Optimization of two-alternative batch data processing

A V Kolnogorov
Yaroslav-the-Wise Novgorod State University, Velikiy Novgorod, Russia

Email: Alexander.Kolnogorov@novsu.ru

Abstract. We consider optimization of batch data processing if there are two alternative processing methods available with different unknown efficiencies. One should determine more efficient method and provide its predominant usage. Formally, the problem is presented as Gaussian two-armed bandit problem with a priori unknown mathematical expectations and variances of incomes. We consider the problem in robust (minimax) setting. According to the main theorem of game theory, minimax strategy and minimax risk are sought for as Bayesian ones corresponding to the worst-case prior distribution of parameter. We describe the properties of the worst-case prior distribution and present corresponding recursive equations for determining Bayesian risk and expected losses. Some numerical examples are presented. We show that the control performance almost does not depend on the number of processed batches if this number is large enough.

1. Introduction
We consider optimization of the two-alternative batch data processing in the framework of Gaussian two-armed bandit. The name of the problem originates from the slot machine with two arms [1]. The choice of each arm of the two-armed bandit is accompanied with a random income which distribution depends only on chosen arm and is unknown to the agent. The goal of the agent is to maximize the total expected income gained on some control horizon N (or to minimize the total expected income; in this case incomes are usually called losses). To this end, the agent should determine the most profitable arm during the control process and provide its predominant application. The problem has numerous applications in engineering and information sciences, see, e.g. [2-4].

Formally, Gaussian two-armed bandit is a random controlled process $\xi_n$, $n=1,2,...,N$, which values are interpreted as incomes, depend only on currently chosen actions $y_n$ and have Gaussian (normal) probability distribution density $f_{D_n}(x|m_i)$ if currently chosen action is $y_n = \ell (\ell \in \{1,2\})$. Here $f_{D_n}(x|m) = (2\pi D)^{-1/2} \exp\left((x-m)^2/(2D)\right)$. Let’s explain Gaussian distributions of incomes. Let $\eta_t$, $t=1,2,...,T$ be the results of processing of $T=M \times N$ primary data. For example, $\eta_t = 1$ if data item number $t$ is successfully processed and $\eta_t = 0$ otherwise. In this case, the goal is to maximize the total expected number of successfully processed data. Or $\eta_t$ is a duration of processing the $t$-th data item and the goal is to minimize the expected total computer time of processing. We partition all the data into $N$ batches each containing $M$ data items and use the same methods for processing in the same batches. For the control, we use the values of the process $\xi_n = M^{-1/2} \sum_{t=1}^{nM} \eta_t$, $n=1,2,...,N$, i.e. cumulative incomes in the batches. According to the central limit theorem, distributions of $\xi_n$, $n=1,2,...,N$ are close to Gaussian and this implies the universality of considered
setting of the problem. It is important that in many cases processing in the same batches can be implemented in parallel. Therefore, the total duration of the processing depends on the number of batches rather than on the total number of data.

We consider the robust (minimax) setting of the problem. According to this setting the values of the properly assigned loss function must not exceed some limiting magnitude called minimax risk. We use the main theorem of the game theory and search for minimax strategy and minimax risk as for Bayesian ones corresponding to the worst-case prior distribution. This formal setup is presented in Section 2. The worst-case prior distribution is asymptotically uniform. This makes it possible to write a recursive equation for determining in a sequential way corresponding Bayesian strategy and Bayesian risk. The worst-case prior and recursive equations are presented in Section 3. Numerical results are presented in Section 4.

2. Strategy, loss function, minimax and Bayesian risks

At first, let’s assume that mathematical expectations of the incomes \(m_1, m_2\)are unknown and the variances \(D_1, D_2\) are known to the agent. Therefore, the two-armed bandit can be described by a vector parameter \(\theta = (m_1, m_2)\). Later on, the assumption of known \(D_1, D_2\) can be omitted if the number of processed data is large enough. We assume that admissible set of parameters is the following: \(\Theta = \{\theta \mid |m_1 - m_2| \leq 2C\}, 0 < C < \infty\).

A control strategy \(\sigma\) assigns generally a random choice of the action at the point of time \(n\) depending on currently observed history of the process, i.e. cumulative numbers of both actions applications \(n_1, n_2\) and corresponding cumulative incomes \(X_1, X_2\). If the agent knew \(m_1, m_2\) he should always choose the action corresponding to the largest of them, his total expected income on the control horizon \(N\) would thus be equal to \(N \max(m_1, m_2)\). But if he uses some strategy \(\sigma\) his total expected income is less than maximal one by the value

\[
L_N(\sigma, \theta) = N \max(m_1, m_2) - E_{\sigma, \theta} \left( \sum_{n=1}^{N} \xi_n \right)
\]

which is called the loss function. Here \(E_{\sigma, \theta}\) is the sign of mathematical expectation with respect to the measure generated by strategy \(\sigma\) and parameter \(\theta\). Note that sometimes the minimization of total expected income is considered as the goal. Corresponding loss function can be reduced to (1) by considering the process \((-\xi_n)\) \(n = 1, 2, \ldots, N\).

By the loss function (1) the minimax risk is defined as follows

\[
R^M_N(\Theta) = \inf_{\sigma} \sup_{\theta} L_N(\sigma, \theta),
\]

corresponding optimal strategy \(\sigma^M\) is called minimax strategy. Note that if \(\sigma^M\) is applied then inequality \(L_N(\sigma^M, \theta) \leq R^M_N(\Theta)\) is satisfied for all \(\theta \in \Theta\) and this implies robustness of the control.

Let \(\lambda(\theta)\) be a prior distribution density on the set of parameters \(\Theta\). The value

\[
R^B_N(\lambda) = \inf_{\{\sigma\}} \int L_N(\sigma, \theta) \lambda(\theta) d\theta
\]

is called the Bayesian risk, corresponding optimal strategy \(\sigma^B\) is called Bayesian strategy. Bayesian approach makes it possible to find Bayesian strategy and Bayesian risk by solving recursive equation. Minimax and Bayesian approaches are integrated by the main theorem of game theory according to which the following equality holds
\[ R^M_N(\Theta) = R^B_N(\lambda^0) = \sup_{\lambda} R^B_N(\lambda), \]  

i.e. minimax risk is equal to Bayesian one calculated with respect to the worst-case prior distribution of the parameter and minimax strategy coincides with corresponding Bayesian strategy. However, it is important to understand that a direct usage of equality (4) is virtually impossible because of the huge calculation complexity. The analysis of the properties of the worst-case prior makes it possible substantially to simplify the problem.

3. Recursive equations

The results of this section are presented according to [5]. In what follows it is convenient to modify the parameterization. Let’s put \( m_1 = m + \nu, \) \( m_2 = m - \nu. \) Then the set of parameters becomes the following \( \Theta = \{ \theta : |\nu| \leq C \}. \) Asymptotically the worst-case prior distribution density is of the form

\[ \lambda^0(\theta) = \kappa_m(m) \times \rho(\nu), \]  

where \( \kappa_m(m) = (2a)^{-1} \) is a uniform distribution on the interval \( m \in (-a, a) \) and \( a \to \infty. \)

In the sequel, we restrict consideration with the strategies which apply the same actions \( M \) times in succession because just these strategies are used for batch data processing. Denote \( n_{i\ell}^* = n_i \times D_{\ell}, \) \( M_{\ell}^* = M \times D_{\ell}, \) \( n_{i\ell} = n_i / D_{\ell}, \) \( M_{\ell} = M / D_{\ell} \) \( (\ell = 1, 2), \) \( n = n_1 + n_2, \) \( n = n_1 + n_2, \) \( U = (X_1 \times n_2 - X_2 \times n_1) / n. \) Define \( D_g, \) \( D_h \) by conditions \( D_g^2 = D_1 D_2, \) \( D_h^{-1} = 0.5(D_1^{-1} + D_2^{-1}). \) Let us consider the recursive equation

\[ R(U, n_1, n_2) = \min \left( R^{(1)}(U, n_1, n_2), R^{(2)}(U, n_1, n_2) \right), \]  

where \( R^{(1)}(U, n_1, n_2) = R^{(2)}(U, n_1, n_2) = 0 \) if \( n = N \) and then

\[ R^{(1)}(U, n_1, n_2) = Mg^{(1)}(U, n_1, n_2) + R(U, n_1 + M, n_2) \ast f_m n_{1\ell}^2 (n_{1\ell}^\prime) + n_{1\ell}^\prime (n + M_{\ell})^{-1} (U), \]  

\[ R^{(2)}(U, n_1, n_2) = Mg^{(2)}(U, n_1, n_2) + R(U, n_1, n_2 + M) \ast f_m n_{2\ell}^2 (n_{2\ell}^\prime) + n_{2\ell}^\prime (n + M_{\ell})^{-1} (U), \]  

if \( 2M \leq n < N. \) The sign ‘\( \ast \)’ denotes convolution. Here we will obtain:

\[ g^{(1)}(U, n_1, n_2) = \left\{ \begin{array}{ll} 0 & \text{if } 2 < 21 |U| \rho(\nu) d\nu; \\ C & \text{otherwise} \end{array} \right. \]

\[ g^{(2)}(U, n_1, n_2) = \left\{ \begin{array}{ll} 0 & \text{if } 2 < 21 |U| \rho(\nu) d\nu; \\ C & \text{otherwise} \end{array} \right. \]

Bayesian strategy applies actions by turns if \( n \leq 2M \) and chooses the action corresponding to the smaller value of (7), (8) if \( n > 2M. \) Bayesian risk (3) calculated with respect to the prior distribution (5) is then equal to

\[ R^B_N(\rho(\nu)) = M C \left\{ 2 |U| \rho(\nu) d\nu + \int_{-\infty}^{\infty} f_{0.5MD_{g\ell}^2D_h}(U) R(U, M, M) dU \right\}. \]
Let us define the loss function weighted with respect to the prior distribution as follows

$$L_N(\sigma, \lambda) = \int_{\Theta} L_N(\sigma, \theta) \lambda(\theta) d\theta .$$

(10)

Note that the loss function (1) can be obtained as (10) if $\lambda(\theta)$ is concentrated at the parameter $\theta$. We will denote the strategy by $\sigma_i(U, n_1, n_2) = \Pr(\gamma_n = i | U, n_1, n_2)$. Let us consider the recursive equation

$$L(U, n_1, n_2) = \sigma_i(U, n_1, n_2) \times L^{(1)}(U, n_1, n_2) + \sigma_j(U, n_1, n_2) \times L^{(2)}(U, n_1, n_2) ,$$

(11)

where $L^{(1)}(U, n_1, n_2) = L^{(2)}(U, n_1, n_2) = 0$ if $n = N$ and then

$$L^{(1)}(U, n_1, n_2) = M g^{(1)}(U, n_1, n_2) + L(U, n_1 + M, n_2) * f_{M_1}^{-1}(n^2) \times \frac{1}{|n + M_1|} (U) ,$$

(12)

$$L^{(2)}(U, n_1, n_2) = M g^{(2)}(U, n_1, n_2) + L(U, n_1, n_2 + M) * f_{M_2}^{-1}(n^2) \times \frac{1}{|n + M_2|} (U) ,$$

(13)

if $2M \leq n < N$. Expected losses (10) calculated with respect to the prior distribution (5) are then equal to

$$L_N(\rho(v)) = M \int_{-c}^{C} \int_{-\infty}^{\infty} f_{0.5MD_{\tilde{h}}} (U) L(U, M, M) dU .$$

(14)

It was shown in [5] that normalized Bayesian risk $N^{-1/2} R_N(\rho(v))$ and normalized expected losses $N^{-1/2} L_N(\rho(v))$ depend only on the number of processed batches. Normalized minimax risk $N^{-1/2} R_N(\Theta)$ almost does not depend on the number of processed batches if this number is large enough. For example, 50000 primary data can be processed in parallel by batches of 1000 data items in 50 stages with almost the same maximal expected losses as if the data were processed optimally one-by-one.

4. Numerical example
Calculation of Bayesian risk by formulas (6)-(9) was implemented for $N = 50, M = 1, D_1 = 1, D_2 = 0.5$. It was assumed that the worst-case $\rho(v)$ is concentrated at two points $v = -d_1 N^{-1/2}$ and $v = d_2 N^{-1/2}$ with probabilities $\rho$ and $1 - \rho$.

![Figure 1. Normalized expected losses.](image-url)
The worst-case prior corresponds to the maximum of normalized Bayesian risk \( N^{-1/2} \mathcal{R}^B_N(\rho(v)) \) and was determined as \( d_1 = 1.49, \; d_2 = 1.33, \; \rho = 0.54 \). Then normalized expected losses 
\( I_N(d) = N^{-1/2} \mathcal{L}_N(\hat{\rho}(v)) \) were calculated for determined strategy by (11)-(14). It was assumed that \( \hat{\rho}(v) \) is concentrated at \( v = dN^{-1/2} \). Normalized expected losses are presented by the thick line 1 on figure 1. One can see that they have two maxima at \(-d_1\) and \(d_2\) which are approximately equal to 0.56 and this confirms the assumption of the worst-case prior. The large values \( I_N(d) \) at large \(|d|\) are caused by application of both actions by turns at initial two steps. Thick line 2 on figure 1 presents normalized expected losses without those on the first two steps and these losses are small for large \(|d|\). Finally, thin lines above and below thick lines on figure 1 present normalized expected losses corresponding to \( D_1 = 1.05, \; D_2 = 0.55 \) and \( D_3 = 0.95, \; D_4 = 0.45 \) respectively. One can see that normalized expected losses change a little in this case. Given large enough number of primary data, this means that variances can be estimated at the initial stage, when actions are applied by turns, and then used for the control.

Acknowledgments
This work was supported in part by the Ministry of Science and Higher Education of the Russian Federation, project 8.7367.2017/8.9.

References
[1] Berry D A and Fristedt B 1985 Bandit Problems: Sequential Allocation of Experiments (Chapman & Hall: London)
[2] Tsetlin M L 1973 Automaton Theory and Modeling of Biological Systems (New York: Academic)
[3] Sragovich V G 2006 Mathematical Theory of Adaptive Control (Singapore: World Sci.)
[4] Cesa-Bianchi N and Lugosi G 2006 Prediction, Learning, and Games (Cambridge: Cambridge Univ. Press)
[5] Kolnogorov A V 2018 Gaussian two-armed bandit and optimization of batch data processing Problems of Information Transmission 54 84-100