пользователя. При желании пользователя доступ к ней будут иметь также психологи, имеющиеся в системе. Стоит отметить, что в силу использования обученной модели на клиентской стороне удалось реализовать алгоритм на клиентской стороне, сделав систему многопользовательской.

При экспериментальной апробации выявилась зависимость от параметров изображения: освещенность, качество, угол поворота, очки. Указанные параметры не являются существенными, так как легко могут быть изменены при настройке.

Считаю, что результат пригоден для наблюдения за эмоциональным состоянием пользователя, а само приложение может применятся психологами для дистанционного наблюдения за пациентами.

Литература

1. *Экман П.* Узнай лжеца // University of Chicago Press. 1965.
2. *Харин Ю. Зуев Н. Жук Е.* Теория вероятностей, математическая и прикладная статистика // БГУ. 2011. С. 330-333.
3. *S. Milborrow, J. Morkel, F.* Nicolls The MUCT Landmarked Face Database // № 6, 1986.
4. *J. Canny* A Computational Approach to Edge Detection // University of Cape Town. 2008.

ПРОГНОЗИРОВАНИЕ ПАРАМЕТРИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ

С. В. Лобач

ВВЕДЕНИЕ

В наиболее общей постановке проблема статистического анализа данных с пропусками может быть сформулирована следующим образом. Имеется многомерная выборка наблюдений $(x_t, 0_t), t = 1,...,T,$ где $x_t = (x_{1t},...,x_{nt})' \in R^n$ – вектор наблюдений, $0_t = (0_{1t},...,0_{nt})' \in R^n$ – вектор шаблона пропусков, координаты которого имеют значение “1”,если в момент времени t соответсвующая координата вектора x_t наблюдается ; “0”, если соответствующая координата вектора x_t не наблюдается. Проблема состоит в том, чтобы по данным с пропусками построить оптимальные статистические выводы относительно пропущенных значений временного ряда $x_t,\ \ t \in 1,2,...,T,$ а также относительно будущих значений времен-

ного ряда $\{x_\tau\}, \tau > T.$ Данной тематике посвящена обширная литература [1, 2, 3, 4].

Заполнение пропусков – наиболее общий и гибкий метод решения статистических задач при наличии пропусков. Например, если $n-$ вектор x_t распределен по многомерному нормальному закону распределения вероятностей $N_n(\mu,\Sigma)$ с математическим ожиданием μ и ковариационной матрицей Σ, то применяют метод, изложенный в работе [5]. Сначала оценивают μ и Σ выборочным средним и выборочной ковариационной матрицей по имеющимся полным наблюдениям, а затем используют эти оценки для вычисления линейной регрессии пропущенных значений по имеющимся.

EM-АЛГОРИТМ

EM-алгоритм описан в работе [2] и предназначен для вычисления оценок максимального правдоподобия параметров по неполным данным. В дальнейшем он получил широкое распространение и стал применяться для решения других задач статистического анализа данных, например, для разделения смесей, для построения процедур кластерного анализа.

EM-алгоритм состоит из двух шагов:

На шаге "M" вычисляют оценку максимального правдоподобия параметра θ модели временного ряда по имеющимся наблюдениям.

На шаге "E" находят условное математическое ожидание пропущенных наблюдений $\{x_{tjmis}\}$ по имеющимся наблюдениям и при условии, что параметры модели оценены на шаге “ММ”.

Этот процесс продолжается до тех пор, пока не удовлетворяются заданные условия сходимости. До сих пор вопрос о сходимости EM-алгоритма остается открытым, получены строгие результаты по сходимости только для некоторых специальных задач статистического анализа данных с пропусками [6].

ВРЕМЕННЫЕ РЯДЫ НА ОСНОВЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ И ФИЛЬТР КАЛМАНА

Предположим, что имеется временной ряд $x_t,\ \ t \in 1,2,...,T,$ т.е. совокупность случайных наблюдений в последовательные моменты времени векторов x_t .Эти переменные описывают текущее состояние некоторой системы. Обычно, переменные состояния не наблюдаются точно, а реги-

стрируются другие переменные y_t, связанные функциональной или статистической зависимостью с переменными состояния x_t.

Моделью в пространстве состояний называется система, состоящая из двух векторных линейных уравнений :

$$x_{t+1} = Fx_t + \varepsilon_t, \quad \varepsilon_t \sim N_n(0, Q); \tag{1}$$

$$y_t = Hx_t + \eta_t, \quad \eta_t \sim N_m(0, R); \tag{2}$$

$$x_0 \sim N_n(x_0, P_0), \quad t = 1,2,\ldots,T,$$

где F, H, Q, R – матрицы параметров размеров $(n \times n)$, $(m \times n)$, $(n \times n)$, $(m \times m)$ соответственно, в общем случае они неизвестны либо являются известными функциями от параметров. Начальные условия α_0, P_0 предполагаются известными. При таком определении временного ряда на основе модели в пространстве состояний (1), (2) нетрудно показать, что случайные векторы $(x_t, y_t), t = 1,\ldots,T,$ являются дискретным марковским процессом. Уравнение (1) называется уравнением состояний, а уравнение (2) – уравнением измерений. Случайные векторы $(\varepsilon_t, \eta_{te}), t = 1,\ldots,T,$ являются последовательностью независимых гауссовских случайных векторов с нулевыми математическими ожиданиями и ковариационными матрицами Q и R.

Оказывается, что большинство известных параметрических моделей временных рядов могут быть представлены в форме моделей в пространстве состояний (1), (2). Главным ограничением модели (1),(2) является линейность модели. Элементы матрицы F и H могут быть нестационарными, уравнения (1) и (2) могут быт нелинейными.

На практике часто используется ARMA(p,q)–модель временного ряда:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=0}^{q} \beta_j \varepsilon_{t-j}. \tag{3}$$

Временной ряд (3) может быть представлен следующим образом в форме модели в пространстве состояний

$$x_t = \begin{pmatrix} \alpha_1 & 1 & 0 & \cdots & 0 \\ \alpha_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{k-1} & 0 & 0 & \cdots & 1 \\ \alpha_k & 0 & 0 & \cdots & 0 \end{pmatrix} x_{t-1} + \begin{pmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_{k-2} \\ \beta_{k-1} \end{pmatrix} \varepsilon_t, \tag{4}$$

$$y_t = (1, 0, \ldots, 0)\, x_t,$$

где $k = \max\{p, q+1\}$.

Заметим, что представление временных рядов в форме моделей в пространстве состояний является неоднозначным. Модель (4) является формой представления $AR(p)-$модели временного ряда в пространстве состояний, удобной для решения задач прогнозирования. Для решения задачи оценивания параметров временного ряда более удобной является другая форма представления в пространстве состояний:

$$x_t = x_{t-1};$$
$$y_t = (y_{t-1}, y_{t-2,\ldots,y_{t-p}})x_t + \varepsilon_t$$

где $x_t = (\alpha_1, \alpha_2, \ldots, \alpha_p)'$.

АНАЛИЗ НЕПОЛНЫХ ДАННЫХ

Пусть только некоторые компоненты вектора y_t наблюдаются в момент времени t: $i_1(t) < i_2(t) < \ldots < i_{l_t}(t)$, тогда наблюдаемым в момент времени t является вектор $y_t^* = S_t y_t$, где S_t является матрицей размера $l_t \times m$, которая принимает значение “1” в позициях $(1, i_1(t), \ldots, (l_t, i_{l_t}(t)))$ и “0” – в других. В этом случае модель (1),(2) сводится к модели:

$$x_t = Fx_{t-1} + \varepsilon_t;$$

$$y_t^* = H^* x_t + \eta_t^*,$$

где $y_t^* = S_t y_t, H^* = S_t H, \eta_t^* = S_t \eta_t$.

Заметим, что полученная модель является моделью с полными данными, к ней может быть применен фильтр Калмана [7, 8].

Теорема. Для модели временного ряда с пропусками в форме модели в пространстве состояний (1), (2) оптимальная оценка $x_{t|t}$ вектора x_t по наблюдениям y_s^*, $s = 1,2.,,,t$, определяется фильтром Калмана с заменой y_t на y_t^*.

Прогнозирование на шаг вперед осуществляется по формулам :

$$x_{t|t-1} = Fx_{t-1|t-1}, \quad y_{t|t-1} = Hx_{t|t-1}.$$

Литература

1. *Green W.* Econometric Analysis. New York, 2000.
2. *Литтл Р., Рубин Д.* Статистический анализ данных с пропусками. М., 1991.
3. *Shafer J.* Analysis of incomplete data. London, 1997.
4. *Харин Ю.С.* Оптимальность и робастность в статистическом прогнозировании. Мн., БГУ, 2008.
5. *Hurtly H., Hocking R.* The analysis of incomplete data. //Biometrics, 1971, v.27, p.783–808.
6. *Jeff Wu C.F.* On the Convergence Properties of the EM Algorithm//The Annals of statistics, Vol. 11, No 1, 1983, p.95–103.

7. *Kalman R., Busy R.* New results in linear filtering and prediction theory. //Transaction ASME Journal of Basic Engineering. Series D, v.83, p.95–108.
8. *Harvey A., Koopman S., Shepard N.* State space and unobserved component models. Theory and applications. Cambridge University Press, 2004.

К ВОПРОСУ ОБ АВТОМАТИЗАЦИИ РАСПАРАЛЛЕЛИВАНИЯ ПРОГРАММ

В. Д. Львович, А. Н. Перемотова

ВВЕДЕНИЕ

На сегодняшний день практически все высокопроизводительные системы используют преимущества такой технологии, как параллелизм. И это неслучайно, ведь данная технология позволяет наиболее эффективно использовать мощности современных вычислительных систем. Существование большого количества отлаженного и активно используемого последовательного кода «заставляет мечтать» о его автоматической трансформации в параллельный код. Возможность автоматического распараллеливания сулит очевидные выгоды:

- сокращение времени разработки параллельных программ;
- удешевление процесса разработки параллельных программ;
- снижение требований к квалификации программистов;
- повышение надежности разрабатываемых программ;
- упрощение переноса многих разработанных программ на параллельные компьютеры;
- эффективное проектирование программно-аппаратных комплексов с учетом новых методов распараллеливания.

Цель работы – помочь пользователю в разработке параллельного приложения. В результате исследования были разработаны программные инструменты, позволяющие упростить разработку многопоточных программ. Инструменты позволяют автоматически получить параллельную программу одним из двух способов:

- из последовательной программы посредством расстановки директив OpenMP;
- задать параметры задачи поиска и получить параллельные реализации в нескольких вариантах.

ИНСТРУМЕНТ ANTLR

В работе был использован известный и популярный инструмент ANTLR, который представляет собой генератор анализаторов. ANTLR позволяет определять правила, по которым лексический анализатор