Choosing a proper starting point in SGD by exploiting dependence between features — an intuition from resource allocation in event triggered communication

Ce ZHENG

Wireless Network Research Department,
Research and Development Center, Sony, China

November 21, 2021

System Model

We have

- K slots
- N users

Each user can access only one slot at most.

Mission: how to allocate slots to users when $N \gg K$?

System Model

Assume we have N users and K slots

- Activity vector: $\mathbf{X} = (X_1, \ldots, X_N) \in \{0,1\}^N$, user i is active if $X_i = 1$ and inactive otherwise.
- $X_i,\ i = 1, \ldots, N$, are assumed to be **mutually independent**.

Event-triggered communication: X_i and X_j $(i \neq j)$ are **dependent**.

System Model

How to allocate N slots to K users when $N \gg K$, with $\mathbf{X}^1, \mathbf{X}^2, \ldots$?

We need to find the allocation matrix $\mathbf{A}$ where $\mathbf{A}$ is a $N \times K$ matrix.
E.g., $N = 3$ users, $K = 2$ slots

$$\mathbf{A} = \begin{bmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix}$$

Intuition: putting highly dependent users at different slots to avoid collision.

System Model

How to allocate N slots to K users when $N \gg K$, with $\mathbf{X}^1, \mathbf{X}^2, \ldots$?

We need to find the allocation matrix $\mathbf{A}$ where $\mathbf{A}$ is a $N \times K$ matrix.

$$\max \ T(\mathbf{A}) = \mathbb{E}_{\mathbf{X}}[T^{\mathbf{X}}(\mathbf{A})],$$

T can be

- $\sum \mathbf{1}$ ($no\ collision$) — slots without collision;
- $\sum R_i$, where $R_i = W \log(1 + SINR)$ — sum-rate;
- $\sum \mathbf{1}$ ($SINR > \theta$) — successful transmission.

Intuition: putting highly dependent users at different slots to avoid collision.

Stochastic Gradient Ascent

$$\max \ T(\mathbf{A}) = \mathbb{E}_{\mathbf{X}}[T^{\mathbf{X}}(\mathbf{A})],$$

- $T(\mathbf{A})$: cost function
- $\mathbf{X}^1, \mathbf{X}^2, \ldots$: sample vectors
- X_i in $\mathbf{X} = [X_1, \ldots, X_N]$: features
- $X_i = 1$: User i is active

In event-triggered communication, X_i and X_j are highly dependent (features are dependent).

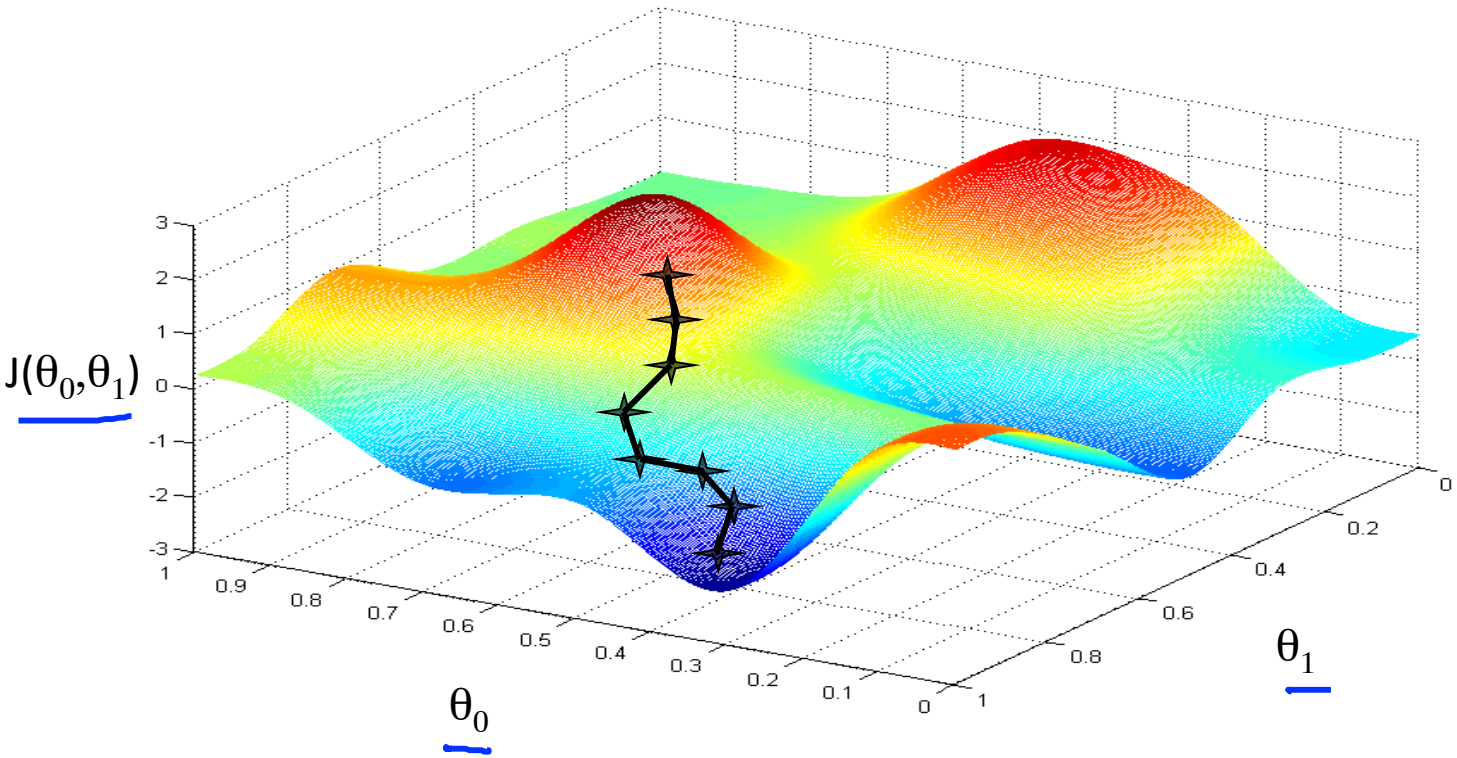
Stochastic Gradient Ascent

$$\max \ \ T(\mathbf{A}) = \mathbb{E}_{\mathbf{X}}[T^{\mathbf{X}}(\mathbf{A})],$$

The performance of Stochastic gradient ascent is mainly decided by

1. the learning rate or step size;
2. **the initial value**.

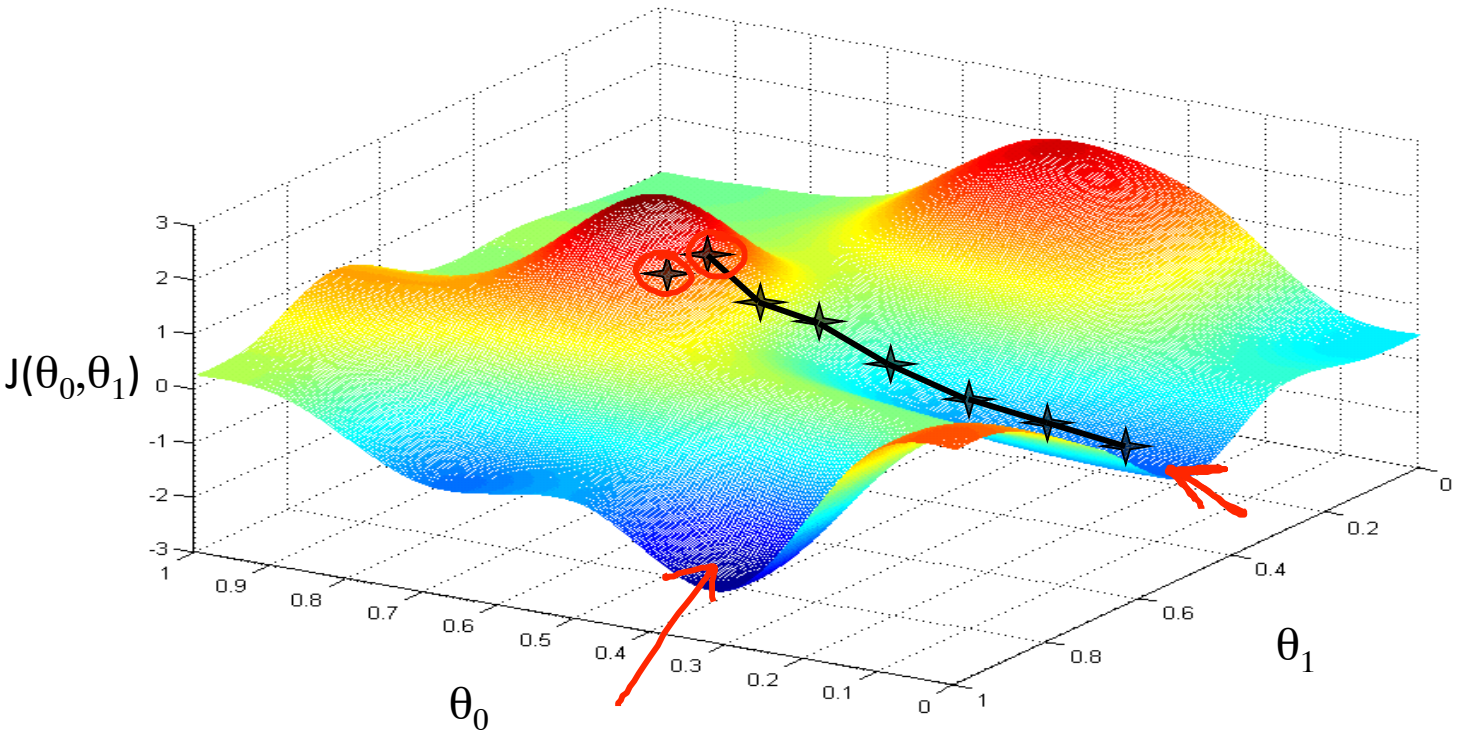
Impact of Initial Value



Andrew Ng

Figure 1: Impact of initial value for SGA

Impact of Initial Value



Andrew Ng

Figure 1: Impact of initial value for SGA

Linear Dependence — Pairwise Correlation

We only consider pairwise correlation for the ease of implementation:

$$E[X_i X_j]$$

where $X_i = 1$ denotes user i is active.

$E[X_i X_j]$ — The probability user i and j are active

Non pairwise correlation:

$$E[X_i X_j X_k \cdots]$$

Impact of Initial Value

Taking into account the dependence of user activity (especially the linear dependence, i.e. the correlation), we set the initial value of $\mathbf{A}$. E.g. two highly dependent users are put in different slots.
The performance of SGA algorithm is improved.

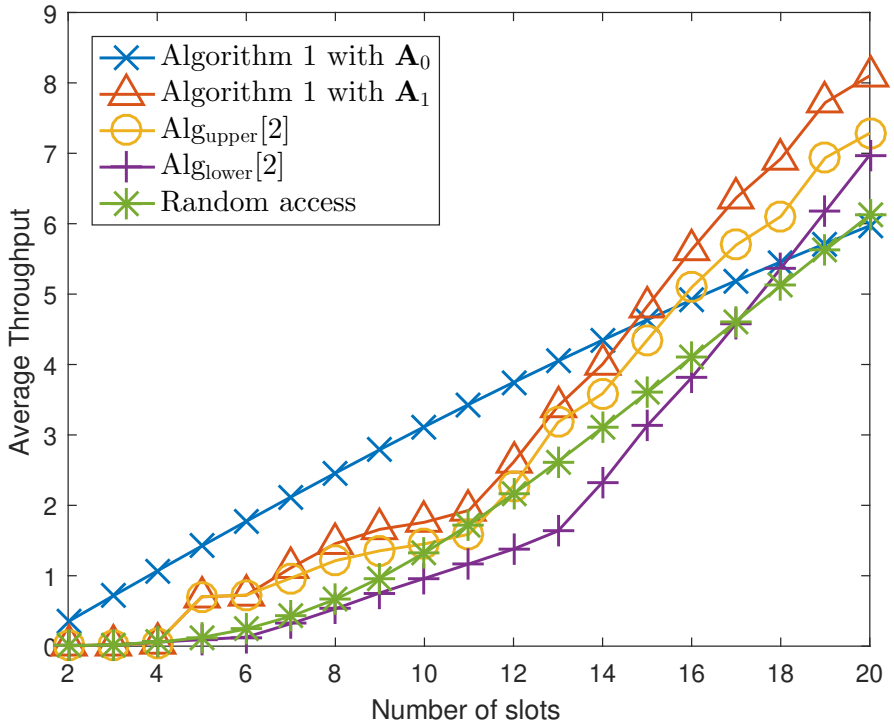


Figure 2: Throughput with $N = 50$ for varying K.

Summary

The performance of SGA algorithm is improved when set the initial value considering the dependence of user activity. [1]

[1] **Ce Zheng**, Malcolm Egan, Laurent Clavier, Petar Popovski, Anders Ellersgaard Kalør, "Stochastic Resource Allocation for Outage Minimization in Random Access with Correlated Activation" In Proc. Global Communications Conference (GLOBECOM) 2021 (Under Review)

Ce Zheng, Malcolm Egan, Laurent Clavier, Petar Popovski, Anders Ellersgaard Kalør, "Stochastic Resource Optimization of Random Access for Transmitters with Correlated Activation", *IEEE Communication Letters*, 2021

Extension — nonlinear dependence

X is symmetric and $Y = X^2$

The correlation is

$$E[XY] = 0$$

X and Y are linearly independent but highly dependent

One promising solution is **copula**

References

- **Ce Zheng**, Malcolm Egan, Laurent Clavier, Petar Popovski, Anders Ellersgaard Kalør, "Stochastic Resource Optimization of Random Access for Transmitters with Correlated Activation", *IEEE Communication Letters*, 2021
- **Ce Zheng**, Malcolm Egan, Laurent Clavier, Petar Popovski, Anders Ellersgaard Kalør, "Stochastic Resource Allocation for Outage Minimization in Random Access with Correlated Activation" In Proc. Global Communications Conference (WCNC) 2021 (Under Review)
- Anders E Kalor, Osama A Hanna, Petar Popovski, "Random access schemes in wireless systems with correlated user activity", In2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)