Extending Projection-Free Optimization Algorithms

Antonio SILVETI-FALLS

Project Description

The Frank-Wolfe [5] or conditional gradient [7] algorithm is an algorithm for solving constrained optimization problems, with many applications in machine learning and data science [2]. It requires only first-order information about the function to be minimized and a linear minimization oracle over the constraint set, which can be much cheaper or easier to compute than the projection onto the constraint set (which is a quadratic minimization oracle). For many common constraint sets founds in practical problems, variants of the Frank-Wolfe algorithm offer a scalable solution to solve constrained optimization problems [4].

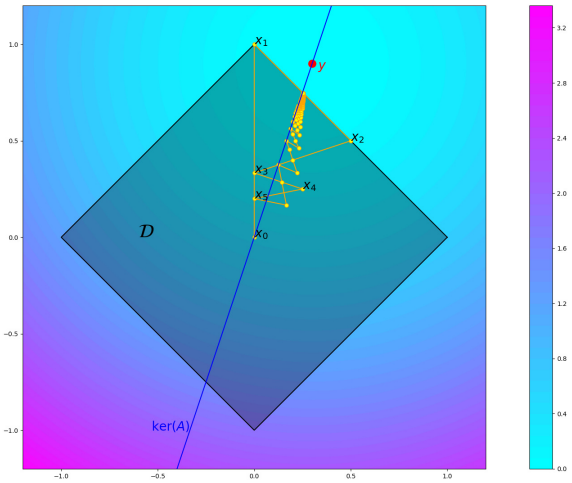

Figure 1: A nonsmooth variant of the Frank-Wolfe algorithm for solving $\min_{x:Ax=0,\|x\|_1\leq 1} \frac{1}{2}\|x-y\|^2$

While classical Frank-Wolfe algorithms are stuck with a sublinear convergence rate due to their cheap oracle, recent innovations like the boosted Frank-Wolfe algorithm [3] have improved the global convergence rate by calling the linear minization oracle multiple times per iteration. However, these advances are limited to the classical Frank-Wolfe setting, e.g., continuously differentiable objective functions with Lipschitz-continuous gradients. Many problems of interest are actually outside of this framework, for instance nonsmooth objective functions like the ℓ^1 norm [9, 11, 6, 1] or stochastic objective functions like $\mathbb{E}_{\xi\sim\Xi}(f_\xi)$ [10, 12, 8].

The objective of this internship is thus to explore extensions of these faster Frank-Wolfe algorithms that will allow for nonsmooth or stochastic objective functions, as well as their numerical implementation in python or MATLAB and their application to problems coming from machine learning and data science.

Keywords

Constrained optimization, convex optimization, stochastic optimization, machine learning, Frank-Wolfe, Conditional Gradient, data science.

Logistics

This internship will be carried out in the Center for Visual Computing (CVN) laboratory at CentraleSupélec. It is not in partnership with any companies. Interested candidates should contact Antonio Silveti-Falls (antonio.silveti-falls@centralesupelec.fr).

References

[1] Andreas Argyriou, Marco Signoretto, and Johan Suykens. Hybrid conditional gradient-smoothing algorithms with applications to sparse and low rank regularization. *Regularization, Optimization, Kernels, and Support Vector Machines*, pages 53–82, 2014.

[2] Gábor Braun, Alejandro Carderera, Cyrille W Combettes, Hamed Hassani, Amin Karbasi, Aryan Mokhtari, and Sebastian Pokutta. Conditional gradient methods. *arXiv preprint arXiv:2211.14103*, 2022.

[3] Cyrille Combettes and Sebastian Pokutta. Boosting frank-wolfe by chasing gradients. In *International Conference on Machine Learning*, pages 2111–2121. PMLR, 2020.

[4] Cyrille W Combettes and Sebastian Pokutta. Complexity of linear minimization and projection on some sets. *Operations Research Letters*, 49(4):565–571, 2021.

[5] Marguerite Frank, Philip Wolfe, et al. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[6] Gauthier Gidel, Fabian Pedregosa, and Simon Lacoste-Julien. Frank-wolfe splitting via augmented lagrangian method. In *International Conference on Artificial Intelligence and Statistics*, pages 1456–1465. PMLR, 2018.

[7] Evgeny S Levitin and Boris T Polyak. Constrained minimization methods. *USSR Computational mathematics and mathematical physics*, 6(5):1–50, 1966.

[8] Geoffrey Négiar, Gideon Dresdner, Alicia Tsai, Laurent El Ghaoui, Francesco Locatello, Robert Freund, and Fabian Pedregosa. Stochastic frank-wolfe for constrained finite-sum minimization. In *international conference on machine learning*, pages 7253–7262. PMLR, 2020.

[9] Antonio Silveti-Falls, Cesare Molinari, and Jalal Fadili. Generalized conditional gradient with augmented lagrangian for composite minimization. *SIAM Journal on Optimization*, 30(4):2687–2725, 2020.

[10] Antonio Silveti-Falls, Cesare Molinari, and Jalal Fadili. Inexact and stochastic generalized conditional gradient with augmented lagrangian and proximal step. *Journal of Nonsmooth Analysis and Optimization*, 2(Original research articles), 2021.

[11] Alp Yurtsever, Olivier Fercoq, Francesco Locatello, and Volkan Cevher. A conditional gradient framework for composite convex minimization with applications to semidefinite programming. In *International Conference on Machine Learning*, pages 5727–5736. PMLR, 2018.

[12] Alp Yurtsever, Suvrit Sra, and Volkan Cevher. Conditional gradient methods via stochastic path-integrated differential estimator. In *International Conference on Machine Learning*, pages 7282–7291. PMLR, 2019.