연구논문집 제 10 권 제 2 호 p. 41～49 1079.

Acyclic Ordering by Design Variable Selection in Chemical Process Design

Kim, Byung-Kyu
Dept. of Industrial Chemistry

<Abstract>

A new method for design variable selection operable on occurrence matrix and producing all sets of design variables lead to acyclic ordering is suggested on the basis of admissible output assignment. Further, systematic method for counting variables and conditions is successfully applied, and resulted in reduction of the occurrence matrix considerably.

설계변수 선택에 의한 공정설계 계산의 Acyclization

김 병 규
공 업 화 학 과

〈요 약〉

admissible output assignment에 기반하여, occurrence matrix상에서 조작 가능하며 acyclic ordering을 가져오게 하는 모든 설계변수의 조합을 얻을 수 있는 새로운 설계변수 선택 algorithm을 고안하였다.

또한 설계문제를 formulation함에 있어서 열역학적 이론에 기초를 둔 체계적인 방법을 이용함으로써 matrix를 다소간 축소시킬 수 있음을 알았다.

I. Introduction

A general characteristic of chemical process design problem is that there exixt more unknown variables than mathematical relations describing the plant or process. And further, usually the system of equations is sparse, random and nonlinear.

In the past, the computational strategy has been determined from experience and engineering intuition. And even in the formulation of problem set, some variables were counted twice, some ignored, and also design relations were not always traced out exactly from independent sources of informations.

When the number of variables is greater than the number of equations, a set of design variables must be assigned numerical values in order that the system be reduced to a determinate system with an equal number of equations and variables. A careless selection of that set of design variables can lead to trouble, magnifying computational difficulties above what be inherent to the system itself.

Solution sequences for algebraic equations can be separated into three classes: acyclic, simultaneous and iterative solution sequences.

Acyclic sequence, as opposed to simultaneous or iterative sequence solves each equation, whether linear or nonlinear, in a one at a time technique and does not require any assumed solution points. And thus acyclic ordering is regarded as the best strategy if only the system does not possess persistent iteration.

Acyclic ordering by careful selection of design variables was demonstrated by Rudd and his coworkers(1, 2) on the bipartite graph. The algorithm, termed LCR algorithm, always gives one possible acyclic ordering if such an ordering is possible.

In the present paper, the LCR algorithm is expanded, and a new algorithm operable on matrix and gives all possible combinations of design variables is suggested. And, in the formulation of problem set a systematic nethod for counting the variables and conditions achieved by Kwauk(3) is employed. Resultantly, the occurrence matrix is reduced to a considerable number of rows and columns.

Ⅱ. The Degree of Freedom in a System

The set of describing equations encountered in process design may be expressed in the symbolic from

$$F(V)=0 \qquad (1)$$

where $F(V)$ is a function of M variables $[V=(v_1, v_2, v_M)]$ and whose elements represent the process design equations

$$f_i(V_j)=0 \qquad (2)$$

wher $i=1, 2, \cdots, N$ and
where $V_j \in [v_1, v_2, \cdots v_M]$ and further
where $M \geq N$.

In general, the set of algebraic functions is unspecified. That is typically there are more variables than equations. Since, if the equations are equal to the variables in number, only certain definite values of the variables satisfy the design equations and accordingly optimization is not usually possible. Additionally the set is generally characterized by being sparse and random. Prior to attempting any solution of the set of equations the equation set must be reduced by an N-by-N set of equations and variables. In alternative expression, it may be said that $F=M-N$, degree of freedom, variables must be preassigned with numerical values, then the other state variables may be calculated from design equations expressing the conditions imposed by the fundamental laws controlling the system.

Ⅲ. Systematic Method for Counting Variables

In the process of design, essential step prior to the reduction to an N-by-N set of equations and variables is to count all of the independent variables and equations neither exclusively nor repeatedly.

In the early stage of process developement, they were not usually examined carefully, and accordingly the design thus obtained did not provide the optimum operation. Typical example for this erroneous design can be found in the design of rectifying column separating binary mixture under the simplified assumptions(4) made on the ordinary McCabe-Thiele constr-action.

The column is to be designed to have one intermediate feed, a partial reboiler with a liquid-bottom-product stream, and a total conenser with a liquid-distillate-product stream

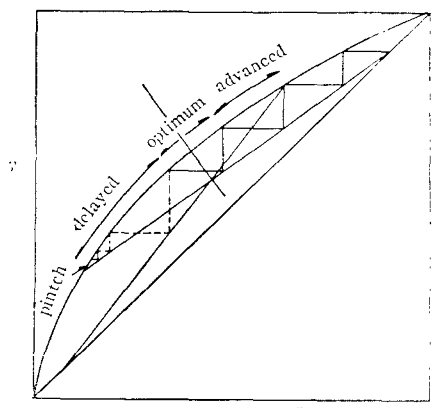


Fig. 1. The erroneous location of feed plate

and it will be operated under degree one of pressure level. When feed composition and conitions, and reflux ratio are specified, it is no longer possible to fix two terminal concentrations of one component and feed plate together. Spceification of concentrations of the overhead and bottoms is sufficient to fix the optimum location of feed plate absolutely. Otherwise, feed entry is delayed or advanced feed, as in Fig. 1, and is not the plate of optimum.

The first systematic method for finding the correct number of independent variables based on thermodynamic laws was initiated by Gilliland and Reed(5), and further discussions are available(6). The laws adopted by Gilliland and Reed are:

1. Law of Conservation of matter
2. The first law of thermodynamics
3. Phase rule

In the system composed of C independent components under equilibrium, law 1 indicates that there are C independent material balances applying of the contacting units. For the flow system in which changes in kinetic energy, potential energy and work done are negligible, the first law of thermodynamics is simplified to straightforward enthalpy or heat balance. Normally heat balance determines one relation in any system. The phase rule gives the degree of feedom of thermodynamics for any single-stage system in which there exists either one or more phases in equilibrium. It must be noted here that the degree of freedom in phase rule represents the difference between total number of thermodynamic intensible variables and the number of equilibrium relations existing among the variables.

The principles Gilliland and Reed employed enabled them to count the exact number of independent variables and conditions. But the systems were restricted to absorption and rectification columms and the procedure was rounding up.

More general and systematic approach thus applicable for wider ranges of stagewise processes was advocated by Kwauk(3) based on the slightly modified principles of Gilliland and Reed. Kwauk resolved the whole process into sinpler component classes i. e., element, complex element, unit and complex unit along the increasing complexity.

In his system of couning variables, phase rule gives all intensible variables for stream, flow rates of phases give extensible variables as many numbers as phases, and the degree of freedom of choosing the amount, or rate of energy exchanged between the system and surrounding gives one more variable.

He also dissected the conditions into two, inherent and necessary. The necessary conditions usually chosen by the designer are composition and flow rate of feed, the pressure of the system, and heat exchanged between system and surrounding. This necessary condition is fixed in advance and this is considered very convinient idea to avoid troublesome often occurred when counting is carried by the previous method(5).

In counting the inherent conditions, total energy balance gives one, material balances give as many conditons as there are components in the system.

Additional conditions inherent in a system are found in equality of certain streams such as those joining different parts of the composite system.

The above mentioned method for counting the variables may be summed up as follows.

Basic relations

Number of in dependent variables are the difference between total variables and inherent conditions plus necessary conditions.

or $M_i = M_v - (N_{c,i} + N_{c,n}) = M_v - N_c$ (3)

Normally fixed variables are those for feed stream, pressure level, and heat exchanged between system and sunounding.

or $$M_X = M_F - M_\pi + M_q \qquad (4)$$

Accordingly, the variables available for optimization are the difference betwen independent variables and normallye fixed variables.

or $$M_a = M_t - M_x = M_V - N_C - M_x \qquad (5)$$

Relations between major and minor classes. (y denotes those for major class)

$$M_v^y = \sum_{minor} M_i - F_\alpha \qquad (6)$$

$$M_t^y = \sum_{minor} M_i - F_\alpha - N_c^y \qquad (7)$$

$$M_x^y = M_F^y - \sum_{minor} (M_x - M_F) \qquad (8)$$

$$M_a^y = \sum_{minor} M_i - F_\alpha - N_c^y - M_F^y - \sum_{minor} (M_x - M_F) \qquad (9)$$

This systematic method for counting variables includes neither equipment nor economic design. But, if the additional equations be algebraic ones, the variable counting method may be retained without the loss of generality.

Fundamentaly, all M's in Eqn. (3)~(9), and all N's in Eqn. (3), (5), (7) and(9) are belong to V and F in Eqn. (1) respectively. Accordingly, they are henceforc treated together without specifications, and a few terms are defined.

The number of independent variables under the inherent conditions alone are termed simply deqree of freedom.

The number of independent variables under some envirnmental(economical, envirnmental, intentional. of control) are promised to be called envirnmental deqree of freedom.

And in the latter case, the variables specified by the envirnments are referred to envirnmental design variables.

This combirnatorial principle enables us to apply for flow diagram(material or iciformational) directly by considering the relation(8)

$$\text{system degree of freedom} = \Sigma F - n \qquad (10)$$

where ΣF denotes the sum of degree of freedom of each component and n denotes interstream relations.

Ⅳ. Design Variable Selection Algorithm

In section Ⅲ, systematic methods for counting variables and conditions were introduced. The next problem confronted with, after the degree of freedom of equation system has been determined, is to choose set of variables lead to solvable set of equations. There are $C_V^M = C_F^M = M!/N!F!$ possible combinations of design variables, not all of which lead to a solvable set of equations. A necessary and sufficient condition for the solution of a set of equations is noted that the Jacobian determint be nonzero (7). But this condition is of little direct use to the process calculation since it is extremely difficult to compute the Jacobian determinant for a large system. Accordingly it is strongly required to device an alternative method for design variable selection which will diminish enormous computational labors associated with process analysis.

The very tool employed for new method for solution was well-known bipartite graph of geometric topology which can expose the structure of equations compactly and lucidly, and which leads directly to their dissection.

The basic concept and property of that graph will be received attentions for the further discussions.

In a logically consistent set of N equations, there will be N unspecified variables, and for each equation it will be possible to denote one variable as the output variable for that equation.

And no variable will be output variable for more than one equation. If such an output assignment is possible, i.e., if the diversity condition of Hall is satisfied, the necessary condition for the existence of a solution is achieved. This condition might be stated thus: There must be at least K unspecified variables associated with each and every group of K equations where $K=1,2,\cdots N$.

This condition will be directly applied for the selection of design variables. If a variable v_j is assigned as the output variable of equation f_i, then the edge connecting node f_i and v_j will be oriented from f_i to v_j and all other edges associated with f_i will be oriented to f_i and accordingly, the original bipartite graph will be directed graph.

At this stage, prior to the deduction of design variable selection algorithm, it might well be noted that the labor of solving a set of equations increases as the cube of the number of equations that must be solved simultaneously (2).

This observation gives the criterion for the selection of design variables i.e., a best set of design variables results in acyclic structure if such an ordering is possible. And further, it may be assumed that all acyclic orderings give the same difficulty compared with the simultaneous set of equations.

If acyclic structure be obtained, there are some important properties pertinent to the graph which might be applied for the establishment of the algorithm reversely.

The graph of acyclic system must contain at least one v node v_j with $\rho(v_j)=1$, and one f node f_i with $\rho(f_i)=1$ where $\rho(v_j)$ and $\rho(f_i)$ denote local degrees of freedom of v_j and f_i respecitvely. This is because every directed graph mut terminate at a v node. Furthermore, since a v node can have only one incomming edge, this v node must have exactly one edge if a path terminate at this v node. If this v node and the f node of which it is an output, are deleted from the graph, the resulting subgraph must also be acyclic, hence, it will also contain at least one v node and one f node with one edge each. Finally it must be noted that any f node connected to only one v node must have that v node as its output.

The first work to base the design variable selection algorithm on bipartite graph was demonstrated by Lee, Christensen and Rudd (1). And later, the algorithm was operated on occurrance matrix (2) which is

1. Make entry unity if there is edge between node v_j and f_i, otherwise 0.
2. Locate *a* column containing only one unity and delete the column and the corresponding row.
3. Repeat step 2 until all equations have been eliminated.

The order in which the equations is to be solved was suggested the reverse of the order in which the equations were deleated.

This and the original algorithm (1) give only one feasible solution for any system. But in practical situations of process designs, we are often limited to strong restrictions as were described in the end of section Ⅲ. For wider

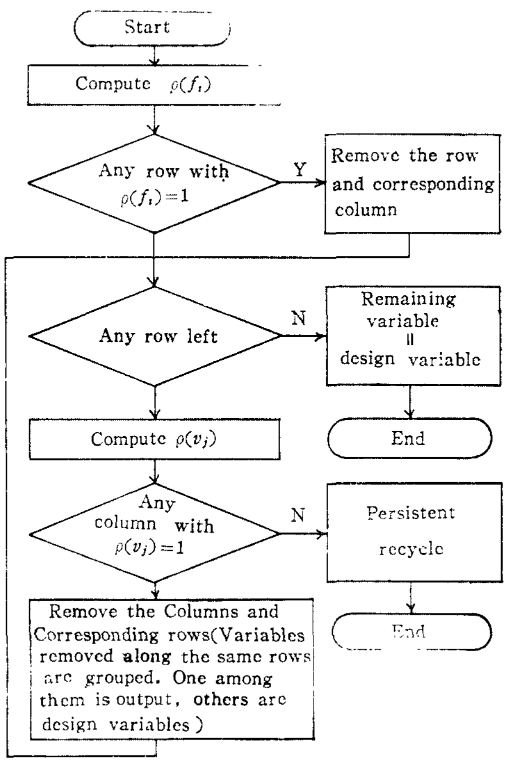


Fig. 2. Design variables selection algorithm.

ranges of applications, it is desired to find all sets of feasible solutions lead to acyclic calculation order.

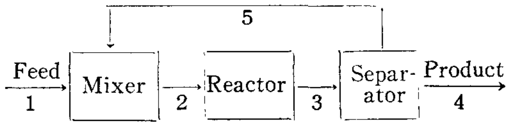
This may be achieved by the expansion of those algorithms (1, 2) described previously on the observation that only one variable can be assigned as the output of any equation. This is quite reasonable from the definition of acyclic structure.

The algorithm newly developed is shown in Fig. 2.

Two examples are shown for the application of the design variable selection algorithm developed.

Example 1.

Mixer-Reactor-Separator system originally suggested for LCR algorithm application.



Design equations

Mixer

f_1 $m_1+m_5-m_2=0$

f_2 $m_1x_1+m_5x_5-m_2x_2=0$

Reactor

f_3 $m_2-m_3=0$

f_4 $m_2x_2-m_3x_3-R(x_3, T. V)=0$

Separator

f_5 $m_3-m_4-m_5=0$

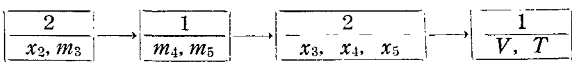
f_6 $m_3x_3-m_4x_4-m_5x_5=0$

f_7 $S=x_5/x_4$

$F=6$

	x_1	m_1	x_2	m_2	x_3	m_3	x_4	m_4	x_5	m_5	V	T	S	$\rho(f_i)$
f_1		1		1						1				3
f_2	1	1	1	1					1	1				6
f_3				1		1								2
f_4			1	1	1	1					1	1		6
f_5						1		1		1				3
f_6					1	1	1	1	1	1				6
f_7							1		1				1	3
$\rho(v_j)$	1	2	2	4	2	4	2	2	3	4	1	1	1	
		1		2	1	3	1	2	1	3				
				1		2		1		1				

Design variables and precedence order



The number in block denotes the number of variables that must be chosen as design variable in each block.

Five more poscible sets of design variables are obtained compared withs that obtained by employing LCR algorithm.

Example 2.

Mixer-Exchanger-Mixer system The following assumptions were made in the developement of the design equations which model the system: steady state operation, pressure effects were ignored, all physical properties were known and constant.

All variables and relations were traced by the systematic method described in section Ⅲ

Variables

stream 1 $(1\phi, 1c)=3$, T_1, P_1, m_1
stream 2 $(1\phi, 1c)=3$, T_2, P_2, m_2
stream 3 $(1\phi, 2c)=4$, T_3, P_3, x_{13}, m_3
stream 4 $(1\phi, 1c)=3$, T_4, P_4, m_4
stream 5 $(1\phi, 2c)=4$, T_5, P_5, x_{15}, m_5
stream 6 $(1\phi, 2c)=4$, T_6, P_6, x_{26}, m_6
stream 7 $(1\phi, 4c)=6$, $T_7, P_7, x_{17}, x_{27}, x_{37}, m_7$
equipment 1 $=2$, τ_1, S_1
equipment 2 $=3$, $Q_2, A_2, (\Delta t)lm$
equipment 3 $=2$, τ_3, S_3

$\therefore \Sigma M_i = 27$(All P's neglected)

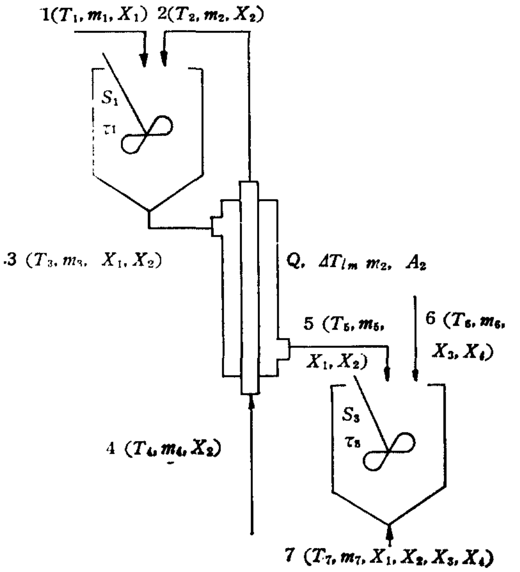


Fig. 3. Process flow diagram for mixer-exchanger-mixer system.

Design relations

Unit 1

f_1 $m_1+m_2=m_3$

f_2 $m_1=x_{13}m_3$

f_3 $m_1\bar{C}_{p1}T_1+m_2\bar{C}_{p2}T_2=m_3\bar{C}_{p3}T_3$

f_4 $S_1=\tau_1 m_3/\rho_3$

Unit 2

f_5 $x_{13}=x_{15}$

f_6 $m_3=m_5$

f_7 $m_4=m_2$

f_8 $m_3\bar{C}_{p3}T_3=m_5\bar{C}_{p5}T_5$

f_9 $m_4\bar{C}_{p4}T_4=m_2\bar{C}_{p2}T_2$

$$f_{10}\quad (\Delta t)lm=\frac{(T_2-T_3)-(T_4-T_5)}{\ln\dfrac{T_2-T_3}{T_4-T_5}}$$

f_{11} $Q_2=U_2A_2(\Delta t)l_m$

Unit 3

f_{12} $m_5+m_6=m_7$

f_{13} $m_5x_{15}=m_7x_{17}$

f_{14} $m_5(1-x_{15})=m_7x_{27}$

f_j \ v_j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	$\rho(f_i)$
1		1		1		1																						3
2		1				1	1																					3
3	1	1	1	1	1	1																						6
4						1															1	1						3
5							1					1																2
6						1					1																	2
7				1					1																			2
8					1	1				1	1																	4
9			1	1				1	1																			4
10			1		1			1		1														1				5
11																							1	1	1			3
12											1			1			1											3
13											1	1					1	1										4
14											1	1					1		1									4
15														1	1		1			1								4
16										1	1		1	1		1	1											6
17											1						1									1	1	4
$\rho(v_j)$	1	3	3	4	3	6	2	2	2	3	7	3	1	3	1	1	6	1	1	1	1	1	1	2	1	1	1	
		2	2	3	2	4	2	2	2	2	3	1		1			1							1				
		2	1	3	1	4	1	1	2	1	2																	
		1		2		2			1																			

Fig. 4. The occurrence matrix and operation of example 2.

f_{15} $m_3x_{36}=m_7x_{37}$

f_{16} $m_5\bar{C}_{p5}T_5+m_3\bar{C}_{p6}T_6=m_7\bar{C}_{p7}T_7$

f_{17} $S_3=\tau_3m_7/\rho_7$

$$\Sigma N_i=17,$$

And $F_\alpha=0$

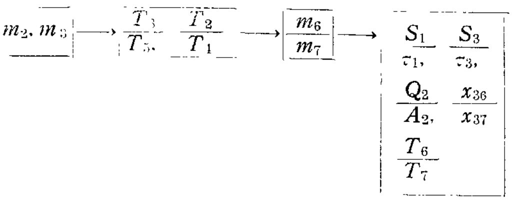
$\therefore M^y_i=10$

For the convinienc of mapping pressure equality relations were removed, and the number of variables are as follows.

1. T_1	2. m_1	3. T_2	4. m_2
5. T_3	6. m_3	7. x_{13}	8. T_4
9. m_4	10. T_5	11. m_5	12. x_{15}
13. T_6	14. m_6	15. x_{36}	16. T_7
17. m_7	18. x_{17}	19. x_{27}	20. x_{37}
21. S_1	22. τ_1	23. Q_2	24. $(\Delta t)l_m$
25. A_2	26. S_3	27. τ_3	

The number of design variables is ten which is coincident with the degree of freedom analyzed in setting the problem.

The precedence order of calculation is to be proceeded as follows.



V. Results

1. All possible variables and conditions were successfully traced by systematic counting method. Accordingly, the operation matrix was reduced considerably.
2. All possible combinations of design variables lead to acyclic ordering were obtained by new algorithm developed. Thus, design under strong restrictions could be rendered acyclic if the system is not subject to persistent iteration.

Notation

A_2=heat transfer area in unit 2 of example 2.

C=number of components

$\bar{C}_{pj}$=molal heat capacity of stream j

F=degree of freedom

vector function of all equations in the set, $f, f, \cdots f_N$

F_α=freedom of choice of number of times which any component could be repeated

M=number of variables

M_a=number of available variables in design

M_F=number of variables in feed stream

M_i=number of independent variables

M_Q=variable accounting for heat exchanged between system and surrounding

M_x=normally fixed variables

M^y=M corresponding to major class

M_π=variable accounting for pressure

m_j=mass or molar flow rate of stream j

N=number of equations

N_c=number of total conditions

$N_{c,i}$=number of inherent conditions

$N_{c,n}$=number of necessary conditions

N^y=N corresponding to major class

n=number of interstream relations

Q_2=rate of heat transfer in unit 2 of example 2.

S_k=volume of unit k

T=temperature of reactor in example 1.

T_j=temperature of stream j

U_2=overall heat transfer coefficient in unit 2 of example 2.

V=reactor volume in example 1.

vector of all variables in the set, $v, v, \cdots v_M$

v_j=variable, an element of V

x_j=composition in stream j

x_{ij}=mole fraction component i in stream j

$\rho(f_i)$=local degree of freedom of equation i

$\rho(v_j)$=local degree of freedom of variable j

ρ_j=molal density of stream j

τ_k=residence time in unit k

Literature Cited

1. W. Lee, J.H. Christensen, and D.F. Rudd, Design Variable Selection to Simplify Process Calculation, AICHE J., 12, 1104 (1965).
2. D.F. Rudd, and C.C. Watson, Strategy of Process Engineering, p.45, 53, J.Wiley, New York (1966).
3. M.Kwauk, A System for Counting Variables in Separation Processes, AIChE J., 2, 240 (1956).
4. See for example, W.L. McCabe, and J.C. Smith, Unit Operations in Chemical Engineering, p.551, McGraw-Hill, New York (1976).
5. E.R. Gilliland, and C.E. Reed, Degree of Freedom in Multicomponent Absorption and Rectification Columns, Ind. Eng. Chem., 34, 551 (1942).
6. Robinson, C.S., and E.R.Gilliland, Elements of Fractional Distillation, p.215, McGraw-Hill, New York (1950).
7. D.V.Wider, Advanced Calculus, p.48, Prentic-Hall, New Delhi (1968).