Bayesian Varying Intercepts and Slopes Modeling of Cognitive Aging Trajectories in Breast Cancer Survivors as Compared to Matched Controls

Yuelin Li[1 2]

[1] Department of Psychiatry & Behavioral Sciences, Memorial Sloan Kettering Cancer Center
[2] Department of Epidemiology & Biostatistics, MSKCC

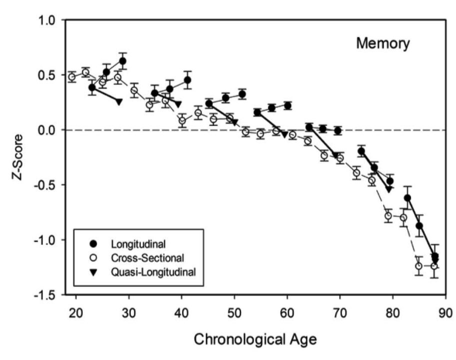
February 3, 2023

1 Purpose

This document summarizes the "varying-intercepts, varying-slopes" model in this example.

2 Problem

The model is motivated by a long-standing challenge in cognitive aging research, seen here from a graph in Salthouse (2019). [1]



If you analyze cognitive assessments cross-sectionally using each person's very first assessment (open circles above), you typically get a cognitive aging trend that shows nearly linear decline with older age. However, if you analyze the

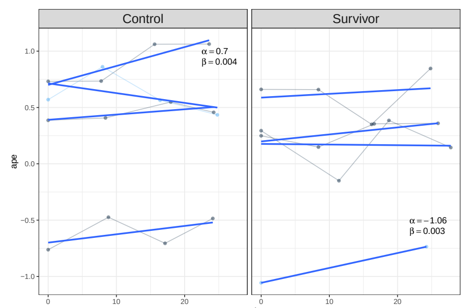
[1] Salthouse (2019). *Psychology and Aging.* `http://dx.doi.org/10.1037/pag0000288`

longitudinal data over repeated assessments (filled circles), you often get an increasing trend over time. Thus, the main problem is that repeated assessments are confounded by the *test exposure* effect (aka 'practice effect'). For pragmatic purposes we should call it practice effect throughout this tutorial.

The cross-sectional data is not affected by practice effect and thus may better explain the true aging trend. However, the cross-sectional aging trend has its own problems. It is affected by the cohort effect, i.e., people born in the 1940's (in their 70's to 80's by 2023) may be different than younger people born in the 1980's. Details on the varous issues are not immediately relevant to our model, but they can be found in the Salthouse paper.

In this tutorial, we use a Bayesian varying-intercepts, varying-slopes model to explain both the cross-sectional aging trend while *simultaneously* fit the longitudinal practice effect. You don't have to analzye cross-sectional and longitudinal data separately.

Before we proceed to data modeling, we should get an intuition on why the model involves the intercepts and slopes. This graph shows the neurocognitive assessments from four randomly-selected cancer survivors and controls.



From each person's 4 assessments, we can fit a personal regression line in blue. Here we show one control participant's regression with an intercept of $\alpha = 0.7$ and a slope of $\beta = 0.004$. The intercept of $\alpha = 0.7$ represents the person's very first assessment that is presumably free of practice effect. This person scored 0.70 standard deviations above the norm at study enrollment. The slope of $\beta = 0.004$ represents this person's expected rate of change per month, and practice effect is subsumed in this slope. Notice how the practice effect is now separated from cross-sectional aging.

3 Preliminary Model

We can now turn to a simpler version of the full model. The preliminary model is divided into 2 levels. In level 1, each person's 4 longitudinal assessments are first distilled into an intercept and a slope. Here we use i to index persons and t to index assessment time points.

$$y_{i[t]} \sim \mathrm{N}\big(\alpha_i + \beta_i \mathrm{Months}_{i[t]}, \sigma_\epsilon^2\big), \text{ for } i = 1, \dots, n; \quad t = 1, \dots, 4, \tag{1}$$

where the assessment collected from the ith person at time t is shown as $y_{i[t]}$. The square brackets show that the tth assessment is nested within the ith person. The varying intercepts and varying slopes are further analyzed in level 2, where the intercepts α_i and slopes β_i are modeled by covariates.

$$\begin{aligned}\alpha_i &= \gamma_{00} + \gamma_{01}\mathrm{Age}_i + \gamma_{02}\mathrm{Age}_i^2 + \gamma_{03}\mathrm{survivor}_i + \\ &\qquad \gamma_{04}\mathrm{survivor}_i \cdot \mathrm{Age}_i + \gamma_{05}\mathrm{survivor}_i \cdot \mathrm{Age}_i^2 \\ \beta_i &= \gamma_{10} + \gamma_{11}\mathrm{survivor}_i + \gamma_{12}\mathrm{age4Q}_i + \gamma_{13}\mathrm{survivor} \cdot \mathrm{age4Q}_i,\end{aligned} \tag{2}$$

the intercepts being a quardratic function of age at enrollment. More specifically, the control cohort follows this quadratic aging trend: $\gamma_{00} + \gamma_{01}\mathrm{Age}_i + \gamma_{02}\mathrm{Age}_i^2$ and the survivor cohort differs from the controls by $\gamma_{03}\mathrm{survivor}_i + \gamma_{04}\mathrm{survivor}_i \cdot \mathrm{Age}_i + \gamma_{05}\mathrm{survivor}_i \cdot \mathrm{Age}_i^2$. Additionally, the intercepts β_i are modeled as a function of 4 age quartiles in the $\mathrm{age4Q}_i$ predictor $\beta_i = \gamma_{10} + \gamma_{11}\mathrm{survivor}_i + \gamma_{12}\mathrm{age4Q}_i + \gamma_{13}\mathrm{survivor} \cdot \mathrm{age4Q}_i$. This effectively fits 8 separate practice effects, 4 for each of the age quartiles in controls and 4 for survivors. Finally, we gather equations (1) – (2) together to yield the following 2-level model:

$$\begin{aligned}&\mathrm{Level1:} \\ &\qquad y_{i[t]} \sim \mathrm{N}\big(\alpha_i + \beta_i \mathrm{Months}_{i[t]}, \sigma_\epsilon^2\big), \text{ for } i = 1, \ldots, n; \quad t = 1, \ldots, 4, \\ &\mathrm{Level2:} \\ &\qquad \alpha_i = \gamma_{00} + \gamma_{01}\mathrm{Age}_i + \gamma_{02}\mathrm{Age}_i^2 + \gamma_{03}\mathrm{survivor}_i + \\ &\qquad\qquad \gamma_{04}\mathrm{survivor}_i \cdot \mathrm{Age}_i + \gamma_{05}\mathrm{survivor}_i \cdot \mathrm{Age}_i^2 \\ &\qquad \beta_i = \gamma_{10} + \gamma_{11}\mathrm{survivor}_i + \gamma_{12}\mathrm{age4Q}_i + \gamma_{13}\mathrm{survivor} \cdot \mathrm{age4Q}_i,\end{aligned} \tag{3}$$

Readers may recognize this model as identical to how Bryk and Raudenbush would specify a 2-level HLM model. This is among the many ways of writing the basic multilevel model. You can plug α_i and β_i back into equation (1) and re-express the 2-level model in one line:

$$\begin{aligned}y_{i[t]} &\sim \mathrm{N}\big(\alpha_i + \beta_i \mathrm{Months}_{i[t]}, \sigma_\epsilon^2\big), \\ &\sim \mathrm{N}\big(\gamma_{00} + \gamma_{01}\mathrm{Age}_i + \gamma_{02}\mathrm{Age}_i^2 + \gamma_{03}\mathrm{survivor}_i + \\ &\qquad \gamma_{04}\mathrm{survivor}_i \cdot \mathrm{Age}_i + \gamma_{05}\mathrm{survivor}_i \cdot \mathrm{Age}_i^2 + \\ &\qquad [\gamma_{10} + \gamma_{11}\mathrm{survivor}_i + \gamma_{12}\mathrm{age4Q}_i + \gamma_{13}\mathrm{survivor} \cdot \mathrm{age4Q}_i] \cdot \mathrm{Months}_{i[t]}, \sigma_\epsilon^2\big).\end{aligned}$$

This one-line model can be easily fitted using the R package `lme4`, SAS PROC MIXED, the HLM package, or Mplus. Using `lme4`, for example, the model can be fitted like this:

```
lmer(y ~ age + I(age^2) + survivor + survivor:age +
    survivor:I(age^2) + months + survivor:months +
    age4Q:months + survivor:age4Q:months + (1 + months |
    study_id))
```

Note that the R syntax matches the math equation. The Bayesian model is exactly the same as a conventional, HLM-like expression. Next we will express the model in a Bayesian way, with priors and all. The differences should become more apparent.

4 Bayesian Model

The Bayesian expression has the same level-1 model. In level 2, however, the varying intercepts and slopes are drawn from a bivariate normal distribution with means μ and covariance Ω.

$$
\begin{aligned}
&\text{Level1}: \\
&\qquad y_{i[t]} \sim \mathrm{N}\big(\alpha_i + \beta_i \text{Months}_{i[t]}, \sigma_\epsilon^2\big), \text{ for } i = 1, \ldots, n; \quad t = 1, \ldots, 4 \\
&\text{Level2}: \\
&\qquad \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \hat{\mu}_{\alpha_i} \\ \hat{\mu}_{\beta_i} \end{pmatrix}, \Omega\right), \\
&\qquad \text{where} \\
&\qquad \begin{aligned}
\hat{\mu}_{\alpha_i} &= \gamma_{00} + \gamma_{01}\text{Age}_i + \gamma_{02}\text{Age}_i^2 + \gamma_{03}\text{survivor}_i + \\
&\qquad \gamma_{04}\text{survivor}_i \cdot \text{Age}_i + \gamma_{05}\text{survivor}_i \cdot \text{Age}_i^2 \\
\hat{\mu}_{\beta_i} &= \gamma_{10} + \gamma_{11}\text{survivor}_i + \gamma_{12}\text{age4Q}_i + \gamma_{13}\text{survivor} \cdot \text{age4Q}_i \\
\Omega &= \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix} \boldsymbol{R} \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix} \\
\boldsymbol{R} &\sim \text{LKJCorr}(2) \\
\boldsymbol{\gamma} &\sim \text{Normal}(\gamma, \sigma_\gamma) \\
\gamma &\sim \text{Normal}(0, 10) \\
\sigma_\gamma &\sim \text{HalfCauchy}(0, 1)
\end{aligned}
\end{aligned}
\tag{4}
$$

The notation using $\hat{\mu}_{\alpha_i}$ and $\hat{\mu}_{\beta_i}$ is a different way of expressing the level-2 equations. The contents are the same as in equation (3).

The covariance matrix ω is constructed by factoring it into separate standard deviations $\sigma_{\alpha_i}, \sigma_{\beta_i}$ and a correlation matrix $\boldsymbol{R}$. The variances $\sigma_{\alpha_i}, \sigma_{\beta_i}$ are for the intercepts and slopes in the diagonals, and a correlation rho between intercepts and slopes. A negative correlation would suggest that a high intercept is associated with a slower growth and vice versa. The correlation matrix $\boldsymbol{R}$ follows an LKJ prior,`https://distribution-explorer.github.io/multivariate_continuous/lkj.html`, which offers more flexibility than the Wishart distribution. An LKJCorr(2) prior is essentially expressing the prior belief of a weak correlation coefficient mostly between -0.5 and +0.5, but it can also be outside that range.`https://psychstatistics.github.io/2014/12/27/d-lkj-priors/` The remaining lines define fixed priors and hyper-priors for the parameters. The correlation between the intercepts and slopes is a noteworthy feature of the model because we can estimate the correlations of these random effects, thereby providing information on the co-occurrences between the cross-sectional aging trend and practice effect. This correlation is omitted when they are analyzed separately.

5 Summary

This completes the model equations. The Bayesian approach is similar to the HLM model, but it contains additional information on the covariance matrix and priors on the model coefficients.