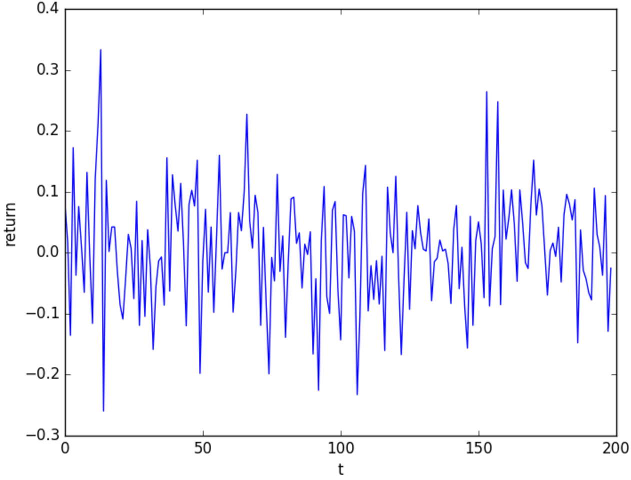
ポートフォリオと高次元データ分析

深澤　風士

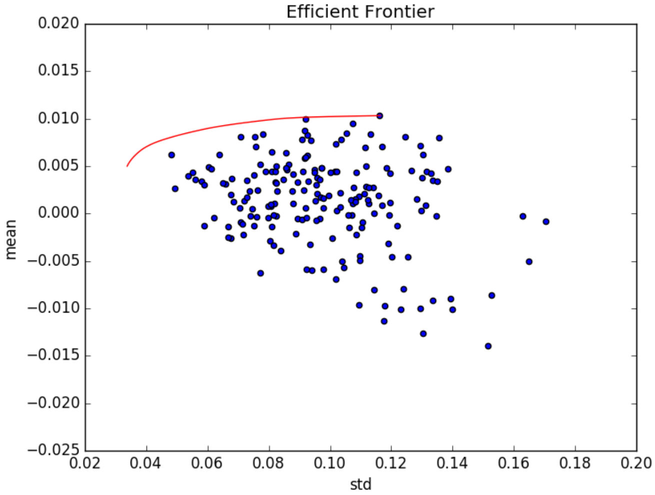
総合政策学部 4 年

2017/1/19

各ストックのデータ



効率的フロンティア


Efficient Frontier
mean
std
0.020
0.015
0.010
0.005
0.000
−0.005
−0.010
−0.015
−0.020
−0.025
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20

マーコヴィッツモデル

Markowitz (1952) によって提唱された.

$$\min \quad \mathbf{x}^T \Sigma \mathbf{x} \tag{1}$$

$$\text{subject to} \begin{cases} \mathbf{x} \geq 0, \\ \mathbf{x}^T \mathbf{1} = 1, \\ \mathbf{x}^T \mathbf{r} = r_0. \end{cases} \tag{2}$$

ここで, Σ は $p \times p$ の分散共分散行列であり, $\mathbf{x}$ は $p \times 1$ のポートフォリオの重みベクトル, $\mathbf{1}$ は 1 を p 個並べたベクトル, $\mathbf{r}$ は $p \times 1$ の各ストックの期待リターンを並べたベクトル, r_0 はポートフォリオの目標リターン（スカラー値）である.

Lasso

Tibshirani (1996) によって提案された.

$$\min \quad \frac{1}{2}\sum_{i=1}^{n}\Bigl(y_i - w_0 - \sum_{j=1}^{p} w_j x_{ij}\Bigr)^2, \tag{3}$$

$$\text{subject to} \quad \sum_{j}^{p} |w_j| \leq \alpha. \tag{4}$$

ここで, $\alpha > 0$ は罰則の度合いを調節するパラメーターである.

Lasso の最適化 1：Coordinate Descent Algorithm

以下の L_1 罰則項付きの目的関数を考える．ここで，
$\mathbf{x}_i \in \mathbb{R}^{p+1} = (1, x_1, \ldots, x_p)^T$, $\mathbf{w} \in \mathbb{R}^{p+1} = (w_0, w_1, \ldots, w_p)^T$ の
$p+1$ 次元ベクトルとする．

$$L(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}(y_i - \mathbf{x}_i^T\mathbf{w})^2 + \lambda\sum_{j=1}^{p}|\mathbf{w}_j| \tag{5}$$

$\lambda > 0$ は正則化の度合いに対するパラメーターである．ここで，

$$\mathrm{Loss}(\mathbf{w}) := \frac{1}{2}\sum_{i=1}^{n}(y_i - \mathbf{x}_i^T\mathbf{w})^2, \tag{6}$$

$$\mathrm{L}_1(\mathbf{w}) := \lambda\sum_{j=1}^{p}|\mathbf{w}_j|. \tag{7}$$

とする．

Lasso の最適化 2

$\mathbf{w}$ 内のあるインデックス d の要素 $\mathbf{w}_d \in \mathbb{R}$ を明示的に書くと以下のように表記できる.

$$L(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}(y_i - \mathbf{x}_{id}\mathbf{w}_d - \mathbf{x}_{\backslash d}^T\mathbf{w}_{\backslash d})^2 + \lambda\Big(|\mathbf{w}_d| + \sum_{j\neq d}|\mathbf{w}_j|\Big) \quad (8)$$

$\mathrm{L}_1(\mathbf{w})$ は微分不可能な点を含むことから, まず $\mathrm{Loss}(\mathbf{w})$ のみを微分することを考える. $\mathrm{Loss}(\mathbf{w})$ を $\mathbf{w}_d$ について微分したものを 0 とおくと,

$$\begin{aligned}
\frac{\partial \mathrm{Loss}(\mathbf{w})}{\partial \mathbf{w}_d} &= \sum_{i=1}^{n}(y_i - \mathbf{x}_{id}\mathbf{w}_d - \mathbf{x}_{\backslash d}^T\mathbf{w}_{\backslash d})(-\mathbf{x}_{id}) \\
&= -\sum_{i=1}^{n} y_i\mathbf{x}_{id} + \sum_{i=1}^{n}\mathbf{x}_{id}^2\mathbf{w}_d + \sum_{i=1}^{n}\mathbf{x}_{\backslash d}^T\mathbf{w}_{\backslash d}\mathbf{x}_{id} = 0
\end{aligned}$$

Lasso の最適化 3

$$\sum_{i=1}^{n} \mathbf{x}_{id}^2 \mathbf{w}_d = \sum_{i=1}^{n} y_i \mathbf{x}_{id} - \sum_{i=1}^{n} \mathbf{x}_{\backslash d}^T \mathbf{w}_{\backslash d} \mathbf{x}_{id}$$

$$\hat{\mathbf{w}}_d = \frac{\sum_{i=1}^{n} \mathbf{x}_{id}(y_i - \mathbf{x}_{\backslash d}^T \mathbf{w}_{\backslash d})}{\sum_{i=1}^{n} \mathbf{x}_{id}^2}. \tag{9}$$

これを用いて, $L(\mathbf{w})$ を書き換える．また, その際に $\mathbf{w}_d$ と関係のない項を const としてまとめておくと,

Lasso の最適化 4

$$L(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}(y_i - \mathbf{x}_{id}\mathbf{w}_d - \mathbf{x}_{\setminus d}^T\mathbf{w}_{\setminus d})^2 + \lambda|\mathbf{w}_d| + \lambda\sum_{j\neq d}|\mathbf{w}_j|$$

$$L(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}\left\{\mathbf{x}_{id}^2\mathbf{w}_d^2 - 2\mathbf{x}_{id}\mathbf{w}_d(y_i - \mathbf{x}_{\setminus d}^T\mathbf{w}_{\setminus d})\right\} + \lambda|\mathbf{w}_d| + \mathrm{const}$$

$$= \frac{1}{2}\mathbf{w}_d^2\sum_{i=1}^{n}\mathbf{x}_{id}^2 - \mathbf{w}_d\Big(\sum_{i=1}^{n}\mathbf{x}_{id}^2\Big)\underbrace{\frac{\sum_{i=1}^{n}\mathbf{x}_{id}(y_i - \mathbf{x}_{\setminus d}^T\mathbf{w}_{\setminus d})}{\sum_{i=1}^{n}\mathbf{x}_{id}^2}}_{\hat{\mathbf{w}}_d} + \lambda|\mathbf{w}_d| + \mathrm{const}$$

Lasso の最適化 5

$$= \frac{1}{2}\mathbf{w}_d^2 \sum_{i=1}^{n} \mathbf{x}_{id}^2 - \mathbf{w}_d(\sum_{i=1}^{n} \mathbf{x}_{id}^2)\hat{\mathbf{w}}_d + \lambda|\mathbf{w}_d| + \mathrm{const}$$

$$= \frac{1}{2}\mathbf{w}_d^2 n \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_{id}^2 - \mathbf{w}_d n(\frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_{id}^2)\hat{\mathbf{w}}_d + \lambda|\mathbf{w}_d| + \mathrm{const},$$

ここで, 各 $\mathbf{x}_d$ が正規化されていると仮定すると, $\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{id}^2 = 1$ となり,

$$L(\mathbf{w}) = \frac{n}{2}\mathbf{w}_d^2 - n\mathbf{w}_d\hat{\mathbf{w}}_d + \lambda|\mathbf{w}_d| + \mathrm{const}. \tag{10}$$

Lasso の最適化 6

これを $\mathbf{w}_d$ について微分して 0 とおくと,

$$\frac{\partial L(\mathbf{w})}{\partial \mathbf{w}_d} = \begin{cases} \mathbf{w}_d - \hat{\mathbf{w}}_d + \gamma = 0 & (\text{if} \quad \mathbf{w}_d > 0) \\ \mathbf{w}_d - \hat{\mathbf{w}}_d - \gamma = 0 & (\text{if} \quad \mathbf{w}_d < 0) \\ \text{undefine} & (\text{if} \quad \mathbf{w}_d = 0) \end{cases} \tag{11}$$

ここで, $\gamma := \frac{\lambda}{2n}$ とおいた.

Lasso の最適化 7

$\mathbf{w}_d > 0$ のとき, $\frac{\partial L(\mathbf{w})}{\partial \mathbf{w}_d} = 0$ となる $\mathbf{w}_d$ は,

$$\mathbf{w}_d = \hat{\mathbf{w}}_d - \gamma \tag{12}$$

であり, $\mathbf{w}_d < 0$ のときは,

$$\mathbf{w}_d = \hat{\mathbf{w}}_d + \gamma \tag{13}$$

となる. しかし, $-\gamma < \hat{\mathbf{w}}_d < \gamma$ のとき, $\mathbf{w}_d > 0$ ならば, $\hat{\mathbf{w}}_d - \gamma > 0$, すなわち $\hat{\mathbf{w}}_d > \gamma$ となって矛盾する. また, $-\gamma < \hat{\mathbf{w}}_d < \gamma$ のとき, $\mathbf{w}_d < 0$ ならば, $\hat{\mathbf{w}}_d + \gamma < 0$ すなわち $\hat{\mathbf{w}}_d < -\gamma$ となって矛盾する. したがって $-\gamma < \hat{\mathbf{w}}_d < \gamma$ のとき, $\mathbf{w}_d = 0$ である.

Lasso の最適化 7

これをまとめると以下のようになる.

$$\hat{\mathbf{w}}_d^{lasso} = S(\hat{\mathbf{w}}_d, \gamma) \equiv \mathrm{sign}(\hat{\mathbf{w}}_d)(|\hat{\mathbf{w}}_d| - \gamma)_+ \tag{14}$$

$$= \begin{cases} \hat{\mathbf{w}}_d - \gamma, & \text{if} \quad \hat{\mathbf{w}}_d > 0 \quad \text{and} \quad \gamma < |\hat{\mathbf{w}}_d|, \\ \hat{\mathbf{w}}_d + \gamma, & \text{if} \quad \hat{\mathbf{w}}_d < 0 \quad \text{and} \quad \gamma < |\hat{\mathbf{w}}_d|, \\ 0, & \text{if} \quad \gamma \geq |\hat{\mathbf{w}}_d|. \end{cases} \tag{15}$$

ここで, 関数 $S(\cdot)$ は soft-threshold 関数と呼ばれる. 各 $\mathbf{x}_d$ が正規化されていれば, 定数となる部分は無視して, 更新のアルゴリズムを以下のように定義できる.

$$\hat{\mathbf{w}}_d^{lasso} \leftarrow S\left(\sum_{i=1}^{n} \mathbf{x}_{id}(y_i - \mathbf{x}_{\backslash d}^T \mathbf{w}_{\backslash d}), \gamma\right) \quad \text{for} \quad d = 1, 2, \ldots, p, 1, 2, \ldots \tag{16}$$

これを, 収束するまで繰り替えすことによって得られる.

シミュレーション：コードについて

python： anaconda3-4.0.0 上で書いた.
https://github.com/kazetof/portfolio

シミュレーション：データについて

- Nikkei 225 に含まれるストックのうち，2000 年 2 月よりデータのある 185 種類.
- 月次データ.
- 対数階差をとってる.

シミュレーション：比較手法 1

1. 経験分散共分散行列

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \tag{17}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})^T. \tag{18}$$

この標本分散は不偏推定量となる.

シミュレーション：比較手法2

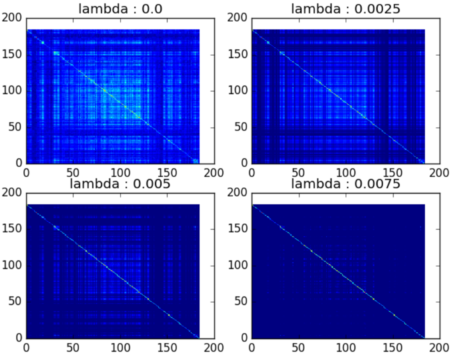
2. 経験分散共分散行列（対角行列を仮定）

$$S = \mathrm{diag}(\mathrm{VAR}[r_1], \ldots, VAR[r_p]) \tag{19}$$

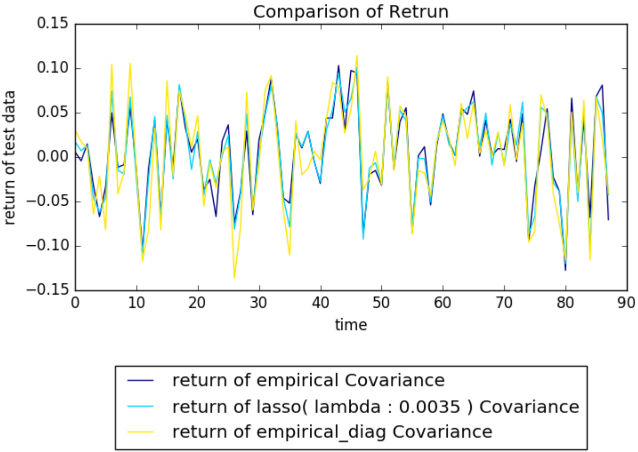
これは，共分散の分だけバイアスを持った推定量である.

シミュレーション：比較手法3

3.Lasso：λ を 0.0035 とした.



結果1：テストリターンのプロット

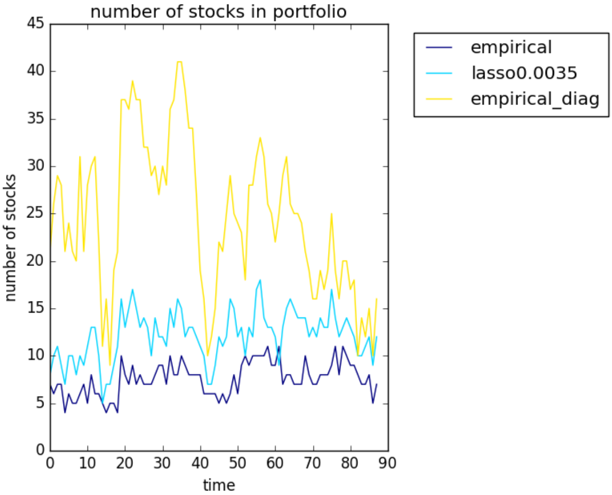


結果2：テストリターンと分散

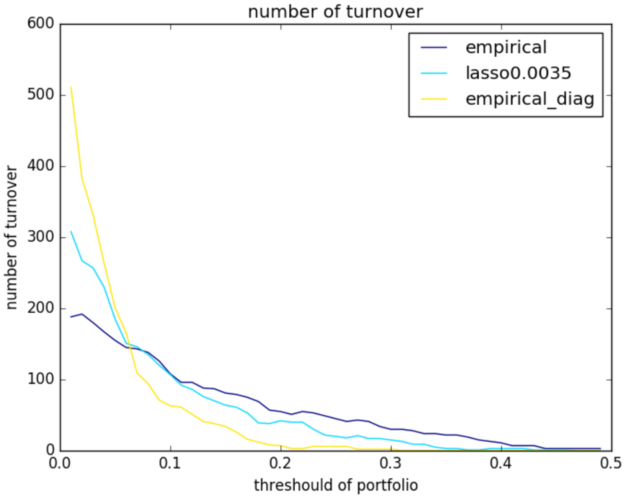
Table: 結果

	経験共分散行列	Lasso	対角要素のみ
期待テストリターン	0.0046	0.0041	0.0017
テストリターンの分散	0.0502	0.0511	0.0584

結果３：ポートフォリオに含まれるストックの数の比較



結果4：ターンオーバーの比較



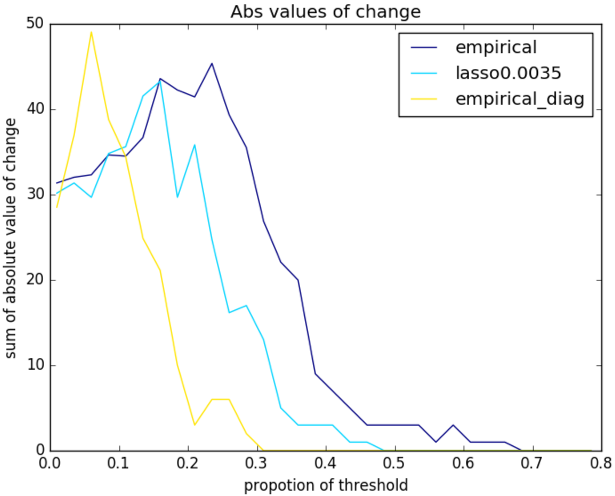
結果4：重みベクトルの変動の比較

重みベクトルの変動の絶対値和により比較する．つまり，

$$D = \sum_{i=2}^{n} \sum_{j=1}^{p} |w_j^i - w_j^{i-1}| \tag{20}$$

に D よって変動を比較する．ここで，w_j^i は i 期におけるポートフォリオの重みベクトルの j 番目の要素を表す．n はテストデータの得られた期間数であり，p はストックの種別数である．

結果 4：重みベクトルの変動の比較



シミュレーション2：比較手法1

等比率ポートフォリオの重みは,

$$w = \frac{1}{p}\mathbf{1}_p \tag{21}$$

として計算できる．これは, 期待リターンやリスクに依存しておらず, 無情報のポートフォリオとして解釈できる.

シミュレーション 2：比較手法 2

William F. Sharpe (1963) によって提案された, 各ストックはマーケット全体の動きとのみ相関があるという仮定を置くモデルである. Single Index Model は以下のように定義される.

$$R_i = \alpha_i + \beta_i R_m + \epsilon_i \quad (i = 1, \dots, p) \tag{22}$$

$$R_m = \alpha_{p+1} + \epsilon_{p+1} \tag{23}$$

ここで, R_i は個別のストックのリターンであり, R_m はマーケット全体のポートフォリオのリターンである. α_i, β_i はパラメータである. $\epsilon_i \quad (i = 1, \dots, p+1)$ は $E[\epsilon_i] = 0, \mathrm{Var}[\epsilon_i] = Q_i$ とする. また, $\epsilon_i, \epsilon_j \quad (i \neq j, \quad i, j = 1, \dots, p)$ について, $\mathrm{Cov}[\epsilon_i, \epsilon_j] = 0$ が成り立つと仮定する.

シミュレーション2：比較手法2

各ストックのパラメータ α_i, β_i を求めると, 各ストックの期待リターン E_i, 分散 V_i, 共分散 C_{ij} は以下のように推定される.

$$E_i = \alpha_i + \beta_i \alpha_{p+1} \tag{24}$$

$$V_i = \beta_i^2 Q_{p+1} + Q_i \tag{25}$$

$$C_{ij} = \beta_i \beta_j Q_{p+1} \tag{26}$$

つまり, $(p \times p)$ 行列である共分散行列 S は, β_i を並べた $(p \times 1)$ ベクトルを β, 各ストックの誤差項の分散 Q_i を並べた $(p \times 1)$ ベクトルを Q として以下のように得られる.

$$S = \beta \beta^T Q_{p+1} + \mathrm{diag}(Q). \tag{27}$$

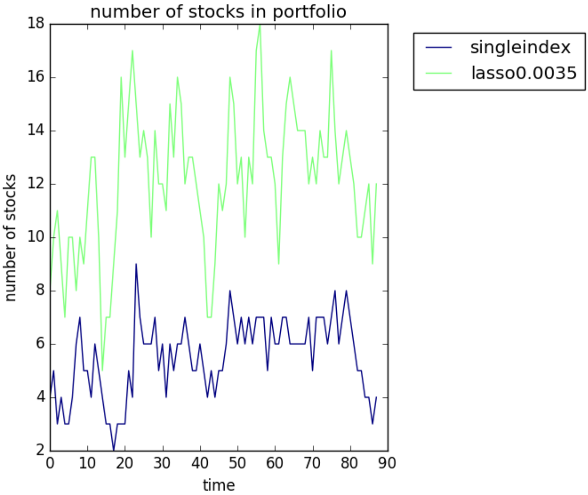
結果1

Table: Lasso とその他モデルの比較

	等比率	Single Index	Lasso
期待テストリターン	0.0051	0.0067	0.0041
テストリターンの分散	0.0572	0.0588	0.0511

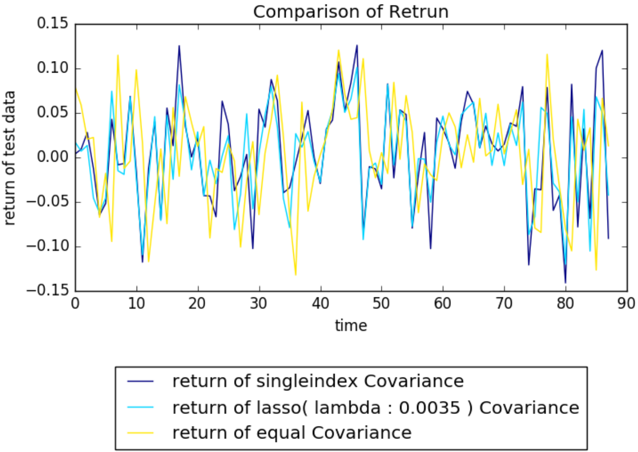
ここで，等比率ポートフォリオとは等比率で 185 品目全て保持するポートフォリオのことである．

結果 1

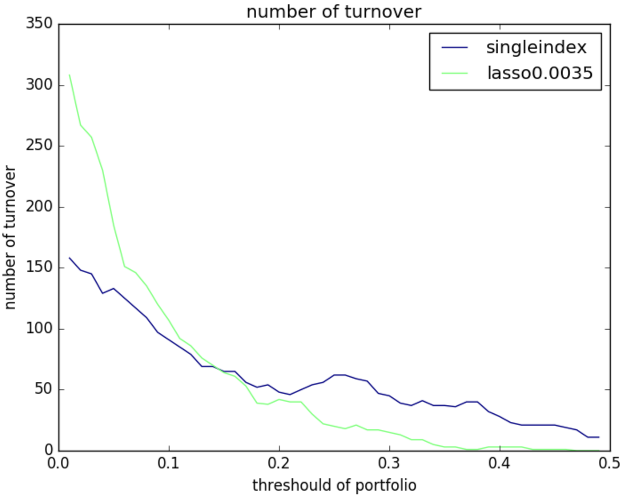


number of stocks in portfolio
number of stocks
time
singleindex
lasso0.0035
18
16
14
12
10
8
6
4
2
0
10
20
30
40
50
60
70
80
90

結果 1

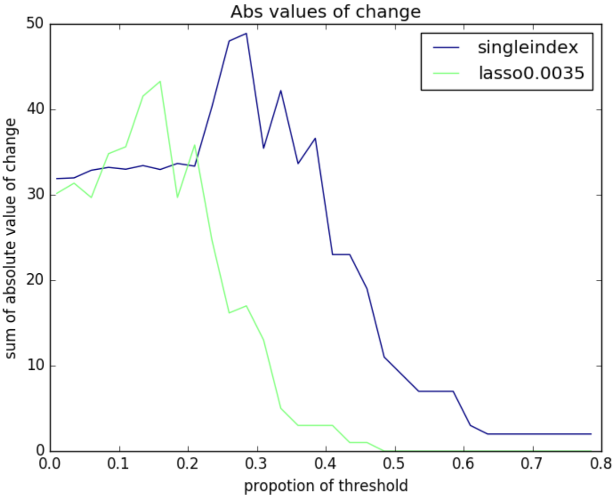

Comparison of Retrun
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
return of test data
0
10
20
30
40
50
60
70
80
90
time
return of singleindex Covariance
return of lasso(lambda : 0.0035) Covariance
return of equal Covariance

結果 1



number of turnover
singleindex
lasso0.0035
number of turnover
threshould of portfolio
350
300
250
200
150
100
50
0
0.0
0.1
0.2
0.3
0.4
0.5

結果 1


Abs values of change
singleindex
lasso0.0035
sum of absolute value of change
propotion of threshold
0
10
20
30
40
50
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8