# 4 non-parametric estimation and mle

0. Load `s4_data.txt` into RStudio. This dataset contains the following series:

| name | description | source |
|---|---|---|
| *nodur – other* | simple returns of 10 industry portfolios | Kenneth French's website |
| *mktrf – hml* | the Fama-French factors | |
| *mkt* | mktrf + rf | |
| *rf* | the risk-free rate | |
| *dol* | return on the trade weighted dollar index (vs. major currencies) | St.Louis Fed |
| *gold* | return on the S&P GSCI gold index | Bloomberg |

All series are in % per day.

1. Let's see if linear models make sense in the world of finance. They are a special form of the following functional relation between two series:

$$y_t = f(x_t) + u_t \tag{1}$$

A way to estimate this relation non-parametrically (i.e. without specifying the function $f(\cdot)$) is called the Nadaraya-Watson estimator. The idea is simple: the best guess for $y(x_t)$ is an average of $y$ in the vicinity of that $x_t$. You will find more details in the Lecture notes, but let's now estimate the relation between *gold* and the market (*mkt*).

2. Gold is claimed to possess safe haven properties. If these claims are justified, it should (1) pay off when things go south on the stock market, and probably (2) pay off even more when the stock market crashes severely. Nadaraya-Watson comes in handy when such non-linearities are to be detected. In order to determine the relashionship between the two variables, plot a least squares line along with the nonparametric regression estimates. Judging by the plot, is a linear relation between *gold* and *mkt* a good enough model? Approximately, what would be the value of *gold* if the market falls by 2% a day and when it does by 4%?

3. Let's shift gears. MLE is a way of estimating a set of parameters, making an assumption about the underlying distribution. Estimate the linear regression model

$$R_t^{\text{enrgy}} = \beta_0 + \beta_1 R_t^{\text{mkt}} + u_t, \tag{2}$$

assuming $u_t \sim N(0, \sigma^2)$, or equivalently $R_t^{\text{enrgy}} \sim N(\beta_0 + \beta_1 R_t^{\text{mkt}}, \sigma^2)$.

You will be effectively estimating three parameters: $\beta_0$, $\beta_1$ and $\sigma$:

$$\mathcal{L}(R^{\text{enrgy}}|\beta_0, \beta_1, \sigma) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(R_t^{\text{enrgy}} - (\beta_0 + \beta_1 R_t^{\text{mkt}}))^2}{2\sigma^2}\}, \tag{3}$$

4. Make sure that the OLS estimation produces similar values.

5. Normality of residuals is just an assumption: maybe $u_t$ is Student $t$-distributed with $\nu$ degrees of freedom? Modify the model and re-estimate the coefficients. What does the best fit for $\nu$ say about the normality of $u_t$? Also, how do the coefficients compare to the previously estimated ones?

6. GARCH models have enjoyed tremendous popularity due to their (1) ability to capture some important stylized facts about stock returns[1], and (2) being easy to estimate with MLE. The idea behind GARCH models is that returns are random numbers generated by the variance. This setting for tomorrow depends upon today's setting and the return generated today. If today's variance is high, tomorrow's variance will be high. If today a huge return is generated (implying a high innovation), tomorrow's variance will be high. Mathematically:

$$\begin{aligned} r_t &= \mu + u_t \\ u_t &\sim N(0, {\sigma_t}^2) \\ \text{(alternative formulation) } u_t &= z_t \sigma_t \text{ where } z_t \sim N(0,1) \\ \sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

   If this is true, the likelihood of observing a sample of $T$ returns is:

$$\mathcal{L}(r|\mu,\omega,\alpha,\beta) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\{-\frac{(r_t-\mu)^2}{2\sigma_t^2}\}, \tag{4}$$

   Estimate this model on *enrgy*.

7. Do the same estimation using the `rugarch` package. Check whether you get similar results.

8. Extract the conditional volatility estimates and plot them on the timeline. Explore also the wide variety of plots offered by `uGARCHfit` to get a deeper look at the underlying structure of the model.

9. The conditional variances can be used for risk management purposes: being in 2017-07-28, estimate the 90% Value-at-Risk for the next day based on the historical sample volatility estimate and on the conditional volatility from your GARCH(1,1) model.

10. It is another stylized fact that volatility surges when markets are bearish. This asymmetric effect is captured in the GJR[2] model. The parametrization in the `rugarch` package differs slightly from the Lecture notes in that $\delta$ takes a value of 1 if the previous shock was *negative*, not positive. With $\gamma < 0$, volatility increases less in response to a negative shock. Estimate the GJR model for *enrgy*.

[1]e.g. volatility clustering and negative correlation between volatility and returns
[2]Glosten, Jagannathan, and Runkle (1993)

homework

1. Load `s4_data.txt` into RStudio. We'll start with estimating the functional relation between the expected return on the *telcm* portfolio and the dollar appreciation rate:

$$R_t^{\text{telcm}} = f(R_t^{\text{dol}}) + u_t \tag{5}$$

   Estimate it using the Nadaraya-Watson estimator and plot the fitted values vs. the regressor, along with an OLS line. Is the absolute effect on *telcm* larger when the dollar appreciates by 1% or when it depreciates by the same amount? **(2 points)**

2. Fill in the table in the answer sheet, indicating if we should see positive or negative returns for different values of the dollar return. **(2 points)**

3. If we chose a higher (wider) bandwidth, what will happen to our plot? **(1 point)**

4. Let's switch to MLE. Say, you believe that the true model of the relation between *telcm* and *mkt* is as follows:

$$\begin{aligned} R_t^{\text{telcm}} &= \beta_0 + \beta_1 R_t^{\text{mkt}} + u_t \\ u_t &\sim L(0, s), \end{aligned} \tag{6}$$

   where $L(0, s)$ is the zero-mean Laplace distribution (link to wiki) parametrized with the scale parameter $s$:

$$p(u_t) = \frac{1}{2s} \exp\left(-\frac{|u_t|}{s}\right). \tag{7}$$

   Estimate the three parameters with MLE, using the log-liklihood. Report $\beta_0$, $\beta_1$, and their respective standard errors. **(3 points)**

5. In order to test for ARCH effects, perform a Ljung-Box test with 12 lags on the squared returns of *telcm*. Give your conclusion at the 5% significance level. **(2 points)**

6. Assuming that innovations follow a normal distribution, estimate an ARCH(5) model on the returns of *telcm*. We will name it **model 1**. Comment on the normal QQ plot (plot `9` using the `uGARCHfit` plot method with the `which` argument). **(2 points)**

   *Hint: Note that an ARCH(5) model is equivalent to a GARCH(5,0) model.*

7. Re-estimate the same model but with a t-distribuition for the residuals instead of a normal one (**model 2**). Report the `shape` coefficient from the estimation output. Relate it to your observation in the previous question. **(2 points)**

8. Finally, estimate a GJR model assuming a t-distribution for the innovations (**model 3**). Does the variance go up or down after a negative return on the previous day? Is this move statistically significant at the 95% level? **(1 point)**

9. You want to assess whether the ARCH effects were correctly modeled. Ideally, the standardized squared residuals should not exhibit autocorrelation. Why? Compare the "ACF of squared standardized residuals" plots (plot `11`) from models 2 and 3. **(1 point)**

10. Taking a look at the conditional volatility plot from model 3, what is approximately the maximum daily 95% Value-at-Risk you could have calculated since 1990 based on the assumption that the returns are conditionally normally distributed with zero mean and the volatility is the one estimated in model 3? **(2 points)**

11. Compare your three models in terms of the AIC. Comment on the ranking obtained. **(2 points)**