State Feedback Control Systems—LQR Problem

System Equation:

$$\dot{X}(t) = A(t)X(t) + Bu(t) \tag{1}$$

State Feedback:

$$u(t) = -K(t)X(t) \tag{2}$$

The closed-loop system becomes:

$$\begin{aligned} \dot{x}(t) &= A(t)X(t) + Bu(t) \\ &= [A(t) - B(t)K(t)]X(t) \\ &= \alpha(t)X(t) \end{aligned} \tag{3}$$

Where, $\alpha(t) = A(t) - B(t)K(t)$

The solution of Eq.(3) can be given as:

$$X(t, t_0) = \phi(t, t_0)X(t_0) \tag{4}$$

Where, $\phi(t, t_0)$ is a state transition matrix satisfying the following:

$$\dot{\phi}(t, t_0) = \alpha(t)\phi(t, t_0) \tag{5}$$
$$\phi(t_0, t_0) = I \tag{6}$$
$$\phi(t, t_0)\phi(t_0, t) = I \tag{7}$$

According to Eq.(7), we have:

$$\dot{\phi}(t, t_0)\phi(t_0, t) = -\phi(t, t_0)\dot{\phi}(t_0, t) \tag{8}$$

But, by making use of Eq.(5), we have:

$$\dot{\phi}(t, t_0)\phi(t_0, t) = \alpha(t)\phi(t, t_0)\phi(t_0, t) \tag{9}$$

Therefore,

$$\alpha(t)\phi(t, t_0)\phi(t_0, t) = -\phi(t, t_0)\dot{\phi}(t_0, t) \tag{10}$$

So that,

$$\alpha(t) = -\phi(t, t_0)\dot{\phi}(t_0, t), \quad and \quad \dot{\phi}(t_0, t) = -\phi(t, t_0)^{-1}\alpha(t) = \phi(t_0, t)\alpha(t) \tag{11}$$

In other words,

$$\dot{\phi}(\tau, t) = -\phi(\tau, t)\alpha(t) \qquad\qquad \dot{\phi}(t, \tau) = -\phi(t, \tau)\alpha(\tau) \tag{12}$$

Consider a performance for the state feedback control system:

$$J(t, t_1, X(t), u(t \in [t, t_1])) = X^T(t_1)Q_1X(t_1) + \int_t^{t_1} [X^T(\tau)QX(\tau) + u^T(\tau)Ru(\tau)]d\tau \tag{13}$$

Since

$$u(t) = -K(t)X(t) \tag{14}$$

Then,

$$\begin{aligned}
J(t, t_1, X(t), u(t \in [t, t_1])) &= J(t, t_1, X(t), K(t \in [t, t_1])) \\
&= X^T(t_1)Q_1X(t_1) + \int_t^{t_1} [X^T(\tau)QX(\tau) + u^T(\tau)Ru(\tau)]d\tau \\
&= X^T(t_1)Q_1X(t_1) + \int_t^{t_1} [X^T(\tau)\tilde{Q}X(\tau)]d\tau \\
&= X^T(t)\phi(t_1, t)Q_1\phi(t_1, t)X(t) + \int_t^{t_1} [X^T(t)\phi^T(\tau, t)\tilde{Q}\phi(\tau, t)x(t)d\tau]dt \\
&= X^T(t)\{\phi^T(t_1, t)Q_1\phi(t_1, t) + \int_t^{t_1} [\phi^T(\tau, t)\tilde{Q}\phi(\tau, t)d\tau]d\tau\}X(t) \\
&= X^T P(t_1, t, K(t \in [t, t_1]))X(t)
\end{aligned} \tag{15}$$

$$P(t_1,t,K(t\in[t,t_1])) = \phi^T(t_1,t)Q_1\phi(t_1,t) + \int_t^{t_1}[\phi^T(\tau,t)\tilde{Q}\phi(\tau,t)d\tau]d\tau \qquad (16)$$

Thus,

$$\begin{aligned}
\dot{P}(t_1,t,K(t\in[t,t_1])) &= +\int_t^{t_1}[\dot{\phi}^T(\tau,t)\tilde{Q}\phi(\tau,t)d\tau]d\tau - \tilde{Q} \\
&\quad + \int_t^{t_1}[\phi^T(\tau,t)\tilde{Q}\dot{\phi}(\tau,t)d\tau]d\tau \\
&\quad + \dot{\phi}^T(t_1,t)Q_1\phi(t_1,t) + \phi(t_1,t)Q_1\dot{\phi}(t_1,t) \qquad (17) \\
\dot{P}(t_1,t,K(t\in[t,t_1])) &= -\tilde{Q} - \int_t^{t_1}\alpha^T(t)\phi^T(\tau,t)\tilde{Q}\phi(\tau,t)d\tau - \\
&\quad - \int_t^{t_1}\phi^T(\tau,t)\tilde{Q}\phi(\tau,t)\alpha(t)d\tau \\
&\quad -\alpha^T(t)\phi^T(t_1,t)Q\phi(t_1,t) - \phi^T(t_1,t)Q\phi(t_1,t)\alpha(t) \\
&= -\tilde{Q} - \alpha^T(t)\{\int_t^{t_1}\phi^T(\tau,t)\tilde{Q}\phi(\tau,t)d\tau + \phi^T(t_1,t)Q\phi(t_1,t)\} \\
&= -\{\int_t^{t_1}\phi^T(\tau,t)\tilde{Q}\phi(\tau,t)d\tau + \phi^T(t_1,t)Q\phi(t_1,t)\}\alpha(t) \\
&\quad -\alpha^T(t)\phi^T(t_1,t)Q\phi(t_1,t) - \phi^T(t_1,t)Q\phi(t_1,t)\alpha(t) \\
&\quad +\underline{\alpha^T(t)\phi^T(t_1,t)Q\phi(t_1,t)} + \underline{\phi^T(t_1,t)Q\phi(t_1,t)\alpha(t)} \\
&= -\tilde{Q} - \alpha^T(t)P(t_1,t,K(t\in[t,t_1])) - P^T(t_1,t,K(t\in[t,t_1]))\alpha(t) \qquad (18)
\end{aligned}$$

If we introduce the definitions of $\alpha(t)$ and $\tilde{Q}$ into Eq.(17), we have:

$$\begin{aligned}
\dot{P}(t_1,t,K(t\in[t,t_1])) &= -Q(t) - k^T(t)R(t)K(t) - [A(t)-B(t)K(t)]^T P(t_1,t,K(t\in[t,t_1])) \\
&\quad -P(t_1,t,K(t\in[t,t_1]))[A(t)-B(t)K(t)]^T \qquad (19)
\end{aligned}$$

or,

$$\begin{aligned}
\dot{P}(t_1,t,K(t\in[t,t_1])) &= -Q(t) - A(t)^T P(t_1,t,K(t\in[t,t_1])) - P(t_1,t,K(t\in[t,t_1]))A(t) \\
&\quad +K^T(t)B^T(t)P(t_1,t) + P(t_1,t,K(t\in[t,t_1]))B(t)K(t) \qquad (20)
\end{aligned}$$

and the optimal state feedback control is to find an optimal $K^*(t)$ such that :

$$\begin{aligned}
J^*(t_1,t,X(t),K^*(t\in[t,t_1])) &= X^T(t)P^*(t_1,t,K^*(t\in[t,t_1]))X(t) \\
&\le X^T(t)P(t_1,t,K(t\in[t,t_1]))X(t); \\
&\qquad for \quad any \quad X(t) \qquad (21)
\end{aligned}$$

As optimal feedback gain K^* must be unique, the avove expression is rewritten as:

$$\begin{aligned} J^*(t_1,t,X(t)) &= X^T(t)P^*(t_1,t,)X(t) \leq X^T(t)P(t_1,t,K(t\in[t,t_1]))X(t); \\ & \quad for \quad any \quad X(t) \end{aligned} \tag{22}$$

Let:

$$\begin{aligned} g[X(t),u(t)] &= X^T(t)QX(t)+u^T(t)Ru(t) \\ h[X(t),u(t)] &= A(t)X(t)+Bu(t) \\ u_t &\cong u(t;t\in[t,t+\delta]) \end{aligned} \tag{23}$$

Let us define:

$$\begin{aligned} f[X(t),T-t] &= Min_{[u(t;t\in[t,T])}J[t,T-t;X(t),u(t;t\in[t,T]) \\ &= J^*[t,T-t;X(t)] \end{aligned} \tag{24}$$

In other wards, the optimal value of the quadratic performance, i.e. f, is a continuous function of C and the control horizon (i.e. $T-t$). Let $[t,T]=[t,t+\delta]\bigcup[t+\delta,T]$ and $X(t)=X_t$, then, according to the *Principle of Optimality*, we can write:

$$f[X_t,T-t]=Min_{u(t;t\in[t,t+\delta])}\{r[X_t,u_t]+f(X_{t+\delta},T-\delta)\} \tag{25}$$

Where, $r[X_t,u_t]$ is a return function of the stage at the beginning $[t,t+\delta]$ interval. From Eq.(1), we have:

$$X_{t+\delta}=X_t+h[X_t,u_t]\delta \tag{26}$$

Then,

$$f(X_t,T-t)=Min_{u_t;t\in[t,t+\delta]}\{g[X_t,u_t]\delta+f(X_t+h[X_t,u_t]\delta,T-t-\delta)\} \tag{27}$$

Let $\tau=T-t$, $f(X_t,T-t)$ becomes $f(X_t,\tau)$. Since $f(X_t,\tau)$ is a continuous function of X_t and τ, we have:

$$\begin{aligned} f(X_t+h[X_t,u_t]\delta,\tau-\delta) &= f(X_t+\delta X_t,\tau-\delta)=f(X_t,\tau)+[\frac{\partial f(X_t,\tau)}{\partial X_t}]\dot{h}[X_t,u_t]\delta \\ & \quad +[\frac{\partial f(X_t,\tau)}{\partial\tau}]\dot{(}-\delta) \end{aligned} \tag{28}$$

$$f[X_t,\tau] = Min_{u(t);t\in[t,t+\delta]}\{g[X_t,u_t]\delta+f(X_t,\tau)+[\frac{\partial f(X_t,\tau)}{\partial X_t}]h[X_t,u_t]\delta+[\frac{\partial f(X_t,\tau)}{\partial \tau}](-\delta)\} \quad (29)$$

Or,

$$[\frac{\partial f(X_t,\tau)}{\partial \tau}] = Min_{u_t,t\in[t,t+\delta]}\{g[X_t,u_t] + [\frac{\partial f(X_t,\tau)}{\partial X_t}]h[X_t,u_t]\} \quad (30)$$

By defining the contents in the main parentheses of the right hand side as H, we have:

$$[\frac{\partial f(X_t,\tau)}{\partial \tau}] = Min_{[u_t;t\in[t,t+\delta]}H[\ X_t,u_t,\frac{\partial f(X_t,\tau)}{\partial X_t}]\ ] \quad (31)$$

The equation given above is known as the *Hamilton-Jacobin* Equation. The necessary condition of Eq.[32] ,if no constraint is set on u_0, is equivalent to :

$$\frac{\partial H}{\partial u_t}\{X_t,u_t^*,\ [\frac{\partial f(X_t,\tau)}{\partial X_t}\} = 0 \quad (32)$$

The only variable to be adjusted to obtain $Min[H]$ is u_t, we have to differentiate H with respect to u_t,i.e.,

$$[\frac{\partial g}{\partial u_t}]_{u_t^*} + [\frac{\partial h}{\partial u_t}]_{u_t^*}[\frac{\partial f(X_t,\tau)}{\partial X_t}]] = 0 \quad (33)$$

By substituting u_t^* into $H[X_t,u_t,\ \frac{\partial f(X_t,\tau)}{\partial X_t}]$, we have:

$$\{H^*[X_t,\tau]\} = Max\{H[\ X_t,u_t^*,\ \frac{\partial f(X_t,\tau)}{\partial X_t}\ ]\} \quad (34)$$

So that,

$$[\frac{\partial f(X_t,\tau)}{\partial \tau}]_{X_t,\tau} - H[\ X_t,u^*{}_t,\ [\frac{\partial f(X_t,\tau)}{\partial X_t}]\ ] = 0 \quad (35)$$

By the definitions of h, g, and H, we have:

$$H[\ X_t,u^*{}_t,\ [\frac{\partial f(X_t,\tau)}{\partial X_t}]\ ] = \frac{1}{2}[X_t^TQX_t + u_t^TRu_t + (\frac{\partial f}{\partial X_t})^T(AX_t+Bu_t)] \quad (36)$$

We can move the origin of time from 0 to t, so that the corresponding state, C, and the corresponding U_0 are replaced by $X(t)$ and $u(t)$. Then, H becomes:

$$H\{\ X_t, u_t,\ [\frac{\partial f(X_t,\tau)}{\partial X_t}]_{X_t,\tau}\ \} = \frac{1}{2}\{X_t^T Q X_t + u_t^T R u_t + (\frac{\partial f}{\partial X})^T (AX_t + Bu_t)\} \tag{37}$$

and,

$$[\frac{\partial f(X_t,\tau)}{\partial \tau}]_{X_t,\tau} = Min_{u_t} H[\ X_t, u_t,\ [\frac{\partial f(X_t,\tau)}{\partial X_t}]_{X_t,\tau}\ ] \tag{38}$$

If T is fixed, $d\tau = -dt$ and $f(X_t,\tau) = f(X_t,t)$

$$-[\frac{\partial f(X_t,t)}{\partial (t)}]_{X_t,t} = Min_{u_t} H[\ X_t, u_t, \frac{\partial f(X_t,t)}{\partial X_t}]_{X_t,t}\ ] \tag{39}$$

or,

$$[\frac{\partial f(X,t)}{\partial (t)}]_{X,t} + Min_{[u(t)]} H[\ X_t, u_t, \frac{\partial f(X_t,t)}{\partial X_t}\ ] = 0 \tag{40}$$

Remember that $\frac{\partial H}{\partial u_t}[X_t, u_t^*, [\frac{\partial f(X_t,t)}{\partial X_t}] = 0$, if H is to be minimized. From Eq.(35), we have:

$$Ru_t^* + B^T \frac{\partial f(X_t,t)}{\partial X_t} = 0 \tag{41}$$

Thus,

$$u_t^* = -R^{-1} B^T \frac{\partial f(X_t,t)}{\partial X_t} \tag{42}$$

Then, substitute u_t^* into H:

$$\begin{aligned} H[X_t, u_t, \frac{\partial f(X_t,t)}{\partial X_t}] &= \frac{1}{2} X_t^T Q X_t + \frac{1}{2} f_{X_t}^T B R^{-1} B^T f_{X_t} + f_{X_t}^T A X_t - f_{X_t}^T B R^{-1} B^T f_{X_t} \\ &= \frac{1}{2} X_t^T Q X_t - \frac{1}{2} f_{X_t}^T B R^{-1} B^T f_{X_t} + f_{X_t}^T A X_t \end{aligned} \tag{43}$$

Where,

$$f_{X_t} = \frac{\partial f(X_t, t)}{\partial X_t}$$

Thus, the Jacobin-Hamiltonian equation becomes:

$$f_t + \frac{1}{2}X_t^T Q X_t - \frac{1}{2}f_{X_t}^T B R^{-1} B f_{X_t} + f_{X_t}^T A X_t = 0 \tag{44}$$

$$f[X_T, 0] = \frac{1}{2}X_T^T Q_1 X_T$$

where

$$f_t = \frac{\partial f(X_t, t)}{\partial t}$$

Since $f[X_t, T-t]$ is a result of the quadratic performance index and a state feedback control input, it can be represented as:

$$\begin{aligned} f[X_t, T-t] &= \frac{1}{2}X_t^T P^*(T-t, K^*(t))X_t \\ &= \frac{1}{2}X_t^T P^*(T-t)X_t \\ &= \leq \frac{1}{2}X_t^T P(T-t, K(t; t\in[t, T-t]))X_t \end{aligned} \tag{45}$$

Noted that:

$$\begin{aligned} f_t &= \frac{1}{2}X_t^T \dot{P}(T-t)X_t \\ f_{X_t} &= P(T-t)X_t \\ P(T-t)A &= \frac{1}{2}[P(T-t)A + (P(T-t)A)^T] + \frac{1}{2}[P(T-t)A - (P(T-t)A)^T] \end{aligned} \tag{46}$$

and,

$$\begin{aligned} X_t^T P(T-t)AX_t &= \frac{1}{2}X_t^T[P(T-t)A + A^T P^T(T-t)]X_t + \frac{1}{2}X_t^T[P(T-t)A - A^T P^T(T-t)]X_t \\ &= \frac{1}{2}X_t^T P(T-t)AX_t + \frac{1}{2}X_t^T A^T P(T-t)X_t \end{aligned} \tag{47}$$

Eq.(44) becomes:

$$\frac{1}{2}X_t^T\{\dot{P}+Q-PBR^{-1}B^TP+PA+A^TP\}X_t=0 \tag{48}$$

Since the above equation holds for any $X(t)$, we conclud that:

$$\dot{P}+Q-PBR^{-1}B^TP+PA+A^TP=0 \quad ; \quad P(T)=Q_1 \tag{49}$$

This equation is known as the Riccati equation for solving the LQ state feedback control problems. The optimal state feedback control becomes:

$$\begin{aligned} u_t^* &= -R^{-1}B^Tf_{X_t} \\ &= -R^{-1}(t)B^T(t)P(t)X_t \\ &= -K^*(t)X_t \end{aligned} \tag{50}$$

In case $T \to \infty$, X and u will converge to zero so that there is no necessary to put penaulty of $X(T)$ in the objective function. Consequently, the matrix Q_1 of Eq.(13) can be given as zero. If the dynamic system in Eq.(1) is a time invariant one, the result of $P(\infty,t)$, or simply $P(t)$, becomes a constant matrix. i.e.,

$$u^*(t)=-K^*X(t) \tag{51}$$

We shall explore some asymptotic properties of such a regulatiory control a little bit further in the following.

1. Extension of the LQR problem to LQ-Servo Problem

For a constant command input,r, let us define:

$$
\begin{aligned}
\tilde{u} &= u_\infty \\
\tilde{X} &= X - X_\infty \\
\tilde{y} &= y - y_\infty
\end{aligned}
\tag{52}
$$

Where, u_∞, X_∞, and y_∞ are the state steady values of u, X, and y corresponding to r. Then, the system in Eq.(1) can be changed to the following:

$$
\begin{aligned}
\dot{\tilde{X}} &= A\tilde{X} + B\tilde{u} \\
\tilde{y} &= C\tilde{X}
\end{aligned}
\tag{53}
$$

The performance index of Eq.(13) is rewritten as:

$$
J(t,\infty) = \int_t^\infty [\tilde{X}^T(\tau)Q\tilde{X}(\tau) + \tilde{u}^T(\tau)R\tilde{u}(\tau)]d\tau \tag{54}
$$

and the corresponding optimal control law is:

$$
\tilde{u}(t) = -K^*\tilde{X}(t)
$$

Thus,

$$
\begin{aligned}
u &= u_\infty - K^*[X(t) - X_\infty\,] \\
&= -K^*X(t) + u_\infty + K^*X_\infty \\
&= -K^*X(t) + u'_\infty
\end{aligned}
\tag{55}
$$

To determine u'_∞, we start with substituting $u(t)$ of the above equation into Eq.(1) and let t approach ∞, we have:

$$
X_\infty = -[A - BK^*]^{-1}Bu'_\infty
$$

$$
\begin{aligned}
y_\infty &= -C[A - BK^*]^{-1}Bu'_\infty \\
or, & \\
u'_\infty &= -[C(A - BK^*)^{-1}B]^{-1}y_\infty \\
&= \{\ C[sI - (A - BK^*)]^{-1}B\ \}^{-1}_{s=0}\ y_\infty \\
& \quad \{\ H_c(o)\ \}^{-1}y_\infty \qquad (56) \\
so & \\
u &= -K^*X(t) + H_c^{-1}(0)y_\infty \qquad (57)
\end{aligned}
$$

2. Robustness of LQR loop

According to the block diagram as shown in Fig.[ ], the loop transfer function of a state feedback system becomes:

$$G_{LQ} = K(sI - A)^{-1}B$$

[Theorem] Kalman Equality

If G_{LQ} is the loop transfer function of an optimal LQR then:

$$[I + G_{LQ}(-s)]^T R[I + G_{LQ}(s)] = R + G^T_{OL}(-s)G_{OL}(s) \qquad (58)$$

$$G_{OL} = C(sI - A)^{-1}B$$

$$G_{CL} = K(sI - A)^{-1}B$$

[**Proof**]

Starting with the Riccati equation:

$$-[\ PA + A^T P + Q - PBR^{-1}B^T P\ ] = 0$$

$$sP - PA - sP - A^T P - Q + PBR^{-1}B^T P\ ] = 0$$

Thus,

$$P[\ sI - A\ ] + [\ -sI - A\ ]^T P - Q + -PBR^{-1}B^T P\ ] = 0$$

Then, premultiply Eq.(56) by $B^T(-sI - A)^{-T}$, and let $\Phi(s) = (sI - A)$, we have:

$$B^T\Phi^{-T}(-s)P\Phi(s) + B^T\Phi^{-T}(-s)\Phi(-s)^T P - B^T\Phi^{-T}(-s)Q$$

$$+ B^T\Phi^{-T}(s)PBR^{-1}B^T P = 0$$

So,

$$B^T\Phi^{-T}(-s)P\Phi(s) + B^T P - B^T\Phi^{-T}(-s)Q + B^T\Phi^{-T}(s)PBR^{-1}B^T P = 0 \tag{59}$$

Again, postmultiply the both sides of the above equation by $\Phi^{-1}(s)B$ to give:

$$B^T\Phi^{-T}(-s)PB + B^T P\Phi^{-1}(s)B - B^T\Phi^{-T}(-s)Q\Phi(s)^{-1}B$$

$$+ B^T\Phi^{-T}(-s)PBR^{-1}B^T P\Phi(s)^{-1}B = 0 \quad (58)$$

Let $Q = C^T C$, then,

$$G_{OL}^T(-s)G_{OL}(s) = B^T\Phi^{-T}(-s)C^T C\Phi^{-1}(s)B = B^T\Phi^{-T}(-s)Q\Phi^{-1}(s)B \tag{60}$$

Also,

$$B^T P\Phi^{-1}(s)B = RR^{-1}B^T P\Phi^{-1}(s)B = RK^*\Phi^{-1}(s)B = RG_{LQ} \tag{61}$$

$$B^T\Phi^{-1}(s)PB = G_{LQ}^T(s)R$$

and

$$PBR^{-1}B^T P = PBR^{-1}RR^{-1}B^T P = K^T RK$$

Thus, Eq.(58) becomes:

$$G_{LQ}^T(-s)R + RG_{LQ}(s) - G_{OL}^T(-s)G_{OL}(s) + G_L Q^T(-s)RG_{LQ}(s) = 0 \tag{62}$$

so that

$$G_{LQ}^T(-s)R + RG_{LQ}(s) + G_L Q^T(-s)RG_{LQ}(s) = G_{OL}^T(-s)G_{OL}(s) \tag{63}$$

and By adding R to the both sides of the above equation, we have:

$$R + G_{LQ}^T(-s)R + RG_{LQ}(s) + G_L Q^T(-s)RG_{LQ}(s) = R + G_{OL}^T(-s)G_{OL}(s) \tag{64}$$

namely,

$$[I + G_{LQ}(-s)]^T R[I + G_{LQ}(s)] = R + G_{OL}^T(-s)G_{OL}(s) \tag{65}$$

[**Theorem**]

$$\underline{\sigma}[\ I + G_{LQ}\ ] \geq 1 \qquad (A)$$

$$\underline{\sigma}[\ I + G_{LQ}^{-1}\ ] \geq \frac{1}{2} \qquad (B)$$

[**Proof**] From the Kalman equality:

$$[I + G_{LQ}(-s)]^T R[\ I + G_{LQ}(s)\ ] = R + G_{OL}^T(-s)G_{OL}(s) \tag{66}$$

IF we substitute s by $j\omega$ and $R = 1$, then:

$$[I + G_{LQ}(j\omega)]^*[\ I + G_{LQ}(j\omega)\ ] = I + G^*_{OL}(j\omega)G_{OL}(j\omega) \tag{67}$$

$$\begin{aligned} \sigma_i[\ I + G_{LQ}\ ] &= \lambda_i\{[\ I + G_{LQ}\ ]^*[\ I + G_{LQ}\ ]\} \\ &= \lambda_i\{I + G^*_{OL}G_{OL}\} \\ &= 1 + \lambda_i\{G^*_{OL}G_{OL}\} \\ &\geq 1 \quad (\quad Because, \lambda_i\{G^*_{OL}G_{OL}\} \geq 0\ \ ) \end{aligned} \tag{68}$$

Let $A = G_{LQ}$,

$$\begin{aligned} I &= (I + A)^{-1}(I + A) \\ &= (I + A)^{-1} + (I + A)^{-1}A \\ &= (I + A)^{-1} + (I + A^{-1})^{-1} \end{aligned} \tag{69}$$

Thus,

$$(I + A^{-1})^{-1} = I - (I + A^{-1})^{-1}$$

Therefore,

$$\begin{aligned} \bar{\sigma}[(I + A^{-1})^{-1}] &= \bar{\sigma}[I - (I + A)^{-1}] \\ &\leq \bar{\sigma}[I] + \bar{\sigma}[(I + A)^{-1}] \\ &= 1 + \bar{\sigma}[\ (I + A)^{-1}\ ] \\ &= 1 + \underline{\sigma}^{-1}[\ I + A\ ] \\ &\leq 2 \end{aligned} \tag{70}$$

Which implies:

$$\underline{\sigma}^{-1}[\ I + A^{-1}\ ] \leq 2$$

or,

$$\underline{\sigma} = \underline{\sigma}[\ I + G^{-1}_{LQ}\ ] \geq \frac{1}{2} \tag{71}$$

From (A), it shows that the LOR controller has an ∞ gain margin. On the other hand, from (B) it shows that the LQR control system guarantees that this system will remain stable to each modeling error for which $|\Delta G(j\omega)/G(j\omega)| = \ell(\omega)$ never exceeds 0.5.

In the LQR problem, it is assumed that all state variables are accessible from the plant. In practical occasions, such an assumption is far from reality. It is thus necessary to incorporate state estimations to facilitate the use of state feedback. One of the state estimation can be accomplished through the use of Luenberger Observer.

Consider a dynamic system:

$$\dot{x} = AX + Bu \quad ; \quad X(t_0) = X_0 y = CX \tag{72}$$

The estimation of state is given as follows:

$$\dot{\hat{x}} = A\hat{X} + Bu + H(y - \hat{X}) \;\; ; \;\; \hat{X}(t_0) = \hat{X}_0 \hat{y} = C\hat{X} \tag{73}$$

Where, H is to be designed. The differences between X and $\hat{X}$ is denoted as the estimation error and is given, by making use of the above equation, as :

$$\dot{e} = (A - HC)e \;\; ; \quad e(t_0) = X_0 - \hat{X}_0 \tag{74}$$

If $[A, C]$ is an observable pair, H can be choosen to give assigned dynamic modes. One of the question raised, in the use of observer, is: *Can we decouple the control problem from the estimation one?* The answer to this question is given by making use of the following: The dynamic of the system is characterized by :

$$\frac{d}{dt}\left[\begin{array}{c} X \\ \hat{X} \end{array} \right] = \left[\begin{array}{cc} A & -BK \\ HC & A - BK - HC \end{array} \right]\left[\begin{array}{c} X \\ \hat{X} \end{array} \right] \tag{75}$$

Let us define a transformation for $[X, \hat{X}]^T$:

$$\left[\begin{array}{c} X \\ Z \end{array} \right] = \left[\begin{array}{cc} I & 0 \\ I & -I \end{array} \right]\left[\begin{array}{c} X \\ \hat{X} \end{array} \right] = T\left[\begin{array}{c} X \\ \hat{X} \end{array} \right] \tag{76}$$

So that,

$$\frac{d}{dt}\left[\begin{array}{c} X \\ \hat{X} \end{array} \right] = T\left[\begin{array}{c} \dot{X} \\ \dot{Z} \end{array} \right] = \left[\begin{array}{cc} A & -BK \\ HC & A - BK - HC \end{array} \right]\left[\begin{array}{cc} I & 0 \\ I & -I \end{array} \right]\left[\begin{array}{c} X \\ Z \end{array} \right] \tag{77}$$

In other words, we have:

$$\frac{d}{dt}\left[\begin{array}{c} X \\ Z \end{array} \right] = \left[\begin{array}{cc} I & 0 \\ I & -I \end{array} \right] \left[\begin{array}{cc} A & -BK \\ HC & A - BK - HC \end{array} \right] \left[\begin{array}{cc} I & 0 \\ I & -I \end{array} \right] \left[\begin{array}{c} X \\ Z \end{array} \right] \tag{78}$$

Which turns out to be:

$$\frac{d}{dt}\left[\begin{array}{c} X \\ Z \end{array} \right] = \left[\begin{array}{cc} A - BK & BK \\ 0 & A - HC \end{array} \right] \left[\begin{array}{c} X \\ Z \end{array} \right] \tag{79}$$

Since the characteristic of the system is not affected by any similar transformation, the characteristic equation of the system can be written as:

$$det\left[\begin{array}{cc} A - BK & BK \\ 0 & A - HC \end{array} \right] = 0 \tag{80}$$

But,

$$det\left[\begin{array}{cc} A - BK & BK \\ 0 & A - HC \end{array} \right] = det[\ A - BK\ ]\ det[\ A - HC\ ] \tag{81}$$

Therefore, zeros of Eq.(83) are zeros of either of the following:

$$det[\ A - BK] = 0$$

or,

$$det[\ A - HC\ ] = 0$$

Thus, design of control and of observer can be fully decoupled. This fact for linear system is known as Separation Theorem.

In many occasions, it is considered that uncertainties exist both in state dynamics and in the ouput measuring, the state estimation under such circumstances is more complicated. If these uncertainties are considered resulting from so called *white noise*, the whole problem can be formulated in terms of **LQG** problem which, conceptually, similar to the LQR excepted that the observer is replaced by a so called **Kalman Filter**. A system dynamic model is given by:

$$\begin{aligned} \dot{X} &= AX(t) + Bu(t) + w(t) \\ y(t) &= CX(t) + v(t) \end{aligned} \tag{82}$$

Where, w and v are white noise with covariance matrix Ω and V,respectively.

The LQG (Linear Quadratic Gaussian) control problem is described as:

given the measure, y(t) for $t\epsilon[0,t]$ *find* $u(\tau)$, $\tau\epsilon[t,t_1]$ *such that*

$$J = E\{X^T(t_1)Q_1X(t_1) + \int_t^{t_1}[X^TQX + u^TRud\tau|y(\eta), 0 \leq \eta \leq t\} \tag{83}$$

is minimized

The optimal state estimaor is called *Kalman filter*,i.e:

$$\dot{\hat{x}} = A\hat{X} + Bu + H(y - \hat{X}) \;\; ; \;\; \hat{X}(t_0) = \hat{X}_0; \quad \hat{y} = C\hat{X} \tag{84}$$

Where,

$$H = PC^TV^{-1}$$

and P is the solution of the Riccati equation:

$$\dot{P}(t) = AP(t) + PA^T(t) + \Omega - P(t)CV^{-1}CP(t) \;\; ; \quad P(t_0) = P_0 \tag{85}$$

$$P_0 = E\{(\hat{X} - X_0)(\hat{X} - X_0)^T\}$$

The state feedback control law becomes:

$$u = -K\hat{X}$$

The system is then augmented as:

$$\frac{d}{dt}\left[\begin{array}{c} X \\ \hat{X} \end{array}\right] = \left[\begin{array}{cc} A & -BK \\ HC & A-BK-HC \end{array}\right]\left[\begin{array}{c} X \\ \hat{X} \end{array}\right] \tag{86}$$

Which is almost the same as Eq.(78). It then can be shown also that **Separation Theorem** holds,too. The LQG system is shown in FIg[ ]. It can be shown that the transfer function that corresponding to the controller in the feedback loop is:

$$G_c(s) = K(sI - A - BK - HC)^{-1}H \tag{87}$$

If G(s) is a minimum phase plant, it has been shown that loop transfer fuction of this LQG system, i.e. $G(s)G_c(s)$ approaches to $C[sI - A)^{-1}H$ which is knon as G_F,i.e,

$$lim_{R\to 0}C(sI - A)^{-1}BK(sI - A - BK - HC)^{-1}H \to C(sI - A)^{-1}H = G_F \tag{88}$$

Thus, one can design H, the filter gain, to achieve a desirable G_F and treat it as a target loop transfer function. Having the target G_F, he can proceed to design K, the state feedback gain, and try to recover this G_F with its loop transfer function, i.e. $G(s)G_c(s)$. This design methodology is known as **LQG-LTR design**.

LQR with constant disturbance

The LTI (linear time-invarying) state model:

$$\dot{X} = AX + Bu + B_d d \quad ; \quad y = CX \tag{89}$$

$$J = \frac{1}{2}\int_{t_0}^{t_1} [\ y^T Q y + u^T R u\ ] dt \tag{90}$$

Where, d is a constant disturbance. The solution to this problem starts with finding a control law in the form of following:

$$u = -R^{-1}B^T P(t)(t)X(t) - R^{-1}B^T \xi(t)$$

Where, P and ξ are obtained from the following equations:

$$\dot{P} + PA + A^T P - PBR^{-1}BP + C^T QC = 0 \quad ; \quad P(t_1) = 0 \tag{91}$$

$$\dot{\xi} + [A^T - PBR^{-1}B^T]\xi - PB_d d = 0 \quad ; \quad \xi(t_1) = 0 \tag{92}$$

IF $t_1 \cong \infty$, P and ξ are steady state solutions:

$$A^T P + PA - PBR^{-1}BP + C^T QC = 0 \tag{93}$$

$$\xi = -(\ A^T - R^{-1}BPB^T\ )^{-1} PB_d d \tag{94}$$

In other words, the control law contains both feedback states and feedforward disturbance, i.e.:

$$u = -KX(t) + K_{FF} d$$

with

$$K = R^{-1}B^T P$$

$$K_{FF} = R^{-1}B^T[\ A^T - PBR^{-1}B^T\ ]^{-1} PB_d$$

Output Feedback

For the system

$$\dot{X} = AX + Bu \quad ; \qquad y = CX \tag{95}$$

There are two alternatives to perform the output feedback control:

1. Constant gain output feedback:

   $$u = -Ky = -KCX$$

   The closed loop system becomes:

   $$\dot{X} = (A - BKC)X$$

2. Dynamic output feedback:

   $$u = Hz + Ny$$

   $$\dot{z} = Fz + GY$$

   The closed loop system is:

   $$\dot{\bar{X}} = \bar{A} + \bar{B}\bar{K}\bar{C}\bar{X}$$

   $$\bar{X} = [\begin{matrix} X \\ z \end{matrix}]$$

   $$\bar{A} = [\begin{matrix} A & 0 \\ 0 & 0 \end{matrix} \qquad \bar{B} = [\begin{matrix} B & 0 \\ 0 & I \end{matrix}] \tag{96}$$

   $$\bar{C} = [\begin{matrix} C & 0 \\ 0 & I \end{matrix}] \qquad \bar{K} = [\begin{matrix} N & H \\ G & F \end{matrix}] \tag{97}$$

To minimize the quadratic performance as given in Eq.(13), the approach via the solution of Riccati equation of LQR problem no longer applies unless KC (or, $\bar{K}\bar{C}$) is a square matrix. As a result, finding optimal $K, (or\bar{K})$ that minimize J becomes a parametric optimization problem. Standard computer routines for numerical parameter optimation are often used in these cases. An alternative way to find KC (or, $\bar{K}\bar{C}$) is to minimize $trace[P]$. Where, P is the cost matrix that defines the quadratic performance from the state feedback problem. [In case of dynamic matrix, P will be replaced by $\bar{P}$ such that:

$$J = \frac{1}{2}\bar{X}(0)^T\bar{P}\bar{X}(0)$$

. There are still many other ways to solve this kind of output feedback problem.

In the previous sections, we have learned the LQR problem for designing a state feedback control system. However, the LQR formulation should based on a model with state space representation. It is thus necessary to know how to realize a transfer function in terms of state space variables. Canonical forms Following are some canonical forms that directly corresponding to linear transfer functions:

1. Controller canonical form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdot & \cdot & \cdot & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} u \tag{98}$$

$$y = \begin{bmatrix} b_n & b_{n-1} & \cdot & \cdot & \cdot & b_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{99}$$

[Illustration]

$$\frac{y(s)}{u(s)} = G(s) = \frac{b_1 s^2 + b_2 s + b_1}{s^3 + a_1 s^2 + a_2 s + a_3} \tag{100}$$

or,

$$\dddot{y} + a_1 \ddot{y} + a_2 \dot{y} + a_3 y = b_1 \ddot{u} + b_2 \dot{u} + b_3 u \tag{101}$$

$$\frac{y(s)}{u(s)} = \frac{y(s)}{z(s)} \frac{z(s)}{u(s)} = \left[b_1 s^2 + b_2 s + b_1 \right] \frac{1}{s^3 + a_1 s^2 + a_2 s + a_3} \tag{102}$$

Therefore,

$$\dddot{z} + a_1\ddot{z} + a_2\dot{z} + a_3 z = u$$

$$b_1\ddot{z} + b_2\dot{z} + b_3 z = y$$

Let

$$\begin{aligned}
x_1 &= z \\
\dot{x}_1 &= \dot{z} = x_2 \\
\dot{x}_2 &= \ddot{z} = x_3 \\
\dot{x}_3 &= \dddot{z} \\
&= -a_1\ddot{z} - a_2\dot{z} - a_3 z + u \\
&= -a_1 x_3 - a_2 x_2 - a_3 x_1 + u \\
y &= b_1 x_3 + b_2 x_2 + b_3 x_1
\end{aligned} \tag{103}$$

so that

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_{2-1} & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \tag{104}$$

$$y = \begin{bmatrix} b_3 & b_2 & b_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{105}$$

2. Observer canonical form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -a_1 & 1 & 0 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_{n-1} & 0 & 0 & 0 & \cdots & 1 \\ -a_n & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} u \tag{106}$$

$$y = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{107}$$

[Illustration]

$$\frac{y(s)}{u(s)} = G(s) = \frac{b_1 s^2 + b_2 s + b_1}{s^3 + a_1 s^2 + a_2 s + a_3} \tag{108}$$

or,

$$\dddot{y} + a_1 \ddot{y} + a_2 \dot{y} + a_3 y = b_1 \ddot{u} + b_2 \dot{u} + b_3 u \tag{109}$$

The above equation is first rewritten as:

$$\begin{aligned} y &= s^{-1}[\ -a_1 y + b_1 u\ ] + s^{-2}[\ -a_2 y + b_2 u\ ] + s^{-3}[\ -a_3 y + b_3 u\ ] \\ &= s^{-1}\left[\ s^{-1}\left[\ s^{-1}(-a_3 y + b_3 u) + (-a_2 y + b_2 u)\right] + (-a_1 y + b_1 u)\right] \end{aligned} \tag{110}$$

Let us define the following:

$$\begin{aligned} \dot{x}_1 &= -a_1 y + b_1 u + x_2 = -a_1 x_1 + b_1 u + x_2 \\ \dot{x}_2 &= -a_2 y + b_2 u + x_3 = -a_2 x_1 + b_2 u + x_3 \\ \dot{x}_3 &= -a_3 y + b_3 u = -a_3 x_1 + b_3 u \\ y &= x_1 \end{aligned} \tag{111}$$

or,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} u \tag{112}$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{113}$$

3. Controllability canonical form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & -a_n \\ 1 & 0 & 0 & 0 & \cdots & -a_{n-1} \\ 0 & 1 & 0 & 0 & \cdots & -a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & 0 & -a_2 \\ 0 & \cdot & \cdot & 0 & 1 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} u \tag{114}$$

$$y = \begin{bmatrix} b_n & b_{n-1} & \cdot & \cdot & \cdot & b_1 \end{bmatrix} \begin{bmatrix} a_{n-1} & a_{n-2} & \cdot & \cdot & a_1 & 1 \\ a_{n-2} & a_{n-3} & \cdot & a_1 & 1 & 0 \\ a_{n-3} & \cdot & a_1 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{115}$$

Let us define $\mathbf{C}[A, B]$ and $\mathbf{O}[A, C]$ as the controllability matrix and observability matrix of the system (A,B,C). We will use the following definition for formulation:

$$\begin{aligned} \mathbf{C}[\bullet, \bullet] &= \textit{controllability matrix constructed by the matrices in the parantheses} \\ \mathbf{O}[\bullet, \bullet] &= \textit{Observability matrix constructed by the matrices in the parantheses} \end{aligned}$$

$$\begin{aligned} &A_c, B_c, C_c \cdots\cdots for\ controller\ canonical\ form \\ &A_o, B_o, C_o \cdots\cdots for\ observer\ canonical\ form \\ &A_{co}, B_{co}, C_{co} \cdots\cdots for\ controllability\ canonical\ form \\ &A_{ob}, B_{ob}, C_{ob} \cdots\cdots for\ observability\ canonical\ form \end{aligned}$$

One important property of this controllability canonical form is that

$\mathbf{C}[A_{co}, Bco] = I$

Assume we have adynamic system represented by $[A, B, C]$. Then we have:

$\mathbf{C}[A, B] = [B, AB, A^2B, A^3B, \cdots, A^{n-1}B]$

If there is a transformation matrix, P, such that:

$$Z = PX$$

Then the triplet matrices that represent the system in terms of Z becomes: $[\bar{A}, \bar{B}, \bar{C}]$:

$$\begin{aligned} \bar{A} &= PAP^{-1} \\ \bar{B} &= PB \\ \bar{C} &= CP^{-1} \end{aligned} \tag{116}$$

It is easy to show that:

$$\mathbf{C}[\bar{A}, \bar{B}] = P\mathbf{C}[A, B]$$

and

$$\mathbf{O}[\bar{A}, \bar{C}] = \mathbf{O}[A, C]p^{-1}$$

Therefore, to contruct the controllabilty form for $[A, B, C]$, we need to set:

$$P = \mathbf{C}[A, B]^{-1} \tag{117}$$

As a special case, if we start with $[A_c, B_c, C_c]$, i.e. a controller canonical form, the transformation that leads $[A_c, B_c, C_c]$ to a controllability form is (for example, n=3):

$$\begin{aligned} P &= \mathbf{C}[A_c, B_c]^{-1} \\ &= [B_c, A_cB_c, A_c^2B_c]^{-1} \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_1 \\ 1 & -a_1 & -a_2 + {a_1}^2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} a_2 & a_1 & 1 \\ a_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned} \tag{118}$$

Thus,

$$
\begin{aligned}
A_{co} &= \begin{bmatrix} a_2 & a_1 & 1 \\ a_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_1 \\ 1 & -a_1 & -a_2 + {a_1}^2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 0 & 0 & -a_3 \\ 1 & 0 & -a_2 \\ 0 & 1 & -a_1 \end{bmatrix} \\
B_{co} &= \begin{bmatrix} a_2 & a_1 & 1 \\ a_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
C_{co} &= \begin{bmatrix} b_3 & b_2 & b_1 \end{bmatrix} \begin{bmatrix} a_2 & a_1 & 1 \\ a_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1}
\end{aligned} \tag{119}
$$

4. Observability canonical form:

Similar to the definition of the controllability canonical form, the obseravability canonical form $[A_{ob}, B_{ob}, C_{ob}]$ has unity $\mathbf{O}[A_{ob}, C_{ob}]$. Since for the trnasformation:

$$Z = PX$$

we have:

$$\mathbf{O}[\bar{A}, \bar{C}] = \mathbf{O}[A, C]P^{-1}$$

so that the required transformation matrix P is given by:

$$P = \mathbf{O}[A, C]$$

and a general obseravability canonical form is as follows:

$$
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdot & \cdot & \cdot & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ a_1 & 1 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-3} & \cdot & a_1 & 1 & 0 & 0 \\ a_{n-2} & a_{n-3} & \cdot & a_1 & 1 & 0 \\ a_{n-1} & a_{n-2} & \cdot & \cdot & a_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} u \tag{120}
$$

$$y = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{121}$$

As a special case, (say, n=3), we have:

$$\begin{aligned} \mathbf{O}[A_o, C_o] &= \begin{bmatrix} C_o \\ C_oA_o \\ C_oA_o{}^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ -a_1 & 1 & 0 \\ a_1^2 - a_2 & -a_1 & 1 \end{bmatrix} \end{aligned} \tag{122}$$

Note that:

$$\begin{aligned} \mathbf{O}[A_o, C_o]^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ -a_1 & 1 & 0 \\ a_1^2 - a_2 & -a_1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & a_1 & 1 \end{bmatrix} \end{aligned} \tag{123}$$

If we premultiply $\mathbf{O}[A_o, C_o]^{-1}$ by $\tilde{I}$, we have:

$$\begin{aligned} \tilde{I}\mathbf{O}[A_o, C_o]^{-1} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -a_1 & 1 & 0 \\ a_1^2 - a_2 & -a_1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} a_2 & a_1 & 1 \\ a_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \mathbf{C}[A_c, B_c]^{-1} \end{aligned} \tag{124}$$

In other words,

$$\mathbf{C}[A_c, B_c] = \mathbf{O}[A_o, C_o]\ \tilde{I}$$

$$\mathbf{C}[A_c, B_c]\ \tilde{I} = \mathbf{O}[A_o, C_o]$$

We should also note that $\mathbf{C}[A_c, B_c]^{-1}$ and $\mathbf{O}[A_o, C_o]^{-1}$ are known as upper triangular and lower triangular Toeplitz matrices. That is:

$$\begin{aligned}
\mathbf{C}[A_c, B_c]^{-1} &= \begin{bmatrix} a_{n-1} & a_{n-2} & \cdot & \cdot & a_1 & 1 \\ a_{n-2} & a_{n-3} & \cdot & a_1 & 1 & 0 \\ a_{n-3} & \cdot & a_1 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
&= \mathrm{a}_+(a_1, a_2, \cdots, a_{n-1}) \\
\mathbf{O}[A_o, C_o]^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ a_1 & 1 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-3} & \cdot & a_1 & 1 & 0 & 0 \\ a_{n-2} & a_{n-3} & \cdot & a_1 & 1 & 0 \\ a_{n-1} & a_{n-2} & \cdot & \cdot & a_1 & 1 \end{bmatrix} \\
&= \mathrm{a}_-(a_1, a_2, \cdots, a_{n-1}) \\
\mathrm{a}_+ &= \tilde{I}\mathrm{a}_- \qquad (125)
\end{aligned}$$

[****] P for $[A, B, C] \to [A_c, B_c, C_c]$

Given a triplet [A, B, C] in the state space X, the transformation P, i.e. $Z = PX$, for $[A, B, C] \to [A_c, B_c, C_c]$ is given by:

$$P = [\mathrm{a}_+(\mathbf{h})]^{-1}\mathbf{C}[A, B]^{-1}$$

Where,

$$\mathbf{h} = [-\hat{a}_n, -\hat{a}_{n-1}, \cdots, -\hat{a}_2]$$

and, $\hat{a}_i$ is the i-th element of last column of $\mathbf{C}[A, B]^{-1}\ A\mathbf{C}[A, B]$.

[Proof]

The transformation is carried out in the sequence of:

$$[A, B, C] \to [A_{co}, B_{co}, C_{co}] \to [A_c, B_c, C_c] \qquad (126)$$

In the first step, we have:

$$P_1 = \mathbf{C}[A, B]^{-1}$$

which will result:

$$A_{co} = \begin{bmatrix} 0 & \hat{a}_1 \\ & \hat{a}_2 \\ \cdot & \cdot \\ I & \hat{a}_n \end{bmatrix} \tag{127}$$

Then, according to eauation (116), $P_2 = [\mathrm{a}_+(-\hat{a}_n, -\hat{a}_{n-2}, \cdot, \hat{a}_2)]^{-1}$ will transform the resulting $[A_{co}, B_{co}, C_{co}]$ into $[A_c, B_c, C_c]$. Thus, $P = P_2\ P_1$:

$$P = \mathrm{a}_+(-\hat{a}_n, -\hat{a}_{n-2}, \cdots, \hat{a}_2)]^{-1}\mathbf{C}[A, B]^{-1}$$

This proves the theorem.

In the same way, we can also conclude that the transformation P for $[A, B, C] \rightarrow [A_o, B_o, C_o]$ is:

$$P = \mathrm{a}_-(-\hat{a}_n, -\hat{a}_{n-2}, \cdots, \hat{a}_2)]\mathbf{O}[A, B]$$

Where, $\hat{a}_i$ is the i-th element of last row of $\mathbf{O}[A, B]\ A\mathbf{C}[A, B]^{-1}$.

5. Summary

1. $[A_c, B_c, C_c] \rightarrow [A_{co}, B_{co}, C_{co}]$

   $$P = \mathbf{C}[A_c, B_c, C_c]^{-1}$$

2. $[A_o, B_o, C_o] \rightarrow [A_{ob}, B_{ob}, C_{ob}]$

   $$P = \mathbf{C}[A_o, B_o, C_o]^{-1}$$

3. $[A_{co}, B_{co}, C_{co}] \rightarrow [A_{ob}, B_{ob}, C_{ob}]$

   $$P = \mathbf{O}[A_{co}, B_{co}, C_{co}]$$

4. $[A_{ob}, B_{ob}, C_{ob}] \rightarrow [A_{co}, B_{co}, C_{co}]$

$$P = \mathbf{C}[A_{ob}, B_{ob}, C_{ob}]^{-1}$$

5. $[A_{co}, B_{co}, C_{co}] \rightarrow [A_c, B_c, C_c]$

$$P = \mathrm{a}_+^{-1}(A_{co})$$

6. $[A_{ob}, B_{ob}, C_{ob}] \rightarrow [A_o, B_o, C_o]$

$$P = \mathrm{a}_-(A_{ob})]$$

7. $[A, B, C] \rightarrow [A_*, B_*, C_*]$

$$[A, B, C] \rightarrow \{[A_{co}, B_{co}, C_{co}] \;\; or \;\; [A_{co}, B_{co}, C_{co}]\} \rightarrow [A_*, B_*, C_*]$$