arXiv:2301.01373v1 [stat.ME] 3 Jan 2023

用于多元时间序列分析的协变量引导型样条专家贝叶斯混合模型

Haoyi Fu
生物统计学系，匹兹堡大学，美国宾夕法尼亚州匹兹堡，haf48@pitt.edu

Lu Tang
生物统计学系，匹兹堡大学，美国宾夕法尼亚州匹兹堡

Ori Rosen数学科学系，德克萨斯大学埃尔帕索分校，美国德克萨斯州埃尔帕索

Alison E. Hipwell
匹兹堡大学精神病学系，美国宾夕法尼亚州匹兹堡

TheodoreJ. Huppert匹兹堡大学电气与计算机工程系，美国宾夕法尼亚州匹兹堡

RobertT. Krafty埃默里大学生物统计学与生物信息学系，美国佐治亚州亚特兰大

摘要

随着测量脑活动和脑结构的技术迅速发展，用于分析现代脑成像数据的统计方法在推动科学进步方面发挥着重要作用。用于测量脑功能的成像数据通常是多元时间序列，并且在成像光源和受试者之间均存在异质性，由此带来了各种统计和计算挑战。本文提出一种基于组的方法，通过平滑样条的贝叶斯混合模型对一组多元时间序列进行聚类。我们假设每个多元时间序列均由多个具有不同混合权重的成分构成。假设时间无关协变量与混合成分相关，并通过专家混合模型的逻辑斯蒂权重将其纳入模型。我们在完全贝叶斯框架下构建这一方法，并采用吉布斯抽样；成分数量则依据偏差信息准则进行选择。我们通过模拟研究将所提方法与现有方法进行比较，并将其应用于一项功能性近红外光谱（功能性近红外光谱）研究，该研究旨在了解婴儿情绪反应性及其从压力中的恢复情况。结果揭示了不同的脑活动模式，以及这些模式与选定协变量之间的关联。

关**键词** 贝叶斯混合模型 · 脑成像 · 功能性近红外光谱 · 基于模型的聚类 ·多元时间序列 · 平滑样条 · 面对面静止面孔

1 引言

时间序列是随机过程的实现。获得估计的时间序列轨迹有助于深入了解许多实际问题。功能性近红外光谱（功能性近红外光谱）是一种非侵入性脑成像技术，利用近红外光测量氧合血红蛋白和脱氧血红蛋白的变化（Jobsis, 1977年）。在功能性近红外光谱中，处理后的数据是非平稳多元时间序列，其均值不恒定且随时间变化幅度较大，这给推断和估计带来了诸多统计挑战。在功能性近红外光谱分析中，不同受试者可能呈现出不同的多元时间序列轨迹模式，这些模式可能与某些临床或人口统计学特征有关。

功能性近红外光谱数据的分析需要采用适当的方法，分析从不同受试者观测到的一组多元时间序列，这通常称为重复多元时间序列设定。
聚类分析常用于解决异质性问题，并从不同受试者的时间序列集合中识别亚组。时间序列聚类已应用于多个科学领域，以发现轨迹模式，从而从复杂且庞大的数据集中挖掘有价值的信息（Liao，2005年）。时间序列聚类将整个数据集合划分为不同的组，并依据某种相似性度量将同质时间序列归为一组。时间序列聚类面临的挑战包括高维数据带来的计算问题，以及如何选择恰当的相似性度量（Lin 等人，2003年；Keogh 和 Pazzani，2000年）。多位学者提出了用于多元时间序列的聚类算法。Kakizawa 等人（1998年）采用 Kullback–Leibler 判别信息作为聚类多元高斯时间序列的最小判别准则。Wang等人 （2007年）基于单变量结构，使用改进的 K-均值聚类算法对多元时间序列进行聚类。许多研究提出了不同的基于模型的多元时间序列聚类方法，例如多元自回归模型（Maharaj，1999年；He 等人，2022年）、隐马尔可夫模型（Li 等人，2001年）和平滑样条（Krafty 等人，2017年；Li 和 Krafty，2019年）。关于时间序列聚类方法的综合综述可参见 Liao（2005年）以及 Maharaj 等人（2019年）。

协变量依赖结构通常可以与时间序列聚类所得到的混合成分相关联。Bertolacci 等人（2022年）利用一种协变量依赖的无限混合模型分析了多个非平稳时间序列，该模型采用逻辑斯蒂棍折权重，混合权重根据协变量计算。专家混合模型（Jacobs等人，1991年）通过协变量依赖的多项逻辑斯蒂模型为每个专家分配权重。Huerta 等人（2003年）利用分层专家混合模型（Jordan 和 Jacobs，1994）研究了基于协变量的时间序列模型混合问题。

平滑样条是一种采用基于粗糙度的惩罚项的非参数方法，已广泛应用于时间序列分析（Wang，2011年；Gu，2013年）。Kimeldorf 和 Wahba（1970）首次探讨了平滑样条的贝叶斯解释。Wahba（1978年）证明，平滑样条目标函数的解等价于采用部分弥散先验的贝叶斯估计。Speckman 和 Sun（2003年）采用完全贝叶斯方法实现平滑样条，对方差分量使用非信息先验，并推导出了后验分布适当性的充要条件。平滑样条需要估计大量系数，这在高维情形下可能并不实用。Gu 和 Kim（2002年）使用一组再生核函数实现低维近似。Wood 等人（2002年）通过对高斯核进行特征分解获得一组基函数。Krafty 等人（2017年）提出了一种张量积模型，用于分析重复观测的多元时间序列；该模型将功率谱分解为单变量结果和频率的乘积。

本文旨在对多元时间序列进行协变量引导的聚类，以捕捉混合成分的轨迹模式，并评估协变量与轨迹模式之间的关系。为此，我们采用平滑样条对每个混合成分进行建模，并通过混合权重将时间无关协变量纳入混合模型。该方法在完整的贝叶斯框架下进行表述。本文其余部分安排如下：第 2 节介绍动机研究。第 3 节和第 4 节分别介绍所提模型和先验分布。第 5 节介绍抽样方案。第 6 节报告不同设置下的模拟结果，第 7 节则通过将所提方法应用于动机研究加以说明。第 8 节总结全文并进行讨论。

2 动机研究

我们的动机研究旨在了解婴儿在一种名为面对面静止面孔（FFSF）的情绪压力探测任务之前、期间和之后的脑活动模式（Tronick等，1978年）。本研究的参与母亲招募自匹兹堡女孩纵向研究（PGS）。该研究以人群为基础，在匹兹堡市招募了2,450名年龄在5至8岁的女孩（Keenan 等，2010年）。2016年，PGS启动了一项大型子研究，以探究环境因素（例如童年期和青春期经历的心理压力源）如何影响之后的孕产妇妊娠和儿童健康。本研究是美国国立卫生研究院儿童健康结局环境影响（ECHO）计划的一部分。该计划考察产前环境暴露在生物、化学、物理和社会领域对后代健康与发育的不同影响（Gillman和Blaisdell，2018年）。PGS-ECHO研究在PGS参与者怀孕或刚分娩活产儿时招募她们。参与者需完成多次产前实验室访视，并对儿童进行随访，年龄范围为6至36个月。实验室方案包括访谈和互动任务，用于评估情境压力源、健康、情绪、生活方式行为，以及子代的行为和情绪发育。

母亲与婴儿之间的面对面互动对婴儿的发育至关重要，尤其关系到沟通和社交技能，以及情绪和气质的调节（Hipwell 等，2019年）。静止脸范式是一种广泛使用的压力任务（即违背对社交互动的预期），可用于生物行为测量婴儿反应和恢复方面的个体差异。静止脸范式包含三个阶段：互动（或基线）、静止脸和恢复（Adamson和Frick，2003年）。在第一阶段，母亲与婴儿进行日常互动，不使用玩具；这一阶段作为基线。在第二阶段，母亲对婴儿采取中性面部表情（静止脸，即不进行面部或口头交流）；随后进入第三阶段，母亲恢复与婴儿的正常互动。在静止脸范式开始前，会将功能性近红外光谱帽戴在婴儿头上，以测量三个阶段中的脑激活水平及其变化。

PGS-ECHO 功能性近红外光谱静止脸数据使用连续近红外光谱成像系统（NIRScout；NIRx Medical Technologies, 柏林，德国）以 7.8125 Hz 的采样率记录，并使用 NIRStart 采集软件。数据在两个波长（760 nm 和 850 nm）下同时测量。如图 1(a) 所示，该功能性近红外光谱探头由 8 个光源和 4 个探测器组成，共有 12 个通道。

在本研究中，我们使用上述功能性近红外光谱探头测量婴儿脑活动（每个阶段的测量时间约为 120 秒）。截至 2021 年底，共收集了 155 名婴儿受试者的功能性近红外光谱静止脸数据。我们还收集了婴儿和母亲的人口统计学变量，例如胎龄、婴儿年龄、性别、出生体重、头围，以及家长对修订版婴儿行为问卷（IBQ-R）（Gartstein 和 Rothbart，2003年）的报告。剔除未完成静止脸范式三个阶段、基于杠杆值存在较大离群值，或在三个静止脸阶段中任一阶段的测量时间过短的婴儿后，共有 82 名受试者具备完整的功能性近红外光谱静止脸数据，可供后续分析。上述质量控制步骤由 MATLAB 中的 NIRS Brain AnalyzIR 工具箱完成（Santosa 等人，2018年）。此外，我们还在 R软件中进行了额外的数据预处理，包括数据插值和重新缩放。最终，处理后的功能性近红外光谱数据中，每名受试者的每个通道均有 1,500 个测量点，其中每个阶段包含 500 个点。所有测量值和采样时间均重新缩放至 0 到 1 之间：互动阶段为 0 至 1/3，静止脸阶段为 1/3 至 2/3，恢复阶段为 2/3 至 1。图 2 展示了两个选定受试者和四个选定通道的处理后功能性近红外光谱时间序列示例。

我们的分析旨在识别由氧合血红蛋白相对浓度表示的多个功能性近红外光谱通道中的不同脑活动轨迹模式，并评估轨迹模式与相关协变量之间的关联。

3 模型

本节将详细介绍我们所提出的协变量引导的样条专家贝叶斯混合模型。该模型由样条分量构成，其混合权重取决于协变量。

3.1 样条混合模型

我们针对多元时间序列提出了一种张量积样条混合模型。对于每位受试者 $i = 1, \ldots, N$，令$\boldsymbol{y}_i = (\boldsymbol{y}'_{i1}, \ldots, \boldsymbol{y}'_{ik}, \ldots, \boldsymbol{y}'_{iK})'$为一个 nK维向量，对应于 $k = 1, \ldots, K$的 K维时间序列，其中 $\boldsymbol{y}_{ik} = \left[y_{ik}(t_1) \ldots, y_{ik}(t_j) \ldots, y_{ik}(t_n)\right]'$包含时间序列第 k个条目在 $j = 1, \ldots, n$的一组 n 时间点上测得的测量轨迹，而 $\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}'_{i1}, \ldots, \boldsymbol{\epsilon}'_{iK})'$是一个 nK维误差向量。遵循Krafty等人（2017年）给出的模型表示，对于分量 g，在给定 $g = 1, \ldots, G$的条件下， K维多元时间序列的张量积模型可写为：

$$\{\boldsymbol{y}_i \mid z_{ig} = 1\} = (\boldsymbol{I}_K \otimes \boldsymbol{X})\boldsymbol{\alpha}_g + (\boldsymbol{I}_K \otimes \boldsymbol{W})\boldsymbol{\beta}_g + \boldsymbol{\epsilon}_i, \tag{1}$$

其中， $\{z_{ig}\}_{g=1}^{G}$为第 3.3 节所述的潜在指标， $\boldsymbol{\alpha}_g = (\boldsymbol{\alpha}'_{g1}, \ldots, \boldsymbol{\alpha}'_{gK})'$为由截距和斜率构成的 $2K$ 维向量，$\boldsymbol{\beta}_g = (\boldsymbol{\beta}'_{g1}, \ldots, \boldsymbol{\beta}'_{gK})'$为第 4.1 节所述的基函数系数 mK 维向量， $\boldsymbol{I}_K$为$K \times K$单位矩阵， $\otimes$表示张量积。矩阵 $\boldsymbol{X}$定义为 $\boldsymbol{X} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_1 & t_2 & \ldots & t_n \end{pmatrix}$，矩阵 $\boldsymbol{W}$的m 列为第 4.1 节所述的平滑样条基函数。我们假设误差向量 $\boldsymbol{\epsilon}_i$服从 MVN$(0, \boldsymbol{\Psi}_g \otimes \boldsymbol{U})$ 分布，其中 $\boldsymbol{U} = \boldsymbol{I}_n$为 $n \times n$ 单位矩阵， $\boldsymbol{\Psi}_g = \text{diag}(\boldsymbol{\sigma}_g^2)$ 是一个 $K \times K$对角矩阵，其误差方差为 $\boldsymbol{\sigma}_g^2 = (\sigma_{g1}^2, \ldots, \sigma_{gK}^2)'$。我们假设每个受试者在所有 K 条目上均具有共同的时间点网格，因此 $\boldsymbol{X}$ 和 $\boldsymbol{W}$ 对所有受试者均相同，尽管所提方法可以推广到受试者在不同时间点网格上接受观测的情形。此外，我们假设对于 $k \neq h$， $\boldsymbol{E}(\boldsymbol{y}_{ik}, \boldsymbol{y}_{ih}) = \boldsymbol{0}_{n \times n}$。

为简化记号，我们令 $\boldsymbol{S} = [\boldsymbol{X}\ \boldsymbol{W}]$ 和 $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}'_{g1},\ \boldsymbol{\beta}'_{g1}、\ \ldots\ldots、\ \boldsymbol{\alpha}'_{gK},\ \boldsymbol{\beta}'_{gK})'$。因此，公式（1）可改写为：

$$\{\boldsymbol{y}_i \mid z_{ig} = 1\} = (\boldsymbol{I}_K \otimes \boldsymbol{S})\boldsymbol{\theta}_g + \boldsymbol{\epsilon}_i. \tag{2}$$

3.2 混合权重模型

专家混合模型（Jacobs等，1991年）被用于为所提方法构建协变量引导结构，其中混合权重是协变量的函数，以多项逻辑函数表示。与 Sun等（2007年）一样，混合权重表示为

$$\pi_{ig}(\boldsymbol{V}_i) = \frac{\exp(\boldsymbol{V}'_i\boldsymbol{\delta}_g + \zeta_{ig})}{\sum_{h=1}^{G}\exp(\boldsymbol{V}'_i\boldsymbol{\delta}_h + \zeta_{ih})}, \tag{3}$$

其中，$\boldsymbol{V}_i = (1,\ V_{i1},\ \cdots,\ V_{iP})'$ 是一个长度为（$P+1$）的向量，包含受试者 i 的 P协变量取值；$\boldsymbol{\delta}_g = (\delta_{g0},\ \delta_{g1},\ \cdots,\ \delta_{gP})'$ 是相应的系数向量。为保证可辨识性，我们令 $\boldsymbol{\delta}_G = \boldsymbol{0}$。公式（3）与传统专家混合模型中的权重略有不同，因为它为每位受试者纳入了一个随机项 ζ_{ig}。该项考虑了观测协变量之外的未测量因素，并提高了模型性能以及对混合权重的推断能力。

3.3 增广似然

为考虑受试者之间的异质性，我们假设多元时间序列 $\boldsymbol{y}_{ik}$ 的第 k 个元素来自一个包含 G 个分量的混合模型，即

$$\boldsymbol{y}_{ik} \sim \sum_{g=1}^{G}\pi_{ig}f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\mu}_{gk}, \sigma^2_{gk}\boldsymbol{I}_n), \tag{4}$$

其中，$f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\mu}_{gk}, \sigma^2_{gk}\boldsymbol{I}_n)$ 是多元正态分布的概率密度函数，其均值向量为 $\boldsymbol{\mu}_{gk} = \boldsymbol{X}\boldsymbol{\alpha}_{gk} + \boldsymbol{W}\boldsymbol{\beta}_{gk}$，协方差矩阵为 $\sigma^2_{gk}\boldsymbol{I}_n$，对应第 g 个分量和第 k 个元素。π_{ig} 是依赖于协变量的混合权重，如第 3.2 节所述。

与混合模型中的常见做法相同，通过引入表示时间序列所属分量的潜变量来增广似然，可以大大简化计算（Dempster等人，1977年）。具体而言，若第 i 个多元时间序列属于第 g 个分量，则令 $z_{ig} = 1$；否则令 $z_{ig} = 0$。令 $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)'$ 表示所有观测到的多元时间序列，并令 $\boldsymbol{\Theta}_{gk}$ 表示分量 g 和元素 k 的所有参数的集合。于是，所有分量和所有元素对应的参数向量记为 $\boldsymbol{\Theta} = (\boldsymbol{\Theta}'_{11}, \ldots, \boldsymbol{\Theta}'_{GK})'$。所有 N 个多元时间序列的增广似然为

$$L(\boldsymbol{\Theta} \mid \boldsymbol{y}, Z) = \prod_{i=1}^{N}\prod_{g=1}^{G}\left[\pi_{ig}\prod_{k=1}^{K}f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{gk})\right]^{z_{ig}}, \tag{5}$$

其中，$f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{gk})$ 是式 (4) 中出现的概率密度函数。根据贝叶斯定理，潜在指标 z_{ig} 的分布为

$$p(z_{ig} = 1 \mid \boldsymbol{y}, \boldsymbol{S}, \boldsymbol{\Theta}, \pi_{ig}) = \frac{\pi_{ig}\prod_{k=1}^{K}f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{gk})}{\sum_{h=1}^{G}\pi_{ih}\prod_{k=1}^{K}f_{hk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{hk})}. \tag{6}$$

4 先验分布

本节介绍模型参数的先验分布。

4.1 平滑样条先验

模型（4）中混合成分的条件期望为 $E(\boldsymbol{y}_{ik} \mid z_{ig} = 1) = \boldsymbol{X}\boldsymbol{\alpha}_{gk} + \boldsymbol{W}\boldsymbol{\beta}_{gk}$。我们对 $\boldsymbol{\beta}_{gk}$ 赋予平滑样条先验，并令 $\boldsymbol{\mathcal{H}}_{gk} = \boldsymbol{W}\boldsymbol{\beta}_{gk}$，其中 $\boldsymbol{\mathcal{H}}_{gk} = \left[\mathcal{H}_{gk}(t_1)\ldots, \mathcal{H}_{gk}(t_n)\right]'$ 是均值为零的高斯过程，其方差协方差矩阵为 $\tau^2_{gk}\boldsymbol{\Phi}$（Wahba，1980年；Wood 等人，2002年），使得 $\operatorname{cov}\left[\mathcal{H}_{gk}(t_r)\ \mathcal{H}_{gk}(t_h)\right] = \tau^2_{gk}\phi_{rh}$，其中 τ^2_{gk} 是分量 g 和条目 k 的平滑参数，而（r, h）阶元素

$\boldsymbol{\Phi}$ 的值为 $\phi_{rh} = \frac{1}{2}t_r^2(t_h = \frac{t_r}{3})$，其中 $t_r \le t_h$。矩阵 $\boldsymbol{\Phi}$ 对所有受试者均相同，因为多元时间序列的所有条目都在共同的时间点被观测到。
如上所示，矩阵 $\boldsymbol{\Phi}$ 是 $n \times n$。为避免 n较大时带来的计算负担，通常采用低秩近似。为便于进行这种近似，我们通过对 $\boldsymbol{\Phi}$进行谱分解来获得基函数，这一方法由 Wood等人 (2002年)提出，并由 Rosen 等人 (2009年、2012年)以及 Krafty等人 (2011年)采用。具体而言，矩阵 $\boldsymbol{W}$由在时刻 $t_1, \ldots, t_n$处求得的 m 个基函数组成，而 $\boldsymbol{\beta}_{gk}$是一个 m维的基函数系数向量。通过对 $\boldsymbol{\Phi}$ 进行谱分解，可得到这些基函数，使得 $\boldsymbol{\Phi} = \boldsymbol{Q}\boldsymbol{\Gamma}\boldsymbol{Q}^T$，其中 $\boldsymbol{Q}$ 是 $\boldsymbol{\Phi}$的特征向量矩阵， $\boldsymbol{\Gamma}$ 是包含 $\boldsymbol{\Phi}$的特征值的对角矩阵。接着，我们令设计矩阵为 $\boldsymbol{W} = \boldsymbol{Q}\boldsymbol{\Gamma}^{1/2}$，并对 $\boldsymbol{\beta}_{gk}$设定正态先验 $N(\boldsymbol{0}, \tau_{gk}^2\boldsymbol{I}_n)$，从而得到如上所述的 $\boldsymbol{\mathcal{H}}^{gk}$或 $\boldsymbol{W}\boldsymbol{\beta}_{gk} \sim N(\boldsymbol{0}, \tau_{gk}^2\boldsymbol{\Phi})$。

利用低秩近似， $\boldsymbol{W}$的列数从 n 减少到 m （$m < n$），在不牺牲模型拟合效果的情况下，大幅减轻了计算负担（Wahba，1980年；Wood，2006年）。Eubank（1999年）指出，对角矩阵 $\boldsymbol{\Gamma}$ 中的特征值会随着 m 的增大而迅速衰减。因此，选取相对较少的 m 个基函数即可获得良好的近似。如第6节所述，在模拟研究中将基函数的数量 m 设为10；已有研究表明（Krafty等人2011年），这可以解释总变异的98%以上。

因此， $\boldsymbol{\theta}_g$的先验为 $\boldsymbol{\theta}_g \sim N$（$\boldsymbol{0}$， $\boldsymbol{D}_g$），其中 $\boldsymbol{D}_g$ =diag($\sigma_{\alpha 1}^2\boldsymbol{1}_2$， $\tau_{g1}^2\boldsymbol{1}_m$，……， $\sigma_{\alpha K}^2\boldsymbol{1}_2$， $\tau_{gK}^2\boldsymbol{1}_m$)是 $\boldsymbol{\theta}_g$的协方差矩阵。向量（$\sigma_{\alpha 1}^2$，……， $\sigma_{\alpha K}^2$）′包含回归系数 $\boldsymbol{\alpha}_{gk}$的固定先验方差，这些方差对所有成分和条目均相同。具体而言，我们固定共同先验方差 $\sigma_\alpha^2 = 100$。向量 $\boldsymbol{\tau}_g^2$ =（τ_{g1}^2，……， τ_{gK}^2）′包含第 g个混合成分的平滑参数，而 $\boldsymbol{1}_m$是一个由 m个1组成的向量。我们假设回归系数 $\boldsymbol{\alpha}_{gk}$ 与基函数系数 $\boldsymbol{\beta}_{gk}$相互独立。

4.2 平滑参数的先验分布

我们假设平滑参数 $\boldsymbol{\tau}_g^2$ =(τ_{g1}^2，……， τ_{gK}^2)′在不同分量 g和条目 k间有所不同。尽管方差参数最常用的先验分布是逆伽马分布，但 Gelman（2006年）和 Wand 等人（2011年）指出，对标准差采用半-t 先验分布能够反映尺度参数信息不足的情况。半-t 属于重尾分布，且具有良好的收缩性能。它可以表示为逆伽马随机变量的尺度混合，其中潜变量服从逆伽马分布（Wand等人，2011年）。因此，我们假设半-t 分布为 $\tau_{gk} \sim t_{\nu_\tau}^+$(0， A_τ)，其中 ν_τ为自由度参数， A_τ为尺度参数。对于所有分量和条目，我们设定 $\nu_\tau = 3$和 $A_\tau = 10$。

4.3 误差方差的先验分布

我们假设 σ_{gk} 独立同分布$\sim t_{\nu_\sigma}^+$(0， A_σ)，并对所有分量和条目设定 $\nu_\sigma = 3$和 $A_\sigma = 10$。

4.4 逻辑斯蒂参数和随机截距方差的先验分布

本节详细介绍逻辑斯蒂权重参数 (3) 的先验分布。为简化记号，我们记 $\boldsymbol{\delta}_g^* = (\boldsymbol{\delta}_g^T, \boldsymbol{\zeta}_g^T)^T$，其中 $\boldsymbol{\zeta}_g = (\zeta_{1g}, \cdots, \zeta_{Ng})^T, g = 1, \ldots, G$。令 $\boldsymbol{V}_i^* = (\boldsymbol{V}_i', \boldsymbol{e}_i')'$，其中 $\boldsymbol{e}_i$是一个除第 i 个位置为 1 外其余元素均为 0 的向量，而 $\boldsymbol{V}^*$是由行 $\boldsymbol{V}_i^{*T}, i = 1, \ldots, N$组成的矩阵。逻辑斯蒂参数采用高斯先验，即 $\boldsymbol{\delta}_g^* \sim N(\boldsymbol{0}, \boldsymbol{B}_g)$，其中 $\boldsymbol{B}_g$ =diag($\sigma_{\delta g}^2\boldsymbol{1}_{\mathrm{P}+1}$, $\kappa_{\zeta g}^2\boldsymbol{1}_{\mathrm{N}}$)；随机截距的先验分布满足 $\boldsymbol{\zeta}_g \sim N(\boldsymbol{0}, \kappa_{\zeta g}^2\boldsymbol{I}_N)$。对于超参数，我们假设所有分量和协变量均为 $\sigma_{\delta g}^2 = 10$，且 $\kappa_{\zeta g} \sim t_{\nu_\kappa}^+(0, A_\kappa)$，其中所有分量均满足 $\nu_\kappa = 3$和 $A_\kappa = 10$。

为抽取逻辑斯蒂参数，Polson等人（2013年）提出了一种引入 Pólya-Gamma潜变量的数据增广方案，从而便于进行 Gibbs 抽样步骤。关于逻辑斯蒂参数抽样的详细信息见补充材料。

5 抽样方案

本节概述了从所有模型参数的条件后验分布中抽样所需的吉布斯步骤。更多详情请参阅补充材料。

5.1 吉布斯抽样步骤

令 ℓ 表示当前的吉布斯抽样迭代，根据以下步骤抽取第 $(\ell+1)$ 次迭代的参数值。

1. 从 $(\boldsymbol{\theta}_{gk}^{(\ell+1)} \mid \boldsymbol{y}, \boldsymbol{S}, \tau^{2(\ell)}{}_{gk}, \sigma_{gk}^{2(\ell)}) \sim N(\boldsymbol{u}_{gk}, \sigma_{gk}^2 \boldsymbol{\Lambda}_{gk})$ 中抽取 $\boldsymbol{\theta}_{gk}^{(\ell+1)}$，其中 $\boldsymbol{u}_{gk}$ 和 $\boldsymbol{\Lambda}_{gk}$ 分别为均值向量和协方差矩阵。2. 从 $(\sigma_{gk}^{2(\ell+1)} \mid \boldsymbol{\epsilon}_{igk}^{(\ell+1)}, a_{\sigma_{gk}}^{(\ell+1)}) \sim IG\Big((nN(\ell)_g + \nu_\sigma)/2, \sum_{i=1}^{N} z_{ig} \boldsymbol{\epsilon}_{igk}' \boldsymbol{\epsilon}_{igk}/2 + \nu_\sigma/a_{\sigma_{gk}}\Big)$ 中抽取 $\sigma_{gk}^{2(\ell+1)}$，其中 $N_g^{(\ell)}$ 是当前第 g 个分量中的受试者数量，$\boldsymbol{\epsilon}_{igk}$ 是第 g 个分量、第 i 个受试者和第 k 个条目对应的误差向量，而 $a_{\sigma_{gk}}$ 是半 -t 分布所依据的 IG 尺度混合中的潜变量。3. 从 $(\tau^{2(\ell+1)}{}_{gk} \mid \boldsymbol{\beta}_{gk}^{(\ell+1)}, a_{\tau_{gk}}^{(\ell+1)}) \sim IG\Big((\nu_\tau + m)/2, \boldsymbol{\beta}_{gk}' \boldsymbol{\beta}_{gk}/2 + \nu_\tau/a_{\tau_{gk}}\Big)$ 中抽取 $\tau^{2(\ell+1)}{}_{gk}$，其中 $a_{\tau_{gk}}$ 是与第 2 步中相同类型的潜变量。4. 从 $(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \omega_{ig}^{(\ell+1)}, \kappa_{\zeta g}^{2(\ell)}) \sim N(\boldsymbol{M}_g, \boldsymbol{\Sigma}_g)$ 中抽取 $\boldsymbol{\delta}_g^{*(\ell+1)}$，其中 $\omega_{ig}^{(\ell+1)}$ 是第 4.4 节所述增广中的 Pólya-Gamma潜变量。5. 从 $(\kappa_{\zeta g}^{2(\ell+1)} \mid \boldsymbol{\zeta}_g^{(\ell+1)}, a_{\kappa_g}^{(\ell+1)}) \sim IG\Big(\nu_\kappa/2, \boldsymbol{\zeta}_g' \boldsymbol{\zeta}_g/2 + (\nu_\kappa + N)/a_{\kappa_g}\Big)$ 中抽取 $\kappa_{\zeta g}^{2(\ell+1)}$，其中 a_{κ_g} 是与第 2 步和第 3 步中相同类型的潜变量。6. 通过根据公式（3）计算 $p(\pi_{ig}^{(\ell+1)} \mid \boldsymbol{V}^*, \boldsymbol{\delta}_g^{*(\ell+1)}, z_{ig}^{(\ell)})$，得到混合权重 $\pi_{ig}^{(\ell+1)}$。7. 根据公式（6）抽取 $z_{ig}^{(\ell+1)} \sim p(z_{ig}^{(\ell+1)} = 1 \mid \boldsymbol{y}, \boldsymbol{S}, \boldsymbol{\theta}_{gk}^{(\ell+1)}, \sigma_{gk}^{2(\ell+1)}, \pi_{ig}^{(\ell+1)})$。

5.2 选择成分数量

Spiegelhalter 等人(2002年)建议使用偏差信息准则（DIC），根据参数的有效数量进行模型选择。
Gelman 等人(2003年)提出了另一种衡量参数有效数量的方法，该方法基于 MCMC 迭代中对数预测密度的方差。与原方法相比，这种衡量方法更稳健、也更准确。此外，它始终为正值，且对重新参数化保持不变（Gelman等人，2003年）。
I本文使用偏差信息准则（DIC）为所提混合模型选择成分数量。

6 模拟研究

为展示所提方法的性能，我们通过在两种情形下从所提模型生成数据集来开展模拟研究：三元时间序列的双成分混合（$G = 2$）（$K = 3$）和二元时间序列的四成分混合（$G = 4$）（$K = 2$）。在每种模拟设置下，我们模拟100次重复，每次包含长度为 $n = 50$的 $N = 150$ 条时间序列。共进行20，000次吉布斯抽样，预热期为4，000次。在所有模拟设置中，超参数均设为相同的值，详见第4节。

6.1 双成分三元模型

在此情形下，我们考虑双成分三变量模型。根据公式 (1)，所提混合模型的第 g 个分量为

$$\{\boldsymbol{y}(t_j) \mid z_{ig} = 1\} = \boldsymbol{\alpha}_{0g} + \boldsymbol{\alpha}_{1g} t_j + \sum_{q=1}^{m} w_q(t_j) \boldsymbol{\beta}_{gq} + \boldsymbol{\epsilon}_{gt_j}, \quad j = 1, \ldots, n, \;\; g = 1, \ldots, G, \tag{7}$$

其中，$\boldsymbol{y}(t_j)$表示在时间 t_j处取值的三元时间序列，$\boldsymbol{\alpha}_{01} = (1, -3, -2)'$、$\boldsymbol{\alpha}_{02} = (5, 4, 3)'$和 $\boldsymbol{\alpha}_{11} = (-2, 2, 0.5)'$、$\boldsymbol{\alpha}_{12} = (1, -1, -0.5)'$分别是各分量相互独立的截距和斜率。向量$\boldsymbol{\beta}_{gq}$包含分量 g中所有变量的第 q个样条系数，而 $w_q(t_j)$是在时间 t_j处取值的第 q个样条基函数。$\boldsymbol{\epsilon}_{gt_j}$是相互独立的零均值误差项，服从$\boldsymbol{\epsilon}_{gt_j} \sim$多元正态分布$\Big(\mathbf{0}, \text{diag}(\sigma_{g1}^2, \sigma_{g2}^2, \sigma_{g3}^2)\Big)$，其中 $\sigma_1^2 = (\sigma_{11}^2, \sigma_{12}^2, \sigma_{13}^2)' = (3, 5, 4.5)'$且 $\sigma_2^2 = (\sigma_{21}^2, \sigma_{22}^2, \sigma_{23}^2)' = (4, 3.5, 4)'$。平滑参数设为 $\tau_1^2 = (\tau_{11}^2, \tau_{12}^2, \tau_{13}^2)' = (3.5, 5, 8.5)'$和 $\tau_2^2 = (\tau_{21}^2, \tau_{22}^2, \tau_{23}^2)' = (6, 2.5, 1.5)'$。

我们考察轨迹和逻辑斯蒂参数（见公式（3））估计值的表现。对于前者，我们计算每个混合成分的平均均方根误差（ARSE） g

$$\text{ARSE}_g = \sqrt{\frac{1}{nK}\sum_{j=1}^{n}\sum_{k=1}^{K}\Big[\mu_{gk}(t_j) - \hat{\mu}_{gk}(t_j)\Big]^2},$$

其中， $\mu_{gk}(t_j)$是根据第 g个分量得到的 $y_k(t_j)$的期望值，而 $y_k(t_j)$是时间序列在时间 t_j处的第 k个条目。$\hat{}\mu_{gk}(t_j)$是 $\mu_{gk}(t_j)$在 $k = 1, \ldots, K$和 $j = 1, \ldots, n$处的后验均值估计值。
为处理混合成分之间可能发生的标签交换，我们使用第 g个分量的估计值以及各组的真实值 $g = 1, \ldots, G$，将所有分量中的最小值作为 ARSE$_g$。通过评估 ARSE 得到正确的分量标签后，我们还报告每个混合成分 g 的平均偏差（A-bias）和偏差方差（V-bias），其中

$$\text{A-bias}_g = \frac{1}{nK}\sum_{j=1}^{n}\sum_{k=1}^{K}\Big[\hat{\mu}_{gk}(t_j) - \mu_{gk}(t_j)\Big],$$

and V-bias$_g$ 通过计算各条目在各时间点上的偏差的样本方差得出 s.
对于每次重复，采用三种方法估计时间序列轨迹：所提方法、R软件包gbmt（Magrini，2022年）以及SAS中的TRAJ过程（Nagin等，2018年）。图3给出了各分量的ARSE、平均偏差和V偏差箱线图。值得注意的是，TRAJ能够通过将基函数视为时变协变量来拟合回归样条模型，而gbmt只能拟合三次模型。所提方法在贝叶斯框架下拟合惩罚样条模型，在两个分量的ARSE和V偏差方面均优于gbmt和TRAJ。三种方法的平均偏差都接近于零，且彼此相当。这些结果表明，由于整个时间序列上的偏差接近于零，三种方法都能合理拟合基于群组的轨迹。从较小的V偏差可以看出，所提方法能够获得更精确的轨迹估计。

为了评估逻辑斯蒂参数的表现，我们使用所提方法和 TRAJ 分别计算每个逻辑斯蒂参数的均方根误差（RMSE）。值得注意的是，gbmt 无法在计算混合权重时纳入协变量。表1给出了各逻辑斯蒂参数的均方根误差。我们还比较了在 $N = 150$,250 和$n = 50$,70 的四种不同组合设置下，所提方法与 TRAJ 的均方根误差。在所有情况下，我们的所提方法得出的逻辑斯蒂参数均方根误差都更小，尤其是截距 δ_0 和第一个协变量 δ_1。这符合预期，因为 TRAJ 使用多项逻辑斯蒂模型，在结果不平衡或出现完全分离时可能导致参数估计膨胀；而我们的所提方法能够通过惩罚方法获得收缩结果。

6.2 四分量双变量模型

在此情景中，我们考虑四分量双变量模型，其中第 g 个分量由公式（7）给出，截距和斜率的取值为 $\boldsymbol{\alpha}_{01} = (1, -2)'$, $\boldsymbol{\alpha}_{02} = (5, 3)'$, $\boldsymbol{\alpha}_{03} = (-3,5.5)'$, $\boldsymbol{\alpha}_{04} = (4, -1)'$,$\boldsymbol{\alpha}_{11} = (-3, 0)'$, $\boldsymbol{\alpha}_{12} = (2, -3.5)'$, $\boldsymbol{\alpha}_{13} = (2.5, 2)'$和 $\boldsymbol{\alpha}_{14} = (-3, 1.5)'$。类比于双成分三元模型，误差 $\boldsymbol{\epsilon}_{gt_j}$ 是相互独立、均值为零的双变量高斯随机变量，服从分布 $\boldsymbol{\epsilon}_{gt_j} \sim$ 多元正态分布$\Big(\mathbf{0}$,diag$(\sigma_{g1}^2,\ \sigma_{g2}^2)\Big)$，其中 $\sigma_1^2 = (\sigma_{11}^2,\ \sigma_{12}^2)' = (6,\ 9)'$， $\sigma_2^2 = (\sigma_{21}^2,\ \sigma_{22}^2)' = (8,\ 7。5)'$，$\sigma_3^2 = (\sigma_{31}^2,\ \sigma_{32}^2)' =$(10，6。5)′和 $\sigma_4^2 = (\sigma_{41}^2,\ \sigma_{42}^2)' = (7,\ 8。5)'$。
该情景下估计轨迹和逻辑斯蒂参数的表现见图4和表2。与第一个情景一样，在所有分量的 ARSE 和 V偏差方面，我们所提方法均优于 gbmt 和 TRAJ。值得注意的是，TRAJ 在若干重复样本中未能得到精确估计，因此平均 ARSE 和 V偏差更大。在逻辑斯蒂参数方面，所提方法表现良好，几乎所有情况下的均方根误差都较小，尤其是对于 δ_0 和 δ_1。基于上述两个情景中 N和 n 不同取值的更多模拟结果见补充材料。

7 真实数据应用

我们将所提方法应用于第 2 节介绍的 fNIRS静止面孔研究的分析。在协变量引导模型中，我们考虑了六个协变量，包括修订版婴儿行为问卷负性情绪评分（IBQ-NE）、修订版婴儿行为问卷努力控制评分（IBQ-EC）、孕龄（天）、

婴儿年龄（月）、头围（厘米）和性别。所有连续协变量均进行中心化和标准化。我们将基函数的数量设为 $m = 20$，并运行总计 30，000 次吉布斯迭代，其中预热期为 6，000 次。超参数的取值与模拟研究中使用的相同。

IBQ-NE 构念综合了以下子量表的数据：悲伤、对限制的痛苦反应、恐惧，以及反应性下降／痛苦恢复速度。IBQ-EC 指抑制占优势反应、转而执行次要反应的能力，研究表明，这种能力有助于抵御多种困难（Gartstein 等，2013年）。最后，数据包含 79 名具有完整功能性近红外光谱和协变量值的受试者。我们展示了基于一组四个通道的分析结果。基于另一组四个通道及全部通道进行分析的其他结果见补充材料。这四个通道为 S1D1、S2D2、S5D3 和 S6D4。通道 S1D1 和 S5D3 位于中央前额叶区域，而通道 S2D2 和 S6D4 分别位于左、右前额叶区域。我们拟合了所提模型，成分数从 2 到 6 不等。根据第 5.2 节中介绍的偏差信息准则（DIC）数值，双分量模型被选为这项四通道分析的最佳模型。

图5展示了拟合四个通道的双分量模型的估计轨迹。我们关注静止脸阶段的脑激活信号，并将互动阶段作为参照水平。对于分量1，四个通道在静止脸阶段均呈现下降轨迹。相比之下，分量2在四个通道的静止脸阶段均呈现上升趋势。拟合混合模型并发现上述轨迹模式后，我们根据静止脸阶段的轨迹模式，将分量1定义为无反应成分，将分量2定义为反应成分。图6展示了双分量模型中所有协变量的逻辑回归参数估计值，其中以分量2为参照。证据表明，两个分量的 IBQ-NE 得分存在差异，因为其 95% 可信区间不包含零。IBQ-NE 的系数为正，表明 IBQ-NE 评分越高，越可能归属于分量1；该分量在四个通道的静止脸阶段均表现为脑激活水平下降。尽管其他逻辑回归系数的 95% 可信区间均包含零，IBQ-EC 得分的后验均值估计为负，仍可能表明较高的 IBQ-EC 与脑激活增强相关，如分量2所示。这些结论与 Gartstein等人（2013年）的研究发现一致，即 IBQ-NE 与 IBQ-EC 呈负相关。Enlow等人（2016年）报告称，在鼓励婴儿进行高水平活动的家庭中，活动水平与 IBQ-NE 呈负相关。此外，婴儿年龄的逻辑斯蒂系数后验均值为负，表明年龄较小的婴儿在静止脸阶段更可能出现脑激活水平下降。

8 讨论

所提出的协变量引导的贝叶斯样条专家混合模型旨在对来自多个受试者的多元时间序列进行基于模型的聚类。该模型的混合成分为惩罚样条，混合权重则纳入协变量。我们通过模拟研究将所提方法与两种常用方法进行比较，结果表明，在不同情景下，所提方法表现更佳。我们将所提方法应用于一项功能性近红外光谱静止面孔研究，发现时间序列轨迹的成分呈现出不同模式，并且 IBQ-NE评分与静止脸阶段脑活动降低的模式相关。据我们所知，这是首项旨在识别轨迹成分的 fNIRS 静止脸研究。

所提方法存在一些局限性。首先，与任何混合模型一样，可能会发生标签交换，尤其是在实际数据应用中。我们采用了 Papastamoulis 和 Iliopoulos（2010年）提出的等价类代表（ECR）算法，使各成分具有可解释性，但也可以考虑其他方法。其次，所提方法假设时间序列各项之间相互独立，因此不允许存在空间依赖。功能性近红外光谱的空间相关，是指根据各光源和探测器的布置与位置而产生的功能性近红外光谱通道之间的相关性。通过考虑时间序列各项之间的空间相关，可以将该方法扩展为多层多元模型。最后，所提方法使用偏差信息准则选择成分数量，这种做法可能并非最优。可以考虑贝叶斯模型平均和可逆跳跃马尔可夫链蒙特卡罗（RJMCMC）方法，但跨维度抽样方法在提供可解释成分方面会带来挑战。

9 软件

R代码形式的软件及示例数据可在 https://github.com/HaoyiFu1993/ CBMOSE 获取。

参考文献

ADAMSON, LAUREN B AND FRICK, JANET E. (2003年). 静止面孔：一种共享实验范式的历史。Infancy 4(4), 451–473. BERTOLACCI, MICHAEL, ROSEN, ORI, CRIPPS, EDWARD AND CRIPPS, SALLY. (2022年). Adaptspec-x：多条非平稳时间序列的协变量依赖型谱建模。计算与图形统计学杂志 31(2), 436–454. DEMPSTER, ARTHUR P, LAIRD, NAN M AND RUBIN, DONALD B. (1977年). 通过EM算法从不完整数据中进行极大似然估计。英国皇家统计学会杂志：B辑（方法学）39(1), 1–22. ENLOW, MICHELLE BOSQUET, WHITE, MATTHEW T, HAILS, KATHERINE, CABRERA, IVAN AND WRIGHT, ROS-ALIND J. (2016年). 修订版婴儿行为问卷：美国文化及社会人口特征多样化样本中的因子结构。Infant Behavior and Development 43, 24–35. EUBANK, RANDALL L. (1999年). 非参数回归与样条平滑。CRC 出版社。GARTSTEIN, MARIA A, BRIDGETT, DAVID J, YOUNG, BRANDI N, PANKSEPP, JAAK AND POWER, THOMAS. (2013年). 努力控制的起源：婴儿与父母的贡献。Infancy 18(2), 149–183. GARTSTEIN, MARIA A AND ROTHBART, MARY K. (2003年). 通过修订版婴儿行为问卷研究婴儿气质。Infant behavior and development 26(1), 64–86. GELMAN, ANDREW. (2006年). 层级模型中方差参数的先验分布（对Browne和Draper文章的评论）。贝叶斯分析 1(3), 515–534. GELMAN, A, CARLIN, JB, STERN, HS, RUBIN, DB and others. (2003年). 贝叶斯数据分析。GILLMAN, MATTHEW W AND BLAISDELL, CAROL J. (2018年). 环境对儿童健康结果的影响：NIH的一项研究计划。Current opinion i n pediatrics 30(2), 260. GU, CHONG. (2013年). 平滑样条方差分析模型，第297卷。施普林格。GU, CHONG AND KIM, YOUNG-JU. (2002年). 惩罚似然回归：一般形式与高效近似。Canadian Journal of Statistics 30(4), 619–628. HE, LINCHEN, WANG, CHAN, HU, JIYUAN, GAO, ZHAN, FALCONE, EMILIA, HOLLAND, STEVEN M, BLASER, MARTIN J AND LI, HUILIN. (2022年). Arzimm：一种新型分析平台，可根据纵向微生物组研究推断微生物相互作用及群落稳定性。Frontiers in genetics 13. HIPWELL, ALISON E, TUNG, IRENE, NORTHRUP, JESSIE AND KEENAN, KATE. (2019年). 母亲童年期压力暴露与婴儿情绪反应性特征之间的跨代关联。Development and psychopathology 31(3), 887–898. HUERTA, GABRIEL, JIANG, WENXIN AND TANNER, MARTIN A. (2003年). 通过层级混合模型进行时间序列建模。Statistica Sinica, 1097–1118. JACOBS, ROBERT A, JORDAN, MICHAEL I, NOWLAN, STEVEN J AND HINTON, GEOFFREY E. (1991年). 局部专家的自适应混合模型。Neural computation 3(1), 79–87. JOBSIS, FRANS F. (1977年). 无创红外监测脑部与心肌氧供及循环参数。Science 198(4323), 1264–1267. JORDAN, MICHAEL I AND JACOBS, ROBERT A. (1994年). 专家层级混合模型与EM算法。Neural computation 6(2), 181–214. KAKIZAWA, YOSHIHIDE, SHUMWAY, ROBERT H AND TANIGUCHI, MASANOBU. (1998年). 多元时间序列的判别与聚类。美国统计学会会刊 93(441), 328–340. KEENAN, KATE, HIPWELL, ALISON, CHUNG, TAMMY, STEPP, STEPHANIE, STOUTHAMER-LOEBER, MAGDA, LOEBER, ROLF AND MCTIGUE, KATHLEEN. (2010年). 匹兹堡女孩研究：概述与初步发现。Journal of Clinical Child & Adolescent Psychology 39(4), 506–521. KEOGH, EAMONN J AND PAZZANI, MICHAEL J. (2000年). 一种用于大型时间序列数据库快速相似性搜索的简单降维技术。载于：亚太知识发现与数据挖掘会议。施普林格。第122–133页。KIMELDORF, GEORGE S AND WAHBA, GRACE. (1970年). 随机过程的贝叶斯估计与样条平滑之间的对应关系。The Annals of Mathematical Statistics 41(2), 495–502. KRAFTY, ROBERT T, HALL, MARTICA AND GUO, WENSHENG. (2011年). 函数型混合效应谱分析。Biometrika 98(3), 583–598。

KRAFTY, ROBERT T, ROSEN, ORI, STOFFER, DAVID S, BUYSSE, DANIEL J AND HALL, MARTICA H. (2017年). 基于重复多元时间序列的条件谱分析及其在夜间生理学中的应用. 美国统计学会会刊**112**(520), 1405–1416. LI, CEN, BISWAS, GAUTAM, DALE, MIKE AND DALE, PAT. (2001年). 使用基于隐马尔可夫模型的时序数据聚类构建生态动态模型——一项初步研究. 载于：智能数据分析国际研讨会. 施普林格. 第53–62页. LI, ZEDA AND KRAFTY, ROBERT T. (2019年). 多元时间序列的自适应贝叶斯时频分析.美国统计学会会刊**114**(525), 453–465. LIAO, T WARREN. (2005年). 时间序列数据聚类——综述. 模式识别**38**(11), 1857–1874. LIN, JESSICA, KEOGH, EAMONN AND TRUPPEL, WAGNER. (2003年). 流式时间序列聚类毫无意义. 载于：第八届*ACM SIGMOD*数据挖掘与知识发现研究问题研讨会论文集. 第56–65页. MAGRINI, ALESSANDRO. (2022年). 欧盟国家农业可持续性评估：一种基于群组的多元轨迹方法. *AStA*统计分析进展, 1–31. MAHARAJ, ELIZABETH ANN. (1999年). 平稳多元时间序列的比较与分类. 模式识别**32**(7), 1129–1138. MAHARAJ, ELIZABEH ANN, D'URSO, PIERPAOLO AND CAIADO, JORGE. (2019年). 时间序列聚类与分类. Chapman and Hall/CRC. NAGIN, DANIEL S, JONES, BOBBY L, PASSOS, VALERIA LIMA AND TREMBLAY, RICHARD E. (2018年). 基于群组的多轨迹建模. 医学研究统计方法**27**(7), 2015–2023. PAPASTAMOULIS, PANAGIOTIS AND ILIOPOULOS, GEORGE. (2010年). 一种基于人工分配的解决方案, 用于解决分布混合贝叶斯分析中的标签交换问题. 计算与图形统计学杂志**19**(2), 313–331. POLSON, NICHOLAS G, SCOTT, JAMES G AND WINDLE, JESSE. (2013年). 使用Pólya–Gamma潜变量进行逻辑斯蒂模型的贝叶斯推断. 美国统计学会会刊**108**(504), 1339–1349. ROSEN, ORI, STOFFER, DAVID S AND WOOD, SALLY. (2009年). 通过贝叶斯平滑样条混合模型进行局部谱分析. 美国统计学会会刊**104**(485), 249–262. ROSEN, ORI, WOOD, SALLY AND STOFFER, DAVID S. (2012年). Adaptspec：非平稳时间序列的自适应谱估计. 美国统计学会会刊**107**(500), 1575–1589. SANTOSA, HENDRIK, ZHAI, XUETONG, FISHBURN, FRANK AND HUPPERT, THEODORE. (2018年). NIRS Brain Analyzir工具箱. 算法**11**(5), 73. SPECKMAN, PAUL L AND SUN, DONGCHU. (2003年). 全贝叶斯样条平滑与内在自回归先验. *Biometrika***90**(2), 289–302. SPIEGELHALTER, DAVID J, BEST, NICOLA G, CARLIN, BRADLEY P AND VAN DER LINDE, ANGELIKA. (2002年). 模型复杂度与拟合度的贝叶斯度量. 英国皇家统计学会杂志：*B*辑（统计方法学）**64**(4), 583–639. SUN, ZHUOXIN, ROSEN, ORI AND SAMPSON, ALLAN R. (2007年). 多元伯努利混合模型及其在精神分裂症死后组织研究中的应用. 生物统计学**63**(3), 901–909. TRONICK, EDWARD, ALS, HEIDELISE, ADAMSON, LAUREN, WISE, SUSAN AND BRAZELTON, T BERRY. (1978年). 面对面互动中婴儿对矛盾信息所致受困状态的反应. 美国儿童精神病学会杂志**17**(1), 1–13. WAHBA, GRACE. (1978年). 非正规先验、样条平滑以及回归中防范模型误差的问题. 英国皇家统计学会杂志：*B*辑（方法学）**40**(3), 364–372. WAHBA, GRACE. (1980年). 对数周期图的自动平滑. 美国统计学会会刊**75**(369), 122–132. WAND, MATTHEW P, ORMEROD, JOHN T, PADOAN, SIMONE A AND FRÜHWIRTH, RUDOLF. (2011年). 复杂分布的平均场变分贝叶斯方法. 贝叶斯分析**6**(4), 847–900. WANG, XIAOZHE, WIRTH, ANTHONY AND WANG, LIANG. (2007年). 基于结构的统计特征与多元时间序列聚类. 载于：第七届*IEEE*国际数据挖掘会议（*ICDM 2007*）. IEEE. 第351–360页. WANG, Y UEDONG. (2011年). 平滑样条：方法与应用. CRC出版社.

WOOD, SALLY A, JIANG, WENXIN 和 TANNER, MARTIN.（2002年）。用于空间自适应非参数回归的样条贝叶斯混合。*Biometrika***89**(3)，513–528。WOOD, SIMON N.（2006年）。广义加性模型：*R*语言入门。Chapman and Hall/CRC。

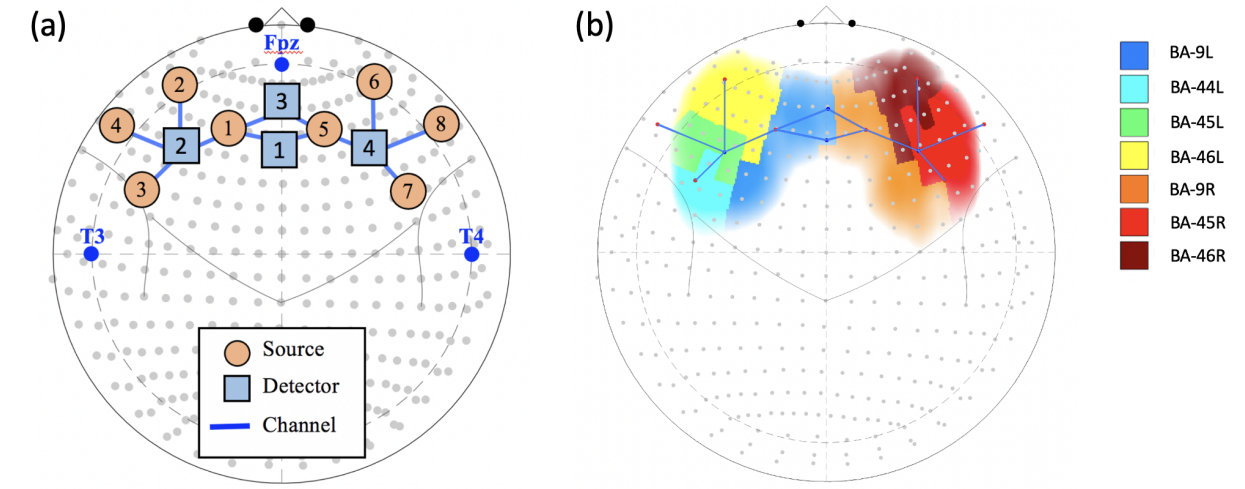

图 1：功能性近红外光谱探头配置。(a) 8 个光源、4 个探测器和 12 个通道的位置。每个通道由一个光源和一个探测器连接（蓝线）。(b) 功能性近红外光谱探头覆盖的布罗德曼分区。

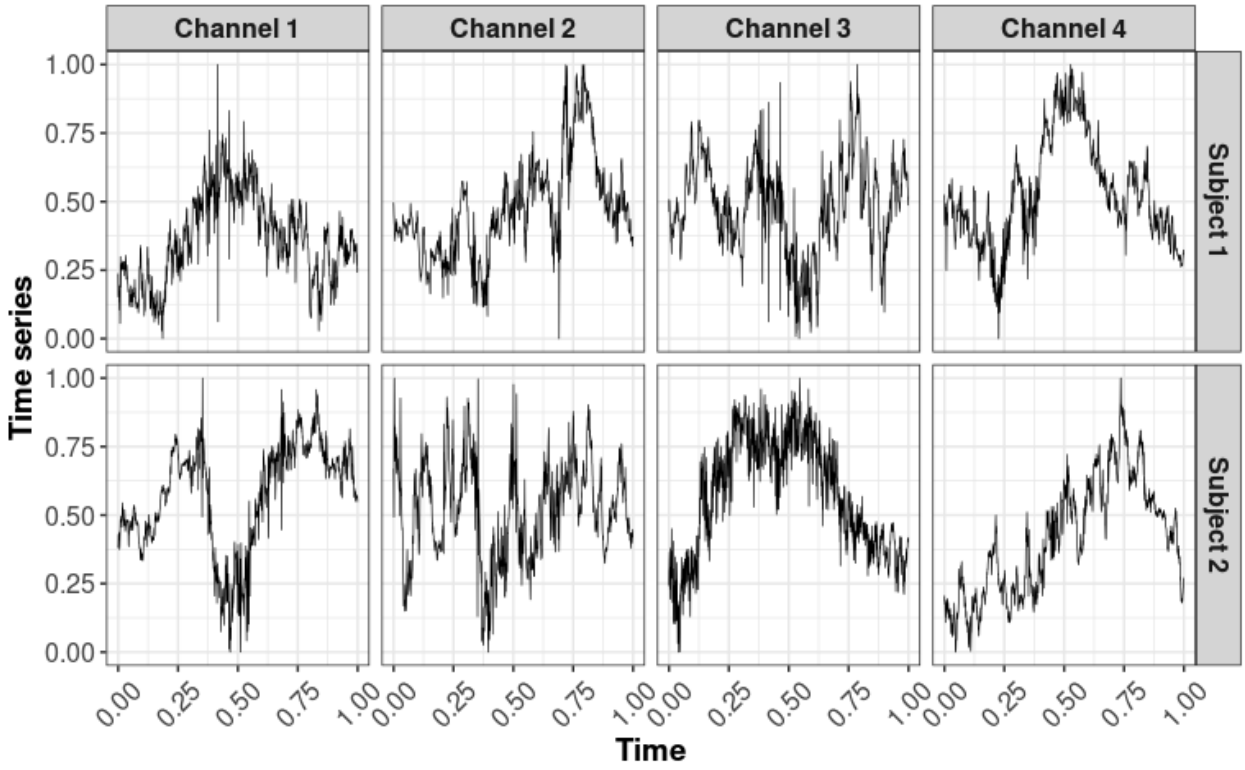

图 2：一个示例　　展示了从两名选定受试者和四个选定通道获得的功能性近红外光谱时间序列。测量值是氧合血红蛋白浓度

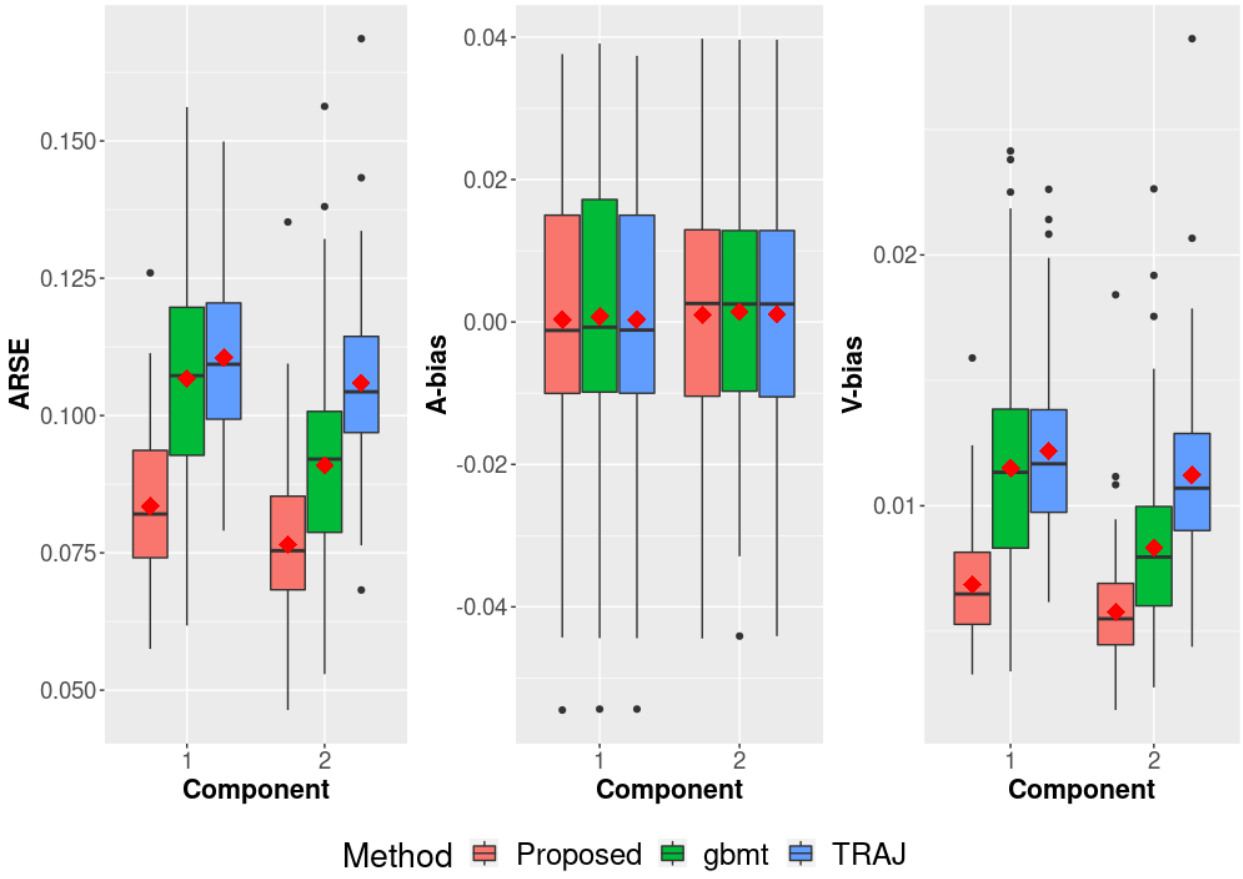

图3：基于100次重复的150组长度为50的双成分三元时间序列，对各分量估计轨迹的平均均方根误差（ARSE）、平均偏差（平均偏差）和偏差方差（V偏差）绘制的箱线图。将所提方法与R软件包gbmt及SAS中的TRAJ过程进行了比较。菱形标记表示各方法和分量的均值统计量。

表1：基于100次重复的 N 组长度为 n的双成分三变量时间序列，双成分三变量模型中各逻辑斯蒂参数的均方根误差（RMSEs）。将所提方法的均方根误差与SAS中的TRAJ过程进行了比较。参数 δ_0、 δ_1、 δ_2 和 δ_3 分别为截距、第一、第二和第三个逻辑斯蒂参数。逻辑斯蒂参数的真实值分别为5， -3.5，1，0。1

n	N	方法	δ_0	δ_1	δ_2	δ_3
50	150	所提方法	0.89	0.52	0.29	0.32
		TRAJ	1.57	0.87	0.36	0.34
70	150	所提方法	0.86	0.50	0.29	0.31
		TRAJ	1.55	0.86	0.36	0.34
50	250	所提方法	0.77	0.40	0.22	0.23
		TRAJ	0.96	0.50	0.23	0.24
70	250	所提方法	0.77	0.41	0.22	0.23
		TRAJ	0.97	0.51	0.24	0.24

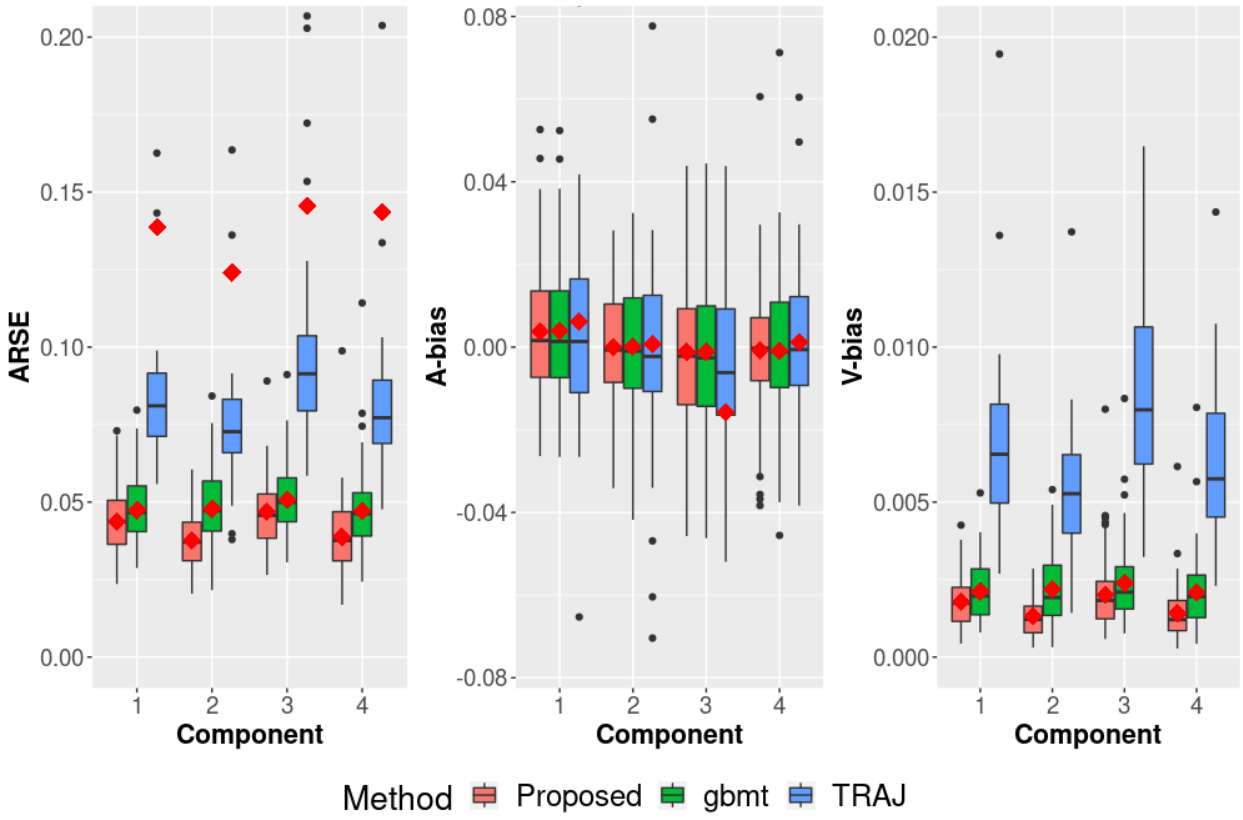

图4：基于100次重复（每次包含150条长度为50的四分量双变量时间序列）得到的各分量估计轨迹的平均均方根误差（ARSE）、平均偏差（A-bias）和偏差方差（V-bias）的箱线图。将所提方法与R软件包gbmt和SAS中的TRAJ过程进行了比较。菱形标记表示各方法和分量的均值统计量。为便于观察，所有箱线图均进行了放大显示。

表2：基于100次重复（每次包含150条长度为50的四分量双变量时间序列）得到的四分量双变量模型中各逻辑斯蒂参数的均方根误差（RMSEs）。将所提方法的均方根误差与SAS中的TRAJ过程进行了比较。参数 δ_0、δ_1、δ_2 和 δ_3 分别为截距、第一、第二和第三逻辑斯蒂参数。第四分量用作参考分量。逻辑斯蒂参数的真实值为5，-3.5，1，0。1（第一分量）、-4，2。5，-2，-0.2（第二分量）、3，-2，0。8，0。2（第三分量）。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。

n	N	方法	比较	δ_0	δ_1	δ_2	δ_3
50	150	所提方法	第一分量与第四分量	0.81	0.53	0.30	0.39
			第二分量与第四分量	1.11	0.46	0.42	0.36
			第三分量与第四分量	0.89	0.42	0.28	0.34
		TRAJ	第一分量与第四分量	1.20	0.74	0.35	0.41
			第二分量与第四分量	3.81	2.27	1.33	0.49
			第三分量与第四分量	2.07	1.33	0.76	0.32

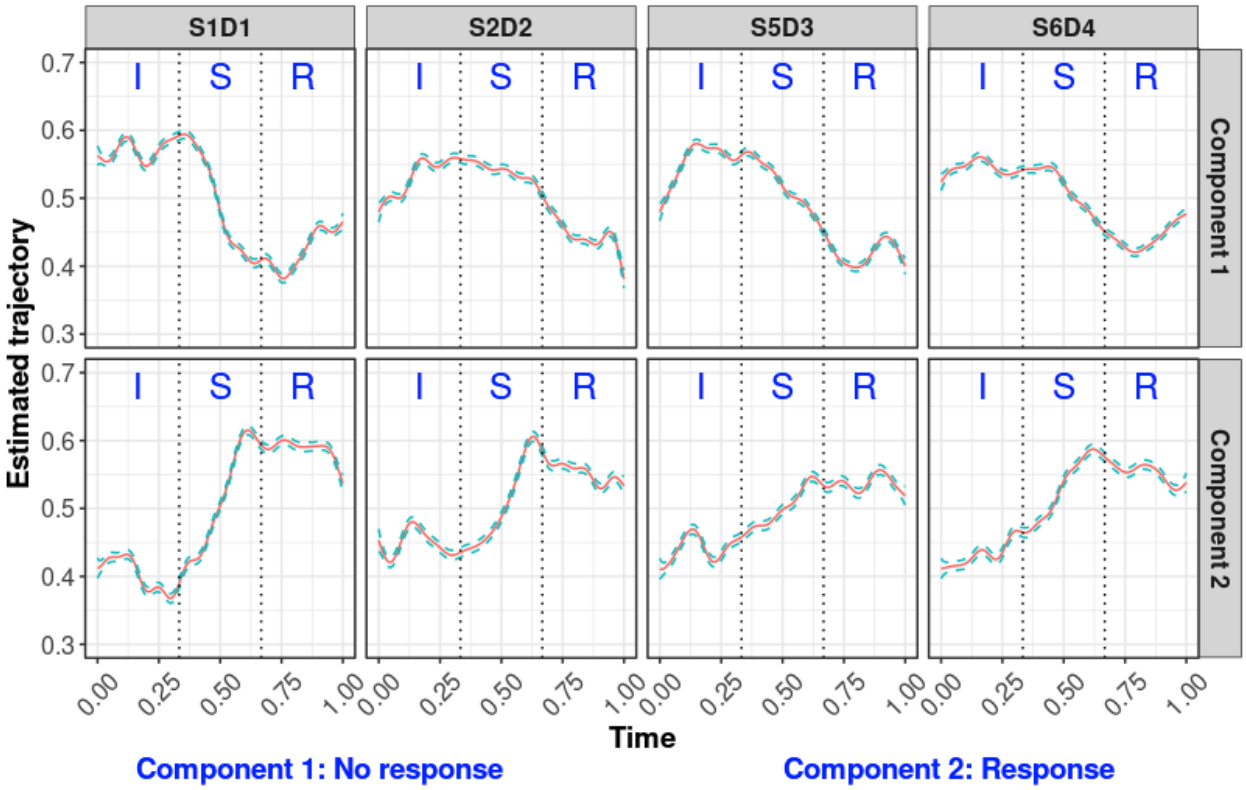

图5：包含四个选定通道的两分量模型的估计轨迹。**I**：互动 **S**：静止脸 **R**：恢复。红色曲线表示后验均值，两条绿色虚线表示95%逐点可信区间。

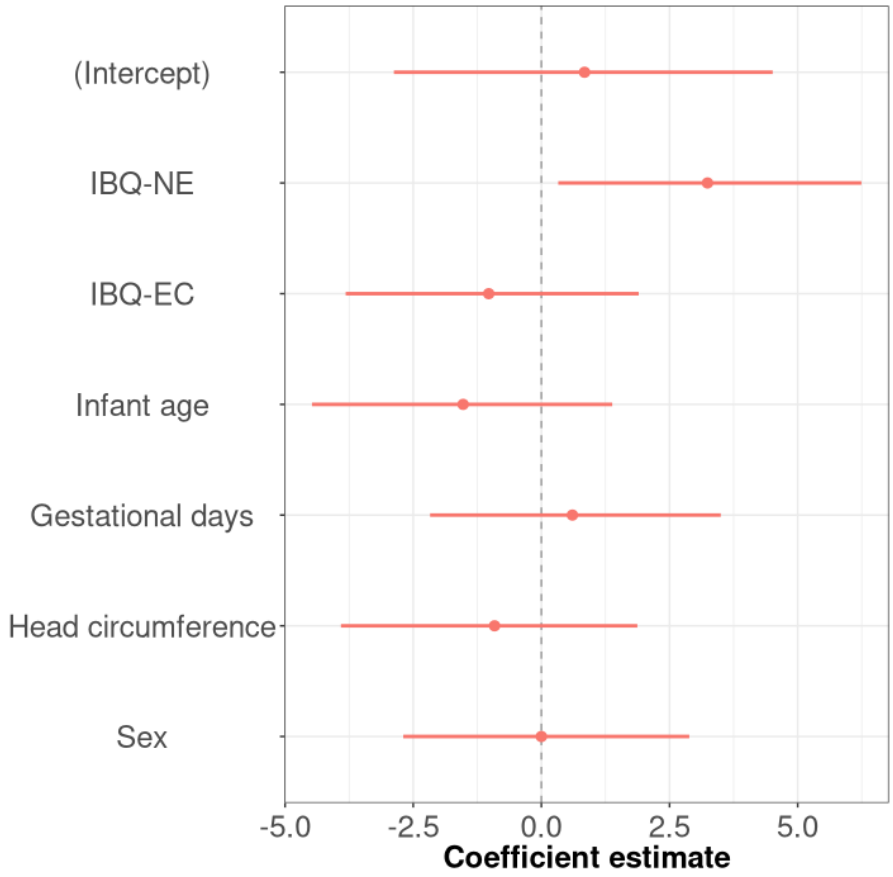

图6：逻辑回归系数 双分量模型中每个协变量的系数估计值和 95% 可信区间。

10 补充材料

附录 A：抽样方案的详细说明

如论文第 5 节所述，采用吉布斯抽样以便进行贝叶斯推断。我们将第 g 个分量和第 k 个条目的参数记为 $\boldsymbol{\Theta}_{gk}=(\boldsymbol{\theta}'_{gk},\ \tau^2_{gk},\ \sigma^2_{gk},\ \boldsymbol{\delta}^{*\prime}_g,\ \kappa^2_{\zeta g})'$，并从相应的条件后验分布中抽取该向量中的参数。令 ℓ 为当前吉布斯抽样迭代；下文给出在第 $(\ell+1)$ 次迭代中抽取参数的详细吉布斯抽样步骤。

1. **基函数系数的抽样**
对于每个分量 g 和时间序列条目 k，根据第 3.3 节中的增广似然，以及第 4.1 节所述关于 $\boldsymbol{\theta}_{gk}=(\boldsymbol{\alpha}'_{gk},\ \boldsymbol{\beta}'_{gk})'$ 的先验分布，$(\boldsymbol{\theta}^{(\ell+1)}_{gk}\mid \boldsymbol{y},\ \boldsymbol{S},\ \tau^{2(\ell)}_{gk},\ \sigma^{2(\ell)}_{gk})$ 的条件后验分布为：

$$
\begin{aligned}
p(\boldsymbol{\theta}^{(\ell+1)}_{gk}\mid \boldsymbol{y},\boldsymbol{S},\tau^{2(\ell)}_{gk},\sigma^{2(\ell)}_{gk}) &\propto p(\boldsymbol{y}\mid \boldsymbol{S},\boldsymbol{\theta}^{(\ell+1)}_{gk},\sigma^{2(\ell)}_{gk})\cdot p(\boldsymbol{\theta}^{(\ell+1)}_{gk}\mid \tau^{2(\ell)}_{gk})\\
&\propto \prod_{i=1}^{N}\Big\{(\sigma^2_{gk})^{-n/2}\exp\big[-\frac{1}{2\sigma^2_{gk}}(\boldsymbol{y}_{ik}-\boldsymbol{S}\boldsymbol{\theta}_{gk})'(\boldsymbol{y}_{ik}-\boldsymbol{S}\boldsymbol{\theta}_{gk})\big]\Big\}^{z_{ig}}\\
&\times|\boldsymbol{D}_{gk}|^{-1/2}\exp\Big(-\frac{1}{2}\boldsymbol{\theta}_{gk}\boldsymbol{D}^{-1}_{gk}\boldsymbol{\theta}_{gk}\Big)\\
&\propto \exp\Big\{-\frac{1}{2\sigma^2_{gk}}\big[\sum_{i=1}^{N}z_{ig}(\boldsymbol{y}_{ik}-\boldsymbol{S}\boldsymbol{\theta}_{gk})'(\boldsymbol{y}_{ik}-\boldsymbol{S}\boldsymbol{\theta}_{gk})+\boldsymbol{\theta}'_{gk}\sigma^2_{gk}\boldsymbol{D}^{-1}_{gk}\boldsymbol{\theta}_{gk}\big]\Big\}\\
&\propto \exp\Big[-\frac{1}{2\sigma^2_{gk}}(\boldsymbol{\theta}_{gk}-\boldsymbol{u}_{gk})'(\boldsymbol{\Lambda}_{gk})^{-1}(\boldsymbol{\theta}_{gk}-\boldsymbol{u}_{gk})\Big]\\
&\sim N(\boldsymbol{u}_{gk},\sigma^2_{gk}\boldsymbol{\Lambda}_{gk}),
\end{aligned}
$$

其中，$\boldsymbol{\Lambda}_{gk}=(N_{(\ell)g}\ \boldsymbol{S}'\boldsymbol{S}+\sigma^2_{gk}\boldsymbol{D}^{-1}_{gk})^{-1}$、$\boldsymbol{u}_{gk}=\boldsymbol{\Lambda}_{gk}\sum_{i=1}^{N}z_{ig}\boldsymbol{S}'\boldsymbol{y}_{ik}$、$N^{(\ell)}_g$表示第 g个分量中当前的受试者数量，$\boldsymbol{D}_{gk}=\mathrm{diag}(\sigma^2_\alpha\mathbf{1}_2,\tau^2_{gk}\mathbf{1}_m)$ 是 $\boldsymbol{\theta}_{gk}$的先验协方差矩阵。因此，对于每个分量 g和条目 k，我们从（$\boldsymbol{\theta}^{(\ell+1)}_{gk}\mid \boldsymbol{y},\boldsymbol{S},\tau^{2}{}_{(\ell)gk},\sigma^{2(\ell)}_{gk}$）$\sim N$（$\boldsymbol{u}_{gk},\sigma^2_{gk}\boldsymbol{\Lambda}_{gk}$）中抽样 $\boldsymbol{\theta}^{(\ell+1)}_{gk}$。

2. **抽样误差方差**
Gelman（2006年）提出使用半-t 分布作为尺度参数的先验。我们遵循 Wand *andothers*（2011年）的做法，将第4.3节中的半-t 先验表示为逆伽马分布的尺度混合，具体如下

$$
(\sigma^2_{gk}\mid a_{\sigma_{gk}})\sim IG\Big(\frac{\nu_\sigma}{2},\frac{\nu_\sigma}{a_{\sigma_{gk}}}\Big),a_{\sigma_{gk}}\sim IG\Big(\frac{1}{2},\frac{1}{A^2_\sigma}\Big).
$$

因此，潜变量$a_{\sigma_{gk}}$ 的条件后验分布为

$$
p(a^{(\ell+1)}_{\sigma_{gk}}\mid \sigma^{2(\ell)}_{gk})\propto\exp\Big[-\frac{1}{a_{\sigma_{gk}}}\Big(\frac{\nu_\sigma}{\sigma^2_{gk}}+\frac{1}{A^2_\sigma}\Big)\Big]\times(a_{\sigma_{gk}})^{-(\frac{1}{2}+1+\frac{\nu_\sigma}{2})},
$$

其中为 $IG\Big(\underbrace{\tfrac{\nu_\sigma+1}{2}}, \underbrace{\tfrac{\nu_\sigma}{\sigma^2_{gk}}+\tfrac{1}{A^2_\sigma}}\Big)$。记 $\boldsymbol{\epsilon}_{igk}$ 为分量 g 的时间序列 $\boldsymbol{y}_{ik}$ 的误差向量，则有 $\boldsymbol{\epsilon}_{igk}=\boldsymbol{y}_{ik}-\underline{\boldsymbol{S}\boldsymbol{\theta}_{gk}}$，其中 $\underline{\boldsymbol{\epsilon}_{igk}}\sim N$（$\mathbf{0}$，$\sigma^2_{gk}\boldsymbol{I}_n$）。误差方差的条件分布为

$$
\begin{aligned}
p(\sigma^{2(\ell+1)}_{gk}\mid \boldsymbol{\epsilon}^{(\ell+1)}_{igk},a^{(\ell+1)}_{\sigma_{gk}})&\propto p(\boldsymbol{\epsilon}^{(\ell+1)}_{igk}\mid \sigma^{2(\ell+1)}_{gk})\cdot p(a^{(\ell+1)}_{\sigma_{gk}}\mid \sigma^{2(\ell+1)}_{gk})\cdot p(\sigma^{2(\ell+1)}_{gk})\\
&\propto\prod_{i=1}^{N}\Big[(\sigma^2_{gk})^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2_{gk}}\boldsymbol{\epsilon}'_{igk}\boldsymbol{\epsilon}_{igk}\Big)\Big]^{z_{ig}}\times(\sigma^2_{gk})^{-(\frac{\nu_\sigma}{2}+1)}\exp\Big(-\frac{\nu_\sigma}{\sigma^2_{gk}a_{\sigma_{gk}}}\Big)\\
&\propto(\sigma^2_{gk})^{-(\frac{n}{2}N^{(\ell)}_g+\frac{\nu_\sigma}{2}+1)}\exp\Big[-\frac{1}{\sigma^2_{gk}}\Big(\frac{\sum_{i=1}^{N}z_{ig}\boldsymbol{\epsilon}'_{igk}\boldsymbol{\epsilon}_{igk}}{2}+\frac{\nu_\sigma}{a_{\sigma_{gk}}}\Big)\Big],
\end{aligned}
$$

即为 $IG\Big(\underbrace{\tfrac{nN^{(\ell)}_g+\nu_\sigma}{2}}, \underbrace{\tfrac{\sum_{i=1}^{N}z_{ig}\boldsymbol{\epsilon}'_{igk}\boldsymbol{\epsilon}_{igk}}{2}+\tfrac{\nu_\sigma}{a_{\sigma_{gk}}}}\Big)$。抽样方案首先对 $(a^{(\ell+1)}_{\sigma_{gk}}\mid \sigma^{2(\ell)}_{gk})$ 进行抽样，然后对 $(\sigma^{2(\ell+1)}_{gk}\mid \underline{\boldsymbol{\epsilon}^{(\ell+1)}_{igk}},\ \underline{a^{(\ell+1)}_{\sigma_{gk}}})$ 进行抽样。

3. **平滑参数的抽样**

平滑参数 τ_{gk}^2的抽取方式与误差方差类似。我们首先抽取 $(a_{\tau_{gk}}^{(\ell+1)} \mid \tau^{2}{}_{(\ell)gk})$ $\sim IG\Big(\frac{\nu_\tau+1}{2}, \frac{\nu_\tau}{\tau_{gk}^2} + \frac{1}{A_\tau^2}\Big)$。

平滑参数的条件后验分布为

$$\begin{aligned} p(\tau_{gk}^{2(\ell+1)} \mid \boldsymbol{\beta}_{gk}^{(\ell+1)}, a_{\tau_{gk}}^{(\ell+1)}) &\propto p(\boldsymbol{\beta}_{gk}^{(\ell+1)} \mid \tau_{gk}^{2(\ell+1)}) \cdot p(a_{\tau_{gk}}^{(\ell+1)} \mid \tau_{gk}^{2(\ell+1)}) \cdot p(\tau_{gk}^{2(\ell+1)}) \\ &\propto (\tau_{gk}^2)^{-\frac{m+\nu_\tau}{2}} \exp\Big[-\frac{1}{\tau_{gk}^2}\Big(\frac{\nu_\tau}{a_{\tau_{gk}}} + \frac{\boldsymbol{\beta}_{gk}'\boldsymbol{\beta}_{gk}}{2}\Big)\Big], \end{aligned}$$

其中为 $IG\Big(\frac{\nu_\tau+m}{2}, \frac{\boldsymbol{\beta}_{gk}'\boldsymbol{\beta}_{gk}}{2} + \frac{\nu_\tau}{a_{\tau_{gk}}}\Big)$。抽样方案首先对 $(a_{\tau_{gk}}^{(\ell+1)} \mid \tau^{2}{}_{(\ell)gk})$ 进行抽样，然后对 $(\tau_{gk}^{2(\ell+1)} \mid \boldsymbol{\beta}_{gk}^{(\ell+1)}, a_{\tau_{gk}}^{(\ell+1)})$ 进行抽样。

4. **逻辑斯蒂参数的抽样**

令 $\boldsymbol{\delta}_g^* = (\boldsymbol{\delta}_g^T, \boldsymbol{\zeta}_g^T)^T$ 表示第 g 个分量的逻辑斯蒂参数及所有随机截距的汇总。根据第 3.2 节的对数几率以及第 4.4 节中所述的相应先验分布，$(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \kappa_{\zeta g}^{2(\ell)})$ 的条件后验分布为

$$\begin{aligned} p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \kappa_{\zeta g}^{2(\ell)}) &\propto p(z_{ig}^{(\ell)} = 1 \mid \boldsymbol{V}^*, \boldsymbol{\delta}_g^{*(\ell+1)}) \cdot p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell)}) \\ &= \prod_{i=1}^{N} \Big[\frac{\exp(\boldsymbol{V}_i^{*\prime}\boldsymbol{\delta}_g^*)}{\sum_{h=1}^{G}\exp(\boldsymbol{V}_i^{*\prime}\boldsymbol{\delta}_h^*)}\Big]^{z_{ig}} p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell)}), \end{aligned}$$

其中，$\boldsymbol{V}^* = (\boldsymbol{V}_1^*, \ldots\ldots, \boldsymbol{V}_N^*)'$是一个 $N \times (P+1)$矩阵，$\boldsymbol{V}_i^*$表示受试者 i的所有协变量（包括截距）。为了从 $p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \kappa_{\zeta g}^{2(\ell)})$的后验分布中进行抽样，我们采用 Polson等人（2013年）提出的 Pólya-Gamma 数据增广策略，引入一个来自 Pólya-Gamma分布的潜变量 ω_{ig}。因此，逻辑斯蒂参数的条件后验分布为

$$\begin{aligned} p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \omega_{ig}^{(\ell+1)}, \kappa_{\zeta g}^{2(\ell)}) &\propto p(z_{ig}^{(\ell)} = 1 \mid \boldsymbol{V}^*, \omega_{ig}^{(\ell+1)}, \boldsymbol{\delta}_g^{*(\ell+1)}) \cdot p(\omega_{ig}^{(\ell+1)} \mid \boldsymbol{V}^*, \boldsymbol{\delta}_g^{*(\ell)}) \\ &\quad \cdot p(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell)}) \\ &\propto \exp\Big(-\frac{\omega_{ig}\eta_{ig}^2}{2}\Big) \cdot p(\omega_{ig} \mid 1, 0) |\boldsymbol{B}_g|^{-P/2} \exp\Big(-\frac{1}{2}\boldsymbol{\delta}_g^{*\prime}\boldsymbol{B}_g^{-1}\boldsymbol{\delta}_g^*\Big), \end{aligned}$$

其中，$\eta_{ig} = \boldsymbol{V}_i^{*\prime}\boldsymbol{\delta}_g^* - C_{ig}$和 $C_{ig} = \log\sum_{h\neq j}\exp(\boldsymbol{V}_i^{*\prime}\boldsymbol{\delta}_h^*)$，$p(\omega_{ig} \mid 1, 0)$ 是 Pólya-Gamma分布 $PG(b, c)$，其中 $b = 1$ 且 $c = 0$，$\boldsymbol{B}_g$是第 4.4 节的先验协方差矩阵，$\boldsymbol{B}_g = \text{diag}(\sigma_{\delta g}^2\mathbf{1}_{P+1}, \kappa_{\zeta g}^2\mathbf{1}_N)$。假设在 $\boldsymbol{\delta}_g^*$上采用共轭先验 $N(\mathbf{0}, \boldsymbol{B}_g)$，Pólya-Gamma潜变量的后验分布为

$$(\omega_{ig}^{(\ell+1)} \mid \boldsymbol{V}^*, \boldsymbol{\delta}_g^{*(\ell)}) \sim PG(1, \eta_{ig}).$$

因此，逻辑斯蒂参数（包括随机截距）的条件分布为

$$(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \omega_{ig}^{(\ell+1)}, \kappa_{\zeta g}^{2(\ell)}) \sim N(\boldsymbol{M}_g, \boldsymbol{\Sigma}_g),$$

其中 $\boldsymbol{\Sigma}_g = (\boldsymbol{V}^{*\prime}\Omega_g\boldsymbol{V}^* + \boldsymbol{B}_g^{-1})^{-1}$，$\boldsymbol{M}_g = \boldsymbol{\Sigma}_g\big[\boldsymbol{V}^{*\prime}(\Omega_g\boldsymbol{C}_g + \boldsymbol{\xi}_g)\big]$，$\Omega_g = \text{diag}(\omega_{1g}, \cdots, \omega_{Ng})$，$\boldsymbol{C}_g = (C_{1g}, \cdots, C_{Ng})'$，以及 $\boldsymbol{\xi}_g = (\xi_{1g}, \cdots, \xi_{Ng})'$，其中 $\xi_{ig} = z_{ig} - \frac{1}{2}$。因此，$\boldsymbol{\delta}_g^{*(\ell+1)}$首先通过抽样 $(\omega_{ig}^{(\ell+1)} \mid \boldsymbol{V}^*, \boldsymbol{\delta}_g^{*(\ell)})$，然后抽样 $(\boldsymbol{\delta}_g^{*(\ell+1)} \mid \boldsymbol{V}^*, z_{ig}^{(\ell)}, \omega_{ig}^{(\ell+1)}, \kappa_{\zeta g}^{2(\ell)})$ 来生成。

5. **随机截距方差的抽样**

类比于误差方差和样条平滑参数的抽样，我们首先抽取 $(a_{\kappa_g}^{(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell)})$ $\sim IG\Big(\frac{\nu_\kappa+1}{2}, \frac{\nu_\kappa}{\kappa_{\zeta g}^2} + \frac{1}{A_\kappa^2}\Big)$。随机截距方差的条件后验分布为

$$\begin{aligned} p(\kappa_{\zeta g}^{2(\ell+1)} \mid \boldsymbol{\zeta}_g^{(\ell+1)}, a_{\kappa_g}^{(\ell+1)}) &\propto p(\boldsymbol{\zeta}_g^{(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell+1)}) \cdot p(a_{\kappa_g}^{(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell+1)}) \cdot p(\kappa_{\zeta g}^{2(\ell+1)}) \\ &\propto (\kappa_{\zeta g}^2)^{-\frac{N+\nu_\kappa}{2}+1} \exp\Big[-\frac{1}{\kappa_{\zeta g}^2}\Big(\frac{\nu_\kappa}{a_{\kappa_g}} + \frac{\boldsymbol{\zeta}_g^T\boldsymbol{\zeta}_g}{2}\Big)\Big], \end{aligned}$$

即为 $IG\Big(\frac{\nu_\kappa+N}{2}, \frac{\boldsymbol{\zeta}_g^T\boldsymbol{\zeta}_g}{2} + \frac{\nu_\kappa}{a_{\kappa_g}}\Big)$。抽样方案首先对 $(a_{\kappa_g}^{(\ell+1)} \mid \kappa_{\zeta g}^{2(\ell)})$ 进行抽样，然后对 $(\kappa_{\zeta g}^{2(\ell+1)} \mid \boldsymbol{\zeta}_g^{(\ell+1)}, a_{\kappa_g}^{(\ell+1)})$ 进行抽样。

6. **计算混合权重**
   在抽取 给定设计矩阵 V_i^*，计算每个分量的混合权重 $\pi_{ig}^{(\ell+1)}$，即 $\boldsymbol{\delta}_g^*$ by

$$p(\pi_{ig}^{(\ell+1)} \mid V_i^*, \boldsymbol{\delta}_g^{*(\ell+1)}) = \frac{\exp(\boldsymbol{V}_i^{*T}\boldsymbol{\delta}_g^*)}{\sum_{h=1}^{G}\exp(\boldsymbol{V}_i^{*T}\boldsymbol{\delta}_h^*)}.$$

7. **抽取潜在指标**
   抽取所有参数并计算混合权重后，最后一个吉布斯步骤是通过抽取潜在指标 z_{ig}，将受试者分配到不同的分量。如第 3.3 节所述，这些指标的条件后验分布为

$$p(z_{ig}^{(\ell+1)} = 1 \mid \boldsymbol{y}, \boldsymbol{S}, \boldsymbol{\Theta}^{(\ell+1)}, \pi_{ig}^{(\ell+1)}) = \frac{\pi_{ig}\prod_{k=1}^{K} f_{gk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{gk})}{\sum_{h=1}^{G}\pi_{ih}\prod_{k=1}^{K} f_{hk}(\boldsymbol{y}_{ik} \mid \boldsymbol{\Theta}_{hk})},$$

   这些指标服从多项分布。

附录B：补充模拟结果

附录B补充了论文正文中的模拟结果。为进一步展示所提方法的性能，我们在两种情形下进行模拟研究：三元时间序列的双成分混合，以及二元时间序列的四成分混合。模型公式见论文第6.1节。我们从估计轨迹和逻辑斯蒂参数两方面考察所提方法的性能。

表3给出了双成分三元模型中各分量的ARSE、平均偏差和V偏差的均值（标准差）。为展示所提方法在不同设置下的性能，我们考察了多元时间序列数量（$N = 150$，250）与每个时间序列长度（$n = 50$，70）的不同组合，并将所提方法与两种现有方法进行比较：R中的gbmt软件包（Magrini，2022年）和SAS中的TRAJ过程（Nagin及其他人，2018年）。表3中 $n = 50$ 和 $N = 150$ 的情形对应正文中的图3。论文表1给出了 n 和 N 取不同值时逻辑斯蒂参数的性能（均方根误差）。

$N = 150$ 长度为 $n = 50$ 的二元时间序列四成分混合中各分量的 ARSE、平均偏差和 V偏差的均值（标准差）见表4，该表对应于论文正文中的图4。该设置下逻辑斯蒂参数的均方根误差列于论文的表2中。表5–10展示了在四分量双变量模型情景下，对于不同时间序列长度 n 和时间序列数量 N的组合，估计轨迹和逻辑斯蒂参数的性能指标。

正如预期的那样，在不同设置下（即 n 和 N取不同值时，对于双分量三变量和四分量双变量两种情景），就各分量的估计轨迹而言，我们的所提方法优于现有的两种方法。所提方法能够获得更小的 ARSE 和 V偏差，而三种方法均能得到偏差非常小的估计轨迹。值得注意的是，在四分量双变量情景下，TRAJ 给出了更大的 ARSE、平均偏差和 V偏差均值，这是由于收敛问题导致若干重复样本的估计不精确所致。就逻辑斯蒂参数而言，在几乎所有比较中，我们的所提方法都优于 TRAJ，尤其是在截距 δ_0 和第一个协变量的斜率 δ_1方面。由于采用了贝叶斯方法，我们的所提方法对逻辑斯蒂参数给出了收缩估计；而 TRAJ 中使用的多项逻辑回归在完全分离和不平衡设计的情况下会产生膨胀的参数估计。

附录 C：更多真实数据结果

附录 C 介绍了主论文第 7 节之外的更多真实数据结果。我们的动机研究见论文第 2 节。图 7 展示了另一组四个通道（S1D3、S3D2、S5D4 和 S7D4）的三成分模型估计轨迹。根据主论文第 5.2 节提出的 DIC 选择准则，三成分模型被选为最佳模型。我们将第二成分称为混合反应成分，因为它涉及静止脸阶段不同通道中脑活动或血红蛋白水平的升高和降低。此外，图 8 展示了各协变量对应的逻辑回归系数估计值和 95% 可信区间。最后一个成分（第三成分）始终作为参照。我们得出了相同的结论：两个成分（成分 1 与 3、成分 2 与 3）的 IBQ-NE 得分估计值均为正，IBQ-EC 得分估计值均为负。

除了四通道分析外，我们还展示了所有通道（十二个通道）的结果。图 9、图 10、图 11展示了使用全部十二个通道的三分量模型中第一、第二和第三成分的估计轨迹

表3：对于长度为 n 的 N 双成分三元时间序列的 100 次重复，每个分量的估计轨迹的平均均方根误差（ARSE）、平均偏差（A-bias）和偏差方差（V-bias）的均值（标准差）。将所提方法与 R 软件包 gbmt 及 SAS 中的 TRAJ 过程进行了比较。C1 和 C2 分别表示第一分量和第二分量。均值通过对 100 次重复的估计值取平均计算得出。标准差为基于 100 次重复估计值计算的蒙特卡洛标准差。每个数值均按 $\times 10^2$ 报告。

n	N	方法	ARSE第一分量	平均偏差 第一分量	V偏差第一分量	ARSE 第二成分	平均偏差 第二成分	V偏差 第二成分
50	150	所提方法	8.35 (1.26)	0.03 (1.83)	0.68 (0.22)	7.65 (1.38)	0.10 (1.76)	0.58 (0.22)
		gbmt	10.67 (1.91)	0.03 (1.83)	1.15 (0.42)	9.08 (1.72)	0.10 (1.76)	0.83 (0.33)
		TRAJ	11.06 (1.48)	0.03 (1.83)	1.22 (0.33)	10.59 (1.52)	0.10 (1.76)	1.12 (0.34)
70	150	所提方法	7.16 (1.04)	0.24 (1.38)	0.51 (0.16)	6.53 (1.07)	-0.11 (1.42)	0.42 (0.14)
		gbmt	9.91 (1.96)	0.24 (1.38)	1.01 (0.40)	8.19 (1.65)	-0.11 (1.42)	0.68 (0.29)
		TRAJ	9.34 (1.10)	0.24 (1.38)	0.87 (0.22)	8.95 (1.13)	-0.11 (1.42)	0.80 (0.20)
50	250	所提方法	6.81 (1.02)	0.07 (1.33)	0.46 (0.14)	6.22 (0.94)	0.02 (1.31)	0.38 (0.12)
		gbmt	9.79 (1.91)	0.07 (1.33)	0.98 (0.40)	8.00 (1.53)	0.02 (1.31)	0.65 (0.26)
		TRAJ	8.70 (1.18)	0.07 (1.33)	0.76 (0.21)	8.20 (1.03)	0.02 (1.31)	0.67 (0.17)
70	250	所提方法	5.65 (0.84)	0.08 (1.00)	0.32 (0.10)	5.27 (0.82)	-0.06 (1.42)	0.27 (0.09)
		gbmt	9.15 (1.96)	0.08 (1.00)	0.87 (0.38)	7.43 (1.60)	-0.06 (1.42)	0.56 (0.26)
		TRAJ	7.18 (0.94)	0.08 (1.00)	0.52 (0.14)	6.80 (0.77)	-0.06 (1.42)	0.45 (0.10)

表4：基于100次重复的150组长度为50的四分量双变量时间序列，对每个分量的估计轨迹计算平均均方根误差（ARSE）、平均偏差（A-bias）和偏差方差（V-bias）的均值（标准差）。将所提方法与R软件包gbmt及SAS中的TRAJ过程进行了比较。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。均值通过对100次重复的估计值取平均计算。标准差为根据100次重复的估计值计算得到的蒙特卡洛标准差。每个数值均报告为 $\times 10^2$。

n	N	方法	ARSE第一分量	平均偏差 第一分量	V偏差第一分量	ARSE 第二分量	平均偏差 第二分量	V偏差 第二分量
50	150	所提方法	4.38 (1.04)	0.38 (1.59)	0.18 (0.08)	3.76 (0.87)	-0.01 (1.37)	0.13 (0.06)
		gbmt	4.75 (1.04)	0.38 (1.59)	0.21 (0.09)	4.79 (1.17)	0.01 (1.65)	0.22 (0.11)
		TRAJ	13.87 (15.59)	0.62 (9.91)	3.39 (9.24)	12.41 (13.42)	0.08 (9.74)	2.41 (5.37)
n	N	方法	ARSE 第三分量	平均偏差 第三分量	V偏差第三分量	ARSE 第四分量	平均偏差 第四分量	V偏差 第四分量
50	150	所提方法	4.69 (1.14)	-0.11 (1.83)	0.20 (0.12)	3.88 (1.15)	-0.09 (1.56)	0.14 (0.08)
		gbmt	5.08 (1.12)	-0.12 (1.83)	0.24 (0.12)	4.70 (1.32)	-0.09 (1.78)	0.21 (0.12)
		TRAJ	14.55 (14.82)	-1.58 (10.31)	3.24 (7.35)	14.36 (17.01)	0.12 (9.92)	3.99 (10.70)

表5：对150组长度为70的四分量双变量时间序列进行100次重复后，各分量估计轨迹的平均均方根误差（ARSE）、平均偏差（A-bias）和偏差方差（V-bias）的均值（标准差）。将所提方法与R软件包gbmt和SAS中的TRAJ过程进行比较。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。均值是对100次重复的估计值取平均计算得出的。标准差为基于100次重复估计值计算的蒙特卡洛标准差。每个数值均报告为 $\times 10^2$。

n	N	方法	ARSE第一分量	平均偏差 第一分量	V偏差第一分量	ARSE 第二分量	平均偏差 第二分量	V偏差 第二分量
70	150	所提方法	3.82 (0.95)	0.44 (1.30)	0.14 (0.07)	3.22 (0.85)	-0.07 (0.95)	0.10 (0.06)
		gbmt	4.05 (0.97)	0.44 (1.30)	0.16 (0.08)	4.11 (1.12)	-0.08 (1.15)	0.17 (0.10)
		TRAJ	13.51 (17.04)	-0.30 (9.34)	3.86 (10.73)	10.00 (10.11)	-0.25 (6.21)	1.64 (4.02)
n	N	方法	ARSE 第三分量	平均偏差 第三分量	V偏差第三分量	ARSE 第四分量	平均偏差 第四分量	V偏差 第四分量
70	150	所提方法	4.12 (0.90)	-0.29 (1.69)	0.15 (0.06)	3.52 (0.85)	0.24 (1.22)	0.12 (0.06)
		gbmt	4.38 (1.01)	-0.29 (1.69)	0.17 (0.07)	4.13 (0.99)	0.27 (1.40)	0.16 (0.09)
		TRAJ	13.03 (17.04)	0.30 (10.49)	3.86 (10.73)	11.77 (13.78)	0.39 (8.49)	2.57 (6.45)

表6：由100次重复中生成的150个长度为70的四分量双变量时间序列所构成的四分量双变量模型中各逻辑斯蒂参数的均方根误差（均方根误差）。将所提方法的均方根误差与SAS中的TRAJ过程进行了比较。参数 δ_0、δ_1、δ_2 和 δ_3 分别表示截距、第一、第二和第三逻辑斯蒂参数。第四分量用作参考分量。逻辑斯蒂参数的真实值为5、-3.5、1、0。1（第一成分）、-4、2。5、-2、-0.2（第二成分）、3、-2、0。8、0。2（第三成分）。C1、C2、C3和C4分别表示第一成分、第二成分、第三成分和第四成分。

n	N	方法	比较	δ_0	δ_1	δ_2	δ_3
70	150	所提方法	第一分量 vs 第四分量	0.81	0.51	0.29	0.41
			第二分量 vs 第四分量	1.42	0.73	0.58	0.36
			第三分量 vs 第四分量	1.05	0.58	0.37	0.31
		TRAJ	第一分量 vs 第四分量	1.13	0.66	0.31	0.45
			第二分量 vs 第四分量	3.12	1.66	0.99	0.55
			第三分量 vs 第四分量	1.15	0.74	0.48	0.35

表7：基于100次重复中250条长度为50的四分量双变量时间序列，计算各分量估计轨迹的平均均方根误差（ARSE）、平均偏差（A-bias）和偏差方差（V-bias）的均值（标准差）。将所提方法与R软件包gbmt及SAS中的TRAJ过程进行了比较。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。均值通过对100次重复的估计值取平均计算。标准差为100次重复估计值的蒙特卡洛标准差。每个数值均报告 $\times 10^2$。

n	N	方法	ARSE第一分量	平均偏差 第一分量	V偏差第一分量	ARSE 第二分量	平均偏差 第二分量	V偏差 第二分量
50	250	所提方法	3.42 (0.78)	0.18 (1.19)	0.11 (0.05)	2.86 (0.61)	-0.14 (0.98)	0.08 (0.04)
		gbmt	3.57 (0.85)	0.18 (1.19)	0.12 (0.06)	3.68 (0.79)	-0.15 (1.20)	0.13 (0.06)
		TRAJ	11.66 (15.03)	0.53 (10.63)	2.50 (6.30)	8.90 (9.38)	1.47 (7.81)	1.05 (2.16)
n	N	方法	ARSE 第三分量	平均偏差 第三分量	V偏差第三分量	ARSE 第四分量	平均偏差 第四分量	V偏差 第四分量
50	250	所提方法	3.93 (0.92)	-0.06 (1.48)	0.14 (0.07)	3.28 (0.76)	-0.10 (1.19)	0.10 (0.05)
		gbmt	4.16 (0.95)	-0.06 (1.49)	0.16 (0.07)	3.83 (0.83)	-0.13 (1.36)	0.14 (0.06)
		TRAJ	10.80 (9.92)	0.49 (5.78)	1.83 (3.70)	10.17 (12.54)	-0.17 (8.83)	1.84 (5.20)

表 8：基于 100 次重复（每次包含 250 个长度为 50 的四分量双变量时间序列）得到的四分量双变量模型各逻辑斯蒂参数的均方根误差。将所提方法的均方根误差与 SAS 中的 TRAJ 过程进行比较。参数 δ_0、 δ_1、 δ_2 和 δ_3 分别为截距、第一、第二和第三逻辑斯蒂参数。第四分量用作参照分量。逻辑斯蒂参数的真实值为 $5, -3.5, 1, 0.1$（第一分量）、 $-4, 2.5, -2, -0.2$（第二分量）、3, −2,0.8,0.2（第三分量）。C1、C2、C3 和 C4 分别表示第一、第二、第三和第四分量。

n	N	方法	比较	δ_0	δ_1	δ_2	δ_3
50	250	所提方法	第一分量 vs 第四分量	0.63	0.41	0.26	0.29
			第二分量与第四分量	1.00	0.46	0.40	0.27
			第三分量与第四分量	0.63	0.33	0.23	0.24
		TRAJ	第一分量与第四分量	0.91	0.56	0.30	0.28
			第二分量与第四分量	1.40	0.86	0.61	0.35
			第三分量与第四分量	2.24	1.40	0.85	0.27

表9：基于250组长度为70的四分量双变量时间序列的100次重复，对每个分量的估计轨迹计算的平均均方根误差（ARSE）、平均偏差（平均偏差）和偏差方差（V偏差）的均值（标准差）。将所提方法与R软件包gbmt及SAS中的TRAJ过程进行了比较。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。均值通过对100次重复的估计值取平均计算得出。标准差是根据100次重复的估计值计算的蒙特卡洛标准差。每个数值均已报告 $\times 10^2$。

n	N	方法	ARSE第一分量	平均偏差 第一分量	V偏差第一分量	ARSE 第二分量	平均偏差 第二分量	V偏差 第二分量
70	250	所提方法	2.94 (0.60)	-0.04 (1.06)	0.08 (0.04)	2.61 (0.57)	-0.01 (0.87)	0.06 (0.03)
		gbmt	3.10 (0.63)	-0.04 (1.06)	0.09 (0.04)	3.18 (0.70)	0.01 (1.05)	0.10 (0.05)
		TRAJ	13.52 (17.70)	-1.58 (11.09)	3.71 (8.80)	11.51 (14.85)	-0.19 (9.98)	2.54 (7.10)
n	N	方法	ARSE 第三分量	平均偏差 第三分量	V偏差第三分量	ARSE 第四分量	平均偏差 第四分量	V偏差 第四分量
70	250	所提方法	3.30 (0.76)	-0.02 (1.21)	0.10 (0.05)	2.85 (0.73)	-0.07 (0.97)	0.08 (0.04)
		gbmt	3.51 (0.79)	-0.01 (1.21)	0.12 (0.06)	3.26 (0.80)	-0.09 (1.06)	0.10 (0.05)
		TRAJ	13.07 (15.21)	1.48 (10.52)	2.90 (7.11)	10.68 (12.93)	0.65 (8.11)	2.16 (5.66)

表10：基于100次重复、每次包含250个长度为70的四分量双变量时间序列，对四分量双变量模型中各逻辑斯蒂参数的均方根误差（RMSEs）。将所提方法的均方根误差与SAS中的TRAJ过程进行比较。参数 δ_0、 δ_1、 δ_2 和 δ_3 分别为截距、第一、第二和第三逻辑斯蒂参数。第四分量用作参考分量。逻辑斯蒂参数的真实值为5、 -3.5、1、0。1（第一分量）， -4、2。5、 -2、 -0.2（第二分量），3、 -2、0。8、0。2（第三分量）。C1、C2、C3和C4分别表示第一、第二、第三和第四分量。

n	N	方法	比较	δ_0	δ_1	δ_2	δ_3
70	250	所提方法	第一分量与第四分量	0.64	0.40	0.26	0.28
			第二分量与第四分量	0.92	0.42	0.41	0.28
			第三分量与第四分量	0.63	0.31	0.23	0.23
		TRAJ	第一分量与第四分量	0.82	0.50	0.27	0.28
			第二分量与第四分量	1.47	0.86	0.61	0.36
			第三分量与第四分量	1.60	0.96	0.57	0.25

通道，分别。根据调整后的DIC，三成分模型被选为十二通道分析的最佳模型。我们将这三个成分分别命名为无响应、混合响应和反应成分。图12展示了各协变量对应的逻辑回归系数估计值和95%可信区间。

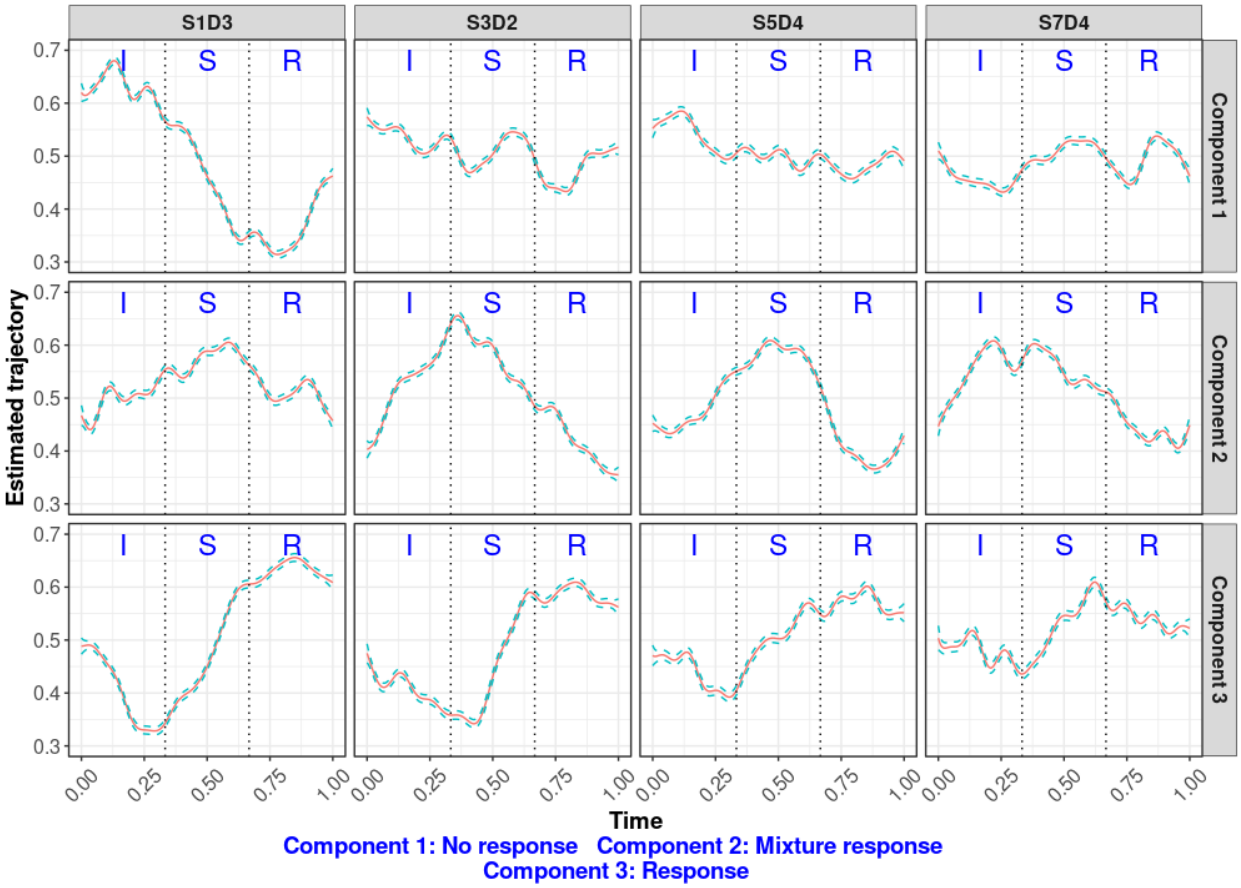

图 7：包含四个选定通道的三分量模型的估计轨迹。**I**：互动 **S**：静止脸 **R**：恢复。红色曲线表示后验均值，两个绿色虚线表示 95%逐点可信区间。

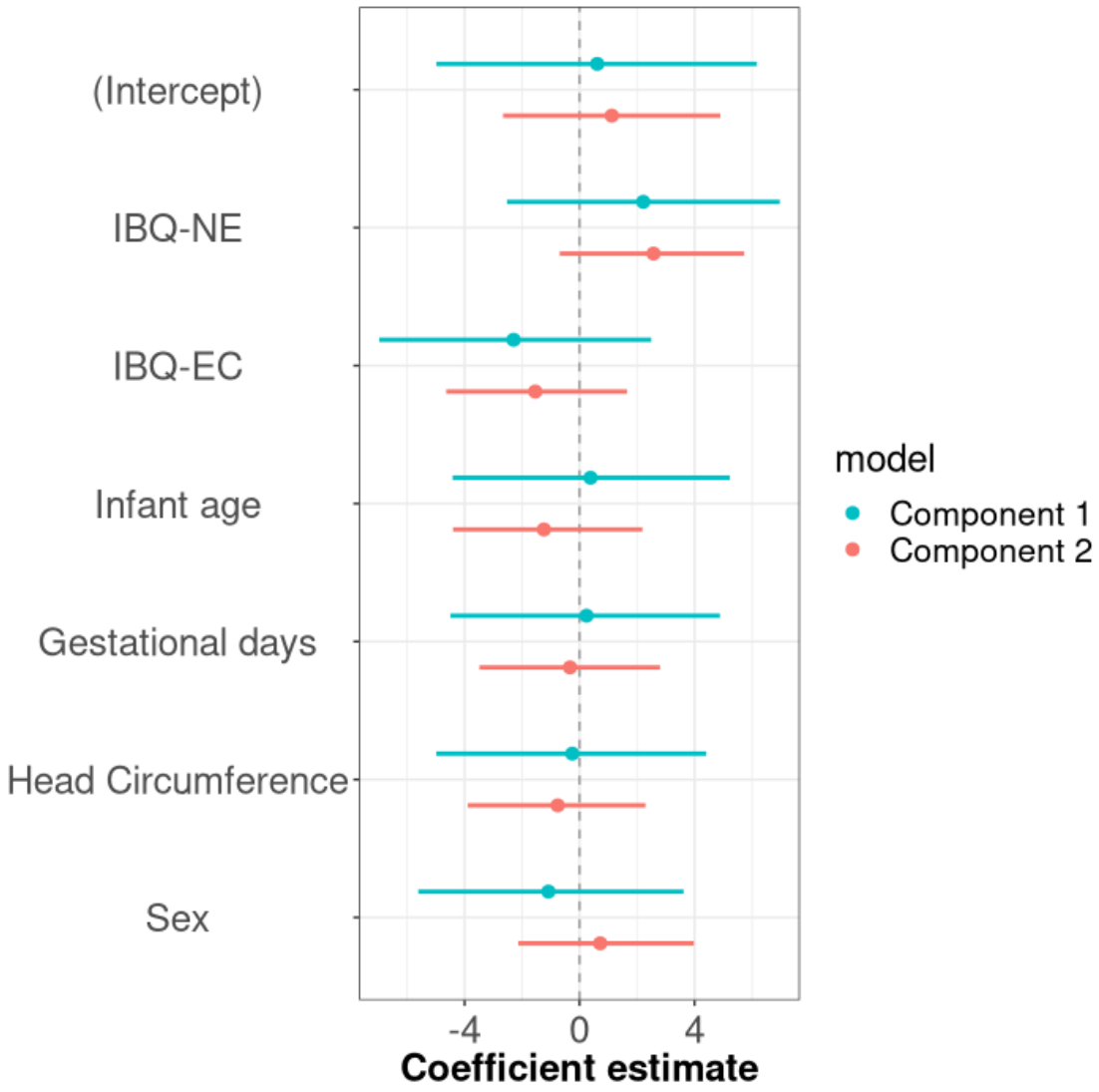

图 8：三分量模型中每个协变量对应的逻辑回归系数估计值和 95% 可信区间。

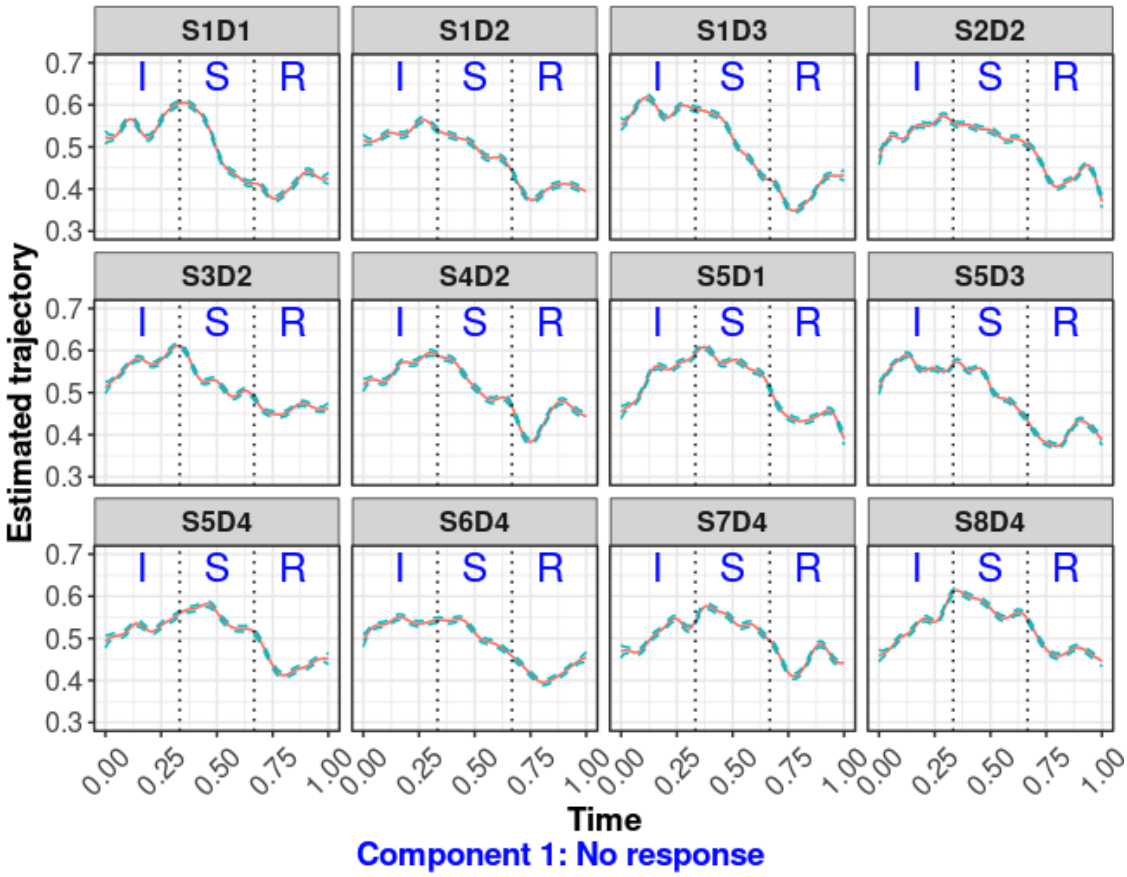

图9：包含全部十二个通道的三成分模型中第一成分的估计轨迹。**I**：互动 **S**：静止脸 **R**：恢复。红色曲线表示后验均值，两条绿色虚线表示95%逐点可信区间。

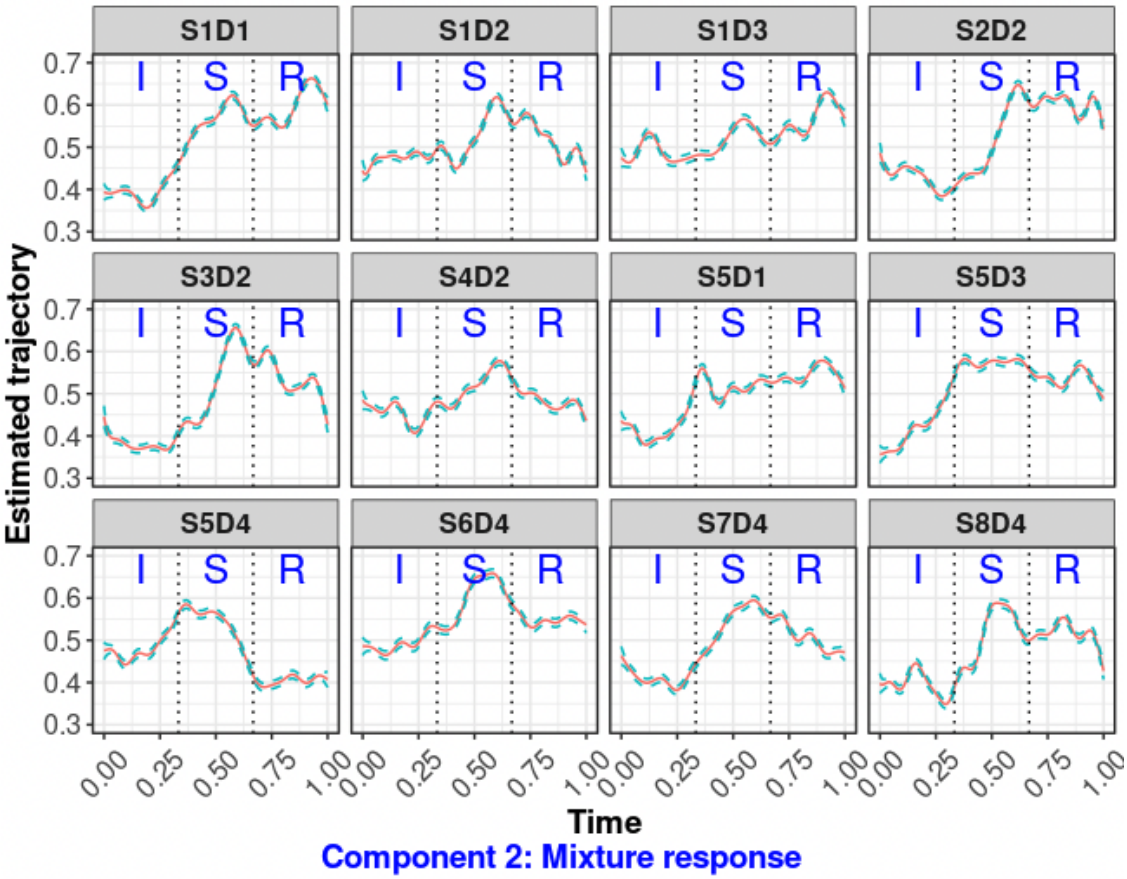

图10：包含全部十二个通道的三成分模型中第二成分的估计轨迹。**I**：互动 **S**：静止脸 **R**：恢复。红色曲线表示后验均值，两条绿色虚线表示95%逐点可信区间。

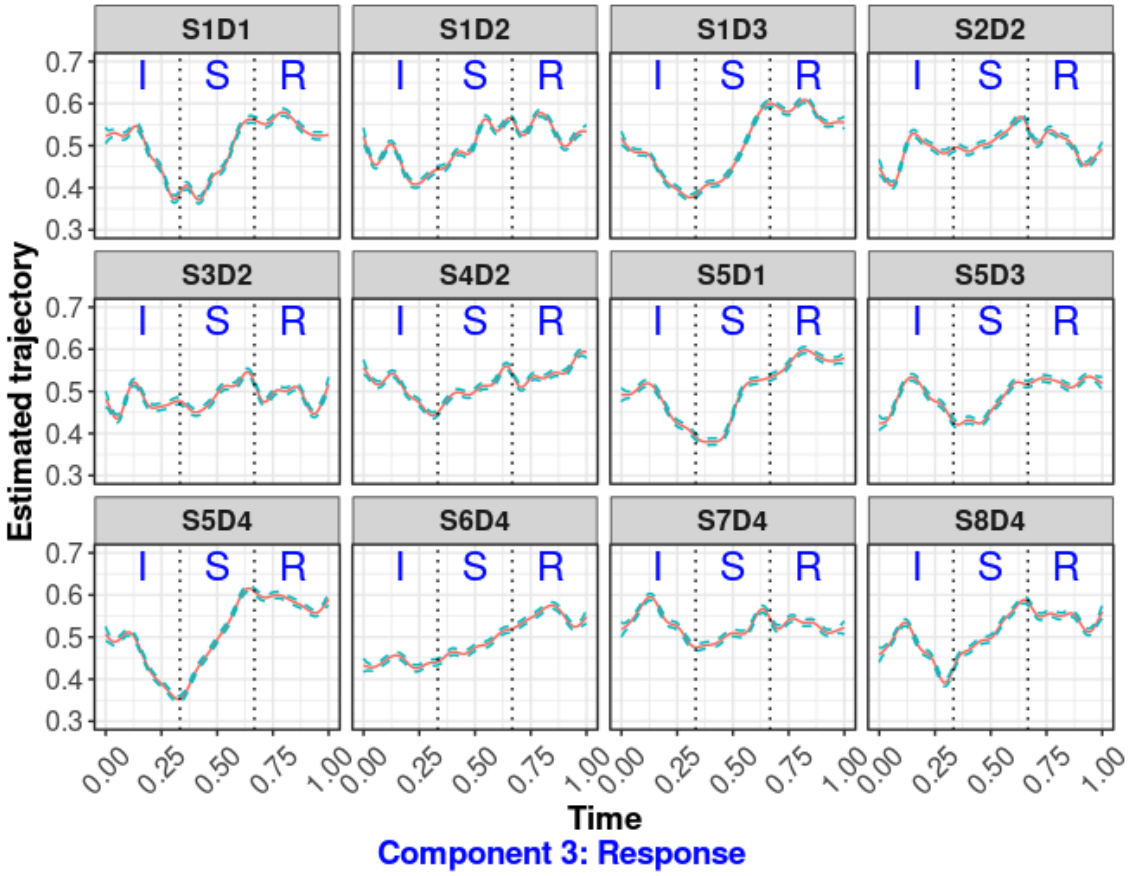

图11：包含全部十二个通道的三成分模型中第三成分的估计轨迹。**I**：互动 **S**：静止脸 **R**：恢复。红色曲线表示后验均值，两条绿色虚线表示95%逐点可信区间。

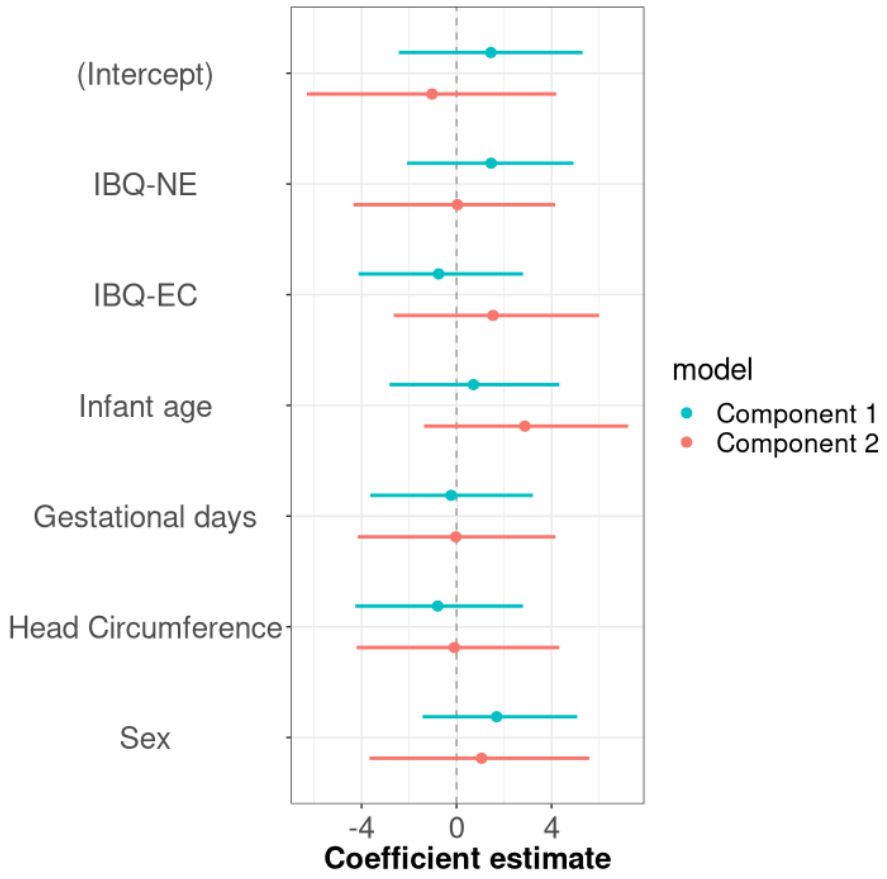

图12：所有十二个通道的三分量模型中各协变量的逻辑回归系数估计值及95% 可信区间。