

教師あり順序付け 手法の比較実験

神鳶敏弘(産総研) 賀沢秀人(NTT) 赤穂昭太郎(産総研)

<http://www.kamishima.net/>

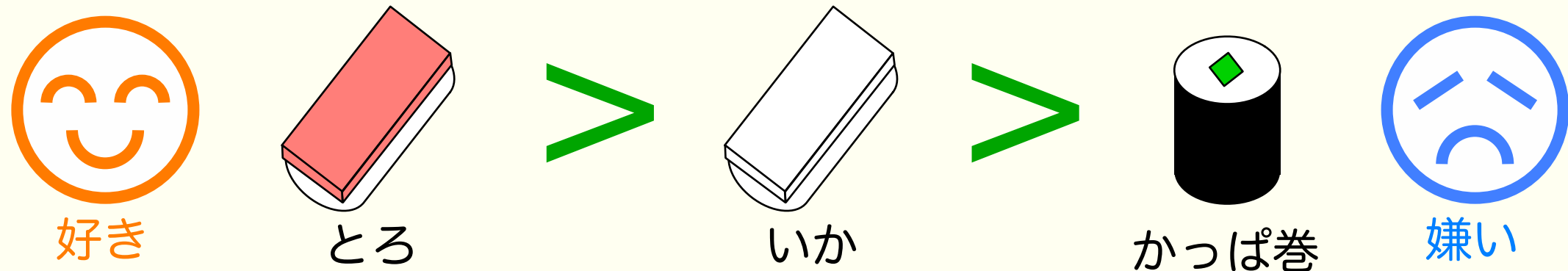
2005年情報論的学習理論ワークショップ(IBIS2005)

概要

訓練事例として与えられた順序とできるだけ一致する
ように対象を整列する規則の学習問題

順序：何らかの基準で対象を整列したもの

例：Aさんが好き[基準]な順に寿司[対象]を整列



“いか” より “とろ” が好き
しかし、どれくらい好きかは不明

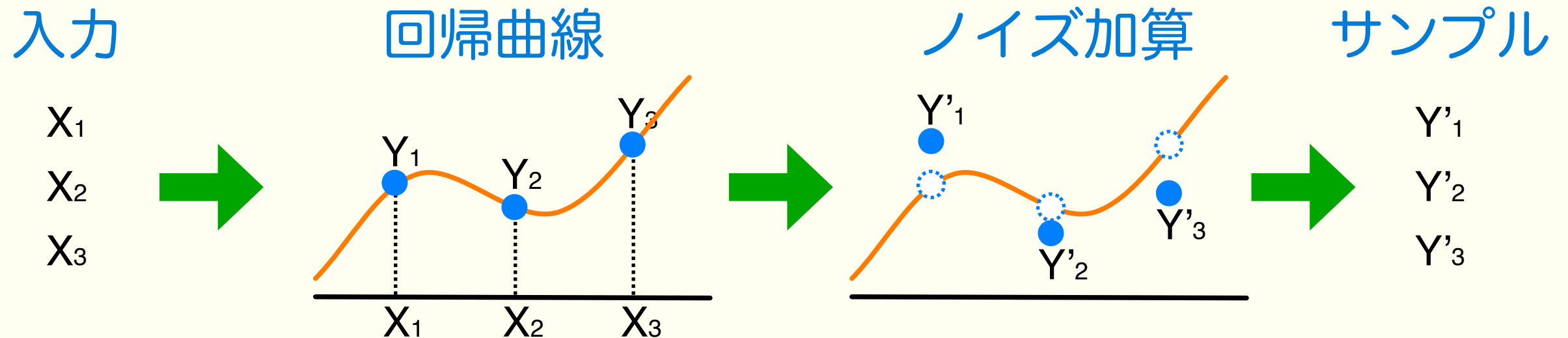
順序を扱う他の問題との関連について議論

5種類の手法について比較実験

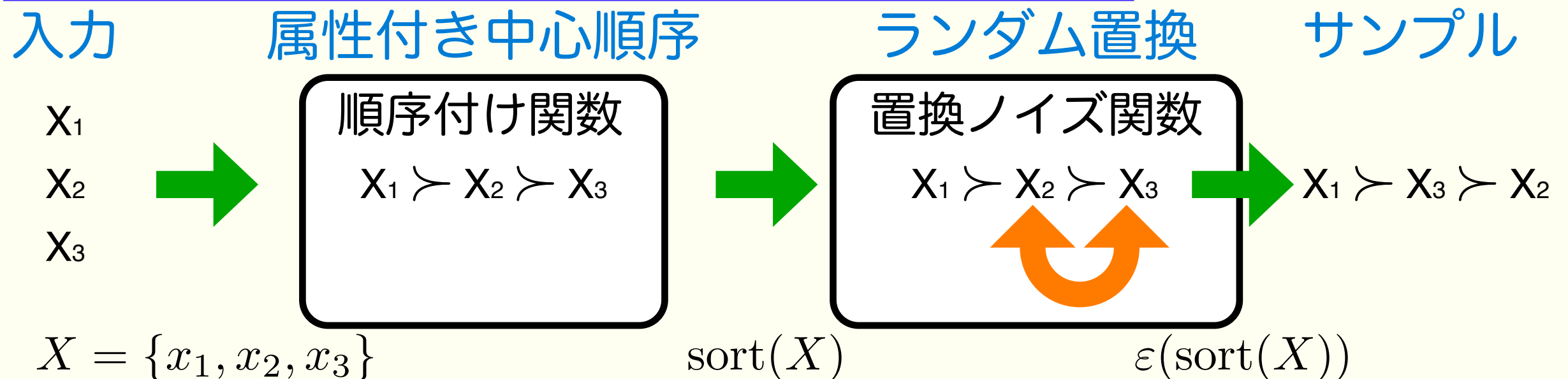
教師あり順序付け

従属変数が順序である回帰問題

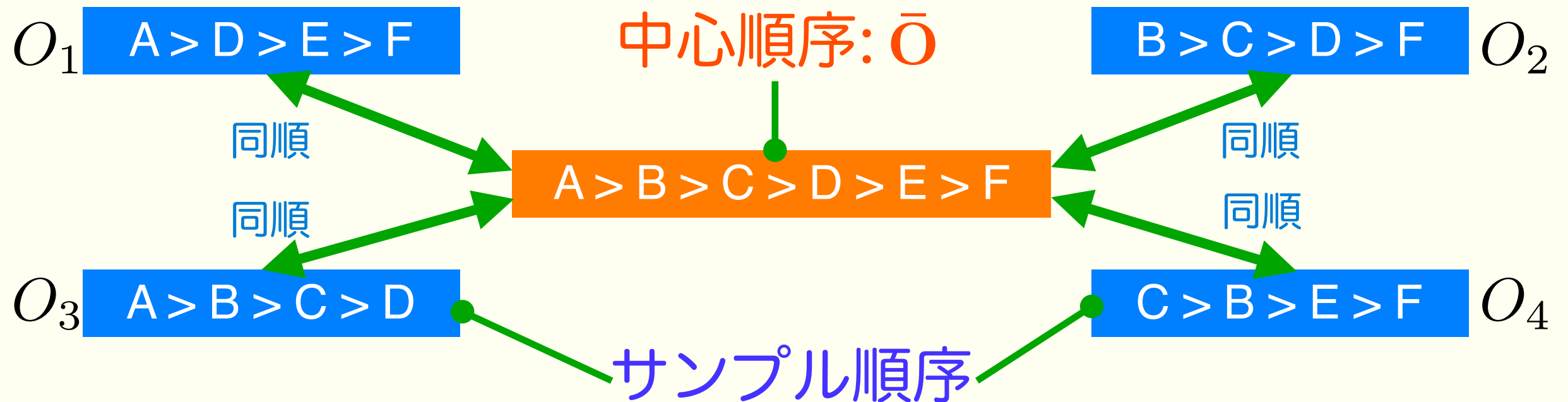
回帰分析のサンプル生成モデル



教師あり順序付けのサンプル生成モデル



中心順序



サンプル順序と平均的に同順

$$\bar{O} = \arg \min_O \sum_{O_i \in S} d(O, O_i)$$

	対象の記述	同順性の考慮対象
中心順序	一意な識別子	与えられたサンプルのみ
属性付き中心順序	属性ベクトル	未知のサンプルも含む

順序回帰

従属変数が順序カテゴリである回帰問題

教師あり順序付けは順序回帰よりより一般的な問題

順序カテゴリ: カテゴリ変数と同様に有限個の値のいずれかをとるが、さらに値に順序関係がある

順序と順序カテゴリの相違点

- ① **順序カテゴリ** 絶対的な情報を部分的に含む
例: 上-中-下で上なら絶対的に上位と分かる
順序 純粹に相対的な情報しか含まない
- ② **順序カテゴリ** 区別可能な順序の段階は有限
例: 上-中-下 なら3段階しか区別できない
順序 何段階でも順序関係を区別可能

順序の確率モデル (1)

Thurstone型

各対象に実数値のスコアを与え, そのスコア順に整列

Thurstoneの比較判断の法則

スコア関数が共通分散の正規分布: $\text{score}(x_i) = \mu_i + N(0, \sigma)$

Gumbel分布だと多段階型モデルと一致

Gumbel分布のCDF: $\exp(-\exp(-a(x_i - \mu_i)))$

一対比較型

一対の対象の一方が上位になる確率を求め次式を最大化する順序

$$\Pr[O] = \prod_{x_i \succ x_j \in \text{Pair}(O)} \Pr[x_i \succ x_j]$$

Babington Smithモデル $\Pr[x_i \succ x_j] = p_{ij}$ (一対比較の飽和モデル)

Bradley-Terry/Mallowsモデル $\Pr[x_i \succ x_j] = \frac{v_i}{v_i + v_j}$

Mallowsモデル $\Pr[x_i \succ x_j] = \frac{\theta^{i-j} \phi^{-1}}{\theta^{i-j} \phi^{-1} + \theta^{j-i} \phi}$
(距離ベース型モデルともみなせるモデル)

順序の確率モデル (2)

距離ベース型

モード順序 O_0 からの距離によって生起確率が決まるモデル

$$\Pr[O] = C(\lambda) \exp(-\lambda d(O, O_0))$$

Mallows ϕ モデル 距離 d をKendall距離としたモデル
一対比較Mallowsモデルで $\theta=1$ としたのと同価

Mallows θ モデル 距離 d をSpearman距離としたモデル
一対比較Mallowsモデルで $\phi=1$ としたのと同価

多段階型

順序の上位から順に決めていくモデル (Plackett-Luceモデル)

N 個の対象最上位の対象が x_{w1} になる確率は $\frac{v_{w1}}{v_1 + \dots + v_N}$

残りの個 $N-1$ 個の中で最上位の対象が x_{w2} ならその確率は $\frac{v_{w2}}{v_1 + \dots + v_{w1-1} + v_{w1+1} + \dots + v_N}$

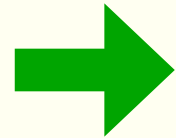
以下同様に N 位までを順に決定

絶対順序付け/相対順序付け

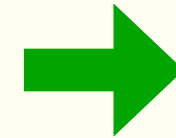
順序付け関数の性質:

もし対象集合を

$\{A, B, C\}$



絶対順序付け関数

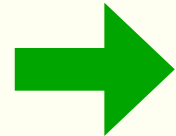


$A > B > C$

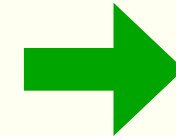
のように整列したなら

たとえば対象Cが消えて
対象Dが加わっても

$\{A, B, D\}$



絶対順序付け関数



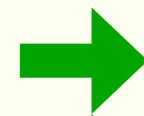
AはBより
必ず上位

つまり $D > A > B$, $A > D > B$, $A > B > D$ のいずれかの順序しか出力しない

文書検索などでは、文書の適合度は対象ごとに独立に決まるので絶対順序付け関数がの望ましい

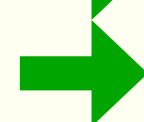
順序の確率モデル:

一対比較型



相対順序付け関数

その他



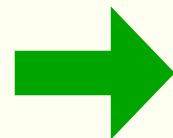
絶対順序付け関数

Cohenの方法 [Cohen 99]

学習: 属性値から対象対の順序関係を推定する確率関数

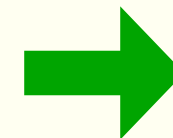
訓練用の順序

$A \succ B \succ C$
 $D \succ E \succ B \succ C$
 $A \succ D \succ C$



一対比較に分解

$A \succ B, A \succ B, B \succ C$
 $D \succ E, D \succ B, D \succ C, \dots$
 $A \succ D, A \succ C, D \succ C$



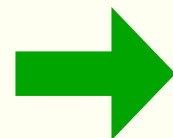
確率関数の学習

$\Pr[x \succ y | x, y]$

整列: 未整列の対象を評価関数に基づき整列

候補順序

$A \succ B \succ C$



一対比較に分解

$\text{Pair} = \{A \succ B, A \succ B, B \succ C\}$

目的関数: $\sum_{x_i \succ x_j \in \text{Pair}} \Pr[x_i \succ x_j]$

候補順序中で目的関数を最大化するものを欲張り探索

RankBoost [Freund 98]

学習: 順序付けのための弱学習器を結合することで, Thurstonianスコア関数を獲得するBoosting

弱学習器: 全対象を含む順序中でのある対象の位置についての部分情報を与える関数

例: 属性値 j の値そのもの, 属性値 j がしきい値 θ 以上なら 1

順位付け損失: 推定した順序とサンプル中の順序の不一致を, 対象対に割り当てた重みで重み付けしたもの

$$rloss_D(H) = \sum_{x_0, x_1} D_t(x_0, x_1) I[H(x_1) \leq H(x_0)]$$

Rankboost: 弱学習器の重み付き線形和 $H(x)$ を学習

$t+1$ 個目の重みを学習するときには, t 回目のときにエラーがあった対の $D_t(x_0, x_1)$ を大きくする

整列: Thurstonianスコアの順に未整列の対象を整列

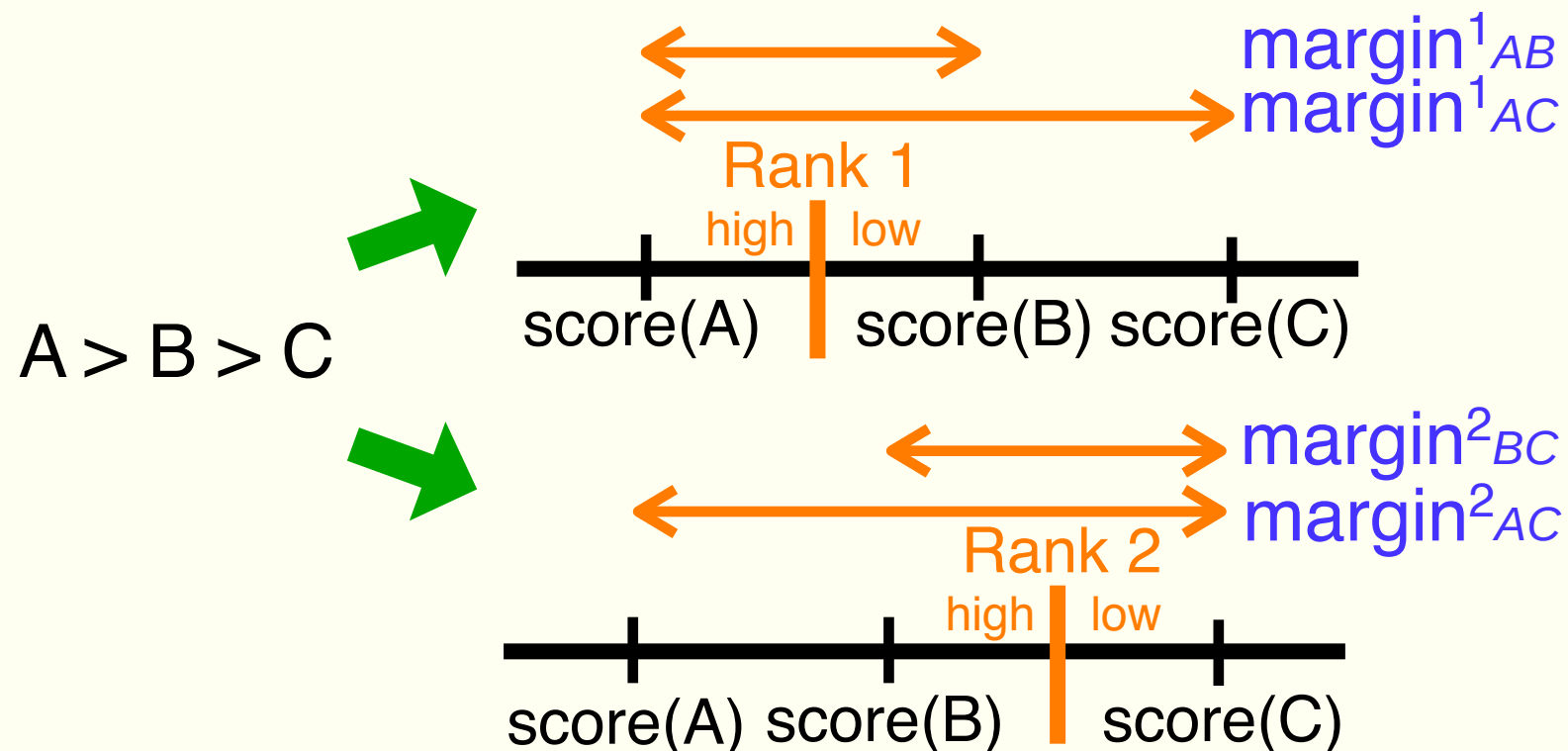
Order SVM [賀沢 03]

学習: ある順序中で上位と下位に最も分離するような
Thurstonianスコア関数を見つける
SVMの一種でカーネルの導入も可能

標本順序

Thurstonianスコアとマージン

目的関数



最大化:

$$\sum_j \sum_{X,Y} \text{margin}^j_{XY}$$

整列: Thurstonianスコアの順に未整列の対象を整列

Support Vector 順序回帰

[Herbrich 98]

学習: 一対の対象を上位と下位に最も分離するような
Thurstonianスコア関数を見つける

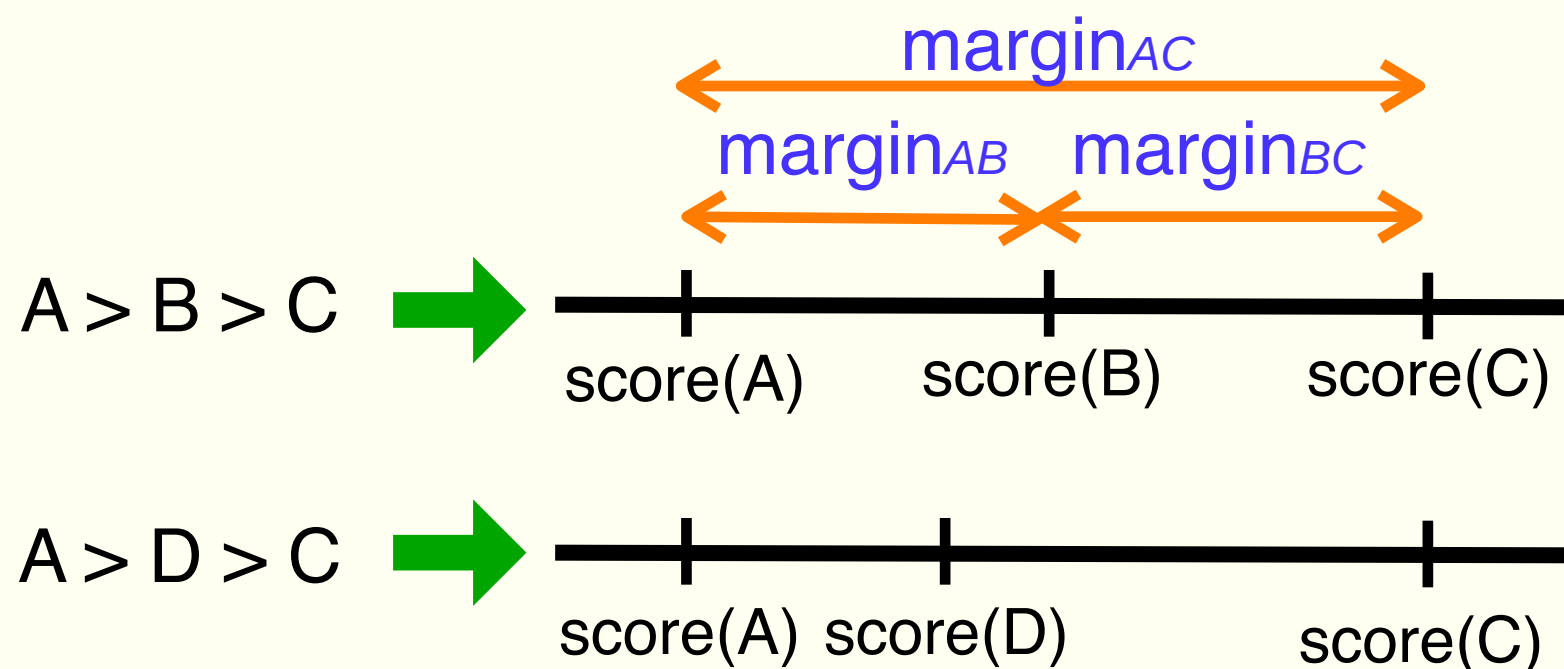
SVOR (Support Vector Ordinal Regression)

SVMの一種でカーネルの導入も可能

標本順序

Thurstonianスコアとマージン

目的関数



最大化:

$$\sum_{X,Y} \text{margin}_{XY}$$

整列: Thurstonianスコアの順に未整列の対象を整列

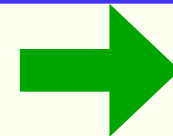
期待順序回帰 [kamishima 05]

学習: 全対象中での期待順位を推定する関数を通常の回帰分析で求めて、Thurstonianスコアとする

全対象を含む順序
(観測されない)

$A < B < C < D < E$

対象は均一ランダムに
サンプリング



サンプル順序
(観測される)

$A < B < D$

この仮定の下で、サンプル順序を観測したあとの、全対象を含む順序中での、サンプル順序中のある対象の順位は次式に比例

$$E[\hat{r}(O_i^*, x_j) | O_i] \propto \underbrace{r(O_i, x_j)}_{\text{サンプル順序中の順位}} / \underbrace{(|X_i| + 1)}_{\text{サンプル順序長}}$$

サンプル順序中の順位

サンプル順序長

この期待順位を推定する関数を最小二乗エラーで求めると、スコア関数が正規分布に従うThurstone型モデルでの最尤推定と等価

整列: Thurstonianスコアの順に未整列の対象を整列

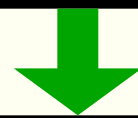
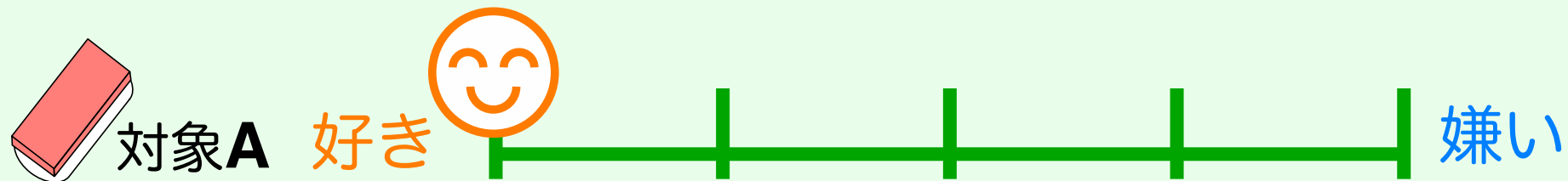
応用例: 主観的な量の計測 (1)

嗜好, 感性, 官能(味覚, 触覚など)の量を順序を用いて計測
教師あり順序付けによって, 未知の対象のこれらの主観的な量を予測

Semantic Differential 法 (SD法)

両端を対義語で表した尺度を用いる

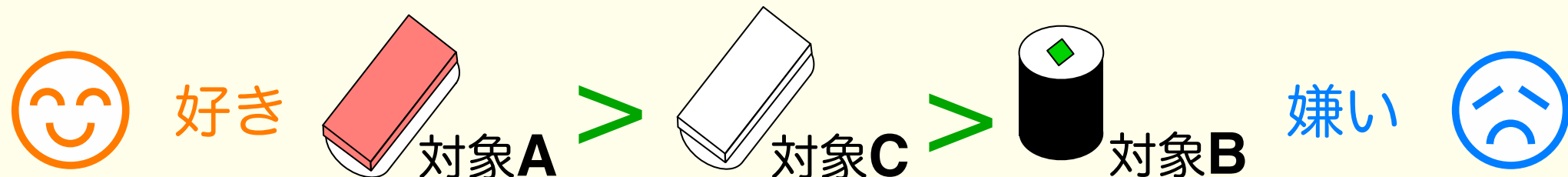
例: 被験者が 対象A を好きならば, 尺度の「好き」を選ぶ



順位法

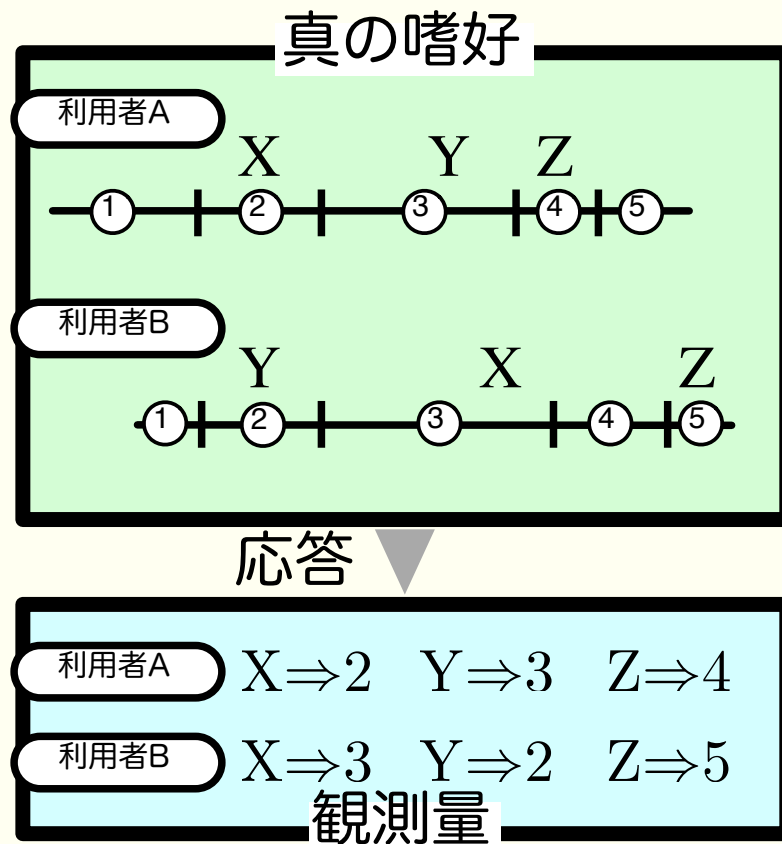
計測する度合いの強さの順に対象を整列

例: 被験者は 対象A が最も好きで, 対象B が最も嫌い



応用例: 主観的な量の計測 (2)

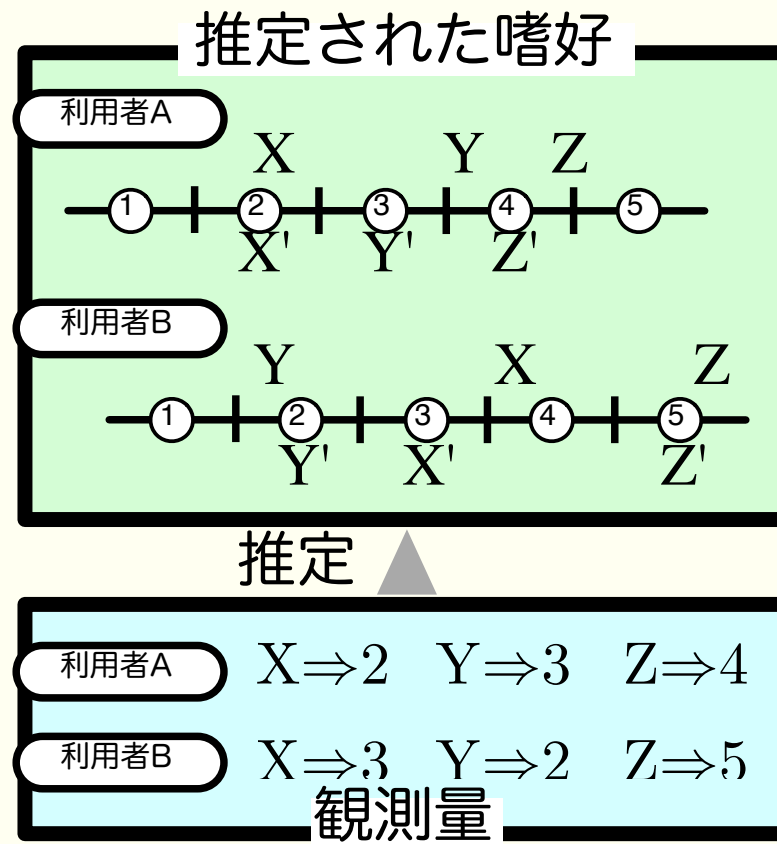
SD法による計測



SD法では、利用者の内部の真の嗜好を、利用者が自身の写像でSDスコアに射影して応答.

写像は利用者ごとに違うので観測量は比較不能.

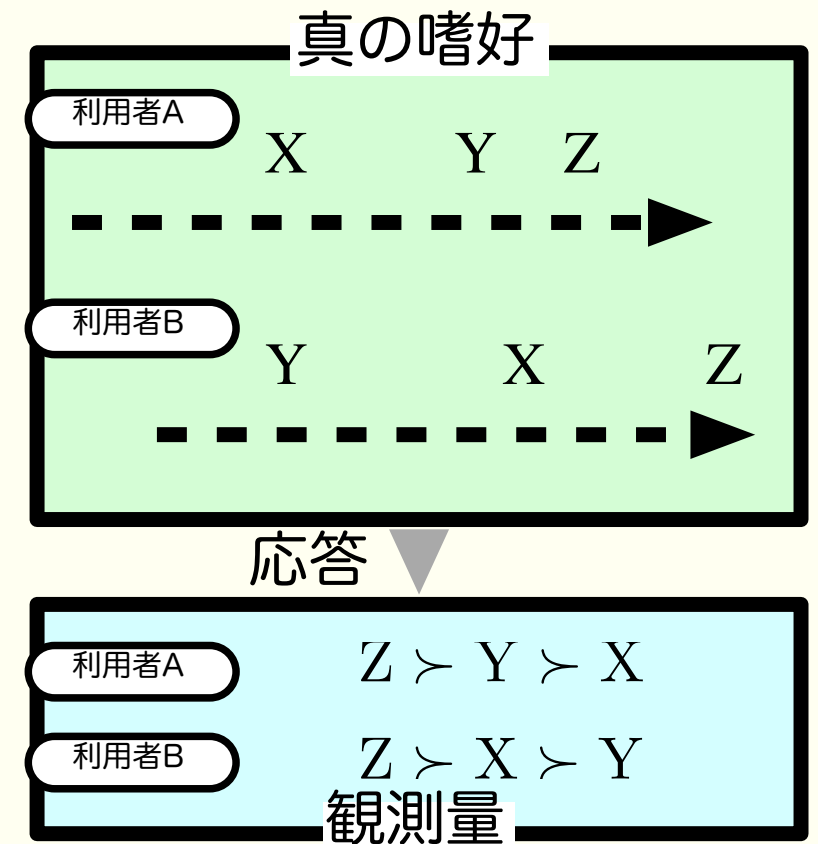
スコアからの嗜好の推定



観測されたSDスコアを比較可能にするため、利用者の写像を揃えるための仮定を導入.

すると、観測量から推定される嗜好はずれてしまう.

順位法による計測



順位法であれば、嗜好の絶対量は無視するので、**全利用者で観測量への写像は共通で、観測量は比較可能.**

応用例: 文書検索のフィードバック

情報検索の結果に対して暗黙的にフィードバックを獲得する[Joachims 02]の方法

クエリQに対する検索結果

1: 検索文書A

2: 検索文書B

3: 検索文書C

4: 検索文書D

5: 検索文書E

利用者が選択

利用者はリストの上位から順に検索文書を見て第3位の文書Cを選択した。

このとき、見たけれども文書AやBは選択しなかったもので、 $C > A$ と $C > B$ のフィードバックが暗黙的に得られる。

文書DやEは見ていないかもしれないので、順序フィードバックは得られない。

教師あり順序付けで、文書のTF-IDFベクトル等から、フィードバックを基に、クエリへの適合順序を予測し、検索結果を再整列する

より詳細な情報

実験結果の詳細は予稿にまとめました

さらに関心のある方はホームページ

<http://www.kamishima.net/>

の参考文献リスト中の

ICDM2005 非公式詳細版

をご覧ください

参考文献

[Cohen 99] W.W.Cohen, R.E.Schapire, Y.Singer. Learning to Order Things. Journal of Artificial Intelligence Research, Vol.10, pp.243-270 (1999)

[Freund 98] Y.Freund, R.Iyer, R.E.Schapire, and Y.Singer. An Efficient Boosting Algorithm for Combining Preferences. In Proc. of The 15th Int'l Conf. on Machine Learning, pp.170-178 (1998)

[Herbrich 98] R.Herbrich, T.Graepel, P.Bollmann-Sdorra, K.Obermayer. Learning Preference Relations for Information Retrieval. In: ICML-98 Workshop: Text Categorization and Machine Learning, pp.80-84, (1998)

[Joachims 02] T.Joachims. Optimizing Search Engines Using Clickthrough Data. In Proc. of The 8th Int'l Conf. on Knowledge Discovery and Data Mining, pp.133-142 (2002)

[Kamishima 05] T.Kamishima, H.Kazawa, and S.Akaho. Supervised Ordering — An Empirical Survey. In Proc. of The 5th IEEE Int'l Conf. on Data Mining (2005) (掲載予定)

[賀沢 03] 賀沢秀人, 平尾努, 前田英作. Order SVM: 一般化順序統計量に基づく順位付け関数の推定. 電子情報通信学会論文誌D-II, Vol.J86-D-II, No.7, pp.926-933 (2003)