Stochastic Analysis

Barış Bostancı

December 2025

1 Introduction

In this paper, we will analyze stochastic properties and behaviors of several stocks from BIST100 index. Our dataset is comprised of 6 stocks, 2 tickers from 3 sectors -banking, food and energy. That would make the total of 6 random time series that we could analyze and create and combine new random variables. We are mainly interested in correlation coefficients within this dataset. Let us briefly present the essential theoretical framework of this study.

1.1 Theory and Definitions

The very nature of our study requires us to work with time series: Stocks, bonds, options or any other financial instrument that has a changing price P_t constitutes a time series. Processes that are random and changing in time are called stochastic processes. The formal definition requires the use of filtration and several underlying assumptions on the process, but we will omit those formalities for the brevity of this report, every process in consideration will assumed to have nice properties which makes them susceptible to analysis.

We will mainly resort to techniques used in time series analysis. Main premise of the time series analysis is to assume a particular form of the stochastic process with yet unknown parameters, and then using the time series data (the realization of the process) to fit the parameters. Since the problem is stochastic, we will be using numerous estimators to fit those parameters. It can be proven that under certain assumptions of the stochastic process, the model is the best estimator in some sense, usually the sense in which the model optimizes certain objective function such as likelihood.

Description so far was informal and vague. Let us formalize some of these notions and introduce some of the models used in time series analysis.

Definition 1 (Discrete Time Stochastic Process)**.** *Discrete time stochastic process is a sequence of random variables* $\{X_t\}_{t\in\mathbb{N}}$*, where* $X_t : \Omega \to \mathbb{R}$*.Given some* $\omega_0 \in \Omega$ *a realization of* X *is the sequence* $X(\omega_0) := \{X_t(\omega_0)\}_{t\in\mathbb{N}}$

Now that we have the definition of what a stochastic process is, how can we estimate its properties. Estimation is a complicated business, not because the mathematical subtleties of the estimator, but rather, given the estimator, how well is it suited for the given stochastic process. Usually the way process of time series analysis unfolds as follows:

1. A realization of some stochastic process $\{X_t\}_{t\in\mathbb{N}}$ is given
2. A particular form is assumed for $\{X_t\}_{t\in\mathbb{N}}$, there are hundreds of models to choose from(this is where subjective beliefs and pattern recognition comes handy; for instance using the stylized facts about volatility of an asset, one can construct heavy tailed models and models that allow volatility clustering) , some prominent models are: AR(n), ARCH(n), GARCH(p,q), VECM, and so on. It is important to make the distinction, however, that some of these models assume the mean dynamics whereas others assume a volatility dynamics.
3. Optimizing the log likelihood of the function. (under which parameters my model predicts best given the data)
4. Feeding the time series data to optimization algorithm and calculating the parameters of the model. (Or using a closed form solution if there is one)

There are some nice theorems about these models, for instance, under the Gauss-Markov assumption, the OLS(ordinary least squares) of the model is BLUE(best linear unbiased estimator).

If one models the prices using the assumption that underlying model is continuous time stochastic process, there is whole another literature on that, which heavily relies on the mathematical framework of stochastic calculus, where the integration is no longer the ordinary Riemann integral, but rather is called Itô integral, and it behaves different accounting for the quadratic variation inherent in Wiener processes. One of the best known model of financial markets that employs the framework of stochastic calculus is Black-Scholes model:

$$dS_t = \mu S_t \ dt + \sigma S_t \ dW_t \tag{1}$$

using this equation one can also estimate the forward-looking volatility of prices by essentially fitting the parameters of BS via optimizing an objective function.

1.2 Discrete-Time Models

We will briefly introduce the models that we will be using in our analysis of BIST100 stocks.

Definition 2 (Autoregressive Models)**.** *An autoregressive model of nth order, called AR(n) of a time series data has the following form*

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots \beta_n y_{t-n} + \epsilon_t \tag{2}$$

where y_t is response variable, and ϵ_t is a noise. This structure allows lagged responses.

An intuitive explanation of this model is that, there are processes with memory, that is, if there is some shock at some time that shock will have an effect on future responses. For instance, if we were modeling a volatility of a log returns of a price, we would need some sort of memory because of the volatility clustering effect. This is just a basic example of how this process can be employed. There are certain technical regulation assumptions so that this model is reasonable, but we will omit those considerations for the sake of brevity.

Definition 3 (ARCH - Autoregressive Conditional Heteroskedasticity)**.** *Autoregressive conditional heteroskedastic models assume the following form for the stochastic process*

$$X_t = \beta_0 + \beta_1 X_{t-1} + \epsilon_t \tag{3}$$

where $\epsilon_t = \sigma_t z_t$, where z_t is white noise and $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$

This model remembers the shocks and models variance of the process, this model mainly used to fit volatility of square log-prices.

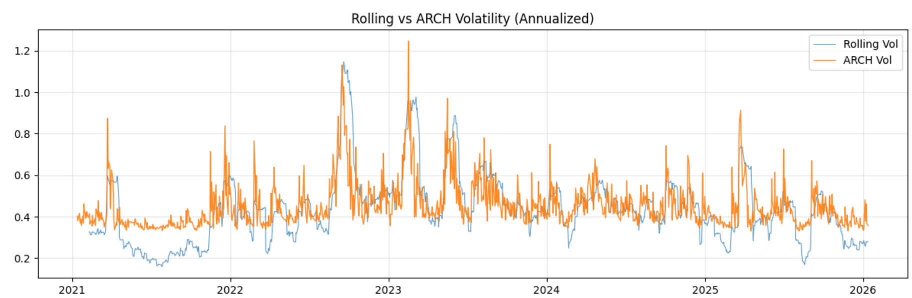

Figure 1: Rolling Volatitiliy vs ARCH(10) Volatility of AKBNK stock prices

1.3 Cointegration & Pairs Trading

So far we were only looking for single stocks, however by doing that we are essentially disregarding very valuable information. One might expect intuitively that two stocks should move together, for instance if A and B are two firms that produces a microchips and both are similar in their qualities, they should (one might heuristically assume) behave similarly, and we can take advantage of that behavior.

Mathematically, we want to define a distance between two time series, or in the continuous case, we want to find a valid metric that measures the similarity of two stocks. It is quite straightforward that one can search space of all stocks and find two that minimizes that metric, and after finding such stocks, one can create a synthetic asset out of those, that is, some special linear combination of two cointegrated stocks and observe that this asset is mean reverting, after that it is just the matter of employing statistical arbitrage to gather the difference.

The question is that, how can we go about finding those stocks, and how can we test whether those are cointegrated given that we find such a pair. For a simple demonstration, simple python script searches the 50 stocks in BIST100, and test each couple using Dickey-Fuller test, and out of those 1225 stocks, the best matching pair was AKBNK and TSKB, below is their spread -It is important to observe that the residual term has a mean reversion property. Let X_t

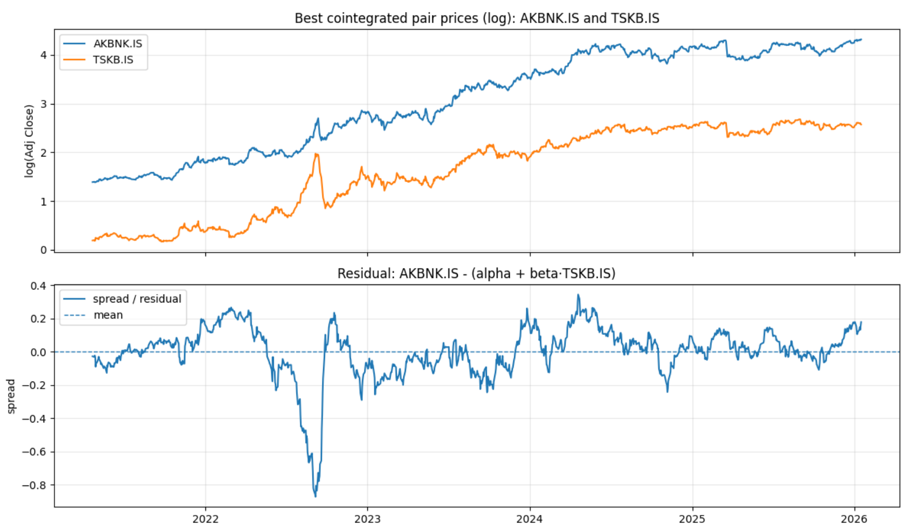

Figure 2: Cointegrated Pairs.

be the process governing the price of AKBNK and Y_t be the process governing the price of TSKB. The idea is that, there might exists a parameter γ for which $Y_t - \gamma X_t \sim \epsilon_t$ where ϵ_t is a stationary distribution. What that means intuitively is that, these stocks follow each other in a peculiar way. In practice, how can we find γ ? The procedure is as follows:

$$Y_t = \hat{\alpha} + \hat{\gamma} X_t + \hat{u}_t \tag{4}$$

we perform least squares to figure out the parameters, and then we obtain an explicit expression for the residual term $\hat{u}$. Afterwards we can test whether the distribution of residual term is stationary using Dickey-Fuller test.

Perform

$$\arg\min\left(\sum_{t\in I}(Y_t-(\hat{\alpha}+\hat{\gamma}X_t))^2\right) \tag{5}$$

and set

$$\hat{u}=Y_t-(\hat{\alpha}+\hat{\gamma}X_t) \tag{6}$$

and apply the adapted Dickey-Fuller test. For the case of AKBNK and TSKB, we performed the OLS to approximate γ, and then use aDF test to check if the residual distribution was stationary. Generalized versions of this model goes by the name **VECM** -vector error correction models- and it performs the cointegration for arbitrarily large number of time series.

1.4 Regime Detection using Hidden Markov Models

Markov chains are extremely ubiquitous objects in applied mathematics, they are extremely nice objects to work with, and easy to implement in numerous algorithms.

Definition 4 (Hidden Markov Models)**.** *A hidden Markov models are augmented versions of Markov chains: It consists of hidden states* $\{H_1,\dots,H_n\}$ *with initial probability distribution* $\pi\in\mathbb{R}^n$*, and observables* $\{O_1,\dots,O_m\}$*, at each time step system moves from one hidden states to another, and the probability of jumping from state* i *to* j *is given by a Markov matrix* $\mathbf{A}\in\mathbb{R}^{n\times n}$*, these probabilities are called* ***transition*** *probabilities, each hidden state corresponds to an observation* O_k*, the probabilities of observations given the hidden states are called* ***emission*** *probabilities, it is denoted by a matrix* $\mathbf{B}\in\mathbb{R}^{n\times m}$*.*

The goal of the hidden Markov model is to provide two answers to two important questions:

1. Given the observations, which set of parameters $\Theta=(\pi,\mathbf{A},\mathbf{B})$ are most likely - $argmax\ P(O_{1:T}|\Theta)$
2. Given the sequence of observations $O_1,\dots,O_T$ which sequence of $H_1,\dots,H_T$ is most likely - $argmax\ P(H_{1:T}|O_{1:T})$

These questions are answered by Viterbi algortihm and by Baum-Welch algorithm, via the use of Expectation-Maximization method -it is an iterative method that converges to some local minima.

We can apply the HMM to markets, hidden states corresponding to whether market is on the rise or on the fall (bear or bull), figure below shows an application of this model to $TUPRS$ stock: Blue stripes represents the hidden state

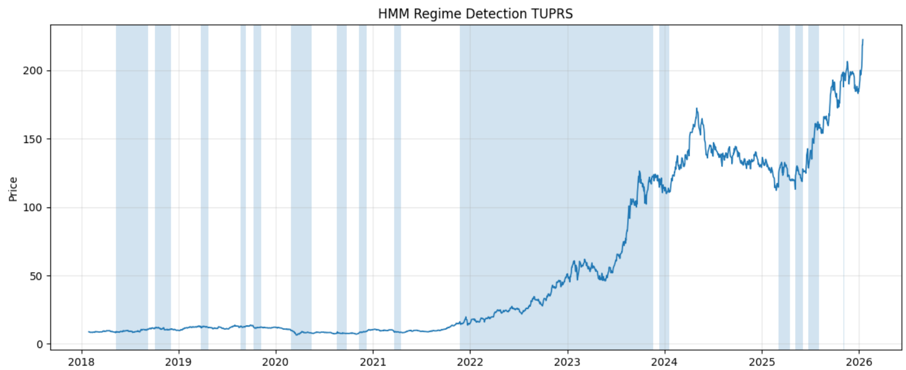

Figure 3: HMM for regime detection

1, and white stripes representing the hidden state 0. Blue stripes corresponds to bullish regimes, and white stripes representing bearish regimes.

2 BIST100 Analysis

We are asked to find the correlation coefficient between daily stock prices and the daily USD/TRY exchange rate. Computing a correlation coefficient amounts to computing a covariance between two random variables. The theoretical definition of covariance requires knowledge of the joint probability law of these variables, which we do not have access to. Therefore, we must resort to an estimator. However, there is an additional layer of difficulty: we are no longer

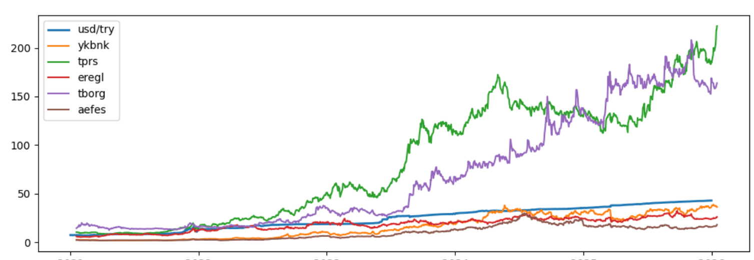

Figure 4: Prices of Several Stocks from BIST100

working with cross-sectional data but with time series. At each time index, we observe a potentially different random variable, whose distribution may change over time. In that case, using the standard sample covariance estimator becomes conceptually unjustified unless we impose additional structural assumptions on the stochastic processes involved.

To make the problem well-defined, one would need to assume that the underlying processes are stationary and ergodic. But for price data, this assumption is clearly violated: real-world price series are not weakly stationary.
Let us try to extract one number for each stock, describing the correlation between the stock and USD/TRY exchange. Given two random variables X, Y, perform

$$\hat{\sigma}_{XY}^2 = \frac{1}{n}\sum_{t=1}^{n}(X_t - \hat{\mu}_X)(Y_t - \hat{\mu}_Y) \tag{7}$$

where $\hat{\mu}_X$ is given by sample mean estimator. The correlation coefficient is given by

$$\hat{\rho}(X,Y) = \frac{\hat{\sigma}_{XY}^2}{\hat{\sigma}_X \hat{\sigma}_Y} \tag{8}$$

X denotes the stocks, and Y denotes the USD/TRY exchange rate. There is a

X	$\hat{\rho}_{XY}$
YKBNK	0.957
AKBNK	0.948
TBORG	0.929
AEFES	0.877
TUPRS	0.963
EREGL	0.930

strong correlation between the USD/TRY exchange rate and the sampled stock prices.

Figure below shows daily price changes in dollars, these time series looks more "nicer" in the sense that it looks much more "stationary" in comparison to price series. In average, investor makes no money, and that means we can model daily price changes as a zero mean stochastic process. The drifts seen in the time series can be described by volatility.

Using the formula (8) we can calculate the correlations between daily price differences, below is a heatmap of correlation matrix between stocks of interest Using these correlations we can estimate price of one asset with another, using the equation

$$\mathbb{E}[\frac{\Delta X}{\sigma_x}] = \rho(X,Y)\frac{\Delta Y}{\sigma_y} \tag{9}$$

Figure(7) depicts this estimation

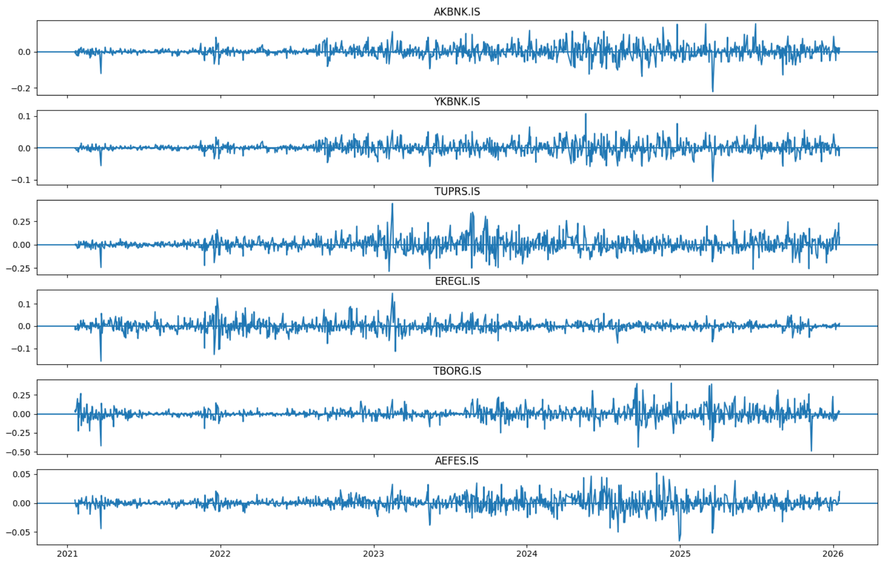

Figure 5: Daily Price Changes in $

2.1 Normal Distribution is a Joke in Finance

In algorithmic trading, time frames change the very dynamics of the system, —so much so that what quantum mechanics is to classical physics, high frequency trading is to algorithmic trading. If one took a glance at the histograms of the monthly log returns of the S&P 500, and then to daily log returns, one would immediately realize that the market disagrees with the normal distribution as the time frame keep getting smaller. One would encounter heavy-tailed distributions, this behavior of the market forces the creation of more realistic models that are using heavy-tailed distributions to model the residuals.

Central limit theorem asserts that contribution of IID random variables converges to a normal distribution, however in finance these random variables are highly unlikely to be IID, and therefore this approach usually fails.

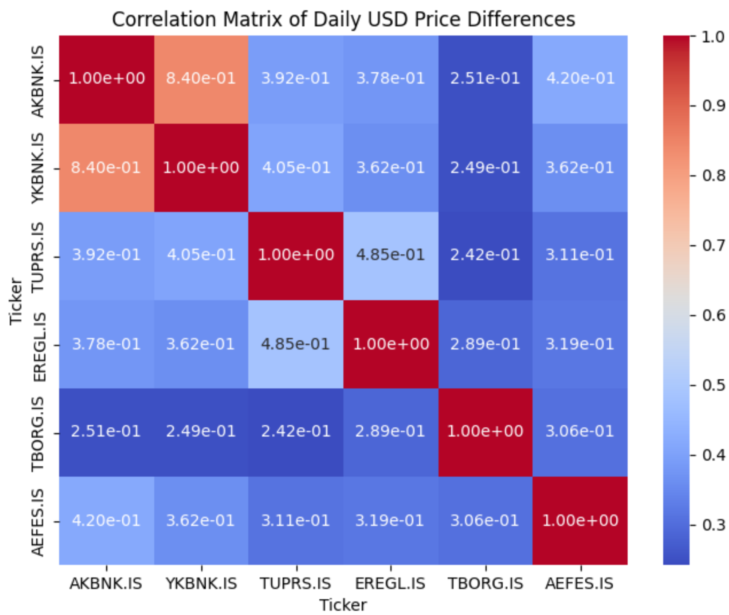

Figure 6: Heatmap of the Correlation Matrix

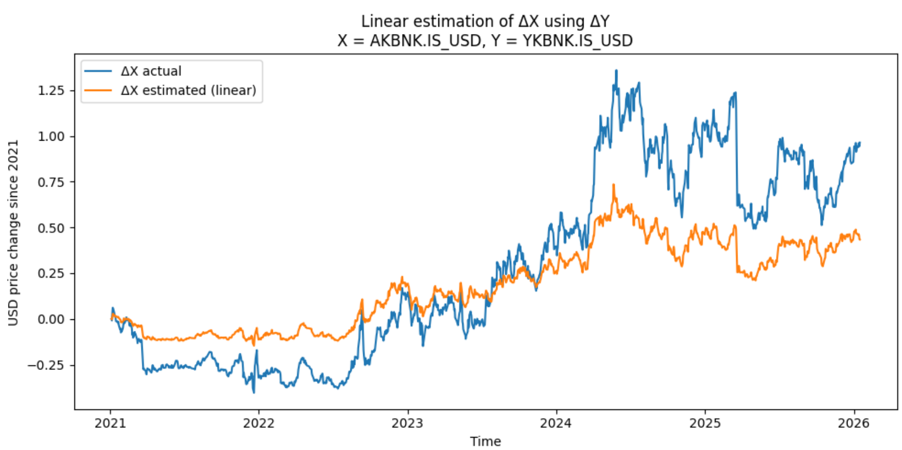

Figure 7: Estimation