

MODEL REDUCTION EXPERIMENTS

PART III: BALANCED TRUNCATION REDUCTION

RYAN NONG*

Abstract. This document describes a set of experiments for students in CAAM 651 at Rice University, Spring 2009, that aims to foster insight into the structure and behavior of a particular dynamical system. Students should complete these exercises for one course credit.

0. Instructions. The third section of CAAM 651 concentrated on balanced truncation techniques for reducing linear time invariant dynamical systems of the usual form

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t),\end{aligned}$$

with $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{C} \in \mathbb{R}^{p \times n}$ and $\mathbf{D} \in \mathbb{R}^{p \times m}$. In particular, we use the Hammarling and Smith methods to solve the controllability and observability Lyapunov equations. We then construct stable reduced-order models using balanced truncation. The exercises described here invite you to investigate how these algorithms perform on a real dynamical system.

This list of tasks is intentionally open-ended. The investigations you conduct will be guided by your own instinct and insight, along with any restrictions imposed by the size and properties of the matrix you select. Be frank if you are observing a phenomenon you cannot understand; feel free to include additional experiments that might give you greater insight. If the size of the matrix you have chosen prohibits you from completing some aspects of the assignment, please adapt the problem to something more reasonable for your system. We are happy to discuss strategies with you to help facilitate large-scale computations.

0.1. A note about your write-up. Please submit a typed report describing the experiments you conduct, as directed by the tasks below. We recommend that you use the L^AT_EX typesetting system to collect and organize your results. In particular, you can download the L^AT_EX template, `modred_tasks3.tex`, from the class website, which you can then edit; this will give our reports a uniform appearance. (This document uses the SIAM journal macros. These macros and supporting files, along with all the graphics files used in this report, are collected in one `.zip` file on the class website).

1. Hammarling Method. Select a system from your research, or from those listed on the course website. Recall that in balanced truncation reduction, we would like to obtain the two infinite system gramians \mathcal{P} and \mathcal{Q} by solving the two associated controllability and observability Lyapunov equations:

$$\begin{aligned}\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^T &= -\mathbf{B}\mathbf{B}^T \\ \mathbf{A}^T\mathcal{Q} + \mathcal{Q}\mathbf{A} &= -\mathbf{C}^T\mathbf{C}.\end{aligned}$$

Consider the controllability Lyapunov equation (the same analysis applies for the observability Lyapunov equation). Since \mathbf{A} is stable and $\mathbf{B}\mathbf{B}^T$ is non-negative

*Department of Computational and Applied Mathematics, Rice University, 6100 Main Street—MS 134, Houston, Texas 77005–1892 (ryannong@caam.rice.edu).

definite, we have that \mathcal{P} is also non-negative definite, and hence possesses a Cholesky factorization

$$\mathcal{P} = \mathbf{U}\mathbf{U}^T,$$

where $\mathbf{U} \in \mathbb{R}^{n \times n}$ is upper triangular with real non-negative diagonal entries. The motivation for considering the Cholesky factors of \mathcal{P} in place of itself is that \mathcal{P} may be considerably more ill-conditioned with respect to inversion than \mathbf{U} . The reason is as follows: Let

$$\text{cond}(\mathcal{P}) = \|\mathcal{P}\|_2 \|\mathcal{P}^{-1}\|_2$$

denote the condition number of \mathcal{P} , which describes the accuracy with which linear systems involving \mathcal{P} can be solved. Then

$$\text{cond}(\mathcal{P}) = \text{cond}(\mathbf{U})^2.$$

Thus, if we could solve the Lyapunov equation directly for \mathbf{U} , then by using \mathbf{U} in place of \mathcal{P} , we may avoid the loss of accuracy associated with the squaring of the condition number as shown above. The Hammarling method aims to fulfill this desire. Similar to the Bartels–Stewart method that was described in class, consider a Schur decomposition of the matrix \mathbf{A} given by

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{Q}^T,$$

where $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is orthogonal and $\mathbf{R} \in \mathbb{R}^{n \times n}$ upper triangular. Then the Lyapunov equation becomes

$$\mathbf{R}\tilde{\mathbf{U}}\tilde{\mathbf{U}}^T + \tilde{\mathbf{U}}\tilde{\mathbf{U}}^T\mathbf{R}^T = -\tilde{\mathbf{B}}\tilde{\mathbf{B}}^T,$$

where $\tilde{\mathbf{U}} = \mathbf{Q}^T\mathbf{U}$ and $\tilde{\mathbf{B}} = \mathbf{Q}^T\mathbf{B}$. We now solve this reduced equation for $\tilde{\mathbf{U}}$. Once $\tilde{\mathbf{U}}$ is obtained, \mathbf{U} is constructed as

$$\mathbf{U} = \mathbf{Q}^T\tilde{\mathbf{U}}.$$

Use the routine `lyapUR.m` to compute the Cholesky factorizations of \mathcal{P} and \mathcal{Q}

$$\mathcal{P} = \mathbf{U}\mathbf{U}^T, \quad \mathcal{Q} = \mathbf{L}\mathbf{L}^T,$$

and study various properties of the infinite gramians and resulting reduced order models.

- 1.1. Produce `semilog` plots of the singular values of \mathcal{P} and \mathcal{Q} .
- 1.2. Compute the singular value decomposition of $\mathbf{U}^T\mathbf{L}$:

$$\mathbf{U}^T\mathbf{L} = \mathbf{P}\mathbf{S}\mathbf{Q}^T.$$

- 1.3. For various values of k , conformally partition the singular value decomposition into

$$\mathbf{P}\mathbf{S}\mathbf{Q}^T = [\mathbf{P}_k \ \hat{\mathbf{P}}_k] \begin{bmatrix} \mathbf{S}_k & \\ & \hat{\mathbf{S}}_k \end{bmatrix} \begin{bmatrix} \mathbf{Q}_k^T \\ \hat{\mathbf{Q}}_k^T \end{bmatrix},$$

where $\mathbf{S}_k \in \mathbb{R}^{k \times k}$, and construct reduced order models

$$\mathbf{A}_k = \mathbf{W}_k^T \mathbf{A} \mathbf{V}_k, \quad \mathbf{B}_k = \mathbf{W}_k^T \mathbf{B}, \quad \mathbf{C}_k = \mathbf{C} \mathbf{V}_k,$$

where $\mathbf{W}_k = \mathbf{L}\mathbf{Q}_k\mathbf{S}_k^{-1/2}$ and $\mathbf{V}_k = \mathbf{U}\mathbf{P}_k\mathbf{S}_k^{-1/2}$.

- 1.4. Are your reduced systems stable? Produce a plot showing the spectral abscissa of \mathbf{A}_k as a function of k .
- 1.5. For several values of k , use the `Mysigma_log.m` routine to produce a plot that compares the transfer function for the original system to the reduced system. (It is probably best to make one plot for each value of k , to reduce clutter.)

The Bartels–Stewart and Hammarling methods are direct methods, based on the QR factorization. These methods ignore any sparsity or low-rank property in the Lyapunov equation and have computational complexity $\mathcal{O}(n^3)$ and storage complexity $\mathcal{O}(n^2)$. Although these methods are considered standard for small and dense Lyapunov equations, their use is limited for large and sparse systems. Next, we consider one of the techniques that targets large and sparse Lyapunov equations.

2. Smith Method. Again, let us investigate the controllability Lyapunov equation:

$$\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^T = -\mathbf{B}\mathbf{B}^T,$$

where \mathbf{A} is assumed stable. Consider a negative real number $\mu \in \mathbb{R}^-$. It was shown in class that this Lyapunov equation can be transformed into a Stein equation of the form

$$\mathcal{P} = \mathbf{A}_\mu \mathcal{P} \mathbf{A}_\mu^T + \mathbf{B}_\mu \mathbf{B}_\mu^T,$$

where $\mathbf{A}_\mu = (\mathbf{A} + \mu \mathbf{I})^{-1}(\mathbf{A} - \mu \mathbf{I})$ and $\mathbf{B}_\mu = \sqrt{2|\mu|}(\mathbf{A} + \mu \mathbf{I})^{-1}\mathbf{B}$. With \mathbf{A} stable and μ negative, the solution of this Stein equation can be shown to be the infinite matrix series

$$\mathcal{P} = \sum_{j=0}^{\infty} \mathbf{A}_\mu^j \mathbf{B}_\mu \mathbf{B}_\mu^T (\mathbf{A}_\mu^j)^T,$$

and can be expressed in factored form as

$$\mathcal{P} = \mathbf{U}\mathbf{U}^T,$$

where $\mathbf{U} = [\mathbf{B}_\mu, \mathbf{A}_\mu \mathbf{B}_\mu, \mathbf{A}_\mu^2 \mathbf{B}_\mu, \dots]$. This infinite series can be approximated by a partial sum

$$\mathcal{P}_m = \sum_{j=0}^m \mathbf{A}_\mu^j \mathbf{B}_\mu \mathbf{B}_\mu^T (\mathbf{A}_\mu^j)^T$$

because it can be shown that the error $\mathcal{E}_m = \mathcal{P} - \mathcal{P}_m \rightarrow \mathbf{0}$ as $m \rightarrow \infty$. An asymptotic convergence rate of the method is obtained from the spectral radius of \mathbf{A}_μ , $\rho(\mathbf{A}_\mu)$, which is strictly less than 1. One would like to select μ to minimize $\rho(\mathbf{A}_\mu)$. This is an active area of research and there are techniques for computing optimal values for μ . For completion of this part of the assignment, pick a few different values for μ and perform experiments to observe their effect on convergence. The routine `smith.m` on the class website can be used to approximately compute the low-rank factors of \mathcal{P} and \mathcal{Q} as in

$$\mathcal{P}_m = \mathbf{U}_m \mathbf{U}_m^T, \quad \mathcal{Q}_m = \mathbf{L}_m \mathbf{L}_m^T.$$

- 2.1. With a prescribed threshold of tolerance (e.g., 10^{-6}), select several values for μ and use `smith.m` to approximately compute the low-rank factors of \mathcal{P} and \mathcal{Q} for each μ . Report the number of steps required for convergence for each μ . In addition, produce `semilog` error plots showing $\|\mathcal{P} - \mathcal{P}_m\|$, where \mathcal{P} is constructed using the dense methods described earlier or the Smith method with a very large m .
- 2.2. Pick the μ that results in the fewest steps for convergence, compute \mathcal{P}_m and \mathcal{Q}_m and plot their singular values.
- 2.3 Compute the singular value decomposition of $\mathbf{U}_m^T \mathbf{L}_m$:

$$\mathbf{U}_m^T \mathbf{L}_m = \mathbf{P} \mathbf{S} \mathbf{Q}^T.$$

Plot the decay of the singular values of $\mathbf{U}_m^T \mathbf{L}_m$, the Hankel singular values.

- 2.4. For various values of k , conformally partition the singular value decomposition into

$$\mathbf{P} \mathbf{S} \mathbf{Q}^T = [\mathbf{P}_k \ \widehat{\mathbf{P}}_k] \begin{bmatrix} \mathbf{S}_k & \\ & \widehat{\mathbf{S}}_k \end{bmatrix} \begin{bmatrix} \mathbf{Q}_k^T \\ \widehat{\mathbf{Q}}_k^T \end{bmatrix},$$

where $\mathbf{S}_k \in \mathbb{R}^{k \times k}$, and construct reduced order models

$$\mathbf{A}_k = \mathbf{W}_k^T \mathbf{A} \mathbf{V}_k, \quad \mathbf{B}_k = \mathbf{W}_k^T \mathbf{B}, \quad \mathbf{C}_k = \mathbf{C} \mathbf{V}_k,$$

where $\mathbf{W}_k = \mathbf{L} \mathbf{Q}_k \mathbf{S}_k^{-1/2}$ and $\mathbf{V}_k = \mathbf{U} \mathbf{P}_k \mathbf{S}_k^{-1/2}$.

- 2.5. Are your reduced systems stable? Produce a plot showing the spectral abscissa of \mathbf{A}_k as a function of k .
- 2.6. For several values of k , use the `Mysigma_log.m` routine to produce a plot that compares the transfer function for the original system to the reduced system. (It is probably best to make one plot for each value of k , to reduce clutter.)
Optional: Fix k and see how the accuracy of \mathcal{P}_m affects the balanced system.