SF2832 Systems Theory - Course Summary

Leo Trolin

08/01-2025

Basic

Definition: (**Basic properties of a system**)
Let $y(t) = f_\Sigma(u(t))$ denote a system.

- The system is relaxed if $f_\Sigma(0) = 0$.
- The system is linear if $f_\Sigma(\alpha u_1(t) + \beta u_2(t)) = \alpha f_\Sigma(u_1(t)) + \beta f_\Sigma(u_2(t))$.
- The system is memoryless if $y(t)$ only depends on the current input $u(t)$.
- The system is time-invariant if $\forall T > 0$, $y_T(t) = f_\Sigma(u_T(t))$, where $u_T(t) = u(t-T)$ if $t - t_0 \geq T$ and 0 else.

Definition: (**Input-output description**)
The input-output description of a linear model is

$$y(t) = \int_{t_0}^{t} G(t,s)u(s)\ \mathrm{d}s + D(t)u(t)$$

Here, G is called the impulse response. (achieved if $u = \delta$)

- The system is always relaxed and linear.
- The system is memoryless if $G = 0$.
- The system is time-invariant if $G(t,s) = G(t-s)$.
- G is finite-dimensional if $G(t,s) = H(t)K(s)$ for some matricies H and K.

Definition: (**State space model**)
A system is a state space model if it is on the form

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

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$.

Definition: (**State transition matrix**)
The state transition matrix $\Phi(t,t_0)$ to a system is the unique solution to

$$\begin{cases}\dot{\Phi}(t,t_0) = A(t)\Phi(t,t_0)\\ \Phi(t_0,t_0) = I\end{cases}$$

↳ Comment: Each column $\phi_i(t)$ in $\Phi(t,t_0)$ satisfies $\dot{\phi}_i(t) = A(t)\phi_i(t)$ and $\phi_i(t_0) = e_i$.

Theorem:
If the system

$$\begin{cases}\dot{x}(t) = A(t)x(t)\\ x(t_0) = a\end{cases}$$

has the solution $x(t)$, then $x(t) = \Phi(t,t_0)a$.

Definition: (**Fundamental matrix**)
A fundamental matrix $\Psi(t)$ to a system is a matrix that contains linearly independent solutions $\psi_i(t)$ in its columns that satisfy $\dot{\psi}_i(t) = A(t)\psi_i(t)$.

↳ Comment: By linear independence, $\Psi(t)$ is non-singular for all t.

Theorem:
$\Phi(t,t_0) = \Psi(t)\Psi^{-1}(t_0)$ for any t and any fundamental matrix $\Psi(t)$.

Theorem:
$\Phi(t,s)^{-1} = \Phi(s,t)$.

Theorem:
$\Phi(t,\tau)\Phi(\tau,t_0) = \Phi(t,t_0)$.

↳ Intuition: We can partition the solution at τ.

Theorem: (**Solution to state space model**)
The state space model

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ x(t_0) = a \end{cases}$$

has the solution

$$x(t) = \Phi(t, t_0)a + \int_{t_0}^{t} \Phi(t, s)B(s)u(s)\ \mathrm{d}s.$$

Theorem: (**Input-output description ↔ State space model**)
The input-output description agrees with a state space model where $G(t, s) = C(t)\Phi(t, s)B(s)$.

Definition: (**Matrix exponential**)
For a matrix A, we define its matrix exponential as

$$e^{At} := \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$$

or equivalently,

$$e^{At} := \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}.$$

Theorem: (**Matrix exponential properties**)
For a matrix A:

- $\left\|e^{At}\right\| < \infty$.
- If A is diagonal, then e^{At} is given by element-wise exponentiation of At.
- $e^{PAP^{-1}t} = Pe^{At}P^{-1}$.
- If A_1 and A_2 commute, then $e^{(A_1+A_2)t} = e^{A_1 t}e^{A_2 t}$.
- $\left(e^{At}\right)^{-1} = e^{-At}$.
- $\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = Ae^{At}$.

Theorem:
For a time-invariant system, the state transition matrix is given by $\Phi(t, s) = \Phi(t - s) = e^{A(t-s)}$.

Controllability & Observability

Definition: (**Controllable**)
A system

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

is controllable if for all x_1 we can find some continuous $u(t)$ such that $x(t_1) = x_1$ for some t_1.

Definition: (**Reachable**)
Coincides with "Controllable", but with $x_0 = 0$.

Definition: (**Null-controllable**)
Coincides with "Controllable", but with $x_1 = 0$.

Theorem:
"Controllable" $\iff$ "Reachable" $\iff$ "Null-controllable" $\iff$ $\forall d \in \mathbb{R}^n\ \exists u(t) : \int_{t_0}^{t_1} \Phi(t, s)B(s)u(s)\ \mathrm{d}s = d$.

Definition: (**Reachability Gramian**)
The reachability Gramian to a system is

$$W(t_0, t_1) := \int_{t_0}^{t_1} \Phi(t_1, s)B(s)B^{\mathrm{T}}(S)\Phi^{\mathrm{T}}(t_1, s)\ \mathrm{d}s.$$

Theorem:
The reachability Gramian $W(t_0, t_1)$ is symmetric and positive definite for all $t_0 < t_1$.

Theorem:
A system is controllable iff $W(t_0, t_1)$ is nonsingular.

Theorem:
The state transfer from $x(t_0) = x_0$ to $x(t_1) = x_1$ is possible iff $x_1 - \Phi(t_1, t_0)x_0 \in \operatorname{im} W(t_0, t_1)$.
The solution $u(t)$ that minimizes $\int_{t_0}^{t_1} u^{\mathrm{T}}(s)u(s)\ \mathrm{d}s$ is given by $u(t) = B^{\mathrm{T}}(t)\Phi^{\mathrm{T}}(t_1, t)a$ where a is a solution to $W(t_0, t_1)a = x_1 - \Phi(t_1, t_0)x_0$.

Theorem:
The rows of $\Phi(t_1, t)B(t)$ are linearly independent over $t \in [t_0, t_1]$ iff $W(t_0, t_1)$ is nonsingular.

Definition: (**Reachability matrix**)
For a time-invariant system, the reachability matrix is defined as $\Gamma := [B, AB, \dots, A^{n-1}B]$. ($A$ is $n \times n$)

Theorem:
For a time-invariant system, $\operatorname{im} W(t_0, t_1) = \operatorname{im} \Gamma$.

↳ Corollary: A time-invariant system is controllable iff $\operatorname{rank} \Gamma = n$. In the case where u is one-dimensional, this means that Γ is nonsingular.

Definition: (**Reachable subspace**)
For a time-invariant system, we denote the reachable subspace as $\mathcal{R} := \operatorname{im} \Gamma$.

Theorem:
The reachable subspace $\mathcal{R}$ is A-invariant, that is $\forall x \in \mathcal{R}$, $Ax \in \mathcal{R}$.

↳ Corollary: Also $A^k x \in \mathcal{R}$ for any $k \in \mathbb{N}$, and $e^{At}x \in \mathcal{R}$ for any t.

Theorem:
A time-invariant can be transferred from $x(t_0) \in \mathcal{R}$ to $x(t_1) \in \mathcal{R}$ in time ε for any $\varepsilon > 0$.

Theorem:
For a time-invariant system, if $x(t_0) \in \mathcal{R}$ then it is impossible for $x(t) \notin \mathcal{R}$ for any t.
If $x(t_0) \notin \mathcal{R}$ then it is impossible for $x(t) \in \mathcal{R}$ for any t.

Definition: (**Observable**)
A state-space model is observable if given $u(t)$, $y(t)$ we can reconstruct $x(t)$.

Theorem:
A state-space model is obserable iff $C(t)\Phi(t, t_0)x_0 = v(t)$ has a unique solution x_0 for all $v(t)$.

Definition: (**Observability Gramian**)
The observability Gramian to a system is

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

Theorem:
A system is observable iff its observability Gramian $M(t_0, t_1)$ is nonsingular.

Theorem:
If $M(t_0, t_1)$ is singular, then two initial states $x(t_0) = a$, $x(t_0) = b$ will produce the same $y(t)$ on $t \in [t_0, t_1]$ iff $a - b \in \ker M(t_0, t_1)$.

Definition: (**Observability matrix**)
For a time-invariant system, the observability matrix is defined as

$$\Omega := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

Theorem:
For a time-invariant system, $\ker M(t_0, t_1) = \ker \Omega$.

↳ Corollary: A time-invariant system is observable iff Ω has full column rank. In the case where y is one-dimensional, this means that Ω is nonsingular.

Definition: (**Unobservable subspace**)
For a time-invariant system, we call the unobservable subspace $\ker \Omega$.

Theorem:
The unobservable subspace is A-invariant.

↳ Corollary: If $x_0 \in \ker \Omega$, then $y = Ce^{At}x_0 = 0$.

Stability

Definition: (**Equilibrium**)
For a general time-invariant system $\dot{x} = f(x)$, we say that x^0 is an equilibrium if $f(x^0) = 0$.

Definition: (**Asymptotically stable**)
The system $\dot{x} = Ax$ is asymptotically stable if $x(t) \to 0$ as $t \to \infty$, for all $x(t_0) \in \mathbb{R}^n$.

Definition: (**Stable, Unstable**)
The system $\dot{x} = Ax$ is stable if $\|x(t)\| < \infty$, for all $x(t_0) \in \mathbb{R}^n$. Otherwise it is unstable.

Theorem:
For $\dot{x} = Ax$:

- Asymptotically stable $\iff$ $e^{At} \to 0$ as $t \to \infty$
- Stable $\iff$ $\|e^{At}\| < \infty, \forall t \geq 0$

Definition: (**Stable matrix**)
A matrix is stable if the real parts of all its eigenvalues are strictly negative.

Theorem:
The system $\dot{x} = Ax$ is asymptotically stable iff A is a stable matrix.

Theorem:
For $\dot{x} = Ax$ where A has eigenvalues $\lambda_\nu = \sigma_\nu + i\omega_\nu$:

- Asymptotically stable $\iff$ $\sigma_\nu < 0 \ \forall \nu$
- Stable $\iff$ $\sigma_\nu \leq 0 \ \forall \nu$; and $J_\nu = [\lambda_\nu]$ for all ν s.t. $\sigma_\nu = 0$ (If the real part of an eigenvalue is zero, its Jordan block must be 1×1)
- Unstable $\impliedby$ $\exists \sigma_\nu > 0$

Definition: (**Input-output stability**)
The system $\dot{x} = Ax + Bu$, $y = Cx$ (assumed: $x(0) = 0$) is input-output stable if for all bounded $u(t)$ (assumed: bounded by 1), $y(t)$ will be bounded.

Theorem:
For $\dot{x} = Ax + Bu$, $y = Cx$, assume that (A, B) is reachable and (C, A) is observable. Then:

$$\text{Input-output stability} \iff A \text{ is a stable matrix}$$

↳ Comment: Without the assumptions, only " $\impliedby$ " holds.

Theorem:
For $\dot{x} = Ax + Bu$, $y = Cx$, if A is a stable matrix then $y(t) \in L_p[0, \infty) \ \forall u(t) \in L_p[0, \infty)$.

↳ Comment: $u(t) \in L_p[0, \infty)$ means that $\int_0^\infty \|u(t)\|^p \ \mathrm{d}t < \infty$.

Definition: (**Positive definite**)
A matrix P is positive definite if $P^{\mathrm{T}} = P$ and all eigenvalues to P are positive.

Theorem:
For a positive definite matrix P with eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$:

$$\lambda_1 \|x\|^2 \leq x^{\mathrm{T}} P x \leq \lambda_n \|x\|^2$$

Definition: (**Lyapunov equation**)
For given matrices A and Q, the Lyapunov equation is

$$A^{\mathrm{T}}P + PA = -Q.$$

Theorem:
For any $Q \in \mathbb{R}^{n\times n}$, if A is a stable matrix, then

$$P = \int_0^\infty e^{A^{\mathrm{T}}t} Q e^{At} \,\mathrm{d}t$$

is the unique solution to the Lyapunov equation.

Theorem:
Let (C, A) be observable. Then the following are equivalent:

- $A^{\mathrm{T}}P + PA = -C^{\mathrm{T}}C$ has a solution $P > 0$.
- A is a stable matrix.

Realizations

Definition: (**Realization, dimension, transfer matrix**)
Matrices (A, B, C) that satisfy $Ce^{At}B = G(t)$ is a realization of $G(t)$.
Alternatively, $C(sI - A)^{-1}B = R(s)$ where $\mathcal{L}\left\{G(t)\right\} = R(s)$ is called the transfer matrix of the system.
The dimension of A is called the dimension of the realization.

↳ Comment: In the case where D is present, we say that $R(s) = C(sI - A)^{-1}B + D$.

Definition: (**Minimal realization**)
A realization is minimal if no other realization has lower dimension.

Theorem:
A transfer function $R(s)$ is realizable iff it is a strictly proper rational matrix.

↳ Comment: If $R(s)$ is just proper, we can let $\bar{R}(s) = R(s) - R(\infty)$ so that $\bar{R}(s)$ is strictly proper and $D = R(\infty)$.

Method: (**Standard reachable realization**)
Let $R(s)$ be a $m \times k$ strictly proper rational matrix.

- Take $\chi(s)$ be the least common denominator of all the elements of $R(s)$. Write $\chi(s) = s^r + a_1 s^{r-1} + \cdots + a_r$.
- Write $\chi(s)R(s) = N_0 + N_1 s + \cdots + N_{r-1}s^{r-1}$.

Then the realization on block form is:

$$A = \begin{bmatrix} 0 & I_k & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & I_k \\ -a_r I_k & \cdots & -a_2 I_k & -a_1 I_k \end{bmatrix} (rk \times rk) \qquad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_k \end{bmatrix} (rk \times k) \qquad C = \begin{bmatrix} N_0 & \cdots & N_{r-1} \end{bmatrix} (m \times rk)$$

Definition: (**Markov parameters**)
For a (strictly proper) transfer matrix $R(s)$, its Markov parameters R_i are given by the Laurent expansion $R(s) = R_1 s^{-1} + R_2 s^{-2} + \ldots$.

Theorem:
Some useful Laurent expansions are:

$$\frac{1}{s-a} = \sum_{k=1}^{\infty} \underbrace{a^{k-1}}_{R_k} s^{-k}, \qquad \frac{1}{(s-a)^2} = \sum_{k=1}^{\infty} \underbrace{a^{k-2}(k-1)}_{R_k} s^{-k}, \qquad \frac{1}{(s-a)(s-b)} = \sum_{k=1}^{\infty} \underbrace{\frac{1}{a-b}\left(a^{k-1} - b^{k-1}\right)}_{R_k} s^{-k}$$

Method: (**Standard observable realization**)
Let $R(s)$ be a $m \times k$ strictly proper rational matrix.

- Take $\chi(s)$ be the least common denominator of all the elements of $R(s)$. Write $\chi(s) = s^r + a_1 s^{r-1} + \cdots + a_r$.
- Find the Markov parameters $R_1, \dots, R_r$ of $R(s)$.

Then the realization on block form is:

$$A = \begin{bmatrix} 0 & I_m & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & I_m \\ -a_r I_m & \cdots & -a_2 I_m & -a_1 I_m \end{bmatrix} \ (rm \times rm) \qquad B = \begin{bmatrix} R_1 \\ \vdots \\ R_r \end{bmatrix} \ (rm \times k) \qquad C = \begin{bmatrix} I_m & 0 & \cdots & 0 \end{bmatrix} \ (m \times rm)$$

Theorem:
For a transfer matrix $R(s)$ with Markov parameters R_i, the matrices (A, B, C) are a realization of $R(s)$ iff $CA^{i-1}B = R_i$ for $i = 1, 2, \dots$.

Method: (**Kalman decomposition**)
Given a realization (A, B, C) of $R(s)$ where A is $n \times n$, we want to find a new realization with lower dimension if possible.

- Let $V_{\bar{o}r} = \operatorname{im} \Gamma \cap \ker \Omega$ (Not observable; Reachable)
- Define V_{or} by $\operatorname{im} \Gamma = V_{\bar{o}r} \oplus V_{or}$ (Observable; Reachable)
- Define $V_{\bar{o}\bar{r}}$ by $\ker \Omega = V_{\bar{o}r} \oplus V_{\bar{o}\bar{r}}$ (Not observable; Not reachable)
- Define $V_{o\bar{r}}$ by $\mathbb{R}^n = V_{\bar{o}r} \oplus V_{or} \oplus V_{\bar{o}\bar{r}} \oplus V_{o\bar{r}}$ (Observable; Not reachable)
- Let $T \in \mathbb{R}^{n \times n}$ be a matrix with the basis vectors in $V_{\bar{o}r}$, V_{or}, $V_{\bar{o}\bar{r}}$, $V_{o\bar{r}}$ as columns.
- Compute

$$\tilde{A} = T^{-1}AT = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & A_{33} & A_{34} \\ 0 & 0 & 0 & A_{44} \end{bmatrix}, \qquad \tilde{B} = T^{-1}B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \qquad \tilde{C} = CT = \begin{bmatrix} 0 & C_2 & 0 & C_4 \end{bmatrix}$$

- The new realization is (A_{22}, B_2, C_2).

↳ Intuition: Note that we build V_{or}, $V_{\bar{o}\bar{r}}$ and $V_{o\bar{r}}$ as complements. There is no "unreachable (vector) space" or "observable (vector) space".

↳ Comment: This new realization will be minimal!

Definition: (**Hankel matrix**)
For a transfer matrix $R(s)$ with Markov parameters R_i, the corresponding Hankel matrix is

$$H_i = \begin{bmatrix} R_1 & \cdots & R_i \\ \vdots & \ddots & \vdots \\ R_i & \cdots & R_{2i-1} \end{bmatrix}.$$

Theorem:
For a transfer matrix $R(s)$, $\operatorname{rank} H_i = \operatorname{rank} H_r \ \forall i \geq r$ where $r = \deg \chi$ is the degree of the least common denominator of all the elements of $R(s)$.

Theorem:
A realization is minimal if and only if it is reachable and observable.

Theorem:
If (A, B, C) is a realization of $R(s)$ and T is a transform, then $(\tilde{A} = TAT^{-1}, \tilde{B} = TB, \tilde{C} = CT^{-1})$ is also a realization of $R(s)$. $(\tilde{x} = Tx)$

Theorem:
Given two minimal realizations (A, B, C) and $(\tilde{A}, \tilde{B}, \tilde{C})$ there exists a transform T such that $\tilde{A} = TAT^{-1}$, $\tilde{B} = TB$, $\tilde{C} = CT^{-1}$. $(\tilde{x} = Tx)$

↳ Comment: This transform is given by $T = (\tilde{\Omega}^{\mathrm{T}}\tilde{\Omega})^{-1}\tilde{\Omega}^{\mathrm{T}}\Omega = \tilde{\Gamma}\tilde{\Gamma}^{\mathrm{T}}(\Gamma\tilde{\Gamma}^{\mathrm{T}})^{-1}$.

Definition: (**McMillan degree**)
The McMillan degree $\delta(R)$ of a transfer matrix $R(s)$ is the dimension of its minimal realization.

Theorem:
$\delta(R) = \operatorname{rank} H_r$ where $r = \deg \chi$ is the degree of the least common denominator of all the elements of $R(s)$.

Theorem:
Let $R(s)$ be a transfer matrix. The minors of $R(s)$ are the determinants of all square matrices contained in $R(s)$ (from all combinations of row and column indices). Let $\rho(R)$ be the least common denominator of all minors of $R(s)$. Then, $\delta(R) = \deg \rho(R)$.

Pole Placement & Observers

Theorem:
Let $\mathcal{R} = \operatorname{im}[B, AB, \dots, A^{n-1}B]$ and $\mathcal{R}_K = [B, (A+BK)B, \dots, (A+BK)^{n-1}B]$. Then $\mathcal{R} = \mathcal{R}_K$.

↳ Intuition: Using the controller $u = Kx + v$ does not affect reachability.

Theorem:
The pole placement problem is solvable iff (A, B) is reachable. That is, for any polynomial $\varphi(s)$ of degree n, it is possible to find a K such that $\det(sI - (A+BK)) = \varphi(s)$.

↳ Intuition: Solving the pole placement means using a controller $u = Kx$ to move the poles of the system, or the eigenvalues of the matrix, to arbitrary locations.

Theorem:
We can assign arbitrary eigenvalues to $A - LC$ by choosing L iff (C, A) is observable.

↳ Intuition: This corresponds to using an observer $\hat{x}$ defined by $\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$. Then, the error $e(t) := x(t) - \hat{x}(t)$ obeys $\dot{e}(t) = (A - LC)e(t)$. Therefore we want to choose the eigenvalues of $A - LC$ to have negative real parts.

Linear Quadratic Optimal Control

Theorem & Notation:
Consider the linear quadratic optimal control problem

$$\min_u J(u) = x(t_1)^{\mathrm{T}} S x(t_1) + \int_{t_0}^{t_1} x(t)^{\mathrm{T}} Q x(t) + u(t)^{\mathrm{T}} R u(t) \; \mathrm{d}t$$

$$\text{such that} \begin{cases} \dot{x} = Ax + Bu \\ x(t_0) = x_0 \end{cases}$$

where $S \geq 0$, $Q \geq 0$, $R > 0$. Then the optimal control is given by

$$u^*(t) = -R^{-1}B^{\mathrm{T}}P(t)x(t) =: -K(t)x(t)$$

where $K(t) = R^{-1}B^{\mathrm{T}}P(t)$ is called the Kalman gain, and $P(t)$ satisfies the Dynamical Riccati Equation (DRE)

$$\begin{cases} \dot{P}(t) = -A^{\mathrm{T}}P - PA + PBR^{-1}B^{\mathrm{T}}P - Q \\ P(t_1) = S \end{cases}$$

and the optimal cost is given by

$$V(x_0) = x_0^{\mathrm{T}} P(t_0) x_0$$

So, the resulting optimal x can be expressed by:

$$\dot{x}(t) = \big(A - BK(t)\big)\, x(t) =: A_K(t)x(t) \implies x(t) = \Phi_K(t_1, t_0)x_0$$

Theorem:
The Dynamical Riccati Equation has a unique solution P on the interval $[t_0, t_1]$ which is positive semidefinite and bounded.

Method: (**Solving DRE**)
Solve

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^{\mathrm{T}} \\ -Q & -A^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}$$
$$\begin{bmatrix} X(t_1) \\ Y(t_1) \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix}$$

Then the solution to the Dynamical Riccati Equation is given by $P = YX^{-1}$.

Definition: (**Infinite time horizon LQ**)
A LQ control problem on infinite time horizon is on the form

$$\min_u J(u) = \int_0^\infty x(t)^{\mathrm{T}}Qx(t) + u(t)^{\mathrm{T}}Ru(t)\ \mathrm{d}t$$
$$\text{such that} \begin{cases} \dot{x} = Ax + Bu \\ x(t_0) = x_0 \end{cases}$$

where $Q \geq 0,\ R > 0$.

Related to this, we have

$$\min_u J(u) = \int_0^{t_1} x(t)^{\mathrm{T}}Qx(t) + u(t)^{\mathrm{T}}Ru(t)\ \mathrm{d}t$$
$$\text{such that} \begin{cases} \dot{x} = Ax + Bu \\ x(t_0) = x_0 \end{cases}$$

where the optimal cost is $x_0^{\mathrm{T}}P(t_1 - 0)x_0$ (abuse of notation: $P(t) \leftrightarrow P(t_1 - t)$).

Definition: (**Feasible**)
For an infinite time horizon LQ problem, a control $u(t)$ is feasible if $J(u)$ is finite.

Theorem:
For an finite time horizon LQ problem with $S = 0$, let $Q = C^{\mathrm{T}}C$. If (C, A) is observable, then $P(t_1 - 0) > 0\ \forall t_1 > 0$.

Theorem:
For an finite time horizon LQ problem with $S = 0$, assume $P > 0$. Then for any feasible control u, we have $x(t) \to 0$ as $t \to \infty$.

Theorem:
Let (A, B, C) be a minimal realization. Then the Algebraic Riccati Equation (ARE)

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

has a unique real positive definite solution P. The optimal control corresponding to the infinite time horizon LQ problem (with $Q = C^{\mathrm{T}}C$) is

$$u = -R^{-1}B^{\mathrm{T}}Px$$

and the optimal cost is given by $x_0^{\mathrm{T}}Px_0$.

Kalman Filtering

Theorem: (**Least squares estimate**)
Let y be a vector of elements $y_1, \ldots, y_N$ in a Hilbert space $\mathcal{H}$. Let $k^{\mathrm{T}} \in \mathbb{R}^N$. Let $x \in \mathcal{H}$. The problem of least squares estimation, that is choosing k to minimize

$$\|x - ky\|^2$$

is given by

$$k^* = x \cdot y^{\mathrm{T}}(y \cdot y^{\mathrm{T}})^{-1} \qquad \Longrightarrow\ \hat{x} = k^*y = x \cdot y^{\mathrm{T}}(y \cdot y^{\mathrm{T}})^{-1}y$$

assuming that $y \cdot y^{\mathrm{T}}$ is invertible, which is true if the components of y are linearly independent.

Definition: ($[y]$)
Let y be a vector of elements $y_1, \ldots, y_N$ in a Hilbert space $\mathcal{H}$. Then we define

$$[y] := \left\{ ky : k^{\mathrm{T}} \in \mathbb{R}^N \right\}$$

i.e. the span of the components of y.

Theorem: (**Orthogonal projection**)
Let $\mathcal{H}$ be a Hilbert space. Let $h \in \mathcal{H}$. Let $M \subset \mathcal{H}$ be a subspace of $\mathcal{H}$ consisting of estimations ("$[y]$"). Then, $\hat{m} \in M$ is the best estimation of h among all points in M if and only if $(h - \hat{m}) \cdot m = 0 \ \forall m \in M$.

Definition: (E^M)
Reusing notations as in the theorem above, we denote $E^M h$ as the orthogonal projection of $h \in \mathcal{H}$ onto the subspace $M \subset \mathcal{H}$, i.e. $\hat{m} =: E^M h$.

Theorem: (**Properties of E^M**)

- $E^M(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 E^M x_1 + \alpha_2 E^M x_2$
- $E^M A x = A E^M x$
- $E^{M \oplus N} x = E^M x + E^N x$ if $M \perp N$.

Definition: (***Kalman Filter setup***)
Consider the time discrete system

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)v(t) \\ y(t) &= C(t)x(t) + D(t)w(t) \end{aligned}$$

where $x \in \mathbb{R}^n$, $v \in \mathbb{R}^p$, $y \in \mathbb{R}^m$, $w \in \mathbb{R}^q$. Here $v(t)$ and $w(t)$ are white noises such that

$$E[v(t)v(s)^{\mathrm{T}}] = Q\delta_{ts}, \quad E[w(t)w(s)^{\mathrm{T}}] = R\delta_{ts}$$

where $Q \geq 0$ and $R > 0$ are covariance matrices (no correlation when $t \neq s$). Also, $E[v(t)] = 0$ and $E[w(t)] = 0$.

Here, $x(t)$ are vectors of random variables in a Hilbert space $\mathcal{H}$ with inner product $x_i(t) \cdot x_j(t) := E[x_i(t)x_j(t)]$.

Definition: (***Things used in Kalman Filtering***)
We are given previous measurements $y_1(0), \ldots, y_m(0), \ldots, y_1(t-1), \ldots, y_m(t-1)$. We want to find the best estimation $\hat{x}(t)$ of $x(t)$ based on these measurements. We minimize $\left\| x_i(t) - \hat{x}_i(t) \right\|^2$ (component-wise).

- $[y_{t-1}] = H_{t-1}y := \operatorname{span}\left\{ y_1(0), \ldots, y_m(0), \ldots, y_1(t-1), \ldots, y_m(t-1) \right\}$
- $\hat{x}(t) := E^{H_{t-1}} x(t)$ or $\hat{x}_i(t) := E^{H_{t-1}} x_i(t)$ for $i = 1, \ldots, n$
- $\tilde{y}(t) := y(t) - E^{H_{t-1}} y(t)$ called the innovation of $y(t)$
- $[\tilde{y}] := \operatorname{span}\{\tilde{y}_1, \ldots, \tilde{y}_m\}$
- $H_t(y) = H_{t-1}(y) \oplus [\tilde{y}]$ (note $H_{t-1}(y) \perp [\tilde{y}]$)
- $\hat{x}_t(t) := E^{H_t(y)} x(t)$
- $\hat{x}(t+1) := E^{H_t(y)} x(t+1)$
- $P(t) := E\left[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^{\mathrm{T}} \right]$
- $\tilde{x}(t) := x(t) - \hat{x}(t)$

Definition: (**Kalman Filter**)
A Kalman Filter is

$$\begin{cases} \hat{x}(t+1) &= (A - AK(t)C)\hat{x}(t) + AK(t)y(t) \\ K(t) &= P(t)C^{\mathrm{T}} \left(CP(t)C^{\mathrm{T}} + DRD^{\mathrm{T}} \right)^{-1} \quad \text{(Kalman gain)} \\ P(t+1) &= AP(t)A^{\mathrm{T}} - AP(t)C^{\mathrm{T}} \left(CP(t)C^{\mathrm{T}} + DRD^{\mathrm{T}} \right)^{-1} CP(t)A^{\mathrm{T}} + BQB^{\mathrm{T}} \end{cases}$$

Given $P(0)$ and $\hat{x}(0)$ (usually $\hat{x}(0) = 0$, because it is the projection onto the empty set) we can recursively estimate the state.

Mathematical results

Theorem:
Consider $x^2 + bx + c$.

- The roots have negative real parts if and only if $b > 0$ and $c > 0$.
- The roots have nonpositive real parts if and only if $b \geq 0$ and $c \geq 0$.
- If $b = 0$ and $c > 0$, then the real parts of the solutions are exactly 0.
- If $b > 0$ and $c = 0$, then one solution is 0 while the other is strictly real negative.

Theorem:
Consider $x^n + a_1 x^{n-1} + \cdots + a_n$. If all roots have negative real parts, then $a_1, \ldots, a_n > 0$.