

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



Advances in Decision Sciences

Volume 23
Issue 1
March 2019

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



Published by Asia University, Taiwan

ADS@ASIAUNIVERSITY

Optimal Solution Techniques in Decision Sciences: A Review*

Kim-Hung Pho

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

Tuan-Kiet Tran

College of Science, Can Tho University, Viet Nam

Thi Diem-Chinh Ho

Faculty of Mathematics and Statistics, University of Natural Sciences, Ho Chi Minh City,
Vietnam

and

General Faculty, Binh Duong Economics & Technology University, Binh Duong City,
Vietnam

Wing-Keung Wong**

Department of Finance, Fintech Center, and Big Data Research Center, Asia University,
Taiwan

and

Department of Medical Research, China Medical University Hospital, Taiwan

and

Department of Economics and Finance, the Hang Seng University of Hong Kong, Hong
Kong

Revised: April 2019

* The authors are grateful to a referee for helpful comments and suggestions.

** Corresponding author: wong@asia.edu.tw

Abstract

Methods to find the optimization solution are fundamental and extremely crucial for scientists to program computational software to solve optimization problems efficiently and for practitioners to use it efficiently. Thus, it is very essential to know about the idea, origin, and usage of these methods. Although the methods have been used for very long time and the theory has been developed too long, most, if not all, of the authors who develop the theory are unknown and the theory has not been stated clearly and systematically. To bridge the gap in the literature in this area and provide academics and practitioners with an overview of the methods, this paper reviews and discusses the four most commonly used methods to find the optimization solution including the bisection, gradient, Newton-Raphson, and secant methods. We first introduce the origin and idea of the methods and develop all the necessary theorems to prove the existence and convergence of the estimate for each method. We then give two examples to illustrate the approaches. Thereafter, we review the literature of the applications of getting the optimization solutions in some important issues and discuss the advantages and disadvantages of each method. We note that all the theorems developed in our paper could be well-known but, so far, we have not seen any book or paper that discusses all the theorems stated in our paper in detail. Thus, we believe the theorems developed in our paper could still have some contributions to the literature. Our review is useful for academics and practitioners in finding the optimization solutions in their studies.

Keywords: bisection method, gradient method, Newton-Raphson method, secant method

JEL: A10, G00, G31, O32

1 Introduction

In many fields of science, economics, finance, and many other areas, the problems related to find the optimization solutions play a very important role in research. In statistics, the problem to find the optimization solution is basically in all models, for example, regression and time series models. In economics and finance, the problem to find the optimization solutions occurs in solving most of the economic and financial problems, for example, interest rate and stock returns. In social science research, this problem often occurs to solve many problems, for example, transportation problems. In marketing, the problem to find the optimization solutions occurs in solving most of the marketing and production problems, for example, cost of goods, etc. Methods to find the optimization solution are fundamental and extremely crucial for scientists to program computational software solving optimization problems efficiently and for practitioners to use it efficiently. Thus, it is very essential to know about the idea, origin, and usage of these methods.

There are numerous methods to find the optimization solutions. The four most ubiquitous approaches include bisection, gradient, Newton-Raphson (N-R), and secant methods. These methods are often applied in numerical analysis and they have many applications in statistics and numerous other fields. About this regard, there are several work have been studied and utilized in many different disciplines, for example, Broyden (1967), Chalco et al. (2015), Chin et al. (2018), Exl et al. (2019), Reddy et al. (2018), Sahu et al. (2018), Wu et al. (2017), Yang et al.(2019).

Until now, getting methods of finding the optimization solutions is still a very important problem. So far, the N-R method is found to be the most commonly utilized method. The N-R method is employed by many statisticians to find the optimization solutions in estimating functions of many models, for example, regression with missing data. For example, Wang et al. (2002) use the N-R method to find the optimization in the logistic regression model with missing covariates. Lukusa et al. (2016) employ the N-R method to find the optimization in the zero-inflated Poisson (ZIP) regression model with missing covariates. Other papers using this method include Hiesh et al. (2009 and 2010), Lee et al. (2012 and 2016). Readers may refer to Pho and Nguyen (2018) for the detail of the algorithm and applications of the N-R method.

Although the methods to find the optimization solutions have been used for very long time and there are numerous articles present about methods to find the optimization solution, the

theory of finding the optimization solutions has been developed too long that most, if not all, of the authors who develop the theory are unknown and the theory has not been stated clearly and systematically. To bridge the gap in the literature in this area and provide academics and practitioners with an overview of the methods to find the optimization solution, our primary goals in this paper is to review and discuss the four most commonly used methods to find the optimization solution including the bisection, gradient, Newton-Raphson, and secant methods. In this study, we will discuss the origin and idea of the methods, discuss how to utilize the methods and discuss its applications.

To do so, in this paper we first introduce the origin and idea of methods to find the optimization solution and develop all the necessary theorems to prove the existence and convergence of the estimate for each method. We then give two examples to illustrate the approaches. Thereafter, we review the literature of the applications of getting the optimization solutions in some important issues by using the methods discussed in our paper and discuss the advantages and disadvantages of each method. We note that all the theorems developed in our paper could be well-known but, so far, we have not seen any book or paper that discusses all the theorems stated in our paper in detail, and thus, we believe the theorems developed in our paper could still have some contributions to the literature. Our review is useful for academics and practitioners in finding the optimization solutions in their studies.

The rest of the paper is structured as follows. In Section 2, we introduce the origin and idea of methods to find the optimization solution and develop all the necessary theorems to prove the existence and convergence of the estimate for each method. Section 3 provides two examples to illustrate the approaches of finding the optimization solution. Section 4 reviews the literature of the applications of getting the optimization solutions in some important issues by using the methods discussed in our paper. The last section gives some concluding remarks and inferences and discuss the advantages and disadvantages of methods being discussed in our paper.

2 The methods to find the optimization solution

In this section, we discuss the origin and idea of four approaches to find the optimization solutions that are most used in the literature. These methods include the bisection method, gradient method, Newton-Raphson method, and secant method. We first discuss the bisection

tion method.

2.1 Bisection method

The bisection method is a naive method based on the following theorem.

Theorem 1

Assuming that $g(u)$ is a continuous function on $[a, b]$ and $g(a)g(b) < 0$, then there exists at least one solution in the (a, b) interval for the equation $g(u) = 0$.

Theorem 1 is also known as Bolzano's theorem. This theorem was first proved by Bernard Bolzano in 1817. Augustin-Louis Cauchy provided an alternative proof in 1821. The procedure to implement the bisection method are as follows: one first finds the (a, b) interval satisfies $g(a)g(b) < 0$. Thereafter, one divides the (a, b) interval into two equal intervals. Let (a_1, b_1) be an one of the above two intervals and satisfies $g(a_1)g(b_1) < 0$. Afterwards, (a_1, b_1) will be divided into two equal intervals. Let (a_2, b_2) be an one of the above two intervals and satisfies $g(a_2)g(b_2) < 0$. Then, one further repeats the above process. If at a_i such that $g(a_i) = 0$ or at b_i satisfies $g(b_i) = 0$, then one can conclude that a_i or b_i is a solution of $g(u)$. The process work of the bisection method is described in Figure 1.

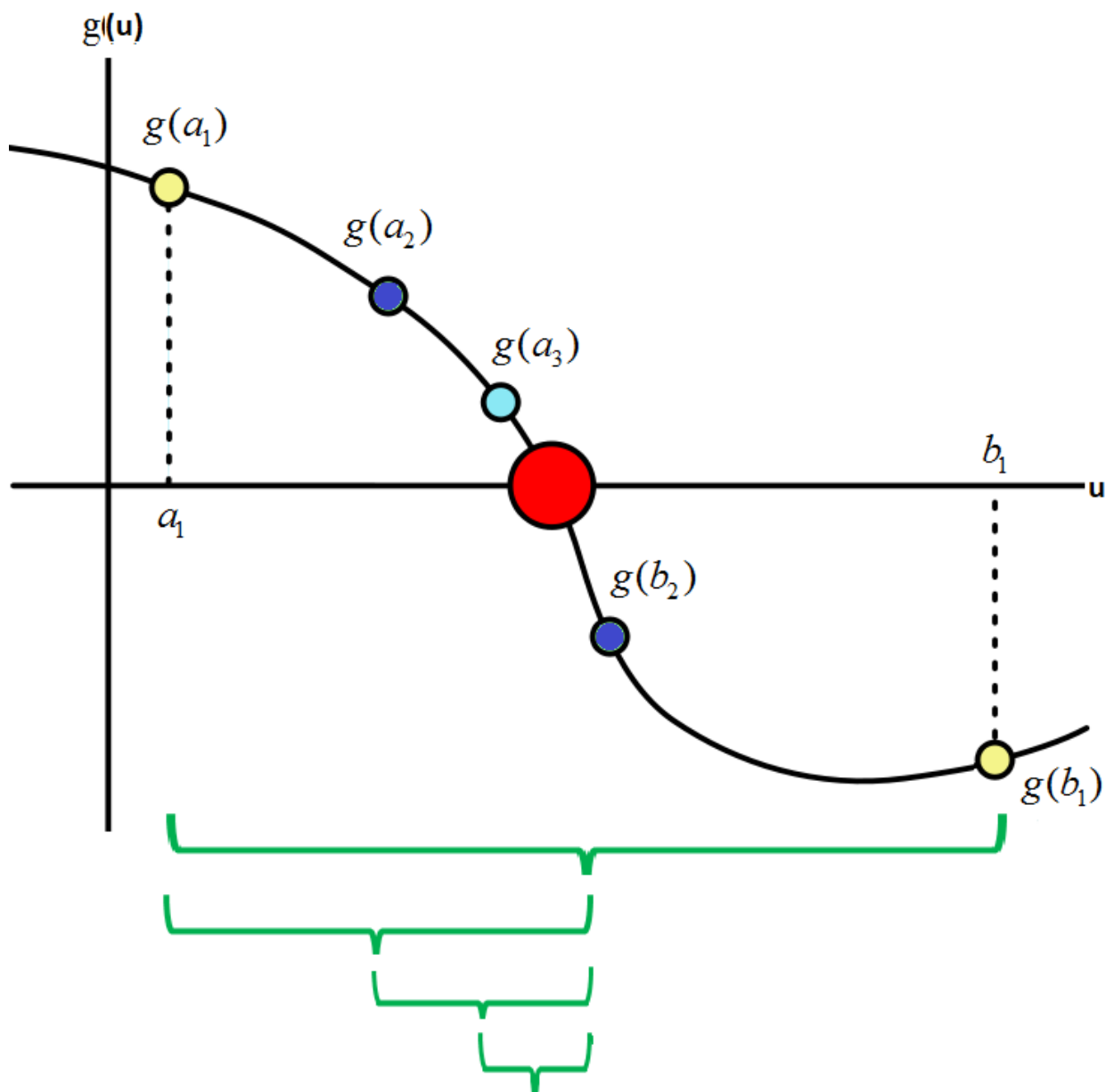


Figure 1: The process work of the bisection method

In Figure 1, we provide a few steps of the bisection method illustrated with the initial $[a_1; b_1]$ interval. The bigger red dot is an optimization solution of $g(u) = 0$.

Theorem 2

In the procedure described in the above, we have $g(a_i) \rightarrow 0$ and $g(b_i) \rightarrow 0$ as $i \rightarrow \infty$.

We note that Theorem 2 does not only hold for the bisection method, but also holds for all the methods discussed in this paper. Since the contents of the theorems for the convergence of the estimates in all other methods are the same, we skip display the theorem for other methods for simplicity.

2.2 Gradient method

Another commonly used approach is the gradient method. This is one of the ubiquitous method to find the optimization solution Terlaky (2013). Our main objective is to obtain the difference between u_k and u_{k+1} and u^* such that the difference between u_k and u_{k+1} is within our acceptance level where u_k is the obtained point after the k^{th} loop. The best situation is that u_k coincides with u^* but usually, this is not the case. Then, we first construct the following theorem which shows the working process of the gradient method:

Theorem 3

Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function, u_k be an obtained point after the k^{th} loop, and u^* be an optimization solution of $g(u)$. We have the following:

1. If $g'(u_k) > 0$, then u_k is on the right of u^* ; and
2. if $g'(u_k) < 0$ then u_k is on the left of u^* .

The gradient method could be used for the one-dimension case and the multi-dimension case.

2.2.1 One-dimension

We first discuss the one-dimension case. The idea of the gradient method in the one-dimension case is based on its derivative. From Theorem 3, it well known that if $g'(u_k) > 0$, then u_k is on the right of u^* . Similarly, if $g'(u_k) < 0$, then u_k is on the left of u^* . In either situations, to the next value of u_k is u_{k+1} that we wish it will be close to u^* . If this is the

situation, then one needs to move to the opposite of the sign of $g'(u_k)$. In general, we have

$$u_{k+1} = u_k + \alpha,$$

where α is a quantity opposite to the sign of $g'(u_k)$. It can be rewritten as follows:

$$u_{k+1} = u_k - \gamma g'(u_k) \tag{1}$$

where $\gamma > 0$ is the learning rate. The minus sign in the above expression shows that one needs to move backwards with the sign of its derivative. The formula (1) is a gradient method in the one-dimension case. The process of the gradient method is described in Figure 2.

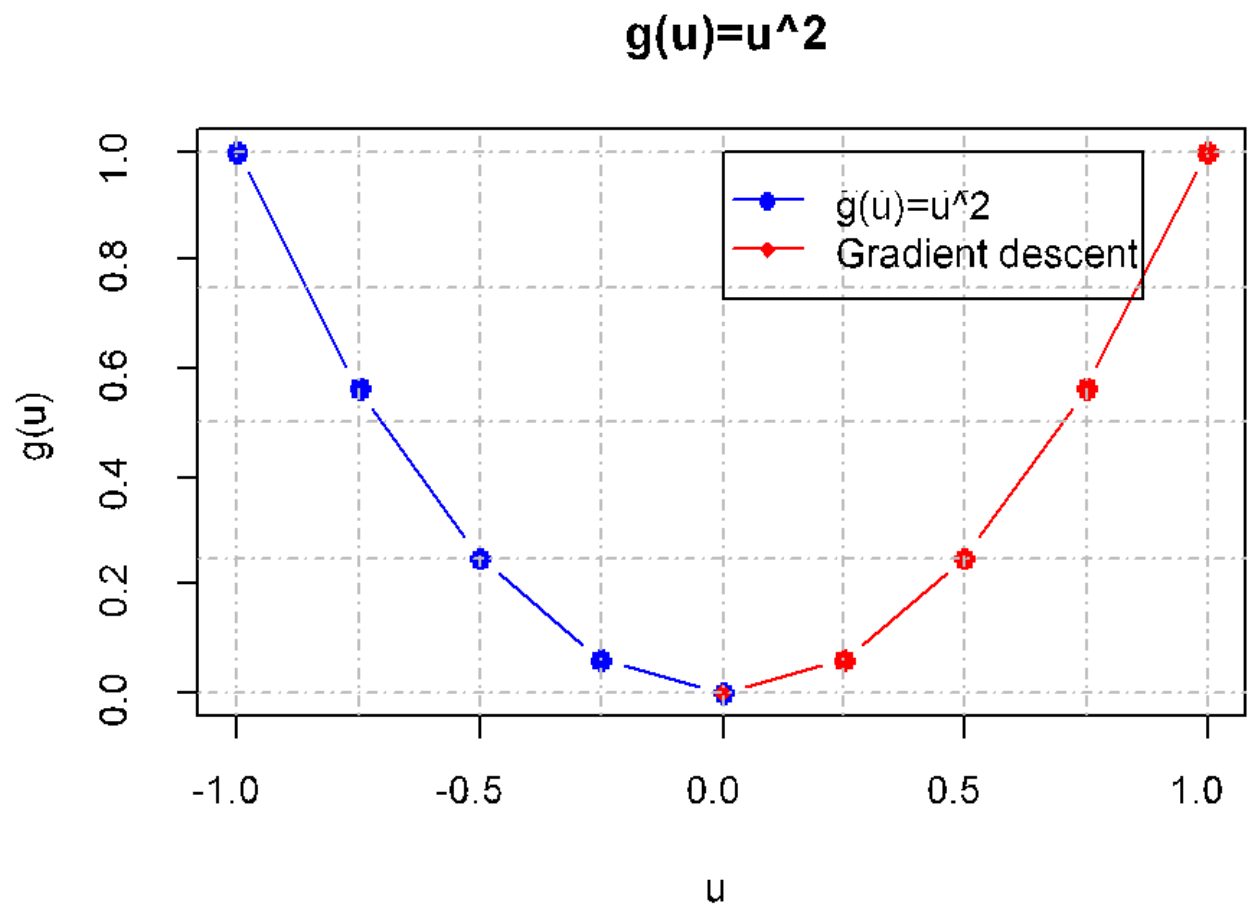


Figure 2: The process work of the gradient method

It can be observed from Figure 2 that, the objective function is $g(u) = u^2$ and the corresponding optimization solution is $u = 0$. If $x_k > 0$, then $f'(x_k) > 0$ as shown as the red dots on the right-hand side of the figure. If $x_k < 0$, then $f'(x_k) < 0$, as shown as the blue dots on the left-hand side of the figure. At point $u = 0$, it is illustrated by a dot with both red and blue color. This is also the optimization solution of $g(u) = u^2$.

2.2.2 Multi-dimension

We turn to discuss the multi-dimension case. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function, *the direction of steepest descent* is the vector $-\nabla g(u_0)$, where u_0 is an initial value. To explain this phenomenon, we consider the following function

$$H(y) = g(u_0 + y\mathbf{a})$$

where \mathbf{a} is a *unit* vector; this means that $\|\mathbf{a}\| = 1$. Using the chain rule, we have

$$\begin{aligned} H'(y) &= \frac{\partial g}{\partial u_1} \frac{\partial u_1}{\partial y} + \cdots + \frac{\partial g}{\partial u_n} \frac{\partial u_n}{\partial y} \\ &= \frac{\partial g}{\partial u_1} a_1 + \cdots + \frac{\partial g}{\partial u_n} a_n \\ &= \nabla g(u_0 + y\mathbf{a}) \cdot \mathbf{a} . \end{aligned}$$

Thus, we obtain

$$H'(0) = \nabla g(u_0) \cdot \mathbf{a} = \|\nabla g(u_0)\| \cdot \cos\alpha , \quad (2)$$

where α is the angle between $\nabla g(u_0)$ and \mathbf{a} . We remark that $H'(0)$ in (2) is minimized when $\alpha = \pi$. Therefore, we get

$$\mathbf{a} = -\frac{\nabla g(u_0)}{\|\nabla g(u_0)\|}, H'(0) = -\|\nabla g(u_0)\| .$$

Hence, the problem of minimizing a function of numerous variables can be reduced to a single variable minimization problem by solving the minimum of $H(y)$ for the choice of \mathbf{a} . To do so, one can find the value of y , ($y > 0$), that minimizes

$$H_0(y) = g(u_0 - y\nabla g(u_0)) .$$

After addressing the minimizer y_0 , we let

$$u_1 = u_0 - y_0 \nabla g(u_0)$$

and continue the process by finding from u_1 in the direction of $-\nabla g(u_1)$ to get u_2 by minimizing $H_1(y) = g(u_1 - y\nabla g(u_1))$, and so on.

Now, we discuss the procedure of the *Method of Steepest Descent*. Given a starting value u_0 , the sequence of iterates $\{u_k\}$ will be calculated by using the following formula:

$$u_{k+1} = u_k - y_k \nabla g(u_k) , \quad (3)$$

where $y_k > 0$ minimizes the following function

$$H_k(y) = g(u_k - y\nabla g(u_k)) .$$

The expression in (3) is a gradient method in multi-dimension.

2.3 Newton-Raphson method

We turn to discuss how to use the Newton-Raphson (N-R) method to find the optimization solution of the equation. The N-R method could also be used for the one-dimension case and the multi-dimension case.

2.3.1 One-dimension

We first discuss the one-dimension case in which the optimization solution of the equation $g(u) = 0$ where $g : D(D \subset \mathbb{R}) \rightarrow \mathbb{R}$. The N-R method is originated from the presumption of the derivative of a function at a point:

$$g'(u_1) = \frac{dg(u_1)}{du} \approx \frac{g(u_2) - g(u_1)}{u_2 - u_1} . \quad (4)$$

One has to find u^* that satisfies $g(u^*) = 0$; this means that $g(u_2) = 0$. Hence,

$$u_2 \approx u_1 - \frac{g(u_1)}{g'(u_1)} . \quad (5)$$

Generally, the formula solution of the N-R method can be expressed as follows:

$$u_{n+1} = u_n - \frac{g(u_n)}{g'(u_n)} , \quad (6)$$

in which with u_0 is a starting value. We reduplicate the above formula until the difference between two contiguous solutions is smaller than α , a very small value we can accept.

We note that using the N-R method to find the optimization solution of the equation $g'(u) = 0$ can be obtained from Equation (6) as follows:

$$u_{n+1} = u_n - \frac{g'(u_n)}{g''(u_n)} . \quad (7)$$

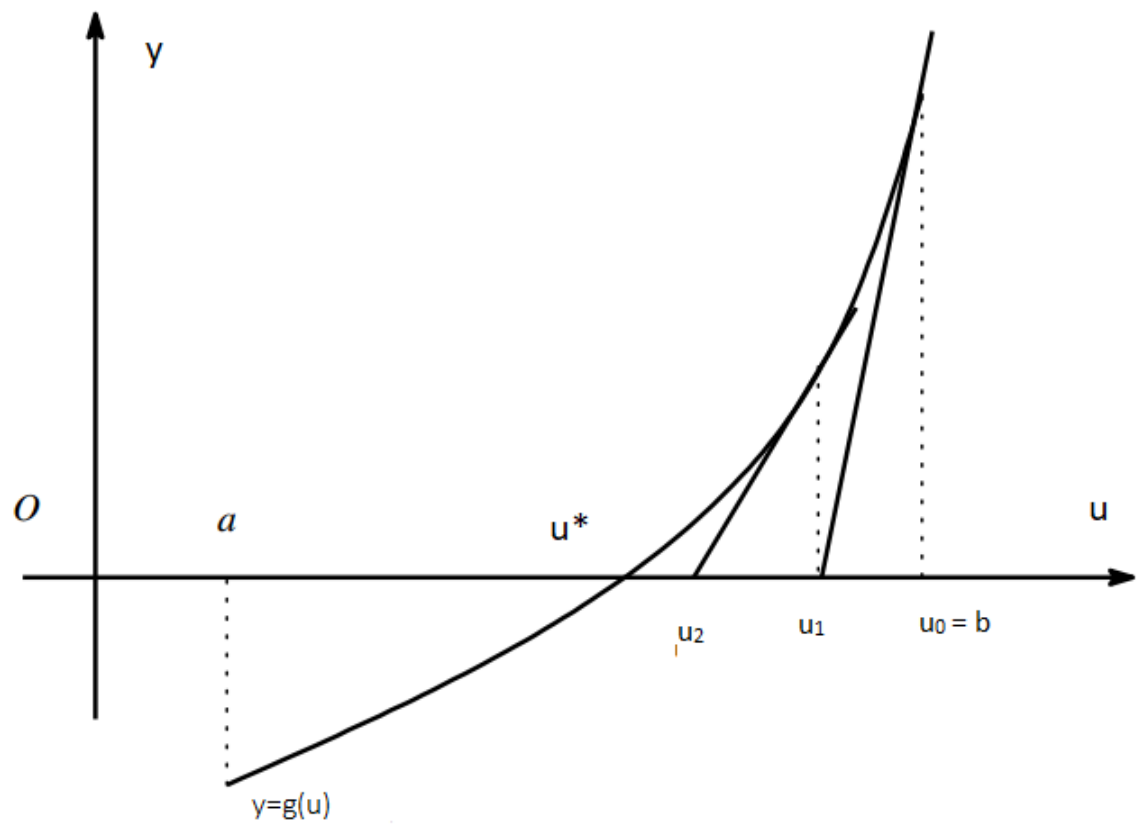


Figure 3: The process work of the Newton-Raphson method

The process work of the Newton-Raphson method is provided in Figure 3. It can be observed from Figure 3 that the optimization solution of $y = g(u)$ is described on the initial $[a, b]$ interval by utilizing the Newton-Raphson method. Choosing an initial value $u_0 = b$, the tangent of $y = g(u)$ at $u_0 = b$ cuts the u axis at point u_1 , the tangent of $y = g(u)$ at u_1 cuts the u axis at point u_2 , and so on, further repeating this process, one can obtain the optimization solution of $y = g(u)$ to be u^* . It can be seen that the Newton-Raphson method utilizes the intersection of the tangent and horizontal axis, and thus, the Newton-Raphson method is also called the tangent method.

2.3.2 Multi-dimension

We turn to discuss the multi-dimension case. To build the N-R method to find the optimization solution of the equation that $g(u) = 0$ in the multi-dimension case, we need to approximate $g(u)$ near the current repetition u_n by the function $h_n(u)$ such that the system of equations $h_n(u) = 0$ is uncomplicated to address. Thereafter, we utilize the solution as the next repetition u_{n+1} and repeat the procedure. An appropriate selection for $h_n(u)$ is a linear approximation of $g(u)$ at u_n , whose graph is the tangent of $g(u)$ at u_n such that:

$$h_n(u) = g(u_n) + \nabla g(u^{(n)}) (u - u_n).$$

Then, we address the following system of equations:

$$\begin{aligned} 0 &= h_n(u_{n+1}) \\ &= g(u_n) + \nabla g(u^{(n)}) (u_{n+1} - u_n) . \end{aligned} \tag{8}$$

Therefore, we do the following:

$$u_{n+1} = u_n - [\nabla g(u^{(n)})]^{-1} g(u_n) . \tag{9}$$

Likewise, in case of the objective function is $g'(u)$, then the formula in (7) can be inferred to have the N-R algorithm in the multi-dimension case by substituting $g'(u)$ by a gradient $\nabla g(u)$, and replacing $g''(u)$ by the inverse of the Hessian matrix $Hg(u)$. The formula solution of the N-R algorithm in the multi-dimension case can be then described as follows:

$$u^{(n+1)} = u^{(n)} - [Hg(u^{(n)})]^{-1} \nabla g(u^{(n)}) , \tag{10}$$

where

$$\nabla g(u) = \left[\frac{\partial g(u)}{\partial u_1}; \frac{\partial g(u)}{\partial u_2}; \dots; \frac{\partial g(u)}{\partial u_n} \right]^T$$

and

$$Hg(u) = \begin{bmatrix} \frac{\partial^2 g(u)}{\partial u_1^2} & \frac{\partial^2 g(u)}{\partial u_1 \partial u_2} & \cdots & \frac{\partial^2 g(u)}{\partial u_1 \partial u_n} \\ \frac{\partial^2 g(u)}{\partial u_2 \partial u_1} & \frac{\partial^2 g(u)}{\partial u_2^2} & \cdots & \frac{\partial^2 g(u)}{\partial u_2 \partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 g(u)}{\partial u_n \partial u_1} & \frac{\partial^2 g(u)}{\partial u_n \partial u_2} & \cdots & \frac{\partial^2 g(u)}{\partial u_n^2} \end{bmatrix}.$$

Similar to the procedure used in the one-dimensional case, to achieve the optimization solution by employing the Newton-Raphson algorithm in the multi-dimension situation, one needs to reduplicate the formula (10) until an adequately correct value is reached.

2.4 Secant method

The last one we discuss is how to use the secant method to find the optimization solution of the equation. The secant could also be used for the one-dimension case and the multi-dimension case.

2.4.1 One-dimension

We first discuss the one-dimension case. It well known that

$$g'(u_n) = \frac{dg(u_n)}{du} \approx \frac{g(u_n) - g(u_{n-1})}{u_n - u_{n-1}}. \quad (11)$$

By replacing Equation (11) into Equation (6), we obtain the following equation:

$$u_{n+1} = u_n - \frac{g(u_n)(u_n - u_{n-1})}{g(u_n) - g(u_{n-1})}. \quad (12)$$

The formula in Equation (12) is referred to the secant method that is the method in which derivative is substituted by its approximation. Thus, it is also called a quasi-Newton method.

2.4.2 Multi-dimension

We turn to discuss the multi-dimension case. Similarly in one-dimension, to find the minimum of $g(u)$, one can use its second-order approximation. Because the Taylor's expansion series is

$$g(u_k + \Delta u) \approx g(u_k) + \nabla g(u_k)^T \Delta u + \frac{1}{2} \Delta u^T H \Delta u,$$

where ∇g is the gradient and H is an approximation of the Hessian matrix, the approximation of the gradient is, thus, given by

$$\nabla g(u_k + \Delta u) \approx \nabla g(u_k) + H \Delta u.$$

Setting the above expression equal to zero, we get

$$\Delta u = -H^{-1}\nabla g(u_k) .$$

Thus, B is selected as follows

$$\nabla g(u_k + \Delta u) = \nabla g(u_k) + H \Delta u .$$

This result is well known in the literature and it is also called the secant equation.

3 Illustrations

In this section, we provide two examples to help readers to overview the methods we discussed in our paper to find the optimization solution to illustrate how to find the optimization solution by utilizing four methods are mentioned in the previous section. We first discuss the following example for the one-dimensional case:

Example 1: To find the optimization solution of $g(u) = u^3 + u^2 + u - 2$ by utilizing the bisection, gradient, Newton-Raphson, and secant methods with a starting value $u_0 = 1$, we performing 4 iterations for each method as follows in the following:

(i) Bisection method

Assuming that $(a_0, b_0) = (0, 2)$, we have $g(a_0)g(b_0) = g(0)g(2) = -24 < 0$. It can be observed that $g(0)g(1) = -2 < 0$ and $g(1)g(2) = 12 > 0$. Thus, we choose $(a_1, b_1) = (0, 1)$. It can be seen that $g(0)g(1/2) = 2.25 > 0$ and $g(1/2)g(1) = -1.125 < 0$. Hence, we select $(a_2, b_2) = (1/2, 1)$. In addition, one can easily find that $g(1/2)g(3/4) = 153/512 > 0$ and $g(3/4)g(1) = -17/640$. Therefore, we choose $(a_3, b_3) = (3/4, 1)$. Thereafter, we have $g(3/4)g(7/8) = -2703/32768 < 0$ and $g(7/8)g(1) = 159/512 > 0$. Thus, we choose $(a_4, b_4) = (3/4, 7/8)$. So, the optimization solution of $g(u)$ belongs to $(a_4, b_4) = (3/4, 7/8)$.

(ii) Gradient method

Since it is easy to get $g'(u) = 3u^2 + 2u + 1$, to use the gradient method, we choose the learning rate $\gamma = 0.009$. Following the expression in Equation (1), the sequence $\{u_n\}$

is defined as follows:

$$u_{k+1} = u_k - 0.009(3u_k^2 + 2u_k + 1)$$

with $u_0 = 1$, we have

$$u_1 = 0.946, g(u_1) = 0.68751$$

$$u_2 = 0.89581, g(u_2) = 0.41715$$

$$u_3 = 0.84902, g(u_3) = 0.18185$$

$$u_4 = 0.80527, g(u_4) = -0.02407 .$$

From the above iterations, one can see that if more iterations are performed, then the value will converge to the exact solution of the original equation.

(iii) Newton-Raphson method

To use the Newton-Raphson method, one can observe that $g'(v) = 3v^2 + 2v + 1$. The sequence $\{v_n\}$ is defined as follows:

$$\begin{aligned} v_{n+1} &= v_n - \frac{g(v_n)}{g'(v_n)} \\ &= v_n - \frac{v_n^3 + v_n^2 + v_n - 2}{3v_n^2 + 2v_n + 1} . \end{aligned}$$

Using $v_0 = 1$, we have

$$v_1 = 0.83333, g(v_1) = 0.10648$$

$$v_2 = 0.81092, g(v_2) = 0.00175$$

$$v_3 = 0.81054, g(v_3) = 5.10^{-7}$$

$$v_4 = 0.81054, g(v_4) = 0 .$$

(iv) Secant method

Last, we illustrate how to use the secant method. Based on the formula in (12), the sequence $\{u_n\}$ is defined as follows:

$$u_{n+1} = u_n - \frac{g(u_n)(u_n - u_{n-1})}{g(u_n) - g(u_{n-1})} .$$

Assuming that $u_0 = 1$ and $u_1 = 0$, we have

$$\begin{aligned} u_2 &= 0.66667, g(u_2) = -0.59259 , \\ u_3 &= 0.94737, g(u_3) = 0.69515 , \\ u_4 &= 0.79584, g(u_4) = -0.06674 , \\ u_5 &= 0.80911, g(u_5) = -0.00652 . \end{aligned}$$

It can be seen that if more iterations are performed, then the value will converge to the exact solution of the original equation.

Example 1 provides an illustration for the one-dimensional case. We turn to provide the following example for the two-dimensional case. To save space, we only discuss getting the optimization solution by utilizing the gradient and Newton-Raphson methods as shown in the following example.

Example 2: To find the optimization solution of $g(u_1, u_2) = 4u_1^2 - 4u_1u_2 + 2u_2^2$ by utilizing the gradient and Newton-Raphson methods with a starting value $u^{(0)} = (1, 1)$, we performing the 6 iterations by utilizing the gradient and Newton-Raphson methods:

(i) Gradient method

We first discuss the gradient method. It can be seen that

$$\nabla g(u_1, u_2) = (8u_1 - 4u_2, -4u_1 + 4u_2) .$$

Thus,

$$\nabla g(u^{(0)}) = \nabla g(1, 1) = (4, 0) .$$

Thereafter, one minimize the following function:

$$\begin{aligned} H(y) &= g((1, 1) - y(4, 0)) \\ &= g(1 - 4y, 1) . \end{aligned}$$

Hence,

$$\begin{aligned} H'(y) &= -\nabla g(1 - 4y, 1) \cdot (4, 0) \\ &= -(4 - 32y, 16y) \cdot (4, 0) \\ &= 128y - 16 . \end{aligned}$$

This strictly convex function has a strict global minimum when $H'(y) = 128y - 16 = 0$, or $y = 1/8$, as can be seen by noting that $H''(y) = 128 > 0$. We now conduct the following iterations:

Iteration 1

In Iteration 1, we first obtain

$$\begin{aligned} u^{(1)} &= u^{(0)} - \frac{1}{8} \nabla g(u^{(0)}) \\ &= (1, 1) - \frac{1}{8}(4, 0) = \left(\frac{1}{2}, 1\right) . \end{aligned}$$

Thereafter, we get

$$\nabla g(u^{(1)}) = \nabla g\left(\frac{1}{2}, 1\right) = (0, 2) .$$

We then let

$$\begin{aligned} H(y) &= g\left(\left(\frac{1}{2}, 1\right) - y(0, 2)\right) \\ &= g\left(\frac{1}{2}, 1 - 2y\right) , \end{aligned}$$

and obtain

$$\begin{aligned} H'(y) &= -(8y, 2 - 8y) \cdot (0, 2) \\ &= 16y - 4 , \end{aligned}$$

and get $H'(y) = 0$ when $y = 1/4$, and $H''(y) = 16$. Thus, this critical point is a strict global minimizer in Iteration 1.

Iteration 2

We turn to Iteration 2. To do so, we first get

$$\begin{aligned} u^{(2)} &= u^{(1)} - \frac{1}{4} \nabla g(u^{(1)}) \\ &= \left(\frac{1}{2}, 1\right) - \frac{1}{4}(0, 2) \\ &= \left(\frac{1}{2}, 1\right) - \left(0, \frac{1}{2}\right) = \left(\frac{1}{2}, \frac{1}{2}\right) . \end{aligned}$$

Thus, we have

$$\nabla g(u^{(2)}) = \nabla g\left(\frac{1}{2}, \frac{1}{2}\right) = (2, 0)$$

and

$$\begin{aligned} H(y) &= g\left(\left(\frac{1}{2}, \frac{1}{2}\right) - y(2, 0)\right) \\ &= g\left(\frac{1}{2} - 2y, \frac{1}{2}\right), \end{aligned}$$

and get

$$\begin{aligned} H'(y) &= -(2 - 16y, 8y) \cdot (2, 0) \\ &= 32y - 4. \end{aligned}$$

Thus, $H'(y) = 0$ when $y = 1/8$ and $H''(y) = 32$, this critical point is a strict global minimizer in Iteration 2.

Iteration 3

We turn to Iteration 3. To do so, we first get

$$\begin{aligned} u^{(3)} &= u^{(2)} - \frac{1}{8} \nabla g(u^{(2)}) \\ &= \left(\frac{1}{2}, \frac{1}{2}\right) - \frac{1}{8}(2, 0) = \left(\frac{1}{4}, \frac{1}{2}\right). \end{aligned}$$

Therefore, we have

$$\nabla g(u^{(3)}) = \nabla g\left(\frac{1}{4}, \frac{1}{2}\right) = (0, 1)$$

and

$$\begin{aligned} H(y) &= g\left(\left(\frac{1}{4}, \frac{1}{2}\right) - y(0, 1)\right) \\ &= g\left(\frac{1}{4}, \frac{1}{2} - y\right). \end{aligned}$$

In addition, we obtain

$$\begin{aligned} H'(y) &= -(4y, 1 - 4y) \cdot (0, 1) \\ &= 4y - 1. \end{aligned}$$

Thus, $H'(y) = 0$ when $y = \frac{1}{4}$ and $H''(y) = 4$, this critical point is a strict global minimizer in Iteration 3.

Iteration 4

We are now doing Iteration 4. To do so, we first get

$$\begin{aligned} u^{(4)} &= u^{(3)} - \frac{1}{4} \nabla g(u^{(3)}) \\ &= \left(\frac{1}{4}, \frac{1}{2}\right) - \frac{1}{4}(0, 1) = \left(\frac{1}{4}, \frac{1}{4}\right) . \end{aligned}$$

Thereafter, we obtain

$$\nabla g(u^{(4)}) = \nabla g\left(\frac{1}{4}, \frac{1}{4}\right) = (1, 0)$$

and

$$\begin{aligned} H(y) &= g\left(\left(\frac{1}{4}, \frac{1}{4}\right) - y(1, 0)\right) \\ &= g\left(\frac{1}{4} - y, \frac{1}{4}\right) , \end{aligned}$$

and get

$$\begin{aligned} H'(y) &= -(1 - 8y, 4y) \cdot (1, 0) \\ &= 8y - 1 . \end{aligned}$$

Thus, we have $H'(y) = 0$ when $y = \frac{1}{8}$ and $H''(y) = 8$, this critical point is a strict global minimizer in Iteration 4.

Iteration 5

Next, we illustrate Iteration 5. To do so, we first obtain

$$\begin{aligned} u^{(5)} &= u^{(4)} - \frac{1}{8} \nabla g(u^{(4)}) \\ &= \left(\frac{1}{4}, \frac{1}{4}\right) - \frac{1}{8}(1, 0) = \left(\frac{1}{8}, \frac{1}{4}\right) . \end{aligned}$$

Hence, we get

$$\nabla g(u^{(5)}) = \nabla g\left(\frac{1}{8}, \frac{1}{4}\right) = \left(0, \frac{1}{2}\right)$$

and

$$\begin{aligned} H(y) &= g\left(\left(\frac{1}{8}, \frac{1}{4}\right) - y\left(0, \frac{1}{2}\right)\right) \\ &= g\left(\frac{1}{8}, \frac{1}{4} - \frac{1}{2}y\right) . \end{aligned}$$

In addition, we have

$$\begin{aligned} H'(y) &= -\left(2y, \frac{1}{2} - 2y\right) \cdot \left(0, \frac{1}{2}\right) \\ &= y - \frac{1}{4} . \end{aligned}$$

Thus, we get $H'(y) = 0$ when $y = \frac{1}{4}$ and $H''(y) = 1$, this critical point is a strict global minimizer in Iteration 5.

Iteration 6

Last, we illustrate Iteration 6. To do so, we first get

$$\begin{aligned} u^{(6)} &= u^{(5)} - \frac{1}{4} \nabla g(u^{(5)}) \\ &= \left(\frac{1}{8}, \frac{1}{4}\right) - \frac{1}{4} \left(0, \frac{1}{2}\right) = \left(\frac{1}{8}, \frac{1}{8}\right). \end{aligned}$$

It has been seen that the method of steepest descent produces a sequence of iterates u_k that is converging to the strict global minimizer of $g(u_1, u_2)$ at $u^* = (0, 0)$. This ends our illustration of the gradient method.

(ii) Newton-Raphson method

We turn to illustrate the Newton-Raphson method. We first get

$$g(u_1, u_2) = 4u_1^2 - 4u_1u_2 + 2u_2^2.$$

Thus, we have

$$\begin{aligned} \nabla g(u_1, u_2) &= (8u_1 - 4u_2, -4u_1 + 4u_2) \\ Hg(u_1, u_2) &= \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix} \end{aligned}$$

with a starting value $u^{(0)} = (1, 1)$, then we get

$$\begin{aligned} \nabla g(u^{(0)}) &= \nabla g(1, 1) = (4, 0) \\ Hg(u^{(0)}) &= \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix} \end{aligned}$$

We now conduct the following iterations:

Iteration 1

In Iteration 1, we first obtain

$$\begin{aligned} u^{(1)} &= u^{(0)} - [Hg(u^{(0)})]^{-1} \nabla g(u^{(0)}) \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}^{-1} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{aligned}$$

we then get

$$\begin{aligned}\nabla g(u^{(1)}) &= \nabla g(0,0) = (0,0) \\ Hg(u^{(1)}) &= \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}.\end{aligned}$$

Iteration 2

In Iteration 2, we first get

$$\begin{aligned}u^{(2)} &= u^{(1)} - [Hg(u^{(1)})]^{-1} \nabla g(u^{(1)}) \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.\end{aligned}$$

Thus, we have

$$u^{(1)} = u^{(2)} = \dots = u^{(6)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

It can be seen from this example that the Newton-Raphson method will converge faster than the gradient method. From Examples 1 and 2, it has been seen that, the Newton-Raphson method provides the most reliable results. If comparing about the Newton-Raphson and gradient method. About memory, each iteration in the Newton-Raphson method needs $O(n^2)$ storage ($n \times n$ Hessian), while each iteration in the gradient method requests storage $O(n)$ (n-dimensional gradient).

For the computation, each iteration in the Newton-Raphson method requests $O(n^3)$ flops (addressing a ponderous $n \times n$ linear system), while each iteration in the gradient method needs to $O(n)$ flops (scaling/adding n-dimensional vectors). Thus, it can be seen that, about the time to output the results of two methods, the gradient method provides the results faster than the Newton-Raphson method. Nevertheless, since the Newton-Raphson method is a robust method, it can work well in case the data set contain missing value. About this regard, there are several scholars utilized the Newton-Raphson method to find parameters in the regression models with missing covariates data. Readers may refer to Hsieh et al. (2009, 2010), Lee et al. (2012, 2016), Lukusa et al. (2016), Wang et al. (2002) for more information. In addition, the Newton-Raphson method can work well in the data set having missing values, where some available function in the statistical software is unworkable (see e.g. Pho et al. (2018)).

4 Applications of Optimization Solutions in Decision Sciences

In order to provide academics and practitioners with an overview of the methods to find the optimization solution, in this article we review of the four most commonly used approaches to find the optimization solution. After applying the methods in the analysis, one can utilize many other statistical models, for example, regression models, to find the optimization solutions in some of problems including interest rate, transportation, the issues involve to the cost of goods, etc. There are many areas in decision sciences that can use the approaches discussed in our paper to get optimization solutions. We discuss a few in this section.

4.1 Portfolio Optimization

The first area in decision sciences that can use the approaches discussed in our paper to get optimization solutions is portfolio optimization. The portfolio optimization is the milestone of modern finance theory for asset allocation, investment diversification, and optimal portfolio construction after Markowitz (1952) introduces the theory. In the procedure, investors select portfolios that maximize profit subject to achieving a specified level of calculated risk or, equivalently, minimize variance subject to obtaining a predetermined level of expected gain.

Bai, Liu, and Wong (2009a) prove that the estimates proposed by Markowitz (1952) is seriously depart from its theoretic optimal return and they call this phenomenon "over-prediction." In order to circumvent this over-prediction problem, they use a new method by incorporating the idea of the bootstrap into the theory of large dimensional random matrix. They develop new bootstrap-corrected estimates for the optimal return and its asset allocation, and prove that these bootstrap-corrected estimates can analytically correct the over-prediction and drastically reduce the error. They also show that the bootstrap-corrected estimate of return and its corresponding allocation estimate are proportionally consistent with their counterpart parameters.

There are some advantages of the approach introduced by Bai, Liu, and Wong (2009a) to obtain the optimal return and its asset allocation. However, the weakness of their approach is that there is no closed form for their estimator. To circumvent the limitation, Leung, Ng, and Wong (2012) develop a new estimator for the optimal portfolio return based on an unbiased estimator of the inverse of the covariance matrix and its related terms, and derive

explicit formulae for the estimator of the optimal portfolio return. Li, Bai, McAleer, and Wong (2016) further improve the estimation by using the spectral distribution of the sample covariance.

They develop the limiting behavior of the quadratic form with the sample spectral corrected covariance matrix, and explain the superior performance to the sample covariance as the dimension increases to infinity with the sample size proportionally. They deduce the limiting behavior of the expected return and risk on the spectral corrected MV portfolio and illustrate the superior properties of the spectral corrected MV portfolio. Literature of using portfolio optimization in their analysis includes Bai, Liu, and Wong (2009b), Abid, Mroua, and Wong (2009, 2013), Abid, Leung, Mroua, and Wong (2014), Hoang, Lean, and Wong (2015), Hoang, Wong, and Zhu (2015), Li, Li, Hui, and Wong (2018) and many others.

4.2 Stochastic Dominance

Another important area in decision sciences that can use the approaches discussed in our paper to get optimization solutions is stochastic dominance (SD) for different types of investors. Readers may refer to Wong and Li (1999), Li and Wong (1999), Wong (2007), Sriboonchitta, Wong, Dhompongsa, and Nguyen (2009), Levy (2015), Chan, Clark, and Wong (2016), Guo and Wong (2016) for the SD theory for risk averters and risk seekers; refer to Levy and Levy (2002, 2004) and Wong and Chan (2008) for the prospect SD (PSD) and Markowitz SD (MSD) to link to investors with the corresponding S-shaped and reverse S-shaped utility functions; and refer to Leshno and Levy (2002), Guo, Zhu, Wong, and Zhu (2013), Guo, Post, Wong, and Zhu (2014), and Guo, Wong, Zhu (2016) for the theory of almost SD.

The approaches discussed in our paper is useful to the SD theory because there are several SD tests that can be used the approaches discussed in our paper to get optimization solutions. For example, Bai, Li, McAleer, and Wong (2015) extend the SD test statistics developed by Davidson and Duclos (2000) to get SD tests for risk averters and risk seekers, Bai, Li, Liu, and Wong (2011) develop the SD test statistics MSD and PSD, and Ng, Wong, and Xiao (2017) develop the SD test by using quantile regressions. In addition, Lean, Wong, Zhang (2008) have conducted simulation and showed that SD tests introduced by Davidson and Duclos (2000) has better size and power performances than two alternative tests. The approaches discussed in our paper is useful to their SD test statistics.

The SD theory can be used in many areas, including indifference curves (Wong, 2006,

2007; Ma and Wong, 2010; Broll, Egozcue, Wong, and Zitikis, 2010), two-moment decision model (Broll, Guo, Welzel, and Wong, 2015; Guo, Wagener, and Wong, 2018), moment rule (Chan, Chow, Guo, and Wong, 2018), economic growth (Chow, Vieito, and Wong, 2018), diversification (Egozcue and Wong, 2010; Egozcue, Fuentes García, Wong, and Zitikis, 2011; Lozza, Wong, Fabozzi, and Egozcue, 2018). It can also be applied to many different assets, including stock (Fong, Lean, and Wong, 2008), fund (Gasbarro, Wong, and Zumwalt, 2007, 2012; Wong, Phoon, Lean, 2008), futures (Lean, McAleer, Wong, 2010; Lean, Phoon, Wong, 2012; Qiao, Clark, Wong, 2012; Qiao, Wong, Fung, 2013; Lean, McAleer, Wong, 2015; Clark, Qiao, Wong, 2016), Warrant (Chan, de Peretti, Qiao, Wong, 2012; Wong, Lean, McAleer, Tsai, 2018), Option (Abid, Mroua, and 2009), wine (Bouri, Gupta, Wong, and Zhu, 2018), warrants (Chan, de Peretti, Qiao, and Wong, 2012), gold (Hoang, Wong, and Zhu, 2015, 2018; Hoang, Zhu, El Khamlichi, and Wong, 2019), property market (Qiao, Wong, 2015; Tsang, Wong, Horowitz, 2016).

In addition, it can also be used to test for anomaly and market efficient (Lean, Smyth, Wong, 2007; Qiao, Qiao, Wong, 2010), examine different trading strategies (Fong, Wong, and Lean, 2005; Wong, Thompson, Wei, Chow, 2006), banking performance (Broll, Wong, and Wu, 2011), study the effects of financial crisis (Vieito, Wong, Zhu, 2015; Zhu, Bai, Vieito, Wong, 2019), and international trade (Broll, Wahl, and Wong, 2006). In addition, it can also be used to measure income inequality (Valenzuela, Wong, and Zhu, 2019).

4.3 Risk Measures

The third important area discussed in our paper related to decision sciences that can use the approaches discussed in our paper to get optimization solutions is risk measures. We include mean-variance rule as one of the risk measures. Readers may refer to Markowitz (1952) and Wong (2007) for the MV rule for risk averters and risk seekers, respectively, refer to Leung and Wong (2008), Wong, Wright, Yam, and Yung (2012), and the references therein for the Sharpe ratio, refer to Ma and Wong (2010) and the references therein for VaR and conditional-VaR (CVaR), refer to Guo, Jiang, and Wong (2017), Guo, Chan, Wong, and Zhu (2018), and the references therein for the Omega ratio, refer to Niu, Wong, and Xu (2017) and the references therein for the n-order Kappa ratio, refer to Guo, Niu, and Wong (2019) and the references therein for the Farinelli and Tibiletti ratio, and refer to Niu, Guo, McAleer, and Wong (2018), Lu, Yang, Wong (2018), Lu, Hoang, and Wong (2019) and the references therein for the economic performance measure of risk and the economic

index of riskiness, refer to Bai, Wang, Wong (2011), Bai, Hui, Wong, Zitikis (2012) for the mean-variance ratio test, refer to Tang, Sriboonchitta, Ramos, Wong (2014), Ly, Pho, Ly, Wong (2019a,b) for Copulas.

Furthermore, there are other risk measures, for example, Guo, Li, McAleer, Wong, (2018), etc. In addition, there are many applications for the risk measures, see, for example, our discussion in Sections 4.1 and 4.2 for the applications.

4.4 Behavioral Models

We first review the utility functions that are the basics of the behavioral models. Utility starts with Bernoulli (1738) who first notes that people are risk averse. However, academics find that people are not always risk averse or even risk neutral; most people have risk-seeking behavior like buying lottery tickets. Hammond (1974), Stoyan (1983), Wong and Li (1999), Li and Wong (1999), Wong (2007), Levy (2015), Guo and Wong (2016), and others consider investors could be risk-averse or risk-seeking. Markowitz (1952), Levy and Levy (2002, 2004), Wong and Chan (2008) suggest investors could follow S-shaped as well as reverse S-shaped utility functions. Broll, Egozcue, Wong, and Zitikis (2010) and Egozcue, Fuentes Garca, Wong, and Zitikis (2011) further study investment behaviors for investors could follow S-shaped as well as reverse S-shaped utility functions. Guo, Lien, and Wong (2016) develop the exponential utility function with a $2n$ -order approximation for any integer n .

Thompson and Wong (1991), Thompson and Wong (1996), Wong and Chan (2004) and others extend the dividend yield plus growth model (Gordon and Shapiro, 1956) by estimating the cost of capital using discounted cash flow (DCF) methods requires forecasting dividends and proving the existence and uniqueness of the reliability. Lam, Liu, and Wong (2010, 2012), Fung, Lam, Siu, and Wong (1998), and Guo, McAleer, Wong, and Zhu (2017) apply the cost of capital model and use Bayesian models to explain investors' behavioral biases by using the conservatism heuristics and the representativeness heuristics.

Readers may ask why behavioral models can use the approaches discussed in our paper to get optimization solutions? Our answer is that after one develops any behavioral model, one may then develop the corresponding econometric models so that the behavioral models can be estimated. For example, Fabozzi, Fung, Lam, and Wong (2013) extend the models developed by Lam, Liu, and Wong (2010, 2012), Guo, McAleer, Wong, and Zhu (2017) and others by developing 3 tests to test for the magnitude effect of short-term underreaction and

long-term overreaction that can use the approaches discussed in our paper to get optimization solutions. On the other hand, Wong, Chow, Hon, and Woo (2018) conduct a questionnaire survey to examine whether the theory developed by Lam, Liu, and Wong (2008, 2010), and Guo, McAleer, Wong, and Zhu (2017) and others that can use the approaches discussed in our paper to get optimization solutions.

There are many other behavior models also. For example, Egozcue and Wong (2010a) and Egozcue, Fuentes Garca, Wong, and Zitakis (2012a) develop an analytical theory to explain the behavior of investors with extended value functions in segregating or integrating multiple outcomes when evaluating mental accounting. Guo, Wong, Xu, and Zhu (2015), Egozcue, Guo, and Wong (2015), and Guo, and Wong (2019) develop models to investigate regret-averse firms' production and hedging behaviors while Guo, Egozcue, and Wong (2019) develop several properties of using disappointment aversion to model production decision.

4.5 Economic and Financial Indicators

Most of economic and financial indicators could be related to decision sciences and can use the approaches discussed in our paper to get optimization solutions. There are many economic and financial indicators that can use the approaches discussed in our paper to get optimization solutions. We only discuss those related to our work.

We have developed some financial indicators and have applied some economic indicators to study some important economic issues that could be related to decision sciences and can use the approaches discussed in our paper to get optimization solutions. For example, Wong, Chew, and Sikorski (2001) develop a new financial indicator to test the performance of stock market forecasts by using the E/P ratios and bond yields. They also develop two test statistics to utilize the indicator and illustrate the tests in several stock markets. Exploring the characteristics associated with the formation of bubbles that occurred in the Hong Kong stock market in 1997 and 2007 and the 2000 dot-com bubble of Nasdaq, McAleer, Suen, and Wong (2016) establish trading rules that not only produce returns significantly greater than buy-and-hold strategies, but also produce greater wealth compared with TA strategies without trading rules.

In addition, Chong, Cao, and Wong (2017) develop a new market sentiment index for the Hong Kong stock market by using the turnover ratio, short-selling volume, money flow, HIBOR, and returns of the U.S. and Japanese markets, the Shanghai and Shenzhen Compos-

ite indices. Thereafter, they incorporate the threshold regression model with the sentiment index as a threshold variable to capture the state of the Hong Kong stock market. Sethi, Wong, and Acharya (2018) examine the sectoral impact of disinflationary monetary policy by calculating the sacrifice ratios for several OECD and non-OECD countries. Sacrifice ratios calculated through the episode method reveal that disinflationary monetary policy has a differential impact across three sectors in both OECD and non-OECD countries.

4.6 Statistical and Econometric Models

Most of statistical and econometric models could be related to decision sciences and can use the approaches discussed in our paper to get optimization solutions. There are many statistical and econometric models that can use the approaches discussed in our paper to get optimization solutions. We only discuss those related to our work.

We first unit root test, cointegration, causality tests, and nonlinearity Tiku and Wong (1998) develop a unit root test to take care of data follow an AR(1) model. Penm, Terrell, Wong (2003) present simulations and an application that demonstrates the usefulness of the zero-non-zero patterned vector error-correction models (VECMs). Lam, Wong, and Wong (2006) develop some properties on the autocorrelation of the k-period returns for the general mean reversion (GMR) process in which the stationary component is not restricted to the AR(1) process but takes the form of a general ARMA process. Bai, Wong, and Zhang (2010) develop a nonlinear causality test in multivariate settings. Bai, Li, Wong, and Zhang (2011) first discuss linear causality tests in multivariate settings and thereafter develop a nonlinear causality test in multivariate settings.

Bai, Hui, Jiang, Lv, Wong, Zheng (2018) revisit the issue by estimating the probabilities and reestablish the CLT of the new test statistic. Hui, Wong, Bai, and Zhu (2017) propose a quick and efficient method to examine whether a time series possesses any nonlinear feature by testing a kind of dependence remained in the residuals after fitting the dependent variable with a linear model. All the above models are be related to decision sciences and can use the approaches discussed in our paper to get optimization solutions. Literature of applying unit root, cointegration, causality and nonlinearity tests includes Wong, Penm, Terrell, and Lim (2004), Wong, Khan, and Du (2006), Qiao, Liew, and Wong (2007), Foo, Wong, and Chong (2008), Qiao, Smyth, and Wong (2008), Qiao, Chiang, and Wong (2008), Chiang, Qiao, and Wong (2009), Qiao, McAleer, and Wong (2009), Qiao, Li, and Wong (2011), Vieito, Wong,

and Zhu (2015), Batai, Chu, Lv, Wong (2017), Chow, Cunado, Gupta, Wong (2018), Chow, Vieito, Wong (2018), Zhu, Bai, Vieito, Wong (2018), Demirer, Gupta, Lv, Wong (2019), Chow, Gupta, Suleman, Wong (2019), and many others.

We next discuss some robust estimation that can use the approaches discussed in our paper to get optimization solutions. We only discuss those related to our work. Firstly, Wong and Bian (1997) develop an alternative approach to estimate regression coefficients while Wong and Bian (2000) introduce the robust Bayesian estimator developed by Bian and Dickey (1996) to the estimation of the Capital Asset Pricing Model (CAPM) in which the distribution of the error component is well-known to be flat-tailed. Tiku, Wong, Vaughan, and Bian (2000) consider AR(q) models in time series with non-normal innovations represented by a member of a wide family of symmetric distributions (Student's t).

Since the ML (maximum likelihood) estimators are intractable, we derive the MML (modified maximum likelihood) estimators of the parameters and show that they are remarkably efficient. We use these estimators for hypothesis testing, and show that the resulting tests are robust and powerful. Tiku, Wong, and Bian (1999a) extend the work by considering AR(q) models in time series with asymmetric innovations represented by gamma and generalized logistic distributions. Tiku, Wong, and Bian (1999b) estimate coefficients in a simple regression model with autocorrelated errors and the underlying distribution is assumed to follow Student's t family. Wong and Bian (2005) extend the results to the case, where the underlying distribution is a generalized logistic distribution.

We have been developing or applying some other statistical and econometric models that can use the approaches discussed in our paper to get optimization solutions. We state a few here. First, Wong and Miller (1990) develop a theory and methodology for repeated time series (RTS) measurements on autoregressive integrated moving average-noise (ARIMAN) process. Second, Bian, McAleer, and Wong (2011) develop a new test, the trinomial test, for pairwise ordinal data samples to improve the power of the sign test by modifying its treatment of zero differences between observations, thereby increasing the use of sample information. The models in the above papers can use the approaches discussed in our paper to get optimization solutions.

Raza, Sharif, Wong, and Karim (2016) have used maximal overlap discrete wavelet transform (MODWT), wavelet covariance, wavelet correlation, continuous wavelet power spectrum, wavelet coherence spectrum and wavelet-based Granger causality analysis to investi-

gate the empirical influence of tourism development (TD) on environmental degradation in a high-tourist-arrival economy (i.e. United States), using the wavelet transform framework. Xu, Wong, Chen, and Huang (2017) analyze the relationship among stock networks by focusing on the statistically reliable connectivity between financial time series, which accurately reflects the underlying pure stock structure. Tsendsuren, Li, Peng, and Wong (2018) examine the relationships among three health status indicators (self-perceived health status, objective health status, and future health risk) and life insurance holdings in 16 European countries. Mou, Wong, and McAleer (2018) analyze core enterprise credit risks in supply chain finance by means of a 'fuzzy analytical hierarchy process' to construct a supply chain financial credit risk evaluation system, making quantitative measurements and evaluation of core enterprise credit risk.

In addition, Pham, Wong, Moslehpour, and Musyoki (2018) suggest an outsourcing hierarchy model based on the concept of the analytic hierarchy process with four levels of the most concerned attributes: competitiveness, human resources, business environment, and government policies and compare between the analytic hierarchy process (AHP) and Fuzzy AHP show some significant differences but lead to similar conclusions. They provide decision makers an outsourcing hierarchy model based on the AHP and Fuzzy AHP approach with the most concerned factors.

We note that it is not only statistical and econometric models related to decision sciences that can use the approaches discussed in our paper to get optimization solutions. There are many other models, for example, probability and mathematical models that can use the approaches discussed in our paper to get optimization solutions. Many probability and mathematical models can use the approaches discussed in our paper to get optimization solutions because first, the models themselves could use the approaches discussed in our paper to get optimization solutions and second, we will further extend the probability and mathematical models to develop the corresponding statistical and econometric models so that the probability and mathematical models can be used in real data analysis while the corresponding statistical and econometric models could use the approaches discussed in our paper to get optimization solutions.

Over here, we give a few examples. Egozcue, Fuentes García, and Wong (2009) derive some covariance inequalities for monotonic and non-monotonic functions. Egozcue, Fuentes García, Wong, and Zitikis (2010) sharpen the upper bound of a Grss-type covariance inequality by incorporating a notion of quadrant dependence between random variables and

also utilizing the idea of constraining the means of the random variables. Egozcue, Fuentes García, Wong, and Zitikis (2011a) show that Grüss-type probabilistic inequalities for covariances can be considerably sharpened when the underlying random variables are quadrant dependent in expectation (QDE).

Moreover, Egozcue and Wong (2010a) extend prospect theory, mental accounting, and the hedonic editing model by developing an analytical theory to explain the behavior of investors with extended value functions in segregating or integrating multiple outcomes when evaluating mental accounting. Egozcue, Fuentes García, Wong, and Zitikis (2012a) develop decision rules for multiple products, which generally call 'exposure units' to naturally cover manifold scenarios spanning well beyond 'products'. All the above models could use the approaches discussed in our paper to get optimization solutions.

Last, we note that there are many other areas in decision sciences that can use the approaches discussed in our paper to get optimization solutions, in this paper we also review some as discussed in the above. For more applications in decision sciences that can use the approaches discussed in our paper to get optimization solutions, readers may refer to Chang, McAleer, and Wong (2015, 2016, 2018, 2018a, 2018b, 2018c) for more information.

5 Concluding Remarks and Discussion

In order to provide academics and practitioners with an overview of the methods to find the optimization solution, in this article we review of the four most commonly used approaches to find the optimization solution, including the bisection, gradient, Newton-Raphson, and secant methods. We have also developed all the necessary theorems to prove the existence and convergence of the estimate for each method in our paper and give two examples to illustrate our approaches. Since there are many areas in decision sciences that can use the approaches discussed in our paper to get optimization solutions, in this paper we also review the literature of the applications of getting the optimization solutions in some important issues by using the methods discussed in our paper and discuss the advantages and disadvantages of each method.

We note that all the theorems developed in our paper could be well-known but so far we have not seen any book or paper that discusses all the theorems stated in our paper in detail, and thus, we believe the theorems developed in our paper could still have some contributions to the literature. Our review is useful for academics and practitioners in finding

the optimization solutions in their studies.

From our discussion and illustration, all four approaches including bisection method, gradient method, Newton-Raphson method, and secant method can find the optimization solution. We note that the bisection method could get convergent very slow, especially for the analysis in the multi-dimensional studies. The disadvantage of using the gradient method is that it not only depends on the starting value, but also relies on the learning rate. Thus, in order to get good use of this approach, one needs to provide both good starting value and good learning rate. On the other hand, both Newton-Raphson and secant methods only depend on the starting value. The secant method needs to have two starting values, while the Newton-Raphson method only needs to have one starting value.

It can be seen that, the Newton-Raphson method is as one of the most powerful tools to find the optimization solution. This is also the reason why most common optimization software uses the Newton-Raphson algorithm as the foundation. Hence, we recommend utilizing the Newton-Raphson method to find the optimization solution in the estimating problem. Nevertheless, in numerous practical applications, it is very difficult to get the first and second derivative of the objective function. Thus, it is not easy to apply the Newton-Raphson method. To solve this issue, we propose to apply the quasi-Newton method (secant method). We have presented the method in our paper. This is also the reason why most common optimization software uses the quasi-Newton algorithm as the foundation. Hence, we recommend utilizing the quasi-Newton method to find the optimization solution in the estimating problem.

In this regard, one can utilize the secant method or the numerical method for the Newton-Raphson method. Readers may refer (see e.g. Argyros et al. (2019), Berardi et al. (2019), Ferreira et al. (2018), Garijo et al. (2018), Lu et al. (2018), Mokhtari et al. (2018)). The idea of these approaches is the approximation of the derivatives in the Newton-Raphson method. Most of the statistical softwares to find the optimization of the estimating function are based on the secant method or the numerical method for the Newton-Raphson method as the foundation. The ubiquitous functions in the statistical software R to solve the issue consist of `optim`, `maxLik`, `nleqslv` function, etc. (see e.g. Hasselman et al. (2009), Henningsen 2011, mebane et al. (2019), Nash (1990), Toomet et al. (2015), Zhu et al. (1997)).

References

- Abid, F., Mroua, M., Wong, W. K.: The impact of option strategies in financial portfolios performance: Mean-variance and stochastic dominance approaches 23(2), 503-526 (2007).
- Abid, F., Mroua, M., Wong, W. K.: Should Americans invest internationally? Meanvariance portfolios optimization and stochastic dominance approaches. *Risk and Decision Analysis*, 4(2), 89-102 (2013).
- Abid, F., Leung, P., Mroua, M., Wong, W.: International diversification versus domestic diversification: Mean-variance portfolio optimization and stochastic dominance approaches. *Journal of Risk and Financial Management*, 7(2), 45-66 (2014).
- Argyros, I., Shakhno, S., Shunkin, Y.: Improved Convergence Analysis of Gauss-Newton-Secant Method for Solving Nonlinear Least Squares Problems. *Mathematics*, 7(1), 99 (2019).
- Bai, Z., Hui, Y., Jiang, D., Lv, Z., Wong, W. K., Zheng, S.: A new test of multivariate nonlinear causality. *PloS one*, 13(1), (2018).
- Bai, Z., Hui, Y., Wong, W. K., Zitikis, R.: Prospect performance evaluation: Making a case for a non-asymptotic UMPU test. *Journal of Financial Econometrics*, 10(4), 703-732 (2012).
- Bai, Z.D., Li, H., Liu, H.X., Wong, W.K.: Test statistics for prospect and Markowitz stochastic dominances with applications. *Econometrics Journal*, 122, 1-26 (2011).
- Bai, Z.D., Li, H., McAleer, M., Wong, W.K.: Stochastic dominance statistics for risk averters and risk seekers: An analysis of stock preferences for USA and China. *Quantitative Finance*, 15(5), 889-900 (2015).
- Bai, Z.D., Li, H., McAleer, M., Wong, W.K.: Spectrally-corrected estimation for high-dimensional Markowitz mean-variance optimization. Tinbergen Institute Discussion Paper, TI 2016-025/III (2016).
- Bai, Z.D., Li, H., Wong, W.K., Zhang, B.Z.: Multivariate Causality Tests with Simulation and Application, *Statistics and Probability Letters*, 81(8), 1063-1071 (2011).
- Bai, Z.D., Liu, H.X., Wong, W.K.: Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance*, 19(4), 639-667

(2009a).

- Bai, Z.D., Liu, H.X., Wong, W.K.: On the Markowitz mean-variance analysis of self-financing portfolios. *Risk and Decision Analysis*, 1(1), 35-42 (2009b).
- Bai, Z.D., Phoon, K.F., Wang, K.Y., Wong, W.K.: The performance of commodity trading advisors: A mean-variance-ratio test approach. *North American Journal of Economics and Finance*, 25, 188-201 (2013).
- Bai, Z.D., Wang, K.Y., Wong, W.K.: Mean-variance ratio test, a complement to coefficient of variation test and Sharpe ratio test. *Statistics and Probability Letters*, 81(8), 1078-1085 (2011).
- Bai, Z.D., Wong, W.K., Zhang, B.Z.: Multivariate Linear and Non-Linear Causality Tests, *Mathematics and Computers in Simulation* 81, 5-17 (2010).
- Batai, A., Chu, A., Lv, Z., Wong, W. K.: China's impact on Mongolian Exchange Rate, *Journal of Management Information and Decision Sciences* 20(1), 1-22 (2007).
- Munkh-Ulzii, B., McAleer, M., Moslehpour, M., Wong, W. K.: Confucius and herding behaviour in the stock markets in China and Taiwan. *Sustainability*, 10(12), 4413 (2018).
- Berardi, M., Difonzo, F., Notarnicola, F., Vurro, M.: A transversal method of lines for the numerical modeling of vertical infiltration into the vadose zone. *Applied Numerical Mathematics*, 135, 264-275 (2019).
- Bian, G., McAleer, M., Wong, W.K.: A Trinomial Test for Paired Data When There are Many Ties, *Mathematics and Computers in Simulation* 81(6), 1153-1160 (2011).
- Bian, G., McAleer, M., Wong, W. K.: Robust estimation and forecasting of the capital asset pricing model. *Annals of Financial Economics*, 8(02), 1350007 (2013).
- Bian, G., Wong, W.K.: An Alternative Approach to Estimate Regression Coefficients, *Journal of Applied Statistical Science*, 6(1), 21-44 (1997).
- Bolzano, B.: Untersuchungen zur Grundlegung der sthetik (Vol. 1). *Athenum-Verlag* (1972).
- Bouri, E., Gupta, R., Wong, W.K., Zhu, Z.Z.: Is Wine a Good Choice for Investment? *Pacific-Basin Finance Journal*, 51, 171-183 (2018).
- Broll, U., Egozcue, M., Wong, W.K., Zitikis, R.: Prospect theory, indifference curves, and hedging risks. *Applied Mathematics Research Express*, 2010(2), 142-153 (2010).

- Broll, U., Guo, X., Welzel, P., Wong, W.K.: The banking firm and risk taking in a two-moment decision model. *Economic Modelling*, 50, 275-280 (2015).
- Broll, U., Wahl, J.E., Wong, W.K.: Elasticity of risk aversion and international trade. *Economics Letters*, 91(1), 126-130 (2006).
- Broll, U., Wong, W.K., Wu, M.: Banking firm, risk of investment and derivatives. *Technology and Investment*, 2, 222-227 (2011).
- Broyden, C. G.: Quasi-Newton methods and their application to function minimisation. *Mathematics of Computation*, 21(99), 368-381 (1967).
- Cauchy, A. L.: Cours d'analyse de l'Ecole royale polytechnique; par m. Augustin-Louis Cauchy... 1. re partie. Analyse algébrique. de l'Imprimerie royale (1821).
- Chalco-Cano, Y., Silva, G. N., Rufin-Lizana, A.: On the Newton method for solving fuzzy optimization problems. *Fuzzy Sets and Systems*, 272, 60-69 (2015).
- Chan, C.-Y., de Peretti, C., Qiao, Z. Wong, W.K.: Empirical test of the efficiency of the UK covered warrants market: Stochastic dominance and likelihood ratio test approach, *Journal of Empirical Finance*, 19(1), 162-174 (2012).
- Chan, R.H., Chow, S.C., Guo, X., Wong, W.K.: Central Moments, Stochastic Dominance, Moment Rule, and Diversification with Application, the International Conference on Scientific Computing, in honor of Professor Raymond Chan's 60th birthday, to be held on December 5-8, 2018 in the Chinese University of Hong Kong (2018).
- Chan, R.H., Clark, E., Wong, W.K.: On the Third Order Stochastic Dominance for Risk-Averse and Risk-Seeking Investors with Analysis of their Traditional and Internet Stocks, MPRA Paper No. 75002. University Library of Munich, Germany (2016).
- Chang, C.L., McAleer, M., Wong, W.K.: Informatics, Data Mining, Econometrics and Financial Economics: A Connection. Technical Report 1, 2015.
Available online: <https://repub.eur.nl/pub/79219/> (accessed on 8 March 2019).
- Chang, C.L., McAleer, M., Wong, W.K.: Big Data, Computational Science, Economics, Finance, Marketing, Management, and Psychology: Connections. *Journal of Risk and Financial Management* 11(1), 15 (2018); <https://doi.org/10.3390/jrfm11010015>
- Chang, C.L., McAleer, M., Wong, W.K.: Big data, computational science, economics, finance, marketing, management, and psychology: connections. *Journal of Risk and Fi-*

- nancial Management* 11: 15 (2018a).
- Chang, C.L., McAleer, M., Wong, W.K.: Decision Sciences, Economics, Finance, Business, Computing, and Big Data: Connections. Available online: <https://papers.ssrn.com/sol3/papers> (2018b).
- Chang, C.L., McAleer, M., Wong, W.K.: Management Information, Decision Sciences, and Financial Economics: A Connection. Decision Sciences, and Financial Economics: A Connection (January 17, 2018). Tinbergen Institute Discussion Paper, p. 4 (2018c).
- Chiang, T.C., Qiao, Z., Wong, W.K.: New Evidence on the Relation between Return Volatility and Trading Volume, *Journal of Forecasting*, 29(5), 502 - 515 (2009).
- Chin, W. S., Yuan, B. W., Yang, M. Y., Lin, C. J.: An efficient alternating newton method for learning factorization machines. *ACM Transactions on Intelligent Systems and Technology* (TIST), 9(6), 72 (2018).
- Chow, S.C., Juncal C., Rangan G., Wong W.K.: Causal Relationships between Economic Policy Uncertainty and Housing Market Returns in China and India: Evidence from Linear and Nonlinear Panel and Time Series Models, *Studies in Nonlinear Dynamics and Econometrics*, 22(2), <https://doi.org/10.1515/snde-2016-0121> (2018).
- Chow, S.C., Rangan G., Tahir S., Wong W.K.: Long-Run Movement and Predictability of Bond Spread for BRICS and PIIGS: The Role of Economic, Financial and Political Risks, *Journal of Reviews on Global Economics*, 8, 239-257 (2019).
- Chow, S.C., Joao P.V., Wong W.K.: Do both demand-following and supply-leading theories hold true in developing countries?, *Physica A: Statistical Mechanics and its Applications* 513, 536-554 (2018).
- Chow, S.C., Vieito, J.P., Wong, W.K.: Do both demand-following and supply-leading theories hold true in developing countries?, *Physica A: Statistical Mechanics and its Applications*, 513, 536-554 (2018).
- Clark, E.A., Qiao, Z., Wong, W.K.: Theories of risk: testing investor behaviour on the Taiwan stock and stock index futures markets. *Economic Inquiry* 54(2), 907-924 (2016).
- Davidson, R., Duclos, J.Y.: Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, 68, 1435-1464 (2000).
- Demirer , R., Gupta, R., Lv, Z.H., Wong, W.K.: Equity Return Dispersion and Stock

- Market Volatility: Evidence from Multivariate Linear and Nonlinear Causality Tests, *Sustainability*, 11(2), 351 (2019); <https://doi.org/10.3390/su11020351>.
- Egozcue M., Luis F.G., Wong W.K.: On some Covariance Inequalities for Monotonic and Non-monotonic Functions, *Journal of Inequalities in Pure and Applied Mathematics*, 10(3), Article 75, 1-7 (2009).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: Grüss-type Bounds for the Covariance of Transformed Random Variables, *Journal of Inequalities and Applications*, Volume 2010, Article ID 619423, 1-10 (2010).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: Do Investors Like to Diversify? A Study of Markowitz Preferences, *European Journal of Operational Research* 215(1), 188-193 (2011).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: Grüss-type bounds for covariances and the notion of quadrant dependence in expectation, *Central European Journal of Mathematics* 9(6), 1288-1297 (2011).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: The covariance sign of transformed random variables with applications to economics and finance, *IMA Journal of Management Mathematics* 22(3), 291-300 (2011).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: The smallest upper bound for the p th absolute central moment of a class of random variables, *The Mathematical Scientist* 37, 1-7 (2012).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: Integration-segregation decisions under general value functions: 'Create your own bundle-choose 1, 2, or all 3! *IMA Journal of Management Mathematics*, 1 of 16, doi:10.1093/imaman/dps024, (2012).
- Egozcue M., Luis F.G., Wong W.K., Zitikis R.: Convex combinations of quadrant dependent copulas, *Applied Mathematics Letters* 26(2), 249-251 (2013).
- Egozcue M., Xu G., Wong W.K.: Optimal Output for the Regret-Averse Competitive Firm Under Price Uncertainty, *Eurasian Economic Review* 5(2), 279-295 (2015).
- Egozcue M., Wong W.K.: Gains from Diversification on Convex Combinations: A Majorization and Stochastic Dominance Approach, *European Journal of Operational Research* 200(3), 893-900 (2010).

- Egozcue M., Wong W.K.: Segregation and Integration: A Study of the Behaviors of Investors with Extended Value Functions, *Journal of Applied Mathematics and Decision Sciences*, Volume 2010, Article ID 302895, 1-8 (2010a).
- Exl, L., Fischbacher, J., Kovacs, A., Oezelt, H., Gusenbauer, M., Schrefl, T.: Preconditioned nonlinear conjugate gradient method for micromagnetic energy minimization. *Computer Physics Communications*, 235, 179-186 (2019).
- Fabozzi, F.J., Fung, C.Y., Lam, K., Wong, W.K.: Market Overreaction and Underreaction: Tests of the Directional and Magnitude Effects, *Applied Financial Economics* 23(18), 1469-1482 (2013).
- Ferreira, O. P., Silva, G. N.: Inexact Newton method for non-linear functions with values in a cone. *Applicable Analysis*, 1-17 (2018).
- Fong, W.M., H.H. Lean, Wong W.K.: Stochastic dominance and behavior towards risk: The market for internet stocks. *Journal of Economic Behavior and Organization*, 68(1), 194-208 (2008).
- Fong, W.M., Wong, W.K., Lean, H.H.: International momentum strategies: A stochastic dominance approach. *Journal of Financial Markets*, 8, 89-109 (2005).
- Fung, E.S., Lam, K., Siu, T.K., Wong, W.K.: A New Pseudo Bayesian Model for Financial Crisis, *Journal of Risk and Financial Management* 4, 42-72 (2011).
- Gasbarro, D., Wong, W.K., Zumwalt, J.K. Stochastic dominance analysis of iShares. *European Journal of Finance*, 13, 89-101 (2007).
- Gasbarro, D., Wong, W.K., Zumwalt, J.K.: Stochastic dominance and behavior towards risk: The market for iShares. *Annals of Financial Economics*, 7(1), 1250005-1-20 (2012).
- Guo, X., Raymond H. C., Wong W.K., Lixing Z.: Mean-Variance, Mean-VaR, Mean-CVaR Models for Portfolio Selection with Background Risk, *Risk Management*, <https://doi.org/10.1057/s41283-018-0043-2> (2018).
- Guo, X., Martin E., Wong W.K.: Optimal Production Decision with Disappointment Aversion under Uncertainty, *International Journal of Production Research*, second revision (2019).
- Guo, X., Jiang, X.J., Wong, W.K.: Stochastic Dominance and Omega Ratio: Measures to Examine Market Efficiency, Arbitrage Opportunity, and Anomaly, *Economies* 5, no. 4:

38 (2017).

- Guo, X., Li, G.-R., McAleer, M., Wong, W.K.: Specification Testing of Production in a Stochastic Frontier Model, *Sustainability*, 10, 3082; doi:10.3390/su10093082 (2018).
- Guo, X., McAleer, M., Wong, W.K., Zhu, L.X. A Bayesian approach to excess volatility, short-term underreaction and long-term overreaction during financial crises, *North American Journal of Economics and Finance* 42, 346-358 (2017).
- Guo, X., Cuizhen N., Wong W.K. Farinelli and Tibiletti ratio and Stochastic Dominance, *Risk Management*, forthcoming (2019).
- Guo, X., Post, T. Wong, W.K., Zhu, L.X.: Moment conditions for almost stochastic dominance. *Economics Letters* 124(2), 163-167 (2014).
- Guo, X., Wagener, A., Wong, W.K.: The Two-Moment Decision Model with Additive Risks, *Risk Management* 20(1), 77-94 (2018).
- Guo, X., Wong, W.K. Multivariate stochastic dominance for risk averters and risk seekers. *RAIRO - Operations Research* 50(3), 575-586 (201).
- Guo, X., Wong, W.K., Qunfang X., Xuehu Z.: Production and Hedging Decisions under Regret Aversion, *Economic Modelling* 51, 153-158 (2015).
- Guo, X., Wong, W.K., Zhu, L.X.: Almost stochastic dominance for risk averters and risk seekers. *Finance Research Letters* 19, 15-21 (2016).
- Guo, X., Zhu, X.H., Wong, W.K., Zhu, L.X.: A note on almost stochastic dominance. *Economics Letters*, 121(2), 252-256 (2013).
- Hammond, J.S.: Simplifying the choice between uncertain prospects where preference is nonlinear. *Management Science*, 20(7), 1047-1072 (1974).
- Hasselman, B.: nleqslv: Solve systems of non linear equations. R package version, 1 (2009).
- Henningsen, A., Toomet, O. maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3), 443-458 (2011).
- Hoang, T.H.V., Lean, H.H., Wong, W.K.: Is gold good for portfolio diversification? A stochastic dominance analysis of the Paris stock exchange, *International Review of Financial Analysis*, 42, 98-108 (2015).
- Hoang, V.T.H., Wong, W.K., Zhu, Z.Z.: Is gold different for risk-averse and risk-seeking

- investors? An empirical analysis of the Shanghai Gold Exchange. *Economic Modelling*, 50, 200-211 (2015).
- Hoang, V.T.H., Wong, W.K., Zhu, Z.Z.: The seasonality of gold prices in China: Does the risk-aversion level matter? *Accounting and Finance*, <https://doi.org/10.1111/acfi.12396> (2018).
- Hoang, V.T.H., Zhu, Z.Z., El Khamlichi, A, Wong, W.K. Does the Shari'ah Screening Impact the Gold-Stock Nexus? A Sectorial Analysis, *Resources Policy*, forthcoming (2019).
- Hsieh, S., Lee, S., Shen, P. S.: Logistic regression analysis of randomized response data with missing covariates. *Journal of Statistical Planning and Inference*, 140(4), 927-940 (2010).
- Hsieh, S. H., Lee, S. M., Shen, P.: Semiparametric analysis of randomized response data with missing covariates in logistic regression. *Computational Statistics and Data Analysis*, 53(7), 2673-2692 (2009).
- Hui, Y.C., Wong, W.K., Bai, Z.D., Zhu, Z.Z. A New Nonlinearity Test to Circumvent the Limitation of Volterra Expansion with Application, *Journal of the Korean Statistical Society*, 46(3), 365-374 (2017).
- Kien P.V., Wong W.K, Moslehpour M., Musyoki D.: Simultaneous Adaptation of AHP and Fuzzy AHP to Evaluate Outsourcing Services in East and Southeast Asia, *Journal of Testing and Evaluation*, <https://doi.org/10.1520/JTE20170420>. ISSN 0090-3973 (2018).
- Lam, K., Liu, T.S., Wong, W.K.: A pseudo-Bayesian model in financial decision making with implications to market volatility, under- and overreaction, *European Journal of Operational Research* 203(1),166-175 (2010).
- Lam, K., Liu, T.S., Wong, W.K.: A New Pseudo Bayesian Model with Implications to Financial Anomalies and Investors' Behaviors, *Journal of Behavioral Finance* 13(2), 93-107 (2012).
- Lam, K., Wong, C.M., Wong, W.K. New variance ratio tests to identify random walk from the general mean reversion model, *Journal of Applied Mathematics and Decision Sciences/Advances in Decision Sciences*, 1-21 (2006).
- Lean, H.H., McAleer, M., Wong, W.K.: Market efficiency of oil spot and futures: A mean-variance and stochastic dominance approach. *Energy Economics*, 32, 979-986 (2010).
- Lean, H.H., McAleer, M., Wong, W.K.: Preferences of risk-averse and risk-seeking investors

- for oil spot and futures before, during and after the Global Financial Crisis. *International Review of Economics and Finance*, 40, 204-216 (2015).
- Lean, H.H., Phoon K.F, Wong W.K.: Stochastic dominance analysis of CTA funds. *Review of Quantitative Finance and Accounting*, 40(1), 155-170 (2012).
- Lean, H.H., Smyth, R. Wong, W.K.: Revisiting calendar anomalies in Asian stock markets using a stochastic dominance approach. *Journal of Multinational Financial Management*, 17(2), 125-141 (2007).
- Lean, H.H., Wong, W.K., Zhang, X.B.: The sizes and powers of some stochastic dominance tests: A Monte Carlo study for correlated and heteroskedastic distributions, *Mathematics and Computers in Simulation* 79, 30-48 (2008).
- Lee, S. M., Hwang, W. H., de Dieu Tapsoba, J.: Estimation in closed capture-recapture models when covariates are missing at random. *Biometrics*, 72(4), 1294-1304 (2016).
- Lee, S. M., Li, C. S., Hsieh, S. H., Huang, L. H.: Semiparametric estimation of logistic regression model with missing covariates and outcome. *Metrika*, 75(5), 621-653 (2012).
- Leshno, M., Levy, H.: Preferred by “all” and preferred by “most” decision makers: Almost stochastic dominance. *Management Science*, 48(8), 1074-1085 (2002).
- Leung, P.L., Ng, H.Y., Wong, W.K.: An Improved Estimation to Make Markowitz’s Portfolio Optimization Theory Users Friendly and Estimation Accurate with Application on the US Stock Market Investment, *European Journal of Operational Research* 222(1), 85-95 (2012).
- Leung, P.L., Wong, W.K.: On testing the equality of the multiple Sharpe ratios, with application on the evaluation of Ishares. *Journal of Risk*, 10(3), 1-16 (2008).
- Levy, H.: Stochastic Dominance: Investment Decision Making Under Uncertainty. Third Edition, Springer, New York (2015).
- Levy, H., M. Levy.: Prospect theory and mean-variance analysis. *Review of Financial Studies*, 17(4), 1015-1041 (2004).
- Levy, M., H. Levy.: Prospect theory: Much ado about nothing? *Management Science*, 48(10), 1334-1349 (2002).
- Li, Z., Li, X., Hui, Y.C., Wong, W.K.: Maslow Portfolio Selection for Individuals with Low Financial Sustainability, *Sustainability* 10(4), 1128; <https://doi.org/10.3390/su10041128>

(2018).

- Li, C.K., W.K. Wong.: Extension of stochastic dominance theory to random variables. *RAIRO - Operations Research*, 33(4), 509-524 (1999).
- Lozza, S.,O., Wong, W.K., Fabozzi, F.J., Egozcue, M.: Diversification versus Optimal: Is There Really a Diversification Puzzle? *Applied Economics*, 50(43), 4671-4693 (2018).
- Lu, H., Zhang, J., Yang, F., Xu, B., Liu, Z., Zheng, Z., Wen, X.: Improved secant method for getting proper initial magnetization in transformer DC bias simulation. *International Journal of Electrical Power & Energy Systems*, 103, 50-57 (2018).
- Lu R., Vu T.H., Wong W.K.: Does Lump-Sum Investing Strategy Outperform Dollar-Cost Averaging Strategy in Uptrend Markets?, *Studies in Economics and Finance*, second revision (2019).
- Lu R., Yang C.C., Wong W.K.: Time Diversification: Perspectives from the Economic Index of Riskiness, *Annals of Financial Economics* 13(3), 1850011 (2018).
- Lukusa, T. M., Lee, S. M., Li, C. S.: Semiparametric estimation of a zero-inflated Poisson regression model with missing covariates. *Metrika*, 79(4), 457-483 (2016).
- Ly, S., Pho, K.H., Ly, S., Wong, W.K.: Determining Distribution for the Product of Random Variables by Using Copulas, *Risks*, 7(1), 23 (2019).
- Ly, S., Pho, K.H., Ly, S., Wong, W.K.: Distribution of Quotient of Dependent and Independent Random Variables Using Copulas, *Journal of Risk and Financial Management*, 12(1), 42 (2019).
- Ma, C., Wong, W.K.: Stochastic dominance and risk measure: A decision-theoretic foundation for VaR and C-VaR. *European Journal of Operational Research*, 207(2), 927-935 (2010).
- Markowitz, H.M.: Portfolio selection. *Journal of Finance*, 7, 77-91 (1952).
- Mebane Jr, W. R.: R Version of GENetic Optimization Using Derivatives (2019).
- Mokhtari, A., Eisen, M., Ribeiro, A.: IQN: An incremental quasi-Newton method with local superlinear convergence rate. *SIAM Journal on Optimization*, 28(2), 1670-1698 (2018).
- Mou W., Wong W.K., McAleer M.: Financial Credit Risk Evaluation Based on Core Enterprise Supply Chains, *Sustainability* 10(10), 3699 (2018);

<https://doi.org/10.3390/su10103699>.

- Nash, J. C.: Compact numerical methods for computers: linear algebra and function minimisation. *CRC press* (1990).
- Ng, P., Wong, W.K., Xiao, Z.J.: Stochastic Dominance Via Quantile Regression, *European Journal of Operational Research* 261(2), 666-678 (2017).
- Niu, C.Z., Guo, X., McAleer, M., Wong, W.K.: Theory and Application of an Economic Performance Measure of Risk, *International Review of Economics & Finance* 56, 383-396 (2018).
- Niu, C.Z., Wong, W.K., Xu, Q.F.: Kappa Ratios and (Higher-Order) Stochastic Dominance, *Risk Management* 19(3), 245-253 (2017).
- Pho, K. H., Nguyen, V. T.: Comparison of Newton-Raphson Algorithm and Maxlik Function. *Journal of Advanced Engineering and Computation*, 2(4), 281-292 (2018).
- Qiao, Z., Clark, E., Wong, W.K.: Investors' preference towards risk: Evidence from the Taiwan stock and stock index futures markets. *Accounting Finance*, 54(1), 251-274 (2012).
- Qiao, Z., Chiang, T.C., Wong, W.K.: Long-run equilibrium, short-term adjustment, and spillover effects across Chinese segmented stock markets, *Journal of International Financial Markets, Institutions & Money* 18, 425-437 (2008).
- Qiao, Z., Li, Y.M., Wong, W.K.: Policy Change and Lead-Lag Relations among China's Segmented Stock Markets, *Journal of Multinational Financial Management* 18, 276-289 (2008).
- Qiao, Z., Li, Y.M., Wong, W.K.: Regime-dependent relationships among the stock markets of the US, Australia, and New Zealand: A Markov-switching VAR approach, *Applied Financial Economics* 21(24), 1831-1841 (2011).
- Qiao, Z., Liew, V.K.S., Wong, W.K.: Does the US IT Stock Market Dominate Other IT Stock Markets: Evidence from Multivariate GARCH Model, *Economics Bulletin*, 6(27), 1-7 (2007).
- Qiao, Z., McAleer, M., Wong, W.K.: Linear and nonlinear causality between changes in consumption and consumer attitudes, *Economics Letters* 102(3), 161-164 (2009).
- Qiao, Z., Qiao, W.W., Wong, W.K.: Examining the Day-of-the-Week Effects in Chinese Stock Markets: New Evidence from a Stochastic Dominance Approach, *Global Economic*

- Review* 39(3), 225-246 (2010).
- Qiao, Z., Smyth, R., Wong, W.K., (2008), Volatility Switching and Regime Interdependence Between Information Technology Stocks 1995-2005, *Global Finance Journal*, 19, 139-156.
- Qiao, Z., Wong, W.K.: Which is a better investment choice in the Hong Kong residential property market: A big or small property? *Applied Economics*, 47(16), 1670-1685 (2015).
- Qiao, Z., Wong, W.K., Fung, J.K.W. Stochastic dominance relationships between stock and stock index futures markets: International evidence. *Economic Modelling*, 33, 552-559 (2013).
- Reddy, S. S., Bijwe, P. R.: An efficient optimal power flow using bisection method. *Electrical Engineering*, 100(4), 2217-2229 (2018).
- Sahu, H. S., Kumar, S., Nayak, S. K.: Maximum Power Point Estimation of a PV Array by Using Improve Bisection Method. In 2018 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific) (pp. 1-5). IEEE (2018).
- Syed A.R., Arshian S., Wong W.K., Mohd Z.A.K.: Tourism Development and Environmental Degradation in United States: Evidence from Wavelet based Analysis, *Current Issues in Tourism*, 20(16), 1768-1790 (2016).
- Sriboonchitta, S., Wong W.K., Dhompongsa S., Nguyen H.T.: Stochastic Dominance and Applications to Finance, Risk and Economics, Chapman and Hall/CRC, *Taylor and Francis*, Boca Raton, Florida, USA (2009).
- Stoyan, D.: Comparison Methods for Queues and Other Stochastic Models, (New York:Wiley) (1983).
- Tang, J., Sriboonchitta, S., Ramos, V., Wong, W.K.: Modelling dependence between tourism demand and exchange rate using copula-based GARCH model, *Current Issues in Method and Practice* 19(9), 1-19 (2014).
- Terlaky, T.: Interior point methods of mathematical programming (Vol. 5). *Springer Science & Business Media* (2013).
- Thompson, H.E., Wong, W.K.: On the unavoidability of “unscientific” judgement in estimating the cost of capital. *Managerial and Decision Economics* 12, 27-42 (1991).
- Thompson, H.E., Wong, W.K.: Revisiting ‘Dividend Yield Plus Growth’ and Its Applicability, *Engineering Economist*, 41(2), 123-147 (1996).

- Toomet, O., Henningsen, A., & Toomet, M. O.: Package ?maxLik? (2015).
- Tiku, M.L., Wong, W.K.: Testing for unit root in AR(1) model using three and four moment approximations, *Communications in Statistics: Simulation and Computation*, 27(1), 185-198 (1998).
- Tiku, M.L., Wong W.K., Bian, G.: Time series models with asymmetric innovations, *Communications in Statistics: Theory and Methods*, 28(6), 1331-1360 (1999).
- Tiku, M.L., Wong W.K., Bian, G.: Estimating Parameters in Autoregressive Models in Non-normal Situations: symmetric Innovations, *Communications in Statistics: Theory and Methods*, 28(2), 315-341 (1999a).
- Tiku, M.L., Wong, W.K., Vaughan, D.C., Bian, G.: Time series models in non-normal situations: Symmetric innovations, *Journal of Time Series Analysis*, 21, 571-96 (2000).
- Tsang, C.K., Wong, W.K., Horowitz, I.: Arbitrage opportunities, efficiency, and the role of risk preferences in the Hong Kong property market. *Studies in Economics and Finance* 33(4), 735-754 (2016).
- Tsendsuren S., Li C.S, Peng S.C., Wong W.K.: The Effects of Health Status on Life Insurance Holdings in 16 European Countries, *Sustainability* 10(10), 3454 (2018); <https://doi.org/10.3390/su10103454>.
- Vieito, J.P., Wong, W.K., Zhu, Z.Z.: Could The Global Financial Crisis Improve The Performance of The G7 Stocks Markets? *Applied Economics* 48(12) 1066-1080 (2015).
- Valenzuela, M.R., Wong, W.K., Zhu, Z.Z.: Is it really that bad? Testing for richness and pooriness in the Philippines, *World Economy*, forthcoming (2019).
- Wang, C. Y., Chen, J. C., Lee, S. M., Ou, S. T.: Joint conditional likelihood estimator in logistic regression with missing covariate data. *Statistica Sinica*, 555-574 (2002).
- Wong, W.K.: Stochastic Dominance Theory for Location-Scale Family, *Journal of Applied Mathematics and Decision Sciences* 2006, 1-10 (2006).
- Wong, W.K.: Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. *European Journal of Operational Research*, 182(2), 829-843 (2007).
- Wong, W.K., Chan, R.: Markowitz and prospect stochastic dominances. *Annals of Finance*, 4(1), 105-129 (2008).

- Wong, W.K., Bian G.: Robust Estimation in Capital Asset Pricing Estimation, *Journal of Applied Mathematics & Decision Sciences*, 4(1), 65–82 (2000).
- Wong, W.K., Bian G.: Estimating Parameters in Autoregressive Models with asymmetric innovations, *Statistics and Probability Letters*, 71(1), 61-70 (2005).
- Wong, W.K., Chan R.: On the estimation of cost of capital and its reliability, *Quantitative Finance*, 4(3), 365 - 372 (2004).
- Wong, W.K., Chew, B.K., Sikorski, D.: Can P/E ratio and bond yield be used to beat stock markets?. *Multinational Finance Journal*, 5(1), 59-86 (2001).
- Wong, W.K., Chow, S.C., Hon, T.Y., Woo, K.Y.: Empirical Study on Conservative and Representative Heuristics of Hong Kong Small Investors Adopting Momentum and Contrarian Trading Strategies, *International Journal of Revenue Management*, 10(2), (2018); <https://doi.org/10.1504/IJRM.2018.091836>.
- Wong, W.K., Lean, H.H., McAleer, M., Tsai, F.T.: Why are Warrant Markets Sustained in Taiwan but not in China?, *Sustainability* 10(10), 3748 (2018); <https://doi.org/10.3390/su10103748>.
- Wong, W.K., C.K. Li.: A note on convex stochastic dominance theory. *Economics Letters*, 62, 293-300 (1999).
- Wong, W.K., Ma, C.: Preferences over location-scale family. *Economic Theory*, 37(1), 119-146 (2008).
- Wong, W.K., Miller R.B.: Analysis of ARIMA-Noise Models with Repeated Time Series, *Journal of Business and Economic Statistics*, 8(2), 243–250 (1990).
- Wong, W.K., Penm, J.H.W., Terrell, R.D. Lim, K.Y.C.: The Relationship between Stock Markets of Major Developed Countries and Asian Emerging Markets, *Advances in Decision Sciences* 8(4), 201-218 (2004).
- Wong, W.K., Phoon, K.F., Lean, H.H.: Stochastic dominance analysis of Asian hedge funds. *Pacific-Basin Finance Journal*, 16(3), 204-223 (2008).
- Wong, W.K., Thompson H.E., Wei S., Chow Y.F.: Do Winners perform better than Losers? A Stochastic Dominance Approach, *Advances in Quantitative Analysis of Finance and Accounting*, 4, 219-254 (2006).
- Wong, W.K., Wright J.A., Yam S.C.P., Yung S.P.: A mixed Sharpe ratio. *Risk and Decision*

Analysis, 3(1-2), 37-65 (2012).

Wu, X., Wen, Z., Bao, W.: A regularized Newton method for computing ground states of Bose-Einstein condensates. *Journal of Scientific Computing*, 73(1), 303-329 (2017).

Xu R.H., Wong W.K., Chen G., Huang S.: Topological Characteristics of the Hong Kong Stock Market: A Test-based P-threshold Approach to Understanding Network Complexity, *Scientific Reports* 7, 41379 (2017); doi:10.1038/srep41379.

Yang, K., Jiang, G. H., Qu, Q., Peng, H. F., Gao, X. W.: A new modified conjugate gradient method to identify thermal conductivity of transient non-homogeneous problems based on radial integration boundary element method. *International Journal of Heat and Mass Transfer*, 133, 669-676 (2019).

Zhu, Z.Z., Bai, Z.D., Vieito, J.P., Wong, W.K.: The Impact of the Global Financial Crisis on the Efficiency of Latin American Stock Markets, *Estudios de Economía*, forthcoming (2019).

Zhu, C., Byrd, R. H., Lu, P., Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software* (TOMS), 23(4), 550-560 (1997).