Date: Tuesday, November 27, 2018.

Project description

→ Write a report for an audience that has no technical time series knowledge.

Group 1.9: Tianqi Wang
Yihan Wang
Jon-Ross Presta

🎂 December 9th: Nathaniel's birthday

🍁 December 11th: Some Canadian statue's anniversary
↳ Also project is due
↳ Also Final Exam
↳ Also MSDS Holiday Party to let loose.

Exponential smoothing techniques
↳ (Holt-Winters Methodology)

lower case

The objective is to predict y_{n+h} given the observed history $\{y_1, y_2, \ldots, y_n\}$ of observations up to time point n. Using exponential smoothing techniques, we do so by using a set of recursive equations that do not require any distributional assumptions.

We will use different sets of recursive equations depending on whether the data has

(1) no trend + no seasonality ← single/simple ES (SES)

(2) trend + no seasonality ← double ES (DES)

(3) trend + seasonality ← Triple ES (TES)
↳ Exponential Smoothing

(1) Single Exponential Smoothing

"Level" Equation: $a_t = \alpha y_t + (1-\alpha) a_{t-1}$, $0 \le \alpha \le 1$ and a_0.

$\hat{Y}_{t+1}$

This is commonly referred to as an exponentially weighted moving average (EWMA).

anything, but $\bar{y}$ is commonly chosen.

- α here is the "smoothing constant".
  - → If $\alpha = 0$, then $a_t = a_0 \ \forall t$, in which case our prediction is a constant. (Straight line -- Extreme smoothing)
  - → If $\alpha = 1$, then $a_t = y_t$ which provides no smoothing. (Shifting today → tomorrow)
  - → So small α provide more smoothing and large α provide less.

- $\alpha = 0.2$ is typically a good choice, but an optimal value of α can be determined from the data. Specifically $\hat{\alpha}$ can be chosen as the value that minimizes one-step-ahead prediction error:

$$\sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n} (y_t - \hat{y}_t)^2$$

"Forecast" equation: $\boxed{\hat{Y}_{t+h} = a_t}$ for $h = 1, 2, 3, \ldots$

flat line
predicting the mean
it makes sense because no trend
+ no seasonality

* Where does the EWMA come from?

$$\begin{aligned}
a_t &= \alpha(y_t) + (1-\alpha) a_{t-1} \\
&= \alpha(y_t) + (1-\alpha)(\alpha y_{t-1} + (1-\alpha) a_{t-2}) \\
&= \alpha y_t + \alpha(1-\alpha) y_{t-1} + (1-\alpha)^2 a_{t-2} \\
&= \alpha y_t + \alpha(1-\alpha) y_{t-1} + (1-\alpha)^2 (\alpha y_{t-2} + (1-\alpha) a_{t-3}) \\
&= \alpha y_t + \alpha(1-\alpha) y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + (1-\alpha)^3 a_{t-3})
\end{aligned}$$

$$\vdots$$

$$= \alpha\left(\sum_{i=0}^{t-1}(1-\alpha)^i y_{t-i}\right) + (1-\alpha)^t a_0. \quad (*)$$

From this formula we can see that a_t is literally an exponential weighted moving average. We also see now, why the choice of a_0 doesn't matter.

(2) Double Exponential Smoothing

"Level" Equation: $a_t = \alpha y_t + (1-\alpha)(a_{t-1} + b_{t-1})$

weighted average of the observed value at time t, and its corresponding prediction from time point t-1

"Trend" Equation: $b_t = \beta(a_t - a_{t-1}) + (1-\beta) b_{t-1}$

weighted average of current observations and previous predicts of trend

"Forecast" equation: $\boxed{\hat{y}_{t+h} = a_t + h b_t}$ for $h = 1, 2, 3, \ldots$

trend to fold linearly

- α and β here are smoothing parameters, both in $[0,1]$, where smaller values provide more smoothing and large values provide less.

Both parameters may be estimated by minimizing the squared error loss function $(*)$.

(3) Triple Exponential Smoothing

"Level" Equation: $a_t = \alpha(y_t - S_{t-m}) + (1-\alpha)(a_{t-1} + b_{t-1})$

weighted average of seasonally adjusted observation and its non-seasonal forecast

"Trend" Equation: $b_t = \beta(a_t - a_{t-1}) + (1-\beta) b_{t-1}$

"Seasonal" Equation: $S_t = \gamma(y_t - a_{t-1} - b_{t-1}) + (1-\gamma) S_{t-m}$

today — previous time period

"Forecast" equation: $\boxed{\hat{y}_{t+h} = a_t + h b_t + S_{t+h-m}}$ for $h = 1, 2, 3, \ldots$

weighted average between the current seasonal index and the seasonal index of the same period of the previous season.

- α, β, γ are smoothing parameters in $[0, 1]$ which behave as usual and m is the period of the seasonal effect.

- this formulation is referred to as "additive", but when heteroscedasticity is present, we may opt to use the "multiplicative" version:

↳ $a_t = \alpha\left(\frac{y_t}{S_{t-m}}\right) + (1-\alpha)(a_{t-1} + b_{t-1})$

$b_t = \beta(a_t - a_{t-1}) + (1-\beta) b_{t-1}$

$S_t = \gamma\left(\frac{y_t}{a_{t-1} + b_{t-1}}\right) + (1-\gamma) S_{t-m}$

$$\boxed{\hat{y}_{t+h} = (a_t + h b_t) S_{t+h-m}}$$