

# *Manejo de los datos experimentales*

## 6.1 INTRODUCCION

En diversas ocasiones, ha sido necesario trazar una línea que pase por los datos experimentales. En el capítulo 2 se comparó una ecuación del modelo con los datos del drenado del tanque, y se calculó, a partir de la pendiente de la recta óptima dibujada uniendo dichos datos, el parámetro requerido para el análisis. Asimismo, en el capítulo 5 se compararon las predicciones del modelo con los datos para el reactor intermitente, y se calculó la constante de velocidad a partir de la pendiente de la recta. La etapa en que se comparan la predicción del modelo y los datos experimentales es fundamental en el proceso del análisis.

En el capítulo 5 fue necesario desarrollar una relación funcional entre la concentración de la sal y la densidad, mientras que en las publicaciones relativas se encontraba solamente una serie de puntos experimentales. De esta manera, se decidió hacer pasar una línea recta por los datos concentración—densidad, y utilizar esta relación lineal en las ecuaciones de conservación. Es importante hacer notar que el criterio seguido para la comparación entre las ecuaciones del modelo y los datos experimentados, ha estado determinado por la teoría y la forma de la ecuación. Cuando se está tratando de adaptar una recta a los datos, con objeto de determinar una relación funcional, simplemente se procede en forma empírica, la forma de la función es arbitraria, y está guiada solamente por la conveniencia. En ambas circunstancias, debido al error en los datos experimentales, las aproximaciones a la teoría o también, debido a que la forma empírica que se ha seleccionado representa solamente en forma aproximada a los datos, en contadas ocasiones podrá trazarse una recta única que pase por los puntos correspondientes a los datos. Por lo tanto,

## 182 Manejo de los datos experimentales

habrá cierta vacilación sobre la pendiente adecuada y la intersección, por tanto, también habrá duda sobre los parámetros que puedan calcularse. Por ello, será necesario desarrollar un procedimiento racional para la evaluación de la mejor recta *óptima* a través de un conjunto de datos experimentales.

### 6.2 CRITERIO DE TRAZO OPTIMO

Suponga que se tiene un conjunto de datos tomados de varios valores de una variable independiente, por ejemplo, el tiempo, y que existen razones para creer que los datos estarán representados por una línea recta que pasa por cero, esto es, mediante la ecuación:

$$y = bt \quad (6.1)$$

Dichos datos son valores medidos de  $y_1$  en el tiempo  $t_1$ ;  $y_2$  en el tiempo  $t_2$ , etc. Por lo general, expresarán el valor medido  $y_k$  en el tiempo  $t_k$ ; suponga que se tiene un total de  $N$  puntos de datos. Se desea encontrar el valor de la pendiente,  $b$ , en la ecuación 6.1 que representa mejor dichos datos.

Claro que se espera tener dispersión en los datos, ninguna línea recta pasará exactamente por cada punto, y por lo tanto, se tendrá un error en cada punto asociado con el uso de la ecuación 6.1. La diferencia entre  $bt_k$ , valor de  $y$  predicho en la ecuación 6.1 en el tiempo  $t_k$ , y el valor  $y_k$  realmente medido en el tiempo  $t_k$  es:

$$\text{'(diferencia)}_k = y_k - bt_k \equiv e_k \quad (6.2)$$

donde la notación  $e_k$  representa el error en el punto  $k$ . Tomando todos los puntos en consideración, no sólo se consideran puntos individuales sino el error *total*. Este error total también deberá definirse.

Será útil considerar el error en un punto en términos de su *magnitud*  $|e_k|$ , ya que de esta forma las diferencias positivas y negativas, igualmente objetables, se mostrarán de la misma manera. Entonces un medio razonable para definir el "trazado óptimo" de una recta con los datos disponibles será, sumar todos los valores absolutos de los errores en un punto. Esta cantidad se expresa mediante  $e_{AV}$ , para el valor absoluto total del error:

$$e_{AV} = \sum_{k=1}^N |e_k| = \sum_{k=1}^N |y_k - bt_k| \quad (6.3)$$

Una alternativa que puede elegirse para este examen se basa en el *cuadrado* del error en un punto, el cual es el mismo, tanto para las diferencias positivas como para las negativas. Esto se expresa mediante  $e_{SS}$ , para la suma total de los cuadrados de errores:

$$e_{SS} = \left[ \sum_{k=1}^N e_k^2 \right]^{1/2} = \left\{ \sum_{k=1}^N [y_k - bt_k]^2 \right\}^{1/2} \quad (6.4)$$

Note que  $e_{SS}$  tiene la forma de una distancia en las coordenadas cartesianas.

Debe hacerse notar que todos los  $y_k$  y  $t_k$  son números conocidos, de tal manera que solamente se requiere la pendiente,  $b$ , con objeto de calcular ya sea  $e_{AV}$  o  $e_{SS}$ . Por lo tanto, se puede evaluar el error total como una función de  $b$ , y encontrar el valor "óptimo" de  $b$  como aquel que contiene el error mínimo. Para ilustrar este aspecto simplemente considere algunos datos representativos del experimento del vaciado del tanque en la tabla 2.1.

$h$	$t$	Media $t$
10	10.9, 11.5, 11.6	11.3
5	43.8, 42.8, 43.8	43.5
1	85.0, 84.0, 85.2	84.7

Como ejemplo se utilizarán solamente estos datos y, de hecho, solamente el valor promedio de  $t$ . La relación teórica que se examina es la ecuación 2.12 para el caso especial  $n = 1/2$ , que puede escribirse como:

$$h_0^{1/2} - h^{1/2} = \frac{k}{2A} t$$

donde  $h_0 = 12.0$ . Se expresa  $h_0^{1/2} - h^{1/2}$  mediante  $y$  y  $k/2A$  mediante  $b$ . Los datos se grafican en la figura 6.1, donde el área sombreada muestra el intervalo posible de pendientes. Note que sólo una línea recta puede pasar por todos lo datos.

Los datos se grafican en la tabla 6.1 en forma apropiada para calcular el error total definido por las ecuaciones 6.3 y 6.4. Entonces el cálculo se continúa en forma directa. Por ejemplo, para  $b = 0.027$ ,

$$\begin{aligned} e_{AV}(0.027) &= |e_1| + |e_2| + |e_3| \\ &= |0.30 - 11.3 \times 0.027| + |1.23 - 43.5 \times 0.027| \\ &\quad + |2.46 - 84.7 \times 0.027| = 0.23 \end{aligned}$$

## 184 Manejo de los datos experimentales

$$\begin{aligned} e_{SS}(0.027) &= [e_1^2 + e_2^2 + e_3^2]^{1/2} \\ &= \{[0.30 - 11.3 \times 0.027]^2 + [1.23 - 43.5 \times 0.027]^2 \\ &\quad + [2.46 - 84.7 \times 0.027]^2\}^{1/2} = 0.18 \end{aligned}$$

La figura 6.2 muestra una gráfica de  $e_{AV}$  y  $e_{SS}$  en función de la pendiente,  $b$ ; el mínimo tiene una definición clara y en cada caso se presenta un valor cercano a 0.029.

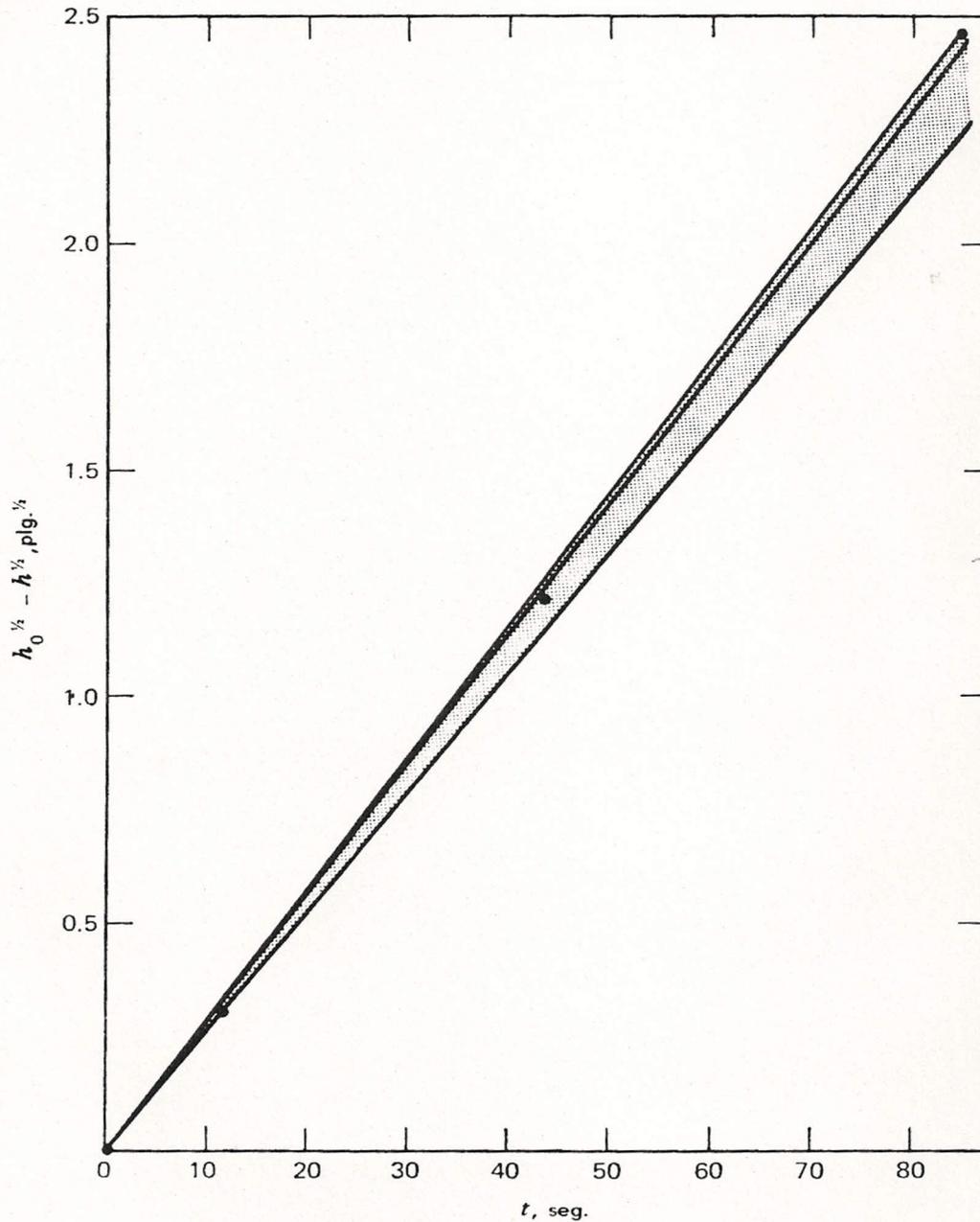


FIGURA 6.1 Datos seleccionados del experimento del vaciado de un tanque. La porción sombreada muestra el intervalo de posibles líneas rectas que pasan a través de los datos.

TABLA 6.1 Error en los puntos seleccionados para el problema del vaciado de un tanque en función de la pendiente,  $b$ .

	$y_k = h_0^{1/2} - h_k^{1/2}$	$t_k$	$e_k = y_k - bt_k$
$k = 1:$	0.30	11.3	$e_1 = 0.30 - 11.3b$
$k = 2:$	1.23	43.5	$e_2 = 1.23 - 43.5b$
$k = 3:$	2.46	84.7	$e_3 = 2.46 - 84.7b$

Debe hacerse notar que aquí hay un cierto grado de ambigüedad. Se han definido *dos* criterios para el trazado óptimo, y aun cuando producen el mismo valor para la pendiente de la recta óptima, las pendientes difieren un poco, aun cuando estos datos tienen escasa dispersión. No se tratará de resolver dicha ambigüedad, ya que no está claro que un criterio sea superior a otro para dicho asunto, ni tampoco ninguno de éstos será superior a varios otros que pudieran proponerse. La suma de cuadrados tiene algún significado en la interpretación estadística de los datos, y probablemente mayor importancia, ya que tiene propiedades analíticas adecuadas (por ejemplo, la diferenciación siempre es posible) de las cuales carece el valor

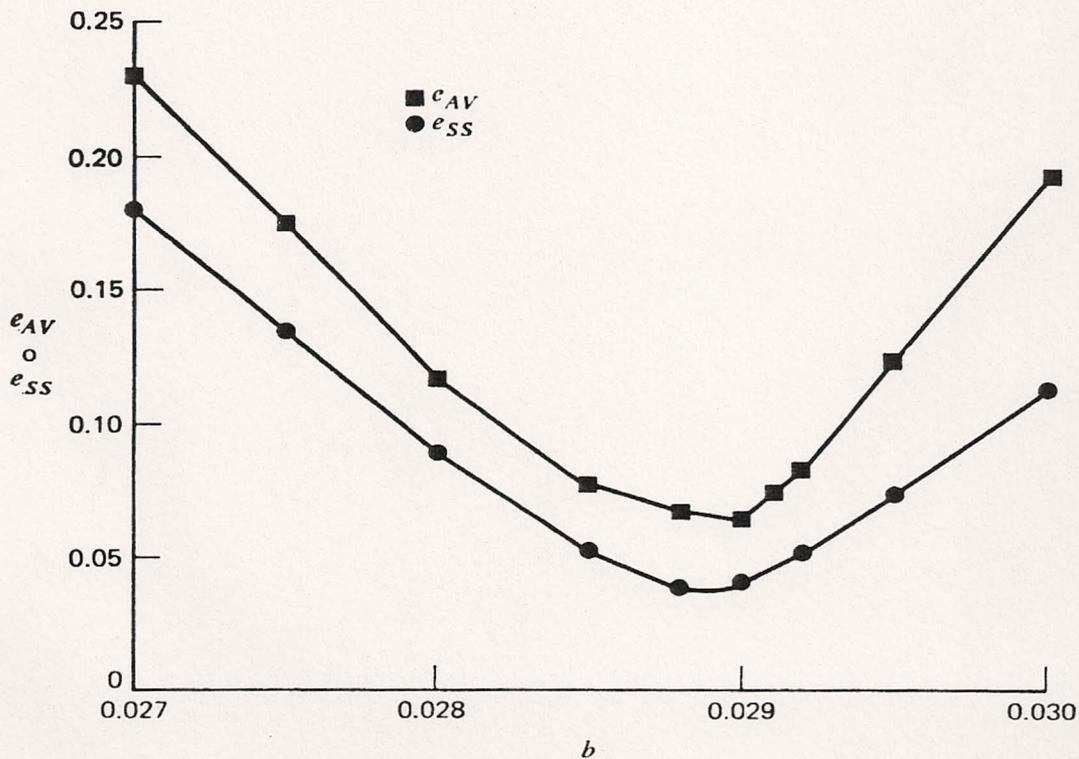


FIGURA 6.2 Valor del error absoluto total ( $e_{AV}$ ) y del error total de la suma de cuadrados ( $e_{SS}$ ) con respecto a la pendiente para los datos seleccionados del experimento del vaciado de un tanque.

## 186 Manejo de los datos experimentales

absoluto. De esta forma, la suma mínima de cuadrados es el criterio en uso común, y es el criterio de trazado óptimo que se utiliza aunque debe recordarse que no es el único.

### 6.3 PENDIENTE OPTIMA - I

Utilizando el criterio de la suma de cuadrados, para el error total, se puede obtener una expresión analítica explícita para la pendiente óptima. Por conveniencia, se eliminará el subíndice "SS". Luego,

$$e^2 = \sum_{k=1}^N [y_k - bt_k]^2 = \sum_{k=1}^N y_k^2 - 2b \sum_{k=1}^N y_k t_k + b^2 \sum_{k=1}^N t_k^2 \quad (6.5)$$

Completando el cuadrado, la ecuación 6.5 puede escribirse:

$$e^2 = \left\{ \sum_{k=1}^N y_k^2 - \frac{\left[ \sum_{k=1}^N y_k t_k \right]^2}{\sum_{k=1}^N t_k^2} \right\} + \left\{ b \left[ \sum_{k=1}^N t_k^2 \right]^{1/2} - \frac{\sum_{k=1}^N y_k t_k}{\left[ \sum_{k=1}^N t_k^2 \right]^{1/2}} \right\}^2$$

El primer término entre paréntesis de la última ecuación es independiente de la pendiente,  $b$ , mientras que el segundo sí depende de ésta. De este modo, se hace  $e^2$  (y por lo tanto  $e$ ) tan pequeño como sea posible se hace lo propio con el segundo término. El segundo término es un cuadrado cuyo valor mínimo es cero, que se obtiene cuando la pendiente es igual a:

$$b = \frac{\sum_{k=1}^N y_k t_k}{\sum_{k=1}^N t_k^2} \quad (6.6)$$

#### Ejemplo 6.1

Los datos seleccionados se grafican en la figura 6.1, se registran en la tabla 6.1 en lo concerniente al vaciado del tanque, y se repiten en la tabla 6.2 con las estimaciones pertinentes para calcular la pendiente óptima conforme a la ecuación 6.6.

Se observa que el resultado está en concordancia con el mínimo de la gráfica de  $e_{SS}$  y  $b$  en la figura 6.2. Esta facilidad para calcular el punto mínimo directamente a partir de los datos, sin cálculos aproximados es lo que hace que el método del criterio de la suma de cuadrados sea el más útil.

TABLA 6.2 Cálculo de la pendiente óptima a través de los puntos seleccionados, mediante la ecuación 6.6, para el problema del vaciado de un tanque.

$y_k$	$t_k$	$t_k^2$	$y_k t_k$
0.30	11.3	128	3.4
1.23	43.5	1892	53.5
2.46	84.7	7174	208.4
		$\sum_k t_k^2 = 9194$	$\sum_k y_k t_k = 265.3$

$$b = \frac{\sum_k y_k t_k}{\sum_k t_k^2} = \frac{265.3}{9194} = 0.0289$$

En la tabla 6.3 se muestra el conjunto completo de datos del problema relativo al vaciado del tanque, incluyendo las tres posibilidades experimentales. Es evidente, que la inclusión de los 33 puntos de datos no hace más complejo el cálculo de la pendiente, que es  $b = k/2A = 0.0287$ , resultado utilizado en el capítulo 2, y que se muestra en la figura 2.6.

En el capítulo 2 se hizo hincapié en que durante un periodo corto de tiempo en relación con el tiempo característico, el comportamiento de todos los modelos es el mismo, y se hicieron estimaciones de las pendientes, solamente con los primeros datos. Calculando la pendiente sólo para los doce primeros puntos a una altura de 8 in.,  $\sum t_k = 2990$ ,  $\sum y_k t_k = 81.0$ , y  $b = 0.0271$ , muy cercano al valor mostrado para la línea superior de la figura 2.6.

(Ver tabla 6.3 en la siguiente página)

### Ejemplo 6.2

Un segundo ejemplo de la determinación del mejor valor de un parámetro a partir de la pendiente de una línea recta experimental, se encuentra en los datos para la reacción del ácido sulfúrico y sulfato de dietilo. La concentración del ácido sulfúrico se expresa como  $c_A$ . Según la ecuación 5.38 una gráfica de:

## 188 Manejo de los datos experimentales

TABLA 6.3 Cálculo de la pendiente óptima para el problema del vaciado de un tanque utilizando la ecuación 6.6.

$h_k$	$y_k = h_0^{1/2} - h_k^{1/2}$	$t_k$	$t_k^2$	$y_k t_k$
11.0	0.15	5.8	33.6	0.87
	0.15	6.1	37.2	0.92
	0.15	5.9	34.8	0.89
10.0	0.30	10.9	118.8	3.27
	0.30	11.5	132.3	3.45
	0.30	11.6	134.6	3.48
9.0	0.46	16.6	275.6	7.64
	0.46	17.8	316.8	8.19
	0.46	17.2	295.8	7.91
8.0	0.64	23.0	529.0	14.7
	0.64	23.5	552.3	15.0
	0.64	23.0	529.0	14.7
7.0	0.82	30.0	900.0	24.6
	0.82	29.2	852.6	23.9
	0.82	29.8	888.0	24.4
6.0	1.01	36.4	1325	36.8
	1.01	35.8	1282	36.2
	1.01	36.4	1325	36.8
5.0	1.23	43.8	1918	53.9
	1.23	42.8	1832	52.6
	1.23	43.8	1918	53.9
4.0	1.46	51.0	2601	74.5
	1.46	50.5	2550	73.7
	1.46	51.6	2663	75.3
3.0	1.73	60.2	3624	104.1
	1.73	59.2	3505	102.4
	1.73	60.6	3672	104.8
2.0	2.05	71.0	5041	145.6
	2.05	69.8	4872	143.1
	2.05	71.4	5098	146.4
1.0	2.46	85.0	7225	209.1
	2.46	84.0	7056	206.6
	2.46	85.2	7259	209.6
			$\sum_{k=1}^{33} t_k^2 = 70,396$	$\sum_{k=1}^{33} y_k t_k = 2019$

$$b = \frac{\sum_k y_k t_k}{\sum_k t_k^2} = \frac{2019}{70,396} = 0.0287$$

$$y = \ln \left\{ \frac{c_A[1 - \sqrt{\beta}] - c_{A0}}{c_A[-1 - \sqrt{\beta}] + c_{A0}} \right\}$$

en función del tiempo debe ser lineal con una pendiente de  $2c_{A0}k_1/\sqrt{\beta}$ . Aquí  $c_{A0} = 5.50$  y  $\sqrt{\beta} = 1.115$ . Los datos de concentración—tiempo de la tabla 5.7 se muestran en la tabla 6.4, junto con los cálculos apropiados para la pendiente óptima y el valor del coeficiente de velocidad directa,  $k_1$ . La línea recta mostrada con los datos de la figura 5.6 es la calculada aquí.

TABLA 6.4 Cálculo del valor óptimo del coeficiente de velocidad directa  $k_1$ , para la reacción entre ácido sulfúrico y sulfato de dietilo, utilizando la ecuación 6.6. Los datos de la concentración con respecto al tiempo se toman de la tabla 5.7.

$t_k$	$c_A$	$y_k = \ln \left\{ \frac{c_A[1 - \sqrt{\beta}] - c_{A0}}{c_A[-1 - \sqrt{\beta}] + c_{A0}} \right\}$	$y_k t_k$	$t_k^2$
0	5.50	0	0	0
41	4.91	0.216	8.9	1681
48	4.81	0.259	12.4	2304
55	4.69	0.312	17.2	3025
75	4.38	0.467	35.0	5625
96	4.12	0.620	59.5	9216
127	3.84	0.818	103.9	16,129
146	3.62	1.009	147.3	21,316
162	3.59	1.039	168.3	26,244
180	3.44	1.200	216.0	32,400
194	3.34	1.325	257.1	37,636
212	3.27	1.423	301.7	44,944
267	3.07	1.774	473.7	71,289
318	2.92	2.156	685.6	101,124
379	2.84	2.442	925.5	143,641
410	2.79	2.676	1097.2	168,100
(∞)	2.60		$\sum t_k y_k = 4509$	$\sum t_k^2 = 684,674$

$$b = \frac{4.509 \times 10^3}{6.847 \times 10^5} = 6.59 \times 10^{-3} = \frac{2c_{A0}k_1}{\sqrt{\beta}}$$

$$c_{A0} = 5.50, \quad \sqrt{\beta} = \frac{c_{A0}}{c_{Ae}} - 1 = \frac{5.50}{2.60} - 1 = 1.115$$

$$k_1 = 6.7 \times 10^{-4}$$

## 6.4 PENDIENTE OPTIMA - II

En la sección anterior se calculó la pendiente óptima mediante procedimientos algebraicos, que es en realidad todo lo que se necesita para este problema. Antes de abordar funciones más generales, es útil resolver el problema de la pendiente óptima utilizando el cálculo. En realidad el cálculo no es necesario para lo que se ha hecho hasta aquí, pero el álgebra puede convertirse en algo tedioso y el uso del cálculo abrevia generalmente la solución.

Se toma en cuenta el hecho de que el mínimo de  $e^2$  con respecto a  $b$  ocurre en un punto, cuando se satisfacen estas dos relaciones:

$$\frac{d}{db} e^2 = 0$$

$$\frac{d^2 e^2}{db^2} > 0$$

De la ecuación 6.5:

$$\frac{de^2}{db} = -2 \sum_{k=1}^N y_k t_k + 2b \sum_{k=1}^N t_k^2$$

$$\frac{d^2 e^2}{db^2} = 2 \sum_{k=1}^N t_k^2 > 0$$

Igualando, la primera derivada a cero, se obtiene:

$$b = \frac{\sum_{k=1}^N y_k t_k}{\sum_{k=1}^N t_k^2}$$

La segunda derivada es siempre positiva.

## 6.5 LA RECTA OPTIMA

Ahora el problema consiste en encontrar la recta óptima que pase por un conjunto de datos, sin la restricción de que la línea deba de pasar a través del origen. O sea, dados los datos  $y_1, y_2, \dots, y_N$  en los puntos  $t_1, t_2, \dots, t_N$ , respectivamente, se busca la recta:

$$y = a + bt \tag{6.7}$$

la cual reduce a un mínimo la suma de los cuadrados de errores. Por tanto, se requiere el valor de *dos* variables, la pendiente,  $b$ , y la intercepción,  $a$ .

Una vez más, el error en cada punto se define como la diferencia entre el valor medido y el valor predicho por la línea recta:

$$e_k = y_k - a - bt_k$$

La suma de los cuadrados de error puede escribirse de la siguiente manera:

$$\begin{aligned} e^2 &= \sum_{k=1}^N e_k^2 = \sum_{k=1}^N [y_k - a - bt_k]^2 \\ &= \sum_{k=1}^N [y_k^2 - 2ay_k + a^2 + 2abt_k - 2by_kt_k + b^2t_k^2] \\ &= \sum_{k=1}^N y_k^2 - 2a \sum_{k=1}^N y_k + Na^2 + 2ab \sum_{k=1}^N t_k - 2b \sum_{k=1}^N y_kt_k + b^2 \sum_{k=1}^N t_k^2 \end{aligned} \tag{6.8}$$

(El término  $Na^2$  resulta de:  $\sum_{k=1}^N a^2 = a^2 + a^2 + \dots + a^2 = Na^2$ .) El mínimo de  $e^2$  con respecto a  $a$  y  $b$ , ocurre cuando las derivadas parciales con respecto a  $a$  y  $b$  desaparecen simultáneamente. Considerando primero la derivada parcial con respecto a  $a$  se tiene:

$$\begin{aligned} \frac{\partial}{\partial a} e^2 &= \frac{\partial}{\partial a} \left[ \sum_{k=1}^N y_k^2 \right] - \frac{\partial}{\partial a} \left[ 2a \sum_{k=1}^N y_k \right] + \frac{\partial}{\partial a} [Na^2] + \frac{\partial}{\partial a} \left[ 2ab \sum_{k=1}^N t_k \right] - \frac{\partial}{\partial a} \left[ 2b \sum_{k=1}^N y_kt_k \right] \\ &= \begin{matrix} \downarrow \\ 0 \end{matrix} - \begin{matrix} \downarrow \\ 2 \sum_{k=1}^N y_k \end{matrix} + \begin{matrix} \downarrow \\ 2Na \end{matrix} + \begin{matrix} \downarrow \\ 2b \sum_{k=1}^N t_k \end{matrix} - \begin{matrix} \downarrow \\ 0 \end{matrix} \\ &\qquad\qquad\qquad + \frac{\partial}{\partial a} \left[ b^2 \sum_{k=1}^N t_k^2 \right] \\ &\qquad\qquad\qquad + \begin{matrix} \downarrow \\ 0 \end{matrix} \end{aligned}$$

En forma semejante la derivada parcial con respecto a  $b$  es:

$$\frac{\partial}{\partial b} e^2 = 2a \sum_{k=1}^N t_k - 2 \sum_{k=1}^N y_kt_k + 2b \sum_{k=1}^N t_k^2$$

Igualando cada una de las derivadas parciales a cero se obtienen estas dos ecuaciones:

$$Na + \left[ \sum_{k=1}^N t_k \right] b = \sum_{k=1}^N y_k$$

## 192 Manejo de los datos experimentales

$$\left[ \sum_{k=1}^N t_k \right] a + \left[ \sum_{k=1}^N t_k^2 \right] b = \sum_{k=1}^N y_k t_k$$

Para el cálculo, algunas veces es más conveniente escribir estas ecuaciones de manera que los coeficientes de  $b$  sean de magnitud semejante a los términos del miembro derecho de las ecuaciones, esto se logra dividiendo entre el coeficiente de  $a$ :

$$a + \left[ \frac{1}{N} \sum_{k=1}^N t_k \right] b = \frac{1}{N} \sum_{k=1}^N y_k \quad (6.9a)$$

$$a + \left[ \frac{\sum_{k=1}^N t_k^2}{\sum_{k=1}^N t_k} \right] b = \frac{\sum_{k=1}^N y_k t_k}{\sum_{k=1}^N t_k} \quad (6.9b)$$

Entonces, los valores óptimos de  $a$  y  $b$  se encuentran mediante la solución simultánea de las ecuaciones 6.9.

### Ejemplo 6.3

Considere el problema de representar la densidad de una solución de cloruro de sodio-agua como una función lineal de la concentración de cloruro de sodio. Las dos primeras columnas de la tabla 6.5 muestran los datos para la gravedad específica de soluciones acuosas de NaCl a 20°C. (La gravedad específica, es la relación de la densidad a la densidad del agua pura a 4°C. La densidad del agua pura a 4°C es 1.000 g/cm<sup>3</sup>, es decir 62.43 lb<sub>m</sub>/pie<sup>3</sup>, de tal manera que la densidad de la solución en lb<sub>m</sub>/pie<sup>3</sup> se obtiene multiplicando el valor en la primera columna por 62.43.) Los datos se grafican en la figura 6.3. Se expresa la gravedad específica  $\rho/\rho_{H_2O}$ , como  $y$  y se sustituye la variable independiente  $t$  por  $c$ , es decir, la concentración. Según las ecuaciones 6.9 son necesarios los valores de la concentración elevados al cuadrado, y el producto de la concentración por la gravedad específica. Estos se muestran en la tercera y cuarta columnas de la tabla 6.5. Calculando la suma según se muestra en la tabla, se obtienen las ecuaciones:

$$a + 9.24b = 1.095$$

$$a + 13.08b = 1.135$$

TABLA 6.5 Gravedad específica (densidad en relación a la densidad del agua a 4°C) de NaCl acuoso a 20°C. Cálculo de la línea recta óptima utilizando la ecuación 6.9.

$y_k = \rho/\rho_{H_2O}$	$c_k$ (lb <sub>m</sub> /ft <sup>3</sup> )	$c_k^2$	$y_k c_k$
1.0005	0.63	0.4	0.63
1.013	1.26	1.6	1.28
1.027	2.56	6.6	2.63
1.041	3.90	15.2	4.06
1.056	5.27	27.8	5.57
1.071	6.68	44.6	7.15
1.086	8.13	66.1	8.83
1.100	9.62	92.5	10.58
1.116	11.15	124.3	12.44
1.132	12.72	161.8	14.40
1.148	14.33	205.3	16.45
1.164	15.99	255.7	18.61
1.180	17.69	312.9	20.87
1.197	19.43	377.5	23.26
$\sum y_k = 15.332$	$\sum c_k = 129.36$	$\sum c_k^2 = 1692.3$	$\sum y_k c_k = 146.76$
$\frac{1}{N} \sum y_k = 1.095$	$\frac{1}{N} \sum c_k = 9.24$	$\frac{\sum c_k^2}{\sum c_k} = 13.08$	$\frac{\sum y_k c_k}{\sum c_k} = 1.135$

con soluciones  $a=0.999$ ,  $b=0.0104$ . De esta manera, los datos se ajustan mejor mediante la ecuación:

$$\rho/\rho_{H_2O} = 0.999 + 0.0104c$$

Esta es la línea mostrada en la figura 6.3. Multiplicando por la densidad del agua a 4°C, 62.43 lb<sub>m</sub>/ft<sup>3</sup>, se obtiene la densidad a 20°C en lb<sub>m</sub>/ft<sup>3</sup>

$$\rho = 62.37 + 0.649c$$

Debe hacerse notar que aun cuando la línea recta es una buena representación de los datos, hay cierta curvatura en los mismos, y la desviación de la línea recta es sistemática no al azar. Por lo tanto, una función con una curvatura determinada podría tener mejores resultados en cuanto a los datos, aun cuando la desviación sistemática observada aquí no significará necesariamente un esfuerzo adicional.

## 194 Manejo de los datos experimentales

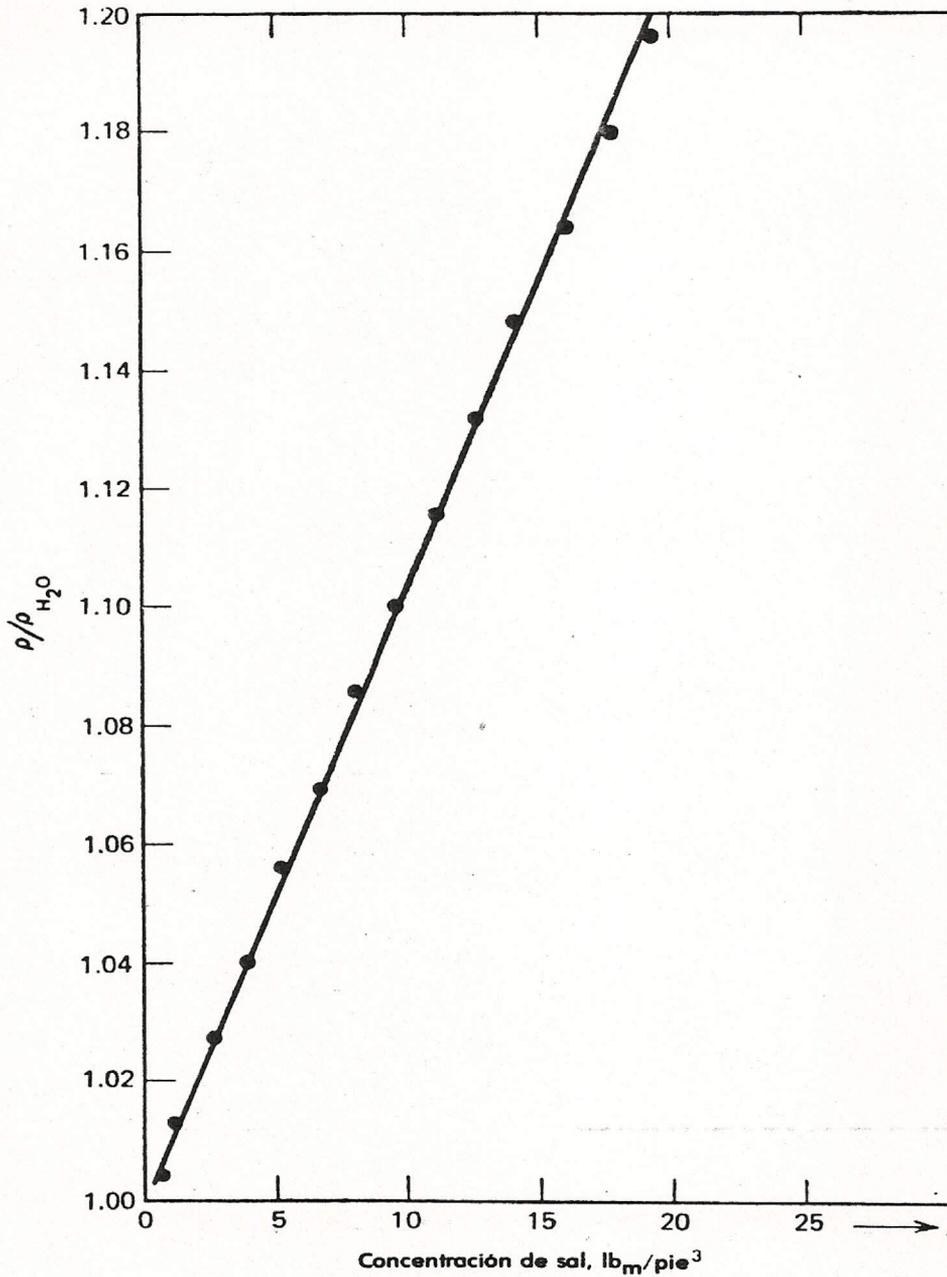


FIGURA 6.3 Gravedad específica de las soluciones de cloruro de sodio-agua a 20°C con respecto a la concentración de cloruro de sodio.  $\rho_{H_2O} = 62.43 \text{ lb}_m/\text{ft}^3$ . La recta óptima se calcula a partir de las ecuaciones 6.9.

### 6.6 CORRELACION DE UNA PROPIEDAD FISICA

Ahora es importante explicar las correlaciones de una propiedad física. Frecuentemente se tiene interés en obtener mucha información sobre propiedades físicas de materiales semejantes. Algunas veces es posible

explotar alguna característica de los constituyentes químicos para representar los datos mediante una ecuación para materiales semejantes. Cuando esto es posible, el problema de almacenar bastante información pormenorizada, se reduce a la consideración de unos cuantos parámetros en la ecuación.

Como ejemplo, considere la siguiente lista de puntos de ebullición para los ocho primeros hidrocarburos parafínicos normales.

Compuesto	Número de átomos de carbono	Punto de ebullición °C.
Metano	1	-162
Etano	2	-88
Propano	3	-42
<i>n</i> -Butano	4	1
<i>n</i> -Pentano	5	36
<i>n</i> -Hexano	6	69
<i>n</i> -Heptano	7	98
<i>n</i> -Octano	8	126

Es posible desarrollar una ecuación para los puntos de ebullición en cuanto al número de átomos de carbono, o en forma equivalente, del peso molecular, que sea suficientemente exacta como para que no sea necesario almacenar todos los datos. Además, la dispersión entre los puntos de ebullición para los isómeros normales excede el intervalo de los puntos de ebullición para los isómeros del mismo peso molecular. Como ejemplo, considere las parafinas con seis átomos de carbono.

Compuesto	Punto de ebullición, °C
<i>n</i> -Hexano	69
3 Metil pentano	64
2 Metil pentano	60
2, 3 Dimetil butano	58
2, 2 Dimetil butano	50

De esta manera, podría buscarse con relativa exactitud una ecuación que relacione el punto de ebullición con el número de átomos de carbono para todas las parafinas.

En la tabla 6.6 se da una lista de los puntos de ebullición para las 35 parafinas con 8 o menos átomos de carbono. El número del compuesto se refiere a la localización en la lista para compuestos orgánicos en el

### 196 Manejo de los datos experimentales

TABLA 6.6 Punto de ebullición (*PE*) en °C de los hidrocarburos parafínicos con ocho átomos de carbono o menos ( $n \leq 8$ ). El cálculo de la recta óptima se hace mediante la ecuación 6.9, y el de la cuadrática óptima mediante la Ecuación 6.12.

Compuesto	$n$	$PE$	$n^2$	$n \cdot PE$	$n^2 \cdot PE$	$n^3$	$n^4$
5584	1	-162	1	-162	-162	1	1
3488	2	-88	4	-176	-352	8	16
7073	3	-42	9	-126	-378	27	81
4950	4	-10	16	-40	-160	64	256
1998	4	1	16	4	16	64	256
2028	5	28	25	140	700	125	625
6481	5	36	25	180	900	125	625
7101	5	10	25	50	250	125	625
2018	6	50	36	300	1800	216	1296
2019	6	58	36	348	2088	216	1296
4556	6	69	36	414	2484	216	1296
6503	6	60	36	360	2160	216	1296
6504	6	64	36	384	2304	216	1296
2041	7	81	49	567	3969	343	2401
4458	7	98	49	686	4802	343	2401
4571	7	90	49	630	4410	343	2401
4572	7	92	49	644	4508	343	2401
6493	7	79	49	553	3871	343	2401
6494	7	90	49	630	4410	343	2401
6495	7	81	49	567	3969	343	2401
6496	7	86	49	602	4214	343	2401
6498	7	94	49	658	4606	343	2401
2040	8	107	64	856	6848	512	4096
4467	8	116	64	928	7424	512	4096
4468	8	122	64	976	7808	512	4096
4469	8	118	64	944	7552	512	4096
4564	8	114	64	912	7296	512	4096
4565	8	110	64	880	7040	512	4096
4566	8	108	64	864	6912	512	4096
4567	8	117	64	936	7488	512	4096
4569	8	119	64	952	7616	512	4096
6245	8	126	64	1008	8064	512	4096
6499	8	114	64	912	7296	512	4096
6500	8	119	64	952	7616	512	4096
6511	8	99	64	792	6336	512	4096

$$\begin{array}{ccccccc} \Sigma n_k = & \Sigma BP_k = & \Sigma n_k^2 = & \Sigma n_k BP_k = & \Sigma n_k^2 BP_k = & \Sigma n_k^3 = & \Sigma n_k^4 = \\ 226 & 2354 & 1574 & 19125 & 145705 & 11362 & 83822 \end{array}$$

$$\frac{1}{N} \Sigma n_k = 6.46 \quad \frac{1}{N} \Sigma BP_k = 67.3 \quad \frac{1}{N} \Sigma n_k^2 = 45.0$$

$$\frac{\Sigma n_k^2}{\Sigma n_k} = 6.96 \quad \frac{\Sigma n_k BP_k}{\Sigma n_k} = 84.6 \quad \frac{\Sigma n_k^3}{\Sigma n_k} = 50.3$$

$$\frac{\Sigma n_k^3}{\Sigma n_k^2} = 7.22 \quad \frac{\Sigma n_k^2 BP_k}{\Sigma n_k^2} = 92.6 \quad \frac{\Sigma n_k^4}{\Sigma n_k^2} = 53.3$$

*Handbook of Chemistry and Physics.*  $n$  se refiere al número de carbonos y  $PE$  al punto de ebullición en grados centígrados. Se busca una ecuación de la forma:

$$PE = a + bn$$

Donde  $n$  desempeña la función de  $t$  y  $PE$  la de  $y$ , según se observó en el análisis previo. La sumas necesarias para el cálculo de  $a$  y  $b$ , se evalúan en las columnas segunda a la quinta de la tabla 6.6, conduciendo a las dos ecuaciones siguientes:

$$a + 6.46b = 67.3$$

$$a + 6.96b = 84.6$$

La solución es  $a = -156.2$ ,  $b = 34.6$  o:

$$PE = -156.2 + 34.6n$$

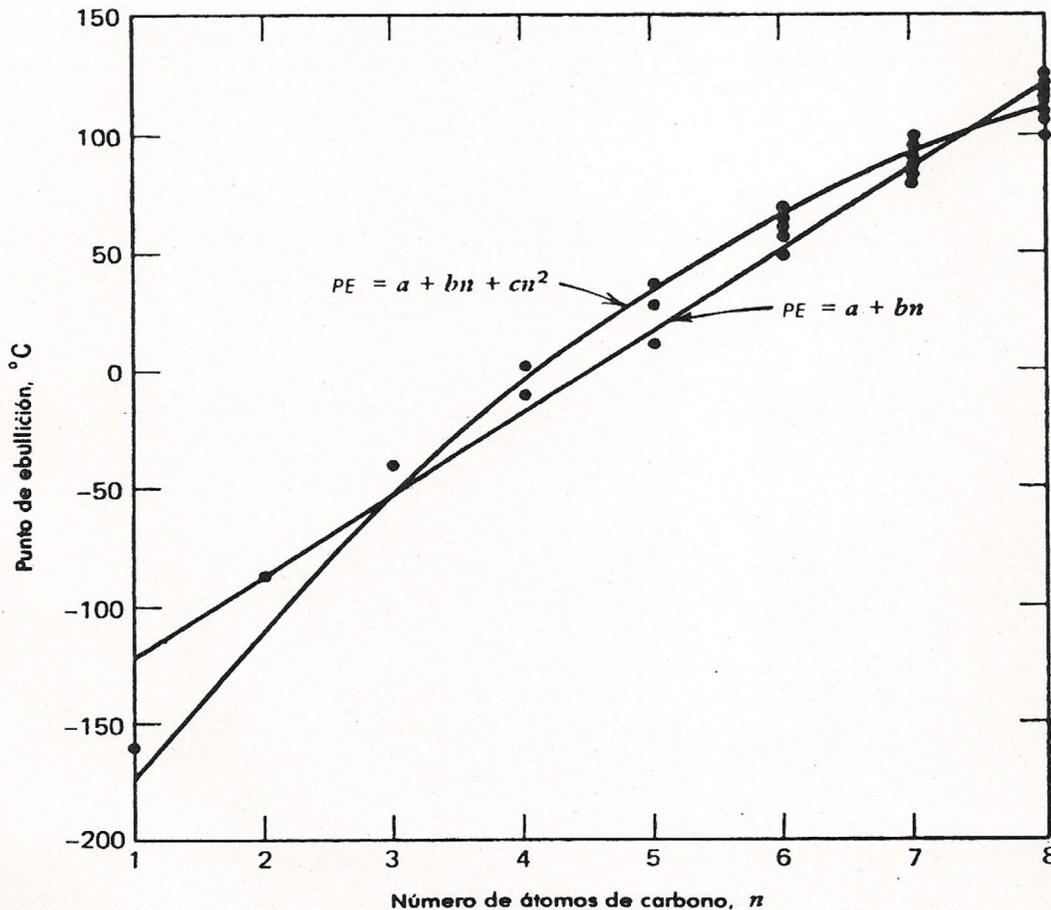


FIGURA 6.4 Puntos de ebullición de los hidrocarburos parafínicos como función del número de átomos de carbono. Las mejores funciones lineal y cuadrática se calculan utilizando las ecuaciones 6.9 y 6.12, respectivamente.

## 198 Manejo de los datos experimentales

En la figura 6.4 se muestran tanto los datos como la línea recta que se calculó. La correlación es suficientemente buena, aun cuando es evidente que la eliminación del metano, cuyo comportamiento químico es único, podría producir un mejoramiento. La tendencia de los datos implica claramente cierta curvatura, y posteriormente, se demostrará cómo se pueden ajustar estos datos mediante una ecuación cuadrática. Una correlación empírica del tipo de la que se desarrolló aquí, es muy común; ya que si se puede tolerar la inexactitud inherente al agrupamiento de los isómeros, es posible reemplazar los 35 puntos de datos por los dos parámetros en una ecuación.

### 6.7 AJUSTE DE ECUACIONES CUADRATICAS

Antes de abarcar la generalidad total, es útil considerar el problema del ajuste de una curva cuadrática a los datos:

$$y = a + bt + ct^2 \quad (6.10)$$

Para los datos  $y_1, y_2, \dots, y_N$  en los puntos  $t_1, t_2, \dots, t_N$  se buscan ahora los tres parámetros,  $a$ ,  $b$  y  $c$ , que minimizan la suma de los cuadrados de errores.

El error en cada punto es la diferencia que existe entre el valor medido y el valor predicho por la línea:

$$e_k = y_k - a - bt_k - ct_k^2$$

El error de la suma de cuadrados es:

$$e^2 = \sum_{k=1}^N e_k^2 = \sum_{k=1}^N [y_k - a - bt_k - ct_k^2]^2 \quad (6.11)$$

Para minimizar  $e^2$  se obtienen derivadas parciales con respecto a cada una de las variables,  $a$ ,  $b$  y  $c$ , y las derivadas se igualan a cero:

$$\begin{aligned} \frac{\partial}{\partial a} e^2 &= \frac{\partial}{\partial a} \sum_{k=1}^N [y_k - a - bt_k - ct_k^2]^2 \\ &= \sum_{k=1}^N \frac{\partial}{\partial a} [y_k - a - bt_k - ct_k^2]^2 \\ &= \sum_{k=1}^N 2[y_k - a - bt_k - ct_k^2] \frac{\partial}{\partial a} [y_k - a - bt_k - ct_k^2] \\ &= \sum_{k=1}^N 2[y_k - a - bt_k - ct_k^2] [-1] \end{aligned}$$

$$= -2 \left[ \sum_{k=1}^N y_k - Na - b \sum_{k=1}^N t_k - c \sum_{k=1}^N t_k^2 \right] = 0$$

Asimismo,

$$\frac{\partial}{\partial b} e^2 = -2 \left[ \sum_{k=1}^N y_k t_k - a \sum_{k=1}^N t_k - b \sum_{k=1}^N t_k^2 - c \sum_{k=1}^N t_k^3 \right] = 0$$

$$\frac{\partial}{\partial c} e^2 = -2 \left[ \sum_{k=1}^N y_k t_k^2 - a \sum_{k=1}^N t_k^2 - b \sum_{k=1}^N t_k^3 - c \sum_{k=1}^N t_k^4 \right] = 0$$

o, reordenando en forma conveniente para el cálculo:

$$a + \left[ \frac{1}{N} \sum_{k=1}^N t_k \right] b + \left[ \frac{1}{N} \sum_{k=1}^N t_k^2 \right] c = \frac{1}{N} \sum_{k=1}^N y_k \quad (6.12a)$$

$$a + \left[ \frac{\sum_{k=1}^N t_k^2}{N} \right] b + \left[ \frac{\sum_{k=1}^N t_k^3}{N} \right] c = \frac{\sum_{k=1}^N y_k t_k}{\sum_{k=1}^N t_k} \quad (6.12b)$$

$$a + \left[ \frac{\sum_{k=1}^N t_k^3}{N} \right] b + \left[ \frac{\sum_{k=1}^N t_k^4}{N} \right] c = \frac{\sum_{k=1}^N y_k t_k^2}{\sum_{k=1}^N t_k^2} \quad (6.12c)$$

Ahora estas tres ecuaciones pueden resolverse despejando las tres incógnitas  $a$ ,  $b$  y  $c$ . Comparando con las ecuaciones 6.9, puede verse que las ecuaciones desarrollan otros términos en el polinomio.

### Ejemplo 6.4

Como ejemplo del cálculo de coeficientes en una cuadrática, se tratará nuevamente la correlación de puntos de ebullición de la sección anterior. Aquí  $PE$  se sustituye por  $y$  y  $n$  por  $t$ . Al final de las tres columnas de la tabla 6.6 se muestran los términos adicionales necesarios, y las adiciones se registran al final de la tabla. Entonces las ecuaciones 6.12 se convierten en:

$$a + 6.46b + 45.0c = 67.3$$

$$a + 6.96b + 50.3c = 84.6$$

$$a + 7.22b + 53.3c = 92.6$$

## 200 Manejo de los datos experimentales

con soluciones  $a = -252.3$ ,  $b = 77.9$ ,  $c = -4.08$ . En este caso, la correlación de puntos de ebullición es:

$$PE = -252.3 + 77.9n - 4.08n^2$$

Esta ecuación es la línea curva graficada en la figura 6.4. Calculando  $e$  a partir de las ecuaciones 6.8 y 6.11, para las formas lineal y cuadrática, respectivamente, se verifica que la forma cuadrática representa mejoramiento sobre la forma lineal. Los resultados son:

$$\text{lineal: } e^2 = \sum_{k=1}^N [y_k - a - bt_k]^2 = 4746$$

$$\text{cuadrática: } e^2 = \sum_{k=1}^N [y_k - a - bt_k - ct_k^2]^2 = 3829$$

En este caso el mejoramiento no es considerable. Para saber si un mejoramiento es suficiente para justificar el uso de términos adicionales, se requiere la aplicación de ciertas pruebas estadísticas elementales que son consecuencia de las propiedades analíticas del error de la suma de cuadrados. Estas no se consideran aquí. Para este ejemplo, podría anticiparse que la complejidad adicional asociada con la forma cuadrática no se justifica, puesto que la diferencia existente entre las funciones lineal y cuadrática en cada valor de  $n$  es, en la mayoría de los casos, menor que la dispersión total en los datos.

Entre otras cosas, esta correlación es conveniente porque permite calcular propiedades físicas que de otra manera no sería posible conocer. Por ejemplo, se puede intentar predecir los puntos de ebullición de las parafinas de nueve carbonos. Para  $n = 9$  se obtiene  $PE = 118$  de la forma cuadrática, y  $PE = 155$  de la forma lineal, comparadas con un valor experimental de 150 para  $n$ -nonano, el de punto de ebullición más alto. Sin embargo, para que no se trate de emplear una correlación empírica de este tipo, en un intervalo muy distinto de aquel para el cual se desarrolló, cabe hacer notar que con la ecuación cuadrática para puntos de ebullición no se obtienen valores menores de  $PE$  cuando  $n > 8$ ; predice un punto de ebullición máximo, y finalmente predice puntos de ebullición inferiores al cero absoluto.

## 6.8 REGRESION LINEAL

Los procedimientos de ajuste de curvas que se han estudiado hasta aquí son casos especiales de lo que se conoce como *regresión lineal*. Este nombre algunas veces causa confusión, ya que evidentemente, la forma cuadrática aplicada en la sección anterior no es una función lineal. Sin

embargo, observe que los *parámetros*  $a$ ,  $b$  y  $c$  que se seleccionaron, son lineales si se considera que no son productos o potencias de los parámetros en la expresión, y los parámetros se encuentran resolviendo un conjunto de ecuaciones algebraicas lineales.

Puede generalizarse hasta cierto punto sin que se pierda la claridad básica de los ejemplos anteriores. Primero suponga que se desea ajustar los datos con una ecuación de la forma:

$$y = af(t) + bg(t) \tag{6.13}$$

donde  $f(t)$  y  $g(t)$  son funciones específicas de  $t$ , tales como potencias, senos o cosenos, exponenciales, etc. Esta expresión también es lineal en  $a$  y  $b$ . Entonces para cada dato:

$$e_k^2 = [y_k - af(t_k) - bg(t_k)]^2$$

y el error total es:

$$e^2 = \sum_{k=1}^N [y_k - af(t_k) - bg(t_k)]^2 \tag{6.14}$$

Entonces el mínimo se da para:

$$\frac{\partial e^2}{\partial a} = -2 \sum_{k=1}^N f(t_k)[y_k - af(t_k) - bg(t_k)] = 0$$

$$\frac{\partial e^2}{\partial b} = -2 \sum_{k=1}^N g(t_k)[y_k - af(t_k) - bg(t_k)] = 0$$

o, al reordenar,

$$a + \left[ \frac{\sum_{k=1}^N f(t_k)g(t_k)}{\sum_{k=1}^N f(t_k)^2} \right] b = \frac{\sum_{k=1}^N f(t_k)y_k}{\sum_{k=1}^N f(t_k)^2} \tag{6.15a}$$

$$a + \left[ \frac{\sum_{k=1}^N g(t_k)^2}{\sum_{k=1}^N f(t_k)g(t_k)} \right] b = \frac{\sum_{k=1}^N g(t_k)y_k}{\sum_{k=1}^N f(t_k)g(t_k)} \tag{6.15b}$$

La ecuación 6.6 es el caso especial obtenido cuando  $f(t) = 0$ ,  $g(t) = t$ ; y las ecuaciones 6.9 son el caso especial en el que  $f(t) = 1$  ( $\sum_{k=1}^n 1 = N$ ) y  $g(t) = t$ . Note que se obtienen aún dos ecuaciones lineales algebraicas para las dos constantes.

## 202 Manejo de los datos experimentales

En general, se trata de encontrar  $M$  constantes para una ecuación de la forma:

$$y = a_1 f_1(t) + a_2 f_2(t) + \cdots + a_M f_M(t) \quad (6.16)$$

donde  $f_1(t), f_2(t), \dots, f_M(t)$  son funciones específicas. Entonces:

$$e^2 = \sum_{k=1}^N [y_k - a_1 f_1(t_k) - a_2 f_2(t_k) - \cdots - a_M f_M(t_k)]^2 \quad (6.17)$$

$$\frac{\partial e^2}{\partial a_1} = -2 \sum_{k=1}^N f_1(t_k) [y_k - a_1 f_1(t_k) - a_2 f_2(t_k) - \cdots - a_M f_M(t_k)]$$

$$\frac{\partial e^2}{\partial a_2} = -2 \sum_{k=1}^N f_2(t_k) [y_k - a_1 f_1(t_k) - a_2 f_2(t_k) - \cdots - a_M f_M(t_k)]$$

etc.

En general, para cualquier  $a_j, j = 1, 2, \dots, M$ ,

$$\begin{aligned} \frac{\partial e^2}{\partial a_j} &= -2 \sum_{k=1}^N f_j(t_k) [y_k - a_1 f_1(t_k) - a_2 f_2(t_k) - \cdots - a_M f_M(t_k)] \\ &= 0, \quad j = 1, 2, 3, \dots, M \end{aligned}$$

o

$$a_1 \left[ \sum_{k=1}^N f_1(t_k) f_1(t_k) \right] + \cdots + a_M \left[ \sum_{k=1}^N f_M(t_k) f_M(t_k) \right] = \sum_{k=1}^N f_j(t_k) y_k$$

$j = 1, 2, \dots, M$

O sea, definiendo  $M^2$  cantidades  $A_{ij}; i = 1, 2, \dots, M; j = 1, 2, \dots, M$ ; por:

$$A_{ij} = \sum_{k=1}^N f_i(t_k) f_j(t_k) \quad \begin{array}{l} i = 1, 2, \dots, M \\ j = 1, 2, \dots, M \end{array}$$

y  $M$  cantidades  $b_i, i = 1, 2, \dots$ , por:

$$b_i = \sum_{k=1}^N f_i(t_k) y_k$$

pueden escribirse las ecuaciones que definen  $a_1, a_2, \dots, a_M$  como:

$$\begin{array}{r} A_{11}a_1 + A_{12}a_2 + \cdots + A_{1M}a_M = b_1 \\ A_{21}a_1 + A_{22}a_2 + \cdots + A_{2M}a_M = b_2 \\ \vdots \\ \vdots \\ A_{M1}a_1 + A_{M2}a_2 + \cdots + A_{MM}a_M = b_M \end{array}$$

o, en notación abreviada,

$$\sum_{j=1}^M A_{ij} a_j = b_i, \quad i = 1, 2, \dots, M \quad (6.18)$$

Por consiguiente, para resolver los problemas de regresión lineal, y obtener más de dos o tres coeficientes, es necesario poder resolver un gran número de ecuaciones lineales algebraicas simultáneas. Este problema de cálculo se considera en la sección 16.2

### 6.9 PAPEL PARA GRAFICAS

En diversas ocasiones al analizar datos, se grafica el logaritmo de una cantidad en función del tiempo. Se tiene disponible papel especial para gráficas con objeto de facilitar esta operación, y se le conoce como papel "semilogarítmico". En éste, un eje está marcado en la escala logarítmica (de base diez) en lugar de la escala aritmética. Como ejemplo, se muestra la figura 6.5, en donde la ordenada 5 debe interpretarse como  $\log 5$ , ó 0.697. De esta forma, puede evitarse la etapa intermedia en el cálculo de los logaritmos, y puede graficar directamente los datos.

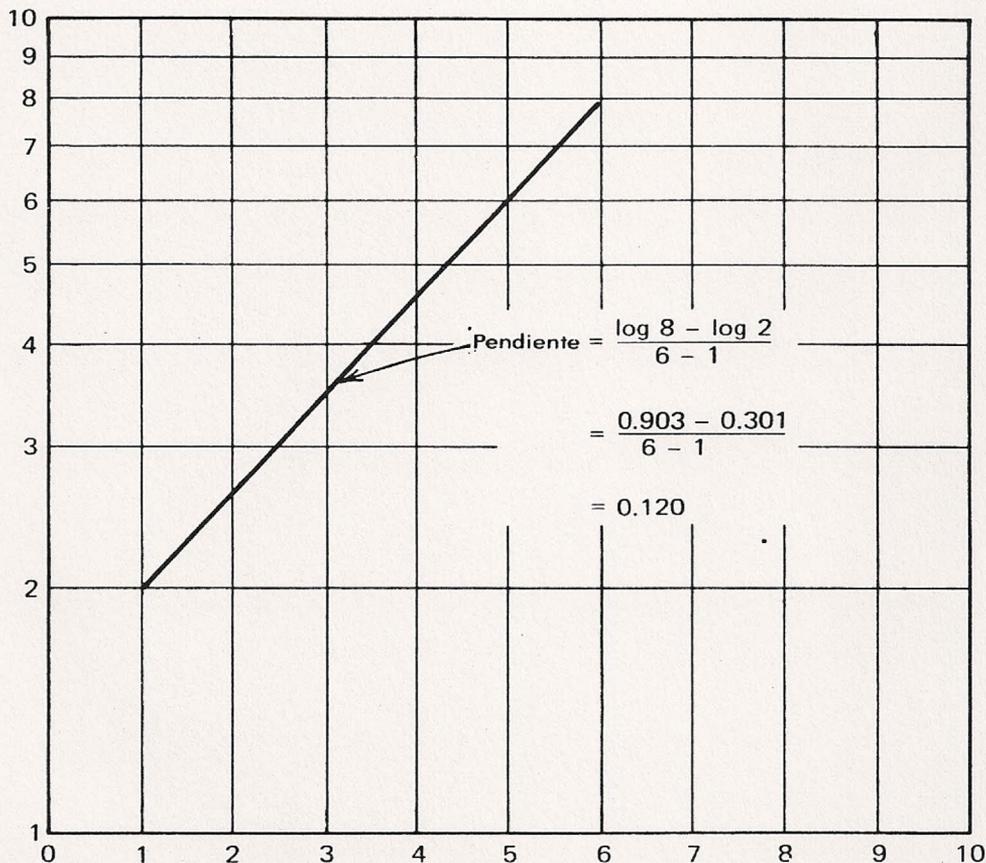


FIGURA 6.5 Papel semilogarítmico para gráficas.

## 204 Manejo de los datos experimentales

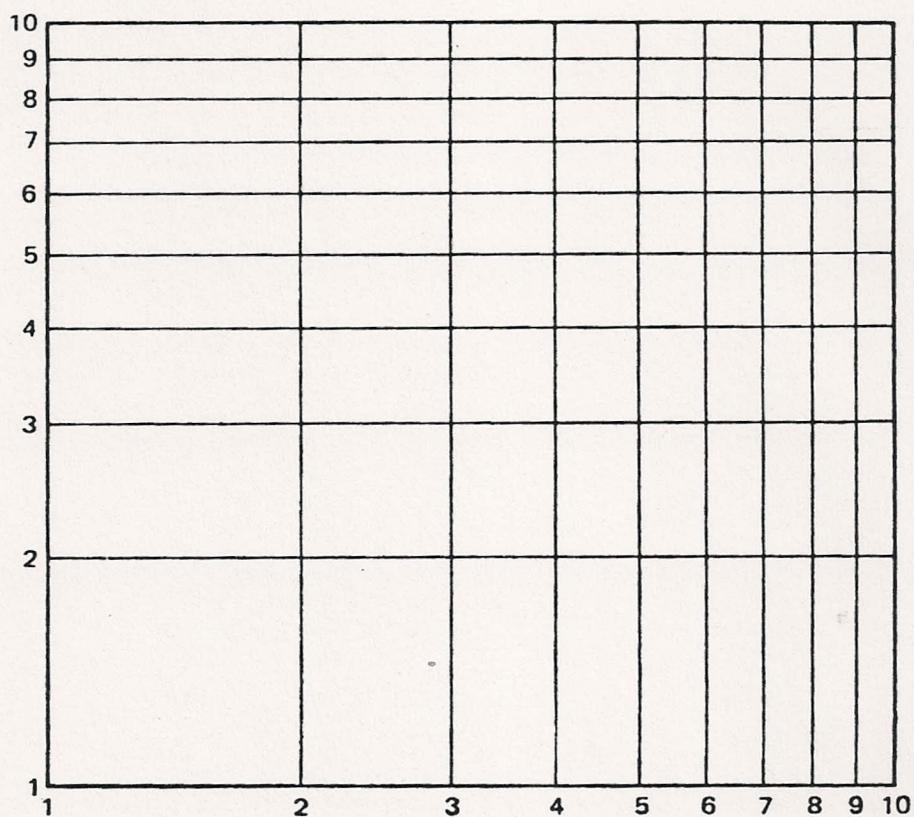


FIGURA 6.6 Papel logarítmico.

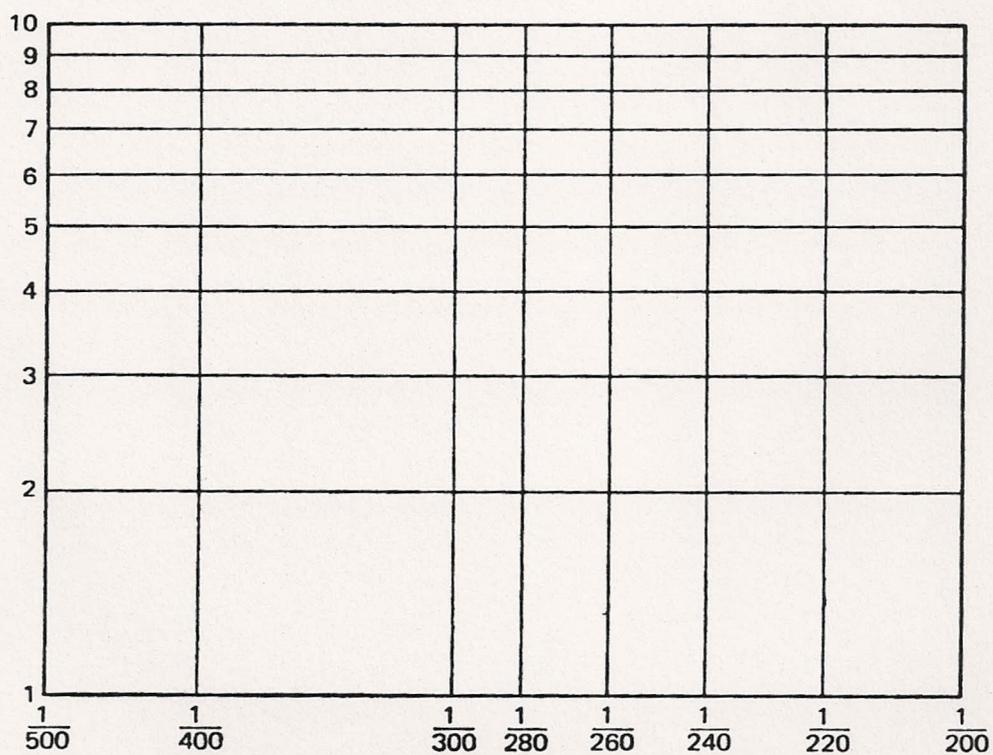


FIGURA 6.7 Papel de logaritmos e inversos para gráficas.

Debe tenerse la siguiente precaución: si se va a evaluar un parámetro a partir de la pendiente de una gráfica de  $\log c$  en función de  $t$ , deberá utilizarse la diferencia en logaritmos, y no las marcas de la escala. Esto se demuestra en la figura 6.5.

Hay otros tipos de papel. Por ejemplo, en las figuras 6.6 y 6.7 se muestran, respectivamente papel logarítmico y el de logaritmos e inversos.

## 6.10 CONCLUSIONES

El objeto primordial de este capítulo es el desarrollo de un procedimiento racional para representar los datos mediante una función. Debe hacerse notar que las técnicas analíticas explicadas, pueden aplicarse solamente cuando la función es lineal en los parámetros cuyos valores se calculan.

Las correlaciones que representan los datos de las propiedades físicas en términos de un pequeño número de parámetros son ampliamente utilizadas. La correlación para el punto de ebullición, que se utiliza aquí como ejemplo, se escogió solamente por conveniencia, y no significa que sea la única.

El análisis estadístico tiene un papel importante para determinar si ha sido bueno el ajuste de los datos mediante la línea de los cuadrados mínimos. Esto se desarrolla en textos como:

6.1 A. Hald, *Statistical Theory with Engineering Applications*, Wiley, Nueva York, 1951.

6.2 D. M. Himmelblau, *Process Analysis by Statistical Methods*, Wiley, Nueva York, 1970.

Hay un pequeño resumen en la sección 2 de *Chemical Engineers' Handbook*.

6.3 J. H. Perry, *Chemical Engineers' Handbook*, 4th ed., McGraw-Hill, Nueva York, 1963.

## 6.11 PROBLEMAS

6.1 Dados los siguiente datos:

$x$	1	2	3	4	5
$y$	5.5	9.0	17.0	17.0	25.5

Encuentre la pendiente  $m$  de la línea recta:

## 206 Manejo de los datos experimentales

$$y = mx$$

que sea la *óptima* en cada uno de los siguientes casos:

- Minimizando la suma de los cuadrados de las desviaciones.
- Minimizando la suma de los valores absolutos de las desviaciones.
- Minimizando la máxima desviación absoluta (es decir, para cada pendiente  $m$ , uno de los puntos de los datos tiene la mayor desviación absoluta. Encuentre la pendiente que hace esta mayor desviación absoluta tan pequeña como sea posible).

- 6.2 Los siguientes datos para la densidad de las mezclas etanol-agua a 20°C fueron tomados de *Chemical Engineers' Handbook*, 4th ed., p. 3-84:

% C <sub>2</sub> H <sub>5</sub> OH, en peso	0	5	10	15
Densidad, g/cm <sup>3</sup>	0.99823	0.98938	0.98187	0.97514

% C <sub>2</sub> H <sub>5</sub> OH, en peso	20	25	30	35
Densidad, g/cm <sup>3</sup>	0.96864	0.96168	0.95382	0.94494

Utilizando el criterio del mínimo de la suma de cuadrados, encuentre la recta óptima a través de los datos, en la forma de la ecuación 4.24:

$$\rho = \rho_0 + bc$$

donde  $c$  es la concentración del alcohol en moles/cm<sup>3</sup>

- Cuando se toma  $\rho_0$  como la densidad del agua pura a 20°C,  $\rho_0 = 0.99823$
- Cuando se buscan los valores óptimos tanto de  $\rho_0$  como de  $b$ , y  $\rho_0$  simplemente es una intercepción arbitraria sin significación física.

(Sugerencia: los datos deben manipularse, dado que la densidad se da en términos de por ciento en peso, no de concentración.)

- 6.3 En el problema 5.6 los datos de la constante de velocidad  $k_1$  para la reacción celulosa a glucosa como una función de la concentración de HCl,  $c_{\text{HCl}}$ , están dados a 160°C. Encuentre los valores óptimos (mínimos cuadrados) de  $k_1'$  y  $n$  en la relación:

$$k_1 = k_1' c_{\text{HCl}}^n$$

- 6.4 Utilizando los datos de la tabla 2.2, obtenga el valor óptimo (mínimos cuadrados) para el exponente  $n$ , en la ecuación del orificio  $q = kh^n$ . Tome el valor medio del logaritmo natural de  $h$  que corresponda a un

valor dado de logaritmo natural  $[-\Delta h/\Delta t]$  como la media de los dos valores de los puntos terminales.

- 6.5 Utilizando los datos en la tabla 5.3 del ejemplo 5.1, obtenga el valor óptimo (mínimos cuadrados) para el orden,  $n$ , de la reacción entre ácido sulfúrico y sulfato de dietilo en solución acuosa a 22.9°C. Tome el valor de  $\ln c_{.1}$  que corresponde a un valor dado de logaritmo natural  $[-\Delta c_{.1}/\Delta t]$  como la media de los dos valores de los puntos terminales.
- 6.6 Utilizando los datos en la tabla 5.4 que se trataron en el ejemplo 5.2, calcule el valor óptimo (mínimos cuadrados) de la constante de velocidad,  $k$ , para la reacción entre el etóxido de sodio, y el yoduro de etildimetilsulfonio en solución en etanol absoluto.
- 6.7 Encuentre las constantes óptimas (mínimos cuadrados)  $a$ ,  $b$  y  $c$  en la ecuación:

$$\rho = a + bT + cT^2$$

para los siguientes datos de densidad de una solución de  $MgSO_4$  al 4% en agua, en relación con la temperatura:

$T, ^\circ C$	0	20	30	40	50	60	80
$\rho, g/cm^3$	1.0423	1.0392	1.0362	1.0326	1.0283	1.0234	1.0118

- 6.8 Algunas veces la capacidad calorífica ( $c_p$ ) se correlaciona con la temperatura absoluta utilizando una ecuación:

$$c_p = a + bT + cT^{-2}$$

- (a) Para el conjunto dado de datos  $\{c_{pi}\}$ , tomados a las temperaturas  $\{T_i\}$ , obtenga la ecuación para calcular los valores óptimos mínimos cuadrados de  $a$ ,  $b$  y  $c$ .
- (b) Aplique el resultado a los siguientes datos para el aluminio:

$T, ^\circ C$	-200	-150	-100	-50	0
$c_p, cal/g^\circ C$	0.076	0.137	0.168	0.191	0.208

$T, ^\circ C$	20	100	300	600
$c_p, cal/g^\circ C$	0.214	0.225	0.248	0.277

## 208 Manejo de los datos experimentales

Note que los datos están en grados centígrados y es necesario convertir a °K.

- 6.9 Algunas veces es útil aproximar una función  $f(t)$ ,  $0 < t < \pi$ , mediante una *serie sinoidal de Fourier*.

$$f(t) \simeq c_1 \sin t + c_2 \sin 2t + c_3 \sin 3t + \dots$$

En el criterio de mínimos cuadrados la suma se efectúa sobre todos los valores de  $t$ , y se convierte en una integral,

$$e^2 = \int_0^\pi [f(t) - c_1 \sin t - c_2 \sin 2t - c_3 \sin 3t - \dots]^2 dt$$

- (a) Obtenga los coeficientes  $c_1, c_2, c_3, \dots$  que minimizan  $e^2$ . Note que puede derivarse una relación general para  $c_n$ .
- (b) Calcule los primeros cuatro términos para la función  $f(t) = 1$  y gráfíquelos. ¿Es adecuado el cálculo aproximado obtenido?
- (c) Calcule y gráfíque los primeros cuatro términos para la función:

$$f(t) = \begin{cases} t, & 0 < t < \pi/2 \\ \pi - t, & \pi/2 < t < \pi \end{cases}$$