The Cooper Union Department of Electrical Engineering
Prof. Fred L. Fontaine
ECE478 Financial Signal Processing
Problem Set V: Financial Time-Series Analysis
April 1, 2022

Here, $\gamma(m)$ denotes the covariance at lag m for a block of data. There are different ways to compute $\gamma(m)$, specifically the scaling factor, so let us use the following convention: assume the data is 0 mean; if not subtract the sample mean first. Then (with data indexed $1 \leq n \leq N$):

$$\gamma(m) = \frac{1}{N} \sum_{n=m+1}^{N} x_n x_{n-m}$$

The (auto-)correlation coefficients are:

$$\rho(m) = \gamma(m) / \gamma(0)$$

For example, the MATLAB function *autocorr*(x) will compute and plot $\rho(m)$ for $0 \leq m \leq 20$, with significance bounds (in the graph, if a displayed value is below the bounds, then it can be considered as 0). For our purposes, it is fine if we use $|\rho(m)| > 0.2$ as a test of "significance". Our formula for $\gamma(m)$ is technically biased, and can be interpreted as a Bartlett (i.e., similar to triangular) windowed form of unbiased estimates. In any case, for our purposes, the lags of interest are $\leq M$, where $M \ll N$ so this distinction is not of interest for us, here.
For pairs of signals, we can define:

$$\gamma_{xy}(m) = \frac{1}{N} \sum_{n=m+1}^{N} x_n y_{n-m}$$

for $m \geq 0$, and similarly for $m < 0$. Then the cross-correlation coefficient is:

$$\rho_{xy}(m) = \frac{\gamma_{xy}(m)}{\sqrt{\gamma_{xx}(0)\, \gamma_{yy}(0)}}$$

and in MATLAB, *crosscorr*(x, y) will compute and graph this, by default for $-20 \leq m \leq 20$.

1. $\boxed{\textbf{Heavy Tail Distributions}}$

   First we are going to explore some distributions. Generate $N = 1e6$ iid samples of $N(0,1)$, Cauchy with $\alpha = 1$, and Students' t-distribution with $\nu = 5$ and $\nu = 10$ degrees of freedom. The variance of Students' t-distribution is finite for $\nu \geq 3$ (well, technically $\nu > 2$ but here we consider only integer values for ν) and is given by:

$$\frac{\nu}{\nu - 2}$$

   For comparison, you should normalize your Students' t-distribution data sets so they have variance 1. Obviously don't normalize Cauchy since it has no variance! However, $\alpha = 0.544$ is a reasonable "match" to $N(0,1)$ in the sense that both yield approximately the same $P(|X|<1) = 68\%$ (which is the probability of being within one standard deviation for the case of $N(0,1)$).

In MATLAB, *trnd* generates the Students' t-distribution. You can generate Cauchy $\alpha = 1$ as a special case of Students' t by setting $\nu = 1$; multiply the result by α to create Cauchy with general α. Another approach is that if U is uniform over (0,1), then $X = \alpha \tan(\pi U)$ is Cauchy. Note in general Cauchy α has pdf;

$$f(x) = \frac{\alpha/\pi}{\alpha^2 + x^2}$$

In any case, at this point you have N samples of "comparably scaled" Gaussian, Cauchy and Students' data. For each, compute the fraction of the time the absolute value exceeds 4. Also, graph the Cauchy data: you will see something interesting (you will not see in the others).

2. **AR and ARMA**

Let us start by synthesizing data using an ARMA$(2,2)$ model:

$$r_t = \frac{(1+0.2z^{-1})(1+0.5z^{-1})}{(1-0.8z^{-1})(1-0.7z^{-1})} v_t$$

where v_t is iid $N(0,1)$. Interpret the above as an operator notation for the appropriate difference equation. Generate $N = 250$ samples. Consider the AR(p) model:

$$r_t = v_t + w_{p1} r_{t-1} + \cdots + w_{pp} r_{t-p}$$

The AR coefficients as given in the notes would be $a_{p0} = 1$ and $a_{pk} = -w_{pk}$, $1 \le k \le p$. As usual we take $a_{p0} = 1$. We can interpret this as a linear regression model with model error v_t. In this way, we can compute the AR coefficients $\{a_{pk}\}_{k=1}^{p}$ using a least-squares fit. Recall a_{pp} is the reflection coefficient κ_p, and $E\left(|v_t|^2\right)$ is P_p, the power in the p^{th} order prediction. The set κ_p and P_p can be found using a Levinson-Durbin recursion instead.

Take maximum order $M = 10$.

(a) Estimate the covariances $\gamma(m)$ for $0 \le m \le M$. Obtain a stem plot of the values $\gamma(m)/\gamma(0)$ for $0 \le m \le M$. Those above say about 0.2 are significant. The number above that level gives an indication that an AR model of that length would be useful, higher order models would give marginal improvement. **Note:** In MATLAB, *autocorr* does all this for you.

(b) Set up the corresponding $(M+1)\times(M+1)$ Toeplitz matrix C, and compute its eigenvalues. They are (hopefully) strictly positive real.

(c) Recall the *Cholesky factorization* of a pd matrix is $C = L_{\text{chol}} L_{\text{chol}}^T$ where L_{chol} is lower triangular with positive entries on the diagonal. Closely related to this is the so-called *LDL* decomposition $C = LDL^T$ where L is lower triangular with 1's on the diagonal, and D is a diagonal matrix with positive entries. If $\{\ell_i\}$ are the diagonal elements of L_{chol}, then $D = diag\{\ell_i^2\}$. You should be able to find LDL decomposition in your function library whether you use MATLAB, Python or really any good scientific computing software; if not and all you have is Cholesky, you can derive D from L_{chol} as suggested, and then $L = L_{\text{chol}} D^{-1/2}$. If you can't find either LDL or Cholesky decomposition, look harder. If it's not there, get a better library! In any case, find the L and D for your covariance matrix C.

(d) Use the Levinson-Durbin recursion (you can use a function from your library, if not then hand code it) to compute the reflection coefficients and prediction error powers P_m up to M. Also compute the FPEF coefficients of all orders up to order M. Pack these FPEFs into a triangular matrix of the form:

$$F = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{11} & 1 & 0 & \cdots & 0 \\ a_{22} & a_{21} & 1 & \cdots & 0 \\ \vdots & & & & \\ a_{MM} & \cdots & & a_{M1} & 1 \end{bmatrix}$$

(e) Compute FCF^T. Compare the diagonal elements of D with the P_m's.Compare F with L^{-1}. In summary, the Levinson-Durbin recursion is tied to the LDL decomposition (though it doesn't give you L and D directly, it gives what turns out to be the more useful L^{-1} and D).

(f) Now use a LS fit to compute the AR coefficients of order M. Compare them with the M^{th} order model you found via Levinson-Durbin.

(g) Comment on the reflection coefficients you found. Consider the following general issue: if a reflection coefficient is close to 1 in magnitude, what does that tell you? if it is close to 0, then what? **Hint:** Recall the formula that relates the P_p's to the κ_p's. Also, recall what holds for the κ's if the data is exactly $AR(M_0)$ for some order M_0.

3. Repeat the above experiment with the modification that you have something close to an ARIMA(2,1,2) model: $r = \mathbf{H}v$ where:

$$H(z) = \left(1 - 0.99z^{-1}\right)^{-1} H_0(z)$$

where $H_0(z)$ is the ARMA model in Problem 2. In other words, your r_t will now exhibit behavior that would suggest a unit-root nonstationarity. Plot the original r_t (one simulated block of data), and now the r_t with this (close to) unit root nonstationarity. After you generate the AR models and do the other analysis, now compute $s_t = r_t - r_{t-1}$, the first difference, and repeat all for s_t. The model should work better for s_t. The question is, can you detect the unit root nonstationarity early on? Go back to where you graphed the partial correlation coefficients, i.e., $\gamma(m)/\gamma(0)$ for r_t and comment.

4. Now repeat both **Problems 2 and 3** using a Students' t-distribution with $\nu = 5$ for v_t (normalized so $E(v_t^2) = 1$ still). The parts to repeat, and the issue of concern, is the numerical stability and apparent fit of the model. In theory, this should change NOTHING! The theoretical values of $\gamma(m)$ are IDENTICAL. There are various ways of examining the changes, see if you can detect anything noticeable (or, in particular, disturbing) happening. If not, good for you!

5. $\boxed{\textbf{Cointegration}}$

Here we look for an example of cointegration, among other effects in time series. Let us start with two underlying signals a_t, b_t, with the first having a unit root nonstationarity and the other stationary:

$$\begin{bmatrix} a_t \\ b_t \end{bmatrix} = \delta + G \begin{bmatrix} a_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ v_t \end{bmatrix}$$

where:

$$\delta = \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix} \quad G = \begin{bmatrix} 0.99 & 0 \\ 0.3 & 0.3 \end{bmatrix}$$

where u_t is ARMA$(2,2)$ driven by an input white noise that is iid $N(0,1)$, with innovations filter:

$$H = \frac{\left(1+0.2z^{-1}\right)\left(1+0.3z^{-1}\right)}{\left(1-0.2z^{-1}\right)\left(1-0.5z^{-1}\right)}$$

and v_t are iid samples with $E(v_t) = 0$, $E\left(v_t^2\right) = \sigma_0^2$ (a constant to be prescribed later), and two possible distributions you will test: $N\left(0, \sigma_0^2\right)$, or Students' t-distribution with $\nu = 5$ scaled to achieve the desired $E\left(v_t^2\right)$.

Next, let:

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$$

and then:

$$r_t = \begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = A \begin{bmatrix} a_t \\ b_t \end{bmatrix}$$

(a) Compute the δ' vector and G' matrix such that we can write:

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \delta' + G' \begin{bmatrix} r_{1t-1} \\ r_{2t-1} \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}$$

where η_{1t}, η_{2t} are correlated disturbances, computed via $\eta_t = A \begin{bmatrix} u_t \\ v_t \end{bmatrix}$.

(b) We want the effect of b_t to be noticeable but not dominant, so do the following. After you generate a_t, take:

$$\sigma_0^2 = \frac{1}{16} E\left(a^2\right) = \frac{1}{16}\left(\frac{1}{N}\sum_{n=1}^{N} a_n^2\right)$$

Then use that to generate b_t. Finally compute r_{1t}, r_{2t}. [In reality, you should do this twice, using the two different distributions for v_t. Do the rest of this exercise, as well, for each case].

(c) Graph u_t and v_t,. separately. Superimpose graphs of a_t, b_t. Superimpose graphs of r_{1t}, r_{2t}. Repeat this, say three times, to see "typical" results.

(d) We now want to detect the unit root stationarity in r_{1t}, r_{2t}, and determine if there is indeed cointegration. Indeed, we can recover a_t, b_t via $A^{-1} r_t$, which shows that a linear combination of r_{1t}, r_{2t} is stationary! Determine the coefficients of cointegration (i.e., the linear combination of these two that yield a stationary result).

(e) Compute and graph the auto-correlation coefficients of r_{1t}, r_{2t}, separately, i.e., $\rho_{11}(m)$, $\rho_{22}(m)$ for $0 \leq m \leq 20$, and compute the cross-correlations $\rho_{12}(m)$ for $-20 \leq m \leq 20$. In MATLAB, *autocorr* and *crosscorr* will do this for you.

(f) With $s_{it} = (1 - z^{-1}) r_{it}$, $i = 1, 2$, repeat the above (i.e., look at the two auto-correlations and cross-correlation coefficients between s_{1t}, s_{2t}). You should see a steeper decay which indicates both r_{1t}, r_{2t} suffer from unit root nonstationarity.

(g) You will note the cross-correlation between r_{1t}, r_{2t} decays slowly. To check this is indicative of co-integration, try creating another signal that is INDEPENDENT of r_{1t} but also has a unit root nonstationarity. For example, let ξ_t be iid $N(0, 1)$, and $c_t = 0.99 c_{t-1} + \xi_t$. Checking the $\rho_c(m)$ coefficients of c should of course confirm it has the unit root nonstationarity. But now also look at the cross-correlation coefficients between c_t and r_{1t}, and between c_t and r_{2t}. You should see significantly less cross-correlation between them.

6. **ARCH/GARCH**

Now we try ARCH/GARCH modeling. Take two years each of daily data for S&P500 and say two stocks that interest you, let's say some time within the last 20 years. (You can get data off Yahoo Finance). Specifically, we are looking at *adjusted closing prices,* and as a first step compute the daily **log-returns** for the period. Each year will produce about 250 data points, but realize the exact number may vary.

First graph the returns and square returns. The returns should appear to be close to constant mean (if not zero mean). Since we do not want to deal with the means here, compute and subtract off the sample means, so from this point we will assume the returns r_t have 0 mean (constant). You may also try multiyear models, say 2007-2009, which captures the 2008 financial crisis, or 2018-2020 (to date) which captures.... hmm, that which will not be mentioned!

Compute and graph the auto-correlation coefficients up to lag 20 for the returns AND the square returns. If there are not significant ρ values for the returns, that suggests our mean model can be simple, i.e., constant mean. If the ρ values for the square returns however show some significance, then it is indicative a conditional heteroskedasticity is in play. Ideally, for this simple experiment, your data should exhibit this (conditional heteroskedasticity and constant mean), and pick different data sets if not. If we assume the conditional expectation is 0 (let's say subtract off the sample mean first, if necessary), our model is simply:

$$\begin{aligned} r_t &= \epsilon_t \\ \epsilon_t &= \sigma_t z_t \\ \sigma_t^2 &= \text{GARCH}(p, q)\left[\epsilon, \sigma^2\right] \end{aligned}$$

where z_t are iid with $E(z_t) = 0$, $E(z_t^2) = 1$.

(a) First some simulation. For a GARCH$(1,1)$ model we can write:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Synthesize r_t according to such a model, with $z_t \sim N(0,1)$, with $\omega = 0.5$, $\alpha = 0.6$, $\beta = 0.4$. Superimpose graphs of r_t and σ_t. You should see variability in the volatility of r_t, that σ_t seems to "track". Now, repeat this with z_t have a Students t-distribution with $\nu = 5$ degrees of freedom.

For each case, fit a GARCH$(1,1)$ model, and an ARCH(2) model, assuming a Gaussian distribution for z_t (this is called QML- quasi-maximum likelihood, since you are assuming a Gaussian likelihood function even when it is not the true one). For the GARCH$(1,1)$ model, do you get back something close to your coefficients? For the ARCH(2) model, superimpose plots of r_t and σ_t as generated by the model, and see if you have a good match for the volatility. What impact, if any, does the Students t-distribution with $\nu = 5$ degrees of freedom have?

(b) Your GARCH fitting library should provide some information regarding the significance of the model coefficients. Take your actual financial data, and fit r_t with a GARCH(1,1) and ARCH(2) and see if either provides good fits, and also plot r_t and σ_t superimposed. If the volatility model is a good fit, the graph of σ_t should appear to be like an “envelope” for the return data.