A New Approach to Estimating the Bellman Function

Jan Zeman

Institute of Information Theory and Automation
Prague, Czech Republic

September 22, 2009

Outline

Dynamic programming

System - part of world interesting for decision maker

Decision maker - human or machine with **aims** according to system

Decisions x_t is designed by decision maker to reach the aims

Output y_t is information about system available to decision maker

Gain function - degree of reaching the aims

$$G : (x_1, \ldots, x_T, y_1, \ldots, y_T) \rightarrow \mathcal{R}_0^+$$

The main aim - search the sequence $\{x_1, \ldots, x_T\}$ to maximize:

$$\max_{\{x_1,\ldots,x_T\}} G = \max_{\{x_1,\ldots,x_T\}} \sum_{k=1}^{T} g_k.$$

Dynamic programming - Bellman function

Off-line optimization at the time t:

$$\mathcal{V}_t = \max_{\{x_t,\dots,x_T\}} \sum_{k=t}^{T} g_k$$

Bellman function (recursive shape):

$$\mathcal{V}_t = \max_{x_t} (g_t + \mathcal{V}_{t+1})$$

Optimal decision:

$$x_t = \arg\max_{x_t} (g_t + \mathcal{V}_{t+1})$$

Dynamic programming - Bellman function

On-line optimization at the time t:

$$\mathcal{V}_t = \max_{\{x_t,\dots,x_T\}} E\left(\sum_{k=t}^{T} g_k \quad \middle| x_1,\dots,x_{t-1},y_1,\dots,y_t\right)$$

Bellman function (recursive shape):

$$\mathcal{V}_t = \max_{x_t} \quad E\left(g_t + \mathcal{V}_{t+1} \quad |x_1,\dots,x_{t-1},y_1,\dots,y_t\right)$$

Optimal decision:

$$x_t = \arg\max_{x_t} \quad E\left(g_t + \mathcal{V}_{t+1} \quad |x_1,\dots,x_{t-1},y_1,\dots,y_t\right)$$

Off-line strategy analysis

Assumption

- The decision x_t does not influence the system.

The optimal strategy obtained at whole dataset:

$$X^T = (x_1^T, x_2^T, \dots, x_T^T)$$

Optimal strategies for shorter horizon:

$$\begin{aligned} X^1 &= (x_1^1) \\ X^2 &= (x_1^2, x_2^2) \\ X^3 &= (x_1^3, x_2^3, x_3^3) \\ &\vdots \\ X^T &= (x_1^T, x_2^T, x_3^T, x_4^T, \dots, x_T^T) \end{aligned}$$

Similarity indexes

Similarity index (number of similar decisions):

$$S_t = \sum_{i=1}^{t} \delta(x_i^t, x_i^T)$$

Strict similarity index (length of non-broken similarity):

$$s_t = \max_i \{i; (\forall j \in \mathcal{N})(j \leq i \Rightarrow x_j^t = x_j^T)\}$$

$$s_t \leq S_t \leq t$$

Example:

	1	2	3	4	5	6	7	8 . . . t	
$X^t =$ {	1	-1	0	1	0	1	0	0 . . . 0	}
$X^T =$ {	1	-1	0	1	1	1	0	1 . . . 1	. . . }

$$s_t = 4, \qquad S_t = 6$$

Estimating the Bellman function

Bellman function (recursive shape):

$$\mathcal{V}_t = \max_{x_t} \quad E\left(g_t + \mathcal{V}_{t+1} | x_1, \dots, x_{t-1}, y_1, \dots, y_t\right)$$

If $s_t \approx t$ and we insert the X^t, we obtain system of functional equations:

$$\mathcal{V}_k = \max_{x_k} \quad E\left(g_k + \mathcal{V}_{k+1} | x_1^t, \dots, x_{k-1}^t, y_1, \dots, y_k\right) \quad \text{for } k \in \{1, \dots s_t\},$$

which can be transformed to system algebraic equations for parametrized shape of Bellman function.

Solution is inserted into on-line Bellman equation and the maximization can be calculated.

Futures trading: task definition

Futures futures contract, obligation to buy a normalized amount of a commodity

Speculator chooses the decision about the future

- **Long** - believe in price increase
- **Short** - believe in price decrease
- **Flat** - do not believe - out of market

Gain function:

$$G = \sum_{k=1}^{T} \underbrace{(y_k - y_{k-1})x_{t-1} - C|x_{k-1} - x_k|}_{g_k},$$

	$y_1, \ldots, y_T$	price of contract
where	$x_1, \ldots, x_T$	count of held contracts
	C	transaction cost per contract

Calculated variables

Five reference price sequences:

- Cocoa (CC)
- Light crude oil (CL)
- U.S. Treasure note (FV2)
- Japanese Yen (JY)
- Wheat (W)

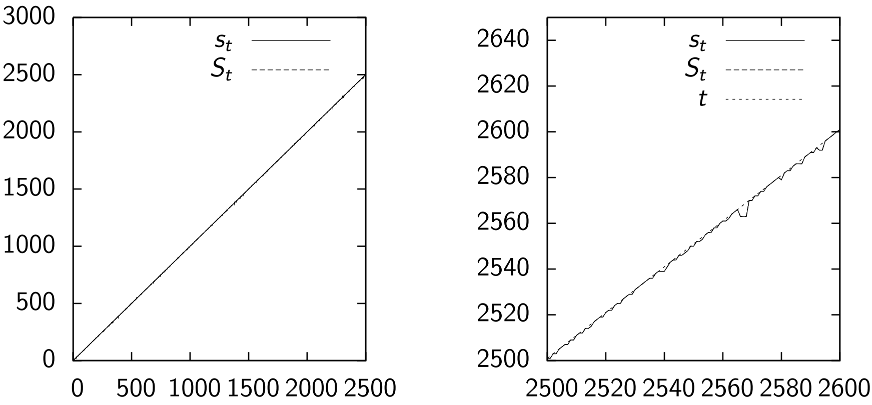
Calculated variables:

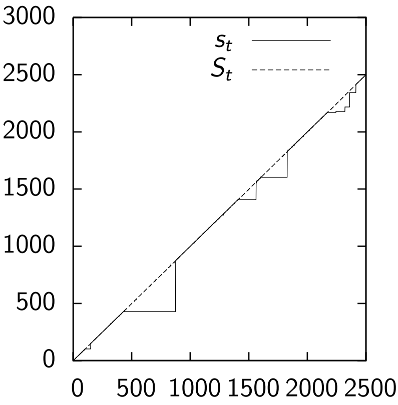
- $c_1 = max_t(t - s_t)$
- $c_2 = max_t(t - S_t)$
- $t_{ch,1}$ and $t_{ch,2}$ - last change of value c_1 and c_2

Table of results

Market	c_1	c_2	$t_{ch;1}$	$t_{ch;2}$	T
CC	6	6	342	342	3822
CL	444	6	847	2205	3863
FV2	8	8	383	383	3766
JY	4	4	50	50	3871
W	7	7	2452	2452	3822

Table: Dominating constants c_1 and c_2


3000
2500
2000
1500
1000
500
0
0
500
1000
1500
2000
2500
s_t
S_t
2640
2620
2600
2580
2560
2540
2520
2500
2500
2520
2540
2560
2580
2600
s_t
S_t
t


3000
2500
2000
1500
1000
500
0
0
500
1000
1500
2000
2500
s_t
S_t
2300
2280
2260
2240
2220
2200
2180
2160
2160
2180
2200
2220
2240
s_t
S_t
t

Experiment setup

Experiment:

- Parametrized shape of Bellman function
- Weighted least squares method
- Prediction using autoregressive model

$\Rightarrow$ iteration spread in time (IST)

Reference results:

- Model predictive method (MPC)
- Predictive model and task setup were same as above

Results

Market	MPC	IST
CC	-6 450	-1 490
CL	-12 350	3 390
FV2	-5 701	10 727
JY	-26 568	-35 247
W	-9 792	-1 923

Table: Results of experiment

Conclusion

- IST is better in 4 of 5 datasets.
- The approach needs further testing.
- The similarity indexes are calculated non-causally.
- The causal criterion of usage the approach should be find.