

多種類のデータに横断的なパターンを抽出

～複合非負値行列因子分解法: NM2F～

どんな研究

人間が処理し切れない多種類のデータをコンピュータで解析し、データの中に隠れた構造を見つけ出す技術が注目を集めています。本展示では多種類のデータを同時に解析する機械学習の手法「複合非負値行列因子分解法: NM2F」とその解析事例を紹介します。

どこが凄い

今回提案するNM2Fは複数のデータを同時に説明できる整合性のとれたパターンを抽出することが出来ます。NM2Fは数値的な解析性能が優れているだけでなく、人間が理解しやすい横断的なパターンを抽出するため、データを多角的に分析できます。

目指す未来

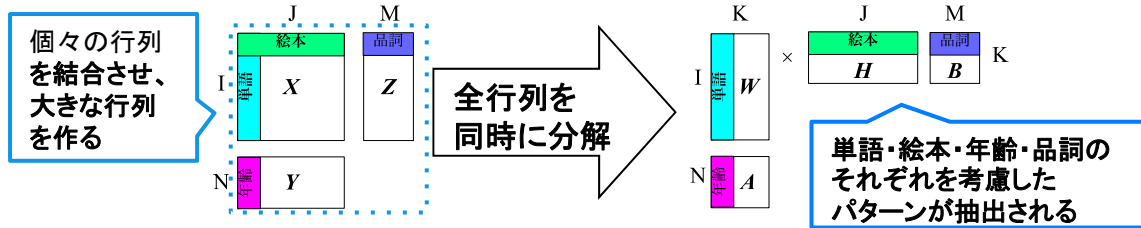
現在、自然気象や人間行動など様々な種類のデータが世界中のあらゆる箇所で観測されています。目的に合わせて膨大な種類のデータから重要なデータを選び出し、社会に役立つパターンを自動で抽出するシステムの発展が期待できます。



多種類のデータを同時に解析することで

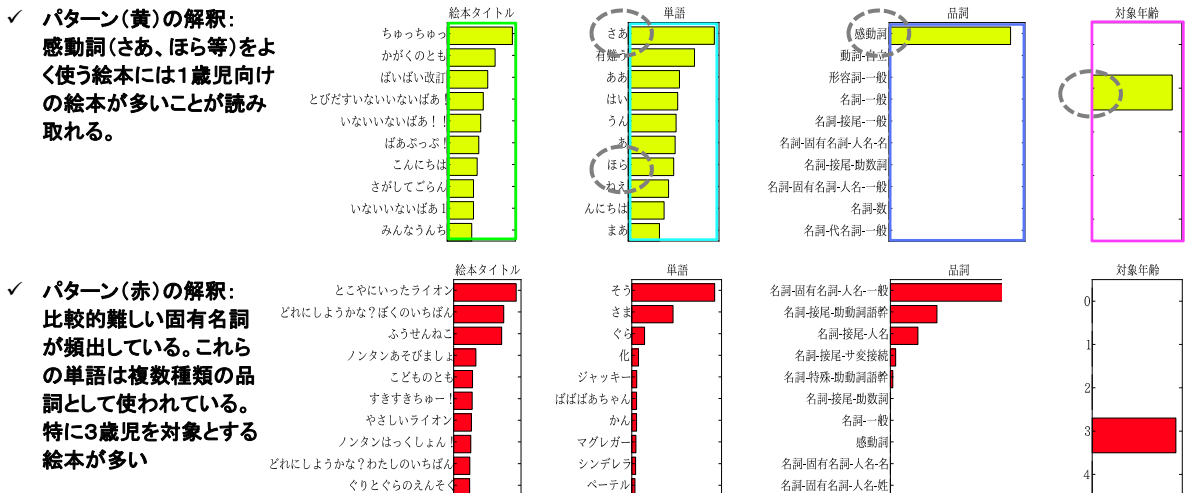
すべてのデータと整合性の高いパターンの抽出が出来ます

- ✓ 複合非負値行列因子分解法 (NM2F: Non-negative Multiple Matrix Factorization)
 - ⇒ 複数種類のデータ行列を1つの複合行列として扱い因子行列に分解する



応用例: 絵本データの潜在構造解析

絵本から抽出された潜在構造を可視化した。対象年齢と品詞が加わることでパターンの可読性が高まった。



関連文献

- [1] 竹内孝, 石黒勝彦, 小林哲生, 藤田早苗, 平博順, “複合非負値行列因子分解(NM2F)による絵本データセットからの多角的パターン抽出,” 人工知能学会全国大会, 2014.
- [2] K. Takeuchi, K. Ishiguro, A. Kimura, H. Sawada, “Non-negative multiple matrix factorization,” in *Proc. IJCAI*, 2013.
- [3] K. Takeuchi, R. Tomioka, K. Ishiguro, A. Kimura, H. Sawada, “Non-negative multiple tensor factorization,” in *Proc. ICDM*, 2013.

連絡先

竹内 孝 (Koh Takeuchi) 協創情報研究部 創発知能環境研究グループ
E-mail: koh.takeuchi[at]lab.ntt.co.jp ({at}の部分をも@に置き換えてください)