Modelling Functional Disability with Hawkes Process

Jiwon Jung

Department of Statistics
Purdue University
jung320@purdue.edu

January 26, 2023

Overview

1. Introduction
2. Three-state health transition model
3. Maximum likelihood estimation
4. Results
5. Conclusion

Introduction

Introduction

- Understanding individual health transitions such as functional disability, recovery and mortality is essential to product pricing and development.
- Prior literature usually assumes Markov property for modelling health transitions, for which the probabilities of transition at each age depend on the current status only (see e.g., Fong et al., 2015; Li et al., 2017; Sherris and Wei, 2021).
- Showing that the probabilities of functional status transitions are duration dependent, other literature (Cai et al., 2006; Biessy, 2017) assumes semi-Markov process model to incorporate not only age and the current status but also on the duration in the current state.
- However, the state and duration effect with respect to the past functional disability experience has been less studied.

Motivation

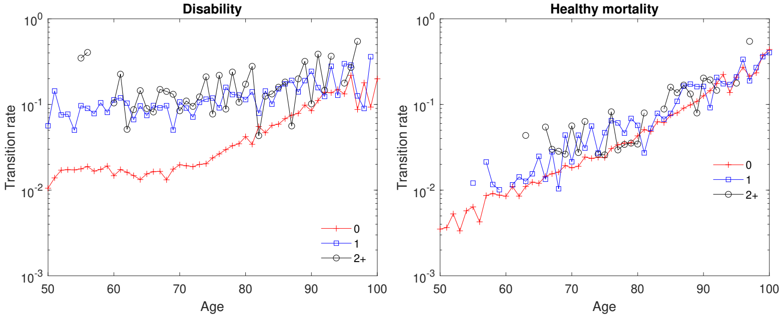

Figure 1. Crude health transition rates of functional disability and healthy mortality. The legend indicates the number of past functional-disabilities during the investigation period.

- Our explanatory data analysis suggest that the elderly who have experienced functional disability have a higher chance of functional disability and mortality than those who were never disabled before.

Hawkes process

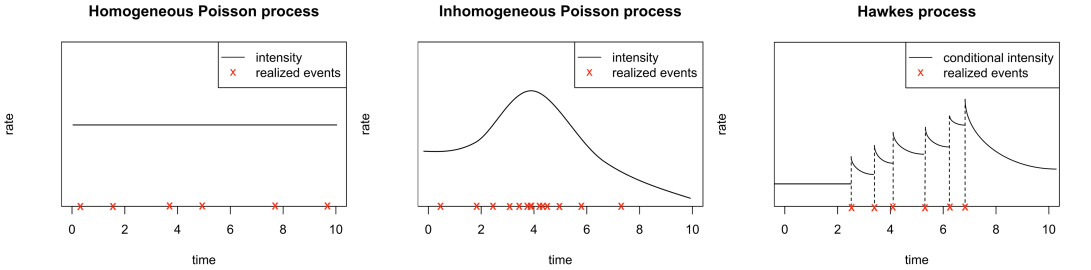

Figure 2. Graphs of a homogenous Poisson process, inhomogeneous Poisson process, and Hawkes process on the real line.

- Hawkes process is a self–exciting point process, in which the occurrence of an event increases the probability of occurrence of another event.

Goal

- Our goal is to estimate the intensity of age and gender-specific transitions by incorporating the impact of past functional disability as well as time spent in the current state using Hawkes process.

Three-state health transition model

Data preparation

- We use the `RAND HRS Data` 1992-2018 from the U.S. Health and Retirement Study (HRS), a nationally representative longitudinal panel survey.[1]
- The HRS is a biennial survey which began in 1992 and follows up with interviews of initially non-institutionalised Americans aged 50 and above.
- The health state is determined by a person's ability to perform activities of daily living (ADLs), such as bathing, toileting, and dressing; An individual needing help in two or more ADLs is functionally disabled.

[1] `https://hrs.isr.umich.edu/data-products`

Hawkes transition intensity

The transition intensity for individual k of transition type s at time t is given by

$$\lambda_s(t) = \underbrace{\phi_s(t)}_{\text{background intensity}} + \underbrace{\mu_s(t - T_t)}_{\text{exciting function}} \cdot \underbrace{\mathbb{1}_F(t)}_{\text{disability indicator}}, \tag{1}$$

where T_t is the latest transition time before time t.

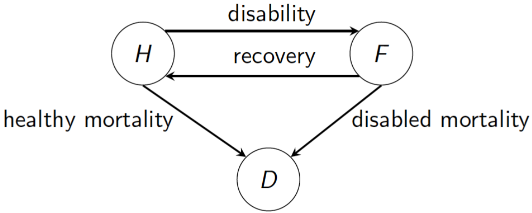

Figure 3. The three-state health transition model. H means healthy; F means functionally disabled; D means dead.

Hawkes transition intensity

- $\phi_s(t)$ is the Gompertz background intensity that captures the impact of observable variates such as (scaled) age $x_k(t)$ and gender indicator F_k at time t

$$\phi_s(t) = \exp(\beta_s^{intercept} + \beta_s^{age} x_k(t) + \beta_s^{female} F_k) \tag{2}$$

- $\mu_s(\cdot)$ is the exciting kernel function that captures the impact of past functional disability and duration in the current state

$$\mu_s(\tau) = \alpha_s \exp\left(-\delta_s \tau\right) \qquad \text{(exponential decay)} \tag{3}$$

- $\mathbb{1}_F(t)$ is the indicator of past functional disability experience

$$\mathbb{1}_F(t) = \begin{cases} 1 & \text{if functionally disabled at least once before time t} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Maximum likelihood estimation

Estimation under left truncation and censoring

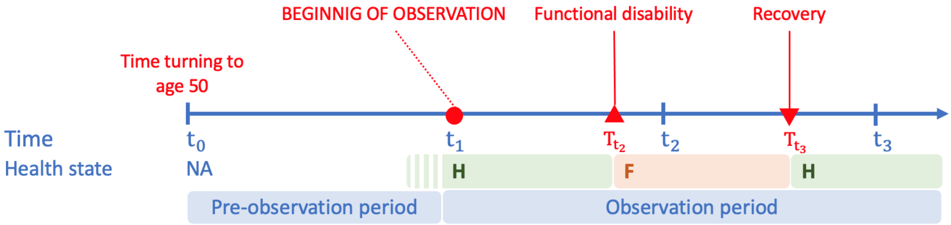

Figure 4. Example of an individual's health transition history. t_j is the time for j^{th} interview.

When an individual was first observed after the age of 50 and he/she was not in the functionally disabled state, we have

1. unknown $\mathbb{1}_F(t_1)$: presence of past functional disability
2. unobserved T_{t_1}: the latest transition time before the first interview (if any)

EM-algorithm

We use the EM algorithm to perform the estimation.

General scheme of EM-algorithm

1. Initialize $\theta^{(1)}$
2. For $i = 1, 2, 3, \ldots,$ Iterate E-step and M-step until convergence
   2.1 **E-step:** Compute the conditional expectation of the log likelihood function given the current estimate
   $$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = \mathbb{E}_{X_{unobserved}|X_{observed},\boldsymbol{\theta}^{(i)}}[l(\boldsymbol{\theta})] \tag{5}$$
   2.2 **M-step:** Update $\boldsymbol{\theta}$
   $$\boldsymbol{\theta}^{(i+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) \tag{6}$$

EM-algorithm

EM-algorithm for Hawkes health transition model:

EM-algorithm for Hawkes process

1. **Initialize** $\theta^{(1)}$: We initialize the parameters assuming no truncation.
2. For $i = 1, 2, 3, \ldots,$ Iterate E-step and M-step until convergence[a]
   - 2.1 **E-step:** Since analytical solution is unavailable, we perform Monte Carlo approximation to obtain the Q value:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = \mathbb{E}_{\mathbb{1}_F, \tau_{trunc}|data, \boldsymbol{\theta}^{(i)}}\left[l(\boldsymbol{\theta})\right] = \mathbb{E}_{\mathbb{1}_F|data, \boldsymbol{\theta}^{(i)}}\left[\mathbb{E}_{\tau_{trunc}|\mathbb{1}_F, data, \boldsymbol{\theta}^{(i)}}\left[l(\boldsymbol{\theta})\right]\right] \tag{7}$$

     We use 10,000 simulated individual's health transition history sampled from $\boldsymbol{\theta}^{(i)}$.
   - 2.2 **M-step:** We use numerical optimization algorithm to obtain the next estimates[b].

[a] Iterate until the difference between the current and previous Q value is less than 10^{-2}

[b] We use `optim` function in `R`

Results

Estimation results

Model	p^*	LRT statistic (df)†	AIC	BIC
Static model (non-Hawkes)	12	-	169,437.66	169,533.87
Single Hawkes: disability	14	2,020.3***(2)	167,421.31	167,533.56
Single Hawkes: recovery	14	213.3***(2)	169,228.31	169,340.56
Single Hawkes: healthy mortality	14	46.8***(2)	169,394.83	169,507.08
Single Hawkes: disabled mortality	14	48.5***(2)	169,393.13	169,505.37
Full Hawkes: all four-transition	20	$\geq$ **336.4***(6)**	**167,096.90**	**167,257.26**

* number of parameters

† Static v. Single Hawkes; Single Hawkes v. Full Hawkes

Table 1. Likelihood ratio test, AIC, and BIC statistics of health transition models.

Estimation results

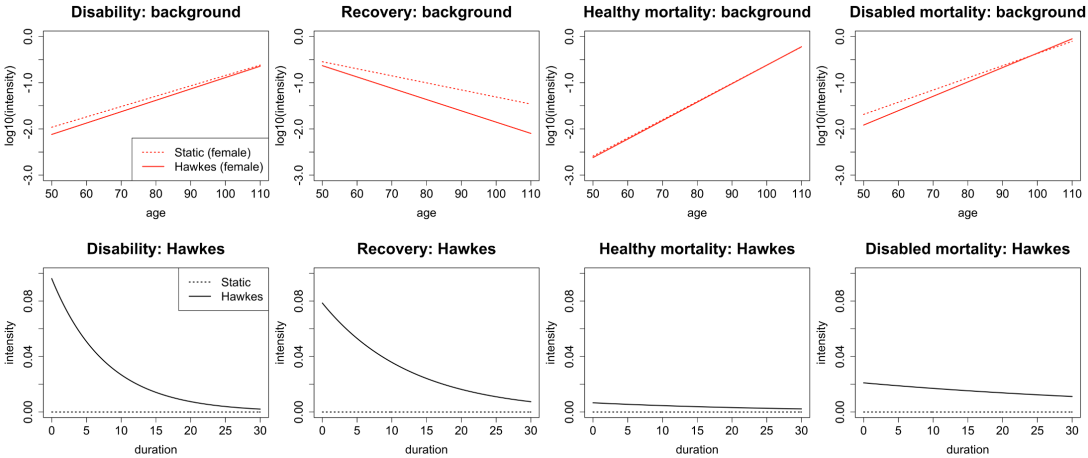

Figure 5. Maximum likelihood estimates of static and Hawkes model parameters. Hawkes parameters are estimated using EM algorithm.

Future lifetime statistics

	Healthy at 65				Healthy at 75			
	Female		Male		Female		Male	
	Static	Hawkes	Static	Hawkes	Static	Hawkes	Static	Hawkes
Healthy proportion	0.84	0.84	0.82	0.82	0.63	0.61	0.57	0.56
Total future lifetime	19.06	19.31	16.36	16.68	12.43	12.58	10.31	10.47
(SE)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
SD	9.06	8.93	8.37	8.33	7.01	6.94	6.18	6.22
Healthy future lifetime	16.21	15.95	14.71	14.73	10.22	10.08	9.09	9.12
(SE)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
SD	8.55	8.67	8.08	8.21	6.51	6.54	5.90	6.03
Disabled future lifetime	2.85	3.36	1.65	1.96	2.21	2.49	1.23	1.35
(SE)	(0.03)	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
SD	4.09	5.04	3.04	3.72	3.39	3.95	2.43	2.74
Healthy over total future lifetime	0.86	0.84	0.90	0.89	0.83	0.82	0.89	0.88
(SE)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
SD	0.20	0.23	0.18	0.20	0.24	0.26	0.21	0.23
Age at onset of disability‡	77.89	78.58	77.03	77.36	83.52	83.80	82.63	82.95
(SE)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
SD	8.39	8.66	7.70	8.07	6.17	6.31	5.54	5.79

‡ Age at onset of disability conditional on becoming disabled after the age of 65 and 75

Table 2. Average future lifetime statistics for 50-year-old healthy individuals who are healthy at 65 and 75, including standard error of the mean in brackets and SD. The maximum attainable age is 110.

Conclusion

Conclusion

- We proposed a multi-state health transition model with Hawkes process which can be numerically estimated via EM algorithm.
- We showed that our Hawkes model successfully incorporated the impact of past functional disability as well as time spent in the current state.
- In particular, our estimation results suggest that the functional disability and mortality intensities significantly increase on the onset of the disability and decay as the duration since the latest transition gets longer.
- Our future work will extend our choice of Hawkes kernels to allow for delaying effect in the speed of decay. (e.g., Rayleigh kernel).

Reference

- Biessy, G. (2017) Continuous-time semi-markov inference of biometric laws associated with a long-term care insurance portfolio. *ASTIN Bulletin: The Journal of the IAA*, **47**(2), 527–561.
- Cai, L., Schenker, N. and Lubitz, J. (2006) Analysis of functional status transitions by using a semi-markov process model in the presence of left-censored spells. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **55**(4), 477–491.
- Fong, J. H., Shao, A. W. and Sherris, M. (2015) Multistate actuarial models of functional disability. *North American Actuarial Journal*, **19**(1), 41–59.
- Li, Z., Shao, A. W. and Sherris, M. (2017) The impact of systematic trend and uncertainty on mortality and disability in a multistate latent factor model for transition rates. *North American Actuarial Journal*, **21**(4), 594–610.
- Sherris, M. and Wei, P. (2021) A multi-state model of functional disability and health status in the presence of systematic trend and uncertainty. *North American Actuarial Journal*, **25**(1), 17–39.

Appendix

Conditional intensity

Definition

(Conditional intensity function) Consider a counting process $N(t)$ with associated histories $\mathcal{F}^N(t), t \geq 0$. If a non-negative function $\lambda(t)$ exists such that

$$\lambda(t) = \lim_{h \to 0} \frac{\mathbb{E}\left[N(t+h) - N(t) | \mathcal{F}^N(t)\right]}{h}, \tag{8}$$

then it is called the conditional intensity function of $N(t)$. Note that this function is also called the hazard function.

Hawkes process

Definition

(Hawkes process) The one-dimensional Hawkes process is a point process $N(t)$ which is characterized by its conditional intensity $\lambda(t)$ with respect to its natural filtration:

$$\lambda(t) = \phi(t) + \int_0^t \mu(t-s)\mathrm{d}N(s), \tag{9}$$

where $\phi(t)$ is the background intensity function, and the $\mu(t)$ is the excitation function satisfying $\int_0^\infty \mu(s)\mathrm{d}s < 1$.

EM-algorithm

E-step (evaluation):

In the presence of left-truncation, we can split the log likelihood terms into two terms with/without truncation;

$$\log L\left(\boldsymbol{\theta} | \{\mathbb{1}_f, \tau_{trunc}\}\right) = \sum_{k=1}^{K} \sum_{s=1}^{S} \left(\sum_{j: t_j < \inf_i \hat{t}_i} l_{k,s,j}\left(\boldsymbol{\theta}\right) \sum_{j: t_j \geq \inf_i \hat{t}_i} l_{k,s,j}\left(\boldsymbol{\theta}\right) \right) \quad (10)$$

Recall that $\inf_i \hat{t}_i$ denote the first transition time.

Complete likelihood function

Suppose there are a total of K individuals, S transition types, and J interview waves. The complete log likelihood function is given by

$$l(\boldsymbol{\theta}) = \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{j=1}^{J-1} l_{k,s,j}(\boldsymbol{\theta}), \tag{11}$$

where $\boldsymbol{\theta}$ denotes the set of parameters to be estimated, and

$$\begin{aligned} l_{k,s,j}(\boldsymbol{\theta}) = Y_{k,s,j} \ln \lambda_{k,s}(\hat{t}_{k,j}) - R_{k,s}(t_{k,j}) \int_{t_{k,j}}^{\min\{\hat{t}_{k,j}, t_{k,j+1}\}} \lambda_{k,s}(u)\mathrm{d}u \\ - R_{k,s}(\hat{t}_{k,j}) \int_{\min\{\hat{t}_{k,j}, t_{k,j+1}\}}^{t_{k,j+1}} \lambda_{k,s}(u)\mathrm{d}u, \end{aligned}$$

Here, we introduce two indicator variables: (1) $Y_{k,s,j} = 1$ if transition type s is observed between the j^{th} and $(j+1)^{\text{th}}$ interviews, and (2) $R_{k,s}(t) = 1$ if the individual is exposed to the risk of transition type s at time t.

EM-algorithm

Now, we can separately calculate the conditional expectation to obtain $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)})$.

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = Q_{trunc}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) + Q_{full}(\boldsymbol{\theta}) \tag{12}$$

i. Q_{trunc}**:**

$$Q_{trunc}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = \sum_{k=1}^{K}\sum_{s=1}^{S}\sum_{j:t_j<\inf_i \hat{t}_i} \mathbb{E}_{\mathbb{1}_f|\boldsymbol{\theta}^{(i)}}\left[\mathbb{E}_{\tau_{trunc}|\mathbb{1}_f,\boldsymbol{\theta}^{(i)}}\left[l_{k,s,j}(\boldsymbol{\theta})\right]\right]$$

$$= \sum_{k=1}^{K}\sum_{s=1}^{S}\sum_{j:t_j<\inf_i \hat{t}_i}\sum_{m=0}^{1} P(\mathbb{1}_f(t_j) = m|\boldsymbol{\theta}^{(i)})\mathbb{E}\left[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_j} \geq \tilde{\tau}_{t_j}, \mathbb{1}_f(t_j) = m, \boldsymbol{\theta}^{(i)}\right]$$

EM-algorithm

Probability of being functionally disabled at least once before t

Note that

$$P(\mathbb{1}_f(t) = 0 | Z_t, \boldsymbol{\theta}^{(i)}) = \begin{cases} \exp\left(-\int_0^t \phi_1(u) + \phi_3(u) du\right) & \text{, if } Z_t = H \\ 0 & \text{, if } Z_t = F \end{cases}$$

Accordingly, $P(\mathbb{1}_f(t) = 1 | Z_t, \boldsymbol{\theta}^{(i)}) = 1 - P(\mathbb{1}_f(t) = 0 | Z_t, \boldsymbol{\theta}^{(i)})$

EM-algorithm

Expectation under truncation

$$\mathbb{E}\left[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_j} \geq \tilde{\tau}_{t_j}, \mathbb{1}_f(t_j) = 0, \boldsymbol{\theta}^{(i)}\right] = Y_{k,s,j} \log \phi_{k,s}(\hat{t}_j) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u)\mathrm{d}u$$
$$- R_{k,s}(\hat{t}_j) \int_{\min\{\hat{t}_j, t_{j+1}\}}^{t_{j+1}} \phi_{k,s}(u)\mathrm{d}u,$$

$$\begin{aligned} &\mathbb{E}[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_1} \geq \tilde{\tau}_{t_1}, \mathbb{1}_f(t_j) = 1, \boldsymbol{\theta}^{(i)}] \\ &= \mathbb{E}\Bigg[Y_{k,s,j} \log\left(\phi_{k,s}(\hat{t}_j) + \mu_s(\tau_{\hat{t}_j})\right) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u) + \mu_s(\tau_u)\mathrm{d}u \\ &\quad - R_{k,s}(\hat{t}_j) \int_{\min\{\hat{t}_j, t_{j+1}\}}^{t_{j+1}} \phi_{k,s}(u) + \mu_s(\tau_u)\mathrm{d}u | \tau_{t_j} \geq \tilde{\tau}_{t_j}, \mathbb{1}_f(t_j) = 1, \boldsymbol{\theta}^{(i)} \Bigg], \end{aligned} \tag{13}$$

EM-algorithm

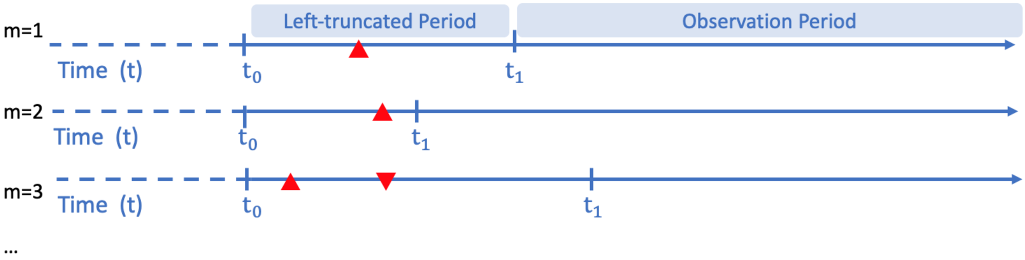

Figure 6. An example of an individual's simulated health transition history.

EM-algorithm

To approximate (13), we use Monte Carlo simulation.

1. Using time-rescaling theorem and Ogata's thinning method, obtain random samples from $T_{t_1}^{(n)} | \mathbb{1}_f(t_1) = 1, \boldsymbol{\theta}^{(i)}$ until we have N number of samples such that $T_{t_1}^{(n)} \leq \tilde{T}_{t_1}$. Note that $\tau_{t_1}^{(n)} = t - T_{t_1}^{(n)}$.

2. Approximate the conditional expectation under truncation

$$(13) \approx \frac{\sum_{n=1}^{N} Y_{k,s,j} \log\left(\phi_{k,s}(\hat{t}_j) + \mu_s(\tau_{\hat{t}_j}^{(n)})\right) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u) + \mu_s(\tau_u^{(n)}) \mathrm{d}u}{N}$$

EM-algorithm

ii. Q_{full}**:**

$$Q_{full}(\boldsymbol{\theta}) = \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{j: t_j \geq \inf_i \hat{t}_i} l_{k,s,j}(\boldsymbol{\theta}) \tag{14}$$

This is the case after the first disability/recovery time (if any). Note that after the first disability/recovery transition, we can remove the expectation since $\mathbb{1}_f(t) = 1$ and $\tau_{t_j} = \tilde{\tau}_{t_j}$ with probability 1.

EM convergence: background intensity parameters

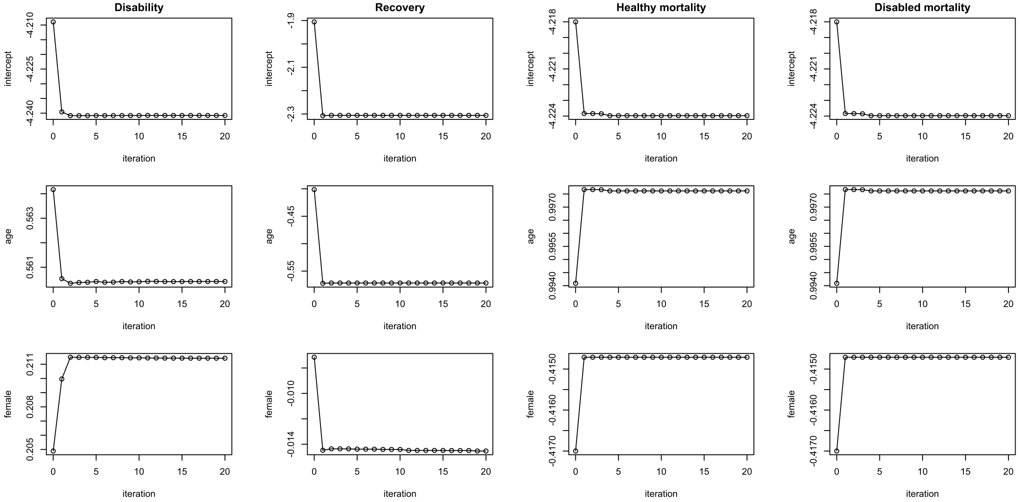

Figure 7. Background intensity coefficients updated by EM algorithm.

EM convergence: Hawkes kernel parameters

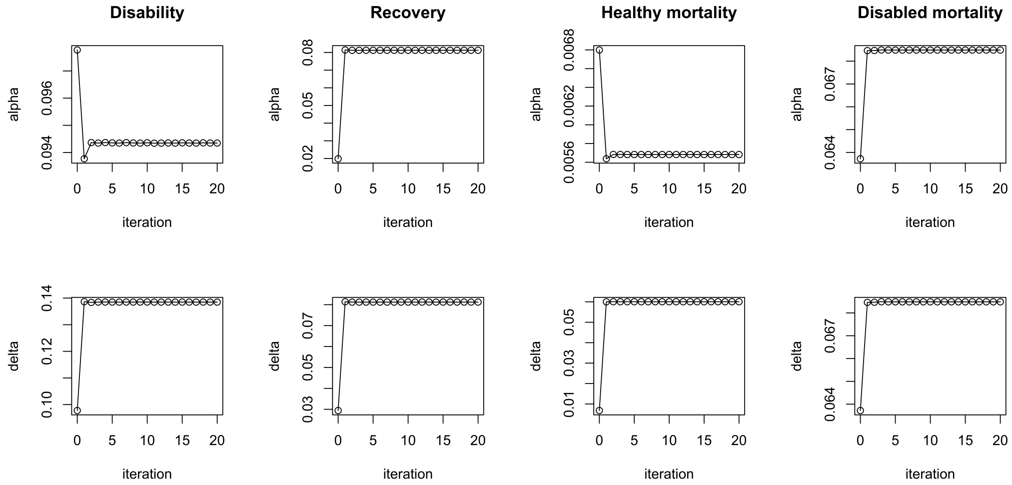

Figure 8. Hawkes kernel coefficients updated by EM algorithm.