Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

TEXTBOOK

Ivan Nagy, Lenka Pavelková, Evgenia Suzdaleva,
Jitka Homolová, Miroslav Kárný

BAYESIAN DECISION MAKING

Theory and Examples

October 2005

ÚTIA AVČR, P.O.Box 18, 182 08 Prague,
Czech Republic

Fax: (+420)286890378, http://www.utia.cas.cz,
E-mail: utia@utia.cas.cz

October 24, 2005

Contents

Chapter 1

Introduction

This text should support learning theory of Bayesian decision making under uncertainty presented at the course of "Adaptive Systems". There are good textbooks for learning both decision-making theory, e.g. [1, 2] and adaptive control, e.g. [3, 4]. The view presented in this text, namely, viewing adaptive systems as an important branch of *dynamic* decision making under uncertainty is dissipated in various textbooks and numerous papers. The classical text [5, 6] is probably closest to this one. Permanently growing computer power available, novel solutions of various theoretical and algorithmic subtasks justify this attempt to update the description of the theory.

1.1 Aim of the work

Theory of statistical decision making under uncertainty [2, 7] provides a unified logical structure for solving problems ranging from parameter estimation, prediction, pattern recognition, learning, testing of hypothesis etc. up to feedback control. The application width is also enormous: inspection, maintenance and control of technologic as well as economic processes; signal processing; information processing in medicine, physics, etc. Huge amount of available particular results, techniques and methods hide the common logical structure and leave often newcomers without a guide.

This *text aims to recall the relatively simple common structure of the mentioned decision tasks and to help the user in focusing on those ingredients that are specific to his/her problem.*

Naturally, the reaching of a satisfactory solution of a particular decision-making problem remains to be an iterative process [8]. The number of trial solutions can be, however, significantly reduced by using the advocated logical structure. This is the main practical outcome of mastering the presented theory.

The logical structure of the decision making under uncertainty is now well understood. The number of open problems, related especially to computational and implementation aspects, is, however, large. Some of them are formulated throughout the text. In this way, we try to *turn the attention of (prospective) researchers to challenging and important topics.*

1.2 Structure

The aim of the work dictates the top-down style of presentation: the logically simple but relatively abstract problem formulation and solution are gradually specialized. The logical skeleton of the theory is filled up by practical details later on in chapters dealing with particular examples.

Chapter 2 summarizes the theory of the design and states the basic analytical results. Usual decision-making tasks are discussed, formulated and solved in Chapter 3. Common evaluation techniques are characterized in Chapter 4.

The rest of the text serves for illustration of the theory. The structure of the remaining chapters is (more or less) uniform. For specific construction elements, they illustrate majority of decision-making tasks that can be related to the considered case. For instance, Chapter 6 deals with a bit academic problem of the controlled tossing of a coin. Simplicity of the calculus related to it helps us to focus on the logical structure to be mastered. Other examples are gradually less and less academic and become more and more of practical importance.

1.3 Readers

It is natural that authors of this text hope to have a wide range of readers. At the same time, it is natural to expect them to be very individual in their attitude, background and interests. So we can distinguish just two groups:

Research inclined readers who are interested in the state and problems of the area and do not worry of relatively abstract notions. They are expected to read the text linearly with occasional skips and complementary reading of references.

Application oriented readers who are searching for a flexible tool-set for their real-life problem. They are supposed to start with demos supplied to particular examples in order to judge whether the presented text is of interest for them at all. Any example, which is judged to be close to the considered application may serve to this purpose. If, as we hope, they will find it worth of interest and energy put in study they are expected to go through Chapter 3 in order to learn the structure of the problem formulation and its solution. Chapter 4 then serves for inspiration on practical evaluation techniques. Both of them contain pointers to the necessary auxiliary notions in remaining "theoretical" Chapters.

All readers are warned that:

- this text is not mathematical one in spite of all mathematical machinery (mis)used: mathematical purity and technical details are sacrificed in order present basic explanations as directly as possible,
- the presentation adapts and sometimes create symbols and terminology with the aim to balance generality and expressiveness: for instance, the notion of probability space with its σ-algebras is replaced by notions like ignorance, experience etc.,
- the attention focuses on problems that can be quantified and potentially solved in as algorithmic way as possible: the relatively straightforward extension to non-numerical problems are not considered, for the sake of conciseness.

1.4 Acknowledgment

This work has specific authors but it could not be created without their predecessors V. Peterka, A. Halousková, A. Peréz, all members and students of Adaptive Systems Department, including visiting members of our Department K. M. Hangos and K. Warwick, applicators P. Ettler, J. Fessl, J. Heřmanská, A. Lízr, A. Lízrová and others.

The long-term support of Institute of Information Theory and Automation as well as various grants have made possible to work in this interesting research area with a great application potential and to create finally this text. The grants that contributed to the current version of this text are Edukalibre (EC Sokrates project 110330-CP-2003-ES-Minerva-M), GA ČR, the project 102/03/0049, AV ČR the project 1ET 100 750 401 and MŠMT ČR 1M6798555601.

Chapter 2

Underlying theory

At a general level, the presented decision-making theory should help the decision maker – typically, human being – to select one of the available options (decisions). These options relate to a part of the real world, to a system. They concern with descriptions of the system and/or its influencing.

This chapter summarizes the design principles and tools exploited later on in our particular task, i.e. operator support. It starts with conventions and notions used throughout, Sections 2.1, 2.2. The considered framework covers a broad range of problems. Inevitably, the adopted symbols and notions have more specific meanings in specific application fields. The reader is kindly asked to be patient especially in this respect.

The adopted principle of the optimal decision-making under uncertainty, inspected in Section 2.3, implies the following important conclusion: incomplete knowledge and randomness have the same operational consequences for decision-making and should be treated in the same way labelled as *Bayesian decision making*. In the same section, the design of optimal decision rules is solved. In Section 2.4 the design of the optimal strategies is derived. These designs work with models that are usually obtained through Bayesian learning described in Section 2.5.

A similarly formulated control design, that is a specific decision-making problem, is presented in [9]. For a detailed explanation of Bayesian learning see [10]. This chapter is a tailored version of a chapter in [11] where our view on decision making is summarized.

2.1 General conventions

The conventions presented here are mostly respected in this work. If some exception is necessary it is named just at the place of its validity. If some verbal notions are introduced within bodies of Propositions, Remarks etc., then they are *emphasized by the print* that differs from that of the surrounding text. Moreover, they appear in Index. Sometimes, important parts of sentences are stressed by underlining them.

f is the letter reserved for probability (density) functions (p(d)f).

The meaning of the p(d)f is given through the name of its argument.

x^* denotes the range of x, $x \in x^*$.

$\mathring{x}$ *denotes the number of members in the countable set* x^* or *number of entries in the vector* x.

$\equiv$ *means the equality by definition.*

x_t *is a quantity* x *at the discrete time labelled by* $t \in t^* \equiv \{1, \ldots, \mathring{t}\}$.

$\mathring{t} \leq \infty$ *is called (decision, learning, prediction, control, advising) horizon.*

$x_{i;t}$ *is an* i*th entry of the array* x *at time* t.

The semicolon in the subscript indicates that the symbol following is time index.

$x(k\_l)$ *denotes the sequence* x_t *with* t *between time moments* $k \leq l$, i.e. $x(k\_l) \equiv x_k, \ldots, x_l$.

$x(t) \equiv x(1\_t)$.

$x(t)$ *is an empty sequence and reflects just the prior information if* $t < 1$.

$\mathrm{supp}\,[\,f(x)]$ *is the support of the pdf* $f : x^* \to [0, \infty]$*, i.e. the subset of* x^* *on which* $f(x) > 0$.

2.2 Basic notions and notations

Below, if need be, a brief characterization of the introduced notion is complemented by explanatory comments (printed differently).

Quantity *is a multivariate mapping.*

The domain and form of the quantity are mostly unused and unspecified. The introduced notion corresponds with random variable used in probability theory. The use of the alternative term should stress that probability serves us as the tool adopted for decision making under uncertainty. The term quantity stresses our orientation on numerical values that arise mostly by observing physical quantities. However, quantities with a discrete range that need not have numerical meaning are also considered.

Realization *is a value of the quantity for its fixed argument.*

Often, the quantity and its realization are not distinguished, as usual. The proper meaning is determined by the context.

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

Decision maker might be a person, group of persons or mechanisms.

The system is specified with respect to the aim that the decision maker wants to reach and with respect to the tools he/she has available. In other words, the boundaries (that can be penetrated) of the system are implied by the decision task.

Decision $a \in a^*$ *is the value of a quantity that can be directly chosen by the decision maker for reaching his/her aims.*

A decision task makes sense only when the decision maker faces real option, i.e. when $\mathring{a} > 1$.

(Decision) experience $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$ *is knowledge about the system available to the decision maker for the selecting the decision* $a \in a^*$.

For example, if just data values D are available for constructing the estimate $\hat{\Theta}$ of an unknown quantity $\Theta \in \Theta^*$ then the experience is $\mathcal{P}_{\hat{\Theta}^*} \equiv D$. Often, experience includes the past data observed.

(Decision) ignorance $\mathcal{F}_{a^*} \in \mathcal{F}^*_{a^*}$ *is knowledge about the system unavailable to the decision maker for the choice of the value of* a.

The unknown value of the quantity Θ to be estimated is a part of the ignorance $\mathcal{P}_{\hat{\Theta}^*}$ of its estimate $\hat{\Theta}$. Often, ignorance contains future, still unobserved data.

(System) behavior $\mathcal{Q}^*$ *consists of all possible realizations (of trajectories)* $\mathcal{Q}$*, i.e. values of all quantities within the time span determined by the horizon of interest that are related to the system and considered by the decision maker.*

The realization $\mathcal{Q}$ can be split with respect to any decision $a \in a^*$ into the relevant experience $\mathcal{P}_{a^*}$ and ignorance $\mathcal{F}_{a^*}$. Formally, $\mathcal{Q} = (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})$.

Note that a single realization $\mathcal{Q}$ can be split differently with respect to a pair of decisions, say a, $\tilde{a}$ with different experience $\mathcal{P}_{a^*} \neq \mathcal{P}_{\tilde{a}^*}$ and different ignorance $\mathcal{F}_{a^*} \neq \mathcal{F}_{\tilde{a}^*}$. $\mathcal{Q} = (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*}) = (\mathcal{P}_{\tilde{a}^*}, \tilde{a}, \mathcal{F}_{\tilde{a}^*})$.

(System) input $u \in u^*$ *is a decision, which is supposed to influence the ignorance part* $\mathcal{F}_{u^*}$ *of the (system) behavior.*

A manipulated valve position influencing a fluid flow is the system input. A point estimate $\hat{\Theta}$ of an unknown (realization of) quantity Θ is an instance of the decision that is not the system input. It is chosen with the aim to describe the system and exhibits no direct influence on it.

(System) output $y \in y^*$ *is an observable quantity that provides the decision maker information about the (system) behavior.*

A pressure measured in an isolated heated system is an instance of the output. A pressure applied on a closed system is an instance of the input.

To be or not to be input or output is a relative property.

Innovation $\Delta_t \in \Delta^*_t$ *contains quantities that are not included in the experience* $\mathcal{P}_{a^*_t}$ *but are included in* $\mathcal{P}_{a^*_{t+1}} \setminus a_t$.

Often, $\Delta_t = y_t =$ the system output at time t.

Decision rule $\mathcal{R} : \mathcal{Q}^* \to a^*$ *is a mapping that assigns a decision* $a \in a^*$ *to the behavior* $\mathcal{Q} \in \mathcal{Q}^*$.

Causal decision rule $\mathcal{R} : \mathcal{P}^*_{a^*} \to a^*$ *is a mapping that assigns the decision* $a \in a^*$ *to its experience* $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$.

In other words, the decision a made by a causal decision rule is uninfluenced by the related ignorance $\mathcal{F}_{a^*}$.

Mostly, we deal with the causal decision rules so that the term "causal" is often dropped.

Estimator is an instance of the causal decision rule $\mathcal{R} : \mathcal{P}^*_{\hat{\Theta}^*} \to \hat{\Theta}^*$ that assigns an estimate $\hat{\Theta}$ of the unknown quantity $\Theta \in \Theta^*$ to the available experience $\mathcal{P}_{\hat{\Theta}^*}$.

Strategy *is a sequence of decision rules* $\{\mathcal{R}_t : \mathcal{Q}^* \to a_t^*\}_{t\in t^*}$.

Causal strategy $\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}$ *is a sequence made of causal decision rules.*

Again, we deal mostly with causal strategies so that the term "causal" is often dropped.

Controller *is a causal strategy assigning inputs* u_t *to experience* $\mathcal{P}_{u_t^*}$, $\forall t \in t^*$.

The proportional controller given by the constant p is an example of the causal control strategy $\{y^*_{t-1} \to u_t^* : u_t = py_{t-1}\}_{t\in t^*}$.

Design *selects the decision rule or strategy.*

The design selecting a single rule is called *static design.* The choice of the strategy is called *dynamic design.*

The person (group) who makes the selection is *designer.* Authors and readers of this text are supposed to be designers. In that sense, the term *we* used within the text should mostly be read: *we designers.*

Agreement 2.1 (Interface between the design and reality) *Physical connections of the design domain to the real world (sensors, transmission lines, actuators etc.) are here always taken as a part of the system. Consequently, all quantities and mappings considered can be and are taken as mathematical entities living in an abstract calculating machine.*

Uncertain behavior *(related to static design) is a behavior whose realizations* $\mathcal{Q}$ *can be decomposed into*

- $\mathcal{Q}_\mathcal{R} \equiv$ *the part that is unambiguously determined by the considered decision rule* $\mathcal{R} \in \mathcal{R}^*$,
- uncertainty $\Upsilon \equiv$ *the part of the behavior that belongs to the ignorance* $\mathcal{F}_{\mathcal{R}(\mathcal{P})^*}$ *of decisions* $\mathcal{R}(\mathcal{P})$ *generated by the admissible rules* $\mathcal{R} \in \mathcal{R}^*$ *and uninfluenced by them, even indirectly.*

With an abuse of notation, we write the corresponding decomposition of realization $\mathcal{Q} = (\mathcal{Q}_\mathcal{R}, \Upsilon) = (\cdot, uncertainty)$.

Uncertain behavior related to dynamic design is encountered if any of its rules faces uncertainty.

Incomplete knowledge of (the realization of) a considered quantity $\Theta \in \Theta^*$ makes the behavior uncertain.

External unobserved noise influencing the system makes its behavior uncertain.

Uncertainty expresses both incomplete knowledge and randomness.

Uncertain behavior in the dynamic design is defined in a similar way by replacing rules with strategies. The related cumbersome notation is avoided in our presentation.

Decision-making *means design and application of a decision rule (strategy).*

Admissible strategy *is a strategy* $\{\mathcal{R}_t\}_{t\in t^*}$ *that*

- is *causal* $\Leftrightarrow$ $\{\mathcal{R}_t\}_{t\in t^*} \equiv \{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}$ *and*

- meets *technological restrictions* $\Leftrightarrow$ *the ranges of its decision rules are in pre-specified subsets of respective decision sets.*

Loss function, $\mathcal{Z} : \mathcal{Q}^* \to [0, \infty]$, *quantifies the degree of achievement of the design aim.*

The loss function measures quality of the realizations $\mathcal{Q}$ and indirectly (partially) orders admissible strategies influencing the behavior.

The smaller the value of $\mathcal{Z}(\mathcal{Q})$ the better.

Negative value of profit might be a loss function in economic applications.

Quadratic deviation $(\hat{\Theta} - \Theta)'(\hat{\Theta} - \Theta)$ ($'$ is transposition) of the estimate $\hat{\Theta}$ and of an unknown vector Θ is a wide spread loss function used for ranking decision rules generating point estimates.

The important case of multi-valued loss function [12] is out of the scope of this text but the approach discussed in Section 2.3 could be relatively simply extended to it by embedding the index of the "decisive" entry of the loss function into uncertainty.

"Expected" loss $\tilde{\mathcal{E}}(\mathcal{Z}) \equiv \tilde{\mathcal{E}}_{\{\mathcal{R}\}}(\mathcal{Z})$ *assigns to the considered loss function* $\mathcal{Z}$ *and strategy* $\{\mathcal{R}\}$ *a value in* $[0, \infty]$ *that is independent of the uncertainty.*

The quotation marks as well as the sign $\tilde{}$ serve us in the discussion that shows that, under widely acceptable conditions, we have to deal with expectation in mathematical sense. Then, they are unused any more.

Optimal design *selects an admissible strategy that leads to the smallest value of the "expected" loss function.*

Practically admissible strategy *is an admissible strategy that respects restrictions limiting complexity of the decision-making.*

The complexity is considered with respect to the computational resources available at the design and application stages. The majority of discussed problems in which the complexity restriction play a role are computationally hard in terms of computer sciences. An intuitive understanding of the computational complexity is sufficient to our purposes.

Practically optimal design *selects a practically admissible strategy giving the smallest values of the "expected" loss.*

The presented optimal design provides optimal admissible strategies and can be simply adapted to provide strategies of a pre-specified complexity by optimizing over a set of simple decision rules, for instance, over proportional controllers only.

Operational formal tools for practically optimal design are not available. It is not known how to make the optimal design of a pre-specified complexity.

We never know whether the selection of the constant determining proportional controller made with use of, say, ten algebraic operations is really the best one possible among all selections that are allowed to perform ten algebraic operations.

This is the *main barrier of the applicability* of the theory describing the optimal design. The optimal design becomes a practical tool by employing sound engineering heuristics. The practical optimum is not guaranteed. This fact is stressed by using the term *suboptimal design.*

2.3 Decision-making under uncertainty

Here, we describe a general way how to understand and face uncertainty that causes incomplete ordering of strategies. In order to avoid a cumbersome notation, we formulate the adopted design principle, related requirements and their consequences for the static design, i.e. the design of a single not necessarily causal decision rule. The obtained results are also valid for the dynamic design, i.e. the choice of decision strategies.

Agreement 2.2 (Uncertainty in decision-making) *Decision-making under uncertainty arises if the optimal decision-making is to be performed and*

- *at least a pair of different decisions can be made, å* > 1,
- *the considered loss function* $\mathcal{Z}(\mathcal{Q}) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)$ *depends on non-void set* Υ^* *of uncertainties.*

Under uncertainty, the function $\mathcal{Z}_{\mathcal{R}}(\Upsilon) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)$ *of uncertainty* Υ*, i.e. the part of realization in ignorance uninfluenced by the rule* $\mathcal{R}$*, is assigned to each considered decision rule* $\mathcal{R} \in \mathcal{R}^*$*. The set of such functions is denoted* $\mathcal{Z}_{\mathcal{R}^*}$

$$\mathcal{Z}_{\mathcal{R}^*} \equiv \{\mathcal{Z}_{\mathcal{R}} : \Upsilon^* \to [0, \infty],\ \mathcal{Z}_{\mathcal{R}}(\Upsilon) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)\}_{\mathcal{R} \in \mathcal{R}^*}\,. \tag{2.1}$$

Under uncertainty, the loss function is insufficient for a *complete ordering* (comparing) of admissible rules, and consequently strategies, in spite of the fact that its values are in the fully ordered set. Each $\mathcal{R}$ is assigned a function in the set (2.1) and not a single number.

For instance, consider a pair of estimates $\hat{\Theta}_1 \neq \hat{\Theta}_2$ of an unknown scalar quantity $\Theta \in \Theta^* \equiv (-\infty, \infty)$. They cannot be compared with the help of the quadratic loss function $(\Theta - \hat{\Theta})^2$ as we do not know whether Θ is in that part of Θ^* where $(\Theta - \hat{\Theta}_1)^2 \leq (\Theta - \hat{\Theta}_2)^2$ or in its complement.

2.3.1 Ordering of decision rules

We rely on the following systematic choice of a good decision rule.

Agreement 2.3 (Design principle: "expectation"-minimization design)

- *A functional* $\tilde{\mathcal{E}}_{\mathcal{R}}$*, called "*expectation*", is selected by designer. It assigns to functions in (2.1) (determined by the considered loss function* $\mathcal{Z}$ *and indexed by inspected decision rules* $\mathcal{R} \in \mathcal{R}^*$*) a non-negative "*expected loss*"* $\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}]$

$$\begin{aligned} & \tilde{\mathcal{E}}_{\mathcal{R}} : \mathcal{Z}_{\mathcal{R}^*} \to [0, \infty] \\ \mathcal{Z}_{\mathcal{R}^*} \quad \equiv \quad & \{\mathcal{Z}_{\mathcal{R}} : \Upsilon^* \to [0, \infty],\ \mathcal{Z}_{\mathcal{R}}(\Upsilon) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)\}_{\mathcal{R} \in \mathcal{R}^*} \end{aligned} \tag{2.2}$$

- *The minimizer of* $\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)] \equiv \tilde{\mathcal{E}}[\mathcal{Z}_{\mathcal{R}}]$ *found in* $\mathcal{R}^*$ *is taken as the optimal decision rule.*

The outcome of this design depends on the introduced "*expectation*" $\tilde{\mathcal{E}}_{\mathcal{R}}$. It has to avoid (at least) unequivocally bad rules. Such bad rules are identified here with dominated decision rules.

Agreement 2.4 (Dominated decision rules; strictly isotonic expectation) *Let a loss function $\mathcal{Z}$ measure the quality of the behavior. The decision rule $\mathcal{R}: \mathcal{Q}^* \to a^*$ is called* dominated *iff there is another decision rule $\tilde{\mathcal{R}}: \mathcal{Q}^* \to a^*$ such that for*

$$\mathcal{Z}_{\mathcal{R}} \geq \mathcal{Z}_{\tilde{\mathcal{R}}} \;\Leftrightarrow\; \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon) \geq \mathcal{Z}(\mathcal{Q}_{\tilde{\mathcal{R}}}, \Upsilon),\ \forall \Upsilon \in \Upsilon^*. \tag{2.3}$$

The dominated decision rule is called strictly dominated *iff there is a non-trivial subset of Υ^* on which the inequality (2.3) is strict.*

The "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ (2.2) is said to be strictly isotonic if for a decision rule $\mathcal{R}$ strictly dominated by a decision rule $\tilde{\mathcal{R}}$ it holds

$$\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \Upsilon)] > \tilde{\mathcal{E}}_{\tilde{\mathcal{R}}}[\mathcal{Z}(\mathcal{Q}_{\tilde{\mathcal{R}}}, \Upsilon)]$$

Now we can specify the dominated decision rules as those to be surely avoided.

Requirement 2.1 (Inadmissibility of strictly dominated decision rules) *The considered "expectation"-minimization design, Agreement 2.3, must not lead to the strictly dominated decision rule.*

We want to select the "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ so that the decision rules are ordered as objectively as possible. It means that the ordering should weakly depend on the inspected set of rules $\mathcal{R}^*$, on the loss function $\mathcal{Z}$ considered and on the "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ selected by the designer.

Requirement 2.2 (Independence of $\mathcal{R}^*$) *The chosen "expectation" should fulfill Requirement 2.1 even if we restrict the set of possible decision rules $\mathcal{R}^*$ to any of its non-trivial subsets that contains at least two different rules while at least one of them gives a finite "expected" loss.*

Proposition 2.1 (Isotonic ordering) *Assume that there is a rule in $\mathcal{R}^*$ for which the "expected" loss is finite. Then, Requirements 2.1, 2.2 are fulfilled, i.e. the chosen rule is not dominated even if the set $\mathcal{R}^*$ is restricted, iff the "expectation" is strictly isotonic, see Agreement 2.4.*

Proof:

1. We prove by contradiction that – with strictly isotonic "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ – the minimizer cannot be strictly dominated.

   Let $\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}]$ be strictly isotonic on its domain $\mathcal{Z}_{\mathcal{R}^*}$ (2.2) and $\mathcal{R}_o \in \mathcal{R}^*$ be a minimizer of the "expected" loss. The minimizer gives necessarily a finite value of the corresponding $\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}]$. Let $\mathcal{R}_d \in \mathcal{R}^*$ dominate it strictly. Then, because of the construction of $\mathcal{R}_o$, the strict dominance and strictly isotonic nature of $\tilde{\mathcal{E}}_{\mathcal{R}}$, we get the following contradictory inequality

   $$\tilde{\mathcal{E}}_{\mathcal{R}_o}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_o}, \Upsilon)] \underbrace{\leq}_{minimum} \tilde{\mathcal{E}}_{\mathcal{R}_d}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_d}, \Upsilon)] \underbrace{<}_{strictly\ isotonic} \tilde{\mathcal{E}}_{\mathcal{R}_o}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_o}, \Upsilon)].$$

2. We prove by contradiction that use of an "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ that is not strictly isotonic leads to violation of Requirement 2.1 when Requirement 2.2 holds.

If $\tilde{\mathcal{E}}_{\mathcal{R}}[\mathcal{Z}]$ is not strictly isotonic on its domain $\mathcal{Z}_{\mathcal{R}^*}$ (2.2) then there is rule $\mathcal{R}_1 \in \mathcal{R}^*$ strictly dominated by a decision rule $\mathcal{R}_d \in \mathcal{R}^*$ such that

$$\tilde{\mathcal{E}}_{\mathcal{R}_d}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_d}, \Upsilon)] \geq \tilde{\mathcal{E}}_{\mathcal{R}_1}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_1}, \Upsilon)].$$

If we restrict the set of decision rules $\mathcal{R}^*$ to the pair of rules $\{\mathcal{R}_d, \mathcal{R}_1\}$ then $\mathcal{R}_1$ can always be taken as the optimal decision rule. Thus, under Requirement 2.2, Requirement 2.1 is not fulfilled with such an "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$.

$\diamondsuit$

Requirement 2.2 guarantees suitability of the "expectation" to a wide range of decision rules. The "full objectivity" of ordering of decision rules is approached if the "expectation" $\tilde{\mathcal{E}}_{\mathcal{R}}$ is made weakly dependent on the loss function considered. We have to adopt rather technical conditions on the set of such loss functions. Essentially, applicability to a very smooth functions and a restricted version of "linearity" of $\tilde{\mathcal{E}}_{\mathcal{R}}$ are required.

Requirement 2.3 (Independence of loss function) *Let us consider various loss functions $\mathcal{Z} \in \mathcal{Z}^*$. The "expectation" $\tilde{\mathcal{E}}$ acts on union of the sets $\mathcal{Z}_{\mathcal{R}^*}$ (2.1) of functions with a common uncertainty set Υ^*. The union is taken over the set of the loss functions*

$$\mathcal{Z}^*_{\mathcal{R}^*} \equiv \cup_{\mathcal{Z} \in \mathcal{Z}^*} \mathcal{Z}_{\mathcal{R}^*} \tag{2.4}$$

The set $\mathcal{Z}^*_{\mathcal{R}^*}$ is required to contain a subset of test loss functions *that are zero out of a compact non-empty subset Ω of $\mathcal{Q}^*$ and continuous on Ω.*

The "expectation" is assumed to be an isotonic, sequentially continuous and boundedly uniformly continuous functional on $\mathcal{Z}^*_{\mathcal{R}^*}$ that is, moreover, additive on loss functions with non-overlapping supports

$$\tilde{\mathcal{E}}[\mathcal{Z}_1 + \mathcal{Z}_2] = \tilde{\mathcal{E}}[\mathcal{Z}_1] + \tilde{\mathcal{E}}[\mathcal{Z}_2] \text{ if } \mathcal{Z}_1\mathcal{Z}_2 = 0,\ \mathcal{Z}_1, \mathcal{Z}_2 \in \mathcal{Z}^*_{\mathcal{R}^*}.$$

This technical Requirement allows us to get an integral representation of the "expectation" searched for. Its proof as well as definitions of the adopted non-common terms can be found in [13].

Proposition 2.2 (Integral form of the "expectation") *Under Requirement 2.3, the "expectation" $\tilde{\mathcal{E}}$ has the form*

$$\tilde{\mathcal{E}}[\mathcal{Z}] = \int_{\Omega} \mathcal{U}(\mathcal{Z}(\mathcal{Q}), \mathcal{Q})\, \mu(d\mathcal{Q}), \quad \text{where} \tag{2.5}$$

μ is a finite regular non-negative Borel measure on Ω. The utility function *$\mathcal{U}$ satisfies $\mathcal{U}(0, \mathcal{Q}) = 0$. It is continuous in values of $\mathcal{Z}(\cdot)$ almost everywhere (a.e.) on Ω, bounded a.e. on Ω for each $\mathcal{Z}$ in the set of the test loss functions.*

Remark(s) 2.1

1. *The test loss functions are widely applicable and their consideration implies no practical restriction. The continuity requirements on $\tilde{\mathcal{E}}$ are also widely acceptable.*

2. *The linearity of $\tilde{\mathcal{E}}$ on functions with non-overlapping support seems to be sound. Any loss function $\mathcal{Z} \in \mathcal{Z}^*_{\mathcal{R}^*}$ can be written as $\mathcal{Z} = \mathcal{Z}\chi_\omega + \mathcal{Z}(1-\chi_\omega) \equiv \mathcal{Z}_1 + \mathcal{Z}_2$, $\mathcal{Z}_1\mathcal{Z}_2 = 0$ with χ_ω denoting an* indicator of a set *$\omega \subset \Omega \subset \mathcal{Q}^*$. The indicator χ equals 1 on ω and it is zero outside of it.*

   The loss "expected" on the set ω and its complement should sum to the loss "expected" on the whole set of arguments.

3. *The utility function $\mathcal{U}$ allows the designer to express his/her attitude to the design consequences: he/she might be risk aware, risk prone or risk indifferent [12].*

4. *The utility function $\mathcal{U}$ and the non-negative measure μ are universal for the whole set of test functions. $\mathcal{U}$ and μ are (almost) "objective", i.e. suitable for a wide range of decision tasks facing the same uncertainty.*

We formulate now our final objectivity oriented Requirement.

Requirement 2.4 (Indifference of the designer)

- *The designer is risk indifferent, i.e. $\mathcal{U}$ is the identity mapping.*
- *The "expectation" preserves any constant loss $\tilde{\mathcal{E}}[constant] = constant$.*
- *The involved measure μ has Radon-Nikodým derivative, [13], $f(\mathcal{Q})$ with respect to a dominating measure denoted $d\mathcal{Q}$.*

  In the treated cases, $d(\cdot)$ is either Lebesgue or counting measure.

Adopting Requirement 2.4, we get the basic representation Proposition that introduces objective "expectation".

Proposition 2.3 (The objective "expectation") *Under Requirement 2.4, the "expectation" $\tilde{\mathcal{E}}$ (2.5) is formally identical with a* mathematical expectation. *The Radon Nikodým's derivative f, [13], has all properties of the* joint probability (density) function *(*p(d)f*) on $\mathcal{Q}^*$.*

Proof: It is sufficient to observe that the preservation of constants implies that μ is a probabilistic measure, i.e. $\mu \geq 0,\ \mu(\mathcal{Q}^*) = 1$. ♢

Remark(s) 2.2

1. *The (mathematical) "expectation" is distinguished by dropping the sign ˜ as well as the quotation symbols " "*

$$\mathcal{E}[\mathcal{Z}] \equiv \int_{\mathcal{Q}^*} \mathcal{Z}(\mathcal{Q}) f(\mathcal{Q})\, d\mathcal{Q} \equiv \int \mathcal{Z}(\mathcal{Q}) f(\mathcal{Q})\, d\mathcal{Q}. \tag{2.6}$$

2. *The first item in Requirement 2.4 is our last step we made in the objectivity direction. It has clear meaning: we support objective, emotionally indifferent designers.*

3. *The last item in Requirement 2.4 is unnecessary but it covers all cases we are here interested in. It helps us to deal with simpler objects, namely, with probability density functions (pdf) or probability functions (pf).*

   Mostly, we make no formal distinction between these cases and we use notations related to pdfs. Only when necessary, we underline that we deal with a pf and write integrals as sums.

4. *The notion of* conditional expectation $\mathcal{E}[\bullet|*]$, *[13] is adopted further on and it is treated in a simplistic way as the integral* $\int \bullet(a) f(a|*)\, da$ *weighted by the* conditional pdf $f(a|*)$.

2.3.2 Calculus with pdfs

The joint pdf f on $\mathcal{Q} \equiv (\alpha, \beta, \gamma)$ is analysed/synthesised using several pdfs related to it. Let us recall meaning of pdfs derived from $f(\mathcal{Q})$.

Agreement 2.5 (Nomenclature of pdfs; Independence) *We work with*

Name	***Meaning***
joint pdf $f(\alpha, \beta\|\gamma)$ *of* α, β conditioned on γ	*a pdf on* $(\alpha, \beta)^*$ *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* γ
marginal pdf $f(\alpha\|\gamma)$ *of* α *conditioned on* γ	*a pdf on* α^* *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* γ *with no information on* β
marginal pdf $f(\beta\|\alpha, \gamma)$ *of* β *conditioned on* α, γ	*a pdf on* β^* *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* α, γ

The conditioning symbol | *is dropped if just trivial conditions are considered.*
The pdf $f(\alpha, \beta)$ *is the* lower dimensional joint pdf *(with respect to the pdf* $f(\alpha, \beta, \gamma)$*),* $f(\beta)$ *is the* marginal pdf.
Quantities α *and* β *are* conditionally independent *under the condition* γ *iff*

$$f(\alpha, \beta|\gamma) = f(\alpha|\gamma) f(\beta|\gamma). \tag{2.7}$$

Our manipulations with the introduced pdfs rely on the following calculus.

Proposition 2.4 (Calculus with pdfs) *For generic* $(\alpha, \beta, \gamma) \in (\alpha, \beta, \gamma)^*$ *it holds:*

Non-negativity $f(\alpha, \beta|\gamma),\ f(\alpha|\beta, \gamma),\ f(\beta|\alpha, \gamma),\ f(\beta|\gamma) \geq 0.$

Normalization $\int f(\alpha, \beta|\gamma)\, d\alpha d\beta = \int f(\alpha|\beta, \gamma)\, d\alpha = \int f(\beta|\alpha, \gamma)\, d\beta = 1.$

Chain rule $f(\alpha, \beta|\gamma) = f(\alpha|\beta, \gamma) f(\beta|\gamma) = f(\beta|\alpha, \gamma) f(\alpha|\gamma).$

Marginalization $f(\beta|\gamma) = \int f(\alpha, \beta|\gamma)\, d\alpha,\ f(\alpha|\gamma) = \int f(\alpha, \beta|\gamma)\, d\beta.$

Bayes rule

$$f(\beta|\alpha, \gamma) = \frac{f(\alpha|\beta, \gamma) f(\beta|\gamma)}{f(\alpha|\gamma)} = \frac{f(\alpha|\beta, \gamma) f(\beta|\gamma)}{\int f(\alpha|\beta, \gamma) f(\beta|\gamma)\, d\beta} \propto f(\alpha|\beta, \gamma) f(\beta|\gamma). \tag{2.8}$$

The proportion sign, $\propto$, *means that the factor independent of* β *and uniquely determined by the normalization is not explicitly written in the equality represented.*

Independence equivalents

$$f(\alpha,\beta|\gamma) = f(\alpha|\gamma)f(\beta|\gamma) \Leftrightarrow f(\alpha|\beta,\gamma) = f(\alpha|\gamma) \text{ or } f(\beta|\alpha,\gamma) = f(\beta|\gamma). \tag{2.9}$$

Proof: For motivation see [10], a more precise and more technical treatment exploits the measure theory [13]. Technically, an intermediate insight can be gained by considering loss functions dependent only on a part of $\mathcal{Q}$ or with some parts of $\mathcal{Q}$ "fixed by the condition", [9]. ♢

Remark(s) 2.3

1. *Duplicates in the presented formulas stress symmetry of many of them.*
2. *The technical fact that the identities, like (2.9), hold only almost everywhere, (a.e.) is mostly omitted in the subsequent explanations.*
3. *The Bayes rule (2.8) is a simple consequence of previous formulas. Its importance in this text justifies the explicit presentation in various forms.*
4. *The symmetric identity (2.9) says that β does not influence the description of α if α and β are conditionally independent for a given γ.*

Often, a pdf $f(\alpha)$ of a multi-variate variable $\alpha \in \alpha^*$ is given and mapped on a variable β by a mapping $T : \alpha^* \to \beta^* \equiv T(\alpha^*)$. We need to specify the corresponding pdf $f(\beta)$. Rules for determining the pdf $f(\beta)$ are implied by the obvious need to preserve the expectation.

Proposition 2.5 (Pdfs of transformed quantities) *Let the expectation $\mathcal{E}$, acting on functions $\mathcal{B} : \beta^* \to (-\infty,\infty)$, be specified by the pdf $f_T(\beta)$, i.e.*

$$\mathcal{E}_T[\mathcal{B}] = \int \mathcal{B}(\beta) f_T(\beta)\, d\beta.$$

This functional expresses the same expectation as $\mathcal{E}[\mathcal{B}] = \int \mathcal{B}(T(\alpha)) f(\alpha)\, d\alpha$ iff

$$\int_{T(A)} f_T(T(\alpha))\, dT(\alpha) = \int_A f(\alpha)\, d\alpha \tag{2.10}$$

for all measurable sets $A \subset \alpha^*$.

Let α be a real vector , $\alpha \equiv [\alpha_1,\ldots,\alpha_{\mathring{\alpha}}]$ and $T = [T_1,\ldots,T_{\mathring{\alpha}}]$ bijection (one-to-one mapping) with finite continuous partial derivatives a.e. on α^*

$$J_{ij}(\alpha) \equiv \frac{\partial T_i(\alpha)}{\partial \alpha_j},\ i,j = 1,\ldots,\mathring{\alpha}, \tag{2.11}$$

for all entries T_i of T and entries α_j of α. Then,

$$f_T(T(\alpha))|J(\alpha)| = f(\alpha), \quad \text{where} \tag{2.12}$$

$|\cdot|$ *denotes* determinant *of the matrix in its argument.*

Proof: Proposition describes substitutions in multi-variate integrals, see [13, 14]. ♢

It is useful to summarize basic properties of expectation, which help us to simplify formal manipulations.

Proposition 2.6 (Basic properties of $\mathcal{E}$) *For arbitrary functions $\mathcal{Z}_1(\cdot)$, $\mathcal{Z}_2(\cdot)$ on which the conditional expectation $\mathcal{E}[\cdot|\gamma]$ is well defined, $\mathcal{E}[\cdot|\gamma]$ has the following properties.*

Isotonic nature of $\mathcal{E}[\cdot|\gamma]$: $\mathcal{Z}_1 \leq \mathcal{Z}_2, cf.\ (2.3), \Rightarrow \mathcal{E}[\mathcal{Z}_1|\gamma] \leq \mathcal{E}[\mathcal{Z}_2|\gamma]$.

Linearity of $\mathcal{E}[\cdot|\gamma]$: $\mathcal{E}[A(\gamma)\mathcal{Z}_1 + B(\gamma)\mathcal{Z}_2|\gamma] = A(\gamma)\mathcal{E}[\mathcal{Z}_1|\gamma] + B(\gamma)\mathcal{E}[\mathcal{Z}_2|\gamma]$
for arbitrary coefficients A, B depending at most on the condition γ.

Chain rule for expectation: $\mathcal{E}\left[\mathcal{E}[\cdot|\gamma,\zeta]|\gamma\right] = \mathcal{E}[\cdot|\gamma]$ *for an arbitrary additional condition ζ.*

Conditional covariance *of a vector α* $\operatorname{cov}[\alpha|\gamma] \equiv \mathcal{E}\left[(\alpha - \mathcal{E}[\alpha|\gamma])(\alpha - \mathcal{E}[\alpha|\gamma])'|\gamma\right]$ *is related to the non-central moments through the formula*

$$\operatorname{cov}[\alpha|\gamma] = \mathcal{E}[\alpha\alpha'|\gamma] - \mathcal{E}[\alpha|\gamma]\mathcal{E}[\alpha'|\gamma], \quad ' \text{ is transposition.} \tag{2.13}$$

Jensen inequality *bounds expectation of a convex function $T_\gamma : \alpha^* \to (-\infty, \infty)$*

$$\mathcal{E}[T_\gamma(\alpha)|\gamma] \geq T_\gamma\left(\mathcal{E}[\alpha|\gamma]\right). \tag{2.14}$$

Proof: All statements can be verified by using the integral expression (2.6) of the expectation. Proof of Jensen inequality can be found e.g. in [15]. ♢

Remark(s) 2.4

1. *Proposition is formulated for the conditional expectation. The unconditional case is formally obtained by omitting the condition used.*
2. *Note that whenever the expectation is applied to a vector (matrix) function V it should be understood as the vector (matrix) of expectations*

$$[\mathcal{E}(V)]_i \equiv \mathcal{E}(V_i). \tag{2.15}$$

2.3.3 Basic decision-making lemma

The optimal selection of *admissible decision rules* relies on the following key proposition that converts minimization over mappings to an "ordinary" minimization.

Proposition 2.7 (Basic decision-making lemma) *The optimal* admissible decision rule $\mathcal{R}^o$

$$\mathcal{R}^o(\mathcal{P}_{a^*}) \equiv a^o(\mathcal{P}_{a^*}),\ \forall \mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$$

minimizing the expected loss (2.6) can be constructed value-wise as follows.

To each $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$, a minimizing argument $a^o(\mathcal{P}_{a^*})$ in

$$\min_{a \in a^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})|a, \mathcal{P}_{a^*}] \tag{2.16}$$

is assigned as the value of the optimal decision rule corresponding to the considered argument. The reached minimum is

$$\min_{\{\mathcal{R}:\, \mathcal{P}^*_{a^*} \to a^*\}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})] = \mathcal{E}\left\{\min_{a \in a^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})|a, \mathcal{P}_{a^*}]\right\}. \tag{2.17}$$

Proof: Let us fix an arbitrary $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$. The definition of minimum implies

$$\mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, \mathcal{R}^o(\mathcal{P}_{a^*}), \mathcal{F}_{a^*})|\mathcal{R}^o(\mathcal{P}_{a^*}), \mathcal{P}_{a^*}] \leq \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})|a, \mathcal{P}_{a^*}], \ \forall a \in a^*.$$

Let an admissible rule $\mathcal{R} : \mathcal{P}^*_{a^*} \to a^*$ assign a decision $a \in a^*$ to the considered $\mathcal{P}_{a^*}$. Then, the previous inequality can be written

$$\mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, \mathcal{R}^o(\mathcal{P}_{a^*}), \mathcal{F}_{a^*})|\mathcal{R}^o(\mathcal{P}_{a^*}), \mathcal{P}_{a^*}] \leq \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*}, \mathcal{R}(\mathcal{P}_{a^*}), \mathcal{F}_{a^*})|\mathcal{R}(\mathcal{P}_{a^*}), \mathcal{P}_{a^*}].$$

We apply unconditional expectation $\mathcal{E}[\cdot]$ acting on functions of $\mathcal{P}_{a^*}$ to this inequality. Due to the isotonic nature of $\mathcal{E}$, the inequality is preserved. The chain rule for expectations, see Proposition 2.6, implies that on the left hand side of the resulting inequality the value of the unconditional expected loss for $\mathcal{R}^o$ is found. $\mathcal{R}^o$ assigns to each $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$ the decision $a^o(\mathcal{P}_{a^*})$. On the right hand side of the discussed inequality, we get the unconditional expected loss corresponding to an arbitrarily chosen $\mathcal{R} : \mathcal{P}^*_{a^*} \to a^*$. ♢

Proposition and its proof imply no preferences if there are more absolutely minimizing arguments $a^o(\mathcal{P}_{a^*})$. We can use any of them or switch between them in a random manner whenever the pdf $f(a_t|\mathcal{P}_{a^*_t})$ has its support concentrated on them. This is an example where a *randomized causal strategy* may occur. We specify it formally as it is used later on more extensively.

Agreement 2.6 (Outer model of randomized decision strategy) *The pdf $f(a|\mathcal{P}_{a^*})$ is called* outer model of the decision rule. *The collection of pdfs $\left\{f(a_t|\mathcal{P}_{a^*_t})\right\}_{t\in t^*}$ forms the* outer model of the decision strategy.

A decision rule $f(a|\mathcal{P}_{a^*})$ is called randomized decision rule *if its support contains at least two different values of a_t. The strategy is called* randomized strategy *if some of its rules is randomized.*

Remark(s) 2.5

1. *We do not enter the technical game with ε-optimum: the existence of the various minimizing arguments is implicitly supposed.*

2. *It is worth repeating that the optimal decision rule is constructed value-wise. Formally, the minimization should be performed for all possible instances of experience $\mathcal{P}_{a^*} \in \mathcal{P}^*_{a^*}$ in order to get the decision rule.*

   Often, we are interested in the optimal decision for a given fixed, say observed, experience. Then, just a single minimization is necessary. This is typically the case of the estimation problem. This possibility makes the main distinction from the dynamic design, when optimal strategy, a sequence of decision rules, is searched for. In this case, discussed in next Section, the construction of decision rules is necessary. This makes the dynamic design substantially harder and mostly infeasible [5, 6].

2.4 Dynamic design

We are searching for the optimal admissible strategy. We consider the usual case, when its individual rules use a non-decreasing sequence of experience. The extending experience models increasing amount of data available for the decision-making.

2.4.1 Dynamic programming

The optimal admissible strategy can be found by using a stochastic version of celebrated *dynamic programming* [16]. It is nothing but a repetitive application of Proposition 2.7.

Proposition 2.8 (Stochastic dynamic programming) *The optimal causal strategy* $\{\mathcal{R}_t^o : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*} \in \{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}$ *acting on non-decreasing sequence of experience* $\mathcal{P}_{a_t^*} \subset \mathcal{P}_{a_{t+1}^*}$ *and minimizing the expected loss function* $\mathcal{E}[\mathcal{Z}(\mathcal{Q})]$ *can be constructed in a value-wise way. For every* $t \in t^*$ *and each* $\mathcal{P}_{a_t^*} \in \mathcal{P}^*_{a_t^*}$*, it is sufficient to take a minimizing argument* $a^o(\mathcal{P}_{a_t^*})$ *in*

$$\mathcal{V}(\mathcal{P}_{a_t^*}) = \min_{a_t \in a_t^*} \mathcal{E}[\mathcal{V}(\mathcal{P}_{a_{t+1}^*})|a_t, \mathcal{P}_{a_t^*}],\ t \in t^* \tag{2.18}$$

as the decision generated by the t*th rule of the optimal strategy, i.e.* $a^o(\mathcal{P}_{a_t^*}) = \mathcal{R}_t^o(\mathcal{P}_{a_t^*})$.

The recursion (2.18) is performed in the backward manner against the course given by the increasing experience. It starts with

$$\mathcal{V}(\mathcal{P}_{a_{\mathring{t}+1}^*}) \equiv \mathcal{E}[\mathcal{Z}(\mathcal{Q})|\mathcal{P}_{a_{\mathring{t}+1}^*}]. \tag{2.19}$$

The reached minimum has the value $\mathcal{E}[\mathcal{V}(\mathcal{P}_{a_1^*})] = \min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{Q})]$.

Proof: Let us define $\mathcal{P}_{a_{\mathring{t}+1}^*} \equiv \mathcal{Q}$ in order to get a uniform notation It is legitimate as $a_{\mathring{t}+1}$ is not optimized. The definition of minimum and Proposition 2.7 imply

$$\min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\mathring{t}+1}^*})] = \min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t<\mathring{t}}} \left\{ \min_{\mathcal{R}_{\mathring{t}} : \mathcal{P}^*_{a_{\mathring{t}}^*} \to a_{\mathring{t}}^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\mathring{t}+1}^*})] \right\} \underbrace{=}_{(2.17)}$$

$$= \min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t<\mathring{t}}} \mathcal{E}\left[\min_{a_{\mathring{t}} \in a_{\mathring{t}}^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\mathring{t}+1}^*})|a_{\mathring{t}}, \mathcal{P}_{a_{\mathring{t}}^*}] \right].$$

Denoting $\mathcal{V}(\mathcal{P}_{a_{\mathring{t}}^*}) \equiv \min_{a_{\mathring{t}} \in a_{\mathring{t}}^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\mathring{t}+1}^*})|a_{\mathring{t}}, \mathcal{P}_{a_{\mathring{t}}^*}]$, we proved the first step of the recursion and specified the start (2.19). The following step becomes

$$\min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t<\mathring{t}}} \mathcal{E}\left[\mathcal{V}(\mathcal{P}_{a_{\mathring{t}}^*})\right].$$

We face the identical situation as above with the horizon decreased by one. Thus, the procedure can be repeated till the initial rule $\mathcal{R}_1$ is constructed. ♢

The optimization relies on our ability to evaluate the expectations

$$\mathcal{E}[\mathcal{V}(\mathcal{P}_{a_{t+1}^*})|a_t, \mathcal{P}_{a_t^*}] = \int_{\Delta_t^*} \mathcal{V}(\mathcal{P}_{a_t^*}, a_t, \Delta_t) f(\Delta_t|a_t, \mathcal{P}_{a_t^*})\, d\Delta_t,\ \forall t \in t^*.$$

The introduced *innovation* Δ_t contains those observable quantities that cannot be used for the choice of a_t. They are not in $\mathcal{P}_{a_t^*}$ but they belong to $\mathcal{P}_{a_{t+1}^*} \setminus a_t$. The pdfs $\left\{f(\Delta_t|a_t, \mathcal{P}_{a_t^*})\right\}_{t\in t^*}$ model the relationships of Δ_t to $\mathcal{P}_t$ and a_t.

Agreement 2.7 (Outer model of the system) *The collection of pdfs*

$$\left\{f(\Delta_t|a_t, \mathcal{P}_{a_t^*})\right\}_{t\in t^*}, \tag{2.20}$$

required for the optimal design, is called outer model of the system.

Remark(s) 2.6

1. *The term outer model of the system is abbreviated to the model of the system or even to the model. The exact meaning is clear from the specific context.*
2. *Often, the innovation $\Delta_t = y_t =$ observable output of the system. The set-point to which the output should be driven by the chosen input u_t has to be included into Δ_{u_t} if its values are uncertain.*

The following agreement is used in a presentation of the most common version of dynamic programming.

Agreement 2.8 (Data driven design; internal quantities) *The design is called* data driven *if realization $\mathcal{Q}$ consists of optional and potentially observable quantities, i.e. with ignorance consisting of unobserved data only*

$$\mathcal{Q} \equiv (\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)) \equiv (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*}). \tag{2.21}$$

Quantities in $\mathcal{Q}$ that are never observed directly are called internal quantities.

Proposition 2.9 (Dynamic programming for additive loss function) *Let us consider data driven design and search for the optimal admissible strategy $\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}$ acting on a non-decreasing sequence of experience $\{\mathcal{P}_{a_t^*}\}_{t\in t^*}$.*

Then, the optimal strategy $\{\mathcal{R}_t^o : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}$ minimizing the expected additive loss function

$$\mathcal{E}\left[\sum_{t\in t^*} z(\Delta(t), a(t))\right] \tag{2.22}$$

can be constructed value-wise. For all $t \in t^*$ and to each $\mathcal{P}_{a_t^*} \in \mathcal{P}^*_{a_t^*}$, a minimizing argument $a^o(\mathcal{P}_{a_t^*}) = \mathcal{R}_t^o(\mathcal{P}_{a_t^*})$ in

$$\mathcal{V}(\mathcal{P}_{a_t^*}) = \min_{a_t \in a_t^*} \mathcal{E}[z(\Delta(t), a(t)) + \mathcal{V}(\mathcal{P}_{a^*_{t+1}})|a_t, \mathcal{P}_{a_t^*}],\ t \in t^* \tag{2.23}$$

is assigned. The recursion (2.23) is performed in the backward manner against the course given by the increasing experience, starting from

$$\mathcal{V}(\mathcal{P}_{a^*_{t+1}}) \equiv 0. \tag{2.24}$$

The reached minimum has the value

$$\min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a^*_{t+1}})] = \mathcal{E}[\mathcal{V}(\mathcal{P}_{a_1^*})].$$

Proof: It follows exactly the line of Proposition 2.8 with a modified definition of the function $\mathcal{V}(\cdot)$

$$\mathcal{V}(\mathcal{P}_{a_t^*}) \equiv \min_{\{\mathcal{R}_\tau : \mathcal{P}^*_{a_\tau^*} \to a_\tau^*\}_{\tau\geq t}} \mathcal{E}\left[\sum_{\tau\geq t} z(\Delta(\tau), a(\tau))|a_t = \mathcal{R}_t(\mathcal{P}_{a_t^*}), \mathcal{P}_{a_t^*}\right]. \tag{2.25}$$

♢

Agreement 2.9 (Partial loss; Bellman function) *The function $z(\Delta(t), a(t))$ is called the* partial loss. *The function $\mathcal{V}(\cdot)$ occurring in dynamic programming is called* Bellman function. *Bellman function $\mathcal{V}(\cdot)$ in (2.25) is also called the* optimal loss-to-go.

2.4.2 Fully probabilistic design

A specific but rather general design is formulated and solved here. It is believed to form a bridge between optimal and practically optimal design.

Here, and in many other places, we need the notion of the *Kullback-Leibler divergence* [17] that measures well proximity of a pair of pdfs.

Agreement 2.10 (Kullback-Leibler divergence) *Let f, g be a pair of pdfs acting on a common set x^*. Then, the Kullback-Leibler divergence $\mathcal{D}(f||g)$ is defined by the formula*

$$\mathcal{D}(f||g) \equiv \int_{x^*} f(x) \ln \left(\frac{f(x)}{g(x)} \right) dx. \tag{2.26}$$

For conciseness, the Kullback-Leibler divergence is referred to as the KL divergence.

Proposition 2.10 (Basic properties of KL divergence)
Let f, g be a pair of pdfs acting on a same set. It holds

1. $\mathcal{D}(f||g) \geq 0$,
2. $\mathcal{D}(f||g) = 0$ *iff* $f = g$ *(a.e.),*
3. $\mathcal{D}(f||g) = \infty$ *iff on a set of a positive measure* $f > 0$ *and* $g = 0$,
4. $\mathcal{D}(f||g) \not\equiv \mathcal{D}(g||f)$*; the KL divergence does not obey triangle inequality.*

Proof: See, for instance, [15] ♢

Now we are ready to formulate and solve the fully probabilistic design problem. A simple version is presented here considering the data-driven design, see Agreement 2.8. In this case, the joint pdf $f(\mathcal{Q}) \equiv f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right))$ describing observable quantities of interest can be factorized by a repetitive use of the chain rule, see Proposition 2.4,

$$f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)) = \prod_{t \in t^*} f(\Delta_t | a_t, \mathcal{P}_{a_t^*}) f(a_t | \mathcal{P}_{a_t^*}). \tag{2.27}$$

The first factors $\left\{ f(\Delta_t | a_t, \mathcal{P}_{a_t^*}) \right\}_{t \in t^*}$ under the product sign describe possible reactions of the system on the decision a_t under the experience $\mathcal{P}_{a_t^*}$. These pdfs form the outer model of the system, see Agreement 2.7. Similarly, $\left\{ f(a_t | \mathcal{P}_{a_t^*}) \right\}_{t \in t^*}$ represent an outer model of a randomized decision strategy to be chosen, see Agreement 2.6. Looking at the joint pdf (2.27), it seems to be "natural" to formulate the design as an attempt to make this pdf as close as possible to some "ideal" joint pdf.

Agreement 2.11 (Fully probabilistic design) *The fully probabilistic, data driven, design specifies its target through an* ideal pdf

$$\lfloor^I f(\mathcal{Q}) = \prod_{t \in t^*} \lfloor^I f(\Delta_t | a_t, d(t-1)) \lfloor^I f(a_t | d(t-1)).$$

The optimal admissible, possibly randomized, decision strategy is defined as a minimizer of the KL divergence (2.26) of $f(d\left(\mathring{t}\right)) = f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right))$ *and* $\lfloor^I f(d\left(\mathring{t}\right)) = \lfloor^I f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right))$

$$\mathcal{D}(f||\lfloor^I f) \equiv \int f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)) \ln\left(\frac{f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right))}{\lfloor^I f(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right))}\right) d(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)). \tag{2.28}$$

Proposition 2.11 (Solution to fully probabilistic design) *The optimal strategy minimizing KL divergence (2.28) has the form*

$$\begin{aligned} f(a_t|d(t-1)) &= \lfloor^I f(a_t|d(t-1)) \frac{\exp[-\omega_\gamma(a_t, d(t-1))]}{\gamma(d(t-1))}, \quad \textit{where} \qquad (2.29)\\ \gamma(d(t-1)) &\equiv \int_{a_t^*} \lfloor^I f(a_t|d(t-1)) \exp[-\omega(a_t, d(t-1))]\, da_t \\ \omega_\gamma(a_t, d(t-1)) &\equiv \int_{\Delta_t^*} f(\Delta_t|a_t, d(t-1)) \ln\left(\frac{f(\Delta_t|a_t, d(t-1))}{\gamma(d(t))\, \lfloor^I f(\Delta_t|a_t, d(t-1))}\right) d\,\Delta_t \\ \gamma(d\left(\mathring{t}\right)) &= 1. \qquad (2.30) \end{aligned}$$

The solution is performed against the time course, starting at $t = \mathring{t}$.

Proof: Using chain rule, we can write the KL divergence in the form

$$\mathcal{D}\left(f||\lfloor^I f\right) = \mathcal{E}\left\{\sum_{t\in t^*} \int f(a_t|d(t-1)) \left[\ln\left(\frac{f(a_t|d(t-1))}{\lfloor^I f(a_t|d(t-1))}\right) + \omega(a_t, d(t-1))\right] da_t\right\} \text{ with}$$

$\omega(a_t, d(t-1)) \equiv \int f(\Delta_t|a_t, d(t-1)) \ln\left(\frac{f(\Delta_t|a_t, d(t-1))}{\lfloor^I f(\Delta_t|a_t, d(t-1))}\right) d\Delta_t$. Let us denote $\ln(\gamma(d(t))) \equiv$

$$\equiv \min_{\{f(a_{\tau+1}|d(\tau))\}_{\tau=t}^{\mathring{t}}} \mathcal{E}\left\{\sum_{\tau=t+1}^{\mathring{t}} \int f(a_\tau|d(\tau-1)) \left[\ln\left(\frac{f(a_\tau|d(\tau-1))}{\lfloor^I f(a_\tau|d(\tau-1))}\right) + \omega(a_\tau, d(\tau-1))\right] da_\tau|d(t)\right\}.$$

Then, this definition implies that $\gamma(d\left(\mathring{t}\right)) = 1$ and

$$-\ln(\gamma(d(t))) \equiv \min_{f(a_{t+1}|d(t))} \int f(a_{t+1}|d(t)) \left[\ln\left(\frac{f(a_{t+1}|d(t))}{\lfloor^I f(a_{t+1}|d(t))}\right) + \omega_\gamma(a_{t+1}, d(t))\right] da_{t+1} \text{ with}$$

$\omega_\gamma(a_{t+1}, d(t)) \equiv \int f(\Delta_t|a_t, d(t-1)) \ln\left(\frac{f(\Delta_t|a_t, d(t-1))}{\gamma(d(t))\, \lfloor^I f(\Delta_t|a_t, d(t-1))}\right) d\Delta_t$. It implies

$$-\ln(\gamma(d(t))) \equiv \min_{f(a_{t+1}|d(t))} \int f(a_{t+1}|d(t))$$

$$\left[\ln\left(\frac{f(a_{t+1}|d(t))}{\frac{\lfloor^I f(a_{t+1}|d(t)) \exp[-\omega_\gamma(a_{t+1}, d(t))]}{\int \lfloor^I f(\tilde{a}_{t+1}|d(t)) \exp[-\omega_\gamma(\tilde{a}_{t+1}, d(t))]\, d\tilde{a}_{t+1}}}\right) da_{t+1} - \ln\left(\int \lfloor^I f(a_{t+1}|d(t)) \exp\left[-\omega_\gamma(a_{t+1}, d(t))\right] da_{t+1}\right)\right].$$

The first term in the above identity is the KL divergence, that reaches its smallest zero value for the claimed pdf. At the same time, it defines the form of the reached minima. An alternative derivations, with more details, see [18, 19]. ♢

Remark(s) 2.7

1. *At a descriptive level, the stochastic dynamic programming consists of a sequence of the evaluation pairs*

   $$(\textit{take conditional expectation, minimize}).$$

   Except of a few numerically solvable cases, some approximation techniques have to be employed. The complexity of the approximated optimum prevents a systematic use of the standard approximation theory and various ad hoc schemes are adopted.

   The fully probabilistic design finds minimizers explicitly and reduces the design to a sequence of conceptually feasible multi-variate integrations.

2. *The found optimal strategy is randomized and obviously causal one. The technological restriction are met trivially if the chosen ideal strategy meets them, if* $\operatorname{supp}\left[{}^{\lfloor I} f(a_t|\mathcal{P}_{a_t^*})\right] \subset a_t^*$, *cf. (2.29).*

2.4.3 Asymptotic of design

The analysis outlined here serves us mainly for interpretation purposes. Thus, all technicalities are suppressed as much as possible.

The asymptotic of the dynamic programming is analyzed for horizon $\mathring{t} \to \infty$ within this section. The data-driven case with an additive loss function (2.22) is considered. Note that the general loss function can always be converted in it by defining

$$z(\Delta(t), a(t))) = \begin{cases} \mathcal{Z}(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)) & \text{if } t = \mathring{t} \\ 0 & \text{otherwise} \end{cases}. \tag{2.31}$$

We deal, however, with a simpler but still useful case by assuming that

- there is a finite-dimensional information state, i.e. $\mathcal{P}_{a_t^*} \equiv x_{t-1} \equiv$ a finite-dimensional vector,
- the partial loss depends on x_t and a_t only $z(\Delta(t), a(t)) \equiv z(x_t, a_t)$, i.e. the considered loss is

  $$\mathcal{Z}(\Delta\left(\mathring{t}\right), a\left(\mathring{t}\right)) = \sum_{t \in t^*} z(x_t, a_t).$$

Agreement 2.12 (Stabilizing strategy) *Let us consider sequence of decision-making problems with the growing horizon $\mathring{t} \to \infty$. The strategy $\{\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to a_t^*\}_{t \in t^* \equiv \{1,\ldots,\mathring{t}\}}$ $\forall \mathring{t} \in \{1, 2, \ldots\}$ is called* stabilizing strategy *if there is a finite constant c such that*

$$\mathcal{E}[z(x_t, a_t)|a_t, \mathcal{P}_{a_t^*}] \leq c < \infty, \ t \in \{1, 2, 3, \ldots\}. \tag{2.32}$$

Proposition 2.12 (Asymptotic design) *Let a stabilizing strategy exist. Then there is an optimal, stationary strategy, formed by a repetitive use of the same rule, for $\mathring{t} \to \infty$ whose decisions are minimizing arguments in the formal analogy of (2.23)*

$${}^{\lfloor \infty}\mathcal{V}(x_{t-1}) + C = \min_{a_t \in a_t^*} \mathcal{E}\left[z(x_t, a_t) + {}^{\lfloor \infty}\mathcal{V}(x_t)|a_t, x_{t-1}\right]. \tag{2.33}$$

with a constant $C \leq c$ and a time invariant Bellman function ${}^{\lfloor \infty}\mathcal{V}(x)$.

Proof: (Outline) Let us take any finite horizon $\mathring{t}$. We denote ${}^{\lfloor\mathring{t}}\mathcal{E}$ the expectation taken for the strategy $\{{}^{\lfloor\mathring{t}}\mathcal{R}_t\}_{t\in t^*}$ optimal for the additive loss specified on this horizon. Similarly, ${}^{\lfloor\mathring{t}}\tilde{\mathcal{V}}(\mathcal{P}_{a_t^*}) \equiv {}^{\lfloor\mathring{t}}\tilde{\mathcal{V}}(x_{t-1})$ denotes the optimal loss-to-go, see Agreement 2.9 and ${}^{\lfloor\mathring{t}}a_t = {}^{\lfloor\mathring{t}}\mathcal{R}_t(x_{t-1})$.

The definition of optimum, the additive form of the loss (2.22) and the assumption (2.32) imply that

$$c \geq {}^{\lfloor\mathring{t}}C \equiv {}^{\lfloor\mathring{t}}\mathcal{E}[\mathcal{Z}(\Delta\left(\mathring{t}\right), {}^{\lfloor\mathring{t}}a\left(\mathring{t}\right))] - {}^{\lfloor\mathring{t}-1}\mathcal{E}[\mathcal{Z}(\Delta(\mathring{t}-1), {}^{\lfloor\mathring{t}-1}a(\mathring{t}-1))] \geq 0.$$

Consequently,
$$\limsup_{\mathring{t}\to\infty}\left\{{}^{\lfloor\mathring{t}}\mathcal{E}[\mathcal{Z}(\Delta\left(\mathring{t}\right), {}^{\lfloor\mathring{t}}a\left(\mathring{t}\right))] - \sum_{t=1}^{\mathring{t}} {}^{\lfloor t}C\right\} \equiv {}^{\lfloor\infty}\mathcal{V} < \infty.$$

If we redefine the partial loss to $z(x_t, a_t) - {}^{\lfloor t}C$ then the optimal strategies for all $\mathring{t}$ do not change and ${}^{\lfloor\mathring{t}}\mathcal{E}\left[\mathcal{Z}\left(\Delta\left(\mathring{t}\right), {}^{\lfloor\mathring{t}}a\left(\mathring{t}\right)\right)\right] = {}^{\lfloor\infty}\mathcal{V} + o\left(\mathring{t}\right)$, where the symbols $o\left(\mathring{t}\right)$ means, as usual, a term converging to zero for $\mathring{t}\to\infty$.

Similar considerations apply to the modified optimal loss-to-go ${}^{\lfloor\mathring{t}}\mathcal{V}(x_t)$, Agreement 2.9, imply that (with the modified partial loss!) it converges for all $x_t \in x_t^*$ to some ${}^{\lfloor\infty}\mathcal{V}(x_t)$.

With the original partial loss and modified loss-to-go, the dynamic programming (2.23) gets the form

$${}^{\lfloor\mathring{t}}\mathcal{V}(x_{t-1}) + {}^{\lfloor t}C = \min_{a_t\in a_t^*} {}^{\lfloor\mathring{t}}\mathcal{E}\left[z(x_t, a_t) + {}^{\lfloor\mathring{t}}\mathcal{V}(x_t)|a_t, x_{t-1}\right] \text{ that converges for } \mathring{t}\to\infty.$$

♢

Remark(s) 2.8

1. *The same proof is directly applicable to the fully probabilistic design as it can be seen as an instance of the additive loss function.*
2. *Solutions to the Bellman equation obtained for a growing finite horizon $\mathring{t}$ can be interpreted as successive approximations for solving its stationary counterpart (2.33).*
3. *So called* iterations in strategy space *[20] is an alternative and efficient way of finding the asymptotic solution. Essentially, a stabilizing stationary, i.e. repeating a single rule, strategy $\{\mathcal{R}\}$ is selected and the linear equation*

   $$\mathcal{V}(x) + C = \mathcal{E}[z(\tilde{x}, \mathcal{R}(x)) + \mathcal{V}(\tilde{x})|\mathcal{R}(x), x]$$

   is solved for the function $\mathcal{V}(\cdot)$ and constant C. Then, a new approximating strategy is found value-wise $\mathcal{R}(x) = \mathrm{Arg}\min_{a\in a^*}\mathcal{E}[z(\tilde{x}, a)+\mathcal{V}(\tilde{x})|a, x]$ with such a $\mathcal{V}(\cdot)$. Under general conditions, the newly found strategy is stabilizing and iterations may be repeated till the guaranteed convergence. Details of this procedure are out of scope of this work but it should be considered when searching for efficient numerical procedures.

2.5 Learning

Generally, considered behavior $\mathcal{Q}^*$ contains quantities that are never observed directly and, in spite of this, we want to describe them or influence them. We called them internal quantities, see Agreement 2.8. Even in this case, the optimal decision-making needs the outer model (2.20), see Proposition 2.8, that should reflect them. Here we describe how to get it. The presented construction is of an independent interest as a good *formal model of learning.*

2.5.1 Bayesian filtration

This section deals with a general case when the construction of the outer model relies on filtering, i.e. learning of *time variant* internal, i.e. directly unobservable, quantities. It requires the following elements.

Requirement 2.5 (Models; natural conditions of decision-making)

1. *The innovations* Δ_t *are related to experience* $\mathcal{P}_{a_t^*}$ *and decisions* a_t *through the* observation model
$$\{f(\Delta_t|a_t, \mathcal{P}_{a_t^*}, \Theta_t)\}_{t\in t^*} \tag{2.34}$$
that is given up to unknown internal quantities $\Theta_t \in \Theta_t^* \subset \mathcal{F}_{a_\tau^*},\ \forall\tau \in t^*$.

2. *The evolution of the quantities* $\Theta\left(\mathring{t}\right) \in \Theta^*\left(\mathring{t}\right)$ *is described by a known collection of pdfs called the* time evolution model
$$\left\{f(\Theta_t|a_t, \mathcal{P}_{a_t^*}, \Theta_{t-1})\right\}_{t\in t^*}. \tag{2.35}$$

3. *The quantities* $\Theta\left(\mathring{t}\right)$ *are unknown to the strategies considered. The* natural conditions of decision-making *(a slight generalization of natural conditions of control [10]) express it formally. They postulate independence of* a_t *and* Θ_t *when conditioned on* $\mathcal{P}_{a_t^*}$
$$f(a_t|\mathcal{P}_{a_t^*}, \Theta_t) = f(a_t|\mathcal{P}_{a_t^*}) \underbrace{\Leftrightarrow}_{\text{Proposition 2.4}} f(\Theta_t|a_t, \mathcal{P}_{a_t^*}) = f(\Theta_t|\mathcal{P}_{a_t^*}). \tag{2.36}$$

4. *The initial values of* Θ_0, $\mathcal{P}_{a_1^*}$ *add nothing new to the prior information so that the* prior pdf
$$f(\Theta_1) \equiv f(\Theta_1|a_1, \mathcal{P}_{a_1^*}, \Theta_0) = f(\Theta_1|a_1, \mathcal{P}_{a_1^*}) \underbrace{=}_{(2.36)} f(\Theta_1|\mathcal{P}_{a_1^*}) \tag{2.37}$$
is the first term in the sequence of pdfs (2.35).

Remark(s) 2.9

1. *Often, the unknown quantities* Θ_t *together with the decision* a_t *are assumed to describe the involved conditional pdfs fully. Then,* $\mathcal{P}_{a_t^*}$ *can be omitted and* Θ_t *can be identified with the* information state.

2. *The natural conditions of decision-making express the assumption that* $\Theta_t \notin \mathcal{P}_{a_\tau^*}\ \forall\tau \in t^*$. *Thus, values of* Θ_t *cannot be used by the decision rules forming admissible strategy. Alternatively, see (2.36), we cannot gain information about* Θ_t *from the decision* a_t *if the corresponding innovation* Δ_t *(the corresponding reaction of the system) is not available.*

   The natural conditions of decision-making are "naturally" fulfilled by strategies we are designing. They have to be checked when data influenced by an "externally chosen" strategy are processed.

Proposition 2.13 (Generalized Bayesian filtering) *Let Requirement 2.5 be met. Then, the outer model of the system (2.20) is given by the formula*

$$f(\Delta_t|a_t,\mathcal{P}_{a_t^*}) = \int_{\Theta_t^*} f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta_t) f(\Theta_t|\mathcal{P}_{a_t^*})\, d\Theta_t. \tag{2.38}$$

The evolution of the pdf $f(\Theta_t|\mathcal{P}_{a_t^*})$, called (generalized Bayesian) filtration of unknown quantities *Θ_t, is described by the following recursion that starts from the prior pdf $f(\Theta_1)$:*

- Data updating

$$f(\Theta_t|\mathcal{P}_{a_{t+1}^*}) = \frac{f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta_t) f(\Theta_t|\mathcal{P}_{a_t^*})}{f(\Delta_t|a_t,\mathcal{P}_{a_t^*})} \propto f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta_t) f(\Theta_t|\mathcal{P}_{a_t^*}) \tag{2.39}$$

  that incorporates the innovation Δ_t and the decision a_t, and

- Time updating

$$f(\Theta_{t+1}|\mathcal{P}_{a_{t+1}^*}) = \int_{\Theta_t^*} f(\Theta_{t+1}|a_{t+1},\mathcal{P}_{a_{t+1}^*},\Theta_t) f(\Theta_t|\mathcal{P}_{a_{t+1}^*})\, d\Theta_t \tag{2.40}$$

  that reflects the time evolution $\Theta_t \to \Theta_{t+1}$.

Proof: Sequential use of marginalization, chain rule, Proposition 2.4, and of the natural conditions of decision making (2.36) implies (2.38)

$$\begin{aligned} f(\Delta_t|a_t,\mathcal{P}_{a_t^*}) &= \int_{\Theta_t^*} f(\Delta_t,\Theta_t|a_t,\mathcal{P}_{a_t^*})\, d\Theta_t = \int_{\Theta_t^*} f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta_t) f(\Theta_t|a_t,\mathcal{P}_{a_t^*})\, d\Theta_t = \\ &= \int_{\Theta_t^*} f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta_t) f(\Theta_t|\mathcal{P}_{a_t^*})\, d\Theta_t. \end{aligned}$$

Data updating coincides with the Bayes rule in which the outer model of the strategy cancels as it does not depend on Θ_t due to the natural conditions of decision making (2.36).

Marginalization, chain rule and natural conditions of decision making imply also the formula for time updating. ♢

Agreement 2.13 (Filtering; predictive pdf) *The process of generating filtration is called (generalized Bayesian)* filtering. *The outer model of the system obtained by filtering is called* predictive pdf.

Remark(s) 2.10

1. *The term* predictive pdf *reflects the way how the outer model of the system has been obtained. It uses the observed experience and extrapolates it into ignorance assuming that the* mechanism of generating Θ_t does not change.

   This accumulation of experience and its extrapolation represent a good formal model of learning.

2. *It has to be stressed that the accumulation of experience can take place only when the rules governing the behavior are not changed during it, i.e. when we can rely on the validity of the underlying models.*

3. *The construction of the predictive pdf is our key motivation for filtering. Its results are often of an independent interest.*

4. *Under the natural conditions of decision-making, the filtering relies on the knowledge of decisions and not on the knowledge of rules $\mathcal{R} : \mathcal{P}^*_{a^*} \to a^*$ generating them. It is practically important when we learn while decision loop is closed, especially, when decisions are made by a human decision maker.*

5. *The time evolution model $f(\Theta_t|\Theta_{t-1}, a_t, \mathcal{P}_{a_t^*})$ as well as the observation model $f(\Delta_t|\Theta_t, a_t, \mathcal{P}_{a_t^*})$ have to result from a theoretical modelling of the system in question. Such a modelling uses both field knowledge, like laws of conservation, and approximation capabilities of the selected family of the models involved. Often, deterministic relationships are modelled and then the "deviations" from an "expected" trajectory are described.*

6. *The prior pdf $f(\Theta_1)$ allows us to introduce an information based on expert knowledge or analogy to situations observed previously.*

7. *The observed data, the only bridge to reality, enter the evaluations in the data updating step only when the newest innovation-decision pair is processed. This simple observation is important for approximation of the time evolution model, see Section 4.2.2.*

8. *In summary, the described Bayesian filtering combines prior information in $f(\Theta_1)$, theoretical knowledge of the specific fields described by $f(\Delta_t|\Theta_t, a_t, \mathcal{P}_{a_t^*})$, $f(\Theta_t|\Theta_{t-1}, a_t, \mathcal{P}_{a_t^*})$ and observed data $d(\mathring{t}) = (\Delta(\mathring{t}), a(\mathring{t}))$ by using coherent deductive rules of the calculus with pdfs.*

   This combination of information sources is a powerful internally consistent framework describing the essence of learning. Due to its deductive structure, the incorrect modelling not an incorrect information processing can be blamed for a failure of the specific learning process.

2.5.2 Bayesian estimation

This section deals with a special version of filtering called *estimation*. It arises when the internal quantities Θ_t are time invariant $\Theta_t = \Theta$, $\forall t \in t^*$. The common value Θ is called *parameter*.

In this case, the time evolution model is $f(\Theta_t|\Theta_{t-1}, a_t, \mathcal{P}_{a_t^*}) = \delta(\Theta_t - \Theta_{t-1})$, where the employed Dirac delta function $\delta(\cdot)$ is a formal pdf of the measure fully concentrated on the zero argument.

The parameterized model, generally defined in (2.34), has the form

$$\{f(\Delta_t|a_t, \mathcal{P}_{a_t^*}, \Theta)\}_{t\in t^*}. \tag{2.41}$$

Proposition 2.14 (Generalized Bayesian estimation) *Let Requirement 2.5 be met with time invariant $\Theta_t = \Theta \in \Theta^* \subset \mathcal{F}_{a_\tau^*}$, $\forall \tau \in t^*$. Then, the outer model of the system (2.20) is given by the formula*

$$f(\Delta_t|a_t, \mathcal{P}_{a_t^*}) = \int_{\Theta^*} f(\Delta_t|a_t, \mathcal{P}_{a_t^*}, \Theta) f(\Theta|\mathcal{P}_{a_t^*})\, d\Theta. \tag{2.42}$$

The evolution of the pdf $f(\Theta|\mathcal{P}_{a_t^*})$, called (generalized Bayesian) parameter estimation *generating parameter estimate coinciding with the* posterior pdf *of* unknown parameter*, is described by the*

recursion identical with data updating (2.39)

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) = \frac{f(\Delta_t|a_t,\mathcal{P}_{a^*_t},\Theta)f(\Theta|\mathcal{P}_{a^*_t})}{f(\Delta_t|a_t,\mathcal{P}_{a^*_t})} \propto f(\Delta_t|a_t,\mathcal{P}_{a^*_t},\Theta)f(\Theta|\mathcal{P}_{a^*_t}). \tag{2.43}$$

It starts from the prior pdf $f(\Theta) \equiv f(\Theta|\mathcal{P}_{a^*_1},a_1) = f(\Theta|\mathcal{P}_{a^*_1})$.

The simplicity of the estimation formula allows us to write down its (non-recursive) batch variant

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) = \frac{\prod_{\tau\le t} f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau},\Theta)f(\Theta)}{\int_{\Theta^*}\prod_{\tau\le t} f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau},\Theta)f(\Theta)\,d\Theta} \equiv \frac{\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}})f(\Theta)}{\mathcal{I}(\mathcal{P}_{a^*_{t+1}})}. \tag{2.44}$$

The introduced likelihood function

$$\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}}) \equiv \prod_{\tau\le t} f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau},\Theta) \tag{2.45}$$

evolves according to the recursion identical with that for the posterior pdf (2.43) but it starts from the $\mathcal{L}(\Theta,\mathcal{P}_{a^*_1})$ *identically equal to 1.*

The normalization factor $\mathcal{I}(\cdot)$ *is defined by the formula*

$$\mathcal{I}(\mathcal{P}_{a^*_{t+1}}) = \int \mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}})f(\Theta)\,d\Theta \propto f(\Delta_t|a_t,\mathcal{P}_{a^*_t}). \tag{2.46}$$

With it, the outer model of the system (2.20) can alternatively be expressed

$$f(\Delta_t|a_t,\mathcal{P}_{a^*_t}) = \frac{\mathcal{I}(\mathcal{P}_{a^*_{t+1}})}{\mathcal{I}(\mathcal{P}_{a^*_t})}. \tag{2.47}$$

Proof: It is again a simple exercise in calculus with pdfs, marginalization, chain rule and Bayes rule, Proposition 2.4 with the natural conditions of decision-making (2.36). ♢

Remark(s) 2.11

1. *The observation model* $f(\Delta_t|a_t,\mathcal{P}_{a^*_t},\Theta)$ *is called* parameterized model *whenever estimation problem is considered. We respect this tradition.*

2. *Note that the recursive evolution of the pdf* $f(\Theta|\mathcal{P}_{a^*_t})$ *allows us to interpret the posterior pdf as the prior one before processing new observations.*

3. *The data inserted into the "objective" parameterized (observation) model correct gradually the subjectively chosen prior pdf* $f(\Theta)$. *The posterior pdf* $f(\Theta|\mathcal{P}_{a^*_t})$ *reflects always both ingredients. If the data are informative enough, the relative contribution of the single subjective factor* $f(\Theta)$ *to the posterior pdf is decreasing with increasing* t *as the likelihood function* $\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}})$ *contains* t *"objective" factors, cf. (2.45).*

4. *Zero values are preserved by multiplication. Thus, the posterior pdf re-distributes probability mass only within the support of the prior pdf, i.e. within the set*

$$\mathrm{supp}\,[\,f(\Theta)] \equiv \{\Theta\in\Theta^* :\ f(\Theta) > 0\}.$$

This fact allows us to introduce hard bounds on possible parameter values but prevents us to "learn" about parameters Θ *out of the support* $\mathrm{supp}\,[\,f(\Theta)]$, *see (2.44).*

5. *Remarks 2.10 related to filtering apply mostly to estimation, too. The parameter estimation is a task on its own; unknown parameters are always in the ignorance of the decision to be chosen; under the natural conditions of decision-making (2.36), decisions are needed and not the strategy* $\left\{\mathcal{R}_t : \mathcal{P}^*_{a^*_t} \to a^*_t\right\}_{t\in t^*}$ *generating them.*

6. *The parameters* Θ_t *are usually assumed to be finite-dimensional in order to avoid technicalities related to measure theory. In exceptional cases, like in description of so called equivalence approach, see Section 4.2.1, we deal with* potentially infinite-dimensional parameter. *It means that the number of unknown quantities is finite but increases without limitations. This case is often called* non-parametric estimation.

2.5.3 Asymptotic of estimation

The analysis outlined here serves us mainly for interpretation purposes. Thus, all technicalities are suppressed as much as possible.

The "objective" pdf $f(\mathcal{Q})$, see Section 2.3.1, describing the system behavior is denoted here ${}^{\lfloor o}f(\mathcal{Q})$. The corresponding outer model of the system $f(\Delta_t|a_t,\mathcal{P}_{a^*_t})$ is denoted ${}^{\lfloor o}f(\Delta_t|a_t,\mathcal{P}_{a^*_t})$. Its relationship to the predictive pdf $f(\Delta_t|a_t,\mathcal{P}_{a^*_t})$ obtained through the parameter estimation, Proposition 2.14, is inspected.

For the analysis, the notion of *entropy rate* $\mathcal{H}_\infty\left({}^{\lfloor o}f||\Theta\right)$ is needed. For each $\Theta\in\Theta^*$, it is defined by the formula

$$\mathcal{H}_\infty\left({}^{\lfloor o}f||\Theta\right) \equiv \lim_{t\to\infty}\sup\mathcal{H}_t\left({}^{\lfloor o}f||\Theta\right) \equiv \tag{2.48}$$
$$\equiv \lim_{t\to\infty}\sup\frac{1}{t}\sum_{\tau\le t}\int_{\Delta^*_\tau} {}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau})\ln\left(\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau})}\right)d\Delta_\tau.$$

Proposition 2.15 (Asymptotic of estimation) *Let the natural conditions of decision-making (2.36) hold and* $0<\underline{C}_\Theta\le\overline{C}_\Theta\le c<\infty$, $\bar{t}_\Theta\in\{1,2,\ldots\}$ *exist, for almost all* $\Theta\in\Theta^*$, *such that*

$$\underline{C}_\Theta f(\Delta_t|a_t,\mathcal{P}_{a^*_t},\Theta)\le {}^{\lfloor o}f(\Delta_t|a_t,\mathcal{P}_{a^*_t})\le\overline{C}_\Theta f(\Delta_t|a_t,\mathcal{P}_{a^*_t},\Theta),\ \forall t>\bar{t}_\Theta,\ \forall\mathcal{P}_{a^*_{t+1}}\in\mathcal{P}^*_{a^*_{t+1}}. \tag{2.49}$$

Then, the posterior pdf $f(\Theta|\mathcal{P}_{a^*_t})$ *(2.43) converges almost surely to a pdf* $f(\Theta|\mathcal{P}_{a^*_\infty})$. *It has the support coinciding with the set of minimizing arguments in*

$$\mathrm{supp}\left[f(\Theta|\mathcal{P}_{a^*_\infty})\right]=\mathrm{Arg}\min_{\Theta\in\mathrm{supp}[f(\Theta)]\cap\Theta^*}\mathcal{H}_\infty\left({}^{\lfloor o}f||\Theta\right). \tag{2.50}$$

Proof: Under the natural conditions of decision making (2.36), the posterior pdf (2.43) can be written in the form

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) \propto f(\Theta)\exp\left[-t\mathcal{H}(\mathcal{P}_{a^*_{t+1}},\Theta)\right] \text{ with} \tag{2.51}$$
$$\mathcal{H}(\mathcal{P}_{a^*_{t+1}},\Theta) = \frac{1}{t}\sum_{\tau\le t}\ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a^*_\tau},\Theta)}\right]. \tag{2.52}$$

This form exploits the fact that non-normalized posterior pdf can be multiplied by any factor independent of Θ.

Let us fix the argument $\Theta \in \Theta^*$ and define

$$e_{\Theta;\tau} \equiv \ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*},\Theta)}\right] - {}^{\lfloor o}\mathcal{E}\left[\ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*},\Theta)}\right]|a_\tau,\mathcal{P}_{a_\tau^*}\right]$$
$$\equiv \ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*},\Theta)}\right] - \int {}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})\ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*},\Theta)}\right] d\Delta_\tau.$$

A direct check reveals that the introduced deviations $e_{\Theta;\tau}$ of values

$$\ln\left[\frac{{}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})}{f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*},\Theta)}\right]$$

from their conditional expectations ${}^{\lfloor o}\mathcal{E}[\cdot|a_\tau,\mathcal{P}_{a_\tau^*}]$, given by ${}^{\lfloor o}f(\Delta_\tau|a_\tau,\mathcal{P}_{a_\tau^*})$, are zero mean and mutually non-correlated. With them,

$$\mathcal{H}(\mathcal{P}_{a_{t+1}^*},\Theta) = \mathcal{H}_t\left({}^{\lfloor o}f||\Theta\right) + \frac{1}{t}\sum_{\tau\le t} e_{\Theta;\tau}.$$

The assumption (2.49) implies that the variance of $e_{\Theta;\tau}$ is bounded. Consequently, the last term in the above expression converges to zero almost surely (a.s.), see [21], page 417. The first term on the right hand side of the last equality is non-negative as it can be viewed as a sum of Kullback-Leibler divergences, see Proposition 2.10. Due to (2.49), it is also finite. Thus, (2.52) converges a.s. to the non-negative value $\mathcal{H}\left({}^{\lfloor o}f||\Theta\right)$. The posterior pdf remains unchanged if we subtract $t\min_{\Theta\in\mathrm{supp}[\,f(\Theta)]\cap\Theta^*}\mathcal{H}_\infty\left({}^{\lfloor o}f\middle\|\,\Theta\right)$ from the exponent of its non-normalized version (2.51). Then, the exponent contains $(-t\times$ an_asymptotically_non-negative_factor$)$. Thus, the posterior pdf $f(\Theta|\mathcal{P}_{a_\infty^*})$ may be asymptotically non-zero on minimizing arguments (2.50) only. ♢

Remark(s) 2.12

1. *The entropy rate can be seen as an extension of the Kullback-Leibler divergence (2.26) that covers well asymptotic and controlled cases. It coincides with the Kullback-Leibler divergence in a range of particular cases.*

2. *The assumption (2.49) can be weakened. It is, however, intuitively well acceptable. It excludes parameterized models, which assign zero belief to data generated by the system with a non-zero probability and vice versa.*

3. *The Bayesian estimation minimizes asymptotically the entropy rate of the objective pdf* ${}^{\lfloor o}f(\Delta_t|a_t,\mathcal{P}_{a_t^*})$ *and a model chosen among candidates* $f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta)$, $\Theta\in\Theta^*$. *In other words, a* best projection *of the objective pdf to the considered parameterized models is asymptotically found. The prior pdf can be interpreted as a prior belief assigned to the individual parameters* $\Theta\in\Theta^*$ *that the corresponding parameterized model is the best projection of the objective pdf [22]: not knowing the reality we do not know the best projection we finally arrive at.*

4. *The posterior pdf concentrates on a point if there is a unique minimizer of the entropy rate. In this case, the model is called* identifiable. *The possibility to identify the model can be influenced by*

- *the considered class of the parameterized models and*
- *the decisions chosen, for instance, by the controller used: e.g. controller generating constant inputs does not allow us to learn their dynamic influence on outputs.*

5. *If the objective pdf* ${}^{\lfloor o}f(\Delta_t|a_t,\mathcal{P}_{a_t^*})$ *coincides with* $f(\Delta_t|a_t,\mathcal{P}_{a_t^*},\Theta)$ *for some* $\Theta = {}^{\lfloor o}\Theta$ *with* $f\left({}^{\lfloor o}\Theta\right) > 0$ *then* ${}^{\lfloor o}\Theta$ *is in the support of the asymptotic posterior pdf* $f(\Theta|\mathcal{P}_{a^*_\infty})$*. If, moreover, the model is identifiable the objective pdf is asymptotically identified by the adopted Bayesian approach. This fact can be expressed in a more appealing form:*

   Bayesian estimate is consistent whenever there is a consistent estimator.

6. *Often, a similar analysis is performed by measuring distance of parameterized models to empirical pdf of data [23]. It gives similar answers if the empirical pdf converges to the objective pdf. Moreover, it provides hints how to approximate the posterior pdf [24], see also Section 4.2.1. On the other hand, the known conditions of such convergence are more restrictive. For instance, an analysis of the controlled case is much harder.*

Problem 2.1 (How to unify statistics?) *Asymptotic analysis and finite-data oriented Bayesian approach are often perceived in an antagonistic way. Their simultaneous and harmonized use still waits for its full exploitation.*

2.5.4 Exponential family

The exponential family of distribution describing the parameterized model plays decisive role in all decision-making tasks. The reason is, that all functional recursions solving the dynamic decision-making tasks change into algebraic ones for their statistics. Thus, the tasks are significantly easier to solve within this class then outside it.

Agreement 2.14 (Exponential family of parameterized models) *The parameterized model belongs to the (dynamic) exponential family iff it can be written in the form*

$$f(\Delta_t|\Theta,\mathcal{P}_{a_t^*},a_t) = A(\Theta)\exp[B'(\Psi_t)C(\Theta)], \text{ where} \tag{2.53}$$

$\Psi_t' \equiv [\Delta_t', \psi_t']$ *is a finite dimensional* data vector *determined by experience* $\mathcal{P}_{a^*_{t+1}}$ *and its values can be recursively updated according to a known formula* $(\Psi_{t-1}, a_t, \Delta_t)^* \to \Psi_t^*$,

$'$ *means transposition.*

$A(\cdot)$ *is a non-negative scalar function defined on* Θ^*,

$B(\cdot)$, $C(\cdot)$ *are vector functions of the same finite fixed dimension. They are defined on respective arguments in* Ψ_t^* *and* Θ^*.

Remark(s) 2.13

1. *The* ψ *part of the data vector* Ψ *is often called* regression vector.

2. *Our definition of the exponential family is non-standard. The additional requirement on the recursive updating of the data vector* Ψ *(2.53) is important for the dynamic situations we deal with.*

3. *Notice that equality is used in (2.53). The normalization of this pdf must not spoil it. Due to it, the allowed form is rather restrictive. Let us indicate it.*

   Often, it is possible to assume that $f(\Delta_t|\Theta,\mathcal{P}_{a_t^*},a_t) = \Phi(\Psi_t,\Theta) > 0$. *Then, a first order Taylor-type expansion of logarithm of this function implies that* $\ln(\Phi(\Psi_t,\Theta)) \approx \tilde{B}'(\Psi_t)\tilde{C}(\Theta)$. *We exponentiate the right hand of this approximate equality and try to normalize it so that integral over* Δ_t *equals to one. Mostly, we find that the normalizing factor* $A(\cdot)$ *will depend both on entries of* ψ_t *and* Θ. *Thus, the resulting pdf lies out of the exponential family.*

Practical significance of the exponential family becomes obvious when we specialize to it Proposition 2.14 describing estimation and prediction.

Proposition 2.16 (Estimation and prediction in exponential family) *Let natural conditions of decision making, Requirement 2.5, be met with time-invariant* $\Theta_t = \Theta \in \Theta^*$. *Let the parameterized model belong to exponential family (2.53). Then, the* predictive pdf *(the outer model of the system) is given by the formula*

$$f(\Delta_t|\mathcal{P}_{a_t^*},a_t) = \frac{\mathcal{I}(V_{t-1}+B(\Psi_t),\nu_{t-1}+1)}{\mathcal{I}(V_{t-1},\nu_{t-1})}, \quad \textit{with} \tag{2.54}$$

$$V_t = V_{t-1} + B(\Psi_t),\ V_0 = 0; \quad \nu_t = \nu_{t-1} + 1,\ \nu_0 = 0, \quad \textit{and} \tag{2.55}$$

$$\mathcal{I}(V,\nu) = \int_{\Theta\in\Theta^*} A^{\nu}(\Theta)\exp[V'C(\Theta)]f(\Theta)\,d\Theta \tag{2.56}$$

where $f(\Theta)$ *is a prior pdf. The Bayesian parameter estimate (posterior pdf) is*

$$f(\Theta|\mathcal{P}_{a_{t+1}^*}) = \frac{A^{\nu_t}(\Theta)\exp[V_t'C(\Theta)]f(\Theta)}{\mathcal{I}(V_t,\nu_t)}, \tag{2.57}$$

i.e. the likelihood function is $\mathcal{L}(\Theta,\mathcal{P}_{a_{t+1}^*}) \equiv \mathcal{L}(\Theta,V_t,\nu_t) = A^{\nu_t}(\Theta)\exp[V_t'C(\Theta)]$.

Remark(s) 2.14

1. *Estimation and prediction within exponential family is extremely simple. The problem of updating of functions (pdfs) converts into the algebraic recursive updating of the finite-dimensional sufficient statistics* V_t *and of the* sample counter ν_t. *The wish to have this recursion complete explains the requirement for possibility to update* Ψ_t *recursively, see Agreement 2.14. Moreover, a single type of the normalization integral* $\mathcal{I}(V,\nu)$ *has to be evaluated.*

2. *An inspection whether there is a wider set of parameterized models with advantageous properties of the exponential family opens just narrow space [25, 26]. Essentially, the exponential family coincides with all parameterized models that are sufficiently smooth functions of* Θ *and with supports independent of* Θ. *Uniform distribution with unknown constant boundaries represents one of a few feasible examples of pdfs out of exponential family.*

3. *The class of models that lead to a finite-dimensional characterization of pdfs occurring in filtering is even more restrictive. Its discussion can be found in [27].*

Chapter 3

Decision-making tasks

Here, we provide practical comments to elements occurring in all decision tasks. Then we list usual decision tasks met. Their specific cases are is the second part of this text. Their presentation follows more or less the structure of this chapter.

3.1 Construction elements

The discussion of "atoms" creating any decision task should help us to learn a good practice and to avoid common mistakes. Their choice, modification and use represent typically specific decision subtasks that have to be harmonized with the final aim considered. It is often very hard task but the following *design golden rule* should be respected as much as possible:

Do not split the overall decision task into subtasks unless necessary. (3.1)

3.1.1 Data

Data connect the artificial world of evaluations with reality. Their information content is crucial for the success of the decision making that use them.

Experimental design

If the designer has an opportunity to influence the information content of data he/she should do that by a proper *experimental design*. At abstract level, it means that the chosen working conditions should suppress ambiguity of the best projections caused by quality of data, see Proposition 2.15. More practically, the optional data (inputs, set-points) have to be chosen so that they "excite" the system in a sufficient way. For instance, we cannot learn the dependence of outputs on inputs when inputs are not varying during the data acquisition. Obviously, data properties influence also speed of the learning.

Thus, if possible, it is reasonable to optimize information content of data. We outline such optimization in the simpler case of estimation.

Mutual information, i.e. Kullback-Leibler divergence (2.26) of the joint pdf of data and parameters to the product of their marginal pdfs, is known to be an adequate quantitative expression

of the informational relationship inspected [15]. Under natural conditions of decision making (2.36), it takes the form

$$\mathcal{I}(\Delta(\mathring{t}),a(\mathring{t})||\Theta) \equiv \tag{3.2}$$
$$\int f(\Delta(\mathring{t}),a(\mathring{t}),\Theta)\ln\left(\frac{f(\Delta(\mathring{t}),a(\mathring{t}),\Theta)}{f(\Delta(\mathring{t}),a(\mathring{t}))f(\Theta)}\right)d(\Delta(\mathring{t}),a(\mathring{t}),\Theta) =$$
$$\sum_{t\in t^*}\int f(\Delta_t|\Theta,\mathcal{P}_{a_t^*},a_t)f(a_t|\mathcal{P}_{a_t^*})f(\Theta|\mathcal{P}_{a_t^*})f(\mathcal{P}_{a_t^*})\ln\left(\frac{f(\Theta|\mathcal{P}_{a_t^*})}{f(\Theta)}\right)d(\Delta(t),a(t),\Theta).$$

This form demonstrates that the mutual information measures an average distance of the posterior and prior pdfs.

In order to learn as much as possible about unknown parameters, the mutual information should be maximized with respect to the optional admissible strategy described by its model $\{f(a_t|\mathcal{P}_{a_t^*})\}_{t\in t^*}$. Formally, dynamic programming, Proposition 2.8, could be used to this purpose.

Remark(s) 3.1

1. *The functional (3.2) depends linearly on pdfs $\{f(a_t|\mathcal{P}_{a_t^*})\}_{t\in t^*}$ describing the optional strategy. Thus, a practically meaningful solution can be obtained only when some restrictions are added. This property is not related to this specific distance but manifests the fact that an infinite stimulating energy, if allowable, may provide the best information about the system inspected.*

2. *The formulated maximization task cannot be generally solved because of its complexity. It is as complex as dual control. It provides us at least the ideal solution to be approximated. For instance, it becomes a useful optimizing guide when only a finite amount of competitive strategies is a priori allowed.*

   Mutual information can also be used for analyzing data when results of parameter estimation are unsatisfactory.

3. *Specific feasible solutions for specific classes of models are well elaborated see, for instance, [28].*

Problem 3.1 (How to approximate experimental design?) *General, numerically feasible approximations of the optimal but infeasible solution to the experimental design are still waiting for their inventors. The problem is even more hard and challenging when the predictive pdfs result from filtering, see Section 2.5.1. It might be reasonable to consider the asymptotic version of this optimization like in Section 2.4.3.*

Data pre-processing

The objective pdf describing reality is (practically) always out of the set considered parameterized models. The estimation searches for the best projection of all aspects of reality which are reflected in measured data, cf. Remarks 2.12. Some aspects may be a priori uninteresting for the problem at hands. The estimation, that always deals with a finite data set, has to handle them, too. Consequently, it might fail to provide a solid information on aspects of interest. Thus, a

wise and commonly used practice is to suppress uninteresting details in data before estimation or filtering, to pre-process data. A similar logic justifies pre-processing in control and other decision-making tasks.

Pre-processing steps could be and ideally should be done using the framework described in this text. Some of them form essence of the presented examples. At the same time, there is a bunch of ready technique [29, 30] that can be *carefully* used.

Typical pre-processing steps are: data transformation linearizing their relationships, data scaling, outliers removal, high frequency noise removal, filling of missing data, data aggregation etc. Each of them makes a pre-projection to a simplified world. It obviously influences quality of the resulting projection. Damages made in the pre-processing phase can hardly be removed in later design phases. Typical errors in pre-processing are:

- a premature reduction of data so that some informative attributes are lost,
- wasting of the information in data by sampling them with too low frequency,
- a significant change of the modelled dynamics by underestimating the fact that the the final model describes coupling of the reality and the used pre-processing mechanism,
- a distortion of the inspected relationships by inadequate substitutions for missing data.

Problem 3.2 (How to harmonize pre-processing with ultimate goal?) *Similarly, as in other places, the optimal pre-processing demands to solve the decision problem in its entirety. It is mostly impossible and splitting the overall task into adequate and harmonized subtasks is left to a "sound" reasoning. It is pleasant as it requires creativity. It is unpleasant as the overall results of the decision making might be spoiled by an improper choice. The problem is severe especially in dynamic decision-making problems in which there is a restricted freedom for an iterative trial-and-error treatment.*

3.1.2 Decisions

Elements

The set a^* of possible decisions contains often elements that can "substitute" each other. This incomplete pre-determination calls for selecting the most suitable elements. For instance, moisture of a produced paper can be influenced both by machine speed and energy used for drying [31]. The choice is first of all based on physical and economic priorities. Then structure estimation task, see Agreement 3.4 and Proposition 3.6, should follow when no other technical priorities can be recognized.

Restrictions

Causality restrictions, see Section 2.2, are always considered in our design. Their impact can be influenced at a pre-design stage. Typically, new sensors can be installed in order to enrich accessible experience. Their efficient choice is far from being trivial and respecting of the design golden rule (3.1) cannot be over-stressed in this instance.

Technological restrictions, i.e. final specification of the decision set $a^*(\mathring{t})$, follows mostly from technological, economic or safety considerations. Generally, they increase complexity of the

design of the optimal strategy and have to be modified in order to make the design and application feasible. When they are relaxed for complexity reasons, they have to be reflected in other elements of the design. Often, the loss function is modified. In a sense, a sort of penalty-based minimization under restrictions is implicitly used.

Moreover, a good and justified optimization practice has to be respected. For instance, when a particular unrestricted optimizing decision is out of the target set a^* a simple clipping at the boundary of a^* should not be used: a proper (near) optimal projection on a^* is needed.

Also, unnecessary restrictions should be avoided. For instance, the restriction to unbiased estimators [32] is well justified in many statistical decision tasks. If it is, however, mechanically required for other tasks, say adaptive control design, it can make the final strategy (much) less efficient than possible.

Problem 3.3 (How to treat technical priorities?) *A systematic analysis of priorities and technical consideration mentioned in previous paragraphs and met on many places should be supported by formal tools. Those available now, like* qualitative modelling and simulation *[33], do not fit well the methodology described here.*

3.1.3 Parameterized model

Again, a non-formalized search for a compromise between the model quality and complexity of its treatment makes the model choice more art than science.

Exponential family

The need for reaching this compromise is responsible for a prominent role of so called *exponential family* [34] and *conjugate prior*. They help us in struggle with the curse of dimensionality [35] and at the same time they are able to cover a range of practical problems.

Majority of the textbook pdfs belong to the exponential family. The following table provides examples of such pdfs in the *static case* when the regression vector is empty, i.e. $\Psi \equiv \Delta \equiv y$ and y is either scalar or its entries are independent (recall Agreement 2.5).

Name	**Parameterized model**	**Innovation**	**Parameter**
Exponential	$\frac{1}{\lambda}\exp\left(-\frac{y}{\lambda}\right)$	$y \geq 0$	$\lambda > 0$
Poisson	$\frac{\mu^y}{\Gamma(y+1)}\exp(-\mu)$	$y \in \{0, 1, \dots\}$	$\mu > 0$
Multinomial	$\prod_{i \in y^*} \Theta_i^{\delta(i,y)}$	$y \in \{1, \dots, \mathring{y}\}$	$\{\Theta_i \geq 0, \sum_{i \in y^*} \Theta_i = 1\}$
Normal	$\frac{1}{\sqrt{2\pi r}}\exp\left[-\frac{(y-\mu)^2}{2r}\right]$	$y \in (-\infty, \infty)$	$\mu \in (-\infty, \infty)$, $r > 0$
Log-normal	$\frac{1}{y\sqrt{2\pi r}}\exp\left[-\frac{\ln^2\left(\frac{y}{\mu}\right)}{2r}\right]$	$y > 0$	$\mu > 0$, $r > 0$

The *Euler gamma function* Γ (7.3) and *Kronecker delta function* δ (7.1) are defined in Chapter 7.

The *dynamic case* with a nonempty ψ is much more narrow (cf. Remarks above). Essentially,

normal (Gaussian) linear-in-parameters model and Markov chains dominate this case.

Name	Parameterized model	Data vector	Parameter
Normal (Gaussian)	$\frac{1}{\sqrt{2\pi r}} \exp\left[-\frac{(y-\theta'\psi)^2}{2r}\right]$	$y \in (-\infty, \infty)$ $\psi \in (-\infty, \infty)^{\mathring{\psi}}$	$\theta \in \theta^* \subset (-\infty, \infty)^{\mathring{\theta}}$ $r > 0$
Markov chain	$\prod_{i \in y^*} \prod_{j \in \psi^*} \Theta_{i\|j}^{\delta(i,y)\delta(j,\psi)}$	$y \in y^* \equiv \{1, \ldots, \mathring{y}\}$ $\psi \in \psi^* \equiv \{1, \ldots, \mathring{\psi}\}$	$\forall j \in \psi^*$, $\{\Theta_{i\|j} \geq 0$, $\sum_{i \in y^*} \Theta_i = 1\}$

Completion of probabilistic models

The presented design relies (with a few exceptions) on a complete specification of the parameterized model. At the same time, the designer has always at disposal a finite amount of facts for its construction. Their necessary *completion to pdf* (an infinite-dimensional object) represents a non-standard and weakly supported interpolation task. We sketch its solution proposed in [36].

Let the designer know the form of the conditional expectation of a vector function $H(\Delta_t)$, i.e.

$$\mathcal{E}[H(\Delta_t, \mathcal{P}_{a_t^*}, a_t)|\Theta, \mathcal{P}_{a_t^*}, a_t] = h(\Theta, \mathcal{P}_{a_t^*}, a_t) \tag{3.3}$$

with a known vector function h. The constructed pdf $f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t)$ should respect just this knowledge. It has to fulfill (3.3). It should be as uncertain as possible in order to leave as much freedom as possible for learning from data. We quantify these requirements by minimizing Kullback-Leibler divergence (2.26) of the constructed $f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t)$ to the uniform pdf under the restriction (3.3). Note that the minimization is equivalent to entropy maximization [15].

Proposition 3.1 (Completion of parameterized models) *The minimizer* $f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t)$ *of the functional*

$$\mathcal{D}(f||\textit{uniform pdf}) = \int f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) \ln[f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t)]\, d\Delta_t$$

within the set of pdfs $f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) \in$

$$\left\{ \int_{\Delta_t \in \Delta_t^*} f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t)\, d\Delta_t = 1,\ \mathcal{E}[H(\Delta_t, \mathcal{P}_{a_t^*}, a_t)|\Theta, \mathcal{P}_{a_t^*}, a_t] = h(\Theta, \mathcal{P}_{a_t^*}, a_t) \right\}$$

has the form

$$f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) = \gamma(\Theta, \mathcal{P}_{a_t^*}, a_t) \exp[\lambda'(\Theta, \mathcal{P}_{a_t^*}, a_t) H(\Delta_t, \mathcal{P}_{a_t^*}, a_t)]. \tag{3.4}$$

The scalar function $\gamma(\cdot)$ *is determined by the normalization, see Proposition 2.4. The vector function* $\lambda(\cdot)$ *has to fulfill the identity*

$$\int_{\Delta_t^*} H(\Delta_t, \mathcal{P}_{a_t^*}, a_t) \gamma(\Theta, \mathcal{P}_{a_t^*}, a_t) \exp[\lambda'(\Theta, \mathcal{P}_{a_t^*}, a_t) H(\Delta_t)]\, d\Delta_t = h(\Theta, \mathcal{P}_{a_t^*}, a_t). \tag{3.5}$$

Proof: It is implied by an elementary calculus of variations. ♢

Remark(s) 3.2

1. *The result is more of a conceptual than of a practical interest as the solution to (3.5) can rarely be found. Moreover, the model found is mostly out of the favorite exponential family.*

2. *Gaussian parameterized model is obtained for* $H(\Delta_t, \mathcal{P}_{a_t^*}, a_t) = [\Delta_t, \Delta_t^2]'$. *This helps us in judging suitability of this popular option.*

Problem 3.4 (How to make the completion practicable?) *The weakness of the presented results indicates how neglected the art of the model completion is. The completion has to be either developed much more or avoided completely. The latter desirable direction would, however, require to built the corresponding decision theory relying only on partial descriptions like (3.3). It is possible that it is impossible.*

Mixtures of models

The model constructed in previous Section exploits available physical knowledge expressed by the expectation (3.3). It is often obtained through modelling valid locally around some working point. Then, an extension to models covering a wider range of working conditions is desirable. *Mixture of models*

$$f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) = \sum_{c\in c^*} \alpha_c f_c(\Delta_t|\Theta_c, \mathcal{P}_{a_t^*}, a_t), \quad \Theta \in \Theta^* \equiv \tag{3.6}$$

$$\left\{\Theta = (\alpha_c, \Theta_c),\ \Theta_c \in \Theta_c^*\ c \in c^* \equiv \{1, \ldots, \mathring{c} < \infty\},\ \alpha \in \alpha^* \equiv \left\{\alpha_c \geq 0,\ \sum_{c\in c^*} \alpha_c = 1\right\}\right\}$$

represents a good universal option for meeting this need as the mixtures of models have often so called *universal approximation property*. It means that the set (3.6) is dense in the set of outer models of the system. This property takes place whenever the number of *components* $\mathring{c}$ is not limited and individual components, i.e. pdfs $f_c(\Delta_t|\Theta_c, \mathcal{P}_{a_t^*}, a_t)$, can be made arbitrarily close to an arbitrarily positioned Dirac delta function [37].

The *exact estimation and prediction* with this model is practically *impossible* as the number of terms in likelihood function $\mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}^*})$ – the product of parameterized models – increases exponentially. Luckily enough, good algorithms for approximate estimation of the mixture of models are available [38]. Even a good approximate *recursive* estimation is possible if the *components* $\{f_c(\Delta_t|\Theta_c, \mathcal{P}_{a_t^*}, a_t)\}_{c\in c^*}$ belong to the exponential family or describe uniform distributions on rectangular boxes. This quasi-Bayes estimation is described in Section 4.2.3.

The quasi-Bayes estimation as well as other algorithms exploit the following alternative interpretation of the mixture of models (3.6). Let $c_t \in c^*$ be the unobservable random selector of the "active" component that takes the value c with probability α_c. Then,

$$f(\Delta_t, c_t|\Theta, \mathcal{P}_{a_t^*}, a_t) = f_{c_t}(\Delta_t|\Theta_{c_t}, \mathcal{P}_{a_t^*}, a_t)\alpha_{c_t}. \tag{3.7}$$

The marginalization rule, see Proposition 2.4, implies that the mixture (3.6) is marginal pdf of the joint pdf (3.7).

3.1.4 Prior pdf

The need for selecting prior pdf is often regarded as the main disadvantage of the adopted Bayesian approach [10]. We contribute positively to the never-ending discussion on its pros and

cons by showing that its use opens a regular gate through which prior "expert" information can be introduced into learning in a systematic way.

The posterior pdf (2.44) is a product of the likelihood function consisting of t factors $f(\Delta_\tau|\Theta,\mathcal{P}_{a^*_\tau},a_\tau)$ and of a single prior pdf $f(\Theta)$. If t is high enough and data bring a sufficient information on Θ then the posterior pdfs obtained for various prior pdfs resemble each other: the role of prior pdf is weak.

The posterior pdf is significantly influenced by the prior pdf when some of the above conditions is not fulfilled. Let us formulate simple Propositions that formalize this.

Proposition 3.2 (Role of the prior pdf)

1. *Parameter values* $\Theta \notin \mathrm{supp}\,[\,f(\Theta)]$, *for which the prior pdf is zero, get the zero posterior pdf, too. Formally,*
$$\mathrm{supp}\left[\,f(\Theta|\mathcal{P}_{a^*_{t+1}})\right] = \mathrm{supp}\left[\,\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}})\right] \cap \mathrm{supp}\,[\,f(\Theta)]\,.$$
2. *The recursive evolution of the likelihood function*
$$\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}}) = f(\Delta_t|\Theta,\mathcal{P}_{a^*_t},a_t)\mathcal{L}(\Theta,\mathcal{P}_{a^*_t}),\quad t\in t^*,\ \mathcal{L}(\Theta,\mathcal{P}_{a^*_1}) = 1, \Theta\in\Theta^*_{\mathcal{L}} \tag{3.8}$$
does not depend on the prior pdf chosen.
3. *The posterior pdf exists iff the product* $\mathcal{L}(\Theta,\mathcal{P}_{a^*_{t+1}})f(\Theta)$ *is integrable.*

Proof: It is a direct consequence of the formula for posterior pdf (2.44). ♢

Remark(s) 3.3

1. *The prior pdf offers a simple and clear way for introducing hard restrictions on parameters.*
2. *The recursion (3.8) is valid even if* $\Theta^*_{\mathcal{L}} \neq \Theta^* \equiv \mathrm{supp}\,[\,f(\Theta)]$. *This trivial statement implies that hard bounds on parameter values must not influence likelihood function, i.e. the "objective" link to data reflecting reality. This fact is easily and repeatedly overlooked in recursive estimation. Instead of restricting the posterior pdf the statistics determining the likelihood are deformed with an adverse effect on the estimation quality.*
3. *Often, a flat prior pdf is chosen in order to model the lack of prior knowledge. Even integrability of the prior pdf is relaxed and the* improper prior pdfs $f(\cdot)\geq 0$, $\int f(\Theta)\,d\Theta = \infty$ *are used. For instance, the posterior pdf is proportional to the likelihood function if we allow the prior pdf be improper and equal 1. Then, however, the posterior pdf might be improper, too. Thus, it is more reasonable to select a flat but proper prior pdf as it serves as a regularizing factor.*

Agreement 3.1 (Conjugate prior pdf) *Let the parameterized model belong to exponential family, see Agreement 2.14,*
$$f(\Delta_t|\Theta,\mathcal{P}_{a^*_t},a_t) = A(\Theta)\exp[B'(\Psi_t)C(\Theta)].$$
The prior pdf is called conjugate prior pdf *if it has the form*
$$f(\Theta)\propto A^{\nu_0}(\Theta)\exp[V_0'C(\Theta)]\chi(\Theta). \tag{3.9}$$
It is determined by the indicator χ *of the set* Θ^* *and "prior statistics"* V_0, ν_0.

Proposition 3.3 (Estimation with conjugate prior pdf) *Let the parameterized system belong to the exponential family and the conjugate prior pdf is chosen. Then, the posterior pdf have the same functional form as the prior one. Its statistics V_t, ν_t evolve according to (2.55) with zero initial conditions replaced by V_0, ν_0:*

$$\begin{aligned} f(\Theta|\mathcal{P}_{a_t^*}) &\propto A^{\nu_t}(\Theta)\exp[V_t'C(\Theta)]\chi(\Theta),\ t\in t^* \\ V_t &= V_{t-1}+B(\Psi_t),\ \nu_t=\nu_{t-1}+1 \\ V_0,\,\nu_0 &\equiv \textit{prior statistics.} \end{aligned} \tag{3.10}$$

Proof: It is a direct consequence of the formula for posterior pdf (2.44) applied to the exponential family. ♢

Agreement 3.2 (Fictitious data) *Let the parameterized model belong to the exponential family and a conjugate prior pdf $f(\Theta)\equiv f(\Theta|K_{t_f})$ respects a piece of knowledge K_{t_f}. The* fictitious data vector Ψ_{t_f} *is said to express this knowledge piece if*

$$f(\Theta|K_{t_f})\propto A^{\bar\nu+1}(\Theta)\exp\left\{[B(\Psi_{t_f})+\bar V]'C(\Theta)\right\}\propto A(\Theta)\exp[B'(\Psi_{t_f})C(\Theta)]\bar f(\Theta). \tag{3.11}$$

The used pdf $\bar f(\Theta)\propto A^{\bar\nu}(\Theta)\exp[\bar V'C(\Theta)]$ is a flat pre-prior pdf.

In other words, the inclusion of the individual knowledge piece K_{t_f} is expressed as an application of the Bayes rule with fictitious data vector inserted into the parameterized model.

Remark(s) 3.4

1. *Instructive examples of fictitious data are in in [39].The construction is also demonstrated on examples of this text. The common procedure is as follows:*

   Algorithm 3.1 (Construction of fictitious data)

   (a) *Ask yourselves what response of the system you expect to some experience (often, past inputs or initial values) when your information piece is correct.*

   (b) *Think about uncertainty of the system response caused by your uncertainty about the treated knowledge piece.*

   (c) *Scale these data so that the response uncertainty resembles the uncertainty caused by random noises within the system.*

2. *It is possible and sometimes reasonable to group fictitious data into mutually consistent blocks that are treated as a block of real data vectors, i.e. modify the flat prior pdf through a plain rule of Bayes rule.*

A collection of knowledge pieces $\{K_{t_f}\}_{t_f\in t_f^*}$ contains generally items of heterogeneous and uncertain nature. The auxiliary notion of fictitious data provides us their unified expression and give us a possibility to reflect their uncertainty. With fictitious data vectors, the knowledge items $K(\mathring{t}_f)$, $\mathring{t}_f<\infty$ can be incorporated into the prior pdf individually. Without knowing their mutual dependencies, like their complementarity, degree of compatibility, repetitions etc., we cannot construct the desirable prior pdf $f(\Theta|K(\mathring{t}_f))$ reflecting the whole collection of available knowledge pieces. Thus, some heuristics is needed. The following simple proposition prepares one of possible and successful formalizations.

Proposition 3.4 (Geometric representant of a pair of pdfs) *Let an unknown pdf $f \in f^* \equiv \{f_1, f_2\}$ be equal to $f_1(x)$ with a probability $\alpha \in \alpha^* \equiv [0,1]$ and equal to $f_2(x)$ with the complementary probability $1-\alpha$. The pdfs f_1, f_2 are supposed to have a common support x^*. Then, the pdf*

$$\overset{\text{o}\rfloor}{f}(x) \propto f_1^{\alpha}(x) f_2^{1-\alpha}(x), \;\; x \in x^* \tag{3.12}$$

minimizes the expected Kullback-Leibler divergence (2.26) of $\hat{f}(\cdot)$ to $f(\cdot)$, i.e. the functional

$$\mathcal{E}[\mathcal{D}(\hat{f}||f)] \equiv \alpha \mathcal{D}(\hat{f}||f_1) + (1-\alpha)\mathcal{D}(\hat{f}||f_2). \tag{3.13}$$

The minimum reached is

$$\omega(\alpha) \equiv \mathcal{E}[\mathcal{D}(\overset{\text{o}\rfloor}{f}||f)] = -\ln \int_{x^*} f_1^{\alpha}(x) f_2^{1-\alpha}(x)\, dx. \tag{3.14}$$

If $f_1 \not\equiv f_2$, the function $\omega(\alpha)$ reaches its maximum on $(0,1)$.

Proof: (3.12) is implied by an elementary calculus of variations. (3.14) follows from it and (3.13). The last statement is implied by the assumption $f_1 \not\equiv f_2$, closeness of α^* and continuity of $\omega(\alpha)$ guaranteed by the common support of f_1, f_2. ♢

The combination or merging of a pair of knowledge pieces may be directly based on this proposition by taking $f_{t_f} \equiv f(\Theta|K_{t_f})$, $t_f = 1,2$ and accepting $\overset{\text{o}\rfloor}{f}$ as $f(\Theta|K(\mathring{t}_f))$. The constructed prior pdf used should not claim misleading confidence thus it makes sense to choose the "worst" (conservative) α in (3.12), i.e. α that maximizes the minimum reached.

It is straightforward to extend Proposition 3.4 for several pooled pdfs. However, the search for the worst probabilistic vector α becomes computationally intensive even for a medium $\mathring{t}_f$.

There is an obvious approximation based on a sequential application of Proposition 3.4 with $f_1 \equiv f(\Theta|K(\mathring{t}_f - 1))$, $f_2 \equiv f(\Theta|K_{t_f})$ and $\overset{\text{o}\rfloor}{f} \equiv f(\Theta|K(\mathring{t}_f))$. The result may, however, strongly depend on the evaluation order. For this reason, an alternative formulation has been proposed in [39] that allows us not only to add a knowledge piece but also to remove superfluous ones.

Let us describe this procedure assuming that $f(\Theta|K(t_f - 1))$ has been gained in some way and that there are two possibilities to built-in a new piece of knowledge K_{t_f}:

Optimistic variant: Bayes rule includes correctly the new knowledge piece

$$f_1(\Theta|K(t_f)) \propto f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f}) f(\Theta|K(t_f - 1)).$$

This variant is assigned the probability $\alpha \in [0,1]$

Cautious variant: Bayes rule can only be applied after removing the new factor $f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f})$ β-times

$$f_2(\Theta|K(t_f)) \propto [f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f})]^{1-\beta} f(\Theta|K(t_f - 1)), \;\; \beta \geq 0.$$

This variant is taken with the probability $1-\alpha$.

Proposition 3.5 (Recursive processing of prior knowledge) *Let $f(\Theta|K(t_f))$ has only the optimistic or cautious variant. Then, the minimizer $\hat{f}(\Theta|K(t_f))$ of the expected value of the Kullback-Leibler divergence (2.26) to it has the form*

$$\hat{f}(\Theta|K(t_f)) \propto [f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f})]^{w} f(\Theta|K(t_f - 1)), \;\; \text{with } w = \alpha + (1-\alpha)(1-\beta).$$

The unique "worst" values of $\alpha \in [0,1]$ *maximizing the reached minimum and*

$$\beta = \left\{ \inf \tilde{\beta} : \int [f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f})]^{-\tilde{\beta}} f(\Theta|K(t_f-1))\, d\Theta = \infty \right\}$$

exist.

Proof: Essentially, Proposition 3.4 is extended. For details see [39]. ♢

In batch processing, when we can combine prior knowledge with the measured data at one shot, it is possible to estimate weights α by using the standard Bayes rule. This straightforward procedure provides directly estimate of the posterior pdf with the available prior knowledge included. We present the result for exponential family.

Agreement 3.3 (Batch processing of prior knowledge) *Let the parameterized model belong to exponential family (2.53) and prior knowledge pieces* $K(\mathring{t}_f)$ *are expressed through fictitious data, see Agreement 3.2. We search for such an approximation* $\hat{f}(\Theta|\mathcal{P}_{a^*_t}, K(\mathring{t}_f))$ *of the unknown posterior pdf* $f(\Theta|\mathcal{P}_{a^*_t}, K(\mathring{t}_f))$ *that*

- *has experience* $\mathcal{P}_{a^*_t}, K(\mathring{t}_f)$,
- *minimizes the expected Kullback-Leibler divergence to projections of* $f(\Theta|\mathcal{P}_{a^*_t}, K(\mathring{t}_f))$ *within the set*

$$\{f(\Theta|\mathcal{P}_{a^*_t}, K_{t_f}\}_{t_f \in t^*_f} \tag{3.15}$$

- *assigns the to the elements of (3.15) the posterior probabilities* $\alpha(K_{t_f}|\mathcal{P}_{a^*_t})$ *corresponding to uniform prior probabilities on* $K(\mathring{t}_f)$.

A straightforward use of Proposition 3.4, definition of fictitious data, Agreement 3.2, estimation in the exponential family, Proposition 3.3, and Bayes rule give:

Algorithm 3.2 (Batch processing of prior knowledge)

1. *Compute the statistics* V, ν *determining the posterior pdf corresponding to the pre-prior pdf only*

$$\bar{V}_t = \sum_{\tau \le t} B(\Psi_\tau) + \bar{V}, \quad \bar{\nu}_t = \bar{\nu} + t \tag{3.16}$$

   where the data vectors Ψ_τ *are constructed from the (real) data available.*

2. *Evaluate the posterior probabilities*

$$\alpha(t_f|\mathcal{P}_{a^*_t}) \propto \frac{\mathcal{I}(\bar{V}_t + B(\Psi_{t_f}), \bar{\nu}_t + 1)}{\mathcal{I}(\bar{V} + B(\Psi_{t_f}), \bar{\nu} + 1)}, \; t_f \in t^*_f \tag{3.17}$$

   where the data vectors Ψ_{t_f} *are fictitious.*

3. *Determine statistics of the final pdf* $\hat{f}(\Theta|\mathcal{P}_{a^*_t}, K(\mathring{t}_f)) \propto A^{\nu_t}(\Theta) \exp[V_t' C(\Theta)]$

$$V_t = \bar{V}_t + \sum_{t_f \in t^*_f} \alpha(t_f|\mathcal{P}_{a^*_t}) B(\Psi_{t_f}), \quad \nu_t = \bar{\nu}_t. \tag{3.18}$$

Remark(s) 3.5

1. *The upper bound on the set of possible β values is chosen so that out of it the removal of the factor $f(\Delta_{t_f}|\Theta, \mathcal{P}_{a^*_{t_f}}, a_{t_f})]^{-\beta}$ shifts the product with $f(\Theta|K(t_f-1))$ out of class of proper pdfs: nothing more can be removed.*

2. *The combination of pdfs with the worst α is tightly related to so called* barycenter *of a pdf group, see [40] and its approximation called pooling [41].*

3. *The results of the batch version do not depend on the order of prior knowledge processing. It is obviously desirable property that supports the use of this version whenever possible.*

Problem 3.5 (Is the knowledge quantification good enough?) *There is a significant progress in this area but the results are not stabilized yet. For instance, there is little experience with a direct use of Proposition 3.5 and Algorithm 3.2. It is worth of being inspected more closely.*

Problem 3.6 (How to choose version of Kullback-Leibler divergence?) *Kullback-Leibler divergence is not symmetric in its arguments. Optimization of the alternative version leads to the arithmetic mean of pdfs instead of the geometric one. There are indications [42] that the chosen version should be preferred but a definite answer does not exist.*

3.1.5 Loss function

The *loss function should express our wishes as precisely as possible*. This obvious demand is often violated. Then, the design provides non-acceptable decision strategies. Let us discuss the common mistakes made.

Common mistakes

A wider than the practically admissible set of decisions a^* is often used in order to decrease complexity of the design. Then, the loss function has to penalize values out of the desirable set. Otherwise, the optimal strategy found may generate useless decisions.

Use of loss functions that do not respect decision making as a dynamic process is another common error. Generally, the decision a_t influences whole future of the coupling system $\leftrightarrow$ strategy. It gives unreasonable strategies if this fact is not respected and sequence of decisions is chosen by optimizing a sequence of unrelated decisions, see Section 4.4.1.

Instability of the closed loop with one-step-ahead controllers applied to non-minimum phase systems is the most known demonstration of this property. However, even stable behavior is achieved with energy wasted. This is a less known manifestation of the same "mistake" [9].

Popular options

A dominating use of a few loss functions reflects an indirect fight with the design complexity. Undoubtedly, the *quadratic loss* is a prominent example of this type. Its general form,

$$\mathcal{Z} = ||\mathcal{Q} - {}^{\lfloor I}\mathcal{Q}||^2, \tag{3.19}$$

where $||\cdot||$ is a quadratic *weighted* norm and $\lfloor^I$ denotes a given ideal (reference, set-point), has numerous variants. They differ in the target ideal, in the variables involved and weighting used. The quadratic loss:

+ is numerically feasible in connection with linear models,

+ is flexible as it allows us to express priorities by a proper choice of weights,

+ approximates majority of smooth loss functions through a second order Taylor expansion.

It is, however, fair to say that:

− the quadratic loss over-stresses significance of large deviations from the ideal,

− the choice of weights and ideal reflecting true aims is not easy,

− usefulness of the quadratic loss as a "suppressor" of the design complexity decreases sharply whenever the underlying model is not linear or hard restrictions are involved.

When hard bounds on admissible decisions have to be considered, loss functions based on l-one norm might be much better. Also, the design based on l-infinity norm is nowadays widely inspected [43]. It enhances the attention paid to requirements on *robustness* of the designed strategy. Note that robustness is an engineering name to a smooth dependence of the strategy on various elements (initial conditions, model quality, etc.) on which it depends.

Ideal in fully probabilistic design

The choice of the loss function should be harmonized with the model adopted. For instance, if large random deviations are expected then the use of a quadratic loss might be doubtful.

The need for the harmonization is an additional argument for our preference to the fully probabilistic design, see Section 2.4.2. In this case, the loss is determined by the ideal pdf $\lfloor^I f(\Delta(\mathring{t}), a(\mathring{t}))$ we would like to approximate. Generally, it is wise to derive the ideal pdf from the outer description of closed loop $f(\Delta(\mathring{t}), a(\mathring{t}))$ by replacing those characteristics we hope to influence by their desirable values. In this way, we get, for instance, an equivalent of a well tailored quadratic performance index for linear systems with Gaussian distribution [18]. Then, the choice of adequate weights in the loss is much simplified. We get also a realistic loss if the well modelled system noise is non-normal.

Problem 3.7 (How to express multiple objectives?) *There are extensive theories how to handle multi-valued loss functions, what properties utility functions should have [44] etc. In spite of this, operational guidelines how to construct the loss function in specific cases are far from being cook-book.*

3.2 Derived elements

3.2.1 Likelihood function

All investigated tasks are based on data measured on the plant we want to decide of. The information extracted from the data measured up to time t (without prior data) is collected in

the likelihood function. It is defined as a product of model pdfs with the data substituted which is viewed as a function of the unknown parameter Θ (2.45)

$$\mathcal{L}(\Theta, \mathcal{P}_{a^*_{t+1}}) \equiv \prod_{\tau \leq t} f(\Delta_\tau | \Theta, \mathcal{P}_{a^*_\tau}, a_\tau) \tag{3.20}$$

Remark(s) 3.6

1. *The form of the likelihood function is often used as a template for the form chosen for the prior pdf. The reason is that reproducibility of the likelihood function when recursively updated, see (3.8).*

3.2.2 The integral $\mathcal{I}$

This integral is defined in (2.46)

$$\mathcal{I}(\mathcal{P}_{a^*_{t+1}}) = \int \mathcal{L}(\Theta, \mathcal{P}_{a^*_{t+1}}) f(\Theta)\, d\Theta. \tag{3.21}$$

It is frequently used in the tasks solved.

3.2.3 Posterior pdf

The posterior pdf represents the most general description the unknown parameter Θ. It yields not only point estimates of parameter values in the form of conditional mean, but it contains also information about its uncertainty. It can be obtained from likelihood and prior pdf

$$f(\Theta | \mathcal{P}_{a^*_{t+1}}) \propto \mathcal{L}(\Theta, \mathcal{P}_{a^*_{t+1}}) f(\Theta) \tag{3.22}$$

with normalization integral $\mathcal{I}(\mathcal{P}_{a^*_{t+1}})$ (3.21) or recursively

$$f(\Theta | \mathcal{P}_{a^*_{t+1}}) \propto f(\Delta_t | \Theta, \mathcal{P}_{a^*_t}, a_t) f(\Theta | \mathcal{P}_{a^*_t}), \tag{3.23}$$

where in the role of the normalization integral appears the predictive pdf (3.24).

3.2.4 Predictive pdf

It is a distribution of the next data item (in time instant $t+1$) conditioned on past and present data items but not parameters. It can be computed from model an posterior pdf by integration

$$f(\Delta_{t+1} | \mathcal{P}_{a^*_{t+1}}, a_{t+1}) = \int_{\Theta^*} f(\Delta_{t+1} | \Theta, \mathcal{P}_{a^*_{t+1}}, a_{t+1}) f(\Theta | \mathcal{P}_{a^*_t}) d\,\Theta = \frac{\mathcal{I}(\mathcal{P}_{a^*_{t+1}})}{\mathcal{I}(\mathcal{P}_{a^*_t})} \tag{3.24}$$

where the integrals $\mathcal{I}(\mathcal{P}_{a^*_{t+1}})$ and $\mathcal{I}(\mathcal{P}_{a^*_t})$ can be found above in (3.21).

Remark(s) 3.7

1. *The pdf (3.24), at time t is the normalization integral for recursive update of posterior pdf (3.23).*

3.3 Static tasks

The *static decision making* deals with the design and use of a single decision rule. Proposition 2.7 describes its optimal design. Specific instances differ in:

- Decomposition of the behavior $\mathcal{Q} = (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*}) \equiv$ (experience, decision, ignorance).
- Admissible rules $\mathcal{R} \in \mathcal{R}^*$ determined by the domain $\mathcal{P}^*_{a^*}$ and range a^*.
- Loss function $\mathcal{Z} : \mathcal{Q}^* \equiv (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*})^* \to [0, \infty]$.

These elements determine the outer model $f(\mathcal{F}_{a^*}|\mathcal{P}_{a^*}, a)$ needed for a direct application of the basic optimization lemma, Proposition 2.7.

3.3.1 Point estimation

Point estimation can be cast in the considered framework as follows.

- $\mathcal{Q} \equiv (D, \hat{\Theta}, \Theta) \equiv$ (data at disposal, point estimate, unknown parameter).
- Admissible rules are of the form $\mathcal{R} : D^* \to \hat{\Theta}^*$, $\Theta^* \subset \hat{\Theta}^*$
- Loss function $\mathcal{Z}$ measures a distance of $\hat{\Theta}$ and Θ. This distance may depend on data D, too.

The ignorance $\mathcal{F}_{\hat{\Theta}^*}$ coincides with the unknown parameter Θ. Thus, the outer model of the system needed for decision making is $f(\Theta|D, \hat{\Theta})$. The adopted natural conditions of decision making (2.36) imply that the decision $\hat{\Theta}$ is superfluous in the conditioning. Thus, the needed model coincides with the (generalized) Bayesian estimate $f(\Theta|D)$ (posterior pdf) determined in Proposition 2.14. For the given D, the optimal point estimate is, Proposition 2.7,

$$\hat{\Theta}(D) \in \operatorname{Arg}\min_{\hat{\Theta}\in\hat{\Theta}^*} \int \mathcal{Z}(D, \hat{\Theta}, \Theta) f(\Theta|D)\, d\Theta. \tag{3.25}$$

Remark(s) 3.8

1. *Note that the needed outer model depends on the specific structuring of realizations $\mathcal{Q}$ not on the loss function chosen.*

   The specific result of the optimization depends of course on the loss function chosen.

2. *The quadratic loss function $\mathcal{Z}(D, \hat{\Theta}, \Theta) = ||\Theta - \hat{\Theta}||^2_{Q(D)} \equiv ||\Theta||^2_{Q(D)} + ||\hat{\Theta}||^2_{Q(D)} + 2(\Theta, \hat{\Theta})_{Q(D)}$ is the most popular option. It is determined by a quadratic norm $||\cdot||_{Q(D)}$ with a positive definite, data-dependent, kernel $Q(D)$ and the related scalar product $(\cdot, \cdot)_{Q(D)}$. For it, the minimized quantity, see Proposition 2.7, is quadratic form in $\hat{\Theta}$, namely,*

$$\mathcal{E}[\mathcal{Z}|D, \hat{\Theta}] = \mathcal{E}[||\Theta||^2_{Q(D)}|D] + ||\hat{\Theta}||^2_{Q(D)} + 2(\mathcal{E}[\Theta|D], \hat{\Theta})_{Q(D)} \tag{3.26}$$

$$\Rightarrow \qquad \hat{\Theta} \equiv \mathcal{E}[\Theta|D],$$

irrespectively of the kernel $Q(\cdot)$.

This general result looks very promising as just the first moment (expectation) of the posterior pdf is needed. However, its evaluation requires mostly a complete knowledge of $f(\Theta|D)$.

Problem 3.8 (How to select the ideal pdf for estimation?) *There is little experience how to specify the ideal of the fully probabilistic design 2.4.2 in the case of parameter estimation. The general guideline, see Section 3.1.5 should be followed in this case, too. A systematic inspection is needed.*

3.3.2 Set estimation

Set estimation is tightly related to the point estimation. It can be cast in the considered framework as follows.

- $\mathcal{Q} \equiv (D, \theta^*, \Theta)$
  (data, set estimate, unknown parameter).
- Admissible rules are of the form

$$\mathcal{R} : D^* \to (\theta^*)^* \subset \left\{ \theta^* \subset \Theta^*, \ \int_{\theta^*} f(\Theta, D)\, d\Theta \geq \alpha \in (0,1) \right\} \tag{3.27}$$

  where α is a pre-specified probability called *credibility level*.
- Loss function $\mathcal{Z}$ measures volume of the θ^*. The volume specification may depend on data D, too.

Structuring of realizations $\mathcal{Q}$ is similar to the point estimation. This, see Remark 1 in the previous section, implies that the outer system model needed is $f(\Theta|D)$. For the given D, the optimal interval estimate is (Proposition 2.7)

$$\theta^*(D) \in \mathrm{Arg} \min_{(\theta^*) \in (\theta^*)^*} \int \mathcal{Z}(D, \theta^*, \Theta) f(\Theta|D)\, d\Theta. \tag{3.28}$$

Remark(s) 3.9

1. *Essentially,* α*-fractiles of* $f(\Theta|D)$ *are searched for.*
2. *Mostly, intervals are taken as the sets considered. Their length is then taken as a measure of their volume. Other choices might be more appropriate when a multivariate* Θ *is considered. For instance, hyper-ellipsoids fit well to normally distributed parameters.*
3. *The sets searched for are called* credibility sets *and they are quite close to so called* confidence intervals *studied in classical statistics.*

Problem 3.9 (Is there a cook-book giving the ideal pdf?) *Again, the experience how to specify the ideal of the fully probabilistic design 2.4.2 is missing. The general guideline, see Section 3.1.5 should be followed but a systematic inspection is needed.*

3.3.3 Testing of hypothesis

Hypothesis testing, a selection of the best variant among several alternatives, can be put within our framework as follows.

- $\mathcal{Q} \equiv (D, \hat{h}, \{H_h\}_{h\in h^*}) \equiv$
  (data at disposal, estimate $\hat{h} \in h^* \equiv \{1, \ldots, \mathring{h} < \infty\}$ of h, {hypothesis list}).
- Admissible rules are of the form $\mathcal{R} : D^* \to h^*$.
- Loss function $\mathcal{Z}$ is a $(\mathring{h}, \mathring{h})$-table with non-negative entries $\mathcal{Z}(D, \hat{h}, h)$, usually with the zero diagonal as no penalty is paid when $\hat{h} = h$.

The outer system model (2.20) needed is $f(h|D, \hat{h})$. The adopted natural conditions of decision making (2.36) imply that the decision $\hat{h}$ is superfluous in conditioning. Thus, we need $f(h|D)$. This probability function (pf, $\mathring{h} < \infty$) is obtained according to Proposition 2.14 as the generalized Bayesian estimate

$$f(h|D) \propto f(D|h)f(h), \ h \in h^* \tag{3.29}$$

where $f(h)$ is the prior pf of h's hypothesis. The pdf $f(D|h)$ relates observed data to individual hypotheses $h \in h^*$. For a given D, the optimal decision in hypothesis testing is, Proposition 2.7,

$$\hat{h}(D) \in \mathrm{Arg}\min_{\hat{h}\in\hat{h}^*} \sum_{h\in h^*} \mathcal{Z}(D, \hat{h}, h) f(h|D)\,. \tag{3.30}$$

Remark(s) 3.10

1. *Unlike in classical hypothesis testing [32], the testing is performed within a completely specified set of alternatives.*
2. *If a pair of hypothesis is compared, $\mathring{h} = 2$, the decision rule constructed according to the described methodology coincides with a celebrated Neymann-Pearson lemma [1].*

   In this case, off diagonal elements of $\mathcal{Z}(D, \hat{h}, h)$ penalize the classical errors of the 1st and 2nd kind [32].
3. *The needed pdfs are rarely obtained directly. Instead, they are predictive pdfs obtained through filtering or parameter estimation and prediction, see Remarks 2.10. If other decisions, like control, influence the observed data then factors describing their generators cancel in (3.29). It follows from natural conditions of acting (2.36) that are supposed to be valid.*
4. *In spite of its formal simplicity, the testing of hypothesis is extremely powerful technique. It is especially true when dealing with predictive pdfs needed when so called compound hypothesis are treated, [32]. It, for instance, brought a whole set of novel and efficient solutions to so called* structure estimation problem *[45, 46, ?] (see below).*
5. *When dealing with predictive pdfs, a special care has to be devoted to the choice of the prior pdf of unknown parameters within the respective hypothesis. Their practical influence may be unexpectedly high.*

The importance of the structure estimation task calls for its formal specification:

Agreement 3.4 (Structure estimation) *Let us consider parameterized models* $\{f(\Delta_t|\Theta_h, \mathcal{P}_{a_t^*}, a_t, h)\}_{h\in h^*; t\in t^*}$ *as candidates for describing a single system. Then, the hypothesis testing about the best value* $h \in h^*$ *is called* structure estimation. *The generator of the additional decisions* a_t *(like system inputs) is supposed to fulfill natural conditions of decision making (2.36).*

Proposition 3.6 (Solution to structure estimation) *Let the data* $D \equiv (\mathcal{P}_{a_{t+1}^*}, a_{t+1})$ *be available for structure estimation. Then, the outer model needed for selecting the best structure estimate according to (3.30) is*

$$\begin{aligned} f(h|\mathcal{P}_{a_{t+1}^*}) &\propto f(\Delta_t|\mathcal{P}_{a_t^*}, a_t, h) f(h|\mathcal{P}_{a_t^*}) \equiv \\ &\equiv \int f(\Delta_t|\Theta_h, \mathcal{P}_{a_t^*}, a_t, h) f(\Theta_h|\mathcal{P}_{a_t^*})\, d\Theta_h f(h|\mathcal{P}_{a_t^*}). \end{aligned} \tag{3.31}$$

The pf $f(h|\mathcal{P}_{a_t^*}, a_t)$ *can be obtained recursively according to the formula (3.31) starting from a given prior pf* $f(h) \equiv f(h|\mathcal{P}_{a_1^*})$. *The needed generalized Bayesian parameter estimate evolves according to Bayes rule,* $h \in h^*$,

$$f(\Theta_h|\mathcal{P}_{a_t^*}) \propto f(\Delta_t|\Theta_h, \mathcal{P}_{a_t^*}, a_t, h) f(\Theta_h|\mathcal{P}_{a_{t-1}^*}) \tag{3.32}$$

starting from a given prior pdf of parameters $f(\Theta_h|h) \equiv f(\Theta_h|h, \mathcal{P}_{a_1^*})$.

The equivalent batch version is, $h \in h^*$,

$$\begin{aligned} f(h|\mathcal{P}_{a_{t+1}^*}) &\propto \frac{\mathcal{I}_h(\mathcal{P}_{a_{t+1}^*})}{\mathcal{I}_h(\mathcal{P}_{a_1^*})} f(h|\mathcal{P}_{a_1^*}) \\ \mathcal{I}_h(\mathcal{P}_{a^*}) &\equiv \int \mathcal{L}(\Theta_h, \mathcal{P}_{a^*}) f(\Theta_h|h)\, d\Theta_h \\ \mathcal{L}(\Theta_h, \mathcal{P}_{a_{t+1}^*}) &= \prod_{\tau=1}^{t} f(\Delta_\tau|\Theta_h, \mathcal{P}_{a_\tau^*}, a_\tau, h),\ \mathcal{L}(\Theta_h, \mathcal{P}_{a_1^*}) \equiv 1. \end{aligned} \tag{3.33}$$

Proof: It uses basic algebra with pdfs, Proposition 2.4, respecting natural conditions of acting adopted for all involved decisions. ♢

Problem 3.10 (How to treat select the list of hypotheses?) *Mechanical ways of generating list of hypotheses make* $\mathring{h}$ *extremely large and consequently the their testing infeasible. GUHA methodology represents a systematic way how to proceed. Nothing, however, has been elaborated which would fit the methodology presented here.*

Problem 3.11 (How to treat extending list of hypotheses?) *Hypotheses are usually created gradually. It opens a question, how to extend the existing set of hypothesis and how to exploit former data so that the new hypothesis is compared in a fair way. A lot of partial steps have been done in this respect but a systematic design and analysis are missing.*

3.3.4 One-step-ahead prediction

One-step-ahead prediction, an extrapolation of experience to the potentially observable data that are part of the ignorance set, is a very basic task that fits in our framework as follows.

- $\mathcal{Q} \equiv (\mathcal{P}_{\hat{\Delta}^*}, \hat{\Delta}, \Delta) \equiv$
  (data at disposal, prediction, potentially observable data).
- Admissible rules are $\mathcal{R} : \mathcal{P}^*_{\hat{\Delta}^*} \to \hat{\Delta}^*$.
- Loss function $\mathcal{Z}$ measures a distance of $\hat{\Delta}$ and Δ and may depend on $\mathcal{P}_{\hat{\Delta}^*}$.

The outer system model (2.20) needed is $f(\Delta|\mathcal{P}_{\hat{\Delta}^*})$ as the natural conditions of decision making (2.36) imply that the decision $\hat{\Delta}$ is superfluous in conditioning. This pdf is obtained according to Proposition 2.14 as the generalized Bayesian prediction. For a given $\mathcal{P}_{\hat{\Delta}^*}$, the optimal point prediction is, Proposition 2.7,

$$\hat{\Delta}(\mathcal{P}_{\hat{\Delta}^*}) \in \mathrm{Arg}\min_{\hat{\Delta}\in\hat{\Delta}^*} \int \mathcal{Z}(\mathcal{P}_{\hat{\Delta}^*}, \hat{\Delta}, \Delta) f(\Delta|\mathcal{P}_{\hat{\Delta}^*})\, d\Delta. \tag{3.34}$$

Remark(s) 3.11

1. *The prediction experience may contain results of other decisions, like system inputs. In this case* $\mathcal{P}_{\hat{\Delta}^*} \equiv (\mathcal{P}_{u^*}, u) \equiv D$, *inputs are integral part of experience and they influence both value of the innovation and its prediction.*
2. *For the popular quadratic loss* $\mathcal{Z}(D, \hat{\Delta}, \Delta) = ||\Delta - \hat{\Delta}||^2_{Q(D)}$ *with quadratic norm* $||\cdot||_{Q(D)}$ *weighted by a positive definite kernel* $Q(\cdot)$, *see Remarks 3.8 and the relation (3.26), the conditional expectation is the optimal prediction for any weight* $Q(\cdot)$,
$$\hat{\Delta}(D) = \mathcal{E}[\Delta|D]. \tag{3.35}$$
3. *The discussed prediction should be called* point prediction. *As in the estimation, set estimates (credibility intervals) might be searched for. Their discussion is identical with that presented in Section 3.3.2.*

3.3.5 One-step-ahead control

One-step-ahead control tries to influence the response of the controlled system in the nearest future by choosing an appropriate input. This problem is formulated within our framework as follows.

- $\mathcal{Q} \equiv [\mathcal{P}_{u^*}, u, (y, x)] \equiv$
  [data at disposal, system input, (system output, system state)].
- Admissible rules are of the form $\mathcal{R} : \mathcal{P}^*_{u^*} \to u^*$.
- Loss function $\mathcal{Z}$ measures distance of (y, x, u) to desirable known reference values (y_r, x_r, u_r) and may depend on $\mathcal{P}_{u^*}$.

The outer system model (2.20) needed is $f(y, x|\mathcal{P}_{u^*}, u) = f(x|\mathcal{P}_{u^*}, u, y) f(y|\mathcal{P}_{u^*}, u)$. The factors creating it are obtained by applying Proposition 2.13 if we

- set $x \equiv \Theta_t$,

- take the filtration result after data updating as the first factor,
- use the output predictor $f(y|\mathcal{P}_{u^*}, u)$ as the second factor.

For the given $\mathcal{P}_{u^*}$, the optimal inputs are (Proposition 2.7)

$$u(\mathcal{P}_{u^*}) \in \operatorname{Arg}\min_{u\in u^*} \int\int \mathcal{Z}(\mathcal{P}_{u^*}, y, x, u, y_r, x_r, u_r) f(x|\mathcal{P}_{u^*}, u, y) f(y|\mathcal{P}_{u^*}, u)\, dxdy. \tag{3.36}$$

Remark(s) 3.12

1. *The input u influences directly the outer model of the system. This property makes the main difference of the control task from estimation.*
2. *This case is a special version of the multi-step control design, see Section 3.4.4, where additional details are presented.*
3. *Often, just the output of the system is controlled. Then the output predictor is needed only.*
4. *Output and state reference values are often called* set points.
5. *Quadratic distance is a popular option used. Then, the conditional expected loss optimized depends on 1st and 2nd conditional moments of the outer model. Often, the conditional variance does not depend on the chosen input. Then, a quadratic form in the conditional expectation $\mathcal{E}\{[y, x]|\mathcal{P}_{u^*}, u\}$ is just needed for determining the optimal input.*
6. *Often, multi-step-ahead control design is approximated by using an extreme version of* receding-horizon strategy*, see Section 4.4.1. The substantial loss of quality up to the closed-loop instability may be encountered in this case [9]. This extreme should be avoided as much as possible.*
7. *Fully probabilistic version, see Proposition 2.11, of this problem may be easily formulated and solved. It is the simples application of Proposition 2.11 and as such is omitted here.*

3.4 Dynamic tasks

Dynamic decision making discussed here form the core of this work. This justified a bit more detailed discussion.

3.4.1 Sequential estimation

Mostly, we consider decision tasks with a fixed horizon $\mathring{t} < \infty$. Sequential estimation is an important exception.

If technically possible, it is often advantageous to decide whether to perform a new measurement before making the final decision. This task makes sense if we respect the price paid in connection with the data acquisition. The optimal decision strategies, called *sequential*, then lead in average to a shorter observation process. It makes the overall decision making cheaper in comparison with strategies that rely on a pre-defined sufficiently long decision horizon. Sequential estimation

serves well for describing the idea. Other variants, like sequential testing of hypothesis might be formulated and solved similarly.

Sequential point estimation can be cast in our framework as follows.

- $\mathcal{Q} \equiv [\mathcal{P}_{a_t^*}, a_t \equiv (\hat{\Theta}_t, s_t), \mathcal{F}_{a_t^*} \equiv (\Theta, \Delta_t)]$
  [data at disposal, (estimate,stopping flag), (unknown parameter,innovations)].
- Admissible strategies consist of rules $\mathcal{R}_t : \mathcal{P}^*_{a_t^*} \to (\hat{\Theta}_t^*, s_t^*)$, $\Theta^* \subset \hat{\Theta}^*$, $s_t^* \equiv$ {stop measuring and estimate Θ, wait for a new measurement} $\equiv \{0, 1\}$,
- Loss function

$$\mathcal{Z} = \begin{cases} \sum_{\tau \le t} c(\mathcal{P}_{a_\tau^*}) + z(\mathcal{P}_{a_t^*}, \Theta, \hat{\Theta}_t) & \text{if} \quad s_t = 0 \text{ and } s_\tau = 1, \ \forall \tau < t \\ \sum_{\tau \le t} c(\mathcal{P}_{a_\tau^*}) & \text{if} \quad s_\tau = 1 \ \forall \tau \le t \end{cases} \tag{3.37}$$

  where $z(\mathcal{P}_{a_t^*}, \Theta, \hat{\Theta}_t)$ measures a distance of Θ and its estimate $\hat{\Theta}$. $c(\mathcal{P}_{a_\tau^*})$ denotes a positive price of τth observation.

Proposition 3.7 (Sequential estimation) *Let us consider the sequential estimation and assume that there is an admissible strategy for which the expected loss is finite. The following inequalities express the sufficient condition for an index t to be the time moment at which observation should be stopped*

$$\mathcal{E}\left[\left(z(\mathcal{P}_{a_t^*}, \Theta, \hat{\Theta}_t) - z(\mathcal{P}_{a^*_{t+k}}, \Theta, \hat{\Theta}_{t+k}) - \sum_{\tau > t}^{t+k} c(\mathcal{P}_{a_\tau^*})\right) | \mathcal{P}_{a_t^*}\right] \le 0, \ \forall k = 1, 2, \ldots. \tag{3.38}$$

In (3.38), $\hat{\Theta}_{t+k}$, $k = 0, 1, 2, \ldots$ denote parameter estimates based on $\mathcal{P}_{a^*_{t+k}}$ minimizing $\mathcal{E}[z(\mathcal{P}_{a^*_{t+k}}, \Theta, \hat{\Theta})]$.

Proof: Let (3.38) be fulfilled. Then, combining the form of the loss function (3.37), the fact that the optimal stopping time has to be determined using its experience and finiteness of the loss for the optimal solution we get, $\forall k = 1, 2, \ldots,$

$$\mathcal{E}\left[\left(z(\mathcal{P}_{a_t^*}, \Theta, \hat{\Theta}_t) + \sum_{\tau=1}^{t} c(\mathcal{P}_{a_\tau^*})\right) | \mathcal{P}_{a_t^*}\right] \le \mathcal{E}\left[\left(z(\mathcal{P}_{a^*_{t+k}}, \Theta, \hat{\Theta}_{t+k}) + \sum_{\tau=1}^{t+k} c(\mathcal{P}_{a_\tau^*})\right) | \mathcal{P}_{a_t^*}\right].$$

Using isotonicity of the expectation (taken over $\mathcal{P}_{a_t^*}$), see Proposition 2.6, we find that the chosen decision cannot be improved by any estimate that uses more measurements than the inspected one. ♢

The outer system model (2.20) needed for this decision making coincides with the generalized Bayesian estimate (posterior pdf) given in Proposition 2.14.

Remark(s) 3.13

1. *The ability to evaluate $\mathcal{E}\left[z(\mathcal{P}_{a^*_{t+k}}, \Theta, \hat{\Theta}_{t+k}) | \mathcal{P}_{a_t^*}\right]$ is decisive for a practical solvability of the problem.*
2. *Stopping rules used for speeding up extensive simulations [47] based on a simple sequential estimation serve as an example of their, still underestimated, usefulness.*
3. *The dependence of the observation price can be effectively exploited when the sequential estimation is performed in an inner loop of some optimization process: the closer we are to the optimum the lower this weight is.*

3.4.2 Multi-step-ahead prediction

Multi-step-ahead prediction, an extrapolation of experience to data in a more distant part of the ignorance set, is an extension of one-step-ahead prediction, Section 3.3.4, and fits in our framework as follows.

- $\mathcal{Q} \equiv (\mathcal{P}_{\hat{\Delta}^*_{t+j|t}}, \hat{\Delta}_{t+j|t}, \Delta_{t+j}) \equiv$
  (data at disposal at time t, prediction of innovations at time $t+j$ based on data at disposal, innovation potentially observable at time $t+j$, $j>1$).
- Admissible strategies are formed by rules $\mathcal{R}_t : \mathcal{P}^*_{\hat{\Delta}^*_{t+j|t}} \to \hat{\Delta}^*_{t+j|t}$.
- Loss function $\mathcal{Z}$ measures $\mathcal{P}_{\hat{\Delta}^*_{t+j|t}}$-dependent distance of $\hat{\Delta}_{t+j|t}$ and Δ_{t+j}.

The outer system model (2.20) needed is $f(\Delta_{t+j}|\mathcal{P}_{\hat{\Delta}^*_{t+j|t}})$ and the optimal point prediction is (Proposition 2.7)

$$\hat{\Delta}_{t+j|t} \in \mathrm{Arg}\min_{\hat{\Delta}\in\hat{\Delta}^*} \int \mathcal{Z}(\mathcal{P}_{\hat{\Delta}^*_{t+j|t}}, \hat{\Delta}, \Delta_{t+j|j}) f(\Delta_{t+j}|\mathcal{P}_{\hat{\Delta}^*_{t+j|t}})\, d\Delta_{t+j}.$$

Let us illustrate the construction of the needed outer model in a special case of the controlled system with $\Delta_t = y_t =$ system output. It is controlled by a randomized controller described by its outer model $f(u_t|\mathcal{P}_{u_t^*})$, Agreement 2.6. The outer model of the system $f(y_t|\mathcal{P}_{u_t^*}, u_t)$ is either given or obtained as the predictive pdf by filtering, Proposition 2.13 or through parameter estimation, Proposition 2.14. The experience of the multi-step predictor is assumed to coincide with $\mathcal{P}_{u_t^*}, u_t$.

Using the calculus with pdfs, Proposition 2.4, the needed outer model can be expressed with the help of the available elements as follows

$$\begin{aligned} f(y_{t+j}|u_t, \mathcal{P}_{u_t^*}) = \\ = \int f(y(t\_t+j), u(t+1\_t+j)|\mathcal{P}_{u_t^*}, u_t)\, d\left(y(t\_t+j-1), u(t+1\_t+j)\right) = \\ = \int \prod_{\tau=t}^{t+j} f(y_\tau|\mathcal{P}_{u_\tau^*}, u_\tau,) f(u_\tau|\mathcal{P}_{u_\tau^*}) d\left(y(t\_t+j-1), u(t+1\_t+j)\right). \end{aligned}$$

Remark(s) 3.14

1. *The need to know the future control strategy makes the main difference of this task from one-step-ahead prediction where just knowledge of inputs is sufficient.*
2. *As a rule, the computed marginal pdf $f(y_{t+j}|u_t, \mathcal{P}_{u_t^*})$, $j>1$ is much flatter than the one-step ahead predictor. Formally, it results from integration (averaging) over intermediate predicted values. It reflects the fact that it is much harder (less reliable) to make a long term prediction. The situation is the more pronounced the longer the prediction horizon j is.*
3. *Notice that integrations over the intermediate quantities is done also over their values in conditioning. This makes multi-step prediction highly non-linear task.*

4. *Sometimes, the parameterized model can be constructed directly with the gap j, i.e. in the form $f(\Delta_{t+j}|\Theta,\mathcal{P}_{\hat{\Delta}^*_{t+j|t}})$. Then, there is no formal difference to the one-step-ahead prediction. The prediction quality is of course worse. If, however, there is a gap in system-output sequence and no gap in system-input sequence the prediction of future inputs is necessary.*

   Predictors of this type are used in connection with so called MUSMAR controllers [48].

5. *The point prediction is discussed here. Of course, the set estimation may be also formulated and solved using our general tools.*

3.4.3 Filtering

Filtering, described in Proposition 2.13, provides filtration and one step-ahead prediction of an unknown internal variable Θ_t and can be extended to multi-step prediction in the same manner as the innovation prediction discussed in Section 3.4.2. The fact that we never observe directly time varying Θ_t calls for a novel task called *smoothing*.

A smoother estimates the internal quantities $\Theta_{t-j},\ j>1$ using also the measured data reflecting newer variables $\Theta((t-j+1)\_t)$. Its construction is cast in our framework as follows.

- $\mathcal{Q}\equiv(\mathcal{P}_{\hat{\Theta}^*_{t-j|t}},\hat{\Theta}_{t-j|t},\Theta_{t-j})\equiv$
  (data at disposal at time t, smoothed estimate of Θ_{t-j} based on data at disposal, unknown internal variable at time $t-j$, $j\geq 1$).
- Admissible rules are of the form $\mathcal{R}:\mathcal{P}^*_{\hat{\Theta}^*_{t-j|t}}\to\hat{\Theta}^*_{t-j|t}$.
- Loss $\mathcal{Z}$ measures distance of $\hat{\Theta}_{t-j|t}$ and Θ_{t-j} and may depend on $\mathcal{P}_{\hat{\Theta}^*_{t-j|t}}$.

Assuming controlled system and $\mathcal{P}_{\hat{\Theta}^*_{t-j|t}}\equiv(u_t,\mathcal{P}_{u^*_t})$, basic decision-making lemma, Proposition 2.7, reads

$$\hat{\Theta}_{t-j|t}\in\mathrm{Arg}\min_{\hat{\Theta}\in\hat{\Theta}^*_{t-j}}\int\mathcal{Z}(u_t,\mathcal{P}_{u^*_t},\hat{\Theta}_{t-j},\Theta_{t-j})f(\Theta_{t-j}|\mathcal{P}_{u^*_t},u_t)\,d\Theta_{t-j} \tag{3.39}$$

Let us assume that we have at disposal both observation $f(y_t|\Theta_t,\mathcal{P}_{u^*_t},u_t)$ and time evolution $f(\Theta_t|\Theta_{t-1},u_t,\mathcal{P}_{u^*_t})$, $t\in t^*$ models used in Proposition 2.13. The applied control strategy meets natural conditions of decision making (2.36). Using the calculus with pdfs, Proposition 2.4, and (2.36), the model needed in the above optimization can be expressed with the help of the available elements as follows

$$\begin{aligned}f(\Theta_{t-j}|\mathcal{P}_{u^*_t},u_t)\propto f(\Theta_{t-j},y(t-j\_t)|u(t-j\_t),\mathcal{P}_{u^*_{t-j}})=\\ =\int\prod_{\tau=t-j+1}^{t-1}f(y_\tau|\Theta_\tau,\mathcal{P}_{u^*_\tau},u_\tau)f(\Theta_\tau|\Theta_{\tau-1},\mathcal{P}_{u^*_\tau},u_\tau)\,d\Theta(t-j+1\_t-1)\times\\ \times f(y_{t-j}|\Theta_{t-j},\mathcal{P}_{u^*_{t-j}},u_{t-j})f(\Theta_{t-j}|\mathcal{P}_{u^*_{t-j}})\end{aligned} \tag{3.40}$$

The integrand contains the available models with available data substituted. The factor $f(\Theta_{t-j}|\mathcal{P}_{u^*_{t-j}})$ results from filtering described by Proposition 2.13. Smoothing is simply derived from filtering.

3.4.4 Multi-step-ahead control

Multi-step-ahead control, and extension of one-stage-ahead control of Section 3.3.5, tries to push quantities characterizing the controlled system, namely (system output, system input, system state)= (y, u, x), to the desired reference values (y_r, u_r, x_r). Their proximity should be enforced on a whole time-interval up to the horizon $\mathring{t}$ by choosing an appropriate sequence of inputs $u(\mathring{t})$.

We consider here the most general combination of tracking and regulation to uncertain reference values. The specialization to the case of (partially) known reference values is straightforward. This problem is formulated in our framework as follows.

- $\mathcal{Q} \equiv [\mathcal{P}_{u_t^*}, u_t, \mathcal{F}_{u_t^*}] \equiv$
  [(observed data $\equiv y(t-1), u(t-1), y_r(t-1), u_r(t), x_r(t-1)$), system input u_t, (future system outputs $y(t\_\mathring{t})$, states $x(\mathring{t})$ and references $y_r(t\_\mathring{t})$, $u_r(t+1\_\mathring{t})$, $x_r(t\_\mathring{t})$)].

  Time t is increasing within the set t^*.

- Admissible strategy consists of a sequence of *control laws* (decision rules) of the form $\{\mathcal{R}_t : \mathcal{P}^*_{u_t^*} \to u_t^*\}_{t \in t^*}$.

- Loss $\mathcal{Z}$ measures distance of $\mathcal{Q}$ to reference values $y_r(\mathring{t}), u_r(\mathring{t}), x_r(\mathring{t})$.

The optimal control strategy can be designed using directly dynamic programming, see the general version in Proposition 2.8, the version with additive loss in Proposition 2.9 or the fully probabilistic version in Proposition 2.11.

The needed outer model is $f(y_t, x_t, y_{r;t}, x_{r;t} | \mathcal{P}_{u_t^*}, u_t, u_{r;t})$. It is obtained via filtering with unknown internal variables $\Theta_t = (x_t, y_{r;t}, x_{r;t})$ using Proposition 2.13.

Remark(s) 3.15

1. *The optimal controller consists of the combination of filtering and controller acting on its results.*

2. *Control of the system with unknown time invariant parameters Θ, the central topic of main stream adaptive control [3], is a special case of the described design with the trivial time evolution $\Theta_t = \Theta_{t-1}$. The parameter estimation, see Proposition 2.13, provides the necessary outer system model in this case.*

3. *It is fair to say that the optimal multi-step-ahead control has the widest gap between the optimal design and practically optimal design. It is rarely analytically or numerically feasible. For this reason, a lot of heuristic approximation techniques have been developed. Some of them are discussed in Chapter* **??**.

4. *The time evolution model $f(x_t, y_{r;t}, x_{r;t} | x_{t-1}, y_{r;t-1}, x_{r:t-1}, \mathcal{P}_{u_t^*}, u_t, u_{r;t})$ has to describe the evolution of both system state and uncertain reference values.*

5. *The typical control design optimizing closed-loop behavior with respect to randomly occurring step changes can be obtained by modelling the output and/or state references as random walks with time-varying dispersion [49].*

6. *The described general cases covers the situation when we try to follow evolution of another uncertain object. Rescue/military interpretations of this case are straightforward. The key message is that the dynamics of the uncertain target has to be modelled for finding the optimal strategy.*

   The situation is simplified whenever target values are (partially) known. Then the corresponding outer models reduce formally to Dirac delta functions on a given known support.

Chapter 4

Evaluation techniques

Here, we review basic evaluation techniques and tools that create *bridge between optimal and practically optimal design*. All of them try to cope with the complexity restrictions. The complexity of the optimal design stems from the fact that the optimal "plan" of dynamic decisions is essentially searched for by comparing all possible behaviors resulting from the interaction of the system and judged strategy. The design complexity is significantly influenced by richness of the inspected space. Its reduction is behind the majority of available approximation schemes including those discussed here.

4.1 Adaptive systems

The ideal solution to the decision making under uncertainty is described by the combination of Bayesian filtering, see Proposition 2.13, and dynamic programming, see Propositions 2.9, 2.11. The functional equations describing them are mostly computationally infeasible. Consequently, their use cannot live without approximations. Even the simpler fully probabilistic design operates on infinite-dimensional objects. The involved multivariate functions should be represented in computer, i.e. in the device that can operate on a high but finite amount of values. Thus, a sort of approximation is needed. Global approximation of functions of many variables we are dealing with is known to be computationally hard. The application of the strategies resulting from the design requires knowledge of the discussed solutions only for the recorded experience. Thus, it is sufficient to know them *locally around the actual experience*. It can often be approximately achieved. Such *local approximations* are known as *adaptive systems* [50]. They natural exploit majority approximation techniques described below.

Remark(s) 4.1

1. *Note that there is no formal definition of adaptive systems. Their operational description is, for instance, in [3].*

2. *We found the coined understanding of adaptive systems as local approximations very useful. It helps us to have a unified view on existing practical strategies and opens a way for designing novel ones [50]. Moreover, it shows that the adaptive systems will be inevitably used in future due to the theoretically provable need for local approximations.*

4.2 Suboptimal estimation

Parameter estimation, Proposition 2.14, is a special case of filtration, Proposition 2.13, with a "trivial" time-evolution model $f(\Theta_{t+1}|\Theta_t, \mathcal{P}_{a^*_{t+1}}, a_{t+1}) = \delta(\Theta_{t+1} - \Theta_t)$. Consequently, majority of techniques suitable for approximate filtering, see Section 4.3, can be used in this case, too. Specific common techniques related to estimation only are outlined here.

4.2.1 Equivalence approach

Under natural conditions of decision making (2.36), the generalized Bayesian estimation (Proposition 2.14) updates the posterior pdfs according to the Bayes rule (2.43)

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) = \frac{f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t) f(\Theta|\mathcal{P}_{a^*_t})}{f(\Delta_t|\mathcal{P}_{a^*_t}, a_t)}, \ t \in t^*. \tag{4.1}$$

Complexity of expressing of this pdf increases quickly with an increasing amount of data, with increasing t. Exponential family (2.53) is essentially the only exception from this rule. This section tries to cope with *recursive estimation out of exponential family.*

Always limited capabilities of a computer that treats non-exponential parameterized models call for a reduced representation of propagated posterior pdfs. It is a peculiar task as the posterior pdfs concentrate quickly on a very narrow support at unknown position in Θ^*, see Proposition 2.15. Thus, a representation on a sufficiently fine grid that does not miss the final position becomes soon computationally prohibitive. The way out has been elaborated in a sequence of papers by Dr. Kulhavý and summarized in [24]. Here, we just outline the essence of this *equivalence approach.*

Proposition 4.1 (Equivalence-preserving mapping) *Let g_t be a finite-dimensional vector representing the posterior pdf $f(\Theta|\mathcal{P}_{a^*_t})$. In other words, g_t is a finite-dimensional image generated by a mapping $G_t : (f(\Theta|\mathcal{P}_{a^*_t}))^* \to g^*_t$. Let the parameterized models be positive on a common, time, data and parameter invariant support.*

Then, the values of g_t can be exactly recursively updated *using only its previous value and the current parameterized model $f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t)$ iff G_t is a time-invariant linear mapping $G \equiv G_t \, t \in t^*$ acting on logarithms of the pdfs involved. The logarithmic pdfs are treated as functions of Θ. G_t has to map Θ-independent elements to zero.*

Proof: To demonstrate necessity is rather hard and the interested reader is referred to [51, 52]. To show that the conditions on $G_t \equiv G$, $t \in t^*$, are sufficient is simple and instructive. They become obvious if we apply G to the logarithmic version of the Bayes rule (4.1) and use both its time-invariance and linearity. The normalizing term $\ln(f(\Delta_t|\mathcal{P}_{a^*_t}, a_t))$ is independent of Θ and as such mapped to zero. The recursion for g_t is then simply

$$g_t = G[\ln(f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t))] + g_{t-1}, \text{ with } g_0 = G(f(\Theta)) \equiv G(\text{prior pdf}). \tag{4.2}$$

♢

Note that (4.2) becomes the true recursion if we need not store whole past observed data for evaluating $f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t)$. Thus, similarly as in the case of exponential family, see Agreement 2.14, we adopt the mild assumption that

$$f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t) \equiv M(\Theta, \Psi_t). \tag{4.3}$$

The data $\Delta_t, a_t, \mathcal{P}_{a_t^*}$ are compressed into a finite dimensional data vector Ψ_t that can be updated recursively $\Psi_{t-1}, a_t, \Delta_t \to \Psi_t$.

In the subsequent discussion, we need:

Agreement 4.1 (Empirical pdf; Riezs representation) *Let parameterized model have the form (4.3). Then,*

$$f_t(\Psi) \equiv \frac{1}{t}\sum_{\tau=1}^{t} \delta(\Psi - \Psi_\tau),\ \ \Psi \in \Psi^* \equiv \bigcup_{t\in t^*} \Psi_t^* \tag{4.4}$$

is the (formal) empirical pdf of Ψ. The used Dirac delta function $\delta(\cdot)$ is defined in (7.2).

We assume that the same representation exists for the linear mapping G introduced in Proposition 6.1, i.e. there is a possibly generalized vector function $G(\Theta)$ such that

$$G(C) \equiv \int_{\Theta^*} C(\Theta)G(\Theta)\, d\Theta \tag{4.5}$$

for any considered function $C(\Theta),\ \Theta \in \Theta^*$.

Assuming (4.3), using the (formal) empirical pdf (4.4) and Riezs representation of G (4.5), we see that

$$g_t = t\int_{\Psi^*} \left[\int_{\Theta^*} \ln[M(\Theta,\Psi)]G(\Theta)\, d\Theta\right] f_t(\Psi)\, d\Psi \tag{4.6}$$

for the vector function $G(\Theta)$ representing functionals chosen. The integral in brackets [] defines the vector function

$$h(\Psi) \equiv \int_{\Theta^*} \ln[M(\Theta,\Psi)]S(\Theta)\, d\Theta. \tag{4.7}$$

With it, we get the equivalent form of (4.6)

$$g_t = t\int_{\Psi^*} h(\Psi)f_t(\Psi)\, d\Psi + g_0. \tag{4.8}$$

Recall, that left hand side of (4.8) is known as it can be update recursively according to (4.2) which has the equivalent "Riezs" form

$$g_t \underbrace{=}_{(4.2)} \int_{\Theta^*} \ln[M(\Theta,\Psi_t)]G(\Theta)\, d\Theta + g_{t-1} \underbrace{\equiv}_{(4.7)} h(\Psi_t) + g_{t-1},\ \ g_0 = G(f(\Theta)). \tag{4.9}$$

Also, for the chosen parameterized model (4.3) and functions $G(\Theta)$ representing the admissible projections to g_t^*, the vector function $h(\Psi)$ is known as well as its values in measured data vectors Ψ_t. The *empirical pdf* $f_t(\Psi)$ (4.4) is the only incompletely known object in (4.8) as we are not able (or willing) to store all measured $\Psi_\tau,\ \tau \le t$.

The posterior pdf, we are interested in, can be expressed in terms of empirical pdf as follows, cf. a similar transformation in connection with Proposition 2.15,

$$f(\Theta|\mathcal{P}_{a_{t+1}^*}) \equiv f(\Theta|f_t) \propto f(\Theta)\exp\left[t\int_{\Psi^*} \ln[M(\Theta,\Psi)]f_t(\Psi)\, d\Psi\right]. \tag{4.10}$$

The empirical pdf is unknown to us, thus this posterior pdf is unknown, too. The selection of a suitable approximating pdf $\hat{f}(\Theta|g_t)$ acting on Θ^* fits to our general decision-making framework:

- $\mathcal{Q} \equiv (\mathcal{P}_{a^*}, a, \mathcal{F}_{a^*}) \equiv (g_t, \hat{f}(\Theta|g_t), f(\Theta|f_t(\cdot)) \equiv$
  (stored statistic, estimate of the posterior pdf, the posterior pdf determined by the unknown empirical pdf $f_t(\Psi)$, $\Psi \in \Psi^*$),
- Admissible decision rule are of the form

$$\mathcal{R} : \; g_t^* \rightarrow \left\{ \hat{f}(\Theta|g_t) \geq 0, \int \hat{f}(\Theta|g_t)\, d\Theta = 1, \; \Theta \in \Theta^* \right\}$$

- Loss function is the Kulback-Leibler divergence (2.26) $\mathcal{D}(\hat{f}(\Theta|g_t)||f(\Theta|f_t))$

Formally, the optimal point estimate of the posterior pdf is, Proposition 2.7,

$$\begin{aligned} \hat{f}(\Theta|g_t) \quad &\in \quad \mathrm{Arg}\min_{\hat{f}\in\hat{f}^*} \mathcal{E}\left[\mathcal{D}(\hat{f}(\Theta|g_t)||f(\Theta||f_t))|g_t\right] = \\ &= \quad \mathrm{Arg}\min_{\hat{f}\in\hat{f}^*} \left[\mathcal{D}(\hat{f}(\Theta|g_t)||f(\Theta)) - t\int_{\Psi^*} \hat{f}(\Theta|g_t)\ln[M(\Theta,\Psi)]\mathcal{E}[f_t(\Psi)|g_t]\, d\Psi\right]. \end{aligned} \tag{4.11}$$

The conditional expectation $\mathcal{E}[\cdot|g_t]$ is taken with respect to the uncertain empirical pdf $f_t(\Psi)$, $\Psi \in \Psi^*$.

The minimizer can be found explicitly as

$$\hat{f}(\Theta|g_t) \propto f(\Theta)\exp\left\{t\int_{\Psi^*} \ln[M(\Theta,\Psi)]\mathcal{E}[f_t(\Psi)|g_t]\, d\Psi\right\}. \tag{4.12}$$

It depends only on the conditional expectation $\mathcal{E}[f_t(\Psi)|g_t]$. The empirical pdf fulfills the identity (4.8) that is linear in it. Consequently, its conditional expectation $\mathcal{E}[f_t(\Psi)|g_t]$ has to, see Proposition 2.6, fulfill it, too:

$$g_t = \int_{\Psi^*} h(\Psi)\mathcal{E}[f_t(\Psi)|g_t]\, d\Psi + g_0. \tag{4.13}$$

For the given g_t, g_0, the solution (4.12) can be used if we specify the expectation

$$E(\Psi) \equiv \mathcal{E}[f_t(\Psi)|g_t], \; \Psi \in \Psi^*. \tag{4.14}$$

For it, let us accept

Agreement 4.2 (Linear span over h-entries; Orthogonality) *The set of real functions defined on Ψ^* and spanned linearly over entries of the vector function $h(\Psi)$, see (4.7)*

$$e^* \equiv \left\{ \sum_{i\in i^*\equiv\{1,\ldots,\mathring{h}\}} \alpha_i h_i(\Psi) \text{ with arbitrary real weights } \alpha_i \right\} \tag{4.15}$$

is embedded into a linear space E^* with the scalar product $< \cdot, \cdot >$. For $A, B \in E^*$ it is defined $< A, B > \equiv \int_{\Psi^*} A(\Psi)B(\Psi)\, d\Psi$.

The orthogonal complement $e^{*\perp}$ of e^* to E^* is the set

$$e^{*\perp} \equiv \{A \in E^* \setminus 0 : < A, B >= 0 \; \forall B \in e^*\}. \tag{4.16}$$

The restriction (4.12) on the constructed expectation $E(\Psi)$ (4.14) can be written in the equivalent form

$$g_{i;t} - g_{i;0} =< h_i, E >, \ \forall i \in i^* \equiv \{1, \ldots, \mathring{h}\}. \tag{4.17}$$

This form and definition (4.16) imply that if some $E(\cdot) \in E^*$ fulfills (4.17) then also $E(\cdot) + B(\cdot)$ does whenever $B(\cdot) \in e^{*\perp}$. All such options are equivalent. At the same time, we have exploited all relevant information sources: the prior pdf $f(\Theta)$, the parameterized model $M(\Theta, \Psi)$ and data compressed into the recursively feasible statistics g_t. We have no reason to prefer any $B(\cdot) \in e^{*\perp}$. Consequently, we consider that $E(\cdot)$ (refexpeE)

$$E(\Psi) \in e^* \Leftrightarrow \mathcal{E}[f_t(\Psi)|g_t] = h'(\Psi)\alpha_t, \text{ for an } \mathring{h}\text{-vector of weights } \alpha_t(g_t). \tag{4.18}$$

This assumption and the condition (4.17) determines the weights as the solution to the set of linear algebraic equations

$$g_t - g_0 = H\alpha_t, \ H_{ij} \equiv< h_i, hj >, \ i, j \in i^*. \tag{4.19}$$

The constant square matrix H is positive semi-definite. Its regularity can simply be guaranteed by a proper choice of $h(\cdot)$. Then, the weights α and consequently of $E(\cdot)$ are uniquely determined. Thus, we can summarize the overall algorithm:

Algorithm 4.1 (Recursive estimation based on equivalence approach) *Off-line phase*

1. *Select the parameterized model* $f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) \equiv M(\Theta, \Psi_t)$.
2. *Select the prior pdf* $f(\Theta)$.
3. *Select the (generalized) vector function* $G(\Theta)$ *such that* $\int G(\Theta)\, d\Theta = 0$.
4. *Prepare evaluation of the function*

$$h(\Psi) \equiv \int \ln(M(\Theta, \Psi))G(\Theta)\, d\Theta. \tag{4.20}$$

5. *Compute the positive definite matrix* $A \equiv [\int h(\Psi)h'(\Psi), d\Psi]^{-1}$.
6. *Prepare evaluation of the function*

$$H(\Theta) \equiv \int \ln(M(\Theta, \Psi))h(\Psi)\, d\Psi. \tag{4.21}$$

7. *Set* $t = 0$ *and* α_0.

On-line use phase for $t \in t^*$

1. *Measure data vector* Ψ_t.
2. *Evaluate* $h(\Psi_t) \equiv \int \ln(M(\Theta, \Psi_t))G(\Theta)\, d\Theta$.
3. *Update weights* $\alpha_t = Ah_t + \alpha_{t-1}$.
4. *Exploit the approximation of the posterior pdf*

$$\hat{f}(\Theta|g_t) \propto f(\Theta) \exp\left[H'(\Theta)\alpha_t\right]. \tag{4.22}$$

Remark(s) 4.2

1. *Zero initial weights α_0 reflect the fact that they are always determined by the increment $g_t - g_0$, see (4.19).*
2. *The term "equivalence" approach stresses the fact that the set of posterior pdfs is reduced to equivalence classes: pdfs with the same representation g cannot be distinguished.*
3. *The chosen form of the expectation (4.18) determining the final solution is definitely not unique. It has, however, the smallest norm among possible options and as such it has tendency to give the most flat estimate of the posterior pdf. This seems to be wise as the approximation involved adds uncertainty to the estimation results.*
4. *The required commutativity of the updating and projecting the posterior pdfs is crucial. The recursion for gs is exact and the approximation errors caused by the use of $\hat{f}(\Theta|g_t)$ instead of $f(\Theta|\mathcal{P}_{a^*_{t+1}}) \equiv f(\Theta|f_t(\cdot))$ do not accumulate! Use of a non-commutative projection $G_t : f^*(\Theta|\mathcal{P}_{a^*_t}) \to g^*_t$ is always endangered by divergence: the estimation described by the Bayes rule can be viewed as a dynamic system evolving $f(\Theta|\mathcal{P}_{a^*_{t+1}})$ at stability boundary.*
5. *The indicated integrations represent computationally the most demanding part of the algorithm. They can be performed in off-line mode if their results can be efficiently stored (the resulting functions interpolated).*

Problem 4.1 (How to choose the mapping G?) *The (generalized) functions G represent the key tuning knobs of the approach. Options leading to discretisation of the function and/or its derivatives, or $M(\Theta, \Psi_i)$ on a grid of Ψ_i have been tried with a success but a deeper insight is needed in order to arrive to a cook-book.*

Problem 4.2 (Does exist an equivalent for filtration?) *There are indications that answer is negative. If it is true, other approaches relying on usual stabilising effect of time-updating (2.40) have to be used. They, however, should have a smooth transition to the estimation as its special case.*

4.2.2 Estimation with forgetting

The case of *slowly varying parameters* is discussed here as an intermediate, widely met, case bridging estimation and filtering. It admits time variations of Θ_t but it assumes that $\Theta_{t+1} \approx \Theta_t$. At the same time, more precise model of changes the is not used.

We summarize an approach called *stabilised forgetting*, [53]. It is based on a flexible problem formulation and a simple application of Proposition 3.4. It proved to be a powerful tool for solving the considered ill-posed problem.

Let $f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a^*_{t+1}})$ be pdf that assumes no parameter changes and $f_a(\Theta_{t+1} = \Theta|\mathcal{P}_{a^*_{t+1}})$ an *alternative pdf* that describes parameters after some changes. Let $\phi \in (0, 1]$ be the probability that the "correct" unknown pdf $f(\Theta) \equiv f(\Theta_{t+1} = \Theta|\mathcal{P}_{a^*_{t+1}})$ has the best projection equal to the former pdf and $1 - \phi$ the probability that the later one is relevant. Then, the best compromise minimizing its expected Kullback-Leibler divergence to the unknown best projection is, see Proposition 3.4,

$$f(\Theta_{t+1} = \Theta|\mathcal{P}_{a^*_{t+1}}) \propto \left[f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a^*_{t+1}})\right]^{\phi} \left[f_a(\Theta_{t+1} = \Theta|\mathcal{P}_{a^*_{t+1}})\right]^{1-\phi}. \tag{4.23}$$

Using this formula, we approximate the time updating step (2.40) in filtering without explicitly specifying the model of time evolution (2.35).

Algorithm 4.2 (Stabilised forgetting) *Select the probability* $\phi \in (0,1]$ *and specify prior pdf* $f(\Theta_1 = \Theta) \equiv f(\Theta_1 = \Theta|\mathcal{P}_{a_1^*})$ *corresponding to the treated parameterized (observation) model* $f(\Delta_t|\Theta_t, \mathcal{P}_{a_t^*}, a_t)$.

Repeat for time $t \in t^*$

1. *Collect the newest data* $\mathcal{P}_{a_{t+1}^*}$, a_{t+1}
2. *Perform data updating*

$$f(\Theta_t = \Theta|\mathcal{P}_{a_{t+1}^*}) \propto f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) f(\Theta_t = \Theta|\mathcal{P}_{a_t^*}).$$

3. *Select or update the alternative pdf* $f_a(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}^*})$.
4. *Approximate time updating (*forget*)*

$$f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}^*}) \propto \left[f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a_{t+1}^*})\right]^{\phi} \left[f_a(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}^*})\right]^{1-\phi}.$$

Remark(s) 4.3

1. *The time-updated pdf (4.23) is a compromise between the posterior pdf obtained under the hypothesis that* Θ_t *is time-invariant and an externally supplied alternative* f_a. *The closer* ϕ *is to unity the slower changes are expected, i.e. the higher weight the posterior pdf corresponding to the time-invariant case gets.*
2. *Note that the* forgetting operation *(4.23) preserves the basic property of time-updating: the posterior pdf on parameters propagates without obtaining any new measured information.*
3. *Let us assume that* $f_a \propto 1$. *Then* $f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}^*}) \propto \left[f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}^*})\right]^{\phi}$, *i.e. the time-updated pdf is a flattened version of the pdf obtained after data-updating (if* $\phi < 1$*). It is intuitively appealing as our uncertainty about parameters can hardly decrease without knowing a good time-evolution model (2.35) and with no new information processed.*
4. *It is instructive to inspect the influence of forgetting on old data built in through the observation model. The older data are the higher flattening is applied to the corresponding model. Consequently, the older data influence the estimation results less than new ones. Data are gradually "forgotten". This explains why the probability* ϕ *is referred as a* forgetting factor,
5. *The alternative pdf assigned to to* Θ^* *expresses our belief where the parameters might move within the time interval* $[t, t+1)$ *while we have no new observable information.*

   Often, the pessimistic uniform alternative pdf ($\propto 1$*) has been used. This special case of stabilised forgetting is called* exponential forgetting. *It allows us to follow relatively fast parameter changes but it forgets the accumulated information with, often to high, exponential rate. For this reason, it is worth to preserve what we feel as a guaranteed information. The prior pdf* $f(\Theta_{t+1} = \Theta)$ *is a typical, reasonably conservative choice of the alternative pdf* $f_a(\Theta_{t+1} = \Theta|\mathcal{P}_{a_t^*})$.

6. *The non-trivial alternative pdf prevents us to forget the "guaranteed" information as it is always incorporated after flattening (exponential forgetting). This stabilises whole learning and reflects very positively in its numerical implementations. Without this, the posterior pdf may become too flat whenever the information brought by new data is not sufficient.*

   Note that insufficient informativeness of new data is more rule than exception. It is true especially in regulation problems in which the controller tries to make closed-loop as quiet as possible, i.e. it tries to suppress any new information in data.

7. *The forgetting factor ϕ can be either taken as tuning knob or estimated. The predictive pdf parameterized by it, however, depends on it in a very complex way so that a partitioned estimation has to be applied when its posterior pdf is estimated on a prespecified grid [54], see Section 4.3.1.*

   Also, the forgetting factor can be chosen in a pessimistic way as the maximiser of the reached minima, see Proposition 3.4.

8. *Practical importance of this particular case of slowly varying parameters cannot be overstressed: vast majority of adaptive systems, see Section 4.1, rely on a version of forgetting combined with a version of receding horizon strategy, see Section 4.4.1. With it, the estimation and prediction are approximated locally within time domain. They rely on data measured nearby the actual time.*

4.2.3 Quasi-Bayes estimation

The quasi-Bayes estimation described here is a special approximation techniques designed for mixtures of models introduced in Section 3.1.3. The standard Bayesian estimation, see Section 2.5.2, is infeasible with this parameterized model

$$f(\Delta_t|\Theta,\mathcal{P}_{a_t^*},a_t)=\sum_{c\in c^*}\alpha_c f_c(\Delta_t|\Theta_c,\mathcal{P}_{a_t^*},a_t),\quad \Theta\in\Theta^*\equiv$$
$$\left\{\Theta=(\alpha_c,\Theta_c),\ \Theta_c\in\Theta_c^*,\ c\in c^*\equiv\{1,\dots,\mathring{c}<\infty\},\ \alpha\in\alpha^*\equiv\left\{\alpha_c\geq 0,\ \sum_{c\in c^*}\alpha_c=1\right\}\right\}$$

as the number of terms in the likelihood function (**??**) (a product of parameterized models) blows up exponentially. The quasi-Bayes estimation is justified as follows.

Let us have the generalized Bayesian estimate in the form

$$f(\Theta|\mathcal{P}_{a_t^*})\propto\prod_{c\in c^*}f(\Theta_c|\mathcal{P}_{a_t^*})\alpha_c^{v_{c;t-1}-1} \tag{4.24}$$

with some pdfs $f(\Theta_c|\mathcal{P}_{a_t^*})$ and scalars statistics $v_{c;t-1}>0$.

Introducing the random unobservable selector $c_t\in c^*$ of components, we defined extension of the parameterized model to the joint pdf of c_t and innovation (3.7). Using it, the assumed form (4.24), Bayes rule (2.8) and definition of Kronecker symbol $\delta(c,\tilde{c})$, (7.1), we get

$$\begin{aligned} f(\Theta,c_t|\mathcal{P}_{a_{t+1}^*}) &\propto f_{c_t}(\Delta_t|\Theta_{c_t},\mathcal{P}_{a_t^*},a_t)\alpha_{c_t}\prod_{c\in c^*}f(\Theta_c|\mathcal{P}_{a_t^*})\alpha_c^{v_{c;t-1}-1}= \\ &= \prod_{c\in c^*}[f_{c_t}(\Delta_t|\Theta_{c_t},\mathcal{P}_{a_t^*},a_t)]^{\delta(c,c_t)}f(\Theta_c|\mathcal{P}_{a_t^*})\,\alpha_c^{v_{c;t-1}+\delta(c,c_t)-1}. \end{aligned} \tag{4.25}$$

The result of the data updating for Θ could be gained from this pdf using marginalisation in order to get rid off unobservable c_t, see Proposition 2.4. It would, however, destroy the product form (4.24) of the posterior pdf. In order to avoid this trap, we simply replace the unknown $\delta(c, c_t)$ by its point estimate. We take the conditional expectation of $\delta(c, c_t)$ as this estimate. It is the optimal point estimate corresponding at least to the quadratic loss, see Remarks 3.8:

$$\begin{aligned} \delta(c, c_t) &\approx w_{c;t} \equiv \mathcal{E}[c_t = c|\mathcal{P}_{a^*_{t+1}}] = f(c_t = c|\mathcal{P}_{a^*_{t+1}}) \propto \\ &\propto (v_{c;t-1} + 1) \int f_c(\Delta_t|\Theta_c, \mathcal{P}_{a^*_t}, a_t) f(\Theta_c|\mathcal{P}_{a^*_t})\, d\Theta_{c_t}. \end{aligned} \tag{4.26}$$

Note that the expectation over unknown α is evaluated using formulae related to Dirrichlet distribution, see Chapter 6.

The substitution of a point estimate preserves the product form (4.24) and whole procedure can be prolonged. We summarize the algorithm for exponential family (2.53) for which it is of practical interest. The results of Propositions 2.16, 3.3 are used.

Algorithm 4.3 (Quasi-Bayes estimation in exponential family)

Off-line phase

1. *Select the number of components and their structure, i.e. the form of* $f_c(\Delta_t|\Theta_c, \mathcal{P}_{a^*_t}, a_t)$.
2. *Select prior statistics for estimation of parameters of individual components* $V_{c,0}$, $\nu_{c,0} > 0$, $c \in c^*$.
3. *Select prior statistics* $v_{c,0} > 0$, $c \in c^*$ *for estimation of the weights* α_c *of components.*

On-line phase, do for $t \in t^*$

1. *do for* $c \in c^*$
   (a) *Collect the newest data* Δ_t, a_t,
   (b) *Construct data vectors* $\Psi_{c,t}$,
   (c) *Compute weights (4.26)*

$$w_{c;t} \propto \frac{\mathcal{I}(V_{c;t-1} + B(\Psi_{c;t}), \nu_{c;t-1} + 1)}{\mathcal{I}(V_{c;t-1}, \nu_{c;t-1})}, \tag{4.27}$$

   (d) *Update statistics*

$$\begin{aligned} V_{c;t} &= V_{c;t-1} + w_{c;t} B(\Psi_{c;t}),\ \nu_{c;t} = \nu_{c;t-1} + w_{c|t} \\ v_{c;t} &= v_{c;t-1} + w_{c;t}. \end{aligned} \tag{4.28}$$

2. *Evaluate required characteristics of the approximate posterior pdf*

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) \propto \prod_{c \in c^*} A^{\nu_{c;t}}(\Theta_c) \exp[V'_{c;t} C(\Theta_c)] \alpha^{v_{c;t}-1}. \tag{4.29}$$

Remark(s) 4.4

1. *The used approximation is intuitively appealing. The more we expect that the innovations Δ_t for given a_t has been generated by some component the higher weight to the increment $B(\Psi_{c;t})$ of the corresponding statistics we assign.*

2. *It is straightforward and computationally cheap to allow slow parameter changes by employing the stabilised forgetting, see Algorithm 4.2.*

3. *Unlike other procedures applied to mixtures, it fits well for* on-line adaptive learning *of dynamic systems (with non-trivial regressors).*

4. *All known procedures are sensitive to initiation, [38]. It applies to the quasi-Bayes estimation even in a wider extent. This is inevitable price for its advantages.*

5. *The components describing multivariate innovations can be always factorized into product of parameterized models describing predicting single entry (single output). Each of these* factor*s may have its structure and may be even of a different type. Thus, we can combine, for instance, regression models with Markov chains. In this way, we get mixture model describing interrelated continuous and discrete data. It is of an extreme practical importance, especially in connection with the universal approximation property of mixtures.*

Problem 4.3 (How to make quasi-Bayes estimation complete?) *The quasi-Bayes algorithm proved to be powerful complement to universal and flexible mixtures. A lot of subproblems still await for their solutions: initiation, structure estimation, multi-step control design etc.*

4.3 Suboptimal filtering

Filtering, see Proposition 2.13, is the basic technique for generating the outer model of the system needed for (sub)optimal decision making. There is a bunch of complementary general techniques that convert optimal learning into practically applicable suboptimal learning. They are discussed here.

4.3.1 Partitioned approach

This approach [54] suits to cases when the problem formulation, determined by the observation $f(\Delta_t|\Theta_t, \mathcal{P}_{a_t^*}, a_t)$ and time evolution $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t^*}, a_t)$ models, has the following structure:

- $\Theta_t = (\Theta, x_t) \equiv$
  (time-invariant parameters, state of the system)
- filtering of x_t is solvable for any fixed $\Theta \in \Theta^*$
- a finite (low-dimensional) grid $\{\Theta_i \in \Theta^*\}_{i\in i^*}$ "covers" sufficiently the space of possible parameter values Θ^*.

The partitioned approach is described by the following algorithm:

Algorithm 4.4 (Partitioned filtering)

Off-line phase
Select prior probabilities $f(\Theta_i|\mathcal{P}_{a_1^*})$ *of individual parameter values on the grid* $\{\Theta_i\}_{i\in i^*}$ *and set time* $t = 0$.

On-line phase, do for $t \in t^*$

1. *Repeat for* $i \in i^*$

   (a) *Collect the newest data* Δ_t, a_t,

   (b) *Perform the conditional* data updating

$$f(x_t|\mathcal{P}_{a_{t+1}^*}, \Theta_i) = \frac{f(\Delta_t|x_t, \Theta_i, \mathcal{P}_{a_t^*}, a_t) f(x_t|\mathcal{P}_{a_t^*}, \Theta_i)}{f(\Delta_t|\mathcal{P}_{a_t^*}, a_t)} \propto$$
$$\propto f(\Delta_t|x_t, \Theta_i, \mathcal{P}_{a_t^*}, a_t) f(x_t|\mathcal{P}_{a_t^*}, \Theta_i). \tag{4.30}$$

   (c) *Perform the conditional* time updating

$$f(x_t|\mathcal{P}_{a_t^*}, \Theta_i) = \int_{x_{t-1}^*} f(x_t|x_{t-1}, \Theta_i, \mathcal{P}_{a_t^*}, a_t) f(x_{t-1}|\mathcal{P}_{a_t^*}, \Theta_i)\, dx_{t-1}. \tag{4.31}$$

   (d) *Evaluate the conditional* prediction

$$f(\Delta_t|\mathcal{P}_{a_t^*}, a_t, \Theta_i) = \int_{x_t^*} f(\Delta_t|x_t, \mathcal{P}_{a_t^*}, a_t, \Theta_i) f(x_t|\mathcal{P}_{a_t^*}, \Theta_i)\, dx_t. \tag{4.32}$$

   (e) *Update* probabilities of individual grid points

$$f(\Theta_i|\mathcal{P}_{a_{t+1}^*}) \propto f(\Delta_t|\mathcal{P}_{a_t^*}, a_t, \Theta_i) f(\Theta_i|\mathcal{P}_{a_t^*}). \tag{4.33}$$

2. *Compute the approximate outer model of the system*

$$f(\Delta_{t+1}|\mathcal{P}_{a_{t+1}^*}, a_{t+1}) = \sum_{i\in i^*} f(\Delta_{t+1}|\mathcal{P}_{a_{t+1}^*}, a_{t+1}, \Theta_i) f(\Theta_i|\mathcal{P}_{a_{t+1}^*}). \tag{4.34}$$

Remark(s) 4.5

1. *Essentially, a* bank of filters *"indexed" by* Θ_i *is run and probabilities of individual predictors are updated in a standard Bayesian way.*

2. *The approach relies on feasibility of individual filters and on the low dimension of the grid. Bootstrap type estimator, [55] and see Section 4.3.2, seems to be an efficient way how to reach it. The time variations of the used grid bring, however, additional troubles.*

Problem 4.4 (How to select the grid and exploit results?) *Acceptable dimensionality of the grid is of a vital importance and it is often reached by adopting additional techniques of a "branch-and-bound" type. It is straightforward to remove* Θ_i *with low probabilities* $f(\Theta_i|\mathcal{P}_{a_t^*})$ *but generating of new promising values* Θ_i *is not easy. Moreover, the exploitation of the filtering results obtained so far has no systematic solution in multi-step control tasks [56].*

4.3.2 Bootstrap type methods

Knowing the posterior pdf or pdfs describing filtration it would be easy to select a low dimensional non-uniform grid sufficient for a good approximation of the whole pdf. This simple observation forms a basis of a whole set approximate filters that use random samples from the current approximate pdfs for generating new approximations and exploitation of new data for data updating [57]. We describe a version of such *bootstrap filtration* methods that exploit interpretation of random samples as a variable grid.

Under natural conditions of decision making (2.36), we search for practically feasible filtration $f_{t|t}(\Theta) \equiv f(\Theta_t = \Theta|\mathcal{P}_{a^*_{t+1}}, a_{t+1}) = f(\Theta_t = \Theta|\mathcal{P}_{a^*_{t+1}})$. The internal quantities Θ_t evolve according to the time evolution model (2.35) $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a^*_t}, a_t)$ and reflect in the observation model $f(\Delta_t|\Theta_t, \mathcal{P}_{a^*_t}, a_t)$, see Requirement 2.5. During the filtering also the prediction $f_{t+1|t}(\Theta) \equiv f(\Theta_{t+1} = \Theta|\mathcal{P}_{a^*_{t+1}}, a_{t+1})$ of Θ_t is computed, see Proposition 2.13. Let $\Theta_{i;t|t-1},\ i \in i^* \equiv \{1, \dots, \mathring{i} < \infty\}$ denote random independent samples of Θ taken from the approximate pdf $f_{t|t-1}(\Theta)$. We assume that the values $f_{t|t-1}(\Theta_{i;t|t-1})$ are known. Then, data updating (2.39) gives simply values

$$f_{t|t}(\Theta_{i;t|t-1}) \propto f(\Delta_t|\Theta_{i;t|t-1}, \mathcal{P}_{a^*_t}, a_t) f_{t|t-1}(\Theta_{i;t|t-1}).$$

We use the pairs $\{\Theta_{i;t|t-1}, f_{t|t}(\Theta_{i;t|t-1})\}_{i\in i^*}$ for a rough approximation of $f_{t|t}(\Theta),\ \Theta \in \Theta^*$. The simplest approximation by empirical pdf is used here

$$f_{t|t}(\Theta) \equiv \sum_{i\in i^*} \delta(\Theta - \Theta_{i;t|t-1}) w_{i;t|t},\ \ w_{i;t|t} \propto f_{t|t}(\Theta_{i;t|t-1}). \tag{4.35}$$

δ is Dirac delta and $\sum_{i\in i^*} w_{i;t|t} = 1$. With the approximation (4.35), the time updating (2.40) is given by the formula

$$f_{t+1|t}(\Theta) = \sum_{i\in i^*} f(\Theta_{t+1}|\Theta_{i;t|t-1}, \mathcal{P}_{a^*_{t+1}}, a_{t+1}) w_{i;t|t}.$$

The filtering recursion is then closed by generating random independent samples $\Theta_{i;t+1|t},\ i \in i^*$ from this pdf. The predictive pdf is obtained simply by approximating the pdf $f_{t+1|t}(\Theta)$ by empirical pdf on the latest samples and using the standard prediction formula (2.38). Let us summarize

Algorithm 4.5 (Bootstrap filtering)

Off-line phase

Generate independent samples $\{\Theta_{i;1|0}\}_{i\in i^*}$ *from the prior pdf* $f_{1|0}(\Theta) \equiv f(\Theta_1|\mathcal{P}_{a^*_1})$ *and set time* $t = 0$.

On-line phase, do for $t \in t^*$

1. *Collect the newest data* Δ_t, a_t.
2. *Perform* data updating *on samples*

$$w_{i;t|t} \propto f_{t|t}(\Theta_{i;t|t-1}) \propto f(\Delta_t|\Theta_{i;t|t-1}, \mathcal{P}_{a^*_t}, a_t) f_{t|t-1}(\Theta_{i;t|t-1}), i \in i^*. \tag{4.36}$$

3. *Perform* time updating

$$f_{t+1|t}(\Theta) = \sum_{i\in i^*} f(\Theta_{t+1} = \Theta|\Theta_{i;t|t-1}, \mathcal{P}_{a^*_{t+1}}, a_{t+1}) w_{i;t}. \tag{4.37}$$

4. *Generate independent samples* $\{\Theta_{i;t+1|t}\}_{i\in i^*}$ *from the pdf* $f_{t+1|t}(\Theta)$ *and evaluate* $w_{i;t+1|t} \propto f_{t+1|t}(\Theta_{i;t+1|t})$.

5. *Compute (approximate outer) system model*

$$f(\Delta_{t+1}|\mathcal{P}_{a^*_{t+1}}, a_{t+1}) = \sum_{i\in i^*} f(\Delta_{t+1}|\Theta_{i;t+1|t}, \mathcal{P}_{a^*_{t+1}}, a_{t+1}) w_{i;t+1|t}. \tag{4.38}$$

Remark(s) 4.6

1. *The algorithm is conceptually simple and universal. It needs a large number of samples. Their generating is generally far from being trivial. For a more complex cases it might make the algorithm infeasible. In other cases sophisticated techniques like Gibbs sampler [58] or adaptive generating [47] might be needed.*

2. *A massive parallelism might be used for making the algorithm computationally efficient.*

3. *Less crude approximations of pdfs than the empirical pdfs are desirable and necessary when time evolution is degenerated and some entries of* Θ_t *are time invariant. Typically, smooth approximations of Dirac delta are used. A narrow Gaussian pdf "sitting" on the sample may be used. This can be interpreted as addition of an external noise to the original samples. Spline-type or mixture approximations might be used in some cases.*

4. *This group of methods is very adaptive, see Section 4.1 It generates approximations locally around measured behavior.*

Problem 4.5 (How to make work bootstrap complete?) *A lot has been reached in development of methods of this type. The state is, however, far from being finished. Problems range from sensitivity to initial values, efficient generating of samples, choice of approximating pdfs up to the choice of the number of samples and numerical (parallel) implementation.*

4.3.3 Extended Kalman filtering

Filtering can be exactly solved if both observation and time-evolution models are Gaussian with mean values depending linearly on the unknown Θ_t and with known covariances. All involved pdfs are Gaussian if the prior pdf is. Then, the filtering reduces to an algebraic evolution of two conditional moments, [10, 59]. In this case, the time-evolution model can be characterized by

$$\begin{aligned} \mathcal{E}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a^*_{t+1}}, a_{t+1}] &= A(\mathcal{P}_{a^*_{t+1}}, a_{t+1})\Theta_t + B(\mathcal{P}_{a^*_{t+1}}, a_{t+1}) \\ \mathrm{cov}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a^*_{t+1}}, a_{t+1}] &= R(\mathcal{P}_{a^*_{t+1}}, a_{t+1}). \end{aligned} \tag{4.39}$$

A is a square matrix, B is a vector and R is positive semi-definite square matrix.

Similarly, the observation model is characterized by

$$\begin{aligned} \mathcal{E}[\Delta_{t+1}|\Theta_{t+1}, \mathcal{P}_{a^*_{t+1}}, a_{t+1}] &= C(\mathcal{P}_{a^*_{t+1}}, a_{t+1})\Theta_{t+1} + D(\mathcal{P}_{a^*_{t+1}}, a_{t+1}) \\ \mathrm{cov}[\Delta_{t+1}|\Theta_{t+1}, \mathcal{P}_{a^*_{t+1}}, a_{t+1}] &= r(\mathcal{P}_{a^*_{t+1}}, a_{t+1}). \end{aligned} \tag{4.40}$$

C is a matrix, D a vector, and r is a positive semi-definite square matrix.

All arrays can be *known functions of available data and time.*

The *extended Kalman filter* is obtained if we approximate the treated conditional expectation $\mathcal{E}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a^*_{t+1}}, a_{t+1}]$ by the affine term of Taylor expansion around the latest point estimate $\hat{\Theta}$ of Θ_t and $\mathrm{cov}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a^*_{t+1}}, a_{t+1}]$ by the absolute term at the same point. Through a similar expansion for the observation model we arrive to forms (4.39), (4.40) and use them for an approximate updating of two moments of the considered filtration by using ordinary Kalman filter. These estimates serve for updating the expansion point.

Remark(s) 4.7

1. *Obviously, the local approximation used in extended Kalman filter is a rather crude. Consequently, a careful start up and various modifications based on a detailed analysis are needed in order to prevent divergence of* $\{\hat{\Theta}_t\}_{t\in t^*}$ *from* $\{\Theta_t\}_{t\in t^*}$*, [60].*
2. *The smoother the model is the higher is chance for success.*

Problem 4.6 (How to combine techniques?) *Naturally, the extended Kalman filtering can be combined with partitioning approach, bootstrap methods or mixture estimation, see Section 4.2.3. A systematic way is (to our best knowledge) not available.*

4.4 Suboptimal design

4.4.1 Receding horizon

The reduction of the design horizon is the most obvious way to a simplified (suboptimal) design. The reduction obtained by planning just one-step-ahead has been popular for a long time. Dynamic decision making, however, means that consequences of a decision are encountered far behind the time moment of its application. Consequently, the decision that is optimal when judged from a shortsighted perspective might be quite bad one from the long-term viewpoint, [9].

This observation has stimulated search for a compromise between the ideal planning over whole horizon of interest and shortsighted, locally optimizing decisions. Knowing that the constructed plan is just an approximation, it is reasonable to correct it permanently. After applying the initial decision(s) implied by the temporary design and after acquiring and exploiting a new measured knowledge about the system state, the design should be repeated. This is essence of the design technique called *receding horizon strategy*. Let us describe its algorithm in the case of additive loss function (2.22), recall (2.31), and for a pre-specified value T of the receding horizon $T < \mathring{t}$.

Algorithm 4.6 (Receding-horizon strategy) *Repeat for* $t = 1, \ldots, \mathring{t}$,

1. *Find the strategy* $\mathcal{R}(t\_t + T) = (\mathcal{R}_t, \ldots, \mathcal{R}_{t+T})$ *approximately minimizing*
$$\mathcal{E}\left[\sum_{\tau=t}^{t+T} z(\Delta(\tau), a(\tau))|\mathcal{P}_{a^*_t}\right]$$
2. *Apply* $a_{t|t-1} \equiv \mathcal{R}_t(\mathcal{P}_{a^*_t})$ *for the available* $\mathcal{P}_{a^*_t}$.
3. *Measure new data* Δ_t, a_t *and use extended experience (perform filtration or estimation) for obtaining improved outer model of the system.*

4.4.2 Super-cautious strategy

Let us consider the receding horizon strategy applied at time t. Then, a substantial degree of its complexity is caused by the use of predictive pdfs $\{f(\Delta_\tau|\mathcal{P}_{a^*_\tau}, a_\tau)\}_{\tau=t}^{t+T}$ obtained through Bayesian estimation. They have the form, see Proposition 2.14,

$$f(\Delta_\tau|\mathcal{P}_{a^*_\tau}, a_\tau) = \int_{\Theta^*} f(\Delta_\tau|\Theta, \mathcal{P}_{a^*_\tau}, a_\tau) f(\Theta|\mathcal{P}_{a^*_\tau})\, d\Theta. \tag{4.41}$$

For a relatively short planning horizon T and a reasonably concentrated parameter estimate $f(\Theta|\mathcal{P}_{a^*_t})$, the predictors (4.41) can be approximated as follows

$$f(\Delta_\tau|\mathcal{P}_{a^*_\tau}, a_\tau) \approx \int_{\Theta^*} f(\Delta_\tau|\Theta, \mathcal{P}_{a^*_\tau}, a_\tau) f(\Theta|\mathcal{P}_{a^*_t})\, d\Theta,\ \tau = t, \dots, t+T. \tag{4.42}$$

This approximation relies on the approximate equality

$$f(\Theta|\mathcal{P}_{a^*_\tau}) \approx f(\Theta|\mathcal{P}_{a^*_t}),\ \tau = t, \dots, t+T. \tag{4.43}$$

The receding horizon strategy combined with the approximation (4.42) is called *super-cautious* as it is based on assumption that we learn nothing about unknown parameters during the planning horizon. The uncertainty of the parameter estimates (4.43) projected into the predictors through the formula (4.42) inhibits excessive decision values. It makes the strategy (typically controller) (super)cautious.

4.4.3 Cautious strategy

The assumption (4.43) is often too strong and it is reasonable to weaken it to

$$f(\Theta|\mathcal{P}_{a^*_\tau}) \approx F(\tau, f(\Theta|\mathcal{P}_{a^*_t})),\ \tau = t, \dots, t+T,\ \Theta \in \Theta^*. \tag{4.44}$$

The mapping $F(\tau, \cdot)$ modifies the estimate $f(\Theta|\mathcal{P}_{a^*_t})$ to other, more concentrated, pdfs. The mapping F is chosen beforehand and does not use data belonging to ignorance $\mathcal{F}_{a^*_t}$. Typically, the generated pdf has the same 1st moment (expectation) as $f(\Theta|\mathcal{P}_{a^*_t})$ but its covariance $C_{\tau|t}$ decreases with increasing distance $\tau - t$ in a pre-specified way. Often, $C_{\tau|t} = b(\tau - t)C_t \equiv b(\tau - t)\times$ covariance computed for $f(\Theta|\mathcal{P}_{a^*_t})$. The following non-negative scalar functions $b(\kappa)$ serve as representative examples

$$b(\kappa) = 1/\kappa \text{ or } b(\kappa) = \begin{cases} 1 & \text{if } \kappa = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{4.45}$$

4.4.4 Certainty-equivalence strategy

This strategy is the most wide spread one. It replaces unknown parameter in the parameterized model by a current point estimate $\hat{\Theta}_t$ of unknown parameters

$$f(\Delta_\tau|\mathcal{P}_{a^*_\tau}, a_\tau) \approx f(\Delta_\tau|\hat{\Theta}_t, \mathcal{P}_{a^*_\tau}, a_\tau),\ \tau = t, \dots, t+T. \tag{4.46}$$

which corresponds to the approximate parameter estimate

$$f(\Theta|\mathcal{P}_{a^*_\tau}) \approx \delta(\Theta - \hat{\Theta}_t),\ \tau = t, \dots, t+T. \tag{4.47}$$

where $\delta(\cdot)$ is Dirac delta function, a formal pdf of the measure concentrated on zero value.

4.4.5 Predictive strategies

Up to now, the reduction of complexity of predictors have been reached by simplifying parameter estimates. There is, however, a whole set of strategies that simplify prediction by working only with point predictions of outputs.

Due to a bit unlucky vocabulary they are known as *predictive strategies* [49, 61]. Obviously, such an approximation may be quite rough when there is a non-negligible uncertainty (lack of knowledge or influence of noise). On the other hand, the gained simplification allows the user to respect hard bounds on behavior. The importance of technologic restrictions implies popularity of such a treatment.

4.4.6 Active strategies

All outlined strategies are *passive*: when planning their decisions they do not care about learning. At the same time, it is known [5, 6] that the optimal strategy is *dual strategy*. It cares both about immediate decisions and learning process in a balanced way. This motivates a use of measures that support learning. The resulting strategies are known as *active strategies*. Two basic ways are used:

- an external stimulating signal is feeded into the closed loop (added to optional quantities like inputs or set points),
- a term is added to the loss function that reflects quality of learning even when a passive-type design is applied [62].

Remark(s) 4.8

1. *The essence of suboptimal designs is demonstrated mostly in combination with estimation. Similar techniques are applicable in conjunction with filtering.*
2. *Quality of approximations depends very much on quality of the parameterized model and on parameter estimates $f(\Theta|\mathcal{P}_{a_t^*}, a_t)$. Thus, the care devoted to the construction of the parameterized model and the prior pdf $f(\Theta)$, see Section 3.1.4, might decide on a success/failure of the decision making. This observation has been at roots of designing of algorithmic and software tools for the prior design of adaptive controllers [?, 63, 64].*
3. *A substantial effort with significant results has been put into analysis of various approximations. Mostly, however, the results are of an asymptotic nature. A little is known about practically significant finite time-properties of such approximations.*
4. *It is fair to say that the "dangerous" certainty equivalent strategies dominate applications. They are simple, understandable and they can be simply extended to the combination of the design with a state estimator.*
5. *For a large $T < \mathring{t}$, the receding horizon strategy gives away majority of computations and it is always in danger that the chosen horizon is too short. Thus, it is reasonable to take the strategy found with the estimate $f(\Theta|\mathcal{P}_{a_t^*})$ as an initial guess of the strategy re-computed for $f(\Theta|\mathcal{P}_{a_{t+1}^*})$. This idea is known as* iterations spread in (learning) time, *[9]. It has been found quite useful but at the same time sensitive to the time variations of the estimated parameters.*

6. *Active strategies are rarely used [3] in spite of known, theoretically proved, danger of passive strategies.*

Problem 4.7 (Is the back of trick a final stage?) *Still, the approximate design can be called more back of tricks than a systematic framework. Naturally, a systematic solution like the equivalence approach is desirable.*

Problem 4.8 (Is the fully probabilistic strategy active?) *The strategy generated by the fully probabilistic design, see Proposition 2.11, is randomized. It can be interpreted as a combination of a deterministic feedback with a dither noise. Thus, there is a chance that it is "naturally" active. This hypothesis, however, has to be inspected in detail.*

4.5 Decomposition of decision making

Repeatedly, splitting of the decision-making task in a chain of subtasks is declared as the way of converting optimal design to an approximation of the practical optimal design. At the same time, it is known that any violation of the golden decision-making rule (3.1) makes the solution to depart from optimality. Lack of the formal tools for the decomposition leaves us with empirical rules in this area.

Here give as an example the list describing such decomposition in the case of adaptive control with learning part based on parameter estimation. Each item in the list has been found as a relatively self-containing decision sub-problem. Whenever possible, we give a reference to the relevant support within the text. The design, as any human activity is iterative. Naturally, the majority of iterations should be concentrated in the off-line phase in order to minimize expenses related to the commission of the controller.

4.5.1 Off-line phase

The following indicative list of subtasks is solved, often with internal iterations, until the potential user is satisfied.

- Formulate the addressed problem, see Chapter 2
  - get the specification of technical control aims,
  - get the specification of the system,
  - get the specification of the available data, system inputs and outputs
  - get the specification of technologic and complexity restrictions
  - collect the knowledge available
- Perform experimental design and collect data.
- Make data pre-processing, see Section 3.1.1
- Select class of parameterized models, see Section 3.1.3
- Quantify prior knowledge, see Section 3.1.4

- Estimate model structure and control period, see Proposition 3.5
- Estimate forgetting factor, see Propositions 3.4, 3.5
- Perform generalized Bayesian estimation based on prior knowledge and available data; the result will be used as the prior and/or alternative pdf in on-line phase, see Algorithm 3.2
- Select the loss function (typically, weights of quadratic loss or ideal in fully probabilistic design, see Sections 3.1.5, 2.4.2
  Do until the results cannot be improved
  - Select suboptimal control design and its parameters
  - Perform prior design of controller and prediction of closed loop (coupling system $\leftrightarrow$ controller) behavior by transforming all uncertain quantities into to user-defined quality indicators, formal solution provides Proposition 2.5, generally Monte Carlo evaluation is needed that becomes feasible using stopping rules, see Section 3.4.1
  - Compare results with user's specification; set estimation is useful to it, see Section 3.3.2

4.5.2 On-line phase

These decision-making subtasks are solved in real time for $t \in t^*$. Here, there is almost no freedom for iterative trial-and-error solutions. Of course, it is wise to store the data collected during the on-line phase and use them for an improved off-line design.

- Collect and pre-process data, see Section 3.1.1
- Generate reference signals
- Estimate with stabilized forgetting, see Section 4.2.2
- Use receding horizon or iterations-spread-in time design, see Section 4.4.1, in an appropriate version, see Section 3.4.1
- Generate input using the designed strategy and measured data
- Check possible discrepancies and make a finer tuning of optional parameters of the design; this supervision can be based on theory of operator control [65], that exploits mixture estimation, see Sections 3.1.3, 4.2.3

Problem 4.9 (Completion of the design support) *The presented description just outlines the complexity of the overall chain. Obviously, some steps are weakly supported by the presented theory and algorithms. Some of them are not supported at all. In spite of this our experience indicates that the presented theory provides the most complete unified support available. Still even a partial completion of this support represents significant research challenge.*

4.6 Some numerics

The presented theory should lead to algorithms implemented in computers. Thus, their numerical realization form a substantial part of the decision making as applied science. Naturally, each decision task requires its specific approach. There is, however, common or wide spread problems that are worth to be mentioned independently. This is done here.

4.6.1 Sharp likelihoods

Probability density functions are the basic object we are dealing with. The main numerical problems with them are encountered in parameter estimation and in testing hypothesis. In these cases, they have tendency to converge to very sharp functions close to Dirac delta. This fact follows from Proposition 2.15. It means that we have to work with logarithms of likelihood functions, $\ln[\mathcal{L}(\Theta, \mathcal{P}_{a_t^*})]$ and before converting them into posterior pdfs subtract $\max_{\Theta\in\Theta^*} \ln[\mathcal{L}(\Theta, \mathcal{P}_{a_t^*})]$. It is possible because the likelihood is determined uniquely up to a Θ independent factor. It is advantageous as it prevents overflows and the dynamical range of the posterior pdfs causes at most underflows for values for which the posterior pdf is negligible.

4.6.2 Monte Carlo techniques

Monte Carlo (MC) techniques evaluate complex formulae by generating samples with a specific distribution and by forming (typically point) estimators of the results of interest. They rely on laws of large numbers that guarantee convergence of such estimators to the target value. There is a numerous variants and problems that have to be resolve in order to make this idea work efficiently. Here, we can only touch this rich, useful and interesting area.

Generating of samples

For MC techniques samples from a pre-specified distribution have to be generated. We need them, for instance, in bootstrap-based filtering, see Section 4.3.2.

We denote $x \sim f$ if the sample x is generated according to the pdf f.

Majority of software tools, like MATLAB, provide (pseudo)random number generators that provide independent samples uniformly distributed on interval (0,1), its pdf we denote $\mathcal{U}(0,1)$, or normally distributed with zero mean and unit variance, its pdf we denote $\mathcal{N}(0,1)$. Thus, we assume further on that samples of this type are at disposal. The results summarized in the following proposition serve often for solution to the task addressed.

Proposition 4.2 (Samples obtained from uniform and normal pdfs) *1. Let $f(x)$ be a pdf of a univariate real variable $x \in x^* \equiv (-\infty, \infty)$ and $r \sim \mathcal{U}(0,1)$. The $x_r \sim f$ if it solve the equation*

$$r = \int_{-\infty}^{x_r} f(x)\, dx. \tag{4.48}$$

2. Let the entries of a vector e be independent and each of them $e_i \sim \mathcal{N}(0,1) \Leftrightarrow e \sim \mathcal{N}(0, I)$ where $I =$ unit matrix. Then,

$$g \equiv \mu + Ge \sim \mathcal{N}(\mu, GG') \tag{4.49}$$

for any vector μ and matrix G of appropriate dimensions.

Proof: ad 1. The equation (4.48) has a unique solution as the integral as a function of the upper bound is continuous non-decreasing function covering interval (0,1] due to the non-negativity and normalization of pdf f, see Proposition 2.4.

Using Proposition 2.5 with $\alpha = x_r$, $\beta = r\, f_T(\beta) = \mathcal{U}(0,1) \equiv 1$ on $(0,1)$, we find $|J(\alpha)| = f(\alpha)$

ad 2. Using Proposition 2.5, it can be verified that the considered affine transformation maps preserves normality. Transformation of moments is implied either by the same Proposition or can be simply verified using basic properties of expectation, see Proposition 2.6. ♢

Usefulness of the transformation (4.48) depends heavily on our ability to solve efficiently this equation. A bunch of other techniques has been developed to cope with (predominant) cases to which it cannot be used.

Prediction of achievable quality

This section presents an example of application of Monte Carlo technique for prior prediction of achievable quality of the optimal strategy [**?**, 63, 66].

The basic scenario is as follows. For a chosen parameterized model, prior pdf $f(\Theta)$ is supposed to be constructed. Let us assume that for each Θ we are able to find the stationary optimal strategy ($\mathring{t} \to \infty$) that is not admissible as it depends on unknown parameter. Without too much loss of generality, we can assume that strategy is determined by a repetitive use of a single rule $\mathcal{R}_\Theta$. With this strategy, the achieved values of the considered stationary loss function become a function, say $Q(\Theta)$ of Θ only. These values represent lower bound on achievable quality. The admissible strategy can achieve this bound only when it successfully learn the unknown parameters.

As we are uncertain about Θ we are uncertain about values of $Q(\Theta)$. This quality indicator, however, has been computed through a very deterministic optimization. Thus, conceptually we can get pdf of Q by using Proposition 2.5. Practically, the proposition cannot be used as the mapping $T : \Theta^* \to Q^*$ is non-linear, many to one and its values are mostly evaluated numerically. Here, Monte Carlo technique can be directly used. Simply, independent samples $\Theta_i \sim f(\Theta)$, $i \in i^*$ are generated and independent samples $Q_i \equiv Q(\Theta_i)$ computed in the mentioned deterministic way. Then, the desired pdf $f(Q)$ is estimated.

The estimation of the pdf $f(Q)$ is simple as Q is scalar variable.

The estimation of the pdf $f(Q)$ is complex as each sample Q result from demanding optimization. For this reason, all measures have to be taken in order to get results within acceptable time. For instance,

- A parameterized version of $f(Q)$ is estimated [**?**].
- Sequential stopping rules [47], see Section 3.4.1, are employed.
- Partial characteristics of $f(Q)$, like credibility intervals, see Section 3.3.2, are searched for.

4.6.3 Quadratic forms

Optimal designs with *q*uadratic loss functions and *G*aussian models with expected value formed by a *l*inear function of parameters and observed data are often *practically* optimal. Both learning and optimization reduce to manipulations with positive semi-definite quadratic forms. This fact is behind the predominant use of this so called *LQG design*. Moreover, other decision problems are often approximated by an LQG one. This explains why we present the summary within this general chapter.

Agreement 4.3 (Quadratic forms) *Quadratic form $\psi' Q \psi$ is the function of a real column vector $\psi \in \psi^*$, specified by a symmetric matrix $Q = Q'$. A quadratic form is said positive semi-definite if it is non-negative for any $\psi \in \psi^*$. It is called positive definite if $\psi' Q \psi > 0$, $\forall \psi \in \psi^* \backslash 0$. The matrix Q is called kernel of the form and the introduced terminology is used also for it. It means that positive (semi)definite kernel specifies positive (semi)definite quadratic form.*

Proposition 4.3 (Factorized kernels) *The kernel Q is positive definite iff it can be uniquely expressed (factorized) as follows*

$$Q = \underline{L}\underline{L}' = \underline{U}\underline{U}' = LDL' = U\underline{D}U', \quad \textit{where} \tag{4.50}$$

$\underline{L}$, $(\underline{U})$ *is lower (upper) triangular matrix with positive diagonal entries*

L, (U) *is lower (upper) triangular matrix with unit diagonal entries*

$D, \underline{D}$ *are diagonal matrices with positive entries.*

Proof: Let the first decomposition exist. Then, $\psi' Q \psi = \sum_i z_i^2 \geq 0$ with $z \equiv L'\psi$. The matrix L is regular and consequently maps non-zero ψ on non-zero z. Thus, the last inequality is sharp for $\psi \neq 0$. The similar arguments apply for the alternative decompositions.

Let Q be positive definite matrix and let us search for the first decomposition. If we write the desirable identity entry-wise exploit the definition of $\underline{L}$ and symmetry of Q, we get formal algorithm for computing $\underline{L}$

Algorithm 4.7 (Choleski algorithm)

$$\begin{array}{lll} i \in i^* & = & \{1, \ldots, \mathring{i}\} \\ & & \underline{L}_{ii} = \sqrt{Q_{ii} - \sum\limits_{k \leq i-1} \underline{L}_{ik}^2} \\ & j & = \{i+1, \ldots, \mathring{i}\} \\ & & \underline{L}_{ji} = \dfrac{Q_{ij} - \sum_{k \leq i-1} \underline{L}_{ik} \underline{L}_{jk}}{\underline{L}_{ii}} \\ & \text{end } j & \\ \text{end } i & & \end{array}$$

The direct inspection shows that all quantities on right-hand side are available in time if we understand sums over empty set as zero. It remains to judge whether square-root and division make sense. If

we, however, assume that it is not true at an ith step then we get the contradiction with positive definiteness by trying the vector ψ which is zero except of the unity on the ith position.

The LD factorization is obtained from the Choleski ones by extracting diagonal entries from $\underline{L}$. Similar considerations verify the factorizations with upper triangular matrices. ♢

Remark(s) 4.9

1. *The expressions $\underline{L}\,\underline{L}' = \underline{U}\,\underline{U}'$ are called* Choleski factorizations *as the name of the algorithm indicate.*

2. *Algorithms providing LD or UD factorizations can be derived in the same way. They are even computationally simpler as they do not need to compute square roots. It gives them alternative name* square-root free factorizations.

3. *LD (UD) factorizations are easily extended (not uniquely) to positive definite Q: whenever $D_{ii} = 0$, ($\underline{D}_{ii} = 0$) the corresponding non-diagonal entries in ith column (row) of L, (U) are set to zero.*

4. *The factorization $Q = AA'$ with a general A is not unique even for positive definite Q. It is sufficient to consider any orthogonal matrix T for which $TT' = I$ =unit matrix. Then, obviously, also AT is a factor of Q if A is.*

Proposition 4.4 (Minimization with factorized kernels) *Let $\psi' = [u', x']$ and $Q = LDL'$ with*

$$L = \begin{bmatrix} L_u & 0 \\ L_{xu} & L_x \end{bmatrix}, \; D = \begin{bmatrix} D_u & 0 \\ & D_x \end{bmatrix}$$

where dimensions of of L_u, D_u are compatible with the dimension of u. Then,

$$u = -(L'_u)^{-1} L'_{xu} x = \arg \min_{u \in u^*} \psi' LDL' \psi'. \tag{4.51}$$

The reached minimum is

$$\min_{u \in u^*} \psi' LDL' \psi' = x' L_x D_x L'_x x' \tag{4.52}$$

Proof: The minimized quadratic form can be written as as a sum of squares of two weighted quadratic norms

$$\psi' LDL' \psi = ||u' L_u + x' L_{xu}||^2_{D_u} + ||x' L_x||^2_{D_x}.$$

Only the first one depends on the optional u. It reaches minimum for the zero argument which gives the unique minimizer. The second one gives the minimum reached. ♢

Remark(s) 4.10

1. *The minimization of a quadratic form is very simple when the kernel is a factorized version. More importantly, it is numerically safe as the work with factors guarantees that the minimized form is always at least positive semi-definite even if L is evaluated with errors. The original Q lacks this property.*

2. *The advantageous numerical properties of factorized versions might be preserved when we do not need to created the complete kernel during manipulations encountered. We indicate the basic algorithm with such a property in case when two kernels should be summed.*

   Let $Q = LDL'$ and l_i, d_i ith column of L and ith diagonal entry of D_i respectively. Then, $Q = \sum_{i\in i^*} l_i d_i l_i'$ where $l_i' = [\underbrace{0,\ldots,0}_{(i-1)},1,x,\ldots,x]$ where x denotes entries with arbitrary values. If the similar notation is used for another kernel $\tilde{Q} = \tilde{L}Q\tilde{L}'$ we have

$$Q + \tilde{Q} = \sum_{i\in i^*} \left(l_i d_i l_i' + \tilde{l}_i \tilde{d}_i \tilde{l}_i' \right) \tag{4.53}$$

   It is straightforward to derive an algorithm for dyadic reduction *recomputing a sum of two weighted dyads with a common leading unit into sum of two dyads where the second gets zero of the leading unity. Taking into account (4.53), it is sufficient to take the first dyad from both decompositions, to reduce one of by the dyadic reduction. Then the reduction is called twice on the triple of dyads having unit on the second position, etc. In this way, LD decomposition of the sum of kernels is directly obtained from factor of summed terms.*

The factorized version of *precision matrix*, i.e. inversion of of covariance matrix of a vector ψ having normal pdf allows us to evaluate simply conditional and marginal pds of its sub-vectors.

Proposition 4.5 (Conditioning and marginalization for normal pdfs) *Let $\psi' = [u', x']$ be normally distributed with zero mean and $Q \equiv [\mathrm{cov}(\psi)]^{-1} = LDL'$ with*

$$L = \begin{bmatrix} L_u & 0 \\ L_{xu} & L_x \end{bmatrix}, \; D = \begin{bmatrix} D_u & 0 \\ & D_x \end{bmatrix}$$

where dimensions of of L_u, D_u are compatible with the dimension of u. Then, $f(u|x)$ is normal with conditional mean and precision matrix

$$\mathcal{E}[u|x] = -(L_u')^{-1} L_{xu}' x, [\mathrm{cov}(u|x)]^{-1} = L_u D_u L_u'. \tag{4.54}$$

The marginal pdf $f(x)$ of x is zero mean with precision matrix

$$[\mathrm{cov}(x)]^{-1} = L_x D_x L_x' \tag{4.55}$$

Proof: It is obvious by writing the quadratic form in exponent of multivariate normal form exactly as in proof of Proposition 4.4 and recall chain rule for pdfs $f(u,x) = f(u|x)f(x)$. ♢

Remark(s) 4.11

1. *Notice that the position of the variable with respect to which minimization or conditioning is performed is important and dependent on the factorization used. This simple observation is often overlooked.*

Chapter 5

Model with normal noise

This chapter illustrates the general theory on a simple example of the normal, single-input single-output, auto-regression system with external input (ARX model).

5.1 Construction elements

5.1.1 Data

The data set is formed by continuous outputs y_t and continuous inputs u_t, measured in discrete time instants $t = 1, 2, \ldots, \mathring{t}$.

Notation

In accordance with the general conventions, we use the following notations.

d_t	$= [y_t, u_t]'$	the data item at time t
$d(t)$	$= [y_t, u_t, y_{t-1}, u_{t-1}, \ldots, y_0, u_0]'$	all data up to time t, including prior data
φ_{t-1}	$= [d_{t-1}, \ldots, d_{t-\mathring{\varphi}}]'$	the regression vector with past data (at time t)
ψ_t	$= [u_t, \varphi'_{t-1}]'$	the regression vector with actual control (at time t)
Ψ_t	$= [y_t, \psi'_t]' = [y_t, u_t, \varphi'_{t-1}]'$	the extended regression vector (data vector)

5.1.2 Decisions

Individual tasks have their own decisions. e.g. input values for optimal control, point estimates of parameters for the task of parameter estimation, optimal output value for the prediction etc. In summary, to be or not to be a decision is an integral part of the problem formulation not an absolute property.

5.1.3 Parameterized model

The scalar autoregressive model with external normal noise (ARX model) is supposed here

$$y_t = b_0 u_t + \sum_{i=1}^{n} (a_i y_{t-i} + b_i u_{t-i}) + c + e_t = \psi'_t \theta + e_t, \tag{5.1}$$

where
$\{b_0, a_i, b_i, c\}, i = 1, \ldots, n$ are unknown regression coefficients
e_t is discrete white normal noise with constant variance r
$\theta = [b_0, a_1, b_1, \ldots, a_n, b_n, c]'$
$n = \mathring{\varphi}$ is the system order.

Equivalently, expressed by the pdf,

$$f(\Delta_t|\Theta, \mathcal{P}_{a_t^*}, a_t) \equiv f(y_t|\Theta, \psi_t) = \tag{5.2}$$

$$= (2\pi r)^{-0.5} \exp\left\{-\frac{1}{2r}\left(y_t - \theta'\psi_t\right)^2\right\} =$$

$$= (2\pi r)^{-0.5} \exp\left\{-\frac{1}{2r}\mathrm{tr}(\Psi_t\Psi_t'[-1,\ \theta']'[-1,\ \theta'])\right\}$$

where
$\Theta = \{\theta, r\}$ are unknown model parameters
tr denotes matrix trace.

This form corresponds to the general form of a model from the exponential family (2.53) with

$$A(\Theta) = (2\pi r)^{-0.5}$$

$$B(\Psi) = \Psi_t\Psi_t'$$

$$C(\Theta) = -\frac{1}{2r}[-1,\ \theta']'[-1,\ \theta']$$

5.1.4 Prior pdf

The parameterized model belongs to the exponential family, see Agreement 2.14. Thus the conjugate prior pdf (Agreement 3.1) has the form

$$f(\Theta) = (2\pi r)^{-0.5\nu_0} \exp\left\{-\frac{1}{2r}[-1,\ \theta']\, V_0\, [-1,\ \theta']'\right\} \chi(\Theta) = \tag{5.3}$$

$$= (2\pi r)^{-0.5\nu_0} \exp \mathrm{tr}\left\{-\frac{1}{2r}[-1,\ \theta']'[-1,\ \theta']V_0\ \right\} \chi(r)$$

where
ν_0, V_0 are prior statistics; ν_0 is scalar V_0 is a $(\mathring{\Psi}, \mathring{\Psi})$-matrix
$\chi(\Theta) = \chi(r)$ is characteristic function and it holds

$$\begin{aligned} \chi(r) &= 1 \quad \text{for } r > 0 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

5.1.5 Loss function

The quadratic loss function (3.19) is used. For example in the case of the control task it means

$$\mathcal{Z} = \sum_{\tau=1}^{T} [(y_\tau - w_\tau)^2 + \omega u_\tau^2] \tag{5.4}$$

where

$T > 1$	is time horizon of the design
w_τ	is user specified required value of the output
$\omega > 0$	is penalization constant.

5.2 Derived elements

5.2.1 Likelihood

It is a product of models with data substituted, see definition (3.20). Here, it has the form

$$\mathcal{L}(\Theta, \mathcal{P}_{a^*_{t+1}}) = \mathcal{L}(\Theta, d(t)) = \prod_{\tau=1}^{t} f(y_\tau|\psi_\tau, \Theta).$$

and for model (5.2) it holds

Proposition 5.1 (Likelihood for general ARX model) *For general ARX model (5.2) the likelihood function is*

$$\mathcal{L}(\Theta, d(t)) = (2\pi r)^{-t/2} \exp\ \mathrm{tr}\left\{-\frac{1}{2r}\sum_{\tau=1}^{t} \Psi'_\tau \Psi_\tau [-1,\ \theta']'[-1,\ \theta']\right\} \tag{5.5}$$

Proof: It follows directly from the likelihood definition. ♢

Example:
The first order system described by $y_\tau = b_0 u_\tau + a_1 y_{\tau-1} + e_\tau$ is considered. It means that $\theta = [b_0, a_1]$, $\Psi = [y_\tau, u_\tau, y_{\tau-1}]$. The likelihood for this model is

$$\mathcal{L}(\Theta, d(t)) = (2\pi r)^{-t/2} \exp\ \{-\frac{1}{2r}\sum_{\tau=1}^{t}(y_\tau^2 + b_0^2 u_\tau^2 + a_1^2 y_{\tau-1}^2 - 2b_0 y_\tau u_\tau + 2b_0 a_1 u_\tau y_{\tau-1} - 2a_1 y_\tau y_{\tau-1})\}$$

Expression in the exponent can be rearranged as follows

$$y_\tau^2 - 2y_\tau(b_0 u_\tau + a_1 y_{\tau-1}) + (b_0 u_\tau + a_1 y_{\tau-1})^2 = (y_\tau - \hat{y}_\tau)^2$$

where
$\hat{y}_\tau = b_0 u_\tau + a_1 y_{\tau-1}$ is the conditional mean of y_τ and the likelihood is then

$$\mathcal{L}(\Theta, d(t)) = (2\pi r)^{-t/2} \exp\left\{-\frac{1}{2r}\sum_{\tau=1}^{t}(y_\tau - \hat{y}_\tau)^2\right\}$$

5.2.2 Integral

The particular form of (3.21) for this case is

$$\mathcal{I}(\mathcal{P}_{a^*_{t+1}}) = \mathcal{I}(d(t))$$

and it holds

Proposition 5.2 (Normalization integral $\mathcal{I}$ for the ARX model) *For likelihood function (5.5) and prior pdf (5.3) the integral $\mathcal{I}$ (3.21) is*

$$\mathcal{I}(d(t)) = \int (2\pi r)^{-\nu_t/2} \exp\left[-\frac{1}{2r} tr\{[-1,\ \theta]'[-1,\ \theta]V_t\}\right] \chi(r>0)d\Theta \tag{5.6}$$

where the statistics ν_t, V_t are updated according to (5.9)

Proof: Simple substitution of (5.5) and (5.3) into the definition (3.21). ♢

After integration the following form is obtained

$$\mathcal{I}(d(t)) = 2\pi^{-0.5(\nu_t-\mathring{\psi})} \frac{\Gamma(\frac{\nu_t-\mathring{\psi}+2}{2})}{\lambda_t^{0.5(\nu_t-\mathring{\psi}+2)}|V_{\psi;t}|^{0.5}} \tag{5.7}$$

where
$V_{\psi;t}$ is part of the statistic V_t according to (5.11)
ν_t is statistic updated by (5.9)
$\mathring{\psi}$ is a length of the regression vector ψ
λ_t is defined in (5.13)
Γ is Gamma function *Proof:* In [10] ♢

5.2.3 Posterior pdf

The fact that our model belongs to the exponential family (2.53) allows us to apply directly Proposition 2.16. The posterior pdf keeps the form

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) = f(\Theta|d(t))$$

and it holds

Proposition 5.3 (Posterior pdf for the normal ARX model and its evolution) *Let the system be described by the parameterized model (5.1), natural conditions of decision making (2.36) hold and a conjugate prior pdf (5.3) be used. Then, for $\Theta \in \Theta^*$ it holds*

$$f(\Theta|d(t)) \propto \mathcal{L}(\Theta, d(t))f(\Theta) = (2\pi r)^{-\nu_t/2} \exp\left[-\frac{1}{2r} tr\{V_t\,[-1,\ \theta']'[-1,\ \theta']\}\right] \chi(r>0) \tag{5.8}$$

where
ν_t, V_t are sufficient statistics, ν_t is scalar, V_t is a $(\mathring{\Psi}, \mathring{\Psi})$-matrix
the normalizing factor is the integral (5.6).

The form (5.8) is self-reproducing with two parameters ν_t and V_t, which can be updated according to (5.9)

$$\begin{array}{rclcl} V_t & = & V_{t-1} + \Psi_t \Psi_t' & = & V_0 + \sum_{\tau=1}^{t} \Psi_t \Psi_t' \\ \nu_t & = & \nu_{t-1} + 1 & = & \nu_0 + t \end{array} \tag{5.9}$$

The recursive form of the posterior pdf is

$$f(\Theta|d(t)) \propto f(y_t|\Theta, u_t, d(t-1)) = (2\pi r)^{-0.5} \exp\left\{-\frac{1}{2r} tr(\Psi_t \Psi_t'[-1,\ \theta']'[-1,\ \theta'])\right\} f(\Theta|d(t-1)), \tag{5.10}$$

starting with the prior pdf $f(\theta)$ (5.3) where in the role of the normalization integral appears the predictive pdf (5.15).

If the matrix V_t is partitioned in the following way

$$V_t = \begin{bmatrix} V_{y;t} & V_{y,\psi;t}' \\ V_{y,\psi;t} & V_{\psi;t} \end{bmatrix} \tag{5.11}$$

where $V_{y;t}$ is scalar, $V_{y,\psi;t}$ of dimension $(\mathring{\Psi}, 1)$, $V_{\psi;t}$ of dimension $(\mathring{\Psi}, \mathring{\Psi})$, then the posterior pdf (5.8) can be given in the form

$$f(\Theta|d(t)) \propto (2\pi r)^{-\nu_t/2} \exp\left[-\frac{1}{2r}[\theta - \hat{\theta}_t]' V_{\psi;t} [\theta - \hat{\theta}_t] + \lambda_t\right] \chi(r) \tag{5.12}$$

where

$$\begin{array}{rcl} \hat{\theta}_t & = & C_t V_{y,\psi;t} \\ C_t & = & V_{\psi;t}^{-1} \\ \lambda_t & = & V_{y;t} - V_{y,\psi;t}' C_t V_{y,\psi;t} \end{array} \tag{5.13}$$

The maximum of the posterior pdf (5.12) lies in the point $\theta = \hat{\theta}_t$, $r = \nu_t^{-1}\lambda_t$. *Proof:* The form (5.12) imply from the completion of squares, proof of the maximum is in [10] ♢

The characteristics $\hat{\theta}_t$, C_t, λ_t can be updated recursively instead of calculating them for each t according to formulae (5.13).

$$\begin{array}{rcl} \hat{\theta}_t & = & \hat{\theta}_{t-1} + \dfrac{1}{1+\zeta_t} C_{t-1} \psi_t \hat{e}_t' \\ \lambda_t & = & \lambda_{t-1} + \dfrac{1}{1+\zeta_t} \hat{e}_t \hat{e}_t' \\ C_t & = & C_{t-1} - \dfrac{1}{1+\zeta_t} C_{t-1} \psi_t \psi_t' C_{t-1} \end{array} \tag{5.14}$$

where
$\zeta_t = \psi_t' C_{t-1} \psi_t$
$\hat{e}_t = y_t - \hat{\theta}_{t-1}' \psi_t$ is prediction error. *Proof:* Evolution of the sub-matrices in V_t and substitution 5.13. For details see [10] ♢

The matrix V_t or equivalently the triad $\{\hat{\theta}_t,\ C_t,\ \lambda_t\}$ is the sufficient statistic for estimating the unknown parameters.

5.2.4 Predictive pdf

The general formula (3.24) has the form

$$f(\Delta_{t+1}|a_{t+1},\mathcal{P}_{a_{t+1}}) = f(y_{t+1}|u_{t+1},d(t)) = \frac{\mathcal{I}(d(t+1))}{\mathcal{I}(d(t))} = \frac{\mathcal{I}(y_{t+1},u_{t+1},d(t))}{\mathcal{I}(d(t))}$$

and it holds

Proposition 5.4 (Predictive pdf for the normal ARX model) *The predictive pdf is given by the general formula (3.24) with the integrals (5.6) and statistics V_t, ν_t defined by (5.9) and partitioned according to (5.11)*

$$f(y_{t+1}|u_{t+1},d(t)) = \frac{\Gamma(\frac{\nu_t-\mathring{\psi}+3}{2})}{\Gamma(\frac{\nu_t-\mathring{\psi}+2}{2})}[\pi(1+\zeta_{t+1})\lambda_t]^{-0.5}\left(1+\frac{\hat{e}'_{t+1}\lambda_t^{-1}\hat{e}_{t+1}}{1+\zeta_{t+1}}\right)^{-0.5(\nu_t-\mathring{\psi}+3)} \tag{5.15}$$

where

$$\begin{aligned}\hat{e}_{t+1} &= y_{t+1}-\hat{\theta}'_t\psi_{t+1}\\ \lambda_t &= V_{y;t}-V'_{y,\psi;t}V^{-1}_{\psi;t}V_{y,\psi;t}\\ \zeta_{t+1} &= \psi'_{t+1}V^{-1}_{\psi;t}\psi_{t+1}\end{aligned}$$

For details see (5.13), (5.14).

5.2.5 Multi-step predictive pdf

For predictions over several steps ahead without current data measurement we need the following multi-step predictive pdf.

Proposition 5.5 (Multi-step predictive pdf for normal ARX model) *For the ARX normal model (5.2) (for which the one-step ahead predictive pdf is given by (5.15)) and for given model $f(u_{t+1}|d(t))$ for evolvement of the input variable u, the k-step ahead predictive pdf can be expressed in the following form*

$$f(y_{t+k}|u_{t+1},d(t)) = \tag{5.16}$$

$$= \int_D \left[f(y_{t+1}|u_{t+1},d(t))\prod_{j=2}^{k} f(y_{t+j}|u_{t+j},d(t+j-1))f(u_{t+j}|d(t+j-1))\right]dD,$$

for $y_{t+k}\in R$, given data $[u_{t+1},\ d(t)]$ and the denotation $D=\{u_{t+k},y_{t+k-1},u_{t+k-1},\dots,u_{t+2},y_{t+1}\}$.

Proof: The result follows immediately if the missing variables u_{t+k}, y_{t+k-1}, u_{t+k-1} $\dots$ y_{t+1} are substituted into the predictive pdf and at the same time they are integrated out. By application of the chain rule to the joint pdf, we get the result. ♢

In the case of the advance inputs and the ARX model given by (5.2) the multi-step predictive pdf is

$$f(y_{t+k}|u_{t+k},u_{t+k-1},\dots,u_{t+1},d(t)) = \tag{5.17}$$

$$= \int_{D_1}[f(y_{t+k},y_{t+k-1},\dots,y_{t+1},\Theta|u_{t+k},u_{t+k-1},\dots,u_{t+1},d(t))]\,dD_1 =$$

$$= \int_{D_1}[f(y_{t+k}|\psi_{t+k},\Theta)f(y_{t+k-1}|\psi_{t+k-1},\Theta)\dots f(y_{t+1}|\psi_{t+1},\Theta)f(\Theta|d(t))]\,dD_1,$$

for $y_{t+k} \in R$, given data $[u_{t+k}, \ldots, u_{t+1},\ d(t)]$ and the denotation $D_1 = \{y_{t+k-1}, \ldots, y_{t+1}, \Theta\}$,
where
$f(y_{t+i}|\psi_{t+i}, \Theta), i = 1, \ldots, k$ are known model pdf's
$f(\Theta|d(t))$ is the estimated pdf of the parameters

Proof: The chain rule (??) and the natural condition of control (??) are applied ♢

5.3 Static tasks

The basic static tasks related to the normal ARX model will be elaborated in this section.

5.3.1 Point estimation

The point estimate is the mean value of the posterior pdf $f(\Theta|d(t))$ (5.8). The estimation within the exponential family 2.16 is simple. The parameter estimates are obtained from the sufficient statistics V_t, ν_t updated according to (5.9) as follows

Proposition 5.6 (Point estimation for the normal ARX model) *For the ARX model (5.2) with the statistics V_t, ν_t defined by (5.9) and V_t partitioned according to (5.11) it holds*

$$\hat{\theta} = V_{\psi;t}^{-1} V_{y,\psi;t} \tag{5.18}$$

$$\hat{r} = \nu_t \lambda_t^{-1} \tag{5.19}$$

where
$\hat{\theta}$, $\hat{r}$ are point parameter estimates
λ_t is computed according to (5.13)

Proof: In [10]. ♢

5.3.2 One-step-ahead prediction

This task is generally formulated in (3.34) and the explicit solution for quadratic criterion is given in (3.35).

In this case

$$\hat{\Delta}(D) = \hat{y}_{t+1} = E[y_{t+1}|u_{t+1}, d(t)] = \int_{y \in y^*} y\, f(y|u_{t+1}, d(t)) dy$$

The predicted output value y_{t+1} is determined on the basis of the previous data items. The predictive pdf (5.15) is used and its mean value provides the desired result.

Proposition 5.7 (Point one-step-ahead prediction for the normal ARX model) *For the normal ARX model (5.2) with the point parameter estimates according to the Proposition 5.6 it holds*

$$\hat{y}_{t+1} = \psi'_{t+1} \hat{\theta}_t \tag{5.20}$$

where
$\hat{y}_{t+1}$ is the one step output prediction in the time t
$\hat{\theta}_t$ is the point parameter estimate

5.3.3 One-step-ahead control

The general formula for generating the optimal control is (3.3.5) with $\mathcal{Z}$ being the loss function. In terms of our case this formula gets the form

$$u_t(d(t-1)) = \arg\min_{u\in u^*} \int_{y\in y^*} \mathcal{Z}(y,u,d(t-1))f(y|u,\varphi_{t-1})dy.$$

For known model parameters, the predictive pdf $f(y|u,\varphi_{t-1})$ is directly given by the model with parameter values substituted and it holds

Proposition 5.8 (One-step control with known parameters) *Let us consider the ARX normal model (5.2) and the loss function $\mathcal{Z}$ defined in (5.4) with $T=1$ where y is modelled variable, u is control, d is vector of past data. Then, the optimal control law is given by the the solution of the minimization task*

$$\mathcal{V}_{u;t}(d(t-1)) = (2\pi r)^{-0.5}\int_{y\in y^*}[(y_\tau - w_\tau)^2 + \omega u_\tau^2]\exp\left\{-\frac{1}{2r}tr(\Psi_t\Psi_t'[-1,\ \theta]'[-1,\ \theta])\right\}$$
$$\hat{u}_t(d(t-1)) = \arg\min_{u\in u^*}\mathcal{V}_{u;t}(d(t-1))$$

Proof: The result is just the general formula above with the model pdf (5.2) and the loss function (5.4) substituted. ♢

Example:
The first order system (5.1) described by $y_t = b_0u_t + a_1y_{t-1} + b_1u_{t-1} + e_t$ is considered. The loss function $\mathcal{Z}$ (5.4) with control horizon $T=1$ and required output value $w_t = 0$ is chosen,

$$\mathcal{Z} = y_t^2 + \omega u_t^2.$$

The optimal control is then

$$u_{t+1} = -(b_0^2+\omega^2)^{-1}b_0(a_1y_t + b_1u_t)$$

Proof: This example is computed in detail in [67] ♢

5.4 Dynamic tasks

5.4.1 Recursive estimation

In this section, sequential decisions about the model parameter values in dependence on currently measured data is evaluated.

Due to the fact, that the normal ARX model (5.2) belongs to the exponential family (2.53) the task of sequential estimation is simple, as it reduces to algebraic recursion for statistics V_t, ν_t for $t \in t^*$. The recursion for data collection and the form of posterior pdf (5.10) of the estimated parameter Θ is given in the following proposition.

Proposition 5.9 (Sequential estimation for the normal ARX model) *For the normal ARX model (5.2) and given prior pdf (5.3) the posterior pdf is (5.10) with the statistics V_t, ν_t according to (5.9).*

Proof: The result is due to the Proposition 5.3. ♢

5.4.2 Multi-step-ahead prediction

The multi-step output prediction for the normal ARX model is the following.

Proposition 5.10 (Multi-step output prediction for the normal ARX model) *For the ARX model (5.2) with unknown parameter θ the multi-step output prediction is*

$$\hat{y}_{t+k} = \mathcal{E}[y_{t+k}|u_{t+1}, d(t)] = \int_{y\in y^*} y_{t+k}\, f(y_{t+k}|u_{t+1}, d(t)), \tag{5.21}$$

where the multi-step predictive pdf is given in (5.17).

Proof: See (5.5). ♢

The computation of the previous general formula is too complex. For the illustration we will consider simple example of the two steps ahead prediction for auto-regression normal model with known parameters and known advanced inputs.

Example:

The first order system (5.2) described by $y_t = a_1 y_{t-1} + e_t$, $e_t \sim \mathcal{N}(0, \sigma^2)$, with known a_1, σ^2 is considered. The corresponding pdf is

$$f(y_t|y_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(y_t - a_1 y_{t-1})^2\right\} \tag{5.22}$$

The task is to compute the output prediction $\hat{y}_{t+2}$. Predictive pdf is

$$f(y_{t+2}|y(t)) = \int f(y_{t+2}|y_{t+1}) f(y_{t+1}|y_t) dy_{t+1} = \tag{5.23}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \int exp\left\{-\frac{1}{2\sigma^2}(y_{t+2} - a_1 y_{t+1})^2\right\} exp\left\{-\frac{1}{2\sigma^2}(y_{t+1} - a_1 y_t)^2\right\} =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2(a_1^2+1)}} exp\left\{-\frac{1}{2\sigma^2(a_1^2+1)}(y_{t+2} - a_1^2 y_t)^2\right\}$$

This is normal distribution with mean value $a_1^2 y_t$ and variance $\sigma^2(a_1^2+1)$.

The point output prediction $\hat{y}_{t+2}$ is then $\hat{y}_{t+2} = a_1^2 y_t$.

Proof: the result is obtained after completing the argument of the integral to the square in y_{t+1}. After this recomputation we obtain argument of the the integral in the form of the product containing two normal distributions - the first one for the variable y_{t+1} and the second one which don't depend on y_{t+1}. Then the integral over the first one is equal to 1 and the second one is the result of the integration. ♢

Remark:

In the practice the multi-step-ahead prediction can be solved as the sequence of several one-step-ahead predictions (see Proposition 5.7) where the output prediction from the previous step is taken as the real output in the next step. The algorithm for the k-steps-ahead prediction is as follows:

For $i = 1 : k$

1. data $d(t+i-1)$ at disposal
2. generation of u_{t+i}
3. one-step-ahead prediction $\hat{y}_{t+i}$
4. if $i < k$, assignment $y_{t+i} = \hat{y}_{t+i}$

5.4.3 Multi-step-ahead control

This task is generally described in the Section 3.4.4. Here suboptimal strategies of receding horizon (4.4.1) and certainty equivalence strategy (4.46) are used.

The principle of receding horizon is that the control variable u is designed N step ahead - we obtain the time sequence $\{u_{t+N}, \ldots, u_{t+1}\}$. Then the u_{t+1} is applied and y_{t+1} is obtained. The control variable is again computed N steps ahead with time index shifted.

Certainty-equivalence strategy replaces unknown parameter in the parameterized model by a current point estimate $\hat{\Theta}_t$ of unknown parameters.

Example:
The ARX model (5.1) described by $y_t = b_0 u_t + a_1 y_{t-1} + e_t$ is considered. The loss function $\mathcal{Z}$ (5.4) with control horizon $T = N$, required output value $w_t = 0$ and penalization constant $\omega > 0$ is chosen,

$$\mathcal{Z} = \sum_{\tau=1}^{T} [y_\tau^2 + \omega u_\tau^2].$$

The optimal control is then

$$u_t = -r_t y_{t-1}, \qquad t = 1, 2, \ldots, N$$

with following control low for $t = N, N-1, \ldots, 1$ and initial condition $s_{N+1} = 0$

$$\begin{aligned} r_t &= \frac{(1+s_{t+1})a_1 b_0}{(1+s_{t+1})b_0^2 + \omega} \\ s_t &= \frac{a_1 \omega}{b_0} r_t \end{aligned} \tag{5.24}$$

Proof: This example is computed in detail in [?] ♢

Chapter 6

Model with discrete variables

This chapter deals with discrete model. It can be interpreted as controlled tossing of a coin with memory. On a solution of the tasks for this model, majority of aspects of decision making can be clearly illustrated.

6.1 Construction elements

The considered system is interpreted as a tossed coin that can fall with head or tail up. The result of the tossing might be influenced by the way in which coin is laid on the tossing hand (control) and by the side on which it fell last times (memory).

6.1.1 Data

A sequence of outputs y_t, $t = 1, 2, \ldots, \mathring{t}$ is observed. The items of the sequence have finite number of different values. The values can be labelled in an arbitrary way that distinguishes different outcomes. We will consider two different values (evoking connection with a tossed coin) and use their binary representation $y \in y^* \equiv \{\mathbf{0}, \mathbf{1}\}$, where **0** and **1** represent two numbers, mostly 0,1 or 2,1. An extension to more different values of the output is straightforward.

Even in this simple case an simplification is used: cases when the coin disappears during the toss or the coin would finally land on its edge are not considered as possible outcomes of the experiment.

Two valued inputs $u \in u^* \equiv \{\mathbf{0}, \mathbf{1}\}$, the position of the coin on the tossing hand before toss are considered. The restriction to two valued variables, representing a coin, is due to better clarity and with respect to the extension to more than two values it is not crucial.

Notation

In accordance with the general conventions we will use the following notations.

Data item at time t

$$d_t = [y_t, u_t]', \tag{6.1}$$

All data up to time t, including prior data

$$d(t) = [y_t, u_t, y_{t-1}, u_{t-1}, \ldots], \tag{6.2}$$

Regression vector with past data (at time t)

$$\varphi_{t-1} = [y_{t-1}, u_{t-1}, \ldots, y_{t-\mathring{\varphi}}, u_{t-\mathring{\varphi}}, \text{the rest}]', \tag{6.3}$$

where "the rest" stands for the remaining items of φ_{t-1} represented either by older inputs or outputs, in dependence on the individual lengths of memories. The number of common delayed items is denoted by $\mathring{\varphi}$, the total number of delayed outputs is $\mathring{\varphi}_y$ and of inputs $\mathring{\varphi}_u$.

Regression vector with actual control substituted (at time t)

$$\psi_t = [u_t, \varphi'_{t-1}]', \tag{6.4}$$

Extended regression vector

$$\Psi_t = [y_t, \psi'_t]' = [y_t, u_t, \varphi'_{t-1}]'. \tag{6.5}$$

Experimental design

Obviously, if we want to learn properties of our system (tossed coin) we should perform proper experiments. We cannot improve our prior estimate of the probability of $y_t = \mathbf{0}$ conditioned on $u_t = \mathbf{0}$ when we do not choose $u_t = \mathbf{0}$ during our experiment. We stress this simple fact as it is often overlooked in more complex situations.

Data pre-processing

Even in this simple case, some data pre-processing is needed. Typically, the cases a priori excluded by problem formulation but practically appearing, like coin ending on its edge, are excluded from the set or treated as missing data: we hope that the coin will eventually fall on one of the considered sides but we do not know to which one.

6.1.2 Decisions

The considered inputs are typical decisions to be taken in control tasks. As it will be seen below, we shall deal with parameter estimation, output prediction, structure estimation etc. Each of those tasks will have its decision (like point estimates of parameters or decision on continuation of a sequential experiment). In summary, to be or not to be a decision is an integral part of the problem formulation not an absolute property.

Restrictions

Everywhere, we insist on information restriction: our decision can be selected using at most the observed data and prior information.

The fact that two-valued input only is allowed is a technological restriction for the control design. Again, it is our option that can be modified by allowing, for instance, any spatial initial

position of the coin. If this restriction is violated by an undisciplined player we have to take an appropriate decision (get rid of him, stop the game, etc.).

It will be seen in Section 6.4.6 that the complexity restrictions are relevant even in this simple case.

6.1.3 Example

A general case of the discrete dynamic model can sometimes be too complex for a clear presentation. That is why a simper example is appended. It is called **CCM problem** and it represents model of a Controlled Coin with one step Memory.

Agreement 6.1 (CCM problem) *By **CCM problem** we will denote the following task:*

1. *The model (6.9) has two-valued outputs $y_\tau \in \{\boldsymbol{0}, \boldsymbol{1}\}$ and inputs $u_\tau \in \{\boldsymbol{0}, \boldsymbol{1}\}$ for all $\tau = 1, 2, \ldots, \mathring{t}$.*

2. *The regression vector is $\psi_\tau = [u_t, y_{t-1}]' \in \psi^*$, i.e. length of the model memory is one.*

3. *The set of admissible parameters Θ^* is a cartesian product of marginals Θ^*_ψ*

$$\Theta^* = \bigotimes_\psi \Theta^*_\psi, \tag{6.6}$$

with $\Theta^*_\psi \subset \{[\Theta_{\boldsymbol{0}|\psi}, \Theta_{\boldsymbol{1}|\psi}] : \Theta_{\boldsymbol{0}|\psi} > 0, \Theta_{\boldsymbol{1}|\psi} > 0, \Theta_{\boldsymbol{0}|\psi} + \Theta_{\boldsymbol{1}|\psi} = 1\}.$

The marginal regions Θ^*_ψ are one dimensional and after substitution they are represented by intervals $(\theta^l_\psi, \theta^u_\psi) \subset (\boldsymbol{0}, \boldsymbol{1})$. For unrestricted model they are intervals $(0, 1)$ for all $\psi \in \psi^*$.

6.1.4 Parameterized model

Two-valued system output $y \in y^* \equiv \{\mathbf{0}, \mathbf{1}\}$ is observed. The system input u is also two-valued, $u^* = y^*$. Outputs are uncertain (random), they depend on current input and lagged values of output and input. We assume that this dependence

- is incompletely known
- is time-invariant
- has finite memory, i.e. the probability of a particular output y_t is determined by the finite-dimensional regression vector ψ_t.

Thus, the general parameterized model (2.41) has the form

$$f(\Delta_t|\Theta, \mathcal{P}_{a^*_t}, a_t) \equiv f(y_t|\Theta, \psi_t). \tag{6.7}$$

The most general finite-dimensional parameterization takes values of the individual probabilities $f(y_t = y|\Theta, \psi_t = \psi) \equiv \Theta_{y|\psi}$ as unknown parameters

$$\Theta \in \Theta^* \subset \{\Theta_{y|\psi} : \Theta_{y|\psi} \geq 0, \sum_{y \in y^*} \Theta_{y|\psi} = 1\}. \tag{6.8}$$

The form of the widest meaningful set Θ^* of Θ is restricted by the fact that $\Theta_{y|\psi}$ are conditional probabilities.

For a compact and useful description of the general parameterized model (6.7), we use the multivariate version of Kronecker delta (7.1). With it,

DISCRETE MODEL

$$f(y_t|\Theta,\psi_t) = \prod_{\psi\in\psi^*}\prod_{y\in y^*} \Theta_{y|\psi}^{\delta(y,y_t)\delta(\psi,\psi_t)}. \tag{6.9}$$

Remark(s) 6.1

1. *The parameterized model belongs to* ***exponential family*** *(Agreement 2.14), as it can be expressed in the following way*

$$f(y_t|\Theta,\psi_t) = \exp\left[\sum_{\Psi\in\Psi^*}\delta(\Psi,\Psi_t)\ln[\Theta_{y|\psi}]\right] \tag{6.10}$$

with $A(\Theta)=1$, $B'(\Psi_t)$ = *row vector with* $\mathring{\Psi}$*-entries* $\delta(\Psi,\Psi_t)$, $\Psi\in\Psi^*$,
$C(\Theta)$ = *column vector with* $\mathring{\Psi}$*-entries* $\ln[\Theta_{y|\psi}]$, $y\in y^*$, $\psi\in\psi^*$.

2. *The use of Kronecker delta in (6.9) helps us to express the parameterized model in the product form that fits well to the product form of Bayes rule (**??**). It is advantageous to use this "trick" whenever possible.*

3. *Logarithms in (6.10) make sense for positive arguments only. We assume only such data vectors $\Psi\in\Psi^*$ for which it holds. Thus, the considered set Θ^* is a proper subset of the widest possible set (6.8): zero transition probabilities are a priori fixed. This seemingly technical remark opens a way to a more parsimonious description of the system. For instance, if we know from physical modelling that some transitions $\psi\to y$ are impossible we can use it efficiently for fighting with the estimation complexity, see [68, 69, 70].*

4. *It is instructive to try an alternative parameterization using the "completion" Proposition 3.1. It is quite hard to stay within the favorable exponential family. Consider, for instance, the case when expected value of the output is a linear function of the regression vector.*

5. *The general notation we are using makes the results directly applicable to all discretely-controlled finite-memory Markov chains [71, 72, 73] for which output and input sets y^*, u^* have an arbitrary but finite amount of possible values.*

Example:

The model (6.9) used for the CCM problem reduces to

$$f(y_t|\psi_t,\Theta) = \prod_{\psi\in\psi^*}\Theta_{\mathbf{0}|\psi}\Theta_{\mathbf{1}|\psi} = \prod_{\psi\in\psi^*}\theta_\psi(1-\theta_\psi) \tag{6.11}$$

where

$y_\tau\in y^*=\{\mathbf{0},\mathbf{1}\}$ and $u_\tau\in u^*=\{\mathbf{0},\mathbf{1}\}$ for all $t\in t^*$ are inputs and outputs,

$\psi_\tau = [u_t, y_{t-1}]' \in \psi^* = \{[\mathbf{0},\mathbf{0}]', [\mathbf{0},\mathbf{1}]', [\mathbf{1},\mathbf{0}]', [\mathbf{1},\mathbf{1}]'\}$ is regression vector – used as a multi-index for indexing items of parameter and statistics,

$\Theta = \Theta_{y|\psi}$ can be described by a table

u_t, y_{t-1}	$y_t = \mathbf{0}$	$y_t = \mathbf{1}$
0 **0**	$\Theta_{\mathbf{0}|\mathbf{00}}$	$\Theta_{\mathbf{1}|\mathbf{00}}$
0 **1**	$\Theta_{\mathbf{0}|\mathbf{01}}$	$\Theta_{\mathbf{1}|\mathbf{01}}$
1 **0**	$\Theta_{\mathbf{0}|\mathbf{10}}$	$\Theta_{\mathbf{1}|\mathbf{10}}$
1 **1**	$\Theta_{\mathbf{0}|\mathbf{11}}$	$\Theta_{\mathbf{1}|\mathbf{11}}$

$$\Theta_{\mathbf{0}|\psi} > 0$$
$$\Theta_{\mathbf{1}|\psi} > 0$$
$$\Theta_{\mathbf{0}|\psi} + \Theta_{\mathbf{1}|\psi} = 1$$

and after introducing the restrictions

u_t, y_{t-1}	$y_t = \mathbf{0}$	$y_t = \mathbf{1}$
0 **0**	$\theta_{\mathbf{00}}$	$1 - \theta_{\mathbf{00}}$
0 **1**	$\theta_{\mathbf{01}}$	$1 - \theta_{\mathbf{01}}$
1 **0**	$\theta_{\mathbf{10}}$	$1 - \theta_{\mathbf{10}}$
1 **1**	$\theta_{\mathbf{11}}$	$1 - \theta_{\mathbf{11}}$

$$\Theta_{\mathbf{0}|\psi} = \theta_\psi$$
$$\Theta_{\mathbf{1}|\psi} = 1 - \theta_\psi$$
$$\theta_\psi \in (0;1)$$

Remark(s) 6.2

From the table expression of the model parameter it can be clearly seen, that the controlled coin with memory can be interpreted as four plain coins indexed by the multi-index ψ – the regression vector of the model (6.11). The individual plain coin parameters occur at rows of the parameter table.

6.1.5 Prior pdf for discrete model

The parameterized model belongs to exponential family. Thus, according to Proposition 3.3, there is a conjugate prior pdf of the form

PRIOR PDF

$$f(\Theta) \propto \prod_\psi \prod_y \Theta_{y|\psi}^{V_{y|\psi;0}-1} \chi(\Theta), \tag{6.12}$$

where the optional scalars $V_{y|\psi;0}$, $y \in y^*$, $\psi \in \psi^*$ create the optional prior statistics V_0 and $\chi(\Theta)$ denotes the indicator of Θ^* according to (6.8).

The term -1 in exponent of (6.12) is used for notational convenience only. The need to normalize the prior pdf determines the admissible range of the optional values V_0. It is implied by the following auxiliary proposition.

Proposition 6.1 (Existence of the prior pdf) *Let a pdf of the form (6.12) be considered with the indicator $\chi(\Theta)$ that is non-zero on the biggest open subset of the set (6.8) of possible parameters. Then, the function is pdf (integrable) iff the statistics V_0 has positive entries. The corresponding normalizing factor, see Proposition 2.16, has the form*

$$\mathcal{I}(V_0, \nu) \equiv \mathcal{I}(V_0) = \int_{\Theta^*} \prod_\psi \prod_y \Theta_{y|\psi}^{V_{y|\psi;0}-1} \, d\Theta = \mathcal{B}(V_0), \tag{6.13}$$

where $\mathcal{B}(V_0)$ is a multivariate beta function (7.10) which for two-valued output equals to a product of normal beta functions.

Remark(s) 6.3

1. *The used conjugate prior pdf (6.12) is known as Dirichlet prior pdf.*

2. *It is known that* $\Gamma(x) = (x-1)!$ *for integer arguments. This indicate that a logarithmic version has to be used whenever numerically evaluated for a bit larger* x*.*

3. *A more restrictive definition of* Θ^* *leads mostly to incomplete gamma and beta functions and thus makes evaluation much harder. For this reason, it is useful to rely more on soft prior restrictions through an appropriate choice of the statistics* V_0*.*

4. *The counter* ν *occurring in the general form of a conjugate prior pdf, see 3.1.4, is not explicitly present here as* $A(\Theta) = 1$*. The sum of entries of* V_0 *is, however, such a (possibly shifted) counter.*

Example:

For the **CCM problem** the prior pdf according to (6.12)

has the form

$$f(\Theta) \propto \prod_{\psi} \Theta_{\mathbf{0}|\psi}^{V_{\mathbf{0}|\psi;0}-1} \Theta_{\mathbf{1}|\psi}^{V_{\mathbf{1}|\psi;0}-1} \chi(\Theta), \tag{6.14}$$

with the prior statistics

$$V_0 = V_{y|\psi;0}: \quad \begin{array}{|cc||c|c|} \hline u_t & y_{t-1} & y_t = \mathbf{0} & y_t = \mathbf{1} \\ \hline\hline \mathbf{0} & \mathbf{0} & V_{\mathbf{0}|\mathbf{00}} & V_{\mathbf{1}|\mathbf{00}} \\ \mathbf{0} & \mathbf{1} & V_{\mathbf{0}|\mathbf{01}} & V_{\mathbf{1}|\mathbf{01}} \\ \mathbf{1} & \mathbf{0} & V_{\mathbf{0}|\mathbf{10}} & V_{\mathbf{1}|\mathbf{10}} \\ \mathbf{1} & \mathbf{1} & V_{\mathbf{0}|\mathbf{11}} & V_{\mathbf{1}|\mathbf{11}} \\ \hline \end{array} \tag{6.15}$$

and indicator function $\chi(\Theta)$ expressing the conditions (6.6).

6.1.6 Loss function

One of the advantages of the discrete case is that there is a finite number of elements and they can be set value-vise. This advantage can be utilized also for construction of the loss function which can be defined as a penalization table

$$\mathcal{Z}(y, u, d) = \omega_{y,u,d}, \tag{6.16}$$

where y is modelled variable, u is control, d is vector of past data and values of $\omega_{y,u,d}$ represent penalties for individual value configurations of y, u, d.

A single common warning should be made: numeric values of inputs and outputs are arbitrary and have to be interpreted as labels only.

Example:

For the **CCM problem** the loss function according to (6.16) can be written as a table

$$\omega_{y|\psi}: \quad \begin{array}{|cc||c|c|}\hline u_t & y_{t-1} & y_t=\mathbf{0} & y_t=\mathbf{1} \\ \hline\hline \mathbf{0} & \mathbf{0} & \omega_{\mathbf{0}|\mathbf{00}} & \omega_{\mathbf{1}|\mathbf{00}} \\ \mathbf{0} & \mathbf{1} & \omega_{\mathbf{0}|\mathbf{01}} & \omega_{\mathbf{1}|\mathbf{01}} \\ \mathbf{1} & \mathbf{0} & \omega_{\mathbf{0}|\mathbf{10}} & \omega_{\mathbf{1}|\mathbf{10}} \\ \mathbf{1} & \mathbf{1} & \omega_{\mathbf{0}|\mathbf{11}} & \omega_{\mathbf{1}|\mathbf{11}} \\ \hline \end{array} \tag{6.17}$$

6.2 Derived elements

6.2.1 Likelihood function

It is a product of models with data substituted, see (3.20). Here, it has the form

$$\mathcal{L}(\Theta, \mathcal{P}_{a^*_{t+1}}) = \mathcal{L}(\Theta, d(t)) = \prod_{\tau=1}^{t} f(y_\tau|\Theta, \psi_\tau).$$

and for model (6.9) it holds

Proposition 6.2 (Likelihood for discrete model) *For general discrete model (6.9) the likelihood function is*

LIKELIHOOD

$$\mathcal{L}(\Theta, d(t)) = \prod_{\psi}\prod_{y} \Theta_{y|\psi}^{\tilde{V}_{y|\psi;t}}, \tag{6.18}$$

Proof: $\prod_{\tau=1}^{t}\prod_{\psi}\prod_{y}\Theta_{y|\psi}^{\delta(y,y_\tau)\delta(\psi,\psi_\tau)} = \prod_{\psi}\prod_{y}\Theta_{y|\psi}^{\sum_{\tau=1}^{t}\delta(y,y_\tau)\delta(\psi,\psi_\tau)}$ and $V_{y|\psi} = \sum_{\tau=1}^{t}\delta(y,y_\tau)\delta(\psi,\psi_\tau)$. ♢

Example:

For the **CCM problem** (Agreement 6.1) the likelihood function has a form

$$\mathcal{L}(\Theta, d(t)) = \prod_{\psi} \Theta_{\mathbf{0}|\psi}^{\tilde{V}_{\mathbf{0}|\psi;t}} \Theta_{\mathbf{1}|\psi}^{\tilde{V}_{\mathbf{1}|\psi;t}}, \tag{6.19}$$

for $\psi \in \{[\mathbf{0},\mathbf{0}]', [\mathbf{0},\mathbf{1}]', [\mathbf{1},\mathbf{0}]', [\mathbf{1},\mathbf{1}]'\}$

6.2.2 The integral $\mathcal{I}$

The particular form of (2.46) for this case is

$$\mathcal{I}(\mathcal{P}_{a^*_{t+1}}) = \mathcal{I}(d(t))$$

and it holds

Proposition 6.3 (Normalization integral $\mathcal{I}$ for discrete model) *For likelihood function (6.18) and prior pdf (6.12) the integral $\mathcal{I}$ (2.46) is*

INTEGRAL

$$\mathcal{I}(d(t)) = \int_{\Theta^*} \prod_{\psi} \prod_{y} \Theta_{y|\psi}^{V_{y|\psi;t}-1} d\Theta = \mathcal{B}(V_t) \tag{6.20}$$

where

$V_t = \tilde{V}_t + V_0$ *according to (6.18) and (6.12) and*
beta function $\mathcal{B} = \prod_{\psi\in\psi^*} \mathcal{B}_\psi$ *with* $\mathcal{B}_\psi = \prod_{y\in y^*} \Gamma(V_{y|\psi})/\Gamma(\sum_{y\in y^*} V_{y|\psi})$ *is defined in (7.10).*

Proof: Simple substitution of (6.18) and (6.12) into the definition (3.21) and using definition of beta function (7.10). ♢

Example:

For the **CCM problem** (Agreement 6.1) the integral is

$$\mathcal{I}(d(t)) = \int_{\Theta^*} \prod_{\psi} \Theta_{\mathbf{0}|\psi}^{V_{\mathbf{0}|\psi;t}-1} \Theta_{\mathbf{1}|\psi}^{V_{\mathbf{1}|\psi;t}-1} d\Theta = \prod_{\psi} \mathcal{B}([V_{\mathbf{0},\psi;t}, V_{\mathbf{1},\psi;t}]) \tag{6.21}$$

where $V_{\psi;t} = [V_{\mathbf{0},\psi;t}, V_{\mathbf{1},\psi;t}]$ is partial statistics, concerning particular choice of the regression vector ψ.

6.2.3 Posterior pdf

The fact that our model belongs to the exponential family (2.53) allows us to apply directly Proposition 2.16. The posterior pdf (3.23) now reads

$$f(\Theta|\mathcal{P}_{a^*_{t+1}}) = f(\Theta|d(t))$$

and it holds

Proposition 6.4 (Posterior pdf for discrete model and its evolution) *Let the system be described by the parameterized model (6.9), natural conditions of decision making (2.36) hold and a conjugate Dirichlet prior pdf (6.12) with $V_0 > 0$ (entry-wise) be used. Then, the posterior pdf (Proposition 2.14) has also Dirichlet form and for $\Theta \in \Theta^*$ it holds*

POSTERIOR PDF

$$f(\Theta|d(t)) \propto \prod_{\psi} \prod_{y} \Theta_{y|\psi}^{V_{y|\psi;t}-1}, \tag{6.22}$$

with the sufficient statistics $V_t = [V_{\mathbf{1}|\psi;t},\ V_{\mathbf{0}|\psi;t}]$ (6.20) whose entries evolve in time according to the recursion

$$V_{y|\psi;t} = V_{y|\psi;t-1} + \delta(\Psi, \Psi_t),\ \ \Psi \in \Psi^* \tag{6.23}$$

starting with V_0.

The normalizing factor of this pdf is $\mathcal{I}(V_t) \equiv \mathcal{B}(V_t) = \prod_\psi \mathcal{B}(V_{\mathbf{1}|\psi;t}, V_{0|\psi;t})$, according to (6.20).

Proof: According to (2.44) the posterior pdf is $f(\Theta|d(t)) \propto \prod_{t=1}^{\mathring{t}} f(y_t|u_t, \varphi_{t-1}) f(\Theta)$. Substituting (6.9) for the model pdfs and (6.12) for the prior pdf and dragging the product sign into the exponent as a sum sign, we directly obtain the form of the posterior as well as the recursion for the statistics evolution. ♢

Example:

For the **CCM problem** (Agreement 6.1) the posterior pdf has the form

$$f(\Theta|d(t)) \propto \prod_{\psi} \Theta_{\mathbf{0}|\psi}^{V_{\mathbf{0}|\psi;t}-1} \Theta_{\mathbf{1}|\psi}^{V_{\mathbf{1}|\psi;t}-1}, \tag{6.24}$$

with the following update of the statistics

$$\begin{aligned} V_{\mathbf{0}|\psi;t} &= V_{\mathbf{0}|\psi;t-1} + \delta(\mathbf{0}, y_t)\delta(\psi, \psi_t) \\ V_{\mathbf{1}|\psi;t} &= V_{\mathbf{1}|\psi;t-1} + \delta(\mathbf{1}, y_t)\delta(\psi, \psi_t) \end{aligned}$$

for $\psi \in \psi^*$ and $t = 1, 2, \ldots, \mathring{t}$.

6.2.4 Predictive pdf

The general formula (3.24) has the form

$$f(\Delta_{t+1}|a_{t+1}, \mathcal{P}_{a_{t+1}}) = f(y_{t+1}|u_{t+1}, d(t))$$

and it holds

Proposition 6.5 (Predictive pdf for discrete model) *The predictive pdf is given by the general formula (3.24) with the integrals (6.20)*

PREDICTIVE PDF

$$f(y_{t+1}|u_{t+1}, d(t)) = \frac{\mathcal{B}(V_{t+1})}{\mathcal{B}(V_t)} = \frac{V_{y_{t+1}|\psi_{t+1};t}}{\sum_{y \in y^*} V_{y|\psi_{t+1};t}}, \tag{6.25}$$

with ψ_{t+1} according to (6.4).

Proof: The first equality follows directly from (3.24) expressing the predictive pdf as a ratio of two integrals and (3.21) where, for the discrete model, the integral is shown to be a multivariate beta function.

Now, as for the second equality. In the appendix, it is shown (7.10) that the multivariate beta function has the form $B(V)$ can be expressed as a product of "partial" beta functions indexed by various regression vectors, i.e. $B(V) = \prod_{\psi \in \psi^*} B(V_\psi)$, where $V_\psi = [V_{\tilde{y}_1|\psi}, V_{\tilde{y}_2|\psi}, \ldots, V_{\tilde{y}_{\mathring{y}}|\psi}]$ is "partial" statistics (by $\tilde{y}$ we denoted values of y). The evolution of the statistics is given by (6.23) which can be written in the form $V_{y|\psi;t+1} = V_{y|\psi;t} + \delta(y, y_{t+1})\delta(\psi, \psi_{t+1})$. From this, it can be clearly seen, that only the partial statistics concerning the regression vector $\psi = \psi_{t+1}$ and thus only the corresponding partial beta function is changed. From it follows, that in the ratio of beta functions all terms are cancelled up to those indexed by $\psi = \psi_{t+1}$. It is $\mathcal{B}(V_{t+1})/\mathcal{B}(V_t) = \mathcal{B}(V_{\psi_{t+1};t+1})/\mathcal{B}(V_{\psi_{t+1};t})$,

where, according to (7.10) the partial beta function is $B(V_\psi) = \prod_{y\in y^*}\Gamma(V_{y|\psi})/\Gamma(\sum_{y\in y^*}V_{y|\psi})$. And again, the update of the partial statistics concerns only the term indexed by $y = y_{t+1}$, it is $V_{y_{t+1}|\psi_{t+1};t+1} = V_{y_{t+1}|\psi_{t+1};t}+\delta(y,y_t)$ the rest remaining unchanged. Thus the ratio of beta functions has the final form, (see also (7.4))

$$\frac{\frac{\Gamma(V_{y_{t+1}|\psi_{t+1};t}+1)}{\Gamma(\sum_{y\in y^*}V_{y|\psi_{t+1};t}+1)}}{\frac{\Gamma(V_{y_{t+1}|\psi_{t+1};t})}{\Gamma(\sum_{y\in y^*}V_{y|\psi_{t+1};t})}} = \frac{V_{y_{t+1}|\psi_{t+1};t}}{\sum_{y\in y^*}V_{y|\psi_{t+1};t}}.$$

This was to be proved. ♢

Example:

For the **CCM problem** (Agreement 6.1), the beta function is (**??**) and the predictive pdf gets a simple form

$$f(y_{t+1}|u_{t+1},d(t)) = \begin{cases} \frac{V_{\mathbf{1}|\psi;t}}{V_{\mathbf{1}|\psi;t}+V_{\mathbf{0}|\psi;t}} & \text{for } y_{t+1}=\mathbf{1}, \\ \frac{V_{\mathbf{0}|\psi;t}}{V_{\mathbf{1}|\psi;t}+V_{\mathbf{0}|\psi;t}} & \text{for } y_{t+1}=\mathbf{0}, \end{cases} \tag{6.26}$$

where $\psi = \psi_{t+1} = [u_{t+1}, y_t]$.

Proof: The result follows directly from the previous general formula. ♢

6.2.5 Multi-step predictive pdf

For predictions over several steps ahead without current data measurement we need the following multi-step predictive pf.

Proposition 6.6 (Multi-step predictive pdf for discrete model) *For the discrete model (6.9) (for which the one-step ahead predictive pf is given by (6.25)) and for given model $f(u_{t+1}|d(t))$ for evolvement of the input variable u, the k-step ahead predictive pdf can be expressed in the following form*

$$f(y_{t+k}|u_{t+1},d(t)) = \tag{6.27}$$
$$\sum_D \left[f(y_{t+1}|u_{t+1},d(t)) \prod_{j=2}^{k} f(y_{t+j}|u_{t+j},d(t+j-1)) f(u_{t+j}|d(t+j-1)) \right],$$

for $y_{t+k} = 1, 2$ and given data u_{t+1}, $d(t)$ and $D = \{u_{t+k}, y_{t+k-1}, u_{t+k-1}, \ldots, u_{t+2}, y_{t+1}\}$.

Proof: The result follows immediately if the missing variables u_{t+k}, y_{t+k-1}, u_{t+k-1} $\ldots$ y_{t+1} are substituted into the predictive pf and at the same time they are summed out. Application of the chain rule to the joint pf, we got, gives the result. ♢

6.3 Static tasks

Here, we elaborate common static tasks related to coin tossing.

6.3.1 Point estimation

The usage of Kronecker delta in (6.10) implies that $V_{y|\psi;t}$ is the number of observed configurations $\Psi = \Psi_\tau,\ \tau \leq t$ increased by $V_{y|\psi;0}$. From it follows

Proposition 6.7 (Moments of the Dirichlet distribution) *Let $J_{y|\psi},\ y \in y^*,\ \psi \in \psi^*$, be an integer multi-index with $J_{y|\psi} \in \{0, 1, 2, \ldots\}$. Then, Jth moment of the Dirichlet distribution is*

POINT ESTIMATES

$$\mathcal{E}\left[\prod_{\psi\in\psi^*}\prod_{y\in y^*}\left(\Theta_{y|\psi}\right)^{J_{y|\psi}}\middle|\,\mathcal{P}_{a^*_{t+1}}\right] = \frac{\mathcal{B}(V_t + J)}{\mathcal{B}(V_t)}. \tag{6.28}$$

For the first moment and second central moments it follows:

$$\mathcal{E}[\Theta_{\tilde{y}|\tilde{\psi}}|\mathcal{P}_{a^*_{t+1}}] = \frac{V_{\tilde{y}|\tilde{\psi}}}{\sum_{y\in y^*} V_{y|\tilde{\psi}}} \tag{6.29}$$

$$\operatorname{cov}[\Theta_{\tilde{y}|\tilde{\psi}}, \Theta_{\hat{y}|\hat{\psi}}|\mathcal{P}_{a^*_{t+1}}] =$$

$$= \begin{cases} 0 & \text{for any } \tilde{y},\ \hat{y} \quad \text{and} \quad \tilde{\psi} \neq \hat{\psi} \\ \mathcal{E}[\Theta_{\tilde{y}|\psi}|\mathcal{P}_{a^*_{t+1}}]\mathcal{E}[\Theta_{\hat{y}|\psi}|\mathcal{P}_{a^*_{t+1}}]\left[\frac{\sum_{y\in y^*} V_{y|\psi}}{\sum_{y\in y^*} V_{y|\psi}+1} - 1\right] & \text{for } \tilde{y} \neq \hat{y} \quad \text{and} \quad \tilde{\psi} = \hat{\psi} = \psi \\ \left(\mathcal{E}[\Theta_{\tilde{y}|\psi}|\mathcal{P}_{a^*_{t+1}}]\right)^2\left[\frac{\sum_{y\in y^*} V_{\tilde{y};t}}{\sum_{y\in y^*}^{\tilde{y}} V_{\tilde{y};t}+1}\ \frac{V_{\tilde{y};t}+1}{V_{\tilde{y};t}} - 1\right] & \text{for } \tilde{y} = \hat{y} \quad \text{and} \quad \tilde{\psi} = \hat{\psi} = \psi \end{cases} \tag{6.30}$$

Proof: The mean value in (6.28) can be expressed in the integral form with posterior (6.22) as follows

$$\mathcal{E}\left[\prod_{\psi\in\psi^*}\prod_{y\in y^*}\left(\Theta_{y|\psi}\right)^{J_{y|\psi}}\middle|\,\mathcal{P}_{a^*_{t+1}}\right] = \frac{1}{\mathcal{B}(V_t)}\int_{\Theta^*}\Theta_{y|\psi}^{J_{y|\psi}}\prod_{\psi\in\psi^*}\prod_{y\in y^*}\Theta_{y|\psi}^{V_{y|\psi;t}-1}d\Theta =$$

$$\frac{1}{\mathcal{B}(V_t)}\int_{\Theta^*}\prod_{\psi\in\psi^*}\prod_{y\in y^*}\Theta_{y|\psi}^{V_{y|\psi;t}+J_{y|\psi}-1}d\Theta = \frac{\mathcal{B}(V_t + J)}{\mathcal{B}(V_t)}$$

The last step is due to the definition of multivariate beta function (7.10).

First moments:

$$\mathcal{E}[\Theta_{\tilde{y}|\tilde{\psi}}|\mathcal{P}_{a^*_{t+1}}] = \frac{1}{\mathcal{B}(V_t)}\int_{\Theta^*}\Theta_{\tilde{y}|\tilde{\psi}}\prod_{\psi\in\psi^*}\prod_{y\in y^*}\Theta_{y|\psi}^{V_{y|\psi;t}-1}d\Theta =$$

$$= \frac{1}{\mathcal{B}(V_t)}\int_{\Theta^*}\prod_{\psi\in\psi^*}\prod_{y\in y^*}\Theta_{y|\psi}^{V_{y|\psi;t}+\delta(y|\psi,\tilde{y}|\tilde{\psi})-1}d\Theta =$$

$$= \frac{1}{\mathcal{B}(V_{\tilde{\psi};t})}\frac{\Gamma(V_{1|\tilde{\psi};t})\Gamma(V_{2|\tilde{\psi};t})\ldots\Gamma(V_{\tilde{y}|\tilde{\psi};t}+1)\ldots\Gamma(\mathring{y}|\tilde{\psi};t)}{\Gamma(V_{1|\tilde{\psi};t}+\ldots+V_{\mathring{y}|\tilde{\psi};t}+1)} =$$

$$= \frac{1}{\mathcal{B}(V_{\tilde{\psi};t})} \frac{\Gamma(V_{1|\tilde{\psi};t})\Gamma(V_{2|\tilde{\psi};t})\dots\Gamma(V_{\tilde{y}|\tilde{\psi};t})(V_{\tilde{y}|\tilde{\psi};t})\dots\Gamma(\mathring{y}|\tilde{\psi};t)}{\Gamma(V_{1|\tilde{\psi};t}+\dots+V_{\mathring{y}|\tilde{\psi};t})(V_{1|\tilde{\psi};t}+\dots+V_{\mathring{y}|\tilde{\psi};t})} = \frac{V_{\tilde{y}|\tilde{\psi};t}}{V_{1|\tilde{\psi};t}+\dots+V_{\mathring{y}|\tilde{\psi};t}}$$

where

$\delta(y|\psi,\tilde{y}|\tilde{\psi})$ denotes matrix of corresponding delta-functions,

$\mathcal{B}(V_{\tilde{\psi};t})$ denotes the term of the multivariate beta function (7.10) corresponding to $\tilde{\psi}$ (the remaining terms got cancelled).

Second moments (co-variance):

For different ψ the terms are independent and so their co-variances are equal to 0. That is why, in the following, we will omit the index ψ and we will consider just a single regression vector. The result for more than one regression vectors are identical (similarly as for mean value, the unchanged terms are cancelled)

Co-variance for common regression vector:

$$\mathrm{cov}[\Theta_{\tilde{y}}\Theta_{\hat{y}}|\mathcal{P}_{a^*_{t+1}}] = \mathcal{E}\left[(\Theta_{\tilde{y}} - \mathcal{E}[\Theta_{\tilde{y}}|d(t)])\,(\Theta_{\hat{y}} - \mathcal{E}[\Theta_{\hat{y}}|d(t)])|\,d(t)\right] =$$

$$= \mathcal{E}\left[\Theta_{\tilde{y}}\,\Theta_{\hat{y}}|\,d(t)\right] - \mathcal{E}\left[\Theta_{\tilde{y}}|\,d(t)\right]\mathcal{E}\left[\Theta_{\hat{y}}|\,d(t)\right] = \frac{\mathcal{B}(V_t+J)}{\mathcal{B}(V_t)} - \mathcal{E}\left[\Theta_{\tilde{y}}|\,d(t)\right]\mathcal{E}\left[\Theta_{\hat{y}}|\,d(t)\right]$$

where V_t is row vector of statistics corresponding to the single regression vector and J is a row vector of zeros up to two ones at positions of $\tilde{y}$ and $\hat{y}$.

Let us notice the first term

$$\frac{\mathcal{B}(V_t+J)}{\mathcal{B}(V_t)} = \frac{1}{\mathcal{B}(V_t)} \frac{\Gamma(V_{1;t})\dots\Gamma(V_{\tilde{y};t}+1)\dots\Gamma(V_{\hat{y};t}+1)\dots\Gamma(V_{\mathring{y};t})}{\Gamma(V_{1;t}+\dots+V_{\mathring{y};t}+2)} = \dagger$$

where in the numerator, the only two terms with arguments increased by one are those indicated. The rest are $V(i,t)$ for $i \neq \tilde{y}$ and $i \neq \hat{y}$. Now, the terms whose arguments are increased can be expressed according to (7.4)

$\Gamma(V_{\tilde{y};t}+1) = \Gamma(V_{\tilde{y};t})V_{\tilde{y};t}$, $\quad\Gamma(V_{\hat{y};t}+1) = \Gamma(V_{\hat{y};t})V_{\hat{y};t}$ and

$\Gamma(V_{1;t}+\dots+V_{\mathring{y};t}+2) = \Gamma(V_{1;t}+\dots+V_{\mathring{y};t})(\sum_{i=1}^{\mathring{y}} V_{i;t}+1)(\sum_{i=1}^{\mathring{y}} V_{i;t})$.

After cancelling with the beta function in denominator we obtain

$$\dagger = \frac{V_{\tilde{y};t}V_{\hat{y};t}}{(\sum_{i=1}^{\mathring{y}} V_{i;t}+1)(\sum_{i=1}^{\mathring{y}} V_{i;t})} = \frac{V_{\tilde{y};t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}} \frac{V_{\hat{y};t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}} \frac{\sum_{i=1}^{\mathring{y}} V_{i;t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}+1} =$$

$$= \mathcal{E}[\Theta_{\tilde{y}}|d(t))]\mathcal{E}[\Theta_{\hat{y}}|d(t))]\frac{\sum_{i=1}^{\mathring{y}} V_{i;t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}+1}.$$

Completing this term with the skipped product of the mean values from the beginning off this derivation we obtain

$$\mathcal{E}[\Theta_{\tilde{y}}|d(t))]\mathcal{E}[\Theta_{\hat{y}}|d(t))]\left(\frac{\sum_{i=1}^{\mathring{y}} V_{i;t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}+1} - 1\right)$$

which is the result to be proved.

Second moments (variance):

Again with omitted regression vector as multi-index we have

$$\mathrm{var}[\Theta_{\tilde{y}}|\mathcal{P}_{a^*_{t+1}}] = \mathcal{E}\left[(\Theta_{\hat{y}} - \mathcal{E}[\Theta_{\hat{y}}|d(t)])^2 \middle| d(t)\right] = \mathcal{E}\left[\Theta^2_{\hat{y}} \middle| d(t)\right] - \mathcal{E}\left[\Theta_{\hat{y}} \middle| d(t)\right]^2$$

With similar steps as in the previous case we can express

$$\mathcal{E}\left[\Theta^2_{\hat{y}} \middle| d(t)\right] = \frac{\mathcal{B}(V_t + J)}{\mathcal{B}(V_t)} = \dagger$$

with J all zeros but a single item 2 at the position of $\tilde{y}$.

$$\dagger = \frac{1}{\mathcal{B}(V_t)} \frac{\Gamma(V_{1;t}) \dots \Gamma(V_{\tilde{y};t} + 2) \dots \Gamma(V_{\mathring{y};t})}{\Gamma(\sum_{i=1}^{\mathring{y}} V_{i;t} + 2)} = \frac{(V_{\tilde{y};t} + 1)(V_{\tilde{y};t})}{(\sum_{i=1}^{\mathring{y}} V_{i;t} + 1)(\sum_{i=1}^{\mathring{y}} V_{i;t})} =$$

$$= \left[\frac{V_{\tilde{y};t}}{\sum_{i=1}^{\mathring{y}} V_{i;t}}\right]^2 \left[\frac{\sum_{i=1}^{\mathring{y}} V_{i;t}}{\sum_{i=1}^{\mathring{y}} V_{i;t} + 1} \frac{V_{\tilde{y};t} + 1}{V_{\tilde{y};t}}\right]$$

The whole expression, with the square of mean values that has been omitted, is

$$(\mathcal{E}[\Theta_{\tilde{y}}|d(t)])^2 \left[\frac{\sum_{i=1}^{\mathring{y}} V_{i;t}}{\sum_{i=1}^{\mathring{y}} V_{i;t} + 1} \frac{V_{\tilde{y};t} + 1}{V_{\tilde{y};t}} - 1\right]$$

which is the proved expression. ♢

Remark(s) 6.4

1. *Notice that the expected values, "natural point estimates", coincide with relative frequencies of occurrence of the particular situation observed in past. It makes the result intuitively appealing. Note that the relative frequency is modified by the prior choice. The influence asymptotically disappear, cf. Proposition 2.15, if the number* $\sum_{\bar{y}\in y^*} V_{\bar{y}|\psi}$ *approaches infinity. For it, this specific condition has to appear infinitely often times. If this condition is not fulfilled the (potentially positive) influence of the prior pdf persists. Note, that under insufficient excitation (when the denominator in moments is finite) there is no estimator that can estimate the unknown parameter consistently.*

2. *Note that the estimation runs independently for a respective fixed conditions. The dependence is introduced just through restrictions on* $\Theta_{y|\psi}$ *that are set independently for each condition* ψ*.*

3. *Restricting of some entries of* $\Theta_{y|u,\tilde{y}}$ *to fixed values (zeros or given numbers) is a non-trivial practically feasible and significant special case that allows us to respect a smoothness of the sampled continuous signals [68].*

6.3.2 Set estimation

According to (3.27), integrals

SET ESTIMATE

$$\int_{\Theta^*} f(\Theta|d(t))d\Theta = \int_{\Theta^*} \prod_{\psi} \prod_{y} \Theta_{y|\psi}^{V_{y|\psi;t}-1} d\Theta \tag{6.31}$$

are necessary to be evaluated. This is generally a difficult task requiring numerical integration.

Example:

The situation gets simpler, if rectangular areas (6.6) for Θ^* are considered. Thus, for the **CCM problem** with model pdf (6.22) it holds

$$\int_{\Theta^*} f(\Theta|d(t))d\Theta = \prod_{\psi} \int_{\theta^*_\psi} \Theta_{\mathbf{0}|\psi}^{V_{\mathbf{0}|\psi;t}-1} \Theta_{\mathbf{1}|\psi}^{V_{\mathbf{1}|\psi;t}-1} d\Theta_\psi = \prod_{\psi} \left[\mathcal{B}_{V_{\psi;t}}(\theta^u_\psi) - \mathcal{B}_{V_{\psi;t}}(\theta^l_\psi)\right], \tag{6.32}$$

where $\mathcal{B}_{V_{\psi;t}}(\theta^u_\psi)$ and $\mathcal{B}_{V_{\psi;t}}(\theta^l_\psi)$ are incomplete beta functions (7.9) and $\Theta^*_\psi \to (\theta^l_\psi, \theta^u_\psi)$.

6.3.3 Testing of hypotheses

Here, the general formulae (3.33) are specified for the considered model (6.9) and cartesian admissible parameter sets (6.6).

Let us suppose, that the structure of the model is valid and we are testing several hypotheses $H_h,\ h = 1, 2, \ldots, \mathring{h}$ about various areas Θ^*_h to which the parameter $\Theta \equiv \Theta_h$ belongs. Thus, the hypotheses are

$$H_h :\ \Theta_h \in \Theta^*_h,\ h = 1, 2, \ldots, \mathring{h}. \tag{6.33}$$

We have parameters $\Theta_{\psi,h}$ for $\psi \in \psi^*$ and $h = 1, 2, \ldots, \mathring{h}$.

Example:

Generally, for arbitrary parameter regions Θ^*_h the task is not analytically solvable and we restrict ourselves to the **CCM problem** with admissible sets of parameters according to (6.6).

For model (6.9), hypotheses (6.33) with Θ^*_h according to (6.6) and with conjugate priors $f(\Theta_{\psi,h})$ (6.12) and prior on hypotheses $f(h)$ it holds

$$f(h|d(t)) = \frac{\mathcal{I}_h(d(t))}{\mathcal{I}_h(d(t-1))} f(h), \tag{6.34}$$

$$\mathcal{I}_h(d(t)) = \prod_{\psi} \left[\mathcal{B}_{V_{\mathbf{1}|\psi;t}, V_{\mathbf{0}|\psi;t}}(\theta^u_{\psi,h}) - \mathcal{B}_{V_{\mathbf{1}|\psi;t}, V_{\mathbf{0}|\psi;t}}(\theta^l_{\psi,h})\right] \tag{6.35}$$

Proof: The formula (6.34) is according to (3.33). ♢

The integrand in (6.35) is a product of likelihood and parameter prior, both according to hypothesis h. The likelihood is (**??**) and the conjugate prior (**??**) which in product gives the form of un-normalized posterior pdf (**??**). For rectangular area of integration (**??**) the integral of product gets product of integrals and it is

$$\mathcal{I}_h(d(t)) = \prod_{\psi} \int_{\theta^*_{\psi,h}} \Theta_{\mathbf{0}|\psi,h}^{V_{\mathbf{0}|\psi;t}} \Theta_{\mathbf{1}|\psi,h}^{V_{\mathbf{1}|\psi;t}} d\Theta_{\psi,h}$$

which gives (6.35).

6.3.4 One-step-ahead prediction

This task is generally formulated in (3.34) and the explicit solution for quadratic criterion is given in (3.35). Its adaptation for this task is

$$\hat{\Delta}(D) = \hat{y}_{t+1} = E[y_{t+1}|u_{t+1}, d(t)] = \sum_{y \in y^*} y\, f(y|u_{t+1}, d(t))$$

Proposition 6.8 (Point one step quadratic prediction) *For the discrete model (6.9) and quadratic criterion the one step prediction* $\hat{y}_{t+1}$ *is*

POINT PREDICTION

$$\hat{y}_{t+1} = \frac{\sum_{y \in y^*} y\, V_{y_{t+1}|\psi_{t+1};t}}{\sum_{y \in y^*} V_{y|\psi_{t+1};t}}. \tag{6.36}$$

Proof: Generally, the prediction is given by (3.34). For quadratic criterion, the point-prediction can be explicitly expressed by the relation (3.35), which, for discrete model (6.9), can be written in the form $\hat{y}_{t+1} = \sum_{y \in y^*} y\, f(y_{t+1}|u_{t+1}, d(t))$, where $f(y_{t+1}|u_{t+1}, d(t)) = \frac{V_{y_{t+1}|\psi_{t+1};t}}{\sum_{y \in y^*} V_{y|\psi_{t+1};t}}$ is predictive pdf (6.25). Substitution for the predictive pdf gives the proved formula $\sum_{y \in y^*} y\, V_{y_{t+1}|\psi_{t+1};t} / \sum_{y \in y^*} V_{y|\psi_{t+1};t}$. ♢

Example:

Point one step prediction for quadratic

For the CCM problem, the one step prediction $\hat{y}_{t+1}$ with quadratic loss function, is

$$\hat{y}_{t+1} = \frac{V_{\mathbf{1}|\psi;t}}{V_{\mathbf{1}|\psi;t} + V_{\mathbf{0}|\psi;t}}. \tag{6.37}$$

Proof: For quadratic loss function assumed, the point prediction is conditional mean value of the output, conditioned by all measured data $d(t)$ (see Remark to (3.34)). The predictive pdf is (6.25). ♢

Point one step prediction for absolute criterion

For the CCM problem, the one step ahead prediction $\hat{y}_{t+1}$, according to with absolute loss function $|y_\tau - \hat{y}_\tau|$, the point prediction $\hat{y}_\tau$ is the value of y_τ with the maximal value of predictive pdf $f(y_\tau|d(\tau-1), u_\tau)$. This gives one-step prediction

$$\hat{y}_{t+1} = \begin{cases} \mathbf{1} & \text{for } V_{\mathbf{1}|\psi;t} \geq V_{\mathbf{0}|\psi;t}, \\ \mathbf{0} & \text{for } V_{\mathbf{1}|\psi;t} < V_{\mathbf{0}|\psi;t}. \end{cases} \tag{6.38}$$

Proof: According to (3.34) for point prediction $\hat{y}$ it holds $\hat{y} = \arg\min_{\hat{y}} \sum_y |\hat{y} - y|$ where the predictive pdf $f(y|\mathcal{P}_y) = [f_{\mathbf{1}}, f_{\mathbf{0}}]$ is (6.25). So we look for minimum of $|\hat{y} - 1| f_{\mathbf{1}} + |\hat{y} - 0| f_{\mathbf{0}}$ which is $f_{\mathbf{0}}$ for $\hat{y} = \mathbf{1}$ and $f_{\mathbf{1}}$ for $\hat{y} = \mathbf{0}$. From it and from (6.25) follows the proved relation. ♢

Remark(s) 6.5

1. *No point estimation of the parameter is needed for making predictions.*
2. *The relative frequency is modified by the prior values of the sufficient statistics n_0. Its influence may be significant especially when the number of observed situation is small. It may happen due to the insufficient richness of the input sequence and not only for small extent of measured data. This situation called* insufficient excitation *is more rule than exception in controlled systems.*

6.3.5 One-step-ahead control

The general formula for generating the optimal control is (3.3.5) with $\mathcal{Z}$ being the loss function. In terms of our case this formula gets the form

$$u_t(d(t-1)) = \arg\min_{u\in u^*} \sum_{y\in y^*} \mathcal{Z}(y,u,d(t-1)) f(y|u,\varphi_{t-1}).$$

For known model parameters, the predictive pdf $f(y|u,\varphi_{t-1})$ is directly given by the model and it holds

Proposition 6.9 (One-step control with known parameters) *Let us consider the discrete model (6.9) and the loss function defined in as a penalization table*

$$\mathcal{Z}(y,u,d) = \omega_{y,u,d}, \tag{6.39}$$

where y is modelled variable, u is control, d is vector of past data and values of $\omega_{y,u,d}$ represent individual penalties (constant for all $t \in t^*$). Then, the optimal control law is given by the the solution of the minimization task

ONE STEP CONTROL

$$\mathcal{V}_{u;t}(d(t-1)) = \sum_{y=\mathbf{1}}^{\mathring{y}} \omega_{y,u,d(t-1)} \Theta_{y|u,\varphi_{t-1}}$$
$$\hat{u}_t(d(t-1)) = \arg\min_{u\in u^*} \mathcal{V}_{u;t}(d(t-1))$$

Proof: The result is just the general formula above with the model pdf (6.9) and the penalization table (6.39) substituted. ♢

We will deal with the **CCM problem** (see Agreement 6.1), again.

Example:

One-step control with known model parameters

For the **CCM problem** with known model parameters and tabular loss function $\mathcal{Z}(d(t)) = \omega_{y_t,u_t}$ we define the control criterion

$$\mathcal{V}_{u;t} = \omega_{\mathbf{0},u}\Theta_{\mathbf{0}|u,\varphi_{t-1}} + \omega_{\mathbf{1},u}\Theta_{\mathbf{1}|u,\varphi_{t-1}}. \tag{6.40}$$

The generator of the optimal control variable $\hat{u}_t$ is then

$$\hat{u}_t = \begin{cases} \mathbf{1} & \text{for } \mathcal{V}_{\mathbf{0};t-1} \geq \mathcal{V}_{\mathbf{1};t-1} \\ \mathbf{0} & \text{for } \mathcal{V}_{\mathbf{0};t-1} < \mathcal{V}_{\mathbf{1};t-1}. \end{cases} \tag{6.41}$$

Proof: The result is just the previous one with $y^* = u^* = \{\mathbf{0},\mathbf{1}\}$. ♢

Example:

One-step control with unknown model parameters

For the **CCM problem** with unknown model parameters and tabular loss function $\mathcal{Z}(d(t)) = \omega_{y_t,u_t}$ the previous result holds with the only difference, that the parameters are substituted by their point estimates

$$\theta_{u,\varphi} \to \hat{\theta}_{u_t,\varphi_{t-1}}$$

according to (6.29). *Proof:* The criterion is $\sum_{y_t} \omega_{y_t|u_t} f(y_t|u_t, d(t-1))$, where the predictive pdf is the ratio of beta functions (6.25) which, according to (6.28) and (6.29) defines point estimate of parameter. ♢

6.4 Dynamic tasks

6.4.1 Sequential estimation

Due to the fact, that the discrete model (6.9) belongs to the exponential family (6.10) the task of sequential estimation is simple, as it reduces to algebraic recursion for statistics $\delta(\Psi, \Psi_t)$ for $t \in t^*$. The recursion for data collection and the form of posterior pdf of the estimated parameter Θ is given in the following proposition.

Proposition 6.10 (Sequential estimation for discrete model) *For the discrete model (6.9) and given prior pdf (6.12) the posterior pdf is (6.22) with the statistics $V_{y|\psi}$, $y \in y^*$, $\psi \in \psi^*$ according to (6.23).*

Proof: The result is due to the Proposition 6.4. ♢

6.4.2 Sequential estimation with stopping rule

The estimation of CCM model parameters is described by the equations (6.50) and (6.51). The stopping rule, in general, is defined in (3.38).

For the example, considered here, we suppose:

1. The loss, connected with data measurements, is fix; i.e. $c(d(t-1)) = c = \text{const}$.
2. The loss, connected with the inaccuracy of parameter estimates, is quadratic; i.e.

$$Z(d(t-1), \Theta, \hat{\Theta}_t) = (\Theta - \hat{\Theta}_t)^2 \tag{6.42}$$

3. Model of input variable u_t in the form

$$f(u_t|y_{t-1}) = {}^{\lfloor u}\theta u_t|y_{t-1} \tag{6.43}$$

4. Indeed, the CCM model (6.9) for output variable is considered.

Under these assumptions, the condition (3.38) can be written in the form

$$\mathcal{E}[(\theta-\hat{\theta}_t)^2-(\theta-\hat{\theta}_{t+k})^2-k.c|d(t-1)]=\mathcal{E}[(\theta-\hat{\theta}_t)^2|d(t-1)]-\mathcal{E}[(\theta-\hat{\theta}_{t+k})^2|d(t-1)]-C, \quad (6.44)$$

with $C = k.c$ =const.

As the optimal point estimate of the parameter θ is conditional mean value, the expectation of the first two terms form the variance and k-step ahead prediction of the variance of the parameter estimate. The first one is given by (6.30), the second one is necessary to compute and the third one is finished.

The result for the second term (prediction of parameter variance) is given by the following proposition

Proposition 6.11 (Prediction of parameter estimates variance) *For the CCM model (6.9) and input generator model via (6.43), the k-step ahead prediction of parameter estimates variance is*

$$\mathcal{E}\left[(\theta-\hat{\theta}_{k+1})^2|u_t,d(t-1)\right]=$$

$$=\sum_D s^2(u_{t+k},d(t+k-1))f(u_{t+k}|d(t+k-1))f(y_{t+k-1}|u_{t+k-1},d(t+k-2))f(u_{t+k-1}|d(t+k-2))\dots$$

$$\dots f(y_{t+1}|u_{t+1},d(t))f(u_{t+1}|d(t))f(y_t|u_t,d(t-1)),$$

where

$D=\{u_{t+k},\ d(t+k-1\_\,t+1),\ y_t\}$.

Proof: Using chain rule, the expectation can be expressed in the form

$$s_p^2=\mathcal{E}\left[(\theta-\hat{\theta}_{k+1})^2|u_t,d(t-1)\right]=\mathcal{E}\left[\mathcal{E}[(\theta-\hat{\theta}_{k+1})^2|u_{t+k},d(t+k-1)]\Big|\,u_t,d(t-1)\right].$$

The inner expectation is variance of parameter estimates at time $t+k$ given by (6.30). We denote it by

$$s^2(u_{t+k},d(t+k-1))=\mathcal{E}[(\theta-\hat{\theta}_{k+1})^2|u_{t+k},d(t+k-1)],$$

and, using chain rule again, we have

$$s_p^2=\mathcal{E}\left[s^2(u_{t+k},d(t+k-1))\Big|\,u_t,d(t-1)\right]=$$

$$=\sum_D s^2(u_{t+k},d(t+k-1))f(u_{t+k}|d(t+k-1))f(y_{t+k-1}|u_{t+k-1},d(t+k-2))f(u_{t+k-1}|d(t+k-2))\dots$$

$$\dots f(y_{t+1}|u_{t+1},d(t))f(u_{t+1}|d(t))f(y_t|u_t,d(t-1)),$$

with D given above. ♢

The computation of the variance prediction can be done according to the following algorithm:

- Generate all possible value combinations of the data sequence D. It can be constructed like all binary numbers of the length equal to the length of D starting form all zeros and sequentially adding 1 till all ones.

- Starting with prior statistics comprising data $\{u_t,\ d(t-1)\}$ evaluate the statistics up to time $t+k-1$ for each of the possible data combination D.
- For each step of the evaluation above compute one step predictive pf (normalized statistics according to (6.25)) and make products indicated in (6.11). Each final product represents a probability of the considered possible path on the prediction horizon.
- Similarly like an element of the predictive pf compute the parameter estimate variance according to (6.30) corresponding to the chosen path.
- The weighted sum of products of predictive pf elements with corresponding variances gives the computed expectation - the prediction of the variance.

Algorithm 6.1 (Sequential estimation with stopping rule) *The sequential estimation of parameters of the CCM model (6.9) using the stopping rule (3.38) can be performed as follows*

1. *Set:*
   - *the maximum data length for estimation (nd),*
   - *the length of parameter estimate variance prediction (np),*
   - *the prize for data on the prediction horizon (C).*
2. *Run the estimation of the CCM parameter according to Proposition 6.10 and at each step of the estimation compute one-step and np-step predictions of the parameter estimate variance according to 6.11.*
3. *Compare the difference of predicted variances and the prize according to (6.44). If the condition is fulfilled, stop the estimation, otherwise, continue by measuring another data pair.*

6.4.3 Multi-step filtering

The multi-step filtering for the CCM model can be specified in the following way. For this task, let us consider unmeasurable time varying state, denoted by x_t. This state develops internally affected by last state value and current input value, only. This development can be monitored through the measured output. The models, describing the state development and the state influence onto the output are both of the CCM type

$$f(x_t|x_{t-1},u_t) = {}^{\lfloor x}\theta_{x_t|x_{t-1},u_t}, \tag{6.45}$$

$$f(y_t|x_t,u_t) = {}^{\lfloor y}\theta_{y_t|x_t,u_t}. \tag{6.46}$$

Proposition 6.12 (CCM multi-step pf for filtering) *For the CCM models (6.45) and (6.46) with known parameters ${}^{\lfloor x}\theta$ and ${}^{\lfloor y}\theta$ the needed pf is*

$$f(x_{t-j}|d(t-1)) = \sum_{x_{t-j+1}}^{x_{t-1}} \prod_{\tau=t-j+1}^{t-1} {}^{\lfloor y}\theta_{y_\tau|x_\tau,u_\tau}\ {}^{\lfloor x}\theta_{x_\tau|x_{\tau-1},u_\tau}\ \times \tag{6.47}$$

$$\times\ {}^{\lfloor y}\theta_{y_{t-j}|x_{t-j},u_{t-j}} f(x_{t-j}|d(t-j-1))$$

where the state pfs are formed by column vectors of dimensions 2.

Proof: The result is due to (3.40) with pfs (6.45) and (6.46) substituted. ♢

Proposition 6.13 (State filtering) *For the CCM models (6.45) and (6.46) with known parameters ${}^{\lfloor x}\theta$ and ${}^{\lfloor y}\theta$ the optimal decision about filtered state value $\hat{x}_{t-j|t-1}$ with optimality criterion (3.39) is*

$$\hat{x}_{t-j|t-1} = \begin{cases} 1 & \text{for } f(x_{t-j}=1|d(t-1)) > f(x_{t-j}=2|d(t-1)), \\ 2 & \text{otherwise.} \end{cases} \tag{6.48}$$

Proof: The result follows directly from the criterion definition (3.39). ♢

6.4.4 Multi-step prediction

The multi-step output prediction for the CCM model is the following.

Proposition 6.14 (CCM multi-step output prediction) *For the CCM model (6.9) with unknown parameter θ (6.8) the multi-step output prediction is*

$$\hat{y}_{t+k} = \mathcal{E}[y_t|u_{t+1}, d(t)] = \sum_{y_{t+k}=1}^{\mathring{y}} y_{t+k}\, f(y_{t+k}|u_{t+1}, d(t)), \tag{6.49}$$

where the multi-step predictive pf is given in (6.27).

Proof: See (6.27). ♢

6.4.5 Multi-step control

Proposition 6.15 (Coin dynamic control) *For the CCM problem Agreement 6.1 with the first order model $f(y_t|u_t, y_{t-1}) = \Theta_{y_t|u_t,y_{t-1}}$ i.e. with $\varphi_{t-1} = y_{t-1}$, known parameters Θ and the control horizon $\mathring{t}$ the general formula (2.23) can be expressed in the following two steps:*

mean value

$$\mathcal{V}_t(u_t, y_{t-1}) = \left[\omega_{\boldsymbol{0}|u_t,y_{t-1}} + \mathcal{V}^*_{t+1}(\boldsymbol{0})\right] \Theta_{\boldsymbol{0}|u_t,y_{t-1}} + \left[\omega_{\mathbf{1}|u_t,y_{t-1}} + \mathcal{V}^*_{t+1}(\mathbf{1})\right] \Theta_{\mathbf{1}|u_t,y_{t-1}} \tag{6.50}$$

and ***minimum***

$$\mathcal{V}^*_t(y_{t-1}) = \min\{\mathcal{V}_t(\boldsymbol{0}, y_{t-1}), \mathcal{V}_t(\mathbf{1}, y_{t-1})\}, \qquad \Rightarrow u^*_t \tag{6.51}$$

for $t = \mathring{t}, \mathring{t}-1, \ldots, 1$, with the initial condition $\mathcal{V}^*_{\mathring{t}+1}(\boldsymbol{0}) = \mathcal{V}^*_{\mathring{t}+1}(\mathbf{1}) = 0$.

Algorithm 6.2 (Coin dynamic control) *The algorithm of the computation of control law and its realization is as follows*

Control law computation

$\mathcal{V}^*_{\mathring{t}+1}(\boldsymbol{0}) = 0, \mathcal{V}^*_{\mathring{t}+1}(\mathbf{1}) = 0$
for $t = \mathring{t} : -1 : 1$

for $y_{t-1} = \boldsymbol{0} : \mathbf{1}$

for $u = \boldsymbol{0} : \mathbf{1}$

$$\mathcal{V}_t(u, y_{t-1}) = [\omega_{\boldsymbol{0}|u,y_{t-1}} + \mathcal{V}^*_{t+1}(\boldsymbol{0})]\Theta_{\boldsymbol{0}|u,y_{t-1}} + [\omega_{\mathbf{1}|u,y_{t-1}} + \mathcal{V}^*_{t+1}(\mathbf{1})]\Theta_{\mathbf{1}|u,y_{t-1}}$$

end *% for* u

$$[\mathcal{V}^*_t(y_{t-1}), u_t(y_{t-1})] = \min\{\mathcal{V}_t(\boldsymbol{0}, y_{t-1}), \mathcal{V}_t(\mathbf{1}, y_{t-1})\}$$

end *% for* y_{t-1}

end *% for* t

Control law realization

for $t = 1 : \mathring{t}$

1. *Measure* y_{t-1}
2. *Select optimal* $u_t = u_t(y_{t-1})$
3. *Apply optimal* u_t

end *% for* t

Derivation

The relations (6.50) and (6.51) are obtained from the general formulae for mean

$$\mathcal{V}_t = E[\omega_{y_t,|u_t,y_{t-1}} + \mathcal{V}^*_{t+1}|u_t, d(t-1)] \tag{6.52}$$

and minimum

$$\mathcal{V}^*_t = \min_{u_t} \mathcal{V}_t \quad \Rightarrow \quad u^*_t(y_{t-1}) \tag{6.53}$$

that follow from the basic optimization Proposition 6.2 for sequential minimization of an additive loss function.

We start minimization in the last time instant $\mathring{t}$. The mean is

$$\mathcal{V}_{\mathring{t}}(u_{\mathring{t}}, y_{\mathring{t}-1}) = \sum_{y=\mathbf{0}}^{\mathbf{1}} \omega_{y|u_{\mathring{t}},y_{\mathring{t}-1}} \Theta^{\delta(y,\mathbf{0})}_{y|u_{\mathring{t}},y_{\mathring{t}-1}} \Theta^{\delta(y,\mathbf{1})}_{y|u_{\mathring{t}},y_{\mathring{t}-1}} = \omega_{\mathbf{0}|u_{\mathring{t}},y_{\mathring{t}-1}} \Theta_{\mathbf{0}|u_{\mathring{t}},y_{\mathring{t}-1}} + \omega_{\mathbf{1}|u_{\mathring{t}},y_{\mathring{t}-1}} \Theta_{\mathbf{1}|u_{\mathring{t}},y_{\mathring{t}-1}}$$

For all possible values of $\psi_{\mathring{t}} = [u_{\mathring{t}}, y_{\mathring{t}-1}]'$ we obtain the criterion in the form of table, i.e. as a discrete function of $u_{\mathring{t}}, y_{\mathring{t}-1}$.

Table of $\mathcal{V}_{\mathring{t}}(\psi_{\mathring{t}})$ **for** $\psi_{\mathring{t}} = [u_{\mathring{t}}, y_{\mathring{t}-1}]'$

configuration of $\psi_{\mathring{t}}$		criterion $\mathcal{V}_{\mathring{t}}(\psi_{\mathring{t}})$
$u_{\mathring{t}}, y_{\mathring{t}-1}$:	$\mathbf{0}, \mathbf{0}$	$\omega_{\mathbf{0}|\mathbf{0},\mathbf{0}}\Theta_{\mathbf{0}|\mathbf{0},\mathbf{0}} + \omega_{\mathbf{1}|\mathbf{0},\mathbf{0}}\Theta_{\mathbf{1}|\mathbf{0},\mathbf{0}}$
	$\mathbf{1}, \mathbf{0}$	$\omega_{\mathbf{0}|\mathbf{1},\mathbf{0}}\Theta_{\mathbf{0}|\mathbf{1},\mathbf{0}} + \omega_{\mathbf{1}|\mathbf{1},\mathbf{0}}\Theta_{\mathbf{1}|\mathbf{1},\mathbf{0}}$
$u_{\mathring{t}}, y_{\mathring{t}-1}$:	$\mathbf{0}, \mathbf{1}$	$\omega_{\mathbf{0}|\mathbf{0},\mathbf{1}}\Theta_{\mathbf{0}|\mathbf{0},\mathbf{1}} + \omega_{\mathbf{1}|\mathbf{0},\mathbf{1}}\Theta_{\mathbf{1}|\mathbf{0},\mathbf{1}}$
	$\mathbf{1}, \mathbf{1}$	$\omega_{\mathbf{0}|\mathbf{1},\mathbf{1}}\Theta_{\mathbf{0}|\mathbf{1},\mathbf{1}} + \omega_{\mathbf{1}|\mathbf{1},\mathbf{1}}\Theta_{\mathbf{1}|\mathbf{1},\mathbf{1}}$

Note: All elements of this matrix can be computed right now. The first two rows hold for $y_{\mathring{t}-1} = \mathbf{0}$ and the second two for $y_{\mathring{t}-1} = \mathbf{1}$. For each couple of rows we take minimum and denote it by $\mathcal{V}^*_{\mathring{t}}(y_{\mathring{t}-1})$. Thus

$$\begin{aligned} \mathcal{V}^*_{\mathring{t}}(\mathbf{0}) &= \min\{\mathcal{V}_{\mathring{t}}(\mathbf{0}, \mathbf{0}), \mathcal{V}_{\mathring{t}}(\mathbf{1}, \mathbf{0})\} \quad \text{first couple} \\ \mathcal{V}^*_{\mathring{t}}(\mathbf{1}) &= \min\{\mathcal{V}_{\mathring{t}}(\mathbf{0}, \mathbf{1}), \mathcal{V}_{\mathring{t}}(\mathbf{1}, \mathbf{1})\} \quad \text{second couple} \end{aligned}$$

The value of $u_{\mathring{t}}$ in the minimum item of each couple defines the optimal control

$$u^*_{\mathring{t}}(y_{\mathring{t}-1}) = \begin{cases} \mathbf{0} \text{ if } \mathcal{V}_{\mathring{t}}(\mathbf{0},\mathbf{0}) \leq \mathcal{V}_{\mathring{t}}(\mathbf{1},\mathbf{0}) \text{ and } \mathbf{1} \text{ if } \mathcal{V}_{\mathring{t}}(\mathbf{0},\mathbf{0}) > \mathcal{V}_{\mathring{t}}(\mathbf{1},\mathbf{0}) & \text{for } \; y_{\mathring{t}-1} = \mathbf{0} \\ \mathbf{0} \text{ if } \mathcal{V}_{\mathring{t}}(\mathbf{0},\mathbf{1}) \leq \mathcal{V}_{\mathring{t}}(\mathbf{1},\mathbf{1}) \text{ and } \mathbf{1} \text{ if } \mathcal{V}_{\mathring{t}}(\mathbf{0},\mathbf{1}) > \mathcal{V}_{\mathring{t}}(\mathbf{1},\mathbf{1}) & \text{for } \; y_{\mathring{t}-1} = \mathbf{1} \end{cases}$$

Note: As we do not know the value of $y_{\mathring{t}-1}$, we are not able to decide about the optimal control, at the moment.

This concludes the last step. The last but one step of minimization is for time $\mathring{t}-1$. The rest from the previous step is $\mathcal{V}^*_{\mathring{t}}(y_{\mathring{t}-1})$. The mean value is – see (6.52)

$$\mathcal{V}_{\mathring{t}-1}(u_{\mathring{t}-1}, y_{\mathring{t}-2}) = \sum_{y=\mathbf{0}}^{\mathbf{1}} [\omega_{y|u_{\mathring{t}-1},y_{\mathring{t}-2}} + \mathcal{V}^*_{\mathring{t}}(y_{\mathring{t}-1})]\, \Theta^{\delta(y,\mathbf{0})}_{y|u_{\mathring{t}-1},y_{\mathring{t}-2}} \Theta^{\delta(y,\mathbf{1})}_{y|u_{\mathring{t}-1},y_{\mathring{t}-2}} =$$

$$= [\omega_{\mathbf{0}|u_{\mathring{t}-1},y_{\mathring{t}-2}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{0})]\Theta_{\mathbf{0}|u_{\mathring{t}-1},y_{\mathring{t}-2}} + [\omega_{\mathbf{1}|u_{\mathring{t}-1},y_{\mathring{t}-2}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{1})]\Theta_{\mathbf{1}|u_{\mathring{t}-1},y_{\mathring{t}-2}}$$

Again, it can be expressed as a table for all possible values of $y_{\mathring{t}-2}$.

Table of $\mathcal{V}_{\mathring{t}-1}(\psi_{\mathring{t}-1})$ for $\psi_{\mathring{t}-1} = [u_{\mathring{t}-1}, y_{\mathring{t}-2}]'$

configuration of $\psi_{\mathring{t}-1}$		criterion $\mathcal{V}_{\mathring{t}}(\psi_{\mathring{t}-1})$
$u_{\mathring{t}-1}, y_{\mathring{t}-2}:$	$\mathbf{0},\mathbf{0}$	$[\omega_{\mathbf{0}\|\mathbf{0},\mathbf{0}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{0})]\Theta_{\mathbf{0}\|\mathbf{0},\mathbf{0}} + [\omega_{\mathbf{1}\|\mathbf{0},\mathbf{0}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{1})]\Theta_{\mathbf{1}\|\mathbf{0},\mathbf{0}}$
	$\mathbf{1},\mathbf{0}$	$[\omega_{\mathbf{0}\|\mathbf{1},\mathbf{0}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{0})]\Theta_{\mathbf{0}\|\mathbf{1},\mathbf{0}} + [\omega_{\mathbf{1}\|\mathbf{1},\mathbf{0}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{1})]\Theta_{\mathbf{1}\|\mathbf{1},\mathbf{0}}$
$u_{\mathring{t}-1}, y_{\mathring{t}-2}:$	$\mathbf{0},\mathbf{1}$	$[\omega_{\mathbf{0}\|\mathbf{0},\mathbf{1}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{0})]\Theta_{\mathbf{0}\|\mathbf{0},\mathbf{1}} + [\omega_{\mathbf{1}\|\mathbf{0},\mathbf{1}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{1})]\Theta_{\mathbf{1}\|\mathbf{0},\mathbf{1}}$
	$\mathbf{1},\mathbf{1}$	$[\omega_{\mathbf{0}\|\mathbf{1},\mathbf{1}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{0})]\Theta_{\mathbf{0}\|\mathbf{1},\mathbf{1}} + [\omega_{\mathbf{1}\|\mathbf{1},\mathbf{1}} + \mathcal{V}^*_{\mathring{t}}(\mathbf{1})]\Theta_{\mathbf{1}\|\mathbf{1},\mathbf{1}}$

As we came to the same form like in the previous step, the recursion is completed and can be expressed in the Algorithm 6.2.

6.4.6 Dual control

Now we come to a complex task of dynamic (n-step ahead) control with simultaneous estimation of parameters. In this task, the optimal control law for time instants $1, 2, \ldots, \mathring{t}$ is to be precomputed, without knowledge of the true values of parameters Θ. We denote the current time $t = 1$ - i.e. we know prior data $d(0)$ and we perform sequential minimization of the loss function for time instants $\tau = \mathring{t}, \mathring{t}-1, \ldots, 1$ in which the needed data $u_\tau, d(\tau-1)$ are unknown, yet.

As the parameters Θ are unknown, the mean value of the actual part of the loss function from (2.23) has to be expressed in terms of the predictive pdf (6.25). This pdf, in difference with the model one, depends on all historical data. That is why the dual control is so complex, that it cannot be practically realized.

Proposition 6.16 (Coin dual control) *For the CCM problem Agreement 6.1 with the first order model $f(y_t|u_t, y_{t-1}) = \Theta_{y_t|u_t,y_{t-1}}$ i.e. with $\varphi_{t-1} = y_{t-1}$, unknown parameters Θ and the control horizon $\mathring{t}$ the general formula (2.23) can be expressed in the following two steps:*

mean value

$$\mathcal{V}_t(u_t, d(t-1)) = \frac{\sum_{y=\mathbf{0}}^{\mathbf{1}} \left[\omega_{y|u_t,\varphi_{t-1}} + \mathcal{V}^*_{t+1}(y, u_t, d(t-1))\right] V_{y|u_t,\varphi_{t-1};t-1}}{\sum_{y=\mathbf{0}}^{\mathbf{1}} V_{y|u_t,\varphi_{t-1};t-1}}, \tag{6.54}$$

with $V_{y_t,u_t,\varphi_{t-1};t-1} = \sum_{\tau=1}^{t-1} \delta(y_t|u_t,\varphi_{t-1};y_\tau|u_\tau,\varphi_{\tau-1}) + V_{y_t,u_t,\varphi_{t-1};0}$.

and ***minimum***

$$\mathcal{V}_t^*(y_{t-1}) = \min\{\mathcal{V}_t(\boldsymbol{0},y_{t-1}),\mathcal{V}_t(1,y_{t-1})\}, \qquad \Rightarrow u_t^* \tag{6.55}$$

for $t = \mathring{t},\mathring{t}-1,\ldots,1$, *with the initial condition* $\mathcal{V}_{\mathring{t}+1}^*(\boldsymbol{0}) = \mathcal{V}_{\mathring{t}+1}^*(\mathbf{1}) = 0$.

Algorithm 6.3 (Coin dual control) *The algorithm of* (*i*) *the computation of control law and* (*ii*) *its realization is as follows*

Control law computation

$\mathcal{V}_{\mathring{t}+1}^*([u_{\mathring{t}},y_{\mathring{t}-1},u_{\mathring{t}-1},\ldots,y_1,u_1]) = 0, \quad \forall$ *configurations of the argument*

for $t = \mathring{t}:-1:1$

$d(t-1) = [u_{t-1},y_{t-2},u_{t-2},\ldots,y_1,u_1]$

for $d(t-1) \in d(t-1)^*$

for $y_{t-1} = \boldsymbol{0}:\mathbf{1}$

for $u = \boldsymbol{0}:\mathbf{1}$

$V_{\boldsymbol{0}|u,y_{t-1};t-1} = V_{\boldsymbol{0}|u,y_{t-1};0} + \sum_{\tau=1}^{t-1} \delta([\boldsymbol{0}|u,y_{t-1}];[y_\tau|u_\tau,y_{\tau-1}])$

$V_{\mathbf{1}|u,y_{t-1};t-1} = V_{\mathbf{1}|u,y_{t-1};0} + \sum_{\tau=1}^{t-1} \delta([\mathbf{1}|u,y_{t-1}];[y_\tau|u_\tau,y_{\tau-1}])$

$\mathcal{V}_t(u,y_{t-1},d(t-1)) =$

$= \frac{[\omega_{\boldsymbol{0}|u,y_{t-1}}+\mathcal{V}_{t+1}^*(\boldsymbol{0},d(t-1))]V_{\boldsymbol{0}|u,y_{t-1};t-1}+[\omega_{\mathbf{1}|u,y_{t-1}}+\mathcal{V}_{t+1}^*(\mathbf{1},d(t-1))]V_{\mathbf{1}|u,y_{t-1};t-1}}{V_{\boldsymbol{0}|u,y_{t-1};t-1}+V_{\mathbf{1}|u,y_{t-1};t-1}}$

end *% for* u

$[\mathcal{V}_t^*(y_{t-1},d(t-1)),u_t(y_{t-1},d(t-1))] =$

$= \min\{\mathcal{V}_t(\boldsymbol{0},y_{t-1},d(t-1)),\mathcal{V}_t(\mathbf{1},y_{t-1},d(t-1))\}$

end *% for* y_{t-1}

end *% for* $d(t-1)$

end *% for* t

Control law realization

for $t = 1:\mathring{t}$

Measure y_{t-1}

Compose $d(t-1) = [u_{t-1},y_{t-2},u_{t-2},\ldots,y_1,u_1]$

Select optimal $u_t = u_t(y_{t-1},d(t-1))$

Apply optimal u_t

end *% for* t

Derivation

The optimization is performed by sequential computation of mean value and minimization according to Algorithm 6.3.

Mean:

$$\mathcal{V}_t(u_t, d(t-1)) = E[\omega_{y_t,|u_t,\varphi_{t-1}} + \mathcal{V}^*_{t+1}(y_t, u_t, d(t-1))|u_t, d(t-1)] \tag{6.56}$$

Minimum:

$$\mathcal{V}^*_t(d(t-1)) = \min_{u_t} \mathcal{V}_t(u_t, d(t-1)) \quad \Rightarrow \quad u^*_t(d(t-1)) \tag{6.57}$$

For computation of the mean value, the predictive pdf (6.26) is needed. It is

$$f(y_t|u_t, d(t-1)) = \frac{V_{y_t|u_t,\varphi_{t-1};t-1}}{\sum_{y=\mathbf{0}}^{\mathbf{1}} V_{y|u_t,\varphi_{t-1};t-1}}, \quad \text{for} \quad y_t = \mathbf{0}, \mathbf{1}$$

where

$$V_{y_t,u_t,\varphi_{t-1};t-1} = \sum_{\tau=1}^{t-1} \delta(y_t|u_t, \varphi_{t-1}; y_\tau|u_\tau, \varphi_{\tau-1}) + V_{y_t,u_t,\varphi_{t-1};0}. \tag{6.58}$$

With this, the mean value in general step t is as stated in Proposition 6.16 with $V_{y|u,\varphi}$ according to (6.58).

After evaluating $\mathcal{V}_t(u_t, d(t-1))$ for all possible configurations of $u_t, d(t-1)$ we can chose optimal $u^*_t(d(t-1))$ for all configurations of $d(t-1)$ (which are numbers for each specific $d(t-1)$). Thus, the rest of the criterion $\mathcal{V}^*_t(y_{t-1}, u_{t-1}, d(t-2))$ is evaluated for each possible configuration of $d(t-1)$.

Starting with $\mathcal{V}^*_{\mathring{t}+1}(y_{\mathring{t}}, u_{\mathring{t}}, d(\mathring{t}-1)) = 0$ the procedure is closed.

Example

For better understanding of the structures mentioned in the previous derivation we will demonstrate an example of dual control synthesis on the control horizon with length 2 - two step ahead dual control; with 1st order model, i.e. model

$$f(y_t|u_t, y_{t-1}) = \theta_{y_t|u_t,y_{t-1}}, \tag{6.59}$$

with regression vector $\varphi_{t-1} = y_{t-1}$. Even in this simple case the computations are rather large. The general formula (**??**) for the last time $t = \mathring{t} = 2$ with prior information $d(0) = y_0$ reads

$$\mathcal{V}_2(u_2, y_1, u_1, y_0) = \frac{\sum_{y=\mathbf{0}}^{\mathbf{1}} \omega_{y|u_2,y_1} V_{y|u_2,y_1;1}}{\sum_{y=\mathbf{0}}^{\mathbf{1}} V_{y|u_2,y_1;1}}, \tag{6.60}$$

with $\mathcal{V}^*_3(y, u_2, y_1, u_1, y_0) = \mathbf{0}$ omitted and $V_{y|u_2,y_1;2} = \delta(y|u_2, y_1; y_1|u_1, y_0) + V_{y|u_1,y_0;0}$.

This expression must be evaluated for each configuration of its argument u_2, y_1, u_1, y_0 so that we could find minimum over u_t: i.e. for each configuration of y_1, u_1, y_0 compare values corresponding to $u_2 = \mathbf{0}$ and $u_2 = \mathbf{1}$ and to select minimum. Thus we obtain minimum depending on y_1, u_1, y_0 that is denoted $\mathcal{V}^*_1(y_1, u_1, y_0)$. Argument of this minima represents the optimal control $u^*_2(y_1, u_1, y_0)$.

But sooner this is said than done. The number of configurations in this simple case (2 steps ahead) is $2^4 = 16$. The possible expressions of the criterion $\mathcal{V}_2(u_2, y_1, u_1, y_0)$ for all possible configuration of u_2, y_1, u_1, y_0 are given in the

Table of $\mathcal{V}_2(u_2, y_1, u_1, y_0)$

configuration of data		criterion
$u_2, y_1, u_1, y_0:$	$\mathbf{0,0,0,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{0}}[1+V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}]+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}/\{[1+V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}]+V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}$
	$\mathbf{1,0,0,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,0,0,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}$
	$\mathbf{1,0,0,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,0,1,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}$
	$\mathbf{1,0,1,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{0}}[1+V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}]+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}/\{[1+V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}]+V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,0,1,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}$
	$\mathbf{1,0,1,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{0}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,1,0,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}$
	$\mathbf{1,1,0,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,1,0,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{1}}[1+V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}]\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+[1+V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}]\}$
	$\mathbf{1,1,0,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,1,1,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}$
	$\mathbf{1,1,1,0}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}$
$u_2, y_1, u_1, y_0:$	$\mathbf{0,1,1,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{1}}V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0}+V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}$
	$\mathbf{1,1,1,1}$	$\{\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{1}}V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{1}}[1+V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}]\}\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0}+[1+V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}]\}$

This table represents a mapping $[y_1, u_1, y_0] \to u_2$ (i.e. control variable u_2 as a discrete function of past variables y_1, u_1 and y_0).

Note: This table can be completely evaluated, as each its row depends on specific values of its variables.

For given (measured) vector $[y_1, u_1, y_0]$ the optimal control u_2 is is given as follows:

- take the two rows of the table corresponding to the vector $[y_1, u_1, y_0]$,
- evaluate them and choose the one with the less value,
- assign to u_2^* the value of u_2 of the chosen row

Note: The optimal value u_2^* cannot be chosen immediately, as we do not know, which values of the variables y_1, u_1, y_0 will be measured.

This ends the computation of the control law at the time instant 2 and we go to the time instant 1. For this (according to (**??**)) it holds

$$\mathcal{V}_{\mathbf{1}}(u_1, y_0) = \frac{\sum_{y=\mathbf{0}}^{\mathbf{1}}[\omega_{y\vert u_1,y_0} + \mathcal{V}_2^*(y, u_1, y_0)]V_{y\vert u_1,y_0;0}}{\sum_{y=\mathbf{0}}^{\mathbf{1}} V_{y\vert u_1,y_0;0}}, \tag{6.61}$$

with $\mathcal{V}_2^*(y, u_1, y_0)$ being the minimum from the previous step and $V_{y\vert u_1,y_0;0}$ the prior statistics. Similarly as for the time instant 2 we can evaluate it for all possible values of its variables u_1, y_0. Here is the corresponding

Table of $\mathcal{V}_{\mathbf{1}}(u_1, y_0)$

data config.		criterion
$u_1, y_0:$	$\mathbf{0},\mathbf{0}$	$\{[\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{0}} + \mathcal{V}_2^*(\mathbf{0},\mathbf{0},\mathbf{0})]V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0} + [\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{0}} + \mathcal{V}_2^*(\mathbf{1},\mathbf{0},\mathbf{0})]V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{0};0} + V_{\mathbf{1}\vert\mathbf{0},\mathbf{0};0}\}$
	$\mathbf{1},\mathbf{0}$	$\{[\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{0}} + \mathcal{V}_2^*(\mathbf{0},\mathbf{1},\mathbf{0})]V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0} + [\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{0}} + \mathcal{V}_2^*(\mathbf{1},\mathbf{1},\mathbf{0})]V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0} + V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$
$u_1, y_0:$	$\mathbf{0},\mathbf{1}$	$\{[\omega_{\mathbf{0}\vert\mathbf{0},\mathbf{1}} + \mathcal{V}_2^*(\mathbf{0},\mathbf{0},\mathbf{1})]V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0} + [\omega_{\mathbf{1}\vert\mathbf{0},\mathbf{1}} + \mathcal{V}_2^*(\mathbf{1},\mathbf{0},\mathbf{1})]V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{0},\mathbf{1};0} + V_{\mathbf{1}\vert\mathbf{0},\mathbf{1};0}\}$
	$\mathbf{1},\mathbf{1}$	$\{[\omega_{\mathbf{0}\vert\mathbf{1},\mathbf{1}} + \mathcal{V}_2^*(\mathbf{0},\mathbf{1},\mathbf{1})]V_{\mathbf{0}\vert\mathbf{1},\mathbf{1};0} + [\omega_{\mathbf{1}\vert\mathbf{1},\mathbf{1}} + \mathcal{V}_2^*(\mathbf{1},\mathbf{1},\mathbf{1})]V_{\mathbf{1}\vert\mathbf{1},\mathbf{1};0}\}/\{V_{\mathbf{0}\vert\mathbf{1},\mathbf{0};0} + V_{\mathbf{1}\vert\mathbf{1},\mathbf{0};0}\}$

This is similar table as from the step 2 but shorter (it represents a function of less variables) and involving nonzero term from the previous minimization $\mathcal{V}_2^*(\cdot,\cdot,\cdot)$. All the table can be evaluated immediately, as it is computed for specific values of the variables.

Note: Even the function $\mathcal{V}_2$ can be evaluated as the previous table can be evaluated and we can immediately decide about the minima of all the couples of rows on which the minimum depends.

Now, when we are at the beginning of the interval, for which the control is planned, which is the current time instant $t = 1$, we have at disposal the measured data $d(0) = y_0$. Thus we can chose the corresponding two rows for minimization (the first two for $y_0 = 0$ and the last two for $y_0 = \mathbf{1}$), and to get the optimal control u_1^* as the value of u_1 in the minimum row.

This value of u_1^* can be applied and we can measure the value of the output y_1. Thus we have the vector $[y_1, u_1, y_0]$ which is necessary for choosing the optimal control u_2^* from the step 2.

In this way we obtained the optimal control on the horizon with length 2.

6.5 Concluding remarks

1. When dealing with discrete models it is not necessary to concern about the problem of linearity. They are naturally nonlinear, nevertheless, the solutions with them are simple.

2. The crucial assumptions for easy solutions are those about the CCM problem (6.1). Most of the solutions presented here are derived just for these assumptions.

3. The simplicity of the solutions with discrete models lies in the construction of algorithms, not in computational time. The dynamic tasks, known from the continuous domain as unsolvable (like dual control), can be easily evaluated here, but can be computed in reasonable time only for short horizon (up to 10-12 steps).

Chapter 7

Appendix

7.1 Kronecker function

The **Kronecker delta function** is defined (for possibly vector arguments x and $\tilde{x}$) as follows

$$\delta(x,\tilde{x}) = \begin{cases} 1 & \text{if } x = \tilde{x} \quad \text{(for all items)}, \\ 0 & \text{otherwise.} \end{cases} \tag{7.1}$$

7.2 Dirac function

The **Dirac delta function** is defined as a functional assigning to (reasonable) functions $g(x)$ their values at zero argument. It has Riezs integral representation, [74],

$$\int_{x^*} g(x)\delta(x)\,dx = g(0) \tag{7.2}$$

where $\delta(x)$ has to be taken as generalized function [75].

7.3 Gamma function

The **gamma function** is defined by the following second type Euler integral

$$\Gamma(x) \equiv \int_0^\infty z^{x-1}\exp(-z)\,dz. \tag{7.3}$$

It is finite for real $x > 0$.

It holds:

$$\Gamma(x+1) = x\Gamma(x) \quad \text{for} \quad x \in \mathrm{R}, \tag{7.4}$$

$$\Gamma(n+1) = n! \quad \text{for} \quad n \in \mathrm{N}. \tag{7.5}$$

7.4 Beta function

The **beta function** is defined by the following first type Euler integral

$$\mathcal{B}(x,y) = \int_0^1 z^{x-1}(1-z)^{y-1}\,dz \tag{7.6}$$

It is finite for $x > 0$ and $y > 0$.
It holds

$$\mathcal{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \tag{7.7}$$

$$\mathcal{B}(x+1,y) = \frac{x}{x+y}\mathcal{B}(x,y). \tag{7.8}$$

The **incomplete beta function** is defined

$$\mathcal{B}_{x,y}(t) = \frac{\int_0^t z^{x-1}(1-z)^{y-1}dz}{\mathcal{B}(x,y)} \tag{7.9}$$

The **multivariate beta function** $\mathcal{B}(V)$, where $V \in y^* \times \psi^*$ with items $V_{y|\psi}$ is defined by the relation

$$\mathcal{B}(V) = \prod_{\psi\in\psi^*} \frac{\prod_{y\in y^*}\Gamma(V_{y|\psi})}{\Gamma(\sum_{y\in y^*} V_{y|\psi})}. \tag{7.10}$$

Recall that **Euler gamma function**, (7.3), is finite for positive real arguments.

Proof: For a pair of different "regressor" indices $\psi \neq \tilde{\psi}$, the corresponding parameters $\Theta_{\cdot|\psi}$ and $\Theta_{\cdot|\tilde{\psi}}$ are defined on unrelated domains of the integrated function. Thus, integration over them can be performed independently and the overall integral is just product of respective factors. Then, we can fix ψ and suppress it in the notation. The integrand of a single factor is $\prod_{y\in y^*}\Theta_y^{V_y-1}$ and we integrate over the set $\{\Theta_y > 0 : \sum_{y\in y^*}\Theta_y = 1\}$. For $\mathring{y} = 2$, it defines directly Euler beta function $\mathcal{B}(n,m) \equiv \int_0^1 z^{n-1}(1-z)^{m-1}\,dz$, that is related to the gamma function through the formula to be proved, [76]. In the general case, we take into account that one entry of the array Θ is on Θ^* fully determined by the other entries. Thus, the integrand becomes $(1-\sum_{\tilde{y}=1}^{\mathring{y}-1}\Theta_{\tilde{y}})^{V_{\mathring{y}}-1}\prod_{y=1}^{\mathring{y}-1}\Theta_y$ and the integration domain is $\left\{\Theta_y > 0,\ \sum_{\tilde{y}=1}^{\mathring{y}-1}\Theta_{\tilde{y}} < 1\right\}$. Then, it is sufficient to extract the term $1-\sum_{\tilde{y}=1}^{\mathring{y}-2}$ and substitute $\Theta_{\mathring{y}-1} \to \frac{\Theta_{\mathring{y}-1}}{1-\sum_{\tilde{y}=1}^{\mathring{y}-2}}$ according to Proposition 2.5, i.e. with Jacobian of transformation equal to $\frac{1}{1-\sum_{\tilde{y}=1}^{\mathring{y}-2}}$. Then, after arranging the terms of the expression, the two-dimensional beta function appears as a factor and the same pattern repeats. ♢

Example to the previous proof:
We are to compute the partial integral $I_\psi = I$ for a fix ψ

$$I = \int \dots \int \Theta_1^{V_1-1}\dots\Theta_n^{V_n-1}\ \ d\Theta_1\dots d\Theta_n,$$

with the conditions $\Theta_i > 0, i = 1,2\dots,n$ and $\sum_{i=1}^n \Theta_i = 1$, which can be rewritten into the form

$$I = \int \dots \int \Theta_1^{V_1-1}\dots\Theta_{n-2}^{V_{n-2}-1}\Theta_{n-1}^{V_{n-1}-1}\left(1-\sum_{i=1}^{n-2}\Theta_i - \Theta_{n-1}\right)^{V_n-1} d\Theta_1\dots d\Theta_{n-1}$$

We pull out the term $1-\sum_{i=1}^{n-2}\Theta_i$ from the bracket and we get

$$I=\int\ldots\int\Theta_1^{V_1}\ldots\Theta_{n-2}^{V_{n-2}}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_n}\Theta_{n-1}^{V_{n-1}}\left(1-\frac{\Theta_{n-1}}{1-\sum_{i=1}^{n-2}\Theta_i}\right)^{V_n}d\Theta_1\ldots d\Theta_{n-1}$$

After substitution $\tilde{\Theta}_{n-1}=\frac{\Theta_{n-1}}{1-\sum_{i=1}^{n-2}\Theta_i}$ with $d\tilde{\Theta}_{n-1}=\frac{1}{1-\sum_{i=1}^{n-2}\Theta_i}d\Theta_{n-1}$ we obtain

$$I=\int\ldots\int\Theta_1^{V_1-1}\ldots\Theta_{n-2}^{V_{n-2}-1}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_n-1}\Theta_{n-1}^{V_{n-1}-1}\left(1-\tilde{\Theta}_{n-1}\right)^{V_n-1}\times$$

$$\times(1-\sum_{i=1}^{n-2}\Theta_i)d\Theta_1\ldots d\Theta_{n-2}d\tilde{\Theta}_{n-1}=$$

$$=\int\ldots\int\Theta_1^{V_1-1}\ldots\Theta_{n-2}^{V_{n-2}-1}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_n-1}\left(\frac{\Theta_{n-1}}{1-\sum_{i=1}^{n-2}\Theta_i}\right)^{V_{n-1}-1}\times$$

$$\times\left(1-\tilde{\Theta}_{n-1}\right)^{V_n-1}(1-\sum_{i=1}^{n-2}\Theta_i)^{V_{n-1}}d\Theta_1\ldots d\Theta_{n-2}d\tilde{\Theta}_{n-1}=$$

$$=\int\ldots\int\Theta_1^{V_1-1}\ldots\Theta_{n-2}^{V_{n-2}-1}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_n-1}\tilde{\Theta}_{n-1}^{V_{n-1}-1}\left(1-\tilde{\Theta}_{n-1}\right)^{V_n-1}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_{n-1}}d\ldots$$

$$=\int\ldots\int\Theta_1^{V_1-1}\ldots\Theta_{n-2}^{V_{n-2}-1}\left(1-\sum_{i=1}^{n-2}\Theta_i\right)^{V_n+V_{n-1}-1}d\Theta_1\ldots d\Theta_{n-2}\times$$

$$\int\tilde{\Theta}_{n-1}^{V_{n-1}-1}\left(1-\tilde{\Theta}_{n-1}\right)^{V_n-1}d\tilde{\Theta}_{n-1}$$

which gives a univariate beta function multiplied by the same expression but one step smaller. So, evidently, the result of repeating of this procedure gives

$$I=\mathcal{B}(V_1,V_2+\ldots+V_n)\mathcal{B}(V_2,V_3+\ldots+V_n)\ldots\mathcal{B}(V_{n-2},V_{n-1}+V_n)\mathcal{B}(V_{n-1},V_n)$$

which, using (7.4), gives the proved (partial) expression

$$I=I_\psi=\frac{\Gamma(V_{1|\psi})\Gamma(V_{2|\psi})\ldots\Gamma(V_{n|\psi})}{\Gamma(V_{1|\psi}+V_{2|\psi}+\ldots+V_{n|\psi})}$$

Bibliography

[1] M.H. DeGroot, *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.

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

[3] K.J. Astrom and B. Wittenmark, *Adaptive Control*, Addison-Wesley, Reading, Massachusetts, 1989.

[4] P.E. Wellstead and M.B. Zarrop, *Self-tuning Systems*, John Wiley, Chichester, 1991.

[5] A.A. Feldbaum, "Theory of dual control", *Autom. Remote Control*, vol. 21, no. 9, 1960.

[6] A.A. Feldbaum, "Theory of dual control", *Autom. Remote Control*, vol. 22, no. 2, 1961.

[7] A. Wald, *Statistical Decision Functions*, John Wiley, New York, London, 1950.

[8] T. Bohlin, *Interactive System Identification: Prospects and Pitfalls*, Springer-Verlag, New York, 1991.

[9] M. Kárný, A. Halousková, J. Böhm, R. Kulhavý, and P. Nedoma, "Design of linear quadratic adaptive control: Theory and algorithms for practice", *Kybernetika*, vol. 21, 1985, Supplement to Nos. 3, 4 ,5, 6.

[10] V. Peterka, "Bayesian system identification", in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239–304. Pergamon Press, Oxford, 1981.

[11] M. Kárný, I. Nagy, and J. Novovičová, "Quasi-Bayes Approach to Multi-Model Fault Detection and Isolation", Tech. Rep. 1974, ÚTIA AV ČR, Praha, 1999.

[12] R.L. Keeny and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, J. Wiley, New York, 1978.

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

[14] V. Jarník, *Integral Calculus II*, Academia, Prague, 1984, (in Czech).

[15] I. Vajda, *Information Theory and Statistical Decision Making*, Alfa, Bratislava, 1982, (in Slovak).

[16] R. Bellman, *Introduction to the Mathematical Theory of Control Processes*, Academic Press, New York, 1967.

[17] S. Kullback and R. Leibler, "On information and sufficiency", *Annals of Mathematical Statistics*, vol. 22, pp. 79–87, 1951.

[18] M. Kárný, "Towards fully probabilistic control design", *Automatica*, vol. 32, no. 12, pp. 1719–1722, 1996.

[19] J. Šindelář and M. Kárný, "Dynamic decision making under uncertainty allows explicit solution", Tech. Rep. 1916, ÚTIA AVČR, POB 18, 18208 Prague 8, CR, 1997.

[20] H. Kushner, *Introduction to Stochastic Control*, Holt, Rinehart and Winston, New York, 1971.

[21] M. Loeve, *Probability Theory*, van Nostrand, Princeton, New Jersey, 1962, Russian translation, Moscow 1962.

[22] L. Berec and M. Kárný, "Identification of reality in Bayesian context", in *Computer-Intensive Methods in Control and Signal Processing: Curse of Dimensionality*, K. Warwick and M. Kárný, Eds., pp. 181–193. Birkhäuser, Boston, 1997.

[23] I.N. Sanov, "On probability of large deviations of random variables", *Matematičeskij Sbornik*, vol. 42, pp. 11–44, 1957, (in Russian), translation in Selected Translations Mathematical Statistics and Probability, I, 1961, 213–244.

[24] R. Kulhavý, *Recursive Nonlinear Estimation: A Geometric Approach*, vol. 216 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, London, 1996.

[25] R. Koopman, "On distributions admitting a sufficient statistic", *Transactions of American Mathematical Society*, vol. 39, pp. 399, 1936.

[26] M. Kárný, "Probabilistic identification and self-reproducing pdfs with application to order estimation", Tech. Rep., UTIA CSAV, Prague, 1976, in Czech, CSc thesis.

[27] E.F. Daum, "New exact nonlinear filters", in *Bayesian Analysis of Time Series and Dynamic Models*, J.C. Spall, Ed. Marcel Dekker, New York, 1988.

[28] M. Zarrop, *Experiment Design for Dynamic System Identification. Lecture Notes in Control and Information Sciences 21*, Springer, New York, 1979.

[29] A.V. Oppenheim and A.S. Wilsky, *Signals and systems*, Englewood Clifts, Jersye, 1983.

[30] R.Todd Ogden, *Essential wavelets for statistical applications and data analysis*, Birkhauser, Boston-Basel-Berlin, 1997.

[31] A. Lízr, "Linear quadratic self-tuning regulators in paper-machine control systems", in *Preprints of the 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing*, pp. 109–112. 1986.

[32] C.R. Rao, *Linear method of statistical inference and their applications*, Academia, Prague, 1987, in Czech.

[33] S. B. Jorgensen and K. M. Hangos, "Qualitative models as unifying modelling tool for grey box modelling", *International Journal of Adaptive Control and Signal Processing*, 1995.

[34] A. R. Barron and C. H. Sheu, "Approximation of density functions by sequences of exponential families", *The Annals of Statistics*, vol. 19, no. 3, pp. 1347–1369, 1991.

[35] K. Warwick and M. Kárný, *Computer-Intensive Methods in Control and Signal Processing: Curse of Dimensionality*, Birkhauser, 1997.

[36] M. Kárný, P. Nedoma, and J. Böhm, "On completion of probabilistic models", in *Preprints of the 2nd European IEEE Workshop on Computer Intensive Methods in Control and Signal Processing, CMP'96*, L. Berec, J. Rojíček, M. Kárný, and K. Warwick, Eds., pp. 59–64. ÚTIA AV ČR, Prague, August 1996.

[37] V. Kurková, "Incremental approximation by neural networks", in *Dealing with Complexity: A Neural networks Approach*, M. Kárný, K. Warwick, and V. Kurková, Eds., pp. 177–188. Springer, London, 1997, ISBN 3 540 76160.

[38] D.M. Titterington, A.F.M. Smith, and U.E. Makov, *Statistical Analysis of Finite Mixtures*, John Wiley, New York, 1985.

[39] M. Kárný and P. Nedoma, "Automatic processing of prior information with application to identification of regression model", *Kybernetika*, 2000, accepted, never finished.

[40] A. Perez and R. Jiroušek, "Constructing an intensional expert system (INES)", in *Medical Decision Making*, pp. 307–313. Elsevier, Amsterdam, 1986.

[41] M. Kárný, J. Šindelář, and M. Valečková, "Towards Bayesian pooling of probabilistic controllers", in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds., Praha, September 1998, pp. 19–24, ÚTIA AV ČR.

[42] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*, John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 2 edition, 1997.

[43] M.A. Kaashoek, J.H van Schuppen, and A.C.M Ran, *Robust Control of Linear Systems and Nonlinear Control*, vol. 2, Birkhauser, Boston, Basel Berlin, 1990, Proceedings of the International Symposium MNTS-89.

[44] P.C. Fishburn, *Utility Theory for Decision Making*, J. Wiley, New York, London, Sydney, Toronto, 1970.

[45] M. Kárný, "Algorithms for determining the model structure of a controlled system", *Kybernetika*, vol. 19, no. 2, pp. 164–178, 1983.

[46] M. Kárný and R. Kulhavý, "Structure determination of regression-type models for adaptive prediction and control", in *Bayesian Analysis of Time Series and Dynamic Models*, J.C. Spall, Ed. Marcel Dekker, New York, 1988, Chapter 12.

[47] J. Rojíček and M. Kárný, "A sequential stopping rule for extensive simulations", in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds., Praha, September 1998, pp. 145–150, ÚTIA AV ČR.

[48] E. Mosca, L. Chisci, and L. Giarré, "Sidestepping the certainty equivalence in 2-DOF adaptive control via multiple implicit identifiers", in *Mutual Impact of Computing Power and Control Theory*, M. Kárný and K. Warwick, Eds., pp. 41–64. Plenum Press, New York, London, 1993.

[49] V. Peterka, "Predictor-based self-tuning control", *Automatica*, vol. 20, no. 1, pp. 39–50, 1984, reprinted in: *Adaptive Methods for Control System Design*, M.M. Gupta, Ed., IEEE Press, New York, 1986.

[50] M. Kárný, "Adaptive systems: Local approximators?", in *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Glasgow, August 1998, pp. 129–134, IFAC.

[51] R. Kulhavý, "A Bayes-closed approximation of recursive nonlinear estimation", *International Journal Adaptive Control and Signal Processing*, vol. 4, pp. 271–285, 1990.

[52] R. Kulhavý, "Recursive Bayesian estimation under memory limitations", *Kybernetika*, vol. 26, pp. 1–20, 1990.

[53] R. Kulhavý and M. B. Zarrop, "On a general concept of forgetting", *International Journal of Control*, vol. 58, no. 4, pp. 905–924, 1993.

[54] R.K. Mehra and D.G. Lainiotis (Eds.), *System Identification – Advances and Case Studies*, Pergamon Press, New York, 1976.

[55] A. F. M. Smith and A. E. Gelfand, "Bayesian statistics without tears: a sampling-resampling perspective", *The American Statistician*, vol. 46, pp. 84–88, 1992.

[56] R. Murray-Smith and T.A. Johansen, *Multiple Model Approaches to Modelling and Control*, Taylor & Francis, London, 1997.

[57] J.S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems", *Journal of the American Statistical Association*, vol. 93, 1998.

[58] C.K. Carter and R. Kohn, "On Gibbs sampling for state space models", *Biometrika*, vol. 81, pp. 541–553, 1994.

[59] J.S. Meditch, *Stochastic Optimal Linear Estimation and Control*, Mc. Graw Hill, 1969.

[60] A.M. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[61] D.W. Clarke, C. Mohtadi, and P.S. Tuffs, "Generalized predictive control", *Automatica*, vol. 23, no. 2, pp. 137–160, 1987.

[62] O.L.R. Jacobs and J.W. Patchell, "Caution and probing in stochastic control", *International Journal of Control*, vol. 16, pp. 189–199, 1972.

[63] M. Kárný and A. Halousková, "Preliminary tuning of self-tuners", in *Lecture Notes: Advanced Methods in Adaptive Control for Industrial Application (Joint UK-CS seminar)*, K. Warwick, M. Kárný, and A. Halousková, Eds., vol. 158. Springer-Verlag, 1991, held in May 1990, Prague.

[64] P. Nedoma, M. Kárný, and J. Böhm, "Software tools for use of prior knowledge in design of LQG adaptive controllers", in *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Glasgow, August 1998, pp. 425–429, IFAC.

[65] M. Kárný, E. Sutanto, P. Ettler, J. Kadlec, A. Quinn, and L. Tesař, "Decision support tool for complex industrial processes based on probabilistic data clustering", Tech. Rep., UTIA AVCR, 1998, deliverable of Esprit project ProDaCTools, No. 25729.

[66] M. Kárný and A. Halousková, "Pretuning of self-tuners", in *Advances in Model-Based Predictive Control*, D. Clarke, Ed., pp. 333–343. Oxford University Press, Oxford, 1994.

[67] V. Peterka, "Digital control of processes with random and uncertain characteristics"', 1975, DrSc thesis, 305 pp. in Czech.

[68] M. Valečková and M. Kárný, "Estimation of Markov chains with reduced parametrisation", in *Preprints of the 2nd European IEEE Workshop on Computer-Intensive Methods in Control and Signal Processing, CMP'96*, L. Berec, J. Rojíček, M. Kárný, and K. Warwick, Eds., pp. 135–140. ÚTIA AV ČR, Prague, 1996.

[69] M. Valečková and M. Kárný, "Markov chains in non-linear identification: Bayesian clustering for reduced parametrisation", in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds., Praha, September 1998, pp. 273–278, ÚTIA AV ČR.

[70] M. Valečková, M. Kárný, and E. L. Sutanto, "Bayesian M-T clustering for reduced parametrisation of Markov chains used for nonlinear adaptive elements", in *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Glasgow, August 1998, pp. 381–386, IFAC.

[71] P.R. Kumar, "A survey on some results in stochastic adaptive control", *SIAM J. Control and Applications*, vol. 23, pp. 399–409, 1985.

[72] M. Kárný, L. Pavelková, A. Halousková, and P. Nedoma, "Estimation of approximate Markov chains", in *IFAC Symposium ACASP'95*, Cs. Bányász, Ed., pp. 317–322. Budapest, 1995.

[73] E. Nováková and Miroslav Kárný, "New approach to control design for Markov chains", in *European Control Conference. /CD-ROM/*, G. Bastin and M. Gevers, Eds., p. 221 kB. EUCA, Brussels, July 1997.

[74] R. Taylor, *Introduction to Functional Analysis*, Academia, Prague, 1973, (Czech translation).

[75] V.S. Vladimirov, *Generalized Functions in Mathematical Physics*, Mir Publishers, Moscow, 1979.

[76] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1972.