The Second International Conference on Physics, Mathematics and Statistics
IOP Conf. Series: Journal of Physics: Conf. Series 1324 (2019) 012087
doi:10.1088/1742-6596/1324/1/012087

The optimization of kink regression with differential evolution (DE)

Pathairat Pastpipatkul¹, Petchaluck Boonyakunakorn², ³ and Songsak Sriboonchitta²

¹Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand
²Puey Ungphakorn Center of Excellence in Econometrics, Chiang Mai University, Chiang Mai, Thailand
³Email: petchaluckecon@gmail.com

Abstract. This paper aims to estimate the kink model with respect to Thailand’s exports to Thai GDP. Another main contribution of this study is to compare the performance of optimization the kink regression model with DE and MLE. The traditional optimization is maximum likelihood estimator (MLE) which has problems in the case where the function is nondifferentiable, or the likelihood is difficult to find. Furthermore, it is difficult to obtain the global maximum in the non-linear model. One of the optimization techniques is DE which is a successful method for searching for the global maximum. To evaluate the estimated performance of DE and MLE, we apply the Monte Carlo simulation. The simulation result indicates that DE has the both lower MSE and bias. We apply the regression model with respect to Thai exports to Thai GDP. For the estimated parameter results show that the kink point is -0.016%. The estimated coefficients in the first part and second part are 1.144% and 0.341%, respectively.

1. Introduction
Optimization is one of the key parts on econometric applications to economic data such as parameter estimation. In the first place, Least Squares (LS) estimation was the most widely used to estimate parameters in a linear model. Later Maximum Likelihood Estimation (MLE) has become widely used in the present. The efficient numerical methods can be an appropriate solution for solving the optimization problem only when the likelihood function is globally convex. Nevertheless, the likelihood function is flat in many cases. Some cases, the likelihood function has multiple local optima. Some of the studies, such as [1] [2] showed that in the case of estimation, it is not appropriate to use the standard numerical approximation algorithms called optim estimating GARCH (1,1).

Differential evolution (DE) was developed by Rainer Storn and Kenneth Price in 1990 called DEoptim in this study. A population of solution vectors are algorithm updated by addition, subtraction and crossover. After that, it chooses the most suitable among the real data and updated population. This approach works simultaneously on an entire set of solutions. The results are more consequently efficiency. Thus, DE should be more appropriate to find the global optimum of a real-valued function of real-valued parameters. The advantage of Global Optimization by DE is the possibility to deal with discontinuous and nondifferentiable functions. Therefore, DE has been widely applied in various fields and found to be a successful tool for finding the global maximum [1]. In the financial area, Byachkova and Simonov [3] used DE to minimize Kullback-Leibler divergence in modelling a financial market.
They found that the optimization procedures are useful to find the parameters which describe the real process of market prices.

Recently, nonlinear models have been widely used in economic data, since there is evidence demonstrate that many economic variables have a nonlinear behaviour. The kink model is also one of the non-linear models that has been used in econometric recently. Kink regression model with an unknown threshold was proposed by Card, et al [4] and Hansen [5]. The regress kink model has been commonly known for the discontinuous regression [3]. The continuous regression function found out that at a threshold point, its slope has discontinuance. The regression kink model is appropriate in the case that the threshold effect come from only one interested variable, and we expect that there is a discontinuous function at the threshold. It also includes estimating the changes in the slope of the relationship between dependent variable and explanatory variable. Thus, it becomes workable for detecting local changes in the slope. Later, the study of Hansen [5] introduced an extended method to estimate and inference in kink regression models combining the point of changing the slope as an unknown parameter with an assumption of continuity of the regression model.

Thai economy heavily relies on Thai exports, accounting for almost 70% of its GDP. Figure 1 shows that both Thai GDP and Thai exports have an upward trend. During this period, Thai exports to ASEAN grew sharply before the global crisis in 2008. After 2009, it recovered with sharp growth again until the financial crisis in 2012. After 2012, the growth of Thailand’s exports to ASEAN became steady with a slight downward trend. From 2016 until now, it has an increasing trend. Investigating the relationship between them is interesting.

Considering estimation of the kink regression model, the LS and tradition MLE was intensively applied. However, in many cases, the likelihood is difficult to find. Some cases, the likelihoods are nondifferentiable. Moreover, the kink model is a non-linear model which is also difficult to obtain the global maximum by traditional optimization techniques. To overcome these issues, we propose to use optimization by DE in the process of the parameter’s estimation. Thus, this paper aims to estimate the regression kink model with respect to Thailand’s exports to Thai GDP. Another main contribution of this study is to compare the performance of optimization of the kink regression model by using DE and MLE.

![Figure 1](image_url). The plot of the values of GDP and exports of Thailand.

This paper is divided into five parts. The first one is an introduction. The second one provides a discussion of the kink model. Next, the simulation study is in the third section. The fourth section is devoted to data and estimation results. And, conclusion is on the last section.

2. Methodology

2.1. Kink model

Considering the Kink model as follows

\[ y_i = \alpha + \phi^1(x_i - r) + \phi^2(x_i - r) + \epsilon_i, \]  

(1)
where \( y_t \) is the interested time series variable for \( t = 1, 2, \ldots, T \). \( \phi^{(1)} \) and \( \phi^{(2)} \) are the coefficients for the lower and higher regime respectively, \( r \) is the non-trivial threshold variable referring to the kink point value. The slope \( \phi^{(1)} \) with respect to the variable \( x_t \) for the value \( x_t \) less than \( r \), and the slope \( \phi^{(2)} \) for the value \( x_t \) higher than \( r \). \( \varepsilon_t \) is a white noise. Equation 1 represents a regression kink model, which refers to the regression function that is continuous in the variable \( x \). However, the slope of \( x \) is referred as discontinuous (has a kink) at \( x = r \).

Considering the log-likelihood function, the kink model is expressed as

\[
\ln L = \ln \prod_i D_i \left( y_t, E(y_t|I_{t-1}) \varepsilon_t \right),
\]

where \( \nu \) is the parameter vector consisting of parameters \( \alpha, \phi^{(1)}, \phi^{(2)}, r \). \( D_i \) is the conditional distribution function which is supposed to be a normal distribution. \( I_{t-1} \) denotes all the information at time \( t-1 \).

After assuming the error term is normal distribution, the log-likelihood can be expressed as

\[
\ln L = \ln \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{1}{2\sigma^2} \varepsilon_t^2 \right),
\]

Then, we maximize this log-likelihood related to the parameter vector. Bearing in mind that \( \varepsilon_t = y_t - \alpha - y_{t-1}\phi^{(1)}I(x_t - r) - y_{t-1}\phi^{(2)}I(x_t - r) \).

### 2.2. The differential evolution (DE)

It was initially proposed by Storn and Price in 1996 to optimize real parameter and real valued function [6]. DE comes from the genetic algorithms (Gas)’s class that are one of the most popular stochastic and heuristic techniques based on the operations of crossover in biology, mutation, and the natural selection to minimize an objective function over the successive generations’ course [7]. In past few years, this method has been performed a simply adjustable probing for worldwide optimization.

DE is also one of the evolutionary algorithms that solves the problem of optimization through evolving a population of candidate solutions using alteration and selection operators. The difference between DE and Gas is that the DE employs floating-point, whereas GAs uses bit-string encoding of population members. Another difference is DE uses arithmetic operations, whereas GAs uses logical operations in mutation.

Suppose \( NP \) represents the number of parameter vectors, the dimension is denoted by \( d \). The initial generation is created by the number of parameters which guesses for the optimal value or applying random values between upper and lower bounds. Each generation which a new population is constructed from the current population members \( \{ x_i | i = 1, \ldots, NP \} \), where \( i \) denotes the vectors that create the population. In order to achieve the solution, the differential mutation of the population members is applied. \( v_i \) represents vector of the initial mutant which is created by selecting randomly among \( x_i, x_j, x_k \). Therefore, the generation of initial mutant parameter vector \( v_i \) can be expressed as

\[
v_i = x_i + F \left( x_j - x_k \right),
\]

where \( F \) represents a factor of positive scale which should be less than one for the effective values. When the first mutation operation is obtained, mutations have been created until \( d \) mutations with a cross probability (\( CR \)) between zero and one. The parameters values’ fraction which are copied from the mutant is controlled by CR. In the case that the trial parameter vector violates the bounds after mutation and crossover, it will reset in order to obtain the new value that respects the bounds. Thus, the values of the objective function related to the children are determined. In the case that the trial parameter vector is equal or lower value of the objective function than the previous vector, it will substitute the previous vector in the population. If not, the previous vector will be remained.
Thus, the scheme’s effect is that the distribution’s shape in the search space is converging regarding with size and direction to areas by great fitness. When the populations are closer to the global optimum, the distribution will be more shrink. Eventually, reinforce the generation of smaller difference vectors. DE has been developed continually. Recently, many researchers have intensely investigated and developed the application of the DEoptim. For example, Feolktistov, et al [8] developed the method to test the genotypes. Yaman, et al [9] studied the convergence of three differential evolution strategies by setting different population sizes. Caraffini and Neri [10] studied the efficiency of a rotation-invariant DE by comparing with the standard counterpart.

3. Simulation study

Table 1. Simulation result of kink model.

| Estimation technique | Method     | Parameters | True Value | MSE    | Bias    |
|----------------------|------------|------------|------------|--------|---------|
| DEoptim              | $\alpha$ 1 | 0.2981     | -0.5444    |        |         |
|                      | $\phi^{(1)}$ -2 | 0.0007*    | -0.0025*   |        |         |
|                      | $\phi^{(2)}$ 0.5 | 0.0015*    | 0.0029*    |        |         |
|                      | $r$ 6 | 0.0553* | 0.2072     |        |         |
| Optim                | $\alpha$ 1 | 0.4940     | 0.0050*    |        |         |
|                      | $\phi^{(1)}$ -2 | 0.0256     | 0.0145     |        |         |
|                      | $\phi^{(2)}$ 0.5 | 0.1161     | -0.0101    |        |         |
|                      | $r$ 6 | 0.2649     | -0.0086    |        |         |
| Nelder-Mead          | $\alpha$ 1 | 0.2611     | 0.0268     |        |         |
|                      | $\phi^{(1)}$ -2 | 0.0238     | 0.0185     |        |         |
|                      | $\phi^{(2)}$ 0.5 | 0.0223     | -0.0027    |        |         |
|                      | $r$ 6 | 0.1647     | -0.0105    |        |         |
| L-BFGS-B             | $\alpha$ 1 | 0.2582*    | 0.0089     |        |         |
|                      | $\phi^{(1)}$ -2 | 0.0204     | 0.0123     |        |         |
|                      | $\phi^{(2)}$ 0.5 | 0.0220     | -0.0018    |        |         |
|                      | $r$ 6 | 0.1618     | -0.0021*   |        |         |

Note: * indicates the minimum value of MSE and Bias.

In this study, we employ Monte Carlo simulation to assess the performance of the estimated results on kink regression model between two main optimizations, namely Optim and DEoptim. For the optim technique, “BFGS”, “Nelder-Mead”, and “L-BFGS-B” method is chosen. For evaluation of the performance of these optimization techniques, Mean Squared Error (MSE), and bias approaches are employed.

In this simulation study, we set the true value for coefficients and the threshold value. We set the intercept $\alpha = 1$, the coefficient for lower part $\phi^{(1)} = -2$, the coefficient for upper part $\phi^{(2)} = 0.5$, and the threshold $r = 6$. The error terms are generated from uniform (0,1). In this simulation study, we simulate 1,000 datasets which could be enough for the simple experiment study. The Bias is calculated as follows

$$Bias = N^{-1} \sum_{i=1}^{N} (\hat{\phi}_i - \phi)$$ (5)

The MSE is calculated as follows.
\[
MSE = M \sum_{i=1}^{n} (\hat{\phi} - \phi)^2,
\]

where \( \hat{\phi} \) is the estimated value and \( \phi \) is the true value.

From Table 1, we can see that the mean squared error (MSE) for \( \phi^{(1)} \), \( \phi^{(2)} \), \( r \) and from DEoptim provided the lowest value. For the intercept \( \alpha \), the lowest MSE is obtained by “L-BFGS-B”. Regarding the Bias, DEoptim provides the lowest bias for \( \phi^{(1)} \), \( \phi^{(2)} \). The lowest bias for intercept \( \alpha \), threshold \( r \) are belonged to “BFGS” and “L-BFGS-B”, respectively. Over all, the DEoptim provides the minimum MSE and bias. Thus, we can conclude that DEoptim provides the better result when compared with Optim technique.

4. Data and empirical result

4.1. Data

The study consists of quarter data which ranges from quarter 1, 2000 to quarter 4, 2017. All data is obtained from DATASTREAM. As stationarity is primarily required by the application of the estimation method, log transformation on Thai GDP and Thai exports is employed. Then, the unit roots test is implemented.

\textbf{Table 2.} Unit root test.

| Method   | GDP   | EXPORT | GDP   | EXPORT |
|----------|-------|--------|-------|--------|
| ADF-test | -1.637| -1.580 | -5.158*| -4.107*|
| PP-test  | -6.989| -4.416 | -71.297*| -59.807*|

Note: * denotes substantial evidence against Ho: There is a unit root is present in a time series.

4.2. Empirical result

Table 2 shows both empirical results before and after applying the earlier processes. To make the statistic inference, we use Minimum Bayes Factors (MBF). To compute MBF, we follow the work of Sellke et al [6] that the value of MBF can be computed by \(-ep \log p\). For the interpretation of the value of MBF, we follow the work of Held and Ott [11] which provided the criteria of the level of evidence against the null hypothesis. The level of strength of evidence can be separated into six groups which are weak, moderate, substantial, strong, very strong, and decisive.

The time-series data of Thai GDP and Thai exports at level are nonstationary for a unit root in both ADF and PP unit root test. The data become stationary with substantial evidence against the null hypothesis when they are transformed by using the first difference.

The model specification

The specification of our model can be expressed as

\[
GDP_t = \alpha + \beta_1 (EXPORT_t - r) + \beta_2 (EXPORT_t - r) + \epsilon_t,
\]

where \( GDP_t \) corresponds to gross domestic product that measures of the monetary value of all final goods and services of Thailand at time \( t \). \( EXPORT_t \) corresponds to the value of Thai exports at time \( t \).

According to Table 3, the empirical result shows that the estimated coefficient of this first part or negative part is 1.144. It interprets that if there is a rise in the growth of Thai GDP by 1%, there will be a rise in growth of Thai exports by 1.144 %. Meanwhile, if Thai GDP growth exceeds the level of kink point at -0.016%, the estimated coefficient is 0.341%, indicating that the effect of the growth of
Thai GDP in the latter part is considerably higher than the other one. In other words, if Thai GDP growth exceeds the level of -0.016%, the growth of Thai exports will increase by 0.341%.

Table 3. Estimating Parameters.

| Parameters | DEoptim |
|------------|---------|
| $\alpha$   | 0.014   |
| $\beta_1$  | 0.144   |
| $\beta_2$  | 0.341   |
| $r$        | -0.016  |

Source: Calculations

Figure 2. The plot of kink point.

Figure 2 shows the relationship between Thai GDP and Thai exports based on regression kink model. The horizon axis on the graph corresponds to the growth of Thai exports with the other axis corresponding to Thailand’s GDP growth on another axis. The fitted regression shows the trend of a positive slope for low Thai GDP. At a kink point at -0.016%, the slope trend sharply went up for Thai GDP growth. This indicates that when the growth of GDP exceeds a threshold at -0.016, the Thai GDP growth tends to move faster.

5. Conclusions

This paper proposes to use Differential Evolution (DE) in estimating the kink regression model. The simulation estimated results show that the DE provides the smallest MSE and Bias compared with the Optim optimization technique with three different methods; “BFGS”, “Nelder-Mead”, and “L-BFGS-B”. Thus, we can conclude that DE tends to be a more appropriate method for estimating the kink model compared to traditional optimization techniques. Furthermore, we also provide the estimated parameter of the non-linear relationship between Thai GDP and Thai export with DE estimation. The result shows that the kink point is -0.016%.

References

[1] Jerrrell M E and Campione W A 2001 Global optimization of econometric functions. *Journal of Global Optimization* **20**(3-4) 273-295
[2] Winker P and Maringer D 2006 The convergence of optimization based GARCH estimators: Theory and application In Compstat 2006-Proceedings in Computational Statistics Physica-Verlag HD 483-494

[3] Byachkova A and Simonov A 2015 Modeling Financial Market Using Percolation Theory In Financial Econometrics and Empirical Market Microstructure Springer, Cham, 47-53

[4] Card D, Mas A and Rothstein J 2008 Tipping and the Dynamics of Segregation The Quarterly Journal of Economics 123(1) 177-218

[5] Hansen B E 2017 Regression kink with an unknown threshold Journal of Business & Economic Statistics 35(2) 228-240

[6] Storn R 1996 On the usage of differential evolution for function optimization In Biennial conference of the North American fuzzy information processing society (NAFIPS), IEEE Berkeley, 519

[7] Sellke T, Bayarri M J and Berger J O 2001 Calibration of p values for testing precise null hypotheses The American Statistician 55(1) 62-71

[8] Feoktistov V, Pietravalle S and Heslot N 2017 Optimal experimental design of field trials using Differential Evolution arXiv preprint arXiv:1702.00815

[9] Yaman A, Iacca G and Caraffini F 2018 A comparison of three Differential Evolution strategies in terms of early convergence with different population sizes

[10] Caraffini F and Neri F 2018 Rotation Invariance and Rotated Problems: An Experimental Study on Differential Evolution. In International Conference on the Applications of Evolutionary Computation Springer, Cham, 597-614

[11] Held L and Ott M 2016 How the maximal evidence of p-values against point null hypotheses depends on sample size The American Statistician 70(4) 335-341