KYBERNETIKA — VOLUME 44 (2008), NUMBER 1, PAGES 53 – 60

STOCHASTIC CONTROL OPTIMAL IN THE KULLBACK SENSE

Jan Šindelář, Igor Vajda and Miroslav Kárný

The paper solves the problem of minimization of the Kullback divergence between a partially known and a completely known probability distribution. It considers two probability distributions of a random vector $(u_1, x_1, \ldots, u_T, x_T)$ on a sample space of $2T$ dimensions. One of the distributions is known, the other is known only partially. Namely, only the conditional probability distributions of x_τ given $u_1, x_1, \ldots, u_{\tau-1}, x_{\tau-1}, u_\tau$ are known for $\tau = 1, \ldots, T$. Our objective is to determine the remaining conditional probability distributions of u_τ given $u_1, x_1, \ldots, u_{\tau-1}, x_{\tau-1}$ such that the Kullback divergence of the partially known distribution with respect to the completely known distribution is minimal. Explicit solution of this problem has been found previously for Markovian systems in Karný [6]. The general solution is given in this paper.

Keywords: Kullback divergence, minimization, stochastic controller

AMS Subject Classification: 60G35, 93E20, 90D60, 49N35

1. INTRODUCTION

The present paper is devoted to the minimization of Kullback divergence between a partially known and a completely known probability distribution. It is motivated by the formulation of the optimal stochastic control problem stated in Karný [6].

The optimality is based on the concept of Kullback divergence [8, 9]. It is widely used in estimation, approximation, filtering and control problems (see [6, 7]), where it plays the role of information-theoretic "distance" between different probability distributions of data.

Two probability distributions are considered, say $P(u_1, x_1, \ldots, u_T, x_T)$ and $Q(u_1, x_1, \ldots, u_T, x_T)$, both defined on a space with even number of dimensions. The distribution Q as well as conditional probability distributions $P(x_\tau \,|\, u_1, x_1, \ldots \ldots, u_{\tau-1}, x_{\tau-1}, u_\tau)$, $\tau = 1, \ldots, T$ are known. The distribution $P(u_1, x_1, \ldots, x_T, u_T)$ is uniquely specified by the remaining conditional probability distributions $P(u_\tau \,|\, u_1, x_1, \ldots, u_{\tau-1}, x_{\tau-1})$, $\tau = 1, \ldots, T$. *Our goal* is to specify these remaining conditional distributions in such a way that the Kullback divergence of P with respect to Q is minimal.

Optimal stochastic control is a well developed field (see e. g. [2, 3, 4]) with numerous applications like control of technological processes and control of economic or biological systems. Real time implementation of the control algorithms requires explicit numerically feasible solutions. The class of problems meeting this requirement consists essentially of

- linear Gaussian systems controlled via minimization of the expected value of a quadratic loss function (see e. g. [1, 3, 13]);
- controlled Markov chains with a finite number of states (see e. g. [3, 10, 11, 12]).

An alternative formulation of the control design has been proposed in [6]. Loosely speaking, optional parts of the probabilistic description of the closed loop are selected so that the Kullback divergence [8, 9] of this description with respect to some ideal description is minimized. The conditional probability distributions $P(u_\tau \,|\, u_1, x_1, \ldots \ldots, u_{\tau-1}, x_{\tau-1})$ define controllers controlling a process represented by the conditional probability distributions $P(x_\tau \,|\, u_1, x_1, \ldots, u_{\tau-1}, x_{\tau-1}, u_\tau)$, $\tau = 1, \ldots, T$. The probability distribution $Q(u_1, x_1, \ldots, u_T, x_T)$ describes an ideal course of the controlled process. The aim is to minimize a distance from the controlled process to its ideal description.

An explicit solution of this optimization problem has been derived for Markovian systems in [6]. Here the general case is solved.

The studied general systems and related mathematical concepts and notations are introduced in Section 2. The minimization problem solved in the paper is rigorously stated in Section 3. Section 4 is devoted to the solution of this problem.

2. MODEL AND NOTATION

Let for all $1 \le \tau \le T$, $(\mathcal{X}_\tau, \mathcal{A}_\tau, \mu_\tau)$ be σ-finite state measure spaces and $(\mathcal{U}_\tau, \mathcal{S}_\tau, \nu_\tau)$ σ-finite control measure spaces. The controls $u_\tau \in \mathcal{U}_\tau$ lead to states $x_\tau \in \mathcal{X}_\tau$ stochastically depending on u_τ and also on the previous history

$$z_{\tau-1} = (y_1, \ldots, y_{\tau-1}) \tag{1}$$

where here and in the sequel,

$$y_\tau \equiv (x_\tau, u_\tau) \in \mathcal{Y}_\tau \equiv (\mathcal{X}_\tau \otimes \mathcal{U}_\tau), \qquad 1 \le \tau \le T. \tag{2}$$

The interplay between controls and states in the time horizon $1 \le \tau \le T$ is thus described by a probability measure Π on $(\mathcal{Z}_T, \mathcal{C}_T)$. Here and in the sequel, for every $1 \le t \le T$ we use the notation

$$(\mathcal{Z}_t, \mathcal{C}_t) \;=\; \otimes_{\tau=1}^{t} (\mathcal{Y}_\tau, \mathcal{B}_\tau), \qquad \mathcal{B}_\tau = \mathcal{A}_\tau \otimes \mathcal{S}_\tau. \tag{3}$$

Let $\mathcal{P}_T$ be the class of all probability measures Π on $(\mathcal{Z}_T, \mathcal{C}_T)$ dominated in the sense

$$\Pi \ll \lambda^T \tag{4}$$

where we put for every $1 \le t \le T$

$$\lambda^t = \mu^t \otimes \nu^t, \;\; \mu^t = \otimes_{\tau=1}^{t} \mu_\tau, \;\; \nu^t = \otimes_{\tau=1}^{t} \nu_\tau. \tag{5}$$

All $\Pi \in \mathcal{P}_T$ are λ^T a. s. uniquely represented by the densities

$$\pi = \frac{\mathrm{d}\Pi}{\mathrm{d}\lambda^T}, \; \pi = \pi(z_T), \; z_T \in \mathcal{Z}_T. \tag{6}$$

We assume more, namely that the densities π of the measures $\Pi \in \mathcal{P}_T$ can be decomposed into products of regular conditional densities as follows

$$\pi\,(z_T) \equiv \pi\,(y_1,\,\dots,\,y_T) = \prod\nolimits_{\tau=1}^{T} \pi\,(y_\tau|z_{\tau-1}) \tag{7}$$

where

$$\pi\,(y_1|z_0) \equiv \pi\,(y_1) = \pi\,(x_1,\,u_1) \tag{8}$$

and that, moreover, for every $1 \le \tau \le T$ takes place the decomposition

$$\begin{aligned} \pi\,(y_\tau|z_{\tau-1}) &\equiv \pi\,(x_\tau,\,u_\tau|z_{\tau-1}) \\ &= \pi\,(u_\tau|z_{\tau-1})\,\pi\,(x_\tau,\,u_\tau|z_{\tau-1})\,. \end{aligned} \tag{9}$$

Here in accordance with (8) it is assumed that

$$\pi\,(x_1,\,u_1|z_0) \equiv \pi\,(x_1,\,u_1) = \pi\,(u_1)\,\pi\,(x_1|u_1)\,. \tag{10}$$

An important *convention* adopted in (6)−(10), as well as everywhere in the sequel, is that the arguments of the densities and conditional densities specify the corresponding unconditional and conditional densities. This convention is common in the literature on information theory, see e. g. [5].

It follows from (7)−(10) that for every $1 \le \tau \le T$ the functions

$$\begin{aligned} \pi\,(z_t) &\equiv \pi\,(y_1,\,\dots,\,y_t) = \prod\nolimits_{\tau=1}^{t} \pi\,(y_\tau|z_{\tau-1}) \\ &= \prod\nolimits_{\tau=1}^{t} \pi\,(u_\tau|z_{\tau-1})\,\pi\,(x_\tau|u_\tau,\,z_{\tau-1}) \end{aligned} \tag{11}$$

are probability densities on the σ-finite measure spaces

$$(\mathcal{Z}_\tau,\,\mathcal{C}_\tau,\,\lambda^\tau)\,. \tag{12}$$

These densities represent interplay between the controls $(u_1,\,\dots,\,u_t)$ and states $(x_1,\,\dots,\,x_t)$ provided that the controls are governed by the *stochastic control rules*

$$\mathbf{\Pi}^{(c)} = (\pi\,(u_\tau|z_{\tau-1}) \;:\; 1 \le \tau \le T) \tag{13}$$

(cf. (10)) and that the controlled system reacts to these controls by the *system dynamics rules*

$$\mathbf{\Pi}^{(s)} = (\pi\,(x_\tau|u_\tau,\,z_{\tau-1}) \;:\; 1 \le \tau \le T) \tag{14}$$

(cf. (10) again).

3. THE PROBLEM

Let $Q \in \mathcal{P}_T$ be a probability measure characterized for every $1 \leq t \leq T$ by the densities

$$q(z_t) = \prod_{\tau=1}^{t} q(u_\tau|z_{\tau-1})\, q(x_\tau|u_\tau,\, z_{\tau-1}) \tag{15}$$

(cf. (11)) corresponding to the system with a given dynamics

$$\mathbf{Q}^{(s)} = (q(x_\tau|u_\tau,\, z_{\tau-1}) \;:\; 1 \leq \tau \leq T) \tag{16}$$

(cf. (13)) and to a given stochastic control

$$\mathbf{Q}^{(c)} = (q(u_\tau|z_{\tau-1}) \;:\; 1 \leq \tau \leq T) \tag{17}$$

(cf. (14)). Further, let $\mathcal{P}_{0,T} \subseteq \mathcal{P}_T$ be a subset of probability measures $P \in \mathcal{P}_T$ characterized for every $1 \leq t \leq T$ by the densities

$$p(z_t) = \prod_{\tau=1}^{t} p(u_\tau|z_{\tau-1})\, p(x_\tau|u_\tau,\, z_{\tau-1}) \tag{18}$$

(cf. (11)) corresponding to a given system dynamics

$$\mathbf{P}^{(s)} = (p(x_\tau|u_\tau,\, z_{\tau-1}) \;:\; 1 \leq \tau \leq T) \tag{19}$$

(cf. (13)) which may be in general different from (16) and to a stochastic control from a given class

$$\mathfrak{P}^{(c)} = \left\{ \mathbf{P}^{(c)} = (p(u_\tau|z_{\tau-1}) \;:\; 1 \leq \tau \leq T) \right\} \tag{20}$$

(cf. (14)) of admissible stochastic controls.

In this paper we solve the problem how to find the probability measure $P_0 \in \mathcal{P}_{0,T}$ characterized for every $1 \leq t \leq T$ by the densities

$$p_0(z_t) = \prod_{\tau=1}^{t} p_0(u_\tau|z_{\tau-1})\, p(x_\tau|u_\tau,\, z_{\tau-1}) \tag{21}$$

corresponding to the given system dynamics (19) and to the desirable stochastic controls

$$\mathbf{P}_0^{(c)} = (p_0(u_\tau|z_{\tau-1}) \;:\; 1 \leq \tau \leq T) \;\in\; \mathfrak{P}^{(c)} \tag{22}$$

(cf. (20)) which are optimal in the sense

$$P_0 = \operatorname{argmin}_{P \in \mathcal{P}_{0,T}} \mathcal{K}(P||Q) \tag{23}$$

where

$$\mathcal{K}(P||Q) = \int_{\mathcal{Z}_T} p(z_T) \ln \frac{p(z_T)}{q(z_T)}\, \mathrm{d}\lambda^T(z_T) \tag{24}$$

is the Kullback divergence of P and Q with the usual convention

$$p \ln \frac{p}{q} \;:=\; \begin{cases} 0 & \text{if} \quad p = 0,\; q \geq 0\,, \\ \infty & \text{if} \quad p > 0,\; q = 0 \end{cases} \tag{25}$$

behind the integral. The problem is thus to find the optimal admissible controls (22).

The solution P_0 of the minimization problem (23) may not exist, e.g. Q may be a boundary point of $\mathcal{P}_{0,\,T}$ not contained in $\mathcal{P}_{0,\,T}$ so that

$$\inf_{P \in \mathcal{P}_{0,\,T}} \mathcal{K}\left(P||Q\right) = 0$$

while $\mathcal{K}\left(P||Q\right) > 0$ for all $P \in \mathcal{P}_{0,\,T}$ due to the assumption $Q \notin \mathcal{P}_{0,\,T}$. Sufficient conditions for the existence as well as the explicit construction rule are given in the next section.

4. SOLUTION

Throughout this section we put $\mathcal{Z}_0 = \{z_0\}$ and for every $(x,\, u,\, z) \in \mathcal{X}_\tau \otimes \mathcal{U}_\tau \otimes \mathcal{Z}_{\tau-1}$ we introduce the symbols

$$\begin{aligned} p_z(u) &\triangleq p(u|z),\ \ p_{uz}(x) \triangleq p(x|u,\, z) \\ q_z(u) &\triangleq q(u|z),\ \ q_{uz}(x) \triangleq q(x|u,\, z) \end{aligned} \tag{26}$$

if $1 \le \tau \le T$ and, in accordance with (10),

$$\begin{aligned} p_{z_0}(u) &\triangleq p(u),\ \ p_{uz_0}(x) = p_u(x) \triangleq p(x|u) \\ q_{z_0}(u) &\triangleq q(u),\ \ q_{uz_0}(x) = q_u(x) \triangleq q(x|u) \end{aligned} \tag{27}$$

if $\tau = 1$ where on the right-hand sides are the above introduced densities w. r. t. μ_τ, ν_τ. Further, we assume that the domination relation

$$\mathcal{P}_{0,\,T} \ll Q \tag{28}$$

holds which means that the densities (26), (27) as well as the densities (15), (18) satisfy the relation

$$q = 0 \Rightarrow p = 0. \tag{29}$$

Therefore, the Kullback divergences of the densities $p,\, q$ in (15), (18) are for all $1 \le \tau \le T$ well defined by the formula

$$\mathcal{K}_\tau(p||q) = \int_{\mathcal{Z}_\tau} p\left(z_\tau\right) \ln \frac{p\left(z_\tau\right)}{q\left(z_\tau\right)}\, \mathrm{d}\lambda^\tau\left(z_\tau\right) \tag{30}$$

and the same divergences of the conditional densities (26), (27) are well defined by the formulas

$$\mathcal{K}_\tau(p_{uz}||q_{uz}) = \int_{\mathcal{X}} p_{uz}\left(x\right) \ln \frac{p_{uz}\left(x\right)}{q_{uz}\left(x\right)}\, \mathrm{d}\mu_\tau\left(x\right) \tag{31}$$

(cf. [14]). By (11), the density π of any $\mathit{\Pi} \in \mathcal{P}_T$ satisfies for

$$z_\tau = (x,\, y,\, z) \in \mathcal{Z}_\tau = \mathcal{X}_\tau \otimes \mathcal{U}_\tau \otimes \mathcal{Z}_{\tau-1} \tag{32}$$

the relation

$$\pi\,(z_\tau) = \pi(u|z)\pi(x|u,\, z)\pi(z). \tag{33}$$

Consequently,

$$\mathcal{K}_\tau(p||q) = \int\limits_{\mathcal{Z}_{\tau-1}} p(z)\left[\int_{\mathcal{U}_\tau} p_z(u)\left(\mathcal{K}\,(p_{uz}||q_{uz}) + \ln\frac{p_z(u)}{q_z(u)}\right)\mathrm{d}\nu_\tau(u) + \ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-1}(z). \tag{34}$$

Let us define and/or suppose for all $(u,\, z) \in \mathcal{U}_\tau \otimes \mathcal{Z}_{\tau-1}$ and for all $1 \le \tau \le T$

$$\alpha_\tau(u,\, z) \;\triangleq\; \varepsilon_\tau(u,\, z) + \mathcal{K}\,(p_{uz}||q_{uz})\,, \tag{35}$$

$$c_\tau(z) \;\triangleq\; \int_{\mathcal{U}_\tau} q_z(u) e^{-\alpha_\tau(u,\, z)}\,\mathrm{d}\nu_\tau(u) \in (0,\, \infty), \tag{36}$$

$$\varepsilon_\tau(u,\, z) \;\triangleq\; -\int_{\mathcal{X}_\tau} \ln c_{\tau+1}(x,\, u,\, z) p_{uz}(x)\,\mathrm{d}\mu_\tau(x) \tag{37}$$

where $c_{T+1}(x,\, u,\, z) \triangleq 1$ and the expressions containing $z_0 \in \mathcal{Z}_0$ are precised in the sense of (27), i. e. we put

$$\alpha_1(u,\, z_0) \;=\; \alpha_1(u) \triangleq \varepsilon_1(u) + \mathcal{K}\,(p_u||q_u)\,, \tag{38}$$

$$c_1(z_0) \;=\; c_1 \triangleq \int_{\mathcal{U}_1} q(u) e^{-\alpha_1(u)}\,\mathrm{d}\nu_1(u) \in (0,\, \infty), \tag{39}$$

$$\varepsilon_1(u,\, z_0) \;=\; \varepsilon_1(u) \triangleq -\int_{\mathcal{X}_1} \ln c_2(x,\, u) p_u(x)\,\mathrm{d}\mu_1(x). \tag{40}$$

Theorem 4.1. Let for every $1 \le \tau \le T$ the conditions (36) hold and let us consider the system of probability densities $\mathbf{P}_0^{(c)}$ of the form (22) defined for every $(u,\, z) \in \mathcal{U}_\tau \otimes \mathcal{Z}_{\tau-1}$ by the formulas

$$p_0(u|z) = p_{0,\,z}(u) \triangleq \frac{q_z(u) e^{-\alpha_\tau(u,\, z)}}{c_\tau(z)}, \qquad 1 \le \tau \le T \tag{41}$$

where α_τ, c_τ are given by (35)–(40). If this system belongs to $\mathfrak{P}^{(c)}$, then the corresponding unconditional probability measure P_0 given by the formula (21) satisfies *(23)*, that is, it solves the problem of Section 3. The minimum achieved in (23) is

$$\mathcal{K}\,(P_0||Q) = -\ln c_1$$

for c_1 given by (39).

P r o o f. Fix any $1 \le \tau \le T$ and put

$$\mathcal{E}_\tau = \int_{\mathcal{Z}_{\tau-1}} p(z)\left[\int_{\mathcal{U}_\tau} \varepsilon_\tau(u,\, z) p_z(u)\,\mathrm{d}\nu_\tau(u)\right]\mathrm{d}\lambda^{\tau-1}(z). \tag{42}$$

We shall minimize over the probability densities $p_z(\cdot) = p(\cdot|z)$, $z \in \mathcal{Z}_{\tau-1}$ on $\mathcal{U}_\tau$ the functional

$$\mathcal{F}_\tau\,(p||q) = \mathcal{K}_\tau\,(p||q) \;+\; \mathcal{E}_\tau$$

for $\mathcal{K}_\tau\left(p||q\right)$ given by (34). By (34), (42) and (35), (41), $\mathcal{F}_\tau\left(p||q\right)$ can be written as the integrals

$$\begin{aligned}
&\int_{\mathcal{Z}_{\tau-1}} p(z)\left[\int_{\mathcal{U}_\tau} p_z(u)\left(\alpha_\tau(u,\, z)+\ln\frac{p_z(u)}{q_z(u)}\right)\mathrm{d}\nu_\tau(u)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-1}(z)\\
=&\int_{\mathcal{Z}_{\tau-1}} p(z)\left[\int_{\mathcal{U}_\tau} p_z(u)\left(-\ln c_\tau(z)+\ln\frac{p_z(u)}{p_{0,\,z}(u)}\right)\mathrm{d}\nu_\tau(u)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-1}(z)\\
=&\int_{\mathcal{Z}_{\tau-1}} p(z)\left[-\ln c_\tau(z)+\mathcal{K}\left(p_z||p_{0,\,z}\right)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-1}(z).
\end{aligned}$$

Therefore

$$\operatorname{argmin}_{p_z}\mathcal{F}_\tau\left(p||q\right)=p_{0,\,z}$$

(cf. (41)) and

$$\begin{aligned}
&\min_{p_z}\mathcal{F}_\tau\left(p||q\right)\\
&=\int_{\mathcal{Z}_{\tau-1}} p(z)\left[-\ln c_\tau(z)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-1}(z)\\
&=\int_{\mathcal{Z}_{\tau-2}} p(z)\left[\int_{\mathcal{U}_{\tau-1}} p_z(u)\left(\alpha_{\tau-1}(u,\, z)+\ln\frac{p_z(u)}{q_z(u)}\right)\mathrm{d}\nu_{\tau-1}(u)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^{\tau-2}(z)
\end{aligned}$$

(cf. (35) – (37)) for $2\le\tau\le T$ and

$$\begin{aligned}
\min_{p_z}\mathcal{F}_2\left(p||q\right) &= \int_{\mathcal{Z}_1} p(z)\left[-\ln c_2(z)+\ln\frac{p(z)}{q(z)}\right]\mathrm{d}\lambda^1(z)\\
&= \int_{\mathcal{U}_1} p(u)\left(\varepsilon_1(u)+\mathcal{K}\left(p_u||q_u\right)+\ln\frac{p(u)}{q(u)}\right)\mathrm{d}\nu_1(u)\\
&= \int_{\mathcal{U}_1} p(u)\left(\alpha_1(u)+\ln\frac{p(u)}{q(u)}\right)\mathrm{d}\nu_1(u)\\
&= -\ln c_1+\int_{\mathcal{U}_1} p(u)\ln\frac{p(u)}{p_0(u)}\,\mathrm{d}\nu_1(u)
\end{aligned}$$

(cf. (38) – (40)) for $\tau=2$. The stated result thus follows by induction for $\tau = T,\, T-1,\, \ldots,\, 2$. □

5. CONCLUSIONS

The present paper addresses the problem of minimization of Kullback divergence of a partially known with respect to a completely known probability distributions. Our effort was motivated by a probabilistic formulation of the optimal stochastic control problem.

The paper presents a rigorous formulation of the problem and proves a theorem that gives explicit solution of the corresponding optimization problem. The result obtained is of general type and shows that dynamic optimization under uncertainty can be solved explicitly.

The result obtained may have a significant theoretical and practical impact on the probabilistic solution of optimal stochastic control problem. The further research is to be focused on the detailed algorithmic elaboration and extensive testing of the found solution.

ACKNOWLEDGEMENT

This work has been partially supported by the Ministry of Education, Youth and Sports of the Czech Republic under grant 2C06001 and by the Grant Agency of the Academy of Sciences of the Czech Republic under grants 1ET 100 750401 and IAA 200 750802.

(Received April 30, 2006.)

REFERENCES

[1] M. Aoki: Optimization of Stochastic Systems: Topics in Discrete-Time Systems. Academic Press, New York – London 1967.
[2] K. J. Åström: Introduction to Stochastic Control. Academic Press, New York – San Francisco – London 1970.
[3] D. P. Bertsekas: Dynamic Programming and Stochastic Control. Second edition. Athena Scientific, Belmont, Mass. 2000.
[4] D. Clark: Advances in Model-Based Predictive Control. Oxford University Press, Oxford 1994.
[5] T. M. Cover and J. A. Thomas: Elements of Information Theory. Second edition. Wiley-Interscience, New York 2006.
[6] M. Kárný: Towards fully probabilistic control design. Automatica *32* (1996), 12, 1719–1722.
[7] R. Kulhavý: A Kullback–Leibler distance approach to system identification. In: Preprints of the IFAC Symposium on Adaptive Systems in Control and Signal Processing (C. Bányász, ed.), Budapest 1995, pp. 55–66.
[8] S. Kullback: Information Theory and Statistics. Wiley, New York and Chapman & Hall, London 1967.
[9] S. Kullback and R. Leibler: On information and sufficiency. Ann. Math. Statist. *22* (1951), 79–87.
[10] P. R. Kumar and P. Varaiya: Stochastic Systems: Estimation, Identification and Adaptive Control. Prentice Hall, Englewood Cliffs, N. J. 1986.
[11] H. Kushner: Introduction to Stochastic Control. Holt, Rinehard and Winston, New York 1971.
[12] J. J. Martin: Bayesian Decision Problems and Markov Chains. Wiley, New York 1967.
[13] J. S. Meditch: Stochastic Optimal Linear Estimation and Control. Mc. Graw Hill, New York 1969.
[14] I. Vajda: Theory of Statistical Inference and Information. Mathematical and statistical methods. Kluwer Academic Publishers, Dordrecht, Boston – London 1989.

Jan Šindelář, Igor Vajda and Miroslav Kárný, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mails: sindelar@utia.cas.cz, vajda@utia.cas.cz, school@utia.cas.cz