Fractal Test for CUSUM

Moo K. Chung, Anass Y El Yaagoubi Bourakna, Hernando Ombao

University of Wisconsin-Madison, KAUST
mkchung@wisc.edu

Abstract. We present the fractal test for cumulative sums (CUSUM). Given time series, the test can determine if it has at least one change point. If there is a change point, CUSUM will increase or decrease till it hits the next change point. Thus, the change of CUSUM can be used to determine the presence of change point. If there is no change point and no trend present, we expect the time series to follow white noise process. The fractal dimension of white noise is 1.5. By measuring the fractal dimension of time series, we can quantify how much the time series is diverging from white noise. This can be used to build a permutation test. The MATLAB code is provided with examples.

1 CUSUM

The cumulative sum (CUSUM) test is a statistical method used to detect structural breaks or changes in a time series data [3]. The CUSUM test is based on the cumulative sum of deviations of the observed values from some threshold w. Suppose we observe time series x_t. We assume x_t are standardized such that

$$z_t = \frac{x_t - \mu}{\sigma}$$

with mean μ and standard deviation σ of the time series over whole time point. The cumulative sum of deviation from some threshold w of standardized times series z_t is given by

$$s_t = \max(0, s_{t-1} + z_t - \omega).$$

We will call this positive-CUSUM. The large s_t indicates the changes in the positive direction. For negative direction, we use

$$s_t = \min(0, s_{t-1} - x_t - \omega).$$

We will call this negative-CUSUM. If we use $\omega = 0$, the pattern should be symmetric between positive-CUSUM and negative-CUSUM.

If z_t is white noise and there is no preference in the direction of change, s_t is also white noise. White noise is some between 1D straight line and 2D plane. Thus, its fractal dimension is 1.5 [2]. Subsequently, if there is no change point, the fractional dimension of CUSUM will be close to 1.5 (Figure 1). However, if there is change point c and z_t begin to change in the positive direction near c, there will be the cumulation of positive changes that will shoot up s_t. Thus s_t will no longer be white noise but form an increasing curvilinear pattern. The fractional dimension of 1D curve is 1. Thus we use the fractal dimension as a test statistic and test its significance under random permutations. We used the Higuchi's algorithm for computing the fractional dimension of time series [1].

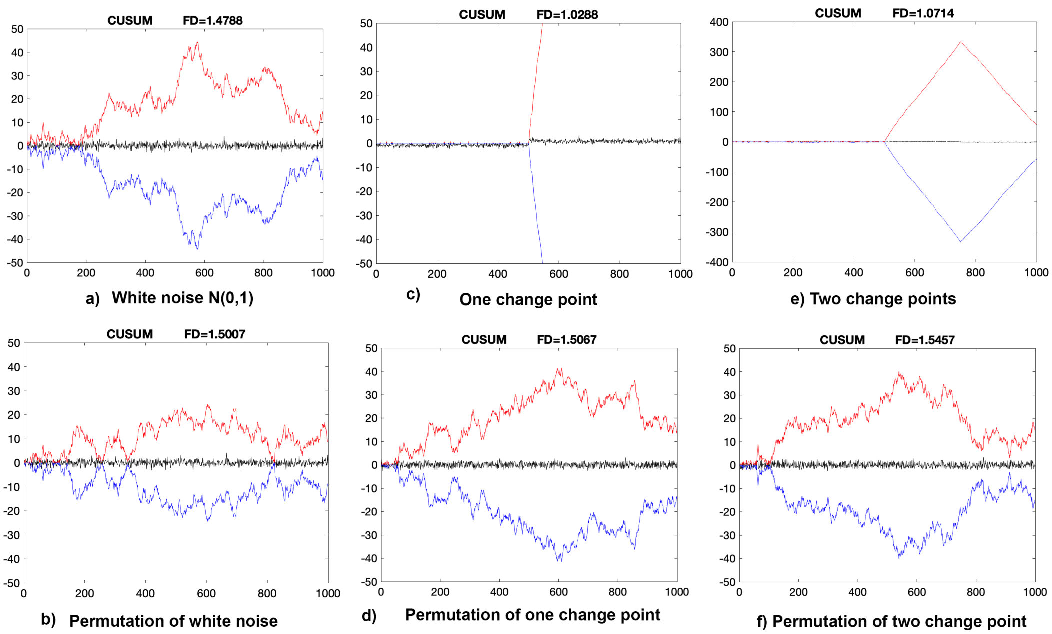


Fig. 1. The fractional dimension of positive- (red) and negative- (blue) CUSUM of signal (black). a) White noise has the fractal dimension of 1.5. Thus simulated signal is close to white noise. b) The permutation of white noise is again white noise so the fractional dimension should be close to 1.5. c) CUSUM on one change point. The fractal dimension of 1D curve is 1. Thus FD=1.0288 shows a strong likelihood there is a change point. d) The permutation will distroy the change point pattern and make the signal white noise. e) CUSUM on one change points. Again FD = 1.0714 shows a strong likelihood there is at least one change point occurred. However, the method cannot detect how many change points are present. f) The permutation will destroy the change point pattern and make the signal white noise.

2 Numerical Implementation

The codes are implemented in MATLAB. Given a vectored time series `x`,

```
z=CUSUM_normalize(y);
```

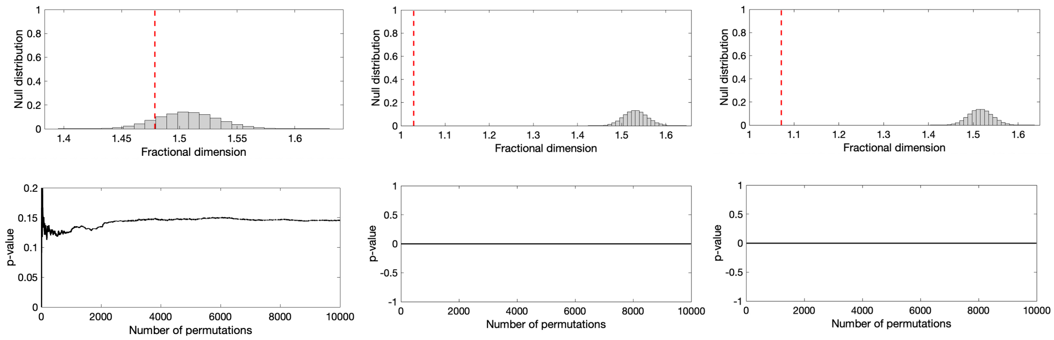
normalize time series from x_t to z_t. The CUSUM is then computed with $\omega = 0$ as

```
w=0
[pos neg] = CUSUM_stat(z,w);
```

which gives postive-CUSUM `pos` and negative-CUSUM `neg`. For more complex time series data, the mean signal w can be time varying, i.e., $w = w(t)$. This is not explored in this paper. The CUSUM is then visualized with

```
CUSUM_display(z,pos,neg)
```

The fractional dimension is then computed as



a) No change point b) One change point c) Two change point

Fig. 2. The empirical null distribution of fractal dimension of signal with no change (a) one change point (b) and two change points (c). Bottom plots show the convergence of the fractal test over the number of permutations used. We obtained p-value of 0.14 indicating it is not likely there is change point present. In (b) and (c), the p-values are all below 0.0001 indicating the existence of at least one change point. As the change point increases, the observed fractal dimension will become close to 1.5. If we have enough data, it may be even possible to develop a method for estimating the the number of change points.

```
oberved = higuchi(pos, 10);
```

This gives the observed fractal dimension. To determine the statistical significance of the fractional dimension, we permute half million times. The fractal dimension is very close to 1.5 for white noise so if the permutation is shuffled well such that no visible pattern is present, we do not need large number of permutations. Using 100000 permutations in 12 seconds in a desktop:

```
for i=1:100000
    zper=z(randperm(n));
    [pos neg] = CUSUM_stat(zper, omega);
    FD = higuchi(pos, 10);
    dist=[dist FD];
end
```

p-values over permutations are then computed as

```
transPval = online_pvalues(dist, observed);
pval = transPval(end)
```

To determine the rate of convergence of the permutation resampling, we do the following. Let p_k be the p-value computed using up to k-th permutation π_k. It is expected the null distribution will be centered around 1.5 while the observed fractal dimension is smaller than 1.5. Thus, we need to perform the one sided test involving the left tail (Figure 2). Note $(k+1)p_{k+1}$ is the total number of permutation resamples that is on the left side of observation. Then p_{k+1} is computed iteratively as

$$(k+1)p_{k+1} = kp_k + \mathcal{I}\big(\texttt{observed} \geq \pi_{k+1}(\texttt{z})\big), \tag{1}$$

for indicator function $\mathcal{I}$ which takes value 1 if the argument is true and 0 otherwise.

References

1. Burns, T., Rajan, R.: Combining complexity measures of EEG data: multiplying measures reveal previously hidden information. F1000Research **4**, 137 (2015)
2. Mandelbrot, B.: The fractal geometry of nature. WH freeman New York (1982)
3. Ploberger, W., Krämer, W., Kontrus, K.: A new test for structural stability in the linear regression model. Journal of Econometrics **40**, 307–318 (1989)