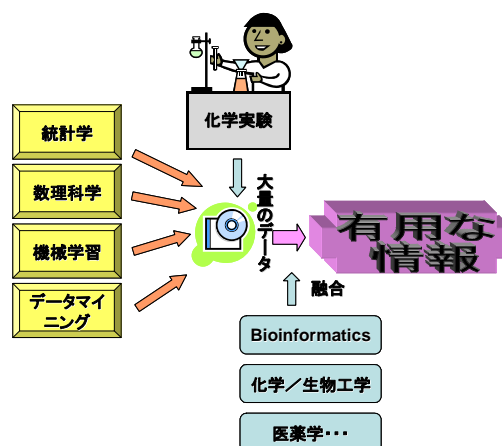


ディビジョン番号	3
ディビジョン名	理論化学・情報化学・計算化学

大項目	2. 情報化学
中項目	2-3. ケモメトリックス
小項目	2-3-1. 情報解析

概要（200字以内）

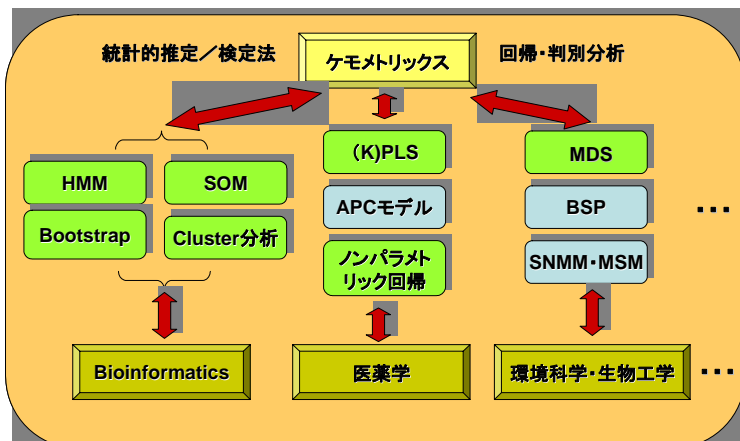
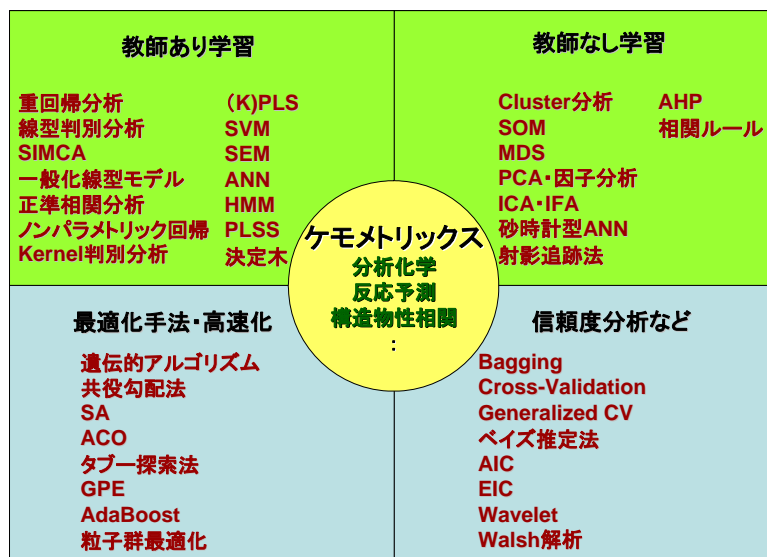
ケモメトリックスは、統計学、機械学習等、数理科学の最前線の手法を化学に応用し、大量で多様なデータをわかりやすく整理し、有用な情報を抽出しようとするものである。分野としては、比較的新しく、主として分析化学分野で発達したが、現在では多様な方法が多様な分野に用いられ、また、Bioinformatics を含む、種々の分野との融合も行われ、更なる発展が最も期待されている分野の一つである。



現状と最前線

ケモメトリックスは、日本では、近赤外分光や食品化学を中心に発達してきたため、calibration の方法論と応用が中心となってきたが、近年は、周辺分野との融合もすすみ、ゲノムデータの解析など、応用範囲が多彩になった。方法論も、SIMCA、PLS、階層型 ANN といったこの世界では古典的な手法から、SOM、Kernel 諸法、ノンパラメトリック回帰、SVM, bagging, boosting (AdaBoost), GA, GPE, ICA, 相関ルール、決定木、Wavelet 変換、Walsh 解析などと、多様化しただけでなく、ICA-PLS[1]、PLSS[2] 等、ケモメトリックスの研究から誕生した手法も少なくない。今後、現在応用が進んでいる多様な手法が評価、融合され、より優れた方法論と、データと手法の対応などが理解されていくことが期待される。同時に、現在応用されている方法の多くは、計量経済学、マーケティング、数理統計学など、化学データを対象としない分野で発展が行われたものであるため、化学データに最適な手法が新たに考案される必要にも迫られている。いわば、この分野が1種の転換点にさしかかったとも言える状況と言っていると思われる。そのためにはケモメトリックス研究者の裾野の広がりが必要だが、国内では少々心許ないのが現状でもある。

また、Bioinformatics だけでなく、計算化学、化学工学、生物工学、医薬学、環境科学など、周辺分野における計量的研究との融合も行われつつある。とりわけ進んでいるのは、医薬品分子設計分野との融合であろう。この傾向は今後ますます進んでいくものと思われる。



参考文献

- 1) Kaneko H., Arakawa M., Funatsu K. *J. Comput. Aided Chem.*, **8**, 41-49 (2007).
- 2) Durand e-F., *Chemomet. Intell. Lab. Syst.*, **58**(2), 235-246 (2001).

将来予測と方向性

- ・ 5年後までに解決・実現が望まれる課題  
 現在使用されている多様な手法の評価と融合・化学データに適した新規手法の開発  
 周辺分野との交流
- ・ 10年後までに解決・実現が望まれる課題  
 国内における研究者の裾野の広がり・実験研究者が利用可能なソフトウェアの幅広い公開  
 プログラミングそのものの最適プログラム化

キーワード

統計学・データマイニング・数理科学・機械学習・多変量解析

(執筆: 高木 達也 )