ECON 4115/5115 Outline of Lecture Notes

Chapter 8. Exponential Smoothing

- Simple exponential smoothing
  - Exponential smoothing forecasts involve taking a weighted average of past observations
  - Weights decline exponentially as observations are more distant
  - Simple exponential smoothing (SES) is a compromise between …
    - the naïve method: $\hat{y}_{T+h|T} = y_T$ &
    - the average method: $\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t$
  - SES: $\hat{y}_{T+h|T} = \alpha y_T + \alpha(1-\alpha)^1 y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots$
  - $0 \le \alpha \le 1$ is the smoothing parameter
    - Distant observations get more weight when α is small (i.e., close to zero)
    - Recent observations get more weight when α is close to one
    - $\alpha = 1$ produces naïve forecasts
  - SES can be written in weighted-average form: $\hat{y}_{T+1|T} = \alpha y_T + (1-\alpha)\hat{y}_{T|T-1}$
  - SES will give flat forecasts: $\hat{y}_{T+h|T} = \hat{y}_{T+1|T}$ for $h = 2,3,4,\ldots$
    - Flat forecasts are reasonable if there's no trend or seasonality
  - You can choose α by minimizing the SSE: $\sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T}(y_t - \hat{y}_{t|t-1})^2$
    - This is a nonlinear optimization problem
- Methods with trend
  - Holt's component trend method

- Forecasting equation: $\hat{y}_{t+h|t} = \ell_t + hb_t$
- Level equation: $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$
- Trend equation: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$

  - α and β^* are smoothing parameters for the level and trend, respectively
  - Forecasts are no longer flat; increasing in b_t
  - Smoothing parameters and initial values, ℓ_0 and b_0, need to be estimated
  - Damped trend model may forecast better
    - Adds a damping parameter, ϕ, to Holt's trend method
    - Forecasting equation: $\hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \phi^h)b_t$
    - Level equation: $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$
    - Trend equation: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$
  - Typically, $0.8 \leq \phi \leq 1$

- Methods with seasonality
  - Holt and Winters extend Holt's method to handle seasonality
  - Seasonality can be quarterly ($m = 4$), monthly ($m = 12$), daily ($m = 7$), etc.
  - Forecasting equation and three smoothing equations:
    - Forecasting equation: $\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$
    - Level equation: $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$
    - Trend equation: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$
    - Seasonal equation: $s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$
  - The three smoothing parameters $(\alpha, \beta^*, \gamma)$ are usually estimated

- The parameter k is the integer part of $(h-1)/m$
- There is also a multiplicative version of Holt-Winters model
- Holt-Winters damped trend method with multiplicative seasonality often works well
  - Forecasting equation: $\hat{y}_{t+h|t} = [\ell_t + (\phi + \phi^2 + \phi^h)b_t]s_{t+h-m(k+1)}$
  - Level equation: $\ell_t = \alpha(y_t/s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$
  - Trend equation: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$
  - Seasonal equation: $s_t = \gamma(y_t/(\ell_{t-1} + \phi b_{t-1}) + (1-\gamma)s_{t-m}$
- Damping parameter is $0 < \phi < 1$

➢ A taxonomy of exponential smoothing methods

- See Table 8.5 with “N”, “A”, “M” and “Ad” notation used in R
- See Table 8.6 for component equations

➢ Innovations state-space models for exponential smoothing

- We now express the exponential smoothing model as a state-space model
- Allows us to generate confidence intervals
- ETS(•,•,•) = ETS(<u>E</u>rror, <u>T</u>rend, <u>S</u>easonal) with components:
  - Error = $\{A, M\}$
  - Trend = $\{N, A, A_d\}$
  - Seasonal = $\{N, A, M\}$
- ETS(A,N,N): Simple exponential smoothing with additive errors
  - Component form:
    - Forecasting equation: $\hat{y}_{t+h|t} = \ell_t$

        - Level equation: $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$
    - State-space form (where $\varepsilon_t \sim NID(0, \sigma^2)$):
        - Measurement or observation equation: $y_t = \ell_{t-1} + \varepsilon_t$
        - State or transition equation: $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$
    - Smoothing parameter (α):
        - If $\alpha = 0$, the level of the series (ℓ_t) doesn't change over time
        - If $\alpha = 1$, the model collapses to a random walk: $y_t = y_{t-1} + \varepsilon_t$
  - See Table 8.7 for ETS additive and multiplicative state-space models

- Estimation and model selection
  - State-space parameters are typically estimated with maximum likelihood methods
  - The AIC, AIC_c, and BIC can be used to select the "best" model
  - See R script #11 for an example of ETS estimation, model selection and forecasting

- Forecasting with ETS models
  - Forecasts in R are given by the *forecast*() function
  - Point forecasts, $\hat{y}_{T+h|T}$, are given by iterating state-space equations forward
  - Prediction intervals are given by $\hat{y}_{T+h|T} \pm c\sigma_h$
    - c depends on the coverage probability
    - σ_h^2 is the forecast variance with formula given in Table 8.8