CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2022-0049

Vol. **4**, No. 4, pp. 672-695
December 2023

Multitask Kernel-Learning Parameter Prediction Method for Solving Time-Dependent Linear Systems

Kai Jiang*, Juan Zhang and Qi Zhou

Key Laboratory of Intelligent Computing and Information Processing of Ministry of Education, Hunan key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan, Hunan 411105, P.R. China.

Received 16 December 2022; Accepted 23 March 2023

Abstract. Matrix splitting iteration methods play a vital role in solving large sparse linear systems. Their performance heavily depends on the splitting parameters, however, the approach of selecting optimal splitting parameters has not been well developed. In this paper, we present a multitask kernel-learning parameter prediction method to automatically obtain relatively optimal splitting parameters, which contains simultaneous multiple parameters prediction and a data-driven kernel learning. For solving time-dependent linear systems, including linear differential systems and linear matrix systems, we give a new matrix splitting Kronecker product method, as well as its convergence analysis and preconditioning strategy. Numerical results illustrate our methods can save an enormous amount of time in selecting the relatively optimal splitting parameters compared with the exists methods. Moreover, our iteration method as a preconditioner can effectively accelerate GMRES. As the dimension of systems increases, all the advantages of our approaches become significantly. Especially, for solving the differential Sylvester matrix equation, the speedup ratio can reach tens to hundreds of times when the scale of the system is larger than one hundred thousand.

AMS subject classifications: 62F15, 62J05, 65F08, 65F45, 65M22

Key words: Multitask kernel-learning parameter prediction, time-dependent linear systems, matrix splitting Kronecker product method, convergence analysis, preconditioning.

1 Introduction

In this paper, we consider the time-dependent linear systems (TDLSs) of the form

$$\dot{\boldsymbol{x}}(t) = \mathcal{L} \circ \boldsymbol{x}(t), \quad t \in [0, T], \tag{1.1}$$

*Corresponding author. *Email addresses:* kaijiang@xtu.edu.cn (K. Jiang), qizhou@smail.xtu.edu.cn (Q. Zhou), zhangjuan@xtu.edu.cn (J. Zhang)

where $\boldsymbol{x}(t):[0,T]\to V, V$ is $\mathbb{R}^n (n\in\mathbb{N})$, $\boldsymbol{x}(0)=\boldsymbol{x}_0\in V$ is an initial value, and $\mathcal{L}$ is a linear operator. TDLSs appear in many branches of science and engineering, such as dynamical systems, quantum mechanics, semi-discretization of partial differential equations, differential matrix equations, etc. [2, 13, 18, 20, 25, 29]. A series of discrete methods suitable for them have been developed, such as linear multistep schemes, Runge-Kutta methods, general linear methods, block implicit methods, and boundary value methods (BVMs) [7,8,10,16,17,19,23]. After temporal discretization, each TDLS can be transformed into a sparse linear system

$$Q\boldsymbol{x}=\boldsymbol{b},\quad Q\in\mathbb{R}^{n\times n}\quad \text{is nonsingular and}\quad \boldsymbol{b}\in\mathbb{R}^n.$$

For solving linear systems, matrix splitting iteration methods play an important role as either solvers or preconditioners. The classic matrix splitting forms are all based on $Q=M-N$, where M is a nonsingular matrix such that a linear system with the coefficient matrix M can easily be solved, such as Jacobi, Gauss-Seidel, and successive over-relaxation iteration methods [30]. Alternating direction implicit (ADI) schemes can effectively improve the performance by alternately updating approximate solution. They were initially designed to solve partial differential equations [15, 26], and were gradually extended to more branches, including numerical algebra and optimization [24, 30]. The typical schemes in numerical algebra are Hermitian and skew-Hermitian splitting type methods [3–5, 31]. Further, a general ADI (GADI) framework has recently been developed to put most existing ADI methods into a unified framework [22].

Matrix splitting iteration methods require coefficient matrix Q into different parts with splitting parameters. The convergence and performance of them are very sensitive to splitting parameters, therefore, choosing the optimal splitting parameters is critical. There have been several approaches to select splitting parameters. Experimental traversal method is limited due to an unbearable computational cost, especially for large-scale systems. A more common approach is using theoretical analysis to estimate the bound of splitting parameters in a case-by-case way [3, 11], while its performance heavily depends on the theoretical bound and the scale of systems. Recently, a data-driven approach, Gaussian process regression (GPR) method [22], has presented to predict optimal splitting parameters by choosing an appropriate kernel. GPR method can efficiently predict one splitting parameter at one time.

However, matrix splitting iteration methods can have two or more splitting parameters, among which there are complicated links. Independently predicting each splitting parameter would inevitably affect the prediction accuracy of the original GPR method. Therefore, it requires improving the GRP method to predict multi-parameters simultaneously. Another critical component that determines GPR's availability is kernel function [34, 35]. The original GRP method chooses kernel function by the problem's properties [22], which might produce an artificial error. Moreover, the chosen kernel in a kind of problems may be difficult to extend to others. Therefore, automatically learning kernel functions based on distinct problems is still a challenge.