





2020 Joint Statistical Meetings

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.

JSM 2020 Lights 1/12

Introductio

Use cases Framework

Submode

Notations

nference

`onclucion

JSM 2020 Lights 1/12

Introduction

Overview Use cases

Framework

Submodel

Notations Likelihood

Penalization

QNMCEM

00110101011

References

I. Introduction

Overview

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

JSM 2020 Lights 2/12

Introduction

Overview

Use cases Frameworl

Submode

Submodel

Notations

Inference

enalization NMCEM

Conclusion

Overview

2/12

JSM 2020

Lights

ntroduction

Overview

Framework

Submode

Notations

Likelihood

Interence

- CONTRACTOR

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

ntroduction

Overview

Frameworl

Submode

Notations

Likelihood

Descliention

- -

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

troduction

Overview

Framework

Submod

Notations

nference

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

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 JSM 2020 Lights 3/12

Introduction

Overviev

Use cases

Model

Submode

Notations

nference

enalization

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

Introduction

Overview

Use cases

Model

Submodels

Notations

nference

QNMCEM

CONCIUSION

 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

Eramous

Model

Submodel

Notations Likelihood

nference

..........

_ .

 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

. A . I . I

Submodels

Notations Likelihood

Panalization

Q14111CEIII

Survival analysis

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

JSM 2020 Lights 4/12

Introduction

Overview Hea cases

Framework

Subn

Notations

nference

RNMCEM

Survival analysis

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

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

JSM 2020 Lights 4/12

Introduction

Overview Use cases

Framework

Submod

Notations

nference 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

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

JSM 2020 Lights 4/12

Introduction

Overview Use cases

Framework

Submodel

Notations

.....

Penalization QNMCEM

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

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

Heterogeneity of the population: latent subgroups

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

JSM 2020 Lights 4/12

ntroduction

Use cases

Framework

/lodel

Notations

Likelihood

Penalization

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

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

▶ 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

Overview

Framework

Model

Notations

Likelihood

Penalization

. . .

JSM 2020 Lights 4/12

Introduction

Use cases

Framev

Model

Submodels Notations

Inference

Penalizatio ONMCEM

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

JSM 2020 Lights 5/12

Introduction

Use caree

Framewo

lodel

Submodels Notations

Likelihood

rerence enalization

Conclusion

- ▶ $h_l(\mathbb{E}[Y^l(t)|b^l,G=k]) = m_k^l(t) = u^l(t)^\top \beta_k^l + v^l(t)^\top b^l$ 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

Framewo

Submodels

Notations

Likelihood

nierence

QIVIVICEIVI

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

Group-specific risk of event

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

JSM 2020 Lights 5/12

ntroduction

Use cases

Framewo

1odel

Submodels

Notations

nference Penalization

enalization NMCEM

onclusion

Group-specific marker trajectories

- ▶ $h_I(\mathbb{E}[Y^I(t)|b^I,G=k]) = m_k^I(t) = u^I(t)^\top \beta_k^I + v^I(t)^\top b^I$ with fixed effect parameters $\beta_k^I \in \mathbb{R}^{q_I}$ and subject-and-longitudinal outcome specific random effects $b^I \in \mathbb{R}^{r_I} \sim \mathcal{N}(0,D_I)$
- $\blacktriangleright \ \mathsf{Cov}[b^I,b^{I'}] = D_{II'}$

Group-specific risk of event

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

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

ntroduction

Use cases

Framewor

Submodels

Notations

ference

QNMCEM

References

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

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

JSM 2020 Lights 6/12

Notations

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

JSM 2020 Lights 6/12

Introduction

Overview

Frameworl

Model

ubmodels

Notations Likelihood

nterence Penalization

Conclusion

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

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

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

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

JSM 2020 Lights 6/12

Introduction

Use cases

Framework

ubmodels lotations

Notations Likelihood

Penalization

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

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

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

Introduction

JSM 2020

Lights 6/12

Overview Use cases

Framework

Submodels Notations

Notations Likelihood

Inference

Conclusion

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

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

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

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

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

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

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

Introduction

Overview Use cases

Framework

Submodels Notations

Likelihood

Inference

QNMCEM

Likelihood

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

JSM 2020 Lights 7/12

Introduction

Overview Use cases

. . . .

Submo

Notations

Likelihood

Inference

Penalization QNMCEM

Conclusion

Likelihood

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

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

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

JSM 2020 Lights 7/12

ntroduction

Use cases

Framewor

Submo

Notations Likelihood

nforonco

rference enalization

Conclusion

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

Survival part:

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

ntroduction

Overview Use cases

Framework

Submodels

Notations Likelihood

> ference enalization

Conclusion

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

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

► Survival part:

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

► Then, the likelihood writes

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

ntroduction

Use cases

Submodels

Notations Likelihood

Penalization

Conclusion

JSM 2020 Lights 7/12

Introduction

Overview Use cases

rramewo

Submodel

Notations Likelihood

Inference

Penalization QNMCEM

Conclusion

References

III. Inference

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

with the elasticnet penalty

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

and the sparse group lasso penalty

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

JSM 2020 Lights 8/12

ntroduction

Overview

Use cases Framework

Submodel

Notations

Inference

Penalization QNMCEM

Conclusion

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

with the elasticnet penalty

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

and the sparse group lasso penalty

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

Resulting optimization problem

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

JSM 2020 Lights 8/12

ntroduction

Overview Use cases

1odel

Submodels Notations

Inforence

Penalization QNMCEM

Conclusion

QNMCEM algorithm (1/2)

JSM 2020 Lights 9/12

ntroduction

Overview

Framewor

Subr

Notations

Penalization QNMCEM

Conclusion

QNMCEM algorithm (1/2)

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

Monte Carlo E-step

JSM 2020 Lights 9/12

Introduction

Overview

Framework

Submod

Notations

Likelihood

Penalizatio

QNMCEM

Inference

QNMCEM

Conclusion

References

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

Monte Carlo E-step

- ▶ Requires to compute expectations of the form

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

for different functions g, where we denote

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

Submodel

Notations

nference

QNMCEM

Conclusion

References

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

Monte Carlo E-step

- ▶ Requires to compute expectations of the form

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

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

Quasi-Newton M-step

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

JSM 2020 Lights 10/12

ntroduction

Overview Use cases

Framewor

Submo

Notations

Penalization

QNMCEM

.

Quasi-Newton M-step

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

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

JSM 2020 Lights 10/12

ntroduction

Overview Use cases

Framework

Submo

Notations

Inference

Penalization QNMCEM

onclusion

Quasi-Newton M-step

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

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

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

JSM 2020 Lights 10/12

introduction

Overview Use cases

Framework

Submode

Notations

Inferenc

QNMCEM

Conclusion

Quasi-Newton M-step

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

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

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

JSM 2020 Lights 10/12

Illuoduction

Overview Use cases

Framework

Submode

Notations

Inferenc

Penalization QNMCEM

Conclusion

JSM 2020 Lights 10/12

Quasi-Newton M-step

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

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

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

- ▶ L-BFGS-B to solve the problem

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

Introduction

Overview Use cases

ramework

Submodels

Notations

Penalization

QNMCEM

Conclusion

Quasi-Newton M-step

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

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

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

- ▶ L-BFGS-B to solve the problem

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

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

Introduction

Overview Use cases

ramework

Submodels

Notations

Penalization

QNMCEM

Quasi-Newton M-step

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

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

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

- ▶ 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 $\gamma_k^{(w+1)}$
- $\begin{array}{l} \blacktriangleright \ \, \text{Predictive marker } \hat{\mathcal{R}}_{ik} = \frac{\pi_{\hat{\xi}_k}(\mathbf{x}_i)\hat{f}(t_i^{\textit{max}},y_i|b_i,G_i=k;\hat{\theta})}{\sum_{k=0}^{K-1}\pi_{\hat{\xi}_k}(\mathbf{x}_i)\hat{f}(t_i^{\textit{max}},y_i|b_i,G_i=k;\hat{\theta})}, \, \text{which} \\ \text{is an estimate of } \mathbb{P}_{\theta}[G_i=k|T_i^{\star}>t_i^{\textit{max}},y_i] \end{array}$

Introduction

Overview Use cases

Framework

Submodels

Notations

nterence

QNMCEM

Conclusion

JSM 2020 Lights 10/12

troduction

Use cases

Framewor

Suh

Notations Likelihood

Interence

Penalization QNMCEM

Conclusion

References

V. 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 JSM 2020 Lights 11/12

Introducti

Overview

.

o

Notations

Likelinood

QNMCEM

Conclusion

deferences

- 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

Introduct

Overview

Framewo

Nodel

Submodels

Likelihood

nference

NMCEM

Conclusion

deferences

- 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

Introductio

Overview

. . . .

iviouei

Notations

Likelihood

nference

QNMCEM

Conclusion

deferences

- 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

Introduction

Overview

1odel

Submodels

Likelihood

nference

Conclusion

Conclusion

Python 3 package

been derived

longitudinal features

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

Prognostic method called lights introduced to deal with the problem of joint modeling of longitudinal data and

censored durations, where a large number of

Penalization of the likelihood in order to perform feature selection and to prevent overfitting

Automatically determines significant prognostic

▶ New efficient estimation algorithm (QNMCEM) has

longitudinal features are available

Conclusion

nference

QNMCEM

Conclusion

References

Conclusion

Prognostic method called lights introduced to d

- 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.

JSM 2020 Lights 11/12

troduction

Use cases

Framewor

Subn

Notations Likelihood

Inference

Penalization QNMCEM

Conclusion

References

Thank you!

- [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

Use cases

Framewor

vlodel

Notations

nference

Statute Fini

Poforoncoc