

제 2 장

데이터의 회귀분석 및 상관 관계

제 1 절 비선형 회귀분석을 사용하는 Antoine 식의 매개인자 추산

1.1 개념 설명

온도에 따른 증기압 데이터를 상관시키기 위하여 Antoine 식의 직접 사용.

1.2 사용된 수치 해법

총괄 분산 및 각 매개인자의 신뢰구간의 결정과 함께 일반적인 대수 표현의 비선형 회귀 분석.

1.3 문제 설명

Antoine 식은 세 매개인자 A , B , C 를 사용하는 증기압 상관식으로 널리 쓰이는 식이다. Antoine 식은 다음과 같이 주어진다.

$$P_v = 10^{(A + \frac{B}{T+C})} \quad (2.1)$$

여기서 P_v 는 증기압(mmHg)이고 T 는 온도(°C)이다.

프로판에 대한 증기압(psia) 대 온도(°F) 데이터는 부록 B의 표 B.5에 주어져 있다. 이 데이터들을 증기압은 mmHg로 온도는 °C로 바꾼 다음에

- (a) 식 (2.1)에 대한 비선형 회귀분석을 이용하여 주어진 데이터들로 부터 Antoine 식의 매개인자와 각 인자들의 95% 신뢰구간을 결정하라.
- (b) Antoine 식에 대한 총괄 분산을 계산하라.
- (c) Antoine 식에 대한 오차 그림을 준비하라.
- (d) 데이터들의 상관식에 대한 Antoine 식의 적절성과 데이터의 정밀도를 평가하라.

1.4 해

이 문제에서 고려될 Antoine 식의 형태는 비선형 형이다. 이에 대한 대안은 이 식의 log 형에 다중 선형 회귀분석을 적용하는 것이나, 이 변환은 문제 1.7에서 논의된 바와 같이 적절하지 않을 수도 없다.

1.5 비선형 회귀분석

일반적인 비선형 회귀분석은 식 (2.2)에 의해 정의되는 것과 같은 양함수적 대수 방정식 또는 모델식의 매개인자를 결정하는데 사용될 수 있다.

$$y = f(x_1, x_2, \dots, x_n; a_1, \dots, a_m) \quad (2.2)$$

여기서 유일한 종속변수는 y 이고, n 개의 독립변수들은 x_1, x_2, \dots, x_n 이고, m 개의 매개인자들은 a_1, \dots, a_m 이다. 일반적으로 앞의 식에서 실험 오차들은 일정한 분산값을 가지며 정규적으로 분포하고 있다고 가정된다. 비선형 회귀분석 알고리즘은 특정 모델에 대하여 다음과 같이 주어지는 최소 제곱 목적함수 (LS)를 최소화하는 매개인자를 결정한다.

$$LS = \sum_{i=1}^N (y_{i(obs)} - y_{i(calc)})^2 \quad (2.3)$$

여기서 N 은 데이터의 갯수이고, (obs)와 ($calc$)는 종속변수의 관측값과 계산값을 의미한다.

모델 분산의 총괄값 σ^2 는 다음과 같은 형태로 계산된다.

$$\sigma^2 = \frac{\sum_{i=1}^N (y_{i(obs)} - y_{i(calc)})^2}{\nu} = \frac{LS}{\nu} \quad (2.4)$$

여기서 ν 는 자유도이고, 데이터의 개수에서 모델 매개인자의 수를 뺀 ($N - m$)으로 정의된다. 따라서 최소 제곱 목적 함수를 최소화하는 알고리즘은 분산도 최소화 한다. 여러 모델 식의 적합성을 비교하는데 분산이 종종 사용된다.

어떤 특정 모델의 난점을 체계적으로 나타내기 위하여 널리 사용되는 도식적 표현을 오차 그림이라 부른다. 이 도표는 다음과 같이 계산되는 모델에서 계산된 오차, ϵ_i 를 종속 변수의 관측값, $y_{i(obs)}$ 에 대하여 그림으로 나타내는 것이다.

$$\epsilon_i = (y_{i(obs)} - y_{i(calc)}) \quad (2.5)$$

(a) Antoine 식의 비선형 회귀분석 우선 psia 단위의 압력을 P로, °F 단위의 온도를 TF로 입력한다. Antoine 식은 mmHg 단위의 압력을 °C 단위의 온도에 대하여 상관지므로, 다음과 같이 mmHg와 °C 단위의 압력과 온도, Pv와 TC를 새로 계산하여 추가한다.

$$Pv = 51.715 * P \quad (2.6)$$

$$TC = (TF - 32)/1.8 \quad (2.7)$$

위의 변수들을 이용하는 일반적인 형태의 비선형 방정식은 다음과 같이 입력된다.

$$Pv = 10 \wedge (A + B/(TC + C)) \quad (2.8)$$

이 식의 매개인자들에 대한 초기 추산값은 타당한 근사에 근거해야 한다. 이 경우에, 표 1.10의 데이터는 $A = 6$, $B = -1000$, $C = 200$ 이 타당함을 제시하여 준다. 그림 2.1은 결과를 보여준다. 표 2.1에 주어진 신뢰 구간은 통계 해석에 근거하여 구해졌다 (신뢰 구간에 대한 자세한 사항은 문제 2.14를 참조하라).

표 2.1: 프로판에 대한 Antoine 식의 매개인자에 대한 95% 신뢰구간

매개인자	값	95% 신뢰구간	하한	상한
A	7.00188	0.0540358	6.94785	7.05592
B	-899.474	29.9126	-929.387	-869.561
C	260.446	4.58219	255.863	265.028

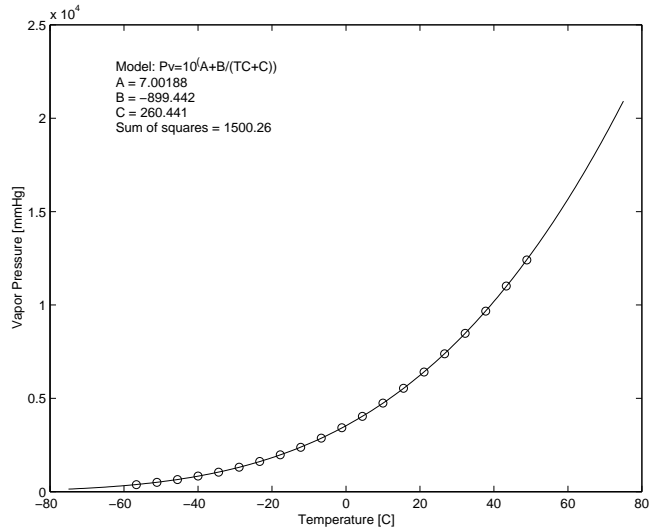


그림 2.1: Antoine 식에 대한 비선형 회귀분석 결과

(b) **Antoine 식에 대한 총괄 분산** 분산은 식 (2.4)로 부터 쉽게 계산된다. 이 식에서 자유도는 $(N - m)$ 으로 주어진다. 이 문제에서, 데이터의 개수, N 은 20이고, 매개인자의 개수 m 은 3이다. LS 는 제공의 함으로 그림 2.1에 1500.26으로 주어져 있다. 따라서 이 모델의 총괄 분산은 $1500.26/17$ 또는 88.251이다.

(c) **Antoine 식에 대한 오차 그림** 식 (2.5)에 의하여 정의된 오차를 나타낸 오차 그림은 그림 2.2에 주어져 있다.

(d) **데이터의 정확도와 Antoine 식의 적절성** 오차 그림은 경향성이 약하지만 증기압 P_v 가 클수록 오차가 커지는 경향을 보여준다. 증기압의 작은 데이터가 오차가 작다는 것은 확실하다. 이는 실험 오차가 측정된 증기압에 무관하지 않다는 것을 나타내는 것일 수도 있다. $\log P$ 를 종속 변수로 사용하는 Antoine 식의 회귀분석이 부가적인 비선형 회귀 분석으로 고려하는데 유용할 수도 있다. (이 경우는 문제 2.2에서 고려되었다.)

비선형 회귀분석의 신뢰구간은 비교적 좁다. 이러한 사실은 현재와 같은 형태의 Antoine 식이 실험 데이터의 범위에 걸쳐서 프로판의 증기압 데이터에 대한 적합한 상관관계를 제공한다는 것을 나타낸다.

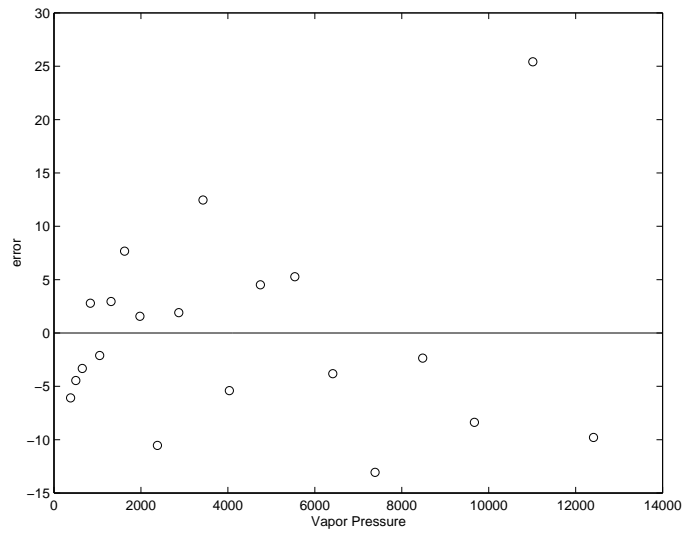


그림 2.2: 프로판에 대한 Antoine 식의 오차 그림



이 문제에 대한 MATLAB 스크립트 파일은 CHAP2의 p201.m에 주어져 있다.

제 2 절 여러 탄화수소에 대한 Antoine 식의 매개인자

2.1 개념 설명

온도에 따른 증기압 데이터를 상관시키기 위하여 Antoine 식의 직접 사용

2.2 사용된 수치 해법

총괄 분산 및 각 매개인자의 신뢰구간의 결정과 함께 일반적인 대수 표현의 비선형 회귀 분석.

2.3 문제 설명

(부록 B의) 표 B.7은 여러 탄화수소에 대하여 온도의 함수로 증기압 데이터를 나타낸다. 어떤 설계 과정에서 이 데이터들을 연관시키는 식이 필요하다.

- (a) 표 B.7에서 탄화수소들 중의 하나를 선택하라. Antoine 식의 매개인자와 각 인자들의 95% 신뢰구간, 총괄 분산을 구하라. 오차를 도시하라. 회귀분석은 문제 2.1에서 논의된 바와 같이 행하여야 한다.
- (b) 종속 변수를 $\log P$ 로 하여 (a)를 반복하라.
- (c) 문제 1.7에서 기술된 선형 회귀분석을 사용하여 (a)를 반복하라.
- (d) 데이터들의 상관식에 대한 (a), (b), (c)로 부터 나오는 Antoine 식의 적절성을 평가하라. 어떤 식이 선택되어야 하는가?



표 B.7의 데이터에 대한 MATLAB 데이터 파일은 tableb07.mat로 주어져 있다.

제 3 절 프로판의 열역학 및 전달현상 물성의 상관관계

3.1 개념 설명

열 용량, 열 전도도, 점도 및 증발 잠열에 대한 상관관계.

3.2 사용된 수치 해법

선형화 및 함수 변환이 포함된 선형 및 비선형 회귀분석.

3.3 문제 설명

표 B.8에서 B.11에는 온도의 함수로 프로판의 여러 물성들(열 용량, 기체의 열 전도도, 점도, 액체의 증발 잠열)이 주어져 있다.

표 B.8에서 B.11에 주어진 물성에 대하여 화학 공학 문헌들에서 추천되는 표현을 사용하여 적절한 상관관계를 구하라.

3.4 해

다양한 회귀분석법이 이 문제에서 사용된다. 이 문제에서 사용될 회귀분석을 설명하는 앞의 문제들은 다항식 맞춤에 대한 문제 1.3, 선형 회귀분석에 대한 문제 1.3과 1.7, 비선형 회귀분석에 대한 문제 2.1등 이다. 분산, 95% 신뢰구간, 오차 그림등과 같은 주요한 통계학적 척도를 사용하여 이용가능한 상관식 가운데서 선택하는 것이 중요하다. 이 문제는 어떻게 선택하는 가를 설명하여 준다.

경험 및 이론적인 고려를 통하여 어떤 형태의 상관관계가 서로다른 물리적 및 열역학적 특성의 온도 의존관계를 나타내는 데 가장 적합한 지를 보여준다. 이와 같은 정보는 문헌에서 얻을 수 있고, 이 문제의 해에서 다양한 상관관계를 전개하는데 사용될 수 있다.

(a) 기체의 열 용량 50K와 1500K 사이의 온도에서 프로판 기체의 열용량은 표 B.8로 주어져 있다. Perry et al. [4]에 따르면 기체의 열용량은 다음과 같은 간단한 다항식으로 흔히 나타내어진다.

$$C_p = a_0 + a_1T + a_2T^2 + a_3T^3 + \dots \quad (2.9)$$

표 B.8의 데이터를 MATLAB에 입력하고 polyfit을 사용하면 그림 2.3에 주어지고 표 2.2에 정리되어 있는 것과 같은 결과를 얻을 수 있다.

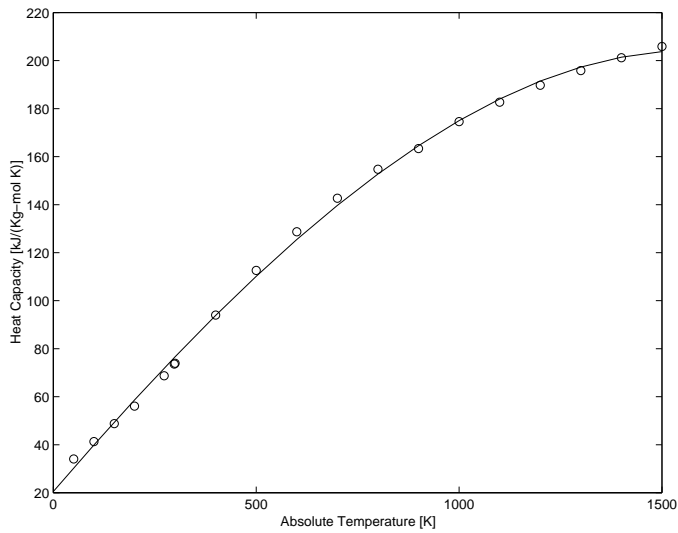


그림 2.3: 프로판의 열 용량에 대한 3차 다항식 맞춤

표 2.2: 프로판에 대한 다항식 계수 및 분산

차수	2	3	4	5
a_0	17.7427	20.5235	26.7577	31.0453
a_1	0.217787	0.196471	0.123020	0.0515068
a_2	-6.164280e-5	2.68480e-5	0.000188971	0.000508754
a_3		1.51289e-8	-2.35697e-7	-7.900072e-7
a_4			7.22918e-11	4.7874e-10
a_5				-1.1e-13
분산	6.81537	6.00692	1.86662	0.662227

1차 직선을 제외한 모든 다항식이 길보기에는 그림 2.3과 비슷한 형태로 데이터를 나타낸다. 하지만 표 2.2의 결과는 5차 다항식의 분산이 다른 다항식의 분산에 비하여 현저히 작은 값을 보여준다. 이는 이 상관식이 데이터를 가장 잘 나타냄을 보여주는 것이다. 5차 다항식의 매개인자 모두가 0과 확실히 다르다는 것은 MATLAB에서 얻을 수 있는 매개인자들에 대한 95% 신뢰구간으로써 입증된다. 매개인자들에 대한 95% 신뢰구간은 표 2.3에 정리되어 있다. 모든 신뢰구간들이 각각의 매개인자 값의 절대값들 보다 작기 때문에 다항식의 모든 항들이 데이터 상관관계에서 유용하다.

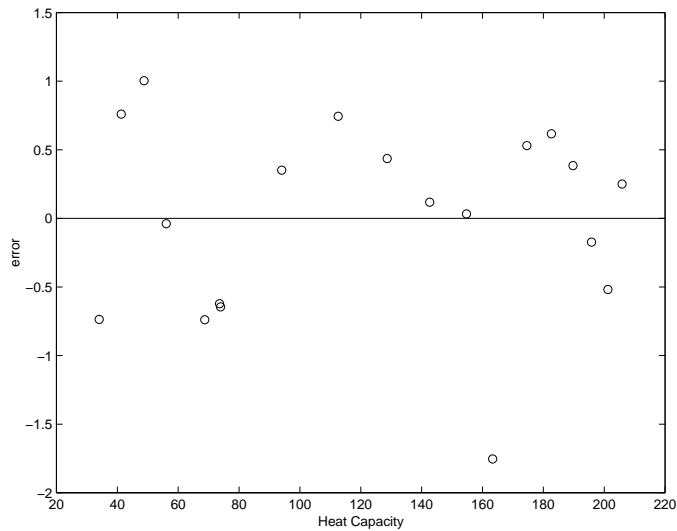


그림 2.4: 5차 다항식으로 나타내어지는 열 용량에 대한 오차 그림

표 2.3: 프로판의 열용량 데이터를 나타내는 5차 다항식의 매개인자 값 및 95% 신뢰구간

매개인자	값	0.95 신뢰구간
a_0	31.0453	2.74902
a_1	0.0515068	0.0353588
a_2	0.000508754	0.000143145
a_3	-7.90007e-7	2.37871e-7
a_4	4.7874e-10	1.7144e-10
a_5	-1.1e-13	4e-14

신뢰구간을 살펴보는 것에 덧붙여서, 그림 2.4에 주어진 것과 같은 5차 다항식에 대한 오차 그림을 살펴보는 것도 아주 유용하다. $C_{p(obs)} - C_{p(calc)}$ 로 주어지는 오차는 그림에서 보듯이 무작위적으로 분포함을 알 수 있다. 이는 오차에는 어떤 특정한 경향이 없으며, 이러한 경향은 측정된 열용량 데이터에 대한 다항식 상관관계의 타당성을 더욱 더 입증하는 것이다.

그림 2.4의 오차 그림을 그림 2.5에 주어지는 3차 다항식에 대한 오차 그림과 비교하는 것은 흥미있는 일이다. 3차 다항식의 상관관계는 오차의 크기가 훨씬 크고, 오차가 진동

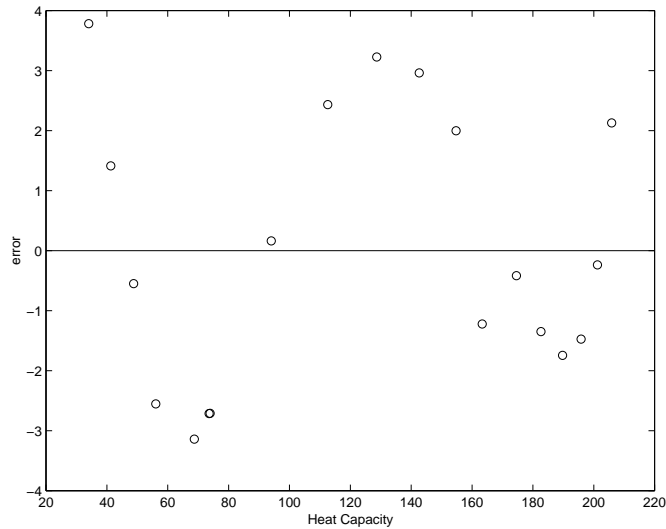


그림 2.5: 3차 다항식으로 나타내어지는 열 용량에 대한 오차 그림

하는 형태를 보여준다. 이는 데이터를 만족스럽게 나타내기 위해서는 더 많은 매개인자가 필요하다는 것을 나타내는 것이다.

(b) 열 전도도 -40°F 와 200°F 사이의 온도에서 프로판 기체의 열 전도도는 표 B.9로 주어져 있다. Perry et al. [4]는 좁은 온도 범위에 걸쳐서 저압 기체의 열 전도도는 1차 다항식인 직선 식으로 상당히 잘 나타낼 수 있음을 언급하였다.

2차, 3차 및 고차 다항식에 대한 도시가 모두가 겹보기에는 열전도도 데이터를 아주 비슷하게 나타내고 있음을 보여준다. 계수 값, 95% 신뢰구간, 1차에서 3차 까지의 다항식에 대한 분산등 표 2.4에 주어진 결과들이 결과들을 감별하기 위한 해석에 필요하다. 이 표는 1차 다항식의 분산과 2차 다항식의 분산 사이에는 현저한 감소가 있음을 보여준다. 2차 다항식과 3차 다항식의 분산 사이에는 약간의 차이 밖에 보이지 않는다. 3차 다항식의 계수 a_3 의 95% 신뢰구간은 매개인자 자체의 절대값 보다 크다. 이러한 사실은 2차 다항식 대신에 3차 다항식을 사용하는 것에 대한 통계학적 정당성이 없다는 것을 나타낸다. 그림 2.6과 2.7은 각각 1차 및 2차 다항식에 대한 오차 그림을 보여준다. 1차 다항식의 경우에는 오차들에는 명백한 경향이 있는 반면에, 2차 다항식의 경우에는 오차들은 무작위적으로 분포한다. 따라서 수치적, 도식적 통계학적 결과들로부터 2차 다항식이 데이터들을 잘 나타낸다고 결론지을 수 있다.

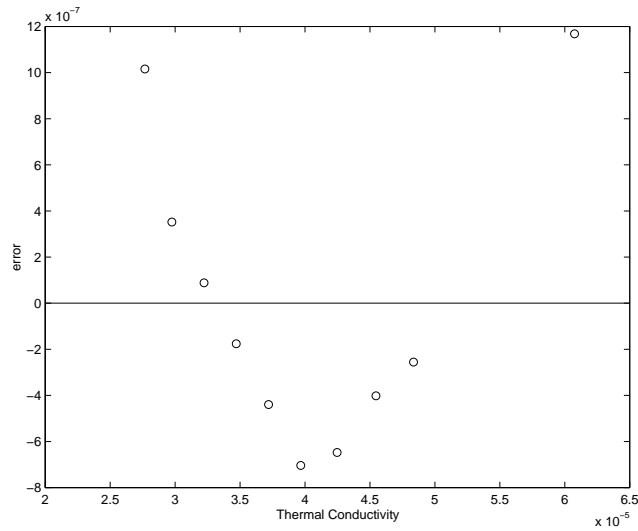


그림 2.6: 1차 다항식으로 나타내어지는 열 전도도에 대한 오차 그림

표 2.4: 기체 열전도도 데이터에 대한 다항식 맞춤 결과

	1차 다항식		2차 다항식		3차 다항식	
	값	95% 신뢰구간	값	95% 신뢰구간	값	95% 신뢰구간
a_0	32.1419	0.656920	32.207448	0.096695	32.1724	0.120205
a_1	0.137198	0.007419	0.119970	0.002347	0.119013	0.003049
a_2			1.14701e-4	1.384e-5	1.4343e-4	5.984e-5
a_3					-1.1924e-7	2.415e-7
σ^2	0.4869		0.009966		0.009353	

넓은 온도 범위에 대해서는 Perry et al. [4]은 열전도도 k 를 $(T)^n$ 에 대하여 상관짓는 것을 추천하였다. 여기서 T 는 절대온도이고 $n \approx 1.8$ 이다. 이 상관식은 c 와 n 이 매개인자인 식 (2.10)을 사용하는 비선형 회귀분석으로 직접 결정된다.

$$k = cT^n \quad (2.10)$$

MATLAB을 사용하여 얻어진 상관식은 그림 2.8에 주어져 있다.

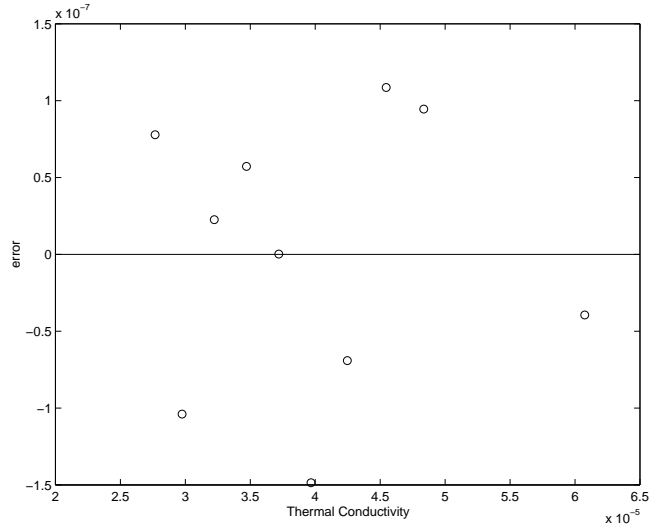


그림 2.7: 2차 다항식으로 나타내어지는 열 전도도에 대한 오차 그림

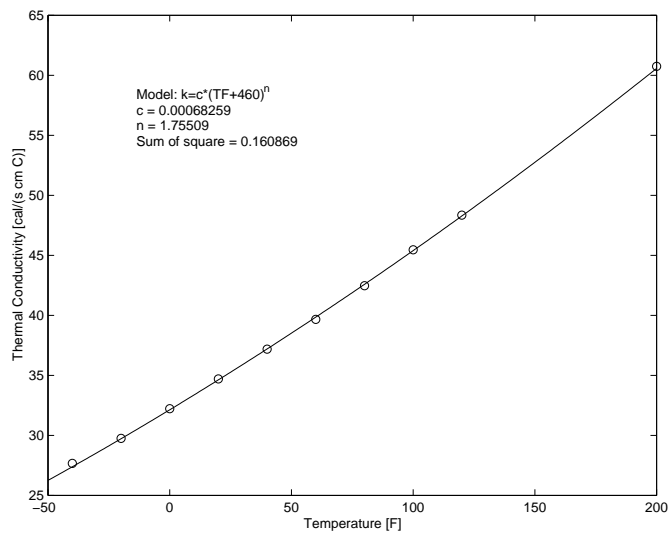


그림 2.8: 프로판의 열전도도에 대한 비선형 상관관계

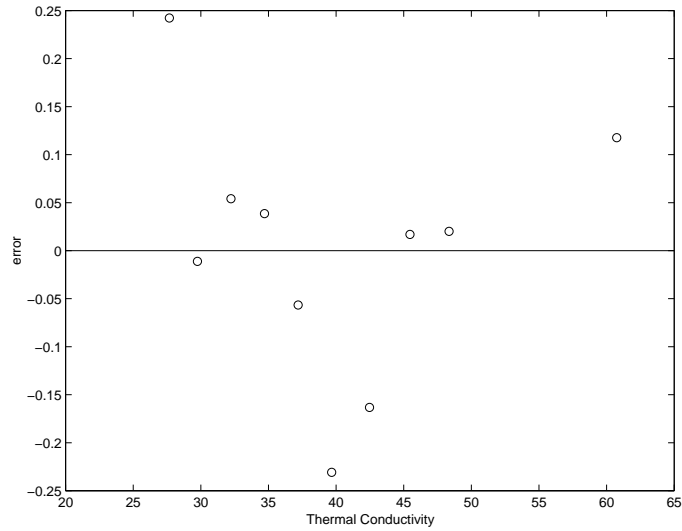


그림 2.9: 비선형 방정식으로 나타내어지는 열 전도도에 대한 오차 그림

지수 n 이 Perry et al. [4]이 제안한 것과 같이 거의 1.8이 되는 것은 흥미있는 일이다. 95%의 신뢰구간을 포함하는 전체 상관식은 다음과 같이 된다.

$$k = (6.8259 \times 10^{-4} \pm 8.119 \times 10^{-5})T^{1.75509 \pm 0.01888} \quad (2.11)$$

그림 2.9는 비선형 방정식에 대한 오차 그림을 보여준다. 그림에서 보듯이 오차들은 무작위적으로 분포하고 있다. 따라서 데이터가 주어진 영역에서는 프로판의 열전도도는 식 (2.11)에 의하여 잘 상관지어 진다고 할 수 있다.

2차 다항식과 비선형 방정식이 모두 적합하므로 이 열전도도 상관식에 대한 두 상관식 사이에서의 선택은 어려운 일이다. 우선 각 상관식의 총괄 분산을 고려하자. 표 2.4로부터 2차 다항식의 분산 0.009966은 비선형 방정식의 분산 0.020109과 비교할 만 하다. 두번째 분산은 그림 2.8의 제공의 합을 자유도인 (10 데이점 - 2 매개인자) 8로 나눈 값을 이용하는 식 (2.4)로부터 계산되었다. 비록 이 분산의 차가 통계학적인 의미를 가지지 않을지 모르지만 2차 다항식이 열전도도 데이터에 대한 더 나은 상관식을 제공한다는 것을 보여준다.

(c) 액체의 점도 Perry et al. [4]에 의하여 추천된 액체의 점도에 대한 상관관계는 증기압에 대한 Antoine 식과 비슷하다.

$$\log(\mu) = A + B/(T + C) \quad (2.12)$$

여기서 μ 는 점도이고, A, B, C 는 매개인자이다. 만약 T 가 Kelvin 온도로 나타내어지면, 매개인자 C 는 $C = 17.71 - 0.19T_b$ 로 근사된다. 여기서 T_b 는 Kelvin 온도로 주어지는 끓는점이다. 프로판의 끓는점은 231K이므로, C 는 -26.18이 된다.

MATLAB의 비선형 회귀분석을 사용하여 표 B.10의 데이터로부터 식 (2.12)를 사용하여 매개인자를 계산할 수 있다. 문제 1.7에서 증기압에 대하여 A, B, C 를 결정했던 것처럼 식 (2.12)를 선형화할 수도 있다. 비선형 회귀분석에서는 매개인자들에 대한 합리적인 초기 추산치가 필요하고 A 와 B 에 대한 추산치들을 얻을 수 없으므로, 이 추산치들을 얻기 위하여 초기 추산치가 필요하지 않은 다중 선형 회귀분석이 사용될 것이다. 계속하여 이 값들을 초기 추산치로 사용하여 비선형 회귀분석이 수행된다.

따라서, 식 (2.12)는 C 값을 -26.18로 사용하여 선형화될 수 있다. MATLAB 프로그램에서 $\log\mu = \log_{10}(\mu)$, $TK = TC + 273.15$, $\text{invTKplusC} = 1/(TK - 26.18)$ 로 주어지는 변수의 변환이 필요하다. 새로운 변수를 사용하는 선형화된 표현은 다음과 같이 된다.

$$\log\mu = A + B * \text{invTKplusC} \quad (2.13)$$

독립변수로 invTKplusC 를, 종속변수로 $\log\mu$ 를 사용하는 선형 회귀분석의 결과는 $A = -1.3439$, $B = 142.07$ 이 된다. 이 값들과 $C = -26.18$ 이 식 (2.12)을 비선형 회귀분석 하는데 필요한 초기 추산치가 된다.

비선형 회귀분석의 결과가 그림 2.10에 주어져 있다. 그림에서 보듯이 비선형 회귀분석의 결과는 꽤 좋은 것처럼 보인다. 하지만 그림 2.11의 오차 그림은 오차 분포에 진동하는 경향을 명확히 보여준다. 이 그림은 3 매개인자 방정식이 넓은 온도 범위에 걸쳐서 점도 데이터를 나타내기에는 부적절하다는 것을 보여준다.

Reid et al. [6]에서 사용되는 다음과 같은 4 매개인자 방정식이 액체의 점도에 대한 또 다른 가능한 방정식을 제공한다.

$$\log(\mu) = A + B/T + CT + DT^n \quad (2.14)$$

위의 식은 비선형 회귀분석을 통하여 맞춰질 수 있는 5개의 매개인자를 포함한다. 위와 같은 식에 대하여 타당한 초기 추산치를 얻기 위하여 선형 회귀분석을 사용하는 것이 유용하다. 지수 n 값을 임의의 값 0.5로 고정하면 선형화는 쉽게 된다. 이와 같은 조작을

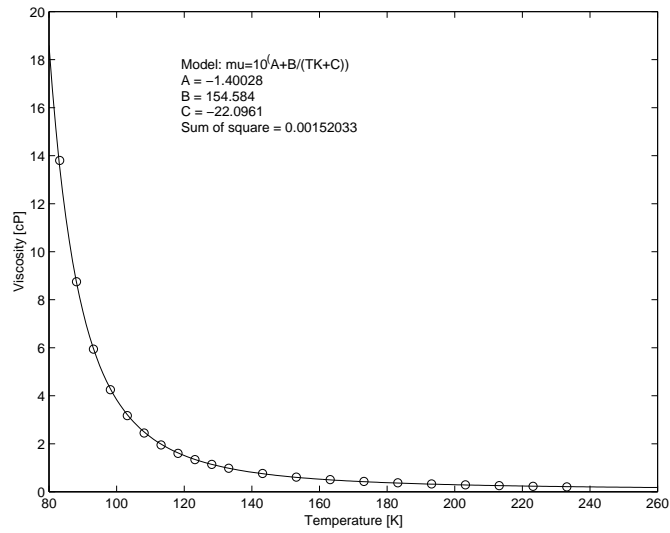


그림 2.10: 점도의 관측값과 계산값. Antoine 식 매개인자들은 비선형 회귀분석을 통하여 얻어졌다.

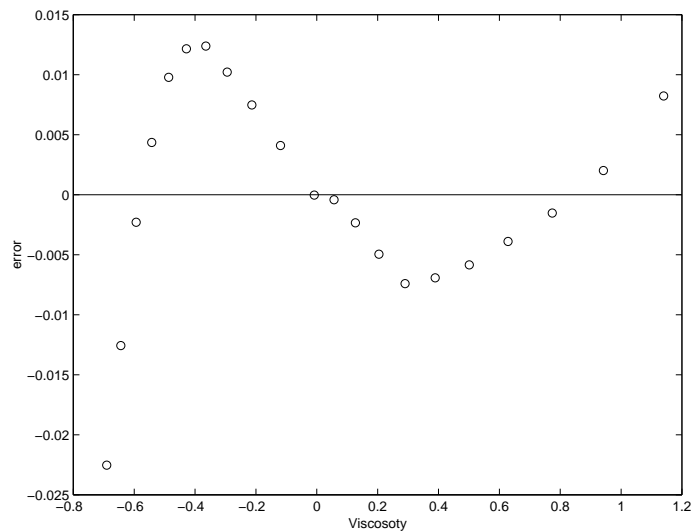


그림 2.11: 비선형 방정식 (2.12)로 나타내어지는 점도에 대한 오차 그림

거친 후의 선형 회귀분석의 최종결과는 $A = -11.1502$, $B = 467.144$, $C = -0.028897$, $D = 0.994892$ 및 $n = 0.5$ 이다. 몇 매개인자의 신뢰구간이 0을 포함하고 있기 때문에 이 값들을 식 (2.14)의 비선형 회귀분석에 사용하기에는 곤란하다.

문헌에서 제안되는 두 상관관계에 대한 결과는 아주 만족스럽지 않고, 이 데이터들은 일반적으로 사용되는 점도의 온도의존성에 대한 관계로는 잘 나타내지 못한다고 결론내릴 수 있다. 이 데이터들에 대해서는 다항식 형태의 회귀분석을 고려하여야 한다.

(d) 증발열 프로판에 대한 증발열 데이터는 표 B.11에 -70°F 에서 120°F 범위에서 주어져 있다. 증발열은 Watson 관계에 근거를 둔 다음과 같은 식으로 상관지어질 수 있다 (Perry et al. [4]):

$$\Delta H = A(T_C - T)^n \quad (2.15)$$

Watson이 제안한 n 값은 0.38이나, n 값은 실험 데이터의 회귀분석을 통하여 결정될 수 있다. 프로판의 임계온도는 666R이다.

위의 식은 직접 비선형 회귀분석에 사용될 수 있고, 양변에 log를 취해서 다음과 같이 선형화 될 수 있다.

$$\log(\Delta H) = \log A + n \log(T_C - T) \quad (2.16)$$

두 회귀분석이 행해질 것이므로, 표 B.11의 원래 데이터는 TF와 deltaH로 입력되고, 변환 변수 $\text{TR}=\text{TF}+460$, $\log\Delta\text{H}=\log_{10}(\Delta\text{H})$, $\log\text{Tc}=\log_{10}(666-\text{TR})$ 이 선형 회귀분석에서 사용된다.

선형 회귀분석의 결과가 그림 2.12에 주어져 있고, 증발열에 대한 결과식은 다음과 같다.

$$\Delta H = 30.333(666 - T)^{0.3248} \quad (2.17)$$

선형 회귀분석 결과를 초기 추산치로 사용하는 비선형 회귀분석의 결과는 다음과 같다.

$$\Delta H = 30.083(666 - T)^{0.3264} \quad (2.18)$$

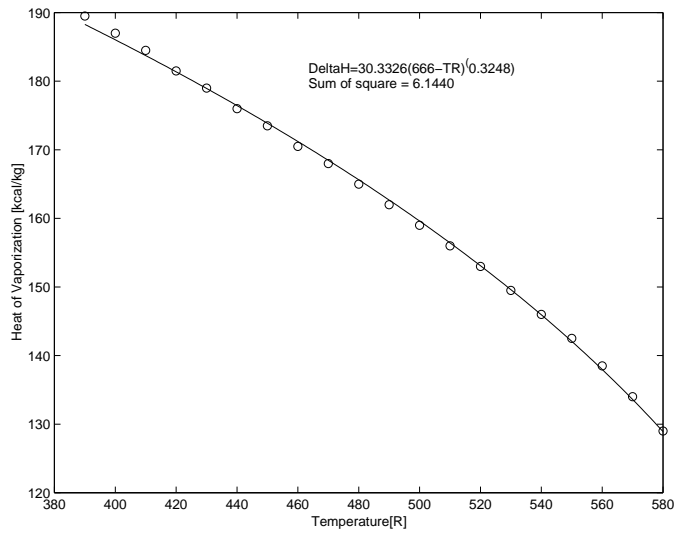


그림 2.12: 선형 회귀분석에 의한 증발열 데이터의 상관관계

두 결과가 아주 유사하다. 매개인자 값들에서의 약간의 차이는 선형 회귀분석에서는 최소 제곱 목적함수가 $\log(\Delta H)$ 에서 계산되는 반면에 비선형 회귀분석에서의 목적함수는 ΔH 로부터 계산되기 때문에 나타난다.

그림 2.13에 주어진 비선형 회귀분석에서의 오차 그림은 주기적이다. 이는 이 식에 의하여 데이터들이 잘 나타내어 질 수 없음을 의미한다. 선형 회귀분석에서도 비슷한 경향이 나타난다.

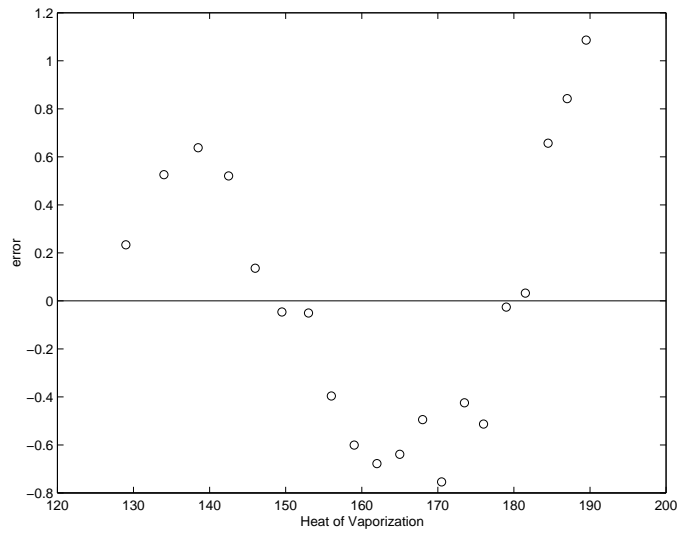


그림 2.13: 비선형 회귀분석에 의한 증발열에 대한 오차 그림



표 B.8에서 B.11까지의 MATLAB 데이터 파일은 CHAP2의 tableb08.mat에서 tableb11.mat로 주어져 있다.

제 4 절 선택된 물성들의 온도 의존성

4.1 개념 설명

열역학 및 전달 현상 물성들의 상관식 및 표시.

4.2 사용된 수치 해법

실험 데이터에 서로 다른 형태의 곡선을 맞춤, 선형 표현을 얻기위한 변수들의 변환.

4.3 문제 설명

표 B.12에서 B.15는 온도의 함수로 여러 화합물의 다양한 물성치 값들을 나타낸다.

이 화합물들의 물성치들의 온도 의존성을 나타내는 최상의 상관식을 구하여라.



표 B.12에서 B.15의 데이터에 대한 MATLAB 데이터 파일은 tableb12a.mat ... tableb12d.mat, tableb13a.mat ... tableb13g.mat, tableb14a.mat ... tableb14f.mat, tableb15a.mat ... tableb15f.mat로 주어져 있다.

제 5 절 차원 해석으로부터 얻어진 열전달 상관식

5.1 개념 설명

무차원 군을 사용하는 열전달 상관식.

5.2 사용된 수치 해법

데이터의 선형 및 비선형 회귀분석 및 변환 함수의 사용.

5.3 문제 설명

공학적 데이터를 상관시키는 중요한 도구는 차원 해석을 사용하는 것이다. 차원 해석은 서로 독립적인 무차원 군을 결정하게 되고, 이 무차원 군은 특정 문제에 중요하게 될 수도 있다. 선형 및 비선형 회귀분석은 실험 데이터에 대한 무차원 수의 상관관계를 결정하는데 아주 유용할 수 있다.

Geankoplis [3]는 Buckingham 법을 사용하여 관 내의 열전달에 대한 처리를 고려하여, Nusselt 수가 Reynolds 수와 Prandtl 수의 함수가 될 것이라는 결과를 얻었다.

$$Nu = f(Re, Pr) \quad \text{or} \quad \frac{hD}{k} = f\left(\frac{Dv\rho}{\mu}, \frac{C_p\mu}{k}\right) \quad (2.19)$$

식 (2.19)에 의하여 제안되는 대표적인 상관식은 다음과 같다.

$$Nu = aRe^bPr^c \quad (2.20)$$

여기서 a, b, c 는 실험 데이터로부터 결정될 수 있는 매개인자들이다.

관 내의 난류 흐름에서의 열전달에 대하여 널리 쓰이는 상관식은 Sider-Tate [7]식이다.

$$Nu = 0.023Re^{0.8}Pr^{1/3}(\mu/\mu_w)^{0.4} \quad (2.21)$$

여기에는 무차원 점도 비가 더해졌다. 이 무차원 점도 비(μ/μ_w)는 유체의 평균 온도에서의 점도와 벽면 온도에서의 점도의 비이다.

표 B.16은 Williams와 Katz [8]가 바깥 지름이 $3/4$ in 관의 외부로의 열전달에 대하여 얻은 데이터들 중의 일부이다. 여기서 Re , Pr , (μ/μ_w) 와 Nu 는 측정된 무차원 수들이다.

- (a) 다중 선형 회귀분석을 사용하여 표 B.16의 데이터를 나타내는 식 (2.20)과 (2.21) 형태의 함수 꼴의 매개인자 값을 결정하라.
- (b) 비선형 회귀분석을 사용하여 (a)를 반복하라.
- (c) 어떤 함수 꼴과 매개인자 값들이 이 데이터에 대한 상관식으로 추천되어야 하는가? 그 이유를 설명하라.

5.4 해

편의상 식 (2.20)과 (2.21)의 함수 꼴들을 다음과 같이 쓰자.

$$Nu = a_i Re^{b_i} Pr^{c_i} \quad (2.22)$$

$$Nu = d_i Re^{e_i} Pr^{f_i} Mu^{g_i} \quad (2.23)$$

여기서 $i = 1$ 은 선형 회귀분석 결과의 매개인자 값들을 나타내고, $i = 2$ 은 비선형 회귀분석 결과의 매개인자 값들을 나타낸다. Mu 는 점도 비를 나타낸다.

(a) 식 (2.22)나 (2.23)의 선형 회귀분석은 선형화된 형태로 변환해야 한다. 각 식의 양변에 자연 로그(\ln)를 취하면 선형화는 쉽게 된다. 변환된 식들은 다음과 같이 된다.

$$\ln Nu = \ln a_i + b_i \ln Re + c_i \ln Pr \quad (2.24)$$

$$\ln Nu = \ln d_i + e_i \ln Re + f_i \ln Pr + g_i \ln Mu \quad (2.25)$$

Re , Pr , Mu 와 Nu 로 정의된 표 B.16의 데이터는 MATLAB으로 직접 입력된다. 선형 회귀분석에서 필요한 추가적인 변수들은 다음과 같이 각 변수에 자연 \log 를 취한 값들이다. $\ln Re = \log(Re)$, $\ln Pr = \log(Pr)$, $\ln Mu = \log(Mu)$, $\ln Nu = \log(Nu)$. 선형 회귀분석의 결과가 표 2.5에 정리되어 있다. a_1 과 d_1 에 대하여 제시된 값들은 이 매개인자 값들의 자연 로그 값을 주는 회귀분석 결과로부터 계산된 값들이므로 주의하라.

표 2.5: 회귀분석 결과 요약

매개인자	선형 회귀분석 식(2.24)와 (2.25)		비선형 회귀분석 식(2.22)와 (2.23)	
	값	95% 신뢰구간	값	95% 신뢰구간
a	0.662295	3.8794	0.165641	0.126393
b	0.539538	0.116063	0.663548	0.064737
c	0.245353	0.113943	0.341377	0.062209
d	0.534712	5.6326	0.149153	0.117505
e	0.558833	0.150668	0.673285	0.067260
f	0.252375	0.122989	0.328563	0.066706
g	-0.067723	0.316322	-0.177681	0.332506

표 2.5에서 식 (2.24)와 (2.25)에 대한 선형 회귀분석의 결과는 매개인자 a 를 d 와, b 를 e 와, c 를 f 와 비교할 때 일관성이 있게 된다. 95% 신뢰구간은 0을 포함하는 큰 값을 가지는 a , d 및 g 를 제외하고는 비교적 작다.

(b) 식 (2.22)와 (2.23)에 대한 직접적인 비선형 회귀분석은 초기 추산값으로 사용할 수 (a)의 선형 회귀분석 결과로부터 수렴된 값을 얻을 때 종료된다. 결과들이 표 2.5에 정리되어 있다. 이 경우 역시, 회귀분석의 결과는 매개인자 a 를 d 와, b 를 e 와, c 를 f 와 비교할 때 일관성이 있게 된다. 95% 신뢰구간은 0을 포함하는 큰 값을 가지는 g 를 제외하고는 비교적 작다.

(c) 가장 적합한 상관식을 선택하는데에는 많은 고려사항이 있다.

신뢰구간 주요 판단 기준은 각 매개인자에 대한 95% 신뢰구간이다. 매개인자에 비하여 신뢰구간이 아주 크면, 이 매개인자는 상관식에서 아마도 중요하지 않고 0으로 두어야 할 것이다. (a)와 (b)에서 행한 선형 및 비선형 회귀분석 두 경우에서 Mu 의 지수인 매개인자 g 의 경우이다. 이에 대한 또다른 설명은 Mu 는 회귀분석의 다른 변수에 의존할 수도 있다는 것이다. 이에 대한 설명은 이절의 뒷 부분에서 기술될 것이다.

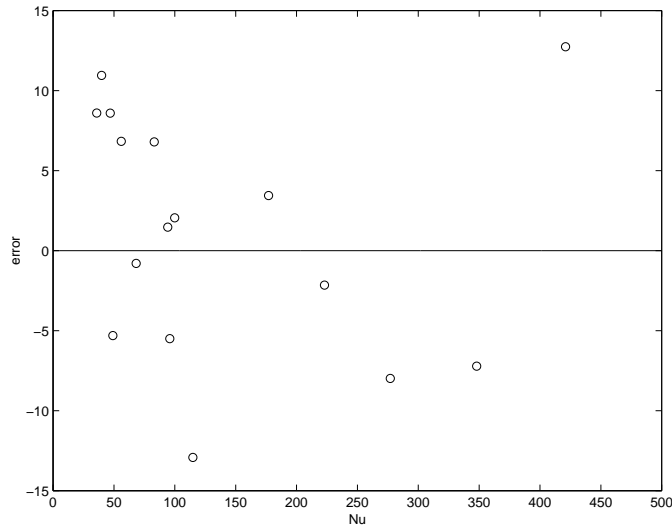


그림 2.14: 식 (2.22)를 사용하는 비선형 회귀분석에 의한 열전달 데이터에 대한 오차 그림

오차 그림 오차들이 무작위적으로 분포해야 되는 것과 같은 어떤 명백한 경향이 있는지를 결정하기 위하여 회귀분석의 오차도시를 살펴보는 것이 아주 유용하다. 그림 2.14는 식 (2.22)를 사용하는 비선형 회귀분석에 대한 대표적인 오차 그림이다. 이 오차도시는 오차들이 무작위적으로 분포하고 있음을 보여준다. 여기서 고려되는 어떤 선형 및 비선형 회귀분석에서도 특별한 오차 형태들이 나타나지 않으므로, 이 경우에는 오차 그림은 상관관계를 선택하는데 도움이 되지 못한다.

분산들의 비교 선형 회귀분석에서의 종속 변수는 $\ln(Nu)$ 이고, 비선형 회귀분석에서의 종속 변수는 Nu 이므로 이 상관관계들에 대한 분산을 비교할 때에는 특히 주의해야 한다. 이 문제에 대한 Nusselt 수에 근거한 분산은 다음과 같이 정의된다.

$$\sigma^2 = \frac{1}{\nu} \sum_{i=1}^N (Nu_{(obs)} - Nu_{(calc)})^2 \quad (2.26)$$

여기서 ν 는 데이터 점의 수에서 매개인자의 수를 뺀 자유도이다 ($N - m$). 이 문제에 대한 상대 오차 분산은 다음과 같이 정의된다.

$$\sigma_r^2 = \frac{1}{\nu} \sum_{i=1}^N \left(\frac{Nu_{(obs)} - Nu_{(calc)}}{Nu_{(obs)}} \right)^2 \quad (2.27)$$

이 분산들은 회귀분석이 끝난 후에 계산된다. 계산 결과들이 표 2.6에 요약되어 있다.

표 2.6: 열전달 상관관계에 대한 분산들

회귀분석 식	분산 σ^2	상대분산 σ_r^2
선형 회귀분석, 식 (2.24)	265.59	0.010278
선형 회귀분석, 식 (2.25)	234.71	0.011346
비선형 회귀분석, 식 (2.22)	68.276	0.016832
비선형 회귀분석, 식 (2.23)	66.685	0.014765

표 2.6의 분산들을 보면, 비선형 회귀분석으로 얻은 두 상관관계가 선형 회귀분석으로 얻은 두 상관관계 보다 좋은 결과를 줌을 알 수 있다. 이는 큰 값의 Nu 가 회귀분석에 더 많은 영향을 미치기 때문이다. 상대 분산의 계산 결과를 보면, 두 선형 회귀분석의 결과가 작은 상대 분산값을 준다. 이는 선형 회귀분석에서는 $\ln(Nu)$ 를 종속 변수로 사용하여서 Nu 값이 큰 데이터 점들의 효과가 줄어들었기 때문이다. 이와 같은 결과로부터 일반적으로 “상대 오차들 최소화 하려면 log 변환이 유용하다”는 결론을 내릴 수 있다.

따라서, 회귀분석 식의 선택은 오차가 상대적인지 또는 측정 Nu 값에 비례하는지에 의존한다. 이 정보가 이 문제에서는 알려져 있지않기 때문에, 분산 및 상대 분산의 계산 만으로는 선택을 할 수가 없다.

변수들 간의 가능한 상호 의존성 선형 및 비선형 회귀분석에서는 독립변수들이 서로 독립적이라고 가정한다. 가정된 독립변수들 사이에서의 가능한 의존성은 한 변수를 다른 변수에 대하여 도시해 봄으로써 알아볼 수 있다. 이 경우에는 0을 포함하는 Mu 의 큰 신뢰구간은 Mu 가 다른 변수들에 연관되어 있음을 알려준다. Mu 의 $\ln(Re)$ 에 대한 회귀분석은 그림 2.15에 보이는 바와 같이 확실한 의존성을 보여준다. 겉보기에는 점도의 비가 실험에서 Re 에 독립적으로 변하지 않으므로, Nu 에 대한 Mu 의 효과는 독립시킬 수 없다. 식 (2.23)의 점도 비의 지수에 대한 결론은 Williams와 Katz [8]의 데이터는 점도 비의 지수를 결정하기에는 불충분하다는 것이다.

최종 상관식 비선형 회귀분석은 Nu 값을 직접 계산할 수 있고, 식 (2.26)에서 계산된 분산이 직접 회귀분석에 적용될 수 있으므로, 이 데이터에 대하여 추천되는 식은 다음과 같

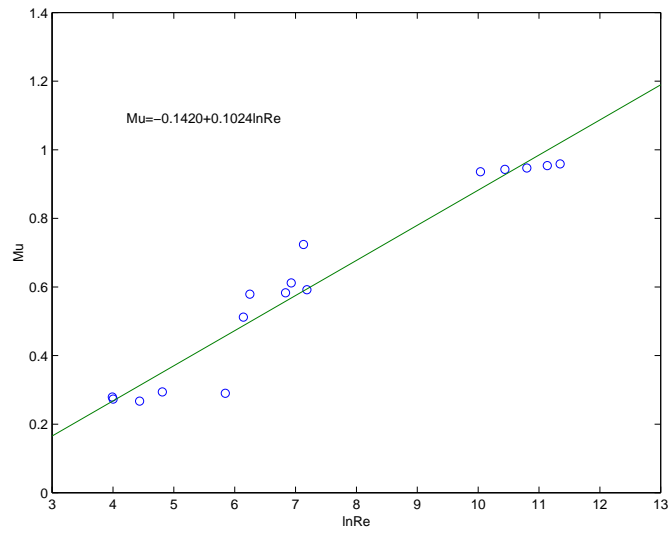


그림 2.15: Reynolds 수의 \ln 값, $\ln(Re)$ 에 대한 점도 비 μ/μ_w 의 도식



표 B.16의 데이터에 대한 MATLAB 데이터 파일은 tableb16.mat로 주어져 있다.

이 주어진다.

$$Nu = (0.1656 \pm 0.126) Re^{0.664 \pm 0.0647} Pr^{0.341 \pm 0.0622} \quad (2.28)$$

Mu 비의 효과를 조사하고, 이 데이터들의 회귀분석에 사용되는 최상의 분산값을 결정하기 위해서는 더 많이 실험을 해야한다.

제 6 절 관에서의 액체의 열전달 상관식

6.1 개념 설명

무차원 군을 사용하는 열전달 데이터의 상관식.

6.2 사용된 수치 해법

선형화 및 변환 함수를 사용하는 데이터의 선형 및 비선형 회귀분석.

6.3 문제 설명

표 B.17에서 B.19는 Sieder와 Tate [7]에 의하여 보고된 0.75-inch No. 16 BWG 안쪽 관과 1.25-inch 바깥쪽 철 관으로 된 동심 원통 열 교환기에서 열전달 데이터이다. 측정은 세계의 서로 다른 종류의 기름을 사용하여 행해졌다. 표에서 Nu , Re , Pr 로 나타내어진 무차원 군은 각각 Nusselt 수, Reynolds 수, Prandtl 수이다. (μ/μ_w) 로 표시된 열은 벽 온도에서의 점도에 대한 평균 온도에서의 점도의 비이다.

- (a) 표 B.17에서 B.19에 있는 열전달 데이터에 대한 가장 적합한 상관식을 결정하라. 다음과 같은 형태의 일반적인 Sieder와 Tate [7] 식을 평가하라.

$$Nu = a_0 Re^{a_1} Pr^{a_2} (\mu/\mu_w)^{a_3}$$

여기서 a_0 , a_1 , a_2 와 a_3 는 전체 데이터로부터 결정되어야 할 매개인자이다.

- (b) 여러분의 결과를 Sieder와 Tate [7]에 의하여 주어진 다음 식과 비교하라.

$$Nu = 0.023 Re^{0.8} Pr^{1/3} (\mu/\mu_w)^{0.14}$$



표 B.17에서 B.19까지의 데이터에 대한 MATLAB 데이터 파일은 tableb17.mat에서 tableb19로 주어져 있다.

제 7 절 유동층 반응기에서의 열전달

7.1 개념 설명

무차원 군을 사용하는 열전달 데이터의 상관식.

7.2 사용된 수치 해법

선형화를 사용하는 데이터의 선형 및 비선형 회귀분석.

7.3 문제 설명

Dow와 Jacob [2]은 유동화된 공기와 고체 혼합물과 수직 관벽 사이의 열전달에 대한 실험적인 데이터를 잘 나타내는 것으로 다음과 같은 무차원 식을 제안하였다.

$$Nu = a_1 \left(\frac{D_t}{L} \right)^{a_2} \left(\frac{D_t}{D_p} \right)^{a_3} \left(\frac{1 - \epsilon \rho_s C_s}{\epsilon \rho_g C_g} \right)^{a_4} \left(\frac{D_t G}{\mu_g} \right)^{a_5} \quad (2.29)$$

여기서

$$Nu = \frac{h_m D_t}{k_g} = \text{Nusselt 수}$$

$$h_m = \text{열전달 계수}$$

$$D_t = \text{관 지름}$$

$$D_p = \text{고체 입자 지름}$$

$$L = \text{가열되는 유동층 길이}$$

$$\epsilon = \text{유동층의 공극률}$$

$$G = \text{기체 질량 속도}$$

$$k_g, \rho_g, C_g, \mu_g = \text{기상의 물성}$$

$$C_s, \rho_s = \text{고상의 물성}$$

- (a) $\log(Nu)$ 에 대한 선형 회귀분석과 Nu 에 대한 비선형 회귀분석을 사용하여 식 (2.29)의 $a_1, a_2 \cdots a_5$ 의 매개인자 값들을 계산하기 위하여 표 B.20의 있는 데이터를 사용하라.
- (b) 데이터를 상관짓는데 어떤 회귀분석법이 가장 유용한가? 여러분의 선택을 신뢰구간, 분산, 오차 그림을 사용하여 정당화하라.
- (c) 이 문제에서 사용된 두가지 형태의 회귀분석법의 차이를 설명하라.



표 B.20의 데이터에 대한 MATLAB 데이터 파일은 tableb20.mat로 주어져 있다.

제 8 절 Margules 식을 사용하는 이성분 활동도 계수의 상관식

8.1 개념 설명

이성분 활동도 계수의 상관식을 위한 Margules 식에서 매개인자 추산.

8.2 사용된 수치 해법

선형 및 비선형 회귀분석, 회귀분석을 위한 데이터의 변환, 신뢰구간, 오차 그림, 제곱의 합의 계산 및 비교.

8.3 문제 설명

이성분 활동도 계수의 상관식을 위한 Margules 식은 다음과 같다.

$$\gamma_1 = \exp[x_2^2(2B - A) + 2x_2^3(A - B)] \quad (2.30)$$

$$\gamma_2 = \exp[x_1^2(2A - B) + 2x_1^3(B - A)] \quad (2.31)$$

여기서 x_1 과 x_2 는 각각 성분 1과 2의 몰분율이고, γ_1 과 γ_2 는 활동도 계수이다. 매개인자 A 와 B 는 특정 이성분 혼합물에 대한 상수이다.

식(2.30)과 (2.31)을 결합하면 다음과 같은 과잉 Gibbs 자유에너지에 대한 표현을 얻을 수 있다.

$$g = G_E/RT = x_1 \ln \gamma_1 + x_2 \ln \gamma_2 = x_1 x_2 (Ax_2 + Bx_1) \quad (2.32)$$

표 벤젠과 n -헵탄의 이성분 계에 대하여 여러 몰 분율에서 활동도 계수들은 식 (2.32)의 g 로부터 계산될 수 있다. 자유 매개인자가 없는 다중 선형 회귀분석이 A 와 B 값을 추산하는데 사용될 수 있다. 또 다른 방법은 식 (2.30)과 (2.32)을 더하고, 이 합에 비선형 회귀분석을 적용하여 A 와 B 를 결정하는 것이다.

표 2.7: 벤젠(1), n -헵탄(2) 계에 대한 활동도 계수

번호	x_1	γ_1	γ_2
----	-------	------------	------------

1	0.0464	1.2968	0.9985
2	0.0861	1.2798	0.9998
3	0.2004	1.2358	1.0068
4	0.2792	1.1988	1.0159
5	0.3842	1.1598	1.3059
6	0.4857	1.1196	1.0676
7	0.5824	1.0838	1.1096
8	0.6904	1.0538	1.1664
9	0.7842	1.0311	1.2401
10	0.8972	1.0078	1.4038

- (a) 벤젠, n -헵탄 계에 대한 Margulus 식의 A 와 B 를 결정하기 위하여 표 2.7의 데이터와 식 (2.32)에 대한 선형 회귀분석을 사용하라.
- (b) 식 (2.30)과 (2.31)의 합인 단일 식에 대한 비선형 회귀분석을 적용하여 A 와 B 를 추산하라.
- (c) 매개인자 신뢰구간, 오차 그림, 오차의 제곱의 합(두 활동도 계수에 대하여 계산된 최소 제곱 합)을 사용하여 (a)와 (b)의 회귀분석 결과를 비교하라.

8.4 해

(a) 과잉 Gibbs 에너지 식에 대한 선형 회귀분석 데이터들이 x_1 , γ_1 , γ_2 로 입력된다. 몰 분율 x_2 는 변환함수 $x_2=1-x_1$ 을 통하여 정의되어야 한다. Gibbs 에너지 함수는 $g=x_1 \log(\gamma_1)+x_2 \log(\gamma_2)$ 로 정의된다. 식 (2.32)는 다음과 같은 선형 형태로 변형 된다.

$$g = Ax_1x_2^2 + Bx_1^2x_2 = a_1X_1 + a_2X_2 \quad (2.33)$$

따라서 다중 선형 회귀분석에 필요한 두 변환 변수는 $X_1 = x_1 * x_2^2$ 과 $X_2 = x_1^2 * x_2$ 로 정의된다.

X_1 을 첫번째 독립변수로, X_2 를 두번째 독립변수로, g 를 종속변수로 하는 다중 선형 회귀분석의 결과가 그림 2.16에 주어져 있다. A 와 B 에 대한 신뢰구간은 표 2.8에 주어져 있

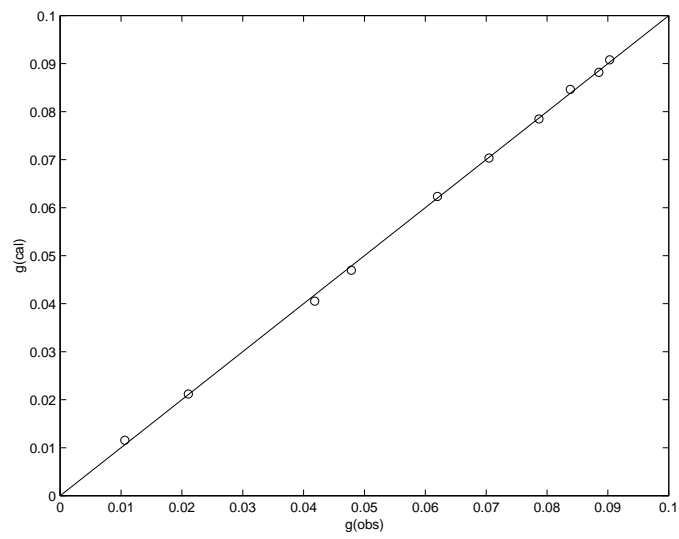


그림 2.16: Margules 식 매개인자에 대한 다중 선형 회귀분석

고, 오차 그림은 그림 2.17에 주어져 있다.

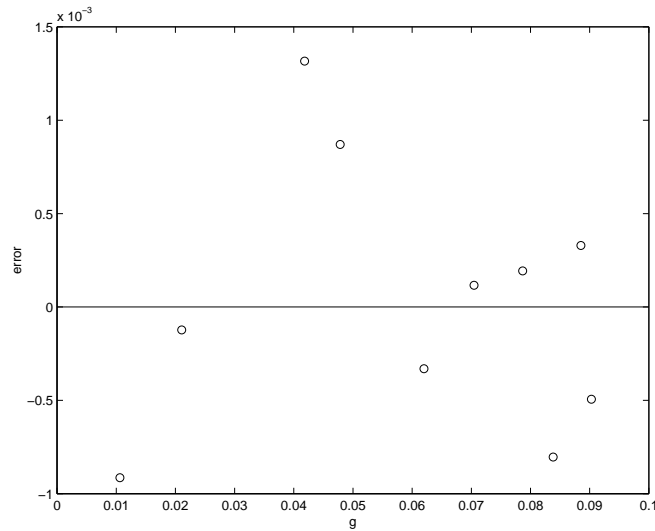


그림 2.17: Margules 식 매개인자에 대한 다중 선형 회귀분석의 오차 그림

표 2.8: Margules 식 매개인자에 대한 다중 선형 회귀분석 결과

매개인자	값	95% 신뢰구간	상한	하한
A	0.251123	0.00827717	0.242846	0.25940
B	0.460917	0.00848425	0.452432	0.469401

식 (2.30)과 (2.31)의 합에 대한 비선형 회귀분석 비선형 회귀분석의 목적함수는 $(\gamma_1 + \gamma_2)$ 이고, 이를 $gsum$ 이라 하면 다음과 같이 표현된다.

$$gsum = \exp [x_2^2(2B - A) + 2x_2^3(A - B)] + \exp [x_1^2(2A - B) + 2x_1^3(B - A)] \quad (2.34)$$

활동도 계수들은 주어진 데이터 영역에서 거의 1에 가깝기 때문에, 이 비선형 목적함수에 대한 γ_1 과 γ_2 의 기여분이 근사적으로 거의 같고, 따라서 비선형 회귀분석은 두 활동도 표현에 거의 같은 가중치를 둔다는 것에 주목하라.

표 2.8에 주어진 매개인자 A 와 B 의 값을 초기 추산치로 사용하는 $gsum$ 에 대한 비선형 회귀 표현은 그림 2.18에 보이는 결과를 준다. 오차 그림은 그림 2.19에 주어져 있고, 신뢰구간은 표 2.9에 주어져 있다.

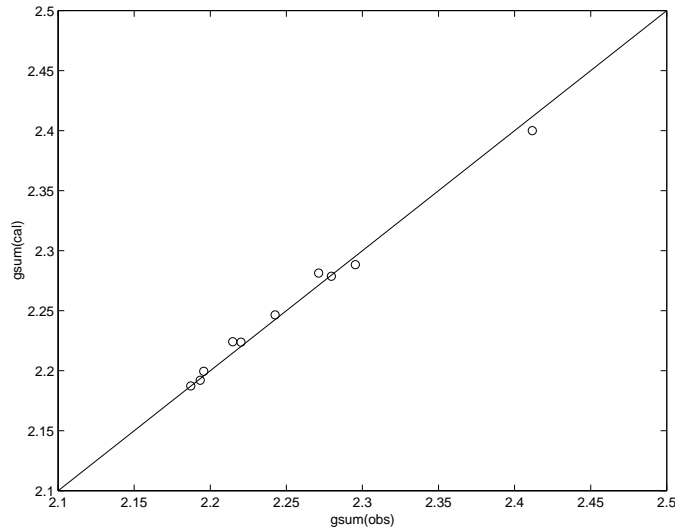


그림 2.18: Margules 식 매개인자에 대한 비선형 회귀분석

표 2.9: Margules 식 매개인자에 대한 비선형 회귀분석 결과

매개인자	값	95% 신뢰구간	상한	하한
A	0.260765	0.01205371	0.248715	0.272822
B	0.451573	0.01564355	0.435928	0.467215

(a)와 (b)의 회귀분석 결과의 비교 두 회귀분석의 기본적인 결과는 거의 같다. 두 상관관계는 입력 데이터를 잘 나타내고, 여러 매개인자들에 대한 신뢰 구간이 작고, 오차 그림은 비교적 무작위적인 분포를 보여준다. 하지만, 자세히 들여다보면 다중 선형 회귀분석이 약간 더 데이터를 잘 나타내고, 더 작은 신뢰구간을 가진다.

두 회귀 표현이 상당히 다르기 때문에, 계산된 분산이나 제곱의 합은 직접 비교될 수 없다. 최종 비교를 위해서는 다음과 같이 주어지는 두 활동도 계수에 대한 오차들의 제곱의 합을 추가적으로 계산하여야 한다.

$$SS = \sum_{i=1}^N \left[(\gamma_{1i(obs)} - \gamma_{1i(calc)})^2 + (\gamma_{2i(obs)} - \gamma_{2i(calc)})^2 \right] \quad (2.35)$$

여기서 (obs)는 관측된 데이터 값을 의미하고, (calc)는 회귀분석 결과로부터 계산된 값을 의미한다. 식 (2.35)를 계산하기 위해서는 각 회귀분석에서 얻어진 A와 B를 값을 사용하여 gam1calc와 gam2calc를 추가적으로 계산하는 MATLAB 명령어를 해야한다. 계산된 SS 값은 다중 선형 회귀분석의 경우에는 $SS = 1.222 \times 10^{-3}$ 이고, 비선형 회귀분석의 경우에는 $SS = 7.994 \times 10^{-4}$ 이다.

따라서 두 회귀분석은 활동도 계수들에 대한 아주 정확한 상관관계를 준다. 비선형 회귀분석이 오차의 제곱으로 파다하며 약간 나은 활동도 계수 값을 주며 바며에 다중 선형

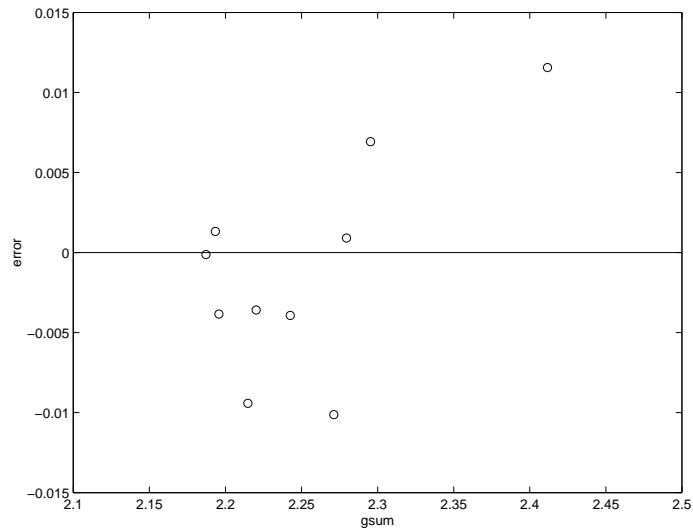


그림 2.19: Margules 식 매개인자에 대한 비선형 회귀분석의 오차 그림



(a),(b)에 대한 MATLAB 스크립트 파일은 CHAP2의 p208a.m과 p208b.m으로 주어져 있다.

제 9 절 Trichloroethane을 포함하는 2성분 계에 대한 Margules 식

9.1 개념 설명

이성분 활동도 계수의 상관식을 위한 Margules 식에서 매개인자 추산.

9.2 사용된 수치 해법

선형 및 비선형 회귀분석, 회귀분석을 위한 데이터의 변환, 신뢰구간, 오차 그림, 제품의 합의 계산 및 비교.

9.3 문제 설명

1,1,1-trichloroethane을 포함하는 이성분 계에 대한 이성분 활동도 계수는 표 B.21에서 B.23에 요약되어 있다.

- (a) 1,1,1-trichloroethane을 포함하는 이성분 계 중의 하나에 대하여 Margules 식의 A와 B를 결정하기 위하여, 식 (2.32)에 다중 선형 회귀분석을 적용하라.
- (b) 식 (2.30)과 (2.31)의 합으로 이루어지는 한 식에 대하여 비선형 회귀분석을 적용하여 A와 B를 결정하라.
- (c) 매개인자 신뢰구간, 오차 그림, 오차의 제품의 합(두 활동도 계수로 계산된 최소 제품 합)을 사용하여 (a)와 (b)의 결과를 비교하라.



표 B.21에서 B.23의 데이터에 대한 MATLAB 데이터 파일은 tableb21.mat에서 tableb23으로 주어져 있다.

제 10 절 촉매 개질 반응에 대한 반응 속도 해석

10.1 개념 설명

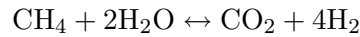
실험 데이터를 위한 촉매 반응 속도의 평가.

10.2 사용된 수치 해법

선형 및 비선형 회귀분석, 회귀분석을 위한 데이터의 변환, 신뢰구간, 오차 그림, 제곱의 합의 계산 및 비교.

10.3 문제 설명

Quanch와 Rouleau [5]는 다음과 같은 가역 촉매 개질 반응에 대한 서로 다른 모델을 조사하였다.



350°C에서 생성물의 분압의 함수로 주어지는 반응 속도의 실험적 결과는 표 2.10에 주어져 있다. 이 촉매 반응에 대한 모델 중의 하나는 다음과 같이 주어지는 메탄이 촉매 표면에 흡착하는 것이다.

$$r_{\text{CO}_2} = \frac{k_s K_{\text{CH}_4} \left(P_{\text{CH}_4} P_{\text{H}_2\text{O}}^2 - \frac{P_{\text{CO}_2} P_{\text{H}_2}^4}{K_P} \right)}{1 + K_{\text{CH}_4} P_{\text{CH}_4}} \quad (2.36)$$

여기서 총괄 평형 상수는 열역학적 계산으로부터 $K_P = 5.051 \times 10^{-5} \text{atm}^2$ 으로 알려져 있다. 따라서 이 촉매 반응 속도식에서 계산될 필요가 있는 매개인자는 단지 두개이다.

표 2.10: 촉매 개질 반응에 대한 반응 속도 데이터

분압 (atm)				CO ₂ 의 반응 속도
CH ₄	H ₂ O	CO ₂	H ₂	g-mol/hr·gm×10 ³
0.06298	0.23818	0.00420	0.01669	0.13717
0.03748	0.26315	0.00467	0.01686	0.15584
0.05178	0.29557	0.00542	0.02079	0.20028
0.04978	0.23239	0.00177	0.07865	0.05700
0.04809	0.29491	0.00655	0.02464	0.20150

0.03849	0.24171	0.00184	0.06873	0.07887
0.03886	0.26048	0.00381	0.01480	0.14983
0.05230	0.26286	0.05719	0.01635	0.15988
0.05185	0.33529	0.00718	0.02820	0.26194
0.06432	0.24787	0.00509	0.02055	0.14426
0.09609	0.28457	0.00652	0.02627	0.20195

더 간단한 다른 모델은 촉매 표면에 흡착하는 성분이 없는 단순히 가역 반응에 대한 것이다. 이 속도식은 다음과 같이 주어진다.

$$r_{CH_4} = k_1 \left(P_{CH_4} P_{H_2O}^2 - \frac{P_{CO_2} P_{H_2}^4}{K_P} \right) \quad (2.37)$$

여기서는 오직 한 매개인자, k_1 만이 존재한다.

반응 속도, r_{CO_2} 는 CO_2 의 생성 속도이고, CH_4 의 생성 속도는 이 속도의 음의 값이 됨에 주의하라.

- (a) 표 2.10의 데이터에 비선형 회귀분석을 적용하여 매개인자 k_s 와 K_{CH_4} 의 값을 계산하라.
- (b) 매개인자 k_1 값을 결정하라.
- (c) 위의 두 속도식 중에 어떤 것이 주어진 데이터를 가장 잘 나타내는가? 이유를 설명하라.



표 2.10의 데이터에 대한 MATLAB 데이터 파일은 p210.mat로 주어져 있다.

제 11 절 속도 데이터의 회귀분석-변수들 사이의 상관 관계 검토

11.1 개념 설명

반응 속도 데이터를 다양한 반응 속도 식으로 상관시킴.

11.2 사용된 수치 해법

매개인자의 신뢰 구간 결정 및 오차 그림을 그리는 다중 선형 회귀분석과 회귀 변수들 사이의 선형 의존 관계 확인.

11.3 문제 설명

표 2.11은 Bacon 과 Downie [1]가 보고한 $A \leftrightarrow R$ 반응에 대한 속도 데이터를 나타낸다. 그들은 두개의 반응 속도 모델로 속도 데이터를 맞추는 것을 제안하였다. 비가역 반응 모델은 다음과 같은 1차 반응 속도식 형태를 갖고,

$$r_R = k_0 C_A \quad (2.38)$$

가역 반응 모델은 다음과 같은 가역 1차 반응 속도식 형태를 갖는다.

$$r_R = k_1 C_A - k_2 C_R \quad (2.39)$$

여기서 r_R 은 성분 R 의 생성 속도($\text{g-mol/dm}^3 \cdot \text{s}$); C_A 와 C_R 은 각각 성분 A 와 R 의 농도(g-mol/dm^3); k_0, k_1, k_2 는 반응 속도 상수(s^{-1})이다.

- (a) 표 2.11의 데이터를 사용하여 두 반응 속도 식의 매개인자를 계산하라.
- (b) 두 모델을 비교하고 어떤 모델이 속도 데이터를 더 잘 상관짓는지를 결정하라.
- (c) 두 변수, C_A 와 C_R 사이에 어떤 상관 관계가 있는지를 결정하라.
- (d) 회귀 변수들 사이의 상관관계의 실용성에 대하여 논의하라.

표 2.11: Bacon 과 Downie [1]의 반응 속도 데이터

실험 번호	r_R g-mol/dm ³ · s × 10 ⁸	C_A g-mol/dm ³ × 10 ⁴	C_R g-mol/dm ³ × 10 ⁴
1	1.25	2.00	7.98
2	2.50	4.00	5.95
3	4.05	6.00	4.00
4	0.75	1.50	8.49
5	2.80	4.00	5.99
6	3.57	5.50	4.50
7	2.86	4.50	5.47
8	3.44	5.00	4.98
9	2.44	4.00	5.99

11.4 해

(a) 반응 속도식의 대한 회귀분석 두 반응 속도식 모두 선형 회귀분석을 위한 표준 형태이므로 선형 회귀분석을 적용할 수 있다.

(b) 반응 속도식 모델의 비교 두 반응 속도식 표현에 대한 선형 회귀분석 결과는 표 2.12에 요약되어 있다. 2-매개인자 모델이 1-매개인자 모델보다 더 작은 분산을 갖는다. 두 식에 대한 오차 그림은 무작위적인 분포를 보일 것 같다. 하지만, k_2 의 신뢰구간이 크고, 0과 음의 값을 포함하고 있으므로, 2-매개인자 모델은 1-매개인자 모델에 비하여 의심스러운 값을 갖는다.

표 2.12: 다중 선형 회귀분석 결과

매개인자	값	95% 신뢰 구간	분산
k_0	6.551e-5	2.74e-6	2.340e-18
k_1	6.867e-5	5.25e-6	1.728e-18
k_2	2.630e-6	3.70e-6	

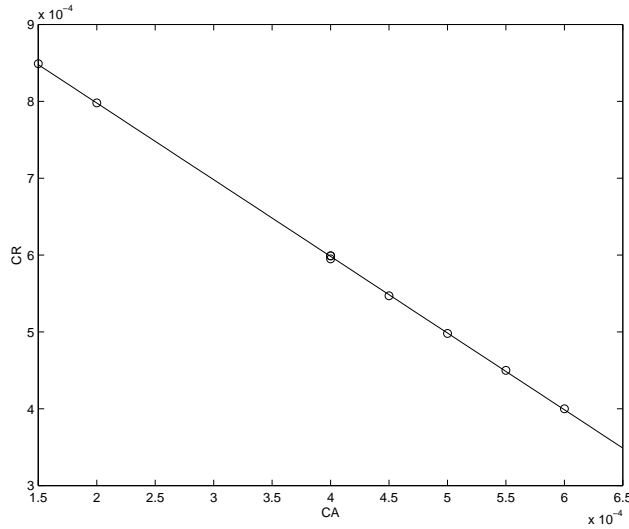


그림 2.20: $A \leftrightarrow R$ 반응에 대한 C_A 에 대한 C_R 의 회귀분석

(c) 변수들 사이의 상관관계에 대한 검토 변수들 사이의 상관관계에 대한 간단한 시험은 한 변수가 다른 변수에 선형적으로 연관되었는지를 결정하기 위하여 선형 회귀분석을 해보는 것이다. 이 경우에는 C_R 을 C_A 에 대하여 회귀분석한 결과는 그림 2.20에 보이는 상관관계이다. 이 그림은 두 실험 변수, C_A 와 C_R 은 선형적으로 연관되어 있음을 명백하게 보여준다.

(d) 회귀 변수들 사이의 상관관계의 중요성 이와 같은 선형 관계는 것은 식 (2.39)의 회귀분석은 모든 변수를 독립변수로 가정하여 행한 것이므로 적합하지 못하다는 것을 의미한다. 새로운 회귀분석에서는 C_R 과 C_A 사이에 알려진 관계는 식 (2.40)으로 간략화되고, 따라서 이 식은 새로운 회귀분석에 이용될 수 있다.

$$C_R = 0.001 - C_A \quad (2.40)$$

식 (2.40)을 식 (2.39)에 대입하면 다음과 같은 결과를 얻을 수 있다.

$$r_R = k_1 C_A - k_2(0.001 - C_A) = (-0.001)k_2 + (k_1 + k_2)C_A \quad (2.41)$$

이 식은 선형관계를 나타낸다. $a_0 = -0.001k_2$, $a_1 = k_1 + k_2$ 로 두면 선형 회귀분석을 적용할 수 있고, 다음과 같은 식에 선형 회귀분석을 적용한 결과는 표 2.13에 주어져 있다.

$$C_R = a_0 + a_1 C_A \quad (2.42)$$

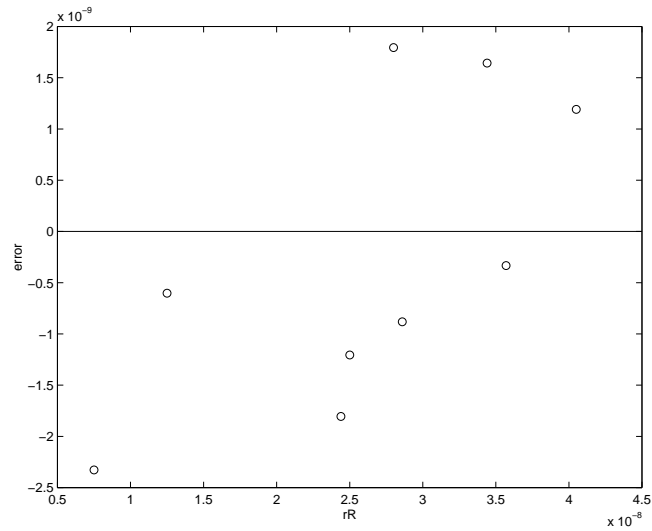


그림 2.21: 식 (2.38)로 나타내어지는 반응 속도 데이터에 대한 오차 그림

표 2.13: 식 (2.42)에 대한 회귀분석 결과

매개인자	값	95% 신뢰 구간
a_0	-2.62633e-9	3.1668e-9
a_1	7.12978e-5	7.37921e-6

이 표에 주어진 매개인자 값들은 a_0 가 0에서 크게 벗어나지 않고 (신뢰구간이 매개인자 값 자체 보다 크다), 따라서 상관관계로부터 소거되어 식 (2.38)의 비가역 반응 모델로 될 수 있다. 표 2.12는 $k_0 = 6.551 \times 10^{-5} \pm 2.74 \times 10^{-6}$ 의 관계를 나타낸다. 오차가 무작위적인 분포를 보이는 이 비가역 모델에 대한 오차 그림은 그림 2.21에 주어져 있다. 이 모든 결과가 비가역 모델이 선호된다는 Bacon과 Downie [1]의 결론을 지지한다.

C_R 과 C_A 사이의 선형관계 때문에 데이터와 회귀분석 결과로부터 어떤 결론도 도출할 수 없다. 이 선형관계 때문에 k_2 에 대한 독립적인 정보를 얻을 수 없고, $a_0 = -0.001k_2$ 관계로부터 k_2 의 계산은 k_2 를 합리적인 추산하지 못한다. 이는 표 2.13에 요약된 a_0 에 대한 신뢰구간 한계가 아주 크기 때문이다.

이 문제는 실험을 계획할 때 가장 중요한 고려사항, 즉 변수들을 서로 독립적으로 변화시켜야 된다는 사실을 끄집어내준다. 그렇지 않으면, 중요한 몇몇 정보는 변수들 사이의

선형 또는 다른 관계때문에 잃게된다. 이 문제에서 농도의 합은 항상 같게 된다. 이 관계는 주어진 온도 압력에서 이성분 기체에 대하여 나타난다. 이 경우에 반응 속도를 측정하는 과정에 회석제의 사용은 아주 적절할 것이다.



이 문제에 대한 MATLAB 스크립트 파일은 CHAP2의 p211.m으로 주어져 있다.

제 12 절 불균질 촉매 반응 속도의 회귀분석

12.1 개념 설명

불균질 촉매 반응 속도 데이터를 반응 속도 식으로 표현.

12.2 사용된 수치 해법

선형화된 표현 및 변수 변환을 사용하는 다중 선형 및 비선형 회귀분석, 회귀 변수들 사이의 가능한 의존 관계 확인.

12.3 문제 설명

표 2.14는 $A \rightarrow B$ 로 나타내어지는 불균질 촉매 반응에 대한 반응 속도 데이터이다.

표 2.14: 불균질 촉매 반응에 대한 반응 속도 데이터

번호	$P_A(\text{atm})$	$P_B(\text{atm})$	$r \times 10^5$	번호	$P_A(\text{atm})$	$P_B(\text{atm})$	$r \times 10^5$
1	1	0	5.1	5	0.6	0.4	6
2	0.9	0.1	5.4	6	0.5	0.5	6.15
3	0.8	0.2	5.55	7	0.4	0.6	6.3
4	0.7	0.3	5.85	8	0.3	0.7	6.45

위의 데이터를 상관시키는데 다음과 같은 식이 제안되었다.

$$r = \frac{k_1 P_A}{(1 + K_A P_A + K_B P_B)^2} \quad (2.43)$$

여기서 k_1 , K_A , K_B 는 회귀분석을 통하여 결정되어야 할 상수이다. 식 (2.43)은 정리하고, 역수를 취하고, 제곱근을 취하면 다음과 같이 선형화 됨에 주의하라.

$$\left(\frac{P_A}{r}\right)^{1/2} = \frac{1}{\sqrt{k_1}} + \frac{K_A}{\sqrt{k_1}} P_A + \frac{K_B}{\sqrt{k_1}} P_B \quad (2.44)$$

- (a) 표 2.14의 데이터를 회귀분석하여 몇 개의 식 (2.43)과 (2.44)의 매개인자가 추산되어야 하는지를 결정하라.
- (b) 선형 및 비선형 회귀분석을 사용하여 매개인자들을 결정하고, 이 두 방법으로 얻은 결과를 비교하라.



이 문제에 대한 MATLAB 데이터 파일은 p212.mat로 주어져 있다.

제 13 절 온도에 따른 반응 속도 상수의 변화

13.1 개념 설명

Arrhenius 식을 사용하여 온도에 따른 반응 속도 상수 변화의 상관관계.

13.2 사용된 수치 해법

분산, 신뢰구간, 오차 그림등을 사용하여 회귀분석 결과를 비교하는 다중 선형 회귀분석 및 비선형 회귀분석.

13.3 문제 설명

Copper magnesia 촉매 상에서 에틸렌의 수소화가 Wynkoop와 Wilhelm [9]에 의하여 여러 온도에서 연속 흐름 관형 반응기를 사용하여 연구되었다. 결과들이 표 B.24에 온도 $T(^{\circ}\text{C})$ 에 대한 반응 속도 상수 $k(\text{g}\cdot\text{mol}/\text{cm}^3 \cdot \text{s} \cdot \text{atm})$ 로 도표화 되어있다.

온도에 따른 k 의 변화는 Arrhenius 식에 의하여 다음과 같이 나타내어 진다.

$$k = A \exp[-E/(RT)] \quad (2.45)$$

여기서 T 는 절대 온도, R 은 기체 상수($1.987\text{cal}/\text{g}\cdot\text{mol}\cdot\text{K}$), E 는 활성화 에너지($\text{cal}/\text{g}\cdot\text{mol}$), A 는 반응 속도의 차원을 갖는 빈도 인자이다.

반응 속도 상수에 대한 Arrhenius 식의 또다른 편리한 형태는 다음과 같다.

$$k = k_0 \exp[E/R(1/T_0 - 1/T)] \quad (2.46)$$

여기서 T_0 는 $k = k_0$ 인 임의의 절대 온도이다. $k_0 = A \exp(-E/RT_0)$ 임에 주의하라.

- (a) 식 (2.45)와 (2.46)에서 Arrhenius 매개인자를 구하기 위하여 선형 및 비선형 회귀분석을 하라. 식 (2.46)에서 $T_0 = 298\text{K}$ 로 고정하라.
- (b) 어떤 식 및 회귀분석법이 가장 정확한 결과를 주는가? 여러분이 선택에 대하여 설명하라.
- (c) 여러분이 구한 가장 정확한 결과를 Wynkoop와 Wilhelm [9]에 주어진 $A = 5690\text{g}\cdot\text{mol}/\text{cm}^3 \cdot \text{s} \cdot \text{atm}$, $E = 13320\text{cal}/\text{g}\cdot\text{mol}$ 과 비교하라.



표 B.24의 데이터에 대한 MATLAB 데이터 파일은 tableb24.mat로 주어져 있다.

제 14 절 선형 회귀분석을 사용하는 Antoine 식의 매개인자 계산

14.1 개념 설명

온도에 따른 증기압 데이터를 상관시키기 위하여 Antoine 식의 직접 사용

14.2 사용된 수치 해법

총괄 분산 및 각 매개인자의 신뢰구간의 결정하는 다중 선형 회귀분석.

14.3 문제 설명

다중 선형 회귀분석 다중 선형회귀분석은 N 개의 데이터를 식 (2.47)에서 주어지는 형태의 선형 함수로 맞추는 것으로 정의될 수 있다.

$$y_{(calc)} = a_0 + a_1x_1 + a_2x_2 + \cdots a_nx_n \quad (2.47)$$

여기서 $x_1, x_2, \cdots x_n$ 은 n 개의 독립변수이고 $a_0, a_1 \cdots a_n$ 은 매개인자, $y_{(calc)}$ 는 종속 변수의 추산값, $y_{(obs)}$ 는 종속 변수의 측정값이다. 식(2.47)의매개인자들은 다음과 같은 선형 연립 방정식을 풀어서 계산될 수 있다.

$$X^T X A = X^T Y \quad (2.48)$$

여기서 X 는 독립 변수의 관측값들로 이루어진 행렬이고, A 는 매개인자 벡터, Y 는 종속 변수의 관측값으로 이루어진 행렬이다. 따라서

$$X = \begin{bmatrix} 1 & x_{1,1} & x_{2,1} & \cdots & x_{n,1} \\ 1 & x_{1,2} & x_{2,2} & \cdots & x_{n,2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1,N} & x_{2,N} & \cdots & x_{n,N} \end{bmatrix} \quad A = \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_n \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y_{1(obs)} \\ y_{2(obs)} \\ \cdot \\ \cdot \\ y_{n(obs)} \end{bmatrix} \quad (2.49)$$

행렬 X 의 첫번째 첨자는 변수 수이고, 두번째 첨자는 관측 데이터 점의 수임에 주의하라. 관측 데이터의 총 수는 N 이다.

분산 σ^2 은 식 (2.4)로부터 계산될 수 있다. 이 식에서 y_{iobs} 는 종속 변수의 관측값이고, y_{icalc} 는 추정값이며, ν 는 $[N - (n + 1)]$ 로 주어지는 자유도(degree of freedom)이다.

$\beta_0, \beta_1 \dots \beta_n$ 으로 표기되는 식 (2.47)의 엄밀한 매개인자 값들은 다음 구간 안에 위치해야 한다.

$$a_{i-1} - \sqrt{\alpha_{ii}}\sigma t_\nu \leq \beta_{i-1} \leq a_{i-1} + \sqrt{\alpha_{ii}}\sigma t_\nu \quad (2.50)$$

여기서 α_{ii} 는 $(X^T X)^{-1}$ 행렬의 대각선 요소이다. t_ν 는 원하는 신뢰도 구간에서 ν 로 주어지는 자유도에 해당하는 통계학적 t 분포 값이고, σ 는 표준 편차(분산의 제곱근)이다. 신뢰구간은 $\sqrt{\alpha_{ii}}\sigma t_\nu$ 항으로 주어진다.

종속 변수의 관측값과 계산값의 차이로 정의되는 오차는 식 (2.5)로부터 계산된다.

식 (2.1)의 Antoine 식의 매개인자와 표 B.5의 프로판의 증기압 데이터에 대한 여러 통계학적 인자를 계산하라. 증기압을 psia로 온도를 °F로 나타내는 경우에 대한 매개인자 값을 구하라. (이 문제는 문제 2.1과 비슷하나 매개인자들이 다른 단위를 갖는다.) 해를 구하는 과정에서 선형 회귀분석에 관한 기본적인 계산을 한다. 다음과 같은 순서로 진행된다.

- (a) Antoine 식의 매개인자를 다중 선형 회귀분석을 통하여 구할 수 있도록 변수들을 변환하라.
- (b) 행렬 $X^T X$ 와 $X^T Y$ 를 구하라.
- (c) 벡터 A 를 구하기 위하여 연립 방정식을 풀어라.
- (d) 분산, $(X^T X)^{-1}$ 의 대각선 요소와 (표 B.6에 주어지는 t 분포 값을 사용하여) 매개인자의 95% 신뢰도 구간을 계산하라
- (e) 오차 그림을 그려라(ϵ_i 를 $y_{i\text{obs}}$ 에 대하여 도시하라).
- (f) 데이터들의 상관식에 대한 Antoine 식의 적절성과 데이터의 정밀도를 평가하라.

14.4 해

문제 풀이의 대부분의 과정이 MATLAB을 사용하여 이루어진다. 우선 표 B.5의 데이터들이 프로그램에 입력되어야 한다.

(a) Antoine 식은 문제 1.7에서 보는 바와 같이 선형화 될 수 있다. 이 식의 또 다른 선형화된 형태는 다음과 같다.

$$T \log P_v = (AC + B) + AT - C \log P_v \quad (2.51)$$

여기서 \log 는 상용 로그를 의미한다. 이 형태가 이 문제에서 사용된다. 원 데이터들이 MATLAB에 입력되어야 하고 표 2.15에 주어진 것과 같이 $y=T \log P_v$, $x_1=T$, $x_2=\log P_v$ 로 변환되어야 한다.

표 2.15: Antoine 식 회귀분석을 위한 변환된 변수들

y	x1	x2
-60.7227	-70	0.867467
-59.26	-60	0.987666
-55.0185	-50	1.10037
-48.3806	-40	1.20952
-39.2249	-30	1.3075
-28.0957	-20	1.40483
-14.9693	-10	1.49693
0	0	1.58206
16.6276	10	1.66276
34.8859	20	1.74429
54.6454	30	1.82151
75.6838	40	1.89209
98.1421	50	1.96284
121.787	60	2.02979
146.54	70	2.09342
172.378	80	2.15473
199.336	90	2.21484
227.184	100	2.27184
256.122	110	2.32838
285.625	120	2.38021

(b) 위의 데이터를 사용하는 다중 선형 회귀분석의 경우에는 행렬 $X^T X$ 와 $X^T Y$ 는 다음과 같이 쓰여질 수 있다.

$$X^T X = \begin{bmatrix} 20 & 500 & 34.51304 \\ 500 & 79000 & 1383.28 \\ 34.5130 & 1383.28 & 63.6854 \end{bmatrix} \quad \text{and} \quad X^T Y = \begin{bmatrix} 1383.78 \\ 159246 \\ 3340.10 \end{bmatrix} \quad (2.52)$$

따라서 문제는 다음과 같은 연립 방정식을 푸는 것으로 변하게 된다.

$$\begin{bmatrix} 20 & 500 & 34.51304 \\ 500 & 79000 & 1383.28 \\ 34.5130 & 1383.28 & 63.6854 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1383.78 \\ 159246 \\ 3340.10 \end{bmatrix} \quad (2.53)$$

(c) 식 (2.53)으로 주어지는 일반식 (2.48)의 해는 MATLAB을 사용하여 구할 수 있다. 결과는 다음과 같다.

$$a_0 = 678.946 \quad a_1 = 5.23328 \quad a_2 = -429.165 \quad (2.54)$$

행렬 $X^T X$ 의 역행렬은 MATLAB에서 $\text{inv}(X'X)$ 를 사용하여 쉽게 구할 수 있다. 또 이 행렬의 대각선 요소들은 $\text{diag}(\text{inv}(X'X))$ 을 사용하여 구해진다. 결과는 $\alpha_{11} = 42.9857$, $\alpha_{22} = 0.00113828$, $\alpha_{33} = 18.3377$ 이다.

(c) 분산은 관측값과 위에서 구한 회귀 곡선을 사용하여 구한 값과 사이의 차로 정의되는 오차 $\text{err} = y - \text{ycal}$ 을 사용하여 다음과 같이 구해진다.

$$\text{var} = \text{err}' * \text{err} / \nu \quad (2.55)$$

여기서 ν 는 자유도로 $\nu = 20 - 3 = 17$ 이고, 계산된 분산 값은 $\sigma^2 = 0.309044$ 이다. 자유도 17, 95% 신뢰도 구간에 해당하는 t 분포 값은 표 B.6에서 $t = 2.1098$ 로 주어진다. 따라서 식 (2.50)으로 주어지는 각 신뢰구간 $\sqrt{\alpha_{ii}\sigma^2}$ 은 다음과 같다.

$$\text{for } a_0 = \sqrt{42.9857} \times 0.555918 \times 2.1098 = 7.68977$$

$$\text{for } a_1 = \sqrt{0.00113828} \times 0.555918 \times 2.1098 = 0.0395708$$

$$\text{for } a_2 = \sqrt{18.3377} \times 0.555918 \times 2.1098 = 5.02255$$

(e) 관측값과 위에서 구한 회귀 곡선을 사용하여 구한 값과 사이의 차로 정의되는 오차 $\text{err} = y - \text{ycal}$ 는 그림 2.22에 y 에 대하여 도시되어 있다.

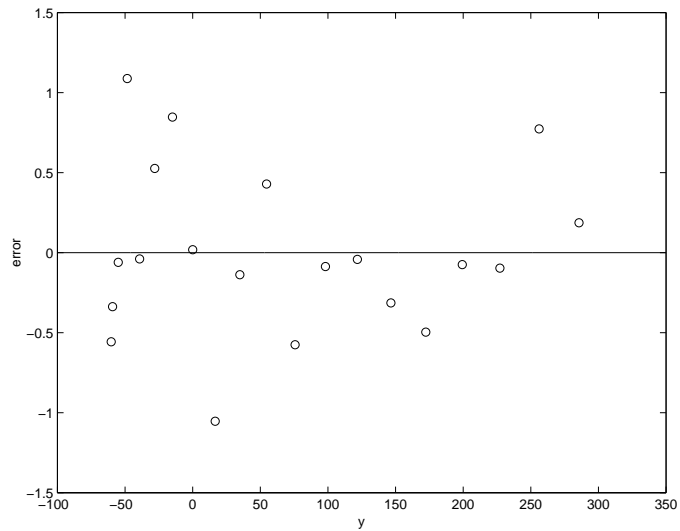


그림 2.22: 선형화된 Antoine 식으로 나타내어지는 증기압 데이터에 대한 오차 그림

(f) 오차 그림은 증기압이 낮은 경우에 해당하는 y 값이 작거나 음인 경우에 오차가 큼을 보여준다. 좀 더 정확한 상관식을 얻기 위해서는 저압(저온) 영역에서 좀 더 정밀하게 측정하여야 한다.

오차들이 0 근처에 균등하게 분포되어 있고, 매개인자의 신뢰 구간들이 작은 것은 Antoine 식이 측정이 행해진 영역에서 프로판의 증기압 데이터들을 적합하게 상관짓고 있음을 나타낸다.



이 문제에 대한 MATLAB 스크립트 파일은 CHAP2의 p214.m에 주어져 있고 이 문제에서 사용한 데이터는 p214.mat로 주어져 있다.

참고문헌

- [1] Bacon, D.W. and Downie, J., *Evaluation of Rate Data-III*, in Crynes, B.L. and Fogler, H.S., eds., *AICHEMI Modular Instruction: Series E Kinetics*, Vol. 2, New York: AICHE, 1981, pp.65-74.
- [2] Dow, W.M. and Jacob, M., *Chem. Eng. Progr.*, 47, 637 (1951).
- [3] Geankoplis, C.J., *Transport Process and Unit Operations*, 3rd ed., Englewood Cliffs, N.J.: Prentice Hall, 1993.
- [4] Perry, R.H., Green, D.W. and Marorey, J.D., Eds, *Perry's Chemical Engineers Handbook*, 7th ed., New York: McGraw Hill, 1997.
- [5] Quanch, Q.P, and Rouleau, J., *Appl. Chem. Biotechnol.*, 25, 445 (1975).
- [6] Reid, R.C., Prausnitz, J.M. and Poling, B.F., *The Properties of Gas and Liquids*, 4th ed., New York: McGraw Hill, 1987.
- [7] Sieder, E.N. and Tate, G.E., *Ind. and Eng. Chem.*, 28, 1429 (1936).
- [8] Williams, R. B. and Katz, D.I., *Trans. ASME*, 74, 1307-1320 (1952).
- [9] Wynkoop, R. and Wilhelm, R.H., *Chem. Eng. Progr.*, 46, 300 (1950).