Trace Minimization Algorithms

Yu Hong Yeung[1], Xin Cheng[1], Di Jin[1]

[1] *Department of Computer Science*

Problem
Given the large sparse symmetric eigenvalue problem $Ax = \lambda Bx$, where A and B are symmetric matrices of order $n = 10^6$, and B is symmetric positive definite. Use the Trace Minimization algorithms: TraceMIN and TraceMIN-Davidson for obtaining the smallest $p \ll n$ eigenpairs, and use TraceMIN for computing **all** the eigenpairs belonging to an interior interval $[a,b]$.

Contents

Introduction

The generalized eigenvalue problem

$$Ax = \lambda Bx, \tag{1}$$

where A and B are $n \times n$ real symmetric matrices with B being positive definite, arises in many applications[1]. Usually A and B are large and sparse and we need only the smallest few or the largest few, or the eigenvalues within an interval. In this project, we implement the trace minimization algorithms that can be used to solve these eigenvalue problems to obtain the smallest p eigenvalues and their corresponding eigenvectors, or all the eigenpairs within a given interval $[a,b]$.

This report is organized as follows. Section 1 presents a brief deivation of the basic trace minimization algortihm and its composite procedures, and the variations of the trace minimization algorithm, namely trace minimization algorithm with deflation and the Davidson-type trace minimization algorithm. Section 2 describes the implementation details of the algorithm: the eigen decomposition methods for dense matrices for solving the eigenvalue problem within a subspace; the iterative solvers used in solving the reduced system, particularly the variation of conjugate gradient method used in the algorithm and its stopping criterion; and also the multisectioning technique to subdivide an interval into subintervals for parallel computations. Section 3 shows our experimental results and discusses about the findings.

1. Algorithms

The trace minimization algorithm is motivated by the following theorem.

Theorem 1 (Sameh and Wisniewski[2]) *Let A and B be as given in problem 1; and let $X*$ be the set of all $n \times p$ matrices X for which $X^TBX = I_p, 1 \le p \le n$. Then*

$$\min_{X \in X*} tr(X^TAX) = \sum_{i=1}^{p} \lambda_i, \tag{2}$$

wher $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ are the eigenvalues of problem 1. The equality holds if and only if the columns of the matrix X, which achieves the minimum, span the eigenspace corresponding to the smallest p eigenvalues.

Hence, if we could find such X that minimizes the trace, we would obtain the smallest p eigenpairs. In the trace minimization algorithm, we start with a random X in the subspace and iteratively refine it by computing a correction term Δ until the trace is minimized.

1.1 Basic Trace Minimization Algorithm

1.2 Trace Minimization Algorithm with Deflation

1.3 The Davidson-type Trace Minimization Algorithm

2. Implementation Details

2.1 Eigen Decomposition for Dense Matrices

2.2 Modified Conjugate Gradient Method

2.3 Multisectioning

3. Results and Discussion

Acknowledgments

References

[1] Alicia M. Klinvex. *Parallel Symmetric Eigenvalue Problem Solvers*. PhD thesis, Purdue University, May 2015.

[2] Ahmed H. Sameh and John A. Wisniewski. A trace minimization algorithm for the generalized eigenvalue problem. *SIAM Journal on Numerical Analysis*, 19(6):1243–1259, 1982.

[3] Ahmed H. Sameh and Zhanye Tong. The trace minimization method for the symmetric generalized eigenvalue problem. *Journal of Computational and Applied Mathematics*, 123(1–2):155–175, 2000. Numerical Analysis 2000. Vol. III: Linear Algebra.

Algorithm 1: Basic Trace Minimization

Input: Subspace dimension $s = 2p$,
$V_1 \in \mathbb{R}^{n\times s}$ with rank s,
$A = A^T; B$ is symmetric positive definite

Output: The smallest p eigenvalues (Θ_k) and their corresponding eigenvectors (Y_k)

for $k = 1 \to$ mat_iter **do**

- B-orthonormalize $V_k \to Z_k$ **:**
  - Compute $\hat{B}_k = BV_k$;
  - Compute all eigenpairs of $V_k^T BV_k$, $V_k^T \hat{B}_k = \Upsilon_k \Sigma_k \Upsilon_k^T$;
  - Compute $Z_k = V_k \Upsilon_k \Sigma_k^{-1/2}$;
  - Compute $\dot{B}_k = \hat{B}_k \Upsilon_k \Sigma_k^{-1/2}$;
- Perform the Rayleigh-Ritz procedure to obtain the Ritz eigenpairs ($AY_k \approx BY_k\Theta_k$) **:**
  - Compute $\hat{A}_k = AV_k \Upsilon_k \Sigma_k^{-1/2}$;
  - Compute all eigenpairs of $Z_k^T AZ_k$, $Z_k^T \hat{A}_k = \Pi_k \Theta_k \Pi_k^T$;
  - Sort the eigenpairs (Θ_k, Π_k) in assending order of Θ_k;
  - Compute $\ddot{B}_k = \dot{B}_k \Pi_k$;
  - Compute $\dot{A}_k = \hat{A}_k \Pi_k$;
- Compute the first p columns of the residual vectors $\Phi_k = \dot{A} - \ddot{B}\Theta_k$;
- Test for convergence **:**
  - $c = 0$;
  - **for** $j = 1 \to p$ **do**
    - **if** $\|\Phi_k e_j\|_2/\theta_j \leq \tau$ **then** $c = c + 1$;
  - **if** $c \geq p$ **then break**;
- Solve the reduced system $P_k A\Delta_k = P_k AY_k$ where $P_k = I - BY_k \left(Y_k^T B^2 Y_k\right)^{-1} Y_k^T B$ **:**
  - Compute the QR decomposition of BY_k, $\ddot{B}_k = [Q_{k1}\ Q_{k2}] \left[R_k^T\ 0\right]^T$;
  - Compute $N_k = \left(I - Q_{k1}Q_{k1}^T\right) \dot{A}_k$;
  - Use conjugate gradient or minimum residual method to solve $\left(I - Q_{k1}Q_{k1}^T\right) A\Delta_k = N_k$ with initial guess $\Delta_k = 0$;
- Compute $V_{k+1} = Y_k - \Delta_k$;