

Journal of Financial and Actuarial Mathematics and Management

A joint edition of the Institute of Information Sciences and Institute of Social Sciences of Perun Academy
Publishing Company: PSP Publishing
ISSN: 2367-5810
Issue 8, Year VIII (2020)
Founding Editor and Editor-in-Chief: Pavel Stoynov

Editorial Board: Hiroyuki Ono (Japan), Yasukazu Yoshizawa (Japan), Naoyuki Ishimura (Japan), Boris Melnikov (Russia), Elena Melnikova (Russia), Goran Karanovic (Croatia), James Osondu (Nigeria)

Contents

Y. Algani. Practical Comparison between Reweighted
Nadaraya-Watson and Local Kernel Mode Estimators……...…….3
P. Stoynov. Databases of anomalous traffic for estimation
of intrusion detection…….……..11
Y. Algani, W. Hibi. Introductory Survey
for Recent Restrictive Approximations and
Applications of Solutions of Initial Boundary Value Probles for PDE……..….14

Practical Comparison between Reweighted Nadaraya-Watson and Local Kernel Mode Estimators

Yousef Methkal Abd Algani,
Department of Mathematics, Sakhnin College, Israel.
Department of Mathematics,
The Arab Academic college for education in Israel
yosefabdalgani@gmail.com
https://orcid.org/0000-0003-2801-5880

Abstract.
The main aim of this research has been to improve and develop the nonparametric kernel estimation of a conditional mode under some regularity conditions. Additionally, the choice of bandwidth depended on n and the location and called local variable bandwidth. Re-Weighted Nadaraya-Watson, $\hat{M}_{RNW}(x)$ and the local variable kernel estimator, $\hat{M}_{L}(x)$ are compared theoretically and practically. The findings of the comparison indicates that $\hat{M}_{L}(x)$ is better than $\hat{M}_{RNW}(x)$, and that's refers to the strong effect of the bandwidth h(x) as a smoothing parameter in the local variable estimation. Several simulated and real data applications were used to prove the accuracy of the results obtained by the Reweighed Nadaraya-Watson estimator and Local Variable Kernel estimators

Key words: Practical Comparison - Reweighted Nadaraya-Watson- Local Kernel Mode Estimators

Introduction

Kernel estimation refers to a general class of techniques for non–parametric estimation of functions. In comparison to parametric estimators where the estimator has a fixed functional form (structure) and the parameters of this function are the only information we need to store, non-parametric estimators have no fixed structure and depend upon all the data points to reach an estimate (Raid B. Salha1 & Hazem I. El Shekh Ahmed**, 2015).**

A kernel is a weighting function used in non-parametric estimation techniques. Kernel density estimation is absolutely dependent on kernels to estimate random variables' density functions, or in kernel regression to estimate the conditional expectation of a random variable. Kernels are also used in time-series applications. (Nan Yang, et al, 2019)

Kernel density estimation is an alternative statistical intensive method, which involves smoothing the data while saving the overall structure. It is a good method of reconstructing an unknown population from a random sample of data, overcomes the problems of histograms, and has many applications in statistics. Kernel density estimation is considered as a most powerful method that gives accurate results and decreases the expected errors (De Gooijer, J.and Zerom, D. 2003).

There are certain standard notations it would be convenient to mention in the presentation to be undertaken here. If, we have a random sample $X_1, X_2, ..., X_n$ from a continuous univeriate distribution with a probability density function (pdf), that function – which we will be try to estimate, will be denoted here by f. The probability density function may be considered to be one of the most important concepts in statistics. Specifying the probability density function f gives a natural description of the distribution of X, and allows probabilities associated with X to be found from the relation:

$$P(a < x < b) = \int_a^b f(x)\, dx \quad \text{for all } a < b \tag{1}$$

The symbol $\hat{f}$ will be used to denote the estimator of the probability density function f.

In addition and without loss of generality we will integrate the functions over the real line R, using the symbol $\int f(.)$.

The research problem

This section discusses the development of statistical methods, leading to the kernel density estimation method and A Comparison of Reweighted Nadaraya-Watson and Local Kernel Mode Estimators. To address this problem, the study tries to answer the following questions

1. What is the Kernel estimation?
2. What is the of Reweighted Nadaraya-Watson?
3. What are differences between Reweighted Nadaraya-Watson and Local Kernel Mode Estimators?

Kernel density estimation

The kernel density estimation involves placing a symmetrical surface over each point, evaluating the distance from the point to a reference location based on a mathematical function, and summing the value of all the surfaces for that reference location. This procedure is repeated for all reference locations (Fan, J. and Q.W. Yao (2005).

Briefly the kernel estimator can be viewed as a sum of bumps placed at the observations. The key point to understand how the nonparametric estimators work for the probability density function is to focus first on the definition of this curve (Fan, J., and T. H. Yim, 2004). The limit concept of ratio between probability mass in the neighborhood of a point and the "size" of that neighborhood plays an important role when considering what a simple histogram is doing (Hall, P., J. S. Racine, and Q. Li,2004)

This idea can be easily extended to obtain the moving histogram (or Naïve estimator) and, finally, the kernel density estimator by just giving different weights to the data according to their proximity to the interest point (Salha, R. and Ioannides, D. 2004). A mathematical expression for the well-known Parzen kernel density estimator can be then introduced:

1. Properties of the Kernels

In this section we discuss some essential properties of kernels and give examples of common kernel functions.

These properties and functions comprise the cornerstone of derivations related to kernel density estimation.

A kernel is a piecewise continuous function, satisfies the condition,

$$\int K(x)\,dx = 1 \tag{3.1}$$

Moreover, the kernel is always considering to be symmetric around zero, and need not have bounded support. In most applications K is considered to be a positive unimodal probability density function on R.

2. Properties of the Kernel Density Estimators

The properties of the kernel density estimator

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \tag{3.2}$$

may be listed as follows:

i. The unknown density function $f(x)$ has continuous second derivative $f''(x)$.

ii. The bandwidth $h = h_n$ satisfies $\lim_{n\to\infty} h = 0,$ and $\lim_{n\to\infty} nh = \infty.$

iii. The kernel K is abounded probability density function of order 2 and symmetric about the origin.

iv. $\int tK(t)\,dt = 0$ and $\int t^2 K(t)\,dt = k_2 \neq 0$. (K<∞)

v. $E(|Yi \mid \delta\ |Xi = U) \leq M < \infty$ for some $\delta > 2$, in a neighborhood of x

The estimator depends on the bandwidth $h > 0$ which acts as a tuning parameter. For large bandwidth h, the estimate $\hat{f}(x)$ tends to be very slowly varying as a function of x, while small bandwidths will produce a more highly variable function estimate. The positivity of the kernel function $K(u)$ guarantees a positive density estimate $\hat{f}(x)$ and the normalization $\int K(x)\,dx = 1$ implies that $\int \hat{f}(x)\,dx = 1$ which is

necessary for $\hat{f}(x)$ to be a density function. Typically, the kernel function $K(u)$ is chosen as a probability density which is symmetric around 0.

3. Kernel Estimation of a Conditional PDF

The goal of conditional density estimation is to construct a density estimate $\hat{f}(y|x)$ for a dependent variable y , conditional on a vector of variables x . In this section, we investigate various types of nonparametric kernel methods for estimating the conditional density function $f(y|x)$ of a random variable y given a random vector x.

Assume that $(X_1,Y_1),(X_2,Y_2),.....,(X_n,Y_n)$\in are an independent identically distributed (i i d) sample from real-valued random variables (X,Y) sitting on a given probability space with random variables that share a joint probability density function pdf $f(x,y)$. The marginal probability density function pdf of X_1 is $g(x)=\int f(x,y)dy$, and the conditional probability density function PDF of Y_1 , given $X_1=x$, is

$$f(y|x)=\frac{f(x,y)}{g(x)} \tag{3.3}$$

Conditional density estimation was introduced by Rosenblatt (1969). A bias correction was proposed by Hyndman, et al. (1996). Fan and Gijbels (1996) proposed a direct estimator based on local polynomial estimation and Yao (2003). Bandwidth selection rules have been proposed by Bashtannyk and Hyndman (2001), Fan and Yim (2004), and Hall et al. (2004). The related problem of conditional distribution estimation is examined in a paper by Hall et al. (1999).

Now, we introduce the basic equation of the kernel conditional density estimation (KCDE). In this case, interpolation must happen in both the x and y directions gives us a double kernel estimator, K_1 and K_2. The Kernel function $K(u)$ is assumed to be a Borel symmetric function, and $\{h_n\}$ is a sequence of positive numbers converging to zero. In this study, we use $\hat{f}(x,y)$, $\hat{g}(x)$ and $\hat{f}(y|x)$ as estimators of $f(x,y)$, $g(x)$ and $f(y|x)$ respectively, where

$$\hat{f}(x,y)=(nh_n^2)^{-1}\sum_{j=1}^{n}K_1\left(\frac{x-X_j}{h_n}\right)K_2\left(\frac{y-Y_j}{h_n}\right)$$

$$\hat{g}(x)=(nh_n)^{-1}\sum_{j=1}^{n}K\left(\frac{x-X_j}{h_n}\right) \tag{3.4.}$$

and,

$$\hat{f}(y|x)=\frac{\hat{f}(x,y)}{\hat{g}(x)}.$$

To construct the proposed conditional density estimator, we first consider two existing kernel-based smoothers of the conditional density, and discuss bandwidth selection for the kernel estimator of conditional density with one explanatory variable. There are several techniques used to investigate some topics of estimation properties. The methods are compared and a practical bandwidth selection strategy which combines the methods is proposed. The methods are compared using two simulation studies and a real data set. The next section deals with methods that intersect and refine the common method used in non-parametric conditional estimation that is widely known as the Nadaraya-Watson (NW) estimator.

Comparison of Re-Weighted Nadaraya-Watson and Local Kernel Mode Estimators

According to Cheng, and Wu (2007) the estimation of the conditional mode is a matter of both theoretical and practical interest. Suppose that $(X_1,Y_1),(X_2,Y_2),...,(X_n,Y_n)$ are iid random variables distributed as (X,Y) with joint pdf $f(x,y)$. The marginal pdf of X is given by $g(x)=\int_{-\infty}^{\infty} f(x,y)dy$. Also suppose that the conditional pdf of Y given $X=x$, $f(y|x)$ is uniformly continuous in y, and possesses a unique conditional mode $M(x)$ which is defined as

$$f(M(x)|x)=\max_{-\infty<y<\infty} f(y|x) \tag{4.1}$$

1. Basic Properties and Formulas

Samanta and Tharaneswarar (1990) considered the problem of estimating the conditional mode using the Nadaraya-Watson (NW) estimator of the conditional density function $f_{NW}(y|x)$ which is defined as

$$\hat{f}_{NW}(y|x)=\frac{\sum_{i=1}^{n} K_h(x-X_i)K_h(y-Y_i)}{\sum_{i=1}^{n} K_h(x-X_i)} \tag{4.2}$$

where K is a kernel function, $h=h_n$ is a sequence of positive number converges to zero and it is called the bandwidth, and where $K_h(.)=\frac{1}{h}K\left(\frac{.}{h}\right)$.

Estimating the conditional density function is such a key aspect in the estimation of the conditional mode that the only way of improving the latter estimation is by improving the former.

Since the Nadaraya-Watson estimator suffers from some disadvantages of producing rather large bias and boundary effects, two new estimators of the conditional mode has been proposed by improving the estimation of the Nadaraya-Watson estimator of the conditional density function.

The first estimator is proposed in this research is obtained by using the Re-Weighted Nadaraya-Watson (RNW) estimator as a weighted version of the NW estimator, which combines the better sides of the local linear LL estimators such as bias reduction and no boundary effects and to preserved the property of the NW estimator is always a distribution function.

The RNW conditional density estimator is defined as

$$\hat{f}_{RNW}(y|x)=\frac{\sum_{i=1}^{n} \tau_i(x)(X_i-x)K_h(y-Y_i)}{\sum_{i=1}^{n} \tau_i(x)K_h(x-X_i)} \tag{4.3}$$

where $\tau_i(x)$ denote the probability like weights with properties that

$$\left.\begin{aligned} \tau_i(x)\geq 0, \quad \sum_{i=1}^{n}\tau_i(x)=1, \\ \sum_{i=1}^{n}\tau_i(x)(X_i-x)K_h(x-X_i)=0 \end{aligned}\right\} \tag{4.4}$$

The role of $\tau_i(x)$ is to adjust the (NW) weights such that the resulting conditional density estimator resembles that from LL estimators.

For more details, see De Gooijer and Zerom (2003).

The second estimator is obtained by using the so-called local variable kernel density estimator, $f_L(y|x)$, which depends on a different bandwidth $h(x)$, for each point x at which $f(y|x)$ is

estimated. The bandwidth $h(x)$ can be chosen such that $h(x) \propto \frac{1}{\hat{f}(x)}$, where $\hat{f}(x)$ is some pilot estimator of $f(x)$ at the point x . See Wand and Jones (1995, pp. 41). It is defined as

$$\hat{f}(y|x) = \frac{\frac{1}{h(y)} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h(x)}\right) K\left(\frac{y - Y_i}{h(y)}\right)}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{h(x)}\right)} \tag{4.5}$$

Two new estimators $\hat{M}_{RNW}$ and $\hat{M}_L$ of the conditional mode function were proposed in

$$f(\hat{M}_{RNW}(x)|x) = \max_{y \in R} \hat{f}_{RNW}(y|x) \tag{4.6}$$

and,

$$f(\hat{M}_L(x)|x) = \max_{y \in R} \hat{f}_L(y|x) \tag{4.7}$$

2. Theoretical Comparison

Let us donate the Re-Weighted Nadaraya-Watson by , $\hat{M}_{RNW}(x)$ and the local variable kernel estimator by, $\hat{M}_L(x)$. And In this section, we offer a brief theoretical comparison of $\hat{M}_{RNW}(x)$ and $\hat{M}_L(x)$ the performance of the estimator has been tested using the mean squared errors (MSE) and by looking at their asymptotic mean squared errors (MSE's) as $n \to \infty$.

The assumptions underlying these results are standard ones.

Let us supposed that the 3rd derivative of $f(y|x)\ exist, then:+\ \partial$
$f^{(i,j)}(y|x) = \partial^{i+j} f(y|x)/\partial x^i \partial y^j$

The asymptotic bias and variance of $\hat{M}_{RNW}(x)$ and $\hat{M}_L(x)$ are listed below.

$$MSE = \left(\hat{M}_L(x)\right) = \frac{1}{2}\left[h^2(x) f^{(2,1)}(M(x)|x) + h^2(y) f^{(0,3)}(M(x)|x)\right] \int u^2 K(u) du$$

$$Var\left(\hat{M}_L(x)\right) = \frac{f(x, M(x)) \int\int \left(K(u) K^{(1)}(v)\right)^2 du dv}{n h(x) h^3(y) \left[f^{(0,2)}(x, M(x))\right]^2}$$

$$MSE\left(\hat{M}_{RNW}(x)\right) = \frac{1}{2} h^2 \left[f^{(2,1)}(M(x)|x) + f^{(0,3)}(M(x)|x)\right] \int u^2 K(u) du$$

$$Var\left(\hat{M}_{RNW}(x)\right) = \frac{f(x, M(x)) \int\int \left(K(u) K^{(1)}(v)\right)^2 du dv}{n h^4 \left[f^{(0,2)}(x, M(x))\right]^2}$$

The two asymptotic biases and variances differ by the bandwidth, thus they are the same when $h = h(x) = h(y)$, as will occur when the choice of the local variable bandwidth is

$$h(x) = \frac{h}{\hat{f}(x)} \quad \text{and} \quad \hat{f}(x) = 1.$$

This means as the value of $\hat{f}(x)$ closer to one, the two asymptotic MSE are closer.

In general, however, the MSE of the two estimators are different and depend on the relation between the local variable and constant bandwidth.

3. Practical Comparison

Our study of the practical differences between the local variable kernel (LVKES), $\hat{M}_L(x)$ and the Re-Weighted Nadaraya-Watson estimator (RNW), $\hat{M}_{RNW}(x)$ of the conditional mode is based on applications to real life data and simulation study. The standard Gaussian kernel function is used throughout the applications.

3.1. Real Life Data

Consider the Ethanol data, which describes the relationship between the predictor E (Ethanol) and the response NOx (Nitric Oxide) (Source of the data is in S-Plus program). Figure 1 indicates the relationship is not linear. The regression relation estimated by using the two estimators $\hat{M}_L(x)$ and $\hat{M}_{RNW}(x)$. A scatter plot of the data together with the graphs of the two estimations is shown in Figure 1. Figure 1 indicates that $\hat{M}_L(x)$ fits the data better than $\hat{M}_{RNW}(x)$.

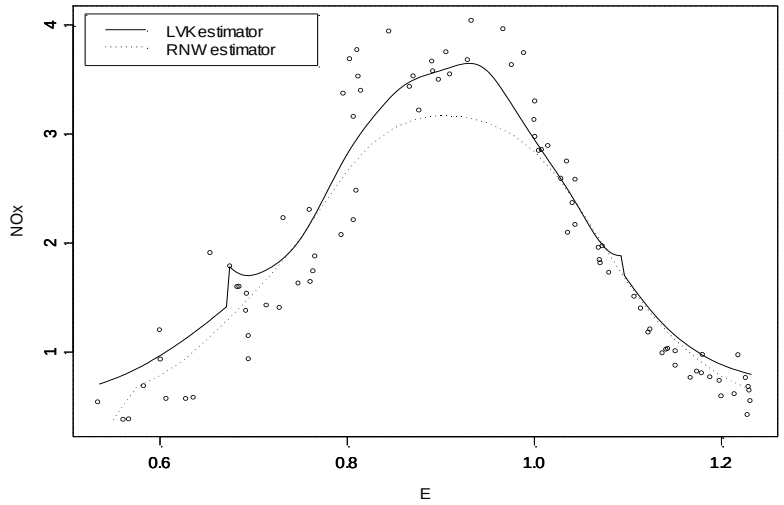


Figure 4.1: (Ethanol Data) Comparison between $\hat{M}_L(x)$ and $\hat{M}_{RNW}(x)$

3.2. Simulation Study

A direct comparison between the two estimators is now given using a simulation data. 200 data points were simulated from the model $y = \sin 2\pi(1-x)^2 + xe$, where $e \approx N(0,1)$, and $x \approx uniform[0,1]$. A perfect smooth would recapture the original signals $y = \sin 2\pi(1-x)^2$. A scatter plot of the simulated data together with the graphs of the perfect smooth and the two estimations are shown in Figure 2. Also, the performance of the two estimations is tested using the mean square error (MSE), which is defined as

$$\text{MSE} = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}$$

where $\hat{y}$, y stand for the predicted values and the actual values respectively and n is the size of the data. The results of this comparison indicates that MSE($\hat{M}_L(x)$)= 0.0145 and MSE($\hat{M}_{RNW}(x)$)=0.0335. This comparison indicates that the $\hat{M}_L(x)$ estimator is better than $\hat{M}_{RNW}(x)$ estimator.

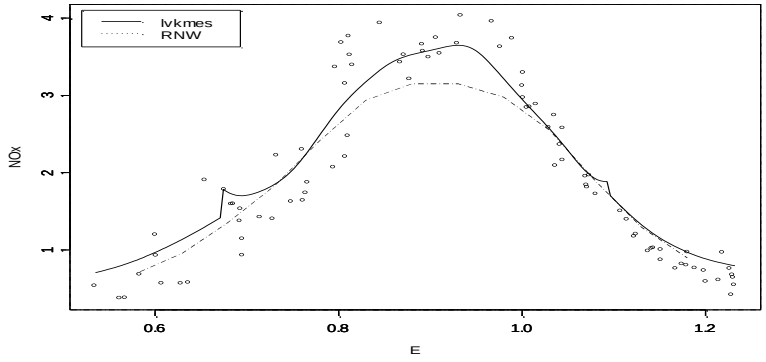


Figure 4.2 (Simulation) Comparison between $\hat{M}_L(x)$ and $\hat{M}_{RNW}(x)$ $\hat{M}_{RNW}(x)$

Conclusion and Future Work

The kernel estimator depends on the bandwidth h and the kernel function, we presented the Reweighted Nadaraya Watson estimator as a refinement of Nadaraya Watson estimator, by reweighing the probability weights $\tau_i(x)$.

The choice of bandwidth depended on n and the location and called local variable bandwidth.

Re-Weighted Nadaraya-Watson, $\hat{M}_{RNW}(x)$ and the local variable kernel estimator, $\hat{M}_L(x)$ are compared theoretically and practically. The comparison indicates that $\hat{M}_L(x)$ is better than $\hat{M}_{RNW}(x)$, and that's refers to the strong effect of the bandwidth h(x) as a smoothing parameter in the local variable estimation.

Several simulated and real data applications were used to prove the accuracy of the results obtained by the Reweighed Nadaraya-Watson estimator and Local Variable Kernel estimators.

In future research of these topics, it may prove fruitful to derive such refined estimators using dependent data, and different types of kernel function formulas.

Also, more work may be done to see if it is possible to obtain a refined estimator by combining the advantages the two previous improvements of the (RNW) and (lvke) estimators in a one new estimator.

References

1. Cheng, M.-Y., L. Peng, and J.-S. Wu (2007): "Reducing Variance in Univariate Smoothing," The Annals of Statistics, 35(2), 522–542.

2. De Gooijer, J.and Zerom, D. (2003). On Conditional Density Estimation. Statistical Neerlandica, Vol. 57, 159-176.

3. Fan J. and Yao Q. (2003), Nonlinear Time Series: Nonparametric and Parametric Methods. Springer-Velage

4. Fan, j. and I. Gijbles (1996), Local Polynomial Modeling and its Applications. London: Chapman and Hall.

5. Fan, J. and Q.W. Yao (2005), Nonlinear Time Series: Nonparametric and Parametric Methods. Springer.

6. Fan, J., and T. H. Yim, (2004). A crossvalidation method for estimating conditional densities. Biometrika, Vol.91, 819-834

7. Hall, P., J. S. Racine, and Q. Li (2004), Cross-validation and the estimation of conditional probability densities. Journal of the American Statistical Association. Vol. 99, (468), 1015-1026

8. Hyndman R. J. and Q. Yao, (2002), Nonparametric Estimation and Symmetry Tests for Conditional Density Functions, *Journal of Nonparametric Statistics*, Vol. 14, Issue 3, 259 – 278

9. Hyndman R.J., D. Bashtannyk, G.K.Grunwald (1996), Estimating and visualizing conditional densities. Journal of Computational and Graphical Statistics, Vol.5,315-336.

10. Nan Yang, et al ,(2019**)** Adaptive Nonparametric Kernel Density Estimation Approach for Joint Probability Density Function Modeling of Multiple Wind Farms**,** Energies 2019, 12, 1356; doi:10.3390/en12071356 www.mdpi.com/journal/energies

11. Raid B. Salha1 & Hazem I. El Shekh Ahmed**, (2015)** Reweighted Nadaraya-Watson Estimator of the Regression Mean**,** International Journal of Statistics and Probability; Vol. 4, No. 1

12. Rosenblatt, M. (1956), Remarks of some nonparametric estimates of a density function, Annals of Mathematical Statistics, Vol. 27, 642 – 669.

13. Salha, R. and Ioannides, D. (2004). Joint asymptotic distribution of the estimated conditional mode at a finite number of distinct points. The national statistical conference, 14-18 Aprile, Lefkada, Greece.

14. Samanta, M., and Thavanesmaran, A. (1990). Nonparametric estimation of the conditional mode. Communications in Statistics. Theory and Methods, Vol. 19, 4515-4524.

15. Wand M. and Jones M., (1995), Kernel Smoothing. Chapman and Hall. ISBN-0:412552701.

Databases of anomalous traffic for estimation of intrusion detection

Pavel Stoynov
Technicla University – Sofia, , Bulgaria
Perun R&D Academy,
Sofia, Bulgaria

Abstract.

Intrusion Detection Systems (IDS) are IT security systems for detecting hostile activity in IT networks with different use (financial, trade, administration, scientific etc.). When these systems also prevent the hostile activity, they are called Intrusion Detection and Prevention Systems (IDPS). The models for intrusion detection can be divided into two groups:

1. Signature verification. In this case, the system detects previously registered intrusion signature.

2. Anomaly detection. In this case, The system looks for abnormal behavior.

Contemporary IDS usually include both signature verification and anomaly detection realized by rule-based expert system and statistical module correspondingly.

Anomaly detection IDS are mainly in research stage today.

This paper aims at presenting some popular databases of anomalous traffic for estimation of Intrusion Detection Systems (IDS). KDD'99, MetroSec and ADFA-IDS are presented.

Key words: Intrusion Dtection Systems (IDS), statistical evaluation of IDS, KDD'99, MetroSec, and ADFA-IDS

Introduction

Intrusion Detection Systems (IDS) are IT security systems for detecting hostile activity in IT networks with different use (financial, trade, administration, scientific etc.). When these systems also prevent the hostile activity, they are called Intrusion Detection and Prevention Systems (IDPS). They can be divided into two groups:

1. Host-based – for detecting hostile activity on host systems. They analyze information from operation systems and application files captured by software agents.

2. Network based – for detecting hostile activity on networks. They evaluate information from network communication, analyzing the stream of packets captured by sensors.

The models for intrusion detection can be divided into two groups (Gotseva, Trifonov and Stoynov, 2019):

1. Signature verification. In this case, the system detects previously registered intrusion signature.

2. Anomaly detection. In this case, The system looks for abnormal behavior.

Most IDS commercial tools follow the signature verification detection system. However, they have some drawbacks (false alarms, signature of the attack can't be easily discovered, IDS should be periodically updated with new signatures).

Contemporary IDS usually include both signature verification and anomaly detection modules.

For statistically evaluating Intrusion Dtection Systems IDS in network information security, researchers need a documented attack database which database generally represents the ground truth (Rindberg et al., 2008). The evaluation itself counts the number of true positives, true negatives, false positives, and false negatives.

This paper aims at presenting some popular databases of anomalous traffic for estimation of Intrusion Dtection Systems (IDS). The databases KDD'99, MetroSec and ADFA-IDS are presented.

Database KDD'99

Up to now, the most used database is KDD'99 originating from DARPA project on the off-line analysis of intrusion detection systems, at the MIT' Lincoln laboratory MIT (2008). A small network was set-up for this purpose, the aim being to emulate a US Air Force base connected to the Internet. The background traffic was generated by injecting attacks in well defined points of the network, and collected with TCPDUMP. Traffic was grabbed and recorded for 7 weeks, and served for IDS calibration. Once calibration was performed, 2 weeks of traces containing attacks were used to evaluate the performance of IDS under study. This DARPA work in 1998 used a wide range of intrusion attempts, tried to simulate a realistic normal activity, and produced results that could be shared between all researchers interested in such topic. After some changes, mainly dealt with the use of more furtive attacks, the use of target machines running Windows NT, the definition of a security policy for the attacked network, and tests with more recent attacks, this database is used nowadays under the name KDD'99.

McHugh (2001) published a strong criticism on the procedures used when creating the KDD'99 database, especially on the lack of verification of the network realism compared to an actual one. Mahoney et Chan (2003) reviewed into detail the database and showed that the traces were far from simulating realistic conditions, and therefore, that even a very simple IDS can exhibit very high performance results, performances that it could never reach in a real environment.

Despite all the disclaimers about this KDD'99 database for IDS evaluation, it is still massively used. This is mainly due to the lack of other choices and efforts for providing new attack databases.

METROSEC project data

The limitations of the KDD'99 database motivated MetroSec project from the French ACI program on Security & Computer science, 2004-2007 (MetroSec, 2008) to produce controlled and documented traffic traces, with or without anomalies, for testing and validating intrusion detection methods Owezarski (2007).

The experimental platform for creating the MetroSec trace database uses the Renater network, the French network for education and research. Renater is an operational network which is used by a significantly large community in its professional activity. The laboratories involved in the generation of traces are ENS in Lyon, LIP6 in Paris, IUT of Mont-de-Marsan, ESSI in Nice, and LAAS in Toulouse. Traffic is captured at these different locations by workstations equipped with DAG cards Owezarski (2007).

Anomalies studied and generated in the framework of the MetroSec project consist of more or less significant increases of traffic in terms of volume. Two kinds of anomalies are considered: anomalies due to legitimate traffic (like flash crowds- FC) and anomalies due to illegitimate traffic, as flooding attacks, generated by several classical attacking tools.

The anomalies due to illegitimate traffic were DDoS attacks, launched by using flooding attacking tools (IPERF, HPING2, TRIN00 et TFN2K), with changes frequently the attack characteristics and parameters (duration, DoS flow intensity, size and rate of packets) in order to create different profiles for attacks to be detected afterwards Owezarski (2007).

Database ADFA-IDS

Another recent dataset is the Intrusion Dataset (IDS) of Australian Defence Force Academy (ADFA) - ADFA-IDS (Creech and Hu, 2017). It is provided by the University of Arizona Artificial Intelligence Lab and intended as representative of modern attack structure and methodology to replace the older datasets KDD and UNM (Creech and Hu, 2013; Tran et al., 2012).

The version of ADFA-IDS used is released on March 27th, 2017 (Haider et al., 2017) as an update to the original ADFA-IDS made publicly available in 2013 (Creech and Hu, 2013; Creech and Hu, 2014; Haider et al., 2015). ADFA IDS includes independent datasets for Linux and Windows environments.

ADFA-LD (Linux dataset) was generated on a Ubuntu Linux 11.04 host OS with Apache 2.2.17 running PHP 5.3.5. FTP, SSH, MySQL 14.14, and TikiWiki were started (Xie and Hu, 2013; Xie et al., 2014).

Table 1 below shows the payloads and vectors used to attack the Ubuntu OS and generate the dataset.

Table 1. The payloads and vectors used to attack the Ubuntu OS and generate the dataset

PAYLOAD/EFFECT	VECTOR
password bruteforce	FTP by Hydra
password bruteforce	SSH by Hydra
add new superuser	Client side poisoned executable
Java based meterpreter	Tiki Wiki vulnerability exploit
Linux meterpreter payload	Client side poisoned executable
C100 Webshell	PHP remote file inclusion vulnerability

ADFA-WD (Windows dataset) was genearted on a Windows XP Service Pack 2 host OS with the XP default firewall enabled for all attacks, and file sharing enabled, a network printer configured, wireless and Ethernet networking (Haider et al., 2016a; Haider et al., 2016b). Norton AV 2013 was used to scan certain payloads. FTP server, web server and management tool, and streaming audio digital radio package were activated.

A target ratio of 1:10:1=normal:validation:attack data was used to guide collection and structuring activities.

The vectors used here are: TCP ports, web based vectors, browser attacks, and malware attachments

The effects are: Bind shell, reverse shell, exploitation payload, remote operation, staging, system manipulation, privilege escalation, data exfiltration, and back-door insertion.

References

Creech, G. and J. Hu (2013) Generation of a new IDS test dataset: Time to retire the KDD collection, 2013 IEEE Wireless Communications and Networking Conference, WCNC 2013, pp.4487-4492.

Creech, G. and J. Hu (2014) A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns, IEEE Transactions on Computers, vol. 63, issue 4, April 2014, pp. 807-819.

Creech, G. and J. Hu (2017) ADFA IDS Dataset, University of Arizona Artificial Intelligence Lab, AZSecure-data, Director Hsinchun Chen.

Gotseva, D., R. Trifonov, P. Stoynov (2019) Neural Networks for Intrusion Detection – Proceedings of the International conference HiTtech-2019, 10-12 October, Sofia, Bulgaria.

Haider, W., J. Hu, and M. Xie (2015) Towards reliable data feature retrieval and decision engine in hostbased anomaly detection systems, Proc. Of the 2015 10th IEEE Conference on Industrial Electronics and Applications, ICIEA 2015, pp. 513-517.

Haider, W., J. Hu, X. Yu, and Y. Xie (2016a) Integer data zero-watermark assisted system calls abstraction and normalization for host based anomaly detection systems, Proc. Of the 2nd IEEE International Conference on Cyber Security and Cloud Computing, 2016, pp. 349-355.

Haider, W., G. Creech, Y. Xie, and J. Hu (2016b) Windows Based Data Sets for Evaluation of Robustness of Host Based Intrusion Detection Systems (IDS) to Zero-Day and Stealth Attacks. Future Internet 8.3 (2016): 29.

Haider, W., J. Hu, S. Slay, B. P. Turnbull and Y. Xie (2017) Generating Realistic Intrusion Detection System Dataset based on Fuzzy Qualitative Modeling," Journal of Network and Computer Applications (JNCA), 2017, DOI: 10.1016/j.jnca.2017.03.018.

Mahoney, M., P. Chan (2003) An analysis of the 1999 darpa/lincoln laboratory evaluation data for network anomaly detection. In Recent Advances in Intrusion Detection (RAID 20003), pages 220–237, September 2003.

McHugh, J. (2001) Testing intrusion detection systems: A critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory. ACM Transactions on Information and System Security, 3(4):262–294, 2001.

Metrosec (2008) http://www.laas.fr/METROSEC.

MIT (2008) Lincoln Laboratory.

http://www.ll.mit.edu/mission/communications /ist/corpora/ideval, 2008.

Rindberg, H., M. Roughan and J. Rexford (2008) The need for simulation in evaluating anomaly detectors. Computer Communication Review, 38(1):55–59, January 2008.

Owezarski, P. (2007) Contribution of the French METROSEC project to traffic anomalies detection, Colloque STIC, Paris, France, 5-7 November 2007.

Tran, Q., F. Jiang, and J. Hu (2012) A real-time NetFlow-based intrusion detection system with improved BBNN and high-frequency field programmable gate arrays, Proc. Of the 11th IEEE International Conference on trust, Security and Privacy in Computing and Communications, TurstCom 2012, 2012, pp. 201-208.

Trifonov, R., G. Tsochev, R. Yoshinov, S. Manolov, G. Pavlova (2017) Conceptual model for cyber intelligence network security system. International Journal of Computers. Vol. 11, 2017.

Xie, M. and J. Hu (2013) Evaluating host-based anomaly detection systems: A preliminary analysis of ADFA-LD, Proc. Of the 6th International Congress on Image and Signal Processing, CISP 2013, pp. 1711-1716.

Xie, M., J. Hu, X. Yu, and E. Chang (2014) Evaluating host-based anomaly detection systems: Applications of the frequency-based algorithms to ADFA-LD," LNCS, vol. 8792, 2014, pp.542-549.

Introductory Survey for Recent Restrictive Approximations and Applications of Solutions of Initial Boundary Value Probles for PDE

Yousef Methkal Abd Algani,
Department of Mathematics, Sakhnin College, Israel.
Department of Mathematics,
The Arab Academic college for education in Israel
yosefabdalgani@gmail.com
https://orcid.org/0000-0003-2801-5880
Wafic Hibi
Department of Mathematics, Sakhnin College, Israel

Abstract.
This research dealt with the restrictive Pade` approximations and restrictive Taylor's approximations, this paper contains the definitions of restrictive approximations and computing applications to many problems of initial and boundary value problems of systems of partial differential equations of parabolic and hyperbolic types, linear and nonlinear problems also that of variable coefficients and singularly perturbed problems of one dimension or more. It used restrictive Pade` approximations and Taylor's approximations which applied on physical phenomena such as convection-diffusion equations. The suggested method is more efficient and more accurate than many known methods, also it is of low cost. The local truncation error in our method is less than that for the classical known methods, in our method it varies from I0-8 to 10-15 while in the classical known methods it varies from 10-2 to10-9. The efficiency of our method is better than that for Douglass and Rachford (D.R.) method because of the ratio of the number of systems to be solved between D.R. method and this method varies from 2: 1 to 10: 1, also their method is more fast than D.R. method because the ratio of the executed time of their suggested method to D.R. method varies from 1: 2 to 1: 15 the suggested method is considered as an acceleration technique for the classical finite difference methods (appending in Elbarbary, 2001).

Key words: Introductory Survey, Recent Restrictive Approximations, Applications, Solutions of Initial Boundary Value.

Introduction

This research dealt with the restrictive Pade` approximations and restrictive Taylor's approximations, this paper contains the definitions of restrictive approximations and computing applications to many problems of initial and boundary value problems of systems of partial differential equations of parabolic and hyperbolic types, linear and nonlinear problems also that of variable coefficients and singularly perturbed problems of one dimension or more, many of the data and results **depended** in **(**Elbarbary, 2001)

1. Approximations and Restrictive Approximations

1.1 Pade` and Taylor Approximations of Scalar Functions

Padé approximations known as PAI, PAII and PAIII has been developed in theoretical physics and mechanics to solve a great number of problems

(Baker G and Morris P, 1981). Several types of the approximations occurred, the famous one is constructed by segment of Taylor's series of a studied function, the use of interpolation approximations drawn through given n-points and using the least squares method which construct the rational function from a given data.

Let f(x) be has a Maclaurin series

$$f(x)=\sum_{i=0}^{\infty} c_i x^i \tag{1.1}$$

where $c_i = \dfrac{f^{(i)}(0)}{i!}$

The Padé Approximation (PA) of the function f(x) is denoted by $[M, N]_{f(x)}(x)$, where

$$[M,N]_{f(x)}(x) = \frac{\sum_{i=0}^{M} a_i x^i}{1+\sum_{i=1}^{N} b_i x^i} \qquad (1.2)$$

Such that $a_M \neq 0$ and $b_N \neq 0$, M and N are positive integers such that

$$f(x)= [M, N]_{f(x)}(x)+ O(x^{M+N+1}).$$

The Pade` approximants to e^{θ}, θ real (Smith G**, 1985):**

Assume that e^{θ} is approximated by $(1+ p_1\theta)/ (1+q_1\theta)$, where p_1 and q_1 are constants. The determination of p_1 and q_1 requires two equations, which will come from the coefficients of θ, θ^2 and so the leading error term will be of order θ^3. Hence

$$e^{\theta} \equiv \frac{1+ p_1\theta}{1+ q_1\theta} + c_3\theta^3 + c_4\theta^4 + \cdots$$

Therefore,

$$(1+q_1\theta)\left(1+\theta+\frac{1}{2}\theta^2+\frac{1}{6}\theta^3+\cdots\right) \equiv 1+ p_1\theta+(1+q_1\theta)(c_3\theta^3+c_4\theta^4+\cdots).$$

Hence,

$$(1+q_1-p_1)\theta+\left(\frac{1}{2}+q_1\right)\theta^2+\left(\frac{1}{6}+\frac{1}{2}q_1-c_3\right)\theta^3 + \text{ higher order terms} \equiv 0.$$

This is satisfied uniquely to terms of order three by

$$p_1=\frac{1}{2}, \quad q_1=-\frac{1}{2} \quad and \quad c_3=-\frac{1}{12}.$$

The rational approximation $\frac{1+0.5\,\theta}{1-0.5\,\theta}$ is called the (1,1) Pade` approximation of order 2 to $exp\theta$ and has a leading error term of order *3*.

The following table gives the first four Pade` approximants to e^{θ} and their leading error terms:

(S,T)	$R_{S,T}(\theta)$	Principal error term
(0,1)	$1+\theta$	$0.5\theta^2$
(0,2)	$1+\theta+0.5\,\theta^2$	$\frac{1}{6}\theta^3$
(1,0)	$\frac{1}{1-\theta}$	$-\frac{1}{2}\theta^2$
(1,1)	$\frac{1+0.5\,\theta}{1-0.5\,\theta}$	$-\frac{1}{12}\theta^3$

In general, it is possible to approximate $exp\theta$ by

$$e^{\theta} = \frac{1+ p_1\theta+ p_2\theta^2 +\cdots+ p_T\theta^T}{1+ q_1\theta+ q_2\theta^2 +\cdots+ q_S\theta^S} + c_{S+T+1}\theta^{S+T+1} + O(\theta^{S+T+2}),$$

where c_{S+T+1} is a constant. The rational function

$$R_{S,T}(\theta) = \frac{1+ p_1\theta+\cdots+ p_T\theta^T}{1+ q_1\theta+\cdots+ q_S\theta^S} = \frac{P_T(\theta)}{Q_S(\theta)}$$

Where $(b_1, b_2, \ldots, b_T)$, $(q_1, q_2, \ldots, q_S)$ can be uniquely determined and T, S is given,

(S, T) is called the Pade` or [S, T] approximant of order (S+T) to e^{θ}.

The classical implicit approximation:

The (1,0) Pade` approximant approximates $V(t+k) = \{exp(Ka)\}V(t)$

by

$$U(t+k)=(I-kA)^{-1}u(t).$$

Premultiplication of both sides by the matrix $(I-kA)$ yields

$$(I-kA)u(t_j+k)=u(t_j),\ j=0,1,2,\ldots,$$

where $u(t_j+k)=[u_{1,j+1},u_{2,j+1},\ldots,u_{N-1,j+1}]^T$ and matrix A is defined as

$$A=\frac{1}{h^2}\begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}$$

In details, for zero boundary –conditions,

$$\begin{pmatrix} (1+2r) & -r & & & \\ -r & (1+2r) & -r & & \\ & & & & \\ & & & (1+2r) & -r \\ & & & -r & (1+2r) \end{pmatrix} \times \begin{pmatrix} u_{1,j+1} \\ u_{2,j+1} \\ \vdots \\ u_{N-2,j+1} \\ u_{N-1,j+1} \end{pmatrix} = \begin{pmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{N-2,j} \\ u_{N-1,j} \end{pmatrix}$$

The i[th] equation gives the implicit or backward-difference scheme

$$-ru_{i-1,j+1}+(1+2r)u_{i,j+1}-ru_{i+1,j+1}=u_{i,j},\quad i=1(1)N-1.$$

This is unconditionally stable for all $r=k/h^2>0$.The leading error terms are of order h^2 in x because of the central-difference approximation to $\frac{\partial^2 u}{\partial x^2}$ and of order k in t.(The leading error term of the (1,0) Pade` approximant to $exp(kA)$ is $O(k^2)$) The method is said to be first order accurate in t .

1.2 Restrictive Approximations of a Scalar Function:

The Restrictive Padé Approximation (RPA) of the Function f(x)

Restrictive Pade` approximation of a function f(x) is constructed in

(Hassan, Ismail and Elbarbary. 1995) with parameters ε_i's to be determined, which if reduces to zero the classical Padé approximation [2] is obtained. The advantage is that it has zero error at certain points and consequently relatively small in between.

The restrictive Padé approximation (RPA) of the function f(x) is

$$RPA[M,\alpha,N]_{f(x)}(x)=\frac{\sum_{i=0}^{M}a_i x^i+\sum_{i=M+1}^{M+\alpha}\varepsilon_{i-M}x^i}{1+\sum_{i=1}^{N}b_i x^i} \tag{1.3}$$

where $\varepsilon_j=0\ \forall j>\alpha$ and $\alpha=0(1)N$ i.e., the positive integer α does not exceed the degree of the numerator, then

$$f(x)=RPA[M,\alpha,N]_{f(x)}(x)+O(x^{M+N+1}) \tag{1.4}$$

and let f(x) have a Maclaurin series (1.1.1)

Then

$$\left(\sum_{i=0}^{\infty}c_i x^i\right)\left(1+\sum_{i=0}^{N}b_i x^i\right)-\left(\sum_{i=0}^{M}a_i x^i\right)-\left(\sum_{i=M+1}^{M+N}\varepsilon_{i-M}x^i\right)=O\left(x^{M+N+1}\right) \tag{1.5}$$

The vanishing of the first $(M+N+1)$ powers of x on the left-hand side of (1.5) implies the system of $(M+N+1)$ equations:

$$a_r=c_r+\sum_{j=1}^{M}c_{r-j}b_j\ ,\ r=0\ (1)M\ ;\qquad (b_j=0\ if\ j>M)$$

$$c_{M+N-s}+\sum_{j=1}^{N}c_{M+N-j-s}b_j=\varepsilon_{N-s}\ ,\ s=0(1);\qquad (c_j=0\ if\ j<0)$$

S.t. we can determine the coefficient $a_i\ and\ b_i$ as a function of ε, $i=1(1)\alpha$, and let $[M,\alpha,N]_{f(x)}(x_j)$ $=f(x_j)$, $j=1(1)\alpha$ when this system of $(M+N+\alpha+1)$ linear equations in $(M+N+\alpha+1)$ unknowns has a solution, it leads to the desired approximation (1.3). If $\varepsilon_i=0, \forall i=1(1)\alpha$ in (1.5) we get the classical Padé approximation (1.2)

The RPA can be arranged in a doubly infinite array, each entry of which represents $[M, \alpha, N]_{f(x)}(x)$, $\alpha \le N$ for $f(x)$.

The case $\alpha = 0$ gives the classical Padé approximation table.

The case $\alpha = 1$ gives, for examples, the following selected elements of RPA:

$$[0, 1, 1]_{f(x)}(x) = \frac{a_0 + \varepsilon_1 x}{1 + b_1 x}$$

where

$$a_o = c_o \text{ , } b_1 = \frac{\varepsilon_1 - c_2}{c_0},$$

$$[1, 1, 1]_{f(x)}(x) = \frac{a_0 + a_1 x + \varepsilon_1 x^2}{1 + b_1 x}$$

where $a_o = c_o$, $a_1 = c_1 + \frac{c_0(\varepsilon_1 - c_2)}{c_1}$, and $b_1 = \frac{\varepsilon_1 - c_2}{c_1}$,

$$[1, 1, 2]_{f(x)}(x) = \frac{a_0 + a_1 x + \varepsilon_1 x^2}{1 + b_1 x + b_2 x^2}$$

where $a_o = c_o$, $a_1 = c_1 + \frac{c_0(c_1c_2 - c_0c_3 - c_1\varepsilon_1)}{c_0c_2 - c_1^2}$, $b_1 = \frac{c_1c_2 - c_0c_3 - c_1\varepsilon_1}{c_0c_2 - c_1^2}$ and $b_2 = \frac{c_1c_3 - c_2^2 - c_2\varepsilon_1}{c_0c_2 - c_1^2}$.

The case $\alpha = 2$, gives, for examples, the following selected elements of RPA:

$$[0, 2, 2]_{f(x)}(x) = \frac{a_0 + \varepsilon_1 x + \varepsilon_2 x^2}{1 + b_1 x + b_2 x^2}$$

where $a_o = c_o$, $b_1 = \frac{\varepsilon_1 - c_1}{c_0}$, and $b_2 = \frac{c_1^2 - c_0c_2 - c_1\varepsilon_1 + c_0\varepsilon_2}{c_0^2}$, and

$$[1, 2, 2]_{f(x)}(x) = \frac{a_0 + a_1 x + \varepsilon_1 x^2 + \varepsilon_2 x^3}{1 + b_1 x + b_2 x^2}$$

where

$a_o = c_o$, $a_1 = \frac{2c_0c_1c_2 - c_0^2c_3 - c_0c_1\varepsilon_1 + c_0^2\varepsilon_2 - c_1^3)}{c_0c_2 - c_1^2}$, $b_1 = \frac{c_1c_2 - c_0c_3 - c_1\varepsilon_1 + c_0\varepsilon_2}{c_0c_2 - c_1^2}$ and $b_2 = \frac{c_1c_3 - c_2^2 - c_2\varepsilon_1 - c_1\varepsilon_2}{c_0c_2 - c_1^2}$ and so on.

For **the Local Truncation Error (LTE) form of RPA** of the function f(x) of an $(n+1)^{th}$ derivative, it can be shown that for every argument $\bar{x}$ there exists a number η in the smallest interval I containing the set of points $\{x_o, x_1, x_2, x_3, \ldots, x_\alpha, \bar{x}\}$, such that

$$R_{M,\alpha,N}(\bar{x}) = f(\bar{x}) - [M,\alpha,N]_{f(x)}(\bar{x}) = \frac{\pi_{\alpha+1}(\bar{x})}{(\alpha+1)!}\left(R_{M,\alpha,N}(\eta)\right)^{(\alpha+1)},$$

where $\pi_{\alpha+1}(x) = x(x-x_1)(x-x_2)(x-x_3)\ldots(x-x_\alpha)$ and $R_{M,\alpha,N}$ is the local truncation error for RPA.

A Restrictive Taylor `s Approximation of the Function f(x) (Hassan, Ismail and Elbarbary.2001):

A new approach for an explicit method to solve the parabolic partial differential equation is developed. This approach will exhibit several advantages features: Highly accurate, fast, and the absolute error still very small whatever the exact solution is too large. The computational results, which are presented, show that computing time and memory space are saved with more accuracy. Stability conditions are obtained and utilized in numerical computations.

One of the strategies for approximation of the function $f(x)$ is by using Taylor's expansion, which approximate $f(x)$ by a polynomial of degree n and its truncation error is of order $n+1$.

$$f(x) = P_{n,f(x)}(x) + R_{n+1}(x).$$

In a more extensive form $f(x)$ can be written out as

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x) \qquad (1.6)$$

where

$$R_{n+1} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \tag{1.7}$$

and ξ lies between a and *x*.

In the paper (Hassan, Ismail and Elbarbary.2001) the authors constructed a restrictive type of Taylor's expansion of the function *f(x)* at the point a, with parameter to be determined such that the error equal to zero at certain point x_o. If the parameter reduces to one, we get the classical Taylor's expansion at the point a. The advantage is that it has exact value at the points *a* and x_o and relatively small in between.

Consider a function *f(x)* defined in a neighborhood of the point *a*, and it has derivatives up to order *(n+1)* in this neighborhood.

Then

$$RT_{n,f(x)}(x,a) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{\varepsilon\, f^{(n)}(a)}{n!}(x-a)^n \tag{1.8}$$

Now, $RT_{n,f(x)}(x_o) = f(x_o)$.It means that the considered approximation is exact at two points *a* and x_o.

Putting

$$f(x) = RT_{n,f(x)}(x) + R_{n+1}(x), \tag{1.9}$$

where $R_{n+1}(x)$ is the remainder term of the restrictive Taylor's series.

In the following Theorem the remainder term $R_{n+1}(x)$ can be expressed in terms of ε, n^{th} and $(n+1)^{th}$ derivatives of the function *f(x)* at a point ξ lies between *a* and *x*.

Consider the function *f(x); f(x)* $\in C^{n+1}$, *x* $\in I$, *I* is the neighborhood of a point *a*. The error for approximation estimated by $R_{n+1}(x)$ is given by the formula (1.9), for which

$$R_{n+1}(x) = \frac{\varepsilon(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi) - \frac{n(\varepsilon-1)(x-a)^{n+1}}{(x-\xi)(n+1)!} f^{(n)}(\xi)$$

where $\xi \in [a,x]$ and ε is a restrictive parameter (Hassan, Ismail and Elbarbary.2001)

Note: $\varepsilon = 1$ gives the classical Taylor's expansion (1.1.6) and remainder term (1.7) and

$$\varepsilon = \frac{-nf^{(n)}(\xi)}{(x_o-\xi)f^{(n+1)}(\xi) - nf^{(n)}(\xi)}$$

gives the exact value at the point x_o, in which $R_{n+1}(x_o) = 0$.

1.3 Restrictive Padé and Restrictive Taylor's Approximation of Matrix Functions:

Padé Approximation of Matrix Functions:

It can be proved the Pade` approximation of the exponential matrix in many references e.g. (Smith G, 1985) when we study methods of finite differences as in that:

$$PA_{e^{Ar}}[1,1] = \left(I - \frac{1}{2}A\,r\right)^{-1}\left(I + \frac{1}{2}A\,r\right). \tag{1.10}$$

Restrictive Pade` Approximation of the Exponential Matrix:

The restrictive Pade` approximation is used to approximate the exponential matrix *exp(rA)*. The advantage is that it has the exact value at certain r

The exponential matrix *exp(rA)* can be formally defined by the convergent power series:

$$Exp\,(rA) = I + rA + \frac{r^2}{2!}A^2 + \cdots = \sum_{n=0}^{\infty}\frac{r^2}{n!}A^n,\ A^o = I$$

where *A* is an *(N-1) × (N-1)* matrix.

In the case of restrictive Pade` approximation of single function the term ε_i in equation (1.3) can be reduced to the square restrictive matrix $\in_{(N-1)\times(N-1)}$.

In the case of restrictive Pade` approximation for matrix function, where

$$\in = \begin{pmatrix} \varepsilon_1 & \varepsilon_1 & & & & 0 \\ \varepsilon_2 & \varepsilon_2 & \varepsilon_2 & & & \\ & \varepsilon_3 & \varepsilon_3 & \varepsilon_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \varepsilon_{N-2} & \varepsilon_{N-2} & \varepsilon_{N-2} \\ 0 & & & & \varepsilon_{N-1} & \varepsilon_{N-1} \end{pmatrix}$$

Without loss of generality, the authors chose $\varepsilon_{i,j}$ to be constant in each row i.e., $\varepsilon_{i,i-1} = \varepsilon_{i,i} = \varepsilon_{i,i+1}$ and $\varepsilon_{i,j} = 0$ otherwise, for example

$$RPA[0,1,1]_{exp(rA)}(r) = \left(I + (\in -\frac{1}{2}A)r\right)^{-1}\left(I + (\in +\frac{1}{2}A)r\right). \qquad (1.11)$$

Restrictive Taylor's Approximation of the Exponential Matrix (Hassan, Ismail and Elbarbary.2001):

Consider the exponential matrix *exp(xA)* in the form (1.10).

In the case of restrictive Taylor's approximation of single function the term ε in (1.1.8) can be reduced to the square restrictive matrix Υ in the case of restrictive Taylor's approximation for matrix function, where

$$\Upsilon = \begin{pmatrix} \varepsilon_1 & & & & 0 \\ & \varepsilon_2 & & & \\ & & \varepsilon_3 & & \\ & & \ddots & \ddots & \ddots \\ & & & \varepsilon_{N-2} & \\ 0 & & & & \varepsilon_{N-1} \end{pmatrix}_{(N-1)\times(N-1)}$$

for example $RT_{\exp(Ax)}(0,1)(x) = I + x\Upsilon A$.

1.4 Uniqueness and Solvability of Restrictive Approximations

Ismail (2004) introduced a general theory for solvability and uniqueness of the restrictive Pade` approximation (RPA) and restrictive Taylor's approximation (RTA). A survey of the individual necessary and sufficient solvability and uniqueness conditions for fifteen examples of these approximations are given.

Solvability and Uniqueness Condition for RPA:

The author introduced the solvability and uniqueness condition for restrictive Pade` approximation (1.3) - (1.5) for the sufficiently smooth function *f(x)* is uniquely solvable for $\prod_{i=1}^{\alpha} x_i \neq 0$ if the following *N×N* restrictive determinant $/D/ \equiv \Delta \neq 0$:

$$\Delta = \begin{vmatrix} c_{M+\alpha} & c_{M+\alpha-1} & \cdots & c_{M+\alpha-N+1} \\ c_{M+\alpha+1} & c_{M+\alpha} & \cdots & c_{M+\alpha-N+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M+N-1} & c_{M+N-2} & \cdots & c_M \\ R_{M+\alpha-1}(x_1) & x_1 R_{M+\alpha-2}(x_1) & \cdots & x_1^{N-1} R_{M+\alpha-N}(x_1) \\ R_{M+\alpha-1}(x_2) & x_2 R_{M+\alpha-2}(x_2) & \cdots & x_2^{N-1} R_{M+\alpha-N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{M+\alpha-1}(x_\alpha) & x_\alpha R_{M+\alpha-2}(x_\alpha) & \cdots & x_\alpha^{N-1} R_{M+\alpha-N}(x_\alpha) \end{vmatrix}$$

where:

1. $R_n(x) = P_n(x) - f(x)$, $P_n(x) = \sum_{i=0}^{n} c_i x^i$, $c_i = \frac{1}{i!} f^{(i)}(0),$.

2. The c_n's in the first $(N-\alpha)$-rows and $R_n(x_i)$'s in the last α-rows such that n decreases toward the right. While the right remaining elements in each row must be zeros. When the last nonzero terms are c_o or $R_o(x)$ respectively. *i.e* $k < 0 \Rightarrow c_k \equiv 0$, $P_k(x_j) \equiv -f(x_j)$, $j = 1(1)\alpha$,

3. In the case of the absence of the mentioned α-restrictions, it is clear the reduction of this condition to the well-known C-table *C(M,N)* in G.A.Bakerwhich discuss the solvability condition for Pade` approximations. Similarly it can be defined "The normal Restrictive Pade` Box" of the tables $C(M,\alpha, N) \equiv \Delta$ if none of these determinants vanish.

Solvability and Uniqueness Condition for RTA:

The restrictive Taylor's approximation is a polynomial satisfying the conditions:

$$RT_{f(x)}(M,\alpha)(x) = \sum_{i=o}^{M} a_i x^i + \sum_{i=1}^{\alpha} \varepsilon_i \, x^{M+i} \ , x \in R \qquad (1.12)$$

Such that :

$$RT_{f(x)}(M,\alpha)(x) - f(x) = O(x^{M+1}), \quad (1.13)$$

$$RT_{f(x)}(M,\alpha)(x_i) = f(x_i) \ \forall x_i \neq 0,\ i=1(1)\alpha. \quad (1.14)$$

where

1. M is a nonnegative integer
2. α is a positive integer
3. $f(x)=\sum_{i=o}^{\infty} c_i (x-x_o)^i$ for simplicity $x_o=0 \Rightarrow c_i = \frac{1}{i!} f^{(i)}(0)$ i.e. $f(x)$ is sufficiently smooth function at $x = 0$.

Uniqueness and solvability of restrictive Taylor's approximation:

Restrictive Taylor's approximation (1.12)-(1.14) is uniquely solvable if:

$$\forall i, j = 1(1)\alpha;\ x_i \neq 0\ ,\ x_i \neq x_j \ \forall i \neq j.$$

(Note that putting $N = 0$ for restrictive Pade` approximation (1.3) - (1.5) will reduce it to restrictive Taylor approximation.)

2. Applications of Approximations for IBVP for PDE

We begin with presenting some details and introduce through the advanced proofs that the used numerical solutions are convergent, from analytical point of view, to the exact solutions.

2.1 Convergence to the exact solution:

In Butcher (2008) proved that LTE $T_{ij/RP}$ for the restrictive Pade` approximation for parabolic type takes the form:

$$\begin{aligned}
&\left[3(u_t)+O(k,h^2)\right]T_{ij/RP}(h,k) = \\
&-\left\{\frac{1}{8}\left[\left(u_{t^2}\right)_{ij}\left(u_t\right)_{i0}-\left(u_{t^2}\right)_{i0}\left(u_t\right)_{ij}\right]h^2\right. \\
&+\frac{1}{4}\left[\left(u_{t^3}\right)_{ij}\left(u_t\right)_{i0}-\left(u_{t^3}\right)_{i0}\left(u_t\right)_{ij}\right]k^2 \\
&+\frac{1}{8}\left[\left(u_{t^4}\right)_{ij}\left(u_t\right)_{i0}-\left(u_{t^4}\right)_{i0}\left(u_t\right)_{ij}+\left(u_{t^3}\right)_{ij}\left(u_{t^2}\right)_{i0}-\left(u_{t^3}\right)_{i0}\left(u_{t^2}\right)_{ij}\right]k^3 \\
&\left.+\frac{1}{8}\left[\left(u_{t^3}\right)_{ij}\left(u_t\right)_{i0}-\left(u_{t^3}\right)_{i0}\left(u_t\right)_{ij}\right]k\,h^2+\cdots\right\}.
\end{aligned}$$

The terms between each square brackets [ ] will take the form:

$$A_{p,m}=\left[\left(u_{t^p}\right)_{ij}\left(u_{t^m}\right)_{i0}-\left(u_{t^p}\right)_{i0}\left(u_{t^m}\right)_{ij}\right],\ p,\ m = 0(1)\infty$$

Using the closed form exact solution: $u(x,t)=e^{-t}\sin x$ where $x=ih,\ t=jk$

$A_{p,m}=(-1)^{p+m}\left[\left(e^{-t}\sin x\right)_{ij}\left(e^{-t}\sin x\right)_{i0}-\left(e^{-t}\sin x\right)_{i0}\left(e^{-t}\sin x\right)_{ij}\right]=0$ i.e it can be shown that: $T_{ij/RP}(h,k) = 0$.

It is proved also that LTE $T_{ij/RP}$ for the restrictive Pade` approximation for hyperbolic type takes the form

$$\begin{aligned}
\left[\left(u_t\right)_{i0}+O(k)\right]T_{ij/RP}(h,k) = &\left\{\frac{a}{6}\left[\left(u_t\right)_{i0}\left(u_{x^3}\right)_{ij}-\left(u_t\right)_{ij}\left(u_{x^3}\right)_{i0}\right]h^2\right. \\
&+\frac{a}{24}\left[\left(u_t\right)_{i0}\left(u_{x^4}\right)_{ij}-\left(u_t\right)_{ij}\left(u_{x^4}\right)_{i0}\right]h^3 \\
&\left.+\frac{a}{48}\left[\left(u_{t^2}\right)_{i0}\left(u_{x^3}\right)_{ij}-\left(u_{t^2}\right)_{ij}\left(u_{x^3}\right)_{i0}\right]h^2k+\cdots\right\}
\end{aligned}$$

The terms between each square brackets [ ] will take the form:

$$A_{p,m}=\left[\left(u_{t^p}\right)_{i0}\left(u_{x^m}\right)_{ij}-\left(u_{t^p}\right)_{ij}\left(u_{x^m}\right)_{i0}\right]$$

The advantages of the restrictive Pade` approximation can be summarized as follows:

i) The method gives the exact solution if it is known at one level of time, for example at $t= k$, *i.e.* $u(x,t)=u(ih,k)\ \ i=1(1)N$.

ii) Without knowing the exact solution at one level, they tried to use an approximate, fast efficient and accurate method with suitable very small step sizes *h* and *k*, to get the needed almost exact solution at specific level, after which the usual restrictive Pade` process is continued.

iii) The needed exact solution at the first level need not be in closed form, i.e., we need only a table of the exact solution at some points.

In **(**Cohn, D, 2013) the authors obtained similar forms-but simpler- to the above forms. in the study of Taylor's approximations

In (Thomas, George B., Jr.; Finney, Ross L. 1996)**,** the authors studied the convergence of the restrictive Pade` approximation to the exact solution of nonlinear IBVP for generalized Burger's-Huxley and generalized Burger's-Fisher equations.

The corresponding LTE for Restrictive Pade method takes the form:

$$T_{i,j/RP} = \frac{1}{k}\left[-\frac{1}{2}r\left(1+\frac{1}{2}h\alpha M^{\delta}\right)u_{i-1}+\left(1+r+r\varepsilon_{i,l}+\frac{1}{2}rh^2\lambda\right)u_i-\frac{1}{2}r\left(1-\frac{1}{2}h\alpha M^{\delta}\right)u_{i+1}\right]_{j+1}$$

$$-\frac{1}{k}\left[-\frac{1}{2}r\left(1+\frac{1}{2}h\alpha M^{\delta}\right)u_{i-1}+\left(1-r+r\varepsilon_{i,l}+\frac{1}{2}rh^2\lambda\right)u_i-\frac{1}{2}r\left(1-\frac{1}{2}h\alpha u_{i,j}^{\delta}\right)u_{i+1}\right]_{j}$$

$$=\frac{1}{k}r\varepsilon_{i,l}\left(u_{i,j+1}-u_{i,j}\right)+T_{i,j/cn}$$

2.2. Restrictive Padé Approximation for Parabolic Types:

In this section we present the papers [20], [6], [5], and [11].

2.2.1 One Dimensional Case:

Restrictive Pade` Approximation for Parabolic PDEs

The derivation of the method is based on the simple heat–conduction standard problem:

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2} \quad 0\le x\le x_0,\ t\ge 0 \tag{2.1}$$

with the initial and boundary conditions, respectively,

$$u(x,0)=g(x)\ 0\le x\le x_0\ ,\ u(0,t)=u(x_0,t)=0,\ t\ge 0. \tag{2.2}$$

The open rectangular domain is covered by a rectangular grid with spacing *h* and *k* in the *x*, *t* directions respectively, the grid point *(x, t)* denoted by *(ih, jk)* and $u(ih, jk) = u_{i,i}$ where *i = 0(1)N*, *j* is a non –negative integer.

The exact solution of grid representation of (2.1) is given by

$$u_{i,j+1}=exp(kD^2_x)\,u_{i,j}\ ,$$

the approximation of the partial derivative D^2_x at the grid point *(ih, jk)* will take the usual form:

$$D^2_x u=\frac{1}{h^2}(u_{i-1,j}-2u_{i,j}+u_{i+1,j}) \tag{2.3}$$

Then

$$U^{j+1}=exp(rA)\,U^j\ ,\ r=\frac{k}{h^2} \tag{.4}$$

where

$$U^j=(u_{1,j}\ ,u_{2,j}\ ,\ldots u_{N-1,j})^T,\quad Nh=x_0$$

and

$$A=\begin{pmatrix} -2 & 1 & & & & 0\\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1\\ 0 & & & & 1 & -2 \end{pmatrix}_{(N-1)\times(N-1)} \tag{2.5}$$

using the equation of the *RPA [0, 1, 1]* $_{exp(rA)}$ to approximate the exponential matrix in (2.4), then the vector form

$$U^{j+1} = \left(I + \left(\varepsilon - \frac{1}{2}A\right)r\right)^{-1}\left(I + \left(\varepsilon + \frac{1}{2}A\right)r\right)U^{j} = BU^{j} \tag{2.6}$$

will have the stability condition: $-1/3 \le r\varepsilon_i \le 1,\ i = 1(1)N-1$.

The desired method is more efficient and more accurate than many known methods; also it is of low cost. The local truncation error in this method is less than that for classical known methods, in this method it varies from 10^{-10} to 10^{-14} *while* in the classical known methods it varies from 10^{-3} to10^{-7} as is done in the enclosed tables in (Ismail and Elbarbary, 2007). The efficiency of RP method is better than that for Douglas method because of the ratio of the number of multiplications and divisions between Douglas method and the constructed method is varies from 10:1 to 61:1, also this method is more fast than Douglas method because the ratio of the executive time of our desired method to Douglas method varies from 1:8 to 1:33.

2.2.2 Chebyshev Approach

Restrictive Chebyshev Rational Approximation and Applications to Heat-Conduction Problems

According to (Silviu-Ioan Filip, 2016), arestrictive type of Chebyshev rational approximation is constructed. It yields more accurate results and exact values at some selected points. Authors used this method to approximate the exponential function. This approach is applied to define a new implicit finite difference scheme to parabolic PDEs. This approach will exhibit several advantages features: highly accurate, fast, and the absolute error still very small whatever the exact solution is too large. Stability conditions are obtained and utilized in numerical computations. Finally, some numerical results to describe the performance of this approach are given.

In an approximation of a function, the property of the error that is of most importance is the maximum error (in magnitude) on the interval. Therefore, the major aim of the approximation process of a function is to make the maximum as small as possible. The rational approximations to a function lead to smaller maximum error than polynomial approximations. Maehly's method is an algorithm which enables one to convert Chebyshev series into a rational expressions involving Chebyshev polynomials. The Chebyshev expansion for the function f(x) is given by

$$f(x) = \frac{1}{2}c_0 + \sum_{j=1}^{\infty} c_i T_j(x),$$

where $c_j = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_j(x)}{\sqrt{1-x^2}}dx,\ j=0,1,\ldots$ and $T_j(x) = \cos(j\cos^{-1}x)$ is the j^{th} degree Chebyshev polynomial.

The Chebyshev rational approximation for the function f(x) will be

$$T_{m,n}(x) = \frac{\sum_{j=0}^{m} a_j T_j(x)}{\sum_{j=0}^{n} b_j T_j(x)},$$

where the a_j*'s* and b_j*'s* are to be determined, so that

$$a_0 = \frac{1}{2}\sum_{i=0}^{n} b_i c_i \text{ and } a_r = \frac{1}{2}\sum b_i (c_{|r-i|} + c_{r+i}),\ r = 1, 2, \ldots, (m+n),\ a_r = 0, \qquad \forall r > m.$$

In this paper authors developed a technique for generating a rational approximation to a function which give the exact values at certain selected points. Furthermore, they used the new approach to solve heat-conduction problems.

Restrictive Chebyshev rational approximation:

The restrictive Chebyshev rational approximation, of the function f(x) due to combining between rational approximation and interpolation. It can be written in the form:

$$T_{m,n}^{p}(x) = \frac{\sum_{i=0}^{m} a_i T_i(x) + \sum_{i=m+1}^{m+p} \varepsilon_{i-m} T_i(x)}{1 + \sum_{i=1}^{n} b_i T_i(x)},$$

where the a_i*'s* and b_i*'s* are to be determined so that

$$f(x)\text{-}T_{m,n}^{p}(x)=\frac{\sum_{i=0}^{\infty}c_iT_i(x)\left(1+\sum_{j=1}^{n}b_iT_j(x)\right)-\sum_{i=0}^{m}a_iT_i(x)-\sum_{i=m+1}^{m+p}\varepsilon_{i-m}T_i(x)}{1+\sum_{i=1}^{n}b_iT_i(x)}.$$

The coefficients of $T_i(x)$, $i=0,1,2,\ldots,(m+n)$ in the numerator of the right-hand side vanish. Hence

$$\sum_{i=0}^{\infty}c_iT_i(x)\left(1+\sum_{j=1}^{n}b_jT_j(x)\right)-\sum_{i=0}^{m}a_iT_i(x)-\sum_{i=m+1}^{m+p}\varepsilon_{i-m}T_i(x)=\sum_{i=m+n+1}^{\infty}h_iT_i(x). \qquad (2.7)$$

In order to get equations for the a_i's, b_i's and ε_i's use the identity,

$$T_i(x)T_j(x)=\frac{1}{2}\left(T_{i+j}(x)+T_{|i-j|}(x)\right)$$

and with this, Eq. (1.2.7) can be rewritten as:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{n}c_ib_j\left(T_{i+j}(x)+T_{|i-j|}(x)\right)-\left(\sum_{i=0}^{m}a_iT_i(x)+\sum_{i=m+1}^{m+p}\varepsilon_{i-m}T_i(x)\right)=\sum_{i=m+n+1}^{\infty}h_iT_i(x), \qquad (2.8)$$

where $b_0=1$. Then it is obtained

$$a_0=\frac{1}{2}\sum_{i=0}^{n}b_ic_i,\ a_r=\frac{1}{2}\sum b_i(c_{r+i}+c_{|r-i|}), \qquad r=1,2,\ldots,m,$$

$$\varepsilon_{r-m}=\frac{1}{2}\sum b_i(c_{r+i}+c_{|r-i|}), \qquad r=m+1,m+2,\ldots,m+n,\ \ \varepsilon_k=0,\ \ \forall k>l,$$

which are $(n+m+1)$ equations in $(m+p+n+1)$ unknowns, then putting the following interpolating equations

$$T_{m,n}^{p}(x_j)=f(x_j), \qquad j=1,2,\ldots,l,$$

which gives a system of l equations. It means that the suggested approximation is exact at l points. Then the $(m+n+p+1)$ coefficients, a_i, $i=0,1,\ldots,m$, b_i, $i=1,2,\ldots,m$ and $\varepsilon_{l,i}=1,2,\ldots,l$ can be obtained from these systems. The points x_j, $j=1,2,\ldots,l$ are chosen such that the error $\left|f(x_j)-T_{m,n}(x_j)\right|$ is maximum.

Chebyshev rational approximation of the exponential matrix:

The approximation of the exponential matrix $exp(rH)$ in terms of Cheby-shev polynomials can be defined as:

$$\exp(rH)=\frac{1}{2}c_0I+\sum_{j=0}^{\infty}c_jT_j(rH),$$

Where $T_0(rH)=I$, $T_1(rH)=rH$, $T_{n+1}(rH)=2rHT_n(rH)-T_{n-1}(rH)$ and we can get c_j by the orthogonality property of Chebyshev approximation.

As an example, the Chebyshev rational approximation $T_{1,1}$ up to 15-digits is given by:

$$T_{1,1}(e^r)=\frac{a_0+a_1r}{1+b_1r}=\frac{1.004818014180699+0.4823222876616465r}{1-0.4622554267229283r} \text{ where } b_1=-\frac{2c_2}{c_1+c_3},$$

$$a_0=\frac{1}{2}(c_0+c_1b_1),\ a_1=c_1+\frac{1}{2}b_1(c_0+c_2),$$

and the restrictive Chebyshev rational approximation $T_{1,1}^{1}$ has the form

$$T_{1,1}^{1}(e^{rA})=(I+b_1rA)^{-1}(a_0I+a_1rA), \qquad (2.9)$$

where

$$b_1=-\frac{2(c_2-\varepsilon_1)}{c_1+c_3},\ a_0=\frac{1}{2}(c_0+c_1b_1),\ a_1=c_1+\frac{1}{2}b_1(c_0+c_2),$$

Restrictive Chebyshev rational approximation for parabolic PDE^s:

The derivation of the desired method is based on the simple heat-conduction standard problem (2.1): with initial and boundary conditions similar to (2.2).

The exact solution of a grid representation of Eq. (2.1) is given by

$$u_{i,j+1} = \exp\left(k \frac{\partial^2}{\partial x^2} \right) u_{i,j} \tag{2.10}$$

where

$$\frac{\partial^2}{\partial x^2} = \frac{1}{h^2}\left(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 + \cdots \right), \qquad r = \frac{k}{h^2}. \tag{2.11}$$

If the difference formula, correct to second difference, is substituted for $\frac{\partial^2}{\partial x^2}$, then Eq. (2.11) can be rewritten as: $U_{i,j+1} = \exp(r\delta_x^2)U_{i,j}$

Then the following finite difference equation is obtained:

$b_1 r U_{i-1,j+1} + (1-2b_1 r)U_{i,j+1} + b_1 r U_{i+1,j+1} = a_1 r U_{i-1,j} + (a_0 - 2a_1 r)U_{i,j} + a_1 r U_{i+1,j}$ (2.12)

The stability condition takes the form

$$\frac{|a_0 - 2a_1 r| + |2a_1 r|}{|1 - 2b_1 r| + |2b_1 r|} \le 1.$$

2.2.3 Two Dimensional Case

According to, D. Ambrosi, D, Arioli, G. and Koch (2012), a new accurate fast implicit method for the finite difference solution of the two dimensional parabolic PDE with first level condition, which may be obtained by any other method. The stability region is discussed. The suggested method is considered as an accelerating technique for the implicit finite difference scheme, which is used to find the first level condition. The obtained results are compared with some famous finite difference schemes and it is satisfactory agreement with the exact solution.

Equations of motion in fluid mechanics are frequently reduced to parabolic formulations. In addition, the unsteady heat conduction equation is also parabolic. Also many of the applications parabolic equations arise as macroscopic descriptions or processes whose microscopic description is essentially probabilistic. Heat flow is related to the random motion of electrons and atoms. Parabolic equations are also used in economic models to give a macroeconomics description of market behavior.

The authors examined the suggested scheme, that can be used to solve a simple model of parabolic PDE^s, that is the heat equation, this equation is called the model equation because it can be used to model the behavior of more complicated parabolic PDE^s. The heat equation can serve as a model equation for other parabolic PDE^s such as the boundary layer equations. The model equation under consideration has the following form:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{2.13}$$

in the rectangular domain bounded by $0 \le x \le x_0$ and $0 \le y \le y_0$ with specific boundary and initial conditions.

The restrictive Pade` approximation is used to approximate the exponential form $\exp(r\delta_x^2)$ where δ_x^2 is the central difference operator.

For example the restrictive Pade` approximation *RPA[1,1]* or the exponential form $\exp(r\delta_x^2)$ is given by

$$RPA[1,1]_{\exp(\delta_x^2)}(r) = \left(1 + \left(\varepsilon_i - \frac{1}{2}\delta_x^2\right)r\right)^{-1}\left(1 + \left(\varepsilon_i + \frac{1}{2}\delta_x^2\right)r\right).$$

Various finite difference approximations can be used to represent the derivatives in Eq. (2.13), The resulting finite difference methods can be classified into two broad categories, explicit or implicit methods. The explicit methods has the advantages of being simple and easy to program, but the restriction of small time step usually requires a very large number of

time step. If we apply the simple explicit method to the two dimensional heat equation, the stability condition reduces to $r \le 0.25$ which is twice as restrictive as the one dimension constraint $r \le 0.5$ and makes this method even more impractical. To alleviate this difficulty various implicit methods have been devised. The resulting system of linear algebraic equations requires substantially more computer time to solve than do a tri-diagonal system. The difficulties described above, which occur attempting to solve the two dimensional heat equation by explicit or implicit algorithms, led to the development of alternating-direction implicit (ADI) methods.

In this paper authors introduced a new implicit method for solving two dimensional parabolic PDE[s], which will exhibit several advantageous features, such as higher accuracy compared with other known methods, the choice of the time step length is sufficiently large compared with that can be used for the classical schemes, this allows them to have the solution at high time level and the accuracy has not been lost when the value of the exact solution is sufficiently large. i.e., the absolute error is sufficiently small whenever the exact solution is relatively large. Their method is considered as acceleration technique for the classical finite difference methods.

Restrictive Pade Approximation for 2-Dimensional Parabolic PDEs

The derivation of the method is based on the unsteady heat conduction Eqs. (2.13). The region to be examined in *(x,y,t)* space is covered by a rectilinear grid with sides parallel to axes with h and k the grid spacing in the distance and time directions respectively. The grid point *(x,y,t)* are given by *(ph, mh, nk)* where *p, m =0(l)N and n = 0, 1, 2,...* . The function satisfying the difference equation at the grid point is $u_{p,m}^{n}$.

The exact difference replacement of Eq, (1.2.13) is given by

$$u_{p,m}^{n+1} = \left(\exp\left(k \frac{\partial^2}{\partial x^2} \right) \exp\left(k \frac{\partial^2}{\partial y^2} \right) \right) u_{p,m}^{n} \tag{2.14}$$

where $\partial^2 / \partial x^2$ is as in (2.11) and so is $\partial^2 / \partial y^2$.

If the difference formulae, correct to second difference are substituted for $\partial^2 / \partial x^2$ and $\partial^2 / \partial y^2$ Eq. (2.14) can be rewritten as:

$$u_{p,m}^{n+1} = \left(\exp\left(r\delta_x^2\right)\exp\left(r\delta_y^2\right)\right) u_{p,m}^{n}, \text{ where } r = \frac{k}{h^2} \tag{2.15}$$

If we use Pade` approximation PA[1, 1] to approximate $\exp(r\delta_x^2)$ and $\exp(r\delta_y^2)$, then we get the Crank-Nicolson formula (1947):

$$(1-0.5r\delta_x^2)(1-0.5r\delta_y^2)u_{p,m}^{n+1} = (1+0.5r\delta_x^2)(1+0.5r\delta_y^2)u_{p,m}^{n} \tag{2.16}$$

Also Eq. (2.15) can be approximate in the form

$$u_{p,m}^{n+1} = \left(1 + k \frac{\delta^2}{\delta x^2} \right) \left(1 + k \frac{\delta^2}{\delta y^2} \right) u_{p,m}^{n} \tag{2.17}$$

by using *PA[0, 1]* for the difference operator (*1 - (1/12)* δ_x^2 *+ (1/90)* δ_x^4 *+…*), $\partial^2 / \partial x^2$ can be approximated in the form:

$$\frac{\partial^2}{\partial x^2} = \frac{1}{h^2} \frac{\delta_x^2}{1 + (1/12)\delta_x^2},$$

similarly

$$\frac{\partial^2}{\partial y^2} = \frac{1}{h^2} \frac{\delta_y^2}{1 + (1/12)\delta_y^2}.$$

Substituting in (1.2.17) and by expanding to given Douglas and Rachford formula:

$$\begin{aligned} &\left(1 - \frac{1}{2}\left(r - \frac{1}{6}\right)\delta_x^2\right)\left(1 - \frac{1}{2}\left(r - \frac{1}{6}\right)\delta_y^2\right) u_{p,m}^{n+1} \\ &= \left(1 + \frac{1}{2}\left(r + \frac{1}{6}\right)\delta_x^2\right)\left(1 + \frac{1}{2}\left(r + \frac{1}{6}\right)\delta_y^2\right) u_{p,m}^{n} \end{aligned} \tag{2.18}$$

Both these methods in Eqs. (1.2.16) and (1.2.18) are unconditionally stable. However, the use of these methods requires the solution of a large number of simultaneous linear algebraic equations at each time step. Now using restrictive Pade` approximation *RPA[1,1]* to approximate $\exp(r\delta_x^2)$, and $\exp(r\delta_y^2)$, then:

$$\begin{aligned} &(1 + \varepsilon_{p,m} r - 0.5r\delta_x^2)(1 + \varepsilon_{p,m} r - 0.5r\delta_y^2) u_{p,m}^{n+1} \\ &= (1 + \varepsilon_{p,m} r + 0.5r\delta_x^2)(1 + \varepsilon_{p,m} r + 0.5r\delta_x^2) u_{p,m}^{n} \end{aligned} \tag{2.19}$$

where the constants $\varepsilon_{p,m}$ are to be determined. Then we must know an additional condition, *(x,y,k)* to be given, i.e., $\varepsilon_{p,m}$ must be given such that the truncation error at certain r is zero, after which, use (2.19) for another level for calculations. Then the method suggested by (2.19) will be stable when $r\varepsilon_{p,m} \geq -1$, *p, m =1(1)N.*

Conclusion

The suggested method is more efficient and more accurate than many known methods, also it is of low cost. The local truncation error in our method is less than that for the classical known methods, in our method it varies from $I0^{-}8$ to 10^{-15} while in the classical known methods it varies from 10^{-2} $to10^{-9}$. The efficiency of our method is better than that for Douglass and Rachford (D.R.) method because of the ratio of the number of systems to be solved between D.R. method and this method varies from 2: 1 to 10: 1, also their method is more fast than D.R. method because the ratio of the executed time of their suggested method to D.R. method varies from 1: 2 to 1: 15 the suggested method is considered as an acceleration technique for the classical finite difference methods.

References

Ahmet Boz and Fevziye Gülsever,*(2016)* Numerical Solution of the Diffusion Equation with Restrictive Pade Approximation , Journal of Applied Mathematics and Physics, 2016, 4, 2031-2037

Baker G. A. Jr. and Morris P. G. “Pade` Approximations” Part I and Part II, Addison-Wesley, (1981).

Butcher, c (2008) Numerical Methods for Ordinary Differential Equations. John Wiley & Sons, Ltd., Chichester, second edition,

Cohn , D , (2013) erasure theory. Birkh¨auser Advanced Texts: Basler Lehrb¨ucher. Birkh¨auser/Springer, New York, second edition

D. Ambrosi, D, Arioli, G. and Koch, A (2012) homoclinic solution for excitation waves on a contractile substratum, SIAM J. Appl. Dyn. Syst. 11, no. 4, 1533–1542.

Elbarbary, E. M. (2001). Accelerating technique for implicit finite difference schemes for two dimensional parabolic partial differential equations. International journal of computer mathematics, 77(1), 135-144.

Hassan N. A Ismail and Elsayed M. E. Elbarbary “Restrictive Pade` Approximation” First National Conference on Physical & Mathematical Science & Applications (PhyMa 1) pp.63-76, Cairo, (1995).

Hassan N. A. Ismail "Unique Solvability of Restrictive Pade` and Restrictive Taylor's Approximations" Appl. Math. and Computation, 152, 89-97, (2004) (U.S.A)

Hassan N. A. Ismail and Elsayed M. E. Elbarbary "Restrictive Taylor Approximation and Parabolic Partial Differential Equations" Int. J. Comp. Math Vol. 78, pp 73-82, (2001).

Ismail and Elsayed M. E. Elbarbary "Restrictive Pade` Approximation and Parabolic Partial Differential Equation " Int. J. Comp. Math., Vol. 66, pp. 343-351,(2007).

Silviu-Ioan Filip. Robust tools for weighted Chebyshev approximation and applications to digital filter design. Signal and Image Processing. Université de Lyon, 2016. English.

Smith G. D. Numerical Solution of Partial Differential Equations Finite Difference Methods (Third Edition) Clarendon Press Oxford, (1985).

Thomas, George B., Jr.; Finney, Ross L. (1996), *Calculus and Analytic Geometry* (9th ed.), Addison Wesley.