Journal of Statistical Software

March 2014, Volume 57, Issue 3. http://www.jstatsoft.org/

Bayesian Semi- and Non-Parametric Models for Longitudinal Data with Multiple Membership Effects in R

Terrance D. Savitsky
U.S. Bureau of Labor Statistics

Susan M. Paddock
RAND Corporation

Abstract

We introduce **growcurves** for R that performs analysis of repeated measures multiple membership (MM) data. This data structure arises in studies under which an intervention is delivered to each subject through the subject's participation in a set of multiple *elements* that characterize the intervention. In our motivating study design under which subjects receive a group cognitive behavioral therapy (CBT) treatment, an *element* is a group CBT session and each subject attends multiple sessions that, together, comprise the treatment. The sets of elements, or group CBT sessions, attended by subjects will partly overlap with some of those from other subjects to induce a dependence in their responses. The **growcurves** package offers two alternative sets of hierarchical models: 1. Separate terms are specified for multivariate subject and MM element random effects, where the subject effects are modeled under a Dirichlet process prior to produce a semi-parametric construction; 2. A single term is employed to model joint subject-by-MM effects. A fully non-parametric dependent Dirichlet process formulation allows exploration of differences in subject responses across different MM elements. This model allows for borrowing information among subjects who express similar longitudinal trajectories for flexible estimation. **growcurves** deploys "estimation" functions to perform posterior sampling under a suite of prior options. An accompanying set of "plot" functions allows the user to readily extract by-subject growth curves. The design approach intends to anticipate inferential goals with tools that fully extract information from repeated measures data. Computational efficiency is achieved by performing the sampling for estimation functions using compiled C++ code.

Keywords: growth curve, Bayesian hierarchical model, conditional autoregressive model, Dirichlet process, R, C++.

1. Introduction

1.1. Motivation

Repeated measures data are commonly used to compare the effectiveness of alternative treatments or interventions, to study persistence of a treatment effect, and to understand the process of symptom attenuation or augmentation. The number of repeated observations per study participant is often constrained by factors such as data collection costs, consideration of substantively meaningful measurement points (e.g., baseline and 1-year follow-up), and data availability when secondary data sources are used.

A common feature of such studies in the behavioral and social sciences is that participants might experience an intervention together, leading to correlated outcomes. A further complication in such studies is that the memberships in or linkages to a set of treatment *elements* might vary. An *element* under the BRIGHT study illustrations to follow is defined to be a cognitive behavioral group therapy session used in a study of mental health treatment effectiveness among clients who experience substance abuse. Each client attends multiple therapy sessions that, together, characterize the treatment. Clients express overlaps in their session attendances that induce correlations among their responses to the treatment. Such studies often employ open enrollment in which individual clients enter and leave a standing therapy group at different sessions (Morgan-Lopez and Fals-Stewart 2006; Paddock, Hunter, Watkins, and McCaffrey 2011). In the education evaluation context, the composition of classrooms may change over time (Hill and Goldstein 1998). In addition, incorporating this membership into the analysis must account for the fact that each participant's post-treatment outcomes reflect the effects of each membership element (e.g., therapy group sessions) attended by that participant during the course of treatment.

1.2. Scope of growcurves

The **growcurves** package for the R statistical software platform (R Core Team 2013) is designed for Bayesian hierarchical modeling of longitudinal repeated measures of continuous outcomes that accounts for dependence among subjects, such as group therapy clients, induced by overlaps in their elements of memberships (such as group therapy sessions). Further, **growcurves** employs a non-parametric Dirichlet process prior on a collection of subject-indexed random effects to borrow strength that permits estimation of flexible, non-linear curves or trajectories. Inference on extracted clusters of subjects also allows for additional inference on subject random effects. Package **growcurves** is available from the Comprehensive R Archive Network at `http://CRAN.R-project.org/package=growcurves` (Savitsky 2014).

It is common for behavioral intervention studies to employ a protocol that may induce correlations across study participants; for example, the design of the BRIGHT study employed as our case illustration in Section 2 induces between-client correlations derived from their overlapping patterns of session attendance induced by open enrollment. To model this dependence requires us to "link" each client to the typically multiple sessions they attend. The **growcurves** package allows the user to employ a multiple membership (MM) model for this purpose that defines a univariate set of random effects, one for each session, which are then mapped to clients to produce an average client effect (Browne, Goldstein, and Rasbash 2001; Hill and Goldstein 1998). We call the collection of these session effects *multiple membership*

(MM) effects. We will model the dependence among clients induced by overlaps among client session attendances and map these effect values back to clients through an MM weight matrix that encodes the session attendances of all clients. Adjacency dependence among the session effects will be modeled using a conditional autoregressive (CAR) formulation specified in Paddock and Savitsky (2013).

There are other software tools and R packages that allow a user to fit sub-models of the overall model described above, such as **MLwiN** by Rasbash, Charlton, Browne, Healy, and Cameron (2012) for fitting parametric hierarchical models that include the MM construction of Hill and Goldstein (1998) and the **DPpackage** of Jara, Hanson, Quintana, Mueller, and Rosner (2011) for fitting DP priors on sets of random effects. Yet, **growcurves** is directly tailored for Bayesian semi-parametric analyses of repeated measures under study protocols that may induce correlations among participants, such as under the open-enrollment structure in the BRIGHT case study, and so combines these separate modeling components and adds new ones. Researchers involved with open-enrollment group therapy studies or similarly structured behavioral or social intervention studies would thus benefit from having a single R package to conduct post-treatment outcomes analyses. Analysis is supported in **growcurves** by incorporating the typical inference performed on repeated measures behavioral intervention data, which generally espouses three goals: 1. To extract by-subject growth curves to distinguish the patterns of persistence, attenuation or augmentation across the population; 2. To compare the fixed effects of two or more treatments. The treatment mean effect may be non-linear in time and composed of interactions of treatment effect with time; 3. To model dependence among subjects induced by study design or any known factors. The **growcurves** package provides a "plot" function for each of these goals that accepts output objects from estimation functions as their inputs and allows additional user settings to directly perform inference and produce associated graphical and data summary outputs not otherwise available in more general packages.

We outline the BRIGHT application study design and data structures that we use to illustrate **growcurves**' functions in Section 2 in order to frame the definition of our Bayesian models that follow. An overview of the input and output data structures for **growcurves** is introduced in Section 3. The section first specifies the additive Bayesian model framework that is composed of fixed, subject and MM random effects. Alternative prior formulations for each of the subject and MM random effects are outlined Section 4 and include illustrations of **growcurves**' functions under our BRIGHT application. A more flexible model that defines a distinct set of MM effects for each subject is introduced in Section 5 and again illustrated for our BRIGHT data set. Computational runtime comparisons are offered for all **growcurves**' estimation functions in Section 6, followed by concluding remarks in Section 7.

2. Case study data set

The building recovery by improving goals, habits and thoughts (BRIGHT) study (Watkins *et al.* 2011) was a community-based effectiveness trial of a group cognitive behavioral therapy (CBT) intervention for reducing depressive symptoms among residential substance abuse treatment clients. An aim of the study was to test whether clients receiving the BRIGHT intervention would have sustained improvements in depressive symptoms following treatment. The BRIGHT study employed a quasi-experimental design under which cohorts of clients in each of four treatment sites that received either treatment as usual (UC) or treatment en-

hanced with the BRIGHT intervention (CBT). Clients were assigned to receive either CBT or UC according to which intervention was offered at their study sites at the time of entry into residential substance abuse treatment. Overall, $n = 299$ subjects enrolled into the study, with 140 assigned to CBT, of which 132 attended CBT sessions, and 159 to UC. A total of $S = 245$ CBT sessions were offered over all the $G = 4$ therapy groups. Data were collected from subjects at a baseline survey administered when subjects enrolled in the BRIGHT study and at two post-treatment follow-up assessments conducted at 3 and 6 months post-baseline, providing a total of $T = 3$ waves of measures. The outcome of interest is client depressive symptoms as measured by the Beck Depression Inventory-II (BDI-II) (Beck, Steer, and Brown 1996). The BDI-II score is a sum across 21 four-level items (scored 0–3), with a higher score indicating a greater level of depressive symptoms. Subjects in the CBT arm could attend the 16 session CBT treatment.

Trends in depressive symptoms scores within-client are non-linear; see Paddock and Savitsky (2013). However, the flexibility to accurately model this non-linearity using parametric procedures widely available in software packages is constrained by having up to just three time points per client. Another feature of BRIGHT is that clients entered an ongoing CBT group on an open-enrollment basis, which allowed for group membership to change session-to-session. This instantiated a complex correlation structure among co-grouped clients. Suppose that clients "Fred" and "Mary" participate in a particular therapy group at session 3. One might expect that their outcomes might be correlated, owing to the interactive nature of group therapy. Now, further suppose that "Fred" leaves the group immediately following session 3 and a new client, "Betty", joins the group at session 5. The depressive score outcomes of Fred and Betty may be correlated, even though they never met. This correlation is allowed through the possibility for Fred to influence Mary, which may, in turn, result in Fred influencing Betty.

3. Overview of package design and input

3.1. Model construction

Our first class of functions that produce posterior estimates of model parameters under a MM data structure is specified in an additive form,

$$\text{subject-time response} = \text{fixed effects} + \text{subject effects} + \text{MM effects} + \text{error}.$$

Stating the same equation in mathematical notation,

$$y_{ij} = \alpha + \mathbf{x}_{ij}^{\top}\boldsymbol{\beta} + \mathbf{z}_{ij}^{\top}\mathbf{b}_i + \mathbf{w}_i^{\top}\boldsymbol{u} + \epsilon_{ij}, \tag{1}$$

where y_{ij} is a continuous outcome observed for subject $i = (1, \ldots, n)$ at repeated measurement event $j = (1, \ldots, m_i)$, and m_i reflects the fact that different numbers of repeated measures may be observed across subjects, so that the data are not required to be balanced. The first two terms on the right-hand side are the intercept and fixed effects, respectively. The fixed effects are parameterized as a polynomial function of time that includes a treatment arm indicator and its interaction with a polynomial function of time in the form of $A + Af(time) + \ldots$ *or* $B + Bf(time) + \ldots$ for each subject, where A and B are treatment arm indicators for this illustration with $K = 3$ arms (with one level excluded as a hold-out for identifiability) and $f(time)$ is a polynomial function of time. This parameterization requires each subject to be

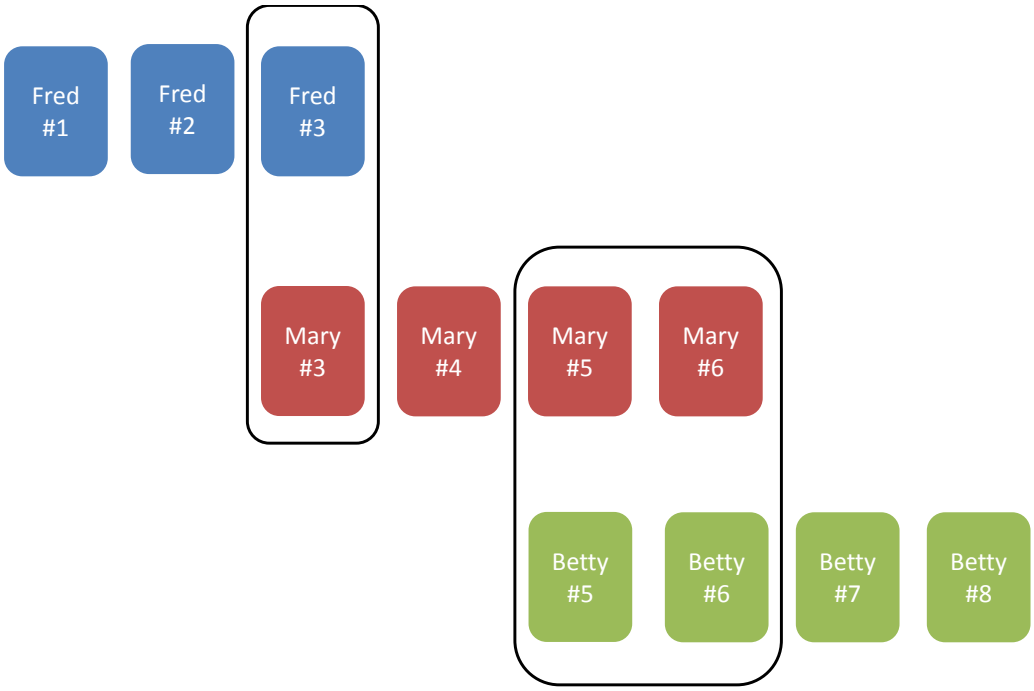

Figure 1: Induced dependence among CBT clients through "chaining" of their session attendance sequences.

assigned to a single treatment arm (including the possibility for assignment to a control arm, if applicable) so that they receive either the fixed effects term for A or B if they are assigned to either of these arms. Our models in **growcurves** allow as many treatment arms as the user may define, though our BRIGHT study illustration includes two arms. The $\{\mathbf{b}_i\}_{i=1,\ldots,n}$ represent subject random effects that permit borrowing of strength among subjects in the growth curve estimation. Finally, $\boldsymbol{u} = (u_1, \ldots, u_S)^\top$ denotes the collection of session random effects.

We next offer more detail on the construction of each term on the right-hand side of Equation 1. Moving left-to-right, α is a fixed intercept term; $\mathbf{x}_{ij}$ is a $p \times 1$ set of fixed effects predictors, with the effect of each predictor on the outcome expressed as $\boldsymbol{\beta}$; $\mathbf{z}_{ij} = \left(1, t_{ij}, t_{ij}^2, \ldots, t_{ij}^q\right)$ is a set of q predictors that multiply the q client random effects, or growth parameters, $\mathbf{b}_i$, where t_{ij} denotes the continuously-valued time at which outcome y_{ij} was observed; and ϵ_{ij} is an observational error term.

We formulate the fixed effects covariates to permit assessment of treatment effects by enumerating predictors to index the treatment arm assignment for subject i, capture the time of repeated measurement and account for their interactions. In particular, we define a vector of fixed effects, $\mathbf{d}_i$, a $1 \times (K-1)$ vector of 0's with a 1 in the position representing the treatment arm assigned to subject i; e.g., $d_{i\ell} = 1$ if subject i is assigned to treatment ℓ and K counts the number of treatment arms (with the first treatment, here labeled as 0, held out for identifiability). The treatment arm(s) design vector, $\mathbf{d}_i$, includes a variable for each arm or level $\in (1, 2, \ldots, (K-1))$. Then this composition allows for more than the usual single treatment and control arms study configuration. The estimation functions of **growcurves** employ an op-

tion, `n.fix_degree` that determines the polynomial order for fixed effects covariates that are functions of time. In particular, we will see in the next section that setting `n.fix_degree = 2` produces the $p = 3K - 1$ fixed effects vector, $\mathbf{x}_{ij} = \left(\mathbf{d}_i, t_{ij}, t_{ij}^2, \mathbf{d}_i t_{ij}, \mathbf{d}_i t_{ij}^2\right)^\top$. It bears mention that under employment of two treatment arms as used in our CBT case study discussed in Section 2, $\mathbf{d}_i$ for subject i reduces to a treatment indicator variable, d_i, such that $d_i = 1$ if subject i is assigned to the treatment arm and $d_i = 0$ if assigned to the control. Then, parameters $\boldsymbol{\beta}$ represent the effects of treatment assignment, time trend and the differential effect of treatment(s) on the average time trend in y.

MM effects, $\boldsymbol{u} = (u_1, \ldots, u_S)^\top$, are indexed by group therapy *session* under the BRIGHT study, producing a set of $S = 245$ group therapy session effects. Clients (who are the "subjects") are assigned to the treatment arm may attend up to 16 sessions of group CBT. To model outcomes for clients who attend multiple sessions, we employ a MM construction to map each element of $\boldsymbol{u}$ to the y's; see Browne *et al.* (2001). Under the MM model, the components of $\boldsymbol{u}$ are mapped to the y_{ij}'s by multiplying $\boldsymbol{u}$ by an $S \times 1$ weight vector, $\mathbf{w}_i$, that is normalized to sum to 1; in particular, S_i may be viewed as the number of sessions attended by client i; $w_{is} = 1/S_i$ if client i participates in session s and $w_{is} = 0$ otherwise. We normalize vector, $\mathbf{w}_i$, to sum to 1 because we intend that the treatment fixed effects reflect the magnitude of the treatment effect, while the MM term expresses dependence among clients induced by the study design. Let $N = \sum_i m_i$ denote the number of repeated measures observed for all study clients. Then $\mathbf{W} = \{\mathbf{w}_i\}$ is the resulting $N \times S$ MM matrix where the rows representing the repeated measures for client i are identical. In the sections to follow, we refer to the set of random effects linked to subjects through an MM matrix as "MM random effects".

3.2. Package input and output structures

Modeling is performed with estimation functions, of which there are 3: `dpgrow`, `dpgrowmm`, and `dpgrowmult`. Working from left-to-right, we will see that each estimation function may be viewed as a simpler, special case of that which follows. Each model is distinguished by the number of included MM terms (which may be 0 or greater) and the prior formulation options chosen for the set of subject-indexed and MM random effects. We will outline the prior formulations for and subsequently illustrate and compare these functions in Section 4.

Inputs to estimation functions

We presently focus on the elements of our data input structure that are common across estimation functions. The data input components are designed to anticipate the structures employed in most applications with repeated measures data. For illustration of the data input vocabulary common to the estimation functions, we select `dpgrowmm`, which contains a single MM term under varied priors and a DP prior on the set of by-subject random effects. Since the construction of growth curves depends on knowing which covariates are associated to treatment and time indices as well as the subject identifier for repeated measures, user input to the estimation functions directly input these as vectors and automatically compose fixed and random effects design matrices from them, as follows:

```
R> library("growcurves")
R> data("datsim", package = "growcurves")
```

```
R> attach(datsim)
R> n.iter <- 20000
R> n.burn <- 10000
R> n.thin <- 10
R> out <- dpgrowmm(y = datsim$y, subject = subject, trt = trt, time = time,
+    n.random = 3, n.fix_degree = 2, group = group, subj.aff = subj.aff,
+    W.subj.aff = W.subj.aff, n.iter = n.iter, n.burn = n.burn,
+    n.thin = n.thin, option = "mmigrp")

[1] "Your chosen option = mmigrp"
```

1. `subject`: A numeric or character vector that provides subject identifiers for each observation. In particular, `subject` is of length, $N = \sum_{i=1}^{n} m_i$, the number of subject-time cases. For example, if there are 3 subjects, `"fred"`, `"barney"` and `"wilma"`, where the first has observations at 2 time points and the latter two clients each hold observations at 3 time points, then

   ```
   R> subject <- c("fred", "fred", "barney", "barney",  "barney", "wilma",
   +    "wilma", "wilma")
   ```

2. `trt`: A numeric or character vector of length N of treatment labels for one or more arms and may be as many as desired. For example, under 2 treatments each entry in `trt` might be $\in \{0, 1\}$ or $\in \{$`"cbt", "uc"`$\}$, where `"uc"` stands for "usual care" in our BRIGHT CBT case study. The `trt` vector assigns each subject to a treatment arm.

3. `time`: A numeric vector of length N specifying the time point for each subject-time observation.

4. `n.random`: The number of random effect terms per subject. `n.random = 2` captures random subject intercept and slope parameters, while `n.random = 3` adds a quadratic term.

5. `n.fix_degree`: The polynomial order for the fixed effects design matrix. As noted above, the fixed effects include treatment and time covariates and treatment-time covariate interactions. If `n.fix_degree = 2`, then there is a quadratic covariate for time and another for time by treatment(s) interaction(s).

6. `subj.aff`: A numeric or character vector of labels for unique subjects connected with the MM effects term. The subjects included in `subj.aff` should be a strict subset of the unique values enumerated in `subject`. The length of `subj.aff` is equal to the number of subjects linked to MM effects, n_{aff}. For example, if only `"fred"` and `"wilma"` are linked to MM effects, such as therapy sessions, then `subj.aff = c("fred", "wilma")`.

7. `W.subj.aff`: A numeric matrix of dimension $n_{\text{aff}} \times S$, where S denotes the number of MM effects when a single set of MM effects is in the model. There is no requirement for the rows to sum to 1, but each row must contain a numerical value that links each MM effect to a subject in `subj.aff`. The identifiers in `subj.aff` align to the rows of `W.subj.aff`. Crafting an example of a row of `W.subj.aff`, if Fred attends the first 5 sessions and there are $S = 10$ total sessions available, then the row of `W.subj.aff` corresponding to Fred will be defined by $(0.2, 0.2, 0.2, 0.2, 0.2, 0, 0, 0, 0, 0)$ (Hill and Goldstein 1998).

8. `group`: A numeric or character vector of length S, providing group identifiers for each of S MM effects. A grouping is composed of a set of disjoint collections of subjects and associated MM effects.

9. `n.iter`: A numerical scalar indicating the number of posterior sampling iterations of the Gibbs sampler.

10. `n.burn`: A numerical scalar for the number of initial iterations to discard.

11. `n.thin`: A numerical scalar indicating the gap in post-burn-in posterior sampling iterations to retain.

12. `option`: A scalar character value that indicates the covariance formulation under a Gaussian prior on the MM effects. Choices must be made within `c("mmi", "mmigrp", "mmcar")`. Values for `option` will vary across estimation functions.

Outputs from estimation functions

The returned object `out` of the `dpgrowmm` estimation function includes parameter summaries, model fit statistics, sampled parameters and a set of standard plots. We extract these summaries with S3 functions. The S3 class is primarily used to define functions that wrap a return object from an estimation function. Invoking the S3 function, `summary()`, on the return object produces parameter summaries; for example, `summary(out)$summary.results`, contains credible intervals for all model parameters. Table 1 highlights some (but not all) of the most useful summary objects returned by `summary(out)`.

Penalized fit statistics, such as the log pseudo marginal likelihood (`"lpml"`), a leave-one-out fit statistic intended to assess out-of-sample fit that marginalizes over the parameters, are particularly useful for mixture models in lieu of the usual *DIC* due to difficulty in estimating $\hat{D}$ to compute the effective number of parameters, pD of the *DIC*. The `"DIC"` object defined as the DIC_3 statistic also marginalizes over the parameters, but the simulations of Celeux *et al.* (2006) suggest it may under-penalize complexity. Since the fixed and random effects design matrices, $\mathbf{X}$ and $\mathbf{Z}$, respectively, are constructed from user data inputs noted above, we return the associated matrices with helpful column names for user inspection.

Posterior sampled values of all model parameters are returned for `(n.iter - n.burn)/n.thin` MCMC iterations with `parms.samples <- samples(out)`. The posterior samples may be used to compose trace plots or to compute effective sample sizes to assess the convergence properties of the sampling chain. Table 2 highlights some of the most useful objects stored in `parms.samples` holding posterior sampled values.

Calling the function `residuals(out)` returns a set of model residual values. A `plot(out, TRUE)` function returns a set of '`ggplot`' (Wickham 2010) objects that include growth curves aggregated by treatment group, a random selection of individual growth curves, plots of 95% credible intervals for MM and by-subject effects, as well as selected trace plots of model precision parameters and number of clusters formed (under the DP prior on subject effects). If the `plot` option is set to `FALSE`, the plots will not be automatically rendered (though they will be generated and stored as '`ggplot`' objects; otherwise, the plots will be displayed in separate plot windows).

Object	Description
`bmat.summary`	Credible intervals for subject random effects. A `list` object of `n.random` components. Each component represents as polynomial order (from 1− `n.random`). Each component contains an $n \times 3$ matrix. Columns are credible intervals; $2.5\%, 50\%, 97.5\%$.
`u.summary`	Credible intervals for MM random effects. An $S \times 3$ `matrix`. Each row indexes an MM element (such as a session). Each component contains an $n \times 3$ matrix. Columns are credible intervals; $2.5\%, 50\%, 97.5\%$.
`beta.summary`	Credible intervals for fixed effects. An $p \times 3$ `matrix`. Each row indexes a fixed effect, e.g., (`time`, `time`^{2}, `trt` × `time`, `trt` × `time`^{2}). Columns are credible intervals; $2.5\%, 50\%, 97.5\%$.
`X`	The $N \times p$ fixed effects design `matrix`. Each row indexes a `subject-time` case observation. Each column includes names for the fixed effects e.g., (`time`, `time`^{2}, `trt` × `time`, `trt` × `time`^{2}), formed from user input of data objects.
`Z`	An $N \times q$ random effects design `matrix`. Each column includes names for the random effects, e.g., (1, `time`, `time`^{2}), formed from user input of data objects.
`lpml`	A leave-one-out penalized fit statistic described in Congdon (2005).
`DIC`	The DIC_3 fit statistic of Celeux, Forbes, Robert, and Titterington (2006).
`Dbar`	The model deviance, $\bar{D}$.

Table 1: Post-burnin posterior summaries for model estimated parameters output from `summary(out)`.

Object	Description
`Beta`	An `n.iter` × p matrix. Columns are labeled with predictor names e.g., (`time`, `time`^{2}, `trt` × `time`, `trt` × `time`^{2}).
`Gamma`	An `n.iter` × S matrix. Each column represents an MM element (e.g., session)
`B`	An `n.iter` × (`n.random` × n) matrix. Polynomial order is the *slow*-moving index Subject is the *fast*-moving index
`bigSmin`	`list` object of M components, where M denotes number of clusters. Component m contains a vector of subjects in cluster m

Table 2: Post-burnin posterior samples for model estimated parameters output from `samples(out)`.

We next review underlying model formulations for each estimation function and explain available modeling `option` choices.

4. Prior formulations for subject and MM effects

4.1. Non-parametric modeling of growth parameters

The `dpgrow` estimation function addresses models that *exclude* an MM term, $\mathbf{w}_{ij}^{\top}\boldsymbol{u}$, from Equation 1, and solely employ the subject effects term, $\mathbf{z}_{ij}^{\top}\mathbf{b}_i$, to borrow estimation strength among subjects. Each subject defines $(q = 3) \times 1$ random effects, $\mathbf{b}_i$, that are multiplied by $\mathbf{z}_{ij} = (1, t_{ij}, t_{ij}^2)^{\top}$ to allow the effect term for each subject to vary with time, t_{ij}. The `option = "lgm"` imposes an independent Gaussian prior, $\mathbf{b}_1, \ldots, \mathbf{b}_n | \boldsymbol{\tau}_b \sim \mathcal{N}_q\left(\mathbf{0}, \boldsymbol{\tau}_b^{-1}\, \mathbf{I}_q\right)$, where $\boldsymbol{\tau}_b = (\tau_1, \ldots, \tau_q)$ for q random effect terms, by subject, each of which receives a further Gamma$(1, 1)$ prior to allow for further variation. All gamma priors employed in this paper specify a construction with rate (rather than scale) hyperparameters. We are restricted to $q = T - 1$, where T counts the number of measurement waves, in the case one wants to avoid weak identification through the prior (accomplished with use of the precision parameter, $\boldsymbol{\tau}_b$ to perform shrinkage estimation); otherwise, the user may select any value for q. The fixed effects parameters, $(\alpha, \boldsymbol{\beta})$ receive non-informative, flat priors in all models.

The DP framework (Ferguson 1973; Escobar and West 1995) generalizes the prior construction on $\mathbf{B} = \left\{\mathbf{b}_i^{\top}\right\}$ to allow estimation of a different form than a specified parametric distribution, such as the Gaussian (Brown and Ibrahim 2003; Müller, Quintana, and Rosner 2007). Using `option = "dp"` in the `dpgrow` estimation function replaces the standard parametric distribution on the random growth parameters, $\mathbf{B} = \left\{\mathbf{b}_i^{\top}\right\}$ with:

$$\mathbf{b}_1, \ldots, \mathbf{b}_n | F \overset{\text{iid}}{\sim} F \tag{2}$$

$$F | c, F_0 \sim \text{DP}(c, F_0), \tag{3}$$

where $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are conditionally independent given F, and are generated from the unknown distribution F. The base measure, F_0, represents the 'best guess' about the form of F prior to observing data; the expected value of F is F_0. The precision parameter, $c > 0$, expresses the degree of confidence that F_0 is the correct generating distribution for $\mathbf{B}$; the higher the value of c, the more F is expected to conform to F_0. We choose $F_0 \equiv \mathcal{N}_q\left(\mathbf{0}, \boldsymbol{\tau}_b^{-1}\, \mathbf{I}_q\right)$, the same as our parametric construction under `option = "lgm"`, to discover how much the estimated model differs from the analogous parametric approach.

The DP prior produces almost surely discrete realizations as can be seen from the conditional prior over $\mathbf{B}$ after using the exchangeability of the DP prior to marginalize over the random measure, F (Blackwell and MacQueen 1973):

$$\mathbf{b}_i | \mathbf{B}_{-i} \sim \frac{1}{n - 1 + c} \sum_{j=1, j \neq i}^{n} \delta_{\mathbf{b}_j}(\mathbf{b}_i) + \frac{c}{n - 1 + c}\, F_0, \tag{4}$$

where $\mathbf{B}_{-i} = (\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n)$ and $\delta_{\mathbf{b}_j}(\mathbf{b}_i)$ is a point mass density that equals 1 if $\mathbf{b}_i = \mathbf{b}_j$ and 0 otherwise. Resulting samples drawn from Equation 4 may express ties or co-clustering as a mechanism for the data to discover the dependence among subject effects, conditional on both the chosen base measure, F_0, and concentration parameter, c, which influence the number of clusters formed. The parameter, M, captures the number of clusters formed for each posterior sampling iteration and the results are returned in both `summary()` and `sample()` functions. Introduce $C_1, \ldots, C_M$, $M \leq n$ to index the clusters,

where $C_j = \{i : \mathbf{b}_i = \mathbf{b}_j^*\}$, where $\mathbf{b}_j^*$ capture unique "locations" or unique values for the subjects assigned to cluster j. Then each posterior sampling iteration produces a clustering defined by $(\mathbf{C}, \mathbf{B}^*)$, which we point out is a sample from the posterior distribution over the space of partitions or clusters. We employ the least squares clustering algorithm of Dahl, Day, and Tsai (2008) in **growcurves** to select a single clustering of subjects from among the clusters sampled over the posterior sampling iterations based on some notion of optimality. This algorithm uses the cluster memberships over the set of posterior samples to compose an $n \times n$ square matrix of pairwise co-clustering probabilities of subjects. The selected cluster minimizes the distance to this pairwise clustering probability matrix in a least squares sense. This chosen clustering may be used to conduct inference on the pattern of dependence among subject effects as we will do in Section 2. We may extract the clustering selected under least squares with `samples(out)$bigSmin`, where `out` is an object of class `c("dpgrow", "dpgrowmm", "dpgrowmult")` from the applicable estimation function.

We impose $c \sim \text{Gamma}(a, 1)$, which encodes the prior expected number of clusters, with higher values for shape parameter, a, producing more clusters. This shape parameter setting may be input to `dpgrowmm` with `shape.dp = 1`. This value defaults to 1 if no input is given by the user.

4.2. Including a single MM term with prior formulations

We next return to the model formulation of Equation 1 under which a single MM term is included. The `dpgrowmm` function performs inference for this model under various prior formulations on the MM term. The simplest construction incorporating MM effects is the standard model in which the effects are assumed to be independent and identically distributed, drawn from a Gaussian distribution:

$$u_1, \ldots, u_S | \tau_u \sim \mathcal{N}_S \left(\mathbf{0}, \tau_u^{-1} \mathbf{I}_S \right), \tag{5}$$

with precision parameter, τ_u. This prior uses `dpgrowmm` under `option = "mmi"`.

The `option = "mmigrp", ..., group = group` allows for a grouping of effect parameters within an MM term. For example, clients assigned to the (group CBT) treatment arm of the BRIGHT study are exclusively assigned to one of $G = 4$ groups to attend therapy sessions. The clients in each group do not receive treatment or "communicate" with clients in other groups such that the grouping divides clients into disjoint collections. Then `option = "mmigrp"` models a possible dependence among client effects in the same group, $g = 1, \ldots, G$ through the location parameter, η_g, where in our present example $\boldsymbol{u} = (u_1, \ldots, u_S)$ captures the set of MM effect parameters over S sessions. Each η_g is exclusively mapped to those sessions belonging to therapy group g by use of an $S \times G$ matrix, $\mathbf{R}$, under the following formulation,

$$\begin{aligned} u_1, \ldots, u_S | \tau_u &\sim \mathcal{N}_S \left(\mathbf{R}\boldsymbol{\eta}, \tau_u^{-1} \mathbf{I}_S \right) && (6) \\ \eta_1, \ldots, \eta_G | \tau_\eta &\sim \mathcal{N} \left(\mathbf{0}, \tau_\eta^{-1} \mathbf{I}_G \right). && (7) \end{aligned}$$

The matrix $\mathbf{R}$ is automatically created by **growcurves** based on the user input vector `group` defined in Section 3.

Choosing `option = "mmcar"` allows for spatial or adjacency-based prior dependence among the MM effects that permits a borrowing of strength across MM effects for estimation. Our

Gaussian CAR formulation follows Besag, York, and Mollié (1991), who enumerate a two-part form for the covariance matrix. Firstly, define an $S \times S$ adjacency matrix, $\mathbf{\Omega}$, to encode an adjacency dependence among neighboring effect covariate labels where we set $\omega_{ss'} \geq 0$ if session s is a neighbor of session s' (*e.g.*,, $s \sim s'$), and 0, otherwise, and ω_{ss} is defined to be 0. Secondly, construct $\mathbf{D} = \text{Diag}(\omega_{s+})$, where $\omega_{s+} = \sum_j \omega_{sj}$ equals the sum of the distances between session s and its neighbors and may be viewed as capturing the relative influence of session s. Then compose covariance matrix, $\mathbf{Q}^- = (\mathbf{D} - \mathbf{\Omega})^-$ of a multivariate Gaussian distribution (where $-$ denotes use of the Moore-Penrose pseudo-inverse since the rows of $\mathbf{Q}$ sum to 0 such that $\mathbf{Q}$ is rank degenerate). Hodges, Carlin, and Fan (2003) show that the rank of this covariance matrix is $S - G$, where G represents the number of distinct groupings or "islands" of MM effects as discussed for `option = "mmigrp"`. Despite the improper joint distribution, this construction specifies a set of proper Gaussian full conditional distributions,

$$u_s | \boldsymbol{u}_{-s}, \tau_u, \mathbf{\Omega} \sim \mathcal{N}\left(\bar{u}_s, (\tau_u \omega_{s+})^{-1}\right), \tag{8}$$

where $\bar{u}_s = \sum_{j \neq s} \omega_{sj} u_j / \omega_{s+}$ is the average of the MM effects for the neighbors of effect s. The adjacency matrix is input as matrix object to `dpgrowmm(..., Omega = Omega, ...)` under `option = "mmcar"`. The precision parameter receives, $\tau_u \sim \text{Gamma}(a, a)$, under all prior options for $\mathbf{u}$, where a is input under `dpgrowmm` with `strength.mm`. If not set by the user, `strength.mm` defaults to 0.1, a value under which the prior is readily changed by the data in estimating the posterior distribution (Banerjee, Wall, and Carlin 2003). It bears noting that all models under `dpgrowmm` employ a DP prior on subject effects, $\mathbf{B}$, such that `dpgrow(..., option = "dp")` may be viewed as a reduced model subset of function `dpgrowmm`. For both `"mmigrp"` and `"mmcar"` options, we input the group structure identifiers for the S effect terms with a numeric or character vector to `dpgrowmm(..., group = group, ...)`. If there is no grouping structure among the sessions (for `option = "mmi"`), this option may be omitted and defaults to a vector of 1's.

4.3. Illustration on BRIGHT data

Producing growth curves

We now illustrate estimation and inference on the model of Equation 1 by loading the BRIGHT data included with **growcurves** that includes the BDI-II response, as well as `(subject, trt, time)` predictors and the MM weight matrix (`W.subj.aff`) with rows identified by affected subjects (`subj.aff`) and group structure the MM session effects with (`group`). The associated $S \times S$ adjacency matrix, `Omega` that will be needed to employ the CAR prior on the MM session effects, is in the BRIGHT data. We load the data with,

```
R> data("datbrghtterms", package = "growcurves")
R> dat <- datbrghtterms
```

where `trt = 1` specifies CBT and `trt = 0` captures UC. We could have also chosen to use character entries, such as {`"cbt"`, `"uc"`}. Study clients are labeled in `subject` and take a value in `1:299`. The `time` field is measured in months, consistent with repeated subject measures taken at $(0, 3, 6)$ month intervals where 0 represents baseline; for example, the first 9 entries for these data vectors are listed. We add a reminder that the data are not required

to be balanced such that subjects may hold observations for varied number of time points or measurement waves.

```
    y subject trt time
1  24       1   1    0
2   3       1   1    3
3   9       1   1    6
4  43       2   1    0
5  17       2   1    3
6   0       2   1    6
7  37       3   1    0
8  31       3   1    3
9  11       3   1    6
```

We first employ an MM model with a CAR prior on the MM random effects using BRIGHT study data. Required inputs include the $(n_{\text{aff}} = 132) \times (S = 245)$ single MM matrix, `dat$W.subj.aff_mat` defined in Section 3, where each row is linked, in order, to a client listed in `subj.aff` and records a positive weight in the columns associated to those sessions attended by that client (if they are among those clients assigned to `trt = 1`) and includes 0 values in the remaining columns (for sessions not attended) such the row entries sum to 1. **growcurves** does not require the rows of `W.subj.aff` to sum to 1. It only requires that a numeric value links each client to a particular session they attend. One recalls that we normalize the rows to sum to 1 because we intend that the treatment fixed effects reflect the magnitude of the treatment effect, while the MM term expresses dependence among clients induced by the study design. As each of the $G = 4$ groups are non-communicating, the resulting structure of input `matrix` object, the structure of `W.subj.aff = dat$W.subj.aff_mat` is block diagonal for session-attending treatment arm clients. We specify estimation function `dpgrowmm` under a CAR prior for the MM session effects with `option = "mmcar"`. A CAR prior is selected for our primary model as we expect adjacency in sessions to exhibit a dependence such that we may borrow strength for estimation of their effects. We must additionally input an $S \times S$ CAR adjacency matrix, `Omega = dat$Omega_mat`, whose form is discussed in Section 4.2, and has $0's$ on the diagonals and $1's$ in those cells where sessions are adjacent or are neighbors. The CAR prior implemented in `dpgrowmm` is targeted to models that employ a single MM term. The `dpgrowmm` function allows input of an $S \times 1$, `group = dat$group_mat` group identifier for each session to account for the sub-grouping structure of these study sessions into disjoint sets of $G = 4$ groups that is then used to set the degrees of freedom for the posterior distributions for session effects, $\mathbf{U}$, and CAR precision parameter, τ_u.

```
R> MMCAR  <- dpgrowmm(y = dat$y, subject = dat$subject, trt = dat$trt,
+    time = dat$time, n.random = 3, n.fix_degree = 2,
+    Omega = dat$Omega_mat, group = dat$group_mat,
+    subj.aff = dat$subj.aff_mat, W.subj.aff = dat$W.subj.aff_mat,
+    n.iter = 40000, n.burn = 15000, n.thin = 10, option = "mmcar")

[1] "Your chosen option  = mmcar"
```

We next employ the simpler prior formulation of Equation 6 that extends the independent Gaussian prior assumption of `option = "mmi"` to account for group dependence by employing unique by-group means with `option = "mmigrp"` with the above syntax to deliver object `MMIGRP`.

Assess the relative contribution of the MM term by next employing models that exclude an MM term. We employ `dpgrow`, but with `option = "dp"` that reduces the effective number of parameters by borrowing strength across clients for their estimation.

```
R> DP  <- dpgrow(y = dat$y, subject = dat$subject, trt = dat$trt,
+    time = dat$time, n.random = 3, n.fix_degree = 2, n.iter = 40000,
+    n.burn = 15000, n.thin = 10, option = "dp")

[1] "Your chosen option  = dp"
```

We select a baseline comparator model under which subjects' random effects receive an independent joint Gaussian prior by utilizing estimation function `dpgrow` with `option = "lgm"` and `n.random = 3` which provides weak identification (through the variance of the prior) since our data only employ $T = 3$ measurement waves. We name the returned object, `LGM`.

We may now employ our returned objects, `MMCAR, MMIGRP, DP, LGM`, in our plot functions to perform graphical analyses and return one or more '`ggplot`' objects and the associated `data.frame` tables used to produce them. In R, the `data.frame` object is a special case of a `list` object where each field or column in a record may be of a distinct data type (e.g., numerical, character, factor), so it is a convenient way of collecting related information across data types. The first goal enumerated in Section 1 to render denoised growth curves is addressed by performing an analysis of the difference in by-subject growth curve shapes and orientations between CBT and UC BRIGHT study arms. We do so with the **growcurves** plot function,

```
R> gc <- growplot(object = MMCAR, compare.objects = list(LGM = LGM, DP = DP),
+    main.label = "MMCAR")
```

which renders two plots. Figure 2 employs a non-parametric loess smoother line through the by-subject posterior mean growth curves aggregated to each treatment arm, (0 = UC, 1 = CBT), and reveals a notable treatment effect focused on an increased attenuation of depressive symptoms between the 0 and 3 month measurement time intervals. We may also examine the characteristics of growth curves for selected subjects. To do so, we use the `growplot` plot function with option, `subjects.subset`, that contains a vector of subjects for which we desire to plot *their* individual growth curves. In this example, we *exclude* the `subjects.subset`, in which case a *random draw* of subjects is selected for by-subject growth curve plotting.

The curves in Figure 3 are smooth because our method performs within-client predictions at multiple time points constrained to the 0–6 month post-treatment measurement window for our BDI depressive symptoms response scores. The curve for subject 283 in the treatment arm reveals that `MMCAR` maps convex (downward facing curves, or bell-shaped) curves more robustly than the `LGM` model. Subject 40 in the CBT arm also demonstrates that `MMCAR` is more shape and orientation adaptive than `LGM`. In particular, one notices that the `LGM` model always estimates concave or U-shaped growth curves because it restricts the client effects to

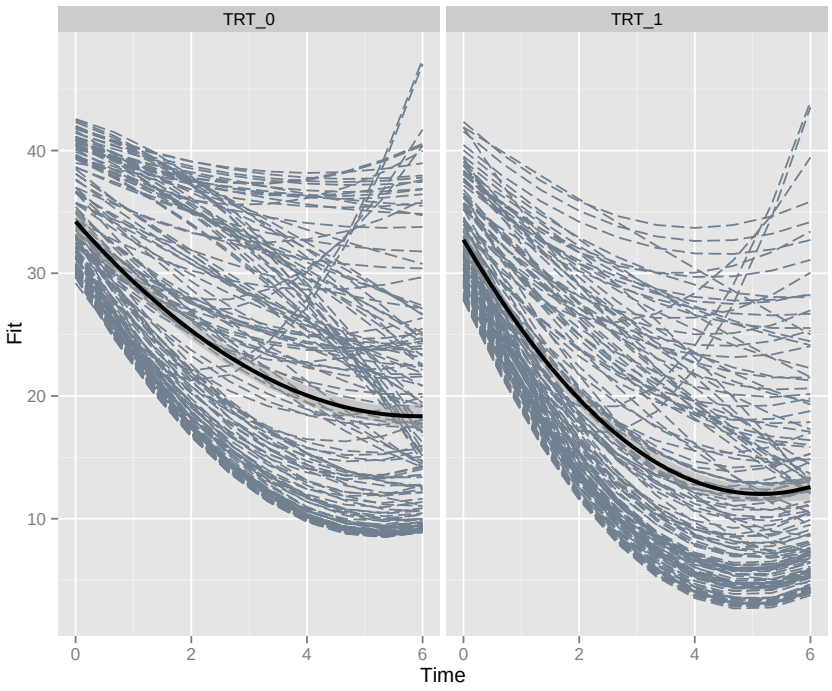

Figure 2: Growth curves aggregated by treatment arm for CAR prior on MM effects.

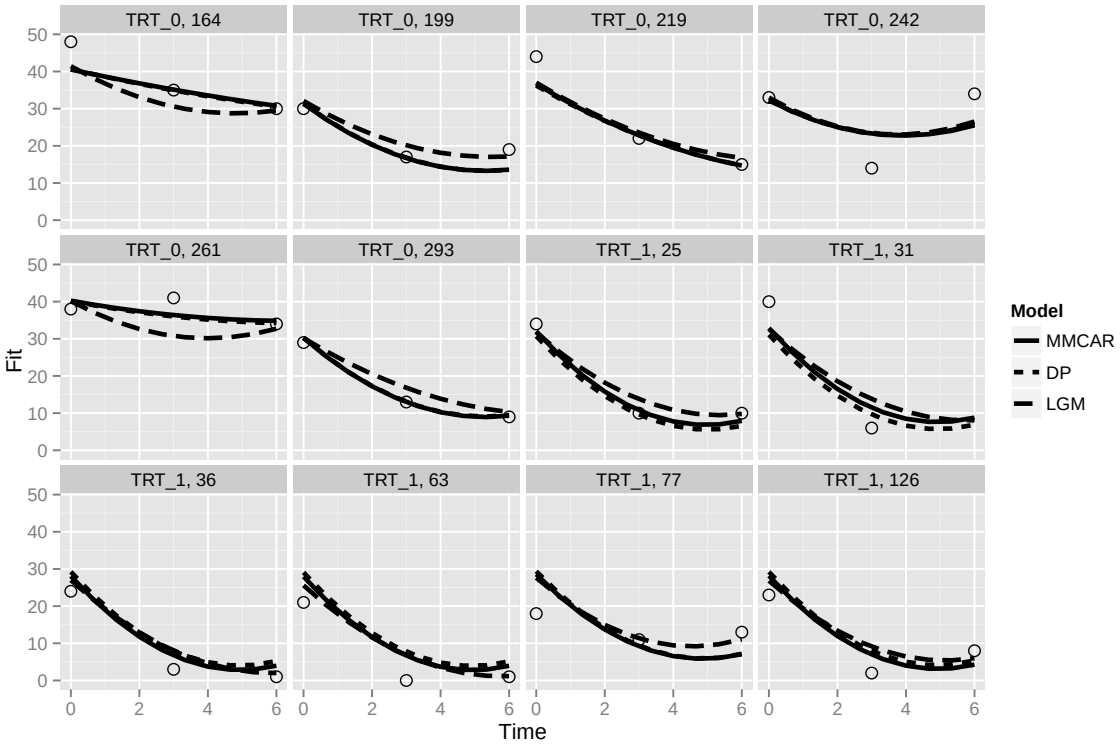

Figure 3: By-subject growth curves comparing MM term under CAR prior to a standard LGM.

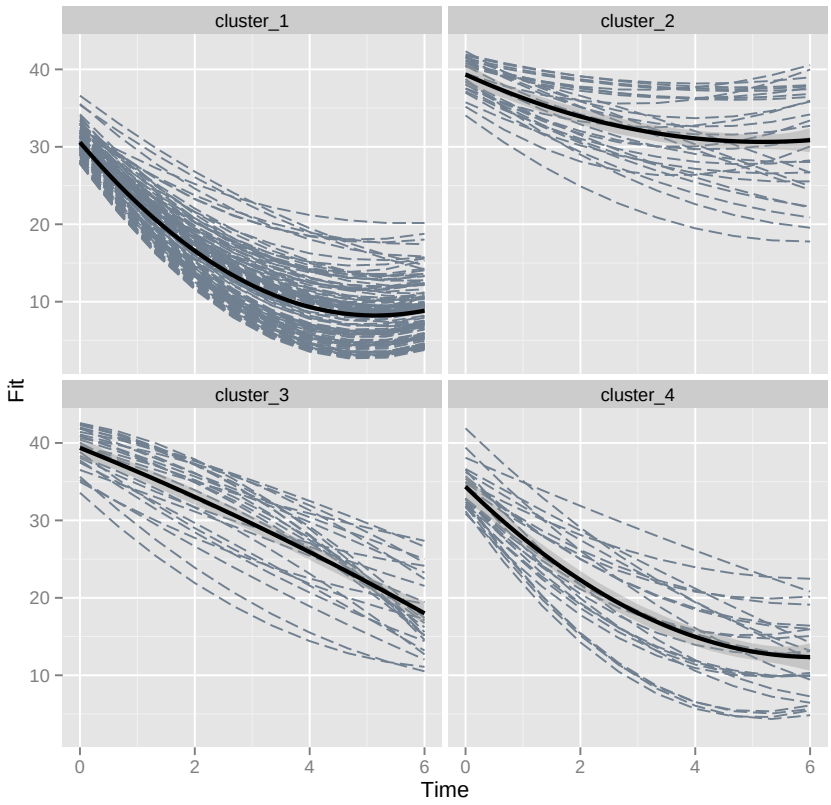

Figure 4: Growth curves aggregated by best-fit cluster under DP prior on subject effects.

be estimated from a common group as contrasted with the *DP* prior used for client effects in the `MMCAR` prior that allows for client effects to draw from multiple clusters, where some clusters may have convex growth curves and others, concave.

We may further utilize `growplot` to examine our model outputs to learn the drivers for this enhanced growth curve adaptability with `MMCAR`. Employ a feature of `growplot` that allows the user to compose a grouping of the by-client growth curves in any chosen fashion through input of `groups.plot` and `subjects.plot`. The former inputs a vector defining a user-defined grouping structure for the growth curves of those clients listed in the latter (such that both input vectors are of the same length). The defaults are to group all clients within treatment arms used for modeling. The DP prior on the subject effects under `option = "mmcar"` allows for a borrowing of strength among clients through clustering their random effect parameters, $\mathbf{b}_i$, to better explore the space of possible growth curve shapes than the usual parametric alternative. We examine whether this feature is responsible for the more adaptive performance of `MMCAR` by using the best fit clustering selected from the least squares clustering algorithm reviewed in Section 4.1 to group our by-subject growth curves with the following script,

```
R> cluster <- samples(MMCAR)$bigSmin
R> c.sizes <- sapply(cluster, length)
R> clusterstoplot <- sort(c.sizes, decreasing = TRUE,
+    index.return = TRUE)$ix[1:4]
```

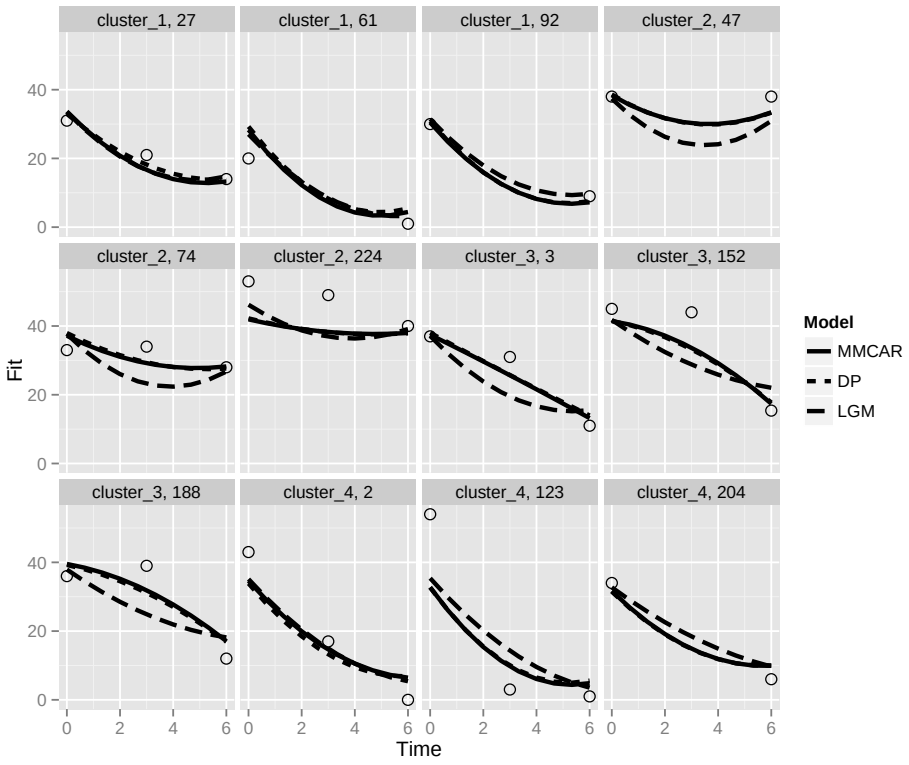

Figure 5: By-subject growth curves, labeled with cluster memberships, comparing MM term under CAR prior to a standard LGM.

```
R> map <- vector(mode = "list", length = length(clusterstoplot))
R> for(i in 1:length(clusterstoplot)) {
+    cluster.i <- cluster[[clusterstoplot[i]]]
+    map[[i]] <- as.data.frame(cbind(cluster.i,
+      paste("cluster", i, sep = "_")), stringsAsFactors = FALSE)
+    names(map[[i]]) <- c("subject", "group")
+  }
R> map <- do.call("rbind", map)
```

where `bigSmin` is a list object of length equal to the number of unique clusters formed that we denote as M, and each list element contains a vector of client identifiers assigned to the cluster represented by that element. The above script selects the 4 largest clusters with size defined by number of client members. The `data.frame` object, `map`, provides us with the subset of clients and their cluster identifiers that we will use to set the `growplot` option. Another instantiation instance of `growplot` is created with,

```
R> gc.2 <- growplot(object = MMCAR,
+    compare.objects = list(LGM = LGM, DP = DP), subjects.plot = map$subject,
+    groups.plot = map$group, main.label = "MMCAR")
```

We see in Figure 4 that cluster 1 captures many of the clients with deeper concave growth curves, confirming our expectation that the basis for clustering is primarily shape and ori-

Model	$\bar{D}$	$-LPML$	DIC_3
LGM	5666	3031	5841
DP	5520	2989	5691
MM(CAR)	5501	2982	5675

Table 3: Model fit comparisons: $\bar{D}$, $-LPML$ and DIC_3 scores for model alternatives. Lower values imply better performance.

entation of curves such that subjects under similarly shaped curves borrow strength in co-estimation. While the effect is not as strong, we also observe a grouping of convex-shaped curves in cluster 2. The clustering by curve shape and orientation indicate that our posterior distribution over the subject effects, $\mathbf{B}$, better spans the space of possible curve shapes and helps explain the improved adaptability we observe with model `MMCAR`. Figure 5 presents a set of by-client growth curves as in Figure 3, only now the labeling and random draw of the clients reflects the new grouping of clients. All plot functions may be assigned to objects in order to recover the '`ggplot`' plot objects and associated `data.frame` objects. They may also be invoked without assignment, in which case the plots are rendered, but the return objects are suppressed.

Recall that we may extract fit statistics from function `summary(out)`, where `out` is an object holding the results from any of the estimation functions. Fit statistics include `DIC3` of Celeux *et al.* (2006), a special case of the deviance information criterion (DIC) of Spiegelhalter, Best, Carlin, and van der Linde (2002), that marginalizes over the parameter space such that it is more robust for mixture models (that includes any model employing a DP prior on random effects) and the log-pseudo marginal likelihood, `lpml` referred to in Congdon (2005), that imposes a stronger model complexity penalty in practice than does `DIC3` by composing the marginal distribution in a leave-one-out fashion. The unpenalized model deviance, `Dbar`, is also available. We present results for the 3 comparison models run with the BRIGHT data in Table 3, where a lower value indicates a better fit for the included fit statistics, and we see that the improved adaptability of the estimated growth curves from employing both client and session effects terms with `MMCAR` produces the best fit.

Analysis of treatment effects

We continue our analysis of object `MMCAR` by testing for a statistically significant differential treatment effect between the CBT and UC arms, which we labeled as the second goal for analysis of repeated measures data in Section 1. Recall that the fixed effects include a CBT treatment indicator, d_i, along with interactions with time, t_{ij}, up to the chosen polynomial order in option `n.fix_degree = q`, where q may be any integer value. We assign `n.fix_degree = 2`. The treatment effects are estimated from predictive margins (Lane and Nelder 1982) where we predicted outcomes as if all clients were in CBT, then predicted outcomes as if all clients were in UC, for each MCMC sample at a chosen time point.

The plot function, `trtplot`, allows us to access this distribution under any desired set of model outputs returned from our estimation functions at a selected set of time points,

```
R> run.objects <- list(LGM = LGM, DP = DP, MMCAR = MMCAR, MMIGRP = MMIGRP)
R> run.models <- c("LGM", "DP", "MMCAR", "MMIGRP")
R> trt.labs <- c(0, 1)
```

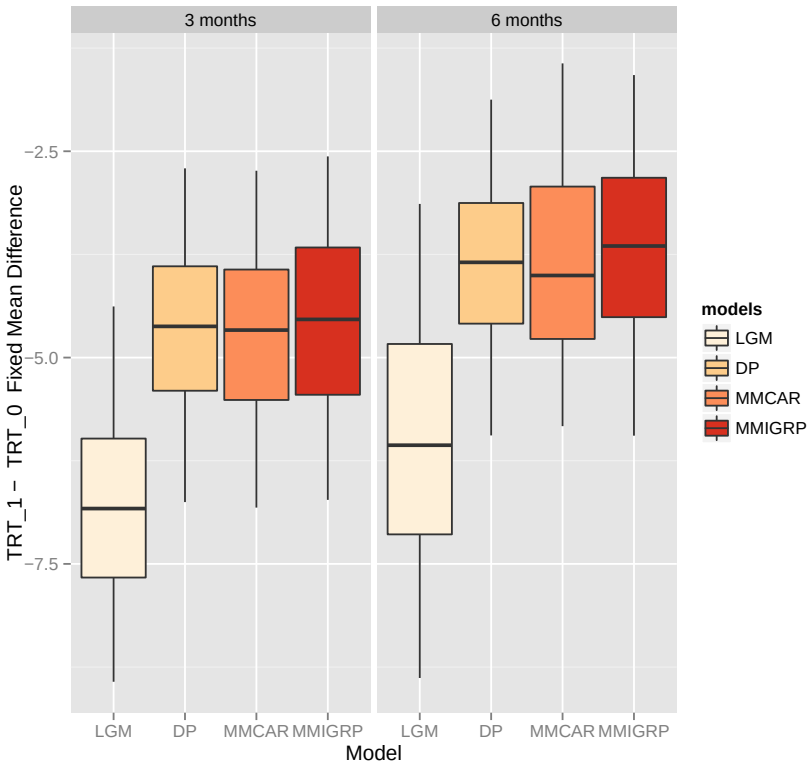

Figure 6: 95% credible intervals for difference between CBT and usual care study subjects.

```
R> time.points <- c(3, 6); time.labels <- c("3 months", "6 months")
R> tp <- trtplot(run.objects = run.objects, run.models = run.models,
+    trt.labs = trt.labs, time.points = time.points,
+    time.labels = time.labels)
```

The mean difference in Figure 6 is located at notably higher ranges of BDI values for the `MMCAR` and `DP` objects, both of which employ a DP prior on the sets of client effects, than for the `LGM` object, which does not. In our analysis of subject growth curves, `LGM` is unable to model the sub-population of subjects who express convex or bell-shaped growth curves, assigning them concave or U-shaped curves. A concave curve produces a more negative fixed effects slope estimate than a convex curve. The result is that the treatment effect for `LGM` is larger than for `DP` and `MMCAR`. Therefore, a more robust modeling of subject growth curves avoids a source of bias expressed in `LGM` estimation of treatment effects. While our case analysis focuses on two treatment arms, all estimation functions and associated plot functions are flexible to handle any number of arms. In particular, `trtplot` may be employed to compare any two treatment arms.

Analysis of MM effects

Although we have employed an MM term of session effects in `MMCAR` and `MMIGRP` to capture the dependence structure among clients, we may also realize an interpretative value by directly studying the posterior summaries of the session random effects to compare their relative effectiveness in reducing depressive symptoms, which is the aim of the goal 3 enumerated in

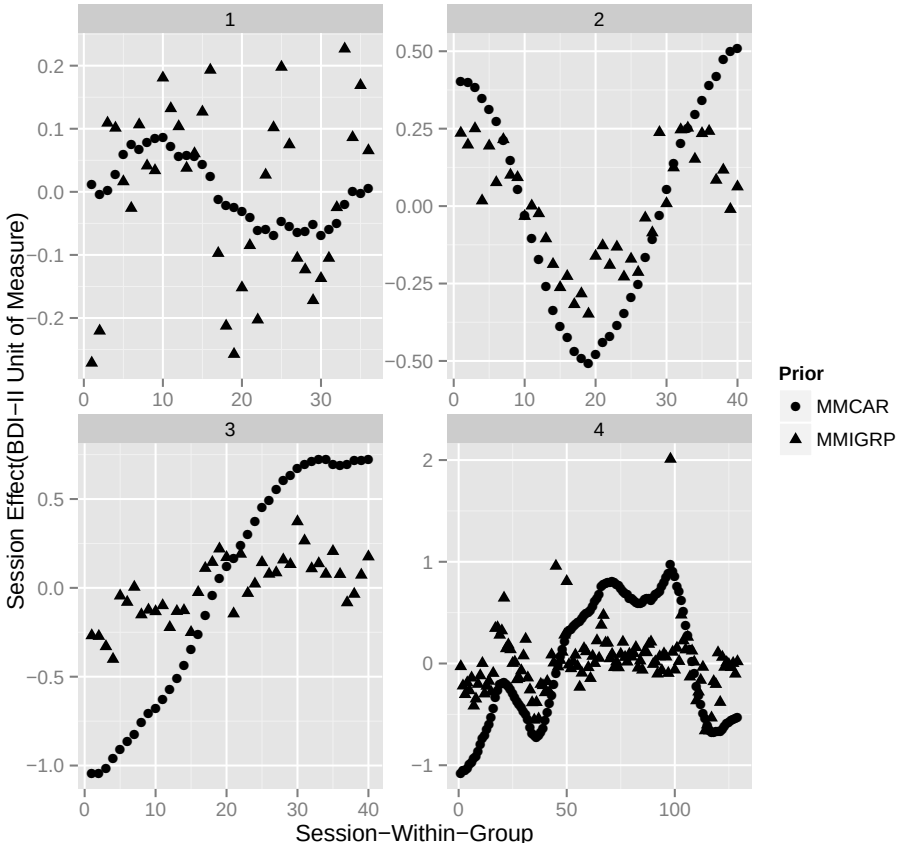

Figure 7: MM effects: posterior mean estimates under varied prior formulations.

Section 1. We employ the plot function `effectsplot` that composes overlaid plots of session effect posterior mean values from one or more prior formulations for the MM term, with the plots faceted by disjoint groups, $g \in (1, \ldots, G)$. As with all discussed plot functions, an associated `data.frame` object used to compose the plots is also returned. We first compare the `MMIGRP` independent prior formulation with by-group means to the `MMCAR` prior that allows for adjacency or spatial dependence among the session effects.

```
R> run.objects <- list(MMCAR = MMCAR, MMIGRP = MMIGRP)
R> prior.labs <- c("MMCAR", "MMIGRP")
R> axis.labs <- c("Session-Within-Group",
+    "Session Effect(BDI-II Unit of Measure)")
R> ep.1 <- effectsplot(objects = run.objects, prior.labs = prior.labs,
+    axis.labs = axis.labs, center = TRUE)
```

Figure 7 aggregates the posterior mean estimates for session effects within each of the 4 groupings of treatment arm clients. We observe how the `"mmcar"` prior construction borrows strength across adjacent sessions, producing estimates that are generally similar to those under `"mmigrp"`, but that provide heightened interpretative value by *smoothing* over `"mmigrp"` results. This clarification under `"mmcar"` may be seen for groups $2, 3$ and 4, in particular.

4.4. Including two or more MM terms

So far our definition of an MM term was restricted to the linear product of a single $N \times S$ MM weight matrix, W, and the associated vector of $S \times 1$ random effects, $\mathbf{u}$, representing all of the S sessions employed for CBT treatment as employed in Equation 1 where N denotes the number of observed subject-time cases. One may find it useful to extend Equation 1 with employment of additional MM terms under two scenarios,

1. There are multiple classes of unrelated exposures under which a subject's weighted exposure pattern in each MM sums to 1. An education example is an experimental design that links a set of teachers to students across years with MM terms indexed by year. A given student may link to a subset of MM terms over the multiple years (Mariano, McCaffrey, and Lockwood 2010).

2. Sub-groupings of clients receive disjoint or non-communicating patterns of exposure (such that modeling under a single MM matrix produces a block diagonal structure). The disjoint collection of groups to which treatment arm clients are assigned for the BRIGHT study offers an example of such a non-communicating exposure pattern. Each subject is constrained to belong to exclusively one MM term of those employed. While this pattern may be addressed under a single MM term (e.g., by using the `"mmigrp"` modeling option in the `dpgrowmm` estimation function), separately modeling each disjoint exposure pattern with its own MM term allows separate prior specifications for each term. We will review this second use of disaggregating disjoint groups into separate MM terms in the BRIGHT case study example that follows.

We perform the extension of the single-MM term formulation of Equation 1 to allow two or more MM terms with,

$$y_{ij} = \alpha + \mathbf{x}_{ij}^{\top}\boldsymbol{\beta} + \mathbf{z}_{ij}^{\top}\mathbf{b}_i + (\mathbf{w}_{ij,1\{i\}})^{\top}\boldsymbol{\gamma}_{q\{i\}} + \ldots + (\mathbf{w}_{ij,L_i\{i\}})^{\top}\boldsymbol{\gamma}_{L_i\{i\}} + \epsilon_{ij}, \tag{9}$$

where $(1\{i\}, \ldots, L_i\{i\})$ denotes those MM terms linked to subject i for a total of $L_i \leq L$ terms and here L denotes the total number of MM terms employed over all subjects. For example, if subject 5 belongs to terms $(2, 6, 7)$ of $L = 10$ terms, then $1\{5\} = 2$, $2\{5\} = 6$, and $3\{5\} = 7$ with $L_5 = 3$. Each MM term, $\mathbf{W}_\ell$, $\ell = 1, \ldots, L$, is an $N_\ell \times S_\ell$ MM matrix specialized to term ℓ that maps S_ℓ effects (where $\sum_{\ell=1}^{L} S_\ell = S$) to N_ℓ repeated measure cases for n_ℓ subjects. Similarly, $\boldsymbol{\gamma}_\ell$ represents an $S_\ell \times 1$ vector of MM (e.g., session) effects associated to MM term ℓ.

The construction for employing more-than-one MM term is modeled with estimation function, `dpgrowmult`, but now inputs `subj.aff` and `W.subj.aff` as *lists* of vector and matrix objects, respectively, with the lengths both `list` objects equal to the number of MM terms. A prior formulation for each MM term is input in a similar manner where now `option` holds a vector of choices equal in length to the utilized number of MM terms. The available option choices include, `("mmi", "mmigrp", "mmcar", "mmdp")`. Each MM term is assumed independent under Equation 9, though the prior constructions for each term may allow the data to learn a dependence structure among the effects defined within that term.

The "mmdp" option employs a DP prior over the effects within the applicable MM term,

$$u_1, \ldots, u_S | H \overset{\text{iid}}{\sim} H \tag{10}$$
$$H | c, H_0 \sim \text{DP}(c, H_0) \tag{11}$$
$$H_0 \equiv \mathcal{N}\left(0, \tau_u^{-1}\right), \tag{12}$$

to allow the data to learn any dependence among the effects. Inputs for Omega and group are now also list objects, the former of length equal to the number of terms under "mmcar" and the latter the number of terms under either of "mmcar" or "mmigrp" prior formulations.

The non-overlapping client attendance patterns in the BRIGHT study across the $G = 4$ groups allow us to model each group with employment of a distinct MM term as an alternative to the block-diagonal construction in a single MM weight matrix employed for W.subj.aff used under a single MM term. We accounted for the group structure of the BRIGHT study CBT arm in the adjacency matrix for the "mmcar" option or with the employment of by-group random mean parameters under "mmigrp". We now extract each block from the block-diagonal structure of the MM weight matrix used under the dpgrowmm into its own MM term in dpgrowmult. We specify $G = 4$ MM terms, one for each group, where the effects for each term receive their own prior specifications. For example, if a "mmcar" or "mmi" prior formulation is specified for each term, then the associated precision parameters, τ_{u_g}, $g = 1, \ldots, G$, will be specific to each term (unlike with the single MM term construction which employs 1 precision term, τ_u). So we can model dependence among the sessions *within* each MM term by using "mmcar" or "mmdp" prior formulations, though the sessions *between* the MM terms are assumed to be independent.

Illustration on BRIGHT data

Our definition of 4 MM terms produces dat$subj.aff and dat$W.subj.aff as list objects, each holding 4 elements. We may view the columns of dat$W.subj.aff[[g]] as a collection of sessions nested under group g. For example, dat$W.subj.aff[[4]] is of dimension 78×129, reflecting the 78 CBT clients who attend any of the 129 unique sessions in this group. Similarly, dat$subj.aff[[4]] is a vector of length 78 which captures the specific client identifiers assigned to group 4. We may mention, however, that the use of 4 MM terms to model the $S = 245$ session effects is about 20% more computationally expensive under our BRIGHT case study runs.

Results displayed in Figure 7 from the modeling of BRIGHT CBT groups under a single MM term (with dpgrowmm) reveal that session effects across the groups express a wide range of values and dependence patterns, suggesting more flexibility might be warranted. We label the object output from running dpgrowmult with MMDCAR, where the "D" denotes the BRIGHT data are modeled under "disjoint" MM terms. Each group is now represented in its own MM term that specifies a prior formulation. In our illustration, we employ two Gaussian prior formulations; CAR ("mmcar"), for groups 2 and 4, and independent ("mmi"), for groups 1 and 3, by setting option = c("mmi", "mmcar", "mmi", "mmcar"). We construct Omega and group as list objects that use the associated constructions only for the groups (2 and 4) receiving the CAR prior as these objects are not required for the independent prior formulations in the following script,

```
R> Omega <- list(dat$Omega[[2]], dat$Omega[[4]])
R> group <- list(dat$group[[2]], dat$group[[4]])
```

```
R> MMDCAR <- dpgrowmult(y = dat$y, subject = dat$subject, trt = dat$trt,
+    time = dat$time, n.random = 3, n.fix_degree = 2, Omega = Omega,
+    group = group, subj.aff = dat$subj.aff, W.subj.aff = dat$W.subj.aff,
+    n.iter = n.iter, n.burn = n.burn, n.thin = n.thin, strength.mm = 1,
+    shape.dp = 4, option = c("mmi", "mmcar", "mmi", "mmcar"),
+    ulabs = c("group_1", "group_2", "group_3", "group_4"))

[1] "Your chosen set of MM term priors  = mmi mmcar mmi mmcar"
```

The option, `ulabs`, allows us to name each MM term such that model outputs, including plots, may reflect this labeling. For example,

```
R> u.summary <- summary(MMDCAR)$summary.results$u.summary
R> u.summary[["group_1"]]

$group_1
       mcmc.low     mcmc.mean mcmc.high
 [1,] -3.674902 -0.1608180162  2.814044
 [2,] -3.432991 -0.1195210419  3.002024

⋮

[36,] -3.052427  0.0554539194  3.475589
```

For comparative purposes, we replace the `"mmcar"` prior formulations for groups 2 and 4 in model run `MMDCAR` with non-parametric `"mmdp"` options in a model we label `MMDDP` to see if the greater flexibility of the non-parametric DP prior to discover dependence among sessions provides a different result from specification of CAR priors.

The `effectsplot` function also permits a comparison of effects in a single, target MM term, of possibly many, across multiple objects produced from `dpgrowmult`. The comparison for a focus term may simultaneously include objects generated from `dpgrowmm` that include all sessions for the single MM term or a selected group within the term (indicated with the `group` option) such that the number of sessions compared across all objects is equal. We illustrate such a comparison of `dpgrowmult` and `dpgrowmm` objects with another run of `effectsplot` focused on group 4, which holds both the largest number of clients and associated sessions. We select MM effects term 4 produced for `MMDCAR` and `MMDDP` using estimation function `dpgrowmult` and also choose `group = 4` for `MMCAR` that was earlier generated using estimation function `dpgrowmm`.

```
R> run.objects <- list(MMDCAR = MMDCAR, MMDDP = MMDDP, MMCAR = MMCAR)
R> mm.terms <- c(4, 4, 4)
R> prior.labs <- c("MMDCAR-G4", "MMDDP-G4", "MMCAR-G4")
R> axis.labs <- c("Session-Within-Group",
+    "Session Effect(BDI-II Unit of Measure)")
R> ep.2 <- effectsplot(objects = run.objects, mm.terms = mm.terms,
+    prior.labs = prior.labs, axis.labs = axis.labs, center = TRUE)
```

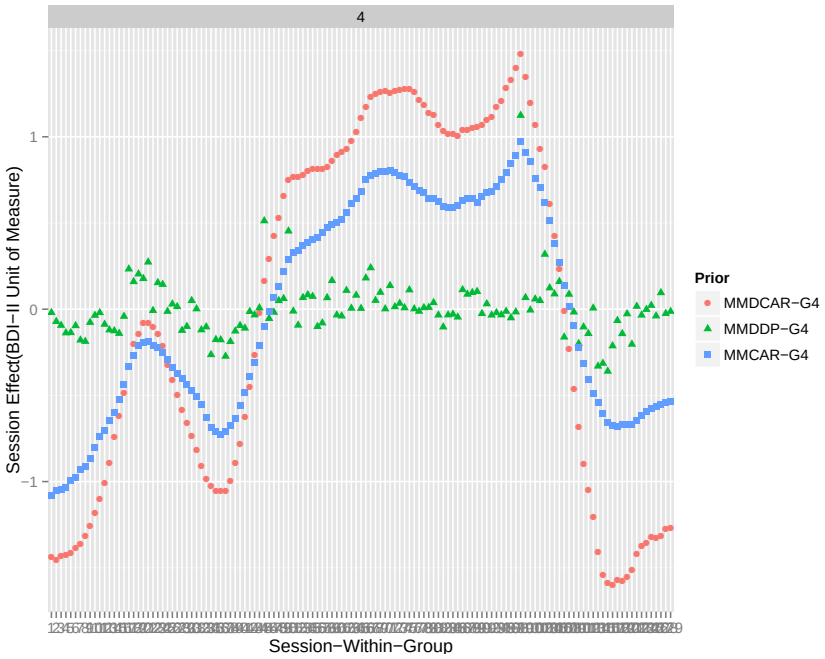

Figure 8: MM session effects: 1-term vs. 4-term model.

The option `mm.terms` identifies the focus MM term for an object generated from `dpgrowmult` or target group for an object of class '`dpgrowmm`' and may either contain the term or group name (set with the `ulabs` option when running `dpgrowmult` or `group` for `dpgrowmm`) or may contain a numerical value indicating the position order of the term or group. Here, we enter 4 for both the two '`dpgrowmult`' and '`dpgrowmm`' objects. The pattern of the MM term 4 posterior means for `MMDCAR`, produced under prior option `"mmcar"`, shown in Figure 8 appears similar to that for the 4th group in Figure 7 for `MMCAR` generated under a single MM term, only the effects for the middle sessions appear now pulled further apart for `MMDCAR`. The magnitudes are relatively higher across the session effects with `MMDCAR` to indicate the data are learning a stronger signal. We may examine the trace plots (not shown here) for the precision parameters, τ_{u_g}, produced for each term of the four terms specified in `MMDCAR` by invoking the statement, `plot(MMDCAR)`. These trace plots reveal that each of 4 MM terms express different values for τ_{u_g} to indicate the data use the flexibility from specializing τ_{u_g}, by CBT group, to discover differing strengths of first order dependence. The `-lpml` fit statistic improves from 2982 for `MMCAR` to 2975 for the `MMDCAR`. We conclude this section by noting that `MMDDP`, which employs a DP prior alternative to CAR for term 4, expresses a similar, though less discernable pattern before session 40 as compared to the other two alternative choices for the priors defined on the session effects, but otherwise appears to estimate relatively little dependence among the session effects.

We make brief mention that if the aggregation of effects across terms (`dpgrowmult`) or groups (`dpgrowmm`) are common for all objects modeled (as they are in this example), an additional plot is generated that compares all effect terms and groups when invoking `effectsplot`.

4.5. Multivariate MM random effects

The univariate MM random effects of Equation 8 may be replaced with a multivariate formulation where the effect of each MM random effect (such as a session effect) is allowed to vary with time.

$$y_{ij} = \alpha + \mathbf{x}_{ij}^{\top}\boldsymbol{\beta} + \mathbf{z}_{ij}^{\top}\mathbf{b}_i + \left(\mathbf{w}_{ij}^{\top}\mathbf{U}\right)\mathbf{z}_{ij} + \epsilon_{ij}, \tag{13}$$

where $S \times q$ matrix of session effects, $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_S)^{\top}$, so that row s defines a multivariate $q \times 1$, $\mathbf{u}_s$ set of random effects for session s that is multiplied by $\mathbf{z}_{ij} = (1, t_{ij}, t_{ij}^2)$ to allow the session effect, s, to vary as a polynomial function of time. We may most easily make the extension of the CAR formulation of Besag *et al.* (1991) by stacking each of the q, $S \times 1$ columns from $\boldsymbol{\Gamma}$ into $qS \times 1$, $\boldsymbol{\mathfrak{G}} = \left(\mathbf{u}_{(1)}, \ldots, \mathbf{u}_{(q)}\right)$ for the $S \times 1$, $\mathbf{u}_{(s)}$. Then compose the multivariate CAR prior,

$$\boldsymbol{\mathfrak{G}} | \boldsymbol{\Lambda}, \boldsymbol{\Omega} \sim \mathcal{N}\left(\mathbf{0}, [(\mathbf{D} - \boldsymbol{\Omega}) \otimes \boldsymbol{\Lambda}]^{-}\right), \tag{14}$$

for $qS \times qS$ precision matrix, $\mathbf{Q} = (\mathbf{D} - \boldsymbol{\Omega}) \otimes \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ describes the dependence among the $q = 3$ polynomial order effects for each MM element or session.

Our implementation restricts the multivariate session effects to be of the same form as the $q \times 1$ subject random effects, $\{\mathbf{b}_i\}$. The $S \times q$ multivariate MM effects, $\mathbf{U}$, are multiplied by the $1 \times S$ client i row for observation j of the MM weight matrix, $\mathbf{W}$, in Equation 13 to produce a $1 \times q$ product that is then multiplied by the $q \times 1$ random effects design vector, $\mathbf{z}_{ij} = (1, t_{ij}, t_{ij}^2)$.

The multivariate MM (session) effects term is invoked with `dpgrowmm(..., multi = TRUE, ...)` (where `multi = FALSE` is the default, which produces a univariate MM (session) effects model). There are two available options for prior specification of the $S \times S$ covariance matrix generating the columns of $\mathbf{U}$, which are set with `option` in a very similar fashion as for Equation 8. The two options are `c("mmi", "mmcar")`. A Wishart prior, $\boldsymbol{\Lambda} \sim \mathcal{W}((q+1), \mathbf{I}_q)$, where $\boldsymbol{\Lambda}$ is the $q \times q$ covariance matrix for the rows of $\mathbf{U}$, completes the prior specification for this model. This prior configuration implies a set of marginally uniform priors on the correlations and we see little sensitivity to alternative specifications in lieu of $\mathbf{I}_q$ for the mean of the Wishart distribution, possibly because q (that specifies the polynomial order) tends to be small for our data applications.

Illustration on BRIGHT data

We next generalize the single MM term to $q = 3$ polynomial effects indexed by time for each MM element. In what follows, we aggregate the 245 sessions from the BRIGHT study into $S = 61$ modules that collect sessions. Each module is associated to a particularly therapeutic topic and collects 4 sessions under that topic. In all, $S = 61$ CBT modules were offered to clients. These 61 modules were divided into the same $G = 4$ CBT open-enrollment therapy groups used to collect the underlying sessions. The number of modules for each of these four groups was 9, 10, 10, and 32 and number of clients enrolled in each open-enrollment group was 17, 21, 19, and 83, respectively. Each client attended modules of only one of the four open-enrollment groups. Other than coarsening our employed MM effects from session to modules, we employ the same BRIGHT study data with identical structure for the study design, but now we will focus on module effects (rather than session effects). The following code loads the BRIGHT study data for use with module MM effects. Then we run `dpgrowmm(..., multi`

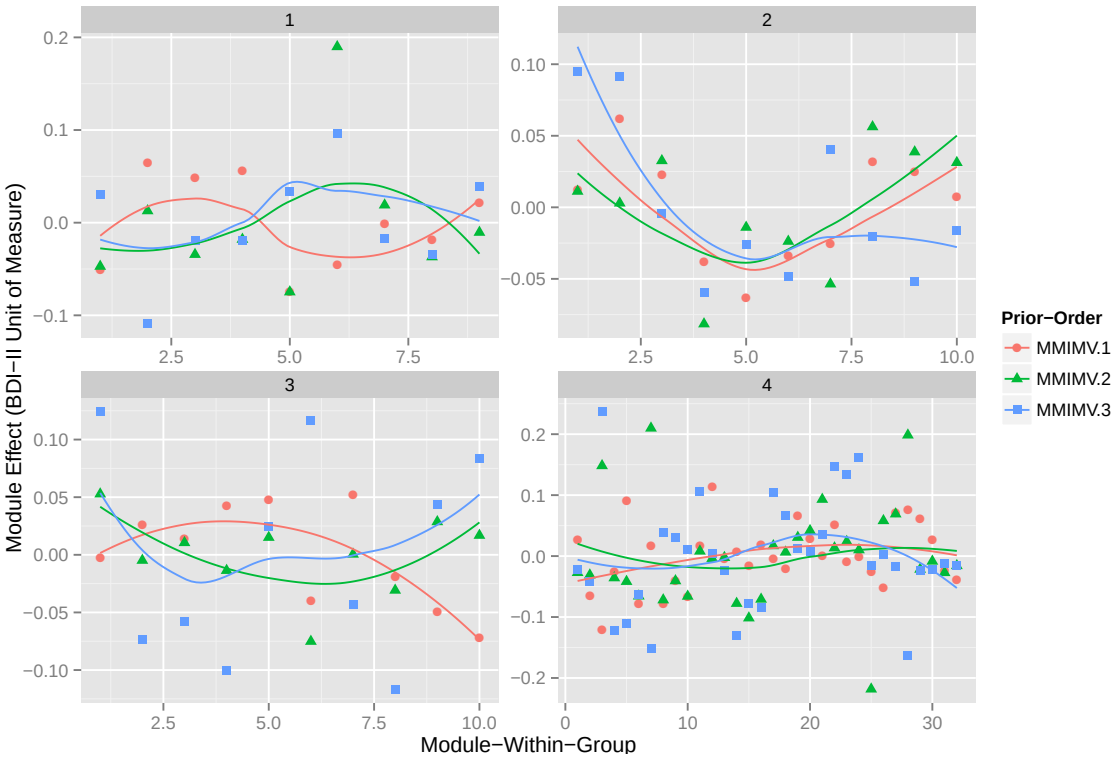

Figure 9: Multivariate module effects: posterior mean intercept (order 1), slope (order 2) and quadratic (order 3) effects for each module.

`= TRUE, n.random = q ...)` to invoke the q polynomial effects. Finally, we compose the effect plots, faceted by CBT group, which is rendered in Figure 9. We first load the data.

```
R> data("datbrghtmodterms", package = "growcurves")
R> dat <- datbrghtmodterms
```

We next run the estimation function

```
R> MMIMV <- dpgrowmm(y = dat$y, subject = dat$subject, trt = dat$trt,
+    time = dat$time, n.random = 3, n.fix_degree = 2,
+    subj.aff = dat$subj.aff_mat, group = dat$group_mat,
+    W.subj.aff = dat$W.subj.aff_mat, multi = TRUE, n.iter = n.iter,
+    n.burn = n.burn, n.thin = n.thin, shape.dp = 4, strength.mm = 1.5,
+    plot.out = TRUE, option = "mmi")

[1] "Your chosen option = mmi for multivariate MM effects"

R> orderto  <- list(1:9, 1:10, 1:10, 1:32)
R> run.objects <- list(MMIMV = MMIMV)
R> prior.labs <- "MMIMV"
R> axis.labs <- c("Module", "Effect Value")
R> ep.mv <- effectsplot(objects = run.objects, prior.labs = prior.labs,
+    axis.labs = axis.labs, smoother = TRUE, orderto = orderto)
```

5. Combined subject-by-MM effects term

The MM models discussed earlier (which are invoked using the `dpgrowmm` and `dpgrowmult` estimation functions) employ *separate* subject and MM terms. The use of separate terms assumes *no* interaction between them such that the subject and MM effects are independent. We will next reformulate Equation 13 that uses a single term with a new formulation to explicitly index the subject random effects by treatment module (for the BRIGHT case study illustration). Each subject will receive their own set of module random effects where each subject is assigned a $q \times (S+1)$ matrix of random effects, in contrast with the set of $q \times 1$, $\{\mathbf{b}_i\}_{i=1,\ldots,n}$ subject random effects and the $S \times 1$ set of MM module effects, $\mathbf{u}$, specified under Equation 13. So each subject in our BRIGHT study application will have their own set of S module effects. Our revised formulation will allow us to explore clusters of subjects whose responses vary across treatment modules. We re-formulate Equation 13 in a more flexible composition,

$$y_{ij} = \alpha + \mathbf{x}_{ij}^\top \boldsymbol{\beta} + \mathbf{z}_{ij}^\top \boldsymbol{\Delta}_i \mathbf{w}_i + \epsilon_{ij} \tag{15}$$

$$\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_n \big| F \overset{\text{iid}}{\sim} F \tag{16}$$

$$F \big| F_0 \sim \text{DP}\left(c, F_0\right). \tag{17}$$

Our construction for $\mathbf{w}$ is equivalent to the MM weight vector of Equation 13, plus an intercept term to incorporate subject mean effects. The $(S+1) \times 1$, $\mathbf{w}$ is composed of values in $[0, 1]$ with $\sum_{s=1}^{S} w_s = 1$ for subjects who are linked to at least one module, and $\sum_{s=1}^{S} w_s = 0$ for subjects who are not. We may see intuitively how (16) expands the DP formulation of (2) by indexing subject effects by modules attended with,

$$\boldsymbol{\Delta}_i = \left[\mathbf{b}_i, \mathbf{a}_{1,i}, \ldots, \mathbf{a}_{S,i}\right], \tag{18}$$

where $\boldsymbol{\Delta}_i$ is composed of a $q \times (S+1)$ set of random effects specific to subject i that includes a $q \times 1$ vector, $\mathbf{b}_i$, and S sets of $q \times 1$ effects, $\{\mathbf{a}_{s,i}\}$, specified for group therapy modules, $s = 1, \ldots, S$. So this formulation estimates a full set of S module effects for *every* client in both the UC and CBT arms, regardless of the particular modules they attended (or not at all). The parameter matrix, $\boldsymbol{\Delta}_i$, is back multiplied by the $(S+1) \times 1$ vector $\mathbf{w}$ that encodes the multiple membership link of subject i to their particular vector of modules they *did* attend. So only those effects, $\{\mathbf{a}_{s(k),i}\}_{k=1,\ldots,K}$, corresponding to the non-zero entries of $\mathbf{w}_i$ will impact the likelihood for $\mathbf{y}_i$, where $(s(1), \ldots, s(K))$ defines a sub-vector of $K \leq S$ MM modules. The way we may interpret the module effects for those modules not attended by a client is as the projected value of their response if they *had* attended those modules.

We refer to the resulting formulation as a "multiple membership dependent Dirichlet process" (MM DDP) because Savitsky and Paddock (2013) show that the construction of Equation 16 may be equivalently specified with a set of prior distributions for the subject random effects, $F_{\mathbf{w}}$, that are indexed by the unique MM weight vectors, $\mathbf{w}$, expressed in the data.

5.1. Illustration on BRIGHT data

We implement our combined model with estimation function,

```
ddpgrow(... numdose = dat$numdose, dosemat = dat$dosemat,
  Omega = dat$Omega, typetreat = c("car", "car", "mvn", "ind"), ...).
```

This function utilizes similar inputs as `dpgrow`, particularly for data structure inputs. Additional inputs include `numdose`, a vector of length equal to the number of distinct sets of MM modules, where each entry contains the number of modules associated to a set. A set contains a collection of modules that may represent a treatment or, in the case of the BRIGHT study data, a CBT therapy group. We specify `numdose = c(9, 10, 10, 32)` for the BRIGHT case study, where each entry represents the number of modules contained within a CBT therapy group and there are 4 total groups. The next input is `dosemat`, an $n \times (S+1)$ MM weight matrix that maps the n clients to S MM modules (across all treatments or groups). Each row of `dosemat` contains, $\mathbf{w}_i$, the MM attendance vector for client i, where $i = 1, \ldots, n$, and the first column is an intercept column of $1's$ to parameterize the client i mean random effect, $\mathbf{b}_i$. A hold-out module is required for identifiability. Under a study design that includes 2 or more interventions, the control arm may serve as the hold-out – which is equivalent to a module "0" or the null module – so that the row entry for all control arm clients is set to $(1, 0, \ldots, 0)$, where the $0's$ in all the module columns indicate a control arm client is not linked to (or did not attend) any modules. The base distribution prior, F_0, for each group of modules is conveyed in `typetreat`, which is of the same length as `numdose` and equal to the number of groups. There are 3 prior options from which to choose for each group, `"car", "mvn"` and `"ind"`. The `"car"` option implements a multivariate CAR formulation similar to Equation 14 used for the multivariate MM module effects term, though the improper CAR formulation is replaced with a very similar proper variant that allows for joint or block sampling. See Savitsky and Paddock (2013) for additional details of the base distribution formulations.

The input, `Omega`, is a list object of length equal to the number of groups of modules for which `"car"` is selected in `typetreat`. This construction is identical to that for `dpgrowmult`, where each list component is a square adjacency matrix dimensioned by the number of modules in the group. So we see that implementing Equation 18 estimation function `ddpgrow` is a non-parametric generalization of `dpgrowmult`. Heuristically, `ddpgrow` may be described as bringing the set of MM module effect terms utilized for `dpgrowmult` "inside" the DP prior specified for client random effects to produce an expanded client-by-module random effects parameterization.

We now run `ddpgrow` with our BRIGHT dataset. Focusing on the code block, below, we first run the `ddpgrow` estimation function, allocating the modules to the 4 CBT groups through the `numdose` input vector. We employ the multivariate CAR base distributions for all of the CBT groups. As earlier, we next extract the least squares clustering of clients (from among the posterior clusterings) and create the matrix, `map`, for faceting our plots of the resulting client-by-module random effects by clusters of clients. These effect plots are accomplished in the last script sequence by invoking plot function, `ddpEffectsplot`. The `orderto` input is an optional vector input of length equal to the number of modules, $S = 61$ that enumerates a specific order for the plotting of sessions within group (with the grouping conveyed in `map.group`). The default is to plot the client-by-module effects in order of increasing value and is overridden by the `orderto` input. This sorting of effects by increasing value may be directly turned off by choosing `re.order = FALSE`, which is done automatically if an input is specified for `orderto`.

```
R> typetreat <- c("car", "car", "car", "car")
R> DDP <- ddpgrow(y = dat$y, subject = dat$subject, trt = dat$trt,
+    time = dat$time, n.random = 3, n.fix_degree = 2, numdose = dat$numdose,
+    dosemat = dat$dosemat, Omega = dat$Omega, n.iter = n.iter,
```

```
+    n.burn = n.burn, n.thin = n.thin, shape.dp = 2, rate.dp = 4,
+    M.init = 15, plot.out = TRUE, typetreat = typetreat, labt = dat$labt)

[1] "Your chosen set of treatment base distributions  = car car car car"
```

We next re-group clients into extracted clustering and focus on 6 most populated clusters.

```
R> cluster <- samples(DDP)$bigSmin
R> c.sizes <- sapply(cluster, length)
R> clusterstoplot <- sort(c.sizes, decreasing = TRUE,
+    index.return = TRUE)$ix[1:6]
R> map <- vector(mode="list", length = length(clusterstoplot))
R> for(i in 1:length(clusterstoplot)) {
+    cluster.i <- cluster[[clusterstoplot[i]]]
+    map[[i]] <- as.data.frame(cbind(cluster.i,
+      paste("cluster", i, sep = "_")), stringsAsFactors = FALSE)
+    names(map[[i]]) <- c("subject", "group")
+  }
R> map <- do.call("rbind", map)
```

We now plot subject-by-module multivariate MM effects averaged over clients within each of the 6 clusters.

```
R> trts.plot <- paste("cbt", 1:4, sep="_")
R> dep <- ddpEffectsplot(DDP, cred.intervals = FALSE,
+    x.axis.label = "Modules", map.group = map, re.order = FALSE,
+    trts.plot = trts.plot, orderto = 1:61)
```

Running `ddpEffectsplot` produces a set of plot objects, including the two we next discuss, which both focus on analyzing the client-by-module effects aggregated to clusters of clients for each of the 4 CBT groups. Figure 10 provides insight for examining the variation in client-by-module effects across clusters of clients and how those effects vary over time. The figure displays $q = 3$ posterior mean polynomial effects for each module averaged up to the 6 clusters of clients, denoted by cluster_1, ..., cluster_6. The 3 polynomial effect values are presented for all modules. Each cluster's trajectories are presented within each of the 4 open-enrollment CBT therapy groups (of modules) along the rows within clusters. They are denoted by cbt_1, ..., cbt_4 in the Figure. These polynomial parameters imply an effect trajectory for each module with the order 1 effect providing the intercept, the order 2 effect the slope and order 3, a non-linear quadratic term. For example, the resulting effects trajectory for a module would be U-shaped if the order 3 term is positive. Scanning from left-to-right across the clusters in Figure 10, we observe a notable variation in first order module effect on depressive symptoms among the clusters, particularly for cbt groups 1–3. If we compare the first order (red) module effects in the third row Figure 10 for group 3 with the same result in the lower left panel in Figure 7 generated from the additive two-term model of Equation 1, we see a similar S-shape, though the magnitude among cluster_1 clients is higher in Figure 10, but progressively diminishes in the other clusters of clients (from left-to-right).

Figure 11 presents module effect trajectories of the BDI-II depressive symptom scores for randomly selected modules, which are composed from performing within sample predictions

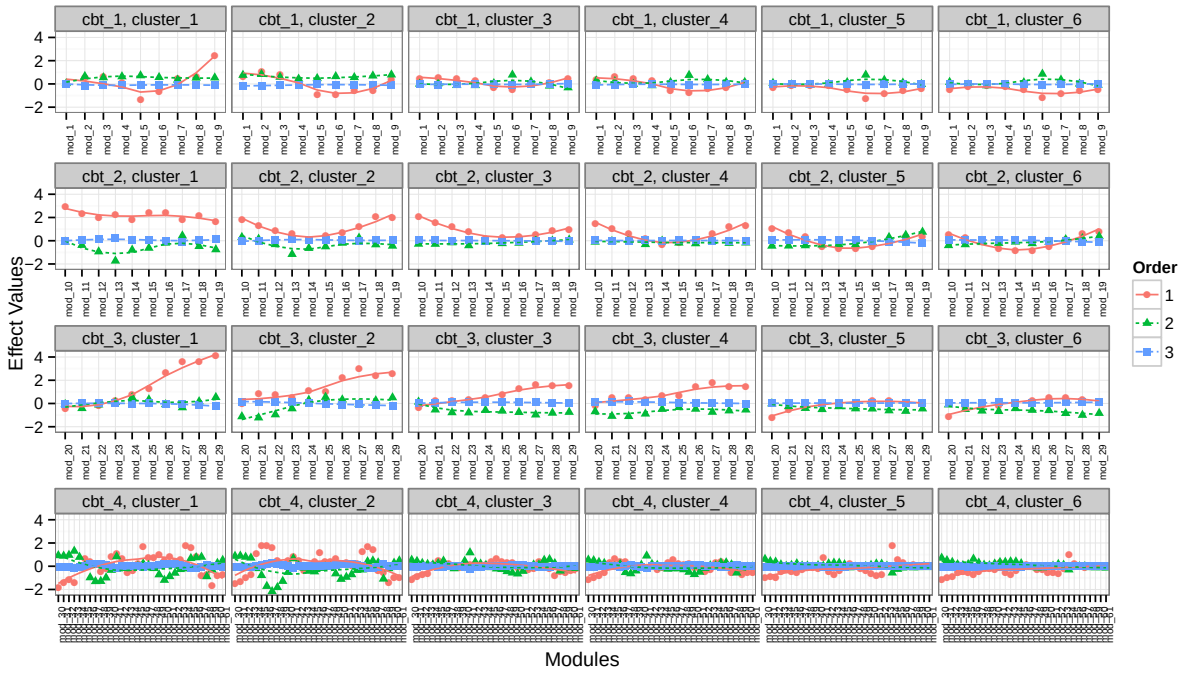

Figure 10: Client-by-module effects: posterior mean intercept (order 1), slope (order 2) and quadratic (order 3) effects. Columns are clusters (of clients) and rows are CBT treatment groups.

Model	$\bar{D}$	$-LPML$	DIC_3
MMCAR	5505	2980	5679
MMIMV	5547	2994	5716
DDP	5079	2929	5302

Table 4: BRIGHT study model fit comparisons: $\bar{D}$, $-LPML$ and DIC_3 scores for model alternatives. Lower values imply better performance.

using the polynomial client-by-module effects. The same display format is used from Figure 10, with clusters of clients across the columns and CBT groups down the rows. Scanning the columns from left-to-right reveals the same marked attenuation in cluster responsiveness to the CBT intervention that we observe in Figure 10. Member clients of clusters 4–6 express much less depressive symptom sensitivity to participation in the modules and, therefore, one notes much less differentiation in effect values among the modules for these clusters. These results reveal the extent to which group therapy clients express distinct underlying tendencies for realizing change in depressive symptoms. The ability for clients to express differing responses to module attendances produces a much better fit for `DDP` as compared to other models, as noted in Table 4 (where our MM effects are denoted by *module*, rather than *session* in our earlier examples using BRIGHT): Even though there are many more parameters in `DDP`, the borrowing of strength among clients reduces the effective parameterization sufficiently such that the dramatically improved deviance drives lower values for both penalized fit statistics.

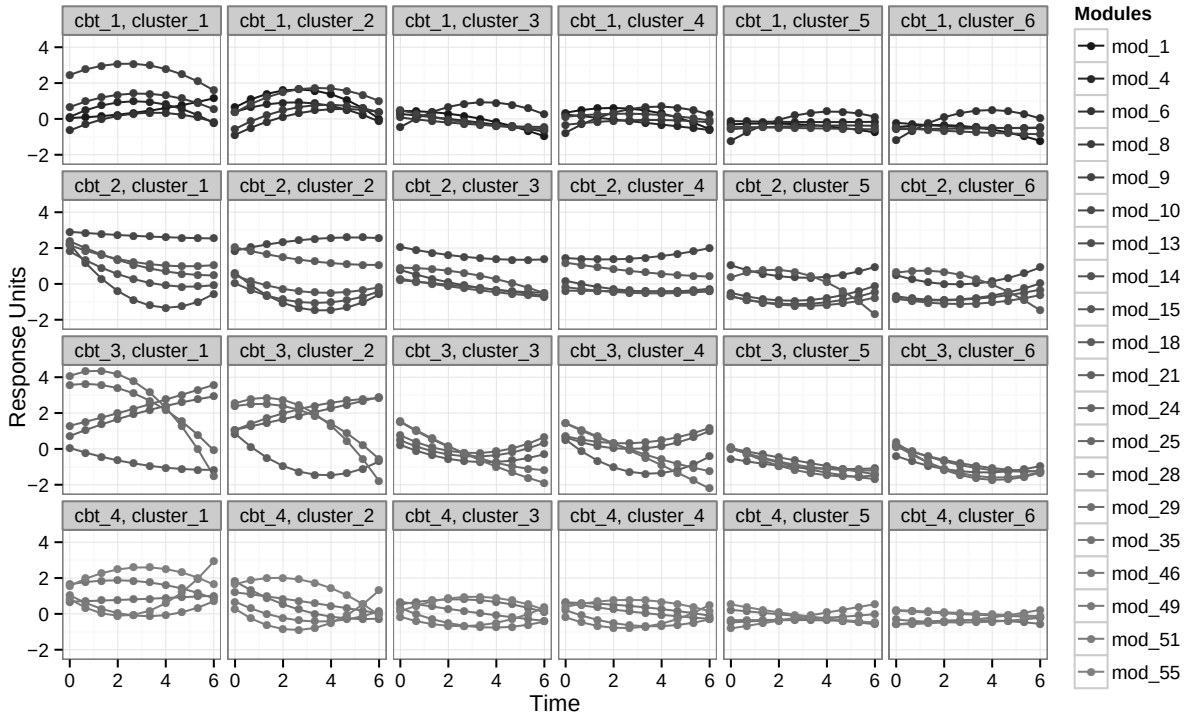

Figure 11: Client-by-module effect curves: within-sample predicted growth curves for client-by-module effects generated from using the 3 effect orders for each module. The columns contain clusters of clients and the rows are CBT treatment groups. Randomly-selected modules are included in each plot cell.

Model	CPU-time (min)
LGM	0.3
DP	2.6
MMCAR	4.6
MMDCAR	5.8
MMIGRP	7.8
DDP	720.0

Table 5: BRIGHT study model computation time comparisons in CPU minutes.

6. Computation

Each of our models illustrated with the BRIGHT study data was run for 40000 iterations on an $i7$ quad core machine configured with 8 threads where only 1 thread was used to compare computation run times in Table 5.

The DDP model computational speed is impacted by the matrix variate objects estimated for each subject and the time is of $\mathcal{O}(n^2)$, where n denotes the number of subjects. While this model experiences a far longer run time than the others, our experience is that it tends to provide the best fit and very useful inference for MM random effects.

7. Concluding remarks

Repeated measures data over a set of subjects drawn from an experimental trial are typically expensive and difficult to acquire. Study protocols on how treatment(s) are administered may induce complex patterns of measurement dependence among subjects. Standard methods for their analysis do not fully borrow strength in estimation of subject effects or account for the often complex set of correlations induced by the manner in which treatments are administered. **growcurves** for R addresses the opportunities with employment of a Bayesian semi-parametric hierarchical framework with employment of MM random effects.

The design of **growcurves** further recognizes the important graphical information inherent to a repeated measures data construction and provides plot functions that supply the user with the flexibility to compose and arrange graphics to suit their inferential tasks. Each of the three plot functions addresses a core analytical need common to inference conducted with repeated measures data for treatment(s) evaluation studies, including generation and aggregation of subject growth curves to extract patterns of persistence, displaying the distribution over treatment fixed mean differences between any two treatment arms, and plots to determine how one or more MM term effects contributes to estimation of the observed response. Then our **growcurves** package construction is well-targeted to a class of inferential problems and intends to save the user much of the effort typically required under application of a general software solution for such analysis.

Future work on the **growcurves** package for R will focus on adding generalized linear model constructions to accommodate dichotomous, polychotomous and event time outcomes. We hope to further focus on more fully modeling sparse matrix formulations in C++ to further reduce the computation times for our estimation functions.

Computational details

For obtaining the results in the manuscript, R version 3.0.2 (R Core Team 2013) was used with packages **growcurves** (Savitsky 2014) and **matlab** (Roebuck 2014). Additionally, **growcurvers** depends on or imports the following packages: **Rcpp** (Eddelbuettel and François 2011), **RcppArmadillo** (Eddelbuettel and Sanderson 2014), **reshape2** (Wickham 2007), **ggplot2** and **scales** (Wickham 2009, 2014a), **testthat** (Wickham 2014b), and **Formula** (Zeileis and Croissant 2010).

Acknowledgments

This study was partly funded by NIH/NIAAA R01AA019663.

References

Banerjee S, Wall MM, Carlin BP (2003). "Frailty Modeling for Spatially Correlated Survival Data, with Application to Infant Mortality in Minnesota." *Biostatistics*, **4**(1), 123–142.

Beck AT, Steer RA, Brown GK (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Besag J, York J, Mollié A (1991). "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *The Annals of the Institute of Statistical Mathematics*, **43**(1), 1–20.

Blackwell D, MacQueen JB (1973). "Ferguson Distributions via Pólya Urn Schemes." *The Annals of Statistics*, **1**(2), 353–355.

Brown ER, Ibrahim JG (2003). "A Bayesian Semiparametric Joint Hierarchical Model for Longitudinal and Survival Data." *Biometrics*, **59**(2), 221–228.

Browne WJ, Goldstein H, Rasbash J (2001). "Multiple Membership Multiple Classification (MMMC) Models." *Statistical Modelling*, **1**(2), 103–124.

Celeux G, Forbes F, Robert CP, Titterington DM (2006). "Reply to Comments on "Deviance Information Criteria for Missing Data Models"." *Bayesian Analysis*, **1**(4), 701–706.

Congdon P (2005). *Bayesian Models for Categorical Data*. John Wiley & Sons.

Dahl DB, Day R, Tsai JW (2008). "Distance-Based Probability Distribution on Set Partitions with Applications to Protein Structure Prediction." *Technical report*, Texas A&M University.

Eddelbuettel D, François R (2011). "**Rcpp**: Seamless R and C++ Integration." *Journal of Statistical Software*, **40**(8), 1–18. URL http://www.jstatsoft.org/v40/i08/.

Eddelbuettel D, Sanderson C (2014). "**RcppArmadillo**: Accelerating R with High-Performance C++ Linear Algebra." *Computational Statistics and Data Analysis*, **71**, 1054–1063.

Escobar MD, West M (1995). "Bayesian Density Estimation and Inference Using Mixtures." *Journal of American Statistical Association*, **90**(430), 577–588.

Ferguson TS (1973). "A Bayesian Analysis of Some Nonparametric Problems." *The Annals of Statistics*, **1**(2), 209–230.

Hill PW, Goldstein H (1998). "Multilevel Modeling of Educational Data with Cross-classification and Missing Identification for Units." *Journal of Educational and Behavioral Statistics*, **23**(2), 117–128.

Hodges JS, Carlin BP, Fan Q (2003). "On the Precision of the Conditionally Autoregressive Prior in Spatial Models." *Biometrics*, **59**(2), 317–322.

Jara A, Hanson TE, Quintana FA, Mueller P, Rosner GL (2011). "**DPpackage**: Bayesian Semi- and Nonparametric Modeling in R." *Journal of Statistical Software*, **40**(5), 1–30. URL http://www.jstatsoft.org/v40/i05/.

Lane PW, Nelder JA (1982). "Analysis of Covariance and Standardization As Instances of Prediction." *Biometrics*, **38**(3), 613–621.

Mariano LT, McCaffrey DF, Lockwood JR (2010). "A Model for Teacher Effects from Longitudinal Data Without Assuming Vertical Scaling." *Journal of Educational and Behavioral Statistics*, **35**(3), 253–279.

Morgan-Lopez AA, Fals-Stewart W (2006). "Analytic Complexities Associated with Group Therapy in Substance Abuse Treatment Research: Problems, Recommendations, and Future Directions." *Experimental and Clinical Psychopharmacology*, **14**(2), 265–273.

Müller P, Quintana FA, Rosner GL (2007). "Semiparametric Bayesian Inference for Multilevel Repeated Measurement Data." *Biometrics*, **63**(1), 280–289.

Paddock SM, Hunter SB, Watkins KE, McCaffrey DF (2011). "Analysis of Rolling Group Therapy Data Using Conditionally Autoregressive Priors." *The Annals of Applied Statistics*, **5**(2A), 605–627.

Paddock SM, Savitsky TD (2013). "Bayesian Hierarchical Semiparametric Modeling of Longitudinal Post-Treatment Outcomes from Open Enrollment Therapy Groups." *Journal of the Royal Statistical Society A*, **176**(3), 795–808.

Rasbash J, Charlton C, Browne WJ, Healy M, Cameron B (2012). ***MLwiN*** *Version 2.02*.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Roebuck P (2014). ***matlab****: MATLAB Emulation Package*. R package version 1.0.1, URL http://CRAN.R-project.org/package=matlab.

Savitsky T (2014). ***growcurves****: Bayesian Semi- and Nonparametric Growth Curve Models that Additionally Include Multiple Membership Random Effects*. R package version 0.2.3.8, URL http://CRAN.R-project.org/package=growcurves.

Savitsky TD, Paddock SM (2013). "Bayesian Nonparametric Hierarchical Modeling for Multiple Membership Data in Grouped Attendance Interventions." *The Annals of Applied Statistics*, **7**(2), 1074–1094.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002). "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society B*, **64**(4), 583–616.

Watkins KE, Hunter SB, Hepner KA, Paddock SM, de la Cruz E, Zhou AJ, Gilmore J (2011). "An Effectiveness Trial of Group Cognitive Behavioral Therapy for Patients with Persistent Depressive Symptoms in Substance Abuse Treatment." *Archives of General Psychiatry*, **68**(6), 1–8.

Wickham H (2007). "Reshaping Data with the **reshape** Package." *Journal of Statistical Software*, **21**(12), 1–20. URL http://www.jstatsoft.org/v21/i12/.

Wickham H (2009). ***ggplot2****: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Wickham H (2010). "A Layered Grammar of Graphics." *Journal of Computational and Graphical Statistics*, **19**(1), 3–28.

Wickham H (2014a). ***scales****: Scale Functions for Graphics*. R package version 0.2.3, URL http://CRAN.R-project.org/package=scales.

Wickham H (2014b). ***testthat****: Testthat Code. Tools to Make Testing Fun :)*. R package version 0.7.1, URL http://CRAN.R-project.org/package=testthat.

Zeileis A, Croissant Y (2010). "Extended Model Formulas in R: Multiple Parts and Multiple Responses." *Journal of Statistical Software*, **34**(1), 1–13. URL http://www.jstatsoft.org/v34/i01/.

Affiliation:

Terrance Savitsky
Office of Survey Methods Research
U.S. Bureau of Labor Statistics
2 Massachusetts Ave. N.E.
Washington, D.C. 20212, United States of America
E-mail: tds151@gmail.com
URL: http://www.rand.org/about/people/s/savitsky_terrance_dean.html

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 57, Issue 3 *Submitted:* 2012-04-16
March 2014 *Accepted:* 2013-09-23