Computer Security Symposium 2019
21 - 24 October 2019

ヒストグラムに対する最適な維持置換攪乱

三浦 尭之[1,a)]　長谷川 聡[1]

概要：本稿では，維持置換攪乱を用いたヒストグラムのランダム化の有用性向上を図った．具体的には，固定したヒストグラムに対して，ε-差分プライバシーというプライバシー保護指標を満たしつつ，変換後のヒストグラムが元のものに最も近くなるような維持置換攪乱を求めるアルゴリズムを考案した．

キーワード：ヒストグラム開示，維持置換攪乱，Pk-匿名性，ε-差分プライバシー，二次計画法

The optimal PRAM for a histogram

Takayuki Miura[1,a)]　Satoshi Hasegawa[1]

Abstract: In this paper, we discuss a way to improve the post randomization method (PRAM) for histograms. By the proposal algorithm, we can obtain the PRAM for a fixed histogram which satisfies ε-differential privacy and outputs the nearest histogram to the original one.

Keywords: disclosure of a histogram, PRAM, Pk-anonymity, ε-differential privacy, quadratic programming

1. はじめに

近年，ICT 技術やデータマイニング技術等の発達により，個人の購買履歴，位置・移動情報，生体情報といった情報の収集，利活用が容易になった．それに伴い，様々な新しいサービスの創出が期待されているが，一方でそのような情報の利活用に伴う個々人のプライバシー侵害のリスクにも関心が集まっている．

収集したデータを第三者に提供する，もしくは公開する場合を考える（これらをまとめて**開示**と呼ぶことにする）．こうした開示の際，データ内の個人のプライバシーは守られなければならない．そのため，開示データから推定できる個人情報を制限するようデータを加工する必要がある．しかし一方で，開示データの利用価値がなくなるほどの加工をしてしまうと開示の意味がなくなってしまう．

このようなトレードオフを前提に，様々なプライバシー保護メカニズムが考案されており [1], [2]，現在最も用いられている保護手法の一つに**維持置換攪乱**という手法がある [3]．これは定められた確率に従い，データベースの各レコードの内容をそのままにしたり（維持），書き換えたり（置換）する保護手法である．従来の維持置換攪乱では，各レコードの情報の維持確率を一律に同じものを採用しており，扱うデータによっては有用性を著しく損ねていないかという懸念があった．

本稿ではこの維持確率に自由度を持たせ，維持置換攪乱の有用性向上を図る検討を行った．具体的には，任意に選んだヒストグラム[*1]に対し，一定のプライバシー保護指標を満たすものの中で最も有用性を保つものを導出するアルゴリズムを提案した．

提案アルゴリズムで得られた維持置換攪乱を従来方式のものと数値実験により比較した結果，各種ヒストグラムで有用性の向上が確認できた（4 節）．このことから，「事前に一定の傾向が知られているようなヒストグラムに対しては，本稿で導出された維持置換攪乱が有効である」ということがわかる．

また，本稿の考察の特色の一つに，解くべき最適化問題

[1] 日本電信電話株式会社 NTT セキュアプラットフォーム研究所
NTT Secure Platform Laboratories, NTT Corporation
[a)] takayuki.miura.br@hco.ntt.co.jp

[*1] ヒストグラムは幅広くデータ分析に利用可能であり，応用上重要なデータ形式である．

を幾何学的な条件に捉え直して考察したことがある．このような観点は今後，維持置換攪乱の理論的な限界を考察し，包括的な理論を完成させる際に役立つと期待できる．

2. 準備

本節では，次節以降の議論の準備として，必要になる諸概念の定義や性質の紹介をする．

2.1 記号と基礎概念

安全性やランダム化などの概念は基本的に数学の言葉を用いて議論される．そのため，まず，テーブルやその保護メカニズムなどの基本概念を数学的な言葉で定式化する．

記号 2.1. 本稿全体を通して用いる数学記号を導入する．

· $\mathbb{Z}_{\geq 0} := \{x \in \mathbb{Z} | x \geq 0\}$：非負の整数全体
· $\mathbb{Z}_{>0} := \{x \in \mathbb{Z} | x > 0\}$：正の整数全体
· $\mathbb{R}_{>0} := \{x \in \mathbb{R} | x > 0\}$：正の実数全体
· $[n] := \{1, 2, \ldots, n\}$ （ここで $n \in \mathbb{Z}_{>0}$）
· $||\cdot||$：$\ell 2$ ノルム
· $\#X$：集合 X の要素数
· $\mathfrak{S}_N := \{\ \sigma : [N] \to [N]$ 全単射 $\}$（N 次対称群）
· $\Pr[A]$：事象 A が起きる確率
· $\Pr[A|B]$：事象 B が起きた状況下で事象 A が起きる確率（条件つき確率）
· $M_d(\mathbb{R})$：実数成分からなる d 次正方行列全体の集合
· $[a, b] := \{x \in \mathbb{R} | a \leq x \leq b\}$：閉区間
· $f^{-1}(y) := \{x \in X | f(x) = y\}$：逆像（$f : X \to Y$ は写像）
· $g \circ f : X \to Z$：写像の合成（$g \circ f(x) := g(f(x))$）

定義 2.2 (テーブル)．$\mathcal{A}$ を属性の集合，各属性 $a \in \mathcal{A}$ に対して，$\mathcal{V}_a$ を属性値の集合とし，$\mathcal{V} := \prod_{a \in \mathcal{A}} \mathcal{V}_a$ とおく．$N \in \mathbb{Z}_{>0}$ とする．このとき，写像 $\tau : [N] \to \mathcal{V}$ をレコード数 N の**テーブル**と呼ぶ．特に，本稿ではヒストグラムのみを考察対象として扱うので，$\#\mathcal{A} = 1$（つまり $\mathcal{V} = [d]$）と仮定する．また，$N, d \in \mathbb{Z}_{>0}$ に対して，テーブル全体の集合を $T_{N,d} := \{\tau : [N] \to [d]\}$ と書くことにする．

定義 2.3 (テーブル間の距離)．二つのテーブル $\tau, \tau' \in T_{N,d}$ に対して，テーブル間の距離 $d : T_{N,d} \times T_{N,d} \to \mathbb{R}$ を $d(\tau, \tau') := \#\{i \in [N] | \tau(i) \neq \tau'(i)\}$ と定める．これは，τ, τ' の中の互いに異なるレコードの個数を意味する．

以降，$T_{N,d}$ のことを単に T と書くことにする．

定義 2.4 (ヒストグラム)．テーブル $\tau : [N] \to [d]$ に対し，ベクトル $(\#\tau^{-1}(1), \ldots, \#\tau^{-1}(d)) \in \mathbb{Z}^d$ のことを τ の**ヒストグラム**という．

定義 2.5 (テーブル保護メカニズム)．写像 $m : T \to \{\ T$-値確率変数 $\}$ を**テーブル保護メカニズム**と呼ぶ．これは入力されたテーブル τ に対して，τ に応じた確率分布に従って，テーブル τ' を出力するメカニズムを表している．

2.2 プライバシー保護指標と維持置換攪乱

k 匿名性は最も基本的なテーブルの安全性指標である．

定義 2.6 (k 匿名性, [4])．$k \in \mathbb{Z}_{>0}$ とする．テーブル $\tau \in T$ が任意の $i \in [N]$ に対して，

$$\#\tau^{-1}(\tau(i)) \geq k$$

を満たすとき，τ は k **匿名性**を持つという．

これは「テーブルのどの個人に対しても，同じ属性をもつものが少なくとも k 人は存在するため，特定を試みても k 人以下には絞り込めない」ということを意味している．

これに対して，「どのような攻撃者に対しても，加工後のテーブルから個人を $\frac{1}{k}$ より大きい確率で特定することができない」ということを意味する安全性指標として，五十嵐らが提案した Pk-匿名性という性質がある．

定義 2.7 (Pk-匿名性, [1])．$k \in \mathbb{Z}_{>0}$ とする．U を一様分布に従う $\mathfrak{S}_N$-値確率変数とする．保護メカニズム m が任意の T-値確率変数 X_T，テーブル $\tau \in T$, $i, j \in [N]$ に対し

$$\Pr[U(i) = j \mid m(X_T) = \tau \circ U] \leq \frac{1}{k}$$

を満たすとき，m は Pk-**匿名性**を満たすという．

k 匿名性がテーブルに関する安全性指標であることに対して，Pk-匿名性が保護メカニズムに関する安全性指標であることを注意しておく．

一方で，Dwork らによって 2006 年に提唱された，差分プライバシーというプライバシー保護指標が，様々な応用，実用とともに現在も盛んに研究されている [2].

定義 2.8 (差分プライバシー, [2])．$\varepsilon \in \mathbb{R}_{>0}$ とする．保護メカニズム m が任意の $d(\tau, \tau') = 1$ になるような $\tau, \tau' \in T$ と任意の $S \subset T$ に対して，

$$\Pr[m(\tau) \in S] \leq e^{\varepsilon} \Pr[m(\tau') \in S]$$

を満たすとき，m は ε-**差分プライバシー**を満たすという．

これは，「テーブルから一人分の情報を削除しても，メカニズムの出力の差があまり区別できない」ということを意味する安全性指標である．

定義 2.9 (遷移行列, [3])．$P = (p_{ij}) \in M_d(\mathbb{R})$ をすべての成分が非負な正方行列とする．任意の $j \in [d]$ に対して，第 j 列の成分の総和が 1，つまり，

$$\sum_{i=1}^{d} p_{ij} = 1$$

となっているとき，P を**遷移行列**とよぶ．ここで，p_{ij} は第 j 属性が第 i 属性に書き換わる確率を表しており，特に p_{ii} は第 i 属性の維持確率を表している．

次を満たす遷移行列 $P \in M_d(\mathbb{R})$ を **A 型**と名付ける．

$p_1, \ldots, p_d \in [0, 1]$ が存在して，

$$p_{ij} = \begin{cases} p_i & \text{if } i = j \\ \dfrac{1 - p_i}{d - 1} & \text{else} \end{cases}$$

が成り立つ．

ここで，A 型というのは全く一般的な用語ではなく，本稿の中のみのものであることを注意しておく．

定義 2.10 (維持置換攪乱, [3])**.** P を遷移行列とする．テーブル保護メカニズム m で，任意のテーブル $\tau \in T$，レコード $r \in [N]$，$j \in [d]$ に対して，$\Pr[m(\tau)(r) = j] = p_{j\tau(r)}$ となるようなものを P に従う**維持置換攪乱**という．

本稿の提案手法と比較するために，一般に用いられる維持置換攪乱の方式を定義する．

定義 2.11 (従来型の維持置換攪乱)**.** $P \in M_d(\mathbb{R})$ を遷移行列，m を P に従う維持置換攪乱とする．このとき，ある $p \in [d, 1]$ が存在して，任意の $i, j \in [d]$ に対して，

$$p_{ij} = \begin{cases} p & \text{if } i = j \\ \dfrac{1-p}{d-1} & \text{else} \end{cases}$$

が成り立つとき，m を**従来型**の維持置換攪乱という．

ここで，$d \leq p$ という条件は，維持確率が攪乱確率以上になる，すなわち $p \geq \frac{1-p}{d-1}$ という条件を表している．

補題 2.12. $v \in \mathbb{Z}^d$ をヒストグラム，$P \in M_d(\mathbb{R})$ を遷移行列とする．P に従う維持置換攪乱を v にかけた際の出力 v' の期待値は，行列とベクトルの積 $Pv \in M_d(\mathbb{R})$ である．

証明. Pv を計算して各成分を見れば確かめられる． □

次の定理は維持置換攪乱における Pk-匿名性と ε-差分プライバシーという二つの保護指標を結ぶ．証明は五十嵐らの論文 [1] と補題 3.2 より導かれるので省略する．

定理 2.13. m を遷移行列 P に従う維持置換攪乱とする．このとき，$\varepsilon := \min_{\varepsilon \in \mathbb{R}_{>0}}\{\ m$ は ε-差分プライバシーを満たす $\}$，$k := \max_{k \in \mathbb{R}_{>0}}\{\ m$ は Pk-匿名性を満たす $\}$ とおく．次の等式が成り立つ．

$$k - 1 = (N - 1) \cdot \frac{1}{e^{2\varepsilon}}$$

この定理により，**一つのテーブルに対する維持置換攪乱においては，プライバシー保護指標を ε として考えることも，k として考えることも等価であると言える．**

3. 提案手法

本節を通して，テーブル $\tau : [N] \to [d]$，プライバシー保護指標 $\varepsilon \in \mathbb{R}_{>0}$ は一つとってあるものとする．$v \in \mathbb{Z}^d$ を τ のヒストグラムとする．

3.1 問題の定式化

本稿では有用性向上を次の問題を解くこととし議論する．

問題 3.1. ε-差分プライバシーを満たす維持置換攪乱を定める A 型の遷移行列 $P \in M_d(\mathbb{R})$ の中で，$||Pv - v||$ を最小化するものを求めよ．

補題 2.12 より，P に従う維持置換攪乱を v にかけたものの期待値は Pv であったので，$||Pv - v||$ は保護処理前後のヒストグラムの差と解釈できる．つまり，$||Pv - v||$ の最小化とは，保護処理による誤差の最小化を表している．これより上記の問題は，ε-差分プライバシーを満たす P の中で，**有用性を最大化する P を求めることと考えられる．**

3.2 問題の言い換え（幾何的な捉え直し）

本項では，問題 3.1 の言い換えをする．

次の補題は Wang らによって示された維持置換攪乱が満たす差分プライバシーを決定する重要な定理である [5].

補題 3.2 ([5], 定理 4.1)**.** $\varepsilon \in \mathbb{R}_{>0}$ を決める．遷移行列 $P = (p_{ij})$ に対して，

$$e^{\varepsilon} \geq \max_{1 \leq i \leq d} \frac{\max_{1 \leq j \leq d} p_{ij}}{\min_{1 \leq j \leq d} p_{ij}}$$

が成り立つとき，P に従う維持置換攪乱は ε-差分プライバシーを満たす．

注意 3.3. この評価は ε に対して，最適なものである．

定義 3.4 (2 次元 ε-d 領域)**.** $\varepsilon \in \mathbb{R}_{>0}, d \in \mathbb{Z}_{>0}$ に対して，$R_{\varepsilon,d} := \{(x, y) \in [0, 1]^2 \mid (x, y)$ は次の 6 つの不等式を満たす $\}$ と定め，**2 次元 ε-d 領域**と呼ぶ．

$$\begin{cases} y \leq e^{\varepsilon} x & (1) \\ y \geq e^{-\varepsilon} x & (2) \\ y \geq e^{-\varepsilon}(d-1)(1-x) & (3) \\ y \leq e^{\varepsilon}(d-1)(1-x) & (4) \\ y \geq 1 - \dfrac{e^{\varepsilon}}{d-1} x & (5) \\ y \leq 1 - \dfrac{e^{-\varepsilon}}{d-1} x & (6) \end{cases}$$

これを平面に図示したものが図 1 である．

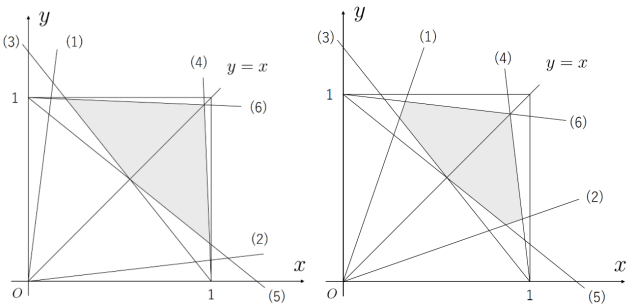

図 1 2 次元 ε-d 領域 $R_{\varepsilon,d}$ （cf. 補題 3.5）

各交点の座標を次のページに記す．(i)-(j) は式 $(i), (j)$ で表される直線の交点を表す．補題 3.5 より $y = x$ に関し対称なので，$y \leq x$ の範囲のみを記す．

補題 3.5. 2 次元 ε-d 領域 $R_{\varepsilon,d}$ は $y = x$ に関して対称であり，$e^{\varepsilon} + e^{-\varepsilon} + 2 > d$ のとき六角形であり，$e^{\varepsilon} + e^{-\varepsilon} + 2 \leq d$ のとき四角形の形をしている．

(2)-(3) $(\frac{d-1}{d}, \frac{d-1}{e^{\varepsilon}d})$
(2)-(4) $(\frac{e^{2\varepsilon}(d-1)}{1+e^{2\varepsilon}(d-1)}, \frac{e^{\varepsilon}(d-1)}{1+e^{2\varepsilon}(d-1)})$
(2)-(5) $(\frac{e^{\varepsilon}(d-1)}{e^{2\varepsilon}+d-1}, \frac{d-1}{e^{2\varepsilon}+d-1})$
(3)-(5) $(\frac{d-1}{e^{\varepsilon}+d-1}, \frac{d-1}{e^{\varepsilon}+d-1})$
(4)-(6) $(\frac{d-1}{e^{-\varepsilon}+d-1}, \frac{d-1}{e^{-\varepsilon}+d-1})$

証明. 6 つの方程式は 3 個ずつ，x, y に関して対称な形であるので，$R_{\varepsilon,d}$ は $y = x$ に関して対称である．

また，図 1 より，2 点 (2)-(4)，(2)-(5) の x 座標の位置関係によって，$R_{\varepsilon,d}$ の形が変わることがわかる．すなわち，

$$\frac{e^{2\varepsilon}(d-1)}{1+e^{2\varepsilon}(d-1)} \leq \frac{e^{\varepsilon}(d-1)}{e^{2\varepsilon}+d-1}$$

のときに四角形になり，そうでないときに六角形になることがわかる．この式を計算すると条件式を得る．

□

定義 3.6 ($(d$ 次元$)\varepsilon$-d 領域$)$. $D_{\varepsilon,d} := \{(x_1, \ldots, x_d) \in [0,1]^d \mid$ 任意の相異なる $i, j \in [d]$ に対して $(x_i, x_j) \in R_{\varepsilon,d}\}$ と定め，これを（d 次元）ε-d **領域**と呼ぶ．

補題 3.7. ε-d 領域 $D_{\varepsilon,d}$ は凸集合[*2]である．

証明. $D_{\varepsilon,d}$ は $\mathbb{R}^d$ 内の有限個の 1 次式と $[0,1]^d$ で区切られた図形である．1 次式で区切られた $\mathbb{R}^d$ の部分集合と $[0,1]^d$ はそれぞれ凸集合である．2 つの凸集合の共通部分は凸集合であるので $D_{\varepsilon,d}$ も凸集合である（cf. 補題 A.2.3）． □

補題 3.8. P を d 次で A 型の遷移行列とする．このとき，$(1-p_1, \ldots, 1-p_d) \in D_{\varepsilon,d}$ が成り立てば，P に従う維持置換攪乱は ε-差分プライバシーを満たす．

証明. 補題 3.2 の条件は，特に，P が A 型のときは以下のように書き換えられる．

任意の相異なる $i, j \in [d]$ に対して，

$$e^{-\varepsilon} \leq \frac{1-p_i}{1-p_j} \leq e^{\varepsilon}, \quad e^{-\varepsilon} \leq \frac{1-p_i}{p_j(d-1)} \leq e^{\varepsilon}$$

が成り立つ．

これを整理すると，上記の (1)〜(6) の不等式になる． □

各 $i \in [d]$ に対して，$q_i := 1 - p_i$ と置く．

$$\begin{aligned} Pv - v &= (P - I_d)v \qquad (I_d\text{は単位行列}) \\ &= \begin{pmatrix} -q_1 & \frac{q_2}{d-1} & \cdots & \frac{q_d}{d-1} \\ \frac{q_1}{d-1} & -q_2 & \cdots & \frac{q_d}{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{q_1}{d-1} & \frac{q_2}{d-1} & \cdots & -q_d \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} \\ &= \begin{pmatrix} -v_1 & \frac{v_2}{d-1} & \cdots & \frac{v_d}{d-1} \\ \frac{v_1}{d-1} & -v_2 & \cdots & \frac{v_d}{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{v_1}{d-1} & \frac{v_2}{d-1} & \cdots & -v_d \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_d \end{pmatrix} \\ &= \frac{-d}{d-1} \begin{pmatrix} v_1 q_1 - \bar{v} \\ v_2 q_2 - \bar{v} \\ \vdots \\ v_d q_d - \bar{v} \end{pmatrix} \end{aligned}$$

ただし，ここで $\bar{v} := \frac{N}{d}$ であり v の平均である．最小化問題において，定数倍の影響は考えなくてよいので，$||Pv - v||$ の最小化は，次の目的関数 f_v の最小化と同値である．

定義 3.9 (目的関数，大域的最適解)**.** $v \in \mathbb{Z}^d$ に対して，写像 $f_v : [0,1]^d \to \mathbb{R}$ で

$$f_v(q_1, \ldots, q_d) := ||(v_1 q_1 - \bar{v}, \ldots, v_d q_d - \bar{v})||$$

となるものを本稿の**目的関数**と呼ぶことにする．一般に，線形写像と $\ell 2$ ノルムの合成は凸関数であるので，f_v も凸関数である．また，$f_v(q_1, \ldots, q_d) = 0$ となる点 $(q_1, \ldots, q_d) \in [0,1]^d$ のことを f_v の**大域的最適解**と呼ぶ．

補題 3.8 より，A 型遷移行列 P が ε-差分プライバシーを満たすことは，$(q_1, \ldots, q_d) \in D_{\varepsilon,d}$ と言い換えることができる．f_v の定義域を $D_{\varepsilon,d}$ と取り換えたものを $f_v|_{D_{\varepsilon,d}}$ とおく．これより，問題 3.1 は次のように言い換えられる．

> **問題 3.10.** $f_v|_{D_{\varepsilon,d}}$ の最小値を与える $(q_1, \ldots, q_d) \in D_{\varepsilon,d}$ を求めよ．

3.3 大域的最適解に関する性質

f_v の大域的最適解が存在するとき，解に自由度があることが多い．そのため，以下の方針を定める．

方針 3.11. 大域的最適解 $(q_1, \ldots, q_d)$ が存在して，自由度があるときは，各成分が最小になるものを選ぶ．q_i を小さくすることで，各維持確率 p_i を大きくすることができる．

$v_i = 0$ となる $i \in [d]$ が存在する場合は，さらに詳細な考察が必要であるため，5 節にあずける．本稿の議論としては，そのような i が存在する場合は，次項 3.4 における方法を適用するものとする．

以降，本項では，任意の $i \in [d]$ に対して $v_i \neq 0$ であり，ヒストグラム v は降順にソートされているものとする（つまり $v_1 \geq \cdots \geq v_d$ が成り立つ）．次の補題は，f_v の大域的最適解を明示的に与えるものである．

補題 3.12. $(q_1, \ldots, q_d) \in [0,1]^d$ に対して次の (i)，(ii) は同値である．

(i) $f_v(q_1, \ldots, q_d) = 0$

(ii) $(q_1, \ldots, q_d) = t \cdot (\frac{v_d}{v_1}, \ldots, \frac{v_d}{v_{d-1}}, 1)$ なる $t \in [0,1]$ が存在．

[*2] 凸集合や凸関数に関する定義や性質は付録 A.2 にまとめて記す．

証明. (ii) ⇒ (i) は計算で確認できる．(i) ⇒ (ii) を示す．条件 (i) より，任意の $i \in [d]$ に対して，$v_i q_i = \bar{v}$ が成り立つ．ゆえに，$s \in \mathbb{R}$ が存在して $v_i q_i = s$ が成り立つ．また，任意の $i \in [d]$ について，$0 \le q_i \le 1$ であるので，$s \le v_i$ である．ここで，$v_1 \ge \ldots \ge v_d$ と仮定しているので，$0 \le s \le v_d$ である．$t = \frac{s}{v_d}$ とおけば (ii) の主張である． □

次はこの解が ε-d 領域 $D_{\varepsilon,d}$ 内にあるかの判定方法である．

補題 3.13. $t \in [0,1]$ に対し，次の 2 点は同値である．

(i) $t \cdot (\frac{v_d}{v_1}, \ldots, \frac{v_d}{v_{d-1}}, 1) \in D_{\varepsilon,d}$

(ii) $\frac{v_1}{v_d} \le e^{\epsilon}$, $\frac{e^{-\epsilon}(d-1)}{\frac{v_d}{v_2} + e^{-\epsilon}(d-1)\frac{v_d}{v_1}} \le t \le \frac{e^{\epsilon}(d-1)}{e^{\epsilon}(d-1) + \frac{v_d}{v_{d-1}}}$

証明. 任意の相異なる $i, j \in [d]$ に対して，$(t \cdot \frac{v_d}{v_i}, t \cdot \frac{v_d}{v_j})$ が 2 次元 ε-d 領域 $R_{\varepsilon,d}$ に入る条件を考える．補題 3.5 より，$R_{\varepsilon,d}$ は $y = x$ に関して対称なので $i < j$ と仮定してよい．$R_{\varepsilon,d}$ の形は i, j の値にかかわらず不変なので，全ての $(t \cdot \frac{v_d}{v_i}, t \cdot \frac{v_d}{v_j})$ を一つの平面において，$R_{\varepsilon,d}$ に入っているかを判定すればよい．全ての点を図示したものが以下の図 2 である．

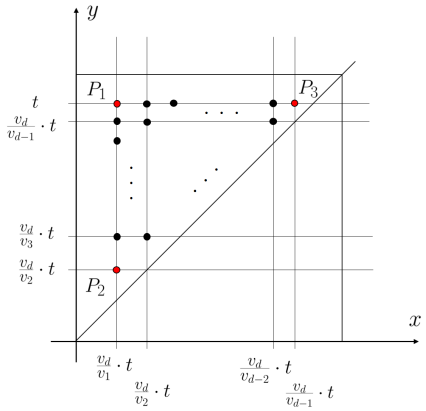

図 2 同一平面に図示した点 $(\frac{v_d}{v_i} \cdot t, \frac{v_d}{v_j} \cdot t)$ の位置関係

この図と $R_{\varepsilon,d}$ の形より，「全ての点が領域に入っているか」は，「もっとも外側の 3 頂点 $P_1(\frac{v_d}{v_1} \cdot t, t), P_2(\frac{v_d}{v_1} \cdot t, \frac{v_d}{v_2} \cdot t), P_3(\frac{v_d}{v_{d-1}} \cdot t, t)$ が入っているか」と同値であることがわかる．いま，$\frac{v_1}{v_d} \le e^{\varepsilon}$ は $P_1 \in R_{\varepsilon,d}$ と同値であることがわかる．また，点の位置関係から，$P_2, P_3 \in R_{\varepsilon,d}$ は「P_2 が直線 (3) より上側にあり，P_3 が直線 (6) より下側にある」と同値であることがわかる．ゆえに，

$$\begin{cases} \frac{v_d}{v_2} \cdot t \ge e^{-\varepsilon}(d-1)(1 - \frac{v_d}{v_1} \cdot t) \\ t \le 1 - \frac{e^{-\varepsilon}}{d-1} \cdot \frac{v_d}{v_{d-1}} \cdot t \end{cases}$$

という条件を考えればよい．これを t について整理すると主張の不等式になる． □

この補題と方針 3.11 より，大域的最適解が存在するときは $t = \frac{e^{-\epsilon}(d-1)}{\frac{v_d}{v_2} + e^{-\epsilon}(d-1)\frac{v_d}{v_1}}$ として取ればいいことがわかる．

3.4 大域的最適解が ε-d 領域に入っていない場合

大域的最適解が ε-d 領域 $D_{\varepsilon,d}$ に入っていない場合を考える．f_v の凸性より，目的関数を最小化する点 $(q_1, \ldots, q_d)$ は $D_{\varepsilon,d}$ の境界上にある（cf. 補題 A.2.4）．一般に，線形不等式で定められた領域内のある種の二次関数には最小値が存在し，それを求めるアルゴリズムも知られている．

定理 3.14 (二次計画法 [6])**.** A を実数値を成分にもつ $m \times d$ 行列とし，$b \in \mathbb{R}^m$ とする．$\mathcal{C} := \{x \in \mathbb{R}^d \mid Ax \le b\}$ とおく．対称行列 $Q \in M_d(\mathbb{R})$ とベクトル $c \in \mathbb{R}^d$ に対して，

$$f : \mathcal{C} \ni x \ \mapsto \ {}^t xQx + {}^t cx \in \mathbb{R}$$

という写像を考える．このとき，f を最小化するような $x^* \in \mathcal{C}$ が存在し，それを求めるアルゴリズムも存在する．

この定理を用いて目的関数 $F := f_v|_{D_{\varepsilon,d}}$ の最小値を与える点を求める．F は非負関数であるため，F を最小化する点と $F^2 := F \cdot F$ を最小化する点は等しい．ゆえに，F^2 に二次計画法を適用することで f_v の最適解も求まる．

対称行列 $Q = (q_{ij}) \in M_d(\mathbb{R})$ を

$$q_{ij} = \begin{cases} \frac{d-1}{d} v_i^2 & \text{if } i = j \\ -\frac{1}{d} v_i v_j & \text{else} \end{cases}$$

と定めると，$F^2(q_1, \ldots, q_d) = {}^t qQq + {}^t 0q$ と表すことができる（ここで 0 はゼロベクトル）．$D_{\varepsilon,d}$ は各組合せ q_i, q_j に対して 6 個ずつある線形不等式によって区切られた凸領域である．合計 ${}_dC_2 \times 6 = 3d(d-1) =: m$ 個の一次不等式は $m \times d$ 行列 A と d 次元ベクトル b を用いて $Aq \le b$ と表現できる．ゆえに，$D_{\varepsilon,d} := \{q \in \mathbb{R}^d | Aq \le b\}$ と記述できる．

以上の表記より，$D_{\varepsilon,d}$ 上で定義された F^2 に定理 3.14 を適用することができる．このようにして得た F に最小値を与える点 $q^* \in D_{\varepsilon,d}$ を CVXOPT(v, ε) と表記する．

$v_i = 0$ となるような $i \in [d]$ が存在するときも，この方法を適用することができる．

注意 3.15. 二次計画法による最適解の導出は，f_v が大域的最適解を $D_{\varepsilon,d}$ 内に持つ場合も適用できるが，本稿ではできる限り明示的な解を得ることを目指すため，前項 3.3 で大域的最適解が存在する場合の考察を行った．

3.5 導出アルゴリズム

3.3, 3.4 項での議論をまとめると，問題 3.10（および 3.1）を解くアルゴリズムを得る．

定理 3.16. τ をテーブル，$\varepsilon \in \mathbb{R}_{>0}$ をプライバシー保護指標とする（定理 2.13 より $k \in \mathbb{Z}_{>0}$ を決めることと等価）．このとき，Algorithm 1 によって問題 3.1 に対する解が得られる．

Algorithm 1 最適な A 型遷移行列 P の導出

Input: τ, ε：テーブル，プライバシー保護指標
Output: P：目的関数を最小化するような A 型遷移行列

v：τ のヒストグラム
v を降順に並び替える
if $v_1 > e^{\varepsilon} \cdot v_d$ **then** ▷ 補題 3.13 （$v_d = 0$ でも成り立つ）
    $(q_1, \ldots, q_d) \leftarrow \mathrm{CVXOPT}(v, \varepsilon)$
else if $e^{\varepsilon}(d-1)(\frac{1}{v_d} - \frac{1}{v_1}) + \frac{1}{v_{d-1}} > \frac{e^{2\varepsilon}}{v_2}$ **then** ▷ 補題 3.13
    $(q_1, \ldots, q_d) \leftarrow \mathrm{CVXOPT}(v, \varepsilon)$
else
    $t \leftarrow \dfrac{e^{-\epsilon}(d-1)}{\frac{v_d}{v_2} + e^{-\epsilon}(d-1)\frac{v_d}{v_1}}$ ▷ 補題 3.13
    $(q_1, \ldots, q_d) \leftarrow (\frac{v_d}{v_1} \cdot t, \frac{v_d}{v_2} \cdot t, \ldots, t)$ ▷ 補題 3.12
end if
$q_1, \ldots, q_d$ を元の v に対応するよう並び替える
$P \leftarrow (1 - q_1, \ldots, 1 - q_d)$：遷移行列を作る

4. 従来手法との比較

4.1 データと比較手法

従来型の維持置換攪乱はテーブル $\tau : [N] \to [d]$ とプライバシー保護指標 $k \in \mathbb{Z}_{>0}$（もしくは $\varepsilon \in \mathbb{R}_{>0}$）により決定される．具体的には，

$$k - 1 = (N-1) \cdot \left(\frac{1-p}{(d-1)p} \right)^2$$

という関係式が成り立つように維持確率 p を選ぶ．

比較に用いたヒストグラムは次の 3 種類である．横軸は縦軸ともに Case 3 の実データに合うよう調節した．

Case 1: **正規分布に従う場合** （図 3）
$N = 46938, d = 74, \max(v) = 850$
$(\varepsilon_2, \varepsilon_{10}, \varepsilon_{100}) = (5.38, 4.28, 3.08)$
· case 3 と違いを出すため，少し底上げしてある

Case 2: **冪乗則に従う場合** （図 4）
$N = 4405, d = 74, \max(v) = 895$
$(\varepsilon_2, \varepsilon_{10}, \varepsilon_{100}) = (4.20, 3.10, 1.90)$
· 様々な自然現象や社会現象で観測される

Case 3: **年齢データ（実データ）** （図 5）
$N = 32561, d = 74, \max(v) = 898$
$(\varepsilon_2, \varepsilon_{10}, \varepsilon_{100}) = (5.20, 4.10, 2.90)$
· Adult data set [7] の年齢データ（17〜90 歳）

これらに対し，従来型と今回検討した維持置換攪乱の結果を比較する．各ヒストグラム v に対して，保護処理を 100 回行い得たヒストグラムを $v^1, \ldots, v^{100}$ とおく．これらの平均として得るヒストグラムを v' とおき（$v' = \frac{1}{100}\sum_{j=1}^{100} v^j$），次の四つの値を比較した．

① $\mathrm{ave}_i := ||v' - v||$ ② $\mathrm{opt}_i := ||Pv - v||$（$P$ は遷移行列）
③ $\mathrm{min}_i := \min ||v^j - v||$ ④ $\mathrm{max}_i := \max ||v^j - v||$

比較結果が表 1 である．また，Case 1 の例を用いて，各種の保護処理を 1 回行った結果が図 6 である．さらに，Case 3 における $k = 2$ のときの保護処理の 100 回の平均ヒストグラム v' が図 7 である．スケールは全て図 3 や図 5 と同じものとなっている（高さはそろえてある）．

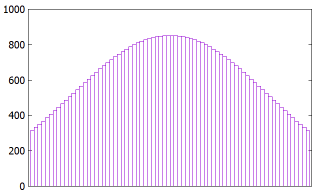

図 3 正規分布に従う場合

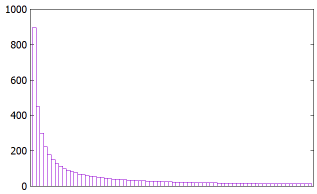

図 4 冪乗則に従う場合

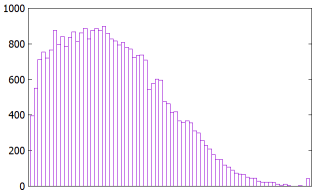

図 5 年齢データ

値	$k=2$		$k=10$		$k=100$	
ave_1	380.9	15.5	764.1	196.9	1168.5	841.0
opt_1	382.7	0.0	763.4	200.5	1169.7	845.1
min_1	371.7	117.9	730.2	237.5	1117.8	803.3
max_1	439.4	195.6	857.6	331.1	1242.3	930.2
ave_2	544.7	493.9	797.6	785.7	952.1	946.6
opt_2	543.5	493.4	796.2	786.7	950.6	947.2
min_2	505.0	465.5	763.5	757.7	930.8	919.3
max_2	574.7	522.2	834.6	818.5	975.9	975.3
ave_3	840.9	736.5	1604.0	1508.7	2338.5	2289.6
opt_3	841.7	736.4	1602.2	1510.2	2340.7	2290.9
min_3	809.5	714.4	1570.9	1470.5	2286.7	2226.8
max_3	890.4	799.7	1671.8	1570.6	2399.1	2337.5

表 1 従来手法と 提案手法 の比較結果

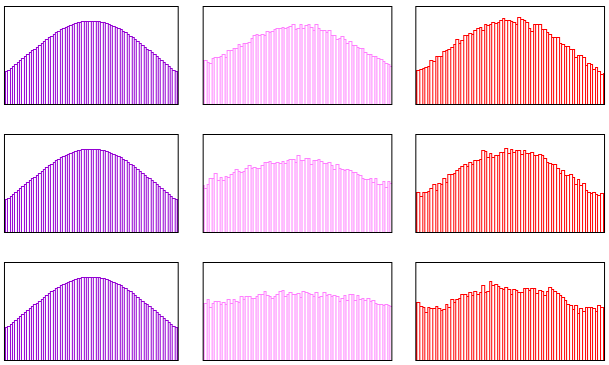

図 6 Case 1 に 1 回保護処理をかけたヒストグラム
左から順に「元データ」「従来手法」「提案手法」
上から順に「$k = 2$」「$k = 10$」「$k = 100$」

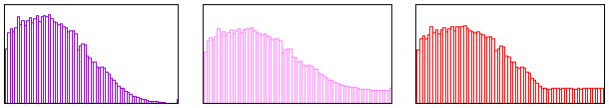

図 7 Case 3，$k = 2$ における平均的ヒストグラム v'
左から順に「元データ」「従来手法」「提案手法」

4.2 結果と考察

各種検証の結果から確認できる傾向を以下にまとめる．

- 表 1 より，提案手法が従来手法と比べて各値で最適化できていることがわかる．特に，Case 1 で $k = 2$ の時は，3.3 項で触れたような大域的最適解が $D_{\varepsilon,d}$ に入っていることが見て取れる．
- 図 6 や図 7 より，提案手法の方が従来手法よりも高いところを高く保つことが見て取れる．
- 図 6 より，プライバシー保護指標 k を厳しくする（大きくする）と攪乱度合いが強くなることが確認できる．
- 図 7 より，提案手法では一定の高さ以下の属性に対して，一律の高さになるような勾配の維持確率が採用されていることがわかる．

以上の傾向を踏まえると次のような主張ができる．

> 提案手法により得た維持置換攪乱は，
> ① 誤差の $\ell 2$ ノルムを最小化するという意味で最適なふるまいをしている．
> ② ヒストグラムの低い部分を一様に隠し，高い部分を従来手法のものよりも正確に出力する．

5. おわりに

本稿では，一つヒストグラムを決めたときに最適な維持置換攪乱を導出する方法を提案し，維持置換攪乱の有用性向上を図った．その結果，従来のものよりも有用性の高い維持置換攪乱を得ることができた．これは例えば，位置情報開示における人口分布など，ある程度事前にヒストグラムの外形がわかっているものに適用できると考えられる．

しかし，一般の場合に対する実際の運用には課題が残されている．今後の検討事項を以下に記す．

5.1 ヒストグラム $\boldsymbol{v}$ に $\boldsymbol{v_i = 0}$ がある場合の考察

f_v の定義より次の補題が成り立つことがわかる．

補題 5.1. $v_i = 0$ とする．このとき，任意の $a, b \in [0, 1]$ に対して，$f_v(q_1, \ldots, q_{i-1}, a, q_{i+1}, \ldots, q_d) = f_v(q_1, \ldots, q_{i-1}, b, q_{i+1}, \ldots, q_d)$ が成り立つ．

これは「$v_i = 0$ ならば q_i の値は自由にとって良い」ということを意味している．これより，$v_i = 0$ になるものに関しては一度，取り除いたものの中で最適解を探し，適当な値を q_i に与えればよいと思われるが，q_i の取りうる値がなくなってしまう可能性がある．そのため，q_i の取りうる値が存在する条件を求める必要がある．

5.2 系としてのプライバシー保護レベルの明示

今回は「入力されたヒストグラム v に対して，最適な有用性を保証する保護メカニズムを出力する方法」を提案した．しかし，実際の運用ではそのメカニズムに v を入力し，出力された v' を開示することになる．ゆえに，v を入力して v' を出力する一連のメカニズムが満たすプライバシー保護指標の評価が必要になる．

5.2.1 入力 $\boldsymbol{v} \in \mathbb{Z}^{\boldsymbol{d}}$ に対する明示的な最適解 $\boldsymbol{P}$ の導出

補題 3.12 である種の条件において P を v によって明示的に表現することができた．本稿で二次計画法を用いた部分に関しても明示的な解が与えられることが期待される．

5.2.2 $\boldsymbol{P}$ の開示を前提とした $\boldsymbol{P}$ に対するランダム化

運用の際，P を公開する可能性も考えることができる．その際，本稿で提案した方法自体に対するランダム化も考える必要がある．方法としては，出力した P にノイズを加える出力摂動法や f_v にノイズを加える目的関数摂動法（cf. [8]）が考えられる．

5.2.3 合成定理などを用いたプライバシー保護指標の計算

P を公開しない場合は，明示的な表現をもとに，定義通り v から v' を出力する系の差分プライバシーを求めればよい．P を公開する場合は，P に対するランダム化と，P によるテーブル保護が満たす差分プライバシーの指標をそれぞれ計算し，合成定理によって足し合わせればよい．

5.3 遷移行列の自由度を上げる

今回は A 型と名付けた遷移行列を考えた．従来型と比べると自由度 1 から d へと増える拡張になっているが，本来，遷移行列 P には $d \times (d-1)$ の自由度がある．その場合の維持置換攪乱を考察することがこのメカニズムの理論的な限界を求めることになる．

謝辞 本稿の執筆にあたり，同研究所の須藤弘貴さんをはじめ多くの同僚の方から，内容や実験のアドバイスをいただきました．この場をかりて御礼申し上げます．

参考文献

[1] Dai Ikarashi, Ryo Kikuchi, Koji Chida, and Katsumi Takahashi. k-anonymous microdata release via post randomisation method. In *International Workshop on Security*, pp. 225–241. Springer, 2015.

[2] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pp. 265–284. Springer, 2006.

[3] Rakesh Agrawal, Ramakrishnan Srikant, and Dilys Thomas. Privacy preserving olap. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pp. 251–262. ACM, 2005.

[4] 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.

[5] Yue Wang, Xintao Wu, and Donghui Hu. Using randomized response for differential privacy preserving data collection. In *EDBT/ICDT Workshops*, Vol. 1558, 2016.

[6] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[7] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.
[8] Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, Vol. 12, No. Mar, pp. 1069–1109, 2011.

付　　録

A.1　ゲーム形式でとらえる Pk-匿名性

本節では，Pk-匿名性をゲーム形式で解釈する方法を紹介する．下記の補題を示す．これは定義 2.7 の際に，「任意の攻撃者」と書いていたところを「任意のテーブルの組み合わせ」へと書き換えられるということを意味している．

補題 A.1.1. $k \in \mathbb{Z}_{>0}$ とする．P を遷移行列とし，m を P に従う維持置換攪乱とする．このとき，m が Pk-匿名性を満たすことと次が成り立つことは同値である．

(★) U を一様分布に従う $\mathfrak{S}_N$-値確率変数とする．任意の $\tau, \tau' \in T$，任意の $i, j \in [N]$ に対して，

$$\Pr[U(i) = j \mid m(\tau) = \tau' \circ U] \leq \frac{1}{k}$$

が成り立つ．

証明. $Pk \Rightarrow$ (★) は任意にとった $\tau, \tau' \in T$ に対して，X_T を $\Pr[X_T = \tau] = 1$ となるようにとれば確かめられる．

$Pk \Leftarrow$ (★) を示そう．X_T を任意の確率変数とする．

$$\begin{aligned}
\mathcal{E} &:= \Pr[U(i) = j \mid m(X_T) = \tau' \circ U] \\
&:= \frac{\Pr[U(i) = j \wedge m(X_T) = \tau' \circ U]}{\Pr[m(X_T) = \tau' \circ U]} \\
&= \frac{\sum_{\tau \in T} \Pr[X_T = \tau]\Pr[U(i) = j \wedge m(\tau) = \tau' \circ U]}{\sum_{\tau \in T} \Pr[X_T = \tau]\Pr[m(\tau) = \tau' \circ U]}
\end{aligned}$$

ここで，$g(\tau) := \dfrac{\Pr[U(i) = j \wedge m(\tau) = \tau' \circ U]}{\Pr[m(\tau) = \tau' \circ U]}$ を最大化する $\tau \in T$ を τ^* とおくと

$$\Pr[U(i) = j \wedge m(\tau) = \tau' \circ U] \leq \Pr[m(\tau) = \tau' \circ U] \cdot g(\tau^*)$$

が成り立つ．ゆえに，

$$\begin{aligned}
\mathcal{E} &\leq \frac{\sum_{\tau \in T} \Pr[X_T = \tau]\Pr[m(\tau) = \tau' \circ U] \cdot g(\tau^*)}{\sum_{\tau \in T} \Pr[X_T = \tau]\Pr[m(\tau) = \tau' \circ U]} \\
&= g(\tau^*) \\
&:= \frac{\Pr[U(i) = j \wedge m(\tau^*) = \tau' \circ U]}{\Pr[m(\tau^*) = \tau' \circ U]} \\
&=: \Pr[U(i) = j \mid m(\tau^*) = \tau' \circ U] \leq \frac{1}{k}
\end{aligned}$$

最後の不等式は，(★) の条件より成り立つ．よって，任意の X_T に対して，不等式が成り立つことが言えたので，m は Pk-匿名性をもつ．　□

この補題によって，メカニズム m が Pk-匿名性を持つことを次のゲームを通して表現することができる．

Pk ゲーム

$\mathcal{A}$ をテーブルの何が何に書き換えられたか推定する攻撃者とし，$\mathcal{C}$ を攻撃に受けて立つ挑戦者とする．
Step 1： $\mathcal{A}$ は一つテーブル τ を用意し，$\mathcal{C}$ に渡す．
Step 2： $\mathcal{C}$ は保護メカニズム m を τ にかけ，加工済みテーブル τ' を一つ得る．
Step 3： また，$\mathcal{C}$ はランダムに選んだテーブル置換 $\sigma \in \mathfrak{S}$ を τ' にかけて $\tau' \circ \sigma$ を $\mathcal{A}$ へ渡す．
Step 4： $\mathcal{A}$ は τ, τ' を使い，$j = \sigma(i)$ となるであろう i, j の組を推測して選ぶ．
このとき，$j = \sigma(i)$ ならば $\mathcal{A}$ の勝ちと定める．

メカニズム m が Pk-匿名性を満たすとは，m が Pk ゲームにおいて，攻撃者 $\mathcal{A}$ が勝つ確率が $\Pr[\mathcal{A} \text{ wins}] \leq \frac{1}{k}$ を満たすことを表す．

A.2　凸領域，凸関数

本稿では，凸領域や凸関数の性質を用いた議論を行うので，凸性の定義と簡単な性質を紹介する．

定義 A.2.1 (凸集合)**.** ユークリッド空間の部分集合 $\mathcal{C} \subset \mathbb{R}^d$ が**凸集合**であるとは，任意の 2 点 $x, y \in \mathcal{C}$ と任意の $t \in [0, 1]$ に対して，$tx + (1-t)y \in \mathcal{C}$ が成り立つことをいう．

これはつまり，任意の 2 点に関して，その 2 点を結ぶ線分が領域にすっぽりと入っていることを表している．

定義 A.2.2 (凸関数)**.** $\mathcal{C} \subset \mathbb{R}^d$ を凸集合とする．このとき関数 $f : \mathcal{C} \to \mathbb{R}$ が**凸関数**であるとは，任意の $x, y \in \mathcal{C}$，任意の $t \in [0, 1]$ に対して，$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$ が成り立つことをいう．

また，これは $\{(x, f(x)) \in \mathbb{R}^d \times \mathbb{R}\} \subset \mathbb{R}^{d+1}$ が凸集合であることと同値である．

補題 A.2.3. 二つの凸集合の共通部分は凸集合である．

証明. $\mathcal{C}, \mathcal{D} \subset \mathbb{R}^d$ を凸集合，$x, y \in \mathcal{C} \cap \mathcal{D}$ とする．すると，任意の $t \in [0, 1]$ に対し，$\mathcal{C}, \mathcal{D}$ は凸集合より，$tx+(1-t)y \in \mathcal{C}$ かつ $tx+(1-t)y \in \mathcal{D}$ である．すなわち，$tx+(1-t)y \in \mathcal{C} \cap \mathcal{D}$ であるので，$\mathcal{C} \cap \mathcal{D}$ は凸集合である．　□

補題 A.2.4. $f : \mathbb{R}^d \to \mathbb{R}$ を非負連続凸関数とする．$\mathcal{C} \subset \mathbb{R}^d$ をコンパクトな凸集合とし（境界も集合に含まれ，有界なものとする），f の最小値を与える点が $\mathcal{C}$ 内には存在しないとする．このとき，$\mathcal{C}$ の境界上には少なくとも一つ，f の定義域を $\mathcal{C}$ に制限した写像の最小値を与える点が存在する．

証明. 紙面の都合上，証明は割愛する．　□