物理学情報処理演習

2015年4月24日

3. 表計算ソフトで数値計算

ver20150424

数値積分

データ処理 最小二乗法

身内賢太朗

レポート提出：fsci-phys-jouhou@edu.kobe-u.ac.jp

本日の作業ファイル名：2015_jouhou_03_学籍番号の下4桁.xlsx

数値計算

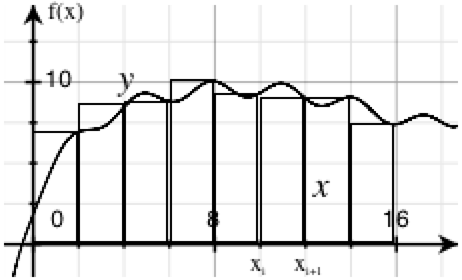
- 計算機が行う演算：数値計算のみ
  - 数値計算：具体的な値を入れて計算する⇔解析解による計算

- 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る)、数学の定義に立ち返って数値的に積分を行う。

$$I=\int_a^b f(x)dx=\sum_{i=1}^{N} f(x_i)\cdot\Delta x=\left[f(x_1)+f(x_2)+\cdots+f(x_N)\right]\cdot\Delta x$$

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$\Delta x=(b-a)/N$とすると、各x_iは、

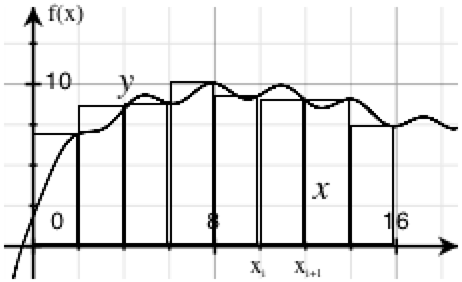
$x_0=a$, $x_1=x_0+\Delta x$, $x_2=x_1+\Delta x$, ..., $x_N=b$

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①：矩形近似

数値積分を行う際、関数$f(x)$をN等分し、関数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。

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

この和を数値的に行えばよい。

演習３－１ ：矩形近似

ファイル名：2015_jouhou_03_学籍番号の下4桁.xlsx
シート名：3-1解答
参考シート：作業例3-1

課題3-1：矩形近似を用いて

$$I = \int_0^{9.4} \cos x dx = \sum_{i=1}^{N} \cos(x_i) \cdot \Delta x = \sin 9.4$$

を計算することで、積分公式

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

確認せよ。Δxを0.05及び0.01としたものを求め、真の値sin9.4と比較すること。幅を0.1、0.05、0.01と変化させた場合にどうなるか考察、シート内A1のセルにに記述せよ。

0.0から 9.4 までのcos x を計算し、幅Δx をかけて、その和をとる

作業例3-1に実際にxの列を作りcos x を計算し、幅0.1をかけて足し合わせることによってsin(9.4)を近似的に求めた例を示す（列B-E）。同様にして、F～IおよびJ～Mで作業を行う。

数値積分②：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（右下図参照）。

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

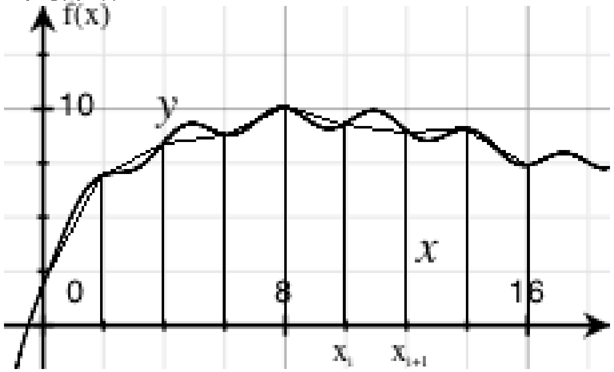
$$I=\int_a^b f(x)dx=\sum_{i=0}^{N-1}\left[\frac{1}{2}f(x_i)+\frac{1}{2}f(x_{i+1})\right]\cdot\Delta x=\left[\frac{1}{2}f(x_0)+\sum_{i=1}^{N-1}f(x_i)+\frac{1}{2}f(x_N)\right]\cdot\Delta x$$

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

$$I=\int_a^b f(x)dx=\sum_{i=0}^{N-2}\left[\frac{1}{3}f(x_i)+\frac{4}{3}f(x_{i+1})+\frac{1}{3}f(x_{i+2})\right]\cdot\Delta x$$

となる。これを**Simpson**則という。



シート名：演習3-2
参考シート：作業例3-2

演習３－２：台形近似

台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、両者を比較する。近似精度の確認のため、刻み値は大きめにとる。

課題3-2：0.1 刻みで 0.0から b までの $x\sin x$ を計算せよ。

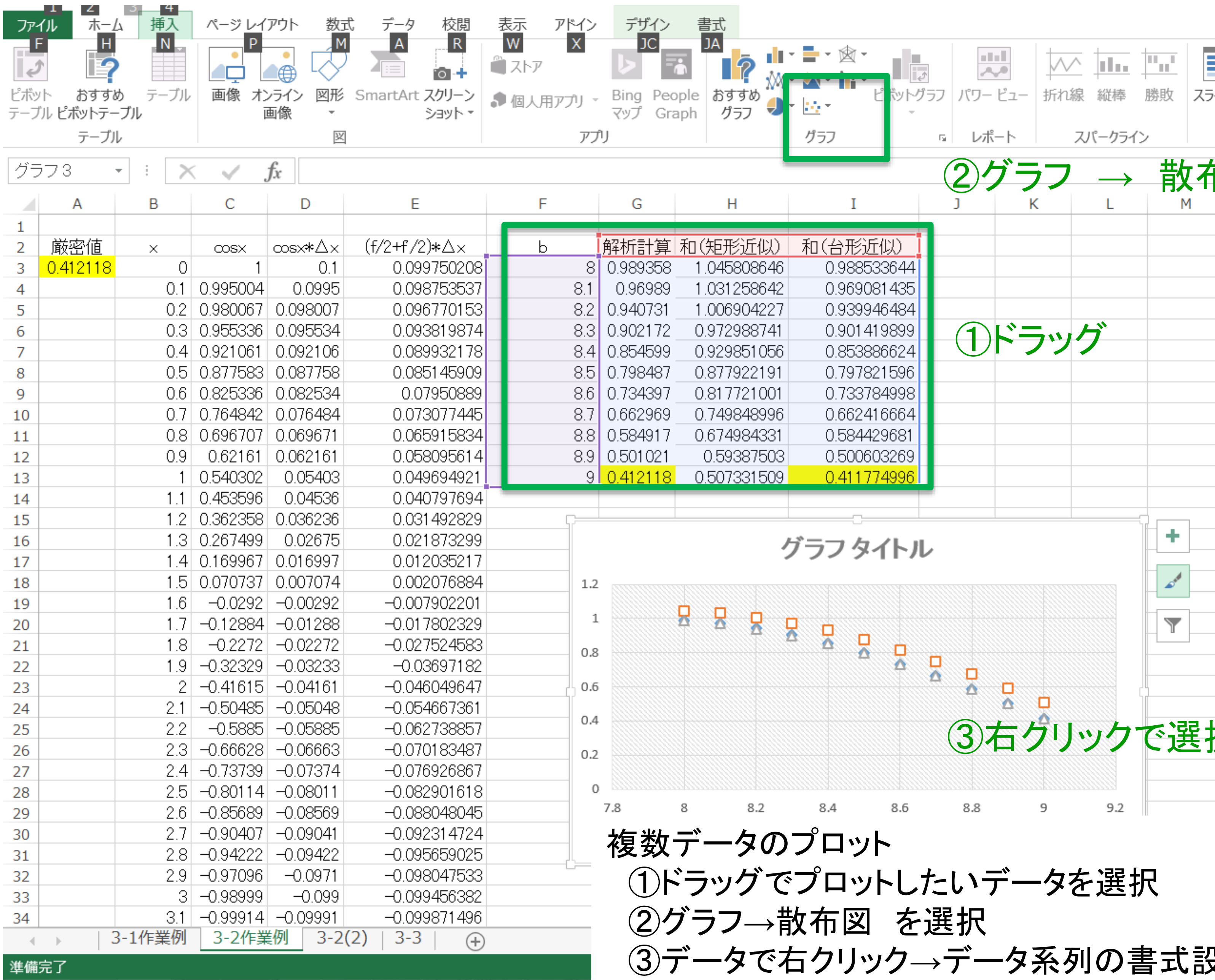
$$I(b) = \int_0^b x\sin x dx = \left[\sum_{i=1}^{N-1} x_i \sin(x_i) + \frac{1}{2}\cdot b\sin(b) \right]\cdot 0.1$$

Excel を使って、実際にxの列を作り、b=8.0, 8.1, ..., 9.0の場合について、矩形近似、台形近似を使って積分値を計算し、$\sin b - b\cos b$ と比較せよ。

結果をグラフ（横軸にb、縦軸に矩形近似値、台形近似値、厳密計算値をプロットしたものとして図示せよ。

作業例3-2には $I(b) = \int_0^b \cos x dx$ についての例を示す。

複数データのプロット



	A	B	C	D	E	F	G	H	I
1									
2	厳密値	x	cosx	cosx*Δx	(f/2+f/2)*Δx	b	解析計算	和(矩形近似)	和(台形近似)
3	0.412118	0	1	0.1	0.099750208	8	0.989358	1.045808646	0.988533644
4		0.1	0.995004	0.0995	0.098753537	8.1	0.96989	1.031258642	0.969081435
5		0.2	0.980067	0.098007	0.096770153	8.2	0.940731	1.006904227	0.939946484
6		0.3	0.955336	0.095534	0.093819874	8.3	0.902172	0.972988741	0.901419899
7		0.4	0.921061	0.092106	0.089932178	8.4	0.854599	0.929851056	0.853886624
8		0.5	0.877583	0.087758	0.085145909	8.5	0.798487	0.877922191	0.797821596
9		0.6	0.825336	0.082534	0.07950889	8.6	0.734397	0.817721001	0.733784998
10		0.7	0.764842	0.076484	0.073077445	8.7	0.662969	0.749848996	0.662416664
11		0.8	0.696707	0.069671	0.065915834	8.8	0.584917	0.674984331	0.584429681
12		0.9	0.62161	0.062161	0.058095614	8.9	0.501021	0.59387503	0.500603269
13		1	0.540302	0.05403	0.049694921	9	0.412118	0.507331509	0.411774996
14		1.1	0.453596	0.04536	0.040797694				
15		1.2	0.362358	0.036236	0.031492829				
16		1.3	0.267499	0.02675	0.021873299				
17		1.4	0.169967	0.016997	0.012035217				
18		1.5	0.070737	0.007074	0.002076884				
19		1.6	−0.0292	−0.00292	−0.007902201				
20		1.7	−0.12884	−0.01288	−0.017802329				
21		1.8	−0.2272	−0.02272	−0.027524583				
22		1.9	−0.32329	−0.03233	−0.03697182				
23		2	−0.41615	−0.04161	−0.046049647				
24		2.1	−0.50485	−0.05048	−0.054667361				
25		2.2	−0.5885	−0.05885	−0.062738857				
26		2.3	−0.66628	−0.06663	−0.070183487				
27		2.4	−0.73739	−0.07374	−0.076926867				
28		2.5	−0.80114	−0.08011	−0.082901618				
29		2.6	−0.85689	−0.08569	−0.088048045				
30		2.7	−0.90407	−0.09041	−0.092314724				
31		2.8	−0.94222	−0.09422	−0.095659025				
32		2.9	−0.97096	−0.0971	−0.098047533				
33		3	−0.98999	−0.099	−0.099456382				
34		3.1	−0.99914	−0.09991	−0.099871496				

②グラフ → 散布図

①ドラッグ

③右クリックで選択

複数データのプロット

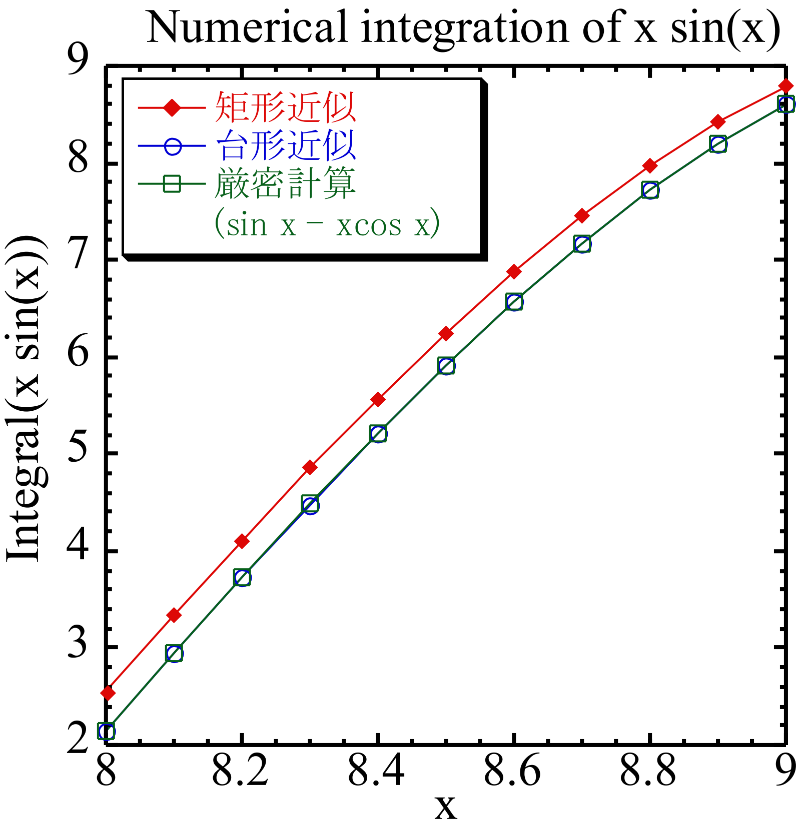
①ドラッグでプロットしたいデータを選択

②グラフ→散布図　を選択

③データで右クリック→データ系列の書式設定　で複数データを見えるように。

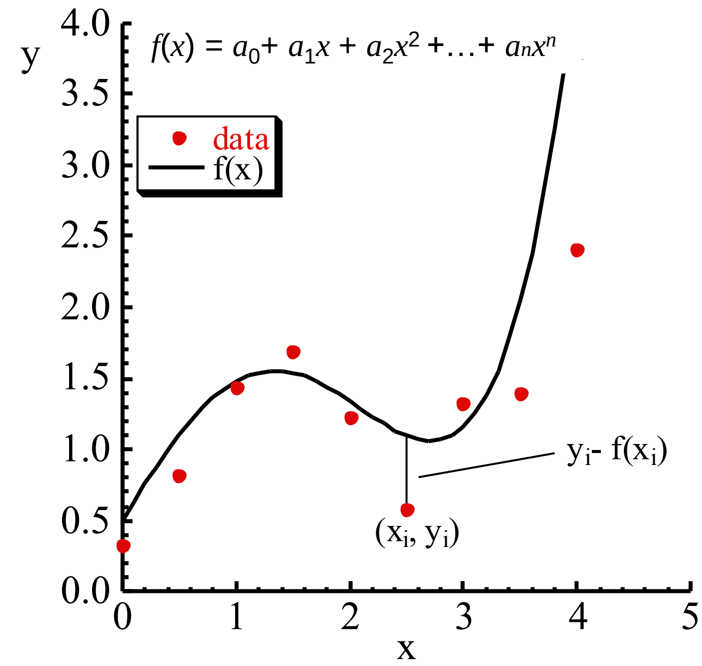
演習３－２ ：台形近似（続き)

右図のように区間[8, 9]においてグラフ化し比較すると、台形近似による計算値と厳密計算値はほぼ等しく、矩形近似による計算値は厳密計算値よりやや大きい。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わかる。



台形近似の誤差は $(\Delta x)^3$ のオーダーである。Simpson 近似の誤差は $(\Delta x)^5$ 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 $(\Delta x)^5$ 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

$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$

とデータ点(x_i, y_i)との距離の2乗は

$$(y_i - f(x_i))^2$$

である。したがって各点での総和は

$$I(a_0, a_1, \ldots, a_n) = \sum_i (y_i - f(x_i))^2 = \sum_i \left\{ \left(y_i - \left(a_0 + a_1x_i + a_2{x_i}^2 + \cdots + a_n{x_i}^n \right) \right)^2 \right\}$$

この$I\ (a_0, a_1, \ldots, a_n)$を最小にする係数の組$(a_0, a_1, \ldots, a_n)$を求めることは、$I$を$(a_0, a_1, \ldots, a_n)$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

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

を満足すればよい。したがって、係数の組を変数としたn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 $f(x) = a_0 + a_1 x$ にフィットさせことを考える。

理論式 $f(x) = a_0 + a_1 x$とデータ点(x_i, y_i)との距離の2乗の総和は

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

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

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。したがって、

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

を満たすa_0, a_1を求めればよい。

最小二乗法（続き）

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

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

であるから、理論式 $f(x) = a_0 + a_1 x$ 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。

* ここでは結果を与えているが、一度は自分で導いてみるのが望ましい。

最小二乗法

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

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

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

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

は1次方程式である。a_0からa_nまであるので、(n+1)元の連立1次方程式を解くことになる。

演習３−３ ：最小二乗法（1次）

シート名：演習3-3
参考シート：作業例3-3

課題3-3：Excelを使って右の11個のデータ点を最小二乗法で直線にフィットし、その係数a_0, a_1を求めよ。さらに、フィットした直線との差の平均、標準偏差を求めよ。データと得られた直線が合っているか、グラフ化することによって確認せよ。

a_0, a_1 は,

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

で差の平均、標準偏差は次の式で与えられる。

$$< residual >= \frac{1}{n} \sum_{i=1}^{n} (f(x_i) - y_i) \qquad \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(f(x_i) - y_i\right)^2}$$

ここで、$f(x_i)$は求めたfitting 関数 $f(x) = a_0 + a_1 x$ 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

課題提出

- 宛先　fsci-phys-jouhou@edu.kobe-u.ac.jp
- 件名　2015-report03_**学籍番号**の下4桁
- 本文　学籍番号と名前
- 添付ファイル：2015_jouhou_03_学籍番号の下4桁.xlsx
  - 3-1解答　3-2解答　3-3解答
    の3枚のシートがあることを確認。
- 締め切り　2014年5月1日（金）13：00