I－44　　確率論的最適化手法の基礎考察

(株)地崎工業 技術開発部　　正会員　須藤 敦史

1．はじめに

施工中に得られる構造物の挙動から物性値を再評価する逆解析[1)-3)]や既知観測値から周辺の未観測点の値を推定する補間[4)-6)]が行われてきている。しかし物性値等は本来の性質により不確定性やばらつきを示す値であり観測値は観測誤差を含む量である。そこで、確率論的最適化手法ではこれらを考慮した問題解法のｱﾙｺﾞﾘｽﾞﾑを構成している。また、離散変数等の組み合わせ最適化問題は、多数の解候補の中から特殊な組み合わせを見つけだす確定的問題であり最適解の探索は極めて難しいため、問題を確率的に考えた組み合わせ最適化手法が用いられてきている[7),8)]。この手法においてもそのｱﾙｺﾞﾘｽﾞﾑに集合･統計等の確率的要素を取り入れることにより、この問題をある程度解決している。

以上のように、確率論を用いた最適化手法はその有用性が認められてきているものの解法の基本的な考え方は明確になっていない。そこで本報告は確率論を用いた最適化手法に対する基本的な考え方の考察を試みている。

2．確率論を用いた最適化手法の分類

確率論を用いた最適化手法では、適用問題とその確率的な考え方に応じて連続型と離散型の大別される。連続型は主に地盤の物性値等の逆問題や推定問題に用いられ、離散型は組み合わせ問題における最適解の探索に用いられている。加えて確率を用いる大きな違いでは、連続型は問題事体が確率的であるのに対し、離散型は解法に確率的な考えを利用している点である。

2.1連続型最適化手法

逆解析･ﾊﾟﾗﾒｰﾀ同定問題では、ばらつきや不確定性を有する状態量(物性値等)を平均や分散などの確率･統計的な特性を有する空間に分布する変数と定義し、かつ観測値も観測誤差を含む値としている。そして地盤試験等により得られた事前分布(分散)を観測値を用いて更新し、その事後分布を推定するものである。言い換えれば、観測値が得られた基で状態量の推定誤差分散を最小にする条件付き最小分散推定である。

ここで連続型最適化手法のうちで代表的なﾍﾞｲｽﾞ推定を例に上げ、そのｱﾙｺﾞﾘｽﾞﾑを考える。ﾍﾞｲｽﾞ推定では状態量の統計的性質が正規分布であると定義している。加えて事前情報(初期期待値･共分散値)が以下に示すように与えられるものとして、事後情報(観測値)が得られる度に確率密度関数を逐次的に更新し、その最適な期待値を推定する手法である。

$$E(x)=\bar{x} \tag{1}$$

$$E[(x-\bar{x})(x-\bar{x})]=M \tag{2}$$

ここで状態量と観測値との間には次式に示すような線形関係があるとする。

$$z=Hx+v \tag{3}$$

$$E\{v\}=0, E\{vv^T\}=R$$

状態量の確率場が正規分布であるとすると事前の確率密度関数は以下の式となる。

$$p(x)=\frac{1}{(2\pi)^{\frac{n}{2}}|M|^{\frac{1}{2}}}\exp\{-\frac{1}{2}(x-\bar{x})^T M^{-1}(x-\bar{x})\} \tag{4}$$

ここで観測値zが得られた時のx_tが起こる条件付き確率密度関数はﾍﾞｲｽﾞ定理より以下の式となる。

$$p(x|z)=\frac{p(z|x)p(x)}{p(z)} \tag{5}$$

$$p(x|z)=\frac{|HMH^t+R|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}|R|^{\frac{1}{2}}|M|^{\frac{1}{2}}}\exp\{-\frac{1}{2}[(x-\bar{x})^T M^{-1}(x-\bar{x})+(z-Hx)^T R^{-1}(z-Hx)$$
$$-(z-H\bar{x})^T(HMH^T+R)^{-1}(z-H\bar{x})]\} \tag{6}$$

A Basic Consideration of Stochastic Optimization method
By Atsushi SUTOH

よって、上式で得られる条件付き確率密度関数を最小にするようなxは最適推定値となり、exp中の第3項は定数となるため、以下の目的関数を最小にすればよい。

$$\theta = \frac{1}{2}\{(x-\bar{x})^T M^{-1}(x-\bar{x}) + (z-Hx)^T R^{-1}(z-Hx)\} \tag{7}$$

したがって、ﾍﾞｲｽﾞ推定では式(7)の右辺第1項(事前情報)と第2項(事後の観測情報)の誤差分散値に応じてxを最小とする最適推定値として繰り返しｱﾙｺﾞﾘｽﾞﾑを構成したものである。

ここで、状態量の最適推定値$\hat{x}$と観測後の推定誤差共分散値Pは以下となる。

$$\hat{x} = \bar{x} + PH^T R^{-1}(z - H\bar{x}) \tag{8}$$

$$P = (M^{-1} + H^T R^{-1} H)^{-1} \tag{9}$$

$$P = E\{(x-\hat{x})(x-\hat{x})^T\}$$

この式(9)において第2項目は非負行列となるので、観測後の推定誤差の共分散値は減少し、状態量の収束時には最小となる。

拡張ｶﾙﾏﾝﾌｨﾙﾀは対象となるｼｽﾃﾑを非線形として、その状態を表現する状態方程式と観測系を表す観測方程式で構成される。

$$\frac{dX_t}{dt} = f(X_t, t) + G_t w_t \tag{10}$$

$$Yt_k = h(Xt_k, t_k) + vt_k \tag{11}$$

X_t, Xt_k:連続型,離散型状態量

w_t, vt_k:ｼｽﾃﾑﾉｲｽﾞ,観測ﾉｲｽﾞ

Yt_k:観測量, G_t:変換行列

拡張ｶﾙﾏﾝﾌｨﾙﾀでは基本式(10),(11)を推定状態量近傍で線形化し、漸化的な最適状態量の推定ｱﾙｺﾞﾘｽﾞﾑを構成したものである。ここで観測値が得られた条件のもとでﾍﾞｲｽﾞ推定と同様な考え方で最適なx_tは次式の目的関数を最小にする値を求めればよい。

$$\theta = \frac{1}{2}\{(x_t-\bar{x}_t)^T M^{-1}(x_t-\bar{x}_t) + (z_t - H_{tk}x_t + h_t(x_{tk}) - H_{tk}x_{tk})^T R^{-1}(z_t - H_{tk}x_t + h_t(x_{tk}) - H_{tk}x_{tk}\} \tag{12}$$

この目的関数を最小にするx_{tk}をｶﾞｳｽﾆｭｰﾄﾝ法により繰り返し推定するｱﾙｺﾞﾘｽﾞﾑを構成したのが、拡張ｶﾙﾏﾝﾌｨﾙﾀの漸化式として次式が得られる。

$$x_{tk+1} = x_{tk} + P_{tk}H_{tk}^T R_t^{-1}(z_t - h_t(x_{tk})) + P_{tk}M_t^{-1}(\bar{x}_t - x_{tk}) \tag{13}$$

$$P_{tk} = (M_t^{-1} + H_{tk}^T R_t^{-1} H_{tk})^{-1} = M_t - K_{tk}H_{tk}M_t \tag{14}$$

$$K_{tk} = M_t H_{tk}^T (H_{tk}M_t H_{tk}^T + R_t)^{-1} \tag{15}$$

K_{tk}:ｶﾙﾏﾝｹﾞｲﾝ

したがって、拡張ｶﾙﾏﾝﾌｨﾙﾀにおいてもﾍﾞｲｽﾞ推定と同様に目的関数の右辺第1項の事前情報と第2項の事後の観測情報を最大とするx_tを最適推定値としている。加えて、事前情報と事後情報をその誤差分散値M, Rに応じた信頼度としている。ここで一般の逆問題では事前情報の信頼度は低い場合が多いため、それらを回避するｱﾙｺﾞﾘｽﾞﾑの提案もなされている[2),3)]。

以上のことより、この手法は事前情報における状態量の確率場(期待値や分散値・共分散値)が不明瞭であるため、観測情報を利用して確率密度関数を更新し、状態量の真の確率場(期待値が推定値)を推定する条件付き最小分散推定である。

一方、既存の観測値を用いて未観測点の値を推定する手法では、対象領域の期待値や分散値・共分散値が何らかの方法で求められるという条件の基で、同様に推定誤差分散値が最小になるように推定値を求める手法である。またこの方法では状態量と観測値の関数関係はない。この手法の中でKrigingを基本としたVariogramを例に上げると、以下に示すように未観測点(x_r)における推定値$z^*(x_r)$は既観測値$z(x_i)$の線形和となる。

$$z^*(x_r) = \lambda_0(x_r) + \sum_{i=1}^{N} \lambda_i(x_r) z(x_i) \tag{16}$$

$\lambda_i(x_r)$:未観測点における値を推定する係数, N:観測点数

ここで、不偏推定式の条件 $E[z^*(x_r)]=E[z(x_r)]$ より以下の式が得られ、また推定値の誤差分散値 $\sigma_\varepsilon^2(x_r)=E[\{z(x_r)-z^*(x_r)\}^2]$ が最小になるように重み係数 $\lambda_i(x_r)$ が決められる。

$$z^*(x_r)=\sum_{i=1}^{N}\lambda_i(x_r)z(x_i) \tag{17}$$

$$\sum_{i=1}^{N}\lambda_i(x_r)=1 \tag{18}$$

したがって、対象領域内における状態量z(x)を本質的に定常性を示す標本場(Intrinsic Random fields)における値と定義し､加えて観測点間の距離相関特性が得られれば未観測点の推定が行える。

よって、この手法は逆解析･ﾊﾟﾗﾒｰﾀ同定と異なり、状態量の確率場(期待値や分散値･共分散値)が既知であるため、不偏推定･最小誤差分散の条件のみで推定値が得られる。

2.2 離散型最適化手法

前記のように組み合わせ最適化問題は、多数の組み合わせの中から特殊な解を選択する確定的な問題であり、最適解の探索は極めて難しい。そこで、解候補を集合と考え、かつその効率的な解法としに確率論的な考え方をｱﾙｺﾞﾘｽﾞﾑに取り入れることにより、この問題をある程度解決している。言い替えれば、解の探索中において各状態の目的関数に確率という重みを与えることにより、効率的な探索を行うｱﾙｺﾞﾘｽﾞﾑである。ここで、組み合わせ最適化が実数目的関数 $f(x)$ を用いて式(19)のように定義する。

$$f(x)\rightarrow\min_{\{x\}} \tag{19}$$

x:状態変数, X:状態空間(有限集合)

このような組み合わせ最適化問題を定式化する際に確率的な考えが導入される。いま、対象となる基本空間をΩ, また確率をPとすると、式(19)に示すように目的関数 $f(x)$ の期待値 $g(x)$ を可能領域 $(x\in H)$ において最大化する問題となる。

$$g(x)\equiv\int_\Omega f(x;\omega)P\{d\omega\} \tag{20}$$

しかし、上式では目的関数の確率分布が既知でないため期待値は求められない。そこで確率論的な考えを導入する。この確率的解法に平衡統計力学の概念を導入した手法としてSimulated Annealing[7]がある。この方法は組み合わせ最適化問題を、式(20)に統計力学における分子の分布確率を導入することで式(21)に示すような確率的重み目的関数 $p(x)$ を用いた最小化問題に置き換えている。

$$\sum_{\{x\}}p(x)f(x)\rightarrow\min_{\{p(x)\}} \tag{21}$$

制約条件: $-\sum_{\{x\}}p(x)\log p(x)(\max)$

$\sum_{\{x\}}p(x)=1,\ 0\le p(x)\le 1,\ \ \forall x\in X$

したがって、式(21)のような問題で確率 $p(x)$ は連続変数であるため、もとの問題が離散的であっても連続変数の最適化問題になっている。このように、最適化手法を確率化することの最大の利点の一つは、離散変数あるいは組み合わせ最適化問題を近似的に連続的な最適化問題に置き換えることが可能となる。

一般にSimulated Annealingにおける確率分布は、そのｴﾝﾄﾛﾋﾟｰを最大にするように決められ、その分布の大方は期待値付近に集中する。そこで期待値を最適解近傍に設定すれば最適解が選択される確率が高くなる。したがって、目的関数値が小さい解候補近傍に分布関数の期待値を移動すれば、最適解が得られる確率は増加する。また確率分布は熱平衡(系の温度 T 一定)のもとで一定分布となる(ｴﾝﾄﾛﾋﾟｰの最大)。ここでｴﾝﾄﾛﾋﾟｰの小さい状態とは制約条件のために起きている特別な状態、言い換えれば可能な多くの組合せの中の限られた組合せ(ﾐｸﾛ状態)である。この制約条件を除けば可能な全てのﾐｸﾛ状態を実現する(ﾏｸﾛ状態)ことになりｴﾝﾄﾛﾋﾟｰは増加し、ｴﾝﾄﾛﾋﾟｰが大きい状態は解候補がﾗﾝﾀﾞﾑに選定されることを意味する。ここで確率分布を大きさを決める条件(ｴﾝﾄﾛﾋﾟｰの大小を表す)として系の温度 T、系(確率分布)の熱平衡条件が重要となる。

一方、著者はｻﾝﾌﾟﾙ値により目的関数の確率分布を最尤法で求め、ﾏﾙｺﾌ過程により部分集合を漸化的に推定するｲﾝﾎﾟｰﾀﾝｽ･ｻﾝﾌﾟﾘﾝｸﾞ法[8]を提案している。

この手法では解候補を集合と考え、その期待値(平均値)を最適解と定義する。また探索の手順の概

要は、まず解候補集合からﾗﾝﾀﾞﾑに解候補をｻﾝﾌﾟﾙし各目的関数を算出する。次に得られた目的関数の中で制約条件を満足するｻﾝﾌﾟﾙで構成される部分集合が最適解に近い集合である。この操作を繰り返し最適解により近い期待値を有しかつ分散が小さい部分集合を漸化的に推定してゆくものである。したがって、この手法で各ｻﾝﾌﾟﾙは一つ前の状態におけるｻﾝﾌﾟﾙ分布に依存するﾏﾙｺﾌ過程となる。

ここでｲﾝﾎﾟｰﾀﾝｽ･ｻﾝﾌﾟﾘﾝｸﾞにおける条件付き確率をﾏﾙｺﾌ過程で表すと以下となる。

$$\pi_{t+1} = \pi_t \cdot P \tag{22}$$

π_t :時刻tにおける状態確率分布$(1 \cdot N)$

$\pi_0 = [y_1, y_2, \cdots, y_N]$, N:解候補数

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

実際のｻﾝﾌﾟﾙより、確率率分布の最大値(P_{ij}:未知ﾊﾟﾗﾒｰﾀ)を対数尤度関数で求めると以下となる。

$$L(P_{y0,y1} \cdot P_{y1,y2} \cdots P_{yk-1,yk}) = \prod_{i=1}^{N}\prod_{j=1}^{N} P_{ij}^{n_{ij}} \tag{23}$$

n_{ij} :状態iからjへの推移回数y_0:Given(ｻﾝﾌﾟﾙ)

$$\log L = \sum_{i=1}^{N}\sum_{j=1}^{N} n_{ij} \log P_{ij} \tag{24}$$

$\log L$:対数尤度関数

上式は制約条件付きの$\log L(F)$最大値問題となり、ｻﾝﾌﾟﾙ(推移)個数n_{ij}から推移確率行列P_{ij}の推定ができる。

ここで、この式(22)を条件付きｴﾝﾄﾛﾋﾟｰで表すと以下となり、状態π_kにおける発生確率(確率分布)は与えられた条件(分布幅の限定)のもとでｴﾝﾄﾛﾋﾟｰを最大にするように決められるという「ｴﾝﾄﾛﾋﾟｰ最大の原理」に基づいている。

$$n_i H(\pi_k | \pi_{k-1}) = -n_i \sum_{i=1}^{N}\sum_{j=1}^{N} n_{ij} \log p_{ij} = -n_i \log L \tag{25}$$

$$n_i = \sum_{j=1}^{N} n_{ij}$$

3．結　論

本報告は確率論を用いた最適化手法の基本的な考え方とその特徴の考察を試みた。確率論を用いた最適化手法はそのｱﾙｺﾞﾘｽﾞﾑに確率的要素を取り入れることにより、逆問題や組み合わせ最適化問題の解法をある程度効率化していると言える。

参考 文献

1) 桜井春輔・岸本修治・藤田修一・末廣匡基:斜面掘削工事における情報化施工管理について,第23回岩盤力学に関するｼﾝﾎﾟｼﾞｳﾑ,pp.207-211,1991.
2) 須藤敦史・星谷勝:EK-WLI法と有限要素法を用いた逆解析,土木学会論文集,No.466/Ⅰ-19,pp.177-185,1992.
3) 斉藤知秀・三上 隆・須藤敦史:拡張ｶﾙﾏﾝﾌｨﾙﾀを用いたﾄﾝﾈﾙ地山の熱定数の予測,土木学会北海道支部論文報告集,第51号,Ⅰ-23,pp.110-115,1995.
4) D.G.Krige: Two-dimensional Weighted Moving Averaging Trend Surfaces for Ore Evaluation, Proc. of Sym. on Math. Stat. and Comput. Appl, Johannesburg, South Africa,pp.13-38,1966.
5) J.P.Delhomme: Kriging in the Hydoroscience, Advance in Water Resources,Vol.1,No.5,pp.251-266,1978.
6) 小宮謙一・山本欣弥・星谷勝:地震時の地盤変位の条件付ｼﾐｭﾚｰｼｮﾝ,土木学会第51回年次学術講演概要集,第1部(B),pp.222-223,1996.
7) E.Aarts and J.Korst: Simulated Annealing and Boltzmann Machines - A Stochastic Approach to Combinatorial Optimization and Neural Computing, John Wiley,1989.
8) 須藤敦史・星谷 勝・宮沢和樹:遺伝的要素を考慮したｲﾝﾎﾟｰﾀﾝｽ･ｻﾝﾌﾟﾘﾝｸﾞによる離散型変数を有するｼｽﾃﾑの最適化,土木学会論文集,No.519/Ⅰ-32,pp.223-232,1995.