report slyvester equation

sicheng4

February 2024

Main

This section contains main theorems and proofs, as well as some errors in original paper.

Proposition 2.1

Let $\boldsymbol{S_U U}_{d+1} = \boldsymbol{Q}_{\boldsymbol{U},d+1} \boldsymbol{T}_{\boldsymbol{U},d+1}$ be a reduced QR decomposition with

$$\boldsymbol{Q}_{\boldsymbol{U},d+1} = \left[\boldsymbol{Q}_{\boldsymbol{U},d}, Q_{\boldsymbol{U},d+1}\right] \text{ and } \boldsymbol{T}_{\boldsymbol{U},d+1} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d} & T_{H,d+1} \\ 0^\top & \boldsymbol{\tau}_{d+1} \end{bmatrix}$$

Then, for the sketched method, the following Arnoldi-like formula holds:

$$\boldsymbol{S_U A U}_d = \boldsymbol{S_U U}_d \left(\boldsymbol{H}_d + R_H E_d^\top\right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top$$

with $R_H = \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d}$ and $Q_{\boldsymbol{U},d+1} \perp \boldsymbol{S_U U}_d$. Similarly, if

$$\boldsymbol{S_V V}_{d+1} = \boldsymbol{Q}_{V,d+1} \boldsymbol{T}_{V,d+1}, \quad \boldsymbol{T}_{V,d+1} = \begin{bmatrix} \boldsymbol{T}_{V,d} & T_{G,d+1} \\ 0^\top & \boldsymbol{\theta}_{d+1} \end{bmatrix}$$

then

$$\boldsymbol{S_V B}^\top \boldsymbol{V}_d = \boldsymbol{S_V V}_d \left(\boldsymbol{G}_d + R_G E_d^\top\right) + Q_{V,d+1} \boldsymbol{\theta}_{d+1} \boldsymbol{g}_{d+1,d} E_d^\top,$$

where $R_G := \boldsymbol{T}_{V,d}^{-1} T_G \boldsymbol{g}_{d+1,d}$.

Proof

$$\begin{aligned}
\boldsymbol{S_U A U}_d &= \boldsymbol{S_U U}_{d+1} \underline{\boldsymbol{H}}_d = \boldsymbol{S_U U}_{d+1} \underline{\boldsymbol{H}}_d = \boldsymbol{S_U U}_{d+1} \left[\boldsymbol{H}_d; H^\top\right] \\
&= \boldsymbol{Q}_{\boldsymbol{U},d+1} \boldsymbol{T}_{\boldsymbol{U},d+1} \left[\boldsymbol{H}_d; H^\top\right] \\
&= \left[\boldsymbol{Q}_{\boldsymbol{U},d}, Q_{\boldsymbol{U},d+1}\right] \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d} & T_{H,d+1} \\ 0^\top & \boldsymbol{\tau}_{d+1} \end{bmatrix} \left[\boldsymbol{H}_d; H^\top\right] \\
&= \begin{bmatrix} \boldsymbol{Q}_{\boldsymbol{U},d} \boldsymbol{T}_{\boldsymbol{U},d} & \boldsymbol{Q}_{\boldsymbol{U},d} T_{H,d+1} + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \end{bmatrix} \left[\boldsymbol{H}_d; H^\top\right] \\
&= \boldsymbol{Q}_{\boldsymbol{U},d} \boldsymbol{T}_{\boldsymbol{U},d} \boldsymbol{H}_d + \boldsymbol{Q}_{\boldsymbol{U},d} T_{H,d+1} H^T + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} H^T \\
&= \boldsymbol{Q}_{\boldsymbol{U},d} \boldsymbol{T}_{\boldsymbol{U},d} (\boldsymbol{H}_d + \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_{H,d+1} H^T) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top \\
&= \boldsymbol{S_U U}_d \left(\boldsymbol{H}_d + \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d} E_d^\top\right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top \\
&= \boldsymbol{S_U U}_d \left(\boldsymbol{H}_d + R_H E_d^\top\right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top
\end{aligned}$$

where letting $R_H = \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d}$ and $Q_{\boldsymbol{U},d+1} \perp \boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{U}_d$ The same process holds for $\boldsymbol{S}_{\boldsymbol{V}}$

The second part is the verification of whitened-sketched Arnoldi relations. The bases are changed to

$$\widehat{\boldsymbol{U}}_d := \boldsymbol{U}_d \boldsymbol{T}_{\boldsymbol{U},d}^{-1}, \quad \widehat{\boldsymbol{V}}_d := \boldsymbol{V}_d \boldsymbol{T}_{V,d}^{-1}$$

. Again, we will only show the $\widehat{\boldsymbol{U}}_d$, since the other one could be derived by same process. By last proposition, we have

$$\boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{A} \boldsymbol{U}_d = \boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{U}_d \left(\boldsymbol{H}_d + \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d} E_d^\top \right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top$$

Times $\boldsymbol{T}_{U,d}^{-1}$ on both side, and just do simple calculations. Note that,

$$\boldsymbol{T}_{\boldsymbol{U},d}^{-1} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d-1}^{-1} & -\boldsymbol{T}_{\boldsymbol{U},d-1}^{-1} T_{H,d-1} \boldsymbol{\tau}_d^{-1} \\ & -\boldsymbol{\tau}_d^{-1} \end{bmatrix}$$

This gives us:

$$\begin{aligned}
\boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{A} \boldsymbol{U}_d \boldsymbol{T}_{U,d}^{-1} &= \boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{U}_d \left(\boldsymbol{H}_d + \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d} E_d^\top \right) \boldsymbol{T}_{U,d}^{-1} + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top \boldsymbol{T}_{U,d}^{-1} \\
&= \boldsymbol{S}_{\boldsymbol{U}} \boldsymbol{U}_d \boldsymbol{T}_{U,d}^{-1} \boldsymbol{T}_{U,d} \left(\boldsymbol{H}_d + \boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_H \boldsymbol{h}_{d+1,d} E_d^\top \right) \boldsymbol{T}_{U,d}^{-1} + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top \boldsymbol{T}_{U,d}^{-1} \\
&= \boldsymbol{S}_{\boldsymbol{U}} \widehat{\boldsymbol{U}}_d \left(\widehat{\boldsymbol{H}}_d + \widehat{H} E_d^\top \right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} E_d^\top \boldsymbol{T}_{U,d}^{-1} \\
&= \boldsymbol{S}_{\boldsymbol{U}} \widehat{\boldsymbol{U}}_d \left(\widehat{\boldsymbol{H}}_d + \widehat{H} E_d^\top \right) + Q_{\boldsymbol{U},d+1} \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1} \boldsymbol{E}_d^\top
\end{aligned}$$

where

$$\widehat{\boldsymbol{H}}_d = \boldsymbol{T}_{\boldsymbol{U},d} \boldsymbol{H}_d \boldsymbol{T}_{\boldsymbol{U},d}^{-1}, \widehat{H} = T_{H,d+1} \boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1}$$

Question1 ??? Note here, the last term is different with the item in paper. I could not figure out why the last term should be $Q_{\boldsymbol{U},d+1} \boldsymbol{h}_{d+1,d} E_d^\top$

Typos on AL1

1. Updae $\widehat{\boldsymbol{H}}_{d+1}$ At d-th iteration, we could get $\widehat{\boldsymbol{H}}_{d+1}$ which would be used in next iteration as$\widehat{\boldsymbol{H}}_d$. Note, $\boldsymbol{T}_{\boldsymbol{U},d+1}^{-1} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d}^{-1} & -\boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_{H,d+1} \boldsymbol{\tau}_{d+1}^{-1} \\ & -\boldsymbol{\tau}_{d+1}^{-1} \end{bmatrix}$
   By $\widehat{\boldsymbol{H}}_{d+1} = \boldsymbol{T}_{\boldsymbol{U},d+1} \boldsymbol{H}_{d+1} \boldsymbol{T}_{\boldsymbol{U},d+1}^{-1}$, we could imply

$$\begin{aligned}
\widehat{\boldsymbol{H}}_{d+1} &= \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d} & T_{H,d+1} \\ & \boldsymbol{\tau}_{d+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{H}_{d+1} & H \\ 0 \ldots 0, \boldsymbol{h}_{d+1,d} & \boldsymbol{h}_{d+1,d+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d}^{-1} & -\boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_{H,d+1} \boldsymbol{\tau}_{d+1}^{-1} \\ & -\boldsymbol{\tau}_{d+1}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d} \boldsymbol{H}_{d+1} + T_{H,d+1} \boldsymbol{h}_{d+1,d} \boldsymbol{E}_d^\top & \boldsymbol{T}_{\boldsymbol{U},d} H + T_{H,d+1} \boldsymbol{h}_{d+1,d+1} \\ \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} \boldsymbol{E}_d^\top & \boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{U},d}^{-1} & -\boldsymbol{T}_{\boldsymbol{U},d}^{-1} T_{H,d+1} \boldsymbol{\tau}_{d+1}^{-1} \\ & -\boldsymbol{\tau}_{d+1}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \widehat{\boldsymbol{H}}_d - T_H \boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1} E_d^\top & -\widehat{H}_{\text{new}} \\ -\boldsymbol{\tau}_{d+1} \boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1} E_d^\top & -\boldsymbol{\tau}_{d+1} \left(\boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1} E_d^\top T_H + \boldsymbol{h}_{d+1,d+1} \right) \boldsymbol{\tau}_{d+1}^{-1} \end{bmatrix}
\end{aligned}$$

   where $\widehat{H}_{\text{new}} = \left(\widehat{\boldsymbol{H}}_d - T_H \boldsymbol{h}_{d+1,d} \boldsymbol{\tau}_d^{-1} E_d^\top \right) T_{H,d+1} \boldsymbol{\tau}_{d+1}^{-1} + \boldsymbol{T}_{\boldsymbol{U},d} H \boldsymbol{\tau}_{d+1}^{-1} + T_{H,d+1} \boldsymbol{h}_{d+1,d+1} \boldsymbol{\tau}_{d+1}^{-1}$ and $H = \left[\boldsymbol{h}_{1,d+1}^\top, \ldots, \boldsymbol{h}_{d,d+1}^\top \right]^\top \in R^{dr \times r}$. A corresponding update for $\widehat{\boldsymbol{G}}_{d+1}$ is performed.

2. $\boldsymbol{E}_d = [zeros((d-1)*r, r); eye(r)]$ and $\boldsymbol{E}_1 = [eye(r); zeros((d-1)*r, r)]$

I could run the code, but the solution is too "bad", the residue satisfies tolerance while the solution NOT