データ外挿の新しい手法に関する研究

九州大学大学院○学生員　水沼　道博
九州大学工学部　正　員　朝位　孝二
九州大学工学部　正　員　小松　利光

1. はじめに

ランダム変動の予測精度は、与えられるデータの性質にかなり左右されるものと思われる。機械・電気系のように、元来人間が設計したシステムの場合には、物理的評価式が既知で変動に混入する雑音もある程度判明しており、カルマンフィルターなどの制御理論の適用によって高精度な予測が可能である。しかし、自然現象や社会現象のように、変動に関与する物理的メカニズムが不明でモデル化が困難であれば、制御理論を適用しても高精度な予測結果は期待できない。本研究は、物理的あるいは統計的手法に基づかない幾何学的な予測手法、すなわち外挿法を用いて簡便にオンライン予測を行う方法を検討したものである。

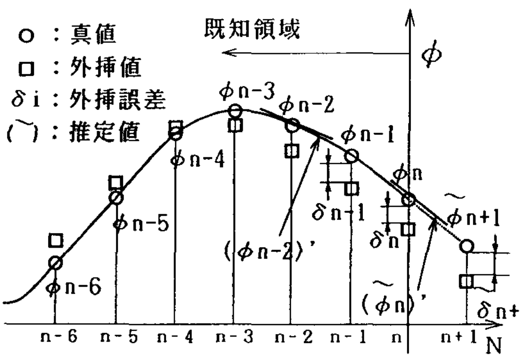
2. 外挿式の誘導

図－１に示す n+1点の値を知るための最も簡単な方法は、n点とn-1点の値を用いて線型外挿を行うことである。ここではやはり線型外挿ではあるが、n点における勾配を 6-point scheme[1] を利用して評価し一次外挿多項式を求めることにする。以下に外挿値$\tilde{\phi}_{n+1}$を計算する方法を示す。

① ：図-1中のn-5～n点に 6-pointスキームを適用し、n-2点での勾配$(\phi_{n-2})'$を求める。

② ：①で求めた勾配と、$\phi_{n-2}, \phi_{n-1}, \phi_n$ の４つの条件により３次多項式を構成する。

③ ：②で求めた３次多項式により n点における勾配$(\tilde{\phi}_n)'$を決定し、その勾配とϕ_nより求まる一次多項式より$\tilde{\phi}_{n+1}$を求める。



図－１　記号説明

上記の方法を図－１中のϕ_{n-5}～ϕ_nを用いて定式化すれば次式を得る。

$$\tilde{\phi}_{n+1}=-0.05633\phi_{n-5}+0.3097\phi_{n-4}-1.033\phi_{n-3}+2.2869\phi_{n-2}-3.4333\phi_{n-1}+2.9256\phi_n \quad (1)$$

(1)式による予測値 $\tilde{\phi}_{n+1}$ は、場合によってはかなりの予測誤差を含んでいると考えられる。そこで、$\tilde{\phi}_n$を以下の手順で補正することを試みた。

①’：ϕ_{n-6}～ϕ_{n-1}の値を用いて(1)式により$\tilde{\phi}_n$の値を求め、真値との誤差δ_nを計算する。

②’：同様の計算を順次行って$\tilde{\phi}_{n-5}$～$\tilde{\phi}_{n-1}$を算出し、δ_{n-5}～δ_{n-1}を求める。（δ_{n-5}を評価するためにはϕ_{n-11}～ϕ_{n-6}の値を必要とするため、計算に使用するデータ点の数はn-11～nまでの12点となる。）

③’：δ_{n-5}～δ_nの値を(1)式に代入して$\tilde{\delta}_{n+1}$を計算する。

④’：③’で求めた$\tilde{\delta}_{n+1}$を補正値として$\tilde{\phi}_{n+1}$に加える。

$\tilde{\delta}_{n+1}$をϕ_{n-11}～ϕ_nを用いて定式化すると以下のようになる。

$$\tilde{\delta}_{n+1}=0.00317\phi_{n-11}-0.0348\phi_{n-10}+0.2123\phi_{n-9}-0.8974\phi_{n-8}+2.8705\phi_{n-7}-7.1809\phi_{n-6} \quad (2)$$
$$+14.1913\phi_{n-5}-22.0573\phi_{n-4}+26.2007\phi_{n-3}-22.3759\phi_{n-2}+11.9925\phi_{n-1}-2.9256\phi_n$$

(1), (2)式を用いることにより、既知領域の境界近傍の12点のデータ値によって境界外１点の値を予測することができる。

3. 外挿適用例

今回導出した外挿法を解析関数に適用した例を図-2に示す。計算に用いた関数は図－２に示すように、ピーク位置 8, 16, 24, ピーク値 1.0, 2.0, -1.5標準偏差 2.0, 3.0, 0.8をもつ３つの正規分布を線形結合したものである。データ間隔は 1.0で、図の左から右へオンラインで情報が与えられるものとして、１ステップ先の値を予測した。なお情報が十分与えられないうちは、関数をそのまま左側に直線で延長し、ダミーデータを与えている。　図－２において、左側２つの正規分布はデータ間隔に対して大きい標準偏差をもつ。つまり変動スケールに対し十分細かいデータ間隔であるため、線形外挿と比較して予測精度は高く、また補正後の値は補正前の値よりもさらに真値に近くなる。これに対し、右側の分布では、データ間隔は変動スケールに対して粗く設定していることになる。この場合本外挿法による予測値は、線形外挿値よりも不安定となる。また、図－３はこのときの補正値をプロットしたものである。補正誤差が統計的に変化しているところでは誤差の予測も精度がよく、良好な補正結果が得られることが分かる。しかしながら、変動スケールが小さい所では補正値が大きく乱れてしまい、誤差予測の精度が低下する。データ間隔が変動スケールに対して十分細かく設定されていない場合には、現在より以前の情報が１ステップ後に起こり得る値に与える影響が小さくなり、１次のマルコフ過程に近い状況となるものと思われる。従ってこのような条件では、線形外挿の方がかえって安定な結果を与えるものと考えられる。

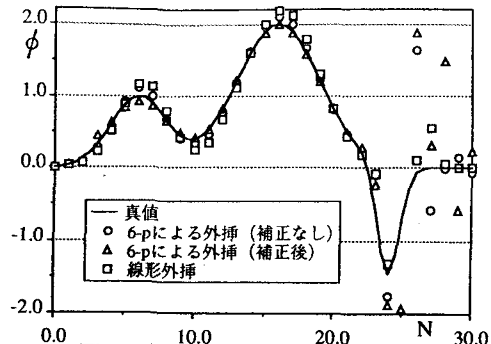


図-２ 解析関数をオンラインで外挿した場合

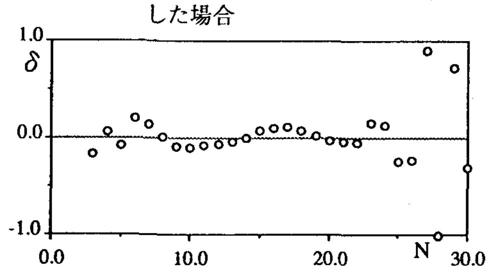


図-３ 図-２の外挿の補正に使用する誤差

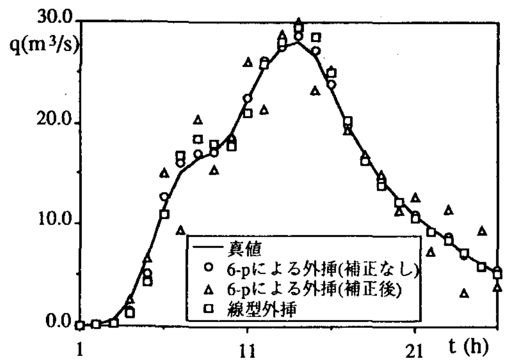


図-４ 流出量の時間変化に外挿法を適用した場合

図－４は、河川流出量[2)]の時間変化の変動に対してオンラインによる外挿を適用したもので、データ間隔は１時間である。この場合の変動に本外挿法を適用したところ、補正誤差が系統的に現れなかったため、補正計算を行うと予測値はかえって真値からずれる結果となった。しかしながら、(1)式のみを用いて計算した場合では、線形外挿値より正確な予測値を与えていることが分かる。

4. おわりに

6-point schemeに基づく線型外挿法及び外挿値の補正の考え方を提案した。外挿（予測）においては、内挿と異なりどんな結果でも主張できる。また予期せぬじょう乱に対しては、いかなる予測法も無力である。これらのことを考慮にいれれば、本研究における外挿法は、変動スケールに対しての制約はあるもののシステムの状態がはっきりしない変動に対しては有用であると考えられる。

＜参考文献＞　1)T.Komatsu,et.al:Numerical Calculation of Pollutant Transport in One and Two Dimensions,JHHE,Vol.3,No2,JSCE,1985

2)日野幹雄：水門流出系予測へのカルマンフィルター理論の適用，土木学会論文報告集 第221号,pp39-47,1974