LOGNORMAL ORDINARY KRIGING METAMODEL

Muzaffer Balaban
Turkish Statistical Institute & Department of Industrial Engineering
Başkent University
Ankara, Turkey
balabanmuzaffer@gmail.com

Berna Dengiz
Department of Industrial Engineering
Başkent University, Ankara, Turkey
bdengiz@baskent.edu.tr

Abstract

This paper provides a comprehensive application of Lognormal Ordinary Kriging (LOK) metamodel to optimization of a stochastic simulation problem. Kriging models have been developed as an interpolation method in geology. Firstly kriging methods have been successfully used for deterministic simulation optimization problem. Recently, Kriging metamodeling has attracted a growing interest, along with other simulation optimization (SO) techniques for deterministic and stochastic systems. SO researchers have started using ordinary kriging as a robust optimization tool in stochastic SO in the last decade. In this paper SO process via lognormal ordinary kriging metamodel is proposed and a lognormal ordinary kriging algorithm is presented.

Keywords
Ordinary kriging, lognormal ordinary kriging metamodel, simulation optimization

1 INTRODUCTION

Simulation Optimization (SO) is a systematic technique to obtain good input/output (I/O) relation models to system design and evaluation when any analytical function is unknown or so complex.

Carson and Maria (1997) defined SO as the process of finding the best input variables values from among all possible combinations without explicitly evaluating each possibility. The main assumption of SO is to estimate the objective function from simulation results when it is not available directly (Fu et all 2005).

Metamodel is defined as a "model of model," which is used for approximation of the input/output (I/O) function for simulation model by Kleijnen (Kleijnen 1979, 2007). Metamodeling is the one of simulation optimization methods and used to find proper functional relationship between outputs and input variables of a simulation model based on the result of experiments when the input variables are continues variables.

Recently kriging models have been used as a global metamodel in SO. Kriging was originally developed as a geostatistical interpolation technique for interpolation of input point data and estimation of a mineral resource model in mining industry by Matheron (1963). Kriging is an optimal spatial regression technique which requires a spatial statistical model, popularly known as a variogram, representing the internal spatial structure of the data (Cressie 1990, 1993).

Sacks et all. first applied kriging models to the deterministic simulation model in 1989 (Van Beers et all. 2002). Mitchell and Morris first mentioned to use of kriging models to stochastic simulation models as one of the metamodel types in 1992 (Ankerman et all 2010). Van Beers and Kleijnen first applied kriging models to stochastic simulation models as SO method in 2003. Although a significant body of SO literature exists, a review finds few examples with ordinary kriging as metamodel for stochastic simulation.

In our paper a Lognormal Ordinary Kriging (LOK) metamodel is proposed for stochastic SO. To show the the performance of SO with LOK the communication network system that was described by Barton and Meckesheimer (2006) was considered in this study. According to the best of our knowledge LOK hasn't been used as a metamodel in random simulation optimization until now. However there are some useful application in geology such as Paul and Cressie (2011), Gilbert and Simpson (1985), Mc Gratha and Zhangb (2004), Lark and Lapworth (2012), etc.

Remaining part of this paper is organized as follows. Kriging techniques are discussed in section 2. Problem statement, experimental design, simulation and application of lognormal ordinary kriging to SO results are explained respectively in section 3. Results and conclusions are given in section 4.

2 KRIGING TECHNIQUES

Kriging is a method of optimal, unbiased prediction of regionalized variables at unsampled locations using the structural properties of the variogram and the initial set of data values (David, 1977). Kriging is also known as a best linear unbiased predictor (BLUP).

The input data are weighted based on the variogram model in the kriging models. Kriging provides the estimation variance at every estimated point, which is an indicator of the accuracy of the estimated value. The effectiveness of kriging depends on the correct specification of the variogram model. More details on the kriging techniques are given in the classical reference books by Journal and Huijbregts (1978), Isaaks and Srivastava (1989) and Cressie (1993).

Objective of the kriging is to predict the value of a random variable Z(x) at unobserved points in a region D from observed data $\{z(x_1), z(x_2), \ldots, z(x_n)\}$. $\hat{Z}(x_0) = \sum_i^n \lambda_i Z(x_i)$, a weighted average of observed data with the weights λ_i. We want the prediction to be unbiased,

$$E\left(\hat{Z}(x_0)\right) = E\left(Z(x_0)\right). \tag{1}$$

2.1 Variogram

Variogram (semivariance or semivariogram) estimation is a very important step of the kriging, because it determines the kriging weights (Genton, 1998). The original semivariogram term was utilized by Matheron because of an half of a variance (Myers 1991).

$Z(x)$, $\{z(x_1), z(x_2), \ldots, z(x_n)\}$ is sample data satisfying second order and intrinsic stationary. Matheron (1965) defined that a logical estimator of experimental variogram as

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}(Z(x_i) - Z(x_i + h))^2 \tag{2}$$

where N(h) is the number of pairs $Z(x_i)$, $Z(x_i + h)$ and h is the distance between the experiments (Myers 1991). Cresssie and Hawkins (1980) proposed a robust variogram estimator as follows:

$$\gamma(h) = \frac{1}{2}\left[\frac{1}{N(h)}\sum_{i=1}^{N(h)}\left(Z(x_i) - Z(x_i + h)\right)^{\frac{1}{2}}\right]^4 / \left(0.457 + \frac{0.494}{N(h)}\right) \tag{3}$$

When calculation of the kriging weights for each point, we need a theoretical variogram model. Theoretical variogram model must fit to the experimental variogram. Myers (1991) has listed four variogram models as a function of distance (h).

In this study below exponential variogram model is used.

$$\gamma(h) = \begin{cases} a_o + a_1\left[1 - \exp\left(\frac{-h}{a_2}\right)\right] & \text{if } h \neq 0 \\ 0 & \text{if } h = 0 \end{cases} \tag{4}$$

2.2 Ordinary Kriging

Z(x) random variable satisfies second order and intrinsic stationary in Matheron (Cressie 1990, 1993 and Myers 1991).

$$Z(x) = \mu + \varepsilon(x) \tag{5}$$

$E(Z(x)) = \mu$ is constant and unknown. Kriging predictor is

$$\hat{Z}(x_0) = \sum_i^n \lambda_i Z(x_i) \text{ with } \sum_i^n \lambda_i = 1 \tag{6}$$

Mean square prediction error is

$$(Z(x_0) - \hat{Z}(x_0))^2 = (Z(x_0) - \sum_i^n \lambda_i Z(x_i))^2 = -\sum_i^n \sum_j^n \lambda_i \lambda_j \gamma(x_i - x_j) + 2\sum_i^n \lambda_i \gamma(x_0 - x_i) \tag{7}$$

Minimizing mean square prediction error using Lagrange multipliers method we get the following system of linear equations.

$$\sum_j^n \lambda_j \gamma(x_i - x_j) + m = \gamma(x_o - x_i)\text{, for i = 1, 2, . . . , n} \tag{8}$$

$$\sum_i^n \lambda_i = 1. \tag{9}$$

We can write the system of equations in matrix form as

$$\Gamma_o \lambda_o = \gamma_o \tag{10}$$

where,

$$\Gamma_o = \begin{vmatrix} \gamma(x_1 - x_1) & \gamma(x_1 - x_2) & \dots & \gamma(x_1 - x_n) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma(x_n - x_1) & \gamma(x_n - x_2) & \dots & \gamma(x_n - x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{vmatrix} \tag{11}$$

$$\gamma_o = (\gamma(x_1 - x_o), \gamma(x_2 - x_o), \gamma(x_3 - x_o), \dots\dots, \gamma(x_n - x_o), 1) \tag{12}$$

$$\lambda_o = (\lambda_1, \lambda_2, \lambda_3, \dots\dots, \lambda_n, m) \tag{13}$$

Thus,

$$\lambda_o = \Gamma o^{-1} \gamma_o \tag{14}$$

is calculated and prediction of $Z(x_0)$ is obtained by

$$\hat{Z}(x_0) = \lambda\, Z \tag{15}$$

We only need to recalculate γ_o for a new point x_0. The matrix Γ does not change for new positions of x_0. The minimized prediction variance is also called the kriging variance by some authors and denoted by $\sigma^2(x_0)$ is

$$\sigma^2(x_0) = \sum_i^n \lambda_i \gamma(x_i - x_0) + m \tag{16}$$

2.3 Lognormal Kriging

If the data are so much skewed, estimation errors are not normally distributed. A solution of this problem is to transform the data (Roth 1998).

Let, Z(x), x Є D, is a random processes that Y(x) = ln Z(x), x Є D, is normally distributed. The objective is to predict $Z(x_o)$ random variable. The main step is to transform the problem from Z to the intrinsically stationary normally distributed Y.

Predictor of $Y(x_o)$ is;

$$\hat{Y}(x_0) = \sum_i^n \lambda_i \ln Z(x_i) = \sum_i^n \lambda_i\, Y(x_i). \tag{17}$$

Back-transformation of $\hat{Y}(x_0)$ to $\hat{Z}(x_0)$ is unbiased using following equation.

$$\hat{Z}(x_0) = \exp\left(\hat{Y}(x_0) + \frac{\sigma^2_{Y,k}}{2} - m_Y\right) \tag{18}$$

Where $\sigma^2_{Y,k}$ is kriging variance of Y and m_Y is Lagrange multiplier value (Cressie 1993). In literature there are several studies on lognormal kriging such as Journell and Huijbregts (1978) and Dowd (1982).

3 PROBLEM STATEMENT AND APPLICATION

3.1 Problem Statement

The problem studied in this paper is to find input factors (message routing percentages) that minimize the total cost (c) of the system having random arriving messages through a communication network system. This problem that was described by Barton and Meckesheimer (2006) was considered in this study with 1000 random messages which are routing through 3 networks.

A message goes to network 1 with p_1 %, network 2 with p_2%, remaining messages goes to network 3 as shown in Figure 1. The system cost are computed considering the assumptions such as $0.005 per time unit for each message and a message processing costs, a_i, are $ 0.03, $ 0.01, $ 0.005 for three networks, respectively . The interarrival time of messages has an exponential distribution with a mean of 1 time unit. The network transit times have an triangular distributions with a mean of $E(S_i)=i$ and its limits +/- 0.5 for each network i.

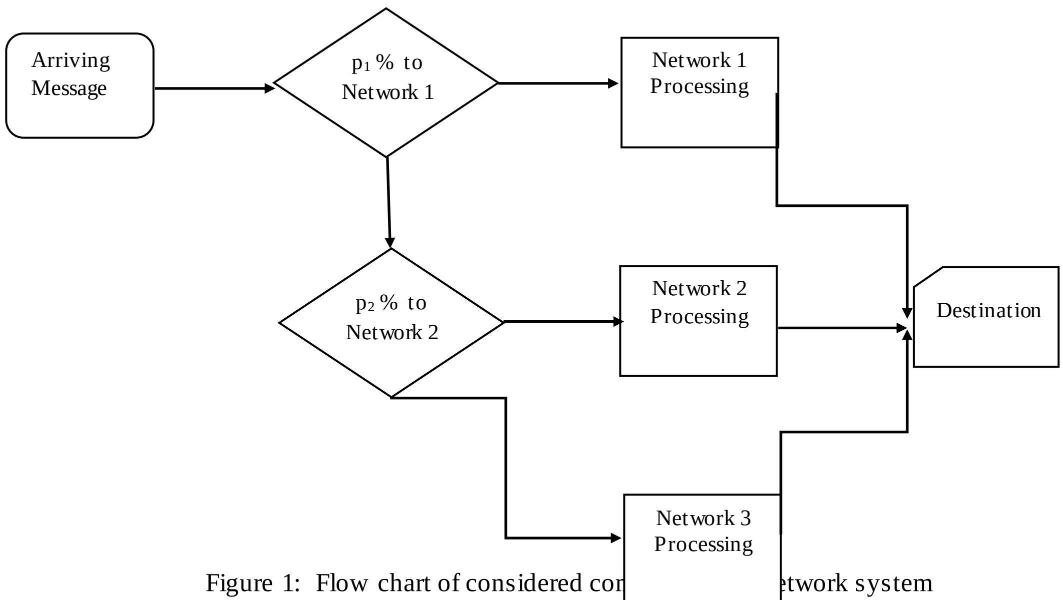

Figure 1: Flow chart of considered communication network system

The objective is to find routing percentages (p_1, p_2) which minimize the total cost of the system p_1 and p_2 are input factors of SO problem. Afterwards x_1, x_2 are used instead of p1, p2 and Z(x) to indicate the total cost of the system.

A simulation model of the system was constructed in Arena platform. The validation was done comparing the two system results from the simulation model of Barton and Meckesheimer (2006) with this simulation model.

3.2 Experimental Design

The objective of the experimental design is the selection of the input factors combinations where the outputs should be evaluated.

Latin hypercube designs (LHD) are often used to find fitted kriging metamodel in the I/O simulation data (Kleinen 2007). LHD designs are especially well suited for kriging because they can cover the design space (Biles et all 2007) LHD was first described for computer experiments by McKay et al. in 1979. LHD are produced by Latin hypercube sampling, dividing each factor axis into n equally intervals (Santner 2003).

A randomly generated LHD may cause two problems. First, the factors may be perfectly correlated. Second, a large area may be not explored in the experimental region. There are some studies in the literature to avoid these problems (Joseph 2008). It is firstly needed to decide how many design points to be used in LHD. The range of each input variable is divided into k equally probable intervals. We obtain a nxn design matrix for two input factors. Each column and row has one design point (Van Beers and Kleijnen 2004).

In our study LHD is used for two dimensional input variables as experimental design strategy. The design region is assumed as x_1= (35, 80), x_2= (35, 80), intervals of each input variable is 3 and the number of the design points is 16.

A random LHD algorithm is used to obtain the design points and the design region is divided in four blocks to ensure that the design points distributed evenly to avoid the perfect correlation between input variables.

The random LHD algorithm is coded C++ and simulation model runs with 10 replications for these 16 designed points.

3.3 Lognormal Ordinary Kriging Calculation

In this section using the methodologies at section 2.1, 2.2 and 2.3 we have developed following algorithms to find the input factors (message routig probabilities) that minimize the total Z(x) of the system.

Lognormal Ordinary Kriging Algorithm for SO

Step 1: *Input:* $x_i=(x_{1i}, x_{2i})$, $Z(x_i)$, $Y(x_i)=ln(Z(x_i))$ *for i=1to n*

Step 2 ***a)*** *Calculate semivariance of each pairs of experiments*
$\gamma(i,j)=0.5*Var(Y(x_i), Y(x_j))$
b) *Calculate distance of each pairs of experiments.*
$h(i,j)=\sqrt{(x_{1i}-x_{1j})^2+(x_{2i}-x_{2j})^2}$

Step 3 ***a)*** *Create Experimental Variogram,* Γ_e
b) *Find Fitted Theoretical Variogram Model,* $\gamma(h)$
c) *Create Theoretical Variogram,* Γ_o
d) *Calculate Inverse of Theoretical Variogram,* Γ_o^{-1}

Step 4: *Set* Z^*, $x_1{}^*$, $x_2{}^*$ *initial optimal values*

Step 5 ***a)*** *Set factors values* $x_o(x_1, x_2)$ *for a new estimation point,* x_1*: from 35 to 80 and* x_2*: from 35 to 80*
b) *Calculate distance between the new estimation point and the experiments.*
$h(0,i)=\sqrt{(x_1-x_{1i})^2+(x_2-x_{2i})^2}$
c) *Calculate Variogam vector using Theoretical Variogram Model,* $\gamma(0)$

Step 6 ***a)*** *Calculate OK coefficient*
$\lambda_i = \Gamma_t^{-1}\gamma(0)$
b) *Calculate the cost for the new factors.*
$Y(x_0) = \sum_{i=1}^{n}\lambda_i Y(x_i)$ *and* $Z(x_0)=exp(Y(x_0)+\frac{\sigma_{Y,k}^2}{2}-m_Y)$

Step 7 ***a)*** *If* $Z(x_0) < Z^*$ *set* $x_1{}^* = x_1$, $x_2{}^* = x_2$, $Z^* = Z(x_0)$
b) *If all combination of factors are evaluated go to step 8*
Otherwise go to step 5

Step 8: *Report the kriging results.* Z^*, $x_1{}^*$, $x_2{}^*$

Step 9: *Calibrate the kriging result using simulation model*

DACE toolkit is used generally in kriging metamodeling for simulation optimization (Van Beers an Kleijnen 2002, Kleijnen 2007, Biles et all 2007). In our study our own c++ routine of LOK algorithm used because there are no available routines for LOK modeling and optimization.

Logarithmic transformation was applied to data obtained with LHD because the high variation among replications as discussed at section 2.3. Exponential variogram model is selected as fitted variogram model $\gamma(h)=0.3*(1-exp(-h/20))$. Then a searching algorithm was used for the whole response region which is 45x45 input grids of x_1 and x_2 to find the minimum Z using LOK metamodel. Therefore a minimum output of OK model is predicted as $Y^*=3.4988$ at the input variables setting $x_1{}^*, x_2{}^*=\{55, 63\}$ with kriging variance, $\sigma_{Y,k}^2$=0.0353 and lagrange multiplier, $m_Y=0$.

Back-transformation of ln Z^* to Z^* is Z^* = exp (3.4988 + 0.0353/2) = 33.6647.

The last step of the SO is the calibration of results using the simulation model of the system. This step is important to obtain actual outputs of model to eliminate the kriging errors. Simulation model replications give average output as Z^*=32.99 with variance $\hat{\sigma}^2$ = 0.1702 for the input factors setting, $x_1{}^*, x_2{}^*$ ={55,63}.

4 RESULTS

As mentioned at section 3.1, this problem originally has been studied by Barton and Meckesheimer (2006). They used a regression metamodel to optimize the total cost of communication network system to determine routing percentages (p_1, p_2). Their strategy is to sequentially explore local subregions of the experimental region and use line searches to find a new experimental subregion closer to the optimum. This strategy has applied to the problem in five iterations. A disadvantage of the method is that automated versions of the algorithm are not available. Global metamodel present an opportunity for optimization using a single metamodel, rather than a sequence of fitted local metamodels. Kriging as a global metamodel can be fitted once, based on a set of simulation runs from a global experiment design, and then the optimization can proceed iteratively using the same metamodel (Barton and Meckesheimer 2006).

The following is a summary of the results obtained from the kriging and regression metamodels to answer the question stated above. Based on the SO with regression and Kriging metamodels of the considered communication network system, the average cost are estimated to be 33.04 and 32.99, respectively as given in Table 1. This result shows that kriging finds the best routing percentages (p_1, p_2), minimizing overall system cost. Although there is no statistically significant difference between two results the confidence intervals of LOK metamodel strategy are more tighten than regression metamodel strategy. This means that LOK finds close solution that had been found by Barton and Meckesheimer (2006) using a regression metamodel for the communication network system. On the other hand LOK algorithm provides more quality solutions than regression from optimization aspect.

Table 1: SO results with kriging and regression metamodel

Metamodels	p_1	p_2	Cost	Confidential Intervals of Cost with 0.95
LOK	55	63	32.99	[32.68, 33.30]
Regression	54	63.8	33.04	[32.67, 33.41]

5 CONCLUSIONS

This paper provides step by step how we implement a kriging metamodel in SO. We have firstly developed a LHD trying to prevent high correlation among experiments keeping with the random structure of the design. Therefore, a LOK algorithm has been developed and applied to the simulation results. At neighbors search stage we successfully adopted directional search to run simulation model in direction with better solution.

As a result, we say that LOK metamodel can be used for SO using suitable variogram model. This paper provides step by step how we implement a LOK algorithm to built metamodel in SO. Future research may be concentrate on searching better solution with proposed LOK among solution neighbors.

References

Ankenman B., B.L. Nelson, and J. Staum, “Stochastic Kriging for simulation metamodeling”, *Proceedings of the 2008 Winter Simulation Conference*, 2008.

Barton R. R. and M. Meckesheimer, “Metamodel-based simulation optimization in Simulation”, *Handbooks in Operations Research and Management Science vol. 13.*, Amsterdam: Elsevier B.V.,2006.

Van Beers W. and J.P.C. Kleijnen, “Kriging for interpolation in random simulation”, *Journal of the Operational Research Society*, no. 54, pp. 255-262, 2003.

Van Beers W. and J.P.C. Kleijnen, “Kriging for Interpolation in Random Simulation”, *Technical report*, Department of Information Management, Tilburg University, Netherlands, 2002.

Van Beers W., J.P.C. Kleijnen, “Kriging Interpolation In Simulation: A Survey”, *Technical report*, Department of Information Management, Tilburg University, Netherlands, 2004.

Biles W. E., J. P. C. Kleijnen, W.C. M. van Beers, I. Van Nieuwenhuyse, “Kriging Metamodeling In Constrained Simulation Optimization: An Explorative Study”, *Proceedings of the 2007 Winter Simulation Conference*, 2007.

Carson Y., A. Maria, “Simulation Optimization: Methods and Applications”, *Proceedings of the 1997 Winter Simulation Conference*, 1997.

Cressie N.A.C., “The origins of kriging”, *Mathematical Geology*, 22, pp.239–252, 1990.

Cressie N.A.C., “Statistics for spatial data”, revised edition, New York, A Wiley-Interscience publication, 1993.

David M., "Geostatistical Ore Reserve Estimation", Amsterdam: Elsevier, 1977.
Dowd P.A., 1982, "Lognormal kriging—The general case", *Math. Geology*, v. 14, no. 5, p. 474-500
Fischer M. M., A. Getis, "Handbook of Applied Spatial Analysis", Springer-Verlag Berlin Heidelberg 2010.
Fu M.C., F.W. Glover, J. April, "Simulation optimization: A review, new developments, and applications", *Proceedings of the 2005 Winter Simulation Conference*, 2005.
Genton M. G., "Highly Robust Variogram Estimation", *Mathematical Geology*, Vol.30, No. 2, 1998.
Gilbert R.O. and J.C. Simpson, "kriging for Estimation Spatial Pattern of Contaminants: Potential and Problems", *Environmental Monitoring and Assessments*, Vol.5, pp.113-135, June 1985.
Isaaks E.H., and R.M. Srivastava, "An Introduction to Applied Geostatistics", Oxford University Press, Toronto. 1989.
Journel A., G., C. J. Huıjbregts, "Mining Geostatistics", Academic Press, London, 1978.
Joseph V.R., and Y. Hung, "Orthogonal-Maximin Latin Hypercube Designs", *Statistica Sinica* 18, pp.171- 186, 2008.
Kleijnen J.P.C., "Regression metamodels for generalizing simulation results", *IEEE Tranactions on systems, man and cybernetics*, vol. SMC-9, No.2, February 1979.
Kleijnen J.P.C., "Kriging metamodeling in simulation: a review", *Technical report*, Department of Information Management, Tilburg University, Netherlands, 2007.
Lark R.M., D.J Lapworth, "Quality measures for soil surveys by lognormal kriging", *Geoderma*, 173-17, pp 231-240, 2012.
Matheron, " Principles of Geostatistics", *Economic Geology*, 58, pp. 1246-1266, 1963.
Mc Kay M G..D., R. J. Beckman, and W.J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code", *Technometrics*, 21, 239– 245, 1979.
Mc Gratha D., C. Zhangb, O. T. Carton, "Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland", *Environmental Pollution* vol. 127, pp. 239–248, 2004.
Myers D.E., "On Variogram Estimation", *The frontiers of statistical scientific theory & industrial applications*,1991.
Paul R. and N. Cressie, "Lognormal Block Kriging for Contaminated Soil", *Europan Journal of Soil Science*, vol.62, pp.337-345, June 2011.
Roth C., "Is Lognormal Kriging Suitable for Local Estimation?", *Mathematical Geology*, Vol. 30, No. 8, 1998.
Santner T.J., B.J. Williams and W.I. Notz, "The Design and Analysis of Computer Experiments", Springer-Verlag, NY, USA, 2003.

Biography

Include author bio(s) of 200 words or less.

Muzaffer Balaban is a statistical expert at Turkish Statistical Institute. He is a candidate of PhD in Industrial Engineering at Başkent University. His research interests is kriging models in simulation optimization. His email address is balabanmuzaffer@gmail.com.

Berna Dengiz is the dean of Engineering at Baskent University. Her field of study is mainly simulation modeling and optimization of complex large sized systems besides heuristic optimization. She has received research funding for her collaborative studies from the NATO-B2 program, TUBITAK (The Scientific and Technical Research Council of Turkey), Government Planning Center of Turkey and National Science Foundation (NSF) of the USA. She has worked as visiting professor at the University of Pittsburgh and Auburn University. Her web page can be reached at http://www.baskent.edu.tr/~bdengiz . Dr. Dengiz's email address is bdengiz@baskent.edu.tr.