Ph.D. thesis abstract

Ph.D. candidate: Simone Rebegoldi

School of graduate studies: Mathematics

Scientific Disciplinary Sector: MAT/08 (Numerical Analysis)

Thesis title: Variable metric line–search based methods for nonconvex optimization

Advisor: Dr. Marco Prato

Co-advisor: Dr. Silvia Bonettini

Academic years: 2014/2016

The aim of this thesis is to propose novel iterative first order methods tailored for a wide class of nonconvex nondifferentiable optimization problems, in which the objective function is given by the sum of a differentiable, possibly nonconvex function and a convex, possibly nondifferentiable term. Such problems have become ubiquitous in scientific applications such as image or signal processing, where the first term plays the role of the fit-to-data term, describing the relation between the desired object and the measured data, whereas the second one is the penalty term, aimed at restricting the search of the object itself to those satisfying specific properties. Our approach is twofold: on one hand, we accelerate the proposed methods by making use of suitable adaptive strategies to choose the involved parameters; on the other hand, we ensure convergence by imposing a sufficient decrease condition on the objective function at each iteration.

Our first contribution is the development of a novel proximal–gradient method denominated Variable Metric Inexact Line–search based Algorithm (VMILA). The proposed approach is innovative from several points of view. First of all, VMILA allows to adopt a variable metric in the computation of the proximal point with a relative freedom of choice. Indeed the only assumption that we make is that the parameters involved belong to bounded sets. This is unusual with respect to the state-of-the-art proximal–gradient methods, where the parameters are usually chosen by means of a fixed rule or tightly related to the Lipschitz constant of the problem. Second, we introduce an inexactness criterion for computing the proximal point which can be practically implemented in some cases of interest. This aspect assumes a relevant importance whenever the proximal operator is not available in a closed form, which is often the case. Third, the VMILA iterates are computed by performing a line–search along the feasible direction and according to a specific Armijo-like condition, which can be considered as an extension of the classical Armijo rule proposed in the context of differentiable optimization.

The second contribution is given for a special instance of the previously considered optimization problem, where the convex term is assumed to be a finite sum of the indicator functions of closed, convex sets. In other words, we consider a problem of constrained differentiable optimization in which the constraints have a separable structure. The most suited method to deal with this problem is undoubtedly the nonlinear Gauss-Seidel (GS) or block coordinate descent method, where the minimization of the objective function is cyclically alternated on each block of variables of the problem. In this thesis, we propose an inexact version of the GS scheme, denominated Cyclic Block Generalized Gradient Projection (CBGGP) method, in which the partial minimization over each block of variables is performed inexactly by means of a fixed number of gradient projection steps.

The novelty of the proposed approach consists in the introduction of non Euclidean metrics in the computation of the gradient projection. As for VMILA, the sufficient decrease of the function is imposed by means of a block version of the Armijo line–search.

For both methods, we prove that each limit point of the sequence of iterates is stationary, without any convexity assumptions. In the case of VMILA, strong convergence of the iterates to a stationary point is also proved when the objective function satisfies the Kurdyka–Łojasiewicz property. Extensive numerical experience in image processing applications, such as image deblurring and denoising in presence of non-Gaussian noise, image compression, phase estimation and image blind deconvolution, shows the flexibility of our methods in addressing different nonconvex problems, as well as their ability to effectively accelerate the progress towards the solution of the treated problem.