TIF20216: Advanced Time Series Econometrics

Assignment Weeks 3–4

Instruction:

- This is a mandatory, individual assignment (not a group assignment).
- This assignment consists of 2 parts: one for each week. The deadline for handing both parts is Tuesday, February 17, 23:59.
- The assignment solution should be submitted as a single zip file, including computer codes and written reports no longer than 4 pages (one for each part). Matlab, Python and R code is acceptable. The code must be clear and well-documented. Please be specific about any packages, modules or toolboxes you use and their source.
- Late submission, up to 3 days, is allowed subject to penalty. We deduct 1% of the grade for each hour delayed.

Part 1: High-Dimensional Volatility Models

Carry out a Monte Carlo experiment to study the need and succes of shrinkage estimators of correlation / variance estimators, where we focus on data with an **equicorrelated** structure, i.e., where the (unconditional) correlation matrix R of a vector process $\{y_t\}_{t=1}^T$ satisfies $\mathrm{corr}(y_{it}, y_{jt}) = \rho$ for all $i \neq j$. It is easily checked that the population eigenvalues of such a matrix are $\nu_1 = \ldots = \nu_{N-1} = 1 - \rho$, and $\nu_N = 1 + (N-1)\rho$ (i.e., the eigenvalue $1 - \rho$ has a multiplicity of $N - 1$). We will assume that all unconditional variances are equal to 1, such that $\Sigma = R$.

1. Using $N = 50$ and $T = 100$, simulate the $T \times N$ data matrix Y of i.i.d. vectors with mean zero and (equicorrelated) variance matrix Σ; you may choose a more or less realistic value for equity returns for ρ, e.g. $\rho = 0.5$. For each replication, calculate the sample variance matrix $S(Y)$, and its ascending eigenvalues $\lambda_1 \leq \ldots \leq \lambda_N$. Average the resulting vector of eigenvalues over B Monte Carlo replications (e.g. $B = 1000$) and compare the result with the population eigenvalues in a plot (with a logarithmic scale for the vertical axis) of the eigenvalues $(\nu_i, \bar{\lambda}_i)$ against i. Comment on what you find: is shrinkage needed, and would linear shrinkage be sufficient, or is non-linear shrinkage needed?
2. Apply the linear shrinkage method of Ledoit and Wolf (2004)[1], in each of the Monte Carlo replications, and take again the Monte Carlo average. Compare the resulting eigenvalues with the uncorrected sample eigenvalues: is there an improvement?
   If you prefer, you may replace linear shrinkage by the non-linear shrinkage method of Ledoit and Wolf (2012). You may use the code made available at Michael Wolf's webpage[2], and adapt it to your needs. You do not need to apply *both* shrinkage methods.
3. If we *knew* that the data displays equicorrelation, then clearly we could have exploited that knowledge: an obvious choice would be to replace the first $N - 1$ eigenvalues by their

[1] `https://doi.org/10.1016/S0047-259X(03)00096-4`
[2] `https://www.econ.uzh.ch/en/people/faculty/wolf/publications.html#Programming_Code`

average (over $i = 1, \ldots, N-1$), and leave the largest eigenvalue unshrunk. Investigate the effectiveness of this method, which is arguably an oracle result (it requires knowledge of the structure of Σ).

4. Now embed each of these correction methods in a scalar BEKK/VEC model

$$\Sigma_t = (1 - \alpha - \beta)\Sigma + \alpha y_{t-1} y_{t-1}' + \beta \Sigma_{t-1},$$

where α and β are scalar parameters satisfying $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$. In a new Monte Carlo experiment, generate data $\{y_t\}_{t=1}^T$ from this model with $\alpha = 0.05$, $\beta = 0.93$, and Σ having an equicorrelation structure with unit unconditional variances. For each replication, estimate the unconditional variance matrix Σ by each of the methods used earlier, and estimate α and β using QML with $\widehat{\Sigma}$ fixed to the sample variance matrix (with or without shrinkage). Next, calculate the sample variance of the minimum-variance portfolio $\widehat{w}_t' y_t$, where $\widehat{w}_t$ is obtained from your fitted $\widehat{\Sigma}_t$, for each of the different methods to estimate $\widehat{\Sigma}_t$; see Engle, Ledoit and Wolf (2019) for the definition of w_t from Σ_t. Average again over the Monte Carlo replications, and compare the results. Is shrinkage needed for good dynamic portfolio optimization? Experiment with N and T (and their ratio $c = N/T < 1$), to see how sensitive your conclusions are to N and T.

Part 2: High-Frequency Financial Econometrics

Download the historical time series for daily volatility of the S&P 500 index (SPY) from Dacheng Xiu's Risklab website:

`https://dachxiu.chicagobooth.edu/#risklab`

Note that the data-set does not contain the original high-frequency data, just the realized measures. You also need to download daily returns on S&P 500 index from CRSP (via WRDS), Yahoo Finance, or other resources you are familiar with. We will refer to $\{r_t\}_{t=1}^T$ as the daily returns series, and $\{\mathrm{RV}_t\}_{t=1}^T$ as the realized variances. Choose a motivated subset of the available data (which should, however, contain enough observations for accurate estimation, so at least 10 years of daily data).

1. Select, estimate and test two GARCH-type models (could also be GJR-GARCH, EGARCH, or a related extension) for r_t;
   - In the first, you should add RV_{t-1}, or a transformation such as $\log \mathrm{RV}_{t-1}$, as explanatory variable to the variance equation for $h_t = \mathrm{Var}_{t-1}(r_t)$.
   - The second model should be nested in this first model, and obtained by removing the RV_{t-1} term(s) from the first model (putting their coefficients to zero). The second model should therefore be a standard GARCH-type model based on returns data only.

2. Report the estimation output from both models, and answer the following questions:
   - Is the likelihood of the extended model significantly higher? In other words, does the likelihood ratio test for the second model, as a restricted version on the first model, reject the null hypothesis? Make sure to state your null hypothesis carefully.
   - Does the inclusion of RV_{t-1} take over the role of ε_{t-1}^2 (or other functions of ε_{t-1}); i.e., are these ε_{t-1}-related terms no longer significant when including RV_{t-1}?

3. Split the dataset into a training and a test set. For instance, you may use the earliest 80% data for training and the latest 20% data for testing. Fit the following RV models to the training set:
   - LogHAR (Corsi, 2009):
   $$\log \mathrm{RV}_{t+1} = \beta_0 + \beta_D \log \mathrm{RV}_t + \beta_W \log \mathrm{RV}_{t-5,t} + \beta_M \log \mathrm{RV}_{t-22,t} + \varepsilon_{t+1},$$
   where $\mathrm{RV}_{t-h,t} = (\mathrm{RV}_{t-h+1} + \ldots + \mathrm{RV}_t)/h$;
   - The nonlinear regression model
   $$\log \mathrm{RV}_{t+1} = f(\mathrm{RV}_t, \mathrm{RV}_{t-5,t}, \mathrm{RV}_{t-22,t}) + \varepsilon_{t+1},$$
   estimating the nonlinear regression function f by a machine learning method, such as a random forest or least-squares boosting.

4. Use the fitted models to make predictions on the test set. Compare the prediction errors (of the logarithm of realized variance) and comment on your findings. Repeat the analysis for the entire sample (from 1996) and the latest ten years' data, respectively: are the conclusions similar, or not? Explain the difference or similarity.