Foundation of Data Science
Lecture 9, Module 1
Fall 2022

Rumi Chunara, PhD

Time Series Discussions

- Overview
- Basic definitions
- Time domain
- Forecasting

Moving Average Model: Overview

- The moving average model uses the last t periods in order to predict demand in period $t+1$
- There can be two types of moving average models:
  - simple moving average
  - weighted moving average

The moving average model assumption is that the most accurate prediction of future demand is a linear combination of past demand

Simple Moving Average

In the simple moving average models the forecast value is:

$$F_{t+1} = \frac{A_t + A_{t-1} + \ldots + A_{t-n}}{n}$$

t is the current period

F_{t+1} is the forecast for next period

n is the forecasting horizon (how far back we look)

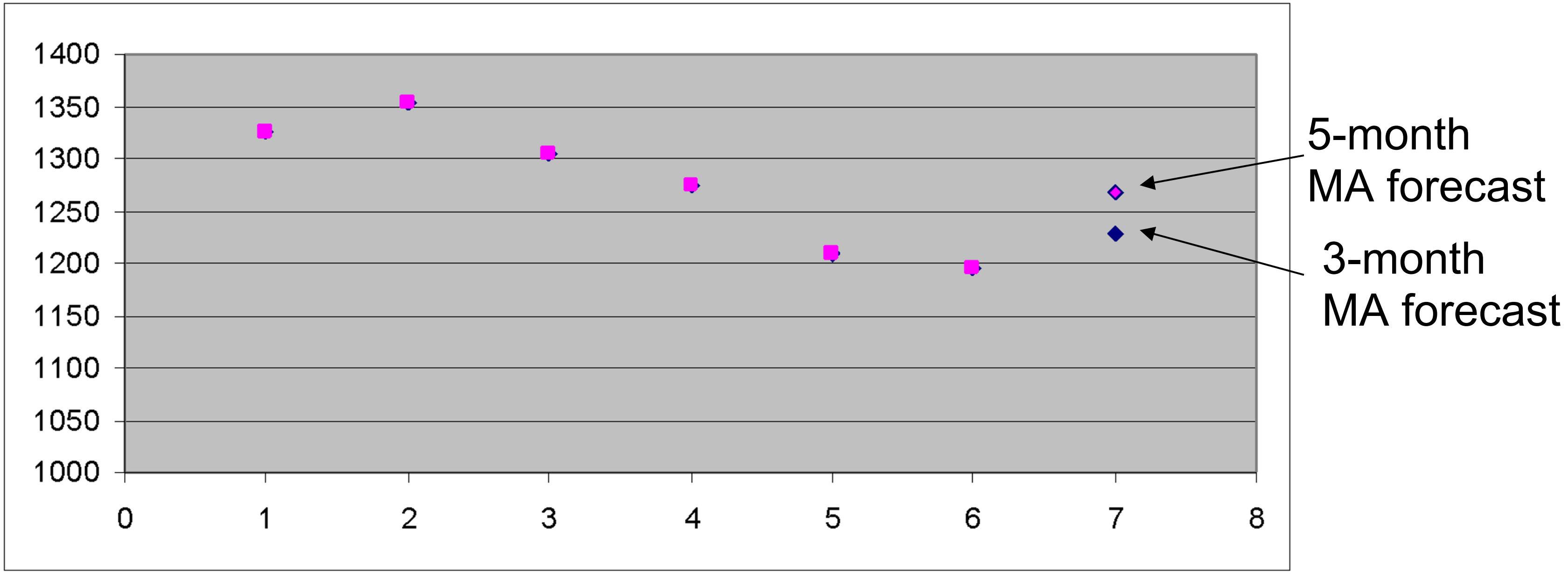
A is the actual amplitude (value) from each period

Example: Forecasting Sales at Kroger

Kroger sells (among other stuff) bottled spring water

Month	Bottles
Jan	*1,325*
Feb	*1,353*
Mar	*1,305*
Apr	*1,275*
May	*1,210*
Jun	*1,195*
Jul	*?*

What will the sales be for July?



- What do we observe?
  - 5-month average smooths data more
  - 3-month average is more responsive

Time series: Weighted Moving Average

We may want to give more importance to some of the data

$$F_{t+1} = \quad w_t A_t + w_{t-1} A_{t-1} + \dots + w_{t-n} A_{t-n}$$

$$\boldsymbol{w_t + w_{t-1} + \dots + w_{t-n} = 1}$$

t is the current period

F_{t+1} is the forecast for next period

n is the forecasting horizon (how far back we look)

A is the actual amplitude (value) from each period

w is the importance (weight) we give to each period

How Do We Choose Weights?

1. Depending on the importance that we feel past data has
2. Depending on known seasonality (weights of past data can also be zero)

WMA is more powerful than SMA because of the ability to vary the weights!

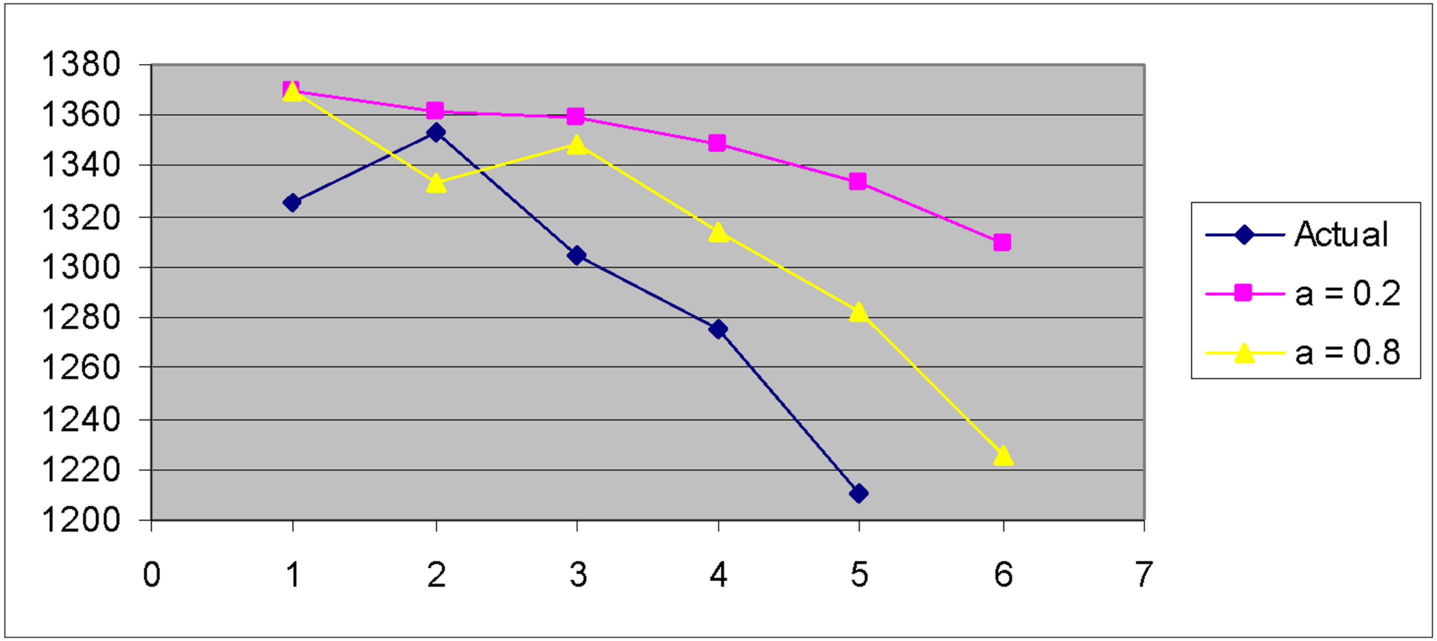
(Single) Exponential Smoothing

- Another linear combination-based forecasting method
- Parameter: smoothing constant α
  - controls the importance of prior times
  - importance decays exponentially
  - Larger α values mean that the model pays most attention to the most recent times, whereas smaller values mean more of the history is taken into account for forecast
  - 0 <= α <= 1

$$\boxed{\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots}$$

https://otexts.com/fpp2/ses.html

Impact of the Smoothing Constant



The actual values dropped “early on”, but when we use $\alpha = 0.2$, we consider the “older”, larger values for many timestamps!

How Can We Compare Models?

We need a metric that provides estimation of accuracy

Forecast Error

Errors can be:

- biased (consistent)
- random

Forecast error = Difference between actual and forecasted value (also known as *residual*)

Measuring Accuracy: MFE

- MFE = Mean Forecast Error (Bias)
  - It is the average error in the observations

$$\mathrm{MFE} = \frac{\sum_{i=1}^{n} A_t - F_t}{n}$$

- **MFE > 0**, model tends to **under-forecast**
- **MFE < 0**, model tends to **over-forecast**

A more positive or negative MFE implies worse performance

The forecast is biased

Measuring Accuracy: MAD

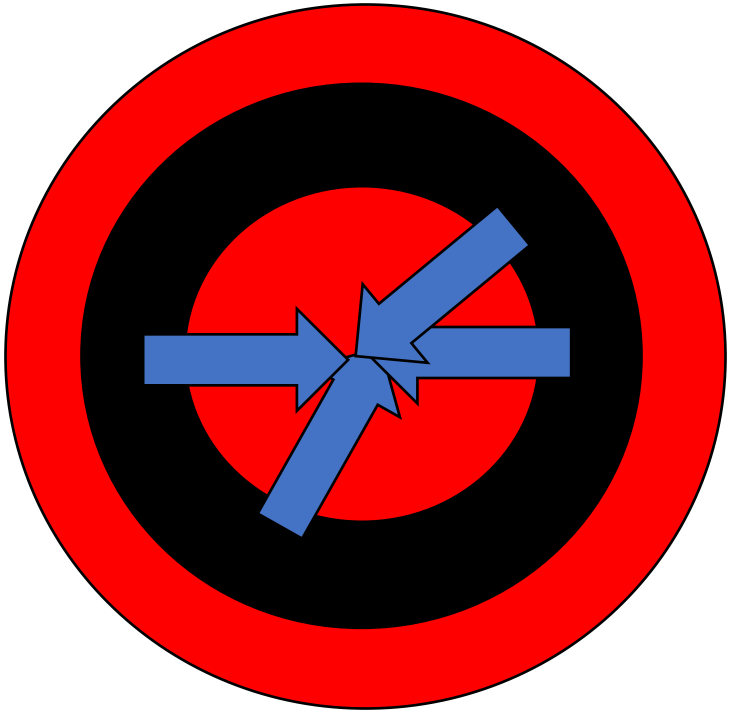
- MAD = Mean Absolute Deviation
    - It is the average absolute error in the observations

$$\text{MAD} = \frac{\sum_{i=1}^{n} \left| A_t - F_t \right|}{n}$$

Higher MAD implies worse performance.

If errors are normally distributed, then std =1.25MAD

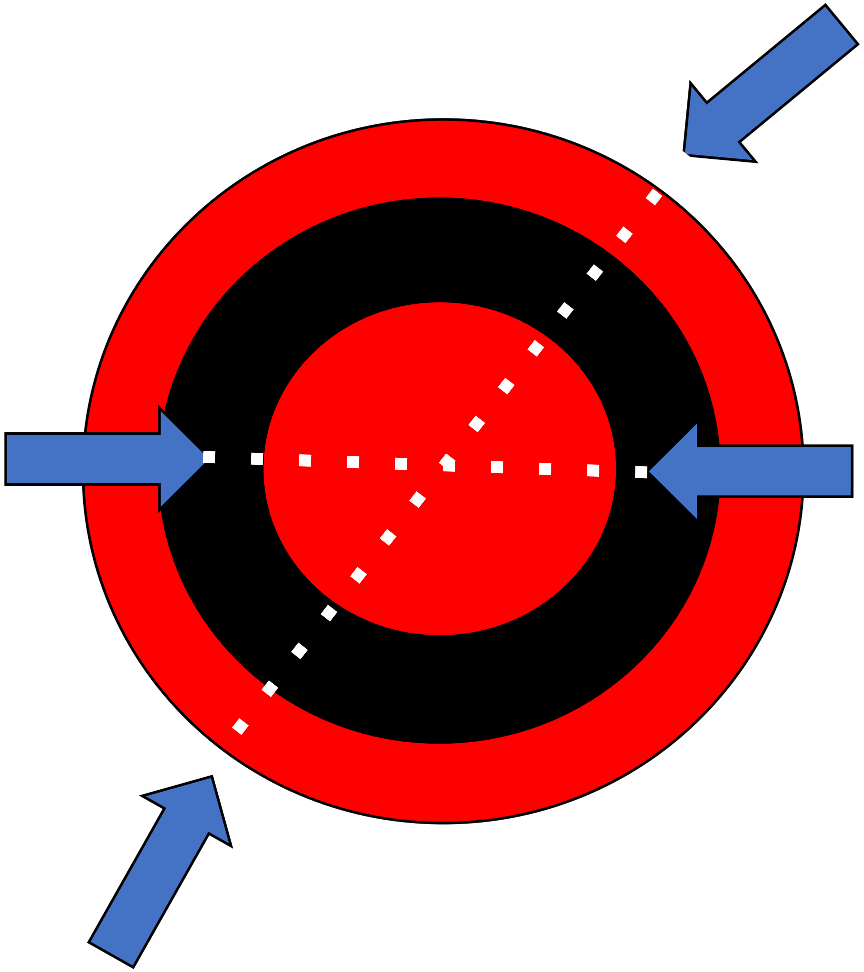
MFE & MAD: A Dartboard Analogy



Low MFE & MAD:

The forecast errors are small & unbiased

An Analogy (continued)

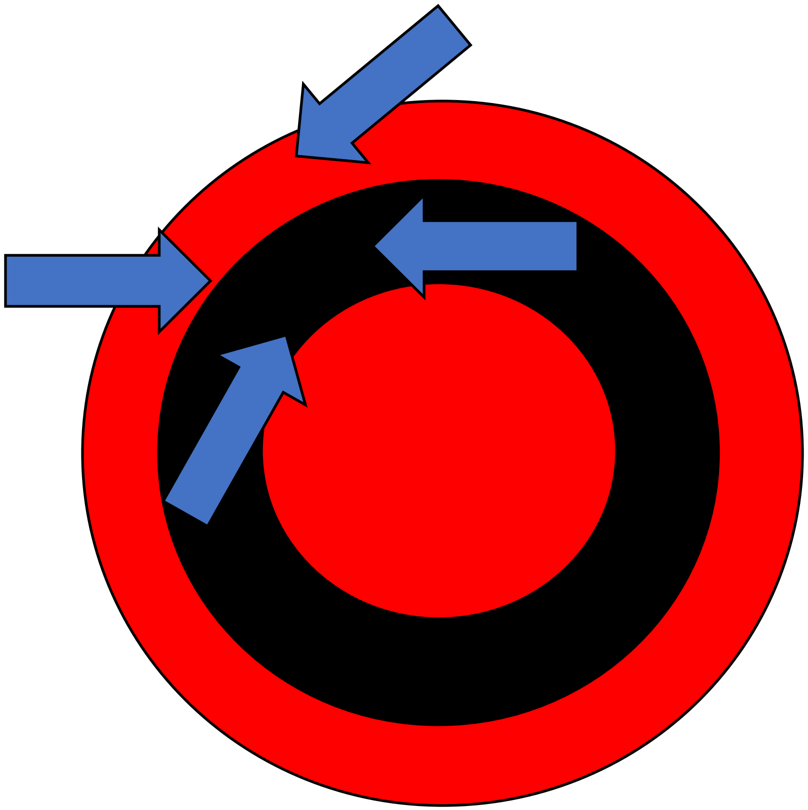


Low MFE but high MAD:

On average, the arrows hit the bullseye (so much for averages!)

How can we get this scenario?

An Analogy (continued)



High MFE & MAD:

The forecasts are inaccurate & biased

Measuring Accuracy: Tracking Signal

The tracking signal is a measure of how often our estimations have been above or below the actual value. It is used to decide whether to use a model.

$$\text{RSFE} = \sum_{i=1}^{n} (A_t - F_t) \qquad \text{TS} = \frac{\text{RSFE}}{\text{MAD}}$$

- Positive tracking signal: most of the time actual values are <u>above</u> our forecasted values
- Negative tracking signal: most of the time actual values are <u>below</u> our forecasted values

How to Decide Which Forecasting Method to Use?

1. Gather the historical data for forecasting
2. Divide data into initiation set and evaluation set
3. Use the first set to develop the models
4. Use the second set to evaluate
5. Compare the residuals, MADs and MFEs of each model