Fully Probabilistic Dynamic Decision Making

Miroslav Kárný

Adaptive Systems Department, ÚTIA AVČR, v.v.i.

supported by AS http://www.utia.cas.cz/AS & GAČR, GA16-09848S

December 11, 2018

Topic and Presentation Approach

Topic: Dynamic Decision Making (DM) Under Uncertainty

The lecture aims to present an applicable prescriptive DM theory.

The theory designs the strategy that uses knowledge to opt an action, which meets as much as possible a given DM aim concerning the behaviour of the closed loop formed by the strategy and the system.

The design faces

- an inevitable uncertainty,
- an incomplete knowledge,
- a limited ability to design and use the optimal decision strategy,
- dynamic dependencies of behaviour constituents.

Theories, Domains, Methodologies and Vocabulary

The theory deals with and touches of an extreme range of

- theories: control [6, 115], artificial intelligence [22, 142], pattern recognition [37, 139], economics [155, 171], social sciences [5, 50], . . .
- methodologies and techniques: cybernetics [166], statistical DM [14, 35, 144, 164], fuzzy DM [161], domain-specific solutions [146], econo-physics [119], universal artificial intelligence [59],. . .
- synonyms: action ↳ vs. decision vs. input ↳; output ↳ vs. response; strategy vs. policy [122] vs. control law [6] vs. decision function [164].

What Is Specific to this Course?

The solution extends Bayesian DM [14, 35, 144, 164]. It is specific by its:

- stress on dynamic decision making requiring a design ↳ of strategies generating sequences of actions,
- systematic use of probabilistic description to all basic DM elements ↳,
- top down presentation of a common logical structure of all DM tasks,
- constructive, problem-driven, approach,
- vocabulary combining terms from various domains.

Audience: Interests, Skills & Take Home Message

The course:

- touches deep DM concepts and targets at Master & PhD students,
- invites researchers interested in criticism, development and non-trivial applications of DM theories,
- is not mathematical text in spite of the (miss)used mathematics,
- relies on knowledge of basics of algebra, analysis & probability theory.

Ideally, the audience should

- gain the ability to specify and quantify the key parts of DM problems,
- recognise multitude of open problems related to DM.

Introduction

Scheme of Overall Decision Making Problem

external influences

SYSTEM: the World part of interest including sensors & actuators

observations | actions

STRATEGY: mapping of accumulated knowledge on actions

observations | strategy

DESIGN: mapping of knowledge, aims & constraints on strategy

observations | quantified knowledge, preferences & constraints; design algorithms

support of the course theory

DESIGNER: mapping user's DM elements on formal ones

noisy, vague knowledge, preferences & constraints incompletely described in domain specific terms

USER: decision maker with a real-life DM task served by designer

The Basic Elements on which DM Operates

aim ↳	the primary reason why a DM is addressed
system ↳	the external World part, which influences DM outcomes
action ↳	the option selected by decision maker for the aim reaching
knowledge ↳	anything what can be used for the action choice
ignorance ↳	the considered but unknown part of DM
uncertainty ↳	the part influencing DM but uninfluenced by actions while preventing to know their consequences
constraint ↳	anything constraining freedom of DM
dynamics ↳	anything making DM consequences manifest gradually

Examples of DM – Course Attending

Example 1 (Enrolling this Course)

aim	*to learn something interesting, to get credits, . . .*
system	*the teacher, school mates, the personal future*
action	*{enter, do not enter} this course*
knowledge	*the course syllabus, a gossip of older students*
ignorance	*the real content of the course*
uncertainty	*the course simplicity, the intellectual state of the teacher, the personal ability to perceive the course content*
constraint	*the spent time and brain effort, the schedule*
dynamics	*the static (one-shot) DM with long-term consequences like the lost time, the usefulness degree in the future life. . .*

Examples of DM - Table-Length Estimation

Example 2 (Estimation of Table Length)

aim	*to provide an information serving for the table displacement using either a small lift or staircase*
system	*the table and space around it*
action	*an upper estimate of the table length*
knowledge	*a personal guess, available observations*
ignorance	*the true length of the table*
uncertainty	*measurement errors*
constraint	*the time spent on DM, the measuring-tape precision*
dynamics	*one-shot decision with consequences like the lost time and energy on measurements or a trial table displacement.*

Example 3 (Control of Metal Thickness)

aim$_{\downarrow}$: *to get the metal of a constant thickness*

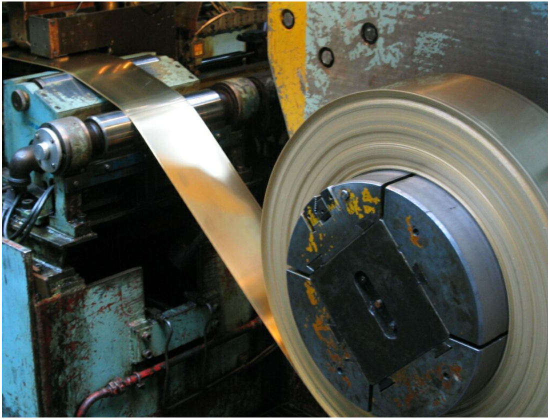
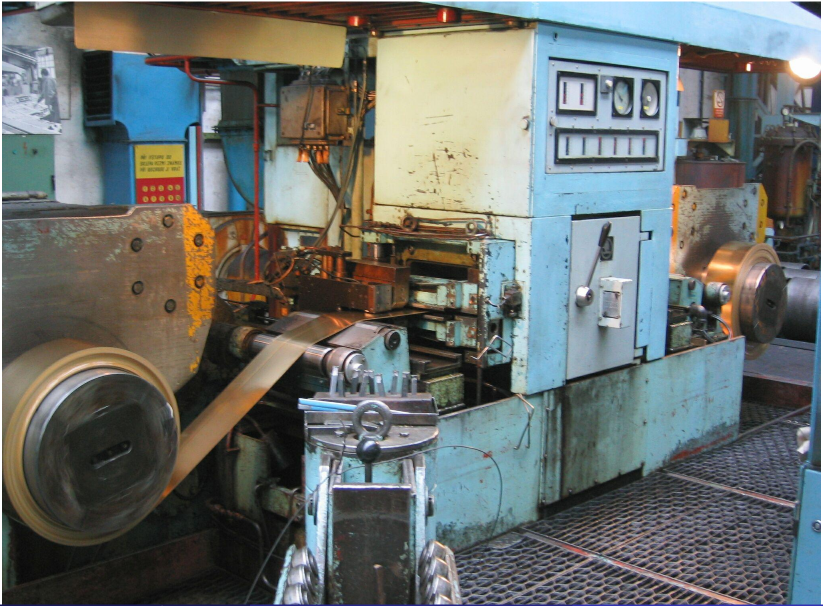


Figure : Rolled Metal Strip

System: Rolling Mill



Actions: rolling force, speed & tensions . . . ∼ 300 options

Other DM Elements Related to Rolling

knowledge	observations – input-output thickness, speeds, tensions, etc. (40 channels) – and personal knowledge (at least 6 months' learning)
ignorance	detailed properties of the mill and of the rolled material (the mill more hammers than rolls)
uncertainty	measurement errors, eccentricity of rolls, responses of actuators to commands, mill aging, etc.
constraint	on forces, tensions & their changes; control period $\approx$ 10 ms precision of the sensors
dynamics	time-delay of input-thickness measurement and a change of rolling force ($\approx$ 20 control periods), actuator dynamics.

Examples of DM – Traffic problem

Example 4 (Control of the Traffic in a Town)

aim ↳ *to exploit fully the available capacity of town roads,*
for instance, to minimise the average travelling time,
or lengths of queues or transportation costs,. . .

System: Traffic Flow in a Town Region

Figure : Part of a Traffic System

Action: Transformation of the System

. . . to the System

Figure : Radical Solution of the Traffic Problem

Figure : Operators of a Traffic System

Other DM Elements

knowledge	off-line statistical data, observations of the traffic intensity & road occupancy, visual inspections, Fig. 7
ignorance	the car flow evolving over the space & time, queue lengths
uncertainty	measurement errors, un-measured number of parking cars, weather, congestions, behaviour of drivers, etc.
constraint	the available capacity of the transportation system, priorities of public transportation, safety regulations, information systems, complexity of evaluations, . . .
dynamics	the traffic is a random spatially-distributed process, minor changes at a cross-road have far reaching influence

Recall personal experience with a strong influence of "green wave" and its violation, accident consequences, priorities of state-guests, etc.

Exercises 1 on DM Elements

Exercises 1 (On DM Elements)

1. *Think over your personal example of dynamic DM.*
2. *Specify its basic DM elements.*
3. *Does you miss some DM element?*
4. *Is to be aim ↳ scalar?*
5. *Who specify the system ↳?*
6. *Why the World is not taken as the system?*
7. *Think about the relation ignorance ↳ and uncertainty ↳.*

Basic Notions

On Notions, Notations & Conventions

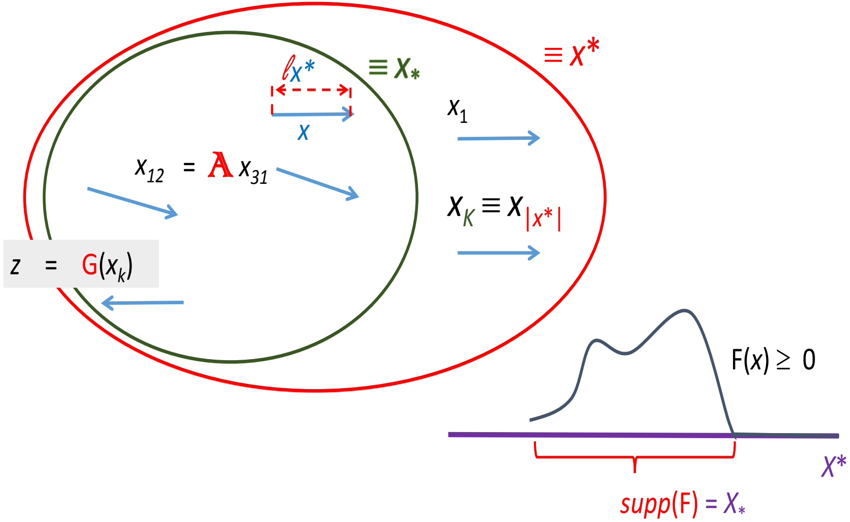
This part summarises basic notions and notations used throughout.

- The conventions listed here are mostly followed in this work. If an exception is necessary, it is introduced at the place of its validity.
- The respective notions are introduced within the text, when they are used for first time. They are emphasised and often cross-referred.
- The presentation starts with general conventions followed by briefly characterised basic notions. Then, the used vocabulary is commented.

Basic Notations and Conventions

- *defining equality* $\equiv$ is the equality by definition.
- *set* $X^\star$ denotes the range of X. Subset $X_\star$ is a part of $X^\star$.
- *cardinality* $|X^\star|$ is the number of members in the set $X^\star$.
- *vector length* $\ell_{X^\star}$ means the number of entries in the vector X.
- *mappings* are marked by sf fonts and matrices by bb fonts e.g. $\mathbb{A}$.
- *quantity* is a multivariate mapping equivalent to a random variable, [138]. Avoiding it stresses that probability serves us as a DM ↳ tool.
- *realisation* is a quantity value for a fixed argument. The quantity ↳ and its realisation ↳ are distinguished by the context only.
- *pd* probability density, f, is Radon-Nikodým derivative [137] of a probability measure. An argument name determines meaning of the pd.
- *support* $\mathrm{supp}\,[\mathsf{f}(X)]$ is the subset $X_\star$ of $X^\star$ on which $\mathsf{f}(X) > 0$.
- *expectation* of $\bullet(X)$ is $\mathsf{E}[\bullet] \equiv \mathsf{E}_{\mathsf{f}}[\bullet] \equiv \int_{X^\star} \bullet(X)\mathsf{f}(X)\,\mathrm{d}X$.

Basic Notation Graphically



Overall Decision Making Problem

external influences

SYSTEM: the World part of interest including sensors & actuators

observations — actions

STRATEGY: mapping of accumulated knowledge on actions

observations — strategy

DESIGN: mapping of knowledge, aims & constraints on strategy

the discussed support of the course theory

observations — quantified knowledge, preferences & constraints; design algorithms

DESIGNER: mapping user's DM elements on formal ones

noisy, vague knowledge, preferences & constraints incompletely described in domain specific terms

USER: decision maker with a real-life DM task served by designer

DM Concerns Interactions

DM concerns interaction of a system↳ and a decision maker↳

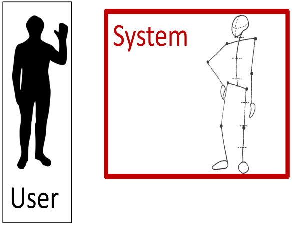
- *system* is a part of the World considered by a decision maker↳ who wants either describe or influence it.
  - The system↳ is specified with respect to the aim↳ of the user↳.
  - The penetrable system boundaries are implied by the decision task.
- *user* is the person or the group, which provides the domain-specific formulation of the solved decision making task.
- *designer* is the person or the group, which selects the strategy↳.
- *decision maker* is the person, the group or the device choosing action↳s.

For simplicity, the user↳, designer↳,decision maker↳ are called it.

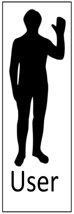
The considered (compound) decision maker↳ is also the user↳ & designer↳.

System Specification Depends on DM Task

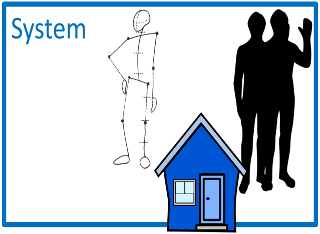
DM: invite to a cinema or pub
with the aim to avoid the bore



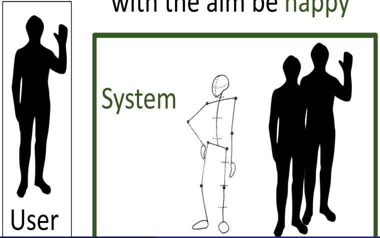
DM: propose not propose the marriage



with the aim be happy and rich



DM: propose or not the marriage
with the aim be happy



Why to Distinguish Roles? Roles of the Course Audience?

The user↳, designer↳ and decision maker↳ may have

- other, often unsaid, aim↳s,
- inherent additional knowledge↳ and/or constraint↳s.

These DM participants have to cooperate to overcome differences in

- knowledge↳, constraint↳s and aim↳s
- language (vocabulary, meaning of the same words,. . .)

This course attendants primarily prepare themselves to be

- the developers of the DM theory and its tools
- designer↳s.

The theory is helpful in other roles, too.

Interface to Real World

The further talk about other DM elements and DM theory relies on:

Agreement 1 (Connections to Reality)

All physical connections of DM elements↳ to World as

- *sensors,*
- *transmission lines,*
- *actuators,*
- *...*

are taken here as a part of the system↳ dealt with.

Thus, all considered quantities and mappings are mathematical entities living in an abstract calculating machine.

Action, Aim, Behaviour & Sequences

- *action* is an irreversible decision with the value $A \in A^{\star}$ directly chosen by the decision maker ↳ for reaching its aim ↳.
- *behaviour* , $B \in B^{\star}$, consists of all quantities and their of realisation ↳s considered by the decision maker ↳ in its DM task.
- *aim* specifies the desired behaviour ↳ of the closed decision loop formed by the decision strategy ↳ and system ↳.

We deal with time-dependent objects up to a time horizon ↳.

- *timed quantity* X_t is labelled by time instant $t \in t^{\star} \equiv \{1, \ldots, h\}$.
- *horizon* $h \leq \infty$ concerns decision, prediction, control, design. . .
- *time index* $X_{t;i}$ at i-th entry of X_t is the first one and separated by **;** .
- *sequence* $X^{\underline{t}:\overline{t}} \equiv (X_t)_{t=\underline{t}}^{\overline{t}}$, $X^t \equiv X^{1:t}$.

  $X^{\underline{t}:\overline{t}}$ adds nothing to prior knowledge ↳ if $\overline{t} < \underline{t}$.

Strategy, Decision Rule & Feedback

The prescriptive theory should support any decision maker↳ in its targeted choice of action↳s for any behaviour↳ realisation↳, i.e. in the design↳ of:

- *strategy* $\equiv$ decision strategy S is a sequence of mappings from the behaviour set $B^\star$ to the action set $A^\star$, $\mathsf{S} \equiv (\mathsf{S}_t : B^\star \to A^\star_t)_{t\in t^\star}$.
- *decision rule* S_t is a mapping $\mathsf{S}_t : B^\star \to A^\star_t$ that assigns the action↳ $A_t \in {A_t}^\star$ to the behaviour↳ $B \in B^\star$ at time t.

  The notation neglects that the randomised decision rule S_t maps behaviours on subsets of $A^\star_t$.

If the used strategy influences the system↳, then it makes it via

- *feedback*, which is the process in which strategy↳ maps behaviour on an action influencing the behaviour, influencing the next action etc.

Knowledge & Ignorance

An applicable strategy uses knowledge & ignores ignorance realisations

- *knowledge* $\equiv$ decision knowledge, $\mathbf{K}_{A^\star} \in \mathbf{K}_{A^\star}{}^\star$, is the behaviour part available to the decision maker for the choice of the action $A \in A^\star$.

Shortcut $\mathbf{K}_{t-1} \equiv \mathbf{K}_{A_t^\star}$ reflects an inevitable delay in gain & use of knowledge.

- *ignorance* $\equiv$ decision ignorance, $\mathbf{G}_{A^\star} \in \mathbf{G}^\star_{A^\star}$, is the behaviour part unavailable to the decision maker for the choice of the action $A \in A^\star$.

  Often, ignorance contains yet unobserved (future) quantities.

Shortcut $\mathbf{G}_t = \mathbf{G}_{A_t^\star}$ attaches the inevitable time delay to strategy.

Example 5 (Knowledge, Action & Ignorance in Parameter Estimation)

The observed data $D \equiv \mathbf{K}_{\hat{\Theta}^\star}$ is the knowledge for selecting the action $A = \hat{\Theta}$, which is an estimate of an unknown quantity $\Theta \in \Theta^\star$.
The estimated Θ is in the ignorance $\mathbf{G}_{\hat{\Theta}}$ of an estimate $\hat{\Theta} \equiv A$.

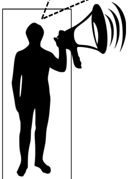
The Action Structures and Changes the Behaviour

DM: invite to a cinema or pub with the aim to avoid the bore

K* = {  is {nice,teribble} if the

film is {horror, love story}

pub is {empty, crowded}}

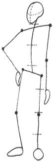


User

Lets go to ∈

A*= {Lucerna, U Fleků}

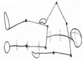
System



G* = {{ bore, fun} in {Lucerna, U Fleků}}

B* = (G*,A*,K*) internally evolves: before acting, by acting,

when spending the time at the chosen place

B is uncertain due N* = {film type, pub population,  mood}

Knowledge Enriching, Ignorance Shrinking & Observation

- The choice of an action $A \in A^{\star}$ splits behaviour $B \in B^{\star}$ into ignorance $\mathbf{G}_{A^{\star}}$ and knowledge $\mathbf{K}_{A^{\star}}$, $B = (\mathbf{G}_{A^{\star}}, A, \mathbf{K}_{A^{\star}})$.

  This is shortcut of the "choice mapping" $\mathrm{CH}_{A^{\star}}(B) \equiv (\mathbf{G}_{A^{\star}}, A, \mathbf{K}_{A^{\star}})$.

- The timed action choice $A_t \in A_t{}^{\star}$ splits behaviour $B \in B^{\star}$ into ignorance $\mathbf{G}_t \equiv \mathbf{G}_{A_t{}^{\star}}$ & knowledge $\mathbf{K}_{t-1} \equiv \mathbf{K}_{A_t{}^{\star}}$,

$$\text{behaviour} = B = (\mathbf{G}_t, A_t, \mathbf{K}_{t-1}) = (\text{ignorance, action, knowledge}). \quad (1)$$

- *observation* $\Delta_t \in \Delta_t{}^{\star}$ enriches knowledge; it consists of quantities in the ignorance $\mathbf{G}_t$ of A_t & in the knowledge $\mathbf{K}_t$ of A_{t+1}.

Remark 1 (Knowledge, Ignorance & Observation)

- *For a fixed behaviour set, knowledge enriches & ignorance shrinks.*
- *The observation Δ_t is often the system output measured at time t.*
- *The void observation Δ_t preserves $\mathbf{K}_t = \mathbf{K}_{t-1}$, $\mathbf{G}_{t+1} = \mathbf{G}_t$.*

Informationally Causal Decision Rules and Strategies

Agreement 2 (Informational Causality)

The applicable decision rules have to be informationally causal.

- *causal decision rule* S_t *maps knowledge* $\mathbf{K}_{t-1} \in \mathbf{K}^{\star}_{t-1}$ *on action* $A_t \in A^{\star}_t$
- *causal strategy* $\mathsf{S} \equiv (\mathsf{S}_t : \mathbf{K}^{\star}_{t-1} \to A^{\star}_t)_{t \in t^{\star}}$ *consists of causal rules.*

The term "causal" is dropped as the causal strategies are only used.

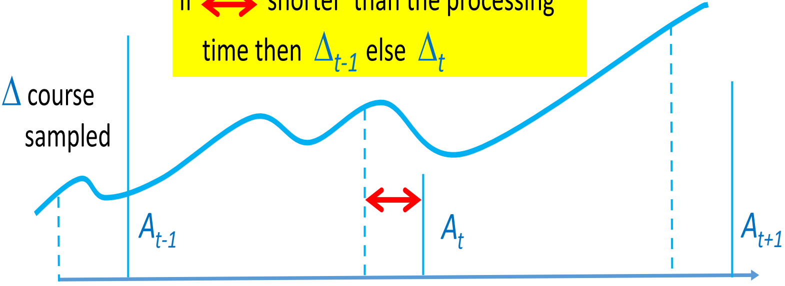
Example 6 (Some Causal Decision Rules)

- *estimator is a causal decision rule* $\mathsf{S}_t : \mathbf{K}^{\star}_{t-1} \to \hat{\Theta}^{\star}_t$ *assigning an estimate* $\hat{\Theta}_t$ *of an unknown quantity* $\Theta_t \in \Theta_t{}^{\star}$ *to the knowledge* $\mathbf{K}_{t-1}$.
- *P-controller is a causal decision rule generating inputs proportional to the deviation of the controlled variable from the targeted value.*
- *classifier is a causal decision rule assigning class labels to objects using knowledge available about them.*
- *tax system is a collection of causal decision rules assigning tax values to the knowledge available about the tax payer.*

Information Causality Does Not Imply Causality!

$$G_{t-1}\, A_{t-1}\, K_{t-2} \xrightarrow[\text{ignorance shrinking}]{\text{knowledge enriching}} G_t\, A_t\, K_{t-1}$$

if $\leftrightarrow$ shorter than the processing time then Δ_{t-1} else Δ_t



t-1

at time t action A_t is chosen & applied

Our Aim & Topic Stressed

The presented prescriptive theory helps the compound decision maker↳ to select the proper – from its view-point – strategy↳ selecting a good action↳ among available options. Ours description uses the notions

- *design* selects a decision rule↳ or a strategy↳.
- *dynamic design* chooses a strategy↳.
- *dynamics* means any circumstance that calls for the dynamic design↳.

Within this text

- *DM* is dynamic DM covering both the design↳ and use of a strategy↳.

The designer↳s form primary audience of this text and "we" means

We designer↳s serve to user↳ whose aim↳ should be reached in the best way.

Recall: decision maker↳, designer↳ and user↳ may mean single person.

Coin Tossing as Overall DM Problem

wind, dirt on coin

SYSTEM: tossed coin with tossing hand & recording eyes

fallen coin side

tossing

STRATEGY: putting coin head or tail on tossing hand & tossing

fallen coin side

strategy

DESIGN: mapping of knowledge, aims & constraints on strategy

fallen coin side

physics & experience, fallen heads,
number of games; filtering & FPD

DESIGNER: mapping user's DM elements on coin model & ideal pd

noisy, vague knowledge of tossing process, desired fallen
heads, game rules all described verbally & qualitatively

USER: player in unfair coin-tossing game served by designer

Engineers Should Approach DM in Its Entirety

An engineer should cover in a systematic algorithmic way as many activities in the overall DM problem, Fig. 29, as possible and to select:

- *admissible strategy* $\mathsf{S} \equiv (\mathsf{S}_t)_{t \in t^\star} = \mathsf{S}^h$, which is causal strategy $\mathsf{S} \equiv (\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to A^\star_t)_{t \in t^\star}$ and respects given constraints.
- *constraint* is any circumstance restricting the set of strategies $\mathsf{S}^\star$ among which the designer can choose.
- *physical constraints* limit sets of possible actions $({A_t}^\star)_{t \in t^\star}$.
- *informational constraints* define the knowledge $\mathbf{K}_{t-1}$ used & the ignorance $\mathbf{G}_t$ unused for the choice of the action $A_t \in {A_t}^\star$, $\forall t \in t^\star$.

Practically Admissible Strategies

An applicable prescriptive DM theory is to provide the best

- *practically admissible strategy* is an admissible strategy ↳ that respects constraint ↳ limiting the complexity of the DM.

The complexity refers to

- the computational resources — computational time and memory — spent at the design ↳ and use stages,
- the deliberation effort of the compound decision maker ↳, i.e. ,designer ↳, user ↳, strategy implementing persons,. . .

Majority of the complexity-affected problems are computationally hard in terms of computer sciences [51]. Its intuitive understanding suffices to us.

The lack of methodology how to select the best practically admissible strategy is the common drawback of known prescriptive theories!

Uncertainty is the Key Cause of DM Complexity

- *uncertainty* occurs if the strategy$_\llcorner$ S does not determine uniquely the behaviour$_\llcorner$ $B \in B^\star$. Then, for a fixed S, there is a mapping
$$\mathsf{W}(\mathsf{S}, \cdot) : N^\star \to B^\star. \tag{2}$$
The argument $N \in N^\star$ is called uncertainty.
- *uncertain behaviour* arises if the uncertainty set $N^\star \neq \emptyset$.

Remark 2 (On Uncertainty)

- *Ideally, the properly modelled uncertainties should not be influenced by the chosen strategy. Only then, it is possible to recognise the influence of the optional strategy and to select the best strategy.*
- *The mapping* W *is used later on. Its deeper study leads to conjecture (uninspected here) that DM should use non-commutative probability operating on quantum logic [38]. This probability is used in micro-world of quantum mechanics [124] but there are strong indicators that it models better cognition and DM in macro-world [25]*

DM View Wipes off Importance of Uncertainty Origins

Possible uncertainty origins are

- randomness modelled by probability theory [138]
- vagueness modelled by fuzzy theory [169]
- incompletely specified beliefs [168]
- problem complexity [45]
- incomplete knowledge [144]
- noise, imprecision, non-robustness, chaotic behaviour, . . .

Any of them prevents to specify uniquely the realised behaviour ↳ for a given strategy ↳ and represents the same obstacle for DM.

- An unknown quantity ↳ $\Theta \in \Theta^\star$ in behaviour ↳ $B \in B^\star$ makes it uncertain in the operationally same way as an unobserved noise.

Control Exemplifies Domain-Dependent Vocabularies

Control theory [7, 18, 42, 165] addresses DM but differs in vocabulary

- *input* $\equiv$ system input, $U \in U_t^\star$, is an action influencing ignorance $\mathbf{G}_t$.
    - A manipulated valve position influencing a fluid flow is the input.
    - An estimate $\hat{\Theta}$ of an unknown quantity Θ is an action, which is not a system input as it describes the system but has no influence on it.
- *output* $\equiv$ system output $Y \in Y^\star$ is a part of observation that informs the decision maker about the behaviour.

  The possibility to manipulate the quantity directly or indirectly distinguishes input from output. For instance, the pressure
    - applied to a system is an input
    - measured on a heated system is the output
- *controller* is a strategy assigning the input U_t to knowledge $\mathbf{K}_{t-1}$.

  The P-controller, given a proportionality constant p, is a causal control strategy $(\mathbf{K}_{t-1}^\star \equiv Y_{t-1}^\star \to U_t^\star : U_t = -pY_{t-1})_{t \in t^\star}$.

Artificial intelligence is well related to DM in [59].

Exercises 2 (On Strategy and Dynamics)

1. *Why is necessary to work with strategy and not just with a sequence of actions?*
2. *Select your favourite domain and try to find relationships between its and ours vocabularies.*
3. *Select your personal example of dynamic DM and delimit all introduced notions on it.*
4. *Think over how you would solve your DM problem in its entirety.*
5. *Specify where your technical knowledge is insufficient for a systematic solution of your dynamic DM problem.*

Formalisation of DM Under Uncertainty

Quest for Prescriptive Theory of DM Under Uncertainty

The targeted prescriptive DM theory should

- help the user ↳ to select action ↳ s — describing or influencing foreseeable realisations of behaviour ↳ and uncertainty ↳ — so that user's preferences, formulated as desirable behaviours, are met as closely as possible;
- be as widely applicable and universal ("objective") as possible.

Content of this part

- The presentation describes a general way how to understand and face uncertainty ↳ that causes incomplete ordering of strategies even when the preferential ordering ↳ of possible behaviours is complete.
- The proposed complete ordering ↳ of strategies, harmonised with the preferential ordering ↳ of behaviours, makes the theory prescriptive.

Ordering of Behaviours

A DM design makes sense if the user prefers some behaviours.

- *preferential ordering* of the user is an ordering $\preccurlyeq_{B^\star}$ of behaviours $B \in B^\star$. It is a relation $\preccurlyeq_{B^\star}$ of pairs $\left({}^aB, {}^bB\right) \in B^\star \times B^\star$

$$
{}^aB \preccurlyeq_{B^\star} {}^bB \quad \text{reads} \quad {}^aB \text{ is preferred against } {}^bB. \tag{3}
$$

Recall A relation is a subset of Cartesian product of the involved sets.

- The desirable consistency of preferences restricts $\preccurlyeq_{B^\star}$ to be transitive

$$
\left({}^aB \preccurlyeq_{B^\star} {}^bB \wedge {}^bB \preccurlyeq_{B^\star} {}^cB\right) \Rightarrow {}^aB \preccurlyeq_{B^\star} {}^cB. \tag{4}
$$

Real DM often violates (4). A prescriptive theory should insist on it.

Completion of Behaviour Ordering

- The preferential ordering $\preccurlyeq_{B^\star}$ is often a partial ordering as the decision maker is unable or unwilling to compare all behaviours.

  An ordering $\preccurlyeq_{B^\star}$ can always be completed by extending behaviour by
- *preferential pointer* labels alternative completions of $\preccurlyeq_{B^\star}$.

The preferential pointer is in ignorance & allows the assumption that $\preccurlyeq_{B^\star}$ is:

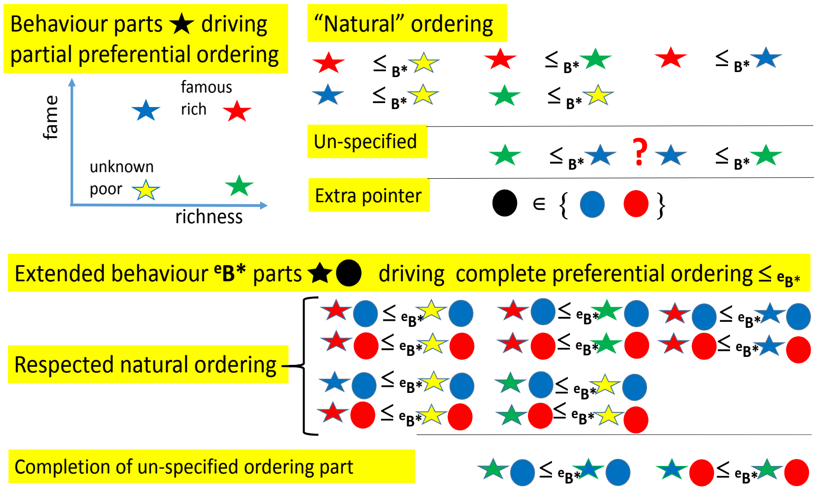
- *complete ordering* $\preccurlyeq_{B^\star}$ of the behaviour set $B^\star$ enables preferential comparison of all behaviour pairs ${}^{a}B,\ {}^{b}B \in B^\star$

$$\text{either} \quad {}^{a}B \preccurlyeq_{B^\star} {}^{b}B \quad \text{or} \quad {}^{b}B \preccurlyeq_{B^\star} {}^{a}B.$$

- The complete ordering $\preccurlyeq_{B^\star}$ induces the strict preferential ordering $\prec_{B^\star}$ and the preferential equivalence $\approx_{B^\star}$

$$\begin{aligned} {}^{a}B \prec_{B^\star} {}^{b}B &\Leftrightarrow {}^{a}B \preccurlyeq_{B^\star} {}^{b}B \wedge \neg({}^{b}B \preccurlyeq_{B^\star} {}^{a}B) \\ {}^{a}B \approx_{B^\star} {}^{b}B &\Leftrightarrow {}^{a}B \preccurlyeq_{B^\star} {}^{b}B \wedge {}^{b}B \preccurlyeq_{B^\star} {}^{a}B \end{aligned} \tag{5}$$

Example of Preferential-Ordering Completion



Exercises 3 (On Preferences)

1. *Think over your case of DM and specify your preferential ordering* ↳.
2. *Make your preferential ordering* ↳ *complete.*
3. *Why the ordering transitivity is so important?*
4. *Do you know real-life example of the transitivity violation?*
5. *Does exist a preferential ordering* ↳, *which operates on a proper extension of the behaviour set?*
6. *How you would handle the preferential ordering* ↳ *if your answer to the previous question is affirmative?*

Numerical Representation of Preferential Ordering

The aimed algorithmic solution of DM↳ tasks is enabled by a numerical representation of the preferential ordering↳ $\preceq_{B^\star}$ via a real-valued loss↳ Z.

- *loss* $\mathsf{Z} : B^\star \to [-\infty, \infty]$ quantifies the degree of the aim achievement if it is strictly isotonic with the preferential ordering↳, i.e.

$$\begin{aligned} {}^aB \prec_{B^\star} {}^bB \quad &\Leftrightarrow \quad \mathsf{Z}({}^aB) < \mathsf{Z}({}^bB) \\ {}^aB \approx_{B^\star} {}^bB \quad &\Leftrightarrow \quad \mathsf{Z}({}^aB) = \mathsf{Z}({}^bB). \end{aligned} \tag{6}$$

Remark 3 (Ordering, Loss & Reward)

- *The loss use motivates the adopted notation for preferential ordering↳. The smaller is the loss↳ value the better behaviour is.*
- *The loss↳ measures a posteriori the quality of each realisation↳ B.*
- *The used maximised reward formally coincides with a negative loss.*

Example 7 (Estimation of Table Length, cf. Example 2)

The aim is to provide an information for a table displacement via a lift of $\bar{L}$-depth or staircase. A prior guess $\hat{L}_0$ completes the knowledge $\mathbf{K} \equiv (\bar{L}, \hat{L}_0)$.

The action A is an upper bound $\hat{L}$ on the unknown table length L – the action ignorance $\mathbf{G}$. *The ordering $\preceq_{B^\star}$ is represented by the loss* $\mathsf{Z}(B) =$

$$\mathsf{Z}(\mathbf{G}, A, \mathbf{K}) = \begin{cases} 1 & \text{if} \quad \max(L, \hat{L}) < \bar{L} \quad \textit{lift is successfully used} \\ 2 & \text{if} \quad \min(L, \hat{L}) \geq \bar{L} \quad \textit{staircase is properly used} \\ 3 & \text{if} \quad \hat{L} < \bar{L} \leq L \quad \textit{lift is tried, stairs is used} \\ 4 & \text{if} \quad L < \bar{L} \leq \hat{L} \quad \textit{stairs is unnecessarily used} \end{cases} \tag{7}$$

Remarks

- ${}^{a}B \approx_{B^\star} {}^{b}B$ iff they only differ in prior guesses ${}^{a}\hat{L}_0$, ${}^{b}\hat{L}_0$.
- Users may well differ in valuations (ordering) of rows 3, 4 in (7).

Existence of Non-Unique Numerical Representation

- The real line $[-\infty, \infty]$ with the usual, transitive and complete, strict ordering $<$ has the topology generated by open intervals [24].

  The real line has a countable $<$-dense subset of rational numbers:

  for any $a < b$ exists a rational $c \in (a, b) = \{x : a < x < b\}$.
- The intuitively appealing condition that a loss$_↳$ may exist if the topology $\preccurlyeq_{B^\star}$, is not richer than that of $<$ is proved in [34, 46].

Proposition 1 (Existence of a Loss Representing $\preccurlyeq_{B^\star}$)

A loss$_↳$ Z representing $\preccurlyeq_{B^\star}$ (6) exists iff a countable $\prec_{B^\star}$-dense set exists.

Remarks

- The loss$_↳$ $Z(\cdot)$ is not unique as any increasing real function U of a real argument provides $U(Z(\cdot))$ representing the same ordering $\preccurlyeq_{B^\star}$ as $Z(\cdot)$.
- The representation freedom can be meaningfully restricted, for instance, by requiring it to be continuous in the $\preccurlyeq_{B^\star}$-topology, [46].

DM design selects the "best" strategy $^{o}\mathsf{S}$ with respect to a consistent, i.e. transitive, ordering $\preccurlyeq_{\mathsf{S}_\star}$. The choice is made among

- *compared strategies* form a subset $\mathsf{S}_\star \subset \mathsf{S}^\star$ of admissible strategies.
- *ordering of strategies* $\preccurlyeq_{\mathsf{S}_\star}$ is interpreted for comparable $^{a}\mathsf{S},\ ^{b}\mathsf{S} \in \mathsf{S}_\star$

$$^{a}\mathsf{S} \preccurlyeq_{\mathsf{S}_\star} {}^{b}\mathsf{S} \quad \Leftrightarrow \quad {}^{a}\mathsf{S} \text{ is better than } {}^{b}\mathsf{S} \tag{8}$$

A DM theory can claim to be prescriptive if it is universal. Primarily, it

- allows to take any subset $\mathsf{S}_\star$ of $\mathsf{S}^\star$ containing at least one comparable pair $^{a}\mathsf{S} \neq {}^{b}\mathsf{S}$ as the set of compared strategies,
- defines $\preccurlyeq_{\mathsf{S}_\star}$ as a restriction o $\preccurlyeq_{\mathsf{S}^\star}$.

Proposition 2 (Completeness of Strategy Ordering)

The ordering of strategies $\preccurlyeq_{\mathsf{S}^\star}$ of the universal DM has to be complete.

Proof It suffices to take $\mathsf{S}_\star = \{{}^{a}\mathsf{S},\ {}^{b}\mathsf{S}\}$ for any $^{a}\mathsf{S},\ ^{b}\mathsf{S} \in \mathsf{S}^\star$. $\square$

The theory must harmonise the strategy $\preccurlyeq_{S^\star}$ & preferential $\preccurlyeq_{B^\star}$ orderings!

Requirement 1 (The Basic Requirements on Prescriptive DM Theory)

A prescriptive DM theory is to use the ordering of strategies ↳ $\preccurlyeq_{S^\star}$, which is

- *complete in order to serve to all DM tasks differing in sets $S_\star$ of compared strategies ↳ $S_\star \subset S^\star$, where any $S \in S^\star$ is admissible strategy ↳*
- *represented numerically in order to allow algorithmic optimisation*
- *harmonised with the user's preferential ordering ↳ $\preccurlyeq_{B^\star}$.*

Remarks

- The formalisation what the harmonisation means is postponed.
- The ordering of strategies ↳ will be represented by an "expected" loss.
- Under widely acceptable conditions, it will be shown that the representation is to be mathematical expectation of the loss ↳ shaped by a utility ↳. Then quotation marks will be dropped.

Ordering Interpretation Implies Design Principle

The meaning of the strategy ordering implies the definition

- *optimal strategy* ${}^{o}\mathsf{S}$ on $\mathsf{S}_\star \subset \mathsf{S}^\star$ fulfills ${}^{o}\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} \mathsf{S}\ \forall\, \mathsf{S} \in \mathsf{S}_\star$.

The "expected" loss represents preferences among strategies. This gives

- *optimal design* selects an admissible strategy among compared strategies $\mathsf{S}_\star$ that leads to the smallest value of the "expected" loss.

  The optimal strategy ${}^{o}\mathsf{S}$ on $\mathsf{S}_\star$ minimises the "expected" loss on $\mathsf{S}_\star$.

Agreement 3 (Design Conditions)

- *The ordering $\preccurlyeq_{\mathsf{S}^\star}$ has the smallest element in $\mathsf{S}^\star$.*
  This assumption – simplifying the formal treatment – also applies to the ordering $\preccurlyeq_{\mathsf{S}^\star}$ constrained to sets $\mathsf{S}_\star$ of compared strategies.
- *The configuration of the decision maker – system, determining behaviour and its structure, are fixed during design.*

Uncertainty Makes the Choice of Strategy Ordering Hard

We start with a preferential ordering↳ $\preccurlyeq_{B^\star}$ represented by a fixed loss↳ Z.

- The substitution $B = \mathsf{W}(\mathsf{S}, N)$ of the mapping (2) W, relating the strategy↳ S and the uncertainty↳ N to the behaviour↳ B, into the loss ↳ Z converts the loss↳ into a function $\mathsf{Z}_\mathsf{S}(N)$ of uncertainty↳ $N \in N^\star \neq \emptyset$

$$\mathsf{Z}_\mathsf{S}(N) \equiv \mathsf{Z}(\mathsf{W}(\mathsf{S}, N)), \quad \text{for any } \mathsf{S} \in \mathsf{S}_\star, \quad \text{and} \quad \forall N \in N^\star. \tag{9}$$

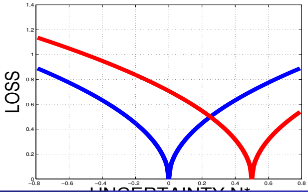
- Compared strategies $\mathsf{S} \in \mathsf{S}_\star$ yield the set of uncertainty functions

$$\mathsf{Z}_{\mathsf{S}_\star} \equiv \left\{ \mathsf{Z}_\mathsf{S} : \; N^\star \to [-\infty, \infty], \; \mathsf{Z}_\mathsf{S}(N) \equiv \mathsf{Z}(\mathsf{W}(\mathsf{S}, N)) \right\}_{\mathsf{S} \in \mathsf{S}_\star}. \tag{10}$$

Example 8 (Hardness of the $\preccurlyeq_{\mathsf{S}^\star}$ Choice)

Let $\mathsf{S}_\star \equiv \{\mathsf{S}, \mathsf{S}\}$, $\mathsf{Z}_{\mathsf{S}_\star}$:

Which one is better?



Ordering of Strategies Orders Possible Loss Realisations

- Recall, $\preccurlyeq_{\mathsf{S}^\star}$ is a complete ordering of all admissible strategies $\mathsf{S} \in \mathbf{S}^\star$ restricted to complete orderings of $\mathsf{S} \in \mathbf{S}_\star \subset \mathbf{S}^\star$ of compared strategies

$$ {}^{a}\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} {}^{b}\mathsf{S}, \ \ {}^{a}\mathsf{S}, {}^{b}\mathsf{S} \in \mathbf{S}_\star, \ \Leftrightarrow \ {}^{a}\mathsf{S} \ \text{ is better than } \ {}^{b}\mathsf{S}. \tag{11} $$

It defines the complete ordering $\preccurlyeq_{\mathsf{Z}_{\mathsf{S}^\star}}$ of functions from the set $\mathsf{Z}_{\mathsf{S}^\star}$ (10)

$$ \mathsf{Z}_{{}^{a}\mathsf{S}} \preccurlyeq_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^{b}\mathsf{S}} \ \Leftrightarrow \ {}^{a}\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} {}^{b}\mathsf{S}. \tag{12} $$

- Let a representation $\mathsf{T}: \ \mathsf{Z}_{\mathsf{S}^\star} \to [-\infty, \infty]$ of $\preccurlyeq_{\mathsf{Z}_{\mathsf{S}^\star}}$ exist, Prop. 1,

$$ \mathsf{Z}_{{}^{a}\mathsf{S}} \preccurlyeq_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^{b}\mathsf{S}} \ \Leftrightarrow \ \mathsf{T}(\mathsf{Z}_{{}^{a}\mathsf{S}}) \le \mathsf{T}(\mathsf{Z}_{{}^{b}\mathsf{S}}). \tag{13} $$

Then, the functional T (13) represents the strategy ordering $\preccurlyeq_{\mathsf{S}^\star}$ due to (12)

$$ {}^{a}\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} {}^{b}\mathsf{S} \ \Leftrightarrow \ \mathsf{T}(\mathsf{Z}_{{}^{a}\mathsf{S}}) \le \mathsf{T}(\mathsf{Z}_{{}^{b}\mathsf{S}}). \tag{14} $$

- The functional T is an “expectation” of the loss $\mathsf{Z}_{\mathsf{S}}(N) = \mathsf{Z}(\mathsf{W}(\mathsf{S}, N))$.

The Path to an Integral Representation of “Expectation”

1. A set of real functions of uncertainty $Z^\star_{S^\star}$ generated by all strategies and *all* losses is taken as the domain of the quantifying functional T.

2. The functional T is expressed in an integral form under rather technical conditions. Essentially, applicability to any smooth loss, smoothness and a local version of “linearity” of T are required.
   - The “expectation” for a specific preferential ordering $\preceq_{B^\star}$, quantified by the specific loss $Z(B)$, is taken as the restriction of T on the set (10)

$$Z_{S_\star} \equiv \left\{ Z_S : N^\star \to [-\infty, \infty],\ Z_S(N) \equiv Z(W(S, N)) \right\}_{S \in S_\star}.$$

   - The numerical representation of the strategy ordering is then

$$^aS \preceq_{S_\star} {}^bS \quad \Leftrightarrow \quad T\left(Z_{^aS}\right) \le T\left(Z_{^bS}\right).$$

3. Functionals T harmonising $\preceq_{S^\star}$ with $\preceq_{B^\star}$ are characterised.

4. The probabilistic nature of uncertainties is stressed.
   The obtained formalised DM task is summarised.

Requirement 2 ("Expectation" Domain)

The "expectation" T *acts on the union* $\mathsf{Z}^\star_{\mathsf{S}^\star}$ *of the sets* $\mathsf{Z}_{\mathsf{S}^\star}$ *(10) of functions with a common uncertainty set* $N^\star \neq \emptyset$

$$\begin{aligned} \mathsf{Z}^\star_{\mathsf{S}^\star} &\equiv \cup_{\mathsf{Z}\in\mathsf{Z}^\star}\mathsf{Z}_{\mathsf{S}^\star} \\ &= \{\mathsf{Z}_\mathsf{S} : \mathsf{Z}_\mathsf{S}(N) = \mathsf{Z}(\mathsf{W}(\mathsf{S}, N)) \in [-\infty, \infty], N \in N^\star\}_{\mathsf{S}\in\mathsf{S}^\star,\mathsf{Z}\in\mathsf{Z}^\star} \end{aligned} \quad (15)$$

and the set $\mathsf{Z}^\star_{\mathsf{S}^\star}$ *is required to contain a subset of*

- *test losses, which are zero out of a compact subset* $\emptyset \neq N_\star$ *of* $N^\star$*, continuous on* $N_\star$*. Supremum norm defines topology & continuity.*
- T *is a sequentially & uniformly continuous on the test losses*↳*.*
- T *is locally additive, i.e. additive on losses with disjoint supports:*

$$\mathsf{T}[{}^a\mathsf{Z} + {}^b\mathsf{Z}] = \mathsf{T}[{}^a\mathsf{Z}] + \mathsf{T}[{}^b\mathsf{Z}] \quad \textit{if} \quad {}^a\mathsf{Z} \times {}^b\mathsf{Z} = 0, \ {}^a\mathsf{Z}, \ {}^b\mathsf{Z} \in \mathsf{Z}^\star_{\mathsf{S}^\star}. \quad (16)$$

Proposition 3 ("Expectation" Form)

Under Requirement 2, the "expectation" T *of* $\mathsf{Z} \in \mathsf{Z}^{\star}_{\mathsf{S}^{\star}}$ *reads*

$$\mathsf{T}[\mathsf{Z}] = \int_{N^{\star}} \mathsf{U}(\mathsf{Z}(N), N)\, \mu(dN), \qquad (17)$$

specified by a finite regular nonnegative Borel measure μ *and by utility* U

- *utility is the mapping* U *in (17), which*
  - *meets* $\mathsf{U}(0, N) = 0$ *almost everywhere (a.e.)*
  - *is continuous in values of* $\mathsf{Z}(\cdot)$ *a.e. on* $N_{\star}$
  - *is bounded a.e. on* $N_{\star}$ *for each* Z *in the set of test losses*$_{\llcorner}$.

Proof The exact formulation and proof of this proposition are in [137], Theorem 5, Chapter 9. □

- The consideration test losses represents no practical restriction. The "expectation" (17) is well defined for a much wider subset of $\mathsf{Z}^{\star}_{\mathsf{S}^{\star}}$ (15).
- The continuity requirements on T are also widely acceptable.
- The local additivity of T on functions with non-overlapping support seems to be sound. Indeed, any loss $\mathsf{Z} \in \mathsf{Z}^{\star}_{\mathsf{S}^{\star}}$ can be written as

$$\mathsf{Z} = \mathsf{Z}\chi_{\omega} + \mathsf{Z}(1-\chi_{\omega}) \equiv {}^{a}\mathsf{Z} + {}^{b}\mathsf{Z} \quad \Rightarrow \quad {}^{a}\mathsf{Z} \times {}^{b}\mathsf{Z} = 0, \text{ with}$$

- *indicator* χ_{ω} of the set $\omega \subset N_{\star} \subset N^{\star}$ is a smooth function that equals 1 within ω and it is zero outside of it.

Local additivity requires the "expected" loss on a set ω & its complement to sum to the loss "expected" on the whole set $N^{\star}$ of uncertainty.

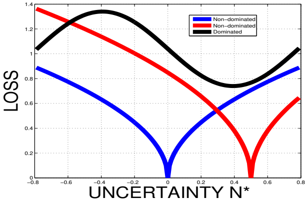
- The utility↳ U shapes the original loss↳ & allows the decision maker↳ to express its attitude toward design consequences and their risks.
- The decision maker↳ can be, [95],
  - risk aware – higher loss values are amplified by U
  - risk prone – higher loss values are attenuated by U
  - indifferent to risk – identity is used as U.
- The utility↳ U and the nonnegative measure μ are universal (objective) by serving to a range of decision tasks differing in compared strategies and losses but facing the same uncertainty↳.
- The use of the set description of uncertain events leads to Kolmogorov's type probability. More general projector-based direction used in quantum physics [124] is a hot emerging research direction [25]. It deals with an alternative modelling of uncertainties.

- The introduced ordering of strategies $\preccurlyeq_{S^\star}$ has been up to now unrelated to the preferential ordering $\preccurlyeq_{B^\star}$ of behaviours $B \in B^\star$.

What makes $\preccurlyeq_{S^\star}$ reasonably harmonised with $\preccurlyeq_{B^\star}$?

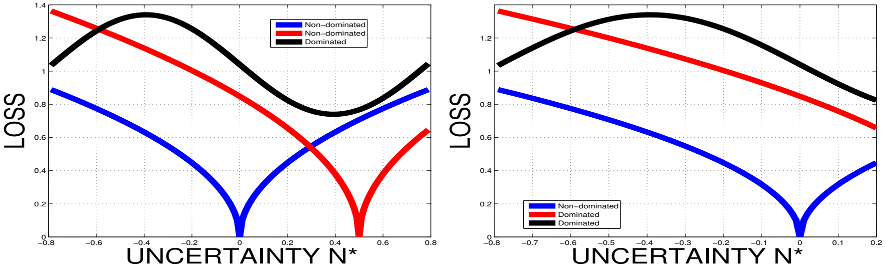
- We have to avoid undoubtedly bad orderings of strategies, which takes a strategy leading surely to a bad behaviour as the optimal one.

  The bad strategy will be identified with dominated strategy, for which a strategy exists uniformly outperforming it for all uncertainties



We are not aware another universal harmonisation requirement!

- The extent of dominated strategies depends on the uncertainty extent.



- The uncertainty extent depends on the relevance of the behaviour content with respect to the DM task, i.e. on modelling quality.

Dominance Ordering

ad step 3

- The set $\mathsf{Z}_{\mathsf{S}^\star} = \left\{\mathsf{Z}_{\mathsf{S}} : N^\star \to [-\infty, \infty],\ \mathsf{Z}_{\mathsf{S}}(N) \equiv \mathsf{Z}(\mathsf{W}(\mathsf{S}, N))\right\}_{\mathsf{S}\in\mathsf{S}^\star}$ (10) of functions of uncertainty is equipped with the partial ordering

$$ {}^{a}\mathsf{Z} \preceq^{d}_{\mathsf{Z}_{\mathsf{S}^\star}} {}^{b}\mathsf{Z} \quad \Leftrightarrow \quad {}^{a}\mathsf{Z}(N) \leq {}^{b}\mathsf{Z}(N),\ \forall N \in N^\star. \tag{18} $$

- The ordering acts on the subset $\mathsf{Z}_{\mathsf{S}^\star}$ (10) of the set $\mathsf{Z}^\star_{\mathsf{S}^\star}$, given by a fixed loss Z, and induces the partial dominance ordering $\preceq^{d}_{\mathsf{S}^\star}$ on $\mathsf{S}^\star$
- *dominance ordering* $\preceq^{d}_{\mathsf{S}^\star}$ means

$$ \left({}^{a}\mathsf{S} \preceq^{d}_{\mathsf{S}^\star} {}^{b}\mathsf{S} \ \Leftrightarrow\ {}^{a}\mathsf{S} \text{ dominates } {}^{b}\mathsf{S}\right) \quad \text{iff} \quad \mathsf{Z}_{{}^{a}\mathsf{S}} \preceq^{d}_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^{b}\mathsf{S}} \tag{19} $$

$$ \left({}^{a}\mathsf{S} \prec^{d}_{\mathsf{S}^\star} {}^{b}\mathsf{S} \ \Leftrightarrow\ {}^{a}\mathsf{S} \text{ strictly dominates } {}^{b}\mathsf{S}\right) \quad \text{iff} $$

$\mathsf{Z}_{{}^{a}\mathsf{S}} \prec^{d}_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^{b}\mathsf{S}}$ i.e. (18) is sharp on $N_\star \subset N^\star$ with $\mu(N_\star) > 0$, cf. (17).

A strictly dominated strategy is the bad one as it leads to a higher loss than another admissible strategy irrespectively of the uncertainty!

Dominated Strategies Have to Be Avoided ad step 3

Requirement 3 (Quest for Non-Dominance)

The "expected" loss must be chosen so that the optimal design performed on any nontrivial set $\mathsf{S}_\star \subset \mathsf{S}^\star$ *of compared strategies must not lead to the optimal strategy strictly dominated by a strategy from* $\mathsf{S}_\star$.

Proposition 4 (Isotonic "Expectation")

Let a strategy $\bar{\mathsf{S}} \in \mathsf{S}^\star$ *exist for which the "expected" loss is finite. Then, Requirement 3 is fulfilled iff the "expectation" is strictly isotonic on* $\mathsf{Z}_{\mathsf{S}^\star}$ *with respect to the strict dominance ordering of strategies* $\prec^d_{\mathsf{S}^\star}$.

... the following proof does not need a specific representation of $\mathsf{T}(\mathsf{Z}_\mathsf{S})$.

Isotonic "Expectation" Avoids Dominance

Proof Both directions are proved by contradiction.

- Let $\mathsf{T}[\mathsf{Z}]$ be strictly isotonic on $\mathsf{Z}_{\mathsf{S}^\star}$ (10) and thus on $\mathsf{Z}_{\mathsf{S}_\star} \subset \mathsf{Z}_{\mathsf{S}^\star}$.

  Let ${}^d\mathsf{S} \in \mathsf{S}_\star$ strictly dominate a minimiser ${}^o\mathsf{S} \in \mathsf{S}_\star$ of the "expected" loss↳, which necessarily gives $\mathsf{T}[\mathsf{Z}_{{}^o\mathsf{S}}] \leq \mathsf{T}[\mathsf{Z}_{\bar{\mathsf{S}}}] < \infty$.

  Then, the construction of ${}^o\mathsf{S}$ via minimisation, the strict dominance and the strictly isotonic nature of T, imply the following contradiction

$$\mathsf{T}[\mathsf{Z}_{{}^o\mathsf{S}}] \underbrace{\leq}_{\text{minimum}} \mathsf{T}[\mathsf{Z}_{{}^d\mathsf{S}}] \underbrace{<}_{\text{strictly isotonic}} \mathsf{T}[\mathsf{Z}_{{}^o\mathsf{S}}] < \infty.$$

- If $\mathsf{T}[\mathsf{Z}_\mathsf{S}]$ is not strictly isotonic on $\mathsf{Z}_{\mathsf{S}^\star}$ (10) then there is a strategy ${}^a\mathsf{S} \in \mathsf{S}^\star$ strictly dominated by a strategy ${}^d\mathsf{S} \in \mathsf{S}^\star$ & $\mathsf{T}[\mathsf{Z}_{{}^d\mathsf{S}}] \geq \mathsf{T}[\mathsf{Z}_{{}^a\mathsf{S}}]$.

  For the set of compared strategies $\mathsf{S}_\star \equiv \{{}^d\mathsf{S}, {}^a\mathsf{S}\}$, ${}^a\mathsf{S}$ can always be taken as the optimal strategy↳ and T violates Requirement 3. □

Increasing Utility Makes the "Expectation" Isotonic ad step 3

Proposition 5 (Utility Must Be Increasing Function of Loss)

The optimal design avoids dominated strategies iff the utility in the "expectation" (17) is increasing in its first argument.

Proof

- If the utility $\mathsf{U}(z, N)$ is increasing in z then T is isotonic with respect to the strict dominance ordering of strategies & Prop. 4 applies.
- If real numbers ${}^{a}z > {}^{b}z$ exist such that $\mathsf{U}({}^{a}z, N) \le \mathsf{U}({}^{b}z, N)$ for $N \in N_{\star} \subset N^{\star}$ with $\mu(N_{\star}) > 0$. Then by assigning to strategies ${}^{a}\mathsf{S}, {}^{b}\mathsf{S} \in \mathsf{S}_{\star} \subset \mathsf{S}^{\star}$, $c \in \{a, b\}$, loss functions

$$\mathsf{L}_{{}^{a}\mathsf{S}}(N) = \begin{cases} {}^{a}z & \text{on } N_{\star} \\ 0 & \text{otherwise} \end{cases}, \quad \mathsf{L}_{{}^{b}\mathsf{S}}(N) = \begin{cases} {}^{b}z & \text{on } N_{\star} \\ 0 & \text{otherwise} \end{cases},$$

we arrive to the contradiction as ${}^{a}\mathsf{S}$ is jointly dominated and optimal. $\square$

- The "expectation" scaled by a positive factor preserves the ordering. Thus, μ (17) can be normalised to the probabilistic measure.

  This choice preserves the constant utilities $\mathsf{T}[constant] = constant$.

- Back-substitution $B = \mathsf{W}(\mathsf{S}, N)$ (2) into (17) gives for ${}^a\mathsf{S}$, ${}^b\mathsf{S}$, $\mathsf{S} \in \mathsf{S}^\star$

$$
\begin{aligned}
{}^a\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} {}^b\mathsf{S} &\Leftrightarrow \mathsf{T}(\mathsf{Z}_{{}^a\mathsf{S}}) \le \mathsf{T}(\mathsf{Z}_{{}^b\mathsf{S}}), \quad \mathsf{Z}_\mathsf{S} \in \mathsf{Z}_{\mathsf{S}^\star} = \left\{\mathsf{Z}(\mathsf{W}(S, N))_{N \in N^\star}\right\}_{\mathsf{S} \in \mathsf{S}^\star} \\
&\Leftrightarrow \mathsf{T}_{{}^a\mathsf{S}}(\mathsf{Z}) \le \mathsf{T}_{{}^b\mathsf{S}}(\mathsf{Z}), \quad \mathsf{Z} \in \mathsf{Z}^\star = \{\mathsf{Z} : B^\star \to [-\infty, \infty]\} \\
\mathsf{T}_\mathsf{S}(\mathsf{Z}) &\equiv \int_{B^\star} \mathsf{U}\left(\mathsf{Z}(B), \mathsf{W}^{-1}(\mathsf{S}, B)\right) \mu_\mathsf{S}(\,\mathrm{d}B), \qquad (20)
\end{aligned}
$$

where the probabilistic measure μ_S is the image of the measure μ

$$\mu_\mathsf{S}(B_\star) = \mu(N_\star), \;\; B_\star \equiv \mathsf{W}(\mathsf{S}, N_\star) \subset B^\star, \;\; N_\star \subset N^\star$$

via the bijective mapping $\mathsf{W}(\mathsf{S}, \cdot)$ (2) modelling the uncertainty influence.

Closed-Loop Model & Performance Index

The measures μ_{S}, $\mathsf{S} \in \mathsf{S}^\star$ are assumed to have probability density, i.e.

- *Radon–Nikodým derivative* (pd ↳) $\mathsf{f}_{\mathsf{S}}(B)$ of $\mu_{\mathsf{S}}(B)$, which is defined with respect to a product dominating measure $\mathrm{d}B$, [137].
- The "expectation" T becomes the expectation ↳

$$\mathsf{E}_{\mathsf{S}}[\mathsf{I}_{\mathsf{S}}] \equiv \int_{B^\star} \mathsf{I}_{\mathsf{S}}(B)\mathsf{f}_{\mathsf{S}}(B)\mathrm{d}B = \mathsf{T}_{\mathsf{S}}(\mathsf{Z}), \quad \text{determined by} \tag{21}$$

- *closed-loop model* , which is the S-dependent pd $\mathsf{f}_{\mathsf{S}}(B)$ on $B^\star$.

The expectation E_{S} is applied to a strategy-dependent

- *performance index* $\mathsf{I}_{\mathsf{S}}(B) \equiv \mathsf{U}(\mathsf{Z}(B), \mathsf{W}^{-1}(\mathsf{S}, B)).$ (22)

The universality of the construction allows us to define

- *objective expectation* is the expectation (21) that serves to all DM ↳ tasks with a common uncertainty ↳ delimited by $\mathsf{W}(\mathsf{S}, N)$, $N \in N^\star$, (2).
- *objective pd* is the pd specifying the objective expectation ↳.

A Few Comments

ad step 4

Remark 4 (On Standard & Fully Probabilistic Designs)

- *The dependence of the performance index↳ on strategy↳* S *arises through the non-standard second argument of utility↳.*

  The dependence of the performance index on S *does not occur in*
- *standard design uses* S*-independent performance index* $\mathsf{I} = \mathsf{I}_{\mathsf{S}}$.
- *The standard design↳ is addressed up to Sec. 16, where its generalisation helps us to justify the fully probabilistic design (FPD↳).*

Remark 5 (On Probability Densities)

- *The assumed existence of pds helps us to deal with simpler objects.*
- *Under our notation, the pds* $\mathsf{f}(N)$ *and* $\mathsf{f}(B)$ *are different functions.*
- *The dominating measure is mostly unspecified & the Lebèsgue notation is used.*

- *formalised DM design* searches for

$$
\begin{array}{lrcl}
\text{optimal strategy} & {}^{o}\mathsf{S} & \in & \mathrm{Arg}\min_{\mathsf{S}\in\mathsf{S}^\star} \mathsf{E}_{\mathsf{S}}[\mathsf{I}_{\mathsf{S}}] \text{ with} \\
\text{admissible strategy} & \mathsf{S} & \equiv & (\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to A_t{}^\star)_{t\in t^\star} \in \mathsf{S}^\star \\
\text{performance index} & \mathsf{I}_{\mathsf{S}} & : & B^\star \to [-\infty,\infty] \text{ evaluating} \\
\text{behaviour} & B & \equiv & (\mathbf{G}_t, A_t, \mathbf{K}_{t-1}) \\
 & & \equiv & (\text{ignorance}, \text{action}, \text{knowledge}) \\
\text{expectation} & \mathsf{E}_{\mathsf{S}}[\mathsf{I}_{\mathsf{S}}] & \equiv & \int_{B^\star} \mathsf{I}_{\mathsf{S}}(B)\mathsf{f}_{\mathsf{S}}(B)\,\mathrm{d}B \\
\text{closed-loop model} & \mathsf{f}_{\mathsf{S}} & : & B^\star \to [0,\infty],\ \int_{B^\star} \mathsf{f}_{\mathsf{S}}(B)\,\mathrm{d}B = 1.
\end{array}
\tag{23}
$$

A course alumnus should

- ✓ understand background of this DM design
- ✓ be able formalise her/his specific DM task
- ✓ learn tools for formalisation & design

Exercises 4: Home Work

Exercises 4 (Formalisation of a Dishonest Coin Tossing as DM)

The problem description

- *You can toss the coin h-times.*
- *You get 1Kč whenever the head will be up otherwise you get nothing.*
- *The result depends on your tossing style, specifically, whether you will put head up on your tossing hand or not.*

The task for you

- *Formulate this problem as the formalised DM design↳.*

The key requirement

- *Specify all DM elements↳ entering the formalised DM design, including sets on which it operates.*

The Formalisation Gaps

We aim at the optimal design$_{\lrcorner}$ but we can perform only

- *practically optimal design* selects a practically admissible strategy$_{\lrcorner}$ gained by the design$_{\lrcorner}$ with limited resources allocated it.
- *formal* tools for practically optimal design$_{\lrcorner}$ are unavailable; the attempts to create them lead to infinite regress. The algorithmic
    - fixing of DM elements, e.g. the behaviour set $B^{\star}$, is the meta-DM task operating on their candidates enforcing the choice of candidates . . .
    - limiting the design$_{\lrcorner}$ complexity leads to a more complex meta-DM task whose complexity can be limited by a more complex DM task . . . , [151]
- the problem can be diminished but not fully removed by
    - sequential DM with stopping rules [164, 127] to meta-DMs
    - evolution gradually accumulating the solved DM prototypes
    - the involvement of clever designer$_{\lrcorner}$s: you!

Exercises 5 (On DM Complexity)

1. *You search for maximum of a function while you*
   - *select sequentially its argument and evaluate the function and compare with previous values at the unit cost*
   - *pay a unit cost for considering any idea how to choose the argument*
   - *have the budget equal to 221.*

   Notice
   - *to have 221 ideas is surely unproductive*
   - *to make 221 random choices is probably unproductive.*
2. *The problem is closely related to*
   - *secretary (marriage) problem [43],*
   - *no free lunch theorems [167].*

   Learn more about them and think over other real-life situations fitting the scheme.

An Outline of the Further Explanations

- Tools needed for the optimisation (23) are prepared in Secs. 8, 9, 10.
- The solution of the standard design↳ of the optimal strategy↳ is found, Sec. 11. Learning needed for it is summarised in Secs. 13, 14.
- Sects. 12, 15 inspect asymptotic properties of the design and learning.
- General, the fully probabilistic design of decision strategies (FPD↳), dealing with a strategy-dependent performance index↳ is in Sec. 16.
- FPD is connected with standard design↳ in Sec. 17 & solved in Sec. 19.

  Among others, FPD↳ provides tools for a construction of DM elements↳ used in design↳. This construction forms the next part of this text, which is finely structured before its start.

Basic Design Tools

About this Part

This part provides basic tools serving us for solving DM tasks, specifically:

- Sec. 8 recalls elementary calculus with pd ↳s.
- Sec. 9 summarises the used properties of conditional expectation.
- Sec. 10 proves the basic DM lemma on which the design ↳ relies.
  It also explicates the DM elements ↳ on which the design ↳ operates.

Calculus with PDs Related to Joint PD $f(B) \equiv f(\alpha, \beta, \gamma)$

The joint pd $f(B) \equiv f(\alpha, \beta, \gamma)$ determines the following derived pds.

- *joint pd* $f(\alpha, \beta|\gamma)$ of α, β conditioned on γ is the pd on $(\alpha, \beta)^\star$ projecting the pd $f(\alpha, \beta, \gamma)$ on the cross-section of $B^\star$ given by a fixed γ.
- *marginal pd* $f(\alpha|\gamma)$ of α conditioned on γ is the pd on $\alpha^\star$ projecting $f(\alpha, \beta, \gamma)$ on the cross-section of $B^\star$ given by γ with no information on β.
- *unconditional pd* is formally obtained if just a trivial condition is considered. Then, the conditioning symbol $|$ is dropped.

Pd $f(\alpha, \beta)$ is marginal pd of pd $f(\alpha, \beta, \gamma)$ & lower-dimensional joint pd.

- *conditionally independent* quantities α, β, under the condition γ, meet

$$f(\alpha, \beta|\gamma) = f(\alpha|\gamma)f(\beta|\gamma) \Leftrightarrow f(\alpha|\beta, \gamma) = f(\alpha|\gamma). \tag{24}$$

Properties of PDs & Basic Operations with PDs

Proposition 6 (Properties of PDs)

For any $B \equiv (\alpha, \beta, \gamma) \in B^{\star}$, *the following properties hold*

- *non-negativity: all variants of pds are non-negative.*
- *normalisation: pds have unit integral over the domain of quantities before the conditioning sign* $|$.
- *chain rule for pds: it holds* $\mathsf{f}(\alpha, \beta|\gamma) = \mathsf{f}(\alpha|\beta, \gamma)\mathsf{f}(\beta|\gamma)$.
- *marginalisation: the marginal pd* $\mathsf{f}(\beta|\gamma) = \int_{\alpha^{\star}} \mathsf{f}(\alpha, \beta|\gamma)\, \mathrm{d}\alpha$.
- *Bayes' rule:*

$$\mathsf{f}(\beta|\alpha,\gamma) = \frac{\mathsf{f}(\alpha|\beta,\gamma)\mathsf{f}(\beta|\gamma)}{\mathsf{f}(\alpha|\gamma)} = \frac{\mathsf{f}(\alpha|\beta,\gamma)\mathsf{f}(\beta|\gamma)}{\int_{\beta^{\star}} \mathsf{f}(\alpha|\beta,\gamma)\mathsf{f}(\beta|\gamma)\, \mathrm{d}\beta} \propto \mathsf{f}(\alpha|\beta,\gamma)\mathsf{f}(\beta|\gamma). \quad (25)$$

- *proportionality* $\propto$ is equality with an implicit presence of a unique normalisation factor independent of the pd's argument before the sign $|$.

Proof outlined Relations of pds follow from the required invariance of the expected value with respect to operations like seeing $\mathsf{I}(\alpha)$ as performance index$_{\hookrightarrow}$ evaluating a behaviour (α, β) only according to α, see [83].

Let, for instance, $\mathsf{I}(\alpha, \beta) = \mathsf{I}(\beta)$, then

$$\begin{aligned} \mathsf{E}[\mathsf{I}] &= \int_{\alpha^\star}\int_{\beta^\star} \mathsf{I}(\alpha,\beta)\mathsf{f}(\alpha,\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta = \int_{\beta^\star} \mathsf{I}(\beta)\left\{\int_{\alpha^\star} \mathsf{f}(\alpha,\beta)\,\mathrm{d}\alpha\right\}\mathrm{d}\beta \\ \overbrace{=}^{!}\mathsf{E}[\mathsf{I}] &= \int_{\beta^\star} \mathsf{I}(\beta)\left\{\mathsf{f}(\beta)\right\}\mathrm{d}\beta \overbrace{\Rightarrow}^{\{\ \}} \mathsf{f}(\beta) = \int_{\alpha^\star} \mathsf{f}(\alpha,\beta)\,\mathrm{d}\alpha. \end{aligned}$$

Bayes' rule (25) applies twice the chain rule$_{\hookrightarrow}$ and then uses marginalisation

$\hookrightarrow$

$$\mathsf{f}(\alpha,\beta|\gamma) = \mathsf{f}(\beta|\alpha,\gamma)\mathsf{f}(\alpha|\gamma) = \mathsf{f}(\alpha|\beta,\gamma)\mathsf{f}(\beta|\gamma) \qquad \square$$

Remark 6 (Calculus with PDS)

- *A rigorous treatment of pds uses measure theory [137].*
- *A nice engineering interpretation is in [132].*

Remark 7 (Bayes' Rule & Quantum Probability)

- *Importance of Bayes' rule↳ for the adopted DM under uncertainty cannot be exaggerated, cf. Props. 16, 17 on Bayesian learning.*
- *Bayes' rule exploits commutativity inherent to Kolmogorov's probability.*
- *There are strong indicators that whole DM machinery can be generalised to exploit non-commutative (quantum) probability [124].*

  For single decision maker↳, this generalisation already improves modelling of human cognition [25, 135].

  Strong contributions to the strategy design↳ are expected in DM of multiple interacting decision makers.
- *The foreseen extension will use the required independence of uncertainty↳ and strategy in conjunction with Gleason's theorem characterising measures on Hilbert spaces [38].*

Smooth Transformation of PDs

Proposition 7 (PDs of Smoothly Transformed Quantities)

Let α *be a real vector ,* $\alpha \equiv [\alpha_1, \ldots, \alpha_{\ell_{\alpha^\star}}]$ *and* $\Upsilon = [\Upsilon_1, \ldots, \Upsilon_{\ell_{\alpha^\star}}]$ *bijection with finite continuous partial derivatives almost everywhere on* $\alpha^\star$

$$J_{ij}(\alpha) \equiv \frac{\partial \Upsilon_i(\alpha)}{\partial \alpha_j}, \;\; i, j = 1, \ldots, \ell_{\alpha^\star}, \tag{26}$$

for all entries Υ_i *of* Υ *and entries* α_j *of* α*. Then,*

$$\mathsf{f}_\Upsilon(\Upsilon(\alpha))|\mathbb{J}(\alpha)| = \mathsf{f}(\alpha), \tag{27}$$

where $|\mathbb{J}|$ *denotes absolute value of the determinant of the matrix* $\mathbb{J}$*.*

Proof Prop. describes substitutions in multivariate integrals; see, for instance, [62, 137]. □

Proposition 8 (Basic Properties of Expectation E)

For real functions ${}^{a}\mathsf{I}(B)$, ${}^{b}\mathsf{I}(B)$, $\mathsf{I}(B)$, $B \in B^{\star}$, *for which the conditional expectations* $\mathsf{E}[{}^{a}\mathsf{I}|\gamma]$, $\mathsf{E}[{}^{b}\mathsf{I}|\gamma]$, $\mathsf{E}[\mathsf{I}|\gamma]$ *exist, it holds*

- *expectation linearity*

$$\mathsf{E}[a(\gamma)\,{}^{a}\mathsf{I} + b(\gamma)\,{}^{b}\mathsf{I}|\gamma] = a(\gamma)\mathsf{E}[{}^{a}\mathsf{I}|\gamma] + b(\gamma)\mathsf{E}[{}^{b}\mathsf{I}|\gamma] \qquad (28)$$

  for arbitrary γ*-dependent real coefficients* $a(\gamma)$, $b(\gamma)$.

- *chain rule for expectation* $\quad \mathsf{E}\left[\mathsf{E}[\mathsf{I}|\beta,\gamma]|\gamma\right] = \mathsf{E}[\mathsf{I}|\gamma] \qquad (29)$

  for an arbitrary additional condition β.

Proof The integral form of E & Fubini theorem provide the chain rule (29):

$$\begin{aligned}
\mathsf{E}[\mathsf{I}|\gamma] &= \int_{\alpha^{\star},\beta^{\star}} \mathsf{I}(\alpha,\beta,\gamma)\mathsf{f}(\alpha,\beta|\gamma)\,\mathrm{d}\alpha\,\mathrm{d}\beta \\
&= \int_{\beta^{\star}} \left[\int_{\alpha^{\star}} \mathsf{I}(\alpha,\beta,\gamma)\mathsf{f}(\alpha|\beta,\gamma)\,\mathrm{d}\alpha\right] \mathsf{f}(\beta|\gamma)\,\mathrm{d}\beta = \mathsf{E}\left[\mathsf{E}[\mathsf{I}|\beta,\gamma]|\gamma\right] \qquad \square
\end{aligned}$$

Second Moments & Jensen's Inequality

Proposition 9 (Covariance & Jensen's Inequality)

- *conditional covariance of a vector* α

$$\mathrm{cov}[\alpha|\gamma] \equiv \mathsf{E}\left[(\alpha - \mathsf{E}[\alpha|\gamma])(\alpha - \mathsf{E}[\alpha|\gamma])'|\gamma\right] \geq 0 \textit{ (positive semidefinite)}$$

relates to the non-central moments through the formula

$$\mathrm{cov}[\alpha|\gamma] = \mathsf{E}[\alpha\alpha'|\gamma] - \mathsf{E}[\alpha|\gamma]\mathsf{E}[\alpha'|\gamma], \quad ' \textit{ is transposition}. \tag{30}$$

- *Jensen's inequality bounds expectation of a convex function* $\mathsf{I}_\gamma : \alpha^\star \to [-\infty, \infty]$

$$\mathsf{E}[\mathsf{I}_\gamma(\alpha)|\gamma] \geq \mathsf{I}_\gamma\left(\mathsf{E}[\alpha|\gamma]\right). \tag{31}$$

Proof The integral expression of the expectation verifies the first part. Proof of Jensen's inequality is, e.g., in [162]. □

On Calculus with Expectations

Remark 8 (On Expectations)

- *The expectation of an array* $V = (V_i)_{i \in i^\star}$ *is the expectation array*

$$(\mathsf{E}[V])_i \equiv \mathsf{E}[V_i], \quad i \in i^\star.$$

- *The recalled properties are formulated for the conditional expectation. The unconditional case is obtained by omitting the condition.*
- *Mostly, it suffices to take the expectation as an integral weighted by a pd* $\mathsf{f}(\cdot|\gamma)$*. The book [137] offers a nice rigourous treatment.*
- *Use of the symbol* E *instead of integrals simplifies manipulations. It may be misleading when the used pd is not obvious. Then:*
  - *The symbol* E_{f} *is used to stress the expectation-defining pd* f.
  - *The symbol* E_{S} *is used to stress that the expectation-defining pd depends on strategy*↳ S *influencing the used pd.*
  - *An explicit use of arguments or integral expression helps.*

Basic DM Lemma Optimising Admissible Decision Rule

Proposition 10 (Basic DM Lemma of Standard Design)

The optimal admissible decision rule $^o\mathsf{S}$ *in the standard design (23) with a performance index* $\mathsf{I} \equiv \mathsf{I}(B) = \mathsf{I}(\mathbf{G}_{A^\star}, A, \mathbf{K}_{A^\star})$ *can be chosen as the deterministic one* $^o\mathsf{S}(\mathbf{K}_{A^\star}) \equiv {}^oA(\mathbf{K}_{A^\star})$ *in the following value-wise way.*

- *The optimal action* $^oA(\mathbf{K}_{A^\star})$, *assigned to an arbitrary* $\mathbf{K}_{A^\star} \in {\mathbf{K}_{A^\star}}^\star$, *is*

$$^oA(\mathbf{K}_{A^\star}) \in \mathrm{Arg} \min_{A \in A^\star} \mathsf{E}[\mathsf{I}|A, \mathbf{K}_{A^\star}]. \tag{32}$$

- *The reached minimum is*

$$\min_{(\mathsf{S}:\mathbf{K}_{A^\star} \to A^\star)^\star} \mathsf{E}[\mathsf{I}] = \mathsf{E}\left[\min_{A \in A^\star} \mathsf{E}[\mathsf{I}|A, \mathbf{K}_{A^\star}]\right]. \tag{33}$$

- The design of the decision rule in the standard design (23) reduces the minimisation over mappings to an "ordinary" function minimisation.
- Recall that the existence of the minimiser is implicitly supposed.

Proof of the Basic DM Lemma

- Let an arbitrary $\mathbf{K}_{A^\star} \in \mathbf{K}_{A^\star}{}^\star$ be fixed. The definition of a minimiser ${}^oA(\mathbf{K}_{A^\star})$ and the assignment ${}^oA(\mathbf{K}_{A^\star}) = {}^o\mathsf{S}(\mathbf{K}_{A^\star})$ imply that $\forall A \in A^\star$

$$\mathsf{E}\left[\mathsf{I}|\, {}^oA(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}\right] = \mathsf{E}\left[\mathsf{I}|\, {}^o\mathsf{S}(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}\right] \le \mathsf{E}[\mathsf{I}|A, \mathbf{K}_{A^\star}]. \quad (34)$$

- Let a decision rule S assign $A \in A^\star$ to the fixed $\mathbf{K}_{A^\star}$. Then (34) reads

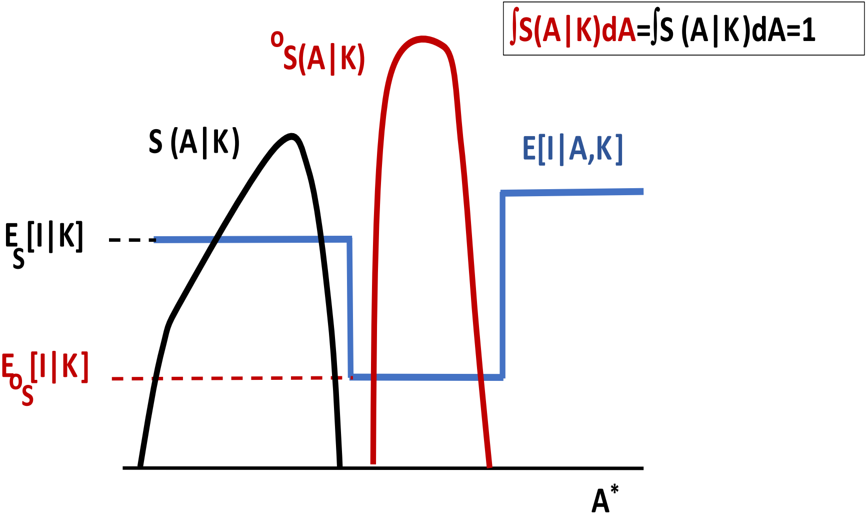
$$\mathsf{E}[\mathsf{I}|\, {}^o\mathsf{S}(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}] \le \mathsf{E}[\mathsf{I}|\mathsf{S}(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}].$$

- Let us apply unconditional expectation $\mathsf{E}[\cdot]$, acting on functions of $\mathbf{K}_{A^\star}$, to this inequality. The isotonic $\mathsf{E}[\cdot]$ preserves the inequality, i.e.

$$\mathsf{E}\left[\mathsf{E}[\mathsf{I}|\, {}^o\mathsf{S}(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}]\right] \le \mathsf{E}\left[\mathsf{E}[\mathsf{I}|\mathsf{S}(\mathbf{K}_{A^\star}), \mathbf{K}_{A^\star}]\right]. \quad (35)$$

- The chain rule for expectation (29) implies that the right-hand side of (35) is the unconditional expectation of I given by an arbitrary S.
- The left-hand side of (35) is the unconditional expected loss for ${}^o\mathsf{S}$. Thus, ${}^o\mathsf{S} : \mathbf{K}_{A^\star} \to {}^oA(\mathbf{K}_{A^\star})$ is the optimal decision rule. $\square$

Illustration of the Basic Decision Lemma



The optimal rule ${}^{o}S$ can be non-unique and randomised.

The achieved minimum value $E_{{}^{o}S}[I|K]$ is unique.

Randomised Strategy & Its Model

- Prop. 10 and its proof imply no preferences among multiple globally minimising arguments $^{o}A(\mathbf{K}_{A^\star})$.

  The optimal action can be selected randomly according to the pd $\mathsf{f}(A_t|\mathbf{K}_{t-1})$, which has its support↳ concentrated on these minimisers.

This useful extension of deterministic rules motivates the next definitions.

- *model of decision rule* is a pd $\mathsf{f}(A|\mathbf{K}_{A^\star}) = \mathsf{S}(A|\mathbf{K}_{A^\star})$.
- *randomised decision rule* is described by a pd with support containing more than one action.
- *model of decision strategy* is the collection of pds $(\mathsf{f}(A_t|\mathbf{K}_{t-1}))_{t\in t^\star}$.

  Strategies with the same model provide the same closed-loop model↳. Thus, strategy↳ can be and will be identified with its model.
- *randomised strategy* has at least one randomised decision rule.

Example 9 (The Point Estimator of an Unknown Parameter $\Theta \in \Theta^\star$)

- *The behaviour* is $B = (\mathbf{G}_{A^\star}, A, \mathbf{K}_{A^\star}) = (\Theta, \hat{\Theta}, D) =$
  (unknown parameter, parameter estimate, known data).
- *The optimal estimator, is searched among rules* $\mathsf{S} : \mathbf{K}_{A^\star}{}^\star \to A^\star$.
- *The performance index* $\mathsf{I}(B) = \mathsf{I}(\Theta, \hat{\Theta}, D)$ *is a "distance" of* Θ *to* $\hat{\Theta}$, *which is strictly convex in* Θ *with the minimum at* $\hat{\Theta} = \Theta$, $\forall D \in D^\star$.

Solution

- *The optimal* ${}^o\hat{\Theta}(D) \in \mathrm{Arg}\min_{\hat{\Theta}\in\hat{\Theta}^\star} \mathsf{E}[\mathsf{I}(\Theta, \hat{\Theta}, D)|\hat{\Theta}, D]$, *Prop. 10.*
- *Jensen's inequality* $\Rightarrow$ $\mathsf{E}[\mathsf{I}(\Theta, \hat{\Theta}, D)|\hat{\Theta}, D] \geq \mathsf{I}(\mathsf{E}[\Theta|\hat{\Theta}, D], \hat{\Theta}, D)$ $\Rightarrow$

$$ {}^o\hat{\Theta}(D) = \mathsf{E}[\Theta|\,{}^o\hat{\Theta}(D), D] = \int_{\Theta^\star} \Theta \mathsf{f}(\Theta|\,{}^o\hat{\Theta}(D), D)\,\mathrm{d}\Theta. $$

- *The estimate* $\hat{\Theta}(D)$ *has no influence on the parameter* Θ: $\hat{\Theta}(D)$ *and* Θ *are conditionally independent, i.e.* $\mathsf{E}[\Theta|\hat{\Theta}, D] = \mathsf{E}[\Theta|D] = {}^o\hat{\Theta}(D)$.

Complexity Jump Between Static & Dynamic Design

Remark 9 (Static vs. Dynamic Design)

- *It is worth stressing that the optimal decision rule ↳ is constructed value-wise. To get it, the minimisation must be performed for all possible instances of the knowledge* $\mathbf{K}_{A^\star} \in \mathbf{K}_{A^\star}{}^\star$.
- *One minimisation suffices if the optimal action for a fixed (observed) knowledge ↳ is needed. This often happens in estimation tasks.*

  This possibility is the main distinction of the static design from the dynamic design ↳, which searches for a sequence of optimal actions & needs the construction of decision rules as functions, see Sec. 11.

  The need to construct functions makes the dynamic design much harder and, mostly, exactly infeasible [18, 42, 149].

Exercises 6 on Actions

Exercises 6 (On Actions in DM)

1. *What is the mathematical reason that the optimal rule can always be deterministic?*
2. *Provide a performance index giving a parameter estimate* $\neq \mathsf{E}[\Theta|\mathbf{K}_{\hat{\Theta}}]$.
3. *Formulate and solve interval estimation as DM task.*
4. *Estimator selects action*$_\lrcorner$*, which does not influence the system*$_\lrcorner$*. Provide an example of DM task with action*$_\lrcorner$*s influencing the system*$_\lrcorner$*.*
5. *Medical doctor has to decide whether to prescribe antibiotics to patient or not. Formulate and solve this as static (one-shot) DM task.*
6. *Judge has to decide whether the accused is guilty or not. Formulate and solve this as static DM task.*
7. *What are possible dynamic consequences of the above tasks?*

Dynamic Design

Recall

- The dynamic design selects actions $A^h = (A_t)_{t=1}^h$ for a horizon $h > 1$.
- The standard design means that the performance index $\mathsf{I}_{\mathsf{S}}(B) = \mathsf{I}(B)$ evaluates behaviour $B \in B^\star$ not the strategy $\mathsf{S} \in \mathsf{S}^\star$.
- The searched optimal admissible strategy $^{o}\mathsf{S}$ consists of decision rules acting at least on the same knowledge as their predecessors.

  This extending knowledge reflects a non-decreasing amount of data available for the DM while no accumulated knowledge is discarded.
- Unless said otherwise, functions $\mathsf{V}(\mathbf{K}_{t-1})$ and $\mathsf{V}(\mathbf{K}_t)$, $t \in t^\star$, differ, i.e. their dependence on time t is implicit.

Extending Knowledge & Dynamic Programming

- The dynamic design ↳ deals with the knowledge ↳ extended by the observation ↳ Δ_t and by the made action A_t, $t \in t^\star$,

$$\mathbf{K}_t{}^\star = (\Delta_t{}^\star, A_t{}^\star) \cup \mathbf{K}^\star_{t-1} = D_t{}^\star \cup \mathbf{K}^\star_{t-1}. \quad (36)$$

- *data record* $D_t = (\Delta_t, A_t)$ =(observation,action) extends the knowledge $\mathbf{K}_{A_t^\star} = \mathbf{K}_{t-1}$ to the knowledge $\mathbf{K}_{A_{t+1}^\star} = \mathbf{K}_t$.

Agreement 4 (Implicit Presence of Prior Knowledge)

All mappings are (mostly implicitly) conditioned on prior knowledge $\mathbf{K}_0$.

This allows to identify $$\mathbf{K}_{t-1} = D^{t-1}, \quad (37)$$

which applies for subsequent text that switches between the labels (37).

- The optimal admissible strategy ↳ is below found by via a stochastic version of the celebrated dynamic programming [10], Prop. 11.
  Dynamic programming "naturally" meets informational constraints ↳.
  This makes its optimisation role exceptional.

Dynamic Programming Solving Dynamic DM Tasks

Proposition 11 (Dynamic Programming)

The optimal strategy ${}^{o}\mathsf{S} \equiv \left({}^{o}\mathsf{S}_t : \mathbf{K}^{\star}_{t-1} \to A^{\star}_t \right)_{t \in t^{\star}}$, minimising the expected performance index $\mathsf{E}[\mathsf{I}(B)]$ while using the extending knowledge $\mathbf{K}_t{}^{\star} = D_t{}^{\star} \cup \mathbf{K}^{\star}_{t-1}$, is constructed in the next value-wise way.

- *For $t \in t^{\star}$ and any $\mathbf{K}_{t-1} \in \mathbf{K}^{\star}_{t-1}$, it takes a minimiser ${}^{o}A(\mathbf{K}_{t-1})$ in*

$$\mathsf{V}(\mathbf{K}_{t-1}) = \min_{A_t \in A_t{}^{\star}} \mathsf{E}[\mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}], \ \ t \in t^{\star}, \tag{38}$$

  as its t-th action ${}^{o}A(\mathbf{K}_{t-1}) = {}^{o}\mathsf{S}_t(\mathbf{K}_{t-1})$.

- *The functional recursion (38) runs against the time course given by the extending knowledge. The recursion starts at the horizon h with*

$$\mathsf{V}(\mathbf{K}_h) \equiv \mathsf{E}[\mathsf{I}(B) | \mathbf{K}_h]. \tag{39}$$

  There, $\mathbf{K}_h$ is the knowledge available for the unmade choice of A_{h+1}.

- *Under Agreement 4, $\mathsf{V}(\mathbf{K}_0) = \min_{\mathsf{S}^{\star} \equiv \left\{ \left(\mathsf{S}_t : \mathbf{K}^{\star}_{t-1} \to A_t{}^{\star} \right)_{t \in t^{\star}} \right\}} \mathsf{E}[\mathsf{I}(B)]$.*

- Let $\mathbf{K}_h$ be the available knowledge when reaching the horizon h. The definition $\mathsf{V}(\mathbf{K}_h) \equiv \mathsf{E}[\mathsf{I}(B)|\mathbf{K}_h]$ (39) and the chain rule for expectation (29) imply $\mathsf{E}[\mathsf{I}(B)] = \mathsf{E}[\mathsf{E}[\mathsf{I}(B)|\mathbf{K}_h]] \equiv \mathsf{E}[\mathsf{V}(\mathbf{K}_h)]$.
- The definition of minimum and Prop. 10 imply

$$\min_{(\mathsf{S}_t\colon Kt-1^\star \to A_t{}^\star)^\star_{t\in t^\star}} \mathsf{E}[\mathsf{V}(\mathbf{K}_h)] = \min_{\left(\mathsf{S}_t\colon \mathbf{K}^\star_{t-1} \to A_t{}^\star\right)^\star_{t<h}} \min_{\{\mathsf{S}_h\colon \mathbf{K}_{h-1}{}^\star \to A_h{}^\star\}} \mathsf{E}[\mathsf{V}(\mathbf{K}_h)] \quad (40)$$

$$\overbrace{\underbrace{=}_{(33)}}^{Prop.\ 10} \min_{\left(\mathsf{S}_t\colon \mathbf{K}^\star_{t-1} \to A_t{}^\star\right)^\star_{t<h}} \mathsf{E}\left[\min_{A_h \in A_h{}^\star} \mathsf{E}[\mathsf{V}(\mathbf{K}_h)|A_h, \mathbf{K}_{h-1}]\right].$$

- The definitions $\mathsf{V}(\mathbf{K}_{h-1}) \equiv \min_{A_h \in A^\star_h} \mathsf{E}[\mathsf{V}(\mathbf{K}_h)|A_h, \mathbf{K}_{h-1}]$ and ${}^o A_h = {}^o\mathsf{S}_h(\mathbf{K}_{h-1})$ verify the first backward step with the start (39).
- The next step becomes $\min_{\{\mathsf{S}_t\colon \mathbf{K}_{t-1} \to A_t{}^\star\}^\star_{t<h}} \mathsf{E}\left[\mathsf{V}(\mathbf{K}_{h-1})\right]$.

  Thus, the minimisation (40) is faced with the decreased horizon.
  The recursion prolongs until the optimal rule ${}^o\mathsf{S}_1$ is found. □

Value/Bellman/Loss-to-Go Function & Loss/Reward

Multitude of terminology occurs in connection with dynamic programming

- *value function* is a widely accepted name for the function $\mathsf{V}(\cdot)$ evolved in the general dynamic programming, Prop. 11.
- *Bellman function* is an alternative name of the value function.
- *loss-to-go* is another name of the value function↳. It is mostly used in the case with an additive performance index↳.
- These notions are also used in equivalent formulations of DM considering a reward for ordering of behaviours instead of the loss↳. Obviously, $\min_{\mathsf{S}\in\mathsf{S}^\star}$ loss↳ $= \max_{\mathsf{S}\in\mathsf{S}^\star}(-\text{loss}) \equiv \max_{\mathsf{S}\in\mathsf{S}^\star}\text{reward}$.

Optimal Design Needs Predictor of Observations

- The optimisation (38) needs to evaluate the expectations, $\forall t \in t^\star$,

$$\mathsf{E}[\mathsf{V}(\mathbf{K}_t)|A_t,\mathbf{K}_{t-1}] = \int_{\Delta_t{}^\star} \mathsf{V}(\underbrace{\Delta_t, A_t, \mathbf{K}_{t-1}}_{\mathbf{K}_t})\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1})\,\mathrm{d}\Delta_t.$$

- The freedom of users in the choice of the performance index↳ implies richness of the function set $\mathsf{V}^\star(\mathbf{K}_t) = \mathsf{V}^\star(\Delta_t, A_t, \mathbf{K}_{t-1})$. Thus, pds $(\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}))_{t\in t^\star}$ predicting the observable responses Δ_t of the system↳ on action↳s A_t under the knowledge↳ $\mathbf{K}_{t-1}$ are needed:
- *predictor* of observation↳s, used by the optimal design↳, is given by pds

$$(\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}))_{t\in t^\star}\,. \tag{41}$$

Instead of the term predictor, the next alternative notion is often used

- *predictive pd* refers to (41) stressing that the prediction concerns pds not only their characteristics as expectations or conditional covariance↳s.

Exercises 7 (On Dynamic Programming)

1. *Can you find optimal strategy↳ without dynamic programming? How?*
2. *Is the dynamic programming applicable in time-less DM tasks?*

   Think, for instance, about DM task concerning a static spatial object, say, a detection of a cat on an image.

   - *How you would handle them if you answer is no?*
   - *How you would handle them if you answer is yes?*
3. *With no uncertainty↳, DM is an ordinary optimisation. Still, it is useful to take it as a limit of stochastic DM. Why?*

Data-Driven Design of the Optimal Decision Strategy

Generally, the behaviour includes

- *hidden quantity* is a considered but never observed part of behaviour B.

  - Non-void data record $D = (\Delta, A)$=(observation,action) is always in B.
  - Availability of data record $D_t = (\Delta_t, A_t)$ moves it from ignorance $\mathbf{G}_t$ to knowledge $\mathbf{K}_t$ but hidden quantity stays in ignorance $\mathbf{G}_t$, $\forall t \in t^\star$.

The predictive pd is the only model needed if $\mathsf{I}(B) = \mathsf{I}\left(D^h\right)$, if we face

- *data-driven design* deals with a data-dependent performance index

$$\mathsf{I}(B) \equiv \mathsf{I}\big(\Delta^h, A^h\big) = \mathsf{I}\big(D^h\big) \equiv \mathsf{I}(\mathbf{K}_h). \tag{42}$$

- Hidden quantities influence the optimal design "only" through the terminal condition $\mathsf{V}(\mathbf{K}_h) = \mathsf{E}[\mathsf{I}(B)|\mathbf{K}_h]$ (39) of dynamic programming. Sec. 7 provides the model needed for evaluation of $\mathsf{V}(\mathbf{K}_h)$.

For $t < h$, $\mathsf{V}(\mathbf{K}_t)$ acts as a performance index in the data-driven design.

Proposition 12 (Data-Driven Design for Additive Performance Index)

The optimal admissible strategy ${}^{o}\mathsf{S} \equiv \left({}^{o}\mathsf{S}_t : \mathbf{K}^{\star}_{t-1} \to A_t{}^{\star} \right)_{t \in t^{\star}}$ *with knowledge* $\mathbf{K}_t{}^{\star} = D^{\star}_t \cup \mathbf{K}^{\star}_{t-1}$ *is searched for the data-driven design and*

- *additive performance index*

$$\mathsf{E}\left[\mathsf{I}\left(D^h\right)\right] \equiv \mathsf{E}\left[\sum_{t \in t^{\star}} \mathsf{z}(\Delta^t, A^t)\right] \equiv \mathsf{E}\left[\sum_{t \in t^{\star}} \mathsf{z}(D^t)\right] \equiv \mathsf{E}\left[\sum_{t \in t^{\star}} \mathsf{z}(\mathbf{K}_t)\right] \quad (43)$$

- *partial performance index is* $\mathsf{z}(\Delta^t, A^t) = \mathsf{z}(D^t) = \mathsf{z}(\mathbf{K}_t) \geq c > -\infty$.

The optimal strategy ${}^{o}\mathsf{S}$ *can be constructed in the value-wise way*

- $\forall\ \mathbf{K}_{t-1} \in \mathbf{K}^{\star}_{t-1}$, $t \in t^{\star}$, *a minimising argument* ${}^{o}A(\mathbf{K}_{t-1})$ *in*

$$\mathsf{V}(\mathbf{K}_{t-1}) = \min_{A_t \in A_t{}^{\star}} \mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}], \ \ t \in t^{\star}, \quad (44)$$

  is the optimal action, ${}^{o}A(\mathbf{K}_{t-1}) = {}^{o}\mathsf{S}_t(\mathbf{K}_{t-1})$.

- *The recursion (44) runs against the knowledge extension, starting from* $\mathsf{V}(\mathbf{K}_h) = 0$. *Under Agreement 4, the reached minimum is* $\mathsf{V}(\mathbf{K}_0)$.

Proof of Dynamic Programming: Data-Driven Additive Case

- Let $\mathbf{K}_h = D^h$ be the knowledge↳ available at the horizon↳ h. Let us define loss-to-go↳ $\mathsf{V}(\mathbf{K}_h) = 0$, for $t = h$ and for $t < h$,

$$\mathsf{V}(\mathbf{K}_{t-1}) \equiv \min_{\left(\mathsf{S}_\tau : \mathbf{K}^\star_{\tau-1} \to A_\tau{}^\star\right)^\star_{\tau \geq t}} \sum_{\tau=t}^{h} \mathsf{E}\left[\mathsf{z}(\mathbf{K}_\tau) | \mathbf{K}_{t-1}\right] \tag{45}$$

$$= \min_{\left(\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to A_t{}^\star\right)^\star} \mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | \mathbf{K}_{t-1}]$$

$$\overbrace{=}^{Prop.\ 10} \mathsf{E}\Big[\overbrace{\min_{A_t \in A_t{}^\star} \mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}]}^{\text{function of } \mathbf{K}_{t-1}} \Big| \mathbf{K}_{t-1}\Big]$$

$$\overbrace{=}^{(28)} \min_{A_t \in A_t{}^\star} \mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}], \text{ which proves (44).}$$

- The chain rule for expectations (29), the definition (45) for $t = 1$ & no influence of the chosen strategy on the prior knowledge $\mathbf{K}_0$ imply

$$\min_{\left(\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to A_t{}^\star\right)^\star_{t \in t^\star}} \mathsf{E}\left[\mathsf{E}\left[\sum_{t \in t^\star} \mathsf{z}(\mathbf{K}_t) \middle| \mathbf{K}_0\right]\right] = \mathsf{E}[\mathsf{V}(\mathbf{K}_0)] = \mathsf{V}(\mathbf{K}_0), \text{ Agr. 4.} \qquad \square$$

Example: Markov Decision Process (MDP, [122, 41])

- *MDP* is the DM task in which the observation $\Delta_{t-1} \in \Delta^\star$ forms the knowledge $\mathbf{K}_{t-1}$. It uses additive performance index given by the partial performance index $\mathsf{z}(\mathbf{K}_t) = \mathsf{z}(\Delta_t, A_t, \Delta_{t-1})$.

Algorithm 1 Dynamic Programming for MDP with $|A^\star|$, $|\Delta^\star|$, $h < \infty$

Inputs: cardinalities $|A^\star|, |\Delta^\star|$; horizon h, partial performance index $\mathsf{z}(\tilde{\Delta}, A, \Delta) = \mathsf{z}(\Delta_t = \tilde{\Delta}, A_t = A, \Delta_{t-1} = \Delta)$, predictors $\mathsf{f}_t(\tilde{\Delta}|A, \Delta)$

Evaluations:

set $\mathsf{V}_h(\tilde{\Delta}) = 1, \ \forall \tilde{\Delta} \in \Delta^\star$,

for $t = h, h-1, \ldots, 1$ **do**

  for $\Delta \in \Delta^\star$ **do**

    find $\,^{o}\mathsf{A}_t(\Delta) \in \mathrm{Arg}\min_{A \in A^\star} \mathsf{E}[\mathsf{z}(\tilde{\Delta}, A, \Delta) + \mathsf{V}_t(\tilde{\Delta})|A, \Delta]$

    set $\mathsf{V}_{t-1}(\Delta) = \mathsf{E}[\mathsf{z}(\tilde{\Delta}, \,^{o}\mathsf{A}_t(\Delta), \Delta) + \mathsf{V}_t(\tilde{\Delta})|\,^{o}\mathsf{A}_t(\Delta), \Delta]$

  end for

end for

Outputs: the optimal rules $\,^{o}\mathsf{S}_t(\,^{o}A_t(\Delta)|\Delta) = 1$, otherwise 0, $t \in t^\star$

Exercises 8 on MDP

Exercises 8 (On Markov Decision Process)

1. *Code Algorithm 1 in your favourite software or find it elsewhere. Play a bit with it.*
2. *Consider a finely discretised continuous real-valued function defined on three-dimensional unit box. Formulate the search for its minimum as Markov decision process.*
3. *For a very fine discretisation you will encounter problems known as curse of dimensionality, [10]. Where it manifests in your program?*
4. *Think over the foreseen problems even if you avoid programming.*
5. *Think over how to suppress the influence of the dimensionality curse.*

Additive Performance Index & Hidden $X^h = (X_t)_{t\in t^\star} \in \mathbf{G}_h$

Proposition 13 (General Design for Additive Performance Index)

The optimal admissible strategy ${}^o\mathsf{S} \equiv \left({}^o\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to A_t{}^\star\right)_{t\in t^\star}$ *with knowledge*↳ $\mathbf{K}_t{}^\star = D^\star_t \cup \mathbf{K}^\star_{t-1}$ *is searched for the optimal design*↳ *with*

- *general additive performance index*↳

$$\mathsf{E}\Big[\mathsf{I}\Big(X^h, D^h\Big)\Big] \equiv \mathsf{E}\Big[\sum_{t\in t^\star} \mathsf{z}(X^t, \mathbf{K}_t)\Big] \equiv \mathsf{E}\Big[\sum_{t\in t^\star} \mathsf{z}(X^t, D^t)\Big] \qquad (46)$$

- *general partial performance index* $\mathsf{z}(X^t, \mathbf{K}_t) = \mathsf{z}(X^t, D^t) \geq c > -\infty$.

The optimal strategy↳ ${}^o\mathsf{S}$ *can be constructed in the value-wise way*

- $\forall\ \mathbf{K}_{t-1} \in \mathbf{K}^\star_{t-1}$, $t \in t^\star$, *a minimising argument* ${}^oA(\mathbf{K}_{t-1})$ *in*

$$\mathsf{V}(\mathbf{K}_{t-1}) = \min_{A_t \in A_t{}^\star} \mathsf{E}[\mathsf{z}(X^t, \mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t)|A_t, \mathbf{K}_{t-1}],\ t \in t^\star, \qquad (47)$$

  is the optimal action, ${}^oA(\mathbf{K}_{t-1}) = {}^o\mathsf{S}_t(\mathbf{K}_{t-1})$.
- *The recursion (47) runs against the knowledge extension, starting from* $\mathsf{V}(\mathbf{K}_h) = 0$. *The reached minimum is* $\mathsf{E}[\mathsf{V}(\mathbf{K}_0)]$.

Proof of Dynamic Programming: General Additive Case

- Let $\mathbf{K}_h = D^h$ be the knowledge available at the horizon h and let us define loss-to-go: for $t = h$, $\mathsf{V}(\mathbf{K}_h) = 0$, while for $t < h$,

$$\mathsf{V}(\mathbf{K}_{t-1}) \equiv \min_{(\mathsf{S}_\tau:\, \mathbf{K}^\star_{\tau-1} \to A_\tau{}^\star)^\star_{\tau \geq t}} \sum_{\tau=t}^{h} \mathsf{E}\left[\mathsf{z}(X^\tau, \mathbf{K}_\tau) | \mathbf{K}_{t-1}\right] \tag{48}$$

$$= \min_{(\mathsf{S}_t:\, \mathbf{K}^\star_{t-1} \to A_t{}^\star)^\star} \mathsf{E}[\mathsf{z}(X^t, \mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | \mathbf{K}_{t-1}]$$

$$\underbrace{=}_{(33)} \mathsf{E}\Big[\underbrace{\min_{A_t \in A_t{}^\star} \mathsf{E}[\mathsf{z}(X^t, \mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}]}_{\text{function of } \mathbf{K}_{t-1}} \Big| \mathbf{K}_{t-1}\Big]$$

$$\overbrace{=}^{(28)} \min_{A_t \in A_t{}^\star} \mathsf{E}[\mathsf{z}(X^t, \mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}], \text{ which coincides with (47).}$$

- The chain rule for expectations (29), the definition (48) for $t = 1$ & no influence of the chosen strategy on the prior knowledge $\mathbf{K}_0$ imply

$$\min_{(\mathsf{S}_t:\, \mathbf{K}^\star_{t-1} \to A_t{}^\star)^\star_{t \in t^\star}} \mathsf{E}\left[\mathsf{E}\left[\sum_{t \in t^\star} \mathsf{z}(X^t, \mathbf{K}_t) \Big| \mathbf{K}_0\right]\right] = \mathsf{E}[\mathsf{V}(\mathbf{K}_0)] \qquad \square$$

On Additive Performance Index

- The definition of the partial performance index ↳

$$z(X^t, \mathbf{K}_t) = \begin{cases} \mathsf{I}(B) = \mathsf{I}(X^h, \mathbf{K}_h) & \text{if} \quad t = h \\ 0 & \text{otherwise} \end{cases}$$

  implies generality of the additive performance index.
- Any design requires the predictor ↳ (41) $(\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}))_{t \in t^\star}$, which suffices for the data-driven design ↳.
- The pds $(\mathsf{f}(X^t|A_t, \mathbf{K}_{t-1}))_{t \in t^\star}$ are needed if the general additive performance index ↳ is considered, see Prop. 13.

  They compress the knowledge about hidden X^t gained before time t.
- The value function ↳ $\mathsf{V}(\mathbf{K}_h) = \mathsf{E}[\mathsf{I}(B)|\mathbf{K}_h]$ (39) uses the pd $\mathsf{f}(X^h|\mathbf{K}_h)$.
- Construction of these pds, known as learning, is discussed in Section 7 after discussing asymptotic properties of the design ↳.

Addressed task

- Analyse sequence of DM designs with the horizon $h \to \infty$, i.e. with extending sets ${}^{h}t^{\star} \equiv \{1, \ldots, h\}$ of time indices.
  This analysis serves to an approximate design, cf. Sec. 32.

Conditions

- The data-driven design with an additive performance index (43) considered, while relying on existence of:
- *information state*, which is an observed finite-dimensional array $\mathcal{X}_{t-1}$ expressing sufficiently knowledge K_{t-1}, $\mathsf{E}[\bullet|A_t, \mathsf{K}_{t-1}] = \mathsf{E}[\bullet|A_t, \mathcal{X}_{t-1}]$.
  The partial performance index $\mathsf{z}(\mathcal{X}_t) \geq 0$ is assumed to depend only on the information state $\mathcal{X}_t$, giving

$$\mathsf{I}(\mathsf{K}_h) = \mathsf{I}(D^h) = \mathsf{I}(\Delta^h, A^h) = \sum_{t \in t^\star} \mathsf{z}(\mathcal{X}_t). \tag{49}$$

Stabilising Strategy

Asymptotic analysis makes sense if a solution of the DM design exists for the unbounded decision horizon. It needs existence of

- *stabilising strategy* consists of an infinite sequence of decision rules

$$(\mathsf{S}_t : \mathbf{K}_{t-1} \to A_t^{\star})_{t \in {}^{\infty}t^{\star} \equiv \{1,2,\ldots,\}}$$

for which a $C < \infty$ bounds the expected performance index

$$\mathsf{E}\left[\sum_{t=1}^{h} \mathsf{z}(\mathcal{X}_t)\right] \le hC, \; C < \infty, \; h \in {}^{\infty}t^{\star} \equiv \{1, 2, \ldots\}. \tag{50}$$

Remark 10 (Performance Index with Growing Horizon)

- *The expected performance index generically grows to infinity at least linearly with the growing decision horizon h.*
- *A change of finite number of decision rules forming the strategy has no influence on the expected performance index for $h \to \infty$.*

Asymptotically Optimal Strategy is Stationary

For $h \to \infty$, the influence of DM-rules' changes on the expected performance index diminishes and the optimal strategy is stationary

- *stationary strategy* repeatedly uses the same decision rule.

Proposition 14 (Asymptotic Design)

- *Let a stabilising strategy exist, with bounding $C < \infty$ and let partial performance index $\mathsf{z}(\mathcal{X}_t) \geq 0$, depend on a finite-dimensional information state $\mathcal{X}_t$.*

  Then, for $h \to \infty$, the non-unique optimal strategy is stationary.

- *Actions generated by the decision rule defining it are minimising arguments in the formal analogy of (44)*

$$^{\infty}\mathsf{V}(\mathcal{X}_{t-1}) + {}^{\infty}C = \min_{A_t \in A_t^{\star}} \mathsf{E}\left[\mathsf{z}(\mathcal{X}_t) + {}^{\infty}\mathsf{V}(\mathcal{X}_t) \middle| A_t, \mathcal{X}_{t-1}\right] \qquad (51)$$

with a constant $^{\infty}C \leq C$ & a time-invariant value function $^{\infty}\mathsf{V}(\mathcal{X})$.

Proof

- For horizon $h < \infty$, let us denote ${}^{h}\tilde{V}(\mathbf{K}_{t-1}) \equiv {}^{h}\tilde{V}(\mathcal{X}_{t-1})$ the optimal loss-to-go and ${}^{h}C$ as the smallest value for which expectation of

modified loss-to-go $\qquad {}^{h}V(\mathcal{X}_t) \equiv {}^{h}\tilde{V}(\mathcal{X}_t) - (h-t)\,{}^{h}C \qquad (52)$

is bounded from above for $h \to \infty$ and a fixed $t \in {}^{\infty}t^{\star}$, $\mathcal{X}_t \in \mathcal{X}^{\star}$.

- The optimal strategy has a lower expected performance index than any stabilising strategy. Thus, ${}^{h}C \leq C$ and $\overline{\lim}_{h\to\infty}\,{}^{h}C = {}^{\infty}C$ exists.
- The optimisation is uninfluenced if we subtract the value ${}^{h}C$ from each partial performance index and the modified loss-to-go fulfills

$$ {}^{h}V(\mathcal{X}_{t-1}) + {}^{h}C = \min_{A_t \in A_t{}^{\star}} \mathsf{E}\left[z(\mathcal{X}_t) + {}^{h}V(\mathcal{X}_t) | A_t, \mathcal{X}_{t-1} \right]. \qquad (53) $$

- For arbitrary fixed $t \leq h, \mathcal{X}_t$, the modified loss-to-go ${}^{h}V(\mathcal{X}_t)$ (52) is bounded difference of a pair of h-index monotonous sequences. Thus, a finite ${}^{\infty}V(\mathcal{X}_t) = \lim_{h\to\infty} {}^{h}V(\mathcal{X}_t)$ exists and meets (53) for $h \to \infty$.
- The same optimisation is performed for each $t < \infty$ providing the same decision rule. Thus, the optimal strategy is a stationary one. $\square$

- Solutions of the Bellman equation for a growing horizon h represent successive approximations for solving its stationary version (51).
- Algorithm 2 provides an alternative way. It is faster than successive approximations but its steps are more demanding, [116].

Algorithm 2 Iterations in the Strategy Space

Input partial performance index $\mathsf{z}(\mathcal{X})$, model $\mathsf{f}(\tilde{\mathcal{X}}|A,\mathcal{X})$ of the information state, and a decision rule S of an stabilising stationary strategy

while the guaranteed convergence is not encountered **do**

Solve the linear equation for the function $\mathsf{V}(\mathcal{X})$ and constant C

$$\mathsf{V}(\mathcal{X}) + C = \mathsf{E}[\mathsf{z}(\tilde{\mathcal{X}}) + \mathsf{V}(\tilde{\mathcal{X}})|\mathsf{S}(\mathcal{X}),\mathcal{X}]. \tag{54}$$

Evaluate the new stabilising stabilising rule

$$\mathsf{S}(\mathcal{X}) \in \mathrm{Arg}\min_{A\in A^\star} \mathsf{E}[\mathsf{z}(\tilde{\mathcal{X}}) + \mathsf{V}(\tilde{\mathcal{X}})|A,\mathcal{X}] \text{ for V solving (54)}$$

end while

Output the optimal stationary rule $\mathsf{S}(\mathcal{X})$, the minimum per step C

Stationary Markov Decision Process

Example 10 (Stationary MDP – aka Controlled Markov Chain [116])

- *Consider the data-driven design with observation $\Delta^\star$ and action $A^\star$ spaces having finite cardinalities.*
- *A time-invariant partial performance index $\mathsf{z}(\Delta_t, A_t)$, $\Delta_t \in \Delta^\star$, $A_t \in A^\star$, a finite table, determines the additive performance index.*

The information state is observation $\mathsf{E}[\bullet|A_t, \mathbf{K}_{t-1}] = \mathsf{E}[\bullet|A_t, \Delta_{t-1}]$.

- *Props. 12, 14 directly imply that the loss-to-go is a finite time-invariant table* $\mathsf{V}(\Delta_t)$, $\Delta_t \in \Delta^\star$.
- *The optimal decision rule* $\mathsf{f}(A_t|\mathbf{K}_{t-1}) = \mathsf{f}(A_t|\Delta_{t-1})$ *of the stationary strategy concentrates on minimising argument* ${}^{o}A_t = {}^{o}A(\Delta_{t-1})$ *in*

$$\mathsf{V}(\Delta_{t-1}) + C = \min_{A \in A^\star} \mathsf{E}[\mathsf{z}(\Delta_t, A_t) + \mathsf{V}(\Delta_t)|A_t, \Delta_{t-1}], \ \forall \Delta_{t-1} \in \Delta^\star$$

Practical Examples Use Safe Handling Squared Norms

- *Squared norm* $||x||^2 = x'x$, $'$ transposes the column vector x.

Proposition 15 (Completion of Squares)

Let $x \in x^\star, y \in y^\star$ with fixed-dimensional real $x^\star, y^\star$. Let real matrices $\mathbb{U}, \mathbb{V}, \mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}$ on left-hand side of (55) are compatible.

$$\phi(x,y) \equiv \left|\left|\mathbb{U}[\mathbb{A},\mathbb{B}]\begin{bmatrix} x \\ y \end{bmatrix}\right|\right|^2 + \left|\left|\mathbb{V}[\mathbb{C},\mathbb{D}]\begin{bmatrix} x \\ y \end{bmatrix}\right|\right|^2 = \left|\left|\mathbb{S}\begin{bmatrix} x \\ y \end{bmatrix}\right|\right|^2. \quad (55)$$

Then, the upper triangular real matrix $\mathbb{W}$ exists meeting (55) on $x^\star$, $y^\star$

$$\mathbb{W} = \begin{bmatrix} \mathbb{W}_{x^\star} & \mathbb{W}_{x^\star y^\star} \\ 0 & \mathbb{W}_{y^\star} \end{bmatrix}, \ \dim(\mathbb{W}_{x^\star}) = (\ell_{x^\star}, \ell_{x^\star}), \dim(\mathbb{W}_{y^\star}) = (\ell_{y^\star}, \ell_{y^\star})$$

$$\Rightarrow \ {}^o x \equiv -\mathbb{W}_{x^\star}^{-1}\mathbb{W}_{x^\star y^\star} y \in \operatorname{Arg}\min_{x \in x^\star} \phi(x,y) \quad (56)$$

Proof $\phi(x,y) \equiv \left|\left|\begin{bmatrix} \text{unitary transformation} \\ \text{mapping rectangular} \\ \text{on triangular matrix} \end{bmatrix}\begin{bmatrix} \mathbb{U}[\mathbb{A},\mathbb{B}] \\ \mathbb{V}[\mathbb{C},\mathbb{D}] \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right|\right|^2$ □

Example 11 (Linear Quadratic Data-Driven Design)

- *Quadratic partial performance index ↳ in the additive loss is considered*

$$z(\Delta_t, A_t) = || \sqrt{\mathbb{Q}}_{\Delta^\star} \Delta_t||^2 + || \sqrt{\mathbb{Q}}_{A^\star} A_t||^2, \text{ where} \tag{57}$$

  - *observation ↳ Δ_t is $\ell_{\Delta^\star}$-dimensional real column vector*
  - *action ↳ is A_t is $\ell_{A^\star}$-dimensional real column vector*
  - $\sqrt{\mathbb{Q}}_{\Delta^\star}$*, regular* $\sqrt{\mathbb{Q}}_{A^\star}$ *are square-roots of penalisation matrices*

This performance index ↳ expresses aim ↳ to push Δ_t and A_t to zero. The weights $\sqrt{\mathbb{Q}}_{\Delta^\star}$, $\sqrt{\mathbb{Q}}_{A^\star}$ define significance of deviations from this target.

- *The conditional moments of observations are given by $\mathbf{K}_{t-1} = \Delta_{t-1}$*

$$\mathsf{E}[\Delta_t|A_t, \Delta_{t-1}] = \mathbb{A}\Delta_{t-1} + \mathbb{B}A_t, \ \mathsf{cov}[\Delta_t|A_t, \Delta_{t-1}] = \sqrt{\mathbb{R}}' \sqrt{\mathbb{R}} \geq 0, \tag{58}$$

  where matrices $\mathbb{A}$, $\mathbb{B}$ of appropriate dimensions are known.

The possibly unknown covariance matrix $\sqrt{\mathbb{R}}' \sqrt{\mathbb{R}}$ is independent of data.

Design of Optimal Strategy in LQ Data-Driven Case

In the use of Prop. 12, the value function with $\mathbf{K}_t = \Delta_t$ is quadratic

$$\mathsf{V}(K_t) = \mathsf{V}(\Delta_t) = \gamma_t + ||\sqrt{}\mathbb{S}_t\Delta_t||^2,\ \gamma_t \geq 0,\ \ \sqrt{}\mathbb{S}_t'\sqrt{}\mathbb{S}_t \geq 0. \qquad (59)$$

At $t = h =$ horizon $\mathsf{V}(\Delta_h) = 0$. Thus, (59) holds with $\sqrt{}\mathbb{S}_h = 0,\ \gamma_h = 0$.

- In an inductive step $t < h$, dynamic programming (44) evaluates

$$\mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t)|A_t, \mathbf{K}_{t-1}] = \left\{ \sqrt{}\tilde{\mathbb{S}}_t'\tilde{\mathbb{S}}_t \equiv \sqrt{}\mathbb{S}_t'\sqrt{}\mathbb{S}_t + \sqrt{}\mathbb{Q}_{\Delta^\star}'\sqrt{}\mathbb{Q}_{\Delta^\star} \right\}$$

$$= \gamma_t + \mathsf{E}[||\sqrt{}\tilde{\mathbb{S}}_t\Delta_t||^2|A_t, \mathbf{K}_{t-1}] + ||\sqrt{}\mathbb{Q}_{A^\star}A_t||^2$$

$$\overbrace{=}^{(58)} \overbrace{\gamma_t + \mathrm{tr}[\sqrt{}\tilde{\mathbb{S}}_t'\sqrt{}\tilde{\mathbb{S}}_t\sqrt{}\mathbb{R}'\sqrt{}\mathbb{R}]}^{\gamma_{t-1}} + ||\sqrt{}\tilde{\mathbb{S}}_t(\mathbb{A}\Delta_{t-1} + \mathbb{B}A_t)||^2 + ||\sqrt{}\mathbb{Q}_{A^\star}A_t||^2$$

$$= \left|\left| \begin{bmatrix} \sqrt{}\mathbb{S}_{A^\star} & \sqrt{}\mathbb{S}_{\Delta^\star} \\ 0 & \sqrt{}\mathbb{S}_{t-1} \end{bmatrix} \begin{bmatrix} A_t \\ \Delta_{t-1} \end{bmatrix} \right|\right|^2 \text{ is smallest if } A\text{-dependent part} = 0.$$

- The minimiser $A_t = {}^{o}A(\Delta_{t-1})$ gives the optimal linear decision rule

$${}^{o}A(\mathbf{K}_{t-1}) = -\sqrt{}\mathbb{S}_{A^\star}^{-1}\sqrt{}\mathbb{S}_{\Delta^\star}\Delta_{t-1} = {}^{o}\mathbb{L}_t'\Delta_{t-1}\ \&\ \sqrt{}\mathbb{S}_{t-1} \text{ in (59) at } t-1. \qquad (60)$$

Remarks on Linear Quadratic Design

- LQ problem is simple as the functional recursion of dynamic programming, Prop. 12, for the value function reduces to the algebraic recursion on square-root ${}^{\surd}\mathbb{S}_t$ of the Riccati matrix.

  It can be evaluated without having past observations.

- This is one a few cases in which only conditional moments are needed.

- Properties of the solution depends on configurations of matrices $\mathbb{A}$, $\mathbb{B}$, ${}^{\surd}\mathbb{Q}_{\Delta^\star}$, ${}^{\surd}\mathbb{Q}_{A^\star}$ whole art concerning notions like controllability, eigenvalue assignment and many others notions [3, 120].

  It needs a nice and difficult analysis of pairs of matrices.

- The involved matrices may depend on time. This fact is exploited when LQ design results as an approximation of non-linear cases.

- Asymptotic analysis, Prop. 14, implies that exists $\sqrt{\mathbb{S}}_\infty = \lim_{t\to\infty} \sqrt{\mathbb{S}}_t$ if the stabilising strategy exists. The optimal strategy is the time-invariant feedback ${}^o\!A(\mathbf{K}_{t-1}) = -\mathbb{L}'_\infty \Delta_{t-1}$ (60).
- The Riccati equation (60) can be given the form $\sqrt{\mathbb{S}}'_{t-1} \sqrt{\mathbb{S}}_{t-1} =$

$$
\begin{aligned}
&= (\mathbb{A} - \mathbb{B}\, {}^o\mathbb{L}'_{t-1})'(\sqrt{\mathbb{S}}'_t \sqrt{\mathbb{S}}_t + \sqrt{\mathbb{Q}}'_{\Delta^\star} \sqrt{\mathbb{Q}}_{\Delta^\star}) \overbrace{(\mathbb{A} - \mathbb{B}\, {}^o\mathbb{L}'_{t-1})}^{\textit{closed-loop transition matrix}} \qquad (61)\\
&+ {}^o\mathbb{L}_{t-1}\mathbb{Q}_{A^\star}\, {}^o\mathbb{L}'_{t-1} \equiv \mathsf{R}(\sqrt{\mathbb{S}}'_t \sqrt{\mathbb{S}}_t, {}^o\mathbb{L}_\infty)
\end{aligned}
$$

Algorithm 3 Iterations in strategy space for LQ Design

Input a stabilising $\mathbb{L}$ that leaves eigenvalues of $\mathbb{A} - \mathbb{B}\mathbb{L}'$ in unit circle
while the guaranteed convergence is not encountered **do**
  Solve the linear matrix equation $\mathbb{S} = \mathsf{R}(\mathbb{S}, \mathbb{L})$ (61) for $\mathbb{S} = \sqrt{\mathbb{S}}' \sqrt{\mathbb{S}}$.
  Evaluate the new (inevitably stabilising) minimiser $\mathbb{L}$
end while
Output the optimal feedback $\mathbb{L}$, the minimum $\mathrm{tr}(\sqrt{\mathbb{R}}' \sqrt{\mathbb{R}} \sqrt{\tilde{\mathbb{S}}}' \sqrt{\tilde{\mathbb{S}}})$

Exercises 9 (On Design)

1. *Provide an example of a system ↳ lacking stabilising strategy ↳.*
2. *Consider an unstable linear system ↳, which blows up for zero actions. Can it be stabilised?*
3. *Design the optimal stationary strategy ↳ for for dishonest-coin-tossing problem, Exercises 4.*
4. *Your home heating system is controlled by setting heating on or off according to the available indication whether the room temperature is below, in or above the desired range.*
   Formulate the optimal heating as MDP ↳ and design the optimal stationary strategy.
5. *Perform LQ design without using square roots. Recognise the numerically sensitive operation(s).*

Learning

Why and When Learning Is Needed?

$\boxed{\text{Learning is needed iff behaviour } B \in B^{\star} \text{ includes a hidden quantity}_{\hookrightarrow}\ X}$.

It is never directly observed but it influences preferential ordering↳.

$\boxed{\text{Learning provides pds used by dynamic programming}}$ see Prop. 11

- the predictor↳ (41) $(\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}))_{t\in t^{\star}}$ generally needed and suffices for the data-driven design↳.
- the pds $(\mathsf{f}(X^t|A_t, \mathbf{K}_{t-1}))_{t\in t^{\star}}$, compressing the knowledge on hidden quantities. The design↳ with the general additive performance index↳ needs them, see Prop. 13.

This part gives the predictor↳ & accumulated knowledge ↳ on hidden quantities.

- The solution, known as Bayesian filtering↳ [63], is of an independent interest. It is a consistent formal prescriptive way of learning.

The closed-loop pd $\mathsf{f}_{\mathsf{S}}(B) \equiv \mathsf{f}_{\mathsf{S}}(B|\mathbf{K}_0)$ models behaviour $B = (X^h, \Delta^h, A^h)$.

$$\mathsf{f}_{\mathsf{S}}(B) = \prod_{t \in t^\star} \mathsf{f}(\Delta_t|X^t, A_t, \mathbf{K}_{t-1}) \mathsf{f}(X_t|X^{t-1}, A_t, \mathbf{K}_{t-1}) \mathsf{f}(A_t|X^{t-1}, \mathbf{K}_{t-1}) \mathsf{f}(X_0|\mathbf{K}_0).$$

The factors' interpretation and adopted assumptions are

- *observation model* relates the observation Δ_t to hidden quantities X^t, to the action A_t and the knowledge $\mathbf{K}_{t-1}$

$$\left(\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) \equiv \mathsf{f}(\Delta_t|X^t, A_t, \mathbf{K}_{t-1})\right)_{t \in t^\star}. \tag{62}$$

This model contains an hidden quantity $(X_t \in X_t^\star \subset \mathbf{G}_\tau^\star)_{\forall \tau \in t^\star}$ unlike the predictor.

- *time-evolution model* interrelates the hidden quantities $X^h \in X^{h\star}$

$$\left(\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1}) \equiv \mathsf{f}(X_t|X^{t-1}, A_t, \mathbf{K}_{t-1})\right)_{t \in t^\star}. \tag{63}$$

The assumed independence (62), (63) is met by the re-definition $X_t \equiv X^t$.

The strategy is described by

- *strategy model*, consists of the pds $\left(f(A_t|X^{t-1}, \mathbf{K}_{t-1})\right)_{t\in t^\star}$.

The last factor, which initiates the filtering, is

- *prior pd* $f(X_0)$ expresses the prior knowledge $\mathbf{K}_0$ about X_0

$$f(X_0) \equiv f(X_0|\mathbf{K}_0). \tag{64}$$

- *modelling* gives observation model, time-evolution model & prior pd.

The designed strategy fulfils informational constraints and meets

- *natural conditions of DM* , which formalise that X^h are unknown to the designed strategies [132]. They postulate independence of A_t and X^{t-1} when conditioned on the knowledge $\mathbf{K}_{t-1}$

$$f(A_t|\mathbf{K}_{t-1}) = f(A_t|X^{t-1}, \mathbf{K}_{t-1}) \overbrace{\Leftrightarrow}^{\text{Prop. 6}} f(X^{t-1}|A_t, \mathbf{K}_{t-1}) = f(X^{t-1}|\mathbf{K}_{t-1}). \tag{65}$$

Are Natural Conditions of DM Natural?

Remark 11 (Natural Conditions of DM)

- *The natural conditions of DM↳ express the assumption that $X_t \notin \mathbf{K}_\tau\ \forall\tau,\ t \in t^\star$. Thus, values of X^h cannot be used by the decision rules forming the admissible strategy↳.*

  Alternatively, knowledge about X^{t-1} cannot be gained from the action↳ A_t without knowing the system↳ reaction, the observation↳ Δ_t.

- *The natural conditions of DM↳ are "naturally" fulfilled by strategies we are designing. They have to be checked when the data record↳s influenced by an "externally chosen" strategy↳ are processed.*

- *The hidden quantities X_τ, $\tau \geq t$, can be influenced by A_t.*

- *The hidden quantity↳ and the action↳ often determine the observation and the next hidden quantity. Then, $\mathbf{K}_{t-1}$ is omitted.*

Proposition 16 (Bayesian Prediction and Filtering)

The predictor↳ (41) reads

$$f(\Delta_t|A_t, \mathbf{K}_{t-1}) = \int_{X_t^\star} f(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) f(X_t|A_t, \mathbf{K}_{t-1}) dX_t. \quad (66)$$

Under natural conditions of DM↳, it uses:

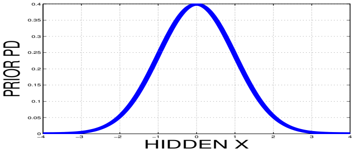
- *filtering means the evolution of the pd* $f(X_t|A_t, \mathbf{K}_{t-1})$ *from the prior pd↳* $f(X_0)$*. The updating consists of the evolution pair*
- *data updating extends the knowledge* $\mathbf{K}_{t-1}$ *by the data record↳* $D_t = (\Delta_t, A_t) =$ *(observation↳,action↳)*

$$\begin{aligned} f(X_t|\mathbf{K}_t) &= \frac{f(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) f(X_t|A_t, \mathbf{K}_{t-1})}{f(\Delta_t|A_t, \mathbf{K}_{t-1})} \\ &\propto f(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) f(X_t|A_t, \mathbf{K}_{t-1}). \end{aligned} \quad (67)$$

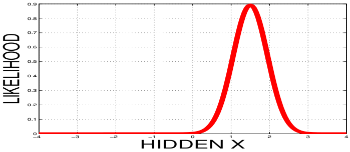
- *time updating reflects the evolution* $X_t \to X_{t+1}$*, if* A_{t+1} *is applied,*

$$f(X_{t+1}|A_{t+1}, \mathbf{K}_t) = \int_{X_t^\star} f(X_{t+1}|X_t, A_{t+1}, \mathbf{K}_t) f(X_t|\mathbf{K}_t)\, \mathrm{d}X_t. \quad (68)$$

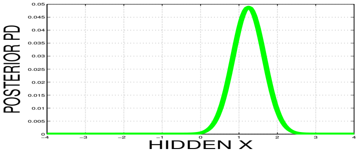
Transformations in One Complete Filtering Step



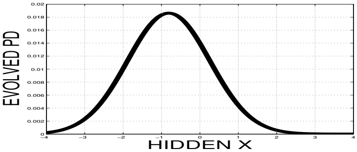
$\mathrm{f}(X_t|A_t, \mathbf{K}_{t-1})$
(pd after previous steps)



$\mathrm{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})$
(observation model with $\mathbf{K}_t$ fixed)



$\mathrm{f}(X_t|\mathbf{K}_t)$
(pd after data updating)



$\mathrm{f}(X_{t+1}|A_{t+1}, \mathbf{K}_t)$
(pd after time updating)

Proof of the Bayesian Prediction and Filtering

- The marginalisation, chain rule and Prop. 6 imply

$$\begin{aligned}
\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}) &= \int_{X_t{}^\star} \mathsf{f}(\Delta_t, X_t|A_t,\mathbf{K}_{t-1})\,\mathrm{d}X_t \\
&= \int_{X_t{}^\star} \mathsf{f}(\Delta_t|X_t, A_t,\mathbf{K}_{t-1})\mathsf{f}(X_t|A_t,\mathbf{K}_{t-1})\,\mathrm{d}X_t.
\end{aligned}$$

- The data updating is Bayes' rule that uses $\mathbf{K}_t = (\Delta_t, A_t, \mathbf{K}_{t-1})$

$$\mathsf{f}(X_t|\mathbf{K}_t) \propto \mathsf{f}(\Delta_t|X_t, A_t,\mathbf{K}_{t-1})\mathsf{f}(X_t|A_t,\mathbf{K}_{t-1}).$$

- The time updating results from the marginalisation, the chain rule, and the natural conditions of DM giving $\mathsf{f}(X_t|A_{t+1},\mathbf{K}_t) = \mathsf{f}(X_t|\mathbf{K}_t)$

$$\mathsf{f}(X_{t+1}|A_{t+1},\mathbf{K}_t) = \overbrace{\int_{X_t^\star} \underbrace{\mathsf{f}(X_{t+1}|X_t, A_{t+1},\mathbf{K}_t)\,\underbrace{\mathsf{f}(X_t|A_{t+1},\mathbf{K}_t)}_{\text{superfluous } A_{t+1}}}_{\text{chain rule}}\,\mathrm{d}X_t}^{\text{marginalisation}} \qquad \square$$

Example 12 (Prediction & Filtering for Finite Amount of Behaviours)

- *The sets of observations $\Delta^\star$, actions $A^\star$ and hidden quantities $X^\star$ have finite cardinalities.*
- *Prior pd $\mathsf{f}(X_0|\mathbf{K}_0)$, a past-data-independent time-evolution model $\mathsf{f}(X_t|A_t, X_{t-1})$ and an observation model $\mathsf{f}(\Delta_t|A_t, X_t)$ are given. They are tables describing pds with respect to a counting measure.*

The predictive pd and posterior pd evolve for $t \in t^\star$ as follow (Prop. 16)

$$
\begin{aligned}
\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}) &= \sum_{X_t\in X^\star} \mathsf{f}(\Delta_t|A_t,X_t)\mathsf{f}(X_t|A_t,\mathbf{K}_{t-1}) && \textit{(predictive pd)}\\
\mathsf{f}(X_t|\mathbf{K}_t) &= \frac{\mathsf{f}(\Delta_t|A_t,X_t)\mathsf{f}(X_t|A_t,\mathbf{K}_{t-1})}{\sum_{X_t\in X^\star}\mathsf{f}(\Delta_t|A_t,X_t)\mathsf{f}(X_t|A_t,\mathbf{K}_{t-1})} && \textit{(data updating)}\\
\mathsf{f}(X_{t+1}|A_{t+1},\mathbf{K}_t) &= \sum_{X_t\in X^\star} \mathsf{f}(X_{t+1}|A_{t+1},X_t)\mathsf{f}(X_t|\mathbf{K}_t) && \textit{(time updating)}.
\end{aligned}
$$

Exercises 10 (On Filtering)

1. *Derive the pd* $\mathrm{f}(X^h|\mathbf{K}_h)$ *needed in (39) for initiating dynamic programming described in Prop. 16.*
2. *Perform prediction and filtering for the time-invariant Markov chain with* $X^\star = \Delta^\star = \{0,1\}$, $A^\star = \emptyset$, *the observation model and time-evolution model given by* $\alpha, \beta, a, b \in [0,1]$

$\mathrm{f}(\Delta_t\|X_t)$	$\Delta_t = 0$	$\Delta_t = 1$
$X_t = 0$	α	$1-\alpha$
$X_t = 1$	β	$1-\beta$

$\mathrm{f}(X_{t+1}\|X_t)$	$X_{t+1} = 0$	$X_{t+1} = 1$
$X_t = 0$	a	$1-a$
$X_t = 1$	b	$1-b$

3. *Complexity of the discrete-valued case grows with* $|X^\star|$ *and* $|A^\star|$. *Can it be decreased when many transitions have zero probability?*
4. *How is the previous problem related with Bayesian networks [64, 65]?*

Recall

- *statistic* is a measurable mapping of the knowledge ↳, $\mathsf{W} : \mathbf{K}^\star \to \mathsf{W}^\star$. Notice that no conditions are put on the set $\mathsf{W}^\star$!
- *sufficient statistic* W fulfills $\mathsf{f}(\Delta_{t+1}|A_t, \mathbf{K}_t) = \mathsf{f}(\Delta_{t+1}|A_{t+1}, \mathsf{W}(\mathbf{K}_t))$, $\mathsf{f}(X_t|\mathbf{K}_t) = \mathsf{f}(X_t|\mathsf{W}(\mathbf{K}_t))$: it suffices to store the statistic values.

Implications

- The described prediction and filtering ↳ provide statistic. They simply accumulate and ideally compress the knowledge ↳.

  Prediction and filtering are mostly understood as specific DM ↳ tasks giving estimates of the future observation and hidden quantities, [63].

A non-bijective transformation of the statistic causes the knowledge loss!

- Dynamic programming, Prop. 11, implies that pds resulting from filtering ↳ form sufficient statistic ↳ for solving dynamic DM ↳ tasks.

Kalman Filtering: An Example of Bayesian Filtering

- *Kalman filter* yields conditional moments fully describing pds gained by the Bayesian filtering for linear normal systems.
- It is one of a few cases allowing the exact Bayesian filtering.
- Kalman filter ↳ is the "machine" computing posterior pds. Often, the gained moments are wrongly understood as point state estimates.
- It has an extreme application width: technological processes, satellites, weapons, traffic control, economy and ecology.
- It underlies approximate filtering techniques like extended Kalman filtering, unscented filtering, particle filtering, ensemble filtering.
- Reference samples are [2, 33, 49, 63, 66, 29, 140]

DM Elements Used in Kalman Filtering

- Aim to process data to get pds needed for the optimal design
- Actions in $A^\star = \{\text{fixed-dimensional real vectors}\}$ entering the system are generated by strategy meeting natural conditions of DM
- Observations belong to $\Delta^\star = \{\text{fixed-dimensional real vectors}\}$
- Hidden quantities are states $X \in X^\star = \{\text{fixed-dimensional real vectors}\}$
- Behaviour $B = (X_0^h, \Delta^h, A^h, \mathbf{K}_0)$
- Ignorance $\mathbf{G}_t = (X_0^h, \Delta_t^h, A_{t+1}^h)$
- Knowledge $\mathbf{K}_{t-1} = (\Delta^{t-1}, A^{t-1}, \mathbf{K}_0)$
- *normal pd* of $X \in X^\star$ reads

$$\mathcal{N}_X(\bar{X}, \mathbb{R}_{X^\star}) = |2\pi\mathbb{R}_{X^\star}|^{-0.5} \exp[-0.5(X - \bar{X})'\mathbb{R}_{X^\star}^{-1}(X - \bar{X})]. \quad (69)$$

- *linear normal system model* (LN) has

✓ the observation & state expectations linear in realised behaviour

✓ behaviour-independent covariances, and

✓ normal observation & time-evolution models.

Specification of Linear-Normal (LN) System Model

- Prior knowledge $\mathbf{K}_0$ specifies the system model $\hookrightarrow$ M

$$
\begin{array}{rl}
\text{Evolution pd} & \mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1}) = \mathcal{N}_{X_t}(\mathbb{A}X_{t-1} + \mathbb{B}A_t, \mathbb{R}_{X^\star}) \\
\text{Observation pd} & \mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) = \mathcal{N}_{\Delta_t}(\mathbb{C}X_t + \mathbb{D}A_t, \mathbb{R}_{\Delta^\star}) \\
\text{Prior pd} & \mathsf{f}(X_0) = \mathcal{N}_{X_0}(\hat{X}_{0|0}, \mathbb{P}_{0|0}).
\end{array}
\tag{70}
$$

$$
\mathbf{K}_0 = \{\text{matrices } \mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}, \mathbb{R}_{X^\star} > 0, \mathbb{R}_{\Delta^\star} > 0, \mathbb{P}_{0|0} > 0 \\
\text{vector } \hat{X}_{0|0}, \text{ a bit algebra, filtering theory}\}.
$$

- All matrices may depend on $t, A_t, \mathbf{K}_{t-1}$ if the dependence is known.

Example 13 (Kalman Filtering)

Under natural conditions of DM & the prior knowledge (70), the predictor is

$$f(\Delta_t|A_t, \mathbf{K}_{t-1}) = \mathcal{N}_{\Delta_t}(\hat{\Delta}_{t|t-1}, \mathbb{R}_{\Delta^\star t|t-1}) \tag{71}$$

$$\begin{aligned} \hat{\Delta}_{t|t-1} &\equiv \mathbb{C}\hat{X}_{t|t-1} + \mathbb{D}A_t,\ \mathbb{R}_{\Delta^\star t|t-1} \equiv \mathbb{R}_{\Delta^\star} + \mathbb{C}\mathbb{P}_{t|t-1}\mathbb{C}' \\ f(X_t|\mathbf{K}_t) &= \mathcal{N}_{X_t}(\hat{X}_{t|t}, \mathbb{P}_{t|t}), \quad f(X_{t+1}|A_{t+1}\mathbf{K}_t) = \mathcal{N}_{X_{t+1}}(\hat{X}_{t+1|t}, \mathbb{P}_{t+1|t}). \end{aligned}$$

- *data updating reduces to*

$$\begin{aligned} \hat{X}_{t|t} &= \hat{X}_{t|t-1} + \mathbb{G}_t \overbrace{(\Delta_t - \mathbb{C}\hat{X}_{t|t-1} - \mathbb{D}A_t)}^{\varepsilon_t,\ \textit{prediction error}} \\ \mathbb{G}_t &\equiv \mathbb{P}_{t|t-1}\mathbb{C}'(\mathbb{C}\mathbb{P}_{t|t-1}\mathbb{C}' + \mathbb{R}_{\Delta^\star})^{-1} \ldots\ldots \textit{ Kalman gain} \\ \mathbb{P}_{t|t} &= \mathbb{P}_{t|t-1} - \mathbb{P}_{t|t-1}\mathbb{C}'(\mathbb{C}\mathbb{P}_{t|t-1}\mathbb{C}' + \mathbb{R}_{\Delta^\star})^{-1}\mathbb{C}\mathbb{P}_{t|t-1} \end{aligned} \tag{72}$$

- *time updating* $\hat{X}_{t+1|t} = \mathbb{A}\hat{X}_{t|t} + \mathbb{B}A_{t+1},\ \mathbb{P}_{t+1|t} = \mathbb{A}\mathbb{P}_{t|t}\mathbb{A}' + \mathbb{R}_{X^\star}$ (73)
- *The recursion starts from the prior statistics* $\hat{X}_{0|0}$, $\mathbb{P}_{0|0}$.

Simplified notation used $X = X_{t-1}, A = A_t, \hat{X} = \hat{X}_{t|t-1}, \mathbb{P} = \mathbb{P}_{t|t-1}$.

$$\mathrm{f}(\Delta|A, \mathbf{K}_{t-1}) = \int_{X^\star} \mathcal{N}_\Delta(\mathbb{C}X + \mathbb{D}A, \mathbb{R}_{\Delta^\star}) \mathcal{N}_X(\hat{X}, \mathbb{P})\, \mathrm{d}X$$

$$\propto \int_{X^\star} \exp\left\{ -0.5 \underbrace{\left[\underbrace{(\Delta - \mathbb{C}X - \mathbb{D}A)'}_{\mathrm{a}(X)} \mathbb{R}_{\Delta^\star}^{-1} \mathrm{a}(X) + \underbrace{(X - \hat{X})'}_{\mathrm{b}(X)} \mathbb{P}^{-1} \mathrm{b}(X) \right]}_{\mathrm{q}(X)} \right\} \mathrm{d}X$$

$$\mathrm{q}(X) = X' \underbrace{\left(\mathbb{C}'\mathbb{R}_{\Delta^\star}^{-1}\mathbb{C} + \mathbb{P}^{-1}\right)}_{\mathbb{P}_{t|t}^{-1}} X - 2X'\mathbb{P}_{t|t}^{-1} \underbrace{\mathbb{P}_{t|t}\left[\mathbb{C}'\mathbb{R}_{\Delta^\star}^{-1} \underbrace{(\Delta - \mathbb{D}A)}_{\delta} + \mathbb{P}^{-1}\hat{X}\right]}_{\hat{X}_{t|t}}$$

$$+\hat{X}_{t|t}'\mathbb{P}_{t|t}^{-1}\hat{X}_{t|t} + \delta'\mathbb{R}_{\Delta^\star}^{-1}\delta + \hat{X}'\mathbb{P}^{-1}\hat{X} - \hat{X}_{t|t}'\mathbb{P}_{t|t}^{-1}\hat{X}_{t|t}$$

$$= \underbrace{(X_t - \hat{X}_{t|t})'}_{\nu_t,\ \text{filtering error}} \mathbb{P}_{t|t}^{-1}\nu_t + \underbrace{(\Delta_t - \overbrace{(\mathbb{C}\hat{X}_{t|t-1} + \mathbb{D}A_t)}^{\hat{\Delta}_{t|t-1}})'}_{\varepsilon_t,\ \text{prediction error}} \overbrace{(\mathbb{R}_{\Delta^\star} + \mathbb{C}\mathbb{P}_{t|t-1}\mathbb{C}')^{-1}}^{R_{\Delta^\star t|t-1}^{-1}} \varepsilon_t$$

- Re-expression of $\hat{X}_{t|t}$

$$\hat{X}_{t|t} = \underbrace{\overbrace{(\mathbb{C}'\mathbb{R}^{-1}_{\Delta^\star}\mathbb{C} + \mathbb{P}^{-1})^{-1}}^{\mathbb{P}_{t|t}} [\mathbb{C}'\mathbb{R}^{-1}_{\Delta^\star}}_{\mathbb{G}_t} \underbrace{(\Delta - \mathbb{C}\hat{X} + \mathbb{C}\hat{X} - \mathbb{D}A)}_{\varepsilon_t + \mathbb{C}\hat{X}} + \mathbb{P}^{-1}\hat{X}]$$

- Matrix Inversion Lemma, $\mathbb{U}, \mathbb{V} > 0$, (think about column $\mathbb{W}$, scalar $\mathbb{V}$)

$$(\mathbb{W}\mathbb{V}\mathbb{W}' + \mathbb{U})^{-1} = \mathbb{U}^{-1} - \mathbb{U}^{-1}\mathbb{W}(\mathbb{U}^{-1} + \mathbb{W}'\mathbb{V}^{-1}\mathbb{W})^{-1}\mathbb{W}'\mathbb{U}^{-1} \quad (74)$$

serves for verifying the final form of $\mathbb{R}_{\Delta^\star t|t-1}$ and $\mathbb{P}_{t|t}$.

- Data updating preserves normality, thus moments suffice.
The chain rule for expectation $_\Lsh$ implies

$$\begin{aligned}\hat{X}_{t+1|t} &= \mathsf{E}[X_{t+1}|A_{t+1}, K_t] = \mathsf{E}[\mathsf{E}[X_{t+1}|X_t, A_{t+1}, K_t]|A_{t+1}, K_t] \\ &= \mathsf{E}[\mathbb{A}X_t + \mathbb{B}A_{t+1}|A_{t+1}, K_t] = \mathbb{A}\hat{X}_{t|t} + \mathbb{B}A_{t+1}.\end{aligned}$$

The second moment is gained similarly. $\square$

Exercises 11 (On Kalman Filtering)

1. *Could be Kalman filter applied to matrix observations?*
2. *Derive Kalman filter with all details.*
3. *Derive a square-root version of Kalman filter using Prop. 15.*
4. *Think about a real-life case where Kalman filter could be useful.*
5. *How you would approach irregularly obtained observations?*
6. *Do you have always a chance to get a nontrivial data-based knowledge about the hidden quantity?*
   Think over about the case $\mathbb{R}_{\Delta^\star}, \mathbb{R}_{X^\star} \approx 0 \approx 0$ *and recall Hamilton-Cayley theorem.*

Remark 12 (Linear Influence of Strategy on Closed Loop)

- *Under natural conditions of DM↳, the closed-loop model↳ factorises*

$$\mathsf{f}_{\mathsf{S}}(B) = \underbrace{\overbrace{\mathsf{f}(X_0)}^{\textit{prior pd}} \prod_{t \in t^\star} \overbrace{\mathsf{f}(\Delta_t, X_t | X_{t-1}, A_t, \mathbf{K}_{t-1})}^{\textit{observation} \times \textit{time evolution pds}}}_{\textit{system model } \mathsf{M}} \underbrace{\prod_{t \in t^\star} \mathsf{f}(A_t | \mathbf{K}_{t-1})}_{\textit{strategy } \mathsf{S}} = \mathsf{M}(B)\mathsf{S}(B) \tag{75}$$

reflecting that the compared strategies↳ work with a common

- *system model* $\mathsf{M} = \mathsf{f}(X_0) \prod_{t \in t^\star} \mathsf{f}(\Delta_t, X_t | X_{t-1}, A_t, \mathbf{K}_{t-1})$,
  the strategy↳ model can be identified with the strategy↳, Agr. 1.

  Strategy S *enters closed-loop model* f_{S} *linearly*

$$\mathsf{f}_{\mathsf{S}} = \mathsf{MS}. \tag{76}$$

Linearity of $\mathsf{E_S[I]}$ in S Implies Basic DM Lemma 10

- For the standard DM design, given by the performance index $\mathsf{I_S = I}$, the formalised DM design defines the optimal strategy

$$^{o}\mathsf{S} \in \mathrm{Arg}\min_{\mathsf{S}\in\mathsf{S}^{\star}} \mathsf{E_S[I]} = \int_{B^{\star}} \mathsf{I}(B)\, \underbrace{\mathsf{M}(B)\mathsf{S}(B)}_{\mathsf{f_S}(B)}\, \mathrm{d}B.$$

- The optimal strategy $^{o}\mathsf{S}$ intuitively concentrates whole probabilistic mass on small values of $\mathsf{I}(B)\mathsf{M}(B)$, while respecting causality.

  This explains the deterministic nature of the optimal strategy.

On Prediction & Filtering in Connection with Reality

- Recall that under the natural conditions of DM↳, filtering↳ relies on the knowledge of actions and not of the strategy↳ S generating them.

  It is important when we learn while the decision loop is closed by a strategy not designed by us, but, say, by a human operator.

- The observation model↳ $f(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})$ and time-evolution model ↳ $f(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})$ result from a theoretical system modelling↳.

  Modelling uses both field knowledge, like conservation laws, e.g., [99], and approximation capabilities [56] of a model family, Sec. 40.

- The prior pd↳ $f(X_0)$ quantifies expert knowledge or analogy, Sec. 22.

- The observations, the only bridge to reality, enter the data updating↳ step only where the newest (action↳,observation↳) pair is processed.

  This is used when approximating time-evolution model↳, Sec. 31.

Summarising Comments on Prediction and Filtering

The Bayesian prediction and filtering combine

- the prior knowledge quantified by the prior pd $\mathsf{f}(X_0)$
- the theoretical knowledge of the specific fields transformed into the observation model $\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})$, the time-evolution model $\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})$, and
- the data records $D^h = (\Delta^h,\, A^h)$ observed on real world

by using coherent deductive calculus with pds.

- This combination of information sources is a powerful, internally consistent, framework describing the essence of learning.
- Due to its deductive structure, an important assurance is gained

  the incorrect modelling or non-informative data can only be blamed for a possible failure of the learning process.

$\Rightarrow$ The errors caused by a bad choice of the learning method are avoided.

This section deals with a special version of filtering called estimation.

- *parameter estimation* is filtering arising for time-invariant hidden X_t

$$X_t = \Theta,\ \forall t \in t^\star. \tag{77}$$

- *unknown parameter* Θ is the value of time-invariant hidden quantities.
- The time-evolution model of the unknown parameter $X_t = \Theta$ is

$$\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1}) = \delta(X_t, X_{t-1}), \quad \text{where} \tag{78}$$

- *Dirac delta* $\delta(\cdot,\cdot)$ is a formal pd of the measure fully concentrated on equal arguments; for a correct handling consult [163].
- A direct application of Prop. 16 with the time-evolution model (78) solves the Bayesian prediction, the evaluation of predictive pd $\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})$, and Bayesian estimation, the evaluation of $\mathsf{f}(\Theta|\mathbf{K}_t)$.

Proposition 17 (Bayesian Estimation)

Let natural conditions of DM↳ *be met & hidden quantity* $X_t = \Theta \in \Theta^\star \subset \mathbf{G}_\tau$, $\forall t, \tau \in t^\star$, *be time invariant. Then, the predictive pd* ↳ *reads*

$$\mathrm{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) = \int_{\Theta^\star} \mathrm{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathrm{f}(\Theta|\mathbf{K}_{t-1})\,\mathrm{d}\Theta. \tag{79}$$

It uses the Bayesian parameter estimation↳, *which evolves*

- *posterior pd* $\mathrm{f}(\Theta|\mathbf{K}_{t-1})$, $t > 1$. *It is the sufficient statistic*↳ *for constructing parameter estimators.*

The data updating↳ *(67) gives its evolution, independent of* $\mathrm{f}(A_t|\mathbf{K}_{t-1})$,

$$\mathrm{f}(\Theta|\mathbf{K}_t) = \frac{\mathrm{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathrm{f}(\Theta|\mathbf{K}_{t-1})}{\mathrm{f}(\Delta_t|A_t, \mathbf{K}_{t-1})} \propto \mathrm{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathrm{f}(\Theta|\mathbf{K}_{t-1}) \tag{80}$$

initiated by the prior pd↳ $\mathrm{f}(\Theta) \equiv \mathrm{f}(\Theta|A_1, \mathbf{K}_0) = \mathrm{f}(\Theta|\mathbf{K}_0)$.

Batch (Non-Recursive) Parameter Estimation

Proposition 18 (Batch Parameter Estimation)

Under natural conditions of DM↳, the parameter estimation↳ allows the batch evaluation of the posterior pd↳

$$f(\Theta|\mathbf{K}_t) = \frac{\prod_{\tau\le t} f(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1}) f(\Theta)}{\int_{\Theta^\star} \prod_{\tau\le t} f(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1}) f(\Theta)\, d\Theta} \equiv \frac{L(\Theta, \mathbf{K}_t) f(\Theta)}{J(\mathbf{K}_t)}. \quad (81)$$

- *likelihood* $L : \Theta^\star \to [0, \infty]$ *is defined*

$$L(\Theta, \mathbf{K}_t) \equiv \prod_{\tau\le t} f(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1}) \text{ for a fixed knowledge}_{\lrcorner}\ \mathbf{K}_t. \quad (82)$$

- *The normalisation factor* $J(\cdot)$ *is*

$$J(\mathbf{K}_t) = \int_{\Theta^\star} L(\Theta, \mathbf{K}_t) f(\Theta)\, d\Theta \;\Rightarrow\; f(\Delta_t|A_t, \mathbf{K}_{t-1}) = \frac{J(\mathbf{K}_t)}{J(\mathbf{K}_{t-1})}. \quad (83)$$

Proof

- The batch form applies repeatedly data updating (Bayes' rule)

$$\mathsf{f}(\Theta|\mathbf{K}_t) \propto \mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathsf{f}(\Theta|\mathbf{K}_{t-1})\ldots \propto \mathsf{L}(\Theta, \mathbf{K}_t)\mathsf{f}(\Theta)$$

Normalisation $\mathsf{J}(\mathbf{K}_t) \equiv \int_{\Theta^\star} \mathsf{L}(\Theta, \mathbf{K}_t)\mathsf{f}(\Theta)\,\mathrm{d}\Theta$ makes $\int_{\Theta^\star} \mathsf{f}(\Theta|\mathbf{K}_t)\,\mathrm{d}\Theta = 1$.

- Predictive pd

$$\begin{aligned}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) &= \int_{\Theta^\star} \mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathsf{f}(\Theta|\mathbf{K}_{t-1})\,\mathrm{d}\Theta \\ &= \int_{\Theta^\star} \frac{\overbrace{\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathsf{L}(\Theta, \mathbf{K}_{t-1})}^{\mathsf{L}(\Theta|\mathbf{K}_t)}\mathsf{f}(\Theta)}{\mathsf{J}(\mathbf{K}_{t-1})}\,\mathrm{d}\Theta = \frac{\mathsf{J}(\mathbf{K}_t)}{\mathsf{J}(\mathbf{K}_{t-1})}. \qquad \square\end{aligned}$$

To get the predictive pd as a function of the newest data, the knowledge increment at time t is to be taken as variable!

On the Batch Estimation

Remark 13

- *parametric model is an alternative name of the observation model* $\mathrm{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$. *It is used if the parameter estimation is considered, i.e. when the hidden quantities* X_t *are time invariant.*
- *The recursive evolution of the pd* $\mathrm{f}(\Theta|\mathbf{K}_t)$ *allows us to interpret the posterior pd as the prior pd before processing new data records.*
- *The data inserted into the parametric model corrects the subjectively chosen prior pd* $\mathrm{f}(\Theta)$. *The posterior pd* $\mathrm{f}(\Theta|\mathbf{K}_t)$ *reflects both objective and subjective knowledge pieces.*
- *If the data are informative, the relative contribution of the single subjective factor* $\mathrm{f}(\Theta)$ *to the posterior pd decreases with increasing* t *as the likelihood* $\mathrm{L}(\Theta, \mathbf{K}_t)$ *contains* t *"objective" factors (82).*

Qualitative Role of the Prior PD

Proposition 19 (Role of the Prior PD)

- *Parameter values* $\Theta \notin \mathrm{supp}\,[\,\mathsf{f}(\Theta)]$, *for which the prior pd* *is zero, get the zero posterior pd, too. Formally,*

$$\mathrm{supp}\,[\,\mathsf{f}(\Theta|\mathbf{K}_t)] = \mathrm{supp}\,[\,\mathsf{L}(\Theta,\mathbf{K}_t)] \cap \mathrm{supp}\,[\,\mathsf{f}(\Theta)]\,.$$

- *The recursive evolution of the likelihood*

$$\begin{aligned} \mathsf{L}(\Theta,\mathbf{K}_t) \quad &\equiv \quad \prod_{\tau\le t} \mathsf{f}(\Delta_\tau|\Theta,A_\tau,\mathbf{K}_{\tau-1}) = \mathsf{f}(\Delta_t|\Theta,A_t,\mathbf{K}_{t-1})\mathsf{L}(\Theta,\mathbf{K}_{t-1}) \\ & \mathsf{L}(\Theta,\mathbf{K}_0) = 1,\ \Theta \in \mathrm{supp}\,[\,\mathsf{L}(\Theta,\mathbf{K}_t)]_{t\in t^\star} \equiv \Theta^\star_{\mathsf{L}} \qquad (84) \end{aligned}$$

does not depend on the prior pd chosen.

- *The posterior pd* *exists iff the product* $\mathsf{L}(\Theta,\mathbf{K}_t)\mathsf{f}(\Theta)$ *is integrable.*

Proof It is a direct consequence of the formula for posterior pd (80). $\square$

Role of the Prior PD

Remark 14 (Role of the Prior PD)

- *The prior pd*$_{\hookrightarrow}$ *allows to introduce hard bounds on parameters.*
- *The recursion (84) is valid even if* $\Theta_{\mathsf{L}}^{*} \neq \Theta^{\star} \equiv \mathrm{supp}\,[\mathsf{f}(\Theta)]$, *i.e. the prior hard bounds on parameter values does not influence likelihood*$_{\hookrightarrow}$.

  This is repeatedly overlooked in recursive estimation. Instead of restricting the posterior pd$_{\hookrightarrow}$, *the likelihood statistic – comprising data – are deformed with an adverse effect on the estimation quality.*
- *Often, a flat prior pd models the lack of prior knowledge. Even integrability of the prior pd is relaxed and*
- *improper prior pd* $\mathsf{f}(\Theta) \geq 0, \ \int_{\Theta^{\star}} \mathsf{f}(\Theta)\,\mathrm{d}\Theta = \infty$ *is used.*

Further Remarks on Parameter Estimation

Remark 15 (Improper Prior PD; Non-Parametric Estimation)

- *For the improper prior uniform pd, the posterior pd is proportional to the likelihood. Then, the posterior pd may be improper, too.*
- *A flat, but proper, prior pd regularises estimation, [20, 156].*
- *The unknown parameter is always in the estimator ignorance.*
- *Under the natural conditions of DM (65), the action values are used in estimation, not the strategy generating them.*
- *The parameter Θ is usually finite-dimensional. Exceptionally, we deal with potentially infinite-dimensional parameter. It means that the number of hidden quantities is finite but increases without bounds. This case is often called nonparametric estimation.*

Asymptotic of Estimation

Estimation asymptotic analyses the posterior pd↳ for a growing data amount. The analysis serves for:

- an interpretation of estimation results when none of the considered parametric model↳s determines the objective pd↳ denoted ${}^{o}\mathsf{f}(B)$.
- a construction of an approximate estimation.

The analysis relates the observation predictor↳, given by the objective pd↳ ${}^{o}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})$, to the parametric model↳ $\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$.

- The relation is studied the support of the posterior pd↳ $\mathsf{f}(\Theta|\mathbf{K}_t)$ obtained via the parameter estimation↳, see Prop. 17.
- The analysis uses Kullback-Leibler divergence (aka relative entropy).
- Technicalities are suppressed as in the asymptotic design, Sec. 12.

Kullback-Leibler Divergence (KLD) of PDs

The asymptotic analysis of the Bayesian parameter estimation ↳ and the fully probabilistic design of decision strategies (FPD ↳) need to measure the proximity of a pair of pds modelling the same $X \in X^\star$.

The analysis and FPD ↳ delimit as an appropriate proximity measure:

- *KLD* Kullback–Leibler divergence, [113], $\mathrm{D}(\mathrm{f}||\mathrm{g})$ compares a pair of pds f, g acting on a common domain $X^\star$. KLD is defined

$$\mathrm{D}(\mathrm{f}||\mathrm{g}) \equiv \int_{X^\star} \mathrm{f}(X) \ln\left(\frac{\mathrm{f}(X)}{\mathrm{g}(X)}\right) \mathrm{d}X. \tag{85}$$

KLD asymmetry is stressed by referring to it as KLD of f from g.

Basic Properties of KLD

Proposition 20 (Properties of KLD)

Let f, g *be pds on* $X^\star \neq \emptyset$. $\mathrm{D}(\mathsf{f}||\mathsf{g}) \equiv \int_{X^\star} \mathsf{f}(X) \ln\left(\frac{\mathsf{f}(X)}{\mathsf{g}(X)}\right) \mathrm{d}X$ *fulfils*

- $\mathrm{D}(\mathsf{f}||\mathsf{g}) \geq 0$,
- $\mathrm{D}(\mathsf{f}||\mathsf{g}) = 0$ *iff* $\mathsf{f} = \mathsf{g}$ $\mathrm{d}X$*-almost surely,*
- $\mathrm{D}(\mathsf{f}||\mathsf{g}) = \infty$ *iff on a subset* $X_\star \subset X^\star$ *of a positive dominating measure* $\mathrm{d}X$, *it holds* $\mathsf{f} > 0$ *and* $\mathsf{g} = 0$,
- $\mathrm{D}(\mathsf{f}||\mathsf{g}) \neq \mathrm{D}(\mathsf{g}||\mathsf{f})$ *and KLD does not obey triangle inequality,*
- $\mathrm{D}(\mathsf{f}||\mathsf{g})$ *is invariant with respect to a sufficient mapping* $\Upsilon : X^\star \to Y^\star$. *Note that any bijective mapping* Υ *is sufficient.*

Proof See, for instance, [162]. □

Example: KLD of Normal PDs

- Let X be $\ell_{X^\star}$-dimensional real vector modelled by a pair of normal pd ↳s

$$
\begin{aligned}
\mathsf{f}(X) &= \mathcal{N}_X(\bar{X}_\mathsf{f}, \mathbb{R}_\mathsf{f}) \equiv |2\pi\mathbb{R}_\mathsf{f}|^{-0.5} \exp[-0.5(X - \bar{X}_\mathsf{f})'\mathbb{R}_\mathsf{f}^{-1}(X - \bar{X}_\mathsf{f})] \\
\mathsf{g}(X) &= \mathcal{N}_X(\bar{X}_\mathsf{g}, \mathbb{R}_\mathsf{g}) \qquad (86)
\end{aligned}
$$

given by expected values $\bar{X}_\mathsf{f}, \bar{X}_\mathsf{g}$ and covariance matrices $\mathbb{R}_\mathsf{f} > 0, \mathbb{R}_\mathsf{g} > 0$.

- Then, see e.g. [77],

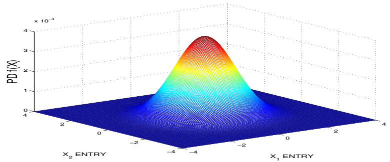
$$2\mathsf{D}(\mathsf{f}||\mathsf{g}) = (\bar{X}_\mathsf{f} - \bar{X}_\mathsf{g})'\mathbb{R}_\mathsf{g}^{-1}(\bar{X}_\mathsf{f} - \bar{X}_\mathsf{g}) + \mathrm{tr}[\mathbb{R}_\mathsf{f}\mathbb{R}_\mathsf{g}^{-1}] + \ln[|\mathbb{R}_\mathsf{f}^{-1}\mathbb{R}_\mathsf{g}|] - \ell_{X^\star}. \quad (87)$$

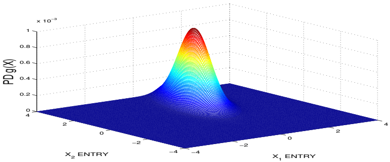
Mutual position, given by $\bar{X}_\mathsf{f} - \bar{X}_\mathsf{g}$, and rotation, given by $\mathbb{R}_\mathsf{f}\mathbb{R}_\mathsf{g}^{-1}$, matter

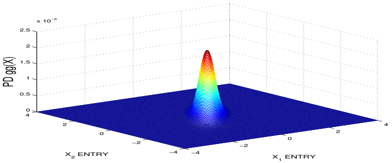
Remember:

The shape similarity of f and g matters, not only their mutual position.

Be Careful in Interpreting the KLD Value







$$\begin{aligned}
f(X) &= \mathcal{N}_X\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\
g(X) &= \mathcal{N}_X\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}\right) \\
gg(X) &= \mathcal{N}_X\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.4 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}\right)
\end{aligned}$$

$$D(f||g) = 4.163 < 6.025 = D(f||gg)$$

Entropy Rate

- The asymptotic analysis exploits the notion of entropy rate↳, an extension of KLD↳ to sequence of pds' pairs, [32].
- Entropy rate measures the proximity of the objective predictor↳ ${}^{o}\mathsf{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1})$ from a parametric model↳ $\mathsf{f}(\Delta_\tau|\Theta,A_\tau,\mathbf{K}_{\tau-1})$. The proximity is judged for realisation↳s $(A_\tau,\mathbf{K}_{\tau-1})_{\tau\le t}$, $t\in t^\star$, and an arbitrary parameter value $\Theta\in\Theta^\star$
- *entropy rate* is defined

$$
\begin{aligned}
&\mathsf{R}_\infty\left({}^{o}\mathsf{f},\Theta\right) \equiv \overline{\lim}_{t\to\infty}\mathsf{R}_t\left({}^{o}\mathsf{f},\Theta\right) \qquad (88)\\
&\equiv \overline{\lim}_{t\to\infty}\frac{1}{t}\sum_{\tau\le t}\int_{\Delta_\tau^\star}{}^{o}\mathsf{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1})\ln\left(\frac{{}^{o}\mathsf{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1})}{\mathsf{f}(\Delta_\tau|\Theta,A_\tau,\mathbf{K}_{\tau-1})}\right)\mathrm{d}\Delta_\tau.
\end{aligned}
$$

- The non-negativity of KLD↳, Prop. 20, implies that the definition (88) is meaningful and that $\mathsf{R}_\infty\left({}^{o}\mathsf{f},\Theta\right)\in[0,\infty]$.

Asymptotic of Estimation

Proposition 21 (Basic Characterisation of the Estimation Asymptotic)

- *Let the natural conditions of DM ↳ (65) hold.*
- *Let exist positive values* $\underline{C}_{\Theta}, \overline{C}_{\Theta}$ *and* $c \in (0, \infty)$ *such that*

$$1/c < \underline{C}_{\Theta} \leq \overline{C}_{\Theta} \leq c < \infty, \quad \textit{for almost all} \;\; \Theta \in \Theta^{\star}, \quad (\mathbf{K}_t \in \mathbf{K}_t{}^{\star})_{t \geq 1},$$

$$\underline{C}_{\Theta} \mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) \leq {}^{o}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \leq \overline{C}_{\Theta} \mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}). \tag{89}$$

- *Then, the posterior pd* $\mathsf{f}(\Theta|\mathbf{K}_t)$ *(80) converges* ${}^{o}\mathsf{f}$*-almost surely to a pd* $\mathsf{f}(\Theta|\mathbf{K}_{\infty})$ *with the support*

$$\mathrm{supp}\left[\mathsf{f}(\Theta|\mathbf{K}_{\infty})\right] = \mathrm{Arg} \min_{\Theta \in \mathrm{supp}[\mathsf{f}(\Theta)]} \mathsf{R}_{\infty}\left({}^{o}\mathsf{f}, \Theta\right). \tag{90}$$

- Under the natural conditions of DM, the posterior pd reads

$$\mathsf{f}(\Theta|\mathbf{K}_t) \quad \propto \quad \mathsf{f}(\Theta)\exp[-t\mathsf{R}(\mathbf{K}_t,\Theta)], \tag{91}$$

$$\mathsf{R}(\mathbf{K}_t,\Theta) = \frac{1}{t}\sum_{\tau\le t}\ln[\eta(\mathbf{K}_\tau,\Theta)], \quad \eta(\mathbf{K}_\tau,\Theta) \equiv \frac{{}^{o}\mathsf{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1})}{\mathsf{f}(\Delta_\tau|\Theta,A_\tau,\mathbf{K}_{\tau-1})}. \tag{92}$$

It exploits that $\bullet = \exp(\ln(\bullet))$ and the fact that the multiplication of right-hand side by a Θ-independent factor defines the same posterior pd.

- Let us fix the argument $\Theta \in \Theta^\star$ and define the deviations

$$\begin{aligned}
\varepsilon_\tau(\Theta) \quad &\equiv \quad \ln(\eta(\mathbf{K}_\tau,\Theta)) - \mathsf{E}_{{}^{o}\mathsf{f}}\left[\ln(\eta(\mathbf{K}_\tau,\Theta))\middle| A_\tau,\mathbf{K}_{\tau-1}\right] \\
&\equiv \quad \ln(\eta(\mathbf{K}_\tau,\Theta)) - \int_{\Delta_\tau{}^\star} {}^{o}\mathsf{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1})\ln(\eta(\mathbf{K}_\tau,\Theta))\,\mathrm{d}\Delta_\tau \\
&\Rightarrow \quad \mathsf{R}(\mathbf{K}_t,\Theta) = \mathsf{R}_t\left({}^{o}\mathsf{f},\Theta\right) + \frac{1}{t}\sum_{\tau\le t}\varepsilon_\tau(\Theta).
\end{aligned}$$

- Deviations $\varepsilon_\tau(\Theta)$ are zero mean & uncorrelated (check directly, [132]).
- The first non-negative summand, Prop. 20, in the decomposition

$$\mathsf{R}(\mathbf{K}_t, \Theta) = \mathsf{R}_t\left({}^{o}\mathsf{f}, \Theta \right) + \frac{1}{t} \sum_{\tau \le t} \varepsilon_\tau(\Theta)$$

is finite and converges for $t \to \infty$ due to the constraints (89).

- The constraints (89) guarantee boundedness the variance of $\varepsilon_\tau(\Theta)$. The 2^{nd} sum converges to zero a.s., see the ergodic result [118], p. 417. Altogether, $\mathsf{R}(\mathbf{K}_t, \Theta)$ converges to $\mathsf{R}_\infty\left({}^{o}\mathsf{f}, \Theta \right) \ge 0$ a.s.
- The subtraction $t \times \min_{\Theta \in \mathrm{supp}[\mathsf{f}(\Theta)]} \mathsf{R}_\infty\left({}^{o}\mathsf{f}, \Theta \right)$ in the exponent of (91) does not change the posterior pd. This pd gets the form

$$\mathsf{f}(\Theta|\mathbf{K}_\infty) \propto \mathsf{f}(\Theta) \exp(-t \times \text{ an asymptotically nonnegative factor})$$

- The pd $\mathsf{f}(\Theta|\mathbf{K}_\infty)$ is nonzero only on Θ-set where the asymptotically non-negative factor is zero. This happens on the set (90). □

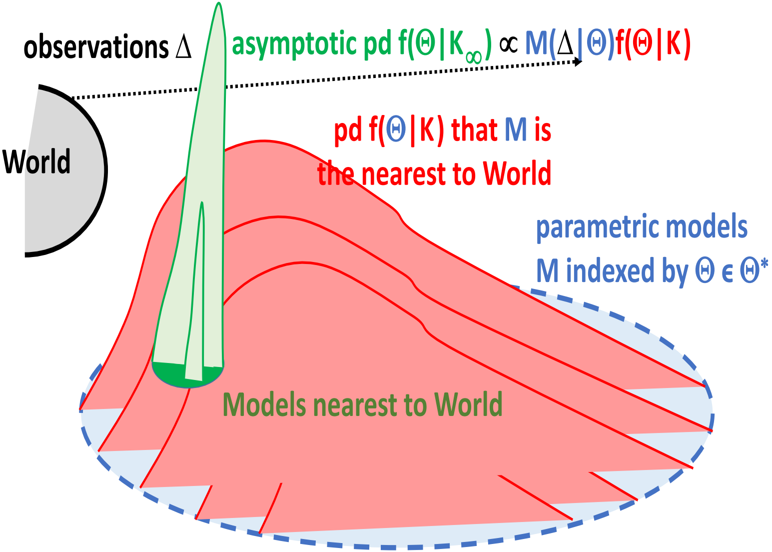
Identifiable Parameter, Entropy Rate and KLD

- Assumption (89) excludes parametric model↳s that do not expect observations generated by the system↳ & vice versa.
- Model is identifiable iff the entropy rate↳ has a unique minimiser. It depends on: a) the parametric-models' set; b) the used actions.
- The entropy rate↳ extends KLD↳ (85). It covers asymptotic and closed-loop cases and often coincides with KLD.

The Bayesian estimation asymptotically finds

- *the best projection* is the minimiser within the set of parametric model↳s $\{(\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}))_{t\in t^\star}\}_{\Theta\in\Theta^\star}$ of the entropy rate↳ of the objective pd↳ ${}^{o}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})$ from it.
- The pd $\mathsf{f}(\Theta|\mathbf{K})$ quantifies the belief that the parametric model↳ given by $\Theta \in \Theta^\star$ is the best projection of the objective pd↳ [12].

$\boxed{\text{Not knowing } {}^{o}\mathsf{f}(\Delta|A, \mathbf{K}), \text{ we do not know the best projection.}}$


observations Δ
asymptotic pd f(Θ|K∞) ∝ M(Δ|Θ)f(Θ|K)
World
pd f(Θ|K) that M is
the nearest to World
parametric models
M indexed by Θ ϵ Θ*
Models nearest to World

Remark 16 (Consistency of the Bayesian Estimation)

- *If the objective pd*↳ ${}^{o}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})$ *coincides with* $\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$ *for some* $\Theta = {}^{o}\Theta \in \Theta^{\star}$ *with* $\mathsf{f}\,({}^{o}\Theta) > 0$ *then* ${}^{o}\Theta$ *is in the support of the asymptotic posterior pd* $\mathsf{f}(\Theta|\mathbf{K}_{\infty})$.
- *If, moreover, the model is identifiable, then the objective pd*↳ *is gained by the Bayesian estimation. This has the appealing expression*

  Bayesian estimator is consistent whenever there is a consistent one.

- *A similar analysis was performed by measuring the distance of the parametric model*↳ *to the empirical pd of data [143]. It gives similar answers if the empirical pd converges to the objective pd*↳.

  It hints how to approximate posterior pd↳ *[109], see Sec. 30.*

  The known conditions of such convergence are more restrictive and make the analysis of the estimation in closed decision loop harder.

Exercises 12 (On Estimation)

1. *What is difference between estimation, estimator and estimate?*
2. *Consider tossing of a coin with unknown probability of head. Select parametric model↳ and prior pd↳.*
   Provide Bayesian parameter estimation↳ of this probability.
3. *Assume one toss and compare point estimates: maximum likelihood one and the posterior expectation.*
4. *Consider a simple linear normal Markov parametric model. Select the version, which is not identifiable due to*
   (a) bad parametrisation,
   (b) chosen action↳s.

Fully Probabilistic Design

On Content of this Part

This part goes beyond the standard design↳ by inspecting a strategy-dependent performance index↳ $\mathsf{I}_{\mathsf{S}}(B)$. It:

- justifies the fully probabilistic design (FPD↳) of DM↳ strategies
- finds the relation of the FPD to the standard Bayesian DM↳
- summaries results on the FPD described in [55, 70, 81, 89, 153]
- generalises approaches known as KL-control [53, 160] having connections with brain research [157, 158]
- prepares a powerful tool for creating DM supporting tools.

Strategy-Optimality Recalled

- DM preferences are expressed by a preferential ordering $\,^{a}B \preccurlyeq_{\mathsf{B}^\star} \,^{b}B$ quantified by a real-valued loss $\mathsf{Z}(\,^{a}B) \leq \mathsf{Z}(\,^{b}B)$, $\,^{a}B, \,^{b}B \in B^\star$.
- Uncertainty $N \in N^\star \neq \emptyset$ models that a chosen strategy $\mathsf{S} \in \mathsf{S}^\star$ does not determine unambiguously the behaviour $B \in B^\star$.

  It enters the bijective mapping (2) $\quad \mathsf{W}(\mathsf{S}, \bullet) : \; N^\star \to B^\star$.
- The preference $\preccurlyeq_{\mathsf{S}^\star}$ between compared strategies $\,^{a}\mathsf{S}, \,^{b}\mathsf{S} \in \mathsf{S}_\star$ is quantified by the mapping T_S (20)

$$\,^{a}\mathsf{S} \preccurlyeq_{\mathsf{S}^\star} \,^{b}\mathsf{S} \Leftrightarrow \mathsf{T}_{\,^{a}\mathsf{S}} \leq \mathsf{T}_{\,^{b}\mathsf{S}}$$

$$\mathsf{T}_\mathsf{S} = \int_{B^\star} \underbrace{\overbrace{\mathsf{U}}^{utility \nearrow in\, \mathsf{Z}} \left(\mathsf{Z}(B), \mathsf{W}^{-1}(S, B)\right)}_{performance\ index\ \mathsf{I}_\mathsf{S}(B)} \quad \underbrace{\mathsf{f}_\mathsf{S}(B)}_{closed-loop\ pd} \quad \mathrm{d}B \equiv \mathsf{E}_\mathsf{S}[\mathsf{I}_\mathsf{S}].$$

Its minimisation over $\mathsf{S}_\star$ gives the non-dominated optimal strategy.

Questions on Strategy-Dependent Performance Indices

In summary, ordering of strategies $\preccurlyeq_{\mathsf{S}^\star}$, (17) is quantified by

$$
\begin{aligned}
\mathsf{E}_{\mathsf{S}}[\mathsf{I}_{\mathsf{S}}] &= \int_{B^\star} \mathsf{I}_{\mathsf{S}}(B)\mathsf{f}_{\mathsf{S}}(B)\,\mathrm{d}B, && \text{see (22)}\\
\mathsf{I}_{\mathsf{S}}(B) &= \mathsf{U}(\mathsf{Z}(B), \mathsf{W}^{-1}(\mathsf{S}, B)), && \text{see (2)}\\
B &= \mathsf{W}(\mathsf{S}, N), && N \in N^\star \neq \emptyset.
\end{aligned}
$$

Unlike in the standard design↳, the strategy-dependent performance index↳ is considered. It induces the natural questions:

- How to model the strategy↳ influence on the performance index↳?
- What is gained by considering the general performance index↳ I_{S}?

These questions are answered below under widely acceptable assumptions on the utility↳ U entering the definition of the performance index↳ I_{S}.

Utility Models the Risk Attitude of Decision Makers

Intuitively, utilities should model properly the risk attitude.

Requirement 4 (Utility Properly Modelling the Risk Attitude)

The utility U *respects the risk attitude properly if it meets the implication*

$$\mathsf{f}_\mathsf{S}({}^aB) = \mathsf{f}_\mathsf{S}({}^bB) \quad \textit{for some} \quad {}^aB,\ {}^bB \in B^\star \tag{93}$$
$$\Rightarrow \quad \mathsf{U}(z, \mathsf{W}^{-1}(\mathsf{S}, {}^aB)) = \mathsf{U}(z, \mathsf{W}^{-1}(\mathsf{S}, {}^bB)),\ \forall z \in [-\infty, \infty].$$

Requirement 4 means that equiprobable behaviours with the same loss value z contribute equally to the expected performance index.

Requirement 4 implies that performance index has to have form

$$\mathsf{I}_\mathsf{S}(B) \equiv \mathsf{U}(\mathsf{Z}(B), \mathsf{W}^{-1}(\mathsf{S}, B)) = \mathsf{U}(\mathsf{Z}(B), \mathsf{f}_\mathsf{S}(B)) \equiv \mathsf{I}(B, \mathsf{f}_\mathsf{S}(B)). \tag{94}$$

as any argument of U added to $(z, \mathsf{f}_\mathsf{S})$ must not influence the value of I_S.

Equivalent Performance Indices

The optimal strategy ↳ ${}^{o}\mathsf{S}$ is determined by

- the set $\mathsf{S}_\star \subset \mathsf{S}^\star$ of compared strategies ↳,
- the system model ↳ M entering the closed-loop model ↳, $\mathsf{f}_\mathsf{S} = \mathsf{MS}$,
- the employed performance index ↳, ${}^{e}\mathsf{I}_\mathsf{S}$, $e \in e^\star \equiv \{a, b, \ldots\}$.

$$ {}^{oe}\mathsf{S} \in \mathrm{Arg}\min_{\mathsf{S}\in\mathsf{S}_\star} \int_{B^\star} {}^{e}\mathsf{I}_\mathsf{S}(B)\mathsf{M}(B)\mathsf{S}(B)\,\mathrm{d}B. \tag{95} $$

The minimisation is just a tool as its result matters. This leads to the notion

- *equivalent performance indices* ${}^{a}\mathsf{I}_\mathsf{S}$, ${}^{b}\mathsf{I}_\mathsf{S}$ yield the optimal strategies ${}^{oa}\mathsf{S}$, ${}^{ob}\mathsf{S}$ giving the same closed-loop model ↳

$$ \mathsf{f}_{{}^{oa}\mathsf{S}}(B) = \mathsf{f}_{{}^{ob}\mathsf{S}}(B), \quad \forall B \in B^\star, \tag{96} $$

while the optimisation (95)

- acts on a common set of compared strategies ↳ $\mathsf{S}_\star \subset \mathsf{S}^\star$
- uses a common system model ↳ M.

Ideal PD Delimits a Set of Equivalent Performance Indices

A fixed system model M and a fixed performance index $\mathsf{I_S}$ determine

- *globally optimal strategy* ${}^{\mathsf{I}}\mathsf{S}$ is the optimal strategy found on the widest set of admissible strategies $\mathsf{S}^\star$ for the given M and $\mathsf{I_S}$

$$ {}^{\mathsf{I}}\mathsf{S} \equiv {}^{o}\mathsf{S} \in \mathrm{Arg}\min_{\mathsf{S}\in\mathsf{S}^\star} \int_{B^\star} \underbrace{\mathsf{U}(\mathsf{Z}(B),\mathsf{W}^{-1}(\mathsf{S},B))}_{\mathsf{I_S}(B)} \mathsf{f_S}(B)\,\mathrm{d}B. \tag{97} $$

This leads to the notion of the ideal (best achievable) closed-loop model

- *ideal pd* ${}^{\mathsf{I}}\mathsf{f}(B)$ is the pd of the behaviour of the decision loop closed by the globally optimal strategy (97) gained for the given system model M and performance index $\mathsf{I_S}$:

$$ {}^{\mathsf{I}}\mathsf{f}(B) \equiv \mathsf{f}_{{}^{\mathsf{I}}\mathsf{S}}(B), \quad B \in B^\star. \tag{98} $$

- Any pd f on $B^\star$, including ideal pd, delimits equivalent performance indices. Thus, a choice of the ideal pd replaces the choice of performance indices meeting (97), (98).

Representant of Performance Indices with Given Ideal PD

Let a decision maker have provided an ideal pd ${}^{\mathsf{I}}\mathsf{f}(B)$, not a performance index. A representant of equivalent performance indices is asked to meet:

Requirement 5 (On Representant of Equivalent Performance Indices)

A performance index I_{S} *is accepted as a representant of equivalent performance indices determined by* ${}^{\mathsf{I}}\mathsf{f}(B)$ *if it:*

1. *is representant, i.e. guarantees that the globally optimal strategy*
$$ {}^{\mathsf{I}}\mathsf{S} \in \operatorname{Arg}\min_{S \in S^\star} \int_{B^\star} \mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B))\mathsf{f}_{\mathsf{S}}(B)\,\mathrm{d}B \quad \textit{makes} \quad \mathsf{f}_{{}^{\mathsf{I}}\mathsf{S}}(B) = {}^{\mathsf{I}}\mathsf{f}(B) \qquad (99)$$
2. *has the form* $\mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B))$, *i.e. meets Requirement 4,*
3. *decomposes* $\mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B)) = {}^{a}\mathsf{I}({}^{a}B, {}^{a}\mathsf{f}_{{}^{a}\mathsf{S}}({}^{a}B)) + {}^{b}\mathsf{I}({}^{b}B, {}^{b}\mathsf{f}_{{}^{b}\mathsf{S}}({}^{b}B))$ *when two independent tasks have been artificially connected, i.e.*
   - *the behaviour set* $B^\star = {}^{a}B^\star \times {}^{a}B^\star$,
   - *the system model* $\mathsf{M}(B) = {}^{a}\mathsf{M}({}^{a}B) \times {}^{b}\mathsf{M}({}^{b}B)$,
   - *the ideal pd* ${}^{\mathsf{I}}\mathsf{f}(B) = {}^{\mathsf{I}a}\mathsf{f}({}^{a}B) \times {}^{\mathsf{I}b}\mathsf{f}({}^{b}B)$.

Proposition 22 (Representant of Indices Meeting Requirement 5)

A representant of the performance indices sharing a given ideal pd $^{\mathsf{I}}\mathsf{f}(B) > 0$ on $B^{\star}$, which meets Requirement 5, exists. It has the form

$$\mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B)) \;=\; \ln\left(\mathsf{f}_{\mathsf{S}}(B)/\,^{\mathsf{I}}\mathsf{f}(B)\right), \tag{100}$$

i.e. the expected performance index to be minimised is the KLD

$$\mathsf{E}_{S}[\mathsf{I}_{\mathsf{S}}] \;=\; \mathsf{D}(\mathsf{f}_{\mathsf{S}}||\,^{\mathsf{I}}\mathsf{f}) = \int_{B^{\star}} \mathsf{f}_{\mathsf{S}}(B) \ln\left(\frac{\mathsf{f}_{\mathsf{S}}(B)}{^{\mathsf{I}}\mathsf{f}(B)}\right) \mathrm{d}B. \tag{101}$$

Proof For the fixed $^{\mathsf{I}}\mathsf{f} > 0$, $\mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B)) \equiv \mathsf{Q}(\mathsf{f}_{\mathsf{S}}(B)/\,^{\mathsf{I}}\mathsf{f}(B))$ meets item 2 of Requirement 5. Let $x\mathsf{Q}(x)$ be a continuous convex function mapping $(0, \infty] \to (-\infty, \infty]$ and $\mathsf{Q}(1) = 0$. Then, Jensen's inequality (31) implies that the expectation of $\mathsf{I}(B, \mathsf{f}_{\mathsf{S}}(B))$ is minimised by $\mathsf{f}_{\mathsf{S}}(B)/\,^{\mathsf{I}}\mathsf{f}(B) = 1$. Thus, item 1 is met. $\mathsf{Q}(x) = \ln(x)$ is the only function, which guarantees item 3, see [39]. $\square$

Fully Probabilistic Design of DM Strategies

- The (non-unique) representant of equivalent performance indices assigned to a given ideal pd $^{\mathsf{I}}\mathsf{f} > 0$, meeting Requirement 5, delimits KLD as the expected performance index.

$$\mathsf{E}_S[\mathsf{I}_\mathsf{S}] \quad = \quad \mathsf{D}(\mathsf{f}_\mathsf{S}||\,^{\mathsf{I}}\mathsf{f}) = \int_{B^\star} \mathsf{f}_\mathsf{S}(B) \ln\left(\frac{\mathsf{f}_\mathsf{S}(B)}{^{\mathsf{I}}\mathsf{f}(B)}\right) \mathrm{d}B. \qquad (102)$$

- *FPD*, the fully probabilistic design of DM strategy, is the optimal design with the performance index (102). It takes

$$^{o}\mathsf{S} \in \mathrm{Arg} \min_{\mathsf{S} \in \mathsf{S}_\star} \mathsf{D}(\mathsf{f}_\mathsf{S}||\,^{\mathsf{I}}\mathsf{f}) \ \text{ as the optimal strategy}, \qquad (103)$$

  where $\mathsf{S}_\star \subset \mathsf{S}^\star$ contains the compared strategies.

- The subsequent discussion provides
  - the solution of FPD, a counterpart of dynamic programming
  - the demonstration that FPD is a dense extension of the standard DM.

To learn the FPD, we consider data-driven design with the behaviour

$$D^h = \text{data records sequence} = (\Delta^h, A^h) = (\text{observation}^h, \text{action}^h).$$

Recall that $D^t = \mathbf{K}_t$ is the knowledge available for the choice of A_{t+1}.

- The joint pd $\mathsf{f}_\mathsf{S}(B) \equiv \mathsf{f}_\mathsf{S}(D^h) = \mathsf{f}_\mathsf{S}(\mathbf{K}_h)$ factorises by the chain rule

$$\mathsf{f}_\mathsf{S}(D^h) \equiv \mathsf{f}_\mathsf{S}(\mathbf{K}_h) = \prod_{t \in t^\star} \underbrace{\mathsf{f}(\Delta_t | A_t, \mathbf{K}_{t-1})}_{\text{predictor}} \underbrace{\mathsf{f}(A_t | \mathbf{K}_{t-1})}_{\text{decision rule}} . \tag{104}$$

- The ideal pd determining the FPD can be factorised similarly

$$^{\mathsf{I}}\mathsf{f}(D^h) \equiv {}^{\mathsf{I}}\mathsf{f}(\mathbf{K}_h) = \prod_{t \in t^\star} {}^{\mathsf{I}}\mathsf{f}(\Delta_t | A_t, \mathbf{K}_{t-1}) \, {}^{\mathsf{I}}\mathsf{f}(A_t | \mathbf{K}_{t-1}). \tag{105}$$

- *data-driven FPD* selects the optimal strategy

$$^{\mathsf{o}}\mathsf{S} \in \operatorname{Arg}\min_{\mathsf{S} \in \mathsf{S}^\star} \mathsf{D}\left(\mathsf{f}_\mathsf{S} || {}^{\mathsf{I}}\mathsf{f}\right), \quad {}^{\mathsf{o}}\mathsf{S}(\mathbf{K}_h) = \prod_{t \in t^\star} {}^{\mathsf{o}}\mathsf{f}(A_t | \mathbf{K}_{t-1}), \text{ for } \mathsf{f}_\mathsf{S} \text{ (104)}, \ {}^{\mathsf{I}}\mathsf{f} \text{ (105)}.$$

Proposition 23 (Solution of the Data-Driven FPD ↳)

The randomised decision rule↳s forming the optimal strategy in the data-driven FPD↳ are

$$ {}^{o}\mathsf{f}(A_t|\mathbf{K}_{t-1}) = {}^{I}\mathsf{f}(A_t|\mathbf{K}_{t-1}) \frac{\exp[-\omega(A_t, \mathbf{K}_{t-1})]}{\gamma(\mathbf{K}_{t-1})} \quad (106)$$

$$\gamma(\mathbf{K}_{t-1}) = \int_{A_t^\star} {}^{I}\mathsf{f}(A_t|\mathbf{K}_{t-1}) \exp[-\omega(A_t, \mathbf{K}_{t-1})] \, \mathrm{d}A_t \leq 1 \quad (107)$$

for $t < h$ *while* $\gamma(\mathbf{K}_h) = 1$

$$\omega(A_t, \mathbf{K}_{t-1}) = \int_{\Delta_t^\star} \mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \ln\left(\frac{\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\, {}^{I}\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})}\right) \mathrm{d}\Delta_t. \quad (108)$$

The solution is performed against the time course, starting at $t = h$.

The product forms of the closed-loop model (104) and its ideal counterpart (105) imply that KLD is additive loss with the partial performance index

$$z(\mathbf{K}_t) = \ln\left(\frac{f(\Delta_t|A_t,\mathbf{K}_{t-1})f(A_t|\mathbf{K}_{t-1})}{{}^{I}f(\Delta_t|A_t,\mathbf{K}_{t-1})\,{}^{I}f(A_t|\mathbf{K}_{t-1})}\right).$$

Thus, we face a variation of Prop. 12. Let us express loss-to-go in the form $-\ln(\gamma(\mathbf{K}_t))$. It defines the terminal condition $\gamma(\mathbf{K}_h) = 1$.
The generic term G_t to be minimised over $f(A_t|\mathbf{K}_{t-1})$ reads

$$\int_{D_t^\star} f(\Delta_t|A_t,\mathbf{K}_{t-1})f(A_t|\mathbf{K}_{t-1}) \ln\left(\frac{f(\Delta_t|A_t,\mathbf{K}_{t-1})f(A_t|\mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\,{}^{I}f(\Delta_t|A_t,\mathbf{K}_{t-1})\,{}^{I}f(A_t|\mathbf{K}_{t-1})}\right) \mathrm{d}D_t$$

$$= \int_{A_t^\star} f(A_t|\mathbf{K}_{t-1})\left[\ln\left(\frac{f(A_t|\mathbf{K}_{t-1})}{{}^{I}f(A_t|\mathbf{K}_{t-1})}\right) + \underbrace{\omega(A_t,\mathbf{K}_{t-1})}_{(108)}\right] \mathrm{d}A_t \equiv G_t$$

$$\omega(A_t,\mathbf{K}_{t-1}) \equiv \int_{\Delta_t^\star} f(\Delta_t|A_t,\mathbf{K}_{t-1}) \ln\left(\frac{f(\Delta_t|A_t,\mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\,{}^{I}f(\Delta_t|A_t,\mathbf{K}_{t-1})}\right) \mathrm{d}\Delta_t \geq 0.$$

Finalised Proof of the Solution of the Data-Driven FPD

The generic term of the minimised functional can be re-arranged as follows

$$\begin{aligned}\mathsf{G}_t &= \int_{A_t^\star} \mathsf{f}(A_t|\mathbf{K}_{t-1}) \ln \left(\frac{\mathsf{f}(A_t|\mathbf{K}_{t-1})}{\frac{{}^{\mathsf{I}}\mathsf{f}(A_t|\mathbf{K}_{t-1}) \exp[-\omega(A_t,\mathbf{K}_{t-1})]}{\int_{A_t^\star} {}^{\mathsf{I}}\mathsf{f}(A_t|\mathbf{K}_{t-1}) \exp[-\omega(A_t,\mathbf{K}_{t-1})]\, \mathrm{d}A_t}} \right) \mathrm{d}A_t \\ &- \ln \left(\int_{A_t^\star} {}^{\mathsf{I}}\mathsf{f}(A_t|\mathbf{K}_{t-1}) \exp\left[-\omega(A_t, \mathbf{K}_{t-1})\right] \mathrm{d}A_t \right).\end{aligned}$$

- The first summand is the conditional KLD of $\mathsf{f}(A_t|\mathbf{K}_{t-1})$ from the conjectured optimal decision rule (106). It reaches its minimum zero value for $\mathsf{f}(A_t|\mathbf{K}_{t-1}) = {}^{\mathsf{o}}\mathsf{f}(A_t|\mathbf{K}_{t-1})$.
- The second summand is independent of $\mathsf{f}(A_t|\mathbf{K}_{t-1})$. It defines the loss-to-go as $-\ln(\gamma(\mathbf{K}_{t-1})) \geq 0$ for the next optimisation step. □

Alternative proofs are in [70, 153].

Remark 17 (Approximate Design; Physical Constraints)

- *At a descriptive level, the dynamic programming, Prop. 11, consists of the evaluation pairs (conditional expectation, minimisation).*

  Except of a few cases, an approximation is needed. The complexity of the approximated evaluation prevents a systematic use of a ready approximation theory and various specific schemes are used [18, 149].

- *The FPD↳ minimises explicitly. This reduces the design↳ to a sequence of, conceptually feasible, multivariate integrations.*

- *The found optimal strategy↳ is a randomised and causal one. The physical constraints↳ are met trivially if the used ideal strategy respects them, i.e. if* $\mathrm{supp}\left[\ {}^{I}\mathsf{f}(A_t|\mathbf{K}_{t-1})\right] \subset A_t^{\star}$*, cf. (106).*

Exercises 13 (On FPD and Choice of $^{\mathsf{I}}\mathsf{f} > 0$)

1. *Show that* $-\ln(\gamma(\mathbf{K}_t)) \geq 0 \ \forall t \in t^\star$.
2. *Consider a simple DM problem and try to specify your ideal closed loop model.*
3. *Did you met the condition* $^{\mathsf{I}}\mathsf{f} > 0$*?*
4. *Did you felt that* $^{\mathsf{I}}\mathsf{f} > 0$ *has limited your freedom, when you respected it? How to meet* $^{\mathsf{I}}\mathsf{f} > 0$, *when you violated it?*
5. *How you would select the ideal pd if you do not care about some behaviour constituents, say actions?*

Relating the FPD to the Standard Design

- We shall show that standard design ↳ forms a proper subset of DM ↳ formulated as the FPD ↳.

$$\Downarrow$$

No standard DM task is neglected when considering the FPD

- This allows us to focus on the flexible FPD, which serves us for preparation of general tools for constructing DM elements ↳.
- The inspection of relation of the standard design ↳ and the FPD ↳ starts from the rectification of a flaw in the interpretation of the ideal pd ↳.

The Discrepancy in the Interpretation of the Ideal PD

- Dynamic programming, Prop. 11, which solves the standard design ↳, leads to the optimal deterministic strategy.
- The ideal pd ↳ is interpreted as closed-loop model ↳ $^{\mathsf{I}}\mathsf{f}(B) = \mathsf{f}_{^{\mathsf{I}}\mathsf{S}}(B)$, $B \in B^{\star}$, with the globally optimal strategy ↳ $^{\mathsf{I}}\mathsf{S}$.

$$\Downarrow$$

The positivity of the ideal pd ↳, $^{\mathsf{I}}\mathsf{f}(B) > 0, \; B \in B^{\star}$, required in Prop. 5 and justifying the FPD is violated as

$$^{\mathsf{I}}\mathsf{f}(\mathbf{K}_h) \equiv {}^{\mathsf{I}}\mathsf{f}(D^h) \equiv \mathsf{f}_{^{\mathsf{I}}\mathsf{S}}(D^h) = \prod_{t \in t^{\star}} \mathsf{f}(\Delta_t | A_t, \mathbf{K}_{t-1}) \delta(A_t, {}^{o}A_t),$$

where ${}^{o}A_t \in \mathrm{Arg}\min_{A_t \in A_t^{\star}} \mathsf{E}[\mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}]$.
$\mathsf{V}(\mathbf{K}_t)$ is the value function ↳ and $\delta(\cdot, \cdot)$ is Dirac delta ↳, see Prop. 11.

Randomisation of decision rules resolves the discussed discrepancy.

The FPD as a Proper Dense Extension of the Standard DM

Let on $\mathsf{S}^\star$ us consider any standard design↳ with a fixed behaviour set $B^\star$ modelled by a fixed system model M and a chosen performance index I.

Proposition 24 (The FPD Densely Extends the Standard DM)

1. *There are FPD tasks having no standard counterpart.*
2. *Any standard design↳ can be approximated to an arbitrary precision by a FPD problem with the ideal pd*

$$ {}^{\mathsf{I}\lambda}\mathsf{f} \propto \tilde{\mathsf{f}} \exp[-\mathsf{I}/\lambda] \tag{109} $$

parameterised by $\lambda > 0$ & an arbitrary pd $\tilde{\mathsf{f}} > 0$ on $B^\star$ making ${}^{\mathsf{I}\lambda}\mathsf{f}$ proper.

Proof

ad 1 Randomness of the FPD optimal strategy implies that it is not optimal one for the standard design↳ with non-constant value function↳.

The FPD Approximates Any Standard Design, Proof of 2

The following strategies ${}^{o}\mathsf{S}$, ${}^{\lambda}\mathsf{S}$, $\lambda > 0$, are well defined

$$
\begin{aligned}
{}^{o}\mathsf{S} &\in \operatorname{Arg}\min_{\mathsf{S}\in\mathsf{S}^\star}[\mathsf{E}_\mathsf{S}[\mathsf{I}], && \text{see Prop. 11,} && (110)\\
{}^{\lambda}\mathsf{S} &\in \operatorname{Arg}\min_{\mathsf{S}\in\mathsf{S}^\star}[\mathsf{E}_\mathsf{S}[\mathsf{I}] + \lambda \mathsf{D}(\mathsf{MS}||\tilde{\mathsf{f}})] && && (111)\\
&= \operatorname{Arg}\min_{\mathsf{S}\in\mathsf{S}^\star} \mathsf{D}(\mathsf{MS}||\,{}^{\mathsf{I}\lambda}\mathsf{f}), && \text{see Prop. 23, 26.}
\end{aligned}
$$

Their specifications imply the following inequalities

$$
\begin{aligned}
0 \;&\overbrace{\le}^{(110)}\; \mathsf{E}_{{}^{\lambda}\mathsf{S}}[\mathsf{I}] - \mathsf{E}_{{}^{o}\mathsf{S}}[\mathsf{I}] \;\overbrace{\le}^{\lambda\mathsf{D}(\bullet||\bullet)\ge 0}\; \mathsf{E}_{{}^{\lambda}\mathsf{S}}[\mathsf{I}] + \lambda\mathsf{D}(\mathsf{M}\,{}^{\lambda}\mathsf{S}||\tilde{\mathsf{f}}) - \mathsf{E}_{{}^{o}\mathsf{S}}[\mathsf{I}]\\
&\overbrace{\le}^{(111)}\; \mathsf{E}_{{}^{o}\mathsf{S}}[\mathsf{I}] + \lambda\mathsf{D}(\mathsf{M}\,{}^{o}\mathsf{S}||\tilde{\mathsf{f}}) - \mathsf{E}_{{}^{o}\mathsf{S}}[\mathsf{I}] = \lambda\mathsf{D}(\mathsf{M}\,{}^{o}\mathsf{S}||\tilde{\mathsf{f}}) \rightarrow_{\lambda\to 0} 0.
\end{aligned}
$$

Thus, the expected performance index $\mathsf{E}_{{}^{\lambda}\mathsf{S}}[\mathsf{I}]$ achieved by the FPD optimal strategy ${}^{\lambda}\mathsf{S}$ is arbitrarily close to the expected performance index $\mathsf{E}_{{}^{o}\mathsf{S}}[\mathsf{I}]$ achieved by the standard optimal strategy ${}^{o}\mathsf{S}$. $\square$

On FPD Extending the Standard Design

Remark 18 (On Constrained Entropy)

- *The optimisation (111) with uniform $\tilde{f}$ can be interpreted as the standard design↳ with a constraint on closed-loop entropy, [85, 89]. It well relates to Agreement 1: A chosen strategy↳ is to be embedded & actions transmitted through a real interface in which no deterministic strategy is exactly implementable.*
- *The constraint on entropy relates to constraints on computational complexity or deliberation effort [150], to "rational inattention" [152] or to a numerical design with Boltzman's machine [149].*
- *The formula (109) allows practitioners quantifying the DM aim↳ by a standard performance index↳ to use the FPD↳.*
- *An efficient learning during DM course requires a rich knowledge $\mathbf{K}_t$. This exploration ability is gained by randomised actions [58].*

A Partially Specified Ideal PD – Leave to the Fate Option

- Often, only a part ${}^{i}B$ of the behaviour B influences preferential ordering of the decision maker. The rest ${}^{r}B$ of B serves as a knowledge bearer. Consequently, in the factorised ideal pd,

$$
{}^{\mathsf{I}}\mathsf{f}(B) = {}^{\mathsf{I}}\mathsf{f}({}^{i}B|{}^{r}B)\,{}^{\mathsf{I}}\mathsf{f}({}^{r}B), \tag{112}
$$

  the decision maker is able or willing to specify ${}^{\mathsf{I}}\mathsf{f}({}^{i}B|{}^{r}B)$ only.

- We should *not* enforce anything above the designer's (user's) wishes. Thus, we have to let the design determine the distribution of quantities ${}^{r}B$ – we have to "leave them to their fate" [77]:

- *leave to the fate* means that the ideal pd of $B = ({}^{r}B, {}^{i}B)$ is set

$$
{}^{\mathsf{I}}\mathsf{f}(B) = {}^{\mathsf{I}}\mathsf{f}({}^{i}B|{}^{r}B)\,{}^{\mathsf{I}}\mathsf{f}({}^{r}B) = {}^{\mathsf{I}}\mathsf{f}({}^{i}B|{}^{r}B)\mathsf{f}_{\mathsf{S}}({}^{r}B),
$$

  where the pd $\mathsf{f}({}^{r}B)$ results from the FPD with this ideal pd.

  This option is widely exploited in methodological Part 9.

Proposition 25 (The FPD with Leave to the Fate Option)

The FPD-optimal strategy $^{o}\mathsf{S}$, *dealing with closed models* $\mathsf{f}_{\mathsf{S}}(B)$, $\mathsf{S} \in \mathsf{S}^{\star}$, $B \in B^{\star} = ({}^{r}B^{\star}, {}^{i}B^{\star})$, *and with the ideal* ${}^{I}\mathsf{f}(B) = {}^{I}\mathsf{f}({}^{i}B|{}^{r}B)\mathsf{f}_{\mathsf{S}}({}^{r}B)$ *is:*

$$^{o}\mathsf{S} \in \operatorname{Arg}\min_{\mathsf{S}\in\mathsf{S}^{\star}} \int_{B^{\star}} \mathsf{f}_{\mathsf{S}}({}^{i}B|{}^{r}B)\mathsf{f}_{\mathsf{S}}({}^{r}B) \ln\left(\frac{\mathsf{f}_{\mathsf{S}}({}^{i}B|{}^{r}B)}{{}^{I}\mathsf{f}({}^{i}B|{}^{r}B)}\right) \mathrm{d}({}^{r}B, {}^{i}B).$$

- ⋆ *The variables* ${}^{r}B$ *has no influence iff* ${}^{i}B$ *&* ${}^{r}B$ *are independent.*
- ⋆⋆ *With the leave to the fate option applied to strategy* $\mathsf{S} = {}^{I}\mathsf{S}$, *the FPD reduces to the standard design with the performance index* $\mathsf{I}(B) = \ln(\mathsf{M}/{}^{I}\mathsf{M})$ *and the FPD optimal strategy is deterministic.*

Proof

- ⋆ The common factor cancels in the logarithmic ratio.
- ⋆⋆ The closed-loop model $\mathsf{f}_{\mathsf{S}} = \mathsf{MS}$ factorises into the system model M and strategy S, (75), which after cancellation enters the optimised functional linearly and Prop. 10 applies. □

The General FPD with Hidden Quantities

Proposition 26 (Solution of the General FPD)

The strategy solving the FPD with hidden quantities has the rules, $t \in t^\star$,

$$^{o}\mathsf{f}(A_t|\mathbf{K}_{t-1}) = \frac{\exp[-\omega(A_t, \mathbf{K}_{t-1})]}{\gamma(\mathbf{K}_{t-1})}, \ \gamma(\mathbf{K}_{t-1}) \equiv \int_{A_t^\star} \exp[-\omega(A_t, \mathbf{K}_{t-1})]\, \mathrm{d}A_t$$

$$\begin{aligned}\omega(A_t, \mathbf{K}_{t-1}) = &-\int_{\Delta_t^\star} \mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \ln(\gamma(\mathbf{K}_t))\, \mathrm{d}\Delta_t \\ &-\int_{X_{t-1}^\star} \mathsf{f}(X_{t-1}|\mathbf{K}_{t-1})\, \mathrm{d}X_{t-1} \Big\{ \ln({}^{\mathsf{I}}\mathsf{f}(A_t|X_{t-1}, \mathbf{K}_{t-1})) \\ &+ \int_{X_t^\star} \mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1}) \Big[\ln\left(\frac{\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})}{{}^{\mathsf{I}}\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})}\right) \\ &+ \int_{\Delta_t^\star} \mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1}) \ln\left(\frac{\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})}{{}^{\mathsf{I}}\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})}\right) \mathrm{d}\Delta_t \Big]\, \mathrm{d}X_t \Big\}\, \mathrm{d}X_{t-1}.\end{aligned} \tag{113}$$

The backward run starts with $\gamma(\mathbf{K}_h) \equiv 1$ and uses filtering, Prop. 16.

Proof Let us define the loss-to-go corresponding to the FPD

$$-\ln(\gamma(\mathbf{K}_{t-1})) = \min_{\{\mathsf{f}(A_\tau|\mathbf{K}_{\tau-1})\}_{\tau=t}^{h}} \sum_{\tau=t}^{h} \int_{\Delta_\tau^\star, A_\tau^\star, X_\tau^\star, X_{\tau-1}^\star} \mathsf{f}(\Delta_\tau, A_\tau, X_\tau, X_{\tau-1}|\mathbf{K}_{\tau-1})$$

$$\times \ln\left(\frac{\mathsf{f}(\Delta_\tau, A_\tau, X_\tau, X_{\tau-1}|\mathbf{K}_{\tau-1})}{{}^{\mathsf{I}}\mathsf{f}(\Delta_\tau, A_\tau, X_\tau, X_{\tau-1}|\mathbf{K}_{\tau-1})}\right) \mathrm{d}\Delta_\tau \, \mathrm{d}A_\tau \, \mathrm{d}X_\tau \, \mathrm{d}X_{\tau-1}.$$

Quantities Δ_t, X_t, X_{t-1} in the ignorance $\mathbf{G}_{A_t^\star}$ have to be integrated out to get an admissible decision rule of the FPD optimal strategy. It holds

$$-\ln(\gamma(\mathbf{K}_{t-1})) = \min_{\mathsf{f}(A_t|\mathbf{K}_{t-1})} \int_{A_t^\star} \mathsf{f}(A_t|K_{t-1}) \ln\left(\frac{\mathsf{f}(A_t|\mathbf{K}_{t-1})}{\exp(-\omega(A_t, \mathbf{K}_{t-1}))}\right) \mathrm{d}A_t$$

$$\omega(A_t, \mathbf{K}_{t-1}) = \int_{\Delta_t^\star, X_t^\star, X_{t-1}^\star} \mathsf{f}(\Delta_t, X_t, X_{t-1}|A_t, \mathbf{K}_{t-1})$$

$$\times \ln\left[\frac{\mathsf{f}(\Delta_t, X_t, X_{t-1}|A_t, \mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\, {}^{\mathsf{I}}\mathsf{f}(\Delta_t, X_t, X_{t-1}|A_t, \mathbf{K}_{t-1})}\right] \mathrm{d}\Delta_t \, \mathrm{d}X_t \, \mathrm{d}X_{t-1}$$

This and the final form of $\omega(A_t, \mathbf{K}_{t-1})$ in (113) exploit marginalisation, normalisation, chain rule, natural conditions of DM, Fubini's theorem [137] and KLD minimiser, Prop. 20. □

The General FPD Determines the Needed DM Elements

- *DM elements* are specified for time $t \in t^{\star}$ up to a decision horizon $h < \infty$ and act on behaviour $B = (X^h, D^h) = (X^h, \Delta^h, A^h) = (\text{hidden quantity}_t, \text{observation}_t, \text{action}_t)_{t=1}^{h} \in B^{\star} = (X^{h\star}, \Delta^{h\star}, A^{h\star})$.

observed	data records	$D_t = (\Delta_t, A_t) \in D_t^{\star} = \mathbf{K}_t^{\star} \setminus \mathbf{K}_{t-1}^{\star}$
optimised	admissible strategy	$\mathsf{f}(A_t \vert X^{t-1}, \mathbf{K}_{t-1}) = \mathsf{f}(A_t \vert \mathbf{K}_{t-1})$
given	observation model	$\mathsf{f}(\Delta_t \vert X_t, A_t, \mathbf{K}_{t-1})$
given	time-evolution model	$\mathsf{f}(X_t \vert X_{t-1}, A_t, \mathbf{K}_{t-1})$
given	prior pd	$\mathsf{f}(X_0, \mathbf{K}_0 \vert A_1) = \mathsf{f}(X_0 \vert \mathbf{K}_0)\mathsf{f}(\mathbf{K}_0)$
given	ideal observation model	${}^{\mathsf{I}}\mathsf{f}(\Delta_t \vert X_t, A_t, \mathbf{K}_{t-1})$
given	ideal time-evolution model	${}^{\mathsf{I}}\mathsf{f}(X_t \vert X_{t-1}, A_t, \mathbf{K}_{t-1})$
given	ideal prior pd	${}^{\mathsf{I}}\mathsf{f}(X_0, \mathbf{K}_0 \vert A_1)$
given	ideal strategy	${}^{\mathsf{I}}\mathsf{f}(A_t \vert X^{t-1}, \mathbf{K}_{t-1})$

Table : DM elements.

- The needed predictor $(\mathsf{f}(\Delta_t \vert A_t, \mathbf{K}_{t-1}))_{t \in t^{\star}}$ & pds $(\mathsf{f}(X_t \vert \mathbf{K}_t))_{t \in t^{\star}}$ are gained by filtering, Prop. 16. The ways to other DM elements follow.

Exercises 14 (On FPD)

1. *Consider one-stage-ahead standard design, with $|A^{\star}| < \infty$. How you would exclude the optimal rule be deterministic by constraining entropy of admissible rules?*
2. *Solve 1 for continuous valued actions.*
3. *Think over real-life example(s) when deterministic decision rule is bad and justify why.*
4. *Proof Prop. 26.*
5. *Is it legitimate to use Pro. filtering, 16, when applying the strategy described by Prop. 26?*
6. *Consider a simple DM problem and think over how to specify the respective DM elements.*

Principles and Tools for Creating DM Elements

On this Part

Recall

The use of FPD and standard design assume ability to specify

- system model (75), $\mathsf{M} = \overbrace{\mathsf{f}(X_0)}^{\textit{prior pd}} \prod_{t \in t^\star} \overbrace{\mathsf{f}(\Delta_t, X_t | X_{t-1}, A_t, \mathbf{K}_{t-1})}^{\textit{observation} \times \textit{time evolution pds}}$
- ideal pd $^{\mathsf{I}}\mathsf{f} = {}^{\mathsf{I}}\mathsf{M}\,{}^{\mathsf{I}}\mathsf{S}$ modelling DM preferences and constraints.

This part

- discusses a (still incomplete) construction of DM elements of a supported DM task from user-supplied informal elements,
- provides the key construction tools (of independent interest) as solutions of supporting DM tasks formulated in FPD terms.

The postponed discussion

- the subset of strategies $\mathsf{S}_\star$ where the optimum ${}^{o}\mathsf{S}$ (103) is searched,
- the knowledge $K^\star$, data $D^\star$, internal quantity $X^\star$ and action $A^\star$ sets,
- abilities to evaluate the strategy ${}^{o}\mathsf{S}$, i.e. store, integrate and optimise functions in Prop. 26 and apply ${}^{o}\mathsf{S}$.

Supporting DM Tasks Serving to Real Decision Makers

A real decision maker is imperfect

- *imperfect decision maker* is unable to
  - (a) specify all needed DM elements↳
  - (b) perform all evaluations in allocated time and with available resources.

A real decision maker needs tools that

- convert its knowledge↳ & DM preferences into DM elements↳
- respect its limited cognitive resources.

The next text provides approximation principle↳ (116) & minimum KLD principle↳ (119) for constructing DM elements↳ from incomplete, practically available, knowledge and preference pieces.

- The tools approximate pds, extend (non-)probabilistic knowledge, Sec. 22, and merge pds, Sec. 24. Special tools are in Part 10.

These tools are conjectured to be sufficient for constructing DM elements↳.

Approximation & extension of fragmental pd↳s support the construction of DM elements↳. They are formulated as supporting DM tasks solved by FPD.

Local Notions & Notation

- *fragmental pd* is a common name we use for all possible marginal and conditional pds derived from the joint pd $\mathsf{f}(B)$.
- The DM elements↳ of the supporting DM task are denoted by calligraphic counterparts of those used in the supported DM task.
- Technicalities connected with infinite-dimensional random variables are avoided by assuming the behaviour $B \in B^\star$ of the supported DM to have a finite number of behaviour instances, $|B^\star| < \infty$.
- Thus, the pds in the supported DM are finite-dimensional vectors

$$\mathsf{f} \in \mathsf{f}^\star \subset \mathsf{f}^\star_\triangle \equiv \left\{ \mathsf{f}(B) : \mathsf{f}(B) \geq 0, \int_{B^\star} \mathsf{f}(B)\,\mathrm{d}B = 1 \right\} \qquad (114)$$

and pds like $\mathcal{F}(B, \mathsf{f})$ in the supporting DM make a good sense.

Approximation of a Known PD as Supporting DM Task

- A known pd $\mathsf{f}(B)$, $B \in B^\star$, constructed from the knowledge available in the supported task, is often complex and is approximated by a pd $\hat{\mathsf{f}}$

$$\hat{\mathsf{f}} \in \hat{\mathsf{f}}^\star \subset \mathsf{f}^\star_\triangle = \left\{ \mathsf{f}(B) : \mathsf{f}(B) \geq 0, \int_{B^\star} \mathsf{f}(B)\,\mathrm{d}B = 1 \right\}.$$

- *approximation* is a supporting DM task specified by

action	$\mathcal{A} \equiv \hat{\mathsf{f}} \in \mathcal{A}^\star = \hat{\mathsf{f}}^\star \equiv$ the approximating pds
knowledge	$\mathcal{K}_{\mathcal{A}^\star} = \mathsf{f} \equiv$ the approximated pd
ignorance	$\mathcal{G}_{\mathcal{A}^\star} = B \equiv$ behaviour of the supported DM
behaviour	$\mathcal{B} \equiv (B, \hat{\mathsf{f}}, \mathsf{f}) =$(behaviour of the supported DM, an approximating pd, the approximated pd)
system model	$\mathcal{M}(\mathcal{G}_{\mathcal{A}^\star} \vert \mathcal{A}, \mathcal{K}_{\mathcal{A}^\star}) \equiv \mathcal{M}(B \vert \hat{\mathsf{f}}, \mathsf{f}) = \mathsf{f}(B)$ reflects that the known $\mathsf{f} = \mathcal{K}_{\mathcal{A}^\star}$ models $B = \mathcal{G}_{\mathcal{A}^\star}$ for any $\hat{\mathsf{f}} = \mathcal{A}$
decision rule	$\mathcal{S}(\hat{\mathsf{f}} \vert \mathsf{f})$ acts on $\hat{\mathsf{f}}^\star$ and specifies
closed-loop model	$\mathcal{F}_{\mathcal{S}}(\mathcal{G}_{\mathcal{A}^\star}, \mathcal{A} \vert \mathcal{K}_{\mathcal{A}^\star}) = \mathsf{f}(B)\mathcal{S}(\hat{\mathsf{f}} \vert \mathsf{f})$.

Ideal Closed-Loop Model for Approximation of Known PD

It remains to specify the ideal pd reflecting the DM aim & constraints: to approximate a known pd $\mathsf{f}(B)$ by $\mathcal{A} \equiv \hat{\mathsf{f}}(B) \in \hat{\mathsf{f}}^{\star}$, $\mathcal{G}_{\mathcal{A}^{\star}} \equiv B \in B^{\star}$.

- The ideal closed-loop model for this task is specified by

$$^{\mathsf{I}}\mathcal{F}(\mathcal{G}_{\mathcal{A}^{\star}}, \mathcal{A}|\mathcal{K}_{\mathcal{A}^{\star}}) = {}^{\mathsf{I}}\mathcal{F}(B, \hat{\mathsf{f}}|\mathsf{f}) = {}^{\mathsf{I}}\mathcal{F}(B|\hat{\mathsf{f}}, \mathsf{f})\, {}^{\mathsf{I}}\mathcal{F}(\hat{\mathsf{f}}|\mathsf{f}) = \hat{\mathsf{f}}(B)\mathcal{S}(\hat{\mathsf{f}}|\mathsf{f}). \quad (115)$$

Motivation

- The first two equalities just use meaning of involved symbols and chain rule. The real options are made in the last equality.
- The choice of the first factor in (115) means that the approximating pd should ideally describe the behaviour B of the supported DM task.
- The choice of the second factor in (115) expresses a lack of additional requirements on the constructed decision rule $\mathcal{S}(\hat{\mathsf{f}}|\mathsf{f})$.

  The decision rule $\mathcal{S}(\hat{\mathsf{f}}|\mathsf{f})$ resulting from the design is accepted as the ideal one, the leave to the fate option is used.

Proposition 27 (Approximation of a Known PD as the FPD: Solution)

The supporting FPD task of approximation of a known pd f *leads to*

- *approximation principle* $\quad {}^{o}\hat{\mathsf{f}} \in \mathrm{Arg}\min_{\hat{\mathsf{f}}\in\hat{\mathsf{f}}^{\star}} \mathsf{D}(\mathsf{f}||\hat{\mathsf{f}}). \quad (116)$

Proof For the ideal pd (115), KLD (103) of $\mathcal{F}_{\mathcal{S}}$ from ${}^{\mathsf{I}}\mathcal{F}$ is linear in $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^{\star}}) = \mathcal{S}(\hat{\mathsf{f}}|\mathsf{f})$. Thus, the FPD reduces to standard design.

Basic DM lemma, Prop. 10, gives the optimal deterministic decision rule of the supporting approximation DM task $\mathcal{S}(\mathcal{A}|\mathsf{f}) = \delta(\mathcal{A}, {}^{o}\mathcal{A})$, where $\delta(\cdot,\cdot)$ is Dirac delta and ${}^{o}\mathcal{A} = {}^{o}\hat{\mathsf{f}} \in \hat{\mathsf{f}}^{\star}$ with

$$
\begin{aligned}
{}^{o}\hat{\mathsf{f}} \quad &\in \quad \mathrm{Arg}\min_{\hat{\mathsf{f}}\in\hat{\mathsf{f}}^{\star}} \mathsf{E}\left[\ln(\mathsf{f}/\hat{\mathsf{f}})\middle|\mathcal{A} = \hat{\mathsf{f}}, \mathcal{K}_{\mathcal{A}^{\star}} = \mathsf{f}\right] \\
&= \quad \mathrm{Arg}\min_{\hat{\mathsf{f}}\in\hat{\mathsf{f}}^{\star}} \int_{B^{\star}} \mathsf{f}(B) \ln\left(\frac{\mathsf{f}(B)}{\hat{\mathsf{f}}(B)}\right) \mathrm{d}B. \qquad \square
\end{aligned}
$$

Remark 19 (On Approximation)

- *This approximation principle ↳ was alternatively justified in [15].*
- *The approximation principle ↳ is valid for $|B^\star| = \infty$. The formula (116) does not exploit finiteness of $|B^\star|$. Formally:*
  - *The case $|B^\star| = \infty$ can be inspected by considering behaviours of a finite cardinality as images of discretisation mappings from the space of behaviours having infinite dimension.*
  - *Prop. 1 implies separability of the underlying ordered space. This implies that the discretisations can be infinitely refined with a smooth transition from finite cardinality to infinite one.*

Approximation by Normal PD

Example 14 (Approximation of a Known PD by Normal PD)

Let a known pd $\mathsf{f}(B)$, $B \in B^\star$, *be approximated by normal pd with an expected value* $\bar{B}$ *& a covariance* $\mathbb{R}$

$$\hat{\mathsf{f}} \in \hat{\mathsf{f}}^\star \equiv \{\mathsf{f}(B) : \mathsf{f}(B) \equiv \mathcal{N}_B(\bar{B}, \mathbb{R})\}.$$

Then, the pd ${}^o\hat{\mathsf{f}}(B) = \mathcal{N}_B({}^o\bar{B}, {}^o\mathbb{R})$ *resulting from approximation principle has the same initial moments as the approximated pd* $\mathsf{f}(B)$*:*

$${}^o\bar{B} \equiv \int_{B^\star} B\mathsf{f}(B)\,\mathrm{d}B \equiv \bar{B}_{\mathsf{f}}, \quad {}^o\mathbb{R} \equiv \int_{B^\star} (B - {}^o\bar{B})(B - {}^o\bar{B})'\mathsf{f}(B)\,\mathrm{d}B \equiv \mathbb{R}_{\mathsf{f}}.$$

Proof

$$\begin{aligned}
\operatorname{Arg}\min_{\hat{\mathsf{f}} \in \hat{\mathsf{f}}^\star} \mathsf{D}(\mathsf{f}||\hat{\mathsf{f}}) &= \operatorname{Arg}\min_{\hat{\mathsf{f}} \in \hat{\mathsf{f}}^\star} \left\{ -\int_{B^\star} \ln\left(\hat{\mathsf{f}}\right) \mathsf{f}(B)\,\mathrm{d}B \right\} \\
&= \operatorname{Arg}\min_{\bar{B},\mathbb{R}} \left[\ln(|\mathbb{R}|) + \int_{B^\star} (B - \bar{B})'\mathbb{R}^{-1}(B - \bar{B})\mathsf{f}(B)\,\mathrm{d}B \right] \\
\nabla_{\mathbb{R}} \ln(|\mathbb{R}|) &= \mathbb{R}^{-1}; \;\; \nabla_{\mathbb{R}} \mathrm{tr}\left[\mathbb{R}^{-1}\mathbb{R}_{\mathsf{f}}\right] = -\mathbb{R}^{-2}\mathbb{R}_{\mathsf{f}}, \;\; \nabla_{\mathbb{R}} \equiv \textit{matrix gradient}. \;\square
\end{aligned}$$

Approximation by Uniform PD

Example 15 (Approximation of Known PD by Uniform PD)

Let a known pd $\mathsf{f}(B)$, $B \in B^\star$ *be approximated by a uniform pd* $\mathsf{U}_B(B_\star)$ *with support* $B_\star \subset B^\star$

$$\hat{\mathsf{f}} \in \hat{\mathsf{f}}^\star \equiv \left\{\mathsf{f}(B) : \mathsf{f}(B) \equiv \mathsf{U}_B(B_\star) = \left\{ \begin{array}{lll} \mathsf{vol}^{-1}(B_\star) & \textit{if} & B \in B_\star \subset B^\star \\ 0 & \textit{if} & B \notin B_\star \end{array} \right\}\right. .$$

Then, the pd ${}^o\hat{\mathsf{f}}(B) = \mathsf{U}_B({}^oB_\star)$ *with the set* $B_\star$ *containing* $\mathrm{supp}\,[\,\mathsf{f}]$ *and having the smallest volume* $\mathsf{vol}(B_\star)$ *among such behaviour subsets.*
If $\mathsf{vol}(\mathrm{supp}\,[\,\mathsf{f}]) = \infty$ *then no uniform pd approximating* f *exists.*

Proof KLD $\mathsf{D}(\mathsf{f}||\mathsf{U}(B_\star))$ is infinite if $\mathrm{supp}\,[\,\mathsf{f}]\,(\cap B^\star \setminus B_\star) \neq \emptyset$, see Prop. 20. For $B_\star \supset \mathrm{supp}\,[\,\mathsf{f}]$,

$$\mathsf{D}(\mathsf{f}||\mathsf{U}(B_\star)) = \int_{B^\star} \mathsf{f}(B)\ln(\mathsf{f}(B))\,\mathrm{d}B + \mathsf{vol}(B_\star),$$

which proofs the first statement. The latter follows from the fact that for $B_\star$ covering $\mathrm{supp}\,[\,\mathsf{f}]$, it holds $\mathsf{vol}(B_\star) \supset \mathsf{vol}(\mathrm{supp}\,[\,\mathsf{f}])$. $\square$

Exercises 15 on Approximation Principle

Exercises 15 (On Approximation and KLD)

1. *Consider the given pdf* $\mathsf{f}(B) = \mathsf{f}({}^{a}B, {}^{b}B)$ *and approximating pds of the form* $\hat{\mathsf{f}}({}^{a}B, {}^{b}B) = \hat{\mathsf{f}}({}^{a}B)\hat{\mathsf{f}}({}^{b}B)$. *Approximate optimally* f *by* $\hat{\mathsf{f}}$.

2. *Consider the task 1 but use* $\mathsf{D}(\hat{\mathsf{f}}||\mathsf{f})$ *as the optimised proximity measure. (It is widely used by variational Bayesian methods [154].)*

3. *Provide an example showing that the KLD used by variational Bayes may be unreasonable. (Hint: take the case with* $\mathrm{supp}\left[\hat{\mathsf{f}}\right] \subsetneq \mathrm{supp}\,[\mathsf{f}]$*.)*

4. *What is drawback popular* L_p *norms*

$$||\mathsf{f} - \hat{\mathsf{f}}||_p = \left(\int_{B^\star} |\mathsf{f}(B) - \hat{\mathsf{f}}(b)|^p \,\mathrm{d}B\right)^{1/p}, \quad p \geq 1$$

in the role of proximity measures? (Hint: think about dependence of their values on a bijective transformation of the behaviour.)

Approximation of an *Unknown* PD as FPD: Behaviour $\mathcal{B}$

We consider approximation of unknown pd as the next supporting DM task:

- The supporting action $\mathcal{A} = \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star}) \in \mathcal{F}^\star$ is a pd on the unknown pds $\mathsf{f} \in \mathsf{f}^\star_\Delta$ describing the behaviour $B \in B^\star$ of the supported DM.
- The knowledge $\mathcal{K}_{\mathcal{A}^\star}$ about the approximated pd f is

$$\mathcal{K}_{\mathcal{A}^\star}: \quad \mathsf{f} \in \mathsf{f}^\star \subset \mathsf{f}^\star_\triangle = \left\{ \mathsf{f}(B) : \mathsf{f}(B) \geq 0, \int_{B^\star} \mathsf{f}(B)\,\mathrm{d}B = 1 \right\},$$

$$\mathsf{f}_0 \in \mathsf{f}^\star_\triangle \quad \text{is the best prior (possibly flat) guess of f.} \tag{117}$$

- The ignorance $\mathcal{G}_{\mathcal{A}^\star} = (B, \mathsf{f}) \equiv$(behaviour, its pd) of the supported DM.
- The behaviour $\mathcal{B}$ of the supporting task is

$$\mathcal{B} \equiv [\mathcal{G}_{\mathcal{A}^\star}, \mathcal{A}, \mathcal{K}_{\mathcal{A}^\star}] \equiv [(B, \mathsf{f}), \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star}), (\mathsf{f}_0, \mathsf{f}^\star)] =$$

$= [(B$ of supported DM, its pd f), opted pd on $\mathsf{f}^\star$,(prior guess of $\mathsf{f} \in \mathsf{f}^\star_\triangle$,$\mathsf{f}^\star$)].

Approximation of an *Unknown* PD as FPD: Models

- The considered system model

$$\mathcal{M}(\mathcal{G}_{\mathcal{A}^\star}|\mathcal{A},\mathcal{K}_{\mathcal{A}^\star}) = \mathcal{M}(B,\mathsf{f}|\mathcal{F}(\cdot|\mathcal{K}_{\mathcal{A}^\star}),(\mathsf{f}_0,\mathsf{f}^\star)) = \mathsf{f}(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$$

respects that the pd $\mathsf{f}(B)$ models B and the action $\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$ models f.

- The optimised randomised decision rule $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$ completes the specification of the closed-loop model

$$\mathcal{F}_{\mathcal{S}}(\mathcal{G}_{\mathcal{A}^\star},\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) = \mathsf{f}(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$$

- The ideal closed-loop model is specified as

$$^{\mathsf{I}}\mathcal{F}(\mathcal{G}_{\mathcal{A}^\star},\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) = \mathsf{f}_0(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}). \tag{118}$$

- (118) takes $\mathsf{f}_0(B)$ as the prior ideal model of the behaviour B.
- No other wish exists on the action $\mathcal{A} = \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$ and decision rule $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$ generating it. The leave to the fate option is used.

Approximation of an *Unknown* PD as FPD: Solution

Proposition 28 (Minimum KLD Principle)

The supporting FPD task of approximation of an unknown pd f *leads to:*

- *minimum KLD principle that recommends to complete knowledge*↳ *expressed by the set* $\mathsf{f}^\star$ *and the prior guess* f_0 *of* f *to*

$$^{o}\mathsf{f} \in \operatorname{Arg}\min_{\mathsf{f}\in\mathsf{f}^\star} \mathsf{D}(\mathsf{f}||\mathsf{f}_0). \tag{119}$$

Proof Due to the leave to the fate↳ option, the decision rule↳ enters KLD↳ linearly and the optimum is deterministic $^{o}\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) = \delta(\mathcal{A}, {}^{o}\mathcal{A})$

$$
\begin{aligned}
^{o}\mathcal{A} &= {}^{o}\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star}) \in \operatorname{Arg}\min_{\mathcal{F}^\star} \mathsf{E}[\ln(\mathsf{f}/\mathsf{f}_0)|\mathcal{A},\mathcal{K}_{\mathcal{A}^\star}] \\
&= \operatorname{Arg}\min_{\mathcal{F}^\star(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})} \int_{\mathsf{f}^\star} \mathcal{F}(f|\mathcal{K}_{\mathcal{A}^\star}) \overbrace{\int_{B^\star} \mathsf{f}(B)\ln(\mathsf{f}(B)/\mathsf{f}_0(B))\,\mathrm{d}B}^{\mathsf{D}(\mathsf{f}||\mathsf{f}_0)}\,\mathrm{d}\mathsf{f}
\end{aligned}
$$

The last expression is linear in $\mathcal{F}(f|\mathcal{K}_{\mathcal{A}^\star})$ and its minimum is reached for $^{o}\mathcal{F}(f|\mathcal{K}_{\mathcal{A}^\star})$ with a full mass on $\mathsf{f} \in \mathsf{f}^\star$ minimising KLD $\mathsf{D}(\mathsf{f}||\mathsf{f}_0)$. □

Remark 20 (On Minimum KLD Principle)

- *Minimum KLD principle also applies for* $|B^{\star}| = \infty$*, cf. Remark 19.*
- *The optimal pd* ${}^{o}\mathsf{f} = \mathsf{f}_0 =$ *the prior guess of* f *if no constraints are put on* f*, i.e. if* $\mathsf{f}_0 \in\in \mathsf{f}^{\star} \subset \mathsf{f}^{\star}_{\triangle}$ *(114).*
- *maximum entropy principle coincides with the minimum KLD principle*$_{\downarrow}$ *for uniform prior guess* f_0 *of the unknown approximated pd.*
- *Both principles are axiomatically justified in [148] for the set* $\mathsf{f}^{\star} \subset \mathsf{f}^{\star}_{\triangle}$ *through a value* $\bar{\phi}$ *of a linear mapping on* $\mathsf{f}^{\star}_{\triangle}$ *(114) given by a kernel* ϕ

$$\mathsf{f}^{\star} \equiv \left\{ \mathsf{f} \in \mathsf{f}^{\star}_{\triangle} : \bar{\phi} = \int_{B^{\star}} \phi(B)\mathsf{f}(B)\,\mathrm{d}B \equiv \mathsf{E}_{\mathsf{f}}\left[\phi|\bar{\phi}\right] \right\}. \tag{120}$$

- *The constraints (120) are structural. It means that the value* $\bar{\phi}$ *specifying (120) need not to be known! The resulting pd* ${}^{o}\mathsf{f}$ *is then parameterised by the unknown* $\bar{\phi}$*.*

Example 16 (Normal PD Implied by the Minimum KLD Principle)

Let the knowledge $\mathcal{K}_{\mathcal{A}^\star}$ about an unknown pd $\mathsf{f} \in \mathsf{f}^\star$ *be*

$$\begin{aligned} \mathsf{f}^\star \;\equiv\; & \{\mathsf{f}(B) : B^\star \textit{ is finite-dimensional real space} \\ & \mathsf{E}_\mathsf{f}[B|\bar{B}] = \bar{B},\; \mathsf{cov}_\mathsf{f}[B|\mathbb{R}] = \mathbb{R} > 0\} \\ & \textit{prior guess } \mathsf{f}_0 \textit{ of } \mathsf{f} \textit{ is improper uniform pd on } B^\star. \end{aligned}$$

Then the minimum KLD principle↳ *coincides with the maximum entropy principle*↳ *and recommends to use* ${}^{\mathrm{o}}\mathsf{f}(B|\bar{B},) = \mathcal{N}_B(\bar{B}, \mathbb{R})$.

Proof rearranges the Lagrangian, given by multipliers μ, $\mathbb{Q}$, into KLD

$$\int_{B^\star} \mathsf{f}(B)\ln(\mathsf{f})\,\mathrm{d}B + \mu' \int_{B^\star} B\mathsf{f}(B)\,\mathrm{d}B + \mathrm{tr}\left[\mathbb{Q}\int_{B^\star}(B-\bar{B})(B-\bar{B})'\mathsf{f}(B)\,\mathrm{d}B\right] \quad \square$$

Completion of Common Deterministic Relations

The minimum KLD principle↳ serves for construction of DM elements↳ from incomplete information provided by the imperfect decision maker↳.

The processed information mostly has the form

- Expected or desired deterministic relations, which can be expressed via moments of a finite collection of nonlinear functions

$$0 \approx \phi(B) = \big(\phi_\kappa(B)\big)_{\kappa \in \kappa^\star}, \;\; \kappa^\star \equiv \{1, \ldots, |\kappa^\star|\}, \;\; |\kappa^\star| < \infty. \qquad (121)$$

  They express the domain-specific theoretical models, empirical rules, universal approximations of relations, empirical relations, etc.

This knowledge restricts the simplex $\mathsf{f}^\star_\triangle$ (114) to the set $\mathsf{f}^\star$ in (117) via

$$\mathsf{E}_\mathsf{f}[\phi_\kappa] = \int_{B^\star} \phi_\kappa(B)\mathsf{f}(B)\,\mathrm{d}B = 0, \quad \kappa \in \kappa^\star. \qquad (122)$$

On Richness of the Information Expression (122)

Remark 21 (On Variants of Constraints $\mathsf{E}_{\mathsf{f}}[\phi_\kappa] = 0$)

- *The discussed constraints often concern of a behaviour↳ part, e.g.*

$$\int_{X_t^\star} \phi(X_t, A_t, X_{t-1}, \mathbf{K}_{t-1}) \mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1}) \, \mathrm{d}X_t = 0, \ t \in t^\star, \quad (123)$$

  for which the application of the minimum KLD principle↳ leads to time-evolution model↳ $\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})$.
- *Analogies of (123) and the minimum KLD principle↳ provide observation model↳, prior pd↳ and possibly their ideal counterparts. Thus, they provide majority of DM elements↳.*
- *The functions ϕ_κ can be parameterised by an additional unknown parameter, which becomes a part of the ignorance↳.*
- *The constraints can also delimit the expected or desired range of a behaviour↳ part by taking $\phi_\kappa = \chi_\kappa - 1$, $\chi_\kappa =$ the indicator↳ of this range. A soft specification can be done via behaviour moments.*

Example of Completion of Deterministic Relation

Example 17 (Free Fall of a Ball in Gravitational Field)

Hidden state $X_t =$ *[position,velocity,acceleration]' of a ball in the field with acceleration* A_t *at time* $t = \text{floor}\left[\frac{\textit{real time}}{\textit{sampling period } T}\right]$ *evolves*

$$X_t - \mathbb{A}X_{t-1} - \mathbb{B}A_t = \varepsilon_t \approx 0,\ \mathbb{A} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ 0 & 0 & 0 \end{bmatrix},\quad \mathbb{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

observation $\Delta_t - \mathbb{C}X_t + \mathbb{D} = e_t \approx 0,\quad \mathbb{C} = [1,\, 0,\, 0],\ \mathbb{D} =$ *constant.*

- *Wind, friction, discretisation errors caused by velocity & acceleration variations in* $((t-1)T, tT), \dots$ *make* $\varepsilon_t \neq 0$ *but give* $\mathsf{E}_{\mathsf{f}}[X_t - \mathbb{A}X_{t-1} - \mathbb{B}A_t | A_t, X_{t-1}] = 0$. *This and the precision model* $\mathsf{cov}_{\mathsf{f}}[X_t | A_t, X_{t-1}] = \mathbb{R}(A_t, X_{t-1})$ *give normal time-evolution model.*
- *Observation errors & neglect of a sensor non-linearity make* $e_t \neq 0$. *With* $\mathbb{D}$ *compensating bias, normal observation model is gained.*
- *Constraints on ranges of* ε_t, e_t *yield models with uniform innovations*

Proposition 29 (Minimum KLD Principle for Inequality Constraints)

Let the knowledge $\mathcal{K}_{\mathcal{A}^\star}$ *about* f *incorporates its prior guess* f_0 *and*

$$\mathsf{f} \in \mathsf{f}^\star = \{\mathsf{f} \in \mathsf{f}^\star_{\Delta^\star},\ \mathsf{E}_\mathsf{f}[\phi_\kappa] \le 0,\ \kappa \in \kappa^\star\}. \tag{124}$$

Then, the minimum KLD principle (119) recommends the pd ${}^o\mathsf{f}(B|\mathcal{K})$*:*

$$ {}^o\mathsf{f}(B|\mathcal{K}) \propto \mathsf{f}_0(B) \exp\left[-\sum_{\kappa\in\kappa^\star} \lambda_k \phi_\kappa(B)\right] \in \mathrm{Arg} \min_{\mathsf{f}:\ (\mathsf{E}_\mathsf{f}[\phi_\kappa]\le 0)_{\kappa\in\kappa^\star}} \mathsf{D}(\mathsf{f}||\mathsf{f}_0) \tag{125}$$

with Kuhn-Tucker multipliers $(\lambda_\kappa \ge 0)_{\kappa\in\kappa^\star}$ *chosen so that the constraints specifying* $\mathsf{f}^\star$ *are met. It is not a priori guaranteed that* $\mathcal{K}_{\mathcal{A}^\star}$ *is not*

- *incompatible knowledge makes* $\mathsf{f}^\star \cap \mathrm{supp}\,[\mathsf{f}_0] = \emptyset$, *i.e. no Kuhn-Tucker multipliers meeting the constraints are found.*

Proof It converts, via a simple algebra, the Kuhn-Tucker functional into KLD of the optimised pd to the pd (125). □

Remark 22 (On Outcomes of Minimum KLD principle)

- *The frequent case of incompatible knowledge↳ is addressed later on.*
- *The resulting pd (125) can be too complex for a further handling in DM task. Then, the approximation principle↳ is to be applied.*
- *Often, the decision maker↳s describe the knowledge↳ or preferences, incompletely. Then, the minimum KLD principle↳, only provides fragmental pd↳, which can be exploited after making:*
- *extension of fragmental pds expands fragmental pd↳s offered by a collection of information sources (decision makers), indexed by $\kappa \in \kappa^\star$, to behaviour↳s $B \in B^\star$ containing all quantities considered by them.*

As in all supporting tasks, the FPD formalisation of the extension of fragmental pds is used. In it, the behaviour splits $B = (\mathcal{U}_\kappa, \mathbf{G}_\kappa, \mathbf{K}_\kappa) \equiv$

$\equiv$ the part (uninteresting for, modelled by, known to) κ-th knowledge source.

- The action $\mathcal{A} \in \mathcal{A}^\star$ of the supporting DM task is the pd $\mathcal{F}({}^e\mathsf{f}_\kappa)$ on possible extensions pd ${}^e\mathsf{f}_\kappa(B)$ of the given pd $\mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)$.
- The ignorance $\mathcal{G}_{\mathcal{A}^\star} \in (B^\star, \mathsf{f}^\star_\triangle)$, (114), is

$$\mathcal{G}_{\mathcal{A}^\star} = [B, {}^e\mathsf{f}_\kappa(B)] = [B \text{ of supported DM}, \text{extension of the pd } \mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)].$$

- The knowledge $\mathcal{K}_{\mathcal{A}^\star}$ delimits the set of possible extensions

$$ {}^e\mathsf{f}^\star_\kappa \equiv \left\{\text{pds } {}^e\mathsf{f}(B) \text{ on } B^\star \text{ with } {}^e\mathsf{f}(\mathbf{G}_\kappa|\mathbf{K}_\kappa) = \mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)\right\}. \qquad (126)$$

and a prior guess ${}^e\mathsf{f}_0(B)$ of possible good extensions.

Extension of PDs as Supporting DM Task: Solution

Recall: an extension of fragmental pds $\mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)$ splits $B = (\mathcal{U}_\kappa, \mathbf{G}_\kappa, \mathbf{K}_\kappa)$.

Proposition 30 (The Optimal Extension)

Let pds ${}^{e}\mathsf{f}_0(B) = {}^{e}\mathsf{f}_0(\mathcal{U}_\kappa, \mathbf{G}_\kappa, \mathbf{K}_\kappa)$ *and* $\mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)$ *be given. Then, the pd*

$$ {}^{e}\mathsf{f}_\kappa(B) = {}^{e}\mathsf{f}_0(\mathcal{U}_\kappa|\mathbf{G}_\kappa, \mathbf{K}_\kappa)\mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)\, {}^{e}\mathsf{f}_0(\mathbf{K}_\kappa), \tag{127} $$

is the unique extension by the minimum KLD principle in the set (126). Pds ${}^{e}\mathsf{f}_0(\mathcal{U}_\kappa|\mathbf{G}_\kappa, \mathbf{K}_\kappa)$, ${}^{e}\mathsf{f}_0(\mathbf{K}_\kappa)$ *are derived from the pd* ${}^{e}\mathsf{f}_0(B)$.

Proof

- ${}^{e}\mathsf{f}_\kappa \in {}^{e}\mathsf{f}^\star_\kappa$ factorises according to the split $B = (\mathcal{U}_\kappa, \mathbf{G}_\kappa, \mathbf{K}_\kappa)$.
- The factor ${}^{e}\mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa) = \mathsf{f}_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa)$ is fixed.
- By equating the remaining factors to their counterparts in the given ${}^{e}\mathsf{f}_0$, the KLD $\mathsf{D}({}^{e}\mathsf{f}||\,{}^{e}\mathsf{f}_0)$ reaches its minimal zero value.

Approximation of an unknown pd can be treated as the supporting DM task:

- The supporting action ↳

$$\mathcal{A} = \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star}) \in \mathcal{F}^\star \equiv \text{pds with support on } \mathsf{f}^\star \subset \mathsf{f}^\star_\triangle.$$

- The knowledge ↳ $\mathcal{K}_{\mathcal{A}^\star}$ about the pd ↳ f, alternative to (117), is:

$$\mathcal{K}_{\mathcal{A}^\star}: \quad \mathsf{f}^\star \subset \mathsf{f}^\star_\triangle \tag{128}$$
$$\mathcal{F}_0 = \mathcal{A}_0 \text{ a pd on } \mathsf{f}^\star_\triangle, \text{ a prior guess of the action } \mathcal{A} \equiv \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$$

- The ignorance ↳ $\mathcal{G}_{\mathcal{A}^\star} = (B, \mathsf{f}) \equiv$(behaviour, its pd) of the supported DM.

- The behaviour ↳ $\mathcal{B}$ of the supporting task is

$$\mathcal{B} \equiv [\mathcal{G}_{\mathcal{A}^\star}, \mathcal{A}, \mathcal{K}_{\mathcal{A}^\star}] \equiv [(B, \mathsf{f}), \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star}), (\mathcal{F}_0, \mathsf{f}^\star)]$$
$$= [(B \text{ of supported DM, its pd f}), \text{ opted pd on } \mathsf{f}^\star, (\text{prior guess } \mathcal{F}_0, \mathsf{f}^\star)].$$

Alternative Approximation of an *Unknown* PD: Models

- The considered system model ↳

$$\mathcal{M}(\mathcal{G}_{\mathcal{A}^\star}|\mathcal{A},\mathcal{K}_{\mathcal{A}^\star}) = \mathcal{M}(B,\mathsf{f}|\mathcal{F}(\cdot|\mathcal{K}_{\mathcal{A}^\star}),(\mathcal{F}_0,\mathsf{f}^\star)) = \mathsf{f}(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$$

  respects that the pd $\mathsf{f}(B)$ models B and the action $\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})$ models f.

- The optimised randomised decision rule ↳ $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$ completes the specification of the closed-loop model ↳

$$\mathcal{F}_{\mathcal{S}}(\mathcal{G}_{\mathcal{A}^\star},\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) = \mathsf{f}(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$$

- The ideal closed-loop model ↳ (118) changes to

$$^{\mathsf{I}}\mathcal{F}(B,\mathsf{f},\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) = \mathsf{f}(B)\mathcal{F}_0(\mathsf{f})\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star}) \text{ respecting} \qquad (129)$$

  that $\mathsf{f}(B)$ models B and $\mathcal{F}_0(\mathsf{f})$ is the only available prior model of f.

- The leave to the fate ↳ applies to $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$ as no wish concerns it.

Proposition 31 (Generalised Minimum KLD Principle)

The supporting FPD task of approximation of an unknown pd f *leads to*

- *generalised minimum KLD principle , which completes knowledge expressed by the set* $\mathsf{f}^\star$ *and the prior guess* $\mathcal{F}_0$ *of the action* $\mathcal{F}$ *to*

$$^{o}\mathcal{F} \in \operatorname{Arg}\min_{\mathcal{F}\in\mathcal{F}^\star} \int_{\mathsf{f}^\star} \mathcal{F}(\mathsf{f}) \ln\left(\frac{\mathcal{F}(\mathsf{f})}{\mathcal{F}_0(\mathsf{f})}\right) \mathrm{d}\mathsf{f}, \qquad \operatorname{supp}[\mathcal{F}^\star] = \mathsf{f}^\star. \quad (130)$$

Proof In KLD, the factors left to their fate cancel. In the expectation, given by the closed-loop model $\mathsf{f}(B)\mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{A}^\star})\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$, the integration over $B^\star$ “removes” the unknown $\mathsf{f}(B)$ due to normalisation of pds.

The optional decision rule $\mathcal{S}(\mathcal{A}|\mathcal{K}_{\mathcal{A}^\star})$ enters the minimised functional linearly. Thus, the minimiser is deterministic and concentrated on the smallest value of $\mathsf{D}(\mathcal{F}||\mathcal{F}_0)$ reachable by $\mathcal{F}$ with the given support. □

On Generalised Minimum KLD Principle

Remark 23 (Alternative Prior & Ideal PDs)

- *The use of* $\mathcal{F}_0$ *in the ideal pd seems to be more plausible than that of* f_0 *in the ideal leading to the minimum KLD principle*$_{\looparrowright}$.
- *Intuitively, when* $\mathcal{F}_0(\mathsf{f})$ *concentrates of a prior guess* f_0 *of* f *then* ${}^{o}\mathcal{F}$ *concentrates on the solution* ${}^{o}\mathsf{f}$ *(119) of the minimum KLD principle*$_{\looparrowright}$. *The gained randomisation supports exploration if need be.*
- *[26] proves that the minimum KLD principle*$_{\looparrowright}$ *often coincides with conditioning if* $\mathsf{f}^{\star}$ *restricts* f *linearly. We extend this to nonlinear cases.*

Processing of incompatible knowledge↳, with the incompatibility caused

$$\emptyset = {}^{\emptyset}\mathsf{f}^{\star} \equiv \{\mathsf{f} :\ \mathsf{E}_{\mathsf{f}}[\phi_{\kappa}] \le 0, \kappa \in \kappa^{\star}\},\ \text{see Prop. 29},$$

resolves by using non-empty, often smallest, subsets of these conditions.

This gives a collection of pds $\mathsf{f}_{\kappa} = {}^{o}\mathsf{f}_{\kappa} \in \mathsf{f}^{\star}_{\triangle}$, which are

- *incompletely compatible pds* describe the same behaviour $B \in B^{\star}$ but

$$\mathsf{E}_{\mathsf{f}_{\kappa}}[\phi_{k}] \ne \mathsf{E}_{\mathsf{f}_{\tilde{\kappa}}}[\phi_{k}] \text{ for some } \kappa, \tilde{\kappa} \in \kappa^{\star}$$

.

- If incompletely compatible pds have a single acceptable representant

$$\hat{\mathsf{f}} \in \mathsf{f}^{\star} \subset \mathsf{f}^{\star}_{\triangle} \text{ with } \mathsf{f}^{\star} \ne \emptyset,$$

then it offers as an outcome the incompatible knowledge↳ processing.

This needs

- *merging* is a supporting DM combining several pds into a representant.

Merging of Incompletely Compatible PDs: Formalisation

- The pd $\mathcal{A} = \mathcal{F}(\mathsf{f}|\mathcal{K}_{\mathcal{F}^\star})$ with its support being the set $\mathsf{f}^\star$ of acceptable representants is the opted action.
- The behaviour $B \in B^\star$ and an acceptable representant pd $\mathsf{f} \in \mathsf{f}^\star \subset \mathsf{f}^\star_\triangle$ describing it are unknown, i.e. the pair $\mathcal{G}_{\mathcal{F}^\star} = (B, \mathsf{f})$ forms ignorance↳ of the merging↳ DM task.
- The knowledge↳ $\mathcal{K}_{\mathcal{F}^\star}$ is delimited as:

$$\begin{aligned} \mathcal{K}_{\mathcal{F}^\star}:\ & \mathsf{E}_{\mathcal{F}}[\mathsf{D}(\mathsf{f}_\kappa||\mathsf{f})] \le \beta_\kappa < \infty,\ \kappa \in \kappa^\star,\ |\kappa^\star| < \infty, \\ & \mathsf{f}_\kappa(B) \text{ are given pds in } \mathsf{f}^\star_\triangle \\ & \mathcal{F}_0(\mathsf{f}) \text{ is a prior (flat) guess of the action } \mathcal{F} \in \mathcal{F}^\star. \end{aligned} \tag{131}$$

The constraint $\mathsf{E}_{\mathcal{F}}[\mathsf{D}(\mathsf{f}_\kappa||\mathsf{f})] \le \beta_\kappa$ takes f as an acceptable merger with respect to f_κ iff the pd↳ f is a good approximation of the pd f_κ, cf. (116).

- The constants β_κ reflect reliability or importance of κ-th pd. Their systematic choice is out of our scope. A variant is in [145].

Proposition 32 (Merging via Generalised Minimum KLD principle)

The optimal description $^{o}\mathcal{F}(f|\mathcal{K}_{\mathcal{F}^\star})$ *of acceptable mergers* $\mathsf{f} \in \mathsf{f}^\star$ *implied by the generalised minimum KLD principle*↳ *and knowledge* $\mathcal{K}_{\mathcal{F}^\star}$ *(131) is*

$$^{o}\mathcal{F}(\mathsf{f}) \propto \mathcal{F}_0(\mathsf{f}) \prod_{B\in B^\star} \mathsf{f}(B)^{\rho(B)}, \quad \rho(B) \equiv \sum_{\kappa\in\kappa^\star} \lambda_\kappa \mathsf{f}_\kappa(B).$$

$\lambda_\kappa > 0$ *if constraints (131) are active, otherwise* $\lambda_k = 0$.

For the conjugate prior Dirichlet's pd↳ *[77],*

$$\mathcal{F}_0(\mathsf{f}) \propto \prod_{B\in B^\star} \mathsf{f}(B)^{\nu_0(B)-1} \text{ with } \nu_0(B) > 0, \int_{B^\star} \nu_0(B)\,\mathrm{d}B < \infty,$$

$^{o}\mathcal{F}(\mathsf{f})$ *is Dirichlet's with* $\nu(B) = \nu_0(B) + \rho(B) = \nu_0(B) + \sum_{\kappa\in\kappa^\star} \lambda_\kappa \mathsf{f}_\kappa(B)$.
Its expectation – a point representant of pds f_κ, $\kappa \in \kappa^\star$ *– is a convex combination of the merged pds and normalised* $\nu_0(B)$

$$\hat{\mathsf{f}}(B) = \mathsf{E}_{^{o}\mathcal{F}}[\mathsf{f}(B)] = \frac{\nu_0(B) + \sum_{\kappa\in\kappa^\star} \lambda_k \mathsf{f}_\kappa(B)}{\int_{B^\star} \nu_0(B)\,\mathrm{d}B + \sum_{\kappa\in\kappa^\star} \lambda_\kappa}. \qquad (132)$$

Proof of the Merging Form ($B^\star$ is countable!)

- The optimal action ${}^{o}\mathcal{F} \in \mathcal{F}^\star$ (130) minimises the Kuhn-Tucker functional [102], given by multipliers $\lambda_\kappa \geq 0$ (the term $\int_{B^\star} \sum_{\kappa\in\kappa^\star} \lambda_\kappa \mathsf{f}_\kappa(B) \ln(\mathsf{f}_\kappa(B))\,\mathrm{d}B$ independent of $\mathcal{F}$ is drop)

$$
{}^{o}\mathcal{F} \in \mathrm{Arg}\min_{\mathcal{F}\in\mathcal{F}^\star} \int_{\mathsf{f}^\star} \mathcal{F}(\mathsf{f}) \Big[\ln\left(\frac{\mathcal{F}(\mathsf{f})}{\mathcal{F}_0(\mathsf{f})}\right) - \int_{B^\star} \underbrace{\sum_{\kappa\in\kappa^\star} \lambda_\kappa \mathsf{f}_\kappa(B)}_{\rho(B)} \ln(\mathsf{f}(B))\,\mathrm{d}B \Big]\,\mathrm{d}\mathsf{f}
$$

$$
= \mathrm{Arg}\min_{\mathcal{F}\in\mathcal{F}^\star} \int_{\mathsf{f}^\star} \mathcal{F}(\mathsf{f}) \ln\left(\frac{\mathcal{F}(\mathsf{f})}{\mathcal{F}_0(\mathsf{f}) \exp[\int_{B^\star} \rho(B) \ln(\mathsf{f}(B))\,\mathrm{d}B]}\right) \mathrm{d}\mathsf{f}.
$$

The identity $$\int_{B^\star} \rho(B) \ln(\mathsf{f}(B))\,\mathrm{d}B = \prod_{B\in B^\star} \mathsf{f}^{\rho(B)}(B)$$

and the KLD property $\mathsf{D}(\mathcal{F}||\mathcal{G}) = 0 \Leftrightarrow \mathcal{F} = \mathcal{G}$ imply the form of ${}^{o}\mathcal{F}$.

- The form of the posterior pd directly follows from the Bayes' rule ↳.
- The form of the expectation is the property of Dirichlet's pd combined with the normalisation $\sum_{B\in B^\star} \mathsf{f}_\kappa(B) = 1$. □

Implicit Description of Joint Extension and Merging

- Prop. 32 describes the merger of the constructed extensions and gives the point estimate (132) of the merger

$$^{e}\hat{f}(B) = \frac{\nu_0(B) + \sum_{\kappa\in\kappa^\star} \lambda_\kappa \, ^{e}f_\kappa(B)}{\int_{B^\star} \nu_0(B)\,\mathrm{d}B + \sum_{\kappa\in\kappa^\star} \lambda_\kappa}. \tag{133}$$

- The extensions $^{e}f_\kappa$, $\kappa \in \kappa^\star$, and thus the optimal merger $^{e}\hat{f}(B)$ strongly depend on the prior guess $^{e}f_0(B)$. It is natural to choose

$$^{e}f_0 = \, ^{e}\hat{f}(B). \tag{134}$$

  This choice gives the implicit formula for the optimal merger $^{e}\hat{f}$

$$^{e}\hat{f}(B) = \frac{\nu_0(B) + \sum_{\kappa\in\kappa^\star} \lambda_k \, ^{e}\hat{f}(U_\kappa|\mathbf{G}_\kappa, \mathbf{K}_\kappa) f_\kappa(\mathbf{G}_\kappa|\mathbf{K}_\kappa) \, ^{e}\hat{f}(\mathbf{K}_\kappa)}{\int_{B^\star} \nu_0(B)\,\mathrm{d}B + \sum_{\kappa\in\kappa^\star} \lambda_\kappa}, \tag{135}$$

- The implicit equation (135) is conjectured to have, not necessarily unique, solution with unique fragmental pds $^{e}\hat{f}(\mathbf{G}_\kappa|\mathbf{K}_\kappa)$, $\kappa \in \kappa^\star$.
- The algorithmic solution of (135) and its subtasks is in its infancy.

Remark 24 (On Merging Use & Inspection)

- *The merging↳ of, generally incompatible, pds, gained by the extension of partially described aims, serves directly to construction of ideal pd, to preference elicitation!*
- *The merging↳ is also an efficient tool for solving extremely hard de-centralised DM↳ [16] via knowledge & preference sharing [80].*
- *Works [76, 84] illustrate that the outlined knowledge and preference treatment indeed algorithmically supports imperfect decision maker↳s.*
- *The proposed merging↳ exploits the generalised minimum KLD principle↳. It benefits from the possibility to specify the set* $\mathsf{f}^\star$ *(131), being the support of* $\mathcal{F}$, *by nonlinear constraints on the approximated* f.
- *The derived merging↳, see (132), justifies and extends the heuristically motivated arithmetic pooling [14, 48]. A related derivation called supra-Bayesian merging is in [145].*

Exercises 16 on Minimum KLD Principle

Exercises 16 (Minimum KLD Principle vs. Practice)

1. *Show that the minimum KLD principle leads to uniform pd if* $\mathsf{f}^\star = \{\mathsf{f} : \mathrm{supp}\,[\,f\,] = \textit{a given finite interval}\,\}$.
2. *Show that the minimum KLD principle leads to exponential family if* $\mathsf{f}^\star = \{\mathsf{f} : \ \mathsf{E}_\mathsf{f}[\phi|\bar{\phi}] = \bar{\phi}\}$.
3. *Think about possible forms and contents of knowledge-describing pieces you are meeting in your specific DM problems.*
4. *Do 3 for aim-describing pieces.*
5. *Any idea how to choose constants* $(\beta_\kappa)_{\kappa\in\kappa^\star}$ *in (131)?*
6. *Broadcasting and merging knowledge/aim describing fragmental pds create powerful tool for human-like cooperation. Do you know a "universal" way to broadcast them?*

Specialised DM Tasks for Constructing DM Elements

Presentation Aims

The following DM tasks are used in constructing DM elements ↳:

- Testing of hypotheses, Sec. 25.
- Structure estimation, Sec. 26.
- Sequential decision making, Sec. 27.

Moreover, these DM tasks

- are of an independent, very practical, interest
- serve as examples of the power of the general DM theory
- serve as exercises.

Example 18 (Testing of Hypotheses as DM task)

aim	*to estimate the pointer* $\vartheta \in \vartheta^\star \equiv \{1, \ldots, \lvert\vartheta^\star\rvert\}$, $\lvert\vartheta^\star\rvert < \infty$ *to the hypothesis about system model*
system	*a modelled World part*
action	$\hat{\vartheta} \in \vartheta_\star \subset \vartheta^\star$
knowledge	$\mathbf{K}$ *entering the system model and prior pd* $\mathsf{f}(\vartheta\lvert\mathbf{K})$
ignorance	*unknown pointer* $\vartheta \in \vartheta^\star$ *and observation* $\Delta \in \Delta^\star$
uncertainty	*anything preventing to determine fully* ϑ *from* $\mathbf{K}$
constraint	$\vartheta_\star$, *a feasible subset of* $\vartheta^\star$ *with usually extreme* $\lvert\vartheta^\star\rvert$

$$\mathsf{f}(\Delta|\vartheta, \mathbf{K}) \equiv \mathsf{f}(\Delta|\vartheta, \hat{\vartheta}, \mathbf{K}) \qquad (136)$$

This simple DM tasks serves for selecting significant entries of behaviour like observations & actions as well as of the choice system-model structure.

Hypotheses Testing Reduces to Estimation of an Unknown ϑ

DM elements relevant to the point estimation are

- behaviour $B = [\mathbf{G}_A, A, \mathbf{K}_A]$=[ignorance, action, knowledge]= [(future observation, pointer to the most plausible hypothesis) pointer estimate, knowledge] $= [(\Delta, \vartheta), \hat{\vartheta}, \mathbf{K}]$
- decision rule meets natural conditions of DM $\mathsf{S}(\hat{\vartheta}|\vartheta, \mathbf{K}) = \mathsf{S}(\hat{\vartheta}|\mathbf{K})$
- system model $\mathsf{f}(\Delta|\vartheta, \hat{\vartheta}, \mathbf{K}) = \mathsf{f}(\Delta|\vartheta, \mathbf{K})$ is identified with the ideal pd ${}^{\mathsf{I}}\mathsf{f}(\Delta|\vartheta, \hat{\vartheta}, \mathbf{K})$, leave to the fate option is applied
- prior pd $\mathsf{f}(\vartheta|\mathbf{K})\mathsf{f}(\mathbf{K}) = {}^{\mathsf{I}}\mathsf{f}(\vartheta|\mathbf{K})\,{}^{\mathsf{I}}\mathsf{f}(\mathbf{K})$, leave to the fate option applies
- ideal decision rule ${}^{\mathsf{I}}\mathsf{S}(\hat{\vartheta}|\vartheta, \mathbf{K})$ is a positive $(|\vartheta_\star|, |\vartheta^\star|)$ table with maximum values for $\hat{\vartheta} = \vartheta$.

Solution $${}^{\mathsf{o}}\mathsf{f}(\hat{\vartheta}|\mathbf{K}) \propto \exp\left[\sum_{\vartheta \in \vartheta^\star} \mathsf{f}(\vartheta|\mathbf{K}) \ln\left({}^{\mathsf{I}}\mathsf{S}(\hat{\vartheta}|\vartheta, \mathbf{K})\right)\right].$$

Hypotheses testing reduces to estimation, i.e. evaluation of the pd $\mathsf{f}(\vartheta|\mathbf{K})$ of the discrete-valued parameter $\vartheta \in \vartheta^\star$. Thus, we focus on its specificity.

Remarks on Hypotheses Testing

- Unlike in classical hypotheses testing [136], the testing is performed within a completely specified set of alternatives.
- The hypotheses testing is usually performed with the knowledge (gradually) extended by data D^t. Bayesian estimation, Prop. 17, provides the key posterior pd $\mathsf{f}(\vartheta|\mathbf{K}_t)$.
- Discrete nature of ϑ implies that this pd quickly concentrates on a small subset of $\vartheta_\star$ containing often single point, see Prop. 21.

  Thus, for any reasonable ideal $^{\mathsf{I}}\mathsf{S}(\hat{\vartheta}|\vartheta, \mathbf{K})$, one hypothesis within $\vartheta_\star$ is accepted as the most plausible one even when none of them is correct.
- The needed pds $\{\mathsf{f}(\Delta|\vartheta, \mathbf{K})\}_{\vartheta\in\vartheta^\star}$ are rarely given directly. They are mostly predictive pds obtained through filtering or parameter estimation, Props. 16, 21.

  $\mathsf{f}(\vartheta|\mathbf{K}_t)$ is strongly influenced of prior pds $\mathsf{f}(\Theta_\vartheta|\vartheta, \mathcal{K})$ of the hypothesis-specific parameters $\Theta_\vartheta \in \Theta^\star_\vartheta$.

- Other actions are usually present and mostly meet natural conditions of DM (65). Then, the decision rules generating them cancel in the formula for $f(\vartheta|\mathbf{K}_t)$, see Prop. 17.
- The testing of hypotheses is extremely powerful technique in spite of its formal simplicity. It is especially true when dealing with the predictive pds evaluated for each hypothesis by filtering or estimation.
- Non-Bayesian treatment of such compound hypotheses, given by $(\vartheta \in \vartheta^{\star}, \Theta_{\vartheta} \in \Theta^{\star}_{\vartheta})$, [136], is nontrivial. The Bayesian solution brought a whole set of novel and efficient algorithms for:
- *structure estimation*, which selects the best model among parametric models differing in functional form, order of regression vector or significant variables to be used in the system model, [11, 68, 90].

Structure Estimation

Proposition 33 (Structure Estimation)

Let:

- $\{\mathsf{f}(\Delta_t|\Theta_\vartheta, A_t, \mathbf{K}_{t-1}, \vartheta)\}_{\vartheta\in\vartheta_\star, t\in t^\star}$ *be candidates for parametric model*
- *the respective unknown parameter* Θ_ϑ *be described by a prior pd* $\mathsf{f}(\Theta_\vartheta|\vartheta, \mathbf{K}_0)$ & $\mathsf{f}(\vartheta|\mathbf{K}_0)$ *be prior probabilities of hypotheses* H_ϑ, $\vartheta \in \vartheta_\star$
- *any action* A_t *in parametric model meet natural conditions of DM.*

Then, the posterior pd on hypotheses, needed for hypotheses testing, is

$$\mathsf{f}(\vartheta|\mathbf{K}_t) \propto \frac{\mathsf{J}_\vartheta(\mathbf{K}_t)}{\mathsf{J}_\vartheta(\mathbf{K}_{t-1})}\mathsf{f}(\vartheta|\mathbf{K}_{t-1}), \ t \in t^\star \tag{137}$$

$$\mathsf{J}_\vartheta(\mathbf{K}_t) = \int_{\Theta_\vartheta^\star} \prod_{\tau=1}^{t} \mathsf{f}(\Delta_\tau|\Theta_\vartheta, A_\tau, \mathbf{K}_{\tau-1})\mathsf{f}(\Theta_\vartheta|\mathbf{K}_0)\,\mathrm{d}\Theta_\vartheta.$$

Proof It uses basic algebra with pds and Prop. 17, respecting natural conditions of DM adopted for all involved decisions. □

Search within a Huge Set of Hypotheses

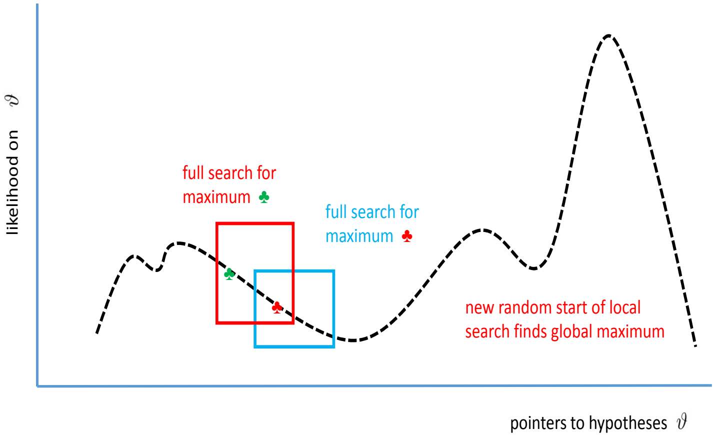


Figure : Full local search starts within blue neighbourhood. Red neighbourhood is defined around the found maximiser,. . . up to no improvement.

Concluding Remarks on Structure Estimation

- Structure estimation uses predictive pd ↳ offered also by filtering ↳.
- Mechanical ways of generating list of hypotheses make $|\vartheta^\star|$ extremely large and consequently the their testing infeasible.
- Hypotheses are usually created gradually and the knowledge may depend on it as well as on the considered model structure.
- The open question arises of how to extend the existing set of hypotheses and at the same time to exploit former data so that the new hypothesis is compared in a fair way.

  Partial attempts have been done but a systematic methodology is missing. It would need DM theory with a varying set $B^\star$ of possible behaviours. Quantum probability operating on projectors to subspaces of lower dimensions seems to be promising even in this respect.

Sequential estimation serves as a bridge to controlling deliberation effort spent to design of DM strategies. It is important on its own.

- *sequential estimation* decides whether to process a new observations or whether to stop and evaluate the final parameter estimate.
- Sequential estimation balances costs connected with acquiring observation with costs induced by imprecisions of the final estimate.
- Unlike majority decision tasks with a fixed horizon $h < \infty$, the sequential estimation deals with a potentially infinite h.
- Hypotheses testing is a specific estimation also treatable sequentially.
- Sequential estimation was at roots of the theory of statistical decision functions, we build on [164].

In the next text, the standard Bayesian sequential estimation is formalised and solved. Prop. 35 describes sequential DM with FPD.

Bayesian Formalisation of Sequential Estimation

Sequential point estimation casts in the standard Bayesian DM as follows.

- Behaviour $B \equiv [\mathbf{G}_t \equiv (\Theta, \Delta_t), A_t \equiv (\hat{\Theta}_t, \Phi_t), \mathbf{K}_{t-1}]$

  =[(unknown parameter,observation↳s),(estimate,stopping flag),data].

- Admissible compared strategies↳ consist of rules

$$\mathsf{S}_t : \mathbf{K}^\star_{t-1} \to (\hat{\Theta}_t \in \hat{\Theta}^\star, \Phi_t \in \Phi^\star) \quad \hat{\Theta}^\star \subset \Theta^\star$$

$\Phi^\star \equiv \{$stop measuring & estimate Θ, make a new observation↳$\} \equiv \{0, 1\}$.

- Loss

$$\mathsf{Z} = \begin{cases} \sum_{\tau \le t} c(\mathbf{K}_{\tau-1}) + z(\Theta, \hat{\Theta}_t, \mathbf{K}_{t-1}) & \text{if} \quad \Phi_t = 0 \ \& \ \Phi_\tau = 1, \ \forall \tau < t \\ \sum_{\tau \le t} c(\mathbf{K}_{\tau-1}) & \text{if} \quad \Phi_\tau = 1 \ \forall \tau \le t \end{cases} \tag{138}$$

  where the partial loss $z(\Theta, \hat{\Theta}, \mathbf{K}_{t-1})$ measures a distance of Θ and its estimate $\hat{\Theta}$. $c(\mathbf{K}_{\tau-1}) > 0$ is a price of τ-th observation.

Optimal Bayesian Sequential Estimation

Proposition 34 (Sequential Estimation)

Let us consider the sequential estimation and assume that there is an admissible strategy for which the expected loss is finite.

Then the next inequalities express the sufficient condition for an index t to be the time moment at which observation should be stopped

$$\mathsf{E}\left[\left(z(\Theta, \hat{\Theta}_t, \mathbf{K}_{t-1}) - z(\Theta, \hat{\Theta}_{t+k}, \mathbf{K}_{t+k-1}) - \sum_{\tau>t}^{t+k} c(\mathbf{K}_{\tau-1})\right) \middle| \mathbf{K}_{t-1}\right] \leq 0$$

$$\forall k = 1, 2, \ldots, \qquad (139)$$

where $\hat{\Theta}_{t+k},\ k = 0, 1, 2, \ldots$ *is parameter estimate* $\hat{\Theta}$ *using* $\mathbf{K}_{t+k-1}$ *and minimising* $\mathsf{E}[z(\Theta, \hat{\Theta}, \mathbf{K}_{t+k-1})]$.

Proof

Let (139) be fulfilled. Then, combining the form of the loss (138), the fact that the optimal stopping time has to be determined using its knowledge and finiteness of the loss for the optimal solution we get, $\forall k = 1, 2, \ldots,$

$$
\begin{aligned}
& \mathsf{E}\left[\left(z(\Theta, \hat{\Theta}_t, \mathbf{K}_{t-1}) + \sum_{\tau=1}^{t} c(\mathbf{K}_{\tau-1})\right)\middle|\mathbf{K}_{t-1}\right] \\
\leq \quad & \mathsf{E}\left[\left(z(\Theta, \hat{\Theta}_{t+k}, \mathbf{K}_{t+k-1}) + \sum_{\tau=1}^{t+k} c(\mathbf{K}_{\tau-1})\right)\middle|\mathbf{K}_{t-1}\right].
\end{aligned}
$$

Using isotonicity of the expectation (taken over $\mathbf{K}_{t-1}$), we find that the chosen decision cannot be improved by any estimate that uses more observations than the inspected one. □

On Sequential Estimation

Remark 25 (Use of Sequential Estimation)

- *The implementation of the optimal decision rule*↳ *with stopping requires the posterior pd*↳ *given in Prop. 17.*
- *The ability to evaluate* $\mathsf{E}\left[z(\Theta, \hat{\Theta}_{t+k}, \mathbf{K}_{t+k-1}) | \mathbf{K}_{t-1}\right]$, $k = 1, 2, \ldots,$ *is decisive for a practical applicability.*
- *Stopping rules used for speeding up simulations [141] based on a sequential estimation illustrates their, still underestimated, usefulness.*
- *The dependence of the observation price on available knowledge was exploited when the sequential estimation runs as an inner optimisation loop. The closer it is to the optimum the lower this price is, [82].*

  The idea offers for stopping Monte Carlo integration in filtering↳.

Sequential DM and Optimisation of Deliberation Effort

Problem

- A decision maker devotes a limited deliberation effort to any particular DM. It causes no problem if its hard bound is fixed. If an additional effort can be expended to get a higher DM quality,

  $\boxed{\text{it is necessary to decide whether it is worthwhile to exert this effort}}$

- [144, 150] and many others claim that any attempt to include this balancing into the optimal design leads to an infinite regress:

  An inclusion of deliberation penalty into the DM loss increases the deliberation effort, which calls for an additional penalisation etc.

- Sequential estimation, [164], indicates that this is generally untrue.

Foreseen Solution

- The sequential DM solved via FPD proposes deliberation aware optimisation, which is expected to avoid the infinite regress.

Formalisation of the Sequential FPD

- Sequential DM deals with the behaviour $B = \left((\Delta_t, A_t, \Phi_t)_{t\geq 1}, \mathbf{K}_0\right)$ containing the additional action entry

  $\Phi_t \in \Phi^\star \equiv \{1, 0\} \equiv \{\text{continue improvements, stop improvements}\}$.
- The quality of the decision strategy is evaluated by the ideal pd. It is the desired closed-loop model when no stopping applies

$$^{\mathsf{I}}\mathsf{f}(\Delta_t|A_t, \Phi_t = 1, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}).$$

- The leave to the fate choice for stopping, $\Phi_t = 0$, gives the ideal

$$\begin{aligned} & ^{\mathsf{I}}\mathsf{f}(\Delta_t|A_t, \Phi_t, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(A_t|\Phi_t, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(\Phi_t|\mathbf{K}_{t-1}) \\ \equiv\ & \left[^{\mathsf{I}}\mathsf{f}(\Delta_t|A_t, \Phi_t = 1, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})\, ^{\mathsf{I}}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\right]^{\Phi_t} \\ \times\ & \left[\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})\mathsf{f}(A_t|\mathbf{K}_{t-1})(1 - {}^{\mathsf{I}}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}))\right]^{1-\Phi_t} \end{aligned} \quad (140)$$

The probability $^{\mathsf{I}}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}) \in (0, 1)$ quantifies the readiness for continuation, i.e. readiness to exert the deliberation cost.

Solution of the Sequential FPD

Proposition 35 (FPD with Stopping: Penalised Deliberation Effort)

With the ideal pd (140), the optimal decision rules ${}^{o}\mathsf{f}$ *have the form*

$$\begin{aligned}
{}^{o}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1}) &= \frac{{}^{I}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1}) \exp[-\omega(A_t, \mathbf{K}_{t-1})]}{\rho(\mathbf{K}_{t-1})} \\
{}^{o}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}) &= {}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\rho(\mathbf{K}_{t-1})/\mathrm{e} \qquad (141) \\
\gamma(\mathbf{K}_{t-1}) &= \exp[-\,{}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\rho(\mathbf{K}_{t-1})/\mathrm{e}]
\end{aligned}$$

$$\omega(A_t, \mathbf{K}_{t-1}) = \int_{\Delta^\star} \mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \ln\left(\frac{\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\,{}^{I}\mathsf{f}(\Delta_t|A_t, \Phi_t = 1, \mathbf{K}_{t-1})}\right) \mathrm{d}\Delta_t.$$

The evaluations (141) run backward. The value function $-\ln(\gamma(\mathbf{K}_t))$ *is zero at the strict upper bound on the decision horizon* $t = T$.

Only the values $\Phi_\tau = 1$, $\tau < t$, *enter the knowledge* $\mathbf{K}_{t-1}$.

Beginning of the Proof of the FPD with Stopping

At the last time $t = T$, where T is a strict upper bound on horizon, the part of KLD influenced by the last optimised decision rule reads

$$
\begin{aligned}
\mathsf{J}_t \;=\;& \mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}) \times \Bigg\{ \ln\left(\frac{\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})}{{}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})}\right) \\
&+ \int_{A^\star} \mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1}) \times \Bigg[\ln\left(\frac{\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})}{{}^{I}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})}\right) \\
&\underbrace{+\int_{\Delta^\star} \mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \ln\left(\frac{\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})}{{}^{I}\mathsf{f}(\Delta_t|A_t, \Phi_t = 1, \mathbf{K}_{t-1})\gamma(\Delta_t, A_t, \Phi_t = 1, \mathbf{K}_{t-1})}\right) \mathrm{d}\Delta_t}_{\omega(A_t, \mathbf{K}_{t-1})} \\
&\Bigg] \mathrm{d}A_t \Bigg\}
\end{aligned}
$$

with $\gamma(\mathbf{K}_t) = \gamma(\Delta_t, A_t, \Phi_t = 1, \mathbf{K}_{t-1})$ depending on $(\Phi_\tau = 1)_{\tau \le t}$.
This holds as $\gamma(\mathbf{K}_t) = 1$ for the considered $t = T$.

A rearrangement of $\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})$-dependent part & minimisation of KLD by identical arguments gives the optimal factor of the DM rule

End of the Proof of the FPD with Stopping

$$
{}^{o}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1}) = \frac{{}^{I}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})\exp[-\omega(A_t, \mathbf{K}_{t-1})]}{\rho(\mathbf{K}_{t-1})}
$$

$$
\rho(\mathbf{K}_{t-1}) = \int_{A^\star} {}^{I}\mathsf{f}(A_t|\Phi_t = 1, \mathbf{K}_{t-1})\exp[-\omega(A_t, \mathbf{K}_{t-1})]\mathrm{d}A_t \in (0,1).
$$

With it, it remains to minimise over $\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}) \in (0,1)$, the function

$$
\min_{\{\mathsf{f}(A_t|\Phi_t=1,\mathbf{K}_{t-1})\}} \mathsf{J}_t = \mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\ln\left(\frac{\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})}{{}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\rho(\mathbf{K}_{t-1})}\right).
$$

Its minimiser and the reached minimum are respectively

$$
\begin{aligned}
{}^{o}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1}) &= {}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\rho(\mathbf{K}_{t-1})/\mathrm{e} \\
\min_{\{\mathsf{f}(A_t,\Phi_t|\mathbf{K}_{t-1})\}} \mathsf{J}_t &= -{}^{I}\mathsf{f}(\Phi_t = 1|\mathbf{K}_{t-1})\rho(\mathbf{K}_{t-1})/\mathrm{e} \equiv \ln(\gamma(\mathbf{K}_{t-1})).
\end{aligned}
$$

The last equality defines $\gamma(\mathbf{K}_{t-1}) \le 1$, which depends on the part of $\mathbf{K}_{t-1}$ containing only $\Phi_\tau = 1$. The situation repeats for decreased t. $\square$

Remark 26 (Infinite Regress of Optimised Deliberation Effort)

- *For the finite strict bound on horizon* $T < \infty$*, Prop. 35, is expected to decrease the number of optimisation steps.*
- *The asymptotic analysis, Prop. 14, indicates that a stationary version of the discussed solution exists.*
- *Both above conjectures have to be (dis)proved via a detailed analysis.*

Approximation and Adaptivity in DM

A Quote on Approximation

Although this may seem a paradox, all exact science is dominated by the idea of approximation.

When a man tells you that he knows the exact truth about anything, you are safe in inferring that he is an inexact man.

Bertrand Russel: The Scientific Outlook, 1931

What the Presented DM Theory Lacks? Way out?

The presented methodology provides a complete formal solution of the optimal DM↳ under uncertainty.

- The solution leading to practically optimal design↳ is missing.

  The design complexity is not under a systematic control.

This part:

- recalls the inherent source of DM complexity
- claims that distributed DM↳ generally handles DM complexity
- outlines a general version of distributed DM↳ coping with the design complexity
- indicates that efficient decision maker↳s form adaptive system↳s.

Parts 8 and 9 provide FPD as the general theory of DM uncertainty.

- The solution operates on the closed-loop model, the pd (75)

$$\mathsf{f}_{\mathsf{S}}(B) = \underbrace{\mathsf{f}(X_0) \prod_{t \in t^\star} \mathsf{f}(\Delta_t, X_t | X_{t-1}, A_t, \mathbf{K}_{t-1})}_{\text{system model } \mathsf{M}(B)} \underbrace{\prod_{t \in t^\star} \mathsf{f}(A_t | \mathbf{K}_{t-1})}_{\text{strategy } \mathsf{S}(B)}.$$

  It describes realisations of the behaviour $B = (X_0^h, \mathbf{K}_h)$ that includes quantities considered by the decision maker up to the horizon h.

- The ideal counterpart ${}^{\mathsf{I}}\mathsf{f}_{\mathsf{S}} = {}^{\mathsf{I}}\mathsf{M}\,{}^{\mathsf{I}}\mathsf{S}$ of f_{S} describes DM preferences.

- Thus, DM deals with a pair of scalar functions acting on the domain $B^\star$ of an extremely high dimension.

  The general impossibility to store exactly and handle functions of many variables makes the optimal DM design exceptional.

Distributed DM as General Complexity Countermeasure

- *distributed DM* splits an unmanageable DM task in DM subproblems dealing with its subparts. Typical splitting is made with respect to
- quantities included into the behaviour ↳ up to the used time horizon ↳,
- domains of quantities in hierarchies (aggregation at higher static levels)
- models and preferences considered (feasible ones and their extensions)
- subproblems faced (learning, design, structure estimation, . . .)

The key induced problems are

- How to make splitting? We mostly rely on "naturally"-evolved splitting and also apply specific solutions for modifying specific tasks.
- How to "glue" parts?. We support decision maker ↳ solving a manageable DM and equip it with FPD-based tools for exploiting knowledge about neighbouring decision makers.

Unexpected even undesirable emergent behaviours are costs for the achieved feasibility. Their inspection and controlling are hard and out of our scope.

Orientation on Adapting Selfish Imperfect Decision Maker

The adopted solution concept relies on distributed DM ↳ with:

- *selfish decision maker* judges DM results according to its preferences.

  This is ethically neutral as selfish preferences may include "happiness" of neighbours.

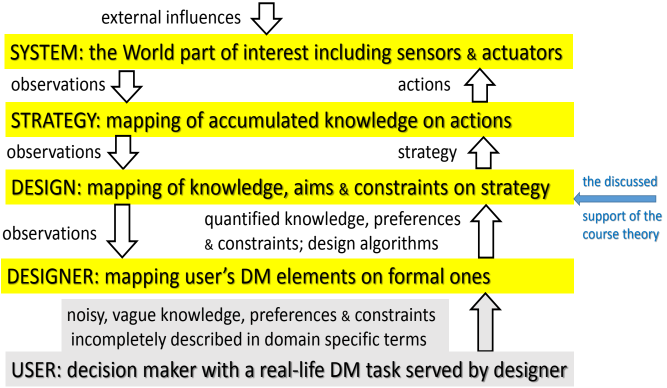
An imperfect decision maker ↳ is considered, which uses FPD ↳ as basic theory and is algorithmically supported in:

- knowledge and preference elicitation
- feasible learning and strategy design
- sharing of knowledge and preferences with neighbouring decision makers of the same type.

The decision maker ↳ supported in this way can become adaptive system ↳.

DM Support & Cooperation Concept

- An in alone

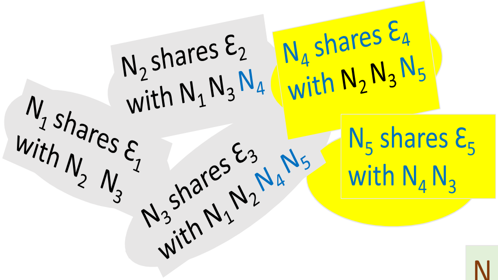


- Knowledge and preference processing merges, see Sec. 24, a partial probabilistic information about DM elements↳ of neighbouring decision maker↳s (for merging). This allows cooperative distributed DM↳.

Cooperation Runs by Sharing DM elements

Cooperation as Sharing Parts of DM Elements among Neighbours

$\mathcal{E}_{\kappa} \equiv [(f_{S\,\kappa}, f_{I\,\kappa}) =$ (closed-loop,ideal) pds acting on behaviour $\equiv B_{\kappa}]$



- communication topology varies
- sharing is asynchronous
- # of jointly reachable neighbours small

N_{κ} shares $\mathcal{E}_{\kappa}$ with reachable neighbours N_{γ} N_{φ} N_{α} N_{λ}

I live a single realisation of my life.

An anonymous author, during his/her life without publishing it.

Decision Maker as Adaptive System

Applications require the strategy be good at the behaviour ↳ realisation ↳.

- Thus, it suffices to know a good approximation of the optimal strategy locally around the actual behaviour realisation.

  It can often be found and converts the decision maker ↳ into:

- *adaptive system* is decision maker ↳ adopting strategy, which is an approximation of the optimal strategy around the behaviour realisation ↳.

  This understanding of adaptive system is in [71]. Its operational description of adaptive systems is, e.g., in [7].

- Even DM of a single decision maker ↳ and its support are demanding. This makes us to adopt practically optimal design ↳, which treats the decision maker as adaptive system ↳.

- Adaptive systems are vital in our concept, but they are useful per se and can be embedded into distributed, say hierarchical, systems.

Feasible and Approximate Learning

On this Part

- The presented theory covers an extreme width of DM tasks.
  Just a few of them are solvable exactly and to some there are well established approximate solution techniques. Both are outlined here.
- The text does not provide a full picture due to the limited
  - coverage by the contemporary research,
  - knowledge of lecturer,
  - presentation time available.
- The presentation focuses predominantly on learning based on parameter estimation with time-invariant hidden quantities.

Factorisation Makes Scalar Parametric Model Sufficient

- The chain rule allows us to decompose any parametric model

$$\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) = \prod_{i=1}^{\ell_{\Delta^\star}} \mathsf{f}(\Delta_{t;i}|\Theta, A_t, \Delta_{t;i+1}, \ldots, \Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1}). \quad (142)$$

- *factor* is the pd modelling a single observation entry $\Delta_{t;i}$ in (142).
- The factor is the basic object we deal with further on as its use
  - is simpler than a model predicting multivariate observations,
  - allows a fine modelling of individual observations,
  - deals with that part of Θ, which enters the specific factor,
  - serves well for modelling of
- *mixed observations* contain both continuous & discrete valued entries.

Simplified Notation of Enriched Factor Knowledge

- In the factorised modelling

$$\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) = \prod_{i=1}^{\ell_{\Delta^\star}} \mathsf{f}(\Delta_{t;i}|\Theta, A_t, \Delta_{t;i+1}, \ldots, \Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1})$$

knowledge $\mathbf{K}_{t-1}$ for predicting the i-th observation entry is *enriched* to $\mathbf{K}_{t-1;i} = \Delta_{t;i+1}, \ldots, \Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1}$.

- The index i & the enrichment $\Delta_{t;i+1}, \ldots, \Delta_{t;\ell_{\Delta^\star}}$ are mostly drop here.
- Thus, we deal with the parametric model $\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$ modelling a scalar observation Δ_t.
- The parametric model, factor, is mostly taken from the dynamic exponential family and mixtures of such models.

Dynamic Exponential Family (EF): Definition

- *exponential family* (dynamic EF) consists of the parametric models

$$f(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) = m(\Psi_t, \Theta) = A(\Theta) \exp \langle B(\Psi_t), C(\Theta) \rangle \text{ with } \quad (143)$$

- *data vector* $\Psi_t' \equiv [\Delta_t, \psi_t']$ with $\ell_{\Psi^\star} < \infty$; $'$ is transposition
- *regression vector* ψ_t, $\ell_{\psi^\star} = \ell_{\Psi^\star} - 1 < \infty$, whose values evolve recursively in a known way

$$(\Psi_{t-1}^\star, D_t^\star) = (\Psi_{t-1}^\star, \Delta_t^\star, A_t^\star) \to \Psi_t^\star \quad (144)$$

- $\langle \cdot, \cdot \rangle$ is the scalar-valued function, linear in the first argument.
- $A(\cdot)$ is a nonnegative scalar function defined on $\Theta^\star$
- $B(\cdot)$, $C(\cdot)$ defined respectively on $\Psi_t^\star$ and $\Theta^\star$ are array functions of $\langle \cdot, \cdot \rangle$ - compatible, finite and fixed dimensions.

Remarks on Definition of Dynamic EF

Remark 27 (On Definition of Dynamic EF)

- *A scalar product is often used as the mapping $\langle \cdot, \cdot \rangle$, typically,*

$$\langle X, Y \rangle = \begin{cases} X'Y & \text{if } X,\ Y \text{ are vectors} \\ \mathrm{tr}[\mathbb{X}'\mathbb{Y}] & \text{if } \mathbb{X},\ \mathbb{Y} \text{ are matrices, } \mathrm{tr} \text{ is trace} \\ \sum_{\iota \in \iota^\star} X_\iota Y_\iota & \text{if } X,\ Y \text{ are arrays with a multi-index } \iota \end{cases} \tag{145}$$

- *The definition of the exponential family$_\llcorner$ requires, unlike textbooks, the recursive updating of the data vector$_\llcorner$ $\Psi_{t-1}, D_t \to \Psi_t$.*

  The recursion is the practically important condition for dynamic DM.

- *Notice that equality is used in (143)*

$$\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) = \mathsf{m}(\Psi_t, \Theta) = \mathsf{A}(\Theta) \exp \langle \mathsf{B}(\Psi_t), \mathsf{C}(\Theta) \rangle \,.$$

  Normalisation of this pd must not spoil the considered exponential form. This makes the dynamic exponential family$_\llcorner$ rather narrow.

Remark 28 (Narrowness of Dynamic EF)

- *Let us try a parametric model gained via a linear expansion of* $\ln f(\Delta|\psi,\Theta) \approx \langle B(\Psi), C(\Theta)\rangle$.
  The function $B(\Psi)$ *contains basis functions of data vectors* $\Psi = [\Delta, \psi']'$. $C(\Theta)$ *collects basis weights. Pd normalisation gives*

$$f(\Delta|\psi,\Theta) = \tilde{A}(\psi,\Theta) \exp\langle B(\Psi), C(\Theta)\rangle, \quad \tilde{A}^{-1} \equiv \int_{\Delta^\star} \exp\langle B(\Psi), C(\Theta)\rangle \, d\Delta,$$

which is in EF iff $\tilde{A}(\psi,\Theta) = A(\Theta)$. *For a non-void* ψ, *this (only?) meets:*

- *normal model* $f(\Delta|\psi, \Theta \equiv (\theta, r)) = N_\Delta(\theta'\psi_t, r)$ *linear in regression coefficients* θ *and with constant variance* r
- *Markov model acting on data vectors* $\Psi \in \Psi^\star$ *with* $|\Psi^\star| < \infty$, *and* $f(\Delta|\psi,\Theta) = \Theta_{\Delta|\psi} \geq 0$, $\sum_{\Delta\in\Delta^\star} \Theta_{\Delta|\psi} = 1$, $\forall \psi \in \psi^\star$.

Textbooks mostly Deal with Static EF

Some members of static exponential family with void regression vector are

Name	**Parametric model**	**Observation Δ**	**Parameter**
Exponential	$\frac{1}{\lambda}\exp\left(-\frac{\Delta}{\lambda}\right)$	$\in(0,\infty)$	$\lambda>0$
Poisson	$\frac{\mu^{\Delta}}{\Gamma(\Delta+1)}\exp(-\mu\Delta)$	$\in\{0,1,\ldots\}$	$\mu>0$
Multinomial	$\prod_{i\in\Delta^\star}\Theta_i^{\delta(i,\Delta)}$	$\in\{1,\ldots,\lvert\Delta^\star\rvert\}$	$\{\Theta_\Delta\geq 0$ $\sum_{\Delta\in\Delta^\star}\Theta_\Delta=1\}$
Normal	$\frac{1}{\sqrt{2\pi r}}\exp\left[-\frac{(\Delta-\mu)^2}{2r}\right]$	$\in[-\infty,\infty]$	$\mu\in[-\infty,\infty],\ r>0$
Log-Normal	$\frac{1}{\Delta\sqrt{2\pi r}}\exp\left[-\frac{\ln^2\left(\frac{\Delta}{\mu}\right)}{2r}\right]$	$\in(0,\infty)$	$\mu>0,\ r>0$

- *Kronecker delta*, discrete Dirac delta,
$$\delta(i,\Delta)=\begin{cases}1 & \text{if } i=\Delta\\ 0 & \text{if } i\neq\Delta\end{cases} \tag{146}$$

- *gamma function*
$$\Gamma(x)\equiv\int_0^\infty z^{x-1}\exp(-z)\,\mathrm{d}x<\infty \text{ for } x>0. \tag{147}$$

Proposition 36 (Estimation and Prediction in Exponential Family)

Let natural conditions of DM (65) hold and let the parametric model belong to EF (143). Then, the predictive pd has the form

$$\mathrm{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) = \frac{\mathrm{J}(V_{t-1} + \mathrm{B}(\Psi_t), \nu_{t-1} + 1)}{\mathrm{J}(V_{t-1}, \nu_{t-1})} \quad (148)$$

$$V_t = V_{t-1} + \mathrm{B}(\Psi_t), \quad V_0 = 0 \quad (149)$$

$$\nu_t = \nu_{t-1} + 1, \quad \nu_0 = 0$$

$$\mathrm{J}(V, \nu) = \int_{\Theta^\star} \mathrm{A}^{\nu}(\Theta) \exp \langle V, \mathrm{C}(\Theta) \rangle \, \mathrm{f}(\Theta) \, \mathrm{d}\Theta, \quad (150)$$

where $\mathrm{f}(\Theta)$ *is a prior pd. The posterior pd is*

$$\mathrm{f}(\Theta|\mathbf{K}_t) = \frac{\mathrm{A}^{\nu_t}(\Theta) \exp \langle V_t, \mathrm{C}(\Theta) \rangle \, \mathrm{f}(\Theta)}{\mathrm{J}(V_t, \nu_t)} \quad \textit{with} \quad (151)$$

$$\textit{likelihood} \quad \mathrm{L}(\Theta, \mathbf{K}_t) \equiv \mathrm{L}(\Theta, V_t, \nu_t) = \mathrm{A}^{\nu_t}(\Theta) \exp \langle V_t, \mathrm{C}(\Theta) \rangle . \quad (152)$$

Proof of Estimation & Prediction Form in EF

Prop. 17 on Bayesian estimation implies

$$\mathrm{f}(\Theta|\mathbf{K}_t) = \frac{\prod_{\tau=1}^t \mathrm{f}(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1})\mathrm{f}(\Theta)}{\int_{\Theta^\star} \prod_{\tau=1}^t \mathrm{f}(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1})\mathrm{f}(\Theta)\,\mathrm{d}\Theta}$$

$$\overbrace{=}^{(143)} \quad \frac{\prod_{\tau=1}^t \mathrm{A}(\Theta)\exp\langle \mathrm{B}(\Psi_\tau), \mathrm{C}(\Theta)\rangle\,\mathrm{f}(\Theta)}{\int_{\Theta^\star} \prod_{\tau=1}^t \mathrm{A}(\Theta)\exp\langle \mathrm{B}(\Psi_\tau), \mathrm{C}(\Theta)\rangle\,\mathrm{f}(\Theta)\,\mathrm{d}\Theta}$$

$$\underbrace{\overbrace{=}^{\langle\cdot,\cdot\rangle}}_{\mathit{linearity}} \quad \frac{\mathrm{A}^t(\Theta)\exp\left\langle \sum_{\tau=1}^t \mathrm{B}(\Psi_\tau), \mathrm{C}(\Theta)\right\rangle\mathrm{f}(\Theta)}{\int_{\Theta^\star} \mathrm{A}^t(\Theta)\exp\left\langle \sum_{\tau=1}^t \mathrm{B}(\Psi_\tau), \mathrm{C}(\Theta)\right\rangle\mathrm{f}(\Theta)\,\mathrm{d}\Theta}$$

$$\equiv \quad \frac{\mathrm{A}^{\nu_t}(\Theta)\exp\langle V_t, \mathrm{C}(\Theta)\rangle\,\mathrm{f}(\Theta)}{\int_{\Theta^\star} \mathrm{A}^{\nu_t}(\Theta)\exp\langle V_t, \mathrm{C}(\Theta),\rangle\,\mathrm{f}(\Theta)\,\mathrm{d}\Theta} = \frac{\mathrm{L}(\Theta, V_t, \nu_t)\mathrm{f}(\Theta)}{\mathrm{J}(V_t, \nu_t)}$$

$$\Leftrightarrow \quad V_t \equiv \sum_{\tau=1}^t \mathrm{B}(\Psi_\tau) = V_{t-1} + \mathrm{B}(\Psi_t), \quad \nu_t \equiv t = \nu_{t-1} + 1$$

$$\mathrm{f}(\Delta|A_t, \mathbf{K}_{t-1}) = \frac{\mathrm{J}(V_t, \nu_t)}{\mathrm{J}(V_{t-1}, \nu_{t-1})} = \frac{\mathrm{J}(V_{t-1} + \mathrm{B}(\Psi_t), \nu_{t-1} + 1)}{\mathrm{J}(V_{t-1}, \nu_{t-1})} \qquad \square$$

Finite-Dimensional Sufficient Statistic

The further discussion needs uses the notions statistic↳ and sufficient statistic↳ It only distinguishes statistic and its value when misunderstanding may arise.

- *finite-dimensional statistic* maps the estimation knowledge into a space whose finite dimension does not grow with observation time t.

Exponential Family and Tractable Parametric Models

- The recursive updating of the data vector Ψ_t is necessary for admitting the estimation recursivity, cf. (149).
- The posterior pd for a parametric model in the exponential family has the fixed functional form (151), which is determined by the value of the finite-dimensional sufficient statistic V_t, ν_t and prior pd $\mathsf{f}(\Theta)$.
- EF exhausts smooth parametric models with Θ-independent support admitting a finite-dimensional statistic, [97].
- The uniform parametric model has Θ-dependent support and admits finite-dimensional sufficient statistic, too.
- The parameter estimation coincides with the data updating part of filtering. It admits a finite-dimensional sufficient statistic if the observation model belongs to EF and the time-evolution model maps, [33],

$$\begin{aligned} \mathsf{f}(X_t|\mathbf{K}_t) \quad &\propto \quad \mathsf{A}(X_t)^{\nu_{t|t}} \exp\left\langle V_{t|t}, \mathsf{C}(X_t)\right\rangle \qquad (153) \\ &\to \mathsf{f}(X_{t+1}|\mathbf{K}_t) \propto \mathsf{A}(X_{t+1})^{\nu_{t+1|t}} \exp\left\langle V_{t+1|t}, \mathsf{C}(X_{t+1})\right\rangle . \end{aligned}$$

Estimation in EF with Conjugate Prior PD

- *conjugate prior* pd $\mathsf{f}(\Theta)$ is from the same set $\mathsf{f}_\star$ of pds as the posterior pd $\mathsf{f}(\Theta|\mathbf{K}_t)$.

  This definition makes a practical sense if the set $\mathsf{f}_\star$ is (substantially) simpler than the set $\mathsf{f}^\star$ of all pds on $\Theta^\star$.

- The pd $$\mathsf{f}(\Theta|V_0,\nu_0) \propto \mathsf{A}^{\nu_0}(\Theta)\exp\langle V_0, \mathsf{C}(\Theta)\rangle\,\chi_{\Theta^\star}(\Theta), \qquad (154)$$

  determined by the finite-dimensional prior statistics V_0, ν_0 and a non-negative function $\chi_{\Theta^\star}(\Theta)$ is conjugate to the exponential family.

  With it, prediction and estimation formulas (148), (151) are valid if
  - V_0, ν_0 replace the zero initial conditions in (149),
  - the function $\chi_{\Theta^\star}(\cdot)$ is formally used as the prior pd.

  Often, $\chi_{\Theta^\star}(\cdot)$ determines hard constraints on Θ and is chosen as indicator, which is equal to 1 on $\Theta^\star$ and it is zero otherwise.

ARX: Normal Autoregressive-Regressive Model Linear in Parameters with External Variables in Regression Vector

- *ARX model* of i-th factor is the parametric model given by the pd

$$
\begin{aligned}
\mathrm{m}(\Psi_{t;i},\Theta_i) &\equiv \mathrm{f}(\Delta_{t;i}|\Theta_i, A_t, \Delta_{t;i+1},\ldots,\Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1}) \qquad (155)\\
&= \mathrm{N}_{\Delta_{t;i}}(\theta_i'\psi_{t;i}, r_i) = (2\pi r_i)^{-0.5}\exp[-0.5 r_i^{-1}(\Delta_{t;i}-\theta'\psi_{t;i})^2]\\
&= \underbrace{(2\pi r_i)^{-0.5}}_{\mathrm{A}(\Theta_i)}\exp\{\mathrm{tr}(\underbrace{\Psi_{t;i}\Psi_{t;i}'}_{\mathrm{B}(\Psi_{t;i})}\underbrace{(-0.5[-1,\theta_i']' r_i^{-1}[-1,\theta_i'])}_{\mathrm{C}(\Theta_i)})\}
\end{aligned}
$$

$$
\begin{aligned}
\Theta_i \equiv (\theta_i, r_i) &= (\text{regression coefficient,noise variance})\\
&= (\ell_{\psi_i^\star}\text{-dimensional vector, positive scalar})\\
\psi_{t;i} &= \psi_i(A_t, \Delta_{t;i+1},\ldots,\Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1}) = \text{regression vector}\\
\Psi_{t;i} &= [\Delta_{t;i}, \psi_{t;i}']' = \text{data vector}
\end{aligned}
$$

- The exceptionally used subscript $_i$ stresses possible variations of quantities entering different factors, which make ARX model flexible.

Justification of ARX Model as the Prominent Feasible Case

It needs the notion

- *innovations*, martingale differences, $\varepsilon_t \equiv \Delta_t - \mathsf{E}[\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}]$ form a zero-mean sequence, uncorrelated with conditioning quantities.

The next construction [132] reveals how an ARX model is obtained

- assuming negligible error of Taylor expansion $\mathsf{E}[\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}] \approx \theta'\psi_t$
- assuming a finite constant conditional variance r of innovations
- selecting the pd of innovations via the maximum entropy principle
- bijectively transforming $\varepsilon_t \leftrightarrow \Delta_t$.

Recall

Regression vector ψ_t of a factor can be any known nonlinear function of the action A_t and the enriched knowledge $\Delta_{t;i+1}, \ldots, \Delta_{t;\ell_{\Delta^\star}}, \mathbf{K}_{t-1}$ that allows a recursive evaluation of the data vector $\Psi_{t-1}, D_t \to \Psi_t$.

Parameter Estimation for ARX Model

- For ARX model, the likelihood ↳ (151) is $\mathsf{L}(\Theta, \mathbf{K}_t) \equiv \mathsf{L}(\theta, r, \mathbb{V}_t, \nu_t) =$ $(2\pi r)^{-0.5\nu_t} \exp(-0.5 r^{-1}[-1, \theta']\mathbb{V}_t[-1, \theta']')$ gives conjugate prior ↳ (156)
- *NiW* Normal-inverse-Wishart (Normal-inverse-Gamma) pd

$$\mathcal{G}i\mathcal{W}(\mathbb{V}_0, \nu_0) \equiv \frac{(2\pi r)^{-0.5(\nu_0 + \ell_{\psi^\star} + 2)} \exp(-0.5 r^{-1}[-1, \theta']\mathbb{V}_0[-1, \theta']')}{\mathsf{J}(\mathbb{V}_0, \nu_0)} \quad (157)$$

$$\mathsf{J}(\mathbb{V}, \nu) = \Gamma(0.5\nu)\left({}^{\Delta}\mathbb{V} - {}^{\psi\Delta}\mathbb{V}' \, {}^{\psi}\mathbb{V}^{-1} \, {}^{\psi\Delta}\mathbb{V}\right)^{-0.5\nu} \left| {}^{\psi}\mathbb{V} \right|^{-0.5} 2^{0.5\nu} (2\pi)^{0.5\ell_{\psi^\star}}$$

The extended information matrix ↳ $\quad \mathbb{V} = \begin{bmatrix} {}^{\Delta}\mathbb{V} & {}^{\psi\Delta}\mathbb{V}' \\ {}^{\psi\Delta}\mathbb{V} & {}^{\psi}\mathbb{V} \end{bmatrix} \quad (158)$

is to be positive definite, $\mathbb{V} > 0$, and degrees $\nu > 0$ to get $\mathsf{J}(\mathbb{V}, \nu) < \infty$.

- Estimation provides posterior pds preserving NiW ↳ form $\mathcal{G}i\mathcal{W}(\mathbb{V}_t, \nu_t)$

$$\mathbb{V}_t = \mathbb{V}_{t-1} + \Psi_t \Psi_t' > 0, \;\; \nu_t = \nu_{t-1} + 1 > 0 \quad (159)$$

initiated by the statistic values of the conjugate prior ↳ (157).

- Let x denote $2r$ multiple of the exponent of the likelihood (156), which is the posterior pd obtained for the flat prior with $\mathbb{V}_0,\ \nu_0 \to 0$,

$$x = \sum_{\tau=1}^{t} \underbrace{(\Delta_\tau - \theta'\psi_\tau)^2}_{\text{prediction error}} = \Lambda_t + (\theta - \hat{\theta}_t)'\mathbb{C}_t^{-1}(\theta - \hat{\theta}_t) \quad (160)$$

$$\mathbb{C}_t = \left(\sum_{\tau=1}^{t} \psi_\tau \psi_\tau'\right)^{-1} \quad \text{LS covariance}$$

$$\hat{\theta}_t = C_t \sum_{\tau=1}^{t} \psi_\tau \Delta_\tau = \mathbb{C}_t\, {}^{\psi\Delta}\mathbb{V}_t \quad \text{LS parameter estimate}$$

$$\Lambda_t = {}^{\Delta}\mathbb{V}_t - {}^{\psi\Delta}\mathbb{V}_t'\, {}^{\psi}\mathbb{V}_t^{-1}\, {}^{\psi\Delta}\mathbb{V}_t = {}^{\Delta}\mathbb{V}_t - \hat{\theta}_t'\mathbb{C}_t^{-1}\hat{\theta}_t' \quad \text{LS remainder}$$

- For the considered flat prior, the moments of the NiW pd are [132]

$$\hat{\theta}_t = \mathsf{E}[\theta|\mathbb{V}_t, \nu_t], \quad \hat{r}_t = \mathsf{E}[r|\mathbb{V}_t, \nu_t] = \frac{\Lambda_t}{\nu_t - 2}, \quad \hat{r}_t\mathbb{C}_t = \mathrm{cov}[\theta|\mathbb{V}_t, \nu_t].$$

Remarks on Parameter Estimation for ARX Model

- *extended information matrix* is the name used for the statistics $\mathbb{V}_t$.
  The recursion $\mathbb{V}_t = \mathbb{V}_{t-1} + \Psi_t \Psi_t'$ algebraically transforms into
- *RLS*, recursive least squares, update $\hat{\theta}_t$, $\mathbb{C}_t$, $\hat{r}_t$, [132].
- Often, $\hat{\theta}_t$ and $\hat{r}_t$ are taken as the best point estimates of θ and r.
  For us, they form a part of sufficient statistic.
- The relation of RLS to the posterior pd and the general asymptotic of estimation, Prop. 21, provide rich asymptotic results for RLS.
- The non-flat prior pd, given by $\mathbb{V}_0 > 0, \nu_0 > 0$, guarantees that $\mathbb{V}_t > 0, \nu_t > 0$: the prior pd regularises the posterior pd.
  Even then RLS is numerically sensitive. The problem is faced by using
- $\mathbb{LDL}'$ *decomposition* of extended information matrix $\mathbb{V} = \mathbb{LDL}'$,
  $\mathbb{L}$ is lower triangular matrix with unit diagonal,
  $\mathbb{D}$ is diagonal matrix with positive diagonal, [19, 52, 77, 132].

Let the data vector $\Psi_t \in \Psi_t{}^\star = (\Delta^\star, \psi^\star)$ of a factor have a finite amount of realisations $|\Psi^\star| < \infty$. Then the most general model is:

- *Markov chain* is the parametric model given by (cf. Kronecker delta):

$$\begin{aligned} f(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) &= \prod_{\Psi\in\Psi^\star} \Theta_{\Delta|\psi}^{\delta(\Psi,\Psi_t)} = \exp\left[\sum_{\Psi\in\Psi^\star} \underbrace{\delta(\Psi,\Psi_t)}_{B_{\Delta|\psi}(\Psi)} \underbrace{\ln(\Theta_{\Delta|\psi})}_{C_{\Delta|\psi}(\Theta)}\right] \\ \Theta \in \Theta^\star &= \left\{\Theta_{\Delta|\psi} > 0, \sum_{\Delta\in\Delta^\star} \Theta_{\Delta|\psi} = 1 \ \forall \psi \in \psi^\star\right\}. \qquad (161) \end{aligned}$$

- *Dirichlet's pd* is conjugate prior to Markov chain defined on $\Theta^\star$ (161)

$$\begin{aligned} f(\Theta) &= \mathcal{D}i_\Theta(\mathbb{V}_0) = \prod_{\psi\in\psi^\star} \frac{\prod_{\Delta\in\Delta^\star} \Theta_{\Delta|\psi}^{\mathbb{V}_{0;\Delta|\psi}-1}}{\mathrm{Be}(\mathbb{V}_{0;\cdot|\psi})} \qquad (162) \\ \mathrm{Be}(\mathbb{V}_{\cdot|\psi}) &= \frac{\prod_{\Delta\in\Delta^\star} \Gamma(\mathbb{V}_{\Delta|\psi})}{\Gamma(\sum \ldots)}, \ \Gamma(\cdot) \text{ is gamma function}. \end{aligned}$$

Parameter Estimation & Prediction for Markov Chain

The self-reproducing Dirichlet's pd $\mathcal{D}i_{\Theta}(\mathbb{V})$ is determined by

- *occurrence matrix* $\mathbb{V} = (\mathbb{V}_{\Delta|\psi} > 0)_{\Psi \in \Psi^{\star}}$, which updates as follows

$$\mathbb{V}_{t;\Delta|\psi} = \mathbb{V}_{t-1;\Delta|\psi} + \delta(\Psi, \Psi_t). \tag{163}$$

- The corresponding predictive pd [77] reads

$$\mathsf{f}(\Delta|A, \psi, \mathbf{K}_{t-1}) = \mathsf{E}[\Theta_{\Delta|\psi}|A, \psi, \mathbb{V}_{t-1}] = \frac{\mathbb{V}_{t-1;\Delta|\psi}}{\sum_{\Delta \in \Delta^{\star}} \mathbb{V}_{t-1;\Delta|\psi}} \tag{164}$$

$\equiv$ relative frequency of occurrence of the data vector $\Psi = [\Delta, \psi']'$.

- (164) relates the frequentist view on probabilities to Bayes' theory.
- The asymptotic learning, Prop. 21, provides conditions under which the relative frequencies converge to unknown probabilities.

Estimation out of Exponential Family

- The exponential family↳ and special uniform pds provide a basic supply of dynamic factors admitting a finite-dimensional sufficient statistic↳.

  What can be done for other parametric models?

- Under natural conditions of DM↳ (65), Bayesian estimation, Prop. 17, updates the posterior pds via Bayes' rule↳ (80)

$$\mathsf{f}(\Theta|\mathbf{K}_t) = \frac{\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\mathsf{f}(\Theta|\mathbf{K}_{t-1})}{\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})}, \ t \in t^\star.$$

  Out of EF (143), the complexity of these pds blows up with t.

- The problem recursive estimation applicable out of EF is here addressed via equivalence approach [107, 108, 109].

The problem formulation

- Parametric models do not admit finite-dimensional sufficient statistic.
- Instead of the exact posterior pd $\mathsf{f}(\Theta|\mathbf{K}_{t-1})$, we deal with its approximation $\hat{\mathsf{f}}(\Theta|V_{t-1})$, a pd of a fixed functional form & determined by a finite-dimensional non-sufficient statistic V_{t-1}.
- $\hat{\mathsf{f}}(\Theta|V_{t-1})$ defines a multi-member equivalence set $\hat{\mathsf{f}}^{\star}(\Theta|V_{t-1}) \equiv \{\mathsf{f}(\Theta) : \text{pds on } \Theta^{\star} \text{ with a common approximating pd } \hat{\mathsf{f}}(\Theta|V_{t-1})\}$.
- We search for $\hat{\mathsf{f}}^{\star}(\Theta|V_{t-1})$ that can be updated recursively and includes the exact posterior pd in the discussed equivalence set.

  The choice of the statistic V_{t-1} & $\hat{\mathsf{f}}^{\star}(\Theta|V_{t-1})$ has to guarantee that Bayes' rule maps $\hat{\mathsf{f}}^{\star}(\Theta|V_{t-1})$ on equivalence set $\hat{\mathsf{f}}^{\star}(\Theta|V_{t})$.

Bayes'-Rule Compatible Non-Sufficient Statistic

Proposition 37 (Equivalence-Preserving Mapping)

Let $\mathsf{f}^{\star}(\Theta|\mathbf{K}_{t-1})$ *be a set of possible posterior pds* $\mathsf{f}(\Theta|\mathbf{K}_{t-1})$ *with a common, time, data, and parameter invariant support.*

Let the mapping

$$\mathsf{K}_{t;\mathsf{f}^{\star}} : \mathsf{f}^{\star}(\Theta|\mathbf{K}_{t-1}) \rightarrow V^{\star}_{t-1} \tag{165}$$

assign to each pd $\mathsf{f}(\Theta|\mathbf{K}_{t-1})$ *from* $\mathsf{f}^{\star}(\Theta|\mathbf{K}_{t-1})$ *a finite-dimensional statistic* $V_{t-1} \equiv \mathsf{K}_{t;\mathsf{f}^{\star}}(\mathbf{K}_{t-1})$ *"representing" it, i.e. determining* $\hat{\mathsf{f}}^{\star}(\Theta|V_{t-1})$.

Then, the value of V_{t-1} *can be exactly recursively updated using only its previous value and the current parametric model* $\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$ *iff:*

- $\mathsf{K}_{t;\mathsf{f}^{\star}}$ *is a time-invariant linear mapping* $\mathsf{K}_{t;\mathsf{f}^{\star}} \equiv \mathsf{K}_{\mathsf{f}^{\star}}$, $t \in t^{\star}$, *acting on logarithms of likelihoods and posterior pds involved.*
- $\mathsf{K}_{\mathsf{f}^{\star}}$ *maps* Θ*-independent elements on zero.*

Proof of the Sufficiency

Proof The difficult necessity proof is in [105, 106]. The sufficiency follows from the use of $\mathsf{K}_{\mathsf{f}^\star}$ on the logarithmic Bayes' rule (80) and exploitation of time-invariance and linearity of $\mathsf{K}_{\mathsf{f}^\star}$. The logarithm of normalising term, $\ln(\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}))$, maps on zero as it is independent of Θ. Thus,

$$\begin{aligned} V_t &= \mathsf{K}_{\mathsf{f}^\star}\left[\ln\left(\mathsf{f}(\Delta_t|\Theta,A_t,\mathbf{K}_{t-1})\right)\right] + V_{t-1}, \quad \text{with} \qquad (166) \\ V_0 &= \mathsf{K}_{\mathsf{f}^\star}[\ln(\mathsf{f}(\Theta))]. \end{aligned}$$ □

Remark 29 (On Model and Mapping $\mathsf{K}_{\mathsf{f}^\star}$)

- *(166) is recursion if we need not store all data for evaluating the likelihood $\mathsf{f}(\Delta_t|\Theta,A_t,\mathbf{K}_{t-1})$. As for dynamic EF, we constrain on*

$$\mathsf{f}(\Delta_t|\Theta,A_t,\mathbf{K}_{t-1}) = \mathsf{m}(\Psi_t,\Theta) \qquad (167)$$

given by a recursively updatable data vector Ψ_t.

- *Riezs representation of $\mathsf{K}_{\mathsf{f}^\star}$, [137], is (with an abused notation)*

$$\mathsf{K}_{\mathsf{f}^\star}(\ln(\mathsf{f}(\Theta|\mathbf{K}_{t-1}))) = \int_{\Theta^\star} \mathsf{K}_{\Theta^\star}(\Theta)\ln(\mathsf{f}(\Theta|\mathbf{K}_{t-1}))\,\mathrm{d}\Theta, \quad \int_{\Theta^\star}\mathsf{K}_{\Theta^\star}(\Theta)\,\mathrm{d}\Theta = 0.$$

Approximation in Data-Vector Space

Values of the non-sufficient statistic V_t compress partial information on

- *empirical pd of data vector* , specified by Dirac delta $\delta(\cdot,\cdot)$,

$$\mathsf{f}_t(\Psi) \equiv \frac{1}{t}\sum_{\tau=1}^{t}\delta(\Psi,\Psi_\tau),\ \Psi \in \Psi^\star \equiv \bigcup_{t\in t^\star} \Psi_t{}^\star, \tag{168}$$

  which is sufficient statistic as $\mathsf{f}(\Theta|\mathbf{K}_{t-1}) = \mathsf{f}(\Theta|\mathsf{f}_t(\cdot))$.

- The pd $\mathsf{f}_t(\Psi)$ determines the non-sufficient statistic V_t (166) via

$$V_t = t\int_{\Psi^\star} \mathsf{f}_t(\Psi) \underbrace{\int_{\Theta^\star} \mathsf{K}_{\Theta^\star}(\Theta)\ln(\mathsf{m}(\Psi,\Theta))\,\mathrm{d}\Theta}_{\mathsf{K}_{\Psi^\star}(\Psi)}\,\mathrm{d}\Psi + V_0. \tag{169}$$

- The statistic values update recursively starting from a given V_0

$$V_t = V_{t-1} + \mathsf{K}_{\Psi^\star}(\Psi_t),\quad \mathsf{K}_{\Psi^\star}(\Psi) \equiv \int_{\Theta^\star} \mathsf{K}_{\Theta^\star}(\Theta)\ln(\mathsf{m}(\Psi,\Theta))\,\mathrm{d}\Theta. \tag{170}$$

Recursively Feasible Approximation of Empirical PD

- The posterior pd can be expressed via empirical pd of data vector $_\downarrow f_t(\Psi)$

$$f(\Theta|\mathbf{K}_t) = f(\Theta|f\,t(\cdot)) \propto f_0(\Theta)\exp\left[t\int_{\Psi^\star} f_t(\Psi)\ln(m(\Psi,\Theta))\,\mathrm{d}\Psi\right]. \quad (171)$$

- This hints to take the approximate posterior pd $_\downarrow$ in the form

$$\hat{f}(\Theta|V_t) \propto f_0(\Theta)\exp\left[t\int_{\Psi^\star}\hat{f}(\Psi|V_t)\ln(m(\Psi,\Theta))\,\mathrm{d}\Psi\right]. \quad (172)$$

- The estimate $\hat{f}(\Psi|V_t)$ of the unknown pd $f_t(\Psi)$ is gained via minimum KLD principle $_\downarrow$, i.e. for a given prior guess $\hat{f}_0(\Psi)$, KLD $D(\hat{f}(\Psi)||\hat{f}_0(\Psi))$ is minimised with respect to $\hat{f}(\Psi)$ under the constraint, cf. (166)

$$\int_{\Psi^\star}\hat{f}(\Psi)K_{\Psi^\star}(\Psi)\,\mathrm{d}\Psi = (V_t - V_0)/t. \quad (173)$$

The solution defining Bayes'compatible approximation (172) has the form

$$\hat{f}(\Psi|V_t) \propto \hat{f}_0(\Psi)\exp[\lambda_t' K_{\Psi^\star}(\Psi)], \quad (174)$$

where the multipliers λ_t are chosen so that (173) is met for $\hat{f}(\Theta|V_t)$.

Summary of Approximate Equivalence Estimation

Off line phase consists of selecting

- parametric,model $\mathsf{f}(\Delta|\Theta, A_t, \mathbf{K}_{t-1}) = \mathsf{m}(\Psi_t, \Theta)$ with data vector Ψ_t
- kernel $\mathsf{K}_{\Theta^\star}(\Theta)$ of Riezs representation of the map $\mathsf{K}_{\mathsf{f}^\star}[\ln(\mathsf{f}(\Theta|\mathbf{K}_t))]$
- algorithm evaluating function $\mathsf{K}_{\Psi^\star}(\Psi) = \int_{\Theta^\star} \mathsf{K}_{\Theta^\star}(\Theta) \ln(\mathsf{m}(\Psi, \Theta))\, \mathrm{d}\Theta$
- a prior pd $\mathsf{f}(\Theta)$ defining $V_0 = \int_{\Theta^\star} \mathsf{K}_{\Theta^\star}(\Theta) \ln(\mathsf{f}(\Theta))\, \mathrm{d}\Theta$
- a prior (flat) guess $\hat{\mathsf{f}}_0(\Psi)$ of the empirical pd of data vector.

On line phase runs for $t \in t^\star$ while Ψ_t are recursively updated

- the stored statistic is updated $V_t = V_{t-1} + \mathsf{K}_{\Psi^\star}(\Psi_t)$
- the empirical pd is approximated by $\hat{\mathsf{f}}(\Psi|V_t) \propto \hat{\mathsf{f}}_0(\Psi) \exp[\lambda_t' \mathsf{K}_{\Psi^\star}(\Psi)]$, where λ_t guarantee $\int_{\Psi^\star} \hat{\mathsf{f}}(\Psi|V_t) \mathsf{K}_{\Psi^\star}(\Psi)\, \mathrm{d}\Psi = (V_t - V_0)/t$, (173)
- The posterior pd $\mathsf{f}(\Theta|\mathbf{K}_t)$ is approximated by
  $\hat{\mathsf{f}}(\Theta|V_t) \propto \mathsf{f}_0(\Theta) \exp\left[t \int_{\Psi^\star} \hat{\mathsf{f}}(\Psi|V_t) \ln(\mathsf{m}(\Psi, \Theta))\, \mathrm{d}\Psi\right]$, (172).

Remarks on Equivalence-Based Recursive Estimation

- The vector of (generalised [163]) functions $\mathsf{K}_{\Theta^\star}(\Theta)$ is the key tuning knob. Options leading to discrete images of $\ln(\mathsf{f}(\Theta|\mathbf{K}_k))$ and/or its derivatives were successfully tried, but a deeper insight is needed.

- Term "equivalence approach" stresses the posterior pds with given V_t are undistinguishable: $\mathsf{f}^\star(\Theta|\mathbf{K}_t)$ splits into equivalence sets $\hat{\mathsf{f}^\star}(\Theta|V_t)$.

- The commutation of the mapping $\mathsf{K}_{\mathsf{f}^\star}$ with the data updating$_\downarrow$ is vital. Recursion for V_t is exact and approximation errors caused by the use of $\hat{\mathsf{f}}(\Theta|V_t)$ instead of $\mathsf{f}(\Theta|\mathbf{K}_t)$ do not accumulate!

- Use of a noncommutative projection, common to many approximation recursive estimations, $\mathsf{K}_{t;\mathsf{f}^\star} : \mathsf{f}^\star(\Theta|\mathbf{K}_t) \to V_t^\star$ is always endangered by a divergence as Bayes' rule is as a dynamic system evolving the functions $\ln(\mathsf{f}(\Theta|\mathbf{K}_t))$ at the stability boundary.

Computational Aspects of Equivalence-Based Estimation

- The integration defining $\mathsf{K}_{\Psi^\star}(\Psi) = \int \mathsf{K}_{\Theta^\star}(\Theta) \ln(\mathsf{m}(\Psi, \Theta))\, \mathrm{d}\Psi$ is the computationally hardest part of the overall algorithm. It can be made off-line but its results must be efficiently stored.
- The solution of the nonlinear equation for Lagrangian multipliers λ_t is also hard but standard problem.
- We would like to get the exact posterior pd if the model belongs to the exponential family (143). This dictates the choice of the mapping $\mathsf{K}_{\mathsf{f}^\star}$ that should lead to $\mathsf{K}_{\Psi^\star}(\Psi)$ with the range including $[\mathsf{B}(\Psi), 1]$.
- The equivalence approach excludes a lot of standard statistics. This motivates contemporary search for approximate learning, which allows errors resulting from non-commutativity but tries to bound them.

Problem – Tracking of Slowly Varying Parameters

- The parameter estimation relies on time-invariance of parameters. If this assumption is violated, the Bayesian filtering ↳ is to be used.
- Filtering requires the time-evolution model ↳ and its exact feasibility is even more restricted than the estimation.

  This stimulated interest in an intermediate case, in
- *parameter tracking*, which is estimation of slowly varying parameters $\Theta_{t+1} \approx \Theta_t$ with a simplified specification of time-evolution model ↳.
- Parameter tracking forms the core of many adaptive systems, which often approximates pds $\mathsf{f}_{t+1|t}(\Theta) \equiv \mathsf{f}(\Theta_{t+1} = \Theta|\mathbf{K}_t)$ via
- *forgetting*, which tries to exploit for estimation of Θ_{t+1} the valid part of $\mathbf{K}_t$ and discard invalid, typically obsolete, knowledge, [28, 74, 103, 104, 110, 111, 112, 121, 132].
- A forgetting version implied by minimum KLD principle ↳ follows.

Tracking as Supporting DM Task: Formulation

The DM elements↳ of this task at $t \in t^\star$ are

- ignorance↳ $\mathbf{G}_{\mathcal{A}_t^\star}$ is time-varying parameter Θ_{t+1} entering the observation model↳ and the posterior pd $\mathsf{f}_{t+1|t}(\Theta) \equiv \mathsf{f}(\Theta_{t+1} = \Theta|\mathbf{K}_t)$
- action↳ $\mathcal{A}_t = \hat{\mathsf{f}}_{t+1|t}(\Theta)$ is an estimate of $\mathsf{f}_{t+1|t}(\Theta) \equiv \mathsf{f}(\Theta_{t+1} = \Theta|\mathbf{K}_t)$
- knowledge↳ $\mathbf{K}_{\mathcal{A}_t^\star}$ consists of:
  - the old supporting action $\mathcal{A}_{t-1} = \hat{\mathsf{f}}_{t|t-1}(\Theta)$
  - likelihood – the observation model↳ $\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})$ in realisation
  - an alternative pd ${}^{\mathcal{A}}\mathsf{f}_{t+1|t}(\Theta)$ a guess of $\mathsf{f}_{t+1|t}(\Theta)$
  - applicability of the assumption on slow parameter changes

$$\begin{aligned} \mathsf{f}_{t+1|t}(\Theta) \quad &\approx \quad \mathsf{f}_{t|t}(\Theta) \equiv \mathsf{f}(\Theta_t = \Theta|\mathbf{K}_t) \propto \mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{\tau-1})\hat{\mathsf{f}}_{t|t-1}(\Theta) \\ \approx \quad &\text{is} \quad \text{expressed by} \quad \mathsf{D}(\mathsf{f}_{t+1|t}||\mathsf{f}_{t|t}) \leq \mathsf{D}(\mathsf{f}_{t+1|t}||\,{}^{\mathcal{A}}\mathsf{f}_{t+1|t}) \end{aligned} \tag{175}$$

Tracking DM Task: Solution via Minimum KLD Principle

Proposition 38 (Tracking by Forgetting)

The minimum KLD principle with the prior guess ${}^{\mathcal{A}}\mathsf{f}_{t+1|t}$ *and the constraint* $\mathsf{D}(\mathsf{f}_{t+1|t}||\mathsf{f}_{t|t}) \leq \mathsf{D}(\mathsf{f}_{t+1|t}||\,{}^{\mathcal{A}}\mathsf{f}_{t+1|t})$ *gives the next solution*

$$\begin{aligned} \hat{\mathsf{f}}_{t+1|t}(\Theta) &\equiv \hat{\mathsf{f}}_{\lambda_t}(\Theta) \propto \mathsf{f}^{\lambda_t}_{t|t}(\Theta)\,{}^{\mathcal{A}}\mathsf{f}^{1-\lambda_t}_{t+1|t}(\Theta) \qquad (176)\\ &= [\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1})\hat{\mathsf{f}}_{t|t-1}(\Theta)]^{\lambda_t}\,{}^{\mathcal{A}}\mathsf{f}^{1-\lambda_t}_{t+1|t}(\Theta), \end{aligned}$$

where $\lambda_t \in [0,1]$ *solves the equality*

$$\mathsf{e}(\lambda) \equiv \int_{\Theta^\star} \hat{\mathsf{f}}_{\lambda_t}(\Theta) \ln\left(\frac{{}^{\mathcal{A}}\mathsf{f}_{t+1|t}(\Theta)}{\mathsf{f}_{t|t}(\Theta)}\right) \mathrm{d}\Theta = 0. \qquad (177)$$

Proof Re-arranging of Kuhn-Tucker functional with a multiplier $\lambda \geq 0$ into KLD gives the solution form. The difference $\mathsf{e}(\lambda)$ (177) of KLDs in the constraint (175) is continuous on [0,1] and meets

$$\mathsf{e}(1) = -\mathsf{D}(\mathsf{f}_{t|t}||\,{}^{\mathcal{A}}\mathsf{f}_{t+1|t}) \leq 0 \leq \mathsf{D}({}^{\mathcal{A}}\mathsf{f}_{t+1|t}||\mathsf{f}_{t|t}) = \mathsf{e}(0).$$

This implies existence $\lambda_t \in [0,1]$ giving $\mathsf{e}(\lambda_t) = 0$. □

On Forgetting Factor and Alternative PD

- Forgetting factor λ controls compromise between the posterior pd obtained under the hypothesis $\Theta_t =$ constant & the alternative pd ${}^{a}\!f$.

  The alternative pd ${}^{a}\!f$ delimits where Θ_{t+1} may move.

- The closer λ is to unity, the slower changes are expected, i.e. the higher weight the posterior pd corresponding to constant Θ gets.

- The observation model, serving for learning from data, is the more "flattened" the older data is processed: the older data records influence the results less than the new, they are gradually "forgotten".

- For ${}^{a}\!f \propto 1$ and $\lambda < 1$, the time evolution reduces to flattening of the pd obtained after data updating. It fits intuition. Uncertainty about unknown Θ can hardly decrease without knowing a relevant time-evolution model (63) & without new information processed.

Stabilised and Exponential Forgetting as Special Cases

- The prior pd is a typical, reasonably conservative, choice of the alternative pd. This alternative pd prevents us to forget the "guaranteed" information and stabilises whole learning.
- The stabilisation reflects positively in numerical implementations of learning. Without this, the posterior pd may become too flat whenever the information brought by new data is poor.
- Non-informativeness of data is more rule than exception. It is true especially for regulation [123] that tries to make the closed decision loop as quiet as possible, which diminishes information brought.
- In the extreme case of uniform alternative, the solution is called exponential forgetting otherwise it is called stabilised forgetting.
- The forgetting operation (176) preserves the basic property of the time updating: the posterior pd on parameters propagates without obtaining any new measured information, Prop. 16.

Forgetting is An Active Research Area

- Introduction of an optional factor $\phi > 0$ into the constraint (175)

$$\mathsf{D}(\mathsf{f}_{t+1|t}||\tilde{\mathsf{f}}_{t+1|t}) \leq \phi \mathsf{D}(\mathsf{f}_{t+1|t}||\,{}^{a}\mathsf{f}_{t+1|t})$$

  preserves the solution form (176). It makes the forgetting factor $\lambda(\phi)$ a tuning knob solving analogy of (177). It can also be estimated.

- The predictive pd depends on ϕ in a complex way. The equivalence approach, Sec. 30, is applicable, say, with the kernel $\mathsf{K}(\phi) = [\delta(\phi, \phi_1) - \delta(\phi, \phi_0), \ldots, \delta(\phi, \phi_{\ell_{\mathsf{K}^\star}}) - \delta(\phi, \phi_0)]$, for fixed grid $0 \leq \phi_0 < \ldots < \phi_{\ell_{\mathsf{K}^\star}} \leq 1$.

- Recently, the proposed forgetting was found as the prominent counter-measure against approximation error accumulation [73]. It also allows to counteract influence of imprecise data.

- Key conclusion: Adaptive learning is impossible without forgetting!

Estimation with Forgetting in Exponential Family

- Let the observation model ↳ belong to exponential family ↳

$$\mathsf{f}(\Delta_t|\Theta, A_t, \mathbf{K}_{t-1}) = \mathsf{f}(\Delta_t|\Theta, \psi_t) = \mathsf{A}(\Theta) \exp \langle \mathsf{B}(\Psi_t), \mathsf{C}(\Theta) \rangle$$

  and the conjugate prior ↳ pd $\mathsf{f}(\Theta|\mathbf{K}_t) \propto \mathsf{A}^{\nu_t}(\Theta) \exp \langle V_t, \mathsf{C}(\Theta) \rangle$ is considered, given by the sufficient statistic ↳ (V_t, ν_t).

- Let us allow slow parameter changes with the forgetting factor $\lambda_t \in [0, 1]$ and the alternative conjugate prior ↳ pd given by the sufficient statistic ↳ ${}^{a}V_t$, ${}^{a}\nu_t$.

- Then, the prediction and estimation formulas, Prop. 36, remain unchanged with statistics evolving according to the recursion

$$\begin{aligned} V_t &= \lambda_t(V_{t-1} + \mathsf{B}(\Psi_t)) + (1-\lambda_t)\,{}^{a}V_t, \quad V_0 \text{ given}, \\ \nu_t &= \lambda_t(\nu_{t-1} + 1) + (1-\lambda_t)\,{}^{a}\nu_t, \quad \nu_0 \text{ given}. \end{aligned}$$

...Omitted Topics

Useful particular examples at least approximately feasible including

- Linear models with restricted support [130]
- Finite mixtures with factor s from exponential family [77]. Their importance is stressed in Sec. 40.

Well established approximation techniques like

- Monte Carlo techniques and their sequential variants known as particle filters [140],
- Variational Bayes exploiting the alternative version of KLD and approximating a joint pds by product of conditionally independent factors [154].

Feasible and Approximate Design

Scope

- This part is the design counterpart of Part 12 dealing with learning.
- FPD ↳ formally covers an extreme width of DM tasks. Just a few of them are solvable exactly and to some there are well established approximate solution techniques. Both are reviewed here.
- The material does not cover full width of theory due to limited
  - coverage by the contemporary research,
  - knowledge of lecturer,
  - available time.
- The part focuses on data-driven design ↳ with parameter estimation ↳.

Evaluation Problem

- Prop. 23 solves the data-driven FPD↳ and provides the optimal randomised decision rule↳s for $t = h, h-1, \ldots, 1$

$$
\begin{aligned}
{}^{o}\mathsf{f}(A_t|\mathbf{K}_{t-1}) &= {}^{I}\mathsf{f}(A_t|\mathbf{K}_{t-1})\frac{\exp[-\omega(A_t,\mathbf{K}_{t-1})]}{\gamma(\mathbf{K}_{t-1})},\ \gamma(\mathbf{K}_h)=1\\
\gamma(\mathbf{K}_{t-1}) &\equiv \int_{A_t^\star} {}^{I}\mathsf{f}(A_t|\mathbf{K}_{t-1})\exp[-\omega(A_t,\mathbf{K}_{t-1})]\,\mathrm{d}A_t \quad \text{if } t<h\\
\omega(A_t,\mathbf{K}_{t-1}) &\equiv \int_{\Delta_t^\star} \mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1})\ln\left(\frac{\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1})}{\gamma(\mathbf{K}_t)\,{}^{I}\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1})}\right) d\Delta_t.
\end{aligned}
$$

- The evaluation of high-dimensional integrals can be conceptually solved via Monte-Carlo techniques, however, the storing high dimensional functions $\omega(A_t,\mathbf{K}_{t-1})$, $\gamma(\mathbf{K}_{t-1})$ is computationally hard.

- The discussion of practical evaluations of this strategy starts from analytically feasible cases followed by common approximation ways.

Good Design has Dual – Exploitive & Explorative – Actions

- The predictive pd is obtain via a feasible parameter estimation, Sec. 28. It gives $\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1}) \approx \mathsf{f}(\Delta_t|A_t, \Psi_{t-1}, V_{t-1}, \nu_{t-1})$.
- The data vector Ψ_{t-1} and the statistic values V_{t-1}, ν_{t-1} form the finite-dimensional and recursively-updatable information state. This defines state-space model $\mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1})$ with an observable state $\mathcal{X}_t$.
- As globally optimal strategy uses $\mathcal{X}_t$, the ideal pd (98) models $A_t, \mathcal{X}_t$.

Obviously

- The functions occurring in FPD depend on the finite-dimensional information state: $\gamma(\mathbf{K}_{t-1}) = \gamma(\mathcal{X}_{t-1})$, $\omega(A_t, \mathbf{K}_{t-1}) = \omega(A_t, \mathcal{X}_{t-1})$.

Less obviously

- The optimal strategy jointly influences in a balanced way both the intended quantities (exploitation) and learning process (exploration).

This duality, recognised in [42], is an important approximation aspect.

Feasible FPD with Finite Number of Information States

If the information state $\mathcal{X}_t$ and actions A_t have finite numbers of realisations then the Prop. 23 provides directly the optimum

$$
\begin{aligned}
{}^{o}\mathsf{f}(A_t|\mathcal{X}_{t-1}) &= {}^{I}\mathsf{f}(A_t|\mathcal{X}_{t-1})\frac{\exp[-\omega(A_t,\mathcal{X}_{t-1})]}{\gamma(\mathcal{X}_{t-1})},\ \gamma(\mathcal{X}_h)=1 \qquad (178)\\
\gamma(\mathcal{X}_{t-1}) &\equiv \sum_{A\in A_t^\star} {}^{I}\mathsf{f}(A_t|\mathcal{X}_{t-1})\exp[-\omega(A_t,\mathcal{X}_{t-1})], \quad \text{if } t<h\\
\omega(A_t,\mathcal{X}_{t-1}) &\equiv \sum_{\mathcal{X}\in\mathcal{X}_t^\star} \mathsf{f}(\mathcal{X}_t|A_t,\mathcal{X}_{t-1})\ln\left(\frac{\mathsf{f}(\mathcal{X}_t|A_t,\mathcal{X}_{t-1})}{\gamma(\mathcal{X}_t)\,{}^{I}\mathsf{f}(\mathcal{X}_t|A_t,\mathcal{X}_{t-1})}\right).
\end{aligned}
$$

- All functions are tables, e.g. $\omega(A_t,\mathcal{X}_{t-1})$ and $\gamma(\mathcal{X}_t)$ with $|A^\star|\times|\mathcal{X}^\star|$ and $|\mathcal{X}^\star|$ entries. The cardinalities $|A^\star|$, $|\mathcal{X}^\star|$ determine evaluation complexity: tables are stored, summing runs over $A^\star$, $\mathcal{X}^\star$, etc.

  FPD is simple for small $|A^\star|$, $|\mathcal{X}^\star|$ & infeasible for large $|A^\star|$, $|\mathcal{X}^\star|$.

- The feasibility is slightly extended if the stationary strategy, obtained for horizon $h\to\infty$, Prop. 14, is considered.

Feasible FPD with LN Information State: Formulation

- Let the observable information state ordered into the vector $\mathcal{X}_t$ be described by the linear normal (LN) model for $t \in t^\star = \{1, \ldots, h\}$

$$\begin{aligned} \mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1}) &= \mathsf{N}_{\mathcal{X}_t}(\mathbb{A}\mathcal{X}_{t-1} + \mathbb{B}A_t, {}^{\mathcal{X}}\mathbb{R}) \\ \mathsf{N}_{\mathcal{X}}(M, \mathbb{R}) &= |2\pi\mathbb{R}|^{-0.5} \exp[-0.5(\mathcal{X} - M)' \, {}^{\mathcal{X}}\mathbb{R}^{-1}(\mathcal{X} - M)] \end{aligned} \tag{179}$$

  determined by known matrices $(\mathbb{A}, \mathbb{B}, {}^{\mathcal{X}}\mathbb{R} \geq 0)$.

- Let the ideal pd be also LN model

$$\begin{aligned} {}^{\mathsf{I}}\mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1}) &= \mathsf{N}_{\mathcal{X}_t}({}^{\mathsf{I}}\mathbb{A}\mathcal{X}_{t-1} + {}^{\mathsf{I}}\mathbb{B}A_t, {}^{\mathsf{I}}\mathbb{R}) \\ {}^{\mathsf{I}}\mathsf{f}(A_t|\mathcal{X}_{t-1}) &= \mathsf{N}_{A_t}({}^{\mathsf{I}}\mathbb{C}\mathcal{X}_{t-1}, {}^{A\mathsf{I}}\mathbb{R}) \end{aligned} \tag{180}$$

  determined by known matrices ${}^{\mathsf{I}}\mathbb{A}$, ${}^{\mathsf{I}}\mathbb{B}$, ${}^{\mathsf{I}}\mathbb{R} \geq 0$, ${}^{\mathsf{I}}\mathbb{C}$, ${}^{A\mathsf{I}}\mathbb{R} > 0$.

FPD with Linear normal Information State: Solution

Proposition 39 (Linear Normal (LN) FPD)

Let the system with finite-dimensional information state $\mathcal{X}_t$ be described by LN model (179) and the ideal pd in FPD be also LN given by pds (180) with ${}^{A\mathsf{I}}\mathbb{R} > 0$. Then, the optimal decision rule is

$$
\begin{aligned}
{}^{\mathsf{o}}\mathsf{f}(A_t|\mathcal{X}_{t-1}) &= \mathsf{N}_{A_t}(\mathbb{L}_t'\mathcal{X}_{t-1},\, {}^{A}\mathbb{R}_t) \text{ with} \qquad (181)\\
{}^{A}\mathbb{R}_t^{-1} &= {}^{A\mathsf{I}}\mathbb{R}^{-1} + \mathbb{B}'\mathbb{S}_t^{-1}\mathbb{B} + (\mathbb{B} - {}^{\mathsf{I}}\mathbb{B})'\,{}^{\mathsf{I}}\mathbb{R}_t^{-1}(\mathbb{B} - {}^{\mathsf{I}}\mathbb{B})\\
\mathbb{L}_t' &= {}^{A}\mathbb{R}_t\left[{}^{A\mathsf{I}}\mathbb{R}^{-1}\,{}^{\mathsf{I}}\mathbb{C} + \mathbb{B}'\mathbb{S}_t^{-1}\mathbb{A} + (\mathbb{B} - {}^{\mathsf{I}}\mathbb{B})'\,{}^{\mathsf{I}}\mathbb{R}^{-1}(\mathbb{A} - {}^{\mathsf{I}}\mathbb{A})\right]
\end{aligned}
$$

They are given by positive semi-definite Riccati matrix $\mathbb{S}_t^{-1}$ that evolves

$$
\mathbb{S}_{t-1}^{-1} = \mathbb{A}\mathbb{S}_t^{-1}\mathbb{A}' + (\mathbb{A} - {}^{\mathsf{I}}\mathbb{A})\,{}^{\mathsf{I}}\mathbb{R}^{-1}(\mathbb{A} - {}^{\mathsf{I}}\mathbb{A})' - \mathbb{L}_t\,{}^{A}\mathbb{R}_t\mathbb{L}_t', \text{ with } \mathbb{S}_h^{-1} = 0. \qquad (182)
$$

Proof It uses Prop. 23. Backward induction gives $\gamma(\mathcal{X}_t) \propto \mathsf{N}_{\mathcal{X}_t}(0, \mathbb{S}_t)$ with $\mathbb{S}_h^{-1} = 0 \Leftrightarrow \gamma(\mathcal{X}_h) = 1$, and finds the decision rule and the recursion for $\mathbb{S}_t^{-1}$. The cumbersome evaluations just manipulate quadratic forms. $\square$

Expected Quadratic Form & Square Completion Used in Proof

Straightforward derivation of LN FPD exploits

- *expected quadratic form*

$$
\begin{aligned}
\mathsf{E}[\mathcal{X}'\mathbb{R}^{-1}\mathcal{X}] &= \mathsf{E}[(\mathcal{X}-\mathsf{E}[\mathcal{X}]+\mathsf{E}[\mathcal{X}])'\mathbb{R}^{-1}(\mathcal{X}-\mathsf{E}[\mathcal{X}]+\mathsf{E}[\mathcal{X}])] \\
&= \mathrm{tr}\{\mathsf{E}[(\mathcal{X}-\mathsf{E}[\mathcal{X}])'\mathbb{R}^{-1}(\mathcal{X}-\mathsf{E}[\mathcal{X}])]\} + 2\mathsf{E}[\mathcal{X}']\underbrace{\mathsf{E}[\mathbb{R}^{-1}(\mathcal{X}-\mathsf{E}[\mathcal{X}])]}_{=0} \\
&+ \mathsf{E}[\mathcal{X}']\mathbb{R}^{-1}\mathsf{E}[\mathcal{X}] = \mathrm{tr}[\mathbb{R}^{-1}\mathrm{cov}(\mathcal{X})] + \mathsf{E}[\mathcal{X}']\mathbb{R}^{-1}\mathsf{E}[\mathcal{X}]
\end{aligned}
$$

- *square completion* for a matrix $\mathbb{R}' = \mathbb{R} > 0$ and vectors Υ, $\mathcal{X}$ of compatible dimensions define $\hat{\mathcal{X}} = \mathbb{R}\Upsilon$, $\zeta = \hat{\mathcal{X}}'\mathbb{R}^{-1}\hat{\mathcal{X}}$. Then,

$$
\begin{aligned}
\mathcal{X}'\mathbb{R}^{-1}\mathcal{X} - 2\mathcal{X}'\Upsilon &= (\mathcal{X}-\hat{\mathcal{X}})'\mathbb{R}^{-1}(\mathcal{X}-\hat{\mathcal{X}}) - \zeta \\
= \quad & \mathcal{X}'\mathbb{R}^{-1}\mathcal{X} - 2\mathcal{X}\underbrace{\mathbb{R}^{-1}\hat{\mathcal{X}}}_{\Upsilon} + \underbrace{\hat{\mathcal{X}}'\mathbb{R}^{-1}\hat{\mathcal{X}} - \zeta}_{=0}
\end{aligned}
$$

- *integral of quadratic form in exponent*

$$
|2\pi\mathbb{R}|^{0.5} = \int_{\mathcal{X}\in\mathcal{X}^\star} \exp[-0.5(\mathcal{X}-\hat{\mathcal{X}})'\mathbb{R}^{-1}(\mathcal{X}-\hat{\mathcal{X}})]\,\mathrm{d}\mathcal{X}.
$$

Relation of LN FPD to Standard Linear Quadratic Design

- Evaluation of LN FPD coincide with linear-quadratic design. It dominates the standard design [120] and optimises an additive quadratic performance index, see Example 11.
- The matrix ${}^{\mathsf{I}}\mathbb{R}^{-1}$ corresponds with state penalisation and ${}^{A\mathsf{I}}\mathbb{R}^{-1}$ with the action penalisation. This qualitative observation allowed to adapt the performance index to observed behaviour: to learn it [79].
- Numerical solution requires a significant care: $\mathbb{L}\mathbb{D}\mathbb{L}'$ type decomposition of the Riccati matrix should be used [83].
- The decision rule has the fixed form with "parameters" $\mathbb{L}_t$, ${}^{A}\mathbb{R}_t$ given by parameters of the system and of the ideal closed-loop model. This:
    - is the main source of feasibility,
    - can deal with time-dependent parameters arising in approximate evaluations (linearisation, adaptation).

Need for Approximate (Suboptimal) Adaptive Design

- Markov chain based FPD with small finite $|A^\star|$, $|\mathcal{X}^\star|$ and LN based FPD (almost) exhaust feasible cases. Otherwise, an approximate (suboptimal) design is used.
- Approximate designs of standard strategies are elaborated in [18, 149].
- Approximations outlined here are connected with adaptive system↳s. They approximate the optimal solutions locally around the behaviour realisation↳ [71].

  Classical references to adaptive systems are [7, 83, 100, 123].

The dynamic design↳ essentially predicts possible behaviours of the system interacting with compared strategies↳ and selects the most favourable one.

This determines complexity causes and possible ways of overcoming them.

Classification of Design Complexity Causes

The complexity of the strategy design stems mainly from:

1. Complexity of the predictive pd originating in complexity of
   - the parametric model or the observation model
   - the time-evolution model, which relates
     the knowledge and the optional action to the ignorance.
2. Richness of the ignorance space inspected when optimising strategy
3. Complexity of full DM consisting of many interconnected subtasks.

Any suboptimal design tries to suppress influence of these complexity causes.

- The described techniques suit to adaptive systems and the main passed message is that approximations can have unexpected consequences, for instance,
  - loss of quality up to in stability
  - passivity, i.e. lack of exploration effort.

Approximations of Predictive PD

- The predictive pd↳s $(\mathrm{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1}))_{\tau=t}^{h}$ gained by Bayesian filtering ↳ or parameter estimation↳, Props 16, 17, serve for "planning" between the real time t & horizon↳ h. Their complexity makes the design complex.
- The predictive pds have the form

$$\mathrm{f}(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1}) = \int_{X_\tau^\star} \mathrm{f}(\Delta_\tau|X_\tau,A_\tau,\mathbf{K}_{\tau-1})\mathrm{f}(X_\tau|A_\tau,\mathbf{K}_{\tau-1})\,\mathrm{d}X_\tau. \quad (183)$$

Their approximations have the following common form

$$\hat{\mathrm{f}}(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) = \int_{X_\tau^\star} \mathrm{f}(\Delta_\tau|X_\tau,A_\tau,\mathbf{K}_{\tau-1})\hat{\mathrm{f}}_\tau(X_\tau|A_\tau,\mathbf{K}_{\tau-1})\,\mathrm{d}X_\tau. \quad (184)$$

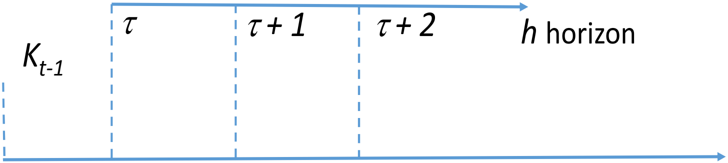
- The approximation (184) covers both right-hand-side approximate integration and use of a simpler pd $\hat{\mathrm{f}}(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) \approx \mathrm{f}(X_\tau|A_\tau,\mathbf{K}_{\tau-1})$.

Timing in Approximations of Predictive PD

ad cause 1

$$\boxed{f(\Delta_\tau \,/\, A_\tau, K_{\tau-1}) = \int_{X^*} f(\Delta_\tau \,|\, X_\tau, A_\tau, K_{\tau-1})\, f(X_\tau \,|\, A_\tau, K_{\tau-1})\, dX_\tau}$$

τ: planning time – dynamic programming runs



t-1 $\quad$ t: time at which action A_t is chosen & applied

t: real time in which knowledge K_{t-1} is collected

Passive and Active Approximations ad cause 1

- The crudest but common approximation of $(\mathsf{f}(X_\tau|A_\tau, \mathbf{K}_{\tau-1}))_{\tau=t}^{h}$ freezes the current knowledge realisation $\mathbf{K}_{\tau-1} = \mathbf{K}_{t-1}$ and it is:
- *passive approximation* selects for $\tau \in [t, h]$

$$\mathsf{f}(X_\tau|A_\tau, \mathbf{K}_{\tau-1}) \approx \hat{\mathsf{f}}(X_\tau|A_\tau, \mathbf{K}_{\tau-1}) = \hat{\mathsf{f}}(X_\tau|A_\tau, \mathbf{K}_{t-1}). \tag{185}$$

  It predicts with the learning stopped while neglecting the action influence on the future knowledge. The antonymous notion is:
- *active approximation* predicts with evolving knowledge. It models influence of actions on knowledge, on future information state.

- *certainty-equivalence approximation* gets the approximate predictive pd by inserting a point estimate $\hat{X}_{\tau|\tau-1}$ of X_τ in the observation model ↳

$$\begin{aligned} f(\Delta_\tau|A_\tau,\mathbf{K}_{\tau-1}) &\approx f(\Delta_\tau|\hat{X}_{\tau|\tau-1},A_\tau,\mathbf{K}_{\tau-1}), \text{ implied by} \qquad (186)\\ f(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) &\approx \hat{f}(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) \equiv \delta(X_\tau,\hat{X}_{\tau|\tau-1}) = \text{Dirac delta}↳. \end{aligned}$$

  The second index of the point estimate $\hat{X}_{\tau|\tau-1}$ means that it is based on A_τ and knowledge↳ $\mathbf{K}_{\tau-1}$, i.e. (186) is active approximation↳.

- The most spread one is passive certainty-equivalence approximation

$$f(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) \approx \hat{f}(X_\tau|A_\tau,\mathbf{K}_{\tau-1}) \equiv \delta(X_\tau,\hat{X}_{\tau|t-1}), \text{ where} \qquad (187)$$

  $\hat{X}_{\tau|t-1}$ is a point estimate of X_τ uses A_τ and "frozen" $\mathbf{K}_{t-1}$ only.

- The certainty-equivalence approximations are legitimate if the pds $f(X_\tau|A_\tau, \mathbf{K}_{\tau-1})$ or $f(X_\tau|A_\tau, \mathbf{K}_{t-1})$ concentrate near $\hat{X}_{\tau|\tau-1}$ or $\hat{X}_{\tau|t-1}$.

Otherwise, information about uncertainty of point estimates $\hat{X}_{\tau|\tau-1}$ or $\hat{X}_{\tau|t-1}$, say, covariance matrices $\mathbb{C}_{\tau|\tau-1}$ or $\mathbb{C}_{\tau|t-1}$ are to be used in

- *cautious approximation* uses the point estimate $\hat{X}_{\tau|\tau-1}$ of unknown X_τ as well as the uncertainty description $\mathbb{C}_{\tau|\tau-1}$

$$f(X_\tau|A_\tau, \mathbf{K}_{\tau-1}) \approx \hat{f}(X_\tau|A_\tau, \hat{X}_{\tau|\tau-1}, \mathbb{C}_{\tau|\tau-1}), \ \tau = t, \ldots, h. \quad (188)$$

- *super-cautious approximation* is a passive cautious approximation with

$$f(X_\tau|A_\tau, \mathbf{K}_{\tau-1}) \approx \hat{f}(X_\tau|A_\tau, \hat{X}_{\tau|t-1}, \mathbb{C}_{\tau|t-1}), \ \tau = t, \ldots, h. \quad (189)$$

The name reflects the pessimism about future learning abilities but to be super-cautious may be dangerous as it may preserve a bad strategy!

Passive approximations prevail as good active ones are too complex.

- The passivity may provably result in a completely bad performance for controlled Markov-chain models (Markov decision process), [114].

  Exploration can be so poor that the posterior pd concentrates on completely wrong parameter estimates.

- Systematic attempts to solve this difficult problem are rare, [44]. Common ways of supporting exploration are
  - A term reflecting learning quality even under a passive-type design is added to the original loss, [61].

    It is numerically demanding and sensitive to the added-term weighting.
  - An external stimulating signal is fed into the closed DM loop.

    This improves learning at the cost of deteriorating the achievable quality.

- Randomised strategy resulting from FPD is conjectured to add the adequate exploratory constituent into the resulting actions.

Simplification of the optimisation space follows two combined directions:

- The influence of a long design horizon ↳ on its complexity is suppressed.
- The value function ↳ is parameterised and its estimation employed.

Remark 30 (FPD Opens New Ways)

The use of FPD ↳ opens a new opportunity to approximate only expectation instead of the standard (minimisation, expectation). This fact is, for instance, seen on Example 10.

- The reduction of the design horizon obviously simplifies the design.

  The reduction obtained by planning just one-step-ahead (a.k.a. greedy or myopic optimisation) has been popular for a long time [131].

  Troubles of Greedy Optimisation

- Dynamic decision making means that consequences of an action are encountered far beyond the time moment of its use.

  Consequently, the action that is optimal when judged from a shortsighted perspective might be quite bad in long terms, [83].

$$\Downarrow$$

This call for a compromise between the planning over the whole horizon of interest & simplicity of short-sighted, locally optimising, strategies.

One vs. Infinity Steps-Ahead LQ Designs [83] ad cause 2

- Systems $\Delta_t \sim \mathrm{N}_{\Delta_t}(x_{t-1} + b_0 A_t, 1)$ with real observation Δ_t, action A_t and $x_t \equiv a\Delta_t + b_1 A_t$. The involved real coefficients are

$$a = 0.99, \quad b_0,\ b_1 \ \text{ are given by } \ \rho = \frac{b_1}{b_0} \quad \wedge \quad \text{static gain } \ \frac{b_0 + b_1}{1-a} = 1.$$

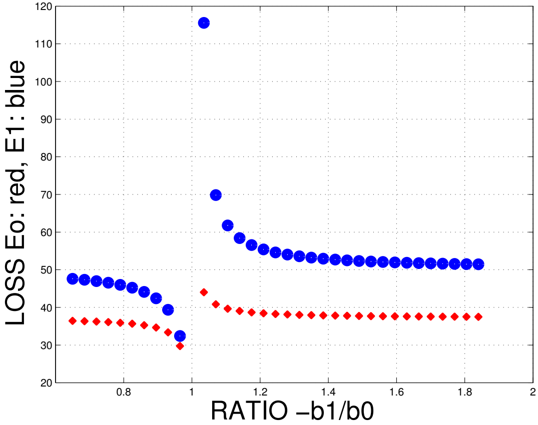
- The expected quadratic loss with horizon $h \to \infty$ quantifies DM aim

$$\mathrm{E}[\mathrm{Z}|\rho] = \lim_{h\to\infty} \frac{1}{h} \mathrm{E}\Big[\sum_{1\le t\le h} (\Delta_t^2 + qA_t^2)|\rho\Big], \quad q = 0.5$$

- Stationary strategies S^1 made of decision rules $A_t = -\frac{b_0}{b_0^2+q} x_{t-1}$ minimising one-step-ahead expected loss $\mathrm{E}\left[\Delta_t^2 + qA_t^2|\rho\right]$ serve for demonstration of adverse effect of short design horizon.

- Optimal stationary strategies S^o minimising the expected loss $\mathrm{E}[\mathrm{Z}]$ have the decision rules $A_t = -\mathbb{L}'[\Delta_{t-1}, A_{t-1}]'$
- Fig. 318 shows dependence of $\mathrm{E}_1 = \mathrm{E}_{\mathrm{S}^1}[\mathrm{Z}|\rho]$ & $\mathrm{E}_o = \mathrm{E}_{\mathrm{S}^o}[\mathrm{Z}|\rho]$ on ρ.

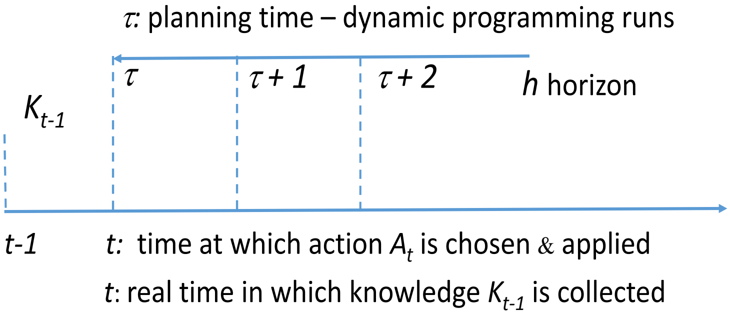
Expected Losses of Optimal and One-Step-Ahead Design



Evaluation of decision rule at time t : $\gamma(\mathbf{K}_h) = 1$

$$\gamma(\mathbf{K}_{\tau-1}) \equiv \int_{A_\tau^\star} {}^{\mathsf{I}}\mathsf{f}(A_\tau|\mathbf{K}_{\tau-1}) \exp[-\omega(A_\tau, \mathbf{K}_{t-1})] \,\mathrm{d}A_\tau \quad \text{if } \tau < h$$

$$\omega(A_\tau, \mathbf{K}_{\tau-1}) \equiv \int_{\Delta_\tau^\star} \mathsf{f}(\Delta_\tau|A_\tau, \mathbf{K}_{\tau-1}) \ln\left(\frac{\mathsf{f}(\Delta_\tau|A_\tau, \mathbf{K}_{\tau-1})}{\gamma(\mathbf{K}_\tau)\,{}^{\mathsf{I}}\mathsf{f}(\Delta_\tau|A_\tau, \mathbf{K}_{\tau-1})}\right) d\Delta_\tau.$$



A suboptimal design only approximates the optimal one. This calls for a redesign after using the initial planned actions. This motivates

- *receding horizon strategy*, which at time t
  - performs the design looking T step ahead, with a small T such that the main dynamic consequences of the action A_t are covered
  - applies the action A_t resulting from this design for the knowledge $\mathbf{K}_{t-1}$
  - acquires the new data record $D_t = (\Delta_t, A_t)$ =(observation,action)
  - updates the knowledge $\mathbf{K}_t = (\Delta_t, A_t, \mathbf{K}_{t-1})$ & performs a learning step
  - increases time and repeats all above steps.

Remark 31 (Common Versions; Model Predictive Control)

Mostly, a passive approximation of the models is used. In an extreme, the widely-spread variant, known as model-based predictive design even runs without the learning while neglecting the future uncertainty. Its wide applicability stems from its ability to cope well with nonlinear systems and hard bounds on the predicted behaviour, [27, 30, 31, 123, 133, 147].

- Dynamic programming nd FPD, Props. 11, 23 are one step ahead designs if the value function↳ is known. This motivates:
- *approximate dynamic programming*, which performs greedy design with an estimated value function↳ [149].

Example 19 (Approximate Dynamic Programming for FPD)

The FPD↳ with information state↳ $\mathcal{X}_t$ may approximate $\gamma(\mathcal{X}_t)$, $\omega(A_t, \mathcal{X}_{t-1})$ by $\gamma(\Theta, \mathcal{X}_t)$, $\omega(\Theta, A_t, \mathcal{X}_{t-1})$ given by a finite dimensional parameter $\Theta \in \Theta^\star$. They should to fulfil, Prop. 23,

$$
\begin{aligned}
\gamma(\Theta, \mathcal{X}_{t-1}) &= \int_{A_t^\star} {}^{\mathsf{I}}\mathsf{f}(A_t|\mathcal{X}_{t-1}) \exp[-\omega(\Theta, A_t, \mathcal{X}_{t-1})] \,\mathrm{d}A_t \qquad (190) \\
\omega(\Theta, A_t, \mathcal{X}_{t-1}) &= \Omega(A_t, \mathcal{X}_{t-1}) - \int_{\mathcal{X}_t^\star} \mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1}) \ln(\gamma(\Theta, \mathcal{X}_t)) \,\mathrm{d}\mathcal{X}_t \\
\Omega(A_t, \mathcal{X}_{t-1}) &\equiv \int_{\mathcal{X}_t^\star} \mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1}) \ln\left(\frac{\mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1})}{{}^{\mathsf{I}}\mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1})}\right) \mathrm{d}\mathcal{X}_t,
\end{aligned}
$$

solvable by successive approximations with Monte Carlo integration.

The need to store multivariate functions limit this way!

The definition of zero-mean, mutually uncorrelated innovations ↳ $\ln(e)$

$$\ln(e(\Theta, A_t, \mathcal{X}_{t-1})) \equiv \ln(\gamma(\Theta, \mathcal{X}_t)) - \int_{\mathcal{X}_t^\star} \mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1}) \ln(\gamma(\Theta, \mathcal{X}_t)) \,\mathrm{d}\mathcal{X}_t$$

gives the regression model (191)

$$\gamma(\Theta, \mathcal{X}_t) = \gamma(\Theta, \mathcal{X}_{t-1}) - \underbrace{\int_{A_t^\star} {}^{\mathsf{I}}\mathsf{f}(A_t|\mathcal{X}_{t-1}) \exp[-\Omega(A_t, \mathcal{X}_{t-1})] \,\mathrm{d}A_t}_{\text{a function given by } {}^{\mathsf{I}}\mathsf{f}(\mathcal{X}_t, A_t|\mathcal{X}_{t-1}), \mathsf{f}(\mathcal{X}_t|A_t, \mathcal{X}_{t-1})}$$

$$+ \underbrace{\int_{A_t^\star} {}^{\mathsf{I}}\mathsf{f}(A_t|\mathcal{X}_{t-1}) e(\Theta, A_t, \mathcal{X}_{t-1}) \,\mathrm{d}A_t}_{\varepsilon(\Theta, \mathcal{X}_{t-1})}$$

Knowledge (191) leads to pd of $\varepsilon(\Theta, \mathcal{X}_{t-1})$ via minimum KLD principle ↳.

- The approximation of the value function ↳ is nonlinear estimation task. The gained randomised FPD strategy makes the learning active.

  A lot is to be done to convert first attempts [72] into a solid solution.

Splitting of the DM task in a chained subtasks converts the optimal design ↳ into a practically optimal design ↳. The experience recommends

- *golden DM rule*: a departure from optimality is the last option.
- The design, as any human activity, is iterative. The majority of trials has to be concentrated in the off-line phase. This minimises expenses related to the commission of the decision strategy ↳.
- The lack of the formal tools for the decomposition forces us to use empirical rules as in a long-term project DESIGNER [8, 82, 92].
- The presented decomposition, which arisen from DESIGNER, concerns adaptive control exploiting parameter estimation ↳.

The subtasks chain is iteratively solved, until the decision maker is happy.

- Formalise the DM problem and collect the available knowledge. Gain the specification of
  - technical decision aims,
  - the system,
  - the available data, actions and observations,
  - the technologic and complexity restrictions.
- Select the set of parametric models & complete behaviour definition.
- Check data-informativeness. If possible, perform experimental design and collect informative data, e.g., [170].
- Pre-process data, i.e. remove outliers, complete missing records, suppress noise, regularise sampling rate,. . . , [101, 129].
- Elicit (quantify) prior knowledge [67, 76, 88, 92].
- Estimate model structure & decision period [13, 60, 69, 90, 92, 117].

Off-Line Phase of Adaptive-DM Design ad cause 3

- Perform parameter estimation, Sec. 14, and estimate forgetting factor, Sec. 31, both using elicited knowledge and off-line data.

  The posterior pd$_{\lrcorner}$ serves as the prior & alternative pd in on-line phase.
- Validate the model quality on validation data, e.g. [134]. Use Bayesian testing of hypotheses, Sec. 25, tailored to dynamic systems and exploiting all off-line data [93].
- Elicit preferences to get the ideal pd$_{\lrcorner}$ or performance index$_{\lrcorner}$, [78].

  Do until the results cannot be improved
  - Select the type of the suboptimal design and its parameters.
  - Perform prior design of the strategy.
  - Predict the closed-loop behaviour [82, 86, 126].
  - Compare the results with decision maker's preferences. □

Data collected in on-line phase may improve future off-line design.

On-Line Phase of Adaptive-DM Design

The following DM subtasks are solved in real time, for $t \in t^\star$.

- Collect and pre-process a data record.
- Generate reference signals followed by the controlled behaviour↳ part.
- Apply data updating↳ and emulate time-updating by forgetting↳.
- Use the chosen suboptimal design, e.g., receding horizon strategy↳.
- Propose actions using the designed strategy↳ and pre-processed data.
- Counteract a violation of constraints on actions by optimised cutting.
- Check and counteract possible discrepancies like an extreme difference of predicted and observed behaviour characteristic by a re-tuning of optional parameters of the parameter estimation↳, filtering↳ & design↳.

No freedom exists for iterative trial-and-error solutions in on-line DM!

A premature stoping stimulated by safety violation is expensive!

Samples of Omitted Topics

- Stationary additive designs leading to linear programming [40, 94].

$$\color{blue}{\mathsf{f}(\mathcal{X},A)} = \int_{\mathcal{X}^\star,A^\star} \overbrace{\mathsf{f}(\mathcal{X}|A,\bar{\mathcal{X}})}^{\text{system m.}} \times \overbrace{\mathsf{f}(A|\bar{\mathcal{X}})}^{\text{DM rule}} \times \overbrace{\color{blue}{\mathsf{f}(\bar{\mathcal{X}},\bar{A})}}^{\text{stationary pd}} \,\mathrm{d}\bar{\mathcal{X}}\,\mathrm{d}\bar{A}$$

$$^{o}\mathsf{f}(A|\bar{\mathcal{X}}) \in \operatorname{Arg} \min_{\substack{\{\mathsf{f}(A|\bar{\mathcal{X}}) \geq 0 \\ \int_{A^\star} \mathsf{f}(A|\bar{\mathcal{X}})\,\mathrm{d}A = 1\}}} \int_{\mathcal{X}^\star,A^\star} \underbrace{\mathsf{z}(\mathcal{X}^\star,A^\star)}_{\text{partial loss}} \color{blue}{\mathsf{f}(\mathcal{X},A)}\,\mathrm{d}\mathcal{X}\,\mathrm{d}A$$

- Handling of constrained actions requiring nonlinear [17], typically quadratic programming [21].
- General approach to learning of performance index [78]
- Some technical details of the decomposed DM are postponed after solving some auxiliary DM tasks.

DM Elements & Related Tasks

Aim and Presentation Way

Presented technical details should help in applications of the DM theory.

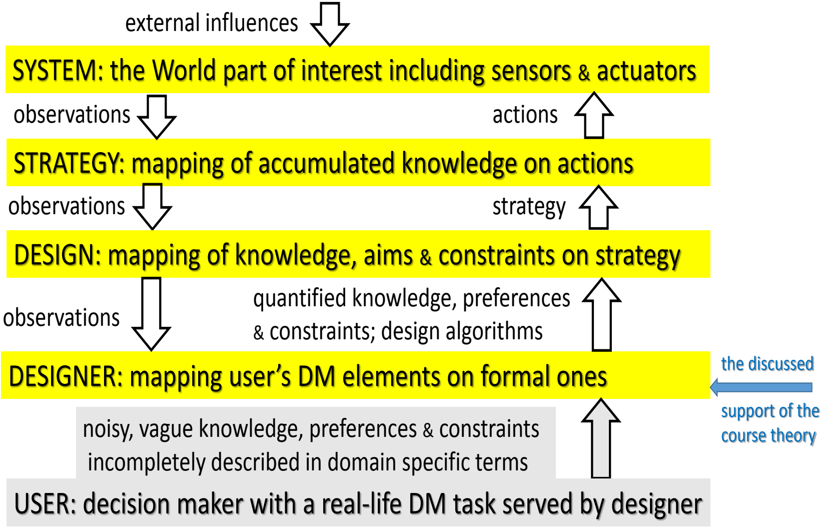
- The discussion of DM elements↳, "atoms" of any decision task, should help us to learn a good practice and to avoid common mistakes.
- The choice, modification and use DM elements↳ are specific decision subtasks that have to be harmonised with the final aim↳ considered.

  It is often very hard work but the golden DM rule↳ is to be respected.
- The presentation describes the design↳ sequentially.

  The interrelations of respective DM elements↳ imply that respective steps are mostly performed in parallel and in an iterative manner.

For instance, the behaviour↳ choice cannot be separated from the DM aim.

Discussion Concerns the Bottom Part of DM Scenario



DM Elements for Optimising PDs $(\mathsf{f}(A_t|X^{t-1},\mathbf{K}_{t-1}) = \mathsf{f}(A_t|\mathbf{K}_{t-1}))_{t\in t^\star}$

1. Behaviour structured $B = [X^h, D^h] = [X^h, \Delta^h, A^h], \quad B \in B^\star$,
   B = [hidden quantity$_t$, observation$_t$, action$_t$]$_{t\in t^\star}$ up to a horizon h
   $B \in B^\star = [X^{h\star}, D^{h\star}] = [X^{h\star}, \Delta^{h\star}, A^{h\star}]$; knowledge $\mathbf{K}_t = (D_t, \mathbf{K}_{t-1})$.
2. Observation model $\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})$ and time-evolution model $\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})$ model of the considered system.
3. Prior pd $\mathsf{f}(X_0, \mathbf{K}_0|A_1) = \mathsf{f}(X_0|\mathbf{K}_0)\mathsf{f}(\mathbf{K}_0)$ describing prior knowledge $\mathbf{K}_0$.
4. Ideal observation ${}^{\mathsf{I}}\mathsf{f}(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})$, strategy ${}^{\mathsf{I}}\mathsf{f}(A_t|X^{t-1}, \mathbf{K}_{t-1})$ and time evolution ${}^{\mathsf{I}}\mathsf{f}(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})$ pds describing DM preferences.

Specific and Derived DM Elements

- Ideal prior pd ${}^{\mathsf{I}}\mathsf{f}(X_0, \mathbf{K}_0)$ does not influence optimisation and thus the leave to the fate option can be always used ${}^{\mathsf{I}}\mathsf{f}(X_0, \mathbf{K}_0) = \mathsf{f}(X_0, \mathbf{K}_0)$.
- Prop. 16 provides the predictor $\mathsf{f}(\Delta|A_t, \mathbf{K}_{t-1})$ and the pd $\mathsf{f}(X_t|\mathbf{K}_t)$ needed for design with a hidden quantity, Prop. 11.

DM ↳ tries to influence closed-loop behaviour ↳ $B \in B^\star$ in harmony with user's preferential ordering ↳ $\preccurlyeq_{B^\star}$.

The designer ↳ has to delimit behaviour ↳ consisting of a sequence of quantities up to optional horizon ↳ h. The t-th behaviour entry is

- available optional action ↳ $A_t \in A_t^\star$
- accessible and potentially useful observation ↳ $\Delta_t \in \Delta_t^\star$
- considered hidden quantity ↳ $X_t \in X_t^\star$.

The delimitation of the behaviour $B \in B^\star$ consists of the choice of the included quantities and their expected as well as desired ranges.

This is a specific DM task exploiting tools as hypotheses testing, Example 18.

Actions

1st input

Availability of an optional action↳ is inevitable precondition of DM↳.

- The domain-specific properties, physical, economical, medical etc., and their knowledge drive the initial choice of possible actions.

  Sometimes, the nature of DM determines possible actions uniquely.

- Often, alternative actions are available: none should be a priori discarded.

  Formally, an optimum on a set dominates that found on a subset. Practically, e.g., think about rewards and punishments in education.

- Freedom in actions determination needs selecting the most suitable ones.

  Bayesian hypotheses testing, Sec. 25, leading to Bayesian structure estimation↳, Prop. 33, serves to the final action selection if the expert-based choice is non-unique or questionable.

  The set of hypothesis has to be specified by knowledgable expert!

An Example of a Knowledgeable Action Choice

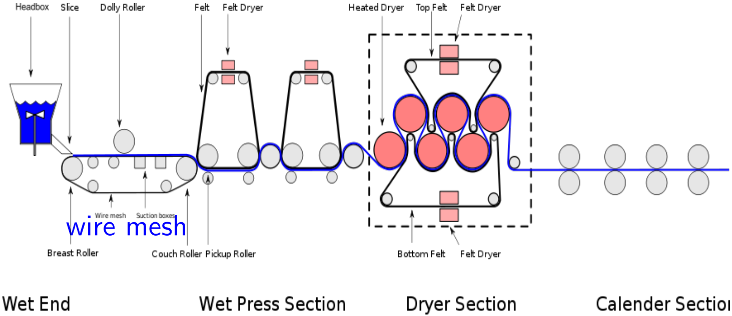


Figure : Paper machine scheme

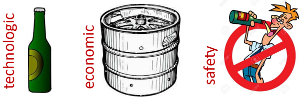
- DM aim to keep the paper wetness in a given range.
- Standard action – variations of energy supply in Dryer Section.
- Wisely chosen action – variations of the wire-mesh speed – spares energy and improves controllability of the system.

Action Domain $(A_t^\star)_{t\in t^\star}$ – Constraints on Actions 1st input

The main source of constraints motivates their name

- *technological constraints* coincide with the action set $(A_t^\star)_{t\in t^\star}$.

Constraints are implied by technologic, economic or safety considerations



Warnings

- Active technological constraints increase complexity of the design.

  They are often relaxed and respected via other DM elements:

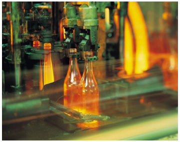
  The performance index enforces a penalty-based optimisation [57] or un-constrained optimal actions are designed and projected on $(A_t^\star)_{t\in t^\star}$.
- A simple clipping at the boundary of $(A_t^\star)_{t\in t^\star}$ should not be used.
- The applied, not the designed, action is to be included into knowledge!

Remarks on Constraints

1st input

- Unnecessary constraints should be avoided. For instance,
  - the restriction to unbiased estimators [136] is often meaningful, but their use in adaptive control [7] decreases the final quality.
  - human-perception-tailored, entry-wise stimulation of a learnt system with a vector action makes learning unnecessarily long.

A joint random excitation of all burners shortened experiments when estimating dynamics of a bottle-producing glass furnace.



- The options related to constraints are mostly done by decision maker.

  Again, it confirms that the theoretical and algorithmic support of the mapping of the practical DM on the formalised one is underdeveloped.

Observation

1st input

Observation connects the theoretical world of the design with reality.

- The strategy acting on more observations is potentially better than that using a narrower set: no observation should be a priori discarded.

  Bayesian hypotheses testing, Sec. 25, helps recognition of whether the information gain brought by an observation entry is over-weighted by the increased approximation error connected with its consideration.

Observations split in

- indicators of decision quality determining preferential ordering
- auxiliary, information bringing, observations: the leave to the fate option is applied to them when selecting ideal pd
- realised actions informing on implementation imprecisions
- *external quantity*, an observation completely uninfluenced by action

External quantities cannot be perceived as actions! Wide-spread violations of this rule hide the need for their prediction in dynamic DM.

- The observation domain is mostly given by its nature & sensors.
- A less rich observation domain potentially decreases modelling errors. This implies good, often implicit, practice: highly improbable realisations are excluded from the considered observation domain.

Examples of Alternative $\Delta^\star$

- coin tossing experiment $\Delta^\star = \{head, tail\}$ vs. $\Delta^\star = \{head, tail, edge\}$
- earthquake degree in Bohemia $\Delta^\star = \{1, \ldots, 5\}$ vs. $\Delta^\star = \{1, \ldots, 9\}$.

Remark 32 (On Desired Observation Ranges)

The observations are not, by their definition, fully determined by the given knowledge realisation and the action chosen.

Thus, wishes on the desired observation ranges cannot be voluntaristic!

- Hidden quantities primarily arise from the wish to influence them.

  Stabilisation of unmeasured temperature by manipulating fuel flow within a large kiln represents a typical example of this type.
- The system modelling is the second "generator" of hidden quantities.

Relations describing the system ↳ are modelled via two complementary ways

- First principles – conservation laws – that relate quantities with a clear physical meaning, which are measured indirectly by imperfect sensors.

  The meaning usually determines well ranges of hidden quantities.
- "Universal" approximation property, [56], the ability to approximate any relation within the behaviour by a member of a rich model class.

  The involved hidden quantity ↳ usually lacks interpretation, which makes the information about its range vague and quite wide.

- Determination of ranges of behaviour ↳ is an integral part of modelling ↳.
- A broader choice improves DM quality but increases modelling errors.
- A priori specified range of unknown parameters is often wide. The prior pd effectively suppresses it as the support of the posterior pd ↳ is in its support, Sec. 14. This fact enhances importance of
- *knowledge elicitation*, which transforms knowledge into the prior pd ↳.
- The design ↳ has to provide causal decision rule ↳s. This constraint is respected by employing dynamic programming, Prop. 11.

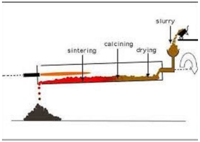
  This relies on clear distinction between observed and hidden quantities, which can be influenced by the choice of sensors.

Need for Active Data Acquisition 1st input

Data connects evaluations to reality ⇒ the information content of data is crucial for the success of the DM ↳ that use them. Ideally, the data acquisition should use

- *experimental design* that selects strategy ↳ used during the data acquisition so that data is as informative as possible.
- The experimental design ↳ should suppress the ambiguity of the best projection caused by data quality, Prop. 21.
- Practically, the optional system inputs, set points etc. have to "excite" sufficiently the inspected system ↳.

  For instance, learning of the dependence of the temperature within a kiln on heating failed as this action was fixed during the data acquisition



- The system ↳ stimulation during the data acquisition influences also rate with which the filtering ↳ or parameter estimation ↳ reach good results.

The need for experimental design arises if hidden quantities $X^h \in X^{h\star} \neq \emptyset$ are in behaviour$_\lrcorner$ $B \in B^\star = [X^{h\star}, D^{h\star}] = [X^{h\star}, A^{h\star}, \Delta^{h\star}]$.

- This DM is driven by the wish to be as far as possible from the worst case in which date pairs D^h are independent of hidden quantities

$$^{\neg\mathsf{I}}\mathsf{f}(X^h, D^h) = {}^{\neg\mathsf{I}}\mathsf{f}(X^h, A^h, \Delta^h) \equiv \mathsf{f}_\mathsf{S}(X^h|A^h)\mathsf{f}_\mathsf{S}(D^h). \tag{192}$$

- An inspection of FPD$_\lrcorner$ axiomatisations implies that in this case the KLD$_\lrcorner$ of the closed-loop model$_\lrcorner$ $\mathsf{f}_\mathsf{S}(X^h, D^h)$ to ${}^{\neg\mathsf{I}}\mathsf{f}(X^h, D^h)$ is to be maximised. Thus, FPD-based experimental design maximises
- *mutual information* between X^h and data pairs D^h

$$\mathsf{D}(\mathsf{f}_\mathsf{S}||\,{}^{\neg\mathsf{I}}\mathsf{f}) = \int_{(X^h,D^h)^\star} \mathsf{f}_\mathsf{S}(X^h, D^h) \ln\left(\frac{\mathsf{f}_\mathsf{S}(X^h, D^h)}{\mathsf{f}_\mathsf{S}(X^h|A^h)\mathsf{f}_\mathsf{S}(D^h)}\right) \mathrm{d}(X^h, D^h) \tag{193}$$

- The formulated optimisation is as complex as the general FPD and faces the same problems as discussed in Section 32.

  A specific simplification can be made by narrowing the set of competitive strategies even to a finite collection used within classical and well-established framework [128, 129].

- Feasible solutions for specific classes of models and strategies are well elaborated see, e.g. the classical reference [170] or survey [47].

- The generic maximisation of the mutual information gives infeasible actions without constraining the support of $\mathsf{f}_{\mathsf{S}}(D^h)$ in (192).

Remarks On Experimental Design

1st input

- The proposed experimental design is non-standard but fits the Bayesian treatment of unknown quantities and FPD ↳.
- This experimental design ↳ neither supposes existence of "true" hidden quantities nor the knowledge of their values describing the best projection ↳ to the set of system models.
- The practical elaboration of the experimental design would be invaluable in many applied domains.
- The mutual information should also be used for analysing data when parameter estimation ↳ results are unsatisfactory.

Data Pre-Processing: the Need for It

1st input

- *pre-processing* maps raw data on pre-processed data used in DM ↳.
- Data pre-processing adds a dynamic mapping into the closed loop so that its common use seems to be illogical and harmful.

  Inevitably approximate DM makes data pre-processing meaningful as
  - the objective pd ↳ is always out of the parametric model ↳ s'set;
  - the parameter estimation ↳ searches for the best projection ↳ of the objective pd ↳ that describes all relations of the behaviour ↳ reflected in measured data;
  - The projector has no information about significance of these relations with respect to the solved DM task.

  $$\Downarrow$$

  Data pre-processing ↳ is to suppress insignificant relations so that adverse influence of the additional dynamics is counteracted by an improved modelling of relations important for the DM task.
- DM tasks operating with the same data and the relevance of modelled relations vary $\Rightarrow$ do not discard raw data!

Typical Data Pre-Processing Procedures 1st input

- *data transformation* linearises nonlinear data relations reflecting their physical nature or sensor properties.
  This enables use of well-supported linear parametric model ↳.
- *data scaling* realises affine data transformation.
  This allows standardisation of priors and suppresses numerical errors.
- *outliers' suppression* removes or cuts outlying observations.
  This makes the system model closer to the well-supported normal pd, whose learning is sensitive to outliers.
- *noise suppression* removes data constituents, typically of a high frequency, reflecting more sensor behaviour than the system dynamics.
- *missing data treatment* substitutes missing data by a guess.
  This preserves informative data by filling gaps in data sequences.
- *re-sampling* standardises sampling rate of the pre-processed data.
  It uses a high frequency of the data acquisition for noise suppression and removes sampling-induced variations of modelled relations.

Common Pre-Processing Mistakes

1st input

Pre-processing significantly influences quality of the resulting projection of the objective pd ↳ on the set of parametric models and thus whole DM.

- Damages made in the pre-processing ↳ phase can hardly be removed in later design phases. Typical errors in pre-processing ↳ are
- a loss of relevant information by a premature reduction of data
- wasting of information by low-frequency sampling of acquired data
- a significant change of the modelled dynamics by the pre-processing block: for instance, introduction too high transportation delay
- a distortion of inspected relations by a bad replacing of missing data.

- Similarly as other sub-tasks, the optimal pre-processing requires – mostly impossible – solution of the overall DM task to which it serves.
- Even the splitting of the pre-processing into adequately chosen and harmonised subtasks is left to a “sound” reasoning.
- The use of “sound” reasoning is pleasant as it requires creativity. It is, however, dangerous as the final DM results may be spoiled by an improper pre-processing choice, which may look nice on its own.
- The problem of the adequate choice of pre-processing is especially difficult in dynamic design ↳ in which there is a restricted freedom for an iterative trial-and-error treatment.

Ways to Model Building

2nd input

Relations describing the system ↳ are modelled via two complementary ways

- first principles – conservation laws – that relate quantities with a clear physical meaning, which are measured indirectly by imperfect sensors.

  It relies on substantial domain specific knowledge provided by experts.

  For instance, the kinematic state of a space shuttle is its position, speed and acceleration while the position is imprecisely measured.

- *universal approximation* which is an expansion with the ability to approximate relations in behaviour by a member of a rich model class.

  The rich class is spanned on a functional basis dense within the class of modelled mappings. Neural-nets community casts [56] this name.

  This choice relies on substantial knowledge of approximation spaces provided by mathematical experts.

The both modelling ways are usually combined and lead to

- *grey box* modelling collects theoretically expected relations between quantities and extend them into the probabilistic parametric model.

  Unknown constants or states always enter the final model.
- The usually deterministic models are extended by minimum KLD principle, Prop. 28, adding often unknown parameters.
- The resulting model is often too complex for further treatment and is to approximated via approximation principle, see Sec. 20.

  The dynamic exponential family provides good candidates of approximating pds as it converts functional recursive estimation into algebraic one, see Prop. 36.
- Bayesian testing of hypotheses, Sec. 25, serves for the choice of the best model among inspected model variants.

White and Black Box Modelling

2nd input

Models obtained solely by first principles are called white box models.

The white or grey box↳ modelling can be either impossible due to the lack of domain knowledge or can lead to unmanageable models approximation of which looses all advantages they bring. Then, we rely on

- *black box* modelling approximates the modelled functional relations by expanding them into a suitable functional basis guaranteeing universal approximation↳ property.
- Usually, the expansion to basis terms applies to moments of the modelled pd. Obviously, an expansion of the pd itself is more complete and systematic and sometimes even simpler [66].
- Finite mixtures of pds are often used as the universally approximating class. They are well-elaborated tool for black box↳ modelling [77, 159], which exploits that any smooth pd↳ can be approximated by a finite mixture↳ of properly parameterised, typically normal, pds.

This prominent black box model is outlined in connection with modelling of a data vector $\Psi'_t = [\Delta'_t, \psi'_t]$ =[observation, regression vector].

- *finite mixture* is the parametric pd of the form

$$f(\Psi_t|\Theta) = \sum_{c \in c^\star} \alpha_c f(\Psi_t|\Theta_c), \; c^\star = \{1, \ldots, |c^\star|\}, \; |c^\star| < \infty, \text{ given by} \quad (194)$$

- *component* , the pd $f(\Psi_t|\Theta_c)$, which is typically (not inevitably) a member of exponential family, and component parameter Θ_c
- *component weight* α_c, whose collection $\alpha = (\alpha_c)_{c \in c^\star}$ has properties of the pd of an unobserved random pointer with values in $c^\star$.
- *pointer to the component* $C_t \in c^\star$, $f(C_t = c|\Theta) = \alpha_c$. C_t is a hidden quantity within the modelled part of the behaviour

$$(C_t, \Theta = (\Theta_c, \alpha_c)_{c \in c^\star}, \text{observed quantities forming } \Psi_t) \quad (195)$$

- Let a pd $\mathsf{f}(\Psi)$ have a compact Hausdorff domain [24]. Then, it can be approximated by a piece-wise function

$$\mathsf{f}(\Psi) \approx \sum_{c=1}^{\infty} \mathsf{f}(\tilde{\Psi}_c)\mathrm{vol_c}\frac{\chi_{\Psi_{\star c}}(\Psi)}{\mathrm{vol_c}}. \tag{196}$$

  There, the indicator $\chi_{\Psi_{\star c}}(\Psi)$ delimits a small neighbourhood $\Psi_{\star c}$ of the grid point $\tilde{\Psi}_c$ with volume $\mathrm{vol_c} = \int_{\Psi^\star} \chi_{\Psi_{\star c}}(\Psi)\,\mathrm{d}\Psi$.
- (196) is countable mixture of uniform pds. Its non-negative weights $\mathsf{f}(\tilde{\Psi}_c)\mathrm{vol_c}$ have to fall to zero as $\int_{\Psi^\star} \mathsf{f}(\Psi)\,\mathrm{d}\Psi \approx \sum_{c=1}^{\infty} \mathsf{f}(\tilde{\Psi}_c)\mathrm{vol_c} = 1$. Thus, a finite mixture of uniform pds approximates $\mathsf{f}(\Psi)$ arbitrarily well.
- Indicators of the exploited decomposition of unity [163] can be approximated by other, even infinitely smooth, non-negative functions having finite integral. They provide other, say normal, basis for creating mixtures and allow to relax the compactness assumption.

- The mixture model (194) describes only a part of the behaviour. Additional assumptions are needed to get its complete parametric description. The wide-spread modelling deals with
- *classic mixture* , which assumes data vectors Ψ_t independent when conditioned on Θ (195), [159].
- The corresponding likelihood

$$\mathsf{L}(\Theta, \mathbf{K}_t) \equiv \prod_{\tau \le t} \mathsf{f}(\Psi_\tau|\Theta) = \prod_{\tau \le t} \sum_{c \in c^\star} \alpha_c \mathsf{f}(\Psi_\tau|\Theta_c) \qquad (197)$$

  is the sum of 2^t different functions of Θ, which causes an extreme complexity of the exact Bayesian estimation.
- The complexity is faced by several ways commented later on.

- The finite mixture (194) generally induces the parametric model

$$\mathsf{f}(\Delta_t|\Theta,\psi_t) \equiv \mathsf{m}(\Psi,\Theta) = \frac{\sum_{c\in c^\star} \alpha_c \mathsf{f}(\Psi_t|\Theta_c)}{\sum_{c\in c^\star} \alpha_c \underbrace{\int_{\Delta_t^\star} \mathsf{f}(\Psi_t|\Theta_c)\,\mathrm{d}\Delta_t}_{\mathsf{f}(\psi_t|\Theta_c)}} \tag{198}$$

$$= \sum_{c\in c^\star} \beta_c(\Theta,\psi_t) \underbrace{\frac{\mathsf{f}(\Psi_t|\Theta_c)}{\mathsf{f}(\psi_t|\Theta_c)}}_{\mathsf{f}(\Delta_t|\Theta_c,\psi_t)} = \sum_{c\in c^\star} \beta_c(\Theta,\psi_t)\mathsf{f}(\Delta_t|\Theta_c,\psi_t),$$

$$\beta_c(\Theta,\psi_t) = \alpha_c \frac{\mathsf{f}(\psi_t|\Theta_c)}{\sum_{c\in c^\star} \mathsf{f}(\psi_t|\Theta_c)}, \Psi_t = \overbrace{[\Delta_t', \quad \underbrace{\psi_t'}_{\textit{regression vector}} \quad]}^{\textit{data vector}}.$$

The finite mixture universally approximates even in the dependent case if the components weights $\beta_c(\Theta,\psi_t)$ may depend on the regression vector ψ_t.

- In fact, the model (198) is ratio of coupled finite mixtures.

- The *exact estimation and prediction* with the mixture model fails as the number of terms in the likelihood ↳ $\mathsf{L}(\Theta, \mathbf{K}_t)$ (197) blows up.
- Good approximations exist if component ↳s are from exponential family ↳.

The available estimation techniques clusters into the following groups

- search for point estimates maximising likelihood ↳, typically, via expectation-maximisation algorithm [36];
- approximation of the intractable prior pd ↳ by the product of approximate prior conjugate pds to respective component ↳s and by heuristic (quasi-Bayesian) [77, 87, 159] or KLD ↳-based projection (116) of the posterior pd on products of pds conjugate to components [4];
- formulation of the learning problem as filtering ↳, which explicitly estimates realisation of the pointer to the component ↳ C_t.

Prior PD

3rd input

- The need for selecting prior pd is often regarded as the main disadvantage of the adopted Bayesian approach.

  The lack of efficient, unambiguous & elicitation-expert independent [48], tools for knowledge elicitation ↳ can be blamed for prior-pd aversion.

- Here, we contribute positively to the never-ending discussion on pros and cons of exploiting prior pds by indicating that the prior "expert" knowledge can be fed into learning in a systematic way.
- The posterior pd ↳ (80) is a product of t factors formed by parametric models $\mathsf{f}(\Delta_\tau|\Theta, A_\tau, \mathbf{K}_{\tau-1})$ and of a single prior pd ↳ $\mathsf{f}(\Theta)$.
- If t is high & data is informative then the posterior pds gained for various prior pds resemble each other: the role of prior pd is weak [35].

  The posterior pd ↳ is significantly influenced by the prior pd when some of the above conditions is not fulfilled.

An automatic processing of different knowledge forms needs a common language expressing them in way serving to any parametric model. It is

- *fictitious data* which is an outcome gedanken experiment on the system.

Examples indicating the conjectured universality of this knowledge form are

- *imprecise data*, which includes
  - *obsolete data* which is measured on a non-identical system.
  - *simulation model*, created as blend of theoretical and empirical domain knowledge, produces fictitious data.
- *experimentally inspected characteristics* designed by experts like: step response – expected observations Δ^h after the unit-step action frequency characteristic – the expected-observation form, $t \in t^\star$, $\mathsf{E}[\Delta_t] = a(\omega)\sin(t\omega + \phi(\omega))$ after applying $A_t = \sin(\omega t)$), ω given
- *theoretical model* predicts consequences of given conditions.
- *expert's prediction* of the form if conditions then the response.

Precision of Fictitious Data – Babel Tower Recurs 3rd input

- Inherent imprecision is common to fictitious data. It must be expressed and respected in a common way to avoid Babel tower problem.

Always partial information on precision of fictitious data ↳ has many forms

- an expected response range
- moments type characteristics like covariance, confidence interval, etc.
- membership functions in fuzzy framework, e.g. [1]
- no precision is assigned to fictitious data, typically, to predictions or simulation data and data vectors are offered as deterministic ones.

  Imprecision is to be respected even in this case!

Prop. 40 adopts a common expression of imprecision of fictitious data ↳ and uses minimum KLD principle ↳ to convert fictitious data ↳ into pds. They are build in prior pd ↳ using Prop. (41).

Learning with imprecise fictitious data causes the same problems as parameter changes, Sec. 31, i.e. it is counteracted similarly.

Proposition 40 (Estimation with Imprecise Fictitious Data)

Let $\mathsf{f}_0(\Theta) = \mathsf{f}(\Theta|\mathbf{K})$ *be prior pd of parameter* $\Theta \in \Theta^\star$ *in the observation model* $\mathsf{f}(\Delta|\Theta, A, \mathbf{K})$. *The posterior pd* $\tilde{\mathsf{f}}(\Theta) \propto \mathsf{f}(\Delta|\Theta, A, \mathbf{K})\mathsf{f}_0(\Theta)$ *gained by using fictitious data in Bayes' rule differs from the unknown correct posterior pd* $\mathsf{f}(\Theta) = \mathsf{f}(\Theta|$*knowledge in fictitious data*$, \mathbf{K})$ *but approximates it*

$$\mathsf{D}(\mathsf{f}||\tilde{\mathsf{f}}) \leq \beta < \infty. \tag{199}$$

With $\mathsf{f}_0(\Theta)$ *being the available guess of the unknown correct posterior pd* $\mathsf{f}(\Theta)$, *the minimum KLD principle respecting this guess and (199) gives*

$$\mathsf{f}(\Theta) \propto \mathsf{f}^{\lambda}(\Delta|\Theta, A, \mathbf{K})\mathsf{f}_0(\Theta), \quad \lambda = \begin{cases} 1 & \textit{if } \mathsf{D}(\mathsf{f}_0||\tilde{\mathsf{f}}) < \beta \\ \in (0,1) & \textit{otherwise} \end{cases} \tag{200}$$

Proof It rearranges Kuhn-Tucker functional to KLD plus a term independent of the opted $\mathsf{f}(\Theta)$ as done repeatedly before. □

Prop. 40 gives the correct pd after using k-th fictitious data item

$$\mathsf{f}(\Theta|\mathbf{K}_k) \propto \mathsf{f}^{\lambda_k}(\Delta_k|\Theta, A_k, \mathbf{K}_{k-1})\mathsf{f}(\Theta|\mathbf{K}_{k-1}), \ \lambda_k \in (0,1]. \qquad (201)$$

- Fictitious data in Bayes' rule is to enter the "flattened" parametric model, where λ_k is determined by the imprecision expressed by (199).
- Sequences of fictitious data are sought to be realised *before* processing real data labelled by $\tau, t \in t^\star$. To stress this, we use
  - *fictitious time* $k \in \{1, \ldots, |k^\star|\}$ marks fictitious data records.
- The weighted Bayesian parameter estimation applied to fictitious data starts at pre-prior pd $\bar{\mathsf{f}}(\Theta)$, which either delimits the range $\Theta^\star$ of unknown parameter Θ or was learnt from previous fictitious data items.
- Further text applies (201) to parametric model given by data vectors Ψ_k

$$\mathsf{f}(\Delta_k|\Theta, A_k, \mathbf{K}_{k-1}) \quad = \quad \mathsf{m}(\Psi_k, \Theta) \qquad (202)$$

$$\Psi_k' = [\Delta_k, \psi_k'] \quad = \quad \text{fictitious [observation, regression vector]}$$

- Knowledge elicitation via the weighted Bayes' rule depends strongly on the chosen weight λ implied by $\mathsf{D}(\mathsf{f}||\tilde{\mathsf{f}}) \leq \beta$ (200).
- The numbers and precisions of fictitious data ↳ expressing knowledge pieces having different knowledge sources may differ substantially.

This motivates to group fictitious data by defining

- *homogenous knowledge piece* , indexed by $g \in g^{\star}$, $|g^{\star}| < \infty$, is expressed by fictitious data Ψ_k, $k \in k_g{}^{\star} \subset k^{\star}$ with a common pd $\mathsf{f}_g(\Psi)$ and a common weight ν_g, which is product of the number of fictitious samples from this sources $|k_g{}^{\star}|$ and the weight λ_g in (201) implied by the imprecision of the processed fictitious data ↳

$$\nu_g \equiv |k_g{}^{\star}|\lambda_g. \tag{203}$$

- Fictitious data vectors Ψ_k corresponding to a homogenous knowledge piece ↳ g are characterised by pd $\mathsf{f}_g(\Psi)$, which is either given or gained by minimum KLD principle ↳, possible extended, cf. Secs. 22, 24, or being Dirac delta ↳ on a given Ψ.

Proposition 41 (Processing of Fictitious Data)

Let us considering $|g^\star| < \infty$ homogenous knowledge pieces characterised by pds $\mathsf{f}_g(\Psi)$ of fictitious data vectors $\Psi_{k;g} \in \Psi^\star$, $k \in k^\star$, and weights (203) $\nu_g < \infty$, $g \in g^\star$. Then, the prior pd $\mathsf{f}(\Theta)$ obtained by the weighted Bayes' rule (201) starting from the pre-prior $\bar{\mathsf{f}}(\Theta)$ pd has the form

$$\mathsf{f}(\Theta) \quad \propto \quad \bar{\mathsf{f}}(\Theta) \exp\left[\sum_{g \in g^\star} \nu_g \int_{\Psi^\star} \mathsf{f}_g(\Psi) \ln(\mathsf{m}(\Psi, \Theta))\, \mathrm{d}\Psi\right]. \qquad (204)$$

Proof By sampling independently fictitious data $\Psi_{k;g} \sim \mathsf{f}_g(\Psi)$ and applying the weighted Bayes' rule (201), we get a sample version of (204), which converges to it by construction of samples. □

- The assumption $\nu_g < \infty$ even for $|k_g{}^\star|$ very large models always limited precision of the knowledge brought by respective sources.
- Formula (204) proposed in [98] is discussed in [75] and used in [76].

Compact Form of Knowledge Elicitation 3rd input

- The choice of the pre-prior pd in the conjugate form
$$\bar{\mathsf{f}}(\Theta) \propto \exp\left[\bar{\nu}\int_{\Psi^\star} \bar{\mathsf{f}}(\Psi)\ln(\mathsf{m}(\Psi,\Theta))\,\mathrm{d}\Psi\right], \qquad (205)$$
given by $\bar{\nu} \geq 0$ and a pd $\bar{\mathsf{f}}(\Psi)$, leads to the compact form of the prior pd$_{\hookrightarrow}$:
$$\mathsf{f}(\Theta) \propto \exp\left[\nu_0\int_{\Psi^\star} \mathsf{f}_0(\Psi)\ln(\mathsf{m}(\Psi,\Theta))\,\mathrm{d}\Psi\right] \qquad (206)$$
$$\nu_0 = \bar{\nu} + \sum_{g\in g^\star}\nu_g \in (0,\infty), \quad \mathsf{f}_0(\Psi) = \frac{\bar{\nu}\bar{\mathsf{f}}(\Psi) + \sum_{g\in g^\star}\nu_g \mathsf{f}_g(\Psi)}{\bar{\nu} + \sum_{g\in g^\star}\nu_g}.$$
The pd $\mathsf{f}_0(\Psi)$ merges pds characterising homogenous knowledge piece$_{\hookrightarrow}$s.
- For dynamic exponential family$_{\hookrightarrow}$ $\mathsf{f}(\Theta)$ becomes conjugate prior$_{\hookrightarrow}$ (154)
$$\mathsf{f}(\Theta) \propto \bar{\mathsf{f}}(\Theta)\mathsf{A}^{\nu_0}(\Theta)\exp\langle V_0, \mathsf{C}(\Theta)\rangle \qquad (207)$$
$$V_0 = \frac{\bar{\nu}}{\nu_0}\bar{V} + \sum_{g\in g^\star}\frac{\nu_g}{\nu_0}V_g, \quad V_g = \int_{\Psi^\star}\mathsf{B}(\Psi)\mathsf{f}_g(\Psi)\,\mathrm{d}\Psi, g\in g^\star, \Rightarrow$$

g-th homogenous knowledge piece$_{\hookrightarrow}$ is to provide its expectation of $\mathsf{B}(\Psi)$.

Choice of Homogenous Groups and Their Weights 3rd input

- Homogenous groups are mostly implied by meaning of the processed knowledge. Their specification does not seem problematic.
- The group weights ν_g can be chosen subjectively to reflect reliability of the knowledge source. This is, however, dangerous as
  - reliability and its guess have a high volatility,
  - knowledge pieces can be mutually dependent, even repeated.

  Thus, it is desirable to choose the group weights more objectively.

Practically, it can be done if some observed data reflecting the current behaviour of the modelled system are available for the choice of weights.

- The choice of $\nu_g, \; g \in g^\star$ is designer↳s' action↳, which should lead to high values of the predictive pd↳ for realised data.

- For a given data D^t, and weights $\nu^{|g^\star|} \geq 0$, the predictive pd is

$$\mathsf{f}(D^t|\nu^{|g^\star|}) = \frac{\mathsf{J}(\nu_t, \mathsf{f}_t)}{\mathsf{J}(\nu_0, \mathsf{f}_0)} \qquad (208)$$
$$\mathsf{J}(\nu, \mathsf{f}) \equiv \int_{\Theta^\star} \exp\left[\nu \int_{\Psi^\star} \mathsf{f}(\Psi) \ln(\mathsf{m}(\Psi, \Theta))\, \mathrm{d}\Theta\right] \mathrm{d}\Psi$$

given by the parametric model $\mathsf{m}(\Psi, \Theta)$. In (208), cf. (206):

$$\mathsf{f}_0 \equiv \mathsf{f}_0(\Psi) = \frac{\bar{\nu}}{\nu_0}\bar{f}(\Psi) + \sum_{g \in g^\star} \frac{\nu_g}{\nu_0}\mathsf{f}_g(\Psi), \qquad \nu_0 = \bar{\nu} + \sum_{g \in g^\star} \nu_g$$
$$\mathsf{f}_t \equiv \mathsf{f}_t(\Psi) = \frac{\nu_0}{\nu_t}\mathsf{f}_0(\Psi) + \frac{t}{\nu_t}\frac{1}{t}\sum_{\tau=1}^{t} \delta(\Psi, \Psi_\tau), \quad \nu_t = \nu_0 + t,$$

with data vectors Ψ_τ made of the observed data D^t, δ is Dirac delta.

- The predictive pd (208) is the ratio of continuous, convex positive functions of $\nu^{|g^\star|}$. It reaches its maximum for non-negative $\nu^{|g^\star|}$.

Open (Knowledge) Elicitation Problems 3rd input

The key open questions related to elicitation tasks are

- does exist a significant class of prior knowledge that cannot be expressed via fictitious data?
- is the conjecture on applicability of the same methodology to filtering valid?
- is the conjecture on applicability to preference elicitation valid?

- The discussion of preference elicitation focuses on the ideal pd $^{\mathsf{I}}\mathsf{f}(B)$ as a descriptor of preferential ordering within FPD.
- It is general enough as any Bayesian DM with strategy-independent performance index (22) $\mathsf{I}(B) = \mathsf{I}_{\mathsf{S}}(B)$ can be converted into the ideal pd using Prop. 24

$$^{\mathsf{I}}\mathsf{f}(B) = \frac{\mathsf{M}(B)\exp\left[-\mathsf{I}(B)/\lambda\right]}{\int_{B^\star}\mathsf{M}(B)\exp\left[-\mathsf{I}(B)/\lambda\right]\,\mathrm{d}B}, \;\; \lambda > 0, \;\; \lambda \approx 0, \quad (209)$$

  where $\mathsf{M}(B)$ is system model recognised in factorisation (75) of the closed-loop model $\mathsf{f}_{\mathsf{S}}(B) = \mathsf{M}(B)\mathsf{S}(B)$.
- Preference elicitation scenarios depend on the ways in which DM quality is reflected in observations.

Preference elicitation in the off-line (design) stage relies on

- fictitious data↳ if DM quality is directly reflected in observations
- minimum KLD principle↳ applied to the ideal closed-loop model.

The DM preferences can be learnt on-line if

- DM quality is directly seen in observations then learning is simply used;
- the unknown preferential pointer↳, used to get complete ordering↳ of behaviours, does not enter the observation model but FPD↳ gives closed-loop model dependent on it;

  Even if the loop is closed non-optimally, its learning is possible with weighted version of Bayes' rule↳ [91] used in lazy-learning [23].
- the form of the optimal strategy is known and the loop is closed by a rational (optimising with respect to its aim) decision maker then optimal-strategy parameters can be estimated [54, 79, 96].

The Choice of the Set of Ideal PDs 4th input

The set of pds ${}^{\mathsf{I}}\mathsf{f}^{\star}$ from which the ideal pd is elicited has to respect that:

- the ideal pd $\equiv {}^{\mathsf{I}}\mathsf{f} \equiv \mathsf{f}_{\mathsf{IS}} =$ closed-loop model with the optimal strategy ${}^{\mathsf{I}}\mathsf{S}$ minimising expected performance index I.
- the ideal pd in the standard Bayesian design is the system model M multiplied by the factor $\exp[-\mathsf{I}(B)/\lambda]$ (209) that can be interpreted as an approximate indicator of the set of desired behaviours $B_\star \subset B^\star$.

Ideal pd should resemble system model restricted to $B_\star$, which implies

- popular quadratic performance index corresponds to normal closed-loop model & should be used if (approximate) normality is reachable;

  It does not suit to systems described by heavy-tailed pds.
- support of the ideal pd is to be the desirable $B_\star$, e.g., admission of large actions' variances can make the optimal strategy useless.

Example: ${}^{\mathsf{I}}\mathsf{f}(\Delta_t|\mathbf{K}_{t-1})$ in Regulation Task 4th input

The regulation task illustrates the use of minimum KLD principle ↳.

- *regulation* is the DM task in which the decision maker selects actions $A_t \in A^\star$ making the observations $\Delta_t \in \Delta^\star$ as close as possible to a given reference $R_t \in \Delta^\star$, $t \in t^\star$, [120].
- The preference elicitation constructs an ideal closed-loop pd, which
  - reflects the verbally and incompletely specified regulation preferences
  - is ambitious but potentially attainable.
- The construction starts with finding the action ${}^{\mathsf{I}}A_t(\mathbf{K}_{t-1})$ making the reference R_t the most probable observation

$$
{}^{\mathsf{I}}A_t(\mathbf{K}_{t-1}) \in \mathrm{Arg} \max_{A_t \in A^\star} \mathsf{f}(\Delta_t = R_t|A_t, \mathbf{K}_{t-1}). \tag{210}
$$

  Note the system model & action set enter the DM formalisation anyway.

- The action ${}^{\mathsf{I}}A_t(\mathbf{K}_{t-1})$ specifies the ambitious but possibly attainable

$$
{}^{\mathsf{I}}\mathsf{f}(\Delta_t|\mathbf{K}_{t-1}) = \mathsf{f}(\Delta_t|\,{}^{\mathsf{I}}A_t(\mathbf{K}_{t-1}), \mathbf{K}_{t-1}).
$$

Example, cont.: ${}^{\mathsf{I}}\mathsf{f}(A_t|\Delta_t, \mathbf{K}_{t-1})$ in Regulation Task

- A pd ${}^{\mathsf{I}}\mathsf{f}_0(A_t|\Delta_t, \mathbf{K}_{t-1})$ with its support on $A^\star$ – either flat or preferring cheap actions – serves as a first guess of the pd ${}^{\mathsf{I}}\mathsf{f}(A_t|\Delta_t, \mathbf{K}_{t-1})$.
- The chain-rule composition ${}^{\mathsf{I}}\mathsf{f}(\Delta_t|\mathbf{K}_{t-1})\, {}^{\mathsf{I}}\mathsf{f}_0(A_t|\Delta_t, \mathbf{K}_{t-1})$ does not suit as closed-loop ideal pd as an adequate joint pd should prefer actions around ${}^{\mathsf{I}}A_t(\mathbf{K}_{t-1})$ defining ${}^{\mathsf{I}}\mathsf{f}(\Delta_t|\mathbf{K}_{t-1}) = \mathsf{f}(\Delta_t|\, {}^{\mathsf{I}}A_t(\mathbf{K}_{t-1}), \mathbf{K}_{t-1})$.
- The simplest expression of the above wish is

$$\left\{ {}^{\mathsf{I}}\mathsf{f}(A_t|\Delta_t, \mathbf{K}_{t-1}) : \int_{A^\star} A_t\, {}^{\mathsf{I}}\mathsf{f}(A_t|\Delta_t, \mathbf{K}_{t-1})\, \mathrm{d}A_t = {}^{\mathsf{I}}A_t(\mathbf{K}_{t-1}) \right\}. \quad (211)$$

- The minimum KLD principle then provides the ideal decision rule

$${}^{\mathsf{I}}\mathsf{f}(A_t|\Delta_t, \mathbf{K}_{t-1}) \propto {}^{\mathsf{I}}\mathsf{f}_0(A_t|\Delta_t, \mathbf{K}_{t-1}) \exp \langle \zeta(\mathbf{K}_{t-1}), A_t \rangle, \quad (212)$$

with the real-valued vector $\zeta(\mathbf{K}_{t-1})$, making the scalar product $\langle \zeta(\mathbf{K}_{t-1}), A_t \rangle$ meaningful, chosen so that equality in (211) is met.

- The system model $\mathsf{f}(\Delta_t|A_t, \mathbf{K}_{t-1})$ is generically obtained as the predictive pd arising from Bayesian estimation, Prop. 17.

  This explains why the action ${}^{\mathsf{I}}A_t(\mathbf{K}_{t-1})$ (210) is not directly applied instead of the above complex indirect construction. The action (210) ${}^{\mathsf{I}}A_t(\mathbf{K}_{t-1})$ is exploitive and FPD adds the needed explorative character.

- The dynamics of DM answers the question why the ideal decision rule (212) is not directly used as a part of the optimal strategy.

  A repetitive use of one-step-ahead-looking rules is far from the optimal strategy, [18], up to closed-loop instability, [83].

The linear-normal case offers an insight into the proposed elicitation

$$
\begin{aligned}
\mathsf{f}(\Delta_t|A_t,\mathbf{K}_{t-1}) &= \mathsf{N}_{\Delta_t}(\mathbb{A}\Delta_{t-1}+\mathbb{B}A_t,\mathbb{Q}) \qquad (213)\\
{}^{\mathsf{I}}\mathsf{f}_0(A_t|\mathbf{K}_{t-1}) &= \mathsf{N}_{A_t}(\mathbb{C}\Delta_{t-1},\mathbb{R})\\
\mathsf{N}_x(\mu,\rho) &= |2\pi\rho|^{-0.5}\exp\left[-0.5(x-\mu)'\rho^{-1}(x-\mu)\right].
\end{aligned}
$$

- The matrices $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{Q} > 0$ of dimensions compatible with vector observations are known from modelling and recursive learning.
- The matrices $\mathbb{C}$, and $\mathbb{R} > 0$ are chosen by the decision maker to have the majority of the probabilistic mass in the desirable action set $A^\star$.
- Let, for simplicity the reference, $R_t = 0$, then the proposed way gives

$$
\begin{aligned}
{}^{\mathsf{I}}\mathsf{f}(A_t|\mathbf{K}_{t-1}) &= \mathsf{N}_{A_t}(\overbrace{-(\mathbb{B}'\mathbb{Q}^{-1}\mathbb{B})^{-1}\mathbb{B}'\mathbb{Q}^{-1}\mathbb{A}\Delta_{t-1}}^{{}^{\mathsf{I}}A(\mathbf{K}_{t-1}),\ (210)},\mathbb{R}) \qquad (214)\\
{}^{\mathsf{I}}\mathsf{f}(\Delta_t|\mathbf{K}_{t-1}) &= \mathsf{N}_{\Delta_t}((I-\mathbb{B}(\mathbb{B}'\mathbb{Q}^{-1}\mathbb{B})^{-1}\mathbb{B}'\mathbb{Q}^{-1})\mathbb{A}\Delta_{t-1},\mathbb{Q}).
\end{aligned}
$$

Recall

FPD with the normal system model and the ideal pd is a randomised version of the widespread design dealing a with linear system and quadratic loss, Sec. 32, [120, 70],

$$\sum_{t \in t^\star} (\Delta_t - R_t)' \mathbb{Q}_{\Delta^\star} (\Delta_t - R_t) + (A_t - R_{t;A})' \mathbb{Q}_{A^\star} (A_t - R_{t;A}),$$

given by set points R_t, $R_{t;A}$, and penalisation matrices $\mathbb{Q}_{\Delta^\star}$, $\mathbb{Q}_{A^\star} > 0$.

- The choice of references R_t, $R_{t;A}$ is easy. The choice of weights $\mathbb{Q}_{\Delta^\star}$, $\mathbb{Q}_{A^\star}$ is difficult repetitively solved problem, (almost) resolved above.
- The matrix $\mathbb{Q}_{\Delta^\star} = \mathbb{Q}^{-1}$ comes from the learnt system model.
- The matrix $\mathbb{Q}_{A^\star} = \mathbb{R}^{-1}$ is determined by the need to "cover" the set $A^\star$ implied by technological constraints by $\mathsf{N}_{A_t}(\mathbb{C}\Delta_{t-1}, \mathbb{R})$.

  This makes the result implicitly dependent on the matrix $\mathbb{C}$.
- The solution extends the DESIGNER line [8] presented in [9, 86, 125].

. . . Omitted Topics

- Preference elicitation is hot topic, especially, in Internet context.
- Serious research of preference elicitation is still in its infancy and inspects a lot of alternatives unmentioned here.
- Even preference-elicitation scenarios already mentioned are insufficiently elaborated and here are not presented.

References

References I

[1] Abbasi, D., Ashrafzadeh, A., Asheghi, R., Merufinia, E.: A survey on the performance of fuzzy-neural network at predicting the average monthly discharge of catchment basin areas having snow regimes. J. of Civil Engineering and Urbanism **4**, 480–484 (2014)

[2] Anderson, B., Moore, J.: Optimal Filtering. Prentice Hall (1979)

[3] Anderson, B., Moore, J.: Optimal Control : Linear Quadratic Methods. Prentice-Hall, Englewood Cliffs, New Jersey (1989)

[4] Andrýsek, J.: Estimation of dynamic probabilistic mixtures. Ph.D. thesis, FJFI, ČVUT, POB 18, 18208 Prague 8, Czech Republic (2005)

[5] Arrow, K.: Social Choice and Individual Values. New Haven: Yale University Press (1963). 2nd. ed.

References II

[6] Åström, K.: Introduction to Stochastic Control. Academic Press, NY (1970)

[7] Åström, K., Wittenmark, B.: Adaptive Control. Addison-Wesley, Massachusetts (1989)

[8] Bůcha, J., Kárný, M., Nedoma, P., Böhm, J., Rojíček, J.: Designer 2000 project. In: Int. Conference on Control '98, pp. 1450–1455. IEE, London (1998)

[9] Belda, K.: Probabilistically tuned LQ control for mechatronic applications (paper version). AT&P J. plus **9**(2 (2009)), 19–24 (2009)

[10] Bellman, R.: Dynamic Programming. Princeton University Press, NY (1957)

References III

[11] Berec, L.: Model Structure Identification: Global and Local Views. Bayesian Solution. Ph.D. Thesis, Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering, Prague (1998)

[12] Berec, L., Kárný, M.: Identification of reality in Bayesian context. In: K. Warwick, M. Kárný (eds.) Computer-Intensive Methods in Control and Signal Processing, pp. 181–193. Birkhäuser (1997)

[13] Berec, L., Rojíček, J.: Control Period Selection: Verification on Coupled Tanks. In: G. Bastin, M. Gevers (eds.) Preprints of European Control Conference ECC'97 (on CD-ROM). ECC, Brussels (1997)

[14] Berger, J.: Statistical Decision Theory and Bayesian Analysis. Springer, NY (1985)

References IV

[15] Bernardo, J.: Expected information as expected utility. The An. of Stat. **7**(3), 686–690 (1979)

[16] Bernstein, D., Amato, C., Hansen, E., Zilberstein, S.: Policy iteration for decentralized control of Markov decision processes. J. of Artificial Intelligence Research **34**, 89 –132 (2009)

[17] Bertsekas, D.: Nonlinear Programming. Athena Scientific, Belmont, Massachusetts (1999)

[18] Bertsekas, D.: Dynamic Programming and Optimal Control. Athena Scientific, US (2001)

[19] Bierman, G.: Factorization Methods for Discrete Sequential Estimation. Academic Press, NY (1977)

[20] Bishop, C.: Training with noise is equivalent to Tikhonov regularization. Neural Computation **7**(1), 108–116 (1995)

[21] Böhm, J.: Adaptive predictive LQ control with constraints. In: S. Krejčí, I. Taufer (eds.) Proc. of the 6th Int. Scientific - Technical Conference. Process Control 2004, pp. 1–12. Univerzity of Pardubice, Pardubice (2004)

[22] Bond, A., Gasser, L.: Readings in Distributed Artificial Intelligence. Morgan Kaufmann (2014)

[23] Bontempi, G., Birattari, M., Bersini, H.: Lazy learning for local modelling & control design. Int. J. of Control **72**(7-8), 643–658 (1999)

[24] Bourbaki, N.: Elements of Mathematics: General Topology. Addison-Wesley (1966)

[25] Busemeyer, J., Bruza, P.: Quantum Models of Cognition and Decision. Cambridge University Press, New York (2012)

References VI

[26] Campenhout, J.V., Cover, T.: Maximum entropy and conditional probability. IEEE Tran. on Inf. Theory **27**(4), 483–489 (1981)

[27] Cannon, M., Kouvaritakis, B., Brooms, A., Lee, Y.: Efficient nonlinear model predictive control. In: Proc. of American Control Conference, June 28-30. Chicago (2000)

[28] Cao, L., Schwartz, H.: A directional forgetting algorithm based on the decomposition of the information matrix. Automatica **36**(11), 1725–1731 (2000)

[29] Chen, Q.: Approximate Kalman Filtering. World Scientific (1993)

[30] Clarke, D.: Advances in Model-Based Predictive Control. Oxford University Press, Oxford (1994)

[31] Clarke, D., Mohtadi, C., Tuffs, P.: Generalized predictive control. Automatica **23**(2), 137–160 (1987)

References VII

[32] Csiszár, I., Körner, J.: Information Theory: Coding Theorems for Discrete Memoryless Systems. Akadémiai Kiadó, Budapest (1986)

[33] Daum, E.: New exact nonlinear filters. In: J. Spall (ed.) Bayesian Analysis of Time Series and Dynamic Models. Marcel Dekker (1988)

[34] Debreu, G.: Representation of a preference ordering by a numerical function. In: R. Thrall, C. Coombs, R. Davis (eds.) Decision Processes. Wiley, NY (1954)

[35] DeGroot, M.: Optimal Statistical Decisions. McGraw-Hill, NY (1970)

[36] Dempster, A.P., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. of the Royal Statistical Society. Series B (Methodological) pp. 1–38 (1977)

[37] Devroye, L., Györfi, L., Lugosi, G.: A Probabilistic Theory of Pattern Recognition. Springer Science & Business Media (2013)

[38] Dvurečenskij, A.: Gleasons Theorem and Its Applications, *Mathematics and Its Applications*, vol. 60. Kluwer Academic Publishers – Ister Science Press, Bratislava, Dordrecht/Boston/London (1993)

[39] Efthimiou, C.: Introduction to functional equations: Theory and problem-solving strategies for mathematical competitions and beyond (2010)

[40] Farias, D.D., Roy, B.: The linear programming approach to approximate dynamic programming. Operations Research **51**(6), 850–865 (2003)

[41] Feinberg, E., Shwartz, A.: Handbook of Markov Decision Processes: Methods and Applications. Kluwer Academic Publishers, Boston, Dordrecht, London (2002)

[42] Feldbaum, A.: Theory of dual control. Autom. Remote Control **21,22**(9,2) (1960,61)

[43] Ferguson, T.S.: Who solved the secretary problem? Statistical Science **4**(3), 282–289 (1989)

[44] Filatov, N., (Ed.), H.U.: Adaptive Dual Control - Theory and Applications. Springer, Berlin (2004)

[45] Fine, T.: Theories of Probability: A examination of foundations. Academic Press, New York, London (1973)

[46] Fishburn, P.: Utility Theory for Decision Making. J. Wiley, New York, London, Sydney, Toronto (1970)

[47] Franceschini, G., Macchietto, S.: Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science **63**(19), 4846 – 4872 (2008). URL http://www.sciencedirect.com/science/article/pii/S0009250907008871

[48] Garthwaite, P., Kadane, J., O´Hagan, A.: Statistical methods for eliciting probability distributions. J. of the American Statistical Association **100**(470), 680–700 (2005)

[49] Gillijns, S., Mendoza, O.B., Chandrasekar, J., Moor, B.L.R.D., Bernstein, D.S., Ridley, A.: What is the ensemble Kalman filter and how well does it work? In: Proceedings of the 2006 American Control Conference, p. FrA08.1. Minneapolis, Minnesota, USA (2006). DOI 10.1109/ACC.2006.1657419

References XI

[50] Glimcher, P., Fehr, E.: Neuroeconomics: Decision Making and the Brain. Academic Press (2013)

[51] Goldreich, O.: Computational Complexity: A Conceptual Perspective. Cambridge University Press (2008)

[52] Golub, G., VanLoan, C.: Matrix Computations. The John Hopkins University Press, Baltimore – London (1996). 3rd edition

[53] Gómez, A.G.V., Kappen, H.: Dynamic policy programming. The J. of Machine Learning Research **30**, 3207–3245 (2012)

[54] Guy, T., Kárný, M., Villa, A., Lintas, A.: Theoretical models of decision-making in the ultimatum game: Fairness vs. reason. In: W. Rubin, P. Xiaochuan (eds.) Advances in Cognitive Neurodynamics (V), pp. 185–191. Springer-Verlag (2016)

[55] Guy, T.V., Kárný, M.: Stationary fully probabilistic control design. In: J. Filipe, J.A. Cetto, J.L. Ferrier (eds.) Proc. of the Second Int. Conference on Informatics in Control, Automation and Robotics, pp. 109–112. INSTICC, Barcelona (2005)

[56] Haykin, S.: Neural Networks: A Comprehensive Foundation. Macmillan, NY (1994)

[57] Horst, R., Tuy, H.: Global Optimization. Springer (1996). 727 pp.

[58] Hůla, F.: Explorative fully probabilistic design of adaptive decision strategies. Master's thesis, FJFI, Czech Technical University, Prague (2018)

[59] Hutter, M.: Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability. Springer, Berlin, Heidelberg, New York (2005)

[60] Ishikawa, M., Moriyama, T.: Prediction of time series by a structural learning of neural networks. Fuzzy Sets and Systems **82**(2), 167–176 (1996)

[61] Jacobs, O., Patchell, J.: Caution and probing in stochastic control. Int. J. of Control **16**, 189–199 (1972)

[62] Jarník, V.: Integral Calculus II. Academia, Prague (1984). In Czech

[63] Jazwinski, A.: Stochastic Processes and Filtering Theory. Academic Press, NY (1970)

[64] Jensen, F.: Bayesian Networks and Decision Graphs. Springer-Verlag, NY (2001)

[65] Jordan, M.: Learning in Graphical Models. MIT Press (1998)

References XIV

[66] Julier, S., Uhlmann, J.: Unscented filtering and nonlinear estimation. Proceedings of the IEEE **90**(3), 401–422 (2004)

[67] Kadane, J., Dickey, J., Winkler, R., Smith, W., Peters, S.: Interactive elicitation of opinions for normal linear models. J. of the American Statistical Association **75**(372), 845–854 (1980)

[68] Kárný, M.: Algorithms for determining the model structure of a controlled system. Kybernetika **19**(2), 164–178 (1983)

[69] Kárný, M.: Estimation of control period for selftuners. Automatica **27**(2), 339–348 (1991). Extended version of the paper presented at 11th IFAC World Congress, Tallinn

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

References XV

[71] Kárný, M.: Adaptive systems: Local approximators? In: Workshop on Adaptive Systems in Control and Signal Processing, pp. 129–134. IFAC, Glasgow (1998)

[72] Kárný, M.: On approximate fully probabilistic design of decision making strategies. In: T. Guy, M. Kárný (eds.) Proceedings of the 3rd International Workshop on Scalable Decision Making, ECML/PKDD 2013. UTIA, Prague (2013). ISBN 978-80-903834-8-7

[73] Kárný, M.: Approximate Bayesian recursive estimation. Infor. Sciences **289**, 100–111 (2014)

[74] Kárný, M., Andrýsek, J.: Use of Kullback-Leibler divergence for forgetting. International Journal of Adaptive Control and Signal Processing **23**(1), 1–15 (2009). Download

[75] Kárný, M., Andrýsek, J., Bodini, A., Guy, T., Kracík, J., Ruggeri, F.: How to exploit external model of data for parameter estimation? International Journal of Adaptive Control and Signal Processing **20**(1), 41–50 (2006). Download

[76] Kárný, M., Bodini, A., Guy, T., Kracík, J., Nedoma, P., Ruggeri, F.: Fully probabilistic knowledge expression and incorporation. Statistics and Its Interface **7**(4), 503–515 (2014)

[77] Kárný, M., Böhm, J., Guy, T.V., Jirsa, L., Nagy, I., Nedoma, P., Tesař, L.: Optimized Bayesian Dynamic Advising: Theory and Algorithms. Springer (2006)

[78] Kárný, M., Guy, T.: Preference elicitation in fully probabilistic design of decision strategies. In: Proc. of the 49th IEEE Conference on Decision and Control (2010)

[79] Kárný, M., Guy, T.: On the origins of imperfection and apparent non-rationality. In: T. Guy, M. Kárný, D. Wolpert (eds.) Decision Making: Uncertainty, Imperfection, Deliberation and Scalability, Studies in Computational Intelligence, pp. 57–91. Springer, International Publishing Switzerland (2015)

[80] Kárný, M., Guy, T., Bodini, A., Ruggeri, F.: Cooperation via sharing of probabilistic information. International Journal of Computational Intelligence Studies pp. 139–162 (2009). Download

[81] Kárný, M., Guy, T.V.: Fully probabilistic control design. Systems & Control Letters **55**(4), 259–265 (2006)

[82] Kárný, M., Halousková, A.: Pretuning of self-tuners. In: D. Clarke (ed.) Advances in Model-Based Predictive Control, pp. 333–343. Oxford University Press, Oxford (1994)

[83] Kárný, M., Halousková, A., Böhm, J., Kulhavý, R., Nedoma, P.: Design of linear quadratic adaptive control: Theory and algorithms for practice. Kybernetika **21** (1985). Supp. Nos 3–6

[84] Kárný, M., Herzallah, R.: Scalable harmonization of complex networks with local adaptive controllers. IEEE Trans. on SMC: Systems **47**(3), 394–404 (2017)

[85] Kárný, M., Hůla, F.: Balancing exploitation and exploration via fully probabilistic design of decision policies. In: NIPS2018 (2018). Submitted

[86] Kárný, M., Jeníček, T., Ottenheimer, W.: Contribution to prior tuning of LQG selftuners. Kybernetika **26**(2), 107–121 (1990)

[87] Kárný, M., Kadlec, J., Sutanto, E.L.: Quasi-Bayes estimation applied to normal mixture. In: J. Rojíček, M. Valečková, M. Kárný, K. Warwick (eds.) Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing, pp. 77–82. ÚTIA AV ČR, Prague (1998)

[88] Kárný, M., Khailova, N., Nedoma, P., Böhm, J.: Quantification of prior information revised. Int. J. of Adaptive Control and Signal Processing **15**(1), 65–84 (2001)

[89] Kárný, M., Kroupa, T.: Axiomatisation of fully probabilistic design. Information Sciences **186**(1), 105–113 (2012). Download

[90] Kárný, M., Kulhavý, R.: Structure determination of regression-type models for adaptive prediction and control. In: J. Spall (ed.) Bayesian Analysis of Time Series and Dynamic Models. Marcel Dekker, New York (1988). Chapter 12

[91] Kárný, M., Macek, K., Guy, T.: Lazy fully probabilistic design of decision strategies. In: 11th Int. Symp. on Neural Networks, no. 8866 in LNCS, pp. 140–149 (2014)

[92] Kárný, M., Nedoma, P., Khailova, N., Pavelková, L.: Prior information in structure estimation. IEE Proc. — Control Theory and Applications **150**(6), 643–653 (2003)

[93] Kárný, M., Nedoma, P., Šmídl, V.: Cross-validation of controlled dynamic models: Bayesian approach. In: Preprints of the 16th IFAC World Congress. IFAC, Prague (2005)

[94] Kárný, M., Šindelář, J., Š. Pírko, Zeman, J.: Adaptively optimized trading with futures. Tech. rep., ÚTIA (2010)

[95] Keeny, R., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. J. Wiley, NY (1978)

[96] Knejflová, Z., Avanesyan, G., T.V.Guy, Kárný, M.: What lies beneath players' non-rationality in ultimatum game? In: T. Guy, M. Kárný (eds.) Proceedings of the 3rd International Workshop on Scalable Decision Making, ECML/PKDD 2013. UTIA AV ČR, Prague (2013)

[97] Koopman, R.: On distributions admitting a sufficient statistic. Trans. of Am. Math. Society **39**, 399 (1936)

[98] Kracík, J., Kárný, M.: Merging of data knowledge in Bayesian estimation. In: J. Filipe, J.A. Cetto, J.L. Ferrier (eds.) Proc. of the Second Int. Conference on Informatics in Control, Automation and Robotics, pp. 229–232. INSTICC, Barcelona (2005)

[99] Králik, J., Stiegler, P., Vostrý, Z., Záworka, J.: Dynamic Modeling of Large-Scale Network with Application to Gas Distribution. Elsevier, Amsterdam – Oxford – New York – Tokyo (1988)

References XXII

[100] Krstić, M., Kannellakopoulos, I., Kokotović, P.: Nonlinear and Adaptive Control Design. Wiley-Interscience, NY (1995)

[101] Kruse, R., Borgelt, C.: Data mining with graphical models. Lecture Notes In Computer Science **2534**, 2–11 (2002)

[102] Kuhn, H., Tucker, A.: Nonlinear programming. In: Proc. of 2nd Berkeley Symposium, pp. 481–492. Univ. of California Press (1951)

[103] Kulhavý, R.: Directional tracking of regression-type model parameters. In: Preprints of the 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing, pp. 97–102. Lund, Sweden (1986)

[104] Kulhavý, R.: Restricted exponential forgetting in real-time identification. Automatica **23**(5), 589–600 (1987)

[105] Kulhavý, R.: A Bayes-closed approximation of recursive nonlinear estimation. Int. J. Adaptive Control and Signal Proc. **4**, 271–285 (1990)

[106] Kulhavý, R.: Recursive Bayesian estimation under memory limitations. Kybernetika **26**, 1–20 (1990)

[107] Kulhavý, R.: Can approximate Bayesian estimation be consistent with the ideal solution? In: Proc. of the 12th IFAC World Congress, vol. 4, pp. 225–228. Sydney, Australia (1993)

[108] Kulhavý, R.: Can we preserve the structure of recursive Bayesian estimation in a limited-dimensional implementation? In: U. Helmke, R. Mennicken, J. Saurer (eds.) Systems and Networks: Mathematical Theory and Applications, vol. I, pp. 251–272. Akademie Verlag, Berlin (1994)

References XXIV

[109] Kulhavý, R.: Recursive Nonlinear Estimation: A Geometric Approach, *Lecture Notes in Control and Information Sciences*, vol. 216. Springer-Verlag, London (1996)

[110] Kulhavý, R., Kárný, M.: Tracking of slowly varying parameters by directional forgetting. In: Preprints of the 9th IFAC World Congress, vol. X, pp. 178–183. IFAC, Budapest (1984)

[111] Kulhavý, R., Kraus, F.J.: On duality of regularized exponential and linear forgetting. Automatica **32**, 1403–1415 (1996)

[112] Kulhavý, R., Zarrop, M.B.: On a general concept of forgetting. Int. J. of Control **58**(4), 905–924 (1993)

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

[114] Kumar, P.: A survey on some results in stochastic adaptive control. SIAM J. Control and Applications **23**, 399–409 (1985)

[115] Kumar, P.R., Varaiya, P.: Stochastic Systems: Estimation, Identification, and Adaptive Control, *Classics of Applied Mathematics*, vol. 75. SIAM (2015)

[116] Kushner, H.: Introduction to Stochastic Control. Holt, Rinehart and Winston, NY (1971)

[117] Ljung, L.: System Identification: Theory for the User. Prentice-Hall, London (1987)

[118] Loeve, M.: Probability Theory. van Nostrand, Princeton, New Jersey (1962). Russian translation, Moscow 1962

[119] Mantegna, R., Stanley, H.: An introduction to econophysics: correlations and complexity in finance. Cambridge University Press (2000)

[120] Meditch, J.: Stochastic Optimal Linear Estimation and Control. Mc. Graw Hill (1969)

[121] Milek, J., Kraus, F.: Time-varying stabilized forgetting for recursive least squares identification. In: C. Bányász (ed.) IFAC Symposium ACASP'95, pp. 539–544. IFAC, Budapest (1995)

[122] Mine, H., Osaki, S.: Markovian Decision Processes. Elsevier (1970)

[123] Mosca, E.: Optimal, Predictive, and Adaptive Control. Prentice Hall (1994)

[124] von Neumann, J.: Mathematical Foundations of Quantum Theory. Princenton U. (1932)

[125] Novák, M., Böhm, J.: Adaptive LQG controller tuning. In: M.H. Hamza (ed.) Proc. of the 22nd IASTED Int. Conference Modelling, Identification and Control. Acta Press, Calgary (2003)

[126] Novák, M., Böhm, J., Nedoma, P., Tesař, L.: Adaptive LQG controller tuning. IEE Proc. — Control Theory and Applications **150**(6), 655–665 (2003)

[127] Novikov, A.: Optimal sequential procedures with Bayes decision rules. Kybernetika **46**(4) (2010)

[128] Ogden, R.: Essential wavelets for statistical applications and data analysis. Birkhauser, Boston-Basel-Berlin (1997)

[129] Oppenheim, A., Wilsky, A.: Signals and Systems. Englewood Clifts, Jersye (1983)

[130] Pavelková, L.: Estimation of models with uniform innovations and its application on traffic data. Ph.D. thesis, Czech Technical University in Prague (2008). Download

[131] Peterka, V.: Adaptive digital regulation of noisy systems. In: Preprints of the 2nd IFAC Symposium on Identification and Process Parameter Estimation. ÚTIA ČSAV, Prague (1970). Paper 6.2

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

[133] Peterka, V.: Predictor-based self-tuning control. Automatica **20**(1), 39–50 (1984). Reprinted in: *Adaptive Methods for Control System Design*, M.M. Gupta, Ed., IEEE Press, New York, 1986

[134] Plutowski, M.: Survey: Cross-validation in theory and practice. Research report, Department of Computational Science Research, David Sarnoff Research Center, Princeton, New Jersey (1996)

[135] Pothos, E., Busemeyer, J.: A quantum probability explanation for violations of 'rational' decision theory. Proceedings of the Royal Society, B pp. 2171–2178 (2009)

[136] Rao, C.: Linear method of statistical inference and their applications. Academia, Prague (1987). In Czech

[137] Rao, M.: Measure Theory and Integration. John Wiley, NY (1987)

[138] Renyi, A.: Probability theory. Academia, Prague (1972). In Czech

[139] Ripley, B.: Pattern Recognition and Neural Networks. Cambridge University Press, London (1997)

[140] Ristic, B., Arulampalam, S.: Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House (2004)

[141] Rojíček, J., Kárný, M.: A sequential stopping rule for extensive simulations. In: J. Rojíček, M. Valečková, M. Kárný, K. Warwick (eds.) Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing, pp. 145–150. ÚTIA AV ČR, Praha (1998)

[142] Sandholm, T.: Distributed rational decision making. In: G. Weiss (ed.) Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence, pp. 201–258 (1999)

[143] Sanov, I.: On probability of large deviations of random variables. Matematičeskij Sbornik **42**, 11–44 (1957). (in Russian), also in Selected Translations Mathematical Statistics and Probability, I, 1961, 213–244

[144] Savage, L.: Foundations of Statistics. Wiley, NY (1954)

[145] Sečkárová, V.: Cross-entropy based combination of discrete probability distributions for distributed decision making. Ph.D. thesis, Charles University in Prague, Faculty of Mathematics and Physics, Dept. of Probability and Mathematical Statistics., Prague (2015). Submitted in May 2015. Successfully defended on 14.09.2015.

[146] Setnes, M., Babuška, R.: Fuzzy decision support for the control of detergent production. Int. J. of Adaptive Control and Signal Processing **15**(8), 769–785 (2001)

[147] Sheng, Y., Tomizuka, M., Ozaki, M.: Dynamic modelling and adaptive predictive control of drilling of composite materials. In: Proc. of American Control Conference, June 28-30. Chicago (2000)

[148] Shore, J., Johnson, R.: Axiomatic derivation of the principle of maximum entropy & the principle of minimum cross-entropy. IEEE Tran. on Inf. Th. **26**(1), 26–37 (1980)

[149] Si, J., Barto, A., Powell, W., Wunsch, D. (eds.): Handbook of Learning and Approximate Dynamic Programming. Wiley-IEEE Press, Danvers (2004)

[150] Simon, H.: A behavioral model of rational choice. The Quarterly J. of Economics **LXIX**, 299–310 (1955)

[151] Simon, H.: Theories of decision-making in economics and behavioral science. The American Economic Review **XLIX**, 253–283 (1959)

[152] Sims, C.: Implications of rational inattention. J. of Monetary Economics **50**(3), 665 – 690 (2003)

References XXXIII

[153] Šindelář, J., Vajda, I., Kárný, M.: Stochastic control optimal in the Kullback sense. Kybernetika **44**(1), 53–60 (2008). Download

[154] Šmídl, V., Quinn, A.: The Variational Bayes Method in Signal Processing. Springer (2005)

[155] Starmer, C.: Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. J. of Economic Literature **XXXVIII**, 332–382 (2000)

[156] Tikhonov, A., Goncharsky, A., Stepanov, V., Yagola, A.: Numerical Methods for the Solution of Ill-Posed Problems. Kluwer, Dordrecht (1995)

References XXXIV

[157] Tishby, N.: Predictive information and the brain's internal time. In: T. Guy, M. Kárný (eds.) Proceedings of the 3rd Int. Workshop on Scalable Decision Making held in conjunction with ECML/PKDD 2013. ÚTIA AVČR, Prague (2013)

[158] Tishby, N., Polani, D.: Information theory of decisions and actions. In: V. Cutsuridis, A. Hussain, J. Taylor (eds.) Perception-Action Cycle, Springer Series in Cognitive and Neural Systems, pp. 601–636. Springer, New York (2011)

[159] Titterington, D., Smith, A., Makov, U.: Statistical Analysis of Finite Mixtures. John Wiley, NY (1985)

[160] Todorov, E.: Linearly-solvable Markov decision problems. In: B. Schölkopf, et al (eds.) Advances in Neural Inf. Processing, pp. 1369 – 1376. MIT Press, NY (2006)

[161] Triantaphyllou, E.: Multi-Criteria Decision Making Methods: A Comparative Study. Kluwer Academic Publishers (2000)

[162] Vajda, I.: Information Theory and Statistical Decision Making. Alfa, Bratislava (1982). In Slovak

[163] Vladimirov, V.: Generalized Functions in Mathematical Physics. Mir Publishers, Moscow (1979)

[164] Wald, A.: Statistical Decision Functions. John Wiley, NY, London (1950)

[165] Wiener, N.: Cybernetics or Control and Communications in the Animal and the Machine. Hermann, Paris (1948)

[166] Wiener, N.: Cybernetics. The MIT Press, John Wiley, NY (1961). 2nd edition

References XXXVI

[167] Wolpert, D., Macready, W.: No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation **1**(1), 67–82 (1997)

[168] Yager, R., Liu, L. (eds.): Classic Works of the Dempster-Shafer Theory of Belief Functions. Springer (2008)

[169] Zadeh, L.: A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. Systems Theory in the Social Sciences pp. 202–282 (1976)

[170] Zarrop, M.: Experiment Design for Dynamic System Identification. Lecture Notes in Control and Information Sciences 21. Springer, NY (1979)

[171] Zopounidis, C., Galariotis, E., Doumpos, M., Sarri, S., Andriosopoulos, K.: Multiple criteria decision aiding for finance: An updated bibliographic survey. European Journal of Operational Research **247**(2), 339 – 348 (2015). DOI http://dx.doi.org/10.1016/j.ejor.2015.05.032

Indices

List of Terms II

List of Terms III

List of Terms VI

List of Terms VIII

List of Terms XV

List of Terms XVI

$\mathbf{1}$, `\1`	vector of units, 1
$\ell_{X^\star}$, `\Cv{X}`	length of vector X, 1
$\lvert X^\star\rvert$, `\S{X}`	cardinality of $X^\star$, 1
$\mathcal{D}i$, `\Di`	Dirrichlet rnd, 1
$\mathsf{D}\left(\mathsf{f}\|\|\mathsf{g}\right)$, `\D{\O{f}}{\O{g}}`	Kullback-Leibler divergence of f on g, 1
$\mathsf{E}\left[X\right]$,, `\Eu{X}`	unconditional expectation of X, 1
$\mathsf{E}\left[X\lvert Y\right]$, `\E{X}{Y}`	expectation of X conditioned on Y, 1
$\mathcal{G}i\mathcal{W}$, `\GiW`	Gauss-inverse-Wishart rnd, 1
$\mathbf{G}_t{=}\mathbf{G}_{A_t}$,, `\G{t}`	ignorance of the action A_t, 1
$\mathbf{K}_{t-1}{=}\mathbf{K}_{A_t}$, `\K{t-1}`	knowledge of the action A_t, 1
$\mathcal{X}$, `\M{X}`	mathcal font used for rare symbols, 1
$\mathcal{N}$, `\N`	normal (Gaussian) rnd, 1
X, `\O{X}`	mathsf font reserved for operators, 1
X, `\R{X}`	bold mathematics font used for realisations, 1
$X^\star$, `\S{X}`	set of Xs, 1

${}^{\surd}X$, `\Us{X}`	left upper square-root index of X, 1
${}^{a}X$, `\U{a}{X}`	left upper non-numerical index of X, 1
$\mathrm{d}X$, `\d{X}`	differential of X, `\d{X}`, 1
$\mathfrak{X}$, `\fr{X}`	$\mathfrak{mathfrakfonts}$, 1
$\mathsf{f}(X)$, `\fu{X}`	unconditional rnd of X, 1
$\mathsf{f}(X\|Y)$, `\f{X}{Y}`	rnd of X conditioned on Y, 1
$X \in X^{\star}$, `\is{X}`	X in the set $X^{\star}$, 1
$\mathbb{X}$, `\m{X}`	matrix, 1
$\mathrm{supp}[X]$, `\su{X}`	support of X, 1
$X_{\star}$, `\s{X}`	subset of $X^{\star}$, 1
$\underline{X}$, `\un{X}`	realisation of X, 1
`\begin{agr}..\end{agr}`	Agreement (definition), labelled `\label{agr:name}`, 1
`\begin{alg}..\end{alg}`	Algorithm, labelled `\label{alg:name}`, 1
`\begin{cor}..\end{cor}`	Corollary, labelled `\label{cor:name}`, 1
`\begin{exa}..\end{exa}`	Example, labelled `\label{exa:name}`, 1

Latex New Commands and Keywords, see mathematics III

`\begin{exc}..\end{exc}`	Exercises, labelled `\label{exc:name}`, 1
`\begin{prb}..\end{prb}`	Problem, labelled `\label{prb:name}`, 1
`\begin{pro}..\end{pro}`	Proposition (weaker theorem), referred `\label{pro:name}`, 1
`\begin{rem}..\end{rem}`	Remark, labelled `\label{rem:name}`, 1
`\begin{req}..\end{req}`	Requirement (condition), labelled `\label{req:name}`, 1
`\begin{thm}..\end{thm}`	Theorem, labelled `\label{thm:name}`, 1

Latex New Commands and Keywords, see mathematics IV

Some Solutions

Solved Exercises 1 on DM Elements

ad 1 Think over your personal example of dynamic DM.

ad 2 Specify its basic DM elements. See Table 11.

ad 3 Does you miss some DM element? Conjecture: nothing is missing at this abstraction level.

ad 4 Is to be aim↳ scalar? No. I want to be rich and fairly treated.

ad 5 Who specify the system↳? Primarily, the user↳ according to the addressed DM task. It should be advised by the cooperating designer↳.

ad 6 Why the World is not taken as the system? It would lead to too complex problem unsolvable with available resources (time, budget, computational power, available data . . .).

ad 7 Think about the relation ignorance↳ and uncertainty↳. The conjectured answer: The uncertainty↳ is a part of ignorance↳, which is uninfluenced by the used strategy↳.

add 1 Why is necessary to work with strategy↳ and not just with a sequence of action↳s? DM theory solves classes of DM problems without knowing the specific knowledge↳ realisation. It prepares processing algorithms (program).

add 2 Select your favourite domain and try to find relationships between its and ours vocabularies.

add 3 Select your personal example of dynamic DM and delimit all introduced notions on it.

add 4 Think over how you would solve your DM problem in its entirety.

add 5 Specify where your technical knowledge is insufficient for a systematic solution of your dynamic DM problem.

Solved Exercises 3 on Preferential Ordering

add 1 Think over your case of DM and specify your preferential ordering ↳.

add 2 Make your preferential ordering ↳ complete.

add 3 Why the ordering transitivity is so important? Even in deterministic case it would led to decision leading to consequences you dislike.

add 4 Do you know real-life example of the transitivity violation? Think about "hidden" aspect of determining your preference. Say, except of the financial gain you care about fairness.

add 5 Does exist a preferential ordering ↳, which operates on a proper extension of the behaviour set? Conjecture: no! If you think about something it becomes a part of the behaviour ↳.

add 6 How you would handle the preferential ordering ↳ if your answer to the previous question is affirmative?

ad 1 You search for maximum of a function while you

- select sequentially its argument and evaluate the function and compare with previous values at the unit cost
- pay a unit cost for considering any idea how to choose the argument
- have the budget equal to 221.

Notice

- to have 221 ideas is surely unproductive
- to make 221 random choices is probably unproductive.

ad 2 The problem is closely related to

- secretary (marriage) problem [43],
- no free lunch theorems [167].

Learn more about them and think over other real-life situations fitting the scheme.

Solved Exercises 6 on Actions

ad 1 What is the mathematical reason that the optimal rule can always be deterministic? The optimised strategy$_{↳}$ enters expectation linearly!

ad 2 Provide a performance index giving a parameter estimate $\neq \mathsf{E}[\Theta|\mathbf{K}_{\hat{\Theta}}]$.

ad 3 Formulate and solve interval estimation as DM task.

ad 4 Estimation selects action$_{↳}$, which does not influence the system$_{↳}$. Provide an example of DM task with action$_{↳}$s influencing the system$_{↳}$. For instance, heating influences temperature in the heated room.

ad 5 Medical doctor has to decide whether to prescribe antibiotics to patient or not. Formulate and solve this as static (one-shot) DM task.

ad 6 Judge has to decide whether the accused is guilty or not. Formulate and solve this as static DM task.

ad 7 What are possible dynamic consequences of the above tasks?

Solved Exercises 8 on MDP

ad 1 Code Algorithm 1 in your favourite software or find it elsewhere. Play a bit with it.

ad 2 Consider a finely discretised continuous real-valued function defined on three-dimensional unit box. Formulate the search for its minimum as Markov decision process.

ad 3 For a very fine discretisation you will encounter problems known as curse of dimensionality, [10]. Where it manifests in your program?

ad 4 Think over the foreseen problems even if you avoid programming.

ad 5 Think over how to suppress the influence of the dimensionality curse.

Solved Exercises 9.5, LQ Data-Driven Design

Example 20 (Linear Quadratic Data-Driven Design)

- *Let us consider quadratic partial performance index in additive loss*

$$z(\Delta_t, A_t) = \Delta_t' \mathbb{Q}_{\Delta^\star} \Delta_t + A_t' \mathbb{Q}_{A^\star} A_t, \text{ where} \tag{215}$$

  - *observation* Δ_t *is* $\ell_{\Delta^\star}$*-dimensional real vector*
  - *action* *is* A_t *is* $\ell_{A^\star}$*-dimensional real vector*
  - *transposition is denoted* $'$
  - *the given penalisation matrices* $\mathbb{Q}_{\Delta^\star}$, $\mathbb{Q}_{A^\star}$ *are positive semi-definite, i.e.* $\mathbb{Q}_{\Delta^\star} \geq 0 \Leftrightarrow \Delta' \mathbb{Q}_{\Delta^\star} \Delta \geq 0, \ \forall \Delta \in \Delta^\star$

  This expresses DM aim *to push* Δ_t *and* A_t *jointly to zero, while the weighting balances significance of deviations from target.*

- *The assumed conditional moments of observations are*

$$\mathsf{E}[\Delta_t | A_t, \mathbf{K}_{t-1}] = \mathbb{A}\Delta_{t-1} + \mathbb{B}A_t, \ \mathsf{cov}[\Delta_t | A_t, \mathbf{K}_{t-1}] = \mathbb{R} \geq 0, \tag{216}$$

  where matrices $\mathbb{A}$, $\mathbb{B}$ *of appropriate dimensions are known and the covariance matrix* $\mathbb{R}$ *– possibly unknown – is independent of data.*

In the use of Prop. 12, we assume that the value function

$$\mathsf{V}(K_t) = \gamma_t + \Delta_t' \mathbb{S}_t \Delta_t, \ \gamma_t \geq 0, \ \mathbb{S}_t \geq 0. \tag{217}$$

- For $t = h =$ horizon, the assumption holds with $\mathbb{S}_h = 0, \ \gamma_h = 0$.
- In a generic inductive step $t < h$, let us apply (44) by evaluating

$$\mathsf{E}[\mathsf{z}(\mathbf{K}_t) + \mathsf{V}(\mathbf{K}_t) | A_t, \mathbf{K}_{t-1}] \underbrace{=}_{\tilde{\mathbb{S}}_t \equiv \mathbb{S}_t + \mathbb{Q}_{\Delta^\star}} \gamma_t + \mathsf{E}[\Delta_t' \tilde{\mathbb{S}}_t \Delta_t | A_t, \mathbf{K}_{t-1}] + A_t' \mathbb{Q}_{A^\star} A_t$$

$$\overbrace{=}^{\textit{moments}\,(58)} \overbrace{\gamma_t + \mathrm{tr}[\tilde{\mathbb{S}}_t \mathbb{R}]}^{\gamma_{t-1}} + (\mathbb{A}\Delta_{t-1} + \mathbb{B}A_t)' \tilde{\mathbb{S}}_t (\mathbb{A}\Delta_{t-1} + \mathbb{B}A_t) + A_t' \mathbb{Q}_{A^\star} A_t$$

$$\overbrace{=}^{{}^2-\textit{completion in } A_t} \gamma_{t-1} + (A_t - {}^o\!A(\mathbf{K}_{t-1}))' \mathbb{S}_{A^\star} (A_t - {}^o\!A(\mathbf{K}_{t-1})) + \Delta_{t-1}' \mathbb{S}_{t-1} \Delta_{t-1}$$

$$\mathbb{S}_{A^\star} \equiv \mathbb{B}' \tilde{\mathbb{S}}_t \mathbb{B} + \mathbb{Q}_{A^\star} > 0 \text{ if } \mathbb{Q}_{A^\star} > 0 \tag{218}$$

$$\mathbb{S}_{t-1} = \mathbb{A}' \left(\tilde{\mathbb{S}}_t - \tilde{\mathbb{S}}_t \mathbb{B} \mathbb{S}_{A^\star}^{-1} \mathbb{B}' \tilde{\mathbb{S}}_t \right) \mathbb{A} \geq 0, \qquad \text{Riccati matrix equation.}$$

- The minimising $A_t = {}^o\!A(\mathbf{K}_{t-1})$ defines the optimal decision rule

$${}^o\!A(\mathbf{K}_{t-1}) = -\,{}^o\mathbb{L}_t' \Delta_{t-1}, \quad {}^o\mathbb{L}_t' \equiv \mathbb{S}_{A^\star}^{-1} \mathbb{B}' \tilde{\mathbb{S}}_t \mathbb{A}, \quad \text{linear feedback}. \tag{219}$$

Prop. 16 provides the predictive pd $f(\Delta_t|A_t, \mathbf{K}_{t-1})$. The recursive evolution of the pd $f(X^h|\mathbf{K}_h)$ needed in (39) uses the same ingredients.

Proposition 42 (PD for Initiating Dynamic Programming)

For given $f(X_0|\mathbf{K}_0)$ *and under natural conditions of DM it holds*

$$f(X^t|\mathbf{K}_t) = f(X^{t-1}|\mathbf{K}_{t-1})\frac{\overbrace{f(\Delta_t|X_t, A_t, \mathbf{K}_{t-1})}^{\text{observation model}}\overbrace{f(X_t|X_{t-1}, A_t, \mathbf{K}_{t-1})}^{\text{time evolution model}}}{\underbrace{f(\Delta_t|A_t, \mathbf{K}_{t-1})}_{\text{predictive pd}}}. \quad (220)$$

Proof Repeat for $t = h, \ldots, 1$, $\quad f(X^h|\mathbf{K}_h) \underbrace{=}_{(25)} \frac{f(X^h, D_h|\mathbf{K}_{h-1})}{f(D_h|\mathbf{K}_{h-1})}$

$$\overbrace{=}^{(62),(63),(65)} f(X^{h-1}|\mathbf{K}_{h-1}) \times \frac{f(\Delta_h|X_h, A_h, \mathbf{K}_{h-1})f(X_h|X_{h-1}, A_h, \mathbf{K}_{h-1})}{f(\Delta_h|A_h, \mathbf{K}_{h-1})}. \quad \square$$

Solved Exercise 10.2, Two-Valued Markov Chain

Prediction and filtering for the time-invariant Markov chain with $X^\star = \Delta^\star = \{0,1\}$, $A^\star = \emptyset$, the observation model and time-evolution model given by $\alpha, \beta, a, b \in [0,1]$

$\mathsf{f}(\Delta_t \vert X_t)$	$\Delta_t = 0$	$\Delta_t = 1$
$X_t = 0$	α	$1-\alpha$
$X_t = 1$	β	$1-\beta$

$\mathsf{f}(X_{t+1} \vert X_t)$	$X_{t+1} = 0$	$X_{t+1} = 1$
$X_t = 0$	a	$1-a$
$X_t = 1$	b	$1-b$

Prediction $\mathsf{f}(X_t = 0|\mathbf{K}_{t-1}) = \mathsf{f}_{t|t-1}$ & estimate $\mathsf{f}(X_t = 0|\mathbf{K}_t) = \mathsf{f}_{t|t}$ of X_t, $\mathsf{f}_{t|t-1}, \mathsf{f}_{t|t} \in [0,1]$, evolve as follow

$$\mathsf{f}_{t|t} = \frac{[1-\alpha-\delta(\Delta_t,1)+2\alpha\delta(\Delta_t,1)]\mathsf{f}_{t|t-1}}{[1-\alpha-\delta(\Delta_t,1)\alpha+2\delta(\Delta_t,1)]\mathsf{f}_{t|t-1}+[1-\beta-\delta(\Delta_t,1)+2\beta\delta(\Delta_t,1)](1-\mathsf{f}_{t|t-1})}$$

$$\mathsf{f}_{t+1|t} = a\mathsf{f}_{t|t} + b(1-f_{t|t}),$$ where δ is Kronecker delta.