Статистика и вычисления
https://doi.org/10.1007/s11222-018-9818-2

Об оценке параметров дисперсии нестандартных обобщенных линейные смешанные модели: применение к штрафному сглаживанию

Мария Хосе Родригес-Альварес [1,2] · Мария Дурбан [3] · Дэ-Джин Ли [1] · Пол ХК Эйлерс [4]

Получено: 16 марта 2018 г. / Принято: 5 июня 2018 г.

Абстрактный
Мы представляем новый метод оценки параметров дисперсии в обобщенных линейных смешанных моделях. Метод имеет корни в Харвилле (J Am Stat Assoc 72 (358): 320–338, 1977), но он может работать с моделями, имеющими точность матрица для вектора случайного эффекта, линейная по отношению к параметрам дисперсии (т. е. параметрам точности). Мы называем метод SOP (разделение перекрывающихся матриц точности). СОП основана на применении метода последовательных приближения к легко вычисляемым обновлениям оценок параметров дисперсии. Эти обновления оценок имеют привлекательный форма: они представляют собой отношение (взвешенной) суммы квадратов к количеству, относящемуся к эффективным степеням свободы. Мы предоставляем достаточные и необходимые условия строго положительности этих оценок. Нарушается важная область применения СОП оценка регрессии моделей, в которых несколько квадратичных штрафов действуют на одни и те же коэффициенты регрессии. Мы обсуждаем в подробно описывают две из этих моделей: сплайны со штрафом для локально адаптивной гладкости и для данных иерархической кривой. Несколько данных представлены примеры в этих настройках.

Ключевые слова Обобщенные линейные смешанные модели · Обобщенные аддитивные модели · Параметры дисперсии · Параметры сглаживания · REML · Эффективные степени свободы

1. Введение

Оценка параметров дисперсии является статистической проблемой. lem, которому уделялось пристальное внимание более 50 годы. Он возник на основе предложенной методологии ANOVA. Фишером в 1920-х гг., где были получены оценки приравнивание среднеквадратичной ошибки к ее ожидаемому значению. Кра- однако результаты, полученные этим методом, не были оптимальными. в некоторых ситуациях, например, при несбалансированном данные. Позже Крамп (1951) применено максимальное правдоподобие (ML) в предположении нормально распределенных ошибок и случайные эффекты. Но только в 1970-х годах, когда оценка параметров дисперсии на основе метода ML ods вызвали интерес. Метод *ограниченного максимума Вероятность* (REML) (Паттерсон и Томпсон1971) дал решение проблемы смещенных оценок переменной параметры. Однако одно из главных препятствий на пути использование этой техники в то время заключалось в том, что расчет оценок ML / REML требует численного решение нелинейной задачи. Паттерсон и Томпсон (1971) предложил итеративное решение с использованием оценки Фишера. алгоритм, но это был Харвилл(1977), предложивший первый численный алгоритм вычисления REML-оценок переменной параметры. Его предложение является источником вдохновения для нашей работы.

За прошедшие годы несколько вычислительных подходов появился с целью улучшения вычислительной бремя решения балльных уравнений для параметра дисперсии Этерс: Смит(1990) предложили использовать алгоритм EM, Graser et al.(1987) предложили использовать симплексный алгоритм. рифм для получения оценок непосредственно из правдоподобия, и Гилмор и др. (1995 г.) разработан метод, основанный на использовании средней информационной матрицы.

В контексте обобщенных линейных смешанных моделей (GLMM), оценка на основе итеративного повторного взвешивания REML был предложен независимо рядом авторов

B Мария Хосе Родригес-Альварес
mxrodriguez@bcamath.org

[1] BCAM - Баскский центр прикладной математики, Аламеда de Mazarredo, 14, 48009 Бильбао, Страна Басков, Испания

[2] ИКЕРБАСК, Баскский фонд науки, Бильбао, Испания

[3] Департамент статистики и эконометрики, Universidad Карлос III де Мадрид, Леганес, Испания

[4] Медицинский центр Университета Эразма, Роттердам, Нидерланды

Страница 2

(например, Schall 1991 г.; Энгель и Кин 1994), как продолжение итеративный алгоритм наименьших квадратов с повторным взвешиванием для обобщенных уточненные линейные модели (GLM, McCullagh и Nelder 1989 г.). Бреслоу и Клейтон(1993 г.) предложил общий метод на основе штрафной квази-правдоподобия (PQL) для оценки фиксированных и случайных эффектов, а также псевдовероятность параметры дисперсии. Как отмечают Энгель и Буист (1996), процедуры оценки, обсуждаемые во всех этих статьях, эквивалентны хороший, хотя и мотивированный с совершенно разных отправных точек.

Большинство упомянутых выше методов требуют сильное ограничение на вектор случайных эффектов: его Матрица дисперсии-ковариации должна быть линейной по дисперсии параметры. Результаты, представленные в этой статье, ослабляют то, что допущение к случаю, когда линейность по параметрам необходимо на матрице точности, а не на дисперсии - ковариационная матрица. Наш вклад мотивирован необходимостью для оценки параметров сглаживания в разрезе штрафных регрессионные модели с нестандартными квадратичными штрафами.

Штрафная сплайн-регрессия (P-сплайны, Эйлерса и Маркса 1996) стал популярным методом оценки мод. els, в котором средний ответ (или линейный предиктор в негауссовский случай) является гладкой неизвестной функцией одного или несколько ковариат. Метод основан на представлении ция гладкой компоненты в терминах базисных функций и оценка параметров путем изменения правдоподобия капюшон с квадратичным штрафом на коэффициенты. Размер штраф контролируется так называемым *параметром сглаживания. тер* . Связь между штрафным сглаживанием и линейным смешанные модели впервые были созданы очень давно (зеленый 1987), и он стал широко использоваться в последние 15 лет. (Карри и Дурбан 2002 г.; Currie et al. 2006 ; Ли 2010 г.; Палочка 2003 г.). Ключевым моментом эквивалентности является то, что сглаживание становится соотношением двух параметров дисперсии. терр. Следовательно, упомянутые выше методы могут быть использованы для непосредственной оценки количества сглаживания, необходимого в модели вместо использования методов, основанных на минимиза- некоторые ошибки предсказания, такие как информация Акаике критерий (AIC), обобщенная перекрестная проверка (GCV) или Mal- lows ' C_p (см., например, Эйлерс и Маркс 1996 ; Древесина 2008 г.). В Кривобокова (2009), асимптотические свойства REML и Рассмотрены оценки параметров сглаживания C_p Маллоуза. эред. В статье показано, что в рамках частотного подхода Параметр сглаживания REML оракула асимптотически суб- оптимально. Однако дисперсия оценки на основе REML оказывается меньше, чем оценка C_p Маллоуса. тор. Также Рейсс и Огден (2009 г.) показывают, что при конечной выборке размеры AIC и GCV склонны к недосглаживанию и более вероятно развитие множественных минимумов, чем REML (см. также Древесина 2017 г., Глава 6). Дополнительный аргумент в пользу REML заключается в том, что включение случайных эффектов в модель регрессии со штрафными санкциями проста.

В контексте сглаживания стандартные методы, основанные на REML может применяться, когда используются простые штрафы, т. Е. на каждый коэффициент регрессии действует один штраф (по одному параметру сглаживания). Однако в некоторых случаях В некоторых случаях штрафы представляют собой перекрывающуюся структуру, с одинаковыми коэффициентами одновременно штрафуются несколько параметров сглаживания. Сюда входят важные случаи такие как многомерные шлицы со штрафом с анизотропным штрафы или адаптивные штрафные сплайны. Методы оценки которые могут справиться с этой ситуацией, были предложены в литература по сглаживанию (например, Wood 2011 ; Wood et al. 2016 г.), но у них есть недостаток в том, что они очень вычислительны - союзник требователен, особенно когда количество сглаживания параметры большие.

Эта работа решает эту проблему и представляет собой быстрое метод оценки параметров дисперсии / сглаживания параметры в обобщенных линейных смешанных моделях / обобщенные аддитивные модели. Метод может использоваться всякий раз, когда предварительная решающая матрица случайной составляющей (или штрафная матрица модели P-сплайна) представляет собой линейную комбинацию, определенную над инверсия параметров дисперсии (параметр сглаживания - терс). Получены простые выражения для оценок параметры дисперсии, которые представляют собой отношения между суммой квадратов и величина, связанная с понятием эффективных степеней свобода в контексте сглаживания (Хасти и Тибширани 1990). Покажем достаточные и необходимые условия того, что гарантировать положительность этих оценок и обсудить несколько Бывают ситуации, когда эти условия могут быть легко проверены. Были введены частные случаи представленного здесь метода. согласно Родригес-Альваресу и др. (2015b), которые решили задача в случае анизотропной многомерной гладкой ing, и в работе Rodríguez-Álvarez et al.(2015a), где результаты для адаптивных P-шлицев. В последнее время, Дерево и Фазиоло(2017) расширил вышеупомянутый работает с более общими моделями сплайнов со штрафом. Про- Обсуждаемый здесь posal имеет два основных преимущества с уважение к Вуду и Фазиоло(2017 г.) подход. Во-первых, оценки параметров сглаживания / дисперсии, описанные в Wood и Фазиоло (2017) полагаются на псевдообратные формулы Мура – Пенроуза матрицы штрафов, которые, по нашему опыту, могут представлять числовые нестабильности. Во-вторых, наше предложение устанавливает явная связь между оценками компонентов дисперсии и эффективные степени свободы, которых не хватает в Wood и Фазиоло(2017 г.). Эффективные степени свободы являются ключевыми компоненты в сглаживающих моделях. Они помогают подвести итоги модели, поскольку частичные эффективные степени свободы являются мерой сложность компонентов модели с сильной интуитивной привлекательностью (см., например, Rodríguez-Álvarez et al. 2018 для примера в сельскохозяйственное поле).

Остальная часть этого документа организована следующим образом: Разд. 2 представляет работу Харвилла(1977), что составляет фундамент представленной здесь работы. Раздел 3 - это Суть статьи: новый метод, получивший название СОП (разделение перекрывающиеся матрицы точности). и кон- связь между СОП и понятием эффективных степеней

Стр. 3

свобода обсуждается. Раздел 4 описывает несколько P-шлицев. модели, к оценке которых можно подойти с помощью СОП. Мы В этой статье основное внимание уделяется адаптивным P-сплайнам и P-сплайнам для данные иерархической кривой. Иллюстрации с примерами данных представлены предоставлено в разд. 5. Обсуждение закрывает статью. Некоторые технические Технические подробности добавлены в виде приложений. Оценка алгоритм подробно описан там.

$$\underbrace{\begin{bmatrix} X^{\top} R^{-1} X & X^{\top} R^{-1} Z \\ Z^{\top} R^{-1} X & Z^{\top} R^{-1} Z + G^{-1} \end{bmatrix}}_{C} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} \text{ знак равно } \begin{bmatrix} X^{\top} R^{-1} z \\ Z^{\top} R^{-1} z \end{bmatrix}, \qquad (2)$$

которые приводят к выражениям в замкнутой форме

$$\hat{\beta} = \left(X^{\top} V^{-1} X \right)^{-1} X^{\top} V^{-1} z,$$

2 Оценка параметров дисперсии в обобщенные линейные смешанные модели: Harville (1977 г.) и расширения

Этот раздел - наша небольшая дань уважения Харвиллу.(1977) статья, что послужило вдохновением для этой работы. Харвилл(1977 г.) В статье рассматриваются подходы ML / REML к параметрам дисперсии. оценки в линейных смешанных моделях (LMM) для гауссовских данные. Тем не менее, оценка GLMM может быть сделана многократное использование методологии LMM на *рабочих-* зависимых переменная (см., например, Schall 1991 ; Энгель и Кин1994 где используются результаты Харвилла 1977 г.). Это подходу, которому мы следуем в этой статье.

Пусть $\boldsymbol{y} = (y_1, \ldots, y_n)$ вектор из n наблюдений. А GLMM можно записать как

$$g(\boldsymbol{\mu}) = \boldsymbol{X\beta} + \boldsymbol{Z\alpha}, \text{ где } \boldsymbol{\alpha} \sim N(\boldsymbol{0}, \boldsymbol{G}), \quad (1)$$

где $\mu_i = \mathrm{E}(y_i \mid \boldsymbol{\alpha})$ и $g(\cdot)$ - функция связи. В модель предполагает, что в зависимости от случайных эффектов наблюдения y_i независимо распределены со средним μ_i и дисперсия Yar $(y_i \mid \boldsymbol{\alpha}) = \varphi v(\mu_i)$. Здесь $v(\cdot)$ - специфическая заданной дисперсионной функцией, а φ - дисперсионный параметр, может быть известным или неизвестным. В модели (1), $\boldsymbol{X}$ и $\boldsymbol{Z}$ представляют разделенные по столбцам матрицы, связанные, соответственно, с фиксированные и случайные эффекты. Мы предполагаем, что $\boldsymbol{X}$ имеет полный ранг, $\boldsymbol{Z} = [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_c]$ и $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_c)$. Каждый $\boldsymbol{Z}_k$ соответствует матрице плана k -й случайной композиции nent $\boldsymbol{\alpha}_k$, где $\boldsymbol{\alpha}_k$ - вектор $(q_k \times 1)$ $(k = 1, \ldots, c)$. Мы далее предположим, что $\boldsymbol{\alpha}_k \sim N(\boldsymbol{0}, \boldsymbol{G}_k)$ и что

$$\boldsymbol{G} = \bigoplus_{k=1} \boldsymbol{G}_k = \bigoplus_{k=1} \sigma^2_k \boldsymbol{I}_{q_k} = \text{диаг}\left(\sigma^2_1 \boldsymbol{I}_{q_1}, \ldots, \sigma^2_c \boldsymbol{I}_{q_c}\right),$$

где $\boldsymbol{I}_m$ - единичная матрица порядка $m \times m$ и $\bigoplus$ обозначает прямая сумма матриц. Обратите внимание, что дисперсия-ковариация матрица $\boldsymbol{G}$ линейна по параметрам дисперсии σ^2_k.

Как отмечалось ранее, оценка модели (1) можно подойти путем итеративного подбора LMM, включающего *рабочую* зависимость переменная z и весовая матрица $\boldsymbol{W}$ (обновляются итерация). Конкретный вид z и $\boldsymbol{W}$ приведен в «Приложении. С ». Если φ и σ^2_k $(k = 1, \ldots, c)$ известны, на каждой итерации обновления для $\boldsymbol{\beta}$ и $\boldsymbol{\alpha}$ следуют из так называемого метода Хендерсона. уравнения (Хендерсон 1963 г.)

$$\boldsymbol{A}_k = \boldsymbol{G}_k \boldsymbol{Z}_k \boldsymbol{Pz} \qquad (k = 1, \ldots, c), \quad (3)$$

где $\boldsymbol{P} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1} \boldsymbol{Икс}\left(\boldsymbol{XV}^{-1}\boldsymbol{Икс}\right)^{-1} \boldsymbol{XV}^{-1}$ с участием $\boldsymbol{V} = \boldsymbol{R} + \boldsymbol{ZGZ}$ и $\boldsymbol{R} = \varphi \boldsymbol{W}^{-1}$. Уравнение Хендерсона- малопригодны, если параметры дисперсии φ и σ^2_k $(k = 1, \ldots, c)$ неизвестны. В своей статье 1977 года Харвилл показывает, как оценить их с помощью REML элегантным итеративным Живой алгоритм. Давайте сначала определим

$$\boldsymbol{T} = \left(\boldsymbol{И} + \boldsymbol{3\ СЗГ}\right)^{-1},$$

где $\boldsymbol{S} = \boldsymbol{R}^{-1} - \boldsymbol{R}^{-1} \boldsymbol{ИксXR}\left(^{-1} \boldsymbol{Икс}\right)^{-1} \boldsymbol{XR}^{-1}$. Мы заметили что $\boldsymbol{T}$ можно разбить следующим образом

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} & \ldots & \boldsymbol{T}_{1в} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} & \ldots & \boldsymbol{T}_{2в} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{T}_{c1} & \boldsymbol{T}_{c2} & \cdots & \boldsymbol{T}_{cc} \end{bmatrix},$$

где $\boldsymbol{T}_{ij}$ - матрицы порядка $q_i \times q_j$. В Харвилле(1977 г.), обновленная оценка σ^2_k $(k = 1, \ldots, c)$ равно

$$\sigma^2_k = \frac{\boldsymbol{A}^{[m]}_k \boldsymbol{A}^{[m]}_k}{\mathrm{ED}^{[m]}_k}, \quad (4)$$

где

$$\mathrm{ED}^{[m]}_k = q_k - \text{след}\left(\boldsymbol{T}^{[m]}_{кк}\right), \quad (5)$$

а верхний индекс $[t]$ обозначает величины, оцененные в текущем оценки параметров дисперсии. Из оценок $\boldsymbol{\beta}$ и $\boldsymbol{\alpha}$ следуют оценку для z : $z = X\beta + Z\ \&\ alpha$; . Остатки являются z - z . Харвилл использует

$$\varphi = \frac{z\ \boldsymbol{W}\left(z - z^{[m]}\right)}{n - \text{ранг}\ (\boldsymbol{X})}, \quad (6)$$

для оценки параметра дисперсии (не всегда требуется в GLMM). Альтернативное выражение (см., Например, Engel1990 ; Родригес-Альварес и др. 2015b)

$$\varphi = \frac{\left(z - z^{[m]}\right) \boldsymbol{W} \left(z - z^{[m]}\right)}{n - \text{ранг}\ (\boldsymbol{X}) - \sum_{k=1}^{c} \mathrm{ED}^{[m]}_k} \quad (7)$$

Здесь rank $(\boldsymbol{X}) + \sum_{k=1}^{c} \mathrm{ED}^{[m]}_k$ можно интерпретировать как эффект Размерность модели. При сходимости уравнения (6) и (7) дайте одинаковые числовые значения.

2.1 Эффективные степени свободы в системе Харвилла метод

Как отмечает Харвилл (1977 г.), итерации, полученные из выражения сион (4) имеют интуитивно привлекательную форму. На каждой итерации σ^2_k оценивается как отношение суммы квадратов оценки для $\boldsymbol{\alpha}_k$ и числа от нуля до q_k . Мы Теперь покажем, что знаменатель в выражении (4) действительно может интерпретироваться как эффективные степени свободы при сглаживании sensu, т. е. как след «шляпной» матрицы (Hastie и Tibshi- рани 1990).

Сначала обратите внимание, что выражение (3) Показывает , что $\boldsymbol{Z}_k \boldsymbol{\alpha}_k$ знак равно $\boldsymbol{Z}_k \boldsymbol{G}_k \boldsymbol{Z}_k \boldsymbol{Pz}$. Таким образом, матрица «шляпа», соответствующая k -я случайная компонента $\boldsymbol{\alpha}_k$ равна

структуры дисперсионно-ковариационной матрицы появляются на практике. tice. Настоящее исследование было мотивировано нашей работой над штрафная сплайн-регрессия. Несмотря на это, метод, который нужно обсуждаемые в этом разделе, не ограничиваются этой областью: это может можно рассматривать как общий метод оценки обобщенных линейных смешанные модели с матрицей точности определенной структуры.

Как и в разд. 2, мы рассматриваем обобщенную линейную смешанную модель

$$g(\boldsymbol{\mu}) = \boldsymbol{X\beta} + \boldsymbol{Z\alpha} = \boldsymbol{X\beta} + \sum_{k=1}^{c} \boldsymbol{Z}_k \boldsymbol{\alpha}_k, \quad (10)$$

где $\boldsymbol{\alpha}_k \sim N(\boldsymbol{0}, \boldsymbol{G}_k)$, $\boldsymbol{\alpha} \sim N(\boldsymbol{0}, \boldsymbol{G})$ и $\boldsymbol{G} = \bigoplus_{k=1}^{c} \boldsymbol{G}_k$.

Основное отличие от разд. 2 это то, что мы не Предположим, что $\boldsymbol{G}_k = \boldsymbol{d}_k$, но мы рассматриваем матрицы точности формы

$$грамм^{-1}_k \text{ знак равно } \sum_{l=1}^{p_k} \sigma^{-2}_{kl} \ldots_{кл}, \quad (11)$$

$H_k = Z_k G_k Z_k P$,

т.е., $H_k z = Z_k \alpha_k$. Теперь покажем, что trace (H_k) = ED $_k$. Это легко проверить, что

$$T = \left(I + Z_t SZG\right)^{-1} = G^{-1}\left(G^{-1} + Z_t SZ\right)^{-1}, \quad (8)$$

где $\left(G^{-1} + ZSZ_t\right)^{-1}$ это разбиение, обратное к C в (2), соответствующий случайному вектору α (Harville 1977 г.; Джонсон и Томпсон 1995). Использование блочной структуры из Z и G, и с использованием результата (8) и (A4) в Джонсоне и Томпсон (1995), мы имеем

$$Z_k PZ_k G_k = I_{q_k} - G_k^{-1} C^*_{kk} = I_{q_k} - T_{kk}, \quad (9)$$

где для облегчения обозначений C^* обозначает обратное к C и C^*_{kk} обозначает это разбиение C^* соответствующий k -й случайная составляющая α_k. Таким образом,

$$\begin{aligned} \text{след}\,(H_k) &= \text{след}\left(Z_k G_k Z_k P\right) = \text{след}\left(Z_k PZ_k G_k\right) \\ &= \text{след}\left(I_{q_k} - T_{kk}\right) = q_k - \text{след}\,(T_{kk}) \\ &= \text{ED}_k. \end{aligned}$$

3 Разделение точности перекрытия матрицы: метод СОП

В предыдущем разделе мы обсудили оценку метод для обобщенных линейных смешанных моделей, где дисперсионно-ковариационная матрица случайной составляющей: линейно по параметрам дисперсии. Однако более сложные

где $\sigma^2_{k_l}$ ($l = 1, ..., p_k$ и $k = 1, ..., c$) - дисперсия параметры и $_{k_l}$ известны симметричные положительные полу- определенные матрицы размерности $q_k \times q_k$. Обратите внимание, что мы не требуют, чтобы $_{k_l}$ было положительно определенным. Единственное требование мы нужно, чтобы G_k^{-1} $(k = 1, ..., c)$ положительно определены, поэтому являются G^{-1} обратное, то ковариационная матрица G.

Выражение (11) заслуживает подробного обсуждения. Во-первых, стоит отметить, что мы не работаем с дисперсией - ковариационные матрицы, но с их обратными, точность матрицы. Как уже было сказано, разработки в этой работе имеют свои origin по штрафным сплайн-методам. В разд. 4, необходимость работать с прецизионными матрицами станет ясно, или, в терминология штрафных шлицев, со штрафной матрицей ces. Во-вторых, в чем состоит основной вклад этого В статье предполагается, что каждая случайная компонента α_k ($k = 1, ..., c$) в модели (10) могут быть "затронуты" (сжаты) из-за несколько параметров дисперсии. Частным случаем будет, когда $p_k = 1\ \forall\ k$, и в этом случае мы находимся в обсуждаемой ситуации в разд. 2.

Для простоты в некоторых случаях перепишем матрица точности G^{-1} следующим образом

$$G^{-1} = \sum_{l=1}^{p} \sigma_l^{-2} \tilde{\Lambda}_l, \quad (12)$$

где $p = \sum_{k=1}^{c} p_k$. Слегка злоупотребляя обозначениями, пусть $_l$ обозначим матрицы, входящие в выражение (11). Матрица $\tilde{\Lambda}_l$ является $_l$ проложенной с нулями. Некоторые конкретные примеры будут быть представлены в разд. 4 ниже. Выражение (12) проясняет что в настоящей работе рассматривается ситуация обобщенного линейные смешанные модели с матрицей точности для случайных составляющая, линейная по параметрам точности σ_l^{-2}.

В следующем разделе представлен предлагаемый метод оценки. что мы называем СОП.

3.1 Метод СОП

Независимо от структуры G, оценка β и α, и, при необходимости параметр дисперсии φ не задает проблема, и это можно сделать, как описано в Разд. 2 выше (см. также «Приложение C» для подробного описания оценки алгоритм). Напомним, что оценки для α_k получаются aŝ $\alpha_k = G_k Z_k Pz$. Матрица шляпы, связанная с k -м случайным Компонента α_k снова равна $H_k = Z_k G_k Z_k P$ и эффективные степени свободы этой компоненты ED $_k$ = след (H_k). Параметры дисперсии $\sigma^2_{k_l}$ ($l = 1, ..., p_k$ и $k = 1, ..., c$), однако нельзя оценить с помощью Подход Харвилла. Это следствие того, что G не линейно по параметрам дисперсии.

Ключ нашего подхода - работа с G^{-1} вместо с G. Учитывая, что G^{-1} линейна по параметрам точности $\sigma^{-2}_{k_l}$, частные производные первого порядка от (приближенной) Функция логарифмического правдоподобия REML может быть явно получена как а также основанные на REML оценки параметра дисперсии эфиры. Сформулируем результат в следующей теореме, в которой доказательство приведено в «Приложении A».

Теорема 1 *Пусть $G = \bigoplus_{k=1}^{c} G_k$ - симметричная положительная деф- бесконечная матрица, с G_k^{-1} знак равно $\sum_{l=1} \sigma^{-2}_{k_l}$ $_{k_l}$ симметричный положительный определенный и $_{k_l}$ известный симметричный положительный полуопределенный. Затем оценка дисперсии на основе REML обновляется. параметры $\sigma^2_{k_l}$ $(l = 1, ..., p_k$ и $k = 1, ..., c)$ даны*

Теорема 2 *Если $\sigma^{2[m]}_{k_l} > 0$, то оценка на основе REML обновления параметров дисперсии, указанных в выражении (13) больше или равны нулю со строгим неравенством, если: (i) ранг $\left(X, Z_k G_k^{[m]} {}_{k_l}\right)$ > ранг (X); и (ii) z (рабочий вектор ответа) не находится в пространстве, занимаемом столбцами из X.*

Доказательства приведены в «Приложении B». Отметим, что con- определение (i) необходимо как для числителя, так и для знаменателя выражения (13) быть строго положительным, а (ii) - только необходимо для числителя. С прикладной точки зрения, несомненно, важно иметь возможность проверить, условия выполняются перед подгонкой модели. Это может непростая задача, так как они зависят от $G_k^{[m]}$ и, таким образом, может варьируются от итерации к итерации. Однако есть общие ситуации, когда условие (i) можно проверить заранее:

- Если $_{k_l}$ имеет полный ранг, то условие (i) упрощается до ранг (X, Z_k)> ранг (X). Отметим, что это условие то же, что обсуждал Харвилл(1977 г.) в лемме 1.
- Если $G_k^{[m]}$ и $_{k_l}$ коммутируют (т. е. $G_k^{[m]}$ $_{k_l}$ = $_{k_l}$ $G_k^{[m]}$), то условие (i) упрощается до ранжирования $\left(X, Z_k {}_{k_l}\right)$ > ранг (X). Примеры, когда $G_k^{[m]}$ а также $_{k_l}$ коммутируют, включают, для Например, когда оба диагональны.

Мы обсуждаем эти ситуации более подробно в разд. 4, где некоторые примеры применения метода СОП предварительно отправлено.

от

$$\sigma^2_{kl} = \frac{A_k^{[m]} \alpha_{kl}^{[m]}}{ED_{kl}^{[m]}} \quad (13)$$

где

$$ED_{kl}^{[m]} = \text{след}\left(Z_k P^{[m]} Z_k G_k^{[m]} \frac{k_l}{\Sigma^{2[m]}_{kl}} \text{грамм} \right), \quad (14)$$

с & alpha ; $P^{[m]}$ и $G^{[m]}$ оценивается по текущим оценкам $\Sigma^{2[m]}_{kl}$ (l = 1 , ..., p_k и k = 1 , ..., c) и, при необходимости, $\Phi^{[m]}$.

Отметим, что когда $G_k = \sigma^2_k I_{q_k}$, выражения (13) и (14) сводятся к выражениям Харвилла [выражения (4) и (5), соответственно тивно].

Важное и желаемое свойство данных обновлений в выражении (13) состоит в том, что они всегда неотрицательны, про- при том, что предыдущие оценки параметров дисперсии неотрицательны. Кроме того, в достаточно слабых условиях эти обновления строго положительные (хотя возможно получить значения, очень близкие к нулю).

3.2 Эффективные степени свободы в методе СОП

В соответствии с методом Харвилла, обсуждаемым в разд. 2, то знаменатель выражения (13) был обозначен как ED $_{(\cdot)}$, от эффективных степеней свободы. Результат (14) упрощает чтобы показать, что сумма ED $_{kl}$ по параметру дисперсии p_k участники G^{-1}_k [см. (11)] соответствует ED $_k$ (эффективные степени свободы α_k)

$$\sum_{l=1}^{p_k} ED_{kl} = \sum_{l=1}^{p_k} \text{след}\left(Z_k PZ_k G_k \frac{k_l}{\sigma^2_{kl}} G_k \right) = \text{след}\left(Z_k PZ_k G_k \right) = \text{след}\,(H_k) = ED_k .$$

Как следствие, при сходимости расчетная эффективная степени свободы, связанные с каждой случайной композицией. в модели (10) получены как побочный продукт СОП метод.

В завершение этой части хотелось бы отметить интересный связь между верхней границей для ED $_{kl}$ (обозначается ED $^{ub}_{kl}$) и условию (i) теоремы 2. Можно показать, что

Стр. 6

$$ED_{kl} \leq \text{ранг}\left(X, Z_k G_{kk_l} \right) - \text{ранг}\,(X) = \text{ранг}\left((I_n - P_X) Z_k G_{kk_l} \right) = ED^{ub}_{k_l},$$

где $P_X = X (X'X)^{-1} X'$. Таким образом, если условие (i) в orem 2 не проверяется, ED $_{kl}$ будет в точности равным нулю. Мы опускаем доказательство предыдущего результата. Его можно получить в сим- Таким же образом, как в статье Cui et al.(2010 г.) (см. Интернет Приложение (f) к этому документу), отмечая, что

$$ED_{kl} = \text{след}\left(Z_k PZ_k G_k \frac{k_l}{\sigma^2_{kl}} G_k \right) = \text{след}\left(Z_k G_k \frac{k_l}{\sigma^2_{kl}} G_k Z_k [(I_n - P_X) V (I_n - P_X)]^{+} \right),$$

где $^{+}$ обозначает псевдообратную матрицу Мура – Пенроуза. Эта эквивалентность была доказана в менее общей ситуации: Родригес-Альварес и др.(2018 г.) (см. веб-приложение D в этой статье). Следуя аналогичным рассуждениям, получаем оценка сверху эффективных степеней свободы k -й случайный компонент

$$ED_k \leq \text{ранг}\,(X, Z_k) - \text{ранг}\,(X) = \text{ранг}\,((I_n - P_X) Z_k) = ED^{ub}_k .$$

Используя известный результат, что ранг матричной суммы не может превышать сумму рангов слагаемых матриц, у нас есть этот ED $^{ub}_k \leq \sum_{l=1}^{p_k}$ ED $^{ub}_{kl}$. Однако в целом это является строгим неравенством, так как в большинстве случаев

$$C\left(Z_k G_{kk_u} \right) \cap C\left(Z_k G_{kk_v} \right) = \{ \mathbf{0} \} \ (u = v),$$

где $C(A)$ обозначает линейное пространство, натянутое на столбцы группы A (см., например, теорему 18.5.7 в Харвилле1997 г.). Интуитивно мы можем интерпретировать это как своего рода соревнование между p_k «Элементы», связанные с k -м случайным компонентом. В ED $_{kl}$ не может «свободно» меняться от 0 до ED $^{ub}_{k_l}$, но у них есть для обеспечения того, чтобы их сумма не превышала ED $^{ub}_k$.

3.3 Вычислительные аспекты

(2)] Вычисляется для оценки β и *альфа* . Таким образом вычисление трасс, необходимых для реализации алгоритма Ритм сводится к вычислению диагональных элементов из $(G_k - C^{*}_{kk})_{\kappa_\lambda}$. Кроме того, в тех случаях, когда k_l равно диагонали, только поэлементное произведение диагоналей $(G_k - C^{*}_{kk})$ и нужен k_l. Это значительно снизит количество требуемых операций и, следовательно, вычисление время.

4 Штрафное сглаживание и метод SOP

В этом разделе обсуждается несколько ситуаций в P-сплайне. структура, в которой к оценке можно подойти с помощью Метод СОП. Как будет видно, этот метод можно использовать, когда: когда на один и тот же коэффициент действуют несколько штрафов. cients. Анизотропные P-сплайны тензорного произведения являются примером накладывающиеся штрафы, и это широко обсуждалось в статье Родригеса-Альвареса и др.(2015b). Тем не менее, множественные штрафы возникают в более широком классе ситуаций. Мы опишите здесь два из них: пространственно-адаптивные P-сплайны и P-сплайны для данных иерархической кривой.

4.1 Пространственно-адаптивные P-шлицы

Рассмотрим задачу обобщенной регрессии

$$g(\mu_i) = f(x_i), \qquad я = 1, \ldots n, \quad (15)$$

где $\mu_i = E(y_i)$, $g(\cdot)$ - функция связи, а $f(\cdot)$ - гладкая и неизвестная функция. Далее мы предполагаем, что Яр $(y_i) = \varphi v(\mu_i)$. В каркасе P-шлицев (Эйлерс и Маркса 1996) неизвестная функция $f(x)$ аппроксимируется линейной комбинацией d B-сплайнов базисных функций, т. е. $f(x) = \sum_{j=1}^{d} \theta_j B_j(x)$. В матричных обозначениях модель (15) является таким образом выражается как

$$g(\mu) = B\theta, \quad (16)$$

где $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$, $\theta = (\theta_1, \theta_2, \ldots, \theta_d)$ и B представляет собой B-сплайн-матрицу регрессии размерности $n \times d$, т. е. $b_{ij} =$

С вычислительной точки зрения оценка выражение, данное в (14) может быть очень дорогостоящим. Однако это вычисления можно упростить, используя результат, приведенный в (9). Для наших целей легко показать, что

$$\text{след}\left(Z_k P Z_k G_{kk l} G_k\right) = \text{след}\left(G_k Z_k P Z_k G_{kk l}\right) = \text{след}\left(\left(G_k - C^*_{\kappa\kappa}\right)_{k l}\right).$$

Этот результат используется в описанном алгоритме оценки. в «Приложении С». Отметим, что C^* [т.е. обратное C в

$B_j(x_i)$ - j-й B-сплайн, вычисленный в x_i. Гладкость достигается за счет наложения штрафа на коэффициенты регрессии $\boldsymbol{\theta}$ в виде

$$\lambda \sum_{\kappa=\partial+1}^{d} \left(\Delta^q \theta_k\right)^2 = \lambda\, \boldsymbol{\theta} \boldsymbol{D}_q \boldsymbol{D}_q \boldsymbol{\theta}, \tag{17}$$

где λ - параметр сглаживания, а Δ^q образует дифф-различий порядка q на соседние коэффициенты, т. е. $\Delta\theta_k = \theta_k - \theta_{k-1}$, $\Delta^2\theta_k = \Delta(\Delta\theta_k) = \theta_k - \theta_{k-1} - (\theta_{k-1} - \theta_{k-2}) = \theta_k - 2\theta_{k-1} + \theta_{k-2}$, и так далее для более высоких q. Наконец, $\boldsymbol{D}_q$ есть просто матричное представление Δ^q.

123

Стр.7

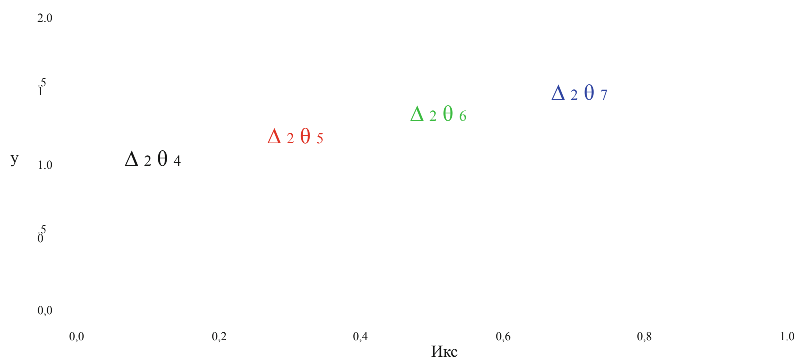


Рис. 1 Графическое представление разностей 2-го порядка на соседние коэффициенты базисных функций кубических B-сплайнов. Обратите внимание на местные и заказанные характер этих различий

Как можно увидеть в уравнении. (17), то же сглаживание параметр λ применяется ко всем разностям коэффициентов, независимо от TIVE от их местоположения (см. 1). Таким образом, модель предполагает что необходимо одинаковое сглаживание вся область ковариаты. Адаптивные Р-шлицы (см., Например, Кривобокова и соавт. 2008 ; Рупперт и Кэрролл2000 среди другие) ослабляют это предположение. Идея проста, заменить глобальный параметр сглаживания параметрами сглаживания которые варьируются локально в зависимости от значения ковариаты. Это может быть достигается путем указания другого параметра сглаживания - ter для каждой разницы коэффициентов (Рупперт и Кэрролл 2000 г.; Древесина 2011)

$$\sum_{\kappa=\partial+1}^{d} \lambda_{k-q} \left(\Delta^q \theta_k\right)^2 = \boldsymbol{\theta} \boldsymbol{D}_q \operatorname{diag}(\boldsymbol{\lambda}) \boldsymbol{D}_q \boldsymbol{\theta}, \tag{18}$$

где $\boldsymbol{\lambda} = \left(\lambda_1, \ldots, \lambda_{d-q}\right)$. Обратите внимание, что этот подход подразумевают столько параметров сглаживания, сколько различаются коэффициенты. (т. е. $d - q$), что может привести к недостаточному сглаживанию и нестабильные вычисления. Учитывая местный и упорядоченный характер разностей коэффициентов (см. 1), мы можем смоделировать параметры сглаживания λ_k как гладкая функция от k (его положения ции) и используйте для этой цели B-сплайны (здесь нет штрафа. предполагается)

$$\boldsymbol{\lambda} = \Psi\boldsymbol{\xi}, \tag{19}$$

где Ψ матрица регрессии B-сплайна размерности $(d - q) \times p$ с $p < (d - q)$ и $\boldsymbol{\xi} = \left(\xi_1, \ldots, \xi_p\right)$ это новый вектор параметров сглаживания. Выполняя некоторые сим-простых алгебраических операций, можно показать, что *адаптивная*

$$\boldsymbol{\theta}\left(\sum_{l=1}^{p} \xi_l \boldsymbol{D}_q \operatorname{diag}\left(\boldsymbol{\psi}_l\right) \boldsymbol{D}_q\right)\boldsymbol{\theta}, \tag{20}$$

где $\boldsymbol{\psi}_l$ обозначает столбец l *таблицы* Ψ. Обратите внимание, что под этим адаптивный штраф, все коэффициенты штрафуются множеством параметры сглаживания, т. е. есть накладывающиеся штрафы.

4.1.1 Повторная параметризация смешанной модели

Оценка Р-сплайн-модели (16) при условии адаптации размер штрафа, определенный в (20) может осуществляться на основе связь между Р-шлицами и смешанными моделями (например, Currie и Дурбан 2002 г.; Жезл 2003). Легко показать, что нуль пространство (т.е. непенализованное функциональное пространство) адаптивного матрица штрафов $\boldsymbol{P}_{Ad} = \sum_{l=1}^{p} \xi_l \boldsymbol{D}_q \operatorname{diag}\left(\boldsymbol{\psi}_l\right) \boldsymbol{D}_q$ не зависит от вмятина $\boldsymbol{\xi}$. Кроме того, заметим, что при $\xi_u = \xi_v = \lambda\ (\forall\ u, v)$ тогда

$$\sum_{l=1}^{p} \xi_l \boldsymbol{D}_q \operatorname{diag}\left(\boldsymbol{\psi}_l\right) \boldsymbol{D}_q = \lambda \boldsymbol{D}_q \text{ диагональ}\left(\sum_{l=1}^{p} \boldsymbol{\psi}_l\right) \boldsymbol{D}_q = \lambda\, Д_\partial Д_\partial = \boldsymbol{P}.$$

Это следствие строк регрессии B-сплайна. матрица с суммой до 1. Таким образом, нулевое пространство $\boldsymbol{P}_{Ad}$ является же , как и $\boldsymbol{P}$. Различные изменения параметров Р-шлица модели предлагались в литературе (см., например, Currie и Дурбан 2002 ; Эйлерс1999), все стремятся разложить модель на штрафную и штрафную части. В Следствием этого разложения является то, что штрафная матрица репараметризованной модели Р-сплайна имеет полный ранг, как и матрица точности соответствующей смешанной модели. Для В нашем приложении мы используем предложение, данное в Eilers (1999). Как будет видно, этот подход обеспечивает диагональную точность.

штраф (18) является

матрицы. Как обсуждалось в разд. 3.3, это очень удобно

с вычислительной точки зрения. Использование Eilers(1999 г.) трансформация, модель (16) переформулируется как

$g(\mu) = B\theta = X\beta + Z\alpha,$

где $X = [1_n | x | \ldots | Икс^{(q-1)}]$ и $Z = BD_q (D_q D_q)^{-1}$,

и матрица точности (штрафа) вектора случайных (штрафные) коэффициенты α становятся

$$грамм^{-1} = \frac{1}{\varphi} FP_{Ad} F = \sum_{l=1}^{n} \sigma_л^{-2} \text{ диагональ } (\psi_l) \text{ знак равно } \sum_{l=1}^{n} \sigma_л^{-2} \breve{}_l, \quad (21)$$

где $F = D_q (D_q D_q)^{-1}$, $\breve{}_l = \mathrm{diag}(\psi_l)$, а $\sigma^2_l = \varphi / \xi_l$. Таким образом, матрица точности линейна по точности параметры $\sigma_л^{-2}$, и поэтому можно использовать метод СОП.

Отметим, что модель имеет единственную случайную составляющую ($c = 1$). Повторная параметризация гарантирует, что rank (X, Z) = ранг (X) + ранг (Z). Таким образом, rank $(X, Z\breve{}_l)$ = rank (X) + ранг $(Z\breve{}_l)$> ранг (X). Используя тот факт, что G и $\breve{}_l$ диагональны, а значит, коммутируют, условие (i) теоремы 2 доволен.

4.2 P-сплайны для данных иерархической кривой

Для простоты предположим, что данные сбалансированной иерархической кривой. Наши данные состоят на $м$ особей каждый с S - разному меа- измерения в моменты времени $t = (t_1, t_2, \ldots, t_s)$. Наш интерес сосредоточен на

$$грамм(\mu_{ij}) = f(t_i) + g_j(t_i), \quad 1 \le i \le s, \; 1 \le j \le m, \quad (22)$$

где $\mu_{ij} \equiv \mathrm{E}(y_{ij})$, y_{ij} - переменная ответа на j -й субъект в момент времени t_i, $f(\cdot)$ - функция, описывающая эффект популяции, а $g_j(\cdot)$ - случайные функции, измеряемые отклонение j-го субъекта от популяции эффект. Как и прежде, Яр $(y_{ij}) \equiv \varphi v(\mu_{ij})$. Простая модель будет заключаться в параметрической спецификации для $f(\cdot)$ и $g_j(\cdot)$, например, $f(t) = \beta_0 + \beta_1 t$ и $g_i(t) = \alpha_{0j} + \alpha_{1j} t$, где $\alpha_0 = (\alpha_{01}, \ldots, \alpha_{0m}) \sim N(0, \sigma_0^2 Я_м)$ и $\alpha_1 = (\alpha_{11}, \ldots, \alpha_{1m}) \sim N(0, \sigma_1^2 Я_м)$.

Более гибкий подход состоит в предположении, что $f(\cdot)$ и $g_j(\cdot)$ гладкая и неизвестная. Важный вклад в Каркас P-сплайна можно найти у Durban et al.(2005 г.) и Ruppert et al.(2003 г.). Оба подхода основаны на моделирование $f(\cdot)$ и $g_j(\cdot)$ с использованием базиса усеченных линий, и она построение основано на штрафных методах и линейной смешанной модели техники. Совсем недавно Джендже и Карри(2010 г.) расширить эти модели и предложить включение дополнительных штраф за отдельные коэффициенты кривой. Авторы утверждают, что это дополнительное наказание необходимо для решения идентифицирующих- при оценке эффекта популяции. Под

так называемая модель M0 в Djeundje и Currie(2010 г.) бумага, модель (22) выражается в матричных обозначениях как

$$грамм(\mu_{\cdot j}) = \underbrace{B\theta}_{f(t)} + \underbrace{\breve{B}\breve{\theta}_j}_{g_j(t)}$$

где $\mu_{\cdot j} = (\mu_{1j}, \ldots, \mu_{sj})$ это средний вектор отклика для j -го индивидуума, а B и $\breve{B}$ являются B-сплайновой регрессией матрицы возможно разного размера для соответственно поп- кривая уляции и индивидуальное отклонение. Вектор θ равен предполагается фиксированным, но с учетом штрафа q-го порядка вида $P = \lambda_1 D_q D_q$, а $\breve{\theta}_j$ - случайный вектор с распределением $N(0, \varphi \breve{P}^{-1})$ где

$$\breve{P} = \lambda_2 \breve{D}_q \breve{D}_q + \lambda_3 \breve{I}_d. \quad (23)$$

Первый член $\lambda_2 \breve{D}_q \breve{D}_q$ отвечает за гладкость кривых индивидов, а $\lambda_3 \breve{I}_d$ обращается к аспект гибкости (см. Djeundje и Currie 2010 для большего подробности). Обратите внимание, что каждый случайный эффект уменьшается (штрафуется) на оба параметра сглаживания λ_2 и λ_3, и, следовательно, точность матрица $\breve{G}^{-1} = 1/\varphi \breve{P}$ линейно по параметрам точности

$$\breve{G}^{-1} \text{ знак равно } \sum_{l=2}^{3} \sigma_л^{-2} \breve{}_л, \quad (24)$$

где $\sigma_2^2 = \varphi/\lambda_2$ и $\sigma_3^2 = \varphi/\lambda_3$ и $\breve{}_2 = \breve{D}_q \breve{D}_q$ и $\breve{}_3 = \breve{I}_d$.

Более компактно выразим модель для всего образец как

$$г(\mu) = [1_м \otimes B] \& thetas; + [я_м \otimes \breve{B}] \breve{} \& thetas;, \quad (25)$$

где $\otimes$ обозначает произведение Кронекера, μ знак равно $(\mu_{\cdot 1}, \ldots, \mu_{\cdot м})$, $\breve{\theta} = (\breve{\theta}_1, \ldots, \breve{\theta}_м)$ а также

$$\breve{\theta} \sim N(0, я_м \otimes \breve{G}).$$

Последний шаг необходим для применения метода СОП. для оценки модели (25): разложение Эффект популяции в штрафной и штрафной частях. Мы используем здесь подход, основанный на разложении на собственные значения. положение (EVD) штрафа. Пусть $D_q D_q = UU$ - EVD $D_q D_q$. Здесь U обозначает матрицу собственных векторов и также диагональная матрица собственных значений. Обозначим также через $U_{+(+)}$ и $U_{0(0)}$ подматрицы, соответствующие ненулевое и нулевое собственные значения соответственно. В таком случае, модель (25) переформулируется как

$g(\mu) = X\beta + Z\alpha,$

где $\boldsymbol{\beta} = \boldsymbol{U}_+ \boldsymbol{\theta}$, $\boldsymbol{\alpha} = \left(\left(\boldsymbol{U}_{0\boldsymbol{\theta}} \right), \breve{\boldsymbol{\theta}} \right)$, $\boldsymbol{X} = [\, \mathbf{1}_м \otimes \boldsymbol{БУ}_0 \,]$

и $\boldsymbol{Z} = [\, \mathbf{1}_м \otimes \boldsymbol{БУ}_+ : \boldsymbol{Я}_м \otimes \breve{\boldsymbol{B}} \,]$. Наконец, $\boldsymbol{\alpha} \sim N(\mathbf{0}, \boldsymbol{G})$,

где

$$\boldsymbol{грамм}^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} + & \mathbf{0}_{d \times (м \breve{} \partial)} \\ \mathbf{0}_{(м \breve{} \partial) \times c} & \boldsymbol{Я}_м \otimes \breve{\boldsymbol{G}}^{-1} \end{pmatrix} = \sum_{l=1}^{3} \sigma_л^{-2} \breve{}_l ,$$

с $\sigma^2_{l} = \varphi \lambda_l$ а также

$$\breve{}_1 = \begin{pmatrix} + & \mathbf{0}_{d \times (м \breve{} \partial)} \\ \mathbf{0}_{(м \breve{} \partial) \times d} & \mathbf{0}_{(м \breve{} \partial) \times (м \breve{} \partial)} \end{pmatrix}$$

$$\breve{}_2 = \begin{pmatrix} \mathbf{0}_{\partial \times \partial} & \mathbf{0}_{d \times (м \breve{} \partial)} \\ \mathbf{0}_{(м \breve{} \partial) \times d} & \boldsymbol{Я}_м \otimes \breve{}_2 \end{pmatrix}$$

$$\breve{}_3 = \begin{pmatrix} \mathbf{0}_{\partial \times \partial} & \mathbf{0}_{d \times (м \breve{} \partial)} \\ \mathbf{0}_{(м \breve{} \partial) \times d} & \boldsymbol{Я}_м \otimes \breve{}_3 \end{pmatrix}.$$

Матрица точности линейна по параметрам точности и поэтому подходит для метода СОП. В этом случай, две случайные составляющие, моделирующие соответственно кривая численности и индивидуальные отклонения, с

$\boldsymbol{Z}_1 = \mathbf{1}_м \otimes \boldsymbol{Б}_+$ и $\boldsymbol{G}_1^{-1} = \sigma_1^{-2} +$,

$\boldsymbol{Z}_2 = \boldsymbol{I}_m \otimes \breve{\boldsymbol{B}}$ и

$$\boldsymbol{грамм}_2^{-1} = \sigma_2^{-2} \; \boldsymbol{Я}_м \otimes \breve{}_2 + \sigma_3^{-2} \; \boldsymbol{Я}_м \otimes \breve{}_3$$

$$= \sigma_2^{-2} \; \boldsymbol{Я}_м \otimes \breve{\boldsymbol{D}}_q^{\,\breve{}} \breve{\boldsymbol{D}}_q + \sigma_3^{-2} \boldsymbol{Я}_м \otimes \boldsymbol{я}_{\breve{} \partial} .$$

Напомним , что $\boldsymbol{X} = \mathbf{1}_м \otimes \boldsymbol{BU}_0$, и заметим , что ранг $(\mathbf{1}_м \otimes \boldsymbol{BU}_0) =$ ранг $(\boldsymbol{BU}_0)$. Теперь покажем, что условие (i) теоремы 2 выполняется для всех параметров дисперсии:

σ_1^2 : По построению $+$ положительно определен и имеет полный ранг, и rank $(\boldsymbol{X}, \boldsymbol{Z}_1) =$ rank $(\boldsymbol{X})$ + rank $(\boldsymbol{Z}_1)$. Отмечая, что ранг $(\boldsymbol{X}, \boldsymbol{Z}_1 \boldsymbol{G}_{1+}) =$ ранг $(\boldsymbol{X}, \boldsymbol{Z}_1) >$ ранг $(\boldsymbol{X})$, состояние проверено.

σ_2^2 : По, например, следствию 18.2.2 в Харвилле(1997), $\boldsymbol{G}_2$ и $\boldsymbol{Я}_м \otimes \breve{\boldsymbol{D}}_q^{\,\breve{}} \boldsymbol{D}_q$ коммутируют. Таким образом, пока $m > 1$, легко показать, что

классифицировать $\left(\boldsymbol{X}, \boldsymbol{Z}_2 \boldsymbol{Я}_м \otimes \breve{\boldsymbol{D}}_q^{\,\breve{}} \boldsymbol{D}_q\right)$

$= $ ранг $\left(\mathbf{1}_м \otimes \boldsymbol{BU}_0, \boldsymbol{I}_м \otimes \breve{\boldsymbol{B}}\, \breve{\boldsymbol{D}}_q^{\,\breve{}} \boldsymbol{D}_q\right) >$ ранг $(\boldsymbol{BU}_0)$.

σ_3^2 : Обратите внимание, что $\boldsymbol{G}_2$ и $\boldsymbol{I}_м \otimes \boldsymbol{I}_{\breve{} \partial}$ добираться. Как и раньше, пока при $m > 1$ легко показать, что

классифицировать $\left(\boldsymbol{X}, \boldsymbol{Z}_2 \boldsymbol{Я}_м \otimes \boldsymbol{I}_{\breve{} \partial}\right) =$ ранг $\left(\mathbf{1}_м \otimes \boldsymbol{BU}_0, \boldsymbol{I}_м \otimes \breve{\boldsymbol{B}}\right)$

$>$ ранг $(\boldsymbol{BU}_0)$.

5 примеров

В этом разделе представлены несколько примеров данных, в которых СОП метод представляет собой мощную альтернативу существующим оценкам Математические процедуры. Мы обсуждаем три различных анализа: первые два примера относятся к пространственно адаптируемым Р-шлицы, но каждый из них имеет дело с разной ситуацией относительно сложности и цели; последний пример посвящен иллюстрирующий наш метод анализа иерархической кривой данные. Все вычисления проводились в (64-битной) R 3.4.4. (Основная команда 2018) и Intel с тактовой частотой 2,30 ГГц ® Core i5 pro- компьютер cessor с 15,6 ГБ ОЗУ и Ubuntu 16.04 Операционная система LTS.

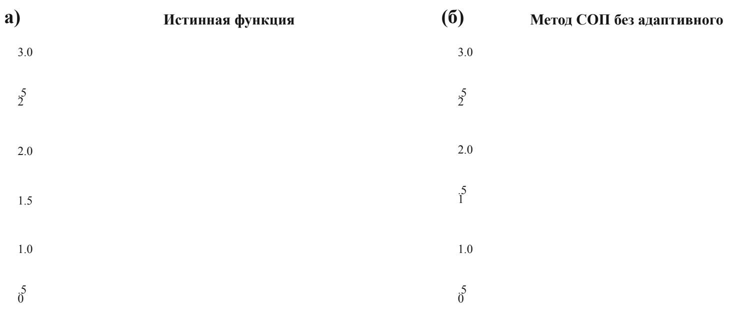
5.1 Доплеровская функция

В нашем первом примере мы рассматриваем функцию Доплера. Это распространенный пример в литературе по адаптивному сглаживанию. ture и обсуждался Рупперт и Кэрролл (2000),), Кривобокова и соавт. (2008 г.) и Тибширани (2014 г.), среди прочих эры. Данные формируются в соответствии с

$y_я$ знак равно грех $(4 / x_я) + 1.5 + \varepsilon_i$, $i = 1, \ldots, n,$

где $x_i \sim U[0, 1]$, $\varepsilon_i \sim N(0, 0.2^2)$ и $n = 1000$. Для подгонки При обработке данных мы предполагаем, что пространственно-адаптивная модель Р-сплайна обсуждается в разд. 4.1. Мы сравниваем производительность Метод СОП с реализованным в R-пакете mgcv, версия 1.8-23, описанная в Wood (2011). Стоит заметив, что оба подхода реализуют, по сути, одно и то же адаптивная Р-сплайн модель; единственная разница - это оценка процедура (и, возможно, репараметризация). Кроме того, мы также подбираем модель, не предполагая адаптивного штрафа. В этом случае метод СОП сводится к подходу Харвилла (см. раздел 2). Во всех случаях мы используем 200 кубических В-шлицев для представляют собой гладкую функцию вместе с дифференциалом второго порядка. особенности. Для адаптивных подходов 15 равноотстоящих кубические В-сплайны используются для параметров сглаживания [см. Уравнение (19)]. Эти значения выбраны для обеспечения достаточной гибкости способность к модели. В этой конфигурации всего 15 параметров дисперсии. Рисунок 2а показывает истинное сим- Вызванная функция Доплера. Рисунок 2b, с показывает, соответственно, расчетные кривые на основе метода СОП без и с адаптивным штрафом. Отображаются результаты с использованием MGCV. на фиг. 2d. Как и ожидалось, оба адаптивных подхода работают по аналогии. В указанной конфигурации они могут захватить 7 циклов функции Доплера. С другой стороны, неадаптивный подход способен уловить только 4 цикла и представляет очень непостоянные оценки, особенно справа - стороны ковариантной области. Что касается ED, то для Модель СОП без адаптивного штрафа, мы получаем общий ED 95 . 8 (из 200). Для моделей с адаптивным сглаживанием мы получили идентичные результаты, т. е. 50 . 2 (с СОП) и 50 . 0 (с

Стр.10



0,0

0,0 0,2 0,4 0,6 0,8 1.0

0,0

0,0 0,2 0,4 0,6 0,8 1.0

(с) **Метод СОП с адаптивным**

3.0

2,5

2.0

1.5

1.0

0,5

0,0

0,0 0,2 0,4 0,6 0,8 1.0

(г) **Пакет mgcv с адаптивным**

3.0

2,5

2.0

1,5

1.0

0,5

0,0

0,0 0,2 0,4 0,6 0,8 1.0

Фиг.2. Для функции Доплера: истинная функция (сплошная линия) и имитировать их обозначенные точки данных (серые точки), **b** оценочная кривая с использованием метода SOP без адаптивного штрафа (сплошная линия) вместе с 95% приблизительным кон- интервалы достоверности (пунктирные линии), **c** оценочная кривая с использованием метода СОП и адаптивный штраф (сплошная линия) вместе с 95% приблизительной конфиденциальностью. интервалы плотности (пунктирные линии); и **d** оценочная кривая с использованием MGCV пакет (сплошная линия) вместе с 95% приблизительным доверительным интервалом (пунктирные линии)

MGCV). Стоит помнить, что при использовании метода СОП общий ED получается путем сложения ED, связанных с каждый параметр дисперсии в модели (плюс размер фиксированная часть). Эти ED являются знаменателем оценки обновить выражения параметров дисперсии. Прирост метод СОП становится понятным, если сравнить вычислительные раз: 1 . 0 второй с нашим подходом (0 . 4 , если не считать адаптивный) и 45 секунд с использованием mgcv. Чтобы получить больше информации в выполнение обоих адаптивных подходов за пределами визуальное сходство, показанное на рис. 2 для одного набора данных, мы запускаем небольшое имитационное исследование. Выполняем в общей сложности 100 реплик- cates с той же спецификацией, описанной выше, и производят сравнения относительно среднеквадратичной ошибки (RMSE) и вычислительное время. В терминах RMSE среднее (стандартное отклонение- ation) на 100 повторов равно 0 . +265 *(* +0 . +0113 *)* для СОП и 0 . 267 *(* 0 . 0121 *)* с использованием mgcv. Что касается времени вычислений, среднее значение (стандартное отклонение) в секундах равно 0 . 84 *(* 0 . 22 *)* и 52 . 79 *(* 30 . 77 *)* для, соответственно, СОП и mgcv.

5.2 Данные дифракции рентгеновских лучей

В этом примере мы используем данные рентгеновской кристаллографии. Физическое радиационное сканирование тонкого слоя оксида индия и олова. Рентгеновский кристаллография позволяет исследовать молекулярные и атомная структура кристаллов. Кристаллографы точно поверните кристалл, указав его желаемую ориентацию, пока он **освещенный** пучком рентгеновских лучей. В зависимости от угла количество дифрагированных фотонов меняется, и они обнаруживаются и **подсчитывается** электронным детектором. Набор данных был проанализирован Дэвис и др.(2008 г., 2013 г.) и его можно найти в R- пакет дифрактометрии как оксид индия. Рисунок 3 показывает такое сканирование дифракции рентгеновских лучей (серые линии). Цель Рентгеноструктурный анализ предназначен для определения (а) базы сигнала - линия (и удалите ее); и (б) количество пиков (и выделить **их** для дальнейшего анализа их положения, высоты, симметрии, и так далее). Этот пример включен исключительно для иллюстрации возможности метода, представленного в этой статье, для анализа

3 00

200

Фотон считает

1 00

0

15 20 25 30 35 год

Угол дифракции

Рис. 3 Для примера рентгеновского излучения: оцененное плавное влияние угла дифракции на рентгеновское излучение с использованием метода SOP (сплошное черная линия). Серые линии представляют необработанные данные

анализ очень сложных данных. Для другого подхода к моделированию применительно к этому набору данных, см. Camarda et al.(2016 г.). Учитывая, что переменная результата представляет данные подсчета, модель Пуассона принят

гарантии могут быть извлечены. FA связана с диффузией воды и, таким образом, к диагностике и прогрессированию рассеянного склероза. Набор данных DTI содержит измерения FA здоровых и больных людей. als, зарегистрированные в нескольких местах вдоль тракта мозолистых волокон журнал (E [y_i]) = журнал (μ_i) = f (x_i),

где y_i и x_i (i = 1 , ..., 2000) обозначают соответственно количество фотонов и угол дифракции. Чтобы обеспечить достаточно гибкость модели, чтобы она могла улавливать пики (см. 3), мы используем 200 кубических B-сплайнов, а вторые - разницы порядка для функции и 80 кубических B-шлицев для адаптивный штраф. Результаты представлены на рис. 3 . Результаты с использованием пакета mgcv почти идентичны нашему pro- посал и не изображены. В этом случае наш метод требует меньше чем 3 с, тогда как MGCV примерно в 750 раз медленнее (33 мин). Что касается ED, всего мы получаем 29 . 5 и 24 . 9 с использованием СОП и МГКВ соответственно.

5.3 Данные сканирования визуализации тензора диффузии

Наш последний пример имеет дело с данными иерархической кривой. Мы ана- yse набор данных DTI, который можно найти в R-пакете возврат (Goldsmith et al. 2016 г.). Подробное описание исследование и данные можно найти в Goldsmith et al.(2011 г.), Goldsmith et al.(2012 г.) и Гревен и Шейпл (2017 г.). В Короче говоря, исследование было направлено на сравнение участков белого вещества в пациенты, страдающие рассеянным склерозом (РС), и здоровые видео. Рассеянный склероз - заболевание центральной нервной системы. система, которая вызывает поражения в трактах белого вещества, таким образом нарушение прохождения нервных импульсов к головному мозгу и от него и спинной мозг. Визуализация тензора диффузии (DTI) - это магнитное измерение метод визуальной визуализации, позволяющий изучать белое вещество трактуется, отслеживая диффузию воды на мозг. На основе DTI-сканирований измеряется фракционная анизотропия (FA). (CCA) и правый кортикоспинальный тракт в головном мозге. Рис. 4a, наблюдаемые измерения FA в разных местах тракта в CAA показаны отдельно для случаев и контроля, т. е. лица, пораженные и не пораженные РС. Каждая строка в эти графики представляют человека, и только первое посещение считается. Обратите внимание, что в целом пациенты с РС имеют более низкую Измерения FA, чем у здоровых людей.

В целях иллюстрации мы представляем два разных ана- исследования и сравнения. В первую очередь мы сосредотачиваем свой интерес на подгруппа лиц, страдающих рассеянным склерозом. В этой группе всего m = 99 особей, каждая с s = 93 Измерения FA в разных местах тракта CAA. СОП Метод используется для оценки модели, описанной в разд. 4.2 (для гауссовских гомоскедастических ошибок) и представлены в Dje- Ундже и Карри (2010). Для сравнения результатов и вычислений раз код, связанный с работой Джёндже и Карри(2010) также тестируется. В этом примере оба реализуют: ментаций используют массивную структуру данные: обобщенные модели линейных массивов (GLAM, Currie et al. 2006) используются для эффективного вычисления внутренних продуктов. для уравнений Хендерсона [Ур. (27)]. Здесь 43 кубических B- сплайновый базис используется для кривой населения, а 23 - для отдельные кривые. Эта конфигурация дает начало модели с 2320 (= 43 + 23 × 99) коэффициентами (оба случайные и исправлено). Методу SOP требуется около 150 секунд, чтобы соответствовать модель, а также Дженджe и Карри(2010) код 14 раз помедленнее. Отметим, что время вычислений может быть больше улучшается, если задействована разреженная структура матриц в модели эксплуатируется. Используя матрицу R-пакета, мы можем сократить время вычислений с помощью SOP до 35 секунд. Рисунок 4b показывает предполагаемый эффект популяции

а)

Органы управления

Случаи

FA − CCA

0,8 0,7 0,6 0,5 0,4 0,3 0,2

0 20 40 60 80

Расположение тракта

(б)

(с)

FA − CCA

0,8 0,7 0,6 0,5 0,4 0,3 0,2

0 20 40 60 80

Расположение тракта

Фиг.4. Для получения данных сканирования диффузии тензорной визуализации: наблюдаемые Ва- интервалы, основанные на стандартных ошибках полного сэндвича, предложенных Хек- ues вдоль тракта CCA. Слева: здоровые элементы управления. Справа: пациенты с рассеянным склерозом; man et al.(2013) Обратная линия - это наблюдаемое среднее, а пунктирная линия **b** расчетный профиль FA в популяции (группе) для лиц, затронутых РС. Сплошная красная линия: метод СОП. Пунктирная синяя линия: Djeundje и Currie (2010) код. Пунктирные красные линии обозначают точечную достоверность 95%. черные линии - эмпирические доверительные интервалы; и **c** оценочный (красный линии) и наблюдаемые (черные линии) индивидуальные профили ТВС для 3 выбранных частные лица. (Цветной рисунок онлайн) используя оба подхода, которые дают очень похожие результаты. 95% точечный доверительный интервал рассчитывается с использованием стандартные ошибки полного сэндвича, предложенные Heckman et al. (2013), но адаптированный к нашему случаю. Рисунок 4с показывает, для нескольких Пациенты с РС, предполагаемые (и наблюдаемые) профили ФА. В с точки зрения ЭД, получаем 35 . 03 (из 43) для населения кривая (включая непенализованную или фиксированную часть), а также общая 2025 . 78 (из 2275 (= 23 × 99 - 2)) для всех индивидуальных кривые. Отметим, что эта сумма соответствует сумме ED, связанные с дисперсиями σ^2_2 и σ^2_3 участвовали в моделирование этих отдельных кривых [см. выражения (23) а также (24) подробнее]. Точнее, методом СОП мы получаем ED 870 . 44 для σ^2_2 и 1155 . 34 для σ^2_3 .

Наш второй анализ рассматривает всех людей, случаи и контролирует. Интересно сравнить профили ТВС на первый визит между этими двумя группами. С этой целью Рассмотрена следующая поэтапная модель взаимодействия.

$$y_{ij} = f_{z_j}(t_i) + g_j(t_i) + \varepsilon_{ij} \quad 1 \le i \le s,\ 1 \le j \le m,$$

где $z_j = 1$, если *j-й* человек подвержен MS (случай) и $z_j = 0$ в противном случае (управление) и $\varepsilon_{ij} \sim N(0, \varphi)$. Примечание что эта модель предполагает различный профиль FA для каждой группы. Для этого анализа всего $m_0 = 42$ контроля и $m_1 = 99$ пациентов с МС $(m = m_0 + m_1 = 141)$, а $s = 93$ различные локации трактов. Подробное описание модели можно найти в «Приложении D». Что касается первого анализа 43 кубический базис B-сплайнов используется для кривых населенности (FA профилей) и 23 для отдельных кривых, что дает общее 3329 (= 43 × 2 + 23 × 141) коэффициентов. Использование GLAM и методы разреженной матрицы (R-пакет Matrix), соответствие занимает 65 секунд. Рисунок 5 а показаны расчетные профили ТВС. как для случаев, так и для контроля, вместе с 95% точечным контролем интервалы достоверности (Heckman et al. 2013). ED для FA Профиль в элементах управления - 32 . 21 год, у больных РС - 35 лет . 55. В В обоих случаях мы включаем фиксированную часть. Что касается инди-

а)

0,8
0,7
0,6
0,5
FA − CCA
0,4
0,3
0,2
0 20 40 60 80
Расположение тракта

(б)

0,8
0,7
0,6
0,5
FA − CCA
0,4
0,3
0,2
0 20 40 60 80
Расположение тракта

Рис. 5 Для данных сканирования визуализации тензора диффузии: расчетная популяция (группа) Профили ТВС. Слева направо: результаты с использованием метода СОП и функциональный регрессионный подход Гревена и Шайпля.(2017). Сплошная красная линия: пациенты с рассеянным склерозом. Сплошная синяя линия: элементы управления. (Цветной рисунок онлайн) а красные линии - точечные 95% доверительные интервалы. Для СОП доверительные интервалы метода основаны на стандарте полного сэндвича ошибки, предложенные Heckman et al.(2013 г.). Сплошная линия спины - это наблюдаемое среднее, а пунктирные черные линии - эмпирическая достоверность интервалы. (Цветной рисунок онлайн)

Визуальные кривые, мы получаем общий ED 2863 . 46, сумма 1263 . 26 и 1600 . 20.

Сравниваем результаты с функциональной регрессией. модель, представленная в статье Гревена и Шейпля (2017 г.) [см. модель (1.1) и рис. 2 в этой статье]. Мы хотели бы обратите внимание, что модель P-сплайна, используемая в этой статье, обсуждалась в Дурбане и Агилера-Морильо(2017) как конкурентоспособный альтернатива функциональному подходу Гревена и Шайпля (2017). Для функционального подхода мы рассматриваем гауссовский гомоскедастических ошибок, и, как полагают авторы, 25 кубический базис B-сплайнов и штраф за разность первого порядка, как а также 8 функций основных функциональных компонентов (FPC). Код для подгонки модели любезно предоставлен авторы. Результаты представлены на рис. 5 б. Как можно заметить, оба подхода дают очень похожие результаты. Однако обратите внимание что точечный 95% доверительный интервал для оцененных Профиль FA у пациентов с РС уже, чем эмпирический доверительный интервал. Эти результаты могут быть объяснены (поз. предположение о гауссовских гомоскедастических ошибках. С точки зрения времени вычислений модель функциональной регрессии требуется 895 с для установки (в отличие от 65 с при использовании SOP). Как известно, что время вычислений обоих подходов (СОП метод и функциональный подход) не полностью сопоставимы поскольку они предполагают разные характеристики модели.

предыдущие подходы, метод SOP подходит для моделей где матрица точности линейна по параметру точности эфиры. Эти матрицы точности распространены, когда их штрафуют. гладкие модели переформулируются как (обобщенные) линейные смешанные модели. Они появляются при наличии нескольких квадратичных штрафы, действующие на те же коэффициенты регрессии. Один частный случай - анизотропные P-сплайны тензорного произведения. Эта сидячая- Эта проблема обсуждалась в статье Rodríguez-Álvarez et al. (2015b), где был предложен алгоритм SAP. Настоящее бумага делает еще один шаг и обобщает SAP на более общий случай. Насколько нам известно, СОП - единственный метод в в этом контексте, когда оценки параметров дисперсии включают «Частичные» эффективные степени свободы. Как следствие, метод обеспечивает, при сходимости, оценочную эффективную пар- основные степени свободы, связанные с каждым гладким / случайным компонент в модели. Это особенно актуально, когда работа с моделями сглаживания, где эффективные градусы свободы каждого гладкого термина дают важную информацию о сложность подобранной функции. Кроме того, мы показываем в статье говорится, что метод СОП обеспечивает неотрицательную оценку сопоставить параметры дисперсии и обсудить условия при котором эти оценки строго положительны.

Мы представляем в статье несколько ситуаций в контексте штрафная сплайн-регрессия, в которой СОП представляет собой

6 Обсуждение

В этой статье представлен новый метод оценки, называемый СОП, для (обобщенных) линейных смешанных моделей. Метод представляет собой расширение предыдущего предложения Харвилла (1977 г.), и gen-сформулировано Шаллом(1991 г.) среди других. В отличие от тех

Хорошая альтернатива существующим методам оценки. В частности, мы показываем использование СОП в случае пространственно-адаптивного P-сплайнов и для оценки профильных кривых в продольные данные. Мы обсуждаем несколько сделок по анализу реальных данных - справиться с этими ситуациями и показать превосходство СОП с точки зрения времени вычислений. Используем простое моделирование ситуации с целью описания метода, модели и примеры подробно, избегая обобщений

Стр. 14

это могло затруднить чтение статьи и затруднить понимание простота предложения. Однако есть несколько другие области применения СОП. Например, перекрытие-Штрафы за пинг появляются в исследованиях изображений мозга (Karas et al. 2017) и оценка производной кривой (Симпкин и Ньюэлл. 2013). Также метод можно использовать для более сложных модели, включающие линейные эффекты (многомерные) гладкие функции, случайные гауссовские эффекты и т. д. Использование других базисные функции вне B-сплайнов также возможны, если квадратичные штрафы складываются.

Предложение и примеры, представленные в этой статье подготовить почву для дальнейших исследований. Например, подход, обсуждаемый для адаптивных P-сплайнов, основан на о сглаживании локально изменяющегося параметра сглаживания на средствами B-шлицев. Это означает, что гладкость исключительно контролируется номером B-сплайнового базиса. Выбор подходящее базовое измерение может быть непростой задачей, с нежелательным последствием является то, что при большом базисном измерении выбранный (больше, чем необходимо), мы можем получить локальную лин-подходит для ушей. Чтобы уменьшить влияние базового измерения, мы будем изучить возможность включения штрафа в коэффициенты (вари-параметры), связанные с базисом B-сплайна. Этот может быть выполнено с помощью иерархической структуры для случайных эффектов (см., например, Кривобокова и др. 2008 г.). Еще одна сложная область - изучение подходящих штрафов. и эффективные методы оценивания адаптивных P-сплайнов в более одного измерения. В то время как некоторые попытки были выполняется в двух измерениях (см., например, Crainiceanu et al. 2007 г.; Кривобокова и соавт. 2008 г.), насколько нам известно, в литературе отсутствуют трехмерные подходы (например, пространство и время). Некоторые предварительные результаты использования СОП: доступный в Rodríguez-Álvarez et al. (2016), но дальнейшие работы еще нужно сделать. Также использование и расширение СОП метод анализа многоуровневых и продольных функций национальные данные (см., например, Greven and Scheipl 2017) представляет захватывающая область исследований. Наконец, для выбора переменных существуют некоторые работы, предлагающие разреженные регрессионные модели. els с использованием локальных квадратичных приближений к L1-норме применяя подход со штрафной вероятностью (см. Фань и Ли 2001 ; Хантер и Ли2005 г.; Цзоу и Ли 2008). В последнее время, (обобщенные) линейные подходы со смешанными эффектами также были предложено (см. Taylor et al. 2012 ; Гролл и Тутц2014) allow-для оценки штрафа одновременно с параметры дисперсии с использованием REML. Мы намерены продлить Метод СОП в этом направлении. В заключение эта статья открывается путь к общему методу оценки, учитывающему как сглаживание, так и выбор переменных в разумных вычислениях. время.

R-код, используемый для представленных примеров реальных данных в разд. 5, а также R-пакет, реализующий СОП метод обобщенных линейных смешанных моделей, спа-полностью адаптивные модели P-сплайна и модели P-сплайна для данные иерархической кривой можно загрузить с https: // bitbucket.org/mxrodriguez/sop.

Благодарности Это исследование было поддержано Баскским Правительство через программу BERC 2018-2021 и Span-иш Министерство экономики и конкурентоспособности MINECO через Аккредитация передового опыта BCAM Северо-Очоа SEV-2013-0323 и через проекты MTM2017-82379-R, финансируемые (AEI / FEDER, UE) и аббревиатура «AFTERAM», MTM2014-52184-P и MTM2014-55966-P. Данные MPT / DTI были собраны в Университете Джона Хопкинса и Кеннеди-Кригера. Мы благодарны Педро Каро и Иэну Кур-гие за полезные обсуждения, Мартину Буру и Кахо тер Брааку за подробное прочтение статьи и их многочисленные предложения, а также Басу Энгель за то, что поделился с нами своими знаниями. Мы также благодарны двум рецензентам за конструктивные комментарии к статье.

Доказательство теоремы 1.

Доказательство. Прежде всего отметим, что частные производные первого порядка (приближенной) функции логарифмического правдоподобия REML может быть выражается как (см., например, Rodríguez-Álvarez et al. 2015b)

$$\frac{\partial l}{\partial \sigma^2_{kl}} = -\frac{1}{2}\,\text{след}\left(\boldsymbol{Z PZ}\frac{\partial \boldsymbol{G}}{\partial \sigma^2_{kl}}\right) + \frac{1}{2}\,\text{A}\,\boldsymbol{G}^{-1}\frac{\partial \boldsymbol{G}}{\partial \sigma^2_{kl}}\,\text{грамм}^{-1}\boldsymbol{\alpha}.$$

Учитывая, что $\boldsymbol{G}$ - положительно определенная матрица, имеем тождество

$$\frac{\partial \boldsymbol{G}}{\partial \sigma^2_{kl}} = -\boldsymbol{G}\frac{\partial \boldsymbol{G}^{-1}}{\partial \sigma^2_{kl}}\boldsymbol{G},$$

и поэтому

$$\frac{\partial \boldsymbol{G}}{\partial \sigma^2_{kl}} \text{ знак равно } \frac{1}{\sigma^4_{kl}}\,\text{диагональ}\left(\boldsymbol{0}^{(1)}, \boldsymbol{G}_{k\,kl}\boldsymbol{G}_k, \boldsymbol{0}^{(2)}\right),$$

где $\boldsymbol{0}^{(1)}$ и $\boldsymbol{0}^{(2)}$ - нулевые квадратные матрицы соответствующих Габаритные размеры.

Частные производные первого порядка от REML log-функция правдоподобия тогда выражается как

$$2\quad \frac{\partial l}{\partial \sigma^2_{kl}} = -\frac{1}{\sigma^4_{kl}}\,\text{след}\left(\boldsymbol{Z}_k\boldsymbol{PZ}_k\boldsymbol{G}_{k\,kl}\boldsymbol{G}_k\right) + \frac{1}{\sigma^4_{kl}}\boldsymbol{\alpha}_\kappa\ {}_{KL}\,\boldsymbol{\&\ alpha}\ ;_{\kappa}.$$

Когда оценки REML положительны, они получаются приравняв первое выражение к нулю (см., например, Engel 1990)

$$\text{след}\left(\boldsymbol{Z}_k\,\boldsymbol{PZ}_k\boldsymbol{G}_{k\,kl}\boldsymbol{G}_k\right)^{\boldsymbol{\alpha}_\kappa\ {}_{\kappa л}\boldsymbol{\alpha}_\kappa} = 1.$$

Теперь умножим обе части на σ^2_{kl} и оцените левый-сторона для предыдущих итераций и правая часть для обновление, получение

Стр. 15

$$\sigma^2_{kl} = \frac{\left(A^{[m]}_k \alpha_k \ldots \alpha^{[m]}_k\right) C \sigma^{2[m]}_{kl}}{\text{след}\; Z_k\, n^{[m]} Z_k G^{[m]}_k\, {}^{[m]}_{kl} G_k}$$

$$\text{знак равно} \frac{\left(A^{[m]}_k \alpha_k \ldots \alpha^{[m]}_k\right)}{\text{след}\; Z_k\, n^{[m]} Z_k G^{[m]}_k\, {}_{kl}\, {}^{[m]}_k \sum_{kl} \sigma^{2[m]}}\; \text{грамм}.$$

⊓⊔

В Доказательство теоремы 2

Доказательство Прежде всего напомним некоторые обозначения и введем некоторые нужны результаты. Обозначим как $P = V^{-1} - V^{-1} X\left(X V^{-1} X\right)^{-1} X V^{-1}$, где $V = R + ZGZ$, $R = \varphi W^{-1}$ с W бытии диагональная матрица весов, участвующая в подсчете очков Фишера алгоритм.

Обозначим через $C(A)$ линейное пространство, натянутое на столбцы множества A, и пусть $P_{XV^{-1}} = X\left(X V^{-1} X\right)^{-1} X V^{-1}$ быть ортогональная проекционная матрица для $C(X)$ относительно V^{-1}. Легко показать, что

$$P = V^{-1}\left(I_n - P_{XV^{-1}}\right) \text{знак равно} \left(I_n - P_{XV^{-1}}\right) V^{-1}\left(I_n - P_{XV^{-1}}\right).$$

По теореме 14.2.9 Харвилля (1997), P является (симметричным) положительная полуопределенная матрица. Кроме того,

$$PX = \mathbf{0},$$

а также

$$\begin{aligned} \text{ранг}(P) &= \text{ранг}\left(V^{-1}\left(I_n - P_{XV^{-1}}\right)\right) \\ &= \text{ранг}\left(\left(I_n - P_{XV^{-1}}\right)\right) \\ &= n - \text{ранг}\left(P_{XV^{-1}}\right) \\ &= n - \text{ранг}(X). \end{aligned}$$

Таким образом,

$$ker(P) = C(X), \tag{26}$$

т.е. $Px = \mathbf{0}$ тогда и только тогда, когда x находится в $C(X)$.

Пусть $_{kl} = UU$ - разложение на собственные значения $= U_+ + U_+$, где U_+ и $+$ - $_{k\,l}$. Отметим, что $_{kl}$ подматрицы, соответствующие ненулевым собственным значениям. потом

$$\alpha_k \quad {}_{kl}\,\&\,alpha\,;\, {}_{k} = \alpha_k U_+ U_+ \alpha_k = z\, PZ_k G_k U_{++} U_+ G_k Z_k Pz \geqslant 0,$$

с равенством тогда и только тогда, когда $U_+ G_k Z_k\; Pz = \mathbf{0}$ (поскольку $_+$ положительно определен). Таким образом, используя результат (26) равенство выполняется, если z принадлежит $C(X)$ или $C(Z_k G_k U_+) \subset C(X)$. По лемме 4.2.2 и следствие 4.5.2 в Харвилле(1997 г.), у нас есть

$$C(Z_k G_k U_+) = C\left(Z_k G_{k\,kl}\right) \subset C(X) \iff \text{ранг}\left(Z_k G_{k\,kl}, X\right) = \text{ранг}(X).$$

Относительно знаменателя оценок на основе REML обновления, мы следуем аналогичным рассуждениям. Использование следствия 14.7.5 (и теорема 14.2.9) в Харвилле (1997), имеем

$$\text{след}\left(Z_k PZ_k G_{k\,kl} G_k\right) = \text{след}\left(U_+ G_k Z_k PZ_k G_k U_{++}\right) \geqslant 0,$$

с равенством тогда и только тогда, когда $U_+ G_k Z_k\; PZ_k G_k U_+ = \mathbf{0}$. Опять же, это равенство выполняется тогда и только тогда, когда $C(Z_k G_k U_+) \subset C(X)$ (т.е. $\iff$ rank $\left(Z_k G_{k\,kl}, X\right) = \text{ранг}(X)$). ⊓⊔

C Алгоритм оценки

В этом приложении кратко описаны этапы алгоритма оценки. рифм для модели (10) на основе метода СОП. Напомним, что интерес заключается в оценочной модели

$$g(\mu) = X\beta + Z\alpha = X\beta + \sum_{k=1}^{c} Z_k \alpha_k,$$

где $Z = [Z_1, \ldots, Z_c]$, $\alpha = \left(\alpha_1, \ldots, \alpha_c\right)$, $\alpha_k \sim N(\mathbf{0}, G_k)$ с G_k^{-1} знак равно $\sum_{l=1}^{p_k} \sigma^{-2}_{kl}\, {}_{kl}$ и $\alpha \sim N(\mathbf{0}, G)$ с $G = \bigoplus_{k=1}^{c} G_k$.

Инициализация Установить начальные значения для параметры дисперсии eters $\sigma^{2[0]}_{kl}$ ($l = 1, \ldots, p_k$ и $k = 1, \ldots, c$). В те ситуации, когда φ неизвестен, устанавливают Начальное значение для этого параметра, $\varphi^{[0]}$. Установите $t = 0$.

Шаг 1 Постройте *рабочий* вектор отклика z и матрица весов W в следующем виде

$$z_i = g\left(\mu^{[m]}_i\right) + \left(y_i - \mu^{[m]}_i\right) g\left(\mu^{[m]}_i\right),$$

$$w_{ii} = \frac{1}{g\left(\mu^{[m]}_i\right)^2 v\left(\mu^{[m]}_i\right)}.$$

Шаг 1.1. Учитывая начальные *оценки* параметров дисперсии, оценить α и β, решив

$$\underbrace{\begin{bmatrix} XR^{[m]-1} X & XR^{[m]-1} Z \\ ZR^{[m]-1} XZR^{[m]-1} & Z + G^{[m]-1} \end{bmatrix}}_{C^{[t]}} \underbrace{\begin{bmatrix} \beta \\ \alpha \end{bmatrix}} \text{знак равно} \begin{bmatrix} XR^{[m]-1} z \\ ZR^{[m]-1} z \end{bmatrix}, \tag{27}$$

Стр.16

где $R^{[T]} = \varphi^{[m]} W^{-1}$. Пусть $\alpha^{[m]}$ и $\beta^{[m]}$ будь этим оценки.

Шаг 1.2 Обновите параметры дисперсии следующим образом

где $z_j = 1$, если *j-й* человек подвержен MS (случай) и $z_j = 0$ в противном случае (контроль). Упорядочим данные с сначала наблюдения за элементами управления, а затем наблюдения за

$$\sigma^2_{kl} = \frac{A^{[m]}_{k} \alpha^{[m]}_{k}}{ED^{[m]}_{kl}}$$

и, при необходимости,

$$\varphi = \frac{\left(z - X\beta^{[m]} - Z\alpha^{[m]}\right) W \left(z - X\beta^{[m]} - Z\alpha^{[m]}\right)}{n - \text{ранг}(X) - \sum_{k=1}^{c} \sum_{l=1}^{p_k} ED^{[m]}_{kl}},$$

с участием

$$ED^{[m]}_{kl} = \text{след}\left(\left(\text{грамм}^{[m]}_{k} \cdot C^{[t]*}_{kk}\right) \frac{k_l}{\sigma^{2[m]}_{kl}}\right).$$

Напомним, что $C^{[t]*}$ обозначает обратное к $C^{[m]}$ в (27), и $C^{[t]*}_{kk}$ - это разбиение $C^{[t]*}$ соответствующий к k -й случайной составляющей α_k.

Шаг 1.3 Повторите шаги 1.1 и 1.2, заменив $\hat{g}\sigma^{2[t]}_{kl}$, а также, если они будут обновлены, φ, до сходимости. Для примеров, представленных в разд. 5, мы используем Отклонение REML как критерий сходимости.

Шаг 2 Повторите шаг 1, заменив параметры дисперсии и фиксированные и случайные эффекты модели (и, следовательно, $M^{[m]}$) на полученные на последней итерации шагов 1.1 – Шаг 1.3, до сходимости.

Стоит отметить, что для удобства записи в примеры, описанные в разд. 4, матрица точности была переписан как $G^{-1} = \sum_{l=1}^{p} \sigma^{-2}_{l} \tilde{}_{l}$, где $p = \sum_{k=1}^{c} p_k$ - количество параметров дисперсии, $\tilde{}_{l}$ - матрица ces $_{kl}$ дополнены нулями. Здесь $\tilde{}_{l}$ s матрицы размерности $q \times q$, где $q = \sum_{k=1}^{c} q_k$ - количество ран- коэффициенты эффектов dom. Для этой спецификации оценка Алгоритм, рассмотренный выше, остается по существу тем же самым, но на шаге 1.2 $\alpha^{[t]}_{k}$, $G^{[m]}_{k}$, $_{kl}$ заменяются на соответственно $\alpha^{[t]}$, $G^{[m]}$ и $\tilde{}_{l}$, и $C^{[t]*}_{kk}$ будет тем разбиением $C^{[t]*}$ соответствующему случайному вектору α (и, следовательно, то же самое для все параметры дисперсии).

D Посторонняя модель иерархической кривой

В этом приложении подробно описывается фактор за кривой зависимости модель действия, обсуждаемая в разд. 5.3, т.е.

$$y_{ij} = f_{z_j}(t_i) + g_j(t_i) + \varepsilon_{ij} \quad 1 \le i \le s,\ 1 \le j \le m,$$

Пациенты с рассеянным склерозом. В матричных обозначениях модель может быть выражена в виде

$$y = [Q \otimes B]\theta + [I_m \otimes \breve{B}]\breve{\theta} + \varepsilon, \tag{28}$$

с B, $\breve{B}$, $\breve{\theta}$, как определено в разд. 4.2, и $\varepsilon = \left(\varepsilon_{\cdot,1}, \ldots, \varepsilon_{\cdot,m}\right)$, с $\varepsilon_{\cdot j} = \left(\varepsilon_{1j}, \ldots, \varepsilon_{sj}\right)$. Матрица Q представляет собой любую подходящую матрицу. траст-матрица размерности $m \times 2$, где $m = m_0 + m_1$, с m_0 - количество элементов управления, а m_1 - количество MS. пациенты. Для нашего приложения мы рассматриваем

$$Q = \begin{pmatrix} 1_{m_0} & 0_{m_0} \\ 0_{m_1} & 1_{m_1} \end{pmatrix},$$

и для f_0 и f_1 предполагается разная степень сглаживания , т.е. матрица штрафа, действующая над вектором коэффициентов θ имеет вид

$$P = \begin{pmatrix} \lambda_1 D_q D_q & 0_{c\times c} \\ 0_{c\times c} & \lambda_2 D_q D_q \end{pmatrix}.$$

Переформулировка смешанной модели может быть выполнена аналогичным образом. мода на описанную в разд. 4.2, в этом случае

$$X = [Q \otimes BU_0],$$

$$Z = [Q \otimes BU_+ : I_m \otimes \breve{B}],$$

а также

$$\text{грамм}^{-1} = \begin{pmatrix} \sigma^{-2}_{1} + & 0_{d\times d} & 0_{d\times (m\breve{d})} \\ 0_{d\times d} & \sigma^{-2}_{2} + & 0_{d\times (m\breve{d})} \\ 0_{(m\breve{d})\times d} & 0_{(m\breve{d})\times c} & Я_m \otimes \breve{G}^{-1} \end{pmatrix},$$

где

$$\breve{G}^{-1} = \sigma^{-2}_{3} \breve{D}_q \breve{D}_q + \sigma^{-2}_{4} \breve{Я}_d.$$

Рекомендации

Бреслоу, Н. Э., Клейтон, Д. Г.: Приближенный вывод в обобщенном линейные смешанные модели. Варенье. Стат. Доц. **88** (421), 9–25 (1993).

Camarda, CG, Eilers, PH, Gampe, J .: Суммы гладких экспонент к разложить сложные серии отсчетов. Стат. Модель. **16** (4), 279–296 (2016)

Крайничану, К.М., Рупперт, Д., Кэрролл, Р.Дж., Джоши, А., Гуднер, Б .: Пространственно адаптивные байесовские штрафные сплайны с гетероскедастикой ошибки. J. Comput. График. Стат. **16** (2), 265–288 (2007).

Крамп, С.Л.: Текущее состояние анализа компонент дисперсии. Био- метрики **7** (1), 1–16 (1951)

Цуй, Й., Ходжес, Дж. С., Конг, Х., Карлин, Б. П.: Степени разделения свобода в иерархических и других моделях с богатыми параметрами. Технометрика **52** , 124–136 (2010)

Карри, И. Д., Дурбан, М .: Гибкое сглаживание с помощью P-шлицев: унифицированное подход. Стат. Модель. **2** (4), 333–349 (2002).

Карри, И. Д., Дурбан, М., Эйлерс, PHC: Обобщенные модели линейных массивов с приложениями к многомерному сглаживанию. JR Stat. Soc. Сер. B (Стат. Метод .) **68** (2), 259–280 (2006)

Дэвис, П.Л., Гатер, У., Мейз, М., Мергель, Д., Милденбергер, Т .: Остаточная локализация и количественная оценка пиков в рентгеновских лучах дифрактограммы. Аня. Прил. Стат. **2** (3), 861–886 (2008).

Дэвис, П.Л., Гатер, У., Мейз, М., Мергель, Д., Милденбергер, Т., Берн- Холт Т., Хофмайстер Т .: дифрактометрия: идентификация базовой линии и разложение пиков для рентгеновских дифрактограмм. Версия пакета R 0,1-10 (2018)

Djeundje, VA, Currie, ID: Соответствующая ковариационная спецификация через штрафы за штрафные шлицы в смешанных моделях для продольных

Хасти, Т.Дж., Тибширани, Р.Дж .: Обобщенные аддитивные модели. Чепмен & Холл, Лондон (1990)

Хекман, Н., Локхарт, Р., Нильсен, Дж. Д.: Штрафная регрессия, смешанная модели эффектов и соответствующее моделирование. Электрон. J. Stat. **7** , 1517–1552 (2013)

Хендерсон, К.Р .: Селекционный индекс и ожидаемый генетический прогресс. Стат. Genet. Порода растений. **982** , 141–163 (1963)

Хантер, Д.Р., Ли, Р.: Выбор переменных с использованием алгоритмов ММ. Аня. Стат. **33** (4), 1617–1642 (2005).

Джонсон, Д.Л., Томпсон, Р .: Ограниченная оценка максимального правдоподобия. определение компонентов дисперсии для одномерных моделей животных с использованием методы разреженных матриц и усредненная информация. J. Dairy Sci. **78** , 449–456 (1995).

Карас, М., Бжиски, Д., Джемидзич, М., Гони, Дж., Карекен, Д.А., Randolph, TW, Harezlak, J .: Информация о подключении мозга методы регуляризации регрессии. Стат. Biosci. (2017). https://doi.org/10.1007/s12561-017-9208-x данные. Электрон. J. Stat. **4** , 1202–1224 (2010).

Дурбан, М., Агилера-Морильо, М.К .: Об оценке функциональных случайные эффекты. Стат. Модель. **17** (1–2), 50–58 (2017)

Дурбан, М., Харезлак, Дж., Ванд, депутат, Кэрролл, Р.Дж .: Простая установка тематические кривые для продольных данных. Стат. Med. **24** (8), 1153–1167 (2005)

Eilers, PHC: обсуждение Verbyla et al. JR Stat. Soc. Сер. С (Прил. Стат.) **48** , 300–311 (1999).

Eilers, PHC, Marx, BD: гибкое сглаживание с помощью B-шлицев и штрафы. Стат. Sci. **11** (2), 89–121 (1996)

Энгель Б .: Анализ несбалансированных линейных моделей с дисперсионным соеди- нением. компоненты. Стат. Neerl. **44** , 195–219 (1990).

Энгель, Б., Буист, В .: Анализ обобщенной линейной смешанной модели: тематическое исследование и результаты моделирования. Биом. J. **38** (1), 61–80 (1996).

Энгель Б., Кин А. Простой подход к анализу обобщения. линейные смешанные модели. Стат. Neerl. **48** (1), 1–22 (1994)

Фан, Дж., Ли, Р .: Выбор переменной с помощью невогнутой штрафной вероятности и его оракульные свойства. Варенье. Стат. Доц. **96** (456), 1348–1360 (2001)

Гилмор, AP, Томпсон, Р., Каллис, Б.Р .: Средняя информация REML: эффективный алгоритм оценки параметров дисперсии в линейных смешанных моделях. Биометрия **51** (4), 1440–1450 (1995)

Голдсмит, Дж., Бобб, Дж., Крайничану, С.М., Каффо, Б., Райх, Д.: Пенитенциарные меры. рекомендованная функциональная регрессия. J. Comput. График. Стат. **20** (4), 830–851 (2011)

Голдсмит, Дж., Крайничану, К.М., Каффо, Б., Райх, Д.: продольный штрафной функциональной регрессии для когнитивных результатов на ней-измерения ронального тракта. JR Stat. Soc. Сер. С (Прил. Стат.) **61** (3), 453–469 (2012)

Голдсмит, Дж., Шайпл, Ф., Хуанг, Л., Вробель, Дж., Геллар, Дж., Харезлак, Дж., Маклин, М. В., Свихарт, Б., Сяо, Л., Крейничану, К., Рейсс, PT: возврат: регрессия с функциональными данными. Версия пакета R 0,1-16 (2016)

Грейзер, Х.-У., Смит, С.П., Таер, Б .: Подход без производных для оценка компонентов дисперсии в моделях на животных ограниченным максимальная вероятность. J. Anim. Sci. **2** (64), 1362–1373 (1987).

Грин, П.Дж.: вероятность общей полупараметрической регрессии с учетом штрафных санкций. модели. Int. Стат. Rev./Revue Internationale de Statistique **55** (3), 245–259 (1987)

Гревен, С., Шайпл, Ф .: Общая основа функциональной регрессии. моделирование. Стат. Модель. **17** (1–2), 1–35 (2017)

Groll, A., Tutz, G .: Выбор переменных для обобщенного линейного смешанного модели по L1-штрафной оценке. Стат. Comput. **24** (2), 137–154 (2014)

Харвилл, Д.А.: Подходы максимального правдоподобия к компоненту дисперсии оценка и связанные с ней проблемы. Варенье. Стат. Доц. **72** (358), 320–338 (1977)

Харвилл, Д.А.: Матричная алгебра с точки зрения статистики. Спрингер, Берлин (1997)

Кривобокова, Т .: Выбор параметра сглаживания в двух кадрах. работает для штрафных шлицев. JR Stat. Soc. Сер. B (Стат. Methodol.) **75** (4), 725–741. https://rss.onlinelibrary.wiley.com/doi / pdf / 10.1111 / rssb.12010 (2009 г.)

Кривобокова Т., Крайничану С.М., Кауэрманн Г .: Быстрая адаптивная штрафные шлицы. J. Comput. График. Стат. **17** (1), 1–20 (2008)

Ли, Д.-Дж .: Смешанная модель сглаживания для пространственных и пространственно-временных данные. Кандидатская диссертация. Департамент статистики Университета Карлоса III де Мадрид, Испания (2010)

Маккаллах, П., Нелдер, Дж .: Обобщенные линейные модели. Чепмен и Монографии Холла / CRC по статистике и прикладной теории вероятностей Серия, 2-е изд. Чепмен и Холл, Лондон (1989)

Паттерсон, HD, Томпсон, Р .: Восстановление межблочной информации когда размеры блоков не равны. Биометрика, **58** (3), 545–554 (1971)

Основная группа R: R: Язык и среда для статистической коммуникации положить. Фонд R для статистических вычислений, Вена, Австрия (2018)

Рейсс П.Т., Огден Р.Т .: Выбор параметра сглаживания для класса полупараметрических линейных моделей. JR Stat. Soc. Сер. B Стат. Методол. **71** (2), 505–523 (2009).

Родригес-Альварес, М.Х., Дурбан, М., Ли, Д.-Дж., Эйлерс, ПМСП: Быстрое оценивание многомерных адаптивных p-сплайн-моделей. В: Friedl, H., Wagner, H. (eds.) Proceedings of 30th International семинар по статистическому моделированию, стр. 330 - 335. arXiv: 1610.06861 (2015a)

Родригес-Альварес, М.Х., Ли, Д.-Дж., Кнейб, Т., Дурбан, М., Эйлерс, PHC: быстрое разделение параметров сглаживания в многомерных обобщенные P-сплайны: алгоритм SAP. Стат. Comput. **25** , 941–957 (2015b)

Родригес-Альварес, М.Х., Дурбан, М., Ли, Д.-Дж., Эйлерс, PHC, Гонсалес, Ф .: Пространственно-временные адаптивные штрафные сплайны с приложение к нейробиологии. В: Dupuy, J.-F., Josse, J. (eds.) Материалы 31-го Международного семинара по статистике Моделирование, стр. 267–272. arXiv: 1610.06860 (2016)

Родригес-Альварес, МХ, Бур, член парламента, ван Иувийк, Ф.А., Эйлерс, PH: Исправление пространственной неоднородности в экспериментах по селекции растений с р-образными шлицами. Плевать. Стат. **23** , 52–71 (2018)

Рупперт Д., Кэрролл Р.Дж.: Пространственно-адаптивные штрафы за подгонку шлицевого соединения. Aust. NZJ Stat. **42** (2), 205–223 (2000)

Рупперт, Д., Ванд, М. П., Кэрролл, Р .: Полупараметрическая регрессия. Cam-Бридж Юниверсити Пресс, Кембридж (2003)

Шалл, Р .: Оценка в обобщенных линейных моделях со случайными эффектами. Биометрика **78** (4), 719–727 (1991).

Симпкин, А., Ньюэлл, Дж .: Подход с аддитивным штрафом p-сплайнов к производной тивная оценка. Comput. Стат. Data Anal. **68** , 30–43 (2013)

Смит С.П .: Оценка генетических параметров в нелинейных моделях. В: Джанола Д., Хаммонд К. (ред.) Достижения в статистических методах для генетического улучшения животноводства. Продвинутая серия в сельском хозяйстве культурологии, т. 18. Springer, Берлин, Гейдельберг (1990)

Тейлор, Дж. Д., Вербила, А. П., Кэвэна, К., Ньюберри, М.: Переменная селек- в линейных смешанных моделях с использованием расширенного класса штрафов. Aust. NZJ Stat. **54** (4), 427–449 (2012).

Тибширани, Р.Дж.: Адаптивная кусочно-полиномиальная оценка через тренд фильтрация. Аня. Стат. **42, №** 1, 285–323 (2014).

Wand, MP: Сглаживающие и смешанные модели. Comput. Стат. **18** (2), 223–249 (2003)

Дерево, SN: быстрая стабильная прямая подгонка и выбор гладкости для генераторов. обобщенные аддитивные модели. JR Stat. Soc. Сер. B (Стат. Метод.) **70** (3), 495–518 (2008).

Вуд, С.Н .: Быстрая стабильная ограниченная максимальная вероятность и предельная оценка правдоподобия полупараметрических обобщенных линейных моделей. JR Stat. Soc. Сер. B (Стат. Метод.) **73** (1), 2–36 (2011)

Вуд, С. Н.: Обобщенные аддитивные модели: введение в R, 2-е место. edn. Chapman & Hall CRC, Лондон (2017)

Вуд, С. Н., Фазиоло, М .: Обобщенный метод Фелльнера-Шалла для оптимизация параметров сглаживания с приложением к Tweedie расположение, масштаб и форма моделей. Биометрия **73** , 1071–1081 (2017)

Вуд, С. Н., Пья, Н., Сефкен, Б.: Параметр сглаживания и выбор модели. для общих гладких моделей. Варенье. Стат. Доц. **111** (516), 1548–1563 (2016)

Цзоу, Х., Ли, Р .: Одноступенчатые разреженные оценки в невогнутой форме со штрафом модели правдоподобия. Аня. Стат. **36** (4), 1509–1533 (2008).

123