EE/AA547-W22 2022.02.02

Topic: Observability/Controllability

Lecturer: L.J. Ratliff

1 Overview

- Controllability/Reachability is the property of a system concerning the ability to steer the state from arbitrary x_0 to arbitrary x_1 on a given time interval $[t_0, t_1]$.
- Observability is the property of a system concerning the ability to uniquely recover the initial state x_0 given the observation y.

1.1 Controllability/Reachability

For simplicity, let $D(\cdot) \equiv 0$ and consider our state space to be $\mathbb{R}^n$.

Definition 1 (Controllable)**.** The system $\mathcal{D} = (A(\cdot), B(\cdot), C(\cdot))$ is controllable on $[t_0, t_1]$ if for all $(x_0, x_1) \in \mathbb{R}^n$, there exists $u_{[t_0,t_1]} \in \mathcal{U}$ which steers x_0 at t_0 to x_1 at t_1.

We tend to break controllability into two different concepts:

- Controllability from the origin (reachability): the reachability map is given by

$$\mathcal{L}_r u(t) = \int_{t_0}^{t} \Phi(t,\tau)B(\tau)u(\tau)\ d\tau$$

  If this map is surjective (i.e., $\mathrm{Im}(\mathcal{L}_r) = \mathbb{R}^n$) then the system is reachable.
- Controllability to the origin (often simply referred to as contorllability, hence, one should take care to identify the precise definition in the given reference they are looking at): the controllability map is given by

$$\mathcal{L}_c u(t) = \int_{t_0}^{t} \Phi(t_0,\tau)B(\tau)u(\tau)\ d\tau$$

  If this map is surjective (i.e., $\mathrm{Im}(\mathcal{L}_c) = \mathbb{R}^n$) then the system is controllable.

We can compute the adjoint of these two maps $\mathcal{L}_r^*$ and $\mathcal{L}_c^*$ respectively, and define the corresponding grammians

$$\mathcal{L}_r\mathcal{L}_r^* = \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)B^*(\tau)\Phi(t_1,\tau)^*\ d\tau$$

and

$$\mathcal{L}_c\mathcal{L}_c^* = \int_{t_0}^{t_1} \Phi(t_0,\tau)B(\tau)B^*(\tau)\Phi(t_0,\tau)^*\ d\tau$$

These are both linear operators and hence by the matrix representation theorem there is a finite dimension matrix representation $W_r \in \mathbb{R}^{n\times n}$ and $W_c \in \mathbb{R}^{n\times n}$, respectively. The finite rank operator lemma let's us show that

$$\mathrm{Im}(\mathcal{L}_r) = \mathrm{Im}(\mathcal{L}_r\mathcal{L}_r^*) = \mathrm{Im}(W_r)$$

and

$$\mathrm{Im}(\mathcal{L}_c) = \mathrm{Im}(\mathcal{L}_c\mathcal{L}_c^*) = \mathrm{Im}(W_c)$$

Moreover, we have the following equivalences:

$$\begin{aligned}(A(\cdot),B(\cdot))\ \text{ controllable on } [t_0,t_1] &\iff \operatorname{Im}(\mathcal{L}_r)=\mathbb{R}^n\\ &\iff \operatorname{Im}(\mathcal{L}_c)=\mathbb{R}^n\\ &\iff \operatorname{Im}(\mathcal{L}_r\mathcal{L}_r^*)=\mathbb{R}^n\\ &\iff \operatorname{Im}(\mathcal{L}_c\mathcal{L}_c^*)=\mathbb{R}^n\\ &\iff \det(W_r)\neq 0\\ &\iff \det(W_c)\neq 0\end{aligned}$$

LTI Systems. Things get much easier in the LTI setting. Cayley Hamilton allows us to show that the condition on the controllability (reachability) grammian can be essentially reduced to checking a rank condition on the so called controllability matrix:

$$\mathcal{C}=\begin{bmatrix}B & AB & A^2B & \cdots & A^{n-1}B\end{bmatrix}\in\mathbb{R}^{n\times nm}$$

Fact 2. The following equality holds:

$$\operatorname{Im}(W_r)=\operatorname{Im}(\mathcal{C})$$

Let $\Delta=t_1-t_0$ for some $t_1>t_0$.

Theorem 3. The following are equivalent:

$$\begin{aligned}&\text{The LTI system is completely controllable on some } [0,\Delta]\\ \iff\ & \operatorname{rank}\left(\begin{bmatrix}B & AB & \cdots & A^{n-1}B\end{bmatrix}\right)=n && \text{(Rank Test)}\\ \iff\ & \operatorname{rank}\left(\begin{bmatrix}sI-A & B\end{bmatrix}\right)=n,\quad \forall\ s\in\mathbb{C} && \text{(PBH Test)}\end{aligned}$$

1.2 Observability

The pair $(A(\cdot),C(\cdot))$ is observable if given output $y(t)$, the initial state x_0 can be uniquely recovered.

Definition 4. The state x_0 is unobservable on $[t_0,t_1]$ if and only if its zero input response is zero on $[t_0,t_1]$.

Analogous to controllability, we can define the observablility map $\mathcal{L}_\circ:\mathbb{R}^n\to\mathcal{Y}_{[t_0,t_1]}$ by

$$\mathcal{L}_\circ x_0(\cdot)=C(\cdot)\Phi(\cdot,t_0)x_0$$

That is, $\mathcal{L}_\circ x_0$ is an operator in $\mathrm{PC}([t_0,t_1]))$ such that

$$(\mathcal{L}_\circ x_0)(t)=y(t)=\int_{t_0}^{t}C(t)\Phi(t,\tau)B(\tau)u(\tau)\ d\tau$$

We have the following equivalences:

$$\begin{aligned}(A(\cdot),C(\cdot)) \text{ is completely observable (CO) on } [t_0,t_1] &\iff \operatorname{Ker}(\mathcal{L}_\circ)=\{0\}\\ &\iff \operatorname{Ker}(\mathcal{L}_\circ^*\mathcal{L}_\circ)=\{0\}\\ &\iff \det(W_\circ)\neq 0\end{aligned}$$

where

$$W_\circ=\int_{t_0}^{t_1}\Phi(\tau,t_0)^*C(\tau)^*C(\tau)\Phi(\tau,t_0)\ d\tau$$

LTI Systems. Just as with controllability, we can use Cayley Hamilton to construct an observability matrix:

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Theorem 5 (LTI Observability Tests)**.** The following are equivalent:

$$\begin{aligned} & \text{The LTI system is completely observable on some } [0, \Delta] \\ \iff & \text{rank}\left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n && \text{(Rank Test)} \\ \iff & \text{rank}\left(\begin{bmatrix} sI - A \\ C \end{bmatrix}\right) = n, \quad \forall\, s \in \mathbb{C} && \text{(PBH Test)} \end{aligned}$$

2 Problems

We will start with some warm up problems.

Problem 1. (Controllability & Observability.) Consider the linear system given by

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -3 & 3 \\ \gamma & -4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 1 \end{bmatrix} x \end{aligned}$$

for some parameter γ.

a. How should we choose γ such that the system is controllable but not observable?
b. How should we choose γ such that the system is observable but not controllable?

solution.

a. Choosing $\gamma = -1$, the system is controllable but not observable since

$$\text{rank}\left(\begin{bmatrix} B & AB \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & -3 \\ 0 & \gamma \end{bmatrix}\right) = 2$$

and

$$\text{rank}\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 1 \\ -3+\gamma & -4 \end{bmatrix}\right) < 2$$

b. Choosing $\gamma = 0$, the system is observable but not controllable:

$$\text{rank}\left(\begin{bmatrix} B & AB \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & -3 \\ 0 & \gamma \end{bmatrix}\right) = 1 < 2$$

and

$$\text{rank}\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 1 \\ -3+\gamma & -4 \end{bmatrix}\right) = 2$$

Problem 2. (Controllability & Observability.) Consider the linear system given by

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix} x$$

Suppose you have the ability to add one sensor and one actuator.

a. Which state should we control with the actuator to make the system controllable?
b. Which state should we measure with the sensor to make the system observable?

solution.

a. The controllability matrix is such that

$$\text{rank}\left(\begin{bmatrix} B & AB \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} b_1 & b_1 \\ b_2 & -b_1 - 2b_2 \end{bmatrix}\right)$$

Hence if we choose to control state 1 but not state 2, we have

$$\text{rank}\left(\begin{bmatrix} B & AB \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\right) = 2$$

On the other hand if we control state 2 but not 1, $\mathcal{C}$ drops rank.

b. The observability matrix is such that

$$\text{rank}\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} c_1 & c_2 \\ c_1 - c_2 & -2c_2 \end{bmatrix}\right)$$

Hence, if we observe state 1 and not state 2 we have

$$\text{rank}\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right) < 2$$

Yet if we observe state 2 and not state 1, we have that

$$\text{rank}\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}\right) = 2$$

Problem 3. (Reachability.) Consider a unit point mass under control of force—i.e.,

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

This is equivalent to $\ddot{x} = u$. Suppose we want to reach $(x(T), \dot{x}(T)) = (1, 0)$ from $(x(0), \dot{x}(0)) = (0, 0)$ using a controller of the form

$$u(t) = \begin{cases} u_0, & 0 \le t < T/10 \\ u_1, & T/10 \le t < 2T/10 \\ \vdots & \vdots \\ u_{10}, & 9T/10 \le t \le T \end{cases}$$

Find $u = [u_1 \ u_2 \ \cdots \ u_{10}]^\top$.

Recall that the reachability map is

$$x(T) = \int_0^T \Phi(T, \tau) B(\tau) u(\tau) \, d\tau$$

Hence, we have that

$$x(T) = \underbrace{\begin{bmatrix} L_1 & L_2 & \cdots & L_{10} \end{bmatrix}}_{\mathcal{L}_r} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{10} \end{bmatrix},$$

where

$$\begin{aligned} L_1 &= \left(\int_0^{T/10} \Phi(T,\tau)B(\tau)\ d\tau \right) = \frac{T}{10}\begin{bmatrix} T \\ 1 \end{bmatrix} \\ L_2 &= \left(\int_{T/10}^{2T/10} \Phi(T,\tau)B(\tau)\ d\tau \right) = \frac{T}{10}\begin{bmatrix} 0.9T \\ 1 \end{bmatrix} \\ &\vdots \\ L_{10} &= \frac{T}{10}\begin{bmatrix} 0.1T \\ 1 \end{bmatrix} \end{aligned}$$

Now to determine if we can drive the state to the desired position, we need to check if $\mathcal{L}_r$ is onto (surjective) which means we need to check the rank of the resulting matrix to see if its equal to two. Indeed, we have

$$\mathcal{L}_r = \frac{T}{10}\begin{bmatrix} T & 0.9T & \cdots & 0.1T \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

Then for example if $T = 10$, we have

$$\mathcal{L}_r = \begin{bmatrix} 10 & 9 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

and

$$W_r = \mathcal{L}_r\mathcal{L}_r^\top = \begin{bmatrix} 385 & 55 \\ 55 & 10 \end{bmatrix}$$

which is full rank. Thus, from the finite rank operator lemma, $\mathrm{Im}(W_r) = \mathrm{Im}(\mathcal{L}_r)$ implies that we can drive the state to the deisred location with a piecewise constant controller.

What is interesting about this example, is that if we make the number of pieces in the piecewise constant controller bigger and bigger then

$$\mathcal{L}_r\mathcal{L}_r^\top \to \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)B(\tau)^\top\Phi(t_1,\tau)^\top\ d\tau$$

Therefore for the continuous time linear system is controllable over the interval $[t_0, t_1]$ if and only if W_r is full rank.

Problem 4. (Minimum Norm Control.) As we saw in the recorded lecture (and its worth re-emphasizing) that the controllability Grammian is related to the cost of control. Given $x(t_0) = x_0$, find a control function or sequence $u(\cdot)$ so that $x(t_1) = x_1$. Let $x_d = x_1 - \Phi(t_1, t_0)x_0$. Then we must have

$$x_d = \mathcal{L}_{r,[t_0,t_1]}(u)$$

where $\mathcal{L}_r$ is the reachability map. Since we saw the continuous time version in the recorded lecture, for some variety let's focus on the discrete time case:

$$L_{r,[k_0,k_1]}(u) = \sum_{k=k_0}^{k_1-1} \Phi(k_1, k+1)B(k)u(k) = \mathcal{C}(k_0,k_1)U$$

where

$$\mathcal{C}(k_0,k_1) \in \mathbb{R}^{n\times(k_1-k_0)m}, \quad \text{and } U = \begin{bmatrix} u(k_0) \\ \vdots \\ u(k_1-1) \end{bmatrix}$$

Since $x_d \in \text{Im}(L_{r,[k_0,k_1]})$ for solutions to exist, we assume $\text{Im}(L_{r,[k_0,k_1]}) = \mathbb{R}^n$. This implies that $W_{r,[k_0,k_1]}$ is invertible. Generally there are multiple solutions. To resolve the non-uniqueness, we find the solution so that

$$J(U) = \frac{1}{2}\sum_{k=k_0}^{k_1-1} u(k)^\top u(k) = U^\top U$$

is minimized.

Solution. This is a constrained optimization problem (with $J(U)$ as the cost to be minimized, and $L_rU - x_d = 0$ as the constraint). It can be solved using the Lagrange Multiplier method of converting a constrained optimization problem into an unconstrained optimization.

Define an augmented cost function, with the Lagrange multipliers $\lambda \in \mathbb{R}^n$ as follows:

$$\tilde{J}(U, \lambda) = J(U) + \lambda^\top (L_rU - x_d)$$

The optimal solution necessarily satifies the first order conditions

$$\nabla_U \tilde{J} = 0 \quad \text{and} \quad \nabla_\lambda \tilde{J} = 0$$

and the constraint. That is,

$$L_rU^* = x_d \quad \text{and} \quad L_r^\top \lambda^* + U^* = 0$$

Solving we get that

$$\lambda^* = -(L_rL_r^\top)^{-1}x_d = -W_r^{-1}x_d \quad \text{and} \quad U^* = -L_r^\top \lambda^* = L_r^\top W_r^{-1}x_d$$

And, the optimal cost of control is

$$J(U^*) = x_d^\top W_r^{-1}L_rL_r^\top W_r^{-1}x_d = x_d^\top W_r^{-1}x_d$$

Thus, the inverse of the reachability Grammian tells us how difficult it is to perform a state transfer from $x = 0$ to x_d. In particular, if W_r is not invertible, for some x_d, the cost is infinite.

Let's look at a Geometric View Now. Geometrically, we can think of the cost as $J = U^\top U$, i.e. the inner product of U with itself. In notation of inner product, this is

$$J = \langle U, U\rangle_R$$

The advantage of the notation is that we can change the definition of inner product, e.g. $\langle U, V\rangle_R = U^\top RV$ where R is a positive definite matrix. The usual inner (dot) product has $R = I$. We say U and V are normal to each other if $\langle U, V\rangle_R = 0$.

Any solution that satisfies the constraint must be of the form

$$(U - U^p) \in \text{Ker}(L_r)$$

where U^p is any particular solution—i.e., $L_rU^p = x_d$.

Fact: Let U^* be the optimal solution, and U is any solution that satisfies the constraint. Then, $(U - U^*) \perp U^*$, i.e.

$$\langle (U - U^*), U^*\rangle_R = 0$$

which is the normal equation for the least norm solution problem.

Problem 5. (Pole Placement.) Consider

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

and the desired characteristic polynomial $p(s) = (s+1)(s+3)$. Design a feedback controller of the form $u = -kx$ to place the poles at those of $p(s)$.

solution. First,

$$\operatorname{rank}\mathcal{C} = \operatorname{rank}\left(\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\right) = 2$$

Then for $u = -kx$,

$$\begin{aligned}\det(sI - A + bk) &= \det\left(\begin{bmatrix}s & 0\\ 0 & s\end{bmatrix} - \begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix}\begin{bmatrix}k_1 & k_2\end{bmatrix}\right)\\ &= \det\left(\begin{bmatrix}s-1+k_1 & k_2\\ k_1 & s-2+k_2\end{bmatrix}\right)\\ &= (s-1+k_1)(s-2+k_2) - k_2k_1\\ &= (s-1)(s-2) + k_1(s-2) + k_2(s-1)\\ &= s^2 - 3s + 2 + k_1 s - 2k_1 + k_2 s - k_2\\ &= s^2 + (k_1 + k_2 - 3)s + 2 - 2k_1 - k_2\end{aligned}$$

So then by equating coefficients of the above and

$$p(s) = s^2 + 4s + 3$$

we get

$$\begin{aligned}4 &= k_1 + k_2 - 3 \implies 7 - k_2 = k_1\\ 3 &= 2 - 2k_1 - k_2 \implies 1 = -2k_1 - k_2 \implies 1 = -2(7 - k_2) - k_2 = -14 + k_2\end{aligned}$$

so that

$$k_1 = -8 \text{ and } k_2 = 15$$

and the closed loop system is thus

$$\dot{x} = (A - BK)x = \begin{bmatrix}9 & -15\\ 8 & -13\end{bmatrix}x$$

Problem 6. (Connections between Observability & Lyapunov.) Let (A, C) be observable and suppose that P is any solution to $A^*P + PA = -C^*C$. Note that we are not assuming that $\mathcal{L}(\cdot)$ is an invertible operator here so there may be more than one solution. Show that $P \succeq 0$ if and only if A is stable.

solution.
($\Longrightarrow$): Suppose $P \succeq 0$ but A is not stable. Then there is a non-trivial (i.e. $v \neq 0$) vector $v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq 0$ and $Av = \lambda v$. For this (λ, v) pair we also have that $v^*A^* = \bar{\lambda}v^*$. Since (A, C) is observable, we have that $Cv \neq 0$. Note that

$$-\|Cv\|^2 = -v^*C^*Cv = v^*(A^*P + PA)v = (\bar{\lambda} + \lambda)v^*Pv = 2\mathrm{Re}(\lambda)v^*Pv$$

Since $\|Cv\| > 0$ and $\mathrm{Re}(\lambda) \geq 0$, it must be the case that $\mathrm{Re}(\lambda) > 0$ and hence from above, we have $v^*Pv < 0$ meaning that P is not positive semidefinite. $\longrightarrow\!\longleftarrow$.

($\Longleftarrow$): Suppose that A is stable. We will prove this implication directly. Since A is stable there is only one solution to the equation $A^*P + PA = -C^*C$; we saw this in the examples on Lyapunov stability—namely that if A stable this is equivalent to $\bar{\lambda}_i + \lambda_j \neq 0$ for all $\lambda_i, \lambda_j \in \mathrm{spec}(A)$ which is in turn equivalent to the operator $\mathcal{L}(\cdot)$ being invertible. We also that the solution is

$$P = \int_0^\infty e^{A^*\tau}C^*Ce^{A\tau}\, d\tau$$

This is clearly a positive semidefinite matrix.

Problem 7. (Controllable Cannonical Form.) We saw in the recorded lecture and the lecture notes that if a system is completely controllable there is a transformation to the controllable canonical form. In many references this form is actually called the "controller canonical form" since its structure is invariant under the use of a feedback controller (for SISO systems this is easy to see). An alternative to this is what is called the "controllability canonical form" which is given by

$$\dot{x} = \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 & -a_{n-1} \\ 1 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 1 & & 0 & -a_{n-3} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & -a_0 \end{bmatrix}}_{A_c} x + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{B_c} u$$

$$y = \underbrace{\begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \end{bmatrix}}_{C_c} x$$

The controllability matrix for this system is the identity $\mathcal{C}_c = I$ and (A_c, B_c) is of course controllable.

Consider another system (A, B) and suppose that it is controllable. Derive a transformation such that $x_c = T^{-1}x$ is in controllable form:

$$A_c = T^{-1}AT, \quad B_c = T^{-1}B$$

solution. First off the transformation of the controllability matrix is given by

$$\mathcal{C} = T\mathcal{C}_c$$

This is because

$$\begin{aligned} \mathcal{C} &= \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \\ &= \begin{bmatrix} TB_c & TA_cT^{-1}TB_c & (TA_cT^{-1})^2TB_c & \cdots & (TA_cT^{-1})^{n-1}TB_c \end{bmatrix} \\ &= T\begin{bmatrix} B_c & A_cB_c & \cdots & A_c^{n-1}B_c \end{bmatrix} \end{aligned}$$

since $(TA_cT^{-1})^k = \underbrace{(TA_cT^{-1})\cdots(TA_cT^{-1})}_{k \text{ times}} = TA_c^kT^{-1}$. Then, since $\mathcal{C}_c = I$, we have that $T = \mathcal{C}$.

Another way to show this is to consider T of the form

$$T = \begin{bmatrix} t_1 & t_2 & \cdots & t_n \end{bmatrix}$$

Then, from $B = TB_c$ we have that $B = t_1$ (since $B_c = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top$). From $AT = TA_c$ we have that

$$\begin{aligned} \begin{bmatrix} At_1 & At_2 & \cdot & At_n \end{bmatrix} &= \begin{bmatrix} t_1 & t_2 & \cdots & t_n \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_{n-1} \\ 1 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 1 & & 0 & -a_{n-3} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & -a_0 \end{bmatrix} \\ &= \begin{bmatrix} t_2 & t_3 & \cdots & -t_1a_{n-1} - t_2a_{n-2} - \cdots - t_na_0 \end{bmatrix} \end{aligned}$$

from which we have that

$$t_1 = B,\ t_2 = At_1 = AB,\ t_3 = At_2 = A^2B,\ \ldots,\ t_n = A^{n-1}B$$

so that $T = \mathcal{C}$.