The Theta Method

Advanced Financial Econometrics: Forecasting

This version: May 4, 2020

May 2020

The Theta Method Overview

- A method closely related to SES with a modified trend model
- Winner of the M3 competition
- Key parameter is θ
  - User Choice
  - Standardard choice is to equally weight models with $\theta = 0$ and $\theta = 2$

References

1. Assimakopoulos, V., & Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International journal of forecasting*, 16(4), 521-530.
2. Hyndman, R. J., & Billah, B. (2003). Unmasking the Theta method. *International Journal of Forecasting*, 19(2), 287-290.

The Theta Method

- Observed time series is $X_t,\ t = 1, 2, \ldots, T$
- Transform to a surrogate series

$$\Delta^2 Y_{t,\theta} = \theta \Delta^2 X_t$$

- This can be shows to relate the the levels as

$$Y_{t,\theta} = a_\theta + b_\theta\,(t-1) + \theta X_t$$

  - a_θ and b_θ are constants
  - $Y_{t,\theta}$ is called the *theta line*

Estimation of constants

- Given θ, a_θ and b_θ can be found using OLS

$$\sum_{t=1}^{T} \left(X_t - Y_{t,\theta}\right)^2 = \sum_{t=1}^{T} \left(X_t - a_\theta - b_\theta\left(t-1\right) - \theta X_t\right)^2$$
$$= \sum_{t=1}^{T} \left(\left(1-\theta\right) X_t - a_\theta - b_\theta\left(t-1\right)\right)^2$$

- This is just the regression

$$\left(1-\theta\right) X_t = a_\theta + b_\theta\left(t-1\right) + \epsilon_t$$

- The estimators are then

$$\hat{a}_\theta = \left(1-\theta\right)\bar{X} - \hat{b}_\theta\left(T-1\right)/2$$
$$\hat{b}_\theta = \frac{6\left(1-\theta\right)}{T^2-1}\left(\frac{2}{T}\sum_{t=1}^{T} tX_t - \left(T+1\right)\bar{X}\right)$$

- $\theta = 0$ corresponds to a standard regression on a time trend

Forecasts

- The original paper forecast the series as

$$\hat{X}_{T+h|T} = \frac{1}{2}\left(\hat{Y}_{T+h|T,0} + \hat{Y}_{T+h|T,2}\right)$$

$$\hat{Y}_{T+h|T,0} = \hat{a}_0 + \hat{b}_0\,(T + h - 1)$$

$$\hat{Y}_{T+h|T,2} = \alpha \sum_{i=0}^{T-1} (1-\alpha)^i\, Y_{T-i,2} + (1-\alpha)^T \underbrace{Y_{1,2}}_{\text{Parameter}}$$

- $\hat{Y}_{T+h|T,2}$ is a standard forecast from a SES (EWMA)
  - α is a smoothing parameter to be estimated
  - Initial value $Y_{1,2}$ is also be estimated

Forecasting

Alternative Expression

- These equations can be shown to be equivalent to

$$\hat{X}_{T+h|T} = \tilde{X}_{T+h|T} + \frac{1}{2}\hat{b}_0\left(h - 1 + \frac{1}{\alpha} - \frac{(1-\alpha)^T}{\alpha}\right)$$

$$\tilde{X}_{T+h|T} = \alpha\sum_{i=0}^{T-1}(1-\alpha)^i X_{T-i} + (1-\alpha)\,X_1$$

- The forecast is then the SES of X_t plus a constant and trend

Assumed Model

- The underlying model can be shown to be

$$X_t = X_{t-1} + b + (\alpha - 1)\,\epsilon_{t-1} + \epsilon_t$$

- Integrated MA(1) with a drift where $b = \hat{b}_0/2$
- Easy to show that one-step forecasts is

$$X_{T+1|T} = X_T + b + (\alpha - 1)\,\epsilon_t$$

- Multistep follows from *inverting* the MA

$$\begin{aligned}
\epsilon_t &= X_t - X_{t-1} - b - (\alpha - 1)\,\epsilon_{t-1} \\
&= X_t - X_{t-1} - b - (1 - \alpha)\,(X_{t-1} - X_{t-2} - b - (1 - \alpha)\,\epsilon_{t-2}) \\
&= \dots \\
&= \underbrace{\alpha \sum_{i=0}^{T-1} (1 - \alpha)^i X_{T-i} + (1 - \alpha)^T X_1}_{\text{SES}} + \frac{b}{\alpha}\left[1 - (1 - \alpha)^n\right] + (1 - \alpha)^n \underbrace{\epsilon_1}_{\text{Asm. } 0} \\
&= \tilde{X}_{T+1|T}
\end{aligned}$$

Forecasting and Prediction Intervals

- Using relationship between IMA and SES

$$X_{T+1|T} = \underbrace{\alpha \sum_{i=0}^{T-1} (1-\alpha)^i X_{T-i} + (1-\alpha)^T X_1}_{\text{SES}} + \frac{b}{\alpha}\left[1-(1-\alpha)^n\right] + (1-\alpha)^n \underbrace{\epsilon_1}_{\text{Asm. 0}}$$

$$= \tilde{X}_{T+1|T} + \frac{b}{\alpha}\left[1-(1-\alpha)^T\right]$$

$$X_{T+h|T} = \tilde{X}_{T+1|T} + b\left[h - 1 + \frac{1}{\alpha} - \frac{(1-\alpha)^T}{\alpha}\right]$$

  - $\alpha \approx 1$ then the trend is bh
  - Smaller values dampen the trend

- Alternative: Estimate parameters using MLE and the MA

$$\Delta X_t = b + (\alpha - 1)\,\epsilon_{t-1} + \epsilon_t$$

Conclusions

- The Theta Method is simple to implement
- Combines SES and a linear trend model
- Two parameters: b_0 and α
  - b_0 estimated using a standard time-trend model
  - α as part of optimizing the SES forecast
  - Alternatively jointly estimate both using MLE
- In practice damps the trend in the forecast