Euler-Maclaurin の和公式の一般化

黒木 玄

2017 年 7 月 24 日作成 *

概 要

論文 [2] では, 実数直線上の確率分布に対して, Bernoulli 多項式の一般化が構成されている. しかし, その論文では, Euler-Maclaurin の和公式の一般化については, 未解決であるとしている ([2], Remark 4). このノートでは彼らが構成した Bernoulli 多項式の一般化を用いて, Euler-Maclaurin の和公式を一般化する. 一般化された公式は積分剰余項型の Taylor の定理と Euler-Maclaurin の和公式と Euler-Boole の和公式を含んでおり, それらを統一する定理になっている.

https://genkuroki.github.io/documents/20170724EulerMaclaurin.pdf

目 次

1 Bernoulli 多項式の一般化

この節では Borwein-Calkin-Manna (2009) [2] に基いて, 実数直線 $\mathbb{R}$ 上の確率分布に対する Bernoulli 多項式の一般化 $P_n(x)$ を構成する. この節で定めた記号は後の節でもそのまま利用する.

*2017 年 7 月 24 日: 作成. 7 月 26 日 (Ver.0.01): まだ書きかけの段階で公開. 今後も気が向いたら更新することにする. 7 月 26 日 (Ver.0.02): 剰余項の評価に関する第 2.3 節を追加した. 他にも色々訂正した.

1.1 Bernoulli 多項式

読者の便のために, 一般化する前の古典的 Bernoulli 多項式について復習しておこう.
Bernoulli 多項式 $B_n(x)$, $n = 0, 1, 2, \dots$ は次の母函数表示によって定義される:

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{z^n}{n!}.$$

たとえば

$$\begin{aligned}
B_0(x) &= 1, \\
B_1(x) &= x - \frac{1}{2}, \\
B_2(x) &= x^2 - x + \frac{1}{6}, \\
B_3(x) &= x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \\
B_4(x) &= x^4 - 2x^3 + x^2 - \frac{1}{30}, \\
B_5(x) &= x^5 - \frac{5}{2}x^4 + \frac{5}{3}x^3 - \frac{1}{6}x, \\
B_6(x) &= x^6 - 3x^5 + \frac{5}{2}x^4 - \frac{1}{2}x^2 + \frac{1}{42}.
\end{aligned}$$

$B_n = B_n(0)$ は **Bernoulli 数**と呼ばれている. $B_n(x)$ の母函数表示で $x = 0$ とおいて $z/2$ を足すと,

$$\frac{z}{e^z - 1} + \frac{z}{2} = \frac{z}{2} \frac{e^{z/2} + e^{-z/2}}{e^{z/2} - e^{-z/2}}$$

と z に関する奇函数になる. このことから, n が 3 以上の奇数のとき, Bernoulli 数 B_n が 0 になることがわかる.

Bernoulli 多項式の母函数表示を用いて次の公式を示すことができる:

$$\begin{aligned}
B_n'(x) &= nB_{n-1}(x), \\
B_n(x+h) &= \sum_{k=0}^{n} \binom{n}{k} B_k(x) h^{n-k}.
\end{aligned}$$

これらの公式より, Bernoulli 多項式の計算は, Bernoulli 数の計算に帰着することがわかる. それらの公式が成立することは以下の計算を見ればわかる:

$$\begin{aligned}
&\frac{d}{dx}\frac{ze^{xz}}{e^z - 1} = z\frac{ze^{xz}}{e^z - 1}, \qquad z\frac{z^{n-1}}{(n-1)!} = n\frac{z^n}{n!}, \\
&\frac{ze^{(x+h)z}}{e^z - 1} = \frac{ze^{xz}}{e^z - 1}e^{hz} = \sum_{k,l=0}^{\infty} \frac{B_k(x) h^l z^{k+l}}{k!\,l!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \binom{n}{k} B_k(x) h^{n-k} \right) \frac{z^n}{n!}.
\end{aligned}$$

このようにして証明される上の公式は形式的には次のようにして得られる:

(1) まず $B_n(x)$ を変数 B の多項式 B^n に置き換える. その B^n を B で微分した後に B^{n-1} を $B_{n-1}(x)$ で置き換える.

(2) まず $B_n(x+h)$ を変数 B の多項式 $(B+h)^n$ で置き換える．二項定理を用いて $(B+h)^n$ を展開した後に B^k を $B_k(x)$ で置き換える．

これはどうしてだろうか?
この疑問は以下の節で解説する文献 [2] の構成によって解消される [1]．

1.2 実数直線上の確率分布

この節では後の節で用いる実数直線上の確率分布の記号法を定める．$\mathbb{R}$ 上の確率測度を P と書き, その累積分布函数を $F(x)$ と書き, そのモーメント母函数を $M(z)$ と書く．詳しくは以下を参照せよ [2]．

P は $\mathbb{R}$ 上の**確率測度** (確率分布) であるとする．

確率測度の例として以下を知っておけばこのノートを読むには十分である [3]．

例 1.1 (デルタ分布). $\int_{\mathbb{R}} \varphi(x)\,P(dx) = \varphi(0)$．このような P は原点に台を持つ**デルタ分布**と呼ばれる．Dirac のデルタ超函数の記号を用いて $P(dx) = \delta(x)\,dx$ と書くこともある:

$$\int_{\mathbb{R}} \varphi(x)\,P(dx) = \int_{\mathbb{R}} \varphi(x)\delta(x)\,dx = \varphi(0).$$

この確率測度は, x の値が確率 1 で 0 になるような確率分布を表現している． □

例 1.2 (コイン投げ分布). $w_i \geqq 0$, $\sum_i w_i = 1$ であるとし, a_i は任意の実数列であるとする．このとき

$$P(dx) = \sum_i w_i \delta(x - a_i)\,dx$$

によって, 離散的確率測度 P を定めることができる．すなわち

$$\int_{\mathbb{R}} \varphi(x)\,P(dx) = \sum_i w_i \varphi(a_i).$$

例えば i は $0, 1$ を動き, $a_0 = 0$, $a_1 = 1$, $w_0 = w_1 = 1/2$ のとき

$$\int_{\mathbb{R}} \varphi(x)\,P(dx) = \frac{1}{2}\int_{\mathbb{R}} \varphi(x)(\delta(x) + \delta(x-1))\,dx = \frac{\varphi(0) + \varphi(1)}{2}.$$

この確率測度は, x の値がそれぞれお確率 $1/2$ で 0 または 1 にになる確率分布を表現している．この確率分布をこのノートでは仮に**コイン投げ分布**と呼ぶことにする． □

例 1.3 (正規分布, 一様分布). $p(x)$ が非負 Lebesgue 可測函数で $\int_{\mathbb{R}} p(x)\,dx = 1$ を満たすとき, $P(dx) = p(x)\,dx$ によって確率測度 P が定義され, $p(x)$ は P の確率密度函数と呼ばれる．このとき

$$\int_{\mathbb{R}} \varphi(x)\,P(dx) = \int_{\mathbb{R}} \varphi(x)p(x)\,dx.$$

---

[1] すぐに答えを知りたい人は第 1.4 節の系 1.16 の証明を参照せよ．
[2] この部分節で解説されている概念と記号法に慣れている読者はこの部分節をとばして先に進んでよい．
[3] 以下の例を知っていれば測度論について知っている必要はない．

この確率測度は x の値が $a \leqq x \leqq b$ となる確率が $\int_a^b p(x)\,dx$ であるような確率分布を表現している. 例えば, **標準正規分布**は確率密度函数

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

で与えられ, $[0,1]$ 区間上の**一様分布**は確率密度函数

$$p(x) = \begin{cases} 1 & (x \in [0,1]) \\ 0 & (x \notin [0,1]) \end{cases}$$

によって与えられる. □

確率測度 P の**累積分布函数** $F(x)$ が次のように定義される:

$$F(x) = \int_{(-\infty,x]} P(dx).$$

例えば P が原点に台を持つデルタ分布のとき

$$F(x) = \begin{cases} 0 & (x < 0) \\ 1 & (x \geqq 0). \end{cases}$$

P が確率密度函数 $p(x)$ で与えられている場合には

$$F(x) = \int_{-\infty}^x p(y)\,dy.$$

例 1.4. 例えば P が $[0,1]$ 区間上の一様分布のとき

$$F(x) = \begin{cases} 0 & (x < 0) \\ x & (x \in [0,1]) \\ 1 & (x > 1). \end{cases}$$

さらに P がコイン投げ分布の場合には

$$F(x) = \begin{cases} 0 & (x < 0) \\ 1/2 & (0 \leqq x < 1) \\ 1 & (1 \leqq x). \end{cases}$$

これらについては $\int_0^1 F(x)\,dx = 1/2$ が成立している. □

これ以後ずっと, 確率測度 P は, 任意の非負の整数 n に対して

$$\int_{\mathbb{R}} |x|^n \, P(dx) < \infty$$

を満たしていると仮定する. このとき確率測度 P の n 次の**モーメント** M_n が

$$M_n = \int_{\mathbb{R}} x^n \, P(dx)$$

によって定義される．$M_0 = 1$ であり，$\int_{\mathbb{R}} x\, P(dx) = M_1$ は確率分布 P の**平均**もしくは**期待値**と呼ばれ，

$$\int_{\mathbb{R}} (x - M_1)^2\, P(dx) = \int_{\mathbb{R}} (x^2 - 2M_1 x + M_1^2)\, P(dx) = M_2 - M_1^2$$

は確率分布 P の**分散**と呼ばれる．

確率分布 P の**モーメント母函数** $M(z)$ が

$$M(z) = \sum_{n=0}^{\infty} M_n \frac{z^n}{n!}$$

と定義される．収束半径が 0 の場合には z の形式べき級数とみなす．$M_0 = 1$ より，モーメント母函数の逆数 $1/M(z)$ も z の形式べき級数とみなせる．べき級数として収束する場合に，モーメント母函数は

$$M(z) = \int_{\mathbb{R}} e^{xz}\, P(dx)$$

とも書ける．大抵の場合，この公式でモーメント母函数が計算される．

注意 1.5. 確率論におけるモーメント母函数 $M(z)$ は統計力学における分配函数に対応する数学的対象である．$z = -\beta$ と書き，$M(z)$ を $Z(\beta)$ と書くと，

$$Z(\beta) = \int_{\mathbb{R}} e^{-\beta x}\, P(dx).$$

このように書けば，統計力学における記号法との対応が付け易いだろう．β は物理的には絶対温度の逆数であり，逆温度と呼ばれる．分配函数は Boltzmann 因子 $e^{-\beta x}$ でベースになる確率分布 P を変形するために使われる．そのときベースになる確率分布 P はミクロカノニカル分布と呼ばれ，Boltzmann 因子で変形された確率分布はカノニカル分布と呼ばれる．カノニカル分布の導出に関する詳しい解説については，筆者による解説ノート [3] を参照せよ． □

例 1.6. 例えば P が原点に台を持つデルタ分布の場合には

$$M(z) = \int_{\mathbb{R}} e^{xz} \delta(x)\, dx = 1$$

となる．平均も分散も 0 になる． □

例 1.7. P が $[0,1]$ 区間上の一様分布の場合には

$$M(z) = \int_0^1 e^{xz}\, dx = \left[\frac{e^{xz}}{z}\right]_{x=0}^{x=1} = \frac{e^z - 1}{z}$$

となり，Bernoulli 数の母函数の逆数に一致する．$M_k = 1/(k+1)$ となり，平均は $1/2$，分散は $1/12$ になる． □

例 1.8. P がコイン投げ分布の場合には

$$M(z) = \frac{e^{0z} + e^{1z}}{2} = \frac{1 + e^z}{2}.$$

$M_k = 1/2$ $(k \geqq 1)$ となり，平均は $1/2$，分散は $1/4$ になる． □

次の節で，原点に台を持つデルタ分布には多項式達 x^n が対応し，$[0,1]$ 区間上の一様分布には Bernoulli 多項式達 $B_n(x)$ が対応することが説明される．コイン投げ分布に対応する多項式達は Euler 多項式達と呼ばれている．

1.3 Bernoulli 多項式の一般化

前節で定めたように P は $\mathbb{R}$ 上の確率測度であるとし, $F(x)$ はその累積分布函数であるとし, $M(x)$ はそのモーメント母函数であるとする.

定義 1.9 (一般化された Bernoulli 多項式). 次の母函数表示で**一般化された Bernoulli 多項式** $P_n(x)$ を定める:

$$\frac{e^{xz}}{M(z)} = \sum_{n=0}^{\infty} P_n(x) \frac{z^n}{n!}.$$

文献 [2] でこの $P_n(x)$ は **Strodt 多項式**と呼ばれている. □

$M(z) = 1 + M_1 z + M_2 z^2/2 + \cdots$ より

$$P_0(x) = 1, \qquad P_1(x) = x - M_1, \qquad P_2(x) = x^2 - 2M_1 x + 2M_1^2 - M_2$$

となることがわかる. $P_2(x)$ はさらに

$$P_2(x) = (x - M_1)^2 - (M_2 - M_1^2)$$

とも書ける. M_1 は確率分布 P の期待値であり, $M_2 - M_1^2$ は分散であることに注意せよ.

例 1.10 (原点に台を持つデルタ分布の場合). P が原点に台を持つデルタ分布のとき, すなわち $P(dx) = \delta(x)\,dx$ のとき, モーメント母函数は $M(x) = 1$ となるので,

$$\frac{e^{xz}}{M(x)} = e^{xz} = \sum_{n=0}^{\infty} x^n \frac{z^n}{n!}$$

より $P_n(x) = x^n$ となる. すなわち x^n は原点に台を持つデルタ分布に対応する一般化された Bernoulli 多項式になっている. □

例 1.11 (Bernoulli 多項式). P が $[0,1]$ 区間上の一様分布のとき, そのモーメント母函数 $M(z)$ は

$$M(x) = \int_0^1 e^{xz}\,dx = \frac{e^z - 1}{z}$$

となるので, 一般化された Bernoulli 多項式 $P_n(x)$ は

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} P_n(x) \frac{z^n}{n!}$$

によって定められる. これは Bernoulli 多項式の定義に一致する. すなわち, $[0,1]$ 区間上の一様分布に対応する一般化された Bernoulli 多項式は Bernoulli 多項式に一致する. □

例 1.12 (Euler 多項式). P がコイン投げ分布のとき, すなわち $P(dx) = (1/2)(\delta(x)+\delta(x-1))\,dx$ のとき, 対応する一般化された Bernoulli 多項式は $E_n(x)$ と書かれ, **Euler 多項式**と呼ばれる. この場合にモーメント母函数は $M(z) = (1+e^z)/2$ になるので Euler 多項式 $E_n(x)$ は

$$\frac{2e^{xz}}{1+e^z} = \sum_{n=0} E_n(x) \frac{z^n}{n!}$$

によって定義される. □

注意 1.13. もしも $e^{xz}/M(z)$ がすべての実数 x について収束していれば, $(e^{xz}/M(z))\,P(dx)$ は確率測度になっていることに注意せよ. 実際,

$$\int_{\mathbb{R}} \frac{e^{xz}}{M(z)}\,P(dx) = \frac{\int_{\mathbb{R}} e^{xz}\,P(dx)}{M(z)} = \frac{M(x)}{M(x)} = 1.$$

確率測度 $(e^{xz}/M(z))\,P(dx)$ は統計力学におけるカノニカル分布に対応している. $z = -\beta$ とおき, $M(z) = Z(\beta)$ と書くと,

$$\frac{e^{xz}\,P(dx)}{M(z)} = \frac{e^{-\beta x}\,P(dx)}{Z(\beta)}.$$

このように書けば統計力学における記号の対応を付け易いだろう. この方面に関する詳しい解説が [3] にある. 物理的には β は絶対温度の逆数である. だから $z = -\beta$ に関するべき級数展開は絶対温度 ∞ における展開である. 一般化された Bernoulli 多項式はカノニカル分布の高温展開によって定義されていると考えられる. □

1.4　一般化された Bernoulli 多項式の特徴付け

この節でも, P は $\mathbb{R}$ 上の確率測度であるとし, $F(x)$ はその累積分布函数であるとし, $M(x)$ はそのモーメント母函数であるとする. 確率分布 P に対応する一般化された Bernoulli 多項式を $P_n(x)$ と表す:

$$\frac{e^{xz}}{M(z)} = \sum_{n=0}^{\infty} P_n(x)\frac{z^n}{n!}, \qquad M(z) = \int_{\mathbb{R}} e^{xz}\,P(dx).$$

函数 $f(x)$ に対して $\mathscr{A}[f](x)$ を次のように定める:

$$\mathscr{A}[f](x) = \int_{\mathbb{R}} f(x+y)\,P(dy).$$

このノートでは, $\mathscr{A}[f]$ を f の確率測度 P による**移動平均 (moving average)** と呼ぶことにする.

例えば, P が原点に台を持つデルタ分布のとき

$$\mathscr{A}[f](x) = \int_{\mathbb{R}} f(x+y)\delta(y)\,dy = f(x)$$

となり, $\mathscr{A}$ は単なる恒等写像になる.

例えば, P が $[0,1]$ 区間上の一様分布ならば

$$\mathscr{A}[f](x) = \int_0^1 f(x+y)\,dy = \int_x^{x+1} f(t)\,dt$$

となり, $\mathscr{A}[f]$ は f の幅 1 の区間にわたる前方移動平均になる.

さらに P がコイン投げ分布ならば, すなわち $P(dx) = (1/2)(\delta(x)+\delta(x-1))\,dx$ ならば

$$\mathscr{A}[f](x) = \frac{1}{2}\int_{\mathbb{R}} f(x+y)(\delta(y)+\delta(y-1))\,dy = \frac{f(x)+f(x+1)}{2}$$

となり, $\mathscr{A}[f]$ は f の離散的な前方移動平均になる.

補題 1.14. 移動平均作用素 $\mathscr{A}$ は x^n を n 次のモニック多項式に移す. ゆえに $\mathscr{A}$ は多項式全体の空間の線形自己同型を与える. さらに $\mathscr{A}$ は微分作用素 d/dx や差分作用素 $f(x) \mapsto f(x+h)$ と可換である.

証明. 作用素 $\mathscr{A}$ は多項式を多項式に移す. なぜならば

$$\mathscr{A}[x^n] = \int_{\mathbb{R}} (x+y)^n \, P(dy) = \sum_{k=0}^{n} \binom{n}{k} x^k \int_{\mathbb{R}} y^{n-k} \, P(dy) = \sum_{k=0}^{n} \binom{n}{k} M_{n-k} x^k.$$

この公式より, 移動平均作用素 $\mathscr{A}$ は x^n 次多項式をモニックな n 次多項式に移すことがわかる. したがって, $\mathscr{A}$ の多項式函数への制限は, 多項式全体の空間の線形自己同型を与える.

$\mathscr{A}$ は d/dx と可換であることは次のようにして確かめられる:

$$\frac{d}{dx}\mathscr{A}[f](x) = \frac{d}{dx}\int_{\mathbb{R}} f(x+y) \, P(dy) = \int_{\mathbb{R}} f'(x+y) \, P(dy) = \mathscr{A}[f'](x).$$

$\mathscr{A}$ は $f(x) \mapsto f(x+h)$ と可換であることは次のようにして確かめられる:

$$\mathscr{A}[f(x+h)] = \int_{\mathbb{R}} f((x+y)+h) \, P(dy) = \int_{\mathbb{R}} f((x+h)+y) \, P(dy) = \mathscr{A}[f](x+h).$$

以上によって上の補題が成立していることがわかった. □

定理 1.15 (一般化された Bernoulli 多項式の特徴付け). 一般化された Bernoulli 多項式 $P_n(x)$ は次の条件によって一意に特徴付けられる.

$$\mathscr{A}[P_n(x)] = \int_{\mathbb{R}} P_n(x+y) \, P(dy) = x^n.$$

証明. 補題 1.14 より, $\mathscr{A}$ は多項式全体の空間の線形自己同型を定めるので, もしも $P_n(x)$ が $\mathscr{A}[P_n(x)] = x^n$ という条件を満たすならばその条件で $P_n(x)$ は一意に特徴付けられる.

$$\mathscr{A}\left[\frac{e^{xz}}{M(z)}\right] = \frac{1}{M(z)}\int_{\mathbb{R}} e^{(x+y)z} \, P(dy) = \frac{e^{xz}}{M(z)}\int_{\mathbb{R}} e^{yz} \, P(dy) = e^{xz}.$$

両辺を z に関して展開すれば $\mathscr{A}[P_n(x)] = x^n$ が成立していることがわかる. □

系 1.16. 一般化された Bernoulli 多項式は以下を満たす:

$$\frac{d}{dx}P_n(x) = nP_{n-1}(x),$$
$$P_n(x+h) = \sum_{k=0}^{n} \binom{n}{k} P_k(x) h^{n-k}.$$

証明. 補題 1.14 より, $\mathscr{A}$ は d/dx と可換なので,

$$\mathscr{A}\left[\frac{d}{dx}P_n(x)\right] = \frac{d}{dx}\mathscr{A}[P_n(x)] = \frac{d}{dx}x^n = nx^{n-1} = \mathscr{A}[nP_{n-1}(x)].$$

$\mathscr{A}$ は多項式全体の空間の線形自己同型なので $P_n'(x) = nP_{n-1}(x)$.

後者の公式を示そう. 二項展開より,

$$(x+h)^n = \sum_{k=0}^{n} \binom{n}{k} x^k h^{n-k}.$$

$\mathscr{A}$ が差分作用素 $f(x) \mapsto f(x+h)$ と可換なことから, 二項展開の公式中の $(x+h)^n$, x^k をそれぞれ $P_n(x+h)$, $P_k(x)$ で置き換えた公式も成立することがわかる. □

系 1.17. 一般化された Bernoulli 多項式 $P_n(x)$ 達は以下の条件によって帰納的に一意に特徴付けられる: $n \geqq 1$ のとき,

$$P_0(x) = 1, \qquad \frac{d}{dx} P_n(x) = n P_{n-1}(x), \qquad \int_{\mathbb{R}} P_n(x)\, P(dx) = 0.$$

証明. $P_0(x) = 1$, $P_n'(x) = nP_{n-1}(x)$ が成立することはすでに示されている. さらに $\mathscr{A}[P_n(x)] = x^n$ は $n \geqq 1$ のとき $x = 0$ で 0 になることから, $\int_{\mathbb{R}} P_n(y)\, P(dy) = 0$ となることもわかる.

$P_n'(x) = nP_{n-1}(x)$ より, $P_n(x)$ は $P_{n-1}(x)$ から積分定数を除いて一意に決定される. そして, その積分定数は $\int_{\mathbb{R}} P_n(x)\, P(dx) = 0$ という条件から一意に決定される. ゆえに, $P_0(x) = 1$ から出発して, $P_n(x)$ 達が帰納的に一意に決定されることがわかる. □

系 1.17 は一般化された Euler-Maclaurin の和公式 (定理 2.2) の証明で使われる補題 2.1 を示すために使われる.

注意 1.18. 以上は文献 [2] の議論の引き写しである. 文献 [2] では, 一般化された Bernoulli 多項式 (Strodt 多項式) $P_n(x)$ に関するこのノートでは扱っていない性質を取り扱っている. 例えば, その最後の節では $n \to \infty$ での $P_n(x)$ の漸近挙動に関する予想を提出している. □

2 Euler-Maclaurin の和公式の一般化

この節では, Taylor の定理と Euler-Maclaurin の定理の一般化を確立する.

P は $\mathbb{R}$ 上の確率分布であり, $F(x)$ はその累積分布函数であり, $M(x)$ はそのモーメント母函数であり, $P_n(x)$ はそれに対応する一般化された Bernoulli 多項式であるとする. さらに $\mathscr{A}$ は次のように定義された移動平均作用素であるとする:

$$\mathscr{A}[f](x) = \int_{\mathbb{R}} f(x+y)\, P(dy).$$

函数 $f(x)$ は十分に滑らかであり, 遠方での増大度も大き過ぎないと仮定する.

2.1 Taylor の定理

Taylor の定理は以下のようにして証明される. 積分型平均値の定理より,

$$f(x+h) = f(x) + \int_0^h f'(x+y_1)\, dy_1.$$

この公式の f, h, y_1 を f', y_1, y_2 で置き換えた結果を右辺の積分の中に代入すると,

$$f(x+h) = f(x) + hf'(x) + \int_0^h dy_1 \int_0^{y_1} f''(x+y_2)\, dy_2.$$

同じことを再度繰り返すと

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \int_0^h dy_1 \int_0^{y_1} dy_2 \int_0^{y_2} f''(x+y_3)\, dy_3.$$

同様の操作を繰り返すことによって次が得られる:

$$f(x+h) = f(x) + hf'(x) + \cdots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x) + R_n,$$
$$R_n = \int_0^h dy_1 \int_0^{y_1} dy_2 \cdots \int_0^{y_{n-1}} f^{(n)}(x+y_n)\, dy_n.$$

剰余項 R_n は, 積分の順序を y_n が一番最後になるように変え, $y_1, \dots, y_{n-1}$ による積分を実行し, y_n を y に置き換えることによって,

$$R_n = \int_0^h \frac{(h-y)^{n-1}}{(n-1)!} f^{(n)}(x+y)\, dy$$

と表されることもわかる. 以上の結果を**積分剰余項型の Taylor の定理**と呼ぶことにする.

すぐ上の剰余項の形は部分積分を繰り返すことによっても導出可能である. 実際, R_n がすでに上の形をしているならば,

$$-\frac{d}{dy}\frac{(h-y)^n}{n!} = \frac{(h-y)^{n-1}}{(n-1)!}$$

を用いた部分積分によって,

$$R_n = \left[-\frac{(h-y)^n}{n!} f^{(n)}(x+y) \right]_{y=0}^{y=h} + \int_0^h \frac{(h-y)^n}{n!} f^{(n+1)}(x+y)\, dy$$
$$= \frac{h^n}{n!} f^{(n)}(x) + \int_0^h \frac{(h-y)^n}{n!} f^{(n+1)}(x+y)\, dy.$$

この結果を使うことによっても, 積分型剰余項型の Taylor の定理を帰納的に証明することができる.

2.2 一般化された Euler-Maclaurin-Taylor の公式

以下では, 計算が書き下し易くなるように,

$$P(dx) = p(x)\, dx$$

と書き, 確率測度 P が確率密度函数 $p(x)$ を持つかのように扱う. 実際に確率密度函数を持つ場合には以下の議論はそのままに正しい. そうでない場合にも適切に訂正すれば以下の議論は正しいとみなせる.

函数 $K(y, y_1)$ を次のように定める:

$$K(y, y_1) = F(y_1) - H(y_1 - y).$$

ここで $F(y_1)$ は確率測度 $P(dy_1) = p(y_1)\, dy_1$ の累積分布函数であり, $H(y_1 - y)$ は Heaviside 函数である. すなわち,

$$F(y_1) = \int_{-\infty}^{y_1} p(y_1)\, dy_1, \qquad H(y_1 - y) = \begin{cases} 0 & (y_1 < y), \\ 1 & (y_1 \geqq y). \end{cases}$$

累積分布函数 $F(y_1)$ は単調増加函数であり, $y_1 \to -\infty$ で $F(y_1) \to 0$ となり, $y_1 \to \infty$ で $F(y_1) \to 1$ となる．このことから, 任意の $y, y_1 \in \mathbb{R}$ に対して $|K(y, y_1)| \leqq 1$ となり, y を固定して $y_1 \to \pm\infty$ とすると $K(y, y_1) \to 0$ となることがわかる.

各 y に対して $K(y, y_1)$ は $y \to \pm\infty$ は十分速く 0 に近付くと仮定する．この仮定は曖昧なので以下で少しコメントしておく.

確率密度函数 $p(x)$ の台が有界で区間 $[a, b]$ に含まれるならば[4], 各 y ごとに y_1 の函数 $K(y, y_1) = F(y_1) - H(y_1 - y)$ の台は区間 $[\min\{a, y\}, \max\{b, y\}]$ に含まれるので, 上の曖昧な仮定は成立しているとみなされる.

確率密度函数 $p(x)$ の台が有界でなくても, $p(x)$ が例えば正規分布の場合には上の曖昧な仮定は成立していると考える．より一般に $p(x)$ が Schwartz の意味で急減少函数であれば上の曖昧な仮定は成立していると考える.

函数 $K(y, y_1)$ は以下を満たしている:

$$
\begin{aligned}
&\frac{\partial}{\partial y_1} K(y, y_1) = p(y_1) - \delta(y_1 - y), \\
&\frac{\partial}{\partial y} K(y, y_1) = \delta(y_1 - y), \\
&\int_{\mathbb{R}} dy\, p(y) K(y, y_1) = F(y_1) - \int_{-\infty}^{y_1} dy\, p(y_1) = 0.
\end{aligned}
$$

積分作用素 $\mathscr{K}$ を次のように定める:

$$
\mathscr{K}[\varphi](y) = \int_{\mathbb{R}} dy_1\, K(y, y_1) \varphi(y_1).
$$

補題 2.1. 積分作用素 $\mathscr{K}$ は以下を満たしている:

$$
\begin{aligned}
&\mathscr{K}[\varphi'](y) = \int_{\mathbb{R}} dy_1\, K(y, y_1) \varphi'(y_1) = \varphi(y) - \int_{\mathbb{R}} dy_1\, p(y_1) \varphi(y_1), \\
&\frac{d}{dy} \mathscr{K}[\varphi](y) = \frac{d}{dy} \int_{\mathbb{R}} dy_1\, K(y, y_1) \varphi(y_1) = \varphi(y), \\
&\int_{\mathbb{R}} dy\, p(y) \mathscr{K}[\varphi](y) = 0, \\
&\mathscr{K}^n[1](y) = \frac{P_n(y)}{n!}.
\end{aligned}
$$

ここで $P_n(x)$ は確率測度 $P(dy) = p(y)\, dy$ に対応する一般化された Bernoulli 多項式 (Strodt 多項式) である.

証明. 部分積分によって

$$
\begin{aligned}
\mathscr{K}[\varphi'](y) &= \int_{\mathbb{R}} dy_1\, K(y, y_1) \varphi'(y_1) \\
&= -\int_{\mathbb{R}} dy_1\, (p(y_1) - \delta(y_1 - y)) \varphi(y_1) = \varphi(y) - \int_{\mathbb{R}} dy_1\, p(y_1) \varphi(y_1).
\end{aligned}
$$

$K_y(y, y_1) = \delta(y_1 - y)$ より

$$
\frac{d}{dy} \mathscr{K}[\varphi](y) = \frac{d}{dy} \int_{\mathbb{R}} dy_1\, K(y, y_1) \varphi(y_1) = \int_{\mathbb{R}} dy_1\, \delta(y_1 - y) \varphi(y_1) = \varphi(y).
$$

[4] これは $x \notin [a, b]$ ならば $p(x) = 0$ となるという意味.

$\int_{\mathbb{R}} dy\, p(y)K(y,y_1)=0$ より

$$
\begin{aligned}
\int_{\mathbb{R}} dy\, p(y)\mathscr{K}[\varphi](y) &= \int_{\mathbb{R}} dy \int_{\mathbb{R}} dy_1\, p(y)K(y,y_1) \\
&= \int_{\mathbb{R}} dy_1 \int_{\mathbb{R}} dy\, p(y)K(y,y_1) = \int_{\mathbb{R}} dy_1\, 0 = 0.
\end{aligned}
$$

系 1.17 より, $\mathscr{K}^n[1](y)=P_n(y)/n!$ を示すためには

$$
\mathscr{K}^0[1](y)=1, \qquad \frac{d}{dy}\mathscr{K}^n[1](y)=\mathscr{K}^{n-1}(y), \qquad \int_{\mathbb{R}} dy\, p(y)\mathscr{K}^n[1](y)=0
$$

を示せば十分である．1 つ目の条件は自明に成立している．2 つ目の条件は上で示した $(d/dy)\mathscr{K}[\varphi](y)=\varphi(y)$ より成立している．3つ目の条件は上で示した $\int_{\mathbb{R}} dy\, p(y)\mathscr{K}[\varphi](y)=0$ より成立している．これで示すべきことがすべて示された． □

定理 2.2 (一般化された Euler-Maclaurin の公式). 以下の公式が成立している:

$$
\begin{aligned}
f(x+h) &= \sum_{k=0}^{n-1} \frac{P_n(h)}{n!} \int_{\mathbb{R}} dy\, p(y) f^{(k)}(x+y) + R_n, \\
R_n &= \int_{\mathbb{R}} dy_1 \cdots \int_{\mathbb{R}} dy_n\, K(y,y_1)K(y_1,y_2)\cdots K(y_{n-1},y_n)f(x+y_n).
\end{aligned}
$$

証明. 補題 2.1 を $\varphi(y)=f(x+y)$ に適用すると,

$$
\begin{aligned}
f(x+h) &= \int_{\mathbb{R}} p(y_1)f(x+y_1) + \int_{\mathbb{R}} dy_1\, K(h,y_1)f(x+y_1) \\
&= \mathscr{A}[f](x) + \mathscr{K}[y \mapsto f'(x+y)](h)
\end{aligned}
$$

Taylor の定理の証明と同様に, この公式中の左辺 $h \mapsto f(x+h)$ を $y \mapsto f'(x+y)$ に置き換えたものを右辺の $f'(x+y)$ に代入すると,

$$
f(x+h) = \mathscr{A}[f](x) + \mathscr{K}[1](h)\,\mathscr{A}[f'](x) + \mathscr{K}^2[y \mapsto f''(x+y)](h).
$$

この式を得るときに, y に関する定数函数 $y \mapsto \mathscr{A}[f'](x)$ に関しては

$$
\mathscr{K}\bigl[y \mapsto \mathscr{A}[f'](x)\bigr](h) = \mathscr{K}[1](h)\,\mathscr{A}[f'](x)
$$

が成立していることを使った．同じことをもう一度実行すると,

$$
\begin{aligned}
f(x+h) = \mathscr{A}[f](x) &+ \mathscr{K}[1](h)\,\mathscr{A}[f'](x) + \mathscr{K}^2[1](h)\,\mathscr{A}[f''](x) \\
&+ \mathscr{K}^3[y \mapsto f^{(3)}(x+y)](h).
\end{aligned}
$$

同様に繰り返せば次が得られる:

$$
\begin{aligned}
f(x+h) &= \sum_{k=0}^{n-1} \mathscr{K}^k[1](h)\,\mathscr{A}[f^{(k)}](x) + R_n, \\
R_n &= \mathscr{K}^n[y \mapsto f^{(n)}(x+y)](h).
\end{aligned}
$$

補題 2.1 より $\mathscr{K}^k[1](h)=P_k(h)/k!$ であるので, 移動平均作用素 $\mathscr{A}$ と積分作用素 $\mathscr{K}$ の定義に戻って, この公式を積分を使って書き直せば求める結果が得られる． □

注意 2.3. 上の定理 2.2 の R_n を**剰余項**と呼ぶことにする. 剰余項 R_n を

$$
\begin{aligned}
R_n &= \int_{\mathbb{R}} dy\, K_n(h,y) f^{(n)}(x+y),\\
K_n(h,y) &= \int_{\mathbb{R}} dy_1 \cdots \int_{\mathbb{R}} dy_{n-1} K(h,y_1)K(y_1,y_2)\cdots K(y_{n-1},y)
\end{aligned}
$$

と表わすこともできる. □

例 2.4 (Taylor の定理). $p(x)=\delta(x)$ の場合には, $P_n(x)=x^n$ であるから, 定理 2.2 より, Taylor の定理

$$
f(x+h) = \sum_{k=0}^{n-1} \frac{x^k}{k!} f^{(k)}(x) + R_n
$$

が得られる. □

$\mathbb{R}$ の部分集合 A に対して, $x \in A$ のとき 1 になり, $x \in \mathbb{R} \setminus A$ のとき 0 になるような函数を $\chi_A(x)$ と書き, A の**特性函数**と呼ぶことにする.

例 2.5 (Euler-Maclaurin の和公式). $p(x)=\chi_{[0,1]}(x)$ ($[0,1]$ 上の一様分布) の場合には,

$$
\begin{aligned}
&\int_{\mathbb{R}} dy\, p(y) f(x+y) = \int_x^{x+1} f(t)\,dt,\\
&\int_{\mathbb{R}} dy\, p(y) f^{(k)}(x+y) = f^{(k-1)}(x+1) - f^{(k-1)}(x) \quad (k \geqq 1)
\end{aligned}
$$

となり, $P_n(x)=B_n(x)$ (Bernoulli 多項式) であったので, 定理 2.2 より, 次の公式が得られる:

$$
f(x+h) = \int_x^{x+1} f(t)\,dt + \sum_{k=1}^{n-1} \frac{B_k(h)}{k!}(f^{(k-1)}(x+1) - f^{(k-1)}(x)) + R_n.
$$

この公式を, 整数 $a<b$ に対して, $x=a,a+1,\dots,b-2,b-1$ について足し上げると,

$$
\sum_{i=a}^{b-1} f(i+h) = \int_a^b f(t)\,dt + \sum_{k=1}^{n-1} \frac{B_k(h)}{k!}(f^{(k-1)}(b) - f^{(k-1)}(a)) + R_{a,b,n}.
$$

ここで $R_{a,b,n}$ は剰余項の和を表わす.

特に $h=0$ のとき, Bernoulli 数 $B_n=B_n(0)$ を使って上の公式を書くと,

$$
\sum_{i=a}^{b-1} f(i) = \int_a^b f(t)\,dt + \sum_{k=1}^{n-1} \frac{B_k}{k!}(f^{(k-1)}(b) - f^{(k-1)}(a)) + R_{a,b,n}.
$$

さらに Bernoulli 数について $B_1=-1/2$, $0=B_3=B_5=B_7=\cdots$ が成立していることを用い, $(n-1)/2$ 以下の最大の整数を m と書き, 両辺に $f(b)$ を加えると,

$$
\sum_{i=a}^{b} f(i) = \int_a^b f(t)\,dt + \frac{f(a)+f(b)}{2} + \sum_{j=1}^{m} \frac{B_{2j}}{(2j)!}(f^{(2j-1)}(b) - f^{(2j-1)}(a)) + R_{a,b,n}.
$$

この公式は **Euler-Maclaurin の公式**と呼ばれている.

これで定理 2.2 が実質的に Euler-Maclaurin の和公式を特別な場合として含んでいることが確かめられた. □

例 2.6 (Euler-Boole の和公式). $p(x) = (\delta(x) + \delta(x-1))/2$ (コイン投げ分布) の場合には,

$$\int_{\mathbb{R}} dy\, p(y) f^{(k)}(x+y) = \frac{1}{2}(f^{(k)}(x) + f^{(k)}(x+1))$$

となり, $P_n(x) = E_n(x)$ (Euler 多項式) であったので, 定理 2.2 より, 次の公式が得られる:

$$f(x+h) = \frac{1}{2}\sum_{k=0}^{n-1} \frac{E_k(h)}{k!}(f^{(k)}(x) + f^{(k)}(x+1)) + R_n.$$

整数 $a < b$ についてこの公式を交代的に足し上げると次が得られる:

$$\sum_{i=a}^{b-1}(-1)^{i-a} f(i) = \frac{1}{2}\sum_{k=0}^{n-1} \frac{E_k(h)}{k!}(f^{(k)}(a) + (-1)^{b-a-1} f^{(k)}(b)) + R_{n,a,b}.$$

ここで $R_{a,b,n}$ は剰余項の和を表わす. 特に $h = 0$ のとき, Euler 数 $E_n = E_n(0)$ を使って書くと,

$$\sum_{i=a}^{b-1}(-1)^{i-a} f(i) = \frac{1}{2}\sum_{k=0}^{n-1} \frac{E_k}{k!}(f^{(k)}(a) + (-1)^{b-a-1} f^{(k)}(b)) + R_{a,b,n}.$$

以上の公式は **Euler-Boole の和公式**と呼ばれている. □

注意 2.7. 以上の例から, 定理 2.2 は積分剰余項型の Taylor の定理と Euler-Maclaurin の和公式と Euler-Boole の和公式の一般化になっていることがわかる [5]. □

2.3 剰余項の絶対値の上からの荒い評価

この節では, 次のように表わされる剰余項 R_n の絶対値の大きさを上から大雑把に評価することを試みよう:

$$\begin{aligned}
R_n(x,h) &= \int_{\mathbb{R}} dy\, K_n(h,y) f^{(n)}(x+y),\\
K_n(h,y) &= \int_{\mathbb{R}} dy_1 \cdots \int_{\mathbb{R}} dy_{n-1} K(h,y_1)K(y_1,y_2)\cdots K(y_{n-1},y),\\
K(y,y') &= F(y') - H(y'-y),\\
F(y') &= \int_{-\infty}^{y'} p(t)\,dt = \int_{(-\infty,y']} P(dt),\\
H(y'-y) &= \begin{cases} 0 & (y' < y),\\ 1 & (y \geqq y). \end{cases}
\end{aligned}$$

ここで $p(t)\,dt = P(dt)$ は $\mathbb{R}$ 上の確率測度を表わす. 定理 2.2 と注意 2.3 より

$$f(x+h) = \sum_{k=0}^{n-1} \frac{P_n(h)}{n!} \int_{\mathbb{R}} dy\, p(y) f^{(k)}(x+y) + R_n(x,h) \qquad (*)$$

[5] 文献 [2] は和公式に関して定理 2.2 のような統一的定理を示しておらず, ケース・バイ・ケースの計算によって類似性を指摘するに留まっている.

が成立している.

以下では, 確率測度 $p(x)\,dx = P(dx)$ の台が有界な場合のみを扱う. 確率測度 $p(x)\,dx = P(dx)$ の台は区間 $[a,b]$ に含まれていると仮定する. さらに $h \in [a,b]$ と仮定する.

$y \in [a,b]$ のとき, y' の函数としての $K(y,y')$ の台は区間 $[a,b]$ に含まれる. ゆえに以上の仮定のもとで

$$R_n(x,h) = \int_a^b dy_1 \cdots \int_a^b dy_n\, K(h,y_1)K(y_1,y_2)\cdots K(y_{n-1},y_n)f^{(n)}(x+y_n).$$

定数 A と $M_n(x)$ を以下のように定める:

$$A = \sup_{y\in[a,b]} \int_a^b dy'\,|K(y,y')|, \qquad M_n(x) = \sup_{y\in[a,b]} |f^{(n)}(x+y)|.$$

これらを使って $|R_n(h,n)|$ を上から評価しよう.

定理 2.8 (剰余項の絶対値の上からの荒い評価). 以上の設定のもとで

$$|R_n(x,h)| \leqq A^n M_n(x).$$

証明. A, M の定義より, $h, y_k \in [a,b]$ のとき

$$\begin{aligned}
&\int_a^b dy_n\,|K(y_{n-1},y_n)|\,|f^{(n)}(x+y_n)| \leqq AM_n(x),\\
&\int_a^b dy_{n-1}\int_a^b dy_n |K(y_{n-2},y_{n-1})|\,|K(y_{n-1},y_n)|\,|f^{(n)}(x+y_n)| \leqq A^2 M_n(x),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&|R_n(h,x)| \leqq \int_a^b dy_1\cdots\int_a^b dy_n\,|K(h,y_1)|\cdots|K(y_{n-1},y_n)|\,|f^{(n)}(x+y_n)| \leqq A^n M_n(x).
\end{aligned}$$

これで示すべきことが示された. □

注意 2.9. 上の定理 2.8 の剰余項の評価はかなり大雑把である. その定理の価値は, 一般的に成立している評価を, 剰余項の具体的な表示を得ることなく, 容易に導けることを示したことにある. □

A は $y = b$ の場合の $0 \leqq \int_a^b dy'\,F(y') \leqq b-a$ と $y = a$ の場合の $\int_a^b dy'\,(1-F(y')) = (b-a) - \int_a^b dy'\,F(y')$ の大きい方になる. ゆえに A は

$$\frac{b-a}{2} \leqq A \leqq b-a$$

を満たす. 特に $[a,b] = [0,1]$ の場合には

$$\frac{1}{2} \leqq A \leqq 1$$

となる. A を小さくすれば剰余項の評価不等式の右辺も小さくなる.

例 2.10. $p(x) = \chi_{[0,1]}(x)$ (区間 $[0,1]$ 上の一様分布, Euler-Maclaurin の場合) と $p(x) = (\delta(x)+\delta(x-1))/2$ (コイン投げ分布, Euler-Boole の場合) のときには, $[a,b]=[0,1]$ かつ $A=1/2$ となる. より一般に r が正の整数で

$$p(x) = \frac{1}{r+1}\sum_{i=0}^{r}\delta\left(x-\frac{i}{r}\right)$$

の場合にも, $[a,b]=[0,1]$ かつ $A=1/2$ になる. これらの場合には定理 2.8 より,

$$|R_n(x,h)| \leqq \frac{1}{2^n}M_n(x).$$

となる. ただし, これはかなり大雑把な評価である. 例えば, Euler-Maclaurin の場合には

$$R_n(x,h) \leqq \frac{4e^{2\pi}}{(2\pi)^n}M_n(x)$$

という評価が知られている ([1, Chapter 25]). □

注意 2.11. 剰余項 $R_n(x,h)$ が $n\to\infty$ で 0 に近付けば公式 $(*)$ で $n\to\infty$ とした公式が成立していることになる. しかし, そのようなことは公式 $(*)$ の典型的な応用場面では成立していない. 例えば $s>0$ でかつ $f(x)=x^{-s}$ の場合には

$$|f^{(n)}(x)| = s(s+1)\cdots(s+n-1)|x|^{-s-n} = \frac{\Gamma(s+n)}{\Gamma(s)}|x|^{-s-n}$$

となり. $n\to\infty$ で $M_n(x)$ が階乗函数のオーダーで増大してしまう. このような場合には近似の誤差を最小にするためには適切な大きさの n を選ばなければいけなくなる[6]. □

参考文献

[1] Victor Kac and Pokman Cheung. Quantum Calculus. Springer (2002), ix+112 pages.

[2] Jonathan M. Borwein, Neil J. Calkin, and Dante Manna. Euler-Boole Summation Revisited. Amer. Math Monthly, Vol. 116, Issue 5, 2009, pp. 387–412. https://scholar.google.co.jp/scholar?cluster=15258475459772769600

[3] 黒木玄. Kullback-Leibler 情報量と Sanov の定理. 私的ノート 2016, 2017, 73 pages. https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf

[4] 黒木玄. Euler-Maclaurin の公式とその数値計算への応用. Julia 言語による数値計算の Jupyter notebook, 2017-07-22. https://gist.github.com/genkuroki/9cb1ffca17caf8f0f3209384fe568efb

[6] 実際の数値計算例については [4] を参照せよ.