Splitting of Quasi-Definite Linear System maintains Inertia

Martin Neuenhofen

November 15, 2018

Abstract

We show that there is a variety of Schur complements that yield a decoupling of a quasi-definite linear system into two quasi-definite linear systems of half the size each. Splitting of linear systems of equations via Schur complements is widely used to reduce the size of a linear system of equations. Quasi-definite linear systems arise in a variety of computational engineering applications.

1 Theorem

Definitions Given $n, m \in \mathbb{N}$, $N = 2 \cdot n$, $M = 2 \cdot m$, $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{B} \in \mathbb{R}^{M \times N}$, $\mathbf{D} \in \mathbb{R}^{M \times M}$, and

$$\hat{\mathbf{K}} := \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & -\mathbf{D} \end{bmatrix},$$

where $\mathbf{A}$ and $\mathbf{D}$ are symmetric positive definite. $\hat{\mathbf{K}}$ is called *quasi-definite* [**?**].

We consider a splitting of $\mathbf{A}, \mathbf{B}, \mathbf{D}$ in blocks of size $n \times n$, $m \times n$, and $m \times m$, respectively:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2^\top \\ \mathbf{A}_2 & \mathbf{A}_4 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{B}_4 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{D}_2^\top \\ \mathbf{D}_2 & \mathbf{D}_4 \end{bmatrix}.$$

Using these blocks, we can reorder $\hat{\mathbf{K}}$ as

$$\mathbf{K} := \left[\begin{array}{cc|cc} \mathbf{A}_1 & \mathbf{B}_1^\top & \mathbf{A}_2^\top & \mathbf{B}_3^\top \\ \mathbf{B}_1 & -\mathbf{D}_1 & \mathbf{B}_2 & -\mathbf{D}_2^\top \\ \hline \mathbf{A}_2 & \mathbf{B}_2^\top & \mathbf{A}_4 & \mathbf{B}_4^\top \\ \mathbf{B}_3 & -\mathbf{D}_2 & \mathbf{B}_4^\top & -\mathbf{D}_4 \end{array}\right] =: \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2^\top \\ \mathbf{K}_2 & \mathbf{K}_4 \end{bmatrix},$$

where $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_4 \in \mathbb{R}^{(n+m)\times(n+m)}$ are defined as indicated by the 2×2 block structure. We then define two matrices:

$$\mathbf{U} := \mathbf{K}_1 - \mathbf{K}_2^\top \cdot \mathbf{K}_4^{-1} \cdot \mathbf{K}_2 \tag{1a}$$

$$\mathbf{V} := \mathbf{K}_4 - \mathbf{K}_2 \cdot \mathbf{K}_1^{-1} \cdot \mathbf{K}_2^\top \tag{1b}$$

Notice that the matrices $\mathbf{U}, \mathbf{V}$ are well-defined since the inverses exist due to quasi-definiteness of $\mathbf{K}_1, \mathbf{K}_4$.

Result The matrices $\mathbf{U}, \mathbf{V}$ have the following 2×2 block-structure

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2^\mathsf{T} \\ \mathbf{U}_2 & -\mathbf{U}_4 \end{bmatrix}, \qquad \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2^\mathsf{T} \\ \mathbf{V}_2 & -\mathbf{V}_4 \end{bmatrix},$$

where $\mathbf{U}_1, \mathbf{V}_1 \in \mathbb{R}^{n \times n}$ and $\mathbf{U}_4, \mathbf{V}_4 \in \mathbb{R}^{m \times m}$ are symmetric positive definite.

Proof All we have to show is that $\mathbf{U}_1$ is symmetric positive definite. The rest then follows by analogy.

Simply by insertion of the blocks into the definition of $\mathbf{U}$, we find the following formula for $\mathbf{U}_1$:

$$\begin{aligned} \mathbf{U}_1 = \quad & (\mathbf{A}_1 + \mathbf{B}_3^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3) \\ & - (\mathbf{A}_2 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3)^\mathsf{T} \cdot (\mathbf{A}_4 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_4)^{-1} \\ & \quad \cdot (\mathbf{A}_2 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3) \end{aligned} \tag{2}$$

We notice that

$$\mathbf{A}_1 + \mathbf{B}_3^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3$$

is symmetric positive definite because $\mathbf{A}_1$ and $\mathbf{D}_4$ are symmetric positive definite.

Hence, $\mathbf{U}_1$ is positive definite if and only if

$$\mathbf{X} := \begin{bmatrix} \mathbf{A}_1 + \mathbf{B}_3^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3 & (\mathbf{A}_2 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3)^\mathsf{T} \\ \mathbf{A}_2 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_3 & \mathbf{A}_4 + \mathbf{B}_4^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \mathbf{B}_4 \end{bmatrix}$$

is positive definite. That is indeed the case, as we now show:

$$\begin{aligned} & \mathbf{x}^\mathsf{T} \cdot \mathbf{X} \cdot \mathbf{x} \\ = \; & \mathbf{x}^\mathsf{T} \cdot \mathbf{A} \cdot \mathbf{x} + \left(\begin{bmatrix} \mathbf{B}_3 \\ \mathbf{B}_4 \end{bmatrix} \cdot \mathbf{x} \right)^\mathsf{T} \cdot \mathbf{D}_4^{-1} \cdot \left(\begin{bmatrix} \mathbf{B}_3 \\ \mathbf{B}_4 \end{bmatrix} \cdot \mathbf{x} \right) > 0 \qquad \forall \mathbf{x} \in \mathbb{R}^{n+m}, \mathbf{x} \neq \mathbf{0} \end{aligned}$$

because $\mathbf{A}$ is positive definite and $\mathbf{D}$ is positive definite, implying that $\mathbf{D}_4$ is positive semi-definite. □

Corollary From "if and only if" follows that the result holds in a strict sense. I.e., $\mathbf{U}_1, \mathbf{V}_1, \mathbf{U}_4, \mathbf{V}_4$ are symmetric positive definite if and only if $\mathbf{A}, \mathbf{D}$ are symmetric positive definite.

Remark The result is easily generalized for complex numbers.

2 Applications

The motivation for considering the matrices $\mathbf{U}, \mathbf{V}$ is that a linear system of equations with $\mathbf{K}$ can be decoupled into two linear systems of equations with $\mathbf{U}, \mathbf{V}$. We show this.

After suitable reordering of a quasi-definite linear system, consider

$$\mathbf{K} \cdot \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} . \tag{3}$$

We find the two decoupled linear systems for $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^{n+m}$:

$$\mathbf{U} \cdot \mathbf{x}_1 = \mathbf{b}_1 - \mathbf{K}_2^\top \cdot \mathbf{K}_4^{-1} \cdot \mathbf{b}_2 \tag{4a}$$

$$\mathbf{V} \cdot \mathbf{x}_2 = \mathbf{b}_2 - \mathbf{K}_2 \cdot \mathbf{K}_1^{-1} \cdot \mathbf{b}_1 \tag{4b}$$

Depending on the structure of $\hat{\mathbf{K}}$, in some cases it is attractive for computational cost to build and solve the decoupled systems; e.g. when $\hat{\mathbf{K}}$ is banded [?]. In doing so, one can exploit that the matrices $\mathbf{U}, \mathbf{V}$ have quasi-definite structure. Also, the result can be applied recursively, as $\mathbf{U}, \mathbf{V}$ can be considered in turn as matrices of format $\hat{\mathbf{K}}$. This recursive applicability can also be used to check the inertia, according to the corollary: The linear systems with $\mathbf{U}, \mathbf{V}$ are split recursively, until eventually the resulting split systems are so small that one can compute the inertia in a direct way to determine if $\mathbf{A}, \mathbf{D}$ were positive definite.

Application to Dense Block-Tridiagonal Linear Systems When $\hat{\mathbf{K}}$ can be reordered into a (dense) block-tridiagonal form, then $\mathbf{U}, \mathbf{V}$ in turn can be reordered into a dense[1] block-tridiagonal form. This is particularly attractive for solving banded linear systems with a symmetric cyclic reduction algorithm [?] and parallel computations. As we show in [?], formulas for the split linear systems use the matrices $\mathbf{U}, \mathbf{V}$. Hence, the theorem holds recursively for all the linear systems generated within the cyclic reduction algorithm, providing following utilities:

- If $\hat{\mathbf{K}}$ is quasi-definite then by induction so will the matrices $\mathbf{U}, \mathbf{V}$ in each iteration of the cyclic reduction.
- Following the induction, the inverses in the cyclic reduction algorithm do all exist, since $\mathbf{K}_1, \mathbf{K}_4$ are quasi-definite (after suitable reordering).
- The inertia of $\mathbf{U}, \mathbf{V}$ after the last iteration of cyclic reduction can be computed to safeguard to check positive definiteness of $\mathbf{A}, \mathbf{D}$.

The last item is particularly useful when the system matrix $\hat{\mathbf{K}}$ arises within an optimization algorithm. For instance, for SQP methods, which typically compute a step-direction by solving a <u>convex</u> quadratic program [?], the matrix $\mathbf{A}$ must be ensured to be positive definite, and must be otherwise manipulated (e.g. by applying a shift to $\mathbf{A}$).

[1] There is fill-in on the band due to the inverses in $\mathbf{U}, \mathbf{V}$.

3 Conclusion and Outlook

We presented a quite general result for quasi-definite matrices. The result has a wide applicability for structured linear systems, cyclic reduction algorithms, and in the field of convex and non-convex programming.

References