

Aprendiendo física con grafos

Quercus Hernández¹, Alberto Badías^{1,2}, Francisco Chinesta³, Elías Cueto¹

¹ Applied Mechanics and Bioengineering (AMB)

Instituto de Investigación en Ingeniería de Aragón (I3A)

Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.

Tel. +34-976762707, e-mail: quercus@unizar.es

² Universidad Politécnica de Madrid (UPM)

³ ENSAM (ParisTech)

Resumen

El objetivo de este trabajo es utilizar inteligencia artificial para aprender las leyes físicas que gobiernan un sistema arbitrario y predecir así su evolución en el tiempo. Para ello se utilizan varios sesgos inductivos, que aseguran la correcta estructura matemática del problema y mejoran su capacidad de generalización.

Introducción

La inteligencia artificial y, en concreto, las redes neuronales profundas (*deep neural networks*) están revolucionando el panorama científico actual y futuro. La creciente capacidad de cálculo y los excelentes resultados en comparación con técnicas tradicionales han propiciado su uso casi obligado en determinados ámbitos científicos.

Sin embargo, a la hora de modelizar fenómenos complejos existen dificultades para el entrenamiento como pueden ser la ingente cantidad de datos necesarios o la inasumible potencia de cálculo. En este sentido, muchos autores están desarrollando nuevas arquitecturas de red que permiten manejar los datos de forma más inteligente y así poder llegar a la solución correcta del problema de una manera más sencilla. A esto se le denomina sesgo inductivo (*inductive bias*).

Un tipo de sesgo inductivo objeto de múltiples líneas candentes de investigación es el denominado *geometric deep learning* [1]. Se trata de tomar ventaja de la estructura geométrica del problema para poder simplificar enormemente los cálculos. Este tipo de técnica revolucionó el procesamiento de imagen y de datos temporales hace varias décadas con el desarrollo de las redes convolucionales y recurrentes. Dichas arquitecturas aprovechan la información estructurada en malla de píxeles y series de datos secuenciales para realizar operaciones sobre pequeños parches locales de información, o convoluciones, en lugar de costosas operaciones sin estructura (*fully-connected*).

El objetivo de este trabajo es utilizar la correcta caracterización geométrica de simulaciones de mecánica del sólido y de fluidos, mediante la utilización de redes neuronales basadas en grafos [2].

Metodología

El trabajo se basa en la imposición de dos sesgos inductivos: uno relacionado con la estructura física y otro relacionado con la estructura geométrica del problema.

Estructura metripléctica

En trabajos anteriores [3], se probó la idoneidad de utilizar integradores físicamente consistentes mediante el formalismo GENERIC [4]. Dicha estructura divide al sistema en dos contribuciones distintas: una conservativa, descrita por la mecánica Hamiltoniana, y otra disipativa, de tal forma que la evolución del vector de estados de un sistema (\mathbf{z}) queda determinado por:

$$\frac{d\mathbf{z}}{dt} = \mathbf{L} \frac{\partial E}{\partial \mathbf{z}} + \mathbf{M} \frac{\partial S}{\partial \mathbf{z}}.$$

Las matrices \mathbf{L} y \mathbf{M} codifican la relación entre el vector de estados \mathbf{z} y la energía E y entropía S del sistema respectivamente. Dicha estructura garantiza la consistencia termodinámica de los resultados, cumpliendo con la primera y segunda ley de la termodinámica.

Estructura geométrica

Un factor común en la mayoría de las simulaciones mecánicas basadas en elementos o diferencias finitas es la discretización del dominio $\Omega \subset \mathbb{R}^d$ en subdominios discretos. Dichos subdominios están formados por nodos y relaciones de adyacencia entre ellos, formando una malla. Dicha malla se puede representar matemáticamente por un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{g})$, donde \mathcal{V} representa los nodos de la malla (vértices), \mathcal{E} las relaciones entre los nodos (aristas) y

$\mathbf{g} \in \mathbb{R}^{F_g}$ las propiedades globales del sistema (ver Figura 1).

Cada nodo y arista tienen asociadas unas variables $\mathbf{v}_i \in \mathbb{R}^{F_v}$ y $\mathbf{e}_{ij} \in \mathbb{R}^{F_e}$ generadas a partir de las variables de estado \mathbf{z}_i del sistema. Dichos vectores se procesan mediante varias redes neuronales compartidas en todo el grafo, que se encargan de extraer la información relevante de la estructura metripléctica del problema (en este caso, las matrices \mathbf{L} y \mathbf{M} y los potenciales E y S). Durante este proceso, la información de los nodos se comparte según la conectividad de la malla en un algoritmo que se denomina *message passing*, que sería el equivalente en grafos a la operación de convolución en imagen.

Aquí es precisamente donde reside la potencia de operar sobre grafos. Esto permite a un nodo detectar los cambios de sus vecinos en base a su conectividad, por lo que puede interpretar y aprender relaciones no-Euclideas arbitrariamente complicadas.

Resultados

Para demostrar la efectividad del método, se han generado datos sintéticos de un flujo de Couette de un fluido viscolástico de tipo Olroyd-B de forma paramétrica con varios números de Weissenberg y Reynolds. Se ha realizado una comparación entre tres métodos: utilizando únicamente la estructura física del problema [3], utilizando únicamente estructura basada en grafos [2] y una combinación de ambas.

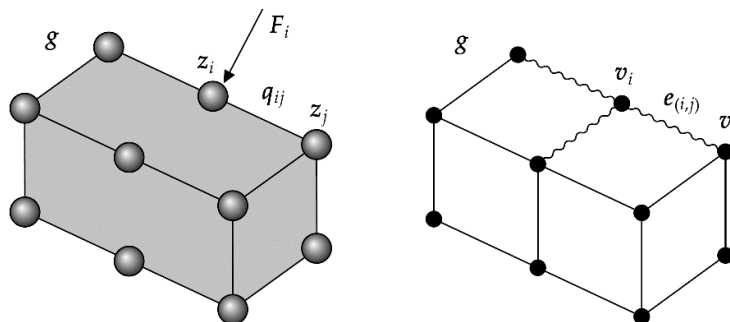


Figura 1. Abstracción de un problema físico en un grafo.

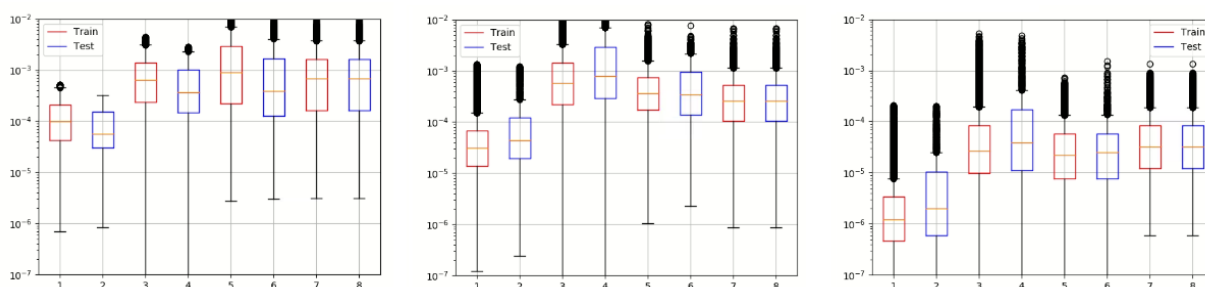


Figura 2. Resultado de aplicar sesgo metripléctico (izquierda), geométrico (centro) y combinados (derecha). El error es cuadrático medio y se representan *train* y *test* por separado.

Como se puede observar en la Figura 2, la combinación de ambos sesgos inductivos reduce el error hasta en un orden de magnitud.

Conclusiones

En este trabajo se demuestra cómo introducir sesgos inductivos tanto de tipo geométrico como físico mejora sustancialmente la calidad de las predicciones de la red neuronal, y mejora la capacidad de generalización del problema.

REFERENCIAS

- [1]. BRONSTEIN, M. M., BRUNA, J., LECUN, Y., SZLAM, A., and VANDERGHEYNST, P., 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4), 18-42.
- [2]. PFAFF, T., FORTUNATO, M., SANCHEZ-GONZALEZ, A., & BATTAGLIA, P., 2020. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*.
- [3]. HERNANDEZ, Q., BADIAS, A., GONZALEZ, D., CHINESTA, F. and CUETO, E., 2020. Structure-preserving neural networks. *Journal of Computational Physics*.
- [4]. GRMELA, M. and ÖTTINGER, H.C., 1997. Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. *Physical Review E*, 56(6), p.6620.