Implementing LCGA and GMM Using `hlme` in R

Mike Bader

The following code snippets show how to estimate three-group (i.e., $K=3$) versions of a latent class growth analysis (LCGA) model and a growth mixture model (GMM) based on a linear growth model.

These snippets come from the R files simulating both types of models (`02a_lcga_simulation.R` and `02b_gmm_simulation.R`) that use a simulated long dataset `d_sim` of decades within neighborhoods with the outcome variable `pwht` simulating the share of White residents in those neighborhoods and a variable `t` measuring decades.

LCGA

A LCGA includes separate parameters for each of the deterministic components (γ_{00} and γ_{10}) for each of the K=3 classes.

$$Y_{ti|c=k} = \gamma_{00|c=k} + \gamma_{10|c=k}\text{time}_t + \epsilon_{ti|c=k}$$

Note that the `mixture` option below contains a formula describing these deterministic components of the model (intercept and slope based on the variable `t`).

```
m3 <- hlme(
    pwht ~ t,           ## Equation of the fixed (deterministic) component
                        ## of the growth trajectory to be estimated; this
                        ## is the `fixed` argument of `hlme`

    mixture = ~ 1 + t,  ## Definition of the components from the fixed
                        ## (deterministic) part of the model to be used to
                        ## define latent trajectories (we can leave off the
                        ## 1 as the intercept is implied, i.e., could
                        ## write `~ t` rather than `~ 1 + t`)

    ng = 3              ## Number of latent classes to estimate

    subject = "i",      ## Name of the variable containing group ID (in quotes)

    data = d_sim,       ## Name of R object containing data

    B = random(m1)      ## Starting values; in this instance use a random draw
                        ## from the results of the model with a single group as
                        ## starting values for the estimation
)
```

GMM

A GMM adds random components for the intercept and slope, ρ_{0i} ρ_{1i}, to the growth trajectory model that represents the class-specific deviation from the class-specific growth parameters. The model assumes that the variance-covariance structure of the random components ($\mathbf{T}$) is unique by class.

$$Y_{ti|c=k} = \gamma_{00|c=k} + \gamma_{10|c=k}\text{time}_t + \rho_{0i} + \rho_{1i}\text{time}_t + \epsilon_{ti|c=k}$$

Note that the `random` option below contains a formula representing the stochastic component of the model above (random intercepts and slope).

The `gridsearch` function wraps `hlme` to specify initial values that are less likely to lead to a local maximum solution. The `minit` option provides the model object of the same model being estimated based on a single group. It can be used for any type of hlme model where K > 1.

```
m3 <- gridsearch(
  rep = 100, maxiter = 30, minit = m1,
  hlme(
    pwht ~ t,          ## Same as above

    mixture = ~t,      ## Same as above (this time omitting the explicit
                       ## intercept in the formula)

    random = ~1 + t,   ## The random (stochastic) components of the model to be
                       ## estimated; this is what makes this model a GMM rather
                       ## than an LCGA

    nwg = TRUE,        ## Should the variance-covariance matrix (Tau) of the
                       ## random effects be estimated separately for each of the
                       ## latent classes? (Note that hlme implements this based
                       ## only on a signle parameter measuring the
                       ## proportional change to a single variance-covariance
                       ## matrix)

    ng = 3,            ## Same as above

    subject = 'i',     ## Same as above

    data = d_sim,      ## Same as above
)
```