Estimating Volatilities and Correlations

Following
Options, Futures, and Other Derivatives, 5th edition
by John C. Hull

Chapter 17

Daniel HERLEMONT

Standard Approach to Estimating Volatility

- **Define σ_n as the volatility per day between day n-1 and day n, as estimated at end of day n-1**
- **Define S_i as the value of market variable at end of day i**
- **Define $u_i = \ln(S_i/S_{i-1})$**

$$\sigma_n^2 = \frac{1}{m-1}\sum_{i=1}^{m}(u_{n-i} - \bar{u})^2$$

$$\bar{u} = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}$$

Daniel HERLEMONT

Simplifications Usually Made

Define u_i as $(S_i - S_{i-1})/S_{i-1}$

Assume that the mean value of u_i is zero

Replace m-1 by m

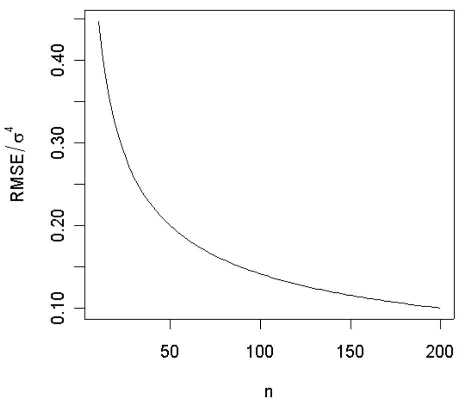
This gives

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}^2$$

Statistical properties

$$ERQM = E\left[(\hat{\sigma}_n^2 - \sigma^2)^2\right] = \frac{2}{n}\sigma^4$$

If prices follow a lognormal process



Example: with 50 historical price, the relative error on variance estimate is still about 20%

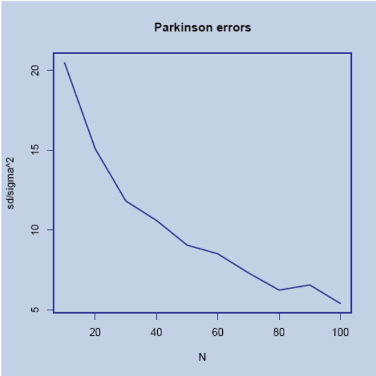
Efficient estimators with Highs and Lows

Parkinson $$\sigma^2_{Parkinson} = \frac{1}{N*4*\log(2)} \sum (h_k - l_k)^2$$

Rogers Satchell $$\sigma^2_{RS} = \frac{1}{N} \sum_{k=1,N} (h_k - o_k)(h_k - c_k) + (l_k - o_k)(l_k - c_k)$$

Garman Klass $$h = \log(high)\ l = \log(low)\ o = \log(open)\ c = \log(close)$$

Exemple: with n=50, relative error is about 10%,
to be compared with the 20% error



Volatility based on Open/Close/High/Low

Source: Malik Magdon-Ismail, Amir F. Atiya
Volatility Estimation Using High, Low, and Close Data

Days used	RMS Prediction Error							
	μ known				μ unknown		$\mu = 0$	
	Close	Park	R–S	ML	Close	ML	G–K	ML
5	0.1597	0.0713	0.0642	0.0621	0.1752	0.0639	0.0591	0.0640
10	0.1090	0.0489	0.0448	0.0426	0.1152	0.0434	0.0399	0.0417
15	0.0900	0.0410	0.0375	0.0353	0.0930	0.0354	0.0337	0.0346
20	0.0781	0.0360	0.0317	0.0303	0.0808	0.0307	0.0292	0.0304
25	0.0702	0.0317	0.0289	0.0273	0.0709	0.0270	0.0260	0.0271
30	0.0645	0.0292	0.0270	0.0246	0.0654	0.0248	0.0233	0.0245
35	0.0605	0.0272	0.0245	0.0230	0.0615	0.0229	0.0224	0.0232
40	0.0556	0.0252	0.0227	0.0215	0.0559	0.0215	0.0205	0.0215
45	0.0526	0.0238	0.0215	0.0200	0.0534	0.0202	0.0192	0.0196
50	0.0499	0.0222	0.0204	0.0192	0.0505	0.0191	0.0186	0.0191

$$\hat{\sigma}_{close} = \sqrt{\frac{1}{NT}\sum_{i=1}^{N}(c_i - o_i - \mu T)^2}$$

$$\sigma_{GK} = \sqrt{\frac{1}{NT}\sum_{i=1}^{N} 0.511(\tilde{h}_i - \tilde{l}_i)^2 - 0.019(\tilde{c}_i(\tilde{h}_i + \tilde{l}_i) - 2\tilde{l}_i\tilde{h}_i) - 0.383\tilde{c}_i^2}$$

$$\hat{\sigma}_{Park} = \sqrt{\frac{1}{4NT\ln 2}\sum_{i=1}^{N}(h_i - l_i)^2}$$

$$\sigma_{RS} = \sqrt{\frac{1}{NT}\sum_{i=1}^{N}(h_i - o_i)(h_i - c_i) + (l_i - o_i)(l_i - c_i)}$$

Difficulty in evaluating and comparing volatility models is due to the fact that volatility is not directly observable.

An approach to solve this particular problem is to compare the volatility forecasts, with squared returns

Weighting Scheme

Instead of assigning equal weights to the observations we can set

$$\sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

where

$$\sum_{i=1}^{m} \alpha_i = 1$$

ARCH(m) Model

In an ARCH(m) model we also assign some weight to the long-run variance rate, V_L:

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

where

$$\gamma + \sum_{i=1}^{m} \alpha_i = 1$$

Daniel HERLEMONT

EWMA Model

In an exponentially weighted moving average model, the weights assigned to the u^2 decline exponentially as we move back through time

This leads to

$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1-\lambda) u_{n-1}^2$$

Daniel HERLEMONT

Attractions of EWMA

Relatively little data needs to be stored

We need only remember the current estimate of the variance rate and the most recent observation on the market variable

Tracks volatility changes

RiskMetrics uses $\lambda = 0.94$ for daily volatility forecasting

GARCH (1,1)

In GARCH (1,1) we assign some weight to the long-run average variance rate

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

Since weights must sum to 1

$$\gamma + \alpha + \beta = 1$$

GARCH (1,1) continued

Setting $\omega = \gamma V$ the GARCH (1,1) model is

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

and

$$V_L = \frac{\omega}{1 - \alpha - \beta}$$

Daniel HERLEMONT

Example

Suppose

$$\sigma_n^2 = 0.000002 + 0.13 u_{n-1}^2 + 0.86 \sigma_{n-1}^2$$

The long-run variance rate is 0.0002 so that the long-run volatility per day is 1.4%

Daniel HERLEMONT

Example continued

Suppose that the current estimate of the volatility is 1.6% per day and the most recent percentage change in the market variable is 1%.

The new variance rate is

$$0.000002 + 0.13 \times 0.0001 + 0.86 \times 0.000256 = 0.00023336$$

The new volatility is 1.53% per day

GARCH (p,q)

$$\sigma_n^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{n-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{n-j}^2$$

Other Models

We can design GARCH models so that the weight given to u_i^2 depends on whether u_i is positive or negative

We do not have to assume that the conditional distribution is normal (e.g. Student residuals)

Variance Targeting

One way of implementing GARCH(1,1) that increases stability is by using variance targeting

We set the long-run average volatility equal to the sample variance

Only two other parameters then have to be estimated

Maximum Likelihood Methods

In maximum likelihood methods we choose parameters that maximize the likelihood of the observations occurring

Example 1

We observe that a certain event happens one time in ten trials. What is our estimate of the proportion of the time, *p*, that it happens?

The probability of the outcome is

$$10p(1-p)^9$$

We maximize this to obtain a maximum likelihood estimate: *p*=0.1

Example 2

Estimate the variance of observations from a normal distribution with mean zero

Maximize: $$\prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi v}}\exp\left(\frac{-u_i^2}{2v}\right)\right]$$

or: $$\sum_{i=1}^{n}\left[-\ln(v)-\frac{u_i^2}{v}\right]$$

This gives: $$v=\frac{1}{n}\sum_{i=1}^{n}u_i^2$$

Daniel HERLEMONT

Application to GARCH

We choose parameters that maximize

$$\sum_{i=1}^{n}\left[-\ln(v_i)-\frac{u_i^2}{v_i}\right]$$

Daniel HERLEMONT

How Good is the Model?

The Ljung-Box statistic tests for autocorrelation

We compare the autocorrelation of the u_i^2 with the autocorrelation of the u_i^2/σ_i^2

Is σ_i^2 a predictor of u_{i+1}^2 ?

Perform a regression of σ_i^2 versus u_{i+1}^2

Daniel HERLEMONT

Forecasting Future Volatility

A few lines of algebra shows that

$$E[\sigma_{n+k}^2] = V_L + (\alpha+\beta)^k(\sigma_n^2 - V_L)$$

The variance rate for an option expiring on day *m* is

$$\frac{1}{m}\sum_{k=0}^{m-1} E\left[\sigma_{n+k}^2\right]$$

Daniel HERLEMONT

Correlations

Define $u_i=(U_i-U_{i-1})/U_{i-1}$ and $v_i=(V_i-V_{i-1})/V_{i-1}$

Also

$\sigma_{u,n}$: daily vol of U calculated on day n-1

$\sigma_{v,n}$: daily vol of V calculated on day n-1

cov_n: covariance calculated on day n-1

The correlation is $\text{cov}_n/(\sigma_{u,n}\,\sigma_{v,n})$

Daniel HERLEMONT

Correlations continued

Under EWMA

$$\text{cov}_n = (1-\beta)\, u_{n-1}v_{n-1}+\beta\, \text{cov}_{n-1}$$

Under GARCH (1,1)

$$\text{cov}_n = \omega + \alpha\, u_{n-1}v_{n-1}+\beta\, \text{cov}_{n-1}$$

Daniel HERLEMONT