Mathematical Problems of Computer Science 32, 65–69, 2009.

On Optimal Identification of Markov Chain Distribution Subject to the Reliability Criterion

Evgueni A. Haroutunian† and Leader Navaei‡

†Institute for Informatics and Automation Problem of NAS of RA
ddag Payame Noor University, Iran
evhar@ipia.sci.am, ashkan_11380@yahoo.com

Abstract

In this paper the identification of distribution of simple homogeneous stationary Markov chain with a finite number of states is studied. The problem has been formulated by Ahlswede and Haroutunian on identification of hypotheses and solved for the case of the sequence of independent observations.

1. Introduction

Ahlswede and Haroutunian [1] (see also [2]) formulated a series of problems on multiple hypotheses testing for many objects and on identification of hypotheses under reliability requirement. The problem of many ($L > 2$) hypotheses testing on distributions of a finite state Markov chain is studied in [3] and also in [4] via large deviations techniques. In this paper we solve the problem of identification of distributions of many hypotheses for the case of Markov chain. In Section 2 we recall main definitions and results of [3], [4], [5], for many hypotheses testing and in Section 3 we present solution of the problem of identification of the Markov chain distribution.

2. Many Hypotheses LAO testing for Markov Chains

We remind the main definitions and results of paper [3]. Let $\mathbf{x} = (x_0, x_1, x_2, \ldots, x_N)$, $x_n \in \mathcal{X} = \{1, 2, \ldots, I\}$, $\mathbf{x} \in \mathcal{X}^{N+1}$, $N = 0, 1, 2, \ldots$, be the vector of observations of a simple homogeneous irreducible stationary Markov chain with finite number I of states. The L hypotheses H_l concern the matrix of the transition probabilities of the chain

$$P_l = \{P_l(j|i),\ i = \overline{1,I},\ j = \overline{1,I}\},\ \ l = \overline{1,L}.$$

The stationarity of the chain provides existence for each $l = \overline{1,L}$ of the unique stationary distribution $Q_l = \{Q_l(i),\ i = \overline{1,I}\}$, such that

$$\sum_i Q_l(i)P_l(j|i) = Q_l(j),\ \ \sum_i Q_l(i) = 1,\ \ i = \overline{1,I},\ \ j = \overline{1,I}.$$

The joint distributions of pairs $(i,j) \in I^2$ are

$$Q_l \circ P_l = \{Q_l(i)P_l(j|i), i=\overline{1,I}, j=\overline{1,I}\},\ l=\overline{1,L}.$$

We denote by $D(Q \circ P \| Q_l \circ P_l)$ the Kullback-Leibler divergence of a joint distribution $Q \circ P = \{Q(i)P(j|i), i=\overline{1,I}, j=\overline{1,I}\}$, from joint distribution $Q_l \circ P_l$,:

$$D(Q \circ P \| Q_l \circ P_l) = \sum_{i,j} Q(i)P(j|i)[\log Q(i)P(j|i) - \log Q_l(i)P_l(j|i)]$$

$$= D(Q\|Q_l) + D(Q \circ P \| Q \circ P_l),$$

where the divergence for stationary distributions is

$$D(Q\|Q_l) = \sum_i Q(i)[\log Q(i) - \log Q_l(i)],\ l=\overline{1,L}.$$

The second order type of Markov vector $\mathbf{x}$ is (see [3]) the square matrix of I^2 relative frequencies $\{N(i,j)N^{-1},\ \ i=\overline{1,I},\ j=\overline{1,I}\}$ of the simultaneous appearance in $\mathbf{x}$ of the states i and j on the pairs of neighbour places. It is clear that $\sum_{ij} N(i,j) = N$. Denote by $\mathcal{T}^N_{Q\circ P}$ the set of vectors $\mathbf{x}$ from $\mathcal{X}^{N+1}$ which have the second order type such that for some joint PD $Q \circ P$

$$N(i,j) = NQ(i)P(j|i),\ \ i=\overline{1,I},\ \ j=\overline{1,I}.$$

The set of joint PD $Q \circ P$ on I^2 we denote by $\mathcal{Q}\circ\mathcal{P}$. Non-randomized test $\phi_N(\mathbf{x})$ accepts one of the hypotheses H_l, $l=\overline{1,L}$ on the basis of the trajectory $\mathbf{x}=(x_0,x_1,\dots,x_N)$ of the $N+1$ observations. We write $\alpha^{(N)}_{l|m}(\phi_N)$ for the probability to erroneously decide the hypothesis H_l when H_m, $m \neq l$, is true. We denote by $\alpha^{(N)}_{m|m}(\phi_N)$ the probability to reject the correct hypothesis H_m, it means that

$$\alpha^{(N)}_{m|m}(\phi_N) = \sum_{l\neq m} \alpha^{(N)}_{l|m}(\phi_N),\ m=\overline{1,L}. \tag{1}$$

To every trajectory $\mathbf{x}$ the test ϕ_N puts in correspondence one from L hypotheses. The space $\mathcal{X}^{N+1}$ will be divided into L parts

$$\mathcal{G}^N_l = \{\mathbf{x},\ \phi_N(\mathbf{x}) = l\},\ l=\overline{1,L},$$

and

$$\alpha^{(N)}_{l|m}(\phi_N) = Q_m \circ P_m(\mathcal{G}^N_l),\ \ \ m,l=\overline{1,L}.$$

We study the matrix of "reliabilities" for infinite sequence of tests ϕ

$$E \triangleq \{E_{l|m}(\phi) \triangleq \overline{\lim_{N\to\infty}} - \frac{1}{N}\log \alpha^(N)_{l|m}(\phi_N),\ m,l=\overline{1,L}\}. \tag{2}$$

It can be proved from definitions (1) and (2) that

$$E_{m|m} = \min_{l\neq m} E_{l|m}. \tag{3}$$

Let P be a given matrix of transition probabilities of a Markov chain and Q be the corresponding stationary PD. For given family of positive numbers $E_{1|1}, E_{2|2}, \dots, E_{L-1|L-1}$, let us define the decision rule ϕ^* by the sets of distributions

$$\mathcal{R}_l \triangleq \{Q \circ P : D(Q \circ P \| Q \circ P_l) \leq E_{l|l},\ D(Q\|Q_l) < \infty\},\ l = \overline{1, L-1}, \quad (4.a)$$

$$\mathcal{R}_L \triangleq \{Q \circ P : D(Q \circ P \| Q \circ P_l) > E_{l|l},\ l = \overline{1, L-1}\}, \quad (4.b)$$

and the functions:

$$E^*_{l|l}(E_{l|l}) \triangleq E_{l|l},\ l = \overline{1, L-1}, \quad (5.a)$$

$$E^*_{l|m}(E_{l|l}) \triangleq \inf_{Q \circ P \in \mathcal{R}_l} D(Q \circ P \| Q \circ P_m),\ m = \overline{1, L},\ l \neq m,\ l = \overline{1, L-1}, \quad (5.b)$$

$$E^*_{L|m}(E_{1|1}, \ldots, E_{L-1|L-1}) \triangleq \inf_{Q \circ P \in \mathcal{R}_L} D(Q \circ P \| Q \circ P_m),\ m = \overline{1, L-1}, \quad (5.c)$$

$$E^*_{L|L}(E_{1|1}, \ldots, E_{L-1|L-1}) \triangleq \min_{l = \overline{1, L-1}} E^*_{l|L}. \quad (5.d)$$

The main result of papers [3] and [4] is:

Theorem 1: Let $\mathcal{X}$ be a finite set, for a family of distinct distributions $P_1, \cdots, P_L$ the following two statements hold. If the positive finite numbers $E_{1|1}, \cdots, E_{L-1|L-1}$ satisfy conditions

$$0 < E_{1|1} < \min[D(Q_m \circ P_m \| Q_m \circ P_1),\ m = \overline{2, L}],$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

$$0 < E_{l|l} < \min[E^*_{l|m}(E_{m|m}),\ m = \overline{1, l-1}, D(Q_m \circ P_m \| Q_m \circ P_l),\ m = \overline{l+1, L}],\ l = \overline{2, L-1},$$

then:

a) there exists a LAO sequence of tests ϕ^*, the reliability matrix of which $E^* = \left\{E^*_{l|m}(\phi^*)\right\}$ is defined in (5), and all elements of it are strictly positive,

b) even if one of conditions (6) is violated, then the reliability matrix of an arbitrary test necessarily has an element equal to zero, the corresponding error probability does not tend exponentially to zero.

3. Problem Statement and Formulation of Result on Reliable Identification of Distribution

Assume that there are $L \geq 2$ hypothetical distributions. Now the question is whether r-th distribution occurred, or not. There are two error probabilities for each $r = \overline{1, L}$, the probability $\alpha^{(N)}_{l \neq r|m=r}$ to accept l different from r, when r is in reality, and the probability $\alpha^{(N)}_{l=r|m \neq r}$ that r is accepted, when it is not correct. The probability $\alpha^{(N)}_{l \neq r|m=r}$ is known (1), it is the probability $\alpha^{(N)}_{r|r}$. The reliability $E_{l \neq r|m=r}$ coincides with $E_{r|r}$, meeting (3). Our aim is to find optimal interdependence between $E_{l=r|m \neq r}$ and $E_{l \neq r|m=r}$. The latter can have values satisfying conditions (6), or now the conditions:

$$0 < E_{r|r} < \min_{l: l \neq r}[D(Q_l \circ P_l \| Q_l \circ P_r),\ r = \overline{1, L}.$$

We shall use the probabilities of hypotheses. Let us assume that the hypotheses H_l, $l = \overline{1, L}$, have positive probabilities $\mathbf{Pr}(l)$, $l = \overline{1, L}$. We shall see, that the formulated below result does not depend on values of $\mathbf{Pr}(l)$, $l = \overline{1, L}$, if they all are strictly positive,(naturally we can disregard hypotheses with probability equal to zero). We can make the following calculations for $r = \overline{1, L}$:

$$\alpha^{(N)}_{l=r|m\neq r} = \frac{\Pr^{(N)}(l = r, m \neq r)}{\Pr(m \neq r)} = \frac{1}{\sum\limits_{m:m\neq r} \Pr(m)} \sum_{m:m\neq r} \Pr(m)\alpha^{(N)}_{r|m},$$

and them for each $r = \overline{1, L}$, we can write:

$$E_{l=r|m\neq r} = \overline{\lim_{N\to\infty}}\left(-\frac{1}{N}\log \alpha^{(N)}_{l=r|m\neq r}\right) =$$

$$= \lim_{N\to\infty} \frac{1}{N}(\log \sum_{m:m\neq r} \Pr(m) - \log \sum_{m:m\neq r} \Pr(m)\alpha^{(N)}_{r|m}) = \min_{m:m\neq r} E^{*}_{r|m} \tag{7}$$

Using (7) by analogy with Theorem 1 we conclude (with $\mathcal{R}_r$ as in (4) for each r including $r = L$ by the values of $E_{r|r}$ from $(0, \min D(Q_l \circ P_l \| Q_l \circ P_r)))$, that

$$E_{l=r|m\neq r}(E_{r|r}) = \min_{m:m\neq r} \inf_{Q\circ P\in\mathcal{R}_r} D(Q \circ P \| Q \circ P_m), \quad r = \overline{1, L}. \tag{8}$$

So we can formulate this result in

Theorem 2: For the model with distinct distributions for the given sample **x** we can find the type $Q \circ P$, of it and when $Q \circ P \in \mathcal{R}_r$, we accept the hypotheses r. Under condition that the probabilities of all L hypotheses are positive the reliability $E_{l=r|m\neq r}$ of such test as function of given $E_{l\neq r|m=r}$ may be calculated by (8).

References

[1] R. F. Ahlswede and E. A. Haroutunian, "On logarithmically asymptotically optimal testing of hypotheses and identification", *Lecture Notes in Computer Science*, vol. 4123. *"General Theory of Information Transfer and Combinatorics" Springer*, pp. 462-478, 2006.

[2] E. A. Haroutunian, "Reliability in multiple testing and identification problems", *NATO Science Series III: Computer and Systems Sciences*, vol. 198, pp. 189–201, IOS Oress, 2005.

[3] E. A. Haroutunian, "On asymptotically optimal testing of hypotheses concerning Markov chain", *(in Russian). Izvestia Acad. Nauk Armenian SSR. Seria Mathem.* vol. 22, no. 1. pp. 76-80, 1988.

[4] L Navaei, "Application of LDT to many hypotheses optimal testing for Markov chain", *Mathematical Problems of Computer Science*, vol. 31, pp.73-78, 2008.

[5] E. A. Haroutunian, M. E. Haroutunian and A. N. Harutyunyan, " Reliability criteria in information theory and in statistical hypothesis testing", *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 2-3, 2008.

Մարկովի շղթայի բաշխման հուսալիության պայմանով օպտիմալ նույնականացման մասին

Ե. Հարությունյան, Լ. Նավայի

Ամփոփում

ՈՒսումնասիրված է վերջավոր թվով վիճակներով համասեռ ստացիոնար Մարկովյան շղթայի վիճակագրական նույնականացումը։ Վարկածների նույնականացման հիմնախնդիրը ձևակերպվել է Ալսվեդեի և Հարությունյանի կողմից և լուծվել անկախ դիտարկումների դեպքի համար։