0266-352X/96/$15.00 + 0.00

Estimation of Parameters in Geotechnical Backanalysis — I. Maximum Likelihood Approach

A. Ledesma, A. Gens & E. E. Alonso

Geotechnical Engineering Department, ETS Ingenieros de Caminos, Technical University of Catalonia (UPC), c/ Gran Capitán s/n, 08034 Barcelona, Spain

ABSTRACT

The estimation of soil and rock parameters based on field instrumentation data is a common procedure in geomechanics. The use of system identification and optimization techniques allows the performance of this type of analyses in a more rational and objective manner. In this paper a probabilistic formulation for the backanalysis problem is presented. The procedure described involves the evaluation of the measurement covariance matrices, which are derived for some geotechnical instruments used in field instrumentation. The algorithm used to solve the mathematical problem of optimization is also presented, as well as its coupling to a finite element code. The algorithm requires the computation of the sensitivity matrix, which can be evaluated "exactly" in terms of the finite element method. Finally, a synthetic example, based on the excavation of a tunnel, is presented in which the elastic modulus E and the K_0 parameter of the material are identified from measured displacements. The effect of the number of measurements and their error structure is also discussed.

INTRODUCTION

Determination of parameters in geotechnical engineering has been traditionally carried out based on the results of laboratory or *in situ* tests [1]. There is merit, however, in using for the same purpose, observations and measurements carried out during the construction of the geotechnical structures themselves. By using field instrumentation results to estimate geotechnical parameters it is possible to take into account the large scale structure of the soil or rock which is outside the possibilities of the other procedures of parameter determination.

In fact, backanalysis of the behaviour of real structures has always been an important part of the best geotechnical engineering practice [2–7]. In recent times, however, the adoption of system identification [8] approaches and optimization techniques has allowed a more systematic and rational approach to this problem.

Although there have been significant contributions in this field using a deterministic point of view, there are important advantages in using a probabilistic approach [9,10]. In this paper, a probabilistic framework based on the concept of maximum likelihood, that represents a conceptual alternative to the more classical Bayesian approach, will be proposed.

The aim of the approach is to compute the parameters that maximize the likelihood of obtaining the field measurements actually observed. The backanalysis problem becomes equivalent to the minimization of an objective function which depends on the values of field observations. Classical function minimization techniques can be applied. In this paper, the constitutive model that defines the behaviour of the material is linear. However, the resulting optimization problem is non-linear so that iterative function minimization procedures must be used. The backanalysis involving non-linear constitutive laws is outside the scope of this paper [11,12].

It is important to note that the process of parameter estimation is carried out in the context of a specified model that includes geometry, boundary conditions and constitutive laws for the materials. Because of the complexity of real engineering situations it is almost unavoidable that a numerical approach is required in practice to define the model. The parameter identification procedure is able to provide information on the structure and capabilities of the model. The topic of model identification, however, is not treated in this contribution [13,14].

The paper presents the basic formulations of the backanalysis problem in a maximum likelihood framework. Appropriate covariance matrices for the field measurement errors for various geotechnical instruments are then computed. Because of the need of defining the geotechnical model numerically, the question of the coupling of the minimization procedure to the finite element method is addressed. Afterwards the reliability of the parameters obtained and its relationships with the measurement errors is discussed. Finally, the performance of the procedure is illustrated by means of a synthetic example.

In a companion paper an application to a real case involving the excavation of a tunnel in overconsolidated materials is described.

BASIC FORMULATION

Assume that a deterministic model relates some unknown parameters, $\mathbf{p}$, and a certain set of variables, $\mathbf{x}$. The measurements are represented by $\mathbf{x}^*$.

Then the differences between measurements and predictions of the model $(\mathbf{x}^* - \mathbf{x})$ are considered as an error, that can be defined in a probabilistic manner.

The best estimation of the parameters is then found by maximizing the likelihood, L, of a hypothesis, $\mathbf{p}$, given some set of error measurements, $(\mathbf{x}^* - \mathbf{x})$. The likelihood of a hypothesis [15, 16] is proportional to the conditional probability of $\mathbf{x}^*$ given a set of parameters $\mathbf{p}$:

$$L = kf(\mathbf{x}^*/\mathbf{p}) \tag{1}$$

where k is a proportionality constant. This formulation has theoretical and conceptual advantages [17]:

— It is not necessary to define the probability of a hypothesis, which has become a controversial concept in probability theory.
— It does not require the model to reproduce the true system exactly [18].
— The model parameters are considered fixed, but uncertain due to lack of information. This allows to introduce in a systematic way the prior information on the parameters available. Examples of this may be seen in Cividini *et al.* [9] and Gens *et al.* [10].

If the model is considered to be correct, differences between field measurements and model predictions are due to an error measurement. Therefore the probability of measuring $\mathbf{x}^*$ given a set of parameters, $\mathbf{p}$, is the probability of reproducing the error measurements $\mathbf{x}^* - \mathbf{x}$. Assuming that probability distribution as multivariate Gaussian, it is possible to write:

$$P(\mathbf{x}^* - \mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{C_x}|}} \exp\left[-\frac{1}{2}(\mathbf{x}^* - \mathbf{x})^t (\mathbf{C_x})^{-1}(\mathbf{x}^* - \mathbf{x}) \right] \tag{2}$$

where

$(\mathbf{x}^* - \mathbf{x})$ = the vector of differences between measured and computed values using a fixed model
$\mathbf{C_x}$ = the measurements covariance matrix, which represents the structure of the error measurements
m = the number of measurements.

$(\ )^t$ is used to indicate a transposed matrix, and the likelihood is now proportional to the value expressed by eqn (2). Maximizing L is equivalent to minimize the support function:

$$S = -2\ln L \tag{3}$$

A. Ledesma et al.

Note that the likelihood is a function on the parameters, taking into account that a model relating state variables **x** and parameters **p** has been specified:

$$\mathbf{x} = M(\mathbf{p}) \tag{4}$$

It follows from the above remarks that

$$S = (\mathbf{x}^* - \mathbf{x})^{\mathrm{t}}\mathbf{C_x}^{-1}(\mathbf{x}^* - \mathbf{x}) + \ln|\mathbf{C_x}| + m\ln(2\pi) - 2\ln k \tag{5}$$

On the assumption that the error structure is fixed, only the first term of eqn (5) has to be minimized, and in this case the objective function becomes:

$$J = (\mathbf{x}^* - \mathbf{x})^{\mathrm{t}}\mathbf{C_x}^{-1}(\mathbf{x}^* - \mathbf{x}) \tag{6}$$

The case in which the error structure is not fixed is outside the scope of this paper, and will be treated elsewhere. If the measurements are independent and its errors have a gaussian distribution of probability with the same variance, the matrix $\mathbf{C_x}$ has the form:

$$\mathbf{C_x} = \sigma^2\mathbf{I} \tag{7}$$

where **I** is the identity matrix. In this case eqn (6) represents a least squares criterion and the objective function is reduced to the simplest expression:

$$J = (\mathbf{x}^* - \mathbf{x})^{\mathrm{t}}(\mathbf{x}^* - \mathbf{x}) \tag{8}$$

If the m measures are obtained from r independent instruments with individual covariance matrices $(\mathbf{C_x})_i$, eqn (6) can be written as

$$J = \sum_{i=1}^{r}(\mathbf{x_i}^* - \mathbf{x_i})^{\mathrm{t}}(\mathbf{C_x})_i^{-1}(\mathbf{x_i}^* - \mathbf{x_i}) \tag{9}$$

which is the most general form of the objective function that will be used in this paper.

COVARIANCE MATRICES

The objective function selected to identify parameters in a geotechnical framework involves the covariance matrix of the measurements, as shown in eqn (6). If an instrument performs independent measurements, then the covariance matrix will be diagonal, as indicated in eqn (7).

However, in some instruments designed for linewise observations [19], the errors of the measurements are not independent. For instance, if an inclinometer device is used to measure horizontal displacements along a borehole, the value of the displacement j (and therefore its error) is based on all the previously measured displacements.

In this section, the covariance matrices for three instruments performing linewise observations — sliding micrometer, inclinometer and deflectometer [19] — are obtained.

In order to do that, it is useful to express the covariance matrix for each instrument as

$$(\mathbf{C_x})_i = \sigma_i^2 (\mathbf{E_x})_i \tag{10}$$

where σ_i^2 is a scale factor which represents the global variance of the measurements made using the instrument i, and $(\mathbf{E_x})_i$ contains the error structure of the instrument which depends on the apparatus itself. Obviously, if the measurements are independent and have the same variance, $\mathbf{E_x}$ is the identity matrix, as shown in eqn (7).

Sliding micrometer

It is assumed that the line over which the measurements are performed is divided in p sections of length l_i. The sliding micrometer will measure the change of length, ϵ_i, between adjacent measuring points. Then, the longitudinal displacement of measurement point n, v_n, will be

$$v_n = \sum_{i=1}^{n} \epsilon_i + A \tag{11}$$

where A is an integration constant. To compute the covariance matrix, it is logical to assume that the primary measurements ϵ_i, are independent, with a variance σ^2. Taken the integration constant A as exactly known (we are interested now in the general structure of $\mathbf{C_x}$), an element of the covariance matrix can be written as

$$(C_x)_{ij} = \operatorname{cov}[v_i, v_j] = \operatorname{cov}\left[\sum_{m=1}^{i} \epsilon_i, \sum_{n=1}^{j} \epsilon_j \right] = \sum_{m=1}^{i} \sum_{n=1}^{j} \operatorname{cov}(\epsilon_i, \epsilon_j)$$

$$= \sigma^2 \sum_{m=1}^{i} \sum_{n=1}^{j} \delta_{mn} = \sigma^2 \min(i, j) \tag{12}$$

and the corresponding element of the matrix $\mathbf{E_x}$:

$$(E_x)_{ij} = \min(i, j) \tag{13}$$

If the variance of the integration constant A is taken into account, then

$$(E_x)_{ij} = \min(i, j) + \left(\frac{\mathrm{var}A}{\sigma^2} \right) \tag{14}$$

Inclinometer

The inclinometer measures slopes α_i in fixed points along a line, and it is used to compute the horizontal displacements along a vertical line. The value α_i is assumed to be small. Then the displacement perpendicular to the line measurement can be computed as

$$u_n = \sum_{i=1}^{n} \alpha_i l_i + B \tag{15}$$

where l_i is the length between two consecutive points of measurement, and B is an integration constant that expresses the horizontal movement of the initial point. Assuming that the value of B is exactly known, the $\mathbf{C_x}$ matrix for an inclinometer will be

$$C_{ij} = \mathrm{cov}\left[u_i, u_j\right] = \mathrm{cov}\left[\sum_{m=1}^{i} \alpha_m l_m, \sum_{n=1}^{j} \alpha_n l_n \right] = \sum_{m=1}^{i} \sum_{n=1}^{j} \mathrm{cov}(\alpha_m, \alpha_n) l_m l_n$$

$$= \sigma^2 \sum_{m=1}^{i} \sum_{n=1}^{j} l_m l_n \delta_{mn} = \sigma^2 \sum_{m=1}^{\min(i,j)} l_m^2 \tag{16}$$

If the integration constant is not fixed, the elements of the $\mathbf{E_x}$ matrix are

$$(E_x)_{ij} = \sum_{m=1}^{\min(i,j)} l_m^2 + \left(\frac{\mathrm{var}B}{\sigma^2} \right) \tag{17}$$

Deflectometer

The deflectometer can measure slope changes κ_i in some points along a line. Then the displacement perpendicular to this line is given by

$$u_n = \sum_{i=1}^{n} \kappa_i l_i \left(\sum_{j=1}^{n} l_j \right) + C \sum_{i=1}^{n} l_i + D \tag{18}$$

where C and D are integration constants. Without these constants, the covariance matrix is in this case:

$$(C_x)_{ij} = \text{cov}[u_i, u_j] = \text{cov}\left[\sum_{m=1}^{i} \kappa_m l_m \left(\sum_{r=m}^{i} l_r\right), \sum_{n=1}^{j} \kappa_n l_n \left(\sum_{s=n}^{j} l_s\right)\right]$$

$$= \sum_{m=1}^{i} \sum_{n=1}^{j} l_m \left(\sum_{r=m}^{i} l_r\right) l_n \left(\sum_{s=n}^{j} l_s\right) \text{cov}[\kappa_m, \kappa_n] = \sigma^2 \sum_{m=1}^{\min(i,j)} l_m^2 \left(\sum_{r=m}^{i} l_r\right) \left(\sum_{s=n}^{j} l_s\right)$$

(19)

and when the integration constants have their own variance,

$$(E_x)_{ij} = \sum_{m=1}^{\min(i,j)} l_m^2 \left(\sum_{r=m}^{i} l_r\right) \left(\sum_{s=n}^{j} l_s\right) + \left(\sum_{k=1}^{i} l_k\right) \left(\sum_{k=1}^{j} l_k\right) \left(\frac{\text{var} C}{\sigma^2}\right) + \left(\frac{\text{var} D}{\sigma^2}\right)$$

(20)

Invariance of J

The covariance matrices described in previous section have been derived assuming that displacements are the variables used in the identification problem. An alternative approach is to choose directly as the basic measurement variables the values of ϵ_i, α_i and κ_i. Whatever approach is used, the objective function to be minimized must remain the same.

Suppose that there is a linear relationship between the two kinds of measurements. For instance, in the inclinometer case:

$$u_i = W_{ij} \alpha_j$$

(21)

where W can be obtained from eqn (15). Assuming, for simplicity, $B = 0$:

$$W_{ij} = \sum_{m=1}^{i} l_m \delta_{mj}$$

(22)

The two objective functions available for the identification problem will be:

$$J_1 = \Delta\mathbf{x}^t \mathbf{C}_{\mathbf{x}}^{-1} \Delta\mathbf{x} \quad ; \quad J_2 = \Delta\boldsymbol{\alpha}^t \left(\frac{1}{\sigma^2}\mathbf{I}\right) \Delta\boldsymbol{\alpha}$$

(23)

where $\Delta\mathbf{x} = (\mathbf{x}^* - \mathbf{x})$ and $\Delta\boldsymbol{\alpha} = (\boldsymbol{\alpha}^* - \boldsymbol{\alpha})$. Using eqn (21):

$$J_2 = \Delta\mathbf{x}^t \left(\sigma^2 \mathbf{W}\mathbf{W}^t\right)^{-1} \Delta\mathbf{x}$$

(24)

and it can be seen that

$$(WW^t)_{ik} = W_{ij}W_{kj} = \sum_{m=1}^{i} \sum_{n=1}^{k} l_m l_n \delta_{mn} \tag{25}$$

which coincides with the measurement covariance matrix for an inclinometer, as indicated in eqn (16).

This proof can be generalized to any other instrument, taking into account the fact that

$$\mathbf{C_x} = \mathrm{cov}[\mathbf{u}, \mathbf{u}] = \mathrm{cov}[\mathbf{W}\alpha, \mathbf{W}\alpha] = \mathbf{WW^t}\mathrm{cov}[\alpha, \alpha] = \mathbf{WW^t}(\sigma^2\mathbf{I}) = \sigma^2\mathbf{WW^t} \tag{26}$$

Hence, $J_1 = J_2$ and the objective function is invariant with respect to using any linear combination of measurements. Usually, displacements are taken as state variables, in order to simplify the coupling of the identification method with a finite element code and, therefore, the appropriate measurement covariance matrix should be used.

NUMERICAL IMPLEMENTATION

Minimization algorithm

The solution of the identification problem defined above requires the minimization of a suitable function. Without losing generality, we can seek the minimum of eqn (6) instead of eqn (9). This can be achieved by means of a wide range of unconstrained optimization algorithms, according to the particular expression of the objective function. In general it is necessary to choose between methods that only need evaluations of the function (i.e. downhill simplex used in [20]) and methods that also require computations of the derivative of that function. In general, algorithms using the derivative of the objective function are expected to be more powerful than those using only the values of the function.

In this work, a Gauss–Newton algorithm has been adopted, because of its good convergence properties when seeking the minimum of functions such as eqn (6), and because the derivatives obtained are also useful in providing information on the reliability of the parameters identified.

Assuming that the relationship between measured variables and parameters [eqn (4)] is in general non-linear, the form of a minimization algorithm will be defined in terms of an iterative procedure:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \Delta\mathbf{p}_k \quad , \quad J(\mathbf{p}_{k+1}) \leqslant J(\mathbf{p}_k) \tag{27}$$

where $\Delta\mathbf{p}_k$ is the advance vector in the parameters space, computed in iteration k using a suitable algorithm.

The Gaus–Newton method [21, 22] is based on an expansion of the objective function into a Taylor series, and gives the value of $\Delta\mathbf{p}$ in each iteration according to:

$$\Delta\mathbf{p} = \left(\mathbf{A}^t\mathbf{C}_x^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^t\mathbf{C}_x^{-1}\Delta\mathbf{x} \qquad (28)$$

where $\Delta\mathbf{x} = (\mathbf{x}^* - \mathbf{x})$ and the matrix

$$\mathbf{A} = \frac{\partial\mathbf{x}}{\partial\mathbf{p}} \qquad (29)$$

is called the sensitivity matrix. If the number of measures is m and the number of parameters is n, the size of this matrix is $m \times n$.

In general Gauss–Newton algorithm tends to exhibit a rapid convergence in this kind of problem, although it can be unstable sometimes. In those cases, an improvement of the algorithm proposed by Levenberg and Marquardt [23] has been used, in which eqn (28) is changed into:

$$\Delta\mathbf{p} = \left(\mathbf{A}^t\mathbf{C}_x^{-1}\mathbf{A} + \mu\mathbf{I}\right)^{-1}\mathbf{A}^t\mathbf{C}_x^{-1}\Delta\mathbf{x} \qquad (30)$$

where μ is an arbitrary real number and $\mathbf{I}$ is the identity matrix. If $\mu \to 0$, the Gauss–Newton procedure is obtained. As μ increases, the parameters' correction provided by eqn (30) becomes smaller and tends towards the gradient direction of the objective function.

The algorithm has to decide on the value of μ in each iteration, depending on the behaviour of the procedure. If the value of J becomes smaller, μ decreases, reaching 0 (or a very small value) at the minimum. However, if J increases, μ is also increased until a smaller value of the objective function is obtained. There is not a common procedure for evaluation of μ and sometimes trial identifications must be performed in advance. In the examples of this paper, an initial value $\mu_i = 10$ was used and for the next values the following criteria was employed: $\mu_{n+1} = \mu_n * 10$ if $J_{n+1} > J_n$, otherwise $\mu_{n+1} = \mu_n/10$, where n is the number of iterations.

Coupling to the finite element method

A general identification procedure for geomechanical problems should be associated with a numerical model relating measurements and parameters. In

this approach, the finite element method has been coupled to the identification techniques described above.

We assume that the measured values are nodal displacements. The process can be generalized to the case in which relative displacements between two arbitrary points are considered.

The finite element method formulated in terms of displacements gives a linear system of equations:

$$\mathbf{Kx} = \mathbf{f} \tag{31}$$

where $\mathbf{K}$ is the global stiffness matrix, $\mathbf{x}$ the vector of nodal displacements, and $\mathbf{f}$ the nodal forces vector. The $\mathbf{K}$ matrix can be computed as

$$\mathbf{K} = \int_V \mathbf{B^t DB} \, \mathrm{d}V \tag{32}$$

where $\mathbf{B}$ is the geometry matrix relating strains and nodal displacements and $\mathbf{D}$ contains the constitutive law as a relationship between stresses and strains. The vector of nodal forces is defined as

$$\mathbf{f} = \int_S \mathbf{N^t} \sigma \, \mathrm{d}S \tag{33}$$

where $\mathbf{N}$ is the shape function matrix, and σ, for an excavation problem, is the vector of stresses acting on the excavation boundary, S.

The displacements can be computed from eqn (31) and then can be compared with the values measured, in order to check the identification process.

Using the finite element method and the Gauss–Newton algorithm, the procedure used to estimate parameters is as follows:

(a) Assume a set of initial parameters $\mathbf{p_0}$.
(b) Calculate the stiffness matrix $\mathbf{K}$.
(c) Solve the direct problem and compute displacements $\mathbf{x}$, $\mathbf{x} = \mathbf{K^{-1}f}$.
(d) Calculate the differences between the measured and the computed values, $\mathbf{x} - \mathbf{x^*}$.
(e) Evaluate a new set of parameters $\mathbf{p}$ by means of eqn (28).
(f) Check convergence. If $\Delta\mathbf{p}$ is still large or the norm of $(\mathbf{x} - \mathbf{x^*})$ has not stabilized, the procedure starts again from step (b).

Computation of the sensitivity matrix

Note that computation of $\mathbf{A} = \partial\mathbf{x}/\partial\mathbf{p}$ is needed in step (e) — eqn (28). One possibility for determining it is to apply a finite differences technique. Alternatively, an "exact" procedure to evaluate the sensitivity matrix, using the

finite element approximation, is presented. It will be shown later that this procedure is also more efficient in terms of computing time.

Deriving eqn (31) with respect to the parameters and rearranging, we obtain:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{p}} = \mathbf{K}^{-1} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} - \frac{\partial \mathbf{K}}{\partial \mathbf{p}} \mathbf{x} \right) \tag{34}$$

It is assumed that the materials involved in the problem are linear isotropic elastic. The parameters to be identified are the Young's modulus E, the Poisson's ratio ν and the coefficient of lateral pressure at rest K_0. For the Young's modulus and the Poisson's ratio [10], $\partial \mathbf{f}/\partial \mathbf{p} = 0$, and

$$\frac{\partial \mathbf{x}}{\partial \mathbf{p}} = \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \mathbf{p}} \mathbf{x} \tag{35}$$

since

$$\frac{\partial \mathbf{K}}{\partial \mathbf{p}} = \int_V \mathbf{B}^t \frac{\partial \mathbf{D}}{\partial \mathbf{p}} \mathbf{B} \mathrm{d} V \tag{36}$$

Hence computing $\partial \mathbf{K}/\partial \mathbf{p}$ is equivalent to find the stiffness matrix substituting the $\partial \mathbf{D}/\partial \mathbf{p}$ matrix for $\mathbf{D}$, which allows an easy implementation in a finite element code.

For instance, in a plane strain problem:

$$\frac{\partial \mathbf{D}}{\partial E} = \frac{1}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1-2\nu}{2} \end{pmatrix} \tag{37a}$$

$$\frac{\partial \mathbf{D}}{\partial \nu} = \frac{E}{(1+\nu^2)(1-2\nu)^2} \begin{pmatrix} 2\nu(2-\nu) & 1+2\nu^2 & 0 \\ 1+2\nu^2 & 2\nu(2-\nu) & 0 \\ 0 & 0 & -\frac{(1-2\nu)}{2} \end{pmatrix} \tag{37b}$$

Note that if ν is assumed known and is not being identified, $\partial \mathbf{D}/\partial E$ is constant and has to be calculated only once.

For the identification of the K_0 parameter, eqn (34) is used, resulting in:

$$\frac{\partial \mathbf{x}}{\partial K_0} = \mathbf{K}^{-1} \frac{\partial \mathbf{f}}{\partial K_0} \tag{38}$$

since for a linear elastic material, $\partial \mathbf{K}/\partial K_0 = 0$.

The excavation is simulated in one phase, by applying on the excavated boundary the opposite nodal forces to the initial stresses. So, using eqn (33), we obtain:

$$\frac{\partial \mathbf{f}}{\partial K_0} = \int_S \mathbf{N}^t \frac{\partial \sigma_0}{\partial K_0} dS = \int_S \mathbf{N}^t \sigma_0^* dS \qquad (39)$$

where $\sigma_0 = (K_0 \sigma_y^0, \sigma_y^0, 0)$, σ_y^0 is the initial vertical stress at depth y and

$$\sigma_0^* = (\sigma_y^0, 0, 0) \qquad (40)$$

Note that to calculate $\partial \mathbf{x}/\partial K_0$, the finite element routines used to find nodal forces due to excavation are employed to compute $\partial \mathbf{f}/\partial K_0$ by changing only the initial stress field from σ_0 to σ_0^*.

This procedure allows us the direct use of the finite element approximation in the computation of the sensitivity matrix, which is expected to reduce the numerical errors. Also it is generally more efficient than a finite differences technique which requires to solve $n + 1$ direct problems to estimate n derivatives, whereas here no additional systems have to be solved if the model is linear elastic.

It should be pointed out that the evaluation of $\mathbf{A}$ in step (e) of the iteration procedure requires to $\mathbf{K}^{-1}$ to be known which has already been computed in (c). Actually, steps (d) to (f) are very fast, especially if the Poisson's ratio ν is not identified (that means, if it is assumed constant), and only the Young's modulus E, and the K_0 parameter are involved in the process. For this particular case,

$$\mathbf{K} = E\mathbf{k} \qquad (41)$$

where the $\mathbf{k}$ matrix depends on the geometry and boundary conditions of the problem, and it is constant in the iterative process. E is the Young's modulus which changes in each iteration. Using eqns (36) and (38), it can be seen:

$$\frac{\partial \mathbf{x}}{\partial E} = \frac{\mathbf{x}}{E} \quad ; \quad \frac{\partial \mathbf{x}}{\partial K_0} = \frac{1}{E} \mathbf{k}^{-1} \frac{\partial \mathbf{f}}{\partial K_0} \qquad (42)$$

where $\mathbf{k}^{-1} \frac{\partial \mathbf{f}}{\partial K_0}$ is constant in each iteration.

RELIABILITY OF THE ESTIMATION

Useful information concerning the parameter identification problem is obtained by factorization of the sensitivity matrix $\mathbf{A}$ by eigenvalues and eigenvectors. This is due to the fact that the rectangular linear system

$$\Delta x = A \Delta p \tag{43}$$

can represent a least squares problem defined by the objective function (8) [24, 25]. A general decomposition of a $m \times n$ matrix can be made in the form

$$A = U \Lambda V^t \tag{44}$$

where Λ is a diagonal matrix containing the k – non-zero eigenvalues (also called singular values) derived from the problem:

$$A v_j = \lambda_j u_j \quad , \quad A^t u_i = \lambda_i v_i \quad \text{(no summation)} \tag{45}$$

where

$$\lambda_i = \lambda_j \ , \quad i = j \ , \quad i \leqslant k \ ; \quad \lambda_i = \lambda_j = 0 \ , \quad i > k, \quad j > k$$

U and V are, respectively, $m \times k$ and $n \times k$ orthogonal matrices containing the measurement-space and the parameter-space eigenvectors corresponding to the non-zero eigenvalues. The "pseudoinverse" matrix of A (in the sense defined above) can be computed using the procedure defined by Lanczos [24] and Lawson and Hanson [25]:

$$H = V \Lambda^{-1} U^t \tag{46}$$

and therefore,

$$\Delta p = H \Delta x = V \Lambda^{-1} U^t \Delta x \tag{47}$$

The eigenvectors can be used for a reparametrization of the problem as follows:

$$\Delta p^+ = V^t \Delta p \quad ; \quad \Delta x^+ = U^t \Delta x \tag{48}$$

Then eqn (47) becomes:

$$\Delta p^+ = \Lambda^{-1} \Delta x^+, \quad \text{i.e. } \Delta p_i^+ = (1/\lambda_i) \Delta x_i^+ \tag{49}$$

Equation (49) shows how a linear combination of the measurements (given by the U eigenvectors) is directly and uniquely related to a linear combination of parameters (given by the V eigenvectors) through the inverse of the corresponding eigenvalues. Furthermore, the value of Λ_i will determine the variance of the corresponding linear combination of parameters. If the measurements are statistically independent, then the variances of the increment of parameters can be directly evaluated:

$$\text{var}(\Delta p_i^+) = \left(\frac{1}{\lambda_i^2}\right)\text{var}(\Delta x_i^+) \tag{50}$$

Therefore

$$\text{var}(\Delta p_i) = \sum_{j=1}^{n}\left(\frac{v_{ij}^2}{\lambda_j^2}\right)\text{var}(\Delta x_j^+) \tag{51}$$

This is, in fact, a lower bound, because of the linearity assumed in the calculation of the variances. Using the decomposition of matrix $\mathbf{A}$ shown in eqn (44), eqn (51) can be written as

$$\text{var}\Delta\mathbf{p} = (\mathbf{A}^t\mathbf{A})^{-1}(\text{var}\Delta\mathbf{x}^+) \tag{52}$$

where $\text{var}\Delta\mathbf{p}$ and $\text{var}\Delta\mathbf{x}$ are vectors of variances of Δp_i and Δx_i, respectively.

The factorization of the sensitivity matrix $\mathbf{A}$ and the results shown above can be generalized if the maximum likelihood criterion is used. In this case matrix $\mathbf{A}'$ instead of $\mathbf{A}$ is used, computed from the factorization:

$$\mathbf{A}^t\mathbf{C_x}^{-1}\mathbf{A} = \mathbf{A}'\mathbf{A}'^t \tag{53}$$

and eqn (52) generalizes to

$$\text{var}\Delta\mathbf{p} = (\mathbf{A}^t\mathbf{C_x}^{-1}\mathbf{A})^{-1}(\text{var}\Delta\mathbf{x}^+) \tag{54}$$

Therefore, in a maximum likelihood approach, the covariance matrix of the identified parameters $\mathbf{C_p}$ can be computed as:

$$\mathbf{C_p} = (\mathbf{A}^t\mathbf{C_x}^{-1}\mathbf{A})^{-1} \tag{55}$$

This matrix agrees with the inverse of the Fisher information matrix [26], and represents the reliability of the parameters estimated in terms of their covariances. It is again a lower bound of the variances, because of the linearity assumed in its computation.

The information density matrix can be defined as $\mathbf{Q} = \mathbf{A}\mathbf{H}$ in the least squares case or $\mathbf{Q} = \mathbf{A}'\mathbf{H}'$ if the generalization to the maximum likelihood criterion is used. The information density matrix represents the interdependency of the observations established by the model and the redundancy of the data. When the number of measurements coincides with the number of parameters, $\mathbf{Q}$ becomes the identity matrix.

Matrices $\mathbf{C_p}$ and $\mathbf{Q}$ are very useful to assess the reliability of the identification result and to examine the relation between measurements and parameters defined by the structure of the model.

SYNTHETIC EXAMPLE

Description of the problem

To illustrate the performance of the formulation presented above, a synthetic problem based on a tunnel excavation case is presented. We suppose that the material is linear elastic, homogeneous and isotropic, with a Poisson's ratio equal to 0.49 in order to simulate undrained conditions and a specific weight of $\gamma = 20$ KN/m^3. The parameters to be identified are the Young's modulus and the K_0 coefficient defined using total stresses. Figure 1 shows the finite element mesh corresponding to this example, and the 12 nodes used as measurement points. Horizontal displacements in points 1 to 7 could represent values obtained by means of an inclinometer device, whereas vertical displacements in points 8 to 12 could represent measurements from a extensometer. Excavation is made in one step, and due to the symmetry, only half of the geometry is considered.
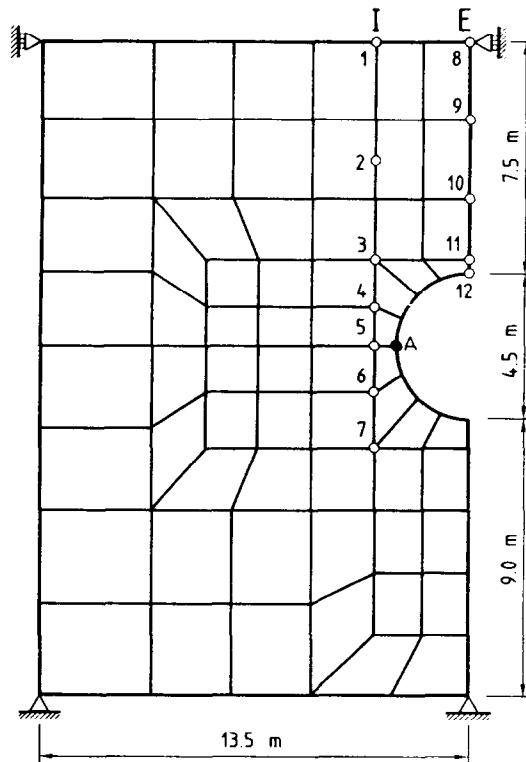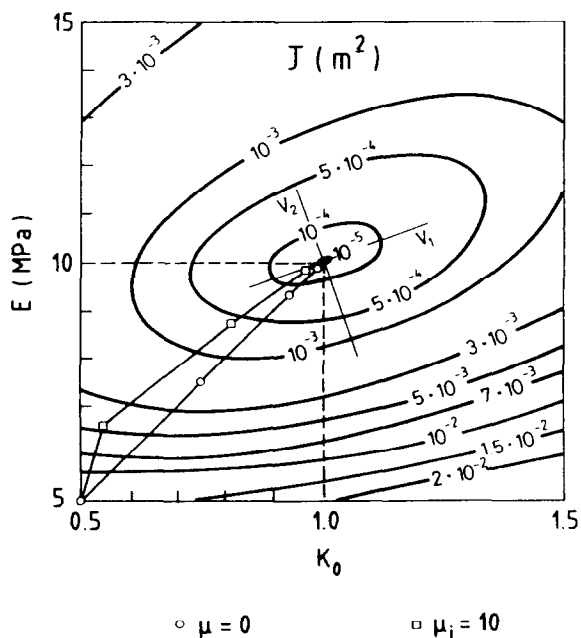

Fig. 1. Finite element mesh of the synthetic example.

For the analysis, the displacements corresponding to the parameters $E = 10$ MPa and $K_0 = 1$ are assumed as input data. The values of these theoretical measurements are presented in Table 1. The identification procedure must be able to obtain those parameters from the information provided by the measurements. As it is a synthetic case, there is no measurement error and the measurements covariance matrix is assumed to be the identity

TABLE 1
Computed Displacements Used as Measurements in the Synthetic Example

Horizontal movements		Vertical movements	
Point	Displacement (cm)	Point	Displacement (cm)
1	0.316	8	−3.598
2	0.556	9	−3.905
3	2.038	10	−4.935
4	3.550	11	−6.327
5	4.550	12	−7.048
6	3.920		
7	2.427		

Fig. 2. Contours of the objective function and paths of the iterative procedure. V_1 and V_2 are the directions of the sensitivity matrix eigenvectors in the minimum.

matrix. The effect of the error structure on the identification process will be considered later.

Starting with the $\mathbf{p_0}$ values: $E = 5$ MPa and $K_0 = 0.5$ the identification procedure has achieved the minimum in 4 or 5 iterations. Figure 2 shows the iterative process on the objective function $J(E,K_0)$, using both, the Gauss–Newton algorithm ($\mu = 0$) and the Levenberg–Marquardt algorithm (initial value of $\mu = 10$).

Figures 3 and 4 show the evolution of the iteration procedure. It can be seen that, in this simple case, there are not too many differences between the algorithms. In general, the Levenberg–Marquardt algorithm is slower but more robust than the Gauss–Newton, especially if the number of parameters to be identified is large.

Finally, Fig. 5 shows the coefficients of the information density matrix of some measurements. The importance of the measurements 5 (maximum horizontal movement) and 12 (maximum vertical movement) is clearly shown,
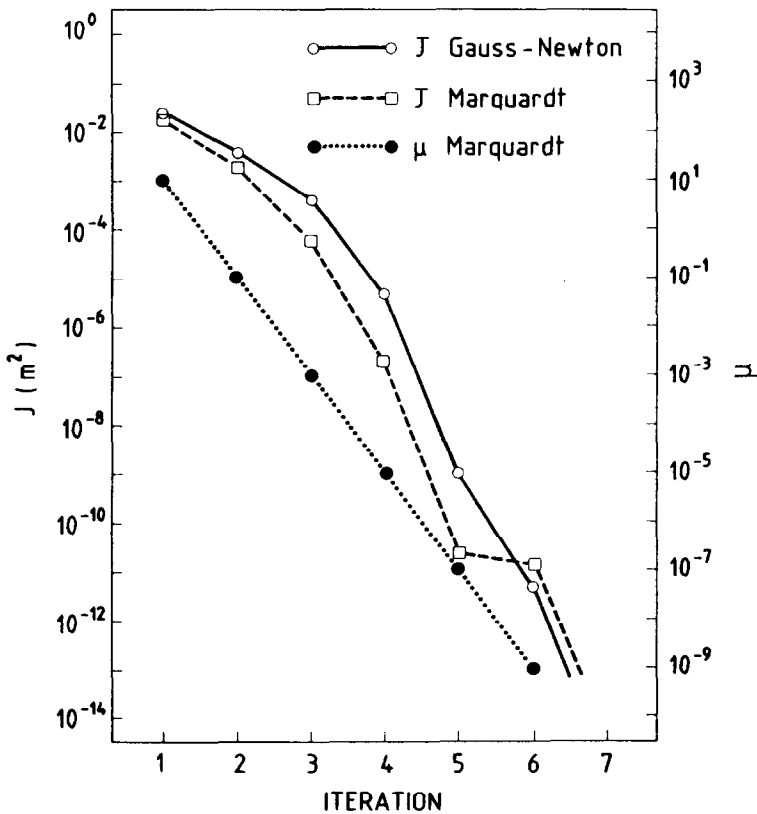

Fig. 3. Evolution of objective function in the iterative process using Gauss–Newton and Marquardt algorithms.
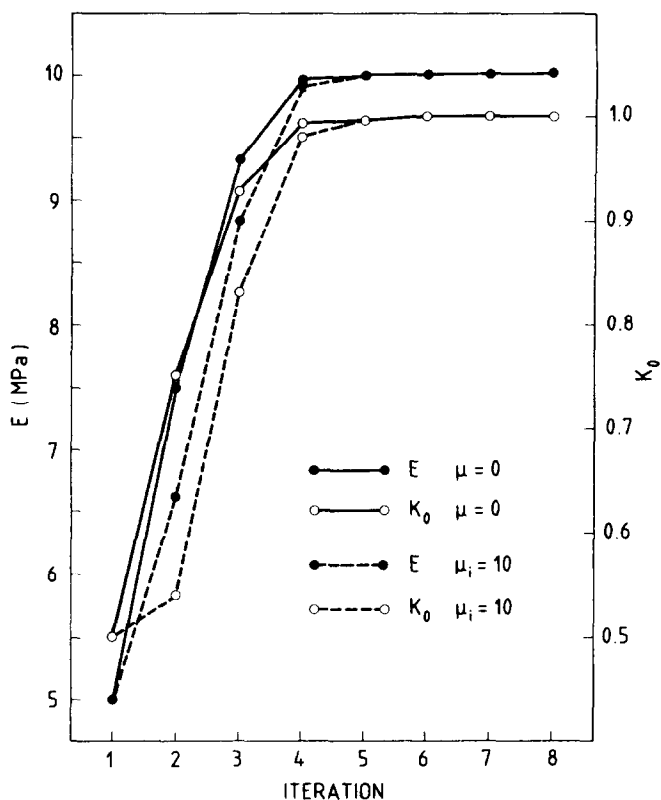
Fig. 4. Evolution of the parameters in the iterative process using Gauss–Newton and Marquardt algorithms.

which means they are the observations that provide more information in this particular case. The study of this matrix can be useful to check which measurements are more important, and at what points the field instrumentation should be installed. It must be noticed, however, that the components of this matrix depend on the value of the parameters. Therefore, it is necessary to have some suitable prior information about the parameters involved in the problem.

Influence of the number of measurements

If the number of measurements in this latest example is increased, we can expect an improvement in the identification results. Obviously that will depend on the kind of information added, i.e. displacements of nodes far from the tunnel will provide little additional information on the problem.

In order to show the effect of the number of measurements involved in the results of the identification procedure, four cases based on the above example
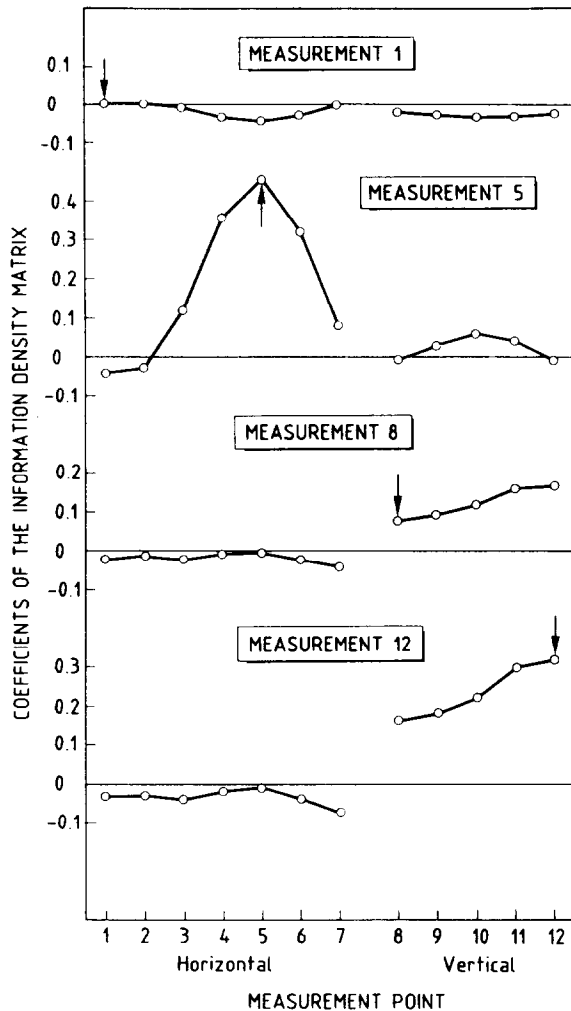
Fig. 5. Coefficients of the information density matrix for some measurements in the minimum of the objective function.

have been considered. All of them use a diagonal measurements covariance matrix:

(a) Only two measurements on the excavation boundary are available: vertical displacement of point 12 and horizontal displacement of point A (Fig. 1). In this case the problem becomes exactly determined, thus only two variables (E and K_0) are unknown.

(b) Twelve measurements are used, corresponding to the example related above.

(c) Twenty four measurements are available, as shown in Fig. 6. Points 1–15 correspond to horizontal displacements and points 16–24 refer to vertical ones.

(d) Fifty five measurements are available: 24 of them are the same as used in case (c), and the rest are horizontal and vertical displacements from points distributed on the excavation boundary.

The estimation of parameters is achieved in a few iterations in all cases. Some differences can be seen in the reliability of the solution for a fixed measurement error, represented by the parameters covariance matrix. As was expected (Fig. 7), the variances of the parameters identified, computed using eqn (55), are smaller when 55 measures are used [case (d)]. The case (a) provides a direct solution (equal number of measurements and unknowns), but the biggest uncertainty on the results.
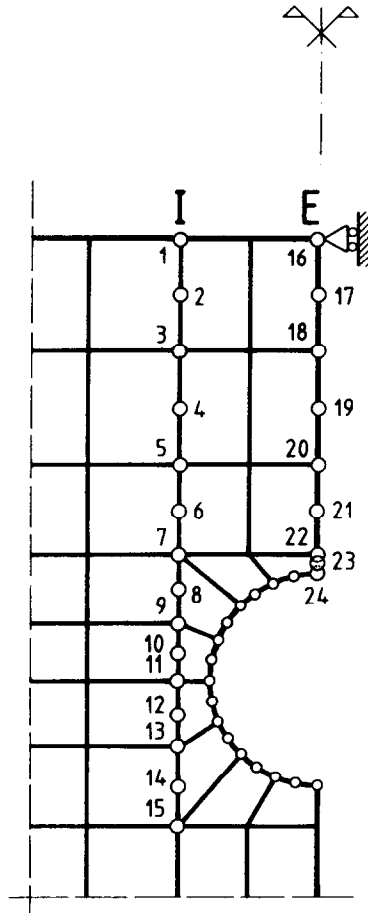

Fig. 6. Position of the measurement points in the synthetic examples (c) and (d).
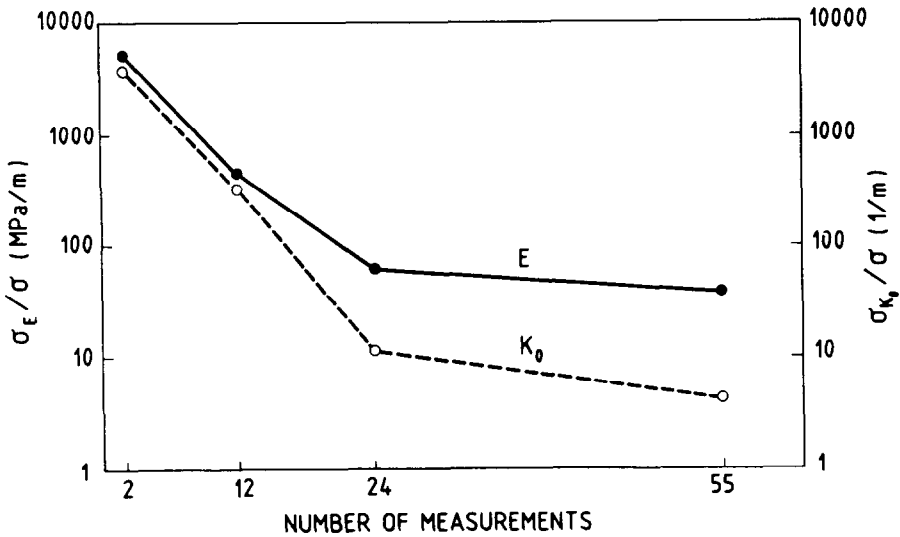
Fig. 7. Standard deviations of E and K_0 for different numbers of measurements used in the analysis. Values are proportional to the measurement's standard deviation.

The coefficients of the information density matrix for measurements A and B have been plotted in Fig. 8(a) and (b). Of course, if only two measures are used, the information is concentrated at points A and B. However, if 55 measurements are considered, the information is distributed along the locations close to the measurement point with a much higher level of inter-dependence between observations.
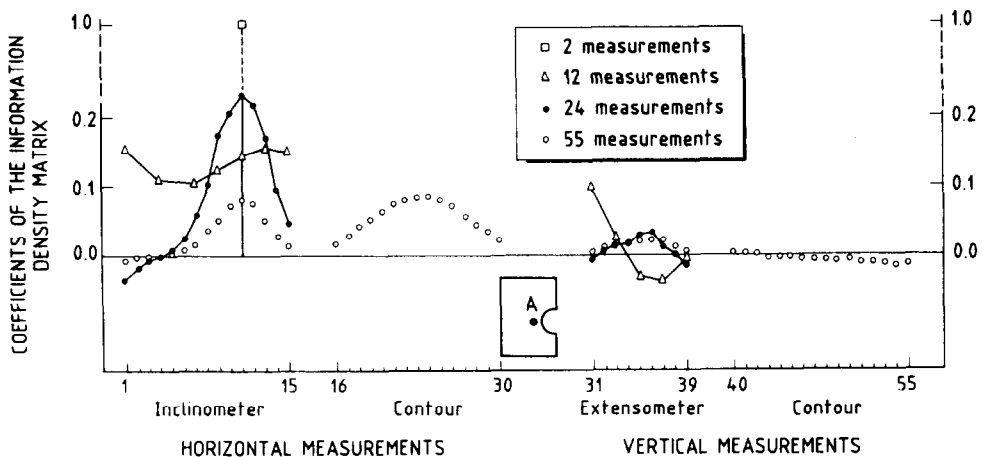

Fig. 8(a). Coefficients of the information density matrix in the minimum corresponding to point 'A', for different numbers of measurements.
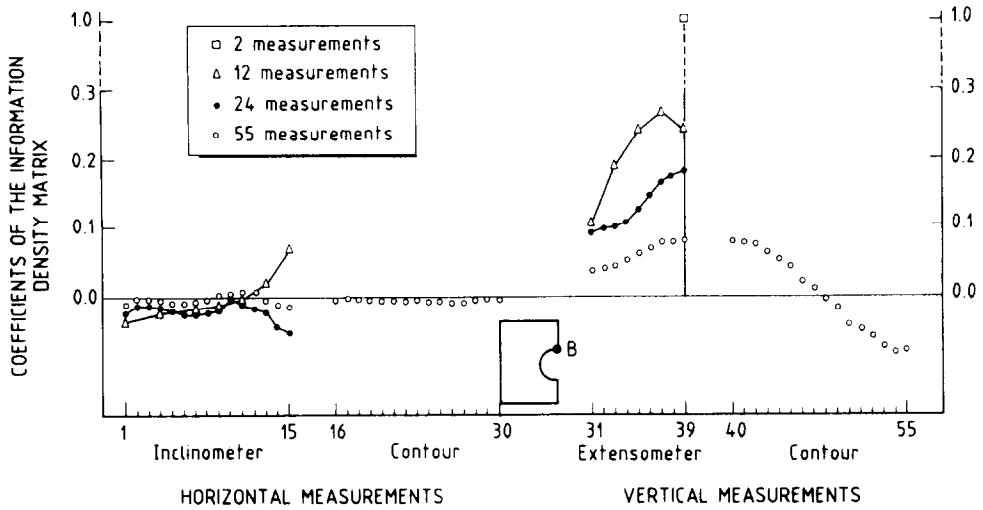
Fig. 8(b). Coefficients of the information density matrix in the minimum corresponding to point 'B', for different numbers of measurements.

Influence of the covariance matrix

The importance of using the complete measurements covariance matrix can be demonstrated by means of this synthetic example. To this end, only a set of 15 horizontal displacements from previous example (c) have been considered. They were perturbed with a noise which was generated taking into account the measurement process of the inclinometer device.

It was assumed that the standard deviation of the slopes measured was $\sigma_\alpha = 0.005$ rad, and that error generated by standard random techniques was propagated to the horizontal displacements.

Two analyses have been carried out using this perturbed input data. The first one uses the full covariance matrix as defined in eqn (17). The second one considers $E_x = I$. As the values used as measurements have the inclinometer error structure, the analysis performed using the full covariance matrix should give (on average) better results.

The parameter estimation process should provide the values $E = 10$MPa, $K_0 = 1$, and the results of the iterative process indicate that the complete covariance matrix provides better results in the parameters identified. In Fig. 9(a) and (b) the objective function using full and identity covariance matrices has been depicted. Note that the minimum obtained using the wrong matrix ($E_x = I$) is far from the actual parameters point. When the full matrix is used, the error structure is well represented and the minimum obtained is close to the real one. Taking into account the formulation described in the
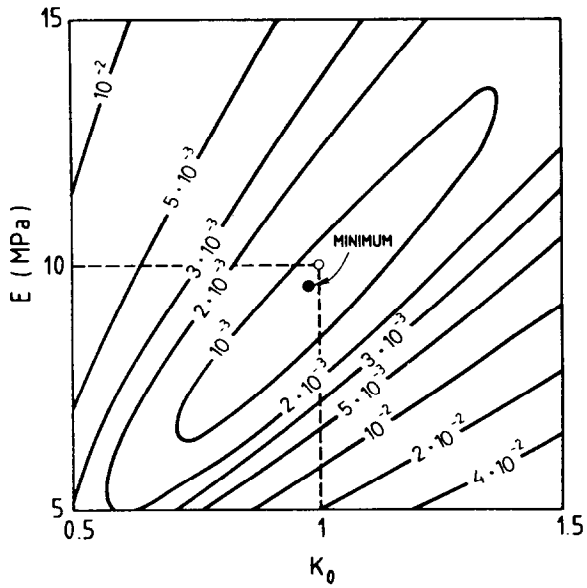
Fig. 9(a). Contour lines of the objective function using the full measurements covariance matrix.
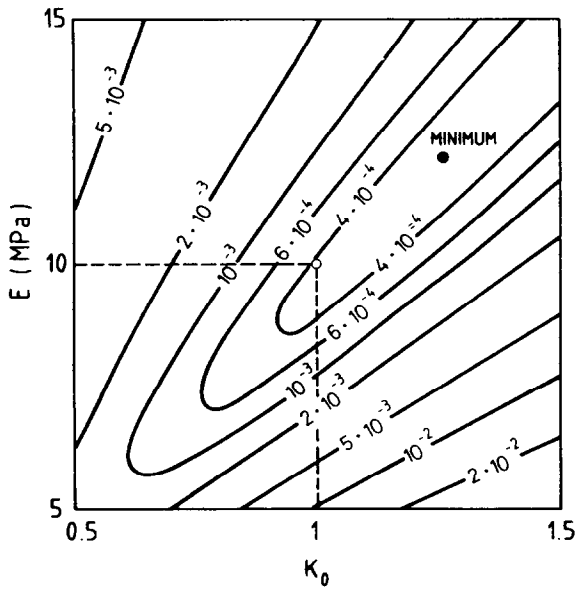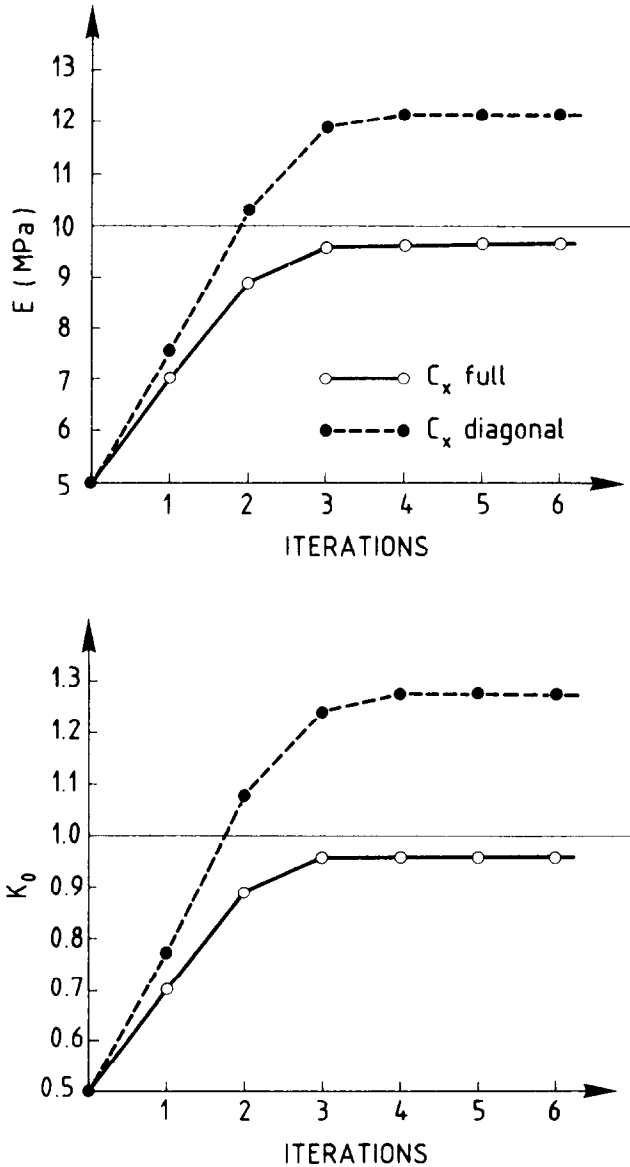

Fig. 9(b). Contour lines of the objective function using the identity measurements covariance matrix.

above sections, this fact should be expected, although only in a probabilistic sense. Figure 10 shows the evolution of the iteration process, starting from $E = 5$ MPa and $K_0 = 0.5$, gives a result in 4 or 5 iterations.

Fig. 10. Evolution of the parameters during the iterative process using full and identity measurements covariance matrices.

In a real backanalysis problem, matrix $\mathbf{C_x}$ can be evaluated using the information available from the measurement devices and the procedure described above. However, an alternative way to compute its value is to introduce it as a new parameter in the identification procedure. This method [10] is outside the scope of this paper and is treated elsewhere.

CONCLUDING REMARKS

The problem of parameter estimation has been formulated in a maximum likelihood framework. This approach has some advantages, i.e. allows the introduction of the error structure of the measurements in a consistent way and provides an estimation of the reliability of the parameters identified. Furthermore, the least squares criterion, commonly used in simple estimation problems, can be considered as a particular case of this formulation.

The probabilistic framework requires the use of the measurements covariance matrix, which has been computed for three instruments used in practical geotechnical problems: sliding micrometer, inclinometer and deflectometer.

This approach results in a mathematical problem of minimization of a suitable objective function, that can be performed by means of any available optimization algorithm. In particular, Gauss–Newton and Levenberg–Marquardt algorithms have been used in this paper, as they have proved to be convenient in these cases.

The identification process requires the computation of the sensitivity matrix $\mathbf{A}$ defined as derivatives of the measured variables (displacements) with respect to the parameters. A procedure to calculate that matrix using the finite element method in linear models is described.

Finally, a synthetic example has illustrated the features of the identification procedure, which provides a consistent framework to solve this kind of problem in a systematic manner, considering the error structure of the measurement process. The effect of the number of measurements and the influence of the covariance matrix on the solution has been examined in this case.

An example, involving the identification of parameters in a real tunnel excavation problem, is described in a companion paper.

REFERENCES

1. Jamiolkowsky, M., Ladd, C.C., Germaine, J.T. & Lancellotta, R., New developments in field and laboratory testing. *11th Int. Conf. SMFE*, San Francisco, CA, Vol. 1, pp. 57–153, 1985.
2. Terzaghi, K. & Peck, R. B., *Soil Mechanics in Engineering Practice*. Wiley, New York, 1967.

3. Maier, G. & Gioda, G., Optimization methods for parametric identification of geotechnical systems. In *Numerical Methods in Geomechanics* (Edited by Martins, J. B.). Nato A.S.I. Series, Reidel, Boston, MA, 1981.
4. Arai, K., Ohta, H. & Kojima, K., Estimation of soil parameters based on monitored movement of subsoil under consolidation. *Soils Foundations*, **24**(4) (1984) 95–108.
5. Gens, A., Ledesma, A. & Alonso, E. E., Maximum likelihood parameter and variance estimation in geotechnical backanalysis. *5th Int. Conf. Appl. Statist. and Probab. in Soil and Struct. Engng*, Vol. 2, pp. 613–621, 1987.
6. Gioda, G. & Sakurai, S., Back analysis procedures for the interpretation of field measurements in Geomechanics. *Int. J. Numer. Meth. Geomech.*, **11** (1987), 555–583.
7. Honjo, Y., Limanhadi, B. & Wen-Tsun, L., Prediction of single pile settlement based on inverse analysis. *Soils Foundations*, **33**(2) (1993) 126–144.
8. Eykhoff, P., *System Identification. Parameter and State Estimation*. Wiley, Chichester, 1974.
9. Cividini, A., Maier, G. & Nappi, A., Parameter estimation of a static geotechnical model using a Bayes' approach. *Int. J. Rock Mech. Mining Sci. Geomech. Abstr.*, **20** (1983) 215–226.
10. Gens, A., Ledesma, A. & Alonso, E. E., Back analysis using prior information. Application to the staged excavation of a cavern. *6th Int. Conf. on Numer. Meth. Geomech.*, Iconmig 88, 1988.
11. Arai, K., Ohta, H. & Kojima, K., Estimation of nonlinear constitutive parameters based on monitored movement of subsoil under consolidation. *Soils Foundations*, **27**(1) (1987), 35–49.
12. Ledesma, A., Identificación de parámetros en Geotecnia. Aplicación a la excavación de túneles. Ph.D. Thesis, Technical University of Catalunya, Spain, 1987.
13. Akaike, M., A new look at statistical model identification. *IEEE Trans. Automat. Control*, **AC-19** (1974), 716–722.
14. Kashyap, R.L., Bayesian comparison of dynamic models. *IEEE Trans. Automat. Control*, **AC-22** (1977), 715–727.
15. Edwards, A.W.F., *Likelihood*. Cambridge University Press, Cambridge, U.K., 1972.
16. Tarantola, A., *Inverse Problem Theory*. Elsevier, Amsterdam, 1987.
17. Carrera, J., State of the art of the inverse problem applied to the flow and solute transport equations. *Nato Res. Work. Adv. in Anal. Numer. Groundwater Flow and Qual. Model*, pp. 549–583. Reidel, Dordrecht, 1988.
18. Baram, Y. & Sandell, N.R., An information theoretic approach to dynamical systems modelling and identification. *IEEE Trans. Automat. Control*, **AC-23** (1978) 61–66.
19. Kovari, K. & Amstad. Ch., Fundamentals of deformation measurements. In *Int. Symp. on Field Meas. in Geomech.*, Zurich, Vol. 1, pp. 219–239, 1983.
20. Gioda, G. & Maier, G., Direct search solution of an inverse problem in elastoplasticity: identification of cohesion, friction angle and in situ stress by pressure tunnel tests. *Int. J. Numer. Meth. Engng*, **15** (1980) 1823–1848.
21. Fletcher, R., *Practical Methods of Optimization. Vol. 1, Unconstrained Optimization. Vol. 2, Constrained Optimization*. Wiley, Chichester, 1981.
22. Scales, L. E., *Introduction to Non-linear Optimization*. Springer, New York, 1986.

23. Marquardt, D. W., An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.*, **11** (1963) 431–441.
24. Lanczos, C., *Linear Differential Operators*. Van Nostrand, London, 1961.
25. Lawson, Ch. L. & Hanson, R. J., *Solving Least Squares Problems*. Prentice–Hall, Englewood Cliffs, NJ, 1974.
26. Bury, K. V., *Statistical Models in Applied Science*. Wiley, New York, 1975.