Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.5, No.9, 2015

www.iiste.org

A Geometric Interpretation and Comparison of the Methods Of Ordinary Least Square (OLS) and Bivariate Lagrange Interpolation

Seyma Caliskan Cavdar[1*], Alev Dilek Aydin[2], Mehmet Fatih Karaaslan[3]

[1] Faculty of Business, Department of International Trade and Business Administration, Halic University, Istanbul, Turkey

[2] Faculty of Business, Department of International Trade and Business Administration, Halic University, Istanbul, Turkey

[3] Faculty of Arts and Sciences, Department of Statistics, Yildiz Technical University, Istanbul, Turkey

Abstract

In this study, the consumer prices, real gross domestic product (GDP) and unemployment for Germany and Turkey between 2006 and 2011 are geometrically interpreted by using Lagrange interpolation and OLS method. The coefficients of linear regression models are obtained by matrix display of OLS method. The Lagrange interpolation polynomial is considered in bivariate situation and we aim a different formulation. Thanks to the considered methods, it will be possible to have an idea about the unemployment rate status in Germany and Turkey. We use two different methods for the prediction of unemployment rates, which are developed by different equations in order to predict third variable by using two variables. Our main purpose is to determine whether an equation gives the correct guess rather than numerical expression. Besides, we have tried to state geometric display to data.

Keywords: Lagrange interpolation, ordinary least square (OLS), regression, geometrical display, matrix display

1. Introduction

In the mathematical field of numerical analysis, interpolation is a method of constructing new data points within the range of a discrete set of known data points. We represent a set of points with a curve that passes exactly through all of the points. Interpolation can also be implemented to functions first by sampling the function at some points. In this way, we can obtain a suitable representation for the data or a function that may be sufficiently precise for practical purposes. It is often required to interpolate the value of that function for a current value of the independent variable. This may be achieved by curve fitting or regression analysis. In fact, the geometry of the set of interpolation points (called nodes) is crucial for determining the solvability of the interpolation problem.

The basis type of interpolation is interpolation by means of univariate polynomials. Multiple solutions for polynomial interpolation have been given, particularly those of Newton and Lagrange. In numerical analysis, multivariate interpolation or spatial interpolation is interpolation on functions of more than one variable. The function to be interpolated is known at given points $(x_i, y_i, z_i, ...)$ and the interpolation problem consists of yielding values at arbitrary points $(x, y, z, ...)$. Interpolation is the process of defining a function that determines the given values at given points. Multivariate interpolation is concerned with interpolation of a function of more than one variable.

The regression analysis and Ordinary Least Square (OLS) and their applications in the non-linear regression analysis is one of the major subjects for the *researchers* in the fields of economics and finance (Arhipova I. and Balina S., 2004). OLS regression is based on determining the dependent variable (Y) by producing unbiased minimum sum of error square in (Y) with regards to the independent variable (X). As Fernandes and Leblanc (2005) mentioned, to obtain unbiased observations, the prediction should be equal to the expected value of the dependent variable for a particular set of data. OLS is based on a set of assumptions such as normality, homogeneity and independence of residuals. As a result of violation of these assumptions, inefficient and biased estimators can be obtained which may lead to an inaccurate estimation of model parameters by OLS (Montgomery et al., 2001).

Unemployment is not only a risk for individual countries but it is also a global problem for almost all of the nations. Current unemployment statistics reveal that unemployment is a very serious problem especially for young population in most developing countries. The most important reason of this situation is that young people have better education levels as compared to old people, however, they do not have the opportunity to use their skills in the current economic situation.

Our study is based on two different countries, in which policy makers respond to reduce high level of unemployment by establishing different policies, namely Germany and Turkey. Germany's economy and the labor market has performed quite well in recent years as compared to Turkey although further improvements for young population, long-term unemployed and women still needs to be provided. The German labor market has also responded very well to the global economic crisis of 2008 whereas the same thing cannot be said for the Turkish labor market. As it was stated in the Employment Plan of Germany (2014), once the economy has started to recover German companies were ready to react quickly increasing the German market share of the global market and thus employing more people. Hence, the major purpose of this paper is to compare a geometric approach that can serve to make further plain on multivariate interpolation and OLS method. This geometric approach is made possible by determining the standard OLS test statistic in terms of geometric display. In this study, we employed observations of Turkish and German unemployment rate covering the period of 2006-2011 on annual basis. We obtained the whole data from the International Financial Statistics (IFS) data base. In this study, we used two different methods in order to compare the geometric perspectives of OLS and Lagrange interpolation. Our main aim of using these methods is to develop a new equation by predicting third variable by using two other variables

This study contributes to the existing literature in two ways. First, it tries to predict the unemployment rates of two different countries with completely different dynamics. In this context, forecast of unemployment rate particularly in a developing country will enrich the existing literature in which few numbers of studies concentrate on the unemployment rates in developing countries. Secondly, it is one of the rare studies in which OLS and bivariate Lagrange interpolation has been used in a single study. In this regard, it will provide a new perspective to the researchers who intend to use different statistical methods and approaches in their researches.

The paper is organised as follows: Section 2 provides the major literature on the comparison of forecasting performances of different linear and non-linear time series models to predict unemployment rates of different countries. Section 3 provides data and methodology. Section 4 presents empirical findings. Section 5 concludes our paper and summarises our findings.

2. Literature Review

Regression is one of the most widely used statistical methods for investigating and modelling the relationships between different macroeconomic variables (Montgomery et al., 2001) and has a wide application in almost every field of studies including economics and finance. Various regression approaches and a number of different research papers have been developed to forecast a variety of different macroeconomic variables and to establish the relationship between these variables. Prediction of unemployment rate constitutes the basis of many economic decisions. This is mainly because unemployment, which is an economic concept, is directly related with the economic situation in a country. As Kwok (2011) states, unemployment has a direct negative influence on the economic development. Different reasons and factors may be affective on the unemployment in different countries.

Various different statistical methods from the OLS to GARCH models have been used to forecast unemployments rates of different countries. Rothman (1998) tries to measure the out-of-sample forecasting accuracy of six non-linear models. On the other hand, Parker and Rothman (1998) tries to develop a model to forecast quarterly adjusted unemployment rate with AR(2) model. Montgomery et al. (1998) conducted a study based on a comparison of forecasting performances of different linear and non-linear time series models by using the U.S. unemployment rate. They particularly emphasized the measuring of forecasting performances during economic expansions and contractions by analyzing the asymmetric cyclical behaviors of unemployment rate and by using additional monthly information to forecast the quarterly rate. The results indicate that significant improvements can be acquired in forecasting accuracy over existing methods. Koop and Potter (1999) are the ones, who used threshold autoregressive (TAR) in order to model and predict the US monthly unemployment rate. Proietti (2001) investigated the out-of-sample forecasting for the US monthy unemployment rate by employing seven linear and non-linear forecasting models. The study concludes that linear models perfom better with a higher persistence.

Several other researches for modelling and predicting the unemployment rates include Johnes (1999), Peel and Speight (2000), Gil-Alana (2001), Brown and Moshiri (2004), Milas and Rothman (2008) and Barnichon and Nekarda (2012). Johnes (1999) compares AR(4), AR(4)-GARCH(1,1), SETAR(3,4,4,4) and neural network to forecast UK monthly unemployment rate by using the sample period of January 1960 and August (1996). He found that SETAR model performs better for short-term forecasts. Peel and Speight (2000) examined whether non-linear time series models of SETAR can obtain superior out-of-sample forecasts of UK unemployment data for February 1971 and September 1991. The study indicates evidence that out-of-sample SETAR has a better forecasting performance as compared to relative AR models. Gil-Alana (2001) utilized a Bloomfield exponential spectral model for developing a model to predict the unemployment rate as an alternative to ARMA models. Brown and Moshiri (2004) developed two artificial neural network (ANN) models to predict unemployment rates in the US, Canada, UK, France and Japan. They compared the out-of-sample forecast results that they obtained by the ANN models with those of several linear and non-linear times series models. They found that ANN models have more accurate prediction perfomances than other time series models. Milas and Rothman (2008) developed multivariate non-linear models to evaluate the US unemployment rate. Barnichon and Nekarda (2012) combined theory of equilibrium unemployment with simple econometric techniques to provide an extremely explanatory analysis of unemployment fluctuations.

3.Data and Methodology

3.1.Polynomial Interpolation

Our empirical analysis is based on the data for one developed and one developing country, namely Germany and Turkey in 2006 and 2011. We obtained the data required for Germany and Turkey from the IFS database.

Let us start with $n+1$ *(MathType)* data points $\{(x_1 y_1),(x_2,y_2),\ldots,(x_{n+1},y_{n+1})\}$ in $\mathbb{R}^2$. For $i=1,2,\ldots,n+1$, consider the Lagrange Basis Polynomials $L_i(x)$ given by the formula

$$L_i(x)=\prod_{j=1,\, j\neq i}^{n+1}\frac{x-x_j}{x_i-x_j}$$

where $L_i(x)$ has the property that

$$L_i(x_j = \begin{cases} 0, & j\neq i \\ 1, & j=i \end{cases}$$

The polynomial $L_i(x)$ is the Lagrange polynomial for the interpolation points

$$\{(x_1,0),(x_2,0),\ldots,(x_{i-1},0)\},(x_i,1),(x_{i+1},0),\ldots,(x_{n+1},0)\}$$

The Lagrange polynomial $L(x)$ for the original interpolation points are now given by the following formula

$$L(x)=\sum_{i=1}^{n+1} y_i . L_i(x)$$

It is clear that this polynomial has the property that $L(x_i)=y_i$ as required.

Note that the Lagrange polynomial, $L(x)$, is unique. If there were two such polynomials, $L(x)$ and $P(x)$, then $L(x)-P(x)$ would be a polynomial of degree $\leq n$ with $n+1$ zeros. Thus, it would have to be identically zero and we must have $L(x)\equiv P(x)$.

The advantage of this form of the interpolation polynomial is that we can write down the interpolant immediately, without computing the coefficients in the function.

Now we want to investigate bivariate form of the Lagrange interpolation polynomial. For this purpose, we will deal with multivariate case and then introduce its formula (Berrut, 2004; Gasca and Sauer, 2000).

3.2. Multivariate Case

Multivariate Lagrange interpolation is an important problem of computational mathematics and approximation theory. The research of the theory and method of multivariate interpolation is developing rapidly in the past few decades to find out some problems such as the calculation of multivariate function, the design of surface, the construction of finite element scheme, computer graphics, numerical quadrature, cubature, numerical solutions to differential equations (Gasca and Sauer, 2000).

Let $f = f(X_1, X_2, ..., X_k)$ be an $k-$variable multinomial function of degree n. Since there are $\binom{n+k}{n} = p$ terms in f, it is a necessary condition that we have p distinct points $(x_{1,i}, x_{2,i}, \ldots, x_{k,i}, f_i) \in \mathbb{R}^{k+1},\ 1 \le i \le p,\ f_i = f(x_{1,i}, \ldots, x_{k,i})$ for f to be uniquely defined. However,

$$f(X_1, ..., X_k) = \sum_{e_i.1 \le n} \alpha_{e_i} X^{e_i}$$

where the α_{e_i} are the coefficients in f, $X = (X_1, \ldots, X_k)$ is the $k-$tuple of independent variables of f, $e_i = (e_{1,i}, e_{2,i}, \ldots, e_{k,i})$ is an exponent vector with nonnegative integer entries consists of an ordered partition of an integer between 0 and n inclusive, $e_i.1 = \sum_{j=1}^{k} e_{ji}$ is the dot product and $X^{e_i} = \prod_{j=1}^{k} X_j^{e_{ji}}$. From Lagrange, we write f in the form $\sum_{i=1}^{p} f_i L_i(X)$, where $L_i(X)$ is a multinomial function in the independent variables $X_1, \ldots, X_k$ with the property that when X is equal to the i^{th} data value, or $X = x_i((X_1, \ldots, X_k) = (x_{1,i}, \ldots, x_{k,i}))$, then $L_i(x_i) = 1$ and $L_j(x_i) = 0\ (j \ne i)$. To do this, consider the system of linear equations $f_i = \sum_{e_j.1 \le n} \alpha_{ej} x_i^{e_j}$, where $1 \le i \le p$. From this system construct the sample matrix $A = [x_i^{e_j}]$;

$$A = \begin{bmatrix} x_1^{e_1} & . & . & . & . & . & x_1^{e_p} \\ . & & & & & & . \\ . & & & & & & . \\ x_i^{e_1} & . & . & . & . & . & x_i^{e_p} \\ . & & & & & & . \\ . & & & & & & . \\ x_p^{e_1} & . & . & . & . & . & x_p^{e_p} \end{bmatrix}$$

We assume here that $\det(A) \ne 0$. Although we can solve for the coefficients α_i in f by inverting A, the utility of Lagrange interpolation can be solved by f without clearly obtaining for its coefficients (Carnicer and Gasca, 2005).

Remark. If A is singular, then the coefficients of f are not adjusted, in which case f is clearly not unique. Therefore, f is unique if and only if its sample matrix is nonsingular. On the other hand, by sembolising the geometric demonstration of the p points, $\det(A) = 0$ occurs to be a complicated problem (Olver, 2006: 201-240).

Let $\Delta = \det(A)$. Now make the substitutions $x_j = X$ in A; this gives the following matrix $A_j(x)$:

$$A_j(X) = \begin{bmatrix} x_1^{e_1} & . & . & . & . & . & x_1^{e_p} \\ . & & & & & & . \\ . & & & & & & . \\ X^{e_1} & . & . & . & . & . & X^{e_p} \\ . & & & & & & . \\ . & & & & & & . \\ x_p^{e_1} & . & . & . & . & . & x_p^{e_p} \end{bmatrix} \leftarrow j^{\text{th}} \text{ row}$$

Let $\Delta_j(X) = \det(A_j(X))$. Next, make the substitutions $X = x_i$ in $A_j(X)$ $(i \neq j)$; this gives the following matrix $(A_j)_i$:

$$(A_j)_i = \begin{bmatrix} x_1^{e_1} & . & . & . & . & x_1^{e_p} \\ . & . & . & . & . & \\ . & . & . & . & . & \\ x_i^{e_1} & . & . & . & . & x_i^{e_p} \\ . & . & . & . & . & \\ . & . & . & . & . & \\ x_i^{e_1} & . & . & . & . & x_i^{e_p} \\ . & . & . & . & . & \\ . & . & . & . & . & \\ x_p^{e_1} & . & . & . & . & x_p^{e_p} \end{bmatrix}$$

Note that the i^{th} row appears twice in $(A_j)_i$. That means $\det((A_j)_i) = 0$. In other words, when $X = x_i$ then $\Delta_j(x_i) = 0\,(i \neq j)$. By construction, morever, $X = x_i \Rightarrow \Delta_i(X) = \Delta$. Hence,

$$L_i(X) = \frac{\Delta_i(X)}{\Delta}$$

and therefore

$$f = \sum_{i=1}^{p} f_i \frac{\Delta_i(X)}{\Delta}.$$

3.3. OLS (Ordinary Least Squares)

Regression analysis is one of the most important approaches in econometrics. For this analysis one can first specify a regression model that determines the relationship between economic variables. The most commonly used specification is the linear model. Regression analysis is a statistical tool for predicting the associations among variables and it involves several techniques for analyzing a number of different variables and modeling.

In statistics, OLS is a technique for calculating the unknown parameters in a linear regression model. This method minimizes the sum of squared vertical distances between the observed responses in the dataset and the responses predicted by the linear approximation. The obtained estimator can be expressed by a simple formula, especially in the case of a single regressor on the right-hand side. The OLS estimator is consistent when the regressors are exogenous and when there is no perfect multicollinearity and the error terms of linear unbiased estimators are homoscedastic and serially uncorrelated. Under these conditions, OLS method provides minimum variance and mean-unbiased estimation when the errors have finite variances. Under the assumption that the errors are normally distributed, OLS is the maximum likelihood estimator. It is used in economics, econometrics, finance, electrical engineering, and political science in addition to many other areas of applications. In statistics and mathematics, linear least squares is an approach which determines whether a mathematical or statistical model complies with the data in cases where the value obtained by the model for any data point is stated linearly with regard to the unknown parameters of the model.

The OLS estimator is rational when the regressors are exogenous and when there is no multicollinearity, and optimal in the class of linear unbiased estimators when the errors are homoscedastic and serially uncorrelated (Amemiya, 1985).

In a linear regression model, the response variable is a linear function of the regressors;

$$y_i = x_i' \beta + \varepsilon_i$$

where $\beta = (\beta_0, \beta_1, ..., \beta_p)$ is a vector of $p+1$ parameters which describes the relationship between some variables $x = (1, x_1, x_2, ..., x_p)$ and ε_i is unobserved random variables (Scott and Holt,1982).

This model can also be written in matrix notation as,

$$y = X\beta + \varepsilon$$

where y and ε are $nx1$ vectors, X is an nxp matrix of regressors (Hayashi, 2000).

This statistical model can be obtained similar to the following;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{nx1} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{k1} \\ 1 & X_{12} & \cdots & X_{k2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & X_{1n} & \cdots & X_{kn} \end{bmatrix}_{nxk} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}_{kx1} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{nx1}$$

This is not the only method for estimating these regression models. However, we can state three good reasons for using OLS;

- The model fit derives from an orthogonal projection onto the model space, which has geometric properties,
- The Gauss-Markov theorem states that $\hat{\beta}$ is the best linear unbiased estimator (BLUE),
- If the errors are independent and normally distributed, the least squares estimate is the maximum likelihood estimate (Plackett, 1950).

OLS is an approximate solution to an overdetermined model of linear equations $X\beta \approx y$, where β is the unknown. If the system cannot be solved precisely (the number of equations n is much larger than the number of unknowns p), we need to look for a solution that could derive the smallest discrepancy between the right-and-left-hand sides (Hayashi, 2000). In other words, we look for the solution,

$$\hat{\beta} = \arg\min_{\beta} \|y - X\beta\|$$

where $\| \|$ is the standard L^2 norm in the n-dimensional Euclidean space R^n. The euclidean space R^n is the vector space composed by all vectors in R^n. The Euclidean norm is also called the Euclidean length, L^2 distance, ℓ^2 distance, L^2 norm, or ℓ^2 norm. The problem: $\min_\beta \|y - X\beta\| \Leftrightarrow \min_\beta \|y - X\beta\|^2$.

$$\hat{Y} = X^T \hat{\beta},$$
$$e = Y - \hat{Y}$$

where e is orthogonal to any point in $S(X)$ to X or $X\hat{\beta}$.

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

where from the orthogonality position $X^T (Y - \hat{\beta}) = 0$ (Sosa-Escudero, 2009; Bourbaki, 1987).

The predicted $X\beta$ is a certain linear combination of the vectors of regressors. Thus, the residual vector $y - X\beta$ will have the smallest length when y is projected orthogonally onto the linear subspace spanned by the columns of X. The OLS estimator in this case can be obtained as the coefficients of vector decomposition of $\hat{y} = Py$ along the basis of X (Hayashi, 2000).

We can show the regression equation geometrically like the diagram in Figure 1. This diagram has been depicted by the error vector $\hat{u}$ which is not orthogonal to the $X\hat{\beta}$ vector (Davidson et al., 2004).

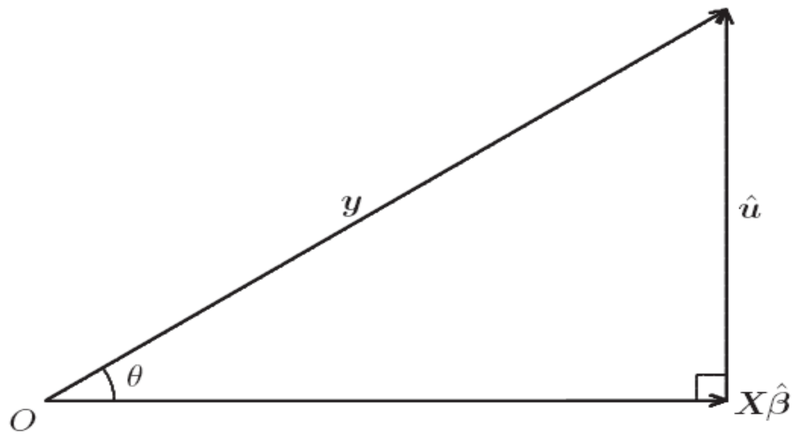

Figure1. Geometry of OLS

4. Empirical Findings

In the section of this study, we try to estimate unemployment rates of Germany and Turkey by using consumer price and real gross domestic product (GDP) data and their geometric displays. We have done the whole calculations by using the MATLAB programme.

Example:

Country	Year	Real GDP	Consumer Price	Unemployement
GERMANY	2006	2.9	1.8	8.1
GERMANY	2007	2.4	2.1	6.5
GERMANY	2008	2	1.8	6.3
GERMANY	2009	-4.7	0.2	7.5
GERMANY	2010	3.3	1.3	7.1
GERMANY	2011	2.0	1.4	7.1
TURKEY	2006	6.1	9.6	10.2
TURKEY	2007	5.0	8.2	10.3
TURKEY	2008	5.3	4.6	11.0
TURKEY	2009	-4.7	6.3	14.0
TURKEY	2010	7.8	8.7	11.0
TURKEY	2011	3.6	5.7	10.7

How can we show the geometric interpretation of the data about Turkey and Germany? What can we say for the values between variables?

Solution:

1) Firstly, we will solve the problem with Ordinary Least Squares.

For Germany, we derive

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$
$$\beta = (X^T X)^{-1}(X^T Y)$$

$$Y = \begin{bmatrix} 8.1 \\ 6.5 \\ 6.3 \\ 7.5 \\ 7.1 \\ 7.1 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 2.9 & 1.8 \\ 1 & 2.4 & 2.1 \\ 1 & 2 & 1.8 \\ 1 & -4.7 & 0.2 \\ 1 & 3.3 & 1.3 \\ 1 & 2 & 1.4 \end{bmatrix}, \quad X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 2.9 & 2.4 & 2 & -4.7 & 3.3 & 2 \\ 1.8 & 2.1 & 1.8 & 0.2 & 1.3 & 1.4 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 6 & 7.9 & 8.6 \\ 7.9 & 55.15 & 20.01 \\ 8.6 & 20.01 & 14.58 \end{bmatrix}, \quad (X^T X)^{-1} = \begin{bmatrix} 2.6515 & 0.373758 & -2.07694 \\ 0.373758 & 0.0888022 & -0.342335 \\ -2.07694 & -0.342335 & 1.7635 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 42.6 \\ 54.07 \\ 60.24 \end{bmatrix}, \quad \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2.6515 & 0.373758 & -2.07694 \\ 0.373758 & 0.0888022 & -0.342335 \\ -2.07694 & -0.342335 & 1.7635 \end{bmatrix} \begin{bmatrix} 42.6 \\ 54.07 \\ 60.24 \end{bmatrix}$$

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}=\begin{bmatrix}8.0481\\ 0.101334\\ -0.754552\end{bmatrix},\quad \beta_0=8.0481,\quad \beta_1=0.101334,\quad \beta_2=-0.754552$$

$$Y_{Germany}=8.0481+0.101334X_1-0.754552X_2$$

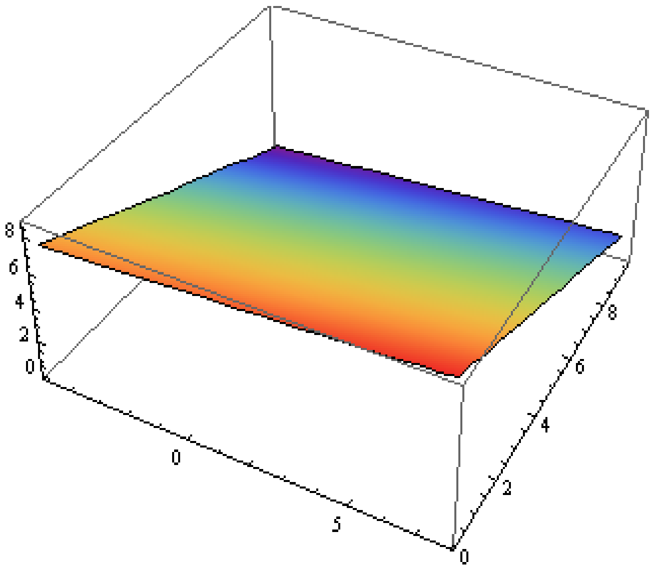

Figure 2. Graphics of Germany for OLS

and for Turkey we have

$$Y=\begin{bmatrix}10.2\\ 10.3\\ 11\\ 14\\ 11\\ 10.7\end{bmatrix},\quad X=\begin{bmatrix}1 & 6.1 & 9.6\\ 1 & 5 & 8.2\\ 1 & 5.3 & 4.6\\ 1 & -4.7 & 6.3\\ 1 & 7.8 & 8.7\\ 1 & 3.6 & 5.7\end{bmatrix},\quad X^T=\begin{bmatrix}1 & 1 & 1 & 1 & 1 & 1\\ 6.1 & 5 & 5.3 & -4.7 & 7.8 & 3.6\\ 9.6 & 8.2 & 4.6 & 6.3 & 8.7 & 5.7\end{bmatrix}$$

$$X^TX=\begin{bmatrix}6 & 23.1 & 43.1\\ 23.1 & 186.19 & 182.71\\ 43.1 & 182.71 & 328.43\end{bmatrix},\quad (X^TX)^{-1}=\begin{bmatrix}2.98623 & 0.0309801 & -0.409119\\ 0.0309801 & 0.0121493 & -0.0108243\\ -0.409119 & -0.0108243 & 0.0627553\end{bmatrix}$$

$$X^TY=\begin{bmatrix}67.2\\ 23.54\\ 477.87\end{bmatrix},\quad \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}=\begin{bmatrix}2.98623 & 0.0309801 & -0.409119\\ 0.0309801 & 0.0121493 & -0.0108243\\ -0.409119 & -0.0108243 & 0.0627553\end{bmatrix}\begin{bmatrix}67.2\\ 23.54\\ 477.87\end{bmatrix}$$

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}=\begin{bmatrix}12.3112\\ -0.289869\\ 0.000666398\end{bmatrix},\quad \beta_0=12.3112,\quad \beta_1=-0.289869,\quad \beta_2=0.000666398$$

$$Y_{Turkey} = 12.3112 - 0.289869 X_1 + 0.000666398 X_2$$

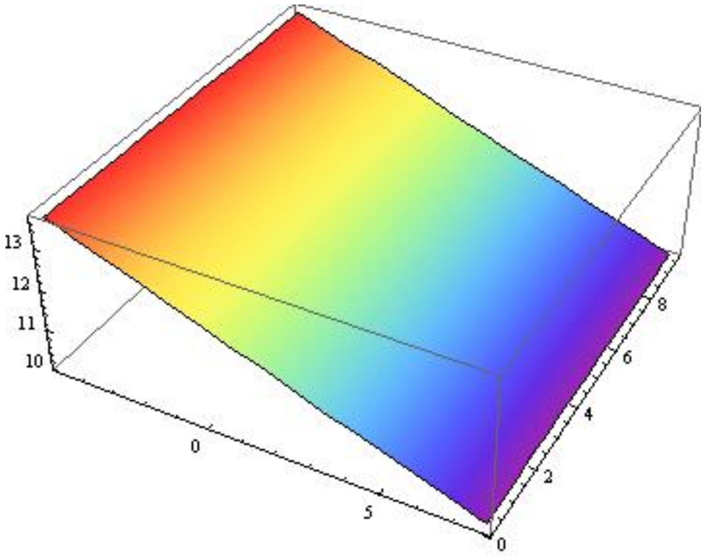

Figure 3. Graphics of Turkey for OLS

2) Now, we can calculate multivariate case of Lagrange interpolation. For Germany, we obtain as follows;

$$z_i = c_1 x_i^{\,2} + c_2 x_i y_i + c_3 y_i^{\,2} + c_4 x_i + c_5 y_i + c_6$$

$$\begin{aligned}
8.1 &= c_1(2.9)^2 + c_2(2.9)(1.8) + c_3(1.8)^2 + c_4(2.9) + c_5(1.8) + c_6 \\
6.5 &= c_1(2.4)^2 + c_2(2.4)(2.1) + c_3(2.1)^2 + c_4(2.4) + c_5(2.1) + c_6 \\
6.3 &= c_1(2)^2 + c_2(2)(1.8) + c_3(1.8)^2 + c_4(2) + c_5(1.8) + c_6 \\
7.5 &= c_1(-4.7)^2 + c_2(-4.7)(0.2) + c_3(0.2)^2 + c_4(-4.7) + c_5(0.2) + c_6 \\
7.1 &= c_1(3.3)^2 + c_2(3.3)(1.3) + c_3(1.3)^2 + c_4(3.3) + c_5(1.3) + c_6 \\
7.1 &= c_1(2)^2 + c_2(2)(1.4) + c_3(1.4)^2 + c_4(2) + c_5(1.4) + c_6
\end{aligned}$$

$$\begin{aligned}
8.1 &= 8.41c_1 + 5.22c_2 + 3.24c_3 + 2.9c_4 + 1.8c_5 + c_6 \\
6.5 &= 5.76c_1 + 5.04c_2 + 4.41c_3 + 2.4c_4 + 2.1c_5 + c_6 \\
6.3 &= 4c_1 + 3.6c_2 + 3.24c_3 + 2c_4 + 1.8c_5 + c_6 \\
7.5 &= 22.09c_1 - 0.94c_2 + 0.04c_3 - 4.7c_4 + 0.2c_5 + c_6 \\
7.1 &= 10.89c_1 + 4.29c_2 + 1.69c_3 + 3.3c_4 + 1.3c_5 + c_6 \\
7.1 &= 4c_1 + 2.8c_2 + 1.96c_3 + 2c_4 + 1.4c_5 + c_6
\end{aligned}$$

and the coefficient matrix is

$$A=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix}$$

$$A_1=\begin{pmatrix} x^2 & xy & y^2 & x & y & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix},\ A_2=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix}$$

$$A_3=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix},\ A_4=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix}$$

$$A_5=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 4 & 2.8 & 1.96 & 2 & 1.4 & 1 \end{pmatrix},\ A_6=\begin{pmatrix} 8.41 & 5.22 & 3.24 & 2.9 & 1.8 & 1 \\ 5.76 & 5.04 & 4.41 & 2.4 & 2.1 & 1 \\ 4 & 3.6 & 3.24 & 2 & 1.8 & 1 \\ 22.09 & -0.94 & 0.04 & -4.7 & 0.2 & 1 \\ 10.89 & 4.29 & 1.69 & 3.3 & 1.3 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{pmatrix}.$$

From here

$$\Delta=3.11894$$
$$\Delta_1=-0.383476x^2+5.83198xy-10.2987y^2-5.15305x+21.2919y-14.1127$$
$$\Delta_2=-0.64224x^2+0.557136xy+13.9221y^2+2.14413x-45.6649y+33.3643$$
$$\Delta_3=1.84317x^2-7.67518xy-13.249y^2+1.31832x+65.5445y-54.3131$$
$$\Delta_4=0.0513x^2+0.042912xy+0.024336y^2-0.328612x-0.163699y+0.51335$$
$$\Delta_5=0.727488x^2-3.98779xy+2.97158y^2+3.61333x-1.53348y-2.64823$$
$$\Delta_6=-1.59624x^2+5.23094xy+6.62972y^2-1.59412x-39.4743y+40.3153$$

and if we use our formula, we get

$$z=8.1\frac{\Delta_1}{\Delta}+6.5\frac{\Delta_2}{\Delta}+6.3\frac{\Delta_3}{\Delta}+7.5\frac{\Delta_4}{\Delta}+7.1\frac{\Delta_5}{\Delta}+7.1\frac{\Delta_6}{\Delta}$$
$$z_{\text{Germany}}=-0.465586x^2+3.73684xy-2.57869y^2-2.44495x-1.2216y+10.1537$$

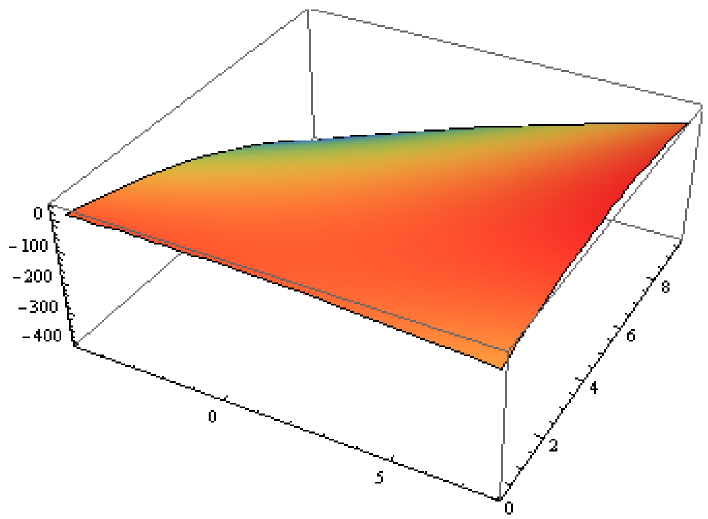

Figure 4. Graphics of Germany for Polynomial Interpolation

and for Turkey, we derive

$$z_i = c_1 x_i^{\ 2} + c_2 x_i y_i + c_3 y_i^{\ 2} + c_4 x_i + c_5 y_i + c_6$$

$$\begin{aligned}
10.2 &= c_1(6.1)^2 + c_2(6.1)(9.6) + c_3(9.6)^2 + c_4(6.1) + c_5(9.6) + c_6 \\
10.3 &= c_1(5)^2 + c_2(5)(8.2) + c_3(8.2)^2 + c_4(5) + c_5(8.2) + c_6 \\
11 &= c_1(5.3)^2 + c_2(5.3)(4.6) + c_3(4.6)^2 + c_4(5.3) + c_5(4.6) + c_6 \\
14 &= c_1(-4.7)^2 + c_2(-4.7)(6.3) + c_3(6.3)^2 + c_4(-4.7) + c_5(6.3) + c_6 \\
11 &= c_1(7.8)^2 + c_2(7.8)(8.7) + c_3(8.7)^2 + c_4(7.8) + c_5(8.7) + c_6 \\
10.7 &= c_1(3.6)^2 + c_2(3.6)(5.7) + c_3(5.7)^2 + c_4(3.6) + c_5(5.7) + c_6
\end{aligned}$$

$$\begin{aligned}
10.2 &= c_1(6.1)^2 + c_2(6.1)(9.6) + c_3(9.6)^2 + c_4(6.1) + c_5(9.6) + c_6 \\
10.3 &= c_1(5)^2 + c_2(5)(8.2) + c_3(8.2)^2 + c_4(5) + c_5(8.2) + c_6 \\
11 &= c_1(5.3)^2 + c_2(5.3)(4.6) + c_3(4.6)^2 + c_4(5.3) + c_5(4.6) + c_6 \\
14 &= c_1(-4.7)^2 + c_2(-4.7)(6.3) + c_3(6.3)^2 + c_4(-4.7) + c_5(6.3) + c_6 \\
11 &= c_1(7.8)^2 + c_2(7.8)(8.7) + c_3(8.7)^2 + c_4(7.8) + c_5(8.7) + c_6 \\
10.7 &= c_1(3.6)^2 + c_2(3.6)(5.7) + c_3(5.7)^2 + c_4(3.6) + c_5(5.7) + c_6
\end{aligned}$$

so that we find

$$\begin{aligned}
10.2 &= 37.21c_1 + 58.56c_2 + 92.16c_3 + 6.1c_4 + 9.6c_5 + c_6 \\
10.3 &= 25c_1 + 41c_2 + 67.24c_3 + 5c_4 + 8.2c_5 + c_6 \\
11 &= 28.09c_1 + 24.38c_2 + 21.16c_3 + 5.3c_4 + 4.6c_5 + c_6 \\
14 &= 22.09c_1 - 29.61c_2 + 39.69c_3 - 4.7c_4 + 6.3c_5 + c_6 \\
11 &= 60.84c_1 + 67.86c_2 + 75.69c_3 + 7.8c_4 + 8.7c_5 + c_6 \\
10.7 &= 12.96c_1 + 20.52c_2 + 32.49c_3 + 3.6c_4 + 5.7c_5 + c_6
\end{aligned}$$

and the coefficient matrix is

$$B = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} x^2 & xy & y^2 & x & y & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}$$

$$B_3 = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}$$

$$B_5 = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ x^2 & xy & y^2 & x & y & 1 \\ 12.96 & 20.52 & 32.49 & 3.6 & 5.7 & 1 \end{pmatrix}, \quad B_6 = \begin{pmatrix} 37.21 & 58.56 & 92.16 & 6.1 & 9.6 & 1 \\ 25 & 41 & 67.24 & 5 & 8.2 & 1 \\ 28.09 & 24.38 & 21.16 & 5.3 & 4.6 & 1 \\ 22.09 & -29.61 & 39.69 & -4.7 & 6.3 & 1 \\ 60.84 & 67.86 & 75.69 & 7.8 & 8.7 & 1 \\ x^2 & xy & y^2 & x & y & 1 \end{pmatrix}$$

and we derive from that

$$\begin{aligned}
D &= -30909.7 \\
D_1 &= 576.391x^2 - 2224.3xy - 4636.93y^2 + 15206.1x + 71383.4y - 272801. \\
D_2 &= -470.686x^2 + 7697.43xy + 3652.23y^2 - 5074.4x - 95503.6y + 456537. \\
D_3 &= -681.218x^2 + 3980.36xy - 2278.94y^2 - 2556.2x + 16665.5y - 1776.2 \\
D_4 &= -320.003x^2 + 362.768xy - 145.308y^2 + 1054.85x - 1.60208y - 2364.06
\end{aligned}$$

$$D_5 = -542.849x^2 - 213602xy + 2177.94y^2 + 129351x - 174043y + 327436$$
$$D_6 = 143837x^2 - 768024xy + 123101y^2 + 47110x + 248606y - 243249$$
$$z = 10.2\frac{D_1}{D} + 10.3\frac{D_2}{D} + 11\frac{D_3}{D} + 14\frac{D_4}{D} + 11\frac{D_5}{D} + 10.7\frac{D_6}{D}$$
$$z_{\text{Turkey}} = 0.0492774x^2 + 0.00700056xy - 0.0112508y^2 - 0.400765x - 0.0738538y + 12.147.$$

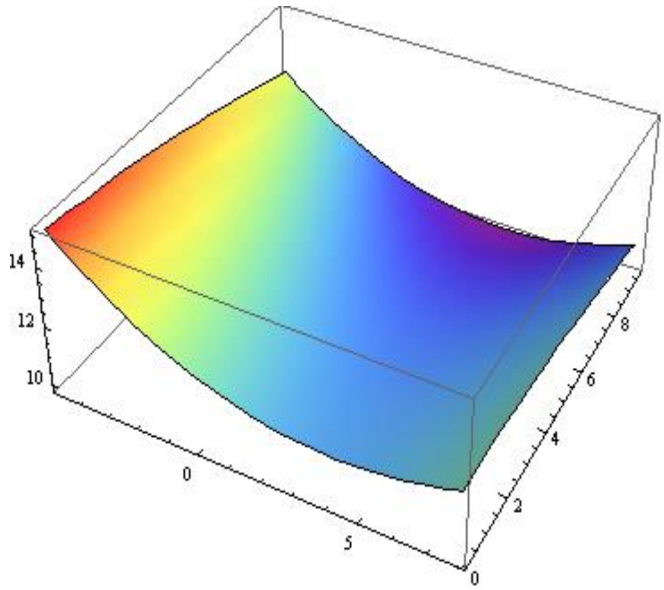

Figure 5. Graphics of Turkey for Polynomial Interpolation

5. Conclusion

Macroeconomic modelling and forecasting the unemployment rate is one of the most widely researched areas in the applied economic literature. Prediction of unemployment rate is one of the most crucial applications for economists and policymakers as it may function as an important tool in development of various different macroeconomic policies. It is one of the most serious economic problems that face different nations. The increase in unemployment rates reduce personal income and increase in the inequality and poverty. Prediction of unemployment rate may undoubtedly help policy makers to take preventive measures on time as well as set the necessary policies to avoid high unemployment rates before it becomes a rigid problem for the nation.

In this study, the consumer prices for Germany and Turkey between 2006 and 2011, real GDP and unemployment data were geometrically interpreted by using Lagrange interpolation and OLS method. These two different methods were used to estimate equations for each country and we obtained four different equations and geometric displays. While estimations from OLS method indicate planar graph, the predictions derived from lagrange interpolation are stated as quadratic equation. When our study is evaluated generally, it can be said that OLS results have linear features.

In our study, which compares OLS and lagrange interpolation methods, we found that both methods have obtained similar results to each other in some ways. The equation obtained by the OLS method for Germany has been calculated as follows;

$$Y_{Germany} = 8.0481 + 0.101334X_1 - 0.754552X_2$$

where the coefficient of GDP is positive (0.101334), and the coefficient of consumer price is negative (-0.754552). This result can be evaluated as a proper result in the economic literature. That is, the more GDP increases, the more unemployment rate will increase in proportion of 0.101334. The more consumer price increases, the more unemployment will decrease at the rate of 0.754552.

In the same way, the results of OLS methods for Turkey have been calculated as follows;

$$Y_{Turkey} = 12.3112 - 0.289869X_1 + 0.000666398X_2$$

where the coefficient of GDP is negative (-0.289869) and the coefficient of consumer price is positive (0.000666398). This result is exactly opposite to the results that we obtained for Germany. In other words, the more GDP increases, the more unemployment rate will decrease at the rate of 0.29 on average. The economic literature points out that for developing countries like Turkey, the rise of consumer price may increase the growth of unemployment rate. On the other hand, the rise of GDP can result in a decrease in the rate of unemployment.

The equation obtained by *Lagrange* interpolation method for Germany has been calculated as follows;

$$\mathrm{z}_{\mathrm{Germany}} = -0.465586\mathrm{x}^2 + 3.73684xy - 2.57869y^2 - 2.44495x - 1.2216y + 10.1537$$

where the coefficient of GDP is negative (-2.44495) and the coefficient of consumer price is negative (-1.2216). The found equation (z_{Germany}) gives the unemployment rate of Germany when the variables are replaced in this equation.

The equation obtained by Lagrange interpolation method for Turkey has been calculated as follows;

$$\mathrm{z}_{\mathrm{Turkey}} = 0.0492774\mathrm{x}^2 + 0.00700056xy - 0.0112508y^2 - 0.400765x - 0.0738538y + 12.147.$$

Where the coefficient of GDP is negative (-0.400765) and the coefficient of consumer price is negative (-0.0738538). The found equation (z_{Turkey}) gives the unemployment rate of Turkey when the variables are replaced in the equation.

While the geometric display of Turkey indicates a convex view, the geometric display of Germany shows a concave view. The results of Lagrange interpolation method are notably remarkable. The geometric demonstrations can be interpreted differently. The graphs can be seen as the planary sight in three-dimensional space for Turkey and Germany. The geometric demonstration, which is obtained by the German data, tends to curve top-down. That is, the unemployment rate of Germany from 2006 to 2011 can be interpreted as in a drop tendency, but the situation is exactly opposite for Turkey.

As a suggestion for further research, this study may be extended by using a larger data set on unemployment rates in Turkey and in Germany and the results may be compared with the current findings. However, more countries from the European Union (EU) or from other emerging countries may also be included into our study. Finally, the method of OLS or bivariate lagrange interpolation can be compared with alternative methods. In this regard, for example the method of ANN can be used as a newly developing method to be compared with these two methods or any other linear or non-linear methods. In other words, the obtained results from ANN can be compared with the obtained results from Lagrange interpolation and OLS in terms of evaluating the performances of these methods.

References

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.

Arhipova I., & Balina S. (2004). *The Problem of Choosing Statistical Hypotheses In Applied Statistics*.

COMPSTAT (2004). 16 th Symposium of IASC, Proceedings in Computational Statistics, Prague/Czech Republic, Physica-Verlag, A Springer Company. 629 – 636.

Barnichon, R. & Nekarda, J.C. (2012). *The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market*. Brookings Papers on Economic Activity, Fall 2012.

Berrut, J-P. & Lloyd N. T. (2004). Barycentric Lagrange Interpolation. *SIAM Review.* 46 (3), 501–517.

Bourbaki, N. (1987). *Topological Vector Spaces*.

Carnicer J. M. & Gasca M. (2005). Generation of Lattices of Points for Bivariate Interpolation, Numer. *Algorithms.* 39(1-3), 69–79.

Davidson, R. & James G. MacKinnon (2004). Econometric Theory and Methods. Oxford University Press.

Employment Plan of Germany (2014). Available at: http://www.bmas.de/SharedDocs/Downloads/DE/Thema-Internationales/g20-employmentplan-germany.pdf?__blob=publicationFile.

Fernandes, R. & Leblanc, S.G. (2005). Parametric (modified least squares) and Nonparametric (Theil–Sen) Linear Regressions for Predicting Biophysical Parameters in the Presence of Measurement Errors. *Remote Sensing of Environment*. 95, 303–316.

Gasca, M. & Sauer T. (2000). Polynomial Interpolation In Several Variables. *Advances in Computational Mathematics*. 12(4), 377–410.

Gil-Alana, L. A. (2001). A Fractionally Integrated Exponential Model for UK Unemployment. *Journal of Forecasting*. 20, 329-340.

Hayashi, F. (2000). *Econometrics*. Princeton University Press.

Johnes, G. (1999). Forecasting Unemployment. *Applied Financial Economics*. 6, 605-607.

Koop, G. & Potter, S. M. (1999). Dynamic Asymmetries in U.S. Unemployment. *Journal of Business and Economic Statistics*. 17, 298-312.

Kwok, A.C.F. (2011). Effects of Economic Performance and Immigration on Unemployment of Low-Skilled Working Force. *Journal of Applied Economics and Business Research*. 1(2): 75-92.

Milas, C. & Rothman, P. (2008). Out-of-sample Forecasting of Unemployment Rates with Pooled STVECM Forecasts. *International Journal of Forecasting*. 24 (1), 101-121.

Montgomery, A., Zarnowitz, V., Tsay, R. & Tiao, G. (1998). Forecasting the U.S. Unemployment Rate. *Journal of the American Statistical Association*. 93, 478–93.

Montgomery, D.C., Peck, E.A., & Vining, G.G. (2001). *Introduction to Linear Regression Analysis.* 3rd edition. New York : John Wiley & Sons, Inc.

Moshiri, S. & Brown, L. (2004). Unemployment Variation Over the Business Cycles: A Comparison of Forecasting Models. *Journal of Forecasting*. 23(7), 497-511.

Olver, P.J. (2006). On multivariate Interpolation. *Studies in Applied Mathematics*. 116(4), 201–240.

Peel, D. A. & Speight, A. E. H. (2000). Threshold Nonlinearities in Unemployment Rates: Further Evidence for the UK and G3 Economies. *Applied Economics*. 32, 705-715.

Plackett, R.L. (1950). Some Theorems in Least Squares. *Biometrika.* 37(1-2), 149–157.

Proietti, T. (2001). *Forecasting the US Unemployment Rate*. Working Paper. Dipartimento di Scienze Statistiche, Universita di Udine, Udine, Italy.

Rothman, P. (1998). Forecasting Asymmetric Unemployment Rates. *Review of Economics and Statistics*. 80, 164-168.

Scott, A. J. & Holt, D.(1982). The Effect of Two-Stage Sampling on Ordinary Least Squares Methods. .*Journal of the American Statistical Association.*77(80), 848-854.

Sosa, E. W. (2009). OLS Geometry. *Econometric Analysis*. Econ 507.

http://en.wikipedia.org/wiki/Ordinary_least_squares

http://en.wikipedia.org/wiki/Lagrange_polynomial