線形システムに対する裾の重い雑音を用いたプライバシー保護

○伊藤 海斗 （京都大学） 河野 佑 （広島大学） 加嶋 健司 （京都大学）

Privacy Preservation with Heavy-tailed Noise for Linear Dynamical Systems

∗K. Ito (Kyoto Univ.), Y. Kawano (Hiroshima Univ.) and K. Kashima (Kyoto Univ.)

Abstract– Differential privacy enables us to quantify privacy levels of data involved in linear dynamical systems, by adding noise for the purpose of data protection. However, information of outliers is vulnerable when Gaussian noise is employed as in usual literature in systems and control. The goal of this paper is to present new differentially private mechanisms which can hide occurrences of outliers. The key idea is to utilize stably distributed noise that is closed under linear dynamics as Gaussian is, and its distribution is heavy-tailed different from Gaussian. We provide conditions for the mechanisms induced by linear dynamical systems and stably distributed noises to be differentially private. Thanks to the former property of the stable distribution, the proposed mechanisms can handle privacy-preserving control problems studied in terms of Gaussian mechanisms. The latter one is beneficial for protecting information of outliers.

Key Words: Differential privacy, stochastic systems, discrete-time linear systems, stable distribution

1 はじめに

IoT(Internet-of-Things) やクラウドコンピューティング技術の急速な発展に伴い，プライバシーやセキュリティの問題が大きく社会で取り上げられている．これらの技術では，ユーザーの個人情報を利用することで，より良いサービスを提供することができるが，その反面で，知られたくない個人情報が漏洩する危険性が増してしまう．よって，このような危険性を十分考慮した，プライバシー保護技術の確立が急務となっており，研究が盛んに行われている [1, 2, 3, 4]．こうした背景のもとで，**差分プライバシー**と呼ばれる，プライバシー保護水準を定量化する概念が提案された [5]．差分プライバシーでは，プライバシー保護をしたい入力データとそれに似たデータを，対応する出力データから識別することの困難さが定量化される．識別困難性を高めるには，公開する前の出力データにラプラス雑音やガウス雑音などの確率雑音が加えられる．

差分プライバシーは従来，静的なデータを対象にしてきたが，近年では動的システムに対しても適用されている [6, 7]．これらの研究では，主にガウス雑音が解析で用いられている．ここで強調すべきなのは，外れ値が含まれるデータを扱う際には，差分プライバシー解析だけでは不十分ということである [8]．外れ値データを隠すには，裾の重い雑音が用いられる．裾の重い雑音では極端に大きな (小さな) 値の標本が，無視できない確率で発生する．この性質により，出力で外れ値が観測されたとき，それが保護対象の入力データに由来するのか，それとも雑音によるものなのか，攻撃者が高い確信度で推定することができなくなる．裾が指数関数的に減衰するガウス雑音では，このような外れ値を隠すことができない．

差分プライバシー解析において，標準的に用いられるもう一つの雑音として，ラプラス雑音がある．ラプラス雑音は，ガウス雑音よりも裾の減衰が遅く，外れ値をモデル化するのに用いられる [9]．しかし，その減衰は，べき乗則での減衰よりは遥かに速く，ゆえにスケールフリー性 [10] をもつ，すなわちべき乗則に従うデータを守る際には，ラプラス雑音ではまだ不十分である．また，ガウス雑音と違い，ラプラス雑音は線形なダイナミクスの下で閉じていないという欠点もある．これは本質的には，ガウス雑音は再生性をもつが，ラプラス雑音はもたないからである．これにより，ラプラス雑音を用いた場合，推定器/制御器設計が困難になってしまう．

これらの問題に対して，本研究では裾の重い分布の一種である安定分布を雑音として用い，線形システムに対するプライバシー保護を達成することを考える．安定分布はべき乗則に従う裾をもっており，べき乗則に従うデータの保護が可能である．さらに安定分布は，ラプラス分布と異なり，線形ダイナミクスの下で閉じている [11]．このようなガウス分布との類似性は，制御系解析や設計の観点から有用である．本稿では特に，安定分布に従う雑音について差分プライバシー解析が可能であることを明らかにする．

本稿の構成は以下の通りである．2 節では，安定分布の導入を行う．3 節で，本稿の主結果である，安定分布に従う雑音を用いたメカニズムが差分プライバシーを満たすための十分条件を与える．最後に 4 節で結論を述べる．

記法

実数の集合，非負整数の集合をそれぞれ $\mathbb{R}, \mathbb{Z}_+$ で記す．虚数単位を j で記す．ベクトル $x_1, \ldots, x_m \in \mathbb{R}^n$ を縦に並べたベクトル $[x_1^\top \ \cdots \ x_m^\top]^\top \in \mathbb{R}^{nm}$ を $[x_1; \cdots; x_m]$ で表す．信号 $u(t) \in \mathbb{R}^m, t \in \mathbb{Z}_+$ に対して，時系列順に時刻 $T \in \mathbb{Z}_+$ まで並べたベクトルを，大文字のアルファベットを用いて，$U_T := [u(0); \cdots; u(T)] \in \mathbb{R}^{(T+1)m}$ と表す．行列 $A \in \mathbb{R}^{n\times n}$ の最大・最小固有値をそれぞれ $\lambda_{\max}(A), \lambda_{\min}(A)$ で表す．行列 A が正定値対称行列であることを $A \succ 0$ と書く．ベクトル $x = [x_1 \ \cdots \ x_n]^\top \in \mathbb{R}^n$ の p ノルム $(p \geq 1)$ を $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$ と定義する．また，行列 $A \succ 0$ で重み付けしたノルムを $\|x\|_A := (x^\top A x)^{1/2}$ で定義する．本稿では以降，ある完備確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ を固定する．$\mathbb{E}$ を確率測度 $\mathbb{P}$ についての期待値とする．$\mathcal{Q}$ 関数を

$$\mathcal{Q}(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty \mathrm{e}^{-\frac{u^2}{2}} du, \ x \in \mathbb{R}$$

で定義する．

2 準備：安定分布

ここで，安定分布の定義や性質を簡単にまとめる．安定分布の定義は以下で与えられる [12].

定義 1 $\mathbb{R}^d$-値確率ベクトル $X = [X_1 \cdots X_d]^\top$ の特性関数がパラメータ $\alpha \in (0, 2]$ と $d \times d$ の行列 $\Sigma \succ 0$ を用いて，

$$\mathbb{E}[\exp(\mathrm{j}\nu^\top X)] = \exp(-(\nu^\top \Sigma \nu)^{\alpha/2}), \ \nu \in \mathbb{R}^d \quad (1)$$

で与えられるとき，X は楕円形安定分布に従うといい，$X \sim \boldsymbol{SG}_d(\alpha, \Sigma)$ で表す． ◁

以降では，楕円形安定分布を単に安定分布と呼ぶ．パラメータ α は分布の非ガウス性の度合いを表す．特に $\alpha = 2$ のとき，安定分布は平均 0，共分散行列 2Σ の多変量ガウス分布となる．ガウス分布の共分散行列と同様に Σ は楕円分布の形状を特徴づける．しかし，$\alpha < 2$ のときは，裾の重さにより共分散行列が存在しないことに注意されたい．安定分布がもつべき乗則は以下で示される [13, Property 1.2.15].

命題 1 $\alpha \in (0, 2), \Sigma \succ 0$ とする．このとき，$X = [X_1 \ \cdots \ X_d]^\top \sim \boldsymbol{SG}_d(\alpha, \Sigma)$ に対し，以下が成り立つ．

$$\lim_{\lambda\to\infty} \lambda^\alpha \mathbb{P}(X_i > \lambda) = \begin{cases} \frac{1-\alpha}{2\Gamma(2-\alpha)\cos(\frac{\pi\alpha}{2})} \Sigma_i^{\alpha/2}, & (\alpha \neq 1) \\ \frac{1}{\pi}\Sigma_i^{\alpha/2}, & (\alpha = 1) \end{cases} \quad (2)$$

ここで，$i \in \{1, \ldots, d\}$ であり，Σ_i は Σ の i 番目の対角要素を表す． ◁

べき乗則の減衰の速さは α で決まることが分かる．すなわち，α が小さいほど，安定分布の裾は重くなる．

また主結果を述べるために，歪度パラメータを含んだ単変量安定分布を導入する [13].

定義 2 実数値確率変数 X の特性関数がパラメータ $\alpha \in (0, 2]\backslash\{1\}$, $\sigma > 0$, $\beta \in [-1, 1]$ を用いて，

$$\mathbb{E}[\exp(\mathrm{j}\nu X)] = \exp\left\{-\sigma^\alpha |\nu|^\alpha \left(1 - \mathrm{j}\beta \mathrm{sgn}(\nu) \tan\frac{\pi\alpha}{2}\right)\right\},$$

$$\mathrm{sgn}(\nu) := \begin{cases} 1 & \text{if } \nu > 0, \\ 0 & \text{if } \nu = 0, \\ -1 & \text{if } \nu < 0, \end{cases}$$

で与えられるとき，X は単変量安定分布に従うといい，$X \sim \boldsymbol{S}(\alpha, \sigma, \beta)$ で表す． ◁

以降では，A_α で

$$A_\alpha \sim \boldsymbol{S}\left(\frac{\alpha}{2}, 2\left(\cos\frac{\pi\alpha}{4}\right)^{2/\alpha}, 1\right), \ \alpha \in (0, 2) \quad (3)$$

を満たす確率変数を表す．

3 安定分布を用いた差分プライバシー解析

3.1 問題設定

本稿では以下の離散時間線形システム

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}, \ t \in \mathbb{Z}_+ \quad (4)$$

を扱う．ここで，$x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m, y(t) \in \mathbb{R}^q$ はそれぞれ，状態，制御入力，出力を表し，A, B, C, D は適切な次元の行列である．(4) を用いると，出力系列 $Y_T \in \mathbb{R}^{(T+1)q}$ は

$$Y_T = O_T x_0 + N_T U_T \quad (5)$$

と書き下すことができる．ここで $x(0) = x_0$ であり，$O_T \in \mathbb{R}^{(T+1)q\times n}, N_T \in \mathbb{R}^{(T+1)q\times(T+1)m}$ は

$$O_T := \left[C^\top \ CA^\top \ \cdots \ (CA^T)^\top\right], \quad (6)$$

$$N_T := \begin{bmatrix} D & 0 & \cdots & \cdots & 0 \\ CB & D & \ddots & & \vdots \\ CAB & CB & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{T-1}B & CA^{T-2}B & \cdots & CB & D \end{bmatrix} \quad (7)$$

で定義される．

ここで，動的システムに対する差分プライバシーを導入する [6, 7]．差分プライバシーでは，公開する前の

出力データに確率雑音を加えることで，対応する入力データのプライバシー保護を行う．つまり，雑音 $w(t)$ を加えた出力 $y_w(t) := y(t) + w(t)$ を公開する．この新たな出力の系列 $Y_{w,T} = [y_w(0); \cdots ; y_w(T)]$ は

$$Y_{w,T} = O_T x_0 + N_T U_T + W_T \tag{8}$$

と書ける．(8) により，写像 $\mathcal{M} : \mathbb{R}^n \times \mathbb{R}^{(T+1)m} \times \mathbb{R}^{(T+1)q} \ni (x_0, U_T, W_T) \mapsto Y_{w,T} \in \mathbb{R}^{(T+1)q}$ が定義される．差分プライバシー解析では，この写像はメカニズムと呼ばれる [5].

差分プライバシーは “似た” データ組を通して定義される．データが似ていることは，具体的には以下で定義される隣接関係を満たしていることを言う．

定義 3 与えられた $c > 0, p \geq 1$ に対して，入力データ組 $((x_0^1, U_T^1), (x_0^2, U_T^2)) \in (\mathbb{R}^n \times \mathbb{R}^{(T+1)m}) \times (\mathbb{R}^n \times \mathbb{R}^{(T+1)m})$ が p ノルムについて，c-隣接関係にあるとは，$\|[x_0^1; U_T^1] - [x_0^2; U_T^2]\|_p \leq c$ が成り立つことをいい，c-隣接関係にある組全体を Adj_p^c で表す． ◁

この隣接関係を用いて，差分プライバシーは以下のように定義される．

定義 4 $\mathcal{B}(\mathbb{R}^{(T+1)q})$ を $\mathbb{R}^{(T+1)q}$ 上のボレル集合族とする．与えられた $\varepsilon > 0, \delta \geq 0$ に対し，メカニズム (8) が，Adj_p^c と有限時刻 $T \in \mathbb{Z}_+$ について，(ε, δ)-差分プライバシーを満たすとは，任意の $((x_0^1, U_T^1), (x_0^2, U_T^2)) \in \mathrm{Adj}_p^c$ に対して，

$$\begin{aligned} &\mathbb{P}(O_T x_0^1 + N_T U_T^1 + W_T \in S) \\ &\leq \mathrm{e}^{\varepsilon} \mathbb{P}(O_T x_0^2 + N_T U_T^2 + W_T \in S) + \delta, \ \forall S \in \mathcal{B}(\mathbb{R}^{(T+1)q}) \end{aligned} \tag{9}$$

が成り立つことを言う． ◁

文献 [7] では，ガウス雑音が差分プライバシー解析に用いられている．本稿では，ガウス分布を特別な場合として含み，かつ裾の重い分布である安定分布に従う雑音を用いる．本稿で扱う主問題を以下にまとめる．

問題 1 与えられた線形システム (4) と $\varepsilon, \delta > 0$ に対し，(9) を満たすような安定分布に従う雑音 W_T を設計せよ． ◁

3.2 出力雑音・入力雑音による差分プライバシー解析

ここで，差分プライバシー解析についての主結果を述べる．以下に述べる定理から分かるように，任意に定められた差分プライバシー水準に対して，それを達成する安定分布に従う雑音が存在する．

定理 1 $\alpha \in (0,2), \varepsilon > 0, \delta \in (0,1)$ とする．$c > 0$ と $T \in \mathbb{Z}_+$，(3) で定まる A_α に対して，行列 $\Sigma \succ 0$ を

$$\lambda_{\max}^{1/2}(\mathcal{O}_{\Sigma,T}) \leq \frac{1}{c} \mathcal{Q}_{\alpha,\varepsilon}^{-1}(\delta), \tag{10}$$

$$\mathcal{O}_{\Sigma,T} := [\ O_T\ N_T\ ]^\top \Sigma^{-1} [\ O_T\ N_T\ ], \tag{11}$$

$$\mathcal{Q}_{\alpha,\varepsilon}(z) := \mathbb{E}\left[\mathcal{Q}\left(\frac{\varepsilon A_\alpha^{1/2}}{z} - \frac{z}{2A_\alpha^{1/2}}\right)\right],\ z > 0, \tag{12}$$

が成り立つように選ぶ．このとき，$W_t \sim \boldsymbol{SG}_{(T+1)q}(\alpha, \Sigma)$ により生成されるメカニズム (8) は Adj_2^c と有限時刻 T について，(ε, δ)-差分プライバシーを満たす． ◁

上で得られた結果はガウス雑音を用いた場合 [7] を $\alpha \to 2$ の極限として含んでいる．実際，法則収束の意味で $A_\alpha \to 2$ $(\alpha \to 2)$ となり，[7, Theorem 2.6] で与えられる条件と一致する．この意味で，定理 1 はガウス雑音で得られていた結果の拡張であると言える．

次に，入力に直接，雑音を印加することで，与えられた差分プライバシー水準を達成することを考える．初期状態と入力に雑音が加わった以下のシステムを考える．

$$\begin{cases} x(t+1) = Ax(t) + B(u(t) + v(t)),\ x(0) = x_0 + v_x, \\ y_v(t) = Cx(t) + D(u(t) + v(t)). \end{cases} \tag{13}$$

出力系列 $Y_{v,T} = [y_v(0); \cdots ; y_v(T)]$ は

$$Y_{v,T} = O_t x_0 + N_T U_T + [\ O_T\ N_T\ ][v_x; V_T] \tag{14}$$

で与えられる．安定分布が線形変換について閉じていることを利用すると，定理 1 から以下の結果を得る．

系 1 $[\ O_T\ N_T\ ]$ は正方行列かつ可逆であるとし，$\alpha \in (0,2), \varepsilon > 0, \delta \in (0,1)$ とする．$c > 0$ と $T \in \mathbb{Z}_+$，(3) で定まる A_α に対して，行列 $\Sigma \succ 0$ を

$$\lambda_{\min}^{1/2}(\Sigma) \geq \frac{c}{\mathcal{Q}_{\alpha,\varepsilon}^{-1}(\delta)} \tag{15}$$

が成り立つように選ぶ．このとき，$[v_x; V_T] \sim \boldsymbol{SG}_{n+(T+1)m}(\alpha, \Sigma)$ により生成されるメカニズム (14) は Adj_2^c と有限時刻 $T \in \mathbb{Z}_+$ について，(ε, δ)-差分プライバシーを満たす． ◁

本結果より，入力に直接雑音を加える場合，差分プライバシー水準はシステムそのものに依らないということが分かる．なお，系 1 では，$[\ O_T\ N_T\ ]$ は正方行列かつ可逆としたが，[7, Corollary 2.16] で用いられている手法により，この仮定は緩めることができる．

4 おわりに

本稿では，線形システムにおける外れ値・非外れ値データ両方のプライバシー保護を目的として，システムの出力，あるいは入力に安定分布に従う雑音を加えるメカニズムを提案した．また，提案メカニズムに対し，差分プライバシー解析を行い，メカニズムが差分プライバシーを満たすための条件を与えた．この条件をもとに，雑音分布の裾の重さが差分プライバシー水準にどのように影響を与えるか，解析することができる．今後の課題は，安定分布に従う雑音の形状パラメータの，外れ値データの保護を考慮した選定方法の検討である．

謝辞

本研究は JSPS 科研費 (JP18H01461) の助成を受けたものです．

参考文献

[1] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," *2010 First IEEE International Conference on Smart Grid Communications*, pp. 327–332, 2010.

[2] F. K. Dankar and K. El Emam, "The application of differential privacy to health data," *ACM International Conference Proceeding Series*, pp. 158–166, 2012.

[3] F. Kargl, A. Friedman, and R. Boreli, "Differential privacy in intelligent transportation systems," *WiSec 2013 - Proceedings of the 6th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 107–112, 2013.

[4] M. Yang, A. Margheri, R. Hu, and V. Sassone, "Differentially private data sharing in a cloud federation with blockchain," *IEEE Cloud Computing*, vol. 5, no. 6, pp. 69–79, 2018.

[5] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of cryptography conference*. Springer, 2006, pp. 265–284.

[6] J. Le Ny and G. J. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2014.

[7] Y. Kawano and M. Cao, "Design of privacy-preserving dynamic controllers," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3863–3878, 2020.

[8] C. C. Aggarwal and S. Y. Philip, "A general survey of privacy-preserving data mining models and algorithms," in *Privacy-preserving data mining*. Springer, 2008, pp. 11–52.

[9] A. Y. Aravkin, B. M. Bell, J. V. Burke, and G. Pillonetto, "An ℓ_1-Laplace robust Kalman smoother," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2898–2911, 2011.

[10] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[11] K. Kashima, H. Aoyama, and Y. Ohta, "Stable process approach to analysis of systems under heavy-tailed noise: Modeling and stochastic linearization," *IEEE Transactions on Automatic Control*, vol. 64, no. 4, pp. 1344–1357, 2019.

[12] J. P. Nolan, "Multivariate elliptically contoured stable distributions: theory and estimation," *Computational Statistics*, vol. 28, no. 5, pp. 2067–2089, 2013.

[13] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, 1994.