On the Uniqueness of Solutions to Rational Expectations Models

Christopher Heiberger[a], Torben Klarl[b], Alfred Maußner[cd]

[a]University of Augsburg, Department of Economics, Universitätsstraße 16, 86159 Augsburg, Germany, christopher.heiberger@wiwi.uni-augsburg.de

[b]University of Augsburg, Department of Economics, Universitätsstraße 16, 86159 Augsburg, Germany, torben.alexander.klarl@wiwi.uni-augsburg.de

[c]Corresponding author, University of Augsburg, Department of Economics, Universitätsstraße 16, 86159 Augsburg, Germany, alfred.maussner@wiwi.uni-augsburg.de

[d]Alfred Maußner acknowledges financial support by the Deutsche Forschungsgemeinschaft within the priority program "Financial market imperfections and macroeconomic performance" under grant number MA 1110/3-1

December 17, 2014

JEL classification: C63, C88, E37

Key Words:

Linear Rational Expectations Models, Schur Decomposition, DSGE Models

Abstract

Klein (2000) advocates the use of the Schur decomposition of a matrix pencil to solve linear rational expectations models. Meanwhile his algorithm has become a center piece in several computer codes that provide approximate solutions to (non-linear) dynamic stochastic general equilibrium models. A subtlety not resolved by Klein is whether or not a certain Schur decomposition could fail to solve the model while a second one would provide a solution. We show that this cannot happen.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have become the workhorse of macroeconomic research. Among the various ways to solve these kind of models (see, e.g., Aruoba et al (2006) and Heer and Maußner (2008)) perturbation methods are the most popular ones. To obtain these solutions, a forward looking system of linear stochastic difference equations must be solved. Blanchard and Kahn (1980) propose to diagonalize the system. A unique solution exists, if there are as many stable and unstable eigenvalues as there are variables with and without given initial conditions, respectively. This approach has two disadvantages: i) the original system must be sufficiently reduced and ii) the Jordan decomposition that leads to the diagonal structure is numerically less reliable than other decompositions.

Paul Klein (2000) proposes to apply the numerically stable Schur decomposition for which state of the art computer algorithms exist. The Fortran and Matlab code written by Paul Klein has been widely used and is part of the code of DYNARE, a popular Matlab toolbox for the solution, simulation, and estimation of (non-linear) DSGE models.[1]

Klein (2000), p. 1419, points to a possible problem:

> "A subtlety in this context is that the generalized Schur form is not unique even if a particular ordering of the eigenvalues is imposed. It is therefore an open question whether there might be two generalized Schur forms of the same matrix pencil, one with Z_{11} invertible and the other with Z_{11} singular. A reasonable conjecture is that this cannot happen, but apparently there is no known proof of this."

In this note we provide a proof of his conjecture.

We depart slightly from Klein (2000) and set up the model in the way Paul Klein does in the latest version of his computer code. The advantage of this approach is that it is not necessary to solve the unstable block of the triangularized model forward. Instead, this is taken care of by the linear algebra package (LAPACK) routine that provides the decomposition. We then prove that any two different Schur forms yield the same solution, given there is one at all.

[1] The respective programs are Solab.f90 and Solab.

From here we proceed with a brief description of the canonical linear rational expectations (RE) model and derive its solution based on the Schur decomposition in the next section. Section 3 provides our proof and section 4 concludes.

2 Analytical Framework

Let $\mathbf{x}_t \in \mathbb{R}^{n(x)}$, $\mathbf{y}_t \in \mathbb{R}^{n(y)}$, and $\mathbf{z}_t \in \mathbb{R}^{n(z)}$, denote a vector of variables with given initial conditions at time t, a vector of not predetermined (jump) variables, and a vector of purely exogenous variables, respectively. The linear RE model that we want to solve is given by:

$$A\mathbb{E}_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = B \begin{bmatrix} \mathbf{x}_t \\ \mathbf{y}_t \end{bmatrix} + C\mathbf{z}_t, \tag{2.1a}$$

$$\mathbf{z}_t = \Phi \mathbf{z}_{t-1} + \sigma\Omega\boldsymbol{\epsilon}_t. \tag{2.1b}$$

A, B, C, Φ, and Ω are given matrices and $\sigma \geq 0$ is a scaling factor. $\mathbb{E}_t$ denotes expectations conditional on information available at time t. $\boldsymbol{\epsilon}_t$ is iid with $\mathbb{E}_t(\boldsymbol{\epsilon}) = \mathbf{0}_{n(z)\times 1}$ and covariance matrix $I_{n(z)}$. The matrix Φ has all eigenvalues within the unit circle so that $\mathbf{z}_t$ is a stationary stochastic process.

The set up of Klein (2000) is more general in terms of the stochastic process $\mathbf{z}_t$ and with respect to what is meant by predetermined variables. We define these as in Blanchard and Kahn (1980) as variables with a given initial condition at time t. Almost all models used in applied research fit in this more restrictive framework, which is also used by algorithms that provide higher order approximate solutions of DSGE models, as, e.g., Schmitt-Grohé and Uribe (2004) and Gomme and Klein (2011).

We rewrite the system (2.1a) as:

$$\begin{aligned}
&\tilde{A}\mathbb{E}_t \begin{bmatrix} \mathbf{w}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \tilde{B} \begin{bmatrix} \mathbf{w}_t \\ \mathbf{y}_t \end{bmatrix}, \\
&\mathbf{w}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{bmatrix}, \\
&\tilde{A} = \begin{bmatrix} A_{11} & 0_{n(x)\times n(z)} & A_{12} \\ A_{21} & 0_{n(y)\times n(z)} & A_{22} \\ 0_{n(z)\times n(x)} & I_{n(z)} & 0_{n(z)\times n(y)} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_{11} & C_1 & B_{12} \\ B_{21} & C_2 & B_{22} \\ 0_{n(z)\times n(x)} & \Phi & 0_{n(z)\times n(y)} \end{bmatrix}.
\end{aligned} \tag{2.2}$$

A_{11} denotes the upper $n(x) \times n(x)$ block of A, A_{12} the upper $n(x) \times n(y)$ block and so forth.

The generalized Schur factorization of the matrix pencil $(B - \lambda A)$ is given by

$$
\begin{aligned}
QSZ^H &= \tilde{A}, \\
QTZ^H &= \tilde{B},
\end{aligned}
\tag{2.3}
$$

where Q and Z are complex unitary matrices and S and T are complex upper triangular matrixes. Z^H is the Hermitian transpose of Z. We define new variables:

$$
\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}}_t \\ \tilde{\mathbf{y}}_t \end{bmatrix} = \begin{bmatrix} \mathbf{w}_t \\ \mathbf{y}_t \end{bmatrix}, \tag{2.4}
$$

so that we can write (2.2) as

$$
\begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} \mathbb{E}_t \begin{bmatrix} \tilde{\mathbf{w}}_{t+1} \\ \tilde{\mathbf{y}}_{t+1} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}}_t \\ \tilde{\mathbf{y}}_t \end{bmatrix}. \tag{2.5}
$$

Assume that for $i = 1, \ldots n(w)$ the diagonal elements of S and T are such that $|s_{ii}| > |t_{ii}| \geq 0$ and that for $i = n(w) + 1, \ldots, n(w) + n(y)$ $0 \leq |s_{ii}| < |t_{ii}|$. Given these assumptions and definitions, the system

$$
S_{22}\mathbb{E}_t\tilde{\mathbf{y}}_{t+1} = T_{22}\tilde{\mathbf{y}}_t
$$

is unstable,[2] and to obtain a definite solution, we must set $\tilde{\mathbf{y}}_t = \mathbf{0}_{n(y)}$ for all t. Thus, from the first line of (2.5)

$$
\tilde{\mathbf{w}}_{t+1} = S_{11}^{-1}T_{11}\tilde{\mathbf{w}}_t.
$$

To get the solution of the original system, we must assume that the matrix Z_{11} is invertible so that the first line of (2.4) can be solved for:

$$
\tilde{\mathbf{w}}_t = Z_{11}^{-1}\mathbf{w}_t, \tag{2.6}
$$

and we get

$$
\mathbf{w}_{t+1} = \underbrace{Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}}_{L_w^w}\mathbf{w}_t.
$$

[2] To see this, consider the last line of this system, which may be written

$$
\mathbb{E}_t\tilde{y}_{n(y),t+1} = \lambda_{n(y),n(y)}\tilde{y}_{n(y),t}, \quad |\lambda_{n(y),n(y)}| = |(t_{n(y),n(y)}/s_{n(y),n(y)})| > 1.
$$

The second line of (2.4) together with (2.6) implies

$$\mathbf{y}_t = \underbrace{Z_{21}Z_{11}^{-1}}_{L_w^y}\mathbf{w}_t.$$

The solved linear model is

$$\mathbf{x}_{t+1} = L_x^x \mathbf{x}_t + L_z^x \mathbf{z}_t, \tag{2.7a}$$

$$\mathbf{y}_{t+1} = L_x^y \mathbf{x}_t + L_z^y \mathbf{z}_t, \tag{2.7b}$$

$$\mathbf{z}_{t+1} = \Phi \mathbf{z}_t + \sigma\Omega\boldsymbol{\epsilon}_{t+1}. \tag{2.7c}$$

where

$$L_w^w = \begin{bmatrix} L_x^x & L_z^x \\ \mathbf{0}_{n(z)\times n(x)} & \Phi \end{bmatrix}, \quad L_w^y = \begin{bmatrix} L_x^y & L_z^y \end{bmatrix}.$$

3 Uniqueness

The Schur decomposition is not unique. Thus, we cannot be sure that the mapping between $\mathbf{w}_t$ and $\tilde{\mathbf{w}}_t$ is unique. Consider the

Proposition. *Let A and B denote two complex $n \times n$ matrices and consider the two decompositions*

$$\begin{aligned} QTZ^H &= A = \tilde{Q}\tilde{S}\tilde{Z}^H, \\ QSZ^H &= B = \tilde{Q}\tilde{T}\tilde{Z}^H, \end{aligned} \tag{3.1}$$

where Q, $\tilde{Q}$, Z, and $\tilde{Z}$ are unitary matrices, $QQ^H = I_n$ while T, $\tilde{T}$, S, and $\tilde{S}$ are upper triangular matrices. Assume:

A.1 The matrix pencil $A - \lambda B$ has finitely many generalized eigenvalues λ.

A.2 None of these eigenvalues lies on the unit circle.

A.3 There is a $k \in \{1, 2, \ldots, n\}$ so that:
$|t_{ii}| > |s_{ii}|$ and $|\tilde{s}_{ii}| > |\tilde{t}_{ii}|$ for $i = 1, 2, \ldots, k$
$|t_{ii}| < |s_{ii}|$ and $|\tilde{s}_{ii}| < |\tilde{t}_{ii}|$ for $i = k+1, \ldots, n$.

A.4

$$Z_{11} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1k} \\ z_{21} & z_{22} & \dots & z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ z_{k1} & z_{k2} & \dots & z_{kk} \end{bmatrix}$$

is invertible.

Then

$$Z_{11}^{-1}\, \textit{exists} \Rightarrow \tilde{Z}_{11}^{-1}\, \textit{exists}, \tag{3.2}$$

$$L_w^w = Z_{11} T_{11}^{-1} S_{11} Z_{11}^{-1} = \tilde{Z}_{11} \tilde{S}_{11}^{-1} \tilde{T}_{11} \tilde{Z}_{11}^{-1}, \tag{3.3}$$

$$L_w^y = Z_{21} Z_{11}^{-1} = \tilde{Z}_{21} \tilde{Z}_{11}^{-1}, \tag{3.4}$$

where X_{11} denotes the $k \times k$ upper left block and X_{21} denotes the $(n-k) \times k$ lower left block of $X \in \{T, S, Z, \tilde{T}, \tilde{S}, \tilde{Z}\}$. □

Remarks.

R.1: A.1 implies $|A - \lambda B| \neq 0$ for at least one $\lambda \in \mathbb{C}$ and excludes $t_{ii} = s_{ii} = 0$ and $\tilde{t}_{ii} = \tilde{s}_{ii} = 0$, so that for $i = 1, \dots, k$, $|t_{ii}|, |\tilde{s}_{ii}| > 0$ according to A.3.

R.2: The existence of the decomposition (3.1) follows from Theorem 7.7.1 in Golub and Van Loan (1996).

As preliminary step, let

$$M = \tilde{Z}^H Z. \tag{3.5}$$

In the Appendix, we prove the following

Lemma.

$$M = \begin{bmatrix} M_{11} & 0_{k \times (n-k)} \\ 0_{(n-k) \times k} & M_{22} \end{bmatrix},$$

M_{11} and M_{22} are unitary matrices, i.e., $M_{11} M_{11}^H = I_k$ and $M_{22} M_{22}^H = I_{n-k}$. □

Proof of the Proposition. The Lemma implies

$$\begin{aligned}
&Z = \tilde{Z}M \\
\Leftrightarrow &\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix} \begin{bmatrix} M_{11} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & M_{22} \end{bmatrix} = \begin{bmatrix} \tilde{Z}_{11}M_{11} & \tilde{Z}_{12}M_{22} \\ \tilde{Z}_{21}M_{11} & \tilde{Z}_{22}M_{22} \end{bmatrix}
\end{aligned}$$

yielding

$$\begin{aligned}
Z_{11} &= \tilde{Z}_{11}M_{11}, \\
Z_{21} &= \tilde{Z}_{21}M_{11}.
\end{aligned} \tag{3.6}$$

Note that (3.6) together with the invertibility of M_{11} implies that there cannot be two transformations Z_{11} and $\tilde{Z}_{11}$, one being invertible and the other not, proving (3.2). Thus, the subtlety raised by Klein (2000) is resolved.

Furthermore, from the first line in (3.1), $\tilde{S} = \tilde{Q}^H QTZ^H\tilde{Z} = \tilde{Q}^H QTM^H$ implies

$$\tilde{S}_{11} = \underbrace{\left(\tilde{Q}^{11}Q_{11} + \tilde{Q}^{12}Q_{21}\right)}_{=:X} T_{11}M_{11}^H \Rightarrow T_{11} = X^{-1}\tilde{S}_{11}M_{11}, \tag{3.7}$$

where $\tilde{Q}^{11}$ and $\tilde{Q}^{12}$ denote the upper left and the upper right block of $\tilde{Q}^H$. The existence of X^{-1} follows from the fact that $\tilde{S}_{11}$, T_{11}, and M_{11}^H are invertible. Analogously, from the second line in (3.1), $\tilde{T} = \tilde{Q}^H QSZ^H\tilde{Z} = \tilde{Q}^H QSM^H$ implies

$$S_{11} = X^{-1}\tilde{T}_{11}M_{11}. \tag{3.8}$$

Using (3.6) to substitute for Z_{11} and Z_{21} on the right-hand side of (3.4), establishes the third part of the proposition. The second statement (3.3) follows from substituting for T_{11}^{-1} and S_{11} on the right-hand side of (3.3) using (3.7) and (3.8). □

A Caveat. The uniqueness of the solution implied by the Proposition is a theoretical result. In practice, i.e., taking into account finite precision computer arithmetic, the particular algorithm that factors $\tilde{A}$ and $\tilde{B}$ comes into play. In a companion paper, Heiberger et al (2014), we show by means of a model from the asset pricing literature that there can be noticable differences in the matrices L_w^w and L_w^y depending on the factorization employed.

4 Conclusion

Popular toolkits, like DYNARE, employ the generalized Schur decompostion of a matrix pencil to generate approximate solutions of non-linear DSGE models. The Schur decomposition, however, is not unique. Therefore, Klein (2000) raises the question whether it could happen that one decomposition fails to compute a solution while another one succeeds. We prove that this cannot happen. Given that the problem at hand satisfies the Blanchard and Kahn (1980) conditions and given that the transformation matrix is invertible, a unique solution exists. In numerical applications, however, solutions may differ, if the involved matrices are sufficiently ill-conditioned.

References

Aruoba, S. Boragan, Jesús Fernández-Villaverde and Juan F. Rubio-Ramí2006. Comparing Solution Methods for Dynamic Equilibrium Economies. *Journal of Economic Dynamics and Control*. Vol. 30. pp. 2477-2508

Blanchard, Oliver J. and Charles M. Kahn. 1980. The Solution of Linear Difference Models Under Rational Expectations. *Econometrica*. Vol. 48. pp. 1305-1311

Golub, Gene H. and Charles F. Van Loan. 1996. Matrix Computations. 3rd. Ed. The Johns Hopkins University Press: Baltimore und London.

Gomme, Paul and Paul Klein. 2011. Second-Order Approximation of Dynamic Models Without the Use of Tensors. *Journal of Economic Dynamics and Control*. Vol. 35. pp. 604-615.

Heer, Burkhard and Alfred Maußner. 2008. Computation of Business Cycle Models: A Comparison of Numerical Methods. *Macroeconomic Dynamics*. Vol. 12. pp. 641-663

Heiberger, Christopher, Torben Klarl, and Alfred Maußner. 2014. On the Numerical Accuracy of First-Order Approximate Solutions to DSGE Models. University of Augsburg. Mimeo.

Klein, Paul. 2000. Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control*. Vol. 24. pp. 1405-1423

Schmitt-Grohé, Stephanie and Martin Uribe. 2004. Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. *Journal of Economic Dynamics and Control*. Vol. 28. pp. 755-775.

Appendix: Proof of the Lemma (Not for publication)

Let $T_{11} = (t_{ij})$ for $i, j = 1, \dots k$ and similarly for the other matrices in (3.1) yielding the partitioning:

$$T = \begin{bmatrix} T_{11} & T_{12} \\ 0_{(n-k)\times k} & T_{22} \end{bmatrix}, \tilde{T} = \begin{bmatrix} \tilde{T}_{11} & \tilde{T}_{12} \\ 0_{(n-k)\times k} & \tilde{T}_{22} \end{bmatrix}, S = \begin{bmatrix} S_{11} & S_{12} \\ 0_{(n-k)\times k} & T_{22} \end{bmatrix}, \tilde{S} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ 0_{(n-k)\times k} & \tilde{S}_{22} \end{bmatrix},$$
$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{21} & \tilde{Q}_{22} \end{bmatrix}, \quad Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}, \quad \tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix}.$$

Next, let $\mathbf{z}_j$ denote the columns of Z,

$$Z = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n],$$

and define the matrix M by

$$M = \tilde{Z}^H Z,$$

with partition

$$M = \begin{bmatrix} m_{11} & \dots & m_{1k} & m_{1k+1} & \dots & m_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m_{k1} & \dots & m_{kk} & m_{kk+1} & \dots & m_{kn} \\ m_{k+11} & \dots & m_{k+1k} & m_{k+1k+1} & \dots & m_{k+1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nk} & m_{nk+1} & \dots & m_{nn} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}.$$

Thus, we may write:

$$A[\mathbf{z}_1, \dots, \mathbf{z}_k] = QTZ^H[\mathbf{z}_1, \dots, \mathbf{z}_k] = QT[\mathbf{e}_1, \dots, \mathbf{e}_k] = Q \begin{bmatrix} T_{11} \\ 0_{(n-k)\times k} \end{bmatrix}$$

and

$$B[\mathbf{z}_1, \dots, \mathbf{z}_k] = QSZ^H[\mathbf{z}_1, \dots, \mathbf{z}_k] = QS[\mathbf{e}_1, \dots, \mathbf{e}_k] = Q \begin{bmatrix} S_{11} \\ 0_{(n-k)\times k} \end{bmatrix},$$
$$= Q \begin{bmatrix} T_{11} \\ 0_{(n-k)\times k} \end{bmatrix} T_{11}^{-1} S_{11}.$$

Therefore:

$$B[\mathbf{z}_1, \dots, \mathbf{z}_k] = A[\mathbf{z}_1, \dots, \mathbf{z}_k] T_{11}^{-1} S_{11}. \tag{A.1}$$

Considering the respective right-hand sides of (3.1), we may also write:

$$A[\mathbf{z}_1,\ldots,\mathbf{z}_k] = \tilde{Q}\tilde{S}\tilde{Z}^H[\mathbf{z}_1,\ldots,\mathbf{z}_k]= \tilde{Q}\tilde{S}\begin{bmatrix}M_{11}\\ M_{21}\end{bmatrix} = \tilde{Q}\begin{bmatrix}\tilde{S}_{11}M_{11}+\tilde{S}_{12}M_{21}\\ \tilde{S}_{22}M_{21}\end{bmatrix},$$

$$B[\mathbf{z}_1,\ldots,\mathbf{z}_k] = \tilde{Q}\tilde{T}\tilde{Z}^H[\mathbf{z}_1,\ldots,\mathbf{z}_k]= \tilde{Q}\tilde{T}\begin{bmatrix}M_{11}\\ M_{21}\end{bmatrix} = \tilde{Q}\begin{bmatrix}\tilde{T}_{11}M_{11}+\tilde{T}_{12}M_{21}\\ \tilde{T}_{22}M_{21}\end{bmatrix}.$$

Employing (A.1) yields:

$$\tilde{Q}\begin{bmatrix}\tilde{T}_{11}M_{11}+\tilde{T}_{12}M_{21}\\ \tilde{T}_{22}M_{21}\end{bmatrix} = \tilde{Q}\begin{bmatrix}\tilde{S}_{11}M_{11}+\tilde{S}_{12}M_{21}\\ \tilde{S}_{22}M_{21}\end{bmatrix}T_{11}^{-1}S_{11}.$$

Since $\tilde{Q}^H\tilde{Q} = I_n$, the second line of the previous matrix equation implies:

$$\tilde{T}_{22}M_{21} = \tilde{S}_{22}M_{21}T_{11}^{-1}S_{11}. \tag{A.2}$$

Note that T_{11}^{-1}, being the inverse of an upper triangular matrix, is itself an upper triangular matrix with diagonal elements $t^{ii} = \frac{1}{t_{ii}}$, $i = 1,\ldots,k$. Let $P = (p_{ij})$, $i,j = 1,\ldots,k$ denote the matrix $P = T_{11}^{-1}S_{11}$ with diagonal elements $p_{ii} = s_{ii}/t_{ii}$. Equation (A.2) yields:

$$\sum_{l=i}^{n}\tilde{t}_{il}m_{lj} = \sum_{l=i}^{n}\tilde{s}_{il}\sum_{h=1}^{j}m_{lh}p_{hj} \text{ for } i = k+1,\ldots,n, j = 1,\ldots,k. \tag{A.3}$$

We use induction over j and i to show $M_{21} = 0_{(n-k)\times k}$.

- $j = 1$:
  - $i = n$:
    In this case (A.3) reduces to
    $$m_{n1} = \underbrace{\frac{\tilde{s}_{nn}}{\tilde{t}_{nn}}\frac{s_{11}}{t_{11}}}_{|\cdot|<1}m_{n1},$$
    where the inequality follows from A.3. Therefore: $m_{n1} = 0$.
  - $m_{i+11} = 0 \Rightarrow m_{i1} = 0$:
    In this case (A.3) reduces to
    $$m_{i1} = \underbrace{\frac{\tilde{s}_{ii}}{\tilde{t}_{ii}}\frac{s_{ii}}{t_{ii}}}_{|\cdot|<1}m_{i1}$$
    so that indeed $m_{i1} = 0$.

- $m_{ij} = 0 \Rightarrow m_{ij+1} = 0$:

  - $i = n$:
    Thus, $m_{n1} = m_{n2} = \cdots = m_{nj} = 0$ so that (A.3) reduces to

    $$m_{nj+1} = \underbrace{\frac{\tilde{s}_{nn}}{\tilde{t}_{nn}} \frac{s_{j+1j+1}}{t_{j+1j+1}}}_{|\cdot|<1} m_{nj+1}$$

    proving the assertion.

  - $m_{nj+1} = m_{n-1j+1} = \cdots = m_{i+1j+1} = 0 \Rightarrow m_{ij+1} = 0$:
    In this case (A.3) reduces to

    $$m_{ij+1} = \underbrace{\frac{\tilde{s}_{ii}}{\tilde{t}_{ii}} \frac{s_{j+1j+1}}{t_{j+1j+1}}}_{|\cdot|<1} m_{ij+1}$$

    proving the assertion.

This results allows us to partition M as:

$$M = \begin{bmatrix} M_{11} & M_{12} \\ 0_{(n-k)\times k} & M_{22} \end{bmatrix}.$$

Since

$$\begin{aligned} M^H M = (Z^H \tilde{Z} \tilde{Z}^H Z) = I_n &= \begin{bmatrix} M_{11}^H & 0_{k\times(n-k)} \\ M_{12}^H & M_{22}^H \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ 0_{(n-k)\times k} & M_{22} \end{bmatrix} \\ &= \begin{bmatrix} M_{11}^H M_{11} & M_{11}^H M_{12} \\ M_{12}^H M_{11} & M_{12}^H M_{12} + M_{22}^H M_{22} \end{bmatrix}, \end{aligned}$$

we get:

$$\begin{aligned} I_k &= M_{11}^H M_{11}, \\ 0_{k\times(n-k)} &= M_{12}, \\ I_{n-k} &= M_{22}^H M_{22} \end{aligned}$$

so that M_{11} and M_{22} are unitary matrices and

$$M = \begin{bmatrix} M_{11} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & M_{22} \end{bmatrix}.$$

□