

# T-1: No Bias Left Behind: Covariate Shift Adaptation for Discriminative 3D Pose Estimation

Makoto Yamada, Leonid Sigal, and Michalis Raptis

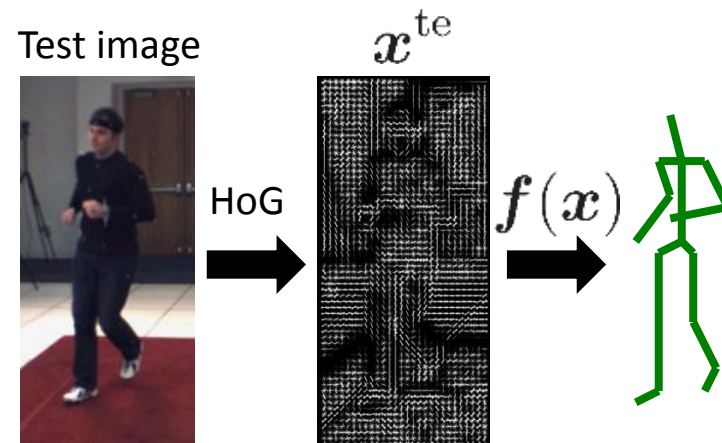


## ■ Discriminative 3D human pose estimation

Given large database of image-pose pairs  $\{(x_i^{\text{tr}}, y_i^{\text{tr}})\}_{i=1}^{n_{\text{tr}}}$ ,  
predict human pose of  $\{x_i^{\text{te}}\}_{i=1}^{n_{\text{te}}}$ .

$$y = f(x) + e$$

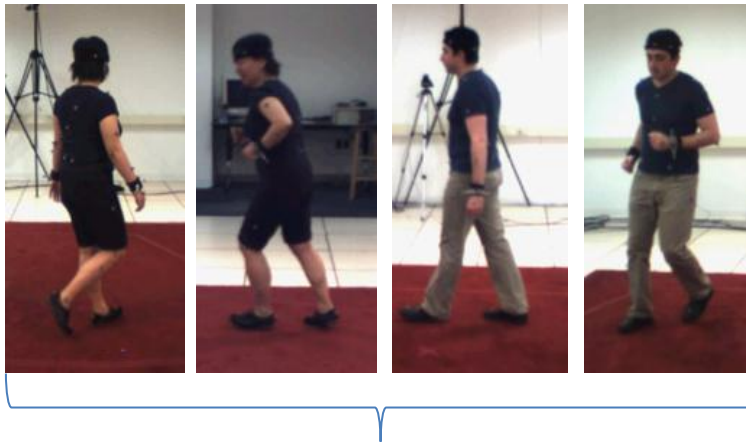
e.g., Nearest Neighbor  
Regression



# Research Motivation

## ■ Dataset Bias

- Sample selection bias
- Test subject is not included in training set
- Lighting, object appearance, viewpoint changes



Training images

Different  
subject



Test image

Pose estimation performance significantly degrades by **dataset bias** 😞  
Thus, we adopt **covariate shift adaptation** to alleviate **dataset bias**!

## T-2

# 大画像の複層ベイズ超解像と位置ずれ推定に関する検討

木下 俊貴  
(関西大大学院)

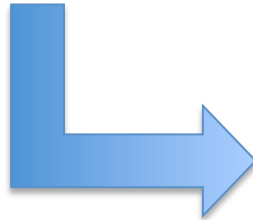
三好 誠司  
(関西大)

- 
- 超解像＝低解像画像から高解像画像を推定
  - ベイズ超解像＝ベイズ推定に基づく超解像 (Tipping&Bishop, 2002)
  - 先行研究
    - 複層ベイズ超解像 (Kanemura, Maeda & Ishii, 2009)
    - フルベイズ超解像 (Katsuki, Torii & Inoue, 2010)
    - ⇒先行研究では小さい画像を扱っていた
  - 今回の発表
    - 画像分割による大画像の複層ベイズ超解像
    - 回転の中心での画素値の変化が乏しい場合の位置ずれ推定

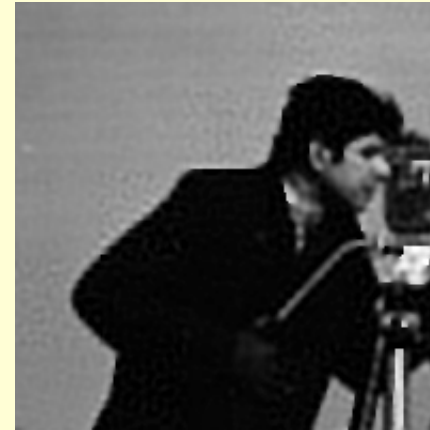
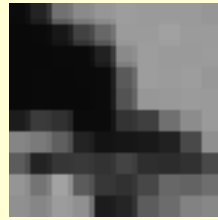


観測画像例

超解像  
処理

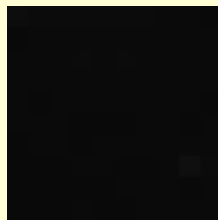


原画像



ISNR= 4.55dB

縦移動: 0.10 px  
横移動: 0.10 px  
回転: 0.003 rad



ISNR= 2.59dB

縦移動: 0.24 px  
横移動: 0.27 px  
回転: 0.004 rad

位置ずれ推定  
に用いた領域

推定画像

位置ずれ推定  
の誤差

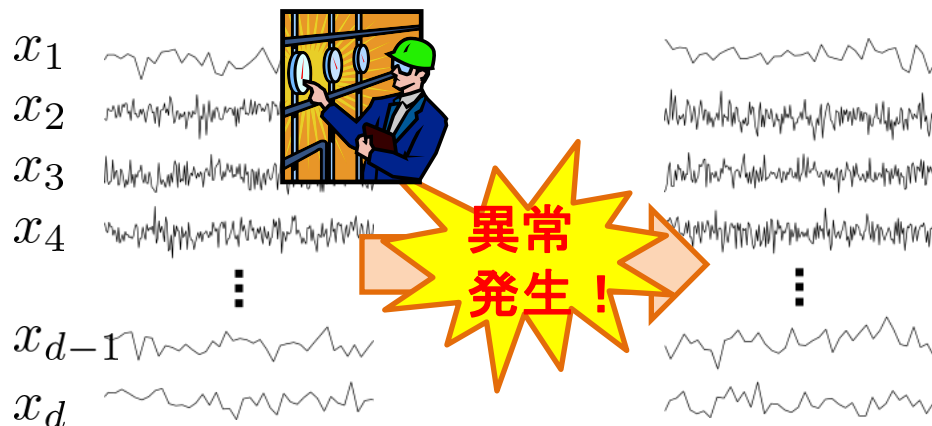


# 異常箇所同定のための グラフィカルモデルの学習

原聡, 鷲尾隆(阪大)

## ■ 異常箇所の同定問題

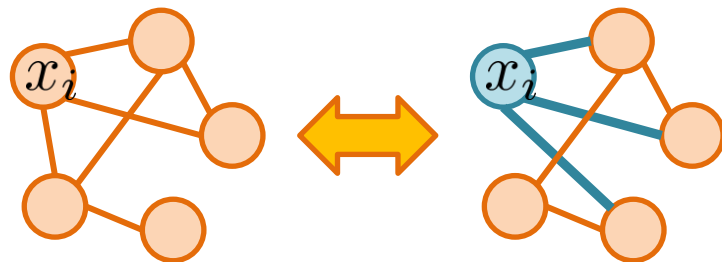
異常前後のデータを比較して  
どの変数に異常があるかを  
同定したい



## ■ グラフィカル・ガウシアン・モデルを用いる方法

- 1) データをグラフィカルモデルに変換 (Idé et al., SDM'09)
- 2) 2つのグラフの対比から異常部位を検出

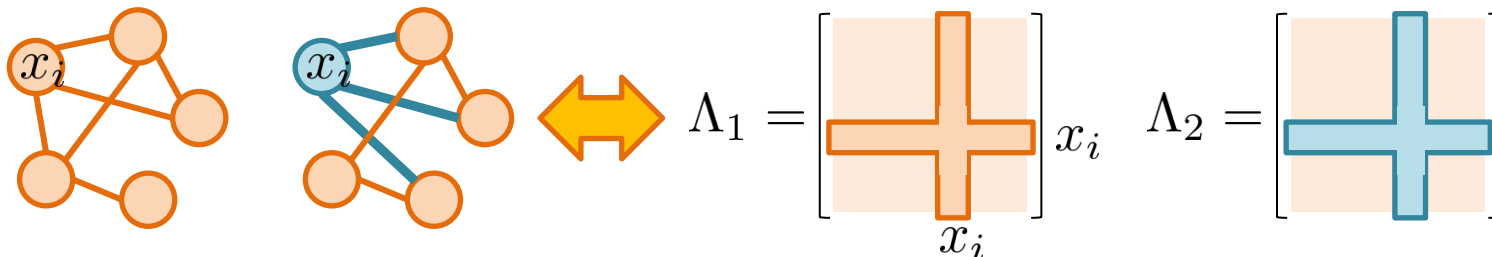
仮定:  
異常変数と正常変数間の  
グラフ構造は変化する.



# 本研究の概要

## ■ 異常箇所同定に適したグラフィカルモデル推定

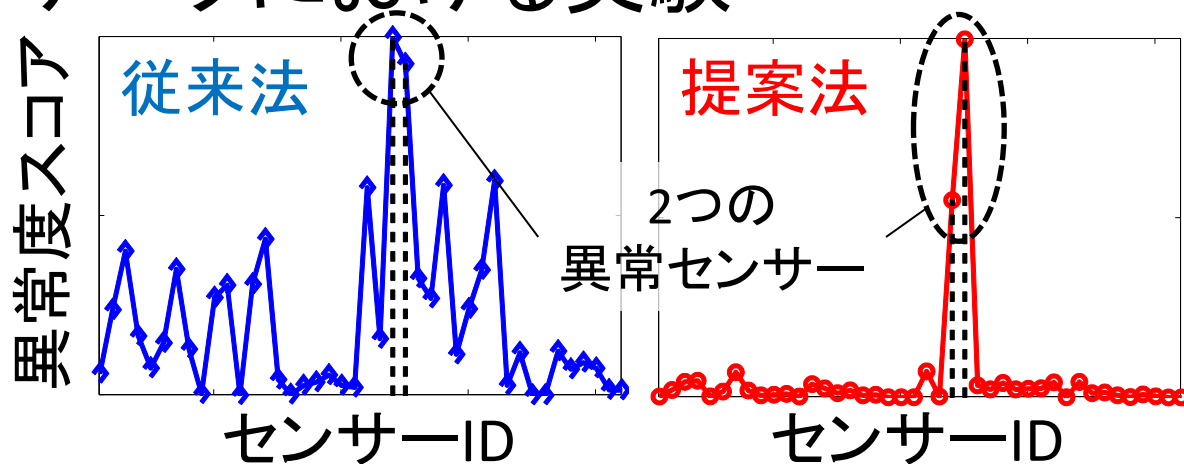
- 異常箇所と精度行列間の行・列ごとの変動が対応



提案法: 行・列の変動への制約を正則化項として導入

## ■ 自動車センサーデータにおける実験

従来法よりも  
異常箇所の  
同定性能が向上



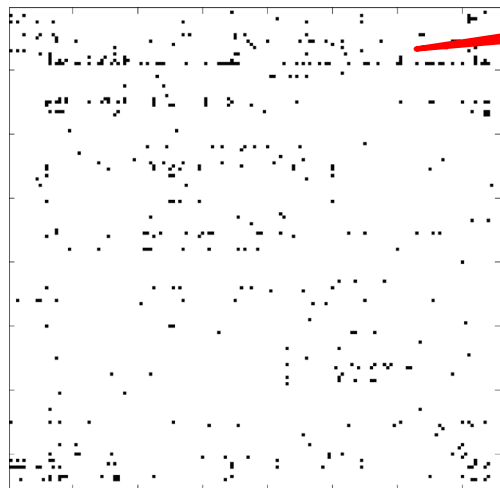
石黒 勝彦、上田 修功、澤田 宏(NTT)

やりたいこと: スパースだったりノイズが多く乗った関係データから、オブジェクトをクラスタリングしたい

オブジェクトj

スパース: 黒い(=1)要素がごく少数  
ノイズ: ランダムな白黒パターン

オブジェクトi



購買履歴{商品, 顧客}

→ 販売実績は行列サイズよりずっと少ない

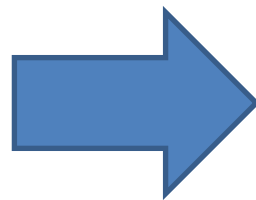
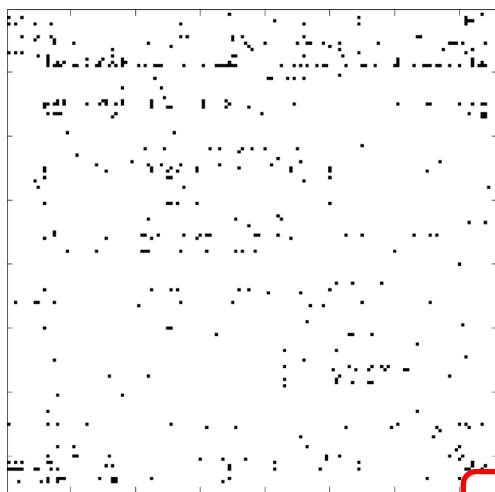
SNSリンク{ユーザ, ユーザ}

→ スпамユーザ、botユーザの存在

仮説：関係データ内のオブジェクトには、  
「relevant=潜在的なクラスタ構造をもつ」オブジェクトと  
「irrelevant=そもそもクラスタ構造をもっていない」オブジェクト  
がいる

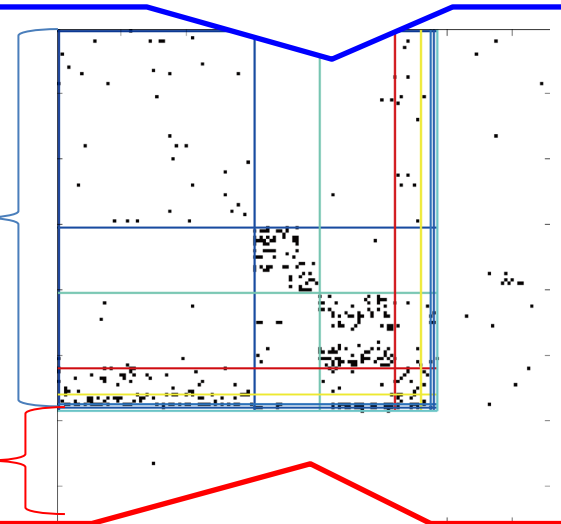
提案法：「クラスタ構造を持つかどうか」も同時に推定する  
クラスタリングモデル

☺ “relevant”なオブジェクトの中では鮮明なクラスタ



relevant

irrelevant



☺ irrelevantなオブジェクトはクラスタリングから除外します

# Perplexity on Reduced Corpora

小林隼人（東芝）

## 学習データが大量にあるときの3つの戦略

### 1. ちょっとずつ処理

分散処理（並列的）、オンライン処理（直列的）

### 2. ちょっとだけ処理

**カットオフ（決定的）**、サンプリング（確率的）

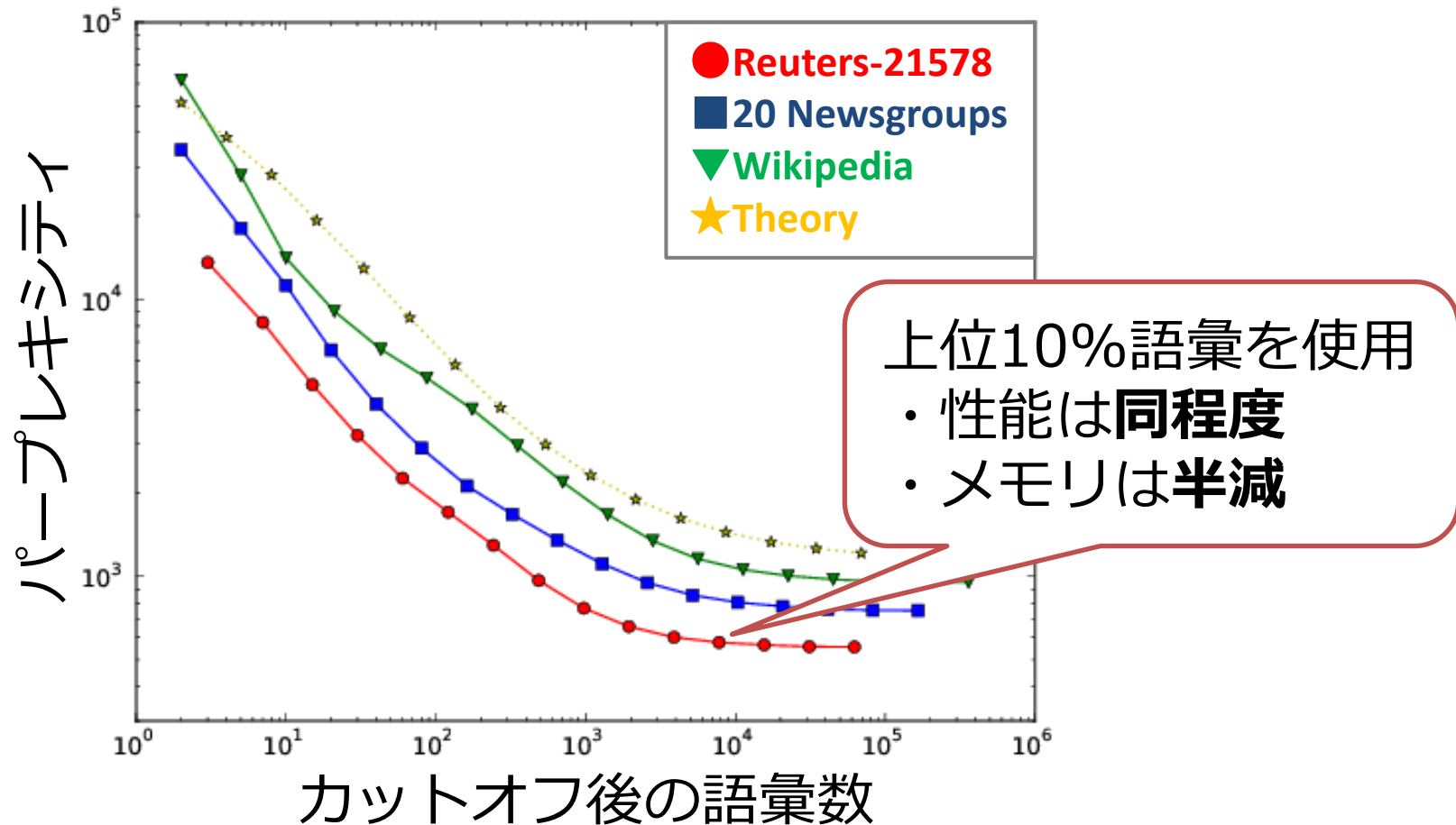
### 3. 処理しない

にげる（戦略的）、すて（エコ）

戦略2のカットオフ（低頻度データを削除）  
でどの程度性能が劣化するか検討しました

# 主結果

データがZipf則に従うとき、LDAのパープレキシティは  
カットオフ後の語彙数に対して**準二次**（対数グラフで二次）



※迷走中につき、ご助言ください

# 混合ベルヌーイ分布の変分ベイズ学習におけるモデル選択

## 大工廻和美・渡辺澄夫(東工大)

### 研究の目的

### WAIC-VBの提案

WAICを計算する方法として、平均場近似された事後分布(変分ベイズ推測)を重点サンプリングにおける提案分布として用いることを提案。

- ①WAIC-VBを混合ベルヌーイ分布の場合について適用し、ベイズ汎化損失の不偏推定に有用であるかどうかを実験。
- ②人工データだけでなく画像を用いた実問題についても実験し、モデル選択時に汎関数分散の値も参考にする意義の確認。

### 原理

### WAIC-VB

WAICを求めるときに、直接サンプルを抽出するのが困難な事後分布の代わりに、平均場近似された事後分布から抽出し、重要度重みによってバイアスを補正する。

$$\text{WAIC} \quad W_n = T_n + \frac{V_n}{n}$$

$T_n$  : 経験損失

$V_n$  : 汎関数分散

## 実験

混合ベルヌーイ分布について変分ベイズ推測を行い、自由エネルギー、汎化誤差、WAICの値を算出する。

### 【人工データ】

入力の次元 $M=6$ ,  
真の分布: $K=3$ の混合ベルヌーイ分布  
学習データ数  $n=200$   
学習モデルのコンポーネント数:  $1 \rightarrow 5$

### 【実データ】

$M=27$ 次元の0または1のデータ  
画像から $3 \times 3$ ピクセルを取り出してRGBのそれぞれの値について閾値処理を行って0と1とする。  
学習データ数  $n=500$   
学習モデルのコンポーネント数:  $1 \rightarrow 8$

## 結果

- 自由エネルギー及びWAICにより、 $K=3$ でモデル選択できる。
- 有効なパラメータの個数を表すは $K=3,4,5$ でほぼ同じ値になっている。
- 自由エネルギー及びWAICにより、 $K=6$ でモデル選択できる。
- 有効なパラメータの個数を表す  $V_n$  は $K=7,8$ でほぼ同じ値になっている。

## まとめ

混合ベルヌーイ分布において、汎化損失の平均とWAIC-VBにより求められたWAICの平均値はほぼ同じになることを、人工データと実データについて実験的に明らかにできた。



# T-7 Robustness of time-consistent Markov decision processes

恐神貴行 (IBM東京基礎研究所)

## Time-consistent MDP

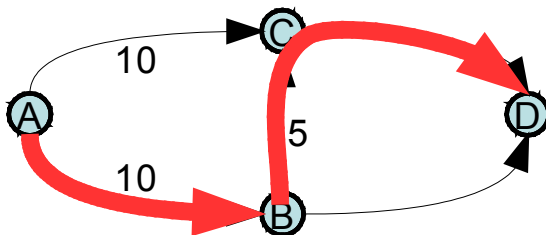
- 😊 Can always make consistent decisions
- 😊 Can represent a wide-range of risk preference
- 😞 Not easy to understand the meaning

## Risk-sensitive MDP

Parameters are known

Minimize a risk measure of cumulative cost

- 😞 Can make inconsistent decisions = Can surely loose infinitely



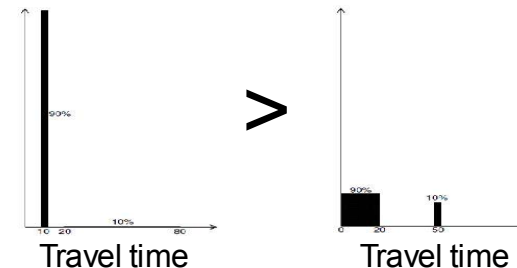
## Robust MDP

Parameters are not known exactly

Minimize expected cumulative cost for the worst case

## Standard risk-sensitive MDP

- 😞 Can only represent limited risk preference



# T-7 Robustness of time-consistent Markov decision processes

## 恐神貴行 (IBM東京基礎研究所)

We show “a certain time-consistent MDP = a certain robust MDP”

- ➔ Can intuitively understand the objective of that time-consistent MDP
- ➔ Can find the optimal policy of that robust MDP efficiently with the knowledge of the equivalence to the time-consistent MDP

**Example.** Optimal policy with respect to the following objective can be found quite efficiently:

$$\min_{\pi} \max_{\substack{p^{\pi} \in \mathcal{P}(p_0^{\pi}) \\ f^{\pi} \in \mathcal{P}(f_0^{\pi})}} \left\{ \mathbb{E}_{p^{\pi}, f^{\pi}} \left[ \sum_{i=0}^{N-1} C_i^{\pi} \right] - \gamma \mathbb{E}_{p^{\pi}} \left[ \sum_{i=0}^{N-2} H(p^{\pi}(\cdot|S_i) || p_0^{\pi}(\cdot|S_i)) + \sum_{i=0}^{N-1} H(f^{\pi}(\cdot|S_i) || f_0^{\pi}(\cdot|S_i)) \right] \right\}$$

Minimize for  
the worst case

Cumulative cost

Penalty for the deviation of the parameters from nominal values  
(H: Kullback-Leibler divergence)

### References

T. Osogami, “Robustness and risk-sensitivity in Markov decision processes,” NIPS 2012

T. Osogami and T. Morimura, “Time-consistency of optimization problems,” AAAI-12

T. Osogami, “Iterated risk measures for risk-sensitive Markov decision processes with discounted cost,” UAI 2011

# T-8 二次形式の大域的最適化によるクラスタリング

広瀬俊亮 (SAS Institute Japan)

## 解きたい問題: クラスタリング

- クラスタ数は既知とする。

## 要請: できるだけ簡単な方法を構築したい

1. 大域的な最適化問題として解ける。
2. パラメータチューニングの手間がいない。

## 解き方:

	クラスタの形	パラメータチューニング
分布を仮定	分布の形に結果が依存	パラメータは推定される
<b>分布を仮定せず</b>	<b>クラスタの形が柔軟</b>	<b>チューニングが必要</b>
	情報量	目的変数への制約
<b>帰属確率を求める</b>	<b>直接利用できる</b>	<b>制約有(負の確率は許されない)</b>
インディケータを求める	直接は使えない	制約の有無は場合に依る

本稿では  
この組を採用。

# T-8 二次形式の大域的最適化によるクラスタリング

## 提案手法：二次形式の大域的最適化を組み合わせたクラスタリング

- データ点：  $\{\mathbf{x}_i\}_{i=1}^n$  クラスタ：  $\{C_l\}_{l=1}^K$

1. **固有値問題**を解いて、 $p(\mathbf{x}|C)$ を求める。

$$\min_{\{\psi\}} \frac{1}{2} \sum_{i,j=1}^n S_{ij} (\psi_l^i - \psi_l^j)^2$$

### 確率振幅の導入

Spectral Clustering等で現れる負の確率の問題を回避。

$$(\psi_l^i)^2 = p(\mathbf{x}_i | C_l)$$

2. **凸二次最適化問題**を解いて、 $p(C)$ を求める。

$$\min_{\mathbf{q}} \mathbf{q}^\dagger Q \mathbf{q}$$

$(q_l = p(C_l) \ (l = 1, \dots, K))$

### 最大エントロピー原理

$P(C)$ を既知とせず、推定する。

$$p(C) = \operatorname{argmin}_{p(C)} \frac{1}{n} \sum_{i=1}^n p(\mathbf{x}_i)^2$$

### 目的関数

$p(C|\mathbf{x})$ を直接求めず、 $p(\mathbf{x}|C)$ と $P(C)$ を求める。

### 提案手法の特長

1. 大域的な最適解が得られる。
2. 求めた確率と情報量(相互情報量)を用いてパラメータ選択が可能。

# T-9 変分ベイズ法の局所解における 自由エネルギーと汎化誤差の関係

東京工業大学

中村文士・渡辺澄夫

## 目的

変分自由エネルギーを最小にする局所解は汎化誤差を最小にするのかどうか調べる

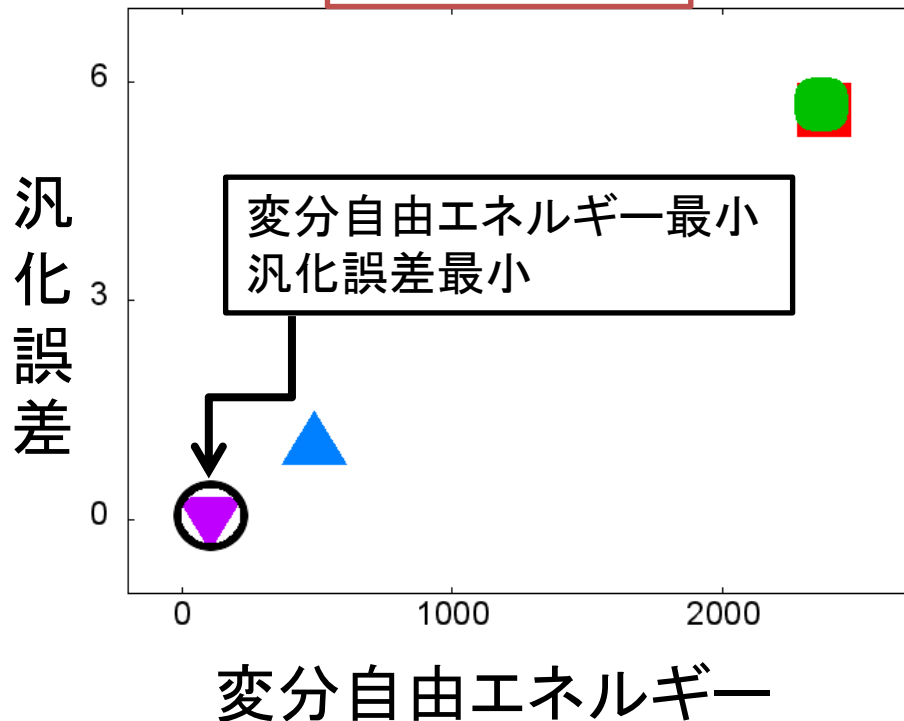


# 実験結果

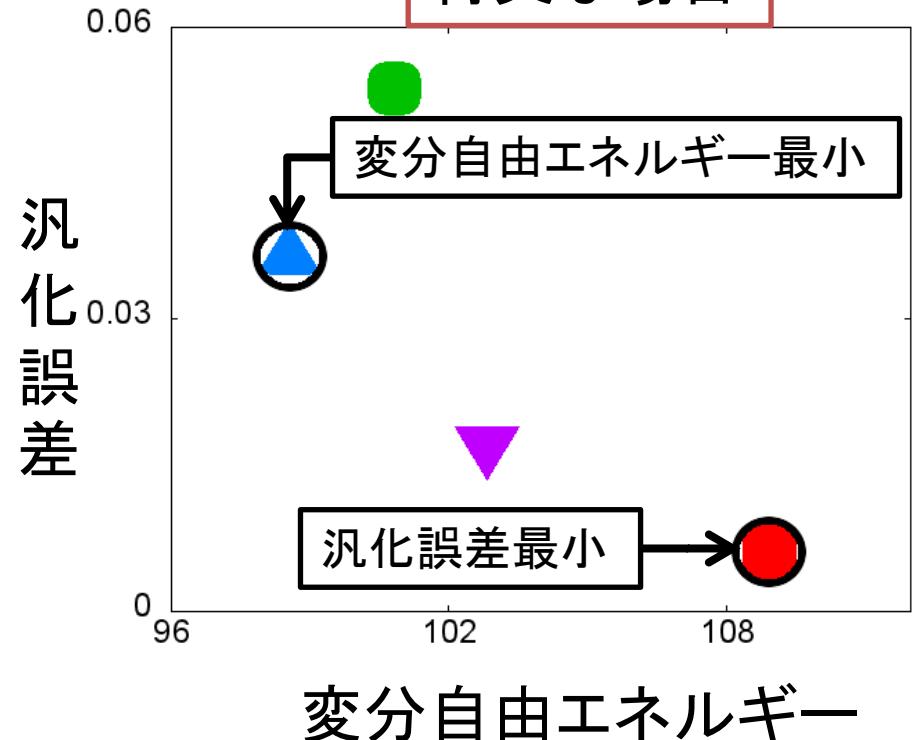
正則な場合: 変分自由エネルギーを最小にする局所解は汎化誤差も同様に最小にする

特異な場合: 変分自由エネルギーを最小にする局所解は汎化誤差を最小にするとは限らない

正則な場合



特異な場合



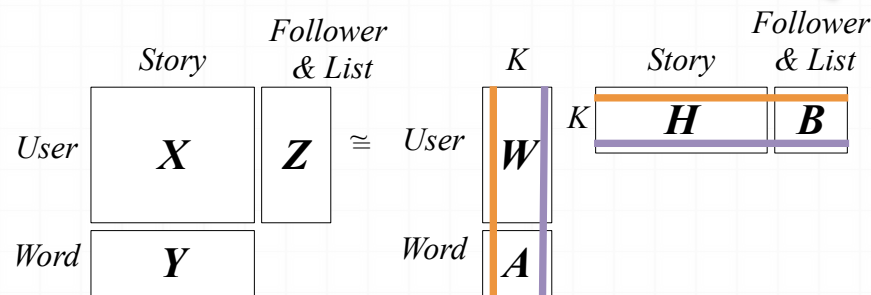
※実データへの応用はポスターで



# ◆ ソーシャルメディアからのオーソリティ発見

つづきは  
T-10で

- ✓ トピックについてオーソリティを発見できれば、信頼性の高い情報をリアルタイムで取得可能になる
- ✓ **Twitter**と**ソーシャルキュレーション**のデータをsNMFで解析し、トピックとオーソリティを発見した



$$X \in \mathbb{R}_+^{1,823,184 \times 235,086} \quad (0.0017\% \text{ non-zero})$$

$$Y \in \mathbb{R}_+^{165,046 \times 235,086} \quad (1.23\% \text{ non-zero})$$

$$Z \in \mathbb{R}_+^{1,823,184 \times 2} \quad (100\% \text{ non-zero})$$

震災トピック

## Topical Words

時震度分地震県情報最大km発生深い震源気象庁福島n3近辺市沖速報PM長官会見総理AM官房日推定アップお知らせいわき首相率線量m4会津若松白河南相馬市郡山番南会津風

Authority	Profile Summary	Follower
kiri_tori	時報bot	166,845
47news	共同通信非公式 bot	108,512
earthquake_jp	地震情報 bot	805,341
Asahi_kantei	朝日新聞官邸クラブ取材班	37,622
Kantei_Saigai	首相官邸公式アカウント	422,307

Hadoopトピック

## Topical Words

使う Hadoop クラウド さん AWS ない 話 データ これいい 開発 No セッション そう PHP de 的 処理 js DB サーバ ヒップホップ er 人 MySQL 会場 Mongo DB 見る Azure 今日 すごい LT 資料 書く 勉強 Ust Windows Cloud IT 中

Authority	Profile Summary	Follower
understeer	Solutions Architect	2,368
yutuki_r	Cloud, Hadoop, NoSQL	800
kimtea	NoSQL, Google AppEngine	2,742
repeatedly	D, Ruby, Python, Opera	4,081
siena	ウェブ技術全般, 情報設計	1,065



# T-11

## 自由エネルギーによる潜在変数推定精度の計算法

山崎啓介(東工大), 渡辺一帆(奈良先端大), 梶大介(コニカミノルタエムジー)

背景： 潜在変数をもつモデルが広く使われている

モデル	混合分布	隠れマルコフモデル	ベイジアンネットワーク
潜在変数	(コンポーネントの) ラベル	隠れ状態	観測不能なノード

近年、ベイズ潜在変数推定の精度が解明された

問題： 潜在変数は見えないため、データから誤差を計算することができない

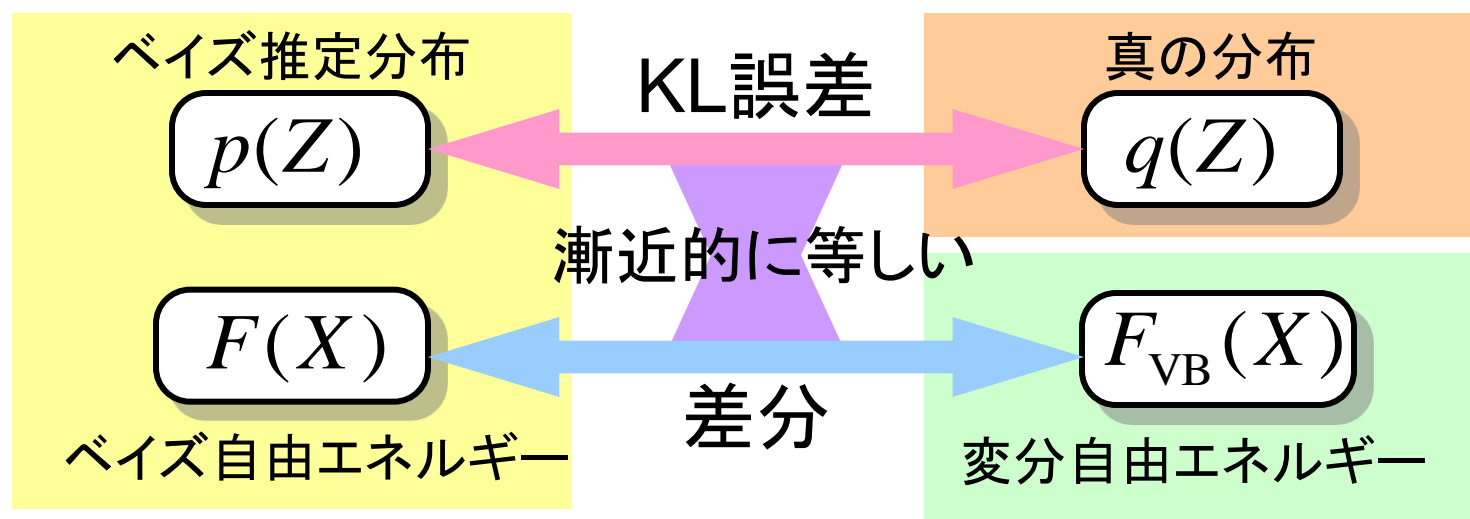
※クロスバリデーションのような数値計算も不可能

# T-11

## 自由エネルギーによる潜在変数推定精度の計算法

山崎啓介(東工大), 渡辺一帆(奈良先端大), 梶大介(コニカミノルタエムジー)

提案法: ベイズ法、変分ベイズ法における自由エネルギー差分と誤差関数の漸近的な等価性を用いて計算を行う



$Z$  : 潜在変数

$X$  : 観測データ

## テキスト分類問題におけるカテゴリ情報を用いた適応的距離学習に関する一考察 三川 健太, 石田 崇, 後藤 正幸, 平澤 茂一 (早稲田大学)

**拡張余弦尺度**: 高次元, スパースな文書データを対象としたメトリック

$$\begin{aligned} \text{sim}_Q(\mathbf{d}_i, \mathbf{d}_j) &= \frac{\mathbf{d}_i^\top Q \mathbf{d}_j}{\|\mathbf{d}_i\| \|\mathbf{d}_j\|} \\ &= \frac{\sum_{k=1}^W \sum_{l=1}^W q_{kl} d_{ik} d_{jl}}{\sqrt{\sum_{k=1}^W (d_{ik})^2} \sqrt{\sum_{k=1}^W (d_{jk})^2}}. \end{aligned}$$

ただし, 文書  $\mathbf{d}_i$ :  $\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{iW})^\top$ ,  $Q = [q_{kl}]$ ,  $\mathbf{d}_i \in \mathcal{R}^W$ ,  $Q \in \mathcal{R}^{W \times W}$  とする.

**拡張余弦尺度を用いた文書分類手法(従来手法)**:

テストデータ  $d$  を  $\arg \max_n \text{sim}_{\hat{Q}}(d, \mathbf{g}_n)$  を満たすクラスへ分類.

ただし, カテゴリ集合を  $\mathcal{C} = \{c_1, c_2, \dots, c_N\}$ , その代表元を  $\mathbf{g}_n = (g_{n1}, g_{n2}, \dots, g_{nW})^\top$  とし, 計量行列  $\hat{Q}$  は学習データから推定する.



カテゴリ間の統計的性質の差異については未考慮

**提案手法**:

カテゴリ毎に計量行列を導入, カテゴリ間の性質の差異を考慮.



カテゴリ毎の統計的性質の差異を表現した計量行列を学習する

**カテゴリ毎の計量行列の学習方法**: 以下の最適化問題の解として導出

$$\hat{H}_n = \arg \max_{H_n} \left\{ \frac{\alpha}{|c_n|} \sum_{\mathbf{d}_i \in c_n} \text{sim}_{H_n}(\mathbf{d}_i, \mathbf{g}_n) - \frac{1-\alpha}{N-1} \sum_{\{n': n' \neq n\}} \frac{1}{|c_{n'}|} \sum_{\mathbf{d}_i \in c_{n'}} \text{sim}_{H_n}(\mathbf{d}_i, \mathbf{g}_{n'}) \right\},$$

subject to  $|H_n| = 1$ .

ただし,  $|H_n|$  は  $H_n$  の行列式,  $|c_n|$  はカテゴリ  $c_n$  に含まれるデータ数

**【定理1】** 上記最適化問題を満たす計量行列  $\hat{H}_n = [\hat{h}_{nkl}]$  は, 以下の式で与えられる.

$$\hat{H}_n = |S_n|^{1/W} S_n^{-1}.$$

ただし,  $S_n = [s_{nkl}]$  は以下の式を満たす.

$$s_{nkl} = \frac{\alpha}{|c_n|} \sum_{\{i: \mathbf{d}_i \in c_n\}} \frac{d_{ik} g_{nl}}{\sqrt{\sum_{k=1}^W (d_{ik})^2} \sqrt{\sum_{l=1}^W (g_{nl})^2}} - \frac{1-\alpha}{N-1} \sum_{\{n': n' \neq n\}} \frac{1}{|c_{n'}|} \sum_{\{i: \mathbf{d}_i \in c_{n'}\}} \frac{d_{ik} g_{n'l}}{\sqrt{\sum_{k=1}^W (d_{ik})^2} \sqrt{\sum_{l=1}^W (g_{n'l})^2}}.$$

**実験:**

新聞記事データ(毎日新聞2000年)を対象, カテゴリ数5, 学習データ数300件/カテゴリ, テストデータ500件/カテゴリとした文書分類実験( $\alpha$ : 0.0~1.0まで0.1ずつ変更).

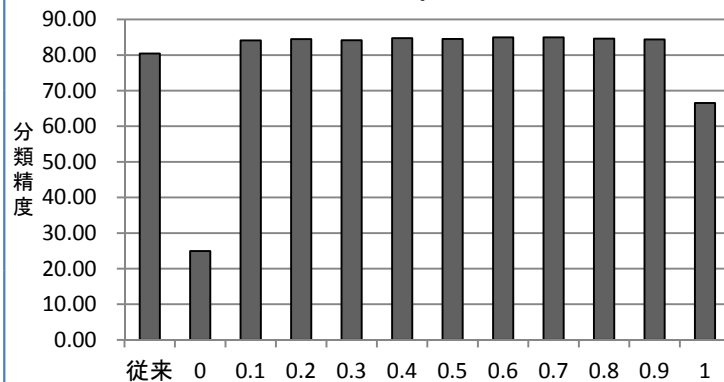


図1 文書分類実験結果

**分類ルール:**

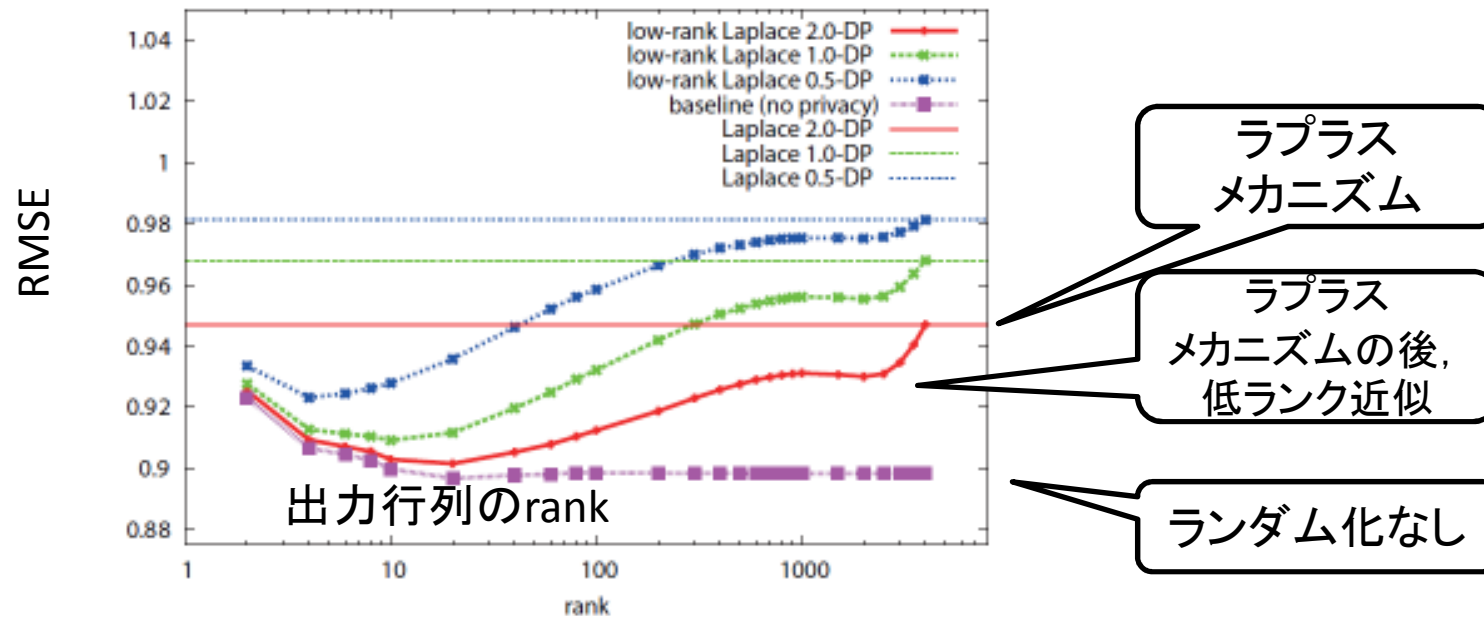
$$\hat{c} = \arg \max_{c_n} \left\{ \alpha \text{sim}_{\hat{H}_n}(\mathbf{d}, \mathbf{g}_n) - \frac{1-\alpha}{N-1} \sum_{\{n': n' \neq n\}} \text{sim}_{\hat{H}_n}(\mathbf{d}, \mathbf{g}_{n'}) \right\}.$$

実験結果より, 従来手法と比較して  
精度の向上を確認した.

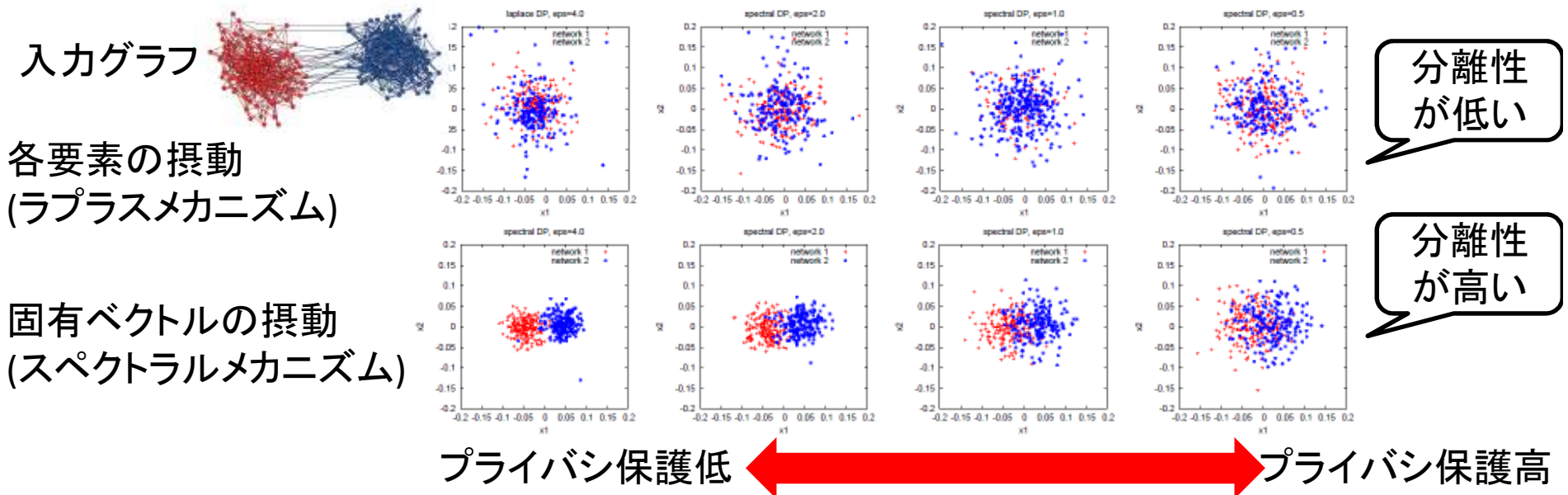
# 半正定値行列の差分プライバシー

- 差分プライバシー(DP: differential privacy)
  - データのランダム化によるプライバシー保護の達成
- 半正定値行列
  - 共分散行列, グラム行列などMLにとって重要な表現
  - 半正定値行列のDPはこれまで議論なし
- 主定理1
  - ランダム化後にrank- $q$ 近似を適用すれば, DPを維持したまま真の行列からの差は $O(q^{1/2}N^{1/2})$
- 主定理2
  - 固有値分解してからランダム化すれば, DP維持に必要なノイズ複雑度は $O(q)$
- 低ランク近似された行列で満足できるならば, DP維持に必要なノイズ複雑度を小さくできる

# 主定理1の数値例: 差分プライベートなユーザ類似度に基づく推薦



# 主定理2の数値例: 差分プライベートなスペクトラル分析



# T-14 交通流の時空図における ZRPのパラメータ推定と能動学習

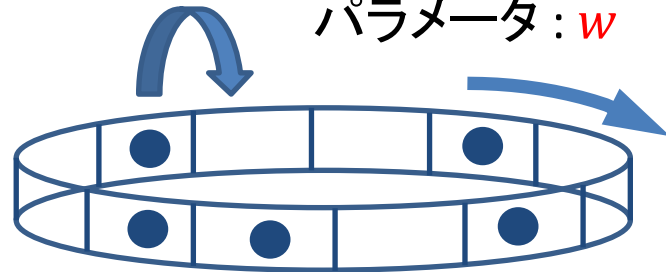
小林 浩一, 山崎 啓介 (東京工業大学)

目的: ZRP (交通流のモデル) におけるパラメータ推定の定式化

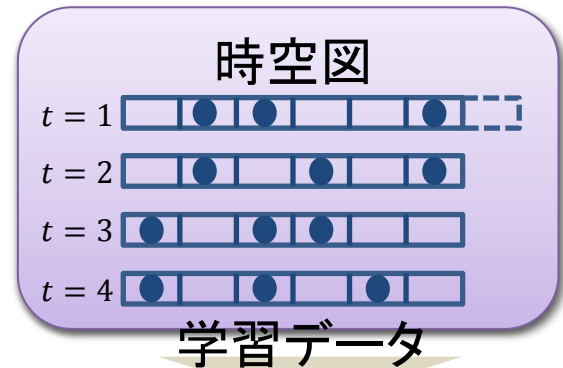
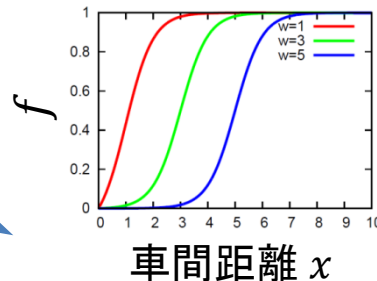
提案: 推定量の二乗誤差を最小化する能動学習アルゴリズムの構築

前進確率  $f(x, w)$

パラメータ:  $w$



ZRPの模式図



各車について

$X$ : 車間距離  
 $Y$ : 進む(1), とどまる(0)

- 分岐などのない1車線の環状な道路を確率セルオートマトンにより再現
- 車の前進確率はパラメータに依存

データの従う分布

$$X, Y \sim p(x, y | w^*, \rho) = p(y | x, w^*) p(x | w^*, \rho)$$
$$p(y | x, w) = f(x, w)^y (1 - f(x, w))^{1-y}$$

# 推定量の二乗誤差

$$\begin{aligned} D(n) &= E_{X^n Y^n} [\|w^* - \hat{w}\|^2] \\ &= \text{Tr} \left[ \left( n I_\rho(w^*) \right)^{-1} \right] + o\left(\frac{1}{n}\right) \end{aligned}$$

$w^*$  : 真のパラメータ  
 $\hat{w}$  : 最尤推定量  
 $I_\rho(w^*)$  : フィッシャー情報量

$\rho$  : 密度  $\left( \frac{\text{車の台数}}{\text{システムサイズ}} \right)$

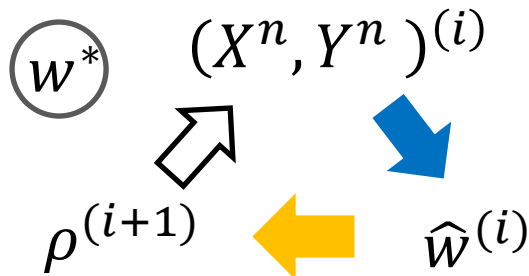
誤差の値が密度に依存!

誤差の推定量  $\text{Tr} \left[ \left( n I_\rho(\hat{w}) \right)^{-1} \right]$  から 誤差が最小となる密度 を決定  
$$\rho = \arg \min \text{Tr} \left[ \left( n I_\rho(\hat{w}) \right)^{-1} \right]$$

結論: ZRPにおいて能動学習を構築・実験により有効性を確認

## 能動学習

$w$  の推定と密度の選択を繰り返す



$$X, Y \sim p(x, y | w^*, \rho) = p(y | x, w^*) p(x | w^*, \rho)$$

$p(x | w^*, \rho)$  は直接最適化できないため  
 $p(x | \hat{w}, \rho)$  を最適化している



# エントロピー最小化によるRestricted Boltzmann Machineの学習の正則化

木脇太一（東大），牧野貴樹（東大），合原一幸（東大）

- 従来のRBMでは過学習が問題となっていた
- データの複雑性に関係なくRBMの全ての隠れ変数が利用されてしまう
- スパース性を導入しても，これは緩和されない

# エントロピー最小化によるRestricted Boltzmann Machineの学習の正則化

木脇太一（東大），牧野貴樹（東大），合原一幸（東大）

- 提案手法
  - RBMの学習に隠れ変数のエントロピーを正則化項として導入してRBMに複雑性の調節機能を取り入れる

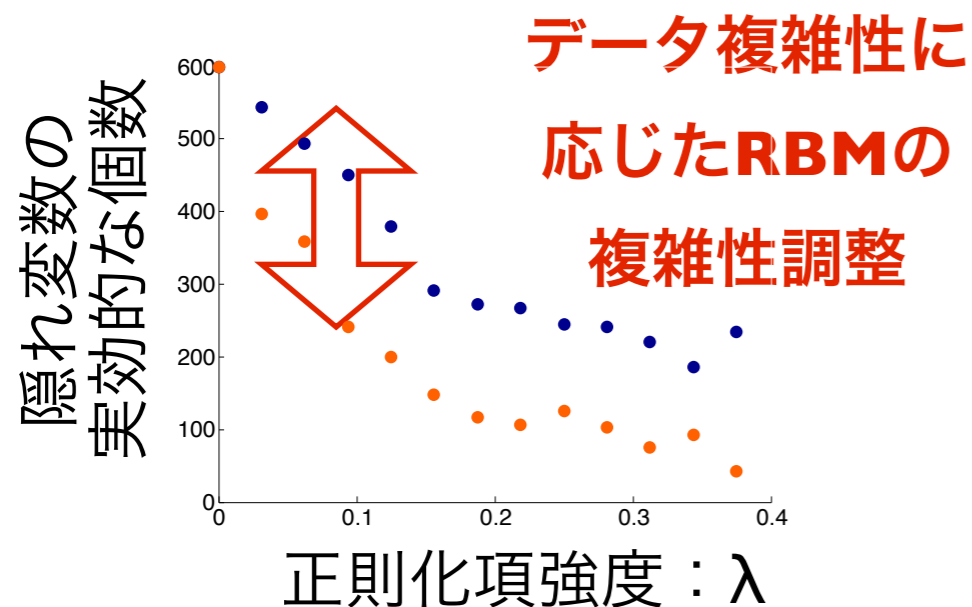
- 目的関数

$$L = \text{KL}(q||p) + \lambda H(p)$$

$q$  : データ分布

$p$  : 再現分布

- 実験結果



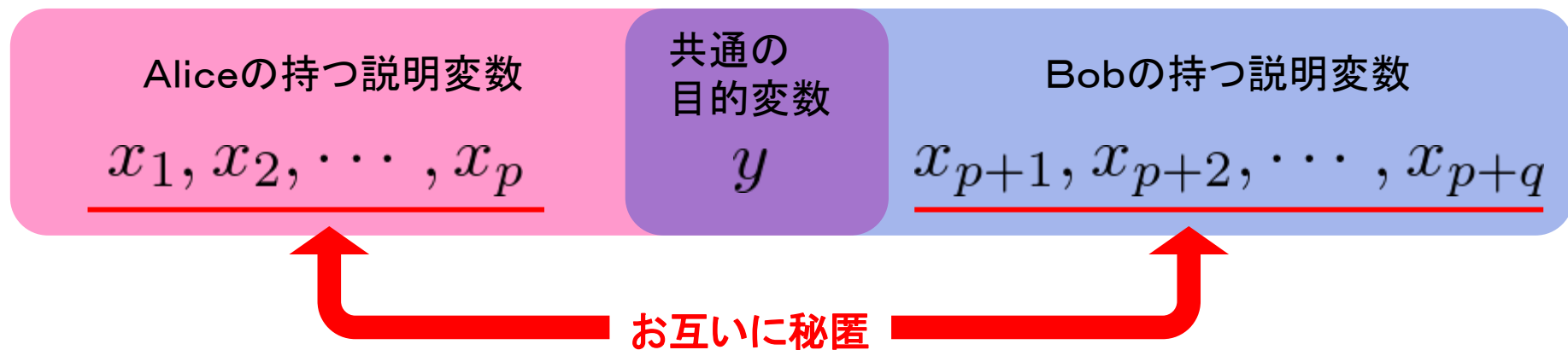
- まとめ
  - 隠れ変数のエントロピーを削減するRBMを正則化項の導入により提案
  - 提案手法ではデータの複雑性に応じてRBMの複雑性が動的に調整されることを実験的に確かめた

# プライバシー保護を目的とした線形回帰モデルにおける最小二乗推定量の分散計算法について

須子統太, 堀井俊佑(早大), 小林学(湘南工科大), 後藤正幸,  
松嶋敏泰, 平澤茂一(早大)

## • 問題設定

- 異なる説明変数と共通の目的変数をもったユーザが協力して、**全ての説明変数を用いた回帰分析**(最小二乗法)を行う
- ただし、**お互い説明変数の内容は秘匿**したい



- 従来研究

- [Du 2004]乱数を用いたプロトコル
  - 2者間のみプロトコル
- [Snail 2004]Powellアルゴリズムを用いたプロトコル
  - 3者以上でないとなん安全性が担保できない

- 本研究

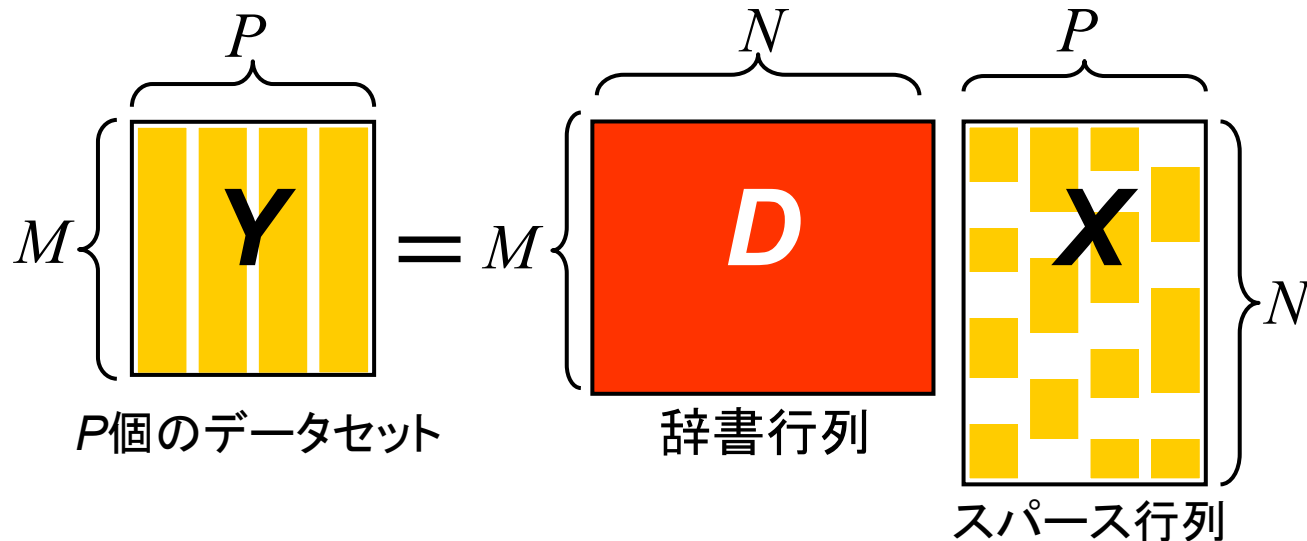
- 繰り返し型の最適化アルゴリズムを用いたプロトコル
- プロトコルの最適性を証明
- 2者以上で安全性を証明
  - それぞれの持つ説明変数が2個以上で安全
- 繰り返し回数の数値実験

# Dictionary Learningにおける サンプル複雑度の典型時解析

坂田綾香・樺島祥介(東工大)

## □ Dictionary Learning

- データがスパースに表現されるための基底(辞書)を学習する



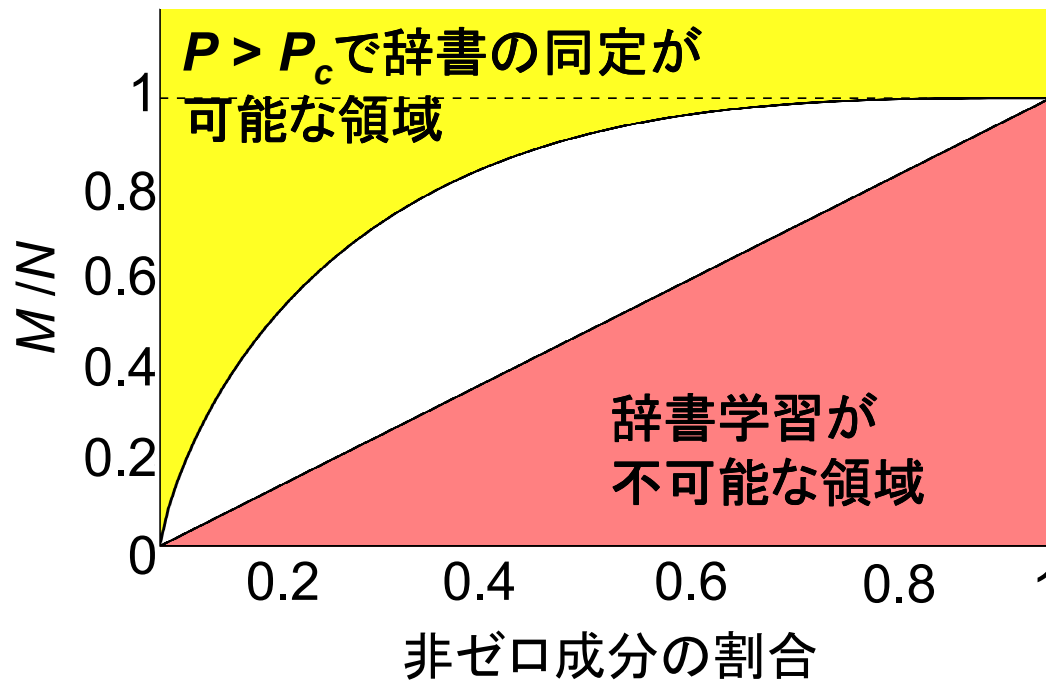
## □ 辞書を同定するために必要なサンプル数 $P_c$ は？

- $P_c$ をサンプル複雑度(Sample complexity)と呼ぶ。

# Dictionary Learningにおける サンプル複雑度の典型時解析

坂田綾香・樺島祥介(東工大)

- 統計力学的解析を通して  $P_c$  の典型時評価を行った結果
  - $M/N$  が十分大きい時、 $P > P_c = \gamma_F N$  で辞書が同定可能。

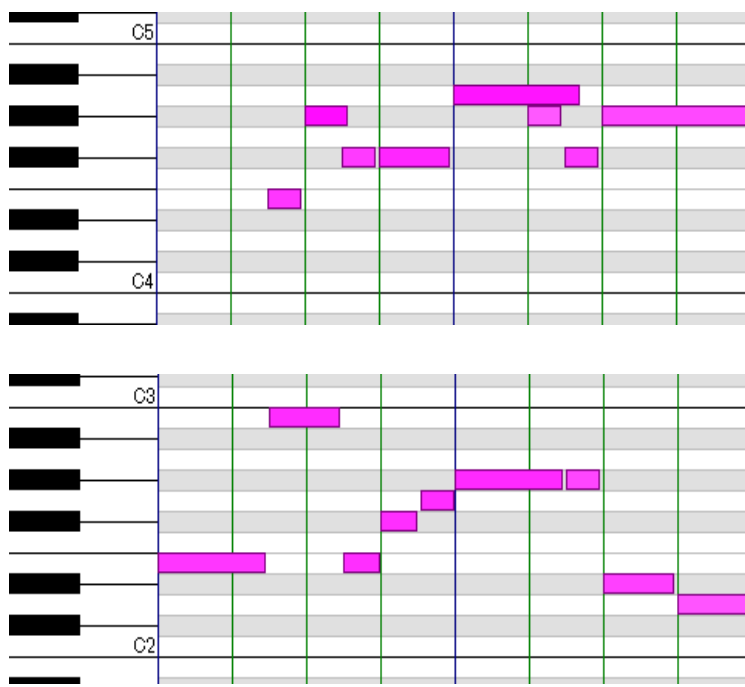


# 音楽音響信号解析のためのガンマ過程に基づく無限複合自己回帰モデル

吉井 和佳 後藤 真孝 (産総研)

## 音楽音響信号を構成するパーツとは何か？

音高 (基本周波数)



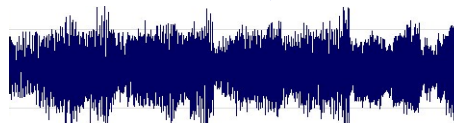
音色 (楽器種)



様々な  
楽器音を  
生成



重ね合わせ



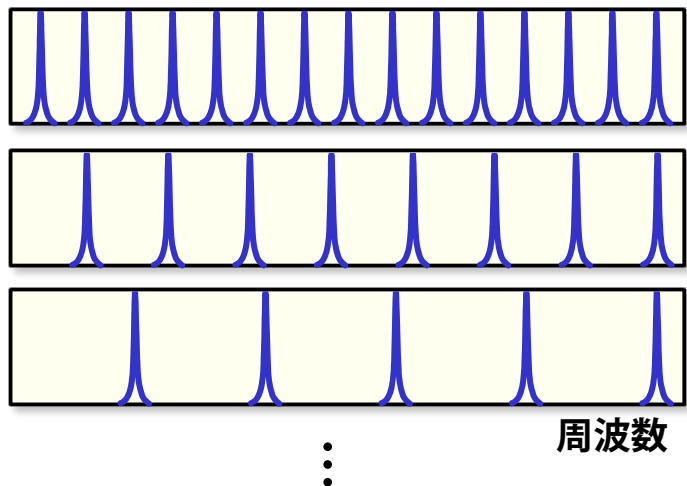
音楽音響信号

# 音楽音響信号解析のためのガンマ過程に基づく無限複合自己回帰モデル

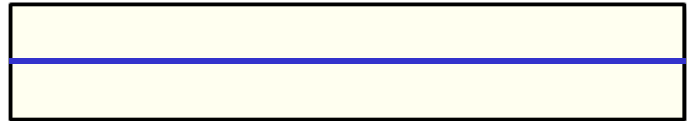
吉井 和佳 後藤 真孝 (産総研)

ノンパラベイズNMF!

スペクトル微細構造



白色雑音



$I$  ソース (音高)

$\infty$  に発散

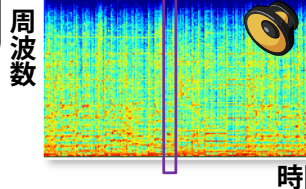
複合

$$\otimes$$

$IJ$  基底

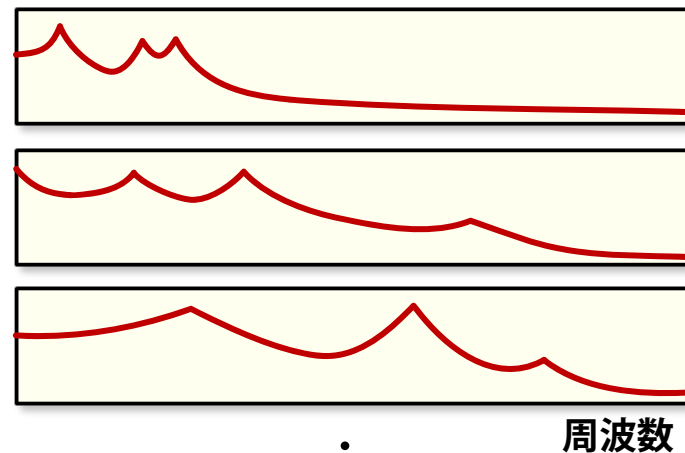
重みつき和

$$\oplus$$



各時刻における  
多重音スペクトル

スペクトル包絡構造



自己回帰系

$J$  フィルタ (音色)

$\infty$  に発散



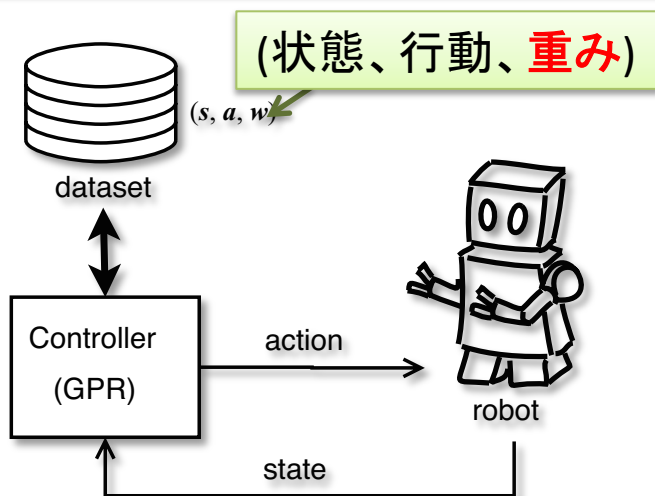
# Gaussian process regression を用いた 確率的方策に対する方策勾配法

中村 泰 (阪大)、石黒 浩 (阪大)

## 概要: ノンパラメトリック法を用いた強化学習

- 仮定: “上手くいった” 行動の抽出・再現  $\equiv$  動作学習
- 強化学習による “良い” 行動ルールの抽出

## 提案手法: GPR を用いた確率的方策



- 状態・行動ペアに対する**重み**付け
- 強化学習による**重み**の学習

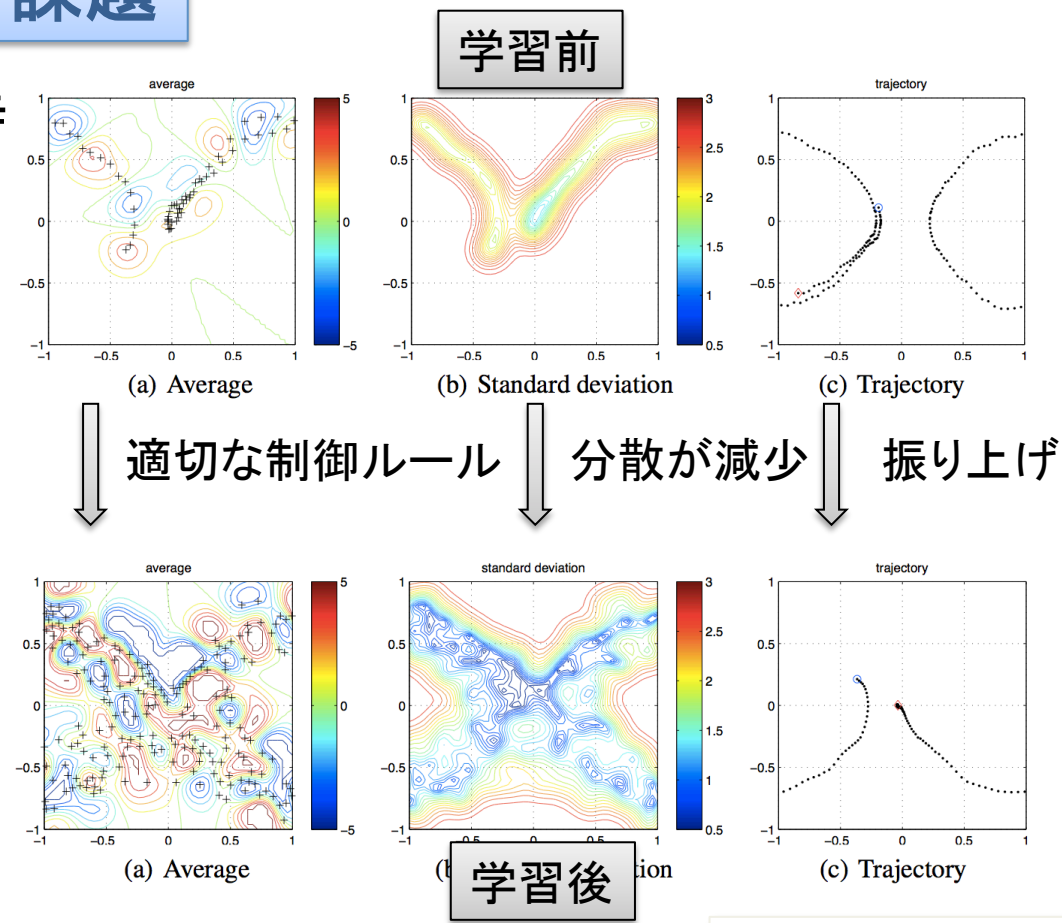
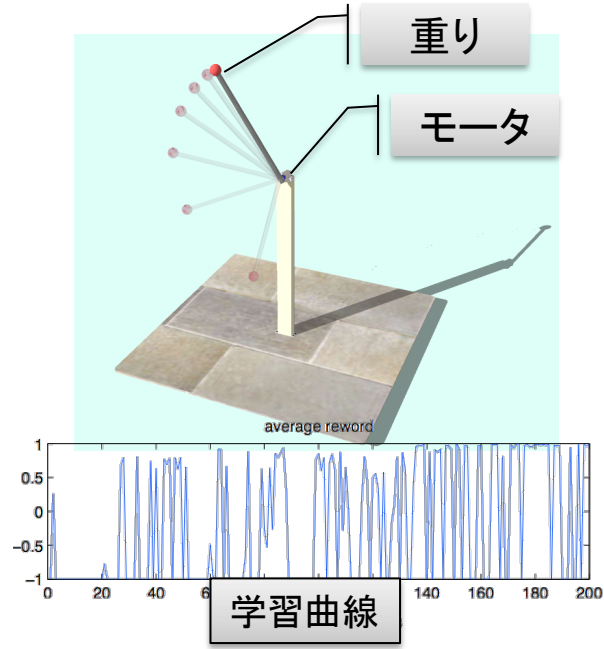
↓ ガウス過程回帰 (GPR)

- 平均: 近傍の “良いペア” が**支配的**
- 分散: 近傍の “良いペア” の密度  $\rightarrow$  分散

自動的に探索搾取

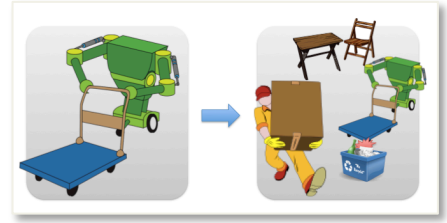
# 実験: 倒立振子の振り上げ課題

- 目的: 振り上げ + 倒立の保持



## まとめと今後の課題

- GPR による方策を用いた制御則を獲得可能
- 計算量が  $O(n^4)$  → 高速化、実問題への適用



学習の困難さは状態変数の次元ではなく、サンプルサイズに依存

# T-20 ベイジアンネットワークを用いた LDAの特徴選択

電気通信大学 大畑亮介, 植野真臣

## ● 背景と問題点

複数の任意の特徴を生成過程に取り入れたトピックモデルが注目されている。しかし、従来研究では特徴の選択方法は考慮されていない（特徴：文書における著者，出版社など）。

## ● 提案手法(1)

どの特徴を学習に使えば良いか？を決定する手法を提案する。

1. 周辺尤度による方法
2. トピックと特徴の相互情報量による方法

上記2手法を用いて実験を行ったが、十分な精度は得られなかった。

## ● 周辺尤度 (ML), 相互情報量 (MI) の問題点

MLの問題点：経験ベイズ法によるハイパーパラメータの学習は最適な特徴を選択できる保証がない。

MIの問題点：統計的一致性を持たない。特徴数が増加すると精度が悪くなる。閾値の設定が恣意的である。

## ● 提案手法 (2)

ハイパーパラメータを変数消去し、ベイジアンネットワークの構造スコアBDeuで評価する。

## ● 実験結果

データ数が十分にあるとき高い精度で特徴選択可能である。  
また、MLよりも計算効率が良い。

○ (真の特徴を選んだ数), + (生成に無関係な特徴を選んだ数)  
- (真の特徴を選ばなかった数)

Structure: 1		BDeu			BDeu1			MI			ML		
K	$N_d$	○	+	-	○	+	-	○	+	-	○	+	-
3(1)	5	63	0	37	22	0	78	89	0	11	60	42	2
	10	98	0	2	49	0	51	99	0	1	35	76	1
	20	100	0	0	64	0	36	100	0	0	32	83	0
	50	100	0	0	87	0	13	100	0	0	24	93	16
	100	87	13	0	84	0	16	100	0	0	35	84	6
Structure: 2		BDeu			BDeu1			MI			ML		
K	$N_d$	○	+	-	○	+	-	○	+	-	○	+	-
5(2)	5	41	0	87	1	0	193	77	0	27	69	31	3
	10	96	0	5	5	0	175	94	0	8	49	60	0
	20	100	0	0	14	0	138	97	0	3	30	86	1
	50	100	0	0	28	0	97	99	1	0	13	133	19
	100	97	3	0	48	0	71	100	0	0	17	114	2
Structure: 3		BDeu			BDeu1			MI			ML		
K	$N_d$	○	+	-	○	+	-	○	+	-	○	+	-
10(5)	5	0	0	41	0	0	500	0	0	458	NaN	NaN	NaN
	10	0	0	6	0	0	500	0	0	386	NaN	NaN	NaN
	20	54	0	0	0	0	500	27	0	159	NaN	NaN	NaN
	50	100	0	0	0	0	500	54	0	61	NaN	NaN	NaN
	100	100	0	0	0	0	500	57	0	52	NaN	NaN	NaN

# T-21: Online Large-margin Weight Learning for First-order Logic-based Abduction

Naoya Inoue, Kazeto Yamamoto, Yotaro Watanabe, Naoaki Okazaki, and Kentaro Inui  
Tohoku University

## Logic-based Abduction

\* **Abduction** is inference to the best explanation

Given:  $(O, B)$ , where:

$O$ : Observation:  $\{get\text{-}gun(John), go\text{-}to\text{-}store(John)\}$

$B$ : Background knowledge:

$(\forall x) hunt(x) \rightarrow get\text{-}gun(x)$

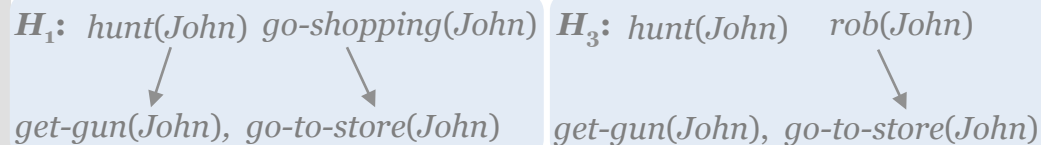
$(\forall x) go\text{-}shopping(x) \rightarrow go\text{-}to\text{-}store(x)$

$(\forall x) rob(x) \rightarrow get\text{-}gun(x)$

$(\forall x) rob(x) \rightarrow go\text{-}to\text{-}store(x)$

Find:  $H$ , where:

$H$ : The best ( $\equiv$  highest-score) explanation of  $O$  w.r.t.  $B$   
( $H \cup B \models O$ , and  $H \cup B$  is consistent)



$score(H_1) = 10.8$

$score(H_2) = \mathbf{13.5}$

$score(H_3) = 4.3$

\* Many applications:

plan recognition, natural language processing, etc.

## Existing approaches for tuning score function

\* Manual tuning

\* Probabilistic logic-based learning  
(e.g. Markov Logic Networks  
[Richardson & Domingos 06])

**Problem:** inference is not scalable;  
learning is even harder

**Scalable tuning algorithm is needed**

## Desiderata for algorithm

\* **Scalability:**

- computationally cheap

\* **Accurateness:**

- discriminately powerful

\* **Usability:**

- learn from partially observed dataset

# Proposed method

## Online large-margin training with latent variables

Training Examples (Observation + Explanation):

$$(O_1 = \{P, Q\}, \tau_1 = \{S\})$$

$$(O_2 = \{X, Y, Z\}, \tau_2 = \{R, T\})$$

$$(O_3 = \{P, Z\}, \tau_3 = \{S, Q\})$$

Background knowledge:

$$B = \{T \wedge S \rightarrow Z, \\ R \rightarrow X, \\ \dots\}$$

**Decoding:** extend efficient abductive reasoner (Inoue and Inui [12]'s Integer Linear Programming-based formulation)

**Goal of weight learning:**

**Feature-based linear scoring model:**

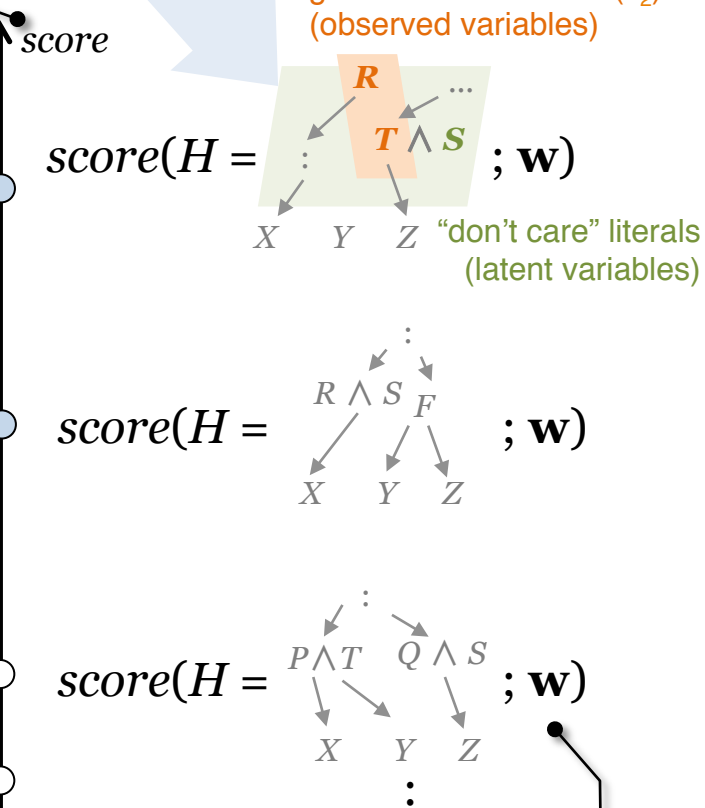
$$score(H; \mathbf{w}) = \mathbf{w} \cdot \Phi(H)$$

$\mathbf{w}$ : weight vector (learned) /  
 $\Phi$ : feature vector

Score of  $H$  that *includes* gold-standard literals

**Large-margin**

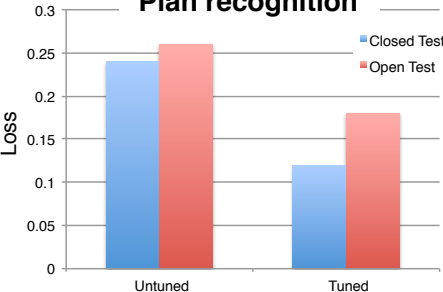
Score of  $H$  that *does not* include gold-standard literals



# Evaluation

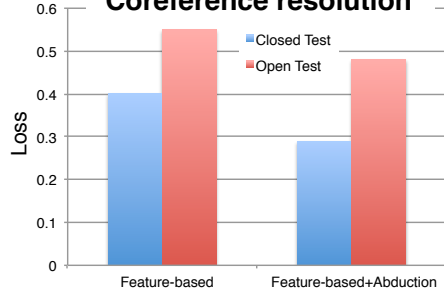
The machine learning-based abductive reasoning helps to reduce the loss in two applications

**Plan recognition**



Untuned: logical abduction  
Tuned: abduction w/ trained weights

**Coreference resolution**



Feature-based: pairwise clustering  
+Abduction: clustering w/ abduction

**Training:** Crammer et al. [06]'s online Passive Aggressive algorithm modified for handling latent variables

# ハッシュ関数を用いた Gaussian Process Regressionの高速化

岡留 有哉(阪大), 中村 泰(阪大), 石黒 浩(阪大)

## 概要

本研究ではLocality Sensitive Hashing (LSH)を用いたデータセット分割とProduct of Experts (PoEs)モデルを用いた確率モデルの統合を用いた高速なガウス過程回帰 (GPR)の計算法を提案する

## LSHを用いたGPRの高速化

以下の処理を行う事によりガウス過程回帰の出力を高速に計算

1. Locality Sensitive Hashing (LSH)によりデータセットを部分集合に分割
2. 入力点の所属する部分集合内の近傍サンプル以外のカーネル関数の値を0と近似
  - ⇒ 共分散行列を**ブロック対角**と近似できる
  - ⇒ 部分空間についてのみ計算を行えばいいため, 高速に計算可能
3. 出力の期待値と分散を計算

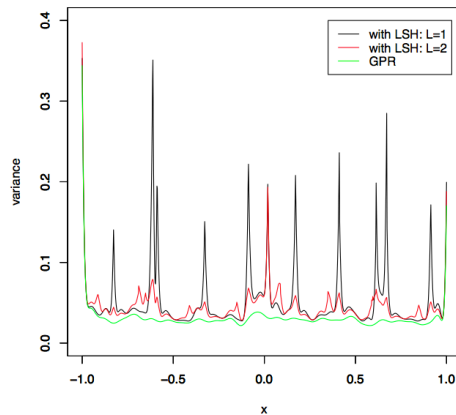
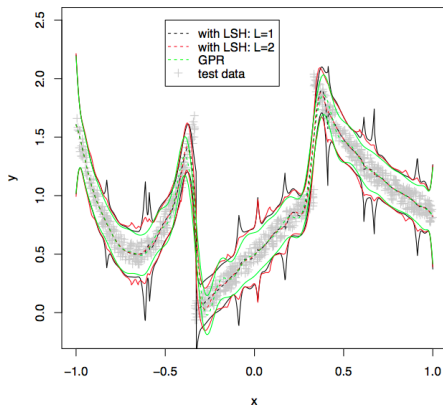


# PoEsモデルを用いた確率モデルの統合

LSHでは部分空間の境界付近において**有効なサンプルの減少**により近似精度が低下すると考えられる

- 異なる複数のハッシュキーを用いて複数パターンの部分空間に分割することで分割境界付近における近似精度の低下を抑制
- 複数の部分空間から出力が得られるため, 出力を**PoEsモデル**を用いて統合

## 関数近似課題



	逆行列	予測
GPR	7.1	96
GPR with LSH: $L = 1$	0.2	1.9
GPR with LSH: $L = 2$	0.4	3.7

予測: 計算時間が最大**1/50**

逆行列: 計算時間が最大**1/35**

ハッシュキーを増やすことで精度上昇

## まとめ

- LSHとPoEsモデルを用いることでガウス過程回帰の計算コストを削減
- 今後は実問題に適用



# T-21 重み付き最尤推定に基づく方策探索法

植野 剛<sup>1</sup>, 林 浩平<sup>2</sup>, 鷲尾 隆<sup>3,1</sup>, 河原 吉伸<sup>3,1</sup>

<sup>1</sup>科学技術振興機構

<sup>2</sup>日本学術振興会 特別研究員

<sup>3</sup>大阪大学 産業科学研究所

## 強化学習の種類

### ■ 価値関数規範:

方策を価値関数を通してモデル化する (例) SARSA, Q-Learning

### ■ 直接方策探索法

方策を直接パラメトリックモデルで書く (例) Policy Search, EM PS

## 既存の直接方策探索法の問題点

### ■ モデル選択

方策モデルの変更 → サンプル生成過程の変更

### ■ 統計的な性質はよく分かってない

## T-21 重み付き最尤推定に基づく強化学習

### 重み付き尤度

$$\ell(\mathbf{X}_1^n, \mathbf{U}_1^n; \theta) \equiv \prod_{i=1}^n \pi(\mathbf{u}_i | \mathbf{x}_i; \theta)^{\hat{Q}_i^\beta} \prod_{j=2}^n p(\mathbf{x}_j | \mathbf{x}_{j-1}, \mathbf{u}_{j-1}), \quad \hat{Q}_i^\beta \equiv \sum_{j'=i}^n \beta^{j'-i} r_{j'}$$

### Weighted Likelihood Policy Search (WLPS)

- 1 モデル  $M'$  を固定してサンプリングする
- 2 モデル  $M_k$  を1つ選択して、重み付き最尤推定する
- 3 次のモデル選択基準を計算する

$$I(\hat{\theta}_n; M_k) = \sum_{i=1}^n \hat{Q}_i^\beta \ln \pi(\mathbf{u}_i | \mathbf{x}_i) - \frac{1}{2} (M_k \text{の次元}) \ln n$$

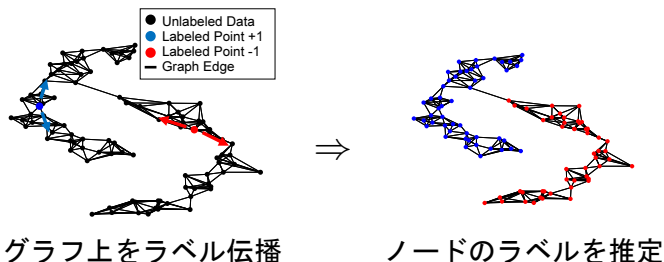
- 4  $I(\hat{\theta}_n; M_k)$  を最大とするモデルを選択して、(1)に戻る

$\beta \rightarrow 1^-$  のとき、WLPSの各反復は漸近的に期待報酬の単調な増加を保証する

# ラベル伝播アルゴリズムにおける 複数グラフのスパース結合法

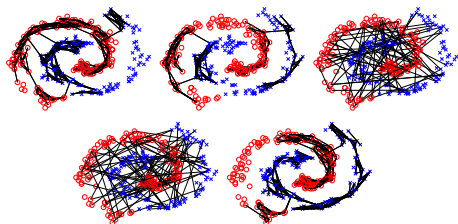
烏山 昌幸, 馬見塚 拓 (京都大学)

- ラベル伝播アルゴリズム: グラフに基づく半教師分類法



- 実践では**グラフが複数あるかも**
  - タンパク質機能予測: 遺伝子発現量, 相互作用ネットワーク, シーケンスデータなどから (時にはそれぞれ複数の) グラフが得られる
- どのグラフが重要か (分類に関係するか) は事前には不明

## 複数グラフをスパースに結合してラベルを伝播するアルゴリズムの提案



人工例: ノイズの混ざったグラフ群



不要なグラフを削除して結合

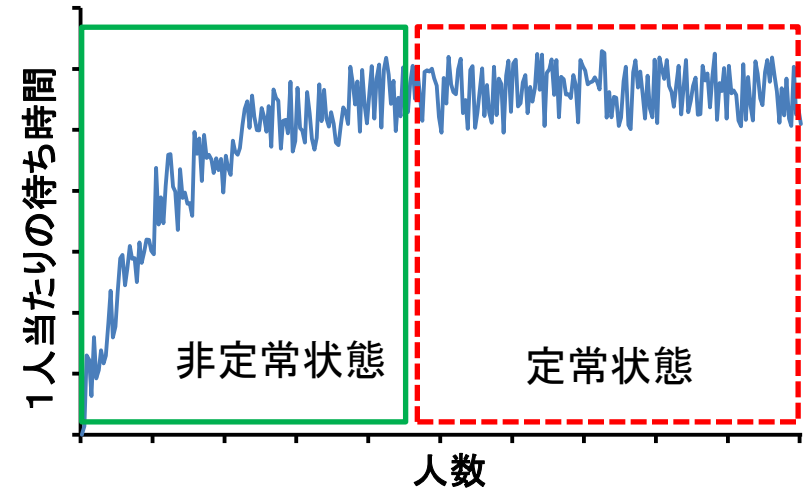
### [提案法の特徴]

- ラベルはグラフ上を滑らかに変化すると仮定のもと、不要なグラフを完全に排除 (スパース性)
- Elastic net に似たグルーピング効果 (似たグラフは似た重みを与えられる)

これにより  
予測精度 と 解釈性 が向上

## 研究背景

既存の待ち行列理論では, 客数を無限とし, システムにおける定常状態で近似した解析を行っている.



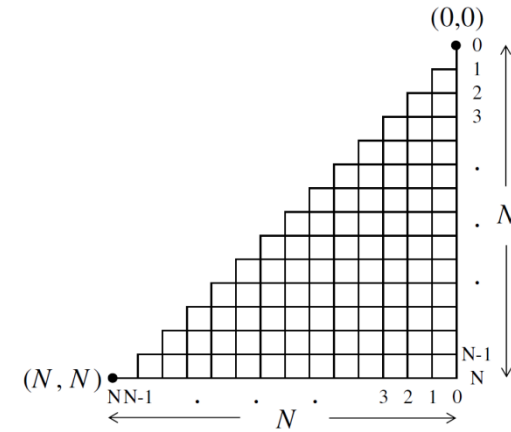
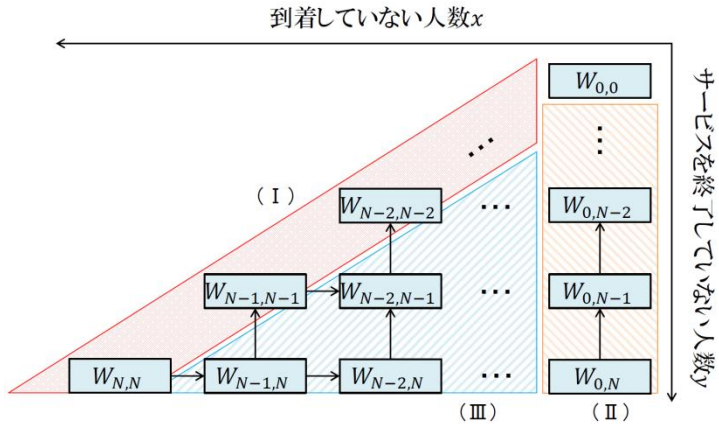
現実の待ち行列システムでは, この客数を無限とした近似を行うことで大きな誤差が生じてしまう場合がある.

## 研究目的

M/M/1 待ち行列を対象とし, 客数を有限としてモデル化し, 漸化式による解と, 閉形式による解を導出することによって非定常状態に対して正確に解析できるような方法を提案する.

# 提案手法

指数分布の無記憶性を利用し有限客数でモデル化する



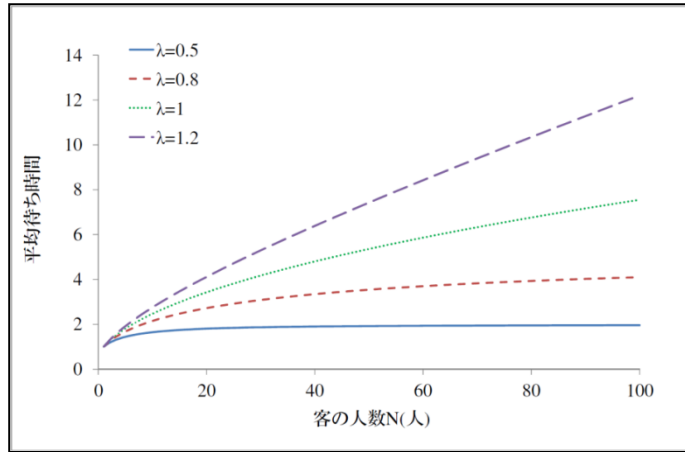
$W_{x,y}$ : 全員の残りの平均待ち時間の総和  
 状態遷移確率に注目して場合分けをし、  
**漸化式**を立てることによって一人当たりの  
 平均待ち時間  $\frac{1}{N} W_{N,N}$  を再帰的に求める。

状態  $(N,N)$  から  $(0,0)$  までの全ての最短経路について、各状態を訪れる期待値と待ち時間の増加量の期待値を合計して**閉形式**によって  $W_{N,N}$  を求める。

$$W_{x,y} \equiv \begin{cases} 0 & \text{if } x = y = 0, \\ W_{x-1,y} & \text{if } x = y \neq 0, \\ W_{x,y-1} + t_2 y & \text{if } x = 0, y \neq 0, \\ p_1 W_{x-1,y} + p_2 W_{x,y-1} + t_3 (y-x) & \text{if } x \neq y, x \neq 0. \end{cases}$$

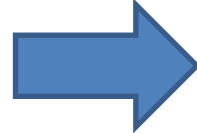
$$W_{x,y} = \sum_{i=1}^x \left( \sum_{j=1}^{i-1} \sum_{k=0}^{x+1-i} R_{x,y,j,i}^{(1)} W_{j,i}^{(1)} P_{x,y,j,i,k} + \sum_{k=0}^{x+1-i} R_{x,y,1,i,k}^{(1)} W_i^{(2)} P_{x,y,0,i,k} \right) + \sum_{i=x+1}^y \left( \sum_{j=1}^x R_{x,y,j,i}^{(2)} W_{j,i}^{(1)} P_{x,y,j,i,0} + R_{x,y,1,i}^{(2)} W_i^{(2)} P_{x,y,0,i,0} \right)$$

## 結果

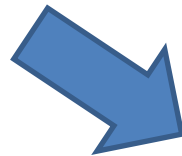


$\mu = 1$ で固定し、 $\lambda$ を変化させたときの  
一人当たりの平均待ち時間

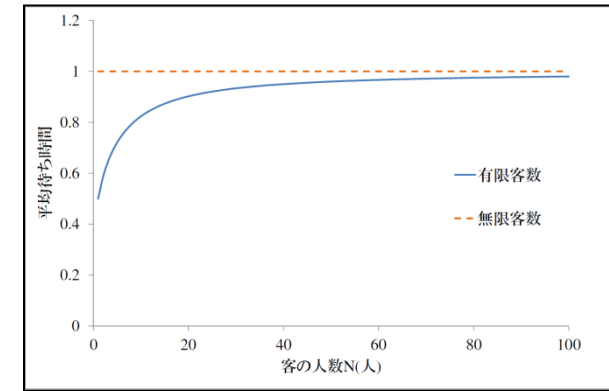
$\lambda = 0.5$ のとき



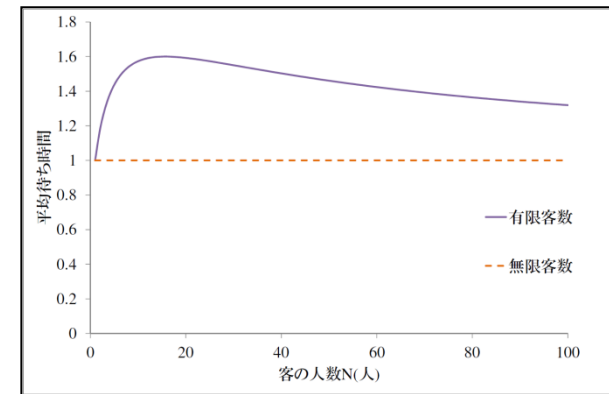
$\lambda = 1.2$ のとき



無限客数の結果を1として比較



既存のM/M/1待ち行列の解析結果との比較



サービス時間・到着間隔一定の  
D/D/1待ち行列との比較

システムの非定常状態を解析することが  
でき、定常状態を持たないようなシステム  
に対しても解析することができるように  
なった.

人数が少ないほど無限客数での解析より  
正確にシステムの評価を行うことができる.

# T-27 類似検索における秘密情報漏えいの評価 及び差分プライバシーの保証

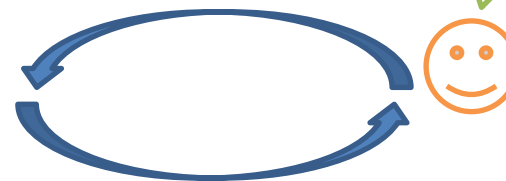
荒井ひろみ(理研) 佐久間淳(筑波大)

個々のレコードが個々の秘密情報であるデータベースを使って、  
クエリに対してレコードのRankingをプライバシーを保護して返したい  
(SNSの友人検索, 類似患者検索などを想定)

Aさん, Bさん, Cさん,  
Dさんの順に似てます

似たような症例の人  
を探したい

氏名	病歴	居住地	職業
Aさん	なし	〇〇市	会社員
Bさん	〇〇病	△△市	自営業
Cさん	〇〇疾患	〇×村	農業
Dさん	〇〇症	△〇区	会社員

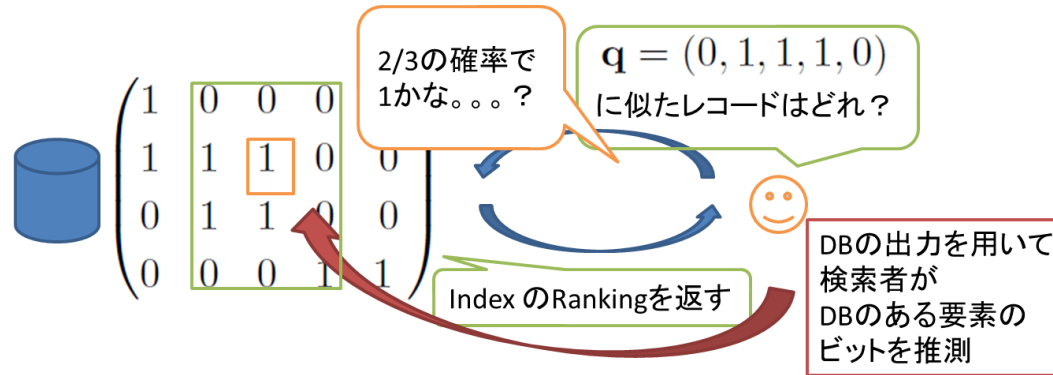


Question: Rankingからレコード  
の情報は洩れるのか? その漏  
えいを防ぐことはできるか



## 本研究の貢献:

### 類似検索のプライバシーモデル/プライバシー漏えい基準の定義



### プライバシー漏えいの理論解析

複数のクエリに回答するとき, どの順位のレコードにも漏えいの可能性がある

### 数値実験による既存のプライバシー保護方法の検証

差分プライバシーを満たすランダム化ではutilityとプライバシー保護がトレードオフ

## まとめ

- 理論解析の結果, レコードの順位を不確定にする必要があった
- 既存手法にみられるトレードオフを解消するランダム化法が必要

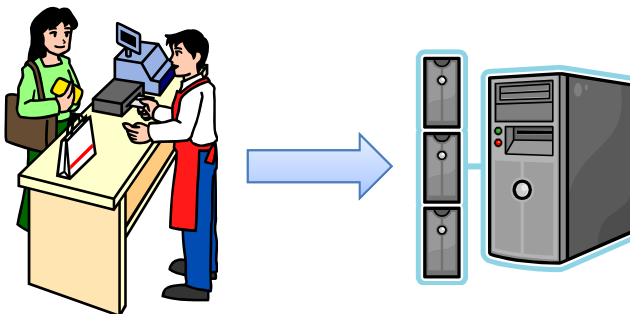
T-28

# 潜在クラス型階層ベイズプロビットモデルによる 大規模購買行動モデル

石垣司、照井伸彦(東北大学)、佐藤忠彦(筑波大学)

## 問題設定

- 購買行動に関するビッグデータの有効活用
- 大規模顧客へのパーソナライゼーション



- 小売業にけるビッグデータの蓄積
- 蓄積のみでは単なる負債

## 提案方法

- 確率的効用モデル ⇒ ビッグデータへの対応が困難
- 帰納的次元圧縮(トピックモデル等) ⇒ 顧客特性をモデル化できない



### 両者の統合モデル

- 大規模顧客・商品のための効用理論に基づいた購買行動モデル
- マーケティング変数を介した購買行動シミュレーションが可能

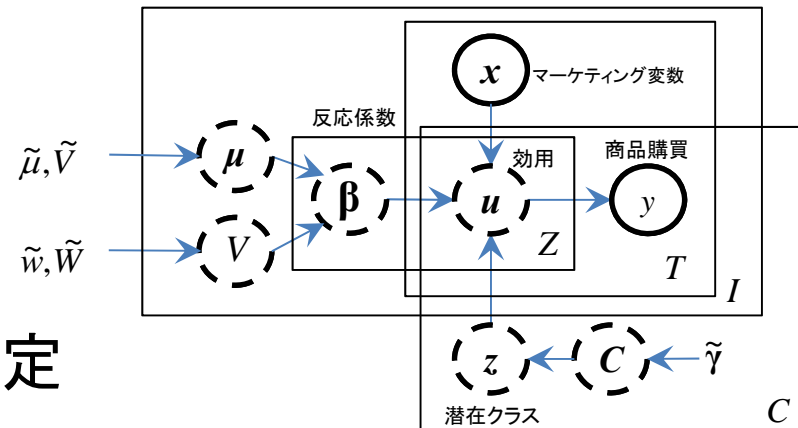
T-28

# 潜在クラス型階層ベイズプロビットモデルによる 大規模購買行動モデル

石垣司、照井伸彦(東北大学)、佐藤忠彦(筑波大学)

## 提案モデル

- 階層ベイズ2項プロビットモデルと潜在クラスモデルの統合モデル
- 個人の各商品に対するマーケティング変数への反応係数を推定



潜在クラス型階層ベイズ2項プロビットモデル

## 数値実験

- スーパーマーケットのID-POSデータ
- マーケティング変数
  - ⇒ 割引、エンド陳列、チラシ掲載の有無
- 予測性能はLDAと比べ良好

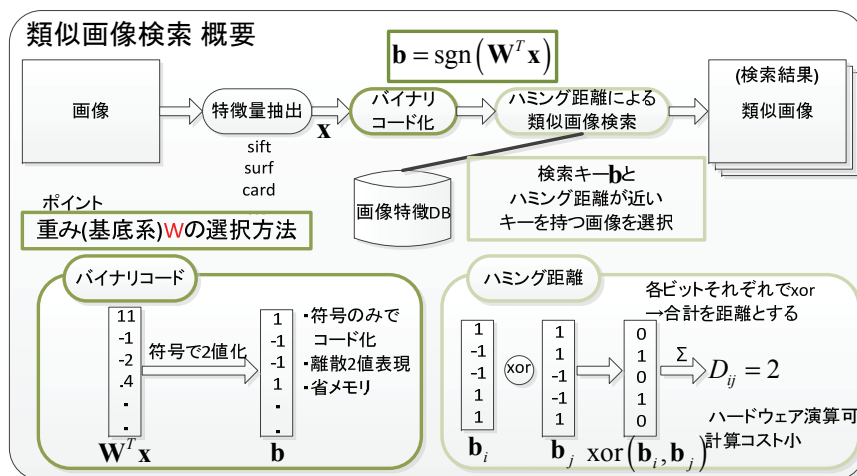
	LDA (Topic model)	提案モデル (潜在クラス)	提案モデル (個人)
RMSE	458.7	239.8	231.8
Hit rate	0.7%	3.5%	5.0%

購買商品の予測性能

# データ分布の2次モーメントのみにもとづく バイナリコーディング手法

鈴木 幸一郎, 安倍 満, 佐藤 育郎 {ksuzuki, manbai, istao} @d-itlab.co.jp  
(株) デンソーアイティラボラトリ

検討目的: 類似画像検索の性能改善



- ・ 類似画像検索: 高次元画像特徴をキーとし, 類似の画像をデータベースから検索
- ・ 高次元画像特徴を低次元の2値符号 (バイナリコード)に変換  
→ データ圧縮可能. 保持, 通信コスト低.
- ・ ハミング距離を特徴間距離とする  
→ ハードウェア演算可、計算コスト小

検討課題: バイナリコードは高次元特徴を線形変換して求める  $b = \text{sgn}(W^T x)$   
2次モーメント(共分散行列)のみを用いて学習  
→ その場合の最適な変換行列  $W$ を求める( $W$ は直交基底系とする)

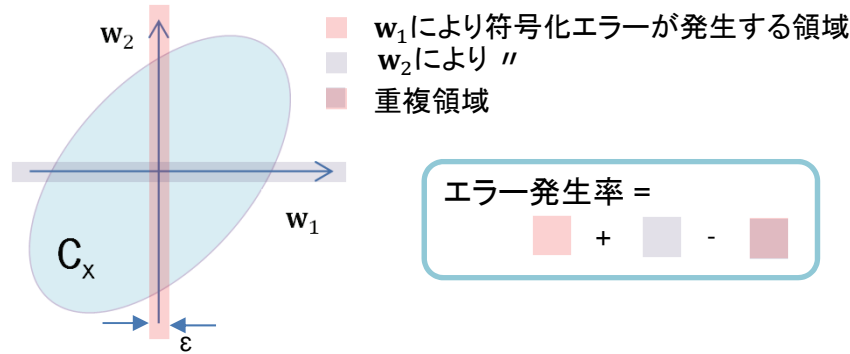
基本戦略: 高次元特徴の近傍は同じバイナリコードに割り当てる. つまり,

$$\text{符号化エラー: } J = \sum_i \|\text{sgn}(W^T x_i) - \text{sgn}(W^T x_{i_c})\|_2^2 \text{ を最小とする } W \text{ を求める}$$

問題:  $\text{sgn}$  関数の取り扱いが困難

アプローチ: 共分散行列が既知なガウス分布に従うとして, エラー率を直接算出.  
エラー率を最小とする  $W$  を求める

# 符号化エラー率 2次元で



B次元に拡張:

$$P_W = \sum_i P_{w_i} - \left( \sum_{ij} P_{w_{ij}} - \left( \sum_{ijk} P_{w_{ijk}} \dots \right) \right)$$

近傍  $\epsilon \rightarrow 0$  の極限で近似:

$$P_W \approx \epsilon \sum P_i - \left( \epsilon^2 \sum P_{i,j} - \left( \epsilon^3 \sum P_{i,j,k} \dots \right) \right)$$

$$\approx \epsilon \sum P_i$$

最小となるのは、各  $P_i$  が全て等しいとき  $P_i = \text{const} / (\mathbf{w}_i^T \mathbf{C}_x \mathbf{w}_i)^{1/2}$   
 → 各基底に対するパワーが等しい

## アルゴリズム

1本ずつ正規直交基底ベクトルを算出する

$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$  画像特徴データ、N次元M本

$\mathbf{C}_X = E \left[ (\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T \right] = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$  データ相関行列推定、展開

信号・雑音分離、LはBCのbit数

$\mathbf{C}_s = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^T, \mathbf{U}_s = [\mathbf{u}_1, \dots, \mathbf{u}_L], \mathbf{\Lambda}_s = \text{diag}(\lambda_1, \dots, \lambda_L)$

$\bar{\lambda} = \sum^K \lambda_i / L$  固有値平均算出

Loop:  $l = 1, \dots, L$  に対し、 $\mathbf{v}_l = \mathbf{U}_s \mathbf{c}_l$  として  $l$  番目の正規直交基底を求める

$\mathbf{c}_l^T \mathbf{\Lambda}_s \mathbf{c}_l = \bar{\lambda}, \mathbf{c}_l^T \mathbf{c}_l = 1$  の両等式を満たすベクトル  $\mathbf{c}$  を算出

今回は、 $\mathbf{c} = \mathbf{1} / \sqrt{L-l+1}$  とした

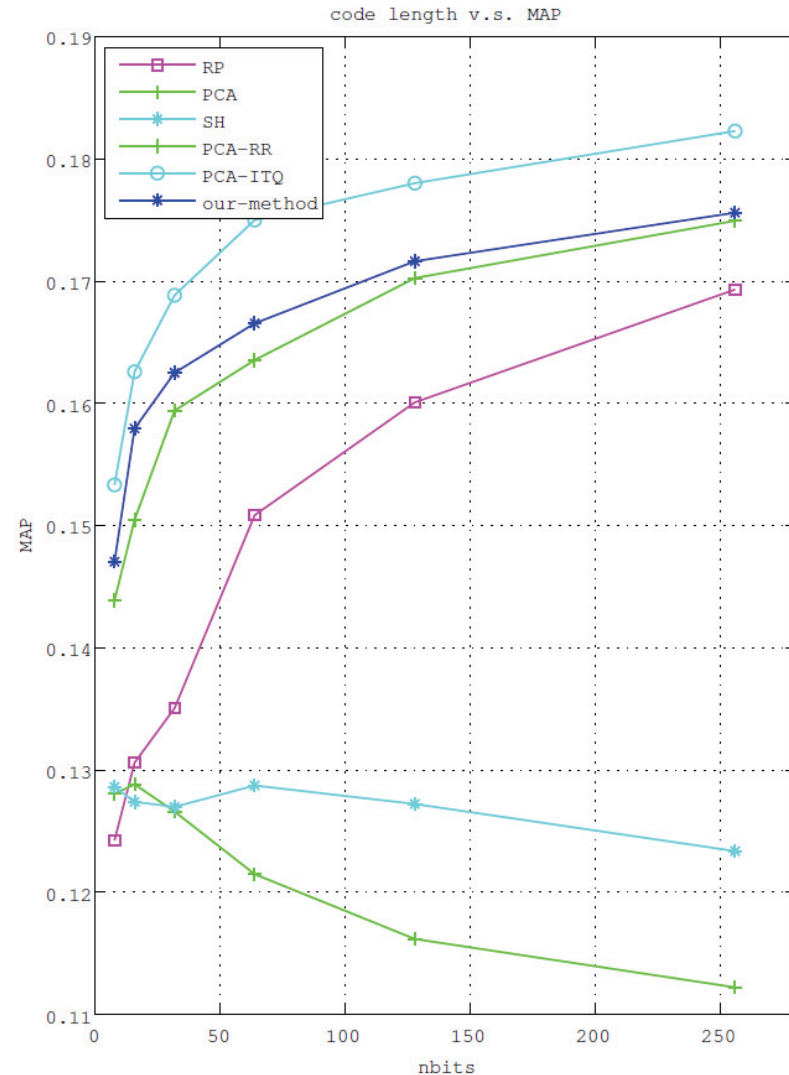
$\mathbf{C}_X (\mathbf{I} - \mathbf{w}\mathbf{w}^T) \mathbf{C}_s (\mathbf{I} - \mathbf{w}\mathbf{w}^T)$ ,

$\mathbf{C}_s = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^T, \mathbf{U}_s = [\mathbf{u}_1, \dots, \mathbf{u}_{L-l+1}], \mathbf{\Lambda}_s = \text{diag}(\lambda_1, \dots, \lambda_{L-l+1})$

Loopに戻る

$\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_L]$  として重みを出カ

# CIFA10データセット・GIST特徴を用いた性能評価



2次モーメントのみにもとづく手法中  
 最もよい性能を示した  
 (PCA-ITQはデータそのものも利用)

# T-30 ユーザのステレオタイピングに基づく推薦

納 竜也<sup>+</sup>, 佐久間 淳<sup>++</sup>

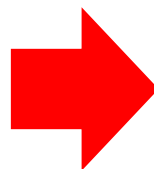
<sup>+</sup>, <sup>++</sup>筑波大学システム情報工学研究科      <sup>++</sup>科学技術振興機構

- 評価値行列を利用する推薦アルゴリズム
  - 同じ属性を持つユーザ → 評価が似ている
- ユーザのステレオタイプ
  - 任意のユーザ属性の組み合わせ
  - 複数のユーザを1つのステレオタイプで表現
    - 評価値行列の疎な割合を減らせる
  - 評価値行列と組み合わせで数値ベクトルに変換
    - 評価値行列を用いる推薦アルゴリズムへの導入が容易

# T-30 ユーザのステレオタイピングに基づく推薦

User-Item

属性	User ID	Item ID				
		1	2	3	4	5
男性, 20代, 学生	1	2	3	-	-	5
女性, 20代, 会社員	2	-	5	-	4	-
女性, 20代, 会社員	3	3	-	-	-	1
女性, 20代, 会社員	4	2	-	4	-	-
男性, 20代, 学生	5	-	-	3	-	4

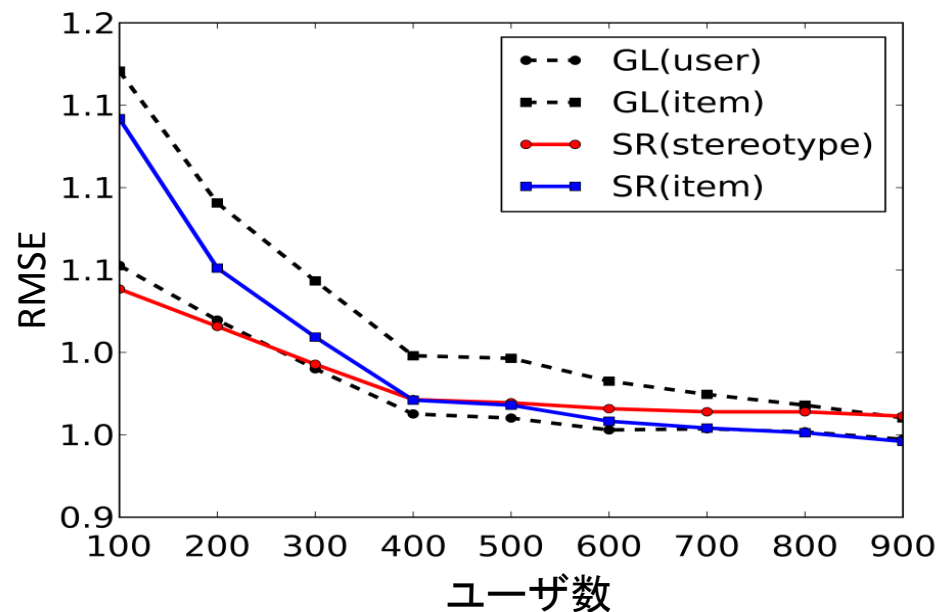


Stereotype-Item

ステレオタイプ	Item ID				
	1	2	3	4	5
男性, 20代, 学生	2	3	3	-	4.5
女性, 20代, 会社員	2.5	5	3.5	4	2.5

- ・欠損値の割合が減少
- ・ステレオタイプによる評価値行列なので推薦アルゴリズムへ導入が容易

- データ: MovieLens(100k)
- 類似度行列を作成して予測
- ユーザ数の変化に対するRMSEを測定
- GL:Group Lens\*の手法, SR:提案手法



P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews" Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94, pp.175-186