解答例

STEP II 2000、第 5 問

純粋数学 • 最小二乗近似

> **STEP II 2000、問 5**
>
> 先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：
>
> step.maths.org/questions/00-s2-q5

鍵となる考え方

平方を展開すれば、最小化すべき量は λ の 2 次式にすぎず、その係数は定数（f を含む積分）である。平方完成（completing the square）を行えば、最適な λ と最小の誤差が一度に得られる。正弦関数の例では、n が大きいとき $\sin(\pi x/n)$ が原点を通るほぼ直線であることが要点であり、小さい角に対する近似（small-angle approximations）によってこれを厳密に扱うことができる。

解答

λ の最適な値

次のようにおく。

$$J = \int_0^1 x f(x)\,\mathrm{d}x, \qquad K = \int_0^1 (f(x))^2\,\mathrm{d}x.$$

これらは λ によらない定数である。平方を展開し、$\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$ を用いると、

$$\begin{aligned} I(\lambda) = \int_0^1 (f(x) - \lambda x)^2\,\mathrm{d}x &= K - 2\lambda J + \tfrac{1}{3}\lambda^2 \\ &= \tfrac{1}{3}(\lambda - 3J)^2 + K - 3J^2. \end{aligned}$$

2 乗の項は決して負にならず、$\lambda = 3J$ のときに限り 0 となる。したがって $I(\lambda)$ が最小となるのはちょうど

$$\lambda = 3\int_0^1 x f(x)\,\mathrm{d}x,$$

のときであり、これが示すべきことであった。（同値な方法として、$I'(\lambda) = -2J + \frac{2}{3}\lambda = 0$ から $\lambda = 3J$ が得られ、$I''(\lambda) = \frac{2}{3} > 0$ により最小値であることが確かめられる。）

残差

R^2 は I の最小値であり、平方完成の式からそれは $K - 3J^2$ である。$J = \lambda/3$ より $3J^2 = \frac{1}{3}\lambda^2$ となるので、

$$R^2 = \int_0^1 (f(x))^2 \,\mathrm{d}x - \tfrac{1}{3}\lambda^2,$$

が示された。（微分によって λ を求めた場合は、$K - 2\lambda J + \frac{1}{3}\lambda^2$ に $J = \lambda/3$ を代入すれば $K - \frac{2}{3}\lambda^2 + \frac{1}{3}\lambda^2$ となり、同じ結果を得る。）

$f(x) = \sin(\pi x/n)$ の場合

計算を見やすくするため $a = \pi/n$ とおく。すると $f(x) = \sin ax$ であり、$n \to \infty$ のとき $a \to 0$ である。$u = x$、$\mathrm{d}v = \sin ax\,\mathrm{d}x$ として部分積分（integration by parts）を行うと、

$$\begin{aligned}\int_0^1 x \sin ax \,\mathrm{d}x &= \Big[-\frac{x\cos ax}{a}\Big]_0^1 + \frac{1}{a}\int_0^1 \cos ax \,\mathrm{d}x \\ &= -\frac{\cos a}{a} + \frac{\sin a}{a^2},\end{aligned}$$

よって

$$\lambda = \frac{3(\sin a - a\cos a)}{a^2}.$$

考えてみよう $\sin\theta$ の近似として単に $\sin\theta \approx \theta$ とするのではなく、θ^3 の項が必要なのはなぜだろうか。

$\sin a \approx a$ と $\cos a \approx 1$ だけを用いると、分子は $a - a = 0$ となり、何もわからない。主要項どうしが打ち消し合い、答えは次の次数に現れるのである。ヒントにある二つの近似をともに用いると、

$$\sin a - a\cos a \approx \Big(a - \tfrac{1}{6}a^3\Big) - a\Big(1 - \tfrac{1}{2}a^2\Big) = \tfrac{1}{3}a^3,$$

したがって、n が大きいとき

$$\lambda \approx \frac{3\cdot\frac{1}{3}a^3}{a^2} = a = \frac{\pi}{n}$$

となる。これが結果 (i) である。

(ii) には K も必要である。$\sin^2 ax = \frac{1}{2}(1 - \cos 2ax)$ を用いると、

$$K = \int_0^1 \sin^2 ax \,\mathrm{d}x = \frac{1}{2} - \frac{\sin 2a}{4a}.$$

$n \to \infty$ のとき $a \to 0$ かつ $\dfrac{\sin 2a}{2a} \to 1$ であるから、$K \to \frac{1}{2} - \frac{1}{2} = 0$ である。また $\lambda \approx \pi/n$ より $\lambda \to 0$ である。ゆえに

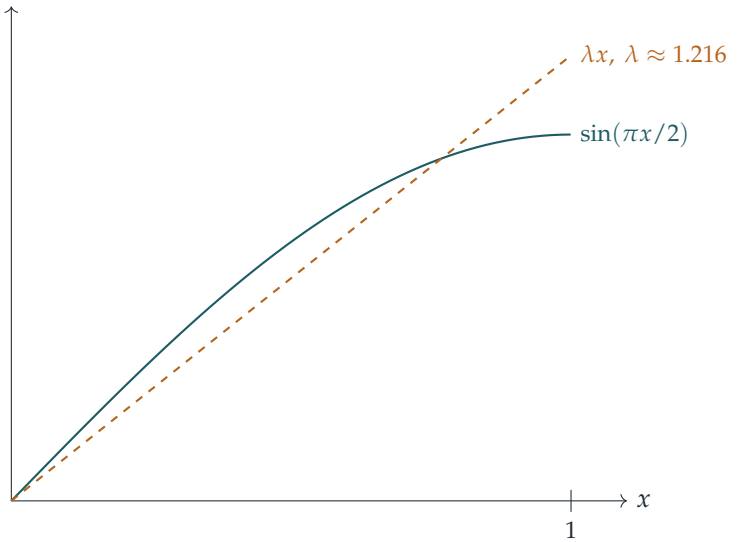
$$R^2 = K - \tfrac{1}{3}\lambda^2 \to 0 - 0 = 0,$$

であり、$R \geq 0$ であるから $\lim_{n\to\infty} R = 0$ を得る。これが結果 (ii) である。

実際には、打ち消し合いはこれよりはるかに強い。$\sin 2a \approx 2a - \frac{4}{3}a^3$ を用いると $K \approx \frac{1}{2} - (\frac{1}{2} - \frac{1}{3}a^2) = \frac{1}{3}a^2$ となり、$\frac{1}{3}\lambda^2 \approx \frac{1}{3}a^2$ でもある。したがって R^2 は単に小さいだけではなく、f 自身の大きさである K と比べても小さい。直線が f のほぼすべてをとらえているのである。

これらの結果が予想される理由

$0 \leq x \leq 1$ において角 $\pi x/n$ は 0 と π/n の間にあり、n が大きければこれは小さい。よって区間全体で $\sin(\pi x/n) \approx \pi x/n$ であり、この関数自体が $\lambda = \pi/n$ とした 1 次関数 λx にきわめて近い。したがって最もよく当てはまる直線の傾きは π/n に近いはずであり、当てはまりは本質的に完全になっていくので、残差は 0 に収束（converge）するはずである。下図から、$n = 2$ の場合でさえ曲線が直線からそれほど離れていないことがわかる。



検算

- はじめから $f(x) = cx$ であれば、$\lambda = 3c\int_0^1 x^2\,\mathrm{d}x = c$ かつ $R^2 = \frac{1}{3}c^2 - \frac{1}{3}c^2 = 0$ となり、当然そうあるべき結果である。
- 数値的には、$n = 10$ のとき厳密な式から $\lambda = 0.31107$ が得られ、これに対して $\pi/10 = 0.31416$ である。$n = 100$ のときは比 $\lambda/(\pi/n)$ が 0.99990 となる。
- $n = 10$ のとき、数値積分（quadrature）により $R^2 \approx 6.0 \times 10^{-7}$ であり、一方 $K \approx 0.032$ であるから、残差は f と比べて確かにきわめて小さい。

よくある間違い

- $\sin a \approx a$、$\cos a \approx 1$ で止めてしまい、$\lambda \approx 0$ を得ること。正しいが役に立たない。3 次の項を残すこと。
- $\lambda = 3\int_0^1 xf(x)\,\mathrm{d}x$ の 3 が $\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$ に由来することを忘れること。

STEP のその先へ

これは最小二乗近似（least-squares approximation）であり、その本質は幾何学である。$[0,1]$ 上の関数を、「内積」（dot product）$\langle f, g\rangle = \int_0^1 fg\,\mathrm{d}x$ をもつベクトルとみなす。すると λx は x が張る直線への f の正射影（orthogonal projection）であり、$R^2 = K - \frac{1}{3}\lambda^2$ はピタゴラス（Pythagoras）の定理にほかならない。展開をもう一段進めると $R^2 \approx a^6/1575$ が示され、これは x^3 の項から x 方向の成分を取り除く射影に由来する。Strang の *Introduction to Linear Algebra*（線形代数入門）では射影と最小二乗法が展開されており、同じ考え方はフーリエ級数（Fourier series）へとつながっていく。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。