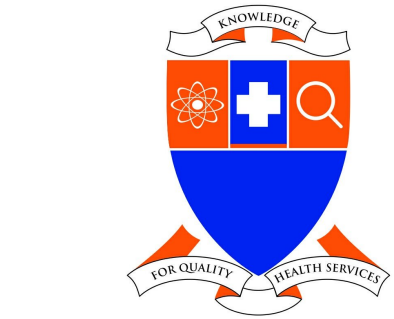
Optimization on Riemannian manifold

Ndlovu PV

School of Science and Technology Department of Mathematics and Applied Mathematics Sefako Makgatho Health Sciences University Mathematics



September 7, 2026

Outline

- Background and motivation
- Optimization in Euclidean space
- Fundamentals of Riemannian optimization
- Classical and Riemannian gradient methods
- References

What Is an Optimization Problem?

At the most basic level, an optimization problem consists of:

a search space B; an objective function. $f : B \to \mathbb{R}$

The general problem is

$$\boxed{\min_{x \in B} f(x)}$$

These two objects completely specify the optimization problem.

What structure does the search space have

$$\boxed{\min_{x \in B} f(x)}$$

Question:

What if $B = \mathbb{R}^n$.

This naturally introduces the Euclidean optimization.

Euclidean optimization

Once our search space is $\mathbb{R}^n$, we inherit a lot of useful structure. We know how to measure distances, angles, directions, and we can calculate gradients and Hessians

Euclidean Optimization

The classical unconstrained problem is

$$\min_{x \in \mathbb{R}^n} f(x)$$

.

When f is differentiable, Euclidean geometry provides:

$\nabla f(x)$ and, when twice differentiable, $\nabla^2 f(x)$.

Euclidean Optimization continuous...

These objects form the foundation of methods such as:

(i) Gradient descent

(ii) Newton's method

(iii) Conjugate-gradient methods

(iv) Momentum and inertial methods

Gradient Descent

$$\boxed{x_{k+1} = x_k - \alpha_k \nabla f(x_k)}$$

The gradient tells us the direction of steepest ascent. Therefore, if our objective is minimization, we move in the opposite direction.

But how far should we move?

Example

$f(x, y) = x^2 + 4y^2$

with

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ 8y \end{bmatrix}.$$

If we are standing at

$$(4, 1)$$

in which direction should we move.

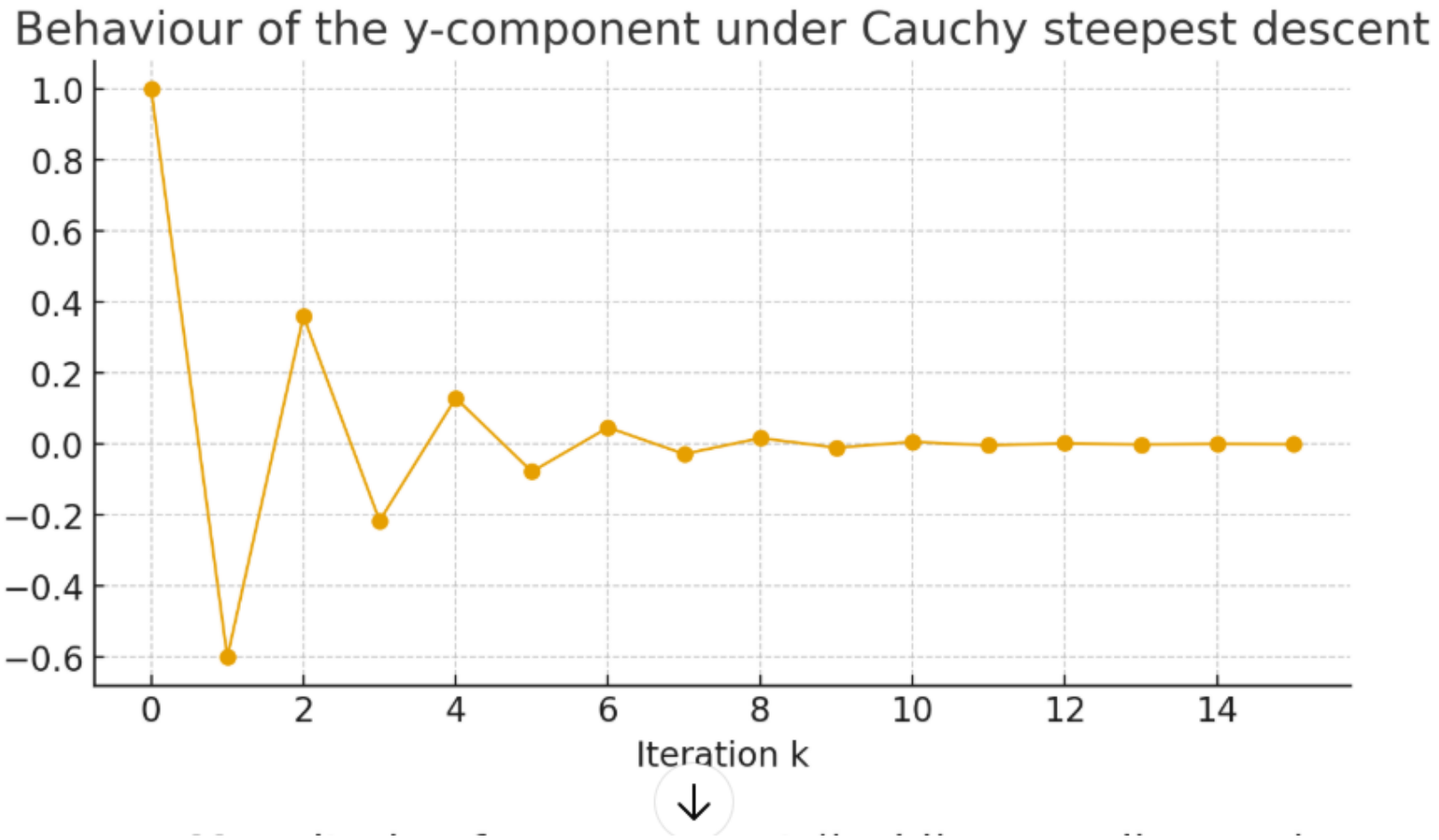
Example continuous

$$\nabla f(4,1) = \begin{bmatrix} 8 \\ 8 \end{bmatrix}.$$

The descent direction is

$$-\nabla f(4,1) = \begin{bmatrix} -8 \\ -8 \end{bmatrix}.$$

The problem with gradient descent



Inertia: learning from the previous step

Can we use information from where we came from?

Polyak [4].

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) + \beta(x_k - x_{k-1}). \tag{1}$$

The last term is an inertial term. Instead of completely forgetting our previous movement, we allow the algorithm to carry some momentum into the next iteration.

Extension of optimization on manifold

The search space is smooth manifold :

$B = M$

Then we can redefine our problem as follows

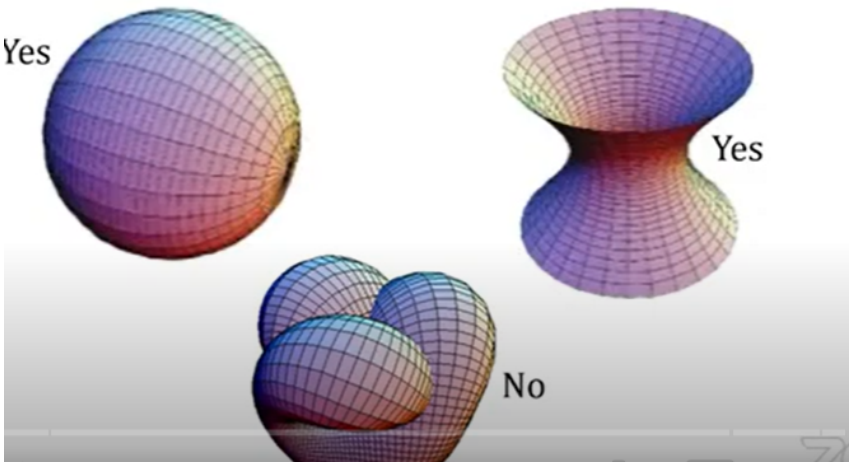
$$\min_{x \in \mathbb{M}} f(x)$$

One thing nice that we can do is to add structure on $\mathbb{M}$ to turn it into Riemannian manifold.

What is manifold

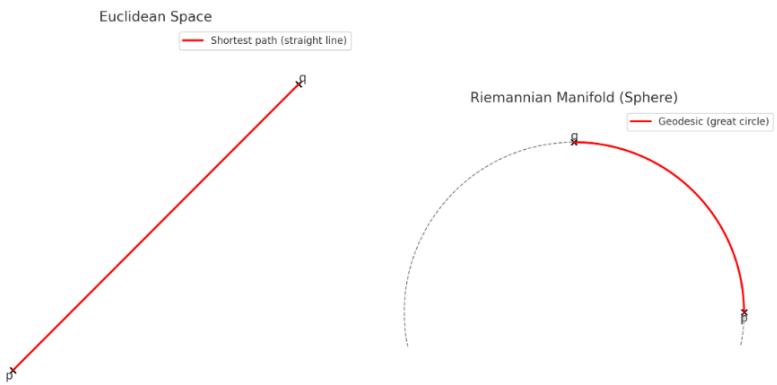
Definition

[9, 8] A manifold is a space that locally looks like Euclidean space.



Riemannian geometry

To clearly explore the Riemannian structure, it is essential to begin by understanding how distance between two points is defined in a Riemannian manifold. Unlike in Euclidean space, where distance is measured by straight lines, a Riemannian manifold may be curved, and the shortest path between two points must follow the manifold's geometry. This path is known as a geodesic, it is determined by the Riemannian metric, which assigns an inner product to each tangent space and enables the computation of lengths and angles.



Tools needed on manifold: Riemannian Metric, Tangent vector, Tangent space, Retraction.

$$R_x : T_xM \to M$$

such that

1. $R_x(0) = x$ (the zero tangent vector maps back to the point itself),

and

2. The differential at 0 satisfies $dR_x(0) = idT_xM$

The sphere as our first example

We now consider the unit sphere, defined by

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : ||x||^2 = x^T x = 1\}$$

What is the tangent space of $\mathbb{S}^{n-1}$

At a point x, the tangent space is the set of all vectors that are perpendicular to x, Mathematically

$$T_x\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n : x^T v = 0\}$$

So, clearly, the tangent space is like a flat ground, and all the operations that we can perform in Euclidean space can also be performed on this space and then mapped back to the manifold.

Retraction on the sphere

Given $x \in \mathbb{S}^{n-1}$ and a tangent vector $v \in T_x M$, a common retraction is

$$R_x(v) = \frac{x+v}{||x+v||}.$$

This maps $x+v$ back into the sphere by normalization. It satisfies $R_x(0) = x$, and its differential at zero is the identity.

How do we move.

We now know the direction in which to move. But there is another problem: the gradient belongs to the tangent space, while our solution must belong to the manifold. How do we get back?

$$\exp_x : T_xM \to M$$

then,

$$x_{k+1} = \exp_{x_k}(-\alpha_k \text{grad}\, f(x_k))$$

Now we have two questions

Can we use previous information to accelerate the method?

Can we choose the step size automatically?

Thank you for you attention....


Ndlovu PV, Jolaoso LO, Aphane M, Abass HA: Viscosity extragradient with modified inertial method for solving equilibrium problems and fixed point problem in Hadamard manifold. Journal of Inequalities and Applications. 2024 Feb 13;2024(1):21.


Hu J, Liu X, Wen ZW, Yuan YX: A brief introduction to manifold optimization. Journal of the Operations Research Society of China. 2020 Jun;8(2):199-248.


Cauchy, A.L. (1847). Methode gene'rale pour la resolution des syste'mes d'equations simultane'es. Comptes Rendus de l'Acade'mie des Sciences, 25, 536-538.


Polyak BT. Some methods of speeding up the convergence of iteration methods. Ussr computational mathematics and mathematical physics. 1964 Jan 1;4(5):1-7.


O.P Ferreira, AN Iusem, and SZ Nemeth: Concepts and techniques of optimization on the sphere. In: TOP 22 (2014), pp. 1148â1170.


Hu, J., Liu, X., Wen, Z.W. and Yuan, Y.X., 2020. A brief introduction to manifold optimization. Journal of the Operations Research Society of China, 8(2), pp.199-248.


Absil PA, Mahony R, Sepulchre R. Optimization algorithms on matrix manifolds. Princeton University Press; 2008 Dec 31.


H. A. Abass, O. K. Oyewole, K. O. Aremu and L. O. Jolaoso, Modified inertial Tseng method for solving variational inclusion and fixed point problems on Hadamard manifolds, *Applicable Analysis*, (2023), 1-24.


Martinet, B., 1970. Breve communication. Regularisation d ’inequations variationnelles par approximations successives.

Revue franÃ§aise d informatique et de recherche operationnelle.

Serie rouge, 4(R3), pp.154-158.