Core Concepts in Financial Econometrics

Romain Lafarguette, Ph.D.

ADIA Quant & IMF External Consultant

Singapore Training Institute, 08 November 2022

Perspective

- ETS (Error, Trend, Season) model was developed in the 1950s as algorithms to produce point forecasts
- ETS combines a "level" (l_{t-1}), a "trend" ((b_{t-1})) and a "seasonal" (s_{t-m}) components to describe a time series
- The combination $f(l_{t-1}, b_{t-1}, s_{t-m})$ can be additive, multiplicative, etc.
- The rate of change of the components are controlled by "smoothing" parameters: α for the level, β for the trend, γ for the seasonal
- The researcher has to:
    1. To choose the best values for the smoothing parameters
    2. The initial state of the parameters
- Equivalent ETS state-space models have been developped in the 1990s and the 2000s

Combining Level, Trend and Seasonal Components

- **Additively:** $y_t = l_{t-1} + b_{t-1} + s_{t-m} + \epsilon_t$
- **Multiplicatively:** $y_t = l_{t-1} \times b_{t-1} \times s_{t-m} \times (1 + \epsilon_t)$
- **Mixed:** $y_t = (l_{t-1} + b_{t-1}) \times s_{t-m} + \epsilon_t$

Notations:

- **Error** can be additive ("A") or multiplicative ("M")
- **Trend** can be None ("N"), additive ("A"), multiplicative ("M") or damped ("Ad" or "Md")
- Seasonality can be None ("N"), additive ("A") or multiplicative ("M")

Point forecasts and forecast distribution

- Models generates point forecasts estimates, based on a given conditional value $y_{t+1|y0} = f_y(y0)$
- However, need to generate the forecast distribution to assess the quality of models and select the best. **Model selection**
- A stochastic data generating process (DGP) can generate an entire forecast distribution
- Core idea: the residual (ϵ) is the only stochastic (=random) element in $y_t = l_{t-1} + b_{t-1} + s_{t-m} + \epsilon_t$
    - Hence, the distribution of the residuals will determine the distribution of the estimator

Level-Only Model ETS(A, N, N) (Simple Smoothing)

Component Form

- Forecast equation: $\hat{y}_{t+h|t} = l_t$
- Smoothing equation: $l_t = \alpha y_t + (1 - \alpha) l_{t-1}$

- l_t is the level (="smoothed value") of the series at time t
- $\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$

Iterate to get exponentially weighted moving average form:
$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1 - \alpha)^j y_{T-j} + (1 - \alpha)^T l_0$

Error Correction Form

Component Form

- Forecast equation: $\hat{y}_{t+h|t} = l_t$
- Smoothing equation: $l_t = \alpha y_t + (1 - \alpha) l_{t-1}$

Forecast error: $e_t = y_t - \hat{y}_{t|t-1} = y_t - l_{t-1}$

Error Correction Form

- $y_t = l_{t-1} + e_t$
- $l_t = l_{t-1} + \alpha * \underbrace{(y_t - l_{t-1})}_{e_t}$
- Intuition: α updates the next-period estimate based on the forecasting error
- Specify probability distribution for e_t, often assumed that $e_t = \epsilon_t \sim \mathcal{N}(0, \sigma^2)$

- Need to choose the best values for α and l_0
- Similarly to regression, choose optimal parameters by minimizing SSE:

$$\text{SSE} = \sum_{t=1}^{T} (y_t - \hat{y}_{t|t-1})^2$$

- Unlike regression, there is no closed form solution: need to use numerical optimization

State-Space Representation for Additive Models

State-Space Representation

- **Measurement equation**: $y_t = l_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$
  - The measurement equation is the relationship between observations (y_t) and state (=structure) l_t
- **State equation** $l_t = l_{t-1} + \alpha * \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$
  - The state equation is evolution of the state variable (l_t) through time

- Both equations have the same error process ϵ

State-Space Representation for Multiplicative Models

- Instead of differential errors, specify relative errors: $\eta_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}}$
- Some easy algebra, substituting $\hat{y}_{t|t-1} = l_{t-1}$ gives

State-Space Representation

- **Measurement equation**: $y_t = l_{t-1}(1 + \epsilon_t)$
- **State equation**: $l_t = l_{t-1}(1 + \alpha * \epsilon_t)$

Implication: Models with additivative and multiplicative errors with the same parameters generate **the same point forecasts but with different prediction intervals**

Holt's Linear Trend

Component Form

- Level: $l_t = \alpha_t + (1 - \alpha)(l_{t-1} + b_{t-1})$
- Trend: $b_t = \beta * (l_t - l_{t-1}) + (1 - \beta)b_{t-1}$
- Forecast: $\hat{y}_{t+h|t} = l_t + hb_t$

- Two smoothing parameters: α and β $(0 \leqslant \alpha,\ \beta \leqslant 1)$
- l_t level: weighted average between y_t and one-step ahead forecast for time t: $l_{t-1} + b_{t-1} = \hat{y}_{t|t-1}$
- b_t slope: weighted average of $(l_t - l_{t-1})$ and b_{t-1}, current and previous estimate of slope
- Choose $\alpha,\ \beta,\ l_0,\ b_0$ to mininize the SSE (sum squared errors)

Damped Trend Method

Component Form

- $l_t = \alpha y_t + (1-\alpha)(l_{t-1} + \phi b_{t-1})$
- $b_t = \beta * (l_t - l_{t-1}) + (1-\beta)\phi b_{t-1}$
- $\hat{y}_{t+h|t} = l_t + (\phi + \phi^2 + \cdots + \phi^h) b_t$

- ϕ is the damping parameter, $0 \leqslant \phi \leqslant 1$
- If $\phi = 1$, the method boils down to Holt's linear trend
- As $h \to \infty$, then $\hat{y}_{T+h|T} \ \to \ l_T + \frac{\phi b_T}{1-\phi}$
- Application: short-run forecasts are trended, long-run forecasts constant

Holt-Winters Seasonal Model

Holt and Winters extended Holt's method to capture seasonality

Component Form

- Level: $l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1})$
- Trend: $b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$
- Season: $\gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$
- Forecast: $\hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)}$

- m is the period of seasonality (e.g. $m = 4$ for quarterly data)

Seasonal Component

- The seasonal component is usually expressed as:

$$s_t = \gamma * (y_t - l_t) + (1 - \gamma) * s_{t-m} \qquad 0 \leqslant \gamma \leqslant 1$$

- By subtitution, we can derive the dynamic of the seasonal term as:

$$s_t = \gamma * (1 - \alpha)(y_t - l_{t-1} - b_{t-1}) + [1 - \gamma(1 - \alpha)]s_{t-m}$$

Forecasting with ETS models

Traditional point forecasts: iterate the equations for $t = T+1, T+2, \ldots, T+h$. By construction, $\epsilon_t = 0 \;\forall\; t > T$

- Equals to $E[y_{t+h}|x_t]$ in the case of additive seasonality only
- Point forecasts for additive ETS are the same as for multiplicative ETS if the parameters are the same

Prediction Intervals

- They can only be generated using the models
- The prediction intervals will differ between models with additive and multiplicative errors
- Some simple ETS models offer exact formula
- For more complex ETS models, the only solution for generating the confidence intervals is by bootstrapping
  - Simulate future sample paths, conditional on the last estimates of the states
  - Obtain the prediction intervals from the percentiles of these simulated future paths

Main Idea: Control the Rate of Change

- α controls the flexibility of the **level**
  - If $\alpha = 0$, the level never updates (stays at the mean)
  - If $\alpha = 1$, the level updates completely (naive, start from yesterday)
- β controls the flexibility of the **trend**
- If $\beta = 0$, the trend is linear
- If $\beta = 1$, the trend changes suddenly at each observation
- γ controls the flexibility of the **seasonality**
  - If $\gamma = 0$ the seasonality is fixed (seasonal mean)
  - If $\gamma = 1$ the seasonality updates completely (seasonal naive)

ARIMA Model

- AR: autoregressive (lagged observations as inputs)
- I: integrated (differencing to make series stationary)
- MA: moving average (lagged errors as inputs)

C

ontrary to an ETS, an ARIMA model is rarely interpretable in terms of visible data structures like trend and seasonality. But it can capture a huge range of time series patterns

Refresher: Intuitive Definition of Stationarity

Intuitive Characterization

If y_t is a stationary time series, then for any period s in the future, the distribution $\{y_t, \ldots, y_{t+s}\}$ doesn't depend on t

A **stationary series** is:

- Roughly horizontal
- Constant variance
- No predictable patterns in the long term