Multiresolution Models of Time Series

Andrea Tamoni

(Bocconi University )

2011
Begin with a univariate time series \( \{x_{t-i}\}_{i \in \mathbb{Z}} \)

For a specified positive integer \( m \), define the process \( x_s^{(m)} \) by

\[
x_s^{(m)} = \frac{1}{m} \sum_{i=0}^{m-1} x_{t+(s-1)m-i}
\]

The \( x_s^{(m)} \) values are the averages of non-overlapping groups of \( m \) consecutive \( x \) values.

Refer to \( m \) as the coarsening window, \( x_t \) as the fine-level process and \( x_s^{(m)} \) as the coarse-level aggregate process.
Time-scale Decomposition

1. Let \( \{x_{t-i}\}_{i \in \mathbb{Z}} \) be the observed time series.
2. Let us focus on block of length \( N = 4 \) (for convenience, we assume \( N = 2^J \) for some \( J \)).

\[
\bar{X}_t = [x_{t-3}, x_{t-2}, x_{t-1}, x_t]^T
\]

3. Consider the two-level (orthogonal) transform matrix, i.e.

\[
\mathcal{W}_2 = \begin{pmatrix}
1 & 1 & 1 & 1 \\
4 & 4 & 4 & 4 \\
-1 & -1 & 1 & 1 \\
-\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\
0 & 0 & -\frac{1}{2} & \frac{1}{2}
\end{pmatrix}
\]
Time-scale Decomposition - Cont’d

- Two-level decomposition is performed as follows

\[ \mathbf{W} = \mathbf{W}_2 \bar{X}_t \]

where \( \mathbf{W} = (a_{2,t}, w_{2,t}, w_{1,t-2}, w_{1,t})^T \)

\[
\begin{align*}
    a_{2,t} &= \frac{(x_t + x_{t-1} + x_{t-2} + x_{t-3})}{4} \\
    w_{2,t} &= \frac{(x_t + x_{t-1} - x_{t-2} - x_{t-3})}{4} \\
    w_{1,t-2} &= \frac{(x_{t-2} - x_{t-3})}{2} \\
    w_{1,t} &= \frac{(x_t - x_{t-1})}{2}
\end{align*}
\]

- It is possible to reconstruct the time series at different levels of aggregation, i.e.

\[
\begin{align*}
    x_t &= a_{2,t} + w_{2,t} + w_{1,t} \\
    x_t^{(2)} &= a_{2,t} + w_{2,t} \\
    x_t^{(4)} &= a_{2,t}
\end{align*}
\]
Note that the $j$-th coefficient is in general obtained as follows

$$w_{j,t} = \frac{\sum_{i=0}^{2^j-1} x_{t-i}}{2^j} - \frac{\sum_{i=0}^{2^{j-1}-1} x_{t-2^{j-1}-i}}{2^{j-1}}$$

Overall $w_{j,t}$ is a band-pass filter with $\left[\frac{1}{2^j-1}, \frac{1}{2^j}\right]$.

Recall that $x_t = \sum_{j=1}^{2} w_{j,t} + a_{2,t}$ and $x_t^{(2)} = w_{2,t} + a_{2,t}$ and $x_t^{(4)} = a_{2,t}$.

Intuitively
- $w_{1,t}$ yields changes on scale 1, i.e. $\ldots t, t+1, t+2, \ldots$
- $w_{2,t}$ yields changes on scale 2, i.e. $\ldots t, t+2, t+4, \ldots$
- $a_{2,t}$ yields changes on scale 2, i.e. $\ldots t, t+4, t+8, \ldots$
Decomposition in time-scale components - Extensions

- Let $\tilde{X}_t = [x_{t-N+1}, \ldots, x_{t-2}, x_{t-1}, x_t]^T$ where $N = 2^J$
- Consider the $J$-level (orthogonal) transform matrix $\mathcal{W}_J$ and construct

$$\mathbf{W} = \mathcal{W}_J \tilde{X}_t$$

- We can partition $\mathbf{W}$ as follows

$$\mathbf{W} = \begin{pmatrix} a_J \\ \mathbf{W}_J \\ \mathbf{W}_{J-1} \\ \vdots \\ \mathbf{W}_1 \end{pmatrix}$$

- We obtain $J + 1$ time series $\{\mathbf{W}_j\}_j, a_J$ each containing $\frac{N}{2^j}$ coefficients, i.e. $\mathbf{W}_j = (w_{j,t-2^j+1}, \ldots, w_{j,t-2^j}, w_{j,t})^T$
- The series $\mathbf{W}_j$ is related to times spaced $2^j$ units apart
Let \( \{x_{t-i}\}_{i \in \mathbb{Z}} \) be a wide-sense stationary process.

Wong (1993) establishes that \( \{w_{j,k}\} \) are wide-sense stationary for a fixed \( j \).

We can write the following representation (understood to be in mean square sense)

\[
x_t = \sum_{j=1}^{J} \sum_{k=0}^{\infty} \tilde{w}_{j,k} \epsilon_{j,t-k \cdot 2^j}
\]

where

\[
\epsilon_{j,t} = x_{t^{(2^j)}} - P_{M_{j,t-2^j}} x_{t^{(2^j)}}
\]

are called Haar innovations and we let \( M_{j,t} = \text{sp} \left\{ x_{t-k2^j}^{(2^j)} \right\}_{k \in \mathbb{N}} \).

The spectral representation (1) holds for a broader class of processes known as “Affine stationary processes” (see Yazici, 1996).
Whenever

\[ \epsilon_{j,t} = \sum_{i=0}^{2j-1} \epsilon_{t-i} \]

\[ x_{t}^{(2j)} - P_{M_{j,t-2j}} x_{t}^{(2j)} = \sum_{i=0}^{2j-1} x_{t-i} - P_{M_{t-i-1}} x_{t-i} \]

where \( M_{j,t} = sp \left\{ x_{t-k2j} \right\}_{k \in \mathbb{N}} \) and \( M_{t} = sp \left\{ x_{t-k} \right\}_{k \in \mathbb{N}} \) then we obtain the standard wold decomposition

The \( sp \left\{ x_{t-k2j} \right\}_{k \in \mathbb{N}} \) is the set of all linear combinations of \( x_{t-k2j} \) and therefore is a subset of the sigma-algebra generated by \( \left\{ x_{t}^{(2j)}, x_{t-2j}^{(2j)}, \ldots \right\} \)
Time series and shocks at different time scales

Scale 1

Scale 2

Scale 3

Scale 4

Scale 5

Scale 6
ARMA time series model for fine-level process

- In this section we not only aggregate data but we also aggregate the model!
- Let \( x_t \) be an ARMA\((p,q)\) process
- Brewer (1973) shows that the time series model governing \( x_s^{(m)} \) is an ARMA\((p^*,q^*)\) where \( p^* = p \) and \( q^* = p + 1 + \frac{(q-p-1)}{m} \) for any finite \( m \)
- Assume \( x_t \) follows an AR(1) process with autoregressive coefficient equal to \( \rho \)
- The pattern of autocorrelations of the temporally aggregated variable \( x_s^{(m)} \) is determined by the autoregressive coefficient \( \rho^m \)
Multi-scale time-series modeling

- What if instead the autocorrelation function at horizon $h$ has a decay governed by $\rho_h > \rho$?
- The question is how can we make the autocorrelation functions at the fine- and coarse-level compatible?
- We use the approach suggested by Dijkerman (1994) and define a multiresolution model on the \( \{W_j\} \) series.
- The multi-scale autoregressive process is defined by an AR(1) process at each level $j$, i.e.

\[
W_{j,t+2^j} = \rho_j W_{j,t} + \epsilon_{j,t+2^j}
\]

for all $j$, where $\epsilon_{j,t+2^j}$ is Gaussian noise with variance equal to $\sigma_j$ independent across scales.
An example

- The fine scale interval is 1 year. Suppose $J = 3$.

\[
a_{3,t} = \rho_4 a_{3,t-8} + \epsilon_{3,t+4}
\]
\[
w_{3,t} = \epsilon_{3,t}
\]
\[
w_{2,t} = \epsilon_{2,t}
\]
\[
w_{1,t} = \rho_1 w_{1,t-2} + \epsilon_{1,t}
\]

- Recall that $x_t^{(8)} = a_{3,t}$ thus $a_{3,t}$ pins down the autocorrelation of the process at coarse level.

- Moreover $x_t = w_{1,t} + w_{2,t} + w_{3,t} + a_{3,t}$. The decay of the autocorrelation function at the fine-level is controlled by the parameters $\sigma_j$, $j = 1, \ldots, J$ (only through the ratio)
Correlation functions on different time scales
Correlation functions on different time scales - Cont’d

Graphs showing simulated correlation functions for fine-level (Multi Resolution) processes and theoretical autocorrelation functions of AR(1) processes with different parameters. The graphs compare the simulated data with the theoretical model predictions for different lags.
Affine Stationary Processes

- A stochastic process $x_{j,t}$ is called affine stationary if it satisfies the following conditions:

$$E[x_{j,t}x_{j',t'}] = R\left(\frac{j}{j'}, \frac{1}{j'}(t - t')\right)$$

- Within the same scale, i.e. $j = j'$ an affine stationary process reduces to an ordinary wide sense stationary process.

- Across the scales for a fixed shift index, i.e. $t = t'$ the process reduces to a class of self-similar processes known as scale stationary process.

- Example: the increments of the fractional Brownian motion