Outline for DTW approach

Riccardo Polo

November 2024

Contents

1 Introduction to Dynamic Time Warping (DTW) 2

2 Key Advantages of DTW in Financial Analysis 2

3 Project Application 2
- 3.1 Project Objectives 2
- 3.2 Concrete Implementation 2
  - 3.2.1 Selection of economic indicators 2
  - 3.2.2 Time Frame 4

4 Methodological Overview for the Project 4
- 4.1 Data Collection 4
- 4.2 Preprocessing 4
- 4.3 Construction of Multivariate Time Series 4
- 4.4 Multivariate DTW Calculation 5
- 4.5 Clustering Approach 5
- 4.6 Visualization 5

1 Introduction to Dynamic Time Warping (DTW)

Dynamic Time Warping (DTW) is a powerful algorithm used to measure the similarity between two time series that may vary in speed or timing. In financial markets, companies' stock prices often react to economic events at different times and rates, making DTW an ideal tool for aligning and comparing such time series.

- **Objective**: Utilize DTW to align and compare the time series of stock prices for individual S&P 500 companies, considering time shifts due to different market reactions. Additionally, integrate economic indicators as temporal factors to enrich the analysis.

2 Key Advantages of DTW in Financial Analysis

- **Temporal Flexibility**: DTW handles timing variations, capturing leading and lagging relationships between financial time series.
- **Pattern Recognition**: Identifies similar patterns even when the data is out of phase, facilitating deeper market behavior understanding.
- **Robust Insights**: Provides a dynamic alternative to static alignment methods, better reflecting the complexities of financial markets.

3 Project Application

The project will apply DTW to analyze the time series of stock prices of S&P 500 companies, integrating economic indicators as temporal factors. This approach enhances the understanding of market movements by accounting for asynchronous behavior and volatility.

3.1 Project Objectives

- **Identify Dynamic Relationships**: Use DTW to map leading and lagging connections between company performance and economic indicators.
- **Clustering Analysis**: Apply DTW-based clustering to discern distinct performance and economic response patterns within the dataset.

3.2 Concrete Implementation

3.2.1 Selection of economic indicators

These are some of the possible economic indicators economic indicators:

1. **Interest Rates and Federal Funds Rate and 10-Year Treasury Yield**
   - Relevance to the S&P 500:
     - **Cost of Borrowing:** Interest rates determine the cost at which businesses can borrow money. Lower rates reduce borrowing costs, encouraging companies to invest and expand, potentially boosting their stock prices.
     - **Investment Decisions:** Interest rates influence investor behavior. When rates are low, stocks often become more attractive compared to bonds, leading to increased demand for equities.
     - **Economic Signals:**The Federal Funds Rate is a primary tool used by the Federal Reserve to manage economic growth and inflation. Changes in this rate signal the Fed's stance on monetary policy, affecting market expectations.
   - Importance in DTW Analysis:
     - **Temporal Relationships:** Interest rate changes can lead or lag market reactions. DTW can help identify these temporal alignments, revealing how quickly or slowly the S&P 500 responds to interest rate fluctuations.

- **Policy Impact Assessment:** By analyzing the alignment between interest rates and the S&P 500, the project can assess the effectiveness of monetary policy decisions on market performance.

2. **Inflation Measures: Consumer Price Index (CPI) and Producer Price Index (PPI)**
   - Relevance to the S&P 500:
     - **Purchasing Power:** Inflation erodes consumer purchasing power, potentially reducing consumer spending, which can negatively impact corporate revenues and profits.
     - **Cost of Goods Sold:** Inflation affects production costs for companies. Rising input prices can squeeze profit margins if companies cannot pass costs onto consumers.
     - **Monetary Policy Influence:** Persistent inflation may prompt the Federal Reserve to raise interest rates, indirectly affecting stock market valuations.
   - Importance in DTW Analysis:
     - **Lag Analysis:** DTW can uncover whether inflation metrics lead or lag stock market movements, providing insights into how inflation expectations are priced into the market.
     - **Sector-Specific Impacts:** Inflation may affect different sectors uniquely. Clustering can help identify which sectors of the S&P 500 are more sensitive to inflationary changes.

3. **Economic Output: Gross Domestic Product (GDP) Growth Rates**
   - Relevance to the S&P500
     - **Economic Health Indicator:** GDP growth reflects the overall economic activity and health of the economy. Strong GDP growth typically signals robust corporate earnings potential.
   - Importance in DTW Analysis
     - **Alignment Patterns:** DTW can reveal how GDP growth rates align temporally with stock market performance, highlighting periods where the market may anticipate or lag economic growth.
     - **Cyclical Analysis:** Understanding the synchronization between GDP cycles and market cycles can aid in predicting future market movements.

4. **Labor Market Data: Unemployment Rates and Non-Farm Payrolls**
   - Relevance to the S&P500
     - **Consumer Spending Power:** Employment levels directly affect consumer income and spending, which drive a significant portion of economic activity and corporate revenues.
     - **Business Investment Decisions:** Labor market tightness can influence wage levels and hiring practices, affecting corporate profitability.

     Importance in DTW Analysis:
     - **Leading/Lagging Indicators:** DTW can help determine if changes in the labor market precede or follow shifts in the stock market.
     - **Recovery Patterns:** Analyzing labor data alignment with the S&P 500 can provide insights into recovery patterns post-economic downturns.

5. **Consumer Sentiment: Consumer Confidence Index**
   - Relevance to the S&P500:
     - **Spending Behavior:** High consumer confidence often translates to increased consumer spending, boosting sales for companies and potentially raising stock prices.
     - **Economic Outlook:** Consumer sentiment reflects public perception of the economy's health, influencing investment decisions and market trends.
   - Importance in DTW Analysis:
     - **Temporal Dynamics:** DTW can identify whether changes in consumer confidence lead market movements or if the market influences consumer sentiment.
     - **Behavioral Insights:** Understanding the alignment can provide valuable insights into behavioral finance aspects impacting the market.

The selected indicators collectively represent critical dimensions of the economy—monetary policy (interest rates), price stability (inflation), economic growth (GDP), labor market conditions (employment data), and consumer behavior (sentiment).

These indicators directly or indirectly affect corporate profitability, which is a fundamental driver of stock prices in the S&P 500.

DTW can uncover complex temporal relationships, such as lag effects between economic changes and market responses. The variety of indicators allows for clustering of time series data based on different economic dimensions, helping to identify distinct patterns of market behavior in response to various economic conditions.

3.2.2 Time Frame

Analyze data over the past 20 years to capture multiple economic cycles, including recessions and expansions.

4 Methodological Overview for the Project

To adapt DTW for the project, the following steps will be undertaken:

4.1 Data Collection

- **Historical Data Collection:** Download daily closing prices of S&P 500 company stocks. Obtain monthly or quarterly data for selected economic indicators.
- **Data Alignment:** Convert all data to a common frequency (e.g., monthly). Align dates to ensure matching between datasets.

4.2 Preprocessing

- **Data Cleaning:** Handle missing values and adjust for stock splits and dividends in stock data.
- **Normalization:**Normalization time series. this allows for comparison between variables on different scales. Use log returns to stabilize variance in financial data. This will stabilizes variance and normalizes distributions in financial data. This mitigates the impact of outliers and extreme values.

4.3 Construction of Multivariate Time Series

- **Data Combination:** Use adjusted closing prices or log returns. Select relevant indicators that may impact the company's performance. Include company fundamentals like earnings per share, revenue, or sector classification.

  Data Structuring: Organize data into a multivariate time series where each time point includes all selected variables. Ensure that the data for each variable is in the same format (e.g., all in log returns). **Example Structure:**
  For company i, the multivariate time series $X^{(i)}$ is:

$$X^{(i)} = \begin{bmatrix} x_{1,1}^{(i)} & x_{1,2}^{(i)} & \cdots & x_{1,d}^{(i)} \\ x_{2,1}^{(i)} & x_{2,2}^{(i)} & \cdots & x_{2,d}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T,1}^{(i)} & x_{T,2}^{(i)} & \cdots & x_{T,d}^{(i)} \end{bmatrix}$$

  - T: Number of time points. - d: Number of variables.

- **Synchronization:** Ensure all variables are temporally aligned. In particular, align all variables to the same time stamps and handle discrepancies due to different data frequencies

4.4 Multivariate DTW Calculation

- **DTW Distance Calculation:** Use Python libraries such as `tslearn` to calculate multivariate DTW. Compute DTW distances between all pairs of multivariate time series of companies.
- **Optimal Warping Path:** Extract the warping path to understand temporal alignment. Analyze leading and lagging relationships.
- **Constraint Application:** Implement a constraint window (e.g., Sakoe-Chiba band) to limit warping Thus limits the maximum temporal shift allowed between the two series and reduces computational complexity and prevents excessive warping that may lead to unrealistic alignments.

4.5 Clustering Approach

- **Choice of Clustering Algorithm:** Use agglomerative hierarchical clustering with DTW distance. Suitable for cases where the number of clusters is not known a priori.
- Experiment with different linkage methods. Examples: Distance between clusters can be the average of all pairwise distances or the maximum pairwise distance or the minimum pairwise distance.
  Distance between clusters is Pass the precomputed DTW distance matrix to the algorithm.
- **Determination of Number of Clusters:** Apply the Elbow Method to find the optimal number of clusters.

4.6 Visualization

- **Time Series Plots:** Plot the raw multivariate time series for selected companies, this helps in observing the inherent differences before alignment.
  Plot the time series after DTW alignment. This could show how DTW adjusts time axes to align similar patterns.
  -Use lines or markers to connect corresponding points.
- **Cluster Visualization:** Reduce high-dimensional data (DTW distances) to 2D or 3D for visualization. We can use the t-SNE Algorithn since it can emphasize local structure, making it suitable for visualizing clusters.
  - Plot the 2D embeddings, coloring points according to cluster assignments.
- **Heatmaps and Distance Matrices:** Create heatmaps of DTW distance matrices. Display hierarchical clustering dendrograms. An implementation can be:
  Input the linkage matrix obtained from hierarchical clustering. Customize the dendrogram: