ML for PDE and Physics

Jérémie Donà

March 2019

Spatio Temporal Problem in Physics

A physical state variable (or vector of state variables) u often obeys a partial differential equation, so we can write:

$$\frac{\partial u}{\partial t} = N(u) \tag{1}$$

where N is a differential operator. This writing is general and may leads to complex PDE systems.
For most linear equation, we can write:

$$\frac{\partial u}{\partial t} = \sum N_i(u)$$

where N_i are differential operators

Separability in PDE:

Suppose the form of (1) is known but the overall state u is not known. For example, it is hard to infer the future position of an object using only is position, we would like to infer the speed of the object using data...

Key Idea:
We can formulate the prediction problem as:

$$u(t) = u(t_0) + \int_{t_0}^{t} N(u(\nu))d\nu \text{ , or if } d\nu \text{ is "small enough" :}$$

$$u(t) \approx u(t_0) + dt * N(u(t_0))$$

$$u(t) \approx u(t_0) + dt * \sum N_i(u(t_0))$$

Also all spatial differentiation can be approximated using an appropriate convolution filter.

Imposing Euler scheme:

Hypothesis: Imposing the form (1) to a PDE can lead to recover the overall state and make accurate prediction.

Why is it interesting ?

- Recovers full state using dynamic
- Proposes a solution to the problem of forecasting under partial observation
- Makes use the physics and not "blindly" apply the ML
- Allows interpretation of the networks

Flourishing litterature:

- Mazari, Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations
- Michael Lutter, Christian Ritter, Jan Peters, Deep Lagrangian Networks: Using Physics as Model Prior for Deep Learning (ICLR 2019)

An example, Diffusion and SST

The data are complex, the problem even more but it can be very roughly designed the following way, u being the sea surface temperature :

$$\frac{\partial u}{\partial t} = w.\nabla u + D * \Delta u$$

Problem setting:

- the full state is (u,w)
- w is unobserved (and D) also it may vary over time

Question: The problem is, can we recover w_t from a sequence past observations of $(u_t)_t$ and use it to perform good prediction ?

$\rightarrow$ Taken's theorem

Two Options: Fixed Euler

(FE) Fixed Euler scheme: Use:

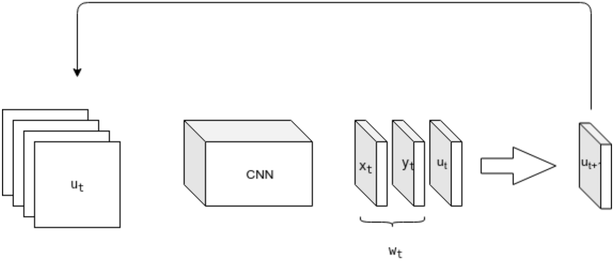
$$u_{t+1} = u_t + dt * (w_t.\nabla u_t + D\Delta u_t) \tag{2}$$

this could lead to:

$$(u_{t+1} - u_t)/dt - D\Delta u_t = (w_t^1.\partial_x u_t + w_t^2.\partial_y u_t) \tag{3}$$

As the forecasting operator, compute the gradients of this (differentiable) operation using supervision and backprop

In the fixed Euler scheme:



Fixed Euler: Elements of results: prediction accuracy

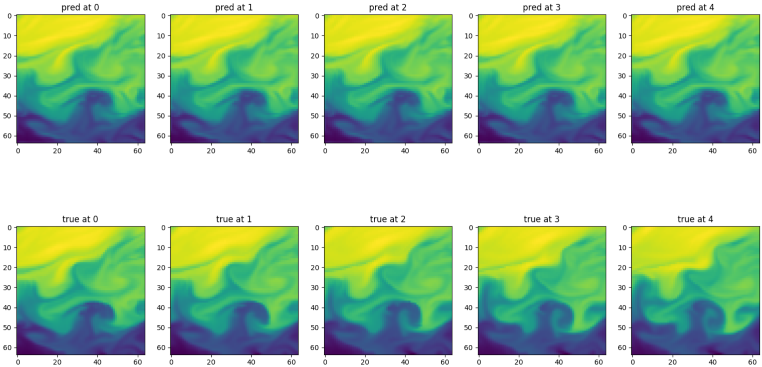
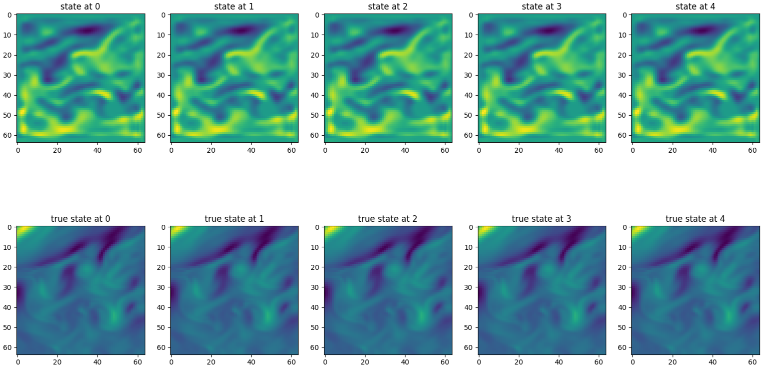


Figure 2: Prediction using Deep Euler Scheme

Fixed Euler: Elements of results, state inference

Not perfect, but not orthogonal



Two Options: Resnet forward scheme: 1/2

This Option (I. Ayed, E. De Bezenac)

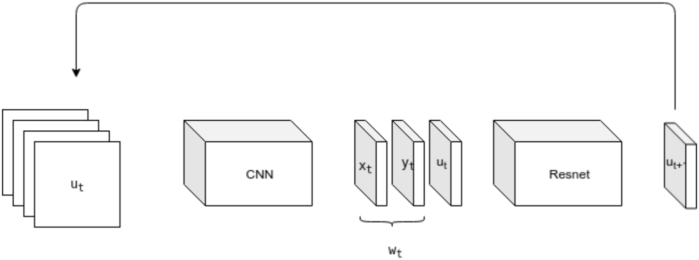


Figure 4: Deep Euler Scheme, using Resnet

Two Options: Resnet forward scheme, 2/2

However, instead of outputting, one vector prediction (one channel), we can decompose the learning into learning 4 vectors, to be summed in order to perform the prediction.

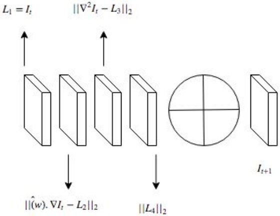


Figure 5: Deep Euler Decomposed Euler Scheme

Why is it interesting ?

- Forcing the appearance of physical term for interpretation
- Helps in prediction

Forward scheme:

Problem:

- which metric to use ? 1- cosine similarity, mse ??
- How many Resblock (meaning) ?
- Does decomposition really helps forecasting ?
- Hard to train: Identity is a good strategy

Experiments:

3 experiments are being carried out:

- Diffusion: SST (Real data: Hard)
- Reaction Diffusion: Turing Equation (Simulated data: Hard)

$$\frac{\partial u}{\partial t} = D_1 * \partial_{xx} u + F(u, v)$$
$$\frac{\partial v}{\partial t} = D_2 * \partial_{xx} v + G(u, v)$$

- Smaller dimension Navier Stokes: Shallow water

Reaction Diffusion

The problem :

$$\frac{\partial u}{\partial t} = D_1 * \partial_{xx} u + F(u, v)$$
$$\frac{\partial v}{\partial t} = D_2 * \partial_{xx} v + G(u, v)$$

may generate very complex data, and the problem may vary a lot according to the initial conditions: `https://www.youtube.com/watch?v=PtPK_xx5Hks`.

This experiment can lead to various hypothesis testings with the same problem setting as previously (on channel either u or v is missing)

- Can we infer F and G using deep neural network
- Can we infer accurately the missing channel using a forward Euler scheme ?
- Does the network generalise well to various initial conditions ?

Problem 2: Climate Recovery and Generation

Problem Setting:

Is it possible to reconstruct the anomaly of the climate using only very few points ?

- data: output of a climate model
- test data : only proxies (i.e very sparse observation of the climate)
- Are those data points informative enough ?

The second problem is that the observations does not concern the anomaly BUT the temperature itself.

$\rightarrow$ Domain transfer

What Does a temperature map looks like ?

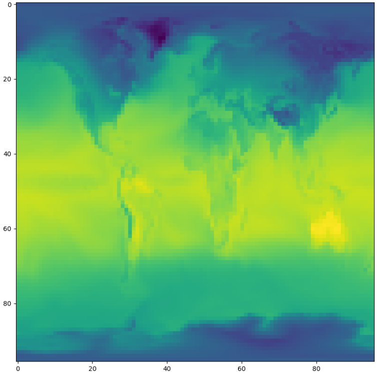


Figure 6: Climate temperature

What Does an anomaly map looks like ?

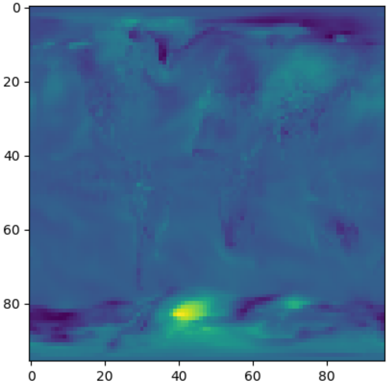


Figure 7: Climate Anomaly at the same point

So how to summarize the problem:

- We have data that comes from a model output
- Domain transfer (from sparse temperature domain D_t to anomaly D_a)
- With missing data

Idea to handle the problem :

- Deep Conditional GAN
- MisGAn ??