Consider the following linear time-invariant system:

\[
\frac{dx}{dt} = Ax + Bu, \quad y = Cx + Du, \quad x(0) = x_o \in R^N \text{ or } C^N, \tag{1}
\]

where the initial state \(x_o\) is to be determined by observing the output \(y(t)\) given by:

\[
y(t) = C \left( \exp(At)x_o + \int_0^t \exp(A(t-\tau))Bu(\tau)d\tau + Du(t) \right). \tag{2}
\]

Let

\[
\tilde{y}(t) \overset{\text{def}}{=} y(t) - C \left( \int_0^t \exp(A(t-\tau))Bu(\tau)d\tau - Du(t) \right) = C \exp(At)x_o.
\]

Since both the output \(y(t)\) and input \(u(t)\) are known quantities, hence \(\tilde{y}(t)\) is a known quantity.

Assume that the system is completely finite-time observable, or the observability matrix \(P = [C^T | A^T C^T | \cdots | (A^T)^{N-1} C^T]\) has rank \(N\). Then the symmetric matrix

\[
S(t) = \int_0^t \exp(A^T\tau)C^T C \exp(At)d\tau
\]

is nonsingular for any \(t > 0\). We have shown earlier (Lecture Notes 10, page 82) that the initial state \(x_o\) can be determined by

\[
x_o = S(t)^{-1} \int_0^t \exp(A^T\tau)C^T \tilde{y}(\tau)d\tau, \quad t > 0. \tag{2}
\]

The right-hand-side of (2) represents an observer or estimator for the initial state \(x_o\). To implement this observer in the form of a differential equation, we substitute \(x_o\) given by (2) into the solution of the original system equation with \(x(t)\) replaced by \(\hat{x}(t)\) representing the state of the observer:

\[
\dot{\hat{x}}(t) = \exp(At)x_o + x_u(t) = \exp(At) \left( S(t)^{-1} \int_0^t \exp(A^T\tau)C^T \tilde{y}(\tau)d\tau \right) + x_u(t),
\]

where

\[
x_u(t) = \int_0^t \exp(A(t-\tau))Bu(\tau)d\tau.
\]

Now, we differentiate the foregoing expression for \(\dot{\hat{x}}(t)\) with respect to \(t\):

\[
\frac{d\hat{x}(t)}{dt} = A \exp(At) \left( S(t)^{-1} \int_0^t \exp(A^T\tau)C^T \tilde{y}(\tau)d\tau \right)
+ \exp(At) \frac{d}{dt} S(t)^{-1} \int_0^t \exp(A^T\tau)C^T \tilde{y}(\tau)d\tau + \exp(At) S(t)^{-1} \exp(A^Tt)C^T \tilde{y}(t) + \frac{dx_u(t)}{dt}.
\]
Using the identity:
\[ \frac{d}{dt} S(t)^{-1} = -S(t)^{-1} \frac{dS(t)}{dt} S(t)^{-1}, \]
we obtain

\[ \frac{d\hat{x}(t)}{dt} = A(\hat{x}(t) - x_u(t)) - \left( \exp(At) S(t)^{-1} \frac{dS(t)}{dt} S(t)^{-1} \int_0^t \exp(A\tau) C^T \tilde{y}(\tau) d\tau \right) \]
\[ + \exp(At) S(t)^{-1} \exp(A^T t) C^T \tilde{y}(t) + \frac{dx_u(t)}{dt}. \]  

(3)

Let \[ P(t) = \exp(At) S(t)^{-1} \exp(A^T t). \]
Then the \((\cdots)\) term in (3) can be rewritten as

\[ (\cdots) = P(t) \exp(-A^T t) (dS(t)/dt) S(t)^{-1} \int_0^t \exp(A^T \tau) C^T \tilde{y}(\tau) d\tau \]
\[ = P(t) \exp(-A^T t) (dS(t)/dt) \exp(-At)(\hat{x}(t) - x_u(t)). \]

Since
\[ \frac{dS(t)}{dt} = \exp(A^T t) C^T C \exp(At), \]
or
\[ \exp(-A^T t) (dS(t)/dt) \exp(-At) = C^T C, \]
hence the \((\cdots)\) term in (3) reduces to

\[ (\cdots) = P(t) C^T C (\hat{x}(t) - x_u(t)). \]

Substituting the above expression into (3) leads to

\[ \frac{d\hat{x}(t)}{dt} = A(\hat{x}(t) - x_u(t)) - P(t) C^T C (\hat{x}(t) - x_u(t)) + P(t) C^T \tilde{y}(t) + \frac{dx_u(t)}{dt}, \]
which in view of the relations:
\[ \tilde{y}(t) = y(t) - Cx_u(t) - Du(t), \]
\[ \frac{dx_u(t)}{dt} = Ax_u(t) + Bu(t), \]
takes on the following form:

\[ \frac{d\hat{x}(t)}{dt} = (A - P(t) C^T C) \hat{x}(t) + P(t) C^T y(t) \]
\[ + (A - A + P(t) C^T C - P(t) C^T C) x_u(t) - P(t) C^T Du(t) + Bu(t) \]
\[ = A\hat{x}(t) + P(t) C^T (y(t) - C\hat{x}(t)) + (B - P(t) C^T D)u(t). \]  

(4)
The term \((y(t) - C\hat{x}(t))\) in (4) corresponds to the deviation between the model output and the actual output. The term \((B - P(t)C^TD)u(t)\) is introduced to cancel the effect of the input. The foregoing observer or state estimator can be implemented as follows:

We observe that the above observer actually generates an estimate of the instantaneous state of the actual system. Before using the above observer, it is necessary to generate the matrix \(P(t)\). We shall derive a differential equation for \(P(t)\).

Consider again

\[
P(t) = \exp(At)S(t)^{-1}\exp(A^Tt).
\]

Differentiating the above expression with respect to time \(t\) gives

\[
dP(t)/dt = A\exp(At)S(t)^{-1}\exp(A^Tt) + \exp(At)(dS(t)^{-1}/dt)\exp(A^Tt) \\
            + \exp(At)S(t)^{-1}A^T\exp(A^Tt) \\
            = AP(t) + P(t)A^T - P(t)C^TCP(t),
\]

which is a nonlinear matrix differential equation known as the matrix Riccati differential equation. To solve this equation we need to specify an initial condition. We note that

\[
P(0) = \exp(At)S(0)^{-1}\exp(A^T0) = S(0)^{-1}.
\]

But \(S(0)^{-1}\) does not exist! To overcome this difficulty, we may compute \(P(\Delta t)\) for some small \(\Delta t\) off-line, and then use the matrix Riccati differential equation (5) to generate \(P(t)\) for \(t > \Delta t\).
Finally, we need to set the initial state $\hat{x}(0)$ for the system model. For some small $\Delta t > 0$, we may set $\hat{x}(\Delta t)$ by

$$\hat{x}(\Delta t) = S(\Delta t)^{-1} \int_0^{\Delta t} \exp(A^T \tau)C^T\bar{y}(\tau)d\tau,$$

where

$$S(\Delta t) = \int_0^{\Delta t} \exp(A^T \tau)C^TC\exp(A \tau)d\tau$$

is computed off-line.

It is evident that the implementation of the foregoing observer or state estimator is quite complicated. This motivates the development of observers having simpler structures such as the one discussed in the following section.

**Asymptotic Observers or Estimators.**

Consider again the system

$$\frac{dx}{dt} = Ax + Bu,$$
$$y = Cx + Du.$$ 

We seek an observer of the form:

$$\frac{dz}{dt} = Hz + Gy + Qu = Hz + G(Cx + Du) + Qu.$$ 

The observer is driven by $u$ and the output of the system.

![Observer or State Estimator](image)

**Figure 2: Observer or State Estimator**

**Problem:** Determine $H, G, Q,$ and $T$ such that $z(t) \to Tx(t)$ as $t \to \infty$, where $T$ is a nonsingular linear transformation.

Suppose that $z(t) = Tx(t)$. Then

$$\frac{dz}{dt} = T \frac{dx}{dt} = TAx + TBu = Hz + G(Cx + Du) + Qu = (HT + GC)x + (GD + Q)u.$$ 

Thus,

$$TAx + TBu = (HT + GC)x + (GD + Q)u$$

for all $x$ and $u$.

This equation is satisfied, if

$$TA = HT + GC, \quad TB = GD + Q.$$

(6)
Now, consider
\[
\frac{d}{dt}(z(t) - Tx(t)) = Hz + G(Cx + Du) + Qu - T(Ax + Bu)
\]
\[
= Hz - TAx + GCx + (GD + Q - TB)u.
\]
\[
= Hz - (TA - GC)x = Hz - HTx = H(z - Tx)
\]
or
\[
\frac{d}{dt}(z - Tx) = H(z - Tx).
\]
Hence,
\[
z(t) = Tx(t) + \exp(HT)(z(0) - Tx(0))
\]

**Note:** If we choose \( z(0) = Tx(0) \), then \( z(t) = Tx(t) \) for all \( t \geq 0 \).

If we choose \( H \) such that the transients due to \( \exp(HT)(z(0) - Tx(0)) \) die out rapidly, then \( z(t) \cong Tx(t) \) for sufficiently large \( t \).

**Remarks:**

(R-1) The \( D \) matrix is a feedforward matrix which is at our discretion. Since \( u(t) \) is known, we can always modify \( D \) by feeding \( u(t) \) forward to \( y(t) \).

(R-2) The observer for a system with \( u(t) \neq 0 \) is driven by \( y = Cx + Du \). By choosing \( D \) and \( Q \) such that \( D + Q = TB \), the effect of the input \( u \) is cancelled. Thus, without loss of generality, we can design the observer by assuming that the system is free.

(R-3) The foregoing observer is not useful for estimating \( x(t) \) unless \( T \) is nonsingular, so that \( x(t) \cong T^{-1}z(t) \) for sufficiently large \( t \). We shall develop conditions for which \( T \) is invertible.

(R-4) In many systems, some components of \( x(t) \) are directly measurable. Thus, it is only necessary to estimate the remaining components of \( x(t) \). This can be done by using an observer with dimension \(< N = \dim(x) \).

**State Estimator.**

For simplicity, consider the free system:
\[
dx/dt = Ax, \quad y = cx,
\]
where \( c \) is a \( N \)-dimensional row vector. We shall seek an observer of the form:
\[
dz/dt = Hz + gy = Hz + gcx,
\]
where \( g \) is an undetermined \( N \)-dimensional column vector. For state estimation, we set \( T = I \). Thus, from (6),
\[
A = H + gc \quad \text{or} \quad H = A - gc.
\]
Hence, the observer is:
\[ \frac{dz}{dt} = (A - gc)z + gy(t). \]
If we choose \( g \) such that \( H \) or \( (A - gc) \) is a stable matrix. Then, \( \exp(Ht) \to 0 \) rapidly as \( t \to \infty \).

**Block Diagram of System:**

![Block Diagram of System](image)

Figure 3:

Now, we consider the question: Is it always possible to choose a \( g \) such that \( (A - gc) \) has eigenvalues with negative real parts so that \( z(t) \to x(t) \) as \( t \to \infty \)? We expect that observability will enter the answer to this question.

We shall obtain results for the more general case where \( T \neq I \).

**Lemma.** Let \( A \) and \( H \) be \( N \times N \) matrices with no common eigenvalues. Assume that \( g \) and \( c \) satisfy

\[
\text{Rank}[c^T|A^Tc^T|\ldots|(A^T)^{N-1}c^T] = N \quad ((A, c) \text{ is completely observable});
\]

and

\[
\text{Rank}[g|Hg|\ldots|H^{N-1}g] = N \quad ((H, g) \text{ is completely controllable}).
\]

Let \( T \) be the unique solution of \( TA - HT = gc \). Then, \( T \) is invertible.

**Proof:** (For the case where \( A \) has distinct eigenvalues only). In this case, \( A \) may be taken to be a diagonal matrix \( \Lambda \).

Consider

\[
TA - HT = gc.
\]

Let

\[
T = [t_1|\ldots|t_N]; \quad g = \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}; \quad c = [c_1 \ldots c_N].
\]
Then,

\[ T \Lambda = [t_1 | \ldots | t_N] \begin{bmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_N \end{bmatrix} = [\lambda_1 t_1 | \ldots | \lambda_N t_N] \]

\[ H \Lambda = [H t_1 | \ldots | H t_N]; \quad gc = \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix} \begin{bmatrix} c_1, \ldots, c_N \end{bmatrix} = [c_1 g | \ldots | c_N g]. \]

Thus,

\[ [\lambda_1 t_1 - H t_1 | \ldots | \lambda_N t_N - H t_N] = [c_1 g | \ldots | c_N g] \]

or

\[ \lambda_i t_i - H t_i = c_i g, \quad i = 1, \ldots, N \Rightarrow t_i = (\lambda_i I - H)^{-1} c_i g. \]

**Note:** \((\lambda_i I - H)^{-1}\) exists, if \(\lambda_i\) is not an eigenvalue of \(H\). To show that \(T\) is invertible, it is sufficient to show that \([t_1, \ldots, t_N]\) is linearly independent.

Consider a linear combination of \(t_i\)'s, i.e.

\[ \sum_{i=1}^{N} \alpha_i t_i = \sum_{i=1}^{N} \alpha_i (\lambda_i I - H)^{-1} c_i g. \]

By the complete observability assumption, \(c_i \neq 0\) for all \(i\). Hence \(c_i g \neq 0\). We shall establish the linear independence of \([t_1, \ldots, t_N]\) by contradiction.

Assume that \([t_i]\) is linearly dependent, then there exist a set of scalars \(\alpha_i, i = 1, \ldots, N\), not all zero, such that

\[ \sum_{i=1}^{N} \alpha_i (\lambda_i I - H)^{-1} c_i g = 0. \]

Multiply the above expression by \(\prod_i (\lambda_i I - H)\), we have

\[ \sum_{i=1}^{N} \alpha_i \prod_{j \neq 1}^{N} (\lambda_j I - H) c_i g = 0. \]

This follows from the fact that \((\lambda_i I - H)\) and \((\lambda_j I - H)\) commute. So we can move \((\lambda_i I - H)\) to the last term in the product. Thus, we have

\[ P(H)g = 0, \]

where \(P\) is a polynomial in \(H\) of degree \((N - 1)\) or less, or there exists a set of coefficients \(\{\delta_i\}\), not all zero, such that

\[ P(H)g = \sum_{i=0}^{N-1} \delta_i H^i g = 0. \]

This contradicts the complete controllability hypothesis that

\[ \text{Rank}[g|Hg| \ldots |H^{N-1}g] = N. \]
Theorem. Let $S : dx/dt = Ax, y = cx, c = [c_1, \ldots, c_N]$ be a $N$-dimensional completely observable system. Let $\{\mu_1, \ldots, \mu_N\}$ be a set of scalars distinct from the eigenvalues of $A$ (If $\mu_i$ is complex, then $\bar{\mu}_i$, the complex conjugate of $\mu_i$ is also in the set). Then a $(N - 1)$-dimensional observer of the form:

$$\frac{dz}{dt} = Hz + gy, \quad z = Tx$$

can be built for $S$ such that the eigenvalues of $H$ correspond to $(N - 1)$ of the $\mu$’s.

Proof: Let the observer be of the form:

$$\frac{dz}{dt} = \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_N \end{bmatrix} z + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} y, \quad (7)$$

where the $\mu$’s are distinct and $\mu_i \neq \lambda_k$ (eigenvalue of $A$) for all $i, k$. Now, let $z(0) = Tx(0)$ and $z(t) = Tx(t)$, where

$$T = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_N \end{bmatrix}, \quad (\tau_i \text{ is the } i\text{th row of } T).$$

Note that observer (7) is obviously completely controllable with respect to the input $y$. From the foregoing Lemma, the $N$ rows of $T$ are linearly independent, since $T$ is invertible. We claim that there exists one $\tau_i$ which may be replaced by the output matrix $c$ so that $\{\tau_1, \ldots, \tau_{i-1}, c, \tau_{i+1}, \ldots, \tau_N\}$ is linearly independent. To show this, we know that

$$\sum_{k=1}^{N} \alpha_k \tau_k = 0 \Rightarrow \alpha_k = 0 \text{ for all } k = 1, \ldots, N.$$ 

Thus, we can write

$$c = \sum_{k=1}^{N} \beta_k \tau_k.$$ 

Suppose that $\beta_i \neq 0$, hence

$$\tau_i = \frac{1}{\beta_i} c - \sum_{k=1, k \neq i}^{N} (\beta_k/\beta_i) \tau_k,$$

and

$$\sum_{k=1}^{N} \alpha_k \tau_k = \sum_{k=1, k \neq i}^{N} \alpha_k \tau_k + \alpha_i \left( \frac{1}{\beta_i} c - \sum_{k=1, k \neq i}^{N} (\beta_k/\beta_i) \tau_k \right) = \sum_{k=1, k \neq i}^{N} (\alpha_k - \alpha_i \beta_k/\beta_i) \tau_k + \frac{\alpha_i}{\beta_i} c = 0$$

implies $\alpha_i = 0$ for $k = 1, \ldots, N$. Thus,

$$(\alpha_k - \alpha_i \beta_k/\beta_i) = 0, \quad \alpha_i/\beta_i = 0,$$
which implies that \( \{ \tau_1, \ldots, \tau_{i-1}, c, \tau_{i+1}, \ldots, \tau_N \} \) is linearly independent.

Now, we can remove the \( i \)th dynamic element

\[
dz_i/dt = \mu_i z_i + y
\]

from the observer. Let

\[
\frac{d\hat{z}}{dt} = \begin{bmatrix}
\mu_1 & & & & \\
& \ddots & & & \\
& & \mu_{i-1} & & \\
& & & \mu_{i+1} & \\
& & & & \mu_N
\end{bmatrix}
\begin{bmatrix}
z_1 \\
\vdots \\
z_i \\
\vdots \\
z_N
\end{bmatrix}
+ \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix} y,
\end{array}
\]

\[
\begin{bmatrix}
\hat{z}(t) \\
y(t)
\end{bmatrix}
= \begin{bmatrix}
\tau_1 \\
\vdots \\
\tau_{i-1} \\
\tau_{i+1} \\
c
\end{bmatrix}
\]

\[
x(t) = \hat{T}x(t).
\]

Since \( \hat{T} \) is invertible, therefore \( x(t) \) is given by

\[
x(t) = \hat{T}^{-1} \begin{bmatrix}
\hat{z}(t) \\
y(t)
\end{bmatrix}.
\]

**Remark:** Following the same arguments, we can show that for the \( N \)-dimensional completely observable system: \( dx/dt = Ax, y = Cx \), where \( y \) is a \( M \)-dimensional column vector, a \( (N-M) \)-dimensional observer can be built. This is a generalization of the foregoing theorem.

**Example:** Consider the following two-dimensional system:

\[
\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad y = [1 \ 0] x.
\]

We wish to construct an observer for the state of this system. Here, since \( y = x_1 \), we only need to estimate \( x_2 \). Let the observer be of the form:

\[
dz/dt = \lambda z + k_1 y + k_2 u, \quad z(t) = Tz(t),
\]

where \( T \) is a \( 1 \times 2 \) matrix satisfying

\[
TA - HT = k_1 c \Rightarrow [T_{11} \ T_{12}] \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} - \lambda [T_{11} \ T_{12}] = [k_1 \ 0]
\]

or

\[
[-(2 + \lambda)T_{11} \ T_{11} - T_{12} - \lambda T_{12}] = [k_1 \ 0] \Rightarrow -(2 + \lambda)T_{11} = k_1, \quad T_{11} - T_{12} - \lambda T_{12} = 0. \quad (8)
\]

Also, we have

\[
k_2 = Tb \Rightarrow k_2 = [T_{12} \ T_{12}] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = T_{12} \Rightarrow k_2 = T_{12}.
\]

111
Let $k_1 = 1, \lambda = -10$ (so that the observer transient dies out rapidly). So the equations (8) become

$-(2 - 10)T_{11} = 1$ or $T_{11} = 1/8$. $T_{12} = T_{11}/(1 - 10) = -1/72$.

A block diagram of the observer system is shown below:

**Figure 4: Observer or State Estimator**

Note: The observer is not unique.

**Relationship Between the Exact Finite-time Observer and Asymptotic Observers:**

Consider again the exact finite-time observer given by (3), i.e.

$$d\hat{x}(t)/dt = (A - P(t)C^TC)\hat{x}(t) + P(t)C^Ty(t) + (A - A + P(t)C^TC - P(t)C^TC)x_u(t) - P(t)C^TDu(t) + Bu(t)$$

$$= A\hat{x}(t) + P(t)C^T(y(t) - C\hat{x}(t)) + (B - P(t)C^TD)u(t),$$

where $P(t)$ satisfies the matrix Riccati differential equation given by (5), i.e.

$$dP(t)/dt = AP(t) + P(t)A^T - P(t)C^TP(t).$$

We let the estimation time $\to \infty$ so that $P(t) \to P_\infty$ (a constant matrix), where $P_\infty$ satisfies the steady-state matrix Riccati equation:

$$AP_\infty + P_\infty A^T = P_\infty CC^TP_\infty.$$

This suggests that for the linear time-invariant system:

$$dx/dt = Ax + Bu, \quad y = Cx + Du,$$
we may seek an asymptotic state estimator of the form:

$$\frac{d\hat{x}}{dt} = A\hat{x} + PC^T(y - C\hat{x}) + (B - PC^TD)u,$$

(9)

where $P$ is a constant matrix. This estimator has the same form as that of the asymptotic observer of the form:

$$\frac{dz}{dt} = Hz + Gy + Qu$$

derived earlier, if we set

$$z = \hat{x}, \quad T = I, \quad H = A - GC, \quad G = PC^T, \quad Q = B - GD.$$

Let the state estimation error be defined by $e(t) = x(t) - \hat{x}(t)$. Then, $e(t)$ satisfies:

$$\frac{de}{dt} = He = (A - PC^TC)e.$$

Thus,

$$e(t) = (\exp(A - PC^TC)t)e(0), \quad t \geq 0.$$

We wish to choose $P$ such that $(A - PC^TC)$ is stable (i.e. all its eigenvalues have negative real parts). This is possible if $(A, C)$ is completely observable (Why?).

We shall consider the single-input single-output system described by the scalar differential equation:

$$\frac{d^Ny}{dt^N} - \sum_{j=0}^{N-1} a_j\frac{d^jy}{dt^j} = u(t),$$

where $y$ denotes the output.

Let $w = (y, dy/dt, \ldots, d^{N-1}y/dt^{N-1})^T$. Then the above equation can be rewritten in the following phase-variable canonical form:

$$\frac{dw}{dt} = \tilde{A}w + \tilde{b}u, \quad y = \tilde{c}w,$$

(10)

where

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \vdots \\ a_0 & a_1 & \cdots & a_{N-1} & \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \tilde{c} = [1, 0, \ldots, 0].$$

By the linear transformation

$$w = Tx = \sum_{j=0}^{N-1} x_j(\tilde{A}^j\tilde{b}), \quad x = (x_1, \ldots, x_N)^T,$$

(10) can be transformed into the first controllability canonical form:

$$\frac{dx}{dt} = Ax + bu, \quad y = cx,$$

(11)
where $\mathbf{A} = \mathbf{\tilde{A}}^T$, $\mathbf{b} = [1, 0, \ldots, 0]^T$, and $\mathbf{c} = [0, \ldots, 0, 1]$. Note that since

$$\text{Rank}[\mathbf{\tilde{b}} | \mathbf{\tilde{A}}\mathbf{\tilde{b}} | \ldots | \mathbf{\tilde{A}}^{N-1}\mathbf{\tilde{b}}] = N,$$

the transformation $\mathbf{T}$ is nonsingular.

Now, we shall construct an asymptotic state estimator for (11) of the form:

$$\frac{d\hat{\mathbf{x}}}{dt} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{Pc}^T(\mathbf{y} - \mathbf{c}\hat{\mathbf{x}}) + \mathbf{bu}.$$

We shall show that we can always construct a $N \times N$ constant matrix $\mathbf{P}$ such that $(\mathbf{A} - \mathbf{Pc}^T\mathbf{c})$ is stable.

By direct computation, we have

$$\mathbf{c}^T\mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ [0, \ldots, 0, 1] \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 1 \end{bmatrix};$$

$$\mathbf{Pc}^T\mathbf{c} = \begin{bmatrix} p_{11} & \cdots & p_{N1} \\ \vdots & \vdots & \vdots \\ p_{N1} & \cdots & p_{NN} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & p_{1N} \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & p_{NN} \end{bmatrix};$$

$$\mathbf{A} - \mathbf{Pc}^T\mathbf{c} = \begin{bmatrix} 0 & 0 & \cdots & a_0 - p_{1N} \\ 0 & 0 & \cdots & a_1 - p_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Evidently, the characteristic polynomial of $(\mathbf{A} - \mathbf{Pc}^T\mathbf{c})$ is

$$p(\lambda) = \lambda^N + (a_{N-1} - p_{NN})\lambda^{N-1} + \ldots + (a_0 - p_{1N}) = 0.$$

Thus, we can always choose suitable $p_{ij}$ such that $(\mathbf{A} - \mathbf{Pc}^T\mathbf{c})$ is stable.