CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.2021.nla.01

Vol. **2**, No. 2, pp. 195-262
June 2021

Linear Constrained Rayleigh Quotient Optimization: Theory and Algorithms

Yunshen Zhou[1], Zhaojun Bai[2,*] and Ren-Cang Li[3]

[1] *Department of Mathematics, University of California, Davis, CA 95616, USA.*
[2] *Department of Computer Science and Department of Mathematics, University of California, Davis, CA 95616, USA.*
[3] *Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, USA.*

Received 16 July 2020; Accepted 31 December 2020

Abstract. We consider the following constrained Rayleigh quotient optimization problem (CRQopt):

$$\min_{v\in\mathbb{R}^n} v^{\mathrm{T}}Av \quad \text{subject to } v^{\mathrm{T}}v=1 \text{ and } C^{\mathrm{T}}v=b,$$

where A is an $n\times n$ real symmetric matrix and C is an $n\times m$ real matrix. Usually, $m\ll n$. The problem is also known as the constrained eigenvalue problem in literature since it becomes an eigenvalue problem if the linear constraint $C^{\mathrm{T}}v=b$ is removed. We start by transforming CRQopt into an equivalent optimization problem (LGopt) of minimizing the Lagrangian multiplier of CRQopt, and then into another equivalent problem (QEPmin) of finding the smallest eigenvalue of a quadratic eigenvalue problem. Although these equivalences have been discussed in literature, it appears to be the first time that they are rigorously justified in this paper. In the second part, we present numerical algorithms for solving LGopt and QEPmin based on Krylov subspace projection. The basic idea is to first project LGopt and QEPmin onto Krylov subspaces to yield problems of the same types but of much smaller sizes, and then solve the reduced problems by direct methods, which is either a secular equation solver (in the case of LGopt) or an eigensolver (in the case of QEPmin). We provide convergence analysis for the proposed algorithms and present error bounds. The sharpness of the error bounds is demonstrated by examples, although in applications the algorithms often converge much faster than the bounds suggest. Finally, we apply the new algorithms to semi-supervised learning in the context of constrained clustering.

AMS subject classifications: 65F15, 65F50, 90C20

Key words: Quadratic programming, quadratic eigenvalue problems, Lanczos algorithms, constrained clustering.

*Corresponding author. *Email addresses:* yshzhou@ucdavis.edu (Y. Zhou), zbai@ucdavis.edu (Z. Bai), rcli@uta.edu (R.-C. Li)

1 Introduction

In this paper, we are concerned with the following *linear constrained Rayleigh quotient* (CRQ) optimization:

$$\text{CRQopt:} \quad \begin{cases} \min\ v^{\mathrm{T}}Av, & \text{(1.1a)} \\ \ \text{s.t.}\ v^{\mathrm{T}}v=1, & \text{(1.1b)} \\ \qquad C^{\mathrm{T}}v=b, & \text{(1.1c)} \end{cases}$$

where $A\in\mathbb{R}^{n\times n}$ is symmetric, $C\in\mathbb{R}^{n\times m}$ has full column rank, and $b\in\mathbb{R}^m$. Necessarily $m<n$ but often $m\ll n$. We are particularly interested in the case where A is large and sparse and $b\neq 0$.

CRQopt (1.1) is also known as *the constrained eigenvalue problem*, a term coined in 1989 [10]. However, it had appeared in literature much earlier than that [15]. In that sense, CRQopt is a classical problem. However, past studies are fragmented with some claims, although often true, not rigorously justified or needed conditions to hold. In this paper, our goal is to provide a thorough investigation into this classical problem, including rigorous justifications of statements previously taken for granted in literature and addressing the theoretical subtleties that were not paid attention to. We also present a quantitative convergence analysis for the Krylov type subspace projection method, which we will also call the Lanczos algorithm, for solving large scale CRQopt (1.1).

1.1 Related works

CRQopt (1.1) has found a wide range of applications, such as ridge regression [5, 12], trust-region subproblem [27, 33], constrained least square problem [9], spectral image segmentation [6, 36], transductive learning [19], and community detection [28].

The first systematic study of CRQopt (1.1) belongs to Gander, Golub and von Matt [10]. Using the full QR and eigen-decompositions, they reformulated CRQopt (1.1) as an optimization problem of finding the minimal Lagrangian multiplier via solving a secular equation (in a way that is different from our secular equation solver in Appendix A). Alternatively, they also turned CRQopt (1.1) into an optimization problem of finding the smallest real eigenvalue of a quadratic eigenvalue problem (QEP). However, the equivalence between the QEP optimization and the Lagrangian multiplier problem was not rigorously justified in [10].

Numerical algorithms proposed in [10] are not suitable for large scale CRQopt (1.1) because they require a full eigen-decomposition of A. Later in [14], Golub, Zhang and Zha considered large and sparse CRQopt (1.1) but only with the homogeneous constraint, i.e, $b=0$. In this special case, CRQopt (1.1) is equivalent to computing the smallest eigenvalue of A restricted to the null space of C^{T}. An inner-outer iterative Lanczos method was proposed to solve the homogeneous CRQopt (1.1). In [41], Xu, Li and Schuurmans proposed a projected power method for solving CRQopt (1.1). The projected power method is an iterative method only involving matrix-vector products, and thus it is suitable for