Def. The impulse response matrix of a linear, lumped, time-varying system is a matrix map

\[ H(\bullet, \bullet) : R \times R \rightarrow R^{r \times m} \]

given by

\[ H(t, \tau) = [h_1(t, \tau), \ldots, h_m(t, \tau)] \]

where each column \( h_i(t, \tau) \) represents the response of the system to the impulsive input

\[ u(t) = \delta(t - \tau) \eta_i, \quad \eta_i \in R^m, \]

\[ \eta_i = [0 \ldots 0 1 0 \ldots 0]^T \] (1 occurs as the \( i \)th component), \( \tau \) is the time of application of the input and \( t \) is the observation time.

Recall that if \( H(t, \tau) \neq 0_{r \times m} \) for any given \( t, \tau \) and \( t < \tau \), then the system is noncausal. On the other hand, if \( H(t, \tau) = 0_{r \times m} \) for \( t < \tau \), then the system is causal.

In block diagram form,

\[ \begin{align*}
&u(t) \quad \text{H(t,\tau)} \quad y(t) \\
&\text{Hence,} \\
&y(t) = \int_{-\infty}^{\infty} H(t,\tau)u(\tau)d\tau
\end{align*} \]

Consider the linear, time-varying system with state model

\[ \dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(-\infty) = 0 \]

\[ y(t) = C(t)x(t) + D(t)u(t) \]
Theorem: The impulse response matrix for the above system is given by

$$H(t, \tau) = \begin{cases} C(t)\Phi(t, \tau)B(\tau) + D(t)\delta(t - \tau) & t \geq \tau \\ 0 & t < \tau \end{cases}$$

Proof: The response of the above time-varying system to an input $u$ is given by

$$y(t) = C(t)\Phi(t,-\infty)x(-\infty) + C(t)\int_{-\infty}^{t} \Phi(t, q)B(q)u(q)dq + D(t)u(t)$$

But, $x(-\infty) = 0$, implies that

$$y(t) = C(t)\int_{-\infty}^{t} \Phi(t, q)B(q)u(q)dq + D(t)u(t)$$

Let $u(t) = \delta(t - \tau)\eta_i$, $\eta_i = [0 \ldots 0 1 0 \ldots 0]$, where the 1 appears at the $i^{th}$ location, then for $t \geq \tau$

$$y(t) = C(t)\int_{-\infty}^{t} \Phi(t, q)B(q)\delta(q - \tau)\eta_i dq + D(t)\delta(t - \tau)\eta_i = C(t)\Phi(t, \tau)b_i(\tau) + d_i(t)\delta(t - \tau)$$

$i = 1, 2, \ldots, m$, i.e. we pick up the contribution due to the $i^{th}$ input.

Recall that, $y(t) = \int_{-\infty}^{\infty} H(t, q)u(q)dq$
If the system is causal, then \( H(t, \tau) = 0 \) for \( \tau > t \). Thus, for \( t \geq \tau \)

\[
y(t) = \int_{-\infty}^{t} H(t, q)u(q)dq = \int_{-\infty}^{t} \left[ h_1(t, q) \ h_2(t, q) \ \cdots \ h_m(t, q) \right] u(q)dq
\]

and \( y(t) = 0 \), for \( t < \tau \).

If \( u(t) = \delta(t - \tau)\eta_i \), then

\[
y(t) = \begin{cases} h_i(t, \tau) & t \geq \tau \\ 0 & t < \tau \end{cases}
\]

Hence, \( H(t, \tau) = C(t)\Phi(t, \tau)B(\tau) + D(\tau)\delta(t - \tau), \quad t \geq \tau \)

\[
= 0 \quad t < \tau
\]

If \( A(\bullet), B(\bullet), C(\bullet) \) and \( D(\bullet) \) are constant matrices, then

\[
H(t, \tau) = H(t - \tau, 0) \equiv H(t - \tau) = \begin{cases} Ce^{A(t-\tau)}B + D\delta(t - \tau) & t \geq \tau \\ 0 & t < \tau \end{cases}
\]

Consider again the time-invariant system state model

\[
\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0
\]

\[
y(t) = Cx(t) + Du(t)
\]
In the s-domain,
\[
\mathcal{L}\{\dot{x}(t)\} = sX(s) - x_0 = \mathcal{L}\{Ax(t) + Bu(t)\} = AX(s) + BU(s)
\]
\[
(sI - A)X(s) = x_0 + BU(s)
\]
Hence,
\[
X(s) = (sI - A)^{-1} x_0 + (sI - A)^{-1} BU(s)
\]
\[
Y(s) = C(sI - A)^{-1} x_0 + [C(sI - A)^{-1} B + D]U(s)
\]
If the system is initially at rest, i.e., \(x_0 = 0\), then
\[
Y(s) = \left[C(sI - A)^{-1} B + D\right]U(s) = H(s)U(s)
\]
**Def.** The transfer function matrix of the time-invariant state model is given by
\[
H(s) = C(sI - A)^{-1} B + D
\]
**Def.** (The Leverrier Algorithm). Let the polynomial
\[
\pi_A(\lambda) \equiv \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n
\]
be the characteristic polynomial of the matrix \(A \in \mathbb{R}^{n \times n}\), then the coefficients \(a_i\), \(i = 1, 2, \ldots, n\) can be computed as follows:
\[
N_1 = I, \quad a_1 = - \text{tr}(A)
\]
\[
N_2 = N_1 A + a_1 I, \quad a_2 = (-1/2) \text{tr}(N_2 A),
\]
\[ N_3 = N_2 A + a_2 I, \quad a_3 = (-1/3) \text{tr}(N_3 A), \]
\[ \vdots \]
\[ N_i = N_{i-1} A + a_{i-1} I, \quad a_i = (-1/i) \text{tr}(N_i A), \quad i \leq n \]
and \[ N_{n+1} = 0 = N_n A + a_n I, \] where \( \text{tr}(M) \) is the trace of the matrix \( M \).

**Example:** Let a dynamic system have the feedback matrix \( A \) be given by
\[
A = \begin{bmatrix}
-2 & 0 & 1 \\
1 & -2 & 0 \\
1 & 1 & -1
\end{bmatrix}
\]
Then, \( N_1 = I, \quad a_1 = -\text{tr}(A) = -(-5) = 5, \quad N_2 = N_1 A + a_1 I = A + a_1 I = \begin{bmatrix} 3 & 0 & 1 \\
1 & 3 & 0 \\
1 & 1 & 4 \end{bmatrix} \]
\[
a_2 = -\frac{1}{2} \text{tr}(N_2 A) = -\frac{1}{2} \text{tr}\left(\begin{bmatrix}
-5 & 1 & 2 \\
1 & -6 & 1 \\
3 & 2 & -3
\end{bmatrix}\right) = 7, \quad \Rightarrow N_3 = N_2 A + a_2 I = \begin{bmatrix} 2 & 1 & 2 \\
1 & 1 & 1 \\
3 & 2 & 4 \end{bmatrix}
\]
and
\[ a_3 = -\frac{1}{3} \text{tr}(N_3 A) = -\frac{1}{3} \text{tr} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = 1 \]

To make sure the calculations are correct, check that \( N_4 = N_3 A + a_3 I = 0 \)

Finally, the characteristic polynomial of \( A \) is \( \pi_A(\lambda) = \lambda^3 + 5\lambda^2 + 7\lambda + 1 \).

**Faddeev-Leverrier Algorithm for Computing \((sI - A)^{-1}\)**

Let \((sI - A)^{-1} \equiv \frac{\text{Adj}(sI - A)}{\det(sI - A)} = \frac{R(s)}{\pi_A(s)} = \frac{N_1 s^{n-1} + N_2 s^{n-2} + \cdots + N_{n-1}s + N_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n} \)

where \( \pi_A(s) \) is the characteristic polynomial of \( A \), \( R(s) \) is the adjoint matrix of \( sI - A \).

**Proposition:** \( \mathcal{L}\{\Phi(t)\} = (sI - A)^{-1} \), where \( \mathcal{L}\{\cdot\} \) is the Laplace transform operator.

**Proof:** For the time-invariant case and \( u(t) = 0 \),
\[
\dot{\Phi}(t) = A\Phi(t), \quad \Phi(0), \quad t \geq 0
\]
\[ s\mathcal{L}\{\Phi(t)\} - \Phi(0) = A\mathcal{L}\{\Phi(t)\}. \] But, \( \Phi(0) = I \), thus \( \mathcal{L}\{\Phi(t)\} = (sI - A)^{-1} \).

Clearly, \( \Phi(t) = \mathcal{L}^{-1}\{(sI - A)^{-1}\} \).

Cayley-Hamilton Theorem:

Let \( \pi_A(\lambda) \equiv \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \)
be the characteristic polynomial of \( A \). Then

\[ \pi_A(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI = 0 \]

This implies that

\[ A^n = -\sum_{i=1}^{n} a_i A^{n-i} \]

Likewise, \( A\pi_A(A) = 0 \Rightarrow A^{n+1} = -\sum_{i=1}^{n} a_i A^{n-i+1} \)

or

\[ A^{n+1} = -a_1A^n - \sum_{i=2}^{n} a_i A^{n-i+1} \]

\[ = -a_1 \left[ -\sum_{i=1}^{n} a_i A^{n-i} \right] - \sum_{i=2}^{n} a_i A^{n-i+1} = a_1^2 A^{n-1} + \sum_{i=2}^{n} [a_1 I - A]a_i A^{n-i} \]
Observation 1: $A^{n+1}$ is also a linear combination of $I, A, A^2, \ldots , A^{n-1}$.

Observation 2: Every polynomial of $A \in R^{n \times n}$ can be expressed as a linear combination of $I, A, A^2, \ldots , A^{n-1}$, i.e.,

$$f(A) = \beta_0 I + \beta_1 A + \ldots + \beta_{n-1} A^{n-1}.$$ 

Furthermore, if a polynomial $f(\lambda)$ is such that $\deg(f(\lambda)) > \deg(\pi_A(\lambda))$, then

$$f(\lambda) = q(\lambda)\pi_A(\lambda) + h(\lambda), \quad \deg(h(\lambda)) < n \quad \bigotimes$$

$$\Rightarrow \quad f(A) = q(A)\pi_A(A) + h(A) = q(A). \ 0 + h(A) = h(A)$$

$$\Rightarrow \text{if } \deg(f(\lambda)) > \deg(\pi_A(\lambda)), \text{ we can solve for } h(\lambda) \text{ directly from } \bigotimes, \text{ i.e.,}$$

Let $h(\lambda) \equiv \beta_0 + \beta_1 \lambda + \ldots + \beta_{n-1} \lambda^{n-1}$, then if the eigenvalues of $A$ are distinct, the $\beta_j$’s can be computed from the $n$ linear equations

$$f(\lambda_i) = q(\lambda_i)\pi_A(\lambda_i) + h(\lambda_i) = h(\lambda_i), \ i = 1, 2, \ldots, n$$

Example: Calculate $f(A) = A^{10} + 3A$ with $A = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$
The characteristic polynomial of $A$ is $\pi_A(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 \implies \lambda_1 = \lambda_2 = -1$.

Let $h(\lambda) = \beta_0 + \beta_1 \lambda$. Then the scalar polynomial is $f(\lambda) = \lambda^{10} + 3\lambda$, and

$$f(\lambda_1) = f(-1) = (-1)^{10} + 3(-1) = -2 = h(\lambda_1) = h(-1) = \beta_0 - \beta_1.$$ 

But, $f(\lambda_2) = f(\lambda_1)$! For the repeated eigenvalue case, the solution procedure is modified as follows:

Let $\pi_A(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i} \ni n = \sum_{i=1}^{m} n_i$

Since $\deg(h(\lambda)) \leq n - 1$, the coefficients $\beta_j$, $j = 0, 1, \ldots, n - 1$, can now be obtained from the following set of $n$ equations:

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i), \quad l = 0, 1, \ldots, n_i - 1; \quad i = 1, 2, \ldots, m,$$

where $f^{(l)}(\lambda_i)$ and $h^{(l)}(\lambda_i)$ are the $l^{th}$ derivatives of $f(\lambda)$ and $h(\lambda)$, respectively, evaluated at $\lambda = \lambda_i$.

Going back to the previous example, $n_1 = 2$ and

$$f^{(1)}(\lambda) \bigg|_{\lambda = \lambda_1} = 10\lambda^9 + 3 \bigg|_{\lambda = -1} = 10(-1)^9 + 3 = -7 = h^{(1)}(\lambda) \bigg|_{\lambda = \lambda_1} = \beta_1.$$
We must solve the equations $\beta_0 - \beta_1 = -2$ and $\beta_1 = -7$. This implies that $\beta_0 = -9$.

Moreover,

$$f(A) = A^{10} + 3A = h(A) = \beta_0 I + \beta_1 A = -9I - 7A = \begin{bmatrix} -9 & -7 \\ 7 & 5 \end{bmatrix}$$

Suppose again the eigenvalues of $A$ are distinct. Then the partial fraction expansion of $(sI - A)^{-1}$ is given by

$$(sI - A)^{-1} = \frac{R(s)}{\pi_A(s)} = \frac{R_1}{s - \lambda_1} + \frac{R_2}{s - \lambda_2} + \cdots + \frac{R_n}{s - \lambda_n}$$

where $\pi_A(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n = \prod_{i=1}^n (s - \lambda_i)$

and $R_i = \lim_{s \to \lambda_i} \left[ \frac{(s - \lambda_i) R(s)}{\pi_A(s)} \right]$ is the $i$th residue of matrix of $(sI - A)^{-1} = \frac{R(s)}{\pi_A(s)}$

Consequently, for $t \geq 0$

$$\Phi(t) = \mathcal{L}^{-1}\{(sI - A)^{-1}\} = \sum_{i=1}^n \mathcal{L}^{-1}\left\{ \frac{R_i}{s - \lambda_i} \right\} = \sum_{i=1}^n R_i e^{\lambda_i t} = e^{At}$$
Example: Consider the system \( \dot{x}(t) = Ax(t) \). Obtain \( \Phi(t) \) using the partial fraction expansion method with the system matrix \( A \) given by

\[
A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}
\]

Now,

\[
sI - A = \begin{bmatrix} s & -1 \\ 2 & s + 3 \end{bmatrix} \Rightarrow R(s) = \begin{bmatrix} s + 3 & 1 \\ -2 & s \end{bmatrix}
\]

and the characteristic polynomial of \( A \) is

\[
\pi_A(s) = \det(sI - A) = s^2 + 3s + 2 = (s + 1)(s + 2) = (s - \lambda_1)(s - \lambda_2)
\]

The residue matrices are found as follow:

\[
R_1 = \lim_{s \to -1} \left( (s + 1) \frac{R(s)}{\pi_A(s)} \right) = \lim_{s \to -1} \left[ \frac{1}{s + 2} \begin{pmatrix} s + 3 & 1 \\ -2 & s \end{pmatrix} \right] = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}
\]

\[
R_2 = \lim_{s \to -2} \left( (s + 2) \frac{R(s)}{\pi_A(s)} \right) = \lim_{s \to -2} \left[ \frac{1}{s + 1} \begin{pmatrix} s + 3 & 1 \\ -2 & s \end{pmatrix} \right] = \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix}
\]
The inverse of $sI - A$ is equal to

$$(sI - A)^{-1} = \frac{R(s)}{\pi_A(s)} = \frac{R_1}{s+1} + \frac{R_2}{s+2} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix}$$

and the state transition matrix of the system is

$$\Phi(t) = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} \mathcal{L}^{-1} \left\{ \frac{1}{s+1} \right\} + \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix} \mathcal{L}^{-1} \left\{ \frac{1}{s+2} \right\} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} e^{-t} + \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix} e^{-2t}$$

or

$$\Phi(t) = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

Suppose now that the matrix $A$ contains repeated eigenvalues, then the adjoint matrix $R(s)$ and $\pi_A(s)$ may have common factors. Consider, for example, the matrix

$$A = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$$
Then,
\[
(sI - A)^{-1} = \frac{R(s)}{\pi_A(s)} = \frac{\begin{bmatrix}
(s - \lambda_1)^2 & s - \lambda_1 & 0 \\
0 & (s - \lambda_1)^2 & 0 \\
0 & 0 & (s - \lambda_1)^2 \\
\end{bmatrix}}{(s - \lambda_1)^3} = \frac{\begin{bmatrix}
s - \lambda_1 & 1 & 0 \\
0 & s - \lambda_1 & 0 \\
0 & 0 & s - \lambda_1 \\
\end{bmatrix}}{(s - \lambda_1)^2}
\]

We know from the Cayley-Hamilton theorem that if \( \pi_A(\lambda) \) is the characteristic polynomial of the \( n \times n \) matrix \( A \), then \( \pi_A(A) = 0 \).

**Def.** A polynomial \( p(\lambda) \) such that \( p(A) = 0 \) is called an annihilating polynomial of the matrix \( A \).

**Def.** The monic polynomial of least degree which annihilates the matrix \( A \) is called the **minimal polynomial** of \( A \) and is denoted by \( \psi_A(\lambda) \).

Suppose \( g(\lambda) \) is a polynomial of arbitrary degree, then \( p(\lambda) = g(\lambda)\pi_A(\lambda) \) is also an annihilating polynomial of \( A \).

**Theorem:** For every \( n \times n \) matrix \( A \), the minimal polynomial \( \psi_A(\lambda) \) divides the characteristic polynomial \( \pi_A(\lambda) \). Moreover, \( \psi_A(\lambda) = 0 \) if and only if \( \lambda \) is an eigenvalue of \( A \), so that every root of \( \psi_A(\lambda) = 0 \) is also a root of \( \pi_A(\lambda) = 0 \).
Proof: If $\pi_A(\lambda)$ annihilates $A$ and if $\psi_A(\lambda)$ is a monic polynomial of minimum degree that annihilates $A$, then $\deg (\psi_A(\lambda)) \leq \deg(\pi_A(\lambda))$.

By the Euclidean algorithm there exists polynomials $q(\lambda)$ and $r(\lambda)$ such that

$$\pi_A(\lambda) = \psi_A(\lambda)q(\lambda) + r(\lambda) \text{ and } \deg(r(\lambda)) < \deg(\psi_A(\lambda))$$

But, $0 = \pi_A(A) = \psi_A(A)q(A) + r(A) = 0 \cdot q(A) + r(A) \implies r(A) = 0$.

However, $\deg(r(\lambda)) < \deg(\psi_A(\lambda))$, and by definition $\psi_A(\lambda)$ is the polynomial of minimum degree such that $\psi_A(A) = 0 \implies r(\lambda) \equiv 0 \implies \psi_A(\lambda)$ divides $\pi_A(\lambda)$.

This result implies that every root of $\psi_A(\lambda) = 0$ is a root of $\pi_A(\lambda) = 0$ and hence every root of $\psi_A(\lambda) = 0$ is an eigenvalue of $A$.

If $\lambda \in \sigma(A)$ and if $x \neq 0$ is its corresponding eigenvector, then $Ax = \lambda x$ and $0 = \psi_A(A) x = \psi_A(\lambda) x \implies \psi_A(\lambda) = 0$.

If $A$ has $\sigma$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_\sigma$, $\lambda_i \neq \lambda_k, i \neq k$, $i, k = 1, 2, \ldots, \sigma$, $\sigma < n$, each of which is repeated $m_i$ times, then for $m_1 + m_2 + \ldots + m_\sigma \leq n$ the minimal polynomial $\psi_A(\lambda)$ has the structure

$$\psi_A(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_\sigma)^{m_\sigma}$$
Theorem: Let \( A \in \mathbb{R}^{n \times n} \), then
\[
(sI - A)^{-1} = R(s) = R(s) = \sum_{i=1}^{\sigma} \sum_{j=1}^{m_i} \frac{R_i^j}{(s - \lambda_i)^j}
\]
where
\[
R_i^j = \lim_{s \to \lambda_i} \frac{1}{(m_i - j)!} \left[ \frac{d^{m_i-j}}{ds^{m_i-j}} (s - \lambda_i)^{m_i} (sI - A)^{-1} \right]
\]
In this particular case, for \( t \geq 0 \) the state transition matrix is given by
\[
\Phi(t) = \sum_{i=1}^{\sigma} \sum_{j=1}^{m_i} R_i^j \frac{t^{j-1}}{(j-1)!} e^{\lambda_i t}
\]
This is true because
\[
\Phi(t) = \mathcal{L}^{-1}\{ (sI - A)^{-1} \} = \sum_{i=1}^{\sigma} \sum_{j=1}^{m_i} R_i^j \mathcal{L}^{-1} \left\{ \frac{1}{(s - \lambda_i)^j} \right\} = \sum_{i=1}^{\sigma} \sum_{j=1}^{m_i} R_i^j \frac{t^{j-1}}{(j-1)!} e^{\lambda_i t}
\]
Example: Consider a dynamic system with the $A$ matrix

$$
A = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-2 & -5 & -4
\end{bmatrix}
$$

Then

$$
\pi_A(s) = \det(sI - A) = (s + 1)^2(s + 2) = s^3 + 4s^2 + 5s + 2
$$

$$
\Rightarrow a_1 = 4, a_2 = 5, a_3 = 2
$$

Using the Faddeev-Leverrier algorithm, we get

$$
N_1 = I
$$

$$
N_2 = N_1 A + a_1 I = \begin{bmatrix}
4 & 1 & 0 \\
0 & 4 & 1 \\
-2 & -5 & 0
\end{bmatrix}
$$

$$
N_3 = N_2 A + a_2 I = \begin{bmatrix}
5 & 4 & 1 \\
-2 & 0 & 0 \\
0 & -2 & 0
\end{bmatrix}
$$
So,
\[ R(s) = N_1 s^2 + N_2 s + N_3 = \begin{bmatrix} s^2 + 4s + 5 & s + 4 & 1 \\ -2 & s^2 + 4s & s \\ -2s & -5s - 2 & s^2 \end{bmatrix} \]
and
\[ (sI - A)^{-1} = \frac{R(s)}{(s + 1)^2(s + 2)} = \frac{N_1 s^2 + N_2 s + N_3}{(s + 1)^2(s + 2)} = \frac{Is^2 + N_2 s + N_3}{(s + 1)^2(s + 2)} \]

In this case, the minimal polynomial is the same as the characteristic polynomial, i.e., \( \psi_A(s) = \pi_A(s) = s^3 + 4s^2 + 5s + 2 \). This is because there are no cancellations.

The inverse of \( sI - A \) is

\[ (sI - A)^{-1} = \frac{R_1^1}{s + 1} + \frac{R_1^2}{(s + 1)^2} + \frac{R_2^1}{s + 2} \]

where the residue matrices are

\[ R_1^1 = \frac{1}{1!} \lim_{s \to -1} \left\{ \frac{d}{ds} (s + 1)^2(sI - A)^{-1} \right\} = \lim_{s \to -1} \left\{ \frac{d}{ds} \left( \frac{Is^2 + N_2 s + N_3}{s + 2} \right) \right\} = -3I + 2N_2 - N_3 \]

\[ = \begin{bmatrix} 0 & -2 & -1 \\ 2 & 5 & 2 \\ -4 & -8 & -3 \end{bmatrix} \]
\[ R_1^2 = \lim_{s \to -1} \left\{ (s + 1)^2 (sI - A)^{-1} \right\} = \lim_{s \to -1} \left\{ \frac{Is^2 + N_2 s + N_3}{s + 2} \right\} = I - N_2 + N_3 = \begin{bmatrix} 2 & 3 & 1 \\ -2 & -3 & -1 \\ 2 & 3 & 1 \end{bmatrix} \]

\[ R_2^1 = \lim_{s \to -2} \left\{ \frac{Is^2 + N_2 s + N_3}{(s + 1)^2} \right\} = 4I - 2N_2 + N_3 = \begin{bmatrix} 1 & 2 & 1 \\ -2 & -4 & -2 \\ 4 & 8 & 4 \end{bmatrix} \]

Therefore, for \( t \geq 0 \), the state transition matrix is given by

\[ e^{At} = \Phi(t) = R_1^1 e^{-t} + R_2^1 te^{-t} + R_2^1 e^{-2t} \]

Consider now the following block diagonal \( \hat{A} \) matrix

\[ \hat{A} = T^{-1}AT = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \], where \( T \) is a nonsingular similarity transformation.

Then, its characteristic polynomial is the same as that of the last example, i.e.,

\[ \pi_{\hat{A}}(s) = \pi_A(s) = (s + 1)^2 (s + 2) \]
In this case,

\[
e^{At} = \begin{bmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-2t} \end{bmatrix}
\]

But,

\[A = T \hat{A} T^{-1} \Rightarrow e^{At} = T e^{\hat{A}t} T^{-1}\]

provided that a nonsingular \(T\) can be constructed.

Suppose that

\[
J = \begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & \vdots \\ \vdots & \ddots & \ddots \\ 0 & \cdots & \cdots & J_p \end{bmatrix}
\]

where

\[
J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda_i \end{bmatrix}, \quad J_i \in \mathbb{R}^{n_i \times n_i}
\]
then

\[
e^{Jt} = \begin{bmatrix}
e^{J_{1t}} & 0 & 0 & 0 \\
0 & e^{J_{2t}} & \cdots & \\
\vdots & \ddots & \ddots & \\
0 & \cdots & \cdots & e^{J_{pt}}
\end{bmatrix}
\]

where

\[
e^{J_i t} = \begin{bmatrix}
e^{\lambda_{i1} t} & te^{\lambda_{i1} t} & \frac{t^{n_i - 1}}{(n_i - 1)!}e^{\lambda_{i1} t} \\
0 & e^{\lambda_{i2} t} & \cdots & \\
\vdots & \ddots & \ddots & \\
0 & \cdots & \cdots & e^{\lambda_{ip} t}
\end{bmatrix}
\]

In this case, matrix $\otimes$ is said to be in the Jordan canonical form.

Let $A$ have eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$, each with multiplicity $n_i$, $i = 1, \ldots, p$, $n_1 + n_2 + \ldots + n_p = n$. Suppose $A$ has $p$ independent eigenvectors $e_1, e_2, \ldots, e_p$ associated with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$. If $q_i = n - \text{rank}(\lambda_i I - A) = 1$, then
Then the set of eigenvectors \( \{ e_1^1, ..., e_1^{n_1}, ..., e_p^1, ..., e_p^{n_p} \} \) generated by

\[
(A - \lambda_i I) e_i^1 = 0 \\
(A - \lambda_i I) e_i^2 = e_i^1 \\
(A - \lambda_i I) e_i^3 = e_i^2 \\
\vdots \\
(A - \lambda_i I) e_i^{n_i} = e_i^{n_i-1}
\]

\( i = 1, 2, \ldots, p \) forms a basis for the \( n \)-dimensional space.

**Theorem:** The generalized eigenvectors \( \{ e_1^1, ..., e_1^{n_1}, ..., e_p^1, ..., e_p^{n_p} \} \) of \( A \) associated with the eigenvalues \( \{ \lambda_1, ..., \lambda_p \} \) each with multiplicity \( n_i, i = 1, 2, \ldots, p \), are linearly independent.

**Example:** Let \( A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix} \). We already know that \( \sigma(A) = \{-1, -1, -2\} \).
Clearly, $n_1 = 2$ and $n_2 = 1$. Moreover,

\[
(A - \lambda_1 I)e_1^1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -2 & -5 & -3 \end{bmatrix} \begin{bmatrix} e_{11}^1 \\ e_{12}^1 \\ e_{13}^1 \end{bmatrix} = 0 \Rightarrow e_1^1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}
\]

\[
(A - \lambda_2 I)e_2^1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ -2 & -5 & -2 \end{bmatrix} \begin{bmatrix} e_{21}^1 \\ e_{22}^1 \\ e_{23}^1 \end{bmatrix} = 0 \Rightarrow e_2^1 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}
\]

Clearly, $e_1^1$ and $e_2^1$ are two linearly independent eigenvectors.

Furthermore, $q_1 = n - \text{rank}(\lambda_1 I - A) = 3 - 2 = 1$, and

\[
(A - \lambda_1 I)e_1^2 = e_1^1 \Rightarrow e_1^2 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T
\]

and $\text{det}\begin{bmatrix} e_1^1 & e_1^2 & e_2^1 \end{bmatrix} = 1$ implies that the three eigenvectors are linearly independent.

Let us construct the similarity transformation $T$ as follows: $T = \begin{bmatrix} e_2^1 & e_1^1 & e_1^2 \end{bmatrix}$
In other words,

\[
T = \begin{bmatrix}
1 & 1 & 1 \\
-2 & -1 & 0 \\
4 & 1 & -1
\end{bmatrix}
\quad \text{and} \quad
T^{-1} = \begin{bmatrix}
1 & 2 & 1 \\
-2 & -5 & -2 \\
2 & 3 & 1
\end{bmatrix}
\]

The new system matrix, in Jordan canonical form is given by

\[
\hat{A} = T^{-1}AT = \begin{bmatrix}
-2 & 0 & 0 \\
0 & -1 & 1 \\
0 & 0 & -1
\end{bmatrix} = J
\]

The matrix exponential of the equivalent system is

\[
e^{\hat{A}t} = e^{JT} = \begin{bmatrix}
e^{-2t} & 0 & 0 \\
0 & e^{-t} & te^{-t} \\
0 & 0 & e^{-t}
\end{bmatrix}
\]
The matrix exponential of the original system is therefore given by

\[ e^{At} = T e^{Jt} T^{-1} = \begin{bmatrix}
    e^{-2t} + 2t e^{-t} & 2e^{-2t} - (2 + 3t) e^{-t} & e^{-2t} - (1 + t) e^{-t} \\
    -2e^{-2t} + 2(1-t)e^{-t} & -4e^{-2t} + (5 - 3t) e^{-t} & -2e^{-2t} + (2 - t) e^{-t} \\
    4e^{-2t} - (4 - 2t)e^{-t} & 8e^{-2t} - (8 - 3t) e^{-t} & 4e^{-2t} - (3 - t) e^{-t}
\end{bmatrix} \]

**Discrete-Time Systems:**

Consider the linear time-invariant discrete-time system

\[
    x(k + 1) = Ax(k) + Bu(k), \quad x_0 = x(0) \\
    y(k) = Cx(k) + Du(k)
\]

Taking the one-sided Z transform, yields

\[
    X(z) = z(zI - A)^{-1} x_0 + (zI - A)^{-1} B U(z) \\
    y(z) = zC(zI - A)^{-1} x_0 + \left[ C(zI - A)^{-1} B + D \right] U(z)
\]

The state transition matrix is given by \( \Phi(k) = A^k = Z^{-1}\{z(zI-A)^{-1}\} \).
The transfer function matrix of a discrete-time linear dynamic system is defined by

\[ H(z) = C(zI - A)^{-1}B + D \]

On the other hand, the unit sample response matrix is then given by

\[ h(k) = \mathcal{Z}^{-1}\{H(z)\}. \]

**Stability of Dynamic Systems**

Dynamic system stability is a very important property. It enables the tracking of desired signals or the suppression of undesired signals. System stability is described either in terms of input-output stability or in terms of internal stability. The last stability concept is very important because a system may be internally unstable and yet input-out stable, i.e. it may have unstable internal modes which may not be observed at the output of the system.
Input-output Stability:
A single input, single output (SISO) linear, time-invariant (LTI), continuous time dynamic system is bounded-input, bounded-output (BIBO) stable if and only if its impulse response $h(t)$ is absolutely integrable, i.e.,

$$\int_{0}^{\infty} |h(t)| dt \leq M < \infty$$

where $M$ is a real constant.

Proof: Let the input $u(t)$ be bounded, i.e., $|u(t)| \leq k_1 < \infty$, $\forall \ t \geq 0$. Then

$$|y(t)| = \left| \int_{0}^{\infty} h(\tau)u(t-\tau) d\tau \right| \leq \int_{0}^{\infty} |h(\tau)||u(t-\tau)| d\tau \leq \int_{0}^{\infty} |h(\tau)| k_1 d\tau = k_1 \int_{0}^{\infty} |h(\tau)| d\tau \leq k_1 M < \infty$$

$\Rightarrow y(t)$ is bounded.

Suppose $h(t)$ is not absolutely integrable. Then for a causal, linear time-invariant system, with $u(t) = k_1 > 0$ and $h(t) > 0$, $t \geq 0$, with nondecreasing envelope.
the output is given by $y(t) = \int_0^t h(\tau)u(t - \tau)d\tau$

For $t \geq 0$, this implies that

$$|y(t)| = k_1\int_0^t |h(\tau)| d\tau$$

as $t \to \infty$, $\int_0^t |h(\tau)| d\tau \to \infty$

This implies that $y(t)$ is not bounded even when $u(t)$ is bounded. Thus, $h(t)$ must be absolutely integrable.

**Theorem:** A SISO LTI, continuous-time dynamic system is BIBO stable if and only if every pole of its transfer function $H(s)$ lies on the left-half of the $s$-plane.

**Proof:** Let $H(s)$ be a proper rational function of $s$, then if every pole located at $s = -p_i, \quad p_i > 0$, has multiplicity $n_i$, such that

$$\sum_{i=1}^m n_i = n$$
Then
\[ H(s) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{r_{ij}}{(s + p_i)^j} \]
and the impulse response is given by
\[
h(t) = \mathcal{L}^{-1} \{ H(s) \} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} k_{ij} \mathcal{L}^{-1} \left\{ \frac{1}{(s + p_i)^j} \right\} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{r_{ij}}{(j-1)!} t^{j-1} e^{-p_i t}
\]
t \geq 0, which is absolutely integrable.

Consider a multiple-input, multiple-output (MIMO), LTI, continuous-time dynamic system described by the impulse response matrix \( H(t) = [h_{ij}(t)] \). Such a system is BIBO stable if and only if \( \forall i, j \),
\[
\int_0^\infty |h_{ij}(t)| \, dt = K_{ij} < \infty
\]
Alternatively, a MIMO, LTI, continuous-time dynamic system described by the proper rational transfer function matrix \( H(s) = \mathcal{L}\{H(t)\} = [H_{ij}(s)] \) is BIBO stable if and only if every pole of \( H_{ij}(s) \) is located on the left half of the \( s \)-plane.
As we already know, the solution of the state equation is given by

\[ \tilde{x}(t) = e^{At} \tilde{x}(0) + \int_{0}^{t} e^{A(t-\tau)} B\tilde{u}(\tau)d\tau \]

Moreover, in the s-domain we get

\[ \tilde{X}(s) = (sI - A)^{-1} \tilde{x}(0) + (sI - A)^{-1} B\tilde{U}(s) \]

Let the input vector be identically zero and the initial state be nonzero, i.e.,

\[ \tilde{u}(t) = \tilde{0} \quad \text{or} \quad \tilde{U}(s) = \tilde{0} \]

and

\[ \tilde{x}(0) \neq \tilde{0} \]

then

\[ \tilde{X}(s) = (sI - A)^{-1} \tilde{x}(0) \]

and

\[ \tilde{x}(t) = e^{At} \tilde{x}(0) \]

Let \( H_{in}(s) \equiv (sI - A)^{-1} \) be the internal transfer function matrix of some continuous time LTI dynamic system. Then, with \( u(t) = 0 \) and for some

\[ \|\tilde{x}(0)\|_2 = k < \infty \]
we get \( \|\tilde{x}(t)\|_2 = \|e^{At}\tilde{x}(0)\|_2 < \infty, \ t \geq 0 \)

and

1. The unforced system is marginally stable if and only if the poles of \( H_{in}(s) \) (or the eigenvalues of \( A \)) have either zero or negative real parts, and those with zero real parts are simple roots of the minimal polynomial of \( A \).

2. The unforced system is asymptotically stable if and only if all the poles of \( H_{in}(s) \) (all eigenvalues of \( A \)) lie strictly on the left half of the \( s \)-plane.

Asymptotic stability also implies that

\[ \lim_{t \to \infty} \tilde{x}(t) = \tilde{0} \]

Observation: These two concepts deal with internal stability only.

Example: Let an unforced dynamic system be described by

\[
\dot{\tilde{x}}(t) = \begin{bmatrix} -2 & \frac{1}{2} \\ 2 & -\frac{1}{2} \end{bmatrix} \tilde{x}(t)
\]
then

\[
H_{in}(s) = \begin{bmatrix}
\frac{s + \frac{1}{2}}{s(s + \frac{5}{4})} & \frac{\frac{1}{2}}{s(s + \frac{5}{4})} \\
\frac{2}{s(s + \frac{5}{4})} & \frac{s + 2}{s(s + \frac{5}{4})}
\end{bmatrix},
\]

which shows that the poles of \(H_{in}(s)\) are located at \(s = 0\) and \(s = -\frac{5}{4}\). Hence, the system is marginally stable.

Suppose now that the input stimulus is nonzero and that the initial condition is zero, i.e., \(\tilde{x}(0) = \tilde{0}\) and \(\tilde{u}(t) \neq \tilde{0}\).

Then

\[
\tilde{X}(s) = (sI - A)^{-1}B\tilde{U}(s)
\]

and

\[
\tilde{x}(t) = \int_{0}^{t} e^{A(t-\tau)}B\tilde{u}(\tau)d\tau
\]

Furthermore,

\[
\tilde{Y}(s) = \left[ C(sI - A)^{-1}B + D \right] \tilde{U}(s) = H(s)\tilde{U}(s)
\]
Clearly, every pole of $H(s)$ is an eigenvalue of $A$. Thus, if every eigenvalue of $A$ has a negative real part, then all poles of $H(s)$ lie on the left-half of the $s$-plane. Therefore, the system described by

$$\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = \tilde{0} \\
\tilde{y}(t) &= C\tilde{x}(t) + D\tilde{u}(t)
\end{align*}$$

is BIBO stable. However, because of possible pole-zero cancellation, not every eigenvalue of $A$ is a pole of $H(s)$.

Hence, while the unforced system $\dot{x}(t) = Ax(t)$ may be unstable, $\otimes$ may not!

**Example:** Consider a LTI dynamic system described by the equations

$$\begin{align*}
\dot{x}_1(t) &= x_2(t) - 2u(t) \\
\dot{x}_2(t) &= x_1(t) + 2u(t) \\
y(t) &= x_2(t)
\end{align*}$$
In block diagram form,

Let \( \tilde{x}(t) = [x_1(t) \ x_2(t)]^T \)

\[
\dot{\tilde{x}}(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} -2 \\ 2 \end{bmatrix} u(t)
\]

\[
y(t) = [0 \ 1] \tilde{x}(t)
\]
Thus, \( \pi_A(\lambda) = \det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ -1 & \lambda \end{bmatrix} = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1) \) means the system is internally unstable.

The transfer function of the system, however, is given by

\[
H(s) = C(sI - A)^{-1}B = \frac{1}{s^2 - 1} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} s & 1 \\ 1 & s \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \frac{-2 + 2s}{s^2 - 1} = \frac{2(s - 1)}{(s + 1)(s - 1)} = \frac{2}{s + 1}
\]

Hence, the impulse response is

\[
h(t) = 2e^{-t}u_s(t)
\]

where \( u_s(t) = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \)

Therefore, the system is BIBO stable. However, as already pointed out, the system is internally unstable because of the eigenvalue at \( \lambda = 1 \)!
In fact,

\[ \Phi(t) = \mathcal{L}^{-1}\left\{ (sI - A)^{-1} \right\} = \mathcal{L}^{-1}\left\{ \begin{bmatrix} \frac{s}{s^2 - 1} & \frac{1}{s^2 - 1} \\ \frac{1}{s^2 - 1} & \frac{s}{s^2 - 1} \end{bmatrix} \right\} = \mathcal{L}^{-1}\left\{ \begin{bmatrix} \frac{1/2}{s + 1} + \frac{1/2}{s - 1} & \frac{1/2 - 1/2}{s - 1} \\ \frac{1/2}{s - 1} - \frac{1/2}{s + 1} & \frac{1/2 + 1/2}{s + 1} \end{bmatrix} \right\} \]

\[ = \begin{bmatrix} \frac{1}{2} (e^{-t} + e^t) & \frac{1}{2} (e^t - e^{-t}) \\ \frac{1}{2} (e^t - e^{-t}) & \frac{1}{2} (e^{-t} + e^t) \end{bmatrix} \]

**Lyapunov stability**

Consider the following LTI, continuous time, unforced dynamic system,

\[ \dot{x}(t) = Ax(t) \]

\[ A \in \mathbb{R}^{n \times n}. \] Then the system \((A)\) is asymptotically stable if every eigenvalue of \(A\) has a negative real part.
Theorem: Let \( \sigma(A) = \{\lambda_1, \ldots, \lambda_n\} \) be the spectrum of \( A \). Then \( \text{Re}\{\lambda_i\} < 0, i = 1, \ldots, n \), if and only if for any given real \( N = N^T > 0 \), the Lyapunov equation

\[
A^T M + M A = -N
\]

has a unique solution \( M = M^T > 0 \).

Consider the weighted energy function of the unforced system \( V(x(t)) = x^T(t)Mx(t) \). Then, if we take the derivative with respect to time, we get

\[
\frac{dV(x(t))}{dt} = \frac{dx^T(t)}{dt} M x(t) + x^T(t) M \frac{dx(t)}{dt}
\]

\[
= (Ax(t))^T M x(t) + x^T(t) M A x(t)
\]

\[
= x^T(t) A^T M x(t) + x^T(t) M A x(t)
\]

\[
= x^T(t) (A^T M + M A) x(t)
\]

Now, if the unforced system is asymptotically stable and \( V(x(t)) \) is a valid energy function, then

\[
\frac{dV(x(t))}{dt} < 0 , \text{ namely, for } \Delta t > 0 , \text{ } V(x(t + \Delta t)) < V(x(t)) \text{ if } V(x(t)) \text{ is a valid energy function. The only way to make } \frac{dV(x(t))}{dt} < 0 \text{ is if } \frac{dV(x(t))}{dt} = -x^T(t) N x(t) , \text{ where } N \text{ is a real symmetric positive definite matrix.}
This implies that \( x^T(t)\left(A^T M + MA\right)x(t) = -x^T(t)Nx(t) \) or 
\[
A^T M + MA = -N
\]

Finally, \( V(x(t)) = x^T(t)Mx(t) \) is a valid energy function if \( M \) is real and \( M = M^T > 0 \).

**Example:** Let the system matrix be described by 
\[
A = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-6 & -11 & -6
\end{bmatrix}
\]
then \( \sigma(A) = \{-1, -2, -3\} \), which means that the system is asymptotically stable.

Let 
\[
N = \begin{bmatrix}
1 & 0 & 0 \\
0 & 2 & 0 \\
0 & 0 & 3
\end{bmatrix}
\]

then \( N = N^T > 0 \), since \( \sigma(N) = \{1, 2, 3\} \).

Solving the Lyapunov equation, yields 
\[
M = \begin{bmatrix}
1.858 & -0.5 & -1.108 \\
-0.5 & 1.108 & -1 \\
-1.108 & -1 & 3.192
\end{bmatrix} = M^T > 0,
\]
since \( \sigma(M) = \{0.182, 2.005, 3.971\} \).

**Discrete-Time System Stability**

Consider a SISO discrete-time, causal, LTI system, then if \( h(n) \) is the unit sample response and \( u(n) \) the input applied to it,

\[
y(n) = \sum_{k=0}^{n} h(n-k)u(k) = \sum_{k=0}^{n} h(k)u(n-k)
\]

**Theorem**: Suppose \( |u(n)| \leq K < \infty, \ n = 0, 1, \ldots \). Then a SISO discrete time, causal, LTI system is BIBO stable if and only if \( h(n) \) is absolutely summable, i.e.,

\[
\sum_{n=0}^{\infty} |h(n)| \leq M < \infty
\]

**Theorem**: Let a SISO, discrete-time, causal LTI system be described by a proper rational transfer function \( H(z) = Z\{h(n)\} \), where \( Z \) is the z-transform operator. Then the system is BIBO stable if and only if every pole of \( H(z) \) has magnitude strictly less than 1 (all poles are located inside the unit circle on the \( z \)-plane).
Theorem (Internal Stability): An unforced discrete time, LTI dynamic system described by
\[ x(k + 1) = Ax(k), \quad x_0 = x(0). \]

1) Is marginally stable if and only if the eigenvalues of \( A \) have magnitude less than or equal to 1, and those equal to 1 are simple roots of the minimal polynomial of \( A \).

2) Is asymptotically stable if and only if all eigenvalues of \( A \) have magnitude less than 1.

Lyapunov Stability Theorem: All eigenvalues of \( A \in \mathbb{R}^{nxn} \) have magnitude less than 1 if and only if for any given \( N = N^T > 0 \) the discrete Lyapunov equation
\[ M - A^TMA = N \]
has a unique solution \( M = M^T > 0 \).

Example: Consider a discrete-time, LTI dynamic system described by
\[
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-0.005 & 0.01 & 0.5
\end{bmatrix}
\]
\[ x(k + 1) = A x(k), \quad x(0) = x_0. \]
Then $\sigma(A) = \{-0.1, 0.1, 0.5\}$. Furthermore, if $N = \text{diag}(1, 2, 3)$, then the discrete Lyapunov equation has the solution

$$ M = \begin{bmatrix} 7.021 & 2.025 & 1.063 \\ 2.025 & 6.021 & 2.025 \\ 1.063 & 2.025 & 4.021 \end{bmatrix} = M^T $$

Now, $\sigma(M) = \{2.753, 4.872, 9.437\}$ implies that $M > 0$. Thus, all eigenvalues of $A$ must lie inside the unit circle (which we already knew it).

**Def.** A state $x_0 = x(t_0) \in \mathbb{R}^n$ is controllable over $[t_0, t_1]$ if there exists an input $u(\bullet)$ defined over $[t_0, t_1]$ such that

$$ x(t_1) = 0 = \Phi(t_1 - t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1 - \tau)Bu(\tau)d\tau $$

**Def.** The state model $dx/dt = Ax + Bu$ is said to be controllable (completely controllable) if and only if every state $x_0 \in \mathbb{R}^n$ is controllable.
Def. The set of all controllable states is the controllable subspace.

Proposition: The controllable subspace is a linear subspace, i.e., let \( V \subseteq \mathbb{R}^n \) be the controllable subspace, then if \( x_1 \) and \( x_2 \in V \), the vector \( x = \alpha_1 x_1 + \alpha_2 x_2 \in V \) for any \( \alpha_1, \alpha_2 \in \mathbb{R} \).

Lemma: For any integer \( p > n \), \( \text{rank } [B \ AB \ldots A^{p-1}B] = \text{rank } [B \ AB \ldots A^{n-1}B] \).

Corollary: Let \( Q = [B \ AB \ldots A^{n-1}B] \). Then for any \( x \in \text{Col-sp } [Q] \), \( Ax \in \text{Col-sp } [Q] \), i.e., the column space of \( Q \) is \( A \)-invariant.

Let \( \text{rank}[Q] = p < n \), \( U_1 \) be an \( n \times p \) matrix whose columns form a basis for the column space of \( Q \) and let \( U_2 \) be an \( n \times (n-p) \) matrix whose columns together with those of \( U_1 \) form a basis for \( \mathbb{R}^n \), i.e., \( \text{Col-sp } [U_1 \mid U_2] = \mathbb{R}^n \).

Proposition: Given \( \frac{dx}{dt} = Ax + Bu \), the state transformation \([U_1 \ U_2]z = x\), yields

\[
\begin{bmatrix}
\dot{z}_1 \\
\dot{z}_2
\end{bmatrix} =
\begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & \tilde{A}_{22}
\end{bmatrix}
\begin{bmatrix}
z_1 \\
z_2
\end{bmatrix} +
\begin{bmatrix}
B_1 \\
0
\end{bmatrix} u
\]
where \( \text{rank } [B_1 \quad \tilde{A}_{11} B_1 \quad \cdots \quad \tilde{A}_{11}^{p-1} B_1] = p \), and eq. \( \otimes \) is known as the Kalman controllable form.

**Proof:** \( [U_1 \quad U_2] \dot{z} = \dot{x} = Ax + Bu = A[U_1 \quad U_2]z + Bu = [AU_1 \quad AU_2]z + Bu \)

Now, \( AU_1 \in \text{Col-sp } [Q] \) and \( AU_1 = [Au_1 \quad Au_2 \quad \cdots \quad Au_p] \) implies that each column of \( AU_1 \) is a linear combination of the columns of \( U_1 \) or \( AU_1 = U_1 \tilde{A}_{11} \) for an appropriate \( \tilde{A}_{11} \in R^{p \times p} \). Each column in \( AU_2 \) is in \( R^n \) since the columns of \( [U_1 \quad U_2] \) are a basis in \( R^n \), therefore, there exists matrices \( \tilde{A}_{12} \) and \( \tilde{A}_{22} \) such that

\[
AU_2 = [U_1 \quad U_2] \begin{bmatrix}
\tilde{A}_{12} \\
\tilde{A}_{22}
\end{bmatrix} = U_1 \tilde{A}_{12} + U_2 \tilde{A}_{22}
\]

Thus, \( [U_1 \quad U_2] \dot{z} = [U_1 \quad U_2] \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & \tilde{A}_{22}
\end{bmatrix} z + Bu \)

or \( \dot{z} = \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & \tilde{A}_{22}
\end{bmatrix} z + [U_1 \quad U_2]^{-1} Bu \)
By construction, \( \text{Col-sp} \, [Q] = \text{Col-sp} \, [B \, AB \, ... \, A^{n-1}B] \supseteq \text{Col-sp} \, [B] \), thus, there exists a \( p \times m \) matrix \( B_1 \) such that \( B = U_1B_1 \) or

\[
B = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} U_1 & U_2 \end{bmatrix}^{-1} B = \begin{bmatrix} U_1 & U_2 \end{bmatrix}^{-1} \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}
\]

or

\[
\dot{z} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix} z + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u
\]

Lemma: The controllability Grammian matrix

\[
K = \int_{0}^{t_1} e^{-\tilde{A}_{11} \tau} B_1 B_1^T e^{-\tilde{A}_{11}^T \tau} d\tau
\]

is nonsingular whenever \( \text{rank} \begin{bmatrix} \tilde{Q} \end{bmatrix} = p \),

where \( \tilde{Q} = [B_1 \quad \tilde{A}_{11} B_1 \quad \cdots \quad \tilde{A}_{11}^{p-1} B_1] \)
Proof: Suppose that rank \( [\tilde{Q}] = p \) and that there exists \( \mathbf{v} \neq \mathbf{0}, \mathbf{v} \in R^p \) such that \( \mathbf{v}^T \mathbf{K} = \mathbf{0}^T \), then

\[
\mathbf{v}^T \mathbf{K} \mathbf{v} = 0 = \int_{0}^{t_1} \mathbf{v}^T e^{-\tilde{A}_1 \tau} \mathbf{B}_1 \mathbf{B}_1^T e^{-\tilde{A}_1^T \tau} \mathbf{v} \, d\tau
\]

Let \( \mathbf{c}(\tau) \equiv \mathbf{B}_1^T e^{-\tilde{A}_1 \tau} \mathbf{v} \equiv [c_1(\tau) \ldots c_p(\tau)]^T \), then

\[
\mathbf{v}^T \mathbf{K} \mathbf{v} = \int_{0}^{t_1} \mathbf{c}^T(\tau) \mathbf{c}(\tau) \, d\tau = \int_{0}^{t_1} \left[ c_1^2(\tau) + \cdots + c_p^2(\tau) \right] \, d\tau = 0
\]

if and only if \( c_i(\tau) \equiv 0 \) for \( \tau \in [0, t_1] \).

Now, if \( c_i(\tau) = 0 \) for \( \tau \in [0, t_1] \) then \( \frac{d}{d\tau} j c^T(\tau) = \mathbf{0}, \quad \forall j \)

Clearly, for \( j = 0 \) and \( \tau = 0 \), \( \mathbf{c}^T(0) = \mathbf{v}^T \mathbf{B}_1 = \mathbf{0}^T \).

For \( j = 1 \) and \( \tau = 0 \),

\[
\left. \frac{d}{d\tau} c^T(\tau) \right|_{\tau=0} = \mathbf{v}^T \mathbf{e}^{-\tilde{A}_1 \tau} (-\tilde{A}_{11}) \mathbf{B}_1 \bigg|_{\tau=0} = -\mathbf{v}^T \tilde{A}_{11} \mathbf{B}_1 = \mathbf{0}^T
\]
For $j \geq 1$

$$\frac{d^j c^T(\tau)}{d\tau^j}\bigg|_{\tau=0} = (-1)^j v^T \tilde{A}_{11}^j B_1 = 0^T$$

Thus,

$$v^T \tilde{Q} = v^T [B_1 \tilde{A}_{11} B_1 \ldots \tilde{A}_{11}^{p-1} B_1] = \begin{bmatrix} v^T B_1 & v^T \tilde{A}_{11} B_1 & \ldots & v^T \tilde{A}_{11}^{p-1} B_1 \end{bmatrix} = 0^T$$

This implies that rank $[\tilde{Q}] \neq p$, which contradicts the hypothesis that rank $[\tilde{Q}] = p$, hence, $K$ is nonsingular.

**Example**: Consider the dynamic system described by

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

Clearly, the system is not controllable, as the input $u$ and $x_2$ do not affect $x_1$. Now,

$$Q = [B \ AB] = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$$

rank $[Q] = 1 \neq 2 \Rightarrow$ system is not controllable.

Let $U_1 = [0 \ 1]^T$. Since $[0 \ 1]^T$ spans Col-sp $[Q]$, then, if $U_2 = [1 \ 1]^T$, the columns of $[U_1 \ U_2]$ span $R^2$. 
Now, 
\[
[U_1 \quad U_2]^{-1} A [U_1 \quad U_2] = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}
\]
and 
\[
[U_1 U_2]^{-1} B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.
\]

Therefore, the equivalent system in the Kalman controllable canonical form is described by
\[
\dot{z} = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u
\]

**Theorem:** The following statements about controllability are equivalent:

1. The pair \((A, B)\) is controllable
2. Rank \([\lambda_i I - A \mid B] = n\) for each eigenvalue \(\lambda_i\) of \(A\)
3. Rank \([Q] = n\)
4. Rank \([e^{-At}B] = n\), i.e., there are \(n\) linearly independent row functions of \(e^{-At}B\) for \(t \in [0, \infty)\)
5. The matrix \( \hat{K} = \int_{t_0}^{t_1} e^{A(t_1-\tau)} BB^T e^{A^T(t_1-\tau)} \, d\tau \) is positive definite. Furthermore, the input \( u(t) = -B^T e^{A^T(t_1-t)} \hat{K}^{-1} [e^{A(t_1-t_0)} x(t_0) - x(t_1)] \) transfers \( x(t_0) \) to \( x(t_1) \).

Proof: \( 1 \Rightarrow 3 \), since \((A, B)\) is controllable \( \Rightarrow \) rank \([Q] = \text{rank} \begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix} = n \).

Suppose there exists \( \lambda_i \) such that rank \([\lambda_i I - A] \mid B] < n\), then there exists a \( v \in \mathbb{R}^n \), \( v \neq 0 \), such that, \( v^T[(\lambda_i I - A) \mid B] = [v^T(\lambda_i I - A) \mid v^T B] = 0^T \Rightarrow v^T(\lambda_i I - A) = 0^T \) and \( v^T B = 0^T \).

But, \( v^T(\lambda_i I - A) = 0^T \Rightarrow v^T = w_i^* \), a left eigenvector of \( A \) associated with \( \lambda_i \).

Now,
\[
v^T A^k = w_i^* A^k = (w_i^* A) A^{k-1} = \lambda_i w_i^* A^{k-1} = \lambda_i (w_i^* A) A^{k-2} = \lambda_i^2 w_i^* A^{k-2} = \cdots = \lambda_i^k w_i^*
\]

\( \therefore v^T Q = [v^T B \ v^T AB \ldots \ v^T A^{n-1}B] = [v^T B \ \lambda v^T B \ldots \lambda^{n-1} v^T B] = [0^T \ 0^T \ldots \ 0^T] = 0^T \)

Hence, rank \([Q] < n\), which contradicts the hypothesis that rank \([Q] = n\), hence, \( 3 \Rightarrow 2 \). The proof that \( 2 \Rightarrow 3 \) basically uses the same type of arguments.
Assume that \( \text{rank } [e^{-At}B] \neq n \), i.e., there exists \( v \neq 0 \) such that \( v^T e^{-At}B = 0^T \).

But, 
\[
\frac{d^k}{dt^k} (v^T e^{-At} B) \bigg|_{t=0} = (-1)^k v^T A^k B = \tilde{0}^T, \ \forall \ k \geq 0 \Rightarrow v^T Q = 0^T
\]
which contradicts the assumption that \( \text{rank } [Q] = n \). Thus, \( \text{rank } [Q] = n \Rightarrow \text{rank } [e^{-At}B] = n \), i.e., \( e^{-At}B \) has \( n \) linearly independent rows.

Suppose \( \hat{K} \) is not positive definite. From the previous lemma, \( \hat{K} \) is also not negative definite, thus, \( \hat{K} \) can only be singular (positive semi-definite). Therefore, there exists \( v \neq 0 \) such that \( v^T \hat{K} = 0^T \Rightarrow v^T \hat{K} v = 0 \) or
\[
0 = v^T \hat{K} v = \int_{t_0}^{t_1} v^T e^{A(t_1 - \tau)} B B^T e^{A^T(t_1 - \tau)} v d\tau
\]
let \( \alpha = t_1 - \tau \), then
\[
v^T \hat{K} v = \int_{0}^{t_1-t_0} v^T e^{A\alpha} B B^T e^{A^T\alpha} v d\alpha = 0
\]
This implies that the integrand \( \equiv 0 \). \( \Rightarrow v^T e^{A\alpha} B = 0^T \Rightarrow \) rows of \( e^{At}B \) are linearly dependent which contradicts that \( \text{rank}[e^{At}B] = n \). Therefore, \( \text{rank } [e^{At}B] = n \Rightarrow \hat{K} \) is positive definite.
Finally,

\[ x(t_1) = e^{A(t_1-t_0)} x(t_0) + \int_{t_0}^{t_1} e^{A(t_1-\tau)} Bu(\tau) d\tau \]

\[ = e^{A(t_1-t_0)} x(t_0) - \int_{t_0}^{t_1} e^{A(t_1-\tau)} BB^T e^{A^T(t_1-\tau)} \hat{K}^{-1} \left[ e^{A(t_1-t_0)} x(t_0) - x(t_1) \right] d\tau \]

\[ = e^{A(t_1-t_0)} x(t_0) - \int_{t_0}^{t_1} e^{A(t_1-\tau)} BB^T e^{A^T(t_1-\tau)} d\tau \hat{K}^{-1} \left[ e^{A(t_1-t_0)} x(t_0) - x(t_1) \right] \]

\[ = e^{A(t_1-t_0)} x(t_0) - \hat{K} \hat{K}^{-1} \left[ e^{A(t_1-t_0)} x(t_0) - x(t_1) \right] = x(t_1) \]

**Example:** Consider the following linear time-invariant dynamical system

\[
\dot{x} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} x + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} u
\]

then,

\[
Q = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & -2 & -2 \end{bmatrix}
\]
Hence, rank $[Q] = 3 = n$ (because columns 1, 3 and 5 are linearly independent). Moreover, the spectrum of $A$ is $\sigma(A) = \{-0.4302, -0.7849 \pm j 1.3071\} = \{\lambda_1, \lambda_2, \lambda_3\}$, which means that

$$\text{rank } [(\lambda_1 I - A) \mid B] = \text{rank } \begin{bmatrix} 0.5698 & 0 & 1 & 1 & 1 \\ 0 & -0.4302 & -1 & 0 & 0 \\ -1 & 1 & 0.5698 & 0 & 0 \end{bmatrix} = 3$$

since the third column is a linear combination of the first and second columns. Likewise, $\text{rank } [(\lambda_i I - A) \mid B] = 3$, $i = 2, 3$.

**Discrete-Time Systems**

Let $x(k+1) = Ax(k) + Bu(k)$, $A \in R^{n \times n}$ and $B \in R^{n \times m}$.

**Def.** A state $x_c \in R^n$ is controllable (reachable) if and only if there exists a finite $N \in N$ and an input sequence $\{u(0), u(1), \ldots, u(N-1)\}$ such that if $x(0) = 0$, then $x(N) = x_c$. 
Example: Consider the following linear shift-invariant discrete-time dynamic system

\[ x(k + 1) = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) \]

then if \( x(0) = [0 \ 0]^T \) and \( u(0) = \alpha \), \( x(1) = [0 \ \alpha]^T \), is a controllable state (because it can be reached in one step when \( N = 1 \)).

Now, \( Q = [B \ AB] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \)

\[ \Rightarrow \text{rank} \ [Q] = 1! \text{ i.e., system is not completely controllable. However, the subspace described by the basis vector } [0 \ 1]^T \text{ is controllable.} \]

Theorem: The following statements about discrete-time controllability are equivalent:

1. There exists a finite index \( N \leq n \) such that \( x(k) = 0 \) can be driven to \( x_c = x(k+N) \) by some input sequence \( \{u(k), u(k+1), \ldots, u(k+N-1)\} \).

2. \( x_c \in \text{Col-sp} \ [Q] \), \( Q = [B \ AB \ldots A^{n-1}B] \)
3. If \( x_{c1} \) and \( x_{c2} \in \text{Col-sp} [Q] \), then there exists \( N \) such that \( x_{c1}(k) \) can be driven to \( x_{c2}(k+N) \) by some input sequence \( \{u(k), u(k+1), \ldots, u(k+N-1)\} \).

### Pole Placement By State Feedback

Consider the linear time-invariant continuous time dynamic system described by

\[
\dot{x} = Ax + Bu
\]

Let \( u = u_c + u_r \), where \( u_c \) is the control signal and \( u_r \) is the external reference signal.

Let \( u_c = Fx \) be the state feedback control signal, then \( \dot{x} = (A + BF)x + Bu_r \) is the new closed-loop system. Given a pre-specified symmetric spectrum (set of complex numbers) \( \Lambda \), construct the feedback matrix \( F \) such that \( \sigma(A + BF) = \Lambda \). If the system under consideration is in the controllable canonical form, i.e.,

\[
\dot{x} = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
-\alpha_n & \cdots & -\alpha_1 & \cdots & 0
\end{bmatrix} x + \begin{bmatrix}
0 \\
0 \\
\vdots \\
1
\end{bmatrix} u
\]
Then the characteristic polynomial is given by

\[ \pi_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \]

Now, if \( u_c = Fx = [f_n \ldots f_1]x \), then

\[
\dot{x} = (A + BF)x + Bu_r = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & & \\
& & \ddots & \ddots & \\
-(a_n - f_n) & \cdots & -(a_1 - f_1)
\end{bmatrix} \begin{bmatrix}
x_0 \\
x_1 \\
\vdots \\
x_n
\end{bmatrix} + \begin{bmatrix}
0 \\
0 \\
\vdots \\
1
\end{bmatrix} u_r
\]

and the closed-loop characteristic polynomial is given by

\[ \pi_{A+BF}(\lambda) = \lambda^n + (a_1 - f_1)\lambda^{n-1} + \cdots + (a_{n-1} - f_{n-1})\lambda + (a_n - f_n) \]

Let the desired spectrum be given by \( \sigma(A + BF) = \Lambda = \{\lambda_{1d}, \lambda_{2d}, \ldots, \lambda_{nd}\} \), then choose \( F \) such that

\[ \pi_{A+BF}(\lambda) = \prod_{i=1}^{n} (\lambda - \lambda_{id}) \]
Now,

\[
\pi_{A+BF}(\lambda) = \prod_{i=1}^{n} (\lambda - \lambda_{id}) = \lambda^n + a_{1,d} \lambda^{n-1} + \cdots + a_{n-1,d} \lambda + a_{n,d}.
\]

\[
= \lambda^n + (a_1 - f_1) \lambda^{n-1} + \cdots + (a_{n-1} - f_{n-1}) \lambda + (a_n - f_n)
\]

Then, we equate coefficients of equal powers, i.e.

\[
a_{i,d} = a_i - f_i, \quad i = 1, 2, \ldots, n.
\]

to compute the feedback gains

\[
f_i = a_i - a_{i,d}, \quad i = 1, 2, \ldots, n.
\]

We select the desired closed-loop eigenvalues \( \lambda_{i,d}, i = 1, 2, \ldots, n \) according to the design performance criteria, which can be speed of response of the system, amount of control effort, etc.

When the state model of the physical system is not in the controllable canonical form, but is completely controllable, we can transform it into the controllable canonical form using a similarity transformation.
General Single Input Case

\((A, B)\) controllable \(\Rightarrow \) rank \([Q] = n, \ Q \in \mathbb{R}^{n \times n} \Rightarrow Q^{-1}\) exists. Let \(v\) be the last row of \(Q^{-1}\) and \(z = Vx\), where the similarity transformation \(V\) is given by

\[
V = \begin{bmatrix}
v \\
vA \\
\vdots \\
vA^{n-1}
\end{bmatrix}.
\]

Then, the new system is given by

\[
\dot{z} = VAV^{-1}z + VBu = A_cz + B_cu
\]

Claim 1: \(VB = B_c = [0 \ 0 \ \ldots \ 0 \ 1]^T\).

Proof:

\[
VB = \begin{bmatrix}
v \\
vA \\
\vdots \\
vA^{n-1}
\end{bmatrix}B = \begin{bmatrix}
vB \\
vAB \\
\vdots \\
vA^{n-1}B
\end{bmatrix}
\]
But,

\[ Q^{-1}Q = I \implies vQ = v[B \ AB \ \cdots \ A^{n-1}B] = [vB \ vAB \ \cdots \ vA^{n-1}B] = [0 \ \cdots \ 0 \ 1] \]

since \( v \) is the last row of \( Q^{-1} \). Therefore, \( VB = B_c = [0 \ \cdots \ 0 \ 1]^T \).

Claim 2: \( VAV^{-1} = A_c = \)

\[
\begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \vdots & \\
\vdots & \ddots & \ddots & 1 \\
-a_n & \cdots & -a_1
\end{bmatrix}
\]

Proof: \( VAV^{-1} = A_c \implies VA = A_cV \). Now,

\[
VA = \begin{bmatrix} v \\ vA \\ \vdots \\ vA^{n-1} \end{bmatrix} \quad A = \begin{bmatrix} vA \\ vA^2 \\ \vdots \\ vA^n \end{bmatrix} = A_cV = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \vdots & \\
\vdots & \ddots & \ddots & 1 \\
-a_n & \cdots & -a_1
\end{bmatrix} V
\]
Now,

\[
A_c V = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
-\alpha_n & -\alpha_{n-1} & \cdots & -\alpha_1 & 1 \\
\end{bmatrix}
\begin{bmatrix}
v \\
vA \\
vA^2 \\
\vdots \\
vA^{n-1} \\
\end{bmatrix} = \begin{bmatrix}
vA \\
vA^2 \\
-\alpha_n v - \alpha_{n-1} vA - \cdots - \alpha_1 vA^{n-1} \\
\end{bmatrix}
\]

From Cayley-Hamilton,

\[
\pi_A(A) = A^n + a_1 A^{n-1} + \ldots + a_1 A + a_n I = 0.
\]

\[
\Rightarrow A^n = -a_n I - a_{n-1} A - \cdots - a_1 A^{n-1} \Rightarrow vA^n = -a_n v - a_{n-1} vA - \cdots - a_1 vA^{n-1}.
\]

Hence, the right-hand side is the same as the left-hand side.

Clearly, the above state transformation \( z = Vx \) changes \( A \) and \( B \) into \( A_c \) and \( B_c \), where \( A_c \) and \( B_c \) are in the controllable canonical form.
Design Procedure:

Transform \((A, B)\) into \((A_c, B_c)\)

Assign \(\Lambda_d = \{\lambda_{1d}, \lambda_{2d}, \ldots, \lambda_{nd}\}\) by \(F_c\), i.e., \(\sigma(A_c + B_c F_c) = \Lambda\)

Find \(F\) so that \(\sigma(A + BF) = \Lambda\), i.e., \(A + BF = V^{-1}(A_c + B_c F_c)V = A + BF_c V\)

\(\Rightarrow F = F_c V\) is the feedback gain matrix in the physical domain.

Example: Consider the system described by

\[
\dot{x} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -4 & 2 \\ 0 & 0 & -10 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} u
\]

Let the desired spectrum be \(\Lambda = \{-5, -20, -25\}\).

Now,

\[
Q = [B \ AB \ A^2B] = \begin{bmatrix} 1 & -1 & -1 \\ 0 & -2 & 28 \\ -1 & 10 & -100 \end{bmatrix}
\]
Since \( Q \) is of full rank, we can compute its inverse, i.e.

\[
Q^{-1} = \begin{bmatrix}
1.60 & 2.20 & 0.60 \\
0.56 & 2.02 & 0.56 \\
0.04 & 0.18 & 0.04 \\
\end{bmatrix}
\]

Moreover,

\[
V = \begin{bmatrix}
v \\
vA \\
vA^2 \\
\end{bmatrix} = \begin{bmatrix}
0.04 & 0.18 & 0.04 \\
-0.04 & -0.68 & -0.04 \\
0.04 & 2.68 & -0.96 \\
\end{bmatrix}
\]

Hence,

\[
A_c = VAV^{-1} = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-40 & -54 & -15 \\
\end{bmatrix}
\]

\[
\Rightarrow \pi_{A_c}(\lambda) = \pi_A(\lambda) = \lambda^3 + 15\lambda^2 + 54\lambda + 40
\]

and the spectrum of \( A \) is given by \( \sigma(A) = \{-1, -4, -10\} \).

Let \( F_c = [f_{c3} \quad f_{c2} \quad f_{c1}] \).
Then the closed-loop system characteristic polynomial is given by

\[ \pi_{A_c + B_c F_c}(\lambda) = \lambda^3 + (15 - f_{c1})\lambda^2 + (54 - f_{c2})\lambda + (40 - f_{c3}) = (\lambda + 5)(\lambda + 20)(\lambda + 25) \]

\[ = \lambda^3 + 50\lambda^2 + 725\lambda + 2500 \]

or \( 15 - f_{c1} = 50 \Rightarrow f_{c1} = -35; \ 54 - f_{c2} = 725 \Rightarrow f_{c2} = -671; \ 40 - f_{c3} = 2500 \)

\[ \Rightarrow f_{c3} = -2460. \]

Thus, \( F_c = [-2460 -671 -35] \) and \( F = F_c V = [-72.96 -80.32 -37.96] \).

Finally, the closed-loop system is given by,

\[ \dot{x} = (A + BF)x + Bu_r = \begin{bmatrix} -73.96 & -79.32 & -37.96 \\ 0 & -4 & 2 \\ 72.96 & 80.32 & 27.96 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} u_r \]

and its spectrum is \( \sigma(A+BF) = \{-5, -20, -25\} \).
Observability Revisited:
Consider the dynamic system described by

\[ \begin{align*}
\dot{x} &= Ax + Bu \\
y &= Cx + Du
\end{align*} \]

Def. \( x(t_0) \in \mathbb{R}^n \) is observable if and only if there exists \( t_1 > t_0 \) such that the knowledge of \( y(t) \) and \( u(t) \) over \([t_0, t_1]\) and of the system matrices \( A, B, C \) and \( D \) is sufficient to determine \( x(t_0) \).

Def. The pair \((C, A)\) in eq. \( \otimes \) is (completely) observable if and only if every \( x \in \mathbb{R}^n \) is observable.

Theorem: For the system described by \( \otimes \), the following are equivalent:

1. The pair \((C, A)\) is observable

2. \( \text{Rank} \left[ \begin{array}{c} C \\ \lambda_i I - A \end{array} \right] = n \), for each eigenvalue \( \lambda_i \) of \( A \)
3. Rank $[\varphi] = n$, where $\varphi$ is the observability matrix $\varphi = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$

4. Rank $[Ce^{At}] = n$, i.e., $Ce^{At}$ has $n$ linearly independent columns each of which is a vector-valued function of time defined over $[0, \infty)$

5. The observability Grammian matrix \( W_0(t_0, t_1) = \int_{t_0}^{t_1} e^{A^T \tau} C^T Ce^{A\tau} d\tau \) is nonsingular $\forall t_1 > t_0$.

Example: Consider the linearized model of an orbital satellite described by

\[
\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.75 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} u
\]
The observability matrix for this system is given by

\[
y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x
\]

Clearly, \( \text{rank } [\varphi] = 4 \). Moreover, \( \sigma(A) = \{0, 0, j0.5, -j0.5\} \).

For \( \lambda_1 = \lambda_2 = 0 \),\( \text{rank } \begin{bmatrix} C \\ \lambda_i I - A \end{bmatrix} = 4 \),
Since

\[
\begin{bmatrix}
C & \lambda_i I - A
\end{bmatrix} = \begin{bmatrix}
C & -A
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 \\
-0.75 & 0 & 0 & -1 \\
0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0
\end{bmatrix}
\]

Def. Suppose rank [\varphi] < n, then the null space of \varphi, N[\varphi] is the unobservable subspace of the state space. Therefore, if \(x_0 \in N[\varphi]\) then we cannot reconstruct \(x_0\) from input-output measurements.

Proposition: Let \(x_0 \in N[\varphi]\), then \(Ax_0 \in N[\varphi]\). (The unobservable subspace is \(A\) invariant).
Proof: If $x_0 \in N[\phi]$, then $\phi x_0 = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_0 = 0 \Rightarrow CAx_0 = \ldots = CA^{n-1}x_0 = 0$

Now, $\phi A x_0 = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^n \end{bmatrix} x_0 = \begin{bmatrix} CAx_0 \\ CA^2x_0 \\ \vdots \\ CA^nx_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

But, $A^n = -a_1 A^{n-1} - \ldots - a_n I \Rightarrow CA^n x_0 = -a_1 CA^{n-1} x_0 - \ldots - a_n C x_0 = 0 \Rightarrow Ax_0 \in N[\phi]$.

Proposition: Let $x_0 \in \text{Im}[\phi^T]$, then $x_0$ can be reconstructed with certainty from input-output measurements and $\text{Im}[\phi^T]$ is called the observable subspace of the state space.

Proposition: Each $x \in R^n$ has the decomposition $x = x_{ob} + x_{unob}$, where $x_{ob} \in \text{Im}[\phi^T]$ and $x_{unob} \in N[\phi]$. 

Corollary: \( x \in \mathbb{R}^n \) is observable if and only if \( x_{\text{unob}} = 0 \).

Let the matrix \( T_1 \) form a basis for the observable subspace and \( T_2 \) be a matrix whose columns together with those of \( T_1 \) form a basis for \( \mathbb{R}^n \). Then \( z = Tx \),

\[
T = \begin{bmatrix}
T_1^T \\
\vdots \\
T_2^T
\end{bmatrix}
\]

results in the Kalman observable form

\[
\begin{bmatrix}
\dot{z}_o \\
\dot{z}_{uo}
\end{bmatrix}
= \begin{bmatrix}
\bar{A}_o & 0 \\
\bar{A}_{21} & \bar{A}_{uo}
\end{bmatrix}
\begin{bmatrix}
z_o \\
z_{uo}
\end{bmatrix}
+ \begin{bmatrix}
\bar{B}_o \\
\bar{B}_{uo}
\end{bmatrix} u
\]

\[
y = \begin{bmatrix}
\bar{C}_o & 0
\end{bmatrix}
\begin{bmatrix}
z_o \\
z_{uo}
\end{bmatrix}
\]

where \( z_o \) is observable and \( z_{uo} \) is unobservable, i.e., the pair \((\bar{C}_o, \bar{A}_o)\) is observable.
Example: Consider the dynamic system described by

\[
\begin{bmatrix}
2 & 0 & 1 \\
0 & -2 & 0 \\
1 & 0 & 2 \\
\end{bmatrix} \dot{x} = \begin{bmatrix} 0 \\
1 \\
1 \end{bmatrix} u
\]

\[
y = \begin{bmatrix} 1 & 0 & 1 \\
1 & 0 & -1 \\
\end{bmatrix} x
\]

The observability matrix is given by

\[
\varphi = \begin{bmatrix} 1 & 0 & 1 \\
1 & 0 & -1 \\
3 & 0 & 3 \\
1 & 0 & -1 \\
9 & 0 & 9 \\
1 & 0 & -1 \\
\end{bmatrix}
\]

Clearly, the rank of the observability matrix is 2 \(\Rightarrow\) system is not observable.
Let the observable subspace be described by $T_1 = C^T$ and $T_2 = [1 \ 1 \ 1]^T$. Then, the similarity transformation $T$ is given by:

$$T = \begin{bmatrix}
1 & 0 & 1 \\
1 & 0 & -1 \\
1 & 1 & 1
\end{bmatrix}.$$ 

Application of this similarity transformation results in the equivalent system

$$\begin{bmatrix}
\dot{z}_o \\
\dot{z}_{uo}
\end{bmatrix} = \begin{bmatrix}
3 & 0 & \vdots & 0 \\
0 & 1 & \vdots & 0 \\
5 & 0 & \vdots & -2
\end{bmatrix} \begin{bmatrix}
z_o \\
z_{uo}
\end{bmatrix} + \begin{bmatrix}
0.5 \\
1 \\
-0.5
\end{bmatrix} u$$

where $I_2$ is a 2 by 2 identity matrix and $z_0 \in \mathbb{R}^2$. 

$$y = \begin{bmatrix}
I_2 & \vdots & 0
\end{bmatrix} \begin{bmatrix}
z_o \\
z_{uo}
\end{bmatrix},$$
Theorem: The model \((A, B, C)\) is completely controllable (observable) if and only if \((A^T, C^T, B^T)\) is completely observable (controllable).

Proof: If \((A, B, C)\) is completely controllable, then \(\text{rank } [Q] = \text{rank } [B \ AB \ldots A^{n-1}B] = n\).

But,
\[
\text{rank } \begin{bmatrix} B^T \\ B^T A^T \\ \vdots \\ B^T (A^T)^{n-1} \end{bmatrix} = \text{rank } [Q^T] = n \Rightarrow (A^T, C^T, B^T) \text{ is completely observable.}
If \((A, B, C)\) is completely observable, then \(\text{rank } [\varphi] = \text{rank } \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n\).

Now, \(\text{rank } [C^T A^T C^T \ldots (A^T)^{n-1} C^T] = \text{rank } [\varphi^T] = n \Rightarrow (A^T, C^T, B^T)\) is completely controllable.

Suppose now that \((A^T, C^T, B^T)\) is completely observable, then \(\text{rank } [\varphi_1] \) is equal to \(\text{rank } \begin{bmatrix} B^T \\ B^T A^T \\ \vdots \\ B^T (A^T)^{n-1} \end{bmatrix} = \text{rank } [\varphi_1] = n\). Now, \(\text{rank } [B \ AB \ldots A^{n-1}B] = \text{rank } [\varphi_1^T] = n\) implies that \((A, B, C)\) is completely controllable.

Suppose now that \((A^T, C^T, B^T)\) is completely controllable, then \(\text{rank } [Q_1] = \text{rank } [C^T A^T C^T \ldots (A^T)^{n-1} C^T] = n\).
But, $\text{rank}[Q_i^T] = \text{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n \Rightarrow (A, B, C)$ is completely observable.

**Dynamic Observer Design (Deterministic Case)**

Consider the scalar state model

$$\dot{x} = \alpha x + \beta u$$

$$y = \xi x$$

Suppose also that the scalars $\alpha, \beta, \xi, u$ and $y$ are known and that $\xi = 1$. We would like to construct an estimate $x_e$ such that $x_e(t) \to x(t)$ as $t$ increases.

Let $\dot{x}_e \equiv \alpha x_e + \beta u$ and $\epsilon \equiv x - x_e$, then $\dot{\epsilon} = \alpha \epsilon \Rightarrow \epsilon(t) = e^{\alpha t} \epsilon(0)$

If $\epsilon(0) = 0$, i.e., $x_e(0) = x(0)$, then $x_e(t) = x(t) \forall t$ regardless of $\alpha$. If, on the other hand, $\epsilon(0) \neq 0$, then depending on $\alpha$, $|\epsilon(t)|$ may or may not go to zero as $t$ increases, i.e., $x_e(t)$ may or may not converge asymptotically to $x(t)$.
More realistically, however, $\xi \neq 1$ or $y \propto x$ (measurement is proportional to the state).

Let us modify the estimator model by including a measurement error, i.e.

$$\dot{x}_e = \alpha x_e + k \left( y - \xi x_e \right) + \beta u$$

then, if $\varepsilon = x - x_e$ we get

$$\dot{\varepsilon} = \dot{x} - \dot{x}_e = \alpha \varepsilon - k(y - \xi x_e) = \alpha \varepsilon - k\xi(x - x_e)$$

$$= (\alpha - k\xi)\varepsilon \Rightarrow \varepsilon(t) = e^{(\alpha-k\xi)t}\varepsilon(0)$$

$\Rightarrow \varepsilon(t) \to 0$ as $t \to \infty$ if $k$ is chosen properly, i.e. if $\alpha - k\xi < 0$

Suppose that $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, and that the system dynamics are described by

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

then if the pair $(C, A)$ is observable, we can asymptotically reconstruct $x$ with the observer (dynamic state estimator)

$$\dot{x}_e = Ax_e + K(y - y_e) + Bu = Ax_e + K(y - Cx_e) + Bu = (A - KC)x_e + Ky + Bu,$$

where $K \in \mathbb{R}^{n \times p}$. 
Schematically, this is depicted as follows:

with this choice of observer structure, \([x(t) - x_e(t)] = e^{(A-KC)t} [x(0) - x_e(0)]\) and in order for \(x_e \to x\) as \(t\) increases, the feedback gain matrix \(K\) must be chosen in such a way that the eigenvalues of \(A - KC\) lie strictly on the left half of the \(s\)-plane.
Theorem: The pair \((C, A)\) is completely observable \(\Rightarrow \sigma(A - KC)\) can be arbitrarily assigned by the proper choice of \(K\).

Proof: \((C, A)\) observable \(\Rightarrow (A^T, C^T)\) controllable (by duality) \(\Rightarrow \sigma(A^T + C^T K_1)\) can be arbitrarily assigned by properly choosing \(K_1\). Let \(K = -K_1^T\), then \(\sigma(A - KC)\) can be arbitrarily assigned because 
\[
\sigma(A^T + C^T K_1) = \sigma(A - KC). 
\]

Consider the system
\[
\dot{x} = A^T \hat{x} + C^T \hat{u}. 
\]

From duality, if the pair \((C, A)\) of the original system is observable, then the pair \((A^T, C^T)\) is controllable and we can find a feedback gain matrix \(K\) and \(\hat{u} = -K^T \hat{x}\) such that the spectrum of \(A - KC\) can be arbitrarily assigned, i.e.
\[
\Lambda_d = \sigma(A^T - C^T K^T) = \sigma(A - KC) = \{\lambda_{1d}, \ldots, \lambda_{nd}\}. 
\]

For the single output case, let \(\hat{A} = A^T\) and \(\hat{B} = C^T\), then \(\dot{x} = \hat{A} \hat{x} + \hat{B} \hat{u}\) and for design purposes, transform \(\hat{A}\) into \(\hat{A}_c\) and \(\hat{B}\) into \(\hat{B}_c\) using the transformation \(\hat{z} = V \hat{x}\),
where \( V = \begin{bmatrix} v & v \hat{A} & \cdots & v \hat{A}^{n-1} \end{bmatrix} \) and \( v \) is the last row of the inverse of \( \hat{Q} = \begin{bmatrix} \hat{B} | \hat{A} \hat{B} | \cdots | \hat{A}^{n-1} \hat{B} \end{bmatrix} \).

Specifically,

\[
\dot{z} = V \dot{x} = V \hat{A} V^{-1} \dot{z} + VB \dot{u} = \hat{A}_c \dot{z} + \hat{B}_c \dot{u} \quad \text{or}
\]

\[
\dot{z} = \hat{A}_c \dot{z} + \hat{B}_c \dot{u} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1 \\
-a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} \dot{z} + \begin{bmatrix} 0 \\
0 \\
\vdots \\
0 \\
1 \end{bmatrix} \dot{u}.
\]

Let \( \dot{u} = -K^T \hat{z} = -\begin{bmatrix} k_{c,n} & k_{c,n-1} & \cdots & k_{c,1} \end{bmatrix} \hat{z} \), then
$\hat{z} = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1 \\
-(a_n + k_{c,n}) & -(a_{n-1} + k_{c,n-1}) & \cdots & -(a_2 + k_{c,2}) & -(a_1 + k_{c,1})
\end{bmatrix} \hat{z}$.

Now,

$$\pi_{\hat{A}_c - \hat{B}_c K_c^T}(\lambda) = (\lambda + \lambda_{1d})(\lambda + \lambda_{2d}) \cdots (\lambda + \lambda_{nd})$$

$$= \lambda^n + a_{1,d} \lambda^{n-1} + a_{2,d} \lambda^{n-2} + \cdots + a_{n-1,d} \lambda + a_{n,d}$$

$$= \lambda^n + (a_1 + k_{c,1}) \lambda^{n-1} + (a_2 + k_{c,2}) \lambda^{n-2} + \cdots + (a_{n-1} + k_{c,n-1}) \lambda + a_n + k_{c,n}$$

Thus,

$$a_1 + k_{c,1} = a_{1d} \Rightarrow k_{c,1} = a_{d1} - a_1$$

$$a_2 + k_{c,2} = a_{2d} \Rightarrow k_{c,2} = a_{d2} - a_2$$

$$\vdots$$

$$a_n + k_{c,n} = a_{dn} \Rightarrow k_{c,n} = a_{dn} - a_n$$
Then, $K^T = K_c^T V$ and we can implement the estimator

$$
\dot{x}_e(t) = (A - KC)x_e(t) + Ky(t) + Bu(t)
$$

**Example**: Consider the continuous-time linear time-invariant system described by

$$
\dot{x}(t) = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} u(t).
$$

$$
y(t) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} x(t)
$$

Then the state estimator is described by

$$
\dot{x}_e(t) = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} x_e(t) + K \left[ y(t) - y_e(t) \right] + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} u(t).
$$

$$
y_e(t) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} x_e(t)
$$
Now, 

\[
\lambda I - A = \begin{bmatrix}
\lambda & 0 & 0 \\
0 & \lambda & 0 \\
0 & 0 & \lambda \\
\end{bmatrix} - \begin{bmatrix}
1 & 2 & 0 \\
1 & 2 & 0 \\
-1 & 0 & 1 \\
\end{bmatrix} = \begin{bmatrix}
\lambda & -2 & 0 \\
-1 & \lambda - 2 & 0 \\
1 & 0 & \lambda - 1 \\
\end{bmatrix}
\]

and 

\[
\det(\lambda I - A) = \det\begin{bmatrix}
\lambda & -2 \\
-1 & \lambda - 2 \\
\end{bmatrix} \cdot \det [\lambda - 1] = \lambda^3 - 3\lambda^2 + 2 = (\lambda - 1)(\lambda - 2.732)(\lambda + 0.732)
\]

Furthermore, 

\[
Q = \begin{bmatrix}
B & AB & A^2B \\
0 & 2 & 4 \\
1 & 2 & 6 \\
1 & 1 & -1 \\
\end{bmatrix} \text{ and } \det(Q) = 10 \Rightarrow \text{rank}(Q) = 3, \\
\varphi = \begin{bmatrix}
C \\
CA \\
CA^2 \\
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 1 \\
-1 & 2 & 1 \\
1 & 2 & 1 \\
\end{bmatrix} \text{ and } \det(\varphi) = -4 \Rightarrow \text{rank}(\varphi) = 3
\]
Which means that the system is completely controllable and observable.

Consider the system

\[
\dot{x} = \hat{A}x + \hat{B}\hat{u} = A^T\hat{x} + C^T\hat{u} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \hat{u}.
\]

Then

\[
\hat{Q} = \begin{bmatrix} \hat{B} | \hat{A}\hat{B} | \hat{A}^2\hat{B} \end{bmatrix} = \begin{bmatrix} C^T | A^TC^T | (A^T)^2C^T \end{bmatrix} = \varphi^T = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 2 \\ 1 & 1 & 1 \end{bmatrix},
\]

which implies that the system is controllable (because the pair \((C,A)\) is observable).

Let us construct the transformation \(V = \begin{bmatrix} v \\ v\hat{A} \\ v\hat{A}^2 \end{bmatrix}\), where \(v\) is the last row of \(\hat{Q}^{-1}\), i.e.
\[ \mathbf{v} \mathbf{Q} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 2 \\ 1 & 1 & 1 \end{bmatrix} = [v_1 + v_3, -v_1 + 2v_2 + v_3, v_1 + 2v_2 + v_3] = [0, 0, 1]. \]

Solving the 3 simultaneous equations \( v_1 + v_3 = 0 \), \(-v_1 + 2v_2 + v_3 = 0 \) and \( v_1 + 2v_2 + v_3 = 1 \), yields

\[ \mathbf{v} = \begin{bmatrix} 0.5 & 0.5 & -0.5 \end{bmatrix}. \]

Hence,

\[ \mathbf{V} = \begin{bmatrix} 0.5 & 0.5 & -0.5 \\ 1 & 1.5 & -1 \\ 3 & 4 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{V}^{-1} = \begin{bmatrix} 4 & -4 & 1 \\ -4 & 2 & 0 \\ -2 & -2 & 1 \end{bmatrix}. \]

We can easily show that

\[ \mathbf{\dot{z}} = \mathbf{V} \mathbf{A} \mathbf{V}^{-1} \mathbf{\dot{z}} + \mathbf{V} \mathbf{B} \mathbf{\dot{u}} = \hat{A}_c \mathbf{\dot{z}} + \hat{B}_c \mathbf{\dot{u}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 0 & 3 \end{bmatrix} \mathbf{\dot{z}} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{\dot{u}}. \]
Let \( \hat{u} = -K_c^T \hat{z} = -\left[ k_{c,3} \ k_{c,2} \ k_{c,1} \right] \hat{z} \), then

\[
\hat{A}_c - \hat{B}_c K_c^T = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-(2 + k_{c,3}) & -k_{c,2} & -(k_{c,1} - 3)
\end{bmatrix}.
\]

Suppose \( \Lambda_d = \sigma \left( A^T - C^T K_c^T \right) = \sigma (A - KC) = \{-1, -2, -3\} \). Then

\[
\pi_{\hat{A}_c - \hat{B}_c K_c^T}(\lambda) = (\lambda + 1)(\lambda + 2)(\lambda + 3)
= \lambda^3 + 6\lambda^2 + 11\lambda + 6
= \lambda^3 + (-3 + k_{c,1})\lambda^2 + k_{c,2}\lambda + 2 + k_{c,3}
\]

implies that \(-3 + k_{c,1} = 6\), \( k_{c,2} = 11\), and \( 2 + k_{c,3} = 6 \) or that \( K_c^T = \begin{bmatrix} 4 & 11 & 9 \end{bmatrix} \). Finally,

\[
K^T = K_c^T V = \begin{bmatrix} 0.5 & 0.5 & -0.5 \\ 1 & 1.5 & -1 \\ 3 & 4 & -2 \end{bmatrix} \begin{bmatrix} 40 \\ 54.5 \\ -31 \end{bmatrix}
\]

and

\[
K = \begin{bmatrix} 40 \\ 54.5 \\ -31 \end{bmatrix}.
\]

Once again, one can easily show that \( \sigma(A - KC) = \{-1, -2, -3\} \).
**Performance Evaluation**: The following figures show how well the state estimator tracks the original system states when $u(t) = 0$, $\mathbf{x}(0) = [0.01 \ 0 \ 0.1]^T$ and $\mathbf{x}_e(0) = [-0.05 \ 0 \ 0.01]^T$.
Example: Obtain a full-order state estimator for the 2nd order system

\[ \dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \]

\[ y = [0 \ 1] x \]

such that the desired estimator spectrum is \( \sigma(A - KC) = \{-5, -5\} \).

\[ \varphi = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \implies \text{rank } [\varphi] = 2 \implies \text{system is observable.} \]

Now, the physical system spectrum is \( \sigma(A) = \{-1, 1\} \) \( \implies \text{system is unstable.} \)
Let
\[ \dot{x}_e = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_e + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} [y - (0 \ 1) x_e] + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \]
then,
\[ \dot{\epsilon} = \left[ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \right] \epsilon = \begin{bmatrix} 0 & 1 - k_1 \\ 1 & -k_2 \end{bmatrix} \epsilon = \begin{bmatrix} 0 & -(k_1 - 1) \\ 1 & -k_2 \end{bmatrix} \epsilon \]
and \( \pi_{A-KC}(\lambda) = \lambda^2 + k_2 \lambda + (k_1 - 1) = (\lambda + 5)^2 = \lambda^2 + 10\lambda + 25 \)
\[ \Rightarrow k_1 - 1 = 25, \text{ or } k_1 = 26 \text{ and } k_2 = 10. \]
Hence, \( K = \begin{bmatrix} 26 \\ 10 \end{bmatrix} \).

Let the initial condition of the state estimator be \( x_e(0) = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix} \) and let the reference input be a step input, i.e. \( u(t) = u_r(t) = 1, t \geq 0 \).

Then, the performance of the state estimator is shown in the following figures:
Feedback From State Estimates:

Let a linear time-invariant dynamic system be described by

\[
\begin{align*}
\dot{x} &= Ax + Bu \\
y &= Cx + Du
\end{align*}
\]

Let the system described by \(\otimes\) be controllable and observable, then, if the state variables are not available for feedback, we can design a state estimator and feedback the estimates in lieu of the actual state variables.

Let the input signal be given by \(u(t) = -Fx_e(t) + u_r(t)\).

Then the closed-loop system is given by

\[
\begin{align*}
\dot{x}(t) &= Ax(t) - BFx_e(t) + Bu_r(t)
\end{align*}
\]

and the state estimator is now described by

\[
\begin{align*}
\dot{x}_e(t) &= Ax_e(t) + K(y(t) - Cx_e(t)) + B(-Fx_e(t) + u_r(t))
\end{align*}
\]
In block diagram form, the complete system is shown below

This is called the controller-estimator configuration.
It can be shown that the above system has the same eigenvalues and the same transfer function as the system with input signal (control + reference signal)

\[ u(t) = -F x(t) + u_r(t) \]

In augmented form,

\[
\begin{bmatrix}
\dot{x} \\
\dot{x}_e
\end{bmatrix} =
\begin{bmatrix}
A & -BF \\
KC & A - KC - BF
\end{bmatrix}
\begin{bmatrix}
x \\
x_e
\end{bmatrix}
+ \begin{bmatrix}
B \\
B
\end{bmatrix} u_r
\]

\[ y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ x_e \end{bmatrix} \]

Consider the following nonsingular similarity transformation

\[
\begin{bmatrix}
x \\
x_e
\end{bmatrix} = \begin{bmatrix}
x \\
x - x_e
\end{bmatrix} = \begin{bmatrix}
I & 0 \\
I & -I
\end{bmatrix} \begin{bmatrix}
x \\
x_e
\end{bmatrix} = T \begin{bmatrix}
x \\
x_e
\end{bmatrix}
\]
Then, the new equivalent system is given by

\[
\begin{bmatrix}
\dot{x} \\
\dot{\varepsilon}
\end{bmatrix} =
\begin{bmatrix}
A - BF & BF \\
0 & A - KC
\end{bmatrix}
\begin{bmatrix}
x \\
\varepsilon
\end{bmatrix} +
\begin{bmatrix}
B \\
0
\end{bmatrix} u_r(t)
\]

\[
y =
\begin{bmatrix}
C & 0
\end{bmatrix}
\begin{bmatrix}
x \\
\varepsilon
\end{bmatrix}
\]

To show that the estimator does not affect the location of the eigenvalues of the original state feedback system, we calculate the augmented system’s transfer function.

Let \( A_{\text{overall}} \equiv \begin{bmatrix} A - BF & BF \\ 0 & A - KC \end{bmatrix} \), \( B_{\text{overall}} \equiv \begin{bmatrix} B \\ 0 \end{bmatrix} \) and \( C_{\text{overall}} \equiv \begin{bmatrix} C & 0 \end{bmatrix} \).

Then, \( Y(s) = C_{\text{overall}}(sI - A_{\text{overall}})^{-1} B_{\text{overall}} U(s) \),
where

$$\left(sI - A_{overall}\right)^{-1} = \begin{pmatrix} \left(sI_1 - (A - BF)\right)^{-1} & \left(sI_1 - (A - BF)\right)^{-1}BF\left(sI_1 - (A - KC)\right)^{-1} \\ 0 & \left(sI_1 - (A - KC)\right)^{-1} \end{pmatrix}$$

So,

$$Y(s) = C\left(sI_1 - (A - BF)\right)^{-1}BU(s) = H(s)U(s).$$

Which establishes that the transfer function of the original closed-loop system does not change with the introduction of the state estimator in the loop.

**Example**: Consider the same system of the last example. This system is unstable with eigenvalues at -1 and 1. Let the desired closed-loop system have eigenvalues at -2 ± j2, i.e., the desired characteristic polynomial is given by

$$\pi_{A-BF}(\lambda) = (\lambda + 2 + j2)(\lambda + 2 - j2) = \lambda^2 + 4\lambda + 8$$

$$= \det(\lambda I - (A - BF)) = \det\begin{pmatrix} \lambda & -1 \\ f_1 - 1 & \lambda + f_2 \end{pmatrix}$$

$$= \lambda^2 + f_2\lambda + f_1 - 1$$
This implies that $f_1 = 9$ and $f_2 = 4$.

Let us now apply the feedback control based on the estimates rather than on the actual values of the states, i.e.,

$$u(t) = -[9 \ 4]\begin{bmatrix} x_{e1}(t) \\ x_{e2}(t) \end{bmatrix} + u_r(t).$$

Then, the performance of the estimator-based closed-loop system is shown in the following figures when the reference input is a unit step and the initial values of the system state and that of the estimator are

$$x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } x_e(0) = \begin{bmatrix} .5 \\ -.2 \end{bmatrix},$$

respectively.
System with estimated state feedback control

Estimator state feedback controlled system with unit step reference input
System with actual state feedback control
Feedback Control Design for Multiple-Input Linear Time Invariant Systems

Let an open-loop linear time-invariant dynamic system be described by

\[
\frac{dx(t)}{dt} = Ax(t) + Bu(t), \; x(t) \in \mathbb{R}^n, \; u(t) \in \mathbb{R}^m.
\]

Suppose the system is completely controllable. Then a feedback gain matrix \( F \) can be found, such that the application of the input \( u(t) = Fx(t) + u_r(t) \) results in a closed-loop system

\[
\frac{dx(t)}{dt} = (A + BF)x(t) + Bu_r(t), \; x(t) \in \mathbb{R}^n, \; u_r(t) \in \mathbb{R}^m
\]

that is asymptotically stable.

Now, the characteristic polynomial of the closed-loop system is given by

\[
\pi_{A+BF}(\lambda) = \det(\lambda I - A - BF).
\]
**Problem Statement:** Find the feedback gain matrix $F$ such that the equation

$$\pi_{A+BF}(\lambda) = \det(\lambda I - A - BF) = 0 \quad (*)$$

is satisfied by $\lambda = \lambda_{di} \in \Lambda_d$, $\Lambda_d = \{\lambda_{d1}, \lambda_{d2}, \ldots, \lambda_{dn}\}$.

If equation $(*)$ is true, then there exists a nonzero vector $v_i$ such that

$$(\lambda_{di} I - A - BF)v_i = 0 \quad (#)$$

Equivalently,

$$(A + BF)v_i = \lambda_{di}v_i$$

says that $v_i$ is an eigenvector of the closed-loop system matrix $A + BF$ associated with the desired closed-loop eigenvalue $\lambda_{di}$. 
Equation (\#) can be rewritten as

\[
\begin{bmatrix}
\lambda_i I - A & B
\end{bmatrix}
\begin{bmatrix}
v_i \\
-Fv_i
\end{bmatrix} = 0_{n+m}. \quad (+)
\]

At this stage, both \(v_i\) and \(F\) are unknown. Define the unknown vector \(w_i \in \mathbb{C}^{n+m}\) as

\[
w_i \equiv \begin{bmatrix}
v_i \\
-Fv_i
\end{bmatrix}.
\]

Then the problem can be solved by finding the set of vectors \(w_i\) (possibly more than one) that satisfy equation (+), namely,

\[
\begin{bmatrix}
\lambda_i I - A & B
\end{bmatrix}w_i = 0_{n+m}.
\]

This is equivalent to computing the null space of \(\begin{bmatrix}
\lambda_i I - A & B
\end{bmatrix}\). If a sufficient number of linearly independent vectors is found, then a feedback gain matrix \(F\) can be obtained.
Note that the first $n$ elements of $w_i$ correspond to the elements of the vector $v_i$ and the remaining $m$ elements of $w_i$ correspond to the elements of the vector $-Fv_i$.

Now, the null space of $[\lambda_{di}I - A \mid B]$ has maximal dimension $m$, namely, it is described by a set of $m$ linearly independent columns of a matrix $U(\lambda_{di})$ with dimension $(n+m)\times m$, in other words,

$$U(\lambda_{di}) = [w_{i1} \ w_{i2} \ \cdots \ w_{im}].$$

Let $w_{ij} \equiv \begin{bmatrix} v_{ij} \\ z_{ij} \end{bmatrix}$, then

$$U(\lambda_{di}) = \begin{bmatrix} v_{i1} & v_{i2} & \cdots & v_{im} \\ -Fv_{i1} & -Fv_{i2} & \cdots & -Fv_{im} \\ z_{i1} & z_{i2} & \cdots & z_{im} \end{bmatrix} = \begin{bmatrix} V(\lambda_{di}) \\ -F \cdot V(\lambda_{di}) \\ Z(\lambda_{di}) \end{bmatrix}, \ i = 1, 2, \ldots, n ,$$

The bottom row can be written as

$$-F \begin{bmatrix} v_{i1} & v_{i2} & \cdots & v_{im} \end{bmatrix} = -F V(\lambda_{di}) = Z(\lambda_{di}), \ i = 1, 2, \ldots, n$$
Hence, $F$ can be found by solving the matrix equation

$$-F \begin{bmatrix} V(\lambda_{d_1}) & V(\lambda_{d_2}) & \cdots & V(\lambda_{d_n}) \end{bmatrix} = \begin{bmatrix} Z(\lambda_{d_1}) & Z(\lambda_{d_2}) & \cdots & Z(\lambda_{d_n}) \end{bmatrix}. \quad (\@)$$

Since $m > 1$, the system of equations (\@) is overdetermined and cannot be solved directly. However, if the system $(A, B)$ is completely controllable, a set of $n$ linearly independent columns, one for each $\lambda_{di}$, can be found in each of the matrices

$$\begin{bmatrix} V(\lambda_{d_1}) & V(\lambda_{d_2}) & \cdots & V(\lambda_{d_n}) \end{bmatrix}$$

and

$$\begin{bmatrix} Z(\lambda_{d_1}) & Z(\lambda_{d_2}) & \cdots & Z(\lambda_{d_n}) \end{bmatrix},$$

respectively. Let the selected $n$ columns from the left-hand side $\begin{bmatrix} V(\lambda_{d_1}) & V(\lambda_{d_2}) & \cdots & V(\lambda_{d_n}) \end{bmatrix}$ be the matrix $G$ and the selected columns from the right-hand side $\begin{bmatrix} Z(\lambda_{d_1}) & Z(\lambda_{d_2}) & \cdots & Z(\lambda_{d_n}) \end{bmatrix}$ be the matrix $H$, then $-FG = H$, or $F = -HG^{-1}$ is a feedback gain matrix that results in the desired set of closed-loop eigenvalues

$$\Lambda_d = \{\lambda_{d_1}, \lambda_{d_2}, \ldots, \lambda_{d_n}\}.$$ 

The $F$ matrix just computed is not unique, since the columns of both matrices $G$ and $H$ are selected arbitrarily, with the only requirement that the set of $n$ columns be linearly independent. Thus, many solutions exist.
Example: Let the dynamics of a continuous-time LTI system be described by

\[
\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 2 \\ 0 & 3 \end{bmatrix} x(t) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u(t).
\]

The open-loop system has eigenvalues 0 and 3 (system is open-loop unstable).

Let the set of desired closed-loop eigenvalues be \( \Lambda_d = \{\lambda_{d1}, \lambda_{d2}\} = \{-3, -5\} \). Now,

\[
[\lambda_{di}I - A \mid B] = \begin{bmatrix} \lambda_{di} & -2 & 1 & 0 \\ 0 & \lambda_{di} -3 & 0 & 1 \end{bmatrix}.
\]

If \( \lambda_{d1} = -3 \), \[
\begin{bmatrix} \lambda_{d1} & -2 & 1 & 0 \\ 0 & \lambda_{d1} -3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 & -2 & 1 & 0 \\ 0 & -6 & 0 & 1 \end{bmatrix}
\]

and the null space of \([ -3I - A \mid B ]\) is described by the matrix

\[
U(-3) = \begin{bmatrix}
3.1531 \times 10^{-1} & -1.0101 \times 10^{-1} \\
1.5014 \times 10^{-3} & 1.6351 \times 10^{-1} \\
\cdots & \cdots \\
9.4894 \times 10^{-1} & 2.3991 \times 10^{-2} \\
9.0084 \times 10^{-3} & 9.8106 \times 10^{-1}
\end{bmatrix}.
\]
If $\lambda_{d_2} = -5$, 
\[
\begin{bmatrix}
\lambda_{d_2} & -2 & 1 & 0 \\
0 & \lambda_{d_2} - 3 & 0 & 1
\end{bmatrix} = \begin{bmatrix}
-5 & -2 & 1 & 0 \\
0 & -8 & 0 & 1
\end{bmatrix}
\]
and the null space of 
\[
[-5I - A | B]
\]
is described by
\[
U(-5) = \begin{bmatrix}
1.9510 \times 10^{-1} & -4.8140 \times 10^{-2} \\
3.1009 \times 10^{-4} & 1.2389 \times 10^{-1} \\
9.8060 \times 10^{-1} & 7.0754 \times 10^{-3} \\
2.4807 \times 10^{-3} & 9.9110 \times 10^{-1}
\end{bmatrix}.
\]
Let us construct $G_1$ and $H_1$ by selecting the first column of $U(-3)$ and $U(-5)$, i.e.
\[
G_1 = \begin{bmatrix}
3.1531 \times 10^{-1} & 1.9510 \times 10^{-1} \\
1.5014 \times 10^{-3} & 3.1009 \times 10^{-4}
\end{bmatrix}
\quad\text{and}\quad
H_1 = \begin{bmatrix}
9.4894 \times 10^{-1} & 9.8060 \times 10^{-1} \\
9.0084 \times 10^{-3} & 2.4807 \times 10^{-3}
\end{bmatrix}.
\]
Both $G_1$ and $H_1$ have rank 2, which means that the inverse of $G_1$ exists. Thus,
\[
F_1 = -H_1G_1^{-1} = \begin{bmatrix}
-6 & 627.03 \\
-4.7388 \times 10^{-3} & -5.005
\end{bmatrix}
\quad\text{and}\quad
\Lambda(A + BF_1) = \{-3, -5\}.
Let us construct $G_2$ and $H_2$ by selecting the second column of $U(-3)$ and $U(-5)$, i.e.

\[
G_2 = \begin{bmatrix}
-1.0101 \times 10^{-1} & -4.8140 \times 10^{-2} \\
1.6351 \times 10^{-1} & 1.2389 \times 10^{-1}
\end{bmatrix}
\quad \text{and} \quad
H_2 = \begin{bmatrix}
2.3991 \times 10^{-2} & 7.0754 \times 10^{-3} \\
9.8106 \times 10^{-1} & 9.9110 \times 10^{-1}
\end{bmatrix}.
\]

Both $G_2$ and $H_2$ have rank 2, which means that the inverse of $G_2$ exists. Thus,

\[
F_2 = -H_2 G_2^{-1} = \begin{bmatrix}
3.9102 \times 10^{-1} & 9.4832 \times 10^{-2} \\
-8.7268 & -11.391
\end{bmatrix}
\quad \text{and} \quad
\Lambda(A + BF_2) = \{-3, -5\}.
\]

There are many more feedback gain matrices $F$ that could be computed from selecting different columns of $U(-3)$ and $U(-5)$ or linear combinations of them to construct the matrices $G$ and $H$. As a consequence, there are several methodologies to select matrices $G$ and $H$ to satisfy different design criteria. For example, Matlab uses a method which selects a feedback gain matrix $F$ for a given desired set of eigenvalues described by

\[
\Lambda_d = \{\lambda_{d1}, \lambda_{d2}, \ldots, \lambda_{dn}\}
\]

such that the closed-loop system is robust in terms of reducing the sensitivity problems associated with large gains.
Robust Pole Placement for Multiple-Input Systems

Given real matrices \((A, B), A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}\) and a symmetric set of desired closed-loop eigenvalues \(\Lambda_d, \Lambda_d = \{\lambda_{d1, \lambda_{d2}, \cdots, \lambda_{dn}\}\}\), find a real matrix \(F \in \mathbb{R}^{m \times n}\) such that the eigenvalues of \(A + BF\) are as insensitive to perturbations as possible.

Let \(x_i\) and \(y_i, i = 1, 2, \cdots, n\), be the right and left eigenvectors of the closed-loop system matrix \(M = A + BF\) associated with the eigenvalue \(\lambda_{di}\), namely,

\[
Mx_i = \lambda_{di}x_i, \quad y_i^TM = \lambda_{di}y_i^T.
\]

If \(M\) is non-defective (it has a full set of \(n\) linearly independent eigenvectors), then \(M\) is diagonalizable. Moreover, the sensitivity of the eigenvalue \(\lambda_{di}\) to perturbations in the components of \(A, B,\) and \(F\) depends on the magnitude of the condition number \(c_i\), which is defined as

\[
c_i = \frac{1}{s_i} \equiv \frac{\|y_i\|_2 \|x_i\|_2}{|y_i^T x_i|} \geq 1,
\]
where $s_i$ is the sensitivity of eigenvalue $\lambda_{di}$, which is defined as the cosine of the angle between the right and left eigenvectors corresponding to $\lambda_{di}$.

A bound on the sensitivities of the eigenvalues is given by

$$
\max_i \{c_i\} \leq k_2(X) \equiv \|X\|_2 \|X^{-1}\|_2,
$$

where $k_2(X)$ is the condition number of the matrix $X = [x_1 \ x_2 \ \cdots \ x_n]$ of eigenvectors. Furthermore, the condition numbers achieve the minimum value of 1 for all $i$, $i = 1, 2, \cdots, n$, if and only if $X$ is a normal matrix, i.e. $X^*X = XX^*$. In this case the eigenvectors of $X$ may be scaled to give an orthonormal basis for $\mathbb{C}^n$, and then the matrix $X$ is perfectly conditioned with $k_2(X) = 1$.

Let us now re-state the robust pole placement problem as follows: Given real matrices $(A, B)$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ and a symmetric set of desired closed-loop eigenvalues $\Lambda_d$, $\Lambda_d = \{\lambda_{d1}, \lambda_{d2}, \cdots, \lambda_{dn}\}$, find a real matrix $F \in \mathbb{R}^{m \times n}$ and a non-singular matrix $X$ that satisfy the equation

$$(A + BF)X = XD,$$
where $D = \text{diag} \{ \lambda_{d_1}, \lambda_{d_2}, \cdots, \lambda_{d_n} \}$, such that some measure $\xi$ of the conditioning or robustness of the eigenproblem is optimized.

**Theorem:** Given $D = \text{diag} \{ \lambda_{d_1}, \lambda_{d_2}, \cdots, \lambda_{d_n} \}$ and $X$ non-singular, then there exists $F$ which is a solution to $(A + BF)X = XD$ if and only if

$$U_1^T (AX - XD) = 0,$$

where $B = \begin{bmatrix} U_0 & U_1 \end{bmatrix} \begin{bmatrix} Z \end{bmatrix}$ with $U = \begin{bmatrix} U_0 & U_1 \end{bmatrix}$ an orthogonal matrix with

$U_0 \in \mathbb{R}^{n \times m}$ and $U_1 \in \mathbb{R}^{n \times (n-m)}$, and $Z \in \mathbb{R}^{m \times m}$ non-singular matrix. Then $F$ is explicitly given by

$$F = Z^{-1} U_0^T \left( XDX^{-1} - A \right).$$
Proof: If $B$ is of full rank, then each column of $B$ is a linear combination of the columns of $U_0$ and $Z$ is non-singular. Now, $(A + BF) X = XD$ can be rewritten as $BFX = XD - AX$. Post multiplying both sides of the last equation by $X^{-1}$, yields

$$(BFX = XD - AX) X^{-1} = BF = XDX^{-1} - A.$$ 

Let us now pre-multiply the last equation by $U^T$, i.e.

$$U^T BF = U^T (XDX^{-1} - A).$$

The left hand side is equal to

$$U^T BF = U^T U \begin{bmatrix} Z \\ - - \\ 0 \end{bmatrix} F = \begin{bmatrix} Z \\ - - \\ 0 \end{bmatrix} F = \begin{bmatrix} ZF \\ - - \\ 0 \end{bmatrix}.$$ 

The right hand side is equal to

$$U^T (XDX^{-1} - A) = \begin{bmatrix} U_0^T \\ - - \\ U_1^T \end{bmatrix} (XDX^{-1} - A) = \begin{bmatrix} U_0^T \left(XDX^{-1} - A\right) \\ - - \\ U_1^T \left(XDX^{-1} - A\right) \end{bmatrix}.$$
Hence,

\[ ZF = U_0^T \left( XDX^{-1} - A \right) \text{ or } F = Z^{-1} U_0^T \left( XDX^{-1} - A \right) \]

and

\[ 0 = U_1^T \left( XDX^{-1} - A \right). \]

Now, \( BF = XDX^{-1} - A \) implies that \( F \) exists if and only if

\[ \text{Im} \left\{ XDX^{-1} - A \right\} \subseteq \text{Im} \{ B \} \equiv \text{Im} \{ U_0 \}. \]

Hence, \( \text{Im} \left\{ XDX^{-1} - A \right\} \perp \text{Null} \{ B \} \equiv \text{Im} \{ U_1 \}. \)

**Corollary:** The eigenvector \( x_i \) of the closed-loop system matrix \( A + BF \) associated with desired eigenvalue \( \lambda_{di} \in \Lambda \) must belong to the space \( S_i = \text{Null} \left\{ U_1^T \left( A - \lambda_{di} I \right) \right\} \) of
dimension \( m + k_{\lambda_{di}} \), where \( k_{\lambda_{di}} = \dim \left( \text{Null} \left\{ [B | A - \lambda_{di} I]^T \right\} \right) \).
The robust pole placement problem reduces to the problem of selecting linearly independent vectors \( x_i \in S_i, \ i = 1, 2, \ldots, n \) such that the eigenproblem \((A + BF)X = XD\) is as well-conditioned as possible.

When the pair \((A, B)\) is completely controllable, \( k_{\lambda_{di}} = 0, \ \forall \lambda_{di} \) and the multiplicity of the desired eigenvalue \( \lambda_{di} \in \Lambda_d \) cannot exceed \( m \), since the maximum number of independent eigenvectors which can be chosen to correspond to \( \lambda_{di} \in \Lambda \) is equal to \( \dim(S_i) = m \).

**Theorem:** The feedback gain matrix \( F \) and the zero-input \((u_r(t) = 0 \ \text{or} \ u_r(k) = 0)\) state response \( x(t) \) or \( x(k) \) of the closed-loop continuous- or discrete-time system

\[
\frac{dx(t)}{dt} = (A + BF)x(t) \quad \text{or} \quad x(k + 1) = (A + BF)x(k)
\]

with initial value \( x(0) \), satisfy the inequalities

\[
\|F\|_2 \leq \left( \|A\|_2 + \max_i \left\{ \left| \lambda_{di} \right| \cdot k_2(X) \right\} \right) \frac{\sigma_m \{B\}}{\sigma_m \{B\}}, \quad \text{where} \quad \sigma_m \{B\} \equiv \min \{\sigma \{B\}\}
\]

and

\[
\|x(t)\|_2 \leq k_2(X) \cdot \max_i \left\{ \left| e^{\lambda_{di} t} \right| \right\} \cdot \|x(0)\|_2 \quad \text{or} \quad \|x(k)\|_2 \leq k_2(X) \cdot \max_i \left\{ \left| \lambda_{di}^k \right| \right\} \cdot \|x(0)\|_2.
\]
**Example:** Consider again the second-order system of the last example, i.e.

\[
\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 2 \\ 0 & 3 \end{bmatrix} x(t) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u(t),
\]

with desired eigenvalues \( \Lambda_d = \{\lambda_{d1}, \lambda_{d2}\} = \{-3, -5\} \).

Matlab implements the methodology to compute the feedback gain matrix \( F \) just described and uses the command \( "K=place(A,B,p)" \), where \( p \) is a vector that contains the desired eigenvalues \( \{\lambda_{d1}, \lambda_{d2}, \ldots, \lambda_{dn}\} \) and \( K = -F \).

Using the above Matlab command with \( p = [-3, -5] \), we get the feedback gain matrix

\[
K = \begin{bmatrix} 3 & 2 \\ 0 & 8 \end{bmatrix} \quad \text{or} \quad F = \begin{bmatrix} -3 & -2 \\ 0 & -8 \end{bmatrix}
\]

and \( \lambda(A + BF) = \{-3, -5\} \).
APPENDIX

Singular value decomposition (SVD)
Let \( H \in \mathbb{R}^{m \times n} \) and define \( M = H^T H \in \mathbb{R}^{n \times n} \). Then \( M = M^T \geq 0 \). Let \( r \) be the total number of positive eigenvalues of \( M \), then we may arrange them such that

\[
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n
\]

Let \( p = \min\{m, n\} \), then the set \( \{ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_p \} \) is called the singular values of \( H \), where \( \sigma_i = \sqrt{\lambda_i} \) and \( r = \text{rank}(H) \).

Example: Let a rectangular matrix \( H \) be given by

\[
H = \begin{bmatrix}
2 & 1 \\
-1 & 2 \\
2 & 4
\end{bmatrix}
\]
Then
\[ M = H^T H = \begin{bmatrix} 9 & 8 \\ 8 & 21 \end{bmatrix} \]

Now, \( \det(\lambda I - M) = \lambda^2 - 30\lambda + 125 \Rightarrow \) eigenvalues \((M) = \{25, 5\} \) and the singular values of \( H \) are the square root of the eigenvalues of \( M \), i.e., \( \{5, \sqrt{5}\} \).

**Example:** Let \( H \) now be described by
\[
H = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}
\]

Then
\[
M = H^T H = \begin{bmatrix} 16 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix}
\]

and
\[
\det(\lambda I - M) = \det\begin{bmatrix} \lambda - 16 & 0 & 0 \\ 0 & \lambda - 4 & -2 \\ 0 & -2 & \lambda - 1 \end{bmatrix} = (\lambda - 16)(\lambda - 5)\lambda
\]
which implies that the set of eigenvalues of $M$ is $\{16, 5, 0\} \Rightarrow$ singular values of $H$
are $\{4, \sqrt{5}\}$, since $\min\{m, n\} = \min\{2, 3\} = 2$. Also, $\text{rank}(H) = 2$.

**Theorem:** Let $H \in \mathbb{R}^{m \times n}$, then $H = RSQ^T$ with $R^TR = RR^T = I_m$, $Q^TQ = QQ^T = I_n$, and $S \in \mathbb{R}^{m \times n}$ with the singular values of $H$ on its main diagonal and such that $Q^TH^THQ = D = S^TS$ with $D$ a diagonal matrix with the squared singular values of $H$ on its main diagonal.

**Example:** Let $H$ be given by $H = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

then the eigenvalues of $M = H^TH$ $\{16, 5, 0\}$ give rise to the normalized eigenvectors

$$\tilde{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T$$
$$\tilde{q}_2 = \begin{bmatrix} 0 \\ \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}^T$$
$$\tilde{q}_3 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}^T$$
Thus,

\[ Q = [\tilde{q}_1 \quad \tilde{q}_2 \quad \tilde{q}_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} \]

\[ S = HQ = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{5} & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} \]

\[ S^T S = \begin{bmatrix} 4 & 0 \\ 0 & \sqrt{5} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 16 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \]