Equations for Hidden Markov Models

Alexander Schönhuth

Pacific Institute for the Mathematical Sciences
Simon Fraser University

December 2008
Guideline

Introduction

Hidden Markov Models: Alternative Formulation

String Functions

Stochastic Processes as String Functions
Dimension of Random Processes
Finite-dimensional Models
Invariants

Finite String Length Complexity

Degree of Hidden Markov Model Invariants
Hidden Markov Chains

- Initial probabilities \( \pi = (0.8, 0.2)^T \)
- Transition probabilities
  \[
  M = (m_{ij} := P(i \rightarrow j))_{i,j=1,2} = \begin{pmatrix}
  0.3 & 0.7 \\
  0.5 & 0.5 
\end{pmatrix}
  \]
- Emission probabilities, e.g. \( e_{1b} = 0.5, e_{2c} = 0.45 \).
Hidden Markov Chains

- Initial probabilities $\pi = (0.8, 0.2)^T$
- Transition probabilities
  \[ M = (m_{ij} := P(i \rightarrow j))_{i,j=1,2} = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix} \]
- Emission probabilities, e.g. $e_{1b} = 0.5, e_{2c} = 0.45.$

Random source $(X_t)$ with values in $\Sigma = \{a, b, c\}$:

\[
P_X(X_1 = a, X_2 = b) = \pi_1 e_{1a}(a_{11} e_{1b} + a_{12} e_{2b}) + \pi_2 e_{2a}(a_{21} e_{1b} + a_{22} e_{2b})
\]
Alternative Formulation

- Initial probabilities \( \pi = (0.8, 0.2)^T \)
- Transition probabilities
  \[
  M = (m_{ij} := P(i \to j))_{i,j=1,2} = \begin{pmatrix}
  0.3 & 0.7 \\
  0.5 & 0.5 
\end{pmatrix}
  \]
- Emission probabilities, e.g. \( e_{1b} = 0.5, e_{2c} = 0.45 \).

Alternative formulation: Let

\[
T_v := M^T \begin{pmatrix} e_{1v} & 0 \\ 0 & e_{2v} \end{pmatrix}, \quad v = a, b, c, \quad C := \begin{pmatrix} 1 \\ 1 \end{pmatrix} (\pi_1, \pi_2)
\]

then

\[
P_X(X_1 = v_1, \ldots, X_n = v_n) = \text{tr} \ T_{v_n} \cdot \ldots \cdot T_{v_1} \cdot C.
\]
Hidden Markov Models: Alternative Formulation

In the following, we write $a, b \in \Sigma$ for single letters and $v = v_1 \ldots v_n$, $w = w_1 \ldots w_m \in \Sigma^* = \bigcup_{k \geq 0} \Sigma^k$ for strings of finite length over an alphabet $\Sigma$. 
Hidden Markov Models: Alternative Formulation

In the following, we write \( a, b \in \Sigma \) for single letters and \( \nu = \nu_1...\nu_n, \)
\( w = w_1...w_m \in \Sigma^* = \cup_{k \geq 0} \Sigma^k \) for strings of finite length over an alphabet \( \Sigma \).

Hidden Markov model for \( l \) hidden states and strings of length \( n \) over the \( \Sigma \):

\[
f_{n,l} : \mathbb{C}^{l(l-1)+l|\Sigma|} \rightarrow \mathbb{C}^{|\Sigma|^n}
\]

\[
((T_a = M^T O_a)_{a \in \Sigma}, \pi) \mapsto (\text{tr } T_{\nu_1}...T_{\nu_n} (1, ..., 1)^T \pi)_{\nu_1...\nu_n \in \Sigma^n}.
\]

where \( M = (m_{ij} = P(i \rightarrow j)) \in \mathbb{C}^{l^2} \) and \( O_a \in \mathbb{C}^{l^2} \) is defined by

\[
(O_a)_{ij} = \begin{cases} e_{ia} & i = j \\ 0 & i \neq j \end{cases}
\]
Hidden Markov Models: Alternative Formulation

In the following, we write $a, b \in \Sigma$ for single letters and $v = v_1...v_n$, $w = w_1...w_m \in \Sigma^* = \bigcup_{k \geq 0} \Sigma^k$ for strings of finite length over an alphabet $\Sigma$.

Hidden Markov model for $l$ hidden states and strings of length $n$ over the $\Sigma$:

$$f_{n,l} : \mathbb{C}^{l(l-1)+l|\Sigma|} \rightarrow \mathbb{C}^{|\Sigma|^n}$$

$$((T_a = M^T O_a)_{a \in \Sigma}, \pi) \mapsto (\text{tr } T_{v_n}...T_{v_1} (1, ..., 1)^T \pi)_{v=v_1...v_n \in \Sigma^n}.$$

where $M = (m_{ij} = P(i \rightarrow j)) \in \mathbb{C}^{l^2}$ and $O_a \in \mathbb{C}^{l^2}$ is defined by

$$(O_a)_{ij} = \begin{cases} e_{ia} & i = j \\ 0 & i \neq j \end{cases}$$

Wanted: New equations describing

\[ \text{image } (f_{n,l}) \]
String Functions

Identify random processes \((X_t)\) with values in \(\Sigma\) as string functions \(p_X : \Sigma^* \to \mathbb{R}\) by

\[ p_X(v = v_1...v_n) := \mathbb{P}(X_1 = v_1, ..., X_n = v_n). \]
String Functions

Identify random processes \((X_t)\) with values in \(\Sigma\) as string functions 

\[ p_X : \Sigma^* \rightarrow \mathbb{R} \text{ by} \]

\[ p_X(v = v_1...v_n) := \mathbb{P}(X_1 = v_1, ..., X_n = v_n). \]

Theorem: A string function \( p : \Sigma^* \rightarrow \mathbb{R} \) is associated with a discrete random process iff the following conditions hold.

(a) \( p(v) \geq 0 \) for all \( v \in \Sigma^* \).

(b) \( \sum_{a \in \Sigma} p(va) = p(v) \) for all \( v \in \Sigma^* \).

(c) \( p(\Box) = 1 \) where \( \Box \in \Sigma^0 \) is the empty word.
String Functions

Identify random processes \((X_t)\) with values in \(\Sigma\) as string functions \(p_X : \Sigma^* \rightarrow \mathbb{R}\) by

\[
p_X(v = v_1...v_n) := \mathbb{P}(X_1 = v_1, ..., X_n = v_n).
\]

**Theorem:** A string function \(p : \Sigma^* \rightarrow \mathbb{R}\) is associated with a discrete random process iff the following conditions hold.

(a) \(p(v) \geq 0\) for all \(v \in \Sigma^*\).

(b) \(\sum_{a \in \Sigma} p(va) = p(v)\) for all \(v \in \Sigma^*\).

(c) \(p(\square) = 1\) where \(\square \in \Sigma^0\) is the empty word.

**Remark:** (b) in combination with (c) implies

\[
\forall n \geq 0 : \sum_{v \in \Sigma^n} p(v) = 1.
\]

In case that only (a) and (b) apply, \(p\) is referred to as **unconstrained** random process.
Dimension of String Functions

The Hankel Matrix

- Let \( wv = w_1 \ldots w_m v_1 \ldots v_n \in \Sigma^{m+n} \) be the concatenation of two strings \( w = w_1 \ldots w_m \in \Sigma^s, v = v_1 \ldots v_n \in \Sigma^t \).

- Consider the (infinite-dimensional) Hankel matrix

\[
P_p := [p(wv)]_{v,w \in \Sigma^*} \in \mathbb{R}^{\Sigma^* \times \Sigma^*} \cong \mathbb{R}^{N \times N}.
\]

for a string function \( p : \Sigma^* \to \mathbb{R} \).
Dimension of String Functions

The Hankel Matrix

- Let $wv = w_1...w_m v_1...v_n \in \Sigma^{m+n}$ be the concatenation of two strings $w = w_1...w_m \in \Sigma^s$, $v = v_1...v_n \in \Sigma^t$.

- Consider the (infinite-dimensional) Hankel matrix

$$P_p := [p(wv)]_{v,w \in \Sigma^*} \in \mathbb{R}^{\Sigma^* \times \Sigma^*} \cong \mathbb{R}^{N \times N}.$$ 

for a string function $p : \Sigma^* \rightarrow \mathbb{R}$.

Example: Let $\Sigma = \{0, 1\}$.

$$P_p = \begin{pmatrix}
p(\square) & p(0) & p(1) & \ldots \\
p(0) & p(00) & p(10) & \ldots \\
p(1) & p(01) & p(11) & \ldots \\
p(00) & p(000) & p(100) & \ldots \\
p(01) & p(001) & p(101) & \ldots \\
\vdots & \vdots & \vdots & \ddots 
\end{pmatrix}$$
Dimension of String Functions

The Hankel Matrix

- Let $wv = w_1...w_m v_1...v_n \in \Sigma^{m+n}$ be the concatenation of two strings $w = w_1...w_m \in \Sigma^s$, $v = v_1...v_n \in \Sigma^t$.  

- Consider the (infinite-dimensional) Hankel matrix 

$$
\mathcal{P}_p := [p(wv)]_{v,w \in \Sigma^*} \in \mathbb{R}^{\Sigma^* \times \Sigma^*} \cong \mathbb{R}^{N \times N}.
$$

for a string function $p : \Sigma^* \to \mathbb{R}$.

Example: Let $\Sigma = \{0, 1\}$.

$$
\mathcal{P}_p = \begin{pmatrix}
  p(\square) & p(0) & p(1) & \ldots \\
  p(0) & p(00) & p(10) & \ldots \\
  p(1) & p(01) & p(11) & \ldots \\
  p(00) & p(000) & p(100) & \ldots \\
  p(01) & p(001) & p(101) & \ldots \\
  \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

We define the dimension of $p$ to be

$$
\dim p := \text{rk } \mathcal{P}_p \in \mathbb{N} \cup \{\infty\}.
$$
Finite-dimensional String Functions

**Theorem:** [AS, Jaeger, 2007]
Let \( p : \Sigma^* \to \mathbb{R} \). Then the following conditions are equivalent.

(i) \[ \dim p = \operatorname{rk} \mathcal{P}_p \leq d. \]
Finite-dimensional String Functions

**Theorem:** [AS, Jaeger, 2007]
Let $p : \Sigma^* \rightarrow \mathbb{R}$. Then the following conditions are equivalent.

(i) \[ \dim p = \text{rk } P_p \leq d. \]

(ii) There exist vectors $x, y \in \mathbb{R}^d$ as well as matrices $T_a \in \mathbb{R}^{d \times d}$ for all $a \in \Sigma$ such that \[
\forall v \in \Sigma^* : p(v = v_1...v_n) = \langle y | T_{v_n}...T_{v_1} | x \rangle.
\]
Finite-dimensional String Functions

**Theorem:** [AS, Jaeger, 2007]
Let $p : \Sigma^* \rightarrow \mathbb{R}$. Then the following conditions are equivalent.

(i) \[ \dim p = \text{rk } \mathcal{P}_p \leq d. \]

(ii) There exist vectors $x, y \in \mathbb{R}^d$ as well as matrices $T_a \in \mathbb{R}^{d \times d}$ for all $a \in \Sigma$ such that
\[ \forall v \in \Sigma^* : \quad p(v = v_1 \ldots v_n) = \langle y | T_{v_n} \ldots T_{v_1} | x \rangle. \]

**Remark:** Let $C := xy^T \in \mathbb{R}^{d \times d}$:
\[ \langle y | T_{v_n} \ldots T_{v_1} | x \rangle = \text{tr } T_{v_n} \ldots T_{v_1} C. \]
Finite-dimensional String Functions

**Theorem:** [AS, Jaeger, 2007]
Let $p : \Sigma^* \to \mathbb{R}$. Then the following conditions are equivalent.

(i) $\dim p = \text{rk } P_p \leq d$.

(ii) There exist vectors $x, y \in \mathbb{R}^d$ as well as matrices $T_a \in \mathbb{R}^{d \times d}$ for all $a \in \Sigma$ such that

$$\forall v \in \Sigma^* : \quad p(v = v_1...v_n) = \langle y | T_{v_n}...T_{v_1} | x \rangle.$$ 

**Remark:** Let $C := xy^T \in \mathbb{R}^{d \times d}$:

$$\langle y | T_{v_n}...T_{v_1} | x \rangle = \text{tr } T_{v_n}...T_{v_1} C.$$ 

**Corollary:** A hidden Markov process on $l$ hidden states has dimension at most $l$. 
Finite-dimensional Models

**Definition:** Let

\[ g_{n,d} : \mathbb{C}^{\left|\Sigma\right|d^2 + 2d} \to \mathbb{C}^{\left|\Sigma\right|^n} \]

be the finite-dimensional model for strings of length \( n \) and dimension at most \( d \).
Finite-dimensional Models

**Definition:** Let

\[
g_{n,d} : \mathbb{C}^{|\Sigma|^d^2 + 2d} \rightarrow \mathbb{C}^{|\Sigma|^n}
((T_a)_{a \in \Sigma}, x, y) \mapsto (\text{tr} T_{v_n} \cdots T_{v_1} xy^T)_{v = v_1 \cdots v_n \in \Sigma^n}.
\]

be the finite-dimensional model for strings of length \(n\) and dimension at most \(d\).

**Observation:** Let \(f_{n,l}\) be the hidden Markov model for strings of length \(n\) and \(l\) hidden states. Then

\[
\text{image } (f_{n,l}) \subset \text{image } (g_{n,l}),
\]

that is, the hidden Markov models are submodels of the finite-dimensional models.
Hankel Matrices of Random Processes

∀v ∈ Σ* : p(v) = \sum_{a∈Σ} p(v a).

\[ P_p = \begin{pmatrix}
  p(\square) & p(0) & p(1) & \ldots \\
  p(0) & p(00) & p(10) & \ldots \\
  p(1) & p(01) & p(11) & \ldots \\
  p(00) & p(000) & p(100) & \ldots \\
  p(01) & p(001) & p(101) & \ldots \\
  \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix} \]

= \begin{pmatrix}
  p(0) + p(1) & p(00) + p(01) & p(10) + p(11) & \ldots \\
  p(00) + p(01) & p(000) + p(001) & p(100) + p(101) & \ldots \\
  p(10) + p(11) & p(010) + p(011) & p(110) + p(111) & \ldots \\
  p(000) + p(001) & p(000) & p(100) & \ldots \\
  p(010) + p(011) & p(001) & p(101) & \ldots \\
  \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix} \]
Hankel Matrices of Random Processes

\[
\forall \nu \in \Sigma^*, \ t \in \mathbb{N} : \sum_{\nu' \in \Sigma^t} p(\nu \nu') = p(\nu).
\]

\[
\mathcal{P}_p = \begin{pmatrix}
p(\square) & p(0) & p(1) & \ldots \\
p(0) & p(00) & p(10) & \ldots \\
p(1) & p(01) & p(11) & \ldots \\
p(00) & p(000) & p(100) & \ldots \\
p(01) & p(001) & p(101) & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\sum_{\nu' \in \Sigma^3} p(\nu') & \sum_{\nu' \in \Sigma^2} p(0\nu') & \sum_{\nu' \in \Sigma^2} p(1\nu') & \ldots \\
\sum_{\nu' \in \Sigma^2} p(0\nu') & \sum_{\nu' \in \Sigma} p(00\nu') & \sum_{\nu' \in \Sigma} p(10\nu') & \ldots \\
\sum_{\nu' \in \Sigma 2} p(1\nu') & \sum_{\nu' \in \Sigma} p(01\nu') & \sum_{\nu' \in \Sigma} p(11\nu') & \ldots \\
\sum_{\nu' \in \Sigma} p(00\nu') & p(000) & p(100) & \ldots \\
\sum_{\nu' \in \Sigma} p(01\nu') & p(001) & p(101) & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]
Finite-dimensional Model: Invariants

**Definition:** Let \( \mathcal{P}_{p,n,d} \) be the partial Hankel matrix that is filled with all values \( p(wv) \) such that \(|v|, |w| \leq d, |wv| \leq n\).

**Theorem:** Let \( n \geq 2d - 1 \). Then

\[
(p(v))_{v \in \Sigma^n} \in \text{image } (g_{n,d})
\]

if and only if the following two conditions apply:

(a) 

\[
\det [p(w_jv_i)]_{1 \leq i,j \leq d+1} = 0
\]

for all choices of words \( v_1, \ldots v_{d+1}, w_1, \ldots w_{d+1} \) of length at most \( d - 1 \) \( (\Leftrightarrow \text{rk } \mathcal{P}_{p,n,d-1} \leq d) \).

(b) Partial rows resp. columns in \( \mathcal{P}_{p,n,d} \) referring to strings \( v \) resp. \( w \) where \( |v|, |w| = d \) are linearly dependent of their counterparts in \( \mathcal{P}_{p,n,d-1} \).
**Invariants: Example**

**Example:** $n = 4$, $d = 2$

$$P_{p,4,2} = \begin{pmatrix}
  p(\square) & p(0) & p(1) & p(00) & p(01) & p(10) & p(11) \\
  p(0) & p(00) & p(10) & p(000) & p(010) & p(100) & p(110) \\
  p(1) & p(01) & p(11) & p(001) & p(011) & p(101) & p(111) \\
  p(00) & p(000) & p(100) & p(0000) & p(0100) & p(1000) & p(1100) \\
  p(01) & p(001) & p(101) & p(0001) & p(0101) & p(1001) & p(1101) \\
  p(10) & p(010) & p(110) & p(0010) & p(0110) & p(1010) & p(1110) \\
  p(11) & p(011) & p(111) & p(0011) & p(0111) & p(1011) & p(1111)
\end{pmatrix}$$
Invariants: Example

**Example**: Condition (a):

\[ P_{p,4,2} = \begin{pmatrix}
p(\square) & p(0) & p(1) & p(00) & p(01) & p(10) & p(11) \\
p(0) & p(00) & p(10) & p(000) & p(010) & p(100) & p(110) \\
p(1) & p(01) & p(11) & p(001) & p(011) & p(101) & p(111) \\
p(00) & p(000) & p(100) & p(0000) & p(0100) & p(1000) & p(1100) \\
p(01) & p(001) & p(101) & p(0001) & p(0101) & p(1001) & p(1101) \\
p(10) & p(010) & p(110) & p(0010) & p(0110) & p(1010) & p(1110) \\
p(11) & p(011) & p(111) & p(0011) & p(0111) & p(1011) & p(1111)
\end{pmatrix} \]

\[ \det \begin{pmatrix}
p(\square) & p(0) & p(1) \\
p(0) & p(00) & p(10) \\
p(1) & p(01) & p(11)
\end{pmatrix} = 0. \]
Invariants: Example

Example: Row condition (b):

\[
\mathcal{P}_{p,4,2} = \begin{pmatrix}
    p(00) & p(0) & p(1) & p(00) & p(01) & p(10) & p(11) \\
    p(0) & p(00) & p(10) & p(000) & p(010) & p(100) & p(110) \\
    p(1) & p(01) & p(11) & p(001) & p(011) & p(101) & p(111) \\
    p(00) & p(000) & p(100) & p(0000) & p(0100) & p(1000) & p(1100) \\
    p(01) & p(001) & p(101) & p(0001) & p(0101) & p(1001) & p(1101) \\
    p(10) & p(010) & p(110) & p(0010) & p(0110) & p(1010) & p(1110) \\
    p(11) & p(011) & p(111) & p(0011) & p(0111) & p(1011) & p(1111)
\end{pmatrix}
\]
Invariants: Example

\textbf{Example:} Column condition (b):

\[
\mathcal{P}_{p,4,2} = \begin{pmatrix}
    p(\square) & p(0) & p(1) & p(00) & p(01) & p(10) & p(11) \\
    p(0) & p(00) & p(10) & p(000) & p(010) & p(100) & p(110) \\
    p(1) & p(01) & p(11) & p(001) & p(011) & p(101) & p(111) \\
    p(00) & p(000) & p(100) & p(0000) & p(0100) & p(1000) & p(1100) \\
    p(01) & p(001) & p(101) & p(0001) & p(0101) & p(1001) & p(1101) \\
    p(10) & p(010) & p(110) & p(0010) & p(0110) & p(1010) & p(1110) \\
    p(11) & p(011) & p(111) & p(0011) & p(0111) & p(1011) & p(1111)
\end{pmatrix}
\]
Finite-dimensional Model: Invariants

**Conjecture:** Let $n \geq 2d - 1$ Then

$$(p(v))_{v \in \Sigma^n} \in \text{image } (g_{n,d})$$

if and only if

$$\det [p(w_j v_i)]_{1 \leq i,j \leq d+1} = 0$$

for all choices of words $v_1, \ldots, v_{d+1}, w_1, \ldots, w_{d+1}$ such that $|w_j v_i| \leq n$. 
Hidden Markov Processes

In the following, let

$$C_p := \text{span}\{\rho^v \mid v \in \Sigma^* \}$$

be the column space of the Hankel matrix $P_p$ of a string function $\rho$. 
Hidden Markov Processes

In the following, let

$$C_p := \text{span}\{p^v | v \in \Sigma^*\}$$

be the column space of the Hankel matrix $P_p$ of a string function $p$.

**Theorem:** A stochastic process $p : \Sigma^* \rightarrow \mathbb{R}$ is associated with a hidden Markov process if and only if there are stochastic processes $p_i \in C_p$, $i = 1, ..., l$ s.t.

(a) $p \in \text{cone}\{p_i | i = 1, ..., l\}$,

(b) $\forall v \in \Sigma^*: (p_i)^v \in \text{cone}\{p_i | i = 1, ..., l\}$,
Hidden Markov Processes

In the following, let

\[ C_p := \text{span}\{p^v \mid v \in \Sigma^*\} \]

be the column space of the Hankel matrix \( P_p \) of a string function \( p \).

**Theorem:** A stochastic process \( p : \Sigma^* \to \mathbb{R} \) is associated with a hidden Markov process if and only if there are stochastic processes \( p_i \in C_p \), \( i = 1, \ldots, l \) s.t.

(a) \( p \in \text{cone} \ \{p_i \mid i = 1, \ldots, l\} \),

(b) \( \forall v \in \Sigma^* : (p_i)^v \in \text{cone} \ \{p_i \mid i = 1, \ldots, l\} \),

**Corollary:** A (unconstrained) stochastic process \( p \) s.t. \( \dim p \leq 2 \) is associated with a (unconstrained) hidden Markov process (but not necessarily with 2 hidden states).
Motivation

**Conjecture:** The ideal of invariants of \( f_{n,2} \) over the alphabet \( \Sigma = \{0, 1\} \) is generated by linear and quadratic polynomials for large \( n \).
Motivation

**Conjecture:** The ideal of invariants of $f_{n,2}$ over the alphabet $\Sigma = \{0, 1\}$ is generated by linear and quadratic polynomials for large $n$.

**Conjecture:** The maximum degree of the generators of the ideal of invariants of $f_{n,d}$ over arbitrary alphabets does not increase for $n \geq 2d$. 
String Length Complexity

**Definition:** Let $\mathcal{M} \subseteq \mathbb{R}^{\Sigma^*}$ be a class of string functions. We define the *string length complexity* of $\mathcal{M}$ to be

$$\text{SLC} (\mathcal{M}) := \inf\{N \in \mathbb{N} \mid p_1, p_2 \in \mathcal{M} : p_1(v) = p_2(v), |v| \leq N \Rightarrow p_1 = p_2\}.$$
String Length Complexity

**Definition:** Let $\mathcal{M} \subset \mathbb{R}^{\Sigma^*}$ be a class of string functions. We define the string length complexity of $\mathcal{M}$ to be

$$\text{SLC} (\mathcal{M}) := \inf\{N \in \mathbb{N} \mid p_1, p_2 \in \mathcal{M} : p_1(v) = p_2(v), |v| \leq N \Rightarrow p_1 = p_2\}.$$

**Remark:**

- Clearly,

$$\mathcal{M}^* \subset \mathcal{M} \Rightarrow \text{SLC} (\mathcal{M}^*) \leq \text{SLC} (\mathcal{M}).$$
String Length Complexity

**Definition:** Let $\mathcal{M} \subset \mathbb{R}^{\Sigma^*}$ be a class of string functions. We define the string length complexity of $\mathcal{M}$ to be

$$\text{SLC} (\mathcal{M}) := \inf \{ N \in \mathbb{N} \mid p_1, p_2 \in \mathcal{M} : p_1(v) = p_2(v), |v| \leq N \Rightarrow p_1 = p_2 \}.$$  

**Remark:**
- Clearly, $\mathcal{M}^* \subset \mathcal{M} \Rightarrow \text{SLC} (\mathcal{M}^*) \leq \text{SLC} (\mathcal{M})$.
- If $\mathcal{M}$ is a class of (unconstrained) random processes such that $\text{SLC} (\mathcal{M}) = N < \infty$ then $p \in \mathcal{M}$ is uniquely determined by the values $p(v), |v| = N$.  

Dimension and String Length Complexity

**Theorem:** Let

\[ M_d := \{ p : \Sigma^* \to \mathbb{R} \mid \dim p \leq d \}. \]

Then

\[ SLC (M_d) = 2d - 1. \]
Dimension and String Length Complexity

**Theorem:** Let
\[ \mathcal{M}_d := \{ p : \Sigma^* \to \mathbb{R} \mid \dim p \leq d \}. \]

Then
\[ \text{SLC} (\mathcal{M}_d) = 2d - 1. \]

**Corollary:** A hidden Markov model on \( l \) hidden states is uniquely determined by the values
\[ p(v), \quad |v| = 2l - 1. \]
Finite String Length Complexity and Invariants

In the following:

\[ p^w : \Sigma^* \rightarrow \mathbb{R} \]

\[ v \mapsto p(wv) \]

is the column vector of the Hankel matrix \( \mathcal{P}_p \) for the string \( w \in \Sigma^* \).
Finite String Length Complexity and Invariants

In the following:

- \( p^w : \Sigma^* \to \mathbb{R} \)
  \[ v \mapsto p(wv) \]
  is the column vector of the Hankel matrix \( \mathcal{P}_p \) for the string \( w \in \Sigma^*. \)

- \( \mathcal{M}_n := \{(p(v))_{v \in \Sigma^n} \mid p \in \mathcal{M}\} \subset \mathbb{R}^{|\Sigma|^n} \)
  is the set of distributions over strings of length \( n \) induced by \( \mathcal{M} \).
Finite String Length Complexity and Invariants

Theorem: Let $\mathcal{M}$ be a class of unconstrained random processes such that

(i) 

$$\text{SLC} (\mathcal{M}) \leq n - 1 < \infty.$$
Finite String Length Complexity and Invariants

**Theorem:** Let $\mathcal{M}$ be a class of unconstrained random processes such that

(i) \[ \text{SLC} (\mathcal{M}) \leq n - 1 < \infty. \]

(ii) \[ p \in \mathcal{M} \Rightarrow \forall a \in \Sigma : \ p^a \in \mathcal{M}. \]
Finite String Length Complexity and Invariants

**Theorem:** Let $\mathcal{M}$ be a class of unconstrained random processes such that

(i) 

$$\text{SLC (}\mathcal{M}\text{)} \leq n - 1 < \infty.$$

(ii) 

$$p \in \mathcal{M} \Rightarrow \forall a \in \Sigma : p^a \in \mathcal{M}.$$

Then it holds that

$$(p(u), u \in \Sigma^{n+1}) \in \mathcal{M}_{n+1} \iff \begin{cases} (p(av), v \in \Sigma^n) \in \mathcal{M}_n \quad \forall a \in \Sigma \\ (p(v), v \in \Sigma^n) \in \mathcal{M}_n \end{cases}$$

where $p(v) = \sum_{a \in \Sigma} p(va)$. 
Degree of Hidden Markov Model Invariants

Let $\mathcal{M}^l$ be the class of hidden Markov processes on $l$ hidden states. Then it holds that

\[
SLC (\mathcal{M}^l) \leq 2l - 1
\]

\[
p \in \mathcal{M}^l \implies p^a \in \mathcal{M}^l, \forall a \in \Sigma.
\]
Degree of Hidden Markov Model Invariants

Let $\mathcal{M}^l$ be the class of hidden Markov processes on $l$ hidden states. Then it holds that

$$ SLC (\mathcal{M}^l) \leq 2l - 1 $$

$$ p \in \mathcal{M}^l \Rightarrow p^a \in \mathcal{M}^l, \forall a \in \Sigma. $$

**Conjecture:** Let $f_{n,l}$ be the hidden Markov model for strings of length $n$ on $l$ hidden states where $n \geq 2l$. Then it holds that

$$ d(n, l) \leq d(2l, l) $$

where $d(n, l)$ is defined to be the maximum degree of the generators of the ideal of invariants of $f_{n,l}$. 
Odds and Ends: Trace Algebras

**Definition:** A string function \( p : \Sigma^* \rightarrow \mathbb{R} \) is called traceable of order \( r \) if there are matrices \( X_a \in \mathbb{R}^{r \times r}, a \in \Sigma \) such that

\[
p(v = v_1...v_n) = \text{tr} \ X_{v_n}...X_{v_1}.
\]
Odds and Ends: Trace Algebras

**Definition:** A string function $p : \Sigma^* \rightarrow \mathbb{R}$ is called **traceable of order** $r$ if there are matrices $X_a \in \mathbb{R}^{r \times r}$, $a \in \Sigma$ such that

$$p(v = v_1...v_n) = \text{tr} X_{v_n}...X_{v_1}.$$

**Theorem:** Let $p \in \mathbb{R}^{\Sigma^*}$ be **traceable of order** $r$. Then

$$\dim p \leq r^2.$$
Thanks for the attention!