物理学情報処理演習

3. 表計算ソフトで数値計算

数値積分

データ処理 最小二乗法

参考文献

Numerical Recipes in C

W. H. Press他著　　技術評論社

Advanced Engineering Mathematics

Erwin Keryszig　　John Wiley & Sons, Inc.

大久保晋

E-mail: buturi-johoshori@tiger.kobe-u.ac.jp

http://extreme.phys.sci.kobe-u.ac.jp/staffs/okubo/lectures/Programming/index.html

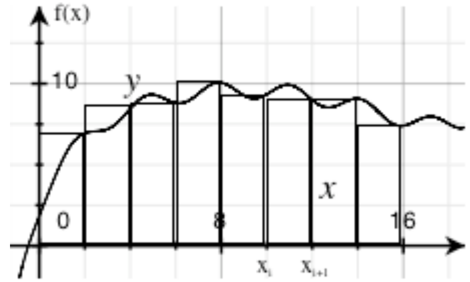
数値積分

- 計算機が行える演算は、基本的に数値計算のみ
- 具体的な値を入れて計算する
- 数値解は解析的な解ではなく、数値を入れて最も近い値を探す方法

- 解析的な解がなくても近似解は得られる。特に、一般的な式に対する積分の一般解はない。
- 適切な評価を行えば、適切な解が得られる反面評価や計算方法によっては全く異なる解を与えることに注意する。

数値積分

関数 f(x) の区間 [a, b] の積分をすることを考えよう。



計算機は、積分公式を知らないので（もちろん知っている人間が与えるという手はあります）、数学の定義に立ち返って数値的に積分を行う。

$$I = \int_a^b f(x)dx = \sum_{i=1}^{N} f(x_i)\cdot\Delta x = \Delta x\left[f(x_0) + f(x_1) + \cdots + f(x_N)\right]$$

この和を数値的に行えばよい。

便宜的に、積分区間[a, b]を等間隔でN等分し、その幅をh=(b-a)/Nとすると、各点は

x_0=a, x_1=x_0+h, …, x_N=b

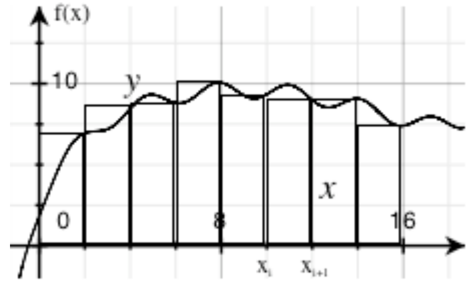
と表される。

これらを関数f(x) に代入し、足し合わせればよい。

数値積分：矩形近似

数値積分を行う際、関数f(x)をN等分し、関数をの積分を矩形で近似する方法。右図参照

関数の変化に比べ、分割する数が多ければ近似としてて成り立つが、分割が少なければ悪い近似となる。



$$I = \int_a^b f(x)dx = \sum_{i=1}^{N} f(x_i)\cdot\Delta x = h\left[f(x_0) + f(x_0 + \Delta x) + \cdots + f(x_N)\right]$$

この和を数値的に行えばよい。

数値積分：台形近似

数値積分を行う際、関数f(x)をN等分し、関数をの積分を台形で近似する方法。

分割区間の両端で直線近似するので、矩形近似よりよい近似になっているが、関数の変化に比べ分割点が十分にあることが必要

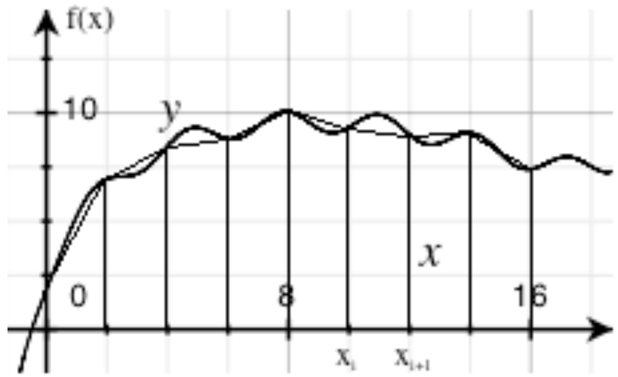
$$I = \int_a^b f(x)dx = \sum_{i=1}^{N}\left[\frac{1}{2}f(x_i) + \frac{1}{2}f(x_{i+1})\right]\cdot \Delta x = h\left[\frac{1}{2}f(x_0) + \sum_{i=1}^{N-1} f(x_i) + \frac{1}{2}f(x_N)\right]$$

この和を数値的に行えばよい。この式を台形公式と呼ぶ。

台形近似は、各区間でf(x)を直線近似したことになる。各区間でf(x)を２次曲線で近似すると積分は

$$I = \int_a^b f(x)dx = \sum_{i=1}^{N}\left[\frac{1}{3}f(x_i) + \frac{4}{3}f(x_{i+1}) + \frac{1}{3}f(x_{i+2})\right]\cdot \Delta x$$

となる。これをSimpson則という。



演習３－１ ：矩形近似

矩形近似を用いて積分公式

$$\int \cos x dx = \sin x$$

を確かめてみよう。

0.01 刻みで 0.0から 9.42 までのcos x を計算し、幅0.01 でかけて、その和をとる

$$I = \int_0^{9.42} \cos x dx = \sum_{i=1}^{N} \cos(x_i) \cdot 0.01$$

演習：刻み幅を0.1 と0.01 と変えたときに積分した結果がどの程度変わるか確認しよう。誤差は累積的に大きくなるので、0.0から9.42 までの各0.01づつの積分値がどの程度厳密解 sin x とずれているか、矩形近似の計算と厳密解のグラフを並べて比較しよう。（差をとって比較するのも一つの手）

A: OpenOffice を使って、実際にxの列を作り、cos x を計算させて、幅0.1をかけて、足し合わせる。その結果が sin (9.42)と比較し、どの程度誤差が生じているか確認する。

B: 同様に幅0.01としたときには誤差が縮まっているか確認する。

演習３－２：台形近似

矩形近似を用いて積分公式

$$\int x\sin x dx = \sin x - x\cos x$$

を確認すると共に矩形近似でも計算し、矩形近似のときと台形近似の際の違いを見てみよう。近似精度の確認のため刻み値は大きめにとる。

0.1 刻みで 0.0から 9.42 までのsin x - x cos x を台形公式で計算する。

$$I = \int_0^{9.42} \cos x dx = \frac{1}{2}\cos(a) + \sum_{i=2}^{N-1} \cos(x_i)\cdot 0.1 + \frac{1}{2}\cos(b)$$

演習：演習３－１と同様に累積的に誤差が広がるのをグラフを使って比較する。

A: OpenOffice を使って、実際にxの列を作り、x cos x を計算させて、幅0.1をかけて、足し合わせる。その結果が sin (9.42) – 9.42 cos (9.42)となっていることを確認しよう。

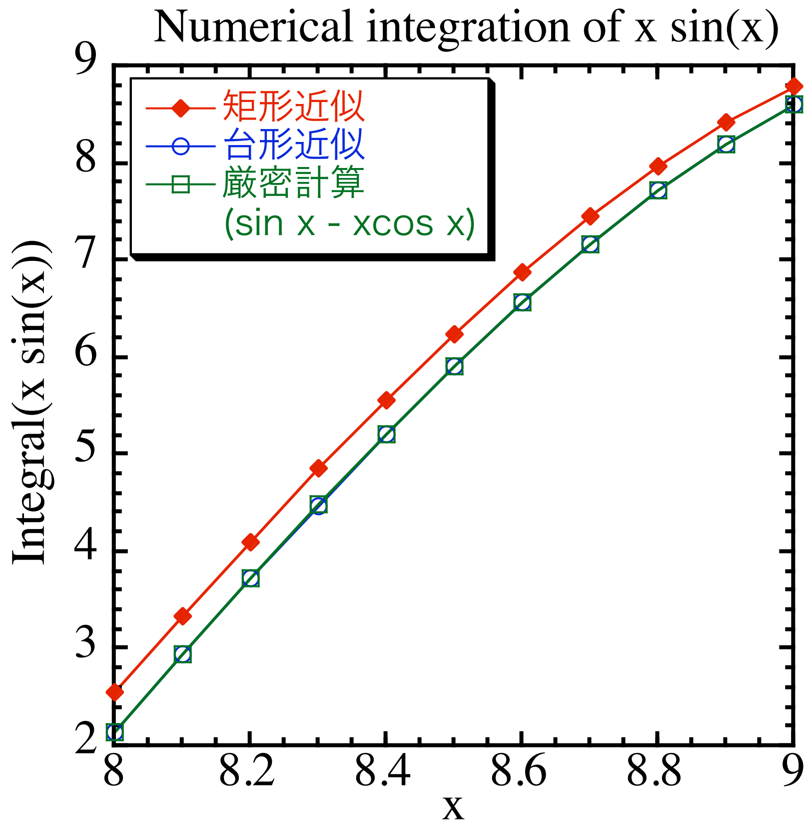
B: 同様のことを矩形近似で行ったときと比べてどの程度誤差が小さくなっているか比較しよう。

演習３-２：台形近似

計算して比較した結果は右図の通り。
区間[8, 9]を拡大した。

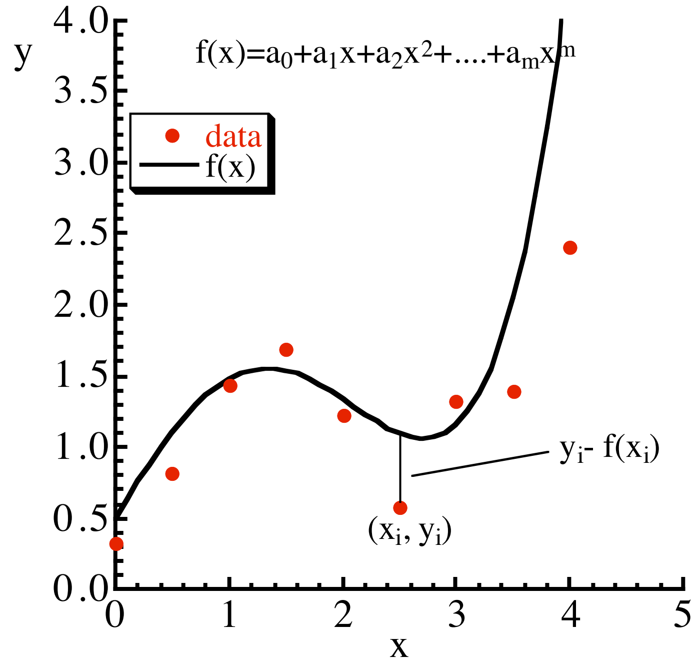
台形近似のプロット点が厳密計算の真上に乗っていることが確認できる。一方、矩形近似は厳密計算のやや上方にプロット点がきている。

これからも、台形近似は計算量が少ない割によい近似を与えることが判る。



台形近似の誤差は h^3 のオーダーである。Simpson 近似の誤差は h^5 のオーダーであるが、多数の区間分割で数値計算をする必要があるので、丸め誤差が支配的になり、必ずしも h^5 のオーダーにはならない。そのため実用的には台形近似で十分である。

最小二乗法



データ点が与えられているとき、データ点を最も再現する理論式f(x)を求める方法。

理論式$f(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m$とデータ点$(x_i, y_i)$との距離は

$$\left(y_i - f(x_i)\right)^2$$

である。したがって各点での総和は

$$I(a_0,a_1,\ldots,a_m) = \sum_i \left(y_i - f(x_i)\right)^2 = \sum_i \left\{ \left(y_i - \left(a_0 + a_1x_i + a_2x_i^{\ 2} + \cdots + a_ix^m\right)\right)^2 \right\}$$

この$I(a_0,a_1,\ldots,a_m)$ を最小にする係数の組$(a_0, a_1, \ldots, a_m)$を求めることは、Iを$(a_0, a_1, \ldots, a_m)$の関数として見て、この関数の最小値を求める問題である。関数の最小値は変数の微分がゼロになる点であるから

$$\frac{\partial I(a_0,a_1,\ldots,a_m)}{\partial a_i} = 0$$

を満足すればよい。したがって、係数の組を変数としたm個の連立方程式の解を求める問題に帰着する。

最小二乗法

m個のデータ点(x_i, y_i)を一次式の理論式$f(x)=a_0+a_1x$にフィットさせることを考えよう。

理論式$f(x)=a_0+a_1x$とデータ点(x_i, y_i)との距離の総和は

$$I(a_0,a_1) = \sum_i \left(y_i - f(x_i)\right)^2 = \sum_i \left\{\left(y_i - \left(a_0 + a_1 x_i\right)\right)^2\right\}$$

$$= \sum_{i=1}^{m} \left\{\left(y_i - a_0\right)^2 - 2a_1\left(y_i - a_0\right)x_i + a_1^2 x_i^2\right\}$$

この$I(a_0,a_1)$ は係数(a_0, a_1)の２次式である。このIの最小値は２次式の極値であるから微分がゼロであることから

$$\frac{\partial I(a_0,a_1)}{\partial a_0} = 0 \qquad \frac{\partial I(a_0,a_1)}{\partial a_1} = 0$$

を満たすa_0, a_1を求めればよい。

最小二乗法

$$\frac{\partial I(a_0, a_1)}{\partial a_0} = ma_0 + \left(\sum_i x_i\right)a_1 - \sum_i y_i = 0$$

$$\frac{\partial I(a_0, a_1)}{\partial a_1} = a_0\sum_i x_i + a_1\sum_i x_i^2 - \sum_i x_i y_i = 0$$

であるから、理論式f(x)=a_0+a_1x のa_0, a_1は

$$a_0 = \frac{\sum_i y_i \sum_i x_i^2 - \sum_i x_i \sum_i x_i y_i}{m\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \qquad a_1 = \frac{m\sum_i x_i y_i - \sum_i x_i \sum_i y_i}{m\sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

となる。

最小二乗法

より一般的なn次多項式 $f(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m$ を用いても、n次指数関数の多項式 $f(x)= a_0+a_1e+a_2e^2+\ldots+a_me^m$ を用いても二乗差

$$\left(y_i - f(x_i)\right)^2$$

は係数の組$(a_0, a_1, \ldots, a_m)$の２次関数である。したがって、

$$\frac{\partial I(a_0,a_1,\ldots,a_m)}{\partial a_i} = 0$$

は１次方程式である。a_iが0～mまであるので、m+1の連立１次方程式を解くことになる。

演習３－３ ：最小二乗法（1次）

OpenOfficeを使って右の11個のデータ点を最小二乗法で直線に近似し、その係数a0, a1を求めよ。

さらに、近似した直線との誤差の平均、誤差の標準偏差を求めよ。右のプロットと得られた直線をプロットして合っているか確認しよう。

a0, a1は

$$a_0 = \frac{\sum_i y_i \sum_i x_i^2 - \sum_i x_i \sum_i x_i y_i}{m\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \qquad a_1 = \frac{m\sum_i x_i y_i - \sum_i x_i \sum_i y_i}{m\sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

であり誤差平均、標準偏差は次の式で与えられる。

$$\overline{error} = \frac{1}{n}\sum_{i=1}^{n} f(x_i) - y_i \qquad \sigma_{error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - y_i\right)^2}$$

ここで、f(x_i)は求めたfitting 関数f(x)=a0+a1*xである。

x	y
0.0000	-0.27832
1.0000	3.3703
2.0000	3.7018
3.0000	4.7366
4.0000	7.6784
5.0000	12.226
6.0000	13.992
7.0000	18.176
8.0000	22.618
9.0000	25.301
10.000	28.685

演習３－３：最小二乗法（1次）

直線近似した結果は右図のようになる

解は以下の通り

a0= -1.9105, a1= 2.9314,
error=0.00018666, σ_{error}= 0.44734