補間法(ラグランジュ補間とスプライン補間)

山本昌志*

2003 年 11 月 25 日

1 はじめに

実験やシミュレーションを行っていると $(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3)$ のように、離散的にデータが得られるのは普通のことである。この得られた値から、任意の x に対して y の値を推測しなくてはならないことがしばしば生じる。工学実験では曲線定規を使って値を推測したが、ここでは計算機を用いて値を推測することを学習する。

値を計算するためには、$y = f(x)$ を示す関数 f を決めればよい。この関数の決め方には、いろいろな方法がある。計算機応用の講義では、2 つの方法を学習する。

補間法 これは、離散的な点、全てを通過する関数を求め、値を推測する方法です (図 1)。データ数が多くなると、問題が生じることがあります。

最小 2 乗法 データをある特定の関数に近似して、値を推測する方法です (図 2)。通常、関数形を決めるパラメーターよりもデータの数が多く、全てのデータを通る曲線が得られるわけではありません。最も誤差が少なくなるように、関数のパラメーターを決めます。

今回の講義では、前者の補間法の学習を行います。補間法にもいろいろ有りますが、ここでは最も基本的なラグランジュ補間とスプライン補間を習得してもらいます。これらを学習した後、最小 2 乗法について説明をします。

一般的には、補間とは得られたデータの範囲内での値の推定のことを言います。範囲外の推定は補外と言います[1]。例えば、図 3 のように (x_0, y_0)〜(x_3, y_3) のデータがあり、それらを通る曲線が得られたとしよう。データの範囲 $[x_1, x_3]$ の推定を補間、それ以外を補外といる。ただし、どちらも同じ関数を用います。

補間に比べて、補外の方が近似の精度が悪くなる場合が多いです。このことは、証券価格のグラフを考えると良く分かります。本日までのデータを補間することはそんなに難しくありませんが、明日以降の価格を推定することは極めて難しくなります。もし、良い精度で明日以降の価格がわかれば、大金持ちになれるでしょう。

*国立秋田工業高等専門学校　電気工学科

[1]補間のことを内挿、補外のことを外挿ということもある。

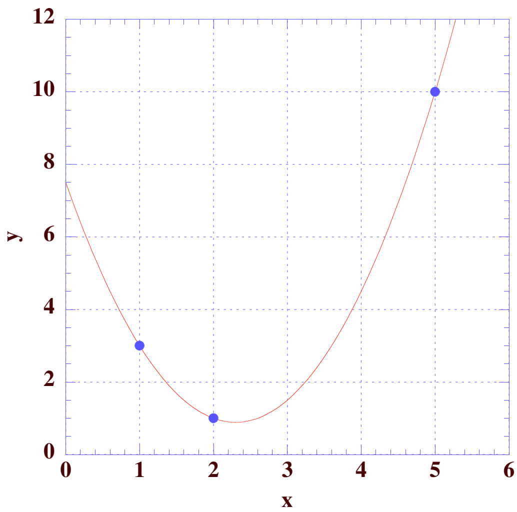

図 1: 補間 (ラグランジュ補間)

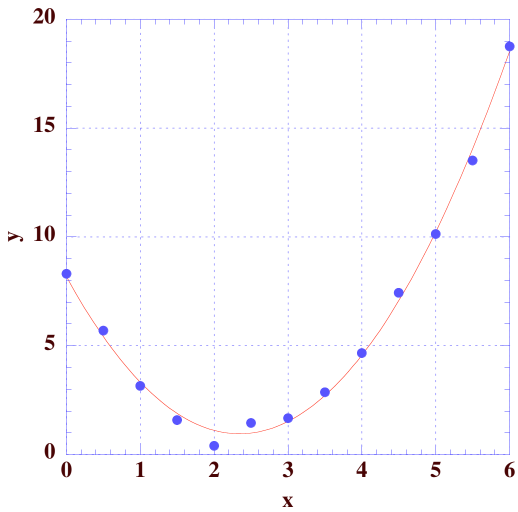

図 2: 最小 2 乗近似 (2 次関数)

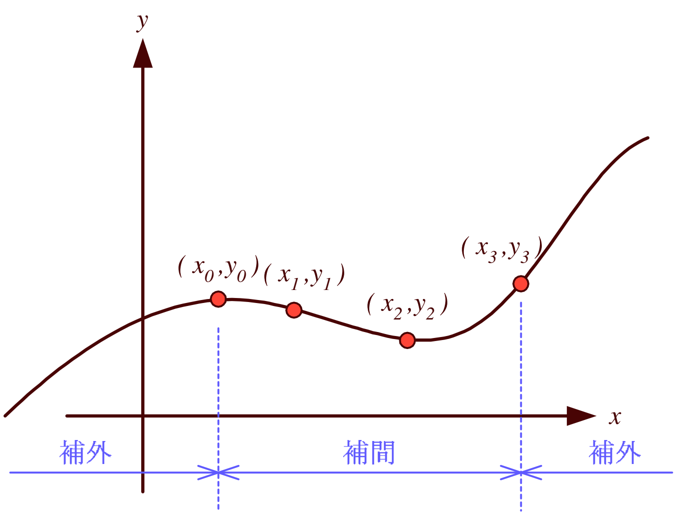

図 3: 補間と補外

2　ラグランジュ補間

2.1　基本的な考え方

ある物理量を測定して N+1 個の値が得られたとしよう。それらは、(x_0, y_0) , (x_1, y_1), (x_2, y_2), $\cdots$, (x_N, y_N) の 2 次元座標で示すことができるでしょう。この全ての点を通る関数を求めることが補間法の課題です。N 次関数を使えばその目的が達成できると容易に分かります。すなわち、

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_i x^i + \cdots + a_N x^N \tag{1}$$

と補間するわけです。この係数 a_i を求めれば、補間の関数が求められたことになります。この係数は、N+1 元の連立 1 次方程式を解くことにより求めることができます。

連立方程式の計算は時間がかかります。それに代わるもっと良い方法があります。ここでは N 次関数で表現できれば良いわけで、以下がそれになります。

$$\begin{aligned}
y = &\frac{(x-x_1)(x-x_2)(x-x_3)\cdots(x-x_N)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)\cdots(x_0-x_N)}y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)\cdots(x-x_N)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)\cdots(x_1-x_N)}y_1 \\
&+ \frac{(x-x_0)(x-x_1)(x-x_3)\cdots(x-x_N)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)\cdots(x_2-x_N)}y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)\cdots(x-x_N)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)\cdots(x_3-x_N)}y_3 \\
&\cdots + \frac{(x-x_0)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_N)}{(x_k-x_0)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_N)}y_k + \cdots \\
&+ \frac{(x-x_0)(x-x_1)(x-x_3)\cdots(x-x_{N-1})}{(x_N-x_0)(x_N-x_1)(x_N-x_2)\cdots(x_N-x_{N-1})}y_N
\end{aligned} \tag{2}$$

この式 (2) を見ると、

- 各項の分子は定数である。分母は N 次関数です。このことから、全ての項は N 次関数になっているので、この式は N 次関数 (N 次多項式) です。

- x に $x_0, x_1, x_2, \cdots, x_N$ を代入すると、y の値は $y_0, y_1, y_1, \cdots, x_N$ になることが分かります。これは、データ点 (x_0, y_0) , (x_1, y_1), (x_2, y_2), $\cdots$, (x_N, y_N) の全てを通過していることを示してます。

となっている。これは、表現こそ違うものの式 (1) と同じです。これは、式 (1) の a_i を求めて補間の関数を求める必要は無く、式 (2) を使えばよいということです。この補間をラグランジュの補間多項式 (Lagrange's interpolating polynomial) と言います。式 (2) をもうちょっと格好良く書けば、

$$\begin{aligned}
L(x) &= \sum_{k=0}^{N} L_k(x) y_k \\
\text{ただし、}\quad L_k(x) &= \prod_{j=0}^{N(j\neq k)} \frac{x-x_j}{x_k-x_j}
\end{aligned} \tag{3}$$

となります。

2.2 問題点

補間の点数が増えてくると、ラグランジュの補間には問題が生じます。具体例を図 4 に示します。これを見ると分かるように、補間の関数が振動しています。ラグランジュの補間では、補間の点数が増えてくると大きな振動が発生して、もはや補間とは言えなくなります。ラグランジュの補間には常にこの問題が付きまといますので、データ点数が多い場合は使えません。ほかの補間を考える必要があります。

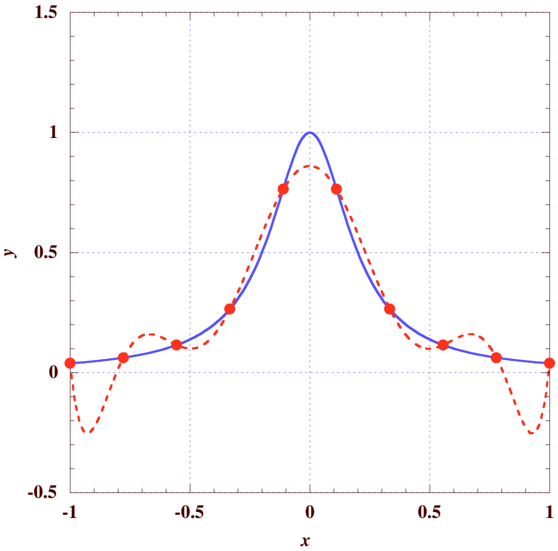

図 4: ラグランジュ補間の問題点。$y = \frac{1}{1+25x^2}$ を 10 点で補間 (点線) したが、両端で振動する。

3 スプライン補間

3.1 区分多項式

ラグランジュの補間は、データ点数が増えてくると関数が振動し問題が発生します。そこで、補間する領域をデータ間隔 $[x_i, x_{i+1}]$ に区切り、その近傍の値を使い低次の多項式で近似することを考えます。区分的に近似関数を使うわけですが、上手に近似をしないと境界でその導関数が不連続になります。導関数が連続になるように、上手に近似する方法がスプライン補間 (spline interpolation) です。

ここでは、通常よくつかわれる 3 次のスプライン補間を考えます。補間する関数が 3 次関数を使うためそう呼ばれているのです。これ以降の説明は、文献 [1] を参考にしました。

補間をするデータは、先と同じように (x_0, y_0) , (x_1, y_1), (x_2, y_2), $\cdots$, (x_N, y_N) とします。そして、区間 $[x_j, x_{j+1}]$ で補間をする関数を $S_j(x)$ とします。この様子を図 5 に示します。

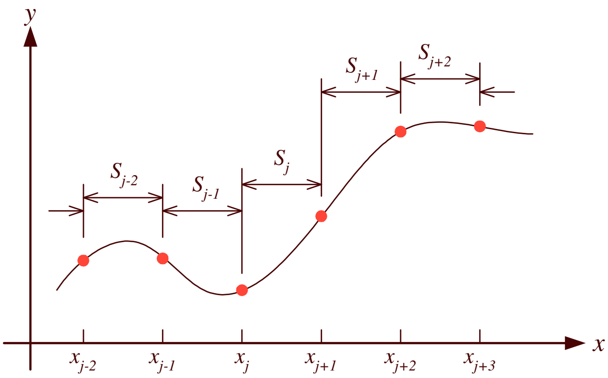

図 5: スプライン補間の区分

3.2　係数が満たす式

3 次のスプライン補間を考えるので、

$$S_j(x) = a_j(x - x_j)^3 + b_j(x - x_j)^2 + c_j(x - x_j) + d_j \qquad (j = 0, 1, 2, 3, \cdots, N - 1) \tag{4}$$

となります。この a_j, b_j, c_j, d_j を決めなくてはなりません。

これらの未知数は、4N 個あります。従って、4N 個の方程式が必要になります。そのために、3 次のスプライン補間に以下の条件を課します。

- 全てのデータ点を通る。各々の $S_j(x)$ に対して両端での値が決まるため、2N 個の方程式ができます。
- 各々の区分補間式は、境界点の 1 次導関数は連続とする。これにより、N-1 個の方程式ができます。
- 各々の区分補間式は、境界点の 2 次導関数は連続とする。これにより、N-1 個の方程式ができます。

以上の条件を課すと方程式は 4N-2 個の方程式で表現できます。未知数は 4N 個なので、2 個方程式が不足しています。この不足を補うために、いろいろな条件が考えられますが、通常は両端 x_0 と x_N での 2 次導関数の値を 0 とします。すなわち、$S_0''(x_0) = S_{N-1}''(x_N) = 0$ です。これを自然スプライン (natural spline) と言います。自然スプライン以外には、両端の 1 次導関数の値を指定するものもあります。

これで全ての条件が決まりました。あとは、この条件に満たす連立方程式を求めるだけです。まずは、2 次導関数が区分関数の境界で等しいという条件からです。$x = x_j$ における 2 次導関数の値を u_j とします。すなわち、

$$u_j = S''(x_j) \qquad (j = 0, 1, 2, \cdots, N) \tag{5}$$

です。$u_j = S_{j-1}''(x_j) = S_j''(x_j)$ としますので、2 次導関数の条件は満足されたことになります。この式から、

$$u_j = S_j''(x_j) = 2b_j \qquad (j = 0, 1, 2, \cdots, N - 1) \tag{6}$$

となります。これから、

$$b_j = \frac{u_j}{2} \tag{7}$$

が、直ちに導けます。ここで、スプライン補間の係数、すなわち計算で求めるべき b_j を u_j で表した理由があります。以降の議論を見ると分かるように、u_j を連立方程式で計算することにより、他の係数を求めることができます。そのようなわけで、できるだけ u_j で表現するようにします。

さらに 2 次導関数の計算から、

$$u_{j+1} = S_j''(x_{j+1}) = 6a_j(x_{j+1} - x_j) + 2b_j \qquad (j = 0, 1, 2, \cdots, N-1) \tag{8}$$

が導けます。この式から、a_j を計算すると、

$$\begin{aligned} a_j &= \frac{u_{j+1} - 2b_j}{6(x_{j+1} - x_j)} \\ &= \frac{u_{j+1} - u_j}{6(x_{j+1} - x_j)} \qquad (j = 0, 1, 2, \cdots, N-1) \end{aligned} \tag{9}$$

となります。これで、2 次導関数の条件は終わりです。

つぎに、全てのデータ点上を通過する (最初の条件) という条件を考えます。まずは、区分の左端の点から、

$$d_j = y_j \tag{10}$$

が直ちに導けます。つぎに、区分の右端の点から

$$a_j(x_{j+1} - x_j)^3 + b_j(x_{j+1} - x_j)^2 + c_j(x_{j+1} - x_j) + d_j = y_{j+1} \tag{11}$$

が導けます。式 (7),(9),(10) を用いると、

$$\begin{aligned} c_j &= \frac{1}{x_{j+1} - x_j}\left[y_{j+1} - a_j(x_{j+1} - x_j)^3 - b_j(x_{j+1} - x_j)^2 - d_j\right] \\ &= \frac{1}{x_{j+1} - x_j}\left[y_{j+1} - \frac{u_{j+1} - u_j}{6(x_{j+1} - x_j)}(x_{j+1} - x_j)^3 - (x_{j+1} - \frac{u_j}{2}x_j)^2 - y_i\right] \\ &= \frac{y_{j+1} - y_i}{x_{j+1} - x_j} - \frac{1}{6}(x_{j+1} - x_j)(2u_j + u_{j+1}) \end{aligned} \tag{12}$$

となります。

これで、a_j と b_j、c_j、d_j が x_j と y_j、u_j で表せました。x_j と y_j はデータ点なので、値はわかっています。したがって、u_j が分かれば、補間に必要な係数が全て分かります。

3.3 連立方程式

それでは、u_j をどうやって求めるのでしょうか?。これは、まだ使われていない条件、1 次導関数が境界点で等しいことを使います。次の式を使います。

$$S'(x_{j+1}) = S_j'(x_{j+1}) = S_{j+1}'(x_{j+1}) \qquad (j = 0, 1, 2, \cdots, N-2) \tag{13}$$

これと式 (4) から、

$$3a_j(x_{j+1} - x_j)^2 + 2b_j(x_{j+1} - x_j) + c_j = c_{j+1} \tag{14}$$

となります。あとは、この式の a_j と b_j、c_j を x_j と y_j、u_j で表して、u_j の連立方程式にするだけです。最終的に式は、

$$(x_{j+1}-x_j)u_j+2(x_{j+2}-x_j)u_{j+1}+(x_{j+2}-x_{j+1})u_{j+2}=6\left[\frac{y_{j+2}-y_{j+1}}{x_{j+2}-x_{j+1}}-\frac{y_{j+1}-y_j}{x_{j+1}-x_j}\right] \tag{15}$$

となります。この方程式は、$j=0,1,2,\cdots,N-2$ で成り立ちます。従って、式の数は、N-1 個です。u_j の数は N+1 個ですが、$u_0=u_N=0$ ですので、未知の u_j は N-1 個となります。式 (15) を解くことにより、全ての u_j が決定できます。これが決まれば、a_j と b_j、そして c_j が計算できます。

$u_0=u_N=0$ を代入した連立 1 次方程式は、

$$\begin{pmatrix} 2(h_0+h_1) & h_1 & & & & & \\ h_1 & 2(h_1+h_2) & h_2 & & \text{\huge 0} & & \\ & h_2 & 2(h_2+h_3) & h_3 & & & \\ & & & \ddots & & & \\ & & & h_{j-1} & 2(h_{j-1}+h_j) & h_j & \\ & \text{\huge 0} & & & & \ddots & \\ & & & & & h_{N-2} & 2(h_{N-2}+h_{N-1}) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_j \\ \vdots \\ u_{N-1} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_j \\ \vdots \\ v_{N-1} \end{pmatrix} \tag{16}$$

となります。ただし、h_j と v_j は以下のとおり。

$$h_j=x_{j+1}-x_j \qquad (j=0,1,2,\cdots,N-1) \tag{17}$$

$$v_j=6\left[\frac{y_{j+1}-y_j}{h_j}-\frac{y_j-y_{j-1}}{h_{j-1}}\right] \qquad (j=0,1,2,\cdots,N-1) \tag{18}$$

4 練習問題

- 教科書の例題 1(p.47) をラグランジュ補間して、区間 $[-1,5]$ のグラフを書け。
- ラグランジュ補間の問題点を示した関数 $y=\frac{1}{1+25x^2}$ のデータ点数と補間関数の関係を調べる。区間 $[-1,1]$ で等間隔にデータ数を 5,11,21 と変化させて、ラグランジュ補間のグラフを書け。
- 先のラグランジュ補間の問題点を示した関数について、同様のことをスプライン補間せよ。

参考文献

[1] 高橋大輔. 数値計算. 岩波書店, 1996.