Block Matrix Solutions

Aaron M. Johnson

June 21, 2024

In analyzing the continuous and impulsive dynamics of a robot that is contacting the world, a common structure arises. These dynamics are characterized by a set of dynamic differential equations involving the mass matrix M and constraint forces/impulses, described by the matrix A^T, as well as an algebraic constraint equation that depends on A. Solving this combined DAE system (differential-algebraic equations) can be done in many different ways, but it is helpful to combine the M and A terms into an invertible block matrix. This can lead to better numerical conditioning and also enable solutions for systems with massless limbs.

1 Block Matrix Inverse

Consider an invertible matrix that is decomposed into several sub-blocks, here labeled E, F, G, and H. If the upper left block E is invertible then the inverse of that matrix can be written as,

$$\boxed{\begin{bmatrix} E & F \\ G & H \end{bmatrix}^{-1} = \begin{bmatrix} E^{-1} + E^{-1}FS_E^{-1}GE^{-1} & -E^{-1}FS_E^{-1} \\ -S_E^{-1}GE^{-1} & S_E^{-1} \end{bmatrix}} \qquad \text{Block matrix inverse} \tag{1}$$

where

$$\boxed{S_E := H - GE^{-1}F} \qquad \text{Schur complement} \tag{2}$$

is called the *Schur complement* of the block E. This formula can be readily validated by testing that this expression is a left and right inverse of the original matrix (see the Practice Problems).

In dynamics we will be using the special case of this general formula, where $E = M$, $F = A^T$, $G = A$, and $H = 0$ (sometimes called the Lagrangian matrix of coefficients),

$$S_M = 0 - AM^{-1}A^T \tag{3}$$

$$\begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} M^{-1} - M^{-1}A^T(AM^{-1}A^T)^{-1}AM^{-1} & M^{-1}A^T(AM^{-1}A^T)^{-1} \\ (AM^{-1}A^T)^{-1}AM^{-1} & -(AM^{-1}A^T)^{-1} \end{bmatrix} \tag{4}$$

2 Dagger Terms

The decomposition of the block matrix inverse in (1) is useful if the first block is invertible. But, we can still think about the block components of the matrix inverse (when it exists) even without this definition,

$$\boxed{\begin{bmatrix} M^\dagger & A^{\dagger T} \\ A^\dagger & \Lambda^\dagger \end{bmatrix} := \begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix}^{-1}} \qquad \text{Dagger terms} \tag{5}$$

If the size of our state is n and the number of constraints is m, such that $M \in \mathbb{R}^{n\times n}$ and $A \in \mathbb{R}^{m\times n}$, and the block of zeros will be an $m \times m$ matrix. The dagger terms have the same sizes, so $M^\dagger \in \mathbb{R}^{n\times n}$, $A^\dagger \in \mathbb{R}^{m\times n}$, and $\Lambda^\dagger \in \mathbb{R}^{m\times m}$. Note that we have not yet defined a non-dagger Λ – for consistency we will define Λ to mean the Schur complement $\Lambda = S_M = -AM^{-1}A^T$ when it exists, such that $\Lambda^\dagger = \Lambda^{-1} = S_M^{-1}$. However, note that other texts may use Λ to mean the Delassus operator $AM^{-1}A^T$ (which is $-S_M$), the contact space or apparent inertia matrix $(AM^{-1}A^T)^{-1}$ (which is $-\Lambda^\dagger$ or $-S_M^{-1}$) [1,2], or instead of $\Lambda^\dagger$ itself [3].

These "dagger terms" will show up in our continuous dynamics and impact laws. **They exist whenever the whole block matrix is invertible, even if M is not invertible**. But, if the leading block (M) is also invertible, they line up with the definition in (4),

$$M^\dagger = M^{-1} - M^{-1}A^T(AM^{-1}A^T)^{-1}AM^{-1} \tag{6}$$

$$A^\dagger = (AM^{-1}A^T)^{-1}AM^{-1} \tag{7}$$

$$\Lambda^\dagger = -(AM^{-1}A^T)^{-1} \tag{8}$$

From the definition of matrix inverse, multiplying the left and right block matrices in (5) together in either order, we can observe the following properties and identities (where $I_{n\times n}$ is an $n\times n$ identity matrix and $0_{n\times m}$ is an $n\times m$ zero matrix):

$$\begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} M^\dagger & A^{\dagger T} \\ A^\dagger & \Lambda^\dagger \end{bmatrix} = \begin{bmatrix} M^\dagger & A^{\dagger T} \\ A^\dagger & \Lambda^\dagger \end{bmatrix}\begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix} = \begin{bmatrix} I_{n\times n} & 0_{n\times m} \\ 0_{m\times n} & I_{m\times m} \end{bmatrix} \tag{9}$$

$$MM^\dagger + A^TA^\dagger = I_{n\times n} \qquad M^\dagger M + A^{\dagger T}A = I_{n\times n} \tag{10}$$

$$MA^{\dagger T} + A^T\Lambda^\dagger = 0_{n\times m} \qquad A^\dagger M + \Lambda^\dagger A = 0_{m\times n} \tag{11}$$

$$AM^\dagger = 0_{m\times n} \qquad M^\dagger A^T = 0_{n\times m} \tag{12}$$

$$AA^{\dagger T} = I_{m\times m} \qquad A^\dagger A^T = I_{m\times m} \tag{13}$$

so **in particular $A^\dagger$ is a left-inverse of A^T, but note that $M^\dagger$ is not a left- or right-inverse of M.**

Note also that if M is symmetric positive (semi-)definite, then $M^\dagger$ is symmetric positive semi-definite and $\Lambda^\dagger$ is symmetric negative (semi-)definite. We can see from (12) that the rank of $M^\dagger$ can be no more than $n-m$ (since A must have full rank of m), and so is always singular for a constrained system. The rank of $M^\dagger$ must be at least $n-m$ if the block matrix is invertible as $A^\dagger$ is rank m, thus the rank of $M^\dagger$ must be exactly $n-m$.

Furthermore, if M is singular (i.e. only semi-definite), then $\Lambda^\dagger$ is as well. If there is only a single constraint, $\Lambda^\dagger = \begin{bmatrix}0\end{bmatrix}$. In fact, M and $\Lambda^\dagger$ have the same nullity, i.e. if M has rank $n-d$ then $\Lambda^\dagger$ has rank $m-d$. To see this, by (11),

$$\forall \dot{q}_i \in \mathbb{R}^n : M\dot{q}_i = 0_{n\times 1} \Rightarrow A^\dagger M\dot{q}_i + \Lambda^\dagger A\dot{q}_i = 0_{m\times n}\dot{q}_i \tag{14}$$

$$0_{m\times 1} + \Lambda^\dagger(A\dot{q}_i) = 0_{m\times 1} \tag{15}$$

Thus $A\dot{q}_i$ is in the null space of $\Lambda^\dagger$, as $A\dot{q}_i \neq 0$ if the block matrix is full rank. In summary,

$$M^{\dagger T} = M^\dagger \qquad M^\dagger \geq 0 \qquad \mathrm{null}(M^\dagger) = m \tag{16}$$

$$\Lambda^{\dagger T} = \Lambda^\dagger \qquad \Lambda^\dagger \leq 0 \qquad \mathrm{null}(\Lambda^\dagger) = \mathrm{null}(M) \tag{17}$$

3 Discussion

$$\underset{(\ddot{q})}{\mathcal{T}^2\mathcal{Q}} \underset{M^\dagger}{\overset{M}{\rightleftarrows}} \underset{(\Gamma)}{\mathcal{T}^*\mathcal{Q}} \underset{A^T}{\overset{A^\dagger}{\rightleftarrows}} \underset{(\lambda)}{\mathcal{T}^*\mathcal{C}} \underset{\Lambda^\dagger}{\overset{\Lambda}{\rightleftarrows}} \underset{(\ddot{c})}{\mathcal{T}^2\mathcal{C}}$$

$$\underset{(\dot{q})}{\mathcal{T}\mathcal{Q}} \underset{A^{\dagger T}}{\overset{A}{\rightleftarrows}} \underset{(\dot{c})}{\mathcal{T}\mathcal{C}}$$

Figure 1: Functional mapping of the dagger terms. Note that this diagram does not necessarily commute, as the dagger elements are not always inverses of their non-dagger versions.

We can think of these various matrices as functions between the state and constraint space velocities, accelerations, and forces. If the state space is $q \in \mathcal{Q}$, then the space of state velocities is $\dot{q} \in \mathcal{T}\mathcal{Q}$ (the tangent space), the space of accelerations is $\ddot{q} \in \mathcal{T}^2\mathcal{Q}$, and the space of forces on the state space is $\Gamma \in \mathcal{T}^*\mathcal{Q}$ (the cotangent space). With this notation, one way to think about the mass matrix is $M : \mathcal{T}^2\mathcal{Q} \to \mathcal{T}^*\mathcal{Q}$, as it maps accelerations to forces ($M\ddot{q}$ equals the applied forces). We also have the space of constraints $c \in \mathcal{C}$, where we have mostly thus far talked about the constraint velocities $\dot{c} \in \mathcal{T}\mathcal{C}$, i.e. we have said that the velocity in constrained directions must be zero. Thus the velocity constraint $A : \mathcal{T}\mathcal{Q} \to \mathcal{T}\mathcal{C}$, as it maps the state velocity to the constraint velocity (which we want to be zero, $A\dot{q} = 0$). It is always true that the transpose of a linear function of tangent spaces maps back from the corresponding cotangent spaces – in our case $A^T : \mathcal{T}^*\mathcal{C} \to \mathcal{T}^*\mathcal{Q}$, as $A^T\lambda = \Gamma$ is the effect of the constraint forces on the state space.

With this notation, we can now think of the dagger terms as mapping $M^\dagger : \mathcal{T}^*\mathcal{Q} \to \mathcal{T}^2\mathcal{Q}$, $A^\dagger : \mathcal{T}^*\mathcal{Q} \to \mathcal{T}^*\mathcal{C}$, $A^{\dagger T} : \mathcal{T}\mathcal{C} \to \mathcal{T}\mathcal{Q}$, and $\Lambda^\dagger : \mathcal{T}^2\mathcal{C} \to \mathcal{T}^*\mathcal{C}$. However, remember that **these are not necessarily inverse functions of the non-dagger versions** even if they map back to the same spaces. For example, we have in (13) that $A^{\dagger T}$ is only a right-inverse of A. So if we have a constraint velocity $\dot{c}$ that we mapped through $A^{\dagger T}$ to get $\dot{q} = A^{\dagger T}\dot{c}$, and then mapped it back through A we would get that $A\dot{q} = AA^{\dagger T}\dot{c} = \dot{c}$. However, it is not a left-inverse, so if we started with a $\dot{q}$, mapped it through A to get $\dot{c} = A\dot{q}$, and the back we would have $A^{\dagger T}A\dot{q} \neq \dot{q}$ (in particular, any constrained motion of the system will have $A\dot{q} = 0$, which will map to $\dot{c} = 0$ and therefore $A^{\dagger T}A\dot{q} = 0 \neq \dot{q}$). This is easy to see when we look at the dimensions of the spaces and matrices – since $\dot{q} \in \mathbb{R}^n$ is larger than $\dot{c} \in \mathbb{R}^m$, we will inevitably lose information when we go down to the smaller space first.

Since $A^{\dagger T}$ is a right-inverse of A, but not the only possible right-inverse of A, what value of $\dot{q}$ do we get when we map $A^{\dagger T}\dot{c} = \dot{q}$? Consider a case where we have a velocity that violates the constraint, so $A\dot{q} = \dot{c} \neq 0$. Knowing how much we are violating the constraint, $\dot{c}$, what change to $\dot{q}$ should we make? Clearly there are many different changes to $\dot{q}$ we could make that would result in the same change in constraint velocity $\dot{c}$. Which of these should we choose? A natural choice (and one that agrees with physics) is that we should choose the one that requires the least energy. Our kinetic energy, $\frac{1}{2}\dot{q}^T M\dot{q}$, defines a metric or norm, which we can write as $\|\dot{q}\|_M$. Thus we would like to find the change to $\dot{q}$ that achieves the desired change in constraint velocity $\dot{c}$ and minimizes the norm $\|\dot{q}\|_M$. This least-norm solution to a linear system can be solved by the **mass-weighted right Moore–Penrose pseudoinverse** $M^{-1}A^T(AM^{-1}A^T)^{-1}$, which is exactly the quantity $A^{\dagger T}$. If we take $\dot{q} = A^{\dagger T}\dot{c}$ as the solution, we can verify that $A\dot{q} = AA^{\dagger T}\dot{c} = \dot{c}$ as desired, by (13). Similarly, $A^\dagger$ is the mass-weighted left pseudoinverse of A^T.

If $M^\dagger$ is not an inverse of M, then what does it represent? Starting from (10), we see that $M^\dagger M = I - A^{\dagger T}A$. We can think of $M^\dagger$ as the **inverse of M in the constrained space**, since if we had a $\dot{q}$ that was consistent with the constraints (i.e. $A\dot{q} = 0$), then $M^\dagger M\dot{q} = I\dot{q} - A^{\dagger T}A\dot{q} = \dot{q}$. Another interpretation is that it is the inverse of M that removes the constrained component (hence the null space (16)), since if we apply A to the output we get $AM^\dagger M = AI - AA^{\dagger T}A = A - A = 0$ and so passing a velocity through $M^\dagger M$ removes any component that is not consistent with the constraints. This will be useful when defining impact laws, as we will need to remove the component of the velocity into the constraint.

4 Block Matrix Dynamics

Starting from the constrained equations of motion,

$$M\ddot{q} + C\dot{q} + N + A^T\lambda = \Upsilon \tag{18}$$

$$A\ddot{q} + \dot{A}\dot{q} = 0 \tag{19}$$

we can take this set of equations and factor out the block matrix from (5),

$$\begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \ddot{q} \\ \lambda \end{bmatrix} = \begin{bmatrix} \Upsilon - C\dot{q} - N \\ -\dot{A}\dot{q} \end{bmatrix} \tag{20}$$

$$\begin{bmatrix} \ddot{q} \\ \lambda \end{bmatrix} = \begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} \Upsilon - C\dot{q} - N \\ -\dot{A}\dot{q} \end{bmatrix} = \begin{bmatrix} M^\dagger & A^{\dagger T} \\ A^\dagger & \Lambda^\dagger \end{bmatrix} \begin{bmatrix} \Upsilon - C\dot{q} - N \\ -\dot{A}\dot{q} \end{bmatrix} \tag{21}$$

Thus our acceleration dynamics can be solved for as,

$$\boxed{\begin{aligned} \ddot{q} &= M^\dagger(\Upsilon - C\dot{q} - N) - A^{\dagger T}\dot{A}\dot{q} \\ \lambda &= A^\dagger(\Upsilon - C\dot{q} - N) - \Lambda^\dagger\, \dot{A}\dot{q} \end{aligned}} \quad \text{Dagger solution to constrained dynamics} \tag{22}$$

Note that if M is invertible, this solution to the dynamics is equivalent to the traditional solution of (e.g. [4, Eqn 6.5, 6.6]),

$$\ddot{q} = M^{-1}(\Upsilon - C\dot{q} - N - A^T\lambda) \tag{23}$$

$$\lambda = (AM^{-1}A^T)^{-1}\left(AM^{-1}(\Upsilon - C\dot{q} - N) + \dot{A}\dot{q}\right) \tag{24}$$

as we can see by using (6)–(8),

$$\ddot{q} = M^{-1}\left(\Upsilon - C\dot{q} - N - A^T(AM^{-1}A^T)^{-1}\left(AM^{-1}(\Upsilon - C\dot{q} - N) + \dot{A}\dot{q}\right)\right) \tag{25}$$

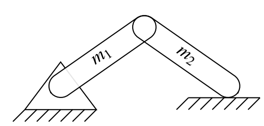
$$= \left(M^{-1} - M^{-1}A^T(AM^{-1}A^T)^{-1}AM^{-1}\right)(\Upsilon - C\dot{q} - N) - \left(M^{-1}A^T(AM^{-1}A^T)^{-1}\right)\dot{A}\dot{q} \tag{26}$$

$$= M^\dagger(\Upsilon - C\dot{q} - N) - A^{\dagger T}\dot{A}\dot{q} \tag{27}$$

$$\lambda = \left((AM^{-1}A^T)^{-1}AM^{-1}\right)(\Upsilon - C\dot{q} - N) + \left((AM^{-1}A^T)^{-1}\right)\dot{A}\dot{q} \tag{28}$$

$$= A^\dagger(\Upsilon - C\dot{q} - N) - \Lambda^\dagger\dot{A}\dot{q} \tag{29}$$

5 Massless Dynamics



Consider the mass matrix for a 2-link manipulator, where each link i is length l_i, mass m_i, and inertia I_i,

$$M = \begin{bmatrix} m_1\frac{l_1^2}{4} + m_2 l_1^2 + m_2\frac{l_2^2}{4} + m_2 l_1 l_2 \cos(\theta_2) + I_1 + I_2 & m_2\frac{l_2^2}{4} + m_2 l_1 \frac{l_2}{2}\cos(\theta_2) + I_2 \\ m_2\frac{l_2^2}{4} + m_2 l_1\frac{l_2}{2}\cos(\theta_2) + I_2 & m_2\frac{l_2^2}{4} + I_2 \end{bmatrix} \tag{30}$$

If $m_2 \ll m_1$, it would be nice to simplify this expression by saying $m_2 \approx 0$ and therefore $I_2 \approx 0$,

$$M \approx \begin{bmatrix} m_1\frac{l_1^2}{4} + I_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{31}$$

The challenge is obviously that this mass matrix is singular and so we cannot use M^{-1} in any of our calculations, e.g. (23). And indeed this would be a problem if the system were unconstrained, as $M\ddot{q}$ would zero out all terms related to $\ddot{\theta}_2$. However, if the system is constrained, e.g. as in the "crank-slider" configuration shown, then the system can be solved for just in terms of $\ddot{\theta}_1$ and then we can calculate $\ddot{\theta}_2$ from there.

In general, note that as long as the block matrix in (5) is invertible, we can use the dagger terms to solve our dynamics as in (22). When is this true? We need a condition on rank, essentially that any rank deficiencies of the inertia tensor M must be "corrected" by velocity constraints in A such that any motion still excites some momentum (i.e. pushes against some mass) [3, Assumption A5]:

$$\boxed{\text{The block matrix in (5) is invertible if } M\dot{q} \neq 0_{n\times 1} \text{ for all } \dot{q} \neq 0_{n\times 1} \text{ such that } A\dot{q} = 0_{m\times 1}.} \tag{32}$$

Note that this requirement is equivalent to requiring that the mass matrix in reduced coordinates, $\tilde{M}$, be invertible [3, Lemma 4] (i.e. in this example, changing coordinates to only consider θ_1).

Looking at the example in (31), we see that a velocity of just the second link, $\dot{q} = \begin{bmatrix} 0 & \dot{\theta}_2 \end{bmatrix}^T$, would result in no momentum, however this would not satisfy the constraint. Assume for simplicity that $l_1 = l_2$, then the constraint can be simplified to $A = \begin{bmatrix} 2 & -1 \end{bmatrix}$. Thus, any valid $\dot{q}$ would have some component of motion in both joints and therefore move the mass of the first link. Looking at the block matrix inverse, and setting $m_1 l_1^2/4 + I_1 = 1$ for simplicity,

$$\begin{bmatrix} M & A^T \\ A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{33}$$

$$M^\dagger = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \qquad A^\dagger = \begin{bmatrix} 0 & -1 \end{bmatrix}, \qquad \Lambda^\dagger = \begin{bmatrix} 0 \end{bmatrix} \tag{34}$$

We see that the block matrix is now full rank (all rows and columns are linearly independent), as the constraint A has "covered up" the zeros from M. Also, note here that $M^\dagger$ has rank 1, as per (16), and $\Lambda^\dagger = \begin{bmatrix} 0 \end{bmatrix}$, as per (17)).

What if we do not have enough constraints? For example, if the two link robot considered here were not touching any surfaces. In these cases, we must remove the singular part of the system (here, the second link) and consider it using separate, decoupled dynamics (e.g. holding constant position) [3, Assumption A6].

Finally, note that even in cases where part of the system is not not truly massless but simply small, and so M is not singular but close to singular, computing the dynamics with the block matrix solution may provide better numerical conditioning [5, Sec. 5.1.1].

6 Practice Problems

1) Validate the expression for the block matrix inverse in (1).

2) Prove that $\Lambda^\dagger = -A^\dagger M A^{\dagger T}$.

3) Considering the matrices we discussed in this chapter as functions, which of them are one-to-one (injective), which are onto (surjective), and which are both (bijective)? Check $M, A, A^T, M^\dagger, A^\dagger, A^{\dagger T}, \Lambda^\dagger$. Assume for this problem that M is full rank, and that the system is not fully constrained (i.e. the number of constraints m is less than the dimensionality of the state space n).

References

[1] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987.

[2] K. M. Lynch and F. C. Park, *Modern robotics*. Cambridge University Press, 2017.

[3] A. M. Johnson, S. E. Burden, and D. E. Koditschek, "A hybrid systems model for simple manipulation and self-manipulation systems," *International Journal of Robotics Research*, vol. 35, no. 11, pp. 1354–1392, September 2016.

[4] R. M. Murray, Z. Li, S. S. Sastry, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 1994.

[5] A. M. Johnson, "Self-manipulation and dynamic transitions for a legged robot," Ph.D. dissertation, Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA, 2014.