Maximum Likelihood Estimation

Lê Văn Phong

Phổ Dụng

Ngày 20 tháng 7 năm 2024

1 MLE

MLE là gì? I

Định nghĩa 1.1 (MLE)

Maximum Likelihood Estimation (MLE) là một phương pháp thống kê được sử dụng để **ước lượng các tham số** của một mô hình phân phối xác suất **dựa trên dữ liệu quan sát được**.
Mục tiêu chính của MLE là tìm ra các giá trị tham số sao cho mô hình có khả năng tạo ra dữ liệu quan sát được là cao nhất, hay nói cách khác, là **mô hình phù hợp nhất với dữ liệu thực tế**.

Ứng dụng MLE trong đầu tư? I

1. **Dự báo biến động giá và rủi ro**: Các mô hình như GARCH sử dụng MLE để ước lượng các tham số, từ đó *dự báo biến động giá trong tương lai* và giúp nhà đầu tư quản lý rủi ro hiệu quả hơn.
2. **Đánh giá hiệu quả danh mục đầu tư**: MLE được sử dụng trong mô hình CAPM để *ước lượng hệ số beta*, giúp *đánh giá hiệu quả của một danh mục đầu tư* so với thị trường chung.
3. **Định giá các sản phẩm phái sinh**: Mô hình Black-Scholes, một công cụ quan trọng trong định giá quyền chọn, sử dụng MLE để *ước lượng biến động ngầm định của tài sản cơ sở.*
4. **Phân tích kỹ thuật**: Nhiều chỉ báo kỹ thuật phổ biến như đường trung bình động (MA) sử dụng MLE để *xác định xu hướng thị trường và tạo ra các tín hiệu giao dịch*.
5. **Xây dựng các mô hình đầu tư phức tạp**: Như ARIMA, Copula, mạng nơ-ron, để *dự báo giá, phân tích rủi ro, tối ưu danh mục đầu tư,...*

Lưu ý

MLE là một phương pháp ước lượng tham số thống kê dựa trên việc tìm các giá trị tham số tối đa hóa hàm likelihood (khả năng). Trong đó **Hàm likelihood** *đo lường mức độ phù hợp của một mô hình thống kê với dữ liệu quan sát được.*

Giả sử ta có một mẫu dữ liệu **độc lập và phân phối giống nhau** là $X = \{x_1, x_2, \ldots, x_n\}$ với *hàm mật độ xác suất* là $f(x, \theta)$, trong đó θ là vector tham số cần ước lượng. Hàm **likelihood** được định nghĩa bởi

$$L(\theta \mid X) := \prod_{i=1}^{n} f(x_i, \theta).$$

Phương pháp MLE sẽ tìm giá trị

$$\theta^* = \arg\max L(\theta \mid X)$$

để tối đa hóa hàm likelihood nói trên.

Trong thực tế, ta thường tính toán với hàm **log-likelihood** vì hàm ln là một hàm đơn điệu tăng, nên nó không làm thay đổi điểm cực đại và còn dễ tính toán hơn,

$$l(\theta \mid X) := \ln L(\theta \mid X) = \sum_{i=1}^{n} \ln f(x_i, \theta).$$

Ta giải phương trình sau giúp tìm θ^* để $l(\theta \mid X)$ đạt max

$$\frac{\partial l(\theta \mid X)}{\partial \theta} = 0.$$

Ví dụ 1.2 (Phân phối Bernoulli)

- **Mô hình:** Tung một đồng xu n lần, với xác suất thành công, (tức ra mặt ngửa) là p. Số lần thành công X tuân theo phân phối Bernoulli.
- **Hàm mật độ xác suất**: $f(x, p) = p^x(1-p)^{1-x}$ với $x \in \{0, 1\}$.
- **Mẫu dữ liệu:** Giả sử chúng ta tung đồng xu 1000 lần và thu được 550 lần ngửa ($x = 1$) và 450 lần sấp ($x = 0$).
- **Hàm log-likelihood**: $l(p \mid X) = 550 \ln p + 450 \ln(1-p)$.
- **Giải phương trình** $\dfrac{\partial l(p \mid X)}{\partial p} = \dfrac{550}{p} - \dfrac{450}{1-p} = 0$ cho ta $p^* = 0.55$.

Ý nghĩa: Ước lượng MLE này cho biết, dựa trên dữ liệu quan sát, xác suất đồng xu ra mặt ngửa là 0.55.

Ví dụ 1.3 (Phân phối Poisson)

- **Mô hình:** Số lượng khách hàng đến một cửa hàng trong một giờ tuân theo phân phối $Poisson(\lambda)$.
- **Hàm mật độ xác suất**: $f(x, \lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ với $x \in \{0, 1\}$.
- **Mẫu dữ liệu:** Giả sử chúng ta quan sát số lượng khách hàng trong 5 giờ và thu được dữ liệu: $X = \{3, 5, 2, 7, 10\}$.
- **Hàm log-likelihood**: $l(\lambda \mid X) = \sum\limits_{i=1}^{5} (-\lambda + x_i \ln \lambda - \ln x_i!)$.
- **Giải phương trình** $\dfrac{\partial l(\lambda \mid X)}{\partial \lambda} = -5 + \dfrac{\sum\limits_{i=1}^{5} x_i}{\lambda} = 0$ cho ta $\lambda^* = 5.4$.

Ý nghĩa: Ước lượng MLE này cho biết, dựa trên dữ liệu quan sát, ước tính trung bình có khoảng 5.4 khách hàng đến mỗi giờ.

Ví dụ 1.4 (Phân phối chuẩn)

- **Mô hình:** Giả sử lợi nhuận hàng ngày của một cổ phiếu tuân theo phân phối chuẩn $N(\mu, \sigma^2)$. Với mẫu dữ liệu trong 30 ngày, $X = \{x_1, \ldots, x_{30}\}$.
- **Hàm mật độ xác suất**: $f(x; \mu, \sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\dfrac{-(x-\mu)^2}{2\sigma^2}\right\}$.
- $l(\mu, \sigma^2 \mid X) = \displaystyle\sum_{i=1}^{30} \left(-\frac{(x_i - \mu)^2}{2\sigma^2} - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}\ln 2\pi \right)$.
- **Giải** $\dfrac{\partial l(\mu, \sigma^2 \mid X)}{\partial \mu} = \dfrac{1}{\sigma^2} \sum_{i=1}^{30} (x_i - \mu) = 0$ và $\dfrac{\partial l(\mu, \sigma^2 \mid X)}{\partial \sigma^2} = \dfrac{-n}{2\sigma^2} + \dfrac{1}{2\sigma^4} \sum_{i=1}^{30} (x_i - \mu)^2 = 0$ ta được $\mu^* = \left(\sum_{i=1}^{30} x_i\right)/30, (\sigma^2)^* = \left(\sum_{i=1}^{30} (x_i - \mu)^2\right)/30.$

Kết quả: Ước lượng MLE cho μ và σ^2 lần lượt là trung bình mẫu và phương sai mẫu của dữ liệu.
Ý nghĩa: Ước lượng này cho biết lợi nhuận trung bình và độ biến động của cổ phiếu dựa trên dữ liệu lịch sử.
Ví dụ ta có dữ liệu lợi nhuận hàng ngày của cổ phiếu trong 30 ngày, sau khi tính toán ta được

- trung bình mẫu $\mu^* = 0.5\%$,
- phương sai mẫu $(\sigma^2)^* = (1\%)^2$.

Khi đó chúng ta kỳ vọng cổ phiếu sẽ tăng giá trung bình 0.5% mỗi ngày và lợi nhuận của cổ phiếu có độ biến động khoảng 1% mỗi ngày.
Giả định phân phối chuẩn có thể **không hoàn toàn chính xác trong thực tế**. Tuy nhiên để có được một mô hình dự báo chính xác hơn, chúng ta cần phải xem xét thêm các yếu tố khác như *xu hướng thị trường, tin tức, và các sự kiện kinh tế vĩ mô*.

Ước lượng MLE có các thuộc tính thống kê tốt như

1. **Tính nhất quán** (Consistency)
2. **Tính hiệu quả** (Efficiency)
3. **Tính bất biến** (Invariance)

Lưu ý

- Các thuộc tính thống kê tốt của MLE chỉ đúng trong một số điều kiện nhất định. Ví dụ, mô hình phải được chỉ định đúng, dữ liệu phải được lấy mẫu một cách độc lập và đồng nhất,...
- Trong thực tế, các điều kiện này không phải lúc nào cũng được đáp ứng hoàn toàn. Tuy nhiên, MLE vẫn là một phương pháp ước lượng rất hữu ích và được sử dụng rộng rãi trong nhiều lĩnh vực.

Tính nhất quán

- **Ý nghĩa**: Khi kích thước mẫu (số lượng quan sát) tăng lên, ước lượng MLE sẽ hội tụ về giá trị thực của tham số cần ước lượng. Nói cách khác, nếu chúng ta có đủ dữ liệu, ước lượng MLE sẽ ngày càng chính xác hơn.
- **Giải thích**: Khi chúng ta có nhiều dữ liệu hơn, thông tin về tham số thực sự sẽ được thể hiện rõ hơn trong dữ liệu. MLE, với mục tiêu tối đa hóa khả năng xảy ra của dữ liệu, sẽ tận dụng thông tin này để đưa ra ước lượng ngày càng gần với giá trị thực.
- **Ví dụ**: Giả sử chúng ta muốn ước lượng xác suất thành công (p) của một loại thuốc mới. Chúng ta thực hiện một thử nghiệm lâm sàng với 100 bệnh nhân và thấy rằng 60 người thành công. Ước lượng MLE cho p là 0.6. Tuy nhiên, nếu chúng ta mở rộng thử nghiệm lên 1000 bệnh nhân và thấy 650 người thành công, ước lượng MLE sẽ thay đổi thành 0.65, gần hơn với giá trị thực của p (giả sử là $p = 0.67$).

Tính hiệu quả

- **Ý nghĩa**: Trong số các *ước lượng không chệch* (unbiased estimators), **ước lượng MLE có phương sai nhỏ nhất**. Điều này có nghĩa là ước lượng MLE *ít biến động hơn và tập trung hơn xung quanh giá trị thực của tham số.*
- **Giải thích**: MLE tận dụng tối đa thông tin có trong dữ liệu để đưa ra ước lượng. Do đó, ước lượng MLE thường có độ chính xác cao hơn so với các phương pháp ước lượng khác.
- **Ví dụ**: Giả sử chúng ta muốn ước lượng giá trị trung bình của một quần thể. Chúng ta có thể sử dụng cả ước lượng trung bình mẫu và ước lượng MLE. Cả hai ước lượng đều không chệch, nhưng ước lượng MLE thường có phương sai nhỏ hơn, đặc biệt khi kích thước mẫu lớn.

Tính bất biến

- **Ý nghĩa**: Nếu $g(\theta)$ là một hàm của tham số θ, thì ước lượng MLE của $g(\theta)$ là $g(\theta^*)$, trong đó θ^* là ước lượng MLE của θ. Điều này có nghĩa là chúng ta có thể dễ dàng tìm được ước lượng MLE của một hàm của tham số mà không cần phải tính toán lại từ đầu.
- **Giải thích**: Tính bất biến là một thuộc tính rất hữu ích của MLE, giúp chúng ta tiết kiệm thời gian và công sức tính toán.
- **Ví dụ**: Giả sử chúng ta đã ước lượng được giá trị trung bình μ và độ lệch chuẩn σ của một quần thể bằng MLE. Nếu chúng ta muốn ước lượng phương sai σ^2, chúng ta chỉ cần bình phương ước lượng MLE của độ lệch chuẩn $(\sigma^2)^*$ mà không cần phải tính toán lại từ dữ liệu gốc.

MLE trong đầu tư I

Định nghĩa 1.5 (Mô hình GARCH)

Mô hình GARCH (Generalized Autoregressive Conditional Heteroskedasticity) là một mô hình thống kê được phát triển bởi Robert F. Engle vào những năm 1980 để mô hình hóa sự biến đổi không đều của dữ liệu tài chính. Mô hình này được sử dụng rộng rãi trong lĩnh vực tài chính để phân tích và dự đoán rủi ro.
Mô hình GARCH kết hợp các thành phần autoregressive (AR) và moving average (MA) để mô hình hóa sự phụ thuộc của phương sai có điều kiện vào các giá trị đã qua xử lý cũng như các giá trị độ biến động trước đó.
Một mô hình $GARCH(p,q)$ có thể được biểu diễn như sau:

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

Trong đó

- σ_t^2 là phương sai có điều kiện của biến ngẫu nhiên tại thời điểm t.
- ε_t là biến ngẫu nhiên có kỳ vọng bằng 0 và **phương sai đẳng nhất**(homoscedasticity), nghĩa là phương sai của sai số (residuals) là không đổi đối với tất cả các giá trị của biến độc lập.
- $\omega, \alpha_i, \beta_j$ là các tham số của mô hình, thông thường các $\alpha_i, \beta_j \in [0, 1]$.
- p là số **lags** cho thành phần *autoregressive* (AR) và q là số lags cho thành phần *moving average* (MA).

Mô hình GARCH được ước lượng thông qua các phương pháp tối đa hóa hàm hợp lý hoặc phương pháp bình phương tối thiểu. Các giá trị ước lượng của các tham số được xác định để tối thiểu hóa sai số giữa dữ liệu quan sát và dữ liệu được dự đoán từ mô hình GARCH.
Mô hình GARCH được sử dụng rộng rãi trong các ứng dụng tài chính như dự báo giá trị tài sản, quản lý rủi ro tài chính, và đánh giá tác động của biến động thị trường lên các quyết định đầu tư và giao dịch.

Ví dụ 1.6 (Dự báo biến động giá cổ phiếu với mô hình GARCH(1,1))

- **Mô hình toán học**: Mô hình **GARCH(1,1)** mô tả biến động (*độ lệch chuẩn*) của lợi suất tài sản theo thời gian, với công thức:

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

  Trong đó
  - σ_t^2 là phương sai (biến động) của lợi suất tại thời điểm t.
  - ε_{t-1}^2 là bình phương sai số (residual) tại thời điểm $t-1$.
  - ω, α, β là các tham số của mô hình.
- **Mục tiêu**: Tìm các giá trị của tham số mô hình (ω, α, β trong trường hợp $GARCH(1,1)$) sao cho khả năng (likelihood) của việc quan sát được dữ liệu (lợi suất logarit) là cao nhất.

- Hàm likelihood của mô hình GARCH(1,1) được tính như sau:

$$L(\omega, \alpha, \beta) = \prod_t \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{\varepsilon_t^2}{2\sigma_t^2}\right\}$$

- Hàm log-likelihood

$$l(\omega, \alpha, \beta) = \sum_t \left(\frac{-\ln(2\pi)}{2} - \frac{\ln \sigma_t^2}{2} - \frac{\varepsilon_t^2}{2\sigma_t^2}\right).$$

- Đây là một bài toán tối ưu hóa không có nghiệm giải tích. Do đó, chúng ta cần sử dụng các phương pháp số học để tìm nghiệm gần đúng. Cụ thể ta sẽ sử dụng ngôn ngữ R để giải quyết.

Ý nghĩa các kết quả:

- ω: Mức biến động cơ bản của lợi suất.
- α: Mức độ ảnh hưởng của biến động ngắn hạn (shock) đến biến động trong tương lai.
- β: Mức độ ảnh hưởng của biến động dài hạn đến biến động trong tương lai.

Cài đặt và nạp gói rugarch
install.packages("rugarch")
library(rugarch)
Dữ liệu giả lập lợi suất logarit hàng ngày trong năm 2024
returns <- c(0.00588, -0.00819, 0.00471, 0.0123, -0.0095, 0.0187, -0.0062, 0.0105, -0.0138, 0.0156, 0.0039, -0.0074, 0.0145, -0.0112, 0.0089, -0.0023, 0.0065, -0.0109, 0.0172, -0.0145, 0.0056, -0.0032, 0.0118, -0.0091, 0.0163)
Đặc tả mô hình GARCH(1,1)
spec <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)), mean.model = list(armaOrder = c(0, 0), include.mean = TRUE), distribution.model = "norm")
Ước lượng mô hình
fit <- ugarchfit(spec, returns)
In ra kết quả ước lượng print(fit)

Trích xuất các tham số ước lượng

omega <- coef(fit)[1]

alpha <- coef(fit)[2]

beta <- coef(fit)[3]

Dự báo biến động cho ngày tiếp theo

last_volatility <- fit@fitsigma[length(fit@fitsigma)]

Lấy biến động của ngày cuối cùng

$forecast\_volatility <$

$-sqrt(omega + alpha * last\_volatility^2 + beta * last\_volatility^2)$

In ra biến động dự báo

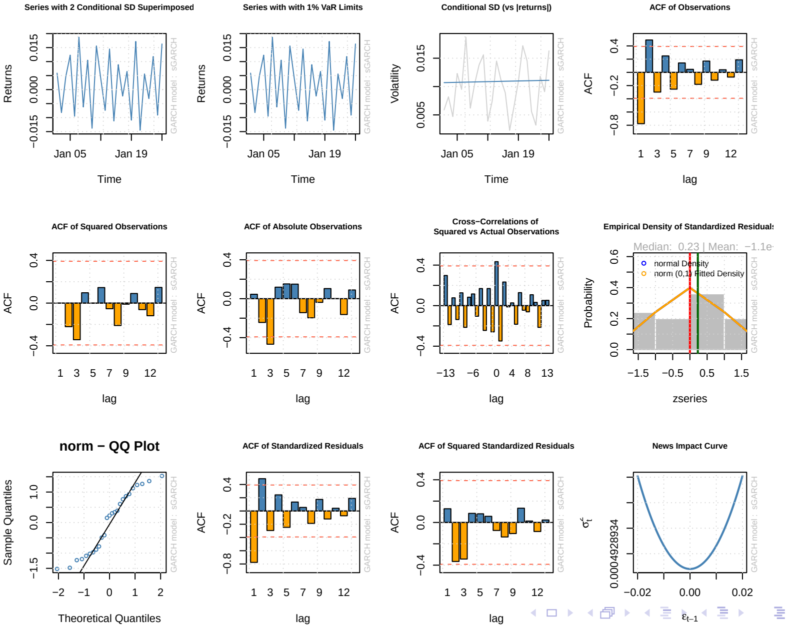
cat("Biến động dự báo cho ngày tiếp theo:", forecast_volatility)

Mở một cửa sổ đồ thị mới

dev.new(width = 10, height = 6)

Vẽ đồ thị biến động

plot(fit, which = "all")


Series with 2 Conditional SD Superimposed
Series with with 1% VaR Limits
Conditional SD (vs |returns|)
ACF of Observations
ACF of Squared Observations
ACF of Absolute Observations
Cross−Correlations of
Squared vs Actual Observations
Empirical Density of Standardized Residuals
Median: 0.23 | Mean: −1.1e
normal Density
norm (0,1) Fitted Density
norm − QQ Plot
ACF of Standardized Residuals
ACF of Squared Standardized Residuals
News Impact Curve
Returns
Volatility
ACF
Probability
Sample Quantiles
Time
lag
zseries
Theoretical Quantiles
Jan 05
Jan 19
GARCH model : sGARCH

Lưu ý

- Hàm **ugarchfit** thực hiện ước lượng mô hình GARCH(1,1) bằng phương pháp MLE.
- Hàm **coef(fit)** trả về một vector chứa các giá trị ước lượng của ω, α, β. Chúng ta lưu các giá trị này vào các biến tương ứng.
- Mô hình GARCH(1,1) với phương sai có điều kiện là **sGARCH** và mô hình trung bình là **ARMA(0,0)**.
  - **Phương sai có điều kiện là sGARCH**: sGARCH là viết tắt của standard GARCH. Đây là một dạng đặc biệt của mô hình GARCH, trong đó phương sai có điều kiện được tính theo công thức trên. sGARCH là dạng đơn giản nhất của mô hình GARCH và thường được sử dụng làm điểm khởi đầu cho việc phân tích.
  - **Mô hình trung bình là ARMA(0,0)**:ARMA là viết tắt của Autoregressive Moving Average. Mô hình ARMA(0,0) là mô hình đơn giản nhất của ARMA, trong đó lợi suất trung bình được giả định là không đổi theo thời gian.

Lưu ý

- **Ý nghĩa:**
  - Mô hình GARCH(1,1) với phương sai có điều kiện là sGARCH và mô hình trung bình là ARMA(0,0) có nghĩa là:
  - Biến động của lợi suất tài sản phụ thuộc vào cả biến động trong quá khứ (thông qua thành phần GARCH) và các cú sốc (shock) ngẫu nhiên (thông qua thành phần ARCH).
  - Lợi suất trung bình của tài sản là không đổi theo thời gian.
- **Ứng dụng:**Mô hình GARCH(1,1) thường được sử dụng để:
  - Dự báo biến động giá của các tài sản tài chính (cổ phiếu, trái phiếu, tỷ giá hối đoái, ...).
  - Tính toán Value at Risk (VaR), một thước đo rủi ro quan trọng trong quản lý danh mục đầu tư.
  - Xây dựng các chiến lược giao dịch dựa trên biến động.