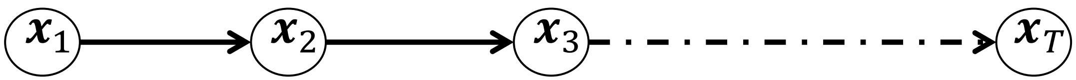
Hidden Markov Models (HMM)

EM in Hidden Markov Models

Markov Chains with Discrete Random Variables



Let's assume we have discrete random variables (e.g., taking 3 discrete values $\boldsymbol{x}_t = \{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\}$)

Markov Property: $p(\boldsymbol{x}_t|\boldsymbol{x}_1, .., \boldsymbol{x}_{t-1}) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$

e.g. $p(\boldsymbol{x}_t = \begin{bmatrix}1\\0\\0\end{bmatrix} | \boldsymbol{x}_{t-1} = \begin{bmatrix}0\\1\\0\end{bmatrix})$

Stationary, Homogeneous or Time-Invariant if the distribution $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ does not depend on t

Markov Chains with Discrete Random Variables

$$p(\boldsymbol{x}_1, .., \boldsymbol{x}_T) = p(\boldsymbol{x}_1) \prod_{t=2}^{\mathrm{T}} p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1})$$

What do we need in order to describe the whole procedure?

(1) A probability for the first frame/timestamp etc $p(\boldsymbol{x}_1)$. In order to define the probability we need to define the vector $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_{\mathrm{K}})$

$$p(\boldsymbol{x}_1 | \boldsymbol{\pi}) = \prod_{c=1}^{K} \pi_c{}^{x_{1c}}$$

(2) A transition probability $p(\boldsymbol{x_t} | \boldsymbol{x_{t-1}})$. In order to define it we need a KxK transition matrix $\boldsymbol{A} = [a_{ij}]$

$$p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1}, \boldsymbol{A}) = \prod_{j=1}^{K} \prod_{k=1}^{K} a_{jk}{}^{x_{t-1j} x_{tk}}$$

Hidden Markov Models

A casino has two dice (our latent variable is probably the dice ☺)

$$\boldsymbol{z} = \{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$$

One is fair and one is not

Each die has 6 sides.

$$\boldsymbol{x} = \{\begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}\}$$

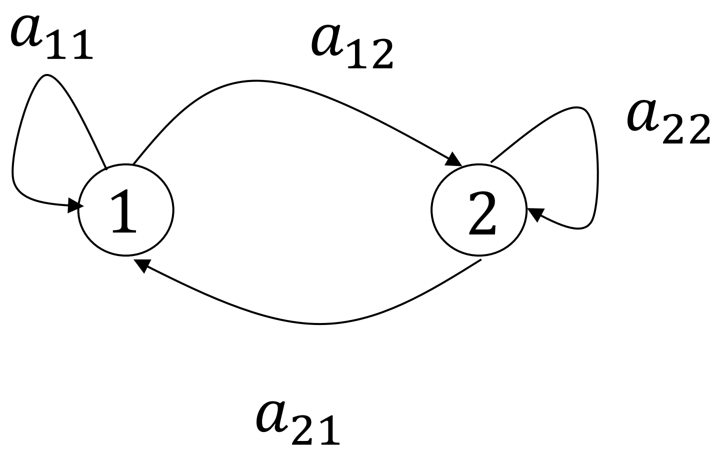
(1) Fair die $p\left(x_j = 1 | z_1 = 1\right) = \frac{1}{6}$ for all $j = \{1, .., 6\}$

(2) Loaded die $p\left(x_j = 1 | z_2 = 1\right) = b_{j2}$ for $j = \{1, .., 5\}$

and $p(x_6 = 6 | z_2 = 1) = 1 - \sum_{j=1}^{5} b_{j2}$ Emission probability

Hidden Markov Models

A casino player switches back & forth between fair and loaded die



$$p(z_{11}=1)=\pi_1$$

$$p(z_{12}=1)=\pi_2$$

$$p(\boldsymbol{z}_1|\boldsymbol{\pi})=\prod_{c=1}^{2}\pi_c^{\,z_{1c}}$$

$$p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1},\boldsymbol{A})=\prod_{j=1}^{2}\prod_{k=1}^{2}a_{jk}^{\,z_{t-1j}z_{tk}}$$

$$p(\boldsymbol{x}_t|\boldsymbol{z}_t)=\prod_{j=1}^{6}\prod_{k=1}^{2}b_{kj}^{\,x_{tj}z_{tk}}$$

EM in HMMs

Given a set of strings of observations (or even one) and the above model:

664153216162115234653214356634261655234232315142464156663246

Find the parameters $\boldsymbol{A}$, $\boldsymbol{\pi}$, b_{j2} by maximizing the probability

$$p(D_1, \ldots, D_N, Z_1, \ldots, Z_N | \theta)$$

$$= \prod_{l=1}^{N} p\left(\boldsymbol{x}_1{}^l, \boldsymbol{x}_2{}^l, \cdots, \boldsymbol{x}_T{}^l, \boldsymbol{z}_1{}^l, \boldsymbol{z}_2{}^l, \cdots, \boldsymbol{z}_T{}^l | \theta\right)$$

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} p(\boldsymbol{x}_t{}^l | \boldsymbol{z}_t{}^l)\, p(\boldsymbol{z}_1{}^l) \prod_{t=2}^{\mathrm{T}} p(\boldsymbol{z}_t{}^l | \boldsymbol{z}_{t-1}{}^l)$$

EM in HMMs

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} p(\boldsymbol{x}_t{}^l | \boldsymbol{z}_t{}^l)\, p(\boldsymbol{z}_1{}^l) \prod_{t=2}^{\mathrm{T}} p(\boldsymbol{z}_t{}^l | \boldsymbol{z}_{t-1}{}^l)$$

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} \prod_{j=1}^{6} \prod_{k=1}^{2} b_{jk}{}^{x_{tj}{}^l z_{tk}{}^l} \prod_{k=1}^{2} \pi_k{}^{z_{1k}{}^l} \prod_{t=2}^{\mathrm{T}} \prod_{j=1}^{2} \prod_{k=1}^{2} a_{jk}{}^{z_{t-1j}{}^l z_{tk}{}^l}$$

$$\underset{ln}{\Rightarrow} = \sum_{l=1}^{N} \sum_{t=1}^{T} \sum_{j=2}^{6} \sum_{k=1}^{2} x_{tj}{}^l z_{tk}{}^l \ln b_{jk} + \sum_{l=1}^{N} \sum_{k=1}^{2} z_k{}^l \ln \pi_k$$

$$+ \sum_{l=1}^{N} \sum_{t=2}^{T} \sum_{j=1}^{2} \sum_{k=1}^{2} z_{t-1j}{}^l z_{tk}{}^l \ln a_{jk}$$

EM in HMMs

Taking the expectations with regards to the posterior

$$= \sum_{l=1}^{N} \sum_{t=1}^{T} \sum_{j=2}^{6} \sum_{k=1}^{2} x_{tj}{}^{l} E[z_{tk}{}^{l}] \ln b_{jk}$$

$$+ \sum_{l=1}^{N} \sum_{k=1}^{2} E[z_{1k}{}^{l}] \ln \pi_k$$

$$+ \sum_{l=1}^{N} \sum_{t=2}^{T} \sum_{j=1}^{2} \sum_{k=1}^{2} E[z_{t-1j}{}^{l} z_{tk}{}^{l}] \ln a_{jk}$$

EM in HMMs

$$E[{z_{1k}}^l] = \sum_{{z_{1k}}^l} {z_{1k}}^l \, p\left({z_{1k}}^l \middle| {\boldsymbol{x}_1}^l, \cdots, {\boldsymbol{x}_T}^l\right) = p\left({z_{1k}}^l = 1 \middle| {\boldsymbol{x}_1}^l, \cdots, {\boldsymbol{x}_T}^l\right)$$

$$E[{z_{tk}}^l] = \sum_{{z_{tk}}^l} {z_{tk}}^l \, p\left({z_{tk}}^l \middle| {\boldsymbol{x}_1}^l, \cdots, {\boldsymbol{x}_T}^l\right) = p\left({z_{tk}}^l = 1 \middle| {\boldsymbol{x}_1}^l, \cdots, {\boldsymbol{x}_T}^l\right)$$

$$E\left[{z_{t-1j}}^l {z_{tk}}^l\right] = \sum_{{z_{t-1j}}^l} \sum_{{z_{tk}}^l} {z_{t-1j}}^l {z_{tk}}^l \, p\left({z_{t-1j}}^l {z_{tk}}^l \middle| {\boldsymbol{x}_1}^l, {\boldsymbol{x}_2}^l, \cdots, {\boldsymbol{x}_T}^l\right)$$

$$= p\left({z_{t-1j}}^l = 1, {z_{tk}}^l = 1 \middle| {\boldsymbol{x}_1}^l, {\boldsymbol{x}_2}^l, \cdots, {\boldsymbol{x}_T}^l\right)$$

EM in HMMs

$$p\left(z_{1k}{}^{l} = 1 \middle| \boldsymbol{x}_1{}^{l}, \cdots, x_T{}^{l}\right)$$
$$= \frac{p\left(\boldsymbol{x}_1{}^{l}, z_{1k}{}^{l} = 1\right) p(\boldsymbol{x}_2{}^{l}, \cdots, \boldsymbol{x}_T{}^{l} | z_{1k}{}^{l} = 1)}{p(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l})} = \frac{\alpha\left(z_{1k}{}^{l}\right)\beta(z_{1k}{}^{l})}{p(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l})}$$

$$p\left(z_{tk}{}^{l} = 1 \middle| \boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l}\right)$$
$$= \frac{p\left(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_t{}^{l}, z_{tk}{}^{l} = 1\right) p(\boldsymbol{x}_{t+1}{}^{l}, \cdots, \boldsymbol{x}_T{}^{l} | z_{tk}{}^{l} = 1)}{p(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l})} = \frac{\alpha\left(z_{tk}{}^{l}\right)\beta(z_{tk}{}^{l})}{p(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l})}$$

$$p\left(z_{t-1j}{}^{l} = 1, z_{tk}{}^{l} = 1 \middle| \boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l}\right)$$
$$= \frac{\alpha\left(z_{t-1j}{}^{l}\right) \prod_{r=1}^{6} b_{kr}{}^{x_{tr}{}^{l}} a_{jk} \beta(z_{tk}{}^{l})}{p(\boldsymbol{x}_1{}^{l}, \cdots, \boldsymbol{x}_T{}^{l})}$$

EM in HMMs

$$L(b_{j2}) = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{j=2}^{6}\sum_{k=1}^{2} x_{tj}{}^{l} E[z_{tk}{}^{l}] \ln b_{jk} + \sum_{l=1}^{N}\sum_{k=1}^{2} E[z_{1k}{}^{l}] \ln \pi_k$$

$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{2}\sum_{k=1}^{2} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right] \ln a_{jk} + \lambda\left(\sum_{j=1}^{6} b_{j2} - 1\right)$$

$$b_{j2} = \frac{\sum_{l=1}^{N}\sum_{t=1}^{T} E[z_{t2}{}^{l}]\, x_{tj}{}^{l}}{\sum_{l=1}^{N}\sum_{t=1}^{T} E[z_{t2}{}^{l}]}$$

EM in HMMs

$$L(\pi_k) = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{j=2}^{6}\sum_{k=1}^{2} x_{tj}{}^{l} E[z_{tk}{}^{l}] \ln b_{jk} + \sum_{l=1}^{N}\sum_{k=1}^{2} E[z_{1k}{}^{l}] \ln \pi_k$$
$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{2}\sum_{k=1}^{2} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right] \ln a_{jk} + \lambda\left(\sum_{k=1}^{2} \pi_k - 1\right)$$

$$\pi_k = \frac{\sum_{l=1}^{N} E[z_{1k}{}^{l}]}{\sum_{l=1}^{N}\sum_{r=1}^{2} E[z_{1r}{}^{l}]}$$

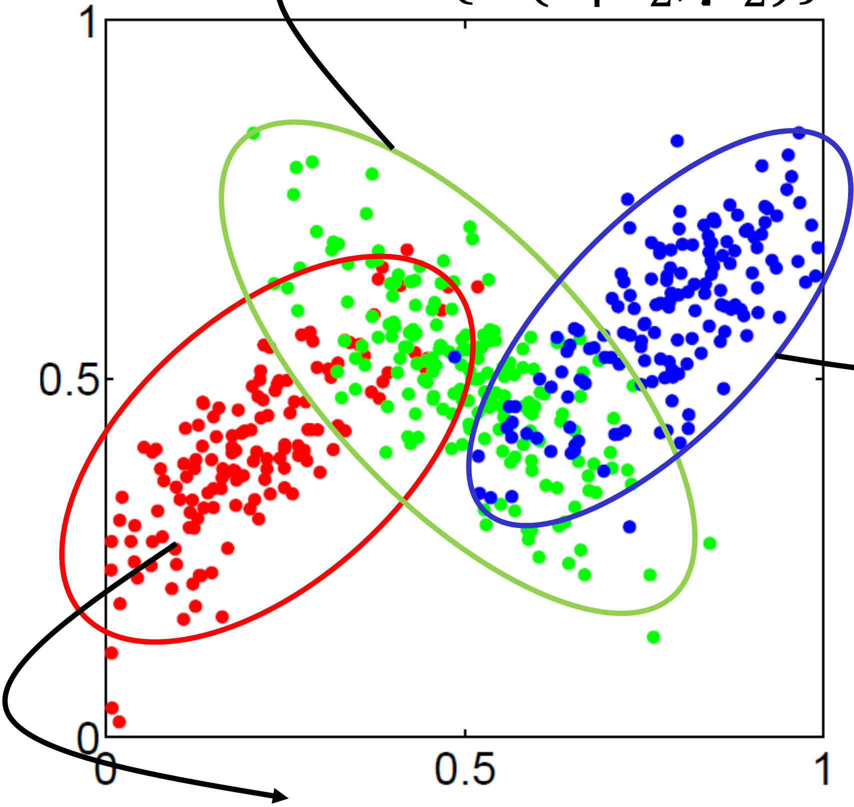
EM in HMMs

$$L(a_{jk}) = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{j=2}^{6}\sum_{k=1}^{2} x_{tj}{}^{l} E[z_{tk}{}^{l}] \ln b_{jk} + \sum_{l=1}^{N}\sum_{k=1}^{2} E[z_{1k}{}^{l}] \ln \pi_k$$
$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{2}\sum_{k=1}^{2} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right] \ln a_{jk} + \lambda\left(\sum_{k=1}^{2} a_{jk} - 1\right)$$

$$a_{jk} = \frac{\sum_{l=1}^{N}\sum_{t=2}^{T} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right]}{\sum_{r=1}^{2}\sum_{l=1}^{N}\sum_{t=2}^{T} E\left[z_{t-1j}{}^{l} z_{tr}{}^{l}\right]}$$

Latent Variables

$\{N(\boldsymbol{x}|\boldsymbol{\Sigma}_2, \boldsymbol{\mu}_2)\}$



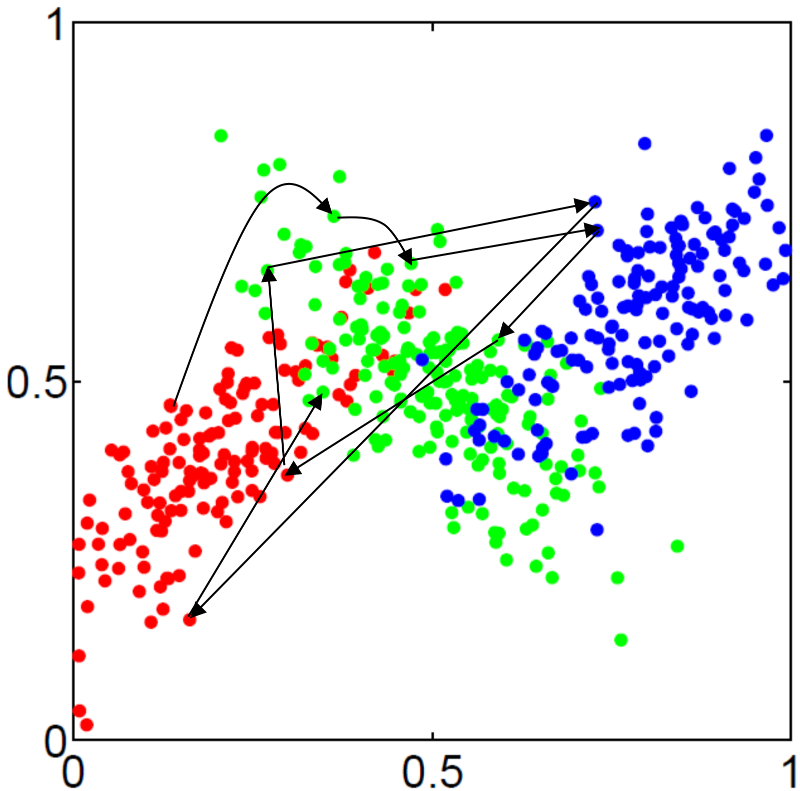
$\{\mathrm{N}(\boldsymbol{x}|\boldsymbol{\Sigma}_3, \boldsymbol{\mu}_3)\}$

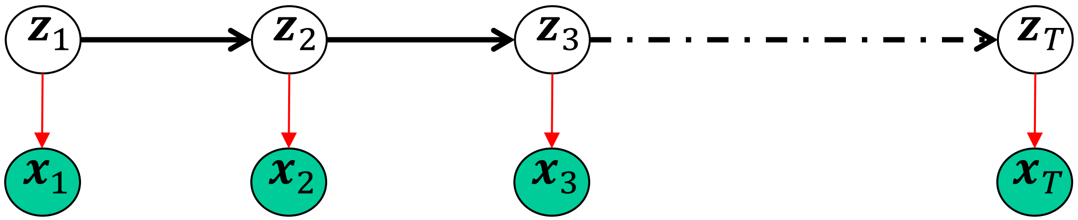
$\{\mathrm{N}(\boldsymbol{x}|\boldsymbol{\Sigma}_1, \boldsymbol{\mu}_1)\}$

$$\boldsymbol{z}_t = \begin{bmatrix} z_{t1} \\ z_{t2} \\ z_{t3} \end{bmatrix} \in \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$p(\boldsymbol{X}, \boldsymbol{Z}|\theta)$$
$$= p\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_T, \boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_T|\theta\right)$$
$$= \prod_{t=1}^{T} p(\boldsymbol{x}_t|\boldsymbol{z}_t, \theta_x) \prod_{t=1}^{T} p(\boldsymbol{z}_t|\theta_z)$$

Latent Variables in a Markov Chain





$$p(\mathbf{z}_1, .., \mathbf{z}_T) = p(\mathbf{z}_1) \prod_{t=2}^{\mathrm{T}} p(\mathbf{z}_t | \mathbf{z}_{t-1})$$

$$p(\boldsymbol{X}, \boldsymbol{Z} | \theta)$$
$$= p\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_T, \mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_T | \theta\right)$$
$$= \prod_{t=1}^{T} p(\boldsymbol{x}_t | \mathbf{z}_t, \theta_x)\, p(\mathbf{z}_1) \prod_{t=2}^{\mathrm{T}} p(\mathbf{z}_t | \mathbf{z}_{t-1})$$

EM HMM with real valued observations

Given a number of sets of observations (or even sequence)

Find the parameters $\boldsymbol{A}, \boldsymbol{\pi}, \boldsymbol{\Sigma}_j, \boldsymbol{\mu}_j$ that maximize the probability

$$p(D_1, \dots, D_N, Z_1, \dots, Z_N | \theta)$$

$$= \prod_{l=1}^{N} p\left(\boldsymbol{x}_1{}^l, \boldsymbol{x}_2{}^l, \cdots, x_T{}^l, z_1{}^l, z_2{}^l, \cdots, z_T{}^l | \theta\right)$$

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} p(\boldsymbol{x}_t{}^l | \boldsymbol{z}_t{}^l)\, p(\boldsymbol{z}_1{}^l) \prod_{t=2}^{\mathrm{T}} p(\boldsymbol{z}_t{}^l | \boldsymbol{z}_{t-1}{}^l)$$

EM HMM with real valued observations

What is different and what is the same?

Same: $$p(\boldsymbol{z}_1|\boldsymbol{\pi}) = \prod_{k=1}^{3} {\pi_k}^{z_{1k}}$$

$$p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1}, \boldsymbol{A}) = \prod_{j=1}^{3}\prod_{k=1}^{3} {a_{jk}}^{z_{t-1j}z_{tk}}$$

Different: $$p(\boldsymbol{x}_t|\boldsymbol{z}_t) = \prod_{k=1}^{3} N(\boldsymbol{x}_t|\boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa)^{z_{tk}}$$

EM HMM with real valued observations

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} p(\boldsymbol{x}_t{}^l | \boldsymbol{z}_t{}^l)\, p(\boldsymbol{z}_1{}^l) \prod_{t=2}^{\mathrm{T}} p(\boldsymbol{z}_t{}^l | \boldsymbol{z}_{t-1}{}^l)$$

$$= \prod_{l=1}^{N} \prod_{t=1}^{T} \prod_{k=1}^{3} N(\boldsymbol{x}_t | \boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa)^{z_{tk}{}^l} \prod_{k=1}^{3} \pi_k{}^{z_{1k}{}^l} \prod_{t=2}^{\mathrm{T}} \prod_{j=1}^{3} \prod_{k=1}^{3} a_{jk}{}^{z_{t-1j}{}^l z_{tk}{}^l}$$

$$= \sum_{l=1}^{N} \sum_{t=1}^{T} \sum_{k=2}^{3} z_{tk}{}^l \ln N(\boldsymbol{x}_t | \boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa) + \sum_{l=1}^{N} \sum_{k=1}^{3} z_{1k}{}^l \ln \pi_k$$

$$+ \sum_{l=1}^{N} \sum_{t=2}^{T} \sum_{j=1}^{3} \sum_{k=1}^{3} z_{t-1j}{}^l z_{tk}{}^l \ln a_{jk}$$

EM HMM with real valued observations

Taking the expectations with regards to the posterior

$$L = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{k=1}^{3} E[{z_{tk}}^{l}] \ln N(\boldsymbol{x}_t | \boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa)$$

$$+ \sum_{l=1}^{N}\sum_{k=1}^{3} E[{z_{1k}}^{l}] \ln \pi_k$$

$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{3}\sum_{k=1}^{3} E[{z_{t-1j}}^{l}{z_{tk}}^{l}] \ln a_{jk}$$

EM HMM with real valued observations

$$E[z_{1k}{}^l] = \sum_{z_{1k}{}^l} z_{1k}{}^l \, p(z_{1k}{}^l | \boldsymbol{x}_1{}^l, \cdots, \boldsymbol{x}_T{}^l) = p(z_{1k}{}^l = 1 | \boldsymbol{x}_1{}^l, \cdots, \boldsymbol{x}_T{}^l)$$

$$E[z_{tk}{}^l] = \sum_{z_{tk}{}^l} z_{tk}{}^l \, p(z_{tk}{}^l | \boldsymbol{x}_1{}^l, \cdots, \boldsymbol{x}_T{}^l) = p(z_{tk}{}^l = 1 | \boldsymbol{x}_1{}^l, \cdots, \boldsymbol{x}_T{}^l)$$

$$E\left[z_{t-1j}{}^l z_{tk}{}^l\right] = \sum_{z_{t-1j}{}^l} \sum_{z_{tk}{}^l} z_{t-1j}{}^l z_{tk}{}^l \, p\left(z_{t-1j}{}^l z_{tk}{}^l | \boldsymbol{x}_1{}^l, \boldsymbol{x}_2{}^l, \cdots, \boldsymbol{x}_T{}^l\right)$$
$$= p\left(z_{t-1j}{}^l = 1, z_{tk}{}^l = 1 | \boldsymbol{x}_1{}^l, \boldsymbol{x}_2{}^l, \cdots, \boldsymbol{x}_T{}^l\right)$$

EM HMM with real valued observations

$$p\left(z_{1k}{}^{l}=1\middle|\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l}\right)$$
$$=\frac{p\left(\boldsymbol{x}_1{}^{l},z_{1k}{}^{l}=1\right)p(\boldsymbol{x}_2{}^{l},\cdots,\boldsymbol{x}_T{}^{l}|z_{1k}{}^{l}=1)}{p(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l})}=\frac{\alpha\left(z_{1k}{}^{l}\right)\beta(z_{1k}{}^{l})}{p(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l})}$$

$$p\left(z_{tk}{}^{l}=1\middle|\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l}\right)$$
$$=\frac{p\left(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_t{}^{l},z_{tk}{}^{l}=1\right)p(\boldsymbol{x}_{t+1}{}^{l},\cdots,\boldsymbol{x}_T{}^{l}|z_{tk}{}^{l}=1)}{p(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l})}=\frac{\alpha\left(z_{tk}{}^{l}\right)\beta(z_{tk}{}^{l})}{p(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l})}$$

$$p\left(z_{t-1j}{}^{l}=1,z_{tk}{}^{l}=1\middle|\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l}\right)$$
$$=\frac{\alpha\left(z_{t-1j}{}^{l}\right)N(\boldsymbol{x}_t|\boldsymbol{\mu}_\kappa,\boldsymbol{\Sigma}_\kappa)a_{jk}\beta(z_{tk}{}^{l})}{p(\boldsymbol{x}_1{}^{l},\cdots,\boldsymbol{x}_T{}^{l})}$$

EM in HMMs

$$\frac{dL(\boldsymbol{\mu}_k)}{d\boldsymbol{\mu}_k} = 0 \Rightarrow \quad \boldsymbol{\mu}_k = \frac{\sum_{l=1}^{N}\sum_{t=1}^{T} E[{z_{tk}}^l]\, {\boldsymbol{x}_t}^l}{\sum_{l=1}^{N}\sum_{t=1}^{T} E[{z_{tk}}^l]}$$

$$\frac{dL(\boldsymbol{\Sigma}_k)}{d\boldsymbol{\Sigma}_k} = 0 \Rightarrow \boldsymbol{\Sigma}_k = \frac{\sum_{l=1}^{N}\sum_{t=1}^{T} E[{z_{tk}}^l]\, ({\boldsymbol{x}_t}^l - \boldsymbol{\mu}_k)({\boldsymbol{x}_t}^l - \boldsymbol{\mu}_k)^T}{\sum_{l=1}^{N}\sum_{t=1}^{T} E[{z_{tk}}^l]}$$

EM in HMMs

$$L(\pi_k) = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{k=1}^{3} E[{z_{tk}}^{l}] \ln N(\boldsymbol{x}_t|\boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa) + \sum_{l=1}^{N}\sum_{k=1}^{3} E[{z_{1k}}^{l}] \ln \pi_k$$

$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{3}\sum_{k=1}^{3} E\left[{z_{t-1j}}^{l}{z_{tk}}^{l}\right] \ln a_{jk} + \lambda\left(\sum_{k=1}^{3} \pi_k - 1\right)$$

$$\pi_k = \frac{\sum_{l=1}^{N} E[{z_{1k}}^{l}]}{\sum_{l=1}^{N}\sum_{r=1}^{2} E[{z_{1r}}^{l}]}$$

EM in HMMs

$$L(a_{jk}) = \sum_{l=1}^{N}\sum_{t=1}^{T}\sum_{k=1}^{3} E[z_{tk}{}^{l}] \ln N(\boldsymbol{x}_t|\boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa) + \sum_{l=1}^{N}\sum_{k=1}^{3} E[z_{1k}{}^{l}] \ln \pi_k$$

$$+ \sum_{l=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{3}\sum_{k=1}^{3} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right] \ln a_{jk} + \lambda\left(\sum_{k=1}^{3} a_{jk} - 1\right)$$

$$a_{jk} = \frac{\sum_{l=1}^{N}\sum_{t=2}^{T} E\left[z_{t-1j}{}^{l} z_{tk}{}^{l}\right]}{\sum_{r=1}^{3}\sum_{l=1}^{N}\sum_{t=2}^{T} E\left[z_{t-1j}{}^{l} z_{tr}{}^{l}\right]}$$