Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

INFORMATION DIRECTED SAMPLING FOR STOCHASTIC ROOT FINDING

Sergio Rodriguez
Michael Ludkovski

Department of Statistics and Applied Probability
University of California, Santa Barbara
Santa Barbara, CA 93106-3110, USA

ABSTRACT

The Stochastic Root-Finding Problem (SRFP) consists of finding the root x^* of a noisy function. To discover x^*, an agent sequentially queries an *oracle* whether the root lies rightward or leftward of a given measurement location x. The oracle answers truthfully with probability $p(x)$. The Probabilistic Bisection Algorithm (PBA) pinpoints the root by incorporating the knowledge acquired in oracle replies via Bayesian updating. A common sampling strategy is to myopically maximize the mutual information criterion, known as Information Directed Sampling (IDS). We investigate versions of IDS in the setting of a non-parametric $p(x)$, as well as when $p(\cdot)$ is not known and must be learned in parallel. An application of our approach to optimal stopping problems, where the goal is to find the root of a timing-value function, is also presented.

1 PROBABILISTIC BISECTION FOR STOCHASTIC ROOT FINDING

Let x^* be the realized value of a random variable X^* with density f_0 supported over $[0,1]$. To learn x^*, points (X_n) are sequentially measured to observe the random sequence $(Z_n(X_n))$ so that $Z(x) = g(x) + \varepsilon(x)$; where ε is a zero-median stochastic noise term and g is *monotone* on $[0,1]$. The PBA as in Waeber et al. (2013) considers $Y_n(x) = \text{sign}(Z_n(x)) \in \{-1,1\}$ as noisy oracle replies which informs whether the root x^* lies to the left or right of x with probability $p(x)$ of this *direction* being correct. Let $\mathscr{F}_n = \sigma(X_n, Y_n(X_n))$ be the σ-algebra generated by the sequence of n sampling points and oracle replies. Based on $\mathscr{F}_n$, the primary objective is to decide at which site x_{n+1} to query the oracle next, such that the long-run uncertainty about X^* is minimized. The latter is quantified through the the posterior density of X^*, $f_n(u) \equiv \mathbb{P}[X^* \in du | \mathscr{F}_n]$. Given a prior f_0 of X^*, Waeber et al. (2013) show that f_n can be updated sequentially using Bayesian methods. Two practically relevant metrics of learning X^* are the posterior entropy $\text{Entr}(f_n)$ and its inter-quartile range $\text{IQR}(f_n)$. The median or the mean of f_n can also be used to obtain a point estimate of X^*.

1.1 Information Directed Sampling

IDS is a myopic policy which queries sites that maximize the *conditional mutual information* $I_n(x) := I(Y_n(x); X^* | X_n = x, f_n)$ between X^* and oracle replies, i.e. $x_{n+1} \in \arg\max_{x\in[0,1]} I_n(x)$. If $p(x) \equiv p \in (1/2, 1)$ the latter criterion is equivalent to sampling at the median of f_n (Jedynak et al. 2012). However, sampling at the median is not suitable when $p(\cdot)$ depends on x (even when $p(\cdot)$ is known and observable) since typically $p(x) \to 1/2$ as $x \to x^*$, causing oracle responses to provide minimal new information about X^* and, consequently, a poor reduction in the overall uncertainty of X^*. In contrast, we show below that IDS avoids this difficulty by staying away from regions where $p(x) \simeq 1/2$.

2 ROOT FINDING IN OPTIMAL STOPPING PROBLEMS

Let $X \equiv X_{1:T}$ be a discrete-time real-valued Markov process. Let $\mathscr{G} = (\sigma(X_{1:t}))$ be the filtration generated by X and $\mathscr{S}$ the collection of all $\mathscr{G}$-stopping times smaller than $T < \infty$. The *Optimal Stopping Problem*

(OSP) consists of maximizing the expected reward $h_\tau(X_\tau)$ over $\tau \in \mathscr{S}$. Define the value function $V(t,x) := \sup_{\tau \geq t, \tau \in \mathscr{S}} \mathbb{E}[h_\tau(X_\tau)|X_t = x]$ for $0 \leq t \leq T$. We have that $V(t,x) = h_t(x) + \max\{T(t,x), 0\}$ where $T(t,x) := \mathbb{E}[V(t+1, X_{t+1})|X_t = x] - h_t(x)$ is the *timing value*. Gramacy and Ludkovski (2015) show that solving the OSP at stage t is equivalent to finding the *roots* of $T(t,x)$; frequently a priori structure implies a unique root. Moreover, a simulated path $x_{t:T}$ and corresponding path-wise stopping time $\tau \equiv \tau(t+1, x_{t:T})$ yields a realization of $z_t(x_t) := h_\tau(x_\tau) - h_t(x_t) = T(t, x_t) + \varepsilon(t, x_t)$.

The latter equality offers a stochastic sampler that maps inputs x into random outputs $h_\tau(X_\tau) - h_t(x)$ centered around the true conditional expectation $T(t,x)$. We use PBA to find the root of $T(t,\cdot)$ and hence construct a novel algorithm for OSP.

Figure 1 shows an application of PBA in the context of an American Put Option problem where $h_t(x) \equiv e^{-rt}(K-x)_+$ and X is a log-normal random walk (classical Black-Scholes model). In the Figure $K = 40$, the true root (known as the stopping boundary for the Put) is $x^* \simeq 36.00$, and we implemented the IDS and median-sampling policies, assuming a known, but non-parametric $x \mapsto p(x)$ setting and a uniform prior $X^* \sim \text{Unif}[25, 40]$. As can be seen, the IDS policy is successful in learning about X^*, with $\text{IQR}(f_N) = 0.000166$ after $N = 1{,}000$ oracle calls. In contrast to the IDS policy that keeps $p(x_n)$ away from 1/2 to consistently gain information on X^*, sampling at the median fails to shrink the posterior IQR as $p(x_n)$ rapidly goes to 1/2 after a few iterations.

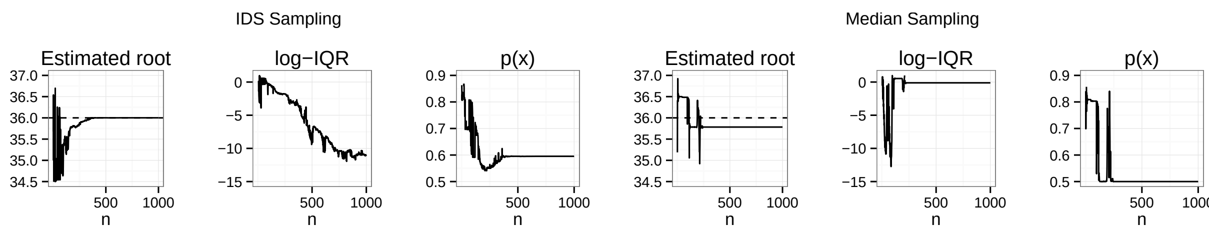

Figure 1: Sampling policy comparison: first three panels IDS; last three panels sampling at median(f_n).

3 CONCLUSIONS AND ONGOING RESEARCH

As seen above, the IDS policy performs well in shrinking posterior uncertainty about the root location. However, the basic definition of IDS relies heavily on knowing $p(x)$, i.e. it is too greedy. In the realistic context of unknown $p(\cdot)$, we propose several extensions of IDS to better handle the exploration aspect, namely sampling at new locations in order to further learn $p(\cdot)$. To this end, we have designed (i) randomized policies that enforce exploration by selecting x_{n+1} non-deterministically and actively (e.g. sampling randomly at the quantiles of f_n, i.e. $x_{n+1} = F_n^{-1}(q)$ where $q \sim \text{Unif}[0,1]$ and $F_n(\cdot)$ the cdf of f_n), as well as (ii) *batched* sampling that repeatedly queries a fixed site x M-times to simultaneously learn $p(x)$ and to update f_n. We also observe that typically the mutual information function $x \mapsto I_n(x)$ has two local maxima on each side of x^*, allowing to approximate its maximization via a simple criterion of the form $x_{n+1} = \arg\max\{I_n(F_n^{-1}(q_1)), I_n(F_n^{-1}(q_2))\}$, with each quantile $F_n^{-1}(q_i)$ chosen (randomly) to straddle the median of f_n. Extensive numerical experiments (*work in progress*) will be presented to illustrate and compare these proposals both on synthetic data and for the American Put application above.

REFERENCES

Gramacy, R. B., and M. Ludkovski. 2015. "Sequential design for optimal stopping problems". *SIAM Journal on Financial Mathematics* 6 (1): 748–775.

Jedynak, B., P. I. Frazier, R. Sznitman et al. 2012. "Twenty questions with noise: Bayes optimal policies for entropy loss". *Journal of Applied Probability* 49 (1): 114–136.

Waeber, R., P. I. Frazier, and S. G. Henderson. 2013. "Bisection search with noisy responses". *SIAM Journal on Control and Optimization* 51 (3): 2261–2279.