Paper Review: Revisiting Frank-Wolfe: Projection Free Sparse Convex Optimization

Brief Summary:

As the paper title suggests, the authors primarily focus on using the Frank-Wolfe (FW) algorithm and its variants to solve constrained convex optimization problems. The authors begin by defining a surrogate duality gap for the constrained convex optimization problem using the convexity property for the function to be minimized and suggest the use of the duality gap as a stopping criterion for the iterative algorithm. The authors then introduce three variants of the original Frank-Wolfe algorithm that solve the linear subproblems in the iterative steps approximately and could fasten the iterative minimizing process. Then, by defining the curvature constant (C_f) for a convex and differentiable function, the authors provide proof of primal convergence for the original FW algorithm and its variants. However, for practical applications, the optimal solution and curvature constant are unknown; therefore the authors introduce a theorem that bounds the surrogate duality gap showing that it must reduce over iterations in order to converge close to the optimal value.

Since the curvature constant is invariant to affine transformations, the authors are also able to show that the FW algorithm, its variants and the convergence analysis hold true even when such transformations are applied. Furthermore, the authors also show that the solution obtained through the FW algorithm is optimal in terms of sparsity by providing a lower bound on the primal and dual errors. The authors then provide applications of the iterative FW algorithms to optimize over atomic sets, vectors as well as norms with structural constraints. As a unique application, the authors use the FW algorithms for convex optimization over bounded matrix trace norm and show that for convex optimization over matrix factorizations, the FW algorithm provides a low-rank solution at every iteration.

Paper Contributions:

The contributions of this paper are multi-fold:

1. The authors prove convergence for the original and the variants of the FW algorithm by introducing a surrogate duality gap and guaranteeing its small magnitude.
2. The authors show that the convergence analysis as well as the variants of the FW algorithm are invariant to affine transformations applied to the optimization problem.
3. The FW algorithms are shown to deal with relatively less expensive linear subproblems for each iteration
4. The authors directly apply the FW algorithm to solve convex optimization problems in the atomic domain.
5. The authors are able to successfully apply the FW algorithms to solve convex optimization problems over the bounded matrix-norm unlike previous approaches such as proximal methods and gradient descent.
6. The authors prove that for convex optimization over matrix factorizations, each FW algorithm iteration outputs a low-rank update

Additional Comments:

While the original FW algorithm was not introduced by the authors, they were successfully able to guarantee convergence through the introduction of a surrogate duality gap and provide applications for the algorithm to obtain low rank and sparse solutions using only linear subproblem approximations.