

OPTIMISATION SCHEMES ON MANIFOLDS OBTAINED VIA DISCRETISATIONS OF CONFORMAL HAMILTONIAN SYSTEMS

Marta Ghirardelli

NTNU, Norway

marta.ghirardelli@ntnu.no

Over the last few decades, optimisation methods for problems on manifolds have been gaining more and more popularity [1], [2]. In many cases, they can be seen as an alternative to solving constrained optimisation problems in Euclidean spaces. More precisely, the manifold setting allows to include the constraints in a more natural geometric way. When considering applications to deep learning, constraints on the parameter space may sometimes be useful to prevent the problem of exploding or vanishing gradients, which often occurs e.g. during the training of recurrent neural networks. It would then be desirable to choose compact manifolds, such as the Stiefel manifold or the orthogonal group.

Still in the fields of machine learning and deep learning, Gradient Descent (GD) algorithms are widely used as they require only first order information of the function f to be minimised, and they are easy to implement. They are hence suitable for minimising e.g. cost or loss function in such contexts, where one has to deal with large amounts of data. Unfortunately, these methods are proved to achieve linear convergence only, as long as f is smooth and strongly convex, and can consequently be rather slow. Accelerated versions have been studied, especially in the Euclidean setting where we find Nesterov's accelerated gradient, Polyak's heavy ball and a relativistic gradient descent method. The last two of these, in particular, can be seen as the discretisation of some conformal Hamiltonian systems [3] where f is seen as the potential energy. We want to consider the corresponding systems when f is defined on a manifold M . To do so we first formulate the Hamiltonian function as the sum of the potential energy f and a momentum dependent kinetic energy. The related Hamiltonian vector field is defined on the cotangent bundle of M , namely T^*M , where an exact natural symplectic form ω can be defined. The manifold (T^*M, ω) is then called exact, and the conformal Hamiltonian system is recovered by adding to the conservative vector field, a dissipative one, the Liouville vector field [4]. The flow of this vector field can then be approximated via splitting methods, composing the conservative flow with the dissipative one.

[1] P-A Absil, Robert Mahony, and Rodolphe Sepulchre. Optimisation algorithms on matrix manifolds, 2009.

[2] Nicolas Boumal. An introduction to optimisation on smooth manifolds, 2023.

[3] Guilherme França, Jeremias Sulam, Daniel Robinson, and René Vidal. Conformal symplectic and relativistic optimisation, 2020.

[4] Robert McLachlan and Matthew Perlmutter. Conformal hamiltonian systems, 2001.