MA8702 Advanced Computer-intensive Statistics

Department of Mathematical Sciences,
NTNU, Trondheim, Norway.

Hidden Markov models & Sequential Monte Carlo methods
Part I: Categorical HMM

Henning Omre, 01.04.2021

HIDDEN MARKOV MODELS / Sequential Monte Carlo

→ Focus: temporal state n-vector

$\underset{n\times 1}{s_t}$; $t \in T_s : \{0, 1, \dots, t, \dots, T, T+1\}$ — categorial / continuous

└ current time

Define:

$s = [s_0, \dots, s_t, \dots, s_T, s_{T+1}] \to s_{i:j} = [s_i, \dots, s_j]$ $i \le j$

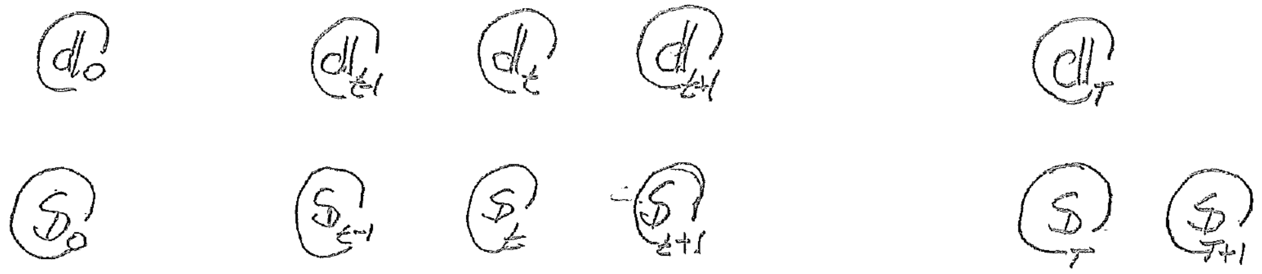
→ Available: temporal observation m-vector

d_t ; $t \in T_d : \{0, 1, \dots, t, \dots, T\}$

└ current time

$d : [d_0, \dots, d_t, \dots, d_T] \to d_{i:j} = [d_i, \dots, d_j]$ $i \le j$

→ Assess s from $d \to [s|d]$.



→ Bayesian framework:

$$p(\mathbf{s}|\mathbf{d};\theta_{\ell p}) = \left[p(\mathbf{d};\theta_{\ell p})\right]^{-1} p(\mathbf{d}|\mathbf{s};\theta_{\ell})\, p(\mathbf{s};\theta_p)$$

posterior   likelihood   prior

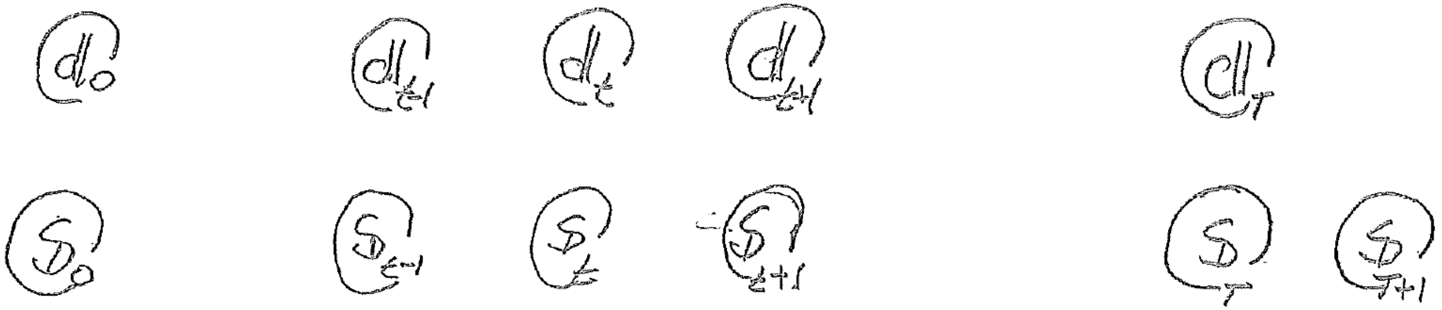
$$\int p(\mathbf{d}|\mathbf{s};\theta_\ell)\, p(\mathbf{s};\theta_p)\, d\mathbf{s}$$

→ Distinguish:

Smoothing $p(\mathbf{s}_{0:T}|\mathbf{d}_{0:T})$ – $p(\mathbf{s}_0|\mathbf{d}_{0:T})$ inversion

$p(\mathbf{s}_t|\mathbf{d}_{0:T})$ marg smoothing

Filtering $p(\mathbf{s}_T|\mathbf{d}_{0:T})$ – $p(\mathbf{s}_{T+1}|\mathbf{d}_{0:T})$ forecasting



Cappe et al [2005]

Standard Model Assumptions

Likelihood model:

$$[\mathbf{d}|\mathbf{s}] \Rightarrow p(\mathbf{d}|\mathbf{s}) = \prod_t p(d_t|\mathbf{s}) = \prod_t p(d_t|s_t)$$

cond. indep

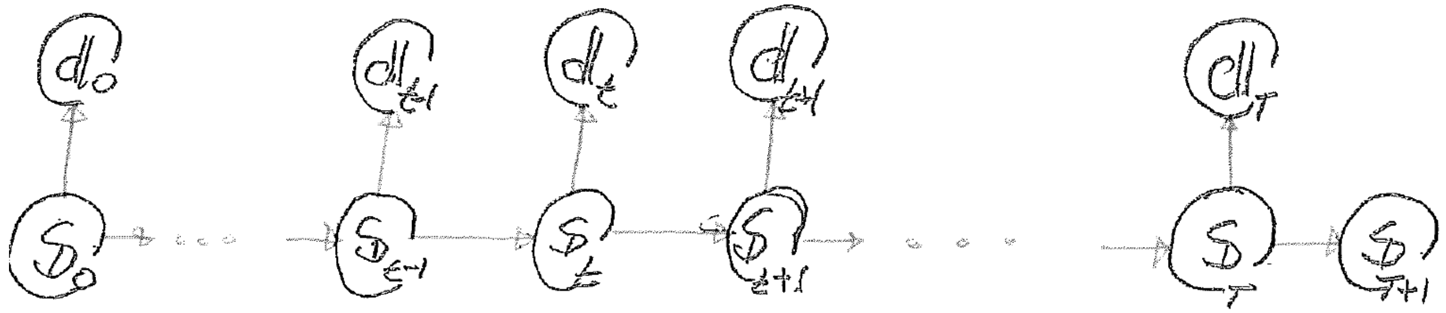
single site response

Prior model:

$$\mathbf{s} \Rightarrow p(\mathbf{s}) = p(s_0)\prod_t p(s_t|s_{t-1:0})$$

$$= p(s_0)\prod_t p(s_t|s_{t-1})$$

Markov assump



Posterior model:

Moja et al [2019]

const

$$[s|d] \Rightarrow p(s|d) = [p(d)]^{-1} p(d|s)\, p(s)$$

$$= [p(d)]^{-1} p(d_0|s_0)\, p(s_0) \prod_t p(d_t|s_t)\, p(s_t|s_{t-1})$$

$$= p(s_0|d_{0:T}) \prod_t p(s_t|s_{t-1}, d_{t:T})$$

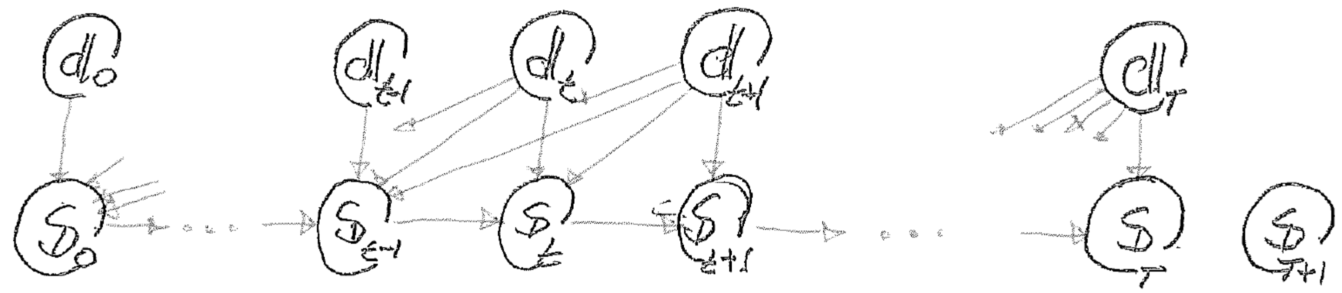
non-stat Markov mod

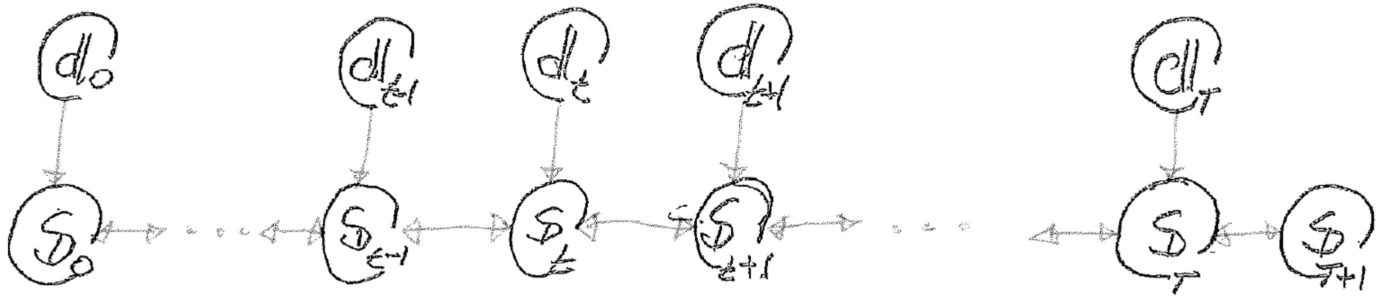
NOTE:

$$[s_t|s_{-t}, d] \Rightarrow p(s_t|s_{-t}, d)$$

$$= p(s_t|s_{t+1}, s_{t-1}, d_t)$$

Markov mod





$\longrightarrow$ Demonstration - Posterior Markov model

- Always correct:

$$p(s|d) = p(s_0|d) \prod_t p(s_t|s_{t-1:0}, d)$$

- Consider:

$$p(s_t|s_{t-1:0}, d) = \frac{p(s_{t:0}, d)}{p(s_{t-1:0}, d)}$$

$$= \frac{\int_{t+1} \cdots \int_T p(d|s)p(s)\, ds_T \cdots ds_{t+1}}{\int_t \int_{t+1} \cdots \int_T p(d|s)p(s)\, ds_T \cdots ds_{t+1}\, ds_t}$$

$$= \frac{p(d_0|s_0)p(s_0) \times \prod_{u=1}^{t-1} p(d_u|s_u)p(s_u|s_{u-1})}{\bullet} \longrightarrow 1$$

$$+ \frac{p(d_t|s_t)p(s_t|s_{t-1})}{\int_t \bullet}$$

$$+ \frac{\int_{t+1} \cdots \int_T \prod_{u=t+1}^{T} p(d_u|s_u)p(s_u|s_{u-1})\, ds_T \cdots ds_{t+1}}{\int_{t+1} \cdots \int_T \bullet} \rightarrow h_t(s_t, d_{t+1:T})$$

$$ds_T \rightarrow h_{t-1}(s_{t-1}, d_{t:T})$$

$$= \frac{h_t(s_t; d_{t+1:T})}{h_{t-1}(s_{t-1}, d_{t:T})} \cdot p(d_t|s_t)\, p(s_t|s_{t-1})$$

$$= p(s_t|s_{t-1}, d_{t:T})$$

→ Assessing the Posterior model:

Consider:

$$p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{s}_{t+1}, \mathbf{d}) = \overbrace{\left[p(\mathbf{s}_{t+1} \mid \mathbf{s}_{t-1}, \mathbf{d})\right]^{-1}}^{\text{const}} p(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{d})\, p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{d})$$

$$p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{s}_{t+1}, \mathbf{d}_t) = \text{const} \times p(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{d}_{t+1:T})\, p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{d}_{t:T})$$

$$p(\mathbf{d}_t \mid \mathbf{s}_t)\, p(\mathbf{s}_t \mid \mathbf{s}_{t-1})\, p(\mathbf{s}_{t+1} \mid \mathbf{s}_t) = \text{const}_+\, p(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{d}_{t+1:T})\, p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{d}_{t:T})$$

$$\downarrow$$

$$p(\mathbf{s}_t \mid \mathbf{s}_{t-1}, \mathbf{d}_{t:T}) = \text{const}_+ \times \frac{p(\mathbf{d}_t \mid \mathbf{s}_t)\, p(\mathbf{s}_t \mid \mathbf{s}_{t-1})\, p(\mathbf{s}_{t+1} \mid \mathbf{s}_t)}{p(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{d}_{t+1:T})}$$

Reverse recursion!

NOTE:

$$p(\mathbf{s}_T \mid \mathbf{s}_{T-1}, \mathbf{d}) = p(\mathbf{s}_T \mid \mathbf{s}_{T-1}, \mathbf{d}_T) = \left[p(\mathbf{d}_T)\right]^{-1} p(\mathbf{d}_T \mid \mathbf{s}_T)\, p(\mathbf{s}_T \mid \mathbf{s}_{T-1})$$

simple inversion!

Posterior model

Reverse Algorithm!

must be made in batch!

Hidden Markov model:

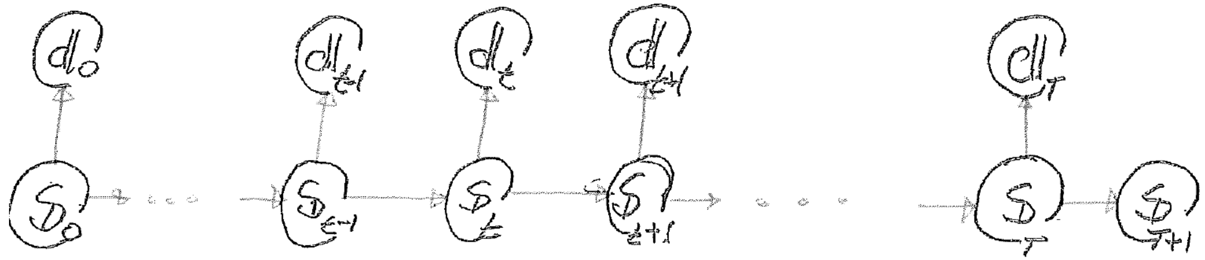
Smoothing:

$$p(s_{0:T}|d_{0:T}) = p(s|d) = p(s_0|d_{0:T}) \prod_t p(s_t|s_{t-1}, d_{t:T})$$

Posterior model

Marginal Smoothing

$p(s_t|d_{0:T})$ – marginalization $p(s|d)$



NOTE:
Forward Recursions
can be made in real time!

Filtering:

$p(s_T|d_{0:T})$

Recursion PF:

$p(s_t|d_{0:t-1})$

$$p(s_t|d_{0:t}) = \left[p(d_t|d_{0:t-1})\right]^{-1} p(d_t|s_t) \int p(s_t|s_{t-1}) p(s_{t-1}|d_{0:t-1}) ds_t$$

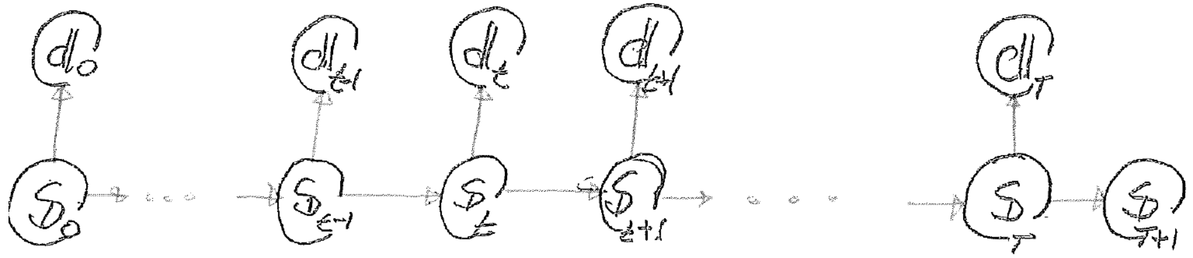
conditioning * forwarding

Recursion EnKF

$$p(s_{t+1}|d_{0:t}) = \int p(s_{t+1}|s_t) \left[p(d_t|d_{0:t-1})\right]^{-1} p(d_t|s_t) p(s_t|d_{0:t-1}) ds_t$$

forwarding * conditioning

Hidden Markov model:



Special Case – Trad Seq. Monte Carlo

Likelihood model

$$p(d|s) = \prod_t p(d_t|s_t)$$

Prior model:

$$p(s) = p(s_0)\prod_t p(s_t|s_{t-1}) = p(s_0)\prod_t \delta(s_t|s_{t-1})$$

Dirac

hence

$$[s_t|s_{t-1}] = s_{t-1} = s_0$$

Posterior model:

$$p(s_0|d) = p(s_0)\times\prod_t p(d_t|s_0)$$

NOTE:

Sequential conditioning

smoothing

filtering

identical!

simplest

Hidden Markov model

Viterbi [1967]

Posterior model:

$$p(\mathbf{s}|\mathbf{d}) = p(s_0|\mathbf{d}_{0:T}) \prod_t p(s_t|s_{t-1}, \mathbf{d}_{t:T})$$

Usually interested in:

Simulations:

$\mathbf{s}^i \leftarrow p(\mathbf{s}|\mathbf{d})$; $i = 1, \dots, n_s$ - Sequential Alg

Predictions:

$\hat{\mathbf{s}}_E = E\{\mathbf{s}|\mathbf{d}\}$ - use $E_t\{s_t\} = E_{t-1}\{E_t\{s_t|s_{t-1}\}\}$ - Seq. Alg

$\hat{\mathbf{s}}_{MAP} = MAP\{\mathbf{s}|\mathbf{d}\}$ - use Viterbi Alg

$\hat{\mathbf{s}}_{MMAP} = MMAP\{\mathbf{s}|\mathbf{d}\}$ - Sequential Alg

Prediction Precision:

$\sigma_E^2 = Diag\{Var\{\mathbf{s}|\mathbf{d}\}\}$

- use $Var_t\{s_t\} = E_{t-1}\{Var_t\{s_t|s_{t-1}\}\}$
$+ Var_{t-1}\{E_t\{s_t|s_{t-1}\}\}$ - Seq. Alg

$\underset{n\times 1}{\mathbf{p}_s} = Marg\{p(\mathbf{s}|\mathbf{d})\}$ - Sequential Alg

Hidden Markov model

Model parameter inference:

Criterion - maximum (marginal) likelihood

$$\hat{\Theta}_{lp} = \underset{\Theta_{lp}}{\operatorname{argmax}} \{ p(d; \Theta_l, \Theta_p) \}$$

NOTE:

$$p(d; \Theta_{lp}) = \int_0 \cdots \int_T p(d|s; \Theta_l)\, p(s; \Theta_p)\, ds_T \cdots ds_0$$

$$= \int_0 \cdots \int_T p(d_0|s_0;\cdot)\, p(s_0;\cdot) \prod_t p(d_t|s_t;\cdot)\, p(s_t|s_{t-1};\cdot)\, ds_T \cdots ds_0$$

$$= \int_0 p(d_0|s_0)\, p(s_0) \int_1 p(d_1|s_1)\, p(s_1|s_0) \int \cdots$$

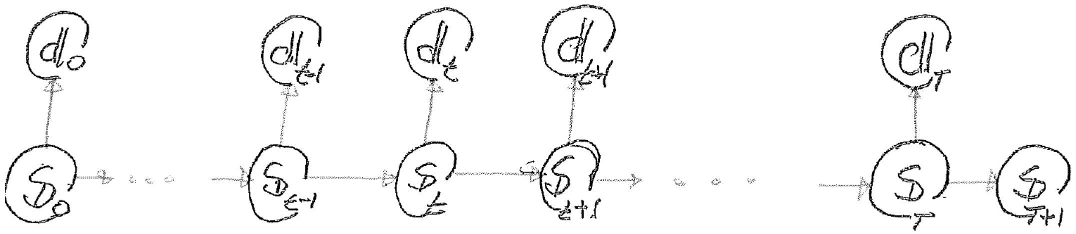
$$\boxed{p(d_{t:T}|s_{t-1})}$$

$$\times \int_T p(d_T|s_T)\, p(s_T|s_{T-1})\, ds_T \cdots ds_0$$

Sequential integration!

NOTE:

Must calculate $p(d; \Theta_l, \Theta_p)$ for each value of $[\Theta_l, \Theta_p]$.



Hidden Markov model

→ Categorical variables

$$\underset{n\times 1}{S_t} \longrightarrow \underset{1\times 1}{l_t} \in \mathbb{L} : \{W, \Psi, R, \dots, \underset{n_L}{B}\}$$ – discrete non-ordered 'colors'

$$\mathbf{l} = [l_0, \dots, l_t, \dots, l_T, l_{T+1}]$$

→ Likelihood models

$$[d_t|l_t] = \mu_{l_t} + \underbrace{\epsilon_t}_{\varphi_1(\epsilon;0,\sigma_\epsilon^2)} \Rightarrow \underbrace{\varphi_1(d; \mu_{l_t}, \sigma_\epsilon^2)}_{p(d_t|l_t)}$$

$$[d_t|l_t] = \Psi_{d_t|l_t} \quad \begin{array}{c|c} & d_t \\ l_t & \begin{array}{c} \\ \end{array} \end{array} \begin{array}{c} \quad W \qquad\qquad B \\ \begin{array}{c} W \\ \\ B \end{array}\begin{bmatrix} \circ & & \\ & & \\ & & \end{bmatrix} \end{array}$$ ← $p(d_t|l_t)$ mis-classification matrix

Algorithm 1: Reverse Algorithm - Calculating the posterior model

for *all* $l_n, l_{n-1} \in \mathbb{L}$ *and* d_n **do**
  $p(l_n|l_{n-1}, d_n) = \text{const} \times p(d_n|l_n)p(l_n|l_{n-1})$
end
$\text{const} = \left[\sum_{l'_n \in \mathbb{L}} p(d_n|l'_n)p(l'_n|l_{n-1})\right]^{-1}$
for *all* $i = n-1, \dots, 2$ **do**
  for *all* $l_i, l_{i-1} \in \mathbb{L}$ *and arbitrary* $l_{i+1} \in \mathbb{L}$ *and* $\mathbf{d}_{i:n}$ **do**
    $p(l_i|l_{i-1}, \mathbf{d}_{i:n}) = \text{const} \times \frac{p(d_i|l_i)p(l_i|l_{i-1}, l_{i+1})}{p(l_{i+1}|l_i, \mathbf{d}_{(i+1):n})}$
  end
  $\text{const} = \left[\sum_{l'_i \in \mathbb{L}} \frac{p(d_i|l_i)p(l'_i|l_{i-1}, l_{i+1})}{p(l_{i+1}|l'_i, \mathbf{d}_{(i+1):n})}\right]^{-1}$
end
for *all* $l_1 \in \mathbb{L}$ *and arbitrary* $l_2 \in \mathbb{L}$ *and* $\mathbf{d}$ **do**
  $p(l_1|\mathbf{d}) = \text{const} \times \frac{p(d_1|l_1)p(l_1|l_2)}{p(l_2|l_1, \mathbf{d}_{2:n})}$
end
$\text{const} = \left[\sum_{l'_1 \in \mathbb{L}} \frac{p(d_1|l_1)p(l'_1|l_2)}{p(l_2|l'_1, \mathbf{d}_{2:n})}\right]^{-1}$

Fjelltreit [2021]

→ Prior model

$$l_0 \Rightarrow p(l_0) = \underset{n_L\times 1}{q_{l_0}}$$

$$[l_t|l_{t-1}] \Rightarrow p(l_t|l_{t-1}) = \underset{n_L\times n_L}{q_{l_t|l_{t-1}}} \quad \begin{array}{c} \quad l_t \\ \begin{array}{c} \\ l_{t-1} \end{array} \end{array} \begin{array}{c} W \qquad\qquad B \\ \begin{array}{c} W \\ \\ B \end{array}\begin{bmatrix} \circ & & \\ & & \\ & & \circ \end{bmatrix} \end{array}$$ forward matrix

→ Posterior model – smoothing

$$[\mathbf{l}|\mathbf{d}] \Rightarrow p(\mathbf{l}|\mathbf{d}) = p(l_0|\mathbf{d}_{0:T}) \prod_t p(l_t|l_{t-1}, \mathbf{d}_{t:T})$$

Reverse Algorithm

– filtering

$$[l_T|\mathbf{d}_{0:T}] \Rightarrow p(l_T|\mathbf{d}_{0:T}) \longrightarrow$$ sequential Alg simple!

Hidden Markov model

→ Categorical variables $\xi_t \to l_t \in \mathbb{L} : \{W, Y, R, \dots, B\}$

→ Posterior model

$$[\mathbb{l}|\mathbb{d}] \leadsto p(\mathbb{l}|\mathbb{d}) = p(l_0|d_{0:T}) \prod_t p(l_t|l_{t-1}, d_{t:T})$$

→ Interested in:

- Simulation

$$\mathbb{l}^i \leadsto p(\mathbb{l}|\mathbb{d}); \; i=1,\dots,n_s \quad \text{- Sequential Alg}$$

Viterbi [1967]

Baum et al. [1970]:

Fjellvreit [2021]

- Prediction

$$\hat{\mathbb{l}}_{MAP} = MAP\{\mathbb{l}|\mathbb{d}\} = \underset{\mathbb{l}}{\operatorname{argmax}}\{p(\mathbb{l}|\mathbb{d})\}$$

- Viterbi Alg

$$\hat{\mathbb{l}}_{MMAP} = MMAP\{\mathbb{l}|\mathbb{d}\} = \{\hat{l}_t = \underset{l_t}{\operatorname{argmax}}\{p(l_t|\mathbb{d})\}; t \in \mathcal{T}_s\}$$

- Sequential Alg

- Prediction Precision

$$P_l = Marg\{p(\mathbb{l}|\mathbb{d})\} = \{P_t(l_t = l|\mathbb{d}); l \in \mathbb{L}, t \in \mathcal{T}_s\}$$

- Sequential Alg

- Model parameter inference

$$\hat{\Theta}_{lp} = \underset{\Theta_{lp}}{\operatorname{argmax}}\{p(\mathbb{d}; \Theta_l, \Theta_p)\} \quad \text{- Sequential + Optimization}$$

$[\mu_l, \sigma_\epsilon^2, \rho]$

Baum&Welch Alg

Algorithm 3: The Viterbi algorithm - MAP predictor

for *all* $l_2 \in \mathbb{L}$ **do**
  $\max_{l_1} p(\mathbf{l}_{1:2}|\mathbf{d}) = \max_{l_1} \left[p(l_1|\mathbf{d}) \times p(l_2|l_1, \mathbf{d}_{2:n})\right]$
end
for *all* $i = 2, \dots, n-1$ **do**
  for *all* $l_{i+1} \in \mathbb{L}$ **do**
    $\max_{\mathbf{l}_{1:i}} p(\mathbf{l}_{1:(i+1)}|\mathbf{d}) = \max_{l_i} \left[\max_{\mathbf{l}_{1:(i-1)}} p(\mathbf{l}_{1:i}|\mathbf{d}) \times p(l_{i+1}|l_i, \mathbf{d}_{(i+1):n})\right]$
  end
end
$\hat{l}_n = \arg\max_{l_n} \left[\max_{\mathbf{l}_{1:(n-1)}} p(\mathbf{l}|\mathbf{d})\right]$
for *all* $i = n-1, \dots, 1$ **do**
  $\hat{l}_i = \arg\max_{l_i} \left[\max_{\mathbf{l}_{1:(i-1)}} p(\mathbf{l}_{1:i}|\mathbf{d}) \times p(\hat{l}_{i+1}|l_i, \mathbf{d}_{(i+1):n})\right]$
end

Example – HMM Categorical

Fjellvreit [2021]

→ Focus on:

$$\mathbf{l} = [l_0, l_1, \dots, l_t, \dots, \underset{\angle 100}{l_T}]; \quad l_t \in \mathbb{L} : \{W, G, B\}$$

True profile: $\mathbf{l}^T$

→ Available observations:

$$\mathbf{d} = [d_0, d_1, \dots, d_t, \dots, d_T]; \quad d_t \in \mathbb{R}^T \quad (A)$$

$$\phantom{\mathbf{d} = [d_0, d_1, \dots, d_t, \dots, d_T]; \quad} d_t \in \mathbb{L} \quad (B)$$

→ Likelihood models:

(A) $[d_t | l_t] \Rightarrow p(d_t | l_t) = \varphi_1(d_t; \mu_{l_t}, 0.5^2)$

$$\mu_{l_t} = \begin{cases} 1 & - \; l_t = W \\ 0 & - \; l_t = G \\ -1 & - \; l_t = B \end{cases}$$

(B) $[d_t | l_t] \Rightarrow p(d_t | l_t) = Q_{d|l} \rightarrow$

$$\begin{array}{c|ccc} & W & G & B \\ \hline W & .85 & .10 & .05 \\ G & .10 & .80 & .10 \\ B & .05 & .10 & .85 \end{array}$$

Figure 7: The true image $\mathbf{l}^T$.

Figure 16: Observations from a Gaussian likelihood model with $\sigma^2 = 0.5^2$.

→ Prior model:

$$[l_t|l_{t-1}] \sim p(l_t|l_{t-1}) = P_{l|l} \rightarrow \begin{array}{c|ccc} & W & G & B \\ \hline W & 1-\beta_{WB} & \beta_{WB} & 0 \\ G & \beta_G/2 & 1-\beta_G & \beta_G/2 \\ B & 0 & \beta_{WB} & 1-\beta_{WB} \end{array}$$

$\beta_{WB} < \beta_G$

↓ reparam

$\beta_1 = 1.0$

$\beta_2 = 0.5$

→ Results:

- Model parameter inference – ML – Sequential Alg
- Posterior model – Markov model – Reverse Alg
  - Realizations – Sequential Alg
  - MAP prediction – Viterbi Alg
  - MMAP prediction – Sequential Alg
  - Prob profiles – Sequential Alg

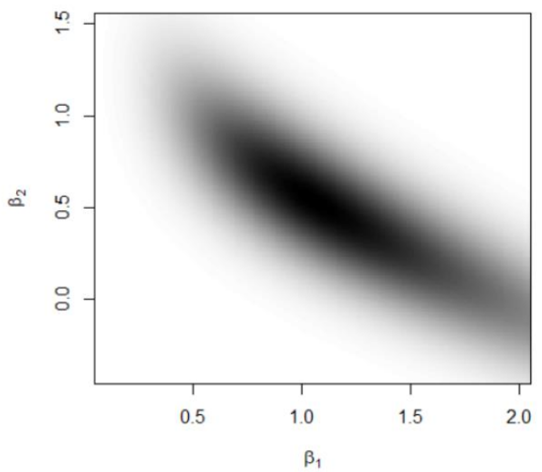
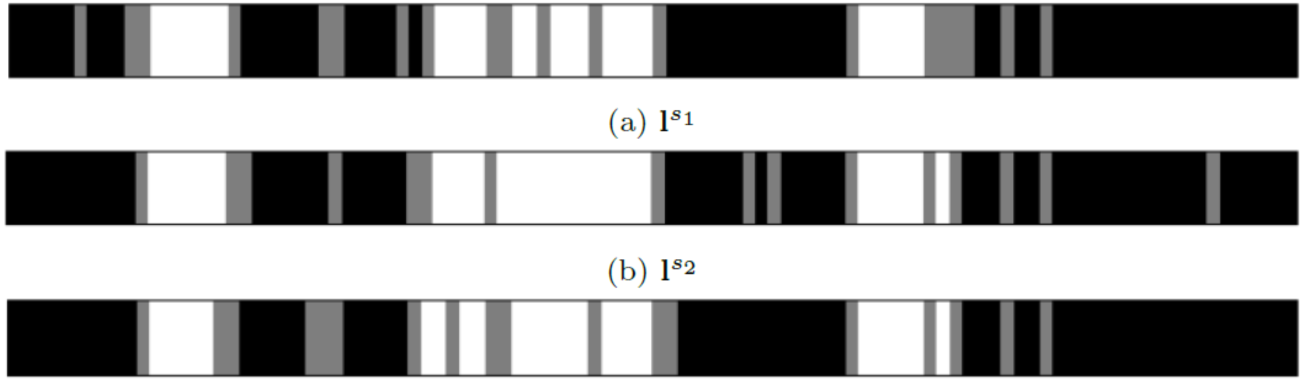


Figure 17: Density plot of the marginal likelihood model with respect to the parameters.



(a) $\mathbf{l}^{s_1}$

(b) $\mathbf{l}^{s_2}$

(c) $\mathbf{l}^{s_3}$

Figure 18: Three realizations of the posterior model.

→ Prior model:

$$[l_t | l_{t-1}] \sim p(l_t | l_{t-1}) = P_{l|l} \rightarrow \begin{array}{cc} & \begin{array}{ccc} W & G & B \end{array} \\ \begin{array}{c} W \\ G \\ B \end{array} & \begin{bmatrix} 1-\beta_{WB} & \beta_{WB} & 0 \\ \beta_G/2 & 1-\beta_G & \beta_G/2 \\ 0 & \beta_{WB} & 1-\beta_{WB} \end{bmatrix} \end{array}$$

$\beta_{WB} < \beta_G$

↓ reparam

$\beta_1 = 1.0$

$\beta_2 = 0.5$

→ Results:

- Model parameter inference – ML – Sequential Alg
- Posterior model – Markov model – Reverse Alg

  Realizations – Sequential Alg

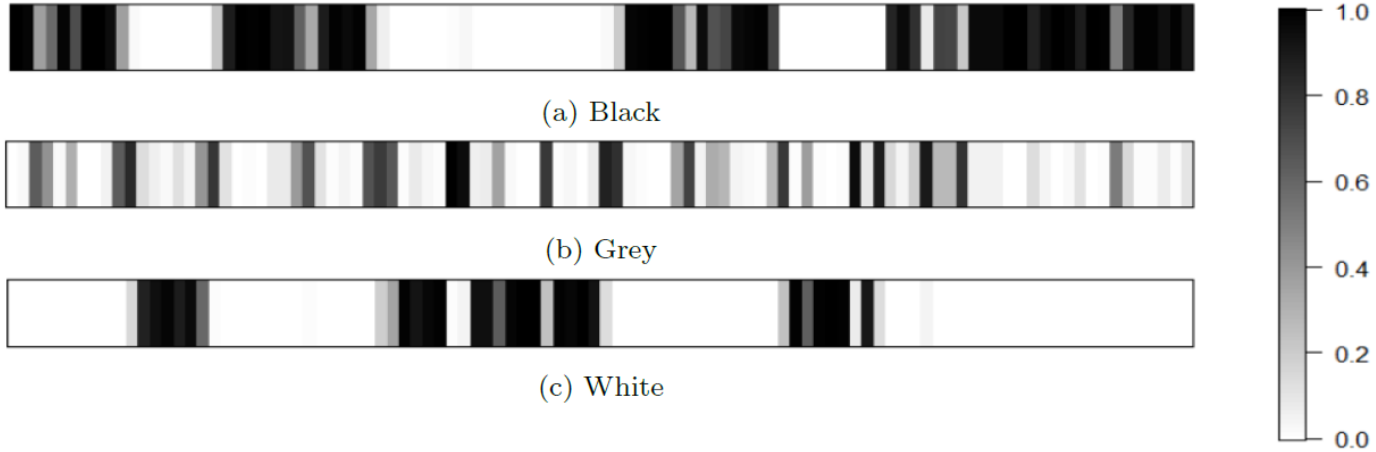
  MAP prediction – Viterbi Alg

  MMAP prediction – Sequential Alg

  Prob profiles – Sequential Alg

Figure 19: MAP predictor by the Viterbi algorithm.

Figure 20: MMAP predictor by Algorithm 4.



(a) Black

(b) Grey

(c) White

Figure 22: Probability profiles for the possible states {black, grey, white}.

Example – HMM Categorical

→ Focus on:

$$\mathbf{l} = [l_0, l_1, \ldots, l_t, \ldots, l_T]; \quad l_t \in \mathbb{L} : \{W, G, B\}$$

$\quad L_{100}$

True profile: $\mathbf{l}^T$

→ Available observations:

$$\mathbf{d} = [d_0, d_1, \ldots, d_t, \ldots, d_T]; \quad d_t \in \mathbb{R}^T \ \text{(A)}, \quad d_t \in \mathbb{L} \ \text{(B)}$$

→ Likelihood models:

(A) $[d_t | l_t] \Rightarrow p(d_t | l_t) = \varphi_1(d_t; \mu_{l_t}, 0.5^2)$

$$\mu_{l_t} = \begin{cases} 1 & - \; l_t = W \\ 0 & - \; l_t = G \\ -1 & - \; l_t = B \end{cases}$$

(B) $[d_t | l_t] \Rightarrow p(d_t | l_t) = Q_{d|l} \rightarrow$

$$\begin{array}{c|ccc} & W & G & B \\ \hline W & .85 & .10 & .05 \\ G & .10 & .80 & .10 \\ B & .05 & .10 & .85 \end{array}$$

Figure 7: The true image $\mathbf{l}^T$.

Figure 23: Observation from a misclassification based likelihood model.

→ Prior model:

$$[l_t | l_{t-1}] \to p(l_t | l_{t-1}) = P_{l|l} \to \begin{array}{c|ccc} & W & G & B \\ \hline W & 1-\beta_{WB} & \beta_{WB} & 0 \\ G & \beta_G/2 & 1-\beta_G & \beta_G/2 \\ B & 0 & \beta_{WB} & 1-\beta_{WB} \end{array}$$

$\beta_{WB} < \beta_G$

↓ reparam

$\beta_1 = 1.0$

$\beta_2 = 0.5$

→ Results:

- Model parameter inference - ML - Sequential Alg
- Posterior model - Markov model - Reverse Alg
  - Realizations - Sequential Alg
  - MAP prediction - Viterbi Alg
  - MMAP prediction - Sequential Alg
  - Prob profiles - Sequential Alg

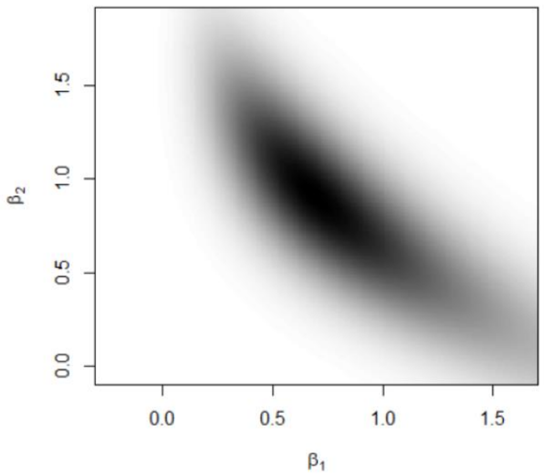
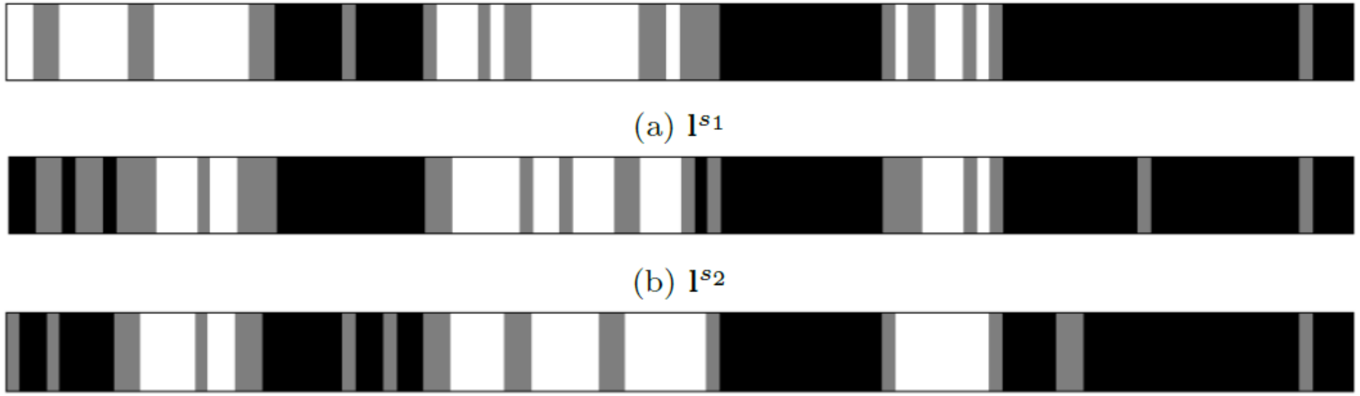


Figure 24: Density plot of the marginal likelihood model with respect to the parameters



(a) $\mathbf{l}^{s1}$

(b) $\mathbf{l}^{s2}$

(c) $\mathbf{l}^{s3}$

Figure 25: Three realizations of the posterior model.

→ Prior model:

$$[l_t | l_{t-1}] \sim p(l_t | l_{t-1}) = P_{l|l} \rightarrow \begin{array}{c|ccc} & W & G & B \\ \hline W & 1-\beta_{WB} & \beta_{WB} & 0 \\ G & \beta_G/2 & 1-\beta_G & \beta_G/2 \\ B & 0 & \beta_{WB} & 1-\beta_{WB} \end{array}$$

$\beta_{WB} < \beta_G$

$\downarrow$ reparam

$\beta_1 = 1.0$

$\beta_2 = 0.5$

→ Results:

- Model parameter inference - ML - Sequential Alg
- Posterior model - Markov model - Reverse Alg

  Realizations - Sequential Alg

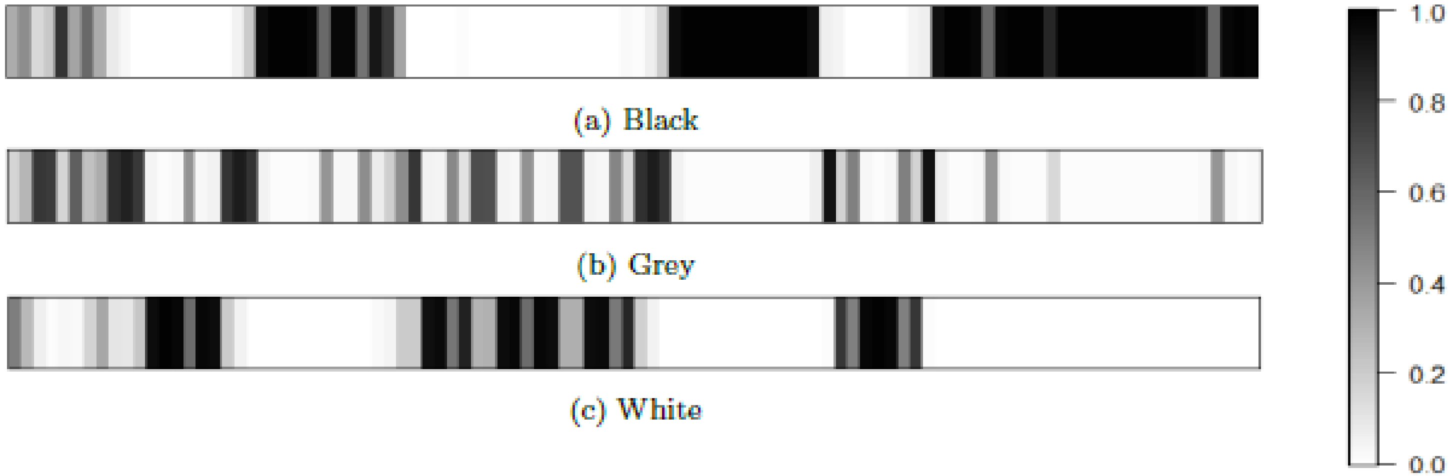
  MAP prediction - Viterbi Alg

  MMAP prediction - Sequential Alg

  Prob profiles - Sequential Alg

Figure 26: MAP prediction generated by the Viterbi algorithm (Algorithm 3).

Figure 27: MMAP prediction generated by Algorithm 4.



(a) Black

(b) Grey

(c) White

Figure 29: Probability profiles for the possible states {black, grey, white}

References:

Baum et al. [1970]: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains; Ann. Math. Statist; Vol. 41, No. 1, pp 164-171

Cappe et al [2005]: Inference in Hidden Markov Models; Springer Series in Statistics; Springer.

Fjelltreit [2021]: Image Analysis - Bayesian Inversion in Hidden Markov Models; MSc Thesis; IMF/NTNU, Trondheim

Moja et al [2019]: Bayesian inversion in hidden Markov models with varying marginal proportions, Math Geoscien; Vol. 51, pp. 463-484.

Scott [2002]: Bayesian methods for hidden Markov models: Recursive computing in the 21st century; JASA; Vol. 97, No. 457, pp 337-351

Viterbi [1967]: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm; IEEE Trans on Information Theory, Vol. 13, No. 2, pp. 260-269.