Dynamische Systeme und Zeitreihenanalyse

Modelle für bedingte Heteroskedastizität
Kapitel 13

Statistik und Mathematik – WU Wien

Michael Hauser

Lernziele

- Autokorrelationsmuster in den Varianzen von Renditen
- Heteroskedastizität versus Homoskedastizität
- Tests auf Heteroskedastizität
- GARCH Modell für bedingte Heteroskedastizität
- Schätzung des GARCH Modells

Stylized facts

Unbedingte Verteilung von Renditen

Renditen von Aktien. y_t, (Tagesschlußkurse, P_t)

$$y_t = \log(P_t/P_{t-1}) = \log(P_t) - \log(P_{t-1})$$

- zeigen oft eine Kurtosis wesentlich größer 3.
  Die Verteilungen sind lepokurtisch.
- Zeichnet man die unbedingte Verteilung, so liegen viele Beobachtungen nahe um das Mittel, aber auch viele weit außen, symmetrisch um das Mittel. Es treten mehr “Ausreißer” auf als unter einer Normalverteilung.
- Für Monats- oder Wochendaten ist der Effekt aber wesentlich schwächer ausgeprägt.

Autokorrelierte Varianzen

- Renditen weisen ausgeprägte autokorrelierte Varianzen auf. Das Korrelogramm fällt langsam ab.
- Tritt eine große Abweichung vom Mittel auf, so zeigen auch die nachfolgenden Beobachtungen betragsmäßig große Abweichungen.
- Es gibt Perioden mit großer Volatilität und Perioden mit geringer Volatilität.
- Die quadrierten Renditen zeigen bei (Tages-)Finanzreihen deutlich größere Autokorrelationen als die Renditen selbst.

Begründungen

Verhaltenstheoretisch können folgende Gründe dafür angeführt werden:

- Veränderungen in den Informationsflüssen:
  Neue Informationen können bei den einzelnen Akteuren eine größere Unsicherheit darüber auslösen, wohin sich der Markt bewegen wird.
  Allmählich setzt sich eine Marktlinie durch. Die Unsicherheit, gemessen als Varianz, nimmt ab.
- In Zeiten erhöhter Unsicherheit treten weitere Akteure auf den Markt.
  Unter der Annahme, dass mehr Akteure auch eine höhere Unsicherheit bedeuten, steigt die Varianz.

Begründungen (Fs)

- Veränderungen im Handelsvolumen:
Volumen und die Anzahl der Akteure sind bis zu einem gewissen Grad austauschbar.
- Veränderungen in der Popularität:
Popularitätsveränderungen bewirken ein Wandern der Akteure von einem Markt zu einem anderen. Dadurch kann ebenfalls Autokorrelation in den Varianzen entstehen.

Bedingte Erwartung, bedingte Varianz

Konstanz der unbedingten Varianz

Für die Stationarität eines Prozesses verlangen wir, dass

- die unbedingte Erwartung, $E(y_t) = \mu$,
- die unbedingte Varianz $y_t = E(y_t - \mu)^2 = \sigma_y^2$, und
- die unbedingten Autokovarianzen, $\gamma(s)$,

konstant sind.

Wir verlangen *nicht*, dass die entsprechenden *bedingten* Werte konstant sind.

Für unsere bisher betrachteten Modelle ist das - zumindest für den AR-Teil - evident.

Beispiel: AR(1) $\quad (1 - \alpha_1 L)(y_t - \mu) = \epsilon_t$, gilt:

$$E(y_t) \neq E[y_t | y_{t-1}] = \alpha_1 (y_{t-1} - \mu) + \mu.$$

Modellierung des bedingten Varianz

Analog zur Modellierung des bedingten Mittels kann die bedingte Varianz von y_t, bzw die bedingte Varianz der Residuen, ϵ_t, modelliert werden.

Man modelliert die bedingte Varianz des Residuums zum Zeitpunkt t, h_t, als lineare Funktion

- von quadrierten vergangenen Residuen,
- von quadrierten vergangenen Werten der Beobachtungen und
- von vergangenen bedingten Varianzen.

Als Verteilung wird

- die Normalverteilung oder
- eine breitschwänzige Verteilung (zB die t-Verteilung mit einer kleinen Anzahl von Freiheitsgraden, $df \leq 7$) verwendet.

Das GARCH Modell

Das Modell

Das allgemeine Modell:

$$\epsilon_t \sim N(0, h_t), \qquad \epsilon_t \ldots \text{unkorreliert}$$

$$h_t = a_0 + a(L)\epsilon_t^2 + b(L)h_t$$

Die Lag-Polynome $a(L)$ und $b(L)$ haben hier *die Form*

$$d_w(L) = d_1 L + d_2 L^2 + \ldots + d_w L^w,$$

mit $d_i \geq 0$, iA. $a_0 > 0$.

- Im Gegensatz zu AR- und MA-Polynomen fehlt hier das konstante Glied.
- Nur verzögerte Werte gehen in das Modell ein.

ARCH, GARCH

(i) Das Modell mit unkorrelierten ϵ_t

$$\epsilon_t \sim N(0, h_t) \quad h_t = E(\epsilon_t^2 | I_{t-1}) \quad h_t = a_0 + a_r(L)\epsilon_t^2$$

heißt **ARCH(r) Modell, autoregressive conditional heteroscedasticity**.

(ii) Das Modell mit unkorrelierten ϵ_t

$$\epsilon_t \sim N(0, h_t) \quad h_t = E(\epsilon_t^2 | I_{t-1}) \quad h_t = a_0 + a_r(L)\epsilon_t^2 + b_s(L)h_t$$

heißt **GARCH(r, s) Modell, generalized autoregressive conditional heteroscedasticity**.

GARCH - Beispiele

Ein ARCH(1) Modell ist

$$\epsilon_t \sim N(0, h_t) \quad h_t = E(\epsilon_t^2 | I_{t-1}) \quad h_t = a_0 + a_1\, \epsilon_{t-1}^2$$

Ein ARCH(2) Modell ist

$$\epsilon_t \sim N(0, h_t) \quad h_t = E(\epsilon_t^2 | I_{t-1}) \quad h_t = a_0 + a_1\, \epsilon_{t-1}^2 + a_2\, \epsilon_{t-2}^2$$

Das GARCH(0,1) Modell

$$\epsilon_t \sim N(0, h_t) \quad h_t = E(\epsilon_t^2 | I_{t-1}) \quad h_t = a_0 + b_1\, h_{t-1}$$

ist *nicht geeignet*. Es beschreibt eine deterministisch ab(zu)nehmende Varianz für $0 < b_1 < (>)1$.

IGARCH(1,1)

Das Modell heißt **IGARCH(1,1), integrierter GARCH**, wenn in

$$y_t = c + \epsilon_t \qquad \epsilon_t \sim N(0, h_t)$$

$$h_t = a_0 + a_1\, \epsilon_{t-1}^2 + b_1\, h_{t-1}$$

$a_1 + b_1 = 1$ gilt.

Bemerkungen

- Oft wird nur von ARCH gesprochen, wenn irgend ein Modell für bedingte Heteroskedastizität gemeint ist.
- Die Positivitätsrestriktion für die Parameter und die Einbeziehung der Quadrate der verschiedenen Variablen hat zum Zweck, dass die Positivität der (bedingten) Varianz, h_t, gewährleistet ist.
  Im Speziellen hat a_0 immmer größer als Null zu sein: $a_0 > 0$!

Eigenschaften von ARCH Modellen

Damit ϵ_t ein stationärer Prozess ist, muss die unbedingte Varianz von $\epsilon_t,\ E\epsilon_t^2$ (da $\mu_\epsilon = 0$), existieren.

$$E(\epsilon_t^2) = \lim_{j \to \infty} E(\epsilon_t^2 | I_{t-j})$$

Der Grenzwert für $j \to \infty$ zeigt an, dass die vergangene Information weniger wichtig, also vergessen wird.

Im Prinzip ist das Modell eine Differenzengleichung, von der wir das Verhalten für $t \to \infty$, die Stabilitätseigenschaften, untersuchen.

Eigenschaften des GARCH(1,1)

Im Beispiel vom GARCH(1,1) Modell gilt für den Grenzwert, falls er existiert:

$$\sigma_\epsilon^2 = a_0 + a_1 \sigma_\epsilon^2 + b_1 \sigma_\epsilon^2$$

Wir lösen nach σ_ϵ^2 auf.

$$\sigma_\epsilon^2 = \frac{a_0}{1 - (a_1 + b_1)}$$

Die unbedingte Varianz existiert, solange $a_1 + b_1 < 1$ gilt.

Die höheren (geraden) Momente von GARCH Prozessen hängen von der Wahl der Parameter a_1 und b_1 ab.

Eigenschaften von ARCH Modellen

ARCH Modelle haben iA eine Kurtosis deutlich größer als 3. Sie können leptokurtisch Verteilungen modellieren.

Die Bedingungen für die Existenz der höheren Momente sind in der Figur aus Bollerslev(1986), *Journal of Econometrics*, 31, pp.307-327, graphisch dargestellt.

Eigenschaften von ARCH Modellen

Wie entdeckt man ARCH ?

- Muster in den Varianzen oder im Absolutbetrag der Reihe.
- Nach Berücksichtigung von Effekten in der bedingten Erwartung der Reihe (lineare Effekte) ist die Autokorrelation in den quadrierten Residuen groß.

  Dafür ist der Portmanteau Test von Box-Pierce analog zum ARMA Fall anwendbar. Die Statistik ist ebenfalls asymptotisch χ^2-verteilt mit der Anzahl der Summanden als Freiheitsgrad.

Schätzen

Schätzung von ARCH

Liegen lineare Effekte vor (ie Effekte in der bedingten Erwartung), ist zuerst ein ARMA Modell anzupassen.
Die Residuen sind anschließend als ARCH zu modellieren.

Für die Schätzung geht man von den **standardisierten Residuen**

$$\epsilon_t / \sqrt{h_t}$$

aus. Wir nehmen sie als (gemeinsam) standardnormalverteilt, $N(0,1)$, und unkorreliert an.

$$\epsilon_t / \sqrt{h_t} \;\sim\; N(0,1)$$

Likelihood des ARCH Modells

Die Likelihood ergibt sich dann einfach als Produkt der Dichten für jeden einzelnen Zeitpunkt.

Die **Likelihood eines ARCH Modells** bei Annahme einer bedingten Normalverteilung ist

$$L(\theta|\epsilon) = \prod_{t=1}^{T} f_N(\theta|\epsilon) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\sqrt{h_t}} e^{-\frac{1}{2}\frac{\epsilon_t^2}{h_t}}$$

μ ist hier Null.
(Man bereinigt die Reihe y_t zuvor um das Mittel: $(y_t - \overline{y})$.)

h_t ist eine Funktion der Parameter, $h_t = h_t(\theta)$.
$\theta = (a_0, a_1, \ldots, b_1, \ldots, c_1, \ldots)'$.

Eigenschaften des Schätzers

Der **ML Schätzer** ist

- asymptotisch erwartungstreu
- asymptotisch normalverteilt und
- asymptotisch effizient

Einzelne Parameter können analog zum ARMA Modell mit Hilfe

- des t-Tests

getestet werden.

Erweiterungen und Diskussion

ARCH-in-Mean Modell

Das Modell mit $\mu_1 > 0$

$$y_t = \mu(h_t) + \epsilon_t \quad \text{bzw} \quad y_t = \mu_0 + \mu_1 h_t + \epsilon_t \quad \text{mit} \quad \epsilon_t \sim N(0, h_t)$$

heißt **ARCH-M Modell, ARCH-in-Mean Modell**.

Angenommen y_t bezeichnet die Rendite einer Finanzreihe.
Die Motivation für ein ARCH-M Modell liegt darin, dass risikoaverse Akteure ein höhere durchschnittliche Rendite erwarten müssen um in Perioden höherer Volatilität (bedingtes Risiko) im Markt aktiv zu sein.

Die durchschnittliche Rendite ist eine Funktion der Volatilität.
Es liegt ein Trade-off zwischen erwarteter Rendite und Risiko vor.

Missspezifikationsanalyse

Angenommen das wahre Modell ist ein ARMA(p, q)-ARCH(r) Modell.

$$\alpha_P(L)(y_t - \mu) = \beta_q(L)\epsilon_t, \quad \epsilon_t \sim N(0, h_t), \quad \Psi(L) = \beta_q(L)/\alpha_p(L)$$

(i) Angenommen wir spezifizieren es als ARMA(p, q)

$$\alpha_p(L)(y_t - \mu) = \beta_q(L)\epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

und vernachlässigen den ARCH Teil.

So ist die bedingte Erwartung richtig spezifiziert (die bedingte Varianz nicht).

$E(y_t - \mu | I_{t-1}) = \sum_{j=1}^{\infty} \psi_j \, \epsilon_{t-j}$,

$V(y_t | I_{t-1}) = \sum_{j=1}^{\infty} \psi_j^2 \, \sigma_\epsilon^2 \neq \sum_{j=1}^{\infty} \psi_j^2 \, h_{t-j}$ (iA)

Missspezifikationsanalyse

(ii) Wir spezifizieren es nur als ARCH(r)

$$y_t - \mu = \eta_t \quad \eta_t \sim N(0, h_{\eta,t})$$

und vernachlässigen den ARMA Teil.

So sind die bedingte Erwartung und die bedingte Varianz falsch spezifiziert.

$E(y_t - \mu | I_{t-1}) = 0 \neq \sum_{j=1}^{\infty} \psi_j \epsilon_{t-j}$,

$V(y_t | I_{t-1}) = V(\eta_t | I_{t-1}) = h_{\eta,t} \neq \sum_{j=1}^{\infty} \psi_j^2 h_{t-j}$ (iA)

Dh: Die Spezifikation des ARMA Teils, falls vorhanden, ist wichtig. Zuerst ist also der ARMA Teil zu modellieren, im zweiten Schritt der GARCH Teil, oder beide simultan.

Die ACF unter bedingter Heteroskedastizität

Liegt Autokorrelation oder bedingte Heteroskedastizität vor, so kann die Varianz der Autokorrelationskoeffizienten beträchtlich größer sein, als die oben genannten $\frac{1}{T}$ unter homoskedastischen white noise.

Heteroskedastizitätsrobuster Schätzer der Varianz der Autokorrelationsfunktion
Der Prozess $\{y_t\}$ sei mittelwertstationär, $\mu_t = \mu$, unkorreliert und habe eine (multiple) symmetrische Verteilung.

Dann ist

$$v_\tau = \frac{\sum_{t=1}^{T-\tau} (y_t - \bar{y})^2 (y_{t+\tau} - \bar{y})^2}{[\sum_{t=1}^{T} (y_t - \bar{y})^2]^2}$$

ein konsistenter Schätzer für die Varianz von ρ_τ, des τ-ten Autokorrelationskoeffizienten.

Die ACF unter bedingter Heteroskedastizität

Es gilt asymptotisch:

$$\lim_{T \to \infty} v_\tau = \frac{1}{T} [1 + (kurtosis - 1) \rho_\tau^2]$$

Existiert die Kurtosis nicht, so existiert auch die Varianz von ρ_τ nicht.

Modellwahl

Hier können die Informationskriterien, AIC und SIC, genauso verwendet werden wie bei ARMA Modellen.