Nonparametric Estimation of Change Points and Stationarity in Finance

David S. Matteson
Department of Statistical Science
Cornell University

matteson@cornell.edu
www.stat.cornell.edu/~matteson

Joint work with: Nicholas A. James (ORIE), William B. Nicholson (STAT) and Louis C. Segalini (FE), Cornell University

Sponsorship: National Science Foundation

2013 May 17

Change Point Analysis

The process of detecting distributional changes within time ordered data

- Changes is mean, variance, skew, tail, correlation,...

Framework:

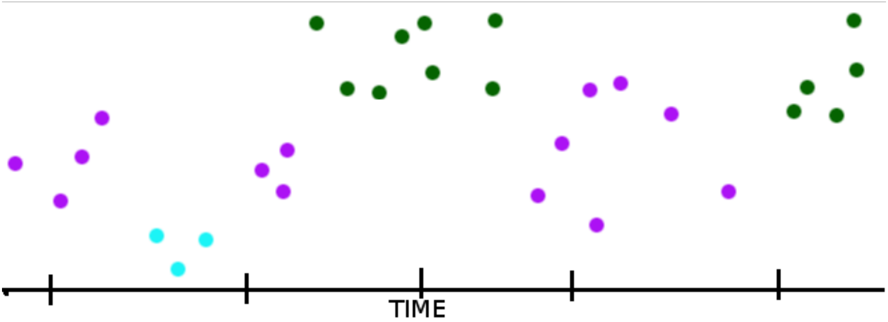
- Retrospective, offline analysis
  - Dynamic Programming for adaptive, online analysis
- Multivariate observations
- Estimation: number of change points and their positions
- Hierarchical algorithms
  - **Divisive:** Data are divided
  - **Agglomerative:** Data are merged

Cluster Analysis

Change point analysis is similar to cluster analysis

In cluster analysis we also wish to partition the observations into homogeneous subsets

- Subsets may not be contiguous in time without some constraints

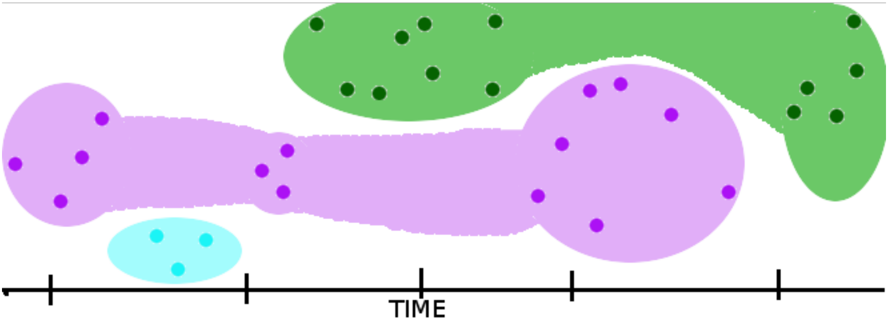


Cluster Analysis

Change point analysis is similar to cluster analysis

In cluster analysis we also wish to partition the observations into homogeneous subsets

- Subsets may not be contiguous in time without some constraints

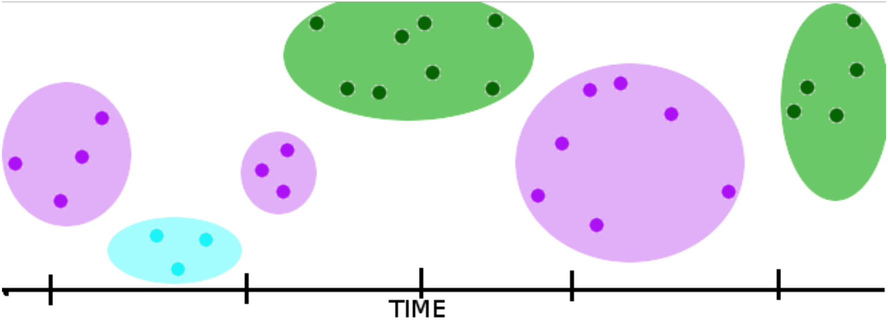


Cluster Analysis

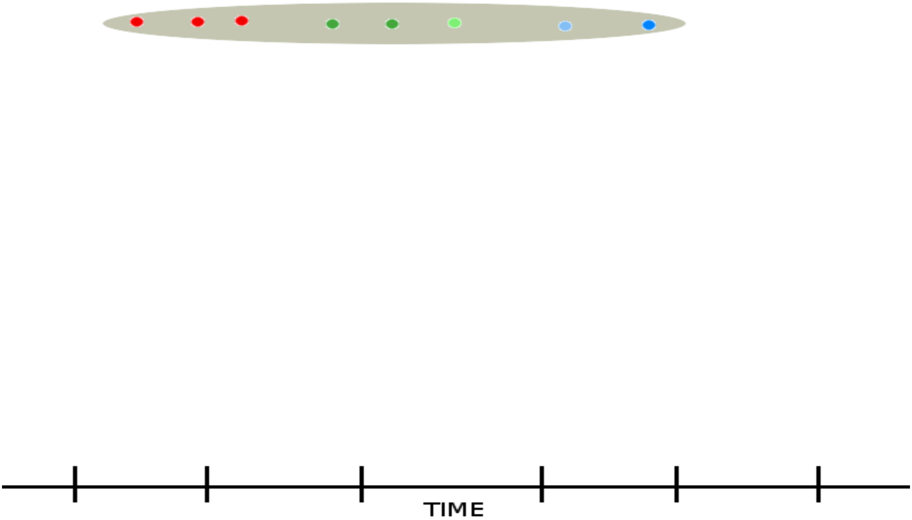
Change point analysis is similar to cluster analysis

In cluster analysis we also wish to partition the observations into homogeneous subsets

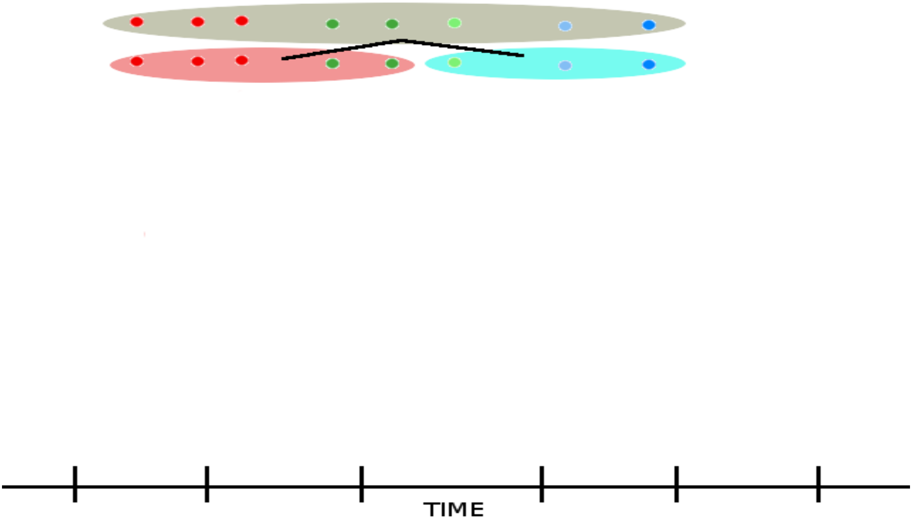
- Subsets may not be contiguous in time without some constraints



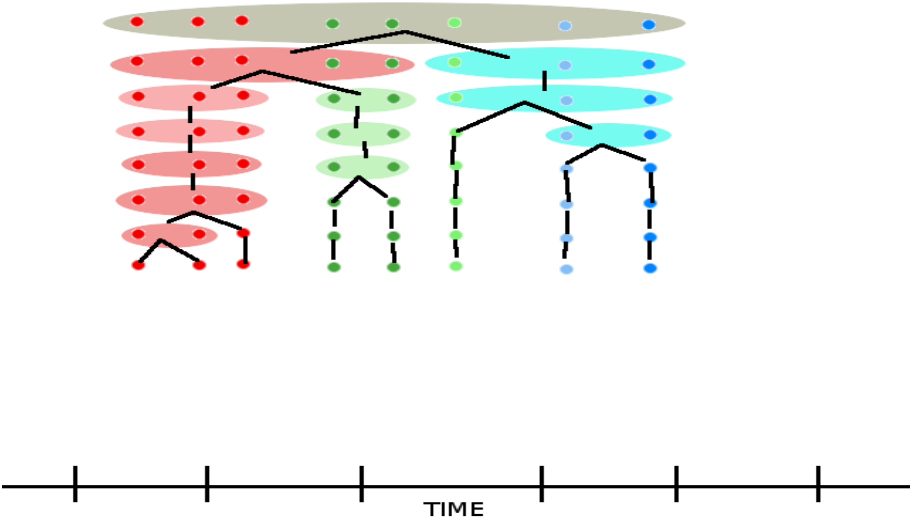
Hierarchical Estimation: Divisive Progression



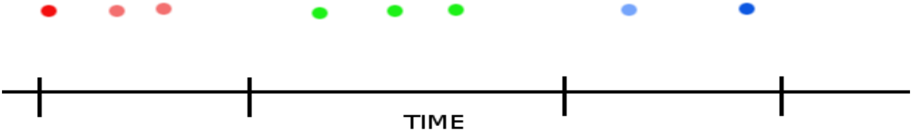
Hierarchical Estimation: Divisive Progression



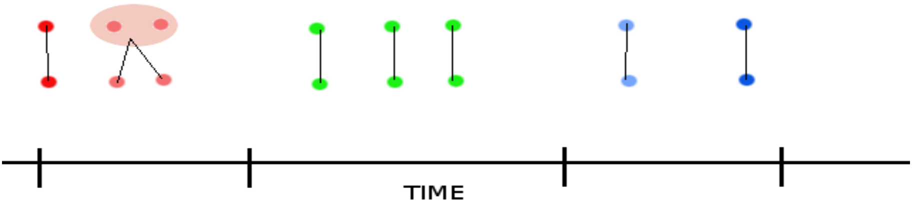
Hierarchical Estimation: Divisive Progression



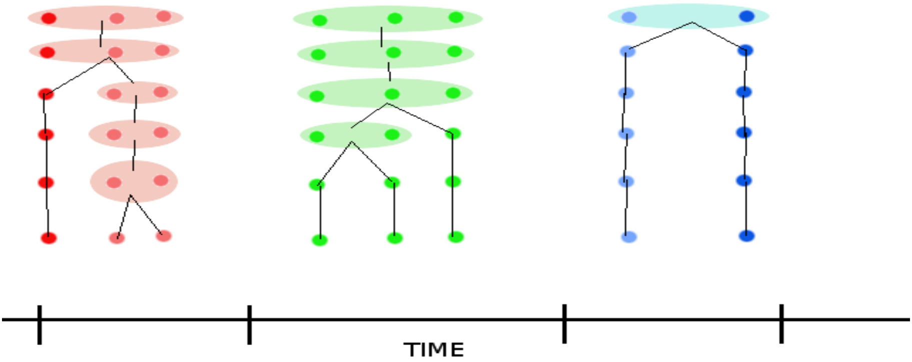
Hierarchical Estimation: Agglomerative Progression



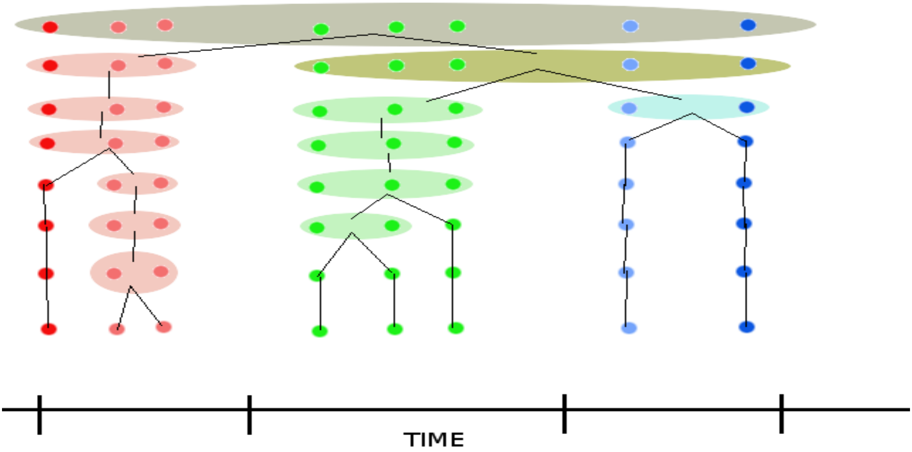
Hierarchical Estimation: Agglomerative Progression



Hierarchical Estimation: Agglomerative Progression



Hierarchical Estimation: Agglomerative Progression

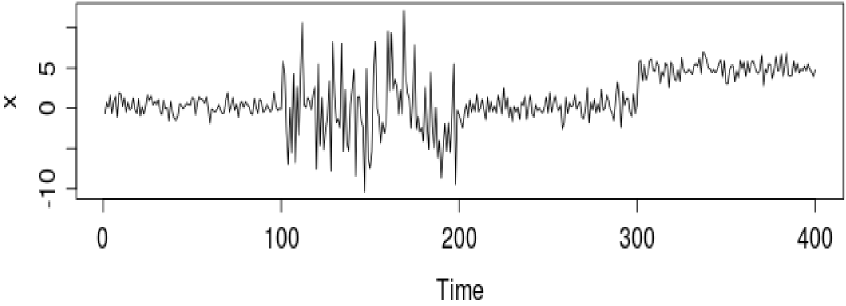


Change Point Analysis

Given independent, time ordered observations $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$

Partition into k homogeneous, temporally contiguous subsets

- k is unknown
- Size of each subset is unknown

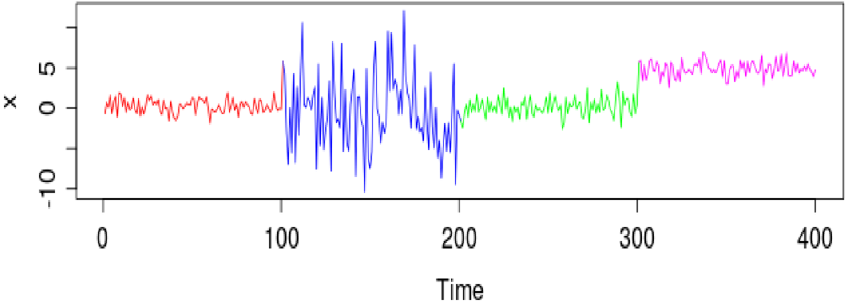


Change Point Analysis

Given independent, time ordered observations $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$

Partition into k homogeneous, temporally contiguous subsets

- k is unknown
- Size of each subset is unknown

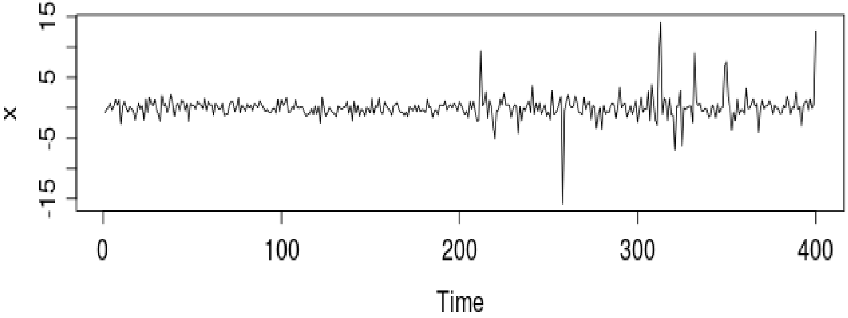


Change Point Analysis

Given independent, time ordered observations $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$

Partition into k homogeneous, temporally contiguous subsets

- k is unknown
- Size of each subset is unknown

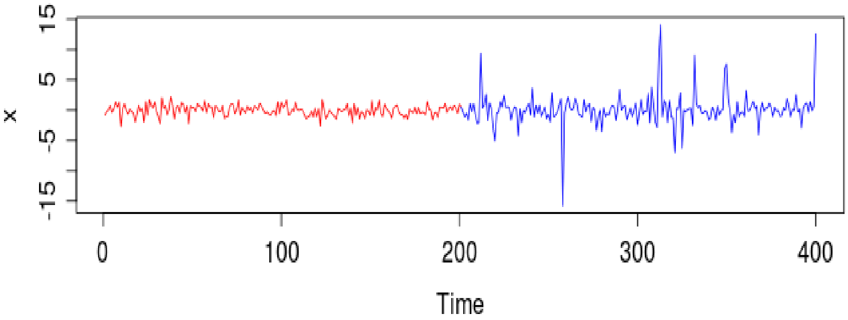


Change Point Analysis

Given independent, time ordered observations $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$

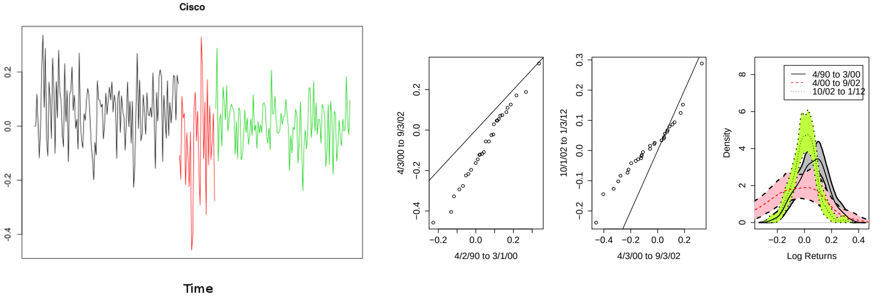
Partition into k homogeneous, temporally contiguous subsets

- k is unknown
- Size of each subset is unknown



Cisco Systems Inc.

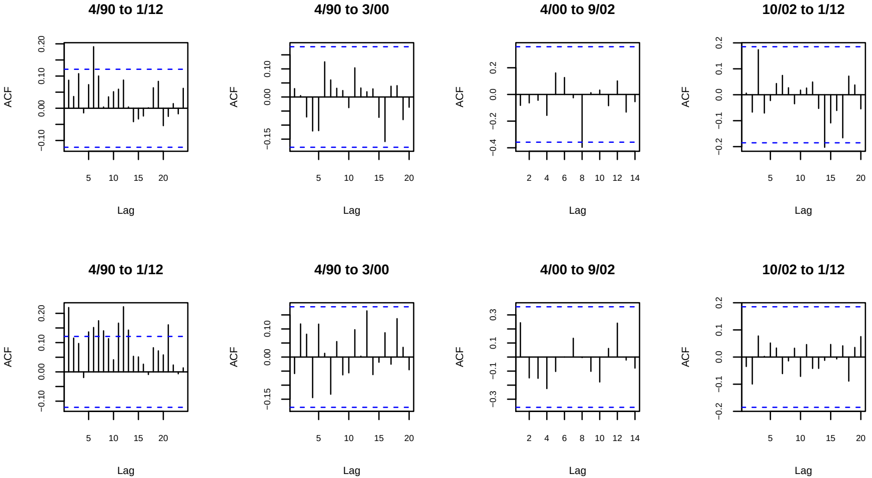
Monthly log returns from April 1990 to January 2010



Change points estimated at April 2000 and October 2002

Cisco Systems

Within cluster diagnostics: $\widehat{ACF}(r_{i,t})$ (top) and $\widehat{ACF}(r_{i,t}^2)$ (bottom)



Measuring Multivariate Homogeneity

Suppose $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$ with $\mathbf{X} \sim F_x \perp\!\!\!\perp \mathbf{Y} \sim F_y$

Let $\phi_x(t) = \mathbb{E}\big(e^{i\langle t, \mathbf{X}\rangle}\big)$ and $\phi_y(t) = \mathbb{E}\big(e^{i\langle t, \mathbf{Y}\rangle}\big)$ characteristic functions

Define a divergence between F_x and F_y as

$$\mathcal{E}(\mathbf{X}, \mathbf{Y}; w) = \int_{\mathbb{R}^d} |\phi_x(t) - \phi_y(t)|^2 \, w(t) \, dt,$$

$w(t)$ denotes an arbitrary positive weight function, for which $\mathcal{E}$ exists

Equivalent Divergence Measures

Let $\mathbf{X}$ and $\mathbf{Y}$ be independent, and $(\mathbf{X}', \mathbf{Y}')$ be an iid copy of $(\mathbf{X}, \mathbf{Y})$

Theorem

Suppose that $\mathbb{E}(|\mathbf{X}|^\alpha + |\mathbf{Y}|^\alpha) < \infty$, *for some* $\alpha \in (0, 2]$, *then*

$$\begin{aligned}
\mathcal{E}(\mathbf{X}, \mathbf{Y}; \alpha) &= \int_{\mathbb{R}^d} |\phi_x(t) - \phi_y(t)|^2 \left(\frac{2\pi^{d/2}\Gamma(1-\alpha/2)}{\alpha 2^\alpha \Gamma((d+\alpha)/2)} |t|^{d+\alpha} \right)^{-1} dt \\
&= 2\mathbb{E}|\mathbf{X} - \mathbf{Y}|^\alpha - \mathbb{E}|\mathbf{X} - \mathbf{X}'|^\alpha - \mathbb{E}|\mathbf{Y} - \mathbf{Y}'|^\alpha \\
&< \infty
\end{aligned}$$

- *If* $0 < \alpha < 2$ *then* $\mathcal{E}(\mathbf{X}, \mathbf{Y}; \alpha) = 0$ *if and only if* $\mathbf{X}$ *and* $\mathbf{Y}$ *are identically distributed*
- *If* $\alpha = 2$ *then* $\mathcal{E}(\mathbf{X}, \mathbf{Y}; \alpha) = 0$ *if and only if* $\mathbb{E}\mathbf{X} = \mathbb{E}\mathbf{Y}$ *(equal means)*

An Empirical Measure (U-statistics)

Let $\mathbf{X}_n = \{X_i : i = 1, \ldots, n\}$ and $\mathbf{Y}_m = \{Y_j : j = 1, \ldots, m\}$ be independent iid samples from the distribution of $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$, respectively, such that $E|\mathbf{X}|^\alpha, E|\mathbf{Y}|^\alpha < \infty$ for some $\alpha \in (0, 2]$

Define

$$\widehat{\mathcal{E}}(\mathbf{X}_n, \mathbf{Y}_m; \alpha) =$$

$$\frac{2}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} |X_i - Y_j|^\alpha - \binom{n}{2}^{-1} \sum_{1 \le i < k \le n} |X_i - X_k|^\alpha - \binom{m}{2}^{-1} \sum_{1 \le j < k \le m} |Y_j - Y_k|^\alpha$$

and

$$\widehat{\mathcal{Q}}(\mathbf{X}_n, \mathbf{Y}_m; \alpha) = \frac{mn}{m+n} \widehat{\mathcal{E}}(\mathbf{X}_n, \mathbf{Y}_m; \alpha)$$

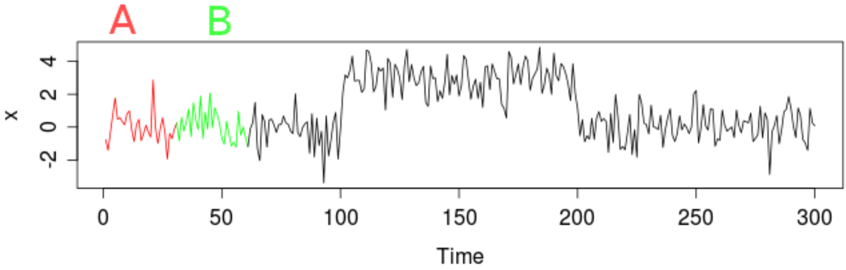
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \; \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets **A** and **B**:



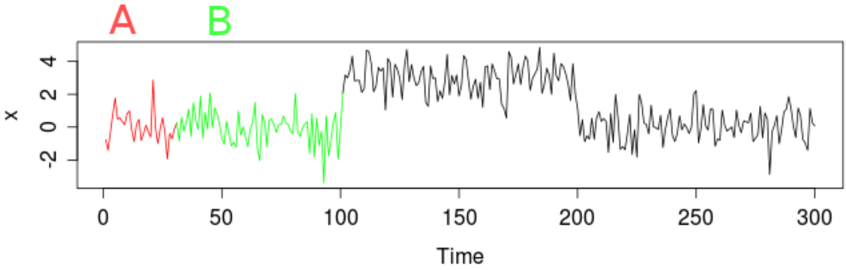
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \; \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets $\mathbf{A}$ and $\mathbf{B}$:



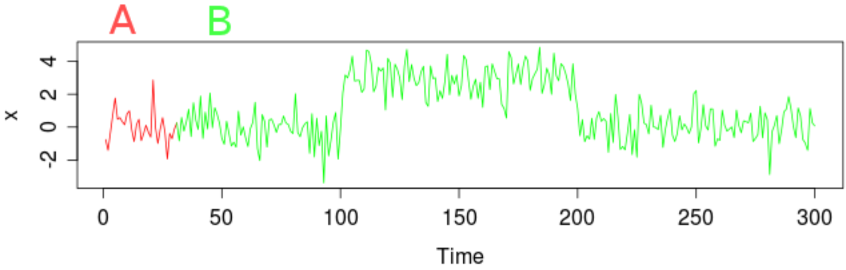
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \ \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets **A** and **B**:



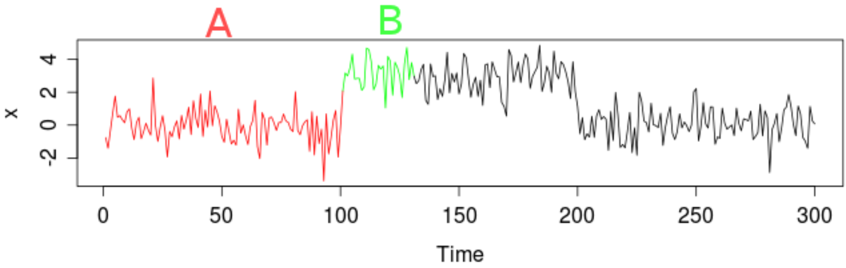
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \; \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets $\mathbf{A}$ and $\mathbf{B}$:



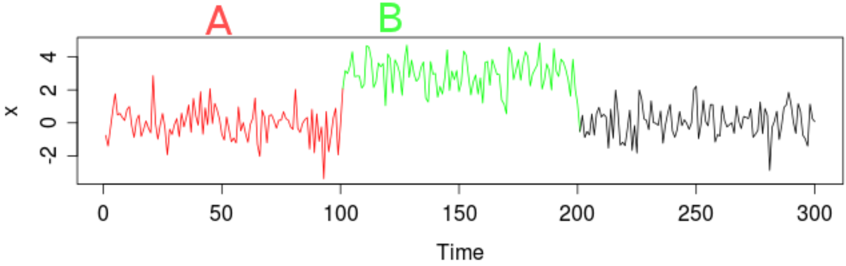
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \ \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets $\mathbf{A}$ and $\mathbf{B}$:



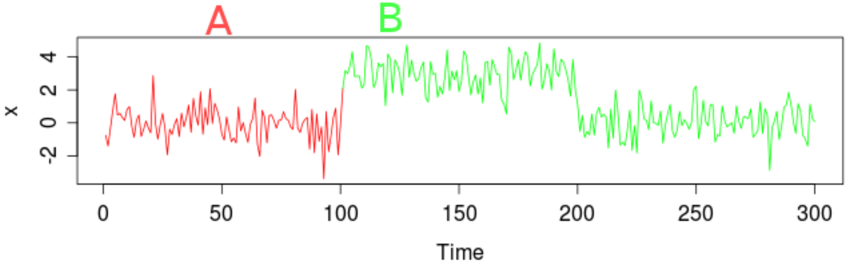
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \; \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets $\mathbf{A}$ and $\mathbf{B}$:



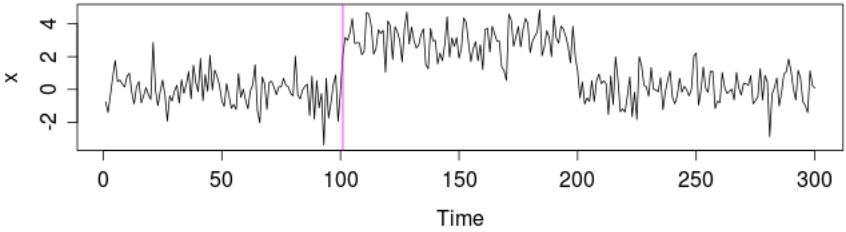
The E-Divisive Algorithm: Estimating Location

$\mathbf{A}_\tau = \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\mathbf{B}_\tau(\kappa) = \{Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_\kappa\}$

A change point location $\hat{\tau}$ is estimated as

$$(\hat{\tau}, \hat{\kappa}) = \underset{(\tau,\kappa)}{\operatorname{argmax}} \ \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)$$

Thus, we maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$ for all subsets $\mathbf{A}$ and $\mathbf{B}$:

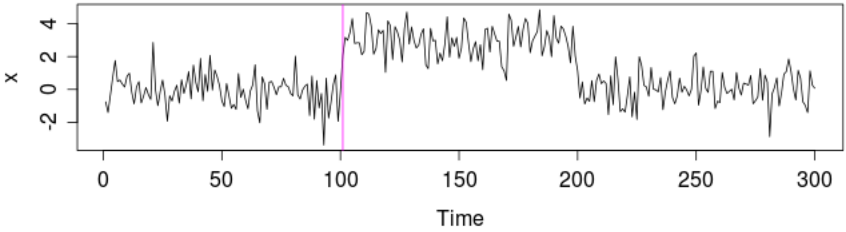


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

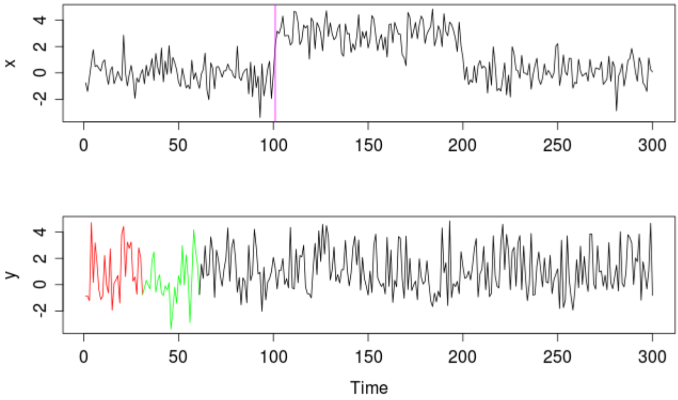


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

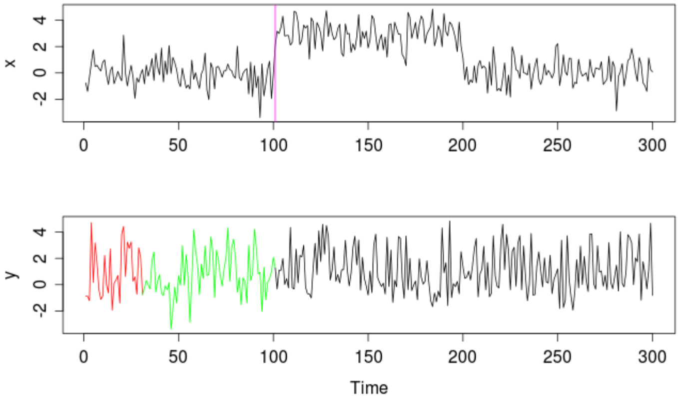


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

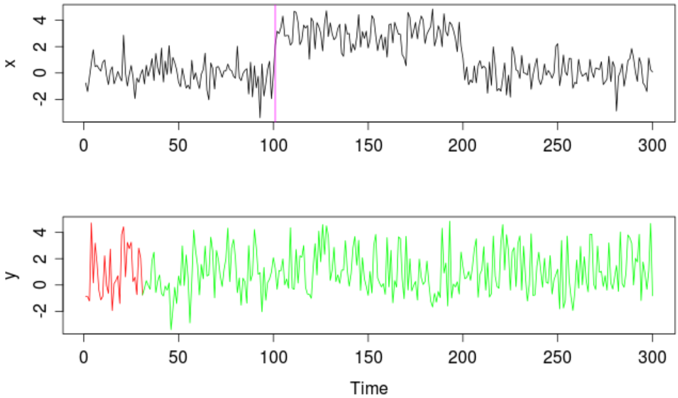


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

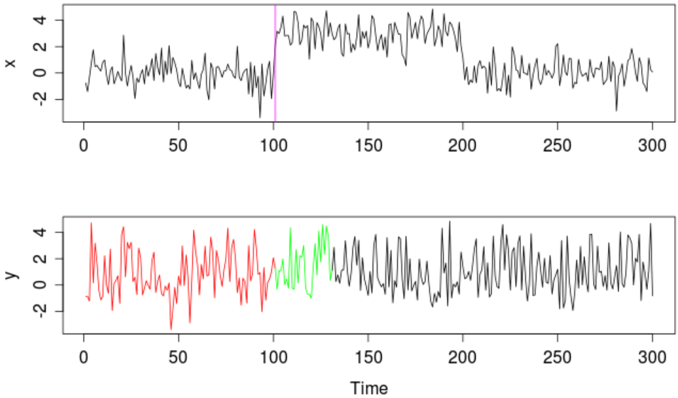


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\Big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

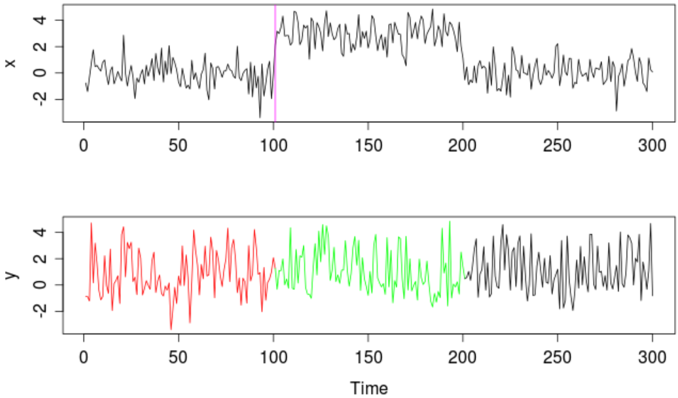


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:

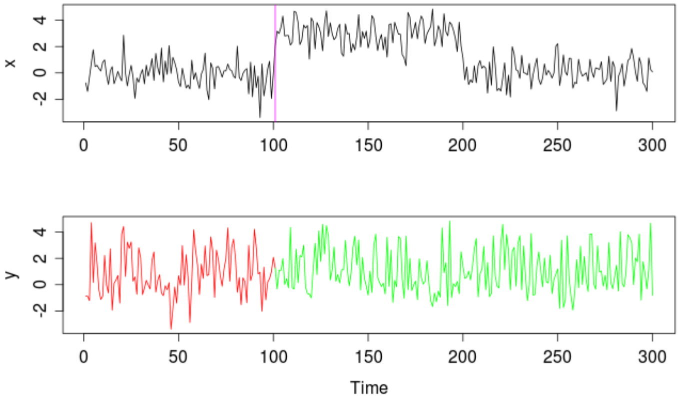


The E-Divisive Algorithm: Inference via Permutation Test

Distribution of test statistic $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is unknown

Significance of proposed change point measured via permutation test

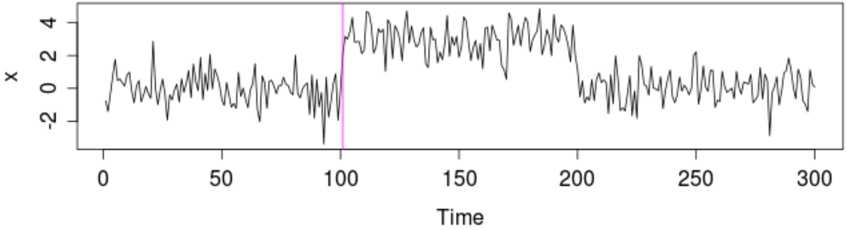
Randomly permute series, maximize $\frac{mn}{n+m}\widehat{\mathcal{E}}(\mathbf{A}, \mathbf{B}; \alpha)$, record and repeat:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

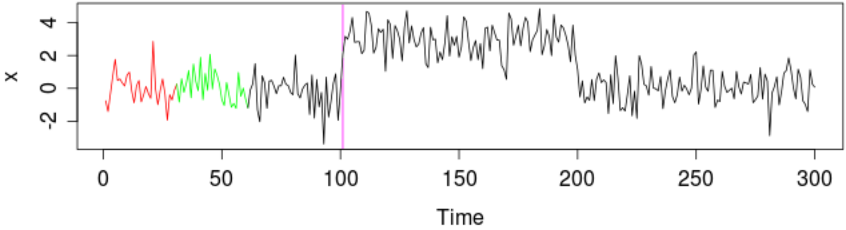
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

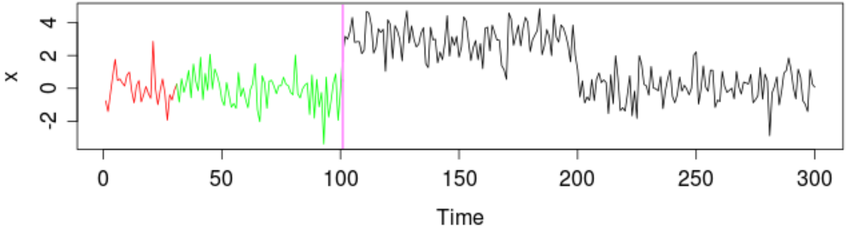
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\Big|_{\tau=\hat{\tau}}$ is insignificant: STOP

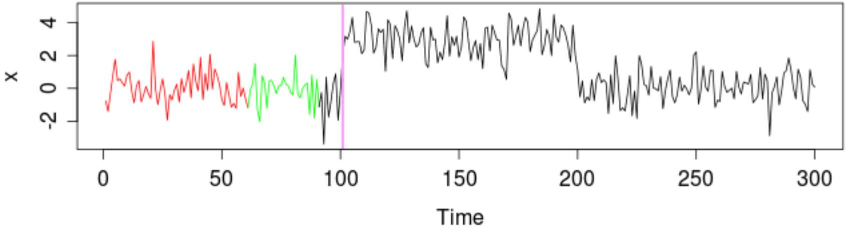
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

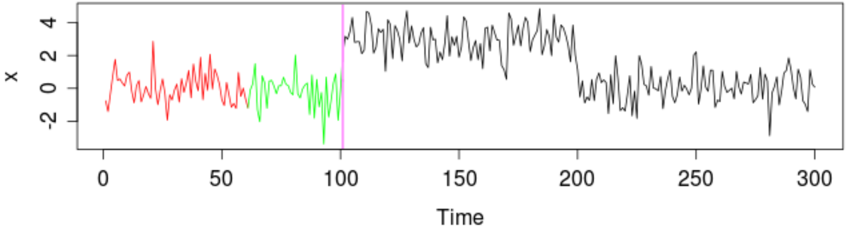
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

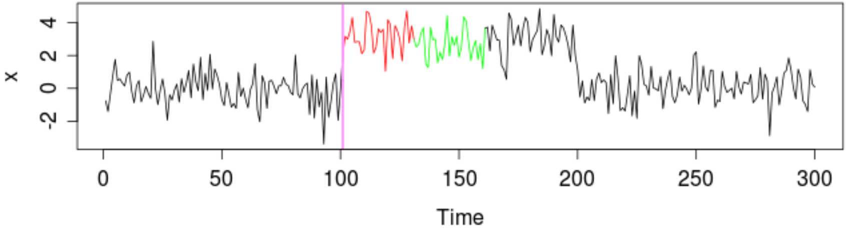
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

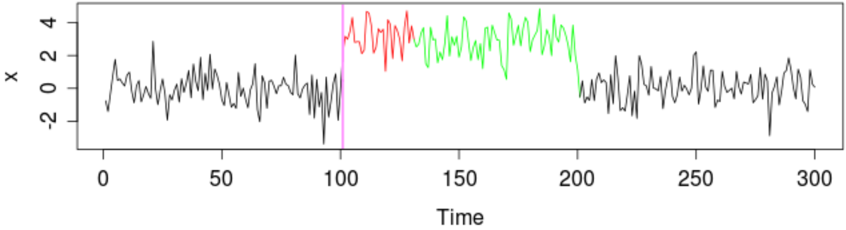
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

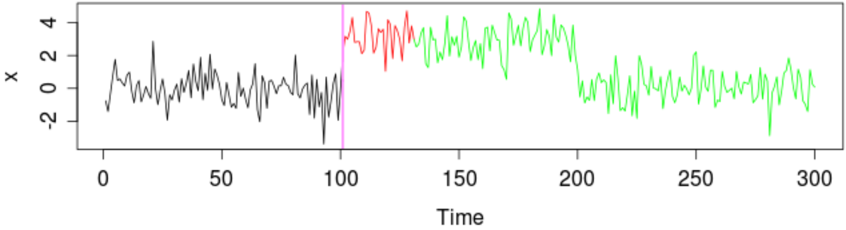
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\Big|_{\tau=\hat{\tau}}$ is insignificant: STOP

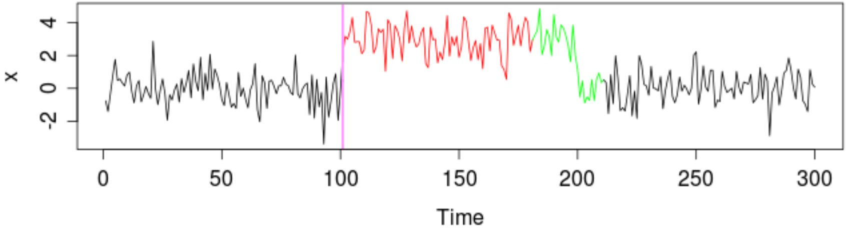
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

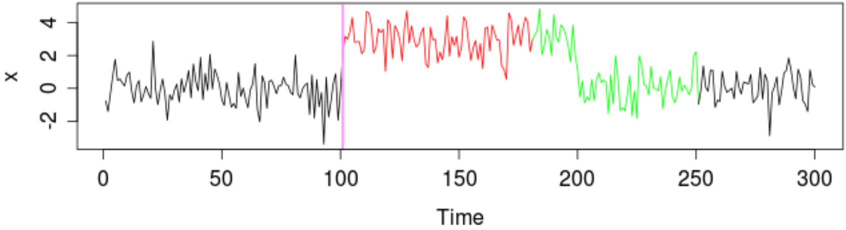
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

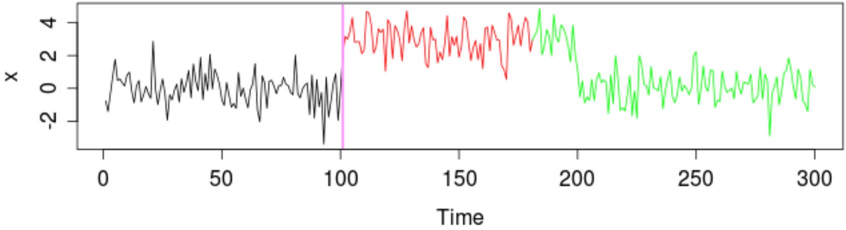
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

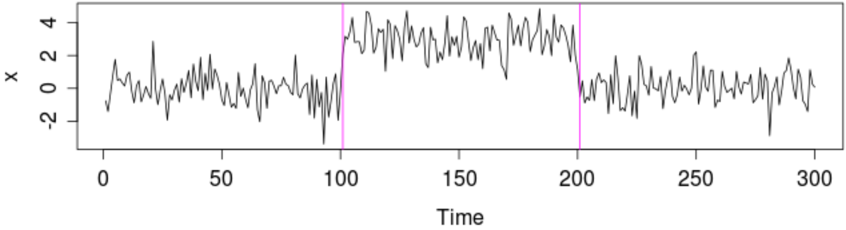
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

If $\hat{q}^* = \widehat{\mathcal{Q}}(\mathbf{A}_\tau, \mathbf{B}_\tau(\kappa); \alpha)\big|_{\tau=\hat{\tau}}$ is insignificant: STOP

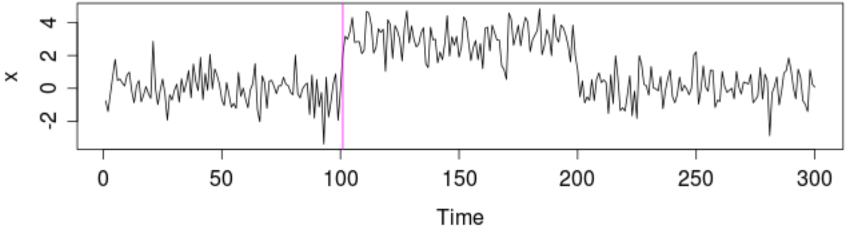
If significant, condition on location, and repeat within clusters:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

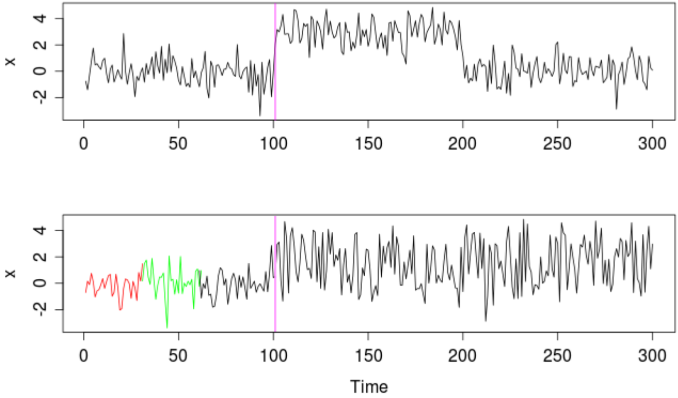
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

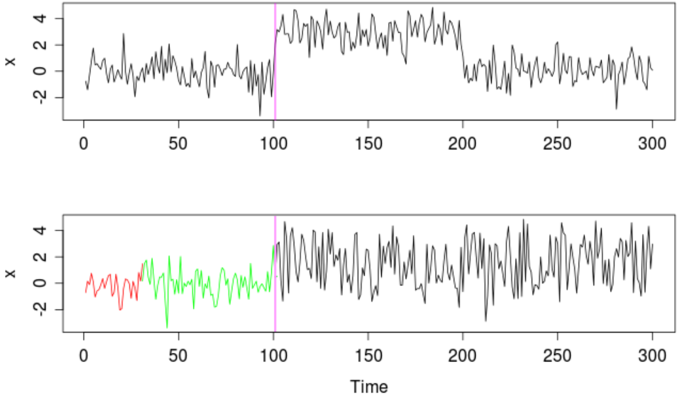
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

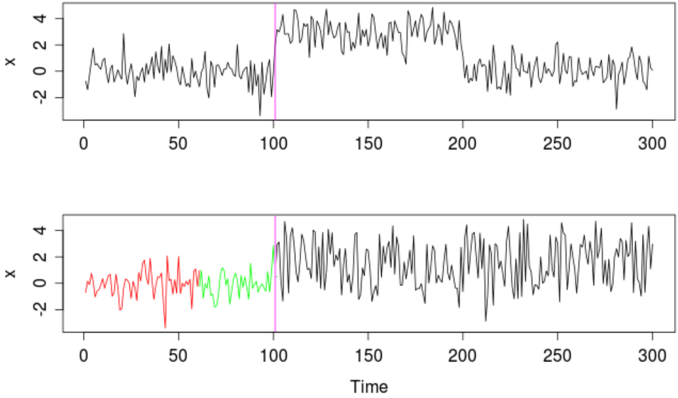
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

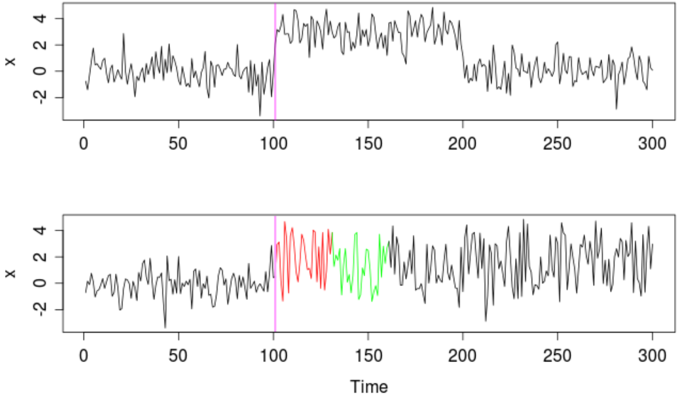
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

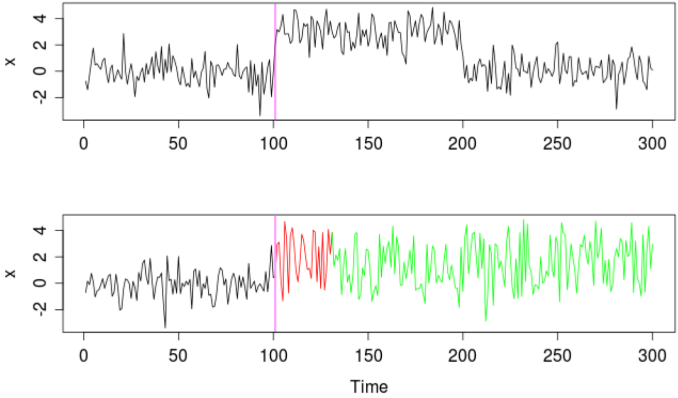
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

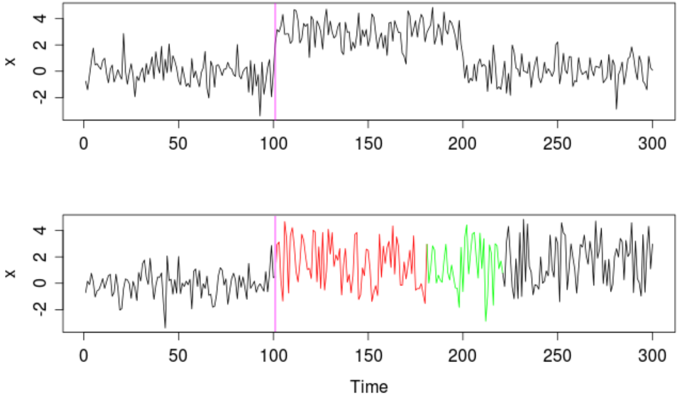
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

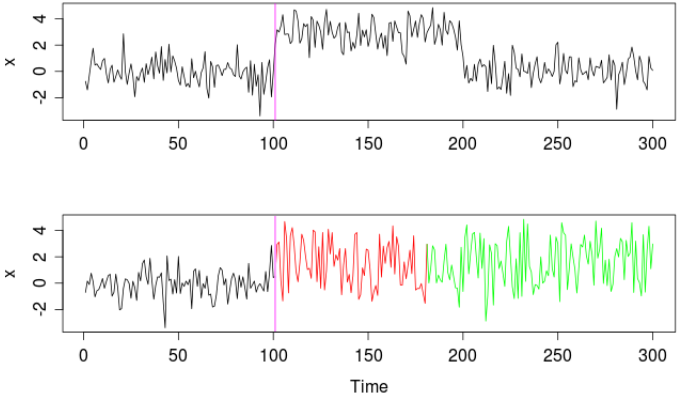
However, only permute within each cluster:



The E-Divisive Algorithm: Multiple Change Points

Once again, perform permutation test

However, only permute within each cluster:



The 'ecp' R package (CRAN)

Signature:

```
e.divisive(X, sig.lvl=0.05, R=199, k=NULL, min.size=30, alpha=1)
```

Arguments:

- `X` - A $T \times d$ matrix representation of a length T time series, with d-dimensional observations.
- `sig.lvl` - The significance level used for the permutation test.
- `R` - The maximum number of permutations to perform in the permutation test.
- `k` - The number of change points to return. If this is NULL only the statistically significant estimated change points are returned.
- `min.size` - The minimum number of observations btw change points.
- `alpha` - The index for test statistic.

The ‘ecp’ R package (CRAN)

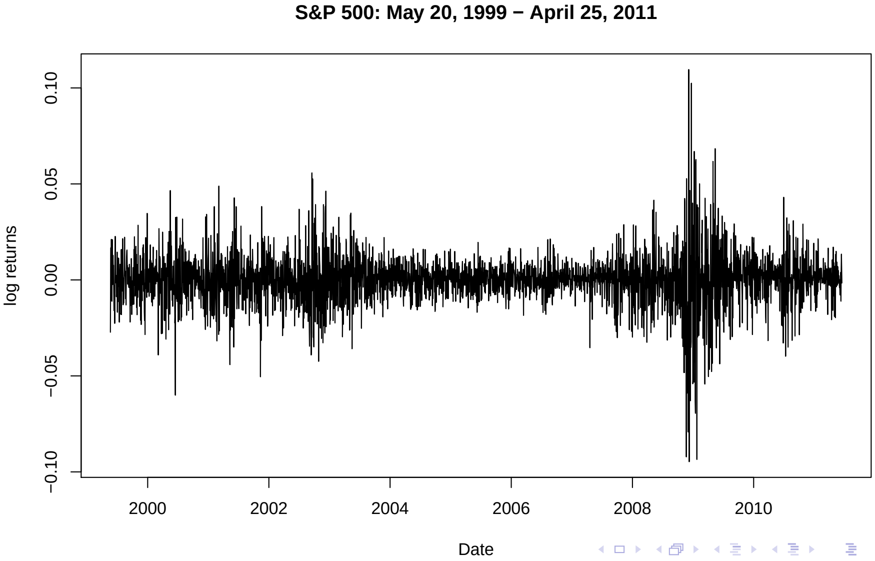
Returned list:

- `estimates` - The vector containing the estimated change point locations.
- `cluster.number` - The number of segments created by the estimated change points.
- `considered.last` - The location of the last estimated change point that was not deemed statistically significant.
- `order.found` - The order in which the change points were estimated.
- `Pvalues` - The approximate p-values returned by the permutation test.

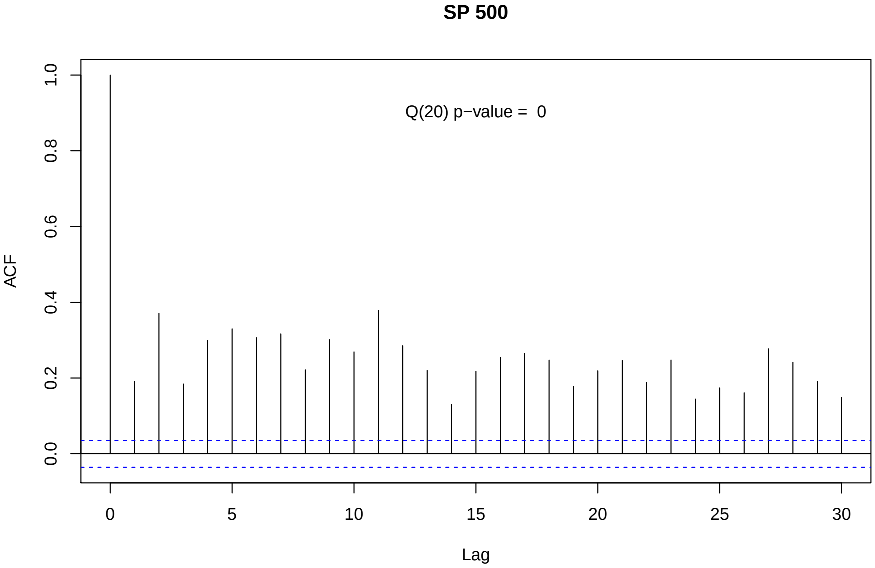
Complexity is $\mathcal{O}(kT^2)$

S&P 500 Index

The E-divisive procedure was applied to daily S&P 500 Index log returns

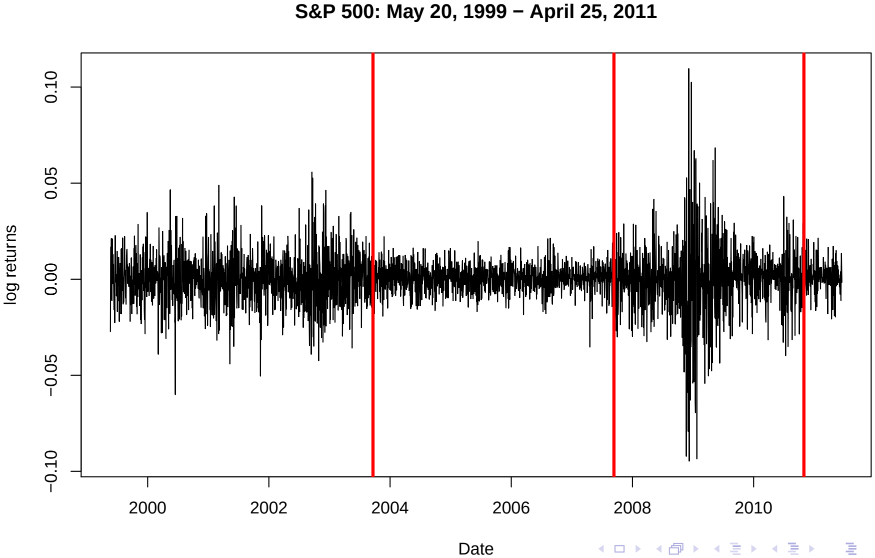


S&P 500 Index: Very significant ARCH effect overall

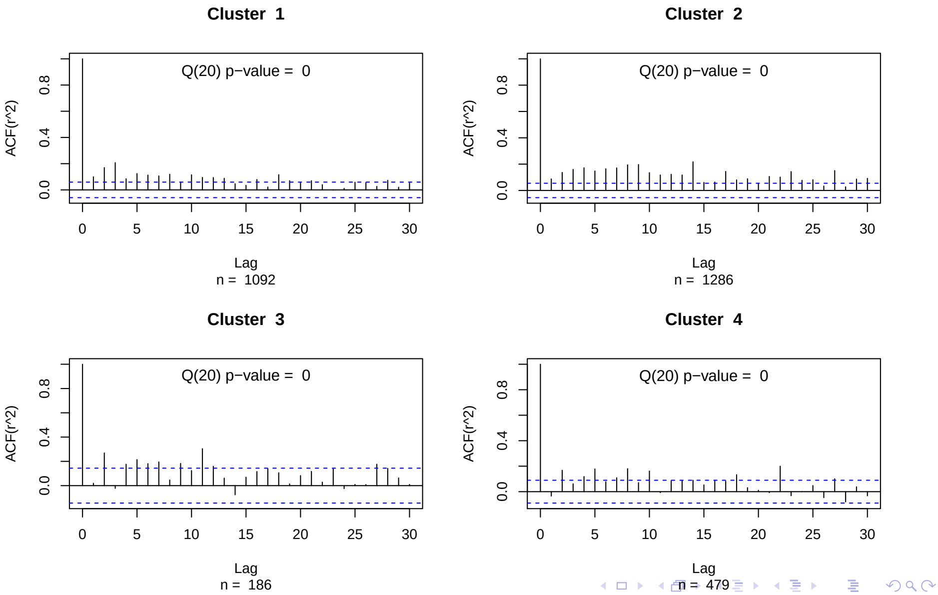


S&P 500 Index: Regime-Switching GARCH model

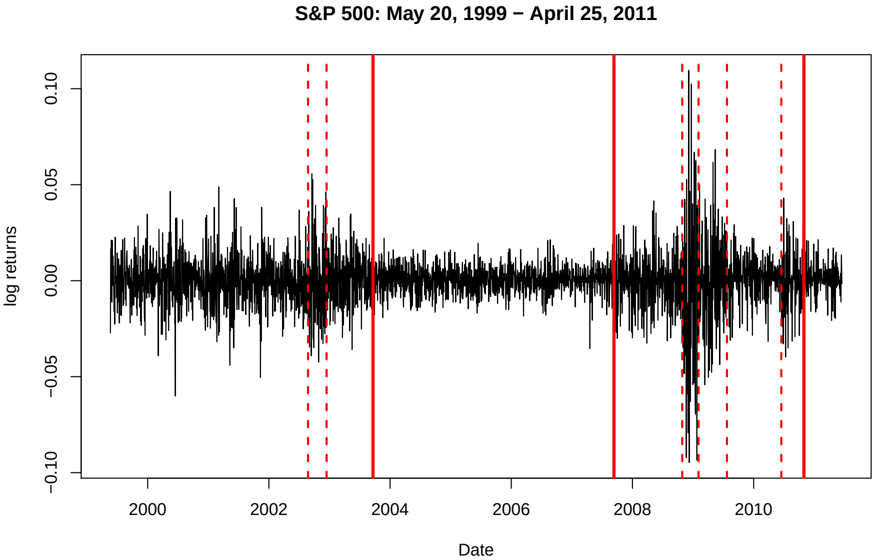
The E-divisive procedure was applied to daily S&P 500 Index log returns



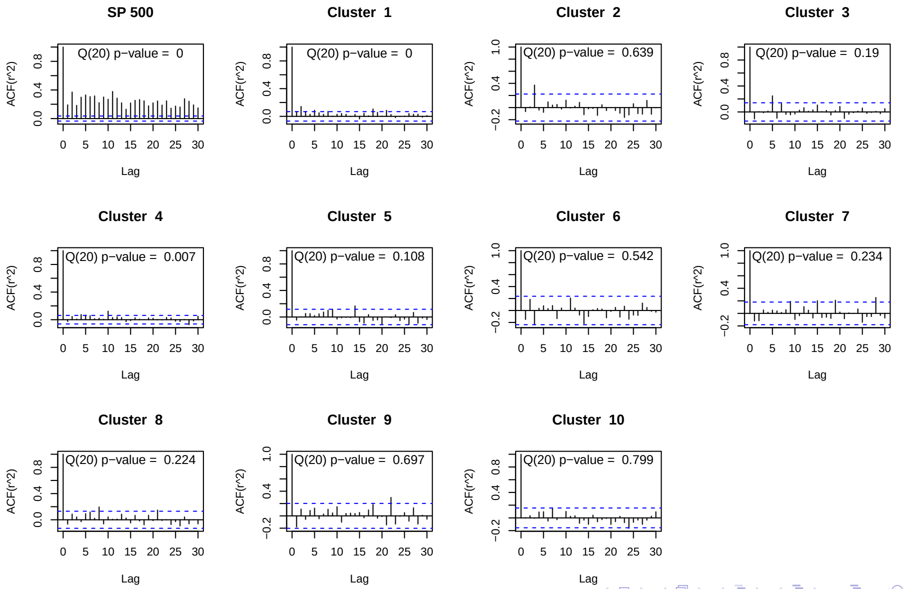
S&P 500 Index: Regime-switching GARCH model



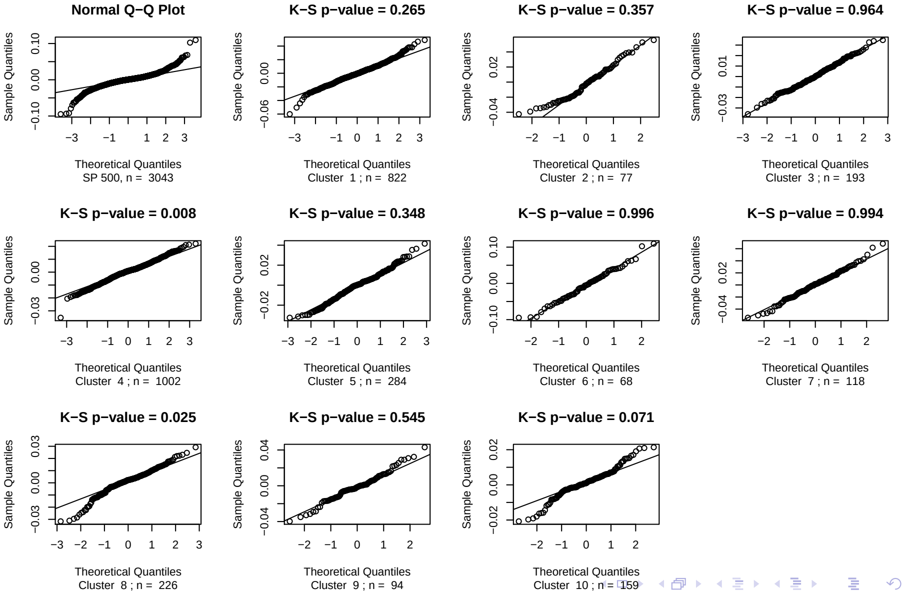
S&P 500 Index: Apply E-divisive Algorithm



S&P 500 Index: No ARCH effect within clusters!



S&P 500 Index: Gaussian distribution within clusters?



Multivariate Change in Correlation

1,000 simulations, 2 CP: $N_d(0,\Sigma)$, $N_d(0,I)$, $N_d(0,\Sigma)$

$$\Sigma_{w.o./noise} = \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix} \qquad \Sigma_{w/noise} = \begin{pmatrix} 1 & \rho & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

E-Divisive		Average Rand Index	
T	d	**Without Noise**	**With Noise**
300	2	0.767	0.774
	5	0.912	0.736
	9	0.970	0.736
600	2	0.817	0.836
	5	0.993	0.631
	9	0.998	0.666
900	2	0.970	0.968
	5	0.998	0.644
	9	0.999	0.612

Local Stationarity

A d-dimensional process $\{Y_t\}_{t=1}^{T}$ is strictly stationary if

$$F_{y_1,\ldots,y_k}(Y_1,\ldots,Y_k) = F_{y_{1+\tau},\ldots,y_{k+\tau}}(Y_{1+\tau},\ldots,Y_{k+\tau})$$

$\forall k, \tau \in \mathbb{N}$, in which F denotes a joint distribution function

Local Stationarity

A d-dimensional process $\{Y_t\}_{t=1}^{T}$ is strictly stationary if

$$F_{y_1,\ldots,y_k}(Y_1,\ldots,Y_k) = F_{y_{1+\tau},\ldots,y_{k+\tau}}(Y_{1+\tau},\ldots,Y_{k+\tau})$$

$\forall k, \tau \in \mathbb{N}$, in which F denotes a joint distribution function

---

An equivalent condition is that $\forall k, \tau \in \mathbb{N}$

$$\phi_{y_1,\ldots,y_k}(s) = \phi_{y_{1+\tau},\ldots,y_{k+\tau}}(s),\ \forall s \in \mathbb{R}^{d\times k}$$

in which ϕ denotes a joint characteristic function

Local Stationarity

A d-dimensional process $\{Y_t\}_{t=1}^{T}$ is strictly stationary if

$$F_{y_1,\ldots,y_k}(Y_1,\ldots,Y_k) = F_{y_{1+\tau},\ldots,y_{k+\tau}}(Y_{1+\tau},\ldots,Y_{k+\tau})$$

$\forall k, \tau \in \mathbb{N}$, in which F denotes a joint distribution function

---

An equivalent condition is that $\forall k, \tau \in \mathbb{N}$

$$\phi_{y_1,\ldots,y_k}(s) = \phi_{y_{1+\tau},\ldots,y_{k+\tau}}(s),\ \forall s \in \mathbb{R}^{d\times k}$$

in which ϕ denotes a joint characteristic function

Piecewise stationary:

$\forall t$, $\exists w_t \geq 0$, such that all observations in $[t - w_t, t]$ are stationarity process

$\underline{w_t}$ defines a one-sided window of homogeneity at t

Cointegration

Bivariate process $X_t = (x_{1,t}, x_{2,t})'$ is cointegrated if

1. Each component is I(1) (unit root nonstationary)
2. $\exists \beta \neq 0$ such that $\underline{x_{1,t} - \beta x_{2,t}}$ is I(0) (unit root stationary)

Cointegration

Bivariate process $X_t = (x_{1,t}, x_{2,t})'$ is cointegrated if

1. Each component is I(1) (unit root nonstationary)
2. $\exists \beta \neq 0$ such that $\underline{x_{1,t} - \beta x_{2,t}}$ is I(0) (unit root stationary)

Error Correction Model (ECM):

For bivariate I(1) process X_t, with cointegrating vector $\boldsymbol{\beta} = (1, -\beta)'$ one-lag ECM

$$\Delta X_t = \boldsymbol{\mu} + \boldsymbol{\alpha}\boldsymbol{\beta}' X_{t-1} + \boldsymbol{\Phi}\Delta X_{t-1} + \varepsilon_t$$

$\Delta X_t = X_t - X_{t-1}$; $\varepsilon_t \overset{\text{iid}}{\sim} (\mathbf{0}, \boldsymbol{\Sigma})$; and $\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\Phi}, \boldsymbol{\Sigma}$ are constant matrices

Local Cointegration

A bivariate process X_t is locally cointegrated with respect to a window of homogeneity w_t if $\forall t$

1. X_t is I(1), within the interval $[t - w_t, t]$ (unit root nonstationary)
2. $\exists \beta_t \neq 0$ such that $\underline{u_t = x_{1,t} - \beta_t x_{2,t}}$ is I(0) (unit root stationary)

Local Cointegration

A bivariate process X_t is locally cointegrated with respect to a window of homogeneity w_t if $\forall t$

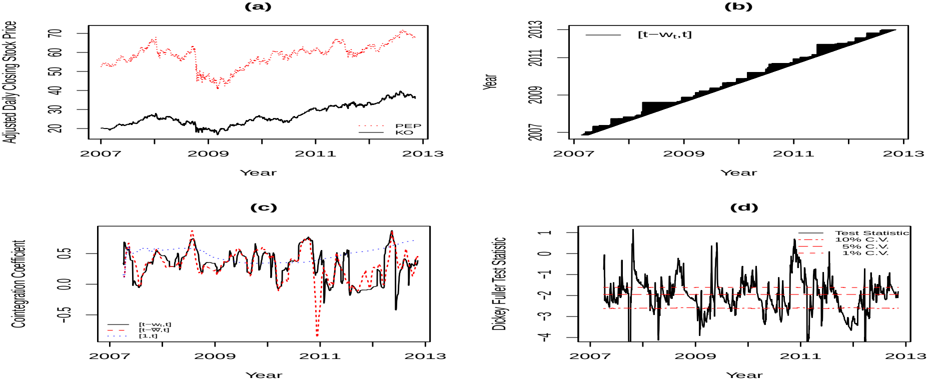
1. X_t is I(1), within the interval $[t - w_t, t]$ (unit root nonstationary)
2. $\exists \beta_t \neq 0$ such that $\underline{u_t = x_{1,t} - \beta_t x_{2,t}}$ is I(0) (unit root stationary)

At time t, find largest δ such that $(Z_t, u_t)'$ is stationary over $[t - \delta + 1, t]$

- $Z_t = \Delta X_t$
- $u_t = x_{1t} - \beta x_{2t}$
- β estimated via OLS over $[t - \delta + 1, t]$

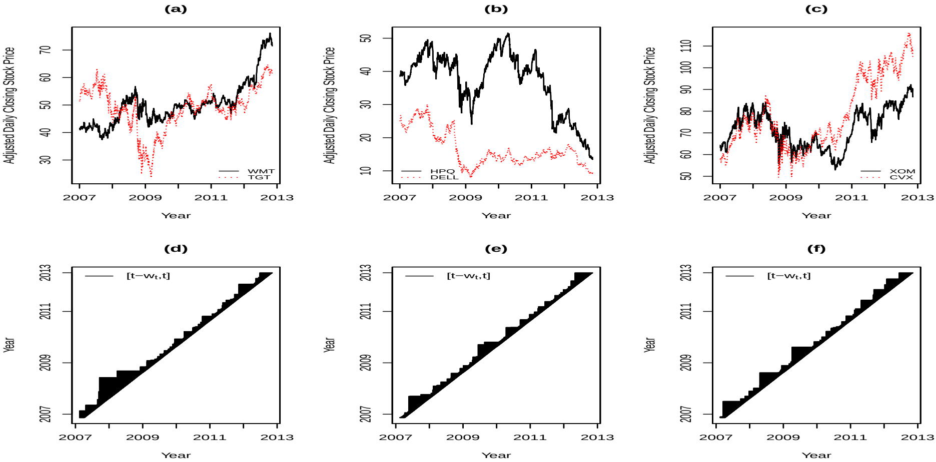
(a) Pepsi (PEP) and Coca-Cola (KO) adjusted daily closing stock prices January 2007 through November 2012

(b) $[t - w_t, t]$, estimated window of local stationarity at times t



(c) Cointegrating coefficient $\hat{\beta}(t, w_t)$ over $[t - w_t, t]$ (adaptive); $\hat{\beta}(t, \bar{w} = 68)$ over $[t - 68, t]$ (fixed); and $\hat{\beta}_t$, over $[1, t]$ (cumulative)

(d) Dickey-Fuller test statistic over $[t - w_t, t]$ (1, 5, & 10 % critical values) (abc) Adjusted daily closing prices from 1/2007 through 11/2012 of Walmart (WMT) & Target (TGT), Hewlett Packard (HPQ) & Dell (DELL), and Exxon Mobil (XOM) & Chevron (CVX)



(def) Estimated window of local stationarity $[t - w_t, t]$, at times t

Pairs Trading: Trading the Spread $\hat{u}_t = x_{1,t} - \hat{\beta}_t(\hat{w}_t) x_{2,t}$

ENTER at $|2\hat{\sigma}_t|$

EXIT at $|\frac{1}{2}\hat{\sigma}_t|$ (gain), $|3\hat{\sigma}_t|$ (loss), or $T = t + \delta_{min}$ (time limit)

Mean Return (per trade)

Window/Pair	KO-PEP	HPQ-DELL	WMT-TGT	XOM-CVX
Fixed	-7.0%	2.5%	-6.6%	0.9%
Cumulative	12.0%	-0.9%	-0.1%	-1.8%
Adaptive	21.0%	4.4%	1.1%	43.0%

Table: Mean return (per trade) for the three window methods on each pair

Mean Trade Duration (days)

Window/Pair	KO-PEP	HPQ-DELL	WMT-TGT	XOM-CVX
Fixed	7.8	9.7	9.7	8.5
Cumulative	16.9	23.0	13.4	17.8
Adaptive	7.8	8.2	8.3	10.6

Table: Mean trade duration (days) for the three window methods on each pair

Bibliography

http://www.stat.cornell.edu/~matteson/

- Matteson, D.S., and James, N.A. (2012+), "A Nonparametric Approach for Multiple Change Point Analysis of Multivariate Data."
- Matteson, D.S., James, N.A., Nicholson, W.B. and Segalini, L.C. (2012+), "Locally Stationary Vector Processes and Adaptive Multivariate Modeling."
- James, N.A. and Matteson, D.S. (2013), "ecp: An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data."
- James, N.A. and Matteson, D.S. (2013) http://cran.r-project.org/web/packages/ecp/