Estimation and Forecast of non-constant Volatility[1]. Testing of Black-Scholes formula.

L. Goldentouch F. Klebaner R. Liptser

March 2007

[1] math.PR/0404277

1. Black-Scholes model for financial markets

The stock price $S(t)$ is modeled as a Geometric Brownian motion

$$dS(t) = rS(t)dt + \sqrt{v(t)}S(t)dB_t,\ 0 \le t \le T,$$

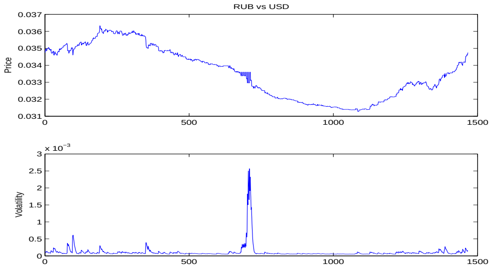
where
- B_t is a Brownian motion
- r is non-risky rate
- $v(t)$ is volatility (risky rate; positive and deterministic).

The averaged volatility $\frac{1}{T-t}\int_t^T v(s)ds$ involves in Black-Scholes formula for "prediction" $\mathrm{E}([S_T - K]^+ | S_t)$ of stock prices, where T and K are the maturity time and strike respectively.

2. Examples

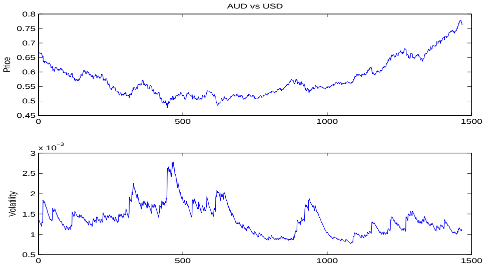
L. Goldentouch, F. Klebaner, R. Liptser

Exchange rate RUB vs USD. Fast oscillations with small amplitudes of the price provide a jumps of volatility



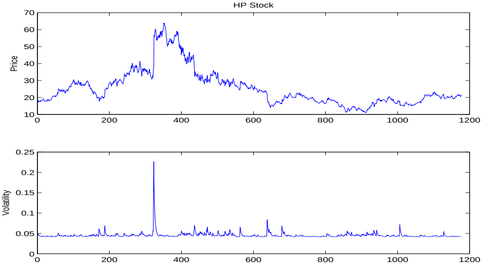
L. Goldentouch, F. Klebaner, R. Liptser

Exchange rate Australian vs USA dollars

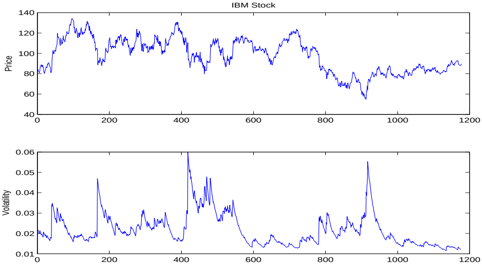


L. Goldentouch, F. Klebaner, R. Liptser

Hewlett-Packard stock



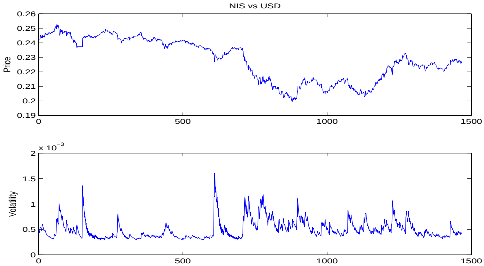
IBM stock


IBM Stock
Price
140
120
100
80
60
40
0
200
400
600
800
1000
1200
Volatility
0.06
0.05
0.04
0.03
0.02
0.01

6. Examples

L. Goldentouch, F. Klebaner, R. Liptser

Exchange New Israeli Shekel vs USA dollar



7. Volatility as unknown function

Estimation of $v(t)$ via the sample $X_i = \frac{1}{\triangle}\log^2\left(\frac{S(t_i)}{S(t_{i-1})}\right)$, where $i = 0, 1, \dots, n, \quad t_i - t_{i-1} = \frac{T}{n}(=: \triangle)$.

Due to the Itô formula

$$\log\left(\frac{S(t_i)}{S(t_{i-1})}\right) = \int_{t_{i-1}}^{t_i} 0.5\big(2r - v(s)\big)ds + \int_{t_{i-1}}^{t_i} \sqrt{v(s)}dB_s,$$

and, for small $\triangle$,

$$X_i = \frac{1}{\triangle}\log^2\left(\frac{S(t_i)}{S(t_{i-1})}\right) \overset{\mathbb{L}^2}{=} \frac{1}{\triangle}\Big(\int_{t_{i-1}}^{t_i} \sqrt{v(s)}dB_s\Big)^2 + O\Big(\sqrt{\triangle}\Big)$$

8. $v(t)$ as smooth function

Then $v_{i-1} = \frac{1}{\triangle}\int_{t_{i-1}}^{t_i} v(s)ds$ is a reasonable approximation for $v(t)$, $t \in [t_{i-1}, t_i)$. Denote

$$\xi_i = \frac{1}{\sqrt{v_{i-1}\triangle}}\int_{t_{i-1}}^{t_i}\sqrt{v(s)}dB_s.$$

$(\xi_i)_{i\geq 1}$ is a sequence of i.i.d. $(0,1)$-Gaussian random variables.

We have

$$\underbrace{X_i}_{\text{measurement}} = \underbrace{v_{i-1}}_{\text{volatility}} + \underbrace{v_{i-1}(\xi_i^2 - 1) + O(\sqrt{\triangle})\xi_i}_{\text{“ noise“}\equiv\eta_i} + \underbrace{O(\triangle)}_{\text{small shift}}$$

9. GARCH TRACKING ALGORITHM

L. Goldentouch, F. Klebaner, R. Liptser

GARCH = Generalized Autoregressive Conditional Heteroscedasticity

Following Bollerslev and Engle, (p, q)-GARCH tracking algorithm estimates v_i's with the help of recursive procedure (with appropriate initial conditions)

$$\widehat{v}_i = K + \sum_{j=1}^{p} g_j \widehat{v}_{i-j} + \sum_{m=1}^{q} a_m X_{(i+1)-m}.$$

Parameters $\boxed{K, g_1, \dots, g_p, a_1, \dots, a_q, p, q}$ are chosen by minimization of the innovation difference (here $n = \frac{T}{\triangle}$):

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - \widehat{v}_{i-1})^2.$$

10. Alternative approach to GARCH algorithm

At first glance the GARCH filter corresponds is a version of the Wiener filter related to a stationary process with rational spectra.

We attempt to find an approach for creating GARCH-type filters based on ideas from

Nonparametric Statistic Inferences

It is known from Ibragimov - Khasminskii and Stone, there exists a kernel type estimate $\widehat{v}_i$ such that

$$\sup_{i \le n} \mathrm{E}(v_i - \widehat{v}_i)^2 \le O\big(n^{-2(1+k)/(2k+3)}\big)$$

provided that $v(t)$ possesses k derivative $v^{(j)}(t), j = 1, \dots, k$ and k-th derivative is Lipschitz continuous.

11. Alternative approach to GARCH algorithm

L. Goldentouch, F. Klebaner, R. Liptser

Any GARTH filter requires an initial conditions implying some boundary layer, where the rate in n guaranteeing by nonparametric estimate might be lost. However, this boundary layer is not large: $i \leq O\big(n^{-1/(2k+3)} \log n\big)$, (Khasminskii and Liptser) and out of it: $i > O\big(n^{-1/(2k+3)} \log n\big)$,

$$\sup_{i \leq n} \mathrm{E}(v_i - \widehat{v}_i)^2 \leq O\big(n^{-2(1+k)/(2k+3)}\big).$$

This result can be obtained with GARCH of Kalman's type filter.

12. GARCH as Kalman filter

It is known from Khasminskii, Liptser that for k-times differentiable function $v(t)$ with the Lipschitz continuous elder derivative the Kalman filter providing the same rate rate in n: $\boxed{\sup_{i\le n}\mathrm{E}(v_i-\widehat{v}_i)^2\le O\big(n^{-2(1+k)/(2k+3)}\big)}$ has the following structure

$$
\begin{aligned}
\widehat{v}_i &= \widehat{v}_{i-1}+\frac{1}{n}\widehat{v}^{(1)}_{i-1}+\frac{q_0(\vartheta)}{n^{(2(k+1))/(2k+3)}}\big(X_i-\widehat{v}_{i-1}\big)\\
\widehat{v}^{(j)}_i &= \widehat{v}^{(j)}_{i-1}+\frac{1}{n}\widehat{v}^{(j+1)}_{i-1}+\frac{q_j(\vartheta)}{n^{(2(k+1)-j)/(2k+3)}}\big(X_i-\widehat{v}_{i-1}\big), j\le k-1\\
\widehat{v}^{(k)}_i &= \widehat{v}^{(k)}_{i-1}+\frac{q_k(\vartheta)}{n^{(k+2)/(2k+3)}}\big(X_i-\widehat{v}_{i-1}\big),
\end{aligned}
$$

for any $q_i(\vartheta)$'s guaranteeing the filter stability.

L. Goldentouch, F. Klebaner, R. Liptser

A remarkable property of GARCH in the form of Kalman's filter is that all coefficients $q_i(\vartheta)$, $i = 0, 1, \ldots, k$ are parameterized by a scalar parameter ϑ

$$q_j(\vartheta) = U_{0j}\vartheta^{j+1/k+1}, \ i = 0, 1, \ldots, k,$$

$U_{0j}, \ j = 0, 1, \ldots, k$ are entries of a matrix U solving the algebraic Riccati equation: $\boxed{aU + Ua^* + B - UA^*AU = 0}$,

where matrices $\boxed{a, A, B}$ are related to "smooth assumptions" of $v(t)$ and a structure of the observation process.

14. Matrices a, A, B

L. Goldentouch, F. Klebaner, R. Liptser

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \end{pmatrix}, \; a = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

of corresponding sizes $1 \times (k+1)$; $(k+1) \times k+1$; $(k+1) \times 1$.

15. U_{0j}

For $k \leq 4$,

k	U_{00}	U_{01}	U_{02}	U_{03}	U_{04}
0	1	NA	NA	NA	NA
1	$\sqrt{2}$	1	NA	NA	NA
2	2	2	1	NA	NA
3	$\sqrt{4+\sqrt{8}}$	$2+\sqrt{2}$	$\sqrt{4+\sqrt{8}}$	1	NA
4	$1+\sqrt{5}$	$3+\sqrt{5}$	$3+\sqrt{5}$	$1+\sqrt{5}$	1

.

16. Global adaptation in ϑ

Measurements:

$$X_i = v_{i-1} + v_{i-1}(\xi_i^2 - 1) + O(\sqrt{\triangle})\xi_i + O(\triangle)$$

We have the filter with one free parameter $\boldsymbol{\vartheta}$:

$$\widehat{v}_i = \widehat{v}_{i-1} + \frac{1}{n}\widehat{v}^{(1)}_{i-1} + \frac{U_{00}\boldsymbol{\vartheta}^{\mathbf{1/k+1}}}{n^{2k+2/2k+3}}\big(X_i - \widehat{v}_{i-1}\big)$$

$$\widehat{v}^{(j)}_i = \widehat{v}^{(j)}_{i-1} + \frac{1}{n}\widehat{v}^{(j+1)}_{i-1} + \frac{U_{0j}\boldsymbol{\vartheta}^{\mathbf{(j+1)/k+1}}}{n^{(2(k+1)-j)/(2k+3)}}\big(X_i - \widehat{v}_{i-1}\big), \; j \le k-1$$

$$\widehat{v}^{(k)}_i = \widehat{v}^{(k)}_{i-1} + \frac{U_{0k}\boldsymbol{\vartheta}}{n^{(k+2)/(2k+3)}}\big(X_i - \widehat{v}_{i-1}\big).$$

17. Global Adaptation in ϑ

$$X_i = \underbrace{v_{i-1}}_{\text{volatility}} + \underbrace{v_{i-1}(\xi_i^2 - 1) + O(\sqrt{\triangle})\xi_i}_{\text{noise}} + \underbrace{O(\triangle)}_{\text{shift}}$$

$\boxed{\text{If } O(\triangle) \equiv 0,}$ the noise variance is bounded. Then, by Ibragimov - Khasminskii

$$\mathrm{E}(v_i - \widehat{v}_i)^2 \le O\big(n^{-2(1+k)/(2k+3)}\big),\ n \to \infty.$$

For $\boxed{O(\triangle) \not\equiv 0,}$ the same upper bound remains valid since

$$O(\triangle) \asymp \frac{1}{n}.$$

18. Adaptation procedure

It is naturale to choose (recall $\widehat{v}_{i-1} = \widehat{v}_{i-1}(\vartheta)$)

$$\vartheta^* = \operatorname*{argmin}_{\vartheta} \frac{1}{n}\sum_{i=1}^{n}(v_{i-1} - \widehat{v}_{i-1}(\vartheta))^2$$

Since $\boxed{v_{i-1}\text{'s}}$ are unknown(!), we are compelled to choose

$$\vartheta^{**} = \operatorname*{argmin}_{\vartheta} \frac{1}{n}\sum_{i=1}^{n}\Big(\underbrace{X_i}_{=v_{i-1}+\text{noise}+O(\triangle)} - \widehat{v}_{i-1}(\vartheta)\Big)^2.$$

Fortunately

$$\mathrm{P}\big(|\vartheta^{**} - \vartheta^*| > \delta\big) \asymp O\big(n^{-(4k+5)/(2k+3)}\big) = o\big(\underbrace{n^{-2(1+k)/(2k+3)}}_{\text{optimal rate in } n\to\infty}\big).$$

19. Filters controlled by a few parameters

$k = 0$ corresponds to Lipschitz continuous $v(t)$. This function can grow in t not faster than a linear function. In this case, we propose the estimator

$$\widehat{v}_i = \widehat{v}_{i-1} + \frac{\vartheta}{n^{2/3}}\big(X_i - \widehat{v}_{i-1}\big).$$

In the real world $v(t)$ is bounded function. As $v(t)$ is Lipschitz continuous and bounded, one may assume that for some positive number $\boldsymbol{a}$:

$$|\dot{v}(t) + \boldsymbol{a}v(t)| \leq \text{const.}$$

20. Filters controlled by a few parameters

The additional information provides a new filter

$$\widehat{v}_i = \widehat{v}_{i-1}\Big(1 - \frac{a}{n}\Big) + \frac{a\overline{X}}{n} + \frac{\vartheta}{n^{2/3}}\big(X_i - \widehat{v}_{i-1}\big),$$

where $\boxed{\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \frac{1}{T}\int_0^T v(t)dt.}$ For a and $\overline{X}$, being sufficiently smaller than n, the rate in n is preserved:

$$\mathrm{E}(v_i - \widehat{v}_i)^2 \le O\big(n^{-2/3}\big).$$

The additional term $\boxed{-\widehat{v}_{i-1}\frac{a}{n} + \frac{\overline{X}}{n}}$ makes tracking error a little bit smaller owing to the boundary layer becomes shorter.

21. Filters controlled by a few parameters

L. Goldentouch, F. Klebaner, R. Liptser

For $k \geq 1$, the filter is transformed to

$$\widehat{v}_i = \widehat{v}_{i-1} + \frac{1}{n}\widehat{v}^{(1)}_{i-1} + \frac{U_{00}\vartheta^{1/k+1}}{n^{2(k+1)/(2k+3)}}\big(X_i - \widehat{v}_{i-1}\big)$$

$$\widehat{v}^{(j)}_i = \widehat{v}^{(j)}_{i-1} + \frac{1}{n}\widehat{v}^{(j+1)}_{i-1} + \frac{U_{0j}\vartheta^{(j+1)/k+1}}{n^{(2(k+1)-j)/(2k+3)}}\big(X_i - \widehat{v}_{i-1}\big)$$

$$j = 1, \dots, k-1$$

$$\widehat{v}^{(k)}_i = \widehat{v}^{(k)}_{i-1}\Big(1 - \frac{a_1}{n}\Big) - \frac{1}{n}\Big(\sum_{\ell=2}^{k-1} a_\ell \widehat{v}^{(k-\ell)}_{i-1} + a_k \widehat{v}_{i-1} + a_k \overline{X}\Big)$$

$$+ \frac{U_{0k}\vartheta}{n^{(k+2)/(2k+3)}}\big(X_i - \widehat{v}_{i-1}\big).$$

22. Tuning procedure: Filter 1($k = 0$) and Filter 2(k=1)

L. Goldentouch, F. Klebaner, R. Liptser

Filter 1

1. Set $\boldsymbol{a} = 0$, $\overline{X} = 0$ and find ϑ^*.
2. Find $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
3. For fixed $\overline{X}$ and ϑ^*, find $\boldsymbol{a}^*$.
4. Local minimization in $(\vartheta, \overline{X}, \boldsymbol{a})$ in a vicinity of $(\vartheta^*, \overline{X}, \boldsymbol{a}^*)$.

Filter 2

The same procedure with $(\vartheta, \overline{X}, \boldsymbol{a_1}, \boldsymbol{a_2})$.

23. Comparison with classical GARCH

Table: Stocks of large computer manufacturers

Asset name	Filter type				
	Garch(1,1)	Garch(2,2)	Filter 0	Filter 1	Filter 2
DIS	4.309e-003	4.314e-003	4.329e-003	4.284e-003	4.285e-003
HPQ	3.741e-002	3.771e-002	3.615e-002	3.608e-002	3.608e-002
IBM	3.229e-003	3.228e-003	3.217e-003	3.199e-003	3.200e-003
INTC	1.232e-002	1.230e-002	1.235e-002	1.223e-002	1.232e-002
MAT	1.899e-002	1.879e-002	1.850e-002	1.817e-002	1.847e-002
SUN	5.144e-004	5.131e-004	5.143e-004	5.135e-004	5.138e-004
TOY	6.134e-003	6.128e-003	6.158e-003	6.098e-003	6.082e-003

A different quality of Filters 1, 2 and GARCH(1,1), (2,2) is provided by different tuning procedures. Filter 1 provides the best quality.

24. Calibration of Black-Scholes Model. Implied volatility.

When $v(t) \equiv v$ (classical Black-Scholes model)

$$\mathrm{E}\big([S_T - K]^+ | S_t\big) = S_t \Phi(h_t) - K e^{-r(T-t)} \Phi\big(h_t - \sqrt{v(T-t)}\big),$$

where Φ is the standard normal distribution function and

$$h_t = \frac{\log(S_t/K) + (r + \frac{1}{2}v)(T-t)}{\sqrt{v(T-t)}}.$$

Since a volatility $\boxed{v}$ is unknown, one can compute $\boxed{\text{implied volatility } v^{C_{T-t}}}$ which solves the equation

$$\mathrm{E}\big([S_T - K]^+ | S_t\big) =: C_t$$

with prescribed stock price C_t.

25. Calibration of Black-Scholes Model for time dependent volatility

For $v(t)$, the calibration can be done for $\boxed{\frac{1}{T-t}\int_t^T v(s)ds}$ since

$$\begin{aligned}&\mathrm{E}\big([S_T-K]^+|S_t\big)\\&=S_t\Phi(\widetilde{h}_t)-Ke^{-r(T-t)}\Phi\left(\widetilde{h}_t-\sqrt{\int_t^T v(s)ds}\right),\end{aligned}$$

where

$$\widetilde{h}_t=\frac{\log(S_t/K)+\int_t^T(r+\frac{1}{2}v(s))ds}{\sqrt{\int_t^T v(s)ds}}.$$

26. Volatility forecasting: historical + implied volatility

For $t = T - \triangle$, we compute the implied volatility $v^{C_\triangle}$.
For $m \le n-2$, we have estimates:
- $\widehat{v}_1, \dots, \widehat{v}_m$ - (estimates of historical volatility)
- $\widehat{v}_n := v^{C_\triangle}$ - (estimates of implied volatility)
We have typical smoothing problem:

$$\widehat{v}_1, \dots, \widehat{v}_m, \underbrace{v_{m+1}, \dots, v_{n-2}, v_{n-1}}_{\text{how to estimates these values?}}, \widehat{v}_n$$

A helpful tool here is a model of corresponding Kalman filter.

27. Example. IBM stock

Two filters have being used for tracking v_i, $1, \ldots, n$:

$$\widehat{v}_i = \left(1 - \frac{a}{n}\right) \widehat{v}_{i-1} + \frac{a\overline{X}}{n} + \text{“innov. difference”}$$

and

$$\widehat{v}_i = \widehat{v}_{i-1} + \frac{1}{n}\widehat{v}^{(1)}_{i-1} + \text{“innov. difference”},$$

$$\widehat{v}^{(1)}_i = \left(1 - \frac{a_1}{n}\right) \widehat{v}^{(1)}_{i-1} - \frac{a_2}{n}\widehat{v}_{i-1} + \frac{a_2\overline{X}}{n} + \text{“innov. difference”}$$

28. Smoother 1

- $n = 1200$
- $m = 1100$

Smoother 1.

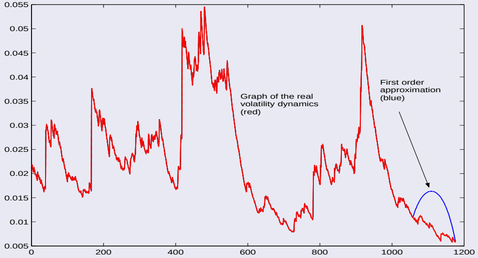
$$\widetilde{v}_i = \left(1 - \frac{a}{n}\right)\widetilde{v}_{i-1} + \frac{a\overline{X}}{n} + u_i, \ \widetilde{v}_m = \widehat{v}_m, \ i > m,$$

where u_i, $i \geq m$ is a control action chosen such

$$\widetilde{v}_n = \widehat{v}_n \quad \text{and} \quad \sum_{i=m+1}^{n} u_i^2 \text{ is minimal.}$$

29. Simulation. Smoother 1 is not successful

L. Goldentouch, F. Klebaner, R. Liptser



30. Smoother 2

L. Goldentouch, F. Klebaner, R. Liptser

- $n = 1200$
- $m = 1100$

Smoother 2. For $i > m$

$$\widetilde{v}_i = \widetilde{v}_{i-1} + \frac{1}{n}\widetilde{v}^{(1)}_{i-1} + u_i$$

$$\widetilde{v}^{(1)}_i = \left(1 - \frac{a_1}{n}\right)\widetilde{v}^{(1)}_{i-1} - \frac{a_2}{n}\widetilde{v}_{i-1} + \frac{a_2\overline{X}}{n},$$

where $\widetilde{v}_m = \widehat{v}_m$ and $\widetilde{v}^{(1)}_m = \widehat{v}^{(1)}_m$ and a control action chosen such

$$\widetilde{v}_n = \widehat{v}_n \quad \text{and} \quad \sum_{i=m+1}^{n} u_i^2 \text{ is minimal.}$$

31. Simulation. Smoother 2 is successful