Mathematical Foundations of the AR(2)-GARCH(1,1) Model

May 2025

Introduction to the AR(2)-GARCH(1,1) Model

Before exploring the details of autoregressive (AR) and generalized autoregressive conditional heteroskedasticity (GARCH) models, it's important to understand the AR(2)-GARCH(1,1) model as a whole. This model is used in finance to study time series data, such as daily stock returns, by modeling both the average value (mean) and the fluctuations (volatility) of the series. It combines two parts:

- **AR(2) Component**: This part models the mean of the time series, predicting the current value based on the two previous values. It's called "AR(2)" because it uses two lags (past values).
- **GARCH(1,1) Component**: This part models the volatility, describing how much the time series fluctuates over time. It's called "GARCH(1,1)" because it uses one lag of past errors and one lag of past variances.

The full AR(2)-GARCH(1,1) model can be written as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t,$$

$$\epsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1),$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where: - y_t: The time series value (e.g., stock return) at time t. - c, ϕ_1, ϕ_2: Parameters for the mean (AR(2) part). - ϵ_t: The random error, scaled by volatility σ_t. - σ_t^2: The conditional variance (GARCH(1,1) part), measuring volatility. - ω, α, β: Parameters for the variance.

Why AR(2)-GARCH(1,1) and Not Just AR(2)? The AR(2) model alone predicts the average return but assumes constant volatility, which is unrealistic for financial data where volatility changes (e.g., during market crashes). The GARCH(1,1) part adds a model for volatility, allowing it to vary based on past shocks and volatility levels. Combining them provides a complete description of both the mean and volatility, making the model more accurate for financial applications.

How Do AR and GARCH Relate? The AR(2) part produces the expected return ($E[y_t]$) and the error ($\epsilon_t = y_t - E[y_t]$), which feeds into the GARCH(1,1) part to calculate volatility (σ_t^2). This interaction ensures the model accounts for both predictable trends and changing risk levels.

Remarks: This model is widely used because it describes both the average behavior and the riskiness of financial time series, like stock returns, which are like daily weather patterns—sometimes calm, sometimes stormy. The following sections explain each component in detail, starting with AR models, then ARCH and GARCH, and how they fit together. By the end, you'll see how these pieces form the AR(2)-GARCH(1,1) framework. Statistical software can estimate the parameters, but the math applies universally.

Autoregressive (AR) Models

Definition 1. An autoregressive model of order p (AR(p)) expresses a time series y_t as a linear combination of its previous p values plus a random error:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where c is a constant, ϕ_i are coefficients, and ϵ_t is a white noise error with mean 0 and variance σ^2.

The AR(2) model, used in AR(2)-GARCH(1,1), is:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t,$$

where $\epsilon_t = \sigma_t z_t$, and $z_t \sim N(0,1)$. This models the mean of the time series, while GARCH(1,1) handles the volatility of ϵ_t.

Mathematical Details: For the AR(2) model to be stable (stationary), the series must not grow or shrink indefinitely. This requires the roots of the characteristic equation:

$$1 - \phi_1 z - \phi_2 z^2 = 0$$

to lie outside the unit circle ($|z| > 1$). Stationarity ensures a constant mean and variance, so the expected return is:

$$E[y_t] = c + \phi_1 E[y_{t-1}] + \phi_2 E[y_{t-2}].$$

For a stationary series, $E[y_t] = E[y_{t-1}] = E[y_{t-2}] = \mu$, so:

$$\mu = c + \phi_1 \mu + \phi_2 \mu \implies \mu(1 - \phi_1 - \phi_2) = c \implies \mu = \frac{c}{1 - \phi_1 - \phi_2}.$$

This mean is finite if $\phi_1 + \phi_2 < 1$.

Common Questions: - **Why use two lags?** Two lags balance simplicity and the ability to model short-term trends or corrections, common in financial data like stock returns. - **Why is AR(2) part of AR(2)-GARCH(1,1)?** AR(2) predicts the mean return, while GARCH(1,1) models the volatility of the errors, together providing a complete model.

Remarks: The AR(2) model predicts the average return based on recent history, like forecasting today's weather using the past two days' conditions. It's the mean component of AR(2)-GARCH(1,1), feeding errors (ϵ_t) into the GARCH part for volatility modeling. This makes it useful for traders forecasting short-term price movements. The next sections build on this by introducing volatility models, showing how they combine with AR(2).

Example 1. Consider an AR(2) model with $c = 0.002$, $\phi_1 = 0.4$, $\phi_2 = -0.2$. If yesterday's return was $y_{t-1} = 0.03$ (3%) and two days ago was $y_{t-2} = 0.01$ (1%):

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t.$$

The expected return is:

$$E[y_t] = c + \phi_1 y_{t-1} + \phi_2 y_{t-2}.$$

Calculate:

- $\phi_1 y_{t-1} = 0.4 \cdot 0.03 = 0.012$ **(1.2%).**
- $\phi_2 y_{t-2} = -0.2 \cdot 0.01 = -0.002$ **(-0.2%).**
- $E[y_t] = 0.002 + 0.012 - 0.002 = 0.012$ **(1.2%).**

Remarks: The 1.2% expected return shows momentum from yesterday's gain, offset by a slight reversal from two days ago. This helps traders anticipate a positive return, though the random ϵ_t adds uncertainty (modeled by GARCH). This example illustrates how AR(2) predicts the mean in the AR(2)-GARCH(1,1) framework, setting the stage for volatility modeling.

Example 2. For a stock with $c = 0.001$, $\phi_1 = 0.3$, $\phi_2 = 0.1$, and returns $y_{t-2} = -0.02$ (-2%), $y_{t-1} = 0.015$ (1.5%):

$$E[y_t] = 0.001 + 0.3 \cdot 0.015 + 0.1 \cdot (-0.02).$$

- $0.3 \cdot 0.015 = 0.0045$ **(0.45%).**
- $0.1 \cdot (-0.02) = -0.002$ **(-0.2%).**
- $E[y_t] = 0.001 + 0.0045 - 0.002 = 0.0035$ **(0.35%).**

Remarks: The small 0.35% return suggests a stable day, as past gains and losses nearly cancel out. This helps investors adjust portfolios, expecting modest returns. The AR(2) model's role in predicting the mean is clear, and its errors will feed into the GARCH(1,1) model to estimate volatility, completing the AR(2)-GARCH(1,1) structure.

ARCH Models

Definition 2. An autoregressive conditional heteroskedasticity model of order q (ARCH(q)) models the conditional variance as:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_q \epsilon_{t-q}^2,$$

where $\omega > 0$, $\alpha_i \geq 0$, and $\epsilon_t = y_t - E[y_t]$ is the error term from the mean model (e.g., AR(2)).

The ARCH(1) model is:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2.$$

Equation Breakdown:

- σ_t^2: The variance at time t, measuring how much returns fluctuate. A high σ_t^2 means larger, riskier price swings, like a stormy market day.
- ω: A positive constant, the minimum variance when past errors are zero. It's like the baseline risk level on a calm day.
- $\alpha_1 \epsilon_{t-1}^2$:
  - $\epsilon_{t-1} = y_{t-1} - E[y_{t-1}]$: The error from the AR(2) model, representing unexpected price changes.
  - ϵ_{t-1}^2: Squaring emphasizes large shocks (e.g., a 5% unexpected drop).
  - α_1: Measures how much yesterday's shock increases today's volatility. If $\alpha_1 = 0.4$, a large error raises risk significantly.

Mathematical Details: The unconditional (average) variance is:

$$E[\sigma_t^2] = E[\omega + \alpha_1 \epsilon_{t-1}^2] = \omega + \alpha_1 E[\epsilon_{t-1}^2].$$

Since $\epsilon_{t-1} = \sigma_{t-1} z_{t-1}$, and $E[z_{t-1}^2] = 1$, we have:

$$E[\epsilon_{t-1}^2] = E[\sigma_{t-1}^2].$$

For a stationary series, $E[\sigma_t^2] = E[\sigma_{t-1}^2]$, so:

$$E[\sigma_t^2] = \omega + \alpha_1 E[\sigma_t^2] \implies E[\sigma_t^2](1 - \alpha_1) = \omega \implies E[\sigma_t^2] = \frac{\omega}{1 - \alpha_1}.$$

This requires $\alpha_1 < 1$ to keep variance finite. The ARCH model describes volatility clustering, where large price movements increase future risk.

Common Questions: - **Why square errors?** Squaring ensures positive contributions to variance and highlights large shocks, which drive financial volatility. - **How does ARCH fit into AR(2)-GARCH(1,1)?** ARCH models the volatility of AR(2) errors, but GARCH extends this by including past variances, as we'll see next. - **What if $\alpha_1 \geq 1$?** Volatility becomes infinite, making the model unstable. Typically, α_1 is small (e.g., 0.2–0.4).

Remarks: The ARCH model shows how volatility changes, unlike AR(2) alone, which assumes constant volatility. It's a stepping stone to GARCH, used in AR(2)-GARCH(1,1) to model the variance of ϵ_t. This is vital for risk management, like estimating potential losses in a stock portfolio. The next section introduces GARCH, which improves on ARCH for the full model.

Example 3. For a stock with $\omega = 0.0001$, $\alpha_1 = 0.2$, and $\epsilon_{t-1} = 0.01$:

$$\sigma_t^2 = 0.0001 + 0.2 \cdot (0.01)^2.$$

- $(0.01)^2 = 0.0001$.
- $0.2 \cdot 0.0001 = 0.00002$.
- $\sigma_t^2 = 0.0001 + 0.00002 = 0.00012$.
- $\sigma_t = \sqrt{0.00012} \approx 0.011$ **(1.1%).**

Remarks: The 1.1% volatility suggests a calm market, useful for pricing options with lower risk. This example shows how ARCH models the volatility of AR(2) errors, a crucial step in the AR(2)-GARCH(1,1) framework, preparing for the more advanced GARCH model.

GARCH Models

Definition 3. A generalized autoregressive conditional heteroskedasticity model of order p, q (GARCH(p,q)) models the conditional variance as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where $\omega > 0$, $\alpha_i, \beta_j \geq 0$.

The GARCH(1,1) model, used in AR(2)-GARCH(1,1), is:

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

Equation Breakdown:

- σ_t^2: The variance at time t, measuring riskiness of returns.
- ω: The baseline variance, ensuring volatility stays positive.
- $\alpha\epsilon_{t-1}^2$: The effect of the previous error (from AR(2)), where α shows how shocks increase volatility.
- $\beta\sigma_{t-1}^2$: The effect of past volatility, where β measures how long high volatility lasts. A high $\beta = 0.8$ means volatility persists, like lingering storm clouds.

Mathematical Details: The unconditional variance is:

$$E[\sigma_t^2] = \omega + \alpha E[\epsilon_{t-1}^2] + \beta E[\sigma_{t-1}^2].$$

Since $E[\epsilon_{t-1}^2] = E[\sigma_{t-1}^2]$ (as $\epsilon_{t-1} = \sigma_{t-1} z_{t-1}$), and $E[\sigma_t^2] = E[\sigma_{t-1}^2]$:

$$E[\sigma_t^2] = \omega + \alpha E[\sigma_t^2] + \beta E[\sigma_t^2] \implies E[\sigma_t^2](1 - \alpha - \beta) = \omega \implies E[\sigma_t^2] = \frac{\omega}{1 - \alpha - \beta}.$$

This requires $\alpha + \beta < 1$ for stationarity, ensuring volatility doesn't grow indefinitely.

Common Questions: - **Why add past variance?** It models persistence, where a volatile day increases future risk, unlike ARCH, which only uses errors. - **How does GARCH fit AR(2)-GARCH(1,1)?** GARCH(1,1) models the volatility of AR(2) errors, completing the model by describing both mean and variance.

Remarks: GARCH(1,1) improves on ARCH by including past volatility, making it more efficient for modeling persistent risk in AR(2)-GARCH(1,1). It's used for option pricing and risk management, as it describes how volatility evolves, like weather patterns shifting over time. This completes the volatility component of the AR(2)-GARCH(1,1) model introduced earlier.

Example 4. For a GARCH(1,1) model with $\omega = 0.00003$, $\alpha = 0.15$, $\beta = 0.8$, if $\epsilon_{t-1} = 0.04$, $\sigma_{t-1}^2 = 0.0005$:

$$\sigma_t^2 = 0.00003 + 0.15 \cdot (0.04)^2 + 0.8 \cdot 0.0005.$$

- $(0.04)^2 = 0.0016$.
- $0.15 \cdot 0.0016 = 0.00024$.
- $0.8 \cdot 0.0005 = 0.0004$.
- $\sigma_t^2 = 0.00003 + 0.00024 + 0.0004 = 0.00067$.
- $\sigma_t = \sqrt{0.00067} \approx 0.0259$ **(2.59%).**

Annualized: $\sqrt{252} \cdot 0.0259 \approx 0.411$ **(41.1%).**

Remarks: The 2.59% daily volatility (41.1% annualized) shows high risk after a shock, useful for setting option prices.

Example 5. Using the same model, if $\epsilon_t = 0.01$, $\sigma_t^2 = 0.00067$:

$$\sigma_{t+1}^2 = 0.00003 + 0.15 \cdot (0.01)^2 + 0.8 \cdot 0.00067.$$

- $(0.01)^2 = 0.0001$.
- $0.15 \cdot 0.0001 = 0.000015$.
- $0.8 \cdot 0.00067 = 0.000536$.
- $\sigma_{t+1}^2 = 0.00003 + 0.000015 + 0.000536 = 0.000581$.
- $\sigma_{t+1} = \sqrt{0.000581} \approx 0.0241$ **(2.41%)**.

Remarks: The drop to 2.41% shows how smaller shocks reduce volatility, a feature of GARCH(1,1) in AR(2)-GARCH(1,1). This helps risk managers adjust strategies, like updating loss estimates. The model's ability to describe changing volatility makes it essential for financial forecasting.

Maximum Likelihood Estimation (MLE)

Definition 4. Maximum likelihood estimation (MLE) estimates parameters θ by maximizing the likelihood function $L(\theta|\text{data})$, or the log-likelihood $\ell(\theta) = \log L(\theta)$.

For AR(2)-GARCH(1,1) with Gaussian errors ($z_t \sim N(0,1)$), the log-likelihood is:

$$\ell(\theta) = -\frac{1}{2} \sum_{t=3}^{T} \left(\log(2\pi) + \log(\sigma_t^2) + \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2})^2}{\sigma_t^2} \right).$$

Equation Breakdown:

- $\sum_{t=3}^{T}$: Sums over time, starting at $t = 3$ because AR(2) needs two past values.
- $\log(2\pi)$: A constant from the Gaussian distribution.
- $\log(\sigma_t^2)$: Accounts for volatility's effect on probability. Higher volatility widens the distribution.
- $(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2})^2$: The squared error ($\epsilon_t^2$), measuring how far the actual return is from the AR(2) prediction.
- $\frac{\epsilon_t^2}{\sigma_t^2}$: Scales the error by volatility, showing how extreme the deviation is relative to risk.

Mathematical Details: The Gaussian density is:

$$f(y_t|y_{t-1}, y_{t-2}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - \mu_t)^2}{2\sigma_t^2}\right),$$

where $\mu_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2}$. The log-likelihood sums the log of this density:

$$\ell(\theta) = \sum_{t=3}^{T} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{(y_t - \mu_t)^2}{2\sigma_t^2}\right].$$

MLE finds $\theta = (c, \phi_1, \phi_2, \omega, \alpha, \beta)$ by maximizing $\ell(\theta)$, typically using numerical methods.

Common Questions: - **Why log-likelihood?** Logs simplify calculations by turning products into sums, avoiding numerical issues. - **How does MLE fit AR(2)-GARCH(1,1)?** It estimates all parameters (AR and GARCH) simultaneously, ensuring the model fits both mean and volatility.

Remarks: MLE ensures the AR(2)-GARCH(1,1) model accurately describes the data by finding the best parameters for both mean and volatility. It's a critical step in applying the model, used in financial forecasting to predict returns and risks. Statistical software can perform MLE, but the concept is universal.