Financial Time Series

MFM Practitioner Module: Quantitative Risk Management

John Dodson

September 26, 2018
Outline

Invariants & Innovations
  Indexes, Generics, & Synthetics
  Investibility & Relevance
  Innovations

Autoregressive Moving Average

Generalized Autoregressive Conditional Heteroskedasticity
  Estimation
  Forecasting
Invariants

We are generally working with financial timeseries data when calibrating models for the future value of financial variables such as the mark-to-market profit/loss on an asset holding.

- In some cases, such as equity shares, this may mean working with market prices (adjusted for dividends and splits).

- In other cases, such as for bonds or derivatives, it may mean working with derived quantities like yield or implied volatility.

Invariants

If we expect today that the meaning of a financial quantity of interest will remain uniform for the foreseeable future, we term it an invariant quantity. For example, the price or yield on a particular derivative or bond is not an invariant because the instrument will expire or mature on a known date.
Invariants

Indexes, Generics, & Synthetics

The challenge of identifying invariants for important classes of financial variables is addressed variously through indexes, synthetics, and generics.

- The S&P 500 equity index purports to represent the performance of typical large-cap U. S. listed equity securities.
- The Fed’s CMT indexes purport to represent the performances of typical nominal U. S. Treasury bonds of particular tenors.
- Bloomberg futures generics represent the performances of the 1st, 2nd, etc. contract of a particular futures product.
- The Cboe VIX index purports to represent the performance of a delta-hedged position in one-month S&P 500 index options.
Investibility & Relevance of Invariants

If we intend to use an invariant as a proxy for an actual asset, it is important to think carefully about how the performance of the proxy can differ from the performance of the asset.

▶ The S&P 500 is an investible index whose performance can be replicated by an instantaneously fixed portfolio of equity shares, its performance is influenced by its dynamic composition and the dynamic correlation between constituents, which is obviously not relevant for individual equities.

▶ Other indexes, such as LIBOR (London interbank offered rate) or OIS (Federal Funds rate overnight index swap), are technically investible, but only by the treasury departments of banks; in particular they are not investible to broker-dealers.
Seasonality

Some financial timeseries exhibit predictable patterns in time, or *seasonality*.

- A futures generic must **roll** whenever new contracts are issued. The actual profit/loss from rolling over a futures position is difficult to predict, and the generic makes no attempt at all.
  - For timeseries analysis purposes, you should omit roll dates from generics for your analysis.

- There may be predictable events, such as earnings announcements or the seasonal consumption patterns of certain commodities, that should be modeled as **regimes**.
  - This is a specialized topic in **econometrics** that we will not cover.
Innovations

We generally only care about the most recent level for a risk factor after our timeseries analysis is finished and we are looking at the loss distribution for a particular portfolio. For the analysis, we are more interested in the periodic innovations of the risk factor, such as the log-returns or simple differences.

You can think of this as the difference operator applied to the index or its (natural) logarithm, $X_t \triangleq \nabla \log S_t$.

Drift
The conditional expected value $E \left[ \nabla \log S_t | \mathcal{F}_{t-1} \right]$ of the log of an index is termed the index drift $\mu_t$.

Volatility
The conditional standard deviation $\sqrt{\text{var} \left[ \nabla \log S_t | \mathcal{F}_{t-1} \right]}$ is termed the index volatility $\sigma_t$.

Note that the drift and volatility are $\mathcal{F}_{t-1}$-measurable.
Drift with ARMA

White noise is a collection of i.i.d. r.v.’s $Z_t$ with zero mean and finite variance $\sigma_t^2$. With $X_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t Z_t$ an innovation of an invariant, we call $\varepsilon_t$ the residual.

**Autoregressive Moving Average**

An ARMA($p$, $q$) process for the drift can be expressed as

$$
\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}
$$

for parameters $\phi_0, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$.

- AR(1) is a simple model for mean reversion for the innovations around a long-run level $\phi_0/(1 - \phi_1)$. 
For financial data there is little to be gained in modeling drifts of timeseries data, because typically $|X_t| \gg \mu_t$.

- Furthermore, if $X_t$ is a log-return, the drift probably ought to include a Jensen term like $-\frac{1}{2} \sigma^2_t$ which certainly does not fit into the ARMA form.

**Generalized Autoregressive Conditional Heteroskedasticity**

A GARCH($p$, $q$) process for the conditional variance can be expressed as

$$\sigma^2_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon^2_{t-i} + \sum_{j=1}^{q} \beta_j \sigma^2_{t-j}$$

for non-negative parameters $\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$. 
Volatility with GARCH

Standardized Residuals
One application of GARCH models is to extract i.i.d. samples from timeseries. We define the standardized residuals as

\[ Z_t = \frac{X_t - \mu_t}{\sigma_t} \]

To the extent that the GARCH model is correct, these are strict white noise.

GARCH(1,1)
By far the most common implementation of this model is GARCH(1,1). An important result about this model is that the unconditional variance is

\[ \sigma^2 \triangleq \text{var} [X_t] = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \]

as long as \( \alpha_1 + \beta_1 < 1 \).
Volatility with GARCH

In fitting GARCH(1,1) to asset returns, one often sees that \( \hat{\alpha}_0 \) is close to zero. From the previous slide, we see that \( \alpha_0 = 0 \) requires that \( \alpha_1 = 1 - \beta_1 \). So the integrated GARCH(1,1) model has only one parameter.

\[
\sigma_t^2 = (1 - \beta_1) \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\]

Exponentially-weighted moving average

Since (with \( \beta_1 = \lambda \)) this is equivalent to

\[
\sigma_t^2 = \frac{\sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-1-i}^2}{\sum_{i=0}^{\infty} \lambda^i}
\]

IGARCH(1,1) is sometimes called the exponentially-weighted moving average (EWMA) model, popularized by RiskMetrics.

\( \lambda \) can be estimated using the technique discussed below.
Volatility with GARCH

GARCH(1,1) Volatility Process

Under GARCH(1,1), *innovations* of the conditional variance are mean-reverting. You can see this because

\[ \nabla \sigma^2_t = (1 - \beta_1 - \alpha_1) (\sigma^2 - \sigma^2_{t-1}) + \alpha_1 \sigma^2_{t-1} \nabla W_{t-1} \]

where

\[ \nabla W_{t-1} = \frac{\varepsilon^2_{t-1}}{\sigma^2_{t-1}} - 1 \]

is stochastic increment uncorrelated to \( \varepsilon_{t-1} \) (if the residuals are unskewed).

- The mean-reversion rate is \( 1 - \beta_1 - \alpha_1 \).
- \( W_t \) is not a Brownian motion but rather a more general martingale.
Estimating GARCH

Let us continue to focus on GARCH(1,1). The principal technique for estimating the parameters of a GARCH process is maximum likelihood, but with several caveats:

- We do not know the marginal densities of the residuals and they are not identical
- We do not know $\varepsilon_0$ or $\sigma_0$ (assume $t = 1$ is the first observed innovation)
- While we may assume that they are i.i.d., we may not know the exact density of the standardized residual $f_Z(\cdot)$

We address these through the quasi-MLE, in which we note that the multivariate density is the product of the conditional densities, and we assume that the residuals are normal:

$$
\log f_{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n | \varepsilon_0, \sigma_0}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = -\frac{1}{2} \sum_{t=1}^{n} \log (2\pi\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}
$$
Estimating GARCH

Variance Targeting

Assuming that the unconditional variance of the innovations exists, it is advisable to set the intercept based on the sample variance.

\[ \alpha_0 = \hat{\sigma}^2 (1 - \alpha_1 - \beta_1) \]

Then you are only using the QMLE to estimate \( \alpha_1 \) and \( \beta_1 \).

N.B.: You should probably put a lower bound on \( \alpha_1 \) in this case, otherwise the \( \beta_1 \) could be degenerate.

Initialization

Assuming \( t = 1 \) is your first innovation, we need a way of determining \( \sigma_1^2 \) in terms of the parameters. That means you need to choose values for \( \varepsilon_0^2 \) and \( \sigma_0^2 \). One choice is to take both to be \( \sigma^2 \). In combination with variance targeting, this means \( \sigma_1^2 = \hat{\sigma}^2 \).
Forecasting GARCH

In terms of forecasting, we already have $\sigma^2_{n+1}$. Say we are interested in $E_n [\sigma^2_{n+2}]$ (the subscript on the expectation represents the sigma-algebra), we can write

$$\sigma^2_{n+2} = \sigma^2 (1 - \alpha_1 - \beta_1) + \sigma^2_{n+1} (\alpha_1 Z^2_{n+1} + \beta_1)$$

so because $Z_{n+1} \sim \text{SWN}(0, 1)$

$$E_n [\sigma^2_{n+2}] = \sigma^2_{n+1} (\alpha_1 + \beta_1) + \sigma^2 (1 - \alpha_1 - \beta_1)$$

Iterating this, we get the general result for integer $m > 0$,

$$E_n [\sigma^2_{n+m}] = \sigma^2_{n+1} (\alpha_1 + \beta_1)^{m-1} + \sigma^2 \left(1 - (\alpha_1 + \beta_1)^{m-1}\right)$$

- The forecasts are a convex combination of the current conditional variance $\sigma^2_{n+1}$ and the unconditional, or long-run, variance $\sigma^2$. 