Student: NINH DO

SID # 25949105

CS 287 - Homework #4

Problem #1

$$P(x=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

The likelihood of samples $x^{(1)}, \dots, x^{(m)}$:

$$L = \prod_{i=1}^{m} P(x = x^{(i)}) = \frac{\lambda^{\sum_1^m x^{(i)}} e^{-m\lambda}}{\prod_1^m (x^{(i)}!)}$$

The log-likelihood:

$$l = \log(L) = \log\left(\frac{\lambda^{\sum_1^m x^{(i)}} e^{-m\lambda}}{\prod_1^m (x^{(i)}!)} \right)$$

$$= \log\left(\lambda^{\sum_1^m x^{(i)}}\right) + \log(e^{-m\lambda}) - \log\left(\prod_1^m (x^{(i)}!)\right)$$

$$= \sum_1^m x^{(i)} \log\lambda - m\lambda - \log\left(\prod_1^m (x^{(i)}!)\right)$$

$$\frac{dl}{d\lambda} = \frac{\sum_1^m x^{(i)}}{\lambda} - m = 0 \Rightarrow \lambda^* = \frac{\sum_1^m x^{(i)}}{m}$$

$$\frac{d^2 l}{d\lambda^2} = -\frac{\sum_1^m x^{(i)}}{\lambda^2} < 0$$

$$\Rightarrow \lambda^* = \frac{\sum_1^m x^{(i)}}{m}$$ maximizes log-likelihood and thereby likelihood

Problem #2

$$\Sigma = E[(X-EX)(X-EX)^T]$$

For any matrix M, we have $MM^T = (MM^T)^T$, i.e MM^T symmetric

Thus $(X-EX)(X-EX)^T$ symmetric

$$\Rightarrow \quad \Sigma = E[(X-EX)(X-EX)^T] \quad \text{symmetric}$$

$$\Sigma_{ij} = \Sigma_{ji} \qquad (1)$$

Consider $\quad x^T \Sigma x \qquad \forall x \in R^n$

$$= x^T E[(X-EX)(X-EX)^T] x$$

$$= E[x^T (X-EX)(X-EX)^T x]$$

$$= E[[(X-EX)^T x]^T [(X-EX)^T x]]$$

let $v = (X-EX)^T x$

$$\Rightarrow [(X-EX)^T x]^T [(X-EX)^T x] = v^T v = \|v\|_2^2 \geqslant 0$$

$$\Rightarrow x^T \Sigma x = E[\|v\|_2^2] \geqslant 0 \qquad (2)$$

From (1) & (2) $\Rightarrow$ Σ is positive semi-definite

Problem #3

pd $A \in R^{n \times n}$, $b \in R^n$, $c \in R$

$$I = \oint_{x \in R^n} \exp\left(-\frac{1}{2} x^T A x - x^T b - c\right) dx$$

Since A is positive definite, there always exists $A^{1/2}$ positive definite

s.t $A = A^{1/2} A^{1/2}$ $\quad (A = Q \Lambda Q^{-1} = Q \Lambda^{1/2} \Lambda^{1/2} Q^{-1}$

$= A^{T/2} A^{1/2}$ $\quad = Q \Lambda^{1/2} Q^{-1} Q \Lambda^{1/2} Q^{-1} = A^{1/2} A^{1/2})$

Consider the exponent:

$$-\frac{1}{2} x^T A x - x^T b - c$$

$$= -\frac{1}{2} x^T A^{T/2} A^{1/2} x - x^T A^{T/2} A^{-1/2} b - \frac{1}{2} b^T A^{-T/2} A^{-1/2} b + \frac{1}{2} b^T A^{-1} b - c$$

$$= -\frac{1}{2} (A^{1/2} x + A^{-1/2} b)^T (A^{1/2} x + A^{-1/2} b)$$

$$= -\frac{1}{2} (x + A^{-1} b)^T A (x + A^{-1} b) + \frac{1}{2} b^T A^{-1} b - c$$

$$\Rightarrow \exp\left(-\frac{1}{2} x^T A x - x^T b - c\right)$$

$$= \exp\left(-\frac{1}{2} (x + A^{-1} b)^T (A^{-1})^{-1} (x + A^{-1} b)\right) \exp\left(\frac{1}{2} b^T A^{-1} b - c\right)$$

let $\mu = -A^{-1} b$, $\Sigma = A^{-1}$

$$\Rightarrow I = \oint_{x \in R^n} \exp\left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right) \exp\left(\frac{1}{2} b^T A^{-1} b - c\right) dx$$

$$= \oint_{x \in R^n} \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \frac{(2\pi)^{n/2}|\Sigma|^{1/2}}{\exp\left(c-\frac{1}{2}b^TA^{-1}b\right)} \, dx$$

$$= \frac{(2\pi)^{n/2}|A^{-1}|^{1/2}}{\exp\left(c-\frac{1}{2}b^TA^{-1}b\right)} \underbrace{\oint_{x \in R^n} \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) dx}_{1}$$

$$= \frac{(2\pi)^{n/2}}{|A|^{1/2}\exp\left(c-\frac{1}{2}b^TA^{-1}b\right)}$$

Problem #4.e

$$x_0 \sim N(\mu_0, \Sigma_0)$$
$$x_{t+1} = Ax_t + w_t$$
$$w_t = 0.3\,w_{t-1} + 0.2\,w_{t-2} + p_{t-1}$$
$$p_t \sim N(0, \Sigma_{pp})$$
$$y_t = Cx_t + v_t$$
$$v_t = 0.8\,v_{t-1} + q_{t-1}$$
$$q_t \sim N(0, \Sigma_{qq})$$
$$p_{-1} = q_{-1} = v_{-1} = w_{-1} = w_{-2} = 0$$

Rewrite everything in the form :

$$\begin{aligned}
x_{t+1} &= Ax_t + w_t \\
w_{t+1} &= 0.3\,w_t + s_t + p_t \\
s_{t+1} &= 0.2\,w_t \\
v_{t+1} &= 0.8\,v_t + q_t \\
y_t &= Cx_t + v_t
\end{aligned}$$

Then,

$$\underbrace{\begin{bmatrix} x_{t+1} \\ w_{t+1} \\ s_{t+1} \\ v_{t+1} \end{bmatrix}}_{x'_{t+1}} = \underbrace{\begin{bmatrix} A & 1 & 0 & 0 \\ 0 & 0.3 & 1 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0.8 \end{bmatrix}}_{A'} \underbrace{\begin{bmatrix} x_t \\ w_t \\ s_t \\ v_t \end{bmatrix}}_{x'_t} + \underbrace{\begin{bmatrix} 0 \\ p_t \\ 0 \\ q_t \end{bmatrix}}_{\varphi_t}$$

$$y_t = \underbrace{\begin{bmatrix} C & 0 & 0 & 1 \end{bmatrix}}_{C'} \begin{bmatrix} x_t \\ w_t \\ s_t \\ v_t \end{bmatrix}$$

The augmented state representation is :

$$x'_t = [x_t \; w_t \; s_t \; v_t]^T$$

The dynamics model :

$$x'_{t+1} = A' x'_t + \varphi_t$$

and the measurement model :

$$y_t = C' x'_t$$

where

$$A' = \begin{bmatrix} A & 1 & 0 & 0 \\ 0 & 0.3 & 1 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0.8 \end{bmatrix} \qquad C' = [C \; 0 \; 0 \; 1]$$

$$\varphi_t = \begin{bmatrix} 0 \\ p_t \\ 0 \\ q_t \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & & \\ & \Sigma_{pp} & \Sigma_{pq} \\ & 0 & \\ & \Sigma_{qp} & \Sigma_{qq} \end{bmatrix} \right)$$

$\Sigma_{pq} = \Sigma_{qp} = 0$ since p_t and q_t are uncorrelated

φ_t is uncorrelated noise