The Korean Journal of Applied Statistics (2016) DOI: http://dx.doi.org/10.5351/KJAS.2016.29.6.1147
29(6), 1147–1154

Detection of multiple change points using penalized least square methods: a comparative study between ℓ_0 and ℓ_1 penalty

Won Son[a] · Johan Lim[a] · Donghyeon Yu[b,1]

[a]Department of Statistics, Seoul National University;
[b]Department of Statistics, Keimyung University

(Received September 26, 2016; Revised September 30, 2016; Accepted September 30, 2016)

Abstract

In this paper, we numerically compare two penalized least square methods, the ℓ_0-penalized method and the fused lasso regression (FLR, ℓ_1 penalization), in finding multiple change points of a signal. We find that the ℓ_0-penalized method performs better than the FLR, which produces many false detections in some cases as the theory tells. In addition, the computation of ℓ_0-penalized method relies on dynamic programming and is as efficient as the FLR.

Keywords: change point detection, complexity penalty, dynamic programing, fused lasso regression

1. 서론

변화점은 연속적으로 관측되는 확률과정의 분포적 성질에 급격한 변화가 발생하는 점으로 지진파를 이용한 지진의 예측, 경제 시계열에 있어서 추세 변동점 탐지, 그리고 생산 공정의 안정적인 관리 등 다양한 분야에 있어 중요한 연구 주제이다 (Kotz 등, 2006). 이러한 변화점과 관련한 연구에서 주요 관심 분야로는 크게 (1) 변화점의 실시간 탐색(on-line detection of a change), (2) 변화점 존재 여부에 대한 검정(off-line hypotheses test), 그리고 (3) 변화점의 위치와 수준에 대한 추정(off-line estimation of the change)로 구분되어진다 (Basseville와 Nikiforov, 1993). 이러한 변화점과 관련한 연구를 위하여 각 목적에 따라 다양한 방법이 제시되었고 몇 가지 대표적 방법들로는 우도비(likelihood ratio)를 이용한 방법, 정보량기준(information criteria)을 이용한 방법, 마코프-몬테칼로에 기반한 베이지안 방법(Bayesian method), 자료의 누적합(cumulative sum; CUSUM)을 이용한 방법, 웨이블렛 근사법(wavelets approximation) 등이 있다. 이들 방법에 대한 보다 자세한 내용은 Carlstein 등 (1994), Csörgö와 Horváth (1997), Chen과 Gupta (2001), Kotz 등 (2006)을 참조하기 바란다.

본 논문에서는 변화점의 위치와 수준을 추정하는 문제, 특히 최근 통계학 분야에서 많은 관심을 받고 있는 벌점 최소제곱법을 이용한 변화점의 추정 방법들을 살펴보고자 한다. 먼저 우리가 다루고자 하는 자

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2015R1C1A1A02036312).
[1]Corresponding author: Department of Statistics, Keimyung University, 1095 Dalgubeol-daero, Dalseo-gu, Daegu 42601, Korea. E-mail: dyu3@kmu.ac.kr

료와 이에 대한 통계적 모형을 살펴보면 다음과 같다. 관측값 $y_1, \ldots, y_n$들이 서로 독립이고 동일한 분포를 따르는(IID) 확률분포로부터 추출되었고 확률밀도함수가 시점 $t = 1, \ldots, n$에 대해 $f(y_t; \theta_t)$로 정의됨을 가정한다. 확률밀도함수의 모수 $\theta_1, \theta_2, \ldots, \theta_n$들은 구간적 상수(piecewise constant)함수를 형성하고 이는 미지의 m과 $k_1, \ldots, k_m$에 대해

$$\theta_1 = \theta_2 = \cdots = \theta_{k_1} \neq \theta_{k_1+1} = \cdots = \theta_{k_m} \neq \theta_{k_m+1} = \cdots = \theta_n$$

임을 의미한다. 따라서 위의 통계적 모형 하에서 본 논문의 주된 문제는 구간적 상수 함수의 추정과 k_i와 θ_j들에 대한 추정에 있다고 할 수 있으며 이를 풀기 위하여 최근 다양한 벌점 최소제곱법들이 제안되었다. 특히 ℓ_1-벌점과 ℓ_0-벌점에 기반한 방법이 많은 주목을 받았다 (Jang 등, 2015; Johnson, 2013; Lim 등, 2012; Rinaldo, 2009; Tibshirani 등, 2005; Ye와 Xie, 2011; Yu 등, 2015a, 2015b).

변화점 검정을 위한 ℓ_1-벌점 최소제곱법은 "fused-라쏘-회귀(fused lasso regression; FLR)" 또는 "fused-라쏘-근사기(fused lasso signal approximator; FLSA)"이라는 이름으로 보다 널리 알려져 있다. 본 논문에서는 의미의 명확성과 명칭의 편의성을 위하여 "FLR"이라는 약어를 사용하고자 한다. FLR은 Tibshirani 등 (2005)에 의하여 처음 제안된 이후 이론적 성질, 빠른 계산, 실제 문제에의 적용 등에 대한 많은 연구가 진행되어 왔다 (Rinaldo, 2009; Rojas와 Wahlberg, 2015; Ye와 Xie, 2011; Yu 등, 2015a). ℓ_1-벌점 최소제곱법은 모수의 총변동(total variation)에 대해 벌점을 부여하는 방식으로 다음과 같이 정의된다.

$$\begin{aligned} &\text{maximize } \sum_{t=1}^{n} \log f(y_t; \theta_t) \\ &\text{subject to } \|\Theta\|_{\mathrm{TV}} \leq \lambda, \end{aligned} \tag{1.1}$$

여기에서 $\|\Theta\|_{\mathrm{TV}} = \sum_{t=1}^{n-1} |\theta_{t+1} - \theta_t|$이다.

FLR이 다양한 문제에 적용되고 사용됨에 따라 최근 많은 연구자들이 FLR의 이론적 성질에 대한 연구를 진행하고 있으며 특히 다중-변화점 탐색과 관련된 이론적 성질을 규명하고자 하는 연구가 활발히 이루어지고 있다. Rinaldo (2009)는 관측값 y_i들의 표준편차 σ_n이 0으로 수렴하는 경우 FLR에 의해 추정된 변화점 $\hat{k}_1, \ldots, \hat{k}_m$들이 점근적 일치성(asymptotically consistency)을 가지고 모수의 추정량 $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$들이 부호-일치성(sign-consistency)을 가진다고 주장하였으나 최근 Rojas와 Wahlberg (2015)를 포함한 여러 연구자들에 의하여 이 주장에 오류가 있음이 밝혀졌다. Rojas와 Wahlberg (2015)는 FLR방법이 모수 $\theta_1, \theta_2, \ldots, \theta_n$들이 계단식(staircase)으로 상승 또는 하강하는 경우에는 무시할 수 없는 확률로 실제 변화점이 아닌 점을 변화점으로 판단하게 됨을 보였다. 또한, 모수 $\theta_1, \theta_2, \ldots, \theta_n$들이 국소적 극단값(local extremum)들로만 구성되어 있는 경우에는 변화점의 추정량 $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$들이 실제 변화점 $k_1, \ldots, k_m$에 가까이 위치함을 보였고 이를 ϵ-부호일치성(ϵ-sign-consistency)이라 명명하였다. 이러한 성질은 이전에 Harchaoui와 Lévy-Leduc (2010)에 의해서도 주장되었다. 한편, Lin 등 (2016)은 ℓ_1-벌점 최소제곱법에 의한 추정량 $\widehat{\Theta}$은 특정 조건하에서 $\|\cdot\|_n = \|\cdot\|_2/\sqrt{n}$로 정의된 $\|\cdot\|_n$에 대해

$$\mathbb{E}\left[\left\|\widehat{\Theta} - \Theta\right\|_n^2\right] = O\left(\frac{\log n \log\log n}{n}\right)$$

을 만족함을 보였고 더불어 이러한 관계가 성립할 때, 변화점의 추정량 $\hat{k}_1, \ldots, \hat{k}_m$들이 실제 변화점 $k_1, \ldots, k_m$에 가까이 위치하게 됨을 증명하였다. Qian과 Jia (2016)는 실제 변화점이 아닌 점을 변화점으로 판단하게 되는 것이 라쏘-회귀분석의 irrepresentable condition (Zhao와 Yu, 2006)이 만족하

지 않음에 기인한다는 점에 착안하여 이 조건을 만족하도록 사전-조건(pre-conditioning)을 부여한 후에 FLR문제를 적용하는 변형된 FLR절차를 제안하였다.

변화점 검정을 위한 ℓ_0-벌점은 복잡도-벌점이라고도 불리우며 Lim 등 (2012)과 Johnson (2013) 등에 의하여 연구 되었다. 먼저 ℓ_0-벌점 최소제곱법은

$$\begin{aligned} &\text{maximize } \sum_{t=1}^{n} \log f\left(y_t;\theta_t\right) \\ &\text{subject to } \Theta \in \mathcal{S}_m \end{aligned} \tag{1.2}$$

로 정의되고 여기에서 집합 $\mathcal{S}_m$은 m개의 변화점을 가진($m+1$개의 연속된 구간(segment)으로 나누어진) 모수의 집합 $\Theta(=\{\theta_1,\ldots,\theta_n\})$로 정의한다.

ℓ_0-벌점 최소제곱법은 활발히 연구가 이루어진 FLR 방법과 비교하여 상대적으로 주목을 받지 못하였는데 이는 통상적으로 복잡도-벌점의 계산이 조합적 최적화(combinatorial optimization)문제를 풀 것을 요구함에 따라 최적화의 계산량의 증대 및 비용이 원인으로 예상된다. 하지만 Lim 등 (2012)와 Johnson (2013)에서 연구 된 것처럼 위의 식 (1.2)의 문제는 좋은 계층적 구조를 지니고 있어 동적프로그래밍(dynamic programming)방법을 이용하여 빠른 계산이 가능함이 알려져 있다.

본 연구에서는 다중-변화점 탐색의 관점에서 FLR과 ℓ_0-벌점 최소제곱법의 성능을 수치적 실험을 통하여 비교함을 목적으로 하며, 특히 변화점 탐색의 성능 측면에서 ℓ_0-벌점 최소제곱법의 우월성을 이야기하고자 한다.

2. 비교 연구

본 절에서는 모의 실험을 통하여 ℓ_0-벌점과 FLR의 다중-변화점 탐색의 성능을 비교하고자 한다. 모의 실험에서는 길이가 100인 관측 값 $\mathbf{Y}=(Y_1,Y_2,\ldots,Y_{100})$에 대하여 다음의 두 가지 변화점 모형을 고려하였다.

- 상황1: Y_t는 평균이 μ_t이고 분산이 $\sigma^2=1$인 정규 난수로 여기서 평균

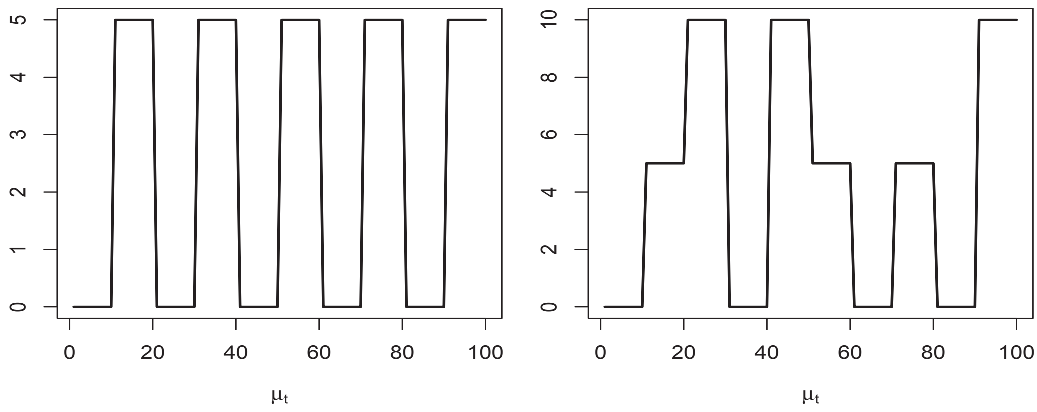
$$\mu_t=\begin{cases} 0, & \text{if } 10(k-1)+1\le t\le 10k,\ k=1,3,5,7,9, \\ 5, & \text{if } 10(k-1)+1\le t\le 10k,\ k=2,4,6,8,10. \end{cases}$$

  따라서 변화점은 $t=11,21,31,41,51,61,71,81,91$로 9개의 변화점이 존재한다. [상황1]의 모든 점들은 국소적 최대/최소(local maximum/minimum)의 형태로만 구성이 되어 있다.

- 상황2: Y_t는 평균이 μ_t이고 분산이 $\sigma^2=1$인 정규 난수로 여기서 평균

$$\mu_t=\begin{cases} 0, & \text{if } 10(k-1)+1\le t\le 10k,\ k=1,4,7,9, \\ 5, & \text{if } 10(k-1)+1\le t\le 10k,\ k=2,6,8, \\ 10, & \text{if } 10(k-1)+1\le t\le 10k,\ k=3,5,10. \end{cases}$$

  [상황2]에서는 [상황1]과 같이 변화점은 $t=11,21,31,41,51,61,71,81,91$로 9개의 변화점이 존재하나 [상황1]과는 다르게 FLR 방법이 이론적으로 변화점을 잘 탐색하지 못한다고 알려진 계단의 형태가 $t\in\{11,12,\ldots,20\}\bigcup\{51,52,\ldots,60\}$에서 발생한다. 이론적으로는 해당 구간에서 많은 위-탐지(false-detection, 비변화점을 변화점으로 탐지)가 발생하게 된다.



(a) Case1 (b) Case2

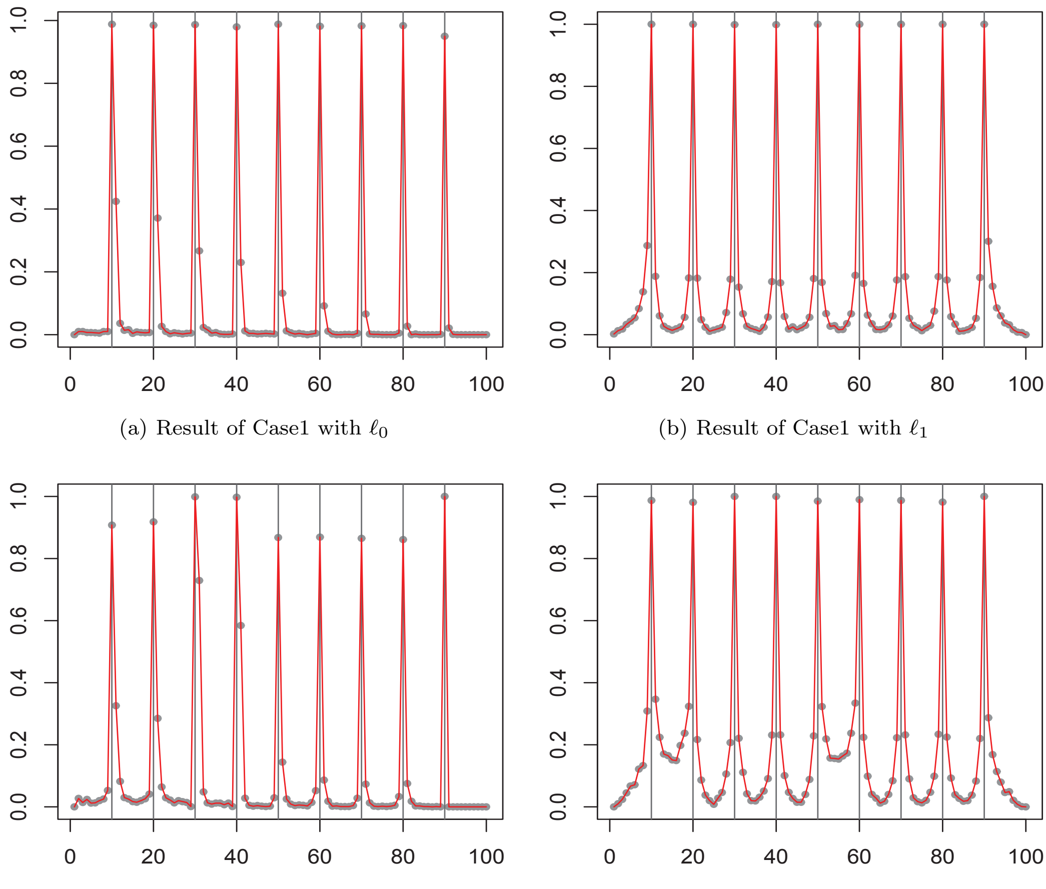
Figure 2.1. True mean signals of simulated data.

독자들의 이해를 돕기 위하여 Figure 2.1에서는 각 상황의 평균 신호를 그림으로 표현하였다.

모의 실험에서는 위의 [상황1]과 [상황2]에 대하여 1,000개의 데이터 셋을 각각 생성하였고 ℓ_0-벌점 최소제곱법은 1절에서 언급한 동적프로그래밍 방법을 R을 이용하여 구현하였고 ℓ_1-벌점 최소제곱법은 R 패키지 "cghFLasso"를 이용하여 구현하였다 (Tibshrani와 Wang, 2008). ℓ_0-벌점 최소제곱법에 대한 동적프로그램의 자세한 절차는 Lim 등 (2012)를 참조하기 바란다. 마지막으로 유한 표본에서 변화점 갯수의 결정과 관련하여 최적의 기준 및 방법 등에 대해 아직 눈에 띄는 이론적인 연구가 발표되지 않았기에 본 모의 실험에서는 ℓ_0-벌점과 ℓ_1-벌점 모두에 있어 경험 법칙(rule of thumb)인 Bayesian Information Criterion(BIC)를 이용하였다.

모의 실험의 결과는 (1) 1,000개의 데이터 셋에서 각 시점 t가 변화점으로 탐색되는 비율, (2) 탐색되는 전체 변화점의 갯수, 두 가지 측도로 정리되었다. 먼저 각 시점 t가 변화점으로 탐색되는 비율은 Figure 2.2로 나타내었으며 Figure 2.2로 부터 다음의 두 내용을 확인 할 수 있다. 첫째, 패널 (a)와 (b)의 비교 그리고 패널 (c)와 (d)의 비교를 통하여 ℓ_0-벌점 최소제곱법이 FLR과 비교하여 위-탐색율(fasle positive rate, 비-변화점을 변화점으로 판단하는 비율)이 상대적으로 낮음을 확인 할 수 있다. 참-탐색률(true positive rate, 실제 변화점을 변화점으로 판단하는 비율) 관점에서는 두 방법 모두 잘 작동하고 차이가 없었다. 둘째, 패널 (b)와 (d)의 비교는 국소적 최대/최소(local maximum/minimum)형태를 지닌 구간 안의 점들보다 계단(stair)형 구간 안의 점들에서 위-탐색율이 높게 나타나고 있음을 보여주고 있다. 이해를 돕기 위하여 계단형 모양의 구간의 점들은 패널 (d)의 $t \in \{11, 12, \ldots, 20\} \bigcup \{51, 52, \ldots, 60\}$이다. 또한 패널 (d) 자체의 국소적 최대/최소 형태의 구간과 계단 형태의 구간 안의 점들의 비교에서도 동일한 결론을 내릴 수 있으며 계단 형태의 구간에서의 높은 위-탐색율과 관련하여 최근 여러 이론적 연구들이 진행 중에 있다 (Rojas와 Wahlberg, 2015; Lin 등, 2016).

다음으로 각 방법으로 부터 탐색된 변화점의 갯수는 Figure 2.3에 히스토그램으로 나타내었으며, 이를 통하여 다음의 세 가지 내용들을 확인 할 수 있다. 첫째, 먼저 [상황1]의 ℓ_0-벌점 방법과 FLR 방법의 차이와 [상황2]의 ℓ_0-벌점과 FLR 방법의 차이를 비교하여 보면 FLR 방법이 보다 많은 점들을 변화점으로 판단하고 있으며 이를 앞의 Figure 2.2의 결과와 연동하여 생각하면 FLR 방법이 비-변화점을 변화점으로 잘못 판단하는 경향이 있음을 확인할 수 있다. 둘째, [상황1]과 [상황2]에서 두 벌점 방법에 의하



(a) Result of Case1 with ℓ_0

(b) Result of Case1 with ℓ_1

(c) Result of Case2 with ℓ_0

(d) Result of Case2 with ℓ_1

Figure 2.2. Empirical detection rates of t as a change point.

여 추가적으로 탐색된 변화점들의 산포를 각각 살펴보면 FLR 방법의 퍼짐의 정도가 ℓ_0-벌점 방법보다 훨씬 크게 나타나는 것을 볼 수 있다. 마지막으로 앞의 Figure 2.2에서와 같이 [상황1](국소적 최대/최소 형태의 구간들로만 구성)과 [상황2](계단 형태의 구간들이 포함되어 구성)에서 FLR 방법의 결과들을 비교하여 보면 계단 형태의 구간들을 포함하고 있는 [상황2]에서 더 많은 변화점들을 감지하고 있고 이러한 추가적인 변화점 탐색은 계단 형태의 구간에서 발생하는 다수의 위-변화점들의 영향이라 판단된다.

추가적으로 변화점의 갯수를 참으로 추정하는 (변화점의 갯수가 9인) 경우에 대하여 참-탐색률을 Table 2.1에 나타내었다. 변화점의 갯수가 참으로 추정된 데이터 셋의 수(#(CPs = 9))는 ℓ_0-벌점 방법에서 상대적으로 크게 나타나며, ℓ_1-벌점 방법에서는 [상황1]에서만 1,000개의 데이터 셋 중에서 5회 추정되었으며, [상황2]에서는 변화점의 갯수를 참으로 추정하는 경우가 나타나지 않았다. 변화점의 갯수가 참으로 추정된 경우, 두 방법 모두 매우 높은 참-탐색률을 나타내는 것을 확인할 수 있다.

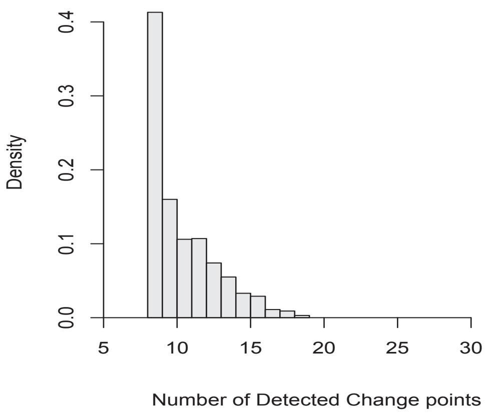
마지막으로 본 절의 모의 실험의 결과를 정리하면 다음과 같다.

1. ℓ_1-벌점 최소제곱법은 ℓ_0-벌점 최소제곱법에 비하여 많은 변화점을 감지하고 있으며 이 중 일부는 비

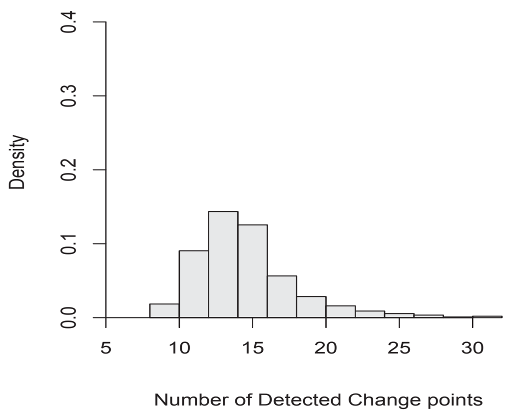
Table 2.1. The true positive rates at the true change points among data sets where the number of change points (CPs) is estimated correctly

Case	Method	#(CPs = 9)	t								
			10	20	30	40	50	60	70	80	90
Case1	ℓ_0	384	98.2	98.2	99.0	96.9	98.2	99.0	99.2	98.7	97.9
	ℓ_1	5	100	100	100	100	100	100	100	100	100
Case2	ℓ_0	157	86.6	91.7	99.4	100.0	87.9	85.4	85.4	88.5	100.0
	ℓ_1	0	0	0	0	0	0	0	0	0	0

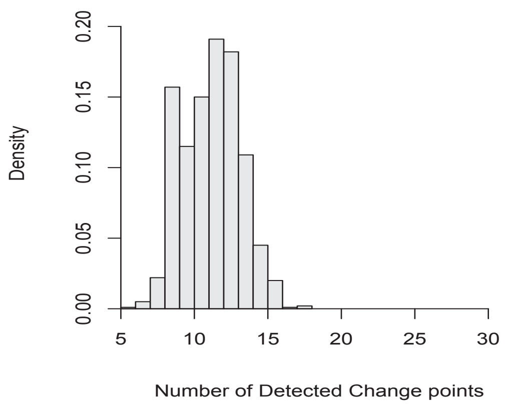
#(CPs = 9) denotes the number of data sets in which each method estimated the number of CPs correctly.



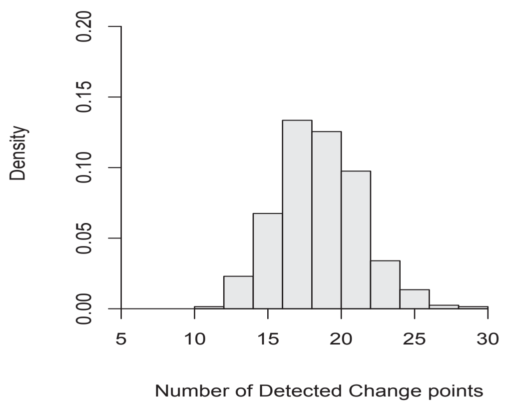
(a) Result of Case1 with ℓ_0



(b) Result of Case1 with ℓ_1



(c) Result of Case2 with ℓ_0



(d) Result of Case2 with ℓ_1

Figure 2.3. Histograms of the number of detected change points (CPs) among 1,000 data sets. The true number of CPs is 9.

변화점을 변화점으로 잘못 판단한 경우이다. 참-변화점의 감지에 있어서는 두 방법 모두 좋은 성능을 보이고 있다.

2. 특히 ℓ_1-벌점 방법은 계단 형태의 구간에 해당하는 점들에 있어 국소 최대/최소 형태의 구간의 점들과 비교하여 위-탐색률이 높아지는 경향이 있다.

3. 결론

본 연구에서는 ℓ_0와 ℓ_1-벌점 최소제곱법을 이용한 다중-변화점 탐색에 대하여 살펴보았다. ℓ_1-벌점 최소제곱법(또는 FLR)은 최근 많은 관심을 받고 있으며 다양한 분야에 적용, 활용되고 있으나 변화점 탐색에 있어서는 특정한 상황 하에서 점근적 일치성을 보장하지 못하는 단점이 있다. 이에 반하여 ℓ_0-벌점 최소제곱법은 동적프로그래밍을 통하여 간편하게 계산이 가능함에도 불구하고 아직까지 많은 연구가 이루어지지는 않았다. 본 연구에서는 ℓ_0-벌점 최소제곱법과 FLR방법을 모의 실험을 통하여 비교하였으며 ℓ_0-벌점 최소제곱법이 변화점의 탐색 관점에서 보다 좋은 성능을 보임을 확인 하였다.

References

Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application* (Vol. 104), Prentice Hall, Englewood Cliffs.

Carlstein, E., Müller, H.-G., and Siegmund, D. (1994). *Change-point Problems*, Institute of Mathematical Statistics, California.

Chen, J. and Gupta, A. K. (2001). On change point detection and estimation, *Communications in Statistics - Simulation and Computation*, **30**, 665–697.

Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*, John Wiley & Sons, New York.

Harchaoui, Z. and Lévy-Leduc, C. (2010). Multiple change-point estimation with a total variation penalty, *Journal of the American Statistical Association*, **105**, 1480–1493.

Jang, W., Lim, J., Lazar, N. A., Loh, J. M., and Yu, D. (2015). Some properties of generalized fused lasso and its applications to high dimensional data, *Journal of the Korean Statistical Society*, **44**, 352–365.

Johnson, N. A. (2013). A dynamic programming algorithm for the fused Lasso and L_0-segmentation, *Journal of Computational and Graphical Statistics*, **22**, 246–260.

Kotz, S., Read, C. B., Balakrishnan, N., Vidakovic, B., and Johnson, N. L. (Eds.) (2006). *Encyclopedia of Statistical Sciences* (2nd ed.), John Wiley & Sons, NJ.

Lim, E., Hahn, K. S., Lim, J., Kim, M., Park, J., and Yoon, J. (2012). Statistical properties of news coverage data, *Communications for Statistical Applications and Methods*, **19**, 771–780.

Lin, K., Sharpnack, J., Rinaldo, A., and Tibshirani, R. J. (2016). Approximate recovery in changepoint Problems, from ℓ_2 estimation error rates, *arXiv preprint, arXiv:1606.06746*.

Qian, J. and Jia, J. (2016). On stepwise pattern recovery of the fused Lasso, *Computational Statistics and Data Analysis*, **94**, 221–237.

Rinaldo, A. (2009). Properties and refinements of the fused lasso, *The Annals of Statistics*, **37**, 2922–2952.

Rojas, C. R. and Wahlberg, B. (2015). How to monitor and mitigate stair-casing in L1 trend filtering, *arXiv preprint, arXiv:1412.0607v1*.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005). Sparsity and smoothness via the fused lasso, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**, 91–108.

Tibshirani, R. and Wang, P. (2008). Spatial smoothing and hot spot detection for CGH data using the fused lasso, *Biostatistics*, **9**, 18–29.

Ye, G.-B. and Xie, X. (2011). Split Bregman method for large scale fused Lasso, *Computational Statistics and Data Analysis*, **55**, 1552–1569.

Yu, D., Won, J., Lee, T., Lim, J., and Yoon, S. (2015a). High-dimensional fused lasso regression using majorization-minimization and parallel processing, *Journal of Computational and Graphical Statistics*, **24**, 121–153.

Yu, D., Lee, S. J., Lee, W. J., Kim, S. C., Lim, J., and Kwon, S. W. (2015b). Classification of spectral data using fused lasso logistic regression, *Chemometrics and Intelligent Laboratory Systems*, **142**, 70–77.

Zhao, P. and Yu, B. (2006). On model selection consistency of Lasso, *Journal of Machine Learning Research*, **7**, 2541–2563.

벌점-최소제곱법을 이용한 다중 변화점 탐색

손원[a] · 임요한[a] · 유동현[b,1]

[a]서울대학교 통계학과, [b]계명대학교 통계학과

(2016년 9월 26일 접수, 2016년 9월 30일 수정, 2016년 9월 30일 채택)

요 약

본 연구에서는 다중 변화점 탐색과 관련하여 최근 많은 관심을 받고 있는 ℓ_0-벌점 최소제곱법과 fused-라쏘-회귀(fused lasso regression; FLR)방법을 모의 실험을 통하여 비교하였다. 모의 실험의 결과로 FLR방법은 비-변화점을 변화점으로 잘못 탐색하는 경향이 ℓ_0-벌점 최소제곱법과 비교할 때 상대적으로 높게 나타났으며 ℓ_0-벌점 최소제곱법이 전반적으로 FLR방법에 비하여 좋은 성능을 보였다. 더불어 ℓ_0-벌점 최소제곱법은 동적프로그래밍을 통하여 FLR 방법과 유사하게 효율적인 계산이 가능하다.

주요용어: 변화점 탐색, 복잡도 벌점, 동적프로그래밍, fused-라쏘-회귀

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2015R1C1A1A02036312).

[1]교신저자: (42601) 대구광역시 달서구 달구벌대로 1095, 계명대학교 통계학과. E-mail: dyu3@kmu.ac.kr