ляется более сложным, например, групповым марковским потоком, когда обслуживание на второй фазе является не показательным, а принадлежит к фазовому типу и т.д.

Литература.

1. Жожикашвили В.А., Вишневский В.М. Сети массового обслуживания. Теория и применение к сетям ЭВМ. М.: Радио и связь, 1988.
2. Дудин А.Н., Клименок В.И. Системы массового обслуживания с коррелированными потоками. Мн.: БГУ, 2000.
3. Dudin A.N., Klimenok V.I. Multi-dimensional quasitoeplitz Markov chains // Journal of Applied Mathematics and Stochastic Analysis. 1999. V.12. N 4. P.393-415.
4. Gnedenko B.V., Koenig D. Handbuch der bedienungstheorie. Berlin: Akademie – Verlag, 1983.

ОБ ОЦЕНИВАНИИ ПАРАМЕТРОВ МНОГОМЕРНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ПРИ НАЛИЧИИ «ПРОПУСКОВ»

Е.В. Ковалевский

Научный руководитель – Ю.С. Харин

Белорусский государственный университет

Введение и постановка задачи.

В практических приложениях при решении задач статистического анализа данных часто оказывается полезной модель многомерного нормального распределения. Статья посвящена проблеме оценивания параметров данного распределения по выборке с пропусками.

Пусть $\{X_t, t = 1,...,T\}$ - независимые одинаково распределенные нормальные случайные вектора, $L(X_t) = N_n(\mu, \Sigma)$. Параметры μ, Σ - неизвестны. В выборке $\{X_t\}$ имеются пропущенные значения. $X_t = \left(X_t^{obs}, X_t^{mis}\right)$, где X_t^{obs} - наблюдаемые, а X_t^{mis} - пропущенные компоненты вектора X_t в момент времени t. Из рассмотрения исключаются случаи, когда отсутствие некоторого значения зависит от пропущенных значений наблюдения, однако не исключается возможность зависимости этого факта от присутствующих значений наблюдения (или, в терминах [1], рассматривается случай, когда данные ОПС либо ОС). Задача заключается в оценивании вектора математического ожидания (МО) μ и ковариационной матрицы Σ.

Проблема обработки данных с пропусками и, в частности, данная модель рассматривалась в [1]. В [1] приводится так называемый ЕМ-алгоритм, позволяющий находить точки в пространстве параметров распределения, потенциально обладающие многими оптимальными свойствами, присущими МП-оценкам. Данный алгоритм является приближенным итеративным алгоритмом поиска локального максимума маргинальной функции правдоподобия, которая, вообще говоря, имеет сложный вид даже в случае многомерного нормального распределения и не поддается аналитической максимизации.

Проблема заключается в выборе начального приближения. В [1] подготовлен достаточно полный обзор литературы по обработке данных с пропусками, однако нет ссылок на алгоритмы генерации начальных значений. Приводятся лишь четыре тривиальных варианта выбора начальных значений, два из которых по разным причинам не могут быть признаны удовлетворительными в общем случае. Аналогичные результаты приводятся, например, в [2]. Автором статьи разработан, реализован и исследован новый алгоритм генерации начальных значений для ЕМ-алгоритма.

Алгоритм генерации начальных значений.

Применим ЕМ-алгоритм для различных начальных приближений и затем выберем ту оценку параметров, на которой значение маргинальной функции правдоподобия наибольшее. Отметим, что тривиальное построение равномерной сетки узлов в пространстве возможных значений параметров неприемлемо ввиду большой размерности данного пространства. Затруднительным является также то, что на значения параметров накладывается требование положительной определенности ковариационной матрицы. Поэтому для поиска глобального максимума функции правдоподобия был разработан следующий алгоритм.

1) Начальное приближение для построения первой оценки выполняется по алгоритму, предложенному в [1]. начальные приближения для дисперсий и компонент вектора МО вычисляются по присутствующим данным, а для ковариаций они полагаются равными нулю. Первая оценка $\left\{\mu^*, \Sigma^*\right\}$ параметров распределения определяется как результат однократного выполнения ЕМ-алгоритма

2) Генерируются $M(M+1)-1$ начальных значений для ЕМ-алгоритма, где M, а также k и β, встречающиеся ниже, являются параметрами метода. Вектора МО генерируется из нормального закона с параметрами $\left\{\mu^*, 1/k\,\Sigma^*\right\}$. Ковариационные матрицы генерируется из распределения Уишарта с параметрами $\left\{\Sigma^*, 1/\beta\right\}$, затем умножаются на β.

3) Выполняется ЕМ-алгоритм для всех сгенерированных начальных параметров. Теперь имеем $M(M+1)$ локально оптимальных оценок.

4) Происходит отбор тех M оценок $\left\{\mu_i^{(1)}, \Sigma_i^{(1)}, i=1,\dots,M\right\}$, которые соответствуют максимальным значениям маргинальной функции правдоподобия. Эти оценки остаются неизменными. Остальные M^2 оценок определяются следующим образом: для каждой из отобранных M оценок генерируется по одному начальному приближению так же, как на шаге 2), оставшиеся $M(M-1)$ начальных приближения генерируются как линейная комбинация из двух различных отобранных оценок, по две штуки для каждой пары отобранных оценок:

$$\alpha\Sigma_i^{(1)} + (1-\alpha)\Sigma_j^{(1)}, 0 < \alpha < 1,$$

где α - случайная величина, распределенная по стандартному равномерному закону.

Начальные приближения для вектора МО генерируются аналогично. Шаги 3) и 4) повторяются до сходимости максимального из значений функции правдоподобия.

Результаты компьютерного моделирования.

В качестве тестового примера использовались следующие параметры многомерного нормального распределения:

$$\Sigma = \begin{pmatrix} 1 & -0.9 & 0 \\ -0.9 & 1 & 0.4 \\ 0 & 0.4 & 1 \end{pmatrix}, \mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Для исследования алгоритма генерировалась выборка с пропусками объема 100. Процесс оценивания повторялся 100 раз, в результате имеем выборку из ста наблюдений (оценок) для каждого параметра распределения.

В таблице 1 приведены результаты для метода однократного применения ЕМ-алгоритма к начальному приближению, построенному по алгоритму, предложенному в [1] и описанному выше. В таблице 2 приведены результаты для разработанного алгоритма.

Таблица 1.

Параметр	Значение			Ср. кв. откл.
	Ср.	Мин.	Макс.	
m1	-0,0219	-0,4795	0,2284	0,1366
m2	-0,0664	-0,3732	0,3213	0,1408
m3	-0,0095	-0,3473	0,3118	0,1359
s11	0,9704	0,4738	1,4378	0,1917
s12	-0,6347	-1,0344	0,0252	0,1804
s13	-0,1068	-0,5342	0,4173	0,1807
s22	0,9155	0,5589	1,4758	0,1976
s23	0,3302	-0,1348	0,757	0,1888
s33	0,9829	0,597	1,4791	0,1884

Таблица 2.

Параметр	Значение			Ср. кв. откл.
	Ср.	Мин.	Макс.	
m1	0,0001	-0,2586	0,3084	0,1075
m2	-0,0385	-0,3979	0,3056	0,1339
m3	0,0082	-0,3036	0,2056	0,1078
s11	0,9705	0,6061	1,5215	0,1578
s12	-0,8114	-1,2247	0,0585	0,1921
s13	-0,084	-0,5472	0,6537	0,1818
s22	0,9571	0,5804	1,7467	0,2037
s23	0,3235	-0,3853	0,9198	0,1955
s33	0,9962	0,5403	1,5521	0,1918

Как и следовало ожидать, смещение оценок уменьшилось, что особенно хорошо заметно по оценке ковариации первой и второй компонент s12.

Литература

1. Литтл Р. Дж. А., Рубин Д. Б. Статистический анализ данных с пропусками / Пер. с англ. – М.: Финансы и статистика, 1990. – 336 с.: ил. – (Математико-статистические методы за рубежом).
2. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей: В 3 т. – М.: Финансы и статистика, 1985 – 487с.

ОЦЕНИВАНИЕ ПАРАМЕТРА РЕГРЕССИИ В СЛУЧАЕ ПРОЦЕССА АВТОРЕГРЕССИИ С УСТОЙЧИВЫМИ ИННОВАЦИЯМИ

А.В. Лис
Научный руководитель – П.М. Лаппо
Белорусский государственный университет

Резюме

В настоящей работе используется предположение, что финансовые индексы подчиняются устойчивому распределению. Рассматривается модель процесса авторегрессии первого порядка с инновациями, являющимися симметричными устойчивыми случайными величинами. В этих предположениях получена оценка параметра регрессии.

Устойчивые случайные величины.

Приведем несколько фактов, касающихся устойчивых распределений, которые будут в дальнейшем использоваться. Более подробно эти и другие свойства описаны в [9].

1. Случайная величина X называется устойчиво распределенной, если существуют параметры $0<\alpha<2$, $\sigma>0$, $|\beta|\le 1$, $\mu\in R$, такие, что ее характеристическая функция имеет вид

$$E\exp(i\cdot t\cdot X)=\begin{cases}\exp\left\{-\sigma^{\alpha}\cdot|t|^{\alpha}\cdot\left(1-i\cdot\beta\cdot\operatorname{sgn}(t)\cdot tg\dfrac{\pi\cdot\alpha}{2}\right)+i\cdot\mu\cdot t\right\}, & \alpha\neq 1\\ \exp\left\{-\sigma\cdot|t|\cdot\left(1+i\cdot\beta\cdot\operatorname{sgn}(t)\cdot\dfrac{2}{\pi}\cdot\ln|t|\right)+i\cdot\mu\cdot t\right\}, & \alpha=1\end{cases}$$

Факт принадлежности X к семейству устойчивых распределений записывают в виде $X\sim S_{\alpha}(\sigma,\beta,\mu)$, где параметр α - индекс устойчивости (характеристическая экспонента), β - коэффициент скошенности, σ - параметр масштаба, μ - пасаметр положения.

2. Случайная величина X является $S\alpha S$ (симметричной α-устойчивой), тогда и только тогда, когда $X\sim S_{\alpha}(\sigma,0,0)$

3. Пусть X и Y являются $S\alpha S$, $\alpha>1$, и Γ является спектральной мерой случайного вектора (X,Y). Ковариационная форма X на Y определяется следующим образом:

$$[X,Y]_{\alpha}=\int_{S_2}s_1\cdot s_2^{<\alpha-1>}\cdot\Gamma(ds),$$

где $a^{<p>}=|a|^{p}\cdot\operatorname{sgn}(a)$, S_2 - единичная сфера в R^2

Ковариационная форма линейна по первому аргументу. Линейность по второму аргументу имеет место при выполнении следующих условий:

Пусть X,Y_1,Y_2 являются $S\alpha S$, $\alpha>1$, и Y_1 и Y_2 являются независимыми. Тогда $[X,Y_1+Y_2]_{\alpha}=[X,Y_1]_{\alpha}+[X,Y_2]_{\alpha}$

4. В пространстве симметричных α-устойчивых случайных величин ковариационная форма порождает норму

$$\|X\|_{\alpha}=([X,X]_{\alpha})^{1/\alpha}.$$

Если $X\sim S_{\alpha}(\sigma,0,0)$, $\alpha>1$, то $\sigma=\|X\|_{\alpha}$.

Авторегрессионная модель первого порядка.