Module 9.3: Time Series Analysis with Python
Fall Term 2023

Week 5:

Volatility Modeling

HSLU Lucerne University of Applied Sciences and Arts

Outline in Weeks

1. Introduction; Descriptive Modeling
2. Returns; Autocorrelation; Stationarity
3. ARMA Models
4. Unit Roots; ARIMA Models
5. Volatility Modeling
6. Value at Risk
7. Cointegration

Outline

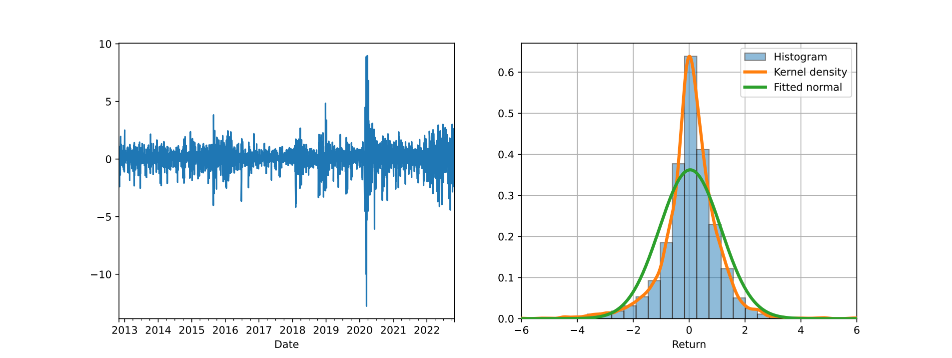
Goal

- Recall these stylized facts about asset returns:
  1. Lack of autocorrelation (efficient market hypothesis)
  2. Volatility clustering
  3. Distribution has heavy tails
  4. Leverage effects
- Goal today: model the last 3 of these, starting with the volatility clustering.

Volatility

- The *volatility* of an investment is a measure of its *risk*. Usually defined as the standard deviation of the return on the investment.
- Volatility is an important ingredient in:
  - portfolio selection;
  - risk management;
  - option pricing.
- Daily financial returns display *volatility clustering*: periods of high volatility alternate with more tranquil periods.
- In other words: large (in absolute value) returns tend to be followed by large (in absolute value) returns.
- This forms the basis for the *autoregressive-conditional heteroskedasticity* model (ARCH; Engle, 1982) and the *generalized ARCH* model (GARCH; Bollerslev, 1986).

Example: Daily Returns on the S&P 500



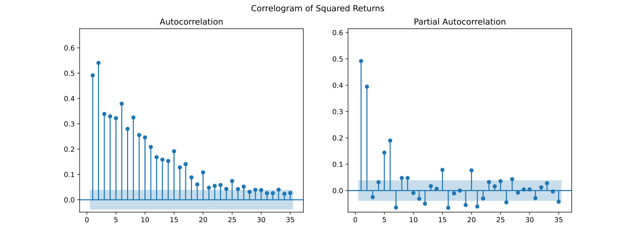
Reminder: Parameters vs. sample values

- We usually write σ for the standard deviation of, e.g., a normally distributed variable.
- σ is a *parameter* and therefore unknown.
- The best we can hope for is to *estimate* it, usually with the *sample standard deviation* s.
- With stock returns, the standard deviation (or *volatility*) changes over time, due to *volatility clustering*.
- We write σ_t for the volatility in period t.
- Note that σ_t is *unobserved*. The best we can do is *estimate* it. We'll write $\widehat{\sigma}_t$ for this estimate.
- Today, we'll mostly discuss different methods of estimation.

Detecting Volatility Clustering (I)

- Since volatility clustering means that large returns tend to be followed by large returns, it is possible to detect it by inspecting the correlogram of the squared returns.

Example: correlogram of squared S&P500 returns.



- Clearly, there is a lot of predictability in squared returns (unlike returns themselves).

Detecting Volatility Clustering (II)

- Besides relying on the correlogram (or the associated Q-tests, see exercises), a formal test for volatility clustering is Engle's *ARCH-LM* test.
- The test is designed to work with regression residuals, not raw returns. Hence, we start by regressing the returns on an intercept (this is equivalent to de-meaning the returns).
- The ARCH-LM test is based on the auxiliary regression

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \ldots + \gamma_m \hat{u}_{t-m}^2 + e_t.$$

- The lag length m is chosen by the user, e.g., 5 for daily data.
- The test statistic is $T \cdot R_{aux}^2$ and has a $\chi^2(m)$ distribution under $H_0 : \gamma_1 = \cdots = \gamma_m = 0$ (no volatility clustering).

Example: ARCH-LM test for the S&P500 (see exercises for details)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.373
Model:                            OLS   Adj. R-squared:                  0.372
Method:                 Least Squares   F-statistic:                     298.0
Date:                Wed, 04 Oct 2023   Prob (F-statistic):          8.83e-251
Time:                        17:19:09   Log-Likelihood:                -7173.1
No. Observations:                2512   AIC:                         1.436e+04
Df Residuals:                    2506   BIC:                         1.439e+04
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3179      0.088      3.618      0.000       0.146       0.490
x1             0.2996      0.020     15.157      0.000       0.261       0.338
x2             0.3959      0.021     19.170      0.000       0.355       0.436
x3            -0.0872      0.022     -3.955      0.000      -0.130      -0.044
x4            -0.0139      0.021     -0.671      0.502      -0.054       0.027
x5             0.1438      0.020      7.272      0.000       0.105       0.183
==============================================================================
```

Here, the dependent variable `y` refers to $\hat{u}_t^2$, the regressor `x1` refers to $\hat{u}_{t-1}^2$, etc. The test statistic is $T \cdot R^2 \approx 937$, much larger than the critical value 11.07. The null of no volatility clustering is thus clearly rejected.

Outline

Historical Volatility

- A first simple estimator is *historical volatility*, i.e., the sample standard deviation of the most recent m observations (often $m = 250$, one year).
- If $r_t = \ln P_t - \ln P_{t-1}$ denotes the daily log-return, then

$$\widehat{\sigma}^2_{t+1,HIST} = \frac{1}{m}\sum_{j=0}^{m-1} r^2_{t-j}.$$

  (Typically the average return is relatively close to zero). This is an estimate of the squared volatility over day $t+1$, made at the end of day t.
- Main disadvantages:
  - either noisy (small m), or reacts slowly to new information (large m);
  - “ghosting” feature: large shock leads to higher volatility for exactly m periods, then drops out.

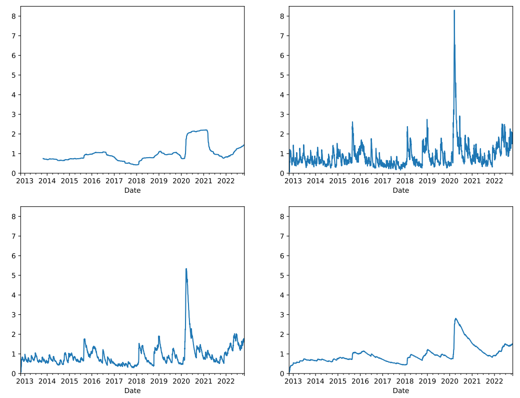
RiskMetrics

- Problems with historical volatility are addressed by replacing equally weighted moving average by an *exponentially* weighted moving average (EWMA), also used in JPMorgan's *RiskMetrics* system:

$$\begin{aligned} \widehat{\sigma}^2_{t+1,EWMA} &= (1-\lambda)\sum_{j=0}^{\infty}\lambda^j r^2_{t-j} \\ &= \lambda\widehat{\sigma}^2_{t,EWMA} + (1-\lambda)r_t^2, \qquad 0<\lambda<1. \end{aligned}$$

- This means that observations further in the past get a smaller weight.
- In practice we do not have $r_{t-\infty}$, but the second equation can be started up by an initial estimate / guess $\sigma^2_{0,EWMA}$.
- The larger λ, the stronger the persistence of shocks (large returns).
- For daily data, RiskMetrics recommends $\lambda = 0.94$.

Example: S&P500 volatility, historical and EWMA ($\lambda = 0.8, 0.94, 0.99$)



Outline

The ARCH Model

- The first-order *autoregressive-conditional heteroskedasticity* (ARCH(1)) model, due to Engle (1982), for a return r_t with mean zero is
$$\sigma_{t+1}^2 = \omega + \alpha r_t^2.$$
- In practice, we need to allow for $\mathbb{E}[r_{t+1}] = \mu_{t+1} \neq 0$. Then $r_{t+1} = \mu_{t+1} + u_{t+1}$, and the model becomes
$$\sigma_{t+1}^2 = \omega + \alpha u_t^2.$$

The ARCH Model

- When trying to estimate ARCH models one might find that more lags are needed, leading to ARCH(q):
$$\sigma^2_{t+1} = \omega + \alpha_1 u_t^2 + \ldots + \alpha_q u^2_{t-q+1}.$$
- *Note*: Variances must be positive, therefore we need to impose $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, q$.
- It can be shown that an ARCH(q) models corresponds to an AR(q) for the squared returns. Thus, we could determine the order from the correlogram of the squared returns: SPACF should cut off after q lags.
- In the example above, we might conclude that we need an ARCH(6) model.

The GARCH Model

- A simpler structure than ARCH(q) is an ARMA(1,1) for r_t^2 or u_t^2, which leads to the *generalized ARCH* model of orders (1,1) (GARCH(1,1)), due to Bollerslev (1986):

$$\sigma_{t+1}^2 = \omega + \alpha u_t^2 + \beta \sigma_t^2, \qquad \omega > 0, \alpha \geq 0, \beta \geq 0.$$

- *Advantage*: Flexible structure with only 3 parameters to estimate.

The GARCH Model

- The GARCH(1,1) model is stationary if the unconditional ("average") variance $\sigma^2 = \mathbb{E}[\sigma_t^2]$ is positive, constant and finite.
- This requires

$$\begin{aligned}\sigma^2 = \mathbb{E}[\sigma_{t+1}^2] &= \omega + \alpha\mathbb{E}[u_t^2] + \beta\mathbb{E}[\sigma_t^2] \\ &= \omega + \alpha\sigma^2 + \beta\sigma^2.\end{aligned}$$

- Hence, provided that $\alpha + \beta < 1$ (the *stationarity condition*),

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}.$$

- The nonstationary model with $\alpha + \beta = 1$ is called *integrated GARCH* (IGARCH): infinite variance, no mean-reversion in volatility.
- Notice that an IGARCH with $r_t = u_t$, $\omega = 0$, $\beta = \lambda$, and $\alpha = (1 - \lambda)$ is just the RiskMetrics model.

The GARCH Model

Some other properties:

- The ACF and PACF of r_t^2 in case of stationary GARCH(1,1) are both exponentially decaying, no cut-off point.
- The *standardized returns*

$$z_{t+1} = \frac{r_{t+1} - \mu_{t+1}}{\sigma_{t+1}}$$

  satisfy $E(z_{t+1}) = 0$ and $\mathrm{var}(z_{t+1}) = 1$. Therefore the model may be formulated as

$$\begin{aligned} r_{t+1} &= \mu_{t+1} + u_{t+1} = \mu_{t+1} + \sigma_{t+1} z_{t+1}, \\ \sigma_{t+1}^2 &= \omega + \alpha u_t^2 + \beta \sigma_t^2. \end{aligned}$$

- Often it is assumed that z_t are i.i.d. as $N(0, 1)$.
- Even if $z_t \sim N(0, 1)$, it can be shown that varying σ_t implies that r_t has non-normal distribution, with higher kurtosis.

The GARCH(p, q) Model

- The GARCH(1, 1) model can be extended to the GARCH(p, q) model

$$\sigma_{t+1}^2 = \omega + \alpha_1 u_t^2 + \cdots + \alpha_q u_{t-q+1}^2 + \beta_1 \sigma_t^2 + \cdots + \beta_p \sigma_{t-p+1}^2$$

  although in practice, this is rarely necessary.
- The model is stationary if $\sum_{i=1}^p \beta_i + \sum_{i=1}^q \alpha_i < 1$, and the unconditional variance is

$$\frac{\omega}{1 - \sum_{i=1}^p \beta_i - \sum_{i=1}^q \alpha_i}.$$

Outline

Estimation of GARCH Models

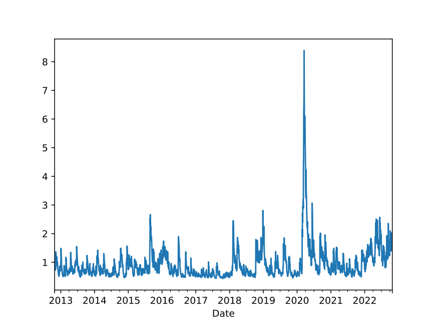
- GARCH cannot be estimated by ordinary least-squares (because σ_t^2 is not observed).
- Such models are estimated by *maximum likelihood*: the joint density of the observations $\{r_1, \ldots, r_T\}$ is maximized with respect to the parameters.
- Maximization of $\log L$ can be done by numerical optimization algorithms. By default, the `arch` package for Python does this under the assumption of normality.
- If we are not sure that the z_t's are normally distributed, then we may still use the same estimation technique. This is called *quasi-maximum likelihood estimator*.
- However, we need to construct standard errors via a more robust method (*Bollerslev-Wooldridge standard errors*). `arch` does this by default.

Example: `arch` output, estimated GARCH model for S&P500

```
                    Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:             log_return   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:               -3119.24
Distribution:                  Normal   AIC:                           6246.48
Method:            Maximum Likelihood   BIC:                           6269.80
No. Observations:                 2517
Date:                Wed, Oct 04 2023   Df Residuals:                     2516
Time:                        17:23:18   Df Model:                            1
Mean Model
==============================================================================
coef    std err          t      P>|t|    95.0% Conf. Int.
------------------------------------------------------------------------------
mu             0.0803  1.410e-02      5.697  1.219e-08 [5.269e-02,  0.108]
Volatility Model
==============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
------------------------------------------------------------------------------
omega          0.0447  9.834e-03      4.551  5.346e-06 [2.548e-02,6.402e-02]
alpha[1]       0.2217  3.185e-02      6.963  3.326e-12     [  0.159,  0.284]
beta[1]        0.7437  2.864e-02     25.970 1.094e-148     [  0.688,  0.800]
==============================================================================

Covariance estimator: robust
```

Example: Estimated GARCH volatility of S&P500

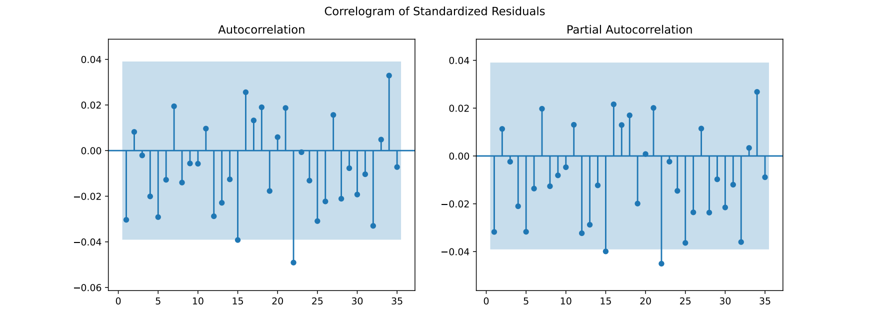


Outline

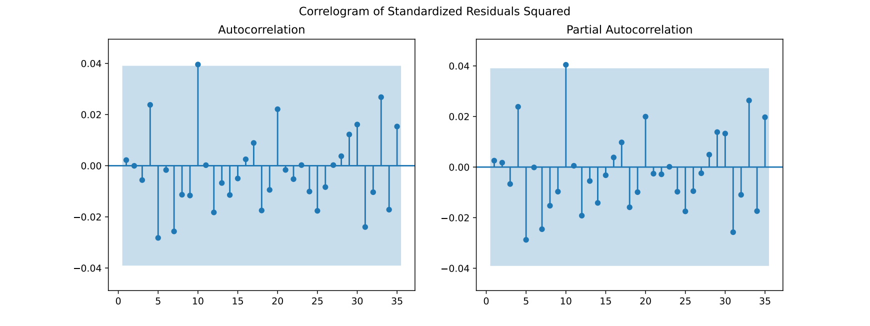
Testing GARCH Models

- Diagnostic tests are based on the *standardized residuals* $\hat{z}_t := \hat{u}_t/\hat{\sigma}_t$. If μ_t and σ_t are correctly specified, we should find no autocorrelation in $\hat{z}_t$ and $\hat{z}_t^2$.
- Therefore, the model can be tested using Q-statistics for $\hat{z}_t$ or $\hat{z}_t^2$.
- Lagrange-Multiplier (LM) test against ARCH, which is obtained by $T \cdot R^2$ in the regression
$$\hat{z}_t^2 = \gamma_0 + \gamma_1 \hat{z}_{t-1}^2 + \ldots + \gamma_m \hat{z}_{t-m}^2 + e_t.$$
- To test for normality of z_t, we can use the Jarque-Bera test based on the skewness and kurtosis of $\hat{z}_t$.

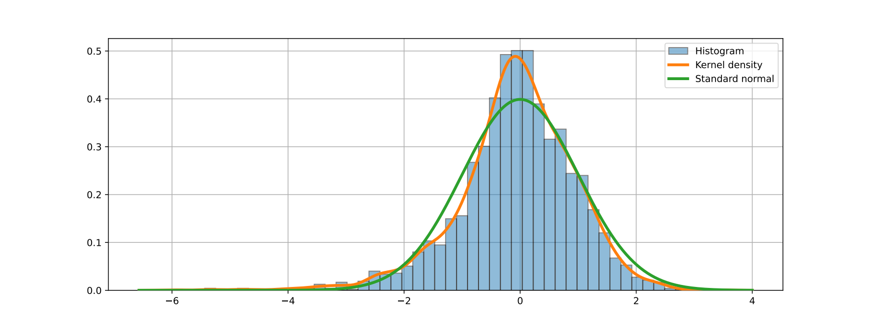
Example: Correlogram of standardized residuals for the S&P500



Example: Correlogram of squared standardized residuals for the S&P500



Example: Histogram of standardized residuals for the S&P500



The Jarque-Beta test rejects with a *p*-value of essentially zero, rejecting normality (see exercises).

Outline

Asymmetry and the News Impact Curve

- The *news impact curve* (NIC) is the effect of u_t on σ_{t+1}^2, keeping σ_t^2 and the past fixed.
- For GARCH(1,1), this is the parabola $NIC(u_t|\sigma_t^2 = \sigma^2) = A + \alpha u_t^2$, with $A = \omega + \beta\sigma^2$. This has a minimum at $u_t = 0$, and is symmetric around that minimum.
- For equity, a large negative shock is expected to increase volatility more than a large positive shock, because of *leverage effect*:

$$
\begin{aligned}
& \downarrow \text{value of firm's stock} \\
\Rightarrow\quad & \downarrow \text{equity value of the firm} \\
\Rightarrow\quad & \uparrow \text{debt-to-equity ratio} \\
\Rightarrow\quad & \text{shareholders (as residual claimants) perceive future cashflows} \\
& \text{as more risky.}
\end{aligned}
$$

- Multiple extensions exist to deal with this issue. Here we focus on Glosten, Jagannathan and Runkle's GJR-GARCH model.

GJR-GARCH (or TARCH, threshold GARCH)

The GJR-GARCH(1,1) model is

$$\sigma_{t+1}^2 = \omega + \alpha u_t^2 + \gamma u_t^2 \mathrm{I}_t + \beta \sigma_t^2.$$

where

$$\mathrm{I}_t = \begin{cases} 1 & \text{if} \quad u_t < 0 \\ 0 & \text{if} \quad u_t \geq 0 \end{cases},$$

and u_t / σ_t has a symmetric distribution.
Properties:

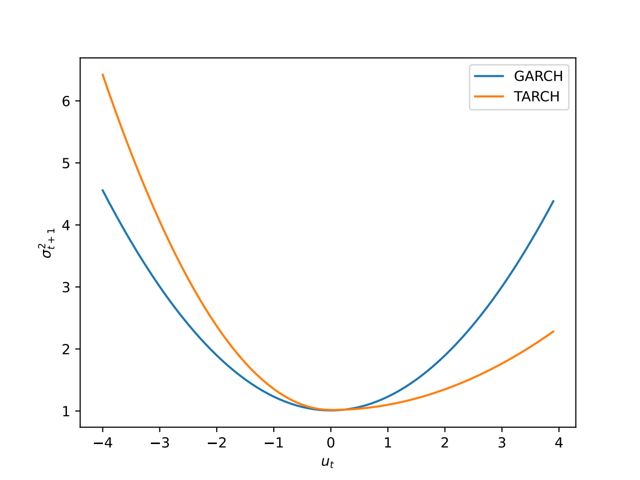
- NIC is asymmetric if and only if $\gamma \neq 0$; leverage effect if $\gamma > 0$;
- σ_t^2 is positive if $\omega > 0$, $\alpha \geq 0$, $\gamma \geq 0$, $\beta \geq 0$;
- u_t^2 is stationary if $0 \leq \alpha + \frac{1}{2}\gamma + \beta < 1$, with unconditional variance $\sigma^2 = \omega / \left[1 - \alpha - \frac{1}{2}\gamma - \beta\right]$.

Example: `arch` output, estimated TARCH model for S&P500

```
                     Constant Mean - GJR-GARCH Model Results
==============================================================================
Dep. Variable:                log_return   R-squared:                       0.000
Mean Model:                Constant Mean   Adj. R-squared:                  0.000
Vol Model:                     GJR-GARCH   Log-Likelihood:               -3088.09
Distribution:                     Normal   AIC:                           6186.18
Method:               Maximum Likelihood   BIC:                           6215.34
No. Observations:                    2517
Date:                   Wed, Oct 04 2023   Df Residuals:                     2516
Time:                           18:38:45   Df Model:                            1
Mean Model
==============================================================================
coef    std err          t      P>|t|      95.0% Conf. Int.
------------------------------------------------------------------------------
mu             0.0475  1.357e-02      3.503  4.609e-04 [2.094e-02,7.415e-02]
Volatility Model
==============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
------------------------------------------------------------------------------
omega          0.0419  8.680e-03      4.832  1.352e-06 [2.493e-02,5.895e-02]
alpha[1]       0.0831  4.212e-02      1.974  4.843e-02   [5.741e-04,  0.166]
gamma[1]       0.2547  5.416e-02      4.703  2.561e-06     [  0.149,  0.361]
beta[1]        0.7569  3.189e-02     23.736 1.541e-124     [  0.694,  0.819]
==============================================================================

Covariance estimator: robust
```

Example: NIC of GARCH and TARCH models for S&P500



Outline

Volatility Forecasting

- GARCH models directly provide forecasts of next day's volatility:

$$\widehat{\sigma}^2_{t+1} = \hat{\omega} + \hat{\alpha}\hat{u}^2_t + \hat{\beta}\widehat{\sigma}^2_t.$$

- Multi-period forecasts can be constructed recursively. In principle, one would use

$$\widehat{\sigma}^2_{t+2} = \hat{\omega} + \hat{\alpha}\hat{u}^2_{t+1} + \hat{\beta}\widehat{\sigma}^2_{t+1},$$

  but $\hat{u}^2_{t+1}$ is unobserved.
- Solution: replace $\hat{u}^2_{t+1}$ with its estimate, $\widehat{\sigma}^2_{t+1}$.
- Result:

$$\widehat{\sigma}^2_{t+2} = \hat{\omega} + \left(\hat{\alpha} + \hat{\beta}\right)\widehat{\sigma}^2_{t+1}.$$

Outline

Learning Goals

Students

- can use appropriate tests to detect volatility clustering,
- are able to estimate, interpret, and forecast the various models (historical volatility, RiskMetrics, (G)ARCH, TARCH), and to apply diagnostic tests to the standardized residuals,
- and understand the concept of leverage, and the NIC.

Homework

- Exercise 5
- Questions 1 and 3 from Chapter 9 of Brooks (2019)