Fast Eigenvalue Optimization for Spectral-Element PDEs

Guillermo Angeris and John Sholar
EE364b: Convex Optimization II Class Project

Introduction

Eigenvalue optimization is a common task in many control problems, inverse design, and, recently, in inference problems on continuous spaces. In this project, we'll attempt to write a fast solver for eigenvalue problems arising from spectral element discretizations of eigenvalue partial differential equations (PDEs).

Spectral Element Methods and Laplacian Construction

A *spectral element method* (SEM)[1] is a finite-element method for solving PDEs which makes use of high-degree polynomials to approximate functions on a compact domain. The idea is to construct a quadrature rule such that all polynomials of lesser degree have exact integrals when sampled at a particular set of points and use these polynomials as a basis to approximate functions.

In general, these SEM-discretized problems—while sparse—are usually slow to solve in modern optimization packages for even a modest number of points in the discretization domain as they lack the tridiagonal structure that many PDE problems have (e.g., when discretizing by assuming $h^2 f''(x) \approx f(x+h) - 2f(x) + f(x-h)$ for fixed $h > 0$). In particular, the matrix generated by the spectral elements method is essentially a block-diagonal matrix with small blocks, each of which overlap with the previous block by exactly one element (see Figure), so no completely trivial decomposition can be applied.[2]

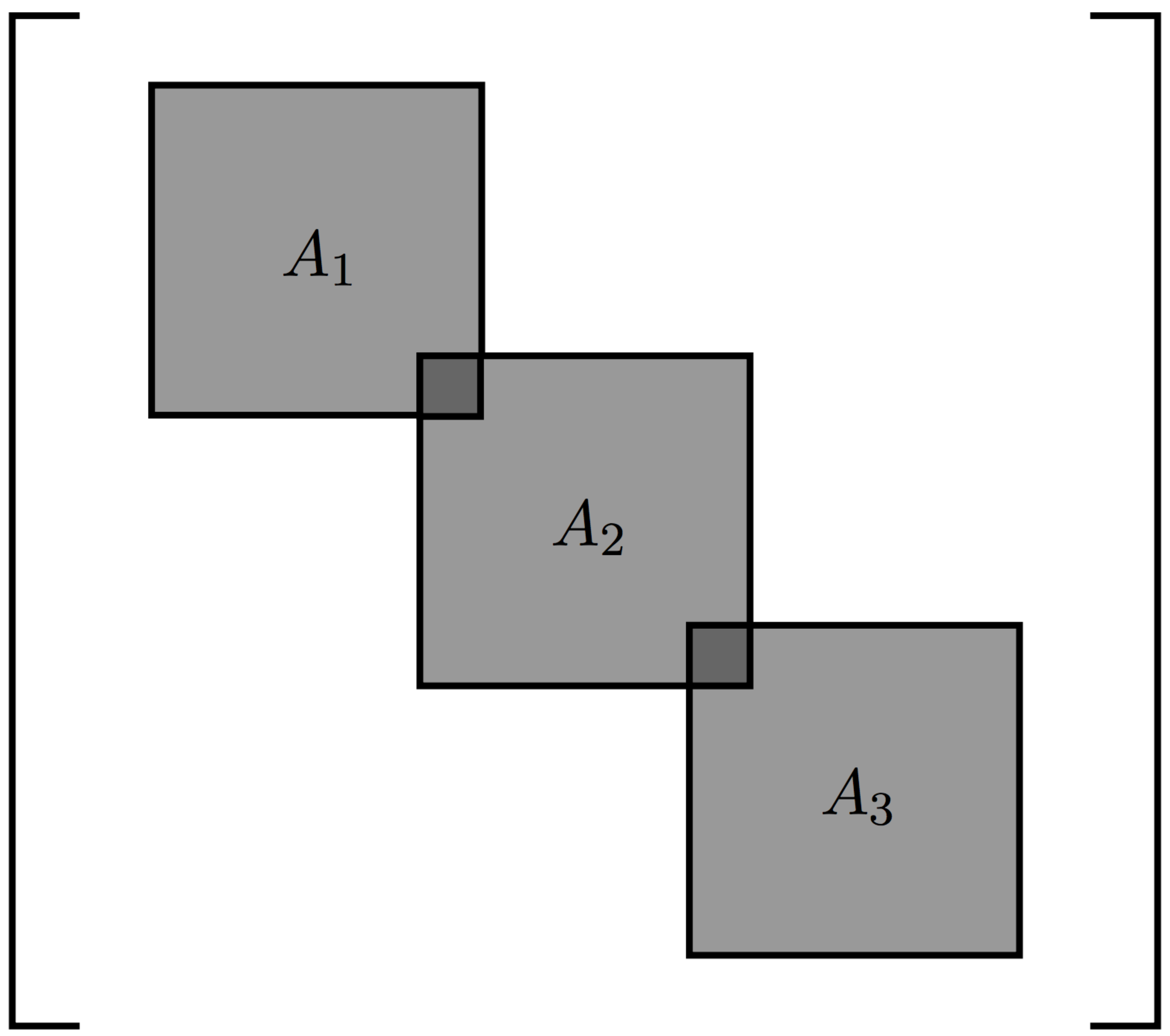

Figure 1: A simple example of the overlapping diagonal elements with three blocks. The darker box indicates a single overlapping element, e.g., $(A_1)_{nn} = (A_2)_{11}$ if $A_1 \in \mathbf{R}^{n\times n}$.

Our project will take advantage of the structure of these approximately-block-diagonal matrices for fast solutions of eigenvalue problems.

Approaches

Let $\mathscr{L} : \mathbf{R}^n \to \mathbf{S}^m$ be affine and have the approximately-block-diagonal structure described in figure and let $x \in \mathbf{R}^n$, then the problem of interest can be written as the dual of a standard-form SDP,

$$\begin{array}{ll} \underset{x,t}{\text{minimize}} & -b^T x + t \\ \text{subject to} & \mathscr{L}(x) \preceq tI, \\ & Cx = d, \\ & x \succeq 0. \end{array} \tag{1}$$

The first inequality is with respect to the semidefinite cone, while the latter is with respect to the positive orthant

This general optimization problem can be solved in several ways. We will attempt to explore three possibilities: interior point methods for solving these large-scale SDPs, first-order cone-splitting methods, and a consensus approach for solving these problems.

Interior Point Methods

For interior point methods, rewriting the problem using a barrier method is straightforward

$$\begin{array}{ll} \underset{x,t}{\text{minimize}} & -b^T x + t - \mu \log\det(tI - \mathscr{L}(x)) \\ \text{subject to} & Cx = d, \\ & x \succeq 0. \end{array}$$

First-Order Cone-Splitting Methods

In the second case, define $E_i \in \mathbf{R}^{d\times m}$ to be the 'selection matrix' given by

$$E_i = \underbrace{\begin{bmatrix} \underbrace{0_{d\times(d-1)} \;\cdots\; 0_{d\times(d-1)}}_{i-1} & I_d & 0_{d\times(d-1)} & \cdots & 0_{d\times(d-1)} \end{bmatrix}}_{b}.$$

where b is the number of sub-blocks of $\mathscr{L}$ and d is their dimension.

This allows us to rewrite (**??**) above to

$$\begin{array}{ll} \underset{x,t}{\text{minimize}} & -b^T x + t \\ \text{subject to} & \mathscr{L}(x) - tI = \sum_i E_i Z_i E_i^T, \\ & Cx = d, \\ & x \succeq 0, \\ & Z_i \succeq 0,\; i \in \{1, 2, \dots, b\}, \end{array}$$

where projecting the much smaller Z_i matrices (if $d \ll m$, as is the case with the motivating example described in section) into the PSD cone is much faster than projecting $\mathscr{L}(x) - tI$.