Aula 3: Modelos de Volatilidade Estocástica (GARCH)

Marcus L. Nascimento

20 de outubro de 2025

Introdução

- Vimos que, utilizando o operador defasagem, um modelo ARMA(p,q) é representado da seguinte forma:

$$\begin{aligned}(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)(y_t-\mu) &= (1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q)w_t \\ \boldsymbol{\phi}(B)(y_t-\mu) &= \boldsymbol{\theta}(B)w_t. \qquad (1)\end{aligned}$$

- Reescrevendo a Equação (1) como um processo de médias móveis (representação de Wold), temos:

$$\begin{aligned}(y_t-\mu) &= \boldsymbol{\phi}(B)^{-1}\boldsymbol{\theta}(B)w_t \\ (y_t-\mu) &= \boldsymbol{\psi}(B)w_t, \qquad (2)\end{aligned}$$

onde $\boldsymbol{\psi}(B)w_t = \boldsymbol{\phi}(B)^{-1}\boldsymbol{\theta}(B)$.

Introdução

- Supondo uma série temporal de tamanho n, a partir da Equação (2), temos:

$$y_{n+l} = \mu + \underbrace{w_{t+l} + \psi_1 w_{t+l-1} + \psi_2 w_{t+l-2} + \ldots}_{\text{desconhecido}} + \underbrace{\psi_l w_n + \ldots}_{\text{conhecido}}.$$

Logo, os erros de previsão podem ser escritos da seguinte forma:

$$\begin{aligned}
e_n(1) &= y_{n+1} - y_n(1) = w_{n+1} \\
e_n(2) &= y_{n+2} - y_n(2) = w_{n+2} + \psi_1 w_{n+1} \\
e_n(3) &= y_{n+3} - y_n(3) = w_{n+3} + \psi_1 w_{n+2} + \psi_2 w_{n+1} \\
&\vdots \\
e_n(l) &= y_{n+l} - y_n(l) = \sum_{i=0}^{l-1} \psi_i w_{n+l-i},
\end{aligned}$$

onde $\psi_0 = 1$.

OBSERVAÇÕES:

- Os erros de previsão podem guardar algum tipo de relação;
- O comportamento dos erros de previsão depende das perturbações w_t:
  - Justificativa plausível para a existência de autocorrelação na variância de w_t;
  - Variâncias dos erros de previsão não são constantes; autocorrelação na variância de $e_t(l)$;
  - Heterocedasticidade.

- Os itens destacados acima são comumente observados ao prevermos séries temporais econômicas (preços de ações, taxas de inflação, taxas de câmbio, entre outras).
  - Capacidade preditiva de modelos estatísticos oscila entre períodos de análise;
  - Períodos de maior **volatilidade** (variância) tendem a perdurar: maior volatilidade hoje pode implicar maior volatilidade amanhã.

Introdução

- A partir da análise da série de inflação do Reino Unido, Engle (1982) propôs um modelo não-linear denominado ARCH (Modelo Autorregressivo para a Heteroscedasticidade Condicional) com intuito de capturar a estrutura de correlação na variância.

- Posteriormente, Bollerslev (1986) propôs uma generalização do modelo ARCH denominada GARCH com intuito de descrever a volatilidade com um número menor de parâmetros que o modelo ARCH.

- Neste estágio do curso, focaremos em estudar técnicas estatísticas que têm por finalidade modelar a volatilidade, ou seja, a variância condicional de uma variável (geralmente, **retorno**).

Fatos estilizados sobre os retornos:

- Em geral, séries temporais de retornos não apresentam autocorrelação;
- Usualmente, agrupamentos de volatilidades ao longo do tempo são observados em séries temporais de retornos;
- Ao analisarmos os quadrados dos retornos é possível observar uma correlação pequena e defasagem 1 e uma queda lenta na correlação para a defasagens subsequentes;
- Distribuição (incondicional) dos retornos é, habitualmente, leptocúrtica e simétrica;
- Presença de não-linearidade em algumas das séries temporais de retornos.

Modelos não-lineares

Modelos não-lineares

- Seja $Y_1, Y_2, \ldots, Y_n$ uma série temporal relativa aos preços de uma ação, denotaremos a série temporal dos retornos através da sequência de variáveis aleatórias $R_1, R_2, \ldots, R_n$, onde $R_t = \log(Y_t/Y_{t-1})$.

- Supondo que as inovações (choques aleatórios) W_t, $t = 1, 2, \ldots, n$, sejam independentes e identicamente distribuídas, um modelo não-linear pode der escrito da seguinte forma:

$$r_t = \mu(r_{t-1}, r_{t-2}, \ldots) + \tau(r_{t-1}, r_{t-2}, \ldots)w_t,$$

onde

  - $\mu(r_{t-1}, r_{t-2}, \ldots)$ é a função média condicional;
  - $\tau(r_{t-1}, r_{t-2}, \ldots)$ é a função variância condicional.

Modelos não-lineares

- Note que:
  - Se $\mu(r_{t-1}, r_{t-2}, \ldots)$ for não linear, então o modelo é não-linear na média;
  - Se $\tau(r_{t-1}, r_{t-2}, \ldots)$ for não linear, então o modelo é não-linear na variância.
- Supondo $\mathcal{J}_{t-1} = \{r_1, r_2, \ldots, r_{t-1}\}$ a informação disponível no tempo $t-1$:
  - A média condicional de R_t é dada por $\mathrm{E}(R_t|\mathcal{J}_{t-1}) = \mathrm{E}_{t-1}(R_t)$;
  - A variância condicional de R_t é dada por $\tau_t^2 = \mathrm{E}\left[(R_t - \mu_t)^2|\mathcal{J}_{t-1}\right] = \mathrm{E}_{t-1}\left[(R_t - \mu_t)^2\right]$;
  - Se $\mu_t = 0$, então $\mathrm{E}(R_t^2|\mathcal{J}_{t-1}) = \mathrm{E}_{t-1}(R_t^2)$.

Processos de Heterocedasticidade Condicional Autorregressiva (ARCH)

Processo ARCH de Ordem 1 - ARCH(1)

- Com o objetivo de estudar algumas das propriedades dos processos ARCH(p), iremos considerar o caso particular onde $p = 1$.

- Seja $\{W_t\}$ ruído branco com média zero e variância igual a um, um processo ARCH(1) é descrito da seguinte forma:

$$r_t = \tau_t w_t,$$

onde $\tau_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$.

Processo ARCH de Ordem 1 - ARCH(1)

- Média condicional de R_t:

$$\begin{aligned}
\mathrm{E}(R_t|\mathcal{J}_{t-1}) &= \mathrm{E}_{t-1}(R_t) \\
&= \mathrm{E}_{t-1}(\tau_t W_t) \\
&= \tau_t \mathrm{E}_{t-1}(W_t) \\
&= 0.
\end{aligned}$$

- Média incondicional de R_t:

$$\begin{aligned}
\mathrm{E}(R_t) &= \mathrm{E}(\mathrm{E}(R_t|\mathcal{J}_{t-1})) \text{ (Regra da torre)} \\
&= \mathrm{E}(\mathrm{E}_{t-1}(R_t)) \\
&= 0.
\end{aligned}$$

- Processo ARCH(1) possui média zero.

Processo ARCH de Ordem 1 - ARCH(1)

- Covariância entre R_t e R_{t-1}:

$$\begin{aligned}\gamma_1 = \gamma_R(t, t-1) &= \mathrm{Cov}(R_t, R_{t-1}) \\ &= \mathrm{E}(R_t R_{t-1}) - \mathrm{E}(R_t)\mathrm{E}(R_{t-1}) \\ &= \mathrm{E}(\mathrm{E}_{t-1}(R_t R_{t-1})) - 0 \\ &= \mathrm{E}(R_{t-1}\mathrm{E}_{t-1}(R_t)) \\ &= 0.\end{aligned}$$

- Seguindo um argumento similar, é possível mostrar que $\gamma_l = 0 \ \forall l \geq 0$.

- Processo ARCH(1) é não correlacionado serialmente.

Processo ARCH de Ordem 1 - ARCH(1)

- Devido à não correlação serial, R_t não pode ser predito através do histórico ($\mathcal{J}_{t-1}$): Evidência da hipótese de eficiência de mercado.
- Variância condicional:

$$\begin{aligned} \text{Var}(R_t|\mathcal{J}_{t-1}) &= \mathrm{E}(R_t^2|\mathcal{J}_{t-1}) - [\mathrm{E}(R_t|\mathcal{J}_{t-1})]^2 \\ &= \mathrm{E}_{t-1}(\tau_t^2 W_t^2) - [\mathrm{E}_{t-1}(R_t)]^2 \\ &= \tau_t^2 \mathrm{E}_{t-1}(W_t^2) - 0 \\ &= \tau_t^2. \end{aligned}$$

- A variância condicional é τ_t^2 que, por definição, é descrita como a seguir:

$$\tau_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2.$$

Logo, R_t^2 pode ser predito (através de R_{t-1}).

Processo ARCH de Ordem 1 - ARCH(1)

- Observe que

$$\mathrm{E}(R_t^2|\mathcal{J}_{t-1}) = \alpha_0 + \alpha_1 r_{t-1}^2. \tag{3}$$

- A partir da Equação (3), temos:
  - R_t^2 segue um processo AR(1);
  - α_0 e α_1 podem ser estimados por uma regressão de r_t^2 em um intercepto e r_{t-1}^2;
  - Processo ARCH(1) pode ser estimado via mínimos quadrados ordinários.

Processo ARCH de Ordem 1 - ARCH(1)

- Variância incondicional:

$$
\begin{aligned}
\text{Var}(R_t) &= \mathrm{E}(R_t^2) - [\mathrm{E}(R_t)]^2 \\
&= \mathrm{E}(\mathrm{E}_{t-1}(R_t^2)) - 0 \\
&= \mathrm{E}(\alpha_0 + \alpha_1 R_{t-1}^2) \\
&= \alpha_0 + \alpha_1 \mathrm{E}(R_{t-1}^2).
\end{aligned}
$$

Sob estacionariedade, temos:

$$
\tau^2 = \text{Var}(R_t) = \frac{\alpha_0}{1 - \alpha_1},
$$

onde $\alpha_0 > 0$ e $0 < \alpha_1 < 1$.

Processo ARCH de Ordem 1 - ARCH(1)

- Dos resultados anteriores, temos:
  - Variância incondicional: $\tau^2 = \frac{\alpha_0}{1-\alpha_1}$;
  - Variância condicional: $\tau_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$.
- Substituindo $\alpha_0 = \tau^2(1-\alpha_1)$, na variância condicional, temos:

$$\begin{aligned} \tau_t^2 &= \tau^2(1-\alpha_1) + \alpha_1 r_{t-1}^2 \\ &= \tau^2 + \alpha_1(r_{t-1}^2 - \tau^2). \end{aligned}$$

- Variância condicional pode ser vista como uma combinação entre a variância incondicional e o desvio entre o retorno ao quadrado e seu valor médio.

Processo ARCH de Ordem 1 - ARCH(1)

- Vimos que um dos fatos estilizados de séries temporais de retornos consiste em caudas mais longas que a normal (curtose maior que 3).

- Supondo um modelo ARCH(1) com $W_t \stackrel{iid}{\sim} N(0,1)$, temos que a curtose é dada por:

$$K = \frac{\mathrm{E(R_t^4)}}{[\mathrm{Var}(R_t)]^2} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3.$$

- Assumindo um modelo ARCH, teremos caudas mais pesadas que as da normal - propriedade conveniente do modelo.

Processo ARCH de Ordem p - ARCH(p)

- Seja $\{W_t\}$ ruído branco com média zero e variância igual a um, um processo ARCH(p) é obtido se R_t^2 segue um processo AR(p):

$$\begin{aligned} r_t &= \tau_t w_t \\ \tau_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2, \end{aligned}$$

onde $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \ldots, p-1$, e $\alpha_p > 0$.

- As restrições anteriores são condições para que $\tau_t^2 > 0$.

Processo ARCH de Ordem p - ARCH(p)

OBSERVAÇÕES:

- Usualmente, assume-se que W_t segue um distribuição
  - Normal padrão;
  - t-Student com baixos graus de liberdade ν ou alguma outra distribuição de caudas pesadas que melhor descreva séries temporais financeiras.
- Os coeficientes α_i devem satisfazer certas condições a depender do tipo de imposição que colocamos sobre o processo R_t.

Processos ARCH generalizados (GARCH)

Processo GARCH(p,q)

- Não raramente, é necessário que p seja grande para que toda correlação serial em R_t^2 seja capturada.

- O processo ARCH generalizado (GARCH) consiste em uma alternativa parsimoniosa ao modelo ARCH(p).

- Seja $\{W_t\}$ ruído branco com média zero e variância igual a um, um processo GARCH(p,q) é descrito da seguinte forma:

$$
\begin{aligned}
r_t &= \tau_t w_t \\
\tau_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \tau_{t-j}^2.
\end{aligned}
$$

Processo GARCH(p,q)

- Como condições para que $\tau_t^2 > 0$, as seguintes restrições são impostas:
    - $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \ldots, p - 1$, e $\alpha_p > 0$;
    - $\beta_j \geq 0$, $j = 1, 2, \ldots, q - 1$, e $\beta_q > 0$;
    - Para $v = \max\{p, q\}$, $\sum_{i=1}^{v}(\alpha_i + \beta_i) < 1$.
- Assim como nos modelos ARCH, é usual supor que W_t tenha distribuição normal ou t-Student, ou ainda uma outra distribuição de cauda pesada.

Processo GARCH(1,1)

- Com o objetivo de estudar alguma das propriedades dos processos GARCH(p,q), consideraremos o caso em que $p = q = 1$:

$$\begin{aligned} r_t &= \tau_t w_t \\ \tau_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \tau_{t-1}^2. \end{aligned}$$

- Variância incondicional:

$$\begin{aligned} \mathrm{Var}(R_t) &= \mathrm{E}(R_t^2) - [\mathrm{E}(R_t)]^2 \\ &= \mathrm{E}(\mathrm{E}_{t-1}(R_t^2)) - 0 \\ &= \mathrm{E}(\alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \tau_{t-1}^2) \\ &= \alpha_0 + \alpha_1 \mathrm{E}(R_{t-1}^2) + \beta_1 \mathrm{E}(\tau_{t-1}^2) \\ \Rightarrow \tau^2 = \mathrm{Var}(R_t) &= \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \text{ (Sob estacionariedade),} \end{aligned}$$

onde $\alpha_0 > 0$ e $0 < \alpha_1 + \beta_1 < 1$.

Processo GARCH(1,1)

- Supondo um modelo GARCH(1,1) com $W_t \overset{iid}{\sim} N(0,1)$, temos que a curtose é dada por:

$$K = \frac{\mathrm{E}(\mathrm{R}_t^4)}{[\mathrm{Var}(R_t)]^2} = 3\left(\frac{1-(\alpha_1+\beta_1)^2}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2}\right) > 3.$$

- Assumindo um modelo GARCH, teremos caudas mais pesadas que as da normal - propriedade desejável do modelo já que retornos normalmente apresentam tal característica.

Processo GARCH(1,1)

- Modelos ARCH podem ser estimados via mínimos quadrados ordinários ou via máxima verossimilhança.
- Modelos GARCH, em contrapartida, são estimados via máxima verossimilhança:
  - Assumindo $W_t \overset{iid}{\sim} N(0,1)$ e $r_0^2 = \tau_0^2 = 0$, a verossimilhança pode ser obtida recursivamente:

$$
\begin{aligned}
\tau_1^2 &= \alpha_0 \\
\frac{r_1}{\tau_1} &\sim N(0,1) \\
\vdots &= \vdots \\
\tau_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \tau_{t-1}^2 \\
\frac{r_t}{\tau_t} &\sim N(0,1).
\end{aligned}
$$

  - O estimador de máxima verossimilhanda para α_0, α_1 e β_1 é obtido maximizando o produto das verossimilhanças.

Material suplementar

Modelo GARCH na média (GARCH-M)

- Um modelo GARCH(1,1)-M é dado por:

$$\begin{aligned} r_t &= \mu + c\tau_t^2 + u_t \\ u_t &= \tau_t w_t \\ \tau_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \tau_{t-1}^2, \end{aligned}$$

  onde μ e c são constantes.

  Observações:

  - A formulação do modelo implica correlação serial nos retornos R_t;
  - c é chamado de parâmetro de risco de prêmio. Valores positivos de c indicam que o retorno é positivamente relacionado com sua volatilidade;
  - Outras especificações para o prêmio de risco presentes na literatura incluem: $r_t = \mu + c\tau_t + u_t$ e $r_t = \mu + c\log(\tau_t^2) + u_t$.

Modelo ARMA-GARCH

- Como ilustração, considere um modelo AR(1)-GARCH(1,1):

$$
\begin{aligned}
r_t &= \phi_0 + \phi_1 r_{t-1} + u_t \\
u_t &= \tau_t w_t \\
\tau_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \tau_{t-1}^2 .
\end{aligned}
$$

Observação:

- Neste caso, está sendo considerado que o retorno pode ser predito tanto em nível quanto em seu quadrado.
- Não há evidência da hipótese de eficiência de mercado.

Modelo GARCH exponencial (EGARCH)

- Nelson (1991) propôs o modelo GARCH exponencial (EGARCH).
- Um modelo EGARCH(p,q) é definido como a seguir:

$$
\begin{aligned}
r_t &= \tau_t w_t \\
\log(\tau_t^2) &= \alpha_0 + \sum_{i=1}^{q} g_i(w_{t-i}) + \sum_{j=1}^{p} \beta_j \log(\tau_{t-j}^2),
\end{aligned}
$$

onde $g_i(w_{t-i})$ permite efeitos assimétricos entre retornos positivos e negativos através de uma inovação ponderada:

$$
g_i(w_t) = \begin{cases} (\alpha_i + \gamma_i) w_{t-i} - \gamma_i \mathrm{E}(|w_t|), & \text{se } w_{t-i} \geq 0 \\ (\alpha_i - \gamma_i) w_{t-i} - \gamma_i \mathrm{E}(|w_t|), & \text{se } w_{t-i} < 0. \end{cases}
$$

Modelo GARCH exponencial (EGARCH)

Observações:

- γ captura o efeito assimétrico dos choques aleatórios, permitindo a diferenciação entre mudanças positivas e negativas.
  - Aplicado na modelagem de assimetrias em dados financeiros.
- Captura efeitos em que choques negativos tendem a aumentar a volatilidade de forma mais significativa do que choques positivos de mesma magnitude.
  - Adequado em casos em que o sentimento dos investidores é significativamente afetado por desenvolvimentos negativos.

Modelo *Threshold* GARCH (TGARCH)

- Proposto por Glosten, Jagannatha, and Runkle (1993) e Zakoian (1994).

- O modelo é descrito da seguinte forma:

$$
\begin{aligned}
r_t &= \tau_t w_t \\
\tau_t^2 &= \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \gamma_i I_{t-i}) |w_{t-i}|^{\delta} + \sum_{j=1}^{p} \beta_j \log(\tau_{t-j}^2),
\end{aligned}
$$

  onde I_{t-i} é uma variável indicadora que assume valor 1 caso $w_{t-i} < 0$ e 0 caso contrário.

- Denominamos o caso em que $\delta = 2$ de GJR-GARCH e o caso em que $\delta = 1$ de TGARCH.

Modelo *Threshold* GARCH (TGARCH)

Observações:

- Diferencia explicitamente os impactos positivos e negativos de choques aleatórios na volatilidade.

- Particularmente, aplicável em mercados em que as perdas podem afetar desproporcionalmente os preços se comparadas a ganhos equivalentes.

Referências

Bollerslev, Tim. 1986. "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 31 (3): 307–327.

Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50 (4): 987–1007.

Glosten, Lawrence R., Ravi Jagannatha, and David E. Runkle. 1993. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance* 48 (5): 1779–1801.

Nelson, Daniel B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59 (2): 347–370.

Zakoian, Jean-Michel. 1994. "Threshold heteroskedastic models." *Journal of Economic Dynamics and Control* 18 (5): 931–955.