

SORT 39 (2) июль-декабрь 2015, 149-186

## Двадцать лет Р-шлицам

Пол Х.К. Эйлерс <sup>1</sup>, Брайан Д. Маркс <sup>2</sup> и Мария Дурбан <sup>3</sup>

### Абстрактный

Р-образные шлицы впервые оказались в центре внимания двадцать лет назад. С тех пор они стали популярными в приложениях и в теоретической работе. Комбинация богатого В-сплайнового базиса и простого дифференциала. Штраф для определения вероятности хорошо поддается множеству обобщений, потому что он основан на регрессии. Фактически, Р-образные шлицы позволяют построить «основу» для «смешивания и согласования» разнообразия. аддитивных компонентов гладкой структуры, предлагая при этом всевозможные расширения: с переменным коэффициентом эффекты, сигнальные (функциональные) регрессоры, двумерные поверхности, ненормальные отклики, квантиль (ожидаемое) моделирование, среди прочего. Сильные связи со смешанными моделями и байесовским анализом. *sis* были установлены. Мы даем обзор многих центральных событий во время первые два десятилетия Р-шлицев.

MSC: 41A15, 41A63, 62G05, 62G07, 62J07, 62J12.

Ключевые слова: В-сплайны, штраф, аддитивная модель, смешанная модель, многомерное сглаживание.

### 1. Введение

Двадцать лет назад *Statistical Science* опубликовала дискуссионный документ под названием «Flex-Возможное сглаживание с помощью В-сплайнов и штрафов» (Эйлерс и Маркс, 1996). Авторы были два статистика с небольшим послужным списком, которые наконец опубликовали рукопись это было отклонено тремя другими журналами. Они пытались с 1992 года продать их детище P-splines (Eilers and Marx, 1992). По-видимому, это имело какую-то ценность, потому что два десятилетия спустя эту статью цитировали более тысячи раз (согласно в *Web of Science*, консервативный источник), как в теоретической, так и в прикладной работе. От Сейчас Р-сплайны стали активной областью исследований, поэтому они будут полезны и, надеюсь, интересно, оглянуться назад и сделать набросок того, что может быть впереди.

<sup>1</sup> Медицинский центр Университета Эразма, Роттердам, Нидерланды, p.eilers@erasmusmc.nl

<sup>2</sup> Департамент экспериментальной статистики, Государственный университет Луизианы, США, bmarx@lsu.edu

<sup>3</sup> Univ. Карлос III Мадрид, Департамент статистики, Легоанес, Испания, mdurban@est-econ.uc3m.es

Поступила: октябрь 2015 г.

P-шлицы упрощают работу О'Салливана (1986). Он заметил, что если мы смоделируем функцию как сумма В-шлицев, знакомая мера шероховатости, интегрированный квадрат Вторая производная может быть выражена как квадратичная функция коэффициентов. P-шлицы делают еще один шаг: они используют В-сплайны с одинаковым интервалом и отбрасывают производную *com*-полностью. Шероховатость выражается как сумма квадратов разностей коэффициентов. Dif-Чрезвычайно легко вычислить, а обобщение до более высоких порядков не вызывает затруднений. вперед.

План работы следующий. В Разделе 2 мы начинаем с описания основных P-сплайны, комбинация базиса В-сплайна и штраф за разность коэффициентов. Штраф - важная часть, и в Разделе 3 мы представляем множество вариантов штрафа. для обеспечения желаемых свойств подогнанных кривых. Штраф настраивается сглаживающей кнопкой. раметр; привлекательно иметь автоматические методы и методы, управляемые данными. Раздел 4 представляет модельную диагностику, которую можно использовать для этой цели, подчеркивая важные роль эффективной размерности модели. Мы представляем основы P-шлицев в контексте штрафных наименьших квадратов и ошибок с нормальным распределением. Для сглаживания ненормальных распределений, легко адаптировать идеи из обобщенных линейных моделей, как это сделано в разделе 5. Здесь мы также устанавливаем связи с GAMLSS (обобщенные аддитивные модели по расположению, масштабу и форме), где не только средства условного распределения смоделированы. Мы увидим, что P-сплайны также привлекательны для квантилей и ожидаемое сглаживание. Первым шагом к множеству измерений является обобщенная дополнительная Реальная модель (раздел 6). Сглаживание можно использовать не только для оценки тенденций в ожидаемой значений (и других статистических данных), но его также можно использовать для нахождения гладких оценок для регрессий. коэффициенты, изменяющиеся со временем, или другая дополнительная переменная. Прототип случай - модель с переменным коэффициентом (VCM). Мы обсуждаем VCM в разделе 7, а также с другими моделями, такими как регрессия сигнала. На современном жаргоне это примеры функций. анализ данных. В разделе 8 мы делаем шаг к полному многомерному сглаживанию, с использованием тензорных произведений В-сплайнов и множественных штрафов. В разделе 9 мы покажем, как все модели из предыдущих разделов могут быть добавлены друг к другу и, таким образом, объединены в одна структура. Здесь снова окупаются корни регресса.

Штраф можно оценить как мощный инструмент. Тем не менее, можно дать это более глубокий смысл. В разделе 10 P-шлицы соединяются со смешанными моделями. Этот приводит к дальнейшим исследованиям, а также к новым алгоритмам поиска разумных значений для параметров штрафа. С точки зрения смешанной модели, это всего лишь небольшой шаг. к байесовскому подходу, интерпретируя штраф как (минус) логарифм предыдущего распределение коэффициентов В-сплайна. Это предмет Раздела 11.

Асимптотикам и бустам нет места в других разделах, поэтому мы положили их вместе в Разделе 12, в то время как вычислительные проблемы и доступность программного обеспечения обсуждается в разделе 13. Мы завершаем статью обсуждением

Насколько нам известно, это первый обзор P-шлицев. Более ранняя работа Ruppert et al. (2009) рассмотрели более широкий взгляд на первые пять лет после выхода их книги. (Ruppert et al., 2003). Мы не пытаемся быть исчерпывающими. Это было бы невозможно (и скучно), учитывая большое количество цитирований. При наличии *Google Scholar*

и коммерческие базы данных цитирования, такие как *Scopus* и *Web of Science*, любой может подробно проследить по следу истории.

Мы добросовестно сделали все, что в наших силах, чтобы дать обзор отрасли, но не

утверждают, что наш выбор статей свободен от субъективности. Появление Р-шлицев привело к значительным достижениям в области сглаживания, и мы активно формировали многие из них. Надеемся, что серьезными упущениями мы не обидим ни одного читателя.

2. Основы Р-сплайна

Двумя компонентами Р-сплайнов являются В-сплайны и дискретные штрафы. В этом разделе мы кратко рассмотрим их, начиная с первого. Мы не вдаемся в технические подробности; об этом см. Eilers and Marx (2010).

2.1. В-шлицы

На рисунке 1 показаны четыре треугольника одинаковой высоты и ширины, средние из которых перекрывают друг друга. со своими двумя соседями. Это линейные В-сплайны, ненулевые части которых состоят из два линейных отрезка. Представьте, что мы масштабируем треугольники на разную величину и добавляем их всех. Это дало бы нам кусочно-линейную кривую. Мы можем создавать множество форм изменяя коэффициенты, и мы можем получить больше или меньше деталей, используя больше или меньше В-шлицы. Если мы указываем треугольники от  $B_j(x)$ , и если  $a_j$  коэффициенты масштабирования, у нас есть  $\sum_{j=1}^n a_j B_j(x)$  в качестве формулы для функции. Это открывает дверь к подгонке данных пары  $(x_i, y_i)$  для  $i = 1, \dots, m$ . Минимизируем сумму квадратов

$$S = \sum_i (y_i - \sum_j a_j B_j(x_i))^2 = \|y - Ba\|^2,$$

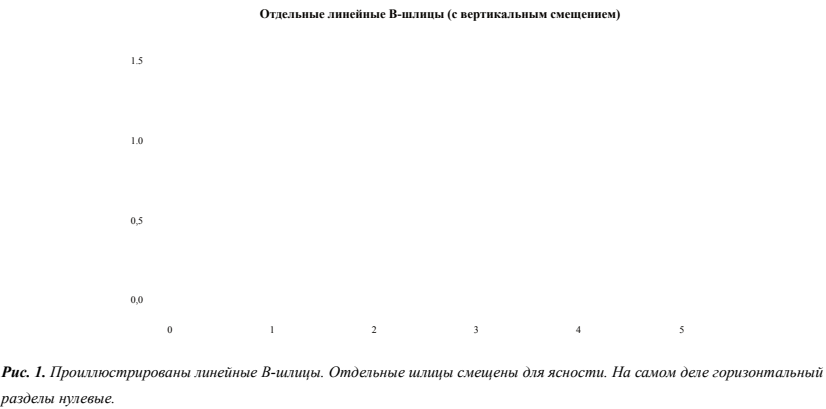


Рис. 1. Проиллюстрированы линейные В-шлицы. Отдельные шлицы смещены для ясности. На самом деле горизонтальный раздели нулевые.



Рисунок 2: Иллюстрированные квадратичные В-шлицы. Для наглядности отдельные шлицы смещены. На самом деле хори-

где  $\mathbf{B} = [b_{ij}]$ , так называемая базисная матрица. Это стандартная задача линейной регрессии и раствор хорошо известно:  $\mathbf{a} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y}$ . Гибкость можно настроить, изменив ширина треугольников (а значит, и их количество).

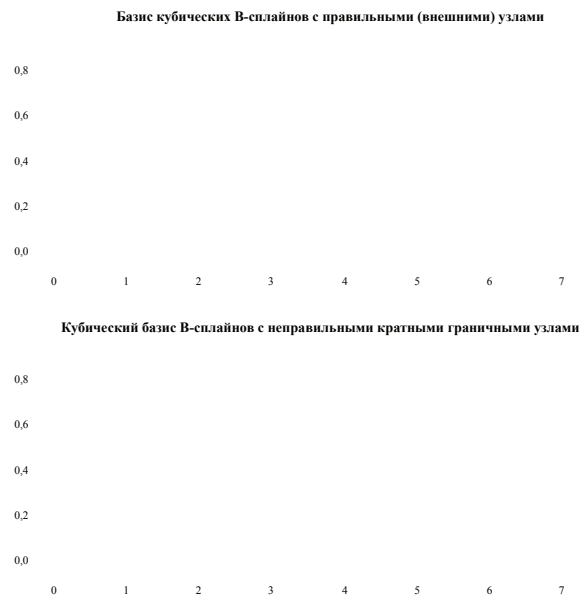
Кусочно-линейная аппроксимация данных может не радовать глаз и не подходить для вычисление производных (которые были бы кусочно-постоянными). На рисунке 2 показаны квадратичные В-шлицы, каждый из которых образован тремя квадратичными сегментами. Сегменты плавно соединяются. В Подобным образом кубические В-шлицы могут быть образованы из четырех кубических сегментов. Рецепт формирования кривой и подгонка коэффициентов к данным остается прежним.

Позиции, в которых соединяются сегменты В-сплайна, называются узлами. В наших иллю- Узлы расположены на одинаковом расстоянии друг от друга, поэтому все В-шлицы имеют одинаковую форму. Это не является обязательным для общих В-шлицев, но, скорее, это осознанный выбор для Р-шлицев, поскольку это упрощает построение штрафов.

При вычислении В-сплайнов следует проявлять осторожность. Верхняя панель рисунка 3 показано основание с равномерно расположенными узлами. Обратите внимание на «неполные» В-шлицы на обоих концы, из которых не все сегменты попадают в область определения  $x$ . На нижней панели показан базис, вычисленный функцией `Rspl`. Имеет так называемые множественные узлы на обоих концах, и поэтому не подходит для Р-шлицев. Во избежание этого следует указывать увеличенный домен, и обрезать шлицы на обоих концах, удалив соответствующие столбцы в базисной матрице. В качестве альтернативы можно использовать код, представленный Eilers и Маркс (2010).

## 2.2. Дискретные штрафы

Имея в основе количество В-сплайнов, мы можем настроить гладкость кривой на данные под рукой. Меньшее количество шлицов дает более плавный результат. Однако это



**Рисунок 3:** Базы В-шлицев с различным выбором узлов. Сверху: равномерно расположенные узлы, подходящая основа для П-образные шлицы. Внизу: несколько узлов на обоих концах области, что является результатом функции `Rspl` и не подходит для Р-шлицев.

это не единственная возможность. Мы также можем использовать большую базу и дополнительно ограничить коэффициенты В-шлицев, чтобы добиться желаемой гладкости. Правильно выбранный штраф достигает этого.

О'Салливану (1986) пришла в голову блестящая идея взять за основу множество В-сплайнов и использовать дискретный штраф. Последнее было получено из интегрального квадрата второго производная кривой. Это был и остается устоявшимся способом измерения шероховатости. кривой  $f(x)$ :

$$R = \int_u^l [f''(x)]^2 dx, \quad (1)$$

где  $l$  и  $u$  обозначают границы области определения  $x$ . Если  $f(x) = \sum_j a_j B_j(x)$ , мы можем получить (полосатую) матрицу  $P$  такую, что  $R = a^T P a$ . Элементы  $P$  вычисляются как интегралы произведений вторых производных соседних В-сплайнов.

О'Салливан предложил минимизировать

$$Q = S + \lambda R = S + \lambda a^T P a = \|y - a\|^2 + \lambda a^T P a, \quad (2)$$

где  $\lambda$  - параметр, определяющий влияние штрафа. Чем больше  $\lambda$ , тем сгладить результат. В пределе вторая производная вынуждена быть очень близкой к нулю. и получится прямая подгонка. Обратите внимание, что нам нужно вычислить  $P$  только один раз. Система будет решенным

$$(B^T B + \lambda P) a = B^T y. \quad (3)$$

Вычисление  $P$  нетривиально и становится довольно утомительным, когда третье или Производная четвертого порядка используется для измерения шероховатости. Палочка и Ормерод (2008) имеют расширил идею О'Салливана на более высокие порядки производной. Они использовали компьютер система алгебры для построения таблицы формул. Р-образные шлицы позволяют избежать этой проблемы. пинга производных и интегралов полностью. Вместо этого они используют дискретную матрицу штрафов. от начала. Его также просто вычислить, поскольку он основан на формулах разности. Позволять  $A a_j = a_j - a_{j-1}$ ,  $A^2 a_j = A(A a_j) = a_j - 2a_{j-1} + a_{j-2}$  и вообще  $A^d a_j = A(A^{d-1} a_j)$ . Пусть  $D_d$  - матрица такая, что  $D_d a = A^d a$ . Если заменить штраф на  $\lambda \|D_d a\|^2 = \lambda a^T D_d^T D_d a = \lambda a^T P a$ , мы получаем конструкцию, аналогичную конструкции О'Салливана, но с минимальным объемом работы. В современных языках, таких как R и Matlab,  $D_d$  можно получить механически. как разность  $d$ -го порядка единичной матрицы.

Удивительно, что от упрощенного штрафа ничего не теряется. Эйлерс и Маркс (1996) показали, сколько полезных свойств можно доказать с помощью нескольких строк простых математика. Ванд и Ормерод (2008) мотивируют свою работу, утверждая, что экстрапо-соединение Р-шлицами идет не так. Они рекомендовали свои «О-образные шлицы» как лучшую альтернативу. родной; см. также (Ruppert et al., 2009). В Приложении А мы представляем небольшое исследование, в котором серьезные сомнения по поводу их заключения.

### 2.3. Сила штрафа

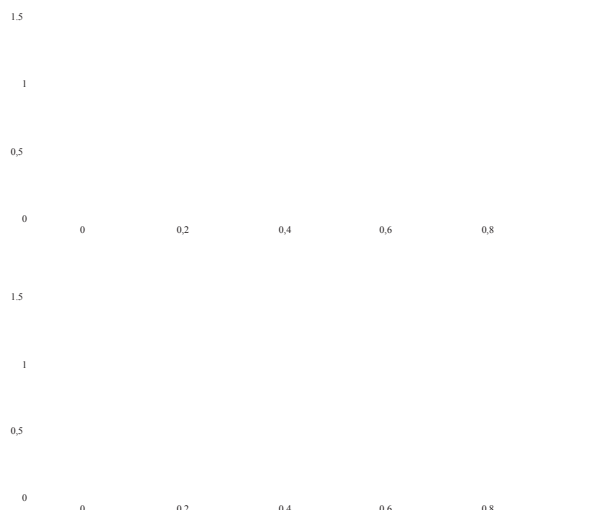
Полезный способ взглянуть на Р-сплайны - это придать коэффициентам центральное положение в виде скелет, с В-образными шлицами, просто накладывая «плоть на кости». Это проиллюстрировано на рисунке 4. Более плавная последовательность коэффициентов приводит к более плавной кривой. Номер шлицев и коэффициентов не имеет значения, если последние гладкие. Роль наказание - сделать это.

Штраф упрощает интерполяцию (Currie et al., 2004; Eilers and Marx, 2010). В позиции, в которых требуются интерполированные значения, вводятся псевдонаблюдения

с  $y = 0$  (или любым произвольным числом) и нулевым весом и решает систему. Верно наблюдения получают вес 1. Один решает

$$(B^T W B + X^T P) = B^T Y a i, \quad (4)$$

где  $W$  - диагональная матрица с весами на диагонали. Плавная интерполяция происходит автоматически. Таким же образом можно осуществить экстраполяцию, вводя вывод псевдонаблюдений за пределы области данных.



**Рисунок 4:** Иллюстрация роли штрафа. Количество В-шлицев одинаково на обеих панелях. в

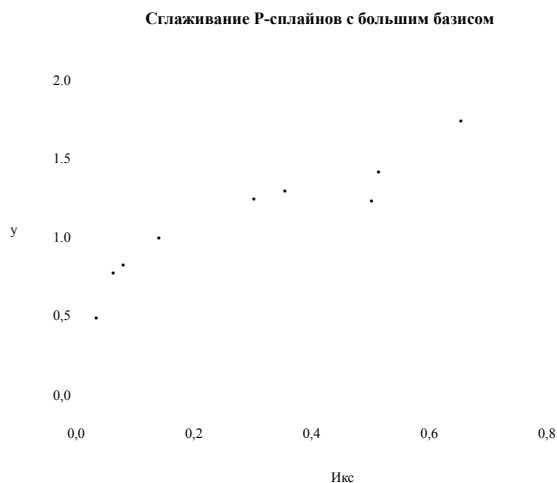
На верхней панели соответствие данным (квадраты) более неравномерно, чем на нижней панели, потому что штраф послабее там. Заштрихованные кружки показывают коэффициенты В-сплайнов. Из-за более сильного штрафа они сформируют плавную последовательность на нижней панели, что приведет к более плавной подгонке кривой.

Количество В-сплайнов может быть (намного) больше, чем количество наблюдений. В штраф делает процедуру примерки хорошо обусловленной. Это следует понимать буквально: даже тысяча сплайнов без проблем уместит десять наблюдений. Такова сила штрафа. Рисунок 5 иллюстрирует это для смоделированных данных. Есть 10 точек данных и 40 (+3) кубических В-шлицев. К сожалению, это свойство Р-сплайнов (и других типов штрафные шлицы) обычно не приветствуются. Но не может быть слишком много В-шлицы. Мудрый выбор - использовать 100 из них, если вычислительные ограничения (в большие модели).

Мы вернемся к этому примеру в разделе 4 после введения эффективной модели размер, а также рассмотрим этот вопрос о многих шлицах в Приложении В.

#### 2.4. Исторические заметки

Название Р-шлицы было придумано Эйлерсом и Марксом (1996) для обозначения комбинации В-сплайнов и штрафа за дискретную разность. Его не всегда использовали с этим конкретное значение. Рупперт и Кэрролл (2000) опубликовали статью о сглаживании, которая также использовали идею богатой базы и дискретного штрафа. Их основа состоит из усеченных степенные функции (TPF), узлы являются квантилями  $x$ , а штраф зависит от размера



**Рисунок 5:** Сглаживание Р-сплайном 10 (смоделированных) точек данных с помощью 43 кубических В-сплайнов.

коэффициенты. Эта работа была расширена в книге Ruppert et al. (2003). Некоторый люди также назвали подход TRF Р-шлицами. Это сбивает с толку и прискорбно потому что TRF уступают исходным Р-шлицам; Эйлерс и Маркс (2010) задокументированы их плохое числовое состояние.

В-сплайны и TRF сильно связаны (Greven, 2008; Eilers and Marx, 2010). Акту-союзные В-сплайны могут быть вычислены как разности TRF, но в эпоху одинарной точности чисел с плавающей запятой этого избегали из-за опасений больших ошибок округления. Эйлерс и Маркс (2010) показали, что этого больше нет. Р-шлицы позволяют выбрать степень В-шлицы и порядок штрафа самостоятельно. С TRF выбора нет: они подразумевают штраф за разницу, порядок которого определяется степенью TRF.

### 3. Варианты наказания

Стандартные Р-шлицы используют штраф, основанный на повторяющихся различиях. Множество вариаций возможны. Как уже говорилось, коэффициенты В-сплайна образуют основу подгонки, поэтому, если мы можем найти другие полезные дискретные штрафы, тогда мы сможем получить кривые с множеством желаемых характеристики. Эйлерс и Маркс (2010) назвали их «дизайнерскими штрафами», и они представили несколько примеров. Мы приводим здесь резюме:

- Круговой штраф соединяет первый и последний элементы вектора коэффициентов, используя различия, чтобы оба конца соединялись плавно. В сочетании с циркулярной

Основа В-сплайна, это правильный инструмент для подгонки периодических данных или круговых наблюдений, как направления.

- При разностях второго порядка  $a_{j-2} - 2a_{j-1} + a_j$ , в штрафной приближается подгонка прямой линии при увеличении  $\lambda$ . Если мы изменим уравнение на  $a_{j-2} - 2\phi a_{j-1} + a_j$  пределом является (ко) синус с периодом  $p$  такой, что  $\phi = \cos(2\pi/p)$ . Фаза (ко) синуса регулируется автоматически, чтобы минимизировать сумму квадратов остатков. Для сглаживания (и интерполяции) сезонных данных (с известным периодом) этот гармонический штраф обычно более привлекателен, чем стандартный.
- Эйлерс и Гоеман (2004) объединили штрафы первой и второй степени, чтобы исключить отрицательные боковые лепестки импульсной характеристики (как в случае только штрафа второго порядка). Это гарантирует, что сглаживание положительных данных никогда не сможет приводить к отрицательным подобранным значениям.
- Как описано, штраф Р-сплайна является квадратичным: он использует норму суммы квадратов. Этот приводит к гладкому результату. Использовались другие нормы. Сумма абсолютных значений (норма  $L_1$ ) разностей первого порядка допускает скачки (Эйлерса и де Менезеса, 2005) между соседними коэффициентами, что делает его пригодным для кусочно-постоянной сглаживания. Эта норма является естественным выбором в сочетании с нормой  $L_1$  на остатки; может использоваться стандартное программное обеспечение для линейного программирования. См. Также Раздел 5. для квантильного сглаживания.
- Скачки, полученные с нормой  $L_1$ , на самом деле не «четкие», но слегка округлый. Причина в том, что норма  $L_1$  выделяется и сжимается. Намного лучшие результаты получены с нормой  $L_0$ , числом ненулевых коэффициентов (Риппе и др., 2012b). Хотя в результате получается невыпуклая целевая функция, на практике она может быть надежно и быстро оптимизируется с помощью итеративно обновляемого квадратичного штрафа.

Другие типы штрафов могут использоваться для обеспечения соблюдения ограничений формы. Примером является монотонно возрастающая кривая (Bollaerts et al., 2006). Второй, асимметричный, вводится штраф  $\kappa \sum_j v_j (A a_j)$ , причем  $v_j = 1$ , когда  $A a_j < 0$ , и  $v_j = 0$  в противном случае. Значение  $\kappa$  регулирует влияние штрафа. Требуются итерационные вычисления, поскольку нужно знать  $v$ , чтобы выполнить сглаживание, а затем знать решение для определения (обновление)  $v$ . На практике, начиная с  $v = 0$  работает хорошо.

Возможны многие вариации, чтобы навязать ограничения по знаку, чтобы гарантировать (увеличение или уменьшение складка) монотонность, либо требовать выпуклой или вогнутой формы. Также можно смешивать и сопоставлять асимметричные штрафы, чтобы реализовать несколько ограничений формы. Эйлерс (2005) использовали их для одномодального сглаживания, в то время как Эйлерс и Боргдорф (2007) использовали их для соответствия смеси логарифмически вогнутых непараметрических плотностей. Эта схема была расширена на два измерения Риппе и др. (2012а) и применительно к генотипированию SNP (мы обсуждаем многомерное сглаживание в разделе 8).

Пия и Вуд (2015) использовали другой подход. Они пишут  $\mathbf{a} = \Sigma \exp(\beta)$  и строят преобразовать матрицу  $\Sigma$  так, чтобы  $\mathbf{a}$  имела желаемую форму для любого вектора  $\beta$ . Для

пример  $\Sigma_{ij} = I(i \geq j)$  с индикаторной функцией  $I(\cdot)$  обеспечивает монотонный рост функция  $\text{ing}$ . Шаблоны для комбинаций ограничений на первую и вторую производную: приведены в таблице в их статье.



#### 4. Диагностика

В отличие от многих других сглаживателей, таких как ядра, локальное правдоподобие и вейвлеты, Р-сплайны используют регрессионную модель с четко определенными коэффициентами. Следовательно, мы можем заимствовать от теории регрессии к вычислению информативных свойств модели. Что мы не делаем `leap` - это выбор хорошего значения для параметра штрафа  $\lambda$ . Только классическая теория учитывает соответствие модели данным и как таковой бесполезен для этой цели. Вместо нам нужно измерить эффективность прогнозирования. В этом разделе мы рассмотрим стандартные ошибки, перекрестная проверка, эффективное измерение и AIC.

Ковариационная матрица коэффициентов сплайна (при фиксированном  $\lambda$ ) имеет вид

$$C_a = \sigma^2 (B^T W B + \lambda D^T D)^{-1}, \quad (5)$$

где  $\sigma^2$  - дисперсия шума наблюдения  $q$  в модели  $y = Ba + q$ . Ковариационная матрица  $C_a$  содержит вычисленный базис В-сплайнов. Соответствие подобранных значений следует проверять, как и в `leap`, при любом выбранном наборе значений  $x$ .

В существующем виде этот  $C_a$  не очень полезен, потому что нам нужно знать  $\sigma$ . Возможно оценивается по невязкам, но для этого нам нужно будет выбрать правильное значение  $\lambda$ , ведущие к надлежащим «степеням свободы».

Перекрестная проверка без исключения (CV) обеспечивает механизм для определения прогнозируемых значений. Эффективность Р-сплайновой модели при любом значении  $\lambda$ . Пусть одно наблюдение  $y_i$  будет опущено и пусть предсказанное значение будет обозначено  $y_{-i}$ . Выполняя это для каждого наблюдения по очереди, мы можем вычислить ошибку предсказания

$$CV = \sqrt{\sum_j (y_j - y_{-j})^2}. \quad (6)$$

Таким образом, CV является естественным критерием выбора  $\lambda$  за счет его минимизации. Используя брут силы, вычисление CV обходится дорого, конечно, когда количество наблюдений большой. К счастью, есть точный ярлык. У нас есть это

$$y = B (B^T W B + \lambda D^T D)^{-1} B^T W y = Hy. \quad (7)$$

Обычно  $H$  называется «шляпной» матрицей. Можно доказать, что

$$y_j - y_{j-1} = (y_j - y_j) / (1 - \eta_j), \quad (8)$$

и диагональ  $H$  может быть вычислена быстро. Вывод можно найти в Приложении Б. Майерс (1989). Неофициальное доказательство выглядит следующим образом. Представьте, что мы меняем элемент  $i$  of  $y$ , чтобы получить новый вектор  $y^*$ ; то  $y^* = Hy^*$ . Теперь считаем, что если мы положим  $y^*_{-i} = y_{-i}$ , то  $y^*_{-i} = y_{-i}$ . Следовательно, мы имеем, что  $y_{-i} - y_j = H_{ii} (y_{-i} - y_j)$ , а  $y_{-i} = y_j$ . После добавления  $y_i - y_i$  в правую часть этого уравнения и переставляя члены, приходим к (8).

Матрица шляпы также дает нам эффективное измерение модели, если следовать Ye (1998): кто предложил

$$ED = \sum_j dy_j / dy_j = \sum \eta_j. \quad (9)$$

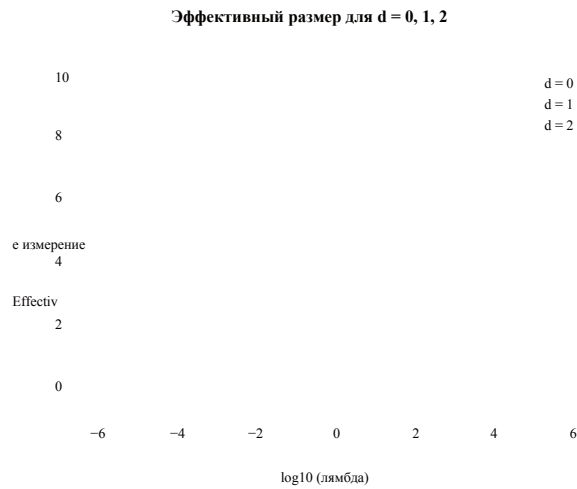
Фактически, мы можем вычислить след  $H$ , фактически не вычисляя его диагональ, используя циклическая перестановка:

$$ED = \text{tr} (H) = \text{tr} [(B^T W B + \lambda D^T D)^{-1} B^T W B]. \quad (10)$$

Дальнейшее упрощение возможно, если отметить, что

$$(B^T W B + P)^{-1} B^T W B = (B^T W B + P)^{-1} (B^T W B + P - P) = I - (B^T W B + P)^{-1} P, \quad (11)$$

где  $P = \lambda D^T D$ . Харвилл (1977) представил аналогичный результат для смешанных моделей. В случае Р-сплайнов выражение очень простое, потому что здесь нет фиксированных эффектов.



**Рисунок 6:** Изменения в эффективном размере модели для сглаживания Р-сплайном 10 (смоделированных) точек данных с 43 кубическими В-шлицами, для разного порядка ( $d$ ) различий в штрафе.

Эффективный размер - отличный способ количественно оценить сложность Р-сплайна. модель. Он суммирует комбинированные влияния размера основы В-сплайна, порядка штрафа и значение параметра сглаживания. Последнее уравнение в (11) аккуратно показывает, что эффективный размер всегда будет меньше  $n$ . На самом деле эффективный размер всегда меньше  $\min(m, n)$ . Иллюстрация представлена на рисунке 6, показывает, как ED изменяется с  $\lambda$  для примера с 10 наблюдениями и 43 В-сплайнами на рисунке 5. При малых  $\lambda$  величина ED приближается к  $m$ , в то время как при больших  $\lambda$  она приближается к  $d$ , порядок различий.

Тот факт, что  $ED < n$ , очевидно из размера системы штрафных вероятностей. уравнения. Эвристический аргумент для  $ED < m$  состоит в том, что  $B(B^T B + \lambda D^T D)^{-1} B^T$  - это  $m$  на  $m$  матрица. Это шляпная матрица со следом меньше  $m$ . Формальное доказательство дано в Приложение Б.

Кроме того, тот факт, что  $ED < m$ , объясняет, почему сглаживание (намного больше) В-сплайнов, чем наблюдения, работает без проблем при любом значении  $\lambda$ . По нашему опыту, многие коллеги не осознают этого факта. Может быть, они боятся сингулярностей и придерживаются маленьких количество базисных функций.

Чтобы оценить  $\sigma^2$ , нужно разделить сумму квадратов остатков на их эффективные степени свободы, то есть количество наблюдений минус эффективная модель. Размер:  $\sigma^2 = \sum (y_j - \hat{y}_j)^2 / (M - ED)$ .

В качестве альтернативы можно использовать информационный критерий Акаике для выбора  $\lambda$ , где  $AIC = -2\ell + 2ED$ , а  $\ell$  - логарифмическая вероятность. Прелесть этой формулы в том, что она показывает баланс между точностью данных и сложностью модели.

Всегда следует быть осторожным при использовании перекрестной проверки или AIC для настройки плавного параметр  $\text{ing}$ . Неявное предположение состоит в том, что наблюдения независимы, при условии на их безмятежные ожидания. Если это не так, как для данных временных рядов,

Серийная корреляция будет рассматриваться как часть плавной составляющей и сильного недо- может произойти сглаживание. Один из способов решения этой проблемы - явное моделирование кор- структура отношения шума. Мы вернемся к этой теме в разделе 10, посвященном смешанным моделям. Недавней альтернативной стратегией является адаптация L-кривой (Hansen, 1992). Это было де- разработан для регрессии гребня, но может быть адаптирован к разнице штрафов. См. Фрассо и Eilers (2015) приведены примеры и вариант, называемый V-образной кривой, который проще в использовании.

В разделе 10 параметр настройки штрафа будет отображаться как отношение дисперсий, и эффективный размер играет важную роль при их оценке.

## 5. Обобщенное линейное сглаживание и расширения.

P-шлицы основаны на линейной регрессии, поэтому их удлинение для сглаживания является обычным делом. в ненормальных наблюдениях, заимствуя структуру обобщенных линейных моделей (GLM). Пусть  $y$  наблюдается, а  $\mu$  - вектор ожидаемых значений. Тогда линейный предсказатель  $\eta = g(\mu) = \mathbf{B}\mathbf{a}$  моделируется B-сплайнами, и для  $y$  выбирается подходящее распределение, учитывая  $\mu$ . Штраф вычитается из логарифма правдоподобия:  $\ell^* = \ell - \lambda \mathbf{D}^T \mathbf{D} / 2$ . Штраф-

Уравнения упорядоченного правдоподобия приводят к  $\mathbf{B}^T (y - \mu) = \lambda \mathbf{D}^T \mathbf{D} \mathbf{a}$ . Это небольшое изменение по сравнению с стандартный GLM, в котором правая часть равна нулю (Dobson, Barnett, 2008).

Уравнения нелинейны, но штраф за максимальное правдоподобие приводит к итеративному решению

$$\mathbf{a}_{m+1} = (\mathbf{B}^T \mathbf{W}_m \mathbf{B} + \lambda \mathbf{D}^T \mathbf{D})^{-1} \mathbf{B}^T \mathbf{W}_m \mathbf{z}_m \quad \text{с } \mathbf{z}_m = \eta_m + \mathbf{B} \mathbf{a}_{m-1} \quad (\mathbf{y} - \hat{\mu}_t), \quad (12)$$

где  $m$  обозначает итерацию и  $\mathbf{W}_m$  и  $\eta_m$  обозначают приближенные решения, а  $\mathbf{z}_m$  является так называемая рабочая переменная. Веса в диагональную матрицу  $\mathbf{W}$  зависят от функция связи и выбранное распределение. Например, распределение Пуассона с  $\eta = \log(\mu)$  имеет  $w = 1/\mu$ .

Мощным приложением обобщенного линейного сглаживания с помощью P-сплайнов является плотность оценка (Эйлерс и Маркс, 1996). Гистограмма с узкими интервалами вычисляется, и счетчики сглаживаются с использованием распределения Пуассона и функции логарифмической связи. Нет опасности, что бункеры будут выбраны слишком узкими: даже если большинство из них содержат только немногочисленные или нулевые хорошие результаты получаются. Величина сглаживания настраивается AIC.

Важно (для любого устройства сглаживания), чтобы достаточное количество бункеров с нулевым счетчиком было включено в концы наблюдаемой области данных, если не известно, что она ограничена (как для внутренне положительные переменные).

P-сплайны сохраняют моменты распределений до порядка  $d-1$ , где  $d$  - порядок о различиях в штрафах. Это означает, что если  $d=3$ , сумма, среднее и дисперсия гладкой гистограммы равна дисперсии необработанной гистограммы, независимо от количество сглаживания (Эйлерс и Маркс, 1996). Напротив, сглаживание ядра увеличивает дисперсия.

Опубликовано множество вариаций на эту тему. Мы уже упоминали одно- и двумерные логарифмически вогнутые плотности в разделе 3. Kauermann et al. (2013) исследовал оценка плотности гибкой связки. Они моделировали плотность непосредственно как сумму тензорных произведения линейных B-сплайнов (тензорные произведения обсуждаются в разделе 8). Чтобы уменьшить количество коэффициентов, они использовали уменьшенные сплайны, которые похожи на вложенные B-сплайны (Ли и др., 2013).

Другой вариант - не моделировать логарифм отсчетов суммой B-сплайнов, а скорее сама плотность с ограничениями на коэффициенты (Шеллхаз и Кауэр-Манн, 2012).

Сглаживание смертности или заболеваемости эквивалентно дискретной оценке плотности с компенсация за выдержки. P-шлицы нашли свое применение в этой области, как для одно-, так и для двумерных

таблицы размеров (Currie et al., 2004; Camarda, 2012); оба документа иллюстрируют автоматическое экстраполяция.

Палитра распределений, которые может использовать обобщенное линейное сглаживание, ограничена. А очень общий подход предлагает GAMLSS: обобщенные аддитивные модели для местоположения, масштаб и форма (Rigby, Stasinopoulos, 2005). Примером может служить нормальное распределение с плавно изменяющимся средним значением и дисперсией в сочетании с (изменяющимся) преобразованием Бокса-Кокса.

форма переменной ответа. Можно подобрать множество непрерывных и дискретных распределений. алгоритмом GAMLSS, а также в сочетании со смесями, цензурированием и случайным составные части.

Вместо использования параметрического распределения можно оценить гладкую условную количественную оценку. плитки, минимизируя асимметрично взвешенную сумму абсолютных значений остатков. Bollaerts et al. (2006) объединили это с ограничениями формы, чтобы усилить монотонность. К избегать пересечения гладких квантильных кривых, рассчитанных индивидуально, Шнабель и Эйлерс (2013) представили квантильный лист, поверхность в области, образованную пояснительным переменной и уровень вероятности.

По сравнению с явными решениями задач наименьших квадратов (со штрафом, взвешиванием), квантильное сглаживание немного менее привлекательно для числовой работы, поскольку оно приводит к линейному программированию или квадратичному программированию, если задействованы квадратичные штрафы. В отличие, ожидаемые значения используют асимметрично взвешенные суммы квадратов и приводят к простым итеративным алгоритмы (Schnabel, Eilers, 2009). Сobotка и Кнейб (2012) расширенный прогноз сглаживание к пространственному контексту, в то время как Сobotка и др. (2013) обеспечивают уверенность в-интервалы. Шнабель и Эйлерс (2013) предложили масштабную модель для пересечения ожидаемые кривые.

При анализе отсчетов с помощью обобщенной линейной модели часто используется распределение Пуассона. Предполагается, что  $\mu = \exp(\eta)$  для ожидаемых значений. Когда счетчики сгруппированы или В совокупности модель составных звеньев (CLM) Томпсона и Бейкера (1981) более соответствующий. В нем говорится, что  $\mu = C \exp(\eta)$ , где матрица  $C$  кодирует агрегацию или шаблон смешивания. В штрафной CLM гладкая структура для  $\eta$  моделируется с помощью P-шицы (Эйлерс, 2007). Это мощная модель для групповых подсчетов (Ламберт и Эйлерс, 2009; Ламберт, 2011; Rizzi et al., 2015), но он также нашел применение в неправильной классификации. выбор и предпочтение цифр (Camarda et al., 2008; Azmon et al., 2014). de Rooi et al. (2014) использовали его для удаления артефактов при сканировании дифракции рентгеновских лучей.

## 6. Обобщенные аддитивные модели.

Обобщенная аддитивная модель (GAM) строит линейный предиктор как сумму гладкие условия, каждый из которых основан на своей ковариате (Hastie and Tibshirani, 1990). В модель  $\eta = \sum_j f_j(x_j)$ ; его можно интерпретировать как многомерное сглаживание без взаимодействия.

GAM с P-сплайнами, или P-GAM, был предложен Марксом и Эйлерсом (1998). Мы проиллюстрировать основную идею в двух измерениях. Позволять

$$\eta = f_1(x_1) + f_2(x_2) = [B_1 | B_2] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = Ba. \quad (13)$$

Объединив две базы в одну матрицу и сцепив коэффициенты в одном векторе мы вернулись к стандартным настройкам регрессии.

Штраф за шероховатость  $\lambda_1 \|D_1 a_1\|_2$  и  $\lambda_2 \|D_2 a_2\|_2$  (где здесь индексы пере-  
к переменным, а не к порядку разностей), что приводит к двум матрицам штрафов

$P_1 = \lambda_1 D_1^T D_1$  и  $P_2 = \lambda_2 D_2^T D_2$ , которые можно объединить в блочно-диагональную матрицу  $P$ . В результате получаются уравнения правдоподобия со штрафом:  $B^T (y - \mu) = Pa$ , которые имеют в точности тот же вид, что и для обобщенного линейного сглаживания. Уравнение взвешенной регрессии все последуют немедленно. То же верно и для ковариационной матрицы оцененных коэффициенты, перекрестная проверка и эффективный размер.

Первоначально для GAM использовалась переоборудование (Hastie and Tibshirani, 1990), обновив каждый компонент функционирует по очереди, используя любой тип сглаживания. Схождение может быть медленным а диагностику получить сложно или невозможно. Прямая установка P-образными шлицами не имеют эти недостатки.

Представленная модель не поддается идентификации, поскольку произвольный сдвиг  $f_1(x_1)$  вверх может быть скомпенсирован равным смещением вниз  $f_2(x_2)$ . Решение состоит в том, чтобы ввести (без штрафных санкций) перехватить и ограничить каждый компонент нулевым средним значением.

P-GAM имеет несколько параметров сглаживания, поэтому оптимизация AIC, скажем, простой поиск по сетке требует много работы. Neim et al. (2007) предложили стратегию поиска который циклически перебирает поиск по одномерной сетке. В качестве более принципиального подхода Вуд (2004) представили алгоритмы численной оптимизации при перекрестной проверке. Его книга (Wood, 2006a) содержит обширный материал о GAM. См. Также Раздел 13 для Программное обеспечение mgcv.

В разделе 10 мы представим алгоритм Шалла для оценки дисперсии. Это привлекает-Инструмент для настройки нескольких параметров штрафа.

## 7. Коэффициенты гладкой регрессии.

В предыдущих разделах P-сплайны использовались для моделирования ожидаемых значений наблюдений. Существует еще один класс моделей, цель которых - моделировать коэффициенты регрессии как кривая или поверхность. В этом разделе мы обсуждаем различные модели коэффициентов (Hastie и Tibshirani, 1993), регрессия сигнала со штрафными санкциями (Marx and Eilers, 1999) и обобщения. На современном жаргоне это все случаи функционального анализа данных (Рамзи и Сильверман, 2003).

Модели с переменными коэффициентами (VCM) были впервые представлены Хасти и Тибширани. (1993). Они позволяют коэффициентам регрессии взаимодействовать с другой переменной, изменяя плавно. Простейшая форма -  $E[y(t)] = \mu(t) = \beta(t)x(t)$ , где  $y$  и  $x$  наблюдаются и  $\beta$  необходимо оценить и заставить медленно изменяться с  $t$ . Модель предполагает, что  $y$  пропорционален  $x$  с различным наклоном линии регрессии. Если мы введем B-сплайн-базис  $B$  и запись  $\beta = Ba$ , получаем  $\mu = XBa$ , где  $X = \text{diag}(x)$ . С разницей штраф на  $a$  у нас есть знакомая структура P-сплайна, только с измененным базисом  $XB$ . А можно добавить переменное смещение:  $E[y(t)] = \mu(t) = \beta(t)x(t) + \beta_0(t)$ . Это имеет форму аддитивная модель. Строя  $\beta_0$  с помощью P-сплайнов, мы фактически получаем P-GAM.

Этот простой VCM можно расширить, добавив больше аддитивных или переменных коэффициентов. термины. Для ненормальных данных мы моделируем линейный предиктор и выбираем правильный ответ. распределение.

VCM с P-шлицами были предложены Эйлерсом и Марксом (2002). Лу и др. (2008) изучил их тоже и представил процедуру Ньютона-Рафсона, чтобы минимизировать перекрестные Ошибка проверки. Андрияна и др. (2014) представили квантильную регрессию в VCM, используя П-образные шлицы. Кауэрманн (2005b) и Кауэрманн и Хомски (2006) разработали P-сплайн модели выживания и опасности, соответственно, с учетом переменных коэффициентов. Ван и другие. (2014) использовали VCM для продольных данных (с ошибками в переменных) с байесовским анализом. П-образные шлицы. Heim et al. (2007) использовали 3D VCM для визуализации мозга.

Модели модуляции для сезонных данных представляют собой интересное приложение VCM (Eilers et al., 2008; Маркс и др., 2010). Амплитуда синусоидальной (или более сложной) форма волны медленно меняется с течением времени. Это предполагает, что период известен. Если это не так, или когда оно не является постоянным, можно оценить оба изменения амплитуда и фаза синусоидальной волны (Eilers, 2009).

В VCM  $y$  и  $x$  - параллельные векторы, заданные в одних и тех же позициях выборки во времени или космос. В регрессии сигнала со штрафом (PSR) у нас есть набор векторов  $x$  и соответствующие скаляры по  $y$ , и цель состоит в том, чтобы предсказать последнее. Если векторы  $x$  образуют строки матрицы  $X$  имеем линейную регрессию  $E(y) = \mu = X\beta$ . Проблема некорректна, потому что  $X$  имеет гораздо больше столбцов, чем строк. Возьмем, к примеру, измеренные оптические спектры. гарантированный многими сотнями длин волн. Элементы  $y$  - известные концентрации вещества. Поскольку столбцы  $X$  упорядочены, имеет смысл заставить  $\beta$  быть гладким, накладывая на него штраф за разницу, тем самым делая задачу хорошо поставленной (Маркс и Эйлерс, 1999).

В принципе, нет необходимости вводить P-сплайны, написав  $\beta = Ba$  и положив штраф на  $a$ , но он снижает вычислительную нагрузку, когда  $X$  имеет много столбцов. Ef- Фактически мы получаем штрафную регрессию на основе  $U = XB$ . После этого шага машина для перекрестной проверки становятся доступны стандартные ошибки и эффективный размер. Уведомление что  $a$  вынуждено быть гладким, но  $\mu$  вовсе не обязательно должно быть гладким. Также не ряды из  $X$  сглаживаются, но коэффициенты регрессии.

Ли и Маркс (2008) предложили усиление сигнала для улучшения внешнего предсказания за счет включая веса PLS.

Обширный обзор функциональной регрессии был представлен Моррисом (2015).

Стандартная модель PSR неявно предполагает функцию связи идентичности. Однако это можно согнуть через  $\mu = f(X\beta) = f(Ua)$ , где  $f(\cdot)$  неизвестно. Мы называем эту модель регрессия одноиндексного сигнала (SISR), которая тесно связана с преследованием проекции (Eilers et al., 2009). Для оценки  $f$  используется второй базис B-сплайнов и соответствующие коэффициенты вводятся. Домен является то, что  $Ua$  и должен быть стандартизированы (например, средний ноль и дисперсия 1), чтобы сделать модель идентифицируемой. Для заданных коэффициентов производная функции  $f(Ua)$  можно вычислить и вставить в разложение Тейлора. Используя это,  $a$  и коэффициенты для  $f$  обновляются по очереди до сходимости.

P-сплайны были реализованы в других типах одноиндексных моделей, например, см. Yu. и Ruppert (2002) и Lu et al. (2006). Leitenstorfer и Tutz (2011) использовали усиление и Антониадис и др. (2004) использовали байесовский подход.

В следующем разделе мы рассмотрим основы тензорного произведения, которые позволяют PSR ex-напряжения в двух измерениях. Например, Эйлерс и Маркс (2003) и Маркс и Эйлерс

Эрс (2005) расширил PSR, чтобы разрешить взаимодействие с дискретной переменной и двумя размерный случай, когда каждый  $x$  не вектор, а матрица. В этих моделях есть нет понятия времени. Когда каждый элемент  $y$  не является скаляром, а является временным рядом, как и  $x$ , историческая функциональная линейная модель (HFLM) предполагает, что в принципе все предыдущие  $x$  может влиять на элементы  $y$  (Malfait and Ramsay, 2003). Harezlak et al. (2007) представила Р-шлицевую технологию для HFLM.

Зеркальное отображение HFLM - это структура процентных ставок, оценивающая ожидаемые будущий курс процентных ставок; см. Jagow et al. (2004) и Кривобокова и др. (2006).

Кроме того, Маркс и др. (2011) расширил SISR до двух измерений, тогда как Маркс (2015) представили гибридную одноиндексную модель с переменным коэффициентом. В SISR взвешенный сумма  $x(t)$  формируется и преобразуется. McLean et al. (2014) пошли еще дальше:  $E(y_i) = \mu_i = \int F(x_i(t), t) dt$ . Это можно интерпретировать как первое преобразование  $x$  (с разным отдельную функцию для каждого  $t$ ), а затем добавление результатов или, наоборот, «преобразование и сложение» «добавлять и преобразовывать».

## 8. Многомерное сглаживание.

Естественным распространением Р-сплайнов на более высокие измерения является формирование тензорных произведений одно-размерных оснований В-шлицев и для добавления штрафов за разницу по каждому измерению (Eilers и Маркс, 2003). Рисунок 7 иллюстрирует эту идею, показывая одно тензорное произведение  $T_{jk}(x, y) = B_j(x) \cdot B_k(y)$ . Фиг.8 иллюстрирует «прореженный» участок основы тензорного произведения; для ясности показаны не все тензорные произведения. Матрица коэффициентов определяет высоту каждая «гора»:  $A = [a_{kl}]$ ,  $k = 1, \dots, n$  и  $l = 1, \dots, \sim n$ . Ситуация полностью аналогичен рисунку 4, но расширен до двух измерений. Шероховатость элементов of  $A$  определяет, насколько гладкой будет поверхность. Чтобы настроить шероховатость, каждый столбец и каждая строка  $A$  наказываются.

Можно выбрать использование одного параметра штрафа для обоих направлений (изотропный гладкий ing) или отдельными (анизотропное сглаживание). В последнем случае оптимизация суммы сглаживания порождает гораздо больше работы. Многие полезные свойства одномерного Р-шлицы переносятся на более высокие размеры. Взвешивание (отсутствующих) наблюдений и хорошо работают интерполяция и экстраполяция. Эффективный размер модели и быстрая перекрестная проверка доступна. Их также можно использовать в качестве строительного блока в гладких конструкциях. (см. следующий раздел).

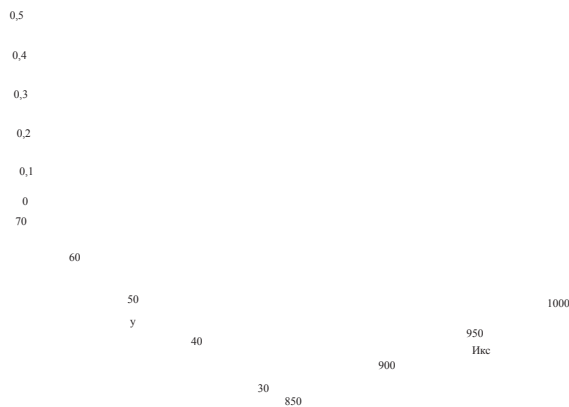
Технически многомерные Р-шлицы сложны. Главный вопрос в том, чтобы чтобы оценить  $A$  с помощью обычных матрично-векторных операций, нам нужно записать его как

вектор и поместить тензорные произведения в соответствующую базисную матрицу. При тщательной организации вычислений это можно решить элегантно (Eilers et al., 2006).

0,5  
0,4  
0,3  
0,2  
0,1  
0  
10



**Рисунок 7:** Строительный блок тензорного произведения.



**Рисунок 8:** Разреженная часть полного базиса В-сплайна тензорного произведения.

Естественным применением многомерных Р-сплайнов является сглаживание данных на сетка. Для более крупных сетей требования к памяти и времени вычислений могут стать слишком большими. нужны большие и специальные алгоритмы. См. Подробности в Разделе 13.

Многомерные Р-сплайны имеют хорошие числовые характеристики, в отличие от усеченных. степенные функции. Плохое численное состояние последнего становится практически непреодолимым. способный в более высоких измерениях. Сторонники TRF избежали этой проблемы, применив радиальные базисные функции (Ruppert et al., 2003; Kammann, Wand, 2003). Однако это не привлекательная схема: используется сложный алгоритм размещения центров базисные функции.

Подчеркнем, что множество тензорных произведений не обязательно должно быть прямоугольным, а!- хотя рисунок 8 может произвести такое впечатление. При работе, скажем, с данными в форме кольца домена, мы можем удалить все тензорные произведения, которые не перекрываются с кольцом, и уменьшить матрица штрафов соответственно. Это может сэкономить много вычислений, а в случае кольцо также более реалистично, потому что оно предотвращает работу штрафа через внутреннюю область, край.

Как мы показали для одномерного сглаживания, количество базисных элементов здесь тензорные произведения без проблем могут быть больше, чем количество наблюдений, благодаря штрафам.

## 9. Аддитивные гладкие структуры

Как мы видели для обобщенной аддитивной модели и модели с переменными коэффициентами, использование Р-сплайны приводят к набору (модифицированных) базисных матриц В-сплайнов, которые можно комбинировать бок о бок в одну большую матрицу. Штрафы приводят к блочно-диагональной матрице. Этот



идея распространяется на другие компоненты модели, такие как регрессия сигнала и тензорные произведения. Стан- Также можно добавить линейную регрессию и факторные члены. Это приводит к аддитивной гладкости конструкции. Эйлерс и Маркс (2002) предложили GLASS (обобщенное линейное аддитивное гладкое структур), в то время как Брезгер и Ланг (2006), ссылаясь на Fahrmeir et al. (2004), предложено STAR (структурированная аддитивная регрессия). Белиц и Ланг (2008) представили одновременное выбор переменных и параметров сглаживания в структурированных аддитивных моделях.

Геоаддитивная модель привлекла большое внимание; он образован добавлением одномерных гладких компонентов и двумерного пространственного тренда. Часто пространственная составляющая моделируется как условная авторегрессионная модель. Брезгер и Ланг (2006) представили байесовскую версию GLASS / STAR, также использующую 2D P-сплайны для выявления пространственных эффектов в полиномиальной логит-модели состояния леса. Фармейр и Кнейб (2009) построены на байесовских моделях STAR за счет включения геоаддитивных функций. и марковские случайные поля, обращаясь к неправильным априорным распределениям. Также считается Используя геодобавочную структуру, Kneib et al. (2011) расширенная и унифицированная байесовская STAR модели для дальнейшего учета многомерных ковариат.

Иерархии кривых образуют особый тип аддитивных гладких структур. Например, в данных о росте детей мы можем ввести общую кривую среднего и два дополнительных

кривые, которые показывают разницу между мальчиками и девочками. Более того, мы можем иметь гладкую кривая на одного ребенка. Дурбан и др. (2005) привел пример (используя усеченную степень функций), а Bugli и Lambert (2006) использовали правильные P-сплайны в байесовском контексте.

## 10. P-шлицы как смешанная модель.

Впервые была установлена связь между непараметрической регрессией и смешанными моделями. был выпущен более 25 лет назад Green (1987) и Speed (1991), но только в конце 1990-е годы, прежде чем он стал «горячей» темой исследований (Wang, 1998; Zhang et al., 1998; Verbyla et al., 1999), отчасти из-за разработок программного обеспечения для смешанных моделей. Эти первоначальные ссылки были основаны на использовании сглаживающих сплайнов. В контексте штрафных сплайнов несколько авторов быстро расширили формулировку модели до смешанной модели (Brumback et al., 1999; Coull et al., 2001; Жезл, 2003). Они использовали усеченные степенные функции как основы регрессии, поскольку они интуитивно связаны со смешанной моделью. Как- однако, как упоминалось ранее, числовые свойства TPFs плохие по сравнению с П-образные шлицы. В коротком комментарии, который по большей части остался незамеченным, Эйлерс (1999) показал, как интерпретировать P-сплайны как смешанную модель. Карри и Дурбан (2002) использовали этот подход и расширил его для обработки гетероскедастического или автокоррелированного шума. Работа по общему плану был также предложен метод представления сглаживающих устройств с квадратичным штрафом в смешанной модели. представлены в Fahrmeir et al. (2004).

При  $\lambda = \sigma^2 / \sigma_a^2$ , задача минимизации в (2) эквивалентна:

$$Q^* = \|y - Ba\|_2^2 / \sigma^2 + a^T Pa / \sigma_a^2, \quad (14)$$

с  $\sigma_a^2$  обозначает дисперсию случайных эффектов  $a$  и  $\sigma^2$  как дисперсию ошибки. В Фактически, это критерий минимизации в модели случайных эффектов вида:

$$y = Ba + \varphi, a \sim N(0, \sigma_a^2 I), \varphi \sim N(0, \sigma^2 I). \quad (15)$$

Как представлено, штрафы за различие порядка  $d$  не штрафуют степени  $x$  до degree  $d-1$ . Следовательно,  $P$  сингулярно ( $d$  собственных значений равны нулю), и, следовательно,  $a$  имеет вырожденный распределение. Одно из решений - переписать модель как  $Ba = X\beta + Zu$ , чтобы  $d$  столбцы  $X$  охватывают полиномиальное пустое пространство  $P$ , а  $(n-d)$  столбцы  $Z$  охватывают его дополнение. В этой презентации случайные эффекты  $u$  имеют невырожденное распределение.

тион. Этот тип повторной параметризации может быть выполнен разными способами. Эйлерс (1999) предложил  $Z = BD; (DD^T)^{-1}$  (где  $D$  - разностная матрица). Более принципиальный подход (который может использоваться для любого квадратичного штрафа) был введен Currie et al. (2006) и основан на разложении по сингулярным числам  $D = U \Sigma V^T$ , что дает  $Z = BU \Sigma^{-1}$ . В любом случае эквивалентная смешанная модель

$$y = X\beta + Zu + \epsilon, u \sim N(0, \sigma^2_u I) \quad \epsilon \sim N(0, \sigma^2_\epsilon I). \quad (16)$$

Вместо одного параметра сглаживания теперь у нас есть две дисперсии, и мы можем получить прибыль от стабильных и эффективных алгоритмов и программного обеспечения, доступных для смешанных моделей. Особенно в сложных моделях с несколькими гладкими компонентами, этот подход может быть более привлекательно, чем оптимизация перекрестной проверки или AIC. Но какой подход (на основе ошибки предсказания или максимальное правдоподобие) оптимально для выбора сглаживания параметр? За прошедшие годы появилось несколько статей по этой теме, но единой было достигнуто: Кауэрманн (2005а) показал, что оценка ML имеет тенденция к недостаточному сглаживанию и методы прогнозирования ошибок дают лучшую производительность, чем подходов, основанных на максимальном правдоподобии. Gu (2002) также обнаружил, что ML обеспечивает более грубую оценок, чем GCV, в то время как Ruppert et al. (2003) с помощью моделирования обнаружили, что REML дает более гладкую посадку, чем GCV (аналогичный вывод также был обнаружен в Кон и др., 1991). Кроме того, Wood (2011) пришел к выводу, что оценка REML / ML предпочтительнее. возможность GCV для полупараметрических GLM из-за его лучшей устойчивости к перетяжке, меньше изменчивость в расчетных параметрах сглаживания и снижение склонности к многократному tiple minima. Итак, ясно, что однозначного ответа на этот вопрос нет, так как разные При различных сценариях можно сделать разные выводы. Более того, за критериями, используемыми для при выборе параметра сглаживания возникает, на наш взгляд, более глубокий вопрос: справедливо ли использовать методологию смешанных моделей для оценки и вывода, когда смешанная модель представление P-сплайна можно считать просто «уловкой» для облегчения параметра оценка? Это вопрос, на который у нас нет ответа; у исследователей разные (и сильные) мнения о смешанном модельном подходе (даже авторы этой статьи не всегда соглашались по этому поводу), но правда в том, что это стало революцией, которая дала невероятные успехи за очень короткое время. Это, безусловно, помогло сделать штрафные шлицы «salonfähig»: в настоящее время они приемлемы и даже привлекательны для большая часть статистического сообщества.

Оценка фиксированных и случайных эффектов основана на максимизации совместная плотность  $(y, u)$  для  $\beta$  и  $u$ , что приводит к хорошо известным уравнениям Хендерсона (Хендерсон, 1975):

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X Z & Z^T Z + \lambda I \end{bmatrix} \begin{bmatrix} \beta \\ u \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix}, \quad (17)$$

где  $\lambda = \sigma^2_\epsilon / \sigma^2_u$ . Решение этих уравнений выходов  $\beta$  и  $u$ . Компонент дисперсии  $\sigma^2_\epsilon$  и  $\sigma^2_u$  как правило, оцениваются с помощью REML (ограниченного максимального правдоподобия), см. Паттерсон и Томпсон (1971)), а решения получают численно оптимальным методом. мизация.

Могут быть использованы и другие подходы, среди которых стоит упомянуть алгоритм данные Schall (1991), который оценил случайные эффекты и параметры дисперсии с Из-за необходимости указывать их распространение. Ключевым моментом является то, что каждый компонент дисперсии подключен к эффективному измерению. Сумма квадратов соответствующих случайных

коэффициенты равны их дисперсии, умноженной на их соответствующую эффективную размерность. Этот факт может быть использован в итеративном алгоритме. После каждого цикла сглаживания суммы квадратов и эффективных размеров пересчитываются, которые затем используются для обновления отклонения для следующего раунда. См. Marx (2010) и Rodriguez-Alvarez et al. (2015) для деталей и для расширений к многомерному сглаживанию.

Важно отметить факт предсказания. Хотя подогнанная модель такая же независимо от параметризации (т.е. смешанная модель или нет), стандартные ошибки для предварительных предсказанных значений не инвариантны. Это происходит потому, что изменчивость случайных эффектов учитывается в случае смешанной модели (а не в другом). Уверенность в-интервале, полученный из исходной параметризации  $F(\mathbf{x}) \pm 2\sigma\sqrt{(\mathbf{H}\mathbf{H})^{-1}}$  (где  $\mathbf{H}$  такая что  $\mathbf{y} = \mathbf{H}\mathbf{u}$ ). Этот доверительный интервал охватывает  $E[e(\mathbf{x})]$ , а не  $\phi(\mathbf{x})$ , так как  $\phi(\mathbf{x})$  не является несмещенная оценка  $f(\mathbf{x})$ . В то время как в смешанной модели структуры,  $e(\mathbf{x})$  является несмещенной из-за случайной  $\mathbf{U}$ , и смещенным скорректированным доверительного интервала  $F(\mathbf{x}) \pm 2\sigma\sqrt{(\mathbf{H})^{-1}}$  (Ruppert et al., 2003).

Конечно, интерес к представлению Р-сплайнов в смешанной модели был очень велик. с возможностью включения сглаживания в более крупный класс моделей. По факту, за последние 15 лет произошел взрыв моделей: от оценочных предметные кривые в продольных данных (Durbán et al., 2005), чтобы расширить классические модели в экономике Basile et al. (2014), чтобы сыграть ключевую роль в последних достижениях в функциональный анализ данных (Scheipl et al., 2015; Brockhaus et al., 2015) и другие.

### 10.1. Р-лицы и коррелированные ошибки

Хотя подход смешанной модели позволил обобщить многие существующие В моделях есть область, в которой он сыграл ключевую роль: данные с последовательной корреляцией. На протяжении многих лет основная сложность при примерке гладкой модели при наличии Корреляция - это совместная оценка параметров сглаживания и корреляции. Известно, что стандартные методы выбора параметров сглаживания (на основе минимизация среднеквадратичной ошибки прогноза) обычно недостаточно сглаживают данные в наличие положительной корреляции, поскольку плавный тренд плюс коррелированный шум может быть рассматривается как менее плавная тенденция плюс белый шум.

Решение состоит в том, чтобы явно учесть корреляционную структуру, т.е.  $\text{Var}(\mathbf{q}) = \sigma^2 \mathbf{V}$ , где  $\mathbf{V}$  может зависеть от одного или нескольких параметров корреляции. Дурбан и Карри (2003) представили стратегию выбора параметра сглаживания и оценки корреляции. ция на основе REML. Кривобокова и Кауэрманн (2007) показали, что максимальное подобие оценка правдоподобия параметра сглаживания является надежной, даже при умеренно некорректных характеристиках. Установленная корреляция. Этот метод позволил включить временные нелинейные тренды. и фильтрация временных рядов (Kauermann et al., 2011).

В последнее время и мотивирован необходимостью повысить скорость и стабильность прогнозов. модели, Вуд и др. (2015) разработали эффективные методы подбора аддитивных моделей. к большим наборам данных с коррелированными ошибками.

Корреляция проявляется и в более сложных ситуациях, например, в случае спатифактические данные. Ли и Дурбан (2009) объединили двумерные Р-сплайны и случайные эффекты с CAR (условной авторегрессивной) структурой для оценки пространственных трендов, когда данные географически распределены по точкам на карте. Другие авторы заняли разные подходы; они объединили аддитивные смешанные модели с пространственными эффектами, представлены марковскими или гауссовскими случайными полями (Kneib, Fahrmeir, 2006; Fahrmeir и др., 2010).

### 10.2. Многомерные Р-сплайны как смешанные модели

Многомерные Р-шлицы также можно рассматривать как смешанную модель. Первая попытка была сделано Ruppert et al. (2003) с использованием радиальных базисных функций. Curtie et al. (2006) проанализированы тензорное произведение Р-сплайны как смешанные модели. Здесь сингулярное разложение Матрица штрафа (как в одномерном случае) используется для построения матриц смешанной модели. Этот подход работает для любой суммы квадратичных штрафов (Wood, 2006a). Однако когда штраф выражается как сумма произведения Кронекера маргинальных базисов (сумма штрафов), представление в виде смешанной модели основано на репараметризации маргинальных баз. Важным побочным продуктом этой параметризации является то, что трансформированная матрица штрафов (т. е. ковариационная матрица случайных эффектов) и смешанная матрицы модели приводят к интересному разложению подобранных значений в виде суммы основных эффекты и взаимодействия (Ли и Дурбан, 2011):

$$E(Y) = f_1(x_1) + f_2(x_2) + f_3(x_1, x_2).$$

Это разложение тесно связано с работой, предложенной Гу (2002) по сглаживанию сплайновый дисперсионный анализ.

Модель теперь имеет несколько параметров сглаживания, что позволяет их оценивать. менее эффективен, если использовалась численная оптимизация. Было предпринято несколько шагов, чтобы сделать вычисления эффективными. Вуд (2011) использовал приближение Лапласа для получения приближенный REML подходит для эффективного прямого приближения. Ли и др. (2013) улучшено вычислительная эффективность за счет использования вложенных баз В-сплайнов и изменение штрафа так эту оптимизацию можно проводить с помощью стандартного статистического программного обеспечения. Вуд и Шайпл (2013) предложили сглаживание промежуточного низкого ранга. В последнее время алгоритм Шалла был распространен (Rodriguez-Alvarez et al., 2015) на случай многомерных гладких. Эта работа также показывает фундаментальную роль эффективных размеров компоненты модели.

### 11. Байесовские Р-сплайны.

От скрытого распределения в смешанной модели к априорному в байесовской модели - небольшой шаг. интерпретация. Байесовские Р-сплайны были предложены Лангом и Брезгером (2004), и они были доступны с помощью соответствующего программного обеспечения (Brezger et al., 2005). Их подход основан на моделировании цепи Маркова Монте-Карло (ММС). Что касается смешанной модели,

штраф приводит к сингулярному распределению. Это решается путем моделирования с использованием случайного ходя того же порядка, что и различия.

Также можно начать со смешанного модельного представления. Crainiceanu et al. (2007) сделали это в одном измерении, используя усеченные степенные функции. Они избегают тензора произведения TPF и перейти к радиальным базисным функциям для пространственного сглаживания. Эти авторцы также допускали различную (хотя и изотропную) гладкость и гетероскедастические-это Джуллион и Ламберт (2007) предложили байесовскую модель адаптивного сглаживания.

В качестве альтернативы МСМС используется интегрированное вложенное приближение Лапласа (INLA). Мощный и быстрый, и он набирает популярность (Rue et al., 2009). INLA избегает стохастическое моделирование для параметров точности и вместо этого использует численное интегрирование. По сути INLA использует параметр для каждого наблюдения, поэтому базис регрессии (В-сплайн) должен быть реализуется косвенным образом, как матрица ограничений (Fraaije et al., 2015).

INLA - прекрасный выбор для анизотропного сглаживания. Работая с суммой матриц точности он может обрабатывать эквивалент смешанной модели с перекрытием пенальти (Родригес-Альварес и др., 2015).

## 12. Варя

В этом разделе мы обсуждаем некоторые темы, которые не находят естественного дома ни в одном из предыдущих разделов. Мы рассмотрим асимптотические свойства Р-сплайнов и бустинг.

Несколько авторов изучали асимптотическое поведение Р-сплайнов. См. Ли и Руперт (2008); Claeskens et al. (2009); Кауэрманн и др. (2009); Wang et al. (2011). Хотя мы восхищаемся техническим уровнем этих вкладов, мы не полностью видим их практическая значимость. Проблема в их очень ограниченной интерпретации увеличения количество наблюдений: речь идет о большем количестве наблюдений в той же области. В этом В случае, если выясняется, что количество узлов должно расти как малая степень числа наблюдений. Тем не менее, вся идея Р-шлицев состоит в том, чтобы использовать слишком много узлов и позволить штраф делать свою работу. Попытка оптимизировать количество узлов, как это сделал Руперт (2002), - это не стоит. Он сообщает о некоторых случаях, когда большее количество узлов увеличивает ошибку оценки, но цифры не драматичны. Его анализ был основан на усеченных степенных функциях, которые он сбивает с толку, называя Р-сплайнами, с узлами в квантилях наблюдаемого  $x$ . Нет Понятно, как этот дизайн влияет на результаты. Для правильного анализа равномерно расположенные узлы должны быть использованы.

Большинство асимптотических анализов штрафных сплайнов используют структуру усеченной степени функции (TPF). Есть нештатная часть, многочлен от  $x$  той же степени

как у TPF. Штраф на TPF зависит от размера коэффициентов, а не от их отличия. Это упрощает аналитическую работу. В разделе 10 мы представили два альтернативных представления Р-сплайнов, которые имеют непенизованную полиномиальную часть и штраф размера по другим базисным функциям. Мы считаем, что такие подходят больше, чем TPF, хотя бы из-за развязки степени шлифовки и порядка штраф. Бур (2015) недавно представил вариант, при котором базис остается разреженным.

В работах по асимптотической теории пренебрегают тем, что часто приходится иметь дело с наблюдением во времени или пространстве, где большее количество наблюдений приводит к пропорциональному неравенству увеличение размера домена.

В разделе 4 мы показали, что нет опасности использовать много сплайнов, даже если соответствует только нескольким точкам данных. Следовательно, всегда можно использовать много сплайнов, а не побеспокоиться об оптимизации их количества. Поэтому мы рекомендуем использовать 100 В-шлицев, а безопасный выбор.

Повышение для сглаживания в основном работает следующим образом (Tutz and Binder, 2006; Schmid and Hothorn, 2008): (1) сглаживание с очень сильным штрафом и сохранение результата, (2) сгладить остатки и добавить (часть) этого результата к предыдущему результату, (3) повторно

шаг торфа (2) много раз. Результат становится более гибким с каждой итерацией. Итак, у одного есть остановиться в какой-то момент, используя AIC или другой критерий. У бустинга есть много энтузиазма сторонников, и его использование было распространено на ненормальные данные и аддитивные модели и другие гладкие структуры (Maug et al., 2012). Нам сложно увидеть его преимущества, особенно когда мы сравниваем его с алгоритмом Шалла для настройки нескольких параметров сглаживания. Параметры, которые мы представили в разделе 10. С другой стороны, повышение позволяет выбрать соответствующие переменные в модели и использование нестандартных целевых функций.

### 13. Вычисления и программное обеспечение

Для многих приложений стандартные R-шлицы не создают вычислительных проблем. В размер базиса B-сплайна будет умеренным, и многие тысячи наблюдений могут быть обрабатываться с легкостью. Если данные наблюдаются на эквидистантной сетке и только сглаживаются значения в этой сетке необходимы, можно также использовать единичную матрицу в качестве основы. Это приводит к более сглаживанию Уиттекера (Whittaker, 1923; Eilers, 2003). Количество коэффициенты будут равны количеству наблюдений, но в сочетании с разреженными Матричные алгоритмы очень быстро сглаживаются.

Разреженные матрицы также привлекательны, когда длинные ряды данных необходимо сглаживать с помощью большой базис B-сплайна (de Rooi et al., 2014). Несмотря на то, что базовая матрица разреженная, одна должен позаботиться о том, чтобы избежать вычисления плотных матриц по пути, как в случае, когда с использованием усеченных степенных функций. Ключ состоит в том, чтобы признать, что  $B_{j+k}(x) = B_j(x - s_k)$ , где  $s$  - расстояние между узлами. Каждый  $x_i$  сдвигается на  $-s_k$  в выбранную под-область, и на ней вычисляется базис только из четырех (кубических) B-сплайнов. В последний шаг строится разреженная матрица со сдвигом столбцов строки  $i$  вправо на  $k_i$ . Дополнительным преимуществом является то, что числовое округление сведено к минимуму (Eilers and Marx, 2010).

Когда параметр штрафа велик, формирование  $B^T B + \lambda D^T D$  явно для решения штрафные нормальные уравнения не являются оптимальными, и могут возникнуть проблемы с округлением. Это лучше использовать расширенную версию  $B$ , дающую  $\tilde{B}$ , где  $\tilde{B} = [B^T \sqrt{\lambda} D^T]^T$ , и расширенную  $y$  как  $\tilde{y} = [y; 0]^T$  и выполнить линейную регрессию  $\tilde{y}$  на  $\tilde{B}$  с помощью QR-разложения. Здесь  $0$  обозначает вектор нулей с длиной, равной числу строк  $D$ . Видеть Wood (2006a) за советы по стабильным вычислениям.

Требования аддитивных моделей к памяти компьютера и времени вычислений часто скромные. Однако очень большие наборы данных нуждаются в особой обработке, если они не подходят в оперативной памяти. Wood et al. (2015) описали такое приложение, прогнозируя потребление электроэнергии во Франции. Они разработали специализированный алгоритм, который является частью пакета R (mgcv) в качестве функции bam.

При двумерном сглаживании больших сеток с использованием тензорных произведений можно запустить в еще одну проблему. Данные и одномерные базы могут легко вписаться в память, но (внутренние продукты) продукции Кронекера не могут быть обработаны. Два-размерный базис  $\tilde{B} \otimes B$  имеет  $m_1 m_2$  строк и  $n_1 n_2$  столбцов. При сглаживании большого 1000 на 1000 изображений, используя  $n_1 n_2 \approx 1000$ , в базисе один миллиард элементов, беря 8 байта каждый, и поэтому не уместится в 8 Гб основной памяти. Даже если бы это было так, вычисления-Использование внутренних продуктов будет чрезвычайно обременительным. Обратите внимание, что в этом случае у нас есть около 1000 коэффициентов, так что это не размер окончательной системы штрафных нормальных уравнений. В этом проблема. К счастью, переставив расчеты, можно избежать явные произведения Кронекера и выигрывают на порядки в скорости вычислений и использовании памяти (Currie et al., 2006; Eilers et al., 2006). Этот алгоритм массива, так называемый GLAM допускает произвольные веса и поэтому подходит для обобщенного линейного сглаживания.

Когда веса не задействованы, возможны даже большие улучшения, используя «Более гладкий сэндвич» (Xiao et al., 2013). Основная идея состоит в том, что сначала можно применить одно-размерное сглаживание до строк матрицы, а затем до столбцов (или наоборот

вокруг). Порядок несущественен, как легко увидеть из явного уравнения для  $A = (B^T B + \lambda D^T D)^{-1} B^T Y \sim B (\tilde{B}^T \tilde{B} + \tilde{\lambda} \tilde{D}^T \tilde{D})^{-1}$ , матрица коэффициентов. Аналогичный подход за ним последовали Eilers и Goeman (2004), использовавшие модифицированный сглаживатель Whittaker.

В настоящее время довольно распространено публиковать компьютерный код на веб-сайтах или в качестве дополнительных приложений. Дополнительный материал, сопровождающий статистические статьи. Это, конечно, верно и для литературы на П-шлицах. Мы не пытаемся описывать эти отдельные усилия. Вместо этого мы указываем на некоторые пакеты для R (R Core Team, 2015) с довольно широкой сферой применения.

Первоначально разработан для использования в обобщенной аддитивной модели и сопутствующей модели Wood (2006b), MGCv превратилась в швейцарский армейский нож сглаживания. Он предлагает разнообразие базисные функции и их тензорные произведения для многомерного сглаживания. Более того, он может соответствовать моделям с переменными коэффициентами и регрессии сигнала, и он может смешивать и согласовывать компоненты аддитивным способом. Он предлагает множество дистрибутивов для обработки (чрезмерно разогнан) ненормальные данные.

Мы описали модель GAMLSS в разделе 5. Доступен очень обширный пакет. способный. Его ядро удачно названо gamlss; его можно расширить набором надстроек для цензурированные данные, смешанные модели и множество непрерывных и дискретных распределений.

Пакет MortalitySmooth ориентирован на сглаживание счетов за один и два. габариты (Камарда, 2012). Это также хороший источник данных о смертности из нескольких страны.

BayesX (Brezger et al., 2005) - это отдельная программа для Windows и Linux. Это охватывает все модели, которые вписываются в обобщенную линейную аддитивную гладкую структуру (или структурную структуру). Байесовские алгоритмы основаны на Маркове сеть Монте-Карло. Он также предлагает алгоритмы на основе смешанных моделей. Есть пакеты R установить BayesX и общаться с ним.

Также можно использовать пакет R-INLA (Rue et al., 2009) для подбора присадки. модели с П-образными шлицами. См. Fraaije et al. (2015) и сопутствующее программное обеспечение.

Пакет mboost предлагает усиление для различных моделей, включая P-шлицы и обобщенные аддитивные модели (Hofner et al., 2014). С расширением gamboostLSS один может применять усиление к моделям для определения местоположения, масштаба и формы, аналогично GAMLSS.

Для оценки гладких кривых или поверхностей доступен пакет expectreg.

## 14. Обсуждение

В статье Эйлерса и Маркса (1996), с которой все началось, содержалась «оценка потребителя. карта », сравнивая различные алгоритмы сглаживания. П-образные шлицы получили лучшие оценки и их изобретатели пришли к выводу, что они должны быть более плавным выбором. Два десятилетия спустя отродно видеть, что это мнение разделяют многие статистики и другие ученые. После того, как известные инструменты, такие как сглаживание ядра и локальная вероятность, постепенно исчезают в безвестности.

За двадцать лет методология P-сплайна была расширена во многих направлениях. В аналогия со смешанными моделями используется в полной мере, как и байесовский подход. подход, ведущий к новым интерпретациям штрафов и мощным рецептам оптимизации количество сглаживания. Многомерное сглаживание тензорными произведениями стало практично и быстро, благодаря алгоритмам массива. Регрессия по (многомерным) сигналам также стало практичным. Гладкие аддитивные структуры позволяют комбинировать различные модели. Ключевым моментом является сочетание богатой регрессии В-сплайна и простой шероховатости. штраф. На самом деле штрафы являются ядром, и было разработано множество вариантов, при этом основа В-сплайна не изменилась. Мы ожидаем увидеть захватывающие события в ближайшее будущее. Для начала набросаем некоторые аспекты, которые, как мы надеемся, привлекут много внимания.

Мы писали, что штраф формирует каркас, а В-шлицы накладывают плотность на кости. Это означает, что необходимо разработать новые идеи штрафов. Один многообещающий

авеню - приложение к дифференциальным уравнениям. Решение можно записать в виде суммы В-сплайны (метод коллокации) и использовать дифференциальное уравнение (DE) в качестве штрафа (Рамзи и др., 2007). В этом свете обычный штраф за сглаживание сплайнов эквивалентен к дифференциальному уравнению, которое говорит, что вторая производная решения равна нулю везде. О'Салливан (1986) сделал шаг от непрерывного штрафа к дискретному. Это также можно сделать с помощью штрафа на основе DE. Однако если коэффициенты ДЭ равны

не фиксируется (например, оценивается на основе данных), тогда это создает значительные вычислительные нагрузки. Будет полезно изучить (почти) эквивалентные дискретные штрафы, основанные на разнице уравнения.

Примечательно, что при одномерном сглаживании кригинг практически отсутствует. Альтман (2000) сравнили сплайны и кригинг и обнаружили, что серийная корреляция является ключевой проблемой. Если он присутствует и игнорируется, шлицы не работают. Есть способы справиться с корреляциями, как обсуждалось в Разделе 10. В анализе пространственных данных кригинг по-прежнему доминирует. Мы считаем, что для многих приложений Р-сплайны с тензорным произведением будут намного лучше. Выбор, особенно если человек больше заинтересован в оценке тенденции, чем в выполнении пространственной интерполяции. Может показаться, что попытка оценить ковариационную структуру Работа с данными стоит того, но на практике часто приводит к нестабильным процедурам. Обработка нестандартных данных с помощью кригинга обременительна. Напротив, Р-шлицы накладывают относительно простую ковариационную структуру и на практике выполняют работу очень стабильно, как показал наш опыт анализа сельскохозяйственных полевых испытаний. Сглаживание данные на больших сетках проблематичны для кригинга, но Р-сплайны и алгоритмы массивов обрабатывают такие данные с легкостью. В некоторых случаях может быть даже интересно обобщить данные (как считает и суммирует) на сетке перед анализом. В сочетании с подходом PS-ANOVA (Lee et al., 2013), что позволяет избежать детального моделирования взаимодействий более высокого порядка, Могут быть разработаны полные инструменты для больших наборов данных.

В некоторых приложениях очень важна экстраполяция. Данные о смертности являются первоочередными достаточно. Порядок отличий в штрафах определяет результат: для первого порядка разности - константа, для второго порядка прямая и взвешенная комбинация обоих дает экспоненциальную кривую. Задача состоит в том, чтобы определить, какой штраф использовать и чтобы установить параметры настройки для оптимальной экстраполяции. В одном измерении экстраполяции не влияет на соответствие данным. Это неверно в двух измерениях, например, с таблицы дожития. Штрафы за возраст и измерение времени взаимодействуют и размер область экстраполяции также имеет влияние. Необходимы дополнительные исследования, чтобы лучше понять эти вопросы.

В нескольких местах этой статьи мы встречали эффективную размерность (comp-компоненты) модели. Это важный параметр при оптимизации штрафов. Тем не менее, это де-обслуживает больше внимания само по себе. Определение Йе (1998) в (10) очень мощный. Вклад в  $ED = E_y dY_i / D y_j$  компонентом аддитивной модели можно четко определить, проследив за изменением  $y_i$  вниз по модели коэффициентов, и оттуда обратно снова к соответствующему изменению в  $y_j$ . Частичный эффективный размер-Таким образом можно рассчитать ссуды; они представляют собой важные резюме вкладов компоненты модели.

В этой статье мы попытались дать представление о многих достопримечательностях, созданных в последние 20 лет. Это было коллективное достижение, результат работы многих исследователи, которые верили в силу Р-шлицев. Мы видим прекрасное и захватывающее будущее впереди, поскольку предстоит решить множество проблем, новые сложные данные для моделирования, и особенно новое поколение ярких статистиков, которые уже показывают, что у Р-сплайнов много больше, чтобы внести свой вклад в этот век, век данных.



## Приложение А. О-образные шлицы или П-образные шлицы?

Ванд и Ормерод (2008) представили шлицы О'Салливана, или для краткости О-образные шлицы. Они были не совсем довольны чистым дискретным штрафом Р-шлицев и вернулись к интеграл от квадрата второй (или более высокой) производной подобранной функции. Это может быть неприятно, особенно когда узлы базиса В-сплайна расположены неравномерно. Есть случаи, когда это может быть очень ценно. Например, Whitehorn et al. (2013) представлены пример многомерного сглаживания с тензорными произведениями в физике высоких энергий для моделирования реакции детектора. В этом случае требовалось больше деталей в центре, чем у границ. Однако это не было мотивацией Ванда и Ормерода.

Они предпочитают использовать квантили  $x$  для узлов.

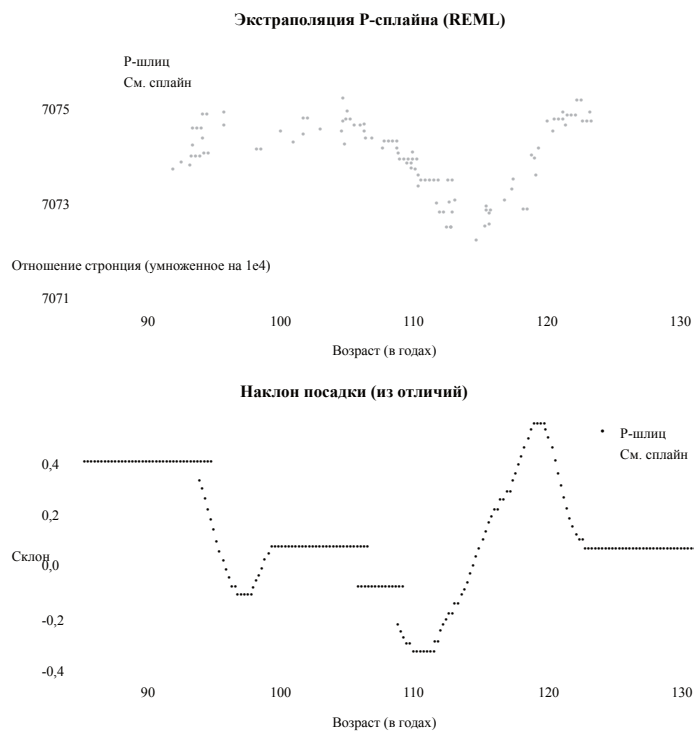
В статье утверждается, что Р-сплайны плохо экстраполируются по сравнению с сглаживающий сплайн. Следовательно, следует отдавать предпочтение О-образным шлицам. Это утверждение было повторено Ruppert et al. (2009). На статью Wand and Ormerod (2008) цитируют более чем 50 раз, так что видимо сообщение не потерялось.

Мы были обеспокоены этим анализом, потому что основание с несколькими узлами на доменные границы использовались для О-сплайнов. Если также использовалось несколько узлов для Р-шлицев (аналогично изображенному на нижней панели рисунка 3) артефакты могли иметь произошел. Поэтому мы решили взять пример данных из их статьи (ископаемое в наборе данных в R пакет SemiPar) и повторно проанализируйте их. Скачиваем файл WandOrmerod08.Rs с личной веб-страницы Мэтта Ванда. Для подгонки мы извлекли раздел, вызывающий функция lme в пакете nlme для оценки эквивалентной смешанной модели. Этот Программа адаптирована для Р-сплайнов путем изменения базиса и матрицы штрафов. Для Для сравнения мы используем функцию smooth.spline, которая является стандартной частью R. Она настраивает количество сглаживания данных автоматически с использованием перекрестной проверки.

На рисунке 9 показаны посадки Р-шлицевого и сглаживающего шлицев. Мы использовали 40 В-шлицев в домене от 85 до 130. Между двумя сплайнами существует сильное соответствие. Это также верно и для оцененной производной, которая была аппроксимирована с помощью различных ences.

Удивительно, но, как показано на Рисунке 10, О-образные шлицы работают не так хорошо, как Р-образные. Особенно это видно по производным. Похоже, что О-образный сплайн не справляется. около  $x = 100$ . Причина в том, что плотность узлов там мала, потому что низкая локальная плотность данных. Что еще более беспокоит, так это то, что производные экстраполированной части не постоянным, как и должно быть для линейного результата.

Мы полагаем, что anomalous поведение О-сплайнов вызвано выбором основание. Множественные узлы не идут вместе с дискретным штрафом на коэффициенты сплайна. Корень всех зол - это выбор использовать квантили  $x$  для узлов; есть абсолютно в этом нет необходимости.



**Рисунок 9:** Верхняя панель: Р-сплайн и сглаживающий сплайн соответствуют данным окаменелостей. Подгонка с обеих сторон экстраполирована автоматически. Нижняя панель: производные обоих сплайнов, рассчитанные на основе первых разностей.

## Приложение Б. Доказательство того, что эффективный размер меньше $m$

Формальное доказательство начинается с упрощенного случая с  $d = 0$  и общим базисом  $Z$ , где система уравнений, которую необходимо решить:

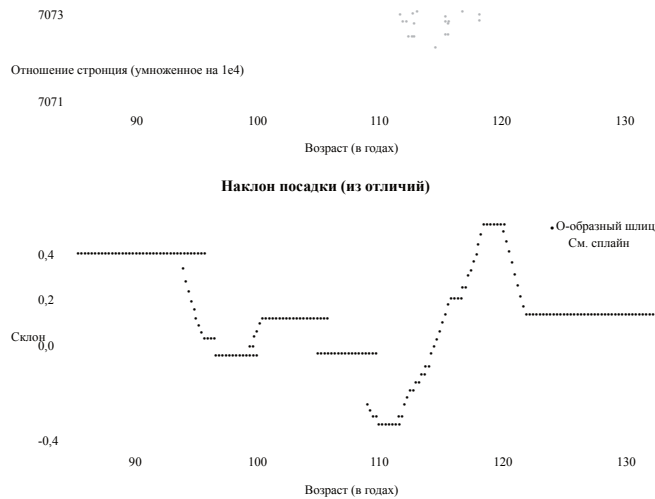
$$(Z^T Z + \lambda I) u = Z^T y. \quad (18)$$

Разложение  $Z$  по сингулярным числам дает:  $Z = USV^T$ , где  $U^T U = I_m$  и  $V^T V = I_n$ .

Путем замены:

$$(B^T U S V^T + \lambda I) y = (V S^2 B^T + X I) = B^T y \quad (19)$$





**Рис. 10:** Верхняя панель: О-образный и сглаживающий сплайн, соответствующие данным окаменелостей. С обеих сторон посадка экстраполируется автоматически. Нижняя панель: производные обоих шлицев, рассчитанные на основе первых разностей.

Теперь предположим, что  $u = V\gamma$ . Заполните и умножьте на  $V^T$ :

$$V^T (VS_2 V^T + \lambda I) V \gamma = V^T V S U^T y. \quad (20)$$

Следовательно

$$(S_2 + \lambda I_m) \gamma = S U^T y. \quad (21)$$

Это система с  $m$  уравнениями с  $m$  неизвестными. Матрица системы диагональна и неособое число.

Штраф здесь является частным случаем, но в разделе 10 было показано, что Р-сплайны могут трансформироваться в смешанную модель, в частности с  $Ba = X\beta + Zu$  и гребнем штраф на  $u$ . Уравнения Хендерсона (17) содержат часть для  $\beta$ . Теперь у нас есть

$$Z^T X^T \beta + (Z^T Z + \lambda I) u = Z^T y, \quad (22)$$

или же

$$(S_2 + \lambda I_m) u = S U^T (y - Z^T X \beta). \quad (23)$$

Обратите внимание, что значение  $\beta$  несущественно, так как оно не меняет свойств системы. матрица и изменяется только правая часть уравнения.

## Подтверждение

Часть работы первого автора была поддержана грантом кафедры передового опыта от Университет Карлоса III в Мадриде, Испания.

## Рекомендации

Альтман, Н. (2000). Теория и методы: Криг, гладкий, оба или ни один? *Австралийский и новозеландский журнал статистики*, 42, 441–461.

- Андрияна Ю., Гиджельс И. и Верхасселт А. (2014). Оценка квантильной регрессии Р-сплайнов в различных модели коэффициентов. *Тест*, 23, 153–194.
- Антониадис А., Грегуар Г. и Маккинг И. (2004). Байесовское оценивание в одноиндексных моделях. *Статистика Sinica*, 14, 1147–1164.
- Азмон, А., Фэйс, К. и Хенс, Н. (2014). Об оценке количества репродукций на основе неправильной записи. перенесенные эпидемические данные. *Статистика в медицине*, 33, 1176–1192.
- Базиле, Р., Дурбан, М., Мингес, Р., Монтеро, Дж. И Мур, Дж. (2014). Моделирование региональной экономической динамики ис: пространственная зависимость, пространственная неоднородность и нелинейности. *Журнал экономики и динамики Контроль*, 48, 229–245.
- Белиц, К. и Ланг, С. (2008). Одновременный выбор переменных и параметров сглаживания в структурированном аддитивные регрессионные модели. *Вычислительная статистика и анализ данных*, 53, 61–81.
- Боллаерте, К., Эйлерс, РНС и Aerts, М. (2006). Квантильная регрессия с ограничениями монотонности с использованием Р-шлицы и  $l_1$ -норм. *Статистическое моделирование*, 6, 189–207.
- Брезгер А. и Ланг С. (2006). Обобщенная структурированная аддитивная регрессия на основе байесовских Р-сплайнов. *Вычислительная статистика и анализ данных*, 50, 967–991.
- Брезгер А., Кнейб Т. и Ланг С. (2005). BayesX: анализ байесовской структурированной аддитивной регрессии модели. *Журнал статистического программного обеспечения*, 14.
- Брокгауз, С., Шайпл, Ф., Торстен, Х., Гревен, С. (2015). Функциональная модель линейного массива. *Статистическая Моделирование*, 15, 279–300.
- Брамбак Б., Рупперт Д. и Ванд М. (1999). Комментарий к: выбор переменной и оценка функции в аддитивной непараметрической регрессии с использованием априорной оценки на основе данных. *Журнал американской статистики Ассоциация*, 94, 794–797.
- Бугли, К. и Ламберт, П. (2006). Функциональный дисперсионный анализ со случайными функциональными эффектами: приложение к моделированию событийных потенциалов для анализа электроэнцефалограмм. *Статистика в медицине*, 25, 3718–3739.
- Камарда, CG (2012). MortalitySmooth: пакет R для сглаживания счета Пуассона с помощью Р-сплайнов. *Journal-журнал статистического программного обеспечения*, 50, 1–24.
- Камарда, К.Г., Эйлерс, РНС и Гамп, Дж. (2008). Моделирование общих закономерностей предпочтения цифр. *Статистическое моделирование*, 8, 385–401.

- Клаескенс Г., Кривобокова Т. и Опсомер Дж. Д. (2009). Асимптотические свойства оценки сплайнов со штрафом. *Биометрика*, 96, 529–544.
- Коулл Б., Шварц Дж. И Ванд М. (2001). Здоровье дыхательных путей и загрязнение воздуха: аддитивная смешанная модель анализа. *Биостатистика*, 2, 337–349.
- Крейничану К., Рупперт Д. и Кэрролл Р. (2007). Пространственно-адаптивные байесовские штрафные сплайны с гетероскедастическими ошибками. *Журнал вычислительной и графической статистики*, 17, 265–288.
- Карри, И., Дурбан, М. и Эйлерс, П. (2004). Сглаживание и прогнозирование показателей смертности. *Статистический моделинг*, 4, 279–298.
- Карри, И. Д. и Дурбан, М. (2002). Гибкое сглаживание с помощью р-сплайнов: единый подход. *Статистическая Моделирование*, 2, 333–349.
- Карри, И. Д., Дурбан, М. и Эйлерс, РНС (2006). Обобщенные модели линейных массивов с приложениями для многомерное сглаживание. *Журнал Королевского статистического общества, серия В-Статистический метод осы*, 68, 259–280.
- де Рой, Дж. Дж., ван дер Перс, Нью-Мексико, Хендрикс, RWA, Делхез, Р., Боттгер, А. Дж. и Эйлерс, РНС (2014). Сглаживание данных дифракции рентгеновских лучей и исключение К альфа (2) с использованием штрафной вероятности и модель составного звена. *Журнал прикладной кристаллографии*, 47, 852–860.
- Добсон, А. и Барнетт, А. (2008). *Введение в обобщенные линейные модели*, 3-е изд. CRC Press.
- Дурбан М. и Карри И. (2003). Примечание об аддитивных моделях Р-сплайна с коррелированными ошибками. *Вычислительная Статистика*, 18, 251–262.
- Дурбан, М., Харезлак, Дж., Ванд, М. и Кэрролл, Р. (2005). Простая подгонка тематических кривых для продольные данные. *Статистика в медицине*, 24, 1153–1167.
- Эйлерс, П. (1999). Обсуждение: анализ запланированных экспериментов и продольных данных с использованием гладкого  $\text{ing}$  шлицы. *Журнал Королевского статистического общества. Серия С-Прикладная статистика*, 48, 307–308.
- Эйлерс, П. (2005). Унимодальное сглаживание. *Журнал хеометрики*, 19, 317–328.
- Эйлерс, П. (2009). Гладкий комплексный логарифм и квазипериодические модели. В Т. Kneib and G. Tutz, *editors, статистическое моделирование и регрессионные структуры*. Springer.
- Эйлерс, П. и де Менезес, Р. (2005). Квантильное сглаживание массива данных CGH. *Биоинформатика*, 21, 1146–1153.
- Эйлерс П., Карри И. и Дурбан М. (2006). Быстрое и компактное сглаживание больших многомерных сеток. *Вычислительная статистика и анализ данных*, 50, 61–76.
- Эйлерс, РНС (2003). Идеально гладкий. *Аналитическая химия*, 75, 3631–3636.

- Эйлерс, РНС (2007). Проблемы, связанные с подсчетами, моделью составной связи и штрафной вероятностью. *Статистическое моделирование*, 7, 239–254.
- Эйлерс, РНС и Боргдорф, MW (2007). Непараметрические бревенчато-вогнутые смеси. *Вычислительная статистика и анализ данных*, 51, 5444–5451.
- Эйлерс, РНС и Goeman, JJ (2004). Улучшение диаграмм рассеяния со сглаженными плотностями. *Биоинформатика*, 20, 623–628.
- Эйлерс, РНС и Маркс, Б.Д. (1992). Обобщенные линейные модели с P-сплайнами. В *трудах GLIM 92 и 7-й Международной семинар по статистическому моделированию*.
- Эйлерс, РНС и Маркс, Б.Д. (1996). Гибкое сглаживание с помощью B-сплайнов и штрафов. *Статистическая наука*, 11, 89–102.
- Эйлерс, РНС и Маркс, Б.Д. (2002). Обобщенные линейные аддитивные гладкие структуры. *Журнал Сопри-Статистическая и графическая статистика*, 11, 758–783.
- Эйлерс, РНС и Маркс, Б.Д. (2003). Многомерная калибровка с температурным взаимодействием с использованием двух размерная регрессия сигнала со штрафом. *Хеометрия и интеллектуальные лабораторные системы*, 66, 159–174.
- Эйлерс, РНС и Маркс, BD (2010). Сплайны, узлы и штрафы. *Междисциплинарные обзоры Wiley: Вычислительная статистика*, 2, 637–653.

- Эйлерс, РНС, Гамп, Дж., Маркс, Б.Д. и Pay, P. (2008). Модели модуляции для сезонных временных рядов и таблицы заболеваемости. *Статистика в медицине*, 27, 3430–3441.
- Эйлерс, РНС, Ли Б. и Маркс Б.Д. (2009). Многомерная калибровка с одноиндексной регрессией сигнала. *Хеометрика и интеллектуальные лабораторные системы*, 96, 196–202.
- Фармейр, Л. и Кнейб, Т. (2009). Приемлемость апостериоров в моделях структурированной аддитивной регрессии: теория и эмпирические данные. *Журнал статистического планирования и вывода*, 139, 843–859.
- Фармейр, Л., Кнейб, Т. и Ланг, С. (2004). Штрафная структурированная аддитивная регрессия для пространственно-временных данных: а Байесовская перспектива. *Statistica Sinica*, 14, 731–761.
- Фармейр, Л., Кнейб, Т. и Конрат, С. (2010). Байесовская регуляризация в структурированной аддитивной регрессии: а унифицированный подход к усадке, сглаживанию и выбору предикторов. *Статистика и вычисления*, 20, 203–219.
- Fraaije, RGA, ter Braak, C.J.F, Verduyn, B., Breeman, LBS, Verhoeven, JTA and Soons, MB (2015). Ранние этапы набора растений задают шаблон для развития структуры вегетации. по гидрологическому градиенту. *Функциональная экология*, 29, 971–980.
- Фраско и Эйлерс (2015). L- и V-образные кривые для оптимального сглаживания. *Статистическое моделирование*, 15, 91–111.
- Грин, П.Дж. (1987). Правдоподобие со штрафом для общих полупараметрических регрессионных моделей. *Международная Статистический обзор / Revue Internationale de Statistique*, 245–259.
- Гревен, С. (2008). *Нестандартные задачи вывода для аддитивных и линейных смешанных моделей*. Cuvillier Verlag.
- Гу, С. (2002). *Сглаживающие сплайновые модели ANOVA*. Springer.
- Хансен, PC (1992). Анализ дискретных некорректных задач с помощью L-кривой. *SIAM Обзор*, 34, 561–580.
- Harezlak, J., Coull, BA, Laird, NM, Magari, SR и Christiani, DC (2007). Штрафные решения для проблемы функциональной регрессии. *Вычислительная статистика и анализ данных*, 51, 4911–4925.
- Харвилл, Д. (1977). Подходы максимального правдоподобия к оценке компонентов дисперсии и к связанным проблемам. *Журнал Американской статистической ассоциации*, 72, 320–338.
- Хасты, Т. и Тибширани, Р. (1990). *Обобщенные аддитивные модели*. Чепмен и Холл.
- Хасты, Т. и Тибширани, Р. (1993). Модели с переменными коэффициентами. *Журнал Королевского статистического общества. Серия В - Статистическая методология*, 55, 757–796.
- Хайм, С., Фармейр, Л., Эйлерс, РНС и Marx, BD (2007). Трехмерные пространственно-изменяющиеся модели коэффициентов с приложением к диффузионной тензорной визуализации. *Вычислительная статистика и анализ данных*, 51, 6212–6228.
- Хендерсон, К. (1975). Лучшая линейная объективная оценка и прогноз в рамках модели выбора. *Биометрия*, 31, 423–447.
- Хофнер Б., Майр А., Робинзонов Н. и Шмид М. (2014). Модельно-ориентированное усиление в R: практическое занятие учебник с использованием пакета R mboost. *Вычислительная статистика*, 29, 3–35.
- Джарроу Р., Рунперт Д. и Ю. Ю. (2004). Оценка временной структуры процентной ставки корпоративного долга с Полупараметрическая сплайн-модель со штрафом. *Журнал Американской статистической ассоциации*, 99, 57–66.
- Джуллион А. и Ламберт П. (2007). Надежная спецификация априорного распределения штрафа за шероховатость в сплайново-адаптивные байесовские модели P-сплайнов. *Вычислительная статистика и анализ данных*, 51, 2542–2558.
- Камманн, Э. и Ванд, М. (2003). Геоаддитивные модели. *Журнал серии Королевского статистического общества C-Прикладная статистика*, 52, 1–18.
- Кауэрманн, Г. (2005a). Замечание о выборе параметра сглаживания для сглаживания сплайнов со штрафом. *Журнал*

- Кауэрманн, Г., Кривобокова, Т. и Фармейр, Л. (2009). Некоторые асимптотические результаты об обобщенных штрафах. сглаживание сплайнов. *Журнал Королевского статистического общества. Серия В-Статистическая методология* , 71, 487–503.
- Кауэрманн, Г., Кривобокова, Т., Семмлер, В. (2011). Фильтрация временных рядов со штрафными сплайнами. *Исследования в области нелинейной динамики и эконометрики* , 15, 1–26.
- Кауэрманн, Г., Шеллхаз, К. и Рупперт, Д. (2013). Гибкая оценка плотности копулы со штрафными иерархическими В-сплайнами. *Скандинавский статистический журнал* , 40, 685–705.
- Кнейб Т. и Фармейр Л. (2006). Структурированная аддитивная регрессия для категориальных пространственно-временных данных: смешанный модельный подход. *Биометрия* , 62, 109–118.
- Кнейб Т., Конрат С. и Фармейр Л. (2011). Высокоразмерные структурированные аддитивные регрессионные модели: Байесовская регуляризация, сглаживание и прогнозирование. *Журнал Королевского Статистического Общества С-Прикладная статистика* , 60, 51–70.
- Кон, Р., Ансли, К. и Тарм, Д. (1991). Эффективность перекрестной проверки и максимальное правдоподобие капотные оценки параметров сглаживания сплайнов. *Журнал Американской статистической ассоциации* , 86, 1042–1050.
- Кривобокова, Т., Кауэрманн, Г. (2007). Замечание о сглаживании сплайнов со штрафом с коррелированными ошибками. *Журнал Американской статистической ассоциации* , 102, 1328–1337.
- Кривобокова Т., Кауэрманн Г. и Архонтакис Т. (2006). Оценка временной структуры процентных ставок с использованием штрафных шлицев. *Статистические статьи* , 47, 443–459.
- Ламберт, П. (2011). Гладкая полупараметрическая и непараметрическая байесовская оценка двумерных плотностей из данных двумерной гистограммы. *Вычислительная статистика и анализ данных* , 55, 429–445.
- Ламберт П. и Эйлерс, РНС (2009). Оценка байесовской плотности на основе сгруппированных непрерывных данных. *Сот-Предполагаемая статистика и анализ данных* , 53, 1388–1399.
- Ланг, С. и Брезгер (2004). Байесовские Р-сплайны. *Журнал вычислительной и графической статистики* , 13, 183–212.
- Ли, Д.-Дж. и Дурбан, М. (2009). Смешанные модели Smooth-CAR для пространственных данных подсчета. *Вычислительная статистика и анализ данных* , 53, 2968–2979.
- Ли, Д.-Дж. и Дурбан, М. (2011). Р-сплайн модели взаимодействия типа ANOVA для пространственно-временного сглаживания. *Статистическое моделирование* , 11, 49–69.
- Ли, Д.-Дж., Дурбан, М. и Эйлерс, П. (2013). Эффективное двумерное сглаживание с помощью Р-сплайна ANOVA смешанные модели и вложенные базы. *Вычислительная статистика и анализ данных* , 61, 22–37.
- Leitenstorfer, F. и Tutz, G. (2011). Оценка одноиндексных моделей на основе методов повышения. *Статистическое моделирование* , 11, 203–217.
- Ли Б. и Маркс Б. Д. (2008). Усиление регрессии сигнала Р-сплайна. *Статистическое моделирование* , 8, 367–383.
- Ли Й. и Рупперт Д. (2008). Об асимптотике штрафных сплайнов. *Биометрика* , 95, 415–436.
- Лу Х., Чен Г., Сингх Р. и Сонг П. (2006). Класс частично линейных одноиндексных моделей выживания. *Канадский статистический журнал* , 34, 97–112.
- Лу Ю., Чжан Р. и Чжу Л. (2008). Штрафная оценка сплайна для моделей с переменными коэффициентами. *Коммунальный Сообщения в статистике, теории и методах* , 37, 2249–2261.
- Мальфейт, Н. и Рамзи, Дж. (2003). Историческая функциональная линейная модель. *Канадский статистический журнал* , 31, 185–201.
- Маркс, Б.Д. (2010). Модели с переменным коэффициентом Р-сплайна для сложных данных. В редакции Г. Тутца и Т. Кнейба, *Статистическое моделирование и регрессионные структуры* , 19–43. Springer.
- Маркс, Б.Д. (2015). Одноиндексная регрессия сигнала с переменным коэффициентом. *Хемомерия и интеллектуальная лаборатория-операторские системы* , 143, 111–121.
- Маркс, Б. Д. и Эйлерс, РНС (1998). Прямое обобщенное аддитивное моделирование со штрафной вероятностью. *Вычислительная статистика и анализ данных* , 28, 193–209.

- Маркс, Б.Д. и Эйлерс, РНС (1999). Обобщенная линейная регрессия для дискретизированных сигналов и кривых: а Р-шлицевой подход. *Технометрика*, 41, 1–13.
- Маркс, Б.Д. и Эйлерс, РНС (2005). Многомерная регрессия сигнала со штрафами. *Технометрика*, 47, 13–22.
- Маркс, Б.Д., Эйлерс, РНС, Гамп, Дж. И Рау, Р. (2010). Модели билинейной модуляции для сезонных таблиц отсчетов. *Статистика и вычисления*, 20, 191–202.
- Маркс, Б.Д., Эйлерс, РНС и Ли, Б. (2011). Многомерная одноиндексная регрессия сигнала. *Химии-метрики и интеллектуальные лабораторные системы*, 109, 120–130.
- Майр А., Фенске Н., Хофнер Б., Кнейб Т. и Шмид М. (2012). Обобщенные аддитивные модели для локальных размеры, масштаб и форма для данных большого размера. Гибкий подход, основанный на бустинге. *Журнал Королевского статистического общества, серия С*, 61, 403–427.
- Маклин, М.В., Хукер, Г., Стайку, А.-М., Шайпл, Ф. и Рупперт, Д. (2014). Функционально-обобщенный Аддитивные модели. *Журнал вычислительной и графической статистики*, 23, 249–269.
- Моррис, Дж.С. (2015). Функциональная регрессия. *Ежегодный обзор статистики и ее применение*, 2, 321–359.
- Майерс, Р. (1989). *Классическая и современная регрессия с приложениями*. PWS-KENT.
- О'Салливан, Ф. (1986). Статистический взгляд на некорректно поставленные обратные задачи (с обсуждением). *Статистическая Наука*, 1, 505–527.
- Паттерсон, Х. и Томпсон, Р. (1971). Восстановление межблочной информации при неравных размерах блоков. *Биометрика*, 58, 545–554.
- Пя, Н. и Вуд, С. (2015). Аддитивные модели с ограничением формы. *Статистика и вычисления*, 25, 543–559.
- R Core Team (2015). *R: Язык и среда для статистических вычислений*. Фонд R для статистического Computing, Вена, Австрия.
- Рамзи, Дж. И Сильверман, Б. (2003). *Функциональный анализ данных, 2-е изд.* Springer.
- Рамзи, Дж. О., Хукер, Г., Кэмпбелл, Д. и Цао, Дж. (2007). Оценка параметров дифференциальных уравнений: обобщенный подход сглаживания. *Журнал Королевского статистического общества: серия В (Статистическая Методология)*, 69, 741–796.
- Ригби Р. и Стасинопулос Д. (2005). Обобщенные аддитивные модели для расположения, масштаба и формы. *Журнал Королевской вычислительной статистики и прикладной статистики*, 54, 507–544.
- Риппе, R.C.A., Меулман, Дж. Дж. И Эйлерс, РНС (2012a). Надежное однокристальное генотипирование с полупараметрическими бревенчато-вогнутыми смесями. *Плос Одн*, 7.
- Риппе, R.C.A., Меулман, Дж. Дж. И Эйлерс, РНС (2012b). Визуализация геномных изменений по сегментам сглаживания с использованием штрафа  $L_0$ . *PLoS ONE*, 7.
- Рицци, С., Гамп, Дж. И Эйлерс, РНС (2015). Эффективное оценивание гладких распределений из грубых сгруппированные данные. *Американский журнал эпидемиологии*, 182, 138–147.
- Родригес-Альварес, М., Ли, Д., Кнейб, Т., Дурбан, М. и Эйлерс, П. (2015). Параметр быстрого сглаживания разделение в многомерных обобщенных Р-сплайнах: алгоритм SAP. *Статистика и вычисления*, 25, 941–957.
- Рю, Х., Мартино, С. и Шопен, Н. (2009). Приближенный байесовский вывод для латентных гауссовских моделей с использованием интегрированных вложенных аппроксимаций Лапласа (с обсуждением). *Журнал Королевской Статистической Общества, Series B*, 71, 319–392.
- Рупперт, Д. (2002). Выбор количества узлов для штрафных шлицев. *Журнал вычислительной и Графической статистики*, 11, 735–757.
- Рупперт Д. и Кэрролл Р. (2000). Пространственно-адаптивные штрафы за установку шлицев. *Австралия и Новая Зеландия Журнал статистики*, 42, 205–223.
- Рупперт Д., Ванд М. и Кэрролл Р. (2003). *Полупараметрическая регрессия*. Кембридж.
- Рупперт, Д., Палочка, член парламента и Кэрролл, Р.Дж. (2009). Полупараметрическая регрессия в 2003–2007 гг. *Электронный Статистический журнал*, 3, 1193–1256.
- Шалл Р. (1991). Оценка в обобщенных линейных моделях со случайными эффектами. *Биометрика*, 78, 719–727.

- Шейпл, Ф., Стайку, А. и Гревен, С. (2015). Функционально-аддитивные смешанные модели. *Журнал вычислительных и графическая статистика*, 24, 477–501.
- Schellhase, С. и Kauermann, G. (2012). Оценка плотности и сравнение с оптрафованной смесью подход. *Вычислительная статистика*, 27, 757–777.
- Шмид, М. и Хотхорн, Т. (2008). Улучшение аддитивных моделей с помощью покомпонентных Р-сплайнов. *Современная международная статистика и анализ данных*, 53, 298–311.
- Шнабель, С.К. и Эйлерс, РНС (2009). Оптимальное ожидаемое сглаживание. *Вычислительная статистика и данные Анализ*, 53, 4168–4177.
- Шнабель, С.К. и Эйлерс, РНС (2013). Одновременная оценка кривых квантилей с использованием листов квантилей. *ASTA-Достижения в статистическом анализе*, 97, 77–87.
- Соботка Ф., Кнейб Т. (2012). Геоаддитивная ожидаемая регрессия. *Вычислительная статистика и анализ данных*, 56, 755–767.
- Соботка, Ф., Кауэрманн, Г., Вальтруп, Л.С., Кнейб, Т. (2013). О доверительных интервалах для полупараметров регрессии. *Статистика и вычисления*, 23, 135–148.
- Скорость, Т. (1991). Комментарий: что BLUP - это хорошо: оценка случайных эффектов. *Статистическая Наука*, 6, 15–51.
- Томпсон Р. и Бейкер Р. (1981). Функции составного зацепления в обобщенных линейных моделях. *Прикладная статистика*, 30, 125–131.
- Тутц, Г. и Биндер, Х. (2006). Обобщенное аддитивное моделирование с неявным выбором переменных по вероятности-усиление на основе. *Биометрия*, 62, 961–971.
- Вербилла А., Куллис Б., Кенвард М. и Уэллхэм С. (1999). Анализ запланированных экспериментов и продольные данные с использованием сглаживающих сплайнов. *Журнал Королевского статистического общества, серия C-Applied Statistics*, 48, 269–312.
- Жезл М. (2003). Сглаживающие и смешанные модели. *Вычислительная статистика*, 18, 223–249.
- Wand, МР и Ормерод, JT (2008). О полупараметрической регрессии со штрафными сплайнами ОСалливана. *Статистический журнал Австралии и Новой Зеландии*, 50, 179–198.
- Ван Х., Шен Дж. и Рупперт Д. (2011). Об асимптотике сглаживания штрафных сплайнов. *Электронный Журнал статистики*, 5, 1–17.
- Ван, Х.-Ф., Ху, Б., Ван, Б. и Фанг, К. (2014). Байесовские обобщенные модели с переменными коэффициентами для продольные пропорциональные данные с ошибками в ковариатах. *Журнал прикладной статистики*, 41, 1342–1357.
- Ван, Ю. (1998). Смешанные эффекты сглаживания сплайнового дисперсионного анализа. *Журнал Королевской Статистической Общества. Серия В - Статистическая методология*, 60, 159–174.
- Уайтхорн, Н., ван Сантен, Дж. И Лафевре, С. (2013). Штрафные сплайны для плавного представления высоких размерные наборы данных Монте-Карло. *Computer Physics Communications*, 184, 2214–2220.
- Уиттакер, Э. (1923). О новой методике градуировки. *Труды Эдинбургского математического общества*, 41, 63–75.
- Вуд, С. (2004). Стабильная и эффективная оценка параметров множественного сглаживания для обобщенной аддитивной модели. *Журнал Американской статистической ассоциации*, 99, 673–686.
- Вуд, С. (2006а). *Обобщенные аддитивные модели: Введение с R*. CRC Press.
- Вуд, С. (2011). Быстрая стабильная ограниченная оценка максимального и предельного правдоподобия полупараметрические обобщенные линейные модели. *Журнал Королевского статистического общества. Серия В-Статистическая Методология*, 73, 3–36.
- Вуд, С., Шейпл, Ф. и Фарауэй, Дж. (2013). Прямое сглаживание тензорного произведения промежуточного ранга в смешанных моделях. *Статистика и вычисления*, 23, 341–360.
- Вуд, С. Н. (2006б). О доверительных интервалах для обобщенных аддитивных моделей на основе штрафной регрессии шлицы. *Статистический журнал Австралии и Новой Зеландии*, 48, 445–464.

- Вуд, С.Н., Гуд, Ю. и Шоу, С. (2015). Обобщенные аддитивные модели для больших наборов данных. *Журнал Серия С-Прикладная статистика Королевского статистического общества*, 64, 139–155.
- Сяо, Л., Ли, Ю. и Рупперт, Д. (2013). Быстрые двумерные Р-шлицы: более гладкий сэндвич. *Журнал Королевского статистического общества, серия В-Статистическая методология*, 75, 577–599.
- Е, Дж. (1998). Об измерении и корректировке эффектов интеллектуального анализа данных и выбора модели. *Журнал Американская статистическая ассоциация*, 93, 120–131.
- Ю. Ю. и Рупперт Д. (2002). Штрафная оценка сплайна для частично линейных одноиндексных моделей. *Журнал Ассоциации численных статистиков*, 97, 1042–1054.



Чжан Д., Линь Х., Раз Дж. И Соуэрс М. (1998). Полупараметрические стохастические смешанные модели для продольных данные. *Журнал Американской статистической ассоциации*, 93, 710–719.