Wavelet Analysis on financial time series

Tobias Setz
Semester Thesis Winter 2011

Computational Science and Engineering, ETH
Supervised by PD Diethelm Würtz
Presented at the ZurichR on January 19 in 2012
Content

- Motivation
  - Musical Notation
  - About
- Theory
- Artificial Time Series
- Real World Time Series
- Outro
The idea of a wavelet analysis is very similar to the idea of musical notation.

Musical notation shows you:
- Which frequencies are played at which point in time.
- How long are those frequencies played.
- What volume do those frequencies have (amplitude).

We would like to describe (stochastic) time series in the same way.

Wavelet transformation is a method for such an approach.
Motivation

About

- This presentation is based on my semester thesis on ETH Zürich (supervised by PD Diethelm Würtz) in Winter 2011 which had the following goals:
  - Summarizing the theoretical framework of the dplR library to perform morlet wavelet transformations.
  - Creating artificial time series to get a qualitative feeling on how to interpret the outcome of a wavelet transformation.
  - Creating wavelet transformations and plots using financial time series dating back to the time before the great depression (end of 1923).
  - Qualitative interpretation of the real time series wavelet plots.

- The full thesis and a more detailed presentation can be found here:
  - https://www.rmetrics.org/LecturesCourses
Content

- Motivation
- Theory
  - Construction
  - Visualization
  - Interpretation
- Artificial Time Series
- Real World Time Series
- Outro
Theory
Construction - Ingredients

- Time Series
  - Possibly oscillations
  - Possibly peaks and discontinuities

- Morlet Wavelet Function
  - Nonorthogonal and Complex
  - Zero Mean and Unit Energy
  - Localized in time and frequency space

\[ \Psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \]
Theory
Construction - Convolution

- Take the morlet wavelet function having a certain width ($s$).
- Slide it over the time series under consideration and calculate for every point in time ($n$) the scalar product.
- Do this for several different values of the scale ($s$).

For every point in time ($n$) we get for all the considered scales ($s$) a indicator for certain events within the series.
Since the wavelet function is complex, so is the wavelet transform $W_n(s)$.

Therefore we plot the square of the absolute value $|W_n(s)|^2$ (power spectrum).

- Good trade-off between detecting oscillations and peaks or discontinuities.
- The distribution and therefore expectation values are known.

Contour Plot using the matrix given by $|W_n(s)|^2$. 

\[ \lambda = f(s) \]
**Theory**

**Visualization - Axes and Zero Padding**

- The y-axis does not show the **scale** directly but instead the corresponding fourier frequency which is $\lambda=1.03s$.
  - $\lambda$ is the value where the fourier transformation of $\Psi$ has its maximum.
- Also note that those values are growing exponentially as $s_j=2^{j(1/4)}$.
- Since the data is not cyclic, the time series are **padded with zeros**.
The Cone of Influence (COI) describes the region that is influenced by the zero padding.
- $\tau_s$ is the time for a feature within the original time series that it needs to disappear.

The wavelet levels are the empirical quantiles of $|W_n(s)|^2$ for $p \in [0, 0.1, \ldots, 0.1, 1]$.
- E.g. the 10% highest values form one group.
By describing the time series as a univariate lag-1 autoregressive process (AR(1)) we can derive expectation values for the power spectrum (for m=0).

\[ r_n = m + \alpha r_{n-1} + \epsilon_n \]

\[ \frac{|W_n(s)|^2}{\sigma^2 P_s} \sim \frac{1}{2} \chi^2 \]

Which means that:

\[ E[|W_n(s)|^2] = P_s \sigma^2 \]

\[ P_s = f(P_k) \quad P_k = f(\alpha) \]

For white noise (\(\alpha=0\) and m=0): \(P_s=1\).

This is the reason why we only use the returns of a financial time series for wavelet analyses.
There are several reasons to have values above the expected values

- from a mathematical point of view

Hazard

- In a truly random process the significance level holds 5% of all the values

- Feature of the distribution

Oscillations

- Deviations from a constant linear trend
- Usually found within the index
- Feature of the convolution
- Horizontal sign. levels
- Check if $w \geq T_s$ -> hint for true feature
Theory

Interpretation – Events II

- **Jumps and Steps**
  - Deviations from a constant linear trend
  - Structural Breaks within the index
  - Feature of the convolution
  - Vertical sign. levels

- **Volatility Clusters**
  - Dev. from a const. standard deviation
  - Structural Breaks within the returns
  - Feature of the formula above
  - Vertical sign. levels

- Difficult to make a difference by looking at the spectrum
- Consult the original time series
If we have a real world time series, can we detect:

- Seasonal Cycles
- Oscillations
- Structural Breaks
  - Jumps and Steps
  - Volatility Clusters

Can we assign those findings to real world events?

What can we learn from it?

What further applications do wavelets offer?
Idea found in the book of R. Gencay et al. and the paper of J. B. Ramsey.

Think of a scale as a class of traders that has a time horizon similar to the scale.

This raises interesting questions, namely:

- How is the decision of a trader on a low scale influencing the decision of a trader on a high scale and vice versa?
Content

- Motivation
- Theory
- Artificial Time Series
  - Motivation (link to finance)
  - Construction of artificial time series
  - Different signals (with and without noise)
    - Periodic Trends
      - Pure, Super Position, Frequency Modulation, Amplitude Modulation
    - Jumps and Steps
    - Volatility Clusters
- Real World Time Series
- Outro
Financial time series are usually exponentially growing. We describe this as the index which can approximately be simulated by taking exponential linear brown noise.

In a first step it makes sense to have a look at the log-index. This corresponds to linear brown noise.

$$S_t$$

$$Y_t = \ln S_t$$
Artificial Time Series

Motivation (link to finance)

\[ y_t = Y_{t+1} - Y_t \]

\[ s_t = \ln \left( \frac{S_{t+1}}{S_t} \right) = Y_{t-1} - Y_t = y_t \]

- Taking the differences of the log index removes the trend and reveals the underlying noise and possibly a signal (described as the log-returns).
- The returns can approximately be simulated by taking white noise.
- Directly calculating the log-returns of the original index is equivalent.
Artificial Time Series
Construction of artificial time series

\[ \varepsilon_t = N(\mu, \sigma) = s_t = y_t, \quad \mu = 0, \sigma = 0.05 \]

\[ B_{t+1} = B_t + \varepsilon_t \]

- Gauss Distribution (White Noise)
- To start, normally distributed numbers are generated.
- This is an approximation of the log-returns.

- Brownian Motion (Brown Noise)
- The normal distributed numbers are cumulative summed up.
Artificial Time Series
Construction of artificial time series

\[ Y_t = B_t + X_t^l \]
\[ Y_t = B_t + X_t^l + X_t^s \]

- A signal/trend can now be added to the Brownian noise
- Linear brown noise
  - This is an approximation of the log-prices.
- Periodic Trend
  - Business Cycles
Artificial Time Series

Periodic Linear Trend (Log-Prices)

\[ X_t = A \cdot \sin \left( \frac{2\pi}{\lambda} \cdot t \right) + m \cdot t , \quad \lambda = 12 \]

\[ Y_t = B_t + X_t \]

- Significance in general gets lost compared to the pure periodic trend.
- The significance level is different compared to the pure periodic trend.
- Otherwise the wavelet spectrum looks very similar.
Artificial Time Series
Periodic Linear Trend (Log-Returns)

\[ x_t = X_{t+1} - X_t \]

- The time series is shifted up.
- The wavelet spectrum is identical to the one of the pure periodic trend.

\[ y_t = Y_{t+1} - Y_t \]

- The time series is shifted up.
- The wavelet spectrum is identical to the one of the pure periodic trend.
Artificial Time Series

Periodic Trend (Log-Prices)

\[ X_t = A \cdot \sin \left( \frac{2\pi}{\lambda} \cdot t \right), \quad \lambda = 12 \]

\[ Y_t = B_t + X_t \]

- The amplitude A is 10% of the maximal value of the brownian motion.
- The period can be seen very clearly for \( \lambda = 12 \).

Looking at the spectrum shows no significance level anymore for \( \lambda = 12 \).
Artificial Time Series

Periodic Trend (Log-Returns)

\[ x_t = X_{t+1} - X_t \]

\[ y_t = Y_{t+1} - Y_t \]

- The periodic trend is conserved.
- The signal-noise ratio is clearly better, which results in a significance level for \( \lambda = 12 \).
Artificial Time Series

Periodic Trend super positioned (Log-Prices)

\[ X_t = A \cdot \left( \sin \left( \frac{2\pi}{\lambda_1} \cdot t \right) + \sin \left( \frac{2\pi}{\lambda_2} \cdot t \right) + \sin \left( \frac{2\pi}{\lambda_3} \cdot t \right) \right) \]

\[ Y_t = B_t + X_t, \quad \lambda_1 = 12, \quad \lambda_2 = 24, \quad \lambda_3 = 48 \]

- The amplitude \( A \) is 20\% of the maximal value of the brownian motion.
- The periods can be seen very clearly for \( \lambda=12 \), \( \lambda=24 \) and \( \lambda=48 \).
- Looking at the spectrum shows no clear significance level anymore for \( \lambda=12 \), \( \lambda=24 \) and \( \lambda=48 \).
Artificial Time Series
Periodic Linear Trend super positioned (Log-Returns)

\[ x_t = X_{t+1} - X_t \]

\[ y_t = Y_{t+1} - Y_t \]

- Again we can discover the period for \( \lambda = 12 \), \( \lambda = 24 \) and \( \lambda = 48 \).
- Also here \( \lambda = 12 \), \( \lambda = 24 \) and \( \lambda = 48 \) can be assumed.
Artificial Time Series

Periodic Trend frequency modulated (Log-Prices)

\[ X_t = A \cdot \sin \left( \frac{2\pi}{\lambda_1} t + B \cdot \int_0^t A \cdot \cos \left( \frac{2\pi}{\lambda_2} t \right) dt \right) \]

\[ Y_t = B_t + X_t, \quad \lambda_1 = 12, \quad \lambda_2 = \max(t)/3 \]

- The amplitude \( A \) is 20% of the maximal value of the brownian motion. \( B = 2 \).
- The modulation of the frequency can be seen very clearly.

- The significance level gets lost.
Artificial Time Series

Periodic Linear Trend frequency modulated (Log-Returns)

\[ x_t = X_{t+1} - X_t \]

\[ y_t = Y_{t+1} - Y_t \]

- Again we can discover the modulation of the frequency.
- Also here modulation of the frequency can be assumed.
Artificial Time Series

Periodic Trend amplitude modulated (Log-Prices)

\[ X_t = \sin \left( \frac{2\pi}{\lambda_1} \cdot t \right) \cdot \left( A + B \cdot \cos \left( \frac{2\pi}{\lambda_2} \cdot t \right) \right) \]

\[ Y_t = B_t + X_t, \quad \lambda_1 = 12, \quad \lambda_2 = \max(t)/3 \]

- The amplitude A is 20% of the maximal value of the brownian motion. B = A/2.
- The modulation of the frequency can be seen very clearly.

- The significance level gets lost.
Artificial Time Series

Periodic Linear Trend amplitude modulated (Log-Returns)

\[ x_t = X_{t+1} - X_t \]

\[ y_t = Y_{t+1} - Y_t \]

- Again we can discover the modulation of the amplitude.
- Also here the modulation of the amplitude can be assumed.
Artificial Time Series

Jumps and Steps (Log-Prices)

\[ X_t = A \cdot f(t) \]

\[ Y_t = B_t + X_t \]

- The amplitude A grows from 30% of the maximal value of the brownian motion to 120%.
- The function \( f(t) \) is equal to 1 or -1 at arbitrary points on the time series.
- Jumps and steps can be assumed but not significantly.
- Not that a step is just a jump only upwards or downwards (fourth and fifth peak).
Artificial Time Series
Jumps and Steps Linear (Log-Returns)

\[ x_t = X_{t+1} - X_t \]
\[ y_t = Y_{t+1} - Y_t \]

- Again we can discover the jumps and the steps.
- They seem to grow downwards with increasing amplitude \( A \).
- Now the jumps and steps can be discovered significantly.
Artificial Time Series
Volatility Clusters (Log-Prices and Log-Returns)

\[X_t = \varepsilon_t = N(\mu, \sigma), \quad \mu = 0, \sigma = f(t)\]

\[Y_t = B_t + X_t\]

- Generally we have \(\sigma_1 = 0.05\). Where the red line is bigger than 0, we have \(\sigma_i = c(i) \times \sigma_1\), where \(c \in [2, 2.5, 4, 6]\).
- This is what we often discover in financial time series during crises.
Content

- Motivation
- Theory
- Artificial Time Series
- Real World Time Series
  - Introduction
  - Discussion of the Returns
- Outro
**Real World Time Series**

Introduction

- **History**
  - Which real world events correspond to the significant power spectrum areas?

- **Cause**
  - Seasonal Cycles?
    - Oscillations
  - Structural Breaks?
    - Jumps, Steps and Volatility Clusters

- Also keep in mind the trader model.
This the known analysis from the presentation on the Meielisalp Conference. It is the basis for the historical events that could possibly be seen within the wavelet spectrum.
Real World Time Series
Equities: SP500

- The Standard & Poor’s 500 index.
- The log prices of the Standard & Poor’s 500 index.
Real World Time Series

History Equities: SP500

returns (SP500)

- <Great Depression (09.29)
- <Versailles Treaty (03.36)
- <Poland Attack (09.39)
- <D Day (06.44)
- <End of the war (05.45)
- <Barking Panics (09.31)
- <US Recession (12.69)
- <Vietnam War (09.65)
- <Berlin Wall (08.61)
- <Oil Crisis (10.73)
- <S&L crisis (10.79)
- <S&L crisis (10.89)
- <Dot-Com Bubble (04.00)
- <September 11 (09.01)
- <Sub Prime (02.07)
- <Food Bubble (10.07)
Real World Time Series

Interpretation Equities: SP500

- Generally areas with power above the 95% level can be associated with real world events.

- Generally high power areas seem to be caused by volatility clusters.
- There is no hint about seasonal cycles.
- The shocks within the system seem to appear randomly.
Real World Time Series

Bonds: USGOVT

- **USGOVT**: US Government Bonds.
- **log(USGOVT)**: The log prices of the US Government Bonds.
Real World Time Series

History Bonds: USGOVT

returns($USGOVT$)
Real World Time Series

Interpretation Bonds: USGOVT

- The statements made for the SP500 do also count for US Government Bonds.

- But shocks within the system are due to different historic events than the historic events that shocked the SP500. This could be examined more closely.
- Interesting here is that it seems that more power is in higher scales than for the SP500. The trader model could be an explanation.
Real World Time Series
Commodities: CBR

- The Commodity Research Bureau index.
- The log prices of the Commodity Research Bureau index.
Real World Time Series

History Commodities: CBR

returns($CBR$)
The statements made for the SP500 do also count for Commodities.

But shocks within the system are due to different historic events than the historic events that shocked the SP500. This could be examined more closely.
Content

- Motivation
- Theory
- Artificial Time Series
- Real World Time Series
- Outro
  - About Scales
  - Applications
  - Improvements
Not only time series can be shown in time-frequency space but also economic variables and relationships.

The wavelet spectrum shows that these are highly non stationary and have random components.

Note that every vertical slice (scale) can be separately treated which opens a wide field of applications and tailored solutions to a problem.
Outro

Applications

- This work is a qualitative approach to understand wavelets and getting new insights into the properties of financial time series.
- Of course wavelets are very well suited to be applied in a quantitative sense such as:
  - Seasonality Filtering
  - Denoising (suppression versus averaging (smoothing))
  - Fitting
  - Forecasting (see below)
  - Cross Correlation (multiscale)
  - Outlier Detection
- Despite it seems impossible to forecast random shocks there are some questions worth to consider if a shock occurs within time-frequency space:
  - If fluctuations are constrained, can we model them by investigation each scale separately?
  - How is a shock on one scale influencing other scales (think about traders)?
Bibliography I

- Fiona Grimson, Diethelm Würtz, Yohan Chalabi, Tobias Setz, Heiko Bailer and Simon Otziger
  - Presentation hold on the 5th R/Rmetrics Meielisalp Workshop 2011 (Lake Thun, Switzerland)
    - Part of this work was presented there
  - Does History Repeat Itself? - A long term financial market study using stress and stability metrics
Bibliography II

- Christopher Torrence and Gilbert P. Combo, 1997
  - A Practical Guide to Wavelet Analysis

- Christian Blatter, 1998
  - Wavelets - A Primer

- Ingrid Daubechies, 1992
  - Ten Lectures on Wavelets

- Ramazan Gencay, Faruk Selcuk and Brandon Whitcher, 2002
  - An Introduction to Wavelets and Other Filtering Methods in Finance and Economics

- James B. Ramsey, 2002
  - Wavelets in Economics and Finance: Past and Future