

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

IBC 2020 Online Learning Series

Simon Bussy

simon.bussy@gmail.com

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

S. Bussy ^(1,2), A. Barbieri ⁽³⁾, S. Zohar ⁽¹⁾, A.S. Jannot ^(1,4)

⁽¹⁾INSERM, UMRS 1138, Centre de Recherche des Cordeliers, Paris, France.

⁽²⁾Califrais' Machine Learning Lab, Paris, France.

⁽³⁾INSERM, UMR 1219, Bordeaux Population Health Research Center, Univ. Bordeaux, France.

⁽⁴⁾Biomedical Informatics and Public Health Department, EGP, APHP, Paris, France.

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion**References**

I. Introduction

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ 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

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ 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 Quasi-Newton Monte Carlo Expectation Maximization algorithm

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- 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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ 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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ 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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQEM

Conclusion

References

- ▶ 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
- ▶ Customer's satisfaction monitoring context → 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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

► Survival analysis

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- Survival analysis

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

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

High-dimensional framework

- Survival analysis

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

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

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- Survival analysis

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

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

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

- Heterogeneity of the population: latent subgroups

$$G \in \{0, \dots, K-1\}$$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- Survival analysis

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

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

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

- Heterogeneity of the population: latent subgroups

$$G \in \{0, \dots, K-1\}$$

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

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNE

Conclusion

References

[Introduction](#)[Overview](#)[Use cases](#)[Framework](#)[Model](#)[Submodels](#)[Notations](#)[Likelihood](#)[Inference](#)[Penalization](#)[MCQNEM](#)[Conclusion](#)[References](#)

II. Model

Submodels

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 $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0, D_{ll})$

Introduction

[Overview](#)[Use cases](#)[Framework](#)

Model

[Submodels](#)[Notations](#)[Likelihood](#)

Inference

[Penalization](#)[MCQNEM](#)

Conclusion

References

Submodels

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 $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0, D_{ll})$
- ▶ $\text{Cov}[b^l, b^{l'}] = D_{ll'}$ 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

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Submodels

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 $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0, D_{ll})$
- ▶ $\text{Cov}[b^l, b^{l'}] = D_{ll'}$

Group-specific risk of event

- ▶ $\lambda(t|G = k) = \lambda_0(t) \exp \left\{ x^\top \gamma_{k,0} + \sum_{l=1}^L \sum_{a=1}^A \gamma_{k,a}^l{}^\top \varphi_a(t, \beta_k^l, b^l) \right\}$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNE

Conclusion

References

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 $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0, D_{ll})$
- ▶ $\text{Cov}[b^l, b^{l'}] = D_{ll'}$

Group-specific risk of event

- ▶ $\lambda(t|G = k) = \lambda_0(t) \exp \left\{ x^\top \gamma_{k,0} + \sum_{l=1}^L \sum_{a=1}^A \gamma_{k,a}^l \varphi_a(t, \beta_k^l, b^l) \right\}$
- ▶ Functionals $(\varphi_a)_{a \in \mathcal{A}}$

Description	$\varphi_a(t, \beta_k^l, b^l)$	$\frac{\partial \varphi_a(t, \beta_k^l, b^l)}{\partial \beta_k^l}$	Reference
Linear predictor	$m_k^l(t)$	$u^l(t)$	Chi and Ibrahim [2]
Random effects	b^l	$\mathbf{0}_{q_l}$	Hatfield et al. [3]
Time-dependent slope	$\frac{d}{dt} m_k^l(t)$	$\frac{d}{dt} u^l(t)$	Rizopoulos and Ghosh [4]
Cumulative effect	$\int_0^t m_k^l(s) ds$	$\int_0^t u^l(s) ds$	Andrinopoulou et al. [1]

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNE

Conclusion

References

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQEM

Conclusion

References

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQEM

Conclusion

References

- ▶ $\mathcal{D}_n = \{(x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n)\}$ with $y_i^l = (y_{i1}^l, \dots, y_{in_l}^l)^\top \in \mathbb{R}^{n_l^l}$ and $y_{ij}^l = Y_i^l(t_{ij}^l)$
- ▶ $y_i = (y_i^{1^\top} \dots y_i^{L^\top})^\top \in \mathbb{R}^{n_i}$ with $n_i = \sum_{l=1}^L n_l^l$
- ▶ $b_i = (b_i^{1^\top} \dots b_i^{L^\top})^\top \in \mathbb{R}^r$ with $r = \sum_{l=1}^L r_l$
- ▶ Design matrices

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

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

$$\begin{cases} U_{il} = (u_i^l(t_{i1}^l)^\top \dots u_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times q_l}, \\ V_{il} = (v_i^l(t_{i1}^l)^\top \dots v_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times r_l}. \end{cases}$$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNE

Conclusion

References

- ▶ $\mathcal{D}_n = \{(x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n)\}$ with $y_i^l = (y_{i1}^l, \dots, y_{in_l}^l)^\top \in \mathbb{R}^{n_l^l}$ and $y_{ij}^l = Y_i^l(t_{ij}^l)$
- ▶ $y_i = (y_i^{1^\top} \dots y_i^{L^\top})^\top \in \mathbb{R}^{n_i}$ with $n_i = \sum_{l=1}^L n_l^l$
- ▶ $b_i = (b_i^{1^\top} \dots b_i^{L^\top})^\top \in \mathbb{R}^r$ with $r = \sum_{l=1}^L r_l$
- ▶ Design matrices

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

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

$$\begin{cases} U_{il} = (u_i^l(t_{i1}^l)^\top \dots u_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times q_l}, \\ V_{il} = (v_i^l(t_{i1}^l)^\top \dots v_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times r_l}. \end{cases}$$

- ▶ $\beta_k = (\beta_k^{1^\top} \dots \beta_k^{L^\top})^\top \in \mathbb{R}^q$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNE

Conclusion

References

- ▶ $\mathcal{D}_n = \{(x_1, y_1^1, \dots, y_1^L, t_1, \delta_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n)\}$ with $y_i^l = (y_{i1}^l, \dots, y_{in_l}^l)^\top \in \mathbb{R}^{n_l^l}$ and $y_{ij}^l = Y_i^l(t_{ij}^l)$
- ▶ $y_i = (y_i^{1^\top} \dots y_i^{L^\top})^\top \in \mathbb{R}^{n_i}$ with $n_i = \sum_{l=1}^L n_l^l$
- ▶ $b_i = (b_i^{1^\top} \dots b_i^{r^\top})^\top \in \mathbb{R}^r$ with $r = \sum_{l=1}^L r_l$
- ▶ Design matrices

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

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

$$\begin{cases} U_{il} = (u_i^l(t_{i1}^l)^\top \dots u_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times q_l}, \\ V_{il} = (v_i^l(t_{i1}^l)^\top \dots v_i^l(t_{in_l}^l)^\top)^\top \in \mathbb{R}^{n_l^l \times r_l}. \end{cases}$$

- ▶ $\beta_k = (\beta_k^{1^\top} \dots \beta_k^{L^\top})^\top \in \mathbb{R}^q$
- ▶ $M_{ik} = U_i \beta_k + V_i b_i \in \mathbb{R}^{n_i}$

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ $\theta = (\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \text{vech}(D), \lambda_0(t), \gamma_0^\top \cdots \gamma_{K-1}^\top) \in \mathbb{R}^{\vartheta}$
- ▶ Assumptions : (i) $b \perp\!\!\!\perp G$ (ii) $b|T, \Delta, Y \perp\!\!\!\perp G|T, \Delta, Y$
(iii) $Y'|b, G \perp\!\!\!\perp Y''|b, G$ (iv)
 $f(t_i, \delta_i, y_i|b_i, G_i = k; \theta) = f(t_i, \delta_i|b_i, G_i = k; \theta)f(y_i|b_i, G_i = k; \theta)$

Introduction

[Overview](#)[Use cases](#)[Framework](#)

Model

[Submodels](#)[Notations](#)[Likelihood](#)

Inference

[Penalization](#)[MCQNEM](#)

Conclusion

References

- ▶ $\theta = (\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \text{vech}(D), \lambda_0(t), \gamma_0^\top \cdots \gamma_{K-1}^\top) \in \mathbb{R}^{\vartheta}$
- ▶ Assumptions : (i) $b \perp\!\!\!\perp G$ (ii) $b|T, \Delta, Y \perp\!\!\!\perp G|T, \Delta, Y$
 (iii) $Y^l|b, G \perp\!\!\!\perp Y'^l|b, G$ (iv)
 $f(t_i, \delta_i, y_i|b_i, G_i = k; \theta) = f(t_i, \delta_i|b_i, G_i = k; \theta)f(y_i|b_i, G_i = k; \theta)$
- ▶ $f(y_i|b_i, G_i = k) = \exp \left\{ (y_i \odot \Phi_i)^\top M_{ik} - c_\phi(M_{ik}) + d_\phi(y_i) \right\}$ with
 $\Phi_i = (\phi_1^{-1} \mathbf{1}_{n_i^1}^\top \cdots \phi_L^{-1} \mathbf{1}_{n_i^L}^\top)^\top \in \mathbb{R}^{n_i}$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- ▶ $\theta = (\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \text{vech}(D), \lambda_0(t), \gamma_0^\top \cdots \gamma_{K-1}^\top) \in \mathbb{R}^{\vartheta}$
- ▶ Assumptions : (i) $b \perp\!\!\!\perp G$ (ii) $b|T, \Delta, Y \perp\!\!\!\perp G|T, \Delta, Y$
(iii) $Y'|b, G \perp\!\!\!\perp Y''|b, G$ (iv)
 $f(t_i, \delta_i, y_i|b_i, G_i = k; \theta) = f(t_i, \delta_i|b_i, G_i = k; \theta)f(y_i|b_i, G_i = k; \theta)$
- ▶ $f(y_i|b_i, G_i = k) = \exp \left\{ (y_i \odot \Phi_i)^\top M_{ik} - c_\phi(M_{ik}) + d_\phi(y_i) \right\}$ 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:

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

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNE

Conclusion

References

- ▶ $\theta = (\xi_0^\top \cdots \xi_{K-1}^\top, \beta_0^\top \cdots \beta_{K-1}^\top, \phi^\top, \text{vech}(D), \lambda_0(t), \gamma_0^\top \cdots \gamma_{K-1}^\top) \in \mathbb{R}^{\vartheta}$
- ▶ Assumptions : (i) $b \perp\!\!\!\perp G$ (ii) $b|T, \Delta, Y \perp\!\!\!\perp G|T, \Delta, Y$
 (iii) $Y^i|b, G \perp\!\!\!\perp Y^{i'}|b, G$ (iv)
 $f(t_i, \delta_i, y_i|b_i, G_i = k; \theta) = f(t_i, \delta_i|b_i, G_i = k; \theta)f(y_i|b_i, G_i = k; \theta)$
- ▶ $f(y_i|b_i, G_i = k) = \exp \left\{ (y_i \odot \Phi_i)^\top M_{ik} - c_\phi(M_{ik}) + d_\phi(y_i) \right\}$ 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:

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

- ▶ Then, the likelihood writes

$$\ell_n(\theta) = n^{-1} \sum_{i=1}^n \log \int_{\mathbb{R}^r} \sum_{k=0}^{K-1} \pi_{\xi_k}(x_i) f(t_i, \delta_i|b_i, G_i = k; \theta) \\ \times f(y_i|b_i, G_i = k; \theta) f(b_i; \theta) db_i$$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNM

Conclusion

References

III. Inference

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

► Penalized objective

$$\ell_n^{\text{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}l_1,\tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}l_1,\tilde{\eta}}$$

with the elasticnet penalty

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

and the sparse group lasso penalty

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

Introduction

[Overview](#)[Use cases](#)[Framework](#)

Model

[Submodels](#)[Notations](#)[Likelihood](#)

Inference

[Penalization](#)[MCQEM](#)

Conclusion

References

► Penalized objective

$$\ell_n^{\text{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sgl}_1,\tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sgl}_1,\tilde{\eta}}$$

with the elasticnet penalty

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

and the sparse group lasso penalty

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

► Resulting optimization problem

$$\hat{\theta} \in \operatorname{argmin}_{\theta \in \mathbb{R}^{\vartheta}} \ell_n^{\text{pen}}(\theta)$$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNE

Conclusion

References

MCQNEM algorithm (1/2)

IBC 2020

Lights

9/12

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

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

MCQNEM algorithm (1/2)

IBC 2020

Lights

9/12

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

Monte Carlo E-step

$$\blacktriangleright Q_n(\theta, \theta^{(w)}) = \mathbb{E}_{\theta^{(w)}}[\ell_n^{\text{comp}}(\theta) | \mathcal{D}_n]$$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

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

Monte Carlo E-step

► $\mathcal{Q}_n(\theta, \theta^{(w)}) = \mathbb{E}_{\theta^{(w)}}[\ell_n^{\text{comp}}(\theta) | \mathcal{D}_n]$

► Requires to compute expectations of the form

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

for different functions g , where we denote

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

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

- ▶ $\ell_n^{\text{comp}}(\theta) = \ell_n^{\text{comp}}(\theta; \mathcal{D}_n, \mathbf{b}, \mathbf{G})$

Monte Carlo E-step

- ▶ $\mathcal{Q}_n(\theta, \theta^{(w)}) = \mathbb{E}_{\theta^{(w)}}[\ell_n^{\text{comp}}(\theta) | \mathcal{D}_n]$
- ▶ Requires to compute expectations of the form

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

for different functions g , where we denote

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

- ▶ Monte Carlo approximations used for untractable integrals

MCQNEM algorithm (2/2)

IBC 2020
Lights
10/12

Quasi-Newton M-step

$$\blacktriangleright \theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sgf}_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sgf}_1, \tilde{\eta}}$$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}l_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}l_1, \tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}t_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}t_1, \tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$
- ▶ $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}l_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}l_1, \tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$
- ▶ $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$
- ▶ $\xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta}$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}t_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}t_1, \tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$
- ▶ $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$
- ▶ $\xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta}$
- ▶ L-BFGS-B to solve the problem

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

Introduction

[Overview](#)[Use cases](#)[Framework](#)

Model

[Submodels](#)[Notations](#)[Likelihood](#)

Inference

[Penalization](#)[MCQNEM](#)

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}l_1,\tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}l_1,\tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^\top | t_i, \delta_i, y_i]$
- ▶ $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$
- ▶ $\xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\text{en},\eta}$
- ▶ L-BFGS-B to solve the problem

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

- ▶ Similar tricks for $\beta_k^{(w+1)}$ and $\gamma_k^{(w+1)}$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Quasi-Newton M-step

- ▶ $\theta^{(w+1)} \in \operatorname{argmin}_{\theta} \mathcal{Q}_n(\theta, \theta^{(w)}) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta} + \zeta_{2,k} \|\gamma_k\|_{\text{sg}t_1, \tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\text{sg}t_1, \tilde{\eta}}$
- ▶ $D^{(w+1)} = n^{-1} \sum_{i=1}^n \hat{\mathbb{E}}_{\theta^{(w)}} [b_i b_i^T | t_i, \delta_i, y_i]$
- ▶ $P_{n,k}^{(w)}(\xi_k) = -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \log \pi_{\xi_k}(x_i)$
- ▶ $\xi_k^{(w+1)} \in \operatorname{argmin}_{\xi_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\xi_k) + \zeta_{1,k} \|\xi_k\|_{\text{en}, \eta}$
- ▶ L-BFGS-B to solve the problem

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

- ▶ Similar tricks for $\beta_k^{(w+1)}$ and $\gamma_k^{(w+1)}$
- ▶ Predictive marker $\hat{R}_{ik} = \frac{\pi_{\hat{\xi}_k}(x_i) \hat{f}(t_i^{\max}, y_i | b_i, G_i = k; \hat{\theta})}{\sum_{k=0}^{K-1} \pi_{\hat{\xi}_k}(x_i) \hat{f}(t_i^{\max}, y_i | b_i, G_i = k; \hat{\theta})}$, which is an estimate of $\mathbb{P}_{\theta}[G_i = k | T_i^* > t_i^{\max}, y_i]$

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

V. 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

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

- ▶ 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

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

- ▶ 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 (QNMCEM) has been derived

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

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 (QNMCEM) has been derived
- ▶ Automatically determines significant prognostic longitudinal features

Python 3 package

- ▶ Available at <https://github.com/Califrais/lights>

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

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 (QNMCEM) 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 on an arXiv paper.

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References

Thank you!

Introduction

Overview

Use cases

Framework

Model

Submodels

Notations

Likelihood

Inference

Penalization

MCQNEM

Conclusion

References

- [1] 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] 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.

Introduction

Overview
Use cases
Framework

Model

Submodels
Notations
Likelihood

Inference

Penalization
MCQNEM

Conclusion

References