

Wing 結合型 SOM によるマルチモダリティ解析

Multimodality Analysis by Wing-like Coupled Self-Organizing Maps

米田 圭佑[†] 中野 貴理博[‡] 堀尾 恵一[†] 古川 徹生[†]
 Keisuke Yoneda Kirihito Nakano Keiichi Horio Tetsuo Furukawa

1 はじめに

近年、インターネット上でマルチモーダルなデータが大量に流通するようになった。たとえば Youtube の動画データは、映像・音声・テキストの3モダリティから構成される。このようなマルチモダリティデータの解析や可視化手法の開発はますます重要になっている。

マルチモダリティデータ解析の主要なアプローチのひとつに、部分空間法に基づく方法がある*。これはモダリティ間に共通する因子、すなわち潜在変数を見出すものである [1]。代表的な手法として正準相関分析 (Canonical Correlation Analysis: CCA) がある [2]。また非線形手法としてはカーネル CCA や Shared Gaussian Process Latent Variable Model などがある [3][4][5]。

一方、モダリティ間の共通因子だけでなく、各モダリティ固有の因子にも注目する場合がある。このような分析法としてインターバッテリー因子分析 (Inter-Battery Factor Analysis: IBFA) がある [6]。IBFA は全モダリティに共通する因子と、各モダリティ固有の因子に分解する。IBFA は心理学の分野で古くから存在する解析法であるが、近年再び機械学習の分野で注目を浴びている [7][8]。

本研究の目的は自己組織化マップ (Self-Organizing Map: SOM) を拡張し、マルチモダリティデータのインターバッテリー解析法を実現することである。すなわち全モダリティ共通のマップと、各モダリティ固有のマップの組み合わせでデータを可視化する Inter-Battery SOM (IB-SOM) の開発である。本研究では M モダリティデータが M 個の固有因子と1個の共通因子から生成される Wing 型の生成モデルを仮定し (図1)、IB-SOM のアルゴリズムを導出した。さらに IB-SOM をワインデータに応用し、化学成分量データとソムリエによる官能評価データのインターバッテリー解析を行った。

2 問題設定

観測データ集合を $X = \{\mathbf{x}_n\}$ とし、データベクトル \mathbf{x}_n は M モダリティの直和 $\mathbf{x}_n = \mathbf{x}_n^1 \oplus \dots \oplus \mathbf{x}_n^M$ とする。ここで $\mathbf{x}^m \in \mathbb{R}^{D_m}$ である。一方、潜在変数 \mathbf{z}^0 を全モダリティの共通因子、 \mathbf{z}^m を第 m モダリティの固有因子とする。このとき観測データは滑

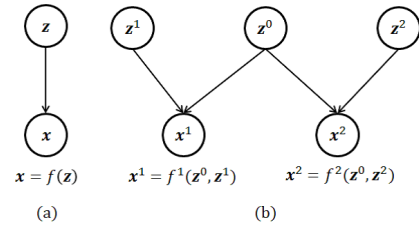


図1 (a) SOM の生成モデル. (b) IB-SOM の生成モデル.

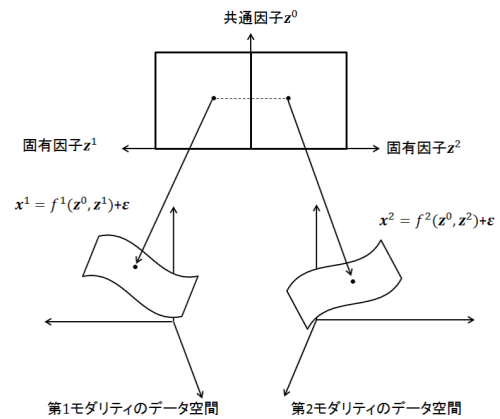


図2 IB-SOM の生成モデルの概念図.

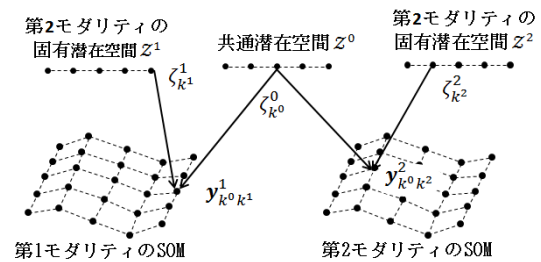


図3 IB-SOM のアーキテクチャ.

らかな位相同型写像により

$$\mathbf{x}^m = f^m(\mathbf{z}^0, \mathbf{z}^m) + \boldsymbol{\varepsilon} \quad (1)$$

として生成されると仮定する (図2)。ここで \mathbf{z}^m と $\mathbf{z}^{m'}$ は互いに独立であるとする。また $\boldsymbol{\varepsilon}$ は観測ノイズであり、本稿では $N(\mathbf{0}, \sigma_m^2 \mathbf{I})$ とする。このようなマルチモダリティデータが観測されたとき、IB-SOM の目的は潜在変数 $\{\mathbf{z}_n^m\}$ ($m = 0, \dots, M, n = 1, \dots, N$) および写像 $\{f^m\}$ を推定することである。

[†] 九州工業大学大学院生命体工学研究科

[‡] 株式会社クラレ

* マルチビューデータ解析とも呼ばれる。正確にはマルチビュー解析とマルチモダリティ解析は区別すべきだが、本稿ではマルチモダリティ解析で統一する。

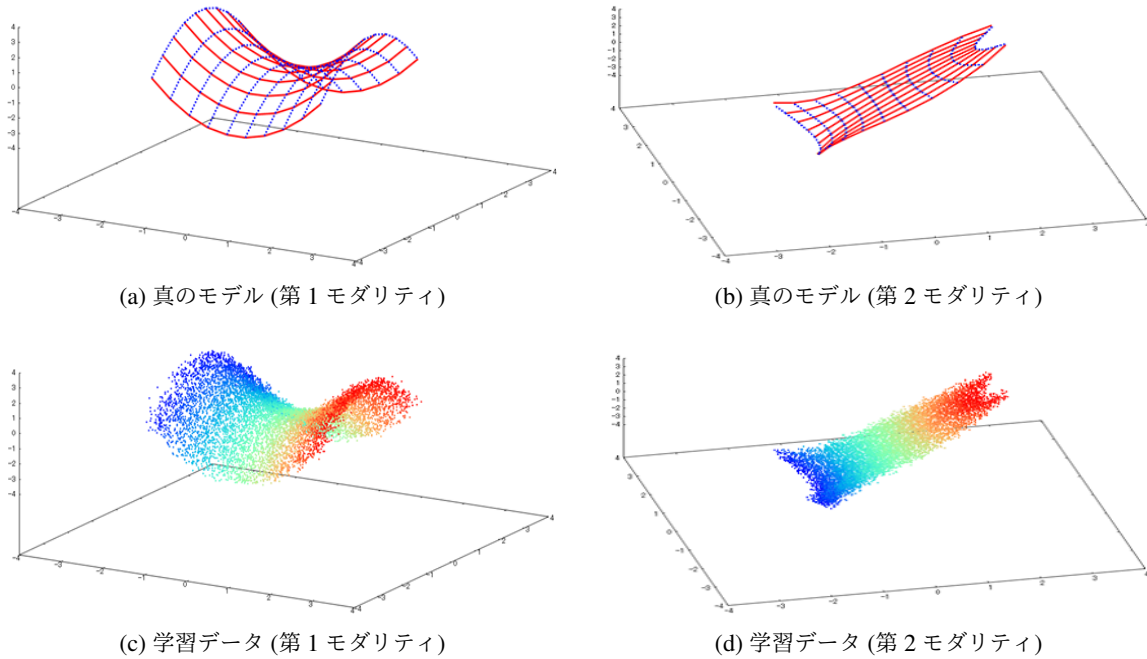


図4 学習データ. 共通因子: 実線(赤), 固有因子: 破線(青).

3 アルゴリズム

SOMでは潜在空間を K 個のノードに離散化する. 今, 第 k ノードの潜在空間における位置ベクトルを ζ_k とし, 観測空間への写像先を $\mathbf{y}_k = f(\zeta_k)$ とする. \mathbf{y}_k は参照ベクトルとも呼ばれる. またデータ \mathbf{x}_n に対してもっとも近い参照ベクトルを持つノードを勝者ノードと呼び, \hat{k}_n と表記する. そして潜在変数 \mathbf{z}_n の推定値を $\hat{\mathbf{z}}_n := \zeta_{\hat{k}_n}$ とする. SOMの目的はデータ $X = \{\mathbf{x}_n\}$ に対する潜在変数の推定値 $Z = \{\hat{\mathbf{z}}_n\}$ を得ること, 写像 $\mathbf{Y} = (\mathbf{y}_k)$ を得ることである.

多くの先行研究により, SOMの目的関数は次式であることが示されている [9][10][11][12][13].

$$L = \sum_{n=1}^N \sum_{k=1}^K H_{\hat{k}_n k} \|\mathbf{x}_n - \mathbf{y}_k\|^2 \quad (2)$$

ここで $H_{kk'}$ は近傍関数である. SOMのアルゴリズムは勝者ノード $\{\hat{k}_n\}$ と参照ベクトル (\mathbf{y}_k) を交互に更新し, 反復的に(2)を最適化する. 具体的には以下の2ステップの交互反復である.

■Step 1: 勝者ノードの決定

まず各データについてもっとも近い参照ベクトルを持つノードを勝者とする.

$$\hat{k}_n = \arg \min_k \sum_{k'=1}^K H_{kk'} \|\mathbf{x}_n - \mathbf{y}_{k'}\|^2 \quad (3)$$

ここで近傍関数 $H_{kk'}$ は

$$H_{kk'} = \exp \left[-\frac{\|\zeta_k - \zeta_{k'}\|^2}{2\sigma^2(t)} \right] \quad (4)$$

である. また σ は近傍半径であり, 学習ステップ t とともに減少させる. なお実際の(3)の計算では写像が十分滑らかであると仮定し, 近似的に

$$\hat{k}_n = \arg \min_k \|\mathbf{x}_n - \mathbf{y}_k\|^2 \quad (5)$$

と求める.

■Step 2: 参照ベクトルの更新

次いで参照ベクトル (\mathbf{y}_k) を更新する.

$$\mathbf{y}_k = \sum_{n=1}^N R_{nk} \mathbf{x}_n \quad (6)$$

ここで $R_{nk} = 1/g_k H_{\hat{k}_n k}$ および $g_k = \sum_{n=1}^N H_{\hat{k}_n k}$ である.

3.1 IB-SOMのアルゴリズム

M モダリティデータに対し, IB-SOMは M 個の固有潜在空間 $\{\mathcal{Z}^m\}$ と1個の共通潜在空間 \mathcal{Z}^0 を持ち, それぞれ K^m 個のノードに離散化する. 本稿では第 m モダリティのノード番号を k^m と表記する.

一方, M 個の写像 $\{f^m(\mathbf{z}^0, \mathbf{z}^m)\}$ は M 組の参照ベクトル $\{\mathbf{Y}^m\} = \{(\mathbf{y}_{k^0 k^m}^m)\}$ により,

$$\mathbf{y}_{k^0 k^m}^m := f^m(\zeta_{k^0}, \zeta_{k^m}) \quad (7)$$

と表される(図3).

IB-SOMの目的関数は(2)を拡張し

$$L = \sum_{n=1}^N \sum_{m=1}^M \sum_{k^0=1}^{K^0} \sum_{k^m=1}^{K^m} \eta^m H_{k^0 k^0}^0 H_{k^1 k^1}^m \|\mathbf{x}_n^m - \mathbf{y}_{k^0 k^m}^m\|^2 + \lambda \sum_{m=1}^M \sum_{m'=1}^M I(\mathbf{z}^m, \mathbf{z}^{m'}) \quad (8)$$

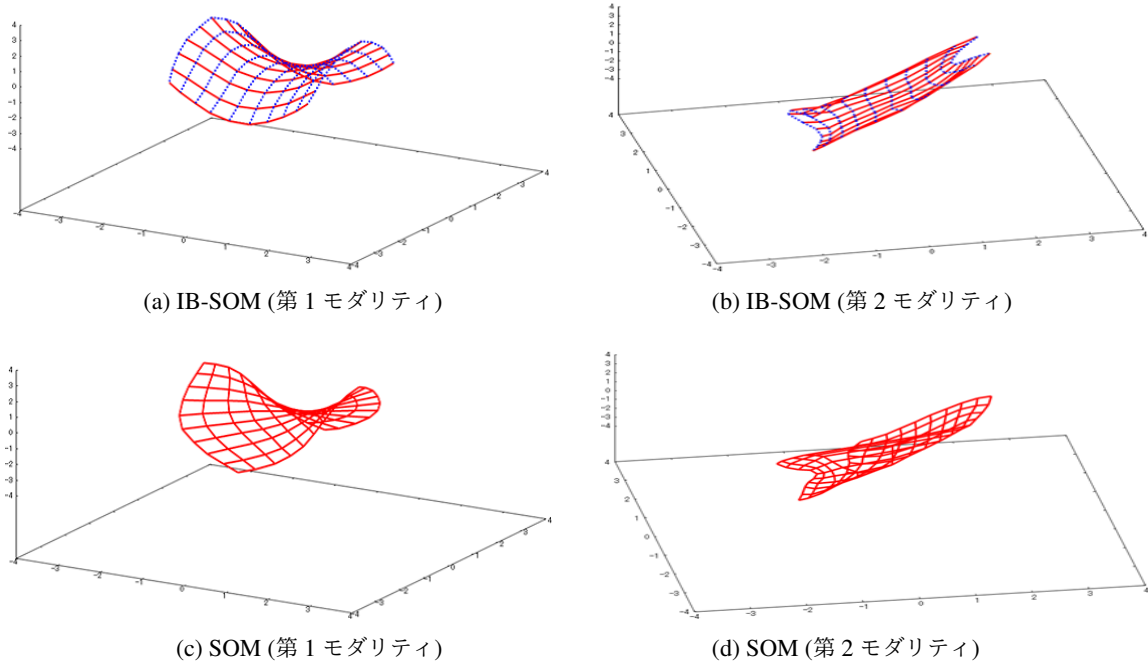


図 5 学習結果. 共通因子: 実線 (赤), 固有因子: 破線 (青).

となる. ここで η^m は第 m モダリティに対する重みであり, 本稿では $\eta^m = 1/D_m$ とした. また $I(\mathbf{z}^m, \mathbf{z}^{m'})$ は相互情報量であり, 固有潜在変数間の独立性が高いほど目的関数は小さくなる. IB-SOM はこの目的関数を反復的に最小化する.

■勝者ノードの決定

データ \mathbf{x}_n に対する勝者ノードを $\hat{\mathbf{k}}_n = (\hat{k}_n^0, \dots, \hat{k}_n^M)$ とする. (8) より

$$\hat{\mathbf{k}}_n = \arg \min_{\mathbf{k}} \sum_{m=1}^M \sum_{k^0=1}^{K^0} \sum_{k^m=1}^{K^m} \eta^m H_{k^0 k^0}^0 H_{k^m k^m}^m \|\mathbf{x}_n^m - \mathbf{y}_{k^0 k^m}^m\|^2 + \lambda \sum_{m=1}^M \sum_{m'=1}^M \ln \left[\frac{P_{k^m k^{m'}}}{P_{k^m} P_{k^{m'}}} \right] \quad (9)$$

となる. ここで P_{k^m} はノード k^m が勝者になった頻度, $P_{k^m k^{m'}}$ はノード対 $(k^m, k^{m'})$ が同時に勝者になった頻度である. さらに (5) 同様の近似を導入し,

$$\hat{\mathbf{k}}_n = \arg \min_{\mathbf{k}} \sum_{m=1}^M \eta^m \|\mathbf{x}_n^m - \mathbf{y}_{k^0 k^m}^m\|^2 + \lambda \sum_{m=1}^M \sum_{m'=1}^M \ln \left[\frac{P_{k^m k^{m'}}}{P_{k^m} P_{k^{m'}}} \right] \quad (10)$$

として勝者を決定する.

■参照ベクトルの更新

参照ベクトルは次式で更新する.

$$\mathbf{y}_{k^0 k^m}^m = \sum_{n=1}^N R_{nk^0 k^m}^m \mathbf{x}_n^m \quad (11)$$

ここで

$$R_{nk^0 k^m}^m = \frac{1}{g_{k^0 k^m}^m} H_{\hat{k}_n^0 k^0}^0 H_{k_n^m k^m}^m \quad (12)$$

$$g_{k^0 k^m}^m = \sum_{n=1}^N H_{\hat{k}_n^0 k^0}^0 H_{k_n^m k^m}^m \quad (13)$$

である.

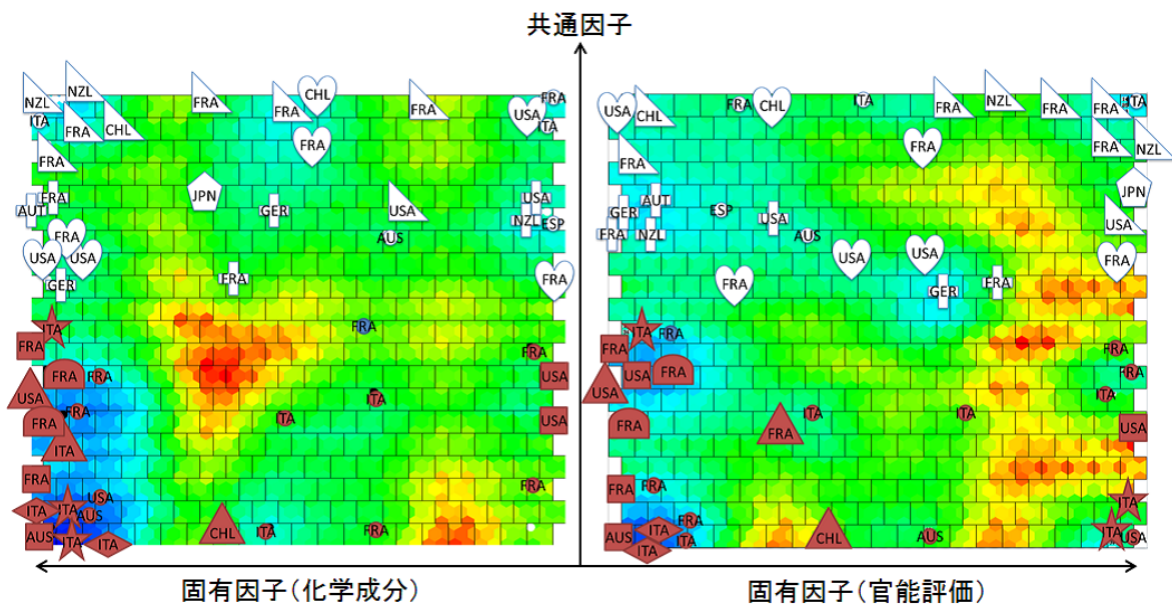
4 シミュレーション

人工データを用いた本アルゴリズムの検証を行った. 学習データの生成に用いた真のモデルを図 4 (a) (b) に, 学習データを図 4 (c) (d) に示す. 共通潜在変数 \mathbf{z}^0 と固有潜在変数 \mathbf{z}^m は $[-1 + 1]$ の範囲で独立に一樣乱数から生成した. 学習データは非線形関数 $f^m(\mathbf{z}^0, \mathbf{z}^m)$ にノイズを加えたものである.

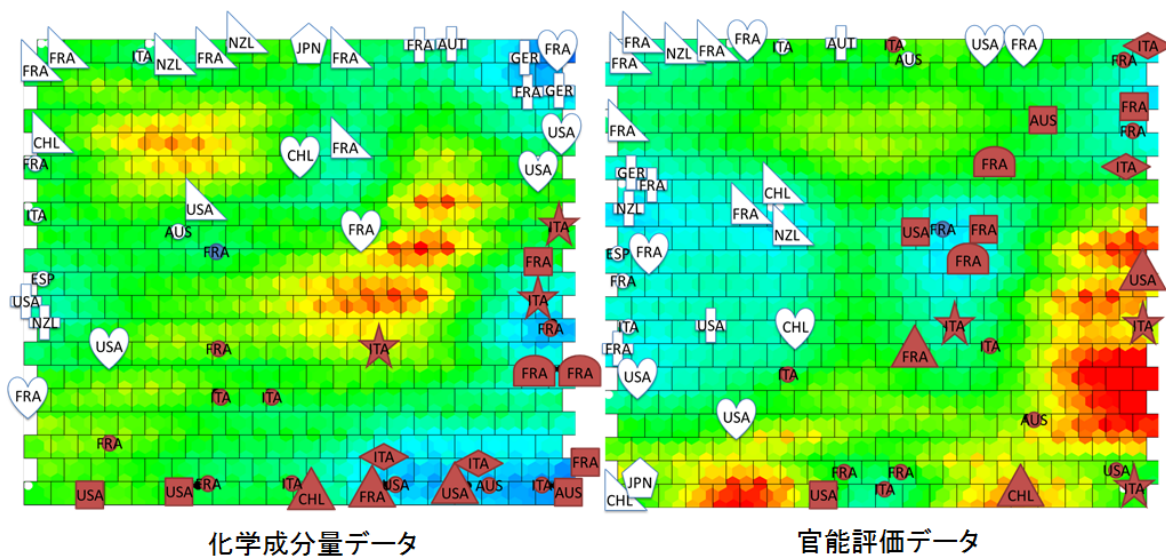
IB-SOM の学習結果と各モダリティを個別に通常の SOM で学習した結果を図 5 に示す. IB-SOM が推定した共通因子と固有因子の張る座標軸 (図 5 (a) (b)) は真のモデルの座標軸を再現できている. 一方 SOM は, 第 1 モダリティのデータを多様体で近似表現できているが, 共通因子と固有因子は推定できていない (図 5 (c)). また第 2 モダリティの SOM の学習結果 (図 5 (d)) では多様体が捻じれており, データを多様体で表現できていない. 従って, IB-SOM は通常の SOM 単独では学習が困難なデータも他方のモダリティデータを活用することで学習できることが分かった. 人工データによる検証は他のモデルから生成したデータでも行い, IB-SOM の再現性を確認した.

| | | | |
|--------|--|-----------------------------------|---------------------------------------|
| ワインの種類 | ● 赤ワイン | ○ 白ワイン | ● ロゼワイン |
| ブドウの品種 | ▲ Cabernet Sauvignon | ◆ nebbiolo | ■ Pinot Noir |
| | ● syrah | ★ Sangiovese | |
| | △ Sauvignon Blanc | ⊕ Riesling | ♡ Chardonnay |
| | ● miscellaneous | ⬠ kosyu | |
| 国名 | USA:アメリカ CHL:チリ ITA:イタリア ESP:スペイン | AUS:オーストラリア FRA:フランス JPN:日本 | AUT:オーストリア GER:ドイツ NZL:ニュージーランド |

(a)



(b)



(c)

図6 (a) ラベル説明. (b) IB-SOMの学習結果. (c) SOMの学習結果. 学習結果はU-matrixで表示.

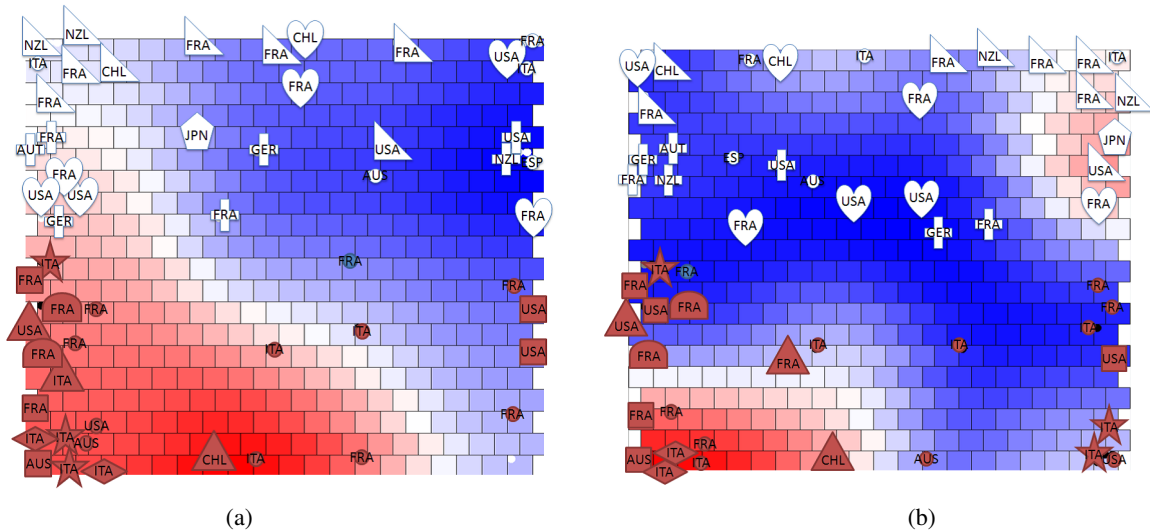


図 7 (a) 1-Hexen-3-ol の含有量. (b) 花系の香り. 赤：強, 青：弱

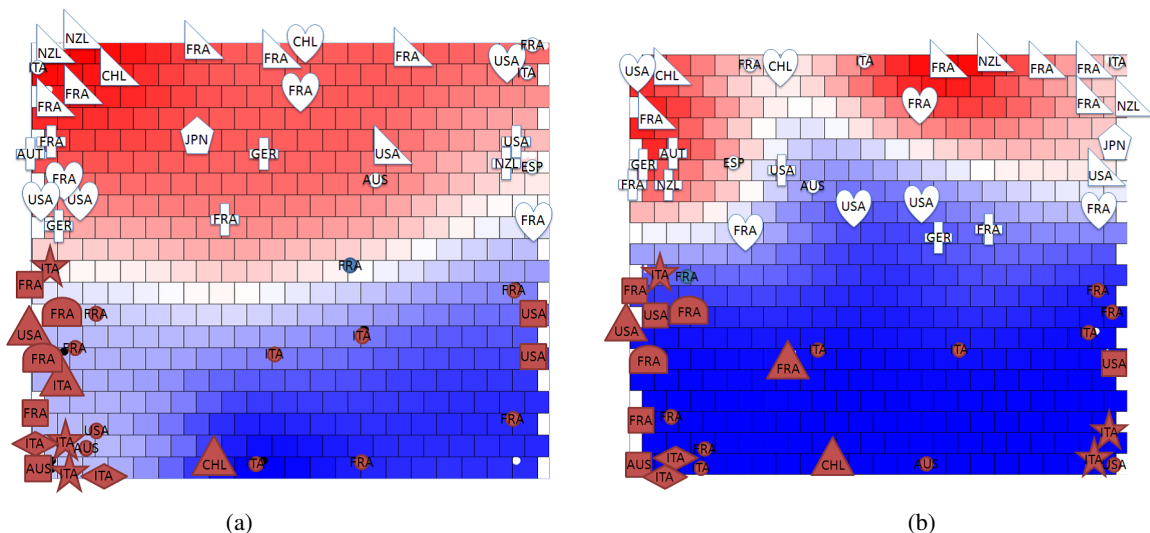


図 8 (a) 1-Hexen-3-ol の含有量. (b) 柑橘系の香り.

5 ワインデータ解析

IB-SOM の実応用として、ワインに関するインターパッテリー解析を行った。ここで用いたのはガスクロマトグラフィーによる化学成分含有量と、ソムリエによる官能評価データの 2 モダリティデータである。

解析対象としたワインは、田崎ワインサロンがソムリエの授業で用いる 55 本である。化学成分の抽出には固相マイクロ抽出法 (Solid Phase Micro Extraction: SPME) とヘッドスペース法を組み合わせた HS-SPME 法を用い、抽出成分の分析はガスクロマトグラフィーにより行った。含有量の分布は対数ガウス分布に従うと仮定して、規格化を行った。

官能評価データの収集にはインターネット上のワインデータベース (Wine-searcher) での、ワイン名と生産年が同一の試料の評価コメントを使用した。ワインの官能評価はアロマホ

ールとして表現され、大分類 13 項目、中分類 29 項目、小分類 94 項目の階層分類となっている。今回は中分類を用い、該当ワインに無かった項目を除く 17 次元ベクトルを学習データとした。そしてアロマが認められる場合は 1、認められない場合は 0 とし、「わずかに」、「ほのかに」等の表現がある場合には 0.5 とした。

IB-SOM のマップ (図 6 (b)) と SOM のマップ (図 6 (c)) のどちらも赤ワインと白ワインが大きく分かれている。また SOM は赤と白の違いが斜め方向に現れているのに対し、IB-SOM では共通因子の方向に現れている。さらに IB-SOM のマップでは、共通因子の方向に赤と白の違いだけでなく、ブドウの品種の違いがわずかに現れている。

また IB-SOM ではコンポーネントプレーン表示により、どの化学成分が香りに関係についても解析できる。フェネチルアルコールの含有量は縦に強弱が分かれており (図 7 (a))、何らかの香りとの相関があることが分かる。実際フェネチルアル

コールはバラ様の香気を有することが知られており [14], 花系の香りでも下の方が強く出ている (図 7 (b)). このように IB-SOM では, どの化学成分と香りが相関があるのかを可視化することができる. また 1-Hexen-3-ol は白ワインに多く含まれており (図 8 (a)), 多くの白ワインは柑橘系の香りを含んでいる (図 8 (b)). つまり 1-Hexen-3-ol は柑橘系の香りに関係していると予想できる. このように IB-SOM を用いることで, クラスタリングだけでなくモダリティ間の関係も解析も可能である.

6 おわりに

本研究では 2 モダリティデータを扱った. 3 モダリティ以上のデータに対しても, 全モダリティの共通因子とモダリティ固有因子のみの場合には本手法は適用できる. しかし各モダリティ同士の共通因子も考慮する場合は, さらなる検討が必要となる. この場合, 推定する因子数がモダリティ数 M に対して $2^M - 1$ になり, 計算コストが指数的に増えてしまう. そのため共通因子が存在するかどうかや不要な因子であるかどうかの, 因子数推定や因子の次元推定が必要となる.

今後の展望として, この因子数推定と因子の次数推定が課題となる. これはマルチモダリティ解析だけの課題でなく, 部分空間学習における潜在空間の次元推定としても重要なテーマである.

参考文献

- [1] Chang Xu, Dacheng Tao, Chao Xu, "A Survey on Multi-view Learning", CoRR abs/1304.5634 (2013)
- [2] H. Hotelling, "Relations between two sets of variates", *Biometrika*, Vol. 28 (1936)
- [3] S. Akaho, "A kernel method for canonical correlation analysis", International Meeting of Psychometric Society (2001)
- [4] Shon, A., Grochow, K., Hertzmann, A., Rao, R., "Learning shared latent structure for image synthesis and robotic imitation", *Neural Information Processing Systems* Vol. 18 (2006)
- [5] Ek, C., Lawrence, N., "Shared Gaussian Process Latent Variable Models", PhD thesis, Oxford Brookes University (2009)
- [6] Ledyard R. Tucker., "An inter-battery method of factor analysis", *Psychometrika*, Vol.23 (1958)
- [7] Archambeau C, Bach F, "Sparse probabilistic projections", *Neural Information Processing Systems* Vol. 21 (2008)
- [8] Andreas Damianou, Carl Ek, Michalis Titsias, Neil Lawrence, "Manifold relevance determination", International Conference on Machine Learning (2012)
- [9] T. Heskes, J.-J. Spanjers, W. Wiegierinck, "Em algorithms for self-organizing maps", International Joint Conference on Neural Networks, Vol. 6 (2000)
- [10] J. Verbeek, N. Vlassis, B. Krose, "Self-organizing mixture models", *Neurocomputing / EEG Neurocomputing*, Vol. 63 (2005)
- [11] S. P. Luttrell, "Derivation of a class of training algorithms", *IEEE Transactions on Neural Networks*, Vol. 1 (1990)
- [12] Y. Cheng, "Convergence and ordering of kohonen's batch map", *Neural Computation*, Vol. 9 (1997)
- [13] T. Graepel, M. Burger, K. Obermayer, "Self-organizing maps: Generalizations and new optimization techniques", *Neurocomputing* Vol. 21 (1998)
- [14] B. C. Rankine, K. F. Pocock, " β -Phenethanol and n-hexanol in wines: Influence of yeast strain, grape variety and other factors and taste thresholds", *Vitis* Vol. 8 (1969)