On application of Newton's method to solve optimization problems in the consumer and production theories

Aleksander Strasburger *

and

Andrzej Zembrzuski

Department of Applied Mathematics
Faculty of Applied Informatics and Mathematics
University of Life Sciences — SGGW
Nowoursynowska 159, 02-776 Warszawa, Poland

Abstract

The aim of the paper is to test possibility of an application of the Newton's approximation method to obtain solutions of classical problems in mathematical economics — determination of the demand function in the consumers theory and the demand function for the production factors in the theory of the firm. As is well known, in most cases these functions are given implicitely as solutions of an optimization problem. The resulting implicit relations only seldom can be resolved to yield a closed explicit form of solution, what necessitates using approximation procedures to obtain a deeper insight into the nature of the solutions. We present here some results in computing these functions by means of the (multi-dimensional) Newton's method as well as discuss the question of the speed of convergence of the approximating sequence.

1 Introduction

In this paper we consider two fundamental problems of mathematical economics, which in the mathematical perspective can be seen as instances of an identical problem of multi-dimensional analysis. The point is to find solutions of a system of nonlinear

*Corresponding author. E-mail: aleksander_strasburger@sggw.pl

in general equations in $k+1$ variables, with k being the number of commodities dealt with by the theory. The equations of the system arise as the Lagrange equations for a suitably formulated extremum problem with constraints (cf. (6) below). Except in a few cases, e.g. when the function is of the Cobb–Douglas or CES type, the equations do not allow a closed form of the solution so a need for determining approximate solutions arises. One of the best known methods of construction of approximate soultions to nonlinear equations is known under the name of ,,Newton's Method" in recognition of its famous ancestor, *the method of the tangent* devised in 1669 by I. Newton, cf. [5, vol. I, p. 317].

In this paper we present results concerning numerical solution of the system for a choice of utility functions with a varied number $k = 2, 4, 10, 20$ of variables (commodities). The algorithm is implemented with use of a program written in the Matlab language. The results demonstrate that the successive approximations obtained from the Newton's Method converge quite rapidly to a solution without regard to the number of variables or a choice of the initial point. In most cases the stabilization of the approximating sequence is achieved with no more then 10 iterations. This seems to be rather satisfactory and shows its advantage over other methods used to approach the problem, e.g. as in [3].

In the formulation of economical problems we follow the monograph (in Polish) by E. Panek *Ekonomia matematyczna*, and also a very reaable exposition of A. P. Barten and V. Böhm *Consumer Theory* in *Handbook of Mathematical Economics*, [1], while the mathematical instruments are mostly taken from the books of W. Walter, *Analysis 1, 2* [5] and S. Krantz and H. Parks, *The Implicit Function Theorem: History, Theory, and Applications* [4].

2 Two optimization problems of mathematical economics

2.1 Consumer theory

In the classical approach to the theory of consumer demand it is assumed that the consumer preferences may be described by means of a $\mathcal{C}^2$ (this symbol stands for the class of functions posessing continuous partial derivatives up to second order) utility function $u : X \ni x \to u(x) \in \mathbb{R}$ defined on the consumption set X, assumed here to be a closed and convex cone with nonempty interior of the nonnegative orthant of the euclidean space $\mathbb{R}^k$ of k dimensions. Here k is of course the number of commodities included into considerations, while vectors $x = (x_1, \ldots, x_k)$ represent commodity bundles with quantities of l-th commodity given by the number x_l, $l = 1, \ldots, k$. We let $p = (p_1, \ldots, p_k) \in \mathbb{R}^k_+$ to denote a price vector, with components p_l denoting the amounts paid in exchange for one unit of the l-th commodity. The inner product

$\langle x, p\rangle = \sum_{i=1}^{k} x_i p_i$ is therefore the value of a commodity bundle x with respect to the given price system p. Further, if the initial income (wealth) of the consumer is denoted by I, then the so called budget set of the consumer is denoted by $D(x, I) = \{ x \in X \mid \langle x, p\rangle \leq I \}$. This represents the set of all consumption bundles which are so to say within the reach of the consumer. In the case when $X = \mathbb{R}^k_+$ the budget set is the solid compact k-dimensional simplex in $\mathbb{R}^k_+$.

Having set these notations, the main question of the consumer theory may be formulated as the problem of determining the best bundle available for the consumer within the budget set, or in the mathematical terms finding the maximum of the utility function over the budget set.

Problem 1 (Maximization of the utility function)**.**
It is sought for:

$$\max\{u(x) \mid x \in X\}$$

under conditions:

$$x_i \geq 0, \quad \text{for } i = 1, \ldots, k, \quad \text{and} \quad \langle x, p\rangle = \sum_{i=1}^{k} x_i p_i \leq I$$

Under suitable assumption on the utility function (monotonicity and strict quasi-concavity) and positivity of the price vector p it can be proved that the problem has a solution described by the following theorem. (cf. Barten and Böhm [1, p.409]).

Theorem 2.1. *Given a positive price vector p and a positive wealth I the Maximization Problem 1 is uniquely solved by a certain vector $x^0 \in D(x, I)$ with positive coordinates, so that:*
There exists a unique positive constant (Lagrange multiplier) λ^0, such that the vektor x^0 is the unique solution of the system of equations

$$\frac{\partial u}{\partial x_i}(x) = \lambda^0 p_i, \qquad \textit{dla } i = 1, \ldots, k$$
$$\langle x, p\rangle = I$$

Equivalently, the system $(x^0, \lambda^0) = (x^0_1, \ldots, x^0_k, \lambda^0) \in \mathbb{R}^{k+1}_+$ is uniquely determined as a solution of the system of equations

$$\operatorname{grad} u(x) - \lambda p = 0 \tag{1}$$
$$I - \langle x, p\rangle = 0 \tag{2}$$

with $k+1$ unknowns $x_1, \ldots, x_k, \lambda$.

It is important to observe that the solution of the equations (1) and (2) depends on the given values of the prices and wealth, and thus the solution should be viewed as giving *the demand function* $f(p, I)$. The study of properties of this function is the main objective of the theory, so it is of paramount importance to be able to solve the system.

2.2 Firm (Production) theory

Here again, the choice situation is described in terms of a production function $f : X \ni x \to f(x) \in \mathbb{R}_+$, whose value $f(x)$ indicates the quantity of the given good produced by utilizing of k production factors in amounts given by the components of the vector $x = (x_1, \ldots, x_k) \in \mathbb{R}^k_+$. In order to preserve analogy with the previous case (consumers theory) we do not include labor into production factors. Now we have the vector $v = (v_1, \ldots, v_k) \in \mathbb{R}^k_+$ of unit prices for the production factors, and the unit price p by which the product is beeing sold. The inner product $\langle x, v\rangle = \sum_{i=1}^{k} x_i v_i$ gives the cost of producing the quantity $f(x)$ of the final product, while the function $\Pi(x) = pf(x) - \langle x, v\rangle$ describes the profit the producer can make by selling the whole amount of the produced good at the price p. It seems reasonable to assume that the firm has only finite resources which it can use for buying production factors, and thus that the analogous bound on the amount of money spent as in the consumer theory may be meanigfully considered. Thus we assume that there is a given wealth I remaining at the disposition of the firm so that it will decide to run its production under condition that $\langle x, v\rangle \leq I$. Similarly as before we shall denote the set of vectors (of production factors) satisfying this ,,budget" constraint by

$$D(x, I) = \{\, x \in \mathbb{R}^k_+ \mid \langle x, v\rangle \leq I \,\}.$$

If we assume that the firm chooses that production plan which brings the maximum profit under the given constraints, then we have the following.

Problem 2 (Maximization of the firms profit)**.**
It is sought for:

$$\max\{\Pi(x) = pf(x) - \langle x, v\rangle \mid x \in \mathbb{R}^k_+\}$$

under conditions:

$$x_i \geq 0, \quad \text{for } i = 1, \ldots, k, \quad \text{and} \quad \langle x, v\rangle = \sum_{i=1}^{k} x_i v_i \leq I$$

Similarly as before, suitable assumptions on the production function (suitable degree of differentiability, strict quasi-concavity and monotonicity) assure the existence of a unique solution to this problem, as formulated in the following theorem.

Theorem 2.2. *With the given set of (positive) prices v and given wealth I The Problem 2 of profit maximization has a unique solution $x^0 \in D(x, I)$ with positive coordinates, so that: for a unique positive constant (Lagrange multiplier) λ^0 the vektor x^0 is the unique solution of the system of equations*

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x) &= \frac{1}{p} v_i + \lambda^0 v_i, \qquad \textit{dla } i = 1, \ldots, k \\ \langle x, v\rangle &= I \end{aligned}$$

Equivalently, the both components of the solution which are x^0 and λ^0 are given as unique solution of the system of equations (with unknowns $x_1, \ldots, x_k, \lambda$)

$$\operatorname{grad} f(x) - (\frac{1}{p} + \lambda)v = 0 \tag{3}$$

$$I - \langle x, v\rangle = 0. \tag{4}$$

3 A uniform formulation of the problem

Similarities visible in the formulation of both problems lead us to describe the situation in the following way. We are given a system of $k+1$ equations,

$$\phi_i(x, \lambda, q, I) = 0, \qquad i = 1, \ldots, k+1 \tag{5}$$

where for uniformity we have put q in place of the price vectors previously denoted p or v, and the functions ϕ_i are given by

$$\phi_i(x, \lambda, q, I) = \begin{cases} \dfrac{\partial u}{\partial x_i}(x) - \lambda q_i, \text{ for } i = 1, \ldots, k, & \text{for the Problem 1} \\ \dfrac{\partial f}{\partial x_i}(x) - (\frac{1}{p} + \lambda)q_i, \text{ for } i = 1, \ldots, k, & \text{for the Problem 2} \end{cases} \tag{6}$$

and

$$\phi_{k+1}(x, \lambda, q, I) = I - \langle x, q\rangle, \qquad \text{for both Problems.} \tag{7}$$

We envisage ϕ_i as components of a vector function $\Phi : \mathbb{R}^{k+1}_+ \times \mathbb{R}^{k+1}_+ \ni (x, \lambda, q, I) \to \Phi(x, \lambda, q, I) \in \mathbb{R}^{k+1}$ defined by

$$\Phi(x, \lambda, q, I) = \begin{pmatrix} \phi_1(x, \lambda, q, I) \\ \vdots \\ \phi_{k+1}(x, \lambda, q, I) \end{pmatrix} \tag{8}$$

what enables us to write down the equations (5) in a form of one vector equation

$$\Phi(x, \lambda, q, I) = 0. \tag{9}$$

Here q has the meaning of the price vector (p or v in the former formulation), I is the wealth and the problem consists in determining (x, λ) from this equation.

Clearly this can be seen as an instance of an “Implicit Function Problem”, the question beeing that of defining the values (x, λ) as functions of (q, I). According to the Implicit Function Theorem, cf. [4], a necessary and sufficient condition for the existence and smoothness of the solution in a neigborhood of a given particular solution x_0, λ_0) is the nonvanishing of the (partial) Jacobian determinant of the function Φ with respect to $y = (x, \lambda)$. It is not difficult to check that the (partial) Jacobian matrix is given by

$$\frac{\partial(\phi_1, \dots, \phi_{k+1})}{\partial(y_1, \dots, y_{k+1})} = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_k} & \frac{\partial \phi_1}{\partial \lambda} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \phi_{k+1}}{\partial x_1} & \dots & \frac{\partial \phi_{k+1}}{\partial x_k} & \frac{\partial \phi_{k+1}}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} \mathcal{H} & -q \\ -q^t & 0 \end{pmatrix}, \tag{10}$$

where $\mathcal{H}$ is the Hessian of the function u, respectively f,

$$\mathcal{H} = \left[\frac{\partial^2 u}{\partial x_i \partial x_j}\right] \qquad \text{respectively} \qquad \mathcal{H} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]$$

and q is an appropriate price vector. One can further show that the Jacobian determinant of (10) is given by

$$\begin{vmatrix} \mathcal{H} & -q \\ -q & 0 \end{vmatrix} = -\mathcal{H}^{\mathrm{adj}}(q, q),$$

where $\mathcal{H}^{\mathrm{adj}}$ is the adjoint matrix to $\mathcal{H}$. By the assumption of strict quasi-concavity the Hessian is negative definite, hence the Jacobian determinant $-\mathcal{H}^{\mathrm{adj}}(q, q)$ is positive.

4 The Newton's Method of solving nonlinear vector equations

By virtue of the Implicit Function Theorem it remains now to establish existence of a particular solution (x_0, λ_0) for the equation (9). Rather than treating the problem abstractly in full generality we have resorted to a numerical procedure of approximating such solution by using the multi-dimensional Newton's Method. It relies on constructing an approximate solution to the equation (9) by means of recurrsively defined sequence

$$(x^{(n+1)}, \lambda^{(n+1)}) = (x^{(n)}, \lambda^{(n)}) - \Phi'(x^{(n)}, \lambda^{(n)})^{-1} \cdot \Phi(x^{(n)}, \lambda^{(n)})$$

with a initial value $(x^{(0)}, \lambda^{(0)})$ suitably chosen. It can be shown that for a large class of functions this sequence is convergent independently of the initial point and that the limit is a solution of (9). In fact, Newton's Method may be used to prove Implicit Function Theorem in its general form, cf. [5, 4].

It is worth mentioning that in many cases the full extend of the Newton's Method can be replaced by the so called Simplified Newton's Method allowing the use of the Banach Contraction Mapping Principle for a proof of convergence and estimates of the errors.

In the remaining part of the paper we investigate the use of the Newton's Method to solve optimization problems formulated above.

5 Numerical calculations

5.1 A class of separable utility functions

In order to avoid trivial complications with the wording of our results they are formulated below in the context of consumer theory. For our numerical calculations we have chosen examples coming from the class of what is called *separable utility functions*, which are given by the formula

$$u(x_1,\, x_2,\, \dots,\, x_k) = x_1^{\alpha_1} + x_2^{\alpha_2} + \dots + x_k^{\alpha_k}, \qquad \text{where } 0 < \alpha_l < 1.$$

The price vector is denoted by $(p_1,\, p_2,\, \dots,\, p_k)$ and the wealth denoted I. It is a matter of a trivial computation to see that the Hessian is a diagaonal matrix with negative entries along the diagonal, so the functions satisfy the necessary assumptions.

The numerical calculations are done for the following utility functions corresponding to $k = 2,\, 4,\, 10$ and 20 commodities, namely:

$$u^2(x) = x_1^{0.4} + x_2^{0.3}, \qquad u^4(x) = x_1^{0.4} + x_2^{0.4} + x_3^{0.3} + x_4^{0.3}, \tag{11}$$

$$u^{10}(x) = \sum_{i=1}^{5} x_i^{0.4} + \sum_{i=6}^{10} x_i^{0.3}, \quad u^{20}(x) = \sum_{i=1}^{5} x_i^{0.5} + \sum_{i=6}^{10} x_i^{0.4} + \sum_{i=11}^{15} x_i^{0.3} + \sum_{i=16}^{20} x_i^{0.2}. \tag{12}$$

The corresponding price vectors are taken to be equal to:

$$p^2 = (10,\, 10)\,, \qquad p^4 = (10,\, 40,\, 10,\, 40)\,, \tag{13}$$

$$p^{10} = (1,\, 5,\, 10,\, 20,\, 40,\, 1,\, 5,\, 10,\, 20,\, 40) \quad \text{and} \quad p^{20} = \big(p^{10},\, p^{10}\big) \tag{14}$$

with the budget constraint

$$\langle x,\, p^a \rangle \le I = 1000, \quad a = 2,\, 4,\, 10,\, 20. \tag{15}$$

5.2 The algorithm

The algorithm based on the Newton's method,

$$h = \Phi'(x^{(n)},\, \lambda^{(n)})^{-1} \cdot \Phi(x^{(n)},\, \lambda^{(n)}), \tag{16}$$

$$\big(x^{(n+1)},\, \lambda^{(n+1)}\big) = \big(x^{(n)},\, \lambda^{(n)}\big) - h, \tag{17}$$

does not ensure that the variables are positive at each stage of iteration. On the other hand, the utility functions (11, 12) and their partial derivatives are defined only for all $x_i > 0$, so the computation might be broken if some x_i becomes negative or zero. To eliminate this problem we slightly modify the method. At each step we check if all x components of the vector $\big(x^{(n)},\, \lambda^{(n)}\big) - h$ are positive. If yes, we follow the algorithm given by eqs. (16, 17), but if not, we divide the value of h by a factor of 2, $h = h/2$, and check components of $\big(x^{(n)},\, \lambda^{(n)}\big) - h$ once again. Repeating this we ensure that all x variables are positive and the algorithm can be used without break.

5.3 The results

The numerical calculations are performed with use of a program written in the Matlab language (see e.g. [6]). The initial values of the Lagrange multiplier are assumed $\lambda^{(0)} = 1$. The results are presented in Fig. 1.

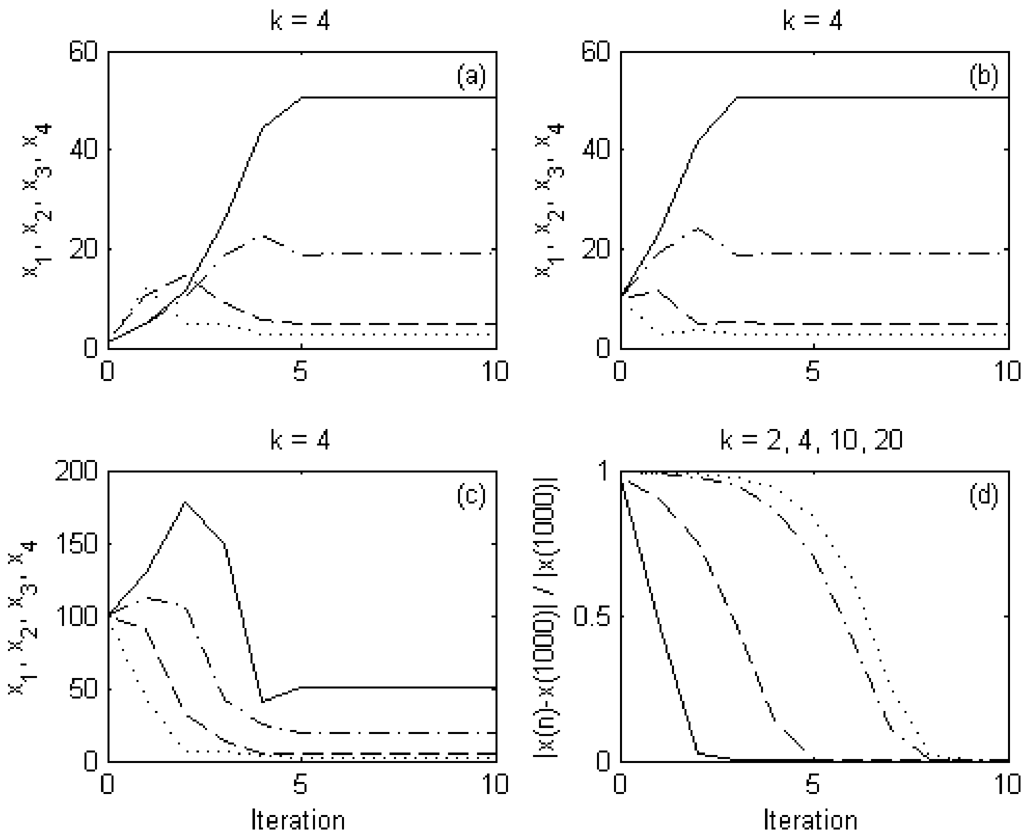

Figure 1: The numerical results of successive iterations versus the number of iteration. The arguments x_1 (solid line), x_2 (dashed line), x_3 (dash-dotted line) and x_4 (dotted line) of the utility function $u^4(x)$ (11) are shown for various starting points: $x_i^{(0)} = 1$ (a), $x_i^{(0)} = 10$ (b) and $x_i^{(0)} = 100$ (c). The relative norms (18) obtained for the utility functions $u^2(x)$ (solid line), $u^4(x)$ (dashed line), $u^{10}(x)$ (dot-dashed line) and $u^{20}(x)$ (dotted line) (11-12) are plotted in (d).

First, we consider the utility function $u^4(x)$ (11) which depends on four variables. In Fig. 1 the values of $x_{i=1,\,2,\,3,\,4}$ obtained in ten iterations are shown as functions of the number of iteration for a few starting points. The initial values $x_i^{(0)} = 1$ (Fig. 1a) satisfy the strict inequality $\langle x,\, p^2 \rangle < I$, the initial values $x_i^{(0)} = 10$ (Fig. 1b) lie on the budget line $\langle x,\, p^2 \rangle = I$, whereas the $x_i^{(0)} = 100$ (Fig. 1c) do not satisfy the budget limitation (15). In all these cases the algorithm achieves the same optimal solution in a few steps, so it seems to converge very quickly.

In order to test the speed of convergence we have devised the following testing method. The discussed algorithm starts from some arbitrary point $x^{(0)}$ and seeks the optimal solution. After sufficient number of steps the components of the vector x stabilise, i.e. achieve values which do not change during next steps, so one can conclude, that the solution was found with an accuracy possible in a numerical computing. We assume, that $x^{(1000)}$ obtained in 1000 iterations is equal to the exact solution of the optimization problem. The norm of the vector $x^{(n)} - x^{(1000)}$ is a measure of the distance between the solution and the vector $x^{(n)}$ obtained in n steps of iteration. To compare results of optimization for our utility functions (11-12) involving different numbers of commodities we consider this norm divided by the norm of the solution:

$$\frac{\left|x^{(n)} - x^{(1000)}\right|}{\left|x^{(1000)}\right|}. \tag{18}$$

The results are shown in Fig. 1d. The initial values of the quantity (18) are close to 1 due to the assumed starting points $x_i^{(0)} = 1$. During successive iterations their values decrease and become close to zero after several steps, what testifies that the vectors $x^{(n)}$ converge quickly to the optimal solutions for all the considered numbers of commodities: $k = 2$, 4, 10 and 20.

Note, that shapes of curves presented in Fig. 1d depend strongly on a choice of initial vectors $x^{(0)}$, exponents in the utility functions (11, 12) and prices (13, 14), but for wide range of values (not shown) the algorithm is numerically stable and lead to the solution of the considered optimization problems.[1]

6 Summary and further outlook

Results concerning numerical solution of the system (9) for a choice of utility functions with a varied number $k = 2, 4, 10, 20$ of variables (commodities) presented here clearly demonstrate the applicability of the Newton's Method for the study of optimization problems of Mathematical Economics. Even with the use of relatively simple devices, the successive approximations obtained from the Newton's Method converge quite rapidly to a solution without regard to the number of variables or a choice of the initial point. In a paper under preparation the authors udertake a numerical study of various expansions paths and Engel curves for a class of separable utility functions.

[1]Numerical problems occur when some component of the vector x has final value many orders of magnitude smaller than the others. In this case one can assume that this component is zero and do not enter the utility function, so an optimization problem with one less commodity should be considered.

References

[1] Handbook of Mathematical Economics, vol. II, K. J. Arrow and M. D. Intriligator, ed. North-Holland 1982

[2] E. Panek, **Ekonomia matematyczna** (in Polish), Akademia Ekonomiczna w Poznaniu, Poznań, 2000

[3] E. Panek (red.), **Podstawy ekonomii matematycznej**, Materiały do ćwiczeń. Wyd. trzecie, Akademia Ekonomiczna w Poznaniu, Poznań, 2001

[4] S. G. Krantz, H. R. Parks, **The Implicit Function Theorem: History, Theory, and Applications,** Birkhäuser, Boston, 2002

[5] W. Walter, **Analysis 1 & 2**, Springer-Verlag, Berlin 1992.

[6] A. Zalewski, R. Cegiełа, **Matlab - obliczenia numeryczne i ich zastosowania** (in Polish), Wydawnictwo NAKOM, Poznań, 2002