Optimization of Chemical Processes Using Surrogate Models Based on a Kriging Interpolation

Natalia Quirante,^{a*} Juan Javaloyes,^a Rubén Ruiz-Femenia,^a José A. Caballero.^a

^aInstitute of Chemical Processes Engineering, University of Alicante. P.O. 99, E-03080 Alicante, Spain.

natalia.quirante@ua.es

Abstract

Superstructure approaches are the solution to the difficult problem which involves the rigorous economic design of a distillation column. These methods require complex initialization procedures and they are hard to solve. For this reason, these methods have not been extensively used. In this work, we present a methodology for the rigorous optimization of chemical processes implemented on a commercial simulator using surrogate models based on a kriging interpolation.

Several examples were studied, but in this paper, we perform the optimization of a superstructure for a non-sharp separation to show the efficiency and effectiveness of the method. Noteworthy that it is possible to get surrogate models accurate enough with up to seven degrees of freedom.

Keywords: process optimization, design, kriging algorithm, modular simulators.

1. Introduction

Computationally efficient process models have been recently demanded in many engineering applications. When we want to optimize a model, the process model has to trade-off the model accuracy against computational efficiency. The problem is that most of those models have a modular structure to which users have limited internal access (they see a “grey box model”) and some of these models introduce numerical noise, so that its derivatives cannot be accurately estimated.

In addition, the rigorous design of distillation column sequences is a difficult problem in chemical process engineering because it includes simultaneous optimization of continuous decisions related to the operational conditions, and discrete decisions allied to the number of trays in each column section or the columns connectivity.

In this work we study the replacement of complex systems of distillation columns by surrogate models based on a kriging interpolation generated from rigorous models, including continuous and discrete variables, and use these metamodels to design the distillation columns. Thereby, we keep the rigor of simulations models, removing most of the numerical problems and increasing the reliability of the synthesis algorithms.

In this work we use simulation data to build models for individual parts of a large system. We focus on kriging metamodels because they are computationally efficient and they use relatively small sampling data. Kriging can be applied to replace a complete system or to substitute one or some of the components of the model.

2. Kriging interpolation

Kriging was developed by a South African mining engineer in his geostatistics Master Thesis (Krige, 1951). Its fitting consists of two parts: a polynomial expression and a deviation from that polynomial:

$$y(x) = f(x) + Z(x) \quad (1)$$

where $Z(x)$ is a stochastic Gaussian process that represents the uncertainty about the mean of $y(x)$ with expected value zero. The covariance for two points x_i and x_j is given by a scale factor σ^2 and by a spatial correlation function $R(x_i, x_j)$. The most common alternative for kriging models is using the extended exponential correlation (Sacks et al., 1989).

$$R(x_i, x_j) = \exp\left(-\sum_{l=1}^d \theta_l (x_{i,l} - x_{j,l})^{p_l}\right) \quad (2)$$

where $\theta_l \geq 0$ and $0 \leq p_l \leq 2$ are adjustable parameters. To estimate the values of the parameters, θ_l and p_l , we maximize the logarithm of the likelihood function of the obtained data y .

$$\log(L) = -\frac{n}{2} \ln(2\mu) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln(|R|) - \left(\frac{-(y - 1\mu)^T R^{-1} (y - 1\mu)}{2\sigma^2}\right) \quad (3)$$

where y is the vector of obtained responses ($n \times 1$), $\mathbf{1}$ is a vector of ones ($n \times 1$), n is the number of sampled points and σ^2 and μ are parameters.

Differentiating Eq.(3) with respect σ^2 and μ , and equating it to zero, we obtain the optimal values for σ^2 and μ .

$$\hat{\mu} = \frac{\mathbf{1}^T R^{-1} y}{\mathbf{1}^T R^{-1} \mathbf{1}} \quad (4)$$

$$\hat{\sigma}^2 = \frac{(y - \mathbf{1}\hat{\mu})^T R^{-1} (y - \mathbf{1}\hat{\mu})}{n} \quad (5)$$

To interpolate a new point x_{new} , we add the point (x_{new}, y_{new}) to the data and compute the augmented likelihood function keeping the parameters constant. The value for y_{new} will be calculated maximizing the augmented likelihood function. With the Eq.(6) is given the final predictor of the kriging method (Sasena, 2002).

$$\hat{y}(x_{new}) = \hat{\mu} + r^T R^{-1} (y - \mathbf{1}\hat{\mu}) \quad (6)$$

where r ($n \times 1$) is the vector of correlations $R(x_{new}, x_i)$ between the sample design points and the points to be correlated.

Kriging allow us to estimate the accuracy with the correlation of the errors, from the formula for the mean-squared error (Sacks et al., 1989).

$$s^2(x_{new}) = \hat{\sigma}^2 \left(1 - r^T R^{-1} r + \frac{(1 - r^T R^{-1} r)^2}{\mathbf{1}^T R^{-1} \mathbf{1}}\right) \quad (7)$$

3. Application of the kriging algorithm in distillation columns

The kriging algorithm implementation is as follows:

- a. First, N points are generated, separated enough to ensure that the noise generated by the simulator has no effect on. These points are sampled.
- b. Kriging metamodel is fitted with the simulation data.
- c. Validate the accuracy of the model using cross-validation (Jones et al., 1998).
- d. Substitute the actual model by the kriging metamodel.
- e. Perform the (MI)NLP optimization. If the accuracy of the model is good, we can finish.
- f. In case that the accuracy was not good, add the optimal point to the set of sampled points, and update and re-optimize kriging. If there is no improvement in two consecutive iterations, go to step g.
- g. In this case, we have to contract the feasible search region within a trust region around the solution obtained and repeat steps d to f, until we can guarantee that the error in gradient is below a given tolerance.

In this work, sampling points are generated distributed throughout the search area using a max-min approach (maximizing the minimum distance between two points). For this, we fix the bounds of all independent variables and then we distribute the rest of the points through the max-min approach.

4. Examples and results

Several examples were studied to verify the effectiveness of the method, but in this paper only an interesting example is presented. It consists of a superstructure for a non-sharp separation. We enumerate the column trays from the top to the bottom of the column. The example was solved on a computer with a 2.60 GHz Pentium® Dual-Core Processor and 4 GB of RAM under Windows 7.

4.1. Objective function

The objective function consists of minimizing the total annualized cost (TAC) of a chemical process. The total annualized cost is function of the capital costs of the column shells, trays and exchangers, and function of the operating cost (we have considered the cost of cooling water and vapor stream).

The objective function is determined by the Eq.(8) (Turton, 2012).

$$\min TAC \left(\frac{\$}{year} \right) = C_{op} + F \cdot C_{cap} \quad (8)$$

where C_{op} is the operating cost per year, F is an annualization factor and C_{cap} is the capital cost. Both are updated by the global CEPCI cost index of 2013. The annualization factor is calculated by the Eq.(9) (Smith, 2005).

$$F = \frac{i(1+i)^n}{(1+i)^n - 1} \quad (9)$$

where n is the horizon time and i is the fractional interest rate per year. We have used a horizon time of 5 years and a fixed interest rate of 10 %.

4.2. Distillation sequences. Nonsharp separations.

We want to separate a multicomponent feed stream given into several desired multicomponent product streams. Figure 1 shows the superstructure that contains all possible alternatives to get the desired products (Aggarwal and Floudas, 1990). The objective is to determine the optimal sequence and number of trays of the columns to minimizing the total annualized cost (TAC) for separating an iso-molar mixture of benzene, toluene and p-xylene. This cost is composed of the operational and fixed costs of both columns. Mixer and splitters costs are neglected. Peng-Robinson equation of state has been used.

Assume that the pressure is fixed ($P = 101.3$ kPa). The independent variables chosen for the rigorous design of the conventional distillation columns are the number of trays in the rectifying section, number of trays in the stripping section and the recoveries of key components for each column.

The problem is formulated as an MINLP. It can be solved by reformulating the disjunctions using a Big-M approach (Grossmann and Ruiz, 2012). Disjunctions are shown in Eq.(10).

$$\left[\begin{array}{l} Y_1 \\ [Q_{Reb}^{C1}, Q_{Cond}^{C1}, D^{C1}] = Kr_1(N_{R1}^{C1}, N_{S1}^{C1}, f_{LK}^{C1}, f_{HK}^{C1}, F5_i) \\ TAC_1 = \Omega(Q_{Reb}^{C1}, Q_{Cond}^{C1}, D^{C1}) \end{array} \right] \vee \left[\begin{array}{l} \neg Y_1 \\ TAC_1 = 0 \\ F5_i = F6_i = F9_i = 0 \forall i \end{array} \right] \quad (10)$$

$$\left[\begin{array}{l} Y_2 \\ [Q_{Reb}^{C2}, Q_{Cond}^{C2}, D^{C2}] = Kr_2(N_{R2}^{C2}, N_{S2}^{C2}, f_{LK}^{C2}, f_{HK}^{C2}, F13_i) \\ TAC_2 = \Omega(Q_{Reb}^{C2}, Q_{Cond}^{C2}, D^{C2}) \end{array} \right] \vee \left[\begin{array}{l} \neg Y_2 \\ TAC_2 = 0 \\ F13_i = F14_i = F18_i = 0 \forall i \end{array} \right]$$

Y_1 and Y_2 are Boolean variables that take the value of True if the corresponding column is selected and False otherwise. The symbol ' f ' makes reference to the key components recoveries.

The minimum objective function ($TAC = 0.3129$ M\$/year) is obtained with the sequence shown in Figure 2.

Figure 1. Superstructure for a three-component system.

Figure 2. Optimal solution for the three-component system.

The column 1 has 26 trays and it is fed on the 14th tray. The column 2 has 31 trays and it is fed on the 15th tray.

A summary of results obtained is shown in Table 1. A contraction step around the optimal solution was performed. Then, a new resampling and kriging calibration was needed.

5. Conclusions

In this work, was presented and applied an algorithm for constrained optimization of distillation columns. This optimization includes implicit grey box functions which are substituted by a kriging metamodel.

Table 1. Summary of results obtained (optimal solution).

Optimal valued (after refining stage)			
Number of sampled points in each columns = 202			
TAC = 0.3129 M\$/year			
Column 1		Column 2	
F5 _{benzene}	13.1490	F13 _{benzene}	3.6504
F5 _{toluene}	11.2260	F13 _{toluene}	9.5000
F5 _{p-Xylene}	3.4158	F13 _{p-Xylene}	12.6589
Number of trays in the rectifying section	13	Number of trays in the rectifying section	14
Number of trays in the stripping section	12	Number of trays in the stripping section	16
Recovery (LK = Benzene)	0.8878	Recovery (LK = Toluene)	0.8903
Recovery (HK = Toluene)	0.8513	Recovery (HK = p-Xylene)	0.8500

The “quality” of the optimum will depend on how accurately the surrogate model represents the actual model. It is interesting remark that though the final model includes discrete variables the kriging surrogate can be generated without taking into account that fact.

The implementation presented in this work only guarantees a local optimum. However, we have checked that it is possible to get accurate global surrogate models with up to seven degrees of freedom, being necessary a high number of sampling points. The problem presented is that if the number of points is large, it will be necessary a large CPU time for calibrating the kriging. A reduced set of sampling points reduce the CPU time for calibrating the kriging and interpolation, but it is likely that we need more contraction, resampling and recalibrating stages.

The algorithm presented has proved to be robust, reliable and also allows a fast interpolation of new values.

Acknowledgements

The authors wish to acknowledge the financial support by the Ministry of Economy and Competitiveness from Spain, under the project CTQ2012-37039-C02-02.

References

- Aggarwal, A., Floudas, C.A., 1990. Synthesis of general distillation sequences-nonsharp separations. *Computers and Chemical Engineering* 14, 631-653.
- Grossmann, I., Ruiz, J., 2012. Generalized Disjunctive Programming: A Framework for Formulation and Alternative Algorithms for MINLP Optimization, in: Lee, J., Leyffer, S. (Eds.), *Mixed Integer Nonlinear Programming*. Springer New York, pp. 93-115.
- Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization* 13, 455-492.
- Krige, D., 1951. A Statistical Approach to Some Mine Valuation and Allied Problems on the Witwatersrand., Masters Thesis, University of Witwatersrand, South Africa.
- Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and Analysis of Computer Experiments. 409-423.
- Sasena, M.J., 2002. Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations. University of Michigan.
- Smith, R., 2005. *Chemical process design and integration*. Wiley.
- Turton, R., 2012. *Analysis, Synthesis, and Design of Chemical Processes*. Prentice Hall.