UNIVERSITEIT GENT

FACULTY OF ENGINEERING

The iHMMpredict algorithm

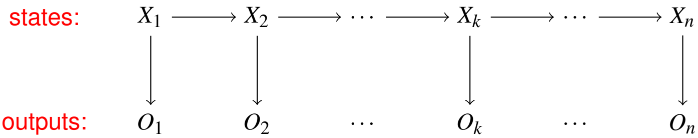
State sequence prediction in imprecise hidden Markov models

Jasper De Bock and Gert de Cooman

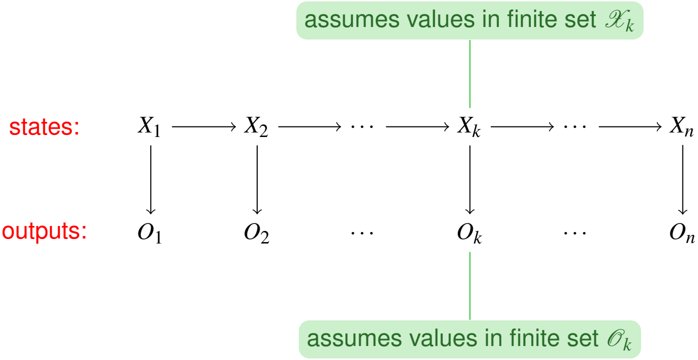
Ghent University, SYSTeMS
{Jasper.DeBock,Gert.deCooman}@UGent.be

IDSIA, 30 November 2010

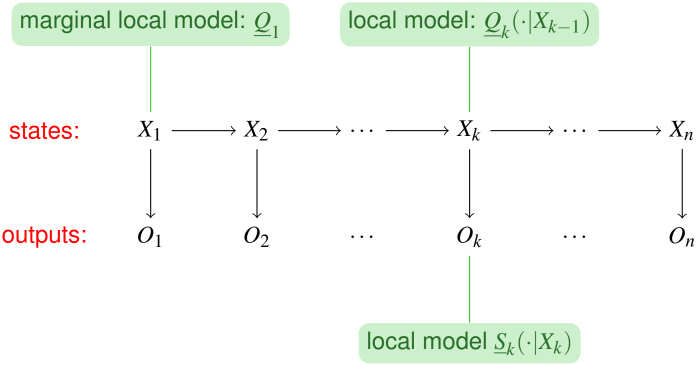
Imprecise Hidden Markov Model (iHMM)



Imprecise Hidden Markov Model (iHMM)

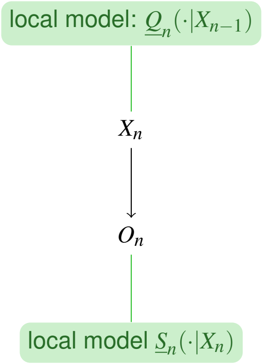


Imprecise Hidden Markov Model (iHMM)



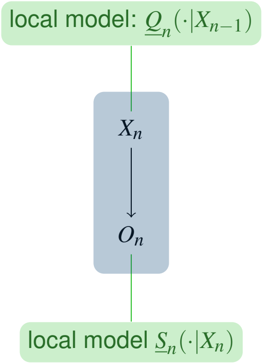
Recursive construction of the joint

First step



Recursive construction of the joint

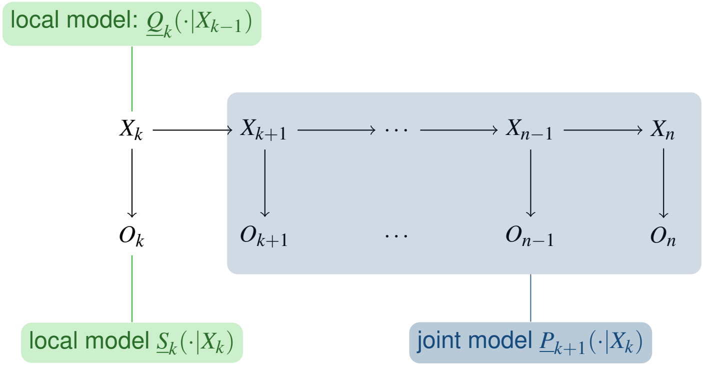
First step



Joint model $\underline{P}_n(\cdot|X_{n-1}) := \underline{Q}_n(\underline{S}_n(\cdot|X_n)|X_{n-1})$

Recursive construction of the joint

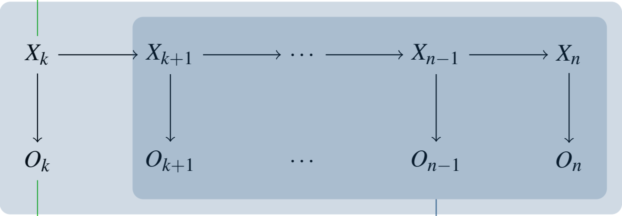
Recursive step



Recursive construction of the joint

Recursive step

local model: $\underline{Q}_k(\cdot|X_{k-1})$



local model $\underline{S}_k(\cdot|X_k)$

joint model $\underline{P}_{k+1}(\cdot|X_k)$

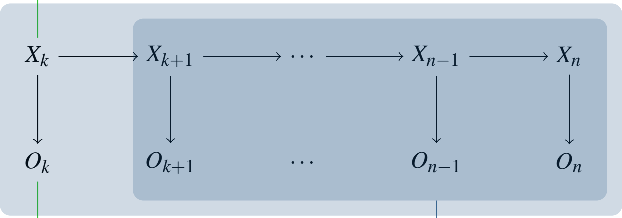
Independent natural extension $\underline{E}_k(\cdot|X_k) := \underline{S}_k(\cdot|X_k) \otimes \underline{P}_{k+1}(\cdot|X_k)$

Joint model $\underline{P}_k(\cdot|X_{k-1}) := \underline{Q}_k(\underline{E}_k(\cdot|X_k)|X_{k-1})$

Recursive construction of the joint

Recursive step

local model: $\underline{Q}_k(\cdot|X_{k-1})$



local model $\underline{S}_k(\cdot|X_k)$

joint model $\underline{P}_{k+1}(\cdot|X_k)$

Independent natural extension $\underline{E}_k(\cdot|X_k) := \underline{S}_k(\cdot|X_k) \otimes \underline{P}_{k+1}(\cdot|X_k)$

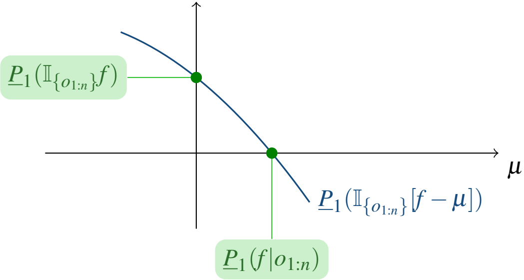
Joint model $\underline{P}_k(\cdot|X_{k-1}) := \underline{Q}_k(\underline{E}_k(\cdot|X_k)|X_{k-1})$

$\underline{E}_k(\cdot|X_k)$ is factorising $\Rightarrow$ very handy recursive expressions!

Updating the joint with the observations $o_{1:n}$

Since all local models are strictly positive: $\underline{P}_1(\{o_{1:n}\}) > 0$.

Generalised Bayes Rule yields uniquely coherent value of $\underline{P}_1(f|o_{1:n})$.



$$\underline{P}_1(f|o_{1:n}) \leq 0 \Leftrightarrow \underline{P}_1(\mathbb{I}_{\{o_{1:n}\}}f) \leq 0.$$

The notion of optimality we use: maximality

We define a partial order $\succ$ on state sequences:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}) > 0$$

The notion of optimality we use: maximality

We define a partial order $\succ$ on state sequences:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1\left(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}\right) > 0$$

A state sequence is optimal if it is undominated, or maximal:

$$\begin{aligned} \hat{x}_{1:n} \in \operatorname{opt}\left(\mathscr{X}_{1:n} | o_{1:n}\right) &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) x_{1:n} \not\succ \hat{x}_{1:n} \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}} | o_{1:n}\right) \leq 0 \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]\right) \leq 0, \end{aligned}$$

The notion of optimality we use: maximality

We define a partial order $\succ$ on state sequences:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1\left(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}\right) > 0$$

A state sequence is optimal if it is undominated, or maximal:

$$\begin{aligned}
\hat{x}_{1:n} \in \operatorname{opt}\left(\mathscr{X}_{1:n} | o_{1:n}\right) &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) x_{1:n} \not\succ \hat{x}_{1:n} \\
&\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}} | o_{1:n}\right) \leq 0 \\
&\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]\right) \leq 0,
\end{aligned}$$

$\underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]\right)$ can be easily determined recursively!

The notion of optimality we use: maximality

We define a partial order $\succ$ on state sequences:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1\left(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}\right) > 0$$

A state sequence is optimal if it is undominated, or maximal:

$$\begin{aligned} \hat{x}_{1:n} \in \operatorname{opt}\left(\mathscr{X}_{1:n} | o_{1:n}\right) &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) x_{1:n} \not\succ \hat{x}_{1:n} \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}} | o_{1:n}\right) \leq 0 \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n}) \underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]\right) \leq 0, \end{aligned}$$

$\underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]\right)$ can be easily determined recursively!

Aim of the algorithm:

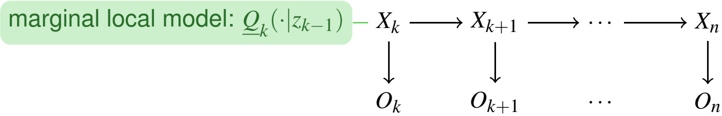
Determine $\operatorname{opt}\left(\mathscr{X}_{1:n} | o_{1:n}\right)$ efficiently.

More general optimality operators

$\mathrm{opt}\left(\mathscr{X}_{k-1}|z_{k-1},o_{k:n}\right)$

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k-1}|z_{k-1},o_{k:n}\right)$$
$$\Leftrightarrow (\forall x_{k:n} \in \mathscr{X}_{k:n})\underline{P}_k(\mathbb{I}_{\{o_{k:n}\}}[\mathbb{I}_{\{x_{k:n}\}} - \mathbb{I}_{\{\hat{x}_{k:n}\}}]|z_{k-1}) \leq 0.$$

These are the optimal sequences for part of the original iHMM:



The corresponding joint lower prevision is precisely $\underline{P}_k(\cdot|z_{k-1})$.

The Principle of Optimality

$$\hat{x}_{k:n} \in \text{opt}\left(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \text{opt}\left(\mathscr{X}_{k+1:n} | \hat{x}_k, o_{k+1:n}\right).$$

And so we can say that:

$$\text{opt}\left(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n}\right) \subseteq S_k := \bigcup_{z_k \in \mathscr{X}_k} z_k \oplus \text{opt}\left(\mathscr{X}_{k+1:n} | z_k, o_{k+1:n}\right).$$

And therefore:

$$\text{opt}\left(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n}\right) = \text{opt}\left(S_k | z_{k-1}, o_{k:n}\right).$$

Checking if a final state $\hat{x}_n$ is optimal

$$\hat{x}_n \in \operatorname{opt}(\mathscr{X}_n|z_{k-1},o_k) \Leftrightarrow (\forall x_n \in \mathscr{X}_n)\underline{P}_n(\mathbb{I}_{\{o_n\}}[\mathbb{I}_{\{x_n\}} - \mathbb{I}_{\{\hat{x}_n\}}]|z_{n-1}) \leq 0.$$

This is OK if $x_n = \hat{x}_n$, so we can use the equivalent condition:

$$(\forall x_n \in \mathscr{X}_n \setminus \{\hat{x}_n\})\underline{Q}_n(\mathbb{I}_{\{x_n\}}\underline{S}_n(\{o_n\}|x_n) - \mathbb{I}_{\{\hat{x}_n\}}\overline{S}_n(\{o_n\}|\hat{x}_n)|z_{n-1}) \leq 0.$$

So for every z_{n-1} we can easily find $\operatorname{opt}(\mathscr{X}_n|z_{n-1},o_n)$.

Checking if a state sequence $\hat{x}_{k:n}$ is optimal

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|z_{k-1}, o_{k:n})$$
$$\Leftrightarrow (\forall x_{k:n} \in \mathscr{X}_{k:n})\underline{P}_k(\mathbb{I}_{\{o_{k:n}\}}[\mathbb{I}_{\{x_{k:n}\}} - \mathbb{I}_{\{\hat{x}_{k:n}\}}]|z_{k-1}) \leq 0.$$

Since $\hat{x}_{k:n} \in S_k$, the inequality holds when $x_k = \hat{x}_k$.

This leads to the equivalent condition:

$$(\forall x_k \neq \hat{x}_k)(\forall x_{k+1:n} \in \mathscr{X}_{k+1:n})\underline{Q}_k(\mathbb{I}_{\{x_k\}}\beta(x_{k:n}) - \mathbb{I}_{\{\hat{x}_k\}}\alpha(\hat{x}_{k:n})|z_{k-1}) \leq 0.$$

where

$$\beta(z_{k:n}) := \underline{S}_k(\{o_k\}|z_k) \prod_{i=k+1}^{n} \underline{S}_i(\{o_i\}|z_i)\underline{Q}_i(\{z_i\}|z_{i-1})$$
$$\alpha(z_{k:n}) := \overline{S}_k(\{o_k\}|z_k) \prod_{i=k+1}^{n} \overline{S}_i(\{o_i\}|z_i)\overline{Q}_i(\{z_i\}|z_{i-1}).$$

Checking if a state sequence $\hat{x}_{k:n}$ is optimal

Equivalent formulations for $\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n})$

$$(\forall x_k \neq \hat{x}_k)(\forall x_{k+1:n} \in \mathscr{X}_{k+1:n})\underline{Q}_k(\mathbb{I}_{\{x_k\}}\beta(x_{k:n}) - \mathbb{I}_{\{\hat{x}_k\}}\alpha(\hat{x}_{k:n})|z_{k-1}) \leq 0$$

is equivalent to

$$(\forall x_k \neq \hat{x}_k)\underline{Q}_k(\mathbb{I}_{\{x_k\}}\beta_k^{\max}(x_k) - \mathbb{I}_{\{\hat{x}_k\}}\alpha(\hat{x}_{k:n})|z_{k-1}) \leq 0,$$

where

$$\beta_k^{\max}(x_k) := \max_{\substack{z_{k:n} \in \mathscr{X}_{k:n} \\ z_k = x_k}} \beta(z_{k:n}).$$

Checking if a state sequence $\hat{x}_{k:n}$ is optimal

Equivalent formulations for $\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|z_{k-1}, o_{k:n})$

$$(\forall x_k \neq \hat{x}_k)\underline{Q}_k(\mathbb{I}_{\{x_k\}}\beta_k^{\max}(x_k) - \mathbb{I}_{\{\hat{x}_k\}}\alpha(\hat{x}_{k:n})|z_{k-1}) \leq 0$$

is equivalent to

$$(\forall x_k \neq \hat{x}_k)\alpha(\hat{x}_{k:n}) \geq \alpha_{threshold}(\hat{x}_k, x_k|z_{k-1}),$$

where

$$\alpha_{threshold}(\hat{x}_k, x_k|z_{k-1}) := \min\left\{\alpha \in \mathbb{R} \colon \underline{Q}_k(\mathbb{I}_{\{x_k\}}\beta_k^{\max}(x_k) - \alpha\mathbb{I}_{\{\hat{x}_k\}}|z_{k-1}) \leq 0\right\}$$

and finally, this is in turn equivalent to:

$$\alpha(\hat{x}_{k:n}) \geq \alpha^{\mathrm{opt}}(\hat{x}_k|z_{k-1}) := \max_{x_k \neq \hat{x}_k} \alpha_{threshold}(\hat{x}_k, x_k|z_{k-1}).$$

Determining the set of optimal $\hat{x}_{k:n}$

by backward-forward recursion

For $k = n \to 1$:

$$\text{opt}\left(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n}\right) = \{\hat{x}_{k:n} \in S_k \colon \alpha(\hat{x}_{k:n}) \geq \alpha^{\text{opt}}(\hat{x}_k | z_{k-1})\}$$

Now recurse forward:

First step: Fix any $\hat{x}_k \in \mathscr{X}_k$.
Then there is some optimal state sequence $z_{k:n} \in \text{opt}\left(\mathscr{X}_{k:n} | z_{k-1}, o_{k:n}\right)$ such that $z_k = \hat{x}_k$ if and only if

$$\max_{\substack{z_{k:n} \in \mathscr{X}_{k:n} \\ z_k = \hat{x}_k}} \alpha(z_{k:n}) =: \alpha_k^{\max}(\hat{x}_k) \geq \alpha^{\text{opt}}(\hat{x}_k | z_{k-1}).$$

Determining the set of optimal $\hat{x}_{k:n}$

by backward-forward recursion

Second step: Fix any $\hat{x}_k$ for which $\alpha_k^{\max}(\hat{x}_k) \geq \alpha^{\text{opt}}(\hat{x}_k|z_{k-1})$.
Consider any first state $\hat{x}_{k+1}$ of some element $\hat{x}_{k+1:n}$ of $\text{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}\right)$ then:

$$\begin{aligned}
&\alpha(\hat{x}_{k:n}) \geq \alpha^{\text{opt}}(\hat{x}_k|z_{k-1}) \\
&\quad \Leftrightarrow \alpha(\hat{x}_{k+1:n})\overline{S}_{k+1}(\{o_{k+1}\}|\hat{x}_{k+1})\overline{Q}_{k+2}(\{\hat{x}_{k+2}\}|\hat{x}_{k+1}) \geq \alpha^{\text{opt}}(\hat{x}_k|z_{k-1}) \\
&\quad \Leftrightarrow \alpha(\hat{x}_{k+1:n}) \geq \alpha^{\text{opt}}(\hat{x}_{k:k+1}|z_{k-1}) := \frac{\alpha^{\text{opt}}(\hat{x}_k|z_{k-1})}{\overline{S}_k(\{o_k\}|\hat{x}_k)\overline{Q}_{k+1}(\{\hat{x}_{k+1}\}|\hat{x}_k)}
\end{aligned}$$

So we see that there is some $z_{k:n} \in \text{opt}\left(\mathscr{X}_{k:n}|z_{k-1}, o_{k:n}\right)$ such that $z_k = \hat{x}_k$ and $z_{k+1} = \hat{x}_{k+1}$ if and only if

$$\alpha_{k+1}^{\max}(\hat{x}_{k+1}) \geq \alpha^{\text{opt}}(\hat{x}_{k:k+1}|z_{k-1}).$$

Go on until you reach the end of the chain ...