ETC3550/ETC5550 Applied forecasting

Week 6: Exponential smoothing

Historical perspective

- Developed in the 1950s and 1960s as methods (algorithms) to produce point forecasts.
- Combine a "level", "trend" (slope) and "seasonal" component to describe a time series.
- The rate of change of the components are controlled by "smoothing parameters": α, β and γ respectively.
- Need to choose best values for the smoothing parameters (and initial states).
- Equivalent ETS state space models developed in the 1990s and 2000s.

ETS models

Exponen**T**ial **S**moothing

Error **T**rend **S**easonal

ETS models

Exponen**T**ial **S**moothing

Error **T**rend **S**easonal

- ETS(A,N,N): **A**dditive Error, **N**o Trend, **N**o Seasonality

ETS models

Exponen**T**ial **S**moothing

Error **T**rend **S**easonal

- ETS(A,N,N): **A**dditive Error, **N**o Trend, **N**o Seasonality
- ETS(M,A,N): **M**ultiplicative Error, **A**dditive Trend, **N**o Seasonality

ETS models

Exponen**T**ial **S**moothing

Error **T**rend **S**easonal

- ETS(A,N,N): **A**dditive Error, **N**o Trend, **N**o Seasonality
- ETS(M,A,N): **M**ultiplicative Error, **A**dditive Trend, **N**o Seasonality
- ETS(M,A,M): **M**ultiplicative Error, **A**dditive Trend, **M**ultiplicative Seasonality

etc

Big idea: control the rate of change

α controls the flexibility of the **level**

- If $\alpha = 0$, the level never updates (mean)
- If $\alpha = 1$, the level updates completely (naive)

β controls the flexibility of the **trend**

- If $\beta = 0$, the trend is linear
- If $\beta = 1$, the trend changes suddenly every observation

γ controls the flexibility of the **seasonality**

- If $\gamma = 0$, the seasonality is fixed (seasonal means)
- If $\gamma = 1$, the seasonality updates completely (seasonal naive)

Models and methods

Methods

- Algorithms that return point forecasts.

Models

- Generate same point forecasts but can also generate forecast distributions.
- A stochastic (or random) data generating process that can generate an entire forecast distribution.
- Allow for “proper” model selection.

ETS(A,N,N): SES with additive errors

State equation	$\ell_t = \ell_{t-1} + \alpha\varepsilon_t$
Observation equation	$y_t = \ell_{t-1} + \varepsilon_t$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$

ETS(A,N,N): SES with additive errors

State equation	$\ell_t = \ell_{t-1} + \alpha\varepsilon_t$
Observation equation	$y_t = \ell_{t-1} + \varepsilon_t$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$

- Forecast errors: $\varepsilon_t = y_t - \ell_{t-1} \sim \text{NID}(0, \sigma^2)$
- "innovations" or "single source of error" because equations have the same error process, ε_t.
- equivalent forecasts:

$$\hat{y}_{t+h|t} = (1-\alpha)^t \ell_0 + \sum_{j=0}^{t-1} \alpha(1-\alpha)^j y_{t-j}$$

ETS(A,A,N): Holt's method with additive errors

State equations	$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$
	$b_t = b_{t-1} + \beta\varepsilon_t$
Observation equation	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + hb_t$

- Forecast errors: $\varepsilon_t = y_t - \ell_{t-1} \sim \text{NID}(0, \sigma^2)$

ETS(A,A,A): Holt-Winters additive method

State equations	$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$
	$b_t = b_{t-1} + \beta\varepsilon_t$
	$s_t = s_{t-m} + \gamma\varepsilon_t$
Observation equation	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$

- Forecast errors: $\varepsilon_t = y_t - \ell_{t-1} \sim \text{NID}(0, \sigma^2)$
- k is integer part of $(h-1)/m$.

ETS(M,N,N): SES with multiplicative errors.

$$\begin{aligned} \text{State equation} \quad & \ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t) \\ \text{Observation equation} \quad & y_t = \ell_{t-1}(1 + \varepsilon_t) \\ \text{Forecast equation} \quad & \hat{y}_{t+h|t} = \ell_t \end{aligned}$$

ETS(M,N,N): SES with multiplicative errors.

State equation	$\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$
Observation equation	$y_t = \ell_{t-1}(1 + \varepsilon_t)$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$

- Relative forecast errors: $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Same point forecasts as ETS(A,N,N) with additive errors and same α.
- Different prediction intervals from ETS(A,N,N).

ETS(M,A,N): Holt's method with multiplicative errors.

State equations	$\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$
	$b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$
Observation equation	$y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t)$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + hb_t$

- Relative forecast errors: $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Same point forecasts as ETS(A,A,N) with additive errors and same α and β
- Different prediction intervals from ETS(A,A,N).

ETS(M,A,M): Holt-Winters multiplicative method

State equations	$\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$
	$b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$
	$s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
Observation equation	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$
Forecast equation	$\hat{y}_{t+h\|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$

- Relative forecast errors: $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- k is integer part of $(h-1)/m$.

ETS models

Additive Error

	Trend Component	**Seasonal Component** N (None)	A (Additive)	M (Multiplicative)
N	(None)	A,N,N	A,N,A	A,N,M
A	(Additive)	A,A,N	A,A,A	A,A,M
A_d	(Additive damped)	A,A_d,N	A,A_d,A	A,A_d,M

Multiplicative Error

	Trend Component	**Seasonal Component** N (None)	A (Additive)	M (Multiplicative)
N	(None)	M,N,N	M,N,A	M,N,M
A	(Additive)	M,A,N	M,A,A	M,A,M
A_d	(Additive damped)	M,A_d,N	M,A_d,A	M,A_d,M

Additive error models

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$
A	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$
$\mathbf{A_d}$	$y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$

Multiplicative error models

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$	$y_t = (\ell_{t-1}+s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}+\alpha(\ell_{t-1}+s_{t-m})\varepsilon_t$ $s_t = s_{t-m}+\gamma(\ell_{t-1}+s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A	$y_t = (\ell_{t-1}+b_{t-1})(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1}+b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}+\beta(\ell_{t-1}+b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1}+b_{t-1}+s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}+b_{t-1}+\alpha(\ell_{t-1}+b_{t-1}+s_{t-m})\varepsilon_t$ $b_t = b_{t-1}+\beta(\ell_{t-1}+b_{t-1}+s_{t-m})\varepsilon_t$ $s_t = s_{t-m}+\gamma(\ell_{t-1}+b_{t-1}+s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1}+b_{t-1})s_{t-m}(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1}+b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}+\beta(\ell_{t-1}+b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A$_\mathbf{d}$	$y_t = (\ell_{t-1}+\phi b_{t-1})(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1}+\phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1}+\beta(\ell_{t-1}+\phi b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1}+\phi b_{t-1}+s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}+\phi b_{t-1}+\alpha(\ell_{t-1}+\phi b_{t-1}+s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1}+\beta(\ell_{t-1}+\phi b_{t-1}+s_{t-m})\varepsilon_t$ $s_t = s_{t-m}+\gamma(\ell_{t-1}+\phi b_{t-1}+s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1}+\phi b_{t-1})s_{t-m}(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1}+\phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1}+\beta(\ell_{t-1}+\phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$