Lecture 15 – Time Series Models

Hamed Hasheminia

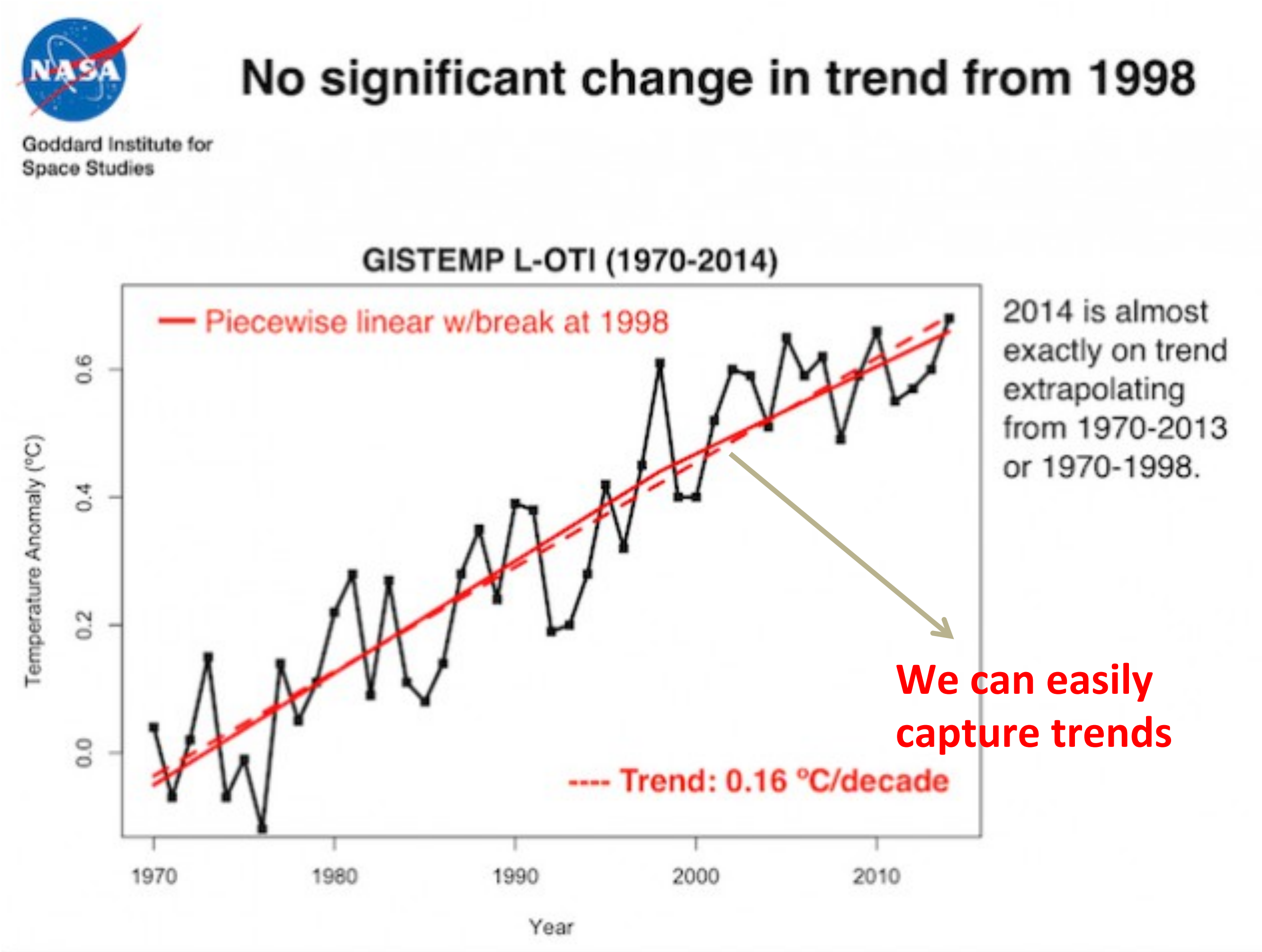
Agenda

- Definition of Time series
- Trend Line and Seasonality
- AR models
- MA Models
- ARMA models
- ARIMA models

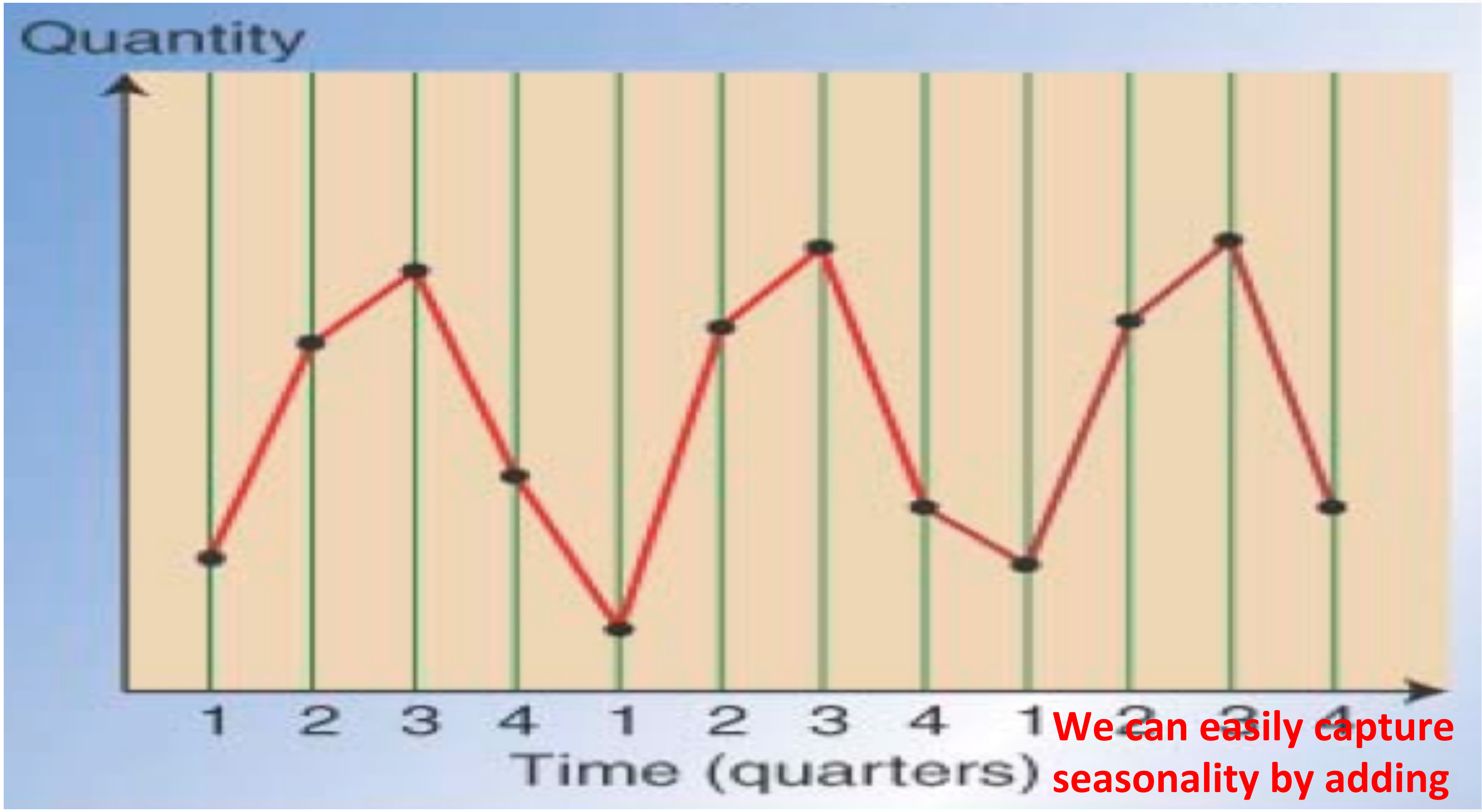
What are time series models?

- Time series models are models that will be used to predict a future value in the time-series.
  - Like other predictive models, we will use prior history to predict the future!
  - Unlike previous models, we will use the *outcome* variables from earlier in time as the *inputs* for prediction.
  - We will want to evaluate on *held-out set or test data* to ensure our model performs well on unseen data.
  - Unlike previous modeling exercises, we won't be able to use standard cross-validation for evaluation!
  - Since there is a time component to our data, we **cannot choose training and test examples at random.**

What can we do with our linear regression lines?



What can we do with our linear regression lines?

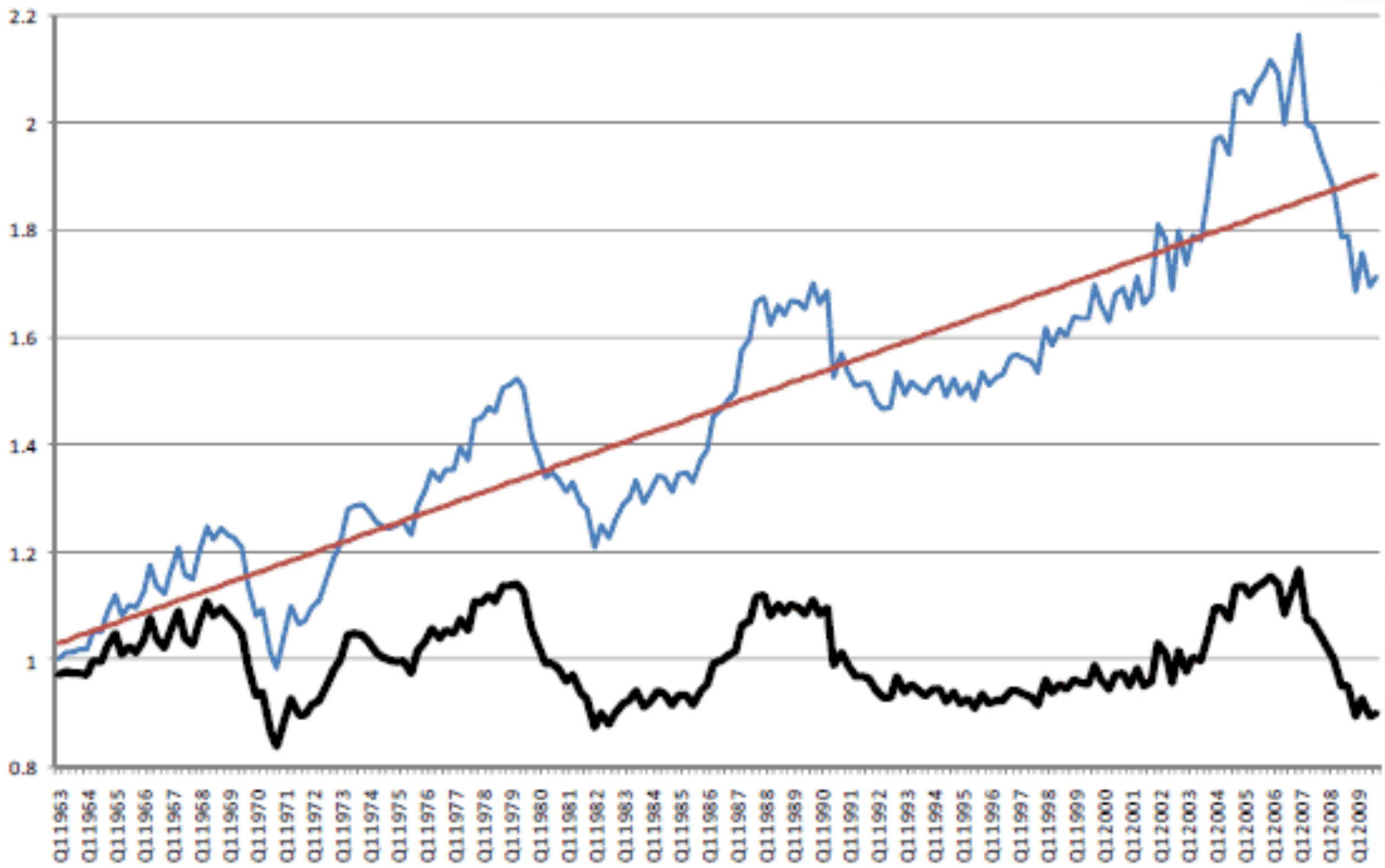


We can easily capture seasonality by adding dummy variables for seasons.

So why do we use time series models?

- Not everything is trend and seasonality!
- Sometimes very recent events affect tomorrow outcomes significantly.
- With time series models we are interested in capturing de-trended patterns.

An example for de-trended data. (Housing Data)



Auto-Regressive Models (AR)

- Autoregressive (AR) models are those are that use data from previous time-points to predict the next time-point. These are very similar to previous regression models, except as input - we'll take some previous outcome.
- If we are attempting to predict weekly sales, we'll use sales from a previous week as our input. Typically, AR models are noted AR(p), where p indicates the number of previous time points to incorporate, with AR(1) being the most common.
- In an autoregressive model, similar to standard regression, we'll learn regression coefficients, where the inputs or features are the previous p values. Therefore, we will learn p coefficients or \beta values.
- For AR models, your de-trended data must be Stationary.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

How do we decide how many lags to use in AR models?

- In time series, the ***partial autocorrelation function*** (PACF) is used to choose number of lags in Auto-Regressive models. For more info please refer to: https://en.wikipedia.org/wiki/Partial_autocorrelation_function
- You stop using lags that have PACFs of zero.
- *One of the most common mistakes is using Autocorrelation function instead of partial autocorrelation functions to choose number of lags in AR models.*

Moving Average Models

- **Moving average models**, as opposed to autoregressive models, do not take the previous outputs (or values) as inputs, but instead take the previous error terms (or shocks in the system). We will attempt to predict the next value based on the overall average and previous shocks.
- Using these as inputs helps model sudden changes by directly incorporating the prior shocks/errors. This is useful for modeling a sudden occurrence - like something going out of stock affecting sales or a sudden rise in popularity.
- As in autoregressive models, we have an order term, *q*, and we refer to our model as *MA(q)*. This moving average model is dependent on the last *q* errors.

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

How do we decide how many lags to use in MA models?

- For moving average we use Autocorrelation function to choose q.
- For more info on how Autocorrelation functions are being computed please refer to:

https://en.wikipedia.org/wiki/Autocorrelation

ARMA Models

- *ARMA*, pronounced 'R-mah', models combine the autoregressive models and moving averages. For an ARMA model, we specify two model settings p and q, which correspond to combining an AR(p) model with an MA(q) model.
- Incorporating both models allows us to mix two types of effects.
- Autoregressive models slowly incorporate changes in preferences, tastes, and patterns. Moving average models base their prediction not on the prior value but the prior error, allowing us to correct sudden changes based on random events - supply, popularity spikes, etc.

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

ARIMA models

- **ARIMA**, pronounced 'uh-ri-mah', is an AutoRegressive Integrated Moving Average model.
- In this model, we learn an ARMA(p, q) to predict not the value of the series, but the difference of the two series.
- An ARIMA model has three parameters and is specified ARIMA(p, d, q), where *p*, is the order of the autoregressive component, *q*, is the order of the moving average component, and *d* is the degree of differencing. In the above, we set *d = 1*.
- diff(X) = ($X_t - X_{t-1}$) → d = 1
- diff(diff(X)) → d = 2
- When d = 1, we basically de-trend linear relationship.

Summary

- We learned what time series are
- We discussed about Trend Line and Seasonality
- Auto Regression Models
- Moving Average Models
- ARMA models
- ARIMA models