Technische
Universität
München

Optimisation Geometry

Paper Review

Emre Faik Gökce, Hamza Imamovic, Wen Perng

Electrical Engineering

15 July 2024

Table of Contents

Classic Optimisation

- Developed to scale well.
- One-time use.

Manifold Optimisation

- Leverage topology of the family of cost functions.
- Offline computation for the whole family of functions.
- Online computation for specific pair of parameters θ.

we are interested in given a θ:

$$\left\{ x_* \in X \mid f(x_*;\theta) = \min_x f(x;\theta) \right\}$$

Problem Formulation

Given input $\theta \in \Theta$ algorithm returns $g(\theta) \in \Theta \times X$

$$f(g(\theta)) = \min_x f(x;\theta)$$

BUT hard to find $g(\theta)$ explicitly

Simplest example:

$$f(x;\theta) = h(x-\theta)$$
$$x = g(\theta) = x_* + \theta$$

Definition 1: Fibre Bundle Optimisation Problem

Let $\mathcal{M}$ be a smooth fibre bundle over the base space Θ with typical fibre X and canonical projection $\pi : \mathcal{M} \to \Theta$. Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. The fibre bundle optimisation problem is to devise an algorithm computing an optimising function $g : \Theta \to \mathcal{M}$ that satisfies:

- $(\pi \circ g)(\theta) = \theta$ for all $\theta \in \Theta$.
- $(f \circ g)(\theta) = \min_{p \in \pi^{-1}(\theta)} f(p)$ for all $\theta \in \Theta$.

During optimisation, we would like to work with nice functions that guarantees convergence to critical points.

Morse Function

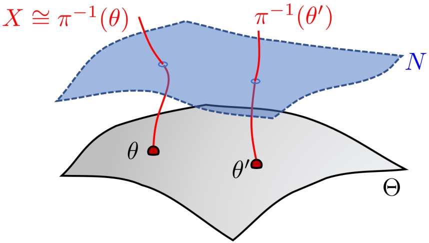
A function is Morse if all its critical points are non-degenerate, i.e. the critical points are isolated.

Def 10. Fibre-Wise Morse Function

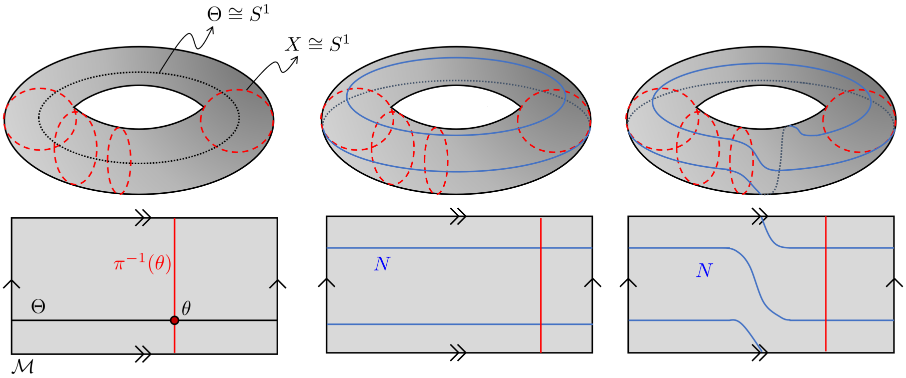
A function $f : \mathcal{M} \to \mathbb{R}$ on the fibre bundle $\pi : \mathcal{M} \to \Theta$ is fibre-wise Morse if its restriction on a fibre $f|_{\pi^{-1}(\theta)}$ is Morse.

Lemma 12

The set N of fibre-wise critical points is a submanifold of $\mathcal{M}$ with the same dimension as Θ, intersecting each fibre transversally.



Consider base $\Theta \cong S^1$ and fibre $X \cong S^1$. A function $f : \Theta \times X \to \mathbb{R}$ has fibre-wise critical points N. N intersects each fibre an equal amount of times: $N \cap \pi^{-1}(\theta)$ is the same size for all θ.



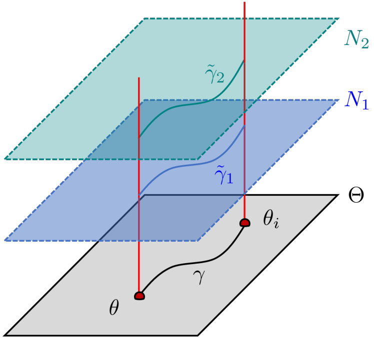
Each connected component of N doesn't change index.

Preparation stage:

1. Produce list of connected components N_i of N.
   Let $\tilde{N}$ be the union of the local minimas in N_i.
2. Choose a set of starting θ_i's, and find $\tilde{N} \cap \pi^{-1}(\theta_i)$.
3. *Compute bounds on certain higher derivatives.*

Optimisation stage: when a θ is given,

1. Determine a starting θ_i and curve $\gamma : [0,1] \to \Theta$ with $\gamma(0) = \theta_i$ and $\gamma(1) = \theta$.
2. For each $p \in \tilde{N} \cap \pi^{-1}(\theta_i)$, lift γ to $\tilde{\gamma} : [0,1] \to \tilde{N}$ such that $\tilde{\gamma}(0) = p$. The lifting uses the bounds.
3. Analyze all $\tilde{\gamma}(1)$ to find global minima.



Topological complexity based on the amount of N_i and $\tilde{\gamma}_i$ tracked.

Lifting is not so easy, we need an approximating method.

Definition 15: Approximate Critical Point

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with a non-degenerate critical point at the origin: $\mathrm{D}h(0) = 0$ and $\mathrm{D}^2 h(0)$ is non-singular.
A point x is an **approximate critical point** if, when started at $x_0 = x$, the Newton iterates x_k at least **double in accuracy** per iteration: $||x_{k+1}|| \leq \frac{1}{2}||x_k||$.

Importance for the homotopy-based algorithm

If a bound $\rho > 0$ can be calculated, such that every point in interval $[-\rho, \rho]$ are approximate critical points are approximate critical points, a fast approximation can be guaranteed.

Lifting Conditions for Step 2 of the Algorithm

For each $p \in \tilde{N} \cap \pi^{-1}(\theta_i)$:

1. Start from p_0 of $\gamma(t_0 = 0)$
2. Use Newton's method to get to $\hat{p}_1$ of $\gamma(t_1)$
3. Repeat 2 until $\hat{p}_N$ of $\gamma(t_N = 1)$

The advantages of this algorithm:

1. p_0's are pre-calculated
2. t_i are so defined that $\hat{p}_i$'s are approximate critical points of p_{i+1}'s
3. 1 and 2 guarantees that p_N can be approximated efficiently

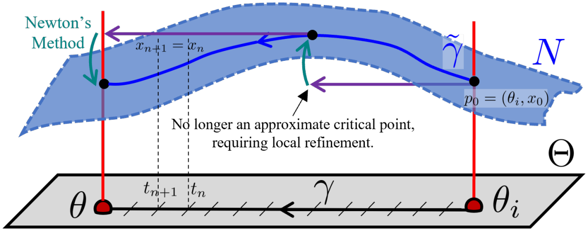


Table of Contents

1. Real-time optimisation problems are formulated as fibre bundle optimisation problems.
$$f : \Theta \times X \to \mathbb{R}$$
2. The fibre-wise critical points lie on a finite number of submanifolds N_i, which allow a smooth search through lifts of γ.
3. It is possible to pre-calculate critical points of a set $\{\theta_1, ..., \theta_n\}$, along with the bounds for approximate critical points, so that the real-time optimisation can be done efficiently.

Technische Universität München

[Man12] Jonathan H. Manton. *Optimisation Geometry*. 2012. arXiv: 1212.1775 [math.OC]. URL: https://arxiv.org/abs/1212.1775.

[PM17] Michael Pauley and Jonathan H Manton. "Optimisation geometry and its implications for optimisation algorithms". In: *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. 2017, pp. 1–5. DOI: 10.1109/CAMSAP.2017.8313169.