Singular Spectrum Analysis for Signal Extraction in Stochastic Volatility Models

Josu Arteche & Javier Garcia
University of the Basque Country UPV/EHU
eman ta zabal zazu

VI$_t$ Workshop in Time Series Econometrics
Zaragoza, April 2016

Introduction

Two main streams for the modelling of conditional heteroscedasticity:

- (G)ARCH and extensions: one series of innovations drive both levels and conditional variances, and the latter (volatilities) are exact functions of the past.
- SV models: Different series of innovations for levels and volatilities.

Introduction

- SV are more flexible and usually provide better fits to financial time series (Carnero et al., 2004, Franses et al. 2008)

Introduction

• SV are more flexible and usually provide better fits to financial time series (Carnero et al., 2004, Franses et al. 2008)

• However, much less popular than (G)ARCH-based models among empirical researchers because:

Introduction

• SV are more flexible and usually provide better fits to financial time series (Carnero et al., 2004, Franses et al. 2008)

• However, much less popular than (G)ARCH-based models among empirical researchers because:

- Difficult estimation.

Introduction

• SV are more flexible and usually provide better fits to financial time series (Carnero et al., 2004, Franses et al. 2008)

• However, much less popular than (G)ARCH-based models among empirical researchers because:

- Difficult estimation.
- Difficult extraction of the volatility component.

Introduction

• SV are more flexible and usually provide better fits to financial time series (Carnero et al., 2004, Franses et al. 2008)

• However, much less popular than (G)ARCH-based models among empirical researchers because:

- Difficult estimation.
- Difficult extraction of the volatility component. $\Rightarrow$

SSA is here proposed for estimation of the volatility in a SV context.

Stochastic Volatility

SV models are defined as

$$z_t = \sigma_t \varepsilon_t$$

where

- $\varepsilon_t \sim iid(0, 1)$,
- $\sigma_t = \sigma \exp(v_t/2)$ for σ a positive constant scale factor, v_t is the volatility component.

Stochastic Volatility

Linearizing, $y_t = \log z_t^2$

$$\boxed{y_t = \mu + v_t + u_t}$$

where

- $\mu = \log \sigma^2 + E \log \epsilon_t^2$
- $u_t = \log \epsilon_t^2 - E \log \epsilon_t^2$ is a mean zero white noise process with finite variance σ_u^2.

Stochastic Volatility

Linearizing, $y_t = \log z_t^2$

$$\boxed{y_t = \mu + v_t + u_t}$$

where

- $\mu = \log \sigma^2 + E \log \epsilon_t^2$
- $u_t = \log \epsilon_t^2 - E \log \epsilon_t^2$ is a mean zero white noise process with finite variance σ_u^2.

Aim of this work: Extract v_t from y_t.

Stochastic Volatility

Characteristics of v_t:

- Persistence in the form of:
  - Weak dependence (e.g. AR(1)).
  - Stationary or nonstationary long memory.
  - Level shifts.

  In all these cases the periodogram shows a peak at the origin.
- Seasonal or cyclical behaviour in intraday data. The periodogram shows peaks at seasonal/cyclical frequencies.

In any case the periodogram shows a distinctive behaviour that allows identification of the signal separatly from the noise.

Stochastic Volatility

Some strategies to extract v_t:

• **Parametric techniques**: Kalman Filter, (Optimal) Wiener-Kolmogorov filter either in the time domain or frequency domain....

Stochastic Volatility

Some strategies to extract v_t:

• **Parametric techniques**: Kalman Filter, (Optimal) Wiener-Kolmogorov filter either in the time domain or frequency domain....

• **Semiparametric techniques**: Semiparametric Wiener-Kolmogorov for long memory SV models....

Stochastic Volatility

Some strategies to extract v_t:

• **Parametric techniques**: Kalman Filter, (Optimal) Wiener-Kolmogorov filter either in the time domain or frequency domain....

• **Semiparametric techniques**: Semiparametric Wiener-Kolmogorov for long memory SV models....

• **Nonparametric technique**: model-free SSA, robust to misspecification and valid for short memory, long memory and level shifts with no need to know the kind of behaviour that produces the peaks in the periodogram.

Singular Spectrum Analysis

- **Main idea**: Decompose the series as the sum of a finite number of components ⇒ Estimate v_t by selecting those components that share spectral characteristics similar to those in the latent signal.

Singular Spectrum Analysis

• **Main idea**: Decompose the series as the sum of a finite number of components ⇒ Estimate v_t by selecting those components that share spectral characteristics similar to those in the latent signal.

• The spectral characteristics of v_t can be assessed by those of y_t because, under full uncorrelation between v_t and u_t:

$$f_y(\lambda) = f_v(\lambda) + \frac{\sigma_u^2}{2\pi}$$

Singular Spectrum Analysis

• **Main idea**: Decompose the series as the sum of a finite number of components ⇒ Estimate v_t by selecting those components that share spectral characteristics similar to those in the latent signal.

• The spectral characteristics of v_t can be assessed by those of y_t because, under full uncorrelation between v_t and u_t:

$$f_y(\lambda) = f_v(\lambda) + \frac{\sigma_u^2}{2\pi}$$

• **Instrument:** Periodogram of y_t.

Singular Spectrum Analysis

- **Step 1**: Let $y_t^* = y_t - \bar{y}$, $t = 1, 2, ..., n$, the trajectory matrix $Y = [Y_1 : ... : Y_K]$ for $Y_j = (y_j^*, ..., y_{j+L-1}^*)'$ where $1 < L < n$, is the window length and $K = n - L + 1$ $(L \leq K)$.
- **Step 2**: Apply the Singular Value Decomposition (SVD) to Y

$$Y = \sum_{j \in J} \sqrt{\mu_j} U_j V_j' \ , \quad V_j = \frac{1}{\sqrt{\mu_j}} Y' U_j \ , \quad J = \{j \ \text{ such that } \mu_j > 0\},$$

  where μ_j and U_j are the j-th eigenvalue and eigenvector of YY'. U_j is an $L \times 1$ vector known as Empirical Orthogonal Function (EOF).
- **Step 3**: **Reconstruction**. Select a subgroup $\daleth$ and form

$$Y_{\daleth} = \sum_{j \in \daleth} \sqrt{\mu_j} U_j V_j'.$$

  $\daleth$ contains the SVD components with EOF's sharing the same spectral characteristics as the latent signal.

Singular Spectrum Analysis

- **Step 4**: Estimate v_t by Hankelization of the matrix $Y_\gimel$ as

$$\hat{v}_t^{ssa} = \begin{cases} \dfrac{1}{t}\displaystyle\sum_{l=1}^{t} c_{l,t-l+1} & 1 \leq t \leq L, \\ \dfrac{1}{L}\displaystyle\sum_{l=1}^{L} c_{l,t-l+1} & L < t \leq K, \\ \dfrac{1}{n-t}\displaystyle\sum_{l=t-K+1}^{L} c_{l,t-l+1} & K < t \leq n, \end{cases}$$

  where $c_{j,k}$ is the (j,k)-th element of the matrix $Y_\gimel$.

- **Step 5**: The estimated variance of the returns conditional on the volatility component in a general SV model is

$$\hat{\sigma}_t^2 = \hat{\sigma}^2 \exp(\hat{v}_t^{ssa}) \ , \ \hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} z_t^2 \exp(-\hat{v}_t^{ssa}).$$

Singular Spectrum Analysis

- **Step 6**: **Validation**. If the signal extraction is correct the standardised series $\widehat{\varepsilon}_t = z_t/\hat{\sigma}_t$, should be close to *i.i.d.*. In particular their squares should be uncorrelated and a portmanteau test in the squared standardized series can be used to check this property. Failure to satisfy this condition may indicate wrong signal extraction, due perhaps to an inadequate choice of the set $\daleth$, and the procedure returns to Step 3. If no improvement is achieved by changing the components in the reconstruction, the SV model should be rejected as an appropriate model for the series under consideration.

Singular Spectrum Analysis

Selection of tuning parameters:

- The window length L is selected close to and smaller than $[n/2]$ and multiple of the periodicity of the series (Golyandina et al. 2001).

- $\gimel$ is selected based on the form of the periodogram of y_t:

$$I_y(\lambda_j) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} y_t \exp(-it\lambda_j) \right|^2$$

at Fourier frequencies $\lambda_j = 2\pi j/n$, $j = 1, 2, ..., [n/2]$.

SSA:Selection of ℶ

Consider a single peak at frequency $\bar{\lambda}$. Define the *accumulated periodogram* in the interval $[\bar{\lambda}-\omega_1, \bar{\lambda}+\omega_2]$, for $\omega_1, \omega_2 \geq 0$ and a general series $a_1, ..., a_n$ as

$$AP_a(\bar{\lambda}, \omega_1, \omega_2) = \sum_{j=-[n\omega_1/2\pi]}^{[n\omega_2/2\pi]} I_a(\bar{\lambda}+\lambda_j)$$

SSA:Selection of $\gimel$

Consider a single peak at frequency $\bar{\lambda}$. Define the *accumulated periodogram* in the interval $[\bar{\lambda} - \omega_1, \bar{\lambda} + \omega_2]$, for $\omega_1, \omega_2 \geq 0$ and a general series $a_1, ..., a_n$ as

$$AP_a(\bar{\lambda}, \omega_1, \omega_2) = \sum_{j=-[n\omega_1/2\pi]}^{[n\omega_2/2\pi]} I_a(\bar{\lambda} + \lambda_j)$$

and the *Relative Spectral Contribution (RSC)* of frequencies in that interval as

$$RSC_a(\bar{\lambda}, \omega_1, \omega_2) = \frac{AP_a(\bar{\lambda}, \omega_1, \omega_2)}{AP_a(0, 0, \pi)}$$

$\gimel$ contains those EOF's with $RSC_{U_j}(\bar{\lambda}, \omega_1, \omega_2) > RSC_y(\bar{\lambda}, \omega_1, \omega_2)$.

SSA:Selection of $\daleth$

Consider a single peak at frequency $\bar{\lambda}$. Define the *accumulated periodogram* in the interval $[\bar{\lambda} - \omega_1, \bar{\lambda} + \omega_2]$, for $\omega_1, \omega_2 \geq 0$ and a general series $a_1, ..., a_n$ as

$$AP_a(\bar{\lambda}, \omega_1, \omega_2) = \sum_{j=-[n\omega_1/2\pi]}^{[n\omega_2/2\pi]} I_a(\bar{\lambda} + \lambda_j)$$

and the *Relative Spectral Contribution (RSC)* of frequencies in that interval as

$$RSC_a(\bar{\lambda}, \omega_1, \omega_2) = \frac{AP_a(\bar{\lambda}, \omega_1, \omega_2)}{AP_a(0, 0, \pi)}$$

$\daleth$ contains those EOF's with $RSC_{U_j}(\bar{\lambda}, \omega_1, \omega_2) > RSC_y(\bar{\lambda}, \omega_1, \omega_2)$.

If $\bar{\lambda} = 0$, $\daleth$ contains EOF's with $RSC_{U_j}(0, 0, \omega_2) > RSC_y(0, 0, \omega_2)$.

SSA:Selection of $\beth$

Consider now that k spectral peaks at frequencies $\bar{\lambda}_i$, $i = 1, .., k$ are identified. The EOFs selected in $\beth$ are those showing some of those peaks in their periodograms.
Define the *Multiple Relative Spectral Contribution (MRSC)* of frequencies in the interval $[\bar{\lambda}_i - \omega_{i1}, \bar{\lambda}_i + \omega_{i2}]$ as

$$MRSC_a(\bar{\lambda}_i, \omega_{i1}, \omega_{i2}) = \frac{AP_a(\bar{\lambda}_i, \omega_{i1}, \omega_{i2})}{AP_a(0, 0, \pi) - \sum_{l \neq i;\ l=1}^{k} AP_a(\bar{\lambda}_l, \omega_{l1}, \omega_{l2})}.$$

Note that if $k = 1$ then $MRSC_a = RSC_a$. The EOFs selected satisfy $MRSC_{U_j}(\bar{\lambda}_i, \omega_{i1}, \omega_{i2}) > MRSC_y(\bar{\lambda}_i, \omega_{i1}, \omega_{i2})$ for some $i = 1, ..., k$.

SSA:Selection of $\daleth$

Consider now that k spectral peaks at frequencies $\bar{\lambda}_i$, $i = 1, .., k$ are identified. The EOFs selected in $\daleth$ are those showing some of those peaks in their periodograms.

Define the *Multiple Relative Spectral Contribution (MRSC)* of frequencies in the interval $[\bar{\lambda}_i - \omega_{i1}, \bar{\lambda}_i + \omega_{i2}]$ as

$$MRSC_a(\bar{\lambda}_i, \omega_{i1}, \omega_{i2}) = \frac{AP_a(\bar{\lambda}_i, \omega_{i1}, \omega_{i2})}{AP_a(0, 0, \pi) - \sum_{l \neq i;\ l=1}^{k} AP_a(\bar{\lambda}_l, \omega_{l1}, \omega_{l2})}.$$

Note that if $k = 1$ then $MRSC_a = RSC_a$. The EOFs selected satisfy $MRSC_{U_j}(\bar{\lambda}_i, \omega_{i1}, \omega_{i2}) > MRSC_y(\bar{\lambda}_i, \omega_{i1}, \omega_{i2})$ for some $i = 1, ..., k$.

Finally, ω_{i1}, ω_{i2} can be selected from among those passing the validation.

Monte Carlo Analysis

$n = 2048$, $\sigma = 1$ and $u_t = \log \chi_1^2$. Processes for v_t:

- Model 1: $(1-B)^d v_t = \sigma_1 w_t$ for $d = 0.4$.
- Model 2: $(1-0.8B)v_t = \sigma_2 w_t$.
- Model 3: $v_t = v_{1t} + \mu_t$ where v_{1t} is the AR(1) in Model 2 and μ_t is a level shift of the form:
  - 3a) Deterministic level shift: $a_t = I_{t>n/4}$ and $\mu_t = a_t - \bar{a}$.
  - 3b) Stochastic level shifts: $\mu_t = \sum_{j=1}^{t} \delta_j \eta_j$.
- Model 4: $(1+B^2)^d v_t = \sigma_1 w_t$ for $d = 0.4$.
- Model 5: $(1-0.8B^2)v_t = \sigma_2 w_t$.
- Model 6: $(1-B)^d v_t = \sigma_1 w_t$ for $d = 0.8$.

w_t and η_j are independent standard normal and $\delta_j \sim B(1, 10/n)$. w_t, η_r and δ_s are mutually independent for all t, r, s. Two $NSR = \pi^2$ and $5\pi^2$ ($\sigma_1 = \sqrt{0.5}$, $\sqrt{0.1}$ and σ_2 is adjusted to equalize variances of v_t).

Monte Carlo Analysis

Tuning parameters:

- $L = 1008$ (other values give similar results).
- $w_1 = w_2 = \frac{2\pi k}{n}$ for $k = 20, 50$ and 100.
  Also an automatic criterion: w_2 is the largest frequency such that the peridogram of y_t at frequencies lower than w_2 are larger than the first (Q_1), second (Q_2) and third (Q_3) quartiles of the periodogram ordinates of y_t.

Monte Carlo Analysis

Criteria for finite sample performance evaluation:

• Global typified MCMSE:

$$MCMSE = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{\sigma_{v,t}^2}\frac{1}{1000}\sum_{k=1}^{1000}(\hat{v}_{t,k}-v_{t,k})^2 \qquad (1)$$

where the subindex t, k indicates observation t in the Monte Carlo replication k and we standardise by the variance of v_t for Models 1 to 5 ($\sigma_{v,t}^2 = var(v_{1t}) + tp(1-p)$ in Model 3b). In Model 6 $\sigma_{v,t}^2$ is the variance of the first differences of v_t.

• Correlation between the true signal and its SSA estimation.

• Validation: number of times that the squared standardised series show no autocorrelation (LB(100)).

Monte Carlo Analysis

Table : MCMSE and correlation between signal and SSA estimates

Low NSR					High NSR				
SSA (k=20)	SSA (k=50)	SSA (k=100)	SSA (Q3)	WK	SSA (k=20)	SSA (k=50)	SSA (k=100)	SSA (Q3)	WK
Model 1									
0.825	0.872	0.984	0.842	0.788	1.676	2.295	3.162	1.707	4.017
(0.505)	(0.521)	(0.516)	(0.477)	(0.545)	(0.306)	(0.293)	(0.276)	(0.297)	(0.289)
[818]	[815]	[750]	[717]	[842]	[669]	[540]	[432]	[731]	[678]
Model 2									
0.770	0.741	0.782	0.859	1.973	1.651	2.154	2.924	2.134	8.536
(0.505)	(0.572)	(0.608)	(0.457)	(0.507)	(0.304)	(0.338)	(0.347)	(0.261)	(0.243)
[402]	[699]	[824]	[265]	[568]	[689]	[611]	[506]	[592]	[732]
Model 3a									
0.788	0.740	0.769	0.847	0.885	1.312	1.819	2.619	1.164	1.311
(0.584)	(0.641)	(0.667)	(0.535)	(0.642)	(0.640)	(0.603)	(0.561)	(0.654)	(0.646)
[337]	[679]	[825]	[225]	[749]	[802]	[752]	[582]	[717]	[713]
Model 3b									
1.036	1.014	1.012	1.043	4.194	1.334	1.354	1.398	1.332	3.418
(0.758)	(0.793)	(0.807)	(0.742)	(0.815)	(0.859)	(0.847)	(0.819)	(0.856)	(0.857)
[188]	[511]	[782]	[172]	[778]	[609]	[780]	[750]	[483]	[721]
Model 4									
0.681	0.790	1.002	0.666	-	2.237	3.239	4.553	1.670	-
(0.539)	(0.540)	(0.522)	(0.517)	-	(0.308)	(0.289)	(0.268)	(0.315)	-
[873]	[741]	[241]	[821]	-	[511]	[175]	[25]	[745]	-
Model 5									
0.791	0.819	0.970	0.837	-	2.448	3.243	4.493	2.225	-
(0.543)	(0.591)	(0.595)	(0.488)	-	(0.309)	(0.326)	(0.314)	(0.261)	-
[684]	[823]	[603]	[478]	-	[481]	[226]	[39]	[557]	-

Monte Carlo Analysis

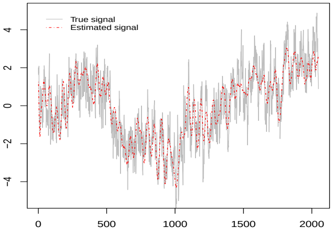
Summary of the results:

- The signal becomes more difficult to estimate as the NSR increases.
- The optimal WK filter (Arteche, 2015, ET) beats the SSA in Models 1 and 6 (long memory SV).
- The SSA beats the WK in the rest of situations.
- Validation: the number of non rejections of no autocorrelation can be close to the desired 950 for some of the selected boundaries.
- Decreasing the degree of persistence from long memory to short memory has little effect on the estimation of the volatility.
- In general, the level shifts have a positive effect on the estimation of the signal.

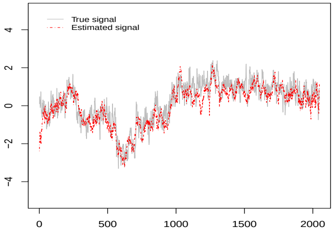
Monte Carlo Analysis: Level shifts

Figure : Signal extraction with level shifts, 1 series

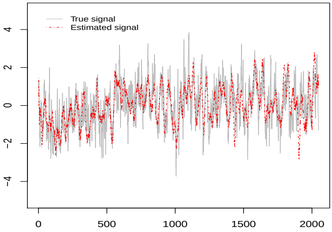
(a) Stochastic level shift (low NSR)



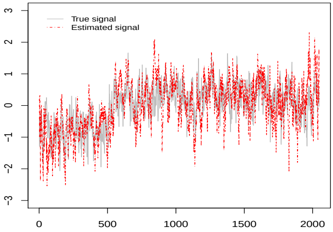
(b) Stochastic level shift (large NSR)



(c) Deterministic level shift (low NSR)



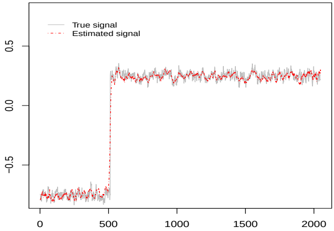
(d) Deterministic level shift (large NSR)



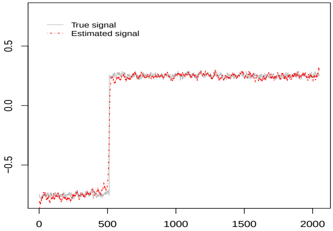
Monte Carlo Analysis: Level shifts

Figure : Deterministic level shifts (Models 3a): Averages

(a) Deterministic level shift (low NSR)



(b) Deterministic level shift (large NSR)



Monte Carlo Analysis: Gains

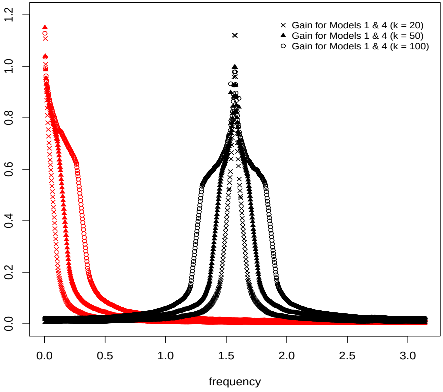
The *gains* are:

$$G(\lambda) = \frac{\hat{f}_{\hat{v}}(\lambda)}{f_y(\lambda)}$$

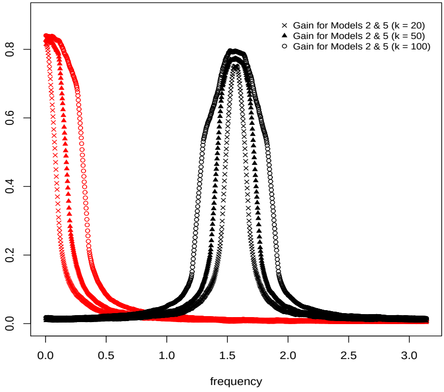
Monte Carlo Analysis: Gains

Figure : Gains of SSA filter

(a) Models 1 and 4 (long memory)



(b) Models 2 and 5 (short memory)



Daily Dow Jones Industrial Index

Sample: 12/12/1996 to 14/11/2015

Figure : Dow Jones returns: Periodograms

(a) Returns (z_t) (b) Log squared returns (y_t)

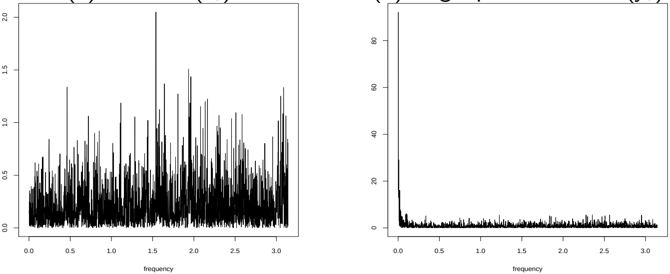
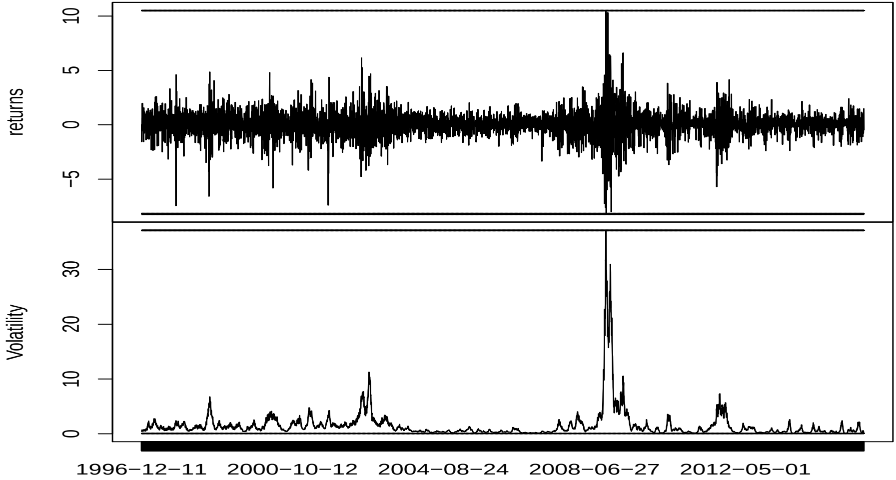


Table : SSA estimation of volatility: parameters used and validation

	L	ω_{02}	p-value LB (100)	p-value LB(200)
DJ Index	2268	$\frac{2\pi 100}{4633}$	0.06	0.29

Daily Dow Jones Industrial Index

Figure : Dow Jones returns and estimated volatility



Ibex35 stock index

Intraday returns (every 90 minutes): 1/10/1993 to 22/03/1996.

Figure : Ibex35 filtered returns: Periodograms

(a) Filtered returns (z_t) (b) Log squared returns (y_t)

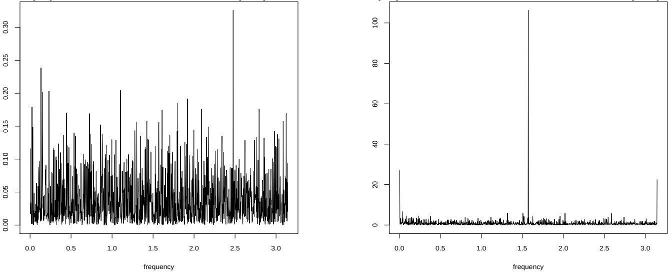
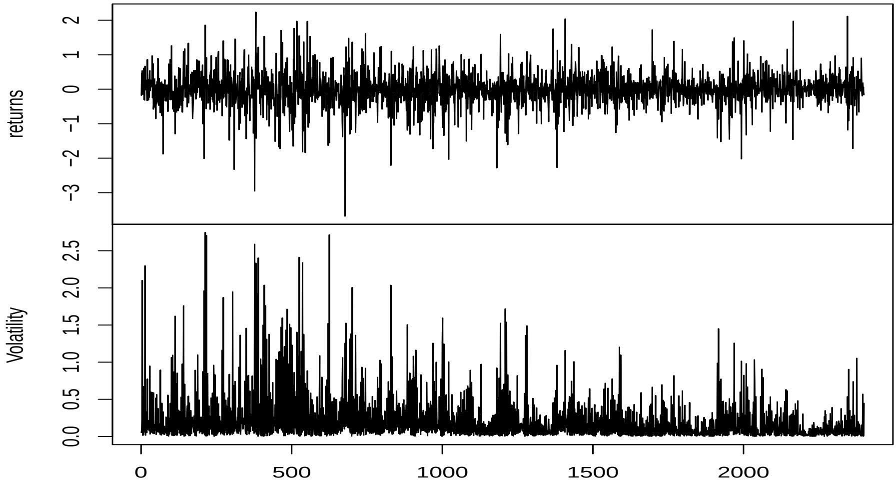


Table : SSA estimation of volatility: parameters used and validation

	L	ω_{02}	ω_{11}	ω_{12}	ω_{21}	p-v. LB (100)	p-v. LB(200)
Ibex35	1024	$\frac{2\pi 10}{2400}$	$\frac{2\pi 100}{2400}$	$\frac{2\pi 100}{2400}$	$\frac{2\pi 70}{2400}$	0.46	0.56

Ibex35 stock index

Figure : Ibex35 filtered returns and estimated volatility



Conclusions

- SSA has been proven to be a powerful tool for estimation of the volatility in SV models.
- It is fully nonparametric: it does not require restrictions on the volatility component (valid for stationary and nonstationary series, weak and strong dependent, level shifts...)
- Tuning parameters can be selected based on a validating criterion.
- It can be used in more general settings for signal extraction.


Thank You