

情報エントロピーによる逆解析のための 観測計画の評価方法に関する基礎的考察

本城勇介¹・工藤暢章²

¹正会員 岐阜大学 工学部 土木工学科 助教授 (〒505-11 岐阜市柳戸1-1)

²学生員 岐阜大学 工学部 土木工学科 博士課程

地盤工学や構造工学で逆解析を成功させるためには、逆解析の目的に応じた観測データの適切な取得が必要であると考えられる。本研究では、逆解析における事前情報、観測情報、予測の過程における情報の流れを、情報理論で広く用いられるシャノンの情報エントロピーにより記述し、これを規準として逆解析を行なうことを前提とした場合の、観測計画の評価を行う方法の基本的な考察を行った。この方法と、従来から実験計画法の分野で提唱されてきた観測計画の評価基準との関係についても明確化した。提案された方法は、比較的簡単な例題に適用され、その有用性が示された。

Key Words: inverse analysis, entropy, information theory, observation planning, geotechnical engineering

1. 序論

地盤工学や構造工学における多くの設計や施工の問題で、技術者は過去の類似構造物における経験、事前の試験や調査、種々の観測結果やそれに基づく予測値などいろいろな種類の情報を得、これにより種々の意思決定を行っている。「情報化施工」等と呼ばれる所以である。

しかし、このような文脈で言われる情報という言葉は極めてあいまいなものであり、従って情報が多い、少ない、また信頼できる、信頼できないなどと言っても、それらはきわめて漠然とした主観的な評価であると言わなければならない。

本論文は、このような地盤工学の設計や施工で言われている情報というものをいくぶんなりとも定量的に表現し、それらの取得や伝達について考察したものである。すなわち本研究では、問題を観測された結果の逆解析と、これに基づいた予測として定式化し、この関係を情報という観点からモデル化し、最終的には必要な情報を制約の中で最大限度得ることの出来る、最適な観測計画を求めることを試みている。さらに、2つの簡単な計算例を加え、提案した方法の理解を助けるようにした。

情報量の定量的な表現としては、シャノンにより定義された情報エントロピー（以下では、ただ「エントロピー」と呼ぶ）を用いた。この概念は、情報

通信の理論では基本的な概念であり、20世紀後半のこの分野の発展に莫大な貢献をした^{1), 2), 3)}。

逆解析における観測計画の問題は、Sunにより論じられている⁴⁾。そこでは、実験計画法で提案された種々の評価規準の適用が論じられている。この点については本論文の3(5)で詳しく述べる。Shoji 他⁵⁾は軟弱地盤上の盛土の弾塑性解析で、求めようとするパラメータの推定感度と計器の埋設位置の関係を検討した。吉田らは、推定パラメータの推定不確実性を小さくすることを目的とした評価規準の提案に始まり、情報エントロピーが評価規準となりうることを述べている^{6), 7)}。しかし、この提案は本論文で述べているほど体系的な提案ではない。一方、地盤定数の推定の規準に情報エントロピーを用いることは、Bakerらにより提案されている⁸⁾。しかしそこでは、逆解析は前提にされていない。また、Chapman⁹⁾は、水文学で扱う時系列データの不確実性をエントロピーで計量することを提案しており、この計量値が無次元量で、普遍性の高いものであることなど、興味深い議論を展開している。

本論文は以下に示すように構成されている。

まず2では、情報エントロピーの定義とその含意について説明した後、離散的な状態量と連続的な状態量についてのエントロピーの意味の違いについて述べ、最後に相互情報量という重要な概念を定義する。

3では、我々が問題としている観測、パラメータ推定、予測という一連の流れをモデル化し、そこにおける情報の流れをエントロピーにより定量化することを試みる。この部分が、本論文の中心をなす部分である。さらにこのモデル化を利用して、我々が意思決定のために知りたい情報エントロピーを最適化するような観測計画の決定で用いることのできる指標について論じる。この章ではさらに、ここで提案しているエントロピーによる観測計画の評価方法と、従来からいろいろな文献で見られる評価方法との関係等について議論する。

4では、前章で展開した理論を例題に適用し、結果を示した後、考察を加えた。まず、モデルパラメータ、予測値ともに1つの変数からなる場合について考察した。さらに、軟弱地盤上の盛土の問題を模擬した簡単な例題を提示し、提案した方法の有効性を例証することを試みた。

5は、結論である。

なお本論文で用いる対数記号はすべて底を e とした自然対数である。しかし4の計算例の結果を表示した図では、解釈を容易にするため底2の対数表示を用いているので注意されたい。

2. エントロピーと相互情報量

(1) エントロピー

情報を表す一つの本質は、我々が関心を持っているある事象の不確実性を減少させてくれるものということが出来るであろう。例えば今、軟弱地盤上に高さ3mの盛土をある施工計画に従って建設しようというとき、その結果が「破壊する」あるいは「破壊しない」という2通りであったとき、この2つの内一つになることを確実に教えてくれるものがあれば、これを我々は情報ということが出来る。

それでは、個々の情報の重要性を定量的に測る尺度はどのように考えれば良いだろうか。情報通信理論の確立者であるシャノンは、上述の盛土の問題を例に取ると、これを次のように考えた^{1) 2) 3)}。今2つのケースを考えよう。ケース1では、盛土が破壊する確率と、破壊しない確率が半々、すなわち{0.5, 0.5}であるとする。一方ケース2では、この確率が、{0.1, 0.9}であったとしよう。換言すると、ケース1では盛土の破壊に関しては五分五分と予想されているが、ケース2ではほとんど壊れないと予想されている。この様な状況で、「盛土は壊れる」という確実な情報もたらされたとき、それはどちらのケースにおいてより重要な情報と判断されるであろうか。それは、壊れる可能性のほとんどな

いと考えられていた、ケース2の場合であると考えられる。

上記のことを定量的に書き表す方法として、情報量 $I(a_i)$ は次のように、定義される。今ある確率事象の全事象を A 、その個々の事象を a_i 、それぞれの事象の生起する確率を p_i とする。事象 a_i が確実に生じると告げる情報の情報量 $I(a_i)$ は、

$$I(a_i) = \log \frac{1}{p_i} = -\log p_i \quad (1)$$

と表すことにする。この定義に従い上記の個々の場合の情報量を計算してみると、

1) ケース1

$$I(\text{盛土安定}) = -\log 0.5 = +0.693$$

$$I(\text{盛土破壊}) = -\log 0.5 = +0.693$$

2) ケース2

$$I(\text{盛土安定}) = -\log 0.9 = +0.105$$

$$I(\text{盛土破壊}) = -\log 0.1 = +2.303$$

従って、情報量はケース2「盛土破壊」、ケース1の両方の場合、ケース2「盛土安定」の順となり、予想される生起確率の低いものに対する確実な生起の情報ほど、情報量は大きい。この事実は、われわれの情報の重要性に関する直感と一致するであろう。

ところで、実際の場面では、個々の事象の生起に関する情報を得るのではなく、個々の事象の生起確率を与えられた上で、全体としてこのような事象を持つ情報量を評価し、より情報量の多い状態に進もうと意図するのである。このような事象全体についての平均的な情報量を表す指標がエントロピーであり、それは次式のように定義される：

$$H(A) = \sum_{i=1}^n -p_i \log p_i \quad (2)$$

すなわち、エントロピーは情報量の期待値である。先のケースについてエントロピーを計算すると：

1) ケース1

$$H(A) = -0.5 \log(0.5) - 0.5 \log(0.5) = 0.693$$

2) ケース2

$$H(A) = -0.1 \log(0.1) - 0.9 \log(0.9) = 0.325$$

エントロピーは、確率が事象間に等分布しているほど高い。したがって、エントロピーが高いほどその事象に関する情報は曖昧である。ちなみに、上記例題で生起確率が{1.0, 0.0}のときのエントロピー

は0である(このとき $0 \log 0 = 0$ と定義する)。

ところで、エントロピーを式(2)のように定義するのは、シャノンが事象に関する情報量の尺度としてのエントロピーを公理的に構成した結果である。すなわち、正規性、展開性、決定性、強加法性、最大性、連続性と言った、情報を記述する上で都合がよいと考えられる性質を満たす関数形は、シャノンが定義したエントロピー関数しか無いことが証明されている(有本⁹⁾; pp. 62-69)。

(2) 離散量と連続量のエントロピー

前項で述べたような確率事象が離散量として定義されているときのエントロピーは、式(2)によって定義されるが、確率事象が変形や圧力と言った連続量で表されるときのエントロピーの計算には、注意が必要である。本項ではこの点を議論する(小沢²⁾; pp. 117-121)。なお本章では一般的な習慣に従って、確率変数やベクトルを大文字で、その実現値を小文字で示した。また \mathfrak{B} で用いる記号とは定義が異なるので注意されたい。

今式(2)にならぬ、連続確率変数 X を n 個の微小区間 Δx に分けて、エントロピーを書くと：

$$\begin{aligned} H(X) &= - \sum_{i=1}^n p(x_i) \Delta x \log(p(x_i) \Delta x) \\ &= - \sum_{i=1}^n \{p(x_i) \log p(x_i)\} \Delta x \\ &\quad - \left(\sum_{i=1}^n p(x_i) \Delta x \right) \log \Delta x \\ &= - \sum_{i=1}^n \{p(x_i) \log p(x_i)\} \Delta x - \log \Delta x \end{aligned} \quad (3)$$

この式で、 $\Delta x \rightarrow 0$ とすると、第1項は有限値に収束するが、第2項は無限大に発散する。そこで連続量に関するエントロピーは、第1項のみで定義する。第1項は $\Delta x \rightarrow 0$ のとき積分の定義となるので：

$$H(X) = - \int_{-\infty}^{\infty} p(x) \log p(x) dx \quad (4)$$

以上の議論より分かるように、連続量のエントロピーは離散量のエントロピーと異なり、情報量の期

待値をある基準値から測った相対的な量であることが分かる。この基準値とは、第2項の無限大のことである。

$H(X)$ が第2項によって、無限大に発散するのは、任意の x の値を言い当てるためには、 x の変域を二つに分けて、そのどちらかを指定するという二者択一を、仮に無限回繰り返したとしても x の連続性の故に正確な答えを得られる保証がないことに対応している。すなわち連続的信号の持つ情報量は全体として本質的に無限大ではあるが、そのほとんどは第2項という確率密度関数 $p(x)$ に無関係な部分によって占められている²⁾。このような連続量に関するエントロピーの性質のため、エントロピーの有用性は、連続量では離散量に比べ劣っている。しかし次節でとりあげる相互情報量は、この問題のある程度克服している^{3), 10)}。

我々がこの論文で考察の対象とするのは、連続量に関する情報であるので、今後は連続量に関するエントロピーを用いて議論を進める。

(3) 相互情報量

確率変数ベクトル X と Y があり、その同時確率密度関数を $p(x, y)$ 、 X と Y それぞれの周辺確率密度関数を $p(x)$ と $p(y)$ とする。このとき、次の量を相互情報量と呼ぶ(相互情報量とエントロピーの関係については式(8)以下の説明を参照されたい)：

$$I(X; Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (5)$$

条件付き確率に関する公式 $p(x, y) = p(x|y)p(y)$ を上式に適用すると、次のようになる：

$$\begin{aligned} I(X; Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \log \frac{p(x|y)}{p(x)} dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \log p(x|y) dx dy \\ &\quad - \int_{-\infty}^{\infty} p(x) \log p(x) dx \end{aligned} \quad (6)$$

ここで条件付きエントロピー $H(X|Y)$ を次式により定義する：

$$\begin{aligned}
 H(X|Y) &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \log p(x|y) dx dy \\
 &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{p(x|y) \log p(x|y)\} p(y) dx dy
 \end{aligned}
 \tag{7}$$

先のエントロピーの定義と、この条件付きエントロピーを用いると、相互情報量は次のようになる：

$$I(X; Y) = H(X) - H(X|Y) \tag{8}$$

$H(X)$ をベイズ推定における X に関する事前情報のエントロピー、 $H(X|Y)$ を Y を観測した後の X の情報(すなわち事後分布)のエントロピーと解釈することが出来るので、結局相互情報量 $I(X; Y)$ は、観測 Y を得ることにより X についての不確実性が減少する量をエントロピーにより計測したものであることになる(有本¹⁾; PP. 87-95)。いうまでもなく、相互情報量が大きいほど、観測により獲得される情報は大きく、 X の不確実性は減少する。

さらに、相互情報量は3つの確率変数ベクトルを含む場合に拡張が可能であり、その一つの場合は次のようになる(有本¹⁾; pp. 89-90)：

$$I(X; Y|Z) = H(X|Z) - H(X|YZ) \tag{9}$$

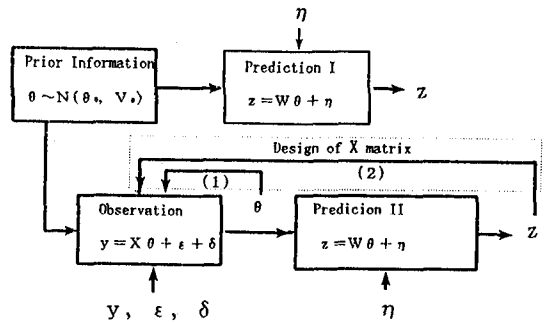
相互情報量は、観測・推定・予測の一連のプロセスを考えるこの研究では、中心的な働きをするようになる。

3. 相互情報量による観測計画の評価

(1) 事前情報・観測・予測のモデル化

本研究では、構造物の設計や施工の作業の中での、情報の流れをモデル化するために、Fig. 1に示すような作業のフローを考える。

建設に入る前の設計に必要なパラメータ・ベクトル θ については、事前の試験や調査、また過去の類似構造物についての経験に基づき情報が得られる。一般の設計作業では、この時点で θ について確実な情報を得ることが重要であり、調査や試験の実施に伴い、 θ の不確実性を確率分布で表現することが出来れば、これよりエントロピーを計算し、この



- (1) Design X based on Model Parameter Mutual Information
- (2) Design X based on Prediction Mutual Information

Fig. 1 Information Flow and Observation Plan

過程での情報量の増加の様子を記述することができる。一方本研究のねらいは、このプロセスではなく、施工中の観測結果を用いて設計パラメータ θ の値を更新して行くときの情報量の増加と、これを効率的に行う観測方法にあるので、事前情報 θ は、ある確率分布により表される不確実性をもって存在するだけ仮定し、議論をすすめる。

この事前情報として与えられる θ を用いて、最終的に我々が予測したい構造物の挙動ベクトル z を求めることが出来る(Fig. 1のPrediction I)。これが一般に設計計算とか、設計のための解析と呼ばれる作業である。このとき θ から、構造物の挙動 z を予測する式を、予測方程式と呼ぶことにする。このとき用いられる事前情報 θ や、これに基づいて予測される z の不確実性は、本研究で議論しようとしている、観測を通じての情報量の増加を考えると、基準となるものである。

θ は、施工中の観測によってもその値が逐次更新される。このとき観測値ベクトル y と設計パラメータベクトル θ を関係付ける式を、観測方程式と呼ぶ(Fig. 1のobservation)。この観測方程式の関係を用いて、 θ の推定値の更新が行われる。構造物の挙動 z の予測は、観測に基づいて推定された θ の値を用いることにより精度を高めることが出来ると考えられる(Fig. 1のPrediction II)。

本章では、まず前章で説明した情報量に関するエントロピーを用いて、Fig. 1の下半分に示したプロセスの情報量の増加の様子を記述する。このとき、2種類の情報量の変化を記述する相互情報量を、観測計画の評価指標として提案した。

一つは、モデルパラメータ θ についての情報量に

着目する場合である。この場合、情報量の増加は、事前パラメータの不確実性に関するエントロピーを基準として、このエントロピーを観測の結果どの程度低減することができたかということにより計測できる。すなわち、この2つの状態の相互情報量を計算する。この定式化を「モデルパラメータ相互情報量」と呼ぶことにし、(3)で詳しく議論する。

情報量の増加を記述するもう一つの方法は、構造物の挙動の予測値 z に関するエントロピーの減少を見るものである。我々が最終的に知りたいのは、当該構造物のある挙動の予測値である。例えば造成地盤で知りたいのは、その最終沈下量であろう。従って、事前情報 θ に基づいた z の予測値のエントロピーと、観測により更新された θ に基づく z の予測値のエントロピーを比較することにより、不確実性の減少を計量できる。これも z に関する相互情報量を求めることになる。この定式化を「予測値相互情報量」と呼び、詳細は(4)項で示す。

この様にプロセスを通じての情報の不確実性の減少を定量化するとき、「モデルパラメータ相互情報量」や「予測値相互情報量」を用いて、不確実性の減少が最も大きくできるような観測計画を考えることができるはずである。これは、上で定義した相互情報量を最大にするように、観測方程式を設定することであり、具体的には線形の問題では観測方程式の中の観測行列 X を設計することである。この観測計画の最適化が本研究の目的である。

以下、次項では観測、予測などの関係が全て線形関係で与えられる場合の基本となる方程式や仮定についてまとめた。(3)では、モデルパラメータ相互情報量について、(4)では、予測値相互情報量について述べ、(5)では相互情報量を観測計画の評価に用いることの利点についてより具体的に考える。最後に(65)では、従来から提案されている観測計画の最適化のための指標と、本研究で提案しているエントロピーに基づく規準の関係について議論している。具体的な計算例は4で示す。

(2)基本モデルと基本仮定

すべてが線形関係で結ばれる場合の各要素は次のような線形方程式で表される。

1)事前情報(Prior Information)

$$\theta \sim N(\theta_0, V_\theta) \quad (10)$$

ここに θ_0 : θ の事前平均値
 V_θ : θ の事前共分散行列
 2)観測方程式(observation equation)

$$y = X\theta + \delta + \epsilon \quad (11)$$

ここに y : 観測値ベクトル (n 次)
 X : 観測行列 ($n \times m$)
 θ : モデルパラメータ (m 次)
 δ : モデル化誤差 (n 次)
 ϵ : 観測値の観測誤差 (n 次)

実際の問題では、モデル化誤差と観測誤差を分離することはできない場合が多いが、ここでの議論は概念的なものなので、一応分離して記述した。

3)予測方程式(prediction equation)

$$z = W\theta + \eta \quad (12)$$

ここに z : 予測値ベクトル (k 次)
 W : 予測行列 ($k \times m$)
 η : モデル化誤差 (k 次)

先にも述べたように、ここでは簡単のためすべてが線形関係で結ばれていると仮定している。観測方程式と予測方程式は実際には同じ現象をモデル化する場合が多いので、 δ と η は類似の統計的な性質を持つと考えられる。

最後に、上で定義したプロセスに導入したいろいろな確率変数ベクトルに関する基本的な仮定を列挙する。導入した確率変数ベクトル θ , δ , ϵ , η はすべて正規確率変数ベクトルであると仮定している。その上で、次のような仮定を導入する：

$$\begin{aligned} E[\theta] &= \theta_0, & E[(\theta - \theta_0)(\theta - \theta_0)^T] &= V_\theta \\ E[\delta] &= 0, & E[\delta\delta^T] &= V_\delta \\ E[\epsilon] &= 0, & E[\epsilon\epsilon^T] &= V_\epsilon \\ E[\eta] &= 0, & E[\eta\eta^T] &= V_\eta \\ E[(\theta - \theta_0)\delta^T] &= 0, & E[(\theta - \theta_0)\epsilon^T] &= 0, \\ E[(\theta - \theta_0)\eta^T] &= 0, \\ E[\delta\epsilon^T] &= 0, & E[\epsilon\eta^T] &= 0, & E[\eta\delta^T] &= 0, \end{aligned} \quad (13)$$

このような仮定は、システム同定等の問題では、標準的なものであると考えられる¹¹⁾。

(3) モデルパラメータ相互情報量

この項及び次の項で、我々はFig. 1に示したプロセスで現れる情報量の増加を前節で定義した相互情報量により記述することを試みる。すなわち、事前情報で得られているモデルパラメータ θ に関する情報量が、観測を通じてどの様に増加するかを相互情報量により記述する。

観測を通じての θ の相互情報量は、式(8)の定義に従い、次のように表すことができる：

$$I(\theta; y) = H(\theta) - H(\theta|y) \quad (14)$$

相互情報量を計算する方法は、 θ と y の同時分布から直接計算する方法と、右辺の2つのエントロピーを計算し、その差から求める方法と2通りあるが、ここでは後者の方法をとることにするが、両者はもちろん一致する。 θ の事前情報のエントロピーは、これが正規分布に従うことを考慮すると、次式のように計算できる^{2), 4)}：

$$H(\theta) = \frac{m}{2}(1 + \log 2\pi) + \frac{1}{2} \log |V_\theta| \quad (15)$$

一方、観測値 y により更新された θ の分布は、周知のようにベイズの定理を用いて、 θ の y に関する条件付き確率として求められ、そのエントロピーは次のように計算される：

$$H(\theta|y) = \frac{m}{2}(1 + \log 2\pi) + \frac{1}{2} \log |(V_\theta^{-1} + X^T(V_\epsilon + V_\delta)^{-1}X)^{-1}| \quad (16)$$

よって、もでのパラメータ相互情報量は次式のようになる：

$$I(\theta, y) = H(\theta) - H(\theta|y) = \frac{1}{2} \log \frac{|V_\theta|}{|(V_\theta^{-1} + X^T(V_\epsilon + V_\delta)^{-1}X)^{-1}|} \quad (17)$$

ここで、事前情報 θ の共分散行列 V_θ と、観測値の誤差とモデル化誤差の共分散行列 V_ϵ と V_δ は、既知で

あり、観測行列 X はある程度設計者が動かすことで出来る行列であるので、この相互情報量を用いて観測計画の評価を行うことが出来る。

(4) 予測値相互情報量

事前情報により得られている θ に基づいた予測値 z のエントロピーと、観測により更新された θ に基づいた予測値 z のエントロピーを比較することにより、この二つの間の相互情報量を求め、この間の情報量の増加を記述する。この関係の相互情報量は式(9)で示した3つの確率変量の間で定義される相互情報量により計算される：

$$I(z; y|\theta) = H(z|\theta) - H(z|y, \theta) \quad (18)$$

この式の右辺第1項は、事前情報 θ のみに基づいた挙動 z の予測値のエントロピーを表し、第2項は観測値 y を考慮することにより更新された θ により予測された z のエントロピーを示す。相互情報量は、観測の結果、予測値 z について低減された不確実性を表している。

予測値相互情報量は、相互情報量の定義に基づき z 、 y と θ の結合確率密度関数からも計算されるし、事前情報 θ に基づく予測値 z の確率密度関数によるエントロピーと、観測値 y により更新された θ に基づいた z の推定値に関するエントロピーとの差を計算することによっても求められる。当然両者は一致する。この結果は、次のようになる：

$$I(z; y|\theta) = \frac{1}{2} \log \frac{|WV_\theta W^T + V_\eta|}{|W(V_\theta^{-1} + X^T(V_\epsilon + V_\delta)^{-1}X)^{-1}W^T + V_\eta|} \quad (19)$$

(5) 相互情報量による観測計画の評価

相互情報量により観測計画を評価する場合、与えられた代替的な観測計画について、モデルパラメータ相互情報量あるいは予測値相互情報量を計算し、その値の大小により、観測計画の善し悪しを判断することになる。いうまでもなく大きな相互情報量を与える計画が良い計画である。このとき、相互情報量を使うことには次のような利点がある：

1) 式(17)や(19)から分かるように、相互情報量は事前と事後の共分散行列の行列式の比である。共分散行列は正定符号行列だから、正值のみよりなる

m個の固有値に分解され：

$$V_0 = U^T \Lambda_m U \quad (20)$$

ここに Λ_m は、m個の固有値を対角要素にもつ対角行列で、Uはそれぞれの固有値に対応した固有ベクトルを列にもつ行列で、 $U^T U = I$ 。よって、

$$|V_0| = |\Lambda_m| = \prod_{i=1}^m \lambda_i \quad (21)$$

従って、 $|V_0|$ を直交化分散値の積と呼ぶことにする。このとき相互情報量について、次の式が成り立つ：

$$\begin{aligned} \text{(相互情報量)} &= \frac{1}{2} \log \frac{\text{(直交化事前分散値の積)}}{\text{(直交化事後分散値の積)}} \\ &= \frac{m}{2} \log \left(\frac{\text{(直交化事前分散値の積)}}{\text{(直交化事後分散値の積)}} \right)^{1/m} \end{aligned}$$

よって

$$\frac{\text{(相互情報量)}}{m} = \log \left(\frac{\text{(直交化事前標準偏差の積)}}{\text{(直交化事後標準偏差の積)}} \right)^{\frac{1}{2}} \quad (22)$$

従って、相互情報量を推定パラメータ数で除した値は、観測の事前と事後のモデルパラメータ推定値や予測値の標準偏差の比のべき平均の対数に等しい。よって、対数の底を2にとれば、相互情報量をパラメータ数で除した値が1増加するごとに、事前と事後の標準偏差の比はだいたい半分づつになってゆくと考えて良い。4.(1)に示す例題は、推定パラメータ数1の場合であり、この関係は厳密に成り立つ。

2) 2つの観測計画の比較のために計算された相互情報量を考えると、この値をパラメータ数で除した値も上と同様に考えて、それらの差は2つの観測計画の事後標準偏差の比のべき平均の対数値である：

$$\begin{aligned} &\frac{\text{(計画1の相互情報量)} - \text{(計画2の相互情報量)}}{m} \\ &= \log \left(\frac{\text{(計画2の直交化事後標準偏差の積)}}{\text{(計画1の直交化事後標準偏差の積)}} \right)^{\frac{1}{2}} \end{aligned} \quad (23)$$

この場合も、対数の底を2にとっていれば、1)と同様

のことが成り立つ。

(6)従来の観測計画に関する指標

計測計画の最適化に関する研究は、実験計画法の一分野として進められてきた¹²⁾。この方法の基礎にあるのは、式(11)で与えられるような観測方程式より θ を推定するときの、好ましい観測行列Xの構成方法をどのようにすべきかという問題である。

事前情報がない場合、モデルパラメータ θ の推定共分散行列は次のように与えられる。

$$V'_0 = (X^T(V_0 + V_e)^{-1}X)^{-1} \quad (24)$$

となる。ここで、 $V_0 + V_e = I$ ($\sigma_0^2 + \sigma_e^2$) とする。ただし、Iは単位行列である。このとき：

$$V'_0 \propto (X^T X)^{-1} \quad (25)$$

さらに、予測方程式(12)を用いて、予測値の共分散行列は：

$$V'_z \propto W(X^T X)^{-1} W^T \quad (26)$$

実験計画法を基礎に提案されている観測計画の最適化の規準の代表的なものは、上記の2種類の共分散行列を元に、次のようなものがある：

1) D-optimum:

$$\min. |(X^T X)^{-1}| \quad (27)$$

2) A-optimum:

$$\min. \{ \text{trace} (X^T X)^{-1} \} \quad (28)$$

3) E-optimum:

$$\min. \{ (X^T X)^{-1} \text{の最大固有値} \} \quad (29)$$

4) G-optimum

$$\min. [W^T (X^T X)^{-1} W \text{の最大対角要素}] \quad (30)$$

これらのどの規準がより優れているかという議論

は、困難であるが、D-optimumは最も一般的なようである^{4),9)}。

一見して明らかのようにD-optimumは、先に導入したモデルパラメータ相互情報量の最小化とほとんど等価である。むしろこの規準の、実験計画法では明らかでなかった内容の深さを、エントロピーを考えることにより明確にしていると言えるであろう。

4. 計算例

(1)一次元モデルを用いた基礎的考察

ここで、定義したモデルパラメータ相互情報量と予測値相互情報量の基礎的な性質について考察するため、モデルパラメータベクトル、予測値ベクトルとも1次元の場合を取り上げ、考察を加える。

式(10)から(12)に示した基本式を、この場合について書き直すと、次のようになる。

1)事前情報：

$$\theta \sim N(\theta_0, \sigma_\theta^2) \quad (31)$$

2)観測方程式：

$$y = x\theta + \epsilon + \delta \quad (32)$$

$$\epsilon \sim N(0, \sigma_\epsilon^2) \quad \delta \sim N(0, \sigma_\delta^2)$$

3)予測方程式：

$$z = w\theta + \eta \quad (33)$$

$$\eta \sim N(0, \sigma_\eta^2)$$

ここに、 x, w は定数である。

なおここで式(32)や(33)で ϵ, δ, η 等を、 y や z の絶対値と関係しない一定値として仮定していることには、後に計算結果を考察するとき注意をしておく必要がある。

この場合の、モデルパラメータ相互情報量は、式(17)に従い次のように計算される：

$$I(\theta; y) = \frac{1}{2} \log_2 \frac{\sigma_\theta^2}{\left(\frac{1}{\sigma_\theta^2} + \frac{x^2}{\sigma_\epsilon^2 + \sigma_\delta^2} \right)^{-1} + \sigma_\theta^2} \quad (34)$$

$$= \frac{1}{2} \log_2 \left(1 + \frac{\sigma_\theta^2}{\sigma_\epsilon^2 + \sigma_\delta^2} x^2 \right)$$

また予測値相互情報量は、式(19)より：

$$I(z; y|\theta) = \frac{1}{2} \log_2 \frac{w^2 \sigma_\theta^2 + \sigma_\eta^2}{w^2 \left(\frac{1}{\sigma_\theta^2} + \frac{x^2}{\sigma_\epsilon^2 + \sigma_\delta^2} \right)^{-1} + \sigma_\eta^2} \quad (35)$$

$$= \frac{1}{2} \log_2 \frac{1 + w^2 \frac{\sigma_\theta^2}{\sigma_\eta^2}}{1 + \frac{1}{J} w^2 \frac{\sigma_\theta^2}{\sigma_\eta^2}}$$

ただしここに：

$$J = 1 + \frac{\sigma_\theta^2}{\sigma_\epsilon^2 + \sigma_\delta^2} x^2 \quad (36)$$

$$= \exp[2 I(\theta; y)]$$

なおここでは、3(5)で述べたように対数の底を2とおいて、計算結果を示している。従って、Fig. 2やFig. 3における相互情報量1の増加は、標準偏差の比が半分になったことを表す。

$I(\theta; y)$ と $x^2 \sigma_\theta^2 / \sigma_\epsilon^2 + \sigma_\delta^2$ の関係をFig. 2を示した。この結果よりモデルパラメータ相互情報量について次のようなことが言える：

1) σ_ϵ^2 が大きくなると、 $I(\theta; y)$ も大きくなる。これは、相互情報量を計算する基準である事前情報の分散が大きくなれば、計算される相互情報量が相対的に大きくなることを表している。当たり前と言えば、当たりの結果である。

2) $\sigma_\epsilon^2 + \sigma_\delta^2$ が小さくなると、 $I(\theta; y)$ が大きくなる。 $\sigma_\epsilon^2 + \sigma_\delta^2$ は、言わば観測に伴う雑音(noise)であり、これが小さければ相互情報量は大きくなるということである。

3) x^2 が大きくなると、 $I(\theta; y)$ も大きくなる。これは、観測方程式で θ に関する情報が x 倍され、信号が雑音に比べて大きくなることによる相互情報量の増加であると解釈される。以上より分かるように $(x^2 \sigma_\theta^2 / \sigma_\epsilon^2 + \sigma_\delta^2)$ は、信号/雑音比と呼ぶことが出来る。

観測計画 x を設計する立場で考えると、この問題では x の絶対値を可能な限り大きくするような計画が、相互情報量を大きくし、従って好ましい観測計画と言える。なお、観測方程式の形からも分かるように、相互情報量は無次元量である。

次に(31)式に示した、予測値相互情報量について Fig. -3は、 $w^2\sigma_\theta^2/\sigma_\eta^2$ を指標として、 $I(z; y|\theta)$ と $x^2\sigma_\theta^2/(\sigma_\epsilon^2+\sigma_\delta^2)$ の関係をプロットしたものである。 $w^2\sigma_\theta^2/\sigma_\eta^2$ はパラメータの事前分散に基づく予測値の分散と、予測式の不確実性を表す分散値との比を示す。Fig. 3より次のことが言える：

1) $\sigma_\eta^2=0$ 、すなわち予測方程式のモデル化誤差がないとき、 $I(z, y|\theta)$ は、 $I(z|\theta)$ に完全に一致する。従って、観測計画の最適化という観点から考えると、 $I(z|\theta)$ についての最適化が行われれば、 $I(z, y|\theta)$ についても最適化されていることが保証される。さらに、(35)式の形より、 σ_η^2 が0でない場合でも、 $I(z|\theta)$ を最大化する観測計画 x は、同時に $I(z, y|\theta)$ を最大化する。ただし、このことは観測方程式、予測方程式とも線形の場合にのみ成り立つのである。

2) J が大きくなれば、 $I(z, y|\theta)$ も大きくなる。すなわち、パラメータが正確に推定されれば、これに応じて予測値の不確実性も低減する。

3) J 一定として、($J > 1$ を考慮すると)、 w が大きくなると、 $I(z, y|\theta)$ も大きくなる。すなわち、予測方程式の係数 w は、 θ に関する不確実性を拡大する係数であるから、エントロピーの定義の性格上、 w が小さいときの方が予測値の分散は小さく、予測値相互情報量も小さくなる。

4) $w^2\sigma_\theta^2/\sigma_\eta^2$ 比は、事前情報による予測値の分散と、予測式のモデル化誤差の分散の比である。この比が大きいほど、 $I(z, y|\theta)$ は大きい。すなわち、予測式のモデル化誤差が相対的に小さいほど、予測値相互情報量は大きくなる。言いかえると、精度の良い予測モデルを持っているときの方が、パラメータを精度よく推定したとき、予測値の不確実性の低減がより効果的に計られる。

5) さらに、それぞれの $w^2\sigma_\theta^2/\sigma_\eta^2$ の値に応じて、予測値相互情報量には、限界がある。すなわちどのように信号/雑音比 ($x^2\sigma_\theta^2/(\sigma_\epsilon^2+\sigma_\delta^2)$)を大きくしても、推定には限界がある。この限界値は、次式で与えられる：

$$\frac{1}{2} \log_2 \left(1 + \frac{w^2 \sigma_\theta^2}{\sigma_\eta^2} \right) \quad (37)$$

以上の考察は、設定された例題の性格上多少定性的であるが、従来あまり議論されてこなかった、モ

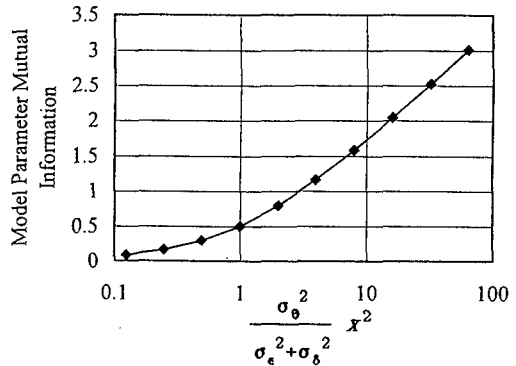


Fig. 2 Model Parameter Mutual Information in one dimensional model

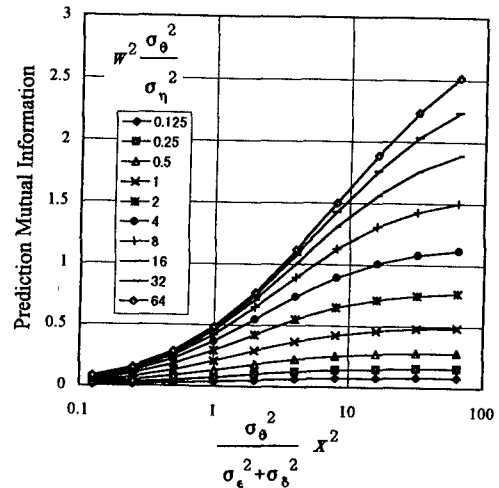


Fig. 3 Prediction Mutual Information in one dimensional model

デル化誤差、観測誤差そして推定誤差の相互関係をかなり明確に説明していると考えられる。

この例題は、観測、予測方程式が1次元線形の場合の結果であるが、方程式が多次元となる一般の場合にも類似の関係が成り立つと考えられる。

(2)線形弾性体地盤の盛土荷重による圧密変形

より現実的な例題として、盛土荷重により圧密変形する線形弾性体地盤における地盤定数推定のための観測計画について考察する。

a)問題の説明

Fig. 4に示すようなヤング率 $E1, E2, E3$ を持つ3層の土層よりなる地盤の盛土荷重による圧密変形問題を考える。排水は、上下面より行われ、また右側も排水境界である。左側は盛土の対称性より非排水境界である。このモデルでは圧密は約100日程度

で90%に達するようにパラメータを設定し(表-1), 載荷は5.1tf/m²を, 10日間で漸加的に行い, その後5日おきに80日まで観測を行う. 載荷開始から, 100日目までのいくつかの点における鉛直変位, 水平変位と間隙水圧の計算結果をFig. 5(a), (b), (c)にそれぞれ示した.

この例題では簡単のため, ヤング率E1, E2, E3のみを推定の対象とし, その他のパラメータ値は既知とする. いろいろな点で鉛直変位, 水平変位や間隙水圧の値を計測し, これよりヤング率を推定しようとする場合のモデルパラメータ相互情報量を計算し, 比較する. 比較は, 次に述べる2つの場合について行った.

Case1: 表-2に示すような計測の量と種類に関する比較を目的としたCase11, 12及び13. 計算の条件となる計測項目の誤差分散値, ヤング係数の事前分散値は表-2の備考欄に示した.

Case2: 計測の種類と位置を比較することを目的とした表-3に示す, Case 21, 22, 23及び24. このケースの計測位置の設定については表-3の備考欄に説明してある.

Case3: 最後に, 予測値相互情報量の例題として上のCase13の場合を取り上げ, このときの中心線上の地盤表面の鉛直変位(y), 盛土法尻の水平変位(x), 中心線上GL. -8mの点の間隙水圧(p)の3つについての予測値相互情報量を, 異なる予測方程式のモデル化誤差の標準偏差の元で計算し, 比較した. 標準偏差は, 表-4の備考欄に示すように, それぞれの基準値に対し3通りの値を設定し, 予測値の標準偏差とした.

なお, 非常に重要な事であるが, 観測方程式と予測方程式が線形で, かつ予測式のモデル化誤差が0の場合, 予測値相互情報量はモデルパラメータ相互情報量と完全に一致する. 一方, この例題の観測方程式, 予測方程式とも今推定の対象としているヤング率に関して非線形ではあるが, これらの式を線形化するためのテーラー展開を, 真値で行っているため, 予測式の感度行列W(式(19)参照)は一定となり, 結果としてモデルパラメータ相互情報量と予測値相互情報量は一致する. この事実のため, 予測値相互情報量の検討はCase3の限られた場合のみ行ったのである.

b) 計算結果と考察

Case1: 計測の量と種類を比較した

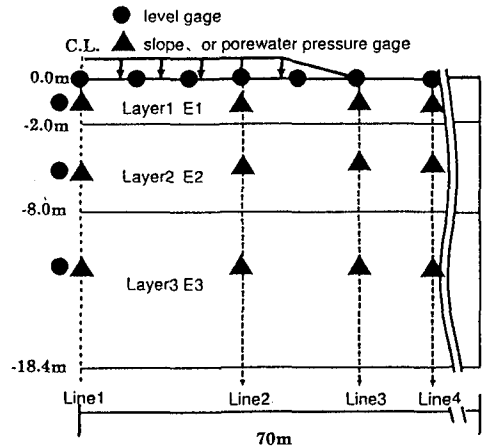
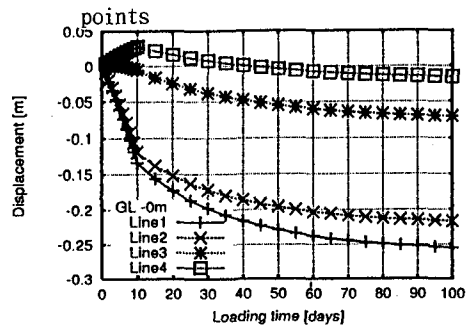
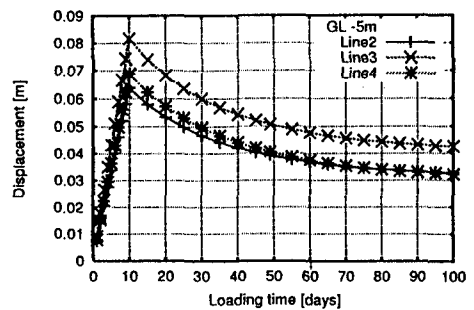


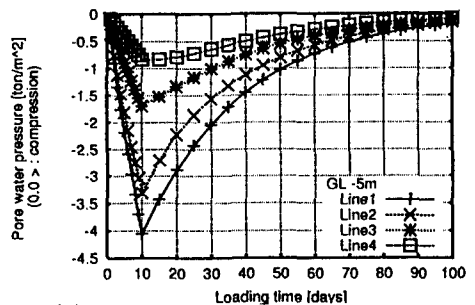
Fig. 4 Linear elastic ground and observation points



(a) vertical displacement change with time



(b) horizontal displacement change with time



(c) pore pressure change with time

Fig. 5 Behavior of the elastic ground during loading

表-1 各層の地盤の物性値

Parameter	Young's Modulus E (tf/m ²)	Poisson's Ratio ν	Permiability k(m/day)
Layer 1	300	0.30	0.003
Layer 2	200	0.30	0.003
Layer 3	400	0.30	0.003

表-2 Case 1の設定条件

Case	計測条件	備考
Case11	・地盤地表面沈下 7点	各計測項目の分散は、鉛直及び水平変位では0.005(m) ² 、間隙水圧0.2(t/m ²) ² である。 各土層のヤング率の事前標準偏差は真値の10%である。
Case12	・地盤地表面沈下 7点 ・盛土中心線上3点の層別鉛直変位	
Case13	・地盤地表面沈下 7点 ・盛土中心線上の層別鉛直変位3点 ・盛土法先の地中水平変位20点	

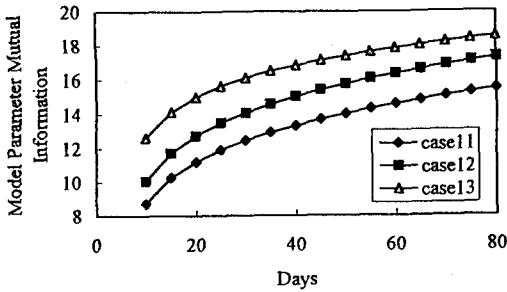
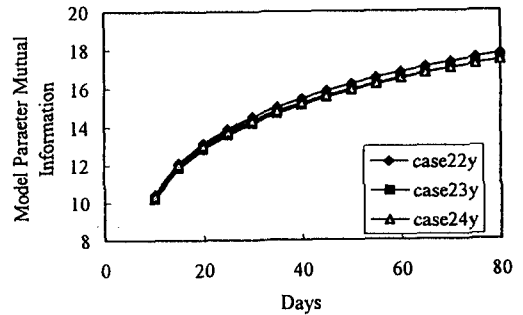


Fig. 6 Model Parameter mutual Information for Case 1

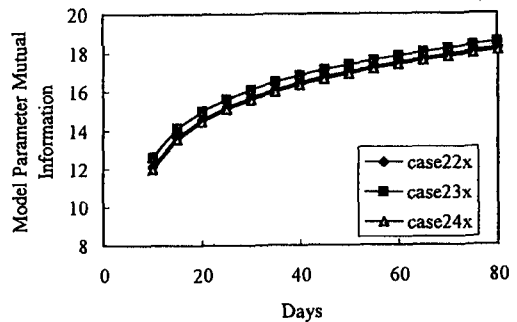
Case11, 12, 13のモデルパラメータ相互情報量の観測量の増加による変化をFig. 6に示した。先に述べたように、対数の底を2とした場合のモデルパラメータ相互情報量をパラメータ数（この場合は3）で除した値の1の増加は、標準偏差が半分になることを意味するので、Case11, 12及び13を比較することにより、計測項目の増加による推定パラメータの不確かがかなり減少することが分かる。さらに時間の経過と共に観測量が増加すると、相互情報量は増

表-3 Case 2の設定条件

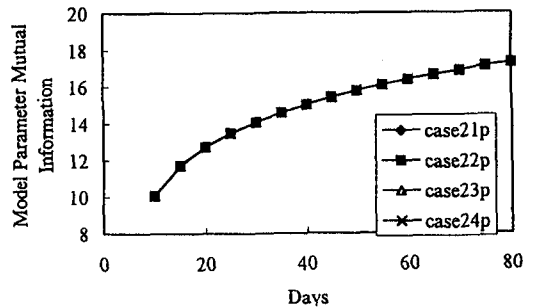
計測位置	水平変位	鉛直変位	間隙水圧
Line 1	-	Case 12	Case 21p
Line 2	Case 22x	Case 22y	Case 22p
Line 3	Case 23x	Case 23y	Case 23p
Line 4	Case 24x	Case 24y	Case 24p
備考	総てのケースで、Case12の計測項目に、それぞれに示したLine上での水平変位、鉛直変位、間隙水圧をそれぞれ加えて計測した。Case12に新たに加えられた計測点数は、鉛直変位と間隙水圧では3個水平変位では傾斜計の性質を考慮し20個である。		



(a) vertical displacement observation



(b) horizontal displacement observation



(c) pore pressure observation

Fig. 7 Model Parameter Mutual Information for Case 2

加しているが、その増加率は時間と共に減少している

Case2: Fig. 7(a), (b), (c)にそれぞれ鉛直変位, 水平変位, 間隙水圧の観測位置がFig. 4に示したLine2, 3及び4と変化したときの, モデルパラメータ相互情報量の変化を示した. 時間の増加と共に情報量が増し, その結果相互情報量が大きくなることは共通している.

ここで重要なのは鉛直変位を計測するなら法尻から少し内側に入ったLine2における相互情報量が大きく, 水平変位ならばちょうど法尻のLine3のとき最大となる. その差は約0.3なので, 推定されるパラメータ値の標準偏差が10%減少する程度である. 一方, 間隙水圧を計測する場合は, どこで計測しても大差はない.

また, これらを比べると水平変位の計測が相互情報量を増すのに最も有効で, 続いて鉛直変位, そして最後に間隙水圧となっていることも興味深い.

先にも述べたように, 予測式のモデル化誤差が0のときは, これらの結果は予測値相互情報量についても全く同一となる.

Case3: Fig. 8に示すように, 設定された条件下では, 予測式の誤差の標準偏差が基準値の10%である場合は, 鉛直変位(case31y), 水平変位(Case31x)及び間隙水圧(Case31p)の予測値相互情報量とも非常に小さく, 観測により予測値の信頼性を事前の情報に基づくものより大きくすることはできない.

一方, 予測誤差が小さくなるに従い, 予測値相互情報量は大きくなり, 0.1%ではかなり大きな相互情報量が得られる. この場合パラメータ数3であるので, 予測値相互情報量3の増加は, 予測値の標準偏差をだいたい半分にしたことを意味する.

予測項目で見ると, 相互情報量の増加が著しいのは鉛直変位であり, 続いて間隙水圧, そして水平変位の順である.

なお, 本論文では予測式のモデル化誤差の分散を, 得られる予測値の絶対的な大きさと無関係に一定値と仮定している. 問題によっては, 予測値の不確実性は予測値の絶対値の大きさに比例する(変動係数一定)といった仮定の方が合理的と思われる場合もあるので, この結果は注意を要する.

5. 結論

本研究では, 構造物の設計・施工過程の, 事前情

表-4 Case 3 の設定条件

Case	予測誤差の設定条件	備考
Case31x Case31y Case31p	水平変位の10% 鉛直変位の10% 間隙水圧の10%	予測は, 中心線上の地盤表面の鉛直変位(x)約0.2m, 盛土法尻の水平変位(y)約0.05m, 中心線上GL. -8mの点の間隙水圧(p)約1.0t/m ² を基準値とし, これらの設定%値を, 予測値の標準偏差とした.
Case32x Case32y Case32p	水平変位の1% 鉛直変位の1% 間隙水圧の1%	
Case33x Case33y Case33p	水平変位の0.1% 鉛直変位の0.1% 間隙水圧の0.1%	

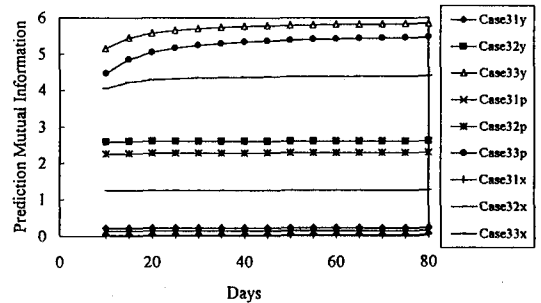


Fig. 8 Prediction Mutual Information for Case 3

報・観測・予測という一連の情報の流れを, 情報エントロピーに概念に基づくモデルパラメータ相互情報量と, 予測値相互情報量という指標で計量することを提案した. そして, これら指標に基づいて観測計画を評価することを提案した. 得られた主要な結論は, 次の通りである:

1) この論文で提案したモデルパラメータ相互情報量と予測値相互情報量は, 得られた情報量によりそれぞれのパラメータや予測値の不確実性がどの程度低減するかを計量する定量的な尺度であり, 相対量であるという欠点はあるが, どのような問題にも同様に適用できる無次元量であるという意味ですぐれた性質を持つ, 観測計画の評価する上で便利な指標である⁹⁾.

2) 相互情報量をパラメータ数で除した値は, 事前と事後の標準偏差の比のべき平均の対数をとったものに等しい. 対数の底を2にとった場合, この値の1の増加は, 標準偏差比のべき平均が半分になった

ことを意味する。

3)モデルパラメータ相互情報量を大きくするには、観測行列 X をうまく設計し、信号／雑音比を高めることが必要である。

4)観測方程式、予測方程式共に線形で、かつ予測方程式のモデル化誤差が0の場合、モデルパラメータ相互情報量と、予測値相互情報量は、完全に一致する。さらに両方程式が非線形でも、本研究の2番目の例題のように、予測方程式の感度行列を固定したモデルパラメータ値でテラー展開により評価する場合は、両者は一致する。(この例題では、行列 X は、常に一定の定数行列である。)

5)設定された仮定条件では、予測値相互情報量には上限値があり、どのように観測行列を改良してもこの上限値を越えることは出来ない。

6)特に2番目の例題を通して、提案された評価指標が、観測の位置や種類を評価する上で有用であることが示された。

7)本研究で提案している評価指標と、従来実験計画法の分野で研究されてきた評価指標との関連が明らかにされた。本研究で提案された評価法の方が、情報全体の流れを明確化した上で提案されている点で、従来の評価指標より優れていると思われる。

本研究の考察はまだ基礎的なものであり、実際の問題に適用してゆくためには、計測あるいは予測した量の価値を経済的に計量し、また観測費用も含めた総合的な最適化を行なう必要がある。本研究は、計測でえられる情報の相対的な大きさを、相互情報量という一つの値に集約して示すことを提案した点に意義がある。

謝辞:本研究の遂行にあたり、貴重なコメントを頂き、また終始暖かい援助を頂いた、岐阜大学工学部教授、宇野尚雄先生、京都大学防災研究所教授、小尻利治先生に深謝の意を表します。

参考文献

- 1) 有本卓：確率・情報・エントロピー、森北出版、1980.
- 2) 小沢一雅：情報理論の基礎、国民科学社、1980.
- 3) 大矢雅則：エントロピー／その様々な顔つき、「エントロピー」、別冊数理科学、pp. 6-15, 1992.
- 4) Sun, N.Z.: Inverse Problems in Groundwater Modeling, Kluwer Academic Publishers, 1994.
- 5) Shoji, M., H. Ohta, K. Arai, T. Matsumoto and T. Takahashi: Two dimensional consolidation back-analysis, Soils and Foundations, Vol. 30, No. 2, pp. 60-78.
- 6) 吉田郁政・豊田耕一・星谷勝：時間更新を考慮した逆解析と観測点配置の評価指標、土木学会論文集, No. 543/I-36, pp. 271-280, 1996.
- 7) 吉田郁政：未知パラメータ空間の基底変換を用いた逆解析、土木学会論文集 I-41/No. 577, pp. 205-215, 1997.
- 8) Baker, R. and D.G. Zeitoun: Soil variability and the maximum entropy principle, Proc. 5th ICASP, pp. 642-649, 1987.
- 9) Chapman, T.G.: Entropy as a measure of hydrologic data uncertainty and model performance, J. of Hydrology, 85 pp. 111-126, 1986.
- 10) 赤池弘次：AICとMDLとBIC、オペレーションズ・リサーチ, pp. 375-378, 1996. 7.
- 11) 中溝高好：信号解析とシステム同定、コロナ社、1988.
- 12) Steinberg, D.M. and W.G. Hunter: Experimental design: review and comment, Technometrics, Vol. 26, No. 2, pp. 71-97, 1984.

(1997. 6. 5受付)

BASIC CONSIDERATION ON EVALUATION OF OBSERVATION PLANNING FOR INVERSE ANALYSIS BASED ON INFORMATION ENTROPY

Yusuke HONJO and Nobuaki KUDO

In order to carry out inverse analyses successfully in geotechnical and structural engineering, it is important to plan observation schemes in an appropriate way. In this study, the information flow among prior information, observation and prediction process is described based on Information Entropy introduced by Shanon which is rather widely used tool in the communication theory; two criteria to evaluate observation schemes are proposed, one is termed model parameter mutual information, and the other prediction mutual information. Some discussions are made to clarify the relationship between the proposed criteria and the conventional observation scheme evaluation criteria developed in the area of experimental design. The proposed criteria are illustrated by two relatively simple but realistic examples.