ISSN 2090-3359 (Print)
ISSN 2090-3367 (Online)

Advances in Decision Sciences

Volume 23
Issue 4
December 2019

Michael McAleer

Editor-in-Chief
University Chair Professor
Asia University, Taiwan

Published by Asia University, Taiwan

ADS@ASIAUNIVERSITY

Applications of the Newton-Raphson Method in Decision Sciences and Education*

Buu-Chau Truong

Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Nguyen Van Thuan

Department of Postgraduate Studies, Vinh University, Vinh City, Vietnam

Nguyen Huu Hau

Department of Training Management, Hong Duc University, Thanh Hoa City, Vietnam

Michael McAleer **

Department of Finance, Asia University, Taiwan
and
Discipline of Business Analytics, University of Sydney Business School, Australia
and
Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam
The Netherlands
and
Department of Economic Analysis and ICAE, Complutense University of Madrid, Spain
and
Institute of Advanced Sciences, Yokohama National University, Japan

December 2019

* For financial support, the fourth author is most grateful to the Australian Research Council and the Ministry of Science and Technology (MOST), Taiwan.
** Corresponding author: michael.mcaleer@gmail.com

Abstract

The Newton-Raphson (NR) method is one of the most important and popular methods to determine an optimal solution in many applications in the decision sciences and education. The NR method can be used for an optimal solution to obtain estimates in regression models, the maxima or minima of many functions in both the one-dimensional and multi-dimensional cases, or to solve systems of equations with many unknowns in both the one-dimensional and multi-dimensional cases. Applications of the NR method cover a wide variety of interesting cases, including the decision sciences and the teaching of mathematics-related subjects. The primary purpose of the paper is to provide a universal approach to the theory and practice of the NR method. In addition, we focus the discussion on applications in decision sciences, statistics, portfolio optimization, and education, among others Interesting potential research directions of the NR method are also discussed.

Keywords: Newton-Raphson method, optimization, missing data, statistics, decision sciences, teaching, education.

JEL: A10, G00, G31, O32.

1. Introduction

The Newton-Raphson (NR) method is one of the most important tools for determining the optimal solutions of many problems in different areas, including statistics, applied mathematics, numerical analysis, economics, finance, management, and marketing, among others. It well known that the NR method is also a common iterative method for finding the roots of an objective function, $f(x)$, that is, to obtain solutions for the equation $f(x) = 0$.

In optimization and statistics, the NR method is one of the most commonly used methods to find the roots of the derivative of a function, $f(x)$. There are many approaches that can be used to find optimal solutions, including the NR, secant, bisection, and gradient methods. Among them, the iterative NR method would seem to be the most common method.

In applications, Broyden (1967) presented the quasi-Newton method for functional minimization, Riks (1972) applied the NR method to the problem of elastic stability, Polyak (2007) developed some properties for the NR method and applied it to optimization, and Irvine (2010) applied the NR method to vibration problems. In addition, Derakhshandeh et al. (2016) introduced high-order Newton-like methods to solve some power flow equations.

Zhu et al. (2017) presented some applications of the NR method to lifetime spectroscopy for extraction of defect parameters. Pho and Nguyen (2018) compared the NR algorithm and Maxlik function to find optimal solutions in the estimation problem. Moreover, Dutto et al. (2019) extended the NR variants and compared their performance in solving some power-flow equations. Pho, Tran, Ho and Wong (2019) reviewed some optimal solution techniques.

Furthermore, many statisticians have applied the method to find optimal solutions in the estimating functions for different regression models with missing data. For instance, Wang et al. (2002) applied the NR method to the logistic regression model with missing covariates. Lukusa et al. (2016) employed the NR method for optimization in the zero-inflated Poisson (ZIP) regression model with missing covariates.

Hiesh et al. (2009, 2010) applied the NR method to semiparametric analysis of randomized response data with missing covariates in logistic regression. Lee et al. (2012) employed the NR method for semiparametric estimation of the logistic regression model with missing covariates and outcomes. Lee et al. (2016) applied the NR method in the estimation of the closed capture-recapture model when covariates are missing randomly.

In this paper, we provide a universal approach to the theory and practice of the NR method. Thereafter, we focus on discussing applications of the approach to various disciplines in the decision sciences such as statistics, portfolio optimization, and education. Furthermore, we also discuss some potential research directions for purposes of using the NR method.

The remainder of the paper is organized as follows. In Section 2, we discuss some algorithms for the NR method and give some examples by using the NR method in both the one- and multi- dimensional cases. Applications of the NR method in decision sciences and education are presented in Sections 3 and 4. Some concluding remarks are given in the final section.

2. The Newton-Raphson Method

In this section, we discuss how to apply the NR method to detect the optimal solutions for the one- and multi-dimensional cases.

a. One-dimension

We first present how to apply the NR method to detect the optimal solution for the following one-dimensional equation:

$$f(x) = 0$$

where $$f : D(D \subset R) \to R \qquad (2.1)$$

The NR method is use to derive the approximation of the following derivative of the function:

f in (2.1) at a point x_1:

$$f'(x_1) = \frac{df(x_1)}{dx} \approx \frac{f(x_1) - f(x_2)}{x_1 - x_2} \qquad (2.2)$$

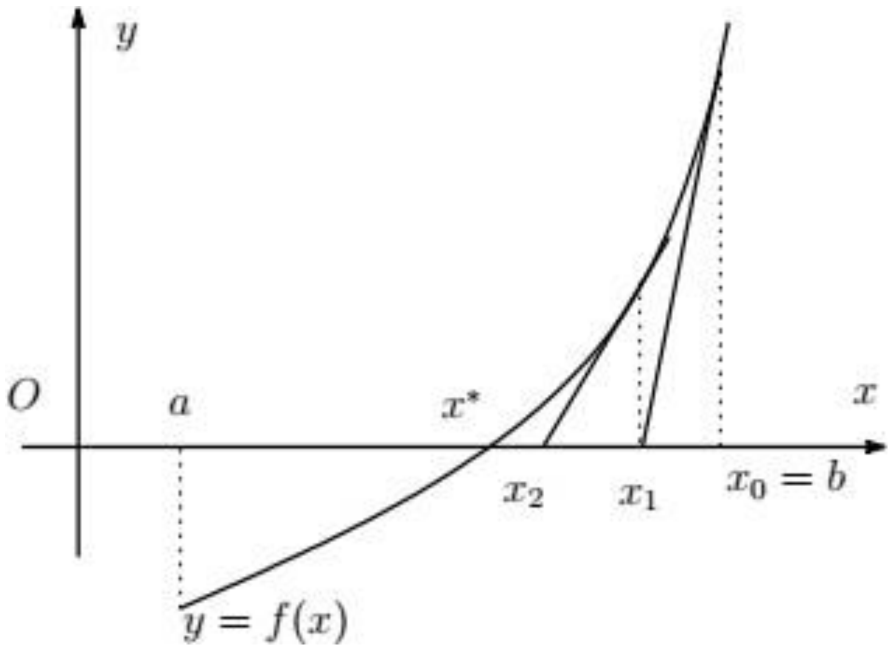

Figure 1: The NR method in the one-dimensional case

The N-R method for the one-dimensional case is illustrated in Figure 1.

Let x^* be an optimal solution of $f(x)$, such that $f(x^*) = 0$. Therefore, in (2.2), one needs to find x_2 such that $f(x_2) = 0$. Thus, we have:

$$x_2 \approx x_1 - \frac{f(x_1)}{f'(x_1)}, \qquad (2.3)$$

In general, the solution of the NR method is given by the following equation:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad (2.4)$$

where x_0 is an initial value. We repeat using equation (2.4) until the difference in two adjacent solutions is less than s, where s is a very small given value.

The tangent of the function $f(x)$ at $(x_k, f(x_k))$ is the line that has the equation $y = f'(x_k)(x - x_k) + f(x_k)$ and intersects the x−axis at $(x_{k+1}, 0)$. For ease of access and application of formula (2.4), we give the following example to illustrate the NR method for the one-dimensional case.

Example 2.1 We use the Newton method to find the optimization solution of the following equation: $f(x) = e^x - x - 20 = 0$.

Solution

The equation $f(x) = e^x - x - 20 = 0$ has a unique root in the $[0, 3.5]$ interval because $f(0).f(3.5) < 0$ and $f'(x) = e^x - 1 \geq 0$, $\forall x \in [0, 3.5]$. Applying the NR method, one can construct the following sequence:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{e^{x_n} - x_n - 20}{e^{x_n} - 1}, \; n = \overline{0, \infty}$$

With $x_0 = 3.5$, it can be seen that:

$x_1 = 3.200597333300494, \; f(x_1) = 0.346591338602703$

$x_2 = 3.143410407193708, \; f(x_2) = 0.039384557255474$

$x_3 = 3.141634952057482, \; f(x_3) = 0.000036517262917$

$x_4 = 3.141633302802458, \; f(x_4) = 0.000000000031477$

$x_5 = 3.141633302801037, \; f(x_5) = 0.000000000000004$

$x_6 = 3.141633302801037, \; f(x_6) = 0.000000000000004.$

As the fifth and sixth iterations have the same value up to the fifteen decimal place, one can conclude that 3.141633302801037 is a correct solution to $f(x) = 0$.

We note that in Example 2.1, if one chooses an initial value $x_0 = 0$, then x_1 will be undefined as $x_1 = 0 - \frac{19}{0}$. Thus, the sequence $\{x_n\}, n = \overline{0, \infty}$ will be also undefined. Therefore, choosing an initial value is very important in the iterations.

When using the NR method, academics and practitioners are interested in the convergence of the method. In order to meet such needs, we introduce a theorem about this issue.

Theorem 2.1 Convergence of the N-R Method in one-dimension

Supposing that f is a continuous real-valued function with continuous derivative f'' defined on $I_\delta = [x^* - \delta; x^* + \delta]$ with $\delta > 0$, satisfies $f(x^*) = 0$ and $f''(x^*) \neq 0$. Suppose further that there exists a positive constant A satisfies:

$$\left| \frac{f''(x)}{f'(x)} \right| \leq A, \forall x \in I_\delta .$$

If $|x^* - x_0| \leq h$ with $h = \min\left\{\delta, \frac{1}{A}\right\}$, then the sequence $\{x_n\}, n = \overline{0, \infty}$ defined in (2.4) will converge quadratically to x^*.

Applying Theorem 2.1, we have the following note.

Note 2.1 If $f'(x)$ is an objective function, then one can also apply the NR method to get its optimal solution. Assuming that x^* is an optimal solution of $f(x)$, one has $f'(x^*) = 0$. Using the NR method, one can set up the following sequence:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} \qquad (2.5)$$

where x_0 is an initial value.

We now provide an application of formula (2.5) in the following example:

Example 2.2 We apply the Newton method to find the optimal solution of the first derivative of the following equation:

$$f(x) = x^4 + 2x^2 - 30x + 20.$$

Solution

One has the objective function as the first derivative of $f(x) = x^4 + 2x^2 - 30x + 20$, with an optimal solution that is defined by the following sequence:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} = x_n - \frac{4x_n^2 + 4x_n - 30}{8x_n + 4} = x_n - \frac{2x_n^2 + 2x_n - 15}{4x_n + 2}.$$

With $x_0 = 2$, it is seen that:

$x_1 = 1.807692307692308$, $f(x_1) = -17.01706654178$, $f'(x_1) = 0.85912608101957$

$x_2 = 1.787811120824951$, $f(x_2) = -17.02566342996$, $f'(x_2) = 0.008542702197822$

$x_3 = 1.787609428983337$, $f(x_3) = -17.0256642915728$, $f'(x_3) = 0.0000087216$

$x_4 = 1.787609408374904$, $f(x_4) = -17.025664291524862$, $f'(x_4) = 0.0000000000016$

$x_5 = 1.787609408374904$, $f(x_5) = -17.025664291524862$, $f'(x_5) = 0$.

As the fourth and fifth iterations have the same value up to the fifteen decimal place, one can conclude that 1.787609408374904 is an optimal solution of $f'(x)$ with the optimal value given as -17.025664291524862.

We now discuss the NR method in a multi-dimensional case in the next sub-section.

b. Multi-dimension

In this section, we present how to construct the NR method to detect the optimal solution in the multi-dimensional case to solve the following vector of equations:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \text{ where } \mathbf{f} : \mathbf{D} \subset \mathbb{R}^n \to \mathbb{R}^n \qquad (2.6)$$

and $\mathbf{0}$ is a zero vector in $\mathbb{R}^n$. In order to solve the vector of equations in (2.6), we first state the following definition:

Definition 2.1 Let $\mathbf{f} = (f_1; f_2; \ldots; f_n)^T : \mathbf{D} \subset \mathbb{R}^n \to \mathbb{R}^n$ be continuous and defined in a neighborhood $\mathbf{N}(\xi)$ of $\xi \in \mathbb{R}^n$, the vector of the first partial derivatives $\frac{\partial f_i}{\partial x_j}; i, j = \overline{1,n}$ exist at ξ, and the Jacobian matrix $J_{\mathbf{f}}(\xi)$ of $\mathbf{f}$ at ξ be the $n \times n$ matrix with elements:

$$\left[J_{\mathbf{f}}(\xi)\right]_{ij} = \frac{\partial f_i}{\partial x_j}(\xi)$$

The recursion is defined by

$$\mathbf{x}_{n+1} = \mathbf{x}_n - [J_{\mathbf{f}}(\mathbf{x}_n)]^{-1} \mathbf{f}(\mathbf{x}_n), n = \overline{0, \infty} \qquad (2.7)$$

where $x_0 \in \mathbf{D}$ (an initial value) is called the NR method for the system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

It is implicitly assumed that $\{\mathbf{x}_n\} \subset$ D, $n = \overline{0, \infty}$ and the matrix $J_{\mathbf{f}}(\mathbf{x}_n)$ is non-singular for each $n \in \mathbb{N}$. The explanation of (2.7) is to approximate $\mathbf{f}(\mathbf{x})$ near current iteration $\mathbf{x}_n$ by a function

$\mathbf{h}_n(\mathbf{x})$ satisfies the system of equations $\mathbf{h}_n(\mathbf{x}) = \mathbf{0}$ his easy to solve. Hence, we can use the solution in the next iteration $\mathbf{x}_{n+1}$, and repeat the process until we obtain the solution.

A suitable choice for $\mathbf{h}_n(\mathbf{x})$ is a linear approximation of $\mathbf{f}(\mathbf{x})$ at $\mathbf{x}_n$, whose graph is the tangent of $\mathbf{f}(\mathbf{x})$ at $\mathbf{x}_n$:

$$\mathbf{h}_n(\mathbf{x}) = \mathbf{f}(\mathbf{x}_n) + J_{\mathbf{f}}(\mathbf{x}_n)(\mathbf{x} - \mathbf{x}_n) \qquad (2.8)$$

Thus we solve the system of equations by using the following equation:

$$0 = \mathbf{h}_n(\mathbf{x}_{n+1}) = \mathbf{f}(\mathbf{x}_n) + J_{\mathbf{f}}(\mathbf{x}_n)(\mathbf{x}_{n+1} - \mathbf{x}_n) \qquad (2.9)$$

Using equation (2.9), one can easily obtain the solution for equation (2.7).

The NR method is presented in (2.7) by using the inverse of the Jacobian matrix. For sake's simplicity. In the practical application, equation (2.7) is usually written in the following form:

$$J_{\mathbf{f}}(\mathbf{x}_n)(\mathbf{x}_{n+1} - \mathbf{x}_n) = -\mathbf{f}(\mathbf{x}_n) . \qquad (2.10)$$

Given a vector $\mathbf{x}_n$, one needs to compute $\mathbf{f}(\mathbf{x}_n)$ and its Jacobian matrix $J_{\mathbf{f}}(\mathbf{x}_n)$, and thereafter, solve the system of equations in (2.10) by using Gaussian elimination. This gives the solution $\mathbf{x}_{n+1} - \mathbf{x}_n$ which is added to $\mathbf{x}_n$ to obtain a new iteration $\mathbf{x}_{n+1}$.

Similar to our discussion in the one-dimensional case, when using the NR method in the multi-dimensional case, academics and practitioners are also interested in the convergence of the method. In order to meet such needs, we introduce the following theorem to state the convergence of the NR method in the multi-dimensional case:

Theorem 2.2 Convergence of the N-R Method in the multi-dimensional case

Suppose that $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$ in some (open) neighbuorhood $\mathbf{N}(\mathbf{x}^*)$ of $\mathbf{x}^*$, in which all the second-order partial derivatives of $\mathbf{f}$ are continuous and $J_{\mathbf{f}}(\mathbf{x}^*)$ of $\mathbf{f}$ is non- singular at the point $\mathbf{x}^*$. Therefore::

1. The sequence $\{\mathbf{x}_n\}, n = \overline{0, \infty}$ is presented in (2.7) converges to the solution $\mathbf{x}^*$, provided that $\mathbf{x}_0$ is sufficiently close to $\mathbf{x}^*$; and

2. The convergence of the sequence $\{\mathbf{x}_n\}, n = \overline{0, \infty}$ to $\mathbf{x}^*$ is at least quadratic.

The NR method can be used to solve the system of nonlinear equations. We now provide an application of this issue in the following example:

Example 2.3 Consider the system of nonlinear equations:

$$\begin{cases} f_1(x; y; z) = x^2 + y^2 + z^2 - 1 = 0 \\ f_2(x; y; z) = x^2 + 2y^2 - 4z = 0 \\ f_3(x; y; z) = x^2 - 4y + 3z^2 = 0 \end{cases}$$

Let $\mathbf{f} = (f_1; f_2; f_3)^T$ and $\mathbf{x} = (x; y; z)^T$, so tjhat we now determine the optimal solution to the system of equations **f** (**x**) = **0** on the domain $\mathbf{D} = \{\mathbf{x} = (x; y; z)^T \in \mathbb{R}^3 : x, y, z \geq 0\}$ in $\mathbb{R}^3$.

Solution

The Jacobian matrix of **f** at $\mathbf{x} = (x; y; z)^T$ is given by:

$$J_{\mathbf{f}}(\mathbf{x}) = \begin{pmatrix} 2x & 2y & 2z \\ 2x & 4y & -4 \\ 2x & -4 & 6z \end{pmatrix}$$

As the first equation illustrates a sphere with radius 1 at center $O = (0; 0; 0)^T$ and the second and third equations are described as elliptic and paraboloid with axes aligned with the coordinate semi-axes $(0; 0; z)^T$, $z \geq 0$, and $(0; y; 0)^T$, $y \geq 0$, respectively, the point of intersection of three surfaces belongs to the unit cube $[0;1]^3$.

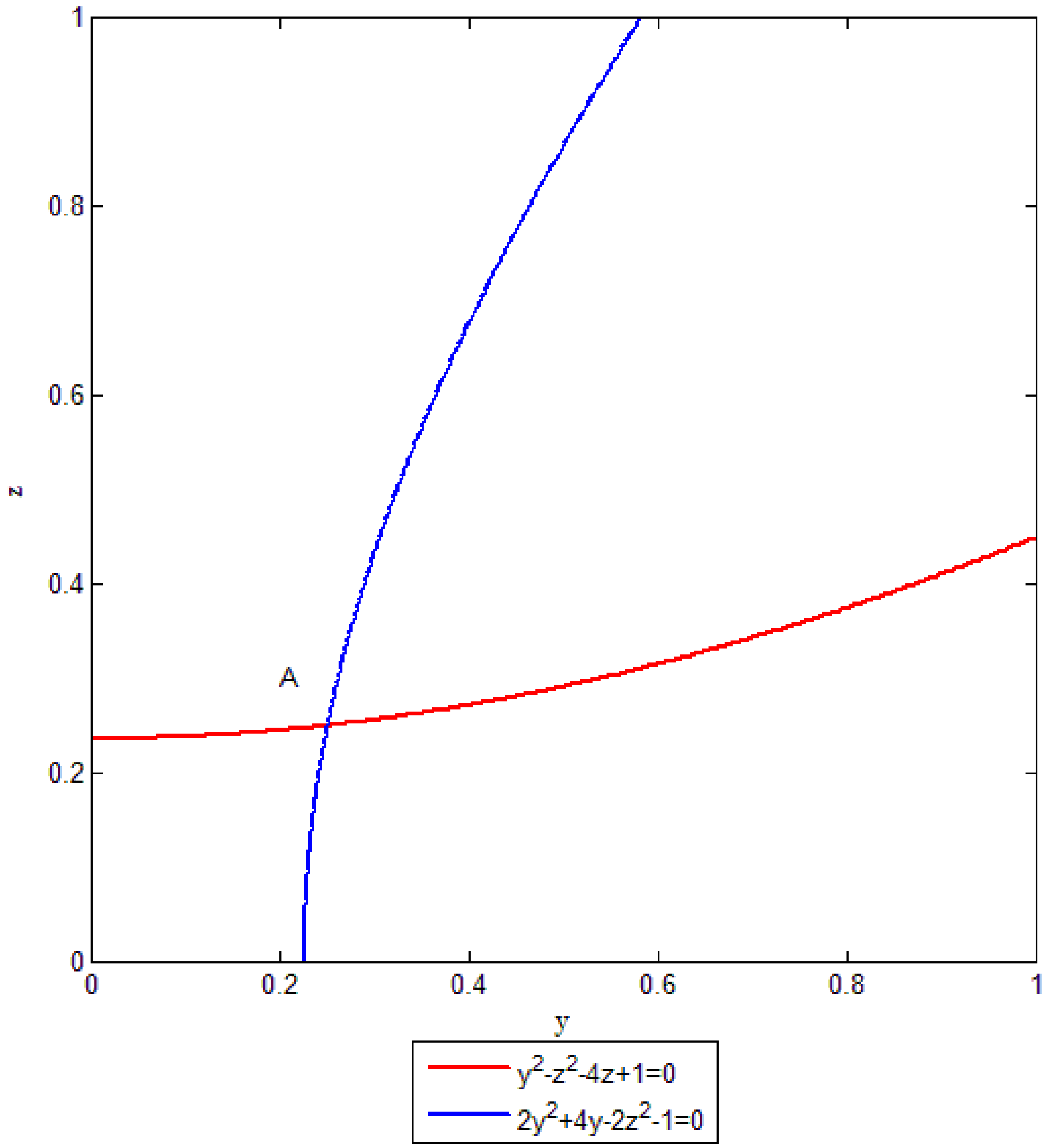

Figure 2: Intersection between $y^2 - z^2 - 4z + 1 = 0$ **and** $2y^2 + 4y - 2z^2 - 1 = 0$

We denote the solution by $\mathbf{x}^*$. It can be observed that the intersection of the first and second surface is a curve whose projection on to the (y; z) plane has the equation $y^2 - z^2 - 4z + 1 = 0$, while the intersection of the first and third surfaces is a curve whose projection on to the (y; z) −plane has the equation: $2y^2 + 4y - 2z^2 - 1 = 0$.

The two curves are shown in Figure 2.

The y − and z − coordinates of A are y ≈ 0.2 and z ≈ 0.2, respectively.

It is reasonable to choose an initial value $\mathbf{x}_0 = (0.9; 0.2; 0.2)^T$. Using the NR method, we construct the following sequence: $\mathbf{x}_{n+1} = \mathbf{x}_n - [J_{\mathbf{f}}(\mathbf{x}_n)]^{-1}\mathbf{f}(\mathbf{x}_n),\ n = \overline{0,\infty}$.

Suppose that, $\mathbf{x}_n = (x_n; y_n; z_n)^T$, for $n = \overline{0,\infty}$.

From (2.11) we have:

$$\mathbf{x}_1 = (0.935527544351074;\ 0.263865546218487;\ 0.251\text{[illegible]})^T$$

$$\mathbf{x}_2 = (0.931075400196031;\ 0.264205777527510;\ 0.251\text{[illegible]})^T$$

$$\mathbf{x}_3 = (0.931064613817713;\ 0.264205809984311;\ 0.251\text{[illegible]})^T$$

$$\mathbf{x}_4 = (0.931064613755233;\ 0.264205809984311;\ 0.251\text{[illegible]})^T$$

$$\mathbf{x}_5 = (0.931064613755233;\ 0.264205809984311;\ 0.251\text{[illegible]})^T$$

As $\mathbf{f}(\mathbf{x}_5) = 10^{-15}(0.222044604925031; 0.166533453693773)^T$, the vector $\mathbf{x}_5$ can be considered to be a satisfactory approximation to the required solution $\mathbf{x}^*$.

We now provide the following note:

Note 2.2 If $\mathbf{f}: \mathbf{D}\ (\mathbf{D} \subseteq \mathbb{R}^n) \to \mathbb{R}$ is an objective function, then one can use the NR method to find its optimal solution after defining the gradient and Hessian matrix of the function $\mathbf{f}$.

We now provide the following definition:

Definition 2.2 (The gradient and Hessian matrix)

Suppose that $\mathbf{f} : \mathbf{D}\ (\mathbf{D} \subset \mathbb{R}^n) \to \mathbb{R}$, if all the first partial derivatives of $\mathbf{f}$ exist on $\mathbf{D}$, then the gradient $\nabla\mathbf{f}$ of $\mathbf{f}$ is a vector defined as: $\nabla\mathbf{f} = \left(\frac{\partial \mathbf{f}}{\partial x_1}; \frac{\partial \mathbf{f}}{\partial x_2}; \ldots; \frac{\partial \mathbf{f}}{\partial x_n}\right)^T$

In addition, if all the second partial derivatives of $\mathbf{f}$ are continuous on $\mathbf{D}$, then the Hessian matrix $\mathbf{H}_\mathbf{f}$ of $\mathbf{f}$ is a square $n \times n$ matrix defined as:

$$\mathbf{H}_\mathbf{f} = \begin{pmatrix} \frac{\partial^2 \mathbf{f}}{\partial x_1^2} & \frac{\partial^2 \mathbf{f}}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \mathbf{f}}{\partial x_1 \partial x_n} \\ \frac{\partial^2 \mathbf{f}}{\partial x_2 \partial x_1} & \frac{\partial^2 \mathbf{f}}{\partial x_2^2} & \cdots & \frac{\partial^2 \mathbf{f}}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 \mathbf{f}}{\partial x_n \partial x_1} & \frac{\partial^2 \mathbf{f}}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \mathbf{f}}{\partial x_n^2} \end{pmatrix}$$

the gradient $\nabla\mathbf{f}$ of $\mathbf{f}$ is a vector defined as: $\nabla\mathbf{f} = \left(\frac{\partial \mathbf{f}}{\partial x_1}; \frac{\partial \mathbf{f}}{\partial x_2}; \ldots; \frac{\partial \mathbf{f}}{\partial x_n}\right)^T$

Assuming that $\mathbf{x}^*$ is an optimal solution of $\mathbf{f}$, one obtains $\nabla\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. In order to find an approximation of $\mathbf{x}^*$, by using (2.7), we construct the following sequence:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - [\mathbf{H}_\mathbf{f}(\mathbf{x}_n)]^{-1} \nabla\mathbf{f}(\mathbf{x}_n), n = \overline{0, \ldots} \qquad (2.12)$$

with $\mathbf{x}_0$ is sufficiently close to $\mathbf{x}^*$ and $\mathbf{x}_0 \in \mathbf{D}$.

In order to find a minimization of $\mathbf{f}$, the NR method approximates $\mathbf{f}$ by its quadratic approximation near $\mathbf{x}_n$:

$$\mathbf{f}_n(\mathbf{x}) = \mathbf{f}(\mathbf{x}_n) + (\mathbf{x}-\mathbf{x}_n)^T \nabla\mathbf{f}(\mathbf{x}_n) + (1/2)(\mathbf{x}-\mathbf{x}_n)^T \mathbf{H}_{\mathbf{f}}(\mathbf{x}_n)(\mathbf{x}-\mathbf{x}_n) \quad (2.13)$$

and then we compute the unique critical point of $\mathbf{f}_n(\mathbf{x})$, which is the unique solution of equation $\nabla\mathbf{f}_n(\mathbf{x}) = \mathbf{0}$. From (2.13) we obtain:

$$\begin{cases} \nabla\mathbf{f}_n(\mathbf{x}) = \nabla\mathbf{f}(\mathbf{x}_n) + \mathbf{H}_{\mathbf{f}}(\mathbf{x}_n)(\mathbf{x}-\mathbf{x}_n) \\ \mathbf{H}_{\mathbf{f}_n}(\mathbf{x}) = \mathbf{H}_{\mathbf{f}}(\mathbf{x}_n) \end{cases} \quad (2.14)$$

Hence, we have:

$$\begin{cases} \nabla\mathbf{f}_n(\mathbf{x}_n) = \nabla\mathbf{f}(\mathbf{x}_n) \\ \mathbf{H}_{\mathbf{f}_n}(\mathbf{x}_n) = \mathbf{H}_{\mathbf{f}}(\mathbf{x}_n) \end{cases} \quad (2.15)$$

If $\mathbf{H}_{\mathbf{f}}(\mathbf{x}_n)$ is a positive definite matrix, then this critical point is also guaranteed to be the unique strict global minimizer of $\mathbf{f}_n(\mathbf{x})$. In addition, we state an important result in the following theorem:

Theorem 2.3

For $n = \overline{0,\infty}$, the point $\mathbf{x}_{n+1}$ defined by (2.12) is the unique strict global minimizer of $\mathbf{f}_n(\mathbf{x})$

The NR method for finding the minimization of $\mathbf{f}$ is illustrated in Figure 3.

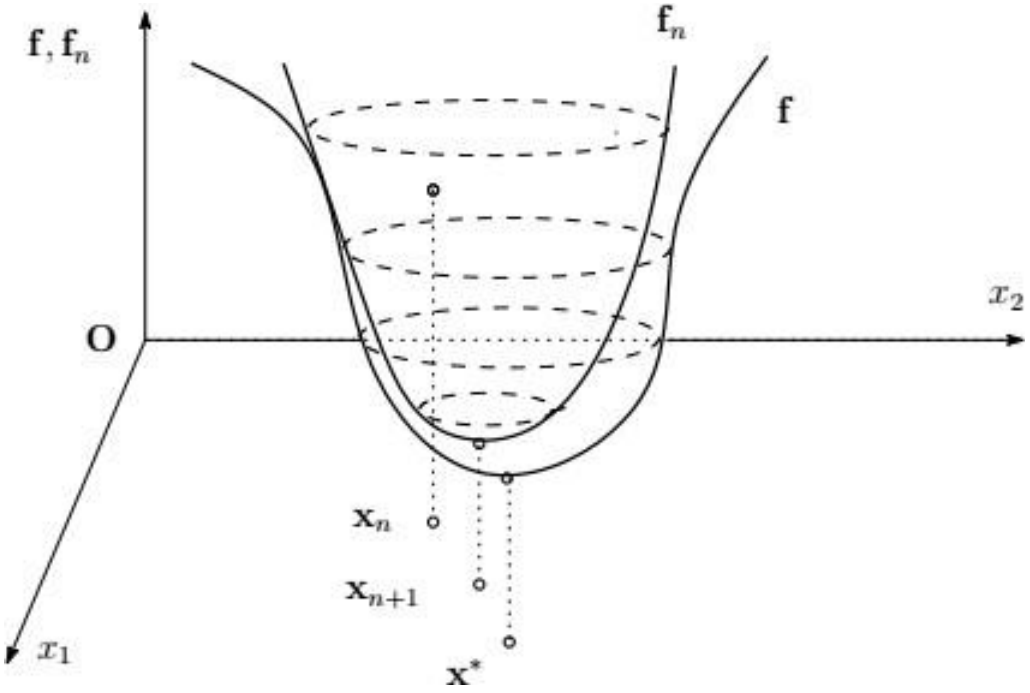

Figure 3: The NR method to find the minimization of f

We provide an example of the issue in the following example:

Example 2.4 We apply the NR method to find the optimal solution of the following equation:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x; y) = 18\sin(x+y) + 4y^2 - 6xy - 3x - 2y + 2.$$

Solution

As $\nabla \mathbf{f}(x; y) = \left(18\cos(x+y) - 6y - 3; 18\cos(x+y) + 8y - 6x - 2\right)^T$,

$$\mathbf{H_f}(x; y) = \begin{pmatrix} -18\sin(x+y) & -18\sin(x+y)+6 \\ -18\sin(x+y)+6 & -18\sin(x+y)+8 \end{pmatrix}$$

based on the contour plot (Figure 4) and graph (Figure 5) of **f** (x; y), we can show that the optimal solution $\mathbf{x}^*$ of $\mathbf{f}(x; y)$ is near the point A(−1; −1). Thereafter, by using the NR method, we construct the following sequence:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - [\mathbf{H_f}(\mathbf{x}_n)]^{-1} \nabla \mathbf{f}(\mathbf{x}_n), n = \overline{0, n} \qquad (2.16)$$

With $\mathbf{x}_0 = (-1; -1)^T$. From (2.16) it can be seen that

$\mathbf{x}_1 = (-1.033302112316977;\ -0.514272333850133)^T$

$\mathbf{x}_2 = (-1.055068707167308;\ -0.523600874500275)^T$

$\mathbf{x}_3 = (-1.055063872828223;\ -0.523598802640667)^T$

$\mathbf{x}_4 = (-1.055063872828369;\ -0.523598802640730)^T$

$\mathbf{x}_5 = (-1.055063872828369;\ -0.523598802640730)^T$.

Hence, we obtain $\mathbf{x}^* \approx \mathbf{x}_5 = (-1.055063872828369;\ -0.523598802640730)^T$.

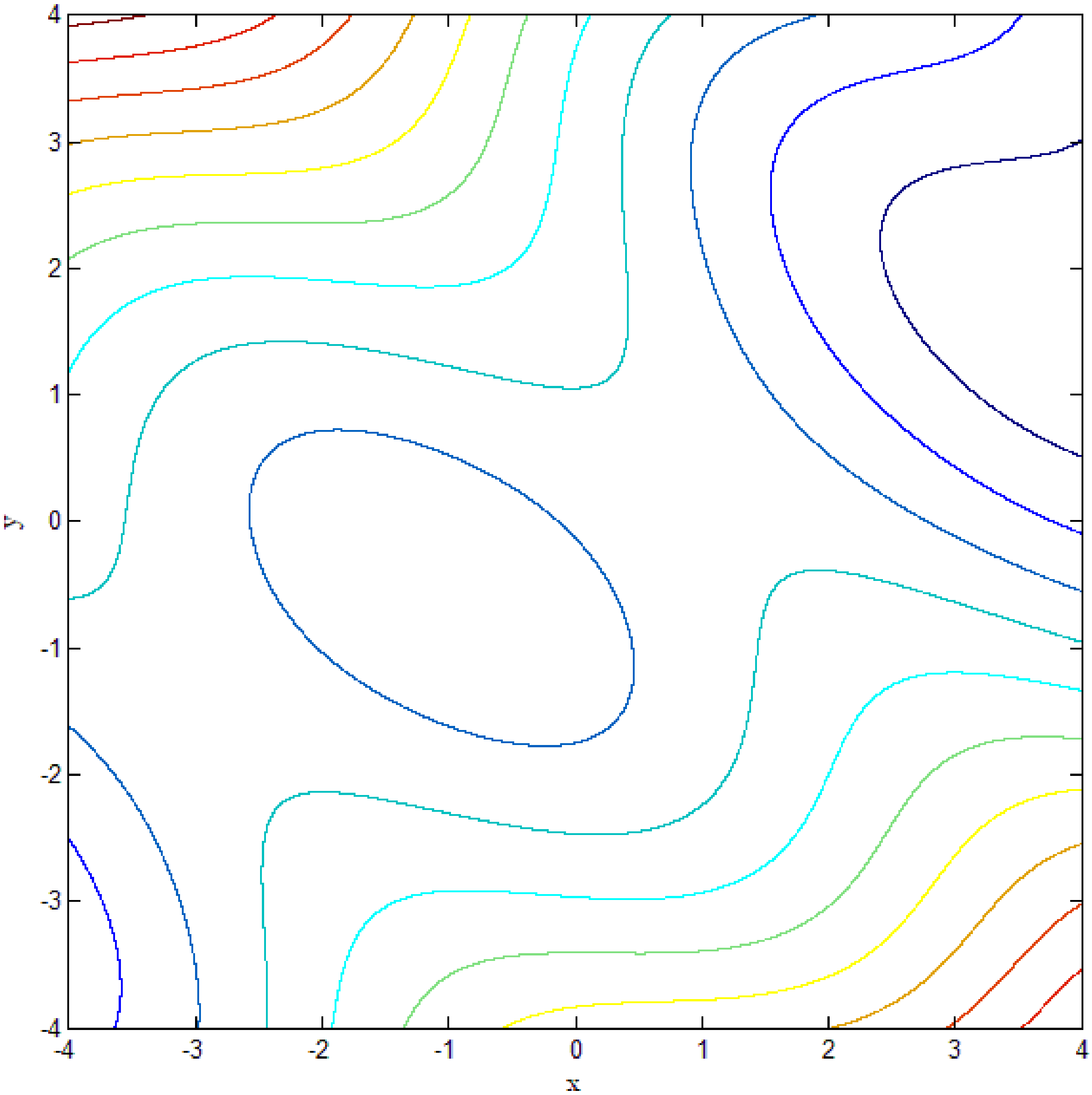

Figure 4: Contour plot of $f(x;y) = 18\,sin(x+y) + 4y^2 - 6xy - 3x - 2y + 2$.

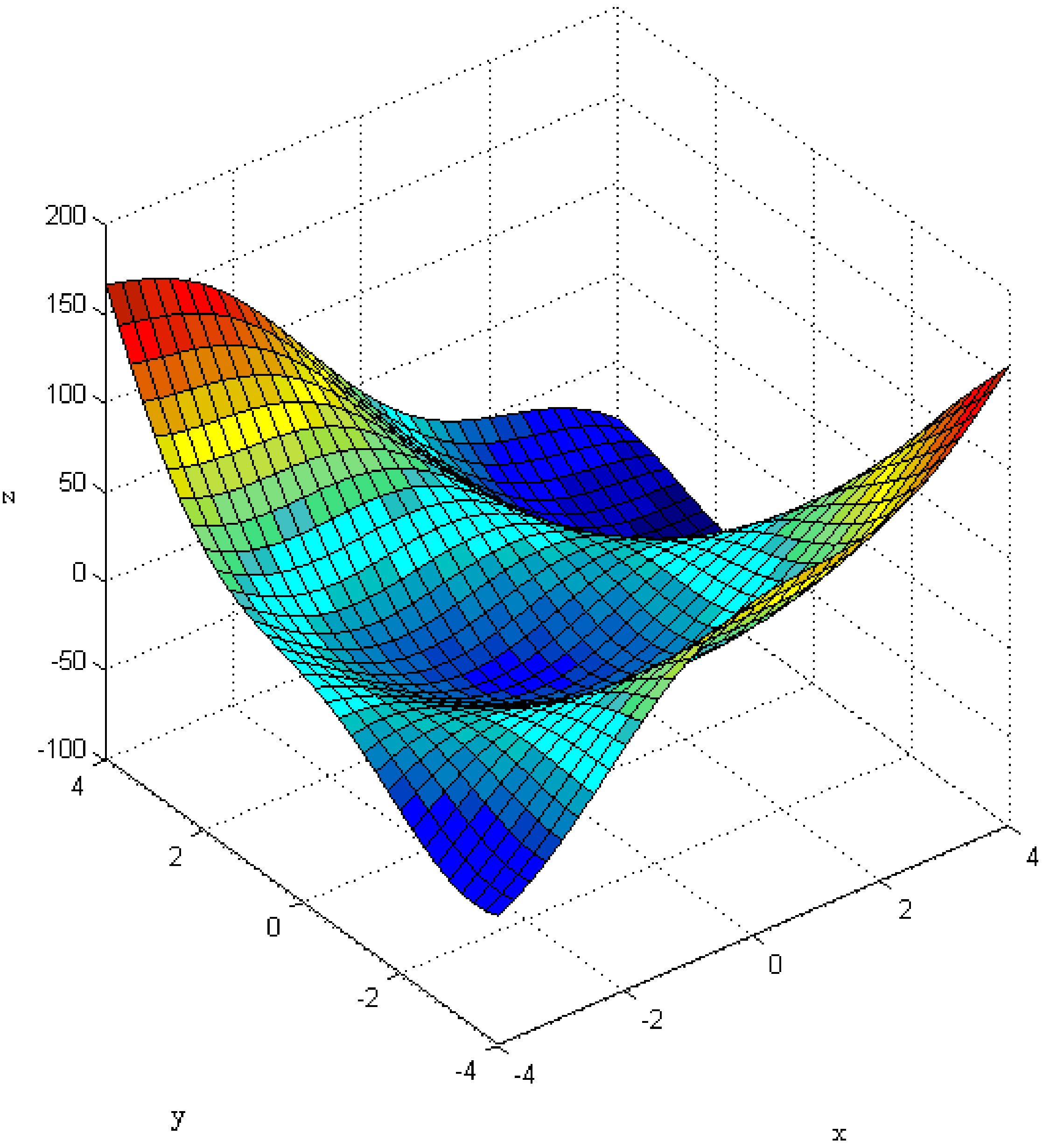

Figure 5: Graph of $f(x;y) = 18\sin(x+y) + 4y^2 - 6xy - 3x - 2y + 2$.

We discuss applications of the NR method in the decision sciences in the next section.

3. Applications of the NR method in Decision Sciences

In this section, we discuss several different applications of the NR method to decision sciences such as statistics, portfolio optimization, and related fields. The first field of application of the NR method in decision sciences is statistics. We consider two cases: the first case is a data set without missing values, and the second case considers the presence of missing values.

a. Statistics

i. Data set without missing values

Applications of the NR method in statistics focusmainly on estimating regression models. For instance, Thall (1988) used the NR method to estimate the parameters in the mixed Poisson likelihood regression models for longitudinal interval count data. Alho (1990) employed the NR method to estimate the parameters in logistic regression in capture-recapture models. Pho and Nguyen (2018) have mentioned how to apply the NR method to estimate the parameters in the zero-inflated binomial (ZIB) regression models in the case where the data set does not contain any missing value.

The NR method has not only been applied to traditional regression models such as linear, logistic, softmax, binomial, negative binomial, and Poisson regressions, but can also be used in regression models with missing data, which that has been gaining traction over the last two decades. The NR method is a robust method to find the optimal solution for the estimating function in the case where the data set contains missing values. We discuss the issue in the next subsection.

ii. Data set have missing values

The problem of missing observations is commonly encountered in the areas of, for example, transportation, health, economics, and finance. There are numerous reasons for missing observations, for example, respondents do not respond to certain items in answering a questionnaire in a survey, non-acceptance as answers, and incomprehensible answers (see Schafer, 2002). Readers may refer to Little (1992) and Pho et al. (2019) for methods to solve the problem of missing observations.

The problems of missing data can be classified as two different types: missing covariates and missing outcomes. The issue of estimating the parameters in regression models with missing data have been largely studied by many academics, and such approaches have been used extensively in many areas.

In this regard, many scholars have employed the method to find the optimal solution in estimating regression models with missing data. For example, Wang et al. (2002) applied the NR method for optimization in the logistic regression model with missing covariates. Lukusa et al. (2016) used the NR method for optimization in the zero-inflated Poisson (ZIP) regression model with missing covariates. In addition, Hsieh et al. (2009, 2010) and Lee et al. (2012, 2016) used this method in their respective analyses.

To the best of our knowledge, estimation of the parameters of regression models with missing data in cases that include the zero-inflated negative binomial (ZINB) regression models, the zero-inflated generalized Poisson (ZIGP) regression model, the zero-inflated power series (ZIPS) model, and the multivariate zero-inflated model have not been investigated in the literature. For example, Lukusa et al. (2016) observed that regression models with missing data have not yet been investigated in the literature. In order to bridge the gap in the literature, we now extend the NR method to estimate the parameters of regression models with missing data.

In addition, there are some other regression models, including the zero-inflated Bernoulli (ZIBer) regression model and the probit regression model, with missing data that have not yet been investigated in the literature. This could be a potential extension of this paper in future research. The structures of these regression models are discussed in Lambert (1992), Hall (2000), Diop et al. (2011), and Diallo et al. (2017), among others.

b. Portfolio Optimization

The applications the NR method in portfolio optimization is diverse and extremely rich. For instance: Chen (2000) introduced NR-type methods for stochastic programming. Agarwal et al. (2006) suggested algorithms for portfolio management based on the NR method. Bartholomew-Biggs et al. (2009) provided a global optimization problem in portfolio selection. Tong et al. (2010) presented a smoothing method for solving portfolio optimization with CVaR and applications in the allocation of generation assets.

Kumar and Mishra (2017) developed portfolio optimization using a novel covariance guided Artificial Bee Colony algorithm. In addition, scholars may refer to Barro and Canestrelli (2005), Calafiore (2007) and Hazan (2016), among others, for further interesting applications.

c. Other Fields

Applications in decision sciences of the NR method are varied and abundant. For instance, Chang, McAleer and Wong (2017) presented the connections of decision sciences with management information and financial economics. Chang, McAleer and Wong (2018) offered the connections of decision sciences to some related fields, such as economics, finance, business, and big data.

Pho, Ho, Tran and Wong (2019) introduced applications of the distribution functions in statistics to decision sciences. Pho, Tran, Ho and Wong (2019) offered applications of the optimization solution to decision sciences. Furthermore, readers may refer to Mahmoudi et al. (2019), Tian et al. (2019), Truong et al. (2019a, b), and Tuan et al. (2019), among others, for related interesting applications.

We discuss applications of the NR method in education in the next section.

4. Applications of the NR method in Education

a. Teaching situation in Vietnam

Education plays an abundantly vital role in the current industrialized era. Higher Education plays a key role in helping students to access science and technology, and to choose a solid future career. Mathematics plays a vital role in the higher education sector, and is a bridge for all scientific disciplines. The NR method is a powerful tool for finding optimal solutions. This is also a fundamental method for scientists to program software to find optimal solutions quickly and accurately.

To the best of our knowledge, the NR method is taught in almost all universities in Vietnam to students who are and are not majoring in mathematics. Universities in Vietnam that have been teaching the NR method include Can Tho University, Hong Duc University, Vinh University, Ton Duc Thang University, Ho Chi Minh University of Industry, and Hutech University Ho Chi Minh, among others. Thus, it can be seen that the NR method has been recognized by universities for its practicality, importance, and influence.

b. Effective teaching methods for students

In today's industrial age, education plays an extremely important and meaningful role. Innovative teaching methods for high school and university students is essential and urgent. The NR method is a powerful tool for finding optimal solutions, and it is taught in most universities in Vietnam and worldwide. The STEM and STEMTech education models have proven effective in the current situation. STEM is an English abbreviation for "Science, Technology, Engineering and Mathematics".

Tsupros et al. (2009) presented "STEM education is a multidisciplinary way to learning where rigorous academic concepts are coupled with real-world lessons as students apply Science, Technology, Engineering, and Mathematics in contexts that make connections between school, community, work, and the global enterprise enabling the development of STEM literacy and with it the ability to compete in the new economy".

Many authors have introduced new teaching methods to meet the current situation such as: Tuan, Pho, Huy and Wong (2019) have improved the STEM model to the STEMTech model, which emphasizes Technology. Hau, Tuan, Thao and Wong (2019) introduced a new trend-based teaching method for high school students by modelling practical problems.

From these ideas, in order to teach effectively the NR method to students, we recommend the following useful steps:

1. Student-centered discussion
2. Make the lessons connected to reality
3. Increase student autonomy
4. Build relationships in the classroom
5. Improve student's reading abilities.

We introduce some of applications of the NR method in practice in the next sub-section.

c. Applications of the NR method in practice

The application of the NR method in practice is varied and abundant. Practical applications of this method makes it easy for students to visualize the diverse applications. Wang et al. (2002) used the NR method to estimate the parameters of the logistic model with a real data set of bladder cancer conducted at the Fred Hutchinson Cancer Research Center in 1990. Lee et al. (2012) employed the NR method to estimate the parameters of the logistic model with data from a 2004 cable TV survey study in Taiwan.

Lukusa et al. (2016) have used the NR method to estimate the parameters in the ZIP regression model with a real data set involving violation of traffic laws in Taiwan in 2007. Tuan, Pho, Lam and Wong (2019) used the NR method to calculate the shooting angles in simulations of a shot machine in the Stemtech model. Hau, Tuan, Thao and Wong (2019) used the NR method to find the solution of a system of equations with four unknowns in finding the location of machines to collect bird's nest.

5. Conclusion

In this paper, we have provided a universal approach to the theory and practice of the NR method. In addition, we focused on discussing applications of the NR approach to decision sciences, such as statistics, portfolio optimization, economics and finance, statistical and econometric models, and education. Furthermore, we also introduced some potential research directions for purposes of using the NR method in future research and teaching.

In addition, to the best of our knowledge, estimation of the parameters of regression models with missing data in cases such as the zero-inflated negative binomial (ZINB) regression model, the zero-inflated generalized Poisson (ZIGP) regression model, the zero- inflated power series (ZIPS) model, and the multivariate zero-inflated model, have not yet been investigated in the literature. In order to bridge the gap in the literature, we extended the NR method to estimate the parameters of such regression models with missing data.

References

Alho, J.M. (1990), Logistic regression in capture-recapture models, Biometrics, 46, 623-635.

Agarwal, A., Hazan, E., Kale, S., Schapire, R.E. (2006), Algorithms for portfolio management based on the Newton method, in Proceedings of the 23rd international conference on Machine Learning, 9-16.

Barro, D., Canestrelli, E. (2005), Dynamic portfolio optimization: Time decomposition using the maximum principle with a scenario approach, European Journal of Operational Research, 163(1), 217-229.

Bartholomew-Biggs, M.C., Kane, S.J. (2009), A global optimization problem in portfolio selection, Computational Management Science, 6(3), 329-345.

Broyden, C.G. (1967), Quasi-Newton methods and their application to function minimization, Mathematics of Computation, 21(99), 368-381.

Calafiore, G.C. (2007), Ambiguous risk measures and optimal robust portfolios. SIAM Journal on Optimization, 18(3), 853-877.

Chang, C.-L., McAleer, M., Wong, W.K. (2017), Management information, decision sciences, and financial economics: A connection, Journal of Management Information and Decision Sciences, 20, 1-19.

Chang, C.L., McAleer, M., Wong, W.K. (2018), Decision sciences, economics, finance, business, computing, and big data: Connections, Advances in Decision Sciences, 22, 1-58.

Chen, X. (2000), Newton-type methods for stochastic programming. Mathematical and computer modelling, 31, 89-98.

Derakhshandeh, S.Y., Pourbagher, R. (2016), Application of high-order Newton-like methods to solve power flow equations, IET Generation, Transmission & Distribution, 10(8), 1853-1859.

Diallo, A.O., Diop, A., Dupuy, J.F. (2017), Asymptotic properties of the maximum-likelihood estimator in zero-inflated binomial regression, Communications in Statistics - Theory and Methods, 46(20), 9930-9948.

Diop, A., Diop, A., Dupuy, J F. (2011), Maximum likelihood estimation in the logistic regression model with a cure fraction, Electronic Journal of Statistics, 5, 460-483.

Dutto, S., Masetti, G., Chiaradonna, S., Di Giandomenico, F. (2019), On extending and comparing Newton-Raphson variants for solving power-flow equations, IEEE Transactions on Power Systems, 34(4), 2577-2587.

Hall, D.B. (2000), Zero inflated Poisson and binomial regression with random effects: a case study, Biometrics, 56(4), 1030-1039.

Hau, N.H., Tuan, B.A., Thao, T.T.T., Wong, W.-K. (2019), Teaching mathematics by practical decision modeling in Vietnam high schools to serve the fourth industrial revolution, Journal of Management Information and Decision Sciences, 22(4), 445-462.

Hazan, E. (2016), Introduction to online convex optimization, Foundations and Trends in Optimization, 2(3-4), 157-325.

Hsieh, S., Lee, S., Shen, P.S. (2010), Logistic regression analysis of randomized response data with missing covariates, Journal of Statistical Planning and Inference, 140(4), 927-940.

Hsieh, S.H., Lee, S.M., Shen, P. (2009), Semiparametric analysis of randomized response data with missing covariates in logistic regression, Computational Statistics & Data Analysis, 53(7), 2673-2692.

Irvine, T. (2010), Application of the Newton-Raphson method to vibration problems, Vibrationdata.com (1999).

Kumar, D., Mishra, K.K. (2017), Portfolio optimization using novel co-variance guided Artificial Bee Colony algorithm, Swarm and Evolutionary Computation, 33, 119-130.

Lambert, D. (1992), Zero-inflated Poisson regression, with an application to defects in manufacturing, Technometrics, 34(1), 1-14.

Lee, S.M., Li, C.S., Hsieh, S.H., Huang, L.H. (2012), Semiparametric estimation of logistic regression model with missing covariates and outcome, Metrika, 75(5), 621-653.

Lee, S.M., Hwang, W.H., de Dieu Tapsoba, J. (2016), Estimation in closed capture-recapture models when covariates are missing at random, Biometrics, 72(4), 1294-1304.

Li, C.S., Lu, J.C., Park, J., Kim, K., Brinkley, P.A., Peterson, J.P. (1999), Multivariate zero-inflated Poisson models and their applications, Technometrics, 41(1), 29-38.

Little, R.J. (1992), Regression with missing X?: A review, Journal of the American Statistical Association, 87(420), 1227-1237.

Lukusa, T.M., Lee, S.M., Li, C.S. (2016), Semiparametric estimation of a zero-inflated Poisson regression model with missing covariates, Metrika, 79(4), 457-483.

Lukusa, T.M., Lee, S.M., Li, C.S. (2017), Review of zero-inflated models with missing data, Current Research in Biostatistics, 7(1), 1-12.

Mahmoudi, M.R., Heydari, M.H., Avazzadeh, Z., Pho, K.H. (2020), Goodness of fit test for almost cyclostationary processes, Digital Signal Processing, 96, 102597.

Pho, K.H., Nguyen, V.T. (2018), Comparison of Newton-Raphson algorithm and Maxlik function, Journal of Advanced Engineering and Computation, 2(4), 281-292.

Pho, K.H., Ho, T.D.C., Tran, T.K., Wong, W.K. (2019), Moment generating function, expectation and variance of ubiquitous distributions with applications in decision sciences: A review, Advances in Decision Sciences, 23(2), 1-85.

Pho, K.H., Ly, S., Ly, S., Lukusa, T.M. (2019), Comparison among Akaike information criterion, Bayesian information criterion and Vuong's test in model selection: A case study of violated speed regulation in Taiwan, Journal of Advanced Engineering and Computation, 3(1), 293-303.

Pho, K.H., Tran, T.K., Ho, T.D.C., Wong, W.K. (2019), Optimal solution techniques in decision sciences: A review, Advances in Decision Sciences 23(1), 1-47.

Polyak, B.T. (2007), Newton-method and its use in optimization, European Journal of Operational Research, 181(3), 1086-1096.

Riks, E. (1972), The application of Newton method to the problem of elastic stability, Journal of Applied Mechanics, 39(4), 1060-1065.

Thall, P.F. (1988), Mixed Poisson likelihood regression models for longitudinal interval count data, Biometrics, 44(1), 197-209.

Tian. Y, Pho, K.H. (2019), A statistical view to study the aphorisms in Nahj al Balaghah, Digital Scholarship in the Humanities, fqz075, https://doi.org/10.1093/llc/fqz075

Tong, X., Qi, L., Wu, F., Zhou, H. (2010), A smoothing method for solving portfolio optimization with CVaR and applications in allocation of generation asset, Applied Mathematics and Computation, 216(6), 1723-1740.

Truong, B.C., Nguyen, V.B., Truong, H.V., Ho, T.D.C. (2019a), Comparison of Optim, Nleqslv and MaxLik to Estimate Parameters in Some of Regression Models, Journal of Advanced Engineering and Computation, 3(4), 532-550.

Truong, B.C., Pho, K.H., Nguyen, V.B., Tuan, B.A., Wong, W.K. (2019b), Graph theory and environmental algorithmic solutions to assign vehicles: Application to garbage collection in Vietnam, Advances in Decision Sciences, 23(2), 1-35.

Tuan, B.A., Pho, K.H., Huy, L.M., Wong, W.K. (2019), Stemtech model in Asean Universities: An empirical research at Can Tho University, Journal of Management Information and Decision Sciences, 22(2), 107-127.

Tuan, B.A., Pudikova, G.N., Mahmoudi, M.R., Pho, K.H. (2009), Statistical approaches in literature: Comparing and clustering the alternatives of love in Divan of Hafiz, Digital Scholarship in the Humanities, fqz069, https://doi.org/10.1093/llc/fqz069

Tsupros, N., Kohler, R., Hallinen, J. (2009), STEM education: A project to identify the missing components, Intermediate Unit, 1, 11-17.

Wang, C.Y., Chen, J.C., Lee, S.M., Ou, S.T. (2002), Joint conditional likelihood estimator in logistic regression with missing covariate data, Statistica Sinica, 12, 555-574.

Zhu, Y., Le Gia, Q.T., Juhl, M.K., Coletti, G., Hameiri, Z. (2017), Application of the Newton- Raphson method to lifetime spectroscopy for extraction of defect parameters, IEEE Journal of Photovoltaics, 7(4), 1092-1097.