新しい情報・統計理論に基づくスペクトル分解
―化学イメージング法の開発

名古屋大学 未来材料・システム研究所
高度計測技術実践センター・電子顕微鏡計測部
武藤　俊介

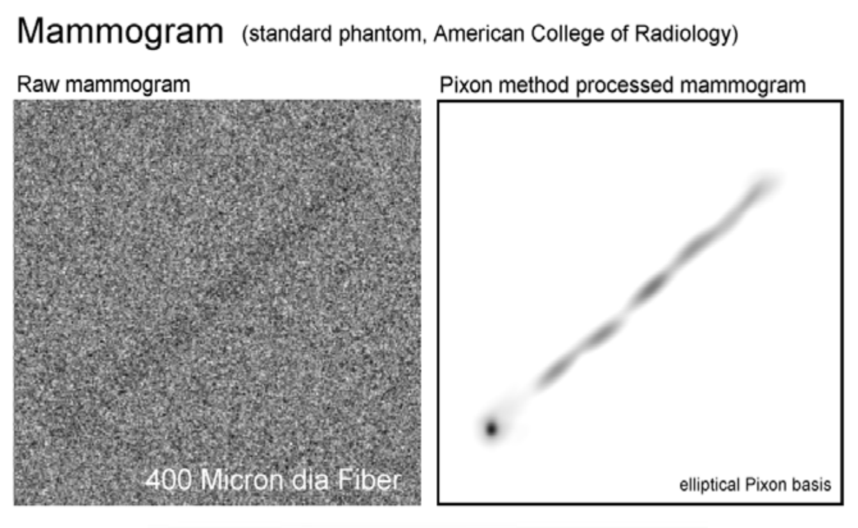
第3回SPring-8材料構造の解析に役立つ計算科学研究会・第10回SPring-8先端利用技術ワークショップ
―放射光利用者のためのマテリアルズ・インフォマティクス入門―
2017年2月27日（月）　(株)ニチイ学館

「いつ」分光データに統計処理が必要か

- 信号強度がノイズレベルに対して同等なとき
  （ランダムノイズの除去）
- 特にデータ点数（評本点の数）が膨大なとき（機械学習）
- さらに分光データにどのような成分が含まれているか不明なとき（複数成分の重なりが予測されるとき）

統計処理が有効なデータ解析例の紹介 1

PIXON法のデモ：画像復元


Mammogram (standard phantom, American College of Radiology)
Raw mammogram
Pixon method processed mammogram
400 Micron dia Fiber
elliptical Pixon basis

- 微弱シグナルの検出への応用

データ解析の現状

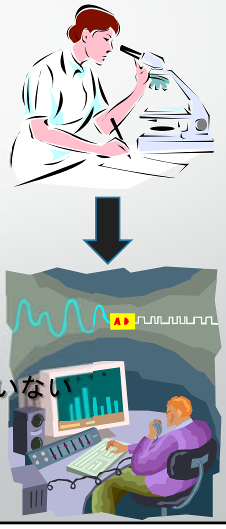
- 分析装置のデジタル化 → PC制御
- 測定の自動化
- ソフトウェアのブラックボックス化

利点：

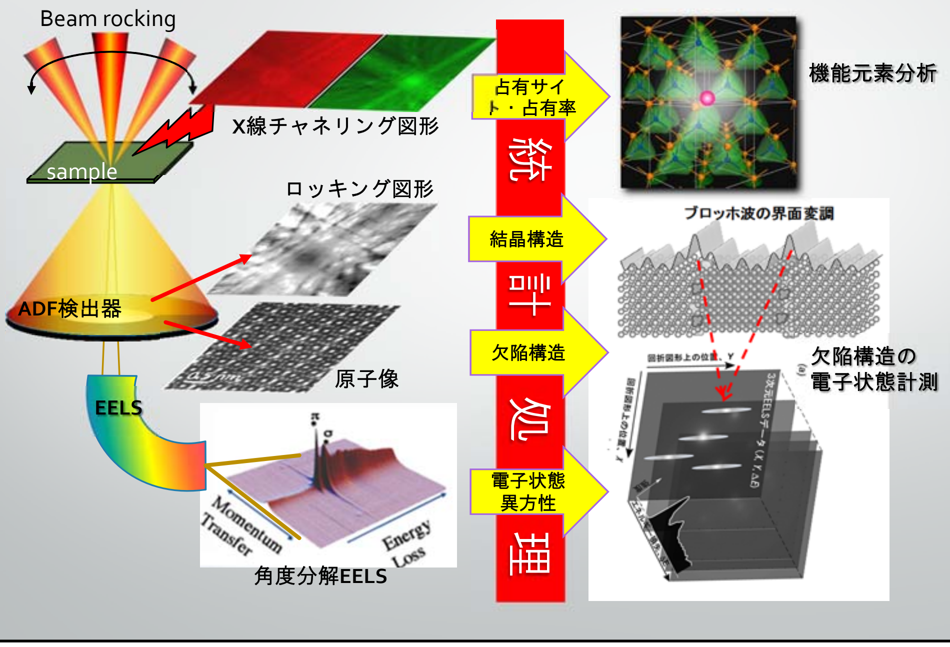
測定の精度向上
同種測定の繰り返しの容易化による網羅的測定

難点：

膨大なデータサイズ
定型解析にしばられ内在する情報を拾い切れていない



複合電子分光による物性イメージング



Beam rocking

sample

X線チャネリング図形

ロッキング図形

ADF検出器

原子像

EELS

Momentum Transfer

Energy Loss

角度分解EELS

占有サイト・占有率

結晶構造

欠陥構造

電子状態異方性

統計処理

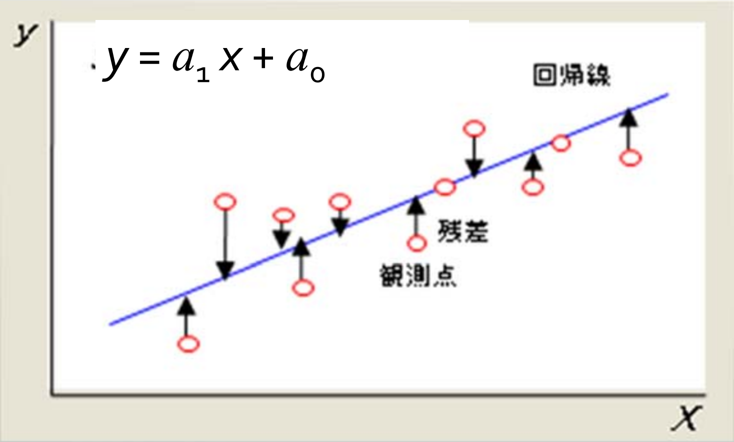
機能元素分析

ブロッホ波の界面変調

欠陥構造の電子状態計測

一種類のデータをモデル化するには？
（一変数から）
→最小二乗法

関係のモデル化：回帰分析：最小二乗法



残差二乗和S＝$\sum$残差^2 ＝ 最小

係数a_1,a_0はどう決められるか

残差二乗和 $S = \sum_{i-1}^{n} \left(y_i - a_1 x_i - a_0 \right)^2$

Sの最小条件 $\frac{\partial S}{\partial a_1} = 0 \quad \frac{\partial S}{\partial a_0} = 0$ ⇔ $\begin{cases} \frac{\partial S}{\partial a_1} = \sum_{i-1}^{n} -2\left(y_1 - a_1 x_i - a_0 \right) x_i = 0 \\ \frac{\partial S}{\partial a_0} = \sum_{i-1}^{n} -2\left(y_i - a_1 x_i - a_0 \right) = 0 \end{cases}$

$$\Rightarrow \underline{\mathrm{X}^{\mathrm{T}}\mathrm{Xa} = \mathrm{X}^{\mathrm{T}}\mathrm{y},} \quad \mathrm{X} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_{\mathrm{n}} & 1 \end{bmatrix}, \quad a = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

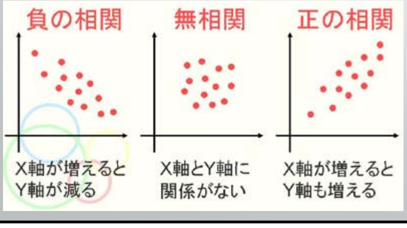
異なる変数間の関係は？
多変量解析

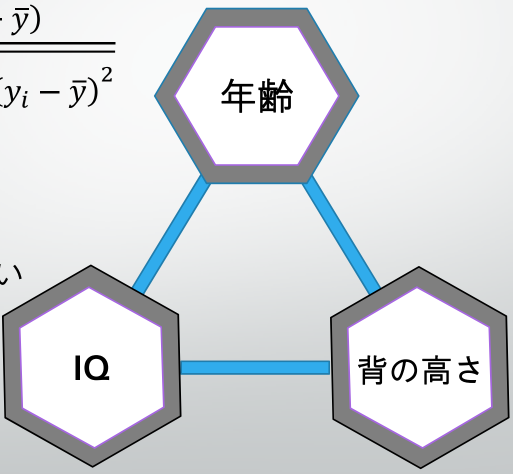
膨大なデータへの統計学的アプローチ：
多変量解析

$$r = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}}$$

相関係数

二つの変数間の関連度合い

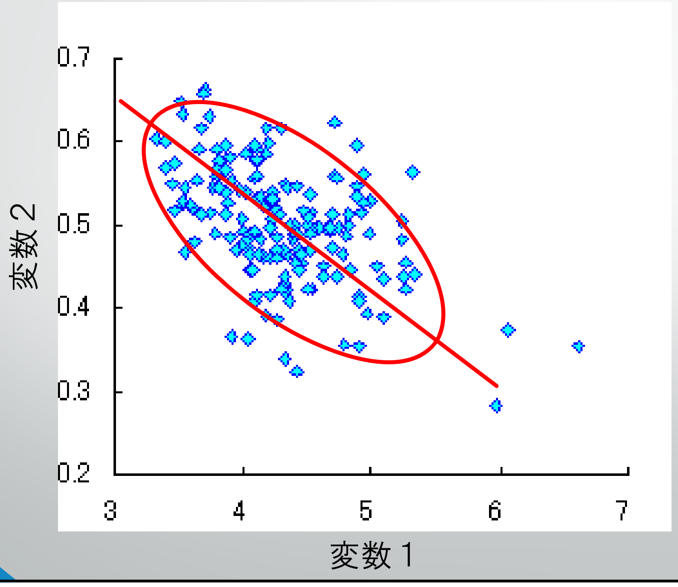



年齢

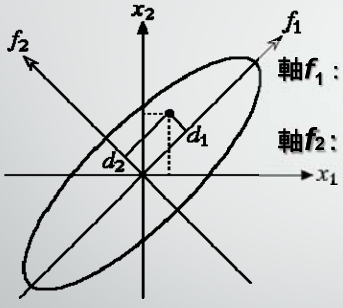
IQ

背の高さ

こんな分布ならどうする？



多次元変数空間：主成分解析（PCA）



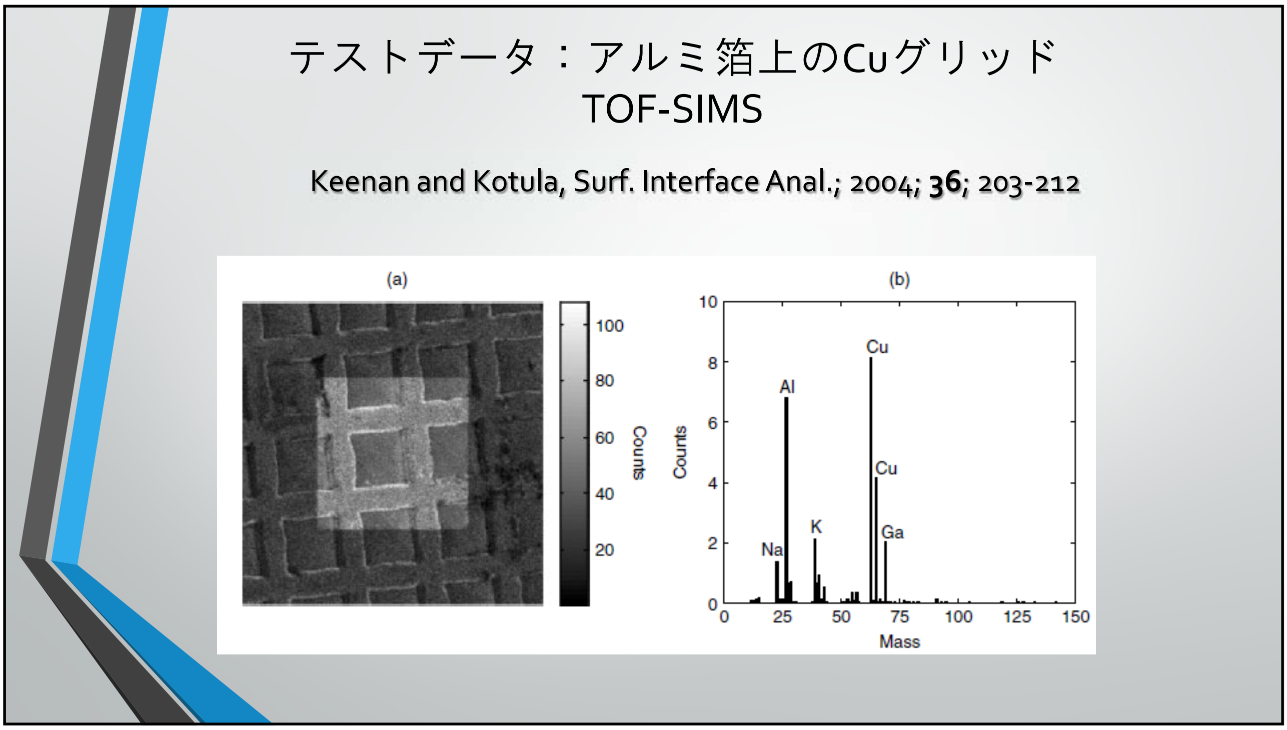
軸f_1：最も分散（情報量）の大きい軸（新しい変数）
変数 x_1, x_2と最も相関が高い
軸f_2：f_1と直交して（f_1と f_2は無相関）, 次に分散の大きい軸

PCA: 特異値分解

$$\mathbf{X} = \mathbf{U}\mathbf{L}^{1/2}\mathbf{V}^T = \mathbf{R}\mathbf{V}^T$$

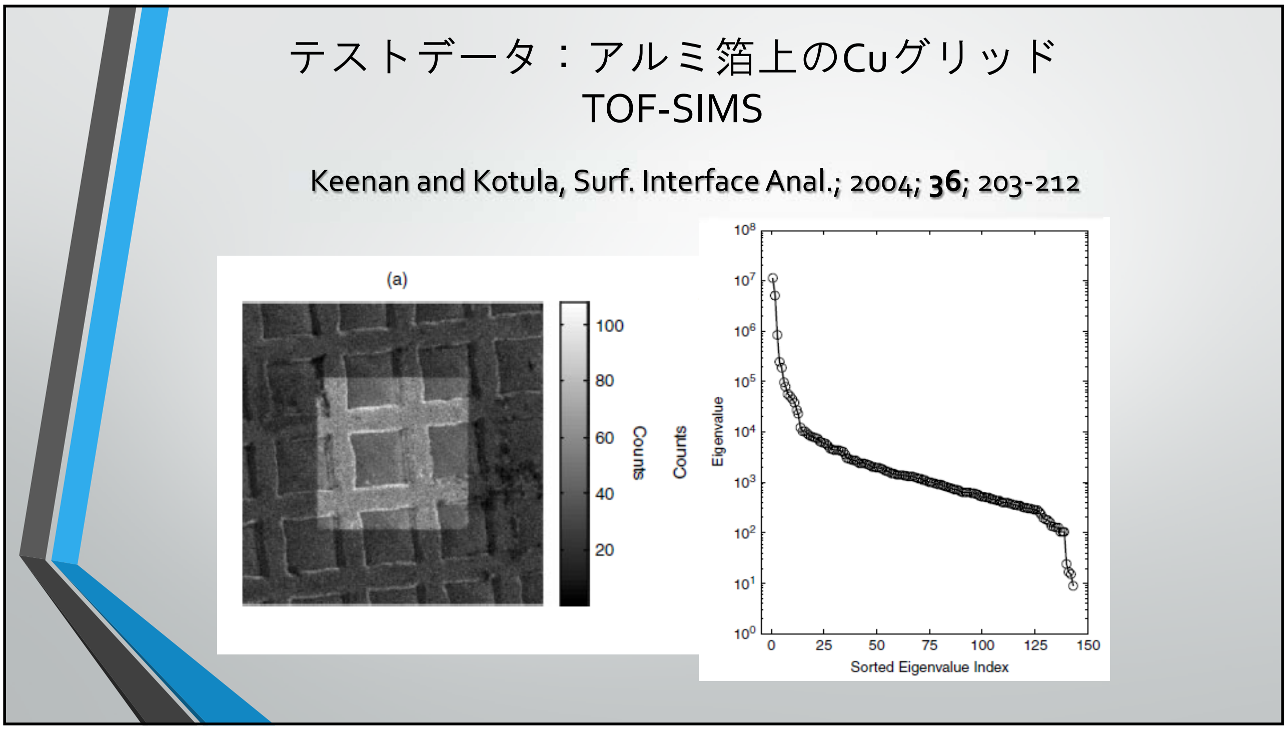
固有ベクトル（ローディング）

固有値行列（スコア）



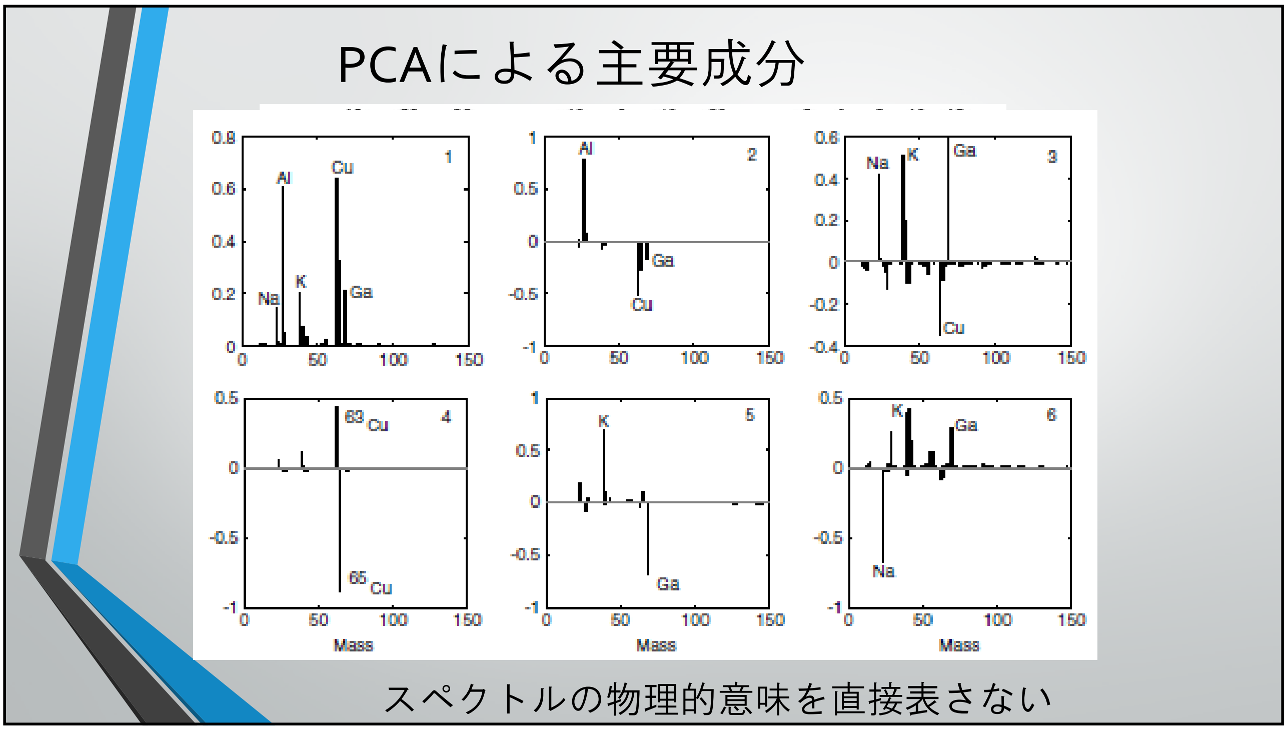
テストデータ：アルミ箔上のCuグリッド TOF-SIMS

Keenan and Kotula, Surf. Interface Anal.; 2004; **36**; 203-212



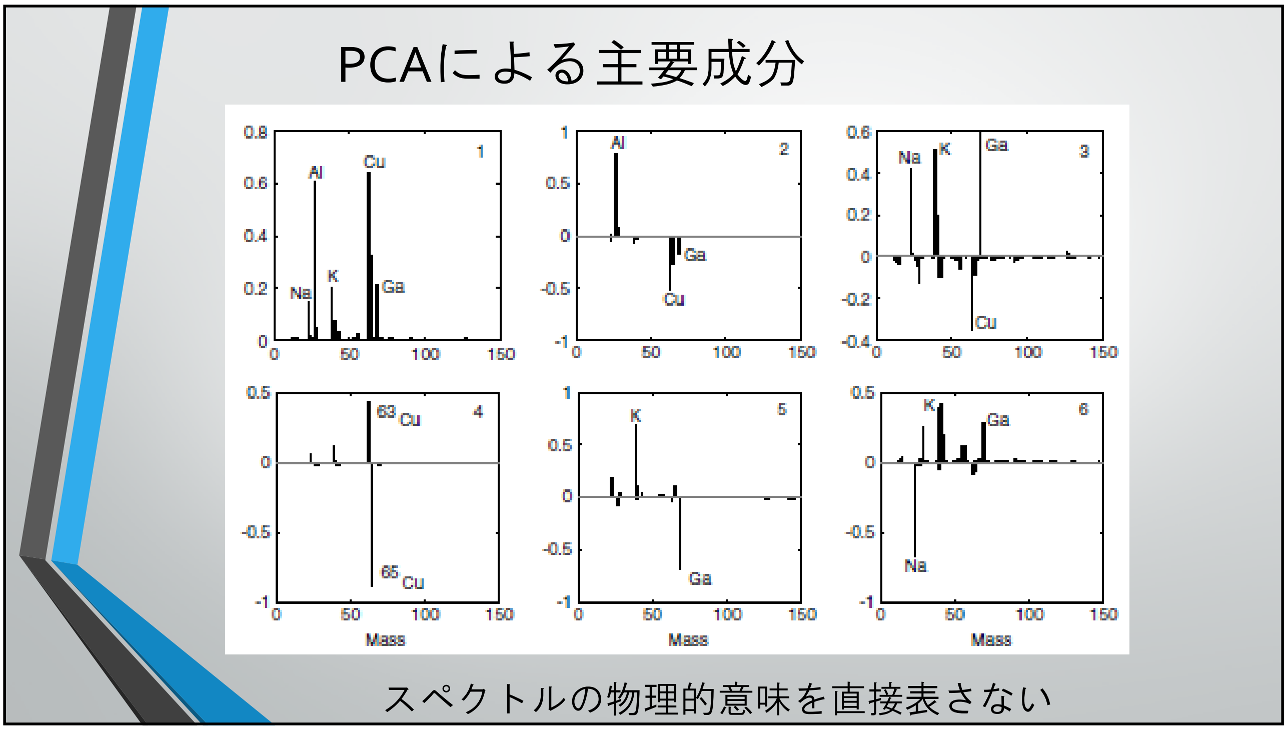
テストデータ：アルミ箔上のCuグリッド TOF-SIMS

Keenan and Kotula, Surf. Interface Anal.; 2004; **36**; 203-212



PCAによる主要成分

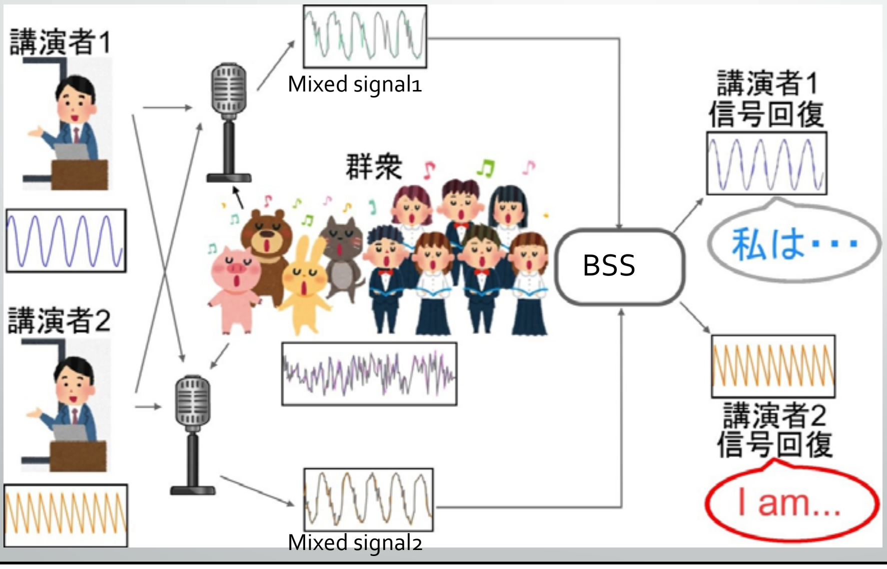
スペクトルの物理的意味を直接表さない



PCAによる主要成分

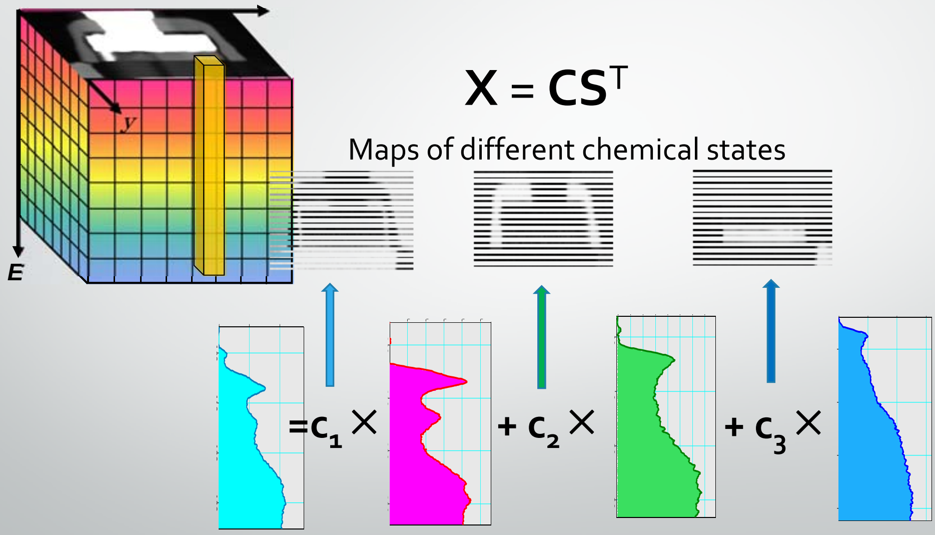
スペクトルの物理的意味を直接表さない

Blind Signal Separation（BSS）


講演者1
Mixed signal1
群衆
講演者2
Mixed signal2
BSS
講演者1
信号回復
私は・・・
講演者2
信号回復
I am...

Datacube：データの数学的構造

Mathematical structure of Two-way Spectral image



$$X = CS^T$$

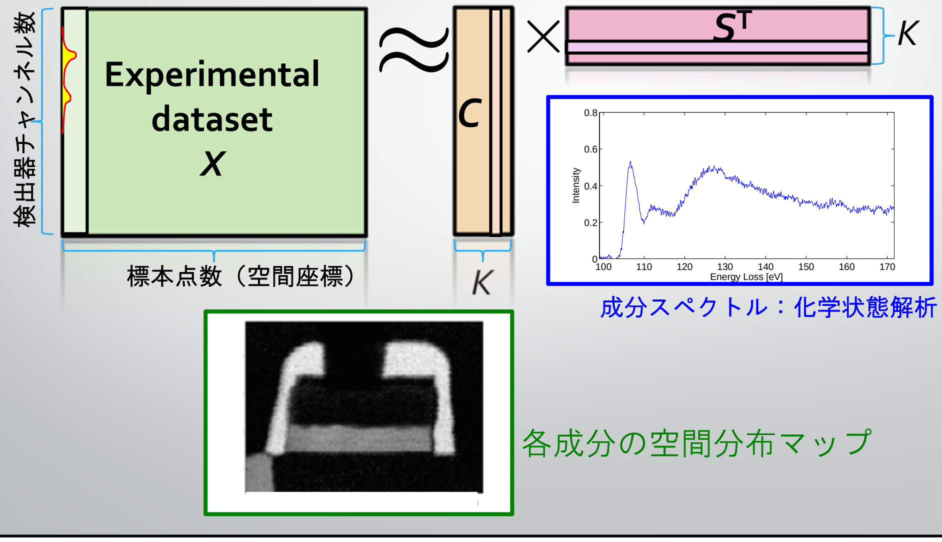
Maps of different chemical states

$$= c_1 \times + c_2 \times + c_3 \times$$

Non-negative entries for all the matrices: NMF

Nonnegative Matrix Factorization (NMF)

Resolve into product of low-rank two matrices
based on non-negativity constraint


検出器チャンネル数
Experimental dataset
X
標本点数（空間座標）
≈
C
K
×
S^T
K
0.8
0.6
0.4
0.2
0
Intensity
100
110
120
130
140
150
160
170
Energy Loss [eV]

成分スペクトル：化学状態解析

各成分の空間分布マップ

多変量スペクトル分解　MCR-ALS

multivariate curve resolution-alternating least-square

***n* experimental spectra, each consisting of *m* points**

⇒ ***m* x *n* matrix, *X***

$$\boldsymbol{X} = \boldsymbol{A}\,\boldsymbol{B} + \boldsymbol{R}$$

Pure component spectra　weights　noise

# of component > # of spectra → non-determinant

$X B^T = ABB^T$ ；B行列を正方行列化

$A = (XB^T)(BB^T)^{-1}$ ；A行列の妥協解

同様に

$B = (A^TA)^{-1}(A^TX)$ ；B行列の妥協解

A,B行列の妥協解が収束するまで以上の繰り返し

主成分解析（PCA）との比較

MCR: $X = CS^T$

PCA: 特異値分解

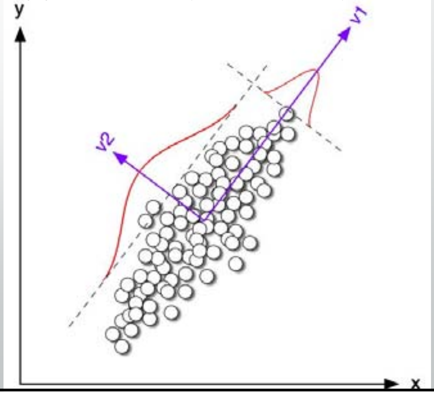
$$X = UL^{1/2}V^T = RV^T$$

固有値行列（スコア）

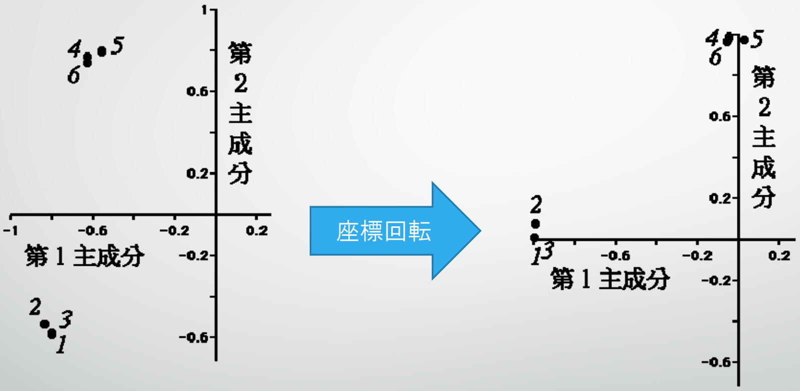
固有ベクトル（ローディング）

$$S = VT,\ C = RT^{-1}$$

MCRとMCAを結びつける
回転行列



具体的な回転行列の意味の一例：Varimax

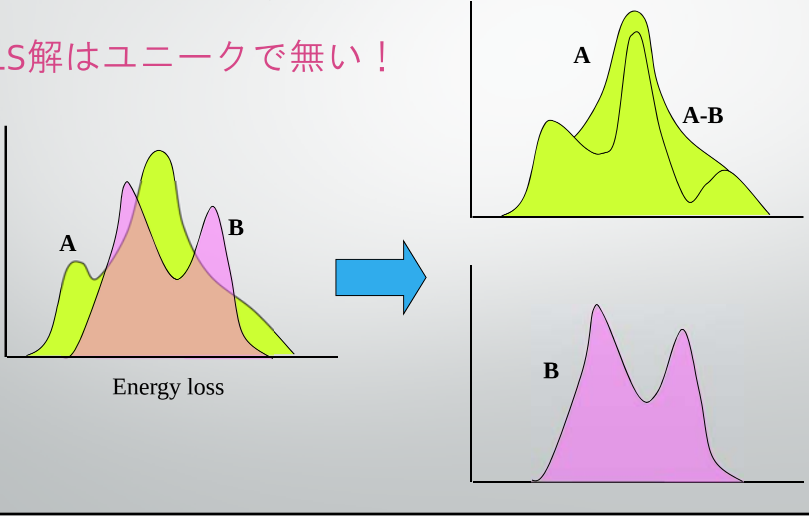


NMFまたはMCRの問題点（特にEELS　不定形）

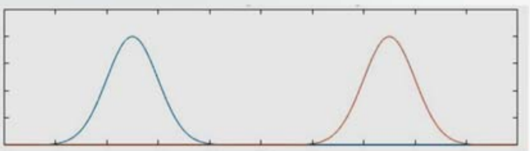
$$\mathbf{X} = \mathbf{C}_{old}\mathbf{S}_{old}^{T} = (\mathbf{C}_{old}\mathbf{T})\,(\mathbf{T}^{-1}\mathbf{S}_{old}^{T}) = \mathbf{C}_{new}\mathbf{S}_{new}^{T}$$

T：arbitrary regular matrix（rotation）

ALS解はユニークで無い！



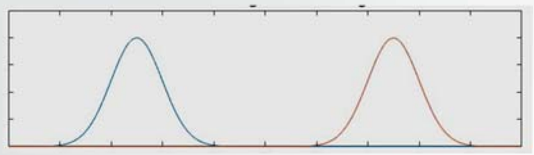
NMF　三つのモード (S or C in $\mathbf{X} = \mathbf{C}\mathbf{S}^{T}$)



Case 1: Non-negative + orthogonal

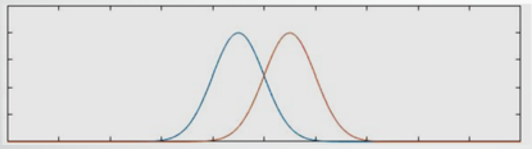
ALS-MCR
（最小二乗法）

NMF　三つのモード (S or C)



Case 1: Non-negative + orthogonal

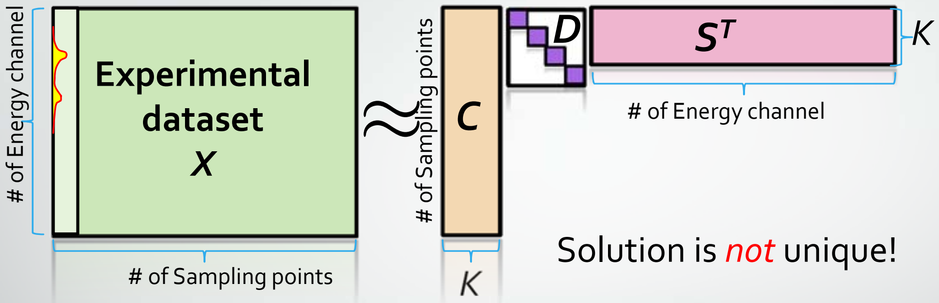
ALS-MCR
（最小二乗法）
VCA



Case 2: Non-negative + partially orthogonal

Soft-orthogonality
VCA

Soft-orthogonality in NMF



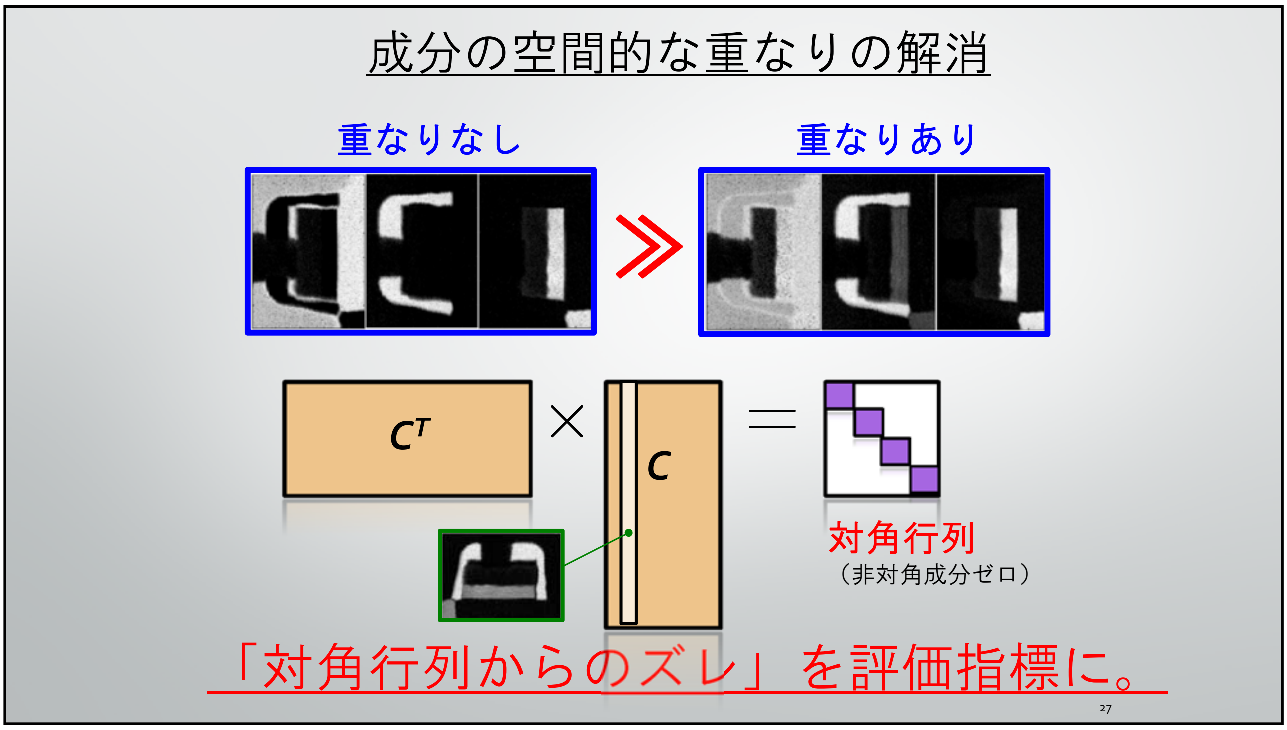
# of Energy channel
Experimental dataset
X
# of Sampling points
≈
# of Sampling points
C
K
D
S^T
K
# of Energy channel

Solution is *not* unique!

Update C and S iteratively
by minimizing “objective function”

【Distance($\left\|\mathbf{X}-\mathbf{C}\mathbf{S}^{\mathrm{T}}\right\|^2$)】
+【Orthogonality($w\lambda\left(\mathbf{C}^{\mathrm{T}}\mathbf{C}-\mathbf{I}\right)$)】＋【sparsity(D)】

Shiga et al, *Ultramicroscopy*, **170** (2016) 43-59.



成分の空間的な重なりの解消

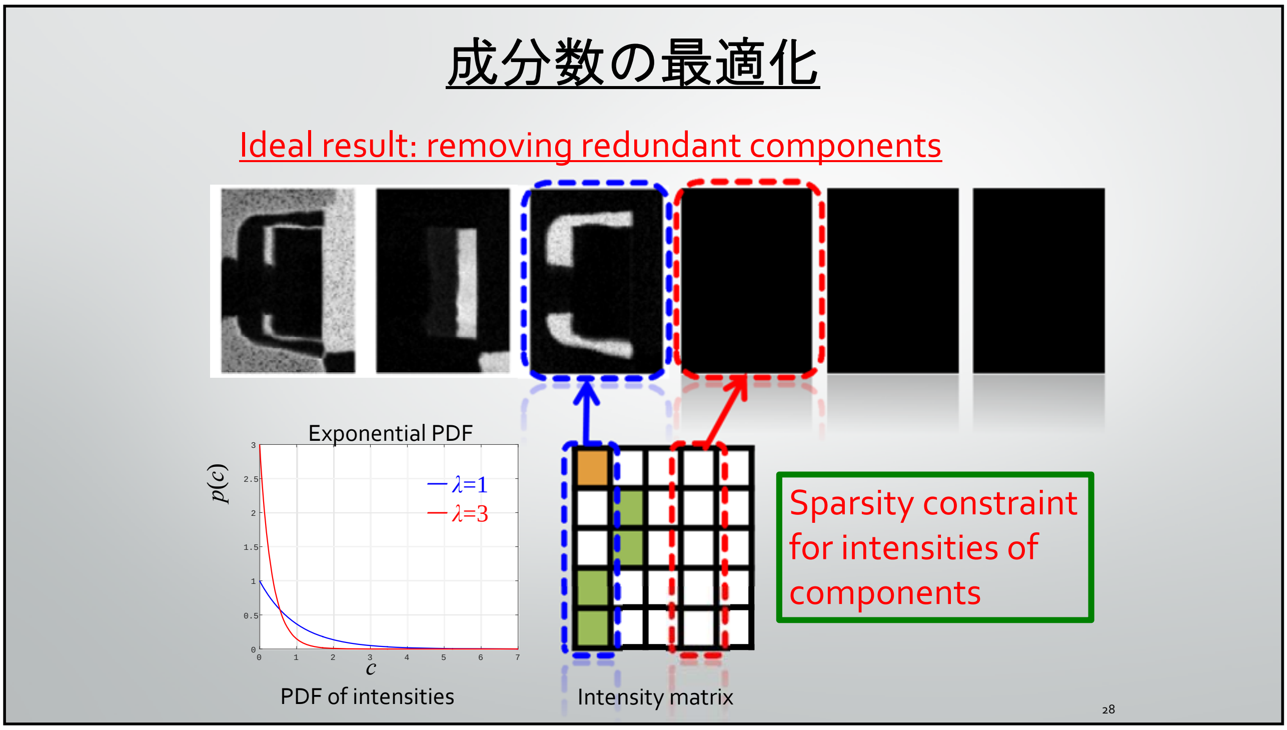
重なりなし ≫ 重なりあり

$C^T \times C =$

対角行列

（非対角成分ゼロ）

「対角行列からのズレ」を評価指標に。



成分数の最適化

Ideal result: removing redundant components

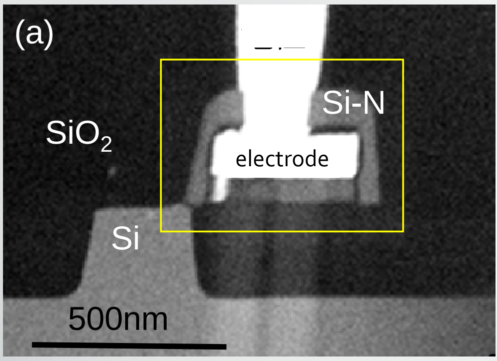
Exponential PDF
$p(c)$
c
— λ=1
— λ=3
0
0.5
1
1.5
2
2.5
3
0 1 2 3 4 5 6 7

PDF of intensities

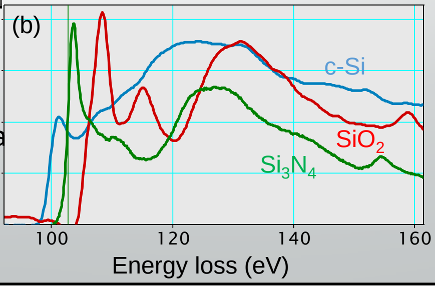
Intensity matrix

Sparsity constraint for intensities of components

ケーススタディ：XSTEM-EELS

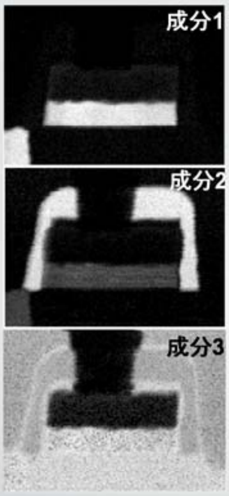

(a)
SiO$_2$
Si-N
electrode
Si
500nm

Reference spectra

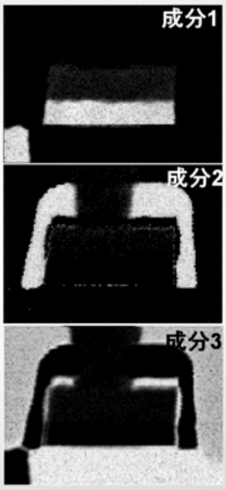


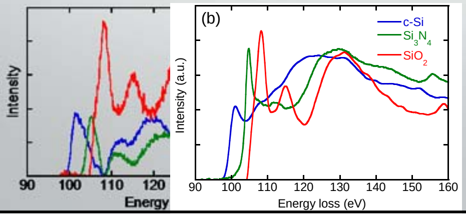
---

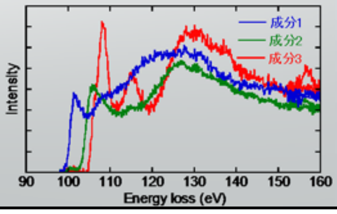
w = 0



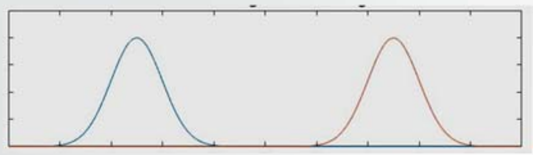
w = 0.1





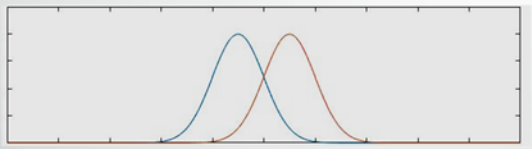


NMF　三つのモード (S or C)



Case 1: Non-negative + orthogonal

ALS-MCR
（最小二乗法）
VCA



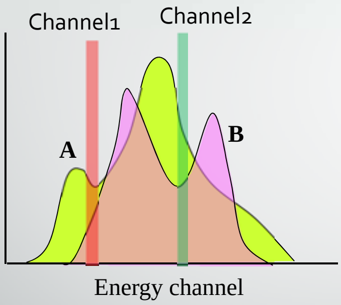
Case 2: Non-negative + partially orthogonal

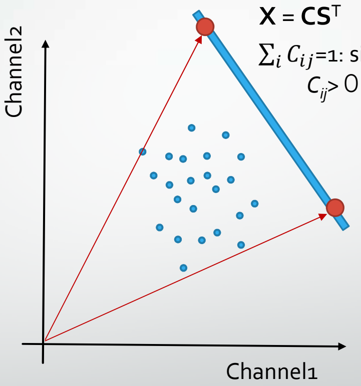
Soft-orthogonality
VCA

Vertex component analysis (VCA)

IEEE Trans. Geoscience and Remote Sensing, 43 (2005) 898.

二成分系


Channel1
Channel2
A
B
Energy channel


Channel2
Channel1

X = **CS**^{T}

$\sum_i C_{ij}=1$: simplex

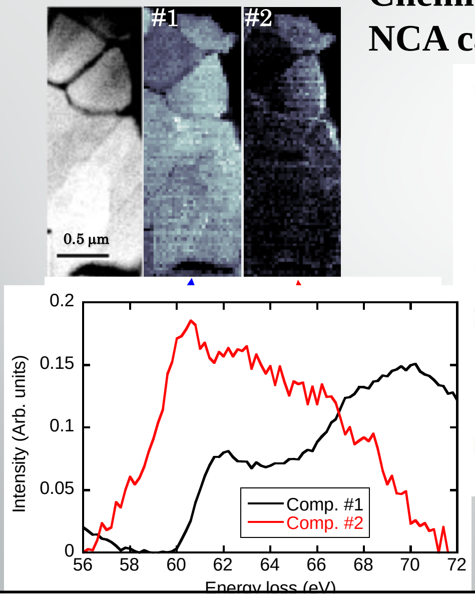
$C_{ij} > 0$

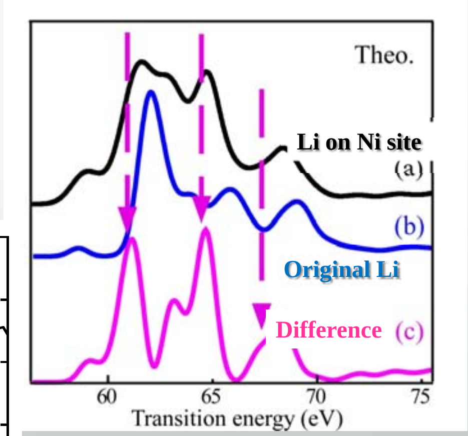
Solution is unique!

前提：pure pixelの存在

Example in NMF-STEM-EELS-SI

Chemical state analysis of Li in NCA cathode

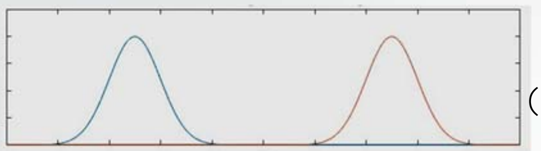




K. Kojima et al, ***J. Power Sources*** (2011).

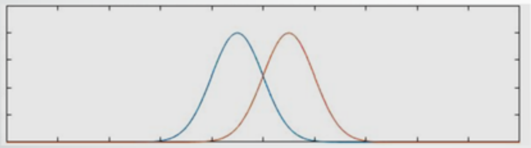
Muto & Tatsumi, *Microscopy*, in press.

NMF　三つのモード (S or C)



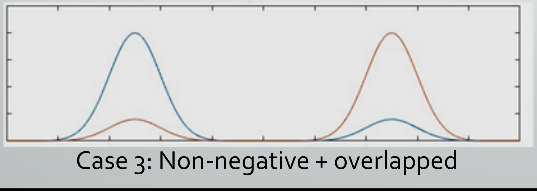
Case 1: Non-negative + orthogonal

ALS-MCR
（最小二乗法）
VCA



Case 2: Non-negative + partially orthogonal

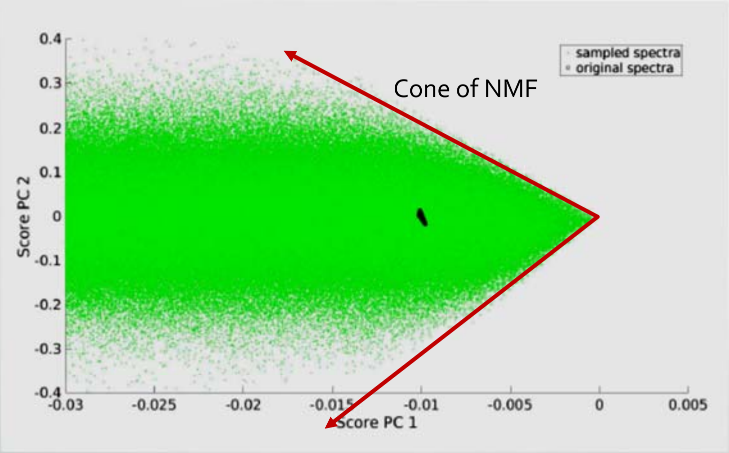
Soft-orthogonality
VCA



Case 3: Non-negative + overlapped

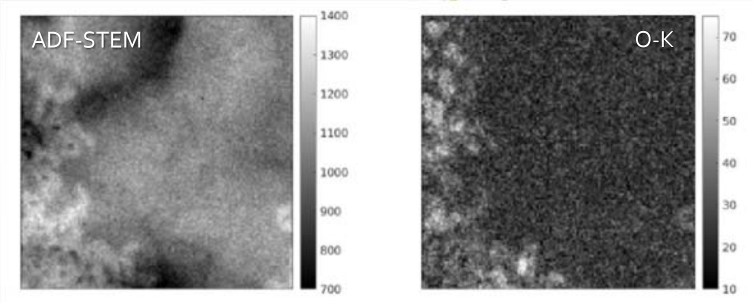
SSS + VCA

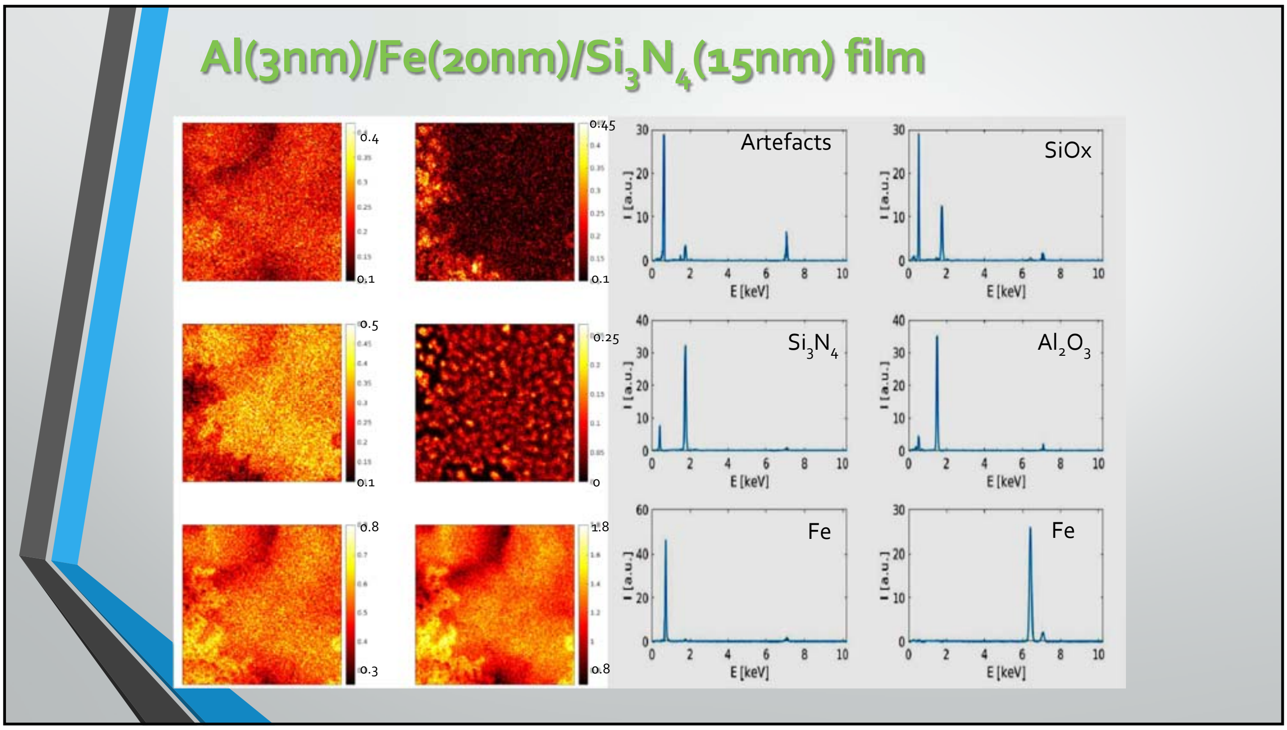
Signal Subspace Sampling in NMF (SSS)



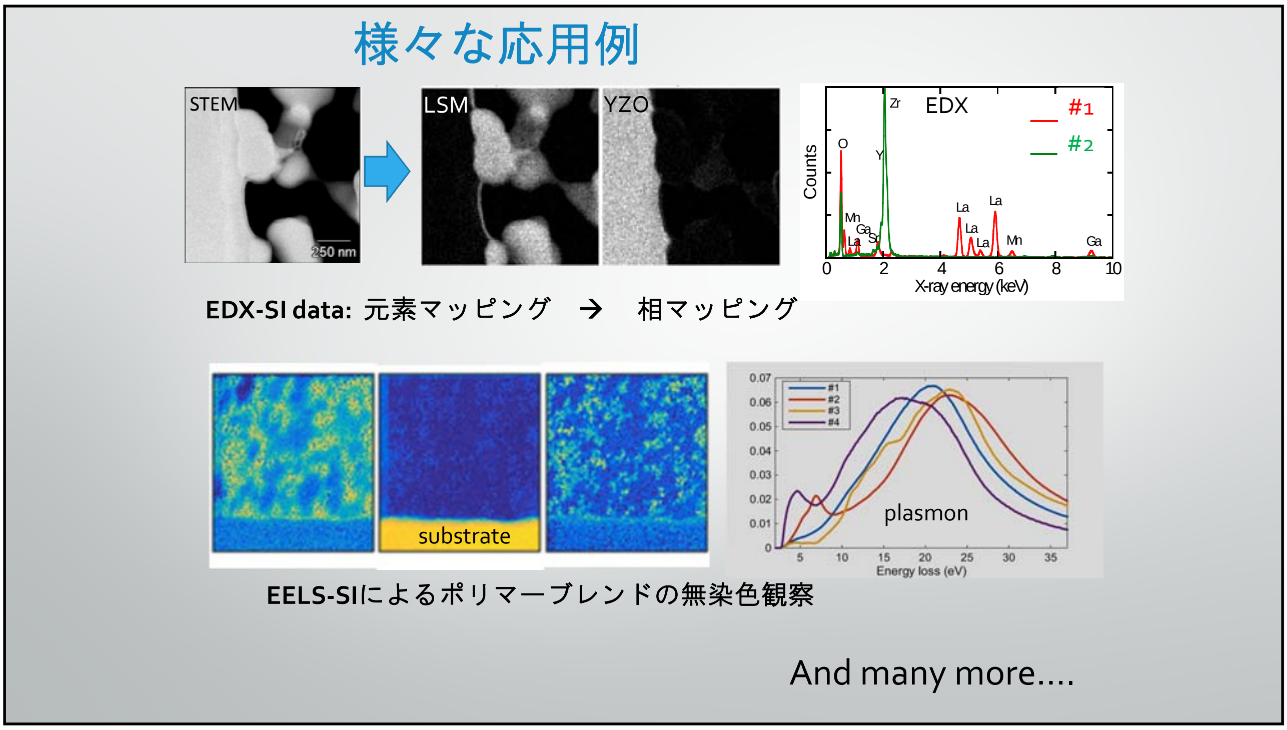
1. Determine the number components by ROB-PCA
2. Select PCs
3. Plot all the linear combinations of PCs **under non-negativitiy constraint**
4. Apply VCA

Al(3nm)/Fe(20nm)/Si_3N_4(15nm) film


ADF-STEM
1400
1300
1200
1100
1000
900
800
700
O-K
70
60
50
40
30
20
10



Al(3nm)/Fe(20nm)/Si$_3$N$_4$(15nm) film



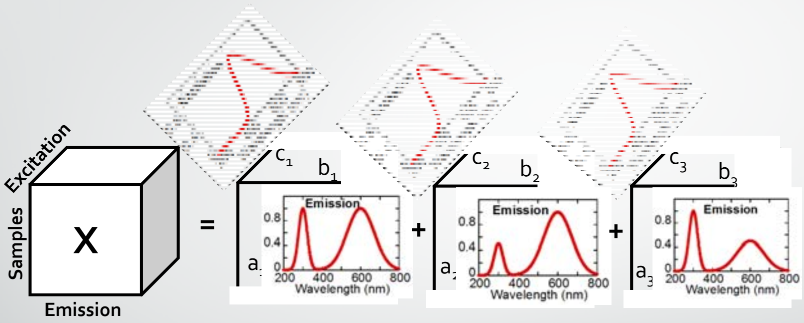
様々な応用例

EDX-SI data: 元素マッピング → 相マッピング

EELS-SIによるポリマーブレンドの無染色観察

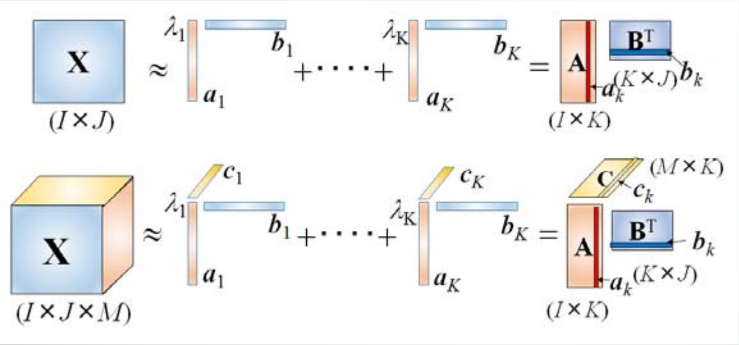
And many more....

Two-way解析からMulti-wayへ


Excitation
Samples
X
Emission
=
c1
b1
a1
+
c2
b2
a2
+
c3
b3
a3
Emission
0.8
0.4
0
200
400
600
800
Wavelength (nm)

Tensor Decomposition:
A. Cichocki et al, *IEEE Signal Processing Magazine* **32,** 145 (2015).

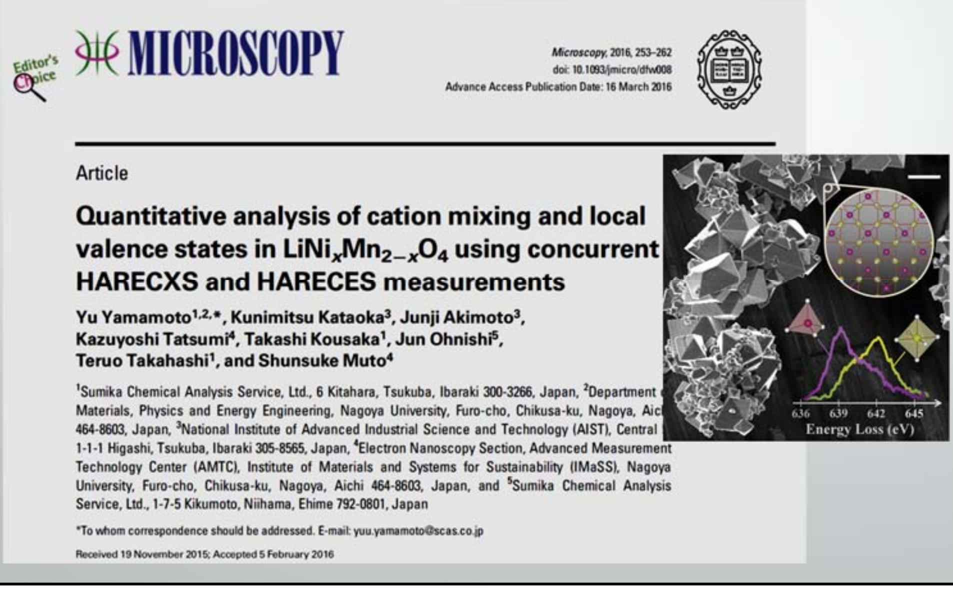
A concept of ‘Tensor Decomposition’


X
(I×J)
≈
λ1
b1
a1
+····+
λK
bK
aK
=
A
B^T
a_k
(K×J)
b_k
(I×K)
X
(I×J×M)
c1
cK
C
c_k
(M×K)

Tucker Decomposition
Canonical Polyadic Decomposition (CPD)

テンソル化の利点: ✓コンパクトな表現
✓各指標毎に独立した拘束条件設定
✓指標間の相関、重み、因果関係の設定または導出

LIB正極材料：$Li_{0.21}Ni_{0.7}Mn_{1.64}O_{4\text{-}\delta}$



Editor's Choice
MICROSCOPY
Microscopy, 2016, 253–262
doi: 10.1093/jmicro/dfw008
Advance Access Publication Date: 16 March 2016

Article

Quantitative analysis of cation mixing and local valence states in $LiNi_xMn_{2-x}O_4$ using concurrent HARECXS and HARECES measurements

Yu Yamamoto[1,2,*], Kunimitsu Kataoka[3], Junji Akimoto[3], Kazuyoshi Tatsumi[4], Takashi Kousaka[1], Jun Ohnishi[5], Teruo Takahashi[1], and Shunsuke Muto[4]

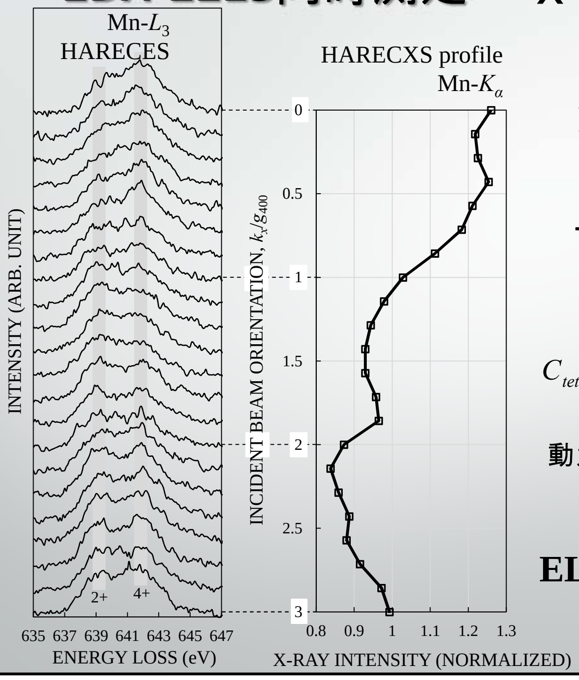
[1]Sumika Chemical Analysis Service, Ltd., 6 Kitahara, Tsukuba, Ibaraki 300-3266, Japan, [2]Department Materials, Physics and Energy Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aic 464-8603, Japan, [3]National Institute of Advanced Industrial Science and Technology (AIST), Central 1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan, [4]Electron Nanoscopy Section, Advanced Measurement Technology Center (AMTC), Institute of Materials and Systems for Sustainability (IMaSS), Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603, Japan, and [5]Sumika Chemical Analysis Service, Ltd., 1-7-5 Kikumoto, Niihama, Ehime 792-0801, Japan

*To whom correspondence should be addressed. E-mail: yuu.yamamoto@scas.co.jp

Received 19 November 2015; Accepted 5 February 2016

EDX-EELS同時測定

$X = CS^{T} \times 2$



HARECXS

↓

サイト占有率

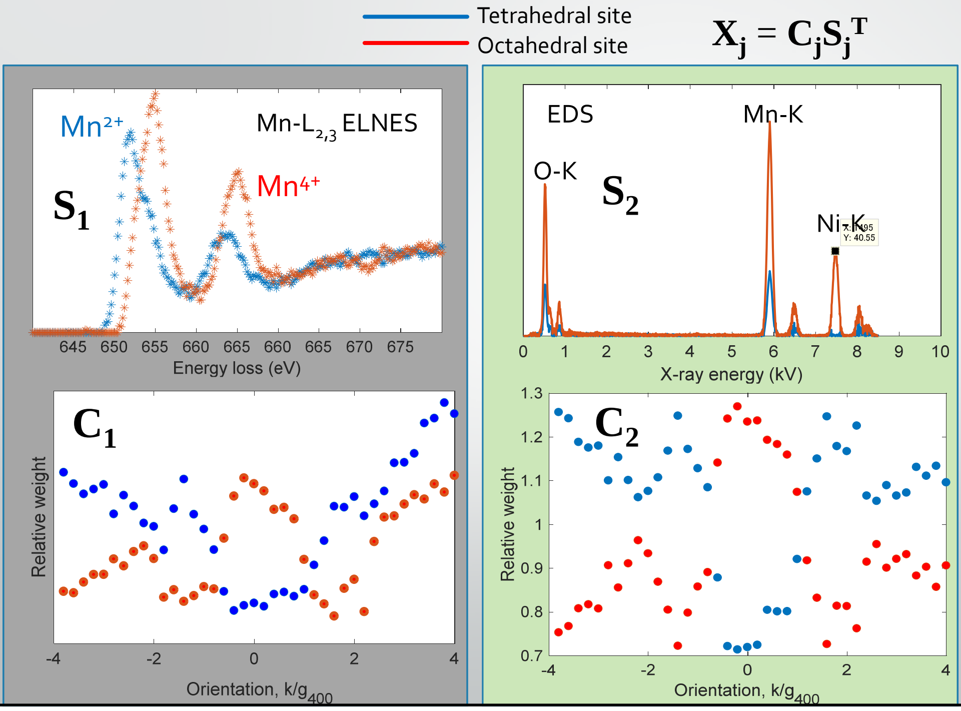
↓

$$C_{tet} = \frac{N_{tet}\sigma_{tet}}{N_{tet}\sigma_{tet} + N_{oct}\sigma_{oct}}$$

動力学電子回折理論計算

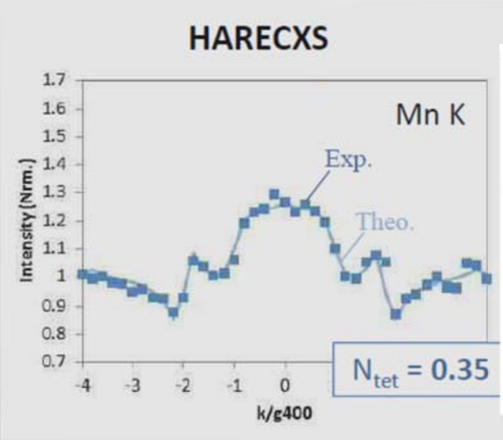
↓

ELNESプロファイル

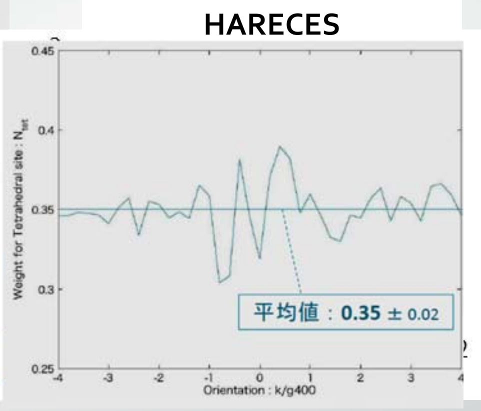


$$\mathbf{X_j = C_j S_j^T}$$

Site occupancies：定量性確認


HARECXS

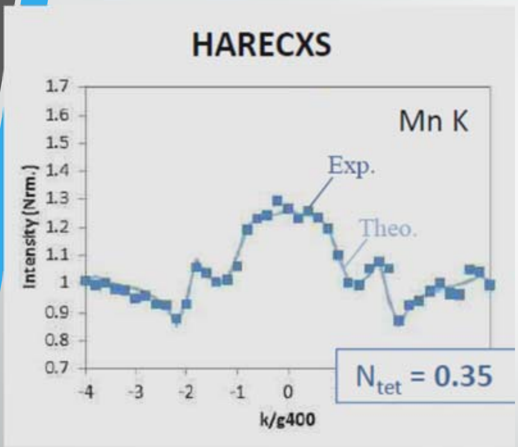
$N_{tet} = 0.35$

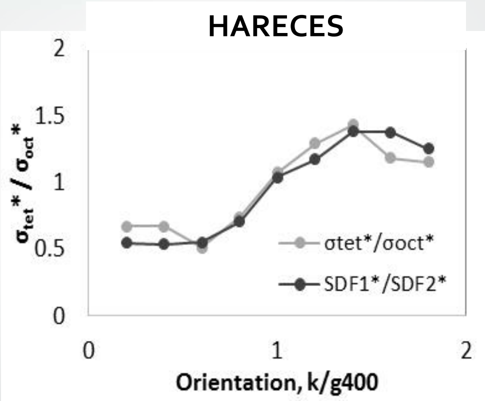

HARECES

平均値：0.35 ± 0.02

$$C_{tet} = \frac{N_{tet}\sigma_{tet}}{N_{tet}\sigma_{tet} + N_{oct}\sigma_{oct}}$$

Site occupancies：定量性確認





$$C_{tet} = \frac{N_{tet}\sigma_{tet}}{N_{tet}\sigma_{tet} + N_{oct}\sigma_{oct}}$$

まとめ：Hyperspectral Image Analysisの進展

- Two-way解析：NMF

  解の唯一性問題の解決

  ①ソフト直交性の導入

  Shiga et al, *Ultramicroscopy*, **170** (2016) 43-59.

  ②SSS+VCAによる強く重畳した場合の分解

  J. Spiegelberg et al, *Ultramicroscopy*, **submitted**.

- Two-way からMulti-way解析へ

  SDFによる複数分光データの取り扱い

装置に左右されない
高度に洗練された新しい分析技術の確立

Machine learning << Human learning

Collaborators

- Dr. Motoki Shiga (Gifu Univ., Japan)
- Dr. Masahiro Ohtsuka (Nagoya Univ., Japan)
- Jakob Spiegelberg (Uppsala Univ., Sweden)
- Dr. Jan Rusz (Uppsala Univ., Sweden)

Dr. M. Shiga
Dr. M. Ohtsuka
J. Spiegelberg
Dr. J. Rusz

Thank you for your kind attention!