Dimensionality reduction of stable matrices using random projections (ランダム射影を用いた安定行列の次元削減)

Dept. of Mathematical Informatics 48216215 Hiroki Sakamoto
Supervisor Lecturer Kazuhiro Sato

1 Introduction

Consider a time-invariant linear stable system:

$$G : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are the state, input, and output of the system, respectively, $A \in \mathbb{R}^{n\times n}$ is a stable matrix, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$. The Lyapunov equation corresponding to (1) is as follows:

$$AX + XA^\top + BB^\top = 0. \tag{2}$$

The system (1) is often found in settings with large values of n [1]. However, when n is large, the analysis and design of system (1) requires an enormous amount of computation and memory storage. In addition, the fast solution of the Lyapunov equation (2) is difficult because the computational complexity of the Bartels-Stewart method, a well-known numerical solution method, is $\mathcal{O}(n^3)$.

For large models with n greater than 10^4, model reduction using the Krylov subspace method is effective [1]. The Krylov subspace method is also efficient for solving the Lyapunov equation [4].

However, the Krylov subspace methods may not guarantee stability. In addition, the computational cost of generating the projection is not negligible. Using the Arnoldi method, the most standard Krylov subspace method, the computational complexity for generating the projection is $\mathcal{O}(nd^2)$.

Based on these considerations, we use random projection, which has been actively studied in the fields of statistics and numerical linear algebra in recent years [6, 5], to reduce the dimension of the system (1) and the Lyapunov equation (2). The computational complexity of generating the random matrix is $\mathcal{O}(nd)$. We show that the method based on the random projection method preserves stability with high probability, and we analyse the spectral abscissa between the original and the reduced matrix. Furthermore, we show the error analysis for solving Lyapunov equation.

2 Preliminaries

In this study, to approximate a d-dimensional identity matrix, we assume that each component of the randomised embedding $R \in \mathbb{R}^{n\times d}$ independently follows a Gaussian distribution with mean zero and variance $\frac{1}{n}$, and call $R \in \mathbb{R}^{n\times d}$ a random matrix.

We denote the spectral radius $\max\{|\lambda| \ | \lambda \text{ is an eigenvalue of } A\}$ of the general matrix A by $\rho(A)$, and the spectral abscissa $\max\{\mathrm{Re}(\lambda) \mid \lambda \text{ is an eigenvalue of } A\}$ by $\eta(A)$.

3 Dimensionality reduction for stable matrices

The dimensionality reduction of the matrix A based on the random projection is performed to reduce the computational cost of projection generation and to preserve stability.

Definition 1. *A matrix $A \in \mathbb{R}^{n\times n}$ is called stable, if all eigenvalues of A are in the open left half of the complex plane.*

Definition 2. *A matrix A is called Metzler if all its off-diagonal elements are non-negative.*

Theorem 1. *Let $A \in \mathbb{R}^{n\times n}$ be stable , $R \in \mathbb{R}^{n\times d}$ be a random matrix, and $d \ll n$. Then, for every $0 < \tilde{\epsilon} < \min\left\{-\frac{2\,\mathrm{tr}\,A}{n}, \frac{2(\lambda_{\min}(A+A^\top)-2\,\mathrm{tr}\,A)}{n}\right\}$, $R^\top AR \in \mathbb{R}^{d\times d}$ is stable, with probability at least $1-\delta_1-\delta_2$, where $\delta_1 := \sqrt{\alpha}\cdot 9^d \exp\left(-\alpha+1-\frac{\tilde{\epsilon}}{2}\right)$, $\delta_2 := \sqrt{\alpha}\cdot 9^d \exp\left(-2\alpha+\frac{1}{2}-\frac{\tilde{\epsilon}}{2}\right)$,and $\alpha := \frac{\mathrm{tr}\,A}{\lambda_{\min}(A+A^\top)}$.*

Theorem 1 holds with high probability if the value of α is sufficiently larger than d.

Note that the proof of Theorem 1 is based on [2, Lemma 3.1]. However, the constant that appears in the upper bound of the probability concentration inequality was not explicitly written in the literature, although the value is practically important. In this paper, we clarify the value of upper bound and the constant for [2, Lemma 3.1].

We then evaluate the error between the largest eigenvalue of the original matrix and the reduced matrix. We present the result for the case where A is a stable and Metzler matrix. The stable and Metzler matrices are used in the study of model reduction problems [3].

Corollary 1. *Let $A \in \mathbb{R}^{n\times n}$ be a stable and Metzler matrix, $R \in \mathbb{R}^{n\times d}$ be a random matrix, and $d \ll n$. Then, for every $0 < \epsilon_1 < \min\left\{-\frac{2\,\mathrm{tr}\,A}{n}, \frac{2(\lambda_{\min}(A+A^\top)-2\,\mathrm{tr}\,A)}{n}\right\}$, every $\epsilon_3 > 0$, and every $h > 0$ such that $A + hI_n$ is a non-negative matrix, with probability at least $(1-\delta_1-\delta_2)\left(1-\delta_3\left(d, -\sigma_{\max}(A), -\frac{\mathrm{tr}\,A}{n}\right)-\delta_3\left(d, \sigma_{\min}(A), \frac{\mathrm{tr}\,A}{n}\right)\right)$, we have*

$$|\eta(R^\top AR) - \eta(A)| \leq \frac{\mathrm{tr}\,A}{n} + \epsilon_3 + h - \rho(A + hI_n),$$

where $\delta_3(d,\sigma,\beta) := 9^d \cdot \exp\left(-\frac{\epsilon_3}{2}\right) \cdot \sqrt{\frac{n}{n+2\sigma}} \exp(\beta)$.

Figure 1 compares the solution times for 100 random matrices and 100 Galerkin projections (matrices obtained by the Arnoldi method) for each d. Figure 2 compares the eigenvalue distribution for $d = 100$.

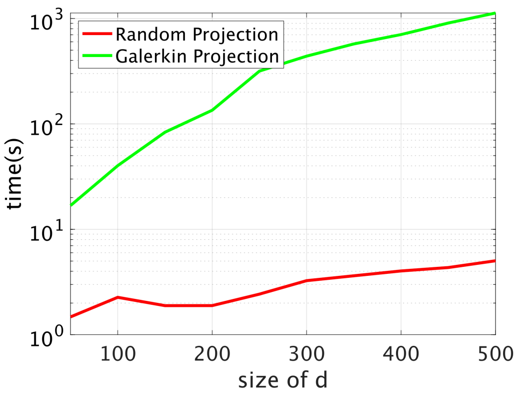

Fig. 1. Computational time to create the projection

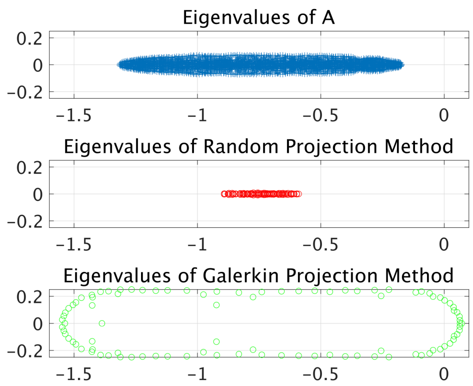

Fig. 2. The eigenvalue distributions of the Metzler matrix A

4 Application to Lyapunov equation

We apply the random projection to the solution of the Lyapunov equation (2). Solving the Lyapunov equation using the usual Arnoldi method has a computational complexity of $\mathcal{O}(nd^2)$, while solving it using a random matrix has a computational complexity of $\mathcal{O}(nd)$.

Let us derive the approximate solution obtained by the random projection. We define the residual r by the approximate solution $\tilde{X}$ to the Lyapunov equation (2) as follows:

$$r := A\tilde{X} + \tilde{X}A^\top + H.$$

We also assume that each column of r is almost orthogonal to a random d-dimensional subspace. From $R^\top R \approx I_d$, we have $R^\top rR \approx 0$. If $\tilde{X} := RY_dR^\top$, we obtain

$$R^\top ARY_d + Y_dR^\top A^\top R + R^\top AR = 0. \qquad (3)$$

Using the solution Y_d of (3), we construct the required solution by $\hat{X} = RY_dR^\top$.

Theorem 2. *Let* $A \in \mathbb{R}^{n\times n}$ *be a stable matrix and* $R \in \mathbb{R}^{n\times d}$ *be a random matrix. Let* $Y_{RP} \in \mathbb{R}^{d\times d}$ *be the solution of* (3) *and* $\lambda_{\min}(Y_{RP}) > -\frac{d}{2}$. *Then, for every* $\epsilon > 0$, *with probability at least* $1 - \delta_3\left(n, -\lambda_{\max}(Y_{RP}), -\frac{\operatorname{tr} Y_{RP}}{n}\right) - \delta_3\left(n, \lambda_{\min}(Y_{RP}), \frac{\operatorname{tr} Y_{RP}}{n}\right)$ *, the following claim holds:*

$$\begin{aligned}&\|AX_{RP} + X_{RP}A^\top + H\|_F^2 \\ &\leq 2\xi^2\|A\|_F^2 + 2\xi^2\|A\|_F\sum_{i=1}^{k}|\sigma_i(A)| \\ &\quad + 4\xi\sum_{i=1}^{n}\|A_i\|\|H^i\| + \|H\|_F^2,\end{aligned}$$

where $\xi = \left(\epsilon + \frac{\operatorname{tr} Y_{RP}}{n}\right)$, $\sigma_i(A)$ *is the* i*-th singular value of matrix* A, $A_i \in \mathbb{R}^n$ *is the* i*-th row vector of* A, *and* $H^i \in \mathbb{R}^n$ *is the* i*-th column vector of* H.

5 Conclusion

We used the random projection to reduce the dimensionality of a stable matrix.

Theoretical and numerical results showed that the reduced matrix with the random projection preserves the stability of the original matrix A with probability. Furthermore, we showed that the probability that the theorems hold changes depending on the properties of A. In addition, we clarified the value of the constant that appears in the theorem, and we showed that the theorem holds with high probability.

A solution method based on the random projection for the Lyapunov equation was proposed and an error analysis was performed. Numerical experiments showed that the proposed method reduces the computational complexity compared to the standard Krylov subspace method.

Bibliography

[1] A. Antoulas, *Approximation of large-scale dynamical systems*, SIAM, 2005.

[2] T. Fuji, P.-L. Poirion, and A. Takeda, *Convexification with bounded gap for randomly projected quadratic optimization*, SIAM Journal on Optimization, 32 (2022), pp. 874–899.

[3] K. Sato, *Reduced model reconstruction method for stable positive network systems*, IEEE Transactions on Automatic Control, (2022).

[4] V. Simoncini, *Computational methods for linear matrix equations*, siam REVIEW, 58 (2016), pp. 377–441.

[5] J. A. Tropp, *Randomized block Krylov methods for approximating extreme eigenvalues*, Numerische Mathematik, 150 (2022), pp. 217–255.

[6] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*, vol. 47, Cambridge university press, 2018.