Hypothesis for Neural Residual PDEs

September 11, 2019

Define a dataset $\mathcal{D} = \{\phi_i, \psi_i, \mathbf{x}_{i,1:T}\}_{i=1}^N$ of boundary conditions ϕ, initial conditions ψ, and a sequence of observations $\mathbf{x}_{1:T}$. We can define an observation in two ways: first, $\mathbf{x}_t = (\mathbf{u}_t, \mathbf{v}_t, \mathbf{p}_t)$ contains two dimensions of momentum and pressure; or second, $\mathbf{x}_t = (\lambda_t^{\mathbf{u}}, \lambda_t^{\mathbf{v}}, \lambda_t^{\mathbf{p}})$, a vector of spectral coefficients for momentum and pressure functions.

Then, the function we want to learn consists of two components — a base function $f \in \mathcal{F}$ where $f_\theta : \mathcal{X} \to \mathcal{X}$ over an observation. This base function is responsible for learning the dynamics over observations. Second, define a residual function $g_\theta : \{\phi, \psi\} \to \mathcal{F}$, tranforming boundary and initial conditions to a function. The base function is shared over all entries in the dataset — we consider the following objective:

$$\min \mathbb{E}_{\phi,\psi,\mathbf{x}_{1:T}\sim p_{\mathcal{D}}} \left[\sum_{t=1}^{T} \mathbf{x}_t - (g_\theta(\phi, \psi) + f_\theta(\mathbf{x}_{t-1})) \right] \tag{1}$$

Note that g_θ is kind of like a hypernetwork. Next question is what does f_θ look like (how does it relate to PDEs/ODEs)?