Notes on the convergence of the ADI method

Nidham Gazagnadou*

January 9, 2020

Based on Heather Wilber's notes on Zolotarev numbers, this note aims at redoing proofs and detailing computations of the convergence of the ADI method when solving the Sylvester equation

$$\mathbf{AX} - \mathbf{XB} = \mathbf{C} \ . \tag{1}$$

1 Convergence of the ADI method

1.1 ADI error of the iterates

The ADI iterations are given by

$$(\mathbf{A} - p\mathbf{I})\mathbf{X}_{j+1/2} = \mathbf{X}_j(\mathbf{B} - p\mathbf{I}) + \mathbf{C} \tag{2}$$
$$\mathbf{X}_{j+1}(\mathbf{B} - q\mathbf{I}) = (\mathbf{A} - q\mathbf{I})\mathbf{X}_{j+1/2} - \mathbf{C} \ . \tag{3}$$

Assuming that there is an unique solution $\mathbf{X}^*$ to the Sylvester equation (1) (for instance if the spectrum of $\mathbf{A}$ and $\mathbf{B}$ are separated enough), this solution satisfies the following equation

$$\begin{aligned}
&(\mathbf{A} - p\mathbf{I})\mathbf{X}^*(\mathbf{B} - q\mathbf{I}) - (\mathbf{A} - q\mathbf{I})\mathbf{X}^*(\mathbf{B} - p\mathbf{I}) = (p-q)\mathbf{C}\\
\Rightarrow &(p-q)(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{C}(\mathbf{B} - q\mathbf{I})^{-1} - \mathbf{X}^* = -(\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{A} - q\mathbf{I})\mathbf{X}^*(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}
\end{aligned} \tag{4}$$

which directly results from developing the left-hand side and using the fact that $\mathbf{X}^*$ satisfies (1).

Let us observe that

$$(\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1} - \mathbf{I} = \left[(\mathbf{A} - q\mathbf{I}) - (\mathbf{A} - p\mathbf{I})\right](\mathbf{A} - p\mathbf{I})^{-1} = (p-q)(\mathbf{A} - p\mathbf{I})^{-1} \tag{5}$$

Now, let us combine (2) and (3) to derive the ADI update in one step

$$\begin{aligned}
&\mathbf{X}_{j+1}(\mathbf{B} - q\mathbf{I}) \overset{(2)+(3)}{=} (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{X}_j(\mathbf{B} - p\mathbf{I}) + (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{C} - \mathbf{C}\\
\iff &\mathbf{X}_{j+1} = (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{X}_j(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} + \left[(\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1} - \mathbf{I}\right]\mathbf{C}(\mathbf{B} - q\mathbf{I})^{-1}\\
\overset{(5)}{\iff} &\mathbf{X}_{j+1} = (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{X}_j(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} + (p-q)(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{C}(\mathbf{B} - q\mathbf{I})^{-1}
\end{aligned} \tag{6}$$

So, we get the ADI error

$$\begin{aligned}
&\mathbf{X}_{j+1} - \mathbf{X}^* \overset{(6)}{=} (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{X}_j(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} + (p-q)(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{C}(\mathbf{B} - q\mathbf{I})^{-1} - \mathbf{X}^*\\
\Rightarrow &\mathbf{X}_{j+1} - \mathbf{X}^* \overset{(4)}{=} (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{X}_j(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} - (\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{A} - q\mathbf{I})\mathbf{X}^*(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}
\end{aligned}$$

Finally, the ADI error is the following

$$\Rightarrow \boxed{\mathbf{X}_{j+1} - \mathbf{X}^* = (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{X}_j - \mathbf{X}^*)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}} \tag{7}$$

*Télécom Paris, France.

1.2 ADI error of the residual

Let us introduce what we call *the residual* of the Sylvester equation (1)

$$R(\mathbf{X}) := \mathbf{C} - (\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B}) \tag{8}$$

We can first observe that

$$R(\mathbf{X}) = \mathbf{C} - (\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B}) = -\mathbf{A}(\mathbf{X} - \mathbf{X}^*) + (\mathbf{X} - \mathbf{X}^*)\mathbf{B}$$

So, the recurrence relation follows

$$\begin{aligned}
&R(\mathbf{X}_{j+1}) \\
&= -\mathbf{A}(\mathbf{X}_{j+1} - \mathbf{X}^*) + (\mathbf{X}_{j+1} - \mathbf{X}^*)\mathbf{B} \\
&\overset{(7)}{=} -\mathbf{A}(\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{X}_j - \mathbf{X}^*)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} + (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{X}_j - \mathbf{X}^*)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}\mathbf{B} \\
&= -(\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{A}(\mathbf{X}_j - \mathbf{X}^*)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} + (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}(\mathbf{X}_j - \mathbf{X}^*)\mathbf{B}(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} \\
&= (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}\left[-\mathbf{A}(\mathbf{X}_j - \mathbf{X}^*) + (\mathbf{X}_j - \mathbf{X}^*)\mathbf{B}\right](\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1} \\
&= (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}R(\mathbf{X}_j)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}
\end{aligned}$$

where in the second equality we use that $\mathbf{A}$, $(\mathbf{A} - q\mathbf{I})$ and $(\mathbf{A} - p\mathbf{I})^{-1}$ commute. The same holds for $\mathbf{B}$, $(\mathbf{B} - p\mathbf{I})$ and $(\mathbf{B} - q\mathbf{I})^{-1}$.

Finally, the ADI error of the residual (equivalent of suboptimality in optimization) is the following

$$\Rightarrow \boxed{R(\mathbf{X}_{j+1}) = (\mathbf{A} - q\mathbf{I})(\mathbf{A} - p\mathbf{I})^{-1}R(\mathbf{X}_j)(\mathbf{B} - p\mathbf{I})(\mathbf{B} - q\mathbf{I})^{-1}} \tag{9}$$

1.3 Conclusion

The error of the ADI iterates and the residual of the ADI iterates have the same behaviour.