Project Proposal

Caleb Ki & Sanjana Gupta

November 15, 2017

Background

Templates for first-order conic solvers, or TFOCS for short, is a software package that provides efficient solvers for various convex optimization problems (Becker, Candès, and Grant 2011). With problems in signal processing, machine learning, and statistics in mind, the creators developed a general framework and approach for solving convex cone problems. The standard routine within the TFOCS library is a four-step process. The notation used below is taken from the paper accompanying the software package.

The first step is to express the convex optimization problem with an equivalent conic formulation. The optimization problem should follow the form:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & \mathcal{A}(x) + b \in \mathcal{K} \end{aligned} \tag{1}$$

In the problem above, $x \in \mathbb{R}^n$ is the vector we are optimizing under the convex objective function f. It is important to note that f does not have to be a smooth function. $\mathcal{A}$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, $b \in \mathbb{R}^m$ is simply a vector, and $\mathcal{K}$ is a convex cone in $\mathbb{R}^m$.

The two main hurdles to developing efficient first-order solutions for convex optimization problems of this form are that f is not restricted to the set of smooth functions and that finding feasible points under this conic constraint is often an expensive computation. These hurdles are circumvented by finding and solving the dual problem of the given convex optimization problem which leads us to the second step of the process, turning the problem into its dual form. The dual of the form given by equation 1 is as follows:

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & g(\lambda) \\ \text{subject to} \quad & \lambda \in \mathcal{K}^* \end{aligned} \tag{2}$$

g is simply the dual function given by $g(\lambda) = \inf_x f(x) - \sum_{i=1}^{m} \langle \mathcal{A}(x) + b, \lambda \rangle$, and $\mathcal{K}^* = \{\lambda \in \mathbb{R}^m : \langle \lambda, x \rangle \geq 0 \, \forall x \in \mathcal{K}\}$ is the dual cone.

The dual problem is not directly solved at this step. This is because the dual function is generally not differentiable for the class of problems we are considering. Further, directly using subgradient methods is not efficient since these methods converge very slowly. In order to convert this to a problem that can be efficiently optimized, we apply a smoothing technique which modifies the primal objective function and instead solves the following problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_\mu(x) \triangleq f(x) + \mu d(x) \\ \text{subject to} \quad & \mathcal{A}(x) + b \in \mathcal{K} \end{aligned} \tag{3}$$

Here μ is a positive scalar and $d(x)$ is a strongly convex function called the *proximity function* which satisfies $d(x) \geq d(x_0) + \frac{1}{2}\|x - x_0\|^2$ for some fixed $x_0 \in \mathbb{R}^n$. The dual of this problem is

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & g_\mu(\lambda) \\ \text{subject to} \quad & \lambda \in \mathcal{K}^* \end{aligned} \tag{4}$$

where g_μ is a smooth approximation of g. In many cases, the dual can be reduced to the following unconstrained problem:

$$\underset{z}{\text{maximize}} \quad -g_{sm}(z) - h(z) \tag{5}$$

This is known as the *composite form*. Here, $z \in \mathbb{R}^m$ is the optimization variable, g_{sm} is a smooth convex function and h is a non-smooth (possibly extended-value) convex function. Finally, we can efficiently solve both the smooth dual and the composite problems using optimal first order methods such as gradient descent.

For both the problems (4), (5), given a sequence of step sizes $\{t_k\}$, the optimization begins with a point $\lambda_0 \in \mathcal{K}^*$ and has the following updating rule respectively:

$$\lambda_{k+1} \leftarrow \underset{\lambda \in \mathcal{K}^*}{\arg\min}\|\lambda_k + t_k \nabla g_\mu(\lambda_k) - \lambda\|_2$$

$$z_{k+1} \leftarrow \underset{y}{\arg\min} \; g_{sm}(z_k) + \langle \nabla g_{sm}(z_k), z - z_k \rangle + \frac{1}{2t_k}\|z - z_k\|^2 + h(z)$$

The approximate primal solution can then be recovered from the optimal value of λ as follows:

$$x(\lambda) \triangleq \underset{x}{\arg\min} \; f(x) + \mu d(x) - \langle \mathcal{A}(x) + b, \lambda \rangle \tag{6}$$

Equation 6 shows that the fundamental computational primitive in this method is the efficient minimization of the sum of a linear term, a proximity function, and a non-smooth function.

Proposal

To summarize the background, TFOCS, when given the correct conic formulation of an optimization problem will apply various first-order solvers to the smoothed dual problem. The goal of this project is to port a portion of the TFOCS software package from Matlab into Python. TFOCS implements six different first-order methods. The default is the single-projection method developed by Auslender and Teboulle (2006). For our project we will be implementing the default algorithm as well as the dual-projection method developed by Lan, Lu, and Monteiro (2011). In particular, we would like our solvers to accept a pythonic version of the calling sequence found in the user manual written by Becker, Candes, and Grant (2014) which is: [x, out] = tfoc(smoothF, affineF, nonsmoothF, x0). Given a smooth function smoothF, an affine form specification affineF, a nonsmooth function nonsmoothF, and an initial point x0, the solver would return the optimal vector x and the primal solution out.

References

Auslender, Alfred, and Marc Teboulle. 2006. "Interior Gradient and Proximal Methods for Convex and Conic Optimization." *SIAM Journal on Optimization* 16 (3). Society for Industrial; Applied Mathematics: 697–725.

Becker, Stephen, Emmanuel J. Candes, and Michael Grant. 2014. "TFOCS User Guide." Stanford.

Becker, Stephen, Emmanuel J. Candès, and Michael Grant. 2011. "Templates for Convex Cone Problems with Applications to Sparse Signal Recovery." *Mathematical Programming Computation* 3 (3). John Wiley & Sons Ltd.: 165.

Lan, Guanghui, Zhaosong Lu, and Renato D. C. Monteiro. 2011. "Primal-Dual First-Order Methods with $\mathcal{O}(1/\epsilon)$ Iteration-Complexity for Cone Programming." *Mathematical Programming* 126 (1). Springer-Verlag: 1–29.