Tutorial 4

Wednesday, April 3, 2019 7:23 PM

7.1

7.1.a.

$$x^T A x = (x_1, x_2 \dots, x_N) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ \vdots & & & \vdots \\ a_{N1} & \dots & & a_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = (x_1 x_2 \dots x_N) \begin{pmatrix} \sum_{j=1}^{N} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{N} a_{Nj} x_j \end{pmatrix} =$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} x_i a_{ij} x_j =$$

$$= x_e^2 a_{ee} + \sum_{i \neq e}^{N} x_i a_{ie} x_e + \sum_{j \neq e}^{N} x_e a_{ej} x_j + \sum_{j \neq e}^{N} \sum_{i \neq e}^{N} x_i x_j a_{ij}$$

$$= 2x_e a_{ee} + \sum_{i \neq e}^{N} a_{ie} x_i + \sum_{j \neq e}^{N} a_{ej} x_j = \sum_{i=1}^{N} a_{ie} x_i + \sum_{j=1}^{N} a_{ej} x_j$$

$$\frac{\partial (x^T A x)}{\partial x_e} = 2x_e a_{ee} + \sum_{i \neq e}^{N} a_{ie} x_i \Rightarrow \frac{\partial (x^T A x)}{\partial x} = \begin{pmatrix} \sum_{i=1}^{N} a_{i1} x_i \\ \vdots \\ \sum_{i=1}^{N} a_{iN} x_i \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{N} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{N} a_{Nj} x_j \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} a_{11} & a_{21} & \dots & a_{N1} \\ \vdots & & & \\ a_{1N} & \dots & & a_{NN} \end{pmatrix}}_{A^T} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \underbrace{\begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & & \\ a_{N1} & \dots & a_{NN} \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

$$= (A^T + A) x .$$

7.1.b.

$$\frac{\partial x^T A x}{\partial A}$$

$$\frac{\partial \left(\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} x_i x_j \right)}{\partial a_{ek}} = x_e x_k$$

$$\Rightarrow \frac{\partial x^T A x}{\partial A} = \begin{pmatrix} x_1^2 & x_1 x_2 & \cdots & x_1 x_N \\ \vdots & & & \\ \vdots & & & \\ x_1 x_N & \cdots & \cdots & x_N^2 \end{pmatrix}$$

$$= \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} (x_1 \cdots x_N) = x x^T.$$

7.1.c.

$$\frac{\partial \det A}{\partial A} = \det(A) (A^{-1})^T$$

the determinant $\det(A) = \sum_{i=1}^{N} a_{ij} (-1)^{i+j} M_{ij}$ (Laplace's Formula)

M_{ij} is a minor. determinant of a matrix that is X without row i and row j

$$\frac{\det(A)}{d a_{ij}} = (-1)^{i+j} M_{ij} = \operatorname{adj}(A)_{ij}$$

$$\Rightarrow \frac{\det(A)}{dA} = \operatorname{adj}(A) \qquad A^{-1} = \frac{(\operatorname{adj}(A))^T}{\det(A)}$$

$$\frac{\det(A)}{dA} = \det(A) (A^{-1})^T$$

$A^{-1})^T$

$$\frac{\quad}{dA} = \ldots$$

7.1.d. $\dfrac{\partial \ln \det(A)}{\partial A} = \dfrac{\partial \ln \det(A)}{\partial \det(A)} \dfrac{\partial \det(A)}{\partial A} = \dfrac{1}{\det(A)} \left(\det(A) (A^{-1})^T \right) = (A$

7.1.e. $\dfrac{\partial \ln \det(A)}{\partial A^{-1}} = -\left(\dfrac{-\partial \ln(\det(A))}{\partial A^{-1}} \right) = -\left(\dfrac{\partial \ln(\det(A^{-1}))}{\partial A^{-1}} \right) =$

$= -(A)^T$

7.4

impulses counted during interval D.

N different cells and using identical stimuli

2 different type of cells $S_n = 1$ with p_1, $S_n = 2$ with p_2

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \vdots \\ x_N \end{pmatrix}$$

$$P(X_n = x_n | S_n = i) = \frac{e^{-c_i D} (c_i D)^{x_n}}{x_n!} \quad i \in \{1, 2\}.$$

c_i (counts/s) depends only on S_n

Use EM

$$\max_{\{c_1, c_2\}} \sum_{i=1}^{N} \log(P(X_n = x_n | c_1, c_2)) = \sum_{i=1}^{N} \log\left(p_1 \frac{e^{-c_1 D} (c_1 D)^{x_i}}{x_i!} + p_2 \frac{e^{-c_2 D} (c_2 D)^{x_i}}{x_i!}\right)$$

$A^{-1})^T.$

$\{c_1, c_2\}$

$$Q(c_1, c_2 \mid c_1^{(t)}, c_2^{(t)}) = \sum_{i=1}^{N} \Big(P(S_i = 1 \mid X_i = x_i, c_1^{(t)}, c_2^{(t)}) \log \Big(p_1 \cdot \frac{e^{-c_1 D} (c_1 D)^{x_i}}{x_i!} \Big)$$

$$+ P(S_i = 2 \mid X_i = x_i, c_1^{(t)}, c_2^{(t)}) \log \Big(p_2 \, \frac{e^{-c_2 D} (c_2 D)^{x_i}}{x_i} \Big)$$

$$P(S_n = i \mid X_n = x_n) P(X_n = x_n) = P(X_n = x_n \mid S_n = i) P(S_n = i)$$

$$P(S_n = i \mid X_n = x_n) = \frac{P(X_n = x_n \mid S_n = i)\, {*}P(S_n = i)}{P(X_n = x_n)} \qquad \gamma_{i,1}^{(t)} \triangleq P(S_i = 1 \mid X_i = x_i, c_1^{(t)}, c_2^{(t)})$$

$$Q(c_1, c_2 \mid c_1^{(t)}, c_2^{(t)}) = \sum_{i=1}^{N} \Big(\gamma_{i,1}^{(t)} \log \Big(p_1 \frac{e^{-c_1 D} (c_1 D)^{x_i}}{x_i!} \Big) + \gamma_{i,2}^{(t)} \log \Big(p_2 \frac{e^{-c_2 D} (c_2 D)^{x_i}}{x_i!} \Big)$$

$$= \sum_{i=1}^{N} \Big(\gamma_{i,1}^{(t)} \big(\cancel{-\log(x_i!)} + \cancel{\log(p_1)} - c_1 D + x_i \log(c_1 D) \big)$$

$$+ \gamma_{i,2}^{(t)} \big(\cancel{-\log(x_i!)} + \cancel{\log(p_2)} - c_2 D + x_i \log(c_2 D) \big) \Big)$$

$$Q(c_1, c_2 \mid c_1^{(t)}, c_2^{(t)}) = \sum_{i=1}^{N} \Big(\gamma_{i,1}^{(t)} \big(x_i \log(c_1 D) - c_1 D \big) + \gamma_{i,2}^{(t)} \big(x_i \log(c_2 D) - c_2 D \big) \Big)$$

$$c_1^{(t+1)} := \underset{c_1}{\arg\max} \sum_{i=1}^{N} \Big(\gamma_{i,1}^{(t)} \big(x_i \log(c_1 D) - c_1 D \big) \Big)$$

$$\Rightarrow \sum_{i=1}^{N} \gamma_{i,1}^{(t)} \Big(x_i \frac{D}{c_1 D} - D \Big) = 0 \qquad \sum_{i=1}^{N} \gamma_{i,1}^{(t)} x_i - c_1 \sum_{i=1}^{N} \gamma_{i,1}^{(t)} D = 0.$$

$$\Rightarrow \sum_{i=1} \gamma_{i,1}^{(t)}\left(x_i \frac{D}{c_1 D} - D\right) = 0$$

$$\boxed{c_1 = \frac{\sum_{i=1}^{N} \gamma_{i,1}^{(t)} x_i}{\sum_{i=1}^{N} \gamma_{i,1}^{(t)} D}}$$

similarly for $c_2 = \dfrac{\sum_{i=1}^{N} \gamma_{i,2}^{(t)} x_i}{\sum_{i=1}^{N} \gamma_{i,2}^{(t)} D}$

$$= \boxed{\frac{\sum_{i=1}^{N} (1-\gamma_{i,1}^{(t)}) x_i}{\sum_{i=1}^{N} (1-\gamma_{i,1}^{(t)}) D}}$$

$$P(S_n = i \mid X_n = x_n) = \frac{P(X_n = x_n \mid S_n = i)\, P(S_n = i)}{P(X_n = x_n)} =$$

$$= \boxed{\frac{p_i e^{-c_i D} (c_i D)^{x_n} / x_n!}{\frac{p_1 e^{-c_1 D} (c_1 D)^{x_n}}{x_n!} + \frac{p_2 e^{-c_2 D} (c_2 D)^{x_n}}{x_n!}}}$$

7.5. observable $\underline{X} = (X_1, \dots, X_t, \dots)$

$\left.\begin{array}{l} \underline{U} = (U_1, \dots, U_t, \dots) \\ \underline{V} = (V_1, \dots, V_t, \dots) \end{array}\right\}$ non observable.

$U_0 = 0$

$$U_t = \begin{cases} +1 & U_{t-1} + V_t \geq 0 \\ -1 & U_{t-1} + V_t < 0 \end{cases}$$

i)

start at T

$\underline{V} = (V_1, \dots, V_t, \dots)$

$\underline{W} = (W_1, \dots, W_t, \dots)$

recall: $P(a, b, c) = P(a \mid b, c)\, P(b \mid c)\, P(c)$.

$$P(\underline{U}, \underline{X} \mid \lambda, c) = P(U_1, \dots, U_T, X_1, \dots, X_T \mid \lambda, c) =$$

$$= P(U_T, X_T \mid U_{T-1}, \dots, U_1, X_{T-1}, \dots, X_1\ \lambda, c)\, P(U_{T-1}, X_{T-1} \mid U_{T-2}, \dots$$

$$\dots$$

$$P(U_i, X_i \mid U_{i-1}, \dots, U_1, X_{i-1}, \dots, X_1, \lambda, c) =$$

$$= P(X_i \mid U_i, U_{i-1}, \dots, U_1, X_{i-1}, \dots, X_1, \lambda, c)\, P(U_i \mid U_{i-1}, \dots$$

$$= P(X_i \mid U_i, c)\, P(U_i \mid U_{i-1})$$

$$\Rightarrow P(\underline{U}, \underline{X} \mid \lambda, c) = \left(\prod_{t=2}^{T} \underbrace{P(X_i = x_i \mid U_i = u_i^c)}_{\mathcal{N}(cU_i, \sigma_W^2)} \cdot P(U_i \mid U_{i-1}) \right)$$

$)$

$, U_1, x_{i-1}, \dots, x_1, \lambda, c) =$

$) \, P(X_1 = x_1 \mid U_1 = u_1^c) \, P(U_1 = u_1 \mid U_0 = 0)$

do not depend on c

$$Q(c|c^{(t)}) = \sum_{\underline{u}} P(\underline{U}=\underline{u} \,|\, \underline{X}=\underline{x}, c^{(t)}) \left(\sum_{t=1}^{T} \log \left(P(X_t = x_t \,|\, U_t = \underline{u}_t \ldots \right.\right.$$

$\underline{u}$ → all combinations.

all the terms that do not appear in the summa…

$$Q'(c|c^{(t)}) = \sum_{t=1}^{T} P(U_t = 1 \,|\, \underline{X} = x, c^{(t)}) \log \left(P(X_t = x_t \,|\, U_t = 1 \ldots \right.$$

$$Q'(c|c^{(t)}) = \sum_{t=1}^{T} \gamma_{1,t} \log \left(\frac{1}{\sqrt{2\pi}\,\sigma_W} e^{-(x_t - c)^2 / 2\sigma_W^2} \right) + \gamma_{2,t} \log \ldots$$

Fix log — and ignore parts that do not depend o…

$$-\sum_{t=1}^{T} \gamma_{1,t} (x_t - c)^2 - \sum_{t=1}^{T} \gamma_{2,t} (x_t + c)^2$$

$$, c)\Big) + \sum_{t=2}^{T} \log\Big(P(U_t = u_t \mid U_{t-1} = u_{t-1})\Big) + \log\Big(P(U_1 = u_1 \mid U_0 = 0)\Big)\Big)$$

...d are marginalized:

$$, c)\Big) + P(U_t = -1 \mid \underline{X} = x, c^{(t)}) \log\Big(P(X_t = x_t \mid U_t = -1, c)\Big)$$

$$\log\left(\frac{1}{\sqrt{2\pi}\,\sigma_W} e^{-(x_t + c)^2 / 2\sigma_W^2} \right)$$

...on c:

T

$t=1$ $\qquad$ $t=1$

concave ✓

take derivative and set to $0 \rightarrow$ $-\sum_{t=1}^{T} \gamma_{1,t}(x_t - c) +$

and therefore: $\sum_{t=1}^{T} \underbrace{(\gamma_{1,t} + \gamma_{2,t})}_{\text{each of them add to } 1} c_1 = \sum_{t=1}^{T} (\gamma_{1,t} - \gamma_{2,t}) x_t$

$$\boxed{c_1 = \frac{1}{T} \sum_{t=1}^{T} (}$$

7.7 HMM with GMM:s

$$\gamma_{im,t} = P[S_t = i \wedge U_t = m \mid x, \lambda] = \gamma_{i,t} \cdot \frac{w_{im}\, g(x_t, \mu_i}{\sum_{k=1}^{M} w_{ik}\, g(x_t,}$$

$$\gamma_{i,t} = P(S_t = i \mid x, \lambda)$$

$$- \sum_{t=1}^{T} \gamma_{2,t} (x_t + c) = 0 .$$

$$\boxed{(\gamma_{1,t} - \gamma_{2,t}) x_t}$$

$$\frac{m , C_{im})}{n_{ik}, C_{ik})}$$

$$\gamma_{i,t} = P(S_t = i \mid x, \lambda)$$

$$P(S_t = i, U_t = m \mid x, \lambda) = P(S_t = i \mid U_t = m, x, \lambda)$$

$$= P(U_t = m \mid S_t = i, x, \lambda$$

$$f(x \mid U_t = m, S_t = i, \lambda) = g(x, \mu_{im}, C_{im})$$

$$P(U_t = m \mid S_t = i, x)\, f(\cancel{x} \mid S_t = i) = f(x \mid U_t = m$$

$$P(U_t = m \mid S_t = i, x) = \frac{g(x_t, \mu_{im}, C}{\sum_{k=1}^{M} w_{ik}\, g(x_t, \mu}$$

$$\Rightarrow \gamma_{im,t} = \gamma_{i,t}.$$

$P(U_t = m \mid x, \underline{\lambda})$

$) \, \underbrace{P(s_t = i \mid x, \lambda)}$

$, S_t = i) \; P(U_t = m \mid S_t = i)$

$im) \, w_{im}$

$ik, a_k)$