

Objetivos Específicos de Aprendizagem

Ao finalizar este Capítulo, você será capaz de:

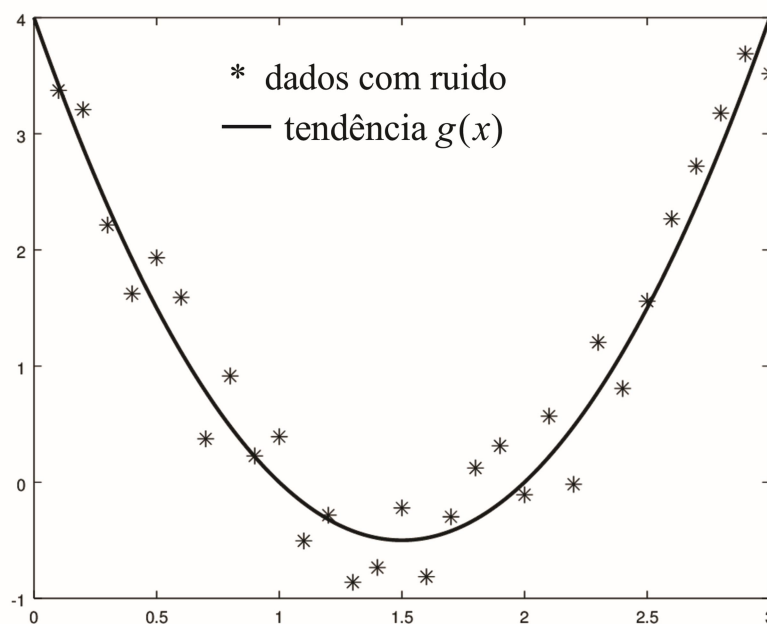
- Obter aproximadoras polinomiais definidas pela tendência de bases de dados;
- Obter aproximadoras não polinomiais com coeficientes lineares e não lineares definidas pela tendência de bases de dados; e
- Fazer uma análise introdutória da qualidade dos resultados.

7 Aproximação pela Tendência via Ajuste de Curvas

No Capítulo 5, abordamos a aproximação de uma função tabelada $(x_k, y_k = f(x_k))$ com $k = 1, 2, \dots, m$ por interpolação polinomial, *splines* cúbicas e curvas de Bézier. Agora veremos que a plotagem de tabelas oriundas de experimentos, de coleta de dados, ou de outros tipos de bases de dados, gera gráficos na forma de nuvens de pontos, também conhecidos como **diagrama de dispersão**, cuja configuração pode não ser uma função, mas globalmente sugira a **tendência** de uma função conhecida, bem como mostre que os valores de y_k podem estar afetados por erros inerentes (definidos no Capítulo 1). Nesse contexto, nenhuma das três técnicas de aproximação de funções tabeladas mencionadas será adequada para obter a função aproximadora sugerida pela **tendência dos dados**. LINK Todos os dados utilizados neste Capítulo e classificados como experimentais são fictícios. FIM DO LINK

Para visualizar como obtemos uma função aproximadora definida pela tendência de dados, observe o Gráfico 7.1.

Gráfico 7.1 – Exemplo de distribuição de dados (x_k, y_k) obtidos experimentalmente e uma função tendência $g(x)$



Fonte: Elaboração própria

Observe que os m pontos do Gráfico 7.1 seguem uma tendência parabólica que caracteriza o seu comportamento. Nesses casos, é preferível escolher uma função que represente essa tendência em vez de **tentar** forçar que uma função interpoladora passe sobre todos os m pontos que genericamente podem nem ser uma função.

A aproximação de configurações desse tipo envolve quatro etapas.

DESTAQUE

Etapa 1: para aumentar o nível de confiança nas conclusões dos experimentos, geralmente temos que coletar a maior quantidade de amostras possível.

Etapa 2: os valores amostrais coletados podem estar afetados por erros inerentes (de observação, calibragem etc.), mas que não devem afetar a função aproximadora. Assim, quanto mais pontos forem coletados, maiores serão as chances de reduzir os efeitos dos erros das medições sobre a função aproximadora.

Etapa 3: a plotagem dos pontos pode sugerir uma tendência para uma função conhecida.

Etapa 4: depois da determinação da função tendência, temos que avaliar a sua real aderência com a base de dados, bem como validar o modelo e a quantidade das amostras utilizadas.

FIM DESTAQUE

Definição 1: uma **função de ajuste** $z = g(x)$ é uma função **de coeficientes parametrizados** que **mais se aproxima** de todos os pontos amostrais (x_k, y_k) , $k = 1, 2, \dots, m$, e não necessariamente os contém.

Para obter a função de ajuste, com o requisito “que mais se aproxima de todos os pontos”, recorreremos à minimização dos **desvios** (erros ou resíduos) entre os pontos **amostrais** observados (tabelados) e a função aproximadora, na qual o desvio de cada ponto é a distância:

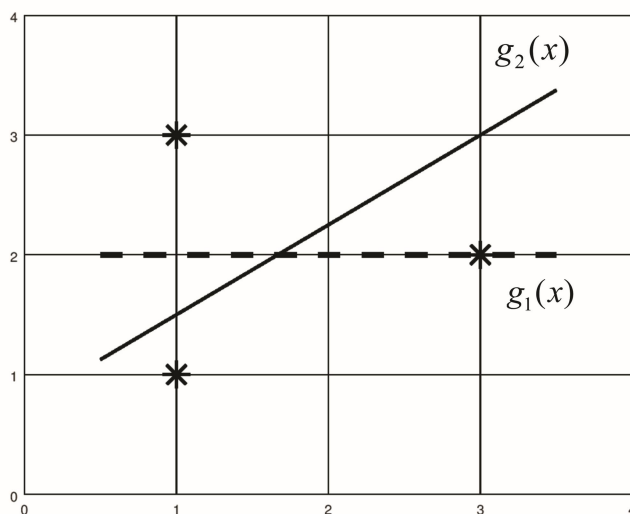
$$d_k = g(x_k) - y_k \quad (1)$$

Uma alternativa seria tentar obter a $g(x)$ obedecendo a condição de aproximação:

$$\text{desvio total } \sum_{k=1}^m d_k \text{ seja mínimo}$$

Nesse critério, os valores dos d_k com sinais opostos compensam-se, podendo resultar num desvio total pequeno, mesmo com desvios individuais altos. **Portanto, não é uma forma adequada, pois os desvios individuais também devem ser pequenos.** Também teríamos ambiguidades, isto é, existiriam várias funções que satisfariam essa alternativa para uma mesma tabela, mas não distinguiríamos o bom ajuste do ruim, conforme o Gráfico 7.2, em que ambas as aproximadoras $g_1(x)$ e $g_2(x)$ geram desvios totais nulos.

Gráfico 7.2. – Duas aproximadoras, $g_1(x)$ e $g_2(x)$, com desvios totais nulos



Para $g_1(x)$:

$$\sum_{k=1}^3 d_k = +1.0 - 1.0 + 0.0 = 0.0$$

Para $g_2(x)$:

$$\sum_{k=1}^3 d_k = +0.5 - 1.5 + 1.0 = 0.0$$

Fonte: Elaboração própria

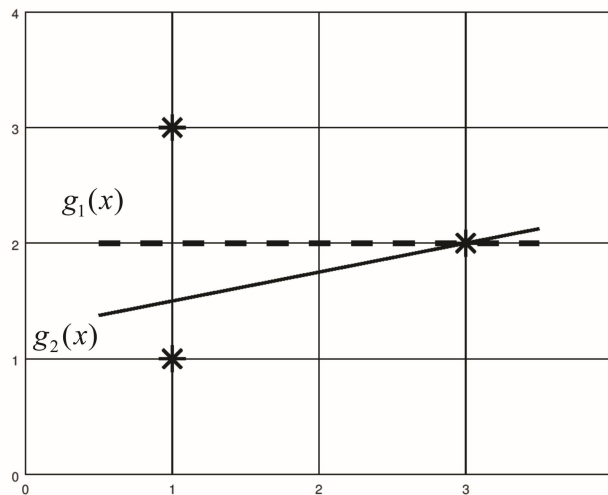
Outra possibilidade seria eliminar o problema dos sinais usando o módulo dos desvios de modo que:

$$\text{desvio absoluto total } \sum_{k=1}^m |d_k| \text{ seja mínimo}$$

Esta alternativa também não permite que possamos distinguir o bom ajuste do

ruim, pois pode ocorrer ambiguidade, conforme o Gráfico 7.3, bem como dificulta a minimização, pois a função módulo não é derivável no ponto de mínimo desvio.

Gráfico 7.3. – Duas aproximadoras, $g_1(x)$ e $g_2(x)$, com a mesma soma dos módulos dos desvios



Para $g_1(x)$:

$$\sum_{k=1}^3 |d_k| = |1.0| + |-1.0| + |0.0| = 2.0$$

Para $g_2(x)$:

$$\sum_{k=1}^3 |d_k| = |0.5| + |-1.5| + |0.0| = 2.0$$

Fonte: Elaboração própria

Note que, nesse critério, a $g_1(x)$ é a melhor aproximadora, pois é a que mais se aproxima de todos os pontos do diagrama de dispersão considerado.

Por fim, temos a alternativa:

soma quadrática dos desvios $\sum_{k=1}^m d_k^2$ seja mínima

Nesse caso, resulta que:

a) desvios negativos (ao quadrado) são adicionados no desvio total;

b) os grandes desvios são enfatizados;

c) os pequenos desvios são minimizados; e

d) o bom ajuste pode ser distinguido do ruim.

Nas duas retas de ajuste do Gráfico 7.3, temos que:

$$\text{Para } g_1(x): \sum_{k=1}^3 d_k^2 = (+1.0)^2 + (-1.0)^2 + (0.0)^2 = 2.0$$

$$\text{Para } g_2(x): \sum_{k=1}^3 d_k^2 = (+0.5)^2 + (-1.5)^2 + (0.0)^2 = 2.5$$

Observe que ocorreu menor desvio quadrático total na função $g_1(x)$, indicando

que essa é a reta que mais se aproxima de todos os pontos da amostra.

Então, a essência do ajuste consiste na **minimização do desvio quadrático total D**, LINK Por isso, a denominação dessa metodologia de aproximação por **mínimos quadrados**. FIM DO LINK cuja expressão algébrica é:

DESTAQUE

$$D = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m (g(x_k) - y_k)^2 \quad (2)$$

FIM DO DESTAQUE

LINK Observe que a condição de aproximação aplicada na interpolação, vista no capítulo 5, é que os erros ou desvios locais sobre cada ponto sejam nulos, enquanto **no ajuste a soma dos quadrados dos erros, ou desvios locais, deve ser mínima**, e normalmente não é nula. FIMLINK

Neste Capítulo, diferentemente dos demais, dispomos de um único método, uma vez que vamos obter o mínimo global da função desvio quadrático total D . As variações ocorrem apenas nos tipos de famílias a que pertencem as funções de ajuste $g(x)$ representativas da tendência do diagrama de dispersão.

A seguir, vamos apresentar o ajuste de bases de dados à família das funções polinomiais e posteriormente estendê-lo para as demais funções.

7.1 Método dos Mínimos Quadrados para Ajuste às Funções Polinomiais

Para um diagrama de dispersão com m pontos:

x_k	x_1	x_2	...	y_m
y_k	y_1	y_2	...	y_m

e um polinômio $P_n(x)$ de grau n ($n < m$),

$$P_n(x) = a_1 + a_2x + \dots + a_nx^{n-1} + a_{n+1}x^n \quad (3)$$

resulta da eq. (1) que os desvios locais são

$$d_k = P_n(x_k) - y_k = a_1 + a_2x_k + \dots + a_nx_k^{n-1} + a_{n+1}x_k^n - y_k$$

e, da eq. (2), que o desvio quadrático total D é

$$D(a_1, a_2, \dots, a_{n+1}) = \sum_{k=1}^m \left[a_1 + a_2 x_k + \dots + a_n x_k^{n-1} + a_{n+1} x_k^n - y_k \right]^2 \quad (4)$$

em que D é uma função com $n+1$ variáveis (coeficientes $(a_1, a_2, \dots, a_{n+1})$), cujo **mínimo global** fornece os coeficientes $(a_1, a_2, \dots, a_{n+1})$ do polinômio ajustador.

Para minimizar $D(a_1, a_2, \dots, a_{n+1})$, primeiramente determinamos o seu **ponto crítico**, obtendo todas as suas derivadas parciais:

$$\frac{\partial D}{\partial a_1} = 0, \quad \frac{\partial D}{\partial a_2} = 0, \quad \dots, \quad \frac{\partial D}{\partial a_{n+1}} = 0$$

e gerando as equações:

$$\begin{aligned} \frac{\partial D}{\partial a_1} &= \sum_{k=1}^m 2 \left[a_1 + a_2 x_k + \dots + a_n x_k^{n-1} + a_{n+1} x_k^n - y_k \right] * 1 = 0 \\ \frac{\partial D}{\partial a_2} &= \sum_{k=1}^m 2 \left[a_1 + a_2 x_k + \dots + a_n x_k^{n-1} + a_{n+1} x_k^n - y_k \right] * x_k = 0 \\ &\vdots \\ \frac{\partial D}{\partial a_n} &= \sum_{k=1}^m 2 \left[a_1 + a_2 x_k + \dots + a_n x_k^{n-1} + a_{n+1} x_k^n - y_k \right] * x_k^{n-1} = 0 \\ \frac{\partial D}{\partial a_{n+1}} &= \sum_{k=1}^m 2 \left[a_1 + a_2 x_k + \dots + a_n x_k^{n-1} + a_{n+1} x_k^n - y_k \right] * x_k^n = 0 \end{aligned}$$

Dividindo as equações anteriores por 2 e aplicando a distributividade dos somatórios em cada equação, temos:

$$\left\{ \begin{aligned} \frac{\partial D}{\partial a_1} &= \sum_{k=1}^m a_1 + \sum_{k=1}^m a_2 x_k^1 + \dots + \sum_{k=1}^m a_n x_k^{n-1} + \sum_{k=1}^m a_{n+1} x_k^n - \sum_{k=1}^m y_k = 0 \\ \frac{\partial D}{\partial a_2} &= \sum_{k=1}^m a_1 x_k^1 + \sum_{k=1}^m a_2 x_k^2 + \dots + \sum_{k=1}^m a_n x_k^n + \sum_{k=1}^m a_{n+1} x_k^{n+1} - \sum_{k=1}^m x_k^1 y_k = 0 \\ &\vdots \\ \frac{\partial D}{\partial a_n} &= \sum_{k=1}^m a_1 x_k^{n-1} + \sum_{k=1}^m a_2 x_k^n + \dots + \sum_{k=1}^m a_n x_k^{2n-2} + \sum_{k=1}^m a_{n+1} x_k^{2n-1} - \sum_{k=1}^m x_k^{n-1} y_k = 0 \\ \frac{\partial D}{\partial a_{n+1}} &= \sum_{k=1}^m a_1 x_k^n + \sum_{k=1}^m a_2 x_k^{n+1} + \dots + \sum_{k=1}^m a_n x_k^{2n-1} + \sum_{k=1}^m a_{n+1} x_k^{2n} - \sum_{k=1}^m x_k^n y_k = 0 \end{aligned} \right. \quad (5)$$

Note que a eq. (5) é um sistema de $n+1$ equações lineares a $n+1$ incógnitas a_i , com $i = 1, 2, \dots, n+1$ que, reescrito na forma matricial, torna-se:

$$\begin{bmatrix} m & \sum x_k^1 & \cdots & \sum x_k^{n-1} & \sum x_k^n \\ \sum x_k^1 & \sum x_k^2 & \cdots & \sum x_k^n & \sum x_k^{n+1} \\ & \vdots & & & \\ \sum x_k^n & \sum x_k^{n+1} & \cdots & \sum x_k^{2n-1} & \sum x_k^{2n} \end{bmatrix} * \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} \sum y_k \\ \sum x_k^1 y_k \\ \vdots \\ \sum x_k^n y_k \end{bmatrix}_{n+1} \quad (6)$$

SAIBA MAIS Na eq. (6), cada somatório $\sum(\dots)$ representa $\sum_{k=1}^m(\dots)$ e $\sum_{k=1}^m a_1 = \underbrace{(a_1 + a_1 + \dots + a_1 + a_1)}_{m \text{ vezes}} = m * a_1$. FIM DO SAIBA MAIS

A solução do sistema linear dado pela eq. (6) fornece os $n+1$ coeficientes do $P_n(x)$ como o ponto crítico do desvio quadrático total $D = \sum_{k=1}^m d_k^2$. Este conjunto de coeficientes é único, sempre que $m \geq n+1$, pois gera um sistema possível e determinado LINK Não pode haver menos dados que coeficientes a determinar para se obter um única função ajustadora. FIMLINK.

Esse ponto crítico é um **ponto de mínimo**, pois a função **desvio quadrático total** D gera matrizes menores principais da Hessiana **de** D com todos os determinantes positivos, de acordo com Chiang (1984), como mostraremos a **seguir para o ajuste polinomial**:

$$H_1 = \left| \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1^2} \right| = m > 0$$

$$H_2 = \begin{vmatrix} \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1^2} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1 \partial a_2} \\ \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_2 \partial a_1} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_2^2} \end{vmatrix} = \begin{vmatrix} m & \sum_{k=1}^m x_k \\ \sum_{k=1}^m x_k & \sum_{k=1}^m x_k^2 \end{vmatrix} > 0$$

$$H_3 = \begin{vmatrix} \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1^2} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1 \partial a_2} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_1 \partial a_3} \\ \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_2 \partial a_1} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_2^2} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_2 \partial a_3} \\ \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_3 \partial a_1} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_3 \partial a_2} & \frac{\partial^2 D(a_1, \dots, a_{n+1})}{\partial a_3^2} \end{vmatrix} = \begin{vmatrix} m & \sum_{k=1}^m x_k & \sum_{k=1}^m x_k^2 \\ \sum_{k=1}^m x_k & \sum_{k=1}^m x_k^2 & \sum_{k=1}^m x_k^3 \\ \sum_{k=1}^m x_k^2 & \sum_{k=1}^m x_k^3 & \sum_{k=1}^m x_k^4 \end{vmatrix} > 0$$

$$\vdots$$

e assim por diante.

DESTAQUE

Como $H_1 > 0$, $H_2 > 0$, ..., $H_{n+1} > 0$, então a matriz **Hessiana** é **positiva definida**, SAIBA MAIS **Hessiana** é a matriz gerada por todas as derivadas parciais de 2ª ordem de uma função de várias variáveis. FIM SAIBA MAIS e o ponto crítico obtido $(a_1, a_2, \dots, a_{n+1})$ é um **ponto de mínimo**.

FIMDESTAQUE

LINK Esta condição, que classifica o ponto crítico como **ponto de mínimo**, pode ser aplicada sobre o desvio quadrático total D para qualquer ajuste, seja com coeficientes lineares, ou não lineares que veremos na sessão 7.2.1.FIMLINK

Os menores principais da matriz Hessiana demonstram que o ponto crítico determinado pela eq. (6) é um ponto de mínimo desvio quadrático total. Por contradição, esse ponto crítico não poderia ser um ponto de **máximo** desvio quadrático, uma vez que sempre seria possível obter uma nova função ajustadora mais afastada dos pontos experimentais, gerando um desvio quadrático total maior.

A seguir, vamos apresentar um exemplo de ajuste de uma função linear, também conhecido como **regressão linear**, a um conjunto de pontos amostrais tabelados.

Exemplo 7.1: ajuste o conjunto de pontos tabelados a uma **reta**. Calcule os valores de $f(5)$, $f(10)$, o desvio quadrático total D obtido, plote o gráfico contendo o diagrama de dispersão e a respectiva reta ajustada.

x_k	1	3	4	6	8
y_k	0	1	2	4	5

Solução:

Temos $m = 5$ pontos e o polinômio ajustador de grau $n = 1 \Rightarrow P_1(x) = a_1 + a_2 x$.

Da eq. (6) com $n + 1 = 2$ incógnitas, resulta o sistema:

$$\begin{bmatrix} 5 & 22 \\ 22 & 126 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 12 \\ 75 \end{bmatrix} \Rightarrow \begin{cases} a_1 = -0.945205479452053 \\ a_2 = +0.760273972602739 \end{cases}$$

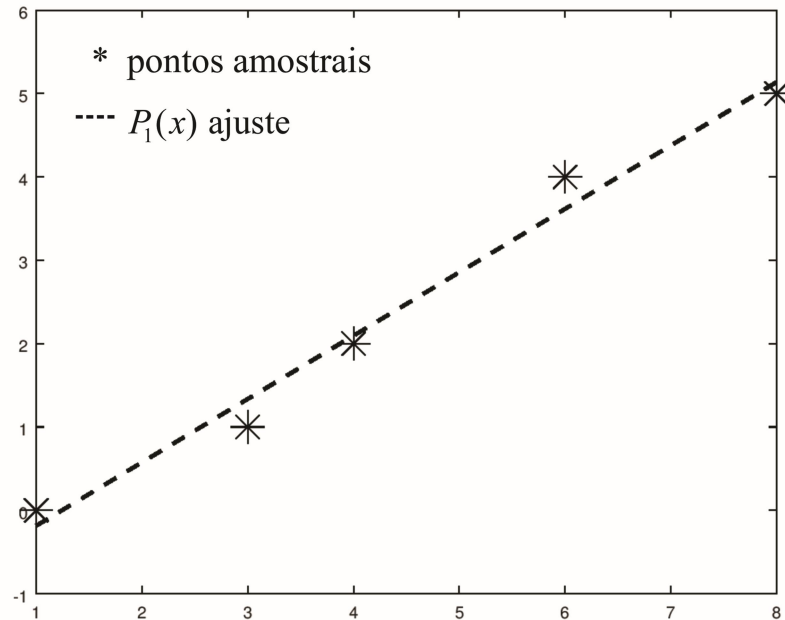
$$P_1(x) = -0.9452 + 0.7603 x$$

$$f(5) \cong P_1(5) = 2.85$$

$$f(10) \cong P_1(10) = 6.65$$

$$D = \sum_{k=1}^5 (P_1(x_k) - y_k)^2 = 0.3219$$

Gráfico 7.4 – Comportamento gráfico do **Exemplo 7.1**



Fonte: Elaboração própria

A seguir, confira algumas considerações importantes sobre o **Exemplo 7.1**.

A primeira questão a ser considerada é: **qual método é mais eficiente para resolver o sistema linear dado pela eq. (6)?**

Observe que esse sistema normalmente é de pequeno porte e formado por uma matriz densa, o que sugere a utilização de um método eliminativo para a resolução. Adicionalmente, note que é um sistema com a matriz dos coeficientes simétrica cujos elementos podem ser gerados pelas seguintes relações:

$$a_{ij} = \sum_{k=1}^m x_k^{i+j-2}$$

$$a_{i,n+1} = \sum_{k=1}^m (x_k^{i-1} * y_k)$$

e é uma matriz **positiva definida**, conforme comprovamos pelas suas Hessianas. Portanto, o método mais eficiente de solução desse tipo de sistema é o de **Cholesky**, que utiliza apenas $O(n^3 / 6)$ operações aritméticas.

Outro ponto importante é que o sistema linear da eq. (6) é mal condicionado, como podemos verificar avaliando o seu grau de condicionamento, conforme vimos no Capítulo 2, por:

$$\|\det(A)\| = \frac{|\det(A)|}{\prod_{i=1}^{n+1} \alpha_i}, \text{ em que } \alpha_i = \sqrt{\sum_{j=1}^m a_{ij}^2}$$

No sistema do **Exemplo 7.1**, $\|\det(A)\| = 0.05059$ é um valor relativamente pequeno se comparado com a unidade, sendo considerado como mal condicionado. Logo, devemos utilizar um método que envolva a menor quantidade de operações aritméticas possível para minimizar os efeitos dos arredondamentos como o método de Cholesky.

Como o $D = \sum_{k=1}^m d_k^2$ é um valor absoluto e, para uma tabela com m valores, podemos obter por mínimos quadrados até $m-1$ polinômios de ajuste, $P_n(x)$, com graus $n = 1, 2, \dots, m-1$, ou ainda a tabela pode ter configuração com tendência não polinomial, devemos também utilizar um critério relativo de aferição de quão bem ajustada está a $g(x)$.

No **Exemplo 7.1**, temos que $\sum_{k=1}^m d_k^2 = 0.3219$, e, como é um valor absoluto, nada podemos afirmar sobre ele ser pequeno ou grande, isto é, se a reta de ajuste é confiável ou não. A única garantia que temos é que: a reta ajustada é a **melhor** entre as infinitas retas existentes.

Para quantificar a qualidade de uma regressão linear $g(x)$, obtida de uma base de dados (x_k, y_k) , $k = 1, 2, \dots, m$, podemos utilizar uma medida relativa denominada **coeficiente de determinação** (BARBETTA; REIS; BORNIA, 2010), cuja expressão na sua forma simplificada é dada por:

$$R^2 = 1 - \frac{\sum_{k=1}^m (g(x_k) - y_k)^2}{\sum_{k=1}^m (\bar{y} - y_k)^2} \quad (7)$$

em que \bar{y} é o valor médio dos valores de y_k , o termo $\sum_{k=1}^m (g(x_k) - y_k)^2$ corresponde à soma D dos quadrados dos desvios (também denominada de SQE), o termo

$\sum_{k=1}^m (\bar{y} - y_k)^2$ corresponde à soma de quadrados totais (SQT), corrigida pela média aritmética, e a diferença $SQR = SQT - SQE$ é a soma de quadrados da regressão (BARBETTA; REIS; BORNIA, 2010).

Podemos interpretar esse coeficiente como a quantidade de variabilidade nos dados, que é explicada pelo modelo ajustado. Portanto, quanto maior o coeficiente de determinação (mais próximo da unidade), maior é a qualidade do ajuste do modelo aos dados (BARBETTA; REIS; BORNIA, 2010).

A seguir, vamos apresentar um exemplo de ajuste de uma função linear e parabólica, para introduzir uma análise de **qualidade** dos ajustes, que é conhecida como **Goodness of fit**. LINK Para saber mais sobre essa análise de **qualidade**, acesse o link: <<http://www.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html>>. Acesso em: 30 nov. 2016. FIM DO LINK

Exemplo 7.2: para a base de dados:

x_k	0.00	0.25	0.50	0.75	1.00
y_k	1.00	1.32	1.79	1.64	1.41

Determine:

- a função de ajuste polinomial linear $n = 1$;
- a função de ajuste polinomial parabólico $n = 2$;
- os coeficientes de determinação de cada função ajustadora em relação ao pontos tabelados;
- os desvios quadráticos totais; e
- o gráfico com os m pontos da tabela, as duas funções ajustadoras e decida qual é o resultado mais confiável.

Solução:

- Ajuste polinomial linear: $n = 1$ e $m = 5 \Rightarrow P_1(x) = a_1 + a_2x$

Montando o sistema de equações dado pela eq. (6), temos

$$\begin{bmatrix} 5.0 & 2.5 \\ 2.5 & 1.875 \end{bmatrix} * \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 7.1600 \\ 3.8650 \end{bmatrix} \Rightarrow \begin{cases} a_1 = 1.2040 \\ a_2 = 0.4560 \end{cases} \quad (\text{pelo Método Cholesky})$$

$$P_1(x) = 1.2040 + 0.4560 x \Rightarrow \text{melhor reta possível para os } m = 5 \text{ pontos.}$$

b) Ajuste polinomial parabólico: $n = 2$ e $m = 5 \Rightarrow P_2(x) = a_1 + a_2x + a_3x^2$

Da eq. (6) temos

$$\begin{bmatrix} 5.0 & 2.5 & 1.875 \\ 2.5 & 1.875 & 1.5625 \\ 1.875 & 1.5625 & 1.3828125 \end{bmatrix} * \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 7.1600 \\ 3.8650 \\ 2.8625 \end{bmatrix}$$

$$\Rightarrow \begin{cases} a_1 = 0.958285714285714 \\ a_2 = 2.421714285714286 \\ a_3 = -1.965714285714284 \end{cases} \quad (\text{Cholesky})$$

$$P_2(x) = 0.958285714285714 + 2.421714285714286 x - 1.965714285714284 x^2$$

\Rightarrow melhor parábola possível para os $m = 5$ pontos

c) Coeficientes de Determinação:

Para a reta $\Rightarrow R_1^2 = 0.350220976608818$

e para a parábola $\Rightarrow R_2^2 = 0.919678467484872$

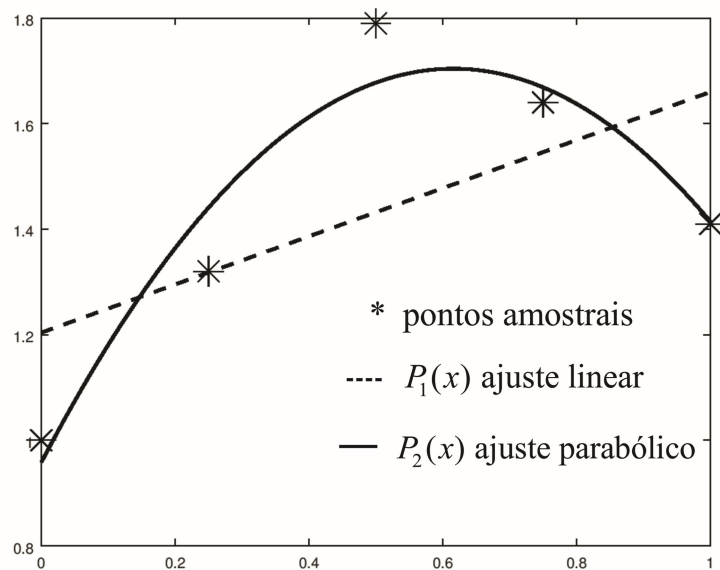
d) Desvios quadráticos totais:

Para a reta $\Rightarrow D_1 = 0.24112$

e para a parábola $\Rightarrow D_2 = 0.0298057142857143$

e)

Gráfico 7.5 – Comportamento dos polinômios ajustados de grau 1 (linha tracejada) e grau 2 (linha contínua) do Exemplo 7.2



Fonte: Elaboração própria

No **Exemplo 7.2**, o ajuste pela parábola resultou em melhores estatísticas (menor desvio quadrático total, maior coeficiente de determinação e gráfico consistente) em comparação com o ajuste pela reta. Lembremos que essa é uma avaliação preliminar, não pode ser considerada conclusiva, pois dispomos de poucos pontos amostrais.

Vários autores ainda recomendam uma análise qualitativa dos resíduos gerados pelos desvios locais d_k entre a função ajustadora e os pontos amostrais. Segundo Barbeta, Reis e Bornia (2010), para validar o modelo de ajuste e o intervalo de confiança que contém os pontos amostrais, seria desejável obter um diagrama de dispersão dos resíduos (desvios locais) com distribuição aleatória para diferentes intervalos da função aproximadora, possibilitando detectar alguma estrutura adicional nesses pontos (o que sugeriria uma função diferente) ou algum ponto discrepante que precisaria ser analisado a parte.

No **Caderno de Algoritmos**, disponível para *download* em <http://sergiopeters.prof.ufsc.br/livro-calculo-numerico-computacional/>, apresentamos o algoritmo **Cap7exemplo.7.2.m**, que ajusta uma base com m amostrais à função polinomial de grau n (n menor que m).

E como muitas bases de dados não possuem tendência polinomial, também vamos

estudar o ajuste de pontos tabelados às funções não polinomiais.

7.2 Ajuste por Mínimos Quadrados às Funções não Polinomiais

Quando a plotagem dos pontos da base de dados:

x_k	x_1	x_2	\dots	y_m
y_k	y_1	y_2	\dots	y_m

sugerir uma tendência não polinomial, temos duas alternativas para obter a função de ajuste: determinação direta da ajustadora não polinomial ou tentar transformar a forma não polinomial em polinomial através de transformações paramétricas.

7.2.1 Determinação Direta da Ajustadora não Polinomial

Como efetuamos para funções polinomiais, podemos deduzir diretamente o modelo matemático que irá fornecer a $g(x)$ representativa do comportamento dos pontos tabelados minimizando a eq. (2), conforme os casos a seguir:

- a) Ajuste de m pontos tabelados a uma função do tipo: $g(x) = a_1 x^2 + a_2 \ln(x)$

Para obter os parâmetros a_1 e a_2 dessa função, procedemos a minimização direta do desvio quadrático total D :

$$D(a_1, a_2) = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m [a_1 x_k^2 + a_2 \ln(x_k) - y_k]^2$$

Para minimizar $D(a_1, a_2)$, obtemos as suas derivadas parciais em relação a a_1 e a_2 , resultando em um sistema **linear**:

$$\frac{\partial D}{\partial a_1} = 0 \quad \text{e} \quad \frac{\partial D}{\partial a_2} = 0$$

em que

$$\frac{\partial D(a_1, a_2)}{\partial a_1} = \sum_{k=1}^m 2[a_1 x_k^2 + a_2 \ln(x_k) - y_k] x_k^2 = 0$$

$$\frac{\partial D(a_1, a_2)}{\partial a_2} = \sum_{k=1}^m 2[a_1 x_k^2 + a_2 \ln(x_k) - y_k] \ln(x_k) = 0$$

Reescrevendo na forma de equações lineares, temos

$$\begin{cases} \left(\sum x_k^4 \right) a_1 + \left(\sum (x_k^2 \ln x_k) \right) a_2 = \sum (y_k x_k^2) \\ \left(\sum (x_k^2 \ln x_k) \right) a_1 + \left(\sum (\ln x_k)^2 \right) a_2 = \sum (y_k \ln x_k) \end{cases}$$

Note que as duas equações formam um sistema linear porque a função aproximadora $g(x) = a_1 x^2 + a_2 \ln(x)$ é **linear em relação aos coeficientes** a_1 e a_2 e podemos tentar resolvê-lo via método de Cholesky, uma vez que é uma matriz simétrica. Apenas não sabemos previamente se é positiva definida. Caso não seja, podemos usar a eliminação Gaussiana.

- b) Ajuste de m pontos tabelados a uma função do tipo $g(x) = a_1 x_k^2 + \text{sen}(a_2 x_k)$.

Procedendo a minimização direta do desvio quadrático total D , conforme a eq. (2), temos

$$D(a_1, a_2) = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m \left[a_1 x_k^2 + \text{sen}(a_2 x_k) - y_k \right]^2$$

Para minimizar $D(a_1, a_2)$, obtemos as derivadas de D em relação a a_1 e a_2 e geramos um sistema **não linear**:

$$\begin{aligned} \frac{\partial D(a_1, a_2)}{\partial a_1} &= 2 \sum_{k=1}^m \left[a_1 x_k^2 + \text{sen}(a_2 x_k) - y_k \right] \left[x_k^2 \right] = 0 \\ \frac{\partial D(a_1, a_2)}{\partial a_2} &= 2 \sum_{k=1}^m \left[a_1 x_k^2 + \text{sen}(a_2 x_k) - y_k \right] \left[x_k \cos(a_2 x_k) \right] = 0 \end{aligned}$$

Ou de forma equivalente:

$$\begin{cases} f_1(a_1, a_2) = \sum_{k=1}^m \left[a_1 x_k^2 + \text{sen}(a_2 x_k) - y_k \right] \left[x_k^2 \right] = 0 \\ f_2(a_1, a_2) = \sum_{k=1}^m \left[a_1 x_k^2 + \text{sen}(a_2 x_k) - y_k \right] \left[x_k \cos(a_2 x_k) \right] = 0 \end{cases}$$

Note que as funções $f_1(a_1, a_2)$ e $f_2(a_1, a_2)$ são compostas de somatórios aplicados sobre todos os m pontos, e a função representativa $g(x) = a_1 x^2 + \text{sen}(a_2 x)$ é **não linear em relação aos coeficientes** a_1 e a_2 . Nesses casos, conforme vimos no Capítulo 4, temos que tentar obter os valores de a_1 e a_2 pelo método de **Newton**, a partir de uma solução inicial $(a_1^{(0)}, a_2^{(0)})$.

A seguir, vamos ver um exemplo de ajuste de um parâmetro $P(v,T)$, dependente de dois parâmetros independentes v e T . Mas note que o ajuste é aplicado apenas ao parâmetro P , portanto trata-se de ajuste em uma única variável.

Exemplo 7.3: a equação de estado de **Redlich–Kwong** permite relacionar propriedades termodinâmicas de gases reais com P (pressão) como propriedade dependente de v (volume específico) e T (temperatura), conforme modelo:

$$P(v,T) = \frac{R \cdot T}{(v - a_2)} - \frac{a_1}{\sqrt{T} \cdot v \cdot (v + a_2)}$$

Em que $R = 8.314 \text{ (J) / (mol} \cdot \text{K)}$ é a constante universal dos gases, e os valores de a_1 e a_2 são parâmetros de cada gás que podem ser determinados a partir de valores de propriedades físicas e de P , v e T medidos experimentalmente.

Como toda medição experimental possui erros inerentes, então procedemos uma série de m medições com o intuito de compensar os erros de uma medida para outra. Nesse exemplo, efetuamos 20 medições do parâmetro pressão P , que é função do volume específico v e da temperatura T . Descartamos 2 medições mais afetadas por erros de medição e então serão consideradas $m = 18$ medições efetivas para a determinação dos parâmetros a_1 e a_2 , conforme segue (medições no formato Octave):

$m = 18$ % Número de pontos experimentais

$v = [5.0 \ 5.2 \ 5.4 \ 5.6 \ 5.8 \ 6.0 \ 6.2 \ 6.4 \ 6.6 \ 6.8 \ 7.0 \ 7.2 \ 7.4 \ 7.6 \ 7.8 \ 8.0 \ 8.2 \ 8.4 \ 8.6 \ 8.8 \ 9.0];$ % (m³/mol)

$T = [300. \ 310.5 \ 321. \ 331.5 \ 342. \ 352.5 \ 363. \ 373.5 \ 384.0 \ 394.5 \ 405.0 \ 415.5 \ 426.0 \ 436.5 \ 447.0 \ 457.5 \ 468. \ 478.5 \ 489. \ 499.5 \ 510.];$ % (K)

$P = [623. \ 614. \ 606. \ 599. \ 592. \ 586. \ 580. \ 575. \ 570. \ 565. \ 561. \ 557. \ 553. \ 549. \ 546. \ 543. \ 540. \ 537. \ 534. \ 532. \ 530.];$ % (Pa)

- Determine os parâmetros a_1 e a_2 pela minimização do desvio quadrático total;
- Calcule também o seu coeficiente de determinação R^2 ;
- Plote um gráfico com a superfície P em função de v e T .

Solução:

a). A função desvio total quadrático de $P(v,T)$, conforme a eq. (2), será:

$$D(a_1, a_2) = \sum_{k=1}^m \left(\frac{R \cdot T_k}{(v_k - a_2)} - \frac{a_1}{\sqrt{T_k} \cdot v_k \cdot (v_k + a_2)} - P_k \right)^2 \quad (8)$$

Para determinar os parâmetros a_1 e a_2 , vamos obter o ponto crítico de D em relação a a_1 e a_2 :

$$\frac{\partial D(a_1, a_2)}{\partial a_1} = 2 \sum_{k=1}^m \left(\frac{R^* T_k}{(v_k - a_2)} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)} - P_k \right)^1 \left(-\frac{1}{\sqrt{T_k} * v_k (v_k + a_2)} \right) = 0 \quad (9)$$

$$\frac{\partial D(a_1, a_2)}{\partial a_2} = 2 \sum_{k=1}^m \left(\frac{R^* T_k}{(v_k - a_2)} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)} - P_k \right)^1 \left(R^* T_k \frac{(-1)(-1)}{(v_k - a_2)^2} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)^2} \right) = 0 \quad (10)$$

Note que as eqs. (9) e (10) formam um sistema de equações não lineares, conforme segue:

$$\begin{cases} f_1(a_1, a_2) = \sum_{k=1}^m \left(\frac{R^* T_k}{(v_k - a_2)} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)} - P_k \right)^1 \left(-\frac{1}{\sqrt{T_k} * v_k (v_k + a_2)} \right) = 0 \\ f_2(a_1, a_2) = \sum_{k=1}^m \left(\frac{R^* T_k}{(v_k - a_2)} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)} - P_k \right)^1 \left(R^* T_k \frac{(-1)(-1)}{(v_k - a_2)^2} - \frac{a_1}{\sqrt{T_k} * v_k (v_k + a_2)^2} \right) = 0 \end{cases} \quad (11)$$

Aplicando o método de Newton, abordado no Capítulo 4, com derivadas calculadas numericamente a partir da solução inicial $(a_1^{(0)}, a_2^{(0)}) = (1.0, 1.0)$ LINK Em alguns sistemas não lineares é necessário efetuar várias tentativas de valores iniciais até se obter um solução. Além disso, esses podem ter mais de uma solução FIMLINK, resulta a solução:

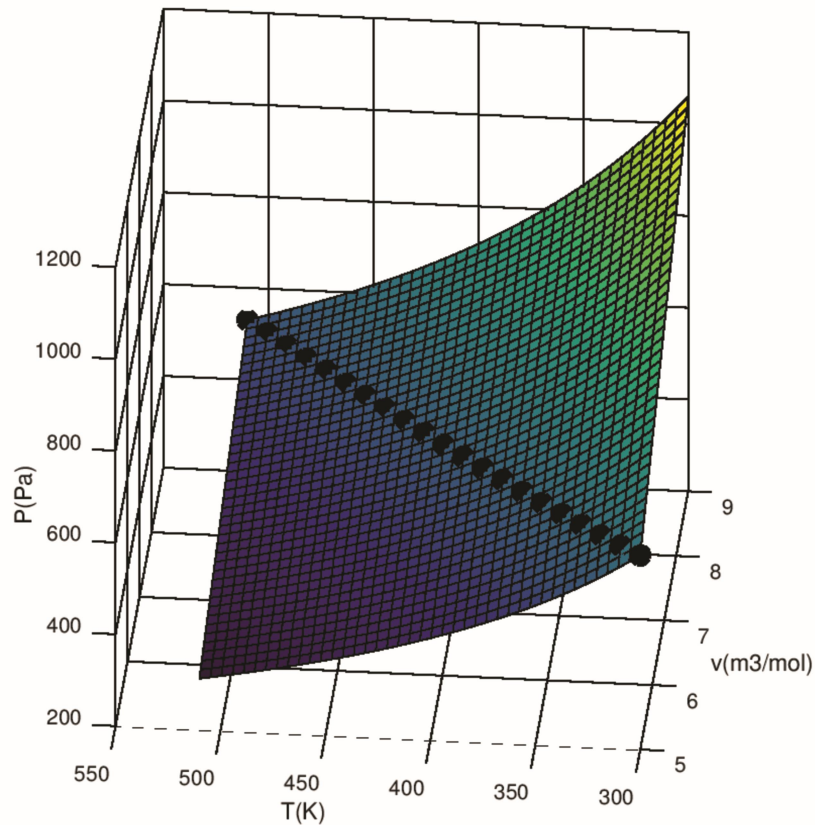
$$a_1 = 496.794871172726 \text{ e } a_2 = 1.00039053653358$$

(obtidos em 5 iterações e critério relativo de parada $7.95e-14$).

b). O coeficiente de determinação obtido aplicando a eq. (7) é: $R^2 = 0.999861930801605$.

c). Para traçar o gráfico da superfície P em função de v e T , usamos o algoritmo disponível no **Caderno de Algoritmos**, no arquivo **Cap7exem7.3.PVT.m**.

Gráfico 7.6 – Gráfico 3D de P em função de v e T do **Exemplo 7.3**



Fonte: Elaboração própria

SAIBA MAIS Se tivéssemos dois pontos com precisão adequada, (P_1, v_1, T_1) e (P_2, v_2, T_2) , poderíamos substituí-los diretamente na equação de estado de **Redlich-Kwong**, fazendo a função $P(v, T)$ passar sobre esses dois pontos, ou seja, com erros locais nulos: $P(v_k, T_k) - P_k = 0$. Desta forma, geramos um sistema de duas equações não lineares, conforme segue:

$$\begin{cases} F_1(a_1, a_2) = \frac{R \cdot T_1}{(v_1 - a_2)} - \frac{a_1}{\sqrt{T_1} \cdot v_1 \cdot (v_1 + a_2)} - P_1 = 0 \\ F_2(a_1, a_2) = \frac{R \cdot T_2}{(v_2 - a_2)} - \frac{a_1}{\sqrt{T_2} \cdot v_2 \cdot (v_2 + a_2)} - P_2 = 0 \end{cases} \quad (12)$$

Assim, poderíamos determinar as duas incógnitas, a_1 e a_2 , por uma interpolação, também usando o método de Newton. Observe também que o sistema dado pelas eqs. (12) tem solução de convergência mais estável e pode determinar valores a_1 e a_2 a partir de dois pontos escolhidos (P_1, v_1, T_1) e (P_2, v_2, T_2) da amostra, cuja solução pode servir de valor inicial $(a_1^{(0)}, a_2^{(0)})$ para as eqs. (11). FIM SAIBA MAIS

Nos casos de ajuste a funções com coeficientes não lineares, como no **Exemplo 7.3**, o R^2 não deve ser o único critério para qualificar o ajuste, uma vez que pode até gerar resultados aparentemente excelentes, mas não ser suficientemente próximo dos pontos amostrais (Gráfico 7.6).

A seguir, vamos mostrar algumas funções que permitem uma transformação paramétrica para a forma polinomial.

7.2.2 Ajuste por Transformações de Forma Paramétrica em Polinomial

Algumas funções não polinomiais permitem a sua transformação para uma forma polinomial através de artifícios algébricos. Para tal, podemos resolver o ajuste polinomial transformado, via eq. (6), e retornar à forma não polinomial de origem posteriormente, conforme alguns casos apresentados a seguir.

7.2.2.1 Ajuste a Exponenciais: $y = a * b^x$

Para tirar a variável x do expoente, aplicamos propriedades dos logaritmos na exponencial:

$$\ln(y) = \ln(a) + x * \ln(b)$$

Fazendo $\ln(y) = z$; $\ln(a) = a_1$; $\ln(b) = a_2$, temos

$$z = a_1 + a_2 * x \text{ (que é um polinômio de 1º grau)}$$

Então efetuamos o ajuste linear ao conjunto de pontos (x_k, z_k) , em que $z_k = \ln(y_k)$, para determinar a_1 e a_2 . Obtidos os coeficientes a_1 e a_2 do polinômio, via eq. (6), recuperamos os valores originais de a e b através da definição de logaritmo pelas relações:

$$\ln(a) = a_1 \rightarrow a = e^{a_1}$$

$$\ln(b) = a_2 \rightarrow b = e^{a_2}$$

Exemplo 7.4: ajuste a tabela

x_k	1	3	4	6
y_k	2.5	13	22	36

a uma curva exponencial $y = a * b^x$. Estime $f(2.5)$, calcule o valor de x correspondente a $y = 50$, faça um gráfico da função ajustada e dos pontos tabelados.

Solução:

$$y = a * b^x \Rightarrow \ln(y) = \ln(a) + x * \ln(b)$$

$$P_1(x) = a_1 + a_2 * x$$

Com $\ln(y) = z$, $\ln(a) = a_1$, $\ln(b) = a_2$.

Montando um sistema para $(x_k, z_k = \ln(y_k))$, temos

x_k	1	3	4	5	6
y_k	2.5	8	13	22	40
$z_k = \ln(y_k)$	0.916290731874	2.0794415416798	2.5649493574615	3.0910424533583	3.688879454

Pela eq. (6) temos

$$\begin{bmatrix} m & \sum x_k \\ \sum x_k & \sum x_k^2 \end{bmatrix} * \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum z_k \\ \sum x_k^n z_k \end{bmatrix}$$

$$\begin{bmatrix} 5 & 19 \\ 19 & 87 \end{bmatrix} * \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 12.34060353848778 \\ 55.00290177823501 \end{bmatrix} \Rightarrow \begin{cases} a_1 = 0.386180730567183 \\ a_2 = 0.547878941350098 \end{cases}$$

Retornando aos valores originais de a e b , temos

$$a = e^{0.386180730567183} = 1.47135056480854$$

$$b = e^{0.547878941350098} = 1.72958058266343$$

Logo,

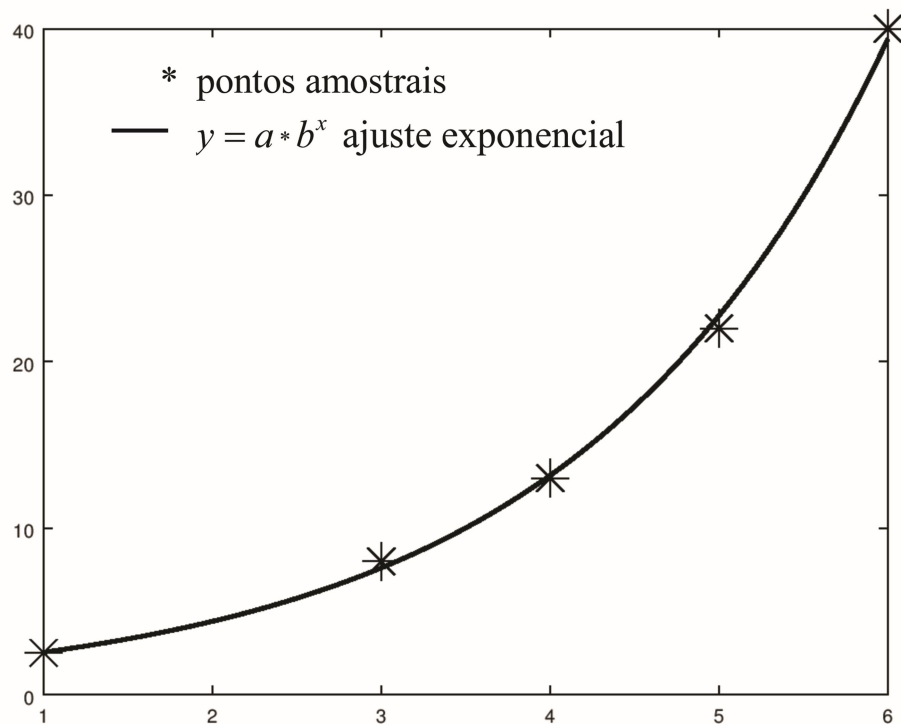
$y = 1.47135056480854 * (1.72958058266343)^x$ é a exponencial ajustadora.

Estimando:

$$f(2.5) = 1.47135056480854 * (1.72958058266343)^{2.5} = 5.78852832481353$$

$$50 = 1.47135056480854 * (1.72958058266343)^x \Rightarrow x = 6.43544040253215$$

Gráfico 7.7 – Ajuste do **Exemplo 7.4**



Fonte: Elaboração própria

7.2.2.2 Ajuste a Curvas Geométricas: $y = a * x^b$, com $b \in \mathbb{R}$

Aplicamos as propriedades dos logaritmos

$$\ln(y) = \ln(a) + b * \ln(x)$$

Fazendo $\ln(y) = z$; $\ln(a) = a_1$; $b = a_2$ e $\ln(x) = t$, temos $z = a_1 + a_2 * t$, que é um polinômio de 1º grau em (t_k, z_k) . Do mesmo modo, efetuamos o ajuste linear ao conjunto de pontos (t_k, z_k) , em que $t_k = \ln(x_k)$ e $z_k = \ln(y_k)$, para determinar a_1 e a_2 . Obtidos os coeficientes a_1 e a_2 do polinômio transformado, recuperamos os valores originais de a e b por meio de suas relações:

$$a = e^{a_1}$$

$$b = a_2$$

7.2.2.3 Ajuste a Curvas Hiperbólicas: $y = \frac{I}{a_1 + a_2 * x}$

Transformamos pelo recíproco,

$$\frac{1}{y} = a_1 + a_2 * x$$

Fazendo $1/y = z$, temos $z = a_1 + a_2 * x$, que é um polinômio de grau $n=1$, e efetuamos o ajuste linear ao conjunto de pontos (x_k, z_k) , em que $z_k = 1/y_k$, para determinar diretamente a_1 e a_2 .

7.2.2.4 Ajuste a Sigmóides: $y = \left(1 + e^{-(a_1 + a_2 * x)}\right)^{-1}$

Transformamos por meio da composição do recíproco e da propriedade dos logaritmos:

$$-\ln\left(\frac{1}{y} - 1\right) = a_1 + a_2 * x$$

Fazendo $-\ln\left(\frac{1}{y} - 1\right) = z$, temos $z = a_1 + a_2 * x$, que é um polinômio de grau $n=1$, e promovemos o ajuste linear ao conjunto de pontos (x_k, z_k) , em que $z_k = -\ln(1/y_k - 1)$, para determinar diretamente a_1 e a_2 .

DESTAQUE O ajuste via transformações lineares pode não gerar o mesmo resultado que o ajuste direto da função aproximadora na sua forma original, como comprovaremos no **Exemplo 7.5**, mas a forma transformada para linear é mais rápida e sempre tem solução, enquanto a solução do sistema não linear, gerado no ajuste direto, pode ter convergência difícil e não garantida. FIM DO DESTAQUE

A seguir, vamos apresentar um exemplo de ajuste efetuado diretamente pela minimização do desvio quadrático total e compará-lo com o ajuste efetuado através de transformação paramétrica para a forma polinomial. Também vamos fazer uma introdução à análise dos resultados.

Exemplo 7.5: a tabela, a seguir, com $m=5$ pontos, obtidos experimentalmente, relaciona o volume V adimensional de álcool gerado em um reator em função da sua temperatura adimensional T de reação:

T_k	0.2	0.4	0.6	0.8	1.0
-------	-----	-----	-----	-----	-----

V_k	0.04	0.14	0.26	0.39	0.50
-------	------	------	------	------	------

Considerando teoricamente que a relação entre essas duas variáveis é modelada pela função não polinomial $V(T) = (a_1 + a_2 / T^2)^{-1}$:

- Determine os parâmetros a_1 e a_2 através de ajuste de curvas, por **transformação paramétrica** em polinomial, de modo a levar em conta todas as $m = 5$ medições experimentais;
- Determine os parâmetros a_1 e a_2 através da **minimização direta** do desvio quadrático total, de modo a considerar as $m = 5$ medições experimentais dadas;
- Calcule o coeficiente de determinação R^2 e a média dos resíduos (desvios locais) em cada caso;
- Plote os ajustes em cada caso;
- Plote os gráficos dos resíduos (desvios locais) em cada caso;
- Avalie os resultados.

Solução:

- Transformando $V(T) = \frac{1}{a_1 + a_2 (1/T^2)}$ na forma polinomial, resulta

$$\frac{1}{V(T)} = a_1 + a_2 (1/T^2)$$

$$y = a_1 + a_2 * x$$

$$\text{Com } y_k = \frac{1}{V_k}, x_k = \frac{1}{T_k^2}$$

Tabela original:

T_k	0.2	0.4	0.6	0.8	1.0
V_k	0.04	0.14	0.26	0.39	0.50

Nova tabela com $x_k = \frac{1}{T_k^2}$ e $y_k = \frac{1}{V_k}$:

x_k	25.0	6.25	2.7778	1.5625	1.00
y_k	25.0	7.1429	3.8462	2.5641	2.00

Montando o sistema pela eq. (6) para (x_k, y_k) ,

$$\begin{bmatrix} m & \sum x_k \\ \sum x_k & \sum x_k^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_k \\ \sum x_k^1 y_k \end{bmatrix}$$

$$A = \begin{bmatrix} 5.0000 & 36.5903 & \vdots & 40.5531 \\ 36.5903 & 675.2200 & \vdots & 686.3330 \end{bmatrix}$$

Resolvemos preferencialmente por Cholesky:

$$a_1 = 1.11383182450055$$

$$a_2 = 0.956099722529505$$

b) Minimizando diretamente o desvio quadrático total via eq. (2), temos

$$D(a,b) = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m \left[\frac{1}{a_1 + a_2 (1/T^2)} - V_k \right]^2$$

$$\frac{\partial D(a,b)}{\partial a} = 2 \sum_{k=1}^m \left[(a_1 + a_2 / T^2)^{-1} - V_k \right] \left[(-1) * (a_1 + a_2 / T^2)^{-2} \right] 1 = 0$$

$$\frac{\partial D(a,b)}{\partial b} = 2 \sum_{k=1}^m \left[(a_1 + a_2 / T^2)^{-1} - V_k \right] \left[(-1) (a_1 + a_2 / T^2)^{-2} \right] (1/T_k^2) = 0$$

Resultam estas duas equações não lineares:

$$\begin{cases} f_1(a,b) = - \sum_{k=1}^m \left[(a_1 + a_2 / T^2)^{-1} - V_k \right] (a_1 + a_2 / T^2)^{-2} = 0 \\ f_2(a,b) = - \sum_{k=1}^m \left[(a_1 + a_2 / T^2)^{-1} - V_k \right] (a_1 + a_2 / T^2)^{-2} (1/T_k^2) = 0 \end{cases}$$

Que resolvemos pelo método de Newton:

$$a_1 = 0.991635399512743$$

$$a_2 = 1.00947137630393$$

c) Calculado os respectivos coeficiente de determinação R^2 e as médias dos resíduos (desvios locais), temos

i. por transformação paramétrica em polinomial:

$$R_1^2 = 0.997397323175315$$

$$\text{Média dos resíduos 1} = -34.2280214515672e-04$$

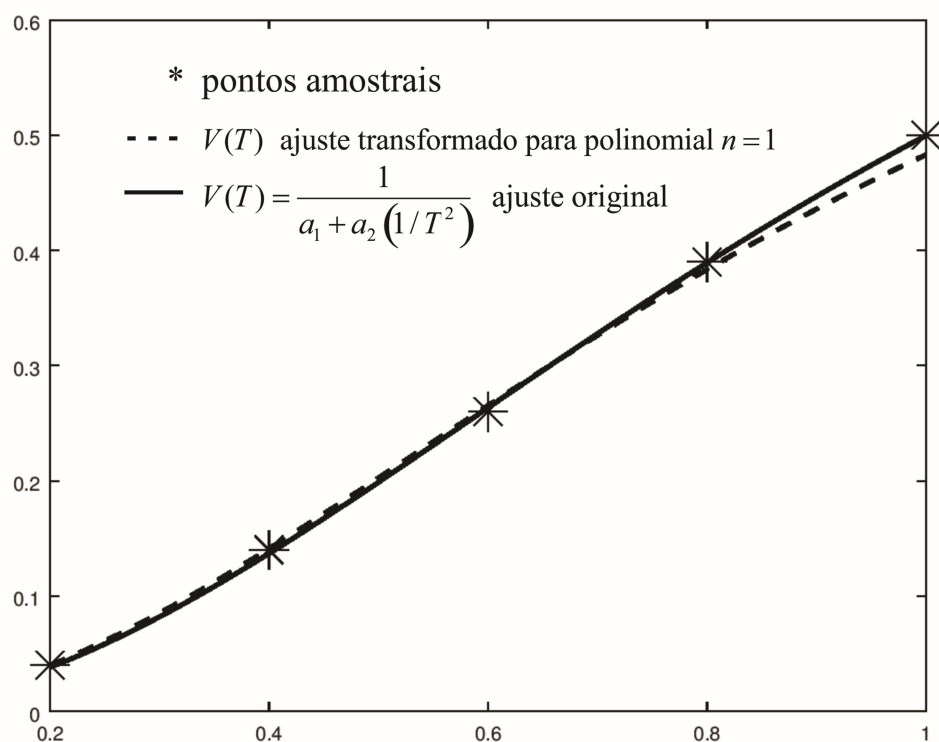
ii. por minimização direta:

$$R_2^2 = 0.999815970365156$$

$$\text{Média resíduos 2} = -4.91670466903520e-04$$

d)

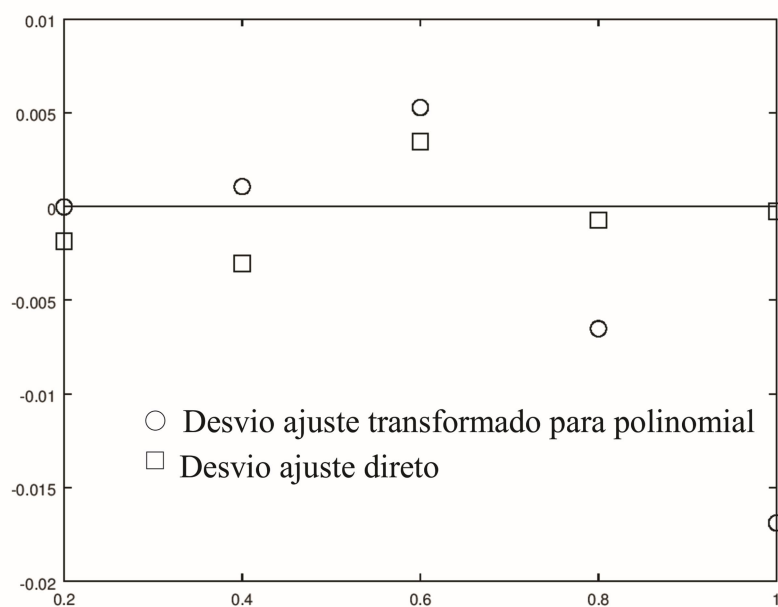
Gráfico 7.8 – Comparativo entre o ajuste transformado em polinomial e o direto do **Exemplo 7.5**



Fonte: Elaboração própria

e)

Gráfico 7.9 – Representação dos resíduos (desvios locais) do **Exemplo 7.5**



Fonte: Elaboração própria

- f) Mesmo com esse número reduzido de amostras, escolhido por motivos didáticos, o ajuste direto gerou um coeficiente de determinação R_2^2 mais próximo da unidade e um resíduo médio mais próximo de zero, em relação ao ajuste parametrizado. Pelo Gráfico 7.8, também verificamos que o ajuste direto, mostrado em linha contínua, representa melhor a tendência dos $m = 5$ pontos experimentais, especialmente no final do intervalo, onde o ajuste parametrizado começa a perder aderência em relação aos pontos tabelados, fato esse comprovado no Gráfico 7.9, onde os resíduos do ajuste direto, representados por quadrados, são menores e mais próximos de uma distribuição aleatória em torno do zero, enquanto os resíduos do ajuste parametrizado, representados por círculos, são maiores e com tendência para valores negativos.

O algoritmo gerador dos resultados do **Exemplo 7.5** está disponível no **Caderno de Algoritmos em Cap7exemplo.7.5.m**.

7.3 Conclusões

Encerramos aqui a nossa abordagem sobre o ajuste de curvas salientando que os tópicos apresentados neste Capítulo são apenas noções básicas sobre o assunto, sempre voltados para implementação em computador. O interessado no aprofundamento desse conteúdo para uso em experimentos reais, com a obtenção do respectivo modelo matemático confiável através do ajuste de curvas, deverá consultar textos sobre a técnica estatística da **análise de regressão**. LINK As obras “Estatística para Cursos de Engenharia e Informática”, de Pedro Alberto Barbetta, Marcelo Menezes Reis e Antonio Cezar Bornia; e “Handbook of Regression Analysis”, de Samprit Chatterjee e Jeffrey S. Simonoff, são referências recomendadas sobre análise de regressão. FIM DO LINK

Sobre a qualidade do ajuste (*Goodness of fit*), recomendamos ainda buscar uma análise mais abrangente dos resultados, para além das medidas citadas neste livro, e usar também técnicas de otimização do ajuste que considerem, por exemplo, o número de parâmetros necessários em cada função ajustadora, como os **Critérios de Informação de Akaike (AIC)**. LINK Para iniciar sua pesquisa sobre AIC, consulte o

endereço eletrônico: <<http://www.portalaction.com.br/analise-de-regressao/2715-aic-e-bic>>. Acesso em: 20 dez. 2016. FIM DO LINK