Wavelet-Based Bootstrapping for Non-Gaussian Time Series

Don Percival

Applied Physics Laboratory
Department of Statistics
University of Washington
Seattle, Washington, USA

overheads for talk available at

http://faculty.washington.edu/dbp/talks.html
Overview

• question of interest: how can we assess the sampling variability in statistics computed from a time series $X_0, X_1, \ldots, X_{N-1}$?

• start with some background on bootstrapping

• review parametric and block bootstrapping (two approaching for handling correlated time series)

• review previously proposed wavelet-based approach to bootstrapping (Percival, Sardy and Davison, 2001)

• describe a new wavelet-based approach that uses ‘trees’ for resampling and is potentially useful for non-Gaussian time series

• demonstrate methodology on time series related to BMW stock

• conclude with some remarks
Motivating Question

• let $\mathbf{X} = [X_0, \ldots, X_{N-1}]^T$ be a portion of a stationary process with autocorrelation sequence (ACS)

$$
\rho_\tau = \frac{s_\tau}{s_0}, \text{ where } s_\tau = \text{cov} \{X_t, X_{t+\tau}\} \text{ and } s_0 = \text{var} \{X_t\}
$$

• given a time series, we can estimate its ACS at $\tau = 1$ using

$$
\hat{\rho}_1 = \frac{\sum_{t=0}^{N-2}(X_t - \overline{X})(X_{t+1} - \overline{X})}{\sum_{t=0}^{N-1}(X_t - \overline{X})^2}, \text{ where } \overline{X} = \frac{1}{N} \sum_{t=0}^{N-1} X_t
$$

• Q: given the amount of data $N$ we have, how close can we expect $\hat{\rho}_1$ to be to the true unknown $\rho_1$?

• i.e., how can we assess the sampling variability in $\hat{\rho}_1$?
Classic Approach – Large Sample Theory

• let $N(\mu, \sigma^2)$ denote a Gaussian (normal) random variable (RV) with mean $\mu$ and variance $\sigma^2$

• under suitable conditions (see, e.g., Fuller, 1996), $\hat{\rho}_1$ has a distribution close to that of $N(\rho_1, \sigma^2_N)$ as $N \to \infty$, where

$$
\sigma^2_N = \frac{1}{N} \sum_{\tau=-\infty}^{\infty} \left\{ \rho^2_\tau (1 + 2\rho^2_1) + \rho_{\tau+1}\rho_{\tau-1} - 4\rho_1\rho_\tau\rho_{\tau-1} \right\}
$$

• in practice, this result is unappealing because it requires
  – knowledge of theoretical ACS (including the unknown $\rho_1$!)
  – ACS to damp down fast, ruling out some processes of interest

• while large sample theory has been worked out for $\hat{\rho}_1$ under certain conditions, similar theory for other statistics can be hard to come by
Alternative Approach – Bootstrapping: I

- if \( X_t \)'s were IID, we could apply ‘bootstrapping’ to assess the variability in \( \hat{\rho}_1 \), as follows
- consider a time series of length \( N = 8 \) that is a realization of a Gaussian white noise process \( (\rho_1 = 0) \):

\[
\hat{\rho}_1 \doteq 0.18
\]
Alternative Approach – Bootstrapping: I

• if \( X_t \)'s were IID, we could apply ‘bootstrapping’ to assess the variability in \( \hat{\rho}_1 \), as follows

• consider a time series of length \( N = 8 \) that is a realization of a Gaussian white noise process (\( \rho_1 = 0 \)):

\[
\hat{\rho}_1 \doteq 0.18
\]

• generate new series by randomly sampling with replacement:
Alternative Approach – Bootstrapping: I

• if \( X_t \)'s were IID, we could apply ‘bootstrapping’ to assess the variability in \( \hat{\rho}_1 \), as follows

• consider a time series of length \( N = 8 \) that is a realization of a Gaussian white noise process \( (\rho_1 = 0) \):

\[
J \rho_S b_{\mathbf{R}} \hat{\rho}_S b_{\mathbf{R}}
\]

\( \hat{\rho}_1 \approx 0.18 \)

• generate new series by randomly sampling with replacement:
Alternative Approach – Bootstrapping: I

• if $X_t$’s were IID, we could apply ‘bootstrapping’ to assess the variability in $\hat{\rho}_1$, as follows

• consider a time series of length $N = 8$ that is a realization of a Gaussian white noise process ($\rho_1 = 0$):

• generate new series by randomly sampling with replacement:

\[ \hat{\rho}_1 \doteq 0.18 \]
Alternative Approach – Bootstrapping: I

• if $X_t$’s were IID, we could apply ‘bootstrapping’ to assess the variability in $\hat{\rho}_1$, as follows

• consider a time series of length $N = 8$ that is a realization of a Gaussian white noise process ($\rho_1 = 0$):

  ![Graph 1](image)

  $\hat{\rho}_1 \doteq 0.18$

• generate new series by randomly sampling with replacement:

  ![Graph 2](image)
Alternative Approach – Bootstrapping: I

• if $X_t$’s were IID, we could apply ‘bootstrapping’ to assess the variability in $\hat{\rho}_1$, as follows

• consider a time series of length $N = 8$ that is a realization of a Gaussian white noise process ($\rho_1 = 0$):

\[
\begin{align*}
\hat{\rho}_1 \doteq 0.18
\end{align*}
\]

• generate new series by randomly sampling with replacement:
Alternative Approach – Bootstrapping: I

- if \( X_t \)'s were IID, we could apply ‘bootstrapping’ to assess the variability in \( \hat{\rho}_1 \), as follows
- consider a time series of length \( N = 8 \) that is a realization of a Gaussian white noise process (\( \rho_1 = 0 \)):

\[
\hat{\rho}_1 \doteq 0.18
\]

- generate new series by randomly sampling with replacement:
Alternative Approach – Bootstrapping: I

- If $X_t$’s were IID, we could apply ‘bootstrapping’ to assess the variability in $\hat{\rho}_1$, as follows

- Consider a time series of length $N = 8$ that is a realization of a Gaussian white noise process ($\rho_1 = 0$):

  $\hat{\rho}_1 \approx 0.18$

- Generate new series by randomly sampling with replacement:
Alternative Approach – Bootstrapping: I

- If $X_t$’s were IID, we could apply ‘bootstrapping’ to assess the variability in $\hat{\rho}_1$, as follows

- Consider a time series of length $N = 8$ that is a realization of a Gaussian white noise process ($\rho_1 = 0$):

  \[
  \hat{\rho}_1 \doteq 0.18
  \]

- Generate new series by randomly sampling with replacement:

  \[
  \hat{\rho}_1^{(1)} \doteq -0.24
  \]
Alternative Approach – Bootstrapping: II

• repeat a large number of times $M$ to get $\hat{\rho}_1^{(1)}, \hat{\rho}_1^{(2)}, \ldots, \hat{\rho}_1^{(M)}$

• plots shows estimated probability density function (PDF) for $\hat{\rho}_1^{(1)}, \hat{\rho}_1^{(2)}, \ldots, \hat{\rho}_1^{(100)}$, along with (a) PDF for $\mathcal{N}(0, \frac{1}{8})$ and (b) approximation to the true PDF for $\hat{\rho}_1$

• can regard sample distribution of $\{\hat{\rho}_1^{(m)}\}$ as an approximation to the unknown distribution of $\hat{\rho}_1$


Alternative Approach – Bootstrapping: III

- quality of approximation depends upon particular time series
- here are bootstrap approximations to PDF of $\hat{\rho}_1$ based upon two other time series of length $N = 8$, along with true PDF

![Graph showing bootstrap approximations and true PDF](image)

- repeating the above for 50 time series yields 50 bootstrap PDFs
- summarize via sample means and standard deviations (SDs):
  
  average of 50 sample means $\bar{x} = -0.127$ (truth $\approx -0.124$)
  average of 50 sample SDs $SD = 0.280$ (truth $\approx 0.284$)
Bootstrapping Correlated Time Series: I

• key assumption: \( X \) contains IID RVs

• if not true (as for most time series!), sample distribution of \( \{\hat{\rho}_1^{(m)}\} \) can be a poor approximation to distribution of \( \hat{\rho}_1 \)

• as an example, consider first order autoregressive (AR) process:

\[
X_t = \phi X_{t-1} + \epsilon_t,
\]

where \( \phi = 0.9 \) and \( \{\epsilon_t\} \) is zero mean Gaussian white noise

• AR time series of length \( N = 128 \) with sample and true ACSs:
Bootstrapping Correlated Time Series: II

- use same procedure as before to get $\hat{\rho}_1^{(1)}, \hat{\rho}_1^{(2)}, \ldots, \hat{\rho}_1^{(100)}$
- bootstrap approximation to PDF of $\hat{\rho}_1$ along with true PDF:

  ![Graph](image)

  vertical line indicates $\hat{\rho}_1$

- bootstrap approximation gets even worse as $N$ increases
- to correct the problem caused by correlation in time series, can use specialized time- or frequency-domain bootstrapping (assuming true ACS damps downs sufficiently fast)
Parametric Bootstrapping: I

- one well-known time-domain bootstrapping scheme is the parametric (or residual) bootstrap
- suppose $X$ is a realization of AR process $X_t = \phi X_{t-1} + \epsilon_t$
- note that $\text{var} \{ X_t \} = \text{var} \{ \epsilon_t \} / (1 - \phi^2)$ and $\rho_\tau = \phi^{\lfloor \tau \rfloor}$
- in particular, $\rho_1 = \phi$, so can estimate $\phi$ using $\hat{\phi} = \hat{\rho}_1$
- since $\epsilon_t = X_t - \phi X_{t-1}$, can form residuals
  \[
  r_t = X_t - \hat{\phi} X_{t-1}, \quad t = 1, \ldots, N - 1,
  \]
  with the idea that $r_t$ will be a good approximation to $\epsilon_t$
- let $r_0^{(1)}, r_1^{(1)}, \ldots, r_{N-1}^{(1)}$ be a random sample from $r_1, r_2, \ldots, r_{N-1}$
- let $X_0^{(1)} = r_0^{(1)} / (1 - \hat{\phi}^2)^{1/2}$ (‘stationary initial condition’)
Parametric Bootstrapping: II

• form

\[ X^{(1)}_t = \hat{\phi} X^{(1)}_{t-1} + r^{(1)}_t, \quad t = 1, \ldots, N - 1, \]

yielding the bootstrapped time series \( X_0^{(1)}, X_1^{(1)}, \ldots, X_{N-1}^{(1)} \)

• AR time series (left-hand plot) and bootstrapped series (right):

\[ \begin{array}{c}
\text{X} \\
-1 \quad 0 \quad 1
\end{array} \quad \begin{array}{c}
\text{X}^{(1)} \\
-1 \quad 0 \quad 1
\end{array} \]

\[ \begin{array}{c}
0 \quad 32 \quad 64 \quad 96 \quad 128
\end{array} \quad \begin{array}{c}
0 \quad 32 \quad 64 \quad 96 \quad 128
\end{array} \]

• use bootstrapped series to compute \( \hat{\rho}^{(1)}_1 \)

• repeat this procedure \( M \) times to get \( \hat{\rho}^{(1)}_1, \hat{\rho}^{(2)}_1, \ldots, \hat{\rho}^{(M)}_1 \)
Parametric Bootstrapping: III

- bootstrap approximation to PDF of $\hat{\rho}_1$ along with true PDF:

- repeating the above for 50 AR time series yields:
  
  average of 50 sample means $\bar{\hat{\rho}} = 0.83$  (truth $\bar{\rho} = 0.86$)
  
  average of 50 sample SDs $\bar{\hat{\rho}} = 0.053$  (truth $\bar{\rho} = 0.048$)
Parametric Bootstrapping: IV

- important assumption: $X$ generated by AR process
- to see what happens if assumption fails, consider a fractionally differenced (FD) process

\[ X_t = \sum_{k=0}^{\infty} \frac{\Gamma(1 - \delta)}{\Gamma(k + 1)\Gamma(1 - \delta - k)} \epsilon_{t-k}, \]

where $\delta = 0.45$ and $\{\epsilon_t\}$ is zero mean Gaussian white noise

- FD time series of length $N = 128$ with sample and true ACSs:
Parametric Bootstrapping: V

- AR process has ‘short-range’ dependence, whereas FD process exhibits ‘long-range’ (or ‘long-memory’) dependence
- bootstrap approximation to PDF of $\hat{\rho}_1$ along with true PDF:

  ![Graph showing bootstrap approximation and true PDF with vertical line indicating $\hat{\rho}_1$]

- repeating the above for 50 FD time series yields:
  
  average of 50 sample means $\hat{\rho}_1 = 0.49$ (truth $= 0.53$)
  
  average of 50 sample SDs $\hat{\rho}_1 = 0.078$ (truth $= 0.107$)
  
  note: $\rho_1 \approx 0.82$ for this FD process; agreement in SD gets worse (better) as $N$ increases (decreases)
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[
\hat{\rho}_1 \doteq 0.78
\]

- generate bootstrapped AR series by randomly sampling blocks:
**Block Bootstrapping: I**

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 = 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

• another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
• break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

• generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

  \[ \hat{\rho}_1 \approx 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into \( B \) blocks (subseries) of equal length:

\[
\hat{\rho}_1 \doteq 0.78
\]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[
\hat{\rho}_1 \doteq 0.78
\]

- generate bootstrapped AR series by randomly sampling blocks:
another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[
\hat{\rho}_1 \geq 0.78
\]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
**Block Bootstrapping:**

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest.
- Break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \doteq 0.78 \]

- Generate bootstrapped AR series by randomly sampling blocks:

14
**Block Bootstrapping: I**

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

  \[ \hat{\rho}_1 \approx 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest
- break time series up into $B$ blocks (subseries) of equal length:

\[ \hat{\rho}_1 \approx 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
Block Bootstrapping: I

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest:
- break time series up into $B$ blocks (subseries) of equal length:

  \[ \hat{\rho}_1 \approx 0.78 \]

- generate bootstrapped AR series by randomly sampling blocks:
**Block Bootstrapping: I**

- another time-domain approach is block bootstrapping, which has many variations, of which the following is the simplest

- break time series up into $B$ blocks (subseries) of equal length:

$$
\hat{\rho}_1 \doteq 0.78
$$

- generate bootstrapped AR series by randomly sampling blocks:

$$
\hat{\rho}_1^{(1)} \doteq 0.63
$$
Block Bootstrapping: II

- bootstrap approximation to PDF of $\hat{\rho}_1$ along with true PDF:

- repeating the above for 50 AR time series yields:
  - average of 50 sample means $\hat{=}$ 0.75 (truth $\hat{=}$ 0.86)
  - average of 50 sample SDs $\hat{=}$ 0.049 (truth $\hat{=}$ 0.048)

- repeating the above for 50 FD time series yields:
  - average of 50 sample means $\hat{=}$ 0.46 (truth $\hat{=}$ 0.53)
  - average of 50 sample SDs $\hat{=}$ 0.082 (truth $\hat{=}$ 0.107)
Frequency-Domain Bootstrapping

- again, many variations, including the following three
- ‘phase scramble’ discrete Fourier transform (DFT)
  \[ x_k = \sum_{t=0}^{N-1} x_t e^{-i2\pi kt/N} = A_k e^{i\theta_k} \]

  of \( X \) and apply inverse DFT to create new series
- periodogram-based bootstrapping: in addition to phase scrambling, evoke large sample result that \(|A_k|’s\) are approximately uncorrelated with distribution related to a chi-square RV with 2 degrees of freedom
- circulant embedding bootstrapping: form nonparametric estimate of spectral density function and generate realizations using circulant embedding (Percival and Constantine, 2006)
Critique of Time/Frequency-Domain Bootstrapping

- time- and frequency-domain approaches are mainly designed for series with short-range dependence (e.g., AR) and are problematic for those exhibiting long-range dependence (e.g., FD)
- parametric and frequency-domain bootstraps work best for series that obey a Gaussian distribution, but can be problematic for non-Gaussian series
- non-Gaussian series better handled by block bootstrapping, but quality of this approach depends critically on chosen size for blocks (ad hoc rule is to set size close to $\sqrt{N}$)
- room for improvement: will consider wavelet-based approaches
Overview of Discrete Wavelet Transform (DWT): I

- DWT is an orthonormal transform $\mathcal{W}$ that reexpresses a time series $X$ of length $N$ as a vector of DWT coefficients $W$:

$$W = \mathcal{W}X,$$

where $\mathcal{W}$ is an $N \times N$ matrix such that $X = \mathcal{W}^T W$

- particular $\mathcal{W}$ depends on the choice of
  - wavelet filter, the most basic of which is the Haar filter (fancier filters include the Daubechies family of ‘least asymmetric’ filters of width $L$ – denoted by $\text{LA}(L)$, with $L = 8$ being a popular choice)
  - level $J_0$, which determines the number of dyadic scales $\tau_j = 2^{j-1}, j = 1, 2, \ldots, J_0$, involved in the transform
Overview of Discrete Wavelet Transform (DWT): II

- DWT coefficient vector $\mathbf{W}$ can be partitioned into $J_0$ sub-vectors of wavelet coefficients $\mathbf{W}_j$, $j = 1, 2, \ldots, J_0$, along with one sub-vector of scaling coefficients $\mathbf{V}_{J_0}$.

- Wavelet coefficients in $\mathbf{W}_j$ are associated with changes in averages over a scale of $\tau_j$, whereas the scaling coefficients in $\mathbf{V}_{J_0}$ are associated with averages over a scale of $2\tau_{J_0}$.

- As a concrete example, let’s look at a level $J_0 = 4$ Haar DWT of the AR time series.
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$
DWT of Autoregressive Process: I

- level \( J_0 = 4 \) Haar DWT of AR series \( X \), with scale \( \tau_1 = 1 \) wavelet coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $\tau_1 = 1$ wavelet coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $\tau_2 = 2$
  wavelet coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $\tau_2 = 2$ wavelet coefficient highlighted
level $J_0 = 4$ Haar DWT of AR series $X$, with scale $\tau_4 = 8$ wavelet coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $\tau_4 = 8$ wavelet coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \times \tau_4 = 16$ scaling coefficient highlighted
level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \times \tau_4 = 16$ scaling coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \times \tau_4 = 16$
- scaling coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \times \tau_4 = 16$ scaling coefficient highlighted
level $J_0 = 4$ Haar DWT of AR series $\mathbf{X}$, with scale $2 \times \tau_4 = 16$ scaling coefficient highlighted
DWT of Autoregressive Process: I

- level \( J_0 = 4 \) Haar DWT of AR series \( \mathbf{X} \), with scale \( 2 \times \tau_4 = 16 \) scaling coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \times \tau_4 = 16$ scaling coefficient highlighted
DWT of Autoregressive Process: I

- level $J_0 = 4$ Haar DWT of AR series $X$, with scale $2 \tau_4 = 16$ scaling coefficient highlighted
Haar DWT of AR series $\mathbf{X}$ and sample ACSs for each $\mathbf{W}_j$ & $\mathbf{V}_4$, along with 95% confidence intervals for white noise
Haar DWT of FD series $\mathbf{X}$ and sample ACSs for each $\mathbf{W}_j$ & $\mathbf{V}_4$, along with 95% confidence intervals for white noise
DWT as a Decorrelating Transform

• for many (but not all!) time series, DWT acts as a decorrelating transform: to a good approximation, each $W_j$ is a sample of a white noise process, and coefficients from different sub-vectors $W_j$ and $W_{j'}$ are also pairwise uncorrelated

• variance of coefficients in $W_j$ depends on $j$

• scaling coefficients $V_{J_0}$ are still autocorrelated, but there will be just a few of them if $J_0$ is selected to be large

• decorrelating property holds particularly well for FD and other processes with long-range dependence

• above suggests the following recipe for wavelet-domain bootstrapping
Recipe for Wavelet-Domain Bootstrapping

1. given $\mathbf{X}$ of length $N = 2^J$, compute level $J_0$ DWT (the choice $J_0 = J - 3$ yields 8 coefficients in $\mathbf{W}_{J_0}$ and $\mathbf{V}_{J_0}$)

2. randomly sample with replacement from $\mathbf{W}_j$ to create bootstrapped vector $\mathbf{W}^{(b)}_j$, $j = 1, \ldots, J_0$

3. create $\mathbf{V}^{(b)}_{J_0}$ using a parametric bootstrap

4. apply $\mathbf{W}^T$ to $\mathbf{W}^{(b)}_1$, \ldots, $\mathbf{W}^{(b)}_{J_0}$ and $\mathbf{V}^{(b)}_{J_0}$ to obtain bootstrapped time series $\mathbf{X}^{(b)}$ and then form corresponding $\hat{\rho}_1^{(b)}$

• repeat above many times to build up sample distribution of bootstrapped autocorrelations
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series X (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series X (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)

25
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series \(X\) (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Illustration of Wavelet-Domain Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and wavelet-domain bootstrap thereof (right-hand)
Wavelet-Domain Bootstrapping of AR Series

• approximations to true PDF using (a) Haar & (b) LA(8) wavelets

• using 50 AR time series and the Haar DWT yields:
  
  average of 50 sample means $\hat{=}$ 0.67 (truth $\doteq$ 0.86)
  
  average of 50 sample SDs $\hat{=}$ 0.071 (truth $\doteq$ 0.048)

• using 50 AR time series and the LA(8) DWT yields:
  
  average of 50 sample means $\hat{=}$ 0.80 (truth $\doteq$ 0.86)
  
  average of 50 sample SDs $\hat{=}$ 0.055 (truth $\doteq$ 0.048)
Wavelet-Domain Bootstrapping of FD Series

- approximations to true PDF using (a) Haar & (b) LA(8) wavelets

- using 50 FD time series and the Haar DWT yields:
  average of 50 sample means $\hat{\mu} = 0.35$ (truth $\hat{\mu} = 0.53$)
  average of 50 sample SDs $\hat{\sigma} = 0.096$ (truth $\hat{\sigma} = 0.107$)

- using 50 FD time series and the LA(8) DWT yields:
  average of 50 sample means $\hat{\mu} = 0.43$ (truth $\hat{\mu} = 0.53$)
  average of 50 sample SDs $\hat{\sigma} = 0.098$ (truth $\hat{\sigma} = 0.107$)
Effect of Non-Gaussianity: I

- wavelet-domain bootstrapping works well if we can assume Gaussianity, but can be problematic if this assumption fails.
- for non-Gaussian series, wavelet-domain bootstraps are typically closer to Gaussianity than original series, which poses a problem for assessing variability in certain statistics.
Effect of Non-Gaussianity: II

- consider Gaussian white noise $X_t$ and $Y_t = \text{sign}\{X_t\} \times X_t^2$:

- right-hand plots show estimated PDFs and true PDFs
Effect of Non-Gaussianity: III

- wavelet-domain bootstraps of $X_t$ and $Y_t = \text{sign}\{X_t\} \times X_t^2$:

- right-hand plots show estimated PDFs and true original PDFs
Tree-Based Bootstrapping

- to preserve non-Gaussianity, consider using groups (‘trees’) of wavelet coefficients co-located across small scales as basic sampling unit for bootstrapping at those scales
- wavelet coefficients at large scales treated in same way as in usual wavelet-domain bootstrap
- scaling coefficients handled using parametric bootstrap
- certain wavelet-based signal denoising schemes for non-Gaussian noise treat small scales in a special way and large scales in the same way as in the Gaussian case (see, e.g., Gao, 1997)
- tree-based structuring of wavelet coefficients is key idea behind denoising using Markov models (Crouse et al., 1998) and notion of wavelet ‘footprints’ (Dragotti and Vetterli, 2003)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and level \( j = 3 \) tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and level \( j = 3 \) tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series \( X \) (left-hand column) and level \( j = 3 \) tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $X$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Illustration of Tree-Based Bootstrapping

- Haar DWT of FD(0.45) series $\mathbf{X}$ (left-hand column) and level $j = 3$ tree-based bootstrap thereof (right-hand)
Tree-Based Bootstraps of Non-Gaussian White Noise

- $Y_t$ (top row) and $j = 1$ Haar tree-based bootstrap (bottom)

- Right-hand plots show estimated PDFs and true original PDF
Tree-Based Bootstraps of Non-Gaussian White Noise

- $Y_t$ (top row) and $j = 2$ Haar tree-based bootstrap (bottom)

- Right-hand plots show estimated PDFs and true original PDF
Tree-Based Bootstraps of Non-Gaussian White Noise

- $Y_t$ (top row) and $j = 3$ Haar tree-based bootstrap (bottom)

- right-hand plots show estimated PDFs and true original PDF
Tree-Based Bootstraps of Non-Gaussian White Noise

- $Y_t$ (top row) and $j = 4$ Haar tree-based bootstrap (bottom)

- right-hand plots show estimated PDFs and true original PDF
Tree-Based Bootstraps of Non-Gaussian White Noise

- $Y_t$ (top row) and $j = 5$ Haar tree-based bootstrap (bottom)

- right-hand plots show estimated PDFs and true original PDF
### Summary of Computer Experiments

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Process</th>
<th>Parm</th>
<th>Block</th>
<th>DWT</th>
<th>Tree</th>
<th>Tree</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>AR</td>
<td>0.86</td>
<td>0.83</td>
<td>0.83</td>
<td>0.84</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>FD</td>
<td>0.58</td>
<td>0.57</td>
<td>0.54</td>
<td>0.55</td>
<td>0.57</td>
<td>0.59</td>
</tr>
<tr>
<td>SD</td>
<td>AR</td>
<td>0.016</td>
<td>0.021</td>
<td>0.025</td>
<td>0.025</td>
<td>0.024</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>FD</td>
<td>0.025</td>
<td>0.042</td>
<td>0.054</td>
<td>0.051</td>
<td>0.055</td>
<td>0.059</td>
</tr>
</tbody>
</table>

- 50 time series of length $N = 1024$ for each $Y_t = \text{sign}\{X_t\} \times X_t^2$
- 100 bootstrap samples from each series, yielding 100 unit lag sample autocorrelations $\hat{\rho}_1^{(m)}$
- mean and SD of 100 sample autocorrelations recorded
- table reports averages of these two statistics over 50 time series
- true values based on 100, 000 generated series for each process
Application to BMW Stock Prices - I

- right-hand plot: log of daily returns on BMW share prices
- left-hand: nonparametric and Gaussian PDF estimates
- series has small unit lag sample autocorrelation: $\hat{\rho}_1 \doteq 0.081$.
- large sample theory appropriate for Gaussian white noise gives standard deviation of $1/\sqrt{N} \doteq 0.013$
Application to BMW Stock Prices - II

- bootstrap estimates of standard deviations:

<table>
<thead>
<tr>
<th></th>
<th>LA(8) j = 2</th>
<th>j = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parm Block DWT Tree Tree Gaussian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD est.</td>
<td>0.012 0.016 0.021 0.019 0.019</td>
<td>0.013</td>
</tr>
</tbody>
</table>

- since $\hat{\rho}_1 = 0.081$, bootstrap methods all confirm presence of autocorrelation (small, but presumably exploitable by traders)
Concluding Remarks

- wavelet-domain & tree-based bootstraps competitive with parametric & block bootstraps for series with short-range dependence and offer improvement in case of long-range dependence
- results to date for tree-based bootstrapping encouraging, but many questions need to be answered, including:
  - are there statistics & non-Gaussian series for which tree-based approach offers more than just a marginal improvement over wavelet-domain approach?
  - what are asymptotic properties of tree-based approach?
  - how can the tree-based approach be adjusted to handle stationary processes that are not well decorrelated by the DWT?
Thanks to . . .

- Ann Maharaj and Giovanni Forchini for opportunity to speak
- CSIRO Mathematics, Informatics and Statistics (CMIS) for visiting scientist position allowing for extended Australian visit
- Crime Scene for 2nd place finish in Melbourne Cup (got $13 in $2 sweep – first winnings ever from a horse race!!!)
References: I


References: II

References: III


References: IV

• B. J Whitcher (2006), ‘Wavelet-Based Bootstrapping of Spatial Patterns on a Finite Lattice,’ *Computational Statistics & Data Analysis*, 50(9), pp. 2399–421