Clustering with bandit feedback: breaking down the computation/information gap

joint work with Alexandra Carpentier[1], Christophe Giraud[2] and Nicolas Verzelen[3]

Victor Thuot[3]

[1]University Potsdam

[2]LMO Orsay

[3]INRAE Montpellier

24th of February 2025, ALT, Milan

1 Setting: Clustering with Bandit feedback Problem

2 Lower bound and benchmark

3 Upper bound

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

- N arms
- each arm $a \in [N] \leftrightarrow$ probability distribution $\mathcal{N}(\mu_a, \sigma^2 I_d)$
- mean-vector $\mu_a \in \mathbb{R}^d$

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

- N arms
- each arm $a \in [N] \leftrightarrow$ probability distribution $\mathcal{N}(\mu_a, \sigma^2 I_d)$
- mean-vector $\mu_a \in \mathbb{R}^d$

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

- N arms
- each arm $a \in [N] \leftrightarrow$ probability distribution $\mathcal{N}(\mu_a, \sigma^2 I_d)$
- mean-vector $\mu_a \in \mathbb{R}^d$
- At each time $t \geqslant 1$,

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

- N arms
- each arm $a \in [N] \leftrightarrow$ probability distribution $\mathcal{N}(\mu_a, \sigma^2 I_d)$
- mean-vector $\mu_a \in \mathbb{R}^d$
- At each time $t \geqslant 1$,
  - the algorithm chooses arm $A_t \in [N]$ (based on passed observations)

Active setting

An algorithm collects sequentially and actively data, by interacting with a stochastic (Gaussian) bandit.

- N arms
- each arm $a \in [N] \leftrightarrow$ probability distribution $\mathcal{N}(\mu_a, \sigma^2 I_d)$
- mean-vector $\mu_a \in \mathbb{R}^d$
- At each time $t \geqslant 1$,
  - the algorithm chooses arm $A_t \in [N]$ (based on passed observations)
  - the algorithm receives X_t, s.t., conditionally on $A_t = a$, $X_t \sim \mathcal{N}(\mu_a, \sigma^2 I_d)$

Active setting

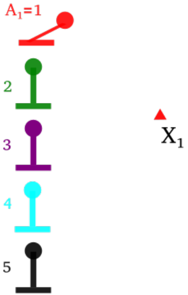


Figure: $A_1 = 1$ and receives $X_1 \sim \mathcal{N}(\mu_{A_1}, \sigma^2 I_d)$

Active setting

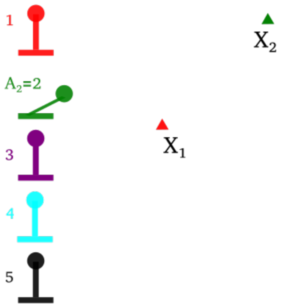


Figure: $A_2 = 2$ and receives $X_2 \sim \mathcal{N}(\mu_{A_2}, \sigma^2 I_d)$

Active setting

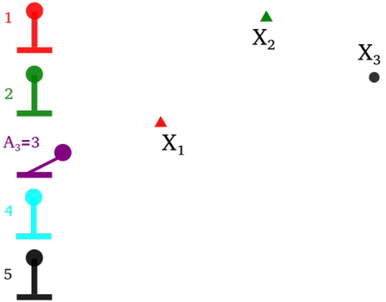


Figure: $A_3 = 3$ and receives $X_3 \sim \mathcal{N}(\mu_{A_3}, \sigma^2 I_d)$

Active setting

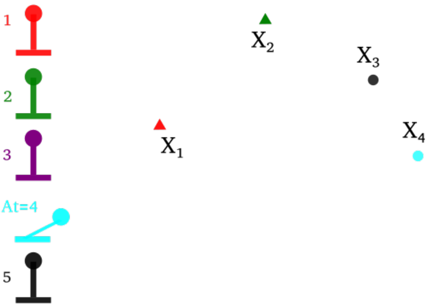


Figure: $A_4 = 4$ and receives $X_4 \sim \mathcal{N}(\mu_{A_4}, \sigma^2 I_d)$

Active setting

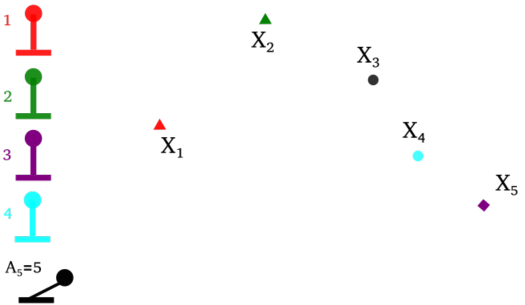


Figure: $A_5 = 5$ and receives $X_5 \sim \mathcal{N}(\mu_{A_5}, \sigma^2 I_d)$

Active setting

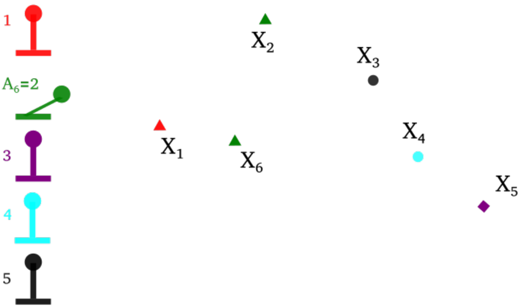


Figure: multiple sampling

Clustering with Bandit feedback Problem

The Active Clustering Problem (ACP) ([Yang et al., 2024]):

- N arms with means $\mu_1, \ldots, \mu_N$

Clustering with Bandit feedback Problem

The Active Clustering Problem (ACP) ([Yang et al., 2024]):

- N arms with means $\mu_1, \ldots, \mu_N$

Assumption 1: hidden partition

There exists a hidden partition G^* of $[N]$ into K groups.

For $a, b \in [N]$, a, b are in the same group $\Leftrightarrow \mu_a = \mu_b$.

Clustering with Bandit feedback Problem

The Active Clustering Problem (ACP) ([Yang et al., 2024]):

- N arms with means $\mu_1, \dots, \mu_N$

Assumption 1: hidden partition

There exists a hidden partition G^* of $[N]$ into K groups.

For $a, b \in [N]$, a, b are in the same group $\Leftrightarrow \mu_a = \mu_b$.

Objective

Recover the unknown partition G^* using as few requests as possible, for that:

Clustering with Bandit feedback Problem

The Active Clustering Problem (ACP) ([Yang et al., 2024]):

- N arms with means $\mu_1, \ldots, \mu_N$

Assumption 1: hidden partition

There exists a hidden partition G^* of $[N]$ into K groups.

For $a, b \in [N]$, a, b are in the same group $\Leftrightarrow \mu_a = \mu_b$.

Objective

Recover the unknown partition G^* using as few requests as possible, for that:

- sample arms until time T (budget)

Clustering with Bandit feedback Problem

The Active Clustering Problem (ACP) ([Yang et al., 2024]):

- N arms with means $\mu_1, \ldots, \mu_N$

Assumption 1: hidden partition

There exists a hidden partition G^* of $[N]$ into K groups.

For $a, b \in [N]$, a, b are in the same group $\Leftrightarrow \mu_a = \mu_b$.

Objective

Recover the unknown partition G^* using as few requests as possible, for that:

- sample arms until time T (budget)
- output $\hat{G}$ estimate of G^*.

Active clustering problem

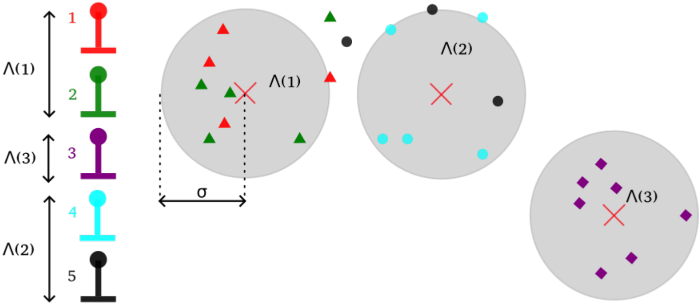


Figure: At time t, estimating $\hat{G}$ or sampling a new point.

Remarks

- assume the groups nonempty and disjoints
- (only for this presentation) assume that the groups have roughly the same size
- K is known (and also N and d)
- σ^2 is also known

PAC setting

δ-PAC algorithm

Given $\delta \in (0, 1)$, an algorithm π for the ACP is said to be δ-PAC on a collection of environments $\mathcal{E}$ if for any ν, then

$$\mathbb{P}_{\pi,\nu}(\hat{G} \sim G^*) \geqslant 1 - \delta \ .$$

where $\sim$ means that the partition is exact up to permutation of the groups.

Parameters of interest

Define the minimal gap Δ_* as

$$\Delta_* = \min_{k \neq k'} \|\Lambda(k) - \Lambda(k')\| > 0 \ .$$

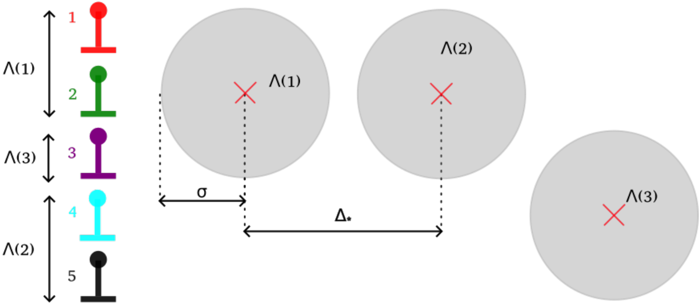


Figure: Here $N = 5$, $K = 3$, $d = 2$.

Information-theoretic lower bound

Define the optimal budget T^* as

$$T^* := N + \frac{\sigma^2}{\Delta^2} N \log\left(\frac{N}{\delta}\right) + \frac{\sigma^2}{\Delta^2}\sqrt{dKN \log\left(\frac{N}{\delta}\right)} \ .$$

Theorem 1

For any δ-PAC algorithm $\mathcal{A}$, there exist an environment with minimal gap Δ_* (and balanced groups) such that :

$$\mathbb{E}[T] \geqslant cT^* \ ,$$

with c a universal constant.

Benchmark: optimal constant if $\delta \to 0$

The problem was introduced in [Yang et al., 2024], they provide BOC algorithm such that

1. Asymptotic optimality (exact constant)
2. For balanced and equidistant groups, budget equal to

$$2\frac{\sigma^2}{\Delta^2}(N+K)\log(1/\delta)$$

Benchmark: uniform sampling

Uniform sampling strategy (US):

1. sample each arm the same number of times
2. use (batch) clustering on the empirical means

Benchmark: uniform sampling

Uniform sampling strategy (US):

1. sample each arm the same number of times
2. use (batch) clustering on the empirical means

- Benchmark ($\delta = 1/N$): uniform sampling is $(1/N) - PAC$ on as long as

$$T \gtrsim \frac{\sigma^2}{\Delta^2}\left[N(\log(N) \vee K) + \sqrt{dKN(\log(N) \vee K)}\right]$$

([Royer, 2017, Giraud and Verzelen, 2019, Vempala and Wang, 2004])

Benchmark: uniform sampling

Uniform sampling strategy (US):

1. sample each arm the same number of times
2. use (batch) clustering on the empirical means

- Benchmark ($\delta = 1/N$): uniform sampling is $(1/N) - PAC$ on as long as

$$T \gtrsim \frac{\sigma^2}{\Delta^2} \left[N(\log(N) \vee K) + \sqrt{dKN(\log(N) \vee K)} \right]$$

([Royer, 2017, Giraud and Verzelen, 2019, Vempala and Wang, 2004])

- Information-computation gap

Upper bound: the ACB algorithm

We derive an algorithm called Active Clustering with Bandit (ACB). Main structure of the algorithm:

1. identify $\hat{S}$ a set of K arms with exactly one arm from each group

Upper bound: the ACB algorithm

We derive an algorithm called Active Clustering with Bandit (ACB).
Main structure of the algorithm:

1. identify $\hat{S}$ a set of K arms with exactly one arm from each group
2. estimate the unknown centers of the groups using $\hat{S}$ (T^*/K times)

Upper bound: the ACB algorithm

We derive an algorithm called Active Clustering with Bandit (ACB). Main structure of the algorithm:

1. identify $\hat{S}$ a set of K arms with exactly one arm from each group
2. estimate the unknown centers of the groups using $\hat{S}$ (T^*/K times)
3. sample uniformly the remaining arms and use a distance-based classifier (T^*/N times)

Upper bound: main theorem

Theorem 2

Let $\delta > 0$. The ACB algorithm is δ-PAC.
If we assume that $N \geqslant K \log(K)$, and the groups are balanced, then

$$\mathbb{E}_{ACB,\nu}[T] \leqslant c \frac{\sigma^2}{\Delta^2} \left[N \log \left(N/\delta \right) + \sqrt{dNK \log \left(N/\delta \right)} \right] = cT^*$$

Upper bound: main theorem

Theorem 2

Let $\delta > 0$. The ACB algorithm is δ-PAC.
If we assume that $N \geqslant K \log(K)$, and the groups are balanced, then

$$\mathbb{E}_{ACB,\nu}[T] \leqslant c \frac{\sigma^2}{\Delta^2} \left[N \log(N/\delta) + \sqrt{dNK \log(N/\delta)} \right] = cT^*$$

- c is independent of all parameters

Upper bound: main theorem

Theorem 2

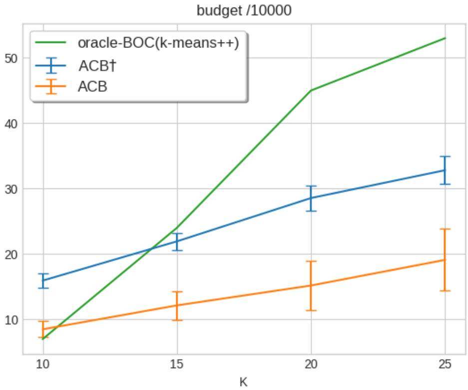
Let $\delta > 0$. The ACB algorithm is δ-PAC.
If we assume that $N \geqslant K \log(K)$, and the groups are balanced, then

$$\mathbb{E}_{ACB,\nu}[T] \leqslant c \frac{\sigma^2}{\Delta^2} \left[N \log\left(N/\delta\right) + \sqrt{dNK \log\left(N/\delta\right)} \right] = cT^*$$

- c is independent of all parameters
- we also give a bound with high probability

Numerical experiment

Figure: Comparison of the necessary budget for ACB and uniform sampling



Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

2. Can we achieve optimality ?

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

2. Can we achieve optimality ?

3. Is there an information-computation gap for ACP?

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

Yes, we provide the ACB Algorithm, which outperforms the uniform sampling strategy.

2. Can we achieve optimality ?

3. Is there an information-computation gap for ACP?

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

Yes, we provide the ACB Algorithm, which outperforms the uniform sampling strategy.

2. Can we achieve optimality ?

Yes, $T^* \simeq \frac{\sigma^2}{\Delta^2}\left[N \log\left(\frac{N}{\delta}\right) + \sqrt{dKN \log\left(\frac{N}{\delta}\right)}\right].$

3. Is there an information-computation gap for ACP?

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

Yes, we provide the ACB Algorithm, which outperforms the uniform sampling strategy.

2. Can we achieve optimality ?

Yes, $T^* \simeq \frac{\sigma^2}{\Delta^2}\left[N\log\left(\frac{N}{\delta}\right) + \sqrt{dKN\log\left(\frac{N}{\delta}\right)}\right].$

3. Is there an information-computation gap for ACP?

No, there is no computational-gap, polynomial time algorithm ACB is optimal in most regimes (e.g., balanced setting).

Take home message

1. Can we improve the budget of a simple uniform sampling strategy ?

Yes, we provide the ACB Algorithm, which outperforms the uniform sampling strategy.

2. Can we achieve optimality ?

Yes, $T^* \simeq \frac{\sigma^2}{\Delta^2}\left[N\log\left(\frac{N}{\delta}\right) + \sqrt{dKN\log\left(\frac{N}{\delta}\right)}\right].$

3. Is there an information-computation gap for ACP?

No, there is no computational-gap, polynomial time algorithm ACB is optimal in most regimes (e.g., balanced setting).

Do you have any question ?

Giraud, C. and Verzelen, N. (2019).
Partial recovery bounds for clustering with the relaxed k-means.
Mathematical Statistics and Learning, 1(3):317–374.

Royer, M. (2017).
Adaptive clustering through semidefinite programming.
Advances in Neural Information Processing Systems, 30.

Vempala, S. and Wang, G. (2004).
A spectral algorithm for learning mixture models.
Journal of Computer and System Sciences, 68(4):841–860.

Yang, J., Zhong, Z., and Tan, V. Y. (2024).
Optimal clustering with bandit feedback.
Journal of Machine Learning Research, 25:1–54.

Sub-Gaussian noise

A mean-zero random variable Z is subGaussian if, for any $t > 0$, we have $\mathbb{E}[\exp(tZ)] \leq \exp(t^2/2)$.

Assumption 2: sub-Gaussian noise

- For any $a \in [N]$, if X is sampled from the arm a,

$$E = \Sigma_a^{-1/2}[X - \mu_a]$$

  is made of independent subGaussian random variables,
- Σ_a is a $d \times d$ symmetric matrix associated to a,
- there exists σ such that $\max_{a\in[N]} \|\Sigma_a\|_{op} \leqslant \sigma^2$.

- Exemples : bounded noise or Gaussian noise.

First step: identification of representatives

Consider for $s \in [r]$ the s-th test performed on the candidate b. For $a \in \hat{S}$, we compute the statistic

$$\left\langle \bar{\mu}_{b,s} - \hat{\mu}_a, \bar{\mu}'_{b,s} - \hat{\mu}'_a \right\rangle$$

- $\hat{\mu}_{b,s}$, $\hat{\mu}'_{b,s}$ are two independent estimation of μ_b computed with n_s samples
- $\hat{\mu}_a$, $\hat{\mu}_a$ are estimates of μ_a computed with n_{max} samples
- the expectation of this statistic is $\|\mu_a - \mu_b\|^2$
- we reject b if it is smaller than $\Delta^2/2$ fo some $a \in \hat{S}$.

We use sub-Gaussian concentration to choose the tuning parameters n_s, n_{max}, r.

Second step: classification

Imagine that $\hat{S} = \{a_1, \dots, a_K\}$ contains one arm from each group.

1. First, for $j \in [K]$, label a_j with j and estimate $\mu(j)$ with two independent means using $2J$ samples.
2. Then, for each $b \in [N] \setminus \hat{S}$, labels b in the group

$$\text{argmin}_{j=1,\dots,K} \left\langle \hat{\mu}_b - \hat{\mu}(j), \hat{\mu}'_b - \hat{\mu}'(j) \right\rangle$$

$\hat{\mu}_b$,$\hat{\mu}'_b$ are computed with $I = KJ/N$ samples