

スパースモデリングの 宇宙物理学への応用

植村誠（広島大学）

@国立天文台談話会 2014.05.23

いきなりですがデモをします

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$
$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

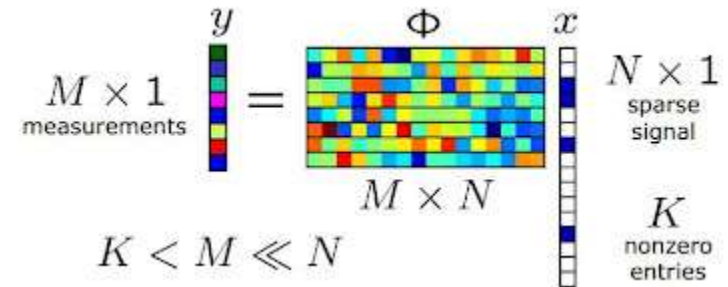
- 問題：連立一次方程式で変数の数が式の数より多い場合
 - $M < N$ だと解が1つに決まらない
 - $M \geq N$ だと解ける
 - x はあらかじめ答えを設定、 A の要素は乱数を入れて実験
 - x が100次元、 y が50次元

少ない情報からより多くの情報を推定できた

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

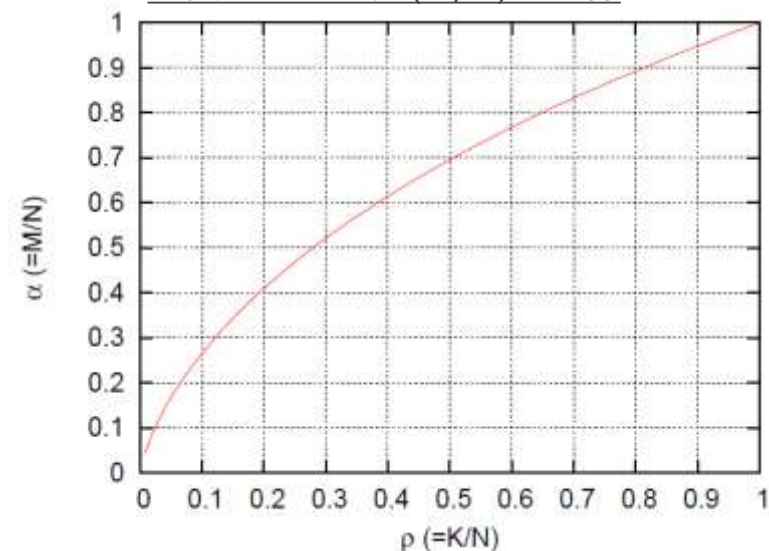
$$\|\mathbf{x}\|_1 = \sum_i |x_i|$$

N個の内K個しか≠0な値をもっていない
K個よりも多いデータ数M



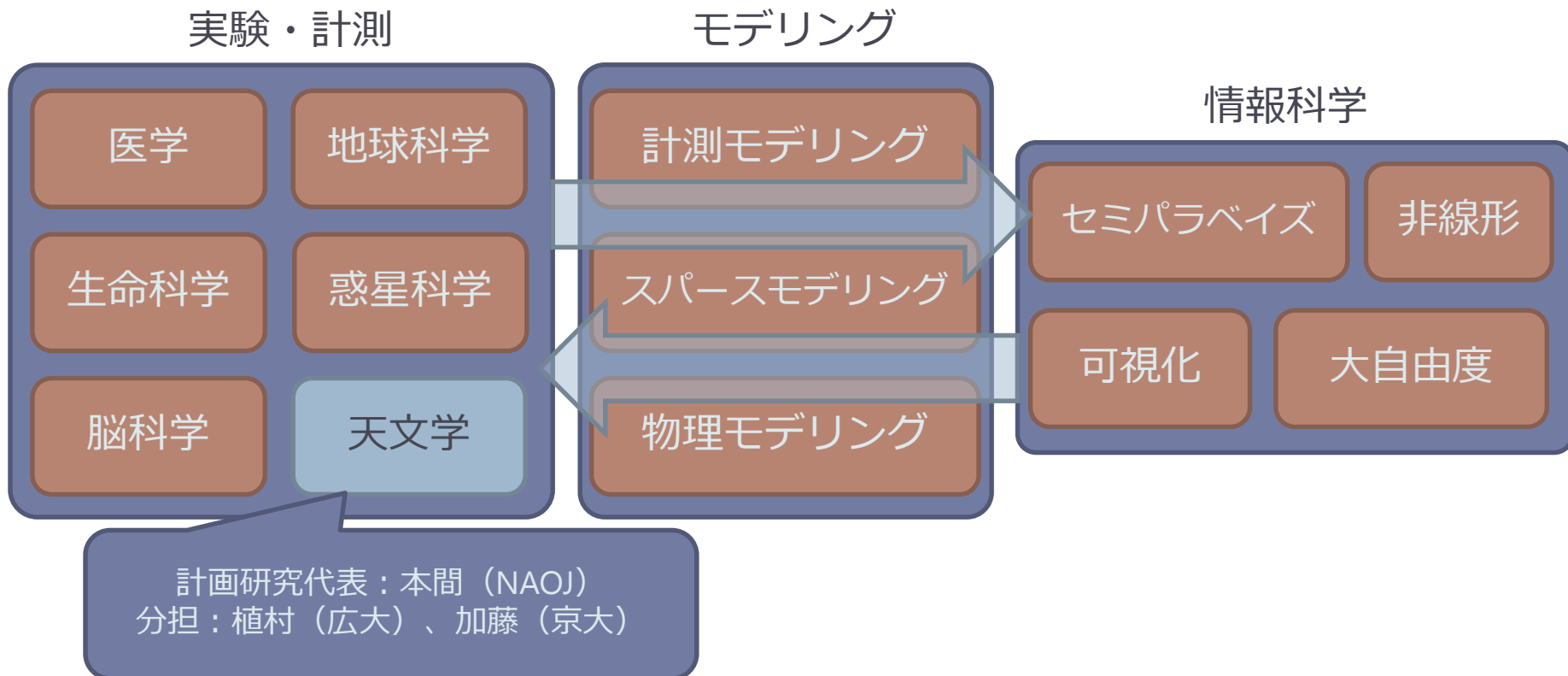
- **スパースな元信号は1次ノルム最小化で再構成できる**
 - 圧縮センシング：圧縮された情報から元信号の完全再構成など（ y の測定誤差は考えないことが多い）（Candes & Tao 2006, Donoho 2006）
 - LASSO： y に測定誤差があり、 A が既知の行列の場合の回帰（Tibshirani 1996）
 - x そのものがスパースでなくても、 $x' = Bx$ で x' がスパースなら（スパースな表現 x' が存在すれば）使える
- **スパースモデリング = 情報の「疎性」を利用することでデータからより多くのand/or 正確な情報を引き出すフレームワーク**

スパースな度合(K/N)と完全再構成に必要なデータ数(M/N)の関係



宇宙物理データにも応用できないか？

- 科研費 新学術領域
「スパースモデリングの高次元データ駆動科学の創成」
- 領域代表：岡田真人（東京大学・大学院新領域創成科学研究科・教授）
- 平成25～29年度



今日の話のアウトライン

- スパースな元情報を少ないデータから高精度に再構成する
 - 変光星の周期解析
 - 電波干渉計
 - ドップラー・トモグラフィー
- データから客観的にモデル（パラメータ）を決定する
 - Ia型超新星の極大光度
- まとめ

変光星の周期解析

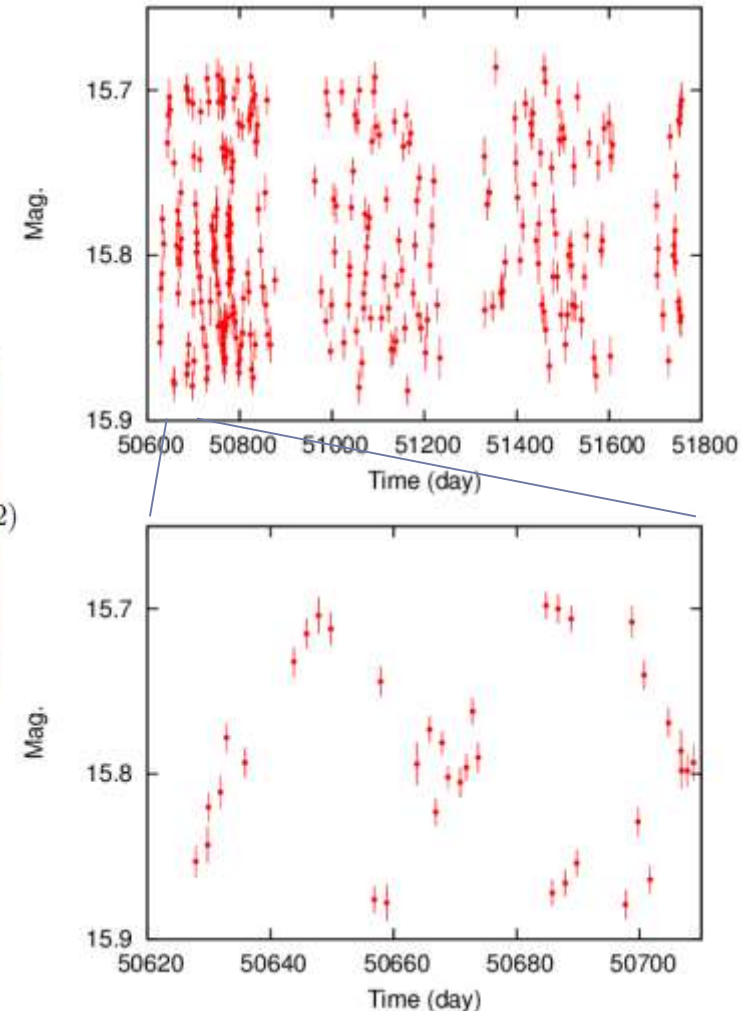
- 観測データから周期を決めたい
- データは不等間隔 → 複雑な窓関数
- 脈動星のように「せいぜい複数周期の重ね合わせ」であれば、周波数空間でスパース

$$\mathbf{y} = \mathcal{F}\mathbf{x}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} \cos(t_1\nu_1) & \cdots & \cos(t_1\nu_N) & \sin(t_1\nu_1) & \cdots & \sin(t_1\nu_N) \\ \cos(t_2\nu_1) & \cdots & \cos(t_2\nu_N) & \sin(t_2\nu_1) & \cdots & \sin(t_2\nu_N) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(t_M\nu_1) & \cdots & \cos(t_M\nu_N) & \sin(t_M\nu_1) & \cdots & \sin(t_M\nu_N) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \\ b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \quad (2)$$

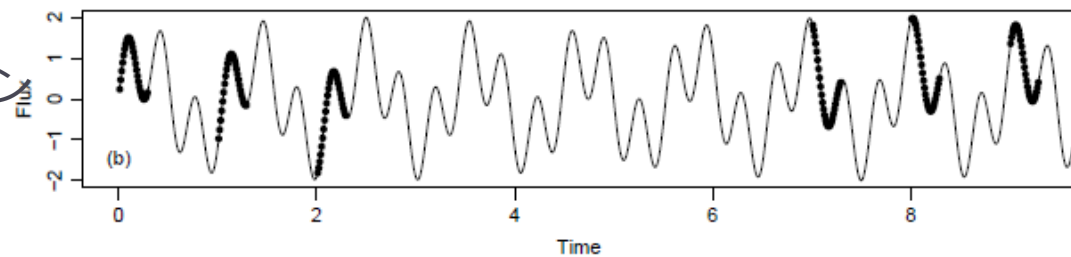
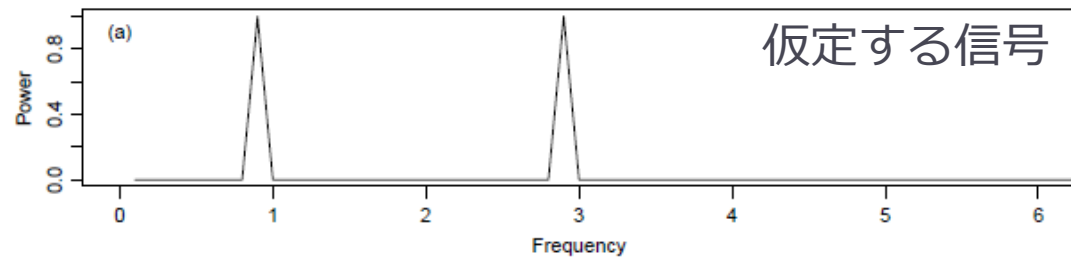
$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathcal{F}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

地上観測で得られる光度曲線の例

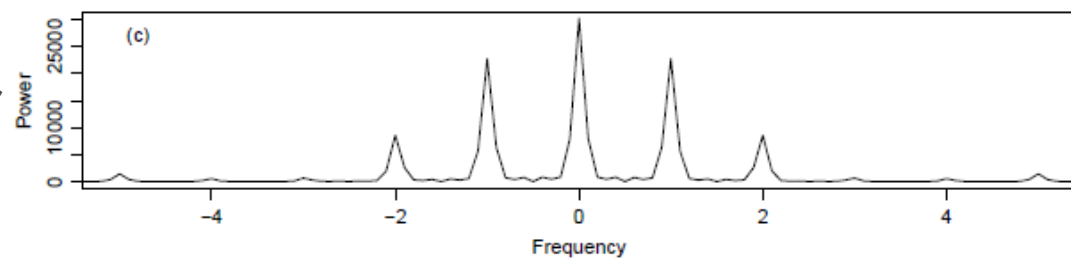


LASSO的周期解析 (実験)

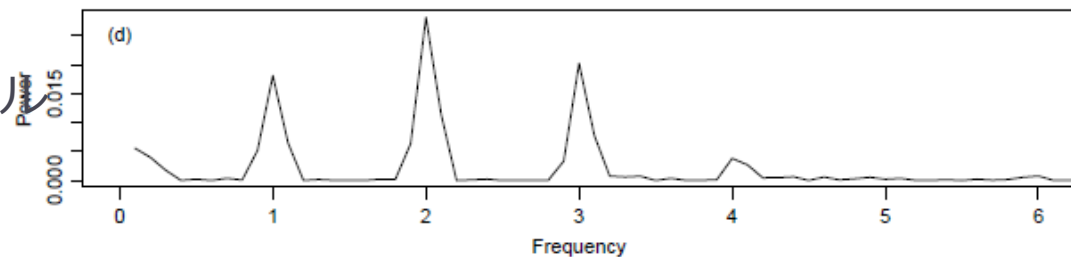
観測のシミュレーション



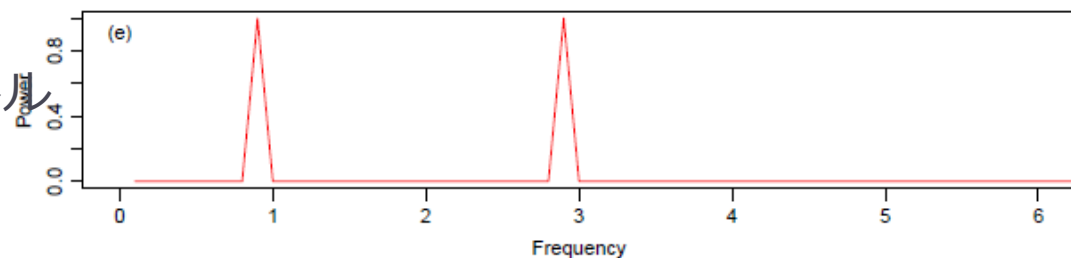
窓関数のパワースペクトル



観測から計算したパワースペクトル



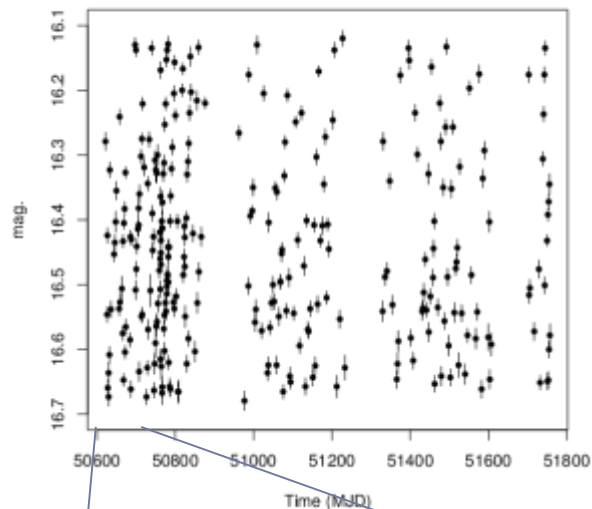
L1最小化で計算したパワースペクトル



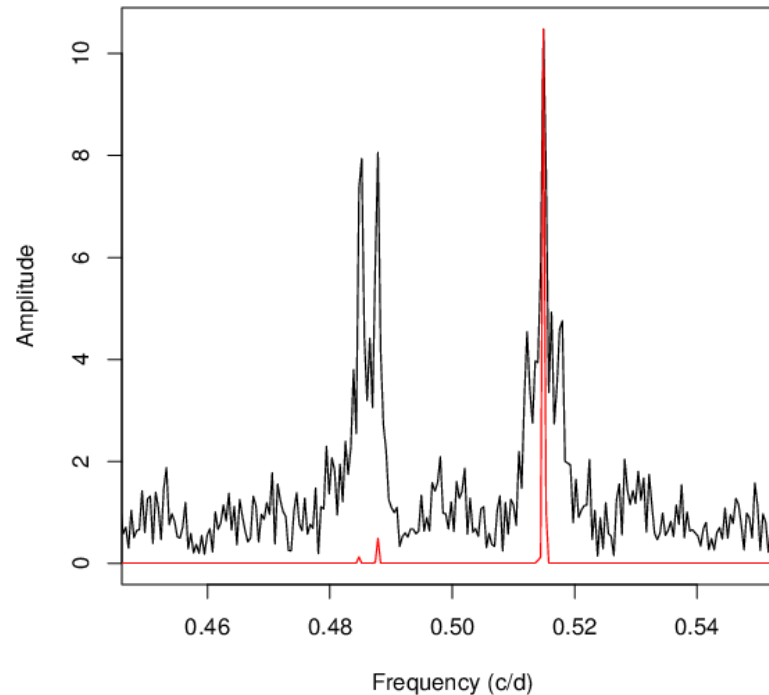
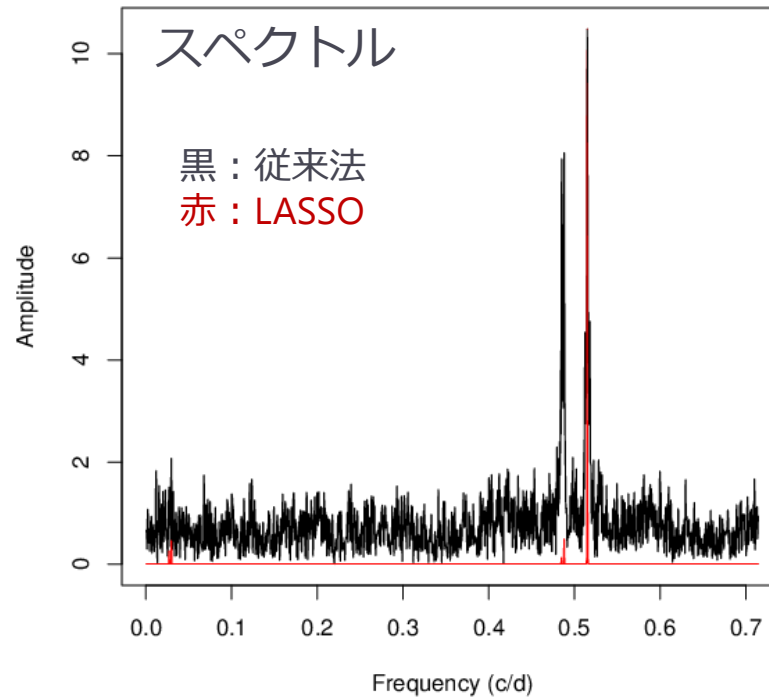
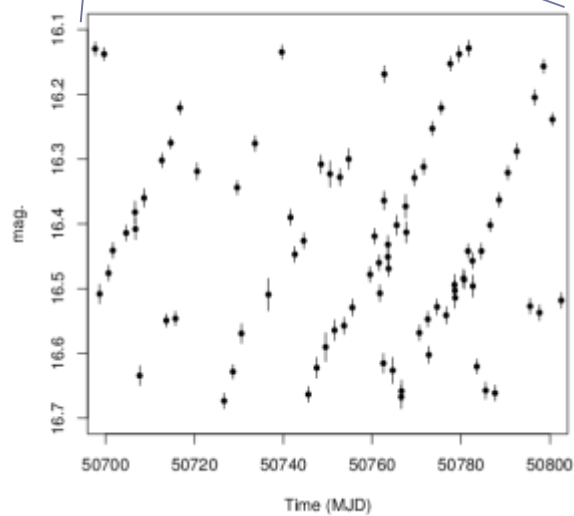
実例

8/22

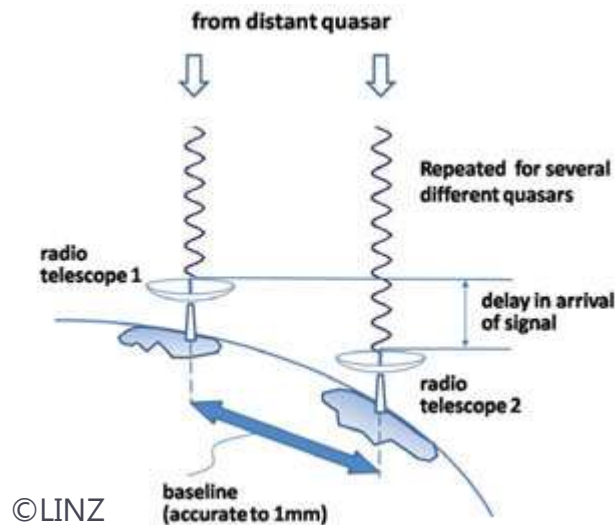
光度曲線



推定

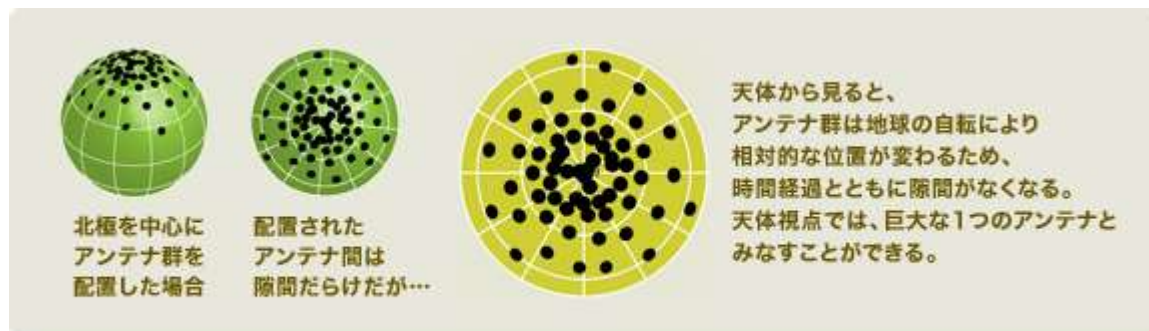


電波干渉計の画像再構成



©LINZ

- 望遠鏡の解像度（角度分解能）
 $\Delta\theta = \text{波長} / \text{望遠鏡口径}$
- 電波干渉計 = 複数の電波望遠鏡で受ける電波を干渉させ、高い角度分解能を得る
- 2次元フーリエ変換



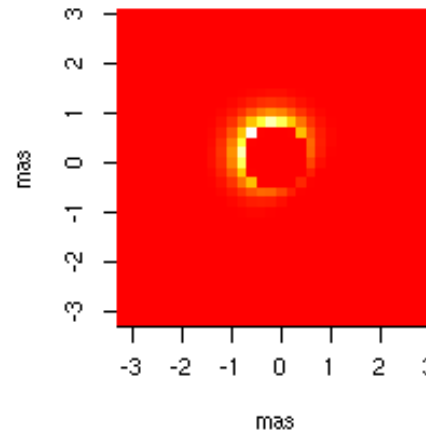
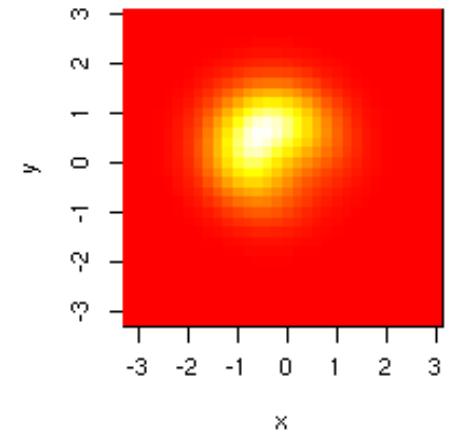
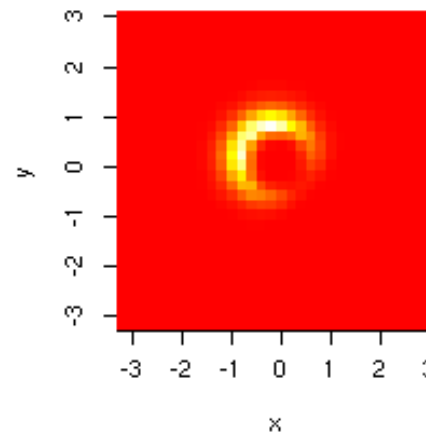
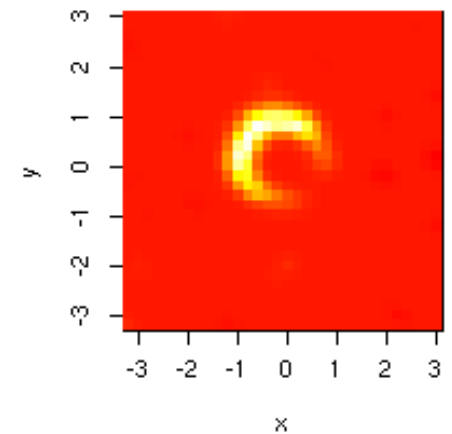
※：実際のアルマ望遠鏡の位置とは異なります。また、これは直観的想像をイメージしたものであり、実際の原理とは少し異なります。

©ALMA

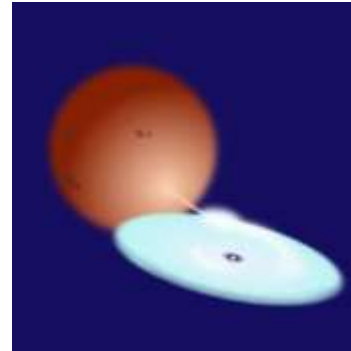
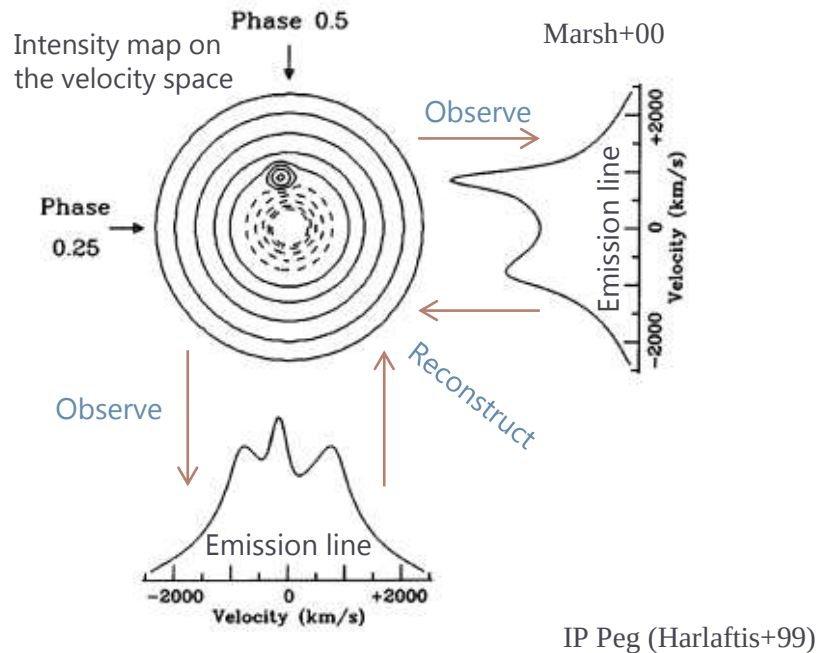
実験

Honma+14

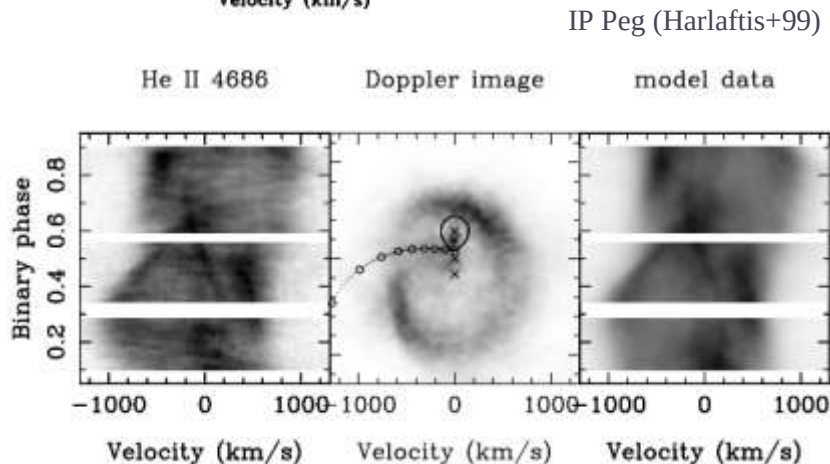
- 2次元平面内で電波源がスパース
- 従来法よりも高い分解能
- 広がった電波源が課題
 - Total Variation Minimization (TVM)やWavelet基底など検討

initial image*beam-convolved initial image**super-resolution initial image**super-resolution reconstructed image*

降着円盤の構造再構成



- Doppler Tomography = 回転する円盤を様々な方向から観測することで再構成する
 - 医療用X線CTと同じ原理
- 従来：最大エントロピー法
- 期待される円盤構造 = 衝撃波など
- Total Variation Minimization (TVM) を使う



ドップラートモグラフィの問題

- 光学的にも幾何的にも薄い円盤放射を仮定すると線形の問題に
 - 自己遮蔽などが効くとダメ
 - 多くはこの仮定でOK

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}}_{\text{Data}} - \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\text{Observation Matrix}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{image}} \right\|_2^2 + \lambda f(\mathbf{x})$$

- Maximum Entropy Method (MEM)
 - Standard method to date

$$S = - \sum_{i=1}^M p_i \ln \frac{p_i}{q_i}.$$

$$q_i = \frac{D_i}{\sum_{j=1}^M D_j},$$

- Best for real Doppler maps?
 - Hot spot and/or shock region may have sharp edges, making entropy low

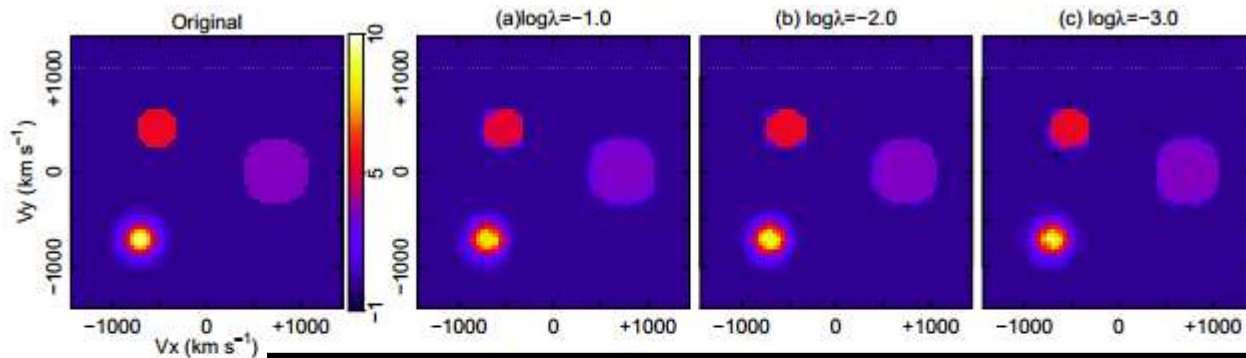
- Total Variation Minimization (TVM)

- Simple prior
- Regularization:

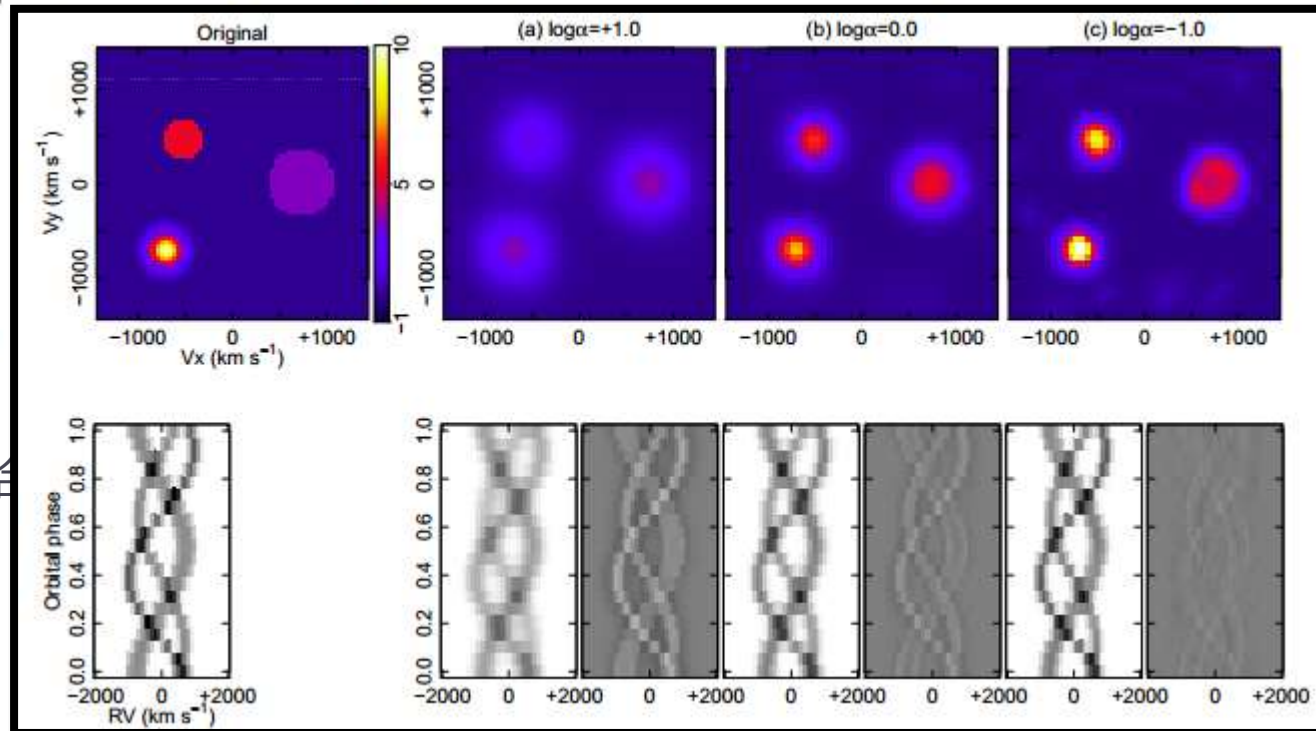
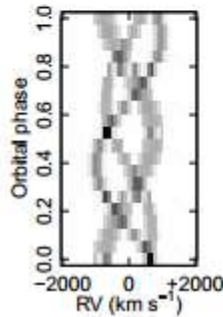
$$TV(\mathbf{x}) = \sum \sqrt{(\Delta^h \mathbf{x})^2 + (\Delta^v \mathbf{x})^2}$$

- Δx : differential operator = $x_{i+1} - x_i$
- Sparse gradient

TVMで解く Doppler tomography

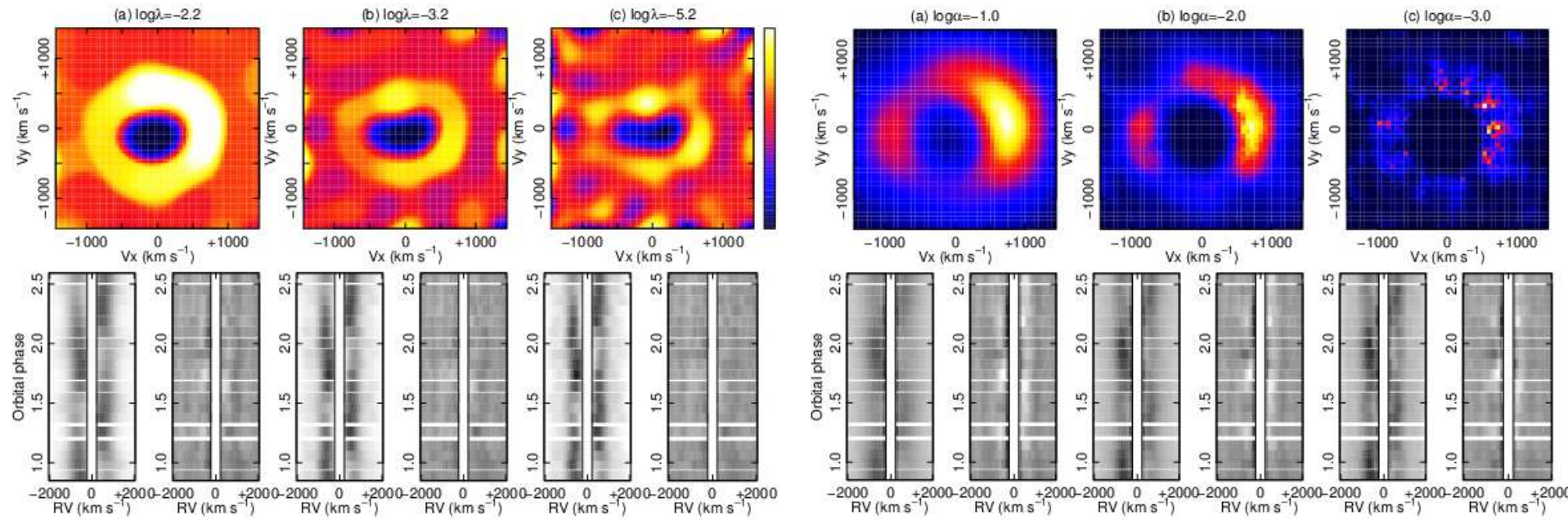


Uemura+14



- データが少ない場合

矮新星 WZ Sge (増光時)



TVM

- ✓ 降着円盤全体が楕円にひしゃげた形状
- ✓ 伴星付近と円盤の一部に局所的な構造
- ✓ データとの残差が小さい

MEM (コードは Spruit+88を使用)

- ✓ 円盤全体はほぼ円
- ✓ 2本腕のスパイラル構造
- ✓ 腕内部は滑らかな構造
- ✓ 回転する残差成分

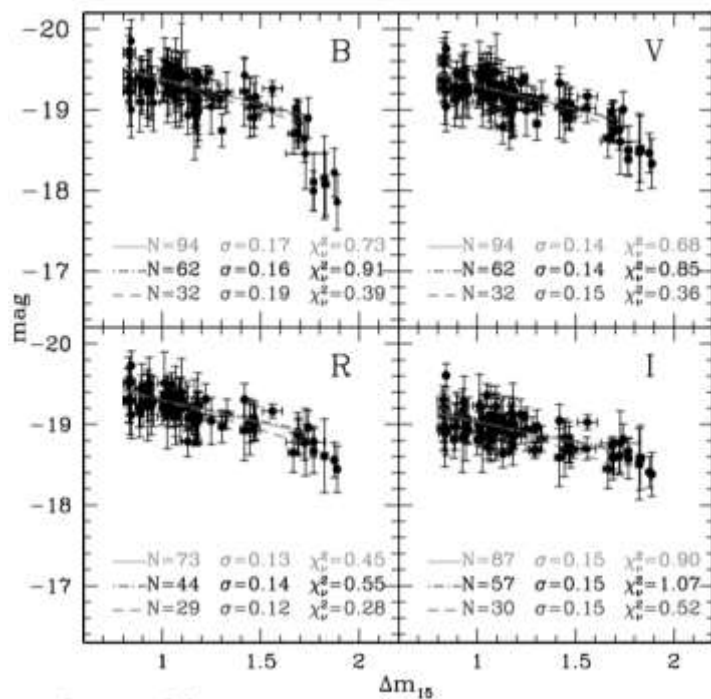
MEMでは再構成しきれなかった局所的な構造をTVMでは再構成可能

今日の話のアウトライン

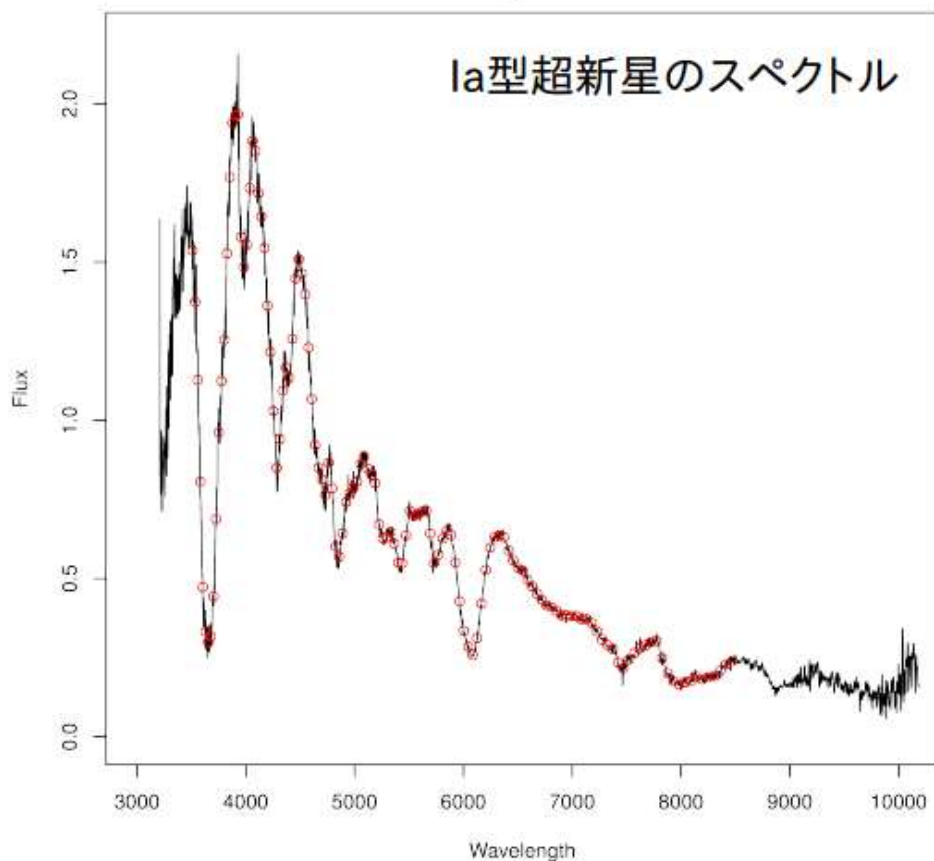
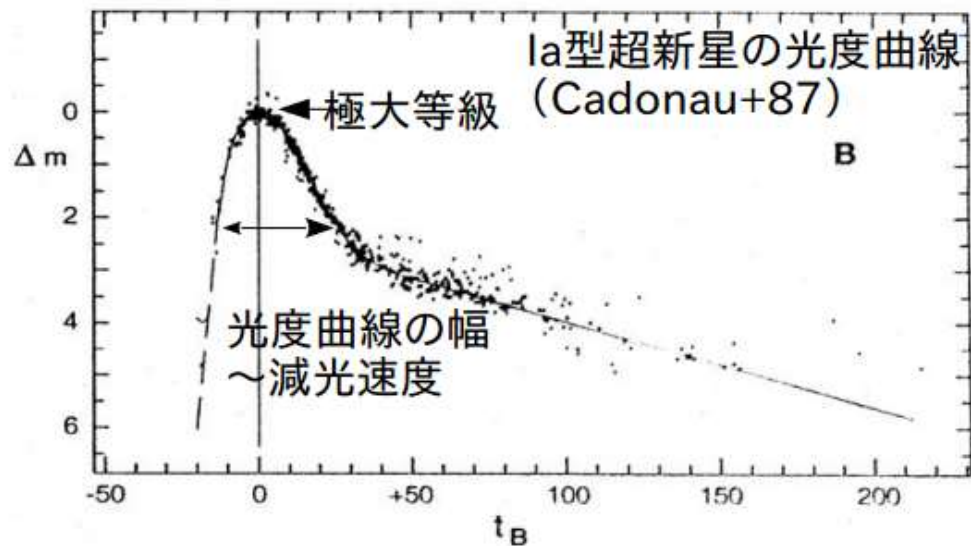
- スパースな元情報を少ないデータから高精度に再構成する
 - 変光星の周期解析
 - 電波干渉計
 - ドップラー・トモグラフィー
- データから客観的にモデル（パラメータ）を決定する
 - Ia型超新星の極大光度
- まとめ

Ia型超新星の極大等級

- 色（= 星間吸収量）と減光速度の関数
 - Phillips 93, Hamuy+96, Prieto+06
- それらを補正することでより正確な距離指標に
- しかし、まだ極大等級ばらつきがある？
- 極大等級の他の説明変数を探す



Prieto+06

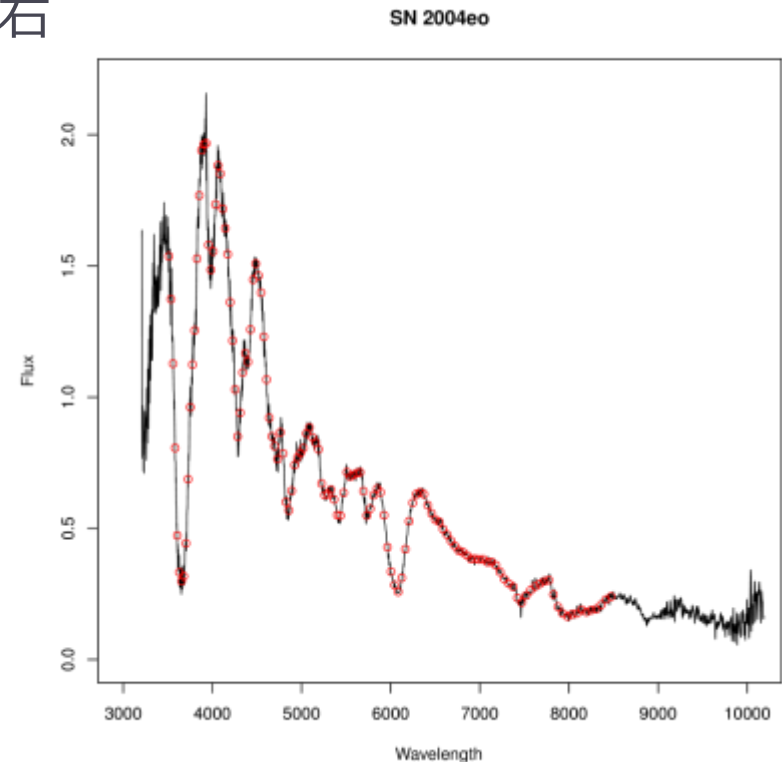


絶対等級を決めるパラメータは何か？

減光速度と色を考慮しても、まだ有意にバラついている → 他にも絶対等級を左右する要因があるのでは？

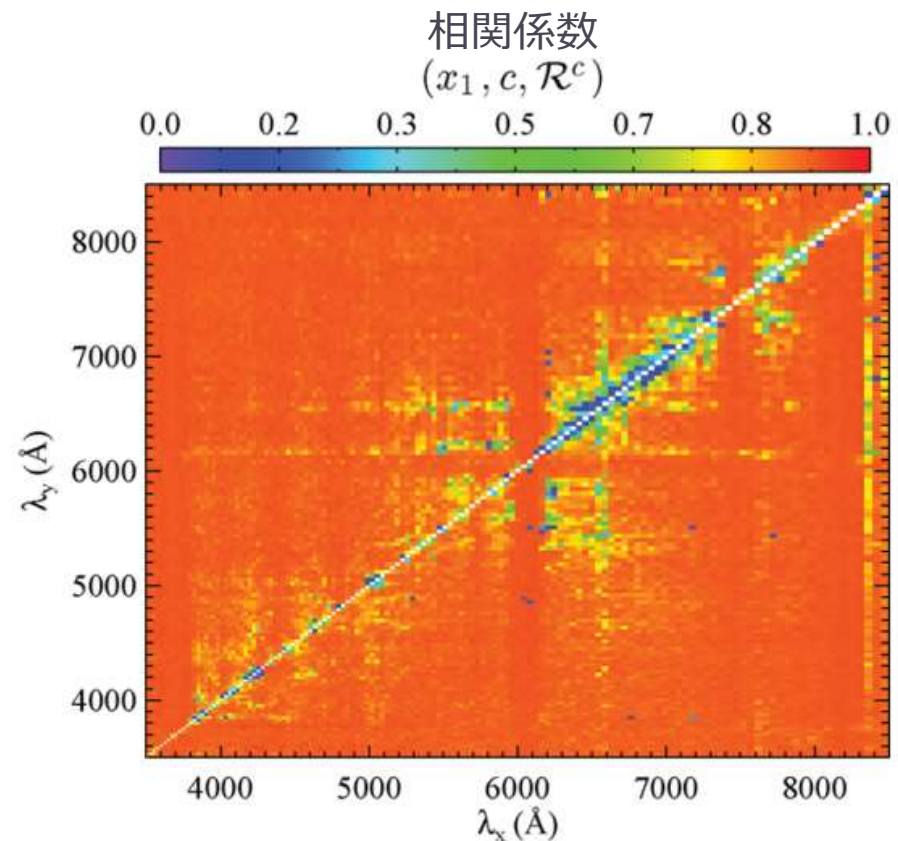
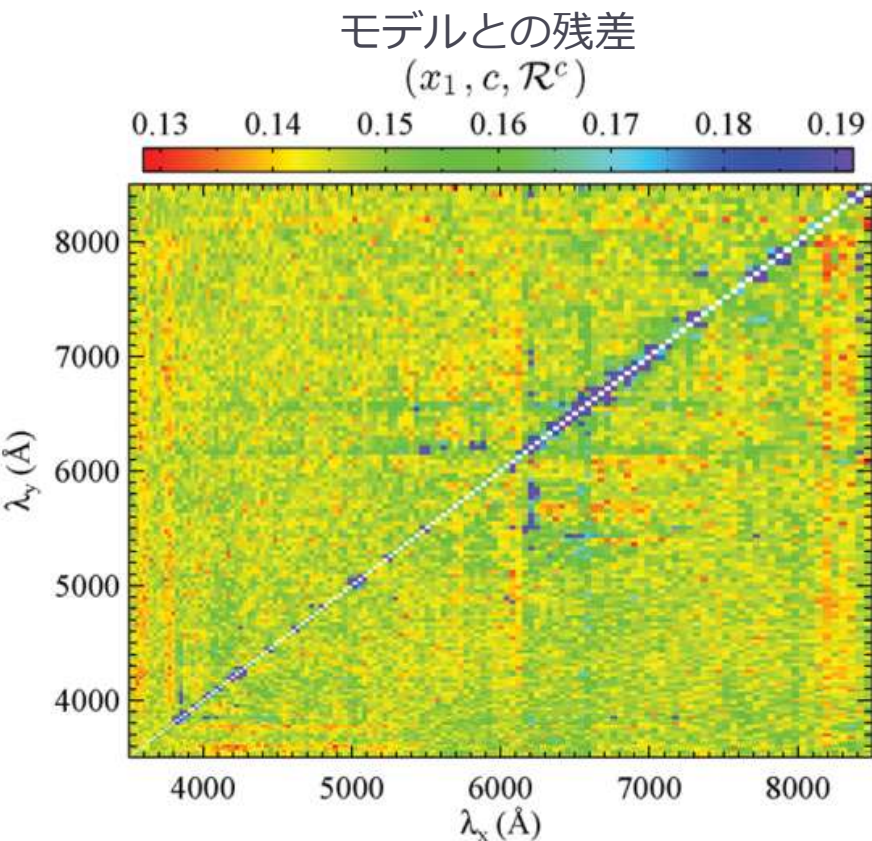
以下、Silverman+12 (BSNIP III) のレビューより。

- Si II 6355 の速度 (Blondin+11)
- Ca II H&K の速度 (Foley&Kasen 11)
- S II “W” の青側の深さ (Blondin+11)
- Si II 4000 の等価幅 (EW) (Arsenijevic+08, Walker+11, Chotard+11, Nordin+11, Walker+11)
- Fe II, Mg II の等価幅 (Nordin+11) 色との相関
- Si II 5972, 6355 の等価幅 (Hachinger+06, Nordin+11)
- Si II EW ratio $EW(5972)/EW(6355)$, Ca II H&K flux ratio (Fr/Fb) (Nugent+95, Hachinger+06)
- SiS flux ratio $Fr(S\ II\ “W”)/Fr(Si\ II\ 6355)$ (Bongard+06)
- SSi EW ratio $EW(S\ II\ “W”)/EW(Si\ II\ 5972)$, SiFe EW ratio $EW(Si\ II\ 5972)/EW(Fe\ II)$ (Hachinger+06)
- 任意の flux ratio から絶対等級と相関の高いものを探す (Bailey+09)



任意のflux比を考える

- Bailey+09, Silverman+12 (BSNIP III)
- $\lambda_{3500-8500}$ を134分割して、134x133通りのflux比 $R(\lambda_1/\lambda_2)$ を計算
- 絶対等級MBと相関の高いRを探す
- $R(3750/4550)$ 、光度曲線の幅(x_1)、色(c)の組み合わせが最も良い (BSNIP III)



要はこういう問題

$$\begin{pmatrix} MB_{\text{SN1994S}} \\ MB_{\text{SN1995E}} \\ \vdots \\ MB_{\text{SN2008s1}} \end{pmatrix} = \begin{pmatrix} x1_{\text{SN1994S}} & c_{\text{SN1994S}} & EW_{\text{SiII4000,SN1994S}} & FWHM_{\text{SiII4000,SN1994S}} \\ x1_{\text{SN1995E}} & c_{\text{SN1995E}} & EW_{\text{SiII4000,SN1995E}} & FWHM_{\text{SiII4000,SN1995E}} \\ \vdots & & & \\ x1_{\text{SN2008s1}} & c_{\text{SN2008s1}} & EW_{\text{SiII4000,SN2008s1}} & FWHM_{\text{SiII4000,SN2008s1}} \end{pmatrix} \begin{pmatrix} c_{x1} \\ c_c \\ c_{EW} \\ c_{FWHM} \\ c_{3535/3512} \\ c_{3558/3512} \\ \vdots \\ c_{8416/8472} \end{pmatrix}$$

- 線形モデルで解きたいけど、サンプル数がパラメータの数より小さいので、普通には解けない
- でも、本来必要なパラメータは数個のはず。変数をデータから選択したい。
- スパースモデリングで解こう！

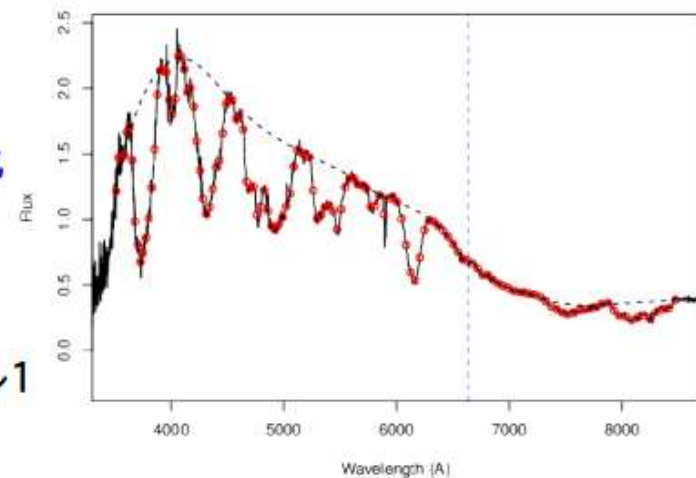
- 1 次ノルム最小化 $\hat{\mathbf{x}} = \arg \min \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$
- データは UC Berkeley のチームのデータベースを使う

結果 (preliminary)

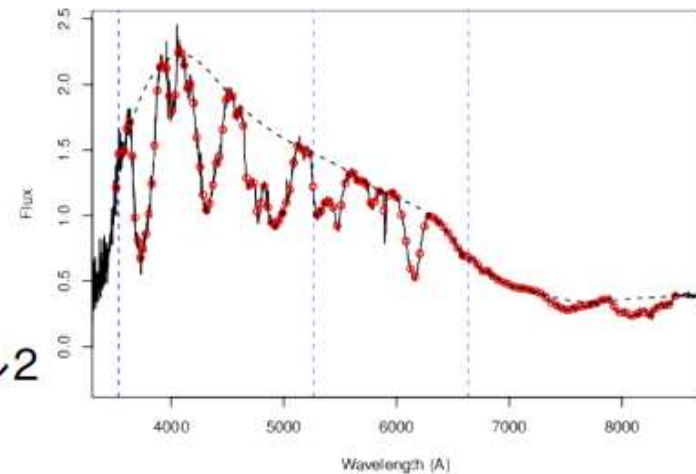
- l1最小化 + AIC(赤池情報量基準) or 交差検定で変数選択
- 今のところ、色 + 減光速度以外にモデルを改善するパラメータはなし
- データから必要な変数やその数まで推定することが可能

選択された波長
赤: 総和で規格化
青: 連続光で規格化
実線: 交差検定
点線: AIC

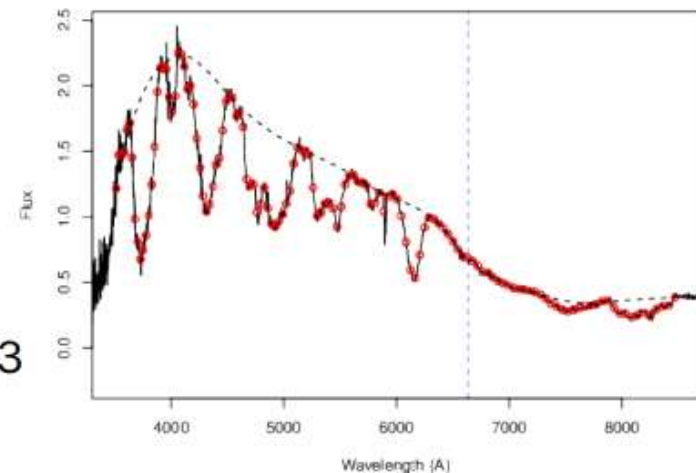
サンプル1



サンプル2



サンプル3



Preliminary

まとめ

- スパースモデリングの宇宙物理学への応用
 - 元情報の疎性を利用して、観測データからより良い情報を引き出す
- スパースな元情報を少ないデータから高精度に再構成する
 - 変光星の周期解析：エイリアス、ノイズ除去
 - 電波干渉計：超解像
 - ドップラー・トモグラフィー：少ないデータでも明確なedge
- データから客観的にモデル（パラメータ）を決定する
 - Ia型超新星の極大光度：データから客観的に変数選択
- **興味のある方、是非ご連絡下さい！**

共同研究者

(敬称略)

- 変光星の周期解析
 - 板（東北大）、松永（東大）、池田（統数研）
- 電波干渉計
 - 本間（PI: NAOJ）、秋山（NAOJ）、池田（統数研）
- ドップラー・トモグラフィー
 - 加藤（京大）、野上（京大）、Mennickent（コンセプション大）
- Ia型超新星の極大等級
 - 前田（京大）、川端（広大）、池田（統数研）

Backup slides

λ をどう決めるか

$$\hat{\mathbf{x}} = \arg \min \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

- λ がモデルのスパース度を決める
 - λ が大きい → 解はスパース、データの適合度は下がる
 - λ が小さい → 解は複雑、データの適合度は上がる
- λ をどう決めるか？
 - 赤池情報量基準(AIC) : $AIC = -2 \ln L + 2K$
 - 交差検証(cross-validation) : データを「訓練用」と「テスト用」に分け、訓練用最適化したモデルをテスト用で検証する

フーリエ変換

- 時系列空間：周波数空間 =
1 : 1 のデータ数であるべき
- 観測に欠測がある場合は実質的に「0」を入れて計算している
- サイドローブ、エイリアスの発生

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \cancel{y_a} \\ \cancel{y_b} \\ \cancel{y_c} \\ y_d \\ y_e \\ \cancel{y_f} \\ \cancel{y_g} \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} \cos(t_1\nu_1) & \cdots & \cos(t_1\nu_{M/2}) & \sin(t_1\nu_1) & \cdots & \sin(t_1\nu_{M/2}) \\ \cos(t_2\nu_1) & \cdots & \cos(t_2\nu_{M/2}) & \sin(t_2\nu_1) & \cdots & \sin(t_2\nu_{M/2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cancel{\cos(t_a\nu_1)} & \cancel{\cdots} & \cancel{\cos(t_a\nu_{M/2})} & \cancel{\sin(t_a\nu_1)} & \cancel{\cdots} & \cancel{\sin(t_a\nu_{M/2})} \\ \cancel{\cos(t_b\nu_1)} & \cancel{\cdots} & \cancel{\cos(t_b\nu_{M/2})} & \cancel{\sin(t_b\nu_1)} & \cancel{\cdots} & \cancel{\sin(t_b\nu_{M/2})} \\ \cancel{\cos(t_c\nu_1)} & \cancel{\cdots} & \cancel{\cos(t_c\nu_{M/2})} & \cancel{\sin(t_c\nu_1)} & \cancel{\cdots} & \cancel{\sin(t_c\nu_{M/2})} \\ \cos(t_d\nu_1) & \cdots & \cos(t_d\nu_{M/2}) & \sin(t_d\nu_1) & \cdots & \sin(t_d\nu_{M/2}) \\ \cos(t_e\nu_1) & \cdots & \cos(t_e\nu_{M/2}) & \sin(t_e\nu_1) & \cdots & \sin(t_e\nu_{M/2}) \\ \cancel{\cos(t_f\nu_1)} & \cancel{\cdots} & \cancel{\cos(t_f\nu_{M/2})} & \cancel{\sin(t_f\nu_1)} & \cancel{\cdots} & \cancel{\sin(t_f\nu_{M/2})} \\ \cancel{\cos(t_g\nu_1)} & \cancel{\cdots} & \cancel{\cos(t_g\nu_{M/2})} & \cancel{\sin(t_g\nu_1)} & \cancel{\cdots} & \cancel{\sin(t_g\nu_{M/2})} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(t_M\nu_1) & \cdots & \cos(t_M\nu_{M/2}) & \sin(t_M\nu_1) & \cdots & \sin(t_M\nu_{M/2}) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{M/2} \\ b_1 \\ b_2 \\ \vdots \\ b_{M/2} \end{pmatrix} \quad (5)$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ 0 \\ 0 \\ 0 \\ y_d \\ y_e \\ 0 \\ 0 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} \cos(t_1\nu_1) & \cdots & \cos(t_1\nu_{M/2}) & \sin(t_1\nu_1) & \cdots & \sin(t_1\nu_{M/2}) \\ \cos(t_2\nu_1) & \cdots & \cos(t_2\nu_{M/2}) & \sin(t_2\nu_1) & \cdots & \sin(t_2\nu_{M/2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(t_a\nu_1) & \cdots & \cos(t_a\nu_{M/2}) & \sin(t_a\nu_1) & \cdots & \sin(t_a\nu_{M/2}) \\ \cos(t_b\nu_1) & \cdots & \cos(t_b\nu_{M/2}) & \sin(t_b\nu_1) & \cdots & \sin(t_b\nu_{M/2}) \\ \cos(t_c\nu_1) & \cdots & \cos(t_c\nu_{M/2}) & \sin(t_c\nu_1) & \cdots & \sin(t_c\nu_{M/2}) \\ \cos(t_d\nu_1) & \cdots & \cos(t_d\nu_{M/2}) & \sin(t_d\nu_1) & \cdots & \sin(t_d\nu_{M/2}) \\ \cos(t_e\nu_1) & \cdots & \cos(t_e\nu_{M/2}) & \sin(t_e\nu_1) & \cdots & \sin(t_e\nu_{M/2}) \\ \cos(t_f\nu_1) & \cdots & \cos(t_f\nu_{M/2}) & \sin(t_f\nu_1) & \cdots & \sin(t_f\nu_{M/2}) \\ \cos(t_g\nu_1) & \cdots & \cos(t_g\nu_{M/2}) & \sin(t_g\nu_1) & \cdots & \sin(t_g\nu_{M/2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(t_M\nu_1) & \cdots & \cos(t_M\nu_{M/2}) & \sin(t_M\nu_1) & \cdots & \sin(t_M\nu_{M/2}) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{M/2} \\ b_1 \\ b_2 \\ \vdots \\ b_{M/2} \end{pmatrix} \quad (6)$$