Polynomial fitting based on least squares approximates for first-order Tracy-Widom distribution

Chong Tian, *Kejun* Lei, *Xiaojuan* Xie, *Yuhao* Tan, and *Xi* Yang[*]

College of Information Science and Engineering, Jishou University, Jishou, Hunan, China

Abstract. Tracy-Widom distribution can primely describe the limit distribution of the largest eigenvalue of noise matrix, so it is widely used in the field of signal processing and wireless communication. The exact expression of this distribution is very complex and difficult to be applied in practice. Therefore, an approximate method of Tracy-Widom distribution based on least squares is proposed, in which the optimal fitting order and optimal fitting coefficient are determined by minimizing the average fitting uncertainty. The simulation results show that the fitting results can accurately predict the largest eigenvalue distribution of the noise matrix, which proves the effectiveness of the polynomial fitting method to approximate the first-order Tracy-Widom distribution.

Keywords: First-order Tracy-Widom distribution, Largest eigenvalue, Least squares, Polynomial fit, Fitting uncertainty.

1 Introduction

The Tracy-Widom distribution is suitable for describing random phenomena of extreme eigenvalues of Wishart matrix, and is considered to have a wide range of applications [1-4]. It plays an important role in related application fields of random matrix, such as the performance analysis of MIMO communication system in fading channel, the design of spectrum detection algorithm based on eigenvalues, etc. [5-7]. Compared with common distributions such as normal distribution and exponential distribution [8, 9], the calculation of Tracy-Widom distribution includes the solution of complex differential equations and integration operations.

Therefore, it is often calculated with the help of special software in practice. For example, some discrete values of the first-order Tracy-Widom distribution CDF (Cumulative Distribution Function) and PDF (Probability Density Function) in the range of [-40,200] were given in [10]. Gamma distribution was used to approximate the Tracy-Widom distribution, and the analytical expressions of the CDF and PDF of first-order Tracy-Widom distribution within a certain range were also given in [11]. However, the approximate expressions obtained in [11] still include complicated integral operations and higher-order power operations, which will inevitably face the following two problems in practical application

[*] Corresponding author: ynkej@163.com

scenarios: (1) The computing abilities of the terminals are extremely limited, therefore the exact value of Tracy-Widom distribution cannot be calculated effectively; (2) Practical scenarios, such as cognitive radio networks and wireless communications, often have strict requirements on timeliness, extra computational complexity cannot meet this requirement [12,13].

In this paper, the polynomial fitting based on least squares is used to obtain the optimal approximates for first-order Tracy-Widom distribution, in which the optimal fitting order and fitting coefficients are determined by minimizing the average fitting uncertainty. On this basis, this paper uses Monte Carlo simulations to obtain the largest eigenvalue distribution of the noise matrix, and compares the predicted distribution with the empirical distribution. The simulation results show that as the matrix dimension changes, the approximation accuracy of the polynomial fitting approximation curve is always maintained at high level.

2 Tracy-Widom distribution and the largest eigenvalue distribution of the noise matrix

2.1 Tracy-Widom distribution

The first-order Tracy-Widom distribution is used to describe the largest eigenvalue distribution of symmetric random matrix under Gaussian ensemble, and its cumulative distribution function can be expressed as:

$$F_1(x) = \exp\left(-\frac{1}{2}\int_x^\infty q(w) + (w-x)q^2(w)\,\mathrm{d}w\right) \qquad (1)$$

Where $q(w)$ is the solution of Painlevé II differential equation:

$$q''(w) = wq'(w) + 2q^3(w) \qquad (2)$$

When $w \to \infty$, the boundary condition is $q(w) \sim \mathrm{Ai}(w)$, where $\mathrm{Ai}(w)$ is the Airy function. The probability density function of the first-order Tracy-Widom distribution is the derivative of the cumulative distribution function. It can be seen from equation (1) that it is very difficult to directly calculate CDF and PDF of first-order Tracy-Widom distribution, which is difficult to meet the requirements of practical application for computability and time limit.

2.2 Largest eigenvalue distribution of the noise matrix

Given a $M \times N$ matrix $\boldsymbol{X} = \left[\boldsymbol{X}_{(1)}, \boldsymbol{X}_{(2)}, \dots \boldsymbol{X}_{(M)}\right]^T$, the elements of the random vector $\boldsymbol{X}_{(i)}$ are real numbers and $\boldsymbol{X}_{(i)}$ is independent and identically distributed with $\mathcal{N}_N\left(0, \ \sigma_x^2 \boldsymbol{I}_{N\times N}\right)$, where $\sigma_x^2 \boldsymbol{I}_{N\times N}$ is the covariance matrix. Then the noise matrix $\boldsymbol{Y} = \boldsymbol{X}^H \boldsymbol{X}$ of order $N \times N$ will obey the white Wishart distribution $\mathcal{W}_N\left(M, \ \sigma_x^2 \boldsymbol{I}\right)$, where $\boldsymbol{X}^H$ is the complex conjugate transpose of $\boldsymbol{X}$. Assume that λ is the largest eigenvalue of the noise matrix $\boldsymbol{Y}$. When $M \to \infty$, $N \to \infty$, and $M / N \to \gamma \geq 1$, we have

$$\left(\lambda / \sigma_x^2 - \mu_{MN}\right)\Big/\sigma_{MN} \xrightarrow{\;\mathcal{D}\;} F_1 \qquad (3)$$

Where F_1 denotes the CDF of the first-order Tracy-Widom distribution. The centre and scaling parameters of the first-order Tracy-Widom distribution are as follows:

$$\mu_{MN} = (\sqrt{M-1} + \sqrt{N})^2 \qquad (4)$$

$$\sigma_{MN} = \sqrt{\mu_{MN}}\left(\frac{1}{\sqrt{M-1}} + \frac{1}{\sqrt{N}}\right)^{\!1/3} \qquad (5)$$

Note that [11] has proved that when $M \to \infty$, $N \to \infty$, and $M < N$, equation (2) and equation (3) are still valid, but the roles of M and N are reversed in equation (2) and equation (3).

3 Fitting of Tracy-Widom distribution

In this section, we use polynomial fitting method based on least squares to approximate the first-order Tracy-Widom distribution. The numeric value of PDF and CDF of the first-order Tracy-Widom distribution obtained from [10] are denoted by f_1 and F_1, respectively. Similarly, the PDF and CDF approximated by the polynomial fitting method based on least squares are denoted by fit_1 and FIT_1, respectively.

Table 1. Data points for fitting the CDF of first-order Tracy-Widom distribution.

i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	-4.8750	0.0005	7	-2.7500	0.1052	13	2.0000	0.9896
2	-4.6250	0.0011	8	-1.9375	0.2921	14	2.4375	0.9951
3	-4.4375	0.0021	9	-1.2500	0.5059	15	2.9375	0.9981
4	-4.1250	0.0053	10	-0.5625	0.7078	16	3.2500	0.9990
5	-3.8750	0.0105	11	0.4375	0.8984	17	3.6250	0.9995
6	-3.1875	0.0493	12	1.0000	0.9514			

The goal of the polynomial fitting method based on least squares is to obtain the optimal coefficients of the polynomial by limited data points so as to minimize the fitting error between the fitting function and the objective function. Suppose the polynomial fitting function expression is:

$$FIT_1(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \qquad (6)$$

Where n denotes the fitting order. In order to avoid over fitting, it is generally required that $1 \le n \le 15$.

In this paper, we select non-uniform points in the range of $x \in [-4.875, 3.625]$, and the number of points in areas where the slope of the curve changes rapidly is more than where the changes are slow. Based on the data given in [10], we use 17 data points in Table 1 to perform polynomial fitting on the CDF of the first-order Tracy-Widom distribution. Assuming that the fitting value at x_i is $FIT_1(x_i)$, the fitting error between the fitting approximation and numeric value in [10] at the data point is expressed as:

$$\delta_i = FIT_1(x_i) - y_i \qquad (7)$$

Least squares means minimizing the sum of fitting error δ_i at each data point:

$$\min \sum_{i=1}^{17} \delta_i^2 = \sum_{i=1}^{17} \left(FIT_1(x_i) - y_i \right)^2 \tag{8}$$

It can be seen from [14] that the minimization problem as in equation (8) can be transformed into finding the coefficient **a** of the following overdetermined equations:

$$Aa = Y \tag{9}$$

where

$$A = \begin{bmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{17}^n & x_{17}^{n-1} & \cdots & x_{17} & 1 \end{bmatrix} \tag{10}$$

$$a = \begin{bmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 \end{bmatrix}^T \tag{11}$$

$$Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_{17} \end{bmatrix}^T \tag{12}$$

Then the coefficient **a** can be obtained by the generalized inverse matrix **K** of matrix **A**.

In order to determine the fitting order n, we use the smallest average fitting uncertainty of sampling point in [14] to find the optimal fitting order. The uncertainty of (x_r, y_r) in the fitting interval is defined as

$$u(n, y_r) = s\sqrt{X_r KK^T X_r^T} \tag{13}$$

where

$$X_r = \begin{bmatrix} x_r^n, x_r^{n-1}, \ldots, x_r^1, 1 \end{bmatrix} \tag{14}$$

$$s = \sqrt{\sum_{i=1}^{17} \delta_i^2 / (17 - n - 1)} \tag{15}$$

The average fitting uncertainty of FIT_1 is defined as

$$u_{ave}(n) = \sqrt{\sum_{r=1}^{l} u^2(n, y_r) / l} \tag{16}$$

Within the fitting interval, l sampling points with step size $\Delta x = 0.0625$ were selected, where $l = \left[3.625 - (-4.875) \right] / \Delta x + 1$.

Using the data points given in Table 1, the average fitting uncertainty of each fitting order for fit_1 is shown in Figure 1. It can be seen from Figure 1 that the smallest average fitting uncertainty of the fitting curve is found when n=13. Therefore, the optimal polynomial fitting

order of CDF is 13. On this basis, the corresponding optimal fitting coefficient can be obtained by equation (9), as shown in Table 2.
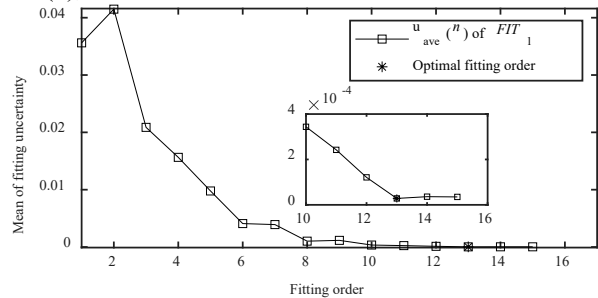

Fig. 1. Mean of fitting uncertainties of FIT_1 with different fitting orders.

Table 2. CDF polynomial fit coefficients for first-order Tracy-Widom distribution.

n	a_n	n	a_n	n	a_n
0	8.3192×10^{-1}	5	-1.6548×10^{-3}	10	6.7255×10^{-7}
1	1.8127×10^{-1}	6	-1.3343×10^{-4}	11	2.2418×10^{-7}
2	-6.9755×10^{-2}	7	1.4919×10^{-4}	12	-1.3876×10^{-8}
3	4.0452×10^{-3}	8	-8.6340×10^{-6}	13	-3.2119×10^{-9}
4	5.5729×10^{-3}	9	-7.4902×10^{-6}		

Similarly, the fitting curve of PDF of the first-order Tracy-Widom distribution can be obtained. Suppose the fitting range is $x \in [-4.875, 3.625]$, and the chosen 19 fitting data points are shown in Table 3. Using the similar method above, the average fitting uncertainty of each order of fitting polynomial fit_1 can be obtained as shown in Figure 2. It can be seen from Figure 2 that the optimal fitting order is also 13, and the corresponding fitting coefficients can be obtained by using equation (9) as shown in Table 4.

Table 3. First-order Tracy-Widom distribution PDF fitted data points.

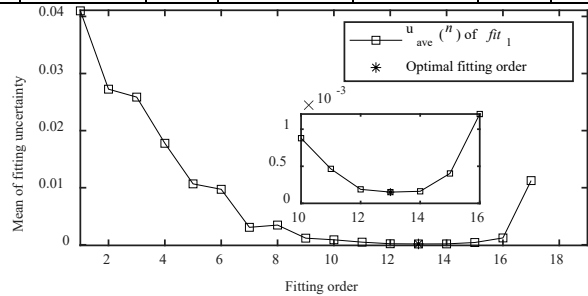| i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	-4.8750	0.0017	8	-1.6825	0.3101	15	1.6250	0.0302
2	-4.7500	0.0026	9	-1.3750	0.3192	16	2.1250	0.0144
3	-4.3125	0.0094	10	-1.0625	0.3084	17	2.7500	0.0051
4	-3.8750	0.0273	11	-0.8750	0.2936	18	3.2500	0.0021
5	-3.1250	0.1035	12	-0.1250	0.1990	19	3.6250	0.0010
6	-2.5625	0.1955	13	0.5625	0.1098			
7	-1.8750	0.2951	14	1.3125	0.0461			

Fig. 2. Mean fitting uncertainties of fit_1 with different fitting orders.

Table 4. PDF polynomial fit coefficients for first-order Tracy-Widom distribution.

n	a_n	n	a_n	n	a_n
0	1.8134×10^{-1}	5	-7.7578×10^{-4}	10	3.7026×10^{-6}
1	-1.3944×10^{-1}	6	1.1302×10^{-3}	11	-3.2423×10^{-8}
2	1.2138×10^{-2}	7	-6.9563×10^{-5}	12	-7.6029×10^{-8}
3	2.2229×10^{-2}	8	-8.3323×10^{-5}	13	-4.8278×10^{-9}
4	-8.4374×10^{-3}	9	5.7204×10^{-6}		

The polynomial fitting expression of CDF of the largest eigenvalue λ of the noise matrix can be derived from equation (3):

$$F_\lambda\left(x\right) = FIT_1\left(\frac{\left(x/\sigma_x^2\right) - \mu_{MN}}{\sigma_{MN}}\right) \tag{17}$$

From the relationship between CDF and PDF, the polynomial fitting expression of PDF of λ is the derivation of equation (17):

$$p_\lambda\left(x\right) = \frac{1}{\sigma_x^2 \sigma_{MN}} fit_1\left(\frac{\left(x/\sigma_x^2\right) - \mu_{MN}}{\sigma_{MN}}\right) \tag{18}$$

Note that the approximate expression of PDF of λ obtained in [11] is:

$$\frac{\left[x - \sigma_x^2\left(\mu_{MN} + x_0\sigma_{MN}\right)\right]^{k-1}}{\left(\sigma_x^2\sigma_{MN}\theta\right)^k \Gamma\left(k\right)} \exp\left[\frac{-\left(x - \sigma_x^2\left(\mu_{MN} + x_0\sigma_{MN}\right)\right)}{\sigma_x^2\sigma_{MN}\theta}\right] \tag{19}$$

Where $x_0 = 9.8029$, $k=46.5651$, $\theta = 0.1850$, $\Gamma\left(\bullet\right)$ is the Gamma function. Comparing equation (18) and equation (19), the expression obtained in [11] includes integral calculation and high-order power operation, and the polynomial fitting method based on least squares proposed in this paper only needs to perform 13-order integer power operation, the computational complexity is much lower than the approximate method in [11].

4 Simulation results

Figure 3 compares the approximation accuracies of CDF and PDF in this paper with the true value of first-order Tracy-Widom distribution in [10], in which the fitting curves of CDF and PDF are denoted as FIT_1 and fit_1, and the CDF and PDF in [10] are denoted F_1 and f_1, respectively. It can be seen from Figure 3 that the proposed fitting curve is highly consistent with the real value in [10] in the range $x \in \left[-4.875, 3.625\right]$, which proves that the proposed approximation method has high approximation accuracy.
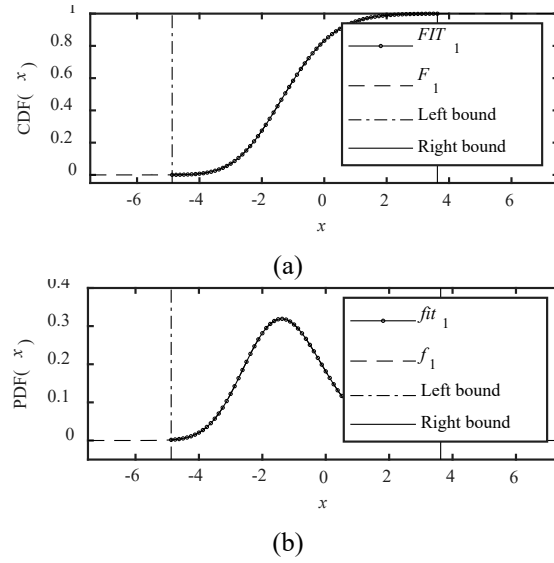
(a)

(b)

Fig. 3. Fitting graphic of first-order Tracy-Widom distribution.

In Figure 4 we use Monte Carlo simulation to evaluate the accuracy of the CDF and PDF fitting curves of the largest eigenvalue λ of the noise matrix Y with σ_x^2=1, M=20, N=40. Here, the CDF and PDF fitting curves are given by equation (17) and equation (18), respectively. In the experiment, the number of Monte Carlo experiments is 10^6. The CDF and PDF curves of Monte Carlo simulation are obtained by setting 100 intervals at equal intervals in the value range [76.8908, 143.0882]. Meanwhile, Figure 4 shows the distribution curve based on Gamma approximation in [11]. It can be seen from Figure 4 that the accuracy of the proposed fitting curve is equivalent to the approximation in [11], and both methods have extremely high approximation accuracy.
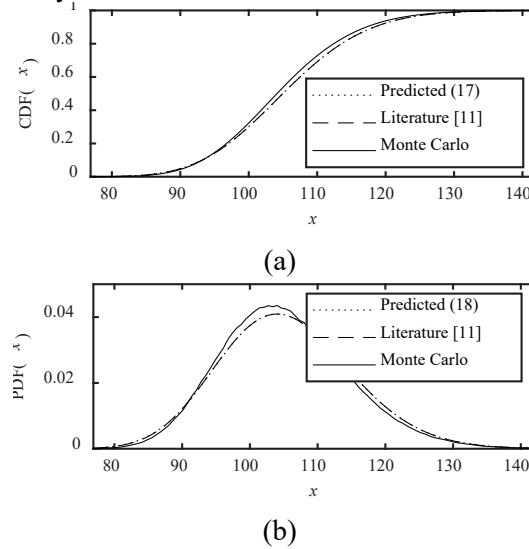

(a)

(b)

Fig. 4. Predicted CDF and PDF graph of largest eigenvalue λ.

In order to quantify the error caused by the polynomial fitting method based on least squares further and compare it with the Gamma approximation method in [11], the following index W^2 is introduced to measure the approximation accuracy:

$$W^2 = \sum_{x=L}^{R} D^2(x) p_{\lambda 1}(x) \tag{20}$$

Where $D(x) = |F_e(x) - F_\lambda(x)|$, $F_e(x)$ is the CDF of the largest eigenvalue λ of the noise matrix Y obtained from the experiment, the range is $x \in [L, R]$, L and R can be calculated by the definition of statistics in equation (3). Note that the smaller W^2 is, the higher the approximation accuracy is. W^2 of proposed approximation and Gamma approximation in [11] are shown in Table 5, where $M=N$, $\sigma_x^2 = 1$, and the number of Monte Carlo experiments for each group (M, N) is 10^6. The fitting PDF corresponding to the proposed approximation method is given by equation (17), and the PDF corresponding to the Gamma approximation method is given by equation (25) in [11]. It can be seen from Table 5 that as the M increases W^2 gradually decreases in general, which indicates that the accuracy of the proposed approximation method generally increases as the dimensions of the matrix increases. Meanwhile, the approximate accuracy of is slightly better than the method in [11]. Note that the computational complexity of the approximate expression of the distribution of λ given in this paper is much lower than that of [11]. Thus, the proposed polynomial fitting method not only has extremely high computational efficiency, but also can give accurate approximate results of the first-order Tracy-Widom distribution. Based on this, the largest eigenvalue distribution of the error matrix can also be accurately predicted.

Table 5. W^2 for the largest eigenvalue λ when $M=N$.

M	W^2	
	λ (17)	λ [11]
20	4.4301×10^{-4}	4.4205×10^{-4}
40	3.6476×10^{-4}	3.6642×10^{-4}
60	3.1816×10^{-4}	3.2020×10^{-4}
80	3.5947×10^{-4}	3.6224×10^{-4}
100	3.2309×10^{-4}	3.2582×10^{-4}
120	3.1019×10^{-4}	3.1290×10^{-4}

5 Concluding

In this paper, the polynomial fitting method based on least squares is used to approximate the CDF and PDF of first-order Tracy-Widom distribution. Here, the minimum average uncertainty method is used to find the optimal fitting order and coefficient of the fitting function with given limited data points. Both the theoretical analysis and the simulation results show that the fitting approximation of the proposed method not only has high computational efficiency, but also can ensure high approximation accuracy. The first-order Tracy-Widom distribution fitting method proposed in this paper is easy to implement and is very suitable for practical Tracy-Widom distribution application scenarios.

References

1. Su Z G 2016 Tracy-Widom distribution and its applications. J. Chinese Journal of Applied Probability and Statistics, 32: 551-580.

2. Bai Z D 1999 Methodologies in spectral analysis of large dimensional random matrices, a review. J. Statistica Sinica, 9: 611-677.

3. Tracy C A and Widom H 1996 On orthogonal and symplectic matrix ensembles. J. Communications in Mathematical Physics, 177: 727-754.

4. Zeng Y H, Cl K and Liang YC 2008 Largest eigenvalue detection: theory and application. In: Proceedings of the Symposium on Wireless Communication and Systems of ICC 2008. Bei Jing. Pp: 855-859.

5. Paysarvi-Hoseini P and Beaulieu N C 2011 Optimal wideband spectrum sensing framework for cognitive radio systems. J. IEEE Transactions on Signal Processing, 59: 1170-1182.

6. Firouzabadi A D and Rabiei A M 2015 Sensing-throughput optimization for multichannel cooperative spectrum sensing with imperfect reporting channels. J. IET Communications,9: 2188-2196.

7. Zhi Q, Cui S and Sayed A H 2008 Optimal linear cooperation for spectrum sensing in cognitive radio networks. J. IEEE Journal of Selected Topics in Signal Processing, 2: 28-40.

8. Aroian Leo A 1996 Handbook of the Normal Distribution. J. Technometrics, 25: 112-115.

9. Forbes C, Evans M, Hastings N and Peacock B 2007 Exponential Distribution. Springer, New York.

10. Michael P and Herber S 2003 Exact scaling functions for one-dimensional stationary KPZ growth. http://www-m5.ma.tum.de/KPZ.

11. Vlok J D and Olivier J C 2012 Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix. J. IET Communications, 6: 1804-1811.

12. Rohit B, Chaurasiya and Rahul S 2020 Fast sensing-time and hardware-efficient eigenvalue-based blind spectrum sensors for cognitive radio network. J. IEEE Transactions on Circuits and Systems I: Regular Papers, 67: 1296-1308.

13. Jimenez D M, Louie R H Y, Mckay M R, Chen Y 2015 Analysis and design of multiple-antenna cognitive radios with multiple primary user signals. J. IEEE Transactions on Signal Processing, 63: 4925-4939.

14. Xu J X and You Q 2020 Uncertainty Calculation for Arbitrary Order Polynomial Least-square Fitting and Analysis of the Best Fitting Order. J. Acta Metrologica Sinica, 41: 388-392.