A GENERAL APPROACH FOR LEARNING CONSTITUTIVE RELATIONSHIPS (PHYSICAL LAWS) FROM NEURAL NETWORKS: A SEVEN STEP PROCESSING FOR SOLVING ANY PROBLEM

ANDREW TEMPLE, MATTHEW MILLER, SUNIL DHAPOLA, AARON STEBNER,
PETER COLLINS, AND BRANDEN KAPPES

1. Methods

Fully dense neural network (NN) architectures, such as the one shown in Figure 1, perform a sequence of affine transformations, $\mathbf{z}_i \leftarrow \boldsymbol{\theta}_i \mathbf{x}^{(i)}$, followed by element-wise functional operations, $\sigma(\mathbf{z}_i)$ to introduce non-linearity at each layer; that is, each layer stretches and distorts the underlying space.

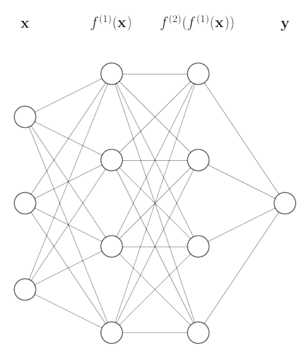

Figure 1. Schematic view of a fully dense neural network. Each sequence of affine and non-linear transformations are captured in the function, $f_i(\mathbf{x}) : \mathbf{x}^{(i+1)} \leftarrow \sigma(\boldsymbol{\theta}_i \mathbf{x}^{(i)})$

The resulting network,

$$f(x) = \sigma(\boldsymbol{\theta}_n \sigma(\boldsymbol{\theta}_{n-1} \sigma(\ldots \boldsymbol{\theta}_2 \sigma(\boldsymbol{\theta}_1 \mathbf{x})))) \tag{1}$$

is an arbitrary function generator, but at present, the network weights $\boldsymbol{\theta}_i$ can not map back to analytic forms that capture and describe the underlying physics. There are, however, many such mappings through polynomial series expansions,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{2}$$

We hypothesize that the physics of a process can be extracted by fitting the polynomial expansions of known physical relationships to the polynomial coefficients of a polynomial series expansion of Equation (1).

Although ReLU (rectified linear units) have become a more common activation function, its discontinuity at $x = 0$ requires an infinite series to fully capture the behavior at this transition. However, the sigmoid function,

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

is a special case of the generating function for the Euler polynomial coefficients,

$$\frac{2e^{xt}}{1 + e^t} = \sum_{n=0}^{\infty} E_n(x) \frac{t^n}{n!} \tag{4}$$

where, for $x = 0$,

$$\sigma(x) = \frac{1}{2} \sum_{n=0}^{\infty} E_n(0) \frac{(-1)^n}{n!}. \tag{5}$$

The Euler polynomials at $x = 0$,

$$E_n(0) = -2(n+1)^{-1} \left(2^{n+1} - 1\right) B_{n+1} \tag{6}$$

where B_n is the n^{th} Bernoulli number. Since Bernoulli numbers of odd index, with the exception of B_1, are zero, $E_i(0) = 0$ for $i = 2, 4, 6, \ldots, 2n$. Therefore, the summand and limits of Equation (5) change to

$$\sigma(x) = \frac{1}{2} - \frac{1}{2} \sum_{n=1}^{\infty} \left(\frac{E_{2n-1(0)}}{(2n-1)!} \right) x^{2n-1}. \tag{7}$$

The series representation of $E_{2n-1}(x)$

$$E_{2n-1}(x) = \frac{(-1)^n 4(2n-1)!}{\pi^{2n+1}} \sum_{k=0}^{\infty} \frac{\cos[(2k+1)\pi x]}{(2k+1)^{2n}} \tag{8}$$

such that,

$$E_{2n-1}(0) = \frac{(-1)^n 4(2n-1)!}{\pi^{2n+1}} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^{2n}} \tag{9}$$

and therefore,

$$\begin{aligned}
\sigma(x) &= \frac{1}{2} - \sum_{n=1}^{\infty} 2\frac{(-1)^n}{\pi^{2n}} \left(\sum_{k=0}^{\infty} \frac{1}{(2k+1)^{2n}} \right) x^{2n-1} && (10)\\
&= \frac{1}{2} - \sum_{n=1}^{\infty} 2\frac{(-1)^n}{\pi^{2n}} \left(4^{-n}\left(4^n - 1\right)\zeta(2n)\right) x^{2n-1} \\
&= \frac{1}{2} - \sum_{n=1}^{\infty} \underbrace{2\left(\frac{-1}{4\pi^2}\right)^n \left(4^n - 1\right)\zeta(2n)}_{a_n} x^{2n-1} \\
&= \sum_{n=0}^{\infty} a_n x^n,\ a_n = \begin{cases} 1/2 & n = 0 \\ -2\left(\frac{-1}{4\pi^2}\right)^{(n+1)/2}\left(4^{(n+1)/2} - 1\right)\zeta(n+1) & n \text{ odd} \\ 0 & n \text{ even} \end{cases} && (11)
\end{aligned}$$

2. Discussion

Many non-trivial problems in materials science, and in science more broadly, are explained not through a single constitutive relationship, but through a superposition of contributing physics.

Figure **??** shows an artificial dataset constructed to replicate the impact of yield stress in a two-phase, solid-solution strengthened alloy system. Using a combination of composite theory for the contribution of flow stress, NAME solid solution [?], and Hall-Petch [?] strengthening, the expected yield stress is

$$\sigma_y = F_v^A \sigma_f^A + F_v^B \sigma_f^B + \sum_i C_i [x_i]^{2/3} + \sum_j k_j d_j^{-1/2} + \dots \quad (12)$$

with free parameters

F_v^i	Volume fraction of phase i
$[x_i]$	Concentration of solute i
d_j	Average grain diameter of phase j

and fixed parameters

σ_f^i	Flow stress of phase i
C_i	Solid solution strengthening coefficient for solute species i
k_j	Hall-Petch strengthening coefficient for phase j

The goal is to iteratively improve on this constitutive model one term at a time, and monitor the effect on the residuals between the predicted yield, $\hat{\sigma_y}$ and the actual yield σ_y.

The analytical form, combining Equations (1) and (11), the estimated output of a two-layer NN can be written as

$$
\begin{aligned}
\mathbf{y}_1 &= \sum_{k=0}^{\infty} a_k (\boldsymbol{\theta}_1^T \mathbf{x})^k \\
\mathbf{y}_2 &= \sum_{k=0}^{\infty} b_k (\boldsymbol{\theta}_2^T \mathbf{y}_1)^k \\
&= b_0 \mathbf{1} + \\
&\quad + b_1(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_2 + (\tilde{a}_3 + (\ldots)\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}}) \\
&\quad + b_2(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_2 + (\tilde{a}_3 + (\ldots)\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}})^2 \\
&\quad \vdots \\
&\quad + b_k(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_2 + (\tilde{a}_3 + (\ldots)\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}})\tilde{\mathbf{x}})^k \\
&\quad \vdots
\end{aligned} \tag{13}
$$

where $\tilde{a}_i = \boldsymbol{\theta}_2^T a_i$ and $\tilde{\mathbf{x}} = \boldsymbol{\theta}_1^T \mathbf{x}$. All $\boldsymbol{\theta}_i$, $\mathbf{x}$, and $\mathbf{y}$ are augmented to include the bias, $\mathbf{b}_i$, that is,

$$
\mathbf{x} : \mathbf{x} \leftarrow \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 1 \\ x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{14}
$$

$$
\mathbf{y} : \mathbf{y} \leftarrow \begin{pmatrix} 1 \\ \mathbf{y} \end{pmatrix} \tag{15}
$$

$$
\boldsymbol{\theta}_i : \boldsymbol{\theta}_i \leftarrow \begin{pmatrix} \mathbf{b}_i & \boldsymbol{\theta}_i \end{pmatrix} \tag{16}
$$

However from Equation (11), $a_i = 0$ for $i = 2, 4, 6, \ldots$, and therefore,

$$
\begin{aligned}
\mathbf{y}_1 &= \sum_{k=0}^{\infty} a_k (\boldsymbol{\theta}_1^T \mathbf{x})^k \\
\mathbf{y}_2 &= \sum_{k=0}^{\infty} b_k (\boldsymbol{\theta}_2^T \mathbf{y}_1)^k \\
&= b_0 \mathbf{1} + \\
&\quad + b_1(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_3 + (\tilde{a}_5 + (\ldots)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}) \\
&\quad + b_2(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_3 + (\tilde{a}_5 + (\ldots)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}})^2 \\
&\quad \vdots \\
&\quad + b_k(\tilde{a}_0 + (\tilde{a}_1 + (\tilde{a}_3 + (\tilde{a}_5 + (\ldots)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}}^2)\tilde{\mathbf{x}})^k \\
&\quad \vdots
\end{aligned} \tag{17}
$$

where the $\tilde{\mathbf{x}}^2 = \tilde{\mathbf{x}} \odot \tilde{\mathbf{x}}$, the Hadamard product of $\tilde{\mathbf{x}}$ with itself.

$$\begin{aligned}
\mathbf{y}_2 &= \sum_{N=0}^{\infty}\sum_{k=0}^{N}\sum_{l=0}^{k}\sum_{m=0}^{l}\dots b_N \binom{N}{k,l,m,\dots} \tilde{a}_0^k \tilde{a}_1^l \tilde{a}_3^m \dots \tilde{\mathbf{x}}^{N-k\dots}(\tilde{\mathbf{x}}^2)^{N-k-l\dots}(\tilde{\mathbf{x}}^2)^{N-k-l-m\dots} \\
&= \sum_{N=0}^{\infty}\sum_{k=0}^{N}\sum_{l=0}^{k}\sum_{m=0}^{l}\dots b_N \binom{N}{k,l,m,\dots} \tilde{a}_0^k \tilde{a}_1^l \tilde{a}_3^m \dots \tilde{\mathbf{x}}^{l+m+n+\dots}(\tilde{\mathbf{x}}^2)^{m+n+\dots}(\tilde{\mathbf{x}}^2)^{n+\dots} \qquad (18)
\end{aligned}$$

where $k + l + m + n + \ldots = N$. Collecting coefficients and terms of power k,

$$\mathbf{y}_2 = \sum_{k=0}^{\infty} c_k \tilde{\mathbf{x}}^k$$

that, having the same form as Equation (11) creates a sequential process for determining the coefficients of the power series expansion of each layer in an ANN. Importantly, the output layer in a ANN regression is a single node with a linear activation, so the final layer, y_f, working from the last hidden layer, $\mathbf{y}_n$, is simply,

$$y_f = \boldsymbol{\theta}_n^T \mathbf{y}_n \qquad (19)$$

Together, this leads to a seven-step process for systematically and incrementally extracting physics information from an ANN:

(1) Collect data–features and targets–for which relationships are expected to exist.
(2) Design and train a fully dense multi-layer perceptron network (ANN).
(3) Build a power series expansion from the architecture of this ANN, using Equations (11) and (18) to populate the coefficients using the trained weights from the neural network.
(4) Hypothesize a constitutive relationship between the feature space and the target space.
(5) Recast the terms in the hypothesis function from #4 as power series expansions, creating power series coefficient generating functions that are functions of the constitutive model fitting parameters. An example of this process is provided below, and a table of select power series expansions relevant to materials research are provided in Table (**??**).
(6) Perform an optimization, *e.g.* least squares, fit to find the fitting parameters from #5
(7) Calculate the residuals of the ANN power series expansion coefficient vector, and from this residual vector, the error in the model. If the accuracy is sufficient for the application, stop; otherwise, expand the constitutive relationship from step #4 and repeat.

TABLE 1. Examples of coefficient generating functions for functional forms commonly found in materials physics.

k	Ca^x	Cx^n	$Ce^{-\beta x}$	$Cx^{-1/2}$
0	1	–	C	C
1	$C \ln a$	–	$-\beta C$	$-\frac{1}{2}C$
2	$\frac{(\ln a)^2}{2}C$	–	$\frac{\beta^2}{2}C$	$\frac{3}{8}C$
$\vdots$			$\vdots$	
n	$\frac{(\ln a)^n}{n!}C$	$\begin{cases} C & \text{if } k = n \\ 0 & \text{otherwise} \end{cases}$	$(-1)^n \frac{\beta^n}{n!}C$	$C \prod_{i=1}^{n} (-1) \frac{2i-1}{2i}$