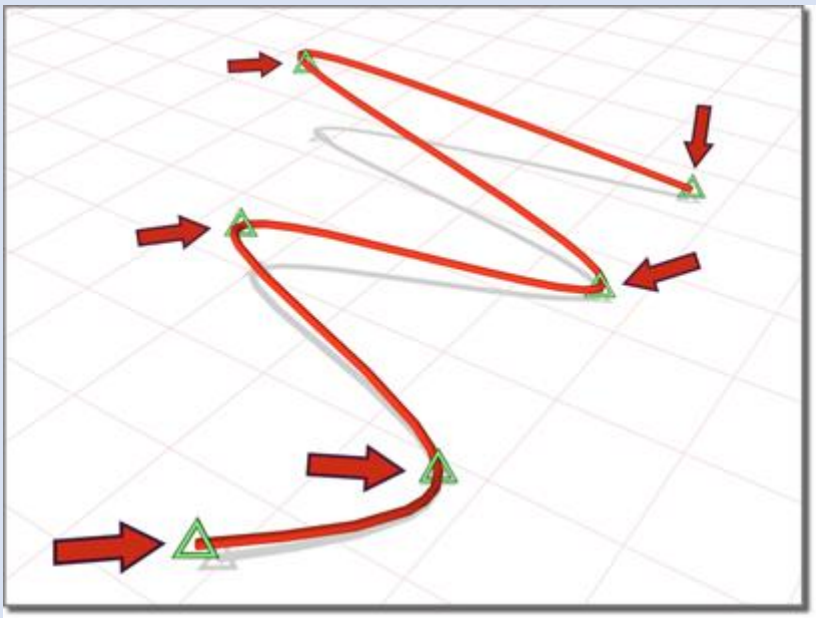
CÁLCULO NUMÉRICO

AJUSTE DE CURVAS



AJUSTE DE CURVAS PELO MÉTODO DOS MÍNIMOS QUADRADOS

INTRODUÇÃO

Nas aulas anteriores estudamos alguns métodos que determinam um polinômio interpolador. Entretanto, existem casos em que o polinômio não se ajusta satisfatoriamente aos valores conhecidos e, assim, devemos utilizar métodos alternativos. E o método mais adequado é o ajuste de curvas, pois consiste em um processo que nos permite encontrar uma função que melhor descreve um certo conjunto de informações experimentais, possibilitando a obtenção de previsões para dados desconhecidos.

Até agora o polinômio de aproximação foi definido de tal maneira a coincidir com o valor da função em pontos definidos na tabela, ou seja, $P(x_k) = f(x_k)$, para k = 0, 1, 2, 3, . . . , n. Já no ajuste de curvas, não existe essa obrigatoriedade da função de ajuste coincidir com os pontos tabelados, isto é, para os pontos tabelados, não é necessário que $P(x_k)$ seja igual a $f(x_k)$. A função de ajuste que será construída a partir dos pontos tabelados, não precisa retornar o mesmo resultado de $f(x)$ tabelado. O gráfico abaixo faz uma comparação entre interpolação e ajuste de curvas.

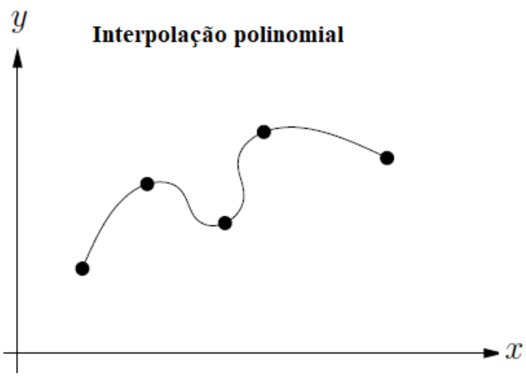


Gráfico 1 – Interpolação

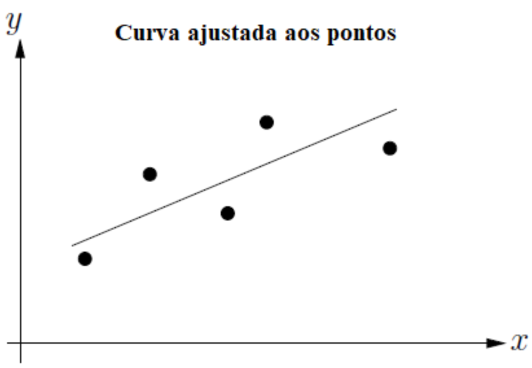


Gráfico 2 – Ajuste de curvas

No Gráfico 1, observe que a curva passa exatamente nos pontos conhecidos de uma função $f(x)$, enquanto que, no Gráfico 2, isto não necessariamente acontece. É pouco provável que haja uma curva que passe exatamente por cada ponto, como é definido na interpolação polinomial, e que descreva fielmente o sistema observado. Por isso é mais viável construir uma função que se aproxime, que se ajuste ao máximo à distribuição dos pontos.

Além disso, a interpolação polinomial se limita a interpolar pontos dentro daquela faixa (ou intervalo) de dados $[a, b]$, não permitindo sair (extrapolar) desse intervalo. Já o ajuste de curvas é muito utilizado para, a partir de dados conhecidos, fazer-se extrapolações, isto é obter um valor para a função $f(x)$ para um dado fora de $[a, b]$.

Assim, o método dos mínimos quadrados é necessário quando:

1- É preciso obter um valor aproximado da função em algum ponto fora do intervalo tabelado, ou seja, quando se quer extrapolar. Essa é uma grande vantagem em relação à interpolação polinomial, onde, na prática, não é aconselhável extrapolar.

2- Os valores tabelados são resultados de algum experimento físico ou de alguma pesquisa. Em geral esses valores poderão conter erros inerentes que, em geral, não são previsíveis. Essa situação é justificada pois ao realizar medidas, os erros são evidentes. Sejam eles devido à precisão do aparelho (menor divisão), defeitos no aparelho (dilatação), interferência de efeitos externos (ação de campos magnéticos, elétricos, temperatura, pressão, etc) bem como a erros cometidos pelo próprio experimentador.

3- Substituição de uma função conhecida por outras funções, como ajustar os pontos distribuídos em uma reta (regressão linear). Em alguns casos, manipular certas funções são tarefas árduas. Neste caso, podemos conhecer a função f(x), mas as manipulações algébricas (derivada, integral, etc) a fim de se chegar em algum resultado são tão difíceis que fica mais conveniente buscar uma função φ(x) pelo ajuste de curvas e trabalhar com ela.

4- O número de pontos que se deseja interpolar, obtidos a partir de um experimento, é muito grande. Por exemplo, se utilizarmos o método de interpolação por sistemas lineares, se desejarmos utilizar 20 pontos tabelados para construir um polinômio, teremos que resolver um sistema linear 20 x 20, o que manualmente se torna inviável.

Dessa forma, para definir uma função analítica que descreva o sistema, não se deve optar por uma forma polinomial interpoladora dos pontos fornecidos, e sim uma curva que melhor se ajusta a esses pontos, levando em consideração a existência de erros que, em geral, não são previsíveis.

Podemos citar abaixo alguns exemplos muito utilizados na prática, onde se deseja extrapolar os pontos conhecidos:

- As empresas que fornecem energia elétrica para uma cidade, conhece os dados de consumo anual da população. A partir destes dados conhecidos, pode-se fazer projeções para o futuro e com isso fazer um planejamento, para que a cidade tenha energia suficiente e adequada nos anos seguintes e não falte energia. Neste caso, deseja-se ajustar uma curva que melhor se ajusta aos dados de consumo disponíveis e assim fazer projeções futuras de consumo.

- Além do consumo de energia elétrica, todos os anos o Brasil sofre com a falta de chuvas, prejudicando os níveis de água dos reservatórios e, consequentemente os cidadãos. Políticas de conscientização por parte do governo são traçadas quando projeções de falta de chuvas são detectadas. Para isso, uma função que relaciona a falta de chuvas (ou o nível de água dos reservatórios) e o consumo pode ser construída e, a partir dela, fazer projeções de falta ou não de água.

- Outro exemplo que se utiliza do ajuste de curvas, muito utilizado pelo Ministério da Educação, são as projeções futuras para a melhoria da educação no Brasil através de indicadores que se utilizam de dados existentes. Por exemplo, o ajuste pode ser aplicado para desenvolver métodos de estimação de demanda demográfica por vagas nas escolas do país.

- O IBGE também se utiliza de ajustes de curvas para fazer projeções de crescimento ou decrescimento de uma população. Muito se pesquisa para saber a populações de jovens daqui a 30 ou 40 anos, ou a população de idosos nos próximos anos para fins de aposentadoria, INSS, etc.

- Outro exemplo da sua aplicação aparece na administração, quando uma empresa pretende relacionar o custo médio com a quantidade de unidades produzidas em um dia.

- Relacionar despesas com a educação e diminuição da violência, relacionar consumo de combustível por um automóvel e adição de aditivos, relacionar crescimento de uma bactéria e temperatura, etc

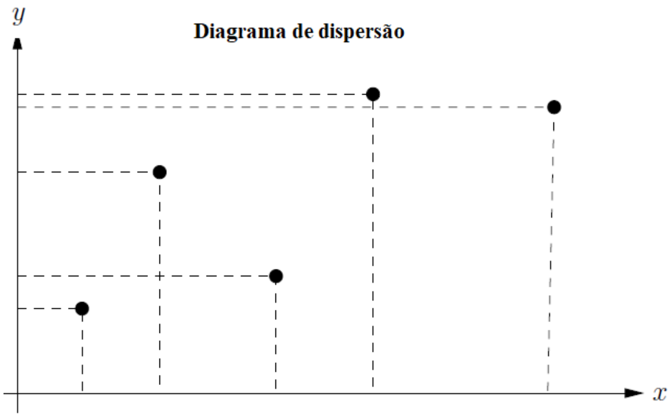
Um outro ganho significativo com o ajuste de curvas é o fato de poder construir outros tipos de funções, além das polinomiais. Em interpolação polinomial, como o próprio nome já diz, construímos apenas polinômios. O ajuste de curvas é usado geralmente para ajustes lineares (polinômio de grau 1: retas), mas em alguns casos pode ser aplicado em outras funções, tais como as funções ***trigonométricas, exponenciais, logarítmicas***, etc.

Como sabemos, existem diversos tipos de funções matemáticas que modelam diversos fenômenos físicos. Por exemplo, o crescimento bacteriano em laboratório fornece um exemplo excelente de crescimento exponencial. No crescimento exponencial, a taxa de crescimento populacional aumenta ao longo do tempo, em proporção ao tamanho da população. Então, se algum modelo matemático for construído para estudar o crescimento de bactérias através do ajuste de curvas, é mais adequado utilizar um ajuste exponencial e não polinomial.

AJUSTE DE CURVAS PELO MÉTODO DOS MÍNIMOS QUADRADOS

Uma das ferramentas matemáticas para encontrar uma curva (função) que melhor se ajusta aos pontos conhecidos é o **Método dos Mínimos Quadrados**. Este método é um dos processos alternativos para determinar a função, não necessariamente polinomial, que se adapte ao conjunto de valores conhecidos.

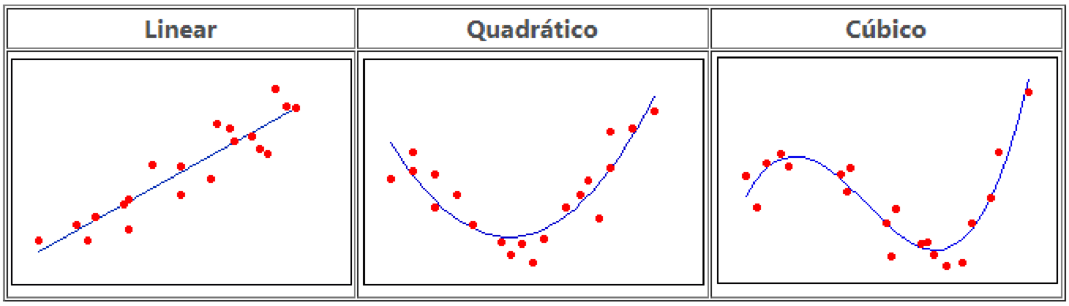
Neste método, é essencial que se plote (marquem) em um sistema de eixos cartesianos, os valores *(x, f(x))* conhecidos. A partir da distribuição dos pontos no gráfico procura-se a função ou combinação de funções que mais se adapte a esta distribuição. Plotando os pontos no plano cartesiano, estes podem revelar uma curva (funções de ajuste) que melhor se adapta ou se ajusta aos pontos tabelados. O gráfico destes pontos é chamado de **diagrama (ou mapa) de dispersão.**



Nestes casos, em geral, são utilizadas as chamadas funções de ajuste de baixa ordem, como por exemplo:

- retas,
- parábolas,
- cúbicas,
- exponenciais,
- trigonométricas, etc..

Abaixo podemos observar quais tipos de curvas se ajustam melhor aos pontos, a partir de um mapa de dispersão. No primeiro quadro é mais adequado um ajuste linear, no segundo um ajuste quadrático e no terceiro um ajuste cúbico.



Ao aplicar o método dos mínimos quadrados devemos levar em conta duas possibilidades:

1- DOMÍNIO DISCRETO - Quando a função referente a distribuição dos pontos não é conhecida. Ou seja, quando a função f é dada por uma tabela de valores.

Neste caso o número de pontos é finito e a função de referência é estabelecida a partir de um modelo matemático suposto para as relações existentes entre as grandezas.

2- DOMÍNIO CONTÍNUO – Procedimento usado quando a função $f(x)$ é conhecida e se deseja uma nova função $\varphi(x)$ que se aproxime suficientemente da função $f(x)$. Neste caso, o número de pontos é então infinito pois podem ser obtidos a partir da própria função.

Vamos estudar apenas o caso discreto.

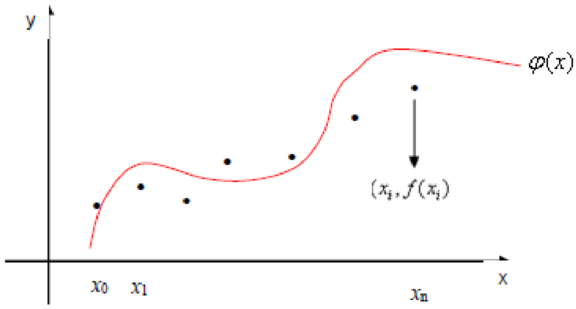
MÉTODO DOS MÍNIMOS QUADRADOS

1- CASO DISCRETO

Neste caso, temos os dados tabelados, com $x_1, x_2, x_3, ..., x_n \in [a, b]$.

x	x_1	x_2	x_3	x_4	......	x_n
f(x)	$f(x_1)$	$f(x_2)$	$f(x_3)$	$f(x_4)$	......	$f(x_n)$

A partir dos dados disponíveis, pode-se construir um *diagrama de dispersão*, que é a representação em gráfico dos dados disponíveis.



Como o objetivo é encontrar uma função $\varphi(x)$ que seja uma boa aproximação para os valores tabelados de $f(x)$ e que nos permita extrapolar com uma certa margem de segurança, devemos escolher n funções $g_1(x), g_2(x), g_3(x), ..., g_n(x)$, contínuas em [$a$. b] e obter n constantes $a_1, a_2, a_3, ..., a_n$ tais que a função

$$\varphi(x) = a_1 . g_1(x) + a_2 . g_2(x) + a_3 . g_3(x) + ... + a_n . g_n(x)$$

se aproxime ao máximo de $f(x)$, ou seja, $f(x) \approx \varphi(x)$.

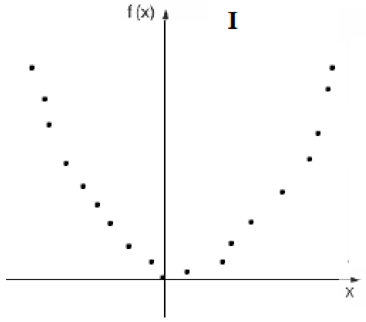
Este modelo matemático é linear pois os coeficientes que devem ser determinados $a_1, a_2, a_3, ..., a_n$ aparecem linearmente, embora as funções $g_1(x), g_2(x), g_3(x), ..., g_n(x)$ possam ser não lineares, como por exemplo $g_1(x) = e^x$ e $g_2(x) = x^2$.

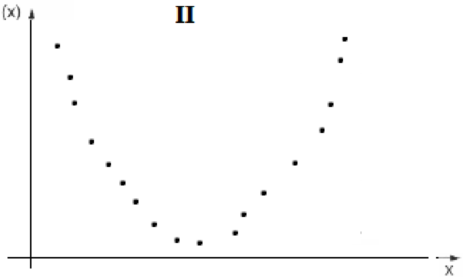
Surge então a primeira pergunta: Como escolher as funções contínuas $g_1(x), g_2(x), g_3(x), ..., g_n(x)$ de modo que $\varphi(x)$ seja uma boa aproximação de f(x)?

Esta pergunta pode ser respondida em dois casos.

CASO 1 - Esta escolha pode ser feita observando o gráfico dos pontos tabelados (diagrama de dispersão).

Ou seja, dada uma tabela de pontos $(x_1, f_1(x)), (x_2, f_2(x)), (x_3, f_3(x)), ..., (x_n, f_n(x))$, deve-se, em primeiro lugar plotar estes pontos num gráfico cartesiano e a partir daí pode-se visualizar a curva que melhor se ajusta aos dados. Por exemplo, os diagramas de dispersão abaixo mostram que esses diagramas se assemelham muito a uma parábola.



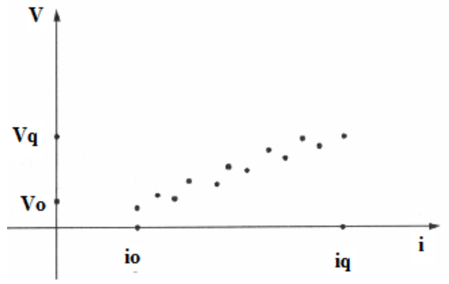


No gráfico **I**, a parábola tem vértice na origem. Portanto, nesse caso, é natural escolhermos apenas uma função $g_1(\mathrm{x})=\mathrm{x}^2$ e procurarmos então $\varphi(x)=a_1.g_1(\mathrm{x})=a_1.\mathrm{x}^2$ (equação geral de uma parábola passando pela origem). Falta apenas descobrir o valor de a_1.

No gráfico **II**, a parábola não tem vértice na origem. Nesse caso, a função $\varphi(x)$ a ser escolhida é uma função quadrática da forma $c+bx+ax^2$. Então, procurarmos uma função $\varphi(x)=a_1.1+a_2.x+a_3.x^2$ **,**onde $g_1(x)=1,\ g_2(x)=x\ \ e\ \ g_3(x)=x^2$. Falta apenas descobrir os valores de $a_1,\ a_2\ \ e\ \ a_3$.

CASO 2 - Baseando-se em fundamentos teóricos do experimento que forneceu a tabela.

Por exemplo, se considerarmos uma experiência onde foram medidos vários valores de corrente elétrica (i) que passa por uma resistência (R) submetida a várias tensões (V), colocando os valores correspondentes de corrente elétrica e tensão em um gráfico podemos ter a figura abaixo:

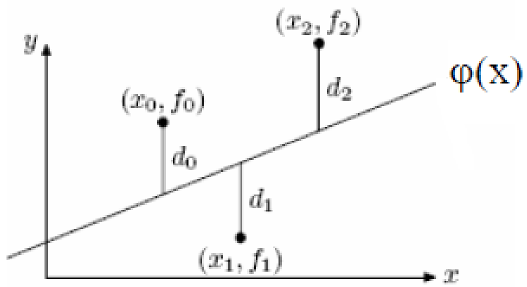


Neste caso, existe uma fundamentação teórica relacionando a corrente com a tensão (V= R i; Lei de Ohm), isto é, V é uma função linear (reta) de i. Assim, $g_1(\mathrm{i})=i\ \ e\ \ \varphi(i)=a_1.i$.

Assim, considerando os dois casos acima, surge uma segunda pergunta:

- qual parábola com equação $\varphi(x)=a_1.\mathrm{x}^2$ melhor se ajusta (ou se aproxima) ao diagrama do **caso 1** e
- qual reta, passando pela origem, melhor se ajusta (ou se aproxima) ao diagrama do **caso 2**?

No caso geral, escolhidas as funções $g_1(\mathrm{x}), g_2(\mathrm{x}), g_3(\mathrm{x}),....,\mathrm{g}_n(x)$ temos que estabelecer o conceito de ***proximidade*** entre as funções $\varphi(x)$ e $f(x)$ para responder as perguntas acima. Uma maneira de estabelecer ***proximidade*** é impor que o desvio $d_k = f(x_k)-\varphi(x_k)$ seja mínimo, para k = 1, 2, ..., n. Ou seja, a distância entre cada ponto da tabela e os pontos da curva $\varphi(x)$ seja a menor possível.



Existem várias formas de impor que os desvios sejam mínimos. O método dos mínimos quadrados é um deles. O ajuste de curvas pelo Método dos Mínimos Quadrados tem por objetivo ajustar $\varphi(x) = f(x)$, de forma que os **desvios quadráticos** sejam mínimos, ou seja, os coeficientes $a_1, a_2, a_3, ..., a_n$ que fazem com que $\varphi(x)$ se aproxime ao máximo de $f(x)$, são os que minimizam os desvios $d_k = f(x_k) - \varphi(x_k)$.

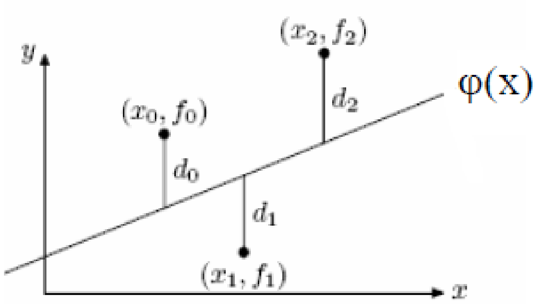
Vamos iniciar com um **ajuste linear**, que é o modelo mais simples que relaciona duas variáveis x e y: ***a reta***. Os casos seguintes seguem o mesmo raciocínio.

AJUSTE LINEAR PELO MÉTODO DOS MÍNIMOS QUADRADOS

Dada uma tabela com ***n** pares* $(x_1, f_1(x)), (x_2, f_2(x)), (x_3, f_3(x)), ..., (x_n, f_n(x))$**,** queremos encontrar a equação da *reta* (y = b + ax) que melhor ajusta a estes pontos. Como o ajuste será feito por uma reta, tomaremos $g_1(x) = 1$ e $g_2(x) = x$ e assim,

$$\varphi(x) = a_1.g_1(x) + a_2.g_2(x) \quad \Rightarrow \quad \varphi(x) = a_1 + a_2.x$$

Assim, pelo método dos quadrados mínimos devemos procurar a_1 e a_2 tal que a soma de todos os desvios $d_k = f(x_k) - \varphi(x_k)$ seja mínima, ou seja, minimizar $\sum_{k=1}^{n} d_k$.



Observe que o valor de cada d_k pode ser positivo ou negativo, e assim, o somatório não seria representativo dos desvios, pois poderíamos ter soma igual a zero. Por exemplo, se as diferenças são:

$$d_1 = 5 \;, d_2 = -3, \quad d_3 = -2, \quad d_4 = 6, \quad d_5 = -3, \quad d_6 = -5, \quad d_7 = 3, \quad d_8 = -1$$

$$\Rightarrow \sum_{1}^{8} d_k = 5 - 3 - 2 + 6 - 3 - 5 + 3 - 1 = 0$$

Uma primeira solução seria utilizar o somatório dos valores absolutos (módulo) de d_k, ou seja $\sum_{k=1}^{m} |d_k|$, já que o módulo de qualquer número é sempre positivo. Entretanto o manuseio de expressões que envolvem valor absoluto é extremamente complexo. A solução mais fácil é a utilização da soma dos desvios ao quadrado (elevar as diferenças ao quadrado), pois todo número ao quadrado é sempre positivo.

Assim, tem-se:

$$D = \sum_{k=1}^{n} d_k^{\,2} = \sum_{k=1}^{n} \left[f(x_k) - \varphi(x_k) \right]^2 \qquad \Rightarrow \qquad D = \sum_{k=1}^{n} d_k^2 = \sum_{k=1}^{n} \left[y_k - \varphi(x_k) \right]^2$$, já que y = f(x).

Substituindo $\varphi(x) = a_1 + a_2 x$ na equação acima, tem-se:

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

Observe que a função D depende de a_1 e a_2. Neste caso, D é uma função de duas variáveis. Do Cálculo II, a condição necessária para que D seja a menor diferença possível (um mínimo) é que as derivadas parciais de D em relação a a_1 e a_2 sejam zero. Essa definição vem do ponto mínimo de uma função, que ocorre sempre no ponto crítico, que é o ponto onde a derivada é igual a zero (C é um ponto crítico se $f'(C)=0$).

Assim, derivando parcialmente em relação a a_1 e a_2 e igualando a zero, tem-se:

$$\frac{\partial D}{\partial a_1} = 2.\sum_{k=1}^{n}\left[y_k - (a_1 + a_2 x_k)\right].1 = 0 \iff \sum_{k=1}^{n}\left[y_k - (a_1 + a_2 x_k)\right] = 0$$

$$\frac{\partial D}{\partial a_2} = 2.\sum_{k=1}^{n}\left[y_k - (a_1 + a_2 x_k)\right].x_x = 0 \iff \sum_{k=1}^{n}\left[y_k - (a_1 + a_2 x_k)\right].x_x = 0$$

Rearranjando as equações, chega-se em:

$$\begin{cases} \sum_{k=1}^{n} y_k - \sum_{k=1}^{n} a_1 - \sum_{k=1}^{n} a_2 x_k = 0 \quad \Rightarrow \quad \sum_{k=1}^{n} y_k - n a_1 - \sum_{k=1}^{n} a_2 x_k = 0 \qquad (I) \\ \sum_{k=1}^{n} x_k y_k - \sum_{k=1}^{n} a_1 x_k - \sum_{k=1}^{n} a_2 x_k^2 = 0 \qquad (II) \end{cases}$$

Isolando as variáveis a_1 e a_2 dos termos constantes, tem-se:

$$\begin{cases} n a_1 = \sum_{k=1}^{n} y_k - \left(\sum_{k=1}^{n} x_k\right) a_2 \quad \Rightarrow \quad a_1 = \dfrac{\sum y_k - \left(\sum_{k=1}^{n} x_k\right) a_2}{n} \qquad (III) \\ \left(\sum_{k=1}^{n} x_k\right) a_1 + \left(\sum_{k=1}^{n} x_k^2\right) a_2 = \sum_{k=1}^{n} x_k y_k \qquad (IV) \end{cases}$$

Substituindo a_1 em **(IV)** e isolando a_2, temos:

$$\left(\sum_{k=1}^{n} x_k\right) \frac{\sum y_k - \left(\sum_{k=1}^{n} x_k\right) a_2}{n} + \left(\sum_{k=1}^{n} x_k^2\right) a_2 = \sum_{k=1}^{n} x_k y_k$$

$$\frac{\sum_{k=1}^{n} x_k . \sum y_k - \sum_{k=1}^{n} x_k . \left(\sum_{k=1}^{n} x_k\right) a_2}{n} + \left(\sum_{k=1}^{n} x_k^2\right) a_2 = \sum_{k=1}^{n} x_k y_k$$

$$\sum_{k=1}^{n} x_k . \sum y_k - \sum_{k=1}^{n} x_k . \left(\sum_{k=1}^{n} x_k\right) a_2 + n.\left(\sum_{k=1}^{n} x_k^2\right) a_2 = n.\sum_{k=1}^{n} x_k y_k$$

$$\sum_{k=1}^{n} x_k . \sum y_k - \left(\sum_{k=1}^{n} x_k\right)^2 a_2 + n.\left(\sum_{k=1}^{n} x_k^2\right) a_2 = n.\sum_{k=1}^{n} x_k y_k$$

$$a_2\left(n.\sum_{k=1}^{n}x_k^2-\left(\sum_{k=1}^{n}x_k\right)^2\right)=n.\sum_{k=1}^{n}x_k y_k-\sum_{k=1}^{n}x_k.\sum_{k=1}^{n}y_k$$

$$a_2=\frac{n.\sum_{k=1}^{n}x_k y_k-\sum_{k=1}^{n}x_k.\sum_{k=1}^{n}y_k}{n.\sum_{k=1}^{n}x_k^2-\left(\sum_{k=1}^{n}x_k\right)^2}$$

Logo, a reta que melhor se ajusta aos pontos do diagrama de dispersão é a reta $\varphi(x)=a_1+a_2.\text{x}$, onde a_1 *e* a_2 são os coeficientes definidos pelas fórmulas acima.

$$a_1=\frac{\sum_{k=1}^{n}y_k-\left(\sum_{k=1}^{n}x_k\right).a_2}{n} \quad e \quad a_2=\frac{n.\sum_{k=1}^{n}y_k x_k-\sum_{k=1}^{n}x_k.\sum_{k=1}^{n}y_k}{n.\sum_{k=1}^{n}x_k^2-\left(\sum_{k=1}^{n}x_k\right)^2}$$

, onde **n** é igual ao número de pontos tabelados

EXEMPLO 1 – Dado a tabela abaixo, faça um ajuste linear que melhor se aproxima dos pontos (x, $f(x)$).

x	14	6	3	6	7	6	7	4	8	7	6	4
f(x)	10	26	41	29	27	27	19	28	19	31	29	33

- Solução-

Para montar a reta $\varphi(x)=a_1+a_2.\text{x}$ temos que encontrar os valores de a_1 *e* a_2, segundo as fórmulas acima.

Para facilitar os cálculos, fica melhor colocar os dados como na tabela abaixo e calcular os termos necessários para calcular a_1 *e* a_2. Observando as fórmulas, precisamos apenas de determinar as somas de *x*.y e *x*². O que precisamos está na última linha, que representa as somas de cada coluna.

x	**y**	**x.y**	**x²**
14	10	140	196
6	26	156	36
3	41	123	9
6	29	174	36
7	27	189	49
6	27	162	36
7	19	133	49
4	28	112	16
8	19	152	64
7	31	217	49
6	29	174	36
4	33	132	16
$\sum x$ = **78**	$\sum y$ = **319**	$\sum x.y$ = **1864**	$\sum x^2$ = **592**

Vamos calcular primeiro a_2 , já que a_1 depende de a_2

$$a_2 = \frac{n\sum_{k=1}^{n} x_k y_k - \sum_{k=1}^{n} x_k . \sum_{k=1}^{n} y_k}{n.\sum_{k=1}^{n} x_k{}^2 - \left(\sum_{k=1}^{n} x_k .\right)^2} = \frac{12\times 1864 - 78\times 319}{12\times 592 - (78)^2} = \frac{22368-24882}{7104-6084} = -2,46$$

Substituindo em a_1, temos:

$$a_1 = \frac{\sum_{k=1}^{n} y_k - \left(\sum_{k=1}^{n} x_k\right) a_2}{n} = \frac{319-(78)(-2,46)}{12} = \frac{319+191,88}{12} = 42,57$$

Assim, a reta que melhor se ajusta aos dados da tabela é:

$$\varphi(x) = 42,57 - 2,46x$$

Observe que, conhecendo a equação da reta que melhor se ajusta aos pontos da tabela, podemos calcular o valor da função *f*(x) para qualquer valor dentro (interpolação) ou fora (extrapolação) da tabela. Por exemplo, podemos querer saber o valor da função *f*(x) quando $x = 5$, que está entre os valores, mas não está na tabela ou quando $x = 20$, que está fora dos valores tabelados. Basta calcular:

$$f(5) = \varphi(5) = 42,57 - 2,46\times 5 = 30,27 \quad e \quad f(20) = \varphi(20) = 42,57 - 2,46\times 20 = -6,63$$

EXEMPLO 2 – Uma empresa está avaliando se existe uma relação entre o investimento anual em propaganda na TV em horário nobre (em milhões de reais) e as vendas (em milhões de reias) da empresa. Para isso foram coletados os dados da tabela abaixo.

GASTO COM PROPAGANDA (X)	2,4	1,6	2,0	2,6	1,4	1,6	2,0	2,2
VENDAS (Y)	75	34	70	90	30	34	36	65

(a) Através do Método de Mínimos Quadrados determine a equação linear $\varphi(x) = a_1 + a_2 .x$ que melhor se ajusta os pontos dados.
(b) Se 3 milhões de reais forem investidos em propaganda anualmente, quanto se espera em vendas?
(c) Para obter vendas de 150 milhões de reais anuais, qual deve ser o valor investido em propaganda?

-Solução-

(a) Construindo a tabela abaixo, temos:

PROPAGANDA (X)	**VENDAS (Y)**	$x.y$	x^2
2,4	75	180	5,76
1,6	34	54,4	2,56
2,0	70	140	4
2,6	90	234	6,76
1,4	30	42	1,96
1,6	34	54,4	2,56
2,0	36	72	4
2,2	65	143	4,84
$\sum x = 15,8$	$\sum y = 434$	$\sum x.y = 919,8$	$\sum x^2 = 32,44$

O que precisamos está na última linha, que representa a soma de cada coluna.

- Calcular a_2.

$$a_2 = \frac{n\sum_{k=1}^{n} x_k y_k - \sum_{k=1}^{n} x_k . \sum_{k=1}^{n} y_k}{n.\sum_{k=1}^{n} x_k{}^2 - \left(\sum_{k=1}^{n} x_k .\right)^2} = \frac{8\times 919,8-(15,8)\times(434)}{8\times 32,44-(15,8)^2} = \frac{7358,4-6857,2}{259,52-249,64} = 50,73$$

Substituindo em a_1, temos:

$$a_1 = \frac{\sum_{k=1}^{n} y_k - \left(\sum_{k=1}^{n} x_k\right) a_2}{n} = \frac{434-(15,8)\times 50,73}{8} = \frac{434-801,53}{8} = -45,94$$

Assim, a reta que melhor se ajusta aos dados da tabela é:

$$\varphi(x) = -45,94 + 50,73x$$

(b) Calculando as vendas, caso o investimento seja de 3 milhões de reais.

$$\varphi(3) = -45,94 + 50,73\times 3 = 106,25 \ \textit{milhões de reais}$$

(c) Para obter 150 milhões em vendas, devemos calcular o valor de x:

150 = -45,94 + 50,73x $\Rightarrow$ *150 + 45,94* = 50,73x $\Rightarrow$ *195,94* = 50,73x

$$x = \frac{195,94}{50,73} \Rightarrow x = 3,86$$

Logo, são necessários, aproximadamente, 3,86 milhões de reais anuais de investimento em propaganda para obter 150 milhões de reais em vendas.

CALCULAR OS COEFICIENTES a_1 *e* a_2 POR SISTEMAS LINEARES

Podemos utilizar sistemas lineares para encontrar os coeficientes a_1 *e* a_2 e construir a reta de ajuste $\varphi(x) = a_1 + a_2.\mathrm{x}$.

Das expressões **(I)** e **(II)** da página 8, rearranjando as equações, tem-se:

$$\begin{cases} \sum_{k=1}^{n} y_k - na_1 - a_2 \sum_{k=1}^{n} x_k = 0 \\ \\ \sum_{k=1}^{n} x_k y_k - a_1 \sum_{k=1}^{n} x_k - a_2 \sum_{k=1}^{n} x_k^2 = 0 \end{cases} \Rightarrow \begin{cases} na_1 + a_2 \sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k \\ \\ a_1 \sum_{k=1}^{n} x_k + a_2 \sum_{k=1}^{n} x_k^2 = \sum_{k=1}^{n} x_k y_k \end{cases}$$

onde obtemos o seguinte sistema linear.

$$\begin{pmatrix} n & \sum_{k=1}^{n} x_k \\ \sum_{k=1}^{n} x_k & \sum_{k=1}^{n} x_k^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} y_k \\ \sum_{k=1}^{n} x_k y_k \end{pmatrix}$$

No exemplo 2 anterior,

$$\begin{pmatrix} 8 & 15{,}8 \\ 15{,}8 & 32{,}44 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 434 \\ 919{,}8 \end{pmatrix} \Rightarrow \begin{cases} 8a_1 + 15{,}8a_2 = 434 \\ 15{,}8a_1 + 32{,}44a_2 = 919{,}8 \end{cases}$$

Resolvendo o sistema, encontramos $a_1 = -45{,}94$ e $a_2 = 50{,}73$ e $\varphi(x) = -45,94 + 50,73x$.

QUALIDADE DO AJUSTE LINEAR

- **COEFICIENTE DE DETERMINAÇÃO**

A qualidade de um ajuste linear pode ser verificada em função do coeficiente de determinação R^2, dado por:

$$R^2 = \frac{\left[\sum x.y - \frac{\sum x.\sum y}{n}\right]^2}{\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right] \times \left[\sum y^2 - \frac{\left(\sum y\right)^2}{n}\right]}$$

sendo $0 \le R^2 \le 1$. Quanto mais próximo de 1, melhor o ajuste. Pode ser expresso em porcentagem também.

O coeficiente de determinação indica a porcentagem de variação nos dados de y que podem ser atribuídos à variação nos dados de ***x***, ou seja, indica a porcentagem dos valores de y que vão variar quando os valores de ***x*** também variarem.

No exemplo acima, podemos dizer que 83,3% dos dados de *y* vão variar quando os valores de ***x*** variarem.

$$R^2 = \frac{\left[919{,}8 - \frac{15{,}8 \times 434}{8}\right]^2}{\left[32{,}44 - \frac{(15{,}8)^2}{8}\right] \times \left[27358 - \frac{(434)^2}{8}\right]} = \frac{3925{,}02}{1{,}235 \times 3813{,}5} = 0{,}833$$

Isto quer dizer que 83,3% da variação das vendas é devida à variação dos gastos em propaganda pela empresa ou, de outra forma, 83,3% das vendas é devido ao investimento em propaganda. Para os outros 16,7% (100% - 83,3%) das variações de *y (vendas),* não podemos afirmar nada, são inexplicáveis e podem ser resultados de erros amostrais ou outros fatores que influenciam nas vendas.

É pouco comum que tenhamos uma correlação perfeita (R^2 =1) na prática, porque existem muitos fatores que determinam as relações entre variáveis na vida real. No caso do exemplo 2, sabemos que não é só propaganda que faz com que um cliente compre determinado produto. Fatores como, novas tendências, papel social da empresa, valor do produto, opinião de outros consumidores, etc também influenciam nas vendas.

Não existe uma escala absoluta para analisar o R^2. Os limites para os quais se considera o ajuste "bom" dependem muito do tipo de dados e situação em que se aplica o ajuste linear. Por exemplo, em Engenharia Ambiental, para equações de volume de árvores individuais em florestas plantadas espera-se R^2 superiores a 0.90. Já para modelos de produção de florestas plantadas é comum se considerar bons valores de R^2 acima de 0.70 ou 0.60.

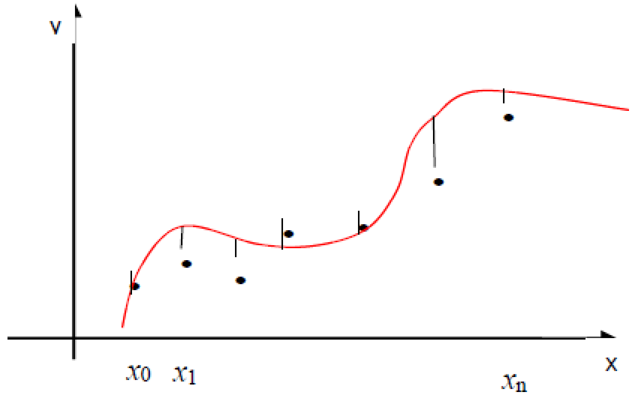
Em algumas áreas, é totalmente esperado que seus valores de R^2 sejam baixos. Por exemplo, qualquer área que tente prever o comportamento humano (como a psicologia) normalmente tem valores de R^2 menores que 50%. Seres humanos são mais difíceis de prever do que, digamos, processos físicos.

Existem outras formas de se avaliar um modelo de ajuste linear, como o coeficiente de correlação e o teste F da estatística, mas não será tratado aqui. Ficaremos apenas com o coeficiente de determinação.

AJUSTE POLINOMIAL

Neste caso, o ajuste dos **n** pontos (x, f(x)) é uma equação de grau n, n > 1, definida por

$$\varphi(x) = a_1.g_1(x) + a_2.g_2(x) + a_3.g_3(x) + \ldots. + a_n.g_n(x)$$



Resolvendo $D = \sum_{k=1}^{n} d_k{}^2 = \sum_{k=1}^{n} \left[f(x_k) - \varphi(x_k) \right]^2$ como na página 8, mas para

$$\varphi(x) = a_1.g_1(x) + a_2.g_2(x) + a_3.g_3(x) + \ldots. + a_n.g_n(x),$$

um polinômio de grau n, n > 1, obtém-se o seguinte sistema linear, que pode ser resolvido pelo método de eliminação de Gauss.

$$\begin{pmatrix} n & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+1} \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \sum x_i^{n+2} & \cdots & \sum x_i^{n+n} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^n y_i \end{pmatrix}$$

EXEMPLO 1 - Ajustar os pontos da tabela a uma expressão quadrática do tipo $y = a_1 + a_2 x + a_3 x^2$

x_i	-2.0	-1.5	0	1	2.2	3.1
y_i	-30.5	-20.2	-3.3	8.9	16.8	21.4

- Solução-

Temos que montar a matriz

$$\begin{pmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 . y_i \end{pmatrix}$$

Calculando os somatórios em uma tabela, temos:

x	y	x^2	x^3	x^4	$x.y$	$x^2.y$
-2	-30,5	4	-8	16	61	-122
-1,5	-20,2	2,25	-3,375	5,0625	30,3	-45,45
0	-3,3	0	0	0	0	0
1	8,9	1	1	1	8,9	8,9
2,2	16,8	4,84	10,648	23,4256	36,96	81,312
3,1	21,4	9,61	29,791	92,3521	66,34	205,654
2,8	**-6,9**	**21,7**	**30,064**	**137,84**	**203,84**	**128,42**

Montando o sistema linear:

$$\begin{pmatrix} 6 & 2{,}8 & 21{,}7 \\ 2{,}8 & 21{,}7 & 30{,}064 \\ 21{,}7 & 30{,}064 & 137{,}84 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -6{,}9 \\ 203{,}5 \\ 128{,}42 \end{pmatrix}$$

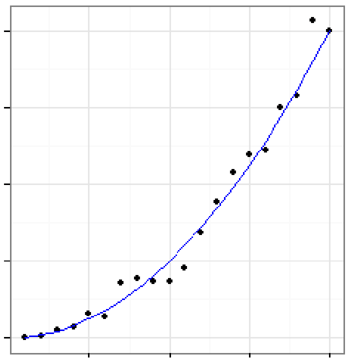
Resolvendo o sistema, encontramos:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{bmatrix} -2.018 \\ 11.332 \\ -1.222 \end{bmatrix}$$

Logo, a expressão quadrática que melhor se ajusta aos pontos da tabela é $y = -2{,}018 + 11{,}332x - 1{,}222x^2$.

AJUSTE EXPONENCIAL

Em alguns casos, a família de funções gi(x) pode ser não-linear nos parâmetros, como, por exemplo, se o diagrama de dispersão de uma determinada função se ajustar a uma **exponencial**.



Vamos trabalhar a seguinte função exponencial:

$$y = f(x) = \alpha_1 . e^{-\alpha_2 . x}$$, com α_1 e α_2 positivos.

O objetivo então consiste em encontrar os valores de α_1 e α_2 para montar a função y = f(x). O método utilizado para encontrar esses dois parâmetros é a aplicação do método dos mínimos quadrados. Mas para aplica-lo deve se efetuar uma linearização da função y = f(x), pois os coeficientes do método dos mínimos quadrados $a_1, a_2, a_3, a_4,$ ...... são lineares.

A linearização é obtida, aplicando a função ***ln*** em ambos os lados de y = f(x). Vejamos:

$$y=\alpha_1.e^{-\alpha_2.x} \quad \Rightarrow \quad \ln(y)=\ln\left(\alpha_1.e^{-\alpha_2.x}\right)$$

Aplicando propriedades de logaritmos, temos:

$$\ln(y)=\ln(\alpha_1)+\ln\left(e^{-\alpha_2.x}\right)= \ln(\alpha_1)-\alpha_2 x.\ln(e)$$
$$\ln(y)= \ln(\alpha_1)-\alpha_2 x$$

Logo, a função y foi linearizada. Neste caso, podemos comparar esta função ln(y) com a função linear de ajuste $\varphi(x)=a_1+a_2.x$, já que ln(y) é linear nos seus parâmetros $\ln(\alpha_1)$ e $-\alpha_2$.

Portanto, a função exponencial *y* definida acima pode ser obtida com as seguintes relações:

$$(1)\ \ln(y)=\varphi(x)$$
$$(2)\ a_1=\ln(\alpha_1)$$
$$(3)\ a_2=-\alpha_2$$

O método dos mínimos quadrados pode então ser aplicado na resolução do problema linearizado. Obtidos os parâmetros a_1 *e* a_2 deste problema, usa-se estes valores para calcular os parâmetros originais α_1 *e* α_2.

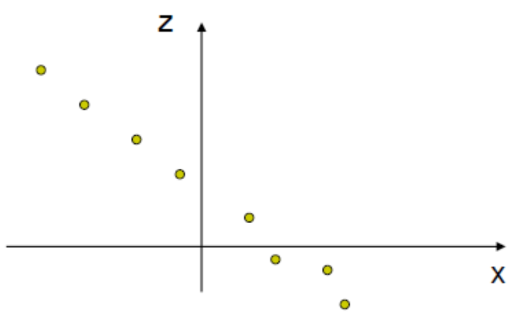
EXEMPLO 1 - Suponha que num laboratório obtivemos experimentalmente os seguintes valores para f(x) sobre os pontos da tabela abaixo.

(a) Faça o diagrama da dispersão dos dados.
(b) Ajuste os dados, usando o Método dos Quadrados Mínimos, para a função $y=\alpha_1.e^{-\alpha_2.x}$;

x_i	−1.0	−0.7	−0.4	−0.1	0.2	0.5	0.8	1.0
$y=f(x_i)$	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246

- Solução-

(a) Faça o diagrama da dispersão dos dados.



(b) Ajuste os dados, usando o Método dos Quadrados Mínimos, para a função $y=\alpha_1.e^{-\alpha_2.x}$.

Queremos construir $y=\alpha_1.e^{-\alpha_2.x}$. Para isso temos que encontrar os valores de α_1 *e* α_2 através das relações obtidas anteriormente.

$$(1)\ \ln(y)=\varphi(x)$$
$$(2)\ a_1=\ln(\alpha_1)$$
$$(3)\ a_2=-\alpha_2$$

Primeiramente temos que calcular $\ln(y)$, já que $\ln(y)=\varphi(x)$. Os resultados estão na terceira linha da tabela abaixo.

X	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
$y=f(x_i)$	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246
$\ln(y)$	3.599	2.849	2.099	1.349	0.599	-0.151	-0.901	-1.402

Com a linha ***x*** e com a linha ln(y) acima, construímos a função de ajuste $\varphi(x)=a_1+a_2.\mathrm{x}$.

x	**z = ln(y)**	$X.Z$	X^2
-1,0	3,599	-3,599	1
-0,7	2,849	-2,0258	0,49
-0,4	2,099	-0,8396	0,16
-0,1	1,349	-0,1349	0,01
0,2	0,599	0,1198	0,04
0,5	-0,151	-0,0755	0,25
0,8	-0,901	-0,7208	0,64
1,0	-1,402	-1,402	1
$\sum x = 0{,}3$	$\sum z = 8{,}041$	$\sum x.z = -8{,}6778$	$\sum x^2 = 3{,}59$

O que precisamos está na última linha, que representa a soma de cada coluna.

- Calcular a_2 e a_1.

$$a_2 = \frac{n.\sum x.z - \sum x.\sum z}{n..\sum x^2 - \left(\sum x\right)^2} = \frac{8.(-8{,}6778) - 0{,}3 \times 8{,}041.}{8 \times 3{,}59 - (0{,}3)^2} = -2{,}5$$

$$a_1 = \frac{\sum z - \left(\sum x\right).a_2}{n} = \frac{8{,}041 - 0{,}3 \times (-2{,}5).}{8} = 1{,}099$$

Assim, como $a_1 = \ln(\alpha_1)$ e $a_2 = -\alpha_2$, podemos calcular α_1 e α_2.

$$a_1 = \ln(\alpha_1) \Rightarrow 1{,}099 = \ln(\alpha_1) \Rightarrow e^{1{,}099} = e^{\ln(\alpha_1)} \Rightarrow 3{,}001 = \alpha_1.\ln(e) = \alpha_1 \Rightarrow \alpha_1 = 3{,}001$$

$$a_2 = -\alpha_2 \Rightarrow -2{,}5 = -\alpha_2 \Rightarrow \alpha_2 = 2{,}5$$

Portanto, a curva que melhor se ajusta aos pontos da tabela é dada pela curva $y = 3{,}001.e^{-2{,}5.x}$.

OUTROS TIPOS DE LINEARIZAÇÃO

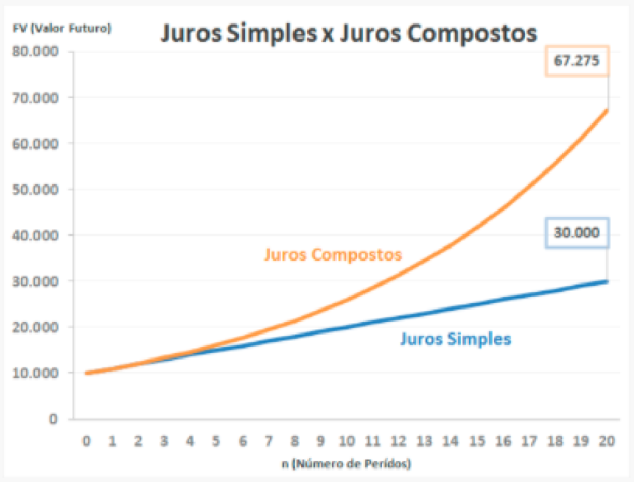
- $y = a\,e^{bx} \longrightarrow \ln y = \ln a + bx$
- $y = a\,b^{x} \longrightarrow \ln y = \ln a + (\ln b)x$
- $y = a\,x^{b} \longrightarrow \ln y = \ln a + b(\ln x)$
- $y = e^{a + bx_1 + cx_2} \longrightarrow \ln y = a + bx_1 + cx_2$
- $y = a x_1^{\,b} x_2^{\,c} x_3^{\,d} \longrightarrow \ln y = \ln a + b\ln x_1 + c\ln x_2 + d\ln x_3$
- $y = \dfrac{1}{a + bx_1 + cx_2} \longrightarrow \dfrac{1}{y} = a + b\,x_1 + cx_2$
- $y = \dfrac{1}{1 + e^{a + bx_1 + cx_2}} \longrightarrow \ln(\dfrac{1}{y} - 1) = a + b\,x_1 + cx_2$

OBSERVAÇÕES SOBRE FUNÇÕES EXPONENCIAIS

A função exponencial é caracterizada pelo crescimento ou decrescimento muito rápido, e por isso é muito utilizada na Matemática e em outras ciências correlacionadas com cálculos, como Química, Biologia, Física, Engenharia, Astronomia, Economia, Geografia, entre outras.

As exponenciais possuem a característica de expressar acentuadas variações em períodos curtos, em razão da presença da incógnita no expoente da expressão. Em razão dessa propriedade, a função exponencial é considerada uma importante ferramenta da Matemática, abrangendo diversas situações cotidianas e contribuindo de forma satisfatória na obtenção de resultados que exigem uma análise quantitativa e qualitativa (Por Marco Noé Pedro da Silva – Mundo Educação). Veja alguns exemplos.

- Na Matemática Financeira, serve para demonstrar o crescimento de um capital aplicado a uma determinada taxa de juros compostos.

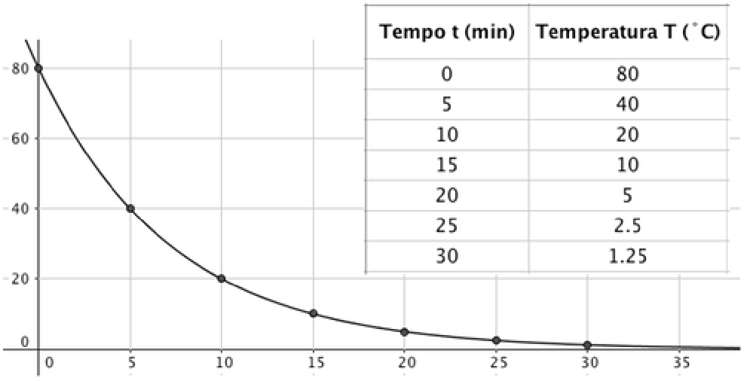


função exponencial dos juros simples X juros compostos

Fonte: Gênio da Matemática

- Na Física, a Lei de Newton para o resfriamento de corpos, que mostra o que acontece quando um corpo é deixado resfriar espontaneamente em um ambiente cuja temperatura possa ser considerada constante.

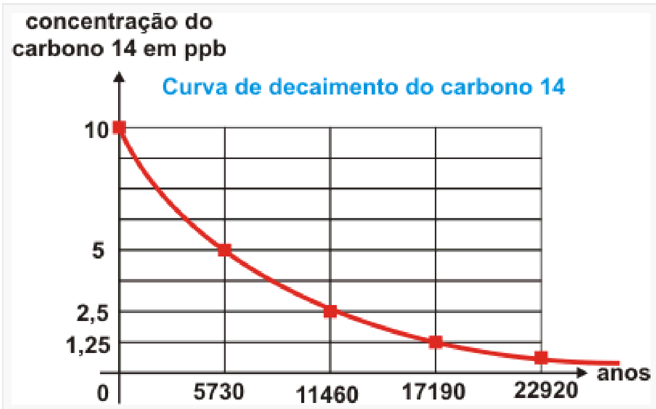
  Se observarmos o resfriamento de um certo volume de água a 80°C em um ambiente a 0°C, por exemplo, notaremos um comportamento mostrado na tabela e no gráfico de uma função exponencial.



Tempo t (min)	Temperatura T (°C)
0	80
5	40
10	20
15	10
20	5
25	2.5
30	1.25

Gráfico: Prof. Marcos A. Simões

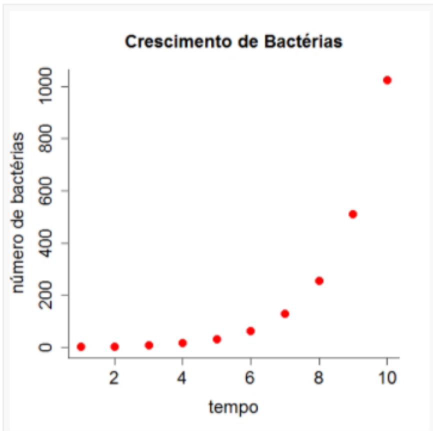
- Na Química está diretamente ligada ao decaimento radioativo.



Decaimento radioativo do carbono 14

Fonte: Gênio da Matemática

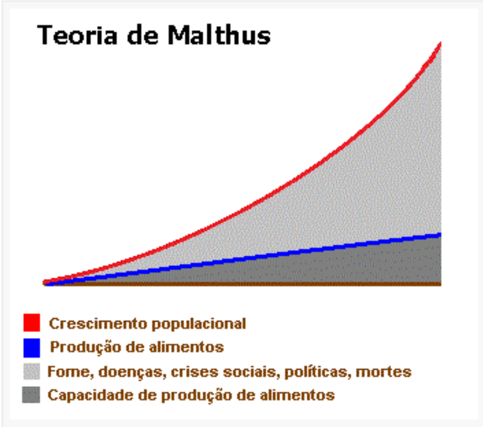
- Na Biologia se apresenta em situações envolvendo o crescimento de bactérias em uma colônia, o crescimento de uma planta, crescimento ou diminuição da população de uma cidade, etc.



Função exponencial com o número de bactérias

Fonte: Gênio da Matemática

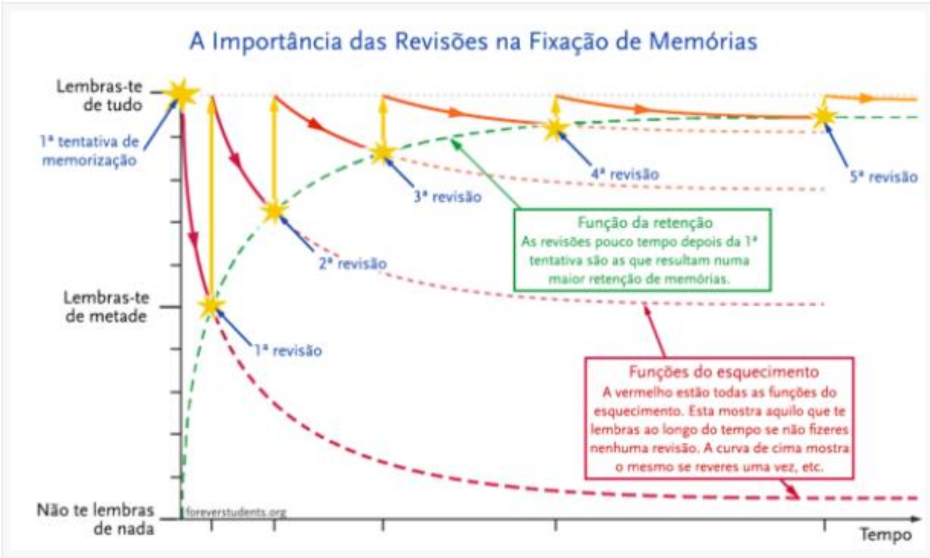
- Usada também na Geografia no intuito de determinar o crescimento populacional.



Função Exponencial de Malthus

Fonte: Gênio da Matemática

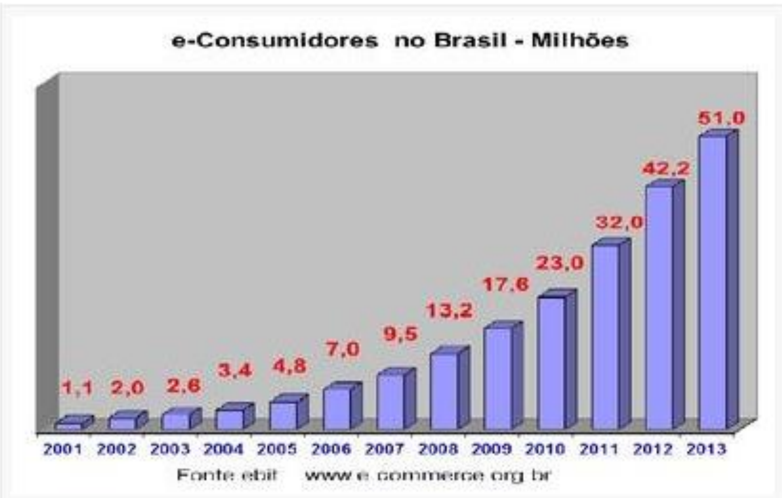
- Na Psicologia e Educação, Psicólogos e Educadores utilizam-se da exponencial a fim de demonstrarem as curvas de aprendizagem.



Curva de esquecimento X número de revisões

Fonte: Gênio da Matemática

- Na Economia, temos um gráfico que mostra o comportamento do tipo de compra e dos valores investidos em e-commerce no Brasil.



Número de pessoas que compraram na internet

Fonte: Gênio da Matemática

EXERCÍCIOS

1- Octanagem é a medida de resistência do combustível à pressão que ele sofre dentro da câmara de combustão do motor. Ou seja, é a capacidade que ele tem de resistir, em mistura com o ar, ao aumento de pressão e de temperatura sem detonar (isso sem que a faísca de vela tenha sido disparada pelo sistema de ignição). Quanto maior a octanagem, maior será a resistência do combustível à detonação. Assim, com maior octanagem é possível que os motores operem com maiores taxas de compressão.

Considere um experimento em que se analisa a octanagem da gasolina (Y) em função da adição de um aditivo (X). Para isto, foram realizados ensaios com os percentuais de 1, 2, 3, 4, 5 e 6% de aditivo. Os resultados seguem na tabela abaixo:

QUANTIDADE % DE ADITIVO (X)	OCTANAGEM DA GASOLINA (Y)
1	80,5
2	81,6
3	82,1
4	83,7
5	83,9
6	85

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x) = a_1 + a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre a adição de aditivos e a octanagem.
(b) Se houver uma adição de 8% de aditivos, qual será a octanagem?
(c) Qual a quantidade de aditivos adicionados para uma octanagem de 85?

Resposta: (a) Y = 0,886.x + 79,7 **(b)** 86,79 **(c)** 5,98

2- Uma construtora dispõe dos seguintes dados: a renda média familiar da população de várias cidades e a quantidade de imóveis vendidos em cada cidade.

Cidade	Renda média (em R$ 1000) (X)	Vendas de Imóveis (Y)
A	5	27
B	10	46
C	20	73
D	8	40
E	4	30
F	6	28
G	12	46
H	15	59

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x) = a_1 + a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre a renda e a venda de imóveis.

Com essa equação da reta, é possível fazer a projeção de valores para outras cidades, verificando a possibilidade de estabelecer novas vendas de imóveis nelas.

(b) Se houver uma cidade com renda 18 (mil reais), qual será a venda esperada?
(c) Qual a renda esperada de uma cidade que tenha 50 imóveis vendidos?

Resposta: (a) Y = 2,9.x + 14,58 **(b)** 66,78 imóveis **(c) R$** 12200,00

3- Uma obra está sendo construída em uma área de 1000 m². Pensando em aumentar a área construída, em no mínimo mais 10 m², o dono da obra deseja verificar qual o custo a mais para aumentar 1 m² de construção, até 15 m². A tabela abaixo mostra a quantidade área a mais construída (em m²) e o custo (em milhares de reais).

Quantidade (x)	10	11	12	13	14	15
Custos (y)	100	112	119	130	139	142

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x)=a_1+a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre a área a mais construída e custo que esse aumento acarreta.

(b) Se ele decidir aumentar em 20 m², qual será o custo esperado?
(c) Quanto ele deve aumentar a área construída, se o custo for de 160 (em milhares de reais)?

Resposta: (a) Y = 8,63.*x* + 15,79 **(b)** 188,39 (milhares de reis a mais) **(c) R$** 16,71 m² a mais.

4-Uma agência da ONU pretende verificar se existe uma relação linear entre a despesa per capita em educação e o grau de criminalidade (assassinatos por mês) numa lista de cinco países.

Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x)=a_1+a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre investimento em educação e taxa de criminalidade.

PAIS	DESPESAS (X) (Milhões)	ASSASSINATOS/MES (Y)
A	5	18
B	8	12
C	12	8
D	13	4
E	15	5

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x)=a_1+a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre despesas em educação e taxa de criminalidade.

(b) Se 10 (Unidades de milhão) forem gastos, quanto se espera de taxa de criminalidade?

Resposta: (a) $\varphi(x)$ = 23,9 - 1,37x **(b)** 10,2 assassinatos por mês

5- Sabendo que a intensidade de campo elétrico no ar, de um ponto em relação a uma carga puntiforme de 650 Coulomb, varia com distância em centímetros de acordo com a tabela abaixo:

Distância em cm (x)	Intensidade de Campo elétrico (y)
5	26
7,5	11,56
10	6,5
12,5	4,16
15	2,88

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x) = a_1 + a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre distância e intensidade de carga elétrica.

(b) Calcule a intensidade de campo elétrico em um ponto situado a 8,5 cm da carga.

Resposta: (a) $\varphi(x)$ = 31,67 - 2,14x **(b)** 13,48 Coulombs

6- Determina-se empiricamente o alongamento de uma mola em milímetros, em função da carga P kgf que sobre ela atua, obtendo-se.

Distância em cm (x)	**Carga (y)**
5	49
10	105
15	172
20	253
25	352
30	473
35	619
40	793

(a) Através do Método de Mínimos Quadrados determine a equação da reta $\varphi(x) = a_1 + a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre alongamento da mola e carga.

(b) Encontrem a carga que produze um alongamento de 12 mm.

Resposta: (a) $\varphi(x)$ = -118,36 + 20,9x **(b)** 132,44 Kgf

7- Após uma regulagem eletrônica um veículo apresenta um rendimento ideal no que tange a consumo de combustível. Contudo, com o passar do tempo esse rendimento vai se degradando. Os dados a seguir representam o rendimento medido mês a mês após a regulagem. Ajuste um modelo linear a esses dados.

Meses após a regulagem (x)	**Rendimento (y)**
1	10,7
2	10,9
3	10,8
4	9,3
5	9,5
6	10,4

Determine a equação da reta $\varphi(x) = a_1 + a_2.x$ que melhor se ajusta os pontos dados, onde seja possível verificar a relação entre tempo após regulagem e rendimento.

Resposta: $\varphi(x)$ = 11,1 - 0,24x

8- Seja a tabela abaixo contendo o tempo de germinação de sementes (dias) em função da temperatura média do solo (ºC) para doze locais de plantio:

Temperatura (ºC)	**Germinação (Dias)**
14	10
6	26
3	41
6	29
7	27
6	27
7	19
4	28
8	19
7	31
6	29
4	33

Determinar a relação entre a temperatura e o tempo de germinação com a relação não linear $y = a.b^x$.

Resposta: $y = (55{,}07).(0{,}887)^x$

9- Ajustar os pontos da tabela abaixo à equação y = $y = a.b^x$

x_i	0.1	1.5	3.3	4.5	5.0
y_i	5.9	8.8	12.0	19.8	21.5

Resposta: $y = 5.6633.e^{0.2646x}$

10-O número de bactérias, por unidade de volume, existente em uma cultura após x horas é dado na tabela abaixo:

número de horas	0	1	2	3	4	5	6
número de bactérias	32	47	65	92	132	190	275

(a) Ajuste os dados acima a curva $y = ae^{bx}$ pelo método dos mínimos quadrados.

(b) Quantas horas seriam necessárias para que o número de bactérias por unidade de volume ultrapasse 2000?

Resposta: (a) $y = 32{,}104.e^{0{,}355x}$ (b) 11,64 horas