NordDesign 2012
August 22 – 24, 2012
Aalborg, Denmark

Cost Optimization of Product Families using Analytic Cost Models

Thomas Ditlev Brunoe & Peter Nielsen
Department of Mechanical and Manufacturing Engineering
tdp@m-tech.aau.dk & peter@m-tech.aau.dk

Abstract
This paper presents a new method for analysing the cost structure of a mass customized product family. The method uses linear regression and backwards selection to reduce the complexity of a data set describing a number of historical product configurations and incurred costs. By reducing the data set, the configuration variables which best describe the variation in product costs are identified. The method is tested using data from a Danish manufacturing company and the results indicate that the method is able to identify the most critical configuration variables. The method can be applied in product family redesign projects focusing on cost reduction to identify which modules contribute the most to cost variation and should thus be optimized.

Keywords: product family optimization, cost drivers, cost analysis, solution space development

Introduction

Product cost is a very important factor for companies since it is determinant for the contribution margin, company profit, competitiveness and ultimately the survival in a particular market. For this reason it is necessary for most companies to continuously ensure that the cost of products is as low as possible without compromising the target functionality, performance and quality. To achieve this it is necessary to evaluate and optimize the product portfolio with regards to new products as well as existing ones. Typically, product development projects are concerned with updating existing products with new features or redesigning for manufacturability and thus reduced cost using design for manufacturing (DFM) methods. However, to focus DFM efforts efficiently it is useful to be able to identify which parts of a product contributes the most to the product cost, i.e. analyze the cost structure to specify which parts of a product should be redesigned.

In any company it is crucial to continuously evaluate the profitability of the product range, however in companies with a significant product variety such as mass customization or engineer to order companies, this is a challenging task. In companies offering customized products, manufacturing cost will depend on a particular configuration of a product and these companies often experience that it is not obvious which product properties drive cost. This makes it nontrivial to identify which products (or combinations of configurations) are profitable and which are not. This evaluation and the resulting development of a product portfolio in mass customizing companies is referred to as solution space development, which is one of three fundamental organizational capabilities which differentiate successful mass customizers from the non successful [10].

In mass customization, where it is unusual to sell and produce more than a few identical products but rather sell high numbers of individually customized products, it makes little sense to evaluate the cost structure of a single product. Instead the solution space must be evaluated as a whole. Evaluating the profitability of the solution space can be approached in several different ways that can be fundamentally separated into qualitative or quantitative approaches. Due to the vast complexity of a mass customization solution space in terms of the number of product features and modules, usually leaving a practically infinite solution space, a qualitative approach seems unfeasible indicating that a quantitative approach should be pursued.

Existing methods

A number of manufacturing process dependent methods have been developed for cost estimation, implying that the particular method of estimating cost can only be applied to certain processes e.g. casting or welding [3,4,11,12,12,16]. Cost estimation methods dependent on the specific product type also exist. In particular much research has been presented within the area of estimating cost during product development for both the finished product and the development process [2,6,7,14,18]. Kingsman & de Souza [5] introduced a general framework for cost estimation and pricing decisions, but no practical methods for estimating cost are presented. Other studies focus primarily on describing mathematically using synthetic models how customized products can be priced, however primarily compared to similar non-customized products [1,13].

One of the deficiencies of the approaches found in literature is that they do not take the large solution spaces into account but rather focus on a single product thus rendering them inapplicable for mass customized products. Furthermore most are synthetic, meaning that a cost and pricing model must be developed in order to evaluate the product profitability which is complicated by a high variety. Finally a number of the approaches described in literature are product specific rather than generic.

Elements of Product Cost for Customized Products

When analysing the manufacturing cost for a standard product, the direct costs are easily identified. Basically the manufacturing costs can be divided into three: 1) component cost, 2) assembly cost and 3) overhead cost [15]. In this paper we will not address overhead cost further. Identifying the component costs for a standard product is typically merely a matter of summing the cost of each component in the bill of materials and based on the cost information for the individual components the most expensive components can be identified as the most interesting for reducing product cost. The assembly cost can be analysed by observing an actual assembly process and measuring the time needed for each operation. However, for mass customized products, this analysis is more complex.

The complexity of analysing the cost structure of a mass customized product depends largely on the variation in product structure and whether the product consists of standard modules which are assembled according to configuration or individually customized components are also included in the product.

As for standard products, the manufacturing cost of a mass customized product will consist of the component costs and the assembly costs. For some mass customized products the manufacturing process will consist entirely of assembling standard modules. In this case, analysing and identifying the modules contributing the most to component costs is trivial since each module type will have a standard cost. However, in cases where the product is assembled partly or entirely from individually customized modules, the cost of each module will vary depending on configuration and it is thus not possible to draw general conclusions on cost distribution from these modules.

The assembly cost is largely labour cost and thus depends on the time required to assemble the product. In many cases different modules will require different times for assembling and the assembly costs will thus vary depending on the combination of modules included in a configuration. Furthermore, it cannot be assumed that the time to assemble a module with a partly assembled product will depend entirely on the type of the module. In some cases the assembly time will depend on both on the module being mounted and the module(s) interfacing to the module being mounted as illustrated in figure 1.

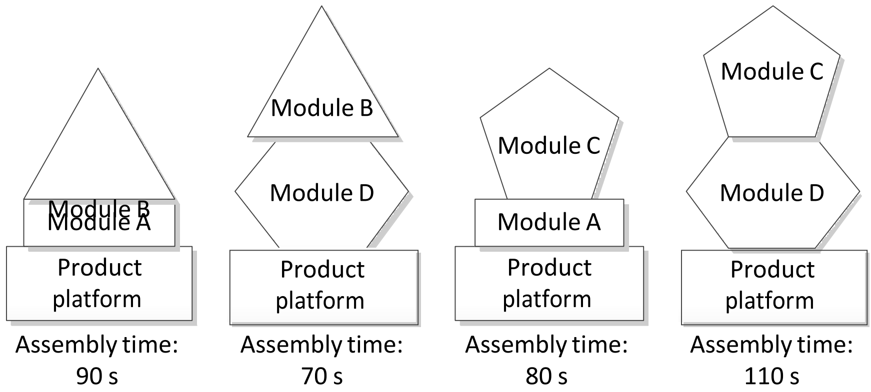

Figure 1. Example of product assembly times depending on combinations of modules

These issues lead to the fact that assembly times are to some extent unpredictable. Since mass customization product families often have a solution space containing billions of possible combinations it is infeasible to evaluate each combination of assembly times. However, identifying which modules or module combinations lead to particularly increased assembly cost is essential in redesigning a product family for reduced manufacturing cost.

Research objectives

When optimizing a product family with respect to manufacturing costs, it is essential to be able to identify the modules or configuration variables which are significant to manufacturing costs, since optimizing those modules will expectedly lead to the greatest cost savings. The objective of this paper is to investigate a method for automatically establishing a model based on historical data describing the relation between product variety and the variation in manufacturing costs. This model should be a generic model and thus cover all possible variety within a product family without having to analyze each possible combination.
The main research questions are:

- How well can a statistical method applied to historical configuration and cost data describe variations in product cost?
- How is such a method applied to optimize product designs in product development projects?

Methods

Since it is desired to establish a simple method for analyzing the cost structure, a simple linear model is proposed. Due to the complexity of the problem and the number of variables considered (large number of components and resources involved) multiple linear regression should be used [17]. However, since some manner relations can be expected between the independent variables all insignificant variables cannot just be removed in one step. One method is to test all alternative combinations of variables must be considered. In the case of 200 independent variables and desired model including 10 of these this would imply testing 8.15×10^{22} linear combinations, making this non practical to implement in real life settings.

Stepwise regression is a particular method for arriving at linear model and is applied when a reduced linear model is desired and some form correlation can exist between the (assumed to be) independent variables. Stepwise regression comes in two forms forward selection and backward elimination. Forward selection is based on stepwise adding the independent variable to the model that accounts for the largest amount of variation in the dependent variable. Backward elimination works in the opposite manner, by starting with a full linear model and stepwise removing the variable that has the least impact on the variation of the dependent variable. If the independent variables are in fact completely independent the model with the same number of variables found using forward selection and backward elimination will be the same [9]. However, when the independent variables are not completely independent the derived models may be different. Using forward selection implies the risk of adding a variable early on that later (due to the inclusion of other variables) becomes insignificant. Using backward elimination implies the risk of removing a variable that later (due to the elimination of other variables) actually becomes significant again [9]. In this particular paper the backward elimination method for stepwise regression is applied. Backward elimination is chosen over forward selection since in the case product configuration it seems better to risk dropping variables that explain a given behaviour (in this case of cost) than having an overly complex model with several variables explaining the same behaviour. However, a potential pitfall of using backward elimination is the need to be able to generate a full linear model, which implies the existence of *n+1* observations where *n* is the number of independent variables. So in cases with a limited number of observations available a forward selection approach may be more suitable with the risk of at arriving at more complex model to explain the same amount of variation in the dependent variable.

The principle of applying backward elimination to the problem is to estimate a (simple) linear model for historical configuration data and corresponding costs, thereby creating a linear model that can predict cost for new configurations [8]. The way backward elimination simplifies the linear model is by iteratively removing variables from a large set of data describing a number of historical product configurations and fitting a model to the reduced set of data. The output of the method is a linear model with few parameters, which can be used to estimate costs with the most significant configuration parameters as inputs [8].

The set of variables describing the configurations can contain different types of variables. Since the method reduces the number of variables significantly it is not necessary to qualitatively select the variables before the method is applied, but merely provide a gross list of variables describing each historical product and a corresponding incurred cost. When a product configuration system is used, a series of data files from this can be used as input to the model estimation process. The configuration variables can include scope options, performance, dimensioning variables, production parameters, feature options etc. Ideally the method takes all available product information. The only mandatory data is a cost for each configuration and more observations than initial parameters.

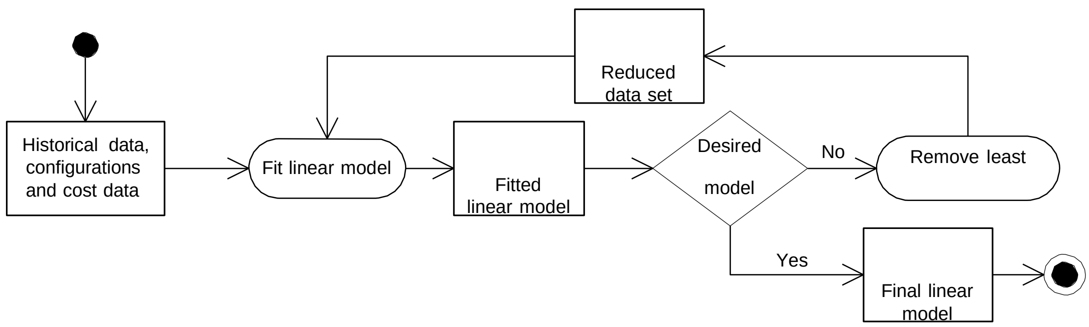

Figure 2. Activity diagram for the backwards selection method

In figure 2 the sequence of activities in the backward elimination method is presented [9]. To ensure that it is possible to fit a model some conditions must be met for the input data. First, all variables that are constant are removed (i.e. features of the product always included). Second, if two variables are perfectly correlated one is removed (i.e. two features that are always sold together), in principle it does not matter which is removed. All these activities (including the data cleansing) are implemented in a script executed in the statistical analysis tool R. The script has a loop that runs until a certain precondition is satisfied. In the first iteration, a linear model is fitted to the full data set, and the least significant configuration variable is removed. In the next iteration, the model is fitted to the reduced data set, and continues until the precondition is satisfied. In this test, the remaining number of variables included in the cost estimation model was used as precondition so that the loop ended when e.g. 5 variables remained. The finished model now consists of the selection of variables most significant for the cost and the linear model to calculate a cost from the values of these variables. In the application any number of termination conditions can be used. Typically residual sum of squares (R^2 or modified R^2), i.e. the amount of variation explained by the model, is used. However, it may be difficult to translate this measure into a business term. For this reason it is chosen to instead use MAPE (Mean Absolute Percentage Error) as evaluation criteria in this paper.

Results

The proposed method has been tested on actual data from a manufacturing company. The case concerns a medium sized company in Denmark producing technical products for domestic water installations which are configured within a predefined solution space. The products share a common structure and are described using typically around 50 variables. The company produces products configured with in a given set of fixed options. The company registers material and salary costs as well as net margin for all sold products. Furthermore a full route and Bill of material is available for all sold configurations. The data uses for this analysis contains about 200 configurations.

The method was applied to the data with the criteria of reducing the data set to the most significant 5, 10, 15 and 20 variables. The reason for reducing the model to different sizes was to analyse how well very simple models could explain variation in manufacturing cost. The results are shown in table 1, where each row represents a separate model with a given complexity, 5, 10, 15 or 20 variables, with the corresponding R^2 and MAPE values. For comparison, the MAPE value of the full model including all configuration variables before the backward elimination was applied was 3.42%.

Table 1 Results from applying backwards selection to configuration data

Variables	R^2	MAPE	Reduction of complexity
5	0.6561	4.56%	90%
10	0.7051	4.08%	80%
15	0.7379	3.58%	70%
20	0.7454	3.42%	60%

Interpretation of results

Based on the MAPE values included in table 1, it seems that a model with 20 variables, i.e. a 60% reduction in configuration variables and thereby model complexity describes the configuration with the same accuracy as the full model, since the MAPE value is the same for the full model and the 20 variable model. This clearly indicates that the remaining 30 variables which were removed do not contribute to the variation in manufacturing cost.

However, perhaps more interesting, going from 20 to 5 variables, i.e. a 75% reduction in complexity only increses the MAPE by 1.14, indicating that those 5 variables by far are those which contribute the most to variation in manufacturing cost.
The variables identified from the analysis was subsequently analysed qualitatively by a product expert from the comapny, who could confirm that those variables for different reason could contribute significantly to variation in cost. All identified configuration variables corresponded to the selection of certain components in the product, indicating that those particular components are significant cost drivers which should be addressed in cost optimization of that product family.
Althoug the results indicate that certain configuration variables, which in this case corresponded to specific components, are major cost drivers, this does not clearly indicate that this cost can in fact be reduced. However it does give an indication of where to focus a qualitative analysis of cost reduction potential.

Applications

Based on the test results, the authors believe that the method proposed can be applied successfully in a number of different ways. When redesigning a product family, the method can be applied to identify cost drivers with respect to the variation in manufacturing costs. This also means that the method does not identify cost driver for manufacturing costs which are static across different configurations. In practice, this means that modules or components which are a part of the product platform and thus always included in the products, are not identified using this method. However, as indicated previously in this paper, platform elements in a product should be addressed using traditional design for manufacturing methods, since variation does not add complexity to the process.
It is important to bear in mind that this method cannot be applied without being complemented by a qualitative analysis. However, this method suggests a number of areas to focus the qualitative analysis thus reducing the effort necessary to identify actual saving potentials. The need for a complementary qualitative analysis is emphasized by the circumstance that for some configuration systems, configuration variables are not representing module choices but rather product functionality or performance requirements, which are eventually translated to a bill of material. In this case, product experts will need to translate the “functional” configuration variables to specific components to achieve the desired result; a gross list of components which seem to be cost drivers.
The method as it was applied in this paper uses the total manufacturing costs as the dependent variable since it is desired to identify both drivers for component and assembly cost. However, if the data is readily available, the method could be applied using only the assembly cost to identify components which are particularly time consuming in the assembly process. In the same manner, the method could be applied to using the component costs as the dependent variable identifying certain components which are driving the component fabrication cost. The latter is only relevant in cases where customized components can be included in a configuration. Generally, the method can be applied to any dependent variable which can be related to specific configurations to serve other purposes, e.g. sales price, lead times, contribution margin, rework rates, customer return rates etc.

Limitations

Apart from the limitations indicated above requiring a supplementary qualitative analysis, there are a number of other limitations due to the fact that the method is based on historical data. In order to apply the method, a sufficient data basis must be available. In practice this means that there must be a higher number of observations than configuration variables. Obviously, the higher number of configurations, the higher validity of the conclusions can be

achieved. Since the methods needs historical data, it can only be applied to existing product families and thus not DFM efforts regarding new product families.
In products where raw material prices have a high variation over the period of the registered configurations, those variations may contribute more to the variation in manufacturing cost than the configuration variables, which the method in its current form does not take into account.
Finally, the requirement for a high data quality is vital for this method to produce valid results. All relevant manufacturing costs must be registered related to the correct configuration on the specific order as well as changes in the configuration over time must be registered to reflect what has actually been produced.

Conclusion

From the results presented it is concluded that the method is able to produce a cost model which with high accuracy can identify the configuration variables which describe the variation in manufacturing costs and can thus be used for identifying cost drivers which can be used to reduce product costs. This is useful for companies manufacturing customized products since it is difficult to identify the relations between product options and manufacturing costs. The method is robust towards different types of data input and can thus be applied to a data set without prior knowledge of the meaning of the configuration variables. Furthermore, the method can also be applied to identify relations between product options and other variables such as contribution margin, lead times, rework rates etc. in order to optimize a product family with regard to other goals.
Further research into this area will include applying the method to data sets for different product types in other industries as well as other types of dependent variables to meet different optimization criteria.

References

[1] Alptekinoglu, A., & Corbett, C. J.: "Mass Customization Vs. Mass Production: Variety and Price Competition", Manufacturing Service Oper.Management, Vol. 10 , pp 204–217, 2008

[2] Ben-Arieh, D., & Qian, L.: "Activity-Based Cost Management for Design and Development Stage", International Journal of Production Economics, Vol. 83 , pp 169-183, 2003

[3] García-Crespo, Á., Ruiz-Mezcua, B., López-Cuadrado, J. L. et al.: "A Review of Conventional and Knowledge Based Systems for Machining Price Quotation", Journal of Intelligent Manufacturing, pp 1-19,

[4] H'mida, F., Martin, P., Vernadat, F.: "Cost Estimation in Mechanical Production: The Cost Entity Approach Applied to Integrated Product Engineering", International Journal of Production Economics, Vol. 103 , pp 17-35, 2006

[5] Kingsman, B. G., & De Souza, A. A.: "A Knowledge-Based Decision Support System for Cost Estimation and Pricing Decisions in Versatile Manufacturing Companies", International Journal of Production Economics, Vol. 53 , pp 119-139, 1997

[6] Layer, A.: "Recent and Future Trends in Cost Estimation", International Journal of Computer Integrated Manufacturing, Vol. 15 , pp 499-510, 2002

[7] Niazi, A., Dai, J. S., Balabani, S. et al.: "Product Cost Estimation: Technique Classification and Methodology Review", Journal of manufacturing science and engineering, Vol. 128 , pp 563, 2006

[8] Nielsen, P., & Petersen, T. D.: "Criticality of Components and Specific Configurations in an Engineer-to-Order Environment", Proceedings of the 14th International Annual EurOMA Conference, 2007

[9] Rawlings, J. O., Pantula, S. G., Dickey, D. A. et al. "Applied regression analysis: A research tool" . Springer New York, NY, US, 1998

[10] Salvador, F., de Holan, M., Piller, F.: "Cracking the Code of Mass Customization", MIT Sloan Management Review, Vol. 50 , pp 70-79, 2009

[11] Shehab, E., & Abdalla, H.: "An Intelligent Knowledge-Based System for Product Cost Modelling", The International Journal of Advanced Manufacturing Technology, Vol. 19 , pp 49-65, 2002

[12] Shtub, A., & Versano, R.: "Estimating the Cost of Steel Pipe Bending, a Comparison between Neural Networks and Regression Analysis", International Journal of Production Economics, Vol. 62 , pp 201-207, 1999

[13] Syam, N. B., & Kumar, N.: "On Customized Goods, Standard Goods, and Competition", Marketing Science, pp 525-537, 2006

[14] Tornberg, K., Jamsen, M., Paranko, J.: "Activity-Based Costing and Process Modeling for Cost-Conscious Product Design: A Case Study in a Manufacturing Company", International Journal of Production Economics, Vol. 79 , pp 75-82, 2002

[15] Ulrich, K. T., & Eppinger, S. D. "Product design and development" . McGraw-Hill, New York, 2004

[16] Verlinden, B., Duflou, J., Collin, P. et al.: "Cost Estimation for Sheet Metal Parts using Multiple Regression and Artificial Neural Networks: A Case Study", International Journal of Production Economics, Vol. 111 , pp 484-492, 2008

[17] Walpole, R. E., Myers, R. H., Myers, S. L. et al. "Probability and statistics for engineers and scientists" . Macmillan New York, 1989

[18] Weustink, I., ten Brinke, E., Streppel, A. et al.: "A Generic Framework for Cost Estimation and Cost Control in Product Design", Journal of Materials Processing Tech., Vol. 103 , pp 141-148, 2000