Update (The neighborhood of retraction convexity)

The current goal is to understand retraction-convexity, in particular a property satisfied by retraction-convex functions over compact manifolds. To begin with the definition of retraction convexity is stated below.

Retraction-Convexity: For a function $f : M \rightarrow R : x \rightarrow f(x)$ on a Riemannian manifold with retraction R define a function $h(t) = f(R_x(t\eta_x));\ x \in M,\ \eta_x \in T_xM$. The function f is retraction convex wrt retraction R in a set $\mathcal{N}$ if for all $x \in \mathcal{N}$, $\eta_x \in T_xM$ and $||\eta_x|| = 1$, $h(t)$ is convex for all t which satisfies $R_x(t\eta_x) \in \mathcal{N}$.

Intuitively the definition of retraction convexity is a relaxation of the definition of geodesic convexity which is a generalization of the standard definition of convexity for functions from $\mathbb{R}^n \rightarrow \mathbb{R}$. Geodesic convexity replaces the straight lines in euclidean spaces to straight line analogue on general smooth manifolds, that is, geodesics. The definition of retraction convexity relaxes this to retraction curves which are locally similar to geodesics upto first order. This is crucial for computations as geodesics are hard to compute.

Let us derive some consequences of the above definition using known properties of convex functions. As per the definition we must have

$$h(t) = f \circ R_x(t\eta_x) : \mathbb{R} \rightarrow \mathbb{R} \tag{1}$$

to be a convex function using the standard notion of convexity for real valued functions. must however hold for every $x \in \mathcal{N}$ and for every such that . We then have by the non negativity of second derivative of convex functions,

$$h^{''}(t) = \eta_x^T \nabla^2 f \circ R_x(t\eta_x)\eta_x \geq 0,\ \ \forall x \in \mathcal{N}\ \ \forall \eta_x \in T_xM \tag{2}$$

Equivalently, the hessian $\nabla^2 f \circ R_x(t\eta_x)$ must be positive semidefinite. Thus, the eigen values of the hessian are non negative and hence $\lambda_{min} \geq 0$. We now look at an important property we are investigating.

Property being studied: For retraction convex functions, it is known that there exists a neighborhood around a critical point where the function $f(x)$ is retraction convex. however the size of the nbh is not known and our goal is to investigate numerically, how big is the size of the nbh on which $h^{''}(t) \geq 0$.

Investigation: We will therefore consider an objective function on a compact manifold whole optimal(critical) point/s are known. We will then consider a retraction curve at the optimal(critical) point and form the function $f \circ R_x$. The idea then is to look at the eigen value of the second derivative or hessian of the function $f \circ R_x$. To study this we consider the following optimization problem on Stiefel manifold whose solution(critical point) is known to us. The problem is the Principle Component Analysis(PCA) problem. Let $X \in \mathbb{R}^{d \times n}$ be the matrix whose columns are data points. Then the task is to find the main directions of variation of the data. In general this involves looking for k principle components(directions) that are mutually orthogonal to each other. So we look for a matrix $U \in \mathbb{R}^{d \times k}$ with k orthonormal columns $u_1, \ldots u_k \in \mathbb{R}^d$.

The set of such matrices is called a Stiefel manifold denoted as $St(d,k)$.

$$St(d,k) = \{U : \mathbb{R}^{d\times k}\ U^T U = I_k\}$$

The PCA amounts to solving the problem

$$\min_{U\in St(d,k)} \sum_{i=1}^{k} \alpha_i \langle XX^T u_i\,, u_i\rangle = \langle XX^T U\,, UD\rangle \qquad (\mathcal{P}_1)$$

where $D \in \mathbb{R}^{k\times k}$ is the diagonal matrix with entires $\alpha_1 > \ldots > \alpha_k$.

Solution: The k top Eigen vectors of XX^T yield the global optimum. So we can say the optimal solution to the optimization problem $\mathcal{P}$ on the Stiefel manifold is the matrix $U \in St(d,k)$ where the k columns of U are the k-eigen vectors of the matrix XX^T.

Retractions on Stiefel Manifold Now there are two popular retractions for Stiefel manifolds. These are the Q-factor retraction given as

$$R_X(V) = Q$$

where $QR = X + V$, $Q \in St(d,k)$ and the polar retraction

$$R_X(V) = (X+V)(I_p + V^T V)^{-\frac{1}{2}}$$

We use the Q factor retraction initially as it is simpler to work out the algebra with the Q-factor retraction.

Composition $f \circ R_x$: So we have the following composition $f \circ R_x$ whose hessian we must check for positive semidefiniteness.

$$f \circ R_x(tV) = \langle XX^T tQ, tQD\rangle \;\; QR = X + V$$

Challenge/Question ?? The challenge is to find the hessian $\nabla^2 f \circ R_x$. Note that at the critical point, that is, the matrix $U \in St(d,k)$ formed by the k-eigen vectors of the matrix XX^T we have the following

$$Hess(f(x)) = Hess(f \circ R_x(0))$$

1. To find the hessian as a function of t.

2. Reverse engineer the size of nbh $\mathcal{N}$, instead of choosing the size of and then find the vales of t, that is an interval $t \in (\epsilon, \epsilon)$ such that $R_x(t\eta_x)$, we could instead choose an interval for t which would give rise to certain nbh $\mathcal{N}$.

I am working towards Dr Wen's suggestions and explanation from 06/08 meeting. As explained, the goal is to look at the eigen values of the hessian of $f(R_x(t\eta))$, where η is any tangent vector at the point x on the manifold. In this problem the points are matrices V and so we have the function $f \circ R_X(tV)$. This is where some insights given in Chapter 7 of Boumal's book come into picture. In this chapter a geometric toolbox has been developed and consolidated for several embedded manifolds. This includes first and second order geometries including a discussion of the hessian of the functions from a given manifold to $\mathbb{R}$, say $f : St(n,p) \to \mathbb{R}$. I am currently focusing on the Stiefel manifold from this chapter and putting all the pieces together. My immediate goal is to use an objective function and use the polar retraction to come up with an expression for the function $f \circ R_x$. The objective function f I am trying to work with is the from the PCA problem as explained on page 20 of Boumal's book. of A crucial skill needed to tackle these problems was a thorough understanding of the notion of second derivative and connections as well as an understanding of the geodesic equation including the calculation of the Christoffel symbols. These took a long time for me to fully understand and grasp. In fact now I am able to appreciate a discussion of these topics from Boothby's book's second last chapter. Some remaining pieces in the puzzle include the tool box for specific manifolds (chapter 7 of Boumals book) and that is what I am finishing now.

Update

Working with the PCA problem as explained on page 20 of Boumal's book.

Goal is to use an objective function and use the polar retraction to come up with an expression for the function $f \circ R_x$.

In parallel, started looking at interior point methods for non convex problems and their implementations.

Hello Nita. This is Tejas here and I am a PhD student at Florida State University. I came across your profile and felt a connection. Feel free to respond back.