PREVISIONE DI VOLATILITÀ: UN NUOVO APPROCCIO NON-PARAMETRICO

Cătălin Stărică[1], Roberto Ferulano[2]

[1] Department of Mathematical Statistics
Chalmers University of Technology, Gothenburg, Sweden.

[2] Dipartimento di Economia, Finanza e Statistica
Università degli Studi di Perugia
(e-mail: ferulano@stat.unipg.it)

ABSTRACT: Si introduce un nuovo modello storico non-parametrico per la previsione di volatilità e viene confrontato con il modello GARCH(1,1). La statistica utilizzata per il confronto dei modelli è il rapporto del MASE (Mean Aggregated Squared Error). Bande di confidenza sono state ottenute tramite simulazioni Monte Carlo. L'applicazione del modello alle serie storiche reali mostra che funziona molto bene per i tassi di cambio.

KEYWORDS: Previsione di volatilità, modelli GARCH, modelli non-parametrici.

1 Introduzione

La previsione di volatilità ha sempre giocato un ruolo significativo nell'ambito finanziario, ad esempio nel pricing delle opzioni.
La volatilità può essere stimata da dati storici (la volatilità storica o volatilità realizzata), o da opzioni plain-vanilla (volatilità implicita). Quest'ultima dovrebbe fornire maggiori indicazioni sulla previsione, in quanto si fonda sulle aspettative del mercato.
In questo articolo ci si concentra sulla stima della volatilità a partire dai dati storici. Si introduce una formula basata su un modello non-parametrico, e se ne confronta la qualità rispetto al modello GARCH(1,1).
L'articolo è organizzato come segue: dopo una breve digressione sulla notazione, nella sezione 2 si introduce il modello GARCH(1,1) e nella sezione 3 si introduce il nuovo modello. I risultati sono stati suddivisi in due sotto-sezioni: nella 4.2 si riportano le simulazioni, mentre nel 4.3 si riporta l'applicazione ai dati reali. Si riportano poi le conclusioni (5).

1.1 Notazione

Sia p_t, $(t = 1, \ldots, T)$, un serie storica dei prezzi definita su $(\Omega, \mathscr{F}_t, \mathbb{P})$, dove $\mathscr{F}_t$ è la filtrazione naturale associata a p_t. Sia $r_t = ln(p_t) - ln(p_{t-1})$ il rendimento continuamente composto sul titolo da $t-1$ a t. Si decompone r_t come

$$r_t = \mu_t + \sqrt{h_t}\varepsilon_t$$

dove μ_t è la media condizionata, h_t è la varianza condizionata e ε_t è un processo stocastico definito sulla stessa filtrazione $\mathscr{F}_t$.
Si definisce la media non condizionata $\mu = \mathbb{E}[r_t]$ e la varianza non condizionata $\sigma^2 = \mathbb{E}[(r_t-\mu)^2]$.

2 GARCH(1,1)

Questo modello è stato introdotto da Bollerslev, 1986. La varianza condizionata è espressa come

$$h_{t+1} = \alpha_0 + \alpha_1 (r_t - \mu_t)^2 + \beta_1 h_t$$

dove α_0, α_1 e β_1 sono positivi. Per soddisfare le ipotesi di stazionarietà, è necessario e sufficiente che $\alpha_1 + \beta_1 \leq 1$. Le innovazioni ε_t sono I.I.D. con media nulla e varianza unitaria.

La previsione della varianza a τ passi si ottiene ricorsivamente e può essere espressa in forma chiusa:

$$h_{t+\tau} = \sigma^2 + (\alpha_1 + \beta_1)^{\tau-1} \left(h_t - \sigma^2\right) \tag{1}$$

dove

$$\sigma^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}. \tag{2}$$

La previsione sul lungo periodo coincide con la varianza non condizionata del modello.
Si definisce la varianza aggregata di previsione a τ passi come:

$$h_{t:t+\tau} = \sum_{i=1}^{\tau} h_{t+i} = \tau\sigma^2 + \frac{(1-\alpha_1-\beta_1)^{\tau}}{1-(\alpha_1+\beta_1)} \left(h_t - \sigma^2\right)$$

3 Modello non-parametrico

Il modello GARCH(1,1) richiede per lo meno 2000 punti[*] per essere stimato correttamente e con bande di confidenza contenute. D'altra parte la formula di previsione della volatilità di un modello GARCH(1,1) (eq. 1) mostra una caratteristica interessante: è una media pesata di una varianza condizionata h_t (o di breve periodo) e una non condizionata (o di lungo periodo) σ^2. Nel nostro modello non-parametrico queste due quantità sono stimate direttamente dai dati. Si fissa invece $\alpha_1 + \beta_1$ (che chiamiamo ρ) ad un valore vicino a 0.9. La stima a τ passi diventa:

$$\sigma^2_{t+\tau} = \sigma_t^{2,long}\left(\rho^{\tau-1} - 1\right) + \rho^{\tau-1}\sigma_t^{2,short}$$

[*]Nel caso di rendimenti giornalieri, questo significa che sono richiesti 8 anni di dati.

dove $\sigma_t^{2,long}$ è ottenuta da un filtro MA, mentre $\sigma_t^{2,short}$ da un filtro esponenziale (EW-MA).

4 Risultati

4.1 Set-up dell'esperimento

Data una serie storica, una finestra mobile seleziona M punti della serie per la stima del modello e si sposta ogni volta di una quantità Δ. Dalla stima dei parametri si ottiene la varianza aggregata di previsione per τ passi, che viene confrontata con la varianza media realizzata $(\bar{r}^2_{t:t+\tau})$. Se la serie è composta di N punti, si ottengono quindi $n = \frac{N-\tau-m}{\Delta}$ errori di previsione. Si calcola quindi il MASE:

$$MASE(\tau) := \frac{1}{n}\sum_{t=1}^{n}\left(\bar{r}^2_{t:t+\tau} - h_{t:t+\tau}\right)^2 \tag{3}$$

La statistica per il confronto tra due modelli è il rapporto tra due MASE: se il rapporto è maggiore di 1 significa che il modello al denominatore è migliore, e viceversa.

4.2 Simulazioni

Sono stati simulate 250 traiettorie di 7500 punti (N) secondo un modello GARCH(1,1). La finestra mobile per la stima del modello seleziona 2000 punti (M), e viene spostata di 10 punti (Δ). La varianza di previsione viene calcolata fino a 500 punti (τ) in avanti. Si ottengono 500 errori (n) di previsione. Simulando 250 traiettorie si è ottenuto una distribuzione Monte Carlo della statistica. Le simulazioni sono state realizzate con nove differenti terne di parametri (in tabella 1 si riportano solo le prime 2).

	σ	α_1	β_1	$\alpha_1+\beta_1$
1	10%	0.05	0.90	0.95
2	22%	0.08	0.90	0.98

Tabella 1. *Terne di parametri GARCH(1,1) scelti per le simulazioni*

In tabella 2 si riporta per ogni terna di parametri e di orizzonte di previsione le bande di confidenza della distribuzione al 90% del rapporto dei MASE.

I risultati che si ottengono sono:

- come era prevedibile, per orizzonti brevi, tre volte le bande di confidenza sono inferiori a 1;
- per orizzonti lunghi (più di 250 giorni) le bande di confidenza contengono sempre il valore 1. Quindi, il nostro modello fornisce valori di previsione confrontabili con quelli del GARCH(1,1).

days	5	21	63	125	250	500
1	0.92 ÷ 0.98	0.88 ÷ 0.97	0.85 ÷ 0.98	0.76 ÷ 0.98	0.61 ÷ 1.00	**0.44** ÷ 1.03
2	0.85 ÷ 1.01	0.79 ÷ 0.99	0.75 ÷ 1.00	0.73 ÷ 1.06	0.63 ÷ 1.12	0.48 ÷ 1.16

Tabella 2. $MASE^{GARCH(1,1)}/MASE^{Model}$. *Il modello GARCH(1,1) è ristimato su una finestra mobile di 2000 punti con passo 10.*

4.3 Serie reali

Il modello è stato applicato su 75 serie storiche reali, di cui 36 sono titoli azionari, 10 sono indici, 15 sono tassi di interesse e 14 tassi di cambio. Facendo uso delle bande di confidenza ottenute precedentemente, si ottengono i seguenti risultati:

- il nostro modello fornisce stime migliori su qualunque orizzonte solo per i tassi di cambio;
- per i titoli azionari il nostro modello è equivalente ai modelli ARCH;
- per i tassi di interesse e gli indici azionari il nostro modello è migliore principalmente per brevi orizzonti.

5 Conclusioni

Il modello proposto risulta fornire delle previsioni di volatilità confrontabili con quelle dei modelli GARCH(1,1).
L'analisi è stata condotta anche per i modelli GJR-GARCH(1,1,1) e si sono ottenuti gli stessi risultati.
Il nostro modello risulta essere ancora migliore se la stima dei modelli ARCH avviene su 1000 e 500 punti. Inoltre i risultati sono robusti anche al cambio di statistica: al posto del MASE, si può usare il MAAE (Mean Aggregated Absolute Error).
Il motivo per cui funziona bene anche un modello non-parametrico è che le serie finanziarie risultano non stazionarie e quindi i modelli di tipo GARCH, senza cambio di regime, tendono a mostrare effetti di Integrated GARCH. Sarebbe interessante vedere cosa succede dal confronto con questi modelli più sofisticati.

References

BOLLERSLEV, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.