合成データ生成のランダム性が持つ Rényi差分プライバシー性の評価

三浦 尭之[1,a)]　紀伊 真昇[1]　芝原 俊樹[1]　市川 敦謙[1]　山本 充子[1]　千田 浩司[2]

概要：合成データ生成によるプライバシー保護は，理論的な安全性を定量的に表現するために，統計量やモデルパラメータに差分プライバシー性を保証させるようなノイズを加えることが主流であるが，データ生成時のランダム性によるプライバシー保護性は考慮されていない．本研究では，元データの平均ベクトルと分散共分散行列を利用した多変量正規分布による合成データ生成方式に対して，データ生成時のランダム性がみたす Rényi 差分プライバシー性を理論的に評価した．具体的には，隣接性が秘匿 n 条件による場合と，公開 n 条件による場合ごとに，$\alpha > 1$ を決めたときの合成データ生成が (α, ε)-Rényi 差分プライバシーを満たすような ε の条件を導出した．特に，秘匿 n 条件で導出した ε は，元データのサンプル数を 1000 万件ほどまで大きくすると，ノイズを足すなどの操作をしなくても，同数のサンプルを出力するメカニズムが $(4, 0.576)$-Rényi 差分プライバシーを満たし，また，従来の (ε, δ)-差分プライバシーに換算しても $(2.72, 10^{-5})$-差分プライバシーを満たすことがわかった．

キーワード：Rényi-差分プライバシー，合成データ生成，プライバシー保護

On Rényi Differential Privacy in Synthetic Data Generation

Takayuki Miura[1,a)]　Masanobu Kii[1]　Toshiki Shibahara[1]　Atsunori Ichikawa[1]
Juko Yamamoto[1]　Koji Chida[2]

Abstract: Most privacy-preserving synthetic data generation techniques guarantee differential privacy by adding noise to statistics and model parameters. However, privacy protection due to randomness in data generation is not considered. In this study, we theoretically evaluate the Rényi differential privacy of randomness in data generation. Specifically, we consider synthetic data generation based on multivariate normal distribution using the mean vector and variance-covariance matrix of the original data. For each of the two adjacency conditions, we derived a condition of ε such that the synthetic data generation satisfies (α, ε)-Rényi differential privacy with fixed $\alpha > 1$. In particular, ε derived under the private n regime satisfies $(4, 0.576)$-Rényi differential privacy when the number of samples of the original data is increased up to about 10 million samples, and the mechanism to output samples of the same size without any operations such as adding noise satisfies the conventional (ε, δ)-differential privacy, and $(2.72, 10^{-5})$-differential privacy.

Keywords: Rényi differential privacy, synthetic data generation, privacy protection

1. はじめに

個人に関わるデータの利活用は様々な分野で期待されているが，これらの取り扱いにはプライバシー保護上の注意が必要である．一方で元データの情報を大きく損ねるような過度なプライバシー保護は，保護済みデータの価値を大幅に削ってしまう．そのため，k 匿名性 [11] や (ε, δ)-差分プライバシー [2,3]，Rényi 差分プライバシー [9] などの定量的な指標に基づいた安全性を保証しながら，有用性と両立させることが重要である．特に，画像や多属性のテーブルデータなど，1 レコードが高次元のデータに対してはそ

[1] NTT 社会情報研究所, NTT Social Informatics Laboratories
[2] 群馬大学, Gunma University
[a)] takayuki.miura.br@hco.ntt.co.jp

の両立が難しいため，高次元データに対しても有用性を保てるプライバシー保護技術として，合成データ生成技術に注目が集まっている [12,14,16].

合成データ生成技術は，保護すべきデータセットから値（本稿ではこれを**生成パラメータ**と呼ぶ）を抽出し，生成パラメータを用いて元のデータセットと同様の特徴を持つレコードやデータセットを生成する技術である．生成パラメータとしてデータセットのもつ統計量を利用した統計量ベースの方法 [15,16] や，データセットを訓練データとし，深層学習によって得られた生成モデルを用いる方法 [5,6,10,13] などが代表的なものとしてあげられる．

これらの方式は，生成パラメータに対して，差分プライバシーやその自然な拡張である Rényi 差分プライバシーを満たすようノイズを加えることによって，その安全性を保証している [8,14]．これは，統計量や生成モデルのパラメータなどの生成パラメータ自体が差分プライベートであれば，そこから生成されるデータも差分プライバシーを満たしているという考え方である（図 1(a)）．しかし，ノイズを加えてランダム性を持たせなくとも合成データ生成には，データを生成する操作自体にランダム性がある．このランダム性により，生成されたデータから生成パラメータや元データに関する情報を推定することは難しい問題となる．実際に，PWS Cup 2020[*1]で合成データ生成を用いた匿名化部門上位チームの保護データは，差分プライバシーなどの保護は用いられていないにも関わらず，他のチームのメンバーシップ推論攻撃に強い耐性を見せた．

こうした事実から，合成データ生成は出力を生成パラメータではなく合成データのみであると考えると，もともと持つデータ生成時のランダム性によって，すでになんらかのプライバシー保護性を有していると考えることができる（図 1(b)）．これらを評価し，従来のノイズを加えることによって保証されるプライバシー保護と適切に組み合わせることで，加えるノイズの量を減らすことができ，同じ安全性の保証のもと，より有用性の高いデータが生成できるようになることが期待される．

こうした観点の先行研究はまだ少ない．Lin ら [7] は，GAN のサンプル生成時のランダム性が持つ確率的差分プライバシー性を理論的に評価した．しかし，評価式には識別器の汎化誤差など具体的な計算には向いていない項も含まれ，また，現実的なサンプルサイズでは無視できない項も訓練データを無限大に増やせば消えるという理屈で無視されている．論文中でもバウンドの具体的な数値を代入した考察や数値実験はされていないため，あくまで研究は理論的な評価式の導出にとどまっていた．また，著者らは 2021 年の CSS で同様の課題に対する研究結果 [17] を報告したが，安全性基準が差分プライバシーを変形した考え方で，さらに生成モデルは 1 次元の正規分布に従うサンプリングというかなり限定的な設定であった．

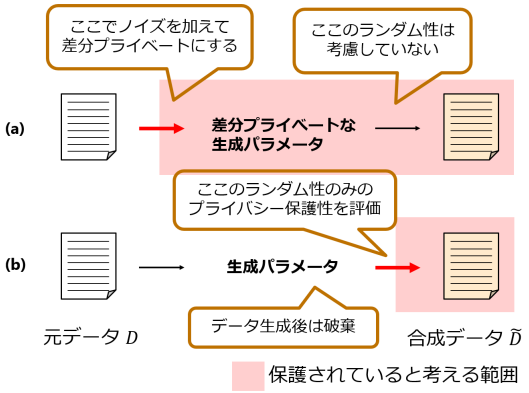

図 1 (a)「出力＝生成パラメータ」：この計算・学習などにランダム性を持たせて差分プライベートにすることで，そこから生成される合成データの安全性を保証．
(b)「出力＝合成データのみ」：ノイズを加えるなどの保護をしない．データ生成後は生成パラメータを破棄する．データ生成時のランダム性がそもそもどの程度のプライバシー保護性を持っているのか評価．

本稿では，任意の次元のレコード（d 次元ベクトルとする）からなるデータセットに対して，それらの平均ベクトルと分散共分散行列を利用した多変量正規分布による合成データ生成方式[*2]のデータ生成時のランダム性がみたす Rényi 差分プライバシー性を理論的に評価した．具体的には，隣接性が秘匿 n 条件による場合と，公開 n 条件による場合ごとに，$\alpha > 1$ を決めたときの合成データ生成が (α, ε)-Rényi 差分プライバシーを満たすような ε の条件を導出した（定理 3.1, 3.2）．さらにこれらの結果に具体的な数値を入れて，観察した結果，秘匿 n 条件の下では元データのサンプル数を 1000 万件ほどまで大きくすると，ノイズを足すなどの操作をしなくても，同数のサンプルを出力するメカニズムが $(4, 0.144)$-Rényi 差分プライバシーを満たし，また，従来の (ε, δ)-差分プライバシーに換算しても $(2.72, 10^{-5})$-差分プライバシーを満たすことがわかった（表 1,2）．これらの値は，Apple 社[*3]や米国センサス局[*4]の事例で用いられている基準よりも小さいため，実用的な水準と考えることができる．

2. 準備

次節以降の議論で必要になる記号・考え方を紹介する．

[*1] https://www.iwsec.org/pws/2020/cup20.html

[*2] 総務省統計局などではこうした手法が用いられている．https://www.nstac.go.jp/sys/files/static/services/ippan/ippan_tebiki.pdf

[*3] https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf

[*4] https://www.census.gov/newsroom/press-releases/2021/2020-census-key-parameters.html

2.1 記号

本稿では正方行列 $A \in \mathbb{R}^{d\times d}$ に対する行列式を $|A| := \det A$ と表すこととする．また，ベクトル $x \in \mathbb{R}^d$ や行列 $A \in \mathbb{R}^{d_1\times d_2}$ に対して，その転置ベクトルや転置行列を ${}^t x$ や ${}^t A$ と表すこととする．各データはテーブル上のデータを想定していて各属性は数値属性でレンジは $[-1,1]$ に正規化してあるとする．1 個人の情報は d 次元のベクトル $x \in [-1,1]^d$ で表現できる．このとき，n 人のレコードからなるデータセットは $D = \{x_i\}_{i=1,\dots,n} \in [-1,1]^{d\times n}$ と表現できる．

2.2 Rényi 差分プライバシー

Rényi 差分プライバシーの定義を紹介する．まず，隣接性の定義をする．

定義 2.1 (隣接データセット)**.** データセット $D, D' \in \mathcal{D}$ が隣接データセットであるとは，「D と D' が 1 レコードのみ異なる」状態であることを表す．ここでデータセットのレコード総数に制約がある場合（**公開 n 条件**と呼ぶ [18]），1 レコードのみが異なるということは 1 レコードの情報が丸々ほかのモノに置き換わることを意味する．そのような制約がない場合（**秘匿 n 条件**と呼ぶ [18]）は，1 個人のデータが追加/削除されている場合を意味する[*5]．

次に Rényi 差分プライバシーを定義するうえで必要となる Rényi divergence を定義する．

定義 2.2 (Rényi Divergence)**.** P, Q を $\mathbb{R}^d$ 上の確率分布とする．$\alpha > 1$ に対して，

$$D_\alpha(P||Q) := \frac{1}{\alpha-1}\log\left(\int_{\mathbb{R}^d} P(x)^\alpha Q(x)^{1-\alpha} dx\right)$$

をオーダー α の **Rényi Divergence** と呼ぶ．

Rényi divergence の基本的な性質については付録 A.1 にまとめたので必要に応じて参照する．

定義 2.3 (Rényi 差分プライバシー [9])**.** プライバシー保護メカニズム（ランダム化関数）$\mathcal{M} : \mathcal{D} \to \mathbb{R}^d$ が実数 $\alpha > 1, \varepsilon > 0$ に対して，(α,ε)-**Rényi 差分プライバシー** ((α,ε)-*RDP*) を満たすとは次が成り立つことを言う．任意の隣接データセット $D, D' \in \mathcal{D}$ に対して，

$$D_\alpha(\mathcal{M}(D)||\mathcal{M}(D')) \le \varepsilon.$$

Rényi 差分プライバシーと従来の差分プライバシーの関係は付録 A.2 に記載するが，定性的な意味として一つ重要なことは「ε **は小さいほど保護が厳しく，α は大きいほど保護が厳しく**なり，また，α を無限大に飛ばすと ε-DP と同じ定義になる」ということである．

また，Rényi 差分プライバシーについても従来の (ε,δ)-差分プライバシー同様に合成則が成り立ち，現在広く使われている (ε,δ)-差分プライバシーへ換算することもできる．

命題 2.4 (Rényi 差分プライバシーの合成則 [9])**.** $\mathcal{M}_1 : \mathcal{D} \to \mathbb{R}^{d_1}$ を (α,ε_1)-*RDP* メカニズム，$\mathcal{M}_2 : \mathcal{D}\times\mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ を (α,ε_2)-*RDP* メカニズムとする．このとき，メカニズム $\mathcal{M} : \mathcal{D} \to \mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$ を $\mathcal{M}(D) = (\mathcal{M}_1(D), \mathcal{M}_1(D, \mathcal{M}_2(D)))$ は $(\alpha,\varepsilon_1+\varepsilon_2)$-*RDP* を満たす．

命題 2.5 ((ε,δ)-差分プライバシーへの換算 [9])**.** メカニズム $\mathcal{M}$ が (α,ε)-*RDP* を満たすとき，$\mathcal{M}$ は任意の $0 < \delta < 1$ に対して，$(\varepsilon + \frac{\log\frac{1}{\delta}}{\alpha-1}, \delta)$-*DP* も満たす．

2.3 平均・分散共分散行列による合成データ生成

本稿では元データセットの平均・分散共分散行列を用いて，それに従うデータをサンプリングするという簡単な合成データ手法を考える．

データセット $D = \{x_i\}_{i=1,\dots,n} \in \mathcal{D}$ に対して，各レコードの平均ベクトル $\mu \in \mathbb{R}^d$ と分散共分散行列 $\Sigma \in \mathbb{R}^{d\times d}$ を計算する．具体的な計算式は

$$\mu := \frac{1}{n}\sum_{i=1}^n x_i, \quad \Sigma := \frac{1}{n}\sum_{i=1}^n x^t x - \mu^t \mu$$

である．これに対して多変量正規分布 $\mathcal{N}(\mu,\Sigma)$ からデータをサンプリングしてレンジ $[-1,1]^d$ に切り戻す．このメカニズムを $\mathcal{M}_G : \mathcal{D} \to [-1,1]^d$ と表す．

入力と同じサイズのデータセットを出力する際の (α,ε) 性は，命題 2.4 の合成則を用いれば評価可能で，単純に ε を入力サンプル数倍すればよい．

3. 主結果

本稿では，合成データ生成メカニズム $\mathcal{M}_G$ が自然にもつ Rényi 差分プライバシー性を評価する．

3.1 主定理

本定理としてはデータセットに対して，分散共分散行列の最小固有値に制約があるとする．すなわち，$\sigma > 0$ を 1 つ固定し，データセット全体の集合を

$$\mathcal{D}_\sigma := \{D \in [-1,1]^{n\times d} \mid z \in S^{d-1}, {}^t z\Sigma_D z \ge \sigma\}$$

と考える．また，表記の簡略化のため $\tau := \frac{4d}{\sigma}$ とおく．

まず，秘匿 n 条件での設定の本研究の結果は次の定理である．元データのレコード数を n，隣接データセットのレコードを $n+1$ とする．

定理 3.1. 秘匿 n 条件の下で，$\alpha > 1$ とする．

$$\frac{n}{n+1} < \tau,\ \alpha < \min\{n+1, \frac{n^2}{\tau(n+1)-n}\}$$

が成り立つとする．このとき，合成データ生成メカニズム $\mathcal{M}_G$ は，後述の $\varepsilon_\alpha := \max\{\varepsilon_{\alpha 1}, \varepsilon_{\alpha 2}\}$ について $(\varepsilon_\alpha,\alpha)$-*RDP* を満たす．ただし，

[*5] この違いはクエリの sensitivity の計算結果に現れ，同じ安全性の下でも加えるノイズ量が変わるなど保護データの品質に影響を与えることがある．

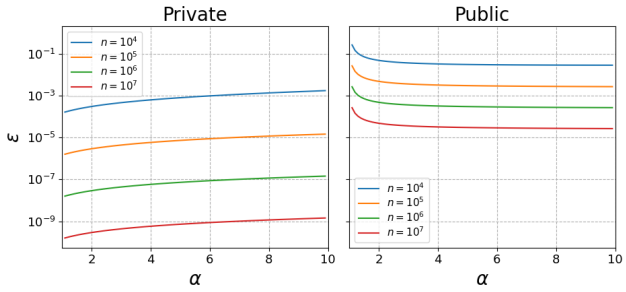

図 2 α - ε 曲線（$d = 6, \sigma = 0.01$）：左が秘匿 n 条件，右が公開 n 条件．縦軸は対数目盛で両グラフ共通．それぞれサンプル数 n を動かして 4 通りの曲線を引いた．

$$\varepsilon_{\alpha 1} = \frac{\alpha}{2} \cdot \frac{\tau}{(n+1)(n+1-\alpha)} + \frac{\alpha d}{2(\alpha-1)} \log \frac{n}{n+1} - \frac{d}{2(\alpha-1)} \log(1 - \frac{\alpha}{n+1}) - \frac{1}{2(\alpha-1)} \log \min\{1, \frac{1 + \alpha \frac{n\tau}{(n+1)(n+1-\alpha)}}{(1 + \frac{\tau}{n+1})^\alpha}\}$$

であり，

$$\varepsilon_{\alpha 2} = \frac{\alpha}{2} \cdot \frac{\tau}{n(n+\alpha) - \alpha(n+1)\tau} + \frac{\alpha d}{2(\alpha-1)} \log \frac{n+1}{n} - \frac{d}{2(\alpha-1)} \log(1 + \frac{\alpha}{n}) - \frac{1}{2(\alpha-1)} \log \min\{1, \frac{1 - \frac{\alpha(n+1)\tau}{(n+\alpha)n}}{(1 - \frac{\tau}{n})^\alpha}\}$$

である．

次に，公開 n 条件での設定では次の定理が成り立つ．

定理 3.2. 公開 n 条件の下で，$\alpha > 1$ とする．

$$\alpha < \frac{n^2}{\tau(n-1)}$$

が成り立つとする．このとき，合成データ生成メカニズム $\mathcal{M}_G$ は $\alpha > 1$ に対して，$(\varepsilon_\alpha, \alpha)$-*RDP* を満たす．ただし，

$$\varepsilon_\alpha = \frac{\alpha}{2} \cdot \frac{\tau}{n^2 - \alpha(n-1)\tau} + \frac{\alpha}{2(\alpha-1)} \log(1 + \frac{(n-1)\tau}{n^2}) - \frac{1}{2(\alpha-1)} \log(1 - \alpha \frac{(n-1)\tau}{n^2})$$

である．

3.2 数値代入

本項では，定理 3.1, 3.2 の式に具体的な値を代入し，その結果を観察する．Adult Dataset [1] の数値属性を取り出して $[-1, 1]$ へと正規化したデータで予備実験したところ $d = 6$ で，$\sigma_{min} = 0.02$ であったため，本数値実験では $d = 6$，$\sigma = 0.01$ とする．

まず，α と ε の関係性をグラフにしたのが，図 2 である．縦軸が対数目盛であることに注意する．秘匿 n 条件は n が大きくなるにつれ，ε が非常に小さくなることが確認できた．一般に Rényi divergence は α に対して単調増加であるが（補題 A.1.2），導出した評価式である α-ε 曲線には減少部があることに注意する．

表 1 入力サンプル数と同数のサンプル数の合成データを出力する場合の ε の値（$\alpha = 4, d = 6, \sigma = 0.01$）

n	10^4	10^5	10^6	10^7
秘匿 n 条件 ε	3535.17	62.5859	**5.80644**	**0.576462**
公開 n 条件 ε	6806.72	3263.22	3205.81	3200.58

表 2 秘匿 n 条件での (ε, δ)-DP 換算時の ε の値（$d = 6, \sigma = 0.01$）

$n = 10^6$ のとき

δ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$\alpha = 2$	7.49906	9.80164	12.1042	14.4068
$\alpha = 4$	**7.34149**	**8.10902**	**8.8766**	**9.64408**
$\alpha = 7$	10.9782	11.3620	11.7458	12.1295
$\alpha = 10$	15.1698	15.4257	15.6815	15.9374

$n = 10^7$ のとき

δ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$\alpha = 2$	4.893309	7.195894	9.498479	11.801064
$\alpha = 4$	2.111518	2.879047	3.646575	4.414104
$\alpha = 7$	**1.776821**	**2.160585**	2.54435	2.928114
$\alpha = 10$	1.954226	2.210069	**2.465912**	**2.721754**

$\alpha = 4$ としたとき，入力と同じサイズを出力するメカニズムが満たす (α, ε)-RDP の ε の値が表 1 である．秘匿 n 条件の行が，定理 3.1 の ε の値を n 倍したもので，公開 n 条件の行が定理 3.2 の ε の値を n 倍したものである．n 倍すれば良いというのは命題 2.4 の合成則に従っている．

数値代入の結果，公開 n 条件の定理 3.2 の評価式は現実的な設定ではあまり意味をなさないバウンドになっていることが分かった．秘匿 n 条件の設定の下では，入力サンプルが $n = 10^6$ 以上の時は $\varepsilon \leq 6$ で現実的な値となっていることが確認できた．特に $n = 10^7$ の時は $\varepsilon = 0.576$ と非常に小さい ε に対しても RDP が成り立つことが確認できた．

また，命題 2.5 より Rényi 差分プライバシーは (ε, δ)-差分プライバシーに換算できる（表 2）．表 2 より，$n = 10^6$ のとき $\alpha = 4$ と考えれば $\varepsilon \leq 10$ で (ε, δ)-差分プライバシーを満たしていることが分かった．さらに，$n = 10^7$ のときは，$\delta = 0.01$ とすれば $\varepsilon = 1.78$ で (ε, δ)-差分プライバシーを満たしていることが分かった．また，一般的に用いられる $\delta = 10^{-5}$ でも $\varepsilon = 2.72$ に対して，(ε, δ)-差分プライバシーを満たしていることが確認できた．

定理 3.2 の代入の結果が非常に大きな値になった原因としては，補題 4.9 の不等式評価がタイトではなかったことがあげられる．改善案としては，この評価を補題 4.5 の証明と同様に $\Sigma_1^{-1} X$ の非 0 固有値（この場合はランクが 2 なので 2 つある）を用いて，L_2 を 2 つの固有値の 2 変数関数として取りうる値の下界を評価する方法が考えられる．

4. 証明

本節では，定理 3.1，定理 3.2 の証明を行う．次の命題

4.1 を前提に，定理 3.1 は補題 4.2, 4.3, 4.4, 4.5 より従う．定理 3.2 は補題 4.6, 4.7, 4.8, 4.9 より従う．

証明の肝となるのは次の一般論である．

命題 4.1 (Manuel ら [4])**.** $\alpha > 1$ とする．このとき，二つの多変量正規分布 $\mathcal{N}(\mu_1, \Sigma_1), \mathcal{N}(\mu_2, \Sigma_2)$ に対して，

$$D_\alpha(\mathcal{N}(\mu_1, \Sigma_1)||\mathcal{N}(\mu_2, \Sigma_2)) = \frac{\alpha}{2}{}^t(\mu_1 - \mu_2)\Sigma_\alpha^{-1}(\mu_1 - \mu_2) - \frac{1}{2(\alpha-1)}\log\frac{|\Sigma_\alpha|}{|\Sigma_1|^{1-\alpha}|\Sigma_2|^\alpha}$$

が成り立つ．ただしここで

$$\Sigma_\alpha := (1-\alpha)\Sigma_1 + \alpha\Sigma_2$$

であり，また条件として

$$T_\alpha := \alpha\Sigma_1^{-1} + (1-\alpha)\Sigma_2^{-1}$$

が正定値である必要がある．

任意に隣接データセット D_1, D_2 をとってきたときの平均ベクトルを μ_1, μ_2，分散共分散行列を Σ_1, Σ_2 とおくと，$D_\alpha(\mathcal{N}(\mu_1, \Sigma_1)||\mathcal{N}(\mu_2, \Sigma_2))$ の上界の値を ε とおくと，Rényi 差分プライバシーの定義からメカニズム $\mathcal{M}_G$ が (α, ε)-Rényi 差分プライバシーを満たすことがわかる．

ここで，

$$L_1 := {}^t(\mu_1 - \mu_2)\Sigma_\alpha^{-1}(\mu_1 - \mu_2), \quad L_2 := \frac{|\Sigma_\alpha|}{|\Sigma_1|^{1-\alpha}|\Sigma_2|^\alpha}$$

とおくと，

$$D_\alpha(\mathcal{N}(\mu_1, \Sigma_1)||\mathcal{N}(\mu_2, \Sigma_2)) = \frac{\alpha}{2}L_1 - \frac{1}{2(\alpha-1)}\log L_2$$

であるので，L_1 の最大値と L_2 の最小値をそれぞれ求めれば ε が得られる．それぞれ秘匿/公開 n 条件の下で証明の具体的な計算が変わるので，それぞれ別に取り組む．流れは共通していて，次の通りである．まず，差分レコードを用いて平均ベクトルの差と分散共分散行列の差を表現する（補題 4.2，補題 4.6）．次に T_α の正定値性を確認する（補題 4.3，補題 4.7）．最後に L_1 の上界を計算し（補題 4.4，補題 4.8），L_2 の下界を計算する（補題 4.5，補題 4.9）．

4.1 秘匿 n 条件

隣接データセット $D_1, D_2 \in \mathcal{D}_\sigma$ を用意する．$\#D_1 = n, \#D_2 = n+s$（追加の場合 $s=1$，削除の場合 $s=-1$）とし，共通部分を $x_1, \ldots, x_n \in [-1,1]^d$，追加/削除されたレコードを $x \in [-1,1]^d$ とする[*6]．このとき，それぞれの平均ベクトルを μ_1, μ_2，分散共分散行列を Σ_1, Σ_2 とおく．全体を通して行列 Σ_1 の最小固有値を σ_{min} と置くが，$\sigma_{min} \geq \sigma$ であることに注意する．

[*6] 一般に $D_\alpha(P||Q) \neq D_\alpha(Q||P)$ であるが，「追加」に対しては $n \to n+1$ として「削除」を考えればよく，「削除」に対しても $n \to n-1$ として「追加」を考えればよいので，この設定で議論を進めても一般性は失われない．

補題 4.2 (秘匿 n 条件での差の表現)**.** 次の 2 式が成り立つ．

$$\mu_d := \mu_2 - \mu_1 = \frac{s}{n+s}x - \frac{s}{n(n+s)}\sum_{i=1}^n x_i$$

$$X := \Sigma_2 - \frac{n}{n+s}\Sigma_1 = \frac{ns}{(n+s)^2}(x-\mu_1)^t(x-\mu_1)$$

証明. 計算で確かめられる． □

X はランクが 1 の行列で，$s=1$ の時は半正定値，$s=-1$ の時は半負定値である．

補題 4.3 (秘匿 n 条件での T_α の正定値性)**.** 次の 2 つの不等式が成り立つとき，T_α は正定値行列である．

$$\frac{n-1}{n} < \tau, \quad \alpha < \min\{n+1, \frac{(n-1)^2}{\tau n - (n-1)}\}.$$

証明. $T_\alpha = \Sigma_1\Sigma_\alpha\Sigma_2 = \Sigma_2\Sigma_\alpha\Sigma_1$ であるので，補題 A.3.2 より，T_α の正定値性は Σ_α の正定値性に帰着する．補題 4.2 より，

$$\Sigma_\alpha = (1-\alpha)\Sigma_1 + \alpha(\frac{n}{n+s}\Sigma_1 + X) = (1 - \frac{s\alpha}{n+s})\Sigma_1 + \alpha X$$

が成り立つ．$s=1$ の時は，X も半正定値であるので，十分条件として $\alpha < n+1$ があげられる．$s=-1$ の時を考える．任意にとったノルム 1 のベクトル $z \in \mathbb{R}^d$ に対して，${}^t z\Sigma_\alpha z$ の最小値が正である条件を求めればよい．ベクトル $x - \mu_1$ は半径 $2\sqrt{d}$ の球の中に入っていると考えられるので，最小値は z が Σ_1 の最小固有値 σ_{min} の固有ベクトルに並行で，また，$x - \mu_1$ が z と平行の場合である．つまり，Σ_α が正定値であるためには，

$$\begin{aligned} {}^t z\Sigma_\alpha z &= (1 + \frac{\alpha}{n-1})\sigma_{min} - \alpha\frac{n}{(n-1)^2}4d \\ &= \sigma_{min} - \alpha\cdot\frac{4dn - (n-1)\sigma_{min}}{(n-1)^2} \\ &= \sigma - \alpha\cdot\frac{4dn - (n-1)\sigma}{(n-1)^2} > 0 \end{aligned}$$

であればよい．主張の 2 つの不等式が成り立つとき，上記の式も成り立つ． □

補題 4.4 (秘匿 n 条件での L_1 の上界)**.** $s=1$ のとき，

$$L_1 \leq \frac{\tau}{(n+1)(n+1-\alpha)}$$

であり，$s=-1$ のとき，

$$L_1 \leq \frac{\tau}{(n-1)(n-1+\alpha) - \alpha n\tau}$$

である．

証明. 補題 4.2 より μ_d は半径 $\frac{2\sqrt{d}}{n+s}$ の球に入っていて，補題 4.3 より Σ_α は正定値であるので，${}^t\mu_d\Sigma_\alpha^{-1}\mu_d$ の最大値は，単位ベクトル $z \in \mathbb{R}$ に対する ${}^t z\Sigma_\alpha z$ の最小値の逆数を $\frac{4d}{(n+s)^2}$ 倍したものである．いま，

$$ {}^tz\Sigma_\alpha z = {}^tz(1-\frac{s\alpha}{n+s})\Sigma_1 z + \frac{s\alpha n}{(n+s)^2}({}^tz(x-\mu_1))^2 $$

である．$s=1$ のときの最小値は

$$ (1-\frac{\alpha}{n+1})\sigma_{min} $$

である．$s=-1$ とすると，いま，$x-\mu_1$ は半径 $2\sqrt{d}$ の球に入っているので，最小値は

$$ (1+\frac{\alpha}{n-1})\sigma_{min} - \frac{\alpha n}{(n-1)^2}\cdot 4d $$

である．ゆえに，主張の不等式を得る． □

補題 4.5 (秘匿 n 条件での L_2 の下界)**.**

$$ L_2 \geq \frac{(1-\frac{s\alpha}{n+s})^d}{(\frac{n}{n+s})^{\alpha d}} \cdot \min\{1, \frac{1+\frac{\alpha n s\tau}{(n+s-s\alpha)(n+s)}}{(1+\frac{s\tau}{n+s})^\alpha}\}. $$

が成り立つ．

証明. まず L_2 は，

$$ \begin{aligned} L_2 &:= \frac{|(1-\frac{s\alpha}{n+s})\Sigma_1 + \alpha X|}{|\Sigma_1|^{1-\alpha}|\frac{n}{n+s}\Sigma_1 + X|^\alpha} \\ &= \frac{(1-\frac{s\alpha}{n+s})^d|I + \frac{n+s}{n+s-s\alpha}\alpha\Sigma_1^{-1}X|}{(\frac{n}{n+s})^{\alpha d}|I+\frac{n+s}{n}\Sigma_1^{-1}X|^\alpha} \end{aligned} $$

と変形できるが，いま，X がランク 1 で Σ_1^{-1} が正則行列なので $\Sigma_1^{-1}X$ もランク 1 の行列である．ゆえに非 0 の固有値は一つしかないため，それを λ とする．また，$A := (1-\frac{s\alpha}{n+s})^d/(\frac{n}{n+s})^{\alpha d}$ とおく．他の固有値はすべて 0 であるので，

$$ L_2 = \frac{1+\frac{n+s}{n+s-s\alpha}\alpha\lambda}{(1+\frac{n+s}{n}\lambda)^\alpha}\cdot A $$

である．これを λ で微分すると

$$ \frac{\partial L_2}{\partial\lambda} = \alpha(\alpha-1)\frac{n+s}{n(n+s-s\alpha)}\cdot\frac{s-(n+s)\lambda}{(1+\frac{n+s}{n}\lambda)^{\alpha+1}}\cdot A $$

が得られる．これより，$\frac{s}{n+s} < \lambda$ のときに $\frac{\partial L_2}{\partial\lambda} > 0$, $\frac{s}{n+s} > \lambda$ のときに $\frac{\partial L_2}{\partial\lambda} < 0$ である．よって，λ は取りうる値のレンジの端で L_2 が最小値となる．

次に $\Sigma_1^{-1}X$ の唯一の非負固有値である λ のレンジを求める．Σ_1 は正定値であるのでスペクトル分解できる；

$$ \Sigma_1 = \sum_{i=1}^d \sigma_i p_i {}^t p_i. $$

ここで，σ_i が Σ_1 の固有値で，p_i がノルムが 1 の固有ベクトル．ここで，p_i たちは $\mathbb{R}^d$ の基底になるので，

$$ x-\mu_1 = \sum_{i=1}^d r_i p_i $$

と表せる．両辺を 2 乗すると $4d \geq \sum_{i=1}^d r_i^2 > 0$ という条件が得られる．このとき，$e_1 := \sum_{i=1}^d \frac{r_i}{\sigma_i}p_i$ とおくと，

$$ \begin{aligned} \Sigma_1^{-1}Xe_1 &= \Sigma_1^{-1}\frac{ns}{(n+s)^2}\sum_{i=1}^d r_i p_i((x-\mu_1)\cdot e_1) \\ &= \frac{ns}{(n+s)^2}((x-\mu_1)\cdot e_1)e_1 \\ &= \frac{ns}{(n+s)^2}(\sum_{i=1}^d \frac{r_i^2}{\sigma_i})e_1 \end{aligned} $$

であるので $\lambda = \frac{ns}{(n+s)^2}\sum_{i=1}^d \frac{r_i^2}{\sigma_i}$ である．このとき，$s=1$ なら，$0 < \lambda \leq \frac{4dn}{(n+1)^2\sigma_{min}} \leq \frac{4dn}{(n+1)^2\sigma}$，$s=-1$ なら $-\frac{4dn}{(n-1)^2\sigma} \leq -\frac{4dn}{(n-1)^2\sigma_{min}} \leq \lambda < 0$ となる．これらの事実より，主張が成り立つ． □

4.2 公開 n 条件の設定

隣接データセット D_1, D_2 を用意する．$\#D_1 = \#D_2 = n$ とし，共通部分を $x_1, \ldots, x_{n-1} \in [-1,1]^d$，残りの一つのレコードは D_1 が $x \in [-1,1]^d$ で D_2 が $y \in [-1,1]^d$ であるとする．このとき，それぞれの平均ベクトルを μ_1, μ_2，分散共分散行列を Σ_1, Σ_2 とおく．秘匿 n 条件の時と同様，$\sigma_{min} \geq \sigma$ であることに注意する．

補題 4.6 (公開 n 条件での差の表現)**.** 次の 2 式が成り立つ．

$$ \mu_d := \mu_2 - \mu_1 = \frac{1}{n}(y-x), $$

$$ X := \Sigma_2 - \Sigma_1 = \frac{n-1}{n^2}(y-\mu')^t(y-\mu') - \frac{n-1}{n^2}(x-\mu')^t(x-\mu'). $$

ただし，ここで $\mu' := \frac{n}{n-1}\sum_{i=1}^{n-1} x_i$ である．

証明. 計算で確かめられる． □

補題 4.7 (公開 n 条件での T_α の正定値性)**.** 次の不等式が成り立つとき，T_α は正定値行列である．

$$ \alpha < \frac{n^2}{\tau(n-1)} $$

証明. 補題 4.3 の証明と同様に，T_α の正定値性は Σ_α の正定値性に帰着する．補題 4.6 より，

$$ \Sigma_\alpha = (1-\alpha)\Sigma_1 + \alpha(\Sigma_1 + X) = \Sigma_1 + \alpha X $$

が成り立つ．単位ベクトル $z \in \mathbb{R}^d$ に対する，${}^tz\Sigma_\alpha z$ の最小値を考察すればよい．よって正定値部分は 0 でよいので $x = \frac{n}{n-1}\mu'$ とでき，補題 4.3 とほぼ同様の議論より，

$$ {}^tz\Sigma_\alpha z = \sigma_{min} - \alpha\frac{n-1}{n^2}4d \geq \sigma - \alpha\frac{n-1}{n^2}4d > 0 $$

が成り立てばよく主張が成り立つ． □

補題 4.8 (公開 n 条件での L_1 の上界)**.** L_1 の最大値は

$$ \frac{\tau}{n^2-\alpha(n-1)\tau} $$

証明. 補題 4.4 の証明と同様の方針で示す．単位ベクトル $z \in \mathbb{R}^d$ に対して

$${}^tz\Sigma_\alpha z = {}^tz\Sigma_1 z + \frac{\alpha(n-1)}{n^2}({}^tz(y-\mu'))^2 - \frac{\alpha(n-1)}{n^2}({}^tz(x-\mu'))^2$$

であるので最小値は

$$\sigma_{min} - \frac{\alpha(n-1)}{n^2} \cdot 4d$$

□

補題 4.9 (公開 n 条件での L_1 の上界)**.**

$$L_2 \geq \frac{1 - \frac{\alpha(n-1)\tau}{n^2}}{(1 + \frac{(n-1)\tau}{n^2})^\alpha}$$

が成り立つ.

証明. $X_1 := \frac{n-1}{n^2}(y-\mu')^t(y-\mu')$, $X_2 := \frac{n-1}{n^2}(x-\mu')^t(x-\mu')$ とおくと $X = X_1 - X_2$ である．いま,

$$L_2 := \frac{|\Sigma_\alpha|}{|\Sigma_1|^{1-\alpha}|\Sigma_2|^\alpha} = \frac{|\Sigma_1 + \alpha X|}{|\Sigma_1|^{1-\alpha}|\Sigma_1 + X|^\alpha}$$

であるが，ここで補題 A.3.5, A.3.6 より,

$$|\Sigma_1 + \alpha X| \geq |\Sigma_1 - \alpha X_2|, |\Sigma_1 + X| \leq |\Sigma_1 + X_1|$$

が成り立つ．ゆえに,

$$L_2 \geq \frac{|\Sigma_1 - \alpha X_2|}{|\Sigma_1|^{1-\alpha}|\Sigma_1 + X_1|^\alpha} = \frac{|I - \alpha\Sigma_1^{-1}X_2|}{|I + \Sigma_1^{-1}X_1|^\alpha}$$

であり，いま，X_1 と X_2 は独立に動かせるので，それぞれ $\Sigma_1^{-1}X_i$ が最大固有値の出すようにすればよいので

$$L_2 \geq \frac{1 - \frac{4\alpha d(n-1)}{\sigma n^2}}{(1 + \frac{4d(n-1)}{\sigma n^2})^\alpha}$$

となり，主張を得る． □

5. まとめ

本研究では，多変量正規分布に基づく合成データ生成技術が自然に持つ Rényi 差分プライバシー性を理論的に評価した．定理 3.1 で得た評価式は入力データ数 $n = 10^7$ 程度の時は，小さい ε に対して RDP 性を保証することを確認できた．また，今後の課題として，Lin らの研究 [7] などと同様に深層学習による生成モデルに対しても同様の評価を行うことがあげられる.

参考文献

[1] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[2] Cynthia Dwork. Differential privacy. In *International Colloquium on Automata, Languages, and Programming*, pp. 1–12. Springer, 2006.

[3] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, Vol. 9, No. 3-4, pp. 211–407, 2014.

[4] Manuel Gil, Fady Alajaji, and Tamas Linder. Rényi divergence measures for commonly used univariate continuous distributions. *Information Sciences*, Vol. 249, pp. 124–131, 2013.

[5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, Vol. 27, , 2014.

[6] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In Yoshua Bengio and Yann LeCun, editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[7] Zinan Lin, Vyas Sekar, and Giulia Fanti. On the privacy properties of gan-generated samples. In *International Conference on Artificial Intelligence and Statistics*, pp. 1522–1530. PMLR, 2021.

[8] Ryan McKenna, Gerome Miklau, and Daniel Sheldon. Winning the nist contest: A scalable and general approach to differentially private synthetic data. *arXiv preprint arXiv:2108.04978*, 2021.

[9] Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275. IEEE, 2017.

[10] Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pp. 1530–1538. PMLR, 2015.

[11] Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 10, No. 05, pp. 557–570, 2002.

[12] Shun Takagi, Tsubasa Takahashi, Yang Cao, and Masatoshi Yoshikawa. P3gm: Private high-dimensional data release via privacy preserving phased generative model. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pp. 169–180. IEEE, 2021.

[13] Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante, and Kalyan Veeramachaneni. Modeling tabular data using conditional gan. In *Advances in Neural Information Processing Systems*, 2019.

[14] Jun Zhang, Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava, and Xiaokui Xiao. Privbayes: Private data release via bayesian networks. *ACM Trans. Database Syst.*, Vol. 42, No. 4, October 2017.

[15] 伊原一, 田中雅行, 北林三就. 一般用ミクロデータ就業構造基本調査版の概要～系統抽出による疑似標本データ～. 日本人口学会第 70 大会, `https://www.nstac.go.jp/services/society_paper/30_06_03.pdf`, 2018.

[16] 岡田莉奈, 正木彰伍, 長谷川聡, 田中哲士. 統計値を用いたプライバシ保護擬似データ生成手法. コンピュータセキュリティシンポジウム 2017 論文集, 第 2017 巻, oct 2017.

[17] 三浦尭之, 紀伊真昇, 芝原俊樹, 市川敦謙, 千田浩司. 合成データ生成のランダム性に内在する安全性の評価. コンピュータセキュリティシンポジウム 2021 論文集, pp. 268–275, oct 2021.

[18] 寺田雅之, 山口高康, 本郷節之. 匿名化個票開示への差分プライバシの適用. 情報処理学会論文誌, Vol. 58, No. 9, pp. 1483–1500, sep 2017.

本稿で参照した URL はすべて著者が 2022 年 8 月 23 日に閲覧可能を確認済みである.

付　　録

A.1 Rényi Divergence の基本性質 [9]

Rényi 差分プライバシーを考える上で重要になる Rényi

divergence の性質を紹介する．本節では共通して，P, Q, R を $\mathbb{R}^d$ 上の確率分布とする．

補題 A.1.1 (α の両端での形)**.** 次の 2 式が成り立つ．

$$D_1(P||Q) := \lim_{\alpha\to 1} D(P||Q)_\alpha = E_{x\sim P} \log \frac{P(x)}{Q(x)},$$

$$D_\infty(P||Q) := \lim_{\alpha\to \infty} D(P||Q)_\alpha = \sup_{x\in suppQ} \log \frac{P(x)}{Q(x)}.$$

いま，$D_1(P||Q)$ は KL-divergence であり，$D_\infty(P||Q)$ は max-divergence である．

特に，$D_\infty(P||Q)$ はこれはポイントワイズな純粋な確率比の上限値であり，log 関数は大小関係を保存するので，これを抑えることは通常の ε-差分プライバシーを考えることと等価となる（補題 A.2.2)．

補題 A.1.2. Rényi divergence は下記 4 つの性質を満たす．

- （非負性）$1 < \alpha$ のとき

$$D_\alpha(P||Q) \geq 0.$$

- （単調性）$1 < \alpha < \beta$ のとき

$$D_\alpha(P||Q) \leq D_\beta(P||Q).$$

- （確率保存）$\alpha > 1$，事象 $A \subset \mathbb{R}^d$ に対して

$$\Pr[X_P \in A] \leq (e^{D_\alpha(P||Q)} \Pr[X_Q \in A])^{\frac{\alpha-1}{\alpha}}.$$

- （弱三角不等式）任意の $\alpha > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ に対して，

$$D_\alpha(P||Q) \leq \frac{\alpha - \frac{1}{p}}{\alpha - 1} D_{p\alpha}(P||R) + D_{q(\alpha-\frac{1}{p})}(R||Q).$$

A.2 RDP と従来の DP の関係

本節では，Rényi 差分プライバシーと，現在広く使われている (ε, δ)-差分プライバシーの関係について説明する．(ε, δ)-差分プライバシー [2] の定義は次の通りである．

定義 A.2.1 (差分プライバシー [2])**.** ランダム化関数 $\mathcal{M} : \mathcal{D} \to \mathcal{Y}$ が (ε, δ)-差分プライバシー（(ε, δ)-*DP*）を満たすとは次が成り立つことをいう．任意の隣接データセット $D, D' \in \mathcal{D}$，任意の $S \subset \mathcal{Y}$ に対して，

$$\Pr[\mathcal{M}(D) \in S] \leq e^\varepsilon \Pr[\mathcal{M}(D') \in S] + \delta.$$

$\delta = 0$ のとき，特に ε-差分プライバシーが成り立つという．

補題 A.1.1 より下記補題が成り立つ．

補題 A.2.2. ランダム化関数 $\mathcal{M}$ が (∞, ε)-*RDP* を満たすことと，ε-*DP* を満たすことは同値である．

また，補題 A.1.2 の単調性より下記補題が成り立つ．

補題 A.2.3. (α, ε)-*RDP* を満たすランダム化関数 $\mathcal{M}$ は，$\alpha' \leq \alpha$，$\varepsilon \leq \varepsilon'$ に対して，$\mathcal{M}$ は (α', ε')-*RDP* を満たす．

これらの二つの補題から，ε **は小さいほど保護が厳しく，** α **は大きいほど保護が厳しくなる**ことが確認できる．α を無限大に飛ばすと ε-DP になることから，逆に，α を小さくすることで ε-DP を緩和しているとみることができる．

A.3 線形代数的な諸性質

証明に用いた線形代数の性質の定義や補題を紹介する．

定義 A.3.1. d 次対称行列 A に対して，次の (1) と (2) は同値であり，どちらか（どちらも）を満たすと**正定値（半正定値）**であるという．

(1) 任意の $x \in \mathbb{R}^d \backslash \{0\}$ で ${}^t xAx > 0$（≥ 0）が成り立つ．

(2) すべての固有値が正（非負）である．

命題 A.3.2. A, B, C を正定値な実数係数対称行列とする．このとき，行列 ABC は対称行列ならば，正定値である．

証明. $D := ABC = CBA$ とおく．いま，C は正定値対称行列なので，スペクトル分解が可能であり，その固有値を正の平方根に置き換えた平方根行列 S を得ることができる．この S は対称行列で $C = S^2$ が成り立つ．このとき，

$$S^{-1}DS^{-1} = S^{-1}AS^{-1}SBS = SBSS^{-1}AS^{-1}$$

と分解できるので，行列 $S^{-1}AS^{-1}$ と SBS に対して，補題 A.3.3 と A.3.4 を適用することにより，$S^{-1}DS^{-1}$ は正定値，すなわち D は正定値であることがわかる． □

補題 A.3.2 に対し，$C = I_d$ とおけば次の補題が成り立つ．

補題 A.3.3. A, B を正定値な実数係数対称行列とする．このとき，AB は対称行列ならば，正定値である．

次の補題も定義より明らか．

補題 A.3.4. A を正定値な実数係数対称行列とする．このとき，任意の同じサイズの正則行列 S に対して，${}^t SAS$ は正定値行列である．

補題 A.3.5. A を正定値，B を半正定値とする．このとき，

$$|A + B| \geq |A| + |B|$$

が成り立つ．

証明. A は正定値対称行列なので，対称行列 S が存在して $A = S^2$．ゆえに

$$\begin{aligned} |A + B| &= |S||I + S^{-1}BS^{-1}||S| \\ &\geq |A|(1 + |S^{-1}BS^{-1}|) \text{ (by positive definite)} \\ &= |A| + |B| \end{aligned}$$

が成り立つ． □

補題 A.3.6. A を正定値，C を半負定値とし，$A + C$ が正定値行列とする．このとき，

$$|A| \geq |A + C|$$

が成り立つ．

証明. $-C$ が半正定値になるので，$A + C$ と $-C$ について補題 A.3.5 を適用すると

$$|A + C + (-C)| \geq |A + C| + |-C| \geq |A + C|$$

が得られる． □