Hamiltonian Learner ReadMe

August 8, 2017

1 Mathematical Background

1.1 Hamiltonian Mechanics

Let $\mathbb{R}^{2n}$ be our phase spaces with coordinates and conjugate momenta x_i and p_i

$$
\begin{aligned}
\frac{dx_i}{dt} &= \{H, x_i\} = \frac{dH}{dp_i} \\
\frac{dp_i}{dt} &= \{H, p_i\} = -\frac{dH}{dx_i}
\end{aligned}
$$

1.2 Algebraic Varieties

Let I be an ideal in $k[x_1 \cdots x_d]$. Demanding that all the functions in that ideal vanish determines a variety in k^d. k will be $\mathbb{R}$ but I don't want to get into the subtleties of base change to algebraic closure so some of the statements aren't strctly true.

1.1 Example (Circle) $I = (x^2 + y^2 - 1) \subset \mathbb{R}[x, y]$

1.3 Lagrangian Correspondences

1.2 Definition (Lagrangian Submanifold)

Consider a pair of symplectic manifolds $X_{in,out}$. This stands for the inputs and outputs of a system. For example, it could be a black boxed electronic circuit. For each of the input wires there is an $\mathbb{R}^2$ for the current and voltage. These form a canonical conjugate pair for each of the wires. So for a circuit with n input wires and m output wires we get $\mathbb{R}^{2n}$ and $\mathbb{R}^{2m}$. To specify what the circuit is at this black boxed level we specify a morphism $X_{in} \to X_{out}$ in the following sense.

1.3 Definition (LagCor) *A morphism between symplectic manifolds in the "category" LagCor is a Lagrangian correspondence. That is a Lagrangian submanifold of $X_{in} \times \bar{X}_{out}$. Composition of L_{12} and L_{23} is given by pullback if possible. That is a transversality condition with the diagonal in $X_2 \times \bar{X}_2$.*
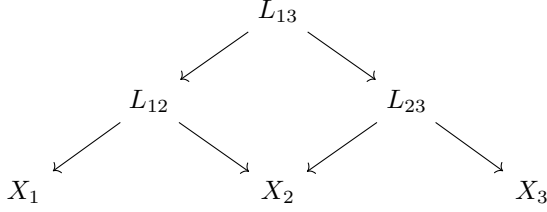
Composition is simply piping some outputs into other inputs. By moving some wires around you are allowed to perform partial gluings so that you can feed some outputs of system 1 into system 2 and leave some outputs as free to go all the way to the end. Similarly for inputs to system 2 that don't come from system 1 but instead straight from the begining.

1.4 Definition (Graphs of symplectomorphisms) *For $X_1 = X_2$ a large class of morphisms $X_1 \to X_2$ is given by the graphs of symplectomorphisms $\phi\ X_1 \to X_2$. A symplectomorphism mean it is a diffeomorphism that preserve ω. But these are special in the sense that if you specify the input, you uniquely know the output. Sets and relations is better than sets and functions.*

1.3.1 Singular Lagrangians

We might have systems where the vanishing locus is no longer a manifold but is stratified. An example that shows up often is hysterisis loops.

1.5 Definition (Magnetic Hysterisis) *Let th symplectic manifold be $\mathbb{R}^2$ with magnetization and applied field. Then a material defines a singular "1-dimensional" subset. A picture describes it best.*

This is thought of as a system with 0 inputs and $\mathbb{R}^2$ output. The dynamics is then constrained to this Lagrangian because once you specify the applied field you are stuck with either 1 or 2 values for the magnetization because of the equations of state. It is not a Lagrangian submanifold only because of those triple points.

2 Down from Infinite Dimensional

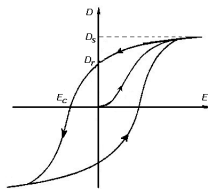Consider the problem for a PDE instead of an ODE. Let the space be M

Figure 1: Hysterisis loop

$$\omega \;=\; \int_M \partial A \wedge \partial B$$

Give a basis of $L^2(\Omega^\bullet(M))$ so that the integration pairing is adapted. Will explain by example on $\Omega^0(S^2)$ and $\Omega^2(S^2)$

$$e_{lm,0} \;=\; Y_l^m(\theta,\phi)^*$$
$$e_{lm,2} \;=\; Y_l^m(\theta,\phi)dvol_{S^2}$$

A complexified 0 form A and a complexified 2 form B can be decomposed in this basis and as can their variations.

$$\omega \;=\; \sum_{lm} \partial_{lm0} A \wedge \partial_{lm2} B + \cdots$$

So this is taking a function on $S^2 \times \mathbb{R}$, decomposing each time slice into spherical harmonics and then asking if we can analyze the associated infinite dimensional Hamiltonian system. Of course realistically, it should be reasonable to approximate with some sufficient L cutoff. There will be some relatively small $n(L)$ modes kept. This can then be fed into the machine with $\mathbb{R}^{2(n(L))}$.

2.1 Definition (Kaluza-Klein Reduction) *This is in particular a reduction from a $M \times \mathbb{R}$ spacetime to just mechanics with only $\mathbb{R}$ for time and no space.*

3 Implementation

3.1 Time Trajectory

Input some traces of trajectories in phase space. That is $x_i(t_j) = x_{ij}$ and $p_i(t_j) = p_{ij}$ for t_j a discetization of a time interval. The loss function is given by comparing to Hamilton's equations.

$$\Delta x_{ij} = \frac{x_{i,j+1} - x_{ij}}{t_{j+1} - t_j} \approx \frac{dx_i}{dt}(t_j)$$

$$\Delta p_{ij} = \frac{p_{i,j+1} - p_{ij}}{t_{j+1} - t_j} \approx \frac{dp_i}{dt}(t_j)$$

$$L = \sum_j \sum_i (\Delta x_i(t_j) - \frac{dH}{dp_i}(x_{ij}, p_{ij}))^2 + (\Delta p_i(t_j) + \frac{dH}{dx_i}(x_{ij}, p_{ij}))^2$$

The possible Hamiltonians are specified as giving the terms that are allowed to show up. For exampe:

$$H_{pol,3} = A + Bx + Cp + Dx^2 + Exp + Fp^2 + Gx^3 + Hx^2p + Ixp^2 + Jp^3$$

$$H_{trig,1} = A \cos Bx + C \sin Dx + E \cos Fx + G \sin Hx$$

$$H_{exp,1} = Ae^{Bx} + Ce^{Dp}$$

These capital letter unknowns are what we optimize over.

3.2 Algebraic Variety Learning

First ignore the dynamical system aspect and consider the problem of giving a series of points x_{ij} where i indexes the space $\mathbb{R}^n$ while j indexes which data point this is. We are told that they come from an algebraic variety of low degree and the task is to determine the defining functions that cut out the variety. For example, if we know it is cut out by a real plane quadric.

$$f = A + Bx_1 + Cx_2 + Dx_1^2 + Ex_1x_2 + Fx_2^2$$

$$\begin{pmatrix} 1 & x_{1,j} & x_{2,j} & x_{1,j}^2 & x_{1,j}x_{2,j} & x_{2,j}^2 \\ \cdots \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \\ F \end{pmatrix} = \begin{pmatrix} 0 \\ \cdots \end{pmatrix}$$

Giving one function means cutting a codimension 1 variety typically. So to give a codimension m variety, we should give m such functions which are independent of each other. Then our variety is the vanishing set $V(I)$ for the ideal $I = (f_1 \cdots f_m)$.

However because of noise we cannot expect exactly 0. Instead we say that we seek to minimize the squares of each of the rows of the RHS. That turns into a quadratic optimization.

3.3 Lagrangian Subvariety Learning

If we are told that we are learning a coisotropic variety of codimension m that amounts to giving m such functios as above but now there is the constraint that all of the f_i Poisson commute. These are quadratic constraints in the capital letter unknown variables.

3.1 Example (Codimension 2 in 4d)

$$\begin{aligned} f_1 &= A_1 + B_1 x_1 + C_1 x_2 + D_1 p_1 + E_1 p_2 + F_1 x_1^2 + G_1 x_2^2 + H_1 x_1 x_2 \\ f_2 &= A_2 + B_2 x_1 + C_2 x_2 + D_2 p_1 + E_2 p_2 + F_2 x_1^2 + G_2 x_2^2 + H_2 x_1 x_2 \\ \{f_1, f_2\} &= B_1 D_2 + C_1 E_2 - D_1 B_2 - E_1 C_2 + \cdots \end{aligned}$$

At the very extreme for a Lagrangian submanifold in $\mathbb{R}^{2n}$ we give a codimension n coisotropic variety. This is the case of an integrable system where the f_i are the conserved quantities that define the action coordinates in action-angle variable system formalism.