Discovery of differential equations using probabilistic grammars: Appendix

Boštjan Gec[1,2], Nina Omejc[1,2][0000−0003−1212−1566], Jure Brence[1,2][0000−0002−1065−9912], Sašo Džeroski[1,2][0000−0003−2363−712X], and Ljupčo Todorovski[1,3][0000−0003−0037−9260]

[1] Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia
[2] Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
[3] Faculty of Mathematics and Physics, University of Ljubljana, Slovenia

Abstract. Online appendix to the paper *Discovery of differential equations using probabilistic grammars* published in the *Proceedings of the Twenty-Fifth International Conference on Discovery Science (DS-2022)* by Springer Nature.

A Learning curve estimation

ProGED performs Monte-Carlo sampling of candidate equations from a distribution defined by a PCFG. We obtain performance estimates, such as learning curves and the computation time required to find the solution, through bootstrapped resampling.

As the result of running ProGED, we have a list of N candidate equations, each with an associated error-of-fit and a probability of the right-hand side expression derived from the grammar. We randomly sample a sequence of N models from this list without repetition and following the expression probabilities. We then calculate the cumulative minimum of error across this sequence and thus obtain a single learning curve. We repeat this procedure many (at least 1000) times and average the learning curves. The resulting curve estimates the expected best error-of-fit ProGED would achieve with a given number of models sampled. In this way, we simulate repeating the random sampling experiment many times.

B Reconstructed equations

Table 1. Each methods' best reconstructed model for the VDP data set.

method	obs.	Reconstructed ODEs
SINDy	XY	$\dot{x} = 0.99998y$ $\dot{y} = -0.99998x + 0.49991y - 0.49991x^2y$
L-ODEfind	XY	$\dot{x} = 0.99990y$ $\dot{y} = -0.99995x + 0.49980y - 0.49986x^2y$
	X	$\ddot{x} = -0.99993x + 0.49918\dot{x} - 0.49961x^2\dot{x}$
	Y	$\ddot{y} = 0.01273 + 0.76615\dot{y} - 0.43317y^3 + 0.80264y^2\dot{y} - 0.06486\dot{y}^3$
GPoM	XY	$\dot{x} = 0.99975y$ $\dot{y} = -1.00014x + 0.50107y - 0.49953x^2y$
	X	$\dot{x} = 1y$ $\dot{y} = -0.99992x + 0.50070y - 0.49965x^2y$
	Y	$\dot{x} = 1y$ $\dot{y} = 0.01552 + 0.44866x - 1.22201y - 0.56297y^3+$ $+ 0.88215x^2y + 0.10693xy^2 + 0.03426y^3$
ProGED	XY	$\dot{x} = 0.99999y$ $\dot{y} = -0.99998x + 0.49991y - 0.49991x^2y$
	XY	$\dot{x} = 1.00001y$ $\dot{y} = -0.99998x + 0.49997y - 0.49998x^2y$
	X	$\dot{x} = 0.24209y$ $\dot{y} = -4.17027x + 0.54101y - 0.54491x^2y$
	Y	$\dot{x} = 0.56267y$ $\dot{y} = -1.78159x + 0.50896y - 1.59903x^2y$

Table 2. The first-order variants of the second-order ODEs reconstructed by the L-ODEfind method and reported in Table 1.

Obs.	Inverted ODEs
X	$\dot{x} = y$ $\dot{y} = -0.99993x + 0.49918y - 0.49961x^2y$
Y	$\dot{x} = 0.01273 + 0.76615x - 0.43317y^3 + 0.80264y^2x - 0.06486x^3$ $\dot{y} = x$