Computational linear algebra aspects of projector based treatment of DAEs

René Lamour[1], Roswitha März[1], Caren Tischendorf [2]

Abstract

The characterization of DAEs by the tractability index concept bases on the investigation of the related matrix sequence. The algorithmic steps of the implementation of the sequence consist of the computation of admissible nullspace projectors, including the widely orthogonal projectors, and ways to perform the differentiation of projectors, if neccessary. Different aspects to compute the required objects are discussed.

KeyWords: DAE, differential algebraic equation, admissible projector, widely orthogonal projector, matrix sequence, tractability index

MCS 2000: 34A09, 65L80

[1]Humboldt-University of Berlin, Dept. of Mathematics
[2]University of Cologne, Mathematical Institute

Computational linear algebra aspects of projector based treatment of DAEs

Originally, the tractability index concept is designed rather for the theoretical investigation of DAEs. However, the resulting clear index criteria by rank conditions let us trust that it has also practical meaning. Moreover, the projectors prove their value when characterizing the different solution components, when looking for consistent initial values and formulating appropriate initial conditions as well. And these are good arguments to implement the associated matrix function sequences.
We study the equation

$$A(Dx)' + Bx = q, \tag{1}$$

with continuous coefficients

$$A \in \mathcal{C}(\mathcal{I}, L(\mathbb{R}^n, \mathbb{R}^k)), \quad D \in \mathcal{C}(\mathcal{I}, L(\mathbb{R}^m, \mathbb{R}^n)), \quad B \in \mathcal{C}(\mathcal{I}, L(\mathbb{R}^m, \mathbb{R}^k)),$$

and an excitation $q \in \mathcal{C}(\mathcal{I}, \mathbb{R}^k)$. $\mathcal{I} \in \mathbb{R}$ is an interval. The coefficients A and D are supposed to be well matched (cf. (2)). The algorithmic realization of a matrix function sequence

$$\begin{aligned} G_{i+1} &= G_i + B_i Q_i, \\ B_{i+1} &= B_i P_i - G_{i+1} D^- (D\Pi_{i+1} D^-)' D\Pi_i \end{aligned}$$

starting with $G_0 = AD,\ B_0 = B$, requires the computation of the involved generalized inverse D^- and the admissible projectors Q_i onto $\ker G_i$ with $P_i = I - Q_i$, $\Pi_i := P_i \Pi_{i-1},\ \Pi_0 := P_0$. (cf. [LMT11], Definitions 2.3).

For a DAE that has the leading term $A(t)(D(t)x(t))'$, it is also important to check whether this leading term is actually properly stated by testing the transversality condition

$$\ker A(t) \oplus \operatorname{im} D(t) = \mathbb{R}^n. \tag{2}$$

The last question is considered in Section 2, whereas basics on the computation of nullspace and image projectors associated to matrices are collected in Section 1. At this point we also bring to mind the detailed Appendix A of [LMT11] on linear algebra issues. Methods to compute a suitable generalized inverse D^- are described in Section 1. In Section 3 we deal with the basic step of the construction of admissible matrix functions, that is, with the step from level i to level $i + 1$ by the computation of an appropriate projector. After that, in Section 4, sequences of matrices with admissible projectors are delivered, first level by level on the background of Section 3, and then by a strongly involved version only for the regular case.
We stress that all the computation are more or less related to matrix decompositions and rank calculations, and, naturally, one has to expect to inherit all the related numerical problems.

1 Image and nullspace projectors

For given $G \in \mathbb{R}^{k \times m}$, with $\operatorname{rank} G = r$, any matrix $Q \in \mathbb{R}^{m \times m}$ that satisfies

$$GQ = 0, \quad Q^2 = Q, \quad \operatorname{rank} Q = m - r$$

is a projector onto $\ker G$. Any matrix $W \in \mathbb{R}^{k \times k}$ that satisfies

$$WG = 0, \quad W^2 = W, \quad \operatorname{rank} W = k - r$$

is a projector along $\operatorname{im} G$.

Clearly, having a basis of the subspace in question, a required projector can immediately be described by these basis elements (cf. Lemma A.6 in [LMT11]). In particular, if $n_1, \ldots, n_{m-r} \in R^m$ form a basis of $\ker G$ and $\Gamma := \begin{bmatrix} n_1 & \cdots & n_{m-r} \end{bmatrix}$, then $Q = \Gamma(\Gamma^*\Gamma)^{-1}\Gamma^*$ represents the orthogonal projector onto this nullspace. If the $n_1, \ldots, n_{m-r}$ form an orthonormal basis, the expression simplifies to

$$Q = \Gamma\Gamma^* = \sum_{i=1}^{m-r} n_i n_i^*.$$

In other words, the knowledge of an orthonormal basis can immediately be used to form an orthogonal projector as the sum of the dyadic product of the basis vectors. For problems of limited dimension a formula manipulation system like *Mathematica*® or *Maple*® can be used to compute a basis. The command in Mathematica is `NullSpace`[G] and in Maple `nullspace`(G).
However, to provide a basis of the nullspace of a given matrix one usually has to carry out a factorization, for instance a singular value decomposition (SVD).

If a generalized reflexive inverse G^- (cf. Appendix B in [LMT11]) is known, we gain at the same time the nullspace projector $Q = I - G^-G$ and the projector along the image $W = I - GG^-$. To compute a generalized inverse of the given matrix G, again a factorization of that matrix serves as an appropriate tool.
Each decomposition

$$G = \mathcal{U} \begin{bmatrix} S & \\ & 0 \end{bmatrix} \mathcal{V}^{-1}, \tag{3}$$

with nonsingular $S \in \mathbb{R}^{r \times r}$, $\mathcal{U} =: \begin{bmatrix} U_1 & U_2 \end{bmatrix} \in \mathbb{R}^{k \times k}$ and $\mathcal{V} =: \begin{bmatrix} V_1 & V_2 \end{bmatrix} \in \mathbb{R}^{m \times m}$, and $U_1 \in \mathbb{R}^{k \times r}$, $V_2 \in \mathbb{R}^{m \times (m-r)}$, delivers immediately the basises $\ker G = \operatorname{span} V_2$ and $\operatorname{im} G = \operatorname{span} U_1$ as well as (3) the family of reflexive generalized inverses of G by

$$G^- = \mathcal{V} \begin{bmatrix} S^{-1} & M_2 \\ M_1 & M_1 S M_2 \end{bmatrix} \mathcal{U}^{-1}, \tag{4}$$

with the free parameter matrices M_1 and M_2 (see Appendix [LMT11, (88)]). The resulting projectors are

$$Q = \mathcal{V} \begin{bmatrix} 0 & \\ -M_1 S & I \end{bmatrix} \mathcal{V}^{-1} \text{ and } \quad W = \mathcal{U} \begin{bmatrix} 0 & -SM_2 \\ & I \end{bmatrix} \mathcal{U}^{-1}.$$

If we are looking for orthogonal projectors, we have to ensure symmetry, that is $\mathcal{U}^{-1} = \mathcal{U}^*$, $\mathcal{V}^{-1} = \mathcal{V}^*$, $M_1 \equiv 0$ and $M_2 \equiv 0$.

There are different ways to generate matrix decompositions (3). Applying the SVD one delivers orthogonal matrices $\mathcal{U}$ and $\mathcal{V}$, and the orthogonal projector Q is given by

$$Q = \begin{bmatrix} \mathcal{V}_1 & \mathcal{V}_2 \end{bmatrix} \begin{bmatrix} 0 & \\ & I \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^* \\ \mathcal{V}_2^* \end{bmatrix} = \mathcal{V}_2 \mathcal{V}_2^*. \tag{5}$$

Also the Householder method is suitable for computing a decomposition (3). The Householder decomposition needs less computational work than the SVD. For a singular matrix G, a Householder decomposition with column pivoting is needed. We obtain

$$G I_{per} = U \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$$

with a column permutation matrix I_{per}, an orthogonal matrix U and a nonsingular upper triangular matrix R_1. The required decomposition (3) has then the structure

$$G = U \begin{bmatrix} R_1 & \\ & 0 \end{bmatrix} \underbrace{\begin{bmatrix} I & R_1^{-1} R_2 \\ & I \end{bmatrix} I_{per}^*}_{=:\mathcal{V}^{-1}}, \tag{6}$$

and hence the nullspace projector

$$Q = I_{per} \begin{bmatrix} I & -R_1^{-1} R_2 \\ & I \end{bmatrix} \begin{bmatrix} 0 & \\ -M_1 R_1 & I \end{bmatrix} \begin{bmatrix} I & R_1^{-1} R_2 \\ & I \end{bmatrix} I_{per}^*$$

and the projector

$$W = U \begin{bmatrix} 0 & -R_1 M_2 \\ & I \end{bmatrix} U^*$$

along the image of G result. The free parameter matrices M_1 and M_2 can be used to provide special properties of the projectors as, for instance, we do in Section 4.
Since the Householder method provides an orthogonal matrix U, choosing $M_2 = 0$ we arrive at an orthoprojector W. If we apply the Householder method to G^* instead of G, we also deliver an orthogonal nullspace projector for G.

In principle also an LU-decomposition of G using the Gaussian method with scaling and pivoting yields a decomposition (3). With a row permutation matrix I_{per} we obtain

$$I_{per} G = LU = \begin{bmatrix} L_1 & \\ L_2 & I \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ & 0 \end{bmatrix}$$

and the decomposition

$$G = \underbrace{I_{per}^* \begin{bmatrix} L_1 & \\ L_2 & I \end{bmatrix}}_{=:\mathcal{U}} \begin{bmatrix} R_1 & \\ & 0 \end{bmatrix} \underbrace{\begin{bmatrix} I & R_1^{-1} R_2 \\ & I \end{bmatrix}}_{=:\mathcal{V}^{-1}}.$$

It is well-known that rank determination by the Gaussian method is not as robust as it is from the Householder method or SVD (cf. [GK65]), which is confirmed by our practical tests. We do not recommend this method here.

2 Matters of a properly stated leading term

Having a pair of matrices A and D one might be interested in making sure whether they are *well matched* in the sense of (2). For instance, one can check pointwise if the DAE has even a proper leading term. This way critical points can be indicated and eventual programming errors in hand-written subroutines as well.
Moreover, when generating the basic matrix function sequences starting pointwise from the given coefficients A, D, and B, the reflexive generalized inverses D^- and the border projector R play their role.

Let the two matrices $A \in \mathbb{R}^{k\times n}$ and $D \in \mathbb{R}^{n\times m}$ be given and $G := AD$. Then the inclusions $\operatorname{im} G \subseteq \operatorname{im} A$ and $\ker D \subseteq \ker G$ are valid. Owing to Lemma A.3 in [LMT11] A and D are well matched, exactly if

$$\operatorname{rank} A = \operatorname{rank} G = \operatorname{rank} D, \tag{7}$$

$$\operatorname{im} G = \operatorname{im} A, \tag{8}$$

$$\ker D = \ker G. \tag{9}$$

The failure of one of these three conditions indicates that A and D miss the mark. Since, in turn, (8), (9) imply the rank condition (7), these two conditions already ensure the well-matchedness.

Let G^- denote a reflexive generalized inverse of G, e.g. provided by a decomposition (4). Then the conditions (8), (9) can be written as

$$(I - GG^-)A = 0, \tag{10}$$

$$D(I - G^-G) = 0, \tag{11}$$

and these conditions are also useful for testing the well-matchedness.

Next we suppose A and D to be well matched, and hence (10) and (11) to be valid. Then

$$D^- := G^-A, \quad A^- := DG^- \tag{12}$$

are reflexive generalized inverses of D and A, and

$$R := DD^- = DG^-A = A^-A$$

is nothing else the projector matrix onto $\operatorname{im} D$ along $\ker A$. Namely, it holds that

$$\begin{aligned} DD^-D &= DG^-AD = DG^-G = D, \\ D^-DD^- &= G^-ADG^-A = G^-A = D^-, \\ AA^-A &= ADG^-A = GG^-A = A, \\ A^-AA^- &= DG^-ADG^- = DG^- = A^-. \end{aligned}$$

It comes out that, decomposing G delivers at the same time a reflexive generalized inverse G^- such that one can first check the conditions (10) and (11), and then, supposed they hold true, form the generalized inverses D^-, A^- and the border projector R.

Stress at this point that an orthogonal projector is often preferable. It can be reached by a SVD applied to G or a Householder factorization applied to G^* (Section 1).

An alternative way to test well-matchedness of A and D and then to provide D^- and R uses factorizations of both matrices A and D. This makes sense, if the factorizations of A and D are given or easily available.
Suppose the decompositions (cf. (3)) of A and D as

$$A = U_A \begin{bmatrix} S_A & \\ & 0 \end{bmatrix} V_A^{-1} \text{ and } D = U_D \begin{bmatrix} S_D & \\ & 0 \end{bmatrix} V_D^{-1}. \tag{13}$$

We can check now the rank conditions $\operatorname{rank} S_A = \operatorname{rank} S_D$ which are necessary for well-matchedness (see (7)). Also $AD = G$ has to have the same rank. The decompositions yield

$$AD = U_A \begin{bmatrix} S_A & \\ & 0 \end{bmatrix} V_A^{-1} U_D \begin{bmatrix} S_D & \\ & 0 \end{bmatrix} V_D^{-1} \tag{14}$$

and, denoting $V_A^{-1} U_D =: H = \begin{bmatrix} H_1 & H_2 \\ H_3 & H_4 \end{bmatrix}$, the necessary rank condition is satisfied iff H_1 remains nonsingular.

The generalized inverses of D and A are not independent of each other, but they have to satisfy the relation $DD^- = A^- A$. Using the given decompositions (13) the reflexive generalized inverses are immediately found (see [LMT11, (88)]) as

$$A^- = V_A \begin{bmatrix} S_A^{-1} & M_{2,A} \\ M_{1,A} & M_{1,A} S_A M_{2,A} \end{bmatrix} U_A^{-1} \text{ and } D^- = V_D \begin{bmatrix} S_D^{-1} & M_{2,D} \\ M_{1,D} & M_{1,D} S_D M_{2,D} \end{bmatrix} U_D^{-1},$$

which leads to

$$DD^- = U_D \begin{bmatrix} I & S_D M_{2,D} \\ 0 & 0 \end{bmatrix} U_D^{-1}, \quad A^- A = V_A \begin{bmatrix} I & 0 \\ M_{1,A} S_A & 0 \end{bmatrix} V_A^{-1}.$$

Using again the denotation $U_D = V_A H$, the relation $DD^- = A^- A$ becomes equivalent with

$$H \begin{bmatrix} I & S_D M_{2,D} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ M_{1,A} S_A & 0 \end{bmatrix} H.$$

This fixes two of the free parameter matrices, namely

$$M_{2,D} = S_D^{-1} H_1^{-1} H_2 \tag{15}$$

and

$$M_{1,A} = H_3 H_1^{-1} S_A^{-1}.$$

The other two parameter matrices $M_{1,D}$ and $M_{2,A}$ can be used to ensure further properties.

Finally in this section, we shortly turn to standard form DAEs given with a leading term of the form $Gx'(t)$. A factorization (3) is then adjuvant in determining a properly stated leading term version. We can define A and D as

$$
\begin{array}{llllll}
\text{(a)} & A & = & \mathcal{U}\begin{bmatrix} S & \\ & 0 \end{bmatrix}, & D & = & \mathcal{V}^{-1}, \\
\text{(b)} & A & = & \mathcal{U}, & D & = & \begin{bmatrix} S & \\ & 0 \end{bmatrix}\mathcal{V}^{-1},
\end{array}
$$

$$
\text{and with } \mathcal{U} =: \begin{bmatrix} \mathcal{U}_1, \mathcal{U}_2 \end{bmatrix} \quad \text{and } \mathcal{V}^{-1} =: \begin{bmatrix} (\mathcal{V}^{-1})_1 \\ (\mathcal{V}^{-1})_2 \end{bmatrix}
$$

$$
\begin{array}{llllll}
\text{(c)} & A & = & \mathcal{U}_1 S, & D & = & (\mathcal{V}^{-1})_1 \in \mathbb{R}^{r\times m}, \\
\text{(d)} & A & = & \mathcal{U}_1, & D & = & S(\mathcal{V}^{-1})_1.
\end{array}
$$

The cases (c) and (d) provide the splitting with full rank matrices A and D, which is advantageous e.g. because the border projector is simply $R = I$.
Analogously one can proceed in case of time-varying terms coefficients $G(t)$, but then one needs a continuous matrix decomposition and a continuously differentiable $D(\cdot)$ as well as its derivative.
Notice that often standard form DAEs are given with separated derivative free equations such that a continuous projector function $I - W(t)$ onto $\operatorname{im} G(t)$ is available at the beginning. Then one can take use of this situation and put $A(t) := I - W(t)$, $D(t) := G(t)$.

3 The basic step of the sequence

Now we consider the basic part of the determination of an admissible matrix function sequence, that is the step from G_i to G_{i+1}. Let a projector $\Pi_i := P_0 \cdots P_i$ be already computed. We are looking for the next admissible projector Q_{i+1}. An admissible projector must satisfy the required properties (cf. [LMT11], Definition 2.3). If we are dealing with matrix functions, the determinations are carried out pointwise for frozen arguments. In the following we suppress the step index i. G complies with $G_{i+1}(z)$ and Π with $\Pi_i(z)$, z is an arbitrary frozen argument.

For a given matrix $G \in \mathbb{R}^{k\times m}$ with $\operatorname{rank} G = r$ and a given projector $\Pi \in \mathbb{R}^{m\times m}$ with $\operatorname{rank} \Pi = \rho$, we seek a new projector matrix Q such that

$$
\begin{aligned}
\operatorname{im} Q &= \ker G, \\
\ker Q &\supseteq X \text{ (cf. (14) in [LMT11])}
\end{aligned}
$$

and X is any complement of $\widehat{N} := \ker \Pi \cap \operatorname{im} Q$ in $\ker \Pi$ (cf. (13) in [LMT11]), which means, Q has to satisfy (cf. Proposition 2.6 (3) in [LMT11]) the conditions

$$
GQ = 0, \ \operatorname{rank} Q = m - r, \tag{16}
$$

$$
\Pi Q(I - \Pi) = 0. \tag{17}
$$

Owing to Lemma A.6 in [LMT11] such a projector Q exists. Denote $N := \ker G$ and $K := \ker \Pi = \operatorname{im}(I - \Pi)$. Condition (16) implies $\operatorname{im} Q = N$. If N and K intersect only trivially, i.e., $K \cap N = \{0\}$, which we call the *regular case*, we can form Q to satisfy

$X = K \subseteq \ker Q$, and then condition (17) takes place. In general a computation of a representation of X is needed. We have to fix a set $X \subseteq K$ such that $K = X \oplus N$. Notice that X is not uniquely defined.
An example illustrates the situation.

Example 3.1 *For* $\Pi = \begin{bmatrix} 0 & & \\ & 0 & \\ & & 1 \end{bmatrix}$ *and* $G = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ & & 0 \end{bmatrix}$, $m = 3$, *we obtain* $K = \ker \Pi = \operatorname{span} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ *and* $N = \ker G = \operatorname{span} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$, *further* $K \cap N = \operatorname{span} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ *and* $K \oplus N = \mathbb{R}^m$. *Any plane given by* $(K \cap N)^c := \operatorname{span} \begin{bmatrix} \cos\alpha & 0 \\ \sin\alpha & -\cos\alpha \\ 0 & \beta \end{bmatrix}$, *with fixed* $\alpha \in (0, \pi)$, *and* $\beta \neq 0$, *is a complement of* $K \cap N$ *in* $\mathbb{R}^m$. *A possible subspace* X *can be given as* $X = K \cap (K \cap N)^c = \operatorname{span} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix}$. *As we can see by the different choices of* α *and* β, *the complement* $(K \cap N)^c$ *as well as* X *are not unique. For reasons of dimensions, in this example, since* $\dim(K + N) = m$, *the projector onto* N *along* X *is uniquely determined as*

$$Q = \begin{bmatrix} 1 & -\frac{\cos\alpha}{\sin\alpha} & \frac{\cos\alpha}{\sin\alpha} \\ & 0 & 1 \\ & & 1 \end{bmatrix}.$$

The Figure 1 shows this case. In general, $R^m = \underbrace{N \oplus X}_{K+N} \oplus (K+N)^c$ *holds with a nontrivial*

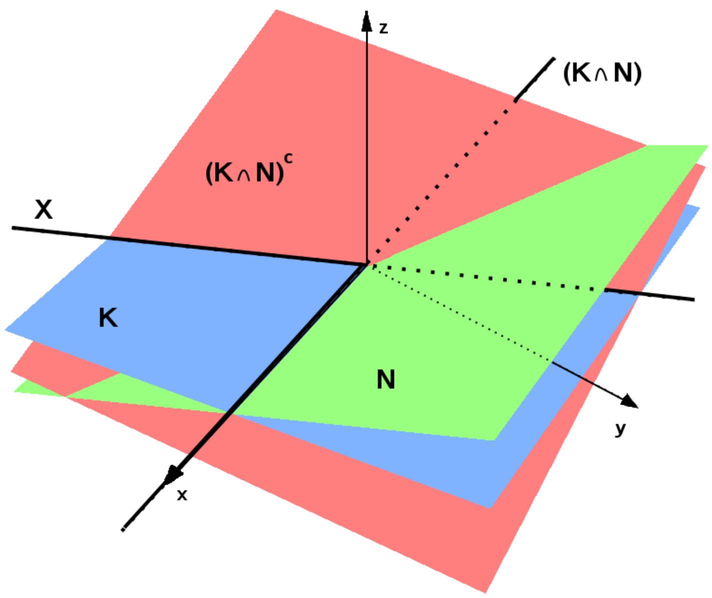

Figure 1: Decomposition of $\mathbb{R}^3$

complement $(K + N)^c$, *which shows that fixing* X *does not completely fix the projector. It is worth mentioning that restricting the choice always to orthogonal complements we* *arrive at the so-called widely orthogonal projectors, and those are uniquely determined. This case corresponds here to the choice $\alpha = \frac{\pi}{2}$ and $\beta = 1$.*

Now we start to discuss several methods of constructing projectors Q.

3.1 Basis representation methods

If a basis $n_1, \ldots, n_{m-r}$ of N and a basis $\chi_1, \ldots, \chi_\sigma$ of a suitable X are available, $X \cap N = 0$, we immediately form a projector Q onto N satisfying $X \subseteq \ker Q$ we are looking for as (cf. Lemma A.6 in [LMT11])

$$Q = H \begin{bmatrix} I & \\ & 0 \end{bmatrix} H^-,$$

whereas $\mathcal{N} = \begin{bmatrix} n_1 \ldots n_{m-r} \end{bmatrix}$, $\mathcal{X} = \begin{bmatrix} \chi_1 \ldots \chi_\sigma \end{bmatrix}$ and $H := \begin{bmatrix} \mathcal{N}, \mathcal{X} \end{bmatrix}$ have full column rank, and H^- is any reflexive generalized inverse of H. Consider different ways to generate suitable basises, and at the same time, a suitable subspace X.

A basis of N is delivered by decomposition (3). We have to provide a basis of a suitable subspace X. Recall that for any matrix $\mathfrak{A}$ the relation $\ker \mathfrak{A}^*\mathfrak{A} = \ker \mathfrak{A}$ is true. Therefore, because of

$$\widehat{N} = N \cap K = \ker \begin{bmatrix} G \\ \Pi \end{bmatrix} = \ker (G^*G + \Pi^*\Pi),$$

by means of a decomposition of $\begin{bmatrix} G \\ \Pi \end{bmatrix} \in \mathbb{R}^{k+m,m}$ or of $(G^*G + \Pi^*\Pi) \in \mathbb{R}^{m,m}$ we can design a projector Z onto $\widehat{N}$. The choice of this projector fixes also a possible complement $\widehat{N}^{\,c} := \operatorname{im} Z$ of $\widehat{N}$. By means of Z we compute a basis of X by one of the relations

$$\ker \begin{bmatrix} Z \\ \Pi \end{bmatrix} = \ker (Z^*Z + \Pi^*\Pi) = (N \cap K)^c \cap K = X.$$

This method to provide the projector Q needs three decompositions including those of matrices with $k+m$ resp. $2m$ rows as well as the computation of expressions like $(G^*G+\Pi^*\Pi)$.

An alternative possibly cheaper way to construct an admissible projector Q is suggested by Lemma A.4 in [LMT11].

$$G = U_G \begin{bmatrix} S_G & \\ & 0 \end{bmatrix} V_G^{-1}, \quad V_G =: [V_{G,1}, V_{G,2}] \tag{18}$$

we obtain $N = \ker G = \operatorname{im} V_{G,2}$, that is, a basis of N. Then, in order to apply Lemma A.4 in [LMT11], we decompose

$$\Pi V_{G,2} = U_{\Pi N} \begin{bmatrix} S_{\Pi N} & \\ & 0 \end{bmatrix} V_{\Pi N}^{-1}, \quad V_{\Pi N} =: [V_{\Pi N,1}, V_{\Pi N,2}],$$

and hence $\ker \Pi V_{G,2} = \operatorname{im} V_{\Pi N,2}$ is valid. Then, owing to Lemma A.4 in [LMT11], $Y := V_{G,2} V_{\Pi N,2} \in \mathbb{R}^{m \times q}$ represents a basis of $\ker G \cap \ker \Pi = N \cap K$. Having the basis Y of

$N \cap K$ we could, as before, compute a projector Z onto $N \cap K$, and put $(N \cap K)^c = \ker Z$, but here we actually do not compute Z, but provide a basis of the nullspace of Z in a different way. We decompose

$$Y = U_Y \begin{bmatrix} S_Y \\ 0 \end{bmatrix}, \quad U_Y =: [U_{Y,1}, U_{Y,2}],$$

with nonsingular U_Y, S_Y. Now, $U_{Y,2} \in \mathbb{R}^{m \times (m-q)}$ serves as a basis of a complement $(N \cap K)^c = \ker Z$, that means $\ker Z = \operatorname{im} U_{Y,2}$. To apply Lemma A.4 in [LMT11] once more we compute a basis of $\ker \varPi U_{Y,2}$ by the further decomposition

$$\varPi U_{Y,2} = U_X \begin{bmatrix} S_X & \\ & 0 \end{bmatrix} V_X^{-1}, \quad V_X =: [V_{X,1}, V_{X,2}],$$

yielding $\ker \varPi U_{Y,2} = \operatorname{im} V_{X,2}$. This finally leads to

$$X = (N \cap K)^c \cap K = \ker Z \cap \ker \varPi = \operatorname{im} U_{Y,2} V_{X,2}.$$

Here, four lower-dimensional matrix decompositions are needed to compute the admissible projector Q.

3.2 Basis representation methods - Regular case

In the regular case, if

$$K \cap N = \{0\}, \tag{19}$$

equation (17) simplifies to

$$Q(I - \varPi) = 0. \tag{20}$$

Condition (19) implies $m - \rho \leq r$. On the background of the decomposition (18) of G, each projector onto N has the form

$$Q = V_G \begin{bmatrix} 0 & 0 \\ -M_1 S_G & I_{m-r} \end{bmatrix} V_G^{-1}. \tag{21}$$

A basis of $\operatorname{im}(I - \varPi) = \ker \varPi$ can be computed by means of the decomposition

$$\varPi = U_\varPi \begin{bmatrix} S_\varPi & \\ & 0 \end{bmatrix} V_\varPi^{-1}, \quad S_\varPi \in \mathbb{R}^{\rho \times \rho} \text{ nonsingular}, \quad V_\varPi =: [V_{\varPi,1}, V_{\varPi,2}]$$

yielding $\operatorname{im}(I - \varPi) = \operatorname{im} V_{\varPi,2}$, $\operatorname{rank} V_{\varPi,2} = m - \rho$. Now condition (20) means $QV_{\varPi,2} = 0$, or $V_{\varPi,2} = PV_{\varPi,2}$, with $P := I - Q$. This leads to

$$V_{\varPi,2} = PV_{\varPi,2} = V_G \begin{bmatrix} I & 0 \\ M_1 S_G & 0 \end{bmatrix} \underbrace{V_G^{-1} V_{\varPi,2}}_{=: \begin{bmatrix} \mathcal{V}_1 \\ \mathcal{V}_2 \end{bmatrix}} = V_G \begin{bmatrix} I & 0 \\ M_1 S_G & I \end{bmatrix} \begin{bmatrix} \mathcal{V}_1 \\ 0 \end{bmatrix}, \tag{22}$$

which shows that $\operatorname{rank} \mathcal{V}_1 = \operatorname{rank} V_{\Pi,2} = m-\rho$, i.e. $\mathcal{V}_1 \in \mathbb{R}^{r,m-\rho}$ has full column rank. The requirement $QV_{\Pi,2} = 0$ results in the condition $-M_1 S_G \mathcal{V}_1 + \mathcal{V}_2 = 0$, which determines M_1. The choice

$$M_1 = \mathcal{V}_2 \mathcal{V}_1^- S_G^{-1} \tag{23}$$

satisfies this relation with an arbitrary generalized reflexive inverse $\mathcal{V}_1^-$, since $\mathcal{V}_1^- \mathcal{V}_1 = I$. If Π is symmetric, V_Π is orthogonal and $\mathcal{V}_1^+$ is the Moore-Penrose inverse, then the choice

$$M_1 = \mathcal{V}_2 \mathcal{V}_1^+ S_G^{-1} \tag{24}$$

generates the widely orthogonal projector Q, which is shown at the end of Subsection 3.3.

3.3 Projector representation method

Now we build the projector Q without using subspace basises. We apply again the decomposition (18) and the general projector representation (21), that is

$$Q = V_G \begin{bmatrix} 0 & 0 \\ -M_1 S_G & I_{m-r} \end{bmatrix} V_G^{-1}.$$

Introducing $\tilde{\Pi} := V_G^{-1} \Pi V_G$ we derive the expression

$$\begin{bmatrix} V_G^{-1} & \\ & V_G^{-1} \end{bmatrix} \begin{bmatrix} \Pi \\ I-Q \end{bmatrix} V_G = \begin{bmatrix} V_G^{-1}\Pi V_G \\ I - V_G^{-1} Q V_G \end{bmatrix} = \begin{array}{c} \begin{bmatrix} \tilde{\Pi}_{11} & \tilde{\Pi}_{12} \\ \tilde{\Pi}_{21} & \tilde{\Pi}_{22} \\ I_r & 0 \\ M_1 S_G & 0 \end{bmatrix} \end{array} \begin{array}{l} \}\, r \\ \\ \\ \\ \end{array} . \tag{25}$$

From $\ker \begin{bmatrix} \Pi \\ I-Q \end{bmatrix} = K \cap N = \widehat{N}$ and $u = \dim \widehat{N}$ it follows that $\dim \ker \begin{bmatrix} \Pi \\ I-Q \end{bmatrix} = m-u$. Regarding this we conclude from (25) that the rank condition $\operatorname{rank} \begin{bmatrix} \tilde{\Pi}_{12} \\ \tilde{\Pi}_{22} \end{bmatrix} = m-u-r$ is valid.

Lemma 3.2 *Given a projector Π and the decomposition* (3) *of a matrix G,* $\operatorname{rank} G = r$. *Then the projector Q defined by* (21) *meets the properties* (16) *and* (17)*, supposed one of the following three conditions is satisfied*

(1) $M_1 = -\begin{bmatrix} \tilde{\Pi}_{12} \\ \tilde{\Pi}_{22} \end{bmatrix}^- \begin{bmatrix} \tilde{\Pi}_{11} \\ \tilde{\Pi}_{21} \end{bmatrix} S_G^{-1}$.

(2) $M_1 = -\tilde{\Pi}_{22}^- \tilde{\Pi}_{21} S_G^{-1}$, *and the reflexive generalized inverse* $\tilde{\Pi}_{22}^-$ *satisfies* $\tilde{\Pi}_{12} = \tilde{\Pi}_{12}\tilde{\Pi}_{22}^-\tilde{\Pi}_{22}$.

(3) $\Pi = \Pi^*$, V_G *is orthogonal, and* $M_1 = -\tilde{\Pi}_{22}^- \tilde{\Pi}_{21} S_G^{-1}$.

Moreover, the special choice of the Moore-Penrose inverse in case (3)*, thus*

$$M_1 = -\tilde{\Pi}_{22}^+ \tilde{\Pi}_{21} S_G^{-1},$$

provides a symmetric ΠQ and a widely orthogonal Q .

Proof: (1) Condition (16) is always given by the construction and it remains to verify (17). We let $M := M_1 S_G = -\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^- \begin{bmatrix}\tilde{\Pi}_{11}\\ \tilde{\Pi}_{21}\end{bmatrix}$ and compute

$$\begin{aligned}
V_G^{-1}\Pi Q(I-\Pi)V_G &= \begin{bmatrix}\tilde{\Pi}_{11} & \tilde{\Pi}_{12}\\ \tilde{\Pi}_{21} & \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}0 & 0\\ -M & I\end{bmatrix}\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\\ -\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}\\
&= \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-M & I\end{bmatrix}\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\\ -\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}\\
&= \begin{bmatrix}-\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}M & \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\end{bmatrix}\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\\ -\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}.
\end{aligned}$$

The relation $0 = \tilde{\Pi}(I-\tilde{\Pi})$ provides

$$\begin{bmatrix}\tilde{\Pi}_{11}\\ \tilde{\Pi}_{21}\end{bmatrix}\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\end{bmatrix} = -\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix},$$

and hence

$$M\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\end{bmatrix} = \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^-\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}$$

Regarding this we finally find

$$\begin{aligned}
V_G^{-1}\Pi Q(I-\Pi)V_G &= \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}M\begin{bmatrix}I-\tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\end{bmatrix} + \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}\\
&= \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^-\left(-\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}\right) + \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}\begin{bmatrix}-\tilde{\Pi}_{21} & I-\tilde{\Pi}_{22}\end{bmatrix}\\
&= 0.
\end{aligned}$$

(2) If we are aware of a $(m-r)\times(m-r)$ submatrix of $\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix} \in \mathbb{R}^{m\times(m-r)}$, which has rank $m-r-u$, a generalized reflexive inverse of $\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}$ can be computed i.e. by a Householder decomposition. We assume without loss of generality that rank $\tilde{\Pi}_{22} = m-r-u$. If the submatrix is distributed over the rows of the matrix, a row permutation leads to the same assumption but at the end the factor U contains row permutations.

Decompose $\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix} = \left[\begin{array}{c|c} Z_{12} & 0\\ \hline Z_{22,1} & 0\\ Z_{22,2} & 0\end{array}\right]U$ with nonsingular $Z_{22,1}\in\mathbb{R}^{m-r-u,m-r-u}$ and orthogonal U, and fix the reflexive generalized inverse $\tilde{\Pi}_{22}^- = U^*\begin{bmatrix}Z_{22,1}^{-1} & 0\\ 0 & 0\end{bmatrix}$. Below we show that a generalized reflexive inverse is given by

$$\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^- := \begin{bmatrix}0 & \tilde{\Pi}_{22}^-\end{bmatrix}. \tag{26}$$

Applying (1) and this special structure of the inverse provides

$$M_1 = -\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^- \begin{bmatrix}\tilde{\Pi}_{11}\\ \tilde{\Pi}_{21}\end{bmatrix} S_G^{-1} = \begin{bmatrix}0 & -\tilde{\Pi}_{22}^-\end{bmatrix} \begin{bmatrix}\tilde{\Pi}_{11}\\ \tilde{\Pi}_{21}\end{bmatrix} S_G^{-1} = -\tilde{\Pi}_{22}^-\tilde{\Pi}_{21}S_G^{-1},$$

which verifies the assertion. It remains to verify that (26) in fact serves as reflexive generalized inverse. The condition

$$\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix} \begin{bmatrix}0 & -\tilde{\Pi}_{22}^-\end{bmatrix} \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix} = \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}$$

is valid because of our assumption concerning the generalized inverse $\tilde{\Pi}_{22}^-$, namely

$$\tilde{\Pi}_{12} = \tilde{\Pi}_{12}\tilde{\Pi}_{22}^-\tilde{\Pi}_{22} \text{ or equivalently } \operatorname{im}(I - \tilde{\Pi}_{22}^-\tilde{\Pi}_{22}) = \ker\tilde{\Pi}_{22} \subseteq \ker\tilde{\Pi}_{12}. \tag{27}$$

(3) The symmetry of Π and $\tilde{\Pi}$ (with orthogonal V_G) yields

$$\tilde{\Pi}_{22} = \begin{bmatrix}\tilde{\Pi}_{21} & \tilde{\Pi}_{22}\end{bmatrix} \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix} = \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}^* \begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}$$

and therefore $\operatorname{rank}\tilde{\Pi}_{22} = m - r - u$ and $\ker\tilde{\Pi}_{22} = \ker\begin{bmatrix}\tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\end{bmatrix}$, i.e. $\ker\tilde{\Pi}_{22} \subseteq \ker\tilde{\Pi}_{12}$. Considering (27), assertion (3) is shown to be a consequence of (2).

Finally we have to verify that taking the Moore-Penrose inverse in case (3) one delivers a widely orthogonal projector Q. By Definition 2.5 in [LMT11] a widely orthogonal projector Q projects onto N along $(K+N)^\perp \oplus X$ with $X = \widehat{N^\perp} \cap K$. Lemma A.6 (7) in [LMT11] describes sufficient conditions. Put $M = \tilde{\Pi}_{22}^+\tilde{\Pi}_{21}$ and derive

$$\Pi Q = V_G\tilde{\Pi}V_G^*V_G\begin{bmatrix}0 & 0\\ -M & I\end{bmatrix}V_G^* = V_G\begin{bmatrix}\tilde{\Pi}_{12}\tilde{\Pi}_{22}^+\tilde{\Pi}_{21} & \tilde{\Pi}_{12}\\ \tilde{\Pi}_{22}\tilde{\Pi}_{22}^+\tilde{\Pi}_{21} & \tilde{\Pi}_{22}\end{bmatrix}V_G^*.$$

The symmetry of $\tilde{\Pi}$ implies the symmetry of $\tilde{\Pi}_{11}, \tilde{\Pi}_{22}$ and $\tilde{\Pi}_{12} = \tilde{\Pi}_{21}^*$. The Moore-Penrose inverse of a symmetric matrix is symmetric itself, therefore $\tilde{\Pi}_{22}\tilde{\Pi}_{22}^+ = \tilde{\Pi}_{22}^+\tilde{\Pi}_{22}$. We consider the matrix blocks of ΠQ which seemingly derange the symmetry,

$$(\tilde{\Pi}_{12}\tilde{\Pi}_{22}^+\tilde{\Pi}_{21})^* = \tilde{\Pi}_{21}^*(\tilde{\Pi}_{22}^+)^*\tilde{\Pi}_{12}^* = \tilde{\Pi}_{12}\tilde{\Pi}_{22}^+\tilde{\Pi}_{21} \text{ and}$$
$$\tilde{\Pi}_{22}\tilde{\Pi}_{22}^+\tilde{\Pi}_{21} = (\tilde{\Pi}_{22}\tilde{\Pi}_{22}^+)^*\tilde{\Pi}_{12}^* = (\tilde{\Pi}_{12}\tilde{\Pi}_{22}\tilde{\Pi}_{22}^+)^* = (\tilde{\Pi}_{12}\tilde{\Pi}_{22}^+\tilde{\Pi}_{22})^* \overset{(27)}{=} \tilde{\Pi}_{12}^* = \tilde{\Pi}_{21},$$

but this shows the symmetry of ΠQ and naturally ΠP with $P = I - Q$.
The last properties we have to show are the symmetry of $P(I - \Pi)$ and the condition $Q\Pi P = 0$ as well. Derive

$$P(I-\Pi) = (I-Q)(I-\Pi) = V_G\begin{bmatrix}I & 0\\ M & 0\end{bmatrix}V_G^*V_G(I-\tilde{\Pi})V_G^* = V_G\begin{bmatrix}I - \tilde{\Pi}_{11} & -\tilde{\Pi}_{12}\\ M(I - \tilde{\Pi}_{11}) & -M\tilde{\Pi}_{12}\end{bmatrix}V_G^*,$$

further

$$M(I - \tilde{\Pi}_{11}) = -\tilde{\Pi}_{22}^+\underbrace{\tilde{\Pi}_{21}(I - \tilde{\Pi}_{11})}_{\tilde{\Pi}_{22}\tilde{\Pi}_{21}} = -\tilde{\Pi}_{21} = -\tilde{\Pi}_{12}^*$$
$$M\tilde{\Pi}_{12} = -\tilde{\Pi}_{22}^+\tilde{\Pi}_{21}\tilde{\Pi}_{12} = -\tilde{\Pi}_{22}^+\tilde{\Pi}_{22}(I - \tilde{\Pi}_{22})$$

which shows the symmetry of $P(I-\Pi)$. Next we compute

$$\begin{aligned} Q\Pi P &= V_G \begin{bmatrix} 0 & 0 \\ -M & I \end{bmatrix} V_G^* V_G \begin{bmatrix} \tilde{\Pi}_{11} & \tilde{\Pi}_{12} \\ \tilde{\Pi}_{21} & \tilde{\Pi}_{22} \end{bmatrix} V_G^* V_G \begin{bmatrix} I & 0 \\ M & 0 \end{bmatrix} V_G^* \\ &= V_G \begin{bmatrix} 0 & 0 \\ -M\tilde{\Pi}_{11} + \tilde{\Pi}_{21} + (-M\tilde{\Pi}_{12} + \tilde{\Pi}_{22})M & 0 \end{bmatrix} V_G^*, \end{aligned}$$

and

$$-M\tilde{\Pi}_{11} + \tilde{\Pi}_{21} + (-M\tilde{\Pi}_{12} + \underbrace{\tilde{\Pi}_{22})M}_{-\tilde{\Pi}_{21}} = \tilde{\Pi}_{22}^{+}(-\underbrace{\tilde{\Pi}_{21}\tilde{\Pi}_{11}}_{(I-\tilde{\Pi}_{22})\tilde{\Pi}_{21}} + \underbrace{\tilde{\Pi}_{21}\tilde{\Pi}_{12}}_{\tilde{\Pi}_{22}(I-\tilde{\Pi}_{22})} \tilde{\Pi}_{22}^{+}\tilde{\Pi}_{21}) = 0,$$

and hence $Q\Pi P = 0$. Now the assertion follows from Lemma A.6 (7) in [LMT11]. □

We are especially interested in the regular case, where $\widehat{N} = \{0\}$. Lemma 3.2 suggests a practical way to compute a widely orthogonal projector for that case. Since $\widehat{N} = \{0\}$ and $u = \dim \widehat{N} = 0$ the matrix $\begin{bmatrix} \tilde{\Pi}_{12} \\ \tilde{\Pi}_{22} \end{bmatrix}$ in (25) has full column rank. Then $\tilde{\Pi}_{22} = \begin{bmatrix} \tilde{\Pi}_{12} \\ \tilde{\Pi}_{22} \end{bmatrix}^* \begin{bmatrix} \tilde{\Pi}_{12} \\ \tilde{\Pi}_{22} \end{bmatrix}$ is nonsingular and $\tilde{\Pi}_{22}^{+} = \tilde{\Pi}_{22}^{-1}$. Moreover, $\tilde{\Pi}_{22}$ is not only nonsingular but positive definite as the following lemma proves.

Lemma 3.3 *Let the symmetric projector* $\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix}$ *have a nonsingular block* Π_{22}. *Then this block* Π_{22} *is positive definite.*

Proof: Π is a projector and therefore $\Pi_{22} = \Pi_{21}\Pi_{12} + \Pi_{22}^2$. It holds $\Pi_{21} = \Pi_{12}^*$ and $\Pi_{22} = \Pi_{22}^*$. We consider $\langle \Pi_{22}x, x\rangle$ for $x \neq 0$.

$$\begin{aligned} \langle \Pi_{22}x, x\rangle &= \langle (\Pi_{21}\Pi_{12} + \Pi_{22}^2)x, x\rangle \\ &= \langle \Pi_{21}\Pi_{12}x, x\rangle + \langle \Pi_{22}^2 x, x\rangle \\ &= \langle \Pi_{12}x, \Pi_{12}x\rangle + \langle \Pi_{22}x, \Pi_{22}x\rangle \\ &\geq \langle \Pi_{22}x, \Pi_{22}x\rangle > 0. \end{aligned}$$

□

Lemma 3.3 suggests to decompose $\tilde{\Pi}_{22}$ by Cholesky decomposition when computing

$$M_1 = -\tilde{\Pi}_{22}^{-1}\tilde{\Pi}_{21}S_G^{-1} \tag{28}$$

for widely orthogonal projectors.

Next we show that, in the regular case formula (24) provides exactly the same M_1 as formula (28), and hence an additional way to compute the widely orthogonal projectors. The projector Π is symmetric and has the decomposition

$$\Pi = V_\Pi \begin{bmatrix} I & \\ & 0 \end{bmatrix} V_\Pi^T =: \begin{bmatrix} V_{\Pi,1} & V_{\Pi,2} \end{bmatrix} \begin{bmatrix} I & \\ & 0 \end{bmatrix} \begin{bmatrix} V_{\Pi,1}^T \\ V_{\Pi,2}^T \end{bmatrix}$$

with an orthogonal matrix V_Π. We obtain $\Pi = I - V_{\Pi,2}V_{\Pi,2}^T$, which leads to

$$\begin{aligned}\tilde{\Pi} = V_G^{-1}\Pi V_G &= I - V_G^{-1}V_{\Pi,2}V_{\Pi,2}^T V_G \quad \text{(cf. (22))}\\ &= I - \begin{bmatrix}\mathcal{V}_1\\ \mathcal{V}_2\end{bmatrix}\begin{bmatrix}\mathcal{V}_1\\ \mathcal{V}_2\end{bmatrix}^T\\ &= \begin{bmatrix} I - \mathcal{V}_1\mathcal{V}_1^T & -\mathcal{V}_1\mathcal{V}_2^T\\ -\mathcal{V}_2\mathcal{V}_1^T & I - \mathcal{V}_2\mathcal{V}_2^T\end{bmatrix}.\end{aligned}$$

Applying (28) we obtain

$$M_1 = \underbrace{(I - \mathcal{V}_2\mathcal{V}_2^T)^{-1}\mathcal{V}_2}_{=\mathcal{V}_2(\mathcal{V}_1^T\mathcal{V}_1)^{-1}}\mathcal{V}_1^T S_G^{-1} = \mathcal{V}_2\underbrace{(\mathcal{V}_1^T\mathcal{V}_1)^{-1}\mathcal{V}_1^T}_{=\mathcal{V}_1^+}S_G^{-1}, \tag{29}$$

which coincides with (23).

4 Matrix function sequences

4.1 Stepping level by level

The admissible sequences of matrix functions are constructed pointwise. We start with matrices (standing for matrix functions with frozen arguments) A, D and B. We compute a generalized inverse of $G_0 := AD$ and fix in that way a projector $Q_0 := I - P_0 = I - G_0^- G_0$ onto $\ker G_0$. The starting matrices of the sequence are $G_0, G_0^-, B_0 := B, \Pi_0 := P_0$.
Let us assume that we have determined the sequence up to level i, that means, G_i. admissible projectors $Q_j, j = 1, \ldots, i$, and the projectors $\Pi_j = P_0 \ldots P_j$ are already computed. Since they are admissible, the condition (cf. (17))

$$\Pi_{j-1}Q_j(I - \Pi_{j-1}) = 0$$

holds for every level $j = 1, \ldots, i$. We have to build $G_{i+1} = G_i + B_iQ_i$ and an nullspace projector Q_{i+1} onto $\ker G_{i+1}$ satisfying

$$X_{i+1} = (N_0 + \cdots + N_i) \ominus \widehat{N_i} \subset \ker Q_{i+1}, \tag{30}$$

or equivalently,

$$\Pi_i Q_{i+1}(I - \Pi_i) = 0. \tag{31}$$

The decomposition

$$G_{i+1} = \mathcal{U}_{i+1}\begin{bmatrix} S_{i+1} & \\ & 0\end{bmatrix}\mathcal{V}_{i+1}^{-1} \tag{32}$$

provides the reflexive generalized inverse

$$G_{i+1}^- = \mathcal{V}_{i+1}\begin{bmatrix} S_{i+1}^{-1} & M_{2,i+1}\\ M_{1,i+1} & M_{1,i+1}S_{i+1}M_{2,i+1}\end{bmatrix}\mathcal{U}_{i+1}^{-1}$$

and the nullspace projector

$$Q_{i+1} = \mathcal{V}_{i+1}\begin{bmatrix} 0 & 0\\ -M_{1,i+1}S_{i+1} & I\end{bmatrix}\mathcal{V}_{i+1}^{-1}.$$

The entry $M_{1,i+1}$ can be computed by means of one of the proposals in Section 3 and $M_{2,i+1}$ can be set to zero.
Since we proceed pointwise with frozen arguments, to ensure continuity of the nullspace projector and then that of the next matrix function, it is recommended to apply widely orthogonal projectors. For widely orthogonal projectors we need an orthogonal matrix $\mathcal{V}_{i+1}$ (see Lemma 3.2 (3)), which requires a decomposition of G_{i+1} by an SVD or by the Householder method (decomposition of G^*_{i+1}). After having generated G_{i+1} and the nullspace projector Q_{i+1} we have to provide also the next

$$B_{i+1} = B_i P_i - G_{i+1} D^-(D\Pi_{i+1} D^-)' D\Pi_i$$

or, in the invariant case,

$$B_{i+1} = B_i P_i.$$

The latter case does not present any difficulty, however, in general the involved derivative of $D\Pi_{i+1}D^-$ represents a serious challenge. In [Lam05] finite differences are used to approximate this derivative, which delivers quite accurate results in lower index cases and if the relevant subspaces are invariant. A more accurate approximation of the derivative by automatic differentiation (cf. [GW08]) is done in [LM10]. The application of automatic differentiation needs highter smoothness assumptions as needed for the tractability index concept itself.

4.2 Involved version for the regular case

A complete new decomposition of G_{i+1} at each level appears to be expensive. In the regular case, a possibility to make better use of results obtained in previous steps is developed in [Lam05]. We use the representation

$$G_{i+1} = G_i + B_i Q_i = (G_i + W_i B_0 Q_i) F_{i+1}$$

with the projector W_i along $\operatorname{im} G_i$ and the nonsingular matrix $F_{i+1} = I + G_i^- B_i Q_i$. For the matrix $G_j,\ j = 0, \dots, i$ we have already the decomposition

$$G_j = \mathcal{U}_j \begin{bmatrix} S_j & \\ & 0 \end{bmatrix} \mathcal{V}_j^{-1}$$

with $\mathcal{U}_j$, S_j and $\mathcal{V}_j$ nonsingular matrices. The other components are given for $j = 0, \dots, i$ by

$$G_j^- = \mathcal{V}_j \begin{bmatrix} S_j^{-1} & M_{2,j} \\ M_{1,j} & M_{1,j} S_j M_{2,j} \end{bmatrix} \mathcal{U}_j^{-1},$$

$$W_j = I - G_j G_j^- = \mathcal{U}_j \begin{bmatrix} 0 & -S_j M_{2,j} \\ & I \end{bmatrix} \mathcal{U}_j^{-1} = \mathcal{U}_j T_{u,j}^{-1} \begin{bmatrix} 0 & \\ & I \end{bmatrix} \mathcal{U}_j^{-1}, \tag{33}$$

$$Q_j = I - G_j^- G_j = \mathcal{V}_j \begin{bmatrix} 0 & \\ -M_{1,j} S_j & I \end{bmatrix} \mathcal{V}_j^{-1} = \mathcal{V}_j \begin{bmatrix} 0 & \\ & I \end{bmatrix} T_{l,j}^{-1} \mathcal{V}_j^{-1} \tag{34}$$

with the upper and lower triangular matrices

$$T_{u,j} := \begin{bmatrix} I & S_j M_{2,j} \\ & I \end{bmatrix} \text{ and } T_{l,j} := \begin{bmatrix} I & \\ M_{1,j} S_j & I \end{bmatrix}.$$

Using the detailed structure of the various matrices we find

$$G_{i+1} = \mathcal{U}_i T_{u,i}^{-1} \left(\begin{bmatrix} S_i & \\ & 0 \end{bmatrix} + \begin{bmatrix} 0 & \\ & I \end{bmatrix} \underbrace{\mathcal{U}_i^{-1} B_0 \mathcal{V}_i}_{\bar{B}_i} \begin{bmatrix} 0 & \\ & I \end{bmatrix} \right) T_{l,i}^{-1} \mathcal{V}_i^{-1} F_{i+1}.$$

If we write $\bar{B}_i = \begin{bmatrix} B_{11}^i & B_{12}^i \\ B_{21}^i & B_{22}^i \end{bmatrix}$ and decompose $B_{22}^i = \tilde{U}_{i+1} \begin{bmatrix} \tilde{S}_{i+1} & \\ & 0 \end{bmatrix} \tilde{V}_{i+1}^{-1}$, we can use this decomposition and obtain

$$G_{i+1} = \underbrace{\mathcal{U}_i T_{u,i}^{-1} \begin{bmatrix} I & \\ & \tilde{U}_{i+1} \end{bmatrix}}_{=:\mathcal{U}_{i+1}} \begin{bmatrix} S_i & & \\ & \tilde{S}_{i+1} & \\ & & 0 \end{bmatrix} \underbrace{\begin{bmatrix} I & \\ & \tilde{V}_{i+1}^{-1} \end{bmatrix} T_{l,i}^{-1} \mathcal{V}_i^{-1} F_{i+1}}_{=:\mathcal{V}_{i+1}^{-1}} . \tag{35}$$

Defining $S_{i+1} := \begin{bmatrix} S_i & \\ & \tilde{S}_{i+1} \end{bmatrix}$ we now have the required decomposition of

$$G_{i+1} = \mathcal{U}_{i+1} \begin{bmatrix} S_{i+1} & \\ & 0 \end{bmatrix} \mathcal{V}_{i+1}^{-1} \tag{36}$$

and

$$G_{i+1}^- = \mathcal{V}_{i+1} \begin{bmatrix} S_{i+1}^{-1} & M_{2,j} \\ M_{1,i+1} & M_{1,i+1} S_{i+1} M_{2,i+1} \end{bmatrix} \mathcal{U}_{i+1}^{-1}.$$

The projector

$$Q_{i+1} = I - G_{i+1}^- G_{i+1} = \mathcal{V}_{i+1} \begin{bmatrix} 0 & 0 \\ -M_{1,i+1} S_{i+1} & I \end{bmatrix} \mathcal{V}_{i+1}^{-1}.$$

is a nullspace projector for each $M_{1,i+1}$.
To fix the projector, the different entries $M_{1,i+1}$ can be determined as described in Section 3, where Π is replaced by Π_i. The computation of M_1 by (23) goes better with the step-by-step computation. Widely orthogonal projectors are computed using the Moore-Penrose inverse of $\mathcal{V}_1$ (see (29)).
The advantage of the involved step-by-step computation of the sequence is that, at each step, we decompose only the matrix $\bar{B}_{22}^i$, whose dimension reduces from step to step.
After having computed G_{i+1} and Q_{i+1} we have to provide

$$B_{i+1} = B_i P_i - G_{i+1} D^- (D\Pi_{i+1} D^-)' D\Pi_i.$$

Here again, the challenge is the differentiation of $D\Pi_{i+1}D^-$.

4.3 Computing characteristic values and index check

The characteristic values of the DAE under consideration, that are (see Definition 2.9 in [LMT11]) the values

$$r_i = \operatorname{rank} G_i,\ u_i = \dim \widehat{N}_i,\ \widehat{N}_i = \ker \begin{bmatrix} \Pi_{i-1} \\ I - Q_i \end{bmatrix} = \ker \begin{bmatrix} \Pi_{i-1} \\ G_i \end{bmatrix} = \ker \left[G_i^* G_i + \Pi_{i-1}^* \Pi_{i-1} \right]$$

are rank values arising as byproducts within the factorizations when generating the matrix sequences as described in the previous subsection.

If one meets a nonzero value u_i, the given DAE fails to be regular, which makes the question

$$\text{“}u_i = 0\ ?\text{”}$$

to serve as a regularity test.

The determination of the tractability index of a regular DAE requires the determination of the matrix sequence up to a nonsingular matrix G_μ. We concentrate on a regular case. At every level $G_i,\ i = 0, 1, \dots$ the characteristic value $r_i = \operatorname{rank} G_i$ is determined checking the nonsingularity of G_i. The step-by-step algorithm of Section 4.2 delivers the characteristic values successively starting from $r_0 = \operatorname{rank} G_0$ and $r_{i+1} = r_i + r_B^i,\ i = 0, 1, \dots$ with $r_B^i := \operatorname{rank} \bar{B}_{22}^i$.

The regularity is implicitly checked computing an admissible projector at every level. In the case of a critical point, we are faced with a rank drop of $\mathcal{V}_1$ if we use (23) or a singular block $\tilde{\Pi}_{22}$ if we apply (28).

The computation of $B_i,\ i > 0$ needs the differentiation of $D\Pi_i D^-$. The factorization $G_{i+1} = (G_i + W_i B_0 Q_i) \underbrace{(I + G_i^- B_i Q_i)}_{\text{nonsingular}}$ allows to determine $r_{i+1} = \operatorname{rank}(G_i + W_i B_0 Q_i)$, which is easier, since one can do without computing the derivative of $D\Pi_i D^-$. The first level where the derivative of $D\Pi_i D^-$ may influences the sequence matrix occurs for $i = 2$. The check of the index-3 property needs only one differentiation, which is accurately realizable by finite differences.

Algorithmic differentiation (AD) to compute the derivative of $D\Pi_i D^-$ is applied in [LM11]. Using an AD tool all computations are made by Taylor polynomials and a derivative is reduced to a shift of the Taylor series. The application of these technique requires higher smoothness assumptions.

For time invariant linear DAEs the tractability index coincides with the Kronecker index (cf. Theorem 6.2 in [LMT11]), i.e., the numerical determination of the characteristic values discloses the inner structure of the DAE.

References

[GK65] G. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis, 2(2):205–224, 1965.

[GW08] A. Griewank and A. Walter. Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation. SIAM, second edition, 2008.

[Lam05] R. Lamour. Index determination and calculation of consistent initial values for daes. Comput. Math. Appl., 50:1125–1140, 2005.

[LM10] R. Lamour and D. Monett Diaz. A new algorithm for the index determination in DAEs by Taylor series using Algorithmic Differentiation. Technical Report 10-04, Humboldt-University, Department of Mathematics, 2010.

[LM11] R. Lamour and D. Monett Diaz. A new algorithm for the index determination in DAEs by Taylor series using Algorithmic Differentiation. Numer. Algorithms, 2011. accepted for publication.

[LMT11] R. Lamour, R. März, and C. Tischendorf. Projector based treatment of linear constant coefficient DAEs. Technical Report 11-15, Humboldt-University, Department of Mathematics, 2011.