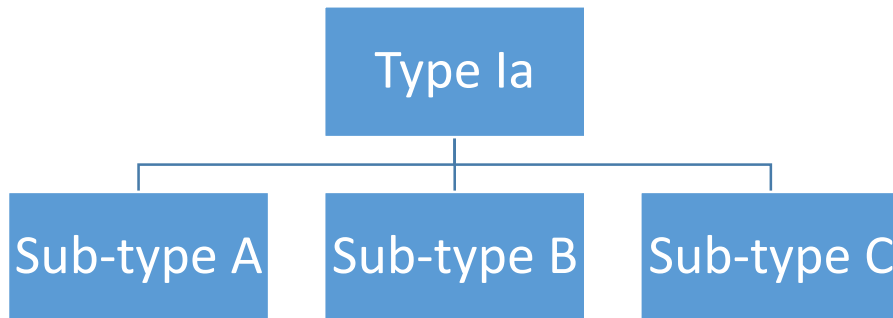


バイクラスタリングの 対話型可視化ツールを用いた Ia型超新星の分類

植村誠 (広島大学), 前田啓一 (京都大学),
Hsiang-Yun Wu(慶應義塾大学), 渡辺一帆 (豊橋技術科学大学),
高橋成雄 (会津大学), 藤代一成 (慶應義塾大学)

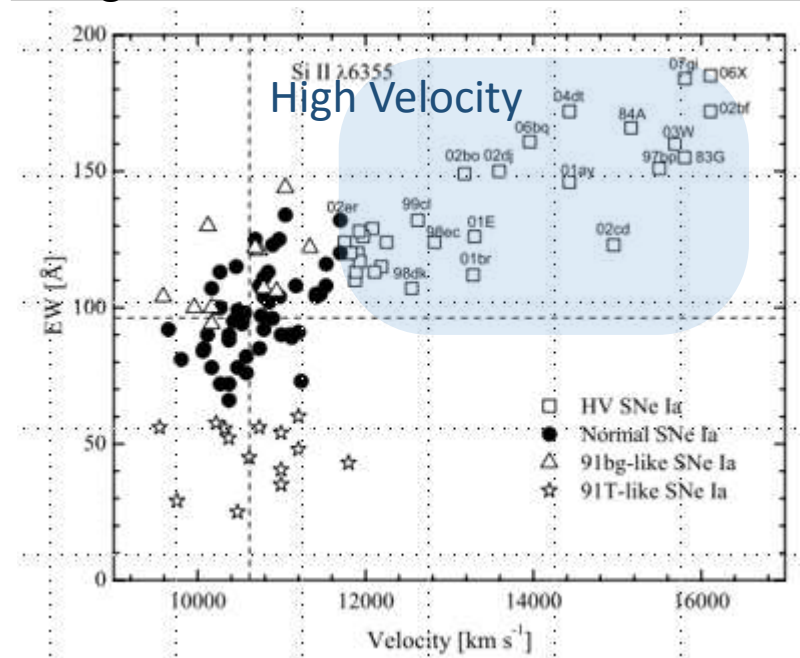
天文学会2015年秋季年会@甲南大学 2015.09.09-11

Ia型超新星の分類



- 爆発機構
- 距離指標の高精度化

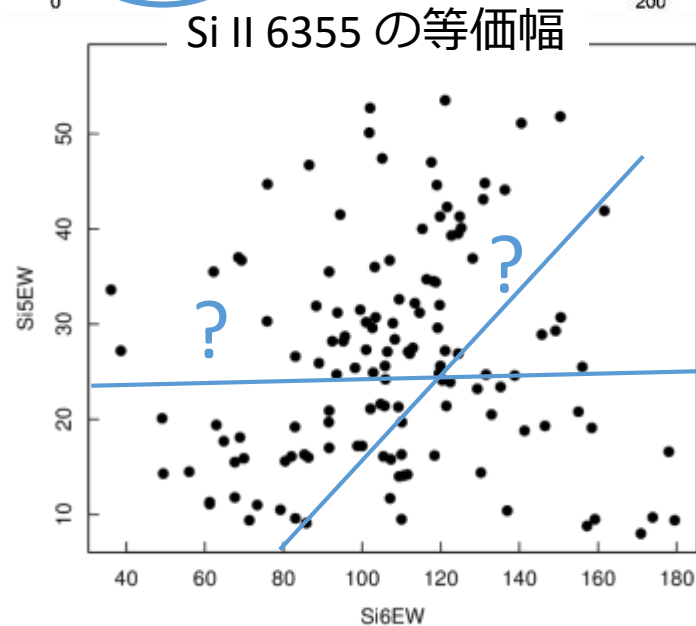
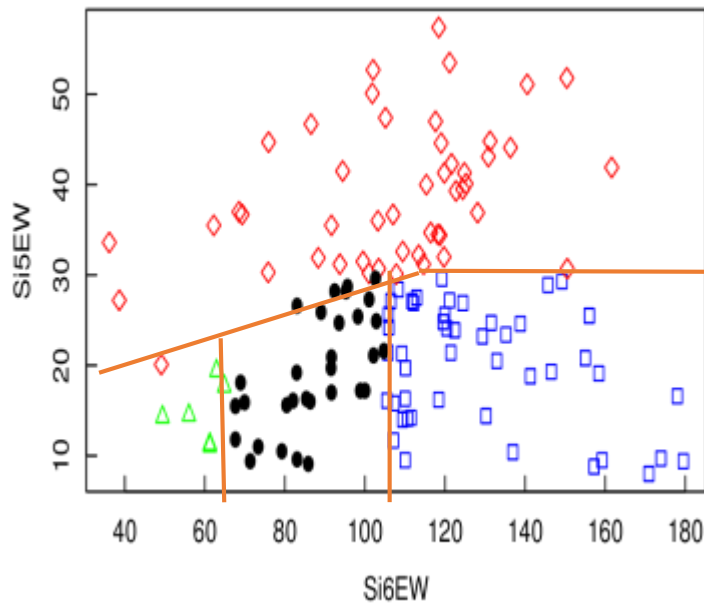
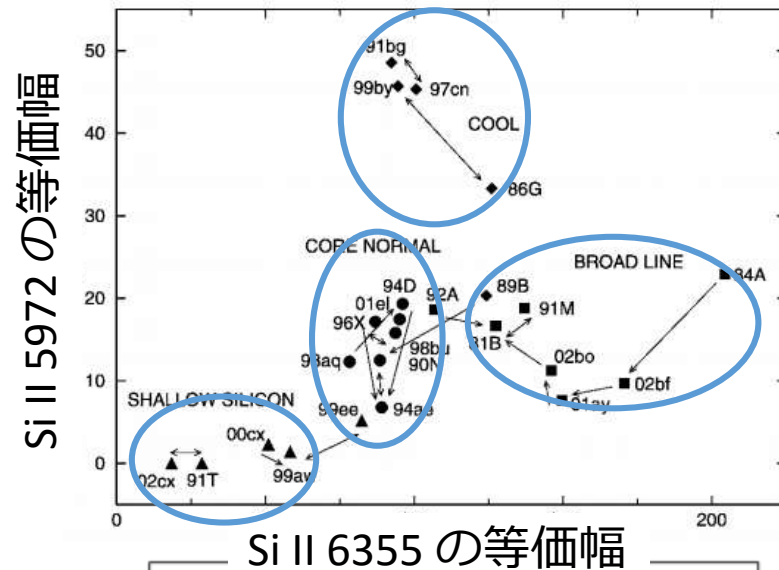
Wang, et al. 2009 の膨張速度による分類



Si II 6355 吸収線の速度

Branch, et al. 2006の分類法

- スペクトルによる分類
- Si II 6355の等価幅と Si II 5972 の等価幅で特徴づけられる。
- より多いサンプル、分類軸ではどうか？



吸収線の速度

Si II 6355

色

吸収線の等価幅

Si II 5972

極大等級

吸収線の半値幅

Si II 4000

減光速度

吸収線の深さ

Ca II, etc...

距離

どれを使う？

問題の難しさ

- どのような sub-type があるのか？
- どの軸で分類するのが妥当か？
 - 軸の候補は現在では多数
 - 以前は観測しやすい軸しか使えなかった

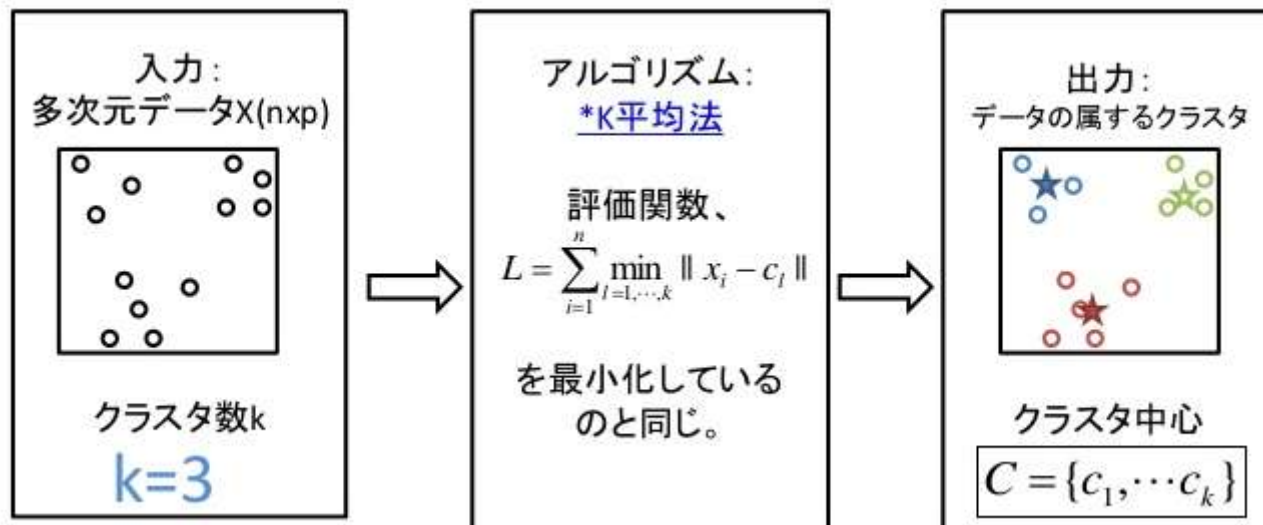
→ 機械学習 (クラスタリング) + 可視化



K-means クラスタリング

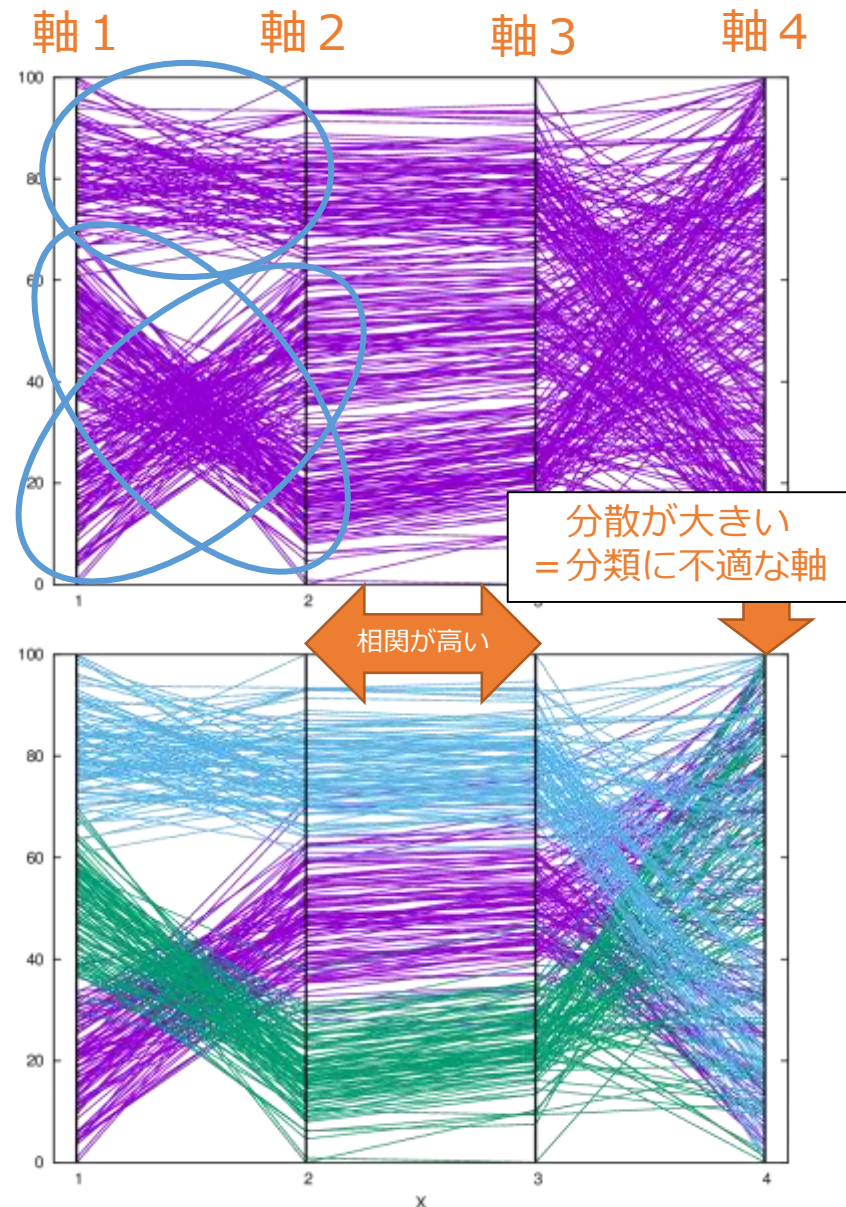
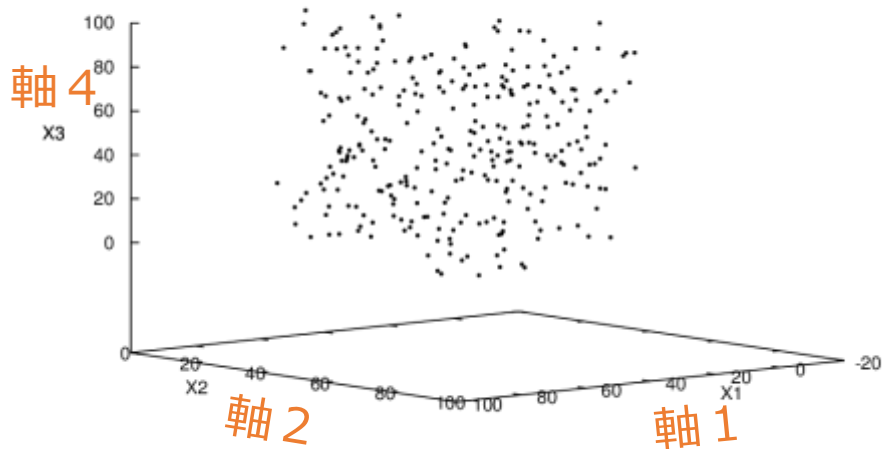
パターン認識02 “k平均法”

<http://www.slideshare.net/sleipnir002/02-kver20>



- 単純なクラスター解析だけだと必要のない軸が入っている場合、精度が落ちる
- クラスタリングしながら、結果を可視化し(visual analytics)、不要な軸を削る

平行座標表示 (データ可視化)

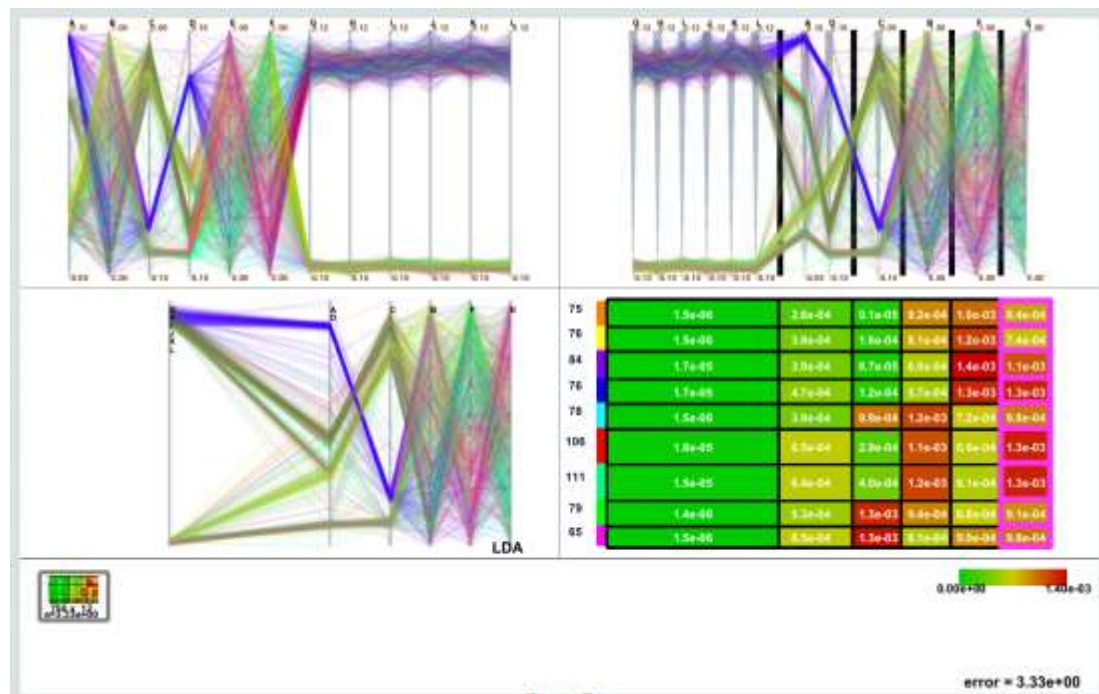


適当な人工データを使った実験

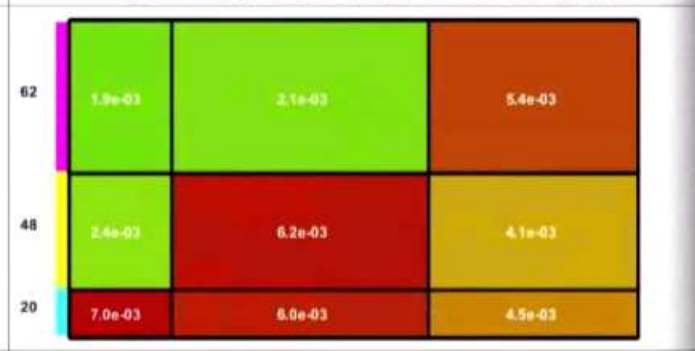
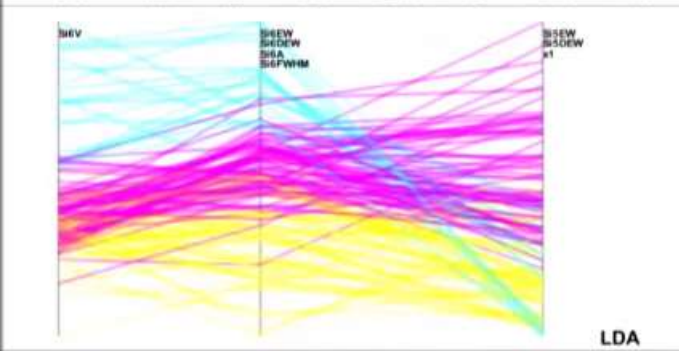
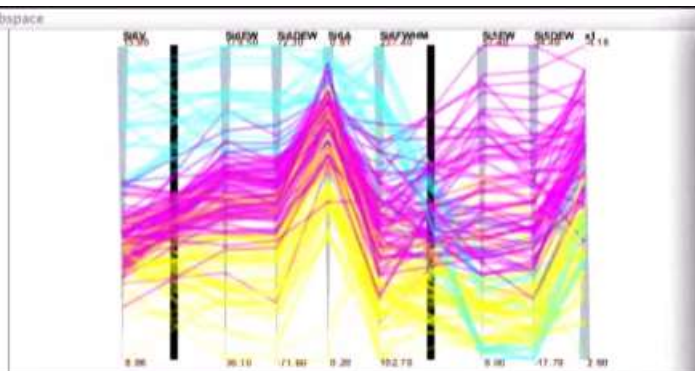
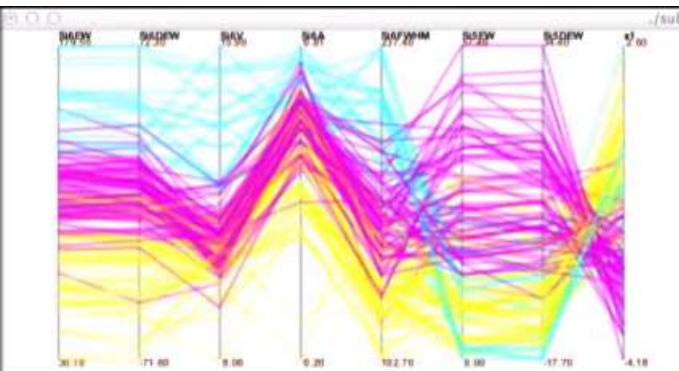
- 4つの軸があったとして、3次元までなら散布図を書いて、潜在構造を見ることが出来る。
(実際は3次元プロットは見にくい)
- 平行座標表示ならサンプルのクラスター構造や軸間の相関が見えることがある。

非対称バイクラスタリングの 対話型可視化ツール

- Watanabe, et al., IEEE Pacific Visualization Symposium 2015
- Ia型超新星の分類に応用



- 使用データ = Berkeley Supernova Database
 - Ia型132サンプル
 - 14軸：Si II 6355 等価幅(EW)、時刻補正した等価幅(DEW)、速度、深さ、半値幅、Si II 5972 等価幅(EW)、時刻補正した等価幅(DEW)、赤方偏移(z)、減光率、色、みかけの等級、絶対等級（補正なし、色補正のみ、色・減光率補正）



error = 4.44e+00

System Menu

Datasets:

Test	Iris	Spain	Titanic
Synthetic (8D)	Car	Wine	Blazar (17D)
Synthetic (12D)	Concrete	Abalone	Blazar (8D)
Food (USDA)	All Wine Qual.	Breast Cancer	Blazar (5D)
Cadaster	Rock	Gesture	Supernova (14D)

Dimension: File name: Load

Label: Message:

Window mode:

- Coordinated view
- Classical PCP
- Clustered PCP
- Contracted PCP
- Block matrix diagram

Bicluster setting:

Data Clusters: Axis Clusters: Biclustering

Block process:

- Cut columns
- Cut rows
- Merge blocks
- Save state

Data coloring:

- Clusters
- Tags

Contraction types:

- LDA
- PCA
- SPMeans

Plot mode:

- PCP
- Scatter Plot

Message:

Message list: 14, 13

Rendering styles:

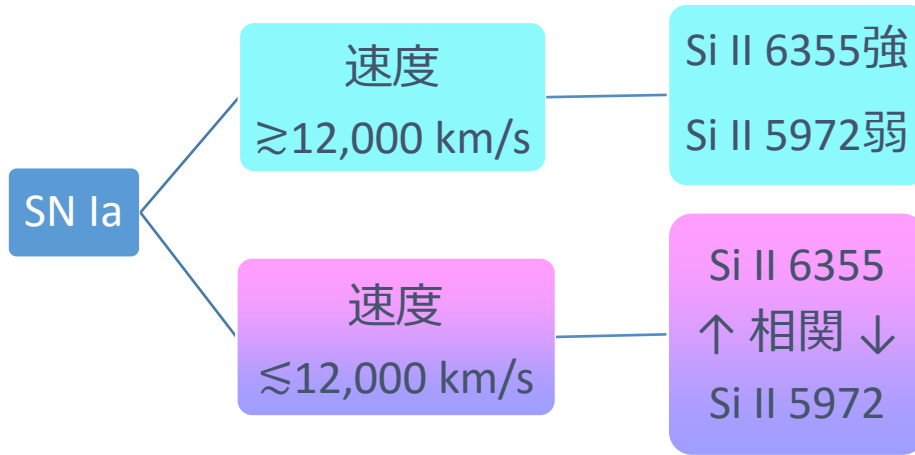
- Polylines
- Edge-bundled
- Strip rendering
- Cluster rendering

Alpha setting:

- Raise alpha
- Lower alpha

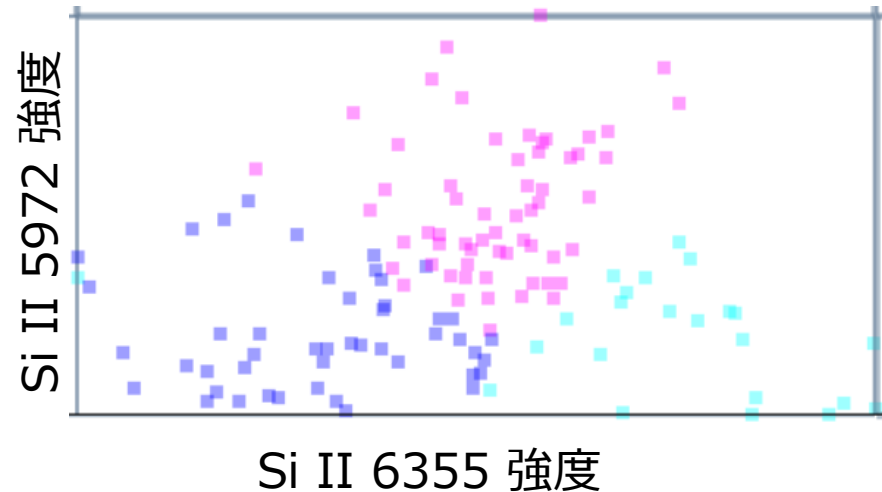
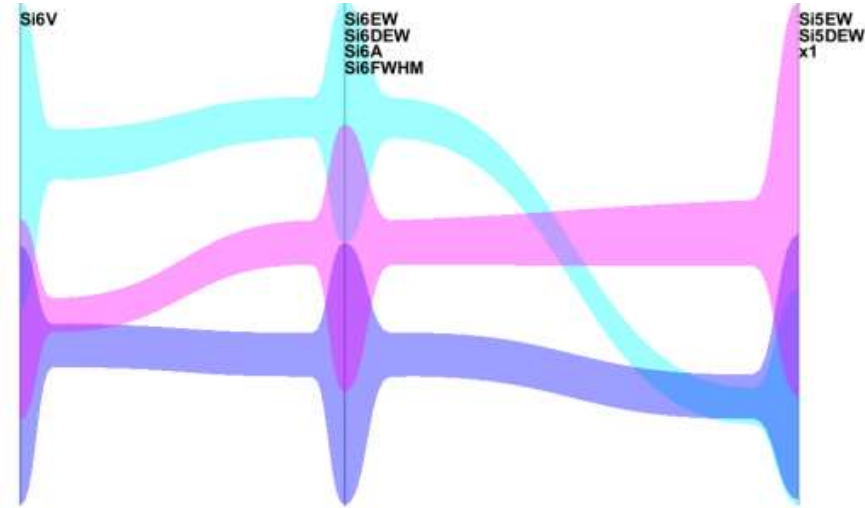
Save CSV Capture Quit

まとめ



- 従来分類軸と一致
 - Branch+06 (Si II 6355, 5972 EW)
 - Wang+09 (Si II 6355 velocity)

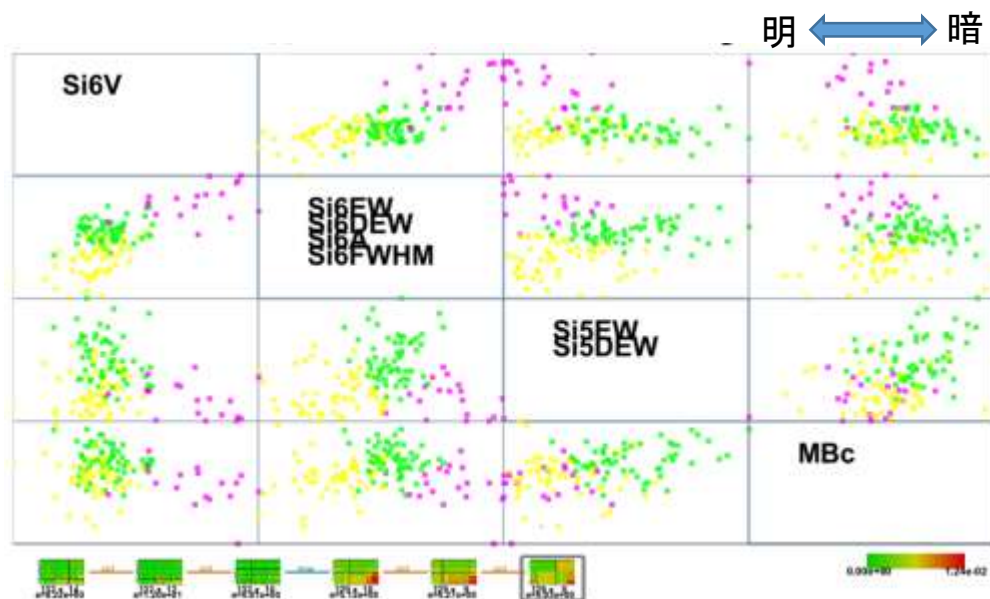
- より多くのデータ、より多くの分類軸候補から機械学習+可視化技術によって、(客観的に) 従来分類の妥当性を検証し、支持する結果。



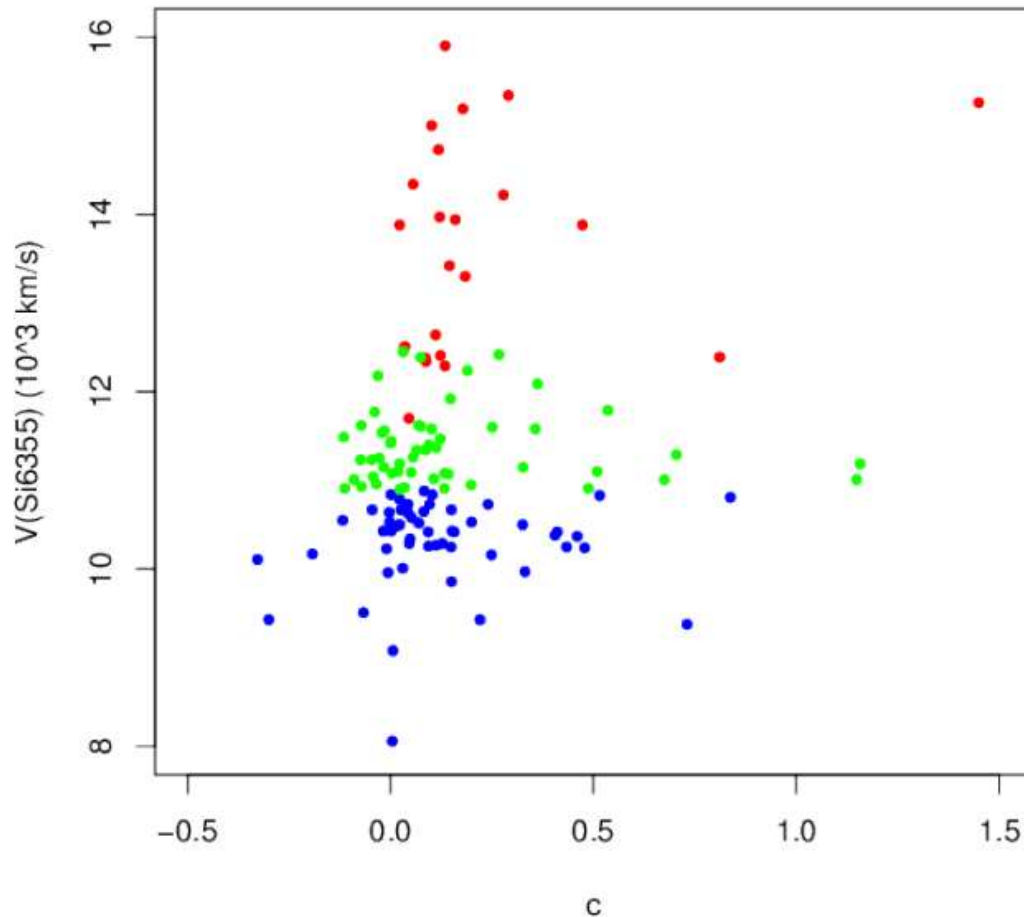
Backup slides

絶対等級での分類について

- 右図：絶対等級を残して解析した結果
- 絶対等級は独立した良い分類軸にはならない
 - 絶対等級は Si II 5972強度と緩い相関
 - Si II 5972強度が同程度のサンプルで見ると、高速度クラスと通常クラスで絶対等級に有意な差はない。
 - sub-luminous な91bg-likeなサンプルは7個と少数だったので、全体に影響しなかった可能性



色”c”と Si II 6355の速度 “V(Si6355)” の散布図



赤:可視化ツール”subspace”で同定された high velocity cluster (HV)。

緑:HVを除いたサンプルで V(Si6355)の中央値よりも大きいもの

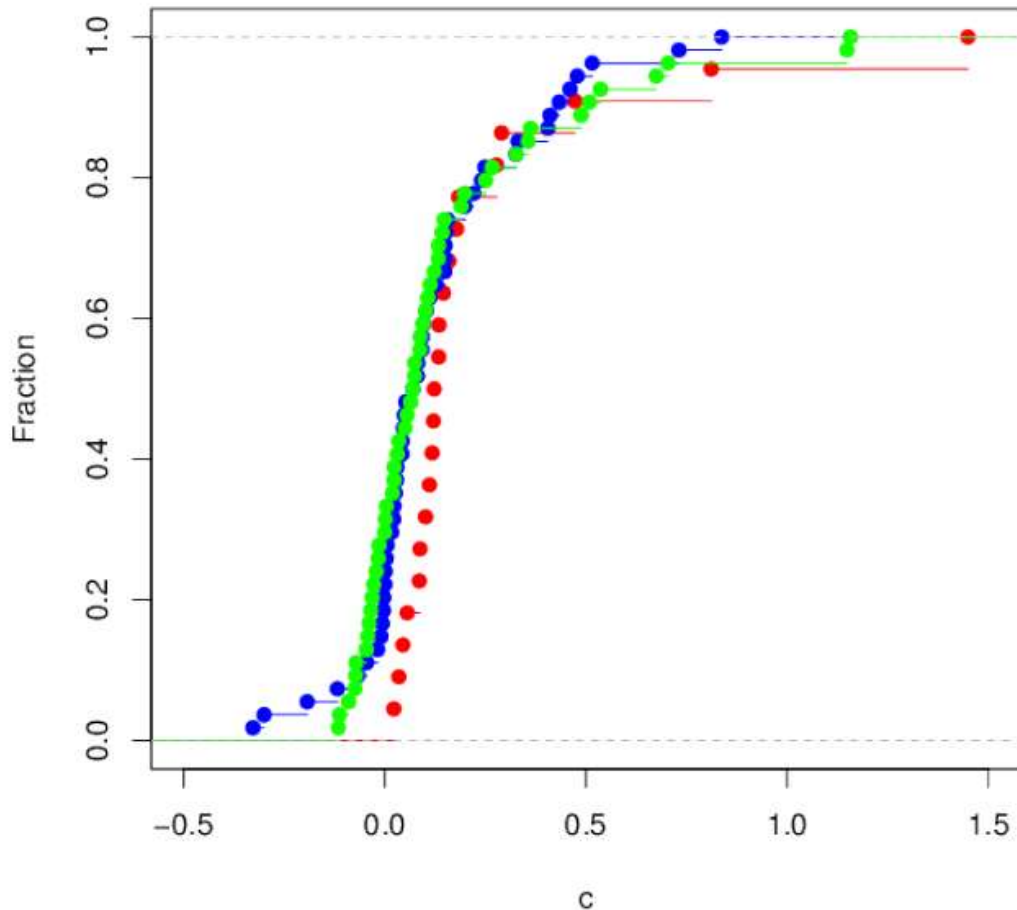
青:HVを除いたサンプルで V(Si6355)の中央値よりも小さいもの

*V(Si6355) と c の間に明確な相関は見られないが、c の分布に差はありそう

->c の累積分布+KS検定(次ページ)

色”c”の累積分布

(Foley & Kasen 2011 の Figure 5 に相当)



赤:可視化ツール”subspace”で同定された high velocity cluster (HV)。

緑:HVを除いたサンプルで V(Si6355)の中央値よりも大きいもの

青:HVを除いたサンプルで V(Si6355)の中央値よりも小さいもの

*2つの分布が同じ母集団からのものである有意確率を調べる(KS検定)

結果:

赤と(緑+青):1.2%

赤と緑:1.5%

赤と青:2.9%

緑と青:59.9%

*HV(赤)が他の天体に比べて系統的に赤い(cが大きい) = Foley&Kasen 11 の結果と consistent

*緑(比較的速度大)と青(比較的速度小)に有意な分布の差は見られない。

→ 速度と共に徐々に色が赤くなるのではなく、velocity>12,000km/s で性質が変わる。← EW(Si6355) v.s. V(Si6355) と同じ傾向