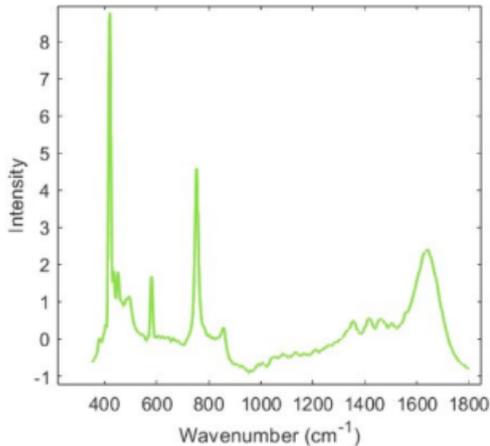


Timegated®ラマン分光解析とケモメトリック・モデリング

このスペクトルから、どのようにして、



3 mM Glucose
27 mM Lactate
2 mM Ammonia

を得るのか？

この画像は、「ラマン分光スペクトルからどのように濃度を推定するのか」という疑問を示しています。

ラマン分光スペクトルは、分子の種類や構造、さらには複雑な分子の特定の部分に対して非常に特異的です。そのため、試料中の成分比率を示す重要な情報を提供します。しかし、スペクトルを単純に視覚的に解析するだけでは、熟練の分光学者であっても正確な情報を得ることは困難です。より詳細な情報を抽出するには、データ解析の手法が不可欠であり、その中でも代表的なのがケモメトリックス（化学計量学）的キャリブレーションです。



Timegate社のオフィスで働くデータアナリスト、**Sofia Lane**。彼女はCHO細胞培養アプリケーションにおけるケモメトリック・モデリングをテーマに修士論文を執筆しました。

Timegated®ラマン分光解析とケモメトリック ・モデリング

ケモメトリック・キャリブレーションとは、試料のラマンスペクトルを用いて、その特性を予測するためのモデルを構築するプロセスを指します。ここで予測する特性（レスポンス変数）には、特定の成分の濃度、質量、体積などが含まれます。この場合、レスポンス変数は連続的な値をとるため、モデルは回帰モデルと呼ばれます。

一方、試料中の汚染物質の有無を検出したり、試料を事前に決められた選択肢の中から分類したりする場合もあります。この場合、レスポンス変数は離散的な値をとるため、モデルは分類モデルと呼ばれます。

ラマンスペクトルとレスポンス変数の関係を定義する方法

ラマン分光データから予測を行うには、スペクトルの変化がレスポンス変数の変化とどのように関連するのかをモデルが理解する必要があります。これを定義する方法として、以下の2つのアプローチが存在します。

データ駆動型モデル

最も一般的なアプローチは、既知の応答変数の値を持つサンプルから測定された大量のスペクトルデータを利用する方法です。これは「トレーニングセット」と呼ばれ、その名の通り、スペクトルと応答変数の関係を学習するために使われます。たとえば、経験から身長や髪の色、肌の質感をもとに誰かの年齢を推測できるのと同じような仕組みです。

このようなモデルは ソフトモデル（データ駆動型モデル） と呼ばれ、統計学や機械学習の手法を活用します。たとえば、最小二乗曲線フィッティング や 多変量最適化 などが用いられます。化学計測（ケモメトリックス）において最も広く使われているソフトモデルは 部分最小二乗回帰（PLS回帰） ですが、サポートベクターマシン（SVM） や ランダムフォレスト などの機械学習手法も、ケモメトリックスのキャリブレーション（校正）に活用されています。

物理ベースのモデル

もう一つのアプローチは、ラマン分光スペクトルがどのように生成されるのかを理論的に理解し、それを利用する方法です。

この手法では、主にピア・ランバートの法則を用います。これは、特定の波数における強度が、その物質の濃度に比例することを示す法則です。混合物のスペクトルは、それを構成する各成分の純粋なスペクトルの加重和として表すことができ、各成分の濃度が重みとして機能します。理論的には、この関係を利用して、混合物のスペクトルと分析対象物質のスペクトルをもとに、その物質の濃度を推定することが可能です。

このようなモデルは ハードモデル(物理ベースのモデル)と呼ばれ、対象物質のスペクトル情報が必要になります。代表的なハードモデルとして、間接ハードモデル(IHM)、古典最小二乗法(CLS)、多変量曲線分解(MCR)などがあります。

ハードモデルとソフトモデルの比較

ケモメトリックスでは、ソフトモデルの方が一般的に使われています。これは、ハードモデルの精度がスペクトルの形成に影響を与えるすべての要因をどれだけ正確に考慮できるかに依存するためです。サンプルが複雑になるほど、これらの要因を完全に把握するのが難しくなります。

一方、ソフトモデルはスペクトルの物理的な背景を深く理解する必要はありませんが、その代わりに大量のトレーニングデータを必要とします。データの収集には時間とコストがかかるため、これも課題となっています。

現在の研究では、これらの課題を克服するために、新しいアルゴリズムの開発やサンプルの前処理、さらには生成AIの活用が進められています。ラマン分光法をプロセス分析でより正確かつ効率的に活用するために、より使いやすく高精度なケモメトリックスモデルの開発が求められています。

詳しくは、CHOパーフュージョン細胞培養におけるラマン分光法の変量モデリングに関する[私の修士論文](#)をご覧ください

Author



This blog was written by Timegate Instruments' Data Analyst **Sofia Lane**. Sofia recently graduated with honors from the Master's program in Mathematics and Operations Research. At Timegate, Sofia works with chemometric calibration methods in biopharmaceutical applications.