# I. Introduction

#### Introduction

Overview Use cases

#### Mode

Submode

Likelihood

#### Inference

QNEM

Evaluation

Conclusion

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

Overview

Use cases

Model

Notations

Notations

Inference

Penalization QNEM

Evaluation

Conclusion

- 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

Model

Notations

Inference

Penalization QNEM

Evaluation

Conclusion

- 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 case

Model

Notations

Inference Penalization

Conclusion

- 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 Expectation Maximization algorithm

Introduction

Overview

Use cases

Model Submodels

Notations Likelihood

Penalization

Evaluation

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 Introduction

Overview Use cases

Use case:

Model

Submodol

Notation:

Inference

Penalization ONEM

 $\mathsf{Evaluation}$ 

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

ntroduction

Use cases

Use case:

Model

Notations

Likelinood

Penalization

Evaluation

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 ntroductio

Use cases

Framewo

C 1 11

Notations

nference

QIVEIVI

Conclusion

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

ntroduction

Use cases

Framewo

Notations

nference

QNEM

. . .

Conclusion

Survival analysis

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

Introduction

Overview

Framework

Model

o . . . .

Notations

Inference

Penalization QNEM

Evaluation

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  $p \gg n$ 

Introduction

Overview

Framework

Model

C., h., ... ... ... ... ... ...

Notations Likelihood

Inference

Penalization QNEM

Conclusion

Survival analysis

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

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

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

Introduction

Overview

Framework

Model

iviodei

Notations

nference

Penalization

Evaluation

Conclusion

Survival analysis

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

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

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

Heterogeneity of the population: latent subgroups

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

Introduction

Overview

Framework

. . . . .

Model

Notations

nference

QNEM

c . .

Conclusion

Survival analysis

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

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

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

► Heterogeneity of the population: latent subgroups

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

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

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

Introduction

Overview

Framework

Model

Notations

nference Penalization

Evaluation

Conclusion

## II. Model

#### Introduction

Overview Use cases

#### Model

Submodels Notations Likelihood

Inference

QNEM

Canalisaian

Deferences

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

Introduction

Overviev

Use cases

Model

Submodels

Notations Likelihoo

> nterence Penalization

QNEM

Conclusion

## 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

Model

Submodels

Notations Likelihood

Penalization

Evaluation

Conclusion

## Group-specific risk of event

Introduction

Overvie

F-----

Mod

Submodels

Notations

Inference

Penalization ONEM

Evaluation

Conclusion

## Group-specific risk of event

- ightharpoonup Fetures  $\Psi$  are extracted from Y by the Python library tsfresh
- ► TODO: Add descriptions

Introduction

Overview

Use case:

Model

Submodels

Notations

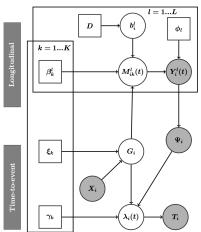
Inference

Penalization QNEM

Lvaldation

Conclusion

 Graphical representation of a joint model of a time-to-event submodel and K-multivariate longitudinal outcomes submodel



ntroduction

Overviev

Framework

/lodel

Submodels Notations

Likelihood

Penalization

Evaluation

Conclusion

$$\mathcal{D}_{n} = \left\{ (x_{1}, y_{1}^{1}, \dots, y_{1}^{L}, t_{1}, \delta_{1}, \Psi_{1}), \dots, (x_{n}, y_{n}^{1}, \dots, y_{n}^{L}, t_{n}, \delta_{n}, \Psi_{n}) \right\}$$
 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

Model

Notations

Likelihoo

nference

QNEM

Conclusion

Conclusion

- $\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1, \Psi_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n, \Psi_n) \right\}$  with  $y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l}$  and  $y_{ij}^l = Y_i^l(t_{ij}^l)$
- $y_i = (y_i^{1^\top} \cdots y_i^{L^\top})^\top \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^L n_i^l$

Introduction

Overview Use case:

Use cases Framewor

Submodels

Notations

lu.fauau aa

Penalization

Evaluation

Conclusion

$$\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1, \Psi_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n, \Psi_n) \right\}$$
 with  $y_i^l = (y_{i1}^l, \dots, y_{in_i^l}^l)^\top \in \mathbb{R}^{n_i^l}$  and  $y_{ij}^l = Y_i^l(t_{ij}^l)$ 

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

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

Introduction

Overview Use cases

Use cases

Submodels

Notations

Likelihood

Inference

QNEM

Conclusion

- $\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1, \Psi_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n, \Psi_n) \right\}$  with  $y_i^l = (y_{i1}^l, \dots, y_{in!}^l)^\top \in \mathbb{R}^{n_i^l}$  and  $y_{ij}^l = Y_i^l(t_{ij}^l)$
- $y_i = (y_i^{1\top} \cdots y_i^{L\top})^{\top} \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^{L} n_i^l$
- $lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$
- 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

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

Introduction

Use cases

Model

Notations

nference Penalization

Evaluation

Conclusion

- $\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1, \Psi_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n, \Psi_n) \right\}$  with  $y_i^l = (y_{i1}^l, \dots, y_{in!}^l)^\top \in \mathbb{R}^{n_i^l}$  and  $y_{ij}^l = Y_i^l(t_{ij}^l)$
- $y_i = (y_i^{1\top} \cdots y_i^{L\top})^{\top} \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^{L} n_i^l$
- $lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$
- 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

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

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

Introduction

Use cases

Submodels

Notations

nference Penalization

Evaluation

Conclusion

- $\mathcal{D}_n = \left\{ (x_1, y_1^1, \dots, y_1^L, t_1, \delta_1, \Psi_1), \dots, (x_n, y_n^1, \dots, y_n^L, t_n, \delta_n, \Psi_n) \right\}$  with  $y_i^l = (y_{i1}^l, \dots, y_{inl}^l)^\top \in \mathbb{R}^{n_i^l}$  and  $y_{ij}^l = Y_i^l(t_{ij}^l)$
- $y_i = (y_i^{1\top} \cdots y_i^{L\top})^{\top} \in \mathbb{R}^{n_i} \text{ with } n_i = \sum_{l=1}^{L} n_i^l$
- $lackbox{b}_i = (b_i^{1^\top} \cdots b_i^{L^\top})^\top \in \mathbb{R}^r \text{ with } r = \sum_{l=1}^L r_l$
- Design matrices

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

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

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

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

Introduction

Use cases

Model

Notations

Inference

QNEM

Evaluation

Conclusion

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

Introduction

Overvie

Eramow

Model

Submode

Likelihood

Inference

enalization

Evaluation

Conclusion

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

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

Introduction

Use cases

Use case:

Model

Notations

Likelihood

Inference

enalization

\_valuation

Conclusion

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

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

Likelihood

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

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

Survival part:

$$f(t_i, \delta_i | G_i = k, \Psi_i; \theta) = \left[ \lambda(t_i | G_i = k, \Psi_i) \right]^{\delta_i}$$

$$\times \exp \left\{ - \int_0^{t_i} \lambda(s | G_i = k, \Psi_i) ds \right\}$$

Then, the likelihood writes

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

$$\times f(y_i | b_i, G_i = k, \Psi_i; \theta) f(b_i; \theta) db_i$$

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

Introduction

Use cases

odel

Notations Likelihood

nference Penalization

Conclusion

## III. Inference

#### Introduction

Overvie

Framewo

#### Model

Notation

Likelihood

#### Inference

Penalizatio

Evaluation

Conclusion

#### Penalization

Penalized objective

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

with the elasticnet penalty

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

and the sparse group lasso penalty

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

Introductio

Overviev Use case

1odel

Notations

Inference

Penalization QNEM

Evaluation

Conclusion

#### Penalization

Penalized objective

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

with the elasticnet penalty

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

and the sparse group lasso penalty

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

Resulting optimization problem

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

Introductio

Overviev Use case

1odel

Notations

Inference

Penalization QNEM

Evaluation

Conclusion

# QNEM algorithm (1/2)

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

## E-step

Introduction

Overvie

Use case:

Model

Notation

Inference

Penalization QNEM

Evaluation

Conclusion

# QNEM algorithm (1/2)

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

## E-step

Introduction

Overview

Use case

Model

Notations

Likelihoo

nterence

Penalization QNEM

Evaluation

Conclusion

# QNEM algorithm (1/2)

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

## 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

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

Introduction

Overview

Use cases Framewor

Model

Notations

Likelihood

Penalization
ONEM

Evaluation

Conclusion

# Quasi-Newton M-step

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

Introduction

Overviev

Casi

Model

viouei

Notations

nterence

Penalization QNEM

Evaluation

Conclusion

## Quasi-Newton M-step

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

### Introduction

Overviev

Use case

Model

viouci

Notations

nference

Penalization QNEM

Evaluation

Conclusion

## Quasi-Newton M-step

- $\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en}, \, \boldsymbol{\eta}} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sgl}_1, \, \tilde{\boldsymbol{\eta}}}$
- $D^{(w+1)} = n^{-1} \sum_{i=1}^{n} \hat{\mathbb{E}}_{\theta^{(w)}}[b_i b_i^{\top} | t_i, \delta_i, y_i]$
- $\begin{array}{l} \blacktriangleright & R_{n,k}^{(w)}(\beta_k) = \\ & -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik} | t_i, \delta_i, y_i] \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik}) | t_i, \delta_i, y_i] \right] \end{array}$

#### Introduction

Overview

Use case:

Model

Submodels

Notations

nference

QNEM

. . .

## Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sgl}_1,\tilde{\eta}}$$

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

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik}|t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik})|t_i, \delta_i, y_i] \right]$$

 $\qquad \qquad \beta_k^{(w+1)} \in \operatorname{argmin}_{\beta_k \in \mathbb{R}^q} \, R_{n,k}^{(w)}(\beta_k)$ 

#### Introduction

Overview

Use case

### Model

Notations

Likelihood

Penalization ONFM

### .....

*c* . .

Deferences

## Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sgl}_1,\tilde{\eta}}$$

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

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik} | t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik}) | t_i, \delta_i, y_i] \right]$$

$$\qquad \qquad \beta_k^{(w+1)} \in \operatorname{argmin}_{\beta_k \in \mathbb{R}^q} R_{n,k}^{(w)}(\beta_k)$$

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

#### ntroduction

Overview

Use case

### Model

Notations

Likelihood

Penalization

QNEM

Conclusion

## Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sgl}_1,\tilde{\eta}}$$

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik}|t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik})|t_i, \delta_i, y_i] \right]$$

$$\qquad \qquad \beta_k^{(w+1)} \in \operatorname{argmin}_{\beta_k \in \mathbb{R}^q} \, R_{n,k}^{(w)}(\beta_k)$$

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

Introduction

Overview

Use case

Model

Notations

Likelihood

Penalization ONFM

Evaluation

Conclusion

## Quasi-Newton M-step

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

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik} | t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik}) | t_i, \delta_i, y_i] \right]$$

$$\qquad \qquad \beta_k^{(w+1)} \in \operatorname{argmin}_{\beta_k \in \mathbb{R}^q} \, R_{n,k}^{(w)}(\beta_k)$$

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

▶ L-BFGS-B to solve the problem

#### Introduction

Overviev

Use case:

### Model

Submodels

Likelihood

Penalization

ONFM

### Evaluation

Conclusion

## Quasi-Newton M-step

$$\qquad \qquad \boldsymbol{\theta}^{(w+1)} \in \mathsf{argmin}_{\boldsymbol{\theta}} \; \mathcal{Q}_{\boldsymbol{n}}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(w)}) + \sum\nolimits_{k=0}^{K-1} \zeta_{1,k} \|\boldsymbol{\xi}_k\|_{\mathsf{en},\eta} + \zeta_{2,k} \|\boldsymbol{\gamma}_k\|_{\mathsf{sgl}_1,\tilde{\eta}}$$

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

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik}|t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik})|t_i, \delta_i, y_i] \right]$$

$$\qquad \qquad \beta_k^{(w+1)} \in \operatorname{argmin}_{\beta_k \in \mathbb{R}^q} \, R_{n,k}^{(w)}(\beta_k)$$

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

- ▶ L-BFGS-B to solve the problem
- Proximal gradient method to estimate  $\gamma_k^{(w+1)}$

Introduction

Overview

Use case

Model

Notations

Inference

QNEM

Conclusion

## Quasi-Newton M-step

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

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

$$P_{n,k}^{(w)}(\beta_k) = \\ -n^{-1} \sum_{i=1}^n \hat{\pi}_{ik}^{\theta^{(w)}} \left[ (y_i \odot \Phi_i^{(w)})^\top \hat{\mathbb{E}}_{\theta^{(w)}}[M_{ik}|t_i, \delta_i, y_i] - \hat{\mathbb{E}}_{\theta^{(w)}}[c_{\phi^{(w)}}(M_{ik})|t_i, \delta_i, y_i] \right]$$

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

- ▶ L-BFGS-B to solve the problem
- Proximal gradient method to estimate  $\gamma_k^{(w+1)}$
- Predictive marker  $\hat{\mathcal{R}}_{ik} = \frac{\pi_{\hat{\xi}_k}(x_i)\hat{f}(t_i^{max},y_i|G_i=k,\Psi_i;\hat{\theta})}{\sum_{k=0}^{K-1}\pi_{\hat{\xi}_k}(x_i)\hat{f}(t_i^{max},y_i|G_i=k,\Psi_i;\hat{\theta})}$ , which is an estimate of  $\mathbb{P}_{\theta}[G_i=k|T_i^{\star}>t_i^{max},y_i]$

Introduction

Overviev

Use case:

Vlodel

Notations

Penalization
ONEM

Evaluation

Conclusion

# V. Evaluation

#### Introduction

Overview

Framewo

### Model

Submode

Notation

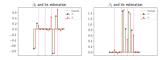
### Inference

Penalization

### Evaluation

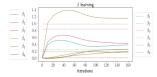
### Conclusion

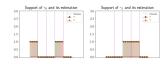
# **Experiments**

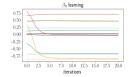












### Introduction

Overview Use cases

### Submodels

Notations Likelihood

### nference

enalization

# Evaluation

### Conclusion

## VI. Conclusion

### Introduction

Overview

Framewo

### Model

Submode

Likelihoo

### Inference

Penalization ONFM

Evaluation

### 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

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

ntroduction

Use case:

Model

Notation

Inference

QNEM

Evaluation

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 (QNEM) has been derived

ntroductio

Use case:

iviodei

Notations Likelihood

Populization

QNEM

Evaluation

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

ntroductio

Use cases

iviodei

Notations Likelihood

Populization

QNEM

Evaluation

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

## Python 3 package

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

ntroductio

Use case:

Model

Notations

nterence

QNEM

Evaluation

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

## Python 3 package

- Available at https://github.com/Califrais/lights
- Applications of the model available soon on an arXiv paper.

ntroductio

Use cases

iviodei

Notations

nterence

QNEM

Lvaldation

Conclusion

# Thank you!

### Introduction

Overview

Framewo

### Model

Submodel Notations

Likelihoo

### Interence

Penalization

Evaluation

### Conclusion

# References

#### Introduction

Overv

Use cas

#### Mode

Submod

Notation

### Inference

Penalizatio ONEM

Evaluation

Conclusion