

質量分析法を用いた差異解析

キーワード：質量分析、差異解析、2群比較、多変量解析

はじめに

質量分析装置により取得したデータには、質量電荷比（質量・構造情報）・溶出時間・強度（ピーク面積）の情報が含まれており、これらが1回の分析で得られます。さらに、研究開発における評価や品質管理において質量分析装置を利用すると、大量の測定データが得られます。これらを活かすためには統計処理により、各サンプルに含まれる特徴的な成分を抽出することが重要になります。特に、良品・不良品判定などのサンプル間比較では、2群比較や多変量解析（主成分分析）といった、多数のデータ・変数を横断的に解釈し、「差」を生む変数の組み合わせを決定する操作が重要になります。本テクニカルシートでは、市販茶サンプルの解析例を紹介します。

サンプルおよび測定条件

市販ティーバッグの緑茶、紅茶、麦茶をサンプルとして使用しました。ティーバッグから取り出した茶葉0.5gをそれぞれ3つずつ計量し、50 mLの湯（98°C）で3分間抽出しました。放冷、遠心分離して得られた上清を10倍希釈し、フィルター濾過して測定に使用しました。

茶サンプル中の成分は ODS カラムを用いた HPLC（移動相：0.1%ギ酸水溶液とアセトニトリルのグラジエント）により分離しました。イオン化はエレクトロスプレーイオン化法（Agilent Jet Stream イオン源）により行い、質量分析装置で検出しました。

測定結果

マススペクトル（MS）モードで取得したトータルイオンクロマトグラム（TIC）を図1に示します。TICで各サンプルの傾向を把握することはできますが、サンプルに特徴的な成分やその情報を抽出するには、解析ソフトウェアを用いてデータ処理する必要があります。

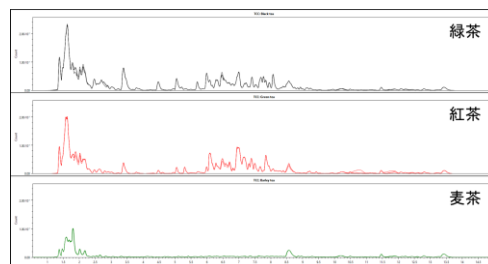


図1 ポジティブイオンモードで測定して得られた各サンプル群のTIC

データの前処理

測定データに含まれる成分の情報を取得するため、化合物ピーク予測アルゴリズムを用いてピーク抽出を行いました。測定データを同一条件で処理したところ、本実験では400成分が予測・抽出されました。

抽出された成分には、再現性の悪い成分など、ノイズとなる情報が含まれています。ノイズを残したまま解析を進めると、各サンプル群に特徴的でない成分を選抜する危険性があります。そこで、「サンプル群の〇〇%に含まれる成分のみ残す」などの条件を指定し、データのフィルタリングを行います。さらに、異常値・外れ値の除去や統計処理を行い、サンプル群の差が有意になる「特徴的な成分」を選抜します。本測定データに対し、「最低1グループで100%出現し、変動係数が25%以下、統計的に有意な差がある」という条件で選抜を行った結果、ピーク抽出された400成分から、グループ間で差が有意な55成分が選抜されました（図2）。

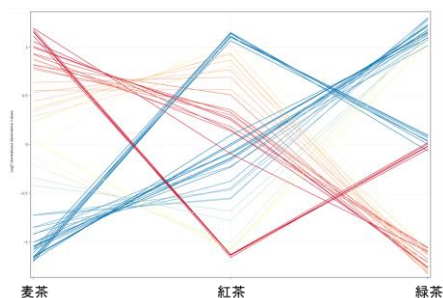


図2 選抜された成分のプロファイルプロット
Y軸は存在量比に、折れ線は成分の増減に対応します。

2 群間比較 (Scatter Plot)

Scatter Plot でのデータ表示により、2 群間の比較がシンプルに行えます。このプロットにより、X 軸成分に特徴的なものほど右下に、Y 軸成分に特徴的なものほど左上にプロットされます。なお、比較したサンプル間で差のない成分は、対角線上にプロットされます。例として、X 軸を緑茶、Y 軸を麦茶として Scatter Plot を行った結果を図 3 に示します。図 3 では、対角線および、両群で 10 倍以上の増減を示す線を表示しています。Scatter Plot の結果、図 3 中赤丸のプロット (m/z 633.1430) は緑茶に特徴的な成分だと分かります。このように、Scatter Plot を行うことで 2 つのサンプルで差がある成分を絞り込むことができます。構造情報を取得する成分の絞り込みにも有効な為、不純物推定などでも活用できます。

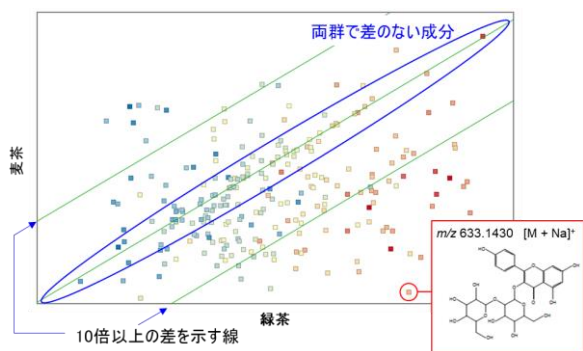


図 3 Scatter Plot による 2 群間の比較例

主成分分析

多変量解析の 1 つである主成分分析はグループ間の傾向をとらえるツールです。サンプル同士の「差」ができるだけ大きくなるような新たな指標 (主成分) を設定し、サンプルが持つ情報を要約します。主成分の軸との相関を示す、数学的に算出される値 (主成分スコア) でサンプル群に固有の範囲が与えられます。そのため、未知サンプルの主成分スコアを算出することで、属する群や傾向を把握できます¹⁾。

例として、茶サンプルの主成分分析および未知サンプルの推定結果を示します。「データの前処理」の項で選抜された 55 成分に対し、部分的最小二乗回帰 (PLS 回帰) アルゴリズムに従って第一主成分、第二主成分を設定しました。X 軸に第一主成分、Y 軸に第二主成分をとり、各主成分スコアに従ってプロットしました (図 4)。緑茶、紅茶、麦茶のグループは第一主成分、第二主成分でほぼ決定され、各グループの違いが明確になっています。また、この主成分分析結

果を基に、2 種類の未知サンプルの解析を行いました。すると、未知サンプル 1 は緑茶と紅茶の中間の性質を持つサンプルと予測されました。また、未知サンプル 2 は、第一主成分は紅茶や麦茶と同様の傾向である一方、第二主成分の傾向は他のサンプル群と全く異なる傾向を示すサンプルだと予測されました。なお、未知サンプル 1 は烏龍茶、未知サンプル 2 はノンカフェインのペットボトル茶です。

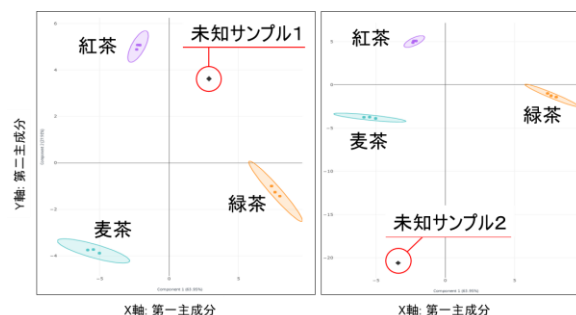


図 4 主成分分析 2D スコアプロット

さらに、検出化合物と主成分の相関を示すローディングプロットを利用すると、各主成分に良く寄与したり、各サンプル群に特徴的な化合物を見つけたりすることができます。図 5 に示した、この主成分分析例で得られたローディングプロットでは、赤の破線で囲った化合物群が第一主成分に良く寄与すると考えられます。

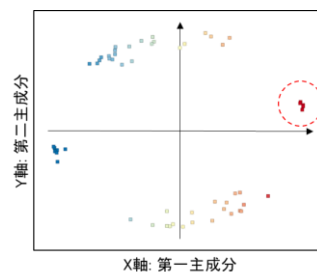


図 5 ローディングプロットの結果

おわりに

2 群比較や多変量解析は、研究開発だけでなく、良品/不良品判定やサイレントチェンジの判別においても強力なツールとなります。実際に解析を行う際、正しくデータを解釈するには、基準となるサンプルや判別対象となるサンプルを複数個分析し、試料成分や測定「ばらつき (曖昧さ)」を減らすことが重要です。さらに、適切な閾値を設定してデータを処理することで分析対象の変化を把握しやすくなります。

1) 栗原伸一, 入門 統計学 検定から多変量解析・実験計画まで, 株式会社 オーム社, 2011

発行日 2021年6月1日

作成者 環境技術研究部

生物・生活材料研究部

有機材料研究部

Phone: 06-6963-8065

環境材料・生物工学研究室

界面活性剤研究室

機能性材料合成研究室

E-mail: hohashi@orist.jp

大橋 博之

佐藤 博文

隅野 修平