Introducción a las ecuaciones diferenciales neuronales

VI semana de la investigación Konrad lorenz

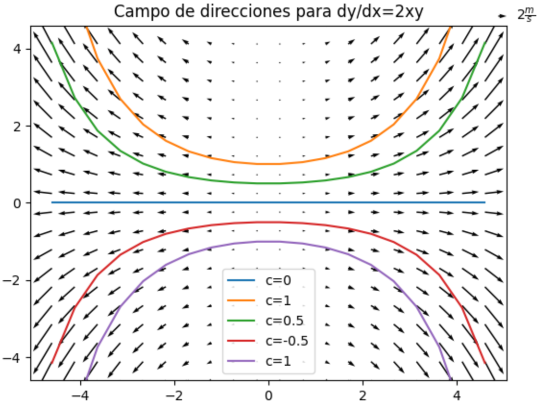
Nathalia Castiblanco Carretero

Fundación Universitaria Konrad Lorenz
Facultad de Matemáticas e Ingeniería

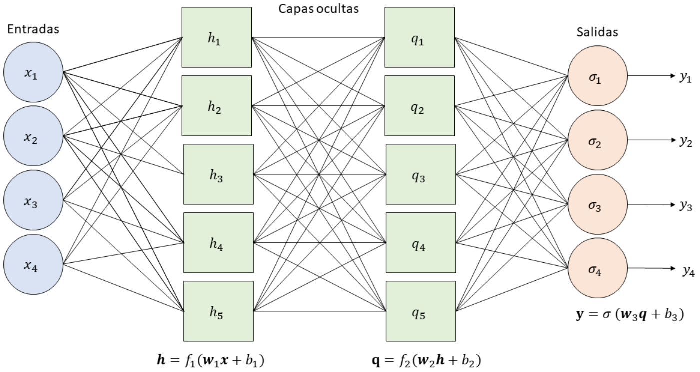
5 de Septiembre de 2024

Tabla de contenidos

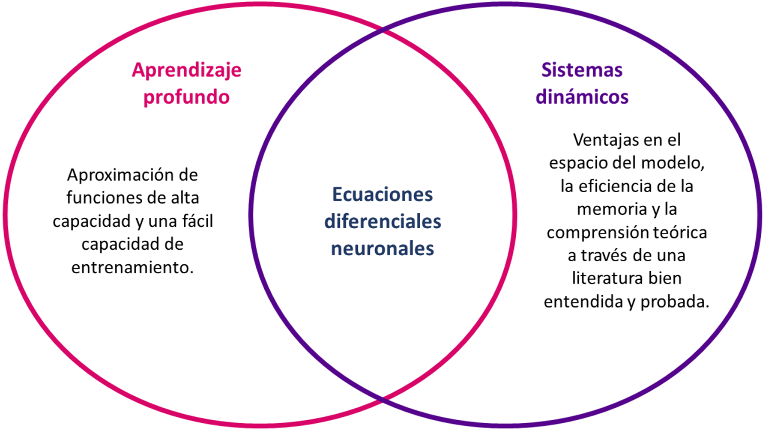
Introducción



Introducción



Introducción



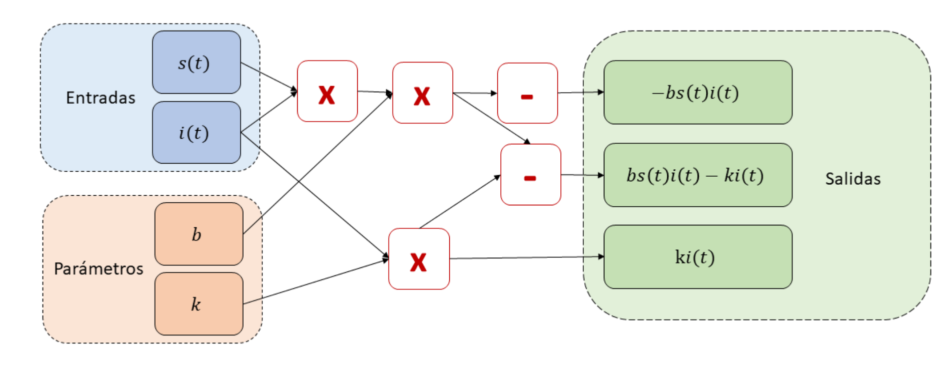
Definición

*Una **ecuación diferencial neuronal** es una ecuación diferencial que utiliza una red neuronal para parametrizar el campo vectorial.*

El ejemplo canónico de una **ecuación diferencial ordinaria neuronal** [1] es

$$\begin{aligned} y(0) &= y_0, \\ y'(t) &= f_\theta(t, y(t)), \end{aligned} \tag{1}$$

donde $y_0 \in \mathbf{R}^{d_1 \times \cdots d_k}$ es un tensor de cualquier dimensión, θ representa algún vector de parámetros aprendidos y $f_\theta : \mathbf{R} \times \mathbf{R}^{d_1 \times \cdots d_k} \to \mathbf{R}^{d_1 \times d_k}$ es una red neuronal. Generalmente, f_θ va a ser una arquitectura neuronal simple, como un perceptron multicapa o una red residual. [2, 1]



Entradas
$s(t)$
$i(t)$
Parámetros
b
k
X
X
X
-
-
$-bs(t)i(t)$
$bs(t)i(t) - ki(t)$
$ki(t)$
Salidas

El método de Euler en:

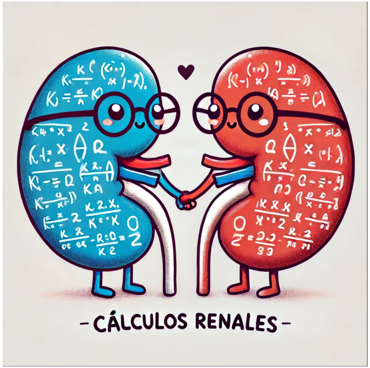
$$y'(t) = f_\theta(t, y(t)),$$

Produce una red neuronal residual:

$$y_{t+1} = y_t + \Delta t f_\theta(y_t, t).$$

Por coincidencia, muchas de las arquitecturas de aprendizaje profundo más efectivas y populares se asemejan a ecuaciones diferenciales. Quizás no deberíamos sorprendernos: las ecuaciones diferenciales han sido el paradigma dominante de modelado durante siglos; no son tan fáciles de derrocar.

- Una de las aplicaciones más relevantes de las matemáticas en la medicina es el modelamiento matemático.
- Con el continuo avance computacional surge otra aplicación de las matemáticas en la medicina relacionada con el uso de las redes neuronales y la inteligencia artificial.



Un tumor es un crecimiento anormal de tejido. Cuando un tejido desconocido aparece en el cuerpo, el sistema inmune trata de identificarlo y, si es posible, tratar de eliminarlo. La dinámica de la respuesta inmunitaria antitumoral *en vivo* es complicada y no se conoce bien, por lo que numerosos modelos matemáticos se han desarrollado acerca del crecimiento de un tumor.

El siguiente modelo planteado en 2003 por Galach [3] que describe la interacción entre las células efectoras (por ejemplo, linfocitos T, citotóxicos) y células tumorales:

$$\begin{cases} \frac{dT}{dt} = rT\left(1 - \frac{T}{K}\right) - nET, \\ \frac{dE}{dt} = \sigma + \mu TE - \eta E, \end{cases} \tag{2.2}$$

- $T(t)$ y $E(t)$ son la concentración de las células tumorales y las células efectoras en un tiempo t, respectivamente.
- r es el coeficiente para el crecimiento máximo de las células tumorales.
- K es la capacidad de carga del entorno biológico para las células tumorales.
- n es el coeficiente de velocidad al que las células efectoras dañan las células tumorales
- σ es la velocidad normal de la entrada de las células efectoras maduras en el sitio del tumor.
- μ es el coeficiente de acción de las células tumorales sobre las células efectoras.
- η es el coeficiente de la tasa de eliminación natural de las células efectoras.

Implementación del Modelo de Crecimiento Tumoral con NeuralODEs

Ahora vamos a entrar la implementación de los modelos de crecimiento tumoral, para esto nos guiamos de los ejemplos de la documentación de diffrax y equinox [4, 5]. Específicamente, investigamos:

1. ¿Puede una NDE resolver el sistema de ecuaciones en una condición inicial nueva y aleatoria?
2. ¿Puede una red más simple, como una red neuronal recurrente (RNN por sus siglas en ingles), resolver el mismo problema? ¿Es necesario usar NDEs para este propósito?

La motivación principal detrás de este estudio es evaluar la eficiencia y precisión de las NDEs y realizar una comparación con las RNNs tradicionales

Implementación del Modelo de Crecimiento Tumoral con NeuralODEs

Implementación con Datos sin Retardo

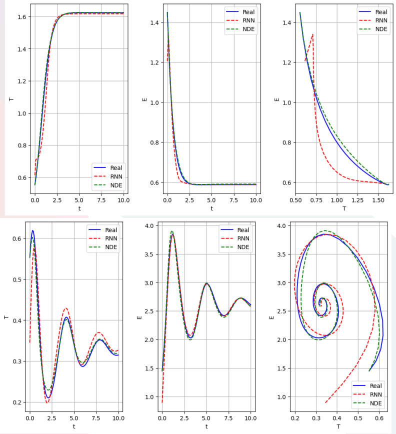
Primero realizamos la implementación para el modelo sin retardo para esto vamos a generar datos sintéticos con base en las ecuaciones del sistema usando diferentes parámetros y condiciones iniciales aleatorias, basándonos en el análisis en sistemas dinamicos que realizamos, luego con estos datos entrenamos una NeuralODE y una RNN y finalmente probamos su desempeño con un nuevo conjunto de datos.

Implementación del Modelo de Crecimiento Tumoral con NeuralODEs

Implementación con Datos sin Retardo

- Generamos los datos simulando el sistema de ecuaciones diferenciales que modela el crecimiento tumoral.
- Construimos la arquitectura de la red basándonos en el ejemplo de NeuralODE de la documentación de diffrax [4], la cual está compuesta de dos clases `Func` que es un perceptron multicapa que funciona como insumo para la clase `NeuralODE`.
  1. **Clase Func:** Define una red neuronal de tipo multi-capa (MLP) que será utilizada como la función dinámica dentro de una ODE (Ecuación Diferencial Ordinaria).
  2. **Clase NeuralODE:** Utiliza la instancia de Func para definir una ODE y resolverla usando el integrador de ecuaciones diferenciales de Diffrax.

- Entrenamos la NeuralODE utilizando el optimizador Adam [6] y un enfoque de descenso de gradiente para minimizar el error cuadrático medio (MSE) entre las predicciones del modelo y los datos reales.
- Realizamos una búsqueda exhaustiva (*grid search*) para determinar la mejor arquitectura para la NeuralODE.
- Para comparar con la NeuralODE, se implementó una RNN guiándonos con el ejemplo de RNN en la documentación de equinox [5] utilizando una arquitectura con dos capas recurrentes y 64 unidades en cada capa.



- **RNN**:
  - Error en T: 0.021728052
  - Error en E: 0.010076428
  - Error medio: 0.01590224
- **NeuralODE**:
  - Error en T: 0.005826596
  - Error en E: 0.00060446496
  - Error medio: 0.0032155307

Conclusiones

Resumen de Resultados

Los resultados de las ecuaciones diferenciales neuronales (NeuralODEs) mostraron un potencial significativo en la modelación de sistemas dinámicos complejos, como el crecimiento tumoral. Las NeuralODEs son capaces de capturar las dinámicas fundamentales del sistema mediante la parametrización de un campo vectorial con una red neuronal, permitiendo una representación continua y diferenciable de las trayectorias temporales. Esta capacidad es particularmente útil en la modelación de fenómenos biológicos, donde las variables cambian de manera continua en el tiempo y pueden ser descritas por ecuaciones diferenciales.

Conclusiones

Perspectivas Futuras

1. Ecuaciones diferenciales neuronales con delay.
2. Arquitectura de la red que permita obtener los parámetros de un modelo.
3. Utilización de NDEs para mitigar los sesgos de las ecuaciones diferenciales.

Actualmente, estamos trabajando en la redacción de un artículo de divulgación, en el cual mostraremos los beneficios y desventajas de aplicar las NDEs en sistemas de alta no linealidad

Referencias I

[1] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud.
Neural Ordinary Differential Equations.
In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[2] Patrick Kidger.
On Neural Differential Equations, February 2022.
arXiv:2202.02435 [cs, math, stat].

Referencias II

[3] Magda Galach.
Dynamics of the tumor-immune system competition - the effect of time delay.
International Journal of Applied Mathematics and Computer Science, 13(3):395–406, 2003.

[4] Patrick Kidger.
diffrax Documentation, 2024.
Accessed: 2024-05-25.

[5] Patrick Kidger.
equinox Documentation, 2024.
Accessed: 2024-05-25.

Referencias III

[6] Diederik P. Kingma and Jimmy Ba.
Adam: A Method for Stochastic Optimization, January 2017.
arXiv:1412.6980 [cs].