Recent applications of Quasi-Newton Methods for solving nonlinear systems of equations

Rosana Pérez * Véra Lúcia da Rocha Lopes †

August 28, 2001

Resumo

Os métodos quase-Newton têm sido amplamente utilizados na resolução de sistemas não lineares que surgem nas mais diversas áreas de aplicações, como na Física, na Engenharia, na Química e na Indústria. Muitas vezes métodos desta família são desenvolvidos e analisados para a resolução de problemas específicos, como é o caso por exemplo, de problemas de complementaridade não linear [8].

Neste trabalho nos propusemos a estudar várias *aplicações recentes* de métodos quase-Newton para resolver sistemas de equações não lineares. É parte fundamental do trabalho, uma minuciosa pesquisa bibliográfica via Bibliotecas e também via Internet, para a escolha das aplicações. Fazemos uma análise crítica das aplicações encontradas e discutimos a eficácia dos métodos quase-Newton utilizados.

*Department of Mathematics, University of Cauca, Popayán, (Cauca), Colombia (rosana@atenea.ucauca.edu.co).

†Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas, SP, Brazil (vlopes@ime.unicamp.br).

No final do trabalho, elaborarmos um diagnóstico sobre o status dos métodos quase-Newton na resolução de problemas práticos. Basicamente nos propomos a responder perguntas do tipo: existem problemas, na pesquisa aplicada, para os quais os métodos quase-Newton são realmente a melhor opção? Quais? Por que?

Abstract

The quasi-Newton methods have been very used in the solution of nonlinear systems that appear in the most applied areas, such as Physics, Engineering, Chemistry and Industry.

Many times some methods of this family are developed and analyzed for a solution of particular problems, for example as in the case of nonlinear complementarity problems [8].

In this work we study several *recent applications* of quasi-Newton methods for solving nonlinear systems of equations. It is a fundamental part of this work, a careful bibliographical research via Libraries and also via Internet, for a selection of the applications. We hope we have made an understandable abstract of the applications chosen and of the quasi-Newton methods used, in each of them.

With this work we believe that we elaborated a diagnosis of the status of the quasi-Newton methods in the solution of real life problems, answering thus, questions like: (i) are there problems, in the applied research, for which the quasi-Newton methods are the best option? (ii) which are they? (iii) why?

1 Introduction

In recent years, quasi-Newton methods for solving square smooth nonlinear systems have been out of the mainstream of numerical analysis research. The SIAM Journal on Numerical Analysis published 4 papers on the subject before 1970, 10 between 1971 and 1980, 11 in the eighties and none from 1991 to 1999. Sometimes, research in a family of numerical techniques becomes out-of-fashion after its incorporation to ordinary practice of problem solvers in Physics, Chemistry, Engineering and Industry. So, promising algorithms are completely forgotten, both in research and in applications.

What is the real situation of quasi-Newton methods for solving nonlinear systems?. The classical paper of Dennis and Moré [3] is cited in most works concerning quasi-Newton methods for nonlinear systems. It had been cited 361 times in indexed scientific journals. The last 100 citations go from 1992 to the present days. 42 of these citations come from non-mathematical journals. It must be warned that, frequently, the Dennis-Moré paper [3] is cited in connection to quasi-Newton methods for minimization problems, and not for nonlinear systems.

Since the everyday practice in Physics, Chemistry and Engineering includes the solution of nonlinear systems using Newton's method, one is tempted to conclude that the penetration of the quasi-Newton technology in applications, although existing, has not been as intense as the potentiality of the technique deserves. But, our bibliographical research shows that actually the quasi-methods are used very often for solving real problems in the areas above mentioned.

In the introduction of most quasi-Newton papers, it is stressed that the main motivation to use them avoid computation of cumbersome derivatives. However, even before the boom of automatic differentiation, practitioners found that, for many of their problems, computing derivatives was not as difficult or costly as stated in the quasi-Newton literature. They also verified that beginning a quasi-Newton process with $B_0 = I$, or some other arbitrary matrix, very often causes disastrous results and, so, the computation of an initial Jacobian is almost always necessary. Moreover, the programming effort of computing

the initial Jacobian is the same as the one necessary for computing all the Jacobians, so the tendency of many practitioners has been to use Newton's method or its stationary variation with refinements. Quasi-Newton methods for solving large-scale nonlinear systems are largely used in applications when both, numerical analysts and potential users, are conscious of their real advantages and limitations. Besides the fact that, in general, quasi-Newton methods do not calculate derivatives, most of them update the iteration matrix in a very simple way.

We consider nonlinear systems of equations

$$F(x) = 0, \tag{1}$$

where $F : I\!R^n \to I\!R^n$ has continuous first partial derivatives. We denote $F = (f_1, \cdots, f_n)$ and $J(x) = F'(x)$ for all $x \in I\!R^n$. All practical algorithms for solving (1) are iterative.

Given an initial approximation $x_0 \in I\!R^n$, Newton's method generate a sequence $\{x_k\}$ of approximations of a solution to (1) by

$$x_{k+1} = x_k - J(x_k)^{-1}F(x_k). \tag{2}$$

The Newton iteration can be costly, since partial derivatives must be computed and the linear system (2) must be solved at every iteration. This fact motivated the development of quasi-Newton methods, which are defined as the generalization of (2) given by

$$x_{k+1} = x_k - B_k^{-1}F(x_k). \tag{3}$$

In quasi-Newton methods, the matrices B_k are intended to be approximations of $J(x_k)$. In many methods, the computation of (3) does not involve computing derivatives at all. Moreover, in many particular methods, B_{k+1}^{-1} is obtained from B_k^{-1} using simple procedures thanks to which the linear algebra cost involved in (3) is much less than the one involved in (2).

The name "quasi-Newton" was used after 1965 to describe also methods of the form (3) such that the equation below is satisfied

$$B_{k+1}s_k = y_k = F(x_{k+1}) - F(x_k). \tag{4}$$

Following [2], most authors call quasi-Newton all the methods of the form (3), whereas the class of methods that satisfy (4) are called "secant methods". Accordingly, (4) is called "secant equation".

Among the secant methods, we have Broyden's method [1] and the Inverse Column Update Method (ICUM) [12] [9]. In the first one, the updating of B_k matrix, is made by

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)s_k^T}{s_k^T s_k},$$

and the second one, the matrix B_k^{-1} is updated by

$$B_{k+1}^{-1} = B_k^{-1} + \frac{(s_k - B_k^{-1} y_k)\mathbf{e}_{j_k}^T}{\mathbf{e}_{j_k}^T y_k},$$

where, $|\mathbf{e}_{j_k}^T y_k| = \|y_k\|_\infty$.

We also intend, in a future work, to implement ICUM to solve some of the nonlinear systems that appear in the papers that we chose in our research. The main motivation for doing it is the recent results [10] about the application of this method in the solution of large-scale nonlinear systems.

2 Applications

The initial part of our work was a bibliographical revision via Internet. For this we used the electronic library program associated to the State University of Campinas, SP, Brazil, named **ProBE**, which permit a rapid and updated electronical research to the complete text of some international journals by the Academic Network of São Paulo. The other electronic tool used was the **Web of Science**, that it is a database made by the Institute for Scientific Information (ISI), with the information about papers published since 1945, in more than 8.400 specialized journals, indexed by ISI, in all areas of knowledge. We also used the non virtual UNICAMP's libraries: BIMECC (Biblioteca do Instituto de Matemática, Estatística e Computação Científica) and BAE (Biblioteca da Área de Engenharia).

We started our electronic search using the term **Broyden**. In the **ProBE** library it appears in 195 articles and in the **Web of Science** we found 100 articles, in both cases published after 1995. These documents include theory development and real applications which deal with this quasi-Newton method. The papers found use Broyden's method, not only for solving nonlinear systems of equations but also for solving minimization problems.

We have two main reasons to have stopped our search with Broyden's method. The first one is that we already had a sufficient number of interesting applications and the second was our interest in comparing the performance of Broyden's method with the Inverse Column Update Method (ICUM) [12] [9], considered by Lukšan and Vlček [10] as the most efficient quasi-Newton method for large-scale problems. In this direction we are already in touch with some of the authors as Lucia Medina [13] and Klaus Werner [17]. We are trying to work together and test the efficiency of the ICUM for their problems.

We chose nine applications among the most interesting in the recent real applications of quasi-Newton methods for solving nonlinear systems of equations that were published between 1999 and 2001. In what follows we present these applications.

2.1 Multiple Target 3D Location Airborne Ultrasonic System (2001) [13]

This work is concerned with extend the air-coupled ultrasonic system[1] which provides 2D measurement of target position, to provide 3D measurement by using an area array.

In order to calculate 3D position of targets, a nonlinear system has to be solved. The authors chose to use Broyden's method in conjunction with coarse beam forming process [13].The latest is used to locate the targets and obtain approximate values of their positions. The Broyden algorithm is then used to get more precise measurements.

[1]Developed at Nottingham University [13]

If a target is present in the volume, an echo is produced after the target has being struck by the transmitted pulse, and is received as a delayed version of the transmitted signal by the receivers. The time of flight, τ, at each receiver is measured as the time where the maximum value of the envelope of the echo amplitude is found. Thus, at nth receiver, τ_n is given by

$$\tau_n = \frac{r_n}{c},$$

where c, the velocity of the medium, is known and r_n is the nth round trip distance of the transmitted pulse: it is the distance transmitter-target, z_T, plus distance target-nth receiving element, R_n, i.e.,

$$r_n = z_t + R_n = z_t + \sqrt{(x_t - x_n)^2 + (y_t - y_n)^2 + z_t^2}, \tag{5}$$

where x_t, y_t and z_t are the unknown target co-ordinates, (x_n, y_n) are the known receiving elements position, located in a plane.

Equation (5) represents a nonlinear system which has three nonlinear unknown variables: the target co-ordinates. Thus, for the nth receiver, it is possible to define the function F_n by

$$F_n(x_t\ ,\ y_t\ ,\ z_t) = z_t + \sqrt{(x_t - x_n)^2 + (y_t - y_n)^2 + z_t^2} - r_n.$$

If the nonlinear system

$$F_n(x_t\ ,\ y_t\ ,\ z_t) = 0$$

is solved, it gives both the range and angular position of the target [13].

The minimum number of receiving elements needed to calculate the unknown are three which gives $N!/[3!\,(N-3)!]$ possible combinations for an N elements planar array.

About the results of composite algorithm for solving the nonlinear system (5), the authors noticed that the 3D measurements accuracy are increased when Broyden's method is applied using the initial values given by the envelope beam-forming algorithm. They also observed that the processing time of the coarse envelope beam forming is reduce drastically.

2.2 Equation-based SPYRO model and solver for the simulation of the steam cracking process (2001) [4]

SPYRO, Technips's proprietary yield prediction program for the steam cracking process. It features the accurate steady-state simulation of complex steam pyrolysis of feedstock ranging from gases to gasoils in all known coil designs.

The core of the SPYRO model is the kinetic reaction scheme[2] and the complete SPYRO model is a system that contain a total of 175 ordinary differential equations and 50 algebraic equations.

The flexibility nowadays required of process models initiated the development of the so-called Open Spyro program. In this, all model equations are written in the open or residual form which allows flexibility and the formulation of the basic equations in their natural form, reducing coding errors. This flexibility and the natural form are created by dividing the SPYRO model into several sub-models.

The ordinary differential equations in the Spyro model[3] have the following general form:

$$\frac{dy}{dz} - f(y, z) = 0,$$

and for solving it, the original SPYRO uses a shooting technique. The solution of the ordinary differential equations in the Open Spyro model is approximated with the collocation technique. The method used is that of Orthogonal Collocation on Finite Elements.

A separated nonlinear solver approximates the solution of the Open Spyro model. Indeed, the authors implemented a fast and globally converging quasi-Newton method based on the update proposed by Broyden [4] in order to solve the Open Spyro model. The idea of the method used is combining the damped Newton method with a secant method. This strategy is used to reduce the

[2]It consists of several types of reactions:radical chain inition, radical decomposition, radical addition, radical chain termination, radical isomerisation and purely molecular.

[3]It is boundary value problem

[4]The authors adapted the method proposed by Broyden citebroy, such that it can be used efficiently for large sparse systems.

computational effort which result to use only the damped Newton method. [4]

The results of the original SPYRO program have proven over the years, to be accurate and reliable. The authors have compared the simulation results of the Open Spyro program to validate the correctness of the implemented model and to test the solving method. For this, they implemented the Open Spyro program with a flexible system of sub-models for the simulation of the steam cracking process. The same reliable results as those for the original SPYRO program are obtained.

2.3 A numerical algorithm for flame propagation in premixed gases (2001) [6]

In this paper, the author take a hyperbolic system of conservation laws as a governing system of equations for reacting gases and propose and algorithm to determinate a wave propagation speed uniquely. The wave speed and states around a flame are computed by solving a Riemann problem near a flame in the phase space. The deflagration wave is a traveling wave solution of parabolic equations and admissible as a solution of hyperbolic conservation laws. Its wave speed is computed based on the fact that the admissible solutions are connecting orbits of critical points in the dynamical system derived from parabolic equations.

He consider the governing equations for reacting gas in one dimension and they reduce this system to hyperbolic conservation laws with a source term. [6] To be admissible as physical solutions of corresponding hyperbolic conservation laws equations, the traveling waves must have viscous profiles. For this they manipulated the equations until they obtain the following ordinary differential equations for the viscous profile of a combustion wave:

$$\begin{aligned} \mu u_x &= -m\,(u-u_0) - m\left(\frac{T}{u-s} - \frac{T}{u-s}\right), \\ \mu u_x &= -m\,c_p\,(T-T_0) + \frac{m}{2}\,(u-u_0)^2 - mq\,(Z-Z_0) \\ &\quad + mR\,(u-s)\,\left(\frac{T}{u-s} - \frac{T}{u-s}\right), \end{aligned}$$

$$\begin{aligned} DY_x &= (u-s)\,(Y-Z), \\ Z_x &= -k\,\frac{Y\,\phi(T)}{u-s}. \end{aligned}$$

With $-m = (u-s)\rho, Z = Y + \rho D Y_x/m$ and $\phi(T)$ is given by Arrhenius term [6]. The variables ρ, u, p, T, and Y are, respectively, the mass density, velocity, pressure, temperature, and reactant mass friction of the gas. The constants μ, λ and D are the viscosity, heat conductivity, and diffusion constants, respectively.

To understand the solution of system (6), the author examine its equilibria. Let the state variables be denoted $U = (u, T, Y, Z)$; setting the right-hand sides of (6) to zero, they find the initial state U_0 and the final state U_1 are related by the jump condition. [6]

Because of a deflagration wave has fewer impinging characteristics on it, it is necessary additional information to determine the evolution of the wave. Taking an experimental flame speed law for the additional information, the approach in this paper consist on to construct the Riemann solution from the exact information coming from the internal structure of the flame.

Solving a Riemann problem near a deflagration wave means finding the flame speed s and the two states U_M and U_N for given Riemann initial data U_L and U_R. He consider the dynamical system (6) and let the right side of the dynamical system be a vector function G. Then U_M and U_N are equilibrim points of the dynamical system which is

$$U' = G(U, s, U_N),$$

and these equilibria are connected by a solution orbit. Since the dynamical system is autonomous, its easy to obtain a new solution. To fix the solution uniquely, he imposes the phase condition [6], which requires that

$$\int_{-\infty}^{\infty} < G(U) - G(U_{ref}), G(U) \cdot G(U) > \, d\xi = 0,$$

for some reference solution U_{ref}. The author discretize the integral using the trapezoid rule. Now he has a complete set of equations, which he solves using a damped Newton method with Broyden update of the Jacobian, to obtain the solution of the Riemann problem, particularity U_M, U_N, and s.

2.4 Comparison of Simulation Algorithms for Accelerated Determination of Periodic Steady State of Switched Networks (2000) [7]

It is the aim of this the paper [7] to implement a wide range of general numerical methods to solve the accelerated steady-state problem in power electronics, and compare them on a common basis.

A power electronic circuit can be described by a general nonlinear differential equation

$$\frac{dx(t)}{dt} = g\left(x(t)\,,\, t\right), \qquad x(t_0) = x_0,\ t > t_0. \tag{6}$$

where x is the state vector and the nonlinear function $g(\cdot,\cdot)$ can be discontinuous on time. From an initial state x_0, and given a sufficiently long time interval, the circuit reaches a periodical steady state when

$$x(t+T) = x(t), \qquad \forall\ t. \tag{7}$$

where T is the period.

The problem of finding steady state for the system described by (6) is to solve (6) for a nontrivial solution $x(t)$ subject to the constraint given by (7). This problem can be reduced to finding the initial value $x_0 = X_{ss}$ and a period T satisfying both (6) and (7). Let an error function $f(\cdot)$ be defined as

$$f(x(t)) = x(t+T) - x(t). \tag{8}$$

The problem can then be solved by finding the roots of $f(x(t))$, which are a set of initial values of the state variables, satisfying both (6) and (7).

The methods addressed in [7] are the following: Newton's method with analytically determined Jacobian; Newton's method with numerically determined Jacobian; Broyden's method; Newton's method with a globally convergent strategy; Bukowski's method and Skelboe's method.

The results on several switching converters suggest that the analytical Newton's method is the most accurate and the fastest. When analytical derivatives are not avaliable, both Broyden's and Skelboe's methods are competitive.

2.5 Efficient Bifurcation Analysis of Periodically-Forced Distributed Parameter Systems (2000) [5]

Changes in the qualitative features of the bifurcation diagrams or the dynamic features or forced periodic systems occur at singular points, which satisfy certain defining conditions. This work presents a new, efficient numerical method for the construction of the loci of these singular points. The procedure uses Fréchet differentiation to simplify the determination of the defining conditions and the Broyden inverse update method to accelerate the iterative steps involved in the shooting method. [5]

A forced periodic system satisfies the following set of equations:

$$F(u_0) \;=\; u_0 - u(u_0\,,\,\tau) \;=\; 0. \tag{9}$$

where the fixed point u_0 is the spatially discretized state vector at $t = 0$ and $u(u_0, \tau)$ is the set of state variables after one period τ.

A Monodromy matrix at the fixed point u_0 is defined as

$$M(u_0\,,\,\lambda) \;=\; \frac{du(\tau)}{du_0},$$

where $u(\tau)$ describes the state vector after one period τ with initial condition u_0.

The defining conditions for a saddle-node point[5] and period doubling point[6] are, respectively

Saddle node			**Period doubling**		
$F(u_0, \lambda)$	$=$	0	$F(u_0, \lambda)$	$=$	0
$J(u_0, \lambda)\, v$	$=$	0	$M(u_0, \lambda)\, v$	$=$	$-v$
$< v\,,\, v >$	$=$	1	$< v\,,\, v >$	$=$	1

[5]In this node, the number of periodic solutions changes, in general by two. The Jacobian matrix of (9) has a zero eigenvalue.

[6]The monodromy matrix has an eigenvalue $\mu = -1$. Following this point the solution returs to the fixed point only after two periods.

The numerical solution procedure of this problem consists of four steps:

- Discretization of the spatial derivatives.
- Evaluation of the discretized conditions using a time integration routine.
- Implementation of a continuation procedure.
- Solution to the nonlinear algebraic equation system using a quasi-Newton method.

For example, the nonlinear algebraic system of equations defining the hysteresis variety [5] is

$$G(X) = \begin{pmatrix} F(u_0, \lambda) \\ Lv_0 \\ L^* y_0 \\ < y_0 \, , \, D^2_{uu} F(u_0, v_0) > \\ < v_0 \, , \, y_0 > -1 \\ s^2 - |Y - Y_{old}|^2 - (\lambda - \lambda_{old})^2 - \sum_{i=1}^{2} (p_i - p_{i,old})^2 \end{pmatrix} = 0.$$

where $Y = (u_0 \, , \, v_0 \, , \, y_0)^T$, $X = (Y \, , \, \lambda \, , \, p_1 \, , \, p_2)^T$. Here, Y is the vector of all spatially discretized variables, λ is the bifurcation parameter end p_1 and p_2 are the continuation parameters. The subscript *'old'* denotes the previous continuation step.

The numerical procedure is illustrated first by the construction of a map of parameters regions with qualitatively different bifurcation diagrams for an reverse-flow reactor[7], the direction of feed to which is changed periodically. After, they construct a map of parameters regions in which a cooled reverse-flow reactor has qualitatively different dynamic features. Both maps revel surprising features. Thus, it can provide useful information needed to avoid pitfalls in the design and/or operation of various periodic processes.

[7] An reverse-flow reactor is a packed-bed catalytic reactor, in which the flow direction is periodically reversed to trap a hot zone within the reactor.

The authors conclude that the success of their numerical procedure is due to the use of Fréchet differentiation, to simplify the defining conditions and the use of Broyden's inverse method, to avoid the repetitive computation of the inverse of the Jacobian.

2.6 Finding all Periodic Orbits of Maps Using Newton Methods: sizes of basins (2000) [14]

The motivation of paper [14] is to find periodic orbits of dynamical systems.

Let $F : I\!R^n \to I\!R^n$ be a $\mathcal{C}^2$ map. Their primary focus is on finding periodic orbits of maps on the plane. A $k-$period poin of F is a point p such that $F^k(p) = p$. Newton type methods can be used to find the periodic k point of F by letting $G = F^k - \mathbf{I}$, where $\mathbf{I}$ is the identity mapping, and solving the system $G(x) = 0$.

One of the maps that they study is the Hénon map, $H : I\!R^2 \to I\!R^2$, defined by

$$H \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2.12 - x^2 - 0.3y \\ x \end{pmatrix}.$$

They actually use variants of Newton's methods that are more robust than the traditional one. For an initial point x, they iterate Newton's method many times. If the process converges to a point p which is a periodic point of F, they say x is in the Newton basin of p for period k. They also investigate the size of the Newton basin and how it depends on p and k.

2.7 The Classical Stellar Atmosphere Problem (1999) [17]

Mathematically, the **classical stellar atmosphere problem** consists on the radiation transfer equations simultaneously with the equations for hydrostatic and radiative equilibrium, together with the statistical equilibrium, or rate equations. [17]

These equations are:

- The radiation transfer equations which are solved for the (angular) mean intensities J_i, $i = 1, \cdots, NF$ on a pre-chosen frequency grid comprising NF points.
- The hydrostatic equilibrium equation which determines the total particle density N.
- The radiative equilibrium equation from which the temperature T follows.
- The statistical equilibrium equations which are solved for the population densities n_i, $i = 1, \cdots, NL$ of the atomic levels allowed to depart from "local thermodynamic equilibrium" (NLTE levels).
- The particle conservation equation, determining the electron density n_e.
- The definition equation for a fictitious massive particle density n_h which is introduced for a convenient representation of the solution procedure.

This set of equations has to be solved at each point d of a grid comprising ND depth points. Thus, they are looking for solutions vectors

$$\Psi'_d = (n_i, n_e, T, n_h, N, J_r),$$

Using the ALI method [17] they eliminate at the outset the explicit occurrence of the mean intensities J_i from the solution scheme by expressing these variables by the current, yet to be determined, occupation densities and temperature,

$$\Psi_d = (n_i, n_e, T, n_h, N).$$

The resulting set of equations for the reduced solution vectors is of course nonlinear. The solution is obtained by linearization and the iteration is performed either with Newton's iteration or by other methods, much faster than Newton methods, like quasi-Newton variants.

The linearized system may be writing as

$$\Psi_d = \Psi_d^0 + \delta\, \Psi_d.$$

where Ψ_0 is the current estimate for the solution vector at depth d and $\delta\Psi_d$ is the correction vector to be computed.

Using a local operator the resulting system for $\delta\Psi_d$ is

$$\delta\,\Psi_d = \beta^{-1}\,c_d,$$

where β is $(NN \times NN)$ matrix where NN is the total number of physical variables, that is, $NN = NL + 4$, and c_d is the residual error in the equations.

If β_k^{-1} is the kth iterate of the inverse Jacobian, then an update can be found from

$$\beta_{k+1}^{-1} = \beta_k^{-1} + \frac{(s_k - \beta_k^{-1} y_k)(s_k^T \beta_k^{-1})}{s_k^T \beta_k^{-1} y_k},$$

where

$$\begin{array}{lcll} s_k & \approx & \delta\,\Psi_k & \text{solution vector of preceding linearization,} \\ y_k & \approx & c_{k+1} - c_k & \text{difference of actual and preceding residuum.} \end{array}$$

The authors also mention another numerical variant, the Kantorovich method. They consider it more simple and straightforward to implement. This method keeps the Jacobian fixed during the linearization cycle. In fact, it turns out to be even more stable than Broyden's method in some cases.

At this moment, we are in contact with Professor Klaus Werner. He told us that in their stellar atmosphere code they implemented Broyden's method but hardly used it. Instead, they used Kantorovich's method, as we had pointed out previously.

We asked him some data about functions, initial points, expected solution, and initial approximation to the Jacobian. He thinks that, it is difficult to give simple answers to these questions, and based on our interest to check how the ICUM works in their "real life" problem, he proposed us to make the implementation of the Inverse Column Updating Method (ICUM) in their stellar atmosphere code.

We consider very interesting for our objectives this joint work. So we sent him the ICUM algorithm and are waiting for his answers.

2.8 Power System Parallel Computation by a Transputer Network (1999) [16]

Load flow and transient stability study are two of the most elementary problems in electric power system computation. Normally, before starting transient stability studies, the load flow program must be run to provide the steady-state operation points. [16]

The load flow problem is described by a system of nonlinear algebraic equations

$$F(X) \;=\; 0, \tag{10}$$

where X is a variable vector of bus voltages. For an n-bus system without the slack bus, (10) can be reformulated as:

$$F_i(X_1\; X_2\,,\,\cdots\,,\,X_n) \;=\; 0, \qquad\qquad i = 1, 2, \cdots, n. \tag{11}$$

where

$$F_i \;=\; (\Delta p_i\,,\,\Delta q_i)^T, \qquad\qquad F_i \;=\; (f_i\,,\,e_i)$$

and

$$\Delta p_i \;=\; p_{is} \;-\; Re\left(V_i \sum_{j=1}^{n} \bar{Y}_{ij}\,\bar{V}_j\right)$$

$$\Delta q_i \;=\; q_{is} \;-\; Im\left(V_i \sum_{j=1}^{n} \bar{Y}_{ij}\,\bar{V}_j\right)$$

In the above equations, $V_i = e_i + jf_i$ represents the complex node voltage, and $S_i = p_{is} + jq_{is}$ the injected power.

Newton's method is widely used for the solution of load flow problem because of their convergence characteristics but it is more time-consuming because it needs a new information and factorization of the Jacobian matrix. For this reason, in many cases it is adequately substituted by Broyden's method.

In this work, the authors present unified parallel algorithm with the coincident parallel transputer system for both load flow and transient stability studies.

2.9 Acceleration of Self Consistent Electronic Structure Calculations: storage saving and multiple-secant implementation of the Broyden method (1999) [15]

In this paper, on the basis of Broyden's method to solve simultaneous nonlinear equations,the authors present efficient computational schemes for acceleration of self-consistent electronic-structure calculations. For this, they choose a Si(011)[8] surface model as the test problem, because it is easy to implement the computational schemes which appear in their study into an existing electronic-structure code.

They adopt the traditional self-consistency strategy where the self-consistent solution is sought in terms of the one-electron potential. They assume that they have an initial guess for the one-electron potential as the input quantity V_{in}.

The self-consistent solution V_{sc} can be defined as the initial potential, V_{in} which satisfies that the functional derivative of the **Kohn-Sham (KS)** energy is zero [15],

$$\frac{\delta E_{KS}}{\delta V_{in}} = \chi_0 V_{diff} = 0,$$

where, χ_0 is an independent-particle polarizability operator and V_{diff} is defined to by the difference between the out potential V_{out} and the initial potential V_{in}.

Since it is computationally demanding to evaluate χ_0 and thus the derivative of the **KS** energy, in practice, the self-consistent solution, V_{sc}, is defined as a initial potential V_{in} such that

$$V_{diff} = V_{out} - V_{in} = 0 \tag{12}$$

is satisfied. Thus the self-consistent calculation is reduced to solve the nonlinear system of equations[9] (12).

[8]The Si (011) model is a rectangular supercell containing five (011) and three empty layers, and thus ten Si atoms.

[9]The equation (12) is implicitly defined as a set of six equations [15].

If the Jacobian operator A, implicitly defined by

$$V_{diff} = -\,A\,V_{in},$$

is available, the Newton-type iterative procedure,

$$V_{in}^{n+1} \;=\; V_{in}^{n} + [A^{n}]^{-1} V_{diff}^{n}$$

will lead to a vector sequence $\{V_{in}^{1}\,,\,V_{in}^{2}\,,\,\cdots\,,\,V_{in}^{n}\}$ rapidly approaching V_{in}.

However, since explicit evaluation and handling of A are as demanding as those of χ_0, when the inverse Jacobian $[A^{n}]^{-1}$ is recursively approximated by B^{n}, using Broyden's updating formula

$$B^{n} = B^{n-1} - \frac{\left\{\delta V_{in}^{n-1} + B^{n-1}\delta V_{diff}^{n-1}\right\}\delta V_{diff}^{n-1}}{\|\delta V_{diff}^{n-1}\|^{2}}$$

with

$$\delta V_{in}^{n-1} \;=\; \delta V_{in}^{n} \;-\; \delta V_{in}^{n-1},$$

and

$$\delta V_{diff}^{n-1} \;=\; \delta V_{diff}^{n} \;-\; \delta V_{diff}^{n-1},$$

then input potential for the next cycle predicted by

$$V_{in}^{n+1} \;=\; V_{in}^{n} + B^{n} V_{diff}^{n}$$

converges toward the solution V_{sc} in fewer iterations than that generated by a simple relaxation formula,

$$V_{in}^{n+1} \;=\; V_{in}^{n} + \beta V_{diff}^{n}$$

where β is a mixing parameter.

In this work, they investigate the computational schemes, which are variants of Broyden's method, for acceleration of self-consistent electronic structure calculations. In particular, they have focused on the storage-saving schemes and they propose two storage-saving schemes where iteration data are partially discarded after a prescribed storage limit is reached.

3 Conclusions

At the end of the bibliographical research about the recent applications of quasi-Newton methods for solving nonlinear systems, we chose nine applications to real problems in different areas of knowledge. The analysis of them turned out to be a little difficult for us because, as mathematicians, our knowledge in these applied areas is quite limited.

In the applications, the authors prefer Broyden's method to Newton's method, essentially because of its computational cost.

We are very happy with the results obtained by our research in terms of answering our initial questions:

(i) Yes, there are many problems in applied research for which the quasi-Newton methods (Broyden's method) are the best option.

(ii) In sections 2.1 to 2.9 we presented applications of them in the Physics, Chemical Engineering, Electronic Engineering, Astrophysics, Electric Engineering and Mechanical Engineering areas.

(iii) Broyden's method is chosen because of low computational cost.

Motivated by the fact that the ICUM [12][9] was considered recently as the most efficient quasi-Newton method for solving large-scale nonlinear systems [10], we are motivated to implement it with some of the applications chosen. Actually, we have already started its application to another real problem.

In order to do the implementation of the ICUM to solve the nonlinear systems that appear in the papers, we wrote to the authors of some of the chosen papers, asking them about initial points used, the functions (in some cases), an expected solution, etc. Until now we have received the answers of Professors Lucia Medina [13] and Klaus Werner [17]. In the first case, we received some data and in the second case the author explained us about the difficulty in giving us the answers. Both of them have considered implementation of ICUM in their "real problems" a very interesting idea. So, the natural continuation of this work is trying to make a joint work with them.

References

[1] BROYDEN, C. G., Luss, D. **[1965] A class of methods for solving nonlinear simultaneus equations**. *Mathematics of Computations* 19, 577-593 .

[2] DENNIS, Jr. J. E., Schnabel R. B. **[1983] numerical methods for unconstrained optimization and nolinear equations**. *Prentice-Hall.*

[3] DENNIS, Jr. J. E., More J. J. **[1977] Quasi-Newton methods, motivation and theory**. *SIAM Review* 19, 46-89.

[4] GOETHEM, M. W. M., Kleinendorst, F. I., and Leeuwen, C. van. **[2001] Equation-based SPYRO model and solver for the simulation of the steam cracking process** . *Computers and Chemical Engineering* 25, 905-911.

[5] KHINAST, J. G., Luss, D. **[2000] Efficient bifurcation analysis of periodically-forced distributed parameter systems**. *Computers and Chemical Engineering* 24, 139-152 .

[6] HWANG, H. C. **[2001] A numerical algorithm for flame propagation in premixed gases**. *Applied Mathematics Letters* 14, 487-493.

[7] LI, D., Tymersky, R. **[2000] Comparison of simulation algorithms for accelerated determination of periodic steady state of switched networks**. *IEEE Transactions on Industrial Electronics* 47, (6) 1278-1285.

[8] LOPES, V. L. R, Martínez , J. M. and Pérez, R. **[1999] On the local convergence of quase-Newton methods for nonlinear complementarity problems**, *Applied Numerical Mathematics* 30, 3-22.

[9] LOPES, V. L. R. ; Martínez, J. M. **[1995] Convergence properties of the inverse Column-Updating method,** *Optimization Methods and Software* 6, 127-144.

[10] LUKŠAN, L., Vlček, J. **[1998] Computational experience with globally convergent descent methods for large sparse systems of nonlineal equations,** *Optimization Methods and Software* 8, 185-199.

[11] MARTINEZ, J. M. **[2000] Practical quasi-Newton methods for solving nonlinear systems,** *Journal of Computational and Applied Mathematics* 124, 97-121.

[12] MARTINEZ, J. M., Zambaldi, M. C. **[1992] An inverse column-updating method for solving large-scale nonlinear systems of Equations,** *Optimization Methods and Software* 1, 129-140.

[13] MEDINA, L., Wykes, C. **[2001] Multiple target 3D location airborne ultrasonic system**. *Ultrasonics* 39, 19-25.

[14] MILLER, J. R., Yorke, J. A. **[2000] Finding all periodic orbits of maps using Newton methods: sizes of basins**. *Physica D* 135, 195-211.

[15] SAWAMURA, A., Kohyama, M., Keishi, T., and Kaji, M. **[1999] Acceleration of self consistent electronic structure calculations: storage saving and multiple-secant implementation of the Broyden method**. *Materials Transactions, JIM* 40, (11) 1186-1192.

[16] WANG, F. Z., Hadjsaid, N., and Sabonnadière, J. D. **[1999] Power system parallel computation by a transputer network**. *Electric Power Systems Research* 52, 1-7.

[17] WERNER, K., Dreizler, S. **[1999] The classical stellar atmosphere problem**. *Journal of Computational and Applied Mathematics* 109, 65-93.