Abstract

The paper deals with an approach to the formulation of the decision-making tasks via such specifying constituents as system, experience, ignorance, admissible strategy, loss function and decision. The paper describes how to construct the offered specification and also shows how this way of the material organizing can be applied for teaching Bayesian decision-making methodology, namely, in e-learning. To demonstrate the approach the fully probabilistic control design in its state-space setting has been taken as an example. The fully probabilistic design describes the controlled closed loop system and its desired behavior in probabilistic terms and uses the Kullback-Leibler divergence as their proximity measure. Because of the state-space setting both solutions of filtering and regulation tasks are shown. The practical example is demonstrated on normal state-space model and normal ideal pdf. In this case the solution of the fully probabilistic design reduces to Kalman filtering and design minimizing the quadratic criterion. In the end the paper shows an implementation of the approach.

Keywords: fully probabilistic design; state-space model; optimal control strategy; e-learning

1   INTRODUCTION

Bayesian decision making plays a significant role in the world of statistics. It is a powerful statistical approach with a wide area of application [1, 2]. A lot of books and articles on Bayesian decision making are available [3, 4, 5].

The fundamental tasks of Bayesian decision making can be conditionally divided among estimation (point and set), testing of hypotheses, prediction and control. These tasks, both static and dynamic, can be specified by six following constituents: *system, experience, ignorance, admissible strategy, loss function* and *decision*. These constituents compose a specification of the decision-making (DM) task and are, in essence, the concepts, which a decision maker should define when formulating a task. They are not always so obvious as it may seem.

In order to determine them a decision maker should know the definitions of the concepts [6]:

- *System* is a part of the world that is of interest for a decision maker who should either describe or influence it.

- *Experience* is knowledge about the system, available to the decision maker for the selection of the particular decision.

- *Ignorance* concerns knowledge about the system, unavailable to the decision maker.

- *Admissible strategy* is a mapping from experience to actions, that meets physical constraints, implied by the system and the actuator.

- *Loss function* quantifies the degree of achievement of the DM aim by assigning to each realization of behavior a number on real line.

- *Decision* is the value of a quantity that can be directly chosen by the decision maker for reaching decision maker's aims.

To demonstrate a specification of the DM task, let's make a simple example from everyday life. Let a decision maker be an usual student. Assume that a student has the following aim: to become a good specialist, who has a good knowledge and excellent diploma. It can be expressed in the mentioned terms by the following way.

Aim: to minimize a divergence between the knowledge of the student and knowledge of a good specialist.

- *System:* is the student, who has the lectures, study material and teachers as the inputs and exams and marks (in principle, the student's knowledge) as the outputs;

- *Experience:* past marks, mapping a student's level of knowledge; student's abilities;

- *Ignorance:* such factor as the exam, which can include the psychological state of student's health, a teacher's behavior, transport problems;

- *Admissible strategy:* time and energy of the student that can be devoted to study, restricted by schedule, number of exams and subjects, other duties;

- *Loss function:* the quality of knowledge expressed in marks; time spent to study;

- *Decision:* how much time and energy the student will devote to study, taking into account all the restrictions.

This example from life is very simple and, probably, it does not stand up to criticism. But the principle of composing the specification is clear now.

What this paper tries is to show how the DM tasks could be specified via these constituents and, simultaneously, how it could be applied for educational purposes in a field of Bayesian DM methodology. The fully probabilistic design (FPD) [8], which can be interpreted as a superstructure of Bayesian DM on the whole and under the corresponding setting can include majority of DM tasks, is a good example to demonstrate the approach. The example of practical solution is shown here on Gaussian models. Whether such formulation of DM tasks was effective, a reader could judge having read this paper.

2   FULLY PROBABILISTIC DESIGN: THEORY AND PRACTICAL EXAMPLE

2.1   Preliminaries

For a start, the notations used through the paper should be pointed out: $\equiv$ is equality by definition; X^* denotes a set of X-values; $\mathring{X}$ means cardinality of a finite set X^*; $f(\cdot|\cdot)$ denotes probability density function (pdf) that is assumed to exist; t labels discrete-time moments, $t \in t^* \equiv \{1,\dots,\mathring{t}\}$; $\mathring{t} < \infty$ is a given control horizon; $d_t = (y_t, u_t)$ is the data record at time t consisting of an observed system output y_t and of an optional system input u_t; x_t is an unobserved system state; d, x are assumed to be finite-dimensional; $X(t)$ denotes the sequence $(X_1, \dots, X_t)$, $X(t) \in \{y(t), u(t), x(t)\}$.

The FPD uses the Kullback-Leibler (KL) divergence [9]

$$\mathcal{D}\left(f||\tilde{f}\right) \equiv \int f(X) \ln\left(\frac{f(X)}{\tilde{f}(X)}\right) dX \qquad (1)$$

to measure the proximity of a pair of pdfs $f, \tilde{f}$ acting on a set X^*. The key property of KL divergence is

$$\mathcal{D}(f||\tilde{f}) \geq 0, \ \mathcal{D}(f||\tilde{f}) = 0 \text{ iff } f = \tilde{f} \text{ almost everywhere on } X^*. \qquad (2)$$

2.2   Aim

The FPD deals with the controlled closed loop system and describes it entirely in probabilistic terms. In order to represent the desired joint distribution of the considered closed loop variables the FPD requires the *ideal pdf* to be introduced.

The <u>aim</u> of the FPD task can be formulated in the following way:

Find admissible control strategy minimizing the KL divergence $\mathcal{D}\left(f_{\mathring{t}}||\,^I f_{\mathring{t}}\right),$

where $f_{\mathring{t}}$ is the joint pdf describing the controlled closed loop and $^I f_{\mathring{t}}$ is the ideal pdf.

Despite the aim is formulated, it remains rather indistinct without precise specification of DM task. It can be seen that the decision maker would need the *loss function* in order to express the aim through the ideal pdf, the *admissible strategies* to denote a set of possible actions and the further specifying constituents. Let the decision maker try to determine them gradually.

2.3   Specification

2.3.1   System

The joint pdf describing the controlled closed loop can be decomposed with the help of the chain rule [10].

$$f(d(\mathring{t}), x(\mathring{t})|x_0) = \qquad (3)$$

$$= f(x_0) \prod_{t \in t^*} f(y_t|u_t, x(t), d(t-1)) f(x_t|u_t, x(t-1), d(t-1)) f(u_t|x(t-1), d(t-1)).$$

From this relation it follows that the system is described by *observation model* $f(y_t|u_t, x(t), d(t-1))$, the *state evolution model* $f(x_t|u_t, x(t-1), d(t-1))$ and the general *randomized controller* $f(u_t|x(t-1), d(t-1))$ [11].

The following <u>assumptions</u> are made about the system models:

- the unobserved system state x_t does not depend on the past history of the system, i.e.

$$f(x_t|u_t, x(t-1), d(t-1)) = f(x_t|u_t, x_{t-1});$$

- the probability distribution of the observed system output y_t is determined only by the current system input u_t and the system state x_t, i.e.

$$f(y_t|u_t, x(t), d(t-1)) = f(y_t|u_t, x_t).$$

The practical example of the FPD will be demonstrated on <u>Gaussian models</u>.

Let the state evolution model be normal $f(x_t|u_t, x(t-1)) = \mathcal{N}_{x_t}(Ax_{t-1} + Bu_t; R)$, where $\mathcal{N}_z(\hat{z}; r)$ denotes normal pdf of z having expectation $\hat{z}$. Let the observation model be also normal $f(y_t|u_t, x_t) = \mathcal{N}_{y_t}(Cx_t + Du_t; r)$. Matrices A, B, C, D, R and r are assumed to be known. More illustrative way is to show the models by equations:

$$x_t = Ax_{t-1} + Bu_t + \omega_t \tag{4}$$

$$y_t = Cx_t + Du_t + e_t \tag{5}$$

The noises ω_t and e_t are supposed to be white and Gaussian with zero mean values and known covariances R and r respectively.

2.3.2 Admissible strategies

The set of admissible actions is restricted by the following <u>assumption</u>:

- admissible control strategies generate the system input u_t from the observed data history $d(t-1)$ and ignore the unobserved states $x(t-1)$, i.e.
$$f(u_t|x(t-1), d(t-1)) = f(u_t|d(t-1)).$$

Hence, the joint pdf (3) reduces to

$$f(d(\mathring{t}), x(\mathring{t})|x_0)f(x_0) = \prod_{t \in t^*} f(y_t|u_t, x_t)f(x_t|u_t, x_{t-1})f(u_t|d(t-1))f(x_0), \tag{6}$$

and the models of the system define the conditional probability distributions

$$f(y_t|u_t, x_t), \ f(x_t|u_t, x_{t-1}), \ f(u_t|d(t-1)).$$

4

2.3.3 Loss function

The loss function for the FPD task follows from its DM aim, namely, from searching the admissible control strategy minimizing the KL divergence (1). It means, the KL divergence between the joint pdf described the closed control loop and the decision maker's ideal pdf will be the loss function.

The decision maker's ideal pdf is constructed in the way analogous to (6)(note an appearance of superscript ^{I})

$$^If(d(\mathring{t}), x(\mathring{t})|x_0) \, ^If(x_0) = \prod_{t \in t^*} {}^If(y_t|u_t, x_t) \, ^If(x_t|u_t, x_{t-1}) \, ^If(u_t|d(t-1))f(x_0),$$

(7)

where pdf $^If(y_t|u_t, x_t)$ describes the ideal model of observation,
$\quad$ $^If(x_t|u_t, x_{t-1})$ – the ideal model of state evolution,
$\quad$ $^If(u_t|d(t-1))$ – the ideal control law.
It is obviously that the prior pdf on initial states x_0^* cannot be influenced by the optimized control strategy, so $^If(x_0) = f(x_0)$.

On <u>Gaussian models</u> it can be shown that the ideal pdfs are defined as

$$^If(x_t|u_t, x_{t-1}) = \mathcal{N}_{x_t}(^I\hat{x}_t; R), \quad ^If(y_t|u_t, x_t) = \mathcal{N}_{y_t}(^I\hat{y}_t; r),$$

while $^If(u_t|d(t-1)) = \mathcal{N}_{u_t}(^I\hat{u}_t; q)$ is required.

With normal models and normal ideal pdfs, the state-space FPD setting of the task reduces to Kalman filtering and design minimizing the quadratic criterion. The penalization matrices are interpreted here as inversions of ideal covariance matrices

$$J = \sum_{t=1}^{N} \left[x_t' R^{-1} x_t + u_t' q^{-1} u_t \right], \; where \; R^{-1} \geq 0, q^{-1} > 0.$$

(8)

2.3.4 Experience

The knowledge about the system, available to the decision maker for achieving the aim, includes

- past outputs,

- model structure,

- and model parameters.

Theoretically, the problem of identification could be included in a list of the tasks for FPD, but here the parameters are supposed to be known.

2.3.5 Ignorance

Such knowledge about the system as

- future inputs,

- future outputs,

- and all the system states.

is not available to the decision maker.

2.3.6 Decision

It has been already mentioned that the FPD is considered in its space-state setting. According to the previous constituent, the knowledge about the system states is unavailable to achieve the DM aim. Therefore, the decision maker should solve the state estimation task.

Bayesian filtering in closed control loop includes the next steps [10] on conditions that the prior pdf $f(x_0)$ is available:

$$f(x_t|u_t, d(t-1)) = \int f(x_t|u_t, x_{t-1}) f(x_{t-1}|d(t-1))\, dx_{t-1} \qquad (9)$$

$$f(x_t|d(t)) = \frac{f(y_t|u_t, x_t) f(x_t|u_t, d(t-1))}{\underbrace{\int f(y_t|u_t, x_t) f(x_t|u_t, d(t-1))\, dx_t}_{f(y_t|u_t, d(t-1))}}, \qquad (10)$$

where (9) is understood as the time updating, while relation (10) as data updating.

The important remark should be made here: the stochastic filtering *does not depend* on the used admissible control strategy $\{f(u_t|d(t-1))\}_{t \in t^*}$ but on the generated inputs only.

Proof: See [11].

Managing the Gaussian models the solution coincides with Kalman filtering [12] that calculates the expectation and covariance for $f(x_t|d(t)) = \mathcal{N}_{x_t}(\hat{x}_t; P)$. The equations fulfilled are shown below

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_t + K(y_t - C\hat{x}_{t-1} - Du_t),$$

$$K = PC'r^{-1}, \ \ P^{-1} = C'r^{-1}C + (APA' + R)^{-1}.$$

Thus the state estimation task built in the FPD is solved. The successive task, necessary to achieve the decision making aim, is the regulation problem.

Let the joint pdf $f(x(\mathring{t}), d(\mathring{t})|x_0)$ and the ideal one $^I f(x(\mathring{t}), d(\mathring{t})|x_0)$ be considered (see Sections 2.3.1 and 2.3.3) and let the assumptions (see Sections 2.3.1 and 2.3.2) be fulfilled.

The optimal admissible control strategy, which minimizes $\mathcal{D}\left(f_{\mathring{t}}\| \, ^I f_{\mathring{t}}\right)$ is randomized one given by the pdfs

$$^o f(u_t|d(t-1)) = \, ^I f(u_t|d(t-1)) \frac{\exp[-\omega(u_t, d(t-1))]}{\gamma(d(t-1))}, \ \ t \in t^*, \qquad (11)$$

$$\gamma(d(t-1)) \equiv \int \, ^I f(u_t|d(t-1)) \exp[-\omega(u_t, d(t-1))]\, du_t.$$

Starting with $\gamma(d(\mathring{t})) \equiv 1$, the functions $\omega(u_t, d(t-1))$ are generated recursively for $t = \mathring{t}, \mathring{t} - 1, \ldots, 1$ in a backward manner, as it is shown below

$$\omega(u_t, d(t-1)) \equiv \int \Omega(u_t, d(t-1), x_{t-1}) f(x_{t-1}|d(t-1)) dx_{t-1}, \qquad (12)$$

where $f(x_t|d(t))$ is updated according to the filtering operations (9)-(10) and

$$\Omega(u_t, d(t-1), x_{t-1}) \equiv \qquad\qquad (13)$$

$$\equiv \int f(y_t|u_t, x_t) f(x_t|u_t, x_{t-1}) \ln\left(\frac{f(y_t|u_t, x_t) f(x_t|u_t, x_{t-1})}{\gamma(d(t))\,^I f(y_t|u_t, x_t)\,^I f(x_t|u_t, x_{t-1})}\right) dy_t\, dx_t.$$

The result is based on the use of the definition of the KL divergence (1), its basic property (2), Fubini theorem on multiple integration [13], marginalization and normalization of pdfs and the chain rule for them [10].

Proof: See [11].

In case of <u>normal models</u> the solution comes to linear quadratic Gaussian (LQG) algorithm and operating with Riccati matrix. The optimal admissible control strategy is also normal here

$$^o f(u_t|d(t-1)) = \mathcal{N}_{u_t}(L\hat{x}_t;\,^o q)$$

and calculated with Riccati equation

$$S_i = A'S_{i-1}A + C'r^{-1}C - L_i'\,^o q L_i + R^{-1}, \qquad (14)$$

$$^o q = (B'S_{i-1}B + D'r^{-1}D + q^{-1}),$$

$$L_i = \,^o q^{-1}(B'S_{i-1}A + D'r^{-1}C),$$

where index i is a number of iterations, $i \in [1, N]$, $S_0 = 0$, $L_N, L_{N-1}, \ldots, L_1$ should be applied backward. The calculation of the optimal control strategy can be shown in more customary way

$$u_{t+i} = L_{N-i+1}x_{t+i-1}, \quad i = 1, 2, \ldots, N. \qquad (15)$$

The form of Riccati equation shown at (14) is caused by the presence of Du_t in the observation model. Without it the mean value of the optimal controller is usual control law obtained in linear quadratic design with state penalization R^{-1} and input penalization q^{-1}.

Now, when six specifying constituents have been introduced for the given task and the example of practical solution on Gaussian models has been offered, it should be said about implementation.

3   IMPLEMENTATION

All said above must find its final form that could be used for teaching, and primarily for e-learning. To achieve this objective, the DM tasks with their

7

specifications, offered solutions and recommended experiments are gathered into collection of illustrative examples, available as .pdf document. Each example is represented by its own main page.

Fig.1 shows the main page of the example for FPD. The structure of the page contains the aim of the task and a brief description of the solution. Here it should be noticed that the important decision making concepts, such as the KL divergence, the ideal pdf and even system inputs and outputs, etc, are the hyperlinks, providing their definitions in new window by clicking on them.

Further the specification of the task with six constituents described is shown. The specifying constituents also enable to see their definitions in order to help to understand such a way of defining the task. Each variable occurred provides its description as well.

All the definitions are collected in the termbase of Bayesian decision-making. The termbase is a set of terms created according to the style document specially developed. Those definitions, which include the further DM terms, also give their definitions with the help of hyperlinks. The termbase is generated automatically, when the example's text is edited. In ready-to-use form it is a .pdf document.

To learn the specialities of the task Fig.1 offers the recommended experiments. The brief comments of system behavior during them and references are shown. Naturally a link to read the theory is available at bottom of the page.

To make the recommended experiments and to show the results it is possible to run the task with the help of link "Run example". This link starts the associated application (MATLAB, Octave) and runs the dialog, which for the FPD case can be seen at Fig.2. The dialog shows the default values of parameters that can be further changed by the decision maker according to recommended experiments. For example, Fig. 3 demonstrates the results of experiment of changing the value of input penalization. The graphs show influence of gradual increase of this value on the system behavior. It can be seen that the higher the value of input penalization is, the more difference is between the system output and its desired value.

In the end it should be said that during the learning the decision maker can use links "Edit m-file" and "Restore m-file".

4   CONCLUSION

In the described way the illustrative examples with the corresponding models and solutions for majority of DM tasks have been presented. The specifying constituents precisely defined for each task are believed to help the decision maker not only in learning, but also in setting the subsequent DM tasks and their solutions.

Since the work is being fulfilled within the bounds of the Edukalibre e-learning open source project (http://edukalibre.org), then the tasks have been implemented in Octave [15], which is a free clone of MATLAB. From this it also follows that the source codes of the material (both LaTeX and Octave) will be

put on the repository free for download and further development. The demo-version of examples and software packages for download are also available at UTIA Moodle environment (http://moodle.utia.cas.cz).

Acknowledgement

References

[1] Stutz J, Cheeseman P. AutoClass - a Bayesian approach to classification. in *Maximum Entropy and Bayesian Methods,* Skilling J, Sibisi S. (eds). Kluwer: Dordrecht, 1995

[2] Mongin P. Consistent Bayesian Aggregation. *Journal of Economic Theory,* 1995 **66**:313–351.

[3] Jensen F.V. *Bayesian Networks and Decision Graphs* Springer-Verlag: New York, 2001.

[4] Berger J.O. *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag: New York, 1985.

[5] Bernardo J.M, Smith A.F.M. *Bayesian Theory* (2st edn). Wiley: Chichester, 1997.

[6] Kárný M, Böhm J, Guy T.V, Jirsa L, Nagy I, Nedoma P, Tesař L. *Optimized Bayesian Dynamic Advising: Theory and Algorithms* (in print). Springer-London, 2005.

[7] Lamport L. *LATEX: A Document Preparation System* (2nd edn). Addison-Wesley, 1994.

[8] Kárný M. Towards fully probabilistic control design. *Automatica* 1996; **12**(32):1719–1722.

[9] Kullback S, Leibler R.. On information and sufficiency. *Annals of Mathematical Statistics* 1951; **22**:79–87.

[10] Peterka V. Bayesian system identification. In *Trends and Progress in System Identification,* Eykhoff P. (eds). Pergamon Press: Oxford, 1981; 239–304.

[11] Kárný M, Guy T.V. Fully probabilistic control design. *Systems & Control Letters* 2005; submitted.

[12] Meditch J.S. *tochastic Optimal Linear Estimation and Control.* AMc. Graw Hill, 1969.

[13] Rao M.M. *Measure Theory and Integration.* John Wiley: New York, 1987.

[14] Welch G, Bishop G. An introduction to the Kalman filter. *UNC-CH Computer Science,* 1995; **95-041**.

[15] Eaton J.W. *GNU Octave Manual. A high-level interactive language for numerical computations* (3st edn). ISBN: 0-9541617-2-6.

Fully Probabilistic Design

Aim:
To find admissible control strategy minimizing the KL divergence (see Section 2.2).

Description:
For this task we consider normal state-space model and normal ideal pdf. The required admissible strategy must provide minimal KL divergence from the user's ideal pdf. Due to normal models the solution of fully probabilistic design comes to operating with linear quadratic Gaussian algorithm with penalization matrices interpreted as inversions of ideal covariance matrices. The solution includes the state estimation and design of the control strategy minimizing the quadratic criterion.

Specification:

System: described by observation and state evolution normal models (see 2.3.1);

Admissible strategy: causal with generating the input u_t from the observed data history $d(t-1)$ and ignoring the unobserved states $x(t-1)$ (see 2.3.2);

Loss function: KL divergence from ideal pdf (see 2.3.3);

Experience: past outputs; model structure; model parameters;

Ignorance: future inputs; future outputs; all states;

Decision: state estimation; optimal regulation (see 2.3.6);

Recommended experiments:
In order to observe a behavior of the controller it is useful to

1. *Try different model structure with parameter stru= [5 3 0], [10 6 0], [5 1 0], [20 1 0], [50 20 0], whose default value is [3 2 0];*

2. *Change value of input penalization by parameter qur=0.00001, 0.01, 0.05, 0.1, 0.5, 0.9. Default value of qur=0.001.*

Comments and references:
The user's compromise between quality and desired input signal is mostly expressed with the help of input penalization *qur*, which is a basic tuning knob. But in practice it is not enough to satisfy all requirements. Usually such tuning knobs as dynamic penalization, a setting of open loop transfer or filtration must be used.

| Run example | Edit m-file | Restore m-file | Read theory | Contents |

Figure 1: FPD example page

```
FULLY PROBABILISTIC DESIGN
=============================================================

--- Models -------------------------------------------------
Choice of system                 ->   S       = 'S01'
Horizon of the criterion         ->   nstep   = 100
Input penalization               ->   qur     = 0.001
Input increment penalization     ->   dqy     = 0
Output increment penalization    ->   dqu     = 0


--- Initial conditions -------------------------------------
Model structure                  ->   stru    = 3   2   0
Sampling period                  ->   To      = 0.1
------------------------------------------------------------


any changes ? ENTER = no, or type a command: >
```
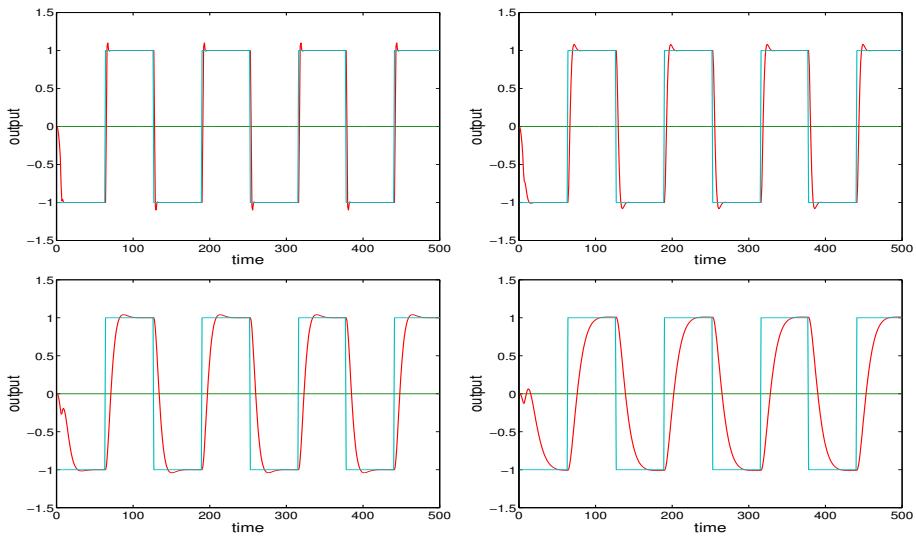
Figure 2: User dialog

Figure 3: Output and its desired value with the input penalization qur=0.00001, 0.001, 0.05, 0.5