数値解析

第8回 関数近似と補間1 ～最小2乗近似～

目的

- よく見る「最小二乗法」による近似ができるようになる

課題

- input.csvにより1行1データの形式で与えられる，4点のデータを2次多項式($ax^2 + bx + c$)で近似し，関数を示せ．可能であれば，データと近似関数をグラフで示せ．

(x, y):　$(-3.00, -26.3)$　$(1.00, 2.05)$　$(1.50, 0.99)$　$(4.70, -29.0)$

関数近似

1. エクセルを開く
2. 課題の数値(4つ)を入れる
3. 散布図を描く
4. 近似曲線を追加する

多項式近似: 次数2

関数近似

- データや複雑な関数を簡単な関数でおおよそあらわすこと.

- 連続型データ(関数)

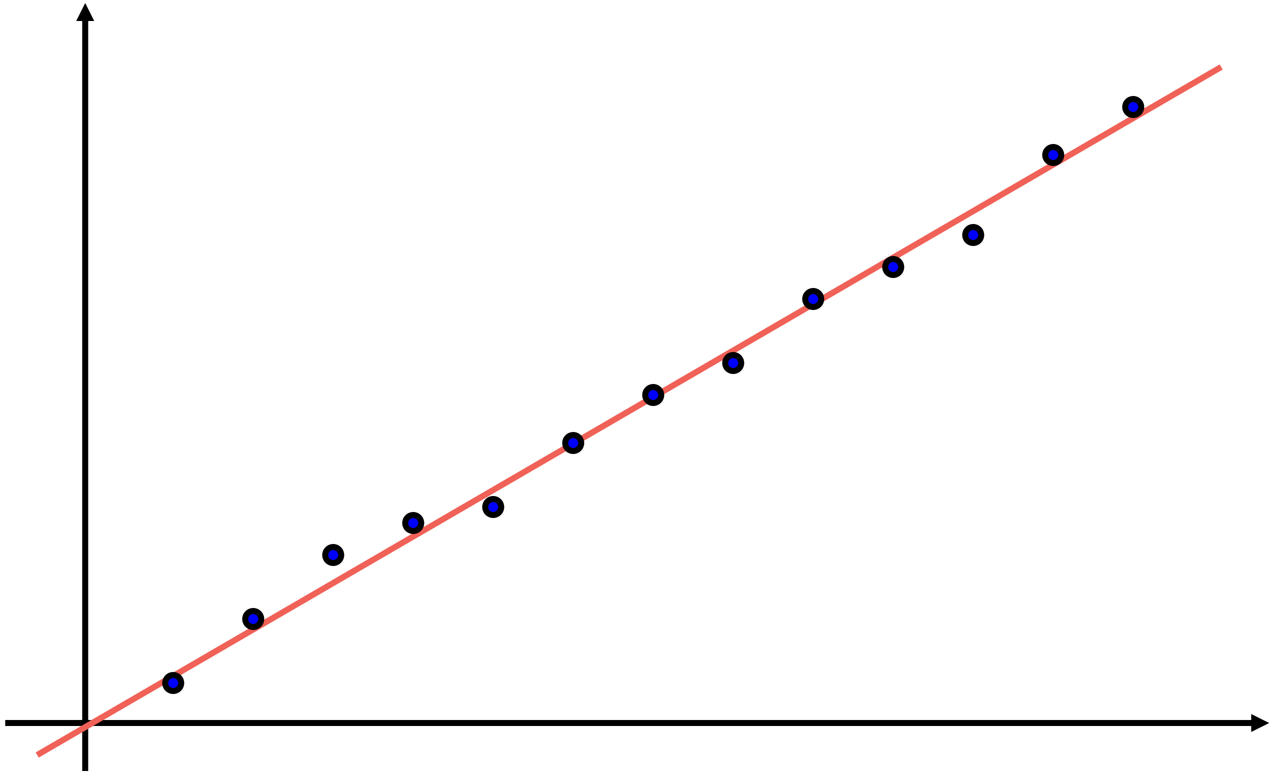
  ex) フーリエ変換: (周期)関数を振動関数の和で表現
  → 後期の実験(実際には離散フーリエ変換)

- 離散型データ(実験データなど)

  数値解析で扱うのはこちら.

関数近似(フィッティング)

- 観測データ群に対して，それらのデータの示す傾向を大まかに知りたい．



最小2乗近似

- 観測データ群と関数によって計算される値との差の2-normを最小化するように関数の重みを求める方法.

m個の観測データ (x_1, y_1), (x_2, y_2), …, (x_m, y_m)

$$\boxed{E = \sum_{k=1}^{m} \left(\boxed{y_k} - \sum_{j=1}^{n} \boxed{a_j} \boxed{\varphi_j} (\boxed{x_k}) \right)^2}$$

n個の1次独立(線形独立)関数φ_j

n個の重み(求めるもの)

数学の復習 (1)

- 線形独立 (1次独立)

$\sum a_j \varphi_j = 0$ となるのが, $a_j = 0$のときのみ

ex. $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$は線形独立だが, $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}$は線形従属

関数の独立も同様の考え

多項式 $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ は線形独立

$a_1 x + a_2 \ \cdot 2x$ は線形従属

フーリエ級数は線形独立

数学の復習 (2)

- 関数が最小値になる条件: 微分が0

$$E=\sum_{k=1}^{m}\left(y_k-\sum_{j=1}^{n}a_j\varphi_j(x_k)\right)^2$$

$$=\sum_{k=1}^{m}\left(y_k^2-2y_k\sum_{j=1}^{n}a_j\varphi_j(x_k)+\left(\sum_{j=1}^{n}a_j\varphi_j(x_k)\right)^2\right)$$

これをaに関して微分(偏微分)

偏微分 (数学の復習続き)

$$E = \sum_{k=1}^{m}\left(y_k^2 - 2y_k \sum_{j=1}^{n} a_j \varphi_j(x_k) + \left(\sum_{j=1}^{n} a_j \varphi_j(x_k) \right)^2 \right)$$

合成関数の微分法

これをa_iに関して微分

$$\frac{\partial E}{\partial a_i} = \sum_{k=1}^{m}\left(-2y_k \varphi_i(x_k) + 2\sum_{j=1}^{n} a_j \varphi_j(x_k)\varphi_i(x_k) \right)$$

$$= -2\sum_{k=1}^{m}\left(y_k \varphi_i(x_k) - \sum_{j=1}^{n} a_j \varphi_j(x_k)\varphi_i(x_k) \right)$$

$$\boxed{\frac{\partial E}{\partial a_i} = 0}$$

$$\sum_{k=1}^{m} y_k \varphi_i(x_k) = \sum_{k=1}^{m}\sum_{j=1}^{n} a_j \varphi_j(x_k)\varphi_i(x_k)$$

係数を求める

$$\sum_{k=1}^{m} y_k \varphi_i(x_k) = \sum_{k=1}^{m}\sum_{j=1}^{n} a_j \varphi_j(x_k)\varphi_i(x_k)$$

: a_iについて偏微分したもの

$$\begin{pmatrix} \sum_{k=1}^{m} y_k \varphi_1(x_k) \\ \sum_{k=1}^{m} y_k \varphi_2(x_k) \\ \vdots \\ \sum_{k=1}^{m} y_k \varphi_n(x_k) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} \varphi_1(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_n(x_k) \\ \sum_{k=1}^{m} \varphi_2(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_n(x_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} \varphi_n(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_n(x_k) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{a}$の解$(\boldsymbol{a})$として係数ベクトルが求まる.

多項式近似

$$y = \sum_{j=1}^{n+1} a_{k-1} x^{k-1}$$

次数(n+1)はデータ数m以下
(一般的には十分小さくとる)

例: 1次(直線) 近似

$(x, y) = (1.0, 5.0), (2.0, 7.1), (3.0, 8.9)$の3点を考えてみる.

$$y = a_0 + a_1 x$$

$$\varphi_1 = 1$$

$$\varphi_2 = x$$

線形近似

左辺値の計算

$$\boxed{\begin{pmatrix} \sum_{k=1}^{m} y_k\varphi_1(x_k) \\ \sum_{k=1}^{m} y_k\varphi_2(x_k) \\ \vdots \\ \sum_{k=1}^{m} y_k\varphi_n(x_k) \end{pmatrix}} = \begin{pmatrix} \sum_{k=1}^{m} \varphi_1(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_n(x_k) \\ \sum_{k=1}^{m} \varphi_2(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_n(x_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} \varphi_n(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_n(x_k) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

φは2つなので2行，データ数3点なのでm=3

φ_1はxの値によらず常に1

$$\begin{pmatrix} 5.0 \cdot 1.0 + 7.1 \cdot 1.0 + 8.9 \cdot 1.0 \\ 5.0 \cdot 1.0 + 7.1 \cdot 2.0 + 8.9 \cdot 3.0 \end{pmatrix} = \begin{pmatrix} 21.0 \\ 45.9 \end{pmatrix}$$

φ_2はxの値そのもの

$(x, y) =$
$(1.0, 5.0),$
$(2.0, 7.1),$
$(3.0, 8.9)$

$y = a_0 + a_1 x$

$\varphi_1 = 1$

$\varphi_2 = x$

線形近似　右辺値の計算

$$\begin{pmatrix} \sum_{k=1}^{m} y_k\varphi_1(x_k) \\ \sum_{k=1}^{m} y_k\varphi_2(x_k) \\ \vdots \\ \sum_{k=1}^{m} y_k\varphi_n(x_k) \end{pmatrix} = \boxed{\begin{pmatrix} \sum_{k=1}^{m} \varphi_1(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_1(x_k)\varphi_n(x_k) \\ \sum_{k=1}^{m} \varphi_2(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_2(x_k)\varphi_n(x_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} \varphi_n(x_k)\varphi_1(x_k) & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m} \varphi_n(x_k)\varphi_n(x_k) \end{pmatrix}} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

φは2つなので2行2列，データ数3点なのでm=3

$$\begin{pmatrix} 1.0\times1.0+1.0\times1.0+1.0\times1.0 & 1.0\times1.0+1.0\times2.0+1.0\times3.0 \\ 1.0\times1.0+2.0\times1.0+3.0\times1.0 & 1.0\times1.0+2.0\times2.0+3.0\times3.0 \end{pmatrix}$$

$$= \begin{pmatrix} 3.0 & 6.0 \\ 6.0 & 14.0 \end{pmatrix}$$

対称行列なのでどちらかの計算でOK

$(x, y) =$
$(1.0, 5.0)$,
$(2.0, 7.1)$,
$(3.0, 8.9)$

$y = a_0 + a_1 x$

$\varphi_1 = 1$

$\varphi_2 = x$

線形近似

$$\begin{pmatrix} \sum_{k=1}^{m} y_k\varphi_1(x_k) \\ \sum_{k=1}^{m} y_k\varphi_2(x_k) \\ \vdots \\ \sum_{k=1}^{m} y_k\varphi_n(x_k) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m}\varphi_1(x_k)\varphi_1(x_k) & \sum_{k=1}^{m}\varphi_1(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m}\varphi_1(x_k)\varphi_n(x_k) \\ \sum_{k=1}^{m}\varphi_2(x_k)\varphi_1(x_k) & \sum_{k=1}^{m}\varphi_2(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m}\varphi_2(x_k)\varphi_n(x_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m}\varphi_n(x_k)\varphi_1(x_k) & \sum_{k=1}^{m}\varphi_n(x_k)\varphi_2(x_k) & \cdots & \sum_{k=1}^{m}\varphi_n(x_k)\varphi_n(x_k) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$$\begin{pmatrix} 21.0 \\ 45.9 \end{pmatrix} = \begin{pmatrix} 3.0 & 6.0 \\ 6.0 & 14.0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

$(x, y) =$
$(1.0, 5.0),$
$(2.0, 7.1),$
$(3.0, 8.9)$

$$y = a_0 + a_1 x$$

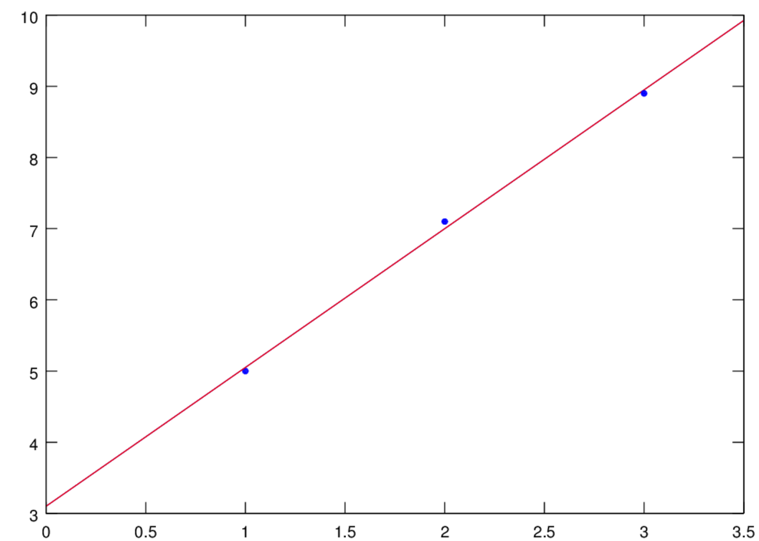
$$\varphi_1 = 1$$

$$\varphi_2 = x$$

ガウスの消去法(部分ピボット選択付)

$$\begin{pmatrix} 6.0 & 14.0 \\ 0.0 & -1.0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 45.9 \\ -1.95 \end{pmatrix}$$

$$\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 3.10 \\ 1.95 \end{pmatrix}$$



演習課題8

- input.csvにより1行1データの形式で与えられる，4点のデータを2次多項式($ax^2 + bx + c$)で近似し，関数を示せ．可能であれば，データと近似関数をグラフで示せ．

(x, y):　$(-3.00, -26.3)$　$(1.00, 2.05)$　$(1.50, 0.99)$　$(4.70, -29.0)$

手順
1. 各係数を求める
2. ガウスの消去法

$$\varphi_1 = 1$$

$$\varphi_2 = x$$

$$\varphi_3 = x^2$$

おまけ

- リダイレクトとパイプ
  - リダイレクト: 標準入出力をファイルへと切り替える.　> <
    ※ 標準エラー出力も変えられます. (標準ストリーム)
  - パイプ: プログラム(プロセス)間を直接接続する　|

パイプを使えば, プログラムの実行結果をgnuplotに渡してグラフが書けます.

ex: hogehoge.exe < indata.txt | gnuplot

- パイプはプログラム中で直接開くこともできます. (ファイルと同じ)

  popen, pclose関数を使う.
  (VSでは_popen, _pclose)

参考コード(抜粋)

```
FILE *pOut = _popen("gnuplot", "w");// VS以外ではpopen関数
fprintf(pOut, "set zeroaxis¥n");
fprintf(pOut, "set xrange [%f:%f]¥n", A[1][1] -0.5, A[M][1]+0.5);
fprintf(pOut, "plot ");
for (i = N+1; i > 1; i--)
    fprintf(pOut,"%2.2f * x**%d + ", a[i], i-1);
fprintf(pOut,"%2.2f * x**%d¥n", a[i], i-1);

fprintf(pOut, "replot '-' title 'data' ¥n");
for (i = 1; i <= M; i++)
{
    fprintf(pOut,"%f¥t%f¥n", A[i][1], A[i][2]);
}
fprintf(pOut,"e¥n");
fflush(pOut);

printf("press enter key to terminate program¥n");
getchar();
_pclose(pOut);// パイプのクローズ
```