Combining the Cost Drivers Based on the Singular Value Decomposition of Matrix

Xudanyang Li and Buxi Li

***Abstract*—Reducing the complexity of activity-based costing system is one of the key factor when successfully implement the activity-based costing system. Based on dimensionality reduction of the singular value decomposition (SVD) of the matrix, this paper proposes and researches the method of SVD-based combination of cost drivers under activity-based costing. Through singular value's decomposition for coefficient matrix in the costing model under activity-based costing, this paper classifies the cost drivers based on singular values and eigenvectors in the right singular matrix. On this basis, using integer programming, representative cost drivers are selected for combination to simplify the activity-based costing model. Numerical examples indicate that the method of SVD-based combination of cost drivers reduces the complexity of activity-based costing system and significantly improves the accuracy of the combined product cost. The result shows, whether reducing model complexity or ensuring product cost accuracy, the method of SVD-based combination of cost drivers is superior to the existing other methods, such as integer programming method, clustering method, etc.**

***Index Terms*—Activity-based costing, product costing model, combination of cost drivers, SVD of matrix.**

I. Introduction

Since the generation of Activity-Based Costing (ABC) in the 1980s, it has received widespread attention because of its ability to more accurately allocate overhead to costing objects such as products, special orders, and customers (Cooper, Kaplan, 1988; Dopuch, 1993; Gupta and Galloway, 2003; Velmurugan, 2010; Tsar *et al*., 2014; Almeida & Cunha, 2017) [1]-[7]. ABC is also gradually applied from the manufacturing industry to service industries such as medical and logistics, and non-profit organizations such as government and universities. ABC's research has also shifted from cost accounting to cost control, production management and so on. When using ABC to calculate costs of product, the costing procedures divided into two stages according to production process activities. The first stage is from resource to activities. In this stage, according to the causal relationship of resource costs incurring, the resource costs incurred in the production process are collected into the activity cost pools, and the activity costs of the activities are formed. The second stage is from activities to cost objects such as product. In this stage, according to the causal relationship of the activity costs incurring, the activity costs are allocated to product, etc., the product cost and other costs are formed. Whether it is the causal relationship of resource costs incurring or activities costs incurring, they are collectively referred to cost drivers. With the diversification of customer needs and increased complexity of production processes, the number of activities involved in the activity-based costing is increasing, resulting to the greatly increased implementing cost when collecting, processing and handling the activity information, so the costing system are becoming more and more complex. This not only adds the burden of activity-based costing system, but also leads to excessive implementation costs (Kaplan & Anderson, 2004) [8], which hinder the application of ABC. Therefore, how to reduce the number of activities, complexity of the activity-based costing system and implementation costs have become the topics of academic research.

II. Literature Review

In order to reduce the complexity of activity-based costing system, many scholars have carried out some useful explorations. Babad & Balachandran (1993) firstly studies the effect of cost drivers' combination on product cost accuracy [9]. Their research shows that the combination of completely correlation cost drivers doesn't compromise the accuracy of product costs. At the same time, they further study the cost drivers' combination by constructing a planning model that balances the information cost savings with the accuracy loss. The results show that the cost drivers' combination will lose the accuracy of product costs. However, their researches are only applicable to the combination of two cost drivers and consider the importance of product in the form of weights in the planning model. Based on the cost drivers and their causal relationships, Schierderjans & Garvin (1997) uses the Analytic Hierarchy Process (AHP) and Zero One Goal Programming (ZOGP) to study the selection of representative cost driver from a set of candidate cost drivers. Their results show that the AHP method is conform to the realistic selection process of cost drivers [10]. Based on Babad & Balachandran's research, Homburg (2001) constructs a combined alternative model where a cost driver is replaced by a set of cost drivers, proving the rationality of multi-cost drivers' combination [11]. Zhao & Ouyang (2009) improves Homburg's model by quantifying the importance of cost drivers and integrating them into the model, therefore achieving the balance of cost drivers' importance and relevance [12]. Li and Wang (2009) note that, although the above optimization model minimizes

Manuscript received March 11, 2019; revised December 12, 2019.

Xudanyang Li is with the Business School, University of Newcastle, Newcastle, CO 2308 Australia. He is on leave in Taiyuan City, Shanxi Province, China (e-mail: 316372139@ qq.com).

Buxi Li is with the School of Economics and Management, Shanxi University, Taiyuan, CO 030006 China (e-mail: buxili@126.com).

doi: 10.18178/joebm.2020.8.1.607

the loss of accuracy directly or indirectly, the minimum accuracy loss may still be very large. They establish an optimization model that maximizes the number of combined cost drivers and has limited error between product costs before and after combination of cost drivers at the same time, thus ensuring the controllability of the product cost costs [13].

Based on the fact that the combination of cost drivers with perfect positive correlation does not change the accuracy of product cost, Wang *et al*. (1999) propose the homogeneity concept of amount of cost drivers, and conclude that “satisfactory accuracy of product cost can be achieved by combining cost drivers with positive and negative correlation” [14]. Li *et al*. (2005) use matrix theory to study the theory of combining cost drivers, and extend the homogeneity concept of the number of cost drivers to the situations that the cost driver ratio of an activity is equal or proportional to linear combination of the cost driver ratio of several activities [15]. Afterwards, Li and Wang (2007) research the theory of the multiple cost drivers’ combination, and further explore the mechanism of “satisfactory accuracy of product cost can be achieved by combining cost drivers with positive and negative correlation”, therefore proposing the concept of the weighted average cost driver [16]. Using above conclusion, Wang *et al*. (2009) proposed a clustering analysis method for cost driver selection and combination [17]. Liu *et al*. (2014) proposed a method of combining cluster analysis with principal component analysis [18]. However, whether it is the clustering method or the combination of clustering analysis and principal component analysis, it is difficult to overcome the problem of uncontrollable cost error by using these methods to combine cost drivers.

In summary, the existing cost driver selection and combination research has proposed various methods to reduce the complexity of activity-based cost system, also consider the accuracy of product costs, but it is difficult to reducing the complexity and ensuring the accuracy of product cost at the same time. In this paper, based on the dimension reduction and denoising roles of singular value decomposition (SVD) [19], it further researches the combination of cost drivers under activity-based costing.

III. Product Costing Model Under ABC and Combination of Cost Drivers

For convenient analysis, product costing model under ABC is established firstly.

Assume there are n kinds of products and m activities. Each activity corresponds to a cost driver. R_{ij} represents the cost driver ratio of activity j consumed by product i to activity j consumed by all products, and is known as coefficient on cost driver j of product i. Obviously, cost driver coefficient R_{ij} satisfy:

$$0 < R_{ij} < 1,\ \sum_{i=1}^{n} R_{ij} = 1,\ j=1,2,\cdots,m.$$

And given S_j ($j=1,2,\cdots,m$) represents resource costs or total activity costs corresponded to cost driver j, C_i ($i=1,2,\cdots,n$) represents total activity cost of product i. The product’s indirect costing model (hereinafter referred as the costing model) under the activity-based costing is：

$$C_i = \sum_{k=1}^{m} S_k R_{ik},\quad i=1,2,\cdots,n.$$

Expressed in matrix form as：

$$C = RS \tag{1}$$

where

$$C=\begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix},\ R=\begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ R_{n1} & R_{n2} & \cdots & R_{nm} \end{pmatrix},\ S=\begin{pmatrix} S_1 \\ S_2 \\ \cdots \\ S_m \end{pmatrix}$$

Matrix R is referred as the coefficient matrix of product cost drivers，it reflects the structural proportion of the cost driver amount consumed by each product in the total amount of different cost drivers. The number of rows of the matrix reflects the number of products, and the number of columns of the matrix reflects the number of activities in the production process. In the coefficient matrix R of product activity driver, the column vector $(R_{1j}\ \ R_{2j}\ \ \cdots\ \ R_{nj})^T$ represents the cost driver ratio of activity j consumed by various products, the row vector $(R_{i1}\ \ R_{i2}\ \ \cdots\ \ R_{im})$ reflects the cost driver ratios of various activities consumed by product i . According to the ratios, the total cost of the product i is calculated. Therefore, the product costing model clearly reflects the forming process of cost from activities to product cost.

When the column number in coefficient matrix (the number of cost drivers) is large, the product costing model becomes very complicated. At the same time, in order to calculate the cost of all products, it is necessary to collect a large amount of relevant information, and the cost of implementing the activity-based costing system is high (Kaplan & Anderson, 2004). Therefore, simplifying the activity-based costing system has become the key to the successful implementation of the activity-based costing method.

How to simplify the activity-based costing model? The common method is to classify and combine the activities or cost drivers represented by the column vector of the coefficient matrix in the costing model (1), which is usually known as cost drivers’ combination. It means the activity cost (or resource cost) which is collected by one or more activities will be combined into another activity cost (or resource cost), and then aggregated cost will be assigned to all products in corresponding cost driver proportion of this another activity, thereby the products cost are calculated. First, before combining the activity cost (or resource cost) of multiple activities, the cost drivers are classified according to certain characteristics. Only the activities that are classified into the same category can be combined. Second, in each category of the cost driver, a cost driver needs to be selected as a representative cost driver. If the representative cost drivers chosen are different, the final calculated cost of each product may different. The above classifying cost drivers and selecting representative cost driver will affect the accuracy of the final calculated product cost.

IV. SVD of Coefficient Matrix and Combination of Cost Drivers

Essentially, the simplification of the costing model under activity-based cost is to simplify the column vector of the coefficient matrix R in the model. The combination of cost drivers is the process of classifying the column vectors of the coefficient matrix R and selecting a column vector in each category to reconstruct the coefficient matrix, which is also the dimension reduction process of the matrix. Therefore, it is necessary to research the characteristics of column vectors in the coefficient matrix. The significance of singular value decomposition of a matrix is that any form of matrix can be decomposed into the product of several simpler matrices. These simple matrices reflect the characteristics of different aspects of the original matrix.

According to the principle of the SVD of the matrix, the $n \times m$ coefficient matrix R in ABC costing model is decomposed as

$$R = U\begin{pmatrix} D & O \\ O & O \end{pmatrix}_{n\times m} V \tag{2}$$

where, U is a $n \times n$ matrix and referred as left singular matrix, it reflects the important characteristics of the product. V is a $m \times m$ matrix and usually known as right singular matrix, it reflects the characteristics of cost drivers. U and V are orthogonal matrices satisfy: $U^T U = E, V^T V = E$. D is a $k \times k$ real diagonal matrix in which the elements on the main diagonal line $d_{ii} = \sigma_i, i = 1, 2, \cdots, k$, satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$ and the other elements are 0，k is the rank of coefficient matrix. In the block matrix, O represents the zero matrix of different order.

It's easy to prove that

$$R^T R = V\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix}_{m\times m} V^T,\ RR^T = U\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix}_{n\times n} U^T$$

It can be seen that V is composed of eigenvectors of matrix $R^T R$, U is composed of eigenvectors of matrix RR^T, $\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2$ are k non-zero eigenvalues the eigenvector matrix V or U.

Because D is a diagonal matrix, singular value decomposition is equivalent to decomposing a $n \times m$ matrix which rank is k to weighted sum of k $n \times m$-order matrices which rank is 1. Where each of the matrices is product of eigenvector u_i and eigenvector v_i, the weight value is the singular value σ_i of the matrix R. Therefore

$$R = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T \tag{3}$$

In formula, u_i and v_i is the i-th column vector of matrix U and V respectively.

In singular value decomposition of matrix, singular values often correspond to important information in the matrix, and the importance of information is positively correlated with the size of singular values. However, those small singular values are usually considered to be caused by random interferences (noise). When these small singular values are set to 0, we can eliminate the interference of random factors. Suppose that the first *s* larger singular values are retained and the other smaller singular values are discarded. It is obtained that

$$R \approx \sum_{i=1}^{s} \sigma_i u_i v_i = U_{n\times s} D_{s\times s} V_{s\times m}^T \tag{4}$$

where $D_{s\times s} = diag(\sigma_1, \sigma_2, \cdots, \sigma_s)$. Therefore

$$RV_{s\times m}^T \approx U_{n\times s} D_{s\times s}$$

It can be seen that order of coefficient matrix R change from $n \times m$ to $n \times s$ because of matrix $V_{m\times s}$, that is, the number of columns of the coefficient matrix is compressed. This can be understood as combination of the cost drivers corresponded to the column vectors. Therefore, the matrix $V_{m\times s}$ consisting of eigenvectors corresponded to larger singular values not only reflects the important features of column vectors in the coefficient matrix R (the cost driver vector), but also eliminates the effect of random factors on each cost drivers. If $V_{m\times s}$ is mapped to *s*-dimensional space, then all row vectors of $V_{m\times s}$ are regarded as m points in *s*-dimensional space. Clustering these *m* points is essentially a classification of the corresponding *m* column vectors (cost driver vectors) in the coefficient matrix R.

Based on the above analysis, the cost drivers' classification and combination steps are as follows:

The step 1: perform singular value decomposition for the coefficient matrix R, obtain a left singular matrix *U*, a right singular matrix *V* and a singular value $\sigma_i, i = 1, 2, \cdots k$.

The step 2: set the smaller singular value to 0 in formula (3), approximate the coefficient matrix.

The step 3: in the matrix *V*, delete the eigenvector corresponded to the smaller singular value and retain the eigenvector corresponded to the larger singular value, then obtain a new matrix $V_{m\times s}$.

The step 4: use the clustering analysis method to classify the row vectors in $V_{m\times s}$ (or the column vectors in $V_{s\times m}^T$).

The step 5: in each type of cost drivers, use the method of the integer programming, select the representative cost drivers which minimize the error degree of all product costs after combination of cost drivers. The error degree (indicated by e) and accuracy (indicated by a) of all product costs are defined as follows (see Wang *et al*., 2009):

$$e = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\varepsilon_i}{C_i}\right)^2}, \quad a \triangleq 1 - e$$

where, ε_i is equal to the cost of product *k* after combination minus cost of product *k* before combination, C_i indicates the cost of product *k* before combination.

The step 6: in the same category, incorporate the activity costs which corresponded to other cost drivers into activity cost which corresponded to representative cost driver. Then allocate aggregated costs to products with the ratio of a representative cost driver and calculate the cost of each product. Finally, complete combination.

V. Numerical Examples

This paper uses the data from the XAMC case used by Jin *et al*. (2005): 6 products are denoted as P_1, P_2, P_3, P_4, P_5, P_6 [20]. The corresponding 14 cost drivers are denoted as j_1, j_2, … , j_{14}. Through calculating the original data, the cost driver ratio of different activities consumed by various products (cost driver coefficient) and the activity cost of each activity are shown in Table I as follow.

TABLE I: Cost Driver Coefficients and Activity Costs

Cost driver	j_1	j_2	j_3	j_4	j_5	j_6	j_7	j_8	j_9	j_{10}	j_{11}	j_{12}	j_{13}	j_{14}
P_1	0.291181	0.555461	0.302548	0.275090	0.267588	0.306452	0.563904	0.387454	0.420958	0.315136	0.290909	0.545455	0.390979	0.337500
P_2	0.299501	0.333248	0.506597	0.052867	0.407286	0.275161	0.304162	0.193727	0.353759	0.147716	0.200000	0.168831	0.302062	0.275000
P_3	0.266223	0.085620	0.075068	0.333333	0.140452	0.306452	0.118122	0.156827	0.105046	0.057545	0.118182	0.140693	0.108247	0.187500
P_4	0.116473	0.020066	0.040378	0.159498	0.058417	0.051613	0.004052	0.129151	0.038877	0.173512	0.090909	0.080087	0.065979	0.100000
P_5	0.026622	0.005605	0.075409	0.179211	0.112186	0.040968	0.005709	0.078413	0.065654	0.238809	0.109091	0.045455	0.085052	0.062500
P_6	0.000000	0.000000	0.000000	0.000000	0.014070	0.019355	0.004052	0.054428	0.015705	0.067282	0.190909	0.019481	0.047680	0.037500
Indirect cost	60701.00	210441.60	54686.24	38446.20	65351.60	3007.00	402906.00	13506.64	72242.40	174010.10	48270.20	27535.20	31622.00	305923.20

A. Singular Value Decomposition for Coefficient Matrix

Using above data, the non-zero singular value, left singular matrix and the right singular matrix are calculated in Table II, Table III and Table IV as follows:

TABLE II: Singular Value

σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
1.90497	0.45198	0.38051	0.29246	0.15761	0.07897

TABLE III: Left Singular Matrix *U*

u_1	u_2	u_3	u_4	u_5	u_6
0.752894	-0.070861	0.650929	0.037224	-0.041972	0.035699
0.548452	0.613212	-0.558569	-0.081399	0.011506	-0.066438
0.295996	-0.569663	-0.437998	0.586003	0.147635	0.175591
0.144801	-0.399224	-0.156543	-0.297257	-0.126840	-0.831082
0.142757	-0.348552	-0.218944	-0.636401	-0.370377	0.517697
0.058234	-0.116467	-0.002796	-0.393983	0.907217	0.069078

TABLE IV: Right Singular Matrix

v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}
0.2535	0.0983	0.3112	-0.3108	0.0374	-0.5796	0.0284	0.1181	-0.3533	-0.3034	-0.2906	-0.0609	0.2567	-0.0976
0.3307	-0.2351	-0.3510	-0.1169	-0.0727	-0.0133	-0.3359	0.4555	0.3001	0.1711	-0.1049	-0.3137	0.3408	0.1669
0.2858	-0.4514	0.3725	0.1572	-0.1830	-0.0531	0.0428	-0.5347	0.0790	0.1915	-0.0783	-0.4119	-0.0437	0.0800
0.2013	0.6706	0.1594	-0.1361	-0.3067	0.3173	-0.3269	-0.1262	-0.1527	0.1687	0.1073	-0.2723	0.0539	-0.0893
0.2581	-0.1919	0.3905	0.1203	-0.1396	0.2236	0.3091	0.5718	-0.1427	0.0508	0.3800	-0.0044	-0.0840	-0.2514
0.2555	0.1431	0.2774	-0.4088	0.1991	0.3308	0.2925	-0.0220	0.5276	-0.1744	-0.2179	0.1542	-0.0456	0.2301
0.3297	-0.1664	-0.3772	-0.2018	-0.0107	0.2600	0.2476	-0.3504	-0.2624	-0.0397	0.1929	0.2507	0.4671	-0.2005
0.2506	0.1841	-0.0993	0.0656	0.0830	-0.4366	-0.0419	-0.1503	0.4638	-0.2157	0.5907	-0.0399	-0.0832	-0.2193
0.2929	-0.1925	-0.0260	0.0379	-0.0831	0.1612	-0.4334	-0.0180	-0.2487	-0.5296	0.1330	0.1903	-0.3323	0.3839
0.2092	0.2763	-0.0467	0.6724	-0.3328	-0.0555	0.2422	0.0373	0.1169	-0.1561	-0.2446	0.1979	0.2644	0.2117
0.2118	0.1368	0.0335	0.3688	0.8172	0.1514	-0.0424	0.0001	-0.1878	0.0205	-0.0043	-0.2451	0.0928	0.0481
0.2961	0.1446	-0.4641	-0.0978	-0.0603	-0.1104	0.4251	0.0435	-0.1860	0.1040	-0.1616	-0.2660	-0.5677	0.0595
0.2712	-0.0759	-0.0244	0.1338	0.0408	0.0682	-0.3075	-0.0175	0.1382	0.0713	-0.4334	0.3050	-0.2510	-0.6581
0.2551	0.0623	0.1195	-0.0540	0.0943	-0.2717	-0.0998	-0.0073	-0.0922	0.6412	0.1318	0.5163	-0.0685	0.3358

B. Cost Driver Classification

Based on SVD principle, the column vector v_1 of the right singular matrix corresponded to the largest singular value σ_1 reflects the most important feature of the column vector (cost drivers) in the coefficient matrix. Therefore, this paper regards v_1 as 14 points in the one-dimensional space (cost drivers) and uses cluster analysis method to classify, thus obtaining categories of all 14 cost driver as Fig. 1 as following.

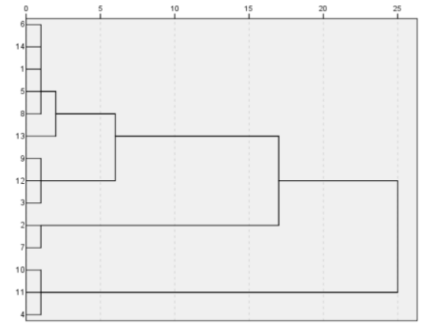

Fig. 1. Cluster tree of 14 cost drivers.

It can be seen that all 14 cost drivers are divided into 5 categories at most, and these 5 categories can be further clustered into 2 categories at least. In the 5 categories, only the cost driver j_{13} is clustered as a separate category, and the other 4 categories contain at least two cost drivers, and the representative cost drivers are selected in each category, as shown in Table V. The bold words in the Table IV indicate selected representative cost drivers.

TABLE V: COST DRIVER CATEGORIES AND SELECTED REPRESENTATIVE COST DRIVERS

Number of Categories	Categories
5	{ j_4, j_{10}, j_{11}}, { j_2, j_7}, { j_3, j_9, j_{12}}, { $j_1, j_5, j_6, j_8, j_{14}$ }, { j_{13}}
4	{ j_4, j_{10}, j_{11}}, { j_2, j_7 }, { j_3, j_9, j_{12}}, { $j_1, j_5, j_6, j_8, j_{14}, j_{13}$}
3	{ j_4, j_{10}, j_{11}}, { j_2, j_7}, { $j_3, j_9, j_{12}, j_1, j_5, j_6, j_8, j_{14}, j_{13}$}
2	{ j_4, j_{10}, j_{11}}, { $j_2, j_7, j_3, j_9, j_{12}, j_1, j_5, j_6, j_8, j_{14}, j_{13}$}

C. Error between Product Costs before and after Combination of Cost Drivers

According to the different number of categories of 14 cost drivers, the different categories and the representative cost drivers in each category, the calculated pre-combination product costs, and the combined product costs, the errors between them, the error degree and accuracy of the total product cost are shown in Table VI. In Table VI, *l* denotes number of categories, *e* and *a* respectively represents the error degree and accuracy of the total product costs before and after the combination.

It can be seen that the eigenvector corresponded to the maximum singular value 1.90497 reflects the most important feature of cost driver (column vector) in the coefficient matrix. According to this eigenvector, the all cost drivers are clustered into five categories. After selecting a representative cost driver for each category of cost drivers that contain multiple cost drivers and combining other activity costs into activity cost corresponded to a representative cost driver in this category, the product costing model is reduced from the original 14 cost drivers to 5 cost drivers, and the complexity of the costing model is greatly reduced. Further clustering, the number of categories of cost drivers can be clustered into 2 categories, and the complexity of the costing model will be further reduced.

TABLE VI: THE RESULTS OF COST DRIVER COMBINATION FOR DIFFERENT NUMBER OF CATEGORIES

	Product	P_1	P_2	P_3	P_4	P_5	P_6	e , a
	Pre-combination cost	642452.70	429899.16	198873.32	99256.30	99248.69	38919.21	-
l=5	Combined cost	656784.98	412599.88	191193.49	100664.88	106626.71	40779.43	e =4.4% a =95.6%
	Absolute error	14332.29	-17299.28	-7679.83	1408.58	7378.02	1860.22	
	Relative error	2.23%	-4.02%	-3.86%	1.42%	7.43%	4.78%	
l=4	Combined cost	655093.87	411744.13	193699.60	101740.70	105913.57	40457.52	e =3.99% a =96.01%
	Absolute error	12641.17	-18155.04	-5173.72	2484.39	6664.88	1538.31	
	Relative error	1.97%	-4.22%	-2.60%	2.50%	6.72%	3.95%	
l=3	Combined cost	642202.63	399578.71	206435.60	111181.99	105426.39	43824.06	e =8.23% a =91.77%
	Absolute error	-250.07	-30320.45	7562.29	11925.69	6177.70	4904.85	
	Relative error	-0.04%	-7.05%	3.80%	12.02%	6.22%	12.60%	
l=2	Combined cost	607487.43	478913.67	145778.18	93637.77	143997.54	37093.60	e =22.21% a =77.79%
	Absolute error	-34965.27	49014.51	-53095.13	-5618.53	44748.85	-1825.61	
	Relative error	-5.44%	11.40%	-26.70%	-5.66%	45.09%	-4.69%	

At the same time, it can be seen from Table VI that when the number of categories is 5, the maximum one of the error between products costs before and after the combination is 7.43%, the error degree of all product costs is 4.4%, and the accuracy is 95.6%. When further clustered into 4 categories, the error degree of all products cost reaches a minimum of 3.99%, and the accuracy reaches the highest, which is 96.01%. After that, with the number of categories of cost drivers decreasing, the error degree of all product costs gradually increases, and the cost accuracy decreases. When the number of categories is 2, the error degree of all products costs reaches the maximum which is 22.21%, and the accuracy reaches the lowest which is 77.79%.

The above results show that the product costing model is simplified from 14 cost drivers to 4 cost drivers, which is the best simplification result. It is superior to the methods of Wang *et al*. (2009) and Liu *et al*. (2014). Therefore, the method of SVD-based combination of cost drivers is better.

D. Discussion: The Number of Non-zero Singular Values Considered for Cost Driver Classification

The above-mentioned cost driver classification is mainly based on the eigenvector in the right singular matrix corresponded to the largest singular value, and discards the eigenvectors corresponded to other non-zero singular values. One question is: When classify the cost drivers, should we discard non-zero singular values except the largest singular values? Or, how many larger singular values should we used to classify the cost drivers? In order to answer this question, we use dividing the cost drivers into four categories as an example because the error degree of all product costs is the lowest when the number of categories is 4. On the basis of classification of eigenvector in the he right singular matrix corresponded to the largest singular value, we increase the eigenvectors in the right singular matrix corresponded to the singular values one by one from large to small, and then cluster the cost drivers.

TABLE VII: CATEGORIES OF ALL 14 COST DRIVERS WHEN VALUES OF S ARE DIFFERENT

S values	Categories
2	$\{j_4\}$, $\{j_3\}$, $\{j_2, j_5, j_7, j_9, j_{13}\}$, $\{j_{10}, j_1, j_6, j_8, j_{11}, j_{12}, j_{14}\}$
3	$\{j_4\}$, $\{j_3, j_5\}$, $\{j_{12}, j_2, j_7, j_9, j_{13}\}$, $\{j_1, j_6, j_8, j_{10}, j_{11}, j_{14}\}$
4	$\{j_4, j_1, j_6\}$, $\{j_3, j_5\}$, $\{j_{10}, j_{11}\}$, $\{j_2, j_7, j_{12}, j_9, j_{13}, j_8, j_{14}\}$
5	$\{j_4\}$, $\{j_{10}\}$, $\{j_{11}\}$, $\{j_3, j_5, j_1, j_6, j_2, j_7, j_{12}, j_9, j_{13}, j_8, j_{14}\}$
6	$\{j_4, j_6\}$, $\{j_{10}\}$, $\{j_{11}\}$, $\{j_3, j_5, j_2, j_7, j_{12}, j_9, j_{13}, j_1, j_8, j_{14}\}$

According to the formula (4), when we cluster the 14 cost drivers based on the 14 column vectors of the matrix $(v_1 \quad v_2 \quad \cdots \quad v_s)^T$ corresponded to the s largest singular values and treat them as 14 points in the s-dimensional space, the 14 cost drivers are clustered into four categories. Table VII gives the clustering results for all 14 cost drivers when the values of s are different. Meanwhile, the bold words in Table VII also show representative cost drivers selected in each category according above step 5. After combining all 14 cost drivers into 4 cost drivers, Table VIII shows the calculated the product costs and the cost errors.

TABLE VIII: THE NUMBER OF SINGULAR VALUES USED FOR CLASSIFICATION AND THE ERROR FF PRODUCTS COST

	Product	P_1	P_2	P_3	P_4	P_5	P_6	e, a
	Pre-combination cost	642452.70	429899.16	198873.32	99256.30	99248.69	38919.21	-
s =2	Combined cost	570169.54	480637.51	217805.32	102059.27	101951.83	36025.91	e =8.44%
	Absolute error	-72283.16	50738.34	18932.01	2802.97	2703.14	-2893.30	a =91.56%
	Relative error	-11.25%	11.80%	9.52%	2.82%	2.72%	-7.43%	
s =3	Combined cost	560532.76	480873.28	221424.46	102639.69	107090.81	36088.37	e =9.65%
	Absolute error	-81919.93	50974.12	22551.14	3383.39	7842.13	-2830.84	a =90.35%
	Relative error	-12.75%	11.86%	11.34%	3.41%	7.90%	-7.27%	
s =4	Combined cost	724583.41	466367.21	152070.95	87675.79	92814.01	36048.67	e =13.05%
	Absolute error	82130.71	36468.04	-46802.37	-11580.51	-6434.68	-2870.55	a =86.95%
	Relative error	12.78%	8.48%	-23.53%	-11.67%	-6.48%	-7.38%	
s =5	Combined cost	604778.37	477790.85	159308.18	89111.72	135444.71	40474.37	e =18.28%
	Absolute error	-37674.33	47891.68	-39565.13	-10144.58	36196.02	1555.15	a =81.72%
	Relative error	-5.86%	11.14%	-19.89%	-10.22%	36.47%	4.00%	
s =6	Combined cost	605639.80	487164.62	159196.13	85119.15	130252.98	41276.69	e =17.43%
	Absolute error	-36812.90	57265.45	-39677.18	-14137.15	31004.30	2357.48	a =82.57%
	Relative error	-5.73%	13.32%	-19.95%	-14.24%	31.24%	6.06%	

It can be seen from Table VIII and Table VI that when the number of non-zero singular on which the cost driver's classification depends is increased from 1 to 2 up to 5, the error degree of all product costs is increased from 3.99% to 8.44%, and finally increased to 18.25%; while the accuracy of all product costs decreases from 96.01% to 91.56%, and finally decreases to 81.72%.When the classification considers all six non-zero singular values, the error degree of all products reaches 17.43% ,the accuracy of all product costs reaches 82.57%.

We also consider the case where all 14 cost drivers are divided into five categories. That is, when the costing model is reduced to five cost drivers, as the number of singular values considered increases from 1 to 2 and gradually increases to 6, the error degree of all product costs increases from 4.4% to 6.68% and gradually increases to 19.6%. The accuracy of all product costs decreased from 95.6% to 93.32%, and gradually decreased to 80.4%. This conclusion is basically the same as when all 14 cost drivers are divided into four categories.

The above discussion shows that, based on the principle of SVD of the matrix, when classifying and combining the cost drivers represented by the column vectors of the coefficient matrix, we only need to consider the eigenvector corresponded to the largest singular values in the right singular matrix. The greater the number of singular values considered, the greater the error degree of all product costs calculated by classifying and combining the cost drivers, and the lower the accuracy of product cost.

VI. CONCLUSION

Reducing the complexity of the activity-based costing system is one of the key factors in the successful implementation of the activity-based costing system. Based on the dimension reduction effect of singular value decomposition of a matrix, this paper proposes and studies the SVD-based combination method of cost driver under activity-based costing. By establishing a product costing model under activity-based costing, the singular value decomposition of the coefficient matrix in the model is carried out. The singular value and the eigenvectors in and the right singular matrix are used to classify the cost drivers. Then the representative cost drivers are selected to perform the combination of cost drivers and simplify the costing model. Numerical examples show that the method of SVD-based combination of cost drivers significantly reduces the error between product costs before and after combination and improves the accuracy of the combined product cost while reducing the complexity of the activity-based costing system. Compared with existing methods of cost driver combination such as integral value planning and clustering methods, the method of SVD-based combination of cost drivers is superior to other methods in reducing model complexity and ensuring accuracy of product cost.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

In this work, Xudanyang Li reviewed current literatures and conducted the research on SVD of the coefficient matrix; Buxi Li constructed the product's costing model under ABC and analyzed the number example; the two authors completed the calculation of the data together; Xudanyang Li wrote the paper; all authors had approved the final version.

ACKNOWLEDGMENT

The authors thank Ms. Xueyan Xu for her help in explanation of the theory of matrix and the principle of singular value decomposition. At the same time, the authors also very appreciate the anonymous reviewers for their guidance on this work.

REFERENCES

[1] R. Cooper and R. S. Kaplan, "Measure cost right: Make the right decision," *Harvard Business Review*, vol. 66, no. 5, pp. 96-103, Sep/Oct 1988.

[2] N. Dopuch, "A perspective on cost drivers," *Accounting Review*, vol. 7, pp. 121-126, July 1993.

[3] G. Foster and D. W. Swenson, "Measuring the success of activity based cost management and its determinants," *Journal of Management Accounting Research*, vol. 9, pp. 109-143, 1997.

[4] M. Gupta and K. Galloway, "Activity-based costing/management and its implication for operations management," *Technovation*, vol. 23, pp. 131-138, Feb. 2003.

[5] M. S. Velmurugan, "The success and failure of activity-based costing systems," *Journal of Performance Management*, vol. 23, no. 2, pp. 3-33, July 2010.

[6] W. H. Tsai, T. S. Tsaur, Y. W. Chou, J. Y. Liu, J. L. Hsu, and C. L. Hsieh, "Integrating the activity-based costing system and life-cycle assessment into green decision-making," *International Journal of Production Research*, vol. 53, pp. 451-465, Sep. 2014.

[7] A. Almeida and J. Cunha, "The implementation of an activity-based costing (ABC) system in a manufacturing company," *Procedia Manufacturing*, vol. 13, pp. 932-939, 2017.

[8] R. S. Kaplan and S. R. Anderson, "Time-driven activity-based costing," *Harvard Business Review*, vol. 82, no. 11, pp. 131-138, Nov 2004.

[9] Y. M. Babad and B. V. Balachandran, "Cost driver optimization in activity-based costing," *The Accounting Review*, vol. 68, no. 3, pp. 563-575, July 1993.

[10] M. J. Schniederjans and T. Garvin, "Using the analytic hierarchy process and multi-objective programming for the selection of cost drivers in activity-based costing," *European Journal of Operational Research*, vol. 100, pp. 72-80, 1997.

[11] C. It can be seen that all 14 cost drivers are divided into 5 categories at most, and these 5 categories can be further clustered into 2 categories at least. In the 5 categories, only the cost driver j_{13} is clustered as a separate category, and the other 4 categories contain at least two cost drivers, and the representative cost drivers are selected in each category, as shown in Table V. The bold words in the Table IV indicate selected representative cost driversHomburg, "A note on optimal cost driver selection in ABC," *Management Accounting Research*, vol. 12, pp. 197-205, June 2001.

[12] J. M. Zhao and L. Y. Ouyang, "Research on the selection of cost drivers under activity-based costing," *Journal of Beijing Jiaotong University* (*Chinese*), vol. 8, no. 4, pp. 78-82, Apr. 2009.

[13] B. XIt can be seen that all 14 cost drivers are divided into 5 categories at most, and these 5 categories can be further clustered into 2 categories at least. In the 5 categories, only the cost driver j_{13} is clustered as a separate category, and the other 4 categories contain at least two cost drivers, and the representative cost drivers are selected in each category, as shown in Table V. The bold words in the Table IV indicate selected representative cost drivers. Li and P. X. Wang, "Research on identification and selection model of multi-cost driver combination," *The Academic Symposium Proceedings of Management Accounting and Application Professional Committee of Accounting Society of China*, Nanjing, pp. 120-121, Sep. 2009.

[14] P. X. Wang, F. J. Wang, and P. Wang. "Research on the homogeneity of cost drivers," *The Journal of Quantitative & Technical Economics* (*Chinese*), no. 5, pp. 44-46, May 1999.

[15] B. It can be seen that all 14 cost drivers are divided into 5 categories at most, and these 5 categories can be further clustered into 2 categories at least. In the 5 categories, only the cost driver j_{13} is clustered as a separate category, and the other 4 categories contain at least two cost driveIt can be seen that all 14 cost drivers are divided into 5 categories at most, and these 5 categories can be further clustered into 2 categories at least. In the 5 categories, only the cost driver j_{13} is clustered as a separate category, and the other 4 categories contain at least two cost drivers, and the representative cost drivers are selected in each category, as shown in Table V. The bold words in the Table IV indicate selected representative cost driversrs, and the representative cost drivers are selected in each category, as shown in Table V. The bold words in the Table IV indicate selected representative cost driversX. Li, P. X. Wang, and L. Chen. "Study on cost driver combination theory in activity-based costing," *System Engineering – Theory & Practice* (*Chinese*), vol. 25, no. 8, pp. 55-61, Aug. 2005.

[16] B. X. Li and P. X. Wang, "Research on multi-cost driver combinations in activity-based costing," *System Engineering – Theory & Practice* (*Chinese*), vol. 27, no. 4, pp. 47-53, Apr. 2007.

[17] F. J. Wang, H. Chang, and K. Huang, "Using clustering analysis for the selection and combination of cost drivers in activity-based costing," *Management Review* (*Chinese*), vol. 21, no. 9, pp. 94-99, Sep. 2009.

[18] X. W. Liu, M. C. Ouyang, and J. Xu, "Systematic approach construction building of the selection, combination and evaluation of activity cost driver," *Chinese Journal of Management Science*, vol. 22, no. 11, pp. 73-78, Nov. 2014.

[19] D. Calman, "A singularly valuable decomposition: The SVD of a matrix," *The College Mathematics Journal*, vol. 27, no. 1, pp. 2-23, Jan. 1996.

[20] Q. L. Jin, P. X. Wang, and L. Thomas, "How an ABC study helped a China stated-owned agricultural machine company to stay competitive," *Journal of Cost Management*, vol. 19, no. 6, pp. 39-47, 2005.

Xudanyang Li was born in Taiyuan, China, in April 1991. He received a bachelor's degree in business from James Cook University Singapore Campus in 2013 and a master's degree in international business from University of Newcastle, Newcastle, Australia. His research interests include accounting for decision making, corporate finance, and management.

Buxi Li was born in Shanxi, China, in June 1963. He received a bachelor's degree in science in 1986 from Shanxi University, Taiyuan, China, the master's degree in Science in 1992 from Peking University, Beijing, China, and a doctor' degree in management in 2007 from Xi'an Jiaotong University, Xi'an, China. He works currently as an associate professor at Shanxi University. His research interests include management accounting, auditing, and capital market.