Об одном подходе к настройке алгоритма Метрополиса-Гастингса для задачи разделения смеси гауссовских компонент*

Ю. А. Дубнов [I, II], А. В. Булычев [I, II]

[I] Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук", г. Москва, Россия
[II] Национальный исследовательский университет "Высшая школа экономики", г. Москва, Россия

Аннотация. В статье рассматривается задача разделения смеси гауссовских компонент, заключающаяся в определении по имеющимся наблюдениям параметров компонент смеси. Предлагается подход к решению данной задачи, основанный на байесовском оценивании с применением наиболее информативных априорных распределений (Maximal Data Information Prior – MDIP). Новизна описанного подхода заключается в использовании выборочных оценок для вычисления априорного распределения и определения настроек алгоритма Метрополиса-Гастингса для семплирования с адаптивной поэтапной подстройкой параметров предложенного распределения.

Ключевые слова: смесь распределений, байесовское оценивание, априорное распределение, алгоритм Метрополиса-Гастингса.

DOI 10.14357/20718632200103

Введение

В современном анализе данных широкое распространение получили вероятностные модели, основанные на смеси различных распределений. Как правило, применяют смеси гауссовских распределений, на таких моделях основаны некоторые методы кластерного анализа [1] и распознавания изображений [2] и текстов [3]. Более продвинутые модели смесей Гамма распределений, Коши или Стьюдента применяются в эконометрике при анализе финаносвых рынков [4].

Задача разделения смеси заключается в оценке параметров входящих в нее распределений. Для смеси нормальных распределений оцениваются математические ожидания и матрицы ковариаций компонент, а также весовые коэффициенты самих компонент. Входными данными для задачи разделения смеси являются наблюдения и предположения о количестве компонент. Подробное описание данной задачи, строгая математическая постановка и качественное сравнение методов её решения приводятся в работе [5].

В данной работе предпринята попытка усовершенствовать метод разделения смеси на основе байесовского оценивания с использованием информативных априорных распределений (Maximal Data Information Prior – MDIP) благодаря предварительной оценке энтропийного интеграла, необходимоого для вычисления

* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проекты 17-29-07079 «офи_м» и 16-29-12901 «офи_м»).

априорного распределения, и детальной настройке алгоритма семплирования.

1. Байесовское оценивание и расчет априорного распределения

Пусть имеется выборка $\boldsymbol{X}$ значений $d-$мерной случайной величины x^d, плотность распределения которой описывается следующим законом:

$$x^d \propto \sum_{i=1}^{k} w_i \mathcal{N}(x^d, \mu_i, \Sigma_i),$$

где $\mu_i = \{\mu_i^1, \mu_i^2, \dots, \mu_i^d\}$ - вектор средних значений i-ой компоненты смеси, Σ_i- его ковариационная матрица, k- количество компонент в смеси, а w_i – положительные коэффициенты этих компонент, причем $\sum_{i=1}^{k} w_i = 1$ и $w_i > 0$.

При фиксированном векторе параметров $\theta = \{\mu_1, \mu_2 \ \dots, \mu_k, \Sigma_1, \Sigma_2, \dots, \Sigma_k, w_1, w_2, \dots, w_{k-1}\}$ и в предположении независимости наблюдений, функция правдоподобия выборки записывается следующим образом:

$$\mathcal{L}(X|\theta) = \prod_{s=1}^{N} p(x_s^d|\theta) = \prod_{s=1}^{N} \sum_{i=1}^{k} w_i \mathcal{N}(x_s^d, \mu_i, \Sigma_i)$$

Задача разделения смеси заключается в оценке параметров θ в уравнении (2) по имеющейся выборке наблюдений, а традиционным методом решения является максимизация правдоподобия [6]. Общее количество искомых параметров квадратично возрастает с ростом размерности задачи, что приводит к необходимости решения многомерной оптимизационной задачи.

Рассмотрим задачу поиска оценок параметров θ с точки зрения теоремы Байеса [7], согласно которой для восстановления апостериорной плотности параметров $p(\theta|X)$ используется функция правдоподобия (2) и некоторая априорная плотность $\pi(\theta)$:

$$p(\theta|X) \propto \pi(\theta)\mathcal{L}(X|\theta)$$

Различают два типа априорных распределений: информативные и неинформативные [8]. Информативные распределения имеют определенную структуру, выбираемую исходя из имеющихся знаний об объекте исследований. Байесовское оценивание в этом случае, фактически, уточняет на основе наблюдений знания, заложенные в априорных распределениях [9]. Другая ситуация с неинформативными априорными распределениями, являющимися, как правило, равномерными или неопределенными в общем виде, но являющимися решением некоторой задачи оптимизации, например, задачи максимизации энтропии [10].

В данной работе используется метод выбора априорных распределений, предложенный А. Зельнером – Maximal Data Information Prior (MDIP) [11]. Согласно методу Зельнера, априорное распределение выбирается следующим образом:

$$\pi(\theta) \propto \exp\big(I(\theta)\big),$$

где энтропийный интеграл $I(\theta)$ вычисляется на основе функции правдоподобия смеси $p(x|\theta)$:

$$I(\theta) = \int_X p(x|\theta) \ln p(x|\theta) dx$$

Проблема использования такого априорного распределения заключается в вычислении интеграла $I(\theta)$, размерность которого совпадает с размерностью смеси, и не имеет аналитического решения при $d \geq 3$. Поэтому, в данной работе предлагается метод поточечной оценки интеграла $I(\theta)$, базирующийся на вспомогательном семплировании нормальных распределений. Подставим функцию правдоподобия одного наблюдения смеси в формулу для $I(\theta)$:

$$\begin{aligned} I(\theta) &= \\ &= \int_X \sum_{i=1}^{k} w_i \mathcal{N}(x^d, \mu_i, \Sigma_i) \ln \sum_{i=1}^{k} w_i \mathcal{N}(x^d, \mu_i, \Sigma_i) dx \\ &= \sum_{i=1}^{k} w_i \int_X \mathcal{N}(x^d, \mu_i, \Sigma_i) \ln \sum_{i=1}^{k} w_i \mathcal{N}(x^d, \mu_i, \Sigma_i)\, dx \\ &= \sum_{i=1}^{k} w_i I_i(\theta), \end{aligned}$$

где $I_i(\theta)$ – одна из компонент интеграла $I(\theta)$.

Таким образом, $I_i(\theta)$ представляет собой интеграл типа $\int f(x)p(x)dx$ для вычисления среднего значения функции $f(x)$ от случайной величины, распределенной по закону $p(x)$. В данном случае

$$I_i(\theta) = \int_X f(x)p(x)dx,$$

$$f(x) = \ln \sum_{i=1}^{k} w_i \mathcal{N}(x^d, \mu_i, \Sigma_i), \ \ p(x) = \mathcal{N}(x^d, \mu_i, \Sigma_i)$$

Соответственно, сгенерировав выборку $X = \{x_i, i = 1, \dots n\}$ из нормального распределения $\mathcal{N}\left(x^d, \mu_i, \Sigma_i\right)$, вместо точного значения $I_i(\theta)$ можно воспользоваться несмещенной и состоятельной оценкой выборочного среднего:

$$\widehat{I_i}(\theta) = \frac{1}{n}\sum_{i=1}^{n} f(x_i).$$

Генераторы многомерных нормальных распределений включены в большинство пакетов статистических расчетов, таких как R, Python, Matlab и пр. Таким образом, задача численного интегрирования для расчета $I(\theta)$ заменяется задачей семплирования нормальных распределений с заданными параметрами.

2. Результаты оценки энтропийного интеграла

Для экспериментов было сформировано несколько модельных наборов данных с заданными параметрами. Оценка интеграла $I(\theta)$ предложенным способом сравнивается с результатом численного интегрирования, доступного при размерности не более 3.

Критерием качества оценки является относительная погрешность и ее стандартное отклонение по результату 500 экспериментов.

$$eps = \frac{\left|I(\theta) - \hat{I}(\theta)\right|}{I(\theta)}$$

Результаты экспериментов представлены в Табл. 1 и Табл.2, и на Рис. 1-Рис. 3. В таблицах приведены результаты оценки энтропийного интеграла для двух- и трехкомпонентных смесей соответственно. В колонках приведены следующие параметры:

- размерности данных;
- объем выборки при семплировании нормального распределения для каждой компоненты;
- средняя относительная ошибка оценки;
- стандартное отклонение относительной ошибки;
- среднее время оценки.

Для наглядности, данные из Табл. 1 представлены в виде графиков на Рис. 1-Рис. 3 для размерности смеси d=1,2,3 соответственно. Графики демонстрируют как убывание средней относительной ошибки оценки с ростом объема выборки для оценивания, так и существенное уменьшение стандартного отклонения оценок, что подтверждает состоятельность таких оценок.

Для сравнения времени вычислений в Табл. 3 приводится время расчета интеграла $I(\theta)$ встроенными средствами численного интегрирования программной среды Matlab 2017a. Для вычислений использовался ПК с 4-х ядерным процессором Intel(R) Core(TM) i7 CPU 920 @ 2.67 GHz. и 12 Gb оперативной памяти, все полученные в ходе выполнения работы результаты являются строго воспроизводимыми при задании начального состояния генератора случайных чисел (rng(2018)).

Как видно, уже при d=2 численное интегрирование требует значительных вычислительных ресурсов. С другой стороны при аналогичных параметрах k и d оценка интеграла на основе

Табл. 1. Результаты оценки энтропийного интеграла для двухкомпонентной смеси нормальных распределений (k=2).

Размерность компонент	Объем выборки	Относительная погрешность	Стандартное отклонение	Время вычислений
d = 1	500	0.000554	0.000201	0.01972
	1000	0.000185	0.000139	0.03843
	5000	9.24E-05	6.5E-05	0.192978
	10000	5.54E-05	4.55E-05	0.385708
d = 2	500	0.000514	0.000206	0.052394
	1000	0.000314	0.000142	0.104876
	5000	0.000114	6.51E-05	0.519604
	10000	5.71E-05	4.59E-05	1.034952
d = 3	500	0.000535	0.000203	0.053446
	1000	0.000424	0.000142	0.105682
	5000	0.000156	6.55E-05	0.527364
	10000	8.92E-05	4.61E-05	1.056118

семплирования выполняется существенно быстрее и со средней относительной погрешностью не более 0,1% (Табл. 2 и Табл. 3) , точность оценки может быть дополнительно повышена посредством увеличения объем выборки, используемой для оценки компонент интеграла.

Табл. 2. Результаты оценки энтропийного интеграла для трехкомпонентной смеси нормальных распределений (k=3)

Размерность компонент	Объем выборки	Относительная погрешность	Стандартное отклонение	Время вычислений
d = 1	500	0.000578	0.00012	0.04442
	1000	0.000129	8.43E-05	0.08817
	5000	0.000129	3.96E-05	0.441074
	10000	0.000193	2.67E-05	0.878998
d = 2	500	0.000464	0.000156	0.117474
	1000	0.000377	0.00011	0.235154
	5000	0.000116	5.02E-05	1.174362
	10000	2.9E-05	3.57E-05	2.352794
d = 3	500	0.000244	0.000144	0.118832
	1000	6.11E-05	0.005037	0.23665
	5000	2.04E-05	0.002258	1.180632
	10000	0.000122	0.001764	2.357184

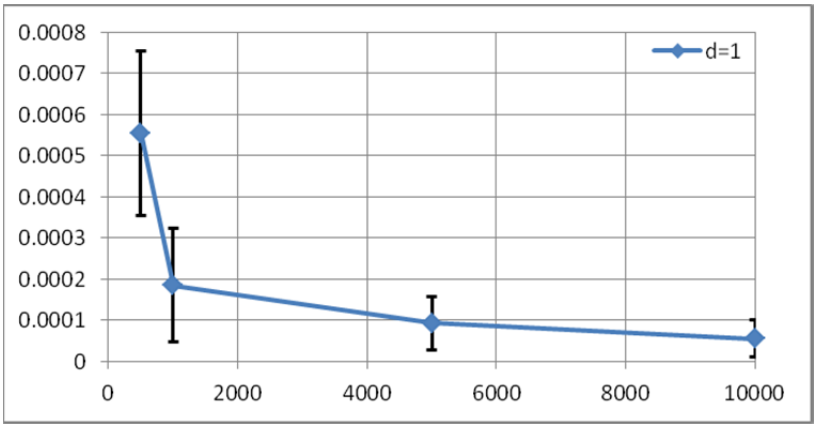

Рис. 1. Относительная погрешность оценки энтропийного интеграла (k=2, d=1)

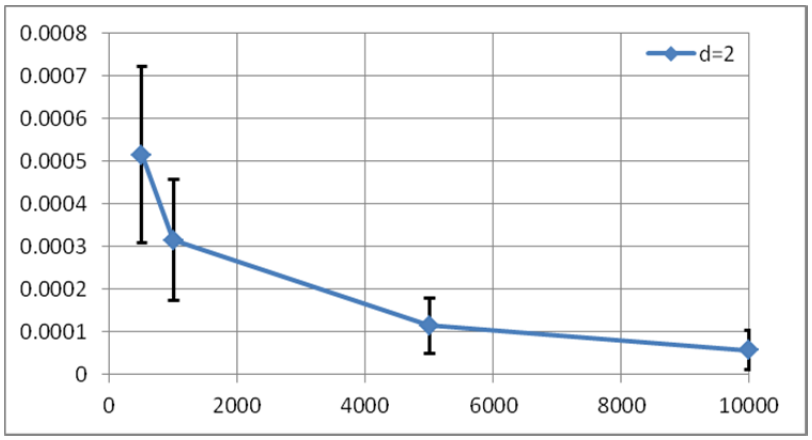

Рис. 2. Относительная погрешность оценки энтропийного интеграла (k=2, d=2)

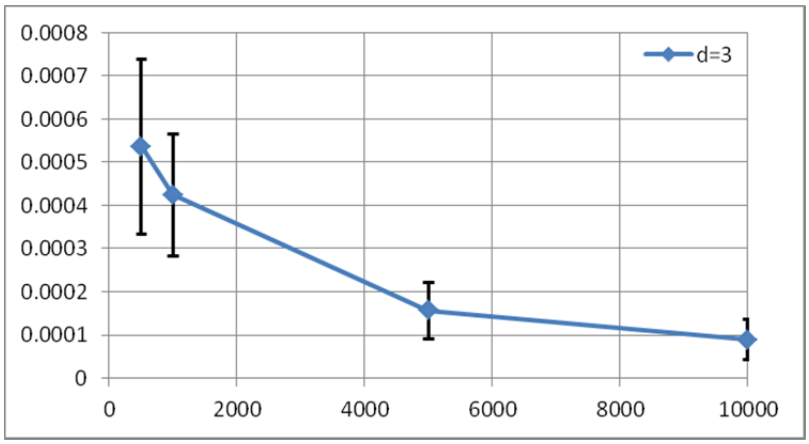

Рис. 3. Относительная погрешность оценки энтропийного интеграла (k=2, d=3)

Табл. 3. Время расчета численных интегралов

Число компонент	Размерность компонент	Время вычислений, сек
k = 2	d = 1	0.094
	d = 2	1.731
	d = 3	128.058
k = 3	d = 1	0.105
	d = 2	5.008
	d = 3	208.817

3. Настройка алгоритма Метрополиса-Гастингса

Предложенный метод оценки энтропийного интеграла используется далее для вычисления априорного распределения параметров, что позволяет реализовать алгоритм поточечного семплирования апостериорного байесовского распределения для разделения компонент смеси.

В качестве алгоритма семплирования используется алгоритм Метрополиса-Гастингса, основанный на построении сходящейся цепи Маркова, на каждой итерации которой происходит генерация новой случайной точки из вспомогательного распределения с последующим решением на основе функции правдоподобия смеси о принятии и переходе к следующей точке.

Основным препятствием к повсеместному применению алгоритма Метрополиса-Гастингса является отсутствие единой универсальной и эффективной стратегии для подбора параметров и предложенного распределения. Данной проблеме при построении марковских цепей в целом и алгоритма Метрополиса-Гастингса в частности посвящено множество статей и монографий [12-14]. Однако единственно верного подхода к ее решению, по-видимому, не существует. В данной работе предлагается метод настройки параметров алгоритма МГ, основанный на итерационном уточнении параметров предложенного распределения.

В первую очередь, для запуска алгоритма Метрополиса-Гастингса необходимо определить начальную точку, с которой начинается построение цепи, а также тип и параметры вспомогательного распределения. И, если выбор начальной точки не представляет особых трудностей, то выбор вспомогательного (предложенного) распределения остается открытым научным вопросом.

Одним из эмпирических подходов к решению данной задачи является выбор в качестве вспомогательного распределения многомерного нормального распределения с параметрами математических ожиданий и ковариационной матрицы, рассчитанными на основе имеющейся выборки наблюдений.

Пусть по выборке наблюдений были рассчитаны выборочные оценки одномерной случайной величины:

$$\hat{\mu} = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad \hat{\sigma}^2 = s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Как известно, для выборки независимых наблюдений оценка среднего является несмещенной и имеет дисперсию

$$D(\bar{x}) = \frac{D(x)}{N} = \frac{\sigma^2}{N} \approx \frac{s^2}{N}.$$

Для того, чтобы рассчитать дисперсию для оценки s^2, вспомним, что нормированная выборочная дисперсия нормальных случайных величин имеет распределение хи-квадрат:

$$(N-1)\frac{s^2}{\sigma^2} \sim \chi^2_{N-1}$$

соответственно,

$$D\left((N-1)\frac{s^2}{\sigma^2}\right) = 2(N-1)$$

или

$$D(s^2) = \frac{2\sigma^4}{N-1} \approx \frac{2(s^2)^2}{N-1}$$

Таким образом, на основе имеющейся выборки из N наблюдений мы вычислим выборочные оценки математического ожидания и дисперсии по каждой компоненте вектора, а также дисперсии самих оценок.

Выборочные оценки используются в качестве стартовой точки для семплирования алгоритмом Метрополиса-Гастингса, в то время как дисперсии выборочных оценок позволяют задать параметры ковариации предложенного распределения, определяющего разброс точек по области, что обеспечивает покрытие всей зоны семплирования.

Следующий этап решения задачи разделения смеси с помощью семплирования заключается в итеративном повторении процедуры семплирования с адаптивной подстройкой параметров предложенного распределения. Идея метода адаптивной подстройки параметров предложенного распределения заключается в итерационной процедуре сравнения эмпирической функции плотности, полученной в результате семплирования выборки, с самим предложенным распределением и изменении его параметров с целью совмещения распределений.

Предположим, что для первого прогона алгоритма МГ использовались стандартные параметры предложенного распределения, например, $\mu_0 = 0$, $\sigma_0 = 1$. В результате семплирования была получена выборка значений случайной величины $X = \{x_i, i = 1,2, \dots N\}$, по которой можно вычислить выборочные оценки математического ожидания и дисперсии генеральной совокупности:

$$\mu_s = E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

$$\sigma_s^2 = D(X) = \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \mu_s)^2.$$

Полученные значения будем использовать в качестве параметров предложенного распределения при следующей генерации, т.е.:

$$q_{s+1}(x) = N(\mu_s, \sigma_s^2).$$

Процесс повторяется до достижения сходимости, т.е. ситуации, при которой выборочные статистки полученной выборки оказываются близки к параметрам использующегося на данном этапе предложенного распределения. Как показывают результаты моделирования, изложенные в следующем разделе, обычно для достижения сходимости достаточно 2-4 итераций. Пример поэтапного изменения предложенного распределения приведен на Рис. 4 для одномерных (слева) и двумерных (справа) распределений.

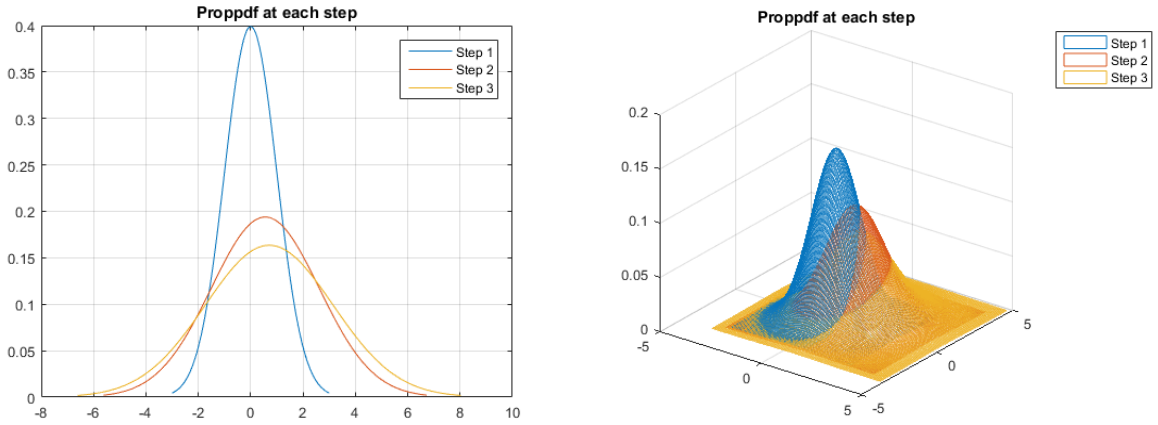

Рис. 4. Пример адаптивной подстройки предложенного распределения в алгоритме МГ для одномерной (слева) и двумерной (справа) задачи

Как показывают эксперименты (Табл. 4, Табл. 5), уже однократное применения описанной подстройки параметров позволяет существенно улучшить оценки максимума, что, в свою очередь, приводит к меньшей относительной ошибке параметров компонент смеси.

4. Результаты разделения смеси с помощью семплирования

Для экспериментов были сгенерированы выборки из смесей нормальных распределений объемом N=30 точек при числе компонент смеси k=2,3 и размерности компонент d=1,2,3, матрицы ковариаций компонент смеси являются диагональными. Объем выборки наблюдений составляет не менее чем $6 * N_\theta$, где N_θ – размерность искомого вектора параметров.

При оценке использовались априорные распределения по методу Зельнера с оценкой энтропийного интеграла по выборке из n=1000 точек для каждой компоненты. Результаты байесовской оценки параметров сравниваются с алгоритмом Expectation Maximization (EM) [15] по средней относительной погрешности определения параметров.

Результаты представлены в Табл. 4 и Табл. 5 для примеров двух- и трехкомпонентных смесей соответственно.

Как видно из результатов экспериментов с модельными данными, EM алгоритм в большинстве примеров выполняется стабильно с приемлемой точностью, что объясняет его всеобщую распространенность. С другой стороны, оценка с помощью семплирования позволяет получить более точный результат и может быть использована в тех приложениях, где время выполнения не является ограничивающим фактором.

Заключение

В работе расмотрена задача разделения смеси нормальных распределений и метод её решения на основе байесовского оценивания и семлирования для поиска точки максимума апостериорного распределения. В качестве алгоритма семплирования используется алгоритм Метрополиса-Гастингса, основанный на построении цепи Маркова по заданному вспомогательному распределению.

Предложены два нововведения при решении данной задачи: во-первых, с целью ускорения вычислений используется оценка энтропийного интеграла, входящего в форму записи информативного априорного распределения, во-вторых, предложен метод адаптивной настройки параметров рассматриваемого распределения для алгоритма Метрополиса-Гастингса.

Табл. 4. Результаты сравнения алгоритмов EM и MH для разделения двухкомпонентной смеси (k=2)

Размерность компонент	Размерность вектора параметров	Объем выборки наблюдений	EM	MH	
				step 1	step 2
d = 1	5	30	0.1409	0.1781	0.0912
d = 2	9	54	0.1027	0.1136	0.0819
d = 3	13	78	0.0941	0.1083	0.0783
d = 5	21	126	0.0819	0.1228	0.0933
d = 10	41	246	0.0403	0.0709	0.0364

Табл. 5. Результаты сравнения алгоритмов EM и MH для разделения трехкомпонентной смеси (k=3)

Размерность компонент	Размерность вектора параметров	Объем выборки наблюдений	EM	MH	
				step 1	step 2
d = 1	8	48	0.3545	0.2705	0.1413
d = 2	14	84	0.2046	0.1879	0.0995
d = 3	20	120	0.1513	0.1394	0.0894
d = 5	32	192	0.0753	0.0948	0.0579
d = 10	62	372	0.0449	0.0644	0.0376

Причем параметры предложенного распределения выбираются только на основе имеющейся выборки, что обеспечивает универсальность предложенного подхода для различных практических задач.

Литература

1. McLachlan, G., and D. Peel. Finite Mixture Models. – Hoboken, NJ: John Wiley & Sons. Inc., 2000.
2. Figueiredo, M.A.T. and Jain A.K. Unsupervised Learning of Finite Mixture Models. // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24(3), pp.381-396, 2012.
3. Reynolds, D.A., Rose, R.C. Robust Text-Independent Speaker Identification using Gaussian Mixture Speaker Models // IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.3(1). pp.72-83, 1995.
4. Brigo Damiano, Mercurio Fabio. Lognormal-mixture dynamics and calibration to market volatility smiles. // International Journal of Theoretical and Applied Finance, vol.5(4), pp.427-452, 2002.
5. Дубнов Ю.А., Булычев А.В. Байесовская идентификация параметров смеси нормальных распределений // Информационные технологии и вычислительные системы. 2017. вып.1. С. 101-111.
6. Dempster A.P., Laird N.M., Rubin D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm.// Journal of the Royal Statistical Society. Series B, vol.39(1), pp.1-38, 1977.
7. John E. Rolph. Bayesian Estimation of Mixing Distributions // The Annals of Mathematical Statistics, vol.39, No.4, pp.1289-1302, 1968.
8. Andrew Gelman. Bayes, Jeffreys, Prior Distributions and the Philosophy of Statistics // Statistical Science, vol.24, No.2, pp.176-178, 2009.
9. Robert E. Kass and Larry Wasserman. The Selection of Prior Distributions by Formal Rules // Journal of the American Statistical Association, vol.91, No.435, pp.1343-1370, 1996.
10. Navid Feroze and Muhammad Aslam. Bayesian Estimation of Two-Component Mixture of Gumbel Type II Distribution under Informative Priors // International Journal of Advanced Science and Technology, vol.53, pp.11-30, 2013.
11. Zellner A. Past and Recent Results on Maximal DataInformation Priors // Texhnical Report, Graduate School of Business, University of Chicago, 1996.
12. Siddhartha Chib, Edward Greenberg. Understanding the Metropolis-Hastings Algorithm // The American Statistician, vol.49, No.4, pp.327-335, 1995.
13. Gilks, W.R. and Roberts, G. O. "Strategies for improving MCMC", in (Gilks, W.R. eds) Markov Chain Monte Carlo in Practice, Chapman & Hall/CRC? 1996.
14. Robert, Christian; Casella, George. Monte Carlo Statistical Methods. Springer, 2004.
15. Dempster A.P., Laird N.M., Rubin D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm.// Journal of the Royal Statistical Society. Series B, vol.39(1), pp.1-38, 1977.

Дубнов Юрий Андреевич. Институт системного анализа Федерального государственного учреждения "Федеральный исследовательский центр "Информатика и управление" Российской академии наук" (ФИЦ ИУ РАН), г. Москва, Россия, Научный сотрудник. Национальный исследовательский университет Высшая школа экономики (НИУ ВШЭ), г. Москва, Россия, Старший преподаватель. Количество печатных работ: 19. Область научных интересов: байесовское оценивание, машинное обучение, принцип максимума энтропии. E-mail: yury.dubnov@phystech.edu

Булычев Александр Викторович. Институт системного анализа Федерального государственного учреждения "Федеральный исследовательский центр "Информатика и управление" Российской академии наук" (ФИЦ ИУ РАН), г. Москва, Россия, Ведущий научный сотрудник. Национальный исследовательский университет Высшая школа экономики (НИУ ВШЭ), г. Москва, Россия, Доцент. Федеральное государственное образовательное бюджетное учреждение высшего образования "Финансовый университет при Правительстве Российской Федерации", Доцент. Кандидат технических наук. Количество печатных работ: 33. Область научных интересов: анализ данных, байесовские методы в статистике и эконометрике. E-mail: bulytchev.isa.ran@gmail.com

On an Approach to Tuning the Metropolis-Hastings Algorithm for the Task of Separating a Mixture of Gaussian Components

Yu. A. Dubnov [I, II], A. V. Boulytchev [I, II]

[I] Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia
[II] Higher School of Economics, National Research University, Moscow, Russia

Abstract The article considers the problem of separating a mixture of Gaussian components, which consists in determining, from available observations, the parameters of the mixture components. An approach to solving this problem is proposed, based on Bayesian estimation using the most informative prior distributions (Maximal Data Information Prior - MDIP). The novelty of the described ap-

proach lies in the use of sample estimates to calculate the prior distribution and determine the settings of the Metropolis-Hastings algorithm for sampling with adaptive stepwise adjustment of the proposed distribution parameters.

Keywords: Gaussian mixture model, Bayesian approach, Prior distribution, Metropolis-Hastings algorithm.

DOI 10.14357/20718632200103

References

1. McLachlan, G., and D. Peel. Finite Mixture Models. – Hoboken, NJ: John Wiley & Sons. Inc., 2000.
2. Figueiredo, M.A.T. and Jain A.K. Unsupervised Learning of Finite Mixture Models. // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24(3), pp.381-396, 2012.
3. Reynolds, D.A., Rose, R.C. Robust Text-Independent Speaker Identification using Gaussian Mixture Speaker Models // IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.3(1). pp.72-83, 1995.
4. Brigo Damiano, Mercurio Fabio. Lognormal-mixture dynamics and calibration to market volatility smiles. // International Journal of Theoretical and Applied Finance, vol.5(4), pp.427-452, 2002.
5. Yu.A. Dubnov, A.V. Boulytchev Bayesian Identification of a Gaussian Mixture Model // Journal of Information Technologies and Computing Systems. 2017. Iss.1. P.101-111.
6. Dempster A.P., Laird N.M., Rubin D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm.// Journal of the Royal Statistical Society. Series B, vol.39(1), pp.1-38, 1977.
7. John E. Rolph. Bayesian Estimation of Mixing Distributions // The Annals of Mathematical Statistics, vol.39, No.4, pp.1289-1302, 1968.
8. Andrew Gelman. Bayes, Jeffreys, Prior Distributions and the Philosophy of Statistics // Statistical Science, vol.24, No.2, pp.176-178, 2009.
9. Robert E. Kass and Larry Wasserman. The Selection of Prior Distributions by Formal Rules // Journal of the American Statistical Association, vol.91, No.435, pp.1343-1370, 1996.
10. Navid Feroze and Muhammad Aslam. Bayesian Estimation of Two-Component Mixture of Gumbel Type II Distribution under Informative Priors // International Journal of Advanced Science and Technology, vol.53, pp.11-30, 2013.
11. Zellner A. Past and Recent Results on Maximal DataInformation Priors // Texhnical Report, Graduate School of Business, University of Chicago, 1996.
12. Siddhartha Chib, Edward Greenberg. Understanding the Metropolis-Hastings Algorithm // The American Statistician, vol.49, No.4, pp.327-335, 1995.
13. Gilks, W.R. and Roberts, G. O. "Strategies for improving MCMC", in (Gilks, W.R. eds) Markov Chain Monte Carlo in Practice, Chapman & Hall/CRC? 1996.
14. Robert, Christian; Casella, George. Monte Carlo Statistical Methods. Springer, 2004.
15. Dempster A.P., Laird N.M., Rubin D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm.// Journal of the Royal Statistical Society. Series B, vol.39(1), pp.1-38, 1977.

Dubnov Yu. A. Federal Research Center "Computer Science and Control" of Russian Academy of Sciences (FRC CSC RAS), Moscow, Russia, Researcher. Higher School of Economics (HSE), National Research University, Moscow, Russia, Senior Lecturer. Scientific publications: 19. Research interests: Bayesian estimation, machine learning, principle of maximum entropy. Corresponding author, E-mail: yury.dubnov@phystech.edu

Boulytchev A. V. PhD. Federal Research Center "Computer Science and Control" of Russian Academy of Sciences (FRC CSC RAS), Moscow, Russia, Leading Researcher. Higher School of Economics (HSE), National Research University, Moscow, Russia, Assistant Professor. Financial University under the Government of the Russian Federation, Assistant Professor. Scientific publications: 33. Research interests: data analysis, bayesian methods in statistics and econometrics. E-mail: bulytchev.isa.ran@gmail.com