وزارة التعليم العالي والبحث العلمي

جامعة الموصل

كلية علوم الحاسوب والرياضيات

طرائق حصينة لانحدار المكونات الرئيسية: دراسة مقارنة مع المحاكاة والتطبيق

إســـراء نجيب سعيد محمود الصراف

رسالة ماجستير

الإحصاء

بإشراف

الأستاذ المساعد الدكتور

بشار عبد العزيز مجيد الطالب

1442 هـ

2021 م

المستخلص

تقوم فكرة الرسالة على اقتراح بعض الطرائق الحصينة في انحدار المكونات الرئيسية لإيجاد مقدرات الانحدار واختيار المقدرات الأكثر حصانة ضد وجود القيم الشاذة، ويعد انحدار المكونات الرئيسية أحد طرائق التقدير المستخدمة التي تعالج مشكلة تعدد العلاقة الخطية والتي تحدث في كثير من النماذج القياسية، وجوهر هذه الطريقة أنها تستبدل المتغيرات المستقلة التي تعاني من مشكلة تعدد العلاقة الخطية بمكونات رئيسية تعبر عنها لتوفيق نموذج الانحدار في حالة التعدد الخطي، ومن ثم استخدام دوال أوزان حصينة الخاصة بمقدري **M** و **MM** الحصين لوزن دالة الهدف وهي (Huber، Hampel، Bisquare) للتعامل مع وجود القيم الشاذة في البيانات، ومن أجل التحقق من كفاءة المقدرات تم إجراء دراسة تجريبية من خلال أسلوب المحاكاة بأحجام عينات ونسب تلويث مختلفة، كما تم تطبيق الطرائق على بيانات حقيقية تم جمعها من ملفات معمل اسمنت بادوش في محافظة نينوى للفترة من (2008-2014)، وقد تم اختبار فيما إذا كانت البيانات تعاني من مشكلة تعدد العلاقة الخطية ومن ثم تطبيق المربعات الصغرى باستخدام المكونات الرئيسية كمتغيرات مستقلة وتقدير النموذج، وقد وجد أن المتغيرات تعاني من مشكلة تعدد العلاقة الخطية، وتمت المعالجة عن طريق تطبيق انحدار المكونات الرئيسية الموزونة بأوزان حصينة وذلك لوجود قيم شاذة في البيانات بالإضافة لمشكلة التعدد الخطي، وقد تم استخدام برنامج R في كتابة برامج التحليل.

UNIVERSITY OF MOSUL

COLLEGE OF COMPUTER SCIENCES AND MATHEMATICS

Robust Methods for Principal Components Regression: A Comparative Study with Simulation and Application

Israa Najeeb Said Mahmoud AL-Sarraf

M.Sc. / Thesis
Statistics

Supervised by

Assist. Prof.

Dr. Bashar Abdul Aziz Majeed Al-Talib

2021A.D. **1443 A.H.**

Abstract

The idea of the thesis is based on an attempt to suggest some robust methods in principal component regression to find estimators of nonlinear regression and to choose the estimators that are most robust to the presence of outliers. The principal component regression is considered as one of the estimation methods to deal with Multi-collinearity that exist in most of a lot of the standard models, This core of this method is to replace the independent variables that suffer from the problem of Multi-collinearity with the principal components that they express to fit the regression model in the case of Multi-collinearity, and then use the weight functions of robust M and MM to weigh the objective function (Huber, Hampel, and Bisquare) to deal with The presence of outliers in the data, In order to verify the efficiency of the estimates, an experimental study was conducted through simulation with different sample sizes and contamination percentages, and the methods were also applied to real data collected from the files of Badoush Cement Factory in Nineveh Governorate for the period from (2008-2014), It has been tested whether the data suffers from the problem of Multi-collinearity and then applying the least squares by using the principal components as independent variables and estimating the model, and it was found that the variables suffer from the problem of Multi-collinearity, and the treatment was done by applying weighted principal components regression with robust weights due to the existence of outliers in the data in addition to the problem of Co-linearity, and R program was used in writing analysis software..