International Journal of Mathematical Analysis and Applications
2014; 1(4): 63-67
Published online October 30, 2014 (http://www.aascit.org/journal/ijmaa)

American Association for Science and Technology

Keywords
Taylor Polynomial,
Analytic Function Approximation,
Square Error

Received: October 08, 2014
Revised: October 18, 2014
Accepted: October 19, 2014

Improving analytic function approximation by minimizing square error of Taylor polynomial

Loc Nguyen

Vietnam Institute of Mathematics, Hanoi, Vietnam

Email address

ng_phloc@yahoo.com

Citation

Loc Nguyen. Improving Analytic Function Approximation by Minimizing Square Error of Taylor Polynomial. *International Journal of Mathematical Analysis and Applications.*
Vol. 1, No. 4, 2014, pp. 63-67.

Abstract

It is very necessary to represent arbitrary function as a polynomial in many situations because polynomial has many valuable properties. Fortunately, any analytic function can be approximated by Taylor polynomial. The quality of Taylor approximation within given interval is dependent on degree of Taylor polynomial and the width of such interval. Taylor polynomial gains highly precise approximation at the point where the polynomial is expanded and so, the farer from such point it is, the worse the approximation is. Given two successive Taylor polynomials which are approximations of the same analytic function in given interval, this research proposes a method to improve the later one by minimizing their deviation so-called square error. Based on such method, the research also propose a so-called shifting algorithm which results out optimal approximated Taylor polynomial in given interval by dividing such interval into sub-intervals and shifting along with sequence of these sub-intervals in order to improve Taylor polynomials in successive process, based on minimizing square error.

1. Introduction to Taylor Polynomial

Given analytic function $f(x)$ and there exists its $n+1^{th}$ derivative, the theorem of Taylor expansion states that $f(x)$ can be approximated by a so-called Taylor polynomial $P(x)$ constructed based on high order derivatives of $f(x)$ with note that $P(x)$ is expanded at arbitrary point x_0 as follows:

$$P(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n + \frac{1}{(n+1)!}f^{(n+1)}(c)(x - x_0)^{n+1}$$

Where c is a real number between x and x_0 and c can be considered as function of x. If n is large enough, the final term is very small and is called truncation error [Burdden 2011 p. 11] denoted $R_n(x)$.

$$R_n(x) = \frac{1}{(n+1)!}f^{(n+1)}(c)(x - x_0)^{n+1}$$

When $R_n(x)$ is very small, the Taylor polynomial is approximated by removing truncate error $R_n(x)$ from it.

$$P(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n$$

Therefore, the quality of approximation is firstly dependent on how large the degree *n* of Taylor polynomial. On the other hand, Taylor polynomial gains highly precise approximation at the point x_0 where the polynomial is expanded. The farer from such point it is, the worse the approximation is. This issues the problem that how to improve quality of a Taylor polynomial in given interval if some other Taylor polynomials expanded at different points are known before. In other words, given two successive Taylor polynomials which are approximations of the same analytic function in given interval, how to improve the later one based on previous one. The problem is solved by method to minimize square error between two successive polynomials described in next section.

2. Improving Taylor Polynomial by Minimizing Square Error

Given two successive Taylor polynomials $P_1(x)$ and $P_2(x)$ which are expansions of the same analytic function $f(x)$ at two distinguish points x_1 and x_2, respectively, it is required to correct $P_2(x)$ so that it is likely that $P_2(x)$ is more approximated to $f(x)$. The square error $s(x)$ of $P_2(x)$ given $P_1(x)$ is the integral of squares of deviations between $P_2(x)$ and $P_1(x)$ over interval $[x_1, x_2]$ as follows:

$$s(x) = \int_{x_1}^{x_2} \left(P_2(x) - P_1(x)\right)^2 \mathrm{d}x$$

Note that $s(x) \geq 0$ for all x. The concept of integral square error is described in [Callahan 2008 p. 669] and it is generality of the concept of sum of square error [Montgomery 2003 p. 379].

The smaller the square error $s(x)$ is, the more approximated to $f(x)$ the $P_2(x)$ is. The polynomial $P_2(x)$ is improved by adding itself by an augmented trinomial $Q(x)$.

$$Q(x) = \alpha x^2 + \beta x + \gamma$$

$$P_2^*(x) = P_2(x) + Q(x)$$

The polynomial $P_2^*(x)$ is an improvement of $P_2(x)$ in the interval $[x_1, x_2]$ or $[x_2, x_1]$ and it is expected that $P_2^*(x)$ is approximated to $f(x)$ better than $P_1(x)$ and $P_2(x)$ are with smaller square error. It is easy to infer that the augmented $Q(x)$ is a factor that makes decrease in square error. The square error $s(x)$ is modified as follows:

$$s(x) = \int_{x_1}^{x_2} \left(P_2(x) + Q(x) - P_1(x)\right)^2 \mathrm{d}x \qquad (1)$$

Note that $s(x) \geq 0$ for all x and all coefficients α, β and γ of $Q(x)$.

The error $s(x)$ is totally determined because the inner part $\left(P_2(x) + Q(x) - P_1(x)\right)^2$ of the integral is also a polynomial with degree k^2 where k is the maximum among $deg(P_2(x))$, $deg(Q(x))$ and $deg(P_1(x))$ where $deg(.)$ denotes degree of given polynomial. It is necessary to minimize $s(x)$ to be as small as possible. If the error $s(x)$ is considered as function of its coefficients α, β and γ, then it is re-written as follows:

$$s(\alpha, \beta, \gamma) = a_1\alpha^2 + a_2\beta^2 + a_3\gamma^2 + a_4\alpha\beta + a_5\alpha\gamma + a_6\beta\gamma + a_7\alpha + a_8\beta + a_9\gamma \qquad (2)$$

Of course, equation (2) is the result of equation (1), where $a_1, a_2, \ldots,$ and a_9 are coefficients associating with variables α, β and γ. The error $s(\alpha, \beta, \gamma)$ is convex function because it is quadratic three-variable function and it is larger than or equal to 0 for all α, β and γ and so it has minimum point $(\alpha^*, \beta^*, \gamma^*)$.

The polynomial $P_2^*(x) = P_2(x) + Q(x)$ has two properties that it goes through points x_1 and x_2 with attention that $P_1(x)$ and $P_2(x)$ is expanded at x_1 and x_2, respectively and so we can infer that:

$$\begin{cases} P_2(x_1) + Q(x_1) = P_1(x_1) \\ P_2(x_2) + Q(x_2) = P_2(x_2) \end{cases}$$

Let b be $P_1(x_1) - P_2(x_1)$, we have:

$$\begin{cases} Q(x_1) = b \\ Q(x_2) = 0 \end{cases}$$

When $P_1(x_1)$ and $P_2(x_1)$ are totally evaluated and the polynomial $Q(.)$ is considered as function of α, β and γ, we have two constraints $h_1(\alpha, \beta, \gamma)$ and $h_2(\alpha, \beta, \gamma)$:

$$\begin{cases} h_1(\alpha, \beta, \gamma) = Q(x_1) - b = x_1^2\alpha + x_1\beta + \gamma - b = 0 \\ h_2(\alpha, \beta, \gamma) = Q(x_2) \qquad = x_2^2\alpha + x_2\beta + \gamma \qquad = 0 \end{cases}$$

The error $s(\alpha, \beta, \gamma)$ is minimized with regard to variables α, β and γ with two constraints $h_1(\alpha, \beta, \gamma)$ and $h_2(\alpha, \beta, \gamma)$; this is problem of convex optimization when $s(\alpha, \beta, \gamma)$ is convex function and $h_1(\alpha, \beta, \gamma)$ and $h_2(\alpha, \beta, \gamma)$ are affine functions.

$$\begin{cases} \text{minimize}_{\alpha,\beta,\gamma}\ s(\alpha, \beta, \gamma) \\ h_1(\alpha, \beta, \gamma) = x_1^2\alpha + x_1\beta + \gamma - b = 0 \\ h_2(\alpha, \beta, \gamma) = x_2^2\alpha + x_2\beta + \gamma \qquad = 0 \end{cases} \qquad (3)$$

Let ∇s, ∇h_1 and ∇h_2 be gradient vectors of $s(\alpha, \beta, \gamma)$, $h_1(\alpha, \beta, \gamma)$ and $h_2(\alpha, \beta, \gamma)$, respectively with convention that these gradient vectors are column vectors, we have:

$$\nabla s = \begin{pmatrix} 2a_1\alpha + a_4\beta + a_5\gamma + a_7 \\ 2a_2\beta + a_4\alpha + a_6\gamma + a_8 \\ 2a_3\gamma + a_5\alpha + a_6\beta + a_9 \end{pmatrix}, \nabla h_1 = \begin{pmatrix} x_1^2 \\ x_1 \\ 1 \end{pmatrix}, \nabla h_2 = \begin{pmatrix} x_2^2 \\ x_2 \\ 1 \end{pmatrix}$$

Suppose $(\alpha^*, \beta^*, \gamma^*)$ is minimum point of $s(\alpha, \beta, \gamma)$ given two constraints $h_1(\alpha, \beta, \gamma)$ and $h_2(\alpha, \beta, \gamma)$, according to Lagrange's theorem [Jia 2013] of convex optimization, there are two real numbers μ_1 and μ_2 so that $(\alpha^*, \beta^*, \gamma^*)$ is solution of following equation:

$$\nabla s + \mu_1\nabla h_1 + \mu_2\nabla h_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

It implies that

$$\begin{cases} 2a_1\alpha + a_4\beta + a_5\gamma + a_7 + \mu_1x_1^2 + \mu_2x_2^2 = 0 \\ 2a_2\beta + a_4\alpha + a_6\gamma + a_8 + \mu_1x_1 + \mu_2x_2 = 0 \\ 2a_3\gamma + a_5\alpha + a_6\beta + a_9 + \mu_1 + \mu_2 \qquad = 0 \end{cases}$$

Adding two more constraints $h_1(\alpha, \beta, \gamma) = 0$ and $h_2(\alpha, \beta, \gamma) = 0$, we have:

$$\begin{cases} 2a_1\alpha + a_4\beta + a_5\gamma + x_1^2\mu_1 + x_2^2\mu_2 = -a_7 \\ a_4\alpha + 2a_2\beta + a_6\gamma + x_1\mu_1 + x_2\mu_2 = -a_8 \\ a_5\alpha + a_6\beta + 2a_3\gamma + \mu_1 + \mu_2 = -a_9 \\ x_1^2\alpha + x_1\beta + \gamma = b \\ x_2^2\alpha + x_2\beta + \gamma = 0 \end{cases} \quad (4)$$

When (4) is the set of five linear equations with five variables α, β, γ, μ_1, and μ_2, it is easy to apply methods such as Gaussian and Cramer [Nguyen 1999 pp. 136-144] into solving (4) or determining whether or not (4) has solution. Suppose $(\alpha^*, \beta^*, \gamma^*)$ is the solution of equation (4), the polynomial $P_2^*(x)$ is totally determined:

$$P_2^*(x) = P_2(x) + Q^*(x)$$

Where,

$$Q^*(x) = \alpha^*x^2 + \beta^*x + \gamma^*$$

Let s^* be minimum value of square error function $s(\alpha, \beta, \gamma)$ at minimum point $(\alpha^*, \beta^*, \gamma^*)$, we have:

$$s^* = s(\alpha^*, \beta^*, \gamma^*)$$

The minimum mean error r is defined as the root of mean of s^*:

$$r = \sqrt{\frac{s^*}{|x_2 - x_1|}}$$

Given a very small threshold ε, if the minimum mean error r is determined and it is smaller than or equal to ε, then $P_2^*(x)$ is the improved version of $P_2(x)$, which results out the optimal approximation of target function $f(x)$. If r is determined and larger than ε, it is impossible to improve $P_2(x)$. If r is not determined, for example, s^* is not found out, then there is no conclusion about whether $P_2^*(x)$ is the improvement of $P_2(x)$ or not.

3. Shifting Algorithm to Approximate Analytic Function

It is required to approximate an analytic function $f(x)$ in a given interval $[u, v]$. Suppose the interval $[u, v]$ is divided into n sub-intervals as follows:

$$[u, v] = [a_0, a_1] \cup [a_1, a_2] \cup \ldots \cup [a_{i-1}, a_i] \cup \ldots \cup [a_{n-1}, a_n]$$

Where,

$$u = a_0 < a_1 < a_2 < \cdots < a_{i-1} < a_i < \cdots < a_n = v$$

Moving from left to right (from a_0 to a_n), the shifting algorithm is to construct current Taylor polynomial at point a_i and improve such current polynomial based on minimizing the aforementioned square error between the current polynomial and previous polynomial expanded at point a_{i-1}. This is shift-and-improve process and the algorithm is stopped when such process reaches point a_n.

Given an error threshold ε, suppose the algorithm moves to point a_i and let $P_{i-1}(x)$ and $P_i(x)$ be Taylor polynomials expanded at point a_{i-1} and a_i, respectively, there are three cases:

1. If the minimum mean error r_i of $P_i(x)$ given $P_{i-1}(x)$ is not determined, then the algorithm moves next. Note that the method to calculate r_i is described in previous section.
2. If the minimum mean error r_i is determined and larger than threshold ε, then the algorithm moves next.
3. If the minimum mean error r_i is determined and smaller than (or equal to) threshold ε, then the *polynomial $P_i(x)$ is replaced by polynomial $P_i^*(x)$* where $P_i^*(x)$ is the improved version of $P_i(x)$ as aforementioned in previous section. After that the algorithm moves next.

Recall that $P_i^*(x)$ is:

$$P_i^*(x) = P_i(x) + Q_i^*(x)$$

Where,

$$Q_i^*(x) = \alpha_i^*x^2 + \beta_i^*x + \gamma_i^*$$

Note that $(\alpha_i^*, \beta_i^*, \gamma_i^*)$ is the minimum point of square error function $s_i(\alpha_i, \beta_i, \gamma_i)$ at point a_i like in equations (1) and (2).

$$s_i(\alpha_i, \beta_i, \gamma_i) = \int_{a_{i-1}}^{a_i} \left(P_i(x) + Q(x) - P_{i-1}(x)\right)^2 \mathrm{d}x$$

The minimum value s_i^* of square error function $s_i(\alpha, \beta, \gamma)$ is:

$$s_i^* = s_i(\alpha_i^*, \beta_i^*, \gamma_i^*)$$

The minimum mean error r_i is:

$$r_i = \sqrt{\frac{s_i^*}{|a_i - a_{i-1}|}}$$

Finally, when the algorithm reaches point a_n, then $P_n(x)$ is the best approximation of target function $f(x)$.

For example, given exponent function $f(x) = e^x$, we apply minimizing square error method and shifting algorithm into approximating $f(x)$ in interval [0, 1] with initial degree 1. For convenience, the interval [0, 1] is kept intact, which means that there is only one sub-interval [0, 1]. Firstly, shifting algorithm visits the first point $x_1 = 0$ and so Taylor polynomial expansion of $f(x)$ at $x_1 = 0$ is:

$$P_1(x) = f(0) + f'(0)x = 1 + x$$

The shifting algorithm moves next and the Taylor polynomial expansion of $f(x)$ at the second $x_2 = 1$ is:

$$P_2(x) = f(1) + f'(1)x = ex$$

Suppose the augmented $Q(x)$ is:

$$Q(x) = \alpha x^2 + \beta x + \gamma$$

Substituting x_1, x_2, $P_1(x)$ and $P_2(x)$ into equation (2), the square error is:

$$s(\alpha,\beta,\gamma)=\int_0^1\left(P_2(x)+Q(x)-P_1(x)\right)^2\mathrm{d}x=\frac{\alpha^2}{5}+\frac{\beta^2}{3}+\gamma^2$$

Let $P_2^*(x)$ be the improvement of $P_2(x)$, we have:

$$P_2^*(x)=P_2(x)+Q^*(x)$$

$$Q^*(x)=\alpha^*x^2+\beta^*x+\gamma^*$$

Where $(\alpha^*, \beta^*, \gamma^*)$ is the solution of equation (4).

After evaluating $P_1(x_1)$ and $P_2(x_2)$, the equation (4) is totally determined and solved as follows:

$$\begin{cases}4\alpha+5\beta+10\mu_1=5-5e\\3\alpha+4\beta+\ \ 6\mu_1=4-4e\\2\alpha+3\beta+\ \ 3\mu_2=3-3e\\\ \alpha\ +\ \beta\qquad\quad\ =-1\\\qquad\qquad\ \gamma\ \ =1\end{cases}\Rightarrow\begin{cases}\alpha^*=\frac{5e-10}{2}\\\beta^*=-\frac{5e-8}{2}\\\gamma^*=1\end{cases}$$

The augmented polynomial $Q^*(x)$, the optimal Taylor polynomial $P_2^*(x)$, the square error s^* and the minimum mean error r are determined as follows:

$$Q^*(x)=\frac{5e-10}{2}x^2-\frac{5e-8}{2}x+1$$

$$P_2^*(x)=P_2(x)+Q^*(x)=\frac{5e-10}{2}x^2-\frac{3e-8}{2}x+1$$

$$s^*=s(\alpha^*,\beta^*,\gamma^*)=\frac{(e-2)^2}{8}\approx 0.0645$$

$$r=\sqrt{\frac{s^*}{|x_2-x_1|}}=\sqrt{\frac{0.0645}{|1-0|}}\approx 0.254$$

Suppose the error threshold is 0.5 which is larger than r, the polynomial $P_2^*(x)$ is exactly the improvement of $P_2(x)$. The shifting algorithm reaches the end point $x_2 = 1$ and the final optimal Taylor polynomial in given interval [0, 1] is:

$$P_2^*(x)=\frac{5e-10}{2}x^2-\frac{3e-8}{2}x+1$$

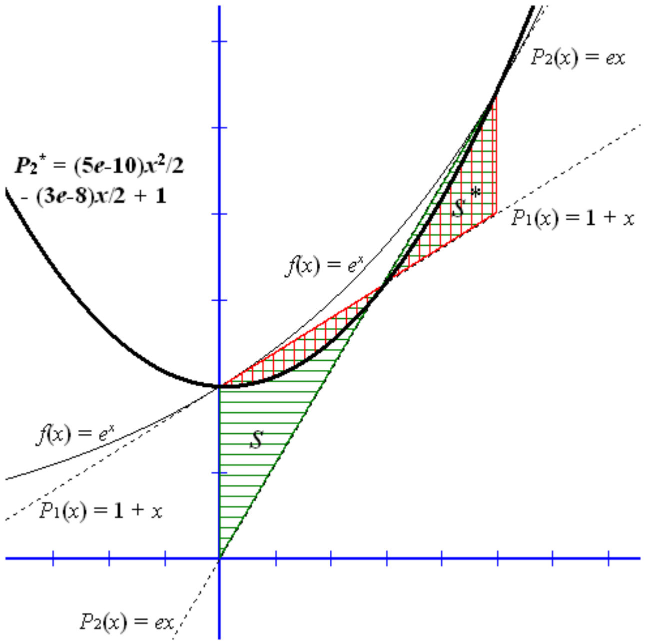

Figure 1. *Optimal Taylor polynomial* $P_2^*(x)$

Figure 1 depicts the optimal Taylor polynomial $P_2^*(x)$ expanded in interval [0, 1].

Note that horizontal shading area represents square error s between $P_1(x)$ and $P_2(x)$ while vertical shading area represents square error s^* between $P_1(x)$ and $P_2^*(x)$. These two areas are overlapped but s^* is much smaller than s. Two polynomial $P_1(x)$ and $P_2(x)$ are drawn as dash lines while the optimal polynomial $P_2^*(x)$ is drawn as bold curve. The target function $f(x) = e^x$ is drawn as normal curve.

4. Conclusion

Given target function $f(x)$, the approximated Taylor polynomial gets more precise at so-called expansion point where it is expanded and it tends to lose accuracy when target function $f(x)$ moves far away from expansion point. The essential idea of proposed method is to keep approximation in precise when Taylor polynomial expansion is moved forward within a given interval. Concretely, suppose $P_1(x)$ and $P_2(x)$ are Taylor polynomials expanded at x_1 and x_2, respectively. Of course, $P_2(x)$ is gain exactly precise approximation at x_2 but its effectiveness at x_1 is lower than $P_1(x)$ expanded x_1. Therefore, $P_2(x)$ is modified by adding itself by an augmented trinomial $Q(x)$, which aiming to minimize the square error between $P_1(x)$ and $P_2(x)$ so that it is likely that $P_2(x)$ keeps approximation in precise within sub-interval $[x_1, x_2]$. The important aspect of proposed method is to determine the trinomial $Q(x)$ by minimizing square error, which is essentially polynomial interpolation. It is possible to imagine that $Q(x)$ is the bridge concatenating two polynomial $P_1(x)$ and $P_2(x)$ together.

Note that the square error is calculated as integral of deviation between $P_1(x)$ and $P_2(x)$, which means that if all Taylor polynomials are bad approximations, the output of proposed method will is also bad approximation of $f(x)$. However, there are two observations:

- Taylor polynomial always results out optimal approximation at which it is expanded, thus, there is no so bad Taylor polynomial. The quality of Taylor polynomial is also dependent on its degree.
- All Taylor polynomials converge to target function $f(x)$ and so the deviations between effective Taylor polynomials approach 0. Therefore, it is feasible to calculate the square error between Taylor polynomials.

If we construct Taylor polynomial with degree k, then the final optimal polynomial resulted from shifting algorithm has degree which is maximum of k and 2. You can modify the proposed method to interpolate $Q(x)$ with high degree (> 2) with expect that getting more accurate approximation but please pay attention to computation cost when equation (4) should not has many variables because of many constraints and it is very complicated to determined the integral in equation (1) with high degree polynomials. Finally, there is an issued problem that how to estimate the initial degree k in order to improve the quality of Taylor expansion, which is solved in another research.

References

[1] [Burdden 2011]. Burdden, R. L., Faires, J. D. Numerical Analysis, ninth edition. Brooks/Cole Cengage Learning Publisher Copyright 2011, 2005, 2001.

[2] [Callahan 2008] James Callahan, J., Hoffman, K., Cox, D., O'Shea, D., Pollatsek, H., Senechal, L. Calculus in Context. The Five College Calculus Project Copyright 1994, 2008.

[3] [Jia 2013]. Jia, Y. B. Lagrange Multipliers. Lecture notes on course "Problem Solving Techniques for Applied Computer Science", Iowa State University of Science and Technology, USA. Last checked 2 September 2014.

[4] [Montgomery 2003]. Montgomery, D. C., Runger, G. C. Applied Statistics and Probability for Engineers, third edition. John Wiley & Sons Publisher Copyright 2003.

[5] [Nguyen 1999]. Nguyen, H. V. H. Linear Algebra, second edition. Hanoi Education Publisher Copyright 1999.