Fracture Conductivity Prediction Based on Machine Learning in Shale – Part I

Ella Meyer

This document is Part I in a series of documents providing ways on predicting fracture conductivity using machine learning.

KEYWORDS: Fracture conductivity, Machine Learning, Shale, Multivariate Regression

Abstract: Hydraulic fracturing extracts oil and gas from deep underground, with fracture conductivity being crucial for efficient production. Traditional lab techniques for measuring conductivity are costly and time-consuming. This paper explores using machine learning, specifically multivariate regression, to predict fracture conductivity based on experimental data like Poisson's ratio and proppant size. Optimizing these models can enhance hydraulic fracturing efficiency in shale formations.

Hydraulic fracturing, often referred to as fracking, is a technique used to extract oil and gas from deep underground formations. This process involves injecting a high-pressure fluid mixture into the rock layer, creating fractures through which hydrocarbons can flow more freely to the production well. A crucial aspect of hydraulic fracturing is fracture conductivity, which refers to the ability of the created fractures to allow the flow of hydrocarbons. High fracture conductivity is essential for maintaining efficient hydrocarbon production. It depends on various factors, including the proppant type and size, the closure stress, and the properties of the fracturing fluid. Optimizing these factors ensures that the fractures remain open and conductive over the lifespan of the well (Montgomery, 2013; Holditch, 2006; Sharma et al., 2014).

Determining fracture conductivity in situ is a very challenging task and conducting lab measurements using Hassler core technique on shale are expensive and time consuming. The Hassler core experiment is a laboratory technique that simulates in-situ conditions to measure the permeability and conductivity of fractures in core samples. This method involves placing a core sample within a Hassler sleeve, applying confining pressure, and injecting fluid to simulate the hydraulic fracturing process. Researchers measure the resultant fracture conductivity, which is a key parameter influencing the efficiency of hydrocarbon extraction from shale formations. (Wu et al., 2017, Wu et al., 2019).

Machine learning is being used to advance scientific computing in a variety of fields, such as fluid mechanics, (Raissi et al., 2019, Zhang et al., 2020, Wang el al., 2017) solid mechanics, (Haghighat et al., 2020, Zheng et al., 2022, Arora et al., 2022) materials science etc (Kim et al., 2021). Machine learning techniques can also be applied to predict the fracture conductivity of new shale formations based on models built using data collected from experiments.

Good experimental data is key to building a well-performing prediction model. In order to build a model to predict fracture conductivity, the following information about various shale samples needs to be recorded while conducting experiments: Poisson's ratio, Young's Modulus, Temperature, Closure pressure, Proppant particle size, and Sand concentration. The wider the range of the above-mentioned parameters, the better the model will be.

There are various machine learning techniques that can be used to build the above model. In this paper we will discuss about multivariate regression.

Multivariate Regression

To build a robust multivariate regression model for predicting fracture conductivity, several mathematical steps and techniques should be adopted:

1. **Data Representation and Feature Matrix Construction**: Begin by organizing the collected data into a matrix Xwhere each row represents an individual shale sample, and each column represents a feature (e.g., Poisson's ratio, Young's Modulus, Temperature, Closure pressure, Proppant particle size, Sand concentration). Let y be the vector of observed fracture conductivity values. Mathematically, this can be represented as:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

   where n is the number of samples and p is the number of features.

2. **Model Formulation**: The multivariate regression model assumes a linear relationship between the dependent variable y and the independent variables X. The model can be expressed as:

$$y = X\beta + \epsilon$$

3. **Parameter Estimation**: The coefficients β can be estimated using the Ordinary Least Squares (OLS) method, which minimizes the sum of the squared residuals. The OLS estimator is given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

4. **Regularization (if necessary):** To prevent overfitting, especially with high-dimensional data, regularization techniques such as Ridge Regression (L2 regularization) or Lasso Regression (L1 regularization) can be applied. Ridge regression minimizes:

$$\hat{\beta}_{\text{ridge}} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} (y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}$$

   and Lasso regression minimizes:

$$\hat{\beta}_{\text{lasso}} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} (y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\}$$

where λ is the regularization parameter.

5. **Model Validation and Performance Evaluation:** After fitting the model, its performance should be evaluated using metrics such as the coefficient of determination (R^2), mean squared error (MSE), and root mean squared error (RMSE). Cross-validation techniques, such as k-fold cross-validation, can provide a more robust assessment of model performance. Mathematically:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\text{RMSE} = \sqrt{\text{MSE}}$$

6. **Model Interpretation and Application**: Once the model is validated, interpret the coefficients to understand the relationship between the features and fracture conductivity. Apply the model to predict fracture conductivity for new shale formations by inputting their corresponding feature values into the regression equation.

By following these mathematical steps, a comprehensive and accurate multivariate regression model can be developed, providing valuable insights for predicting the fracture conductivity of shale formations.

References

Montgomery, C. T. (2013). "Fracturing Fluids." In Hydraulic Fracturing (pp. 67-98). Elsevier.

Holditch, S. A. (2006). "Tight gas sands." Journal of Petroleum Technology, 58(06), 86-93.

Sharma M. M et al. (2014). "Proppant Selection and Its Effect on Fracture Conductivity and Production." SPE Hydraulic Fracturing Technology Conference.

Wu W, Kakkar P, Zhou J. et al. 2017. An Experimental Investigation of the Conductivity of Unpropped Fractures in Shales. Presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, 24-26 January. SPE-184858-MS. https://doi.org/10.2118/184858-MS.

Wu W, Zhou J, Kakkar P et al. An experimental study on conductivity of unpropped fractures in preserved shales. SPE Production & Operations, 2018, 34(2): 280–296

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378: 686–707, 2019. DOI: https://doi.org/10.1016/j.jcp.2018.10.045.

T. Zhang, B. Dey, P. Kakkar, A Dasgupta, and A. Chakraborty. Frequency-compensated pinns for fluid-dynamic design problems. arXiv preprint arXiv:2011.01456, 2020. DOI: https://doi.org/10.48550/arXiv.2011.01456.

J. Wang, J. Wu, and H. Xiao. Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data. Physical Review Fluids, 2(3):034603, 2017. DOI: https://doi.org/10.1103/PhysRevFluids.2.034603.

E. Haghighat, M. Raissi, A. Moure, H. Gomez, and R. Juanes. A deep learning framework for solution and discovery in solid mechanics: linear elasticity. arXiv preprint arXiv:2003.02751, 2020.

B. Zheng, T. Li, H. Qi, L. Gao, X. Liu, and L. Yuan. Physics-informed machine learning model for computational fracture 41 of quasi-brittle materials without labelled data. International Journal of Mechanical Sciences, 223:107282, 2022. DOI: https://doi.org/10.1016/j.ijmecsci.2022.107282.

R. Arora, P. Kakkar, B. Dey, and A. Chakraborty. Physics-informed neural networks for modeling rate- and temperature dependent plasticity. arXiv preprint arXiv:2201.08363, 2022. DOI: https://doi.org/10.48550/arXiv.2201.08363

Y. Kim, C. Yang, K. Park, G. X. Gu, and R. Seunghwa. Deep learning framework for material design space exploration using active transfer learning and data augmentation. npj Computational Materials, 7:140, 2021. DOI: https://doi.org/10.1038/s41524-021-00609-2