BGN: 2191A

P6

Q8

a. A stationary distribution is a probability distribution π over the possible values of a random variable $\cancel{X_t}$ in a time series such that

$$\forall t \geq 0 \ (X_t \sim \pi) \Rightarrow (X_{t+1} \sim \pi)$$

Let t arbitrary

Assume $E_t \sim \text{Normal}(0, \sigma^2)$

$$E_{t+1} = \lambda E_t + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$$\therefore E_{t+1} \sim \lambda \text{Normal}(0, \sigma^2) + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$$= \text{Normal}(0, (\lambda\sigma)^2) + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$$= \text{Normal}(0+0, (\lambda\sigma)^2 + (1-\lambda^2)\sigma^2)$$

$$= \text{Normal}(0, (1-\lambda^2+\lambda^2)\sigma^2)$$

$$= \text{Normal}(0, \sigma^2)$$

$$\sim E_t$$

b. "Memorylessness" refers to the property that the distribution of E_{t+1} depends solely on the value of E_t, not any E_{t-i} for $i \geq 1$

$$\log(\text{lik}) = \log \Pr(e_1, \ldots, e_n \mid e_0)$$

$$= \log\left(\Pr(e_1 \mid e_0)\Pr(e_2 \mid e_1, e_0) \ldots \Pr(e_n \mid e_{n-1}, \ldots e_0)\right)$$

$$= \log\left(\Pr(e_1 \mid e_0)\Pr(e_2 \mid e_1) \ldots \Pr(e_n \mid e_{n-1})\right) \text{ by memorylessness}$$

$$= \sum_{i=1}^{n} \log \Pr(e_n \mid e_{n-1})$$

c. $$Y_{t+1} = \alpha + \beta\sin(2\pi\omega(t+1)) + \gamma(t+1) + E_{t+1}$$

$$= \alpha + \beta\sin(2\pi\omega(t+1)) + \gamma(t+1) + \lambda E_t + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$$= \lambda\alpha - \lambda\alpha + \alpha + \beta\sin(2\pi\omega(t+1)) + \lambda\beta\sin(2\pi\omega t) - \lambda\beta\sin(2\pi\omega t) + \lambda\gamma t - \lambda\gamma t + \gamma(t+1)$$
$$+ \lambda E_t + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$$= \lambda Y_t + \alpha(1-\lambda) + \beta\left(\sin(2\pi\omega(t+1)) - \lambda\sin(2\pi\omega t)\right) + \gamma(t+1-\lambda t) + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

$\therefore Y$ is memoryless

$$\therefore \log(\text{lik}) = \log\left(\prod_{i=0}^{t} \Pr(y_{i+1} \mid y_i)\right)$$

$$= \sum_{i=0}^{t} \log \Pr(y_{i+1} \mid y_i)$$

d. A linear model is one in which the estimated vector is a linear combination of known feature vectors, with the parameters of the model being scalar multipliers.

from (c):

$$Y_{t+1} = \lambda Y_t + \alpha(1-\lambda) + \beta\left(\sin(2\pi\omega(t+1)) - \lambda\sin(2\pi\omega t)\right) + \gamma(t+1-\lambda t) + \text{Normal}(0, (1-\lambda^2)\sigma^2)$$

Let $\Delta_t = Y_{t+1} - \lambda Y_t$

Δ_t is the variable we will be modelling.

Let $f_1 = 1-\lambda$

$f_2 = \sin(2\pi\omega(t+1)) - \lambda\sin(2\pi\omega t)$

$f_3 = t+1-\lambda t$

be feature vectors

$$\Delta_t \triangleq \alpha f_1 + \beta f_2 + \gamma f_3$$