Data analysis for stress measurements by overcoring : new optimization techniques

Armand Lizeur, Hafid Baroudi, V. Baucheron

DATA ANALYSIS FOR STRESS MEASUREMENTS BY OVERCORING: NEW OPTIMIZATION TECHNIQUES

LIZEUR A., Pr. Eng. INERIS, Laboratoire de Mécanique des Terrains
Ecole des Mines, 54042 Nancy, France.
BAROUDI H., Ph. D. INERIS.
BAUCHERON V., Gr. Student, University Nancy I

ABSTRACT

The data processing for in situ measurements by overcoring (or stress-relief in general) is classically performed by the least squares method. This appears well suited, because the relation $y = F(\sigma)$ between stress tensor (σ) and displacements or strain (Y) is linear. In this case, the unknowns are stresses, and the known data of the problem are the measured displacements or strains and the elastic properties of the rock, those are preferably measured in-situ (biaxial test).

However if we wish to determine the elastic properties (E matrix in the general case) from the overcoring test itself, the previous relation becomes $Y = F(\sigma,E)$. This relation is non linear; to resolve it numerous optimization techniques are available. The optimization methods without gradient calculation seems well adapted for this problem.

INTRODUCTION

The knowledge of natural stress state in a rock mass is often an essential datum to solve most modelling problems in geomechanics. Among the quantitative methods of in situ stress measurements at great depth in a borehole, overcoring is usually preferred (fig. 1). Two different types of measurements are obtained: radial displacements with the "U.S.B.M cell", and strains with the "CSIR triaxial strain cell" or the "CSIRO hollow inclusion gauge" which was developped in Australia (fig. 2).

The determination of the stress tensor is done by assuming the rock as perfectly elastic and homogeneous. We obtain six components of the stress tensor expressed in a fixed reference sytem as follow :

(a) First of all, determine the elastic constants of the rock by means of loading and unloading tests on the core obtained from overcoring or by uniaxial compression tests on rock specimens drilled out in the appropriate directions.

(b) Then, find the equations for the measurements (displacements or strains) in terms of the elastic constants and six stress components.

(c) Invert the equations by using the least squares method or similar methods.

After presenting the classical technique for data analysis we will discuss new numerical methods, that allow the determination of elastic properties and stresses from overcoring data.

TECHNIQUES OF DATA ANALYSIS FOR STRESS MEASUREMENT

The determination of the stresses with in situ overcoring measurement leads to resolve a linear system where the second member is the measurement vector (Y) and the unknowns are the six stress tensor components (σ). The number of linear equations is usually greater than six, which involves that we can't inverse the matrix (A). We use in that case the least squares method wich consists of finding the solution which minimizes the residual error :

$AX = Y$ with X estimator of σ

e = residual error = $Y - AX$

$e^t e$ minimum $\Rightarrow$ $X = (A^t A)^{-1} A^t Y$

This method gives many parameters which allows to estimate the quality of the linear regression such as the probability law of X, estimation of the stress tensor in a fixed reference system. But no information is given on the principal stresses because the relation between these, the six directions and shear stress tensor components is highly non-linear. In order to solve this problem, we used a statistical approach. The probability law of the principal stresses and their orientations is determined by using statistical simulations of the law given by the least squares method.

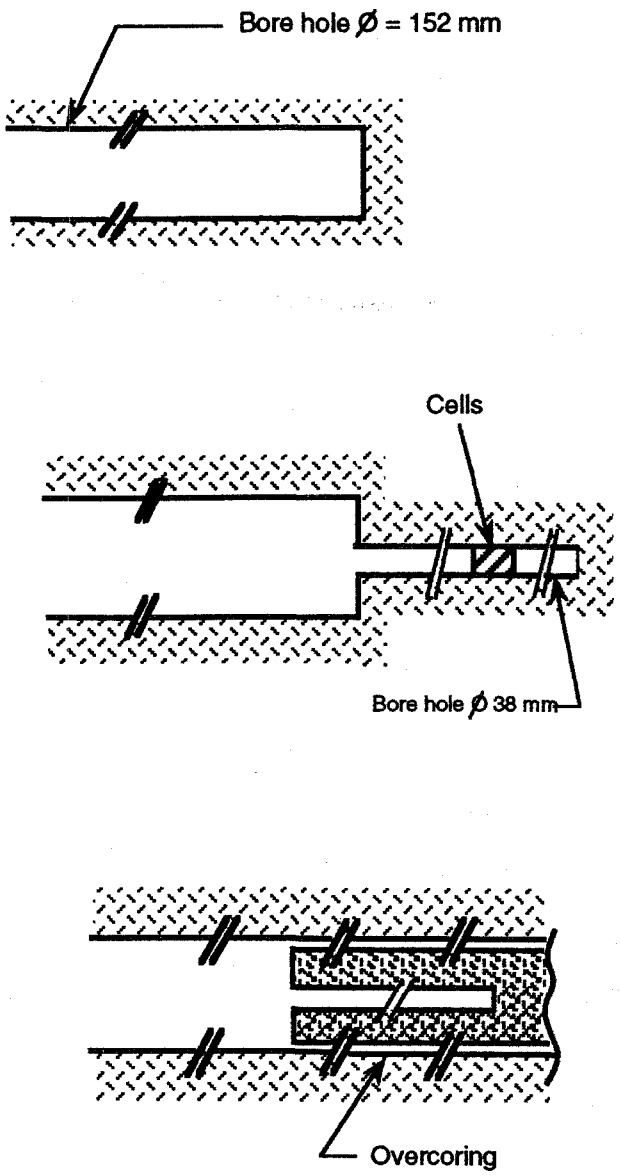

Fig. 1 : Steps of overcoring method

Other similar methods of data analysis for stress measurements like the weighted least squares or the residual values methods [4] could be used but results are almost identical. The probalistic approach developped by Tarantola [8]

and based on the maximum of likelihood gives interesting informations but the results must be carefully interpreted.

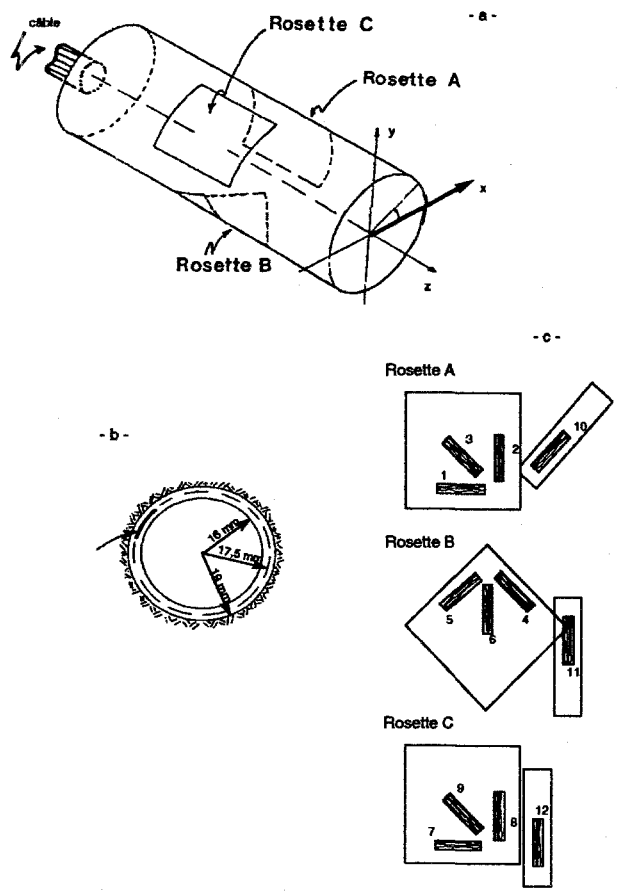

Fig. 2 : CSIRO cell

NEW FORMULATION OF THE PROBLEM

The theory presented here is based on linear elastic behaviour. We suppose that elastic constants and stresses are separate variables this means that the values of the elastic parameters are not depending on the rock stress state. This hypothesis is very important and, in general, implicit when the elastic constants are determined in laboratory.

Classically, these parameters are taken as data of the problem. The disadvantages of this method are that significant variations of the modulus values can occur from point to point in the measurement area, and that the properties of the rock can change between the moments when the in situ measurements are done and the determination of the elastic constants (alteration of the rock, humidity, ...).

The relation between the stress tensor (σ) and the observed measurements (Y_{obs}) is linear. We have :

$Y_{obs} = A.\sigma$ where A is the compliance matrix.

The coefficients of the matrix A depend on the properties of the rock and the orientation of the hole. For special cases, isotropic bodies for example [10], there are known explicitly. In general [1], there are determined from computing. Also, A is a function of the elastic constant vector E.

$$Y_{obs} = A(E).\sigma$$

The idea is to consider E as unknown and determine σ and E so that $Y_{cal} = A(E).\sigma$ the calculated displacements or strains are equivalent to Y_{obs}. This problem can be solved by minimization of the function $f(E,\sigma) = \|Y_{obs} - Y_{cal}\|_p$ where $\|\ .\ \|_p$ designate the norm of the vector $Y_{obs} - Y_{cal}$.

Usually, three vectorial norms are used :

$$\| x \|_1 = \sum_i | x_i |$$

$$\| x \|_2 = \left(\sum_i x_i^2 \right)^{1/2}$$

$$\| x \|_\infty = \max_i | x_i |$$

The dimension of the vector E can be 2,5 or 9 if we consider the rock as isotropic, transversaly isotropic or orthotropic. Also the total number of unknorwns could vary between 8 and 15.

ALGORITHMS OF OPTIMIZATION USED

The classical techniques of optimization using gradients and hessian matrix cannot be used in our case. In general we have no information on the derivability of the function f, that is why we used two algorithms for unconstrained minimization of a function of several variables that does not require the evaluation of partial derivatives : POWELL's and MIFFLIN's algorithm (see Appendix A and B for more details on the methods).

Like many methods of optimization, these algorithms must be initialized. For any results, we must verify that the minimum obtained is not relative and that we have really $f(E,\sigma) = 0$. In our case, there is not only one solution in general. The results given by the two methods are the nearest of the initial values.

The Powell's algorithm [7] is easy to implement. It can be link with the family of methods using conjugate directions. The basic idea is to find the minimum of the function successively in n conjugates directions (n is the number of variables). This method requires at each iteration n one variable minimization. For this we used a golden section method [3,6].

The Mifflin's algorithm [5] is a second order extension of the method of local variations and it does not require any exact one variable minimization. The method retains the local variations property of accumulation points. This extension makes the algorithm an approximate Newton method and its convergence is in general super-linear.

The efficiency in the convergence of the two methods is measured by the number of evaluations of the function F. Mifflin needs in general fewer calculations than Powell. For the same initial values, the results of the algorithms are identical except when there are several solutions in a limited space. In this case, Mifflin gives the result which is the nearest to the initial solution while Powell give a result which could be different.

APPLICATION

We present here an example of use of these two methods. We consider the rock to be isotropic. We used data from the CSIRO HI cell for which the values of the elastic constants and the components of the stress tensor were known. The tensor corresponds to the stress state existing at a depth of 450 m. The minor principal stress σ_3 is almost vertical and its value corresponds to overburden weight (σ_z = 12,1 MPa).

The first test considered was with fixed

Young's modulus and Poisson's ratio. In this case the results obtained with the two methods are exactly identical (table 1) to the exact values, even with a far initial solution. Thus, these methods can be used like the classical least square method when the rock properties are well defined.

Table 2 shows an example where all the parameters are unknown and the initial solution is close to the exact values. In this case, the algorithms converge to the exact values. However, when the initial solution is relatively far from the exact values, the results are physically acceptable but the solution is not good. It shows clearly, that this algorithms must be used carefuly.

If we consider that the principal vertical stress is known (τ_{yz},τ_{xz}, σ_z known), wich is a common assumption in many cases, the results obtained for the unknown parameters are as expected (table 4); the same if ν is fixed and σ_z is unknown (table 5).

parameter	exact values	initial values	MIFFLIN	POWELL
σ_x	18,8	15	18,8	18,8
σ_y	23,1	15	23,1	23,1
σ_z	12,1	15	12,1	12,1
τ_{xy}	2,1	5	2,1	2,1
τ_{yz}	-0,4	5	-0,4	-0,4
τ_{zx}	-1,0	5	-1,0	-1,0
nb evaluations of F			71	1349

Table 1

parameter	exact values	initial values	MIFFLIN	POWELL
E	75000	72000	75734	75309
ν	0,3	0,3	0,30	0,30
σ_x	18,8	18	18,9	19,0
σ_y	23,1	25	23,3	23,4
σ_z	12,1	13	12,2	12,1
τ_{xy}	2,1	3	2,1	2,1
τ_{yz}	-0,4	-0,5	-0,4	-0,4
τ_{zx}	-1,0	-0,5	-1,0	-0,9
nb evaluations of F			778	2189

Table 2

parameter	exact values	initial values	MIFFLIN	POWELL
E	75000	70000	51183	55402
ν	0,3	0,25	0,26	0,28
σ_x	18,8	15	12,3	13,0
σ_y	23,1	15	15,3	15,3
σ_z	12,1	15	7,3	8,3
τ_{xy}	2,1	3	1,3	1,6
τ_{yz}	-0,4	1	0,0	0,2
τ_{zx}	-1,0	1	0,4	0,0
nb evaluations of F			904	2077

Table 3

parameter	exact values	initial values	MIFFLIN	POWELL
E	75000	60000	73865	75001
ν	0,3	0,25	0,30	0,30
σ_x	18,8	15	18,6	18,8
σ_y	23,1	15	22,8	23,1
τ_{xy}	2,1	5	2,1	2,1
nb evaluations of F			387	1247

Table 4

parameter	exact values	initial values	MIFFLIN	POWELL
E	75000	60000	75003	74999
σ_x	18,8	15	18,8	18,8
σ_y	23,1	15	23,1	23,1
σ_z	12,1	15	12,1	12,1
τ_{xy}	2,1	5	2,1	2,1
nb evaluations of F			162	653

Table 5

CONCLUSION AND FORTHCOMING RESEARCH

We tested these two algorithms with data coming from different types of overcoring measurements. We didn't observe differences between results when we changed the norm. Nevertheless the evaluations number of F is less when condidering the euclidean norm.

We are allowed to work with the principal stresses and their orientations (expressed in term of Euler angles) rather than with direction and shear stresses.

The two methods reveal themselves to be very complementary and highly performant when the number of unknowns is less then or equal to six. In the opposite case, the results depend mainly on the initial values, because of the high number of solutions [2].

The different uses of the Powell's and Mifflin's algorithms showed that they fit very well with classic overcoring analysis methods such as least squares method. We now try to ameliorate these methods for future work. At first, the introduction of constraints on the unknowns (such as a variation interval) will allow to better understand the influence of initial values and to direct the parameters research towards given directions. The present approach didn't give any informations about the precision of the results. So we are thinking of a probabilistical approach which will be based on the maximum likelihood : it means that we will work on the probability law of the unknow parameters rather than on their values. Finally, because we want to determine all the solutions of the equation $F(E, \sigma) = 0$, we use a process which combines the Powell's and Mifflin's algorithms with a Monte-Carlo method in order to see more precisely the solutions domains.

REFERENCES

[1] AMADEI B., 1983 - Rock anisotropy and the theory of stress measurements. Lecture notes in Engineering, Brebbia and Orszag. Ed. Springer Verlag.

[2] BAUCHERON V., 1992 - Application de méthodes d'optimisation à la détermination des contraintes et des propriétés géomécaniques d'une roche. Diplôme d'Etudes Supérieures Spécialisées en Ingénierie Mathématique et Outils Informatiques. Université de NANCY I.

[3] BRENT R.P., 1972 - Algorithms for minimization without derivatives. Ed. Prentice Hall, Englewood Cliffs.

[4] CHAMBON C., REVALOR R., 1986 - Statistic analysis applied to rock stress measurements. Internationnal Symposium of Rock Stress and Rock Stress Measurement. Ed. Sentek, Stockholm.

[5] MIFFLIN R., 1974 - A superlinearly convergent algorithm for minimization without evaluation derivates - Mathematical Programming 9, North Holland Publishing Company, pp. 100-117.

[6] MINOUX M., 1983 - Programmation mathématique : Théories et algorithmes. Tome 1. Ed. Dunod, Paris.

[7] **POWELL M.J.D., 1964** - An efficient method for finding the minimum of a function of several variables without calculating derivatives. The computer journal 7 pp.155-162.

[8] TARANTOLA A., 1987 - Inverse problem theory : Methods for data fitting and model parameter estimation. Elsevier Science Publishers.

[9] THEODOR R. , LASCAUX P. , 1986 - Analyse numérique matricielle appliquée à l'art de l'ingénieur. Tome 1. Ed. Masson, Paris.

[10] VAN MEERDEN .L., 1983 - Stress Strain Relations applicable to overcoring techniques in transversaly isotropic rocks. International Journal of Rock Mechanics in Mining Sciences and Geomechanics Abstracts. Vol. 20 - n° 6 pp. 277-282.

APPENDIX A

We describe here the different steps of the Powell's algorithm.

Step 0 : Initialisation : choose a starting point x^0 and n linearly independant directions d_1, d_2,.., d_n, for example the canonic directions.

Step 1 : Determine the series of points x^1, x^2,...,x^n as :

$$F(x^i) = \min F(x^i + \lambda\ d_i) = F(x^{i-1} + \lambda_i\ d_i).$$

This could be done by one variable minimization like a dichotomy or a golden section method.

Step 2 : Define $f_0 = F(x^0)$, $f_1 = F(x^n)$ and $f_2 = F(2x^n - x^0)$. Calculate

$$\Delta = \max\ [\ F(x_i) - F(x_{i-1})\ ].$$

Determine the m index which gives the maximum Δ.

Step 3 : If $f_2 \geq f_0$ or $2(f_0 - 2f_1 + f_2)(f_0 - f_1 - \Delta)^2 \geq \Delta(f_0 - f_2)^2$ then do not change the directions d_1, .. , d_n , set $x^0 = x^n$ and go to step 5.

Step 4 : If $f_2 < f_0$ and $2(f_0 - 2f_1 + f_2)(f_0 - f_1 - \Delta)^2 < \Delta(f_0 - f_2)^2$ then determine the minimum of F in the direction $x^n - x^0$. The obtained point will be taken as new starting point x^0 at the next iteration. Replace the n independant directions in this order :

$$d_1, .. , d_{m-1}, d_{m+1}, .. , d_n, x^n - x^0.$$

Step 5 : If the stopping test is not verify go to step 1.

APPENDIX B

The Mifflin algorithm require positive real numbers α,β,γ,δ and ρ with r<1 and $(2\beta^2 n^2 \gamma) < \rho$. Given these parameters , the detailed algorithm is as follows :

Step 0 : Choose a starting solution point x of R^n and a starting step-size s > 0. Set the sequence index k = 1 and the sequence values $x^1 = x$ and $s_1 = s$.

Step 1 : Compute an n-vector of approximation first partial derivatives ΔF by :

$$\Delta F_i = (1/2s)\ [\ F(x + se_i) - F(x - se_i)]\ \text{for i = 1 to n}$$

and an approximate gradient norm : $\| \Delta F \| = [\ \Sigma\ (\Delta F_i)^2\ ]^{1/2}$.

Set the descent direction indicators σ_i for i=1 to n : $\sigma_i = +1$ if $\Delta F_i \leq 0$ and $\sigma_i = -1$ if $\Delta F_i > 0$.

Define a best axis point x_a by : $F(x_a) = \min F(x + s\sigma_i e_i)$.

Step 2 : Compute an (n,n) matrix of approximate second partial derivatives $\Delta^2 F$ by :

$$\Delta^2 F_{ii} = (1/s^2) [F(x + se_i) + F(x-se_i) - 2F(x)] \text{ for } i = 1,2, .. , n$$

$$\Delta^2 F_{ij} = (\sigma_i\sigma_j/s^2) [F(x + s\sigma_i e_i + s\sigma_j e_j) + F(x) - F(x + s\sigma_i e_i) - F(x + s\sigma_j e_j)] \text{ for } 1 \leq i < j \leq n$$

Define a best corner point x_c by : $F(x_c) = \min F(x + s\sigma_i e_i + s\sigma_j e_j)$,

and a (possible) move point x_m by : $F(x_m) = \min [F(x_a) , F(x_c)]$.

Step 3 : For $1 \leq j \leq i \leq n$, if $|\Delta^2 F_{ij}| > \gamma$, replace $\Delta^2 F_{ij}$ by $\gamma \text{sign}(\Delta^2 F_{ij})$. Using a modified Cholesky factorisation procedure (see [9]) wicn compute the matrices L,D and E such that $LDL^t = \Delta^2 F + E$. Define index q by :

$$D_{qq} - E_{qq} = \min [D_{ii} - E_{ii}] .$$

Step 4 : If $\alpha s > \| \Delta F \|$ and $D_{qq} - E_{qq} \geq 0$, go to step 7 . If $\alpha s \leq \| \Delta F \|$, compute y^1 satisfying $LDL^t y^1 = -\Delta F$ and set p = 1 ; and if $E \neq 0$, set $y^2 = - [\| y^1 \| / \| \Delta F \|] \Delta F$ and p = 2 and if $D_{qq} - E_{qq} < 0$, compute z satisfying $L^t z = e_q$ and set $y^3 = - \text{sign}(z^t \Delta F) [\| y^1 \| / \| z \|] z$ and p = 3 , and define a search direction vector d by :

$$d^t \Delta F + 1/2 d^t \Delta^2 F d = \min [(y^i)^t \Delta F + 1/2 (y^i)^t (LDL^t - E) y^i] \text{ with } 1 \leq i \leq p .$$

Otherwise ($\alpha s > \| \Delta F \|$ and Dqq - Eq < 0) compute z as above and set $d = - \text{sign}(z^t \Delta F) z$.

Step 5 : Compute, if possible, a search point x + td, where t is a positive number satisfying

$$F(x + td) - F(x) \leq \rho t [d^t \Delta F + 1/2 t d^t \Delta^2 F d] .$$

Then redefine x_m by : $F(x_m) = \min [F(x + td) , F(x_m)]$.

Step 6 : If $F(x_m) - F(x) > - (\alpha\beta s)^2$, go to step 7 . If $F(x_m) - F(x) \leq - \beta^2 \| \Delta F \|^2$, choose some reduced stepsize r in the interval (0,s] and go to step 8. Otherwise set r = s and go to step 9.

Step 7 : Set r = s / 2 and $x_m = x$.

Step 8 : If $x \neq x^k$ replace k by k+1. Set the sequence values $x^k = x$ and $s_k = s$.

Step 9 : Replace x by x_m and s by r and go to