Estimation of thresholds in optimal stopping problems via a Cross-Entropy method

T. D. Kularatne  G. Sofronov

Department of Statistics
Macquarie University
Sydney, Australia

International Conference on Robust Statistics (ICORS 2017)
4th July, 2017
Outline

1. Introduction
   - Background and Aims
   - Statement of Problem

2. Cross-Entropy Method
   - General CE Algorithm for Optimization
   - CE Algorithm for Noisy Optimization

3. Numerical Results

4. Future Research

5. References
Background and Aims

- Secretary Problem (Marriage Problem/Dowry Problem)
  - Problem was introduced by Martin Gardner in 1960’s [Ferguson, 1989].

Source(http://previews.123rf.com)
Background and Aims

- **Sequential data**
  - Collection of records which are ordered with respect to time.
  - Observations appear one by one, and the data are analysed as they are collected without fixing the sample size in advance.
  - Decisions are made on already obtained information while future observations are not known yet.

- **Applications of sequential data analysis**
  1. In finance
  2. In industrial quality control
Aim of our study:
- Observe a sequence of random variables, which can be interpreted as the value of an asset at a particular time.
- Decide when we must stop, given that no recall allowed.
- Objective is to maximize an expected reward.

In this study, we focus on constructing a Cross-Entropy method to find an approximate optimal stopping rule using Monte Carlo simulation.
Statement of Problem

Consider an optimal stopping problem for finite discrete time stochastic sequences.
- We observe $X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i$ sequentially and we must decide either to stop and accept the value $x_i$ at time $i$ ($i = 1, \ldots, N - 1$) or continue and observe $X_{i+1}$.

- We denote $V_n(x)$ as the gain from the optimal procedure and $v_n$ as the expected gain with $n$ observations remained.

- Consider ‘sampling without recall’ from a known distribution $F$. 
Statement of Problem

- According to [DeGroot, 1970] the expected gain $v_n$ can be expressed as

$$v_n = E[V_{n-1}(X)] = \int_{-\infty}^{\infty} V_{n-1}(x)dF(x)$$

The gain from the optimal procedure is

$$V_n(x) = \max\{x, v_n\}$$

- We follow sampling without recall therefore, we must stop and accept the final observed value $x$ as our gain. Hence,

$$V_0(x) = x$$

Our aim is to find a stopping rule which maximizes the expected gain $v_N$. 
Statement of Problem

As an example,
- Consider a sequential random sample \( X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N \) taken from \( U(a, b) \).
- If \( v_0 = a \) and \( v_1 = \frac{(a+b)}{2} \), then considering above equations we can write

\[
v_2 = E[V_1(X)] = \int_a^b \max\{x, v_1\} \frac{1}{(b-a)} dx
\]

Since

\[
\max\{x, v_1\} = \begin{cases} x, & x \geq v_1 \\ v_1, & x < v_1. \end{cases}
\]

Therefore, in general we can write the expected gain (value of threshold) at \((n+1)\)-th time as

\[
v_{n+1} = \frac{1}{2(b-a)}[b^2 - 2v_na + v_n^2]
\]
Statement of Problem

- If $X_i \sim U(0, 1)$
- Then $v_0 = 0$ and $v_1 = 0.5$
- Therefore we can obtain

$$v_2 = \frac{1}{2} [1 + (0.5^2)] = 0.625$$

$$v_3 = \frac{1}{2} [1 + (0.625^2)] = 0.695$$

...
The Cross-Entropy (CE) method is an adaptive Monte Carlo (MC) approach introduced by Reuven Y. Rubinstein in 1999.

The method was first developed as an adaptive algorithm for rare event simulation using variance minimization technique.

Then the method was further modified not only for estimating rare event probabilities but also for solving complex combinatorial, continuous and multi-extremal optimization problems.

The CE method can be applied to both estimation and optimization problems.
The CE method is an iterative procedure, which can be summarized using the following three steps.

- **Step 1:** Generate a random sample of objects (vector of parameters) according to a statistical distribution.
- **Step 2:** Obtain the best performing sample of objects using the performance (objective) function.
- **Step 3:** Update the parameters of the statistical distribution in Step 1, using the sample from Step 2, to produce a better sample at the next iteration.
- Consider $\mathcal{X}$, an arbitrary set of states and a real valued performance function $S$ on $\mathcal{X}$.

- The objective of this optimization problem is to find the maximum of $S$ over $\mathcal{X}$ and the corresponding maximizer $\gamma^*$.

$$\gamma^* = \max_{x \in \mathcal{X}} S(x).$$

- Convert deterministic problem to an associate stochastic problem (ASP).
- Let ASP be defined as

\[ P_u(S(X) \geq \gamma) = E_u I\{S(X) \geq \gamma\} \]

- Can be estimated by using a log-likelihood estimator with parameter \( u \),

\[
\hat{u}^* = \arg \max_u \frac{1}{N_{elite}} \sum_{i=1}^{N_{elite}} I\{S(X_i) \geq \gamma\} \ln f(X_i, u).
\]

where \( N_{elite} \leq M \) is the size of the candidate solutions simulated from the statistical distribution \( f(\cdot; u) \) where \( S(X) \geq \gamma. \)
1. Initialization
- Define an initial parameter vector $\hat{u}_0$ of the statistical distribution

2. Adaptive updating of $\gamma_t$
- Generate $X_1, X_2, \ldots, X_M$ from the density $f(\cdot; u)$.
- Calculate the performances $S(X_1), S(X_2), \ldots, S(X_M)$.
- Sort the values in an increasing order $S(1) \leq S(2) \leq \cdots \leq S(M)$. 

Kularatne T.D., Sofronov G. ICORS 2017 14 / 28
Let \( \gamma_t \) be the \((1 - \rho)\)-quantile of \( S(X) \) under \( u_{t-1} \) which satisfies

\[
\Pr_{u_{t-1}}(S \geq \gamma_t) \geq \rho, \\
\Pr_{u_{t-1}}(S \leq \gamma_t) \geq 1 - \rho.
\]

The estimate of \( \gamma_t \) is

\[
\hat{\gamma}_t = S\left(\lceil (1 - \rho)M \rceil \right).
\]
General CE Algorithm for Optimization

3. Adaptive updating of $u_t$

- For fixed $\gamma_t$ and $u_{t-1}$, obtain estimates for $u_t$ applying the CE method to the optimization problem

$$\hat{u} = \arg \max_u \mathbb{E}_{u_{t-1}} I_{\{S(X) \geq \gamma_t\}} \ln f(X, u).$$

- Update the parameters of the statistical distribution using the best performing (elite) sample.

$$\hat{u} = \arg \max_u \frac{1}{N_{\text{elite}}} \sum_{i=1}^{N_{\text{elite}}} I_{\{S(X_i) \geq \hat{\gamma}_t\}} \ln f(X_i, u).$$
4. Smoothed updating of $u_t$ (optional)

For the constant smoothing coefficient defined as $\alpha (0 \leq \alpha \leq 1)$. The smoothed update of $\hat{u}_t$ can be obtained as

$$\hat{u}_t = \alpha \hat{u}_t + (1 - \alpha) \hat{u}_{t-1}$$

5. Repeat until the stopping criterion is met and obtain the solution which maximizes the performance function.
CE Algorithm for Noisy Optimization

- CE method is effective to handle noisy optimization (stochastic optimization).
  i.e. When objective function value obtained via simulation.

- For the maximization problem, consider our performance function \( S(x) \) corrupted with noise.

\[
\hat{S}(x) = \hat{S}(x, Y).
\]

- Our aim is to solve the maximization problem

\[
\max_{x \in X} \mathbb{E} \hat{S}(x, Y).
\]
CE Algorithm for Noisy Optimization

CE method for optimal stopping problem

- Consider a problem of finding a set of thresholds that maximizes the value of the game for an optimal stopping problem.

- For random variables $Y_1, Y_2, \ldots, Y_N$ with a pdf $F$, we can write the maximization problem to obtain the set of thresholds $x^* = (x_1^*, x_2^*, \ldots, x_N^*)$

$$\max_{x \in \mathcal{X}} \mathbb{E} \hat{S}(x, Y).$$
For this study,

- \( N = 7 \) consecutive time points.

- \( K = 1000 \) simulation paths.

- Obtain the set of thresholds \( \nu^* = (\nu_1^*, \nu_2^*, \ldots, \nu_N^*) \) at each time point and the value of the game \( \nu \) with an objective function is to maximise the expected gain.
Numerical Results

- Developed two versions of CE method for the optimal stopping problem.
  - Algorithm for non-ordered thresholds
  - Algorithm for ordered thresholds

- Considered
  - 50 repetitions of different sequences of random variables from $U(0, 1)$.
  - Generate samples of $X$ from Normal (N), Truncated Normal (TN) and Beta (B) distributions.

- The accuracy of the estimated set of thresholds and the value of the game was checked using the root mean squared error (RMSE).
Numerical Results

Graph showing the value of thresholds against the number of observations (n) for different categories: CE-B, CE-N, CE-TN, and True.
Numerical Results

![Graph showing the value of thresholds against n for different methods: CE-Bo, CE-No, CE-TNo, and True. The graph illustrates the trend as n increases.]
Numerical Results

RMSE for the two versions of algorithm
## Numerical Results

### Comparison between two versions of CE algorithm

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Iteration Count</th>
<th>Running time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>CE-B</td>
<td>14</td>
<td>48</td>
</tr>
<tr>
<td>CE-Bo</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td>CE-N</td>
<td>14</td>
<td>33</td>
</tr>
<tr>
<td>CE-No</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>CE-TN</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>CE-TNo</td>
<td>11</td>
<td>33</td>
</tr>
</tbody>
</table>
Future Research

- Develop other numerical techniques including MCMC, linear programming and dynamic programming methods [Cho and Stockbridge, 2002].
- Modify gain function including reward functions or cost functions.
- Consider dependent observations from distributions like Geometric Brownian Motion [Longstaff and Schwartz, 2001].
- Generalizing the study for optimal stopping problems involving more than one stop (multiple stopping problem) [Sofronov, 2016].
- Application of sequential change-point detection problem.
References

Thank You.