

Ia型超新星の極大等級 の変数選択

植村誠（広島大学）

川端弘治（広島大学）、池田思朗（統計数理研究所）、
前田啓一（京都大学）

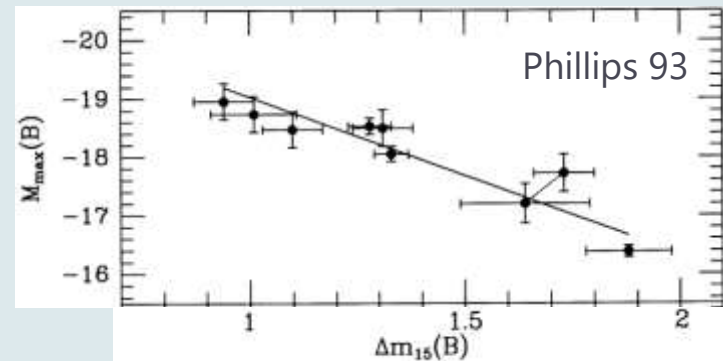
Uemura, et al., 2015, PASJ, 67, 55

連星・変光星・低温度星研究会 2016 @慶応大 日吉

Ia型の絶対等級(M_B)の説明変数は何か？

- $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

- 色 (= 星間赤化・星間減光)
- 減光率 (, or 光度曲線の幅)
- Phillips relation (Phillips 93)



絶対等級を説明す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

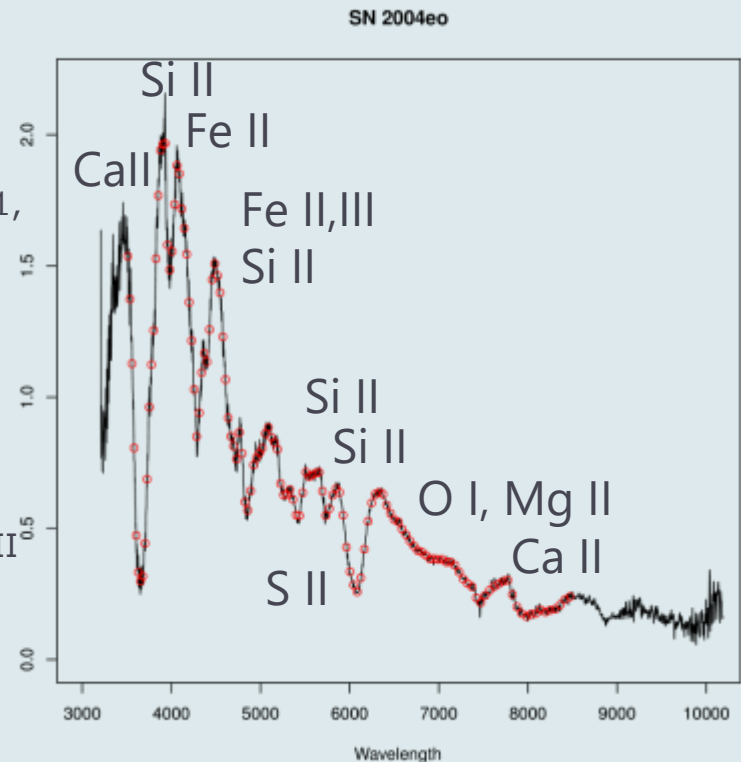
→他にも絶対等級を左右する要因があるのでは？

以下、Silverman+12 (BSNIP III) のレビューより。

$$M_B = x_1(\text{color}) + x_2(\text{decline rate}) + x_3(\text{???})$$

- 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)
- 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 (F_r/F_b) (Nugent+95, Hachinger+06)
- SiS flux ratio $F_r(\text{S II “W”})/F_r(\text{Si II 6355})$ (Bongard+06)
- SSi EW ratio $EW(\text{S II “W”})/EW(\text{Si II 5972})$, SiFe EW ratio $EW(\text{Si II 5972})/EW(\text{Fe II})$ (Hachinger+06)

- 任意のflux ratioから絶対等級と相関の高いものを探す



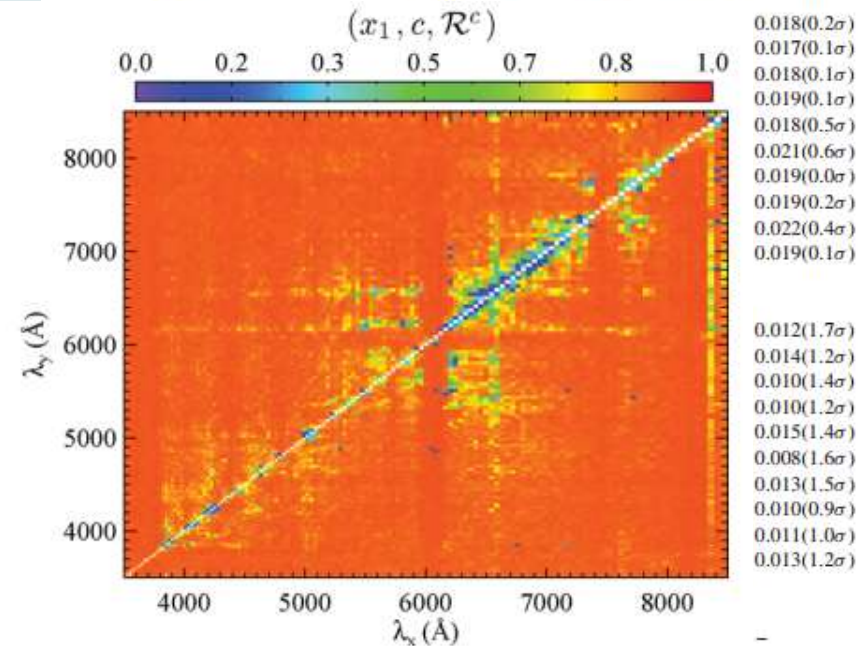
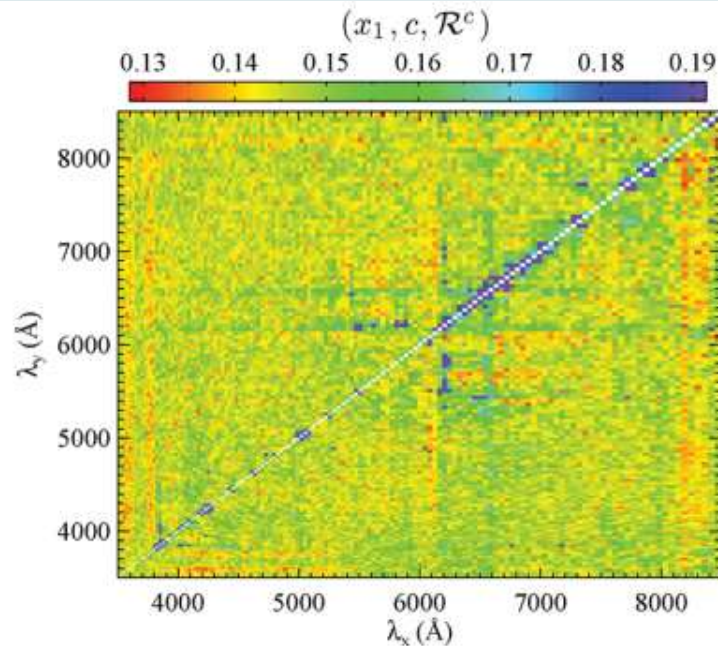
Silverman+12

- データ：Berkeley Supernova Ia Program, 108天体、Shane 3-m telescope at Lick Observatory、 $0.01 < z < 0.1$
- 手法：10-fold cross-validationで予測誤差の小さいものを選ぶ
- 結論：Color + decline rate + $\mathcal{R}^c(3750/4550)$ がベストなモデル

Table 1. Top 10 flux ratios for each model.

Rank	λ_y	λ_x	WRMS (mag)	σ_{pred} (mag)	$\rho_{x_1,c}$	$\Delta_{x_1,c}$
\mathcal{R}						
1	7770	3750	0.218 ± 0.027	0.179 ± 0.023	-0.11	$0.095 \pm 0.030(3.2\sigma)$
2	7670	3750	0.223 ± 0.028	0.189 ± 0.023	0.09	$0.104 \pm 0.030(3.5\sigma)$
3	7720	3750	0.223 ± 0.027	0.188 ± 0.023	-0.05	$0.103 \pm 0.030(3.4\sigma)$
4	7930	3750	0.226 ± 0.030	0.187 ± 0.023	-0.10	$0.104 \pm 0.031(3.4\sigma)$
5	6990	3750	0.227 ± 0.030	0.193 ± 0.023	-0.17	$0.111 \pm 0.031(3.6\sigma)$
6	7880	3750	0.227 ± 0.027	0.190 ± 0.023	-0.06	$0.106 \pm 0.030(3.5\sigma)$
7	7670	3780	0.230 ± 0.030	0.199 ± 0.024	0.08	$0.114 \pm 0.030(3.8\sigma)$
8	6990	3780	0.231 ± 0.029	0.202 ± 0.023	-0.21	$0.121 \pm 0.031(3.9\sigma)$
9	6900	3750	0.233 ± 0.029	0.198 ± 0.024	-0.18	$0.116 \pm 0.031(3.7\sigma)$
10	7040	3780	0.234 ± 0.029	0.208 ± 0.023	-0.15	$0.127 \pm 0.031(4.1\sigma)$
(x_1, \mathcal{R})						
1	6990	3750	0.199 ± 0.022	0.160 ± 0.022	-0.19	$0.080 \pm 0.031(2.6\sigma)$
2	7770	3750	0.200 ± 0.024	0.160 ± 0.022	-0.09	$0.077 \pm 0.030(2.6\sigma)$
3	6720	3750	0.203 ± 0.024	0.166 ± 0.022	-0.28	$0.088 \pm 0.031(2.8\sigma)$
4	6900	3750	0.204 ± 0.024	0.164 ± 0.022	-0.18	$0.083 \pm 0.031(2.7\sigma)$
5	6760	3750	0.206 ± 0.027	0.165 ± 0.022	-0.22	$0.085 \pm 0.031(2.8\sigma)$
6	6950	3750	0.206 ± 0.024	0.166 ± 0.023	-0.19	$0.085 \pm 0.031(2.7\sigma)$
7	6850	3750	0.206 ± 0.024	0.166 ± 0.022	-0.22	$0.085 \pm 0.031(2.8\sigma)$
8	7930	3750	0.209 ± 0.027	0.167 ± 0.022	-0.12	$0.085 \pm 0.030(2.8\sigma)$
9	6590	3750	0.209 ± 0.028	0.166 ± 0.023	-0.23	$0.092 \pm 0.032(2.9\sigma)$
10	6990	3780	0.209 ± 0.027	0.177 ± 0.022	-0.19	$0.097 \pm 0.030(3.2\sigma)$

(c, \mathcal{R}^c)



要はこういう問題

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

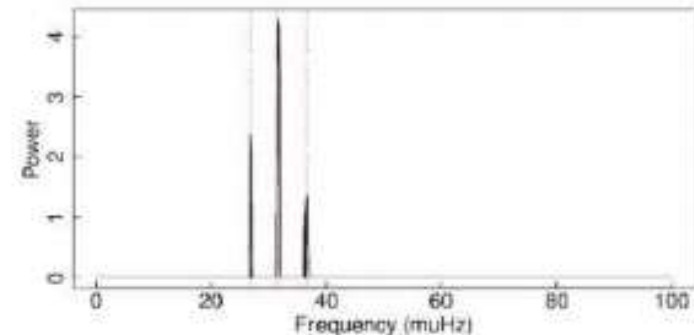
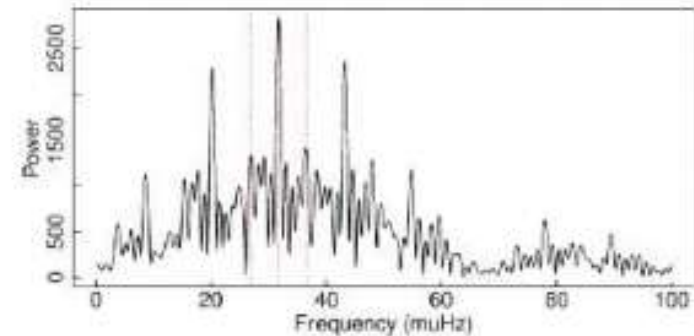
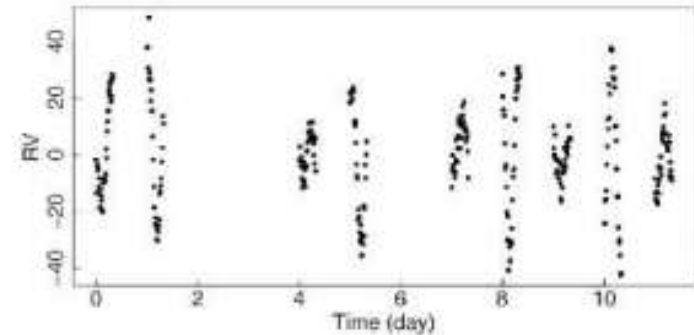
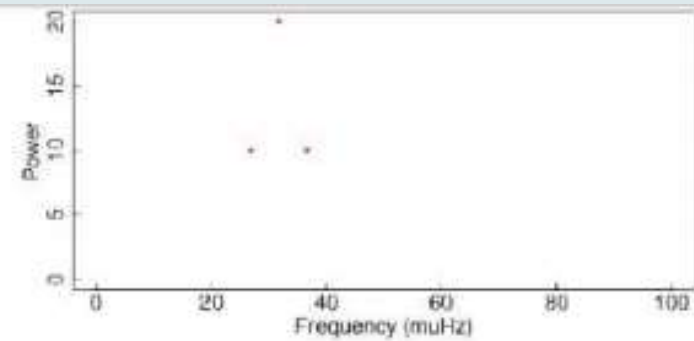
- サンプル数は～100。説明変数は任意のフラックス比を考えると> 1万
- 線形モデルで解きたいけど、サンプル数がパラメータの数より小さいので、普通には解けない
- でも、本来必要なパラメータは数個のはず。変数をデータから選択したい。

スパースモデリング で解こう！

- 1次ノルム最小化

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

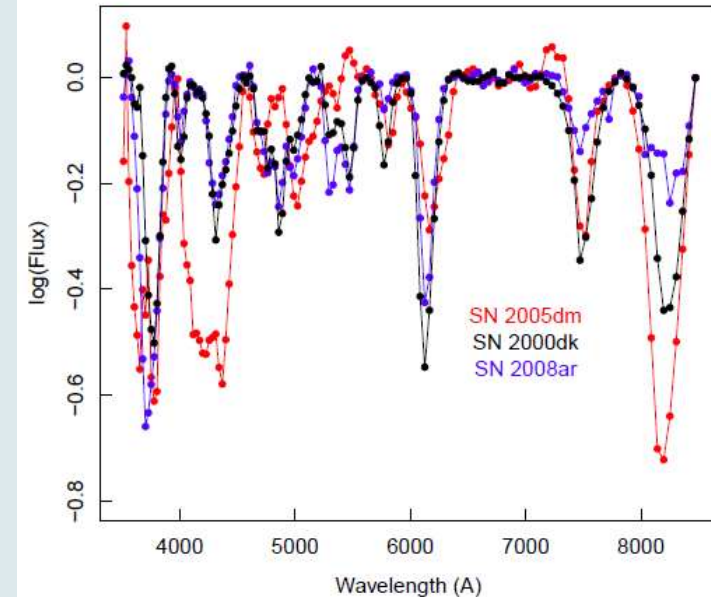
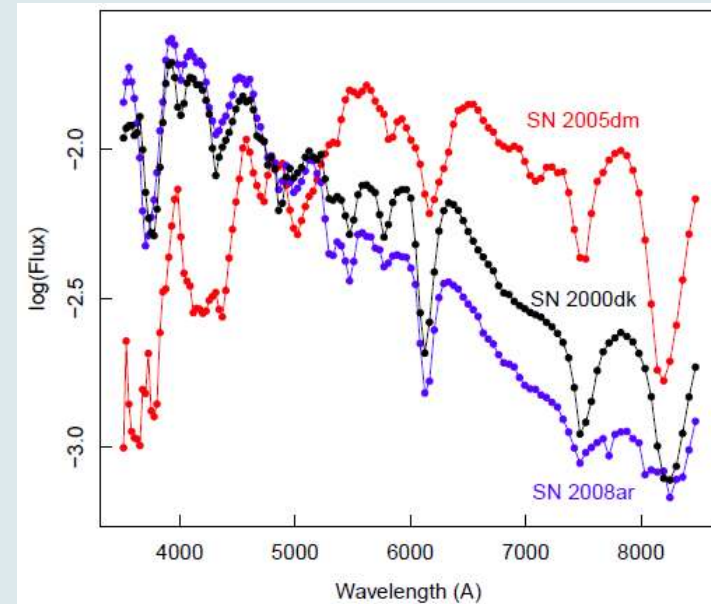
- データは Silverman+12と同じデータ



Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

データと変数

- Berkeley Supernova Database
 - Silverman+12
- 天体数 = 76
- MBの説明変数 (合計276個)
 - 色(c)
 - 光度曲線の幅(x1)
 - 3500 – 8500 Å の総フラックスで規格化したスペクトル
 - 134点/スペクトル
 - 連続光レベルで規格化したスペクトル
 - 134点/スペクトル
 - 先行研究で提案されてきたフラックス比 (6点)
- 任意のフラックス比 ($> 10^4$) と比べて変数の数が減少

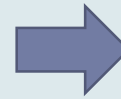


Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

Step 1/3 : 全ての変数を使ったモデル

$$\hat{\beta}_\lambda = \arg \min_{\beta} \{ \|\mathbf{y} - X\beta\|_2^2 + \lambda \|\beta\|_1 \}$$

$$M_B = M_{B,0} + \beta_1 c + \beta_2 x_1 + \beta_3 f_{\text{tot}}(3512) + \beta_4 f_{\text{tot}}(3534) + \dots + \beta_{136} f_{\text{tot}}(8472) + \beta_{137} f_{\text{cnt}}(3512) + \beta_{138} f_{\text{cnt}}(3534) + \dots + \beta_{270} f_{\text{cnt}}(8472) + \beta_{271} \mathcal{R}(3780/4580) + \beta_{272} \mathcal{R}(4610/4260) + \beta_{273} \mathcal{R}(5690/5360) + \beta_{274} \mathcal{R}(6420/4430) + \beta_{275} \mathcal{R}(6420/5290) + \beta_{276} \mathcal{R}(6630/4400) + e. \quad (6)$$



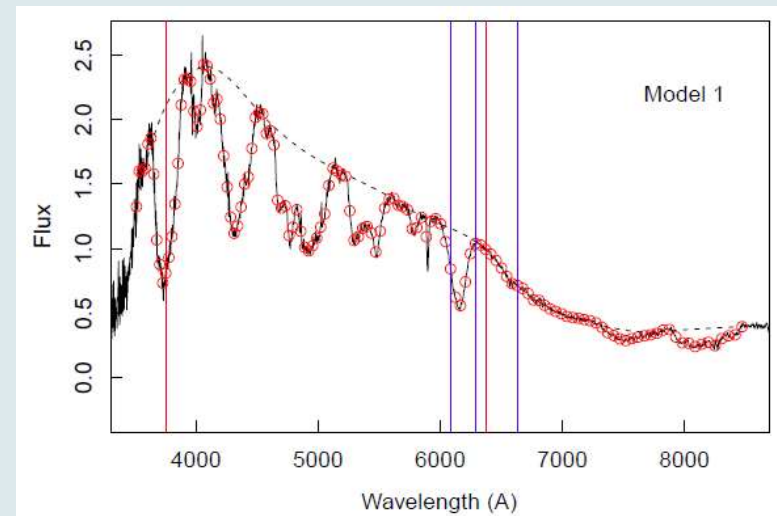
Non-zero elements	coefficients β	p
c	0.376	1.00
$f_{\text{tot}}(6373)$	0.100	1.00
x_1	-0.050	0.98
$f_{\text{cnt}}(6084)$	-0.034	0.98
$f_{\text{cnt}}(6289)$	-0.045	0.95
$f_{\text{cnt}}(6631)$	-0.061	0.80
$\mathcal{R}(3780/4580)$	-0.050	0.74
$f_{\text{tot}}(3752)$	0.063	0.73

$f_{\text{tot}}(6373)$ (= local な色の情報?)

- 色(c)とは違う情報を持っているか?
- ➔ 色(c)であらかじめ M_B を補正して再解析

$f_{\text{cnt}}(6084), (6289)$ (=Si II 6355?)

- 光度曲線の幅(x_1)と相関があるか? (e.g. Hachinger+06)
- ➔ x_1 を目的変数にして再解析

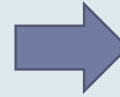


Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

Step 2/3 : 色(c)補正MBを使ったモデル

$$\hat{\beta}_\lambda = \arg \min_{\beta} \{ \|\mathbf{y} - X\beta\|_2^2 + \lambda \|\beta\|_1 \}$$

$$\begin{aligned} M_B = & M_{B,0} + \beta_1 c + \beta_2 x_1 \\ & + \beta_3 f_{\text{tot}}(3512) + \beta_4 f_{\text{tot}}(3534) + \dots + \beta_{136} f_{\text{tot}}(8472) \\ & + \beta_{137} f_{\text{cnt}}(3512) + \beta_{138} f_{\text{cnt}}(3534) + \dots + \beta_{270} f_{\text{cnt}}(8472) \\ & + \beta_{271} \mathcal{R}(3780/4580) + \beta_{272} \mathcal{R}(4610/4260) \\ & + \beta_{273} \mathcal{R}(5690/5360) + \beta_{274} \mathcal{R}(6420/4430) \\ & + \beta_{275} \mathcal{R}(6420/5290) + \beta_{276} \mathcal{R}(6630/4400) + e. \end{aligned} \quad (6)$$



Non-zero elements	coefficients	p
	β	
x_1	-0.020	0.99

• 色補正したMB

- $M_B = M_{B,0} + \beta c$ で $(M_{B,0}, \beta)$ を決定、 $M_B - \beta c$ を目的変数とする
- x_1 のみ選択。ftot(6373)は選ばれない → "c"だけで十分。

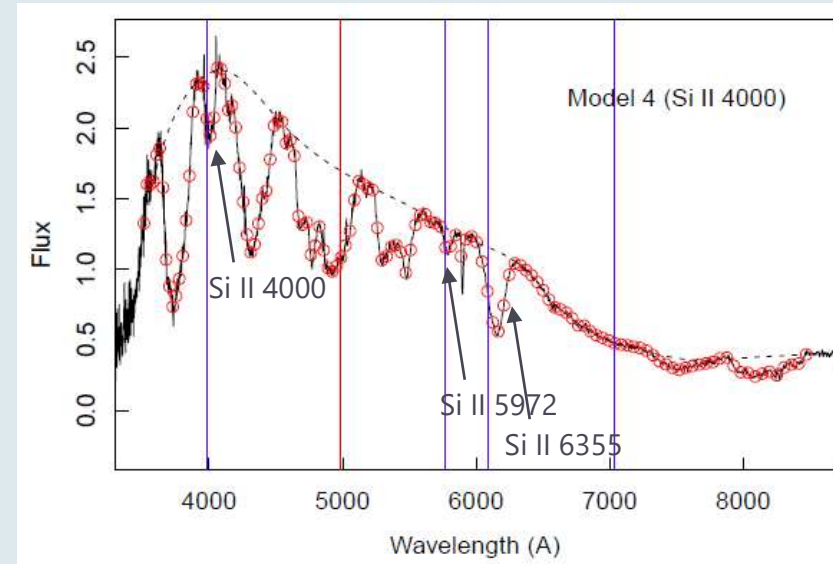
Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

Step 3/3 : (c,x1)補正MBを使ったモデル

- 光度曲線の幅(x1)の説明変数は？

$$x_1 = x_{1,0} + \beta_1 c + \beta_2 f_{tot}(3512) + \dots + e$$

- Si II 4000, 5972, 6355とx1の相関を確認
→ Step 1/3 で Si II 6355付近のフラックスが
選択されたのはx1との相関のため？



- (c,x1)でMBを補正

$$\begin{aligned}
 M_B = & M_{B,0} + \beta_1 c + \beta_2 x_1 \\
 & + \beta_3 f_{tot}(3512) + \beta_4 f_{tot}(3534) + \dots + \beta_{136} f_{tot}(8472) \\
 & + \beta_{137} f_{cnt}(3512) + \beta_{138} f_{cnt}(3534) + \dots + \beta_{270} f_{cnt}(8472) \\
 & + \beta_{271} \mathcal{R}(3780/4580) + \beta_{272} \mathcal{R}(4610/4260) \\
 & + \beta_{273} \mathcal{R}(5690/5360) + \beta_{274} \mathcal{R}(6420/4430) \\
 & + \beta_{275} \mathcal{R}(6420/5290) + \beta_{276} \mathcal{R}(6630/4400) + e. \quad (6)
 \end{aligned}$$

何も選択されない
→ (c,x1)が最適な変数の
組み合わせ

Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

Discussion & Future plan

- スペクトルデータは良い変数に成り得ないか？
 - Broad-band color "c" が選択されたのは測定誤差 (= 測光誤差) が小さいため
 - Flux-calibrationの精度が非常に高ければ(e.g. Bailey+09)スペクトルデータが選択される可能性はある
 - 今回の結果はあくまで「Berkeleyデータだとスペクトルデータはモデルを改善しない」という結論
- 今後は、母銀河の情報等を加える、別のデータサンプルで試す、等々。

まとめ

$$M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$$

- Berkeleyのデータを使い、LASSOで変数選択
- ???に該当する変数はない
- 少ないデータに対して説明変数の候補が多すぎるとノイズを拾ってしまう確率が高くなるので、なるべく変数の数を減らすべき。
- 変数選択についてはLASSO等、統計学・機械学習の分野から良い手法を応用すべき。

Back-up slides

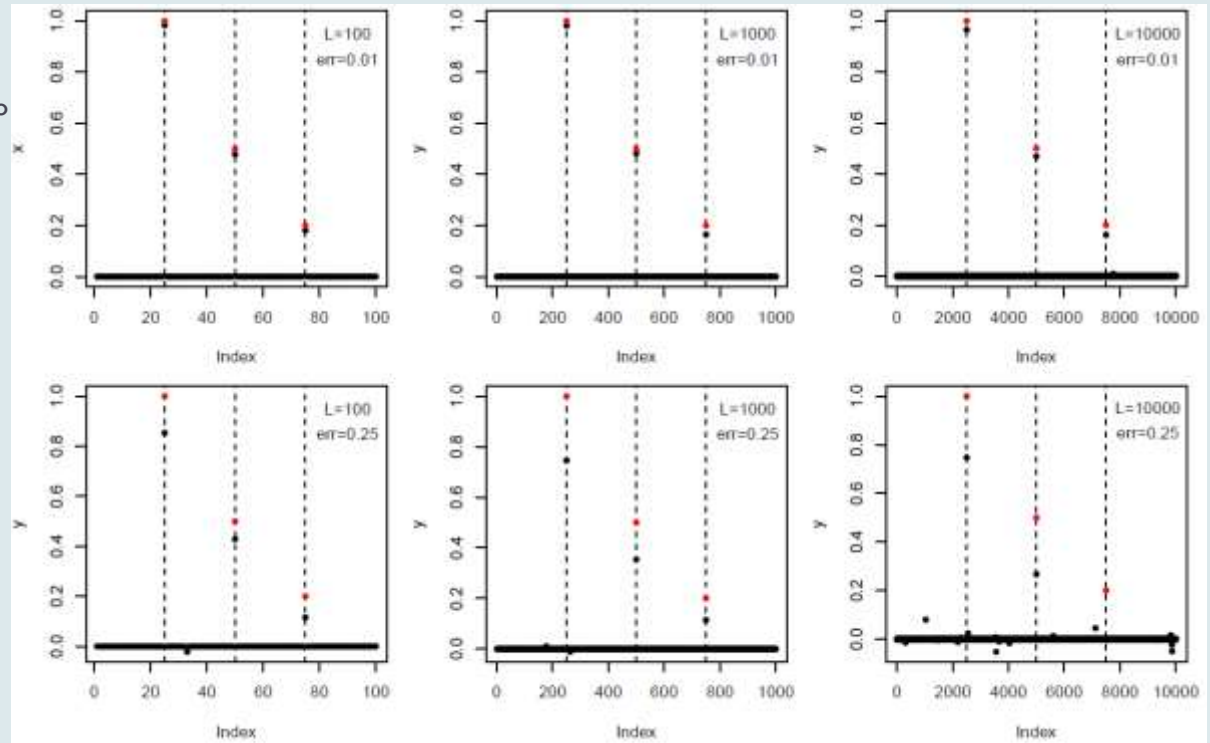
Ia型超新星の極大等級 $M_B = M_{B,0} + \beta_1(\text{色}) + \beta_2(\text{減光率}) + ???$

LASSOによる変数選択

$$\hat{\beta}_\lambda = \arg \min_{\beta} \{ \|\mathbf{y} - X\beta\|_2^2 + \lambda \|\beta\|_1 \} \quad \|\beta\|_1 = \sum_i |\beta_i|$$

例：50個の y から100, 1000, 10000個の β をLASSOで推定。
赤：仮定した3つの信号。
黒：推定値。

ノイズ小



ノイズ大

→ 変数の候補数が多すぎるとノイズを拾う

LASSOのλをどう決めるか？

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

- Cross-validation

