Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 10, n. 1, 2023.

Trabalho apresentado no XLII CNMAC, Universidade Federal de Mato Grosso do Sul - Bonito - MS, 2023

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Uso de aprendizado profundo na seleção do parâmetro de regularização

Emanuel Oliveira Souza[1] Elias Salomão Helou Neto[2]
ICMC-USP, São Carlos, SP

Os problemas inversos surgem em várias áreas das ciências e da engenharia, onde aparecem quando o objetivo é obter as causas desconhecidas a partir dos efeitos observados. Esta classe de problemas tem bastante importância para a sociedade, por ajudarem a obter informações importantes usadas muitas vezes em tomadas de decisões. Por exemplo, na medicina, um problema inverso surge quando o objetivo é reconstruir a estrutura interna de um corpo físico, tendo apenas informações tiradas através dos feixes de raios-x em diferentes posições. Este problema é conhecido popularmente como tomografia por raios-x, utilizada principalmente em diagnóstico médico. Alguns problemas inversos são modelados como

$$T\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{\epsilon}, \tag{1}$$

onde T é um operador linear limitado entre dois espaços de Hilbert $\mathcal{X}$ e $\mathcal{Y}$, $\boldsymbol{x}$ é a solução procurada, ou seja, a causa, $\boldsymbol{y}$ é o efeito observado e $\boldsymbol{\epsilon}$ é o ruído desconhecido. O principal obstáculo dos problemas inversos é o fato que boa parte é mal-posta no sentido de Hadamard [1]. Esta dificuldade vem principalmente da descontinuidade das pseudos-inversas dos operadores, o que as torna instáveis numericamente. Uma condição suficiente para verificar a continuidade das pseudos-inversas no caso que os operadores são compactos, é conhecido como a condição de Picard, que estabelece a existência de soluções estáveis baseada na sequência de autovalores do operador.

Portanto, resolver um problema inverso é obter uma aproximação $\hat{\boldsymbol{x}}$ de $\boldsymbol{x}$, conhecido o operador T e a observação $\boldsymbol{y}$. Uma abordagem bastante usada é utilizar técnicas de regularização [2] para encontrar soluções razoáveis. Uma abordagem possível é a regularização de Tikhonov [3] formulada em sua versão variacional como se segue:

$$\hat{\boldsymbol{x}}_\alpha := \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\min} \left\{ \|T\boldsymbol{x} - \boldsymbol{y}^\delta\|_{\mathcal{Y}}^2 + \alpha\|\boldsymbol{x}\|_{\mathcal{X}}^2 \right\}, \tag{2}$$

onde $\alpha > 0$ é o parâmetro de regularização e $\boldsymbol{y}^\delta = \boldsymbol{y} + \boldsymbol{\epsilon}$ com $\|\boldsymbol{y} - \boldsymbol{y}^\delta\|_{\mathcal{Y}} \leq \delta$, sendo δ conhecido como o nível de ruído nos dados. O parâmetro de regularização α tem o objetivo de fornecer equilíbrio entre a estabilidade e precisão da solução. Quanto maior for o parâmetro, mais peso é colocado sobre a norma da solução e assim perdemos precisão. Em contrapartida, quanto menor for o parâmetro, mais estabilidade é perdida.

A escolha do parâmetro de regularização é uma etapa crucial para encontrar uma solução precisa. Existem diversos métodos usados para selecionar o parâmetro de regularização, por exemplo, o princípio da discrepância [4], a validação cruzada generalizada [5] e a curva-L [6]. Os métodos existentes são desenvolvidos a partir de características do problema, tais como o tipo de ruído contido nos dados e qual a regularização escolhida. Isso dificulta a escolha de um método na prática, pois em alguns casos é necessário ter um conhecimento *a priori* do problema.

[1]emanuel.o.souza@usp.br
[2]elias@icmc.usp.br

O rápido desenvolvimento dos equipamentos de *hardware* e as tecnologias de *software* permitiram o uso continuado e bem-sucedido do aprendizado de máquina em aplicações práticas. Este fato permitiu a criação de novos métodos para selecionar o parâmetro de regularização utilizando conceitos de aprendizado de máquina. Na literatura, é possível encontrar trabalhos recentes que utilizam estes conceitos, pode-se destacar [7] que utiliza um conjunto de dados sintéticos gerados e um modelo de regressão pré-treinado para obter o parâmetro em problemas inversos sísmicos; e [8], que utilizaram redes neurais profundas para selecionar o parâmetro de regularização em diferentes aplicações. Tais abordagens mencionadas obtiveram resultados superiores os métodos tradicionais utilizados, isto reforça que a utilização de técnicas de aprendizado profundo pode fornecer uma melhor aproximação do parâmetro ótimo de regularização.

Com base nestes fatos, o objetivo deste trabalho é desenvolver um novo método de seleção do parâmetro de regularização utilizando técnicas de aprendizado profundo. A estratégia escolhida é dividida em duas etapas: i) criação de um conjunto de dados para cada problema específico; ii) definição da arquitetura de redes que se adapte melhor ao problema a ser resolvido. Desse modo, é esperado que o método desenvolvido possa ser aplicado a quaisquer problemas inversos e que possa servir de inspiração para novas pesquisas na escolha do parâmetro de regularização. Finalmente, tendo os resultados em mãos, a intenção é comparar qualitativamente com os métodos tradicionais usados atualmente em alguns problemas inversos, por exemplo, a tomografia computadorizada, o *deblurring* e a difusão do calor.

Agradecimentos

Agradecemos à CAPES, ao DM Service e ao ICMC-USP pelo apoio prestado.

Referências

[1] Jacques Hadamard. **Lectures on Cauchy's problem in linear partial differential equations**. New York: Dover Publications, 2003. ISBN: 9780486495491.

[2] Heinz Werner Engl, Martin Hanke e Andreas Neubauer. **Regularization of inverse problems**. Vol. 375. Dordrecht, The Netherlands: Springer Science & Business Media, 1996. ISBN: 9780792341574.

[3] Andrey N. Tikhonov e Vasiliy Y. Arsenin. **Solutions of Ill-Posed Problems**. Washington; New York: Winston, 1977. ISBN: 9780470991244.

[4] Vladimir Alekseevich Morozov. "On the solution of functional equations by the method of regularization". Em: **Doklady Akademii Nauk**. Vol. 167. 3. Russian Academy of Sciences. USSR, 1966, pp. 510–512.

[5] Gene H Golub, Michael Heath e Grace Wahba. "Generalized cross-validation as a method for choosing a good ridge parameter". Em: **Technometrics** 21.2 (1979), pp. 215–223. DOI: 10.2307/1268518.

[6] Per Christian Hansen. "Analysis of discrete ill-posed problems by means of the L-curve". Em: **SIAM Review** 34.4 (1992), pp. 561–580. DOI: 10.1137/1034115.

[7] Shihuan Liu e Jiashu Zhang. "Machine-learning-based prediction of regularization parameters for seismic inverse problems". Em: **Acta Geophysica** 69.3 (2021), pp. 809–820. DOI: 10.1007/s11600-021-00569-7.

[8] Babak Maboudi Afkham, Julianne Chung e Matthias Chung. "Learning regularization parameters of inverse problems via deep neural networks". Em: **Inverse Problems** 37.10 (2021), p. 105017. DOI: 10.1088/1361-6420/ac245d.