反応モデルの適否の検定*

目 崎 令 司**

1. はじめに

反応工学における反応モデルの組み立ては(1)反応モデルの想定とその式化、(2)想定したモデルの選判別、(3)モデル中のパラメーターの決定、(4)モデルの適否の検定の四つの工程にわけることができる。この中(1)~(3)の項目に関してはすでに2, 3の報告¹⁻³⁾もあり、特にモデルの選判別法及び非線型モデル中のパラメーターの決定法は高効率の電子計算機の普及ならびにそれに伴った数値解析法の発展により近年急速に開発された。モデルの適否の検定はすでに確立された分野でありその代表的な方法である残差分析や分散分析法などについては数理統計学関係の成書⁴⁻⁶⁾に広く解説されている。しかしながら反応モデルの検討にこれらの方法を取扱った報告はあまり見当たらないので、ここでは主として上記の分析方法を反応工学に関する例と組み合わせて紹介することにする。

2. 従来のモデル検定法

不均一系特に気体の接触反応に工学モデルとして広く使用された Hougen-Watson モデルではそのモデルの開発当初は線型化したモデルに線型最小二乗法を用いて得られた速度パラメーターの符号と実験値と計算値から得られる残差二乗和の値の大小から反応モデルの判別ならびに検定を行なうことが提案されており⁷⁾、また非等温条件で採取された速度データより得られたパラメーター値が Arrhenius の関係式に合うかどうかを調べることも有力なモデルの検定法であった。最近 Hougen-Watson 式のような非線型反応モデル中のパラメーターの算出法は数理統計学的処法により色々検討され、その結果線型化した非線型式から回帰法によって得られたパラメーターの算出値と非線型最小二乗法により直接得られた算出値との間にはかなりの差があることが見いだされてい

る⁸⁾。さらにまた非線型式中のパラメーターの算出値の符号はパラメーターの共信頼領域を調べることによって決定されるべきであることも報告されている⁹⁾。従って上記のモデルの検定基準はもはやあまり適切ではなくなった。1958年に Chou¹⁰⁾は Hougen-Watson 反応モデルの検定に統計的処法を使用することを強調し、また Froment ら¹¹⁾はこれらの方法をアルコールの接触脱水反応の反応モデルの検定に使用している。以上のような検定方式では一般に反応モデルの厳密な検定はできないばかりでなく時には全く誤った結論に至ることもある。算出パラメーターの正負の符号で反応モデルの適否を知るにはパラメーターの変動に対する残差二乗和面の性状を十分検討する必要がある、特に非線型パラメーターの数が多い場合は多次限のパラメーター空間における残差二乗和面を調べることになり、これは簡単に実施できる検定法ではない。

3. 残差分析

残差の分析の理論的考察に関しては Anscombe などの詳細な報告¹²⁻¹²⁾があるが、残差の検査はモデル内部に存在するモデルの欠陥を調べるのに非常に有効である。反応モデルに包含される色々の独立変数に関するモデルの微細な不適性を容易に検出することができる。残差はある応答たとえば反応速度の実測値とその反応速度の反応モデルによる計算値との差であり、簡単な例として今反応速度 r が二成分の分圧 p_1 と p_2 で次のように表わされるとする。

$$r = \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \varepsilon \quad (1)$$

ここで ε は実験誤差であり反応速度 r の計算値 $\hat{r}$ は β の推定値 b を用いて次の(2)式で与えられる。

$$\hat{r} = b_0 + b_1 p_1 + b_2 p_2 \quad (2)$$

推定値 b は最小二乗法などにより算出することができる。この場合残差は $r - \hat{r}$ であり、もし想定したモデルが正しければ実験誤差 ε に等しい。すなわち

$$r - \hat{r} = \varepsilon \quad (3)$$

(3)式からわかるように残差は分圧などの独立変数、

* 昭和44年5月6日 受理

** Reiji Mezaki (エール大学工学部 Department of Engineering and Applied Science, Yale University, New Haven, Connecticut, U.S.A.)

Fig. 1 Residuals and independent variables

Fig. 2 Residuals and independent variable

反応速度などの従属変数のいずれにも無関係で Fig. 1 に示すように零値を平均値としてでたために分布しなければならない。もし Fig. 2~Fig. 4 に示すようにならんかのかたまり、または相関関係が存在する場合にはたとえ後述する分散分析の結果が取り扱っているモデルの適性を証明したとしても、そのモデルは明らかに不適切である。

モデルの判別に残差を使用するには残差の頻度、残差と時間、残差と独立変数、従属変数などとの相関を調べることが必要である。残差の頻度は Fig. 5 に示すように残差の大きさとその発生頻度とをプロットするもので、このプロットより誤差分布が正規分布しているかどうかを知ることができる。もし使用したモデルが正しければ誤差すなわち残差は零点の正負の側にはほぼ等分に分布すべきである。次に残差と時間とのプロットはいわゆる Time-Sequence Plot と呼ばれ Fig. 1~Fig. 4 で横軸が時間の場合である。いうまでもなく実験データに対して採取された時間ならびに順序が解析に必要である。Fig. 2 は使用しているモデルが不適であることよりもむしろ誤差分散が時間とともに変動していることを示すものでモデル中のパラメーターを算出するには誤差分散の均一化を図るため重み関数を使用した最小二乗法を使用すべきであろう。また Fig. 3 は反応モデルに時間に関する一次項を導入する必要があることを物語っている。さらに Fig. 4 は時間に関する一次ならびに二次の項をモデルに付加する必要があることを示す。

Fig. 1~Fig. 4 はその横軸を反応物及び生成物の分圧、反応温度または接触時間などの反応工学で一般に使用される独立変数で置き換えることによって残差と独立変数との相関が得られ、それぞれの図より上記と同様の情報を得ることができる。またこれらの図において横軸を反応速度 r の計算値 $\hat{r}$ とした場合には残差と従属変数

Fig. 3 Residuals and independent variable

Fig. 4 Residuals and independent variables

Fig. 5 Residuals and frequency

との関係が得られ各図より抽出できる情報は残差と時間とのプロット及び残差と独立変数のプロットの場合と同様であるが、ただ Fig. 2 のような関係が得られる時は従属変数の変換、特に対数変換を行なって残差の発散状態を除去することがしばしば可能である。反応速度を従属変数として取り扱う場合は誤差分散が反応速度の絶対値の増加に比例して増加することは多く、従って反応速度の対数変換を行った後、最小二乗法をパラメーターの決定のため用いることが望ましい。白金触媒を使用したプロピレンの水素添加反応¹⁴⁾において不等温条件で得られた反応速度の誤差の不均一性が対数変換によってほぼ均一となった例はすでに報告されている¹⁵⁾。

以上のように残差プロットは反応モデルの適、不適を示すばかりでなく適切な反応モデルを導くためどのような項を追加し、あるいはまた消去すべきかを示すこともあり反応モデルの組み立てには一つの重要な処法であろう。電子計算機で反応データの処理とあわせて残差の計算ならびにプロットを行なう時は計算及び作業時間の増加はデータ数が極端に多い場合は別として全く問題はないが、ただ机上計算機を使用しなければならない場合は作業時間は一般にかなり増加するが、この作業より得られる情報の価値は要した仕事量及び消費した時間を補償することは確実である。

4. 分散分析法

分散分析法は反応速度などの応答の予想値と実験値とのずれと実験データの偶然的な変動とを統計的に比較するものであり、この比較よりモデルと実験点との間の一

致の状態を知ることができる。またその応用としていくつかの項の和からなっている線型モデルのあてはめに対する各項の寄与率を調べることができる。分散分析の詳細は数理統計学の成書にゆずるとしてここでは分析法を簡単に説明することにする。

今仮りに反応速度を応答とし N 個の実験点が採取されたとする。この際データの全変動 S_1 は $\sum_{i=1}^N r_i^2$ である。最小二乗法などの方法で想定したモデル中のパラメーターの算出値が得られたとするとこの反応モデルならびにパラメーター値を用いて各実験点における反応速度の理論値 $\hat{r}_i$ が算出される。この反応モデルに対する理論値の全変動 S_2 は $\sum_{i=1}^N \hat{r}_i^2$ であり、データの全変動 S_1 理論値の変動 S_2 、残差二乗和 S_3 との間に次の関係が成立する。

$$S_1 = S_2 + S_3 \quad (4)$$

すなわち

$$\sum_{i=1}^N r_i^2 = \sum_{i=1}^N \hat{r}_i^2 + \sum_{i=1}^N (r_i - \hat{r}_i)^2 \quad (5)$$

今実験点は N 点であり想定されたモデル中のパラメーター数を p とすると S_1 の自由度は N であり、また、 S_2 、 S_3 の自由度はそれぞれ p 、 $N-p$ である。実験値の誤差の値はある実験条件で行なったいくつかの繰返し実験から得ることができ、 $\bar{r}$ をこれらの全繰返し実験の平均値とすれば誤差二乗和は $\sum_{i=1}^N (r_i - \bar{r})^2$ である。もし残差二乗和がモデルそのものに起因した変動と誤差の二乗和からなっているとすると次の (6) 式が成立し

$$\sum_{i=1}^N (r_i - \hat{r}_i)^2 = \sum_{i=1}^N (r_i - \bar{r})^2 + \sum_{i=1}^N (\hat{r}_i - \bar{r})^2 \quad (6)$$

式中の $\sum_{i=1}^N (\hat{r}_i - \bar{r})^2$ は不適合二乗和 (Lack of fit sum of squares) でこの値を自由度で割った値、すなわち平均平方値と誤差に関する平均平方値の比を相等する F -値と比較することによってモデル全体が実験データによくあてはまっているかどうか検討することができる。簡単な例として今 k 個の異った実験条件で n 回の繰返し実験を行なったとすると誤差二乗和の自由度は $k(n-1)$ であり、また不適合二乗和の自由度は $N-p-k(n-1)$ である。従ってモデルの検定には

$$\frac{\text{不適合の平均平方}}{\text{誤差の平均平方}} - F_{\alpha}[N-p-k(n-1), k(n-1)] \begin{cases} < 0 \\ > 0 \end{cases} \quad (7)$$

を用いるわけである。ここで α は有意水準であり、もし (7) 式の左辺が負であればモデルは使用した有意水準で適切なモデルであり、反対に正の場合はモデルは不適切として棄却すべきであろう。

分散分析法は以上のように想定モデルの良否の決定に用いられる以外に最適項を含む線型モデルの選択にも用いることができる。線型の反応モデルは多くの場合実験

反応モデルであるが反応装置の設計や、その操作の最適化に対する予備的な知識を集めるには十分に役立つことが多い。今反応速度 r が反応及び生成物の濃度、反応接触時間、反応温度などの独立変数 $x_1, x_2, x_3, \dots, x_k$ の関数と考えられる場合、最も正確な実験速度式はいうまでもなくすべての独立変数の一次項ならびにこれらの二次項、交互作用項などを含んだ速度式を考え、これらの各項に付随するパラメーターを実験値より算出することである。しかしながら考慮する独立変数の数が多くなればパラメーターの算出に要する実験数も多くなり、従って速度式の適切度と実験時間の適切度の両者を考え合わせた最適反応モデルを見いだすことが望ましい。

Carr¹⁶⁾ はアルミナ担体金属触媒を用いて正ペンタンとイソペンタンの異性化反応速度を測定して次のような実験式を提出した。全実験点は24点でこの中実験誤差量を算出するため一つの実験条件で4回の繰返し実験が含まれている。

$$r = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 \quad (8)$$

ここで x_1, x_2, x_3 はそれぞれ水素、正ペンタン、イソペンタンの分圧を示す。Table 1 で示すように分散分析の結果は平均平方の比が5% および2.5% 水準の F 値以下であるから (8) 式の反応モデルは適切であると判断される。またこの反応系に対して Carr¹⁶⁾ は2つの Hougen-Watson モデルを提案しており、これらは次の (9)、(10) 式で示される。

$$r = \frac{k K_2 (x_2 - x_3 / K)}{(1 + K_1 x_1 + K_2 x_2 + K_3 x_3)} \quad (9)$$

$$r = \frac{k K_2 (x_2 - x_3 / K)}{(1 + K_1 x_1 + K_2 x_2 + K_3 x_3)^2} \quad (10)$$

(9) 及び (10) 式を線型化すれば

$$\frac{(x_2 - x_3 / K)}{r} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \quad (11)$$

$$\sqrt{\frac{(x_2 - x_3 / K)}{r}} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \quad (12)$$

となり分散分析法でこれらの線型化された反応式のいずれが適切であるかを調べた。その結果は Table 2 及び Table 3 に示す。分散分析の結果からは (12) 式は適切であると結論できるが、(11) 式はその不適合の平均平方と誤差の平均平方値の比が有意水準5% と2.5% との F 値の中間の値を取る所以この反応モデルの適切度には若干疑問がある。次に (8) 式で表わされる線型反応モデルで二乗項及び交互作用項が実際必要であるかを知りたい場合には先ず一次項のみを含んだ次式 (13)

$$r = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \quad (13)$$

を用いて Table 4 に示すように反応速度の回帰二乗和を求め、また二次項を含んだ反応モデル (8) に対して回帰二乗和を計算し両者の差を求めれば、この差は残差二乗和に対する二次項の寄与を示すものである。従って次式

Table 1 Analysis of variance for Eq. (8)

Source	Sum of squares	Degrees of freedom	Mean squares	Ratios	F Statistics	
					95%	97.5%
Crude	640.49	24				
Due to regression	637.35	10	63.74			
Residual	3.14	14	0.225			
Lack of fit	2.85	11	0.260			
Pure error	0.29	3	0.095	2.73	8.80	14.3

Table 2 Analysis of Variance for Eq. (11)

Source	Sum of squares	Degrees of freedom	Mean squares	Ratios	F Statistics	
					95%	97.5%
Crude	28894.02	24				
Due to regression	26896.82	4	6,724.2			
Residual	1997.20	20	99.86			
Lack of fit	1964.56	17	115.56			
Pure error	32.65	3	10.88	10.62	8.70	14.3

Table 3 Analysis of variance for Eq. (12)

Source	Sum of squares	Degrees of freedom	Mean squares	Ratios	F Statistics	
					95%	97.5%
Crude	763.54	24				
Due to regression	754.29	4	188.6			
Residual	9.25	20	0.462			
Lack of fit	8.98	17	0.528			
Pure error	0.27	3	0.090	5.90	8.70	14.3

で示す比の値の大小より (8) 式中の 6 ケの二次項を保持することが必要であるかどうかを決定することができる。

$$\frac{\text{二次項による平均平方}}{\text{残差平均平方}} \quad (14)$$

この例では Table 4 に見られるように平均平方の比は 5% 及び 2.5% 水準の F 統計値より大であるので、二次項は棄却できないことを示している。

分散分析法は以上のような線型モデルの外に非線型モデル中の項の必要度の判定に近似的に使用することもできる。Kabel と Johanson¹⁷⁾ はエタノールのイオン交換樹脂による脱水反応に対して次の反応モデルを提案した。

$$r = \frac{kK_A(p_A^2 - p_E p_W / K_{eq})}{(1 + K_A p_A + K_E p_E + K_W p_W)^2} \quad (15)$$

ここで k は表面反応速度定数であり、 K_A , K_E , K_W はそれぞれエタノール、エーテル、水蒸気の吸着平衡定数、 K_{eq} はこの脱水反応の平衡定数、また p_A , p_E , p_W はそれぞれエタノール、エーテル、水蒸気の方圧を示す。この反応式で Kabel らは K_E は零または零に非常に近いことを報告しているが、分散分析法で $K_E p_E$ 項がデータのあてはめに対してどの程度必要であるか調べることができる。非線型最小二乗法で (15) 式中の速度パラメーターを計算したがその結果を Table 5 に示し

Table 4 Complete analysis of variance for Eq. (8)

Source	Sum of squares	Degrees of freedom	Mean squares	Ratios	F Statistics	
					95%	97.5%
Crude	640.49	24				
Due to first order regression	627.00	4	156.8			
Extra due to second order regression	10.35	6	1.72			
Residual	3.14	14	0.225	7.66	2.90	3.5
Lack of fit	2.85	11	0.260			
Pure error	0.29	3	0.095	2.73	8.80	14.3

Table 5 Parameter estimates for Eq. (15)

Situation	k	K_A [atm ⁻¹]	K_W [atm ⁻¹]	K_E [atm ⁻¹]
With K_E	7.02×10^{-4}	2.91	6.34	8.7×10^{-3}
Without K_E	7.04×10^{-4}	2.89	6.30	

Table 6 Approximate analysis of variance for Eq. (15)

Source	Sum of squares	Degrees of freedom	Mean squares	Ratios	F Statistics	
					95%	
Regression without K_E (due to k , K_A , K_W)	6.411×10^{-8}	3	2.3170×10^{-8}	305.2	2.90	
Regression with K_E (due to K_E)	0.012×10^{-8}	1	0.0102×10^{-8}	1.46	4.15	
Residual	0.231×10^{-8}	33	0.0070×10^{-8}			

た。また近似的な分散分析の結果を Table 6 に示した。Table 6 の平均平方の比の値よりわかるように $K_E p_E$ 項はまず必要でない。

以上に記したような分散分析法を適用するにはいくつかの仮定が必要である。すなわち (1) 誤差は正規分布をしていること、(2) 誤差の分散は一定であること、(3) 誤差は互いに独立であること、が仮定されている。今これらの仮定の一部または全部が成り立たない時は分散分析はどのような結果を与えるかは興味ある問題である。すでにのべたように反応工学では特に (2) の仮定は必ずしも成り立たない場合が多く反応速度が著しく減少または増大すると反応速度の誤差の変動も大きくなることは自然である。このような場合の解析法の詳細は Scheffé⁴⁾ の成書に見られるが、一般に誤差分散の変動より算出される重み関数を使用した加重データを用いて種々の解析を行なえばこの仮定の誤りの影響は小さくなると考えられる。

おわりに

反応工学におけるモデルの組み立てはモデルの想定、その選別、パラメーターの決定、モデルの検定といった4つの同等に重要な連繫作業よりなり、これらの作業に対して物理化学専攻者はモデルの想定に興味を持ち、統計学の専攻者は実験計画やモデルの検定に主として注

意を払い、他方技術者はモデルを早急に決定して反応装置の設計、あるいは反応操作のシミュレーションや最適化などの実際問題の解決にあたらうとする。これら3者が協力して適切なモデルの組み立てという目的に向うことが望ましく、このためには化学工学専攻者は物理化

学、統計学を専攻した者がたびたび使う処法及びその意味を知る必要があるように考えられる。この解説が反応工学におけるモデルの組み立てに広く応用できる一統計的処法の紹介となれば幸いである。

引用文献

- 1) Box, G.E.P. and Lucas, H.L.: *Biometrika*, **46**, 77 (1959)
- 2) Mezaki, R.: *Kagaku Kōgaku (Chem. Eng., Japan)*, **30**, 1081 (1966)
- 3) Mezaki, R.: *Kagaku Kōgaku (Chem. Eng., Japan)*, **31**, 705 (1967)
- 4) Scheffé, H.: "The Analysis of Variance", John Wiley and Sons, New York, 1964.
- 5) Bennett, C.A. and Franklin, N.L.: "Statistical Analysis in Chemistry and the Chemical Industry", John Wiley and Sons, New York, 1966.
- 6) Draper, N.R. and Smith, H.: "Applied Regression Analysis", John Wiley Sons, New York, 1966.
- 7) Hogen, O.A. and Watson, K.M.: "Chemical Process Principles, Part III", John Wiley and Sons, New York, 1947.
- 8) Kittrell, J.R., Hunter, W.G. and Watson, C.C.: *AIChE Journal*, **11**, 1051 (1965)
- 9) Mezaki, R. and Kittrell, J.R.: *Ind. Eng. Chem.*, **50**, 63, May (1967)
- 10) Chou, C.H.: *Ing. Eng. Chem.*, **50**, 799 (1958)
- 11) Franckaerts, J. and Froment, G.F.: *Chem. Eng. Sci.*, **19**, 807 (1964)
- 12) Anscombe, F.J.: *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, Vol. 1 (1961)
- 13) Anscombe, F.J. and Tukey, J.W.: *Technometrics*, **5**, 141 (1963)
- 14) Shabaker, R.H.: Ph. D. Thesis, University of Wisconsin, Madison, Wisconsin, U.S.A. (1965)
- 15) Mezaki, R.: *Kagaku Kōgaku (Chem. Eng., Japan)*, **31**, 400 (1967)
- 16) Carr, N.L.: *Ind. Eng. Chem.*, **52**, 391 (1960)
- 17) Kabel, R.L. and Johanson, L.N.: *AIChE Journal*, **8**, 621 (1962)

資料

西ドイツにおける化学工学*

林 信 也**

西ドイツにおける化学工学という題名を冒頭にかかげてあるが、すべての物はその歴史をたずねなければならないという意味で、ドイツの化学工学ということと、ある意味で同じである。しかし第2次大戦後、ドイツが東西に分割されてよりは、東ドイツに関する情報がほとんど得られていないので、最近の傾向をふくめては西ドイツの化学工学ということになる。

Chemical Engineering はアメリカで最初に体系化された工学として出現し、現在でも化学工学の発展の中に大きな影響をもちつづけている。それはアメリカのヨーロッパからの独立以来の急激な国力と経済の発展、化学工学にとっては、とりわけ化学工業の発展とそれに伴う要求がもたらしたものであって、主として当時ヨーロ

ッパにおいて発達した化学工業技術をアメリカにおいてつくりあげる際に重点のおかれた、新大陸の豊富な工業原料と莫大な需要からくる化学工場建設の大規模化と効率化がその契機になっている。つまりアメリカにおける Chemical Engineering は、設計、開発 (Entwicklung) であって研究 (Forschung) を少なくとも当初は目標にしていなかったという見方も成立してくる。ヨーロッパにおける工学 (Engineering) は、古くより歴史の伝統をうけついで Maschinentechnik, Elektrotechnik, Bautechnik に三大別され、すべての Institut (Lehrstuhl) はこれらのどれかに属していた。当時ドイツは化学工業の世界中での最先進国として君臨していたが、その理由は現象的には工学、化学および物理学のすぐれた技術の寄せ集めでなく、すぐれた独創という指導思考が存在していたからである。Bosch-Haber のアンモニア合成法の確

* 昭和44年5月15日 受理

** Shinya Hayashi 神戸大学工学部 化学工学科