Explaining Success in Sports Competitions: Paired Comparison Methods with Explanatory Variables

Gerhard Tutz und Gunther Schauberger
Ludwig-Maximilians-Universität
München

Padova
June 2017

Collaboration with Andreas Groll

Paired Comparison Data

- Sports competitions, experiments, ...
- Aim: measure unobservable latent trait for set of objects $\{a_1, \ldots, a_m\}$
- Comparison/Competition between two objects a_r and a_s
- Binary response

$$Y_{(r,s)} = \begin{cases} 1 & \text{if } a_r \text{ preferred over } a_s \\ 0 & \text{if } a_s \text{ preferred over } a_r \end{cases}$$

- Ordinal response

$$Y_{(r,s)} = \begin{cases} 1 & \text{if } a_r \text{ strongly preferred over } a_s \\ \vdots & \qquad\qquad \vdots \\ K & \text{if } a_s \text{ strongly preferred over } a_r \end{cases}$$

Bradley-Terry Model

Set of objects $\{a_1, \ldots, a_m\}$

for $Y_{(r,s)} = \begin{cases} 1 & \text{if } a_r \text{ preferred over } a_s \\ 0 & \text{if } a_s \text{ preferred over } a_r \end{cases}$

$$P(Y_{(r,s)} = 1) = \frac{\exp(\gamma_r - \gamma_s)}{1 + \exp(\gamma_r - \gamma_s)}, \quad \sum_{r=1}^{m} \gamma_r = 0$$

γ_r attractivity/strength of object r

γ_s attractivity/strength of object s

From binary to ordinal response

A match between teams a_r and a_s is treated as a paired comparison with ordinal response $Y_{(r,s)}$, with

$$Y_{(r,s)} = \begin{cases} 1 & \text{if team } a_r \text{ wins by at least 2 goals difference} \\ 2 & \text{if team } a_r \text{ wins by 1 goal difference} \\ 3 & \text{if the match ends with a draw} \\ 4 & \text{if team } a_s \text{ wins by 1 goal difference} \\ 5 & \text{if team } a_s \text{ wins by at least 2 goals difference.} \end{cases}$$

$$P(Y_{(r,s)} \leqslant k) = \frac{\exp(\theta_k + \gamma_r - \gamma_s)}{1 + \exp(\theta_k + \gamma_r - \gamma_s)}, \quad k = 1, \dots, 5$$

- θ_k: category-specific threshold parameters, $\theta_1 = -\theta_4$, $\theta_2 = -\theta_3$
- γ_r, γ_s: team-specific abilities, $\sum_{r=1}^{18} \gamma_r = 0$

Assumptions and Derivation

- **Unobservable random utility** U_r that represents ability of team a_r:

$$U_r = \gamma_r + \varepsilon_r,$$

  - γ_r is a fixed value (the fixed ability)
  - ε_r is a random variable (represents noise)

- Assume that $\varepsilon_1, \dots, \varepsilon_m$ are iid random variables with distribution function F_ε.
- Given the pair (a_r, a_s), one observes

$$Y_{(r,s)} = k \quad \Leftrightarrow \quad \theta_{k-1} < U_s - U_r < \theta_k,$$

  - Low categories k indicate dominance of a_r
  - High categories k indicate dominance of a_s

$\Rightarrow Y_{(r,s)}$ is a categorized/coarsened version of the differences in latent abilities.

Ordinal Bradley-Terry Model

From

$$Y_{(r,s)} = k \quad \Leftrightarrow \quad \theta_{k-1} < U_s - U_r < \theta_k$$

we derive

$$\begin{aligned} Y_{(r,s)} \leqslant k \quad &\Leftrightarrow \quad U_s - U_r < \theta_k \\ Y_{(r,s)} \leqslant k \quad &\Leftrightarrow \quad \varepsilon_s - \varepsilon_r < \theta_k + \gamma_r - \gamma_s \end{aligned}$$

and

$$P(Y_{(r,s)} \leqslant k | (r,s)) = F(\eta_{rsk}), \quad \eta_{rsk} = \theta_k + \gamma_r - \gamma_s$$

where $F(\cdot)$ is the distribution of the differences $\varepsilon_s - \varepsilon_r$.

Ordinal Bradley-Terry Model

From

$$Y_{(r,s)} = k \quad \Leftrightarrow \quad \theta_{k-1} < U_s - U_r < \theta_k$$

we derive

$$\begin{aligned} Y_{(r,s)} \leqslant k \quad &\Leftrightarrow \quad U_s - U_r < \theta_k \\ Y_{(r,s)} \leqslant k \quad &\Leftrightarrow \quad \varepsilon_s - \varepsilon_r < \theta_k + \gamma_r - \gamma_s \end{aligned}$$

and

$$P(Y_{(r,s)} \leqslant k | (r,s)) = F(\eta_{rsk}), \quad \eta_{rsk} = \theta_k + \gamma_r - \gamma_s$$

where $F(\cdot)$ is the distribution of the differences $\varepsilon_s - \varepsilon_r$.

With $F(\cdot)$ as the logistic distribution function we get

$$P(Y_{(r,s)} \leqslant k | (r,s)) = \frac{exp(\theta_k + \gamma_r - \gamma_s)}{1 + exp(\theta_k + \gamma_r - \gamma_s)}$$

Ordinal Bradley-Terry Model

From

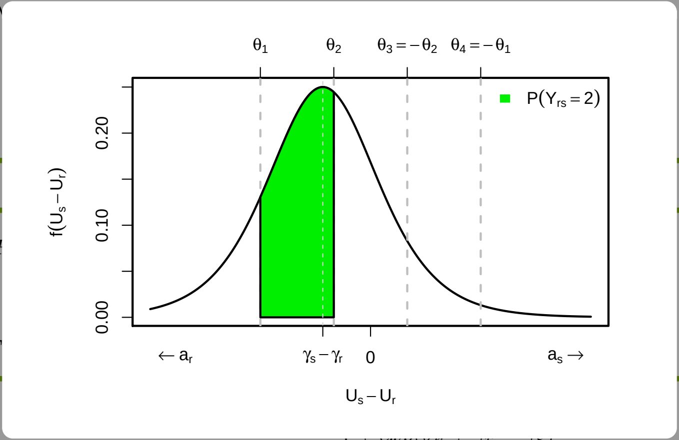
$$Y_{(r,s)} = k \quad \Leftrightarrow \quad \theta_{k-1} < U_s - U_r < \theta_k$$

we deriv

and

where F

With F



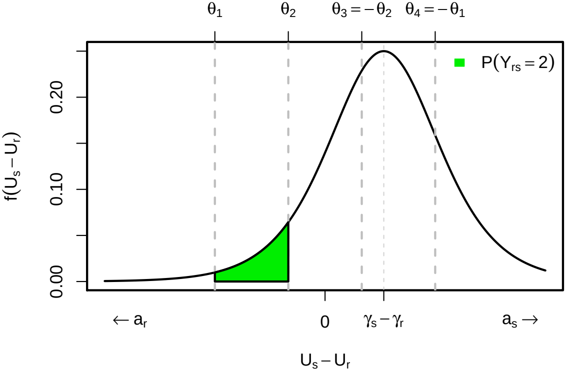
$1 + exp(\theta_k + \gamma_r - \gamma_s)$

Ordinal Bradley-Terry Model

From

$$Y_{(r,s)} = k \quad \Leftrightarrow \quad \theta_{k-1} < U_s - U_r < \theta_k$$

we deriv



and

where F

With F

Restrictions

Symmetric restrictions of threshold parameters:

- $\theta_k = -\theta_{K-k}$, $k = 1, \ldots, [K/2]$
  e.g. $K = 5 \Rightarrow \theta_1 = -\theta_4,\ \theta_2 = -\theta_3$
- (if K is even): $\theta_{K/2} = 0$

That means, that for teams a_r and a_s one obtains

$$P(Y_{(r,s)} = k) = P(Y_{(s,r)} = K + 1 - k).$$

For the special case $K = 5$ one obtains

$$P(Y_{(r,s)} = 1) = P(Y_{(s,r)} = 5)$$

and

$$P(Y_{(r,s)} = 2) = P(Y_{(s,r)} = 4)$$

Ordinal Model With Home/Order Effect

Possible order effects in sports:

- playing at home (football)
- serving (tennis)
- playing with the white pieces (chess)

Simplest case: binary response given pair (a_r, a_s)

$$a_r \quad \text{wins if} \quad U_r > U_s,$$

With home/order effect

$$a_r \quad \text{wins if} \quad U_r + \delta > U_s,$$

$\Rightarrow$ A constant δ is added to the first team (home team).

Ordinal Model With Home/Order Effect

Possible order effects in sports:

- playing at home (football)
- serving (tennis)
- playing with the white pieces (chess)

Simples

In the general case

$$\eta_{rsk} = \delta + \theta_k + \gamma_r - \gamma_s,$$

where $\delta > 0$ represents the order/home effect.

- If $\delta = 0$ no order/home effect
- If $\delta > 0$ large the probability for low categories (dominance of a_r) is increased

With ho

$$a_r \quad \text{wins if} \quad U_r + \delta > U_s,$$

$\Rightarrow$ A constant δ is added to the first team (home team).

Ordinal Model With Home/Order Effect

Possible

- pla
- ser
- pla

Simples

With ho

⇒ A co

$$\eta_{rsk} = \delta + \theta_k + \gamma_r - \gamma_s$$

Season 2015/16

$$\hat{\delta} = 0.265$$

$$\hat{\theta}_1 = -\hat{\theta}_4 = -1.591$$

$$\hat{\theta}_2 = -\hat{\theta}_3 = -0.576$$

Rank	Team	$\hat{\gamma}_r$	Rank($\hat{\gamma}_r$)
1	BAY	1.899	1
2	DOR	1.598	2
3	LEV	0.433	4
4	MGB	0.475	3
5	S04	0.133	5
6	MAI	0.088	6
7	BER	-0.001	7
8	WOB	-0.142	9
9	KOE	-0.045	8
10	HSV	-0.183	10
11	ING	-0.228	11
12	AUG	-0.363	13
13	BRE	-0.361	12
14	DAR	-0.467	15
15	HOF	-0.448	14
16	FRA	-0.623	16
17	STU	-0.699	17
18	HAN	-1.068	18

Explanatory Variables - Effect of Budget

A simple two-step approach:

- Fit a Bradley-Terry Model
- Investigate the dependence of abilities on explanatory variables

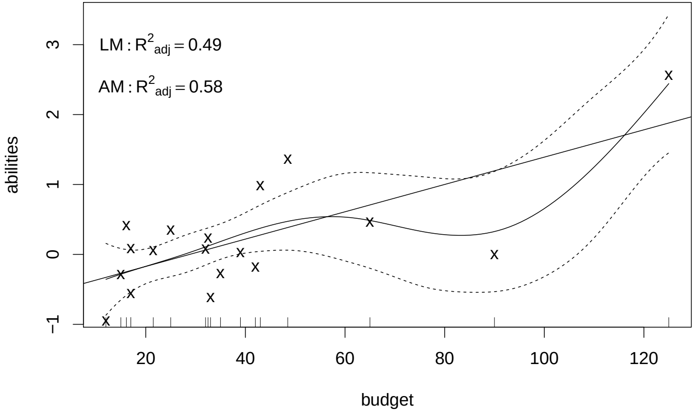


Figure: Budgets (in millions) versus estimated abilities for all teams from the Bundesliga season 2012/2013; lines represent linear and additive model fit

Two-step approach is not satisfactory

- BT-Model fit does not use the explanatory variables
- Inference difficult, in the second step abilities are considered as random variables, uncertainty of estimates ignored

Nowadays various explanatory variables of different types are available

Two-step approach is not satisfactory

- BT
- Inf... ...riables,
  un...

On-field covariates

Bayern Munich				Hamburger SV
Goals	5	:	0	Goals
Shots on goal	23	:	5	Shots on goal
Distance	108.54	:	111.28	Distance
Completion rate	90	:	64	Completion rate
Ball possession	77	:	23	Ball possession
Tackling rate	52	:	48	Tackling rate
Fouls	10	:	12	Fouls
Offside	3	:	0	Offside

Source: German football magazine kicker (http://www.kicker.de/)

Generalize Bradley-Terry Model

- Make use of all kinds of covariates in paired comparisons
- Link explanatory variables to final match outcome
- Set up paired comparison model that can include all types of covariates simultaneously
- Sparse, interpretable model

Generalize Bradley-Terry Model

- Make use of all kinds of covariates in paired comparisons
- Link explanatory variables to final match outcome
- Set up paired comparison model that can include all types of covariates simultaneously
- Sparse, interpretable model

$$P(Y_{(r,s)} = 1) = \frac{\exp(\gamma_r - \gamma_s)}{1 + \exp(\gamma_r - \gamma_s)}$$

$$\Downarrow$$

$$\gamma_r \rightarrow \gamma_{ir}$$

$\Rightarrow$ γ_{ir} can depend on **covariates varying over subjects and/or objects**

$$\Downarrow$$

$$P(Y_{i(r,s)} = 1) = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1 + \exp(\gamma_{ir} - \gamma_{is})}$$

Match-Specific Explanatory Variables

In classical paired comparison experiments several persons choose between objects

Let person i be characterized by covariates $\boldsymbol{x}_i$ (gender, age...). Covariates are subject-specific, do not vary over objects.

$$\gamma_{ir} = \beta_{r0} + \boldsymbol{x}_i^T \boldsymbol{\beta}_r$$

In Sports

- γ_{ir} is the strength of team a_r when meeting under circumstances captured by $\boldsymbol{x}_i$ like
  - Temperature when playing, raining or not,..
  - Time of the year
  - Type of tournament,...
- Each team may react differently to the circumstances, $\boldsymbol{\beta}_r$ is team-specific

Match-Team-Specific Explanatory Variables

More often one has explanatory variables that vary both over matches and team yielding match-team-specific (subject-object-specific) covariates $\boldsymbol{z}_{ir}$.

- Total amount of km run by team a_r in match i,
- Percentage of passes reaching team mates (team a_r in match i)
- ...

Effect of match-team-specific covariates can be modelled

$$\gamma_{ir} = \beta_{r0} + \boldsymbol{z}_{ir}^T \boldsymbol{\alpha}.$$

or

$$\gamma_{ir} = \beta_{r0} + \boldsymbol{z}_{ir}^T \boldsymbol{\alpha}_r.$$

Team-Specific Explanatory Variables

Some interesting explanatory variables characterize the team and do not vary over matches (object-specific covariates).

- Budget of team a_r (fixed over a season)
- Home country of the team/player
- ...

$$\gamma_{ir} = \gamma_r = \beta_{r0} + \boldsymbol{z}_r^T \boldsymbol{\tau}.$$

But: Effects of team-specific explanatory variables are not identifiable !

$$\gamma_r = \beta_{r0} + \boldsymbol{z}_r^T \boldsymbol{\tau} = \underbrace{\beta_{r0} + \boldsymbol{z}_r^T \boldsymbol{c}}_{\tilde{\beta}_{r0}} + \underbrace{\boldsymbol{z}_r^T (\boldsymbol{\tau} - \boldsymbol{c})}_{\boldsymbol{z}_r^T \tilde{\boldsymbol{\tau}}} = \tilde{\gamma}_r$$

Therefore γ_r and $\tilde{\gamma}_r$ yields the same model.

Regularization by Penalty Terms

Estimation via framework of multivariate GLMs

→ number of parameters can be huge!
Example:
$m = 18$ teams, $p = 8$ match-team-specific covariates
$\Rightarrow (m-1)p = 136$ additional parameters

⇒ Penalized likelihood estimation

Replace the usual log-likelihood by

$$l_p(\boldsymbol{\beta}) = l(\boldsymbol{\beta}) - \lambda J(\boldsymbol{\beta})$$

where

- $l(\boldsymbol{\beta})$ is the log-likelihood of a GLM,
- λ is a tuning parameter
- $J(\boldsymbol{\beta})$ is a penalty term

Simple Ridge Type Shrinkage Penalty

$$J(\boldsymbol{\beta}) = \sum_r \beta_r^2$$

Lasso-type Penalty: the Simple Case

Consider the simplest case without covariates

$$P(Y_{(r,s)} \leqslant k) = \frac{exp(\delta + \theta_k + \beta_{r0} - \beta_{s0})}{1 + exp(\delta + \theta_k + \beta_{r0} - \beta_{s0})}$$

Estimate by using the fusion penalty penalty

$$J(\boldsymbol{\gamma}) = \sum_{r>s} w_{rs} |\beta_{r0} - \beta_{s0}|,$$

With weights w_{rs} (optional).

- Shows which teams have different abilities
- Clusters of teams with equal strength are identified
- Reduces number of parameters

Lasso-type Penalty: the Simple Case

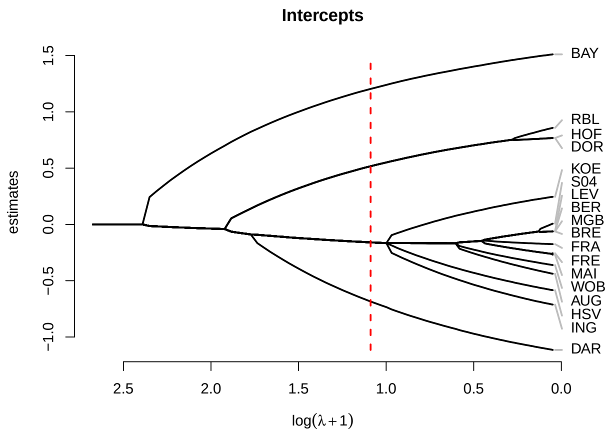
Conside

Estimat

With w

- Sh
- Cl
- Re

Bundesliga 2016/17



Fusion for Match-Team-Specific Covariates

$$P(Y_{i(r,s)} \leqslant k) = \frac{\exp(\delta_r + \theta_k + \gamma_{ir} - \gamma_{is})}{1 + \exp(\delta_r + \theta_k + \gamma_{ir} - \gamma_{is})}$$
$$= \frac{\exp(\delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T \boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T \boldsymbol{\alpha}_s)}{1 + \exp(\delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T \boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T \boldsymbol{\alpha}_s)}$$

- δ_r team-specific home effects of team r
- θ_k category-specific threshold parameters
- β_{r0} team-specific intercepts
- $\boldsymbol{z}_{ir}$ p-dimensional covariate vector that varies over teams and matches
- $\boldsymbol{\alpha}_r$ p-dimensional parameter vector that varies over teams.

Penalty For Fusion and Selection

- Predictor: $\eta_{rsk} = \delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T \boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T \boldsymbol{\alpha}_s$

$$J(\cdot) = J_\delta(\cdot) + J_\alpha(\cdot)$$

combining the penalties

$$J_\delta(\delta_1, \ldots, \delta_m) = \sum_{r<s} |\delta_r - \delta_s| \,,$$

$\Rightarrow$ Fusion of home effects

$$J_\alpha(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m) = \sum_{j=1}^{p} \left(\sum_{r<s} |\alpha_{rj} - \alpha_{sj}| + \sum_{r=1}^{m} |\alpha_{rj}| \right) .$$

$\Rightarrow$ Fusion and selection of covariate effects

Penalty For Fusion and Selection

- Predictor: $\eta_{rsk} = \delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T\boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T\boldsymbol{\alpha}_s$

combini

For $\lambda \rightarrow \infty$:

$\Rightarrow$ Global home effect

$$\delta_1 = \ldots = \delta_{18} = \delta$$

$\Rightarrow$ Elimination of covariate effects

$$\boldsymbol{\alpha}_1 = \ldots = \boldsymbol{\alpha}_{18} = \boldsymbol{0}$$

$\Rightarrow$ One ends up with the basic model

$$P(Y_{i(r,s)} \leqslant k) = \frac{\exp(\delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T\boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T\boldsymbol{\alpha}_s)}{1 + \exp(\delta_r + \theta_k + \beta_{r0} - \beta_{s0} + \boldsymbol{z}_{ir}^T\boldsymbol{\alpha}_r - \boldsymbol{z}_{is}^T\boldsymbol{\alpha}_s)}$$

$$= \frac{\exp(\delta + \theta_k + \beta_{r0} - \beta_{s0})}{1 + \exp(\delta + \theta_k + \beta_{r0} - \beta_{s0})}$$

$\Rightarrow$ F

$\Rightarrow$ Fusion and selection of covariate effects

Selection of Tuning Parameter

- Calculate model for a grid of tuning parameters λ
- Cross-validate model with respect to a certain criterion

- We use the *ranked probability score* (RPS) (Gneiting and Raftery, 2007)
- RPS for ordinal response $y \in \{1, \ldots, K\}$ is

$$RPS(y, \hat{\pi}(k)) = \sum_{k=1}^{K} (\hat{\pi}(k) - \mathbb{1}(y \leqslant k))^2 \quad \text{where } \pi(k) = P(y \leqslant k).$$

→ takes into account the ordinal structure of the response

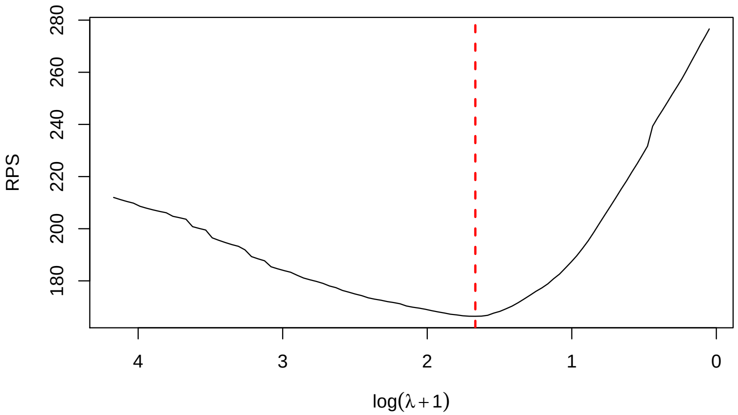
Bundesliga 2015/2016

Position		Team	Goals For	Goals Against	Points
1		Bayern München	80	17	88
2		Borussia Dortmund	82	34	78
3		Bayer 04 Leverkusen	56	40	60
4		Bor. Mönchengladbach	67	50	55
5		FC Schalke 04	51	49	52
6		1. FSV Mainz 05	46	42	50
7		Hertha BSC	42	42	50
8		VfL Wolfsburg	47	49	45
9		1. FC Köln	38	42	43
10		Hamburger SV	40	46	41
11		FC Ingolstadt 04	33	42	40
12		FC Augsburg	42	52	38
13		Werder Bremen	50	65	38
14		SV Darmstadt 98	38	53	38
15		TSG Hoffenheim	39	54	37
16		Eintracht Frankfurt	34	52	36
17		VfB Stuttgart	50	75	33
18		Hannover 96	31	62	25

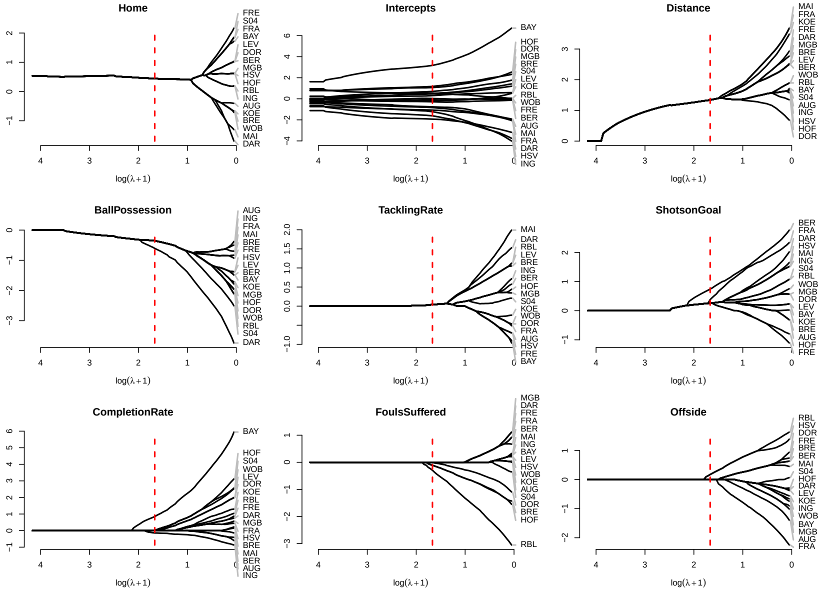
- Distance Total amount of km run
- BallPossession Percentage of ball possession
- TacklingRate Rate of won tacklings
- ShotsonGoal Total number of shots on goal
- CompletionRate Percentage of passes reaching teammates
- FoulsSuffered Number of fouls suffered
- Offside Number of offsides (in attack)

Source: German football magazine kicker (http://www.kicker.de/)

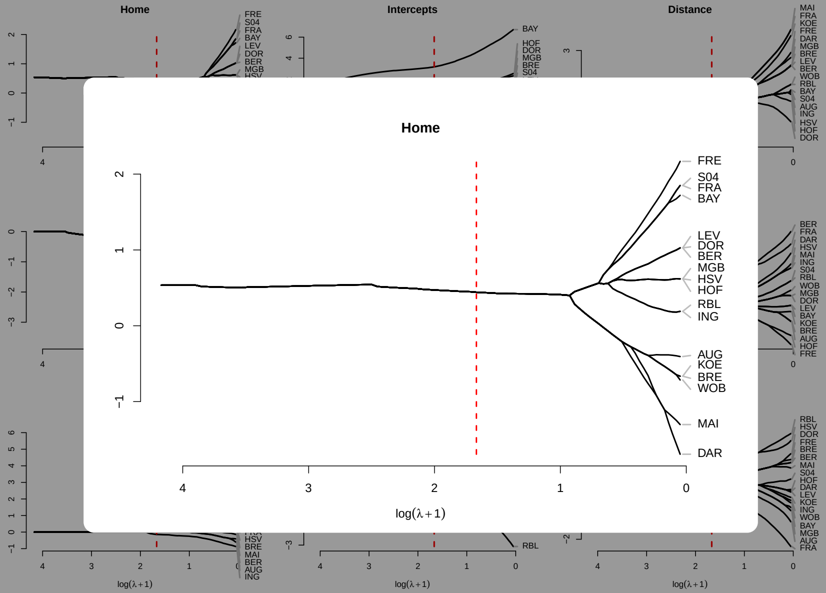
CV error against tuning parameter λ



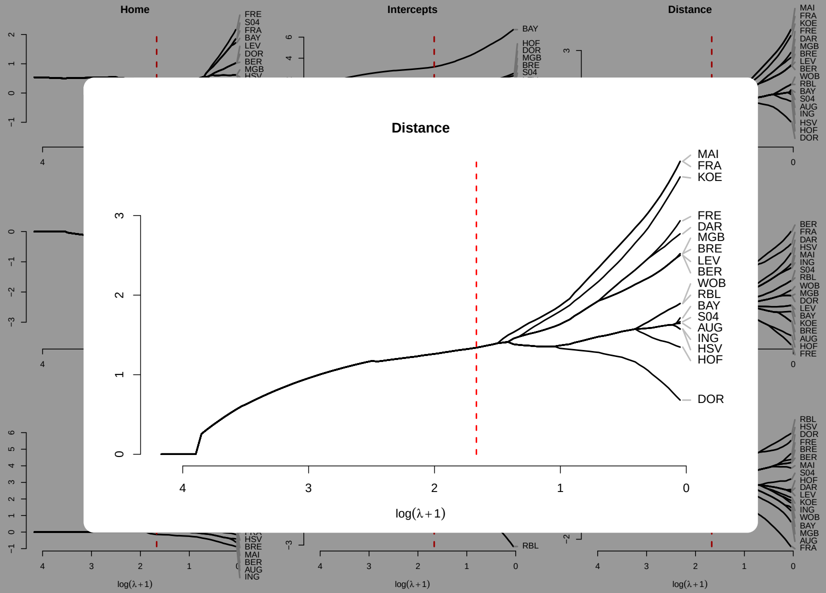
Coefficient Paths



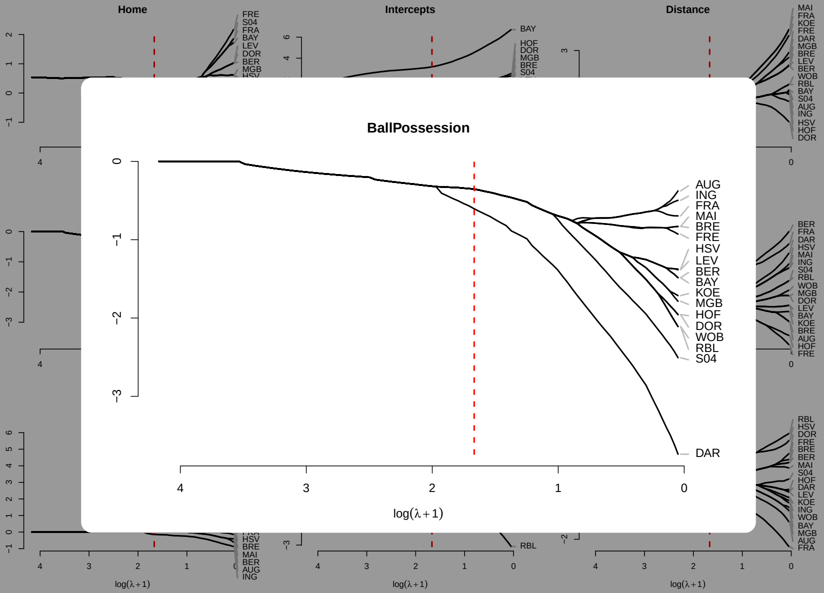
Coefficient Paths



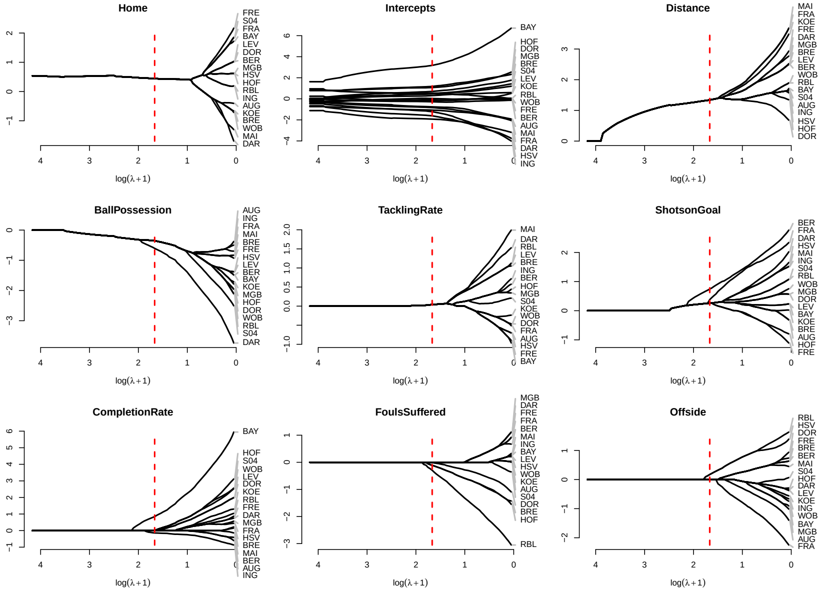
Coefficient Paths



Coefficient Paths



Coefficient Paths



Parameterization

Scale

In order to obtain a common scale covariates are transformed to variance one (over all matches and teams).

Centering

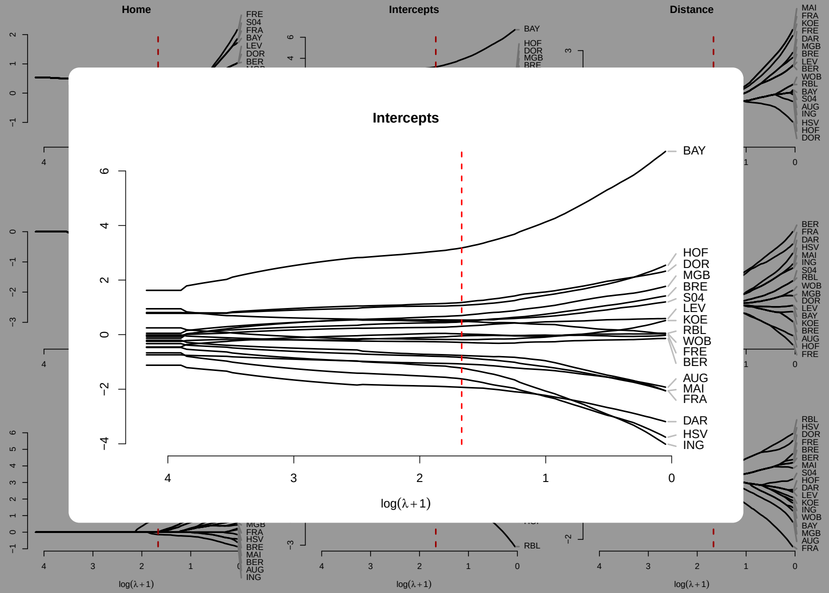
Covariates are centered around $\bar{z}_r$, the team-specific mean

The strength of team r with covariates is

$$\begin{aligned}\tilde{\gamma}_r &= \delta_r + \beta_{r0} + (\boldsymbol{z}_{ir} - \bar{\boldsymbol{z}}_r)^T \boldsymbol{\alpha}_r \\ &= \delta_r + \beta_{r0} - (\bar{\boldsymbol{z}}_r - \bar{\boldsymbol{z}})^T \boldsymbol{\alpha}_r + (\boldsymbol{z}_{ir} - \bar{\boldsymbol{z}})^T \boldsymbol{\alpha}_r\end{aligned}$$

- Effect α_r is the same if one centers around the global mean $\bar{z}$ (over all matches and teams)
- Only the strengths/intercepts are changing

Coefficient Paths



Pure team-Specific Explanatory Variables

Parameterization

$$\gamma_{ir} = \gamma_r = \beta_{r0} + \boldsymbol{z}_r^T \boldsymbol{\tau}.$$

Penalizing only intercepts

$$J(\beta_{10}, \ldots, \beta_{m0}) = \sum_{r<s} |\beta_{r0} - \beta_{s0}|$$

For $\lambda \to \infty$:

- $\beta_{10} = \ldots = \beta_{m0} = 0$, enforces that all variation in strength contained in explanatory variables, no variation left
- reduces number of parameters, identifiable for proper design matrices

Penalizing all parameters

$$J(\beta_{10}, \ldots, \beta_{m0}, \boldsymbol{\tau}) = \sum_{r<s} |\beta_{r0} - \beta_{s0}| + \sum_{r=1}^{m} |\tau_r|.$$

- In addition selects explanatory variables

Pure team-Specific Explanatory Variables

Parameterization

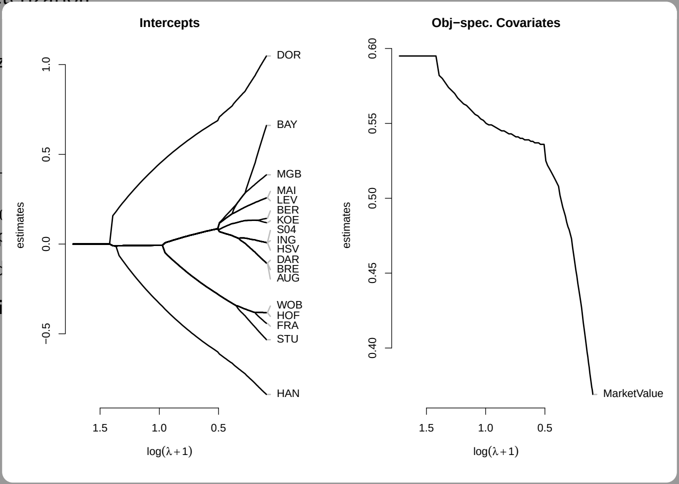
Penaliz

For λ –

- $\beta_{1($
  exp
- rec

Penalizi

- In



R Package BTLLasso

- Provides whole framework to incorporate all types of covariates in paired comparison models
- Specific penalty terms to
  - Select variables
  - Cluster objects
- Main functions:

```
BTLLasso(Y, X = NULL, Z1 = NULL, Z2 = NULL, lambda = NULL,
  control = ctrl.BTLLasso(), trace = TRUE)
```

and

```
cv.BTLLasso(Y, X = NULL, Z1 = NULL, Z2 = NULL, folds = 10,
  lambda = NULL, control = ctrl.BTLLasso(), cores = folds, trace = TRUE,
  trace.cv = TRUE, cv.crit = c("RPS", "Deviance"))
```

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1+\exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir}-\gamma_{is})}{1+\exp(\gamma_{ir}-\gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1 + \exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1 + \exp(\gamma_{ir} - \gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1 + \exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1 + \exp(\gamma_{ir} - \gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1+\exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1+\exp(\gamma_{ir} - \gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1 + \exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1 + \exp(\gamma_{ir} - \gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1+\exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir} - \gamma_{is})}{1+\exp(\gamma_{ir} - \gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^T\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^T\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^T(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^T\boldsymbol{\tau}$	$+\boldsymbol{z}_s^T\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^T\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^T\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^T\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^T\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^T\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^T\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^T\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^T\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Extensions of the Bradley-Terry Model in BTLLasso

$$P(Y_{i(r,s)} = 1 \mid \boldsymbol{x}_i, \boldsymbol{z}_r, \boldsymbol{z}_{ir}, \boldsymbol{z}_{is}) = \frac{\exp(\eta_{i(rs)})}{1+\exp(\eta_{i(rs)})} = \frac{\exp(\gamma_{ir}-\gamma_{is})}{1+\exp(\gamma_{ir}-\gamma_{is})}$$

Covariate type	Effect type	$\gamma_{ir} =$	$\gamma_{is} =$	$\eta_{i(rs)} =$
intercept	object-spec.	β_{r0}	β_{s0}	$\beta_{r0} - \beta_{s0}$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_r$	$+\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_s$	$+\boldsymbol{x}_i^{\mathrm{T}}(\boldsymbol{\beta}_r - \boldsymbol{\beta}_s)$
object-spec. $\boldsymbol{z}_r$	global	$+\boldsymbol{z}_r^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_s^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_r - \boldsymbol{z}_s)^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\tau}$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\tau}$	$+(\boldsymbol{z}_{ir} - \boldsymbol{z}_{is})^{\mathrm{T}}\boldsymbol{\tau}$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$	$+\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$	$+\boldsymbol{z}_{ir}^{\mathrm{T}}\boldsymbol{\alpha}_r$ - $\boldsymbol{z}_{is}^{\mathrm{T}}\boldsymbol{\alpha}_s$
order effect	object-spec.	$+\delta_r$		$+\delta_r$
↳ incl. order effect	global	$+\delta$		$+\delta$

Penalties in BTLLasso

For all model components, specific penalty terms can be applied to keep the model sparse and interpretable

We can apply L_1-penalties for

- clustering of intercept parameters
- clustering and selection of variable effects
- clustering and selection of order effects

Penalties in BTLLasso

For all model components, specific penalty terms can be applied to keep the model sparse and interpretable

We can apply L_1-penalties for

- clustering of intercept parameters
- clustering and selection of variable effects
- clustering and selection of order effects

All penalties can be weighted according to the idea of adaptive lasso proposed by Yuan and Lin (2006)

- weights are calculated as inverse of the absolute values of the penalized term when estimates using ML or small Ridge penalties
- high values (or large differences) in ML estimates are penalized less strong compared to small values (or small differences)

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar

interc

subjec

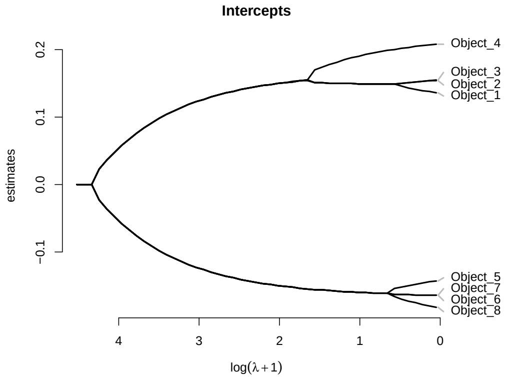
subjec

↳ i

subjec $|\alpha_{rj}|$

order

order



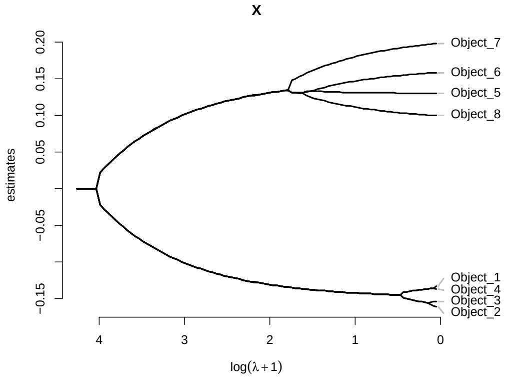
$$P(\cdot) = \sum_{r<s} |\beta_{r0} - \beta_{s0}|$$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar	
interc	
subjec	
subjec	
↳ i	
subjec	$\lvert\alpha_{rj}\rvert$
order	
order	



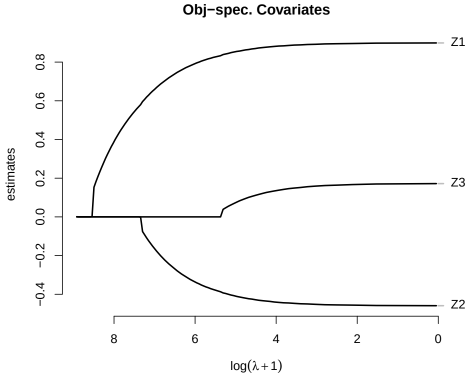
$$P(\cdot) = \sum_{j=1}^{p_x} \sum_{r<s} |\beta_{rj} - \beta_{sj}|$$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar
interc
subjec
subjec
↳ i
subjec $|\alpha_{rj}|$
order
order



$$P(\cdot) = \sum_{j=1}^{p_2} |\tau_j|$$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar

interc

subjec

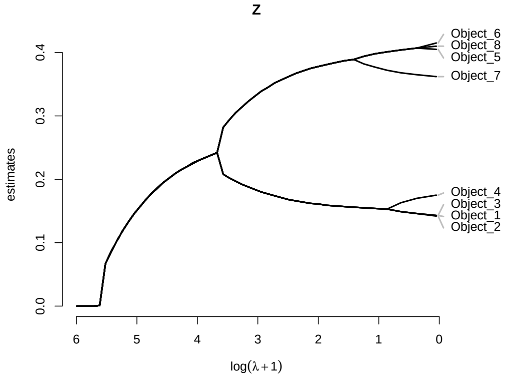
subjec

↳ i

subjec $|\alpha_{rj}|$

order

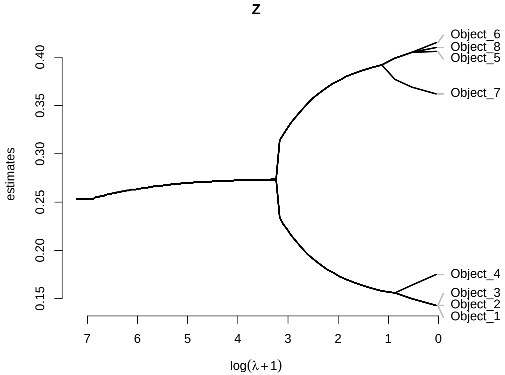
order



$$P(\cdot) = \sum_{j=1}^{p_1} \sum_{r<s} |\alpha_{rj} - \alpha_{sj}|$$

Overview over Penalty Terms in BTLLasso

Covar
interc
subjec
subjec
↳ i
subjec $|\alpha_{rj}|$
order
order



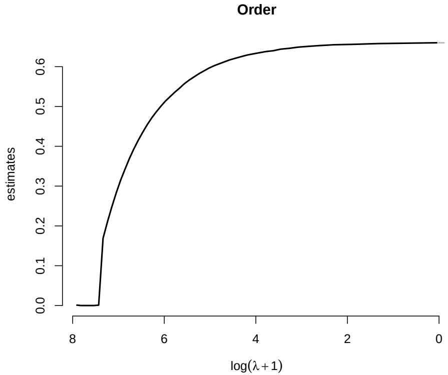
$$P(\cdot) = \sum_{j=1}^{p_1} \sum_{r<s} |\alpha_{rj} - \alpha_{sj}| + \sum_{j=1}^{p_1} \sum_{r=1}^{m} |\alpha_{rj}|$$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar
interc
subjec
subjec
↳ i
subjec $\lvert\alpha_{rj}\rvert$
order
order



$P(\cdot) = |\delta|$

Overview over Penalty Terms in BTLLasso

Covariate type	Effect type	Penalty
intercept	object-spec.	$\sum_{r<s} \lvert\beta_{r0} - \beta_{s0}\rvert$
subject-spec. $\boldsymbol{x}_i$	object-spec.	$\sum_{j=1}^{p_x} \sum_{r<s} \lvert\beta_{rj} - \beta_{sj}\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
↳ incl. object-spec. $\boldsymbol{z}_r$	global	$\sum_{j=1}^{p_2} \lvert\tau_j\rvert$
subject-object-spec. $\boldsymbol{z}_{ir}$	object-spec.	$\sum_{j=1}^{p_1} \sum_{r<s} \lvert\alpha_{rj} - \alpha_{sj}\rvert + \nu_1 \sum_{j=1}^{p_1} \sum_{r=1}^{m} \lvert\alpha_{rj}\rvert$
order effect	global	$\lvert\delta\rvert$
order effect	object-spec.	$\sum_{r<s} \lvert\delta_r - \delta_s\rvert + \nu_2 \sum_{r=1}^{m} \lvert\delta_r\rvert$

Overview over Penalty Terms in BTLLasso

Covar

interc

subje

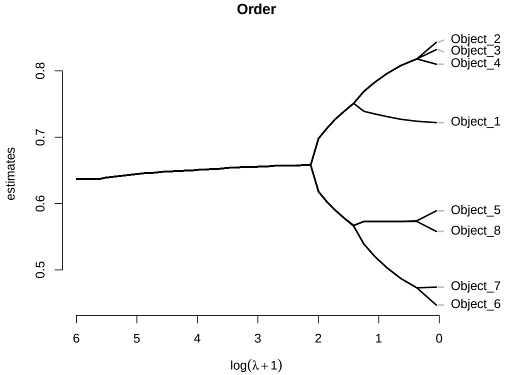
subje

↳ i

subje $|\alpha_{rj}|$

order

order



$$P(\cdot) = \sum_{r<s} |\delta_r - \delta_s|$$

Overview over Penalty Terms in BTLLasso

Covar

interc

subjec

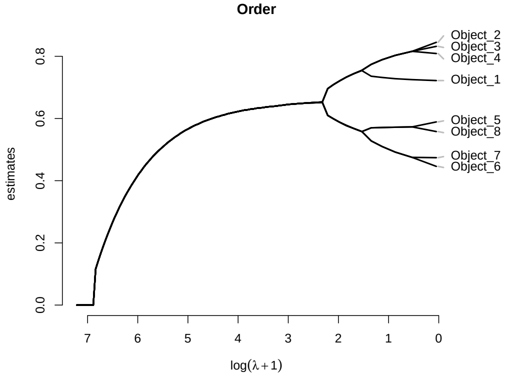
subjec

↳ i

subjec $|\alpha_{rj}|$

order

order



$$P(\cdot) = \sum_{r<s} |\delta_r - \delta_s| + \sum_{r=1}^{m} |\delta_r|$$

Further Functions of BTLLasso

- Methods to
  - `plot()`
  - `coef()`
  - `logLik()`
  - `print()`
  - `predict()`

  for objects created by `cv.BTLLasso()`
- Function

```
boot.BTLLasso(model, B = 500, lambda = NULL, cores = 1, trace = TRUE,
  trace.cv = TRUE, with.cv = TRUE)
```

  for bootstrap intervals of parameter estimates including `plot()` and `print()` methods

Further Functions of BTLLasso

- Methods to
  - plot()

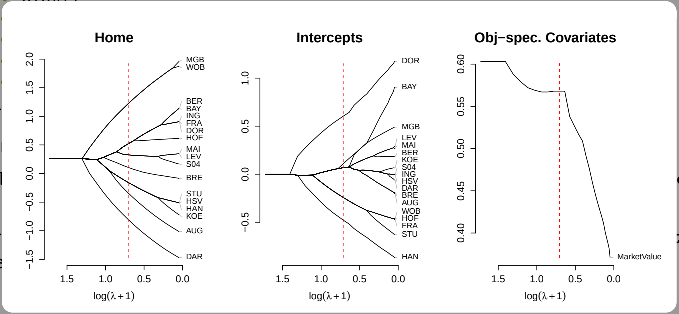
for

- Fu

l e = TRUE,

for ()

me



Further Functions of BTLLasso

- Methods to
  - plot()
  - coef()
  - logLik()
  - print()
  - predict()

  for objects created by cv.BTLLasso()
- Function

```
boot.BTLLasso(model, B = 500, lambda = NULL, cores = 1, trace = TRUE,
  trace.cv = TRUE, with.cv = TRUE)
```

  for bootstrap intervals of parameter estimates including plot() and print() methods

Further Functions of BTLLasso

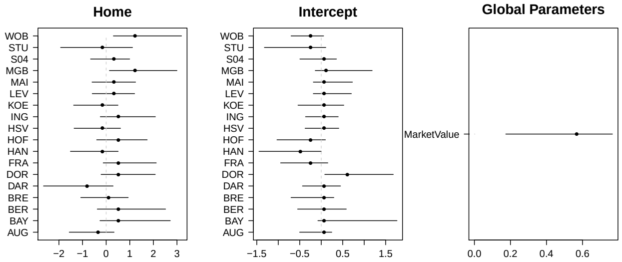
- Methods to
  - plot()

for

- Fu

l e = TRUE,

for ()
me



Further Functions of BTLLasso

- Methods to
  - `plot()`
  - `coef()`
  - `logLik()`
  - `print()`
  - `predict()`

  for objects created by `cv.BTLLasso()`
- Function

```
boot.BTLLasso(model, B = 500, lambda = NULL, cores = 1, trace = TRUE,
  trace.cv = TRUE, with.cv = TRUE)
```

  for bootstrap intervals of parameter estimates including `plot()` and `print()` methods

Implementational Details

- Fitting procedure implemented in C++ to speed-up computation time
  → integrated into R via Rcpp and RcppArmadillo
- Fitting procedure is (penalized) Fisher scoring
- L_1 penalties are approximated by quadratic terms to make them differentiable
  → following Oelker and Tutz (2017)
- Both cv.BTLLasso() and boot.BTLLasso() can be parallelized on several cores to speed-up computation time

References

Bradley, R. A. and M. E. Terry (1952). Rank analysis of incomplete block designs, I: The method of pair comparisons. *Biometrika 39*, 324–345.

Gneiting, T. and A. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102*(477), 359–376.

Oelker, M. and G. Tutz (2017). A uniform framework for the combination of penalties in generalized structured models. *Advances in Data Analysis and Classification, published online (DOI 10.1007/s11634-015-0205-y) 11*, 97– 120.

Schauberger, G. (2017). *BTLLasso: Modelling Heterogeneity in Paired Comparison Data*. R package version 0.1-5.

Schauberger, G., A. Groll, and G. Tutz (2016). Modeling football results in the German Bundesliga using match-specific covariates. Technical Report 197, Department of Statistics, Ludwig-Maximilians-Universität München, Germany.

Schauberger, G. and G. Tutz (2017). Subject-specific modelling of paired comparison data - a lasso-type penalty approach. *Statistical Modelling*, in press.

Tutz, G. and G. Schauberger (2015). Extended ordered paired comparison models with application to football data from german bundesliga. *Advances in Statistical Analysis 99*, 209–227.

Importance of Home Effect

For K=3 (a_r wins, draw, a_s wins) and equal strength, $\gamma_r = \gamma_s$, δ reflects the proportion of odds for winning of (home) team r and winning of team s,

$$\delta = \frac{1}{2} \log \left(\frac{P(Y_{rs} = 1)/(1 - P(Y_{rs} = 1))}{P(Y_{rs} = 3)/(1 - P(Y_{rs} = 3))} \right).$$

The general odds of winning are

$$\frac{P(Y_{rs} = 1)}{1 - P(Y_{rs} = 1)} = e^{\delta} e^{\theta} e^{\gamma_r - \gamma_s}, \frac{P(Y_{rs} = 3)}{1 - P(Y_{rs} = 3)} = e^{-\delta} e^{\theta} e^{\gamma_s - \gamma_r}$$

For K=3 (a_r wins, draw, a_s wins) and equal strength, $\gamma_r = \gamma_s$, δ reflects the proportion of odds for winning of (home) team r and winning of team s,

The ger

Season 2012/2013

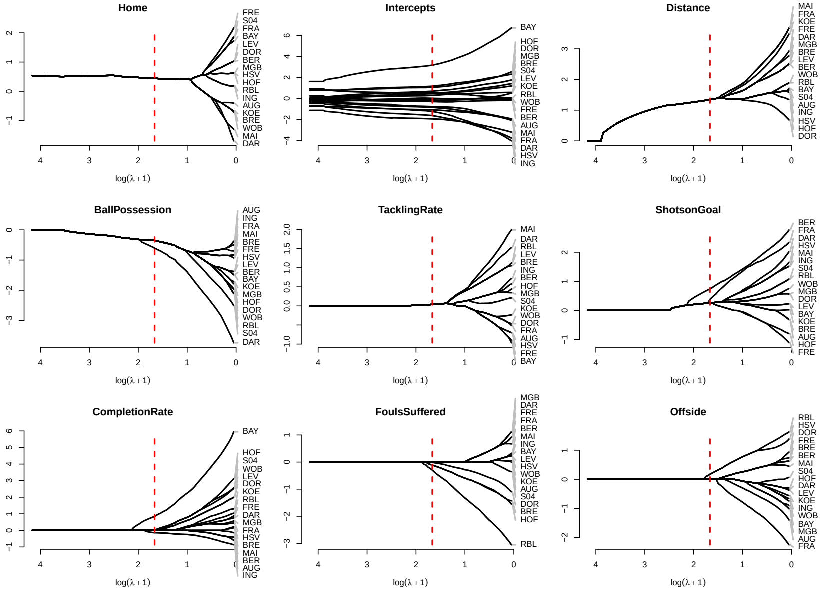
5-point scale: $\hat{\delta} = 0.293$

For two teams with equal abilities, one obtains the probabilities

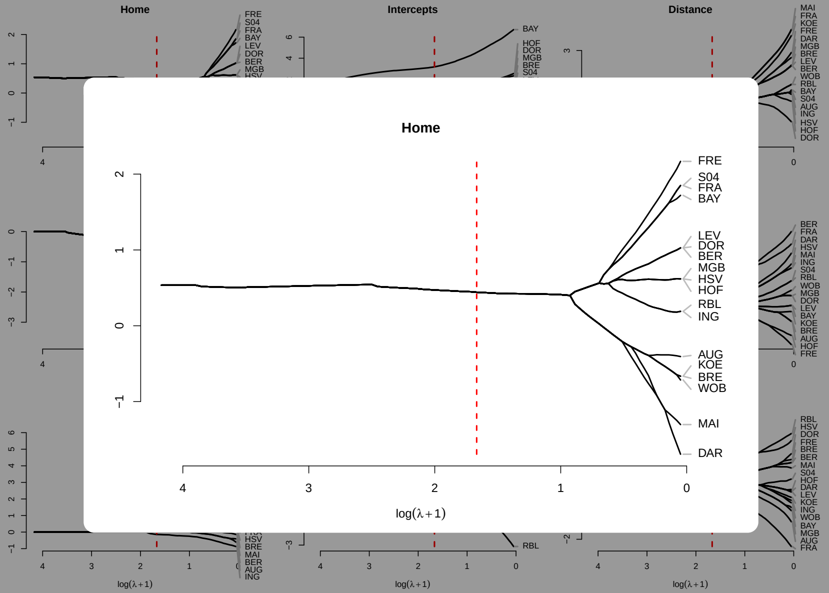
- 0.41 for a victory of the home team
- 0.31 for a draw
- 0.28 for a victory of the away team

42.5% of the matches were won by the home team, 25.5% of the matches ended with a draw and 32% of the matches were won by the away team.

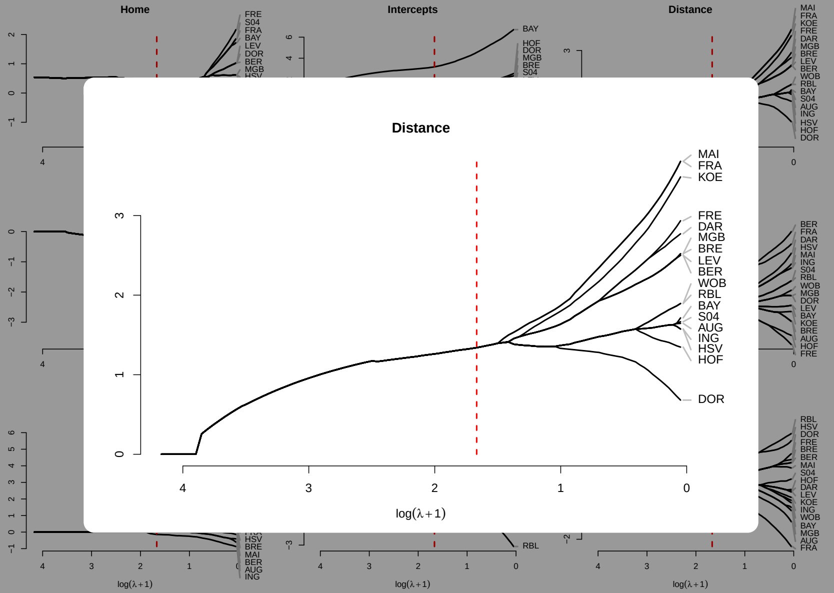
Coefficient Paths



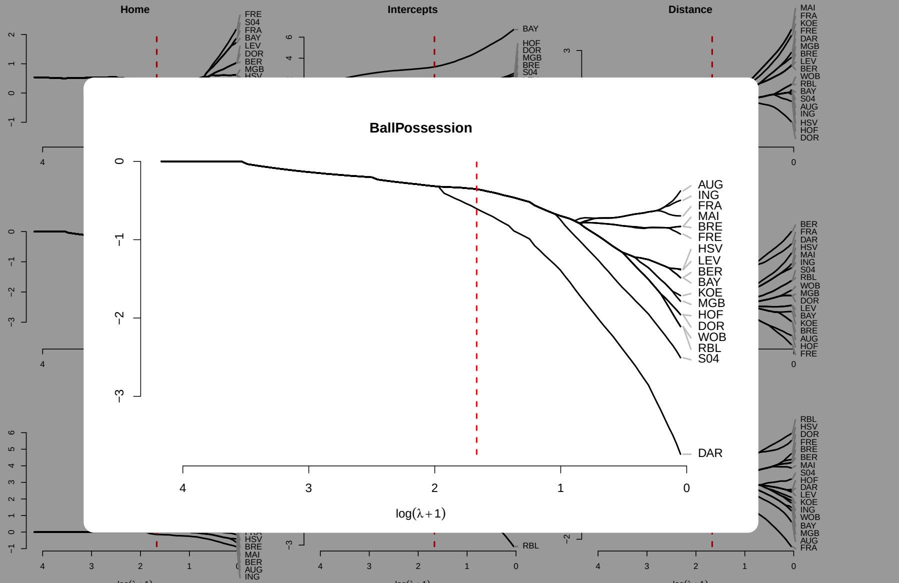
Coefficient Paths



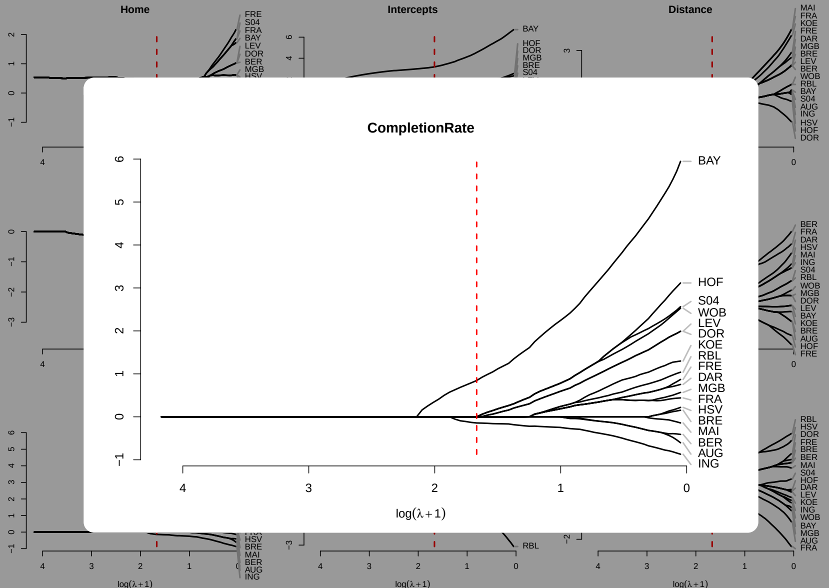
Coefficient Paths



Coefficient Paths



Coefficient Paths



Scale

In order to obtain a common scale covariates are transformed to variance one (over all matches and teams).

Centering

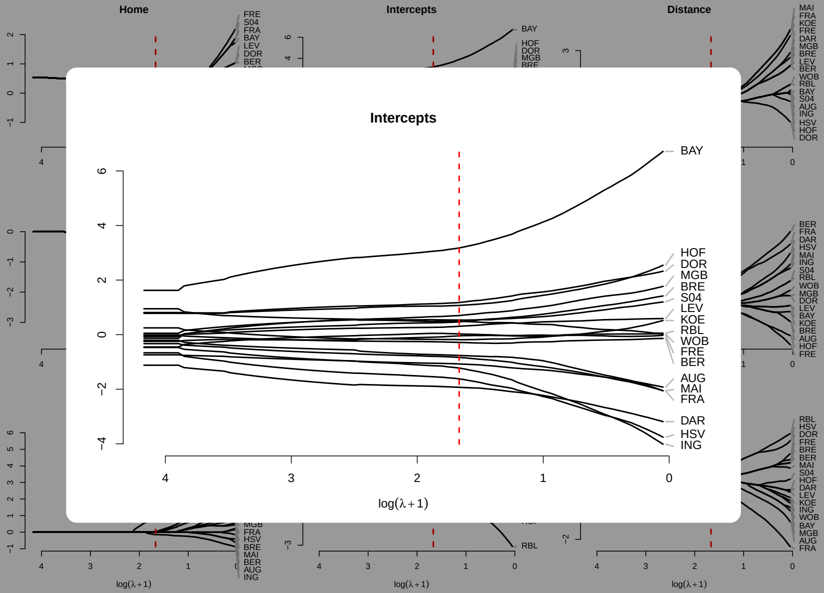
Covariates are centered around $\bar{z}_r$, the team-specific mean

The strength of team r with covariates is

$$\begin{aligned}\tilde{\gamma}_r &= \delta_r + \beta_{r0} + (\boldsymbol{z}_{ir} - \bar{\boldsymbol{z}}_r)^T \boldsymbol{\alpha}_r \\ &= \delta_r + \beta_{r0} - (\bar{\boldsymbol{z}}_r - \bar{\boldsymbol{z}})^T \boldsymbol{\alpha}_r + (\boldsymbol{z}_{ir} - \bar{\boldsymbol{z}})^T \boldsymbol{\alpha}_r\end{aligned}$$

- Effect α_r is the same if one centers around the global mean $\bar{z}$ (over all matches and teams)
- Only the strengths/intercepts are changing

Coefficient Paths



Alternative Approach to Pure team-Specific Explanatory Variables

Turner and Firth (2012) use the parameterization

$$\gamma_r = \beta_{r0} + \boldsymbol{z}_r^T \boldsymbol{\tau},$$

where β_{r0} are iid random effects, $\beta_{r0} \sim N(0, \sigma^2)$

Estimation by quasi likelihood, which can be seen as a penalized estimation with penalty

$$J(\beta_{10}, \ldots, \beta_{m0}) = \sum_{r=1}^{p} \beta_{r0}^2$$

- The main difference is that a ridge type is used instead of a lasso-type penalty (no fusion)
- Effects of explanatory variables are not penalized, therefore no selection of effects
- Random effects model assumes that explanatory variables and random effects are uncorrelated (endogeneity problem)