Implementation of interior-point methods for positive semi-definite Toeplitz cone programming

Matt Staib, Dieterich Lawson

March 21, 2015

1 Introduction

Consider the cone of positive-semidefinite Toeplitz matrices

$$\mathcal{T}_n = \{x \in \mathbf{R}^n \mid T_n(x) \geq 0\}$$

with $T_n(x)$ defined as

$$T_n(x)_{ij} = \begin{cases} x_{|i-j|+1} & 0 \leq |i-j| \leq n-1 \\ 0 & \text{otherwise} \end{cases}.$$

In words, $T_n(x)$ is the square Toeplitz matrix resulting from taking the ith entry of x as the ith sub- and super-diagonal. This cone appears in various applications in signal processing

Flesh out applications

. In this paper we investigate methods for cone programming over $\mathcal{T}$, i.e. we present methods for solving problems of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{T}_n. \end{array} \tag{1}$$

where $x \in \mathbf{R}^n$ is the optimization variable, $c \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, and $b \in \mathbf{R}^m$. The central difficulty for problems of this form is that $\mathcal{T}_n$ is not symmetric, so much of the standard primal-dual machinery becomes unweildly.

The remainder of this paper is organized as follows. We first explore the properties of $\mathcal{T}_n$, deriving its dual and showing that it is not symmetric. We discuss a suitable barrier for $\mathcal{T}_n$ and its use in two different algorithms that make speed gains over the standard primal-dual interior-point method by not calculating the derivatives of the dual barrier. Finally we present some numeric results from these algorithms.

2 The Positive Semi-Definite Toeplitz Cone

To characterize the dual of $\mathcal{T}_n$ we show that $\mathcal{T}_n$ is the dual of the cone of finite autocorrelation sequences. A finite autocorrelation sequence is a vector $x \in \mathbf{R}^n$ with a $y \in \mathbf{R}^n$ such that

$$x_k = \sum_{i=1}^{n-k} y_i y_{i+k} \quad k = 1, \ldots, n.$$

For any positive scalar α, αx can be represented as above by $\alpha^{1/2} y$, so the set of finite autocorrelation sequences forms a cone. The dual of $\mathcal{C}_n$ is

$$\mathcal{C}_n^* = \{z \in \mathbf{R}^n \mid z^T x \geq 0 \text{ for } x \in \mathcal{C}_n\}.$$

Because x is a finite autocorrelation sequence, $z^T x \geq 0$ for all $x \in \mathcal{C}_n$ if and only if

$$\sum_{i=1}^{n} z_i \left(\sum_{i=1}^{n-k} y_i y_{i+k} \right) \geq 0.$$

This summation can be expressed in terms of the Toeplitz form of z — the above is true if and only if $y^T T_n(z) y \geq 0$ for all $y \in \mathbf{R}^n$. Thus $\mathcal{C}_n^*$ is exactly the positive semi-definite Toeplitz cone, $\mathcal{T}_n$. Nonempty, closed, convex cones are reflexive, so $\mathcal{T}_n^* = \mathcal{C}_n$, and $\mathcal{T}_n$ is not symmetric.

The cone $\mathcal{T}_n$ posesses a θ self-concordant logarithmically homogenous barrier function

$$\phi_n(x) = -\log\det T_n(x).$$

A θ logarithmically homogenous function satisfies

$$\phi_n(tx) = -\theta \log t + \phi_n(x).$$

The cone $\mathcal{T}_n$ is of particular practical interest because there are fast methods for evaluating ϕ_n, its gradient, and its Hessian, as detailed in [AV02]. The motivation for this paper is the desire to solve cone problems like (1) without evaluating the gradient or Hessian of the dual barrier, which would be very expensive. Instead, we develop techniques that use only the derivatives of the primal barrier function which we can evaluate quickly using the methods given in [AV02].

3 The Barrier Method

We turn now to interior point methods and their suitability for solving un-symmetric cone problems. Consider the following problem, which is slightly more general than (1)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in K \end{array} \tag{2}$$

where all variables are as in (1), and K is a proper (closed, convex, solid, pointed) cone possessing a self-concordant and θ-logarithmically homogenous barrier function, ϕ. The dual of (2) is

$$\begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & c + A^T \lambda + \nu = 0 \\ & \nu \in K^* \end{array} \tag{3}$$

where $\lambda \in \mathbf{R}^m, \nu \in \mathbf{R}^n$ and K^* is the dual cone of K. We assume that (2) and (3) satisfy the interior point condition, i.e. that there exists a strictly feasible (x^0, λ^0, ν^0) that satisfies

$$Ax^0 = b,\ x^0 \in \mathbf{int}K,\ c + A^T\lambda^0 + \nu^0 = 0,\ \nu^0 \in \mathbf{int}K^*.$$

The *barrier method* is an interior-point method that solves (2) by using ϕ to make the cone constraint implicit in the objective. Using equality-constrained Newton's method, the barrier method solves a series of problems of the form

$$\begin{array}{ll} \text{minimize} & tc^T x + \phi(x) \\ \text{subject to} & Ax = b \end{array} \tag{4}$$

where $t \in \mathbf{R},\ t > 0$. For each $t > 0$, (4) has a unique minimizer [BV04], which defines the *primal central path*

$$x^*(t) = \arg\min_x \{tc^T x + \phi(x) \mid Ax = b\}.$$

Each point on the primal central path is associated with a dual-feasible point which can be used to lower-bound the optimal value of (2). To see this note that any minimizer of (4) $x^*(t)$ must satisfy the optimality conditions

$$tc + \nabla\phi(x^*(t)) + A^T y = 0,\ Ax^*(t) = b,\ x^*(t) \in K$$

for some $y \in \mathbf{R}^m$. It can be shown that if $x \in \mathbf{int}K$ then $\nabla\phi(x) \in K^*$ [AV02]. Thus, from the optimality conditions we see that $\nabla\phi(x^*(t))/t$, y/t are feasible for the dual, (3). Plugging y/t into the dual objective and subtracting it from the primal gives

$$c^T x^*(t) + b^T(y/t) = x^*(t)\nabla\phi(x^*(t))/t = \nu/t$$

This bounds the sub-optimality of $x^*(t)$ as

$$c^T x^*(t) - p^* \le c^T x^*(t) + b^T(y/t) = \theta/t$$

where p^* is the optimal value of (1). So, in summary we can solve a sequence of problems of the form (4) with equality-constrained Newton's method until θ/t is less than our desired accuracy.

A natural improvement would be to use a primal-dual interior point method, but such a method would require evaluating both the primal and dual barrier derivatives which as we discussed earlier would be computationally expensive. The barrier method presented here avoids this, allowing us to only evaluate the gradient and Hessian of the primal barrier which we can do efficiently.

4 A Predictor-Corrector Method

In [Nes06], Nesterov presents an algorithm for solving a general primal-dual cone problem, i.e. when the cone is not necessarily symmetric. This method involves a centering step based on the primal, and uses the dual barrier only in computing an appropriate step size. While the usual gradient and Hessian are needed for the primal barrier, we only need the value of the dual barrier. In fact, per Nesterov, the value of the dual barrier need not be terribly accurate.

In our case, for the standard log determinant barrier function for the semidefinite cone, we need to compute the conjugate

$$f_*(s) = \max_{x \in \mathcal{T}_n} \{ s^T x + \log\det(\text{toeplitz}(x)) \}. \tag{5}$$

For this barrier, in the case of the semidefinite cone, we have an explicit form for the conjugate. When we further constrain x to the Toeplitz cone, though, we no longer know of a closed-form solution. In practice we solve this using Newton's method.

The most computationally-intensive part of Nesterov's algorithm is the computation of a step size α such that $\Omega(z + \alpha\delta z) = \Omega(z) + C$ for a function $\Omega(z)$ involves a sum of the barrier and conjugate barrier. Again, we know of no closed-form solution, so we compute α using the bisection method.

5 Numerical Results

Our implementation is hosted on Github at `github.com/mstaib/toeplitz-ipm`. Unfortunately we were unable to produce a fully working variation of Nesterov's algorithm. We encountered difficulties with poorly-conditioned Hessians among other things. We tested the specific problem with $c = \mathbf{1} \in \mathbb{R}^4$, $A \in \mathbb{R}^{3\times 4}$, and $b = A \cdot (8, 4, 2, 1)^T$. Given a starting feasible point, our method managed 10 iterations in about 2 seconds, approximately converging to an objective value of 12.83 before this inexplicably jumped up to 14.7 on the 10th iteration. Each point produced was strictly feasible. In contrast, CVX took only 0.3 seconds to produce a better objective value of 11.04.

References

[AV02] Brien Alkire and Lieven Vandenberghe. Convex optimization problems involving finite autocorrelation sequences. *Mathematical Programming*, 93(3):331–359, 2002.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[Nes06] Yurii Nesterov. Towards nonsymmetric conic optimization. *Available at SSRN 921788*, 2006.