3I2-OS-20-8

スペクトル差分プライバシーに基づくプライバシー保護推薦アルゴリズム

Privacy-preserving Recommendation with Spectral Differential Privacy

佐久間 淳[*1][*2]
Jun Sakuma

[*1]筑波大学 システム情報系
University of Tsukuba, Faculty of Engineering, Information and Systems

[*2]科学技術振興機構 さきがけ
Japan Science and Technology Agency, PRESTO

Positive semidefinite matrices are important for a number of data mining applications. We consider the problem of differentially private publication of positive semidefinite matrices computed from private information. Differential privacy is typically achieved by adding random noise. However, when the outputs form positive semidefinite matrices, element-wise additive randomization causes problems. First, when not a single element, but the entire matrix is released, the scale of noises to provide differential privacy can be too large. Second, such randomization not only destroys the positive semidefiniteness, but may be statistically denoised in some cases. For these problems, we introduce a new randomization mechanism which separately randomizes eigenvectors and eigenvalues so that the randomization does not completely destroy the spectral features. Furthermore, noting that low-rank approximation preserves useful information of matrices while discarding unnecessarily details, we incorporate low-rank approximation into randomization. We prove that the scale of perturbation required to guarantee differential privacy is inversely proportional to the rank of the output matrices in the proposed randomization mechanism. Thus, if a data analyst does not need the output matrix itself, but needs only a low-rank approximation, the scale of perturbation can be relatively smaller without sacrificing privacy. This is convenient for data mining applications which work well even with lower-rank approximation. We experimentally demonstrate the proposed mechanism with collaborative filtering.

1. はじめに

開示できない個人情報を集めた**データベース**を用いた解析を考える. **解析者**はデータベースにクエリを発行し知識を得ることが許されている. 一方データベースが正直にクエリに応答した場合, 解析者に各個人の情報を推測されるおそれがある. このようなやり取りで引き起こされる情報の漏えいを防ぐために，データベースは応答が情報漏えいをなるべく引き起こさないように応答の内容を加工する必要がある. これと同時に解析者によっての応答の有用性はなるべく損なわれないような配慮が必要である. これを**出力プライバシー**の問題と呼ぶ.

近年, 出力プライバシーのための普遍的な安全性定義として**差分プライバシー** (differential privacy) が注目されている [3]. 直観的には, どの個人がデータベースに入っていようと (あるいはいなかろうと), 応答値が大きく変化しないクエリ応答アルゴリズム (以降, **メカニズム**と呼ぶ) は, 差分プライバシーを満足していると考える.

本稿では, データベースが保持する個人情報を利用して生成された半正定値カーネル行列を出力するメカニズムが差分プライバシーを満たすための条件について理論的に考察する. 半正定値カーネル行列は, データの共分散行列やある種の類似度行列, ネットワークの接続行列などデータ解析にとって重要な様々な行列の表現形式である. 本稿では一例として，類似度行列とこれを用いた推薦アルゴリズムを取り上げる．データベースは, 多数のユーザからの複数のアイテム評価値の集合を保持しているとしよう. ユーザ間あるいはアイテム間の類似度はこれらの評価値集合から共分散行列の形式で評価できる．解析者は推薦などの目的でこの共分散行列を取得したいが，類似度行列の開示は，”だれがどのアイテムを評価したか”などの個人情報の漏えいを引き起こす可能性がある [6]. 類似度行列をどのように加工し，公開すればよいだろうか?

差分プライバシーは主に出力に対する (**加法的摂動**) で実現される．摂動の分散の大きさは，差分プライバシーの定義に基づき, 出力が入力の変化に対し十分鈍感になるよう決定する. このような加法的摂動は, 出力がベクトル形式の場合は問題にならないが, 行列をなす場合には脱ノイズ化が容易な場合がある. 後に実験で示すように, 要素毎の加法的摂動は, 対象の行列の固有値分布を大きく変更し, その正定値性を破壊する. 摂動モデルは秘密にされないことから, 統計的手法により**脱ノイズ**が可能であり, 期待された安全性が保障されないことがある.

これらを考慮し, 半正定値行列の差分プライバシーを満足した開示を実現するための二つのテクニックを導入する. 一つは, 行列を固有値分解し, 行列に摂動を与える代わりに, 固有ベクトルと固有値に個別に摂動を与える. より具体的には, 半正定値行列の固有ベクトルは正規直交基底を張ることを利用し, 固有ベクトルには von Mises-Fishder 分布による摂動を加える. 固有値には Laplace 分布による摂動を加える. このような摂動は行列の固有値分布に大きな変更を与えないため脱ノイズを避けることができ, また行列の半正定値性を保存することができる. もう一つは, 摂動と行列の低ランク近似の統一的な取り扱いである. 低ランク近似はスペクトル上の詳細を排除し抽象化された行列を与える．この抽象化とプライバシー保護のための摂動を組み合わせることによって, 低ランク行列の開示に必要な摂動のサイズをフルランク行列に比べて抑えることができる. 低ランク近似は協調フィルタリングやグラフラプラシアンのスペクトルアルゴリズムなど, 様々なデータマイニングアルゴリズムに有用である. このため提案メカニズムの応用に当たっては, 要求される最小限のランクにおいて低ランク近似された行列を開示対象とすることで, 行列が持つ効用を大きく犠牲にせず差分プライバシーを実現することができる.

連絡先: 佐久間淳，筑波大学システム情報系，茨城県つくば市天王台 1-1-1，jun@cs.tsukuba.ac.jp

以下, 2 章では差分プライバシーを導入し, 3 章では低ランク正定値行列のための新しい摂動法を設計し, その摂動法が差分プライバシーを満たす条件を導出する. 4 章では MovieLens データにおける協調フィルタリングを例に, 提案した摂動法がプライバシー保護と行列の効用の維持を両立可能であることを示す.

2. 正定値行列の差分プライバシー

2.1 差分プライバシー

$D = \{d_1, d_2, \ldots, d_N\} \in \mathcal{D}^N$ をデータベースとする. ただし $d_i \in \mathcal{D}$ は i 番目の個人のデータである. $f : \mathcal{D}^N \mapsto O$ を, データベースを入力に取るクエリ関数とする. 解析者はデータベースにクエリを通じてのみアクセスすることを許されており, クエリの出力からデータベースに含まれる個人の情報を推測しようとする. このような攻撃による個人情報の漏えいを抑制するために, データベースはクエリ出力を開示前にランダム化する. これを (ランダム化) メカニズムと呼ぶ. 差分プライバシー [3] はこのような枠組みでの安全性定義の一種である. 直観的には, 対象とするデータベースに任意の個人が含まれていようがいなかろうが, メカニズムの出力の変化が大きくなければ, メカニズムは差分プライバシーを満足する, と考える.

二つのデータベース $D = \{d_1, \ldots, d_N\}$ および $D' = \{d'_1, \ldots d'_N\}$ について, もし両者が高々一要素のみ異なる場合, $D \sim D'$ と書く. このとき, 一般性を失うことなく $i = 1, \ldots N-1$ について $d_i = d'_i$ および $d_N \neq d'_N$ としてよい. このとき差分プライバシーは以下のように定義される:

Definition 1 *(ϵ-差分プライバシー)* クエリを $f : \mathcal{D}^N \mapsto O$ とする. 全ての $D \sim D'$ なる (D, D') のペアと任意の値域の部分集合 $S \subseteq O$ について, メカニズム $\mathcal{M}_f : O \mapsto O$ は

$$\Pr[\mathcal{M}_f(f(D)) \in S] \leq \exp(\epsilon)\Pr[\mathcal{M}_f(f(D')) \in S]. \quad (1)$$

を満たすならば $\mathcal{M}_f$ は ϵ-差分プライバシーを満たす,

このように, 差分プライバシーは, 任意の $D \sim D'$ について, 出力値の確率密度比が, プライバシーパラメタ ϵ に基づいてバウンドされることを要求する. これを満たす単純な方法は, 出力の加法的摂動である. $\epsilon-$ 差分プライバシーの達成に必要な摂動のスケールはクエリ f の**敏感度**によって定まる. 出力の値域を $\mathbb{R}^M$ としたとき, ℓ_p-**敏感度**は以下のように定義される.

Definition 2 *(ℓ_p-sensitivity) The ℓ_p-sensitivity of function f is defined as follows:* $\Delta_{p,f} = \max_{D\sim D'} \|f(D) - f(D')\|_p$.

ただし $\|\cdot\|_p$ は ℓ_p ノルムである. $\mathrm{Lap}(x; \mu, b)$ をラプラス分布する. ここで μ は平均, b はスケールパラメタである. クエリ関数 f の ℓ_1 敏感度に応じて, **ラプラスメカニズム**は ϵ-差分プライバシーを与える [3].

Theorem 1 *(Laplace mechanism) Let $f : \mathcal{D}^N \mapsto \mathbb{R}^M$ be a query function. Given $D \in \mathcal{D}$, if $b \geq \Delta_{1,f}/\epsilon$, $\mathcal{M}(f(D)) = f(D) + \boldsymbol{r}$ guarantees ϵ-differential privacy, where $\boldsymbol{r}^T = (r_1, r_2, \ldots, r_M)$ and $r_i \sim \mathrm{Lap}(x; 0, b)$ for all i.*

2.2 共分散行列の差分プライバシー

続いて具体的な正定値行列の例として共分散行列の差分プライバシーとそのための摂動について考察する. 簡単のために, 全てのデータ $\boldsymbol{d}_i^T = (d_{i1}, \ldots, d_{iM})$ は中央化 $\sum_j d_{ij} = 0$ されているものとする. このとき共分散の (i, j) 要素は, $\sigma_{ij} = \frac{1}{N}\boldsymbol{d}_i^T \boldsymbol{d}_j$ で与えられる. このとき ℓ_1-敏感度は

$$\Delta_{\mathrm{cov},1} = \max_{D\sim D'} \|\sigma_{ij} - \sigma'_{ij}\|_1 = \max_{D\sim D'} \|\frac{d_{jp}}{N}(d_{ip} - d'_{ip})\|_1 \leq \frac{4\gamma^2}{N}. \quad (2)$$

である. これに基づくメカニズムは

$$\mathcal{M}_{1,\mathrm{cov}}(f(D)) = \frac{1}{N}\boldsymbol{d}_i^T \boldsymbol{d}_j + r_{ij} \quad (3)$$

となる. ここで $r_{ij} \sim \mathrm{Lap}(x; 0, \frac{4\gamma^2}{\epsilon N})$ である. 上記は単一の要素を出力する場合のメカニズムであるが, 行列全体を返す場合には, $r_{ij} \sim \mathrm{Lap}(x; 0, 4\gamma^2/\epsilon)$ とする必要がある.

このメカニズムに基づく出力は確かに差分プライバシーを満足するが, もし $\boldsymbol{X}$ が正定値性や低ランクなど特徴的なスペクトル分布を持つ場合には, $\tilde{\boldsymbol{X}}$ の低ランク近似は原行列 $\boldsymbol{X}$ をある程度復元させる [1]. 図 1 はこれを経験的に評価した結果である. 共分散行列を 1,000 個ランダムに生成しこれを開示対象の原行列としたとき, 図 1 は, それぞれの 1.0-差分プライバシーを保持する共分散行列と, その rank-5 近似のそれぞれについて, 原行列との要素毎 ℓ_1 ノルム $\|\cdot\|_{\mathrm{ew},1}$ のヒストグラムを示している. この尺度において分散後の各要素は原行列からは, おもに 0.07 ほど離れている. しかし低ランク近似による脱ノイズ後はこれが 0.03 に減っている. このことは (少なくともこのケースでは)1.0-差分プライバシーの達成には $\epsilon \leq 0.3$ を必要とすることを示唆している. これは経験的な評価ではあるが, 要素毎のランダム化による差分プライバシーは脱ノイズによって弱められる可能性があることが示された.

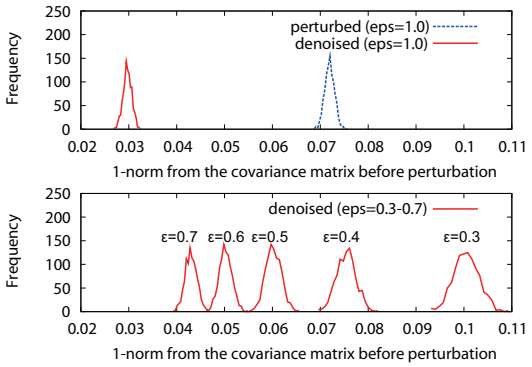

図 1: Top: Histogram of $\|\tilde{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{ew},1}$ (perturbed, $\epsilon = 1.0$) and $\|\tilde{\boldsymbol{\Sigma}}_5 - \boldsymbol{\Sigma}\|_{\mathrm{ew},1}$ (denoised with rank-5 approximation, $\epsilon = 1.0$). Histogram of $\|\tilde{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{ew},1}$ (denoised with rank-5 approximation, $\epsilon = 0.3-0.7$)

3. スペクトル差分プライバシー

3.1 接近法

序章で議論したように, 入力となる行列の固有値分布を無視して要素毎にランダム化しても脱ノイズのリスクがあることはすでに述べた. これを考慮し本章では入力行列を固有値分解し, 固有ベクトルと固有値をそれぞれの性質を考慮したランダム化手法を導入する.

$\mathbb{S}_+^N$ と $\mathbb{S}_+^{N,q}$ を N 次元の正定値行列および rank-q 正定値行列の集合とする. また $\mathrm{St}^{N,q}$ を $N \times q$ 正規直交行列の集合 (Stiefel 多様体) とする. $\boldsymbol{X} \in \mathbb{S}_+^N$ の固有値分解を $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$ とする. ここで $\boldsymbol{U} \in \mathrm{St}^{N,N}$ の列ベクトルは固有ベクトル, $\boldsymbol{\Lambda}$ は対角要素

truncated Laplace mechanism
- Input: $\lambda_i \in \mathbb{R}_+$, privacy parameter ϵ
- Output: $\tilde{\lambda}_i \in \mathbb{R}_+$

1. Evaluate $b = \frac{\epsilon}{\|\Delta \boldsymbol{X}\|_2}$
2. Let $\tilde{\lambda}_i = \lambda_i + r$, where $r \sim \mathrm{Lap}(x; 0, b)$
3. If $\tilde{\lambda}_i < 0$, outputs 0. Otherwise, outputs $\tilde{\lambda}_i$.

図 2: Randomization mechanism for eigenvalues.

を固有値に持つ正対角行列である. 同様に $\boldsymbol{X}$ の rank-q 近似の固有値分解を $\boldsymbol{X}_q = \boldsymbol{U}_q \boldsymbol{\Lambda}_q \boldsymbol{U}_q^T$ とする. 最終的には低ランク近似の差分プライバシーの達成を目指しているため, クエリ関数は $f : \mathcal{D}^N \mapsto \mathbb{S}_+^{N,q}$, ランダム化メカニズムは $\mathcal{M} : \mathbb{S}_+^{N,q} \mapsto \mathbb{S}_+^{N,q}$ で与えるが, 以下では, 固有値分布と正定値性を考慮し, 固有値と固有ベクトルをそれぞれ別個に摂動させ, これらを独立に開示するメカニズムについて議論する.

3.2 固有値の差分プライバシー

$\mathcal{M}_{\mathrm{eval}} : \mathbb{R}_+ \mapsto \mathbb{R}_+$ を固有値のランダム化メカニズムとする. このとき, ϵ_{λ_i} において, 任意の $S_{\mathrm{eval}} \subseteq \mathbb{R}_+$ および $D \sim D'$ なる任意の (D, D') について以下を満たすとき, 定義より $\mathcal{M}_{\mathrm{eval}} : \mathbb{R}_+ \mapsto \mathbb{R}_+$ は半正定値行列の固有値の差分プライバシーを満たす.

$$\Pr\left[\mathcal{M}_{\mathrm{eval}}(\lambda_i) \in S_{\mathrm{eval}}\right] \leq e^{\epsilon_{\lambda_i}} \Pr\left[\mathcal{M}_{\mathrm{eval}}(\lambda_i') \in S_{\mathrm{eval}}\right]. \quad (4)$$

ここで λ_i と λ_i' はそれぞれ $\boldsymbol{X} = f(D)$ と $\boldsymbol{X}' = f(D')$ の i 番目に大きい固有値である. ランダム化メカニズムの設計には, 摂動に対する固有値の敏感度の評価が必要である. 加法摂動行列による摂動が加えられた固有値の敏感度は, その摂動行列のスペクトルノルムでバウンドされる [8].

Theorem 2 *If $\boldsymbol{X}$ and $\boldsymbol{X} + \Delta\boldsymbol{X}$ are $N \times N$ symmetric matrices, then, for $k = 1, ..., N$, $|\lambda_k - \tilde{\lambda}_k| \leq \|\Delta \boldsymbol{X}\|_2$.*

ここで $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \ldots \geq \tilde{\lambda}_N$ は $\tilde{\boldsymbol{X}}$ の固有値である. これに基づけば, 固有値に対するランダム化メカニズムについて, 以下の定理を得る (証明略).

Theorem 3 *Let Λ be a diagonal matrix in which the diagonal elements are eigenvalues of $\boldsymbol{X}$. If $b \geq \frac{\epsilon_{\lambda_i}}{\|\Delta \boldsymbol{X}\|_2}$, then $\mathcal{M}_{eval}$ yields ϵ_{λ_i}-differential privacy.*

直観的には, ラプラス分布のスケールパラメータを $\frac{1}{\|\Delta \boldsymbol{X}\|_2}$ でバウンドすることで差分プライバシーが達成される. ただしラプラス分布による加法的摂動は, 負数を出力する可能性がある. 半正定値行列の固有値は 0 以上であるため, 負数が出た場合はこれを 0 に強制する. この操作を適用しても差分プライバシーは保障される [5]. これらをまとめ、図 2 のアルゴリズムを得る.

3.3 固有ベクトルの差分プライバシー

$\mathcal{M}_{\mathrm{evec}} : \mathrm{St}^{N,q} \mapsto \mathrm{St}^{N,q}$ を q 本の固有ベクトルのランダム化メカニズムとする. 所与のパラメタ ϵ_U において, 任意の $S_{\mathrm{evec}} \subseteq \mathrm{St}^{N,q}$ および (D, D') なる任意の $D \sim D'$ が以下を満たすとき $\mathcal{M}_{\mathrm{evec}}$ は固有ベクトル群の差分プライバシーを満たす.

$$\Pr\left[\mathcal{M}_{\mathrm{evec}}(\boldsymbol{U}) \in S_{\mathrm{evec}}\right] \leq e^{\epsilon_U} \Pr\left[\mathcal{M}_{\mathrm{evec}}(\boldsymbol{U}') \in S_{\mathrm{evec}}\right] \quad (5)$$

ここで $\boldsymbol{U}$ と $\boldsymbol{U}'$ はそれぞれ $\boldsymbol{X} = f(D)$ と $\boldsymbol{X}' = f(D')$ の固有ベクトルである. 固有ベクトルの敏感度の扱いは若干複雑である. ここでは固有ベクトルの代わりに, q 個の固有値 $\lambda_1, \lambda_2, \ldots, \lambda_k$ に対応した q 本の固有ベクトルで張られる不変部分空間の敏感度を扱う. $\boldsymbol{X}$ と $\tilde{\boldsymbol{X}}$ の固有値分解を $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U} = \left(\boldsymbol{U}_q \boldsymbol{U}_{\bar{q}}\right)\begin{pmatrix}\boldsymbol{\Lambda}_q & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Lambda}_{\bar{q}}\end{pmatrix}\begin{pmatrix}\boldsymbol{U}_q^T \\ \boldsymbol{U}_{\bar{q}}^T\end{pmatrix}, \tilde{\boldsymbol{X}} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{U}} = \left(\tilde{\boldsymbol{U}}_q \tilde{\boldsymbol{U}}_{\bar{q}}\right)\begin{pmatrix}\tilde{\boldsymbol{\Lambda}}_q & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{\Lambda}}_{\bar{q}}\end{pmatrix}\begin{pmatrix}\tilde{\boldsymbol{U}}_q^T \\ \tilde{\boldsymbol{U}}_{\bar{q}}^T\end{pmatrix}$ とする. ただし $\boldsymbol{U}_q$ と $\tilde{\boldsymbol{U}}_q$ は $N \times q$ 直交行列であり, $\boldsymbol{U}_{\bar{q}}$ と $\tilde{\boldsymbol{U}}_{\bar{q}}$ は $N \times (N - q)$ 直交行列である. このとき, 不変部分空間の加法的摂動に対する敏感度について以下の定理が知られている [2] .

matrix von-Mises Fisher mechanism
- Input: database $\boldsymbol{D} \in \mathcal{D}^N$, privacy parameter ϵ, rank $q \leq N$
- Output: matrix $\tilde{\boldsymbol{U}}_q \in \mathrm{St}^{N,q}$

1. Evaluate $\boldsymbol{X} = f(D)$, where $\boldsymbol{X}$ is positive semidefinite
2. Compute low-rank approximation $\boldsymbol{X}_q = \boldsymbol{U}_q \boldsymbol{\Lambda}_q \boldsymbol{U}_q^T$
3. Evaluate $\kappa = \frac{\epsilon\delta}{2q\|\Delta \boldsymbol{X}\|_2}$ be the concentration parameter
4. Sample $\tilde{\boldsymbol{U}}_q \sim \mathrm{mvMF}(\boldsymbol{Z}; \boldsymbol{U}_q, \kappa)$. Outputs $\tilde{\boldsymbol{U}}_q$.

図 3: Randomization mechanism for eigenvectors.

Theorem 4 *Let $\mathcal{X}$ and $\tilde{\mathcal{X}}$ be subspaces spanned by column vectors of $\boldsymbol{U}_q$ and $\tilde{\boldsymbol{U}}_q$. Let $\theta_1 \geq \theta_2 \ldots \geq \theta_q$ be the canonical angles between $\mathcal{X}$ and $\tilde{\mathcal{X}}$ and $\Theta = diag(\theta_1, \theta_2, \ldots, \theta_q)$. Assume there is an interval $[\beta, \alpha]$ and a $\delta > 0$ such that the eigenvalues of $\boldsymbol{\Lambda}_q$ lie in $[\alpha, \beta]$, while the eigenvalues of $\boldsymbol{\Lambda}_{\bar{q}}$ in $(0, \beta - \delta)$. Then, $\|\sin\Theta\|_2 \leq \frac{\|\Delta \boldsymbol{X}\|_2}{\delta}$.*

θ_{max} を最大の正準角, $\tilde{\lambda}_1$ を最大固有値とすると, $\sin\theta_{max} = \|\sin\Theta\|_2 \leq \frac{\|\Delta \boldsymbol{X}\|_2}{\delta} = \frac{\tilde{\lambda}_1}{\delta}$. である. このように不変部分空間の敏感度は摂動行列のスペクトルノルムと固有値のギャップでバウンドされることが示唆される. この結果をランダム化メカニズムに利用するには, 正規直交行列を確率変数に取る分布を用いる. ここでは matrix von Mises-Fisher 分布 (mvMF 分布) を用いる [4]. この分布の確率密度は下式で与えられる.

$$\mathrm{mvMF}_{n,r}(\boldsymbol{Z}; \boldsymbol{F}, \kappa) = a(\boldsymbol{F}) \exp(\mathrm{tr}(\kappa \boldsymbol{F} \boldsymbol{Z}^T)) \quad (6)$$

ここで $\boldsymbol{Z} \in \mathrm{St}^{N,q}$ は確率変数, $\boldsymbol{F} \in \mathrm{St}^{N,q}$ と $\kappa > 0$ はパラメタ, $\mathrm{tr}(\cdot)$ はトレースを表す. mvMF は $\boldsymbol{F}$ で最も高い確率密度を持つ. κ は集中度と呼ばれ, この値が大きいほど, 分布の $\boldsymbol{F}$ への集中度が高まる. $a(\boldsymbol{F})$ は正規化定数である. これを用いて固有ベクトルのランダム化メカニズムについて, 以下の定理を得る (証明略). またこれを実現するアルゴリズムを図 3 に示す.

Theorem 5 *Let $\boldsymbol{X} = f(D)$ and $\boldsymbol{X}' = f(D')$ where $D \sim D'$, respectively. Let $\boldsymbol{U}_q \in St^{N,q}$ and $\boldsymbol{U}_q' \in St^{N,q}$ be eigenvectors of rank-q approximation of $\boldsymbol{X}$ and $\boldsymbol{X}'$, respectively. If $\kappa \leq \frac{\epsilon\delta}{q\|\Delta \boldsymbol{X}\|}$, then $\tilde{\boldsymbol{U}}_q \sim mvMF(\boldsymbol{Z}; \boldsymbol{U}_q, \kappa)$ satisfies ϵ-differential privacy.*

集中度のバウンド $\kappa \leq \frac{\epsilon\delta}{q\|\Delta \boldsymbol{X}\|}$ に着目されたい. 分母に出力行列の列数 (=rank) が含まれることから, 出力が低ランクであれば, 必要な摂動のサイズは小さくて済む. 多くのデータマイニングアプリケーションでは原行列の低ランク近似で十分良い精度の結果を返す場合が多いことから提案メカニズムはプライバシー保護と効用の維持の両立にとって有用な性質を持つ.

4. Experiments

この章では, 提案メカニズムがプライバシー保護と効用のトレードオフをランクに応じて達成することを協調フィルタリングタスクを用いて実験的に示す.

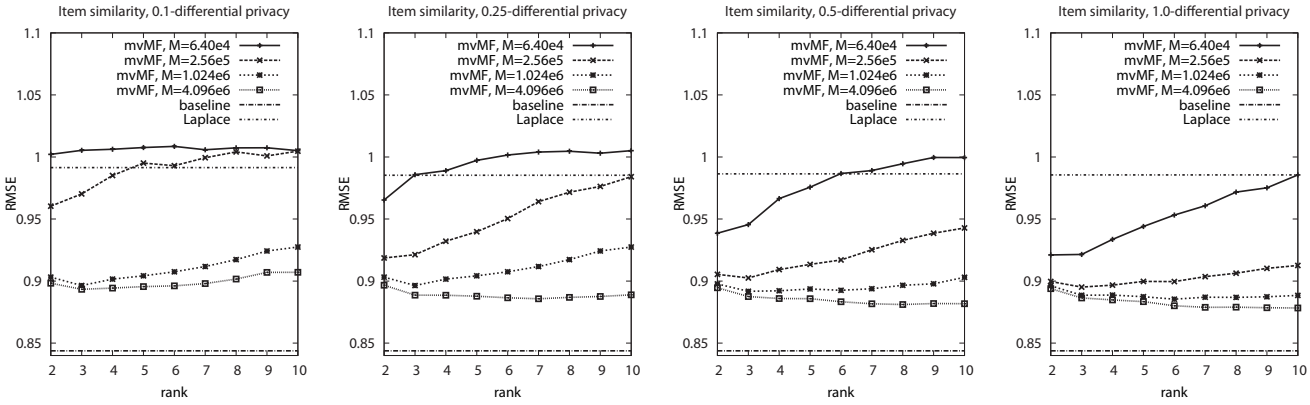

図 4: Rank vs. RMSE of collaborative filtering. From left to right, the privacy parameter was varied as $\epsilon = 0.1, 0.25, 0.5$, and 1.0. The number of items is fixed as $N = 500$. The number of users was varied as $M = 6.4 \times 10^4, 2.56 \times 10^5, 1.024 \times 10^6$, and 4.096×10^6. The RMSE evaluated using non-randomized covariance matrices (baseline) and covariances randomized with the Laplace mechanism (Laplace) are also shown.

データセット. 協調フィルタリングのためのデータとして Movielens (10M dataset) を用いた. 手法はアイテム類似度協調フィルタリングである [7]. この実験では, $N = 500$ ユーザと $M = 64,000$ アイテムを含む 351,323 レイティングのデータを用いた. 評価は 10-fold 交差検定で行い, 80%を類似度行列の作成に利用し, 残りは予測の評価に用いた.

評価. s_{ik} を訓練事例 I_{tr} における i 番目のユーザの映画 k のレイティングとする. テスト事例 I_{ts} の予測レイティングをアイテム類似度で重みづけされたレイティングの評価値の平均値とする: $\hat{s}_{ik} = \frac{\sum_{j=1}^{N} \sigma_{ij} s_{jk}}{\sum_{j=1}^{N} \sigma_{ij}}$. 予測精度はテスト事例の RMSE(最小二乗誤差の平方根) として評価する: RMSE $= \sqrt{\frac{1}{|I_{\mathrm{ts}}|} \sum_{(i,k) \in I_{\mathrm{ts}}} (\hat{s_{ik}} - s_{ik})^2}$.

設定. ベースラインとしてこの類似度行列の低ランク近似を用いて予測した値および要素毎のラプラスメカニズムにおける最小二乗誤差の平方根 (RMSE) を用いた. 提案法の評価には, 提案メカニズムでランダム化した類似度行列による予測値の RMSE を用いた. 提案法およびラプラスメカニズムのプライバシーパラメタは $\epsilon = 0.1, 0.25, 0.5, 1.0$ とした. 提案メカニズムが出力するランダム化類似度行列のランクは $q = 2, 3, \ldots, 10$ とした. ユーザ数の変化が RMSE に与える影響を評価するために $M = 6.4 \times 10^4, 2.56 \times 10^6, 1.024 \times 10^6, 4.096 \times 10^5$ とした. 10M データセットに含まれるユーザ数は高々71,567 であるため, アイテム類似度行列を固定し, メカニズムの入力パラメータのみを変化させた.

実験結果と議論. 図 4 は予測精度の RMSE(10-fold 交差検定) を示している. プライバシー保護をしない場合の RMSE は 0.84 程度である. 類似度行列全体が開示された場合, ラプラス分布のスケールパラメータは $b \in O(1)$ であり, 摂動のスケールはユーザ数が多数であっても小さくはならないため, ラプラスメカニズムの RMSE はどのような ϵ であっても, ユーザ数にかかわらずおおむね 0.98-0.99 程度である. この値は推薦結果としては妥当とは言えない水準である. 提案メカニズムの摂動のスケールは $\kappa^{-1} \in O(1/M)$ であり, 予測精度はユーザ数の増加とともに改善する. 興味深いことに, 提案メカニズムでランダム化した場合, ϵ が小さくとも, ランク q が小さくユーザ数 M が大きければ予測精度はおおむね 0.9 程度以下と比較的良い. また大きい ϵ と大きいユーザ数 M においては, 予測精度は高ランクにおいても比較的良い. このことは, プライバシー保護と効用の維持において, トレードオフが存在し, ランクによりそれを制御可能であることを示唆している.

結論. 本稿では, 半正定値行列のためのランダム化メカニズムを提案した. キーアイディアは以下の二つである; (1) 入力の固有値分布を考慮した, 固有値と固有ベクトルの個別ランダム化, (2) 低ランク近似の差分プライバシーのためのランダム化への組み込み. 協調フィルタリングにおける実験結果では, 低ランク近似が効用の維持とプライバシー保護をバランスさせることができることを示した. 協調フィルタリング以外への提案法の適用は将来の課題である.

謝辞 本研究は JST さきがけ「知の創成と情報社会」および部分的に最先端研究開発プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会サービスの実証・評価」の助成を受けた.

参考文献

[1] Y. Azar, A. Fiat, A. Karlin, F. McSherry, and J. Saia. Spectral analysis of data. In *Symposium on Theory of Computing*, volume 33, pages 619–626, 2001.

[2] C. Davis and W. Kahan. Some new bounds on perturbation of subspaces. *Bull. Amer. Math. Soc*, 75(4):863–868, 1969.

[3] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. *Theory of Cryptography*, pages 265–284, 2006.

[4] A. Gupta and D. Nagar. *Matrix variate distributions, volume 104 of Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2000.

[5] D. Kifer and B. Lin. Towards an axiomatization of statistical privacy and utility. In *Proceedings of the 29th Symposium on Principles of database systems of data*, pages 147–158. ACM, 2010.

[6] F. McSherry and I. Mironov. Differentially private recommender systems: building privacy into the net. In *Proc. of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 627–636. ACM, 2009.

[7] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. In *Proc. of the 10th international conference on World Wide Web*, pages 285–295. ACM, 2001.

[8] G. Stewart and J. Sun. *Matrix perturbation theory*, volume 175. Academic press New York, 1990.