

Lasso estimatoren



The Least Absolute Shrinkage Selection Operator (lasso) løser optimeringsproblemet

$$\hat{\beta}^{\text{lasso}} = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \sum_{i=1}^n \left(y_i - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \right\}, \text{ u.h.t. at } \sum_{j=1}^p |\beta_j| \leq t,$$

som kan omskrives til et lagrange problem

$$\hat{\beta}^{\text{lasso}} = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1 \right\}.$$

Ridge regression estimatoren findes ud fra

$$\begin{aligned} \hat{\beta}^{\text{ridge}} &= \arg \min_{\beta \in \mathbb{R}^p} \left\{ \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_2^2 \right\} \\ &= \left(\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_p \right)^{-1} \mathbf{X}^T \mathbf{y}. \end{aligned}$$

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable
MAOK9 5.2018

1 Lasso og dens generaliseringer

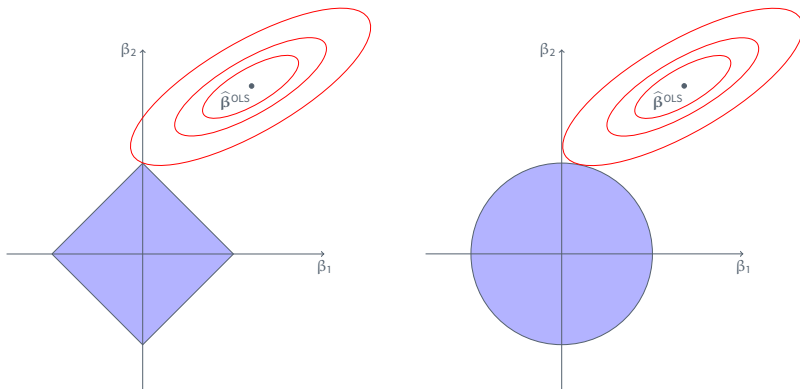
Data

Benchmark modellen

Lasso modellen og dens generaliseringer

Out-of-sample

Lasso estimatoren



Figur: Estimations illustration for lasso (venstre) og ridge regression (højre). De blå arealer er betingelsesområderne $|\beta_1| + |\beta_2| \leq t$ og $\beta_1^2 + \beta_2^2 \leq t^2$, mens de røde ellipser er konturkurver for SSR. Konturkurverne har centrum i OLS estimatoren, $\hat{\beta}^{\text{OLS}}$.

Generaliseringer af lasso estimatoren

Naiv elastisk net



Selvom lasso har vist succes i mange tilfælde, har den også nogle begrænsninger:

- ▶ Hvis $p > n$, da udvælger lasso højst n variable
- ▶ Hvis der eksisterer en gruppe af variable med høj parvis korrelation, da vil lasso blot udvælge én variabel fra denne gruppe og denne variabel udvælges tilfældigt

Naiv elastisk net løser optimeringsproblemet

$$\hat{\beta}^{\text{naivEN}} = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \left[\frac{1}{2}(1 - \alpha)\|\beta\|_2^2 + \alpha\|\beta\|_1 \right] \right\}.$$

- ▶ Hvis $\alpha = 0$, da reduceres det til den kvadrerede ℓ_2 -norm svarende til strafleddet for ridge regression
- ▶ Hvis $\alpha = 1$ reduceres strafleddet til ℓ_1 -normen svarende til strafleddet for lasso.

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable

MAOK9 5.2018

3

Lasso og dens generaliseringer

Data

Benchmark modellen

Lasso modellen og dens generaliseringer

Out-of-sample

Generaliseringer af lasso estimatoren

Group lasso



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

4

Lasso og dens
generaliseringer

Data

Benchmark modellen

Lasso modellen og dens
generaliseringer

Out-of-sample

Lad \mathbf{z}_j og $\boldsymbol{\theta}_j$ være $p_j \times 1$ vektorer som repræsenterer prædiktorerne og deres koefficienter i gruppe j for $j = 1, \dots, J$. Group lasso løser følgende optimeringsproblem

$$\hat{\boldsymbol{\theta}}_j^{\text{group lasso}} = \arg \min_{\boldsymbol{\theta}_j \in \mathbb{R}^{p_j}} \left\{ \frac{1}{2} \sum_{i=1}^n \left(y_i - \sum_{j=1}^J \mathbf{z}_{ij}^T \boldsymbol{\theta}_j \right)^2 + \lambda \sqrt{p_j} \sum_{j=1}^J \|\boldsymbol{\theta}_j\|_2 \right\}.$$

- ▶ Alle indgange i $\hat{\boldsymbol{\theta}}_j^{\text{group lasso}}$ vil være lig nul eller ikke-nul afhængig af λ .
- ▶ Når $p_j = 1$, da har vi, at $\|\boldsymbol{\theta}_j\|_2 = |\boldsymbol{\theta}_j|$, således at alle grupper består af én prædiktor, dermed reduceres optimeringsproblemet til standard lasso.

Generaliseringer af lasso estimatoren

Adaptive lasso



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Antag $\hat{\beta}$ er rod-n konsistent til β^* . Vælg $\gamma > 0$, da er adaptive lasso estimerne givet ved

$$\hat{\beta}^{\text{AL}} = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^p \frac{|\beta_j|}{|\hat{\beta}_j|^\gamma} \right\}.$$

- ▶ Antag $\frac{\lambda_n}{\sqrt{n}} \rightarrow 0$ og $\lambda_n n^{\frac{\gamma-1}{2}} \rightarrow \infty$, da opfylder adaptive lasso orakelegenskaberne:
 - ▶ Konsistent variabeludvælgelsen: $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{A}_n^{\text{AL}} = \mathcal{A}) = 1$.
 - ▶ Asymptotisk normalitet: $\sqrt{n} \left(\hat{\beta}_{\mathcal{A}}^{\text{AL}} - \beta_{\mathcal{A}}^* \right) \xrightarrow{d} N(\mathbf{0}, \sigma^2 \mathbf{C}_{11}^{-1})$.

5

Lasso og dens
generaliseringer

Data

Benchmark modellen

Lasso modellen og dens
generaliseringer

Out-of-sample

► Datasæt fra FRED

- 128 variable
- 1. januar 1959 - 1. november 2017 (707 observationer)
- Opdelt i 8 grupper:
 1. Output og indkomst
 2. Arbejdsmarked
 3. Bolig
 4. Forbrug, ordrer og varebeholdninger
 5. Penge og kredit
 6. Renter og valutakurser
 7. Priser
 8. Aktiemarked

► Transformerede datasæt

- 123 variable
- 1. januar 1960 - 1. juli 2017 svarende til 691 variable
 - Træningsmængde: 1. januar 1960 - 1. december 2005 (552 observationer)
 - Testmængde: 1. januar 2006 - 1. juli 2017 (139 observationer)
- centre responsvariablen og standardiser prædiktorerne

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

6

Data

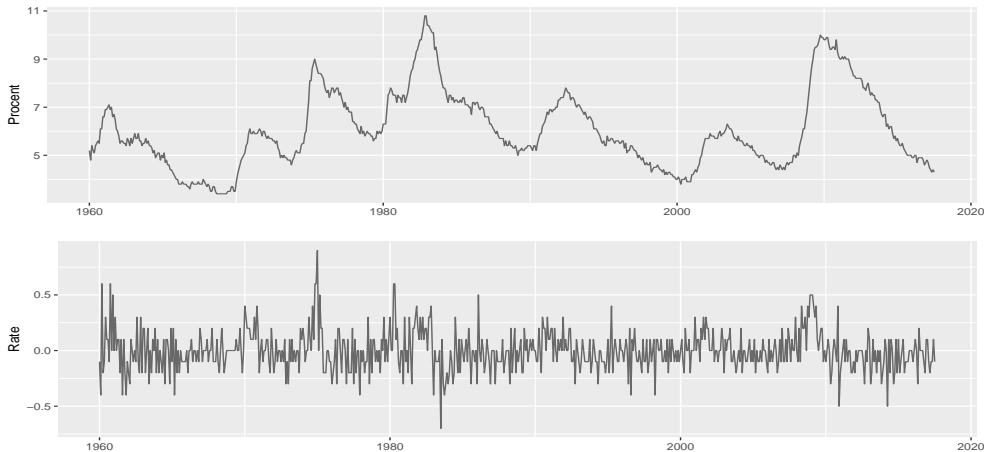
Benchmark modellen

Lasso modellen og dens generaliseringer

Out-of-sample

Data

Arbejdsløshedsraten



Inferens i lasso
modellen med
anvendelse inden for
prædiktions af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

7

Data

Benchmark modellen

Lasso modellen og dens
generaliseringer

Out-of-sample

Figur: Den øverste figur viser arbejdsløshedsraten og den nederste figur illustrerer 1. differensen af arbejdsløshedsraten fra 1. januar 1960 til 1. juli 2017

Benchmark modellen

Den autoregressive model



- ordenen bestemmes ud fra BIC

$\hat{\varphi}_1$	-0.0162
$\hat{\varphi}_2$	0.1992***
$\hat{\varphi}_3$	0.1873***
$\hat{\varphi}_4$	0.1686***
BIC	-3.5651
R^2_{adj}	12.31%
LogLik	211.8617

Tabel: Estimationsresultater for en AR (4), BIC, justeret R^2 og log-likelihood. Det opløftede symbol betegner signifikans ved henholdsvis ***0.1%, **1%, *5% og †10%.

Inferens i lasso modellen med anvendelse inden for prædiktion af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

Data

8

Benchmark modellen

Lasso modellen og dens generaliseringer

Out-of-sample

Benchmark modellen

Den autoregressive model



Inferens i lasso
modellen med
anvendelse inden for
prædiktions af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

9 Benchmark modellen

Lasso modellen og dens
generaliseringer

Out-of-sample

Skewness	0.2666
Kurtosis	1.4773
JB-test	$2.535 \cdot 10^{-13}$
LB ₁₀ -test	0
MAE	0.1312
MSE	0.0272

Tabel: Skewness, excess kurtosis og p -værdier for Jarque-Bera og Ljung-Box testen for de standardiserede residualer af en AR (4). Vi lader LB_{10} betegne Ljung-Box testen med lag = 10. p -værdier $< 2.2 \cdot 10^{-16}$ sættes til 0.

Benchmark modellen

Faktor modellen

- ▶ Lad $k = 1, \dots, k_{\max}$, hvor $k_{\max} = 20$
- ▶ ordenen bestemmes ud fra BIC



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

10 **Benchmark modellen**

Lasso modellen og dens
generaliseringer

Out-of-sample

Benchmark modellen

Faktor modellen



Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

Data

11 Benchmark modellen

Lasso modellen og dens generaliseringer

Out-of-sample

Faktor model (IC_1)				
	Værdi	IC_1	R^2_{adj}	LogLik
k	6	-0.3519	15.79%	224.3621

Faktor model (IC_2)				
	Værdi	IC_2	R^2_{adj}	LogLik
k	11	-0.5314	16.85%	230.3414

Faktor model (IC_3)				
	Værdi	IC_3	R^2_{adj}	LogLik
k	20	-0.6931	17.87%	238.3753

Tabel: Antal faktorer, værdien af informationskriteriet, justeret R^2 samt log-likelihood for faktormodellerne valgt ud fra IC_1 , IC_2 og IC_3 , som betegnes faktor model (IC_1), faktor model (IC_2) og faktor model (IC_3).

Benchmark modellen

Faktor modellen



Inferens i lasso
modellen med
anvendelse inden for
prædiktions af
makroøkonomiske
variable
MAOK9 5.2018

Lasso og dens
generaliseringer

Data

12 **Benchmark modellen**

Lasso modellen og dens
generaliseringer

Out-of-sample

	Faktor model (IC_1)	Faktor model (IC_2)
Skewness	0.0444	-0.0418
Kurtosis	0.5768	0.4612
JB-test	0.0172	0.0712
LB ₁₀ -test	0.729	0.4637
MAE	0.1190	0.1111
MSE	0.0221	0.0187

Tabel: Skewness, excess kurtosis, p -værdier for Jarque-Bera og Ljung-Box testen for de standardiserede residualer fra faktor modellerne valgt ud fra IC_1 og IC_2 . Vi lader LB_{10} betegne Ljung-Box testen med lag = 10.

Fraktion af ℓ_1 -norm ...

Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

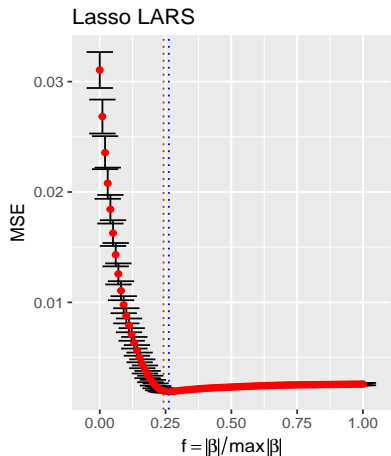
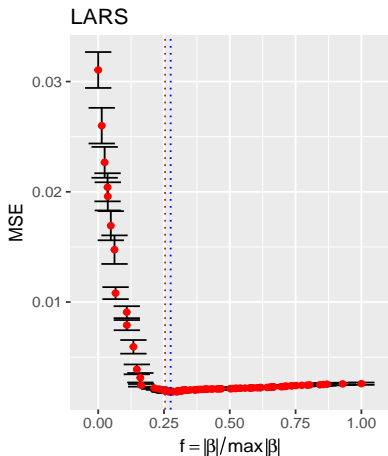
13

Lasso modellen og dens
generaliseringer

Out-of-sample

28

Aalborg universitet



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

14

Lasso modellen og dens
generaliseringer

Out-of-sample

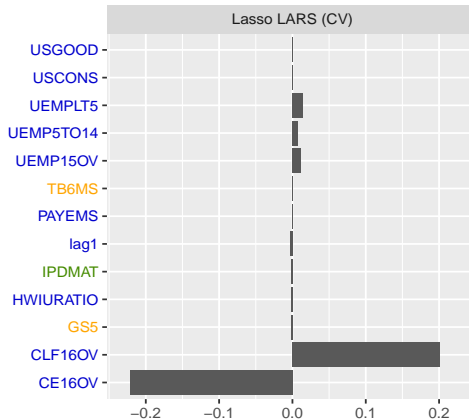
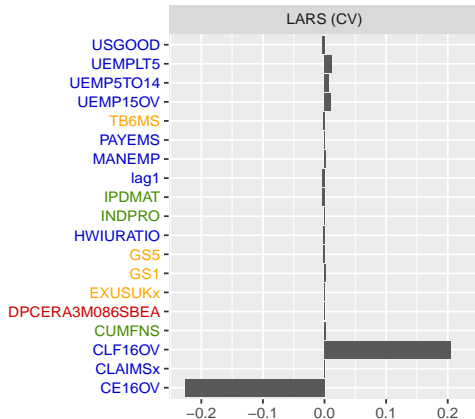
Figur: 10-fold krydsvalideringsfejl som funktion af fraktionen af ℓ_1 -normen LARS og lasso LARS. De stiplede linier indikerer f_{\min} og f_{1sd} .

LARS (CV)						Lasso LARS (CV)					
	Værdi	MSE	p	R^2_{adj}	LogLik		Værdi	MSE	p	R^2_{adj}	Loglike
f_{\min}	0.2753	0.0019	27	94.43%	974.8317	f_{\min}	0.2626	0.0019	21	94.52%	980.0982
f_{1sd}	0.2542	0.0019	19	94.19%	967.2669	f_{1sd}	0.2424	0.0019	13	94.43%	971.6687

Tabel: Værdien af f_{\min} og f_{1sd} , gennemsnitlig krydsvalideringsfejl, som er målt i MSE, antallet af parametre, justeret R^2 og log-likelihood for LARS og lasso LARS. De valgte tuning parametre er markeret med tykt.

- 22 trin udføres for lasso LARS (CV), hvor variablerne ...

Estimerede koefficienter



Inferens i lasso modellen med anvendelse inden for prædiktio af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

Data

Benchmark modellen

16

Lasso modellen og dens generaliseringer

Out-of-sample

Kovarianstesten

Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

17 Lasso modellen og dens
generaliseringer

Out-of-sample



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

18

Lasso modellen og dens
generaliseringer

Out-of-sample

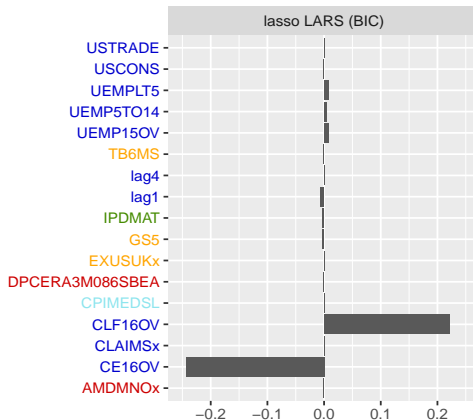
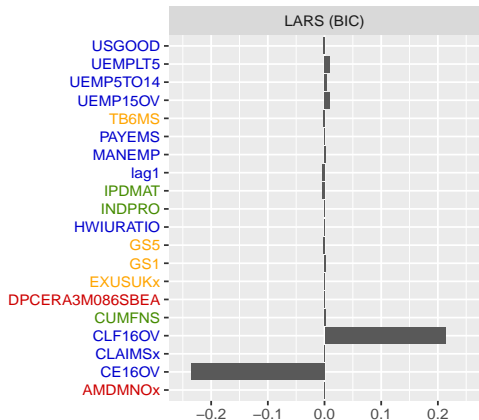
28

Aalborg universitet

LARS (BIC)						Lasso LARS (BIC)					
	Værdi	BIC	p	R^2_{adj}	LogLik		Værdi	BIC	p	R^2_{adj}	LogLik
f_{BIC}	0.2623	-6.0925	20	94.43%	975.2909	f_{BIC}	0.2604	-6.1627	17	94.46%	974.9938

Tabel: Værdien af f_{BIC} , antallet af parametre, BIC, justeret R^2 og log-likelihood for LARS og lasso LARS.

- 32 trin udføres for lasso LARS (BIC) ..



Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

Data

Benchmark modellen

20 Lasso modellen og dens generaliseringer

Out-of-sample

Figur: Estimerede koefficienter for LARS (BIC) og lasso LARS (BIC). Farverne indikerer hvilken gruppe variablerne tilhører

Kovarians testen

Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

21 Lasso modellen og dens
generaliseringer

Out-of-sample

TG testen

Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

22 Lasso modellen og dens
generaliseringer

Out-of-sample

Oversigt over in-sample resultater



Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

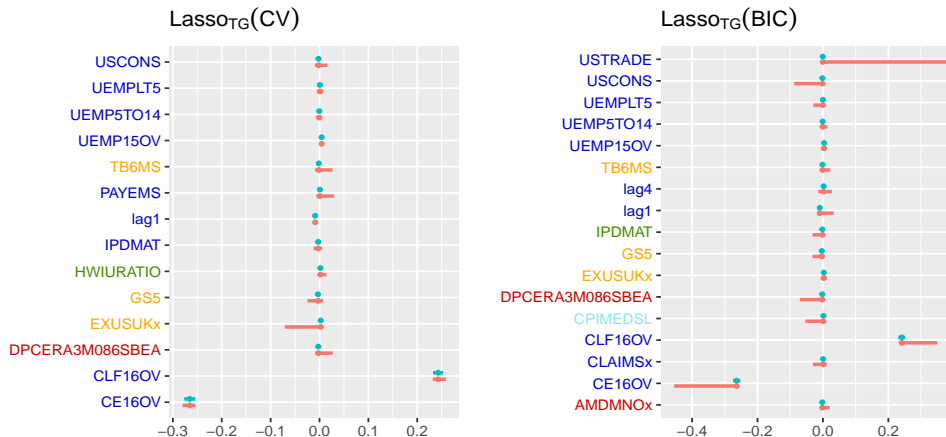
Data

Benchmark modellen

23 Lasso modellen og dens
generaliseringer

Out-of-sample

Oversigt over in-sample resultater



Figur: Betinget 95% konfidensinterval for lasso_{TG} (CV) og lasso_{TG} (BIC) (rød) og 95% konfidensinterval for OLS-estimatorerne for de udvalgte prædiktorer (blå)

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable
MAOK9 5.2018

Lasso og dens generaliseringer

Data

Benchmark modellen

24 Lasso modellen og dens generaliseringer

Out-of-sample

Out-of-sample



	MAE	R^{MAE}	MSE	R^{MSE}
Benchmark model	0.1111	1	0.0187	1
AR(4)	0.1312	1.1811	0.0272	1.454
Faktor model (IC ₁)	0.119	1.0717	0.0221	1.1798
Lasso (CV)	0.032	0.2877	0.0016	0.0876
Lasso (BIC)	0.0308	0.277	0.0015	0.0795
Ridge regression (CV)	0.0582	0.5239	0.0052	0.28
Ridge regression (BIC)	0.0573	0.5155	0.0051	0.2706
Group lasso (CV)	0.0352	0.3168	0.0019	0.1042
Group lasso (BIC)	0.0382	0.3437	0.0022	0.1202
Adap. lasso m. OLS vægte (CV)	0.0304	0.2733	0.0014	0.0729
Adap. lasso m. OLS vægte (BIC)	0.0310	0.2787	0.0014	0.0743
Adap. lasso m. lasso vægte (CV)	0.0298	0.2684	0.0013	0.0716
Adap. lasso m. lasso vægte (BIC)	0.0304	0.274	0.0014	0.0729
Lasso _{TG} (CV)	0.0303	0.2724	0.0014	0.0744
Lasso _{TG} (BIC)	0.031	0.279	0.0014	0.0767
LARS (CV)	0.0307	0.2761	0.0015	0.0802
LARS (BIC)	0.0305	0.2747	0.0015	0.0793
Lasso LARS (CV)	0.0352	0.317	0.002	0.1089
Lasso LARS (BIC)	0.0322	0.2901	0.0017	0.0903
LARS _{TG} (CV)	0.0300	0.2701	0.0014	0.0745
LARS _{TG} (BIC)	0.0301	0.2708	0.0014	0.0750

Tabel: Den gennemsnitlige absolutte og kvadrerede fejl samt gennemsnitlig tabs ratio mellem hver model og benchmark modellen.

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable
MAOK9 5.2018

Lasso og dens generaliseringer

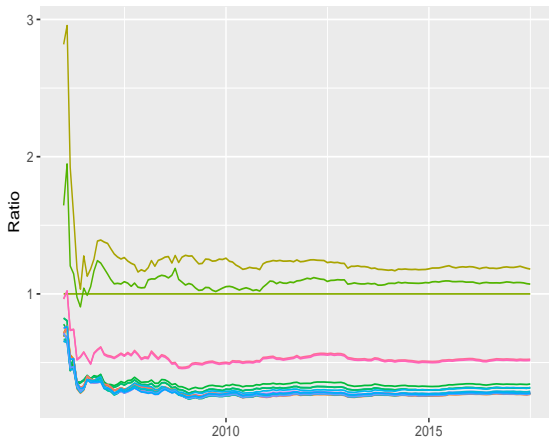
Data

Benchmark modellen

Lasso modellen og dens generaliseringer

25 Out-of-sample

Out-of-sample



- | | |
|-------------------------------|----------------------------------|
| Adaptive lasso m. lasso (BIC) | LARS m. lasso modifikation (BIC) |
| Adaptive lasso m. lasso (CV) | LARS m. lasso modifikation (CV) |
| Adaptive lasso m. OLS (BIC) | LARS TG (BIC) |
| Adaptive lasso m. OLS (CV) | LARS TG (CV) |
| AR(4) | Lasso (BIC) |
| Benchmark model | Lasso (CV) |
| Faktor model (IC1) | Lasso TG (BIC) |
| Group lasso (BIC) | Lasso TG (CV) |
| Group lasso (CV) | Ridge regression (BIC) |
| LARS (BIC) | Ridge regression (CV) |
| LARS (CV) | |

Inferens i lasso
modellen med
anvendelse inden for
prædiktion af
makroøkonomiske
variable

MAOK9 5.2018

Lasso og dens
generaliseringer

Data

Benchmark modellen

Lasso modellen og dens
generaliseringer

26

Out-of-sample

28

Aalborg universitet

Figur: Rullende gennemsnitlig absolut tabs ratio.

Out-of-sample

Diebold-Mariano testen



	Absolutte fejl	Kvadrerede fejl
AR(4)	0.0021	0.0032
Faktor model (IC_1)	0.1692	0.1183
Lasso (CV)	0	$2.933 \cdot 10^{-12}$
Lasso (BIC)	0	$2.728 \cdot 10^{-12}$
Ridge regression (CV)	$6.418 \cdot 10^{-13}$	$3.551 \cdot 10^{-9}$
Ridge regression (BIC)	$2.85 \cdot 10^{-13}$	$2.507 \cdot 10^{-9}$
Group lasso (CV)	0	$5.999 \cdot 10^{-12}$
Group lasso (BIC)	0	$8.845 \cdot 10^{-12}$
Adap. lasso m. OLS vægte (CV)	0	$2.797 \cdot 10^{-12}$
Adap. lasso m. OLS vægte (BIC)	0	$2.905 \cdot 10^{-12}$
Adap. lasso m. lasso vægte (CV)	0	$2.676 \cdot 10^{-12}$
Adap. lasso m. lasso vægte (BIC)	0	$2.814 \cdot 10^{-12}$
Lasso _{TG} (CV)	0	0
Lasso _{TG} (BIC)	0	0
LARS (CV)	0	$2.64 \cdot 10^{-12}$
LARS (BIC)	0	$2.615 \cdot 10^{-12}$
Lasso LARS (CV)	0	$4.694 \cdot 10^{-12}$
Lasso LARS (BIC)	0	$3.328 \cdot 10^{-12}$
LARS _{TG} (CV)	0	0
LARS _{TG} (BIC)	0	0

Tabel: p -værdier for Diebold-Mariano testen for hver model imod benchmark modellen. p -værdier $< 2.2 \cdot 10^{-16}$ sættes til 0.

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable
MAOK9 5.2018

Lasso og dens generaliseringer

Data

Benchmark modellen

Lasso modellen og dens generaliseringer

27

Out-of-sample

28

Aalborg universitet

Out-of-sample

MCS



T_R		T_{max}	
$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.2$
Benchmark model	Benchmark model	Benchmark model	Benchmark model
AR(4)	AR(4)	AR(4)	AR(4)
Lasso (CV)	Lasso (CV)	Faktor (IC ₁)	Lasso (CV)
Lasso (BIC)	Lasso (BIC)	Lasso (CV)	Lasso (BIC)
Group lasso (CV)	Group lasso (CV)	Lasso (BIC)	Ridge regression (CV)
Group lasso (BIC)	Group lasso (BIC)	Ridge regression (CV)	Ridge regression (BIC)
Adap. lasso m. OLS vægte (CV)	Adap. lasso m. OLS vægte (CV)	Ridge regression (BIC)	Group lasso (CV)
Adap. lasso m. OLS vægte (BIC)	Adap. lasso m. OLS vægte (BIC)	Group lasso (CV)	Group lasso (BIC)
Adap. lasso m. lasso vægte (CV)	Adap. lasso m. lasso vægte (CV)	Group lasso (BIC)	Adap. lasso m. OLS vægte (CV)
Adap. lasso m. lasso vægte (BIC)	Adap. lasso m. lasso vægte (BIC)	Adap. lasso m. OLS vægte (CV)	Adap. lasso m. OLS vægte (BIC)
Lasso _{TG} (BIC)	Lasso _{TG} (BIC)	Adap. lasso m. lasso vægte (CV)	Adap. lasso m. lasso vægte (CV)
LARS (CV)	LARS (CV)	Adap. lasso m. lasso vægte (BIC)	Adap. lasso m. lasso vægte (BIC)
LARS (BIC)	LARS (BIC)	Lasso _{TG} (CV)	Lasso _{TG} (CV)
Lasso LARS (CV)	Lasso LARS (CV)	Lasso _{TG} (BIC)	Lasso _{TG} (BIC)
Lasso LARS (BIC)	Lasso LARS (BIC)	LARS (CV)	LARS (CV)
		LARS (BIC)	LARS (BIC)
		Lasso LARS (CV)	Lasso LARS (CV)
		Lasso LARS (BIC)	Lasso LARS (BIC)
		LARS _{TG} (CV)	LARS _{TG} (CV)
		LARS _{TG} (BIC)	LARS _{TG} (BIC)

Tabel: 80% og 90% model confidence set for arbejdsløshedsraten for absolutte og kvadrerede fejl.

Inferens i lasso modellen med anvendelse inden for prædiktions af makroøkonomiske variable

MAOK9 5.2018

Lasso og dens generaliseringer

Data

Benchmark modellen

Lasso modellen og dens generaliseringer

28 Out-of-sample