





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, EGPH, APHP, Paris, France.

IBC 2020 Lights 1/12

ntroduction

Use cases

Submodel

Notations

_

MCQNEM

Circiasion

IBC 2020 Lights 1/12

Introduction

...

Use casi

1101110110

Submode

Notations

Likelihood

I. Introduction

illierence

MCQNEM

Conclusion

Overview

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

IBC 2020 Lights 2/12

Introduction

Overview

Use cases Framewor

Submode

Notations

| :|...|:|......................

nference

MCQNEM

Overview

IBC 2020 Lights 2/12

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

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

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

Use cases

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 IBC 2020 Lights 3/12

Introduction

Use cases

Framew

Model

Submodels

Notations

Informed

enalization

Conclusion

 Provides powerful interpretability by automatically pinpointing significant time-dependent and time-independent features Introduction

Use cases

Use case

Model

Submodels

Notations Likelihood

Penalization

) - f - u - u - u - u

 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

Use cases

rramewo

1odel

Votations

- C----

MCQNEM

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
- 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

Use cases

1odel

Notations

Inference

MCQNEM

Conclusion

Survival analysis

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

IBC 2020 Lights 4/12

Introduction

Ouenieur

Use cases Framework

N / - - | - |

Submodels Notations

Likelihood

Penalization

Conclusion

Survival analysis

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

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

IBC 2020 Lights 4/12

Framework

Survival analysis

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

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

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

IBC 2020 Lights 4/12

Introduction

11-- ----

Framework

Submode

Votations

. .

Penalization

Conclusion

Survival analysis

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

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

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

IBC 2020 Lights 4/12

ntroduction

Use cases

Framework

Submode

Notations

. .

MCQNEM

Conclusion

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

- ▶ 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

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

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

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

ntroduction

Use seem

Framework

Model Submod

Notations

Likelillood

Penalization

C l

IBC 2020 Lights 4/12

Introduction

Use cases

Framewo

Model

Notations Likelihood

Inference

Penalization

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

IBC 2020 Lights 5/12

Introduction

Hen races

Framewor

Лodel

Submodels

Likelihood

nference

enalization ICQNEM

Lonciusion

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

Introduction

Use cases

Framewo

1odel

Submodels

Notations

nference

MCONEM

CONCIUSION

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 $b^l \in \mathbb{R}^{r_l} \sim \mathcal{N}(0,D_l)$
- $\qquad \mathsf{Cov}[b^I,b^{I'}] = D_{II'}$

Group-specific risk of event

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

IBC 2020 Lights 5/12

Introduction

Use cases

Framewor

Submodels

Notations

Likelihood

enalization

onclusion

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

 $\triangleright \mathsf{Cov}[b^I,b^{I'}] = D_{II'}$

Group-specific risk of event

▶ Functionals $(\varphi_a)_{a \in \mathcal{A}}$

Description	$\varphi_{a}(t,\beta_{k}^{I},b^{I})$	$\frac{\partial \varphi_s(t,\beta_k^I,b^I)}{\partial \beta_k^I}$	Reference
Linear predictor	$m_k^l(t)$	u'(t)	Chi and Ibrahim [2]
Random effects	<i>b</i> ′	$0_{q_{I}}$	Hatfield et al. [3]
Time-dependent slope	$\frac{\mathrm{d}}{\mathrm{d}t}m_k^I(t)$	$rac{\mathrm{d}}{\mathrm{d}t}u^I(t)$	Rizopoulos and Ghosh [4]
Cumulative effect	$\int_0^t m_k^I(s) \mathrm{d} s$	$\int_0^t u^l(s) \mathrm{d} s$	Andrinopoulou et al. [1]

ntroduction

Use cases

Framework

Submodels

l-a-ai---

_

enalization

Conclusion

Notations

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

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

IBC 2020 Lights 6/12

Introduction

Use case

Framewor

Submodel

Notations

Inference

Penalization MCQNEM

Conclusion

Notations

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

IBC 2020 Lights 6/12

Introduction

Use cases

Eramous

1odel

Submodel

Notations

Inference

Penalization MCONEM

Conclusion

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 $r = \sum_{l=1}^L r_l$

IBC 2020 Lights 6/12

Introduction

Use cases

Framewor

-

Notations

Inference

Penalizatio MCQNEM

Conclusion

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

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

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

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

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

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

Notations

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

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

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

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

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

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

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

Notations

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

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

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

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

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

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

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

Notations

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

IBC 2020 Lights 7/12

Introduction

Use case:

Framewor

iviouei

Notations

Likelihood

Inference

Penalization MCQNEM

Conclusion

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

IBC 2020 Lights 7/12

Introduction

Use case

Framework

Submodel

Notations

Likelihood

Penalization

Conclusion

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

IBC 2020 Lights 7/12

Introduction

Use cases

Frameworl

/lodel

Notations Likelihood

I... C.....

nference Penalization

Conclusion

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

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

IBC 2020 Lights 7/12

Introduction

Use cases

/lodel

Notations Likelihood

Linciniood

Penalization

Conclusion

Assumptions: (i)
$$b \perp G$$
 (ii) $b \mid T, \Delta, Y \perp G \mid T, \Delta, Y$ (iii) $Y' \mid b, G \perp Y'' \mid b, G$ (iv) $f(t_i, \delta_i, y_i \mid b_i, G_i = k; \theta) = f(t_i, \delta_i \mid b_i, G_i = k; \theta) f(y_i \mid b_i, G_i = k; \theta)$

- $f(y_i|b_i, G_i = k) = \exp\left\{ (y_i \odot \Phi_i)^\top M_{ik} c_\phi(M_{ik}) + d_\phi(y_i) \right\} \text{ with } \Phi_i = (\phi_1^{-1} \mathbf{1}_{n_i^1}^\top \cdots \phi_L^{-1} \mathbf{1}_{n_i^L}^\top)^\top \in \mathbb{R}^{n_i}$
- Survival part:

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

► Then, the likelihood writes

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

IBC 2020 Lights 7/12

Introduction

Use cases Framework

1odel

Submodels Notations

I... C.....

Penalization

Conclusion

IBC 2020 Lights 7/12

minoduction

Use cases

N.A. J. J. I

Submodels

Notations Likelihood

Inference

Penalization MCQNEM

Conclusion

deferences

III. Inference

Penalization

Penalized objective

$$\ell_n^{\mathsf{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\mathsf{sg}I_1,\tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\mathsf{sg}I_1,\tilde{\eta}}$$

with the elasticnet penalty

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

and the sparse group lasso penalty

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

IBC 2020 Lights 8/12

Penalization

$$\ell_n^{\mathsf{pen}}(\theta) = -\ell_n(\theta) + \sum_{k=0}^{K-1} \zeta_{1,k} \|\xi_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\gamma_k\|_{\mathsf{sg}I_1,\tilde{\eta}} + \zeta_{3,k} \|\beta_k\|_{\mathsf{sg}I_1,\tilde{\eta}}$$

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

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

Penalization

MCQNEM algorithm (1/2)

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

IBC 2020 Lights 9/12

ntroduction

Use cases

Frameworl

iviouei

Votations

nference

Penalization MCONEM

Conclusion

MCQNEM algorithm (1/2)

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

Monte Carlo E-step

IBC 2020 Lights 9/12

Introduction

Use cases

Vlodel

Submodels Notations

Likelihood

Penalization MCONEM

Conclusion

Monte Carlo E-step

- ► Requires to compute expectations of the form

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

for different functions g, where we denote

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

ntroduction

Use cases

lodel

Submodels Notations

nference

Penalization MCQNEM

onclusion

Monte Carlo E-step

- ► Requires to compute expectations of the form

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

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

ntroduction

Use cases

Framework

.......

Notations

nference

Penalization MCQNEM

Conclusion

Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta^{(w+1)}} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{\eta}}(\boldsymbol{\theta}, \boldsymbol{\theta^{(w)}}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg} l_1, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg} l_1, \tilde{\boldsymbol{\eta}}}$$

IBC 2020 Lights 10/12

Introduction

Use cases

Submode

Notations

nference

MCQNEM

Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

$$D^{(w+1)} = n^{-1} \sum_{i=1}^{n} \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$$

IBC 2020 Lights 10/12

Introduction

Use cases

rramewori

Submodel

Notations Likelihood

Penalization

MCQNEM

Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

$$D^{(w+1)} = n^{-1} \sum_{i=1}^{n} \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$$

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

IBC 2020 Lights 10/12

MCONEM

Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{\eta}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \gamma_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

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

IBC 2020 Lights 10/12

ntroduction

Use cases

Framework

Submodel

Notations

Inference

Penalization MCQNEM

Conclusion

Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

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

- ▶ L-BFGS-B to solve the problem

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

IBC 2020 Lights 10/12

Introduction

Use cases

Framework

Submodels

Notations Likelihood

Populization

MCQNEM

Conclusion

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

$$D^{(w+1)} = n^{-1} \sum_{i=1}^{n} \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$$

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

- $\qquad \qquad \boldsymbol{\xi}_k^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\xi}_k \in \mathbb{R}^p} P_{n,k}^{(w)}(\boldsymbol{\xi}_k) + \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\operatorname{en},\eta}$
- ▶ L-BFGS-B to solve the problem

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

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

Introduction

Use cases

lodel

Submodels Notations

Inference

Penalization MCQNEM

Conclusion

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \operatorname{argmin}_{\boldsymbol{\theta}} \, \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \, \zeta_{1,k} \| \boldsymbol{\xi}_k \|_{\operatorname{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \| \boldsymbol{\gamma}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}} + \zeta_{3,k} \| \boldsymbol{\beta}_k \|_{\operatorname{sg}l_1, \, \tilde{\boldsymbol{\eta}}}$$

$$D^{(w+1)} = n^{-1} \sum_{i=1}^{n} \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$$

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

- ▶ L-BFGS-B to solve the problem

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

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

ntroduction

Use cases

odel

Submodels Notations

Inference

Penalization MCQNEM

Conclusion

IBC 2020 Lights 10/12

Introduction

Ouomious

Use case:

.

Submodel

Votations

Inferenc

V. Conclusion

Penalizatio

Conclusion

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 IBC 2020 Lights 11/12

Introduction

Use cases

Model

Submodels

Votations

Inference

nalization CQNEM

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

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

Conclusion

- 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

Use cases

/lodel

Notations

ikelihood

nterence

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

Use cases

/lodel

Notations

Likelihood

nference

WCQNEW

Conclusion

IBC 2020 Lights 11/12

Introduction

Hee cases

USE Case:

Model

Notations

Likelihood

ninerence

MCONEM

Conclusion

References

Thank you!

[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.

Johanna JM Takkenberg, and Emmanuel Lesaffre. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. Statistical methods in

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.

[1] Eleni-Rosalina Andrinopoulou, Dimitris Rizopoulos,

[2] Yueh-Yun Chi and Joseph G Ibrahim. Joint models for multivariate longitudinal and multivariate survival data.

medical research, 26(4):1787-1801, 2017.

Biometrics, 62(2):432-445, 2006.