Dedicated to Professor Yeol Je Cho on the occasion of his retirement

A note on the accelerated proximal gradient method for nonconvex optimization

HUIJUAN WANG and HONG-KUN XU

ABSTRACT.

We improve a recent accelerated proximal gradient (APG) method in [Li, Q., Zhou, Y., Liang, Y. and Varshney, P. K., *Convergence analysis of proximal gradient with momentum for nonconvex optimization*, in Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017] for nonconvex optimization by allowing variable stepsizes. We prove the convergence of the APG method for a composite nonconvex optimization problem under the assumption that the composite objective function satisfies the Kurdyka-Łojasiewicz property.

REFERENCES

[1] Attouch, H., Bolte, J., Redont, P. and Soubeyran, A., *Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality*, Math. Operations Research, **35** (2010), No. 2, 438–457

[2] Attouch, H., Bolte, J. and Svaiter, B. F., *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Math. Program., **137** (2013), 91–129

[3] Beck A. and Teboulle, M., *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., **2** (2009), No. 1, 183–202

[4] Bolte, J., Daniilidis, A. and Lewis, A., *The Łojasiewicz inequality for nonsmooth sub-analytic functions with applications to subgradient dynamical systems*, SIAM J. Optim., **17** (2007), 1205–1223

[5] Bolte, J., Sabach, S. and Teboulle, M., *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Math. Program., Ser. A , **146** (2014), 459–494

[6] Combettes, P. L. and Wajs, R., *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., **4** (2005), No. 4, 1168–1200

[7] Frankel, P., Garrigos, G. and Peypouquet, J., *Splitting methods with variable Metric for Kurdyka-Łojasiewicz functions and general convergence rates*, J. Optim. Theory Appl., **65** (2015), No. 3, 874–900

[8] Kurdyka, K., *On gradients of functions definable in o-minimal structures*, Ann. Inst. Fourier (Grenoble), **48** (1998), No. 3, 769–783

[9] Łojasiewicz, S., *Une propriete topologique des sous-ensembles analytiques reels*, in: *Les Euations aux Derivees Partielles*, Editions du centre National de la Recherche Scientifique, Paris, 1963, pp. 87–89

[10] Li, Q., Zhou, Y., Liang, Y. and Varshney, P. K., *Convergence analysis of proximal gradient with momentum for nonconvex optimization*, in Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017

[11] Nesterov, Y. E., *Introductory Lectures on Convex Optimization: a basic course* Kluwer Academic Publishers, Massachusetts, 2004

[12] Rockafellar, R. T. and Wets, R. J.-B., *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften, vol. **317**, Springer, Berlin, 1998

DEPARTMENT OF MATHEMATICS
SCHOOL OF SCIENCE
HANGZHOU DIANZI UNIVERSITY
HANGZHOU 310018 CHINA
E-mail address: 2805065050@qq.com
E-mail address: xuhk@hdu.edu.cn