Overview

Recurrent Neural Networks (RNNs) are a kind of Discrete-Time (DT) systems. Despite remarkable efforts to analyze their behavior from the dynamical systems perspective, just recently a dynamical systems approach was used to design RNN architectures. In Legendre Memory Units (LMUs), the state equation of a DT Linear Time-Invariant (DTLTI) system serves as a memory update policy by orthogonal time projections. Likewise, in Bistable Recurrent Cells (BRCs), a nonlinearity is used to force bistable behavior in the cell, increasing their memory capabilities mimicking biological neurons. So, the close relationship between RNNs and DT systems can be exploited not only for analysis but also for design.

\[
x_k = f(k, x_{k-1}, u_{k-1}) \iff x_k = \sigma(u_k, y_k, x_{k-1})
\]

\[
y_k = g(k, x_k, u_k) \iff y_k = \gamma(h_k)
\]

Relevant Papers


Proposal

Our work exploits a complete state-space representation of a DTLTI system inside an RNN. We employ Laguerre functions, a DTLTI system orthogonal basis, to construct a state update policy and the output equation to perform memory computation thus enhancing RNN expressivity and generalizability. We also introduce a Ladder network, a special case of Laguerre networks that introduces delay dynamics inside the RNN.

State and Memory Update

\[
x_k = Ax_{k-1} + B(u_k + m_{k-1})
\]

\[
m_k = Cx_k
\]

Laguerre Network

The state-space representation is constructed from Laguerre polynomials

\[
\ell_i(t) = \left[\frac{(2p)^i}{(i-1)!} t^i e^{-\alpha t} \right]
\]

Laguerre polynomials can be computed from \(l(t) = e^{At}1(0)\) where

\[
l(t) = \begin{bmatrix} \ell_1(t) & \ell_2(t) & \cdots & \ell_i(t) \end{bmatrix}'
\]

The \(A, B, C\) matrices are \(A, B, C\) where \(B, = I(0)\) and \(C\), is learned during training.

Ladder Network

Let be a delay \(m_i\) (number of samples) set by the designer. The ladder network is

\[
G(z) = \begin{bmatrix} 1 \ 0 \ 0 \\
\end{bmatrix}
\]

\[
0 \ 1 \ z^{-i} \\
\end{bmatrix}
\]

\[
0 \ 0 \ 1 \ z^{-i}
\]

\[
(A, B, C) \text{ obtained through a state-space representation of } G(z)
\]

Results

We validated the performance of our RNN architectures in two dynamical systems-based experiments. In the time-series prediction benchmark, the goal is to predict a nonlinear system’s behavior from a set of available measurements. For the system identification experiment, the learning problem aims at determining the state-space representation of a system from data as a regression task. Both experiments are performed for a pendulum with no friction and a fluid flow system.

<table>
<thead>
<tr>
<th>Model</th>
<th>Pendulum (no noise; linear)</th>
<th>Pendulum (noisy; linear)</th>
<th>Pendulum (no noise; linear)</th>
<th>Pendulum (noisy; linear)</th>
<th>Fluid Flow (noisy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ladder</td>
<td>0.0157 ± 0.00050</td>
<td>0.0167 ± 0.0018</td>
<td>0.1819 ± 0.1041</td>
<td>0.1920 ± 0.1261</td>
<td>0.5135 ± 0.0026</td>
</tr>
<tr>
<td>Laguerre</td>
<td>0.0146 ± 0.0029</td>
<td>0.0168 ± 0.0057</td>
<td>0.1249 ± 0.0060</td>
<td>0.1257 ± 0.0059</td>
<td>0.7973 ± 0.5440</td>
</tr>
<tr>
<td>LMU</td>
<td>0.2261 ± 0.00866</td>
<td>0.1914 ± 0.0817</td>
<td>0.5132 ± 0.2465</td>
<td>0.5012 ± 0.1133</td>
<td>1.3942 ± 0.0225</td>
</tr>
<tr>
<td>BRC</td>
<td>21.8499 ± 1.8473</td>
<td>23.1749 ± 1.8363</td>
<td>14.0479 ± 4.1699</td>
<td>18.0374 ± 1.7233</td>
<td>0.6973 ± 0.1684</td>
</tr>
<tr>
<td>nBRC</td>
<td>2.1914 ± 0.9910</td>
<td>1.7015 ± 0.4043</td>
<td>2.6853 ± 1.5883</td>
<td>3.0102 ± 1.0030</td>
<td>0.5656 ± 0.0481</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Pendulum (no noise; linear)</th>
<th>Pendulum (noisy; linear)</th>
<th>Pendulum (no noise; linear)</th>
<th>Pendulum (noisy; linear)</th>
<th>Fluid Flow (noisy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ladder</td>
<td>0.0157 ± 0.00050</td>
<td>0.0167 ± 0.0018</td>
<td>0.1819 ± 0.1041</td>
<td>0.1920 ± 0.1261</td>
<td>0.5135 ± 0.0026</td>
</tr>
<tr>
<td>Laguerre</td>
<td>0.0146 ± 0.0029</td>
<td>0.0168 ± 0.0057</td>
<td>0.1249 ± 0.0060</td>
<td>0.1257 ± 0.0059</td>
<td>0.7973 ± 0.5440</td>
</tr>
<tr>
<td>LMU</td>
<td>0.2261 ± 0.00866</td>
<td>0.1914 ± 0.0817</td>
<td>0.5132 ± 0.2465</td>
<td>0.5012 ± 0.1133</td>
<td>1.3942 ± 0.0225</td>
</tr>
<tr>
<td>BRC</td>
<td>21.8499 ± 1.8473</td>
<td>23.1749 ± 1.8363</td>
<td>14.0479 ± 4.1699</td>
<td>18.0374 ± 1.7233</td>
<td>0.6973 ± 0.1684</td>
</tr>
<tr>
<td>nBRC</td>
<td>2.1914 ± 0.9910</td>
<td>1.7015 ± 0.4043</td>
<td>2.6853 ± 1.5883</td>
<td>3.0102 ± 1.0030</td>
<td>0.5656 ± 0.0481</td>
</tr>
</tbody>
</table>