











HeKA Staff Meeting

Simon Bussy

simon.bussy@gmail.com

Lights: a generalized joint model for high-dimensional multivariate longitudinal data and censored durations

S. Bussy (1,2), V.T. Nguyen (1,2), A. Barbieri (3), S. Zohar (4), A.S. Jannot (4,5)

(1) LOPF, Califrais' Machine Learning Lab, Paris, France (2) LPSM. UMR 8001, CNRS, Sorbonne University, Paris, France (3) INSERM, UMR 1219, Bordeaux Population Health Research Center, Univ. Bordeaux, France (4) INSERM, UMRS 1138, Centre de Recherche des Cordeliers, Paris, France (5) Biomedical Informatics and Public Health Department, EGPH, APHP, Paris, France

I. Introduction

Introduction

...

Eramouso

6.1

Notations

Likelihood

Inference

prox-QNEM

Evaluation

Application Competitors

Simulations

esuits

Officialion

 Deal with the problem of joint modeling of longitudinal data and censored durations ntroduction

Overview Use cases

Use cases Framewor

Model

Submodels

Votations

Informaco

Penalizatio

Fvaluation

plication

ompetitors

Results

onclusion

deferences

Overview

- Deal with the problem of joint modeling of longitudinal data and censored durations
- ► Large number of both longitudinal and time-independent features are available

troduction

Overview Use cases

4 - 1 - 1

Submodels

Votations

Inference

prox-QNE

Evaluation

pplication

Competitors

Results

onclusion

Overview

- Deal with the problem of joint modeling of longitudinal data and censored durations
- Large number of both longitudinal and time-independent features are available
- Flexibility in modeling the dependency between the longitudinal features and the event time with appropriate penalties

Introduction

Overview Use cases

Framework

Vlodel

Notations

ikelihood

Inference

prox-QNE

Evaluation

Application

Competitors

oimulatioi Results

onclusion

Overview

- Deal with the problem of joint modeling of longitudinal data and censored durations
- Large number of both longitudinal and time-independent features are available
- ▶ Flexibility in modeling the dependency between the longitudinal features and the event time with appropriate penalties
- Inference achieved using an efficient and novel Proximal Quasi-Newton Expectation Maximization algorithm

Overview

Predict the risk for an event of interest to occur quickly, taking into account simultaneously a huge number of longitudinal signals in a high-dimensional context ntroduction

0----i---

Use cases

Framev

Mode

Submodels

Notations

Informa

merenee

prox-QNEN

plication

Competitors

mulations

onclusion

- Predict the risk for an event of interest to occur quickly, taking into account simultaneously a huge number of longitudinal signals in a high-dimensional context
- Provides powerful interpretability by automatically pinpointing significant time-dependent and time-independent features

ntroduction

Overview

Use cases

Model

viouei

lotations

Notations

Inference

orox-QNEM

Evaluation

pplication

imulations esults

onclusion

- Predict the risk for an event of interest to occur quickly, taking into account simultaneously a huge number of longitudinal signals in a high-dimensional context
- Provides powerful interpretability by automatically pinpointing significant time-dependent and time-independent features

Real-time decision support

Medical context → event of interest: survival time, re-hospitalization, relapse or disease progression; longitudinal data: biomarkers or vital parameters measurements troduction

Overview

Use cases

Framewo

Submodels

Votations

Inference

prox-QNEI

Evaluation

Application

Competitors Simulations

resurts

Conclusion

- Predict the risk for an event of interest to occur quickly, taking into account simultaneously a huge number of longitudinal signals in a high-dimensional context
- Provides powerful interpretability by automatically pinpointing significant time-dependent and time-independent features

Real-time decision support

- Medical context → event of interest: survival time. re-hospitalization, relapse or disease progression; longitudinal data: biomarkers or vital parameters measurements
- ightharpoonup Customer's satisfaction monitoring context \rightarrow event of interest: time when a client churns; longitudinal data: the client's activity recorded from account opening throughout the duration of the business relationship

Hee cases

$$T = T^{\star} \wedge C$$
 and $\Delta = \mathbb{1}_{\{T^{\star} \leq C\}}$

ntroduction

Overview

Use cases Framework

Model

Submodels Notations

Likelihood

Inference

prox-QNEM Evaluation

> pplication Competitors

Simulations Results

`onelucion

$$T = T^{\star} \wedge C$$
 and $\Delta = \mathbb{1}_{\{T^{\star} \leq C\}}$

▶ Time-independent features $X \in \mathbb{R}^p$ with $p \gg n$

troduction

Overview

Framework

/lodel

oubmodels Notations

1.6

Penalization prox-QNEM

Evaluation

Application Competitors

Simulations Results

Conclusion

$$\mathcal{T} = \mathcal{T}^\star \wedge \mathcal{C} \quad \text{and} \quad \Delta = \mathbb{1}_{\{\mathcal{T}^\star \leq \mathcal{C}\}}$$

- ▶ Time-independent features $X \in \mathbb{R}^p$ with $p \gg n$
- ▶ L longitudinal outcomes such that $L \gg n$ and

$$Y(t) = \left(Y^1(t), \dots, Y^L(t)\right)^{ op} \in \mathbb{R}^L$$

troduction

Overview

Framework

1odel

Submodels

Notations

. .

merence

prox-QNEM

Application
Competitors

Simulations Results

nclusion

$$\mathcal{T} = \mathcal{T}^\star \wedge \mathcal{C} \quad \text{and} \quad \Delta = \mathbb{1}_{\{\mathcal{T}^\star \leq \mathcal{C}\}}$$

- ▶ Time-independent features $X \in \mathbb{R}^p$ with $p \gg n$
- ▶ L longitudinal outcomes such that $L \gg n$ and

$$Y(t) = \left(Y^1(t), \dots, Y^L(t)
ight)^ op \in \mathbb{R}^L$$

Heterogeneity of the population: latent subgroups

$$\textit{G} \in \{0, \ldots, \textit{K}-1\}$$

troduction

Use cases

Framework

lodel

Notations

Likelihood

nference

prox-QNEN

Evaluation

Competitors Simulations

Results

onclusion

$$T = T^* \wedge C$$
 and $\Delta = \mathbb{1}_{\{T^* \leq C\}}$

- ▶ Time-independent features $X \in \mathbb{R}^p$ with $p \gg n$
- ▶ L longitudinal outcomes such that $L \gg n$ and

$$Y(t) = \left(Y^1(t), \dots, Y^L(t)
ight)^ op \in \mathbb{R}^L$$

► Heterogeneity of the population: latent subgroups

$$\textit{G} \in \{0, \ldots, \textit{K}-1\}$$

► Softmax link function for the latent class membership probability given time-independent features

$$\pi_{\xi_k}(x) = \mathbb{P}[G = k | X = x] = \frac{e^{x^\top \xi_k}}{\sum_{k=0}^{K-1} e^{x^\top \xi_k}}$$

roduction

Use cases

Framework

VIOGEI Submode

Notations

. .

prox-QNEM

Application Competitors

resuits

II. Model

Introduction

Overview

Framewor

Model

Notations

Inference

Penalization prox-QNEM

Application

Simulations

an alvaian

Group-specific marker trajectories

▶ $h_l(\mathbb{E}[Y^l(t)|b^l,G=k]) = m_k^l(t) = u^l(t)^\top \beta_k^l + v^l(t)^\top b^l$ with fixed effect parameters $\beta_k^l \in \mathbb{R}^{q_l}$ and subject-and-longitudinal outcome specific random effects $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0,D_l)$

ntroductio

Use cases

Use cases

Mode

Model

Submodels

Notations

Likelinood

Penalization

prox-QIVEIV Evaluation

pplication

Competitors

Results

onclusion

Group-specific marker trajectories

- ▶ $h_l(\mathbb{E}[Y^l(t)|b^l,G=k]) = m_k^l(t) = u^l(t)^\top \beta_k^l + v^l(t)^\top b^l$ with fixed effect parameters $\beta_k^l \in \mathbb{R}^{q_l}$ and subject-and-longitudinal outcome specific random effects $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0,D_l)$
- $ightharpoonup \operatorname{Cov}[b^I,b^{I'}] = D_{II'}$ and

$$D = \begin{bmatrix} D_{11} & \cdots & D_{1L} \\ \vdots & \ddots & \vdots \\ D_{1L}^\top & \cdots & D_{LL} \end{bmatrix}$$

the global variance-covariance matrix

ntroduction

Use cases

Framewor

Vlodel

Submodels

Notations

nference

prox-QNEN

Lvaluation

Competitors

Results

onclusion

Group-specific marker trajectories

- ▶ $h_l(\mathbb{E}[Y^l(t)|b^l,G=k]) = m_k^l(t) = u^l(t)^\top \beta_k^l + v^l(t)^\top b^l$ with fixed effect parameters $\beta_k^l \in \mathbb{R}^{q_l}$ and subject-and-longitudinal outcome specific random effects $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0,D_l)$
- $\qquad \qquad \mathsf{Cov}[b^{l},b^{l'}] = D_{ll'}$

Group-specific risk of event

$$\lambda(t|\mathcal{Y}(t), G = k) = \lambda_0(t) \exp \left\{ \underbrace{\sum_{l=1}^{L} \sum_{a=1}^{\mathcal{A}} \gamma_{k,a}^l \Psi_a^l(t)}_{\gamma_k^\top \Psi(t)} \right\}$$

with $\mathcal{Y}(t) = \{Y(u), 0 \leq u < t\}$ and for each l-th longitudinal outcome, we consider $\mathcal{A} \in \mathbb{N}_+$ known functionals Ψ_a^l extracted from $\mathcal{Y}^l(t)$ through a given representation mapping

ntrodu

Use cases

Model

Submodels

Votations

nference

prox-QNEM

Application

Simulation Results

onclusion

Group-specific marker trajectories

- ▶ $h_l(\mathbb{E}[Y^l(t)|b^l,G=k]) = m_k^l(t) = u^l(t)^\top \beta_k^l + v^l(t)^\top b^l$ with fixed effect parameters $\beta_k^l \in \mathbb{R}^{q_l}$ and subject-and-longitudinal outcome specific random effects $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0,D_l)$
- $\triangleright \mathsf{Cov}[b^{I},b^{I^{\prime}}] = D_{II^{\prime}}$

Group-specific risk of event

$$\lambda(t|\mathcal{Y}(t), G = k) = \lambda_0(t) \exp\Big\{\underbrace{\sum_{l=1}^{L} \sum_{a=1}^{\mathcal{A}} \gamma_{k,a}^l \Psi_a^l(t)}_{\gamma_k^\top \Psi(t)} \Big\}$$

with $\mathcal{Y}(t) = \{Y(u), 0 \leq u < t\}$ and for each l-th longitudinal outcome, we consider $\mathcal{A} \in \mathbb{N}_+$ known functionals Ψ_a^l extracted from $\mathcal{Y}^l(t)$ through a given representation mapping

• $\Psi(t) \in \mathbb{R}^{LA}$: high-dim representation features vector, highly flexible and independent of any modeling assumption!

ntroduc

Use cases

Model

Submodels

Votations

nference

prox-QNEM Evaluation

Application

Simulations

Conclusion

Generalization of SREMs and JLCMs

► JLCMs [8] : homogeneous latent subgroups sharing the same marker trajectories and the same prognostic

Lights 6/20

Introdu

Use case

NA - J. I

Submodels

lotations

Inforence

prox-QNEN

Evaluation

pplication

ompetitors mulations

esults

onclusion

- ▶ JLCMs [8] : homogeneous latent subgroups sharing the same marker trajectories and the same prognostic
- ➤ SREMs [9]: characteristics of the longitudinal processes (e.g. functions of the random effects) are included as features in the survival model

ntroduct

Use cases

Wodel

Submodels

Votations

. .

....

prox-QNEM Evaluation

pplication

ompetitors imulations

Results

onclusion

- ▶ JLCMs [8] : homogeneous latent subgroups sharing the same marker trajectories and the same prognostic
- ► SREMs [9] : characteristics of the longitudinal processes (e.g. functions of the random effects) are included as features in the survival model
- Usual functionals would be

Description	$\tilde{\Psi}_a^I(t,\beta_k^I,b^I)$	Reference
Linear predictor	$m_k^l(t)$	Chi and Ibrahim [2]
Random effects	b^{I}	Hatfield et al. [3]
Time-dependent slope	$\frac{\mathrm{d}}{\mathrm{d}t}m_k^I(t)$	Rizopoulos and Ghosh [6]
Cumulative effect	$\int_0^t m_k^I(s) \mathrm{d} s$	Andrinopoulou et al. [1]

- ► JLCMs [8] : homogeneous latent subgroups sharing the same marker trajectories and the same prognostic
- ➤ SREMs [9]: characteristics of the longitudinal processes (e.g. functions of the random effects) are included as features in the survival model
- Usual functionals would be

Description	$\Psi_a^I(t,\beta_k^I,b^I)$	Reference
Linear predictor	$m_k^l(t)$	Chi and Ibrahim [2]
Random effects	b^{I}	Hatfield et al. [3]
Time-dependent slope	$\frac{\mathrm{d}}{\mathrm{d}t}m_k^I(t)$	Rizopoulos and Ghosh [6]
Cumulative effect	$\int_0^t m_k^I(s) \mathrm{d} s$	Andrinopoulou et al. [1]

Depend on β_k (then on k) which leads to complicated updates, and on b which leads to untracktable integrals that requires approximation methods : computationally intensive and not scalable

ntroduction

Overview

Use cases

Model

Submodels Notations

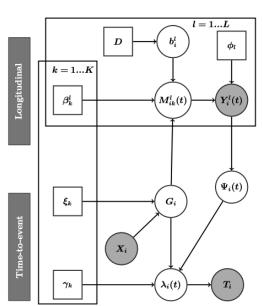
foronco

Penalization prox-QNEM Evaluation

Competitors
Simulations

onclusion

Graphical model



Introdu

Use case

1odel

Submodels

Notations

Inference

prox-QNEM

Applicat

Simulations

nclusion

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n) \right\} \text{ with }$$

$$y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l} \text{ and } y_{ij}^l = Y_i^l(t_{ij}^l)$$

ntroduction

Use case:

Framewo

Submodo

Notations

Likelihood

Inference

prox-QNEI

Application Competitors

Simulations Results

onclusion

deferences

 $\mathcal{D}_n = \{(x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n)\}$ with $y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^ op \in \mathbb{R}^{n_i^l}$ and $y_{ij}^l = Y_i^l(t_{ij}^l)$

 $\mathbf{v}_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$

Notations

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n) \right\} \text{ with }$$

$$y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l} \text{ and } y_{ij}^l = Y_i^l(t_{ij}^l)$$

$$y_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$$

$$lacksquare$$
 $b_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r$ with $r = \sum_{l=1}^L r_l$

troduction

Use cases

Framewor

Model

Submode

Notations

Inforence

prox-QNEN

Evaluation

Competitors

Simulations Results

onclusion

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n) \right\} \text{ with }$$

$$y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l} \text{ and } y_{ij}^l = Y_i^l(t_{ij}^l)$$

$$ightharpoonup y_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$$

$$lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$$

Design matrices

$$U_i = \begin{bmatrix} U_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times q} \text{ and } V_i = \begin{bmatrix} V_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times r}$$

with $q = \sum_{l=1}^{L} q_l$ and where for all l = 1, ..., L, one writes

$$\left\{ \begin{array}{ll} U_{il} &= \left(u_i'(t_{i1}')^\top \cdots u_i'(t_{in_i'}^l)^\top\right)^\top \in \mathbb{R}^{n_i' \times q_l}, \\ V_{il} &= \left(v_i'(t_{i1}')^\top \cdots v_i'(t_{in_i'}^l)^\top\right)^\top \in \mathbb{R}^{n_i' \times r_l}. \end{array} \right.$$

ntroduction Overview

Use cases

Model

Notations

nference

Penalization prox-QNEM Evaluation

Application
Competitors
Simulations

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n) \right\} \text{ with }$$

$$y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l} \text{ and } y_{ij}^l = Y_i^l(t_{ij}^l)$$

$$\mathbf{y}_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$$

$$lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$$

Design matrices

$$U_i = \begin{bmatrix} U_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times q} \text{ and } V_i = \begin{bmatrix} V_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times r}$$

with $q = \sum_{l=1}^{L} q_l$ and where for all l = 1, ..., L, one writes

$$\begin{cases} U_{il} &= \left(u_i^l(t_{i1}^l)^\top \cdots u_i^l(t_{in_i^l}^l)^\top\right)^\top \in \mathbb{R}^{n_i^l \times q_l}, \\ V_{il} &= \left(v_i^l(t_{i1}^l)^\top \cdots v_i^l(t_{in_i^l}^l)^\top\right)^\top \in \mathbb{R}^{n_i^l \times r_l}. \end{cases}$$

troduction Overview

Use cases Framework

Submodel

Notations

nference

Penalization prox-QNEM Evaluation

Application Competitors Simulations

onclusion

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n) \right\} \text{ with }$$

$$y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l} \text{ and } y_{ij}^l = Y_i^l(t_{ij}^l)$$

$$y_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$$

$$lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$$

Design matrices

$$U_i = \begin{bmatrix} U_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times q} \text{ and } V_i = \begin{bmatrix} V_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_{iL} \end{bmatrix} \in \mathbb{R}^{n_i \times r}$$

with $q = \sum_{l=1}^{L} q_l$ and where for all l = 1, ..., L, one writes

$$\left\{ \begin{array}{ll} U_{il} &= \left(u_i^l(t_{i1}^l)^\top \cdots u_i^l(t_{in_i^l}^l)^\top\right)^\top \in \mathbb{R}^{n_i^l \times q_l}, \\ V_{il} &= \left(v_i^l(t_{i1}^l)^\top \cdots v_i^l(t_{in_i^l}^l)^\top\right)^\top \in \mathbb{R}^{n_i^l \times r_l}. \end{array} \right.$$

$$M_{ik} = U_i \beta_k + V_i b_i \in \mathbb{R}^{n_i}$$

troduction verview

Overview Use cases

Model

Notations

nforonco

Penalization prox-QNEM Evaluation

Application Competitors Simulations

onclusion

$$\blacktriangleright \ \theta = \left(\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \mathsf{vech}(D), \lambda_0^\top, \gamma_0^\top \cdots \gamma_{K-1}^\top\right) \in \mathbb{R}^\vartheta$$

ntroduction

Use cases

Framewo

Model

Submodels

Likelihood

Informaco

Penalization

prox-QNEN

Application Competitors

Simulations

$$\blacktriangleright \ \theta = \left(\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \mathsf{vech}(D), \lambda_0^\top, \gamma_0^\top \cdots \gamma_{K-1}^\top\right) \in \mathbb{R}^\vartheta$$

$$f(y_i|b_i, G_i = k) = \exp\left\{ (y_i \odot \Phi_i)^\top M_{ik} - c_\phi(M_{ik}) + d_\phi(y_i) \right\} \text{ with }$$

$$\Phi_i = (\phi_1^{-1} \mathbf{1}_{n_i^1}^\top \cdots \phi_L^{-1} \mathbf{1}_{n_i^L}^\top)^\top \in \mathbb{R}^{n_i}$$

ntroduction

Jverview

Use cases Framework

Model

Submodel

Votations

Likelihood

Inference

Penalization prox-QNEM

prox-QNEM Evaluation

Competitors

Simulations Results

onclusion

Likelihood

$$\blacktriangleright \ \theta = \left(\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \mathsf{vech}(\mathit{D}), \lambda_0^\top, \gamma_0^\top \cdots \gamma_{K-1}^\top\right) \in \mathbb{R}^\vartheta$$

$$f(y_i|b_i, G_i = k) = \exp\left\{ (y_i \odot \Phi_i)^\top M_{ik} - c_\phi(M_{ik}) + d_\phi(y_i) \right\} \text{ with }$$

$$\Phi_i = (\phi_1^{-1} \mathbf{1}_{n_i^1}^\top \cdots \phi_L^{-1} \mathbf{1}_{n_i^L}^\top)^\top \in \mathbb{R}^{n_i}$$

Survival part:

$$f(t_i, \delta_i | G_i = k; \theta) = \left[\lambda(t_i | \mathcal{Y}(t_i), G_i = k)\right]^{\delta_i} \times \exp\left\{-\int_0^{t_i} \lambda(s | \mathcal{Y}(s), G_i = k) ds\right\}$$

ntroduction

Use cases

Model

Submodels

Likelihood

nference

Penalization prox-QNEM

Application

Results

- $\bullet = \left(\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \mathsf{vech}(D), \lambda_0^\top, \gamma_0^\top \cdots \gamma_{K-1}^\top\right) \in \mathbb{R}^{\vartheta}$
- $f(y_i|b_i,G_i=k) = \exp\left\{ (y_i \odot \Phi_i)^\top M_{ik} c_\phi(M_{ik}) + d_\phi(y_i) \right\} \text{ with }$ $\Phi_i = (\phi_1^{-1} \mathbf{1}_{n_1^1}^\top \cdots \phi_l^{-1} \hat{\mathbf{1}}_{n_l^l}^\top)^\top \in \mathbb{R}^{n_i}$
- Survival part:

$$egin{aligned} f(t_i, \delta_i | G_i = k; heta) &= \left[\lambda(t_i | \mathcal{Y}(t_i), G_i = k)
ight]^{\delta_i} \\ & imes \exp \left\{ - \int_0^{t_i} \lambda(s | \mathcal{Y}(s), G_i = k) \mathrm{d}s
ight\} \end{aligned}$$

Then, the log-likelihood (JLCMs type) writes

$$\ell_n(\theta) = n^{-1} \sum_{i=1}^n \log \sum_{k=0}^{K-1} \pi_{\xi_k}(x_i) f_{\theta}(t_i, \delta_i | G_i = k) f_{\theta}(y_i | G_i = k) db_i$$

Likelihood

III. Inference

Introduction

Use case:

Model

Notations

Likelihood

Inference

prox-QNEN

Evaluation

Competitors

Simulations

Results

onclusion

deferences

Penalized objective

$$\ell_n^{\text{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}I_1,\tilde{\eta}}$$

with the elasticnet penalty

$$\|z\|_{\mathsf{en},\eta} = (1-\eta)\|z\|_1 + rac{\eta}{2}\|z\|_2^2$$

and the sparse group lasso penalty

$$||z||_{\operatorname{sg} l_1, \tilde{\eta}} = (1 - \tilde{\eta}) ||z||_1 + \tilde{\eta} \sum_{l=1}^{L} ||z^l||_2$$

troduction

Use cases

Submodels

Votations

Inforence

Penalization prox-QNEM

Evaluation

Competitors
Simulations

onclusion

Penalized objective

$$\ell_n^{\text{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}I_1,\tilde{\eta}}$$

with the elasticnet penalty

$$||z||_{\mathsf{en},\eta} = (1-\eta)||z||_1 + \frac{\eta}{2}||z||_2^2$$

and the sparse group lasso penalty

$$||z||_{\operatorname{sg} I_1, \tilde{\eta}} = (1 - \tilde{\eta})||z||_1 + \tilde{\eta} \sum_{l=1}^{L} ||z^l||_2$$

Resulting optimization problem

$$\hat{ heta} \in \operatorname{argmin}_{ heta \in \mathbb{R}^{artheta}} \ell_n^{\mathsf{pen}}(heta)$$

troduction

Overview Use cases

> ioaei Submodels

oubmodels Votations

Inference

Penalization prox-QNEM

Application
Competitors

Kesults

 $\blacktriangleright \ \ell_n^{\mathsf{comp}}(\theta) = \ell_n^{\mathsf{comp}}(\theta; \mathcal{D}_n, \boldsymbol{b}, \boldsymbol{G})$

ntroduction

Use cases

Trainewor

Submodels Notations

Likelihood

Penalization

prox-QNEM Evaluation

Application
Competitors
Simulations

nelucion

 $\blacktriangleright \ \ell_n^{\text{comp}}(\theta) = \ell_n^{\text{comp}}(\theta; \mathcal{D}_n, \boldsymbol{b}, \boldsymbol{G})$

Monte Carlo E-step

ntroduction

Ouomiou

Use cases Framework

Model

Submodels

Notations

. .

Penalization

prox-QNEM

Evaluation

Competitors

Simulations

nclusion

 $\blacktriangleright \ \ell_n^{\mathsf{comp}}(\theta) = \ell_n^{\mathsf{comp}}(\theta; \mathcal{D}_n, \boldsymbol{b}, \boldsymbol{G})$

Monte Carlo E-step

- ▶ Requires to compute expectations of the form

$$\mathbb{E}_{\theta^{(w)}}[g(b_i, G_i)|t_i, \delta_i, y_i] = \sum_{k=0}^{K-1} \pi_{ik}^{\theta^{(w)}} \int_{\mathbb{R}^r} g(b_i, G_i) f_{\theta^{(w)}}(b_i|t_i, \delta_i, y_i) db_i$$

for different functions g, where we denote

$$\pi_{ik}^{\theta^{(w)}} = \mathbb{P}_{\theta^{(w)}}[G_i = k | t_i, \delta_i, y_i]$$

troduction

Use cases

odel

Submodels Notations

Likelihood

Penalization

prox-QNEM Evaluation

Application
Competitors

Simulations Results

onclusion

Proximal Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\boldsymbol{\vartheta}}} \ \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en}, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sg}l_1, \tilde{\boldsymbol{\eta}}}$$

Lights 12/20

prox-QNEM

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\boldsymbol{\vartheta}}} \ \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en}, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sg}l_1, \tilde{\boldsymbol{\eta}}}$$

ntroduction

Overview

Frameworl

Mode

Submodels

Notations Likelihood

nference

prox-QNEM

\ --!:--+:

Competitors Simulations

· · · ·

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\boldsymbol{\vartheta}}} \ \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_{k}\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\boldsymbol{\gamma}_{k}\|_{\mathsf{sg}l_{1},\tilde{\eta}}$$

$$P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$$

ntroduction

Overview

Frameworl

Framewor

Submode

Notations

Inference

prox-QNEM

Application

Simulations

onclusion

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\boldsymbol{\vartheta}}} \ \mathcal{Q}_{\boldsymbol{\eta}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en}, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sg}l_1, \boldsymbol{\tilde{\eta}}}$$

$$P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$$

$$\blacktriangleright \ \xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\operatorname{en},\eta}$$

ntroduction

Use cases

Frameworl

Mode

Submodels

Notations

Inference

prox-QNEM

Evaluation

Competitors

Simulations

onclusion

Proximal Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{\vartheta}}} \ \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_{k}\|_{\mathsf{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_{k}\|_{\mathsf{sg}l_{1}, \, \tilde{\boldsymbol{\eta}}}$$

$$P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$$

$$\qquad \qquad \xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\operatorname{en},\eta}$$

▶ L-BFGS-B to solve the problem

$$\begin{split} & \text{minimize} & \quad P_{n,k}^{(w)}(\xi_k^+ - \xi_k^-) + \zeta_{1,k} \Big((1-\eta) \sum_{j=1}^p (\xi_{k,j}^+ + \xi_{k,j}^-) + \frac{\eta}{2} \|\xi_k^+ - \xi_k^-\|_2^2 \Big) \\ & \text{subject to} & \quad \xi_{k,j}^+ \geq 0 \text{ and } \xi_{k,j}^- \geq 0 \text{ for } j = 1, \dots, p \end{split}$$

ntroduction

Hen caree

Framework

Submodels

Notations

Likelihood

Penalization

prox-QNEM Evaluation

Competitors

Results

Lonclusion

Proximal Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\vartheta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\mathsf{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\mathsf{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

$$P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$$

▶ L-BFGS-B to solve the problem

minimize
$$P_{n,k}^{(w)}(\xi_k^+ - \xi_k^-) + \zeta_{1,k} \Big((1-\eta) \sum_{j=1}^p (\xi_{k,j}^+ + \xi_{k,j}^-) + \frac{\eta}{2} \|\xi_k^+ - \xi_k^-\|_2^2 \Big)$$
 subject to
$$\xi_{k,j}^+ \geq 0 \text{ and } \xi_{k,j}^- \geq 0 \text{ for } j = 1, \dots, p$$

► Closed-form update for $\beta_k^{(w+1)}$

ntroduction

Use cases

Frameworl

Submodels

Votations

Penalization prox-QNEM

valuation

Competitors

Results

Conclusion

Proximal Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\vartheta}} \ \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en},\, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sg}l_1,\, \boldsymbol{\tilde{\eta}}}$$

$$P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$$

- ▶ L-BFGS-B to solve the problem

$$\begin{array}{ll} \text{minimize} & P_{n,k}^{(w)}(\xi_k^+ - \xi_k^-) + \zeta_{1,k} \Big((1-\eta) \sum_{j=1}^p (\xi_{k,j}^+ + \xi_{k,j}^-) + \frac{\eta}{2} \|\xi_k^+ - \xi_k^-\|_2^2 \Big) \\ \\ \text{subject to} & \xi_{k,j}^+ \geq 0 \text{ and } \xi_{k,j}^- \geq 0 \text{ for } j = 1, \dots, p \end{array}$$

- ▶ Closed-form update for $\beta_k^{(w+1)}$
- Proximal gradient (ISTA) for the $\gamma_k^{(w+1)}$ update, based on Lemma 1 that states $\operatorname{prox}_{\operatorname{sg} l_1, \tilde{\eta}, \zeta} = \operatorname{prox}_{\zeta \tilde{\eta}} \sum_{l=1}^L \|\cdot\|_{2,l} \circ \operatorname{prox}_{\zeta (1-\tilde{\eta})} \|\cdot\|_1$

Introduction

Overview

Framework

Submodels

Votations

Inference

prox-QNEM

Application Competitors

Results

- $\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{\vartheta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\mathsf{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\mathsf{sg} \boldsymbol{h}_1, \, \tilde{\boldsymbol{\eta}}}$
- $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$
- L-BFGS-B to solve the problem

$$\begin{array}{ll} \text{minimize} & P_{n,k}^{(w)}(\xi_k^+ - \xi_k^-) + \zeta_{1,k} \Big((1-\eta) \sum_{j=1}^p (\xi_{k,j}^+ + \xi_{k,j}^-) + \frac{\eta}{2} \|\xi_k^+ - \xi_k^-\|_2^2 \Big) \\ \\ \text{subject to} & \xi_{k,j}^+ \geq 0 \text{ and } \xi_{k,j}^- \geq 0 \text{ for } j = 1, \ldots, p \end{array}$$

- ► Closed-form update for $\beta_k^{(w+1)}$
- Proximal gradient (ISTA) for the $\gamma_k^{(w+1)}$ update, based on Lemma 1 that states $\operatorname{prox}_{\operatorname{sg} h_1, \tilde{\eta}, \zeta} = \operatorname{prox}_{\zeta \tilde{\eta}} \sum_{l=1}^{L} \| \cdot \|_{2,l} \circ \operatorname{prox}_{\zeta (1-\tilde{\eta})} \| \cdot \|_1$

$$\lambda_0^{(w+1)}(t) = \frac{\sum_{i=1}^n \delta_i \mathbb{1}_{\{t=t_i\}}}{\sum_{i=1}^n \sum_{k=0}^{K-1} \hat{\pi}_{ik}^{\theta(w)} \exp\left\{\gamma_k^{(w+1)^\top} \Psi_i(t_i)\right\} \mathbb{1}_{\{t_i \geq t\}}}$$

ntroductio

Overview

Framework

Submodels

Votations

nference

prox-QNEM Evaluation

Application
Competitors
Simulations

onclusion

Real-time prediction paradigm

▶ Practitioners need predictive prognostic tools to be used in real-time

troduction

Overview

Frameworl

viouei

Submodels

Notations

Inference

rox-QNEM

Evaluation

Application

Simulations

Results

onclusion

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data

ntroduction

Overview

Use cases Framework

Model

Submodel

Votations

Penalizatio

Evaluation

valuation

ppiication

mpetitors nulations

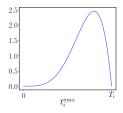
esults

onclusion

Performance evaluation

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data
- On test data : sample $t_i^{max} \sim T_i ig(1 \mathsf{Beta}(2,5)ig)$ the time for subject iwhen one wants to perform the risk prediction - the "present" time



Evaluation

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data
- ▶ On test data : sample $t_i^{max} \sim T_i (1 \text{Beta}(2, 5))$ the time for subject i when one wants to perform the risk prediction the "present" time
- ▶ Truncate y_i up to t_i^{max} to get \tilde{y}_i

ntroduction

Overview

Use cases Framework

Submodels

Votations

Penalization

Evaluation

Application

Competitors

esults

onclusion

Performance evaluation

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data
- ▶ On test data : sample $t_i^{max} \sim T_i (1 \text{Beta}(2, 5))$ the time for subject i when one wants to perform the risk prediction the "present" time
- ▶ Truncate y_i up to t_i^{max} to get \tilde{y}_i

Predictive marker

$$\mathbb{P}_{\hat{\theta}}[G_i = k | T_i > t_i^{max}, \tilde{y}_i] = \frac{\pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}{\sum_{k=0}^{K-1} \pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}$$

troduction

Use cases

ramework

Submodels Notations

nference Penalization

Evaluation

valuation

Application

Simulations

onclucion

) - f - u - u - u - u

Performance evaluation

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data
- ▶ On test data : sample $t_i^{max} \sim T_i (1 \text{Beta}(2, 5))$ the time for subject i when one wants to perform the risk prediction the "present" time
- ▶ Truncate y_i up to t_i^{max} to get \tilde{y}_i

Predictive marker

$$\mathbb{P}_{\hat{\theta}}[G_i = k | T_i > t_i^{max}, \tilde{y}_i] = \frac{\pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}{\sum_{k=0}^{K-1} \pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}$$

▶ In practice K = 2

troduction

Use cases

uhmodels

Votations

Penalization

Evaluation

Applicatio

Competitors Simulations

Conclusion

Real-time prediction paradigm

- Practitioners need predictive prognostic tools to be used in real-time
- ightharpoonup Obtain $\hat{\theta}$ on training data
- ▶ On test data : sample $t_i^{max} \sim T_i (1 \text{Beta}(2, 5))$ the time for subject i when one wants to perform the risk prediction the "present" time
- ▶ Truncate y_i up to t_i^{max} to get \tilde{y}_i

Predictive marker

$$\mathbb{P}_{\hat{\theta}}[G_i = k | T_i > t_i^{max}, \tilde{y}_i] = \frac{\pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}{\sum_{k=0}^{K-1} \pi_{\hat{\xi}_k}(x_i) f_{\hat{\theta}}(t_i^{max}, \delta_i = 0, \tilde{y}_i | G_i = k)}$$

- In practice K=2
- C-index metric $C_{\tau} = \mathbb{P}[\hat{\mathcal{R}}_i > \hat{\mathcal{R}}_j | T_i < T_j, T_i < \tau]$

troduction

Use cases

ubmodele

oubmodels Votations

enalization

Evaluation

Applicatio

Competitors

onclusion

► Landmark Cox model : survival R package [10]

Competitors

- ► Landmark Cox model : survival R package [10]
- ► The time-dependent Cox model [7] : survival R package

Introduction

Use cases

Framewo

Model

Notations

ikelihood

Inference

prox-QNEM

Evaluation

Application Competitors

ompetitors imulations

onclusion

Competing models

- ► Landmark Cox model: survival R package [10]
- ► The time-dependent Cox model [7] : survival R package
- ► Multivariate joint latent class model : R package lcmm [4]

ntroduction

Use cases

Use cases Frameworl

Submodel

Votations

Inference

Penalizatio

Evaluation

plication

Competitors Simulations

onclusion

Competing models

- ► Landmark Cox model: survival R package [10]
- ► The time-dependent Cox model [7] : survival R package
- ▶ Multivariate joint latent class model : R package lcmm [4]
- Multivariate shared random effect model : JMbayes package [5]

ntroduction

Use cases Framework

Submode

Notations Notations

Likelihood

Penalization

prox-QNEN

\ --!:--+:-

Competitors

imulations esults

onclusion

▶ High-risk subjects proportion $\pi_1 \in [0,1]$

Lights 15/20

Introduction

Use case

Framewor

Submodel

otations

Inference

Penalization prox-QNEM

pplicatio

Simulations

imulations

a malicata

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$

Lights 15/20

Introduction

Use cases

/lodel

ubmodels lotations

Inference

prox-QNEM

pplication

Simulations

resuits

- -

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $\blacktriangleright \ [x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$

Lights 15/20

Introduction

Overview

Use cases Framewor

Submodels

otations kelihood

Inference

prox-QNEN

pplicatio

Simulations

.........

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$

Lights 15/20

Introduction

Use cases

Eramous

Mode

Submodel

lotations

Inference

prox-QNEN

pplication

Simulations

.........

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \ldots, s$

Lights 15/20

ntroductio

Overviev

Framewor

Mode

Submodels Notations

Likeliho

Inference

prox-QNEN

ompetitors

Simulations

. .

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$

ntroduction

Overview

Use case

Framewor

Submodels

Notations

Inference

prox-QNEM

mnetitors

Simulations

Results

Conclusion

Lights 15/20

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{ |\pi_1 n| \text{ random samples without replacement} \}$
- $[x_{ii}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- \triangleright $x_{ii} \leftarrow x_{ii} + (-1)^{\mathbb{I}_{\{i \notin \mathcal{H}\}}}$ gap for $i = 1, \dots, s$
- $\blacktriangleright \ \xi = (\underline{\varsigma_1, \ldots, \varsigma_1}, 0, \ldots, 0) \in \mathbb{R}^p$
- $ightharpoonup G_i \sim \mathcal{B}(\pi_{\xi}(x_i))$
- $\triangleright S_k = \left\{ k \left\lfloor \frac{Lr_s}{\kappa} \right\rfloor + 1, \dots, (k+1) \left\lfloor \frac{Lr_s}{\kappa} \right\rfloor \right\}$

Simulations

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $\mathcal{H} = \{ |\pi_1 n| \text{ random samples without replacement} \}$
- $[x_{ii}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- \blacktriangleright $x_{ii} \leftarrow x_{ii} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$
- $ightharpoonup G_i \sim \mathcal{B}(\pi_{\varepsilon}(x_i))$

Simulations

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $\mathcal{H} = \{ |\pi_1 n| \text{ random samples without replacement} \}$
- $[x_{ii}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- \triangleright $x_{ii} \leftarrow x_{ii} + (-1)^{\mathbb{I}_{\{i \notin \mathcal{H}\}}}$ gap for $i = 1, \dots, s$
- $\qquad \qquad \xi = (\underbrace{\varsigma_1, \ldots, \varsigma_1}, 0, \ldots, 0) \in \mathbb{R}^p$
- $ightharpoonup G_i \sim \mathcal{B}(\pi_{\varepsilon}(x_i))$
- $\blacktriangleright \ \beta_k^l \sim \mathbb{1}_{\{l \in \mathcal{S}_k\}} \mathcal{N} \Big(\mu_k, \begin{bmatrix} \rho_3 & 0 \\ 0 & \rho_3 \end{bmatrix} \Big)$
- $Y_i'(t) = \epsilon_i'(t) + \mathbb{1}_{\{l \in S_i\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$

Simulations

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$

- $Y_i'(t) = \epsilon_i'(t) + \mathbb{1}_{\{l \in S_k\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$
- $\qquad \qquad \bullet \quad \left(\tilde{\Psi}_{k,a}(t,\beta_k,b_i^l) \right)_{a \in \{1,2\}} = \left(\beta_{k,1}^l + \beta_{k,2}^l t + b_{i,1}^l + b_{i,2}^l t , b_i^{l^{-1}} \right)$

Introduction

Use cases

lodel

otations

nference

prox-QNEM

Application

Competitors
Simulations

Results

nclusion

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$

- $Y_i'(t) = \epsilon_i'(t) + \mathbb{1}_{\{l \in S_k\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$
- $\qquad \qquad \left(\tilde{\Psi}_{k,a}(t,\beta_k,b_i^l) \right)_{a \in \{1,2\}} = \left(\beta_{k,1}^l + \beta_{k,2}^l t + b_{i,1}^l + b_{i,2}^l t , b_i^l \right)^{\top}$
- $\tilde{\gamma}_{k,a}^I = \varsigma_2 \mathbb{1}_{\{I \in \mathcal{S}_k, a \in \{1,2\}\}}$

ntroduction

Use cases Framework

> Submodels Notations

nference

prox-QNEM Evaluation

Application Competitors

Simulations

onclusion

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \{\lfloor \pi_1 n \rfloor \text{ random samples without replacement} \}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$
- $\blacktriangleright \ G_i \sim \mathcal{B}\big(\pi_{\xi}(x_i)\big)$

- $Y_i'(t) = \epsilon_i'(t) + \mathbb{1}_{\{l \in S_k\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$
- $\qquad \qquad \left(\tilde{\Psi}_{k,a}(t,\beta_k,b_i^l) \right)_{a \in \{1,2\}} = \left(\beta_{k,1}^l + \beta_{k,2}^l t + b_{i,1}^l + b_{i,2}^l t , b_i^{l^\top} \right)$
- $\qquad \qquad \tilde{\gamma}_{k,a}^I = \varsigma_2 \mathbb{1}_{\{I \in \mathcal{S}_k, \ a \in \{1,2\}\}}$

Introduction Overview

Use cases Framework

> ubmodels lotations

nference

prox-QNEM

Application

Simulations

nclusion

- ▶ High-risk subjects proportion $\pi_1 \in [0,1]$
- $ightharpoonup \mathcal{H} = \big\{ \lfloor \pi_1 n \rfloor \text{ random samples without replacement} \big\}$
- $[x_{ij}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$
- $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$
- $\blacktriangleright G_i \sim \mathcal{B}\big(\pi_{\xi}(x_i)\big)$

- $Y_i'(t) = \epsilon_i'(t) + \mathbb{1}_{\{l \in S_k\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$
- $\qquad \qquad \left(\tilde{\Psi}_{k,a}(t,\beta_k,b_i^l) \right)_{a \in \{1,2\}} = \left(\beta_{k,1}^l + \beta_{k,2}^l t + b_{i,1}^l + b_{i,2}^l t , b_i^l \right)^{\top}$
- $\qquad \qquad \tilde{\gamma}_{k,a}^I = \varsigma_2 \mathbb{1}_{\{I \in \mathcal{S}_k, \ a \in \{1,2\}\}}$
- ▶ Gompertz baseline $\lambda_0(t) = \kappa_1 \kappa_2 \exp(\kappa_2 t)$

Introduction Overview

Use cases Framework

Submodels

Inforence

prox-QNEM

Application

Simulations

nclusion

High-dimensional simulation study (1/2)

▶ High-risk subjects proportion $\pi_1 \in [0,1]$

 $\mathcal{H} = \{ |\pi_1 n| \text{ random samples without replacement} \}$

 $| [x_{ii}] \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, \Sigma_1(\rho_1))$

 $ightharpoonup x_{ij} \leftarrow x_{ij} + (-1)^{\mathbb{1}_{\{i \notin \mathcal{H}\}}} gap \text{ for } j = 1, \dots, s$

 $\xi = (\varsigma_1, \ldots, \varsigma_1, 0, \ldots, 0) \in \mathbb{R}^p$

 $ightharpoonup G_i \sim \mathcal{B}(\pi_{\varepsilon}(x_i))$

 $\triangleright S_k = \left\{ k \left\lfloor \frac{Lr_s}{K} \right\rfloor + 1, \dots, (k+1) \left\lfloor \frac{Lr_s}{K} \right\rfloor \right\}$

 $Y_i^l(t) = \epsilon_i^l(t) + \mathbb{1}_{\{l \in S_k\}} \sum_{k=0}^{K-1} \mathbb{1}_{\{G_i = k\}} ((1, t)^\top \beta_k^l + (1, t)^\top b_i^l)$

 $\tilde{\Psi}_{k,a}(t,\beta_k,b_i^l))_{a \in \{1,2\}} = (\beta_{k,1}^l + \beta_{k,2}^l t + b_{i,1}^l + b_{i,2}^l t , b_i^{l^{-1}})$ $\tilde{\gamma}_{k,a}^l = \varsigma_2 \mathbb{1}_{\{l \in \mathcal{S}_k, a \in \{1,2\}\}}$

 $\lambda_{i}(t|G_{i}=k) = \lambda_{0}(t) \exp \left\{ \iota_{i,k,1} + \iota_{i,k,2} t \right\}$

• Gompertz baseline $\lambda_0(t) = \kappa_1 \kappa_2 \exp(\kappa_2 t)$

 $T_i^{\star} | G_i = k \sim \frac{1}{\iota_{i,k,2} + \kappa_2} \log \left(1 - \frac{(\iota_{i,k,2} + \kappa_2) \log U_i}{\kappa_1 \kappa_2 \exp \iota_{i,k,1}} \right)$

roduction

Use cases Framework

> ibmodels otations

> ference

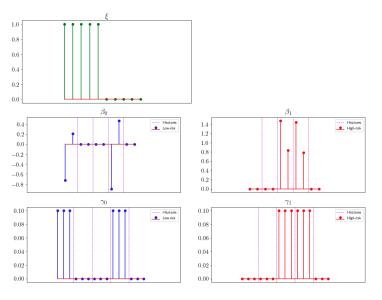
ox-QNEM valuation

Competitors
Simulations

nclusion

oforoncos

High-dimensional simulation study (2/2)



troduction

Use cases

odel

lotations

Inference

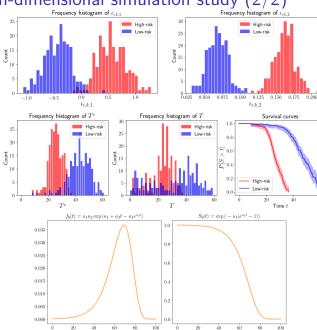
prox-QNEM

Application

Simulations

esults

onclusion



ntroductio

Use cases Framewor

Submodels

Notations Likelihood

Inference

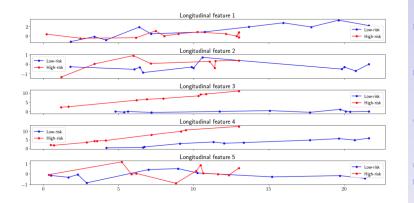
prox-QNEM Evaluation

pplicatio

Simulations

Conclusion

High-dimensional simulation study (2/2)



troduc

Jse cases ramework

odel . ..

otations

nforonco

enalization rox-QNEM

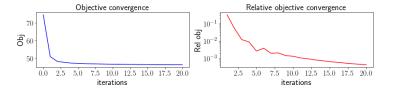
A --- | | |

Competitors

Simulations

esults

CONCIDSION



ntroduction

Use cases

Frameworl

Submodels

Notations Likelihood

Inference

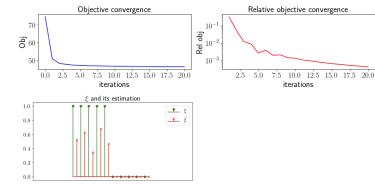
prox-QNEN

Application
Competitors

Results

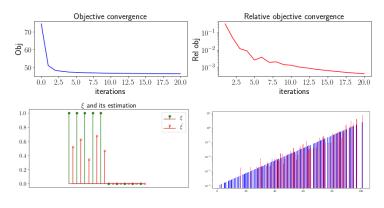
onclusion

First results (1/2)



Results

First results (1/2)



ntroductio

Overview

Framework

Submodels Notations

Notations Likelihood

Inference

prox-QNEM

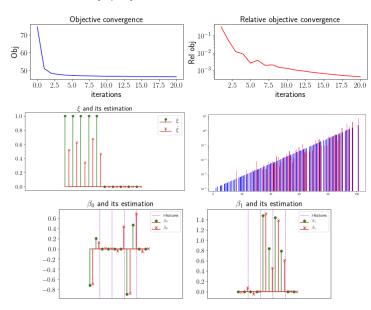
Applicat

Simulations

Results

Conclusion

First results (1/2)



ntroduc

Use cases

-ramework

ubmodels lotations ikelihood

Inference

prox-QNEM Evaluation

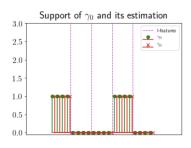
Applicat

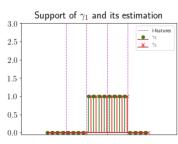
Simulations

Results

Conclusion

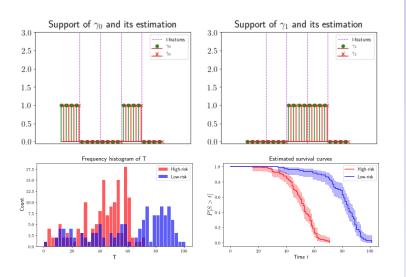
First results (2/2)





Results

First results (2/2)



Introduction

Use cases Framework

odel

lotations

nference

Penalization prox-QNEM Evaluation

Applicati

Competitors

Results

Conclusion

V. Conclusion

Introduction

Overview

Framewor

Notations

Likelihood

Inference

prox-QNEN

Evaluation

Application
Competitors

Simulations

Conclusion

Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available Introduction

Overvier

Framewo

Submodel

Notations

. .

Penalization

prox-QNEN

oplication

ompetitors mulations

Conclusion

- Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available
- Penalization of the likelihood in order to perform feature selection and to prevent overfitting

Introduction

Use case

Framev

/lodel

Submodels

Notations

Inference

prox-QNEI

Evaluation

plication

ompetitors mulations

sults

Conclusion

- Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available
- Penalization of the likelihood in order to perform feature selection and to prevent overfitting
- New efficient estimation algorithm (prox-QNEM) has been derived

Introduction

Use case

Submodel

Votations

Inference

orox-QNEN

Evaluation

Application

ompetitors mulations

Conclusion

- Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available
- Penalization of the likelihood in order to perform feature selection and to prevent overfitting
- New efficient estimation algorithm (prox-QNEM) has been derived
- Automatically determines significant prognostic longitudinal features

Introduction

Overviev

Framewor

Submode

Votations

Inference

orox-QNEN

Evaluation

Applicatio

imulations

esults

Conclusion

- Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available
- Penalization of the likelihood in order to perform feature selection and to prevent overfitting
- ▶ New efficient estimation algorithm (prox-QNEM) has been derived
- Automatically determines significant prognostic longitudinal features

Python 3 package

Available at https://github.com/Califrais/lights

Conclusion

- Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and censored durations, where a large number of longitudinal features are available
- Penalization of the likelihood in order to perform feature selection and to prevent overfitting
- New efficient estimation algorithm (prox-QNEM) has been derived
- Automatically determines significant prognostic longitudinal features

Python 3 package

- ► Available at https://github.com/Califrais/lights
- Applications of the model available soon

Introduction

Overviev Use case

Framewor

Submodel

Notations

nference

rox-QNEN

Evaluation

Applicat

Simulations

Conclusion

Lights 19/20

Thank you!

Introductioi

Overview

Eramous

Model

Submodels

Likelihood

Inference

Penalization prox-QNEM

Evaluation

ompetitors

Simulations

Conclusion

References

- Eleni-Rosalina Andrinopoulou, Dimitris Rizopoulos, Johanna JM Takkenberg, and Emmanuel Lesaffre. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. Statistical methods in medical research, 26(4):1787–1801, 2017.
- [2] Yueh-Yun Chi and Joseph G Ibrahim. Joint models for multivariate longitudinal and multivariate survival data. Biometrics, 62(2):432–445, 2006.
- [3] Laura A Hatfield, Mark E Boye, and Bradley P Carlin. Joint modeling of multiple longitudinal patient-reported outcomes and survival. *Journal of Biopharmaceutical Statistics*, 21(5):971–991, 2011.
- [4] Cécile Proust-Lima, Viviane Philipps, and Benoit Liquet. Estimation of extended mixed models using latent classes and latent processes: The r package lcmm. *Journal of Statistical Software, Articles*, 78(2):1–56, 2017. ISSN 1548-7660. doi: 10.18637/jss.v078.i02. URL https://www.jstatsoft.org/v078/i02.
- [5] Dimitris Rizopoulos. The r package imbayes for fitting joint models for longitudinal and time-to-event data using mcmc. Journal of Statistical Software, Articles, 72(7):1–46, 2016. ISSN 1548-7660. doi: 10.18637/jss.v072.i07. URL https://www.jstatsoft.org/v072/i07.
- [6] Dimitris Rizopoulos and Pulak Ghosh. A bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. Statistics in medicine, 30(12):1366–1380, 2011.
- [7] Glenn T Sueyoshi. Semiparametric proportional hazards estimation of competing risks models with time-varying covariates. *Journal of econometrics*, 51(1-2):25–58, 1992.
- [8] Jeroen K Vermunt and Jay Magidson. Latent class models for classification. Computational Statistics & Data Analysis, 41(3-4):531–537, 2003.
- [9] Michael S Wulfsohn and Anastasios A Tsiatis. A joint model for survival and longitudinal data measured with error. Biometrics, pages 330–339, 1997.
- [10] Zhongheng Zhang, Jaakko Reinikainen, Kazeem Adedayo Adeleke, Marcel E Pieterse, and Catharina GM Groothuis-Oudshoorn. Time-varying covariates and coefficients in cox regression models. Annals of translational medicine, 6(7), 2018.

ntroduction Overview

Use cases

odel

Votations

ikelihood

nference

prox-QNEN

Evaluation

Competitors

Simulations

onclusion

References