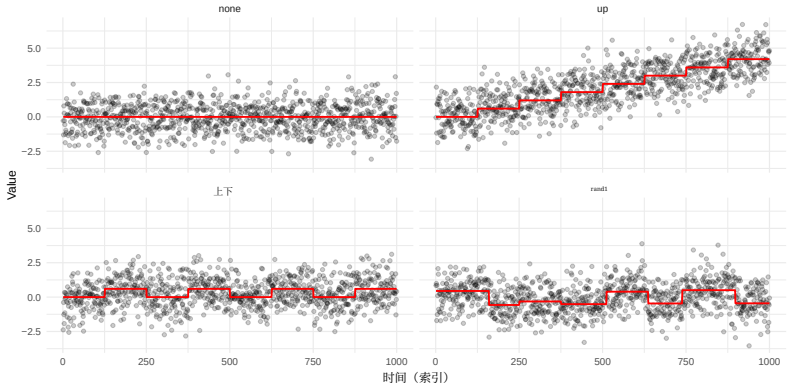


图 1：展示了模拟情景中 4 个变化模式示例，跳跃幅度为 0.6。在所展示的模拟中，我们调整跳跃幅度，并改变重复实验中的噪声。

我们将 SVP 与其 PELT 版本 [26] 中的 OP 进行比较。具体而言，我们使用 2 log(n) 的 BIC 惩罚项将 PELT 与 SVP 进行比较，并使用自定义惩罚项 1.5log(n) 的 SVP 进行比较，以匹配无 情景中的 PELT 假阳性率。

为评估准确性，我们采用基于容差的匹配准则：如果检测到的变点位于真实变点 ± 2.5 个观测值范围内，则视为正确检测。根据这些匹配结果，我们计算精确率、召回率和 F1 分数（两者的调和平均数）。图2展示了各情景和跳跃幅度对应的 F1 分数，附录C中的图6则展示了相应的精确率和召回率曲线。

在所有情景中，对于这两种方法，跳跃幅度越大，检测难度越低。跳跃幅度较小时，PELT 的 F1 分数略高于 SVP。相应的变点位置分布（附录C，图 8）显示，当变化幅度非常小时，PELT 检测到的变点更集中于真实变点周围，这与 F1 曲线所呈现的行为一致。对于中到大幅度的跳跃（约高于 0.9），所有方法在所有情景中的表现实际上都相当，F1 分数接近 1。在这些情况下，PELT 和 SVP 的精确率和召回率都很高，且几乎相同。总体而言，SVP 比 PELT 更保守；在采用相同的 BIC 惩罚项时（见命题 2），SVP 在无 情景下产生的假阳性更少，并且在小跳跃幅度的设置中倾向于避免边缘性检测。

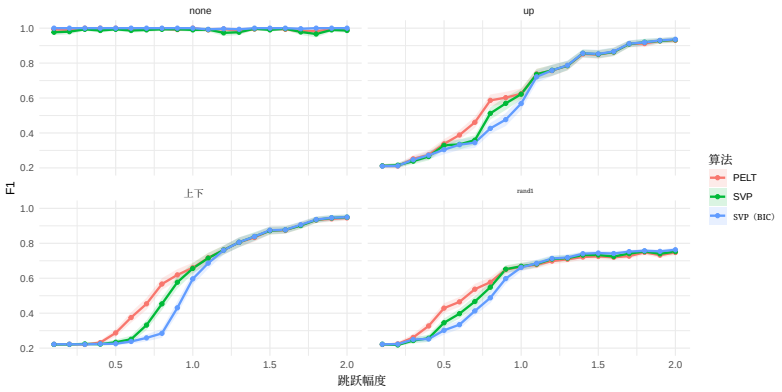


图 2：PELT 和 SVP（使用似然和贝叶斯信息准则惩罚项）在四种模拟情景中的 F1 分数随跳跃幅度的变化。每个点表示 100 次重复实验的平均值。

5.2 重尾观测

我们评估存在重尾噪声时 SVP 的表现。我们考虑与第 5.1 节相同的四种情景，但现在将高斯噪声替换为自由度为 2 的 Student t-分布。

由于 PELT 假设观测值服从高斯分布并使用平方误差损失，因此无法稳健地应对这类重尾污染。因此，我们将 SVP 与 RFPOP 进行比较[36], 。RFPOP 是最优分割算法的一个版本（与 PELT 类似），但采用了对离群值稳健的代价函数（双权损失）。我们测试了第 4.2 节介绍的两种非参数 SVP 变体。对于 RFPOP，我们再次使用贝叶斯信息准则，并采用 ℓ 阈值 3 作为双权损失参数的离群值检测阈值。对于非参数 SVP 变体，威尔科克森方法采用 $1.5\sqrt{(n/K)^3/12}$ 惩罚项（其中K 是真实分段的数量）；中位数方法则采用数据驱动阈值，并将假阳性率控制在 $\alpha = 0.01$ （见第 4.2 节）。

F1 分数结果见图 3，附录C 中还列出了精确率、召回率和变化分布。正如预期，在所有跳跃幅度下，存在重尾噪声时 PELT 的性能都显著下降。与 PELT 相比，RFPOP 的稳健性有所提升，即使在重尾污染下也能成功检测变化。非参数 SVP 变体在所有情景下均表现出强大且稳定的稳健性，在上升和无情景下优于 RFPOP，其中威尔科克森版本略优于 Mood 版本。与其他稳健方法相比，RFPOP 往往会产生更多检测结果（见附录图 7中的召回率图）；但与此同时，其假阳性率也明显高于 SVP 变体。

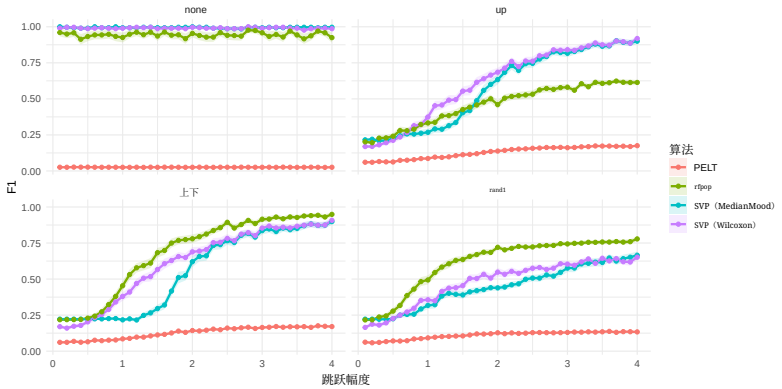


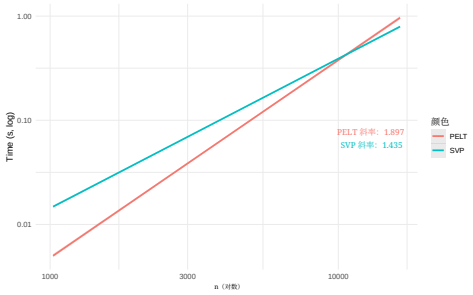
图 3：重尾噪声下四种情景中，鲁棒变点检测方法的 F1 分数随跳跃幅度的变化。每个点表示 100 次重复实验的平均值。

5.3 运行时性能

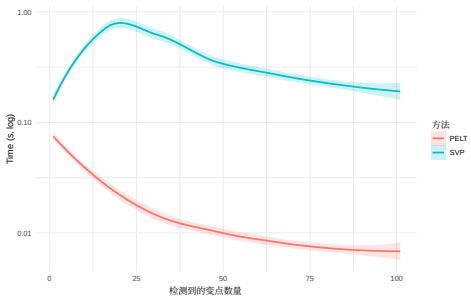
现在，我们将考察采用 FOCuS 基于代价的有效性检验的 SVP 与 PELT 相比的计算效率。我们进行两项实验：首先，在没有任何变点的情况下改变时间序列的长度，以评估该方法在零假设下的扩展性；其次，考察运行时间如何随检测到的变点数量而变化。我们测量了各方法在搭载第13代英特尔 i7-1370P、运行 Ubuntu 22.04 LTS 的设备上的用户执行时间。PELT 的实现来自官方的 `changepointR` 软件包，该软件包基于 Rcpp。

在第一个实验中，我们生成了如下独立同分布的高斯观测序列：$\sigma^2 = 1$ 且均值没有变化，并改变序列长度 n，范围从 1,000 到 10,000。结果以双对数尺度展示在图 4a 中。正如预期，在这种情景下，PELT 的复杂度接近最坏情况复杂度 $\mathcal{O}(n^2)$。SVP 的计算复杂度呈对数线性 $\mathcal{O}(n\log(n))$ 增长，与 FOCuS 算法的时间复杂度相符。对于较短的序列，由于 SVP 的计算开销，PELT 的运行时间更短。

在第二个实验中，我们生成了含有数量不等且间距相等的变点的数据，并测量了各方法在包含 10,000 个观测值的序列上的运行时间。图 4b 展示了运行时间随检测到的变点数量变化的情况。我们观察到，随着检测到的变点数量增加，PELT 的实测运行时间会缩短。另一方面，SVP 在没有变点时运行时间较短；变点数量较少时，运行时间会增加，随后随着变点数量继续增加而再次缩短。这与命题 3的结果一致。



(a) 运行时间与序列长度的关系 n（无变点）。两个坐标轴均采用对数刻度。



(b) 运行时间与检测到的变点数量的关系。Y 轴采用对数刻度。

图 4：SVP 与 PELT 的计算性能比较。

6 测井数据分析

我们在一个用于 [37]变点检测的标准石油时间序列数据集上比较了四种分段算法，该数据集包含许多离群值。结果如图5所示。PELT 显然不适用于这种情形：它产生了许多短分段，这些分段似乎是由离群观测值驱动的。相比之下，稳健 FPOP 得到的划分有意义得多，而 Mood–SVP 方法的表现相近。与稳健 FPOP 相比，Mood–SVP 识别出了相同的主要变化，并在最长的分段内额外识别出三个变点。

7 结论

在本工作中，我们提出了一种新的优化框架，用于在有效性约束下对时间序列进行分段，从而对全局分段施加局部的、分段尺度的控制。这通过采用字典序来实现，该顺序偏好更稀疏的划分，同时要求每个分段满足所选的有效性准则。由此得到的最小有效分割（SVP）可通过动态规划计算，从而得到一种高效算法，其时间复杂度从线性到立方级不等，并强烈依赖于代价与有效性函数的选择。

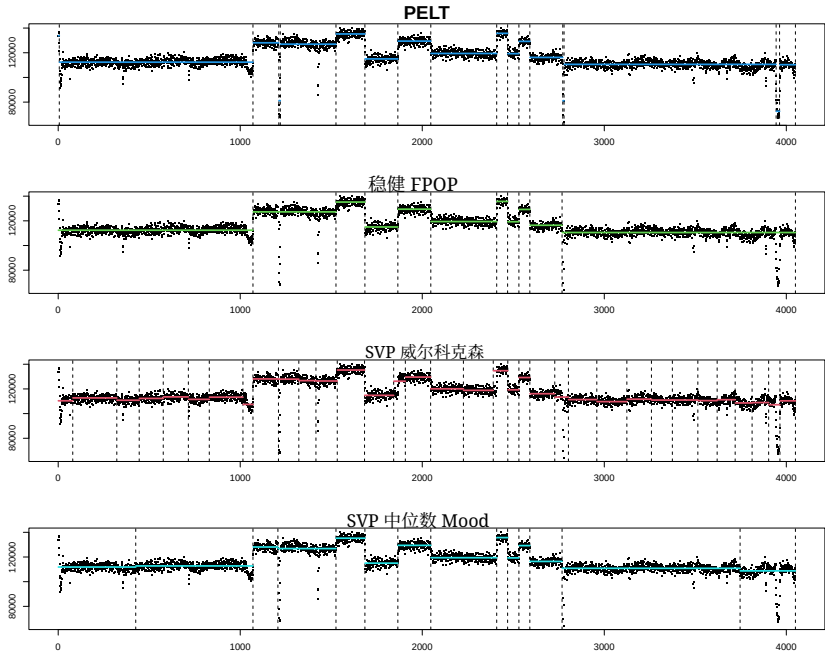


图 5：测井数据分段。采用 4 种不同方法的结果

SVP的灵活设计还支持稳健变体，例如将标准代价与抗异常值的有效性检验相结合。

仍有几个方面可以改进。首先，可以整合基于凸包的方法（例如 FOCuS 算法中使用的方法），以加快有效性更新。其次，我们目前对每个分段使用局部检验的方法仍较为粗略。例如，该检验具有方向性，这意味着按相反顺序运行时间序列算法 $y_{n..1}$，不一定会得到相同的分段。采用更有原则的方法来组合或校准这些局部决策，可以显著提高统计功效和整体分段质量。这将留待未来研究。

致谢。本研究得到了英国工程与自然科学研究委员会资助项目 UKRI2698 的支持。

命题的证明1

我们将所有最后一个变点位于 s 的分段中最优的双点记为 R_t^s：

$$R_t^s := \min_{\preceq} \left\{ \left(K, \sum_{k=0}^{K-1} \mathcal{C}(y_{\tau_k..\tau_{k+1}}) \right) : f(y_{\tau_k..\tau_{k+1}}) \leq \gamma,\ \forall k,\ \tau_{K-1} = s,\ \tau_K = t \right\}.$$

我们回到 R_t 的定义，并分离出最后一段：

$$\begin{aligned}
R_t &= \min_{\preceq} \{R_t^s,\ 0 \leq s < t\} \\
&= \min_{\preceq} \Big\{ \min_{\preceq} \Big\{ \Big(K, \sum_{k=0}^{K-1} \mathcal{C}(y_{\tau_k..\tau_{k+1}}) \Big) : f(y_{\tau_k..\tau_{k+1}}) \leq \gamma, \\
&\qquad\qquad \forall k,\ \tau_{K-1} = s,\ \tau_K = t \Big\}, 0 \leq s < t \Big\} \\
&= \min_{\preceq} \Big\{ \min_{\preceq} \Big\{ \Big(K-1, \sum_{k=0}^{K-2} \mathcal{C}(y_{\tau_k..\tau_{k+1}}) \Big) + (1, \mathcal{C}(y_{s..t})) : f(y_{\tau_k..\tau_{k+1}}) \leq \gamma, \\
&\qquad\qquad \forall k,\ \tau_{K-1} = s,\ \tau_K = t \Big\}, 0 \leq s < t \Big\} \\
&= \min_{\preceq} \Big\{ \min_{\preceq} \Big\{ \Big(K-1, \sum_{k=0}^{K-2} \mathcal{C}(y_{\tau_k..\tau_{k+1}}) \Big) : f(y_{\tau_k..\tau_{k+1}}) \leq \gamma, \\
&\qquad\qquad \forall k,\ \tau_{K-1} = s \Big\} + (1, \mathcal{C}(y_{s..t})) : f(y_{s..t}) \leq \gamma,\ 0 \leq s < t \Big\} \\
&= \min_{\preceq} \Big\{ R_s + (1, \mathcal{C}(y_{s..t})) : f(y_{s..t}) \leq \gamma,\ 0 \leq s < t \Big\}
\end{aligned}$$

B PELT 剪枝规则

命题 4 *A PELT-like pruning rule can be described by the following inequality. If the validity function f is γ^{-1} stable, if for $s < t$, we have:*

$$(K_s, Q_s) + (0, \mathcal{C}(y_{s..t})) \succ (K_t, Q_t)\,.$$

we can prune solution $(K_s,\ Q_s)$ at further iterations $u > t$.

符号 $\succ$ 表示左侧元素在字典序下严格大于右侧元素。

Proof 考虑索引 $u > t$，我们有：

$$\begin{aligned}
(K_s, Q_s) + (1, \mathcal{C}(y_{s..u})) &\succeq (K_s, Q_s) + (1, \mathcal{C}(y_{s..t}) + \mathcal{C}(y_{t..u})) \\
&= (K_s, Q_s) + (0, \mathcal{C}(y_{s..t})) + (1, \mathcal{C}(y_{t..u})) \\
&\succ (K_t, Q_t) + (1, \mathcal{C}(y_{t..u}))\,.
\end{aligned}$$

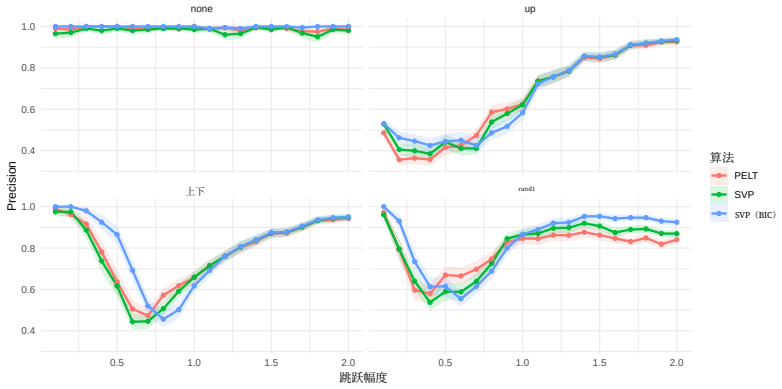
索引 s 不可能在这种情况下成为最优解，因为根据 PELT 条件，索引为 t 的解总是更优。γ^{-1} 稳定性用在什么地方？如果分段 $y_{t..u}$ 无效

且未按字典序使用，我们还需要确保 $y_{s..u}$ 未被使用。这正是 γ^{-1} 稳定性的定义。□

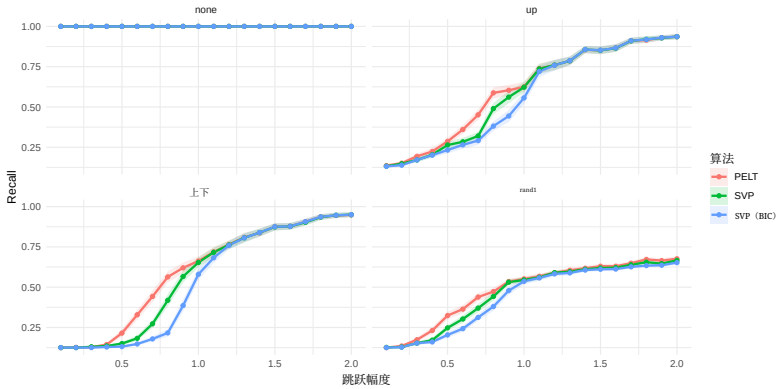
类似 FPOP 的剪枝规则可用函数关系 $Q_t^s(\theta) = \min\left\{q_t^{s'}(\theta) \quad s' \geq s\,, \quad K_s = K_t\right\}$ 表示，其中 $q_t^{s'}(\theta) = Q_{s'} + \mathcal{C}(y_{s'..t};\, \theta)$。

C 附加模拟结果

这里报告一些补充图表，以支持主要的实证模拟研究。

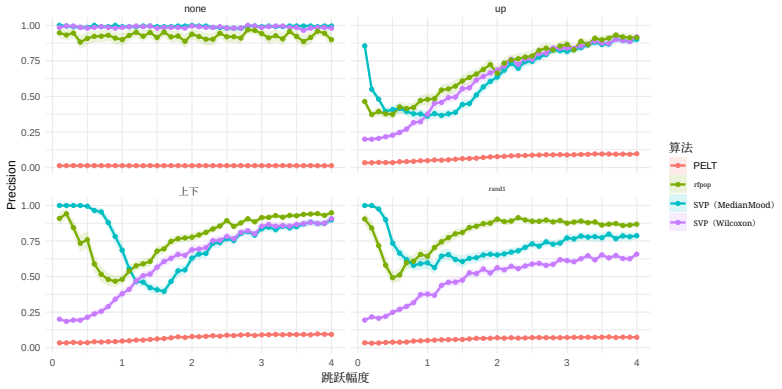


(a) 精确率

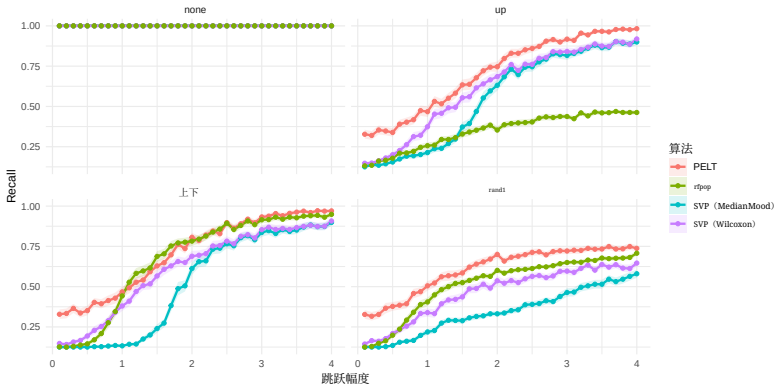


(b) 召回率

图 6：均值变化情形下，四种情景中的性能指标随跳跃幅度的变化。每个点代表 100 次重复实验的平均值。



(a) 精确率



(b) 召回率

图7：重尾噪声下鲁棒变点检测方法在四种情景中的性能指标随跳跃幅度的变化。每个点表示 100 次重复实验的平均值。

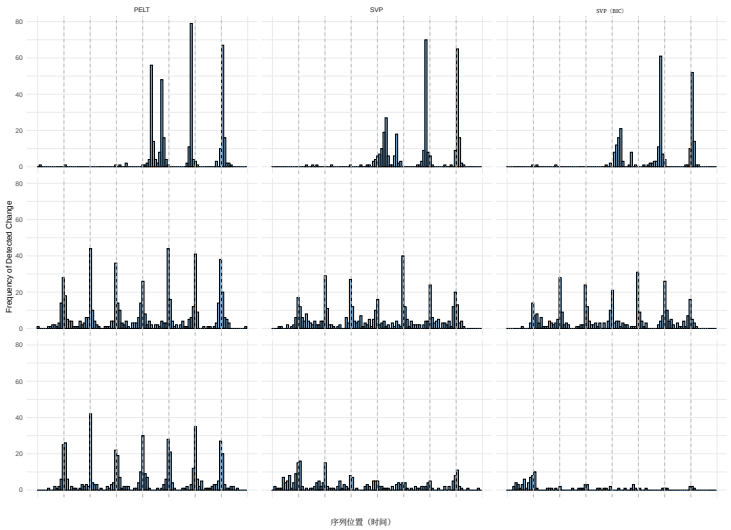


图 8：均值变化研究中，在 0.6 变化幅度情景下所有重复实验检测到的变化的分布。

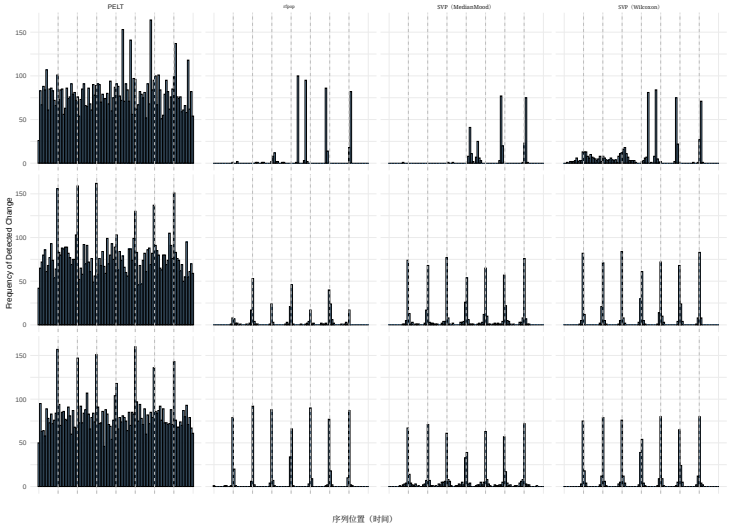


图 9：稳健重尾研究中，在变化幅度为 2 的情景下所有重复实验检测到的变化的分布。

参考文献

[1] Ghosh, B.K., Sen, P.K.：《序贯分析手册》。CRC出版社，???（1991）[2] Siegmund, D.：《序贯分析：检验与置信区间》。施普林格，???

（2013）

[3]Lai, T.L.：序贯分析：若干经典问题与新挑战。中华统计学志，303–351（2001）[4] Series,I.L.N.-M.：变点问题 IMS讲义专著丛书（第23卷，1994）齐性非参数检验：在参数估计中的应用，作者为K. Ghoudi与D. Mcdonald。变点问题 23，149（1994）[5] Brodsky, E., Darkhovsky, B.S.：变点问题中的非参数方法 第243卷。施普林格，???（2013）[6] Cs¨org¨o, M., Horv´ath, L.：变点分析中的极限定理。（无标题）（1997）[7] Chen, J., Gupta, A.K., Gupta, A.：参数统计变点分析 第192卷。施普林格，???（2000）[8] Fryzlewicz, P.：用于多变点检测的野二分分割。《统计年鉴》42(6)，2243–2281（2014）https://doi.org/10.1214/ 14-AOS1245[9] Frick, K., Munk, A., Sieling, H.：多尺度变点推断。《皇家统计学会会刊B辑：统计方法学》76(3)，495–580（2014）[10] Jackson, B., Scargle, J.D., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L., Tsai, T.T.：区间数据最优划分算法。IEEE信号处理快报 12(2)，105–108（2005）[11] Auger, I.E., Lawrence, C.E.：片段邻域最优识别算法。数学生物学通报 51(1)，39–54（1989）[12]Maidstone, R., Hocking, T., Rigaill, G., Fearnhead, P.：关于大规模数据的最优多变点算法。《统计与计算》27，519–533（2017）[13] Rigaill, G.：一种用于恢复含1至k个最大变点的最佳分段的剪枝动态规划算法。Journal de la Soci´et´e Fran˛caise de Statistique 156(4)，180–205（2015）[14] Runge, V., Hocking, T.D., Romano, G., Afghah, F., Fearnhead, P., Rigaill, G.：gfpop：用于单变量图约束变点检测的R软件包。《统计软件杂志》106(6)，1–39（2023）https://doi.org/10.18637/jss. v106.i06[15] Fearnhead, P., Maidstone, R., Letchford, A.：使用l 0惩罚检测斜率变化。计算与图形统计学杂志 28(2)，265–275

（2019
年）
[16]Yao, Y.-C.：基于Schwarz准则估计变点数目。《统计与概率快报》6(3)，181–189（1988）[17] Lebarbier, ´E.：通过模型选择检测高斯过程均值中的多变点。《信号处理》85(4)，717–736（2005）[18] Verzelen, N., Fromont, M., Lerasle, M., Reynaud-Bouret, P.：最优变点检测与定位。《统计年鉴》51(4)，1586–1610（2023）[19] Page, E.S.：连续检验方案。《生物计量学》41(1/2)，100–115（1954）[20] Romano, G., Rigaill, G., Runge, V., Fearnhead, P.：在存在局部波动和自相关噪声时检测突变。《美国统计学会会刊》117(540)，2147–2162（2022）[21] Chen, Y., Wang, T., Samworth, R.J.：高维多尺度在线变点检测。《皇家统计学会会刊B辑：统计方法学》84(1)，234–266（2022）[22] Meier, A., Kirch, C., Cho, H.：mosum：用于变点分析中移动和的软件包。《统计软件杂志》97，1–42（2021）[23] Scott, A.J., Knott, M.：一种用于方差分析中均值分组的聚类分析方法。《生物计量学》，507–512（1974）[24] Fryzlewicz, P.：用于多变点检测的野二分分割。《统计年鉴》42(6)（2014）https://doi.org/10.1214/14-aos1245[25] Kov´acs, S., B¨uhlmann, P., Li, H., Munk, A.：种子二分分割：一种快速且最优的变点检测通用方法。《生物计量学》110(1)，249–256（2023）[26] Killick, R., Fearnhead, P., Eckley, I.A.：以线性计算成本实现变点的最优检测。《美国统计学会会刊》107(500)，1590–1598（2012）[27] Yu, Y., Madrid Padilla, O.H., Wang, D., Rinaldo, A.：关于在线变点检测的注记。《序贯分析》42(4)，438–471（2023）[28] Romano, G., Eckley, I.A., Fearnhead, P., Rigaill, G.：通过函数剪枝累积和统计量实现快速在线变点检测。《机器学习研究杂志》24(81)，1–36（2023）[29] Ward, K., Romano, G., Eckley, I., Fearnhead, P.：针对指数族模型在线变点检测的每次迭代常数似然比检验。《统计与计算》34(3)，99（2024）

[30] Pishchagina, L., Romano, G., Fearnhead, P., Runge, V., Rigaill, G.：在线多元变点检测：利用与计算几何的联系。《皇家统计学会会刊B辑：统计方法学》，046（2025年）[31] Wilcoxon, F.：基于秩方法的个体比较。《生物统计学通报》**1**（6），80–83（1945年）[32] Pettitt, A.N.：变点问题的非参数方法。英国皇家统计学会会刊：C辑（应用统计） **28**（2），126–135（1979年）[33]Gerstenberger, C.：短程依赖下变点位置的稳健Wilcoxon型估计。《时间序列分析杂志》 **39**（1），90–104（2018年）[34] Brown, G.W., Mood, A.M.：关于线性假设的中位数检验。见：第二届伯克利数学统计与概率研讨会论文集，第2卷，第159–167页（1951年）。加利福尼亚大学出版社[35] ˇSid´ak, Z.：多元正态分布均值的矩形置信域。《美国统计学会会刊》 **62**（318），626–633（1967年）[36] Fearnhead, P., Rigaill, G.：存在离群值时的变点检测。《美国统计学会会刊》 **114**（525），169–183（2019年）[37] ´O Ruanaidh, J.J.K., Fitzgerald, W.J.：应用于信号处理的数值贝叶斯方法。《统计与计算》。施普林格，纽约，NY（1996年）。
https://doi.org/10.1007/978-1-4612-0717-7