Linear Optimal Control Notes

Jorge Aurelio Menéndez January 20, 2020

These are my notes of the material in the textbook *Linear Optimal Control Systems* by [Kwakernaak and Sivan, 1972]. All theorems and lemmas are numbered as in the textbook. Elementary linear algebra and linear systems theory is skipped, as well as chapter 2.

Contents

Chapter 1: (Stochastic) Linear Systems

1.1 Linear systems and their solutions

In all the of the following, we'll consider **linear control systems**, which are dynamical systems of the form

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1}$$

where

- $x(t) \in \mathbb{R}^n$ is the **state** variable
- $u(t) \in \mathbb{R}^m$ is the **input** variable
- $A(t) \in \mathbb{R}^{n\times n}$ is the **open-loop coupling matrix** (always assumed to be a continous function of t)
- $B(t) \in \mathbb{R}^{n\times m}$ is the **input matrix** (always assumed to be a piecewise continuous function of t)

We'll generally omit the time-dependence when writing such differential equations. Unless otherwise noted, the results below will hold for the case of continous time-varying matrices $A(t), B(t)$, even if we leave the time-dependence implicit in the notation. When A, B are constant in time, this is referred to as a **linear time-invariant (LTI) system**.

It is easy to see that the following theorem holds:

Theorem 1.1. If $A(t)$ is continuous for all t, then the homogenous differential equation

$$\dot{x}(t) = A(t)x(t)$$
$$x(t_0) = x_0$$

admits a unique solution of the form

$$x(t) = \Phi(t, t_0)x_0$$

where the $n \times n$ **transition matrix** $\Phi(t, t_0)$ is defined implicitly by the matrix differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, t_0) = A(t)\Phi(t, t_0)$$
$$\Phi(t_0, t_0) = I$$

In particular, if the coupling matrix is time-invariant, then this matrix differential equation can be analytically solved to yield the classic exponential solution to linear time-invariant systems:

$$\Phi(t, t_0) = e^{A(t-t_0)}$$

More generally, the transition matrix satisfies the following theorem:

Theorem 1.2. The transition matrix $\Phi(t, t_0)$ defined in theorem 1.1 satisfies the following properties:

(a) $\forall t_0, t_1, t_2 \quad \Phi(t_2, t_1)\Phi(t_1, t_0) = \Phi(t_2, t_0)$

(b) $\forall t_0, t \quad \Phi(t, t_0)^{-1} = \Phi(t_0, t)$

(c) $\forall t_0, t \quad \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t_0, t) = -\Phi(t_0, t)A(t)$

The first property is trivially proven by noting that, for any t_1, the matrix $G(t) = \Phi(t, t_1)\Phi(t_1, t_0)$ satisfies the matrix differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}G(t) = A(t)\Phi(t, t_1)\Phi(t_1, t_0) = A(t)G(t)$$

which, by definition of Φ, admits a single unique solution for all t: $G(t) = \Phi(t, t_0)$. It is easy to see that the second property follows from the first, since

$$\Phi(t, t_0)\Phi(t_0, t) = \Phi(t, t) = I \Leftrightarrow \Phi(t_0, t) = \Phi(t, t_0)^{-1}$$

Note that this property additionally implies that the transition matrix is always non-singular. The last property can be derived as follows:

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t}I &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\Phi(t, t_0)\Phi(t_0, t)\right] \\ &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, t_0)\right]\Phi(t_0, t) + \Phi(t, t_0)\left[\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t_0, t)\right] \\ \Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t_0, t) &= -\Phi(t, t_0)^{-1}\left[\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, t_0)\right]\Phi(t_0, t) \\ &= -\Phi(t_0, t)A(t)\Phi(t, t_0)\Phi(t_0, t) \\ &= -\Phi(t_0, t)A(t) \end{aligned}$$

From theorem 1.1, the general solution to equation 1 can be easily derived:

Theorem 1.3. If $A(t)$ is continuous and $B(t)$, $u(t)$ are piecewise continuous for all t, then the differential equation

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$
$$x(t_0) = x_0$$

admits a unique solution of the form

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t, \tau)B(\tau)u(\tau)\,\mathrm{d}\tau$$

1.2 Controllability

A linear control system is said to be **controllable** (or **reachable**) iff for any initial state $x_0 \in \mathbb{R}^n$ and desired final state $x_1 \in \mathbb{R}^n$, there exists an input $u : [t_0, t_1] \to \mathbb{R}^m$ that will steer the state variable from the initial state $x(t_0) = x_0$ to a final state $x(t_1) = x_1$ within a finite amount of time $t_1 - t_0$. If such a system is controllable, we say that the pair (A, B) is controllable. This definition is due to [Kalman, 1960a], and is a central concept to all of control theory.

In the case of a linear time-invariant system, the following condition for controllability can be easily derived:

Theorem 1.23. An n-dimensional linear time-invariant system with m-dimensional inputs is controllable iff the $n \times nm$ **controllability matrix**

$$P = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$

is full rank n.

To prove this, we use the Taylor expansion of the matrix exponential. Without loss of generality, we let $x_0 = 0$ so that

$$x(t_1) = \int_{t_0}^{t_1} e^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau$$

Using the Taylor expansion of the exponential, we have that

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 +$$

Plugging this into the previous equation entails that $x(t_1)$ is a linear combination of the columns of $B, AB, A^2B, A^3B, \dots$. Critically, if the columns of A^kB are not linearly independent of the columns of all the previous matrices in this sequence, i.e.

$$A^kB = A^{k-1}B\Gamma_{k-1} + A^{k-2}B\Gamma_{k-2} + \dots + B\Gamma_0$$

for some set of coefficient matrices $\Gamma_0, \Gamma_1, \dots, \Gamma_{k-1}$, then neither are those of $A^{k+1}B$, since

$$A^{k+1}B = A(A^kB) = A(A^{k-1}B\Gamma_{k-1} + \dots + B\Gamma_0) = A^kB\Gamma_{k-1} + \dots + AB\Gamma_0$$

Since the columns of these matrices live in $\mathbb{R}^n$, a maximum of n of them can be linearly independent, entailing that, for any $k \geq n$, the columns of A^kB are not linearly independent of those of $B, AB, \dots, A^{n-1}B$. This entails the following important property of the matrix exponential:

$$e^{At}B = \sum_{k=0}^{n-1} A^kB\Gamma_k$$

for some set of $m \times m$ real matrices $\Gamma_0, \dots, \Gamma_{n-1}$, given any $n \times m$ real matrix B and real number t. Plugging this into the above expression for $x(t_1)$ entails that the final state at any time t_1 must live in the column space of P, thus proving theorem 1.23.

If P is not full rank, then $x(t_1)$ is evidently restricted to a subspace of $\mathbb{R}^n$, which we call the **controllable subspace**. It is easy to see that the controllable subspace is invariant under A: if x is in it then Ax is too (lemma 1.3), so that any state in this subspace is reachable from any other state within it (theorem 1.25).

When a linear time-invariant system is not controllable, it is often useful to transform the system into a canonical form that makes the controllable subspace explicit, by invoking an appropriate change in coordinates. We do this by picking a basis $e_1, e_2, \ldots, e_n$ for $\mathbb{R}^n$ such that

▷ $e_1, e_2, \ldots, e_{n_c}$ span the n_c-dimensional controllable subspace

▷ $e_{n_c+1}, e_{n_c+2}, \ldots, e_n$ span the rest of $\mathbb{R}^n$, and are therefore orthogonal to the controllable subspace

Assembling these two sets of basis vectors into a coordinate transform matrix $T = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}$ and letting $\tilde{x}$ denote the state variable in this new coordinate system, we have:

$$x(t) = T\tilde{x}(t) \Leftrightarrow \tilde{x}(t) = T^{-1}x(t)$$
$$\Rightarrow \dot{\tilde{x}} = T^{-1}AT\tilde{x}(t) + T^{-1}Bu$$

We then partition the coordinate transform matrix $T = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ into an $n \times n_c$ block T_1 containing the basis vectors spanning the controllable subspace and an $n \times (n - n_c)$ block T_2 containing the rest. We can similiarly partition the inverse transform matrix $T^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ into an $n_c \times n$ block U_1 and an $(n - n_c) \times n$ block U_2. With this partitioning, we have

$$\dot{\tilde{x}} = \begin{bmatrix} U_1AT_1 & U_1AT_2 \\ U_2AT_1 & U_2AT_2 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} U_1B \\ U_2B \end{bmatrix} u$$

We now note that, by the definition of the matrix inverse, $U_2T_1 = 0$, which implies that U_2 is orthogonal to the controllable subspace. Because B is always in the controllable subspace, this means that $U_2B = 0$. Recalling that the controllable subspace is invariant under A, AT_1 is in the controllable subspace too, meaning that $U_2AT_1 = 0$ as well. The dynamics of the state variable in these coordinates thus simplify to

$$\dot{\tilde{x}} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix} u$$
$$\tilde{A}_{11} = U_1AT_1, \quad \tilde{A}_{22} = U_2AT_2, \quad \tilde{A}_{12} = U_1AT_2, \quad \tilde{B} = U_1B$$

In other words, in these coordinates, the system partitions into an $(n - n_c)$-dimensional autonomous subsystem and an n_c-dimensional controllable subsystem that are coupled only in one direction via $\tilde{A}_{12}$.

Importantly, the eigenvalues of this system – which are equal to the eigenvalues of the original system – are given by the eigenvalues of the two subsystems[1], i.e. the eigenvalues of the diagonal blocks $\tilde{A}_{11}, \tilde{A}_{22}$. For any linear time-invariant system, we thus identify the **controllable** and **uncontrollable modes** with those corresponding to the eigenvalues of $\tilde{A}_{11}$ and $\tilde{A}_{22}$, respectively. We say that such a system is **stabilizable** when all unstable modes are controllable, i.e. when all the uncontrollable modes are stable.

For the more general case of time-varying $A(t), B(t)$, it turns out one can derive the following conditions for controllability, similar to theorem 1.23 but with the controllability matrix effectively replaced with another matrix:

Theorem 1.29. An n-dimensional linear control system is controllable iff for all t_0 there exists a finite $t_1 > t_0$ such that the $n \times n$ **controllability Gramian**

$$W(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)B(\tau)^T\Phi(t_1, \tau)^T \, \mathrm{d}\tau$$

is full rank n, where $\Phi(t, t_0)$ denotes the transition matrix of the system.

The proof for this can be found in [Kalman et al., 1969].

1.3 Reconstructibility

Suppose that, at any given time, we only observe a d-dimensional linear transformation of the full state variable, given by the **output** variable

$$y(t) = C(t)x(t) \tag{2}$$

[1] This follows from the fact that the determinant of a block diagonal matrix is equal to the product of determinants of its diagonal blocks, such that the characteristic polynomial of a block diagonal matrix is just the product of the characteristic polynomials of its diagonal blocks. This can be derived from, for example, Leibniz's determinant formula.

In this case, one might ask whether the information provided by this variable is sufficient to recover the full state variable. This notion is captured by the properties of reconstructibility and observability. In the below, we'll use the notation $y(t; t_0, x_0, u)$ to denote the outputs resulting from the boundary condition $x(t_0) = x_0$ and input u, i.e. (using theorem 1.3),

$$y(t; t_0, x_0, u) = C(t)\Phi(t, t_0)x_0 + C(t)\int_{t_0}^{t_1} \Phi(t, \tau)B(\tau)u(\tau)\,\mathrm{d}\tau$$

A linear control system is said to be **reconstructible** if for any timepoint t_1 there exists a (finite) previous timepoint $t_0 < t_1$ such that the statement

$$\forall t \in [t_0, t_1] \quad y(t; t_0, x_0, u) = y(t; t_0, x_0', u)$$

implies $x_0 = x_0'$. In other words, there is always at least one timepoint in the past at which the state can be in some sense recovered from the history of outputs since that timepoint. This definition is due to [Kalman et al., 1969]. If such a system is reconstructible, we say that that the pair (A, C) is reconstructible. Reconstructibility is complimentary to **observability**, which can be defined similarly but in terms of the future outputs: for any t_0 there exists a finite *future* timepoint $t_1 > t_0$ such that the above conditional statement holds[2].

Intuitively, what is required of a reconstructible system is that its transition matrix $\Phi(t, t_0)$ ensure the initial state x_0 evolve along dimensions that ensure the d-dimensional outputs $\{y(t)\}_{t\in[t_0,t_1]}$ suffice to reconstruct it. This requires that in some sense the sequence of transformations $\{\Phi(t, t_0)\}_{t\in[t_0,t_1]}$ ensure that the mapping from initial state to observations $x_0 \to \{y(t)\}_{t\in[t_0,t_1]}$ is invertible. This intuition is captured by plugging in our above expression for $y(t; t_0, x_0, u)$ and re-arranging the reconstructibility condition to arrive at the following equivalent statement:

$$\forall t \in [t_0, t_1] \quad y(t; t_0, x_0, u) - y(t; t_0, x_0', u) = C(t)\Phi(t, t_0)(x_0 - x_0') = 0$$

implies $x_0 - x_0' = 0$. Thus, a system is reconstructible if and only if for any t_1, there exists a finite $t_0 < t_1$ such that

$$\forall t \in [t_0, t_1] \quad C(t)\Phi(t, t_0)x_0 = 0$$

implies $x_0 = 0$ (theorem 1.31). In other words, it is reconstructible iff the mapping $x_0 \to \{C(t)\Phi(t, t_0)x_0\}_{t\in[t_0,t_1]}$ is in some sense "full rank". This notion underlies the conditions for reconstructibility derived next.

For a linear time-invariant system, reconstructibility is ensured by the following condition:

Theorem 1.32. An n-dimensional linear time-invariant system with d-dimensional outputs $y(t) = Cx(t)$ is controllable iff the $nd \times n$ **reconstructibility matrix**

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

is full rank n.

We first prove that reconstructibility implies Q is full rank. As we did in section 1.2, we exploit the Taylor expansion of the matrix exponential to write

$$Ce^{At} = \sum_{k=0}^{n-1} \Gamma_k CA^k = \Gamma Q$$

for some set of real $d \times d$ coefficient matrices $\Gamma_0, \ldots, \Gamma_{n-1}$ assembled into a $d \times nd$ matrix $\Gamma = \begin{bmatrix}\Gamma_0 & \ldots & \Gamma_{n-1}\end{bmatrix}$. We then have that each observation $y(t) = Ce^{A(t-t_0)}x_0 = \Gamma(t)Qx_0$ is a linear transform of a linear combination of the columns of Q. Therefore, if Q is not full rank, then there exist non-zero x_0 for which $y(t) = 0$, entailing the system is not reconstructible; i.e. if the system is reconstructible, then Q must be full rank. To prove the other direction (that full rank Q implies reconstructibility), we first

[2] This is how reconstructibility and observability are defined in this particular textbook, which I believe is based on Kalman's theory of linear control systems.

assume that $y(t) = 0$ for all $t \in [t_0, t_1]$, given some t_0, t_1. Given the constant value of 0, we additionally have that all temporal derivatives of $y(t)$ are equal to 0 at time t_0. Plugging in $y(t) = Ce^{A(t-t_0)}x_0$ then gives us

$$\begin{aligned} y(t_0) &= Cx_0 = 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}y(t_0) &= CAx_0 = 0 \\ \frac{\mathrm{d}^2}{\mathrm{d}t^2}y(t_0) &= CA^2x_0 = 0 \\ &\vdots \\ \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}y(t_0) &= CA^{n-1}x_0 = 0 \end{aligned}$$

Noting the rows of Q in each of these equations, this set of n equations is equivalently expressed by writing $Qx_0 = 0$. If Q is full rank, then this equation implies $x_0 = 0$. This entails reconstructibility, thus proving the other direction of theorem 1.32.

If Q is not full rank, we can identify the **reconstructible subspace** with its row space. As we did in section 1.2, we can apply a change of coordinates to express an unreconstructible system in a canonical form that makes the reconstructible subspace and the **reconstructible** and **unreconstructible modes** explicit. We say that a linear time-invariant system is **detectable** when all unstable modes are reconstructible, i.e. when all unreconstructible modes are stable.

As we'll make more explicit in section 1.4, all of the notions and tools from section 1.2 used to analyze the controllability of a system can in fact be directly applied to analyzing the reconstructibility of a system. As such, one can derive a similar more general condition for reconstructibility:

Theorem 1.29. An n-dimensional linear control system with outputs $y(t) = C(t)x(t)$ is reconstructible iff for all t_1 there exists a finite $t_0 < t_1$ such that the $n \times n$ **observability Gramian**

$$M(t_0, t_1) = \int_{t_0}^{t_1} \Phi(\tau, t_0)^T C(\tau)^T C(\tau)\Phi(\tau, t_0)\,\mathrm{d}\tau$$

is full rank n, where $\Phi(t, t_0)$ denotes the transition matrix of the system.

The proof for this can be found in [Kalman et al., 1969].

1.4 Duality

It is easy to find parallels and symmetries between the notions and proofs of controllability and reconstructibility in sections 1.2 and 1.3, respectively. In fact, these can be made very explicit by invoking the idea of duality, due to [Kalman, 1960a, Kalman, 1960b].

Given a linear control system of the form

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \end{aligned}$$

we say that its **dual with respect to time** t^* is the linear control system

$$\begin{aligned} \dot{x}^*(t) &= A(t^* - t)^T x^*(t) + C(t^* - t)^T u^*(t) \\ y^*(t) &= B(t^* - t)^T x^*(t) \end{aligned}$$

We then have the following theorem:

Theorem 1.41.

(a) A linear control system is controllable iff its dual is reconstructible.

(b) A linear control system is reconstructible iff its dual is controllable.

This is easily proved for linear time-invariant systems: the controllability matrix of the dual system is the transpose of the reconstructibility matrix of the original system, and vice versa. Duality will prove to be very useful for applying statements and proofs about optimal controllers derived in section 2 to statements and proofs about optimal observers in sections 3 and 4.

1.5 Stochastic processes

In the below, we'll often consider linear control systems in which some source of noise is injected into them, making them a **stochastic process**. For example,

$$\dot{x}(t) = A(t)x(t) + B(t)v(t)$$

where $v(t)$ is a stochastic process with some **mean function** $m(t)$ and **cross-covariance function** $R(t_1, t_2)$, defined as

$$m(t) = \mathbb{E}\left[v(t)\right]$$
$$R(t_1, t_2) = \mathbb{E}\left[\left[v(t_1) - m(t_1)\right]\left[v(t_2) - m(t_2)\right]^T\right]$$

Here, $\mathbb{E}\left[\cdot\right]$ denotes an expectation over the probability distribution implied by the stochastic process. We'll call the matrix $Q(t) = R(t, t)$ the **covariance matrix** at time t.

In particular, we'll only consider **white noise** processes. Without getting into the details of the theory of stochastic differential equations, we will define a white noise process to be a stochastic process that satisfies the following two properties

(i) $m(t) = 0$

(ii) $R(t_1, t_2) = V(t_1)\delta(t_1 - t_2)$

where $\delta(\cdot)$ is the Dirac δ-function. In other words, samples from a white noise process are 0-mean and uncorrelated over time. The covariance matrix $V(t)$ is called the **intensity** of the white noise process.

An important fact about white noise processes is that they can be constructed as the temporal derivative of a stochastic process with uncorrelated increments (e.g. Brownian motion). We first construct such a stochastic process $v(t)$ by assuming the following three properties:

(i) The initial value is fixed to 0: $v(t_0) = 0$

(ii) The increment between any two successive timepoints has 0 mean: $\mathbb{E}\left[v(t_2) - v(t_1)\right] = 0$ for any t_1, t_2 such that $t_0 \leq t_1 \leq t_2$

(iii) Increments in two separate intervals are uncorrelated: $\mathbb{E}\left[\left[v(t_2) - v(t_1)\right]\left[v(t_4) - v(t_3)\right]^T\right] = 0$ for any t_1, t_2, t_3, t_4 such that $t_0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$

From this, one can easily derive that the mean and cross-covariance functions are given by

$$\begin{aligned}
m(t) &= \mathbb{E}\left[v(t)\right] = \mathbb{E}\left[v(t) - v(t_0)\right] = 0 \quad \text{for any } t \geq t_0 \\
R(t_1, t_2) &= \mathbb{E}\left[v(t_1)v(t_2)^T\right] \\
&= \mathbb{E}\left[\left[v(t_1) - v(t_0)\right]\left[v(t_2) - v(t_1) + v(t_1) - v(t_0)\right]^T\right] \\
&= \mathbb{E}\left[\left[v(t_1) - v(t_0)\right]\left[v(t_1) - v(t_0)\right]^T\right] + \mathbb{E}\left[\left[v(t_1) - v(t_0)\right]\left[v(t_2) - v(t_1)\right]^T\right] \\
&= \mathbb{E}\left[\left[v(t_1) - v(t_0)\right]\left[v(t_1) - v(t_0)\right]^T\right] \quad \text{assuming } t_2 \geq t_1 \geq t_0 \\
&= Q(t_1)
\end{aligned}$$

Similarly, if $t_1 \geq t_2 \geq t_0$, we have $R(t_1, t_2) = Q(t_2)$. Taking the derivative of this process $\dot{v} = \frac{\mathrm{d}v}{\mathrm{d}t}$, its mean and cross-covariance are given by

$$\mathbb{E}\left[\dot{v}(t)\right] = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[v\right](t) = 0$$
$$\mathbb{E}\left[\dot{v}(t_1)\dot{v}(t_2)^T\right] = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\mathbb{E}\left[v(t_1)v(t_2)^T\right] = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\begin{cases} Q(t_1) & \text{if } t_2 \geq t_1 \\ Q(t_2) & \text{else} \end{cases} = \dot{Q}(t_1)\delta(t_1 - t_2)$$

which are those of a white noise process with intensity function $V(t) = \dot{Q}(t)$ (assuming $Q(t)$ is differentiable). It is this formulation of white noise that is more rigorously worked out in the theory of stochastic differential equations (cf. [Oksendal, 2013]).

When the probability distribution of the increments is further specified to be a 0-mean Gaussian with diagonal covariance $(t_2 - t_1)I$, then the process $v(t)$ is called a **Wiener process** or a **Brownian motion** process, and its derivative a **Gaussian white noise** process. The Weiner processes is said to be a **Gaussian stochastic process** because the joint distribution $P(v(t_1), v(t_2), \ldots, v(t_n))$ is a multivariate Gaussian for any collection of timepoints $t_1, t_2, \ldots, t_n$.

Given the properties of white noise processes, the following theorem immediately follows:

Theorem 1.52. Consider a dynamical system satisfying

$$\dot{x}(t) = A(t)x(t) + B(t)w(t)$$
$$x(t_0) = x_0$$

where

- ▷ $w(t)$ is a white noise process with intensity $V(t)$
- ▷ the initial condition x_0 is random and independent of $w(t)$ with mean and covariance

$$m_0 = \mathbb{E}[x_0]$$
$$Q_0 = \mathbb{E}\left[(x_0 - m_0)(x_0 - m_0)^T\right]$$

The stochastic process $x(t)$ then has

(a) mean function

$$m(t) = \Phi(t, t_0)m_0$$

(b) cross-covariance function

$$R(t_1, t_2) = \Phi(t_1, t_0)Q_0\Phi(t_2, t_0)^T + \int_{t_0}^{\min(t_1,t_2)} \Phi(t_1, \tau)B(\tau)V(\tau)B(\tau)^T\Phi(t_2, \tau)^T \mathrm{d}\tau$$

(c) covariance function $Q(t)$ that satisfies the matrix differential equation

$$\dot{Q}(t) = A(t)Q(t) + Q(t)A(t)^T + B(t)V(t)B(t)^T$$

with boundary condition $Q(t_0) = Q_0$

From theorem 1.3, we have that

$$m(t) = \mathbb{E}\left[\Phi(t, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t, \tau)B(\tau)w(\tau)\,\mathrm{d}\tau\right] = \Phi(t, t_0)\mathbb{E}[x_0] + \int_{t_0}^{t_1} \Phi(t, \tau)B(\tau)\mathbb{E}[w(\tau)]\,\mathrm{d}\tau = \Phi(t, t_0)m_0$$

so that

$$x(t) - m(t) = \Phi(t, t_0)(x_0 - m_0) + \int_{t_0}^{t_1} \Phi(t, \tau)B(\tau)w(\tau)\,\mathrm{d}\tau$$

We then have that

$$\begin{aligned}
R(t_1,t_2) = \mathbb{E}\Bigg[& \left[\Phi(t_1,t_0)\left(x_0-m_0\right)\right]\left[\Phi(t_2,t_0)\left(x_0-m_0\right)\right]^T \\
& + \left[\Phi(t_1,t_0)\left(x_0-m_0\right)\right]\left[\int_{t_0}^{t_2}\Phi(t_2,\tau)B(\tau)w(\tau)\mathrm{d}\tau\right]^T \\
& + \left[\int_{t_0}^{t_1}\Phi(t_1,\tau)B(\tau)w(\tau)\,\mathrm{d}\tau\right]\left[\Phi(t_2,t_0)\left(x_0-m_0\right)\right]^T \\
& \left[\int_{t_0}^{t_1}\Phi(t_1,\tau)B(\tau)w(\tau)\,\mathrm{d}\tau\right]\left[\int_{t_0}^{t_2}\Phi(t_2,\tau)B(\tau)w(\tau)\mathrm{d}\tau\right]^T\Bigg] \\
= \Phi(t_1,t_0)\mathbb{E}\left[\left(x_0-m_0\right)\left(x_0-m_0\right)^T\right]\Phi(t_2,t_0)^T & \\
& + \Phi(t_1,t_0)\mathbb{E}\left[x_0-m_0\right]\left[\int_{t_0}^{t_2}\Phi(t_2,\tau)B(\tau)\mathbb{E}\left[w(\tau)\right]\mathrm{d}\tau\right]^T \\
& + \left[\int_{t_0}^{t_1}\Phi(t_1,\tau)B(\tau)\mathbb{E}\left[w(\tau)\right]\,\mathrm{d}\tau\right]\left[\Phi(t_2,t_0)\mathbb{E}\left[x_0-m_0\right]\right]^T \\
& + \int_{t_0}^{t_1}\int_{t_0}^{t_2}\Phi(t_1,\tau)B(\tau)\mathbb{E}\left[w(\tau)w(\tau')^T\right]B(\tau)^T\Phi(t_2,\tau)^T\mathrm{d}\tau\mathrm{d}\tau' \\
= \Phi(t_1,t_0)Q_0\Phi(t_2,t_0)^T + 0 + 0 + & \int_{t_0}^{t_1}\int_{t_0}^{t_2}\delta(\tau-\tau')\Phi(t_1,\tau)B(\tau)V(\tau)B(\tau)^T\Phi(t_2,\tau)^T\mathrm{d}\tau'\mathrm{d}\tau \\
= \Phi(t_1,t_0)Q_0\Phi(t_2,t_0)^T + & \int_{t_0}^{\min(t_1,t_2)}\Phi(t_1,\tau)B(\tau)V(\tau)B(\tau)^T\Phi(t_2,\tau)^T\mathrm{d}\tau
\end{aligned}$$

Using the definition of the transition matrix from theorem 1.1 together with Leibniz's integral rule[3], the differential equation satisfied by the covariance function $Q(t) = R(t,t)$ is easily derived.

In the case of a linear time-invariant system driven by white noise with constant intensity $V(t) = V$, the **steady-state covariance** $\bar{Q}$ is given by

$$\bar{Q} \equiv \lim_{t\to\infty} Q(t) = \lim_{t\to\infty}\left[e^{A(t-t_0)}Q_0e^{A^T(t-t_0)} + \int_{t_0}^{t}e^{A(t-\tau)}BVBe^{A^T(t-\tau)}\mathrm{d}\tau\right] = \int_0^\infty e^{A\tau}BVBe^{A^T\tau}\mathrm{d}\tau$$

where the last equality holds only if the coupling matrix A is stable (i.e. all eigenvalues have real part < 0) so that the matrix exponential decays to 0 in the limit. Another way to derive this is by assuming that the covariance matrix $Q(t)$ will hit an equilibrium state $\bar{Q}$ at which point

$$\dot{Q}(t) = 0$$

Plugging in the differential equation satisfied by $Q(t)$ into the left-hand side of this equation results in the following **Lyapunov equation** that $\bar{Q}$ must satisfy:

$$A\bar{Q} + \bar{Q}A^T + BVB^T = 0$$

This equation has a unique positive-definite solution if A is stable, which is given by the above integral. One can see this by taking the derivative of the integrand $M(t) = e^{At}BVB^Te^{A^Tt}$:

$$\dot{M}(t) = AM(t) + M(t)A^T$$

and integrating the left- and right- hand sides

$$\begin{aligned}
\text{LHS: } & \int_0^\infty \dot{M}(t)\,\mathrm{d}t = M(\infty) - M(0) = -BVB^T \\
\text{RHS: } & \int_0^\infty AM(t) + M(t)A^T\,\mathrm{d}t = A\left[\int_0^\infty M(t)\mathrm{d}t\right] + \left[\int_0^\infty M(t)\mathrm{d}t\right]A^T = A\bar{Q} + \bar{Q}A^T
\end{aligned}$$

to recover the above Lyapunov equation (assuming A is stable).

In the below, we'll often be interested in computing expectations of quadratic expressions of the state variable, such as a mean squared error. For this, it will be extremely useful to have the following theorem in hand:

[3] Wikipedia

Theorem 1.54. Consider a dynamical system satisfying

$$\dot{x}(t) = A(t)x(t) + B(t)w(t)$$
$$x(t_0) = x_0$$

where

- ▷ $w(t)$ is a white noise process with intensity $V(t)$
- ▷ the initial condition x_0 is random and independent of $w(t)$ with first and second moments

$$m_0 = \mathbb{E}[x_0]$$
$$Q_0 = \mathbb{E}\left[x_0 x_0^T\right]$$

Let $P_1, R(t)$ be symmetric positive semi-definite matrices and consider the (random) quantity

$$\mathcal{L} \equiv \int_{t_0}^{t_1} x(t)^T R(t) x(t)\,\mathrm{d}t + x(t_1)^T P_1 x(t_1)$$

We then have that the stochastic process $x(t)$ satisfies the following equality:

$$\mathbb{E}[\mathcal{L}] = \mathrm{Tr}\left[P(t_0)Q_0 + \int_{t_0}^{t_1} B(t)V(t)B(t)^T P(t)\,\mathrm{d}t\right]$$

where $P(t)$ is a symmetric positive semi-definite matrix that satsifies the matrix differential equation

$$-\dot{P}(t) = A(t)^T P(t) + P(t)A(t) + R(t)$$

with boundary condition $P(t_1) = P_1$, which has the following solution:

$$P(t) = \int_{t}^{t_1} \Phi(\tau, t)^T R(\tau)\Phi(\tau, t)\mathrm{d}\tau + \Phi(t_1, t)^T P_1 \Phi(t_1, t)$$

To prove this, we begin by plugging in the solution to the differential equation for $x(t)$ into $\mathcal{L}$, re-arranging terms, and eliminating all mean 0 terms as in the above proof of theorem 1.52 to obtain

$$\mathbb{E}[\mathcal{L}] = \mathrm{Tr}\left[MQ_0 + N\right]$$
$$M = \int_{t_0}^{t_1} \Phi(t, t_0)^T R(t)\Phi(t, t_0)\,\mathrm{d}t + \Phi(t_1, t_0)^T P_1 \Phi(t_1, t_0)$$
$$N = \int_{t_0}^{t_1} \left[R(t)\int_{t_0}^{t} \Phi(t, \tau)B(\tau)V(\tau)B(\tau)^T\Phi(t, \tau)^T\mathrm{d}\tau\right]\mathrm{d}t$$
$$+P_1\int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)V(\tau)B(\tau)^T\Phi(t_1, \tau)^T\mathrm{d}\tau$$

To simplify N futher, we'll use a change of order of integration to combine the two terms. We first exploit the cyclic property of the trace operator to equivalently define

$$N = \int_{t_0}^{t_1}\int_{t_0}^{t} B(\tau)V(\tau)B(\tau)^T\Phi(t, \tau)^T R(t)\Phi(t, \tau)\mathrm{d}\tau\mathrm{d}t + \int_{t_0}^{t_1} B(\tau)V(\tau)B(\tau)^T\Phi(t_1, \tau)^T P_1\Phi(t_1, \tau)\mathrm{d}\tau$$

We then note that the first term is an integral over the set

$$\{(t, \tau) : t \in [t_0, t_1], \tau \in [t_0, t]\}$$

which can be equivalently be expressed as

$$\{(t, \tau) : t \in [\tau, t_1], \tau \in [t_0, t_1]\}$$

Thus, we can exchange the order of integration to re-write the above double integral as follows:

$$\begin{aligned} N &= \int_{t_0}^{t_1}\int_{\tau}^{t_1} B(\tau)V(\tau)B(\tau)^T\Phi(t,\tau)^T R(t)\Phi(t,\tau)\mathrm{d}t\mathrm{d}\tau + \int_{t_0}^{t_1} B(\tau)V(\tau)B(\tau)^T\Phi(t_1,\tau)^T P_1\Phi(t_1,\tau)\mathrm{d}\tau \\ &= \int_{t_0}^{t_1} B(\tau)V(\tau)B(\tau)^T \Big[\underbrace{\int_{\tau}^{t_1}\Phi(t,\tau)^T R(t)\Phi(t,\tau)\mathrm{d}t + \Phi(t_1,\tau)^T P_1\Phi(t_1,\tau)}_{P(\tau)}\Big]\mathrm{d}\tau \end{aligned}$$

The differential equation for $P(t)$ then follows from the definition of the transition matrix from theorem 1.1 together with theorem 1.2 and Leibniz's integral rule (cf. footnote 3).

Note that if the system is deterministic, i.e. $V(t) = 0$ and $x(t_0) = x_0$ is fixed, then $\mathcal{L}$ reduces to

$$\mathcal{L} = x_0^T P(t_0) x_0$$

Moreover, if the system is time-invariant, A is stable, and $V(t) = V$, $R(t) = R$ are constant, the matrix $P(t)$ settles at a steady state $\bar{P}$ (at finite t) in the limit of $t_1 \to \infty$:

$$\begin{aligned} \bar{P} \equiv \lim_{t_1\to\infty} P(t) &= \lim_{t_1\to\infty}\int_t^{t_1} e^{A^T(\tau-t)}Re^{A(\tau-t)}\mathrm{d}\tau + e^{A^T(t_1-t)}P_1e^{A(t_1-t)} \\ &= \lim_{t_1\to\infty}\int_0^{t_1-t} e^{A^T\tau}Re^{A\tau}\mathrm{d}\tau \\ &= \int_0^{\infty} e^{A^T\tau}Re^{A\tau}\mathrm{d}\tau \end{aligned}$$

As above, $\bar{P}$ also satisfies a Lyapunov equation:

$$A^T\bar{P} + \bar{P}A + R = 0$$

This suggests the following approximation for the time-invariant case whenever t_1 is large:

$$\mathbb{E}[\mathcal{L}] \approx \mathrm{Tr}\left[\bar{P}Q_0 + (t_1 - t_0)BVB^T\bar{P}\right]$$

In other words, in this regime $\mathbb{E}[\mathcal{L}]$ increases approximately linearly with t_1 at a rate given by $\mathrm{Tr}\left[BVB^T\bar{P}\right]$.

Chapter 3: Optimal State Feedback Control

2.1 Deterministic case: the linear quadratic regulator (LQR)

In this section, we focus on solving the following problem:

> Find the input $u(t)$ that minimizes the cost functional
>
> $$\mathcal{L}_{[t_0,t_1]}[u] = \int_{t_0}^{t_1} x(t)^T R_x(t)x(t) + u(t)^T R_u(t)u(t)\,\mathrm{d}t + x(t_1)^T P_1 x(t_1)$$
>
> where $x(t)$ implicitly depends on $u(t)$ through equation 1.

Evidently, the goal of solving such a problem is to bring the state variable $x(t)$ to 0 as quickly and efficiently as possible given its dynamics. In the next section we'll consider the case where there is some possibly non-zero target state x^* to reach.

The $R_x(t)$ and $R_u(t)$ matrices weight the different components of the state variable and input variable, respectively. This is critical, for example, when the different components of these vectors are quantities with different units and of possibly different magnitudes. Another typical use of the state weighting matrix $R_x(t)$ is to express the a cost on some linear output variable $y(t) = C(t)x(t)$. In this case, a quadratic cost on the output can be expressed as

$$y(t)^T R_y(t)y(t) = x(t)^T C(t)^T R_y(t)C(t)x(t) = x(t)^T R_x(t)x(t)$$

where $R_x(t) = C(t)^T R_y(t) C(t)$ reflects the state dimensions that are relevant to the output variable, and may be low rank (and thus positive semi-definite) if the outputs $y(t)$ are lower-dimensional than the state $x(t)$.

To solve this optimization problem, we'll use the calculus of variations. We begin by expressing the input in terms of the true optimal input $u^*(t)$ and its variation $\delta u(t)$ around it:

$$u(t) = u^*(t) + \epsilon \cdot \delta u(t)$$

for some scalar ϵ. The state variable resulting from the optimal input, which we denote by $x^*(t)$, satisfies

$$\dot{x^*} = Ax^* + Bu^*$$
$$x^*(t_0) = x_0$$

where we've left the time-dependence of the matrices implicit to lighten the notation. It is easy to see that, because of the linearity of linear control systems, this leads to a corresponding partitioning of the state variable as

$$x(t) = x^*(t) + \epsilon \cdot \delta x(t)$$

where

$$\dot{\delta x} = A\delta x + B\delta u$$
$$\delta x(t_0) = 0$$

Because at $x^*(t), u^*(t)$ the cost functional is at a minimum (by definition), we know that the derivative of the cost functional with respect to ϵ must be 0 at $\epsilon = 0$. We can use this fact to derive an equation that $x^*(t), u^*(t), \delta x(t), \delta u(t)$ must satisfy. By expressing the cost functional $\mathcal{L}_{[t_0,t_1]}[u]$ in terms of $x^*, u^*, \delta x, \delta u, \epsilon$, it is straight-forward to derive that

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \epsilon} = 2\Bigg[&\int_{t_0}^{t_1} \delta x(t)^T R_x(t) x^*(t) + \delta u(t)^T R_u(t) u^*(t)\,\mathrm{d}t + \delta x(t_1)^T P_1 x^*(t_1) \\ &+ \epsilon \int_{t_0}^{t_1} \delta x(t)^T R_x(t) \delta x(t) + \delta u(t)^T R_u(t) \delta u(t)\,\mathrm{d}t + \delta x(t_1)^T P_1 \delta x(t_1)\Bigg]\end{aligned}$$

thus giving us

$$\frac{\partial \mathcal{L}}{\partial \epsilon}\Big|_{\epsilon=0} = 2\int_{t_0}^{t_1} \delta x(t)^T R_x(t) x^*(t) + \delta u(t)^T R_u(t) u^*(t)\,\mathrm{d}t + \delta x(t_1)^T P_1 x^*(t_1)$$

Plugging in the solution to the above differential equation for $\delta x(t)$, we then have

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \epsilon}\Big|_{\epsilon=0} &= \int_{t_0}^{t_1} \left[\int_{t_0}^{t} \Phi(t,\tau) B(\tau) \delta u(\tau) \mathrm{d}\tau\right]^T R_x(t) x^*(t) + \delta u(t)^T R_u(t) u^*(t)\,\mathrm{d}t \\ &\quad + \left[\int_{t_0}^{t_1} \Phi(t_1,\tau) B(\tau) \delta u(\tau) \mathrm{d}\tau\right]^T P_1 x^*(t_1) \\ &= \int_{t_0}^{t_1} \int_{t_0}^{t} \delta u(\tau)^T B(\tau)^T \Phi(t,\tau)^T R_x(t) x^*(t) \mathrm{d}\tau \mathrm{d}t + \int_{t_0}^{t_1} \delta u(t)^T R_u(t) u^*(t) \mathrm{d}t \\ &\quad + \int_{t_0}^{t_1} \delta u(\tau)^T B(\tau)^T \Phi(t_1,\tau)^T P_1 x^*(t_1) \mathrm{d}\tau \\ &= \int_{t_0}^{t_1} \int_{\tau}^{t_1} \delta u(\tau)^T B(\tau)^T \Phi(t,\tau)^T R_x(t) x^*(t) \mathrm{d}t \mathrm{d}\tau + \int_{t_0}^{t_1} \delta u(\tau)^T R_u(\tau) u^*(\tau) \mathrm{d}\tau \\ &\quad + \int_{t_0}^{t_1} \delta u(\tau)^T B(\tau)^T \Phi(t_1,\tau)^T P_1 x^*(t_1) \mathrm{d}\tau \\ &= \int_{t_0}^{t_1} \delta u(\tau)^T \left[B(\tau)^T \left(\int_{\tau}^{t_1} \Phi(t,\tau)^T R_x(t) x^*(t) \mathrm{d}t \Phi(t_1,\tau)^T P_1 x^*(t_1)\right) + R_u(\tau) u^*(\tau)\right] \mathrm{d}\tau\end{aligned}$$

where in the third line we exchanged the order of integration (exactly as we did above in the proof of theorem 1.54). Letting

$$p(\tau) = \int_{\tau}^{t_1} \Phi(t,\tau)^T R_x(t) x^*(t) \mathrm{d}t + \Phi(t_1,\tau)^T P_1 x^*(t_1)$$

we then have that

$$\frac{\partial \mathcal{L}}{\partial \epsilon}\Big|_{\epsilon=0} = 0 \Leftrightarrow \int_{t_0}^{t_1} \delta u(\tau)^T \Big(B(\tau)^T p(\tau) + R_u(\tau) u^*(\tau) \Big) \mathrm{d}\tau = 0$$

For this to hold for any variation $\delta u(\tau)$, it must be the case that for any $t \in [t_0, t_1]$,

$$\begin{aligned} & B(t)^T p(t) + R_u(t) u^*(t) = 0 \\ \Leftrightarrow\ & u^*(t) = -R_u(t)^{-1} B(t)^T p(t) \end{aligned}$$

thus giving us a solution for the optimal input $u^*(t)$ in terms of $p(t)$.

Our next step is thus to analyze the so-called **adjoint variable** $p(t)$. Using theorem 1.2 together with Leibniz's integral rule, we first derive the following differential equation for $p(t)$:

$$\dot{p} = -A^T p - R_x x^*$$

We additionally note that at time t_1 the adjoint variable satisfies the boundary condition $p(t_1) = P_1 x^*(t_1)$. Plugging in our solution for $u^*(t)$ into the dynamics of $x^*(t)$ then yields the following $2n$-dimensional autonomous system, termed the **variational equations**

$$\begin{bmatrix} \dot{x^*} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} A & -BR_u^{-1}B^T \\ -R_x & -A^T \end{bmatrix} \begin{bmatrix} x^* \\ p \end{bmatrix}$$

with boundary conditions

$$\begin{aligned} x^*(t_0) &= x_0 \\ p(t_1) &= P_1 x^*(t_1) \end{aligned}$$

Solving for $p(t)$ thus equates to solving this **two-point boundary problem**. Note that the coupling matrix of this system is a Hamiltonian matrix[4].

We can make some analytical progress on this front by considering the transition matrix of this $2n$-dimensional system $\Theta(t, t_0)$. By partitioning it into a set of four $n \times n$ matrices,

$$\Theta(t, t_0) = \begin{bmatrix} \Theta_{11}(t, t_0) & \Theta_{12}(t, t_0) \\ \Theta_{21}(t, t_0) & \Theta_{22}(t, t_0) \end{bmatrix}$$

we note that

$$\begin{aligned} x^*(t) &= \Theta_{11}(t, t_1) x^*(t_1) + \Theta_{12}(t, t_1) p(t_1) \\ &= \Theta_{11}(t, t_1) x^*(t_1) + \Theta_{12}(t, t_1) P_1 x^*(t_1) \\ &= \left[\Theta_{11}(t, t_1) + \Theta_{12}(t, t_1) P_1 \right] x^*(t_1) \Leftrightarrow x^*(t_1) = \left[\Theta_{11}(t, t_1) + \Theta_{12}(t, t_1) P_1 \right]^{-1} x^*(t) \\ p(t) &= \Theta_{21}(t, t_1) x^*(t_1) + \Theta_{22}(t, t_1) p(t_1) \\ &= \left[\Theta_{21}(t, t_1) + \Theta_{22}(t, t_1) P_1 \right] x^*(t_1) \\ &= \left[\Theta_{21}(t, t_1) + \Theta_{22}(t, t_1) P_1 \right] \left[\Theta_{11}(t, t_1) + \Theta_{12}(t, t_1) P_1 \right]^{-1} x^*(t) \end{aligned}$$

implying $p(t)$ has the following form:

$$p(t) = P(t) x^*(t)$$

This form for the adjoint variable in turn entails that the optimal input $u^*(t)$ has the form of a **linear control law**:

$$u^*(t) = K(t) x^*(t)$$

where $K(t) = -R_u(t)^{-1} B(t)^T P(t)$. This optimal input comprises a form of **state feedback**, since the input at time t depends only on the state of the system at time t. In particular, this *optimal* control law is called the **linear quadratic regulator**, since it optimizes a quadratic cost functional. From the aforementioned adjoint variable dynamics, we can additionally derive a differential equation for the matrix $P(t)$:

$$\begin{aligned} \dot{p} &= -A^T p - R_x x^* = (-A^T P - R_x) x^* \\ &= \dot{P} x^* + P \dot{x^*} = (\dot{P} + PA) x^* + PBu^* = (\dot{P} + PA - PBR_u^{-1} B^T P) x^* \\ \Leftrightarrow 0 &= \left(\dot{P} + A^T P + PA - PBR_u^{-1} B^T P + R_x \right) x^* \end{aligned}$$

[4] Wikipedia

For this to hold for all x^* (as it must), it must then be the case that

$$-\dot{P} = A^T P + PA - PBR_u^{-1}B^T P + R_x$$

which is the so-called **matrix Riccati equation**. [Kalman, 1960a] Note that $P(t)$ is symmetric at its boundary condition $P(t_1) = P_1$ and remains so onwards, as its dynamics are symmetric as well. Many numerical methods exist for numerically solving this matrix differential equation, cf. chapter 3.5 of [Kwakernaak and Sivan, 1972].

From theorem 1.54, we can in fact obtain a very direct interpretation of the matrix $P(t)$. First, we note that in this deterministic case $V(t) = 0$. Second, we note that under the optimal input $u^* = -R_u^{-1}B^T Px$ the system can be described using the **closed-loop dynamics**

$$\dot{x} = Ax + Bu^* = (A - BR_u^{-1}B^T P)x$$

Furthermore, the cost of the optimal input can be expressed as

$$\mathcal{L}_{[t_0,t_1]}[u^*] = \int_{t_0}^{t_1} x(t)^T \left(R_x(t) + P(t)B(t)R_u^{-1}(t)B(t)^T P(t)\right) x(t)\,\mathrm{d}t + x(t_1)^T P_1 x(t_1)$$

Plugging this into theorem 1.54 then gives us that the minimum cost from any time t can be written

$$\mathcal{L}_{[t,t_1]}[u^*] = x(t)^T P(t) x(t)$$

Thus, we can think of $P(t)$ as being the **cost-to-go** matrix, since multiplying it with the current state $x(t)$ gives the minimum cost achievable from this state onwards until the terminal time t_1.

We summarize our results with the following theorem:

Theorem 3.4. The optimal input $u^*(t)$ that minimizes the cost functional

$$\mathcal{L}_{[t_0,t_1]}[u] = \int_{t_0}^{t_1} x(t)^T R_x(t) x(t) + u(t)^T R_u(t) u(t)\,\mathrm{d}t + x(t_1)^T P_1 x(t_1)$$

where $x(t)$ implicitly depends on $u(t)$ through equation 1, is given by the linear control law

$$u^*(t) = -F(t)x(t)$$
$$F(t) = R_u^{-1}(t)B(t)^T P(t)$$

where the matrix $P(t)$ satisfies the matrix Riccati equation

$$-\dot{P} = A^T P + PA - PBR_u^{-1}B^T P + R_x$$

Furthermore, the resulting optimal cost from any given starting time t and initial condition $x(t)$ is given by

$$\mathcal{L}_{[t,t_1]}[u^*] = x(t)^T P(t) x(t)$$

An important caveat to this result is that in our above derivation we assumed that a unique optimum existed. However, nothing in the above guarantees that this solution for $u^*(t)$ is the only, or even the best (i.e. global), optimum of the cost functional. However, it turns out that if the following conditions are satisfied

(i) $R_x(t), R_u(t)$ are piecewise continuous functions of t (at least for $t \in [t_0, t_1]$)
(ii) $R_u(t)$ is a positive-definite symmetric matrix (at least for $t \in [t_0, t_1]$), and
(iii) $R_x(t)$ and P_1 are positive semi-definite symmetric matrices (at least for $t \in [t_0, t_1]$)

then there exists a unique optimum of the cost functional, which must therefore be the one we have derived. Moreover, this guarantees that the matrix Riccati equation has a unique solution, given by the matrix derived above in terms of partitions of the transition matrix of the variational equations (for which, under these conditions, the necessary inverse matrix is guaranteed to exist). For the proofs of these existence and uniqueness results, see the paragraph at the end of chapter 3.3 of [Kwakernaak and Sivan, 1972] for references.

Finally we state the following lemma due to [Wonham, 1968], which will be immensely useful in extending these results to stochastic systems:

Lemma 3.1. Consider the $n \times n$ matrix differential equation

$$-\dot{\tilde{P}} = (A - BF)^T \tilde{P} + \tilde{P}(A - BF) + R_x + F^T R_u F$$

with boundary condition $\tilde{P}(t_1) = P_1$, where R_x, R_u, and P_1 satisfy the conditions listed in the above paragraph, and $F(t)$ is an arbitrary continuous matrix function for $t \in [t_0, t_1]$. We then have that for any $t \in [t_0, t_1]$

$$\forall v \in \mathbb{R}^n \qquad v^T \tilde{P}(t) v \geq v^T P(t) v$$

where $P(t)$ satisfies the aforementioned matrix Riccati equation (cf. theorem 3.4) with the same boundary condition $P(t_1) = P_1$. This inequality becomes an equality iff

$$\forall \tau \in [t, t_1] \quad F(\tau) = R_u^{-1}(\tau) B(\tau)^T P(\tau)$$

The proof can be easily gleamed from our above results for the linear quadratic regulator. The first statement assumes the input takes on the form of a linear control law $u(t) = -F(t)x(t)$. The second statement follows from theorem 1.54, which tells us that the quantity $v^T \tilde{P}(t) v$ is the total future cost under this input from time t onwards, given the initial condition $x(t) = v$. Because we know from theorem 3.4 that the minimum such cost is given by $v^T P(t) v$, the statement follows. Finally, the last statement follows from the form of the optimal input u^* given by theorem 3.4, which is also a linear control law.

2.2 Infinite-horizon solution

Here we consider the special case of an **infinite horizon** on the cost function, i.e. the limit of $t_1 \to \infty$. We'll consider only the time-invariant case. See [Kalman, 1960a] for the more general time-varying case.

We first ask what happens to $P(t)$ in this limit. To do so, we recall from theorem 3.4 that

$$\mathcal{L}_{[t,t_1]}[u^*] = x(t)^T P(t) x(t)$$

and ask what happens to $\mathcal{L}$ as $t_1 \to \infty$. First, we note that if all of the unstable modes of A that live in the column space of R_x are controllable, then then the cost functional $\mathcal{L}$ remains finitely bounded even in the limit of $t_1 \to \infty$. This is because, if those unstable modes are controllable, then there exists an input that can bring them to 0 within some finite time, after which they will remain at 0 and the input can be set to 0 as well. The remaining unstable modes outside (i.e. orthogonal to) the column space of R_x are irrelevant, as these don't affect the cost functional. Second, we note that, if $P_1 = 0$, then $\mathcal{L}_{[t_0,t_1]}[u]$ is a non-decreasing function of t_1, since $\frac{\partial \mathcal{L}}{\partial t_1}$ is a quadratic quantity and therefore non-negative. Together, these two facts imply that $\mathcal{L}$ converges to a limit as $t_1 \to \infty$, entailing that $P(t)$ does too through the above equality. We thus have the following theorem:

Theorem 3.7a. In the case of a linear time-invariant system and constant weighting matrices R_x, R_u and no terminal cost $P_1 = 0$, the cost-to-go matrix $P(t)$, with dynamics given by the matrix Riccati equation of theorem 3.4, settles at a steady state $\bar{P}$ as $t_1 \to \infty$ iff all the unstable modes of A that live in the column space of R_x are controllable.

The sufficiency direction of this theorem was just proved. Necessity is easily proved by verifying that if the unstable modes of A that live in the column space of R_x are not controllable, then $\mathcal{L}$ diverges and thus $P(t)$ won't go to a steady state. If it does, then it need be that these modes are controllable.

We can re-state this theorem more succinctly in terms of reconstructibility. Using the eigendecomposition of the symmetric positive semi-definite matrix R_x, we can write

$$R_x = C^T R_y C$$

where the rows of C contain the eigenvectors of R_x with non-zero eigenvalues, and R_y is a diagonal positive definite matrix containing only the non-zero eigenvalues of R_x, which are positive. An interpretation of this decomposition was provided above, in terms of a quadratic cost on an output variable $y(t) = Cx(t)$ with a positive definite weighting matrix R_y such that $x(t)^T R_x x(t) = y(t)^T R_y y(t)$. For any positive definite R_y, the above theorem can then be re-stated in terms of reconstructibility: $P(t)$ goes to a

steady state as $t_1 \to \infty$ iff the unstable modes of the system that are reconstructible through C are also controllable through B.

So what is this steady state? By taking the matrix Riccati equation and simply setting the derivative to 0, we arrive at the following equation for the steady state $\bar{P}$, called the **continuous-time algebraic Riccati equation (CARE)**:

$$0 = A^T\bar{P} + \bar{P}A - \bar{P}BR_u^{-1}B^T\bar{P} + R_x$$

Solving this equation is non-trivial, but several techniques exist for solving the CARE, both analytically (cf. chapter 3.4.4 of [Kwakernaak and Sivan, 1972]) and numerically (cf. Wikipedia page for algebraic Riccati equation).

Note that this equation is quadratic in $\bar{P}$ and therefore may have multiple solutions, reflecting the fact that the dynamics defined by the matrix Riccati equation could in principle have multiple fixed points. However, it turns out we can extend the uniqueness results regarding theorem 3.4 to the current setting. To see this, we again consider the case of $P_1 = 0$, where

$$\begin{aligned}
\bar{u}^*(t) &\equiv \arg\min_{u(\cdot)} \lim_{t_1\to\infty} \int_{t_0}^{t_1} x(t)^T R_x(t)x(t) + u(t)^T R_u(t)u(t)\mathrm{d}t \\
&= \arg\min_{u(\cdot)} \int_{t_0}^{\infty} x(t)^T R_x(t)x(t) + u(t)^T R_u(t)u(t)\mathrm{d}t \\
&= \arg\min_{u(\cdot)} \mathcal{L}_{[t_0,\infty]}[u] \quad \text{with } P_1 = 0 \\
&= -R_u^{-1}B^T\bar{P}x(t)
\end{aligned}$$

where the last line follows from straight-forward application of theorem 3.4 with $t_1 = \infty$, together with the assumption that a steady state $\bar{P}$ exists (i.e. that the conditions of theorem 3.7a hold). Because this is the unique optimum of the cost functional, it follows that $\bar{P}$ is the unique steady state of $P(t)$ under the boundary condition $P(t_1) = 0$. Given that $P(t)$ is always positive semi-definite, this in turn implies that $\bar{P}$ is the unique positive semi-definite solution to the CARE. What if P_1 is a non-zero positive semi-definite matrix? In this case, the optimum remains exactly the same. Intuitively, the infinite integral makes the terminal cost term of the cost function negligible, leaving the cost function effectively equivalent to the case of $P_1 = 0$. More formally, we could postulate the existance of an input u that, under a non-zero P_1, provides a lower value of the cost functional than does the above defined steady state optimum $\bar{u}^*$. In other words, that there exists an input u that satisfies

$$\lim_{t_1\to\infty} \mathcal{L}_{[t_0,t_1]}[u] < \lim_{t_1\to\infty} \mathcal{L}_{[t_0,t_1]}[\bar{u}^*]$$

with a non-zero positive semi-definite terminal cost matrix P_1. Taking these limits, we obtain the following

$$\begin{aligned}
&\Leftrightarrow \lim_{t_1\to\infty} \left[\int_{t_0}^{t_1} x(t)^T R_x x(t) + u(t)^T R_u u(t)\mathrm{d}t + x(t_1)^T P_1 x(t_1)\right] \\
&\quad < \lim_{t_1\to\infty} \left[\int_{t_0}^{t_1} \bar{x}^*(t)^T R_x \bar{x}^*(t) + \bar{u}^*(t)^T R_u \bar{u}^*(t)\mathrm{d}t + \bar{x}^*(t_1)^T P_1 \bar{x}^*(t_1)\right] \\
&\Leftrightarrow \int_{t_0}^{\infty} x(t)^T R_x x(t) + u(t)^T R_u u(t)\mathrm{d}t + \lim_{t_1\to\infty} \left[x(t_1)^T P_1 x(t_1)\right] \\
&\quad < \int_{t_0}^{\infty} \bar{x}^*(t)^T R_x \bar{x}^*(t) + \bar{u}^*(t)^T R_u \bar{u}^*(t)\mathrm{d}t + \lim_{t_1\to\infty} \left[\bar{x}^*(t_1)^T P_1 \bar{x}^*(t_1)\right] \\
&\Leftrightarrow \int_{t_0}^{\infty} x(t)^T R_x x(t) + u(t)^T R_u u(t)\mathrm{d}t + \lim_{t_1\to\infty} \left[x(t_1)^T P_1 x(t_1)\right] \\
&\qquad\qquad < x(t_0)^T \bar{P} x(t_0) + \lim_{t_1\to\infty} \left[\bar{x}^*(t_1)^T P_1 \bar{x}^*(t_1)\right]
\end{aligned}$$

where $x(t), \bar{x}^*(t)$ are the states generated through the linear time-invariant dynamics by the inputs u, u^*, respectively. If all the unstable modes of the system both

(i) live in the column space of R_x, and

(ii) are controllable

then it is easy to anticipate that $\bar{u}^*$ will ensure that $\bar{x}^*(t) \to 0$ as $t \to \infty$: the unstable modes will be stabilized by $\bar{u}^*$ because they directly influence the cost functional and the remaining stable modes will decay to 0 regardless. In this case, then, $\lim_{t_1 \to \infty} \left[\bar{x}^*(t_1)^T P_1 \bar{x}^*(t_1)\right] = 0$ for any P_1 and the above inequality simplifies further to

$$\int_{t_0}^{\infty} x(t)^T R_x x(t) + u(t)^T R_u u(t) \mathrm{d}t + \lim_{t_1 \to \infty} \left[x(t_1)^T P_1 x(t_1)\right] < x(t_0)^T \bar{P} x(t_0)$$

But note that, by definition, the integral on the left-hand side is minimized by $\bar{x}^*, \bar{u}^*$, so that it must be greater than or equal to the right-hand side. Moreover, because P_1 is positive semi-definite, the other term on the left-hand side is non-negative. Together, these two facts entail that the left-hand side is greater than or equal to the right-hand side. This contradiction implies that there exists no input u that satisfies $\lim_{t_1 \to \infty} \mathcal{L}_{[t_0,t_1]}[u] < \lim_{t_1 \to \infty} \mathcal{L}_{[t_0,t_1]}[\bar{u}^*]$, meaning that, when P_1 is a non-zero positive semi-definite matrix and the above conditions (i) and (ii) hold, then there is a unique optimal input again given by $\bar{u}^*$. This in turn implies that there is a unique steady state $\bar{P}$ to which $P(t)$ converges in the limit of $t_1 \to \infty$. It is easy to see that this must be the unique positive semi-definite solution to the CARE, since any such matrix that satisfies this equation will be a steady state of $P(t)$, which, as we have just proven, has a unique steady state.

We can state these results more succinctly in terms of reconstructibility, as we did above, resulting in the following theorem:

Theorem 3.7b,c. Consider a linear time-invariant system and the aforementioned quadratic cost functional $\mathcal{L}_{[t_0,t_1]}[u]$ with constant weighting matrices:

- ▷ positive semi-definite $R_x = C^T R_y C$, where R_y is positive-definite
- ▷ positive-definite R_u

In the limit of $t_1 \to \infty$, we then have that, if

(i) the pair (A, C) is detectable, and

(ii) the pair (A, B) is stabilizable

then the optimal cost-to-go matrix $P(t)$ converges to a unique steady state $\bar{P}$, given by the symmetric positive semi-definite solution to the continuous-time Riccati equation (CARE)

$$0 = A^T \bar{P} + \bar{P} A - \bar{P} B R_u^{-1} B^T \bar{P} + R_x$$

which, under these conditions, is unique.

We also showed that, under the conditions of this theorem, the optimal input is given by the steady-state control law (theorem 3.7f)

$$\bar{u}^*(t) \equiv \arg\min_{u(\cdot)} \left[\lim_{t_1 \to \infty} \mathcal{L}_{[t_0,t_1]}[u]\right] = -R_u^{-1} B^T \bar{P} x(t)$$

for any positive semi-definite terminal cost matrix P_1. In this case, the optimal infinite-horizon cost is given by

$$\lim_{t_1 \to \infty} \mathcal{L}_{[t_0,t_1]}[\bar{u}^*] = x(t_0)^T \bar{P} x(t_0)$$

One last point worth noting is that if the pair (A, C) is detectable, then the optimal steady state control law will enforce asymptotically stable closed-loop dynamics, since all the unstable modes of A influence the cost functional and will thus be stabilized by the optimal input (theorem 3.7e). But if (A, C) are not detectable, this won't necessarily be the case.

2.3 Non-zero setpoints

In all the above, we considered optimizing the inputs of a linear control system so as to drive its state variable to 0 as efficiently as possible. Here we consider the setting in which the final desired state is not the 0 state, but some other non-zero **set point** x_s. We'll consider how to adapt the above tools of the linear quadratic regulator to solve this problem for linear time-invariant systems.

Our approach will consist of first identifying a constant input u_s that will ensure the desired set point can be maintained, i.e. that the desired set point is a steady state of the system. Such an input must satisfy the following equation:

$$0 = Ax_s + Bu_s$$

More generally, we might only care about the steady state y_s of some d-dimensional output variable $y(t) = Cx(t)$. Solving the above equation for x_s, we arrive at the following linear equation for u_s:

$$x_s = -A^{-1}Bu_s \Rightarrow y_s = Cx_s = -CA^{-1}Bu_s$$

Solving this last equation will give us a static input u_s that ensures the steady state of the system leads to the desired output steady state y_s. In solving this linear equation, however, we must consider the following three cases:

- ▷ Inputs and outputs are of same dimension ($m = d$): in this case, the matrix $CA^{-1}B$ is, in general, invertible and a unique solution for u_s exists
- ▷ Inputs are lower-dimensional than the outputs ($m < d$): in this case, there only exist solutions for some settings of y_s (namely, those in column space of $CA^{-1}B$), as you simply don't have enough degrees of freedom to freely set the output steady state
- ▷ Inputs are higher-dimensional than the outputs ($m > d$): in this case, many solutions for u_s exist. This ambiguity can be resolved by either adding dimensions to your output variable or by incorporating constraints on the steady state input u_s or the state variable steady state x_s.

Given u_s, we then ask what *additional* input $u'(t)$ will most efficiently drive the state to the desired set point x_s, at which point the system is settled at a steady state and $u'(t)$ can be shut off. We formalize this by defining the the *shifted* state and input variables

$$x'(t) = x(t) - x_s$$
$$u'(t) = u(t) - u_s$$

where the true input to the system $u(t) = u'(t) + u_s$ is taken to be composed of two components:

- ▷ a *sustained* component u_s that fixes the steady state of the system, and
- ▷ a *transient* component $u'(t)$ that drives the state variable to this steady state

Given the aforementioned relationship between x_s and u_s, we have that the dynamics of this new state variable are the same as the original one:

$$\dot{x'} = Ax + Bu = Ax' + Bu' + Ax_s + Bu_s = Ax' + Bu'$$

Most importantly, our goal is now to find the optimal transient input $u'(t)$ that will drive the shifted state variable $x'(t)$ to 0, since $x'(t) = 0 \Leftrightarrow x(t) = x_s$. Since $x'(t)$ is a time-invariant linear control system with input $u'(t)$, this is exactly the same problem we studied in sections 2.1 and 2.3! We thus have that, given a quadratic cost functional of the form considered in theorem 3.4, the optimal shifted input is given by the linear control law

$$u'(t) = -F(t)x'(t)$$
$$F(t) = R_u^{-1}B^T P(t)$$

where $P(t)$ satisifes the matrix Riccati equation. Moving back to our original state and input variables, this implies that the optimal input to the original system is

$$u(t) = -F(t)x(t) + F(t)x_s + u_s$$

Note that the first term on the right-hand side is just the standard linear quadratic regulator: this feedback term drives the state variable to reach a steady state as efficiently as possible. The other two terms are required to enusre the steady state is the right one; it is easy to verify that the steady state $\bar{x}$ of the resulting closed-loop dynamics

$$\dot{x} = (A - BF)x + B(Fx_s + u_s)$$

is the desired one:

$$\begin{aligned} 0 &= (A - BF)\bar{x} + B(Fx_s + u_s) \\ &= (A - BF)\bar{x} + BFx_s - Ax_s \\ &= (A - BF)(\bar{x} - x_s) \\ \Leftrightarrow \bar{x} &= x_s \end{aligned}$$

where in the second line we plugged in $Ax_s + Bu_s = 0 \Leftrightarrow Bu_s = -Ax_s$ and the last line follows whenever the conditions of theorem 3.7b,c hold, in which case the closed-loop dynamics must be asymptotically stable so all the eigenvalues of $A - BF$ are non-zero and negative (cf. last sentence of section 2.2) thus ensuring $A - BF$ is full rank.

In the infinite horizon limit, $F(t) \to \bar{F}$, in which case the optimal input is a steady state control law of the form

$$u(t) = -\bar{F}x(t) + \bar{F}x_s + u_s$$

It is straight-forward to verify that the sustained component $\bar{u}_s = \bar{F}x_s + u_s$ is equal to the input needed to fix the steady state output of the closed loop system to y_s, i.e. it satisfies $y_s = -C(A - B\bar{F})^{-1}B\bar{u}_s$.

2.4 Stochastic case

We now turn to solving the same problem outlined in section 2.1 but in the presence of white noise. We first state the theorem:

Theorem 3.9. Consider a linear control system with white noise

$$\dot{x} = Ax + Bu + w$$
$$x(t_0) = x_0$$

where $w(t)$ is a white noise stochastic process, i.e.

$$\mathbb{E}\left[w(t)\right] = 0$$
$$\mathbb{E}\left[w(t)w(t')^T\right] = V(t)\delta(t - t')$$

and the initial condition x_0 is a random variable with first and second moments

$$\mathbb{E}\left[x_0\right] = m_0$$
$$\mathbb{E}\left[x_0 x_0^T\right] = Q_0$$

Consider the quadratic cost functional of theorem 3.4, where the input weighting matrix $R_u(t)$ is positive definite and the state weighting matrices $R_x(t), P_1$ are positive semi-definite. The linear control law $u(t) = -F(t)x(t)$ that minimizes the expected value of this cost functional is then given by the that stated in theorem 3.4. The expected cost achieved by this input is given by

$$\mathbb{E}\left[\mathcal{L}_{[t_0,t_1]}[u^*]\right] = \mathrm{Tr}\left[P(t_0)Q_0 + \int_{t_0}^{t_1} P(t)V(t)\,\mathrm{d}t\right]$$

To prove this, all we need are theorem 1.54 and lemma 3.1. Setting $u(t) = -F(t)x(t)$ and using theorem 1.54, we first write:

$$\mathbb{E}\left[\mathcal{L}_{[t_0,t_1]}[u^*]\right] = \mathrm{Tr}\left[\tilde{P}(t_0)Q_0 + \int_{t_0}^{t_1} \tilde{P}(t)V(t)\,\mathrm{d}t\right]$$

where the matrix $\tilde{P}$ satisfies

$$-\dot{\tilde{P}} = (A - BF)^T\tilde{P} + \tilde{P}(A - BF) + R_x + F^T R_u F$$

Given that Q_0 and $V(t)$ are covariance matrices, they are positive semi-definite and can thus be written as sums of such outer products

$$Q_0 = \sum_i v_0^{(i)} v_0^{(i)^T}, \quad V(t) = \sum_i v_w^{(i)}(t) v_w^{(i)}(t)^T$$

(e.g. via their eigendecomposition), entailing that

$$\begin{aligned}\mathbb{E}\left[\mathcal{L}_{[t_0,t_1]}[u^*]\right] &= \sum_i \mathrm{Tr}\left[\tilde{P}(t_0)v_0^{(i)}v_0^{(i)^T}\right] + \sum_i \int_{t_0}^{t_1} \mathrm{Tr}\left[\tilde{P}(t)v_w^{(i)}(t)v_w^{(i)}(t)^T\right] t\,\mathrm{d} \\ &= \sum_i v_0^{(i)^T}\tilde{P}(t_0)v_0^{(i)} + \sum_i \int_{t_0}^{t_1} v_w^{(i)}(t)^T\tilde{P}(t)v_w^{(i)}(t)\,\mathrm{d}t\end{aligned}$$

Therefore, the expected cost is minimized when each of these quadratic terms is minimized. Critically, from lemma 3.1 we have that, for any $v \in \mathbb{R}^n$,

$$v^T \tilde{P} v \geq v^T P v$$

where $P(t)$ satisfies

$$\begin{aligned} -\dot{P} &= (A - BR_u^{-1}B^T P)^T P + P(A - BR_u^{-1}B^T P) + R_x + (R_u^{-1}B^T P)^T R_u R_u^{-1} B^T P \\ &= A^T P - PBR_u^{-1}B^T P + PA - PBR_u^{-1}B^T P + R_x + PBR_u^{-1}B^T P \\ &= A^T P + PA - PBR_u^{-1}B^T P + R_x \end{aligned}$$

Thus the minimum expected cost is achieved when $\tilde{P}(t) = P(t)$, which occurs when $F(t) = R_u^{-1}(t)B(t)^T P(t)$. This is exactly the control law stated in theorem 3.4, thus proving the theorem.

We note that we have only derived the optimal *linear control law*, without showing that this is the true optimal control input $u(t)$. It turns out this is indeed the optimal input only when the white noise $w(t)$ is Gaussian. [Åström, 2012] Under this linear control law, all the results of sections 2.2 and 2.3 naturally extend to this stochastic setting.

2.5 Tracking problems

Here we consider a particular type of problem that can be tackled using these methods. In a **tracking problem**, we are given a **reference** output variable $y^*(t)$ that we'd like to track using an output $y(t) = C(t)x(t)$ of our linear control system. The problem is thus to find the input $u(t)$ that will most efficiently minimize the **tracking error**

$$e(t) = y(t) - y^*(t)$$

In order to derive such an input, a natural cost functional to minimize is the cumulative squared error

$$\mathcal{L}_{[t_0,t_1]}[u] = \int_{t_0}^{t_1} e(t)^T R_e(t) e(t) + u(t)^T R_u(t) u(t) \, \mathrm{d}t$$

given symmetric positive definite weighting matrices $R_e(t), R_u(t)$.

To solve this problem using the linear quadratic regulator of theorems 3.4 and 3.9, we'll assume that the reference variable is the output of a linear system driven by white noise:

$$\begin{gathered} y^*(t) = C^*(t)x^*(t) \\ \dot{x}^* = A^* x + w \\ \mathbb{E}\left[w(t)\right] = 0, \quad \mathbb{E}\left[w(t)w(t')^T\right] = V(t)\delta(t - t') \end{gathered}$$

To massage this system into the form considered in these theorems, we define the augmented state variable

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ x^*(t) \end{bmatrix}$$

resulting in the augmented system

$$\dot{\tilde{x}} = \underbrace{\begin{bmatrix} A & 0 \\ 0 & A^* \end{bmatrix}}_{\tilde{A}} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix}}_{\tilde{B}} + \underbrace{\begin{bmatrix} 0 \\ w \end{bmatrix}}_{\tilde{w}}$$

where $\tilde{w}(t)$ is a white noise process with intensity function

$$\tilde{V}(t) = \begin{bmatrix} 0 & 0 \\ 0 & V(t) \end{bmatrix}$$

The error can now be expressed as an output variable of this augmented system

$$\begin{gathered} e(t) = \tilde{C}(t)\tilde{x}(t) \\ \tilde{C}(t) = \begin{bmatrix} C(t) & -C^*(t) \end{bmatrix} \end{gathered}$$

allowing us to re-write the cost functional as

$$\mathcal{L}_{[t_0,t_1]}[u] = \int_{t_0}^{t_1} \tilde{x}(t)^T R_{\tilde{x}}(t)\tilde{x}(t) + u(t)^T R_u(t)u(t)\,\mathrm{d}t$$

where $R_{\tilde{x}} = \tilde{C}^T R_e \tilde{C}$. Given the stochastic linear dynamics of $\tilde{x}(t)$, the optimal linear control law follows immediately from theorem 3.9:

$$u^*(t) = -F(t)\tilde{x}(t)$$
$$F(t) = R_u^{-1}(t)\tilde{B}(t)^T P(t)$$

where $P(t)$ satisfies the matrix Riccati equation stated in theorem 3.4, with terminal condition $P(t_1) = 0$.

To get some intuition for the nature of this solution, we partition the symmetric $2n \times 2n$ matrix $P(t)$ into $n \times n$ blocks,

$$P(t) = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}$$

so that the optimal linear control law can be expressed as

$$u^*(t) = -F_{\mathrm{fb}}(t)x(t) - F_{\mathrm{ff}}(t)x^*(t)$$
$$F_{\mathrm{fb}}(t) = R_u^{-1}(t)B(t)^T P_1 1(t), \quad F_{\mathrm{ff}}(t) = R_u^{-1}(t)B(t)^T P_1 2(t)$$

where, from the matrix Riccati equation, we have

$$-\dot{P}_{11} = A^T P_{11} + P_{11}A - P_{11}BR_u^{-1}B^T P_{11} + C^T R_e C$$
$$-\dot{P}_{12} = A^T P_{12} + P_{12}A^* - P_{11}BR_u^{-1}B^T P_{12} - C^T R_e C^*$$

We thus note that, much like in section 2.3, the optimal linear control law separates into two components, only one of which depends on the properties of the tracking variable:

- ▷ a *feedback* component $F_{\mathrm{fb}}(t)x(t)$ that, through its dependence on the control system parameters A, B, C, drives the state variable to a steady state as efficiently as possible. Note that this feedback component is exactly the linear quadratic regulator from theorem 3.4, which was designed to do exactly that.
- ▷ a *feedforward* component $F_{\mathrm{ff}}(t)x^*(t)$ that, through its dependence on the control system parameters A, B, C and reference variable parameters A^*, C^*, efficiently guides the state variable towards the right steady state at each moment in time in order to minimize the tracking error

This separation of feedback and feedforward components arises due to the linearity of the systems. This linearity allows one to effectively divide such problems into two: (i) drive the system to the right steady state, and (ii) make it get there quickly. Part (ii) is effectively the problem studied in section 2.1, and its linear solution is the linear quadratic regulator.

Chapter 4: Optimal State Reconstruction

3.1 Observers, full- and reduced- order

3.2 Optimal observers

3.3 The innovation process

Chapter 5: Optimal Output Feedback Control

4.1 The separation principle

LQG

4.2 Reduced-order output feedback controllers

References

[Åström, 2012] Åström, K. J. (2012). *Introduction to Stochastic Control Theory*. Courier Corporation.

[Kalman, 1960a] Kalman, R. E. (1960a). Contributions to the theory of optimal control. *Bol. soc. mat. mexicana*, 5(2):102–119.

[Kalman, 1960b] Kalman, R. E. (1960b). On the general theory of control systems. In *Proceedings First International Conference on Automatic Control, Moscow, USSR*.

[Kalman et al., 1969] Kalman, R. E., Falb, P. L., and Arbib, M. A. (1969). *Topics in Mathematical System Theory*, volume 1. McGraw-Hill New York.

[Kwakernaak and Sivan, 1972] Kwakernaak, H. and Sivan, R. (1972). *Linear Optimal Control Systems*, volume 1. Wiley-interscience New York.

[Oksendal, 2013] Oksendal, B. (2013). *Stochastic Differential Equations: An Introduction with Applications*. Springer Science & Business Media.

[Wonham, 1968] Wonham, W. M. (1968). On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697.