Near-Optimal Target Learning With Stochastic Binary Signals

Mithun Chakraborty, Sanmay Das, Malik Magdon-Ismail,
Department of Computer Science, Rensselaer Polytechnic Institute

The Learning Problem

We wish to learn, within error tolerance \( \varepsilon \), an unknown value \( \bar{V} \in \mathbb{R} \), given access only to sequential binary thresholded observations under an additive Gaussian noise model. At \( t = 0, 1, 2, \ldots \), we maintain a probability distribution \( p_t(x) \) over \( \bar{V} \) and set a threshold \( \theta \). The observation (stochastic binary signal) is

\[
x_t = \text{sign}(V + z_t - \theta)
\]

where

\[
z_t \sim_{i.i.d.} \mathcal{N}(0, \sigma^2_t)
\]

Fig 1. The effect of noise increases (i.e. probability of getting a noisy signal becomes greater than 1/2) as our estimate gets closer to the target.

Starting prior: Gaussian \( \mathcal{N}(0, \theta^2) \). Define the information disadvantage \( \rho_t = \log \frac{p_t(x)}{p_t(\bar{V})} \) (measure of the harshness of the learning environment).

Myopic threshold: \( \theta = \rho_t \), the mean of the belief distribution.

Bayesian setting: The probability estimate is updated as:

\[
p_{t+1}(x) = \int_{-\infty}^{\infty} \frac{p_t(v) \Phi(v|\bar{V} - z_t)}{\Phi(x|\mu_t, \sigma_t)} \, dv
\]

Bayesian Inference on Exact Signals (Optimal)

If \( f_t \) are available, we maintain a Gaussian belief (simplified scalar Kalman filter) with parameters

\[
\mu_{t+1} = \frac{1 - \frac{1}{2} \rho_t}{1 + \frac{1}{2} \rho_t} \mu_t + \frac{1 + \frac{1}{2} \rho_t}{1 + \frac{1}{2} \rho_t} \bar{V},
\]

\[
\sigma_{t+1}^2 = \frac{1}{1 + \frac{1}{2} \rho_t} \sigma_t^2.
\]

Theorem 1. For \( t < \frac{1}{2} \), \( 0 \leq \rho_t < 1 \). For \( t > \frac{1}{2} \), \( \frac{1}{2} \leq \rho_t \leq 1 \).

Bayesian Inference on Stochastic Binary Signals

Exact Bayesian inference on thresholded signals is analytically intractable. Alternative: Numerical Integration.

- Efficiency: \( O(N) \) computations for posterior \( p_t(x) \) with \( N \) quadrature points.
- Numerical instability

Solution: Non-parametric, discrete, finite distribution as a near-exact benchmark.

Update rules

Approximate Gaussian inference

\[
\rho_{t+1} = \rho_t + \frac{1}{N} \left( \frac{1}{2} \sigma_t^2 - 2 \rho_t \mu_t + \bar{V} \right)
\]

Truncation (\( \rho_t > \rho^* \)):

\[
l_t = \frac{1}{N} \left( \frac{1}{2} \sigma_t^2 - 2 \rho_t \mu_t + \bar{V} \right)
\]

Collapse (truncated Gaussian to Gaussian):

\[
\rho_{t+1} = \rho_t + \frac{1}{N} \left( \frac{1}{2} \sigma_t^2 - 2 \rho_t \mu_t + \bar{V} \right)
\]

Convergence properties (for \( \rho_0 < 1 \))

Theorem 2. There exist absolute positive constants \( C > 0 \) and \( \delta \), \( 1 \leq \delta < \pi^2/4 \approx 3.14 \), such that: if \( t > C/|\rho_0^2| \), then \( |V| - |\bar{V}| < \varepsilon \).

Thus, the waiting time for the expected mean belief to get within \( \varepsilon \) of the target \( V \) is \( O(1/\rho_0^2) \) while the corresponding time bound for exact signals is \( O(1/\rho_0^2) \). This shows that the dependence of our algorithm on \( \rho_0 \) is optimal and algorithm is polynomial in \( 1/\varepsilon \).

Experimental results

References
