

Applications of ADI

Jane Du

November 2019

In this memo we detail a few example problems that are well-suited to the ADI method. Accompanying this document is a notebook that generates our example problems.

1 Dynamical Systems

In control theory, a dynamical system is a system in which a function describes the time-dependence of a point in a geometrical space.

Given a state-vector x , an input vector u , and an output vector y , consider the control system given by

$$\dot{x}(t) = Mx(t) + Bu(t), \quad x(0) = x_0, \quad y(t) = Cx(t), \quad (1)$$

where $M \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times k}$, and $C \in \mathbb{R}^{m \times n}$ are time-invariant matrices. In many situations of practical interest neither the initial state x_0 nor the states $x(t)$ are known explicitly. A popular idea is to design control systems related to (1), whose states $\tilde{x}(t)$ approximate $x(t)$.

Luenberger [1] proposed the construction of an approximation $\tilde{x}(t)$ of $x(t)$ as follows:

$$\dot{\tilde{x}}(t) = H^T \tilde{x}(t) + G^T y(t) + Du(t), \quad \tilde{x}(0) = \tilde{x}_0.$$

where $H, G \in \mathbb{R}^{m \times m}$, and $D \in \mathbb{R}^{m \times k}$ are matrices to be determined.

Assuming that the spectra of H and M are disjoint, then the Sylvester equation

$$M^T X - XH = C^T G$$

has a unique solution $X \in \mathbb{R}^{n \times m}$.

We will now show that the difference between $\tilde{x}(t)$ and $X^T x(t)$ converges to zero as $t \rightarrow \infty$, provided that the matrices H and D are suitably selected. Differentiating the difference

$$e(t) = \tilde{x}(t) - X^T x(t)$$

gives

$$\begin{aligned} \dot{e}(t) &= H^T \tilde{x}(t) + G^T y(t) + Du(t) - X^T (Mx(t) + Bu(t)) \\ &= H^T e(t) + (H^T X^T + G^T C - X^T M)x(t) + (D - X^T B)u(t). \end{aligned}$$

Letting $D = X^T B$ shows us that

$$e(t) = e^{Ht}(\tilde{x}_0 - X^T x_0).$$

Assuming that the eigenvalues of H have negative real part, it follows that $e(t) \rightarrow 0$ as $t \rightarrow \infty$.

There are other conditions that are needed so that $\tilde{x}(t)$ is not sensitive to perturbations.

to estimate $x(t)$ with $z(t)$, that the error $e(t) = z(t) - Xx(t)$ approaches 0 with time for any $x(0), z(0), u(t)$. It can be shown that $\dot{e}(t) = e^{Tt}e(0) \rightarrow 0$ if the matrices used to construct z satisfy the following requirements:

- T is a stable matrix (has negative real part),
- $TX - XA = -GC$,
- $H = XB$.

The algorithm to find $z(t)$ involves first choosing T such that it is stable and its spectrum is disjoint from A . Next, G is chosen such that (T, G) is controllable. After this, the main step is solving the following equation for X :

$$XM - TX = GC.$$

If X is nonsingular, H is calculated by $H = XB$, and $z(t)$ can be constructed. The matrix A can easily be dense and very large. Again, T is required to be stable. Since this process is used to calculate an approximation of a vector.

Other current methods include using direct methods for smaller matrices based on the Schur decomposition, the SVD-based algorithm, and a generalization of the Hessenberg algorithm [5].

The last method heavily employs BLAS-3 routines and is well-suited for high-performance computing. Additionally, several algorithms have been developed to be large-scale and parallelizable [5].

2 PDE Discretization

An example of the discretization of PDEs: for example, looking at the 2D Poisson equation on the function f :

$$u' = \alpha \nabla^2 u, u(x, y) = 0 \text{ for } |x| = 1, |y| = 1$$

the discretized version is the Lyapunov equation below:

$$D_2 X + X D_2^T = F.$$

Using finite differences to form this problem, we get tridiagonal D_2 . With $D_2 \in \mathbb{R}^{n \times n}$, Bartels-Stewart would take $O(n^3)$ to solve directly, and using

an iterative method like conjugate gradients would take $O(n^3)$ to solve the Kronecker formulation of the problem:

$$(D_2 \otimes I + I \otimes D_2)\text{vec}(U) = \text{vec}(F).$$

Normally, factored ADI (fADI) would be a good choice since D_2 is symmetric positive definite and triangular. However, applying a randomized linear solve during ADI would destroy this structure, so a more interesting application would be when D_2 is dense, such as with spectral methods.

Time-dependent systems described by PDEs can be approximated with spatial discretizations. One of the examples provided in [3] is the simulation of VLSI circuits, where the models are high-order, with $n \approx 10^5$. Approximation methods that have been employed to solve these problems are typically SVD-based, Krylov-based, or iterative methods combining both.

3 Control Theory

3.1 Controllability of a system

From [4], the Sylvester-controller matrix equation (and its dual, the Sylvester-observer matrix equation) discussed in the first section is an important problem in control theory. Another problem alluded to in the algorithm for state vector estimation is the controllability of a pair of vectors (T, G) .

Using the same discrete-time linear system introduced in dynamical systems:

$$\dot{x} = Ax + Bu, \quad y = Cx$$

Again, x, u, y are state, input, and output vectors and A, B, C are time-invariant.

If A is stable, then the controllability Gramian P is the solution to the Sylvester equation

$$AP + PA^* + BB^* = 0$$

If this matrix is positive definite, then the system is controllable.

3.2 Eigenvalue Assignment

Another matrix equation problem that arises in the design of control systems is the matrix eigenvalue assignment.

The eigenvalue assignment problem can be used to solve the feedback-stabilization function. The goal is to find a feedback matrix F such that the closed-loop matrix $A + BF$ has a preassigned spectrum.

Given the controllable pair (A, B) and the set $\Omega = \{\lambda_1, \dots, \lambda_n\}$, closed under complex conjugation, find a feedback matrix F such that the spectrum of the closed-loop matrix $A + BF$ is the set Ω .

The observability Gramian can be used to construct algorithms for the above problem.

Given a Sylvester-observer equation

$$AX - XT = -BG$$

If T 's spectrum is chosen to be Ω and there is a nonsingular solution X , then choosing the feedback matrix to be $F = GX^{-1}$ ensures that the spectrum of $A + BF$ is the set Ω . It is not guaranteed that there exists a nonsingular solution X , but there are algorithms that implicitly generate nonsingular solutions for the form $XA - BX = R$, where A is arbitrary, B is nonderogatory, and R is known up to its first $n - 1$ rows [4].

Some other problems in control theory include stability analysis, stabilization, optimal control, solution of algebraic Riccati equations, and balancing.

4 Denoising Images

Applying the ADI iteration to restoring noisy images was discussed in [1], and was examined by Cheong and Mongera in detail in [2].

In the problem of restoring noisy images, we represent the image as $\mathbf{f} = \text{vec}(F)$ where F is the pixel map of the image. If the noisy image is $\mathbf{g} = \mathbf{f} + \eta$, where η is the Gaussian noise vector, the minimum mean square error estimate $\hat{\mathbf{f}}$ can be computed by applying the Wiener filter:

$$(I + \Phi_\eta \Phi_f^{-1})\hat{\mathbf{f}} = \mathbf{g}$$

Given white and Gaussian noise, we can represent the covariance matrix of the noise as $\Phi_\eta = \sigma_\eta^2 I$. Additionally, it is assumed that Φ_f is separable:

$$(I + \sigma_\eta^2 \Phi_y^{-1} \otimes \Phi_x^{-1})\hat{\mathbf{f}} = \mathbf{g}$$

This is the Kronecker formation of the Sylvester equation. Re-writing as Sylvester's equation:

$$\hat{F}\Phi_x + \sigma_\eta^2 \Phi_y^{-1} \hat{F} = G\Phi_x$$

Here, the constant matrices in the equation are positive definite and dense. ADI can be applied to estimate the original image F .

Using G as a first guess, it was observed that ADI typically converged faster than conjugate gradients. The authors also discuss Generalized ADI (GADI), in which the linear systems in ADI are not applied in strict alternation. In their experiments, it was shown that applying one "pass" more frequently than the other helps the algorithm to converge faster. The shift parameters used in [2] were chosen using Bagby points, which was generalized to work with GADI.

References

- [1] V. SIMONCINI, *Computational methods for linear matrix equations*, SIAM Rev., 58 (2016), pp. 377–441, <https://doi.org/10.1137/130912839>.
- [2] D. CALVETTI AND L. REICHEL, *Application of ADI iterative methods to the restoration of noisy images*, SIAM Journal on Matrix Analysis and Applications, 17:1 (1996), pp. 165–186, <https://doi.org/10.1137/S0895479894273687>.
- [3] A. ANTOULAS AND D. SORENSEN, *Approximation of large-scale systems: an overview*, Int. J. Appl. Math. Comput. Sci., 11 (2001), pp. 1093–1121, <http://matwbn.icm.edu.pl/ksiazki/amc/amc11/amc1155.pdf>.
- [4] B. N. DATTA, *Linear and numerical linear algebra in control theory: some research problems*, Linear Algebra and its Applications, 197–198 (1994), pp. 755–790, [https://doi.org/10.1016/0024-3795\(94\)90512-6](https://doi.org/10.1016/0024-3795(94)90512-6).
- [5] J. CARVALHO, K. DATTA, AND Y. HONG, *A new block algorithm for full-rank solution of the Sylvester-observer equation*, IEEE Transactions on Automatic Control, 48, (2003), pp. 2223–2228, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1254095>.