Data-Driven FPD – Examples

Vladimíra Sečkárová

14.11.2011

Notation

- B - behaviour – generally consists of observations, actions and hidden quantities part
- Data-Driven case: $B = D^h = (\triangle^h, A^h)$ observable data records
- $\triangle^h$ – a sequence of observations up to time h including
- A^h – a sequence of actions up to time h including
- $\triangle_h$, A_h – an observation, an action respectively, at time point h

Data-Driven FPD: Solution

Proposition 20 (Solution of Data-Driven FPD)

The randomised decision rules of the optimal strategy in the data-driven FPD are

$$
\begin{aligned}
{}^{O}\mathrm{f}(A_t|D^{t-1}) &= {}^{I}\mathrm{f}(A_t|D^{t-1})\frac{\exp[-\omega_\gamma(A_t, D^{t-1})]}{\gamma(D^{t-1})} && (75)\\
\gamma(D^{t-1}) &\equiv \int_{A_t^\star} {}^{I}\mathrm{f}(A_t|D^{t-1})\exp[-\omega(A_t, D^{t-1})]\,\mathrm{d}A_t \leq 1 && (76)\\
&\quad \text{for } t < h \text{ and } \gamma(D^h) = 1 && (77)\\
\omega_\gamma(A_t, D^{t-1}) &\equiv \int_{\Delta_t^\star} \mathrm{f}(\Delta_t|A_t, D^{t-1}) \ln\left(\frac{\mathrm{f}(\Delta_t|A_t, D^{t-1})}{\gamma(D^t)\,{}^{I}\mathrm{f}(\Delta_t|A_t, D^{t-1})}\right) d\Delta_t. && (78)
\end{aligned}
$$

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

- ${}^{O}f(A_t|D^{t-1})$ – (model of) randomized decision rule
- optimal strategy: a sequence of randomized dec. rules indexed by time
- ${}^{I}f(A_t|D^{t-1})$ – ideal (model of) randomized decision rule
- $f(\triangle_t|A_t, D^{t-1})$ – observation model
- ${}^{I}f(\triangle_t|A_t, D^{t-1})$ – ideal observation model

First example

- $A_t = \{0, 1\}$
- $\triangle_t = \{0, 1\}$
- does not depend on D^{t-1}
- ${}^I f(A_t)$, $f(\triangle_t | A_t)$, ${}^I f(\triangle_t | A_t)$ are fixed
- particular choices:

$f(\triangle_t \| A_t)$	Values of A_t	
$\triangle_t$	0	1
0	0.6	0.5
1	0.4	0.5

${}^I f(\triangle_t \| A_t)$	Values of A_t	
$\triangle_t$	0	1
0	0.8	0.8
1	0.2	0.2

- ${}^I f(A_t) = (0.5, 0.5)^T$
- results: ${}^O f(A_t) = {}^O f(\{0\}, \{1\}) = (0.53, 0.47)$

Second example

- $A_t = \{0, 1\}$
- $\triangle_t = \{0, 1\}$
- $f(\triangle_t | A_t, D^{t-1})$ takes previous observation records into account
- ${}^I f(\triangle_t | A_t, D^{t-1}) = {}^I f(\triangle_t | A_t) = (0.82, 0.18, 0.76, 0.24)$ for $[0, 0], [0, 1][1, 0][1, 1]$
- ${}^I f(A_t | D^{t-1}) = {}^I f(A_t) = (0.7, 0.3)$
- optimal strategy for 3 time instants:

	${}^O f(A_t \| D^{t-1})$		
A_t	t=1	t=2	t=3
0	0.89	0.81	0.73
1	0.11	0.19	0.27