

Método de Quadrados Mínimos Fuzzy para Estudo de HIV a Partir de Dados Imprecisos

Ana Luz P. Mendes¹, Yasmin B. Shimizu², Letícia A. Nunes³, Vinícius F. Wasques⁴
 Ilum/CNPEM, Campinas, SP.

Diversos problemas científicos, por mais complexos que sejam, seguem padrões bem definidos que podem ser descritos por ferramentas matemáticas. Ao realizarmos medições de certos dados ao longo do tempo, é possível identificar qual o comportamento do nosso objeto de estudo e, a partir disso, descrever matematicamente o problema de forma que seja possível a realização de previsões futuras, a partir de acontecimentos passados. No entanto, uma grande parte dos fenômenos da natureza possui uma incerteza intrínseca ou não somos capazes de coletar dados o suficiente para afirmar com absoluta certeza o comportamento do determinado fenômeno investigado. Por exemplo, ao informar o número de infectados por HIV no Brasil, é ilógico afirmar que há exatos 50 mil infectados e, na realidade, o mais usual é falarmos “em torno de 50 mil”. Esse fato faz com que tenhamos incerteza tanto no número de infectados quanto no número de suscetíveis. Aqui, essa incerteza será tratada pela teoria de conjuntos fuzzy.

Para isso, precisamos de ferramentas matemáticas que não se restringem a uma natureza binária, na qual variáveis possuam graus de pertinência a um determinado objeto, possibilitando a descrição de diversos fenômenos sem restringi-los a valores únicos. Além disso, pode existir a dependência entre os dados, como ocorre por exemplo com dados longitudinais. Na teoria de conjuntos fuzzy, essa dependência é descrita por uma relação chamada de interatividade, definida por distribuições de possibilidade conjunta e manipulada algebricamente pelo princípio de extensão sup- J . Neste trabalho, a interatividade considerada foi a gerada a partir da distribuição J_0 , proposta por [4]. No entanto, tais ferramentas exigem uma quantia considerável de operações que seriam quase impraticáveis de serem realizadas à mão. Portanto, para isso é necessário o uso de recursos computacionais para realização de cálculos precisos em um curto intervalo de tempo.

Para modelarmos esses dados, ou seja realizarmos o ajuste de uma curva que tenha entradas fuzzy, lidamos com a reinterpretação de ferramentas da matemática clássica como o método dos quadrados mínimos e a decomposição LU. A proposta é produzir um ajuste linear fuzzy que se adeque não apenas a parte dos dados do conjunto real clássico, mas também que apresente uma faixa que nos informe sobre como a incerteza se projeta ao longo dessa curva.

O método dos quadrados mínimos é uma técnica de otimização utilizada com a finalidade de preencher, estimar ou prever informações através de uma aproximação da curva que melhor descreve a dispersão dos dados conhecidos, minimizando o erro quadrado entre cada ponto e a curva prevista. A adaptação desse método para receber números fuzzy resulta em um conjunto de sistemas lineares que possui uma mesma matriz de coeficientes e, para resolvê-los, utilizamos a decomposição LU [1]. A decomposição LU é um método direto de resolução de sistemas lineares do tipo $Ax = b$ na qual se fatora a matriz A em uma matriz triangular inferior com diagonal unitária L e uma matriz triangular superior U . Com essa decomposição, pode-se resolver, quase que imediatamente, qualquer sistema linear que possua a mesma matriz A não singular de coeficientes [2].

¹ana24007@ilum.cnpem.br

²yasmin24023@ilum.cnpem.br

³leticia24013@ilum.cnpem.br

⁴vwasques@outlook.com

A partir da transformação de cada um dos dados incertos em um número fuzzy triangular e da decomposição LU, é possível obter a matriz x que contém os coeficientes fuzzy da função que descreve a variável em função do tempo, bem como a incerteza associada a esta. Assim, além do aprofundamento teórico, foi desenvolvido um algoritmo computacional para a utilização desses métodos e plotagem das curvas. Usando a teoria de conjuntos fuzzy, foi feita a modelagem da incerteza de dados sobre infectados por HIV no Brasil e a previsão do número de infectados nos três anos seguintes. Todas as operações foram realizadas a partir de funções de autoria própria desenvolvidas em *Python* com o suporte das bibliotecas *Numpy* e *Matplotlib*. Com base nas soluções obtidas, foi possível construir gráficos 2D e 3D para a melhor compreensão do fenômeno e visualização da pertinência dos valores. Sobre isso, é relevante mencionar que são geradas até quatro soluções devido ao método interativo utilizado. Dessa forma, também foi necessário realizar uma análise sobre a coerência das curvas de ajuste. No contexto do fuzzy, uma menor incerteza se traduz em uma menor abertura (diâmetro), que é o perfil procurado entre as possíveis curvas. Por fim, apresentamos, ainda, a previsão do número de infectados por HIV até o ano de 2024 [3]. Desse modo, obtivemos os resultados apresentados na Tabela 1. Além disso, uma discussão sobre a modelagem do problema biológico é apresentada.

Tabela 1: Previsão do número de infectados por HIV no Brasil de 2022 a 2024.

Ano	Previsão
2022	(11752.0; 24357.5; 36962.2)
2023	(6123.7; 21032.6; 35940.6)
2024	(495.4; 17707.7; 34919.0)

Comparando o trabalho proposto a um ajuste tradicional, pudemos identificar a relevância do método para análise de dados incertos, assim como a necessidade de identificar a melhor solução para o problema em questão. Com essa abordagem, pode-se modelar as incertezas de dados biológicos, procurando uma melhor faixa que os descreva. Assim, o estudo em relação à regressão linear de dados triangulares fuzzy para análise de dados biológicos foi efetivo. Com isso, foi verificado que as subjetividades que a informação biológica de infectados possui pôde ser modelada por meio da matemática fuzzy, mantendo a linguagem do “em torno de”.

Agradecimentos

O quarto autor agradece o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2023/03927-0.

Referências

- [1] N. J. B. Pinto, V. F. Wasques, E. Esmi e L.C. Barros. “Sistemas lineares fuzzy e problemas de mínimos quadrados aplicados a dados longitudinais”. Em: **IMECC – Unicamp: Biomatemática** 32 (2022), pp. 43–60.
- [2] M. A. G. Ruggiero e V. L. Lopes. **Cálculo Numérico - Aspectos Teóricos e Computacionais**. São Paulo: Editora Makron Books, 1996. ISBN: 9788534602044.
- [3] Ministério da Saúde. **Boletim epidemiológico HIV/AIDS 2021**. Online. Acessado em 25/04/2025, <https://antigo.aids.gov.br/pt-br/pub/2021/boletim-epidemiologico-hivaids-2021>.
- [4] V. F. Wasques. “Fuzzy Differential Equations via Interactive Arithmetic: Applications in Biomathematics”. Tese de doutorado. IMECC/UNICAMP, 2019.