ETC3550/ETC5550 Applied forecasting

Week 6: Non-seasonal 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 and the PBS

ETS models

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

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

ETS models

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

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

Error: Additive, Multiplicative

Trend: None, Additive, Additive damped

Seasonal: None, Additive, Multiplicative

ETS models

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

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

Error: Additive, Multiplicative

Trend: None, Additive, Additive damped

Seasonal: None, Additive, Multiplicative

- 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

Error: Additive, Multiplicative

Trend: None, Additive, Additive damped

Seasonal: None, Additive, Multiplicative

- 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

Error: Additive, Multiplicative

Trend: None, Additive, Additive damped

Seasonal: None, Additive, Multiplicative

- 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

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

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

Note that $\varepsilon_t = y_t - \ell_{t-1}$. So

$$\begin{aligned}
\hat{y}_{t+1|t} = \ell_t &= \ell_{t-1} + \alpha\varepsilon_t \\
&= \ell_{t-1} + \alpha(y_t - \ell_{t-1}) \\
&= \alpha y_t + (1-\alpha)\ell_{t-1} \\
&= \alpha y_t + (1-\alpha)[\alpha y_{t-1} + (1-\alpha)\ell_{t-2}] \\
&= \alpha y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2\ell_{t-2} \\
&\ldots \\
&= \alpha\sum_{j=0}^{t-1}(1-\alpha)^j y_{t-j} + (1-\alpha)^t\ell_0
\end{aligned}$$

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

Note that $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$. So

$$\begin{aligned}
y_{T+h} &= \ell_{T+h-1} + \varepsilon_{T+h} \\
&= \ell_{T+h-2} + \alpha\varepsilon_{T+h-1} + \varepsilon_{T+h} \\
&= \ell_{T+h-3} + \alpha\varepsilon_{T+h-2} + \alpha\varepsilon_{T+h-1} + \varepsilon_{T+h} \\
&\dots \\
&= \ell_T + \alpha \sum_{j=1}^{h-1} \varepsilon_{T+h-j} + \varepsilon_{T+h}.
\end{aligned}$$

$$\text{Therefore} \quad \mathrm{E}(y_{T+h}|y_1,\dots,y_T) = \ell_T$$

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

Note that $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$. So

$$\begin{aligned}
y_{T+h} &= \ell_{T+h-1} + \varepsilon_{T+h} \\
&= \ell_{T+h-2} + \alpha\varepsilon_{T+h-1} + \varepsilon_{T+h} \\
&= \ell_{T+h-3} + \alpha\varepsilon_{T+h-2} + \alpha\varepsilon_{T+h-1} + \varepsilon_{T+h} \\
&\dots \\
&= \ell_T + \alpha\sum_{j=1}^{h-1}\varepsilon_{T+h-j} + \varepsilon_{T+h}.
\end{aligned}$$

Therefore $\quad \text{Var}(y_{T+h}|y_1,\dots,y_T) = \alpha^2\sum_{j=1}^{h-1}\sigma^2 + \sigma^2 = \sigma^2\left[1 + \alpha^2(h-1)\right].$

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: $\varepsilon_t = y_t - \hat{y}_{t|t-1} \sim \text{NID}(0, \sigma^2)$

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).

Summary of models so far

Simple exponential smoothing

ETS(A,N,N) – additive errors

ETS(M,N,N) – multiplicative errors

Holt's linear trend method

ETS(A,A,N) – additive errors, additive trend

ETS(M,A,N) – multiplicative errors, additive trend

ETS(A,Ad,N): Damped trend with additive errors

$$\begin{aligned}
\text{State equations} \qquad \ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t \\
b_t &= \phi b_{t-1} + \beta\varepsilon_t \\
\text{Observation equation} \qquad y_t &= \ell_{t-1} + \phi b_{t-1} + \varepsilon_t \\
\text{Forecast equation} \qquad \hat{y}_{t+h|t} &= \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t
\end{aligned}$$

- Forecast errors: $\varepsilon_t = y_t - \hat{y}_{t|t-1} \sim \text{NID}(0, \sigma^2)$
- Damping parameter $0 < \phi < 1$.
- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1-\phi)$.
- Short-run forecasts trended, long-run forecasts constant.

ETS(M,Ad,N): Damped trend with multiplicative errors

State equations	$\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$
	$b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$
Observation equation	$y_t = (\ell_{t-1} + \phi b_{t-1})(1 + \varepsilon_t)$
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$

- Relative forecast errors: $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Damping parameter $0 < \phi < 1$.
- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1 - \phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Non-seasonal ETS models

Simple exponential smoothing

ETS(A,N,N) – additive errors

ETS(M,N,N) – multiplicative errors

Holt's linear trend method

ETS(A,A,N) – additive errors, additive trend

ETS(M,A,N) – multiplicative errors, additive trend

Damped trend method

ETS(A,Ad,N) – additive errors, damped trend

ETS(M,Ad,N) – multiplicative errors, damped trend

Parameters

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

ϕ controls the rate of **damping**

- If $\phi = 1$, there is no damping (trend is linear)
- If $0 < \phi < 1$, the trend converges to constant

States

ℓ_t = level at time t

b_t = slope at time t