Hands On Continuous Time Modeling With ctsem

Leonard Kosanke & Valentin Kriegmair

Outline

- Recap
- Introduction to ctsem
- Walk-through
- Excercise
- Solution

Recap I

What are the assumptions about time in the general cross lagged panel model (GCLPM)?

1. time prgoresses in discrete steps
2. equally spaced measurement occasions
3. subjects are measured at the same time points

Recap II

What are limitations of treating time as a discrete variable when modeling longitudinal data?

- biased parameter estimates due to:
  - unequally spaced measurement waves in the panel data
  - measurement of individual subjects at each wave spread out over time
- uncomparable parameter estimates across studies

Recap II

Multivariate stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}_{\mathrm{i}}(\mathrm{t}) = (\mathbf{A}\boldsymbol{\eta}_{\mathrm{i}}(\mathrm{t}) + \mathbf{B}\mathbf{z}_{\mathrm{i}} + \mathbf{M}\boldsymbol{\chi}_{\mathrm{i}}(\mathrm{t}))\,\mathrm{dt} + \mathbf{G}\mathrm{d}\mathbf{W}_{\mathrm{i}}(\mathrm{t})$$

Recap II

Multivariate stochastic equation:

$$d\boldsymbol{\eta}_i(t) = (\mathbf{A}\boldsymbol{\eta}_i(t) + \mathbf{B}\mathbf{z}_i + \mathbf{M}\boldsymbol{\chi}_i(t))\,dt + \mathbf{G}d\mathbf{W}_i(t)$$

modelling change in latent variables $\boldsymbol{\eta}_i(t)$ as a function of time over infinitesimal small time steps dt

Recap II

Multivariate stochastic differential equation:

$$d\boldsymbol{\eta}_i(t) = (\mathbf{A}\boldsymbol{\eta}_i(t) + \mathbf{B}\mathbf{z}_i + \mathbf{M}\boldsymbol{\chi}_i(t))\,dt + \mathbf{G}\,d\mathbf{W}_i(t)$$

modelling change as a function of time over infinitesimal small time steps dt

A: drift matrix with auto effects on the diagonal and cross effects on the off-diagonals ($\mathbf{A} \in \mathbb{R}^{v \times v}$).

Recap II

Multivariate stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}_{\mathrm{i}}(\mathrm{t}) = (\mathbf{A}\boldsymbol{\eta}_{\mathrm{i}}(\mathrm{t}) + \mathbf{B}\mathbf{z}_{\mathrm{i}} + \mathbf{M}\boldsymbol{\chi}_{\mathrm{i}}(\mathrm{t}))\,\mathrm{dt} + \mathbf{G}\mathrm{d}\mathbf{W}_{\mathrm{i}}(\mathrm{t})$$

modelling change as a function of time over infinitesimal small time steps dt

$\mathbf{A}$: drift matrix with auto effects on the diagonal and cross effects on the off-diagonals ($\mathbf{A} \in \mathbb{R}^{\mathrm{v} \times \mathrm{v}}$).

$\mathbf{z}_{\mathbf{i}}$: time independent predictors

$\boldsymbol{\chi}_{\mathrm{i}}(\mathrm{t})$: time dependent predictors

$\mathbf{W}_{\mathrm{i}}(\mathrm{t})$: stochastic error term (Wiener process)

$\mathbf{G}$: diffusion matrix

Recap II

Multivariate stochastic differential equation:

$$d\boldsymbol{\eta}_i(t) = (\mathbf{A}\boldsymbol{\eta}_i(t) + \mathbf{B}\mathbf{z}_i + \mathbf{M}\boldsymbol{\chi}_i(t))\,dt + \mathbf{G}d\mathbf{W}_i(t)$$

measurement part:

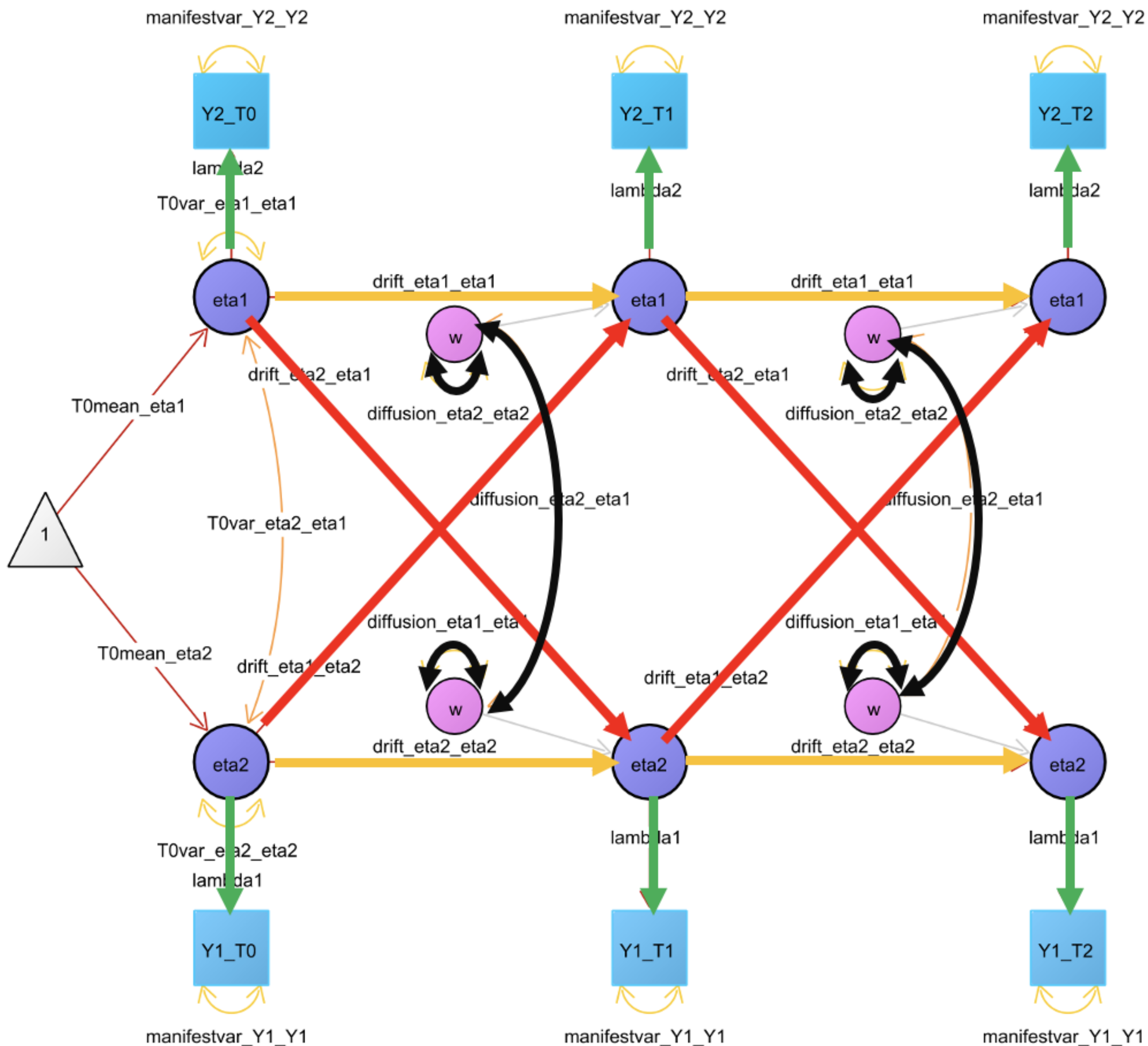
$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\varepsilon}(t)$$

$$\boldsymbol{\varepsilon}(t) \sim N(\mathbf{0}, \boldsymbol{\Theta})$$

$\boldsymbol{\Lambda}$: factor loadings

$\boldsymbol{\tau}$: intercepts

$\boldsymbol{\varepsilon}(t)$: measurement error



Bivariate Process (no Factor Structure)

stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}_i(t) = \left(\mathbf{A}\boldsymbol{\eta}_i(t) + \mathbf{B}\mathbf{z}_i + \mathbf{M}\boldsymbol{\chi}_i(t)\right)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}_i(t)$$

Measurement:

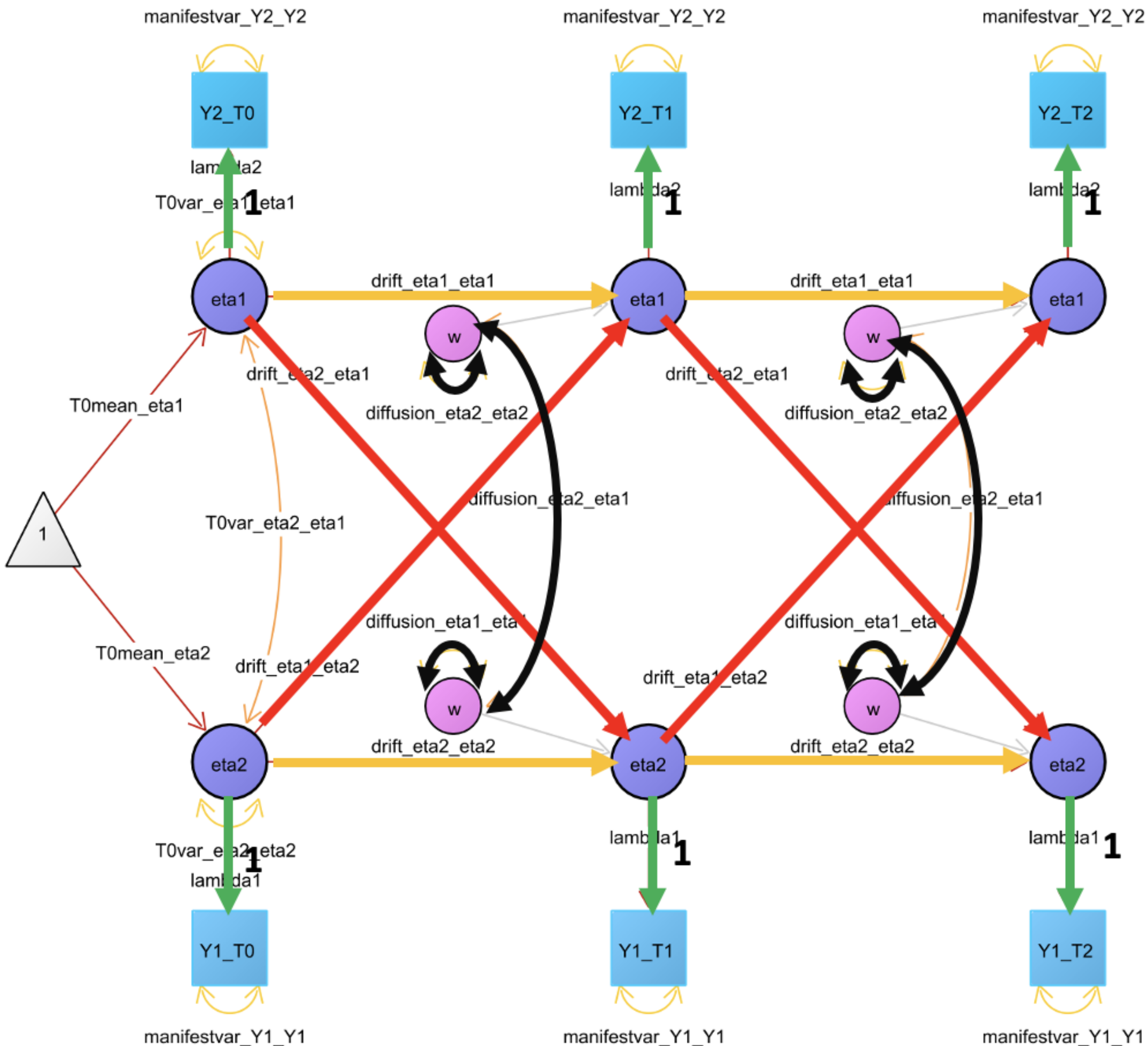
$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\varepsilon}(t) \qquad \boldsymbol{\varepsilon}(t) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Theta})$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$ Drift Matrix

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{11} & 0 \\ 0 & \lambda_{22} \end{pmatrix}$$ Loading Matrix

$$\mathbf{G} = \begin{pmatrix} g_{11} & 0 \\ g_{21} & g_{22} \end{pmatrix}$$ Diffusion Matrix

Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, *77*, 1-35.



Bivariate Process (no Factor Structure)

stochastic differential equation:

$$\mathrm{d}\boldsymbol{\eta}_i(t) = \left(\mathbf{A}\boldsymbol{\eta}_i(t) + \mathbf{B}\mathbf{z}_i + \mathbf{M}\boldsymbol{\chi}_i(t)\right)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}_i(t)$$

Measurement:

$$\mathbf{y}(t) = \boldsymbol{\Lambda}\boldsymbol{\eta}(t) + \boldsymbol{\tau} + \boldsymbol{\epsilon}(t) \qquad \boldsymbol{\epsilon}(t) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Theta})$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$ Drift Matrix

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ Loading Matrix

$$\mathbf{G} = \begin{pmatrix} g_{11} & 0 \\ g_{21} & g_{22} \end{pmatrix}$$ Diffusion Matrix

Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, *77*, 1-35.

Introducing `ctsem` & `ctsemOMX`

- R-packages `ctsem` & `ctsemOMX` to handle the math
- `ctsem` interfaces with Stan
- `ctsemOMX` interfaces with OpenMx
- today we will use `ctsemOMX`

Data Structure

`ctsemOMX` uses wide data with time intervals specified in columns:

```
  Quality_T0 Quality_T1 Quality_T2 dT1 dT2 dT3
1  -8.016139  -8.940450  -8.243518   1   1   1
2  -6.367648  -6.762183  -6.731700   1   1   1
3  -7.927406  -8.518258  -8.907190   1   1   1
4  -5.887097  -7.304229  -5.552340   1   1   1
5  -5.415876 -10.365073  -8.733621   1   1   1
6  -6.408571  -7.434263  -4.577385   1   1   1
```

(time points can also be interindividually varying)

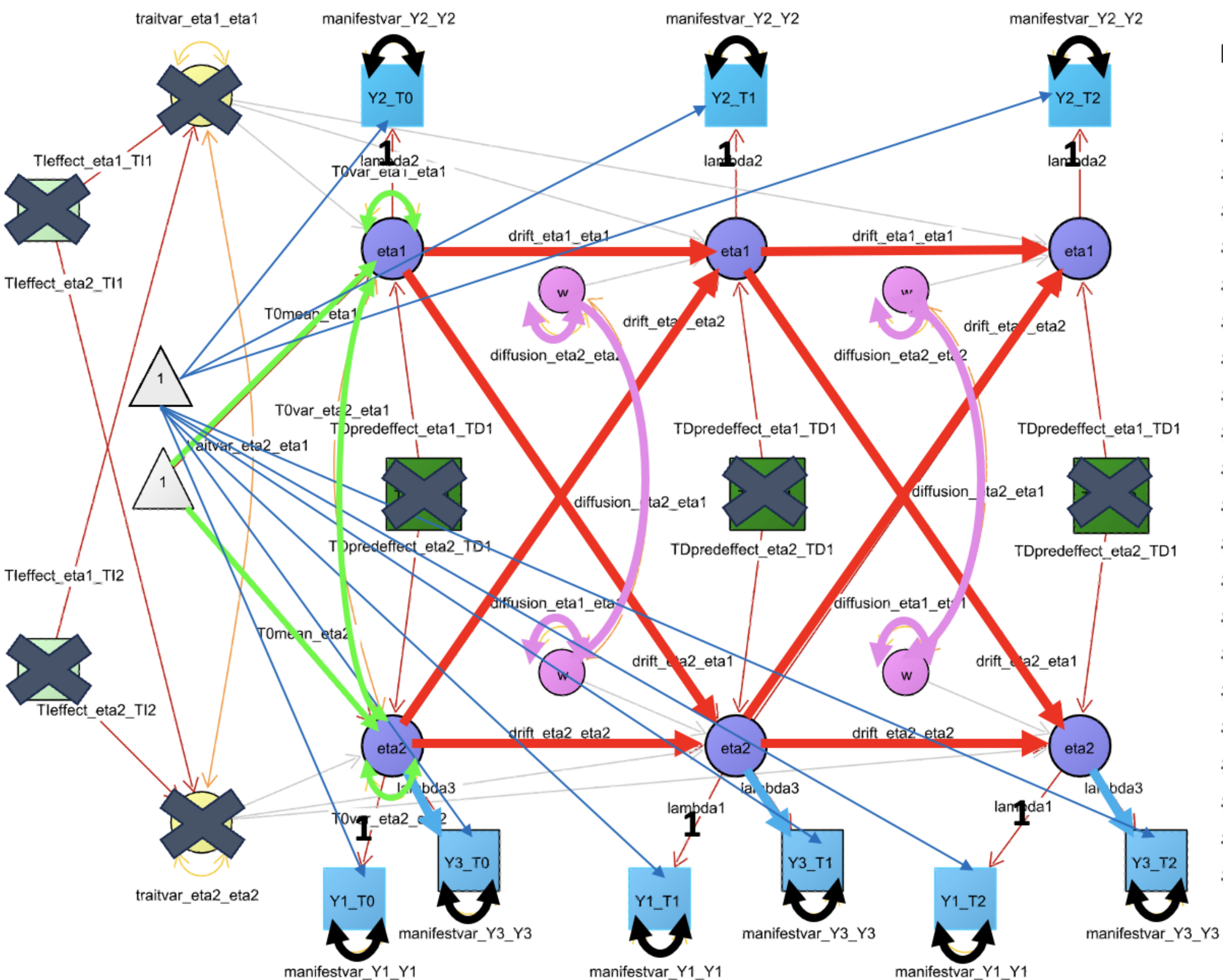
Function to convert to wideformat:

```r
df_wide <- ctLongToWide(datalong = df_long, id = "id",
  time = "time", manifestNames = c("Quality", ...))
```

Function to create time intervals:

```r
df_wide <- ctIntervalise(datawide = wideexample, Tpoints = 3, n
  manifest = 2, manifestNames = c("Quality", ...))
```

Next: Model set-up and interpretation walkthrough



model from exercise 6

```
## $ctparameters
##                        Value          Matrix
## stressLoading    0.002977787          LAMBDA
## T0m_eta1        -3.695248816         T0MEANS
## T0m_eta2       -19.950426226         T0MEANS
## drift_eta1      -0.304066770           DRIFT
## drift_eta2_eta1 -0.227847804           DRIFT
## drift_eta1_eta2  0.186420890           DRIFT
## drift_eta2       0.085819017           DRIFT
## mvarStress       0.229760390     MANIFESTVAR
## mvarStress2      0.269595124     MANIFESTVAR
## mvarQuality      0.301769320     MANIFESTVAR
## diff_eta1        0.792245329       DIFFUSION
## diff_eta2_eta1   0.908630806       DIFFUSION
## diff_eta2        1.194224737       DIFFUSION
## T0var_eta1       1.900422851           T0VAR
## T0var_eta2_eta1  2.455585775           T0VAR
## T0var_eta2       3.488468973           T0VAR
## mm_Stress       11.182382669 MANIFESTMEANS
## mm_Stress2       0.034551688 MANIFESTMEANS
## mm_Quality      13.212576945 MANIFESTMEANS
```

Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, *77*, 1-35.