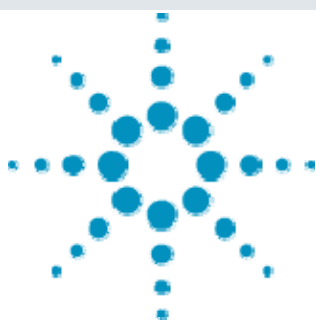
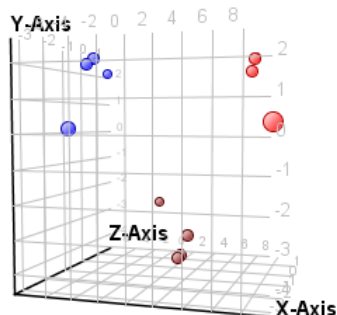




GC/MS 分析における多変量解析の活用 — 日本酒保管試験をモデルとした 判別分析による解析手法 —



＜要旨＞ 日本酒の保管試験の結果を多変量解析ソフトウェア Mass Profiler Professional (MPP)を用いて解析を行いました。経時変化による判別モデルを作成し、既知サンプルを **Unknown** に見立て、正しく判別できていることを確認しました。また、保存状態の異なるサンプルをこの判別モデルに適用することで、どの経時変化の状態にあるのかを調べることも可能でした。さらに、経時変化による判別モデルに寄与する化合物を調べることも可能でした。

Key Words: 多変量解析, MPP, Mass Profiler Professional, 判別分析, PLSD, GC/MS, メタボロミクス, 食品分析, 品質管理

* * * * *

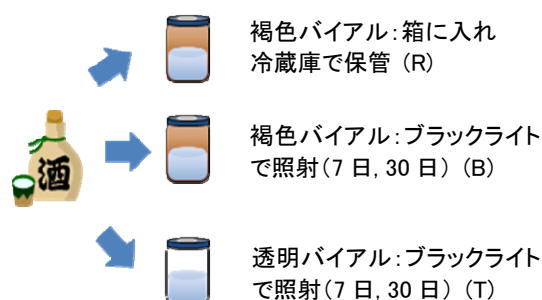
1. はじめに

判別分析 (Class Prediction) とは、既知のサンプルを用いて、あらかじめ判別モデル (品種、等級、嗜好、産地などによるグループ分け) を作成し、それを未知のサンプルに適用することで、どのグループに属するかを判別する手法です。一般的に品種や嗜好などを判別するための化合物は単一成分でないことも多いため、複数の成分によって1つの量の変数 (品種や経時変化など) を予測できる判別分析は、グループ間の差がどの化合物に起因するのかを推測可能にします。

本アプリケーションノートでは、日本酒の保管試験の結果をもとに、Mass profiler Professional (MPP) の判別分析を使って、判別モデルの作成から適用まで行いましたので報告します。

2. 実験方法

清酒を不活性処理済みバイアルビンに入れて、下図の条件で保管したものを、GC/MS で測定しました。



＜前処理方法＞

条件の異なる保管条件の清酒を各 10 μ l ずつ分取

し、オキシム化、トリメチルシリル化を行い試験溶液としました。また、試験は試行回数 $n=5$ で行いました。GC/MS 測定条件は 3. に示しました。

＜ピーク抽出＞

NIST の AMDIS を用いてデコンボリューションを行いました。

3. 測定条件

装置: 7890 GC/5975C TAD MSD

カラム: DB-5ms 30m, 0.25mm, 0.25 μ m, 10m デュラガード

注入量: 1 μ l

注入法: スプリット, 10:1

注入口温度: 250 $^{\circ}$ C

オープン : 60 $^{\circ}$ C (1min) -10 $^{\circ}$ C/min-325 $^{\circ}$ C (10min)

カラム流量: 1.1ml/min (定流量モード)

インターフェース温度 : 290 $^{\circ}$ C

イオン源温度 : 250 $^{\circ}$ C

質量範囲: m/z 50-600

4. 判別モデルの作成

判別能力の高いモデルを作成するには、できる限りグループ間に違いがあり、再現性のよい Entity (化合物) が含まれていることが望ましいとされます。ここでは、ブラックライト照射における経時変化 (対象区、透明バイアルで照射 1 週間、透明バイアルで照射 1 ヶ月) を判別モデルとしました。

まず「Filter by Frequency」を使い、最低 1 グループで 100%出現する (Frequency が 100%) のピークを対象としました。このフィルターによって、ピーク抽出時に得られた 407 成分から、271 成分に絞り込みました。



Agilent Technologies

次に ANOVA (p-value<0.05) でフィルターをかけました。ANOVA は分散分析の手法を用いて、3つのグループで1つ以上の組み合わせで統計的に有意差のある化合物を抽出します。この操作によって、判別モデルを作成するための化合物は271成分から53成分に絞り込まれました。(Fig.1)

Result Summary			
Group Name	[0]	[7]	[30]
[0]	53	29	50
[7]	24	53	22
[30]	3	31	53

Fig.1 Turkey HSD による Post-hoc Analysis (右上の青は有意差のある Entity 数を示す)

判別モデルのアルゴリズムとして、産地判別などで広く一般的に用いられる PLSD (Partial Least Squares Discrimination) を用いました。その結果、Table 1 の結果のように、グループ間で判別能力の高いモデルが作成されたことが分かりました。

Table 1 Validation Algorithm の Confusion Matrix

	[0] (Predicted)	[30] (Predicted)	[7] (Predicted)	Accuracy
(True)[0]	4	0	0	100.000
(True)[7]	0	4	0	100.000
(True)[30]	0	0	4	100.000
				100.000

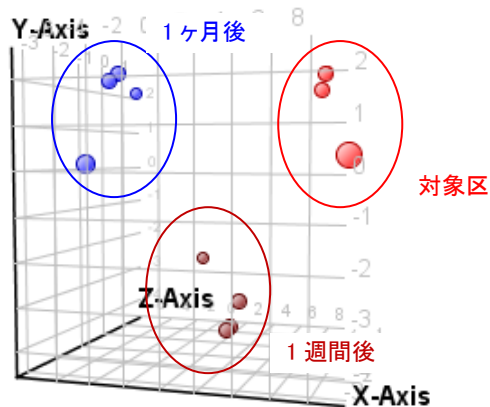


Fig.2 判別モデルと各サンプルのスコア

5. 判別分析

判別モデル作成で使用しなかった試料を未知試料に見立て、作成した判別モデルで正確に判別できるかを検証しました。判別結果を Table 2 に示します。

Table 2 Prediction Results (判別結果)

Identifier	Days	Predicted(Days-model)	Confidence
SAKE_UR	Unknown	[0]	0.672
SAKE_UTW	Unknown	[7]	0.883
SAKE_UTM	Unknown	[30]	0.811

ここで、SAKE_UR は対象区、SAKE_UTW は1週間保管、SAKE_UTM は1ヵ月保管したサンプルでしたが、予測されたグループ(表の [Predicted]) も各々

[0], [7], [30] となり、正確に判別されたことが分かりました。

7. 寄与率の高い化合物を探す

判別モデルを作成したときに、各グループに特徴的な化合物を調べることができます。Table 3 にその結果を表示します。この寄与率の高い化合物は PCA スコアで判定されます。

Table 3 各グループに特徴的な化合物

Compound	[0]	Compound	[7]
144.0@16.47	0.2138874	329.0@23.62	0.4637599
361.0@25.40	0.2097984	176.0@14.30	0.4555686
117.0@12.57	0.2000226	174.0@7.68	0.4541729
267.0@11.89	0.1999953	202.0@18.56	0.4527591
102.0@11.44	0.1995520	105.0@7.50	0.251933

Compound	[30]
174.0@9.76	0.261444
292.0@14.50	0.220899
117.0@13.26	0.187044
204.0@27.88	0.154979
174.0@16.86	0.137363

各グループの Profile plot を Fig.2 に示します。Table 3 より対象区(0)に特徴的とされた 117.0@12.57 の化合物を緑で表示させました。この化合物は対象区では強度を持つのにに対し、7 日後には減少し強度が 0 になることが分かりました。

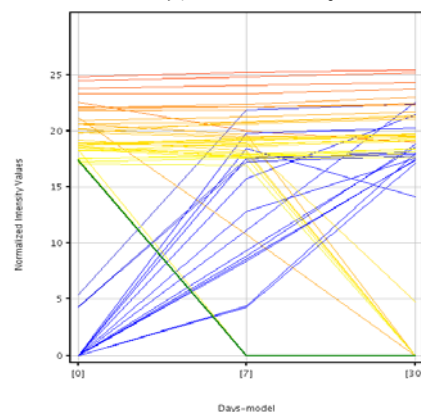


Fig.2 Profile plot (緑の線は 117.0@12.57 の化合物)

7. まとめ

MPP を用いることで、判別分析のためのモデル作成から、未知サンプルの分類予測を行うことができました。また、判別モデルを作成するときに、品種や嗜好などの、機器分析以外の応答変数を組み合わせることで、目的の各グループ(品種や嗜好を知りたい、など)に特徴的な化合物を見つけ出すことが可能でした。

【GCMS-201010SG-004】

本資料に記載の情報、説明、製品仕様等は予告なしに変更することがあります。

アジレント・テクノロジー株式会社

〒192-8510 東京都八王子市高倉町 9-1

www.agilent.com/chem/ja_2



Agilent Technologies