Information Exchange between Bayesian Agents

Jan Kracík

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

Outline

Bayesian Decision-Maker (Agent)

- random quantities (data): actions a_t, innovations Δ_t, experience: $\mathcal{P}(t) = (a_1, \Delta_1, \ldots, a_t, \Delta_t)$
- model of the system: $f(\Delta_t | a_t, \mathcal{P}(t-1), \Theta)$
  Θ - unknown parameter
- prior pdf: $f(\Theta)$

Bayesian Theory:

- increasing experience $\mathcal{P}(t)$ modifies knowledge about Θ

Bayes Theorem

$$f(\Theta|\mathcal{P}(t)) \propto \prod_{\tau=1}^{t} f(\Delta_\tau | a_\tau, \mathcal{P}(\tau-1), \Theta) f(\Theta)$$

- optimal decision strategy minimizes the expected loss

Bayesian Decision-Maker (Agent)

- random quantities (data): actions a_t, innovations Δ_t, experience: $\mathcal{P}(t) = (a_1, \Delta_1, \ldots, a_t, \Delta_t)$
- model of the system: $f(\Delta_t | a_t, \mathcal{P}(t-1), \Theta)$
  Θ - unknown parameter
- prior pdf: $f(\Theta)$

Bayesian Theory:

- increasing experience $\mathcal{P}(t)$ modifies knowledge about Θ

Bayes Theorem

$$f(\Theta|\mathcal{P}(t)) \propto \prod_{\tau=1}^{t} f(\Delta_\tau | a_\tau, \mathcal{P}(\tau-1), \Theta) f(\Theta)$$

- optimal decision strategy minimizes the expected loss

Multiple Agents

- 2 agents
- the same random quantities a_t, Δ_t
- different models $f_1(\Delta_t|a_t, \mathcal{P}(t-1), \Theta_1),\ f_2(\Delta_t|a_t, \mathcal{P}(t-1), \Theta_2)$
- different parameters Θ_1, Θ_2
- posterior pdfs $f_1(\Theta_1|\mathcal{P}_1(t)), f_2(\Theta_2|\mathcal{P}_2(t))$
  $\mathcal{P}_1(t), \mathcal{P}_2(t)$ consist of different realizations; not stored

cooperating agents $\rightarrow$ sharing knowledge

"How to improve $f_1(\Theta_1|\mathcal{P}_1(t))$ by $f_2(\Theta_2|\mathcal{P}_2(t))$?"

Exploiting of Probabilistically Described Knowledge in Bayesian Estimation

system model: $f(\Delta_t|a_t,\Theta)$
expert information: $h(\Delta, a)$

How to incorporate $h(\Delta, a)$ into the posterior pdf 'as a finite number of observations'?

$$f(\Theta|\mathcal{P}(t)) \propto f(\Theta)\prod_{\tau=1}^{t} f(\Delta_\tau|a_\tau,\Theta) =$$

$$= f(\Theta)\exp\left(t\int r(\Delta,a)\ln f(\Delta|a,\Theta)\mathrm{d}\Delta\mathrm{d}a\right)$$

$r(\Delta, a)$ ... empirical pdf from $\mathcal{P}(t)$

$$f(\Theta|h) \propto f(\Theta)\exp\left(t\int h(\Delta,a)\ln f(\Delta|a,\Theta)\mathrm{d}\Delta\mathrm{d}a\right)$$

Exploiting of Probabilistically Described Knowledge in Bayesian Estimation

system model: $f(\Delta_t | a_t, \Theta)$
expert information: $h(\Delta, a)$

How to incorporate $h(\Delta, a)$ into the posterior pdf 'as a finite number of observations'?

$$f(\Theta|\mathcal{P}(t)) \propto f(\Theta) \prod_{\tau=1}^{t} f(\Delta_\tau | a_\tau, \Theta) =$$

$$= f(\Theta) \exp\left(t \int r(\Delta, a) \ln f(\Delta | a, \Theta) \mathrm{d}\Delta \mathrm{d}a \right)$$

$r(\Delta, a) \ldots$ empirical pdf from $\mathcal{P}(t)$

$$f(\Theta|h) \propto f(\Theta) \exp\left(t \int h(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a\right)$$

Exploiting of Probabilistically Described Knowledge in Bayesian Estimation

system model: $f(\Delta_t|a_t, \Theta)$
expert information: $h(\Delta, a)$

How to incorporate $h(\Delta, a)$ into the posterior pdf 'as a finite number of observations'?

$$
\begin{aligned}
f(\Theta|\mathcal{P}(t)) &\propto f(\Theta) \prod_{\tau=1}^{t} f(\Delta_\tau|a_\tau, \Theta) = \\
&= f(\Theta) \exp\left(t \int r(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a \right)
\end{aligned}
$$

$r(\Delta, a)$ ... empirical pdf from $\mathcal{P}(t)$

$$f(\Theta|h) \propto f(\Theta) \exp\left(t \int h(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a \right)$$

Exploiting of Probabilistically Described Knowledge in Bayesian Estimation

system model: $f(\Delta_t|a_t, \Theta)$
expert information: $h(\Delta, a)$

How to incorporate $h(\Delta, a)$ into the posterior pdf 'as a finite number of observations'?

$$
\begin{aligned}
f(\Theta|\mathcal{P}(t)) &\propto f(\Theta) \prod_{\tau=1}^{t} f(\Delta_\tau|a_\tau, \Theta) = \\
&= f(\Theta) \exp\left(t \int r(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a \right)
\end{aligned}
$$

$r(\Delta, a)$ ... empirical pdf from $\mathcal{P}(t)$

$$
f(\Theta|h) \propto f(\Theta) \exp\left(t \int h(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a \right)
$$

Comments

$$f(\Theta|h) \propto f(\Theta) \exp\left(t \int h(\Delta, a) \ln f(\Delta|a, \Theta) \mathrm{d}\Delta \mathrm{d}a\right)$$

- $h(\Delta, a)$ is processed 'data-like'
- $h(\Delta, a)$ is taken as a finite number (t) of observations
- the method is technically feasible
- can be adapted for approximate learning (Quasi-Bayes)

Information Exchange between Bayesian Agents

- Participant 2:
  system model
  posterior pdf
  decision strategy

$$\left.\begin{array}{l}\text{system model}\\ \text{posterior pdf}\\ \text{decision strategy}\end{array}\right\} f_2(\Delta_t, a_t, |\mathcal{P}(t-1)) \rightarrow f_2(\Delta, a, \phi)$$

- Participant 1:

$$f_1(\Delta_t|a_t, \mathcal{P}(t-1), \Theta_1) \equiv f_1(\Delta_t|a_t, \phi_{t-1}, \Theta_1)$$

$$f_1(\Theta_1|f_2) \propto f_1(\Theta) \exp\left(T \int f_2(\Delta, a, \phi) \ln f_1(\Delta|a, \phi, \Theta_1)\right)$$

Information Exchange between Bayesian Agents

- Participant 2:
  system model
  posterior pdf
  decision strategy
  $\left.\begin{matrix}\text{system model}\\ \text{posterior pdf}\\ \text{decision strategy}\end{matrix}\right\} f_2(\Delta_t, a_t, |\mathcal{P}(t-1)) \rightarrow f_2(\Delta, a, \phi)$
- Participant 1:
  $f_1(\Delta_t | a_t, \mathcal{P}(t-1), \Theta_1) \equiv f_1(\Delta_t | a_t, \phi_{t-1}, \Theta_1)$

  $f_1(\Theta_1|f_2) \propto f_1(\Theta) \exp\left(T \int f_2(\Delta, a, \phi) \ln f_1(\Delta|a, \phi, \Theta_1)\right)$

Information Exchange between Bayesian Agents

- Participant 2:
  system model
  posterior pdf
  decision strategy

  $$\left.\begin{array}{l}\text{system model}\\ \text{posterior pdf}\\ \text{decision strategy}\end{array}\right\} f_2(\Delta_t, a_t, |\mathcal{P}(t-1)) \rightarrow f_2(\Delta, a, \phi)$$

- Participant 1:
  $$f_1(\Delta_t|a_t, \mathcal{P}(t-1), \Theta_1) \equiv f_1(\Delta_t|a_t, \phi_{t-1}, \Theta_1)$$

  $$f_1(\Theta_1|f_2) \propto f_1(\Theta) \exp\left(T \int f_2(\Delta, a, \phi) \ln f_1(\Delta|a, \phi, \Theta_1)\right)$$

Concluding Remarks

- information exchange via pdfs quantities in common
- based on a method for incorporating information in form of pdf of data
  - attempt to extend Bayesian theory
  - practically feasible
  - ad hoc
  - theoretical base is missing
- open problems:
  - How to avoid a repeated incorporation of the same information?
  - How to select proper T?
  - How to proceed in case of partially quantities?

Concluding Remarks

- information exchange via pdfs quantities in common
- based on a method for incorporating information in form of pdf of data
  - attempt to extend Bayesian theory
  - practically feasible
  - ad hoc
  - theoretical base is missing
- open problems:
  - How to avoid a repeated incorporation of the same information?
  - How to select proper T?
  - How to proceed in case of partially quantities?