Derivation of Linear Quadratic Regulator (LQR) controller

Consider a linear system \( \dot{x} = Ax + Bu \) where \( x \in \mathbb{R}^n \) and \( u \in \mathbb{R}^m \). Consider a cost functional of the form

\[
V(t, T) = \int_t^T [x^T(\tau)Q(x(\tau)) + u^T(\tau)R(u(\tau))]d\tau + x^T(T)\Omega x(T)
\]  

(1)

where \( Q \) and \( \Omega \) are symmetric positive semi-definite matrices, \( R \) is a symmetric positive-definite matrix, and \( T \) is the terminal (final) time. If \( T = \infty \), this is called the infinite-horizon problem. If \( T \) is a finite quantity, then this is called the finite-horizon problem.

We will show two ways to derive the optimal controller (i.e., the LQR controller) for this cost functional. In the first method, we will consider the set of all possible controllers (which could include nonlinear as well as linear controllers) and we will show that the optimal controller is indeed a linear controller for this cost functional and we will derive the equation for this controller. In the second method, we will start with assuming that we are trying to find a linear controller and simply look for the best linear controller. Since in the second method, we are only considering linear controllers, the derivation will be much simpler than in the first method.

1 Derivation of LQR controller considering all possible (nonlinear/linear) controllers

Let the time signal \( u^*(\tau) \) over the time interval \( \tau \in [t, T] \) be the optimal controller that we are trying to find. If we consider a perturbation of this control input signal as \( u(\tau) = u^*(\tau) + \epsilon \tilde{u}(\tau) \), then the value of the cost functional with the perturbed control input signal \( u(\tau) \) cannot be lower than the value of the cost functional with the optimal control input signal \( u^*(\tau) \). Denote by \( x^*(\tau) \) the state trajectory when the control input signal \( u^*(\tau) \) is applied and denote by \( x(\tau) = x^*(\tau) + \epsilon \tilde{x}(\tau) \) the state trajectory when the control input signal \( u(\tau) \) is applied. Then,

\[
\dot{x}^*(\tau) = A(x^*(\tau) + B(u^*(\tau))
\]

(2)

\[
\dot{x}(\tau) = A(x(\tau) + B(u(\tau)))
\]

(3)

\[
\Rightarrow \dot{x}^*(\tau) + \epsilon \tilde{x}(\tau) = A(x^*(\tau) + \epsilon \tilde{x}(\tau)) + B(u^*(\tau) + \epsilon B(\tilde{u}(\tau)).
\]

(4)

Hence,

\[
\hat{x}(\tau) = A(x^*(\tau) + \epsilon \tilde{x}(\tau)) + B(u^*(\tau) + \epsilon B(\tilde{u}(\tau)).
\]

(5)

The value of the cost functional when the control input signal \( u^*(\tau) \) is applied is denoted by \( V^*(t, T) \) and is given by

\[
V^*(t, T) = \int_t^T [(x^*(\tau))^TQ(x^*(\tau)) + (u^*(\tau))^TR(u^*(\tau))]d\tau + (x^*(T))^T\Omega x^*(T).
\]

(6)

The value of the cost functional when the control input signal \( u(\tau) \) is applied is given by

\[
V(t, T) = \int_t^T [(x^*(\tau) + \epsilon \tilde{x}(\tau))^TQ(x^*(\tau) + \epsilon \tilde{x}(\tau)) + (u^*(\tau) + \epsilon \tilde{u}(\tau))^TR(u^*(\tau) + \epsilon \tilde{u}(\tau))]d\tau
\]

\[
+ (x^*(T) + \epsilon \tilde{x}(T))^T\Omega (x^*(T) + \epsilon \tilde{x}(T))
\]

\[
= V^*(t, T) + 2\epsilon \left\{ \int_t^T [\tilde{x}^T(\tau)Q(x^*(\tau) + \epsilon \tilde{x}(\tau))d\tau + \tilde{x}^T(T)\Omega x^*(T)] + \epsilon^2 \left\{ \int_t^T [\tilde{x}^T(\tau)Q(x^*(\tau) + \epsilon \tilde{u}(\tau)R(u^*(\tau))]d\tau + \tilde{x}^T(T)\Omega x^*(T) \right\} \right\}
\]

(7)

Note that we must have \( V(t, T) \geq V^*(t, T) \) for all values of \( \epsilon \) since otherwise \( u^*(\tau) + \epsilon \tilde{u}(\tau) \) would be a better control input signal (i.e., lower cost functional) than \( u^*(\tau) \) which is supposed to be the optimal control input signal. From (8), it is seen that the only way that we can have \( V(t, T) \geq V^*(t, T) \) for all values of \( \epsilon \) is if the coefficient in the \( 2\epsilon \) \{ \} term is zero. Hence,

\[
\int_t^T [\tilde{x}^T(\tau)Q(x^*(\tau) + \epsilon \tilde{u}(\tau)R(u^*(\tau))]d\tau + \tilde{x}^T(T)\Omega x^*(T) = 0.
\]

(9)
Denoting the transition matrix corresponding to the matrix \( A \) to be \( \Phi(t, t_0) \), we can write
\[
\dot{x}(\tau) = \Phi(\tau, t)\dot{x}(t) + \int_t^\tau \Phi(\tau, \tau_1)B(\tau_1)\dot{u}(\tau_1)d\tau_1.
\] (10)

Since both the signals \( x^*(\tau) \) and \( x^*(\tau) + \epsilon\dot{x}(\tau) \) should start (at time \( t \)) from the current value of \( x \), i.e., \( x(t) \), we see that \( \dot{x}(t) = 0 \). Hence,
\[
\dot{x}(\tau) = \int_t^\tau \Phi(\tau, \tau_1)B(\tau_1)\dot{u}(\tau_1)d\tau_1.
\] (11)

Substituting (11) into (9), we get
\[
0 = \int_t^T \int_t^\tau \dot{u}(\tau_1)B^T(\tau_1)\Phi^T(\tau, \tau_1)Q(\tau)x^*(\tau)d\tau_1d\tau + \int_t^T \dot{u}(\tau)R(\tau)u^*(\tau)d\tau
+ \int_t^T \dot{u}(\tau_1)B^T(\tau_1)\Phi^T(T, \tau_1)dx^*(T).
\] (12)

By rearranging the order of the integrations in the double integral, we can write
\[
\int_t^T \int_t^\tau \dot{u}(\tau_1)B^T(\tau_1)\Phi^T(\tau, \tau_1)Q(\tau)x^*(\tau)d\tau_1d\tau = \int_t^T \int_{\tau_1}^T \dot{u}(\tau_1)B^T(\tau_1)\Phi^T(\tau, \tau_1)Q(\tau)x^*(\tau)d\tau d\tau_1
= \int_t^T \dot{u}(\tau_1)B^T(\tau_1)\int_{\tau_1}^T \Phi^T(\tau, \tau_1)Q(\tau)x^*(\tau)d\tau d\tau_1.
\] (13)

Also, we can write
\[
\int_t^T \dot{u}(\tau)R(\tau)u^*(\tau)d\tau = \int_t^T \dot{u}(\tau_1)R(\tau_1)u^*(\tau_1)d\tau_1
\] (15)
by simply changing the dummy variable utilized in the integration. Using (14) and (15), we can write (12) as
\[
0 = \int_t^T \dot{u}(\tau_1)B^T(\tau_1)\int_{\tau_1}^T \Phi^T(\tau_1, \tau_1)Q(\tau_1)x^*(\tau_1)d\tau_1d\tau + \int_t^T \dot{u}(\tau_1)R(\tau_1)u^*(\tau_1)d\tau_1
+ \int_t^T \dot{u}(\tau_1)B^T(\tau_1)\Phi^T(T, \tau_1)dx^*(T).
\] (16)

Changing the dummy variables \( \tau \) and \( \tau_1 \) in the integration to \( \tau_1 \) and \( \tau \), respectively, we get
\[
0 = \int_t^T \dot{u}(\tau)B^T(\tau)\int_{\tau}^T \Phi^T(\tau_1, \tau_1)Q(\tau_1)x^*(\tau_1)d\tau_1d\tau + \int_t^T \dot{u}(\tau)R(\tau)u^*(\tau)d\tau
+ \int_t^T \dot{u}(\tau)B^T(\tau)\Phi^T(T, \tau)dx^*(T).
\] (17)

Hence,
\[
\int_t^T \dot{u}(\tau) \left\{ B^T(\tau)\int_{\tau}^T \Phi^T(\tau_1, \tau_1)Q(\tau_1)x^*(\tau_1)d\tau_1 + R(\tau)u^*(\tau) + B^T(\tau)\Phi^T(T, \tau)\Omega x^*(T) \right\} d\tau = 0.
\] (18)

Therefore,
\[
\int_t^T \dot{u}(\tau) \left\{ B^T(\tau)p(\tau) + R(\tau)u^*(\tau) \right\} d\tau = 0.
\] (19)
where \( p(\tau) \) is defined as
\[
p(\tau) = \int_{\tau}^T \Phi^T(\tau_1, \tau_1)Q(\tau_1)x^*(\tau_1)d\tau_1 + \Phi^T(T, \tau)\Omega x^*(T).
\] (20)

Since the equation (19) should hold for any signal \( \dot{u} \), we should have
\[
B^T(\tau)p(\tau) + R(\tau)u^*(\tau) = 0.
\] (21)

Hence,
\[
u^*(\tau) = -R^{-1}(\tau)B^T(\tau)p(\tau).
\] (22)
Differentiating (20) with respect to the time variable \( \tau \), we get

\[
\frac{d}{d\tau} p(\tau) = \int_{\tau}^{T} \frac{d}{d\tau} \Phi^T(\tau_1, \tau) Q(\tau_1) x^*(\tau_1) d\tau_1 - Q(\tau) x^*(\tau) + \frac{d}{d\tau} \Phi^T(T, \tau) \Omega x^*(T). \tag{23}
\]

Since \( \Phi \) is the transition matrix corresponding to the matrix \( A \), we know that \( \frac{d}{d\tau} \Phi(\tau, t_0) = A(\tau) \Phi(t_0, \tau) \) and \( \frac{d}{d\tau} \Phi(t_0, \tau) = -\Phi(t_0, \tau) A(\tau) \). Hence, from (23), we get

\[
\frac{d}{d\tau} p(\tau) = -\int_{\tau}^{T} A^T(\tau) \Phi^T(\tau_1, \tau) Q(\tau_1) x^*(\tau_1) d\tau_1 - Q(\tau) x^*(\tau) - A^T(\tau) \Phi^T(T, \tau) \Omega x^*(T) \tag{24}
\]

\[
= -A^T(\tau) p(\tau) - Q(\tau) x^*(\tau). \tag{25}
\]

Also, using (22), we get

\[
\frac{d}{d\tau} x^*(\tau) = A(\tau) x^*(\tau) - B(\tau) R^{-1}(\tau) B^T(\tau) p(\tau). \tag{26}
\]

Using (25) and (26), we get

\[
\frac{d}{d\tau} \begin{bmatrix} x^*(\tau) \\ p(\tau) \end{bmatrix} = \overline{A}(\tau) \begin{bmatrix} x^*(\tau) \\ p(\tau) \end{bmatrix}, \tag{27}
\]

where \( \overline{A}(\tau) \) denotes

\[
\overline{A}(\tau) = \begin{bmatrix} A(\tau) & -B(\tau) R^{-1}(\tau) B^T(\tau) \\ -Q(\tau) & -A^T(\tau) \end{bmatrix}. \tag{28}
\]

Denoting the transition matrix corresponding to \( \overline{A}(\tau) \) by \( \overline{\Phi} \), we get

\[
\begin{bmatrix} x^*(\tau) \\ p(\tau) \end{bmatrix} = \overline{\Phi}(\tau, T) \begin{bmatrix} x^*(\tau) \\ p(\tau) \end{bmatrix}. \tag{29}
\]

From the definition of the signal \( p(\tau) \) in (20), we see that \( p(T) = \Omega x^*(T) \). Hence, the equation (29) can be rewritten as

\[
\begin{bmatrix} x^*(\tau) \\ p(\tau) \end{bmatrix} = \overline{\Phi}(\tau, T) \begin{bmatrix} x^*(\tau) \\ \Omega x^*(T) \end{bmatrix} = \overline{\Phi}(\tau, T) \begin{bmatrix} I_n \\ \Omega \end{bmatrix} x^*(T) \tag{30}
\]

where \( I_n \) denotes the \( n \times n \) identity matrix. From (31), we see that \( p(\tau) = H_1(\tau) x^*(T) \) and \( x^*(\tau) = H_2(\tau) x^*(T) \) where

\[
H_1(\tau) = [0_n, I_n] \overline{\Phi}(\tau, T) \begin{bmatrix} I_n \\ \Omega \end{bmatrix} \tag{32}
\]

\[
H_2(\tau) = [I_n, 0_n] \overline{\Phi}(\tau, T) \begin{bmatrix} I_n \\ \Omega \end{bmatrix} \tag{33}
\]

where \( 0_n \) denotes the zero matrix of dimensions \( n \times n \). Hence, \( p(\tau) \) is of the form \( M(\tau)x^*(\tau) \) with \( M(\tau) = H_1(\tau) H_2^{-1}(\tau) \). To find the differential equation that \( M(\tau) \) must satisfy, we can differentiate the equation \( M(\tau) = H_1(\tau) H_2^{-1}(\tau) \) with respect to the time variable \( \tau \) and utilize the equation \( \frac{d}{d\tau} \overline{\Phi}(\tau, T) = \overline{A}(\tau) \overline{\Phi}(\tau, T) \) to find \( \frac{d}{d\tau} H_1(\tau) \) and \( \frac{d}{d\tau} H_2^{-1}(\tau) \). A somewhat less formal way to find the differential equation for \( M(\tau) \) is to simply substitute \( p(\tau) = M(\tau)x^*(\tau) \) into (25) to see what the condition on \( \frac{d}{d\tau} M(\tau) \) would be. When we substitute \( p(\tau) = M(\tau)x^*(\tau) \) into (25), we get

\[
\left( \frac{d}{d\tau} M(\tau) \right) x^*(\tau) + M(\tau) \left( \frac{d}{d\tau} x^*(\tau) \right) = -A^T(\tau) M(\tau) x^*(\tau) - Q(\tau) x^*(\tau). \tag{34}
\]

Hence,

\[
\left( \frac{d}{d\tau} M(\tau) \right) x^*(\tau) = -M(\tau) [A(\tau) x^*(\tau) - B(\tau) R^{-1}(\tau) B^T(\tau) M(\tau) x^*(\tau)] - A^T(\tau) M(\tau) x^*(\tau) - Q(\tau) x^*(\tau). \tag{35}
\]

Therefore,

\[
\left( \frac{d}{d\tau} M(\tau) \right) x^*(\tau) = [-M(\tau) A(\tau) + M(\tau) B(\tau) R^{-1}(\tau) B^T(\tau) M(\tau) - A^T(\tau) M(\tau) - Q(\tau)] x^*(\tau). \tag{36}
\]
Hence, comparing both sides of equation (36), we see that this equation is satisfied if

$$\left( \frac{d}{dt} M(\tau) \right) = -M(\tau)A(\tau) - A^T(\tau)M(\tau) + M(\tau)B(\tau)R^{-1}(\tau)B^T(\tau)M(\tau) - Q(\tau).$$ (37)

The equation (37) is called the differential Riccati equation. From equation (25), we see that \( p(T) = \Omega x^*(T) \) and therefore \( M(T) = \Omega \). Hence, the equation (37) has to be solved backward in time over the time interval \([t, T]\) starting with the terminal condition \( M(T) = \Omega \).

In the infinite-horizon time-invariant LQR problem, the differential equation (37) for \( M(\tau) \) is replaced by its steady-state solution given as the solution to the algebraic Riccati equation:

$$MA + A^T M - MBR^{-1}B^T M + Q = 0.$$ (38)

### 2 Derivation of LQR controller considering only linear controllers

In the derivation of the LQR controller above, we started with looking for a general control signal \( u(\tau) \) and saw that the optimal controller given the cost functional (1) is actually a linear state-feedback controller \( u = Kx \) with \( K = -R^{-1}B^TM \).

Now that we know that the optimal controller for the cost functional (1) is actually a linear controller, we can derive the optimal feedback gain more easily than in the derivation above. In other words, if we pose the optimal control problem as one of finding the best linear controller, then a simpler derivation of the LQR feedback gain can be found as shown below.

If \( u(\tau) = K(\tau)x(\tau) \), the cost functional (1) can be written as

$$V(t, T) = \int_t^T x^T(\tau)\{Q(\tau) + K^T(\tau)R(\tau)K(\tau)\}x(\tau) d\tau + x^T(T)\Omega x(T)$$ (39)

$$= \int_t^T x^T(\tau)L(\tau)x(\tau) d\tau + x^T(T)\Omega x(T)$$ (40)

where

$$L(\tau) = Q(\tau) + K^T(\tau)R(\tau)K(\tau).$$ (41)

Defining the transition matrix corresponding to the closed-loop system matrix \((A + BK)\) by \( \Phi_c \), we can write \( x(\tau) = \Phi_c(\tau, t)x(t) \). Substituting this expression for \( x(\tau) \) into equation (40), we get

$$V(t, T) = x^T(t) \int_t^T \Phi_c^T(\tau, t)L(\tau)\Phi_c(\tau, t) d\tau + x^T(t)\Phi_c^T(T, t)\Omega\Phi_c(T, t)x(t)$$ (42)

$$= x^T(t)M(t, T)x(t)$$ (43)

where

$$M(t, T) = \int_t^T \Phi_c^T(\tau, t)L(\tau)\Phi_c(\tau, t) d\tau + \Phi_c^T(T, t)\Omega\Phi_c(T, t).$$ (44)

Differentiating the equation (44) and using the equation \( \frac{d}{dt} \Phi_c(\tau, t) = -\Phi_c(\tau, t)A_c(t) \) where \( A_c(t) = A(t) + B(t)K(t) \), we get

$$\frac{d}{dt}M(t, T) = -\int_t^T A_c^T(\tau)\Phi_c^T(\tau, t)L(\tau)\Phi_c(\tau, t) d\tau - \int_t^T \Phi_c^T(\tau, t)L(\tau)\Phi_c(\tau, t)A_c(t) d\tau - L(t)$$

$$- A_c^T(\tau)\Phi_c(T, t)\Omega\Phi_c(T, t) - \Phi_c^T(T, t)\Omega\Phi_c(T, t)A_c(t).$$ (45)

Hence,

$$\frac{d}{dt}M(t, T) = -A_c^T(\tau)M(t, T) - M(t, T)A_c(t) - L(t).$$ (46)

Note that (44) is the solution of the differential equation (46) starting from the terminal condition \( M(T, T) = \Omega \).

For each choice of \( K \), there is a corresponding function \( M(t, T) \) as determined by (44) or (46) and a corresponding cost functional as written in (43).

Now, to find the optimal gain, denote the optimal gain by \( K^*(\tau) \), the corresponding function \( M(t, T) \) as \( M^*(t, T) \), and the corresponding cost functional by \( V^*(t, T) = x^T(t)M^*(t, T)x(t) \). Also, consider a perturbation of the gain as \( K^*(\tau) + \tilde{K}(\tau) \), the corresponding function \( M(t, T) \) as \( M^*(t, T) + \tilde{M}(t, T) \), and the corresponding cost functional by \( V^*(t, T) + \tilde{V}(t, T) = x^T(t)[M^*(t, T) + \tilde{M}(t, T)]x(t) \). If \( K^*(\tau) \) is indeed the optimal gain
Since we should have \( V^*(t,T) + \bar{V}(t,T) \geq V^*(t,T) \), i.e., \( \bar{V}(t,T) \geq 0 \). Hence,
\[
x^T(t)\hat{M}(t,T)x(t) \geq 0,
\]
for all \( x(t) \in \mathbb{R}^n \), which implies that \( \hat{M}(t,T) \) must be a positive semi-definite matrix. Since a differential equation such as (46) must be satisfied (where \( A_c \) denotes \( A + BK \) and \( L = Q + K^TRK \), for the corresponding \( K \) matrix) by both \( M^*(t,T) \) and \( M^*(t,T) + \hat{M}(t,T) \), we get
\[
\frac{d}{dt}M^*(t,T) = -[A(t) + B(t)K^*(t)]^T M^*(t,T) - M^*(t,T) [A(t) + B(t)K^*(t)] - Q(t)
- (K^*(t))^T R(t)K^*(t) 
\]
\[
\frac{d}{dt}[M^*(t,T) + \hat{M}(t,T)] = -[A(t) + B(t)K^*(t) + B(t)\hat{K}(t)]^T [M^*(t,T) + \hat{M}(t,T)]
- [M^*(t,T) + \hat{M}(t,T)][A(t) + B(t)K^*(t) + B(t)\hat{K}(t)] - Q(t)
- [K^*(t) + \hat{K}(t)]^T R(t)[K^*(t) + \hat{K}(t)] 
\]
From (48) and (49), we get
\[
\frac{d}{dt}\hat{M}(t,T) = -[A(t) + B(t)K^*(t) + B(t)\hat{K}(t)]^T \hat{M}(t,T) - [B(t)\hat{K}(t)]^T M^*(t,T)
- \hat{M}(t,T)[A(t) + B(t)K^*(t) + B(t)\hat{K}(t)] - M^*(t,T)[B(t)\hat{K}(t)]
- [\hat{K}(t)]^T R(t)\hat{K}(t) - [K^*(t)]^T R(t)K^*(t) - [\hat{K}(t)]^T R(t)K^*(t) 
\]
Hence,
\[
\frac{d}{dt}\hat{M}(t,T) = -[A(t) + B(t)K^*(t) + B(t)\hat{K}(t)]^T \hat{M}(t,T) - \hat{M}(t,T)[A(t) + B(t)K^*(t) + B(t)\hat{K}(t)] - \hat{L}(t) 
\]
where
\[
\hat{L}(t) = [\hat{K}(t)]^T R(t)\hat{K}(t) + [\hat{K}(t)]^T [R(t)K^*(t) + B^T(t)M^*(t,T)] + [(K^*(t))^T R(t) + M^*(t,T)B(t)]\hat{K}(t) 
\]
Since we should have \( M^*(T,T) = \Omega \) and \( M^*(T,T) + \hat{M}(T,T) = \Omega \), we should have \( \hat{M}(T,T) = 0 \). Since the matrix differential equation (46) has the solution (44) when the terminal condition is specified as \( M(T,T) = \Omega \), the matrix differential equation (51) for \( \hat{M} \) with terminal condition \( \hat{M}(T,T) = 0 \) has a solution of form
\[
\hat{M}(t,T) = \int_t^T \hat{\Phi}_c^T(\tau,t)\hat{L}(\tau)\hat{\Phi}_c(\tau,t)d\tau 
\]
where \( \hat{\Phi}_c \) denotes the transition matrix corresponding to \( A + BK^* + B\hat{K} \). As mentioned above, \( \hat{M}(t,T) \) must be a positive semi-definite matrix to satisfy equation (47). Hence, from (53), we see that \( \hat{L}(\cdot) \) must be a positive semi-definite matrix. From the definition of (52), we see that for \( \hat{L}(t) \) to be a positive semi-definite matrix for all possible \( \hat{K}(t) \), we should have \( R(t)K^*(t) + B^T(t)M^*(t,T) = 0 \). Therefore,
\[
K^*(t) = -R^{-1}(t)B^T(t)M^*(t,T). 
\]
Substituting (54) into (48) and noting that \( M^*(t,T) \) and \( R(t) \) are symmetric matrices, we get the differential Riccati equation
\[
\frac{d}{dt}M^*(t,T) = -M^*(t,T)A(t) - A^T(t)M^*(t,T) + M^*(t,T)B(t)R^{-1}(t)B^T(t)M^*(t,T) - Q(t). 
\]
The optimal controller is given by (54) and (55). The equation (55) has to be solved backward in time over the time interval \( [t,T] \) starting with the terminal condition \( M^*(T,T) = \Omega \). In the infinite-horizon time-invariant LQR problem, the differential equation (55) for \( M^*(t) \) is replaced by its steady-state solution given as the solution to the algebraic Riccati equation:
\[
M^*A + A^TM^* - M^*BR^{-1}B^TM^* + Q = 0. 
\]