519.28: 681.326.06

Moduleco, software for macroeconomic modelization*

PIERRE NEPOMIASTCHY I.N.R.I.A., Domaine de Voluceau, 78150, Rocquencourt, France (Received 2 February 1981)

सार - इस गोघ पत्न में माइयूल्कों प्रणाली का वर्णन किया गया है। इस प्रणाली का निर्माण बड़ी संख्या में (1000 या अधिक समीकरण) गतिक तथा अरेखीय मैंबरोइकानामिक मॉडलों की रचना व उनके उपयोग को सरल करने के लिये किया गया है।

माड्युल्कों प्रणाली में समय-ऋंखला आंकड़ा आधार प्रबन्ध के लिये यंत्रेतर सामग्री मॉडल समीकरण आगत के लिये एक विशेष मॉडलिंग भाषा, कार्यों को चालू करने के लिये एक विशेष आदेशात्मक भाषा, औपचारिक गणना के कई उपकरण तथा आंकड़ों को सुगम आगत व निर्गत तथा सुगम दृश्यावली उत्पादन के लिये एक अन्योन्य किया भाषा शामिल है।

इस शोघ पत्न में उन गणितीय एलगोरिष्मों का वर्णन भी जिन्हें माड्यूल्कों प्रणाली में शामिल किया जाना है, वास्तव में, हमने यह पाया है कि अधिकतर मैक्रोइकानामिक माँडल अपना विभुजीय रूपाकार में रखे जा सकते हैं। सम्भवतः चरांकों तथा समीकरणों के पुनरांकन के बाद चरांकों काएक छोटा समुच्चय लूप चरांक होता है। इस प्रकार उन के प्रदत्त मूल्यों के लिये शेष माँडल तिभुजीय होता है और सीधे ही हल किया जा सकता है। जैसा कि हम पहले ही बता चुके हैं, तिभुजाकार जैसे माँडलों की नकल की जा सकती है तथा इन्हें सामान्य माँडलों की अपेक्षा अधिक तेजी से लाभानुकूलित किया जा सकता है। इस तरह माइयूल्कों प्रणाली में लूप चरांकों की संख्या कम से कम करने के लिए चरांकों तथा समीकरणों के स्वचालित पुनरांकन के लिये विधियां शामिल हैं।

इस तिभुजाकार रूप को ध्यान में रखते हुए तब एल्गोरिथ्मों के प्रतिलिपिकरण व लाभानुकूलन को अनुकृलित किया जाता है । वास्तविक मैंबरोइकानामिक मॉडलों पर किए गए अनुप्रयोगों से प्रस्तावित विधियों की क्षमता स्पष्ट की गई है ।

इसके अतिरिक्त, यादृष्ठिक नियन्त्रण सिद्धांत में सुप्रसिद्ध संलग्न चरांक तकनीक को मैक्रोइकानामिक माँडलों की बनावट के अनुकूल ढाल लिया गया है। फ्रांसीसी स्टार माँडल (139समीकरण) के नमूने पर यह दशिया गया है कि आमतौर पर अर्थशास्त्री इसी तकनीक का प्रयोग करते हैं।

ABSTRACT. The paper describes the Moduleco system which is designed to facilitate the construction and the use of large scale (1000 equations or more) dynamic and non-linear macroeconomic models.

The Moduleco system will include a software for the management of the time-series data base, a special modelling language for the model equations input, a special command language to active the tasks, several tools of formal computation and an interactive language for easy data input-output and for easy scenario generation.

The paper describes also the mathematical algorithms which are to be included in the Moduleco system. Indeed, we have noticed that most of the macroeconomic models can be put in a quasi-triangular form: possibly after renumbering of the variables and equations, there exist a small set of variables, called loop variables, such as for given values of them, the remaining model is triangular and can be solved directly. As we have shown that quasi-triangular models can be simulated and optimized much faster than general ones, the Moduleco system will include methods for automatic renumbering of variables and equations in order to minimize the number of loop variables.

The simulation and optimization algorithms will then be adapted to take into account this quasitriangularity. Experiments made on 4 concrete macroeconomic models have shown the efficiency of the proposed methods.

Moreover, the adjoint variable technique, well known in optimal control theory, has been adapted to the structure of macroeconomic models. On the example of the French STAR model (139 equations), it is shown that this technique is 106 times faster to compute the gradient than the finite difference technique generally used by economists.

^{*}Paper presented in the symposium "Indo-French school on recent advances in Computer Techniques in Metecrology, Biomechanics and applied systems" held at I.I.T., New Delhi, February 1980.

1. Introduction

With the progress of computer science, macroeconomic models have grown in size and complexity. For instance, the DMS model (Fouquet *et al.* 1976), one of the recent models of the French national economy, is a dynamic, non-linear model of 1262 simultaneous equations per period.

For models of this size, it is physically impossible, without the help of a specialized software, to manage the very large amount of data and describe the equations of the model without mistakes. An obvious consequence is that many computer systems for macroeconoemic modelling have been constructed in the last years with a double aim:

- (a) to facilitate the *construction* of the model; in particular, the equations will be given by the user in a language very close to the natural one of economists:
- (b) to facilitate the *use* of the model; for example with commands for easy scenario generation and for readable and convenient outputs.

At the present time, it appears (see Oudet & Ruderman 1979), with the comparative study of 12 systems, and (Ruderman 1978) that there does not exist any system which satisfies all the following criteria which to us seem to be essential:

- 1° transportability on different computers;
- 2° flexible and open structure of the system, making it easy to add, modify or replace any algorithm or module, in order to keep up with scientific progress;
- 3° integration of the software components in a multiprogrammed interactive system;
- 4° adaptability to the specific needs of the user and for example, to his own internal standards for economic data storage;
- 5° easiness of use, for example for scenario generation by non-specialists of computer science;
- 6° efficiency, i.e., including the most recent mathematical algorithms especially designed for dealing with large scale models (more than 1000 equations).

To answer the needs clearly expressed by numerous economists, some French agencies decided in 1977 to support the development of a computer system which aims to the maximal fulfilment of the criteria listed above. Based on Modular programming technique and designed for Economists, this software is called Moduleco. It is built by a non-profit association, the "Club Moduleco", which is briefly presented in in section 4.

(c) to valorize the French research in econometry and applied mathematics by giving to the scientists of these fields the possibility of testing their methods on real models with real data without the laborious tasks of equations and data copy and by giving to them the guarantee that their methods, if they appear to be competitive, will be concretely used (easy

substitution of a module in the Moduleco software).

2. Description of the Moduleco system

The Moduleco system mainly performs the following tasks:

(i) Management of the time-series data base—With multidimensional time-series stored with their identification, passwords, periodicity [a user can design in Moduleco his own calendar such as a year with 15 months (useful in national accounting)], and comments; a special language is developed to read (with or without the comments), write, copy and transform (the series "inflation rate" is obtained from the series "price"), either in conversational or batch mode, the time-series of the base. The relation between the variables of a model and the time-series is established either implicitly by name identification or explicity (variable name, name of the time-series).

For more details on the data base management, see Club Moduleco Assoc. Publ., ref. A09.

- (ii) Input of the model equations This is done with the special modelling language of Moduleco. Its syntax is very close to the conventional writing of econometric equations. Variables are identified by their names (with 16 significants characters). Some predefined operators can be used (such as growth rates) for conciseness and clarity of the model description. Alternative equations are written as in modern programming languages (if condition then set of equations else set of equations). Sectorial or regional models are written with the help of indexed variables, which may have values in symbolic sets [the French G.N.P. can be written GNP (France) in the model equations]. The description of the model also includes its own documentation associated to equations or variables (input DOC GNP produces in conversational mode the comments associated with the variable GNP.
- (iii) Translation of the model The source text of the equations is checked by a translator which eventually outputs error messages and builds the following data: a tree-form representation of the equations, the symbol table of the identifiers (category, type, dimension) and the incidence matrix (see section 3 of this paper).
- (iv) Estimation The first version of Moduleco will only contain well known and simple (such as OLSQ) techniques estimating the model equation by equation. The second version will contain more sophisticated algorithms such as maximum likelihood estimation of simultaneous equations [see Club Moduleco Assoc. Publ., ref. A07 and (Holly 1979)].
- (v) Simulation and optimization The mathematical methods used for these tasks are described in section 3 of this paper. From a software point of view, let us point out the following formal manipulations done by the system: isolation of the predetermined part (expressions of known data and of lagged endogeneous variables) of each equation which are

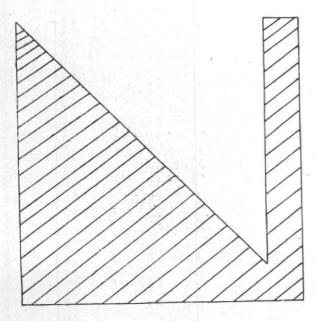


Fig. 1. Incidence matrix of the model

evaluated only once for each simulation period; analytical resolution of each equation by reference to the explained variable (when possible); computation by formal derivation of the analytical expression of the Jacobian matrices of the model.

For more details on these formal manipulation, see Club Moduleco Assoc. Publ., ref. A03.

(vi) Outputs and scenario generation — A command of Moduleco will help the user to find (in interactive mode) the identifier of any variable which then can be printed out, with the comments of the model builder, by a report generator. Graphic outputs will also be available in version 2, as well as an interactive language for easy scenario generation by users which are neither computer scientists nor the model builders such as members of Parliament), as described by Nepomiastchy (1978).

We are not going to describe here the architecture of the Moduleco system; this is done in details in Club Moduleco Assoc. Publ., ref. AO2 and a brief presentation is also given by Tarabout et al. (1980). Let us just say a few words of the command language of Moduleco, described in details in Club Moduleco Assoc. Publ., ref. AO1 and also briefly presented by Tarabout et al. (1980). This language offers a wide variety of standard commands to handle models, programs, data, etc. These commands are hierarchical: they may have subcommands, some of them having in turn subcommands, and so on.

One original aspect of Moduleco is that the abstract type command is available to the users. They can create, modify or destroy commands with the help of primitives, thus the set of commands is extensible in personalized way.

3. Solving and optimizing quasi-triangular econometric models

Let us consider a non-linear and dynamic macroeconomic model with n endogeneous variables. To solve this model, we proposed in 1976 (Nepomiastchy 1977) the following method. A set of s variables, called *loop variables*, and a set of s equations, called *loop equations* are selected such as: for given values of the loop variables, the remaining model with n-s equations and n-s variables is triangular and can be solved directly; then an algorithm is chosen to iterate on the values of the loop variables in order to satisfy the loop equations.

3.1. Structure of the model

For sake of clarity of presentation, we shall simplify the structure of the model (a more general presentation is given by Gabay *et al.* (1980) assuming that the model is described by the following set of equations:

$$k_i x_t^i = f_t^i(x_t, x_{t-1}, \dots, x_t - p, u_{t-1}, \dots, u_{t-q})$$
, (1a)

$$x_t$$
 given for $i = 1, ..., n$ and $t = -p+1, -p+2, ..., -1, 0$ (1b)

where n is the number of endogeneous variables, x_i^i is the value of endogeneous variable i at period t, k_i is a given indicator $(k_i=1)$ if equation i is solved in x_i^i and $k_i=0$ otherwise) x_i is the vector $[x_i^1,\ldots,x_i^n]$, r is the number of control variables, u_i^i is the value of control variable i at period i, i is the vector $[u_i^1,\ldots,u_i^n]$ and i (resp. i) is the maximum lag on endogeneous (resp. control) variables. The given functions i are continuously differentiable and at least for any control taken in a reasonable range, the system (1) has a unique solution.

For solving the model at period t, the vector:

$$e_t = [x_{t-1}, x_{t-2}, \dots, x_{t-p}, u_t, u_{t-1}, \dots, u_{t-q}]$$
 (2)

is known and the problem is to find the solution x_t of the problem:

$$k_i x_t^i = f_t^i (x_t, e_t)$$
 , $i = 1, ..., n$ (3)

After a possible renumbering of equations and variables, the loop equations are the s last equations of the model and $y_t = [x_t^{n-s+1}, \dots, x_t^n]$ is the vecto of the s loop variables. Then, by definition of the loop variables, the system (3) is triangular for any given y_t and (3) can be written:

$$x_t^i = f_t^i(x_t^1, \dots, x_t^{i-1}, y_t, e_t)$$
 , $i = 1, \dots, n-s$ (4a)

$$k_i x_t^i = f_t^i(x_t, e_t)$$
 , $i = n - s + 1, \dots, n$ 4(b)

Note that $k_i=1$ for any $i \le n-s$. The incidence matrix E, the elements e_{ij} of which being equal to 1 if x_t^j occurs in equation i and 0 otherwise, has then the structure (Fig. 1). where the non-zero pattern is hatched.

If s is small in comparison with n, then most of the non-zero elements of the incidence matrix E are under the main diagonal and matrix E is said to be quasitriangular.

TABLE 1
Quasi-triangular structure of 4 models

Model	No. of eqns.	No. of loop variables
R.C. Fair's model	83	6
STAR (French Administr.)	139	3
MPS (Americ. Administer.)	400	20
METRIC (French Administr.)	426	29

It is clear that a renumbering of variables modifies the matrix E, thus the quasi-triangularity is not a specific property of a model. We have developed an algorithm, described by Nepomiastchy et al. (1978), which renumbers variables and equations in order to minimize the number of loop variables. This algorithm has been applied to concrete French and American macroeconomic models with encouraging results (Table 1).

Consequently, we suppose that most of the macroeconomic models can be put in a quasi-triangular form (this is due to the fact that in behavioural equations, the variable is "generally" explained by variables previously introduced) and, for this reason, we have designed solution and optimization algorithms especially adapted to this structure.

3.2. Solving the model

For any given y_t , the system (4) is triangular and can be solved by simple evaluations of functions f_t^i , which give us values of x_t^i for any $i \le n-s$. Since $y_t = [x_t^{n+s-1}, \dots, x_t^n]$, the subsystem can be considered as a mapping $y_t \to x_t(y_t)$. Of course, with an arbitrary value of y_t , the solution $x_t(y_t)$ of (4a) does not satisfy in general the Eqns. (4b); there is an error (which depends on and only on y_t) which shall be denoted by $\phi_t(y_t)$. The problem is then to find y_t in such a way that this error is equal to zero:

$$\phi_i(v_i) = 0$$
 , $i=1,...,s$ (5)

Note that the system (3), of dimension n, is reduced to the system (5), which is only on dimension s (n=139 for the STAR model, and s=3).

In theory, it is possible to eliminate variables x_t^i , $i \le n$ —s, from Eqn. (4) and obtain the analytical expressions of functions $\phi_t^i(y_t)$. For large models, this is obviously impossible and one has to compute using (4a) the numerical value of $\phi_t(y_t)$ for any given numerical value of y_t .

Papers by Gabay at el. (1980), Nepomiastchy & Ravelli (1978) and Rachdi (1979) give comparisons of the the computing times for solving the systm (5) using Gauss-Seidel and several Newton-like methods (classical Newton, modified Newton, discretized Newton, secant Broyden's method and Brown's method). These methods have been compared on two models (STAR and Fair's). It appeared clearly that the Gauss-Seidel issmuch less efficient than any Newton-like method (for STAR, with a required precision of 10⁻⁴, the model is solved

in 0.344 seconds by Gauss-Seidel and in 0.046 seconds by the secant method), that the loop variable technique deduces (at least for the models tested) the computing time by a factor 100 [in the same conditions, the secant methods applied to the "global" system (3) converges in about 15 seconds], and that the secant method seems to be the best among the versions of Newton, without a a big difference with some others.

Some of the methods used for the comparison require the knowledge of the partial derivatives $\partial \phi_i t^i / \partial y_i t^j$ of system (5). If no formal derivation package is available, they are replaced by finite difference approximations (discretized Newton's method). But we have seen in section 2 that Moduleco will-include such a package. In this case, the formal expressions of the Jacobian $(\partial f_i^t / \partial x_i t^j)$ can be computed from the relations:

$$\frac{\partial x_t^i}{\partial y_t^j} = \sum_{k=1}^{i-1} \frac{\partial f_t^i}{\partial x_t^k} \frac{\partial x_t^k}{\partial y^j_t} + \frac{\partial f_t^i}{\partial y_t^j} , i \leqslant n-s, j=1, \dots, s \quad (6)$$

which, for any j, is a triangular system in $\partial x_i^i \partial y_i^j$ and :

$$\frac{\partial \phi_{t^{k}}}{y_{t^{j}}} = k_{i} \, \delta_{ij} - \sum_{k=1}^{n-s} \frac{\partial f_{t}^{n-s+i}}{\partial x_{t^{k}}} \, \frac{\partial x_{t^{k}}}{\partial y_{t^{j}}} - \frac{\partial f_{t}^{n-s+i}}{\partial y_{t^{j}}} ,$$

$$i, j = 1, \dots, s \qquad (7)$$

3.3. Optimizing the model

We have to minimize the following loss function:

$$j(x, u) = \sum_{t=1}^{T} j_t (x_t, x_{t-1}, \dots, x_{t-p}, u_t, u_{t-1}, \dots, u_{t-q})$$
 (8)

where x and u are linked by the model (1). We shall assume that the functions j_t are continuously differentiable.

In this paper, we are only concerned by the search of an efficient method to compute the reduced gradient J'(u) of J(u)=j[x(u), u] where x(u) is the solution of the model (1) associated to the control u. The generally used by (economists) finite difference method is very simple but requires for one gradient evaluation rT model simulations, where r is the control dimension and T the number of periods. We have proposed in (8) to apply the adjoint variable method, well known in optimal control theory (see, for example, Bensoussan et al. 1974, and we have shown in [11] how this technique can be adapted to optimize quasi-triangular models. Let us briefly recall these results. Let δ_t be the indicator:

$$\delta_t = \left\{ \begin{array}{l} 1 \text{ if } t \leqslant T \\ 0 \text{ if } t > T \end{array} \right.$$

and let $\psi = [\psi_1, \dots, \psi_T]$, with $\psi_t \in \mathbb{R}^n$, be the solution of :

$$\psi_{t}^{i} = \sum_{k=0}^{p} \delta_{t+k} \left[\sum_{j=1}^{n} \frac{\partial f_{t+k}^{j}}{\partial x_{t}^{i}} \psi_{j_{t+k}} + \frac{\partial j_{t+k}}{\partial x_{t}^{i}} \right],$$

$$t=1,\ldots,T$$

$$i=1,\ldots,n$$
(10)

Then, it is proved by Gabay et al. (1978) that the gradient J'(u) is deduced from the knowledge of ψ by the formula:

$$\{J_{i}'(u)\}_{i} = \sum_{k=0}^{q} \delta_{t+k} \left[\sum_{j=1}^{n} \frac{\partial f_{t+k}^{j}}{\partial u_{i}^{i}} \psi_{t+k}^{j} + \frac{\partial j_{t+k}}{\partial u_{i}^{i}} \right],$$

$$t=1,\ldots, T$$

$$i=1,\ldots, r$$

$$(11)$$

It is shown by Nepomiastchy and Ravelli (1978) that, solving the adjoint system (10) for t=T, T-1, and so on till t=1, one has to solve for each period a linear system of dimension s which has the same quasi-triangular structure (but transposed) that the model itself and can be reduced by the same technique as above) to a linear system of dimension s (s is still the number of loop variables); moreover, the matrix defining this linear system is the transposed matrix of the Jacobian matrix of the system (5), consequently this matrix is known and the gradient computation is very fast. The numberical experiments given is (11) on the example of the STAR model, with r = 10 et T = 30, show a gradient computing time of 4.90 sec for the finite difference method and of 0.046 sec (106 times faster !) for the adjoint variable technique.

4. The Moduleco club

It was decided from the beginning of the Moduleco project (in 1977) to associate in an appropriate structure various scientists (computer scientists, economists, applied mathematicians, statisticians) implied in the system construction as well as the future users of the system and the scientists interested by the algorithms we are developing. Consequently, it was created in February 1979 a non-profit association called "Club Moduleco", the aims of which are to conceive, realize, operate, maintain and develop the Moduleco system, which will be built in order to satisfy as closely as possible the needs of the Association members.

At the present time, there are about 30 person members of the Association (including foreign scientists) as well as the following French institutions:

Assemblee National;

Centre Interuniversitaire de Calcul de Grenoble :

Direction de la Prevision du Ministere de l'Economie ;

Institut National de la Statistique et des Etudes Economiques ;

Institut National de Recherche d'Informatique et d'Automatique ;

Laboratoire d'Informatique et de Mathematiques Appliquees de Grenoble;

Senat.

Each one of these institutions contributes to the Moduleco project either with its scientists or by financial support. Moreover, the Club has also a financial support from the French Ministry of Industry.

The calendar of the project is a first (experimental) version of the Moduleco system in 1981 and a complete version in 1983.

More information on the Club can be obtained from its Director who is the author of this paper.

5. Conclusion

The objectives of the Moduleco project are the following:

- (a) to built a software for macroeconomic modelization which must be operational, effective, easy-touse even for non-specialists and which can be adapted to the specific needs of the user.
- (b) to valorize the French research in computer science by putting on the market a product built with the most recent programming techniques.
- (c) to valorize the French research is econometry and applied mathematics by giving to the scientists of these fields the possibility of testing their methods on real models with real data without the laborious tasks if equations and data copy and by giving to them the guarantee that their methods, if they appear to be competitive, will be concretely used (easy substitution of module in the Moduleco software.

Acknowledgement

This paper not only presents the own results of its author, but also the work done by the Moduleco team, which is at present time composed of V. Chari, D. Gabay, M. C. Imbert, S. Maurin, P. Nepomiastchy, B. Oudet, B. Pincemaille, M. Rachdi, A. Ravelli, F. Rechenmann, B. Sefsaf, F. Sylvain and A. Tarabout

References

- Fouquet, D., Charipin, J. M., Guillaume, H., Muet, P. A. & Vallet, D., 1976, DMS, Modele de prevision a moyen terme, Economie et Statistique, n°79, 1976 (in French).
- Oudet, B. A. and Ruderman, G., 1979, Etude comparative deslogiciels disponibles, in *Methodes Mathematiques de la Modelisation Economique*, p. Malgrange ed., coll. Les Syntheses du Sesori, IRIA, May 1979 (in French).
- Ruderman, G., 1978, Compartive study of systems currently available in the United States for construction and use of macroeconomic models, proceedings of the Journees sur la Construction et l'Utilisation de Modeles Macroeconomiques, IRIA, October 1978.
- Club Moduleco Association Publications which can be obtained from the author of the present paper (in French) A09.
- Holly, A. 1979, Methodes d'estimation in Methodes Mathematique de la Modelisation Economique, P. Magrange ed., uoll. Les Syntheses du SESORI, IRIA, May 1979 (in French).
- Nepomiastchy, P., 1978. Moduleco, systeme informatique pour la construction et l'utilisation de modeles macro-economiques, presented at the *Colloque sur la Decision Parlemen taire et l'Informatique*, French Senate, May 1978 (in French).
- Tarabout, A., Rechenman, F. and Sefsaf, B., 1980, Moduleco, An Integrated, Modular Computer system for Economists, presented at IFAC World Congress, Warsaw.

- Nepomiastchy, P.. 1977. Methodes d'optimisation adaptees aux, modeles macroeconomiques, pesented at the *International Symposium on "New Trends in Systems Analysis*, Versailles (France, Dec. 1976 and published in "New Trends in System Analysis", Part 2, pp. 626-635, *Lecture Notes in Control and Information Sciences*, Springer Verlag, (in French).
- Gabay, D., Nepomiastchy, P., Rachdi, M., and Ravelli, A., 1980, Numerical Methods for Simulation and Optimal Control of Large Scale Macroeconomic Models, in "Applications of Stochastic Control in Econometrics and Management Science", A. Bensoussan, P. Kleindorfer, C. S. Tapiero eds., North Holland.
- Nepomiastchy, B., Ravelli, A. and Rechenmann, F., 1978, An Automntic Method to Get an Econometric Modelin a Quasi Triangular Form, IRIA, Report n* 313.
- Nepomiastchy, P. and Ravelli, A., 1978, Adapted Methods for solving and Optimizing Quasi-Triangular Econometric Models, Annuals of Economic and Social Measurement, 6/5, pp. 555-562.
- Rachdi, M., 1979, Quelques algorithmes de resolution de Modeles macroeconomiques, These de 3e Cycle, Universite Paris IX, July 1979 (in French).
- Bensoussan, A., Hurst, E. and Naslund, B., 1974, Management Applications of Modern Control Theory, North Holland.
- Gabay, D., Nepomiastchy P., Rachdi, M. and Ravelli, A., 1978, Etude, resolution et optimisation de modeles macroeconomiques, IRIA, Report nº 312, (in French).