

MODELAGEM MATEMÁTICA NO ENSINO DE QUÍMICA: UMA ABORDAGEM PRÁTICA COM REGRESSÃO NO MICROSOFT EXCEL

Joacy B. Lima¹; Cicero W. B. Bezerra¹; Sergiane J. R. Mendonça¹; Paulo S. S. Bezerra¹; Gilvan O. C. Dias¹; Joselene R. J. Santos¹; Janyeid K. C. Sousa²; Victor L. F. Pinheiro³; Walinson F. Martins³; Leonardo B. Cantanhede⁴

¹Departamento de Química, Universidade Federal do Maranhão, São Luís - MA, Brasil

²Coordenação do Curso de Ciência e Tecnologia, Universidade Federal do Maranhão, São Luís - MA, Brasil

³Instituto Federal de Educação, Ciência e Tecnologia do Maranhão, 65030-005 São Luís - MA, Brasil

⁴Instituto Federal de Educação, Ciência e Tecnologia do Maranhão, 65400-000 Codó - MA, Brasil

joacy.lima@ufma.br

Palavras-Chave: linha de tendência, análise de dados, análise estatística

Introdução

A Química, ao investigar a estrutura, a reatividade e as transformações da matéria, lida com sistemas complexos que exigem estratégias quantitativas robustas para sua interpretação. Nesse contexto, as análises de regressão (linear e não linear) constituem ferramentas obrigatórias para modelar relações entre variáveis experimentais, permitindo extrair parâmetros relevantes e compreender os mecanismos que regem os fenômenos observados (OLSEN, 2016; MORAN, 2011; BENNINGA, 2010). Tais regressões consistem em técnicas estatísticas aplicadas à modelagem da dependência entre uma variável resposta (y) e uma variável explicativa (x), de forma linear e não linear, em relação aos parâmetros “ θ ”, como ilustrada na Eq. 1, em que ε_i representa o erro entre os valores experimental e o previsto pelo modelo. A escolha adequada da forma funcional f e o ajuste eficiente dos parâmetros são decisivos para a interpretação confiável de dados (MAZUCHELI & ACHCAR, 2002; MATTOS, 2013; SILVA, E. M., et al, 2019; CHIACCHIO, E. J, 1993; SEBER, G. A. F., 1977; DRAPER, N. R., SMITH, H., 1981; BATES, D. M., WATTS, D. G, 1988).

$$y_i = f(x_i; \theta) + \varepsilon_i \quad i = 1, \dots, n \quad \text{Eq. 1}$$

O método dos mínimos quadrados é o método de ajuste mais comum utilizado nas regressões (CHIACCHIO, E. J, 1993; BENNINGA, S., 2010) e busca minimizar a soma dos quadrados dos resíduos (SQR, Eq. 2).

$$\text{SQR}(\theta) = \sum [y_i - f(x_i, \theta)]^2 = \sum [y_i - \hat{y}_i]^2 = \sum \varepsilon_i^2 = \varepsilon_i^T * \varepsilon_i \quad \text{Eq. 2}$$

As regressões não lineares utilizam algoritmos iterativos: Newton-Raphson, Quasi-Newton, Gauss-Newton, Levenberg-Marquardt, Gradiente descendente, Nelder-Mead (Simplex), etc., sendo o algoritmo de Gauss-Newton o mais simples e mais utilizado em regressões não lineares, equação 3, (LAI, W. H.; KEK, S. L.; TAY, K. G., 2017; SIREGAR, R. W. et al. 2018).

$$\theta_k^{(u+1)} = \theta_k^{(u)} + [J^{T(u)} * J^{(u)}]^{-1} * \{J^{T(u)} * [y_i - f(x_i, \theta_k^{(u)})]\} \quad \text{Eq. 3}$$

Onde: “ J ” é a matriz Jacobiana ou matriz das derivadas da equação modelo em relação aos parâmetros (θ_k), “ T ” é a transposta da matriz Jacobiana e “ u ” o índice de iterações.

O Microsoft Excel é um programa de planilha eletrônica com diversas ferramentas amplamente utilizadas para organização, análise e visualização de dados, oferecendo recursos matemáticos, estatísticos e gráficos que facilitam a interpretação de tendências e relações entre variáveis. Na Química, destaca-se por permitir o tratamento ágil de grandes conjuntos de dados e a identificação de padrões ou anomalias experimentais. Desenvolvido pela Microsoft, tornou-se o padrão em ambientes acadêmicos, corporativos e industriais, consolidando-se como uma das planilhas eletrônicas mais utilizadas no mundo (OLSEN, 2016; WALKENBACH, 2019).

A maioria dos modelos empregados na análise de dados químicos baseia-se em regressões que utilizam linhas de tendência (lineares, exponenciais, logarítmicas, polinomiais ou de potência) ajustadas aos dados experimentais por meio do método dos mínimos quadrados. O Excel oferece suporte direto a esses modelos e permite determinar parâmetros com rapidez e eficiência.

Além disso, dispõe de recursos avançados como resolução de equações, automação via macros em VBA (Visual Basic for Applications) e uma ampla biblioteca de funções matemáticas e estatísticas, o que permite personalizar modelos de regressão conforme a complexidade do sistema analisado (WALKENBACH, 2019; OLSEN, 2016; MORAN, 2011; BENNINGA, 2010).

Isso significa que dentre os modelos matemáticos propostos nesse aplicativo, as regressões lineares são as funções polinomiais do 1º ao 6º grau e a logarítmica do tipo $[y = a \cdot \ln(x) + b]$ e as regressões não lineares composta pela função exponencial do tipo $[y = a \cdot \exp(b \cdot x)]$ e pela função potência do tipo $(y = a \cdot x^b)$. Essas funções representam a maioria dos modelos matemáticos aplicados à química com a vantagem dos dados experimentais serem digitados diretamente na planilha sem nenhum tratamento matemático prévio e sem a necessidade de linearização da equação modelo.

Este trabalho teve como objetivo desenvolver um aplicativo baseado em planilhas eletrônicas, capaz de realizar regressões lineares e não lineares simples, por meio de modelos matemáticos aplicados à Química, com geração de gráficos e análise estatística, compatível com computadores e dispositivos móveis. Considerando sua ampla aceitação e recursos avançados, o Microsoft Excel foi adotado como plataforma para a construção do aplicativo.

Material e Métodos

O aplicativo para a realização de regressões lineares e não lineares foi desenvolvido em uma planilha do Excel utilizando operações matriciais com recursos de funções internas como: `MATRIZ.MULT`; `MATRIZ.INVERSO`; `TRANSPOR` e outras funções como: `INCLINAÇÃO` e `INTERCEPÇÃO` para obter os valores iniciais do processo de iteração, `DESVQ` para calcular a soma dos quadrados dos desvios em relação à média, além de funções lógicas como: `SE`, `SES`, `ÉERRO` e `ÉNÚM` para a verificação de erros e adequação das funções de acordo com o tamanho da amostra, as funções `MÁXIMO` e `MÍNIMO` para obtenção das regressões nos gráficos e as funções estatísticas `INVF`, `INVT`, `DISTF` e `DISTT` para os cálculos do inverso da distribuição de probabilidade F ou F-crítico, o inverso bicaudal da distribuição t de Student para calcular os limites de confiança, a distribuição de probabilidade F ou F de significação e a distribuição t de Student para o cálculo do valor-p, respectivamente.

Em todas as regressões são obtidas as matrizes Jacobianas para calcular a matriz das variâncias e covariâncias necessária para o cálculo do erro padrão na análise estatística, bem como os incrementos dos processos iterativos.

A planilha “dados” contém apenas duas colunas para a digitação dos dados experimentais (variável independente “x” e dependente “y”) e um quadro com o resultado de todas as regressões indicando a melhor regressão com links para visualização dos gráficos e análises estatísticas, figura 1A.

A análise estatística (variância e regressão) foi realizada utilizando equações da literatura (STEVENSON, W. J., 1981).

O aplicativo desenvolvido envolve as áreas de química e informática e pode ser utilizado como ferramenta auxiliar nas análises de medidas experimentais através de regressões lineares e não lineares com visualização gráfica e análise estatística.

Outra grande vantagem desse aplicativo é o uso da análise estatística para prevenir a superestimação dos parâmetros ou coeficientes das funções através do valor F de significância da análise de variância (ANOVA) e do valor-p dos parâmetros.

A avaliação da qualidade dos ajustes foi realizada por meio de métricas estatísticas como a Soma dos Quadrados dos Resíduos (SQR), o coeficiente de determinação (R^2), o teste F e o valor-p, o que permite a escolha objetiva do modelo mais adequado a cada conjunto de dados. Os exemplos utilizados foram simulados ou extraídos da literatura especializada (MONK, P; MUNRO, L. J., 2009), visando contemplar diferentes contextos aplicáveis à Química.

Resultados e Discussão

Este aplicativo na forma de planilhas eletrônicas para a realização de regressões lineares e não lineares simples foi desenvolvido com o propósito de efetuar regressões lineares e não lineares de funções exponenciais, logarítmicas, potência e polinomiais do 1º ao 6º grau, o que torna o aplicativo de grande utilidade na química, visto que essas funções representam a maioria das regressões utilizadas em experimentos químicos.

O aplicativo desenvolvido foi testado utilizando dados simulados e conjuntos de dados extraídos da literatura química, abrangendo situações experimentais representativas como processos de adsorção, cinética enzimática, degradação catalítica, reações de pseudo-primeira ordem, entre outros. Essas aplicações visam não apenas demonstrar a funcionalidade do sistema, mas também ilustrar a diversidade de contextos químicos em que a regressão linear e não linear se fazem necessárias para interpretar dados experimentais.

A figura 1B ilustra uma regressão não linear da função potência do tipo “ $y = a \cdot x^b$ ” de um experimento relacionado ao estudo dos gases ideais de uma transformação isotérmica a partir dos valores da pressão e volume. Os dados foram digitados diretamente na planilha “dados” (figura 1), sem nenhum tratamento matemático e o resultado das nove regressões é mostrado em uma tabela indicando a melhor em coloração verde com links para visualização dos gráficos e análise estatística. Também é disponibilizado ao lado do gráfico um recurso extra para interpolação e extrapolação de dados a partir dos coeficientes obtidos das regressões. Nessa regressão, O coeficiente “a” está relacionado com o produto das constantes “ $n \cdot R \cdot T$ ”. Assim, para uma temperatura qualquer é possível calcular o número de moles do gás em estudo através do coeficiente obtido por regressão. Nesse experimento, $n = 4$, $R = 0,08205 \text{ atm L mol}^{-1} \text{ K}^{-1}$ e $T = 300 \text{ K}$. O coeficiente “a” obtido por regressão foi igual a 98,463 e o coeficiente “b” igual a -1.

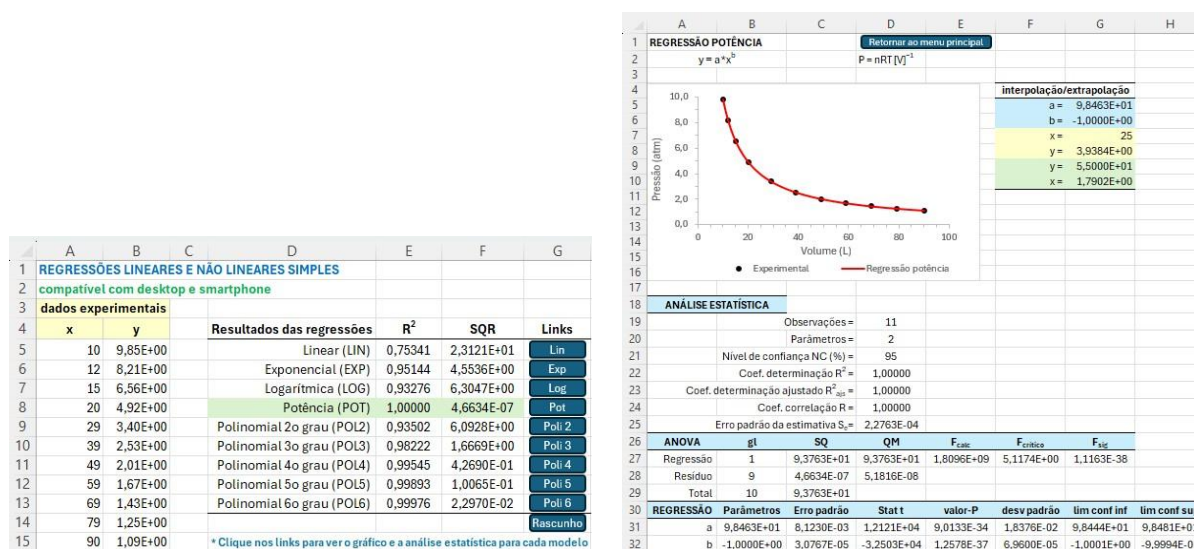
A figura 2A ilustra uma regressão não linear da função exponencial do tipo “ $y = a \cdot \exp(b \cdot x)$ ” de uma cinética química de 1ª ordem monitorada pelo reagente a partir dos dados do tempo e concentração molar do reagente. Os dados experimentais sempre são digitados na planilha “dados” e em todas as regressões não há necessidade de nenhum tratamento matemático para linearização da função. Nessa regressão, o coeficiente “a” está relacionado com a contração molar inicial do reagente e o coeficiente “b” está relacionado com o oposto da constante de velocidade da reação ($-k$). Os valores obtidos de a e b foram, respectivamente, $9,9999 \times 10^{-2}$ e $-5,9999 \times 10^{-3}$. Assim, o valor da concentração inicial foi aproximadamente 0,1 mol/L e a constante de velocidade aproximadamente $6,00 \times 10^{-3} \text{ min}^{-1}$.

A figura 2B ilustra uma regressão linear da função logarítmica do tipo “ $y = a \cdot \ln(x) + b$ ” de um experimento eletroquímico relacionado ao potencial de eletrodo do par Cd(II)/Cd em função da concentração molar do íon Cd^{2+} . Nessa regressão, o coeficiente “a” está relacionado com a constante eletroquímica “ $R \cdot T / (n \cdot F)$ ” e o coeficiente “b” relacionado ao potencial padrão de redução do par Cd(II)/Cd. Onde: R é a constante universal dos gases, T a temperatura absoluta, n o número de elétrons envolvido no processo e F a constante de Faraday. Os coeficientes “a” e “b” obtidos por regressão foram, respectivamente, $1,2837 \times 10^{-2}$ e $-0,400 \text{ V}$, o que confirma o processo envolvendo dois elétrons e o valor do potencial padrão de redução do par Cd(II)/Cd de acordo com a literatura.

O método utilizado para a realização das regressões é o da soma dos mínimos quadrados dos resíduos, SQR, embora seja um método eficiente, existe a possibilidade dos parâmetros ou coeficientes das funções estarem superestimados para produzir valores menores de SQR. Para saber qual função tem a melhor regressão utilizaremos a análise estatística através dos valores de F de significância e do valor-p dos parâmetros.

Os parâmetros estatísticos mais utilizados em análise de regressão são: SQR (soma dos quadrados dos resíduos), R^2 (coeficiente de determinação), F_{sig} (F de significância da tabela ANOVA), erro padrão e valor-p dos coeficientes.

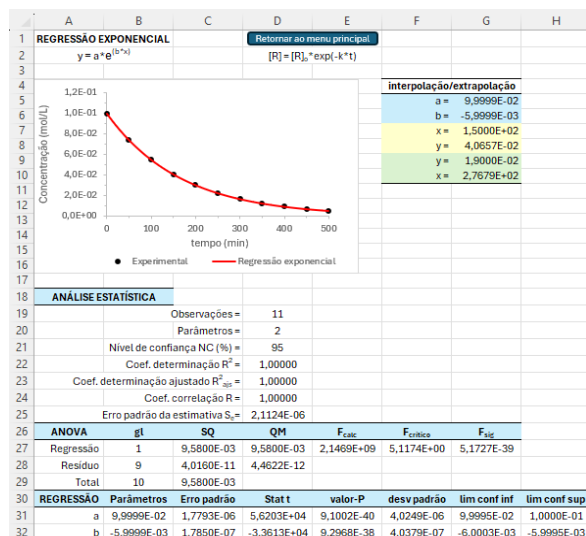
Figura 1 – (a) Planilha “dados” mostrando as colunas para digitação dos dados experimentais e uma tabela com o resultado das regressões indicando a melhor na coloração verde com links para visualização dos gráficos e análise estatística; (b) Planilha “POT” mostrando o gráfico e a análise estatística da regressão não linear da função potência do tipo $y = a \cdot x^b$.



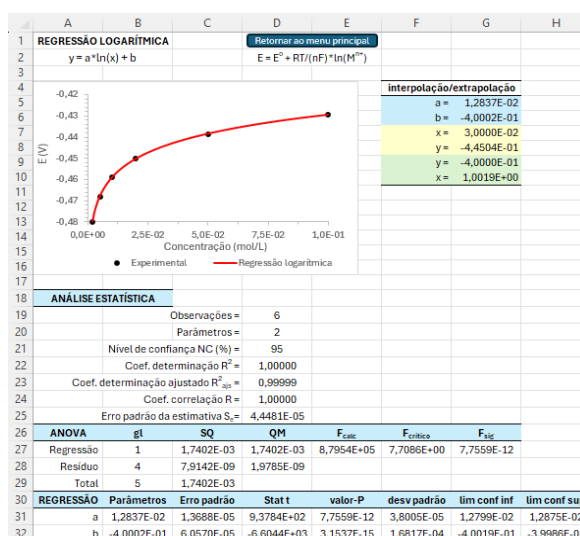
(a)

(b)

Figura 2 – (a) Planilha “EXP” mostrando o gráfico e a análise estatística da regressão não linear da função exponencial do tipo “ $y = a \cdot \exp(b \cdot x)$ ”; (b) Planilha “LOG” mostrando o gráfico e a análise estatística da regressão linear da função logarítmica do tipo “ $y = a \cdot \ln(x) + b$ ”.



(a)



(b)

Conclusões

O aplicativo desenvolvido em planilha eletrônica, utilizando a plataforma Microsoft Excel, mostrou-se uma ferramenta eficiente e acessível para realizar regressões lineares e não lineares simples com análise estatística integrada. Sua interface intuitiva e a capacidade de processar dados sem a necessidade de linearização prévia ou de software especializado representam uma contribuição valiosa para a análise de dados experimentais em Química. A integração de métricas estatísticas robustas, como SQR, R^2 , o teste F e o valor-p permite a escolha objetiva do modelo mais adequado, prevenindo a superestimação de parâmetros e assegurando a confiabilidade dos resultados. Além de ser uma ferramenta prática para a pesquisa, este aplicativo demonstra o potencial do Excel como um recurso didático poderoso no ensino de Química. Ao permitir que os estudantes explorem conceitos de modelagem matemática e análise estatística de forma direta e visual, ele estimula o raciocínio científico, incentiva a autonomia e contribui significativamente para uma aprendizagem mais dinâmica e efetiva.

Agradecimentos

Ao grupo de pesquisas “Mídias no Ensino de Ciências: pesquisa e produção” pelo apoio.

Referências

- BATES, D. M., WATTS, D. G. Nonlinear regression analysis and its applications. Wiley, New York, 1988.
- BENNINGA, S. Principles of Finance with Excel. Oxford University Press, 2010.
- CHIACCHIO, E. J. Regressão não-linear: desenvolvimento de um sistema computacional e aplicações, Dissertação de mestrado, USP – Piracicaba – SP, 1993. [https://www.teses.usp.br/index.php?option=com_jumi&fileid=11&Itemid=76&lang=pt-br]
- DRAPER, N. R., SMITH, H. Applied regression analysis. Wiley, New York, 1981.



- LAI, W. H.; KEK, S. L.; TAY, K. G. Solving Nonlinear Least Squares Problem Using Gauss-Newton Method, *IJISSET - International Journal of Innovative Science, Engineering & Technology*, 4, 1, 2017, [https://ijiset.com/vol4/v4s1/IJISSET_V4_I01_35.pdf].
- MATTOS, T. B. Modelos não lineares e suas aplicações, Monografia, UFJF, Juiz de Fora, MG, Brazil, 2013. [<https://www2.ufjf.br/ppge/dissertacoes/teses/>]
- MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear, *Acta Scientiarum*, 2002 24 6 1761-1770, Maringá – PR.
- MONK, P.; MUNRO, L, J. Matemática para a Química: uma caixa de ferramentas de cálculo dos químicos, segunda edição, LTC, 2009.
- MORAN, M. J. Planilhas Eletrônicas com Excel. Pearson, 2011.
- OLSEN, P. Microsoft Excel Data Analysis and Business Modeling. Microsoft Press, 2016
- SEBER, G. A. F. Linear regression analysis. Wiley, New York, 1977.
- SILVA, E. M., et al. Método de Newton e Gauss-Newton na estimação dos parâmetros de modelo de regressão não linear. *Sigmae* 2019 8, n,2, p. 728-734, Alfenas., ISSN: 2317-0840.
- SIREGAR, R. W. et al. Analysis Local Convergence of Gauss-Newton Method, *IOP Conf. Ser.: Mater. Sci. Eng.* 300, 012044, 2018, [10.1088/1757-899X/300/1/012044, <https://iopscience.iop.org/article/10.1088/1757-899X/300/1/012044/pdf>].
- STEVENSON, W. J. Estatística aplicada à administração, Harper & Row do Brasil, São Paulo – SP, 1981.
- WALKENBACH, J. Excel Bible. Wiley, 2019.