Equivalence of State Representations for Hidden Markov Models

B. Vanluyten K. De Cock J.C. Willems B. De Moor

Department of Electrical Engineering
Katholieke Universiteit Leuven
Belgium

European Control Conference 2007
02-05/07/2007

Outline

Outline

Introduction

- Hidden Markov Models (HMM) are frequently used in many engineering aplications: speech processing, image analysis, bioinformatics.
- Many open problems.
- The realization problem:
  - Realizability
  - Realization algorithms
  - Equivalence problem

Outline

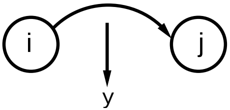
Hidden Markov models (1)

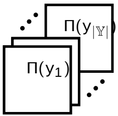
- Moore and Mealy type

  State and output process take values from finite sets

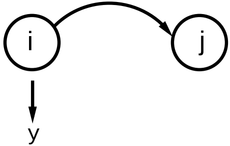
  $\mathbb{X} = \{1, 2, \ldots, |\mathbb{X}|\}$ and $\mathbb{Y} = \{y_1, y_2, \ldots, y_{|\mathbb{Y}|}\}$.

  - Mealy model $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$
    - $\Pi(y)_{ij} = P(x(t+1) = j, y(t) = y | x(t) = i)$,
    - $\pi_i = P(x(1) = i)$.





  - Moore model $(\mathbb{X}, \mathbb{Y}, \Pi_{\mathbb{X}}, \beta, \pi(1))$
    - $(\Pi_{\mathbb{X}})_{ij} = P(x(t+1) = j | x(t) = i)$,
    - $\beta(y)_i = P(y(t) = y | x(t) = i)$ or $B := \left[\begin{array}{ccc} \beta(y_1) & \ldots & \beta(y_{|\mathbb{Y}|}) \end{array} \right]$,
    - $\pi_i = P(x(1) = i)$.



Hidden Markov models (2)

- **String probabilities**
  for Mealy models $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$

$$\mathcal{P}(y_1 y_2 \ldots y_{|y|}) = \pi \Pi(y_1) \Pi(y_2) \ldots \Pi(y_{|y|}) e,$$

with $e := \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$. for Moore models $(\mathbb{X}, \mathbb{Y}, \Pi_{\mathbb{X}}, \beta, \pi(1))$

$$\mathcal{P}(y_1 y_2 \ldots y_{|y|}) = \pi \,\mathrm{diag}(\beta(y_1)) \Pi_{\mathbb{X}} \,\mathrm{diag}(\beta(y_2)) \Pi_{\mathbb{X}} \ldots \mathrm{diag}(\beta(y_{|y|})) \Pi_{\mathbb{X}} e.$$

- **Realization problem**
  - Mealy realization problem:
    Given output string probabilities $\mathcal{P}$, find a Mealy HMM $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$ that realizes $\mathcal{P}$, i.e. such that

$$\mathcal{P}(y) = \pi(1) \Pi(y_1)) \Pi(y_2) \ldots \Pi(y_{|y|}) e, \qquad \forall y = y_1 y_2 \ldots y_{|y|} \in \mathbb{Y}^*.$$

  - Realizations problem very hard due to nonnegativity constraints.
  - Typically one (first) solves the quasi realization problem.

Hidden Markov models (3)

- Quasi Hidden Markov model
  - Analogous to HMM but no nonnegativity constraints on the system vectors and matrices.
    - ★ Mealy type: $(\mathbb{X}_q, \mathbb{Y}, \Pi_q, \pi_q(1), e_q)$

$$\mathcal{P}(\mathrm{y}_1\mathrm{y}_2\ldots\mathrm{y}_{|\mathrm{y}|}) = \pi_q \Pi_q(\mathrm{y}_1)\Pi_q(\mathrm{y}_2)\ldots\Pi_q(\mathrm{y}_{|\mathrm{y}|})e_q,$$

    - ★ Moore type: $(\mathbb{X}_q, \mathbb{Y}, \Pi_{\mathbb{X},q}, \beta_q, \pi_q(1), e_q)$

$$\mathcal{P}(\mathrm{y}_1\mathrm{y}_2\ldots\mathrm{y}_{|\mathrm{y}|}) = \pi_q \,\mathrm{diag}(\beta_q(\mathrm{y}_1))\Pi_{\mathbb{X},q}\ldots\mathrm{diag}(\beta_q(\mathrm{y}_{|\mathrm{y}|}))\Pi_{\mathbb{X},q}e_q.$$

  - In some applications (filtering problems) it suffices to have a quasi realization.

Analogy: linear Gaussian models

- Output $y(t) \in \mathbb{R}^p$ generated by stochastic system (A, C, Q, R, S) of order n (assumed *minimal*)

$$
\begin{aligned}
x(t+1) &= Ax(t) + w(t) \\
y(t) &= Cx(t) + v(t)
\end{aligned}
$$

  with $w(t)$ and $v(t)$ zero mean, white, Gaussian vector sequences with joint covariance matrix:

$$
E \begin{bmatrix} w(t') \\ v(t') \end{bmatrix} \begin{bmatrix} w^\top(t'') & v^\top(t'') \end{bmatrix} = \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \delta_{t't''}.
$$

- General: Mealy; S=0: Moore.

Outline

Equivalence of (Mealy) linear Gaussian models

- (A, C, Q, R, S) gives rise to the autocovariance sequence $\Lambda(\tau) = E(y(t+\tau)y(t)^\top)$ through $P := E(x(t)x(t)^\top)$ and $G := E(x(t+1)y(t)^\top))$, computed as

$$\begin{aligned} P &= APA^\top + Q & \qquad G &= APC^\top + S \\ \Lambda(0) &= CPC^\top + R & \qquad \Lambda(t) &= CA^{t-1}G \end{aligned}$$

- Equivalent models :$\Leftrightarrow$ same autocovariance sequence.
- First type of equivalence:
  Performing a basis change in the state space gives rise to an equivalent model $(TAT^{-1}, CT^{-1}, TQT^\top, R, TS)$.
- Second type of equivalence:
  Every possible $P > 0$ which fullfills

$$\left[\begin{array}{c|c} P - APA^\top & G - APC^\top \\ \hline G^\top - CPA^\top & \Lambda(0) - CPC^\top \end{array}\right] \geq 0,$$

  gives also rise to an equivalent model $(A, C, P - APA^\top, \Lambda(0) - CPC^\top, G - APC^\top)$.

Equivalence of Mealy HMMs (1)

- Equivalent models $:\Leftrightarrow$ same string probabilities.
- Permuting states gives rise to equivalent model ($\approx$ basis change).
- Many more equivalent representations

Theorem (Vidyasagar)

Two minimal quasi Mealy models $(\mathbb{X}_q, \mathbb{Y}, \Pi_q, \pi_q(1), e_q)$ *and* $(\mathbb{X}'_q, \mathbb{Y}, \Pi'_q, \pi'_q(1), e'_q)$ *are equivalent, if and only if there exists a nonsingular matrix S, such that*

$$\begin{aligned} \pi_q &= \pi'_q S^{-1} \\ \Pi_q(y) &= S \Pi'_q(y) S^{-1} \quad \forall y \in \mathbb{Y} \\ e_q &= S e'_q \end{aligned}$$

Equivalence of Mealy HMMs (2)

Theorem

For a (positive) Mealy hidden Markov model $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$ *the following holds. If there exists a nonsingular matrix* T *such that*

- $Te = e$,
- $\pi(1)T^{-1}$ *and* $T\Pi(y)T^{-1} \quad \forall y \in \mathbb{Y}$ *are non-negative*

then $(\mathbb{X}, \mathbb{Y}, T\Pi T^{-1}, \pi(1)T^{-1})$ *is a positive hidden Markov model which is equivalent to the given HMM* $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$.

- If $(\mathbb{X}, \mathbb{Y}, \Pi, \pi(1))$ is minimal as a quasi model, these conditions are necessary and sufficient. In that case the theorem gives a complete description of the set of equivalent models.
- Simulation: Plot π of all possible HMMs equivalent to the given HMM (minimal as quasi HMM).

$$\Pi(0) = \begin{bmatrix} 0.05 & 0.4 & 0.1 \\ 0.2 & 0.15 & 0.2 \\ 0.35 & 0.15 & 0.2 \end{bmatrix}, \quad \Pi(1) = \begin{bmatrix} 0.15 & 0.1 & 0.2 \\ 0.1 & 0.05 & 0.3 \\ 0.05 & 0.15 & 0.1 \end{bmatrix}, \quad \pi = [\ 0.3060 \quad 0.3284 \quad 0.3657\ ].$$

Equivalence of Mealy HMMs (3)

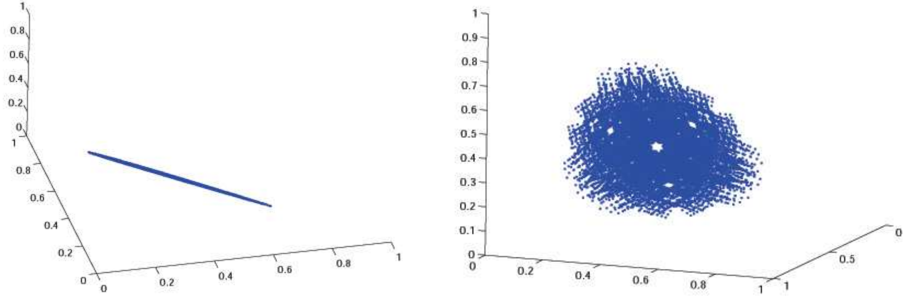


Figure: Plot of π of all possible HMMs which are equivalent to a given HMM (the two subplots give different views of the same three-dimensional plot).

- Problem: Give an accurate description of this set.

Outline

Equivalence of Moore HMMs (1)

Theorem

Given a quasi Moore HMM $(\mathbb{X}, \mathbb{Y}, \Pi_{\mathbb{X},q}, B_q, \pi_q(1), e_q)$ *which is minimal as a quasi Mealy model then the following holds.*

If all states of the Moore model have a different output distribution

and if the state transition matrix $\Pi_{\mathbb{X},q}$ *has full rank,*

then the only equivalent minimal quasi Moore models are obtained by permuting the states of the original model.

Equivalence of Moore HMMs (2)

- A quasi Mealy model with two output symbols can be converted, under general conditions, into an equivalent quasi Moore model with the same number of states.
- Every minimal quasi Moore model with two outputs is minimal as a quasi Mealy model $\rightarrow$ previous theorem can be applied.
- Also useful for HMMs with more than two outputs. According to the theorem, the Moore model below (3 outputs) has no Moore equivalents (except trivial ones).

$$\Pi_{\mathbb{X}} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}, \quad B = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.1 & 0.1 & 0.8 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}, \quad \pi(1) = \begin{bmatrix} 0.5405 & 0.1622 & 0.2973 \end{bmatrix}.$$

Equivalence of Moore linear Gaussian models

Theorem

Given a minimal Moore Gauss Markov model $(A, C, Q, R, 0)$ *then the following holds:*

If C has full column rank

and A has full rank,

then the only equivalent minimal Moore models are obtained by performing a base change in the state space.

Outline

Summary

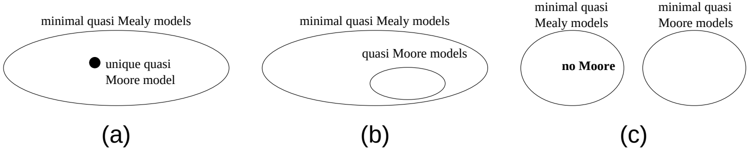


Figure: The three main cases for equivalence classes of HMMs.

a Quasi Moore model minimal as a quasi Mealy model and conditions of Theorem [equivalence of Moore models] fulfilled.
b Quasi Moore model minimal as a quasi Mealy model and conditions of Theorem [equivalence of Moore models] not fulfilled.
c Quasi Moore model not minimal as a quasi Mealy model.

Summary

- We considered the question: Given a (quasi) HMM, what can be said about the set of all equivalent (quasi) HMMs?
- For quasi Mealy HMM, a necessary and sufficient condition for two models to be equivalent was already proven in literature.
- We have proven a sufficient condition for two positive Mealy models to be equivalent.
- We have proven that under certain conditions, the set of equivalent minimal quasi Moore models consists of only one element.
- The situation is very analogous to the situation for Gauss Markov stochastic systems.