A guide to time series analysis

Arthur Wang, Kolin Konjura

Recent developments in big data and artificial intelligence (AI) have increased the benefits of time series analysis. More data, increasingly efficient computation, and greater processing power enhance the ability of proven time series methods to solve business problems through forecasting and process modeling. In addition, these same improvements have led to new ways to apply time series analysis.

What are time series?

A time series is data measured in chronological order. The intervals can be regular or irregular, and can vary in length – for example, time series can be monthly, weekly, daily, or even by the second. Much of the data we collect are time series, such as stock prices, machine telemetry, and temperature readings over a period of time. Granular data can also be converted to a time series; for example, individual clicks on a website can be aggregated to measure daily website activity. Time series analysis includes several techniques and models designed for this kind of data. **The growth of big data has created increasing opportunities to leverage time series analyses**.

What is time series analysis?

Time series analysis is a set of methods that allow us to understand time series data. This can be as simple as understanding when seasonal trends occur, but usually involves deeper analysis to build a more complete model of the underlying processes. Building and implementing a time series model describes underlying correlations within the data. Understanding these correlations generates insight into the data which can be adequately utilized to forecast and predict future values.Why use time series analysis?

Time series analysis is especially useful for forecasting. In forecasting, time series methods often perform better than simpler methods, such as linear regression or simple extrapolation of historical data. **Time series methods capture underlying structures that are common to time series data, such as trend and seasonality[1]**. These aspects of time series data can violate the assumptions of simpler methods, making the simpler methods potentially less powerful. By accounting for common real-world conditions, time series analysis produces forecasts that are generally more accurate.

"By accounting for common real-world conditions, time series analysis produces forecasts that are generally more accurate."

Use cases

Let's discuss a few business cases where time series could be integral to the data presented.

Suppose a company needs to understand household energy levels to identify neighborhoods that don't meet green energy requirements They find that an individual's daily usage is higher during the day and lower during the night, which produces a cyclical pattern within the data. Now suppose the city is part of Minnesota. The climate is relatively mild in the summers, but is unbearably cold in late fall, winter, and early spring. The data may show that household energy levels spike during those cold months, and level out during the summer months. The data then shows cyclical trends throughout the day and seasonal trends throughout the year. Time series analysis would work well for modeling and predicting this situation.

Now suppose a distributor needs to understand local traffic patterns for their supply lines. The company then needs to determine the most efficient times to move their product. The supply distribution company does not want to move product during daily rush hour. If they are moving product on Fridays, rush hour may start at 2pm or 3pm instead of 4pm or 5pm. Or, if there is a sport event occurring near their facilities, they will want to avoid that traffic as well. A time series model can help the distribution company find the optimal time to move their product with efficiency.

Advantages

The nature of time series data can lead to complex and arduous calculations that can make forecasting difficult. Luckily, with modern computational capabilities, implementing a time series forecast can be done with relative ease. Languages such as Python and R have ready-built packages that allow analysts to create and tune accurate time series forecasts with relatively little effort[4].

In addition to providing accurate predictive capabilities, time series models have minimal assumptions, are statistically rigorous, and most importantly, are stable. This means that if a large, unexpected event occurs, the model can react and provide insight to solutions around that event. For example, the COVID-19 pandemic has disrupted operations in several industries. Time series methods are able to adjust forecasts and analysis with the new climate in mind.

Drawbacks

As with most advanced techniques, a time series analysis can be only as good as the data itself. If the historical data is sparse, then the forecasting will not be fruitful and a machine learning approach technique may be better in that instance.

The more complex the data is, the more difficult it can be to implement a time series forecast model. In these instances, it may be that a time series model is only part of the solution[4]. For instance, with the supply distribution example previously explored, time series analysis may give recommendations for optimal driving time, but those times could be dependent on when the supply is ready to be shipped, unforeseen accidents on the road, etc.

The landscape of time series analysis

Today, time series analysis is a mix of proven and emerging techniques. Advances in computing have enabled rapid exploration of different models in order to select the one that performs best[3], rather than painstakingly selecting one up front. Greater processing power and developments in GPUs have also enabled the application of deep learning to time series forecasting, an approach that has recently outstripped traditional methods[2]. Businesses today have a greater range of choices than ever before in how they generate powerful forecasts.

References:

1. https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm

2. https://towardsdatascience.com/introducing-pytorch-forecasting-64de99b9ef46

3. https://robjhyndman.com/hyndsight/tscv/

4. Time Series Analysis with Applications in R by Jonathan D. Cryer and Kung-Sik Chan