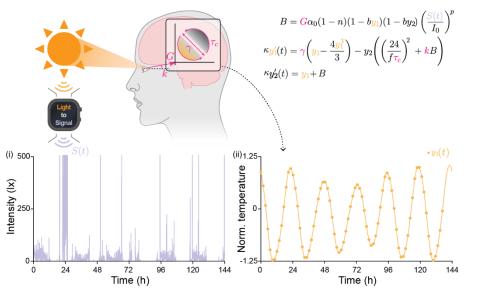
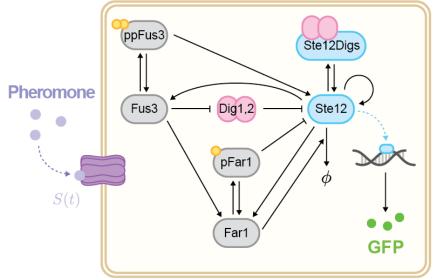
Physics-Informed Neural Networks (PINNs) Fitting a Mathematical Model to Real Data





Korea U. Sejong campus
Division of applied math
Hyeontae Jo
04. April. 2025.



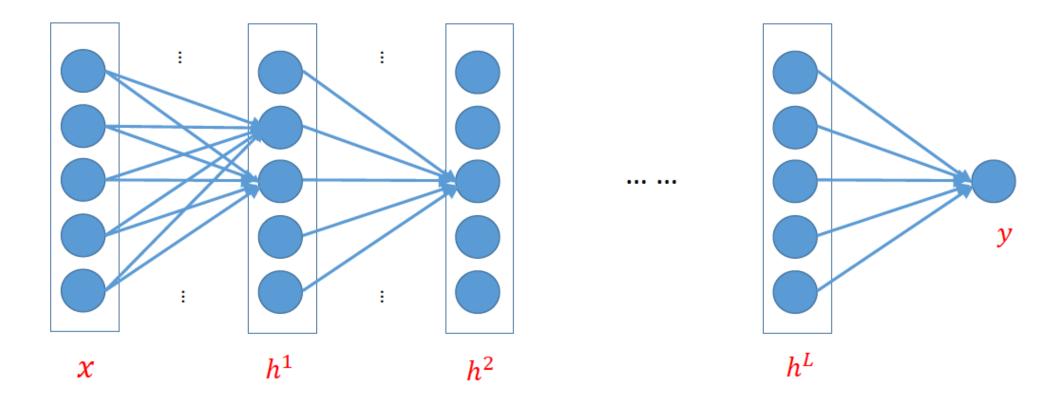


Part 1: Introduction to Neural Networks





Artificial Neural Networks (ANNs) have been utilized for finding a mathematical relationship between input data x and output data y



Typical way to explain the structure of ANNs

How do we build an ANN step by step?

Mathematical description of the NN $(x \in \mathbb{R}, y \in \mathbb{R})$!

←

two variables are a single pair of real numbers

Linear model (transformation)

$$y = w_1 x + b_1$$

With unknown parameters $p = (w_1, b_1)$

NN description (Layer/Node Representation): Each node represents a neuron, and each layer represents a transformation (usually matrix multiplication + activation)



Input layer

Output layer

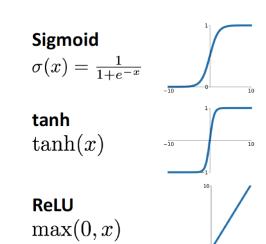
However, our dataset does not have to follow such linear relationship

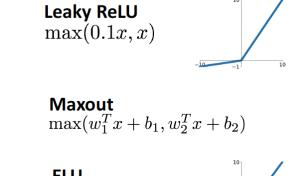
Mathematical description of the NN $(x \in \mathbb{R}, y \in \mathbb{R})$!

To make nonlinearity,...

NN with 3-layers

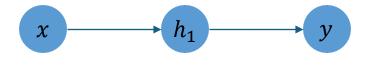
$$z_1 = w_1 x + b_1$$
$$h_1 = \sigma(z_1)$$
$$y = w_2 h_1 + b_2$$





With unknown parameters $p = (w_1, b_1, w_2, b_2)$

NN description (Layer/Node Representation): Each node represents a neuron, and each layer represents a transformation (usually matrix multiplication + activation)



Input layer

Hidden layer

Output layer

Mathematical description of the NN $(x \in \mathbb{R}, y \in \mathbb{R})$!

$$z_1 = w_1 x + b_1$$

$$h_1 = \sigma(z_1)$$

$$z_2 = w_2 h_1 + b_2$$

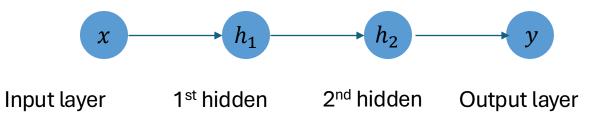
$$h_2 = \sigma(z_2)$$

$$y = w_3 h_2 + b_3$$

NN with 4-layers

With unknown parameters $p = (w_1, b_1, w_2, b_2, w_3, b_3)$

NN description (Layer/Node Representation): Each node represents a neuron, and each layer represents a transformation (usually matrix multiplication + activation)



We can also add a lot of nodes in each hidden layer!

NN with two nodes in a single hidden layer

$$z_1 = w_{11}x + b_{11} \to h_1(z_1) = \sigma(z_1)$$

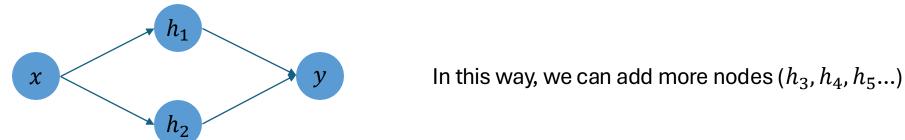
$$x$$

$$y = w_{21}h_1 + w_{22}h_2 + b_2$$

$$z_2 = w_{12}x + b_{12} \to h_2(z_2) = \sigma(z_2)$$

With unknown parameters $p = (w_{11}, w_{12}, b_{11}, b_{12}, w_{21}, w_{22}, b_2)$

NN description (Layer/Node Representation): Each node represents a neuron, and each layer represents a transformation (usually matrix multiplication + activation)



How can we measure the performance of a neural network to determine whether it produces the correct output?

Once weights and biases are given, we can evaluate the L2 errors -> Typical curve-fit procedure

NN with two nodes in a single hidden layer

 χ

$$z_1 = w_{11}x + b_{11}, \qquad h_1(z_1) = \sigma(z_1)$$

$$z_2 = w_{12}x + b_{12}, \qquad h_2(z_2) = \sigma(z_2)$$

$$y = w_{21}h_1 + w_{22}h_2 + b_2 = f(x; p)$$

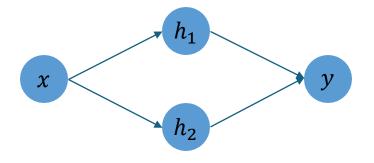
Dataset: $\{(x_i, y_i) | i = 1, ..., n\}$

L2 error $E = \sum_{i=1}^{n} |f(x_i) - y_i|^2$

: The difference between outputs of the NN and data y_i

: a function of p (i.e., E = E(p))

Can you find the optimal p that minimizes the L2 error?



One famous way: gradient descent

 p_0 is chosen randomly,

We can find a sequence of set of parameters $\{p_n\}$ via the following rule: $p_{n+1}=p_n-\delta\nabla_p E(p_n)$

Calculating $\nabla_p E(p)$ would be difficult or complicate..

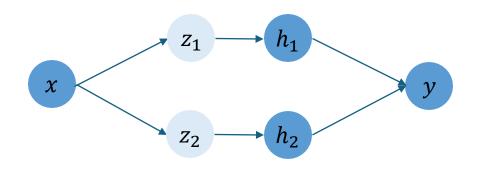
Let's consider forward propagation

NN with two nodes in a single hidden layer

$$x$$
 $z_1 = w_{11}x + b_{11}, \qquad h_1(z_1) = \sigma(z_1)$

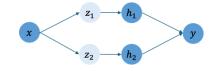
$$z_2 = w_{12}x + b_{12}, \qquad h_2(z_2) = \sigma(z_2)$$

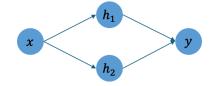
$$y = w_{21}h_1 + w_{22}h_2 + b_2$$



p = (1,1,1,1,1,1),
$$\sigma(x) = x^2$$

{ (x_i, y_i) } = { $(1,2)$ }





Feature	Computational Graph	Neural Network Representation
Node meaning	Variables or operations	Artificial neurons
Edge meaning	Flow of computation	Weighted connections
Purpose	Compute gradients (e.g., for training)	Define network structure and forward computation
Granularity	Fine-grained (add, multiply, etc.)	Coarse-grained (layers of neurons)
Interpretation	Math-level computation	Learning model structure

Computational graph: A graph of operations that represent how computations are done.

Let's consider forward propagation

 $z_2 = 2$

NN with two nodes in a single hidden layer

 $h_2 = 4$

p = (1,1,1,1,1,1),
$$\sigma(x) = x^2$$

{ (x_i, y_i) } = { $(1,2)$ }

Computational graph: A graph of operations that represent how computations are done.

NN with two nodes in a single hidden layer

$$x$$

$$z_1 = w_{11}x + b_{11}, h_1(z_1) = z_1^2$$

$$z_2 = w_{12}x + b_{12}, h_2(z_2) = z_2^2$$

$$y = w_{21}h_1 + w_{22}h_2 + b_2$$

$$p = (1,1,1,1,1,1), \ \sigma(x) = x^{2}$$

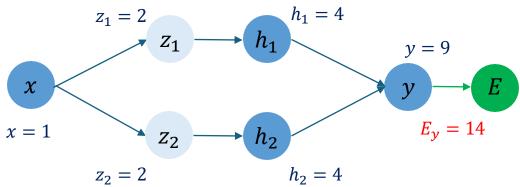
$$\{(x_{i}, y_{i})\} = \{(1,2)\}$$

$$E(p) = (f(1; p) - 2)^{2}$$

$$= (y - 2)^{2}$$

$$\partial_{w_{11}} E(p) = \partial_{w_{11}} (y - 2)^{2} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial w_{11}} \quad \text{Chain Rule}$$

$$= \frac{\partial E}{\partial y} * \frac{\partial y}{\partial w_{11}} = 2(y - 2) * \frac{\partial y}{\partial w_{11}} = 14 * \frac{\partial y}{\partial w_{11}}$$



We can directly calculate $\frac{\partial E}{\partial y}$ via forward propagation

NN with two nodes in a single hidden layer

$$p = (1,1,1,1,1,1), \ \sigma(x) = x^{2}$$

$$\{(x_{i},y_{i})\} = \{(1,2)\}$$

$$E(p) = (f(1;p)-2)^{2}$$

$$z_{2} = w_{12}x + b_{12}, \quad h_{2}(z_{2}) = z_{2}^{2}$$

$$y = w_{21}h_{1} + w_{22}h_{2} + b_{2}$$

$$d_{w_{11}}E(p) = \partial_{w_{11}}(y-2)^{2} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial w_{11}} \quad \text{Chain Rule}$$

$$= \frac{\partial E}{\partial y} * \frac{\partial y}{\partial w_{11}} = 2(y-2) * \frac{\partial y}{\partial w_{11}} = 14 * \frac{\partial y}{\partial w_{11}}$$

$$z_{1} = 2$$

$$x = 1$$

$$z_{2} = 2$$

$$h_{2} = 4$$

$$d_{w_{11}}(y-2) = \frac{\partial E}{\partial y} \frac{\partial F}{\partial w_{11}} = 14 * \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 1 * \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

$$d_{w_{11}}(y-2) = \frac{\partial F}{\partial y} \frac{\partial F}{\partial w_{11}} = 0$$

NN with two nodes in a single hidden layer

$$p = (1,1,1,1,1,1), \sigma(x) = x^{2}$$

$$\{(x_{i}, y_{i})\} = \{(1,2)\}$$

$$z_{1} = w_{11}x + b_{11}, \quad h_{1}(z_{1}) = z_{1}^{2}$$

$$E(p) = (f(1; p) - 2)^{2}$$

$$z_{2} = w_{12}x + b_{12}, \quad h_{2}(z_{2}) = z_{2}^{2}$$

$$y = w_{21}h_{1} + w_{22}h_{2} + b_{2}$$

$$y = w_{21}h_{1} + w_{22}h_{2} + b_{2}$$

$$z_{1} = 2 \qquad h_{1z_{1}} = 4 \qquad h_{1} = 4$$

$$z_{1} = 1 \qquad z_{1} \qquad h_{1} = 4 \qquad y_{h_{1}} = 1 \qquad y = 9$$

$$z_{2} = 2 \qquad h_{2} = 2 \qquad h_{2} = 4$$

$$z_{2} = 2 \qquad h_{2} = 4$$

$$z_{3} = 2 \qquad h_{2} = 4$$

$$z_{2} = 2 \qquad h_{2} = 4$$

$$z_{3} = 2 \qquad h_{2} = 4$$

$$z_{3} = 2 \qquad h_{2} = 4$$

$$z_{4} = 4 \qquad \frac{\partial y}{\partial w_{11}} = \frac{\partial y}{\partial h_{1}} \frac{\partial h_{1}}{\partial w_{11}} + \frac{\partial y}{\partial h_{2}} \frac{\partial h_{2}}{\partial w_{11}} = w_{21} * \frac{\partial h_{1}}{\partial w_{11}} = 1 * \frac{\partial h_{1}}{\partial w_{11}}$$

$$\frac{\partial h_{1}}{\partial w_{11}} = \frac{\partial h_{1}}{\partial z_{1}} * \frac{\partial z_{1}}{\partial z_{1}} * \frac{\partial z_{1}}{\partial z_{1}} = 2z_{1} * x_{1} = 4 * 1$$

NN with two nodes in a single hidden layer

$$p = (1,1,1,1,1,1), \ \sigma(x) = x^2$$

$$\{(x_l, y_l)\} = \{(1,2)\}$$

$$E(p) = (f(1;p) - 2)^2$$

$$z_2 = w_{12}x + b_{12}, \quad h_2(z_2) = z_2^2$$

$$y = w_{21}h_1 + w_{22}h_2 + b_2$$

$$\partial_{w_{11}}E(p) = \partial_{w_{11}}(y - 2)^2 = \frac{\partial E}{\partial y} \frac{\partial y}{\partial w_{11}}$$

$$Chain Rule$$

$$= \frac{\partial E}{\partial y} * \frac{\partial y}{\partial w_{11}} = 2(y - 2) * \frac{\partial y}{\partial w_{11}} = 14 * \frac{\partial y}{\partial w_{11}}$$

$$x = 1$$

$$z_2 = 2$$

$$h_2 = 4$$

$$\partial_{w_{11}}E(p) = \frac{\partial E}{\partial y} * \frac{\partial f(1,p)}{\partial h_1} * \frac{\partial E}{\partial h_1} * \frac{\partial E}{$$

Backpropagation: Separating one heavy task to several number of simple tasks via chain rule

NN with two nodes in a single hidden layer

$$p = (1,1,1,1,1,1), \ \sigma(x) = x^2$$

$$\{(x_i,y_i)\} = \{(1,2)\}$$

$$E(p) = (f(1;p)-2)^2$$

$$= (y-2)^2$$

$$y = w_{21}h_1 + w_{22}h_2 + b_2$$

$$\partial_{w_{11}}E(p) = \partial_{w_{11}}(y-2)^2 = \frac{\partial E}{\partial y} \frac{\partial y}{\partial w_{11}} \text{ Chain Rule }$$

$$= \frac{\partial E}{\partial y} * \frac{\partial y}{\partial w_{11}} = 2(y-2) * \frac{\partial y}{\partial w_{11}} = 14 * \frac{\partial y}{\partial w_{11}}$$

$$x = 1$$

$$x = 1$$

$$x = 1$$

$$x = 2$$

$$x = 2$$

$$x = 2$$

$$x = 2$$

$$x = 3$$

$$x = 4$$

$$x = 1$$

$$x = 2$$

$$x = 4$$

$$x = 1$$

$$x = 2$$

$$x$$

Backpropagation: Separating one heavy task to several number of simple tasks via chain rule

Later, we will design NN using python! → Colab

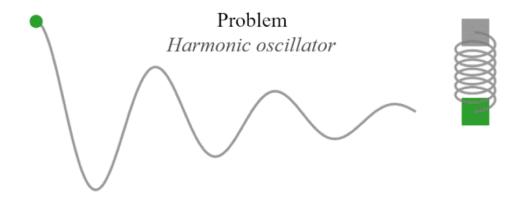
```
import numpy as np
     from numpy random import randn
     N, Din, H, Dout = 64, 1000, 100, 10
     x, y = randn(N, Din), randn(N, Dout)
                                                     Initialize weights and bias
     w1, w2 = randn(Din, H), randn(H, Dout)
     for t in range(10000):
       h = 1.0 / (1.0 + np.exp(-x.dot(w1)))
                                                     Compute L2 loss E = \sum_{i=1}^{n} |f(x_i) - y_i|^2
      y_pred = h.dot(w2)
      loss = np.square(y_pred - y).sum()
10
11
      dy_pred = 2.0 * (y_pred - y)
12
      dw2 = h.T.dot(dy_pred)
                                                     Compute gradients
       dh = dy_pred.dot(w2.T)
13
       dw1 = x.T.dot(dh * h * (1 - h))
14
      w1 = 1e-4 * dw1
15
                                                     Gradient descent w \leftarrow w - \frac{\partial}{\partial w} L
16
      w2 = 1e-4 * dw2
```

Part 2: Application of NNs to Differential Eqns





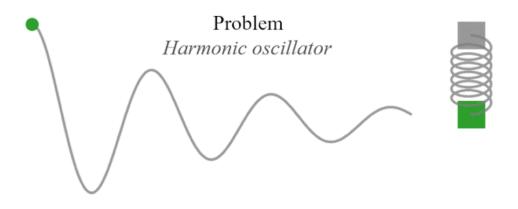
Why PINN?



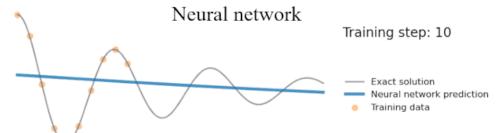
$$m\frac{d^2u}{dx^2} + \mu\frac{du}{dx} + ku = 0$$

Data obtained from Scientific phenomena can be described by a physics law (~Nonlinear)

Conventional neural networks can effectively fit observed data



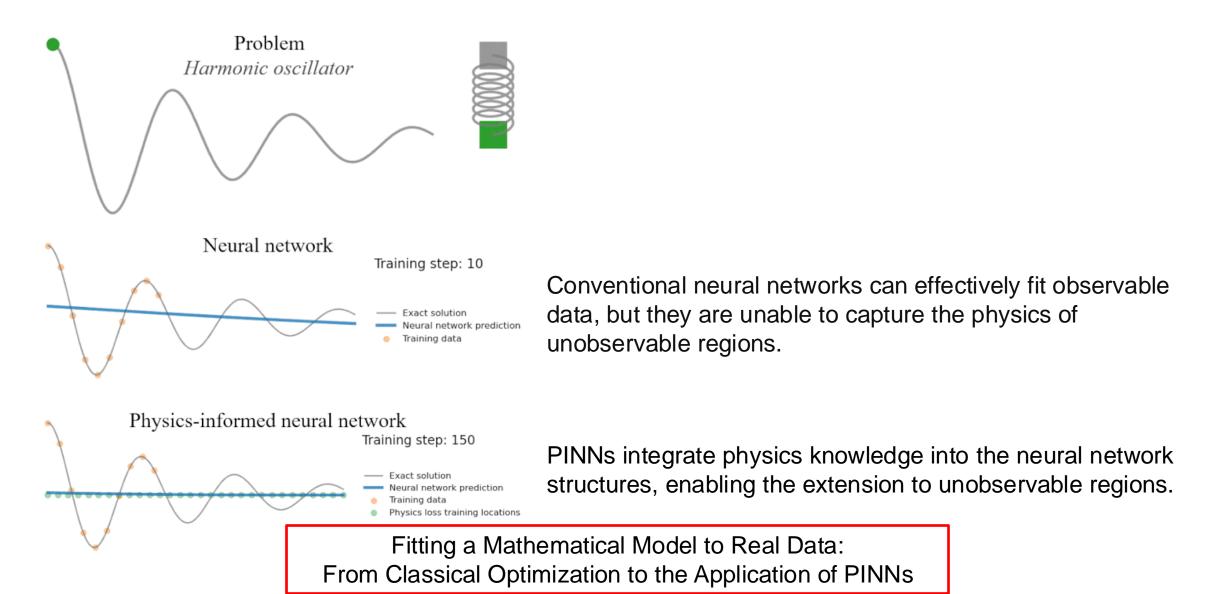
$$m\frac{d^2u}{dx^2} + \mu\frac{du}{dx} + ku = 0$$



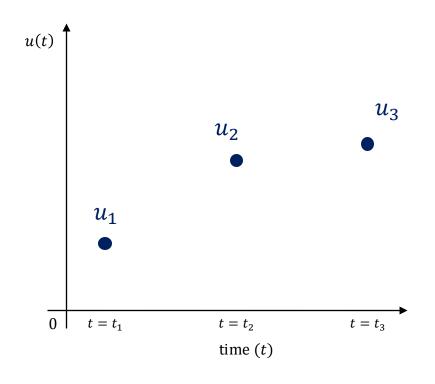
But they are unable to capture the physics of unobservable regions.

$$\min \frac{1}{N} \sum_{i}^{N} (u_{\text{NN}}(x_i; \theta) - u_{\text{true}}(x_i))^2$$

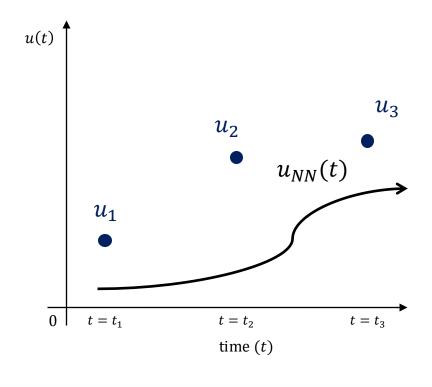
Physics-informed neural networks can be used to data-driven science Scientific discovery



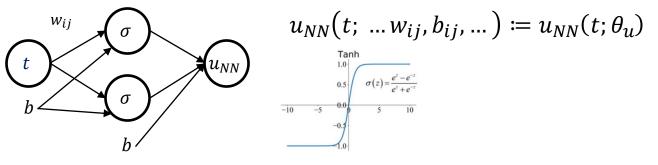
Given data $\{t_i, u_i\}_{i=1}^3$ and differential equation u'(t) = au(t) + b

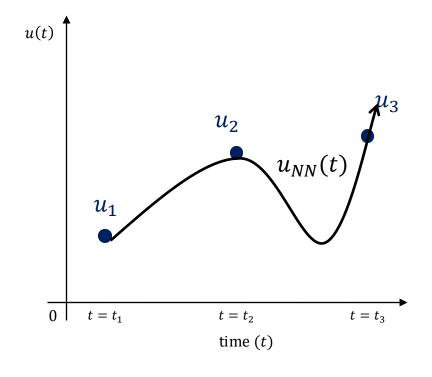


NN: A function with linear transform and nonlinear activation functions

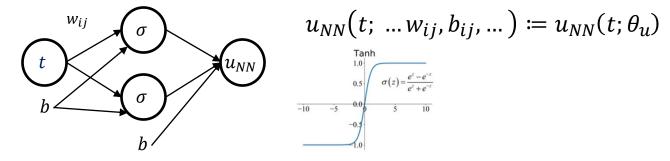


Step 1: Construct an artificial Neural Network (NN), $u_{NN}(t;\theta_u)$



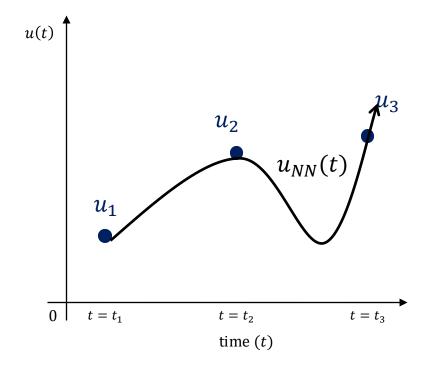


Step 1: Construct an artificial Neural Network (NN), $u_{NN}(t;\theta_u)$

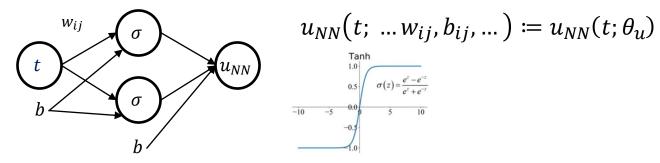


Step 2: Fit the NN that passes through all observation points

(E2 Loss)
$$L_{data}(\theta_u) \coloneqq \sum_{i=1}^3 |u_i - u_{NN}(t_i; \theta_u)|^2 \to 0$$



Step 1: Construct an artificial Neural Network (NN), $u_{NN}(t;\theta_u)$



Step 2: Fit the NN that passes through all observation points

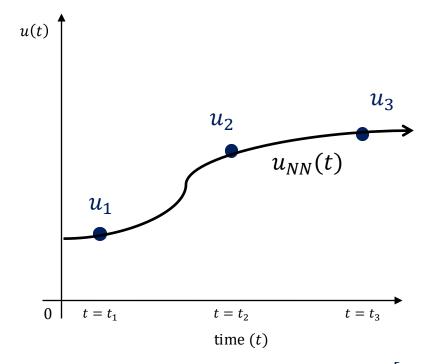
(E2 Loss)
$$L_{data}(\theta_u) \coloneqq \sum_{i=1}^{3} |u_i - u_{NN}(t_i; \theta_u)|^2 \to 0$$

However,

$$u'_{NN}(t) \neq au_{NN}(t) + b$$

We add constraints to u_{NN} to fit both data points and the model simultaneously.

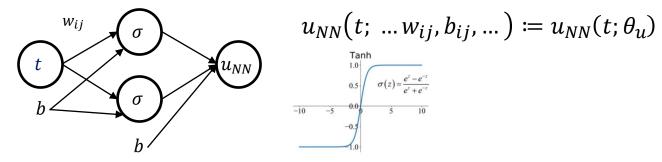
Given data $\{t_i, u_i\}_{i=1}^3$ and differential equation u'(t) = au(t) + b



M number of collocation time points $t_j \in [0, T]$ Sample from [0,T] with equal distance

$$(dt = t_j - t_{j-1})$$
 or,
$$t_j \sim Unif(0, T)$$

Step 1: Construct an artificial Neural Network (NN), $u_{NN}(t; \theta_u)$



Step 2: Fit the NN that passes through all observation points

(E2 Loss)
$$L_{data}(\theta_u) \coloneqq \sum_{i=1}^3 |u_i - u_{NN}(t_i; \theta_u)|^2 \to 0$$

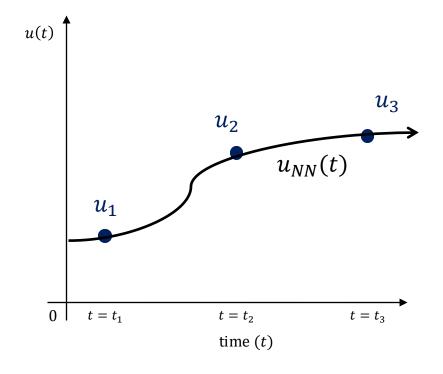
Step 3: Input the NN to the equation

(Regularization)
$$L_{reg}(\theta_u) \coloneqq \sum_{j=1}^M \left| u_{NN}'(t_j;\theta_u) - \left(au_{NN}(t_j;\theta_u) + b\right) \right|^2 \to 0$$
$$u_{NN} \text{ is 'differentiable' with respect to } t$$

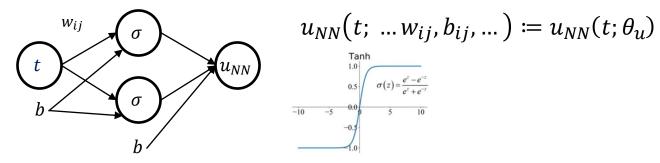
$$u'_{NN}(t) \approx a u_{NN}(t) + b$$

We add constraints to u_{NN} to fit both data points and the model simultaneously.

Given data $\{t_i, u_i\}_{i=1}^3$ and differential equation u'(t) = au(t) + b



Step 1: Construct an artificial Neural Network (NN), $u_{NN}(t;\theta_u)$



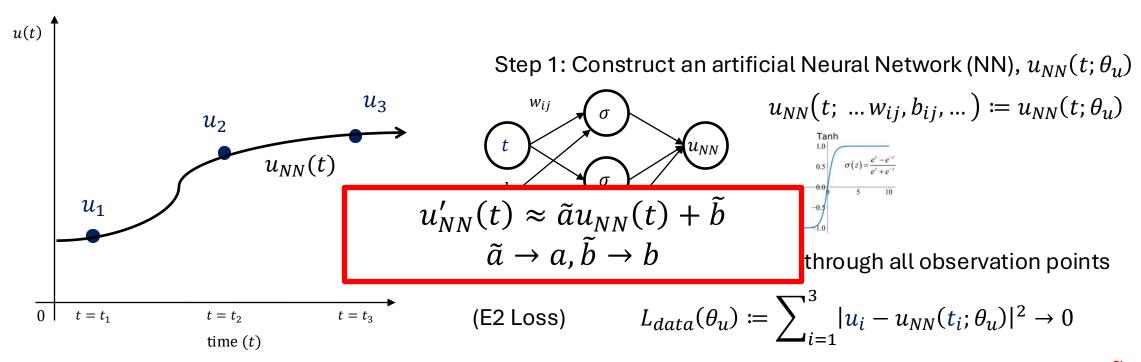
Step 2: Fit the NN that passes through all observation points

(E2 Loss)
$$L_{data}(\theta_u) \coloneqq \sum_{i=1}^{3} |u_i - u_{NN}(t_i; \theta_u)|^2 \to 0$$

Step 3: Input the NN to the equation with initial guesses \tilde{a} and \tilde{b}

(Regularization)
$$L_{reg}(\theta_u; \tilde{\mathbf{a}}, \tilde{\mathbf{b}}) \coloneqq \sum_{j=1}^{M} \left| u'_{NN}(t_j; \theta_u) - \left(\tilde{\mathbf{a}} u_{NN}(t_j; \theta_u) + \tilde{\mathbf{b}} \right) \right|^2 \to 0$$

We add constraints to u_{NN} to fit both data points and the model simultaneously.



Step 3: Input the NN to the equation with initial guesses \tilde{a} and \tilde{b}

(Regularization)
$$L_{reg}(\theta_u; \tilde{\mathbf{a}}, \tilde{\mathbf{b}}) \coloneqq \sum_{j=1}^{M} \left| u'_{NN}(t_j; \theta_u) - \left(\tilde{\mathbf{a}} u_{NN}(t_j; \theta_u) + \tilde{\mathbf{b}} \right) \right|^2 \to 0$$

Thank you!

Let's jump into colab implementation

Korea_htj@korea.ac.kr





Appendix: Application of PINNs...



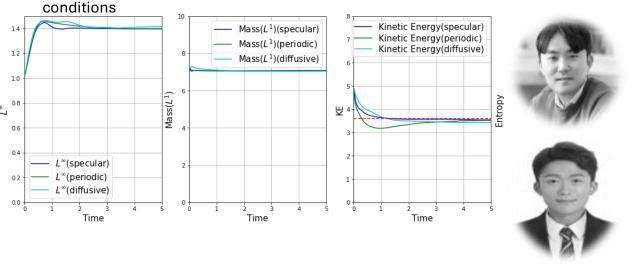


Real-world examples 1: Solving DEs

Solution of Fokker-Planck equation

$$\partial_t u + v \partial_x u = \partial_v (\partial_v u + v u)$$

Physical quantities from u_{NN} with different boundary

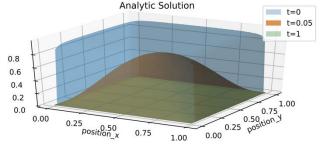


Jin Woo Jang POSTECH

Jae Young Lee Joongang University

Solutions & Parameters of PDE

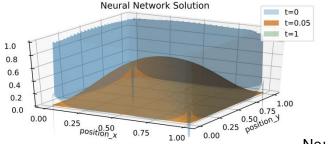
$$\partial_t u - D(\partial_{xx} u - \partial_{yy} u) = 0$$





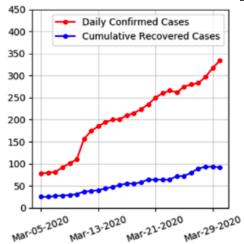
Hwijae Son Konkuk University

Eunheui Kim California State University Long Beach



Fokker-Planck equation via the neural network approach (J. Comput. Phys.)

Real-world examples 2: Time-varying parameters in infectious diseases model.



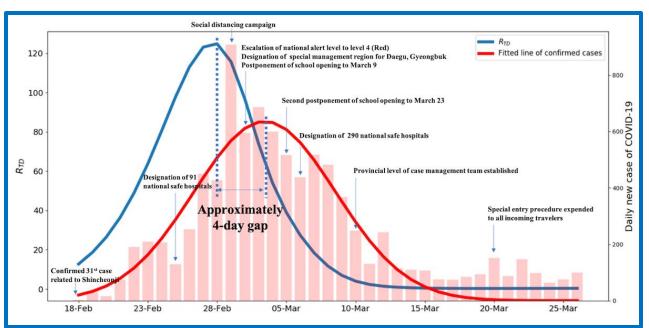
COVID19 report

I(t): The number of daily confirmed cases

R(t): The cumulative number of recovered or deceased cases

(from Korea Disease Control and Prevention Agency)

$$R_{TD}(t) \approx \frac{\beta_{NN}(t)}{\gamma_{NN}(t)}$$



us diseases model.
$$\frac{dS}{dt} = -\beta(t)SI$$

$$\frac{dI}{dt} = \beta(t)SI - \gamma I \quad \beta(t) \leftarrow \text{NN } \beta_{NN}(t)$$

$$\gamma(t) \leftarrow \text{NN } \gamma_{NN}(t)$$
 ses
$$\frac{dR}{dt} = \gamma(t)I$$

Hyung Ju Hwang POSTECH



Se Young Jung Seoul National University Bundang Hospita



Hwijae Son Konkuk University

Real-world examples 3: Density-PINN to infer the cell signaling dynamics