Bootstrap methods for time series

Part 3: Block Bootstrap Methods

Bartosz Majewski

AGH University of Krakow,
Faculty of Applied Mathematics

bartmaje@agh.edu.pl

Syllabus

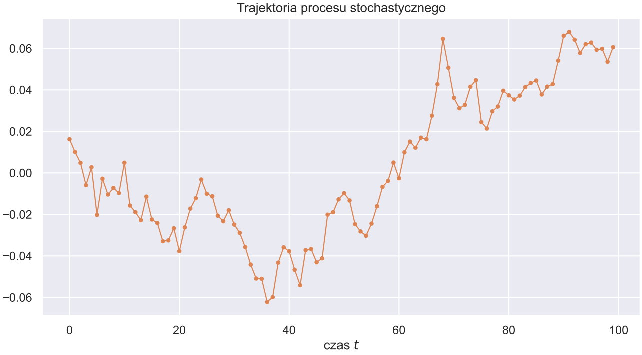
1. Introduction to Bootstrap
2. Bootstrap Confidence Interval
3. **Block Bootstrap Methods**
4. Bootstrap for Data with Periodic Structure

Block Bootstrap Methods

3.1. Motivation

Motivation

If the data is dependent (e.g., time series, spatial data), the random resampling data points destroys the existing correlation structure, leading to incorrect statistical inferences.



Motivation

Block bootstrap methods are designed to handle dependent data, such as time series and spatial data, where observations are correlated.

Motivation

Block bootstrap methods are designed to handle dependent data, such as time series and spatial data, where observations are correlated.

Instead of resampling individual data points, we **resample blocks of observations to preserve the dependence structure.**

Motivation

Block bootstrap methods are designed to handle dependent data, such as time series and spatial data, where observations are correlated.

Instead of resampling individual data points, we **resample blocks of observations to preserve the dependence structure.**

Some block bootstrap methods for stationary time series:

- Moving Block Bootstrap (MBB),
- Circular Block Bootstrap (CBB),
- Nonoverlapping Block Bootstrap (NBB),
- Stationary Bootstrap (SB).

We will discuss the first two methods.

Stationary time series

Let $\{X_t, t \in \mathbb{N}\}$ be a stationary time series. That is, a mean and an autocovariance functions are constant is time

$$\mathbb{E}X_t = \mu$$

and

$$\mathrm{Cov}(X_t, X_{t+h}) = \mathrm{Cov}(X_1, X_{1+h}) = \gamma(h).$$

3.2. Moving Block Bootstrap

MBB

Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,

MBB

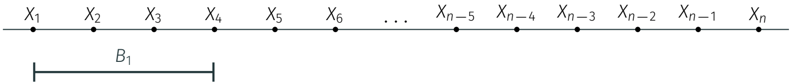
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

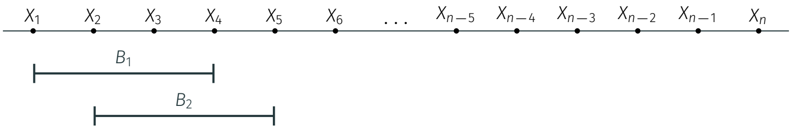
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\mathbf{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

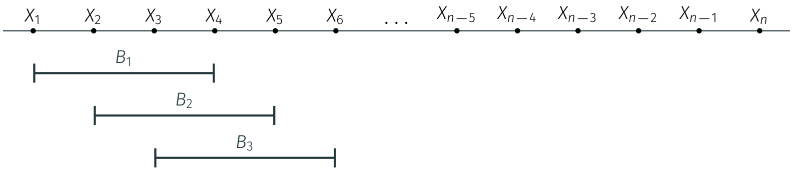
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

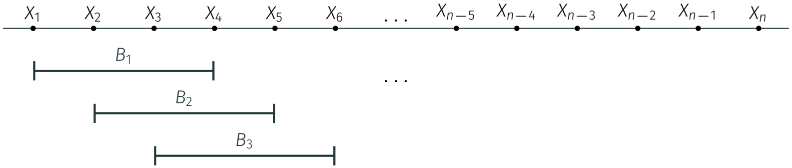
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

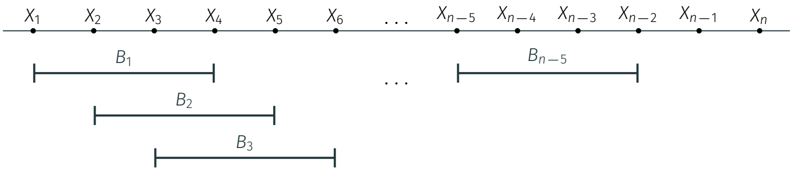
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

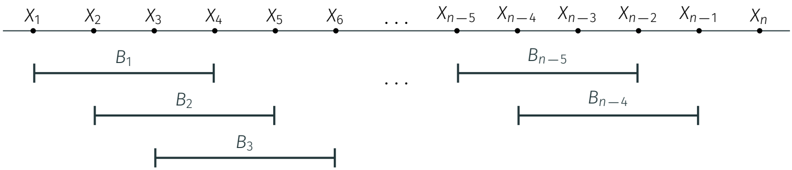
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \ldots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



MBB

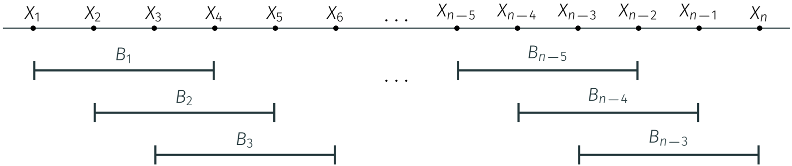
Let $\{X_t, t \in \mathbb{N}\}$ be a time series and $\boldsymbol{X}_n = (X_1, \dots, X_n)$ be a observation from this time series.

Fix an integer number $b < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = (X_1, \dots, X_{i+b-1}), \qquad 1 \leq i \leq n - b + 1,$$

That is,



Algorithm

1. Fix $b < n$.
2. From the set $B = \{B_1, \ldots, B_{n-b+1}\}$, we randomly select, with replacement, l blocks: $B_1^*, \ldots, B_l^*$. The probability of selecting any given block is $\frac{1}{n-b+1}$.
3. We concatenate the selected blocks and take the first n observations to obtain the bootstrap sample $(X_1^*, \ldots, X_n^*)$.

$$X^* = (\underbrace{X_{k_1}, X_{k_1+1}, \ldots, X_{k_1+b-1}}_{B_1^*}, \underbrace{X_{k_2}, X_{k_2+1}, \ldots, X_{k_2+b-1}}_{B_2^*}, \ldots, \underbrace{X_{k_l}, X_{k_l+1}, X_{k_l+b-1}}_{B_l^*})$$

Bias in the MBB method

The observations in the center of the considered sample are present in b different blocks, while **the observations from the beginning and the end of the sample appear more rarely.**

For example, the first observation X_1 is present only in block B_1. This fact **leads to an increase in the bias of the estimator.**

To reduce the bias, Politis and Romano in 1998 proposed **treating the data as wrapped in the circle**. Then each observation is present in the same number of blocks.

3.3. Circular Block Bootstrap

CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$

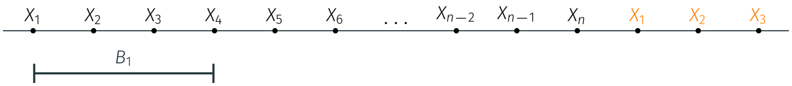
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



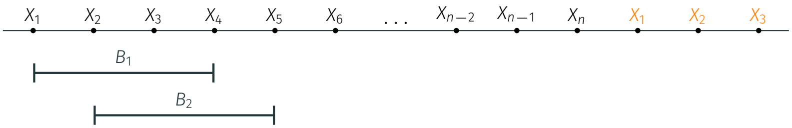
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



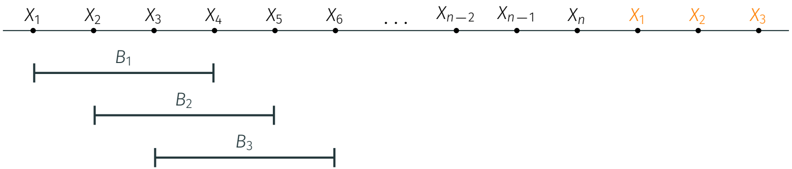
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



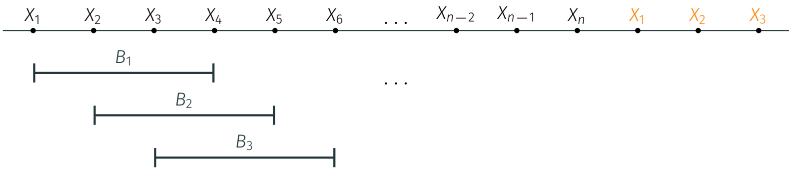
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



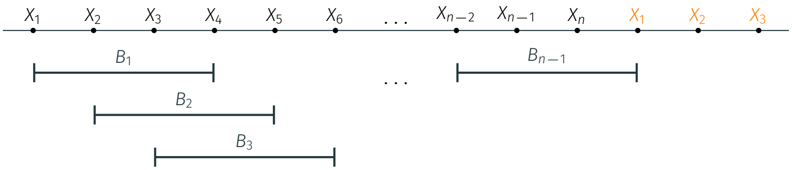
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



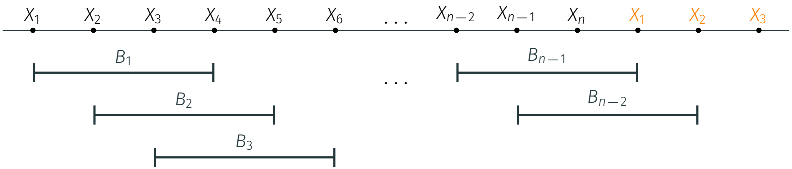
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



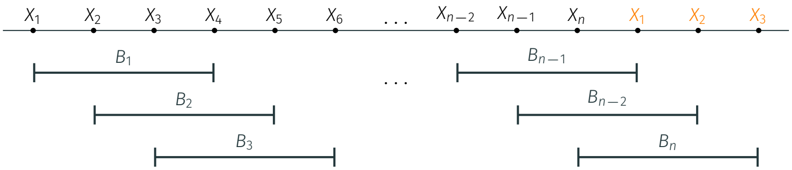
CBB

Let $X_1, X_2, \ldots$ be a time series and $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a observation from this time series.

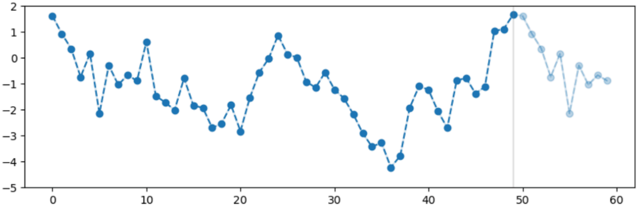
Fix an integer number$b = b_n < n$.

Define a block B_i of size b, starting with X_i:

$$B_i = \begin{cases} (X_1, \ldots, X_{i+b-1}), & i = 1, \ldots, n-b+1, \\ (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1}), & i = n-b+2, \ldots, n. \end{cases}$$



CBB



Algorithm

1. Fix $b < n$.
2. From the set $B = \{B_1, \ldots, B_n\}$, we randomly select, with replacement, l blocks: $B_1^*, \ldots, B_l^*$. The probability of selecting any given block is $\frac{1}{n}$.
3. We concatenate the selected blocks and take the first n observations to obtain the bootstrap sample $(X_1^*, \ldots, X_n^*)$.

$$X^* = (\underbrace{X_{k_1}, X_{k_1+1}, \ldots, X_{k_1+b-1}}_{B_1^*}, \underbrace{X_{k_2}, X_{k_2+1}, \ldots, X_{k_2+b-1}}_{B_2^*}, \ldots, \underbrace{X_{k_l}, X_{k_l+1}, X_{k_l+b-1}}_{B_l^*})$$

3.4. few words about NBB and SB methods

NBB and SB

- **Non-overlapping Block Bootstrap (NBB):** Similar to MBB but uses non-overlapping blocks, meaning each block is drawn independently without overlapping.
- **Stationary Bootstrap (SB):** Uses random-length blocks where block lengths follow a geometric distribution. It preserves the stationary distribution of the original time series.

3.5. Optimal block length

Optimal block length

The optimal choice of the block length b is a difficult problem. By 'optimal' we mean one that minimizes the mean squared error of the bootstrap estimator's variance.

In the case of estimating the mean and variance of a stationary time series, we can show that for MBB and CBB the optimal block length b satisfies

$$b_{opt} = Cn^{\frac{1}{3}},$$

where for $C > 0$ the formula can be determined[1].

[1] See Lahiri, S.N. (2003). Resampling Methods for Dependent Data. Springer. New York.

Literature

Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*. Springer. New York.

Course materials

Go to my website (https://home.agh.edu.pl/∼ bartmaje/) to the section Teaching.

You can use the QR code below:

Thank you!