

# Notes on Approximation Theory

Peter Occil

This version of the document is dated 2026-05-12.

Some notes that may be useful when finding approximation error bounds that are explicit, with no hidden constants and without introducing transcendental or trigonometric functions.

The notes generally relate to error bounds on how close a polynomial is to a single-variable function on a closed interval. The mapping from a function to a function (in this case, from a single-variable function to a polynomial “close” to it) is called an *operator*, and operators involved in these error bounds are often linear operators, whose behavior is relatively simple to examine.

## 1 Contents

- Contents
- Definitions
- “Moments” of Linear Operators
- Taylor Expansion of Linear Operators
- Results on Error Bounds
  - Whitney’s Inequality
  - Lebesgue Inequality for Certain Linear Operators
- Example
- License
- Notes

## 2 Definitions

The *closed unit interval* (written as  $[0, 1]$ ) means the set consisting of 0, 1, and every real number in between.

For definitions of *continuous*, *derivative*, *convex*, *concave*, *Hölder continuous*, and *Lipschitz continuous*, see the definitions section in “**Supplemental Notes for Bernoulli Factory Algorithms**<sup>1</sup>”.

- An *operator* is a mapping from functions to functions.
- An operator  $L$  is *linear* if it satisfies  $L(af) = aL(f)$  and  $L(f + g) = L(f) + L(g)$  for all input functions  $f$  and  $g$  and every real number  $a$ .
- An operator  $L$  is *positive* if it has the property that, if  $f$  is nonnegative on its domain, so is  $L(f)$ , for every input function  $f$ .
- The *operator norm* of an operator  $L$  is the maximum absolute value of  $L(f)$  over all input functions  $f$  with maximum absolute value 1 or less. This assumes  $L$  takes only continuous functions.
- In this document,  $e_i$  is a function such that  $e_i(t) = t^i$ , so that  $e_1(t) = t$ ; as an example, if  $L(f) = f(0) + f(1)$ , then  $L(e_1 - x) = (e_1(0) - x) + (e_1(1) - x) = (0 - x) + (1 - x) = 1 - 2x$ .

---

<sup>1</sup><https://peteroupc.github.io/bernsupp.html#Definitions>

### 3 Bernstein Form and Bernstein Polynomials

In this document, a polynomial  $P(x)$  is written in *Bernstein form of degree  $n$*  if it is written as—

$$P(x) = \sum_{k=0}^n a_k \frac{n!}{(k!((n-k)!)} x^k (1-x)^{n-k},$$

where the real numbers  $a_0, \dots, a_n$  are the polynomial's *Bernstein coefficients*.

The degree- $n$  *Bernstein polynomial* of an arbitrary function  $f(x)$  has Bernstein coefficients  $a_k = f(k/n)$ . In general, this Bernstein polynomial differs from  $f$  even if  $f$  is a polynomial. In this section, the degree- $n$  Bernstein polynomial of  $f$  is denoted  $B_n(f)$ .  $B_n(f)$  is a positive linear operator.

### 4 “Moments” of Linear Operators

To examine the approximation behavior of linear operators, it is helpful to find the so-called “moments” of those operators, that is, the functions they map certain polynomials to.

For a linear operator  $L$ , they are:

- “Raw moments”: The values of  $L(e_i)(x)$  for each integer  $i \geq 0$ .
- “Central moments”: The values of  $L((e_1 - x)^i)$  for each integer  $i \geq 0$ . If the “raw moments”  $L(e_0), \dots, L(e_j)$  are known, then  $L((e_1 - x)^j)$  is also known, thanks to proposition 5.6 of Gonska et al. (2006)<sup>2</sup>.
- “Absolute moments”: The values of  $L(\text{abs}(e_1 - x)^i)(x)$  for each integer  $i \geq 0$ . When  $i$  is even,  $L(\text{abs}(e_1 - x)^i) = L((e_1 - x)^i)$ .

### 5 Taylor Expansion of Linear Operators

Let  $f(\lambda)$  have a continuous  $s$ -th derivative on a closed interval, where  $s$  is zero or a positive integer, and let  $L(f)$  be a linear operator. Then:

$$L_n(f)(\lambda) = L_n(R(f, \lambda)) + \sum_{i=0}^s L_n((e_1 - \lambda)^i)(\lambda) \frac{f^{(i)}(\lambda)}{i!},$$

where  $R(f, \lambda)$  is the remainder of  $f$  after subtracting the degree- $s$  Taylor polynomial of  $f$  centered at  $\lambda$ . (See also Pițul (2007, proof of theorem 5.8)<sup>3</sup>.)

### 6 Results on Error Bounds

Some results on error bounds for certain classes of operators.

---

<sup>2</sup>Gonska, Heiner, Paula Pi ul, and Ioan Rașa. “On differences of positive linear operators.” *Carpathian Journal of Mathematics* (2006): 65-78.

<sup>3</sup>Pițul, P., “Evaluation of the Approximation Order by Positive Linear Operators”, dissertation, Universität Duisberg-Essen, 2007.

## 6.1 Whitney's Inequality

Let  $n$  be zero or a positive integer, let  $f(\lambda)$  be continuous on a closed interval  $[a, b]$ , and let  $P$  be a polynomial of degree  $n$  or less with the least maximum absolute difference between  $f$  and the polynomial on that interval. Then the error of  $P$  in approximating  $f$  is bounded as follows (see Babenko and Kryakin 2019<sup>4</sup>):

$$\|f - P\|_\infty \leq W \cdot \omega_{n+1}\left(f, \frac{b-a}{n+1}\right),$$

where—

- $W$  is:
  - 1 if  $n \leq 7$ .
  - $(2 + \exp(-2)) < 2.13534$  if  $n \geq 8$ .
  - $3/4$  if  $n = 1$  and  $f$  is convex (Singh Kaire and Prymak 2023/2025)<sup>5</sup>.
  - $1/2$  if  $n = 1$ ,  $f$  is convex, and  $a = -b$  (Singh Kaire and Prymak 2023/2025)<sup>6</sup>.
- $\|g\|_\infty$  is the maximum of the absolute value of (the continuous function)  $g$  on  $[a, b]$ , and
- $\omega_n(f, h)$  is the modulus of continuity of  $f$  of order  $n$ , with parameter  $h$ .

Using properties of the modulus of continuity (see Sevy 1991<sup>7</sup>, sec. 2.0.2; Gonska 1985<sup>8</sup>), if  $f$  has a continuous  $(n+1)$ -th derivative on  $[a, b]$ :

$$\|f - P\|_\infty \leq W \cdot \left(\frac{b-a}{n+1}\right)^{n+1} \|f^{(n+1)}\|_\infty,$$

and if  $f$  has a continuous  $n$ -th derivative on that interval:

$$\|f - P\|_\infty \leq W \cdot \left(\frac{b-a}{n+1}\right)^n \omega_1(f^{(n)}, \frac{b-a}{n+1}).$$

## 6.2 Lebesgue Inequality for Certain Linear Operators

Let  $f(\lambda)$  be a continuous function on a closed interval. For any sequence of linear operators  $(L_n)$  that map continuous functions to polynomials and reproduce all polynomials up to degree  $m(n)$  (which depends on  $n$ ), the following error bound (also known as *Lebesgue's lemma* or the *Lebesgue inequality*) holds true for each  $n$ :

$$\text{abs}(L_n(f)(x) - f(x)) \leq (1 + \|L_n\|) \cdot \max_t (\text{abs}(f(t) - P(t))),$$

where  $\|L_n\|$  is the operator norm of  $L_n$ , and  $P$  is a polynomial of degree up to  $m(n)$  with the least maximum absolute difference between  $f$  and the polynomial (see also DeVore and Lorentz (1993)<sup>9</sup>, Cheney (1996, chapter 6)<sup>10</sup>). But this error bound will generally be crude or trivial unless  $L_n$  are nonpositive

<sup>4</sup>Babenko, Alexander G., and Yuriy V. Kryakin. "Special difference operators and the constants in the classical Jackson-type theorems." Topics in Classical and Modern Analysis: In Memory of Yingkang Hu. Cham: Springer International Publishing, 2019. 35-46.

<sup>5</sup>Jaskaran Singh Kaire and Andriy Prymak. "Whitney-type estimates for convex functions." arXiv preprint arXiv:2311.00912 (2023).

<sup>6</sup>Jaskaran Singh Kaire and Andriy Prymak. "Whitney-type estimates for convex functions." arXiv preprint arXiv:2311.00912 (2023).

<sup>7</sup>Sevy, J., "Acceleration of convergence of sequences of simultaneous approximants", dissertation, Drexel University, 1991.

<sup>8</sup>H. H. Gonska, *Quantitative Approximation in  $C(X)$* , Habilitationsschrift, Universität Duisburg, 1985.

<sup>9</sup>R.A. DeVore and G.G. Lorentz, *Constructive Approximation*, 1993.

<sup>10</sup>E. W. Cheney, *Introduction to Approximation Theory*, 1998.

operators. Indeed, the only positive linear operator  $L$  that reproduces all polynomials up to degree 2 is the identity operator  $L = f$ .<sup>11</sup>

Now let  $f$  have a continuous third derivative on the closed unit interval. Combining the previous inequality with the Whitney-type inequalities in the previous section leads to the following error bound for linear operators  $L$  that map continuous functions to polynomials and reproduce all polynomials up to degree 2:

$$\begin{aligned} \text{abs}(L(f)(x) - f(x)) &\leq (1 + \|L\|) \cdot 1 \cdot \left(\frac{1}{3}\right)^3 \|f^{(3)}\|_\infty \\ &= (1 + \|L\|)\|f^{(3)}\|_\infty/27. \end{aligned}$$

## 7 Example

This example shows how to find a linear operator's bounds.

Let  $L_n(f)$  be a (nonpositive) linear operator described as follows. The functions it maps to and from lie on the closed unit interval.

Let  $W_n = 2f - B_n(f)$ . Then  $B_n(W_n(f))$  is a (nonpositive) linear operator that is the iterated Boolean sum of degree- $n$  Bernstein polynomials, with one iteration; see Güntürk and Li (2021a, Theorem 5)<sup>12</sup>. That paper, among others (for example, Micchelli 1973<sup>13</sup>), showed that  $B_n(W_n(f)) = O(1/n^{3/2})$  if  $f$  has a continuous third derivative. (“ $= O(1/n^{3/2})$ ” means the absolute value is no greater than a constant times  $1/n^{3/2}$  for all values of  $n$ .)

$L_n(f)$  is then based on  $W_n(f)$  and is a special case of **a conjecture I have**<sup>14</sup> on polynomial approximation. Let  $k = 2n\lambda$ , where  $0 \leq \lambda \leq 1$ . Let—

$$\begin{aligned} L_n(f)(\lambda) &= \sum_{i=0}^n \left( (2f\left(\frac{k}{2n}\right) - B_{2n}(f)\left(\frac{k}{2n}\right)) - (2f\left(\frac{i}{n}\right) - B_n(f)\left(\frac{i}{n}\right)) \right) \sigma_{n,k,i} \\ &= \mathbb{E} \left[ (2f\left(\frac{k}{2n}\right) - B_{2n}(f)\left(\frac{k}{2n}\right)) - (2f\left(\frac{X_k}{n}\right) - B_n(f)\left(\frac{X_k}{n}\right)) \right] \\ &= \sum_{i=0}^n \left( (2f\left(\frac{k}{2n}\right) + B_n(f)\left(\frac{i}{n}\right)) \right) \sigma_{n,k,i} - \sum_{i=0}^n \left( (2f\left(\frac{i}{n}\right) + B_{2n}(f)\left(\frac{k}{n}\right)) \right) \sigma_{n,k,i} \\ &= LA_n(f)(\lambda) - LB_n(f)(\lambda). \end{aligned}$$

Here,  $\sigma_{n,k,i}$  equals  $\binom{n}{i} \binom{n}{k-i} / \binom{2n}{k}$  and is the probability of getting  $i$ . (This is the probability that a hypergeometric( $2n, k, n$ ) random variable equals  $i$ .) Meanwhile,  $LA_n$  and  $LB_n$  are positive linear operators, making it easier to assess their approximation properties.

<sup>11</sup>Guessab, A., Nouisser, O. & Schmeisser, G. Enhancement of the algebraic precision of a linear operator and consequences under positivity. *Positivity* 13, 693–707 (2009). <https://doi.org/10.1007/s11117-008-2253-4>

<sup>12</sup>Güntürk, C. Sinan, and Weilin Li. “**Approximation with one-bit polynomials in Bernstein form**”, arXiv:2112.09183 (2021); *Constructive Approximation*, pp.1-30 (2022). <https://arxiv.org/pdf/2112.09183>

<sup>13</sup>Micchelli, Charles. “**The saturation class and iterates of the Bernstein polynomials**”, *Journal of Approximation Theory* 8, no. 1 (1973): 1-18. <https://www.sciencedirect.com/science/article/pii/0021904573900282>

<sup>14</sup>[https://peteroupc.github.io/bernsupp.html#A\\_Conjecture\\_on\\_Polynomial\\_Approximation](https://peteroupc.github.io/bernsupp.html#A_Conjecture_on_Polynomial_Approximation)

It will be shown that, if  $f$  has a continuous third derivative, the rate of  $L_n$  is  $O(M/n^{3/2})$ , where  $M$  is the maximum absolute value of  $f$  and its derivatives up to the third derivative. The proof of this relies on exact expressions of  $L_n$ 's “**raw moments**” and “**central moments**”, and those for the combined operator ( $LA_n + LB_n$ ).

The following are some of these values and those for related operators:

- $L_n(e_0)(x) = L_n((e_1 - x)^0)(x) = 0.$
- $L_n(e_1)(x) = L_n((e_1 - x)^1)(x) = 0.$
- $L_n(e_2)(x) = L_n((e_1 - x)^2)(x) = 3x(x - 1)/(2n(2n - 1)) = O(1/n^2).$
- $L_n(e_3)(x) = n^3x^2(2nx - 4x + 3)/(2n - 1).$
- $LA_n((e_1 - x)^2)(x) = -x(3n - 2) \cdot (x - 1)/(n(2n - 1)) = O(1/n).$
- $LB_n((e_1 - x)^2)(x) = -x(6n - 1) \cdot (x - 1)/(2n(2n - 1)) = O(1/n).$
- $(LA_n + LB_n)((e_1 - x)^2)(x) = LA_n(\text{abs}(e_1 - x)^2)(x) + LB_n(\text{abs}(e_1 - x)^2)(x) = -x(12n - 5) \cdot (x - 1)/(2n(2n - 1)) = O(1/n).$

To find values like those just listed, it is useful to calculate raw moments (Wang et al. 2023)<sup>15</sup> and central moments (Weisstein)<sup>16</sup> of hypergeometric random variables. Let  $X_k$  be a hypergeometric( $2n, k, n$ ) random variable. If  $g(y) = W_{2n}(e_r; k/(2n)) - W_n(e_r; y)$  is a polynomial in  $y$  of degree  $r$  or less (which, in this case, it is), then  $L_n(e_r)$  can be found using a Taylor expansion, namely as—

$$\begin{aligned} L_n(e_r) &= \sum_{i=0}^r \mathbb{E}[(X_k/n - \mathbb{E}[X_k/n])^i] \frac{g^{(i)}(\mathbb{E}[X_k/n])}{i!} \\ &= \sum_{i=0}^r \frac{\mathbb{E}[(X_k - \mathbb{E}[X_k])^i]}{n^i} \frac{g^{(i)}(k/(2n))}{i!}, \end{aligned}$$

where the derivatives are taken with respect to  $y$ , and where  $\mathbb{E}[(X_k - \mathbb{E}[X_k])^i]$  is the  $i$ -th central moment of  $X_k$ .

In the following, the notation  $\| \| f \| \|$  means  $\max_{0 \leq \lambda \leq 1} (\text{abs}(f(\lambda)))$ .

The first step is to **find the Taylor expansion** of  $L_n(f)(\lambda)$ . Given that  $L_n((e_1 - x)^0)(x) = L_n((e_1 - x)^1)(x) = 0$ , this becomes:

$$L_n(f)(\lambda) = L_n(R(f, \lambda)) + \sum_{i=2}^3 L_n((e_1 - \lambda)^i)(\lambda) \frac{f^{(i)}(\lambda)}{i!},$$

$$\text{abs}(L_n(f)(\lambda)) \leq \|L_n(R(f, \lambda))\| + \|L_n((e_1 - \lambda)^2)\| \|f^{(2)}\|/2$$

$$+ \|L_n((e_1 - \lambda)^3)\| \|f^{(3)}\|/6.$$

The function  $\text{abs}(L_n((e_1 - x)^3)(x))$  has its maximum at  $x = 1/2 - \sqrt{3}/6$ ; and  $\text{abs}(L_n((e_1 - x)^2)(x))$  has its maximum at  $x = 1/2$ , so:

<sup>15</sup>Wang, Y.Q., Zhang, Y.Y, Liu, J.L., “Expectation identity of the hypergeometric distribution and its application in the calculations of high-order origin moments”, Communications in Statistics–Theory and Methods 52(17), 2023. <https://doi.org/10.1080/03610926.2021.2024235>

<sup>16</sup>Weisstein, Eric W. “Central Moment.” From MathWorld—A Wolfram Resource. <https://mathworld.wolfram.com/CentralMoment.html>

$$\begin{aligned}
\text{abs}(L_n(f)(\lambda)) &\leq \|L_n(R(f, \lambda))\| + \text{abs}\left(\frac{3\lambda(\lambda-1)}{2n(2n-1)}\right)\|f^{(2)}\|/2 \\
&\quad + \|L_n((e_1 - \lambda)^3)\| \|f^{(3)}\|/6 \\
&\leq \|L_n(R(f, \lambda))\| + \frac{3}{8n(2n-1)}\|f^{(2)}\|/2 \\
&\quad + \frac{\sqrt{3}(6n-5)}{24n^2(2n-1)}\|f^{(3)}\|/6.
\end{aligned}$$

Meanwhile the remainder is estimated as follows, using the proof of corollary 2.3 of Gonska et al. (2006)<sup>17</sup>:

$$\|L_n(R(f, \lambda))\| \leq \frac{1}{6}\|f^{(3)}\| \|(LA_n + LB_n)(\text{abs}(e_1 - \lambda)^3)\|.$$

In turn, using Schwarz's inequality (see proof of the same paper's corollary 2.1):

$$\begin{aligned}
\|(LA_n + LB_n)(\text{abs}(e_1 - \lambda)^3)\| &\leq (\|(LA_n + LB_n)((e_1 - \lambda)^4)\|)^{1/2} \\
&\quad \times (\|(LA_n + LB_n)((e_1 - \lambda)^2)\|)^{1/2} \leq \frac{3\sqrt{3}}{8n^{3/2}}.
\end{aligned}$$

(The expression in the middle takes its maximum at  $\lambda = 1/2$ ; the right-hand side is an upper bound of that expression for all positive integers  $n$ .) Altogether:

$$\begin{aligned}
\|L_n(f)\| &\leq \frac{3}{8n(2n-1)}\frac{1}{2}\|f^{(2)}\| \\
&\quad + \left(\frac{3\sqrt{3}}{8n^{3/2}} + \frac{\sqrt{3}(6n-5)}{24n^2(2n-1)}\right)\frac{1}{6}\|f^{(3)}\| = LC_n(f) \\
&\leq 0.1875\frac{\|f^{(2)}\|}{n^{3/2}} + \frac{5\sqrt{3}}{72}\frac{\|f^{(3)}\|}{n^{3/2}} \leq \frac{0.3078M}{n^{3/2}} = O(1/n^{3/2}).
\end{aligned}$$

If  $n \geq 2$  is an integer,  $LC_n(f) \leq 0.2165M/n^{3/2}$ .

## 8 License

Any copyright to this page is released to the Public Domain. In case this is not possible, this page is also licensed under **Creative Commons Zero**<sup>18</sup>.

## 9 Notes

<sup>17</sup>Gonska, Heiner, Paula Pi ul, and Ioan Raşa. "On differences of positive linear operators." Carpathian Journal of Mathematics (2006): 65-78.

<sup>18</sup><https://creativecommons.org/publicdomain/zero/1.0/>