

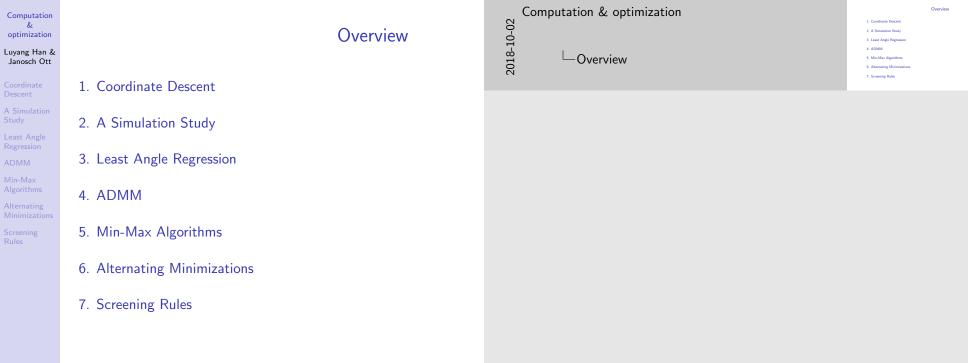
# 2018-10

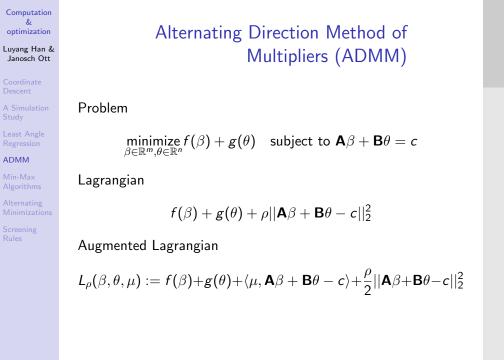
Computation & optimization for Lasso - part 2

Luyang Han & Janosch Ott

22 October 2018

Computation & optimization







Alternating Direction Method of

minimize  $f(\beta) + g(\theta)$  subject to  $\mathbf{A}\beta + \mathbf{B}\theta = c$ 

Augmented: scalar product with  $\mu$  gets added

Computation & optimization

-ADMM



Luyang Han &

# Dual variable update

Janosch Ott

**ADMM** 

 $\beta^{t+1} = \arg\min_{\beta \in \mathbb{R}^m} L_{\rho}(\beta, \theta^t, \mu^t)$ 

 $heta^{t+1} = \arg\min_{ heta \in \mathbb{R}^m} L_{
ho}(eta^{t+1}, heta, \mu^t)$ 

 $\mu^{t+1} = \mu^t + \rho(\mathbf{A}\beta^{t+1} + \mathbf{B}\theta^{t+1} - c)$ 

Computation & optimization 2018-10-02 -ADMM └─Dual variable update

 $\beta^{t+1} = \arg \min_{\alpha, \dots, \mu} L_{\rho}(\beta, \theta^t, \mu^t)$  $\theta^{t+1} = \arg \min_{\rho = 0} L_{\rho}(\beta^{t+1}, \theta, \mu^{t})$  $\mu^{t+1} = \mu^t + \rho(\mathbf{A}\beta^{t+1} + \mathbf{B}\theta^{t+1} - c)$ 

Dual variable update

# Computation & optimization

Luyang Han &

Janosch Ott

ADMM - Why?

- Coordina
- A Simulation Study
- Least Ang

. . . . . . . .

### ADMM

Min-Max Algorithms

Alternating

Screening

- convex problems with nondifferentiable constraints
- blockwise computation
  - sample blocks
  - feature blocks

Computation & optimization COMPUTE COM

convex problems with nondifferentiable constraints
 blockwise computation
 sample blocks
 feature blocks

ADMM - Why?



# ADMM for the Lasso

Computation















# where $S_{\lambda/\rho}(z) = \operatorname{sign}(z)(|z| - \frac{\lambda}{\rho})_+$ .











# Update

Problem in Lagrangian form

 $\theta^{t+1} = \mathcal{S}_{\lambda/\rho}(\beta^{t+1} + \mu^t/\rho)$  $\mu^{t+1} = \mu^t + \rho(\beta^{t+1} - \theta^{t+1})$ 

$$\left\{ eta eta eta _2^2 + \lambda \left\| heta 
ight\|_1 
ight\} \quad ext{such that } eta$$

 $\beta^{t+1} = (\mathbf{X}^T \mathbf{X} + \rho \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{y} + \rho \theta^t - \mu^t)$ 

# $\underset{\beta \in \mathbb{R}^{p}, \theta \in \mathbb{R}^{p}}{\text{minimize}} \left\{ \frac{1}{2} \left\| \mathbf{y} - \mathbf{X}\beta \right\|_{2}^{2} + \lambda \left\| \theta \right\|_{1} \right\} \quad \text{such that } \beta - \theta = 0$

- 2018-10-02 -ADMM

# □ADMM for the Lasso

Computation & optimization

### $\underset{\boldsymbol{\lambda},\boldsymbol{\alpha},\boldsymbol{\theta}}{\text{minimize}} \left\{ \frac{1}{2} \| \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \|_2^2 + \lambda \|\boldsymbol{\theta}\|_1 \right\} \quad \text{such that } \boldsymbol{\beta} - \boldsymbol{\theta} = \mathbf{0}$ $\beta^{t+1} = (\mathbf{X}^T \mathbf{X} + a\mathbf{I})^{-1} (\mathbf{X}^T \mathbf{v} + a\theta^t - u^t)$

ADMM for the Lasso





Computation optimization

Dual Polytope Projection (DPP)

Luyang Han & Janosch Ott

Screening

Rules

Suppose we want to calculate a lasso solution at  $\lambda < \lambda_{max}$ . The DPP rule discards the *i*<sup>th</sup> variable if

$$\left\|\mathbf{x}_{j}^{\mathcal{T}}\mathbf{y}
ight\|<\lambda_{\mathsf{max}}-\left\|\mathbf{x}_{j}
ight\|_{2}\left\|\mathbf{y}
ight\|_{2}rac{\lambda_{\mathsf{max}}-\lambda}{\lambda}$$

Sequential DPP rule

Suppose we have the lasso solution  $\hat{\beta}(\lambda')$  at  $\lambda'$  and want to screen variables for solutions at  $\lambda < \lambda'$ . We discard the  $i^{th}$ variable if

$$\left|\mathbf{x}_{j}^{T}(\mathbf{y}-\mathbf{X}\hat{eta}(\lambda'))
ight|<\lambda'-\left\|\mathbf{x}_{j}
ight\|_{2}\left\|\mathbf{y}
ight\|_{2}rac{\lambda_{\mathsf{max}}-\lambda}{\lambda}$$

Computation & optimization -Screening Rules

□ Dual Polytope Projection (DPP)

Dual Polytope Projection (DPP)

 $\left|\mathbf{x}_{j}^{T}\mathbf{y}\right|<\lambda_{\max}-\left\|\mathbf{x}_{j}\right\|_{2}\left\|\mathbf{y}\right\|_{2}\frac{\lambda_{\max}-\lambda}{\epsilon}$ 

Sequential DPP rule

Suppose we have the lasso solution  $\hat{\beta}(\lambda')$  at  $\lambda'$  and want to screen variables for solutions at  $\lambda < \lambda'$ . We discard the  $i^{ti}$ 

 $|\mathbf{x}_{i}^{T}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}(\lambda'))| < \lambda' - ||\mathbf{x}_{i}||_{2} ||\mathbf{y}||_{2} \frac{\lambda_{\max} - \lambda}{2}$ 



# Global Strong Rule

Screening Rules

Suppose we want to calculate a lasso solution at  $\lambda < \lambda_{\text{max}}$ . The global strong rule discards the  $i^{th}$  variable if

$$\left|\mathbf{x}_{j}^{T}\mathbf{y}\right|<\lambda-\left(\lambda_{\mathsf{max}}-\lambda
ight)=2\lambda-\lambda_{\mathsf{max}}$$

# Sequential Strong Rule

Suppose we have the lasso solution  $\hat{\beta}(\lambda')$  at  $\lambda'$  and want to screen variables for solutions at  $\lambda < \lambda'$ . We discard the  $i^{th}$ variable if

$$\left|\mathbf{x}_{j}^{\mathcal{T}}(\mathbf{y}-\mathbf{X}\hat{eta}(\lambda'))
ight|<2\lambda-\lambda'$$

Computation & optimization -Screening Rules 2018-1 -Global Strong Rule

### Global Strong Rule

 $|\mathbf{x}_{i}^{T}\mathbf{y}| < \lambda - (\lambda_{\text{max}} - \lambda) = 2\lambda - \lambda_{\text{max}}$ Sequential Strong Rule

Suppose we have the lasso solution  $\hat{\beta}(\lambda')$  at  $\lambda'$  and want to  $|\mathbf{x}^T(\mathbf{v} - \mathbf{X}\hat{\boldsymbol{\beta}}(\lambda'))| < 2\lambda - \lambda'$ 

Computation & optimization

Luyang Han &

Paragraphs of Text

Janosch Ott

scent

Study Least Angl

Regressio

lin-Max

ternating

Minimization Screening

Rules

Sed diam enim, sagittis nec condimentum sit amet, ullamcorper sit amet libero. Aliquam vel dui orci, a porta odio. Nullam id suscipit ipsum. Aenean lobortis commodo sem, ut commodo leo gravida vitae. Pellentesque vehicula ante iaculis arcu pretium rutrum eget sit amet purus. Integer ornare nulla quis neque ultrices lobortis. Vestibulum ultrices tincidunt libero, quis commodo erat ullamcorper id.

Sed iaculis dapibus gravida. Morbi sed tortor erat, nec interdum arcu. Sed id lorem lectus. Quisque viverra augue id sem ornare non aliquam nibh tristique. Aenean in ligula nisl. Nulla sed tellus ipsum. Donec vestibulum ligula non lorem

vulputate fermentum accumsan neque mollis.

Computation & optimization

Screening Rules

—Paragraphs of Text

### Paragraphs of Text

Sed izculis dapibus gravida. Morbi sed tortor erat, nec interdum arcu. Sed id lorem lectus. Quisque viverra augue id sem ornare non aliquam nibh tristique. Anntan in ligula nid. Nulla sed tellus ipsum. Donec vestibulum ligula non lorem vulgotate fermentum accumsian nique mollis.

Sed d'am neim, sajittis nec condimentum sit amer, ullancope sit amet libero. Alquam vel dis criç a porta odio. Nullam si suscipit lipsum. Aensan loborisi commodo sem, ut commodo los gavida vitate. Pheltenseque velicita ante isculis sem pretium ratrum eget sit amet parus. Integer ornare nulla quis neque attrices bloristi. Vestibulum strices sincidant libero, quis commodo erast ullamocorper id. Computation optimization

Luyang Han & Janosch Ott

Screening Rules

## **Bullet Points**

- Lorem ipsum dolor sit amet, consectetur adipiscing elit
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu
- Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- Nam cursus est eget velit posuere pellentesque
- Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Computation & optimization 10-02 Screening Rules 2018--Bullet Points

### **Bullet Points**

- . Lorem ipsum dolor sit amet, consectetur adipiscing elit · Aliquam blandit faucibus nisi, sit amet dapibus enim
- . Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- · Nam cursus est eget velit posuere pellentesque
- . Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Computation Blocks of Highlighted Text optimization Luyang Han & Janosch Ott Block 1 Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor. Block 2 Pellentesque sed tellus purus. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Screening Vestibulum quis magna at risus dictum tempor eu vitae velit. Rules Block 3 Suspendisse tincidunt sagittis gravida. Curabitur condimentum, enim sed venenatis rutrum, ipsum neque consectetur orci, sed blandit justo nisi ac lacus.

## Blocks of Highlighted Text

Block 1 Lorem ipsum dolor sit amet, consectetur adipiscing elit. Intege lectus nisl, ultricissi in feugiat rutrum, portitior sit amet augue. Aliquam ut to tortor mauris. Sed volutpat ante purus, quis accumsian dolor.

Pellentesque sed tellus purus. Class aptent taciti sociosqu ad litera torquent per conubia nostra, per inceptos himenaeos. Vestibulum quis magna at risus dictum tempor eu vitae velit

Block 3

Computation & optimization

-Blocks of Highlighted Text

-Screening Rules

Block 3 Suspendisse tincidunt sagittis gravida. Curabitur condimentun enim sed venenatis rutrum, ipsum neque consectetur orci, sed blandit justo nisi ac lacus. Computation & optimization

Luyang Han & Janosch Ott

A Simulation

Regression

Heading

Statement

2 Explanation

3 Example

ADMM

ADIVIN

Min-Max Algorithms

Alternating

Screening Rules

# Multiple Columns

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

Computation & optimization
Screening Rules

Multiple Columns

Heading

Statement
Explanation
Example

Lorem ipsum dolor sit amet, consectetur adipiciolig elit. Integer lectus nisl, ultricios in feugiat rutrum, portitior sit amet augue. Aliquam ut torto mauris. Sed volutpat ante purus, quis accumsan dolor.

Multiple Columns



Janosch Ott

Table

Alternating

Screening

Table: Table caption Rules

Treatment 1 Treatment 2

Treatment 3

**Treatments Response 1** 

Response 2 0.562 0.910 0.296

0.0003262

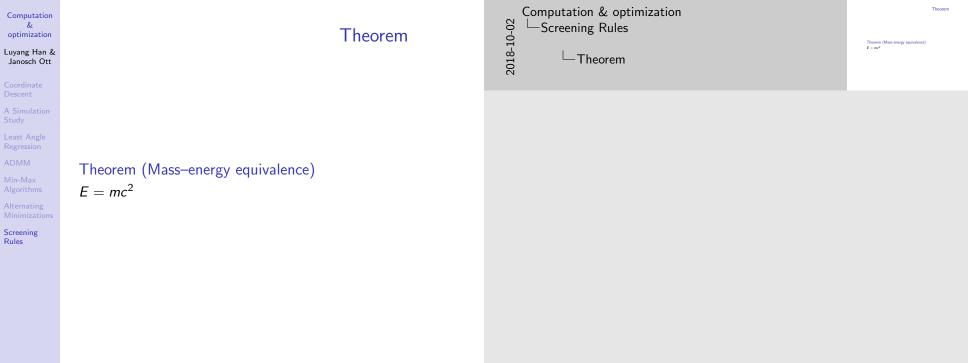
0.0015681

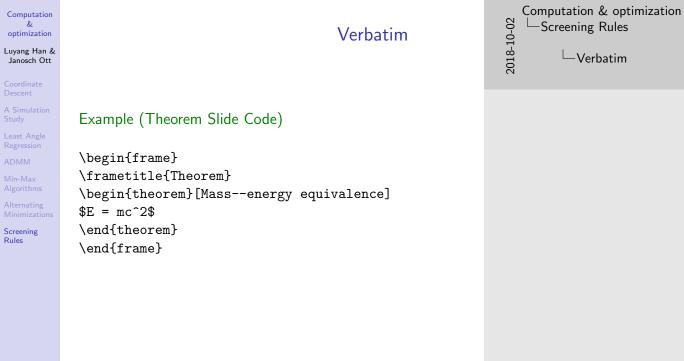
0.0009271

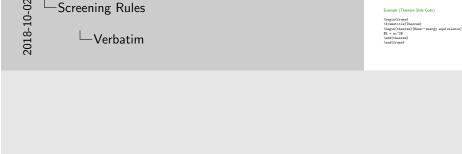
2018-10-02 -Screening Rules └─Table

Computation & optimization

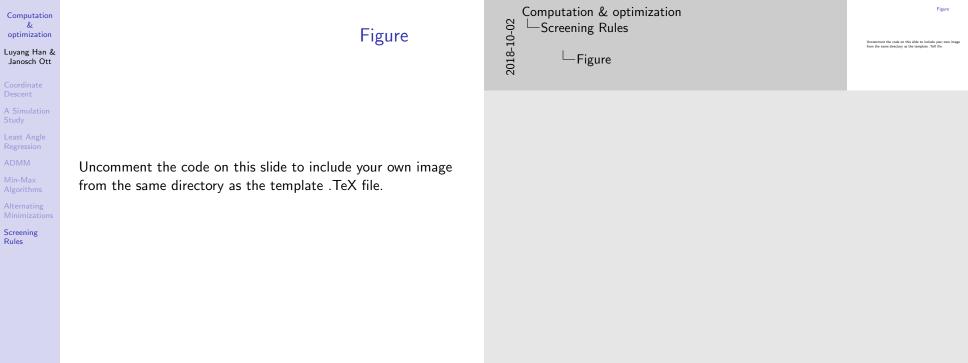
Treatments Response 1 Response 2 Treatment 1 0.0003262 0.562 Treatment 2 0.0015681 0.910 Treatment 3 0.0009271 0.296 Table: Table caption

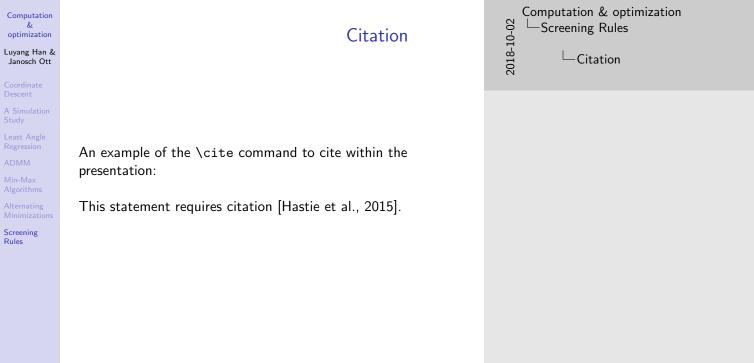






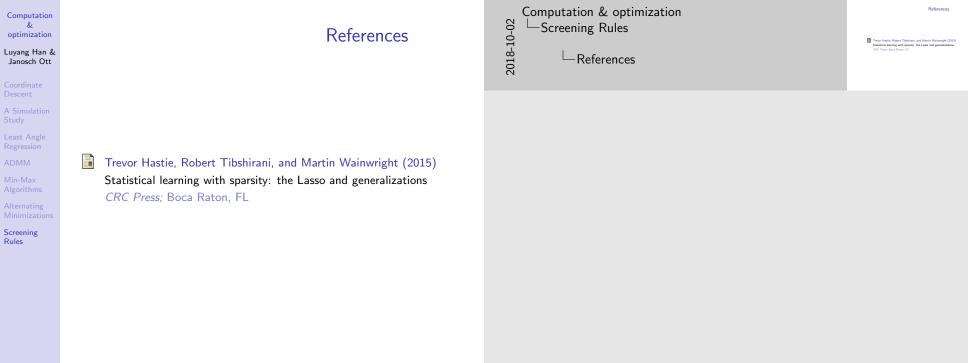
Verbatim





An example of the \cite command to cite within the presentation:

This statement requires citation [Hastie et al., 2015].





Screening Rules

The End

Computation & optimization