ARCH/GARCH Modeller
Forelæsningsnoter til Finansiel Økonometri

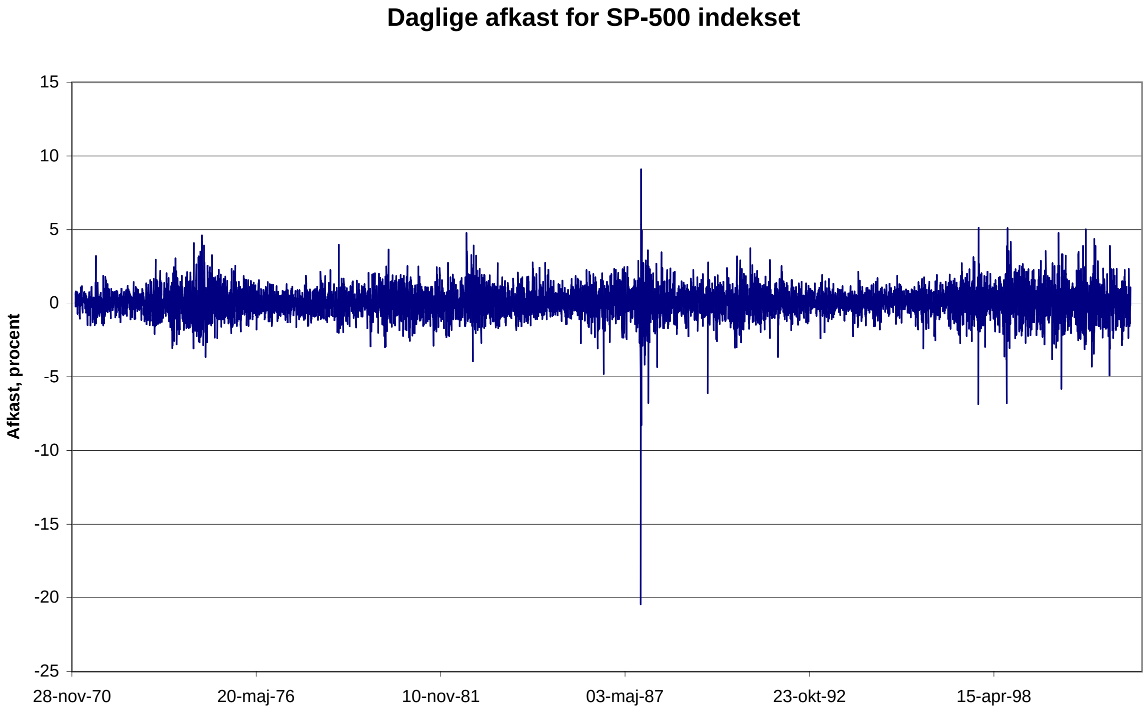
Jesper Lund

mail@jesperlund.com
http://www.jesperlund.com

17. april 2006

1 Indledning

ARCH/GARCH er statistiske modeller for tidsvarierende varianser.[1] Figuren nedenfor viser afkast på SP–500 aktieindekset. Det ses tydeligt at variansen varierer over tid, og at der er positiv autokorrelation i variansen.



Denne forelæsningsnote beskriver to standardmodeller til tidsvarierende varianser, nemlig GARCH(p, q), herunder specialtilfældet ARCH(q), og EGARCH(p, q). Det er primært modellernes matematiske struktur, som beskrives her. De to sidste afsnit kommer dog ind på estimation af modellernes parametre, herunder hvordan dette kan gøres ved hjælp af statistikprogrammet SAS.

2 ARCH(q) modellen

ARCH(q) modellen blev foreslået af Engle (1982) og er således den første model med tidsvarierende varians (betinget heteroskedasticitet). Hvis vi antager, at den betingede

[1] (G)ARCH er en forkortelse for (Generalized) Autoregressive Conditional Heteroskedasticity. I visse dele af literaturen, specielt den ældre del, bruges ARCH alene som navn for den model, som Engle (1982) foreslog, og som beskrives nedenfor i afsnit 2. I andre dele af literaturen, bl.a. Bollerslev et al. (1994), bruges ARCH som en generisk betegnelse for alle modeller, hvor den betingede varians afhænger af laggede fejlled, herunder GARCH og EGARCH modellerne, der i dette terminologi er specifikke ARCH parameteriseringer. En sådan moderat begrebsforvirring er meget typisk i en literatur i rivende udvikling.

fordeling for y_t er normalfordelingen, kan ARCH(q) modellen skrives på følgende form:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,\, \sigma_t^2) \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \tag{2}$$

hvor $\mu_t = E\,(y_t \,|\, Y_{t-1})$ med $Y_{t-1} = (y_{t-1}, y_{t-1}, \ldots, y_1)$ er den betingede middelværdi af y_t. For aktieafkast specificeres μ_t typisk som en konstant, dvs. $\mu_t = \mu$.[2]

ARCH(q) modellen antager, at den betingede varians σ_t^2 er en lineær funktion af tidligere kvadrerede fejlled. Store udsving af y_t medfører således en større varians σ_t^2, og dermed større sandsynlighed for store udsving i fremtiden. Dette matcher udviklingen i figuren med SP–500 aktieafkast i afsnit 1.

I nogle artikler [f.eks. Bollerslev et al. (1994)] ses en opskrivning af ARCH(q) modellen ved hjælp af lag-operatoren L, som er defineret ved egenskaben $L^n y_t = y_{t-n}$. Med lag-operator notationen kan ligning (2) skrives som:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_{t-1}^2, \text{ hvor } \alpha(L) = \left(\alpha_1 + \alpha_2 L + \alpha_3 L^2 + \cdots + \alpha_q L^{q-1}\right). \tag{3}$$

En varians skal per definition være positiv, og for at sikre at $\sigma_t^2 > 0$ altid holder, er det nødvendigt med følgende parameterrestriktioner:

$$\omega > 0 \text{ og } \alpha_i \geq 0 \text{ for alle } i. \tag{4}$$

Hvis $\sum_{i=1}^{q} \alpha_i < 1$, kan det vises at ARCH($q$) modellen er svagt stationær, og den ubetingede varians er givet ved:[3]

$$\mathrm{Var}\,(y_t) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i}. \tag{5}$$

Erfaringerne har vist, at det typisk er nødvendigt med en relativ høj orden q (mange lags) for at fitte udviklingen i den betingede varians for eksempelvis aktieafkast. Det betyder at man skal estimere mange parametre, og jo større q er, desto vanskeligere bliver det at sikre, at restriktionen $\alpha_i \geq 0$ holder. Disse problemer kan undgås ved at bruge GARCH(p, q) modellen, som beskrives i næste afsnit.

3 GARCH(p, q) modellen

GARCH(p, q) er udviklet af Bollerslev (1996), og den er en generalisering af ARCH(q) modellen i den forstand, at ARCH(q) er specialtilfældet GARCH($0, q$). GARCH(p, q) kan skrives på følgende form:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,\, \sigma_t^2) \tag{6}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2. \tag{7}$$

[2] Mere generelt kan man specificere en ARMA(m, s) model for μ_t, eller en regressionsmodel med forklarende variable.

[3] Ved hjælp af lag-operator notationen, kan ligning (5) skrives som $\mathrm{Var}\,(y_t) = \frac{\omega}{1-\alpha(1)}$, hvor $\alpha(1)$ skal læses som $\alpha(L)$ polynomiet i ligning (3) evalueret ved $L = 1$.

Som regel er det tilstrækkeligt med $p = q = 1$, dvs. GARCH$(1,1)$ modellen, hvor den betingede varians er

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2. \tag{8}$$

GARCH$(1,1)$ modellen kan omskrives til en ARCH(∞) model. Det gøres ved gentagen substitution af σ_{t-1}^2 med højresiden i (8). Forudsat at $0 < \beta < 1$, får vi

$$\begin{aligned}
\sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \beta\left(\omega + \alpha\varepsilon_{t-2}^2 + \beta\sigma_{t-2}^2\right) \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \beta\omega + \beta\alpha\varepsilon_{t-2}^2 + \beta^2\left(\omega + \alpha\varepsilon_{t-3}^2 + \beta\sigma_{t-3}^2\right) \\
&= \omega\left(1 + \beta + \beta^2 + \beta^3\right) + \alpha\left(\varepsilon_{t-1}^2 + \beta\varepsilon_{t-2}^2 + \beta^2\varepsilon_{t-3}^2 + \beta^3\varepsilon_{t-4}^2\right) + \beta^4\sigma_{t-4}^2 \\
&= \omega\left(1 + \beta + \beta^2 + \beta^3 + \cdots\right) + \alpha\left(\varepsilon_{t-1}^2 + \beta\varepsilon_{t-2}^2 + \beta^2\varepsilon_{t-3}^2 + \beta^3\varepsilon_{t-4}^2 + \cdots\right) \\
&= \frac{\omega}{1-\beta} + \sum_{i=1}^{\infty}\alpha\beta^{i-1}\varepsilon_{t-i}^2 \equiv \omega^* + \sum_{i=1}^{\infty}\phi_i\varepsilon_{t-i}^2 \qquad (9)
\end{aligned}$$

hvilket er en ARCH(∞) model med $\phi_i = \alpha\beta^{i-1}$. Hvis $\omega > 0$, $\alpha > 0$, $\beta > 0$ og $\alpha + \beta < 1$, opfylder modellen betingelserne for svag stationaritet, og den ubetingede varians kan udregnes til

$$\operatorname{Var}(y_t) = \frac{\omega/(1-\beta}{1 - \alpha\sum_{i=0}^{\infty}\beta^i} = \frac{\omega/(1-\beta}{1-\alpha/(1-\beta)} = \frac{\omega}{1-\alpha-\beta}. \tag{10}$$

Den generelle GARCH(p,q) model kan ligeledes omskrives til en ARCH(∞), men da der sjældent er brug for mere end $p = q = 1$, udelades disse detaljer. Betingelserne for svag stationaritet er $\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$, ligesom for GARCH$(1,1)$. Endvidere er det nødvendigt med nogle parameterrestriktioner for at sikre at $\sigma_t^2 > 0$. Disse betingelser er udledt af Nelson & Cao (1992).[4]

4 IGARCH(1,1) og EWMA modellen

IGARCH$(1,1)$ er et specialtilfælde at GARCH$(1,1)$ hvor $\alpha + \beta = 1$. Denne model er interessant af to årsager. For det første er IGARCH$(1,1)$ et realistisk eksempel på en stokastisk proces, som er stærkt stationær uden at være svagt stationær.[5]

For det andet er IGARCH$(1,1)$ identisk med en variansestimator, som ofte benyttes af praktikere, nemlig EWMA (Exponential Weighted Moving Average) modellen, der

[4]Positiv-varians betingelsen i Nelson & Cao (1992) tager udgangspunkt i ARCH(∞) omskrivningen af GARCH(p,q) modellen. I ARCH(∞) omskrivningen skal $\phi_i \geq 0$ gælde for samtlige i, og ud fra dette krav kan man formulere restriktioner på parametrene α_i og β_j i GARCH(p,q) modellen.

[5]Da $\alpha + \beta = 1$, eksisterer den ubetingede varians ikke, dvs. processen kan ikke være svagt stationær. Til gengæld har Nelson (1990) vist, at IGARCH(1,1) modellen er stærkt stationær. Bemærk at der ikke er nogen modstrid mellem de to udsagn, idet stærk stationaritet kun medfører svag stationaritet, hvis den ubetingede varians eksisterer. For IGARCH$(1,1)$ modellen er $\operatorname{Var}(y_t) = \infty$, dvs. den ubetingede varians eksisterer ikke.

blandt andet bruges i forbindelse med Value-at-Risk beregninger i RiskMetrics™ metoden.[6] EWMA bruger denne formel til at estimere variansen på tidspunkt t:

$$\sigma_t^2 = (1-\lambda)\sum_{i=1}^{\infty} \lambda^i \left(y_{t-i} - \bar{y}\right)^2, \tag{11}$$

hvor $0 < \lambda < 1$ er en smoothing parameter, og $\bar{y}$ er gennemsnittet af tidsserien y_t.[7] Den grundlæggende idé er at lægge større vægt på nye observationer i forhold til gamle, hvilket styres via parameteren λ. Eftersom $y_t - \bar{y}$ svarer til ε_t i de statistiske modeller ovenfor, ses EWMA formlen (11) imidlertid at være identisk med ARCH(∞) formuleringen af en IGARCH$(1,1)$ model, hvor $\omega = 0$ og $\beta = \lambda$.

5 EGARCH(p, q) modellen

EGARCH(p, q), som er en forkortelse for Exponential GARCH, er introduceret af Nelson (1991) som en model for logaritmen til den betingede varians. Det har den fordel, at det ikke er nødvendigt med parameterrestriktioner for at sikre at σ_t^2 er positiv ($\log x$ er kun defineret for $x > 0$, og EGARCH er en model for $\log \sigma^2$). Derudover har EGARCH modellen en asymmetrisk sammenhæng mellem ε_t og fremtidige varianser, mens det i GARCH modellen er ligegyldigt om ε_t er positiv eller negativ. Denne feature er vigtig, når man modellerer tidsvarierende varianser for aktieafkast, idet man her typisk finder at der er negativ korrelation mellem aktieafkastet og den fremtidige varians.

I Nelson (1991) og Bollerslev et al. (1994) defineres en EGARCH(p, q) model med betinget normalfordeling på denne måde:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,\, \sigma_t^2) && (12)\\ \log \sigma_t^2 &= \omega + \sum_{i=0}^{q} \alpha_i \left\{ \theta z_{t-1-i} + \gamma \left(|z_{t-1-i}| - \sqrt{2/\pi} \right) \right\} + \sum_{j=1}^{p} \beta_j \log \sigma_{t-j}^2, && (13) \end{aligned}$$

hvor $z_t = \varepsilon_t / \sigma_t$ er den standardiserede innovation (fejlled). Parameteren α_0 er normaliseret til $\alpha_0 \equiv 1$.[8] Bemærk at $E(|z|) = \sqrt{2/\pi}$, hvis z er normalfordelt med middelværdi 0 og varians 1.

I `SAS/ETS` kaldes ovenstående en EGARCH$(p, q+1)$ model, og der anvendes en anden normalisering, nemlig $\gamma = 1$.[9]

[6] Se `http://www.riskmetrics.com` og "RiskMetrics™— Technical Document" udgivet af J.P. Morgan/Reuters (1996) for mere information om RiskMetrics™og Value-at-Risk.

[7] Det ses jævnligt, at RiskMetrics m.fl. sætter $\bar{y} = 0$ i stedet for at estimere parameteren. Med daglige afkast er der som regel kun en meget beskeden forskel mellem at sætte $\bar{y} = 0$ og estimere parameteren som gennemsnittet af tidsserien y_t (da $\bar{y} \approx 0$).

[8] Uden denne normalisering kan man multiplicere γ og θ med en konstant c, og hvis man dividerer alle α_i med samme konstant c, får man en identisk model. Det betyder at modellens parametre kun er identificeret, hvis der foretages en normalisering, for eksempel $\alpha_0 = 1$.

[9] Der er desværre ikke 100% enighed i literaturen om hvordan EGARCH modellen skal defineres, men det skal understreges, at der naturligvis altid er tale om samme model, blot med en eventuel reparameterisering i forhold til definitionen i (13).

For aktieafkast er estimatet på EGARCH parameteren θ generelt negativt. Det betyder at variansen σ_t^2 er negativt korreleret med laggede (tidligere) aktieafkast. Dette fænomen kaldes "leverage" effekten i literaturen.

6 Estimation af GARCH og EGARCH modeller

Estimation af GARCH og EGARCH modeller gøres vha. maximum likelihood metoden. Begge modeller specificerer den betingede fordeling for y_t givet Y_{t-1}, og dermed kan man bruge *prediction error decomposition* princippet til at opstille log-likelihood funktionen:

$$\log L(\boldsymbol{\psi}) = \sum_{t=1}^{T} \log L_t(\boldsymbol{\psi}), \tag{14}$$

hvor $\boldsymbol{\psi}$ er en vektor som indeholder modellens parametre, og

$$\log L_t(\boldsymbol{\psi}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma_t^2 - \frac{1}{2}\varepsilon_t^2/\sigma_t^2, \quad \varepsilon_t = y_t - \mu_t. \tag{15}$$

Ligning (15) er logaritmen til den betingede tæthed for y_t givet Y_{t-1}. Maksimering af log-likelihood funktionen (14) kræver numerisk optimering, hvilket ligger uden for rammerne af denne forelæsningsnote. Mange statistikpakker, herunder `SAS/ETS`, kan heldigvis estimere GARCH og EGARCH modeller, så man behøver ikke nødvendigvis kende detaljerne i algoritmer til numerisk optimering af ikke-lineære funktioner.

Kovariansmatricen for ML estimatet $\hat{\boldsymbol{\psi}}$ beregnes normalt ved hjælp af denne formel:

$$\mathrm{Cov}\left(\hat{\boldsymbol{\psi}}\right) = \mathcal{H}^{-1}\mathcal{G}\mathcal{H}^{-1}, \tag{16}$$

hvor matricerne $\mathcal{H}$ og $\mathcal{G}$ er givet ved

$$\mathcal{H}\left(\hat{\boldsymbol{\psi}}\right) = \sum_{t=1}^{T} \frac{\partial^2 \log L_t(\boldsymbol{\psi})}{\partial\boldsymbol{\psi}\partial\boldsymbol{\psi}'} \Bigg| \boldsymbol{\psi} = \hat{\boldsymbol{\psi}} \tag{17}$$

$$\mathcal{G}\left(\hat{\boldsymbol{\psi}}\right) = \sum_{t=1}^{T} \frac{\partial \log L_t(\boldsymbol{\psi})}{\partial\boldsymbol{\psi}} \frac{\partial \log L_t(\boldsymbol{\psi})}{\partial\boldsymbol{\psi}'} \Bigg| \boldsymbol{\psi} = \hat{\boldsymbol{\psi}}. \tag{18}$$

Disse matricer kaldes henholdsvis Hessian matricen og OPG (Outer Product of the Gradient) matricen.

Det er muligt at lave ML estimation (MLE) af GARCH modeller med andre betingede fordelingsantagelser for ε_t end normalfordelingen, men det bruges ikke særligt ofte i praksis. MLE baseret på normalfordelingen har nemlig den store fordel, at estimatoren er robust overfor fejlspecifikation af fordelingen, hvis blot den betingede middelværdi $\mu_t = E\left[y_t \,|\, Y_{t-1}\right]$ og den betingede varians σ_t^2 er korrekt specificeret, se Bollerslev & Wooldridge (1992) for en udledning af denne egenskab. Ingen andre betingede fordelinger har denne robuste egenskab.

Ofte ses estimatorerne ligefrem benævnt QMLE, quasi maksimum likelihood estimation, for at understrege at man ikke bruger den korrekte betingede fordeling til at definere log-likelihood fordelingen.[10] Udtrykket for $\mathrm{Cov}(\hat{\boldsymbol{\psi}})$ i ligning (16) er ligeledes korrekt (konsistent i statistisk forstand), blot μ_t og σ_t^2 er korrekt specificeret.[11]

7 Estimation af GARCH og EGARCH i SAS

`PROC AUTOREG` i `SAS/ETS` kan estimere GARCH og EGARCH modeller. Syntaksen for QML estimation af en GARCH$(2,1)$ model med $\mu_t = \mu$ (konstant) er

```
/* Estimation af GARCH(2,1) model */
PROC AUTOREG DATA = mydat;
  MODEL ret = / GARCH=(P=2, Q=1)
                DIST=NORMAL METHOD=ML COVEST=QML;
RUN;
```

I dette eksempel estimeres en GARCH$(2,1)$ model for en variabel ved navn `ret` i SAS datasættet `mydat`. Options `METHOD=ML` og `DIST=NORMAL` udfører maksimum likelihood estimation baseret på normalfordelingen, og med option `COVEST=QML` beregnes robuste standardafvigelser for parameterestimaterne vha. formel (16).

Hvis man ønsker at estimere en EGARCH$(1,1)$ model i den "normale" forstand, altså en EGARCH$(1,2)$ model i `PROC AUTOREG`, bruges følgende programstump:

```
/* Estimation af EGARCH(1,1) model */
PROC AUTOREG DATA = mydat;
  MODEL ret = / GARCH=(P=1, Q=2, TYPE=EXP)
                DIST=NORMAL METHOD=ML COVEST=QML;
RUN;
```

Den eneste forskel er altså angivelsen af option `TYPE=EXP`, som betyder at der estimeres en EGARCH model i stedet for en GARCH model.

`PROC AUTOREG` har flere valgmuligheder ved estimation af GARCH og EGARCH modeller, og der henvises til SAS dokumentationen (online help) for en udtømmende beskrivelse af disse. Man kan blandt andet beregne en tidsserie for den betingede varians σ_t^2 og den standardiserede residual $z_t = \varepsilon_t/\sigma_t$.

[10] Bemærk at det grundlæggende problem er at den "korrekte" betingede fordeling ikke kendes. Hvis man eksempelvis specificerer en $t(\nu)$ fordeling, og dette er den korrekte betingede fordeling, får man en mere efficient (bedre) estimator. Men hvis innovationen $z_t = \varepsilon_t/\sigma_t$ ikke er $t(\nu)$-fordelt, er ML estimatoren ikke længere konsistent. Øget efficiens kommer altså med en risiko for inkonsistens, og dette trade-off falder generelt ud til fordel for den robuste estimator, altså QMLE baseret på normalfordelingen.

[11] Hvis likelihood funktionen er korrekt specificeret, kan det vises at $\mathcal{H} = \mathcal{G}$, hvilket i givet fald kunne bruges til at forenkle beregningen af $\mathrm{Cov}(\hat{\boldsymbol{\psi}})$. Eftersom man generelt estimerer GARCH og EGARCH modeller vha. QMLE og ikke MLE, bruges denne forenkling relativt sjældent i praksis. Husk at forskellen mellem QMLE og MLE alene er fortolkningsmæssig — det er præcist de samme estimater.

Litteratur

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., R.F. Engle & D.B. Nelson (1994), "ARCH Models," in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics, Vol. 4*, North-Holland, Amsterdam, Chapter 49

Bollerslev, T. & J.M. Wooldridge (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," *Econometric Reviews*, 11, 143–172.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987–1008.

Nelson, D.B. (1990), "Stationarity and Persistence in the GARCH(1,1) Model," *Econometric Theory*, 6. 318-334.

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

Nelson, D.B. & C.Q. Cao (1992), "Inequality Constraints in the Univariate GARCH Model," *Journal of Business and Economic Statistics*, 10, 229–335.