Detectability of Changepoints Using the Likelihood Ratio Test Statistic

Katharina Ring[1]; Supervisors: Thomas Grundy[2], Mirjam Kirchner[2]

[1]LMU Munich, Department of Statistics; [2]STOR-i, Lancaster University

Changepoint Detection

A changepoint is the position in a time series where a part of the structure changes as illustrated in Figure 1.

Whether a changepoint occurs can be detected using a Likelihood Ratio Test. For this test, we derive a detection boundary (introduced by Cai et al., 2011), seperating detectable from undetectable changes, in a simulation study.

Structure of simulated data:

- before the change: $y \overset{iid}{\sim} N(0, 1)$
- after the change: $y \overset{iid}{\sim} N(\mu, \sigma^2)$ with $\mu \neq 0$ and/or $\sigma^2 \neq 1$

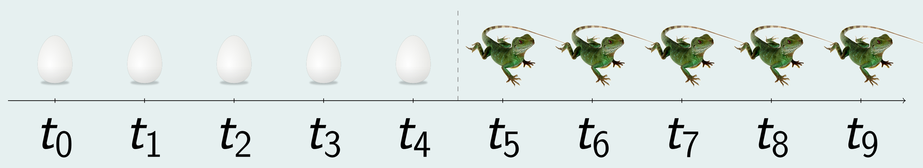


Figure 1: A time series exhibiting a change at t_4

Likelihood Ratio Test

The Likelihood Ratio Test (LRT) can be seen as a binary classifier differentiating between two hypotheses (Eckley et al., 2011).

$$H_0 : \text{no changepoint}$$
$$H_1 : \text{one changepoint}$$

The LRT compares the likelihood without a changepoint to the highest likelihood obtainable with a changepoint. The latter is found by trying all points in time as changepoints and taking the maximum likelihood for these:

$$\lambda = 2\left(\max_{\tau} \underbrace{\left[\log p(y_{1:\tau}|\hat{\theta}_1) + \log p(y_{\tau+1:n}|\hat{\theta}_2)\right]}_{\text{Log-Likelihood with changepoint } \tau} - \underbrace{\log p(y_{1:n}|\hat{\theta})}_{\text{Log-Likelihood without a changepoint}}\right)$$

The null hypothesis is rejected if λ surpasses a given threshold $c \in \mathbb{R}^+$.

Detectability

We define a detectable change as follows: If c is chosen such that the true positive rate (empirical power) is 80 %, the resulting false positive rate (empirical type I error) is at most 5 %.

This corresponds to the *ROC curve*, which gives the false and true positive rates for every c, passing through or above the point (0.05, 0.8).

Here, we chose *points of interest* for the detection boundary as points where the ROC curve passes this point with a distance < 0.01 (the *boundary region*) as illustrated in Figure 2.

The detection boundary can then be estimated by the median of several of these points of interest.

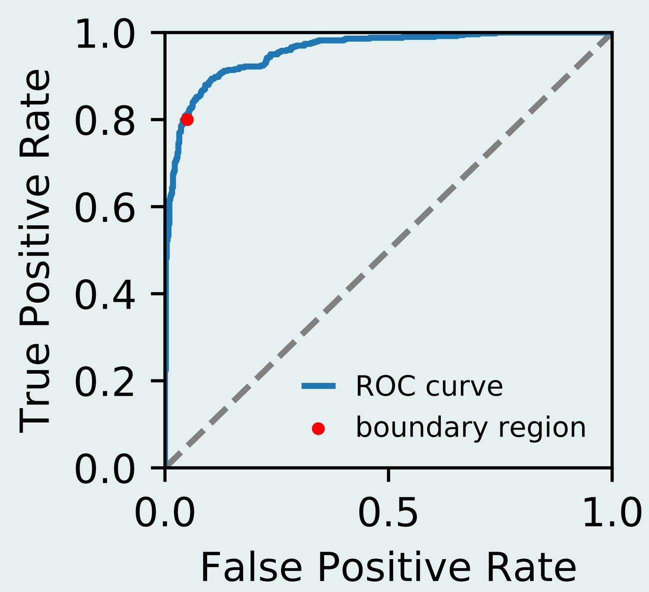


Figure 2: ROC curve example

Surrogate Model Bayesian Optimisation

As 1000 time series were sampled and tested for one ROC curve, evaluating the distance to the ROC curve is computationally expensive.
In order to efficiently find changes where the ROC curve goes through the boundary region, Surrogate Model Bayesian Optimisation was used.

Surrogate modelling (depicted in Figure 3):

1. Evaluate function k times in random positions
2. Fit a surrogate model through current points
3. Use model to determine next evaluation point
4. Considering the new point, jump to 2 unless maximum number of iterations is reached

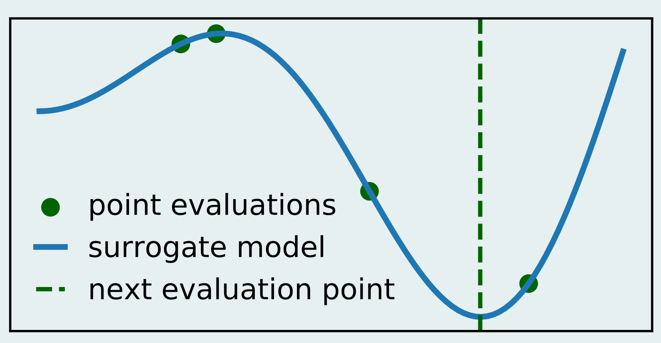
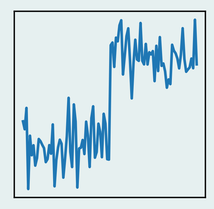


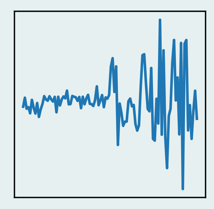
Figure 3: Example of a surrogate model

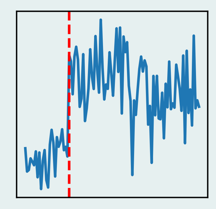
Because of their versatility, we used Gaussian Processes as a surrogate model. The next point to evaluate (Step 3) was determined by the probability of improvement, because of the the convex shape of the function of interest.

Univariate Analysis & Results

In the univariate case, we investigated the influence of the following three variables on detectability:


size of mean change


size of variance change


location of changepoint

The resulting surface separating detectable (above the boundary) from undetectable changes (below the boundary) is shown in Figure 4.

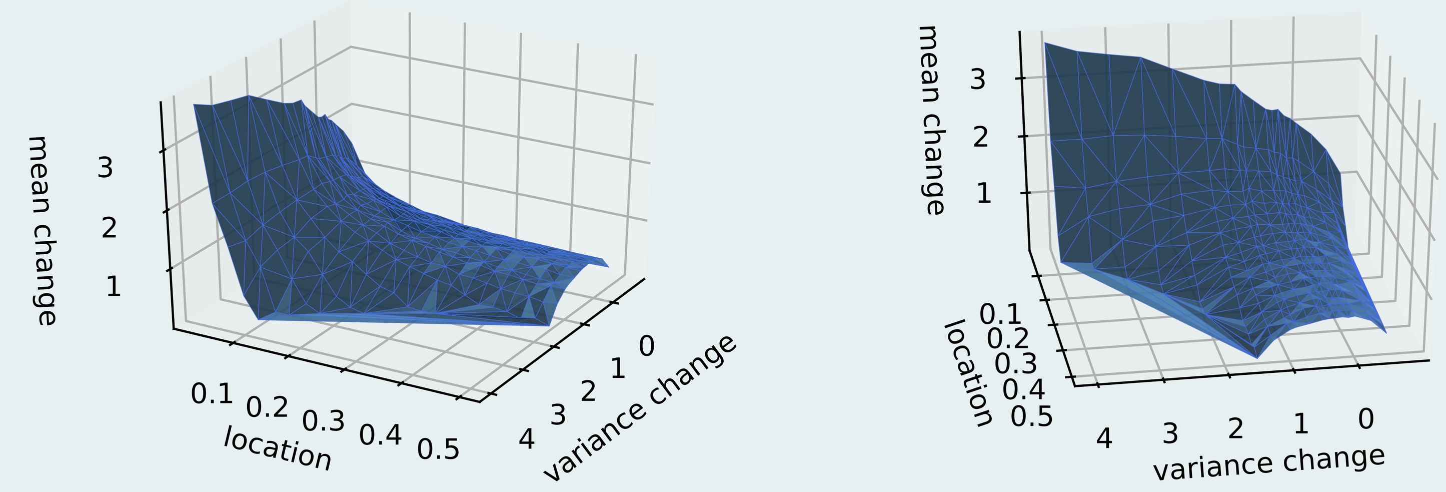
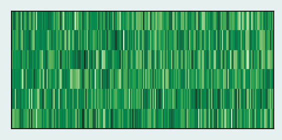
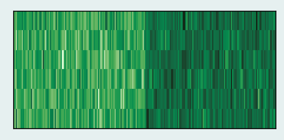


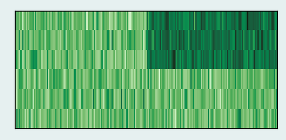
Figure 4: The detection boundary for the univariate case shown from different angles

Multivariate Analysis & Results

In the multivariate case, we investigated the influence of the following three variables on detectability:


number of time series


size of summed mean changes


sparsity of mean changes

The resulting surface separating detectable (above the boundary) from undetectable changes (below the boundary) is shown in Figure 5.

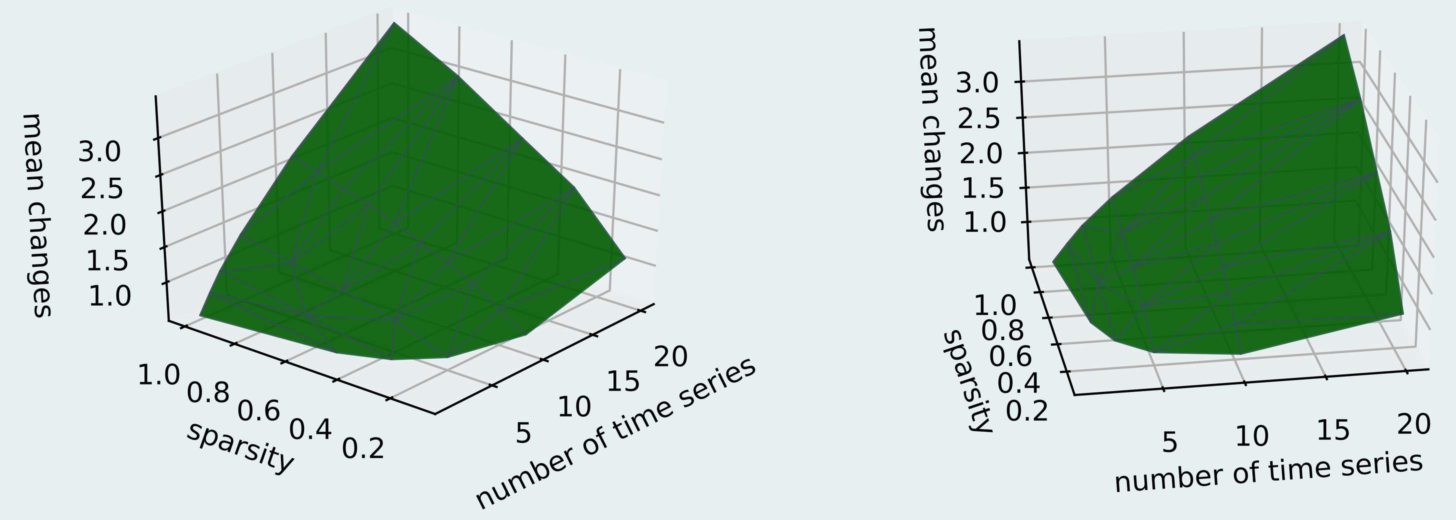


Figure 5: The detection boundary for the multivariate case shown from different angles

Future Work

More dimensions of interest can be investigated:

- univatiate case: length of time series
- multivariate case: length of time series, size of variance change, and location

Additionally, both cases could consider multiple changepoints per time series.

In order to make this computationally feasible, using a *stochastic* surrogate model (e. g. stochastic Gaussian Processes) in the Bayesian optimisation is advisable as it handles the randomness implicitly. Consequently, the optimisation does not have to be repeated to take the median.

A stochastic surrogate model could also improve the results obtained here, especially in the multivariate case, as it directly estimates the true minimum of the stochastic function.

References

Eckley, I., Fearnhead, P., and Killick, R. (2011). Analysis of Changepoint Models. In *Bayesian time series models*, Cambridge University Press.

Cai, T., Jeng, J., and Jin, J. (2011). *J. R. Statist. Soc. B*, 73(5): 629–662. Optimal detection of heterogeneous and heteroscedastic mixtures.