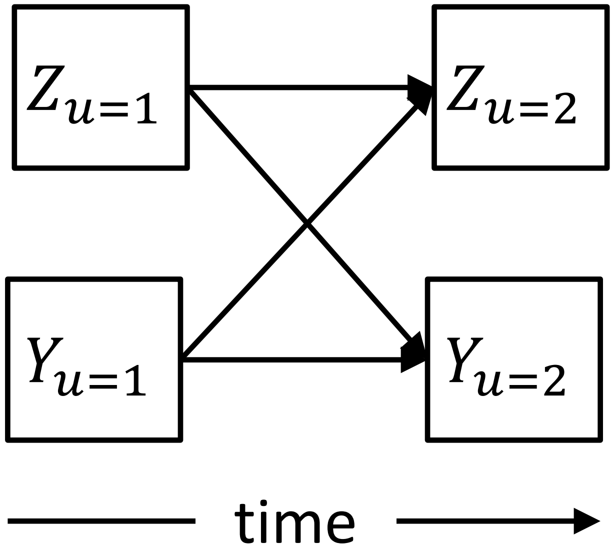
An introduction to continuous time dynamic modeling part 2

(…a closer look and some terminology)

Continuous time models: A closer look and some terminology

Dynamic models for the analysis of change:



$$\mathbf{x}_{t_u} = \mathbf{A}\,\mathbf{x}_{t_{u-1}} + \mathbf{w}_{t_u}$$

Continuous time models: A closer look and some terminology

…here it is!

$$\mathbf{x}_{t_u} = \mathbf{A}\,\mathbf{x}_{t_{u-1}} + \mathbf{w}_{t_u}$$

$$\mathbf{x}_{t_u} - \mathbf{x}_{t_{u-1}} = (\mathbf{A} - \mathbf{I})\mathbf{x}_{t_{u-1}} + \mathbf{w}_{t_u}$$

$$\frac{\Delta \mathbf{x}_{t_u}}{\Delta time} = (\mathbf{A} - \mathbf{I})\;\mathbf{x}_{t_{u-1}} + \mathbf{w}^*_{t_u}$$

Continuous time models: A closer look and some terminology

0. From differences to differentials

$$\frac{\Delta \mathbf{x}_{t_u}}{\Delta time} = (\mathbf{A} - \mathbf{I})\ \mathbf{x}_{t_{u-1}} + \mathbf{w}^*_{t_u} \qquad \lim_{\Delta time \to 0} \left(\frac{\Delta \mathbf{x}_{t_u}}{\Delta time} \right) = \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}time}$$

1. Latent dynamic model (extended Ornstein-Uhlenbeck process)

$$\mathrm{d}\boldsymbol{\eta}(t) = \big(\mathbf{A}\boldsymbol{\eta}(t) + \mathbf{b} + \mathbf{M}\boldsymbol{\chi}(t)\big)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}(t)\big)$$

including the measurement part

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

and

$$\boldsymbol{\chi}(t) = \sum_{u \in U} \mathbf{x}_u \delta(t - t_u)$$

with $\delta()$ denoting the Dirac delta function

Continuous time models: A closer look and some terminology

2. Discrete time solution of the stochastic differential equation

$$\boldsymbol{\eta}_u = \mathbf{A}^*_{\Delta t_u}\boldsymbol{\eta}_{u-1} + \mathbf{b}^*_{\Delta t_u} + \mathbf{M}\mathbf{x}_u + \boldsymbol{\zeta}_u \quad \text{with} \quad \boldsymbol{\zeta}_u \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}^*_{\Delta t_u})$$

and

$$\mathbf{A}^*_{\Delta t_u} = e^{\mathbf{A}(t_u - t_{u-1})}$$

$$\mathbf{b}^*_{\Delta t_u} = \mathbf{A}^{-1}\left(\mathbf{A}^*_{\Delta t_u} - \mathbf{I}\right)\mathbf{b} \quad \text{thus} \quad \mathbf{b}^*_{\Delta t_\infty} = -\mathbf{A}^{-1}\mathbf{b}$$

$$\mathbf{Q}^*_{\Delta t_u} = \mathbf{Q}_{\Delta t_\infty} - \mathbf{A}^*_{\Delta t_u}\mathbf{Q}_{\Delta t_\infty}(\mathbf{A}^*_{\Delta t_u})^{\mathrm{T}}$$

with $\mathbf{Q}_{\Delta t_\infty} = \mathrm{irow}(-(\mathbf{A}\otimes\mathbf{I} + \mathbf{I}\otimes\mathbf{A})^{-1}row(\mathbf{Q}))$

Continuous time models: A closer look and some terminology

Discrete time		Continuous time	
Parameter	Label	Parameter	Label
$\mathbf{A}^*_\Delta$	Autoregression matrix	$\mathbf{A}$	Drift matrix
$\mathbf{A}^*_\Delta[k, k]$	Autoregressive effect	$\mathbf{A}[k, k]$	Auto-effect
$\mathbf{A}^*_\Delta[k, l]$	Cross-lagged effect	$\mathbf{A}[k, l]$	Cross-effect
$\mathbf{Q}^*_\Delta$	Process error[a] matrix	$\mathbf{Q}$	Diffusion covariance matrix
$\mathbf{Q}^*_\Delta[k, k]$	Process error[a] variance	$\mathbf{Q}[k, k]$	Diffusion variance
$\mathbf{Q}^*_\Delta[k, l]$	Process error[a] covariance	$\mathbf{Q}[k, l]$	Diffusion covariance
$\mathbf{b}^*_\Delta$	Dt intercepts	$\mathbf{b}$	Ct intercepts
$\mathbf{\Sigma}^*_{\mathrm{b}\Delta}$	Dt intercepts covariance matrix	$\mathbf{\Sigma}_{\mathrm{b}}$	Ct intercepts covariance matrix
$\mathbf{\Sigma}^*_{\mathrm{b}\Delta}[k, k]$	Dt intercepts variance	$\mathbf{\Sigma}_{\mathrm{b}}[k, k]$	Ct intercepts variance
$\mathbf{\Sigma}^*_{\mathrm{b}\Delta}[k, l]$	Dt intercepts covariance	$\mathbf{\Sigma}_{\mathrm{b}}[k, l]$	Ct intercepts covariance

Note. $k \neq l$; Dt = discrete-time; Ct = continuous-time.
[a]Synonymously "prediction error."

Hecht & Voelkle (2019, p. 3).

Continuous time models: A closer look and some terminology

3. Unit level log likelihood

$$ll = \sum^{U}\left(-\frac{1}{2}(n\ln(2\pi) + \ln|\mathbf{V}_u| + \left(\hat{\mathbf{y}}_{u|u-1} - \mathbf{y}_u\right)\mathbf{V}_u^{-1}\left(\hat{\mathbf{y}}_{u|u-1} - \mathbf{y}_u\right)^{\mathrm{T}})\right)$$

n = number of non-missing observations at measurement occasion *u*.

Continuous time models: A closer look and some terminology

3b. Excursus: To aid estimation, a hybrid continuous-discrete Kalman filter is often used

$$\hat{\boldsymbol{\eta}}_{u|u-1} = \mathbf{A}^{*}_{\Delta t_u}\hat{\boldsymbol{\eta}}_{u-1\,|\,u-1} + \mathbf{b}^{*}_{\Delta t_u} + \mathbf{M}\mathbf{x}_u$$

$$\hat{\mathbf{P}}_{u|u-1} = \mathbf{A}^{*}_{\Delta t_u}\hat{\mathbf{P}}_{u-1\,|\,u-1}(\mathbf{A}^{*}_{\Delta t_u})^{\top} + \mathbf{Q}^{*}_{\Delta t_u}$$

with

$$\hat{\mathbf{y}}_{u|u-1} = \boldsymbol{\Lambda}\hat{\boldsymbol{\eta}}_{u|u-1} + \boldsymbol{\tau}$$

$$\hat{\mathbf{V}}_u = \boldsymbol{\Lambda}\hat{\mathbf{P}}_{u|u-1}\boldsymbol{\Lambda}^{\top} + \boldsymbol{\Theta}$$

$$\hat{\mathbf{K}}_u = \hat{\mathbf{P}}_{u|u-1}\boldsymbol{\Lambda}^{\top}\hat{\mathbf{V}}_u^{-1}$$

$$\hat{\boldsymbol{\eta}}_{u|u} = \hat{\boldsymbol{\eta}}_{u|u-1} + \hat{\mathbf{K}}_u(\mathbf{y}_u - \hat{\mathbf{y}}_{u|u-1})$$

$$\hat{\mathbf{P}}_{u|u} = (\mathbf{I} - \hat{\mathbf{K}}_u\boldsymbol{\Lambda})\hat{\mathbf{P}}_{u|u-1}$$

Driver & Voelkle (2018). *Psychological Methods.*

Continuous time models: A closer look and some terminology

4. Accounting for (unobserved) unit heterogeneity

Frequentist Approach

$$\mathrm{d}\boldsymbol{\eta}(t) = \big(\mathbf{A}\boldsymbol{\eta}(t) + \mathbf{b} + \boxed{\boldsymbol{\kappa}} + \mathbf{M}\boldsymbol{\chi}(t)\big)\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}(t)\big)$$

$$\boldsymbol{\eta}_u = \mathbf{A}^*_{\Delta t_u}\boldsymbol{\eta}_{u-1} + \mathbf{b}^*_{\Delta t_u} + \boxed{\mathbf{H}^*_{\Delta t_u}\boldsymbol{\kappa}} + \mathbf{M}\mathbf{x}_u + \boldsymbol{\zeta}_u$$

$$\boxed{\mathbf{H}^*_{\Delta t_u} = \mathbf{A}^{-1}\big(\mathbf{A}^*_{\Delta t_u} - \mathbf{I}\big)}$$

with

$$\boldsymbol{\Phi}^*_{\boldsymbol{\kappa}\Delta t_u} = \mathbf{H}^*_{\Delta t_u}\boldsymbol{\Phi}_{\boldsymbol{\kappa}}(\mathbf{H}^*_{\Delta t_u})^{\mathrm{T}}$$

$$\boldsymbol{\Phi}^*_{\boldsymbol{\eta}_{u=1},\boldsymbol{\kappa}\Delta t_u} = \boldsymbol{\Phi}_{\boldsymbol{\eta}_{u=1},\boldsymbol{\kappa}}(\mathbf{H}^*_{\Delta t_u})^{\mathrm{T}}$$

Bayesian Approach (fully hierarchical)

$$p(\boldsymbol{\Phi}, \boldsymbol{\mu}, \mathbf{R}, \boldsymbol{\beta} | \mathbf{Y}, \mathbf{z}) \propto p(\mathbf{Y}|\boldsymbol{\Phi})p(\boldsymbol{\Phi}|\boldsymbol{\mu}, \mathbf{R}, \boldsymbol{\beta}, \mathbf{z})p(\boldsymbol{\mu}, \mathbf{R}, \boldsymbol{\beta})$$

(joint posterior distribution)

with $\boldsymbol{\Phi}_i = \mathrm{tform}(\boldsymbol{\mu} + \mathbf{R}\mathbf{h}_i + \boldsymbol{\beta}\mathbf{z}_i)$

$$\mathbf{h}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{1})$$

$$\boldsymbol{\mu} \sim \mathrm{N}(\mathbf{0}, \mathbf{1})$$

$$\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, \mathbf{1})$$

Driver & Voelkle (2018). *Psychological Methods.*

ctsem & ctsemOMX

- To deal with the math behind continuous time modeling, we developed *ctsem.*

- Originally, *ctsem* was a single R package interfacing either to OpenMx or Stan.

- For compatibility reasons, *ctsem* was split in two packages in 2020. The main package *ctsem* (Stan) and OpenMx functions *ctsemOMX*. [loading ctsemOMX will always also load ctsem but not vice versa]

- Both versions are open source and freely available to everyone on CRAN.

Driver, Oud, & Voelkle (2017), *Journal of Statistical Software*

Ctsem: Dealing with unequal time intervals (panel data)

- To illustrate the approach, let us consider an empirical example using data collected within the research project *Group Focused Enmity* (Heitmeyer, 2004).

- N = 2,722 at t_1; computer assisted interviews in 2002, 2003, 2004, 2006, 2008 (U = 5)

- Focus on two constructs: *Anomia/Anomie* and *Authoritarianism*

- Research question: Does anomia cause authoritarianism (Merton, 1949; Srole, 1956) or is anomia caused by authoritarianism (McClosky & Schaar, 1965)?

Voelkle, Oud, Davidov, & Schmidt (2012). *Psychological Methods.*

Ctsem: Dealing with unequal time intervals (panel data)

<u>Preparation:</u> Installation of R and *ctsem* (*ctsemOMX*)

<u>Data:</u> Person-level data (wide format) with time intervals at the end.

t_1 (year 2002)

Missing value

1	2.67	3.50	3.33	3.50	NA	NA	NA	NA	NA	NA	1	1	2	2
2	3.33	3.25	NA	NA	NA	NA	NA	NA	NA	NA	1	1	2	2
3	3.33	2.75	3.33	3.00	3.33	2.50	2.33	3.00	2.33	3.00	1	1	2	2
4	3.33	3.25	NA	NA	NA	NA	NA	NA	NA	NA	1	1	2	2
5	4.00	4.00	NA	NA	NA	NA	NA	NA	NA	NA	1	1	2	2
6	3.67	4.00	NA	NA	NA	NA	4.00	4.00	4.00	4.00	1	1	2	2
7	2.33	3.50	3.33	3.25	3.33	3.25	2.00	3.25	1.67	3.25	1	1	2	2
8	3.00	2.25	NA	NA	NA	NA	NA	NA	NA	NA	1	1	2	2
9	2.33	4.00	3.00	4.00	3.33	3.75	2.67	3.75	3.00	3.50	1	1	2	2
10	4.00	3.25	3.33	3.25	4.00	3.50	4.00	3.50	4.00	3.25	1	1	2	2

Person (ID)

Time intervals (years)

Ctsem: Dealing with unequal time intervals (panel data)

<u>Input:</u> (1) ctModel, (2) ctFit, (3) summary & plot

ctModel {ctsem} R Documentation

Define a ctsem model

Description

This function is used to specify a continuous time structural equation model, which can then be fit to data with function ctStanFit.

Usage

```
ctModel(
  LAMBDA,
  type = "omx",
  n.manifest = "auto",
  n.latent = "auto",
  Tpoints = NULL,
  manifestNames = "auto",
  latentNames = "auto",
  id = "id",
  time = "time",
```

(1) ctModel

Ctsem: Dealing with unequal time intervals (panel data)

Input: (1) ctModel, (2) ctFit, (3) summary & plot

ctStanFit {ctsem} R Documentation

ctStanFit

Description

Fits a ctsem model specified via `ctModel` with type either 'stanct' or 'standt'.

Usage

```
ctStanFit(
  datalong,
  ctstanmodel,
```

(2) ctStanFit or ctFit

Ctsem: Dealing with unequal time intervals (panel data)

Output: Maximum likelihood parameter estimates & standard errors

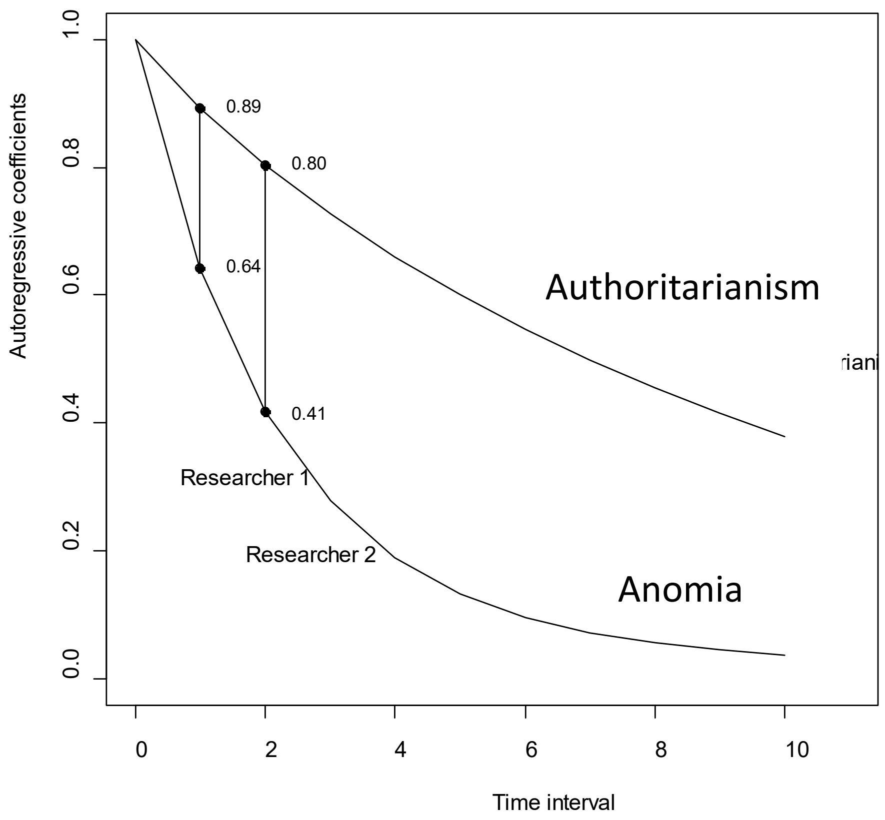
$ctparameters

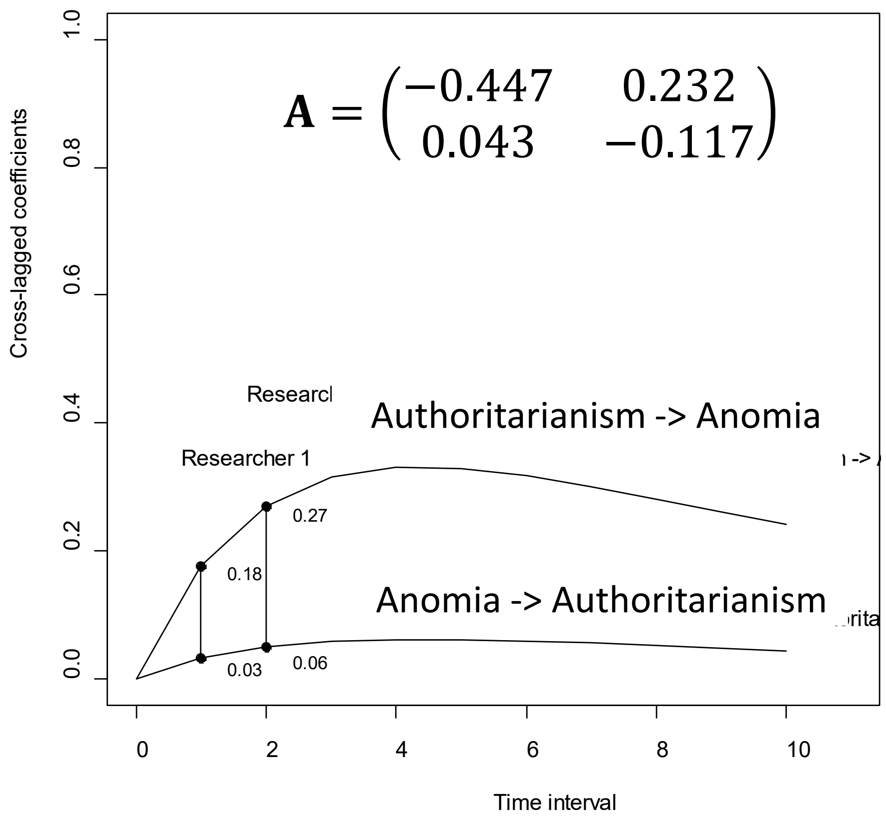
		Value	Matrix	StdError
	T0mean_eta1	2.503428618	T0MEANS	0.015249233
	T0mean_eta2	2.842722176	T0MEANS	0.012965464
Anomia	drift_eta1_eta1	-0.447282102	DRIFT	0.019876762
Anomia -> Authoritarianism	drift_eta2_eta1	0.043292534	DRIFT	0.009630959
Authoritarianism -> Anomia	drift_eta1_eta2	0.232497503	DRIFT	0.018058224
Authoritarianism	drift_eta2_eta2	-0.117466597	DRIFT	0.008651654
	diffusion_eta1_eta1	0.473241990	DIFFUSION	0.015697481
	diffusion_eta2_eta1	-0.004610021	DIFFUSION	0.005495036
	diffusion_eta2_eta2	0.154509667	DIFFUSION	0.003846351
	T0var_eta1_eta1	0.632861176	T0VAR	0.017155897
	T0var_eta2_eta1	0.244730836	T0VAR	0.011330971
	T0var_eta2_eta2	0.457577114	T0VAR	0.012403760
	c1	0.536315947	CINT	0.046289177
	c2	0.220005403	CINT	0.022199716

Drift matrix

Ctsem: Dealing with unequal time intervals (panel data)

Plots: Autoregressive and cross-lagged parameters



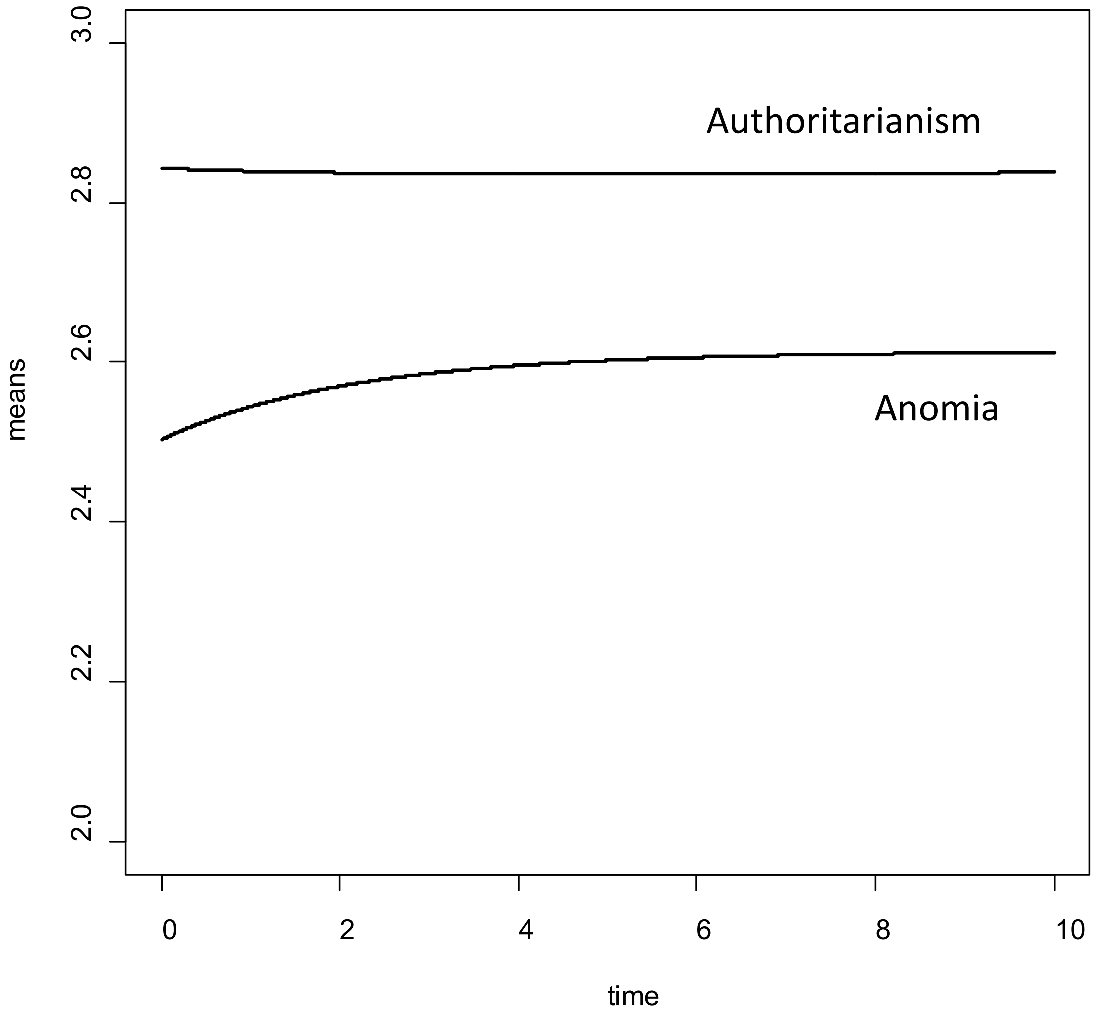


$$e^{\mathbf{A}\cdot\Delta t_i} = e^{\mathbf{A}\cdot 1} = \begin{pmatrix} 0.643 & 0.176 \\ 0.033 & 0.893 \end{pmatrix}$$

$$e^{\mathbf{A}\cdot\Delta t_i} = e^{\mathbf{A}\cdot 2} = \begin{pmatrix} 0.419 & 0.271 \\ 0.050 & 0.804 \end{pmatrix}$$

Ctsem: Dealing with unequal time intervals (panel data)

Plots: Mean trajectories



Ctsem: Dealing with unequal time intervals (panel data)

From theory to practice… let's repeat the analysis together:

```
#install.packages("ctsem")
library(ctsem)
library(ctsemOMX)

#Bivariate CT model
data(AnomAuth)
AnomAuthmodel <- ctModel(       Tpoints = 5, n.latent = 2, n.manifest = 2,
                                LAMBDA = matrix(c(1, 0, 0, 1), nrow = 2, ncol = 2),
                                MANIFESTVAR=diag(0, 2),
                                TRAITVAR = NULL)
AnomAuthfit <- ctFit(AnomAuth, AnomAuthmodel)
summary(AnomAuthfit)
```

01_cl_panelmodel_anomauth.R.

Ctsem: Dealing with unequal time intervals (panel data)

Let's take a look under the hood:

```
#Examine the RAM specification
attributes(AnomAuthfit)
AnomAuthfit$mxobj$matrices$A$labels
AnomAuthfit$mxobj$matrices$M$labels
AnomAuthfit$mxobj$matrices$F$values

#Examine the parameter constraints
AnomAuthfit$mxobj$algebras
AnomAuthfit$mxobj$algebras$discreteDRIFT_T1
AnomAuthfit$mxobj$algebras$discreteDRIFT_i1
```

Ctsem: Dealing with unequal time intervals (panel data)

Constraining parameters & model comparisons

```
#1) LR test for testing whether the effect of anomia (eta1) on authoritarianism (eta2) is
significantly different from zero

AnomAuthmodel_restricted <- AnomAuthmodel
AnomAuthmodel_restricted$DRIFT[2,1] <- 0
AnomAuthfit_restricted <- ctFit(AnomAuth, AnomAuthmodel_restricted)

summary(AnomAuthfit)$DRIFT
summary(AnomAuthfit_restricted)$DRIFT
chi2 <- summary(AnomAuthfit_restricted)$omxsummary$Minus2LogLikelihood-
summary(AnomAuthfit)$omxsummary$Minus2LogLikelihood
pchisq(chi2, 1, lower.tail=F) #p-value
mxCompare(AnomAuthfit$mxobj, AnomAuthfit_restricted$mxobj)
```

Ctsem: Dealing with unequal time intervals (panel data)

Constraining parameters & model comparisons

```
#2) are the two cross effects significantly different from each other?

AnomAuthmodel_restricted2 <- AnomAuthmodel
AnomAuthmodel_restricted2$DRIFT[2,1] <- "cross"
AnomAuthmodel_restricted2$DRIFT[1,2] <- "cross"
AnomAuthfit_restricted2 <- ctFit(AnomAuth, AnomAuthmodel_restricted2)
chi2 <- summary(AnomAuthfit_restricted2)$omxsummary$Minus2LogLikelihood-
summary(AnomAuthfit)$omxsummary$Minus2LogLikelihood
pchisq(chi2, 1, lower.tail=F) #p-value
mxCompare(AnomAuthfit$mxobj, AnomAuthfit_restricted2$mxobj)
```

Study Questions

Question 1:
Match the following discrete-time parameters with their continuous-time counterpart. Describe what each parameter represents in modeling dynamic systems.

Discrete-Time Parameter	Continuous-Time Equivalent	Description
A^* (autoregression)		
Q^* (process error)		
b^* (intercepts)		

Study Questions

Question 2:

Familiarize yourself with ctsemOMX. Check out the help files of **?ctModel** and **?ctFit**.

Question 3:

Examine the code of a bivariate continuous time model using the `ctsemOMX` package and the provided `AnomAuth` dataset. Fit the model and examine the drift parameters. What do the results tell you about the relationship between Anomia and Authoritarianism?

Question 4:

Apply a bivariate continuous time model to the ctExample1 data provided with the package (without a TRAITVAR argument).

Question 5:

Apply a bivariate continuous time model to the ctExample1 data provided with the package. Include TRAITVAR="auto". Compare the results to the previous analysis. What happened?

Study Questions

Question 6 (optional):

Choose a new longitudinal dataset (uni- or bivariate) and analyze the data using a basic dynamic ct model [for this task do not worry about preprocessing the data or whether the model is justified on theoretical grounds; it’s just about the method]. Share your approach/results with the group.