Active Manifolds: A non-linear analogue to Active Subspaces

Robert A. Bridges
PhD, Oak Ridge National Laboratory
bridgesra@ornl.gov

Anthony Gruber, PhD, TTU

Christopher Felder, MS, WUSTL

Miki Verma, BS, ORNL

Chelsey Hoff, BS

Problem

$$f : \mathbb{R}^n \xrightarrow{C^1} \mathbb{R} \text{ so } f(x_1, ..., x_n) \in \mathbb{R}$$

- **Regression:** Given $\{(\mathbf{x}_i, f(\mathbf{x}_i), \nabla f(\mathbf{x}_i))\}_i$ recover f
- **Sensitivity Analysis:** Which coordinate directions matter?

Can we reduce dimension of input space to make this easier?

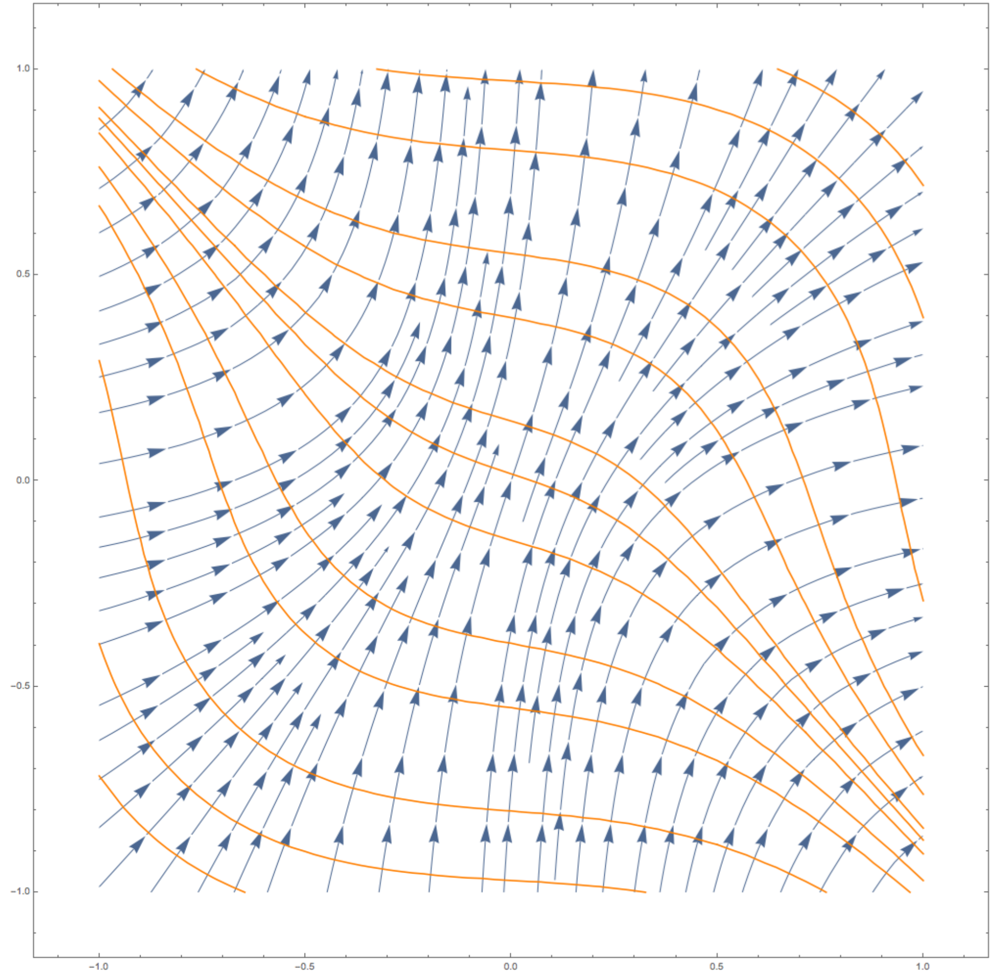
New Approach: Active Manifolds

Intuition: Standing at $\mathbf{x_0}$ in domain of f

- There are $n-1$ directions one can step without changing f.
- There is one, special direction, $\nabla f(\mathbf{x_0})$ in which f changes maximally!

Regression Algorithm Idea:

- Use gradient ascent/descent to walk up/down hill and record values of f—an **active manifold.**
- To approximate $f(\mathbf{x_0})$ walk along a level set to the **active manifold.**



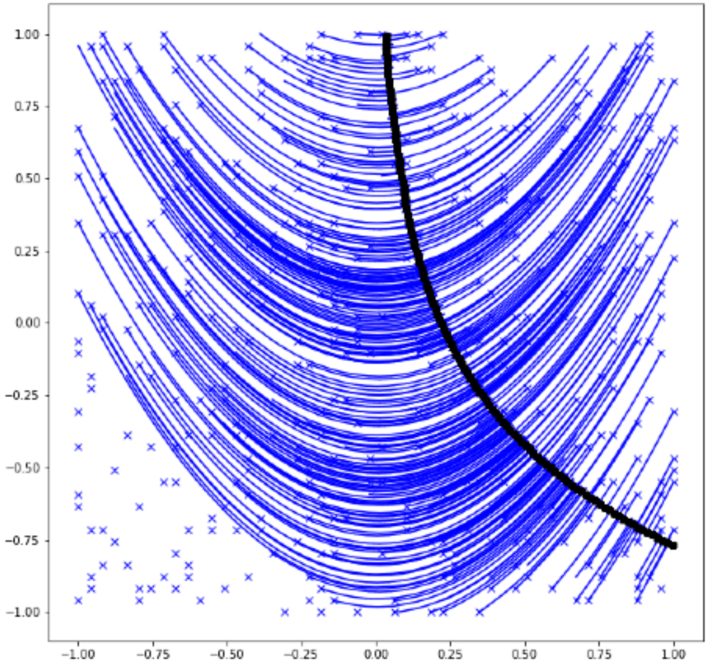
$$f(x,y) = x^3 + y^3 + 0.2x + 0.6y$$

Level sets (orange) and **gradient vector field (blue)** tangent to an **active manifold** at every point.

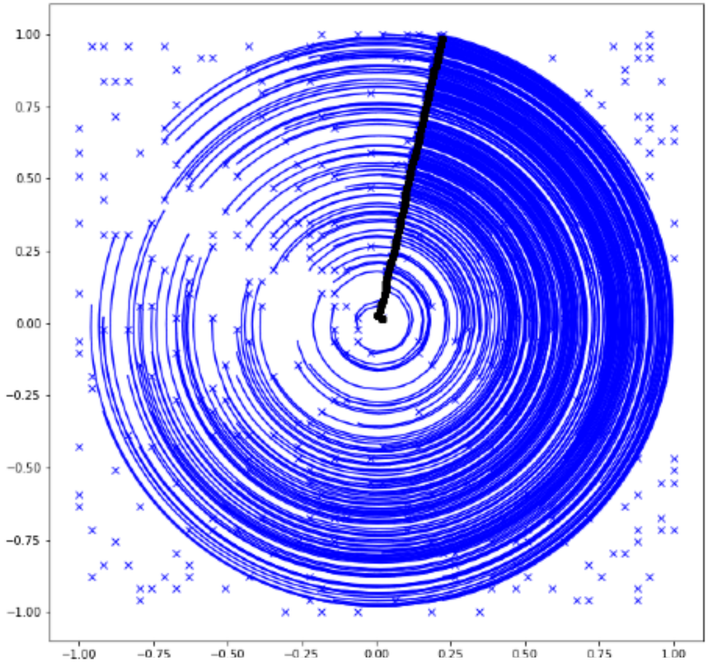
Robert A. Bridges

Examples & Regression Results

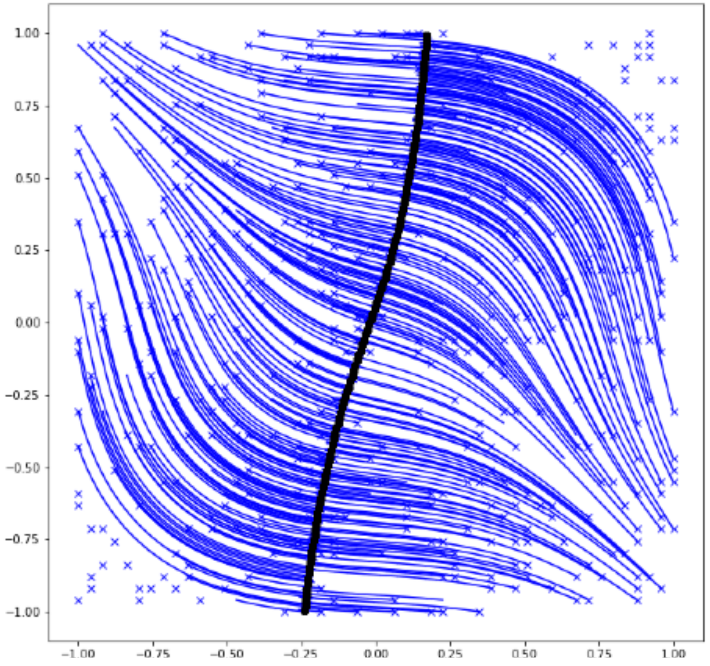
$$f_1(x,y) = e^{y-x^2}$$



$$f_2(x,y) = x^2 + y^2$$



$$f_3(x,y) = x^3 + y^3 + 0.2x + 0.6y$$



Result: AM exhibits order(s) of magnitude less L^1 and L^2 average error and error variance than AS.

		ℓ^1 **mean**	ℓ^1 **std**	ℓ^2 **mean**	ℓ^2 **std**	n/N **mean**
f_1	AM	6.739E-3	6.826E-4	1.879E-4	1.847E-5	86.7%
	AS	0.585	8.130E-3	0.751	8.600E-3	100%
f_2	AM	0.0158	9.697E-4	4.015E-4	2.562E-5	77%
	AS	0.395	5.484E-3	0.488	6.890E-3	100%
f_3	AM	0.0106	8.442E-4	3.154E-4	2.887E-5	92.9%
	AS	0.982	0.018	1.22	0.0224	100%

Mathematical Foundation & Pseudo-Algorithm Presented

1. Active Manifolds (AM)

Here we provide the mathematical foundation for AM and describe a pseudo-algorithm for reducing analysis of the m-dimensional function to its one-dimensional analogue. Examples to illustrate the method are provided, including illustrations of problems or obstructions identified.

1.1. Mathematical Justification:

Recall that the arc length of a C^1 curve $\gamma(t) : [0,1] \to \mathbb{R}^m$ is given by $S(\gamma) = \int_0^1 |\gamma'(t)|\, dt$. Let $U \subset \mathbb{R}^m$ open and assume $f \in C^1(U)$.

We seek

$$\arg\max \int_0^1 \langle \nabla f(\gamma(t)), \gamma'(t) \rangle \, dt$$

over all C^1 curves $\gamma(t) : [0,1] \to U$, such that $|\gamma'| = 1$ (constant speed), where $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean inner product. Note that the integrand satisfies

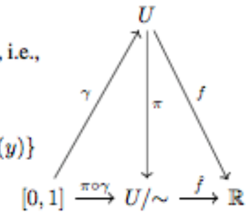
$$\langle \nabla f(\gamma(t)), \gamma'(t) \rangle = |\nabla f(\gamma(t))| \, |\gamma'(t)| \cos\theta$$

where θ is the angle between $\nabla f(\gamma(t))$ and $\gamma'(t)$. Clearly this quantity is maximal when $\theta = 0$, indicating that $\nabla f(\gamma(t))$ and $\gamma'(t)$ point in the same direction; hence, the solution to this optimization problem is

$$\gamma'(t) = \frac{\nabla f(\gamma(t))}{|\nabla f(\gamma(t))|}, \qquad (1)$$

For the following proposition, let f, U be as above and:

- $\mathcal{M} = \operatorname{Im} \gamma(t)$ an active manifold of f
- The relation $\sim$ defined by f, i.e., $\forall x, y \in U$, $x \sim y \iff f(x) = f(y)$
- $[x] = \{y \in \mathbb{R}^m : f(x) = f(y)\}$
- $\mathbb{R}^m/\!\sim\; = \{[x] : x \in \mathbb{R}^m\}$
- $\pi : \mathbb{R}^m \twoheadrightarrow \mathbb{R}^m/\!\sim$



Proposition 1.3. *If $\gamma(t)$ is a solution to Eqn. (1) on an open set U away from points where $\nabla f = 0$, then the following statements hold.*

(i) $\mathcal{M}$ is an immersed C^1 submanifold of $U \subseteq \mathbb{R}^m$.

(ii) $U/\sim$ is a C^1 manifold.

(iii) $\pi \circ \gamma$ is a C^1 embedding of $\mathcal{M}$ in $\mathbb{R}^m/\sim$.

(iv) $f \circ \gamma : [0,1] \to \mathbb{R}$ is strictly increasing.

Proof. (i): Note that $f|_{\mathcal{M}}$ provides a global C^1 chart for $\mathcal{M}$. Further, $\mathcal{M}$ is immersed since $|\gamma'| = 1$, hence γ' does not vanish. (ii): Since f is C^1 and constant on the fibers of π, the map $\hat{f} : U/\!\sim\; \to \mathbb{R}$ defined as $\hat{f}([x]) := f(x)$ is C^1. So $U/\!\sim$ is a C^1 manifold with global chart $\hat{f}$. (iii): $\pi|_{\mathcal{M}}$ is a bijection onto $\pi(\mathcal{M})$ since $\mathcal{M}$ fibers pointwise under π. Since π is linear, it follows that $d\pi|_{\mathcal{M}} = \pi|_{\mathcal{M}}$ is bijective; hence, π is an embedding. (iv): Monotonicity of $f \circ \gamma$ follows directly from the definition: $\forall t\ \gamma(t) \parallel \nabla f(t)$. □

a constant-speed streamline of ∇f. Specifying a starting point, γ_0, uniquely identifies the flow as furnished by the following standard theorem of differential equations.

Lemma 1.1. *Given $f : U \subset \mathbb{R}^m \xrightarrow{C^1} \mathbb{R}$ and an initial value $\gamma_0 \in U$, there exists a unique local solution $\gamma(t)$ to the system of first-order ordinary differential equations described by (1).*

Proof. Choose any compact and convex subset $K \subset U$ containing γ_0. Since f is C^1, ∇f satisfies the Lipschitz condition, $|\nabla f(x_1) - \nabla f(x_2)| \le L_K |x_1 - x_2|$, for $x_1, x_2 \in K$ and $L_K < \infty$ is some Lipschitz constant. By Theorem 1 Ch. 6 from Birkhoff & Rota (1969), these conditions are sufficient for the existence and uniqueness of a local solution $\gamma(t)$ to Eqn. (1) about γ_0 in K, which can be reparametrized to have domain $[0,1]$ as desired. Since K was an arbitrary compact set we have the result. □

Definition 1.2. Let $f : U \subset \mathbb{R}^m \xrightarrow{C^1} \mathbb{R}$. We say that $\mathcal{M} \subset U$ is an **active manifold** defined by f provided there exists a constant-speed parametrization of $\mathcal{M}$, $\gamma(t) : [0,1] \to U$, such that condition (1) is satisfied for all $t \in [0,1]$.

Theorem 1.4. *Suppose the level set $\{f = \alpha\}$ is connected and γ is any active manifold such that $\alpha \in Im(f \circ \gamma)$. Then $\exists!\ t_0$ such that $\gamma \cap \{f = \alpha\} = \{\gamma(t_0)\}$, and $\gamma \perp \{f = \alpha\}$.*

Proof. The Implicit Function Theorem guarantees that for each $\alpha \in \operatorname{Im} f$, the level set $\{x : f(x) = \alpha\}$ is an $(m-1)$-dimensional submanifold of $\mathbb{R}^m$ that is orthogonal to the gradient vector field and therefore to any intersecting active manifold. By hypothesis $\exists\, t_0$ such that $\gamma(t_0) \in \{f = \alpha\}$. Uniqueness follows from monotonicity of $f \circ \gamma$ (Proposition 1.3.*iv*). □

Implication: This theorem implies that if one can recover $f \circ \gamma$ (a 1-D regression problem), then one can recover f on the connected component of any level set touching γ. Concisely, if p is in the component of $A := \{f = f(p)\}$ intersecting γ, one may move freely in the $(m-1)$-dimensional submanifold A transverse to γ without changing f. This motivates our AM pseudo-algorithm.

1.2. Active Manifolds Pseudo-Algorithm:

The AM algorithm has three broad components: (1) Build the active manifold $\mathcal{M} = \operatorname{Im}(\gamma(t))$; (2) Approximate the

Sensitivity Analysis for MHD Generator (Following Glaws et al. 2017)

Induced magnetic field model

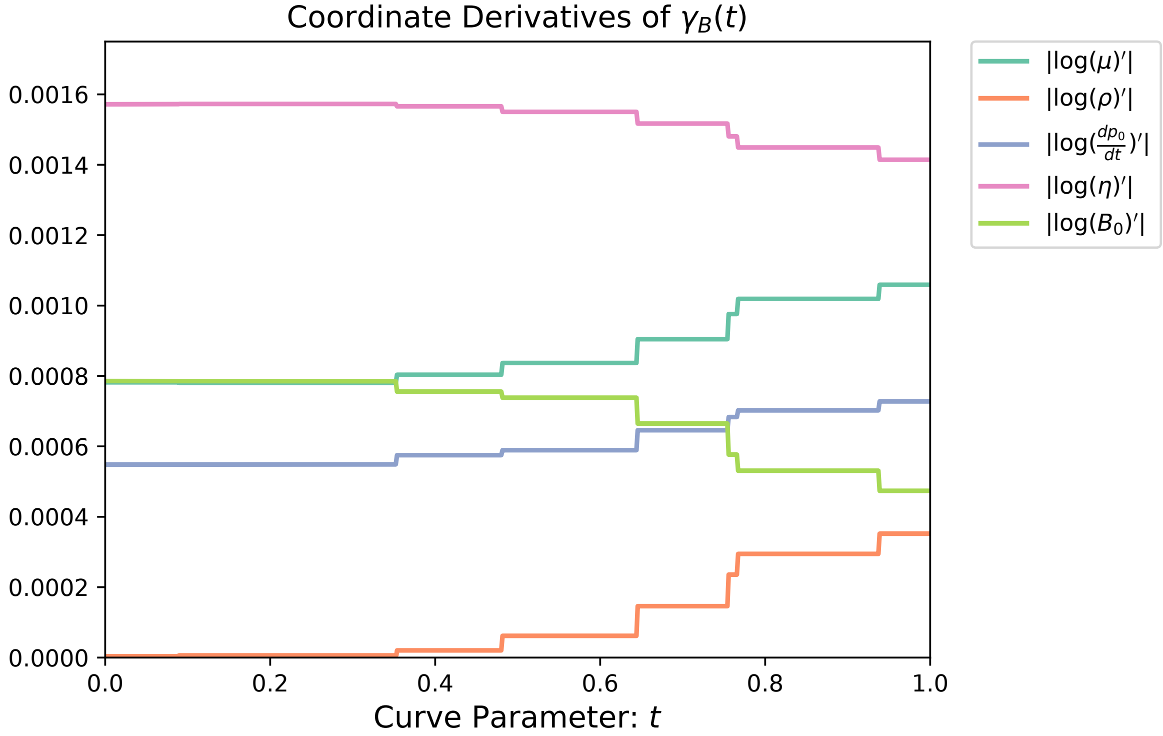
$$B_{ind} = \frac{\partial p_0}{\partial x} \frac{l \mu_0}{2 B_0} \left(1 - 2\frac{\sqrt{\eta\mu}}{B_0 l} \tanh\left(\frac{B_0 l}{2\sqrt{\eta\mu}}\right)\right)$$

Parameters & ranges:

Index	Name	Notation	Interval
1	Fluid viscosity	$\log(\mu)$	$[\log(0.05), \log(0.2)]$
2	Fluid density	$\log(\rho)$	$[\log(1), \log(5)]$
3	Applied pressure gradient	$\log\left(\frac{\partial p_0}{\partial x}\right)$	$[\log(0.5), \log(3)]$
4	Resistivity	$\log(\eta)$	$[\log(0.5), \log(3)]$
5	Applied magnetic field	$\log(B_0)$	$[\log(0.1), \log(1)]$

Result

AM allows visualization to see parameter influence throughout the active manifold:



Active Manifold Benefits

- Reduces n - dimensional analysis to **1 dimension** (computationally more expensive)
- Order of magnitude **greater accuracy** in regression over AS
- **Accessible visualizations** of the function and parameters gradients along the active manifold
- **Permits sensitivity analysis** locally along the active manifold

Questions?

Robert A. Bridges
bridgesra@ornl.gov

Previous Approaches & Related Work

- **Sliced Inverse Regression**[1]: Given $\{(\mathbf{x}_i, f(\mathbf{x}_i))\}_i$ find lower rank matrix B so $f(x) \approx g(Bx)$.
- **Active Subspaces (AS)** [2]**:** Given $\{(\mathbf{x}_i, f(\mathbf{x}_i), \nabla f(\mathbf{x}_i))\}_i$
  - Let $$\mathbf{C} = \frac{1}{N} \sum_{i=1}^{N} \nabla f_{a_i} \nabla f_{a_i}^T = \mathbf{W} \mathbf{\Lambda} \mathbf{W}^T.$$
  - Do SVD on $\mathbf{C}$ to find directions f changes most.
- **ResNet Isosurface Learning**[3]**:** Given $\{(\mathbf{x}_i, f(\mathbf{x}_i), \nabla f(\mathbf{x}_i))\}_i$ find nonlinear, lower rank B (using ResNet) so $f(x) \approx g(Bx)$.

References:
[1] See Li 1991, Duan & Li 1991, Li & Naschtsheim 2006, Coudret et al. 2014
[2] See Russi 2010, Constantine et al. 2014, 2015, Lukaczyk et al. 2014, Constantine & Diaz 2017
[3] See Zhang & Hinkle 2019 (arxiv preprint)