First order methods for relatively-smooth convex optimization

Abstract: The first order methods gained a strong research focus triggered by the recent data scale growth. It is widely known that most of the first order algorithms for convex optimization are based on uniform smoothness assumption on the objective function, i.e. the gradient of the objective function is Lipschitz continuous. However, in many application settings, such as in D-optimal design problem, the differentiable convex objective function is not uniformly smooth. Therefore, we present more relaxed notions like "relative"-smoothness and "relative"-strong convexity, which are determined by a user-specified (computationally tractable) "reference" function. Based on these concepts, it can be shown that many differentiable convex functions are relatively smooth with respect to a correspondingly fairly-simple reference function. We present the extension to this setting of the most basic first order method – the primal gradient scheme. Lastly, we provide some open problems in relation with the acceleration of the gradient scheme and with the dual gradient methods under the "relative"-smoothness assumption.