6. Polynomial Matrices, State Estimation

- Theory of Polynomial Matrices
  - Smith Normal Form & Smith-McMillan Form
  - Poles & Zeros of MIMO Systems, Transmission Zeros
  - Connection to Controllability & Observability
- State Estimation
  - Separation Theorem
- Sylester Equation in Control Theory.
- Riccati Equation

Theory of Polynomial Matrices

- $\mathbb{R}[s]^{p\times m}$ denotes the set of $p \times m$ real polynomial matrices on s.
- A real polynomial matrix is a matrix valued function whose entries are polynomials with real coefficients. Eg. $P(s) = \begin{bmatrix} (s+1) & (s+2) \\ (s+3) & (s+4) \end{bmatrix}$.

Nonsingular & Unimodular Polynomial Matrix

Consider a square $(p = m)$ real polynomial matrix. Then, $P(s)$ is called

- **nonsingular** if $\texttt{det}[P(s)]$ is a non-zero polynomial
- **unimodular** if $\texttt{det}[P(s)] \neq 0$

Determinantal Divisors of $P(s)$: Polynomials $\{D_i(s) \mid i \in [0, r]\}$ with $D_0(s) = 1, D_i(s)$ is the monic gcd of all $\neq 0$ minors of $P(s)$ of order i.

Fact: $r = \texttt{rank}[P(s)]$ drops precisely at the roots of $D_r(s)$.

Rank of a Transfer Function Matrix (TFM)

Normal Rank of a transfer function matrix

Normal rank of a TFM $P(s) \in \mathbf{R}[s]^{p\times m}$, denoted by $\overline{\texttt{rank}}[P(s)]$ is said to be equal to r if $\texttt{rank}[P(s)] = r$ for **almost all** values of s.

For eg., for the TFM $P(s) = \begin{bmatrix} \frac{1}{s+1} & \frac{1}{s} \\ \frac{1}{s} & \frac{1}{s+1} \end{bmatrix}$, $\texttt{det}[P(s)] = \frac{-(2s+1)}{s^2(s^2+2s+1)}$.

$\overline{\texttt{rank}}[P(s)] = 2$. But, $\texttt{det}\left[P\left(\frac{-1}{2}\right)\right] = 0 \implies \texttt{rank}\left[P\left(\frac{-1}{2}\right)\right] = 1$

Facts:

- $\overline{\texttt{rank}}[P(s)] \leq \min(m,p)$.
- If $\overline{\texttt{rank}}[P(s)] = p$, then $\exists$ right inverse of $P(s)$ s.t. $P(s)P^{-R}(s) = I_p$
- If $\overline{\texttt{rank}}[P(s)] = m$, then $\exists$ left inverse of $P(s)$ s.t. $P^{-L}(s)P(s) = I_m$

System Inverse

- When $p = m$, we say system has an inverse $P^{-1}(s)$ satisfying $P(s)P^{-1}(s) = P^{-1}(s)P(s) = I_p$.
- $P^{-1}(s)$ is usually improper and does not have a state space realisation.
- Many control theory approaches tend to construct a rational approximation of $P^{-1}(s)$ which though desirable can lead to robustness issues.
- Suppose that $p = m$ and $\exists D^{-1}$. Then a state space realisation of $P^{-1}(s)$ is given by

$$\dot{z} = (A - BD^{-1}C)z + BD^{-1}v$$

$$w = -D^{-1}Cz + D^{-1}v$$

Rational Matrices

Monic Greatest Common Divisor (GCD)

The monic GCD of a family of polynomials is the monic polynomial of greatest order that divides all the polynomials in the family.

Monic Least Common Denominator (LCD)

The monic LCD of a family of polynomials is the monic polynomial of smallest order that is divided by all the polynomials in the family.

- $\mathbb{R}(s)^{p\times m}$ denotes the set of $p \times m$ real rational matrices on s.
- $G(s) \in \mathbb{R}(s)^{p\times m}$ is a matrix valued function whose entries are ratios of polynomials with real coefficients. Eg. $G(s) = \begin{bmatrix} 1/(s+1) & 1/s^2 \\ 1/(s+3) & 1/s \end{bmatrix}$.
- Any $G(s) \in \mathbb{R}(s)^{p\times m}$ can be written as $G(s) = \frac{1}{d(s)} N(s)$, where $d(s)$ is the monic LCD of all entires of $G(s)$ and $N(s) \in \mathbb{R}[s]^{p\times m}$.

Smith Form

Smith form of a real polynomial matrix $P(s) \in \mathbb{R}[s]^{p\times m}$ is the diagonal $S_P(s) \in \mathbb{R}[s]^{p\times m}$ defined by

$$S_P(s) = \begin{bmatrix} \mathtt{diag}(\epsilon_1(s), \dots, \epsilon_r(s)) & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(p-r)\times r} & \mathbf{0}_{(p-r)\times(m-r)} \end{bmatrix}, \quad \text{where } \mathtt{rank}[P(s)] = r,$$

$\epsilon_i(s) = \dfrac{D_i(s)}{D_{i-1}(s)}, \quad i = \{1, 2, \dots, r\}$ are the **invariant factors** of $P(s)$.

Smith Form Factorisation Theorem

For every $P(s) \in \mathbb{R}[s]^{p\times m}$ with Smith form $S_P(s) \in \mathbb{R}[s]^{p\times m}$, $\exists$ unimodular matrices $L(s) \in \mathbb{R}[s]^{p\times p}, R(s) \in \mathbb{R}[s]^{m\times m}$ which can be found using Gaussian elimination procedure such that

$$P(s) = L(s) S_P(s) R(s).$$

Smith-McMillan Form

Smith-McMillan form of a real rational matrix $G(s) \in \mathbb{R}(s)^{p\times m}$ is the diagonal $SM_G(s) \in \mathbb{R}(s)^{p\times m}$ defined by

$$SM_G(s) = \frac{1}{d(s)} S_N(s) = \begin{bmatrix} \mathtt{diag}\left(\frac{\eta_1(s)}{\psi_1(s)}, \dots, \frac{\eta_1(r)}{\psi_r(s)}\right) & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(p-r)\times r} & \mathbf{0}_{(p-r)\times(m-r)} \end{bmatrix}, \text{ where}$$

- $S_N(s) \in \mathbb{R}[s]^{p\times m}$ denotes the Smith form of $N(s)$
- $\{\eta_i(s), \psi_i(s)\}, i = 1, \dots, r$ are all co-prime.

Smith-McMillan Form Factorisation Theorem

For every $G(s) \in \mathbb{R}(s)^{p\times m}$ with Smith-McMillan form $SM_G(s)$, $\exists$ unimodular matrices $L(s) \in \mathbb{R}[s]^{p\times p}, R(s) \in \mathbb{R}[s]^{m\times m}$ such that

$$G(s) = \frac{1}{d(s)} S_N(s) = L(s)\, SM_G(s)\, R(s).$$

McMillan Degree, Poles & Zeros

Smith-McMillan form is used to define poles & zeros of rational matrices.

Pole & Zero Polynomial

For a real rational matrix $G(s) \in \mathbb{R}(s)^{p\times m}$ with Smith-McMillan form $SM_G(s)$, define the following polynomials

Zero Polynomial of $G(s)$: $z_G(s) := \eta_1(s)\eta_2(s)\ldots\eta_r(s)$

Transmission Zeros of $G(s)$: Roots of $z_G(s)$

Pole Polynomial of $G(s)$: $p_G(s) := \psi_1(s)\psi_2(s)\ldots\psi_r(s)$

McMillan Degree of $G(s)$: degree of $p_G(s)$

Fact: A scalar rational function can't have a pole and zero at the same location. However, the matrix case $G(s) \in \mathbb{R}(s)^{p\times m}$ can have.

General Description of a Linear System

Physical system can be described by linear differential equations with input u, output y and internal physical variables η as follows

$$\begin{cases} P(s)\eta = Q(s)u \\ y = R(s)\eta + W(s)u \end{cases} \iff \underbrace{\begin{bmatrix} P(s) & Q(s) \\ -R(s) & W(s) \end{bmatrix}}_{:=P_\Sigma(s)} \begin{bmatrix} -\eta \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

Then the transfer function is $G(s) = R(s)P^{-1}(s)Q(s) + W(s)$.

Rosenbrock System Matrix

Representation bridging state-space & transfer function matrix form. For a CT LTI system, its Rosenbrock system matrix is given by

$$P_\Sigma(s) = \begin{bmatrix} sI - A & B \\ -C & D \end{bmatrix} \in \mathbb{R}[s]^{(n+p)\times(n+m)}$$

Rosenbrock Matrix & Controllability/Observability

Two systems $P_{\Sigma_1}(s), P_{\Sigma_2}(s)$ are said to be **equivalent** if $\exists$ unimodular matrices $U(s), V(s)$ and polynomial matrices $X(s), Y(s)$ such that

$$\begin{bmatrix} U(s) & 0 \\ X(s) & I \end{bmatrix} \begin{bmatrix} P_1(s) & Q_1(s) \\ -R_1(s) & W_1(s) \end{bmatrix} \begin{bmatrix} V(s) & Y(s) \\ 0 & I \end{bmatrix} = \begin{bmatrix} P_2(s) & Q_2(s) \\ -R_2(s) & W_2(s) \end{bmatrix}$$

State Space Equivalence of Rosenbrock System Matrix

Any Rosenbrock system matrix is equivalent to one in state space form

$$\begin{bmatrix} P(s) & Q(s) \\ -R(s) & W(s) \end{bmatrix} \sim \begin{bmatrix} sI - A & B \\ -C & J(s) \end{bmatrix}.$$

Equivalence

- Controllability $\iff$ (P, Q) or ($\iff$ $(sI - A, B)$) left co-prime
- Observability $\iff$ (P, R) or ($\iff$ $(sI - A, C)$) right co-prime

Output Zeroing Problem & Zero Dynamics

Consider the CT LTI (MIMO) system $\dot{x} = Ax + Bu, y = Cx$.

Output Zeroing Problem

If possible, find a control u & initial state x_0 such that $y(t) = 0, \forall t \geq 0$.

If above problem has a solution, then we can define the zero dynamics.

Zero Dynamics

The dynamics of the CT LTI system restricted to the set of initial conditions defining the o/p zeroing problem is called the **zero dynamics**.

Sylvester Inequality: If A, B are two matrices of the same order n, then

$$\mathtt{rank}(A) + \mathtt{rank}(B) \leq \mathtt{rank}(AB) + n$$

Normal Rank of a Rosenbrock System Matrix

Lemma

$\overline{\texttt{rank}}[P_\Sigma(s)] = n + \overline{\texttt{rank}}[P(s)]$

Proof: We can see that

$$\begin{bmatrix} I & 0 \\ C(sI-A)^{-1} & I \end{bmatrix} \begin{bmatrix} sI-A & B \\ -C & D \end{bmatrix} = \begin{bmatrix} sI-A & B \\ 0 & P(s) \end{bmatrix}.$$

$$\Longrightarrow \ \overline{\texttt{rank}}[LHS] = \overline{\texttt{rank}}[RHS] = n + \overline{\texttt{rank}}[P(s)], \quad \text{Sylvester Inequality}$$

Crucial Assumptions:

1. $\overline{\texttt{rank}}[P(s)] = \min(p,m) \quad \Longleftrightarrow \quad \overline{\texttt{rank}}[P_\Sigma(s)] = n + \min(p,m).$
2. (A,B) is controllable and (A,C) is observable.

Transmission Zeros

Transmission Zero

Suppose all assumptions are satisfied. Then $z \in \mathbb{C}$ is a **transmission zero** of system (A, B, C, D) if $\overline{\texttt{rank}}[P_\Sigma(z)] < n + \min(p, m)$

For scalar systems with $m = p = 1$, we see that

$$P(s) = \frac{N(s)}{D(s)} = \frac{\texttt{det}(P_\Sigma(s))}{\texttt{det}(sI - A)}, \quad \text{and} \quad \texttt{det}\left(\begin{bmatrix} I & 0 \\ C(sI-A)^{-1} & I \end{bmatrix}\right) = 1$$

$$\implies \texttt{det}\left(P_\Sigma(s)\right) = \texttt{det}\left((sI - A)\right)\texttt{det}(P(s)) = \texttt{det}\left((sI - A)\right)P(s).$$

- SISO Transmission zeros = zeros of $P(s)$ (given assumption 1 holds)
- For a MIMO TFM, the TFM loses rank given that it can have poles & zeros at the same s.

Transmission Zeros

- Transmission zeros are associated with modes of behavior wherein the input and states of a system are nonzero, yet the output equals zero.
- Suppose that both assumptions hold true and that z is a transmission zero but not a pole of $P(s)$. Then $\texttt{rank}(P(z)) < \min(p, m)$.

Transmission Blocking Property

Suppose that both assumptions hold true and that z is a transmission zero with $p \geq m$. Then, $\exists u(t) = u_0 e^{zt}, u_0 \neq 0$ and x_0 such that $y(t) = 0, \forall t \geq 0$. Further, if $\texttt{Real}(|\lambda_i(A)|) < 0, \forall i$, then $y(t) \to 0, \forall x_0$.

u_0: Input Zero Direction

x_0: Zero State Direction

Invariant Zeros

- The **invariant zero polynomial** of a state space of the CL LTI is the monic GCD $z_P(s)$ of all nonzero minors of order $r = \texttt{rank}[P_\Sigma(s)]$.
- The roots of $z_P(s)$ are called the **invariant zeros** of the state space.
- Transmission zeros are defined in frequency domain for TFM, while invariant zeros are defined in time-domain for state space realisations.
- Both transmission zeros & invariant zeros have transmission blocking property.

Important Facts

- $\{$Poles of $G(s)\} \subset \{$Eigenvalues of $A\}$
- $\{$Transmission zeros of $G(s)\} \subset \{$Invariant Zeros of LTI$\}$

Transmission Blocking Property

Since, $\texttt{rank}(P(z)) < \min(p, m)$, $\exists$ a non-trivial nullspace of $P_\Sigma(s)$.

$$\begin{bmatrix} zI - A & B \\ -C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Consider the input $u(t) = u_0 e^{zt}$ with $u_0 = -u$ and $x_0 = x$. Then,

$$\begin{bmatrix} zI - A & B \\ -C & D \end{bmatrix} \begin{bmatrix} x_0 \\ -u_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies (zI - A)x_0 = Bu_0, \quad Cx_0 + Du_0 = 0.$$

$$\begin{aligned} X(s) &= (sI - A)^{-1}x_0 + (sI - A)^{-1}B\frac{u_0}{s - z}, \quad \left(\text{substituting } U(s) = \frac{u_0}{s - z}\right) \\ &= (sI - A)^{-1}\left(x_0 + B\frac{u_0}{s - z}\right) = (sI - A)^{-1}\left(x_0 + \frac{(zI - A)x_0}{s - z}\right) \\ &= (sI - A)^{-1}\left((sI - A)\frac{x_0}{s - z}\right) = \frac{x_0}{s - z} \implies x(t) = x_0 e^{zt} \end{aligned}$$

$$y(t) = Cx(t) + Du(t) = Cx_0e^{zt} + Du_0e^{zt} = (Cx_0 + Du_0)e^{zt} = 0.$$

Minimal Realisation & BIBO Stability

Proposition

Assuming the LTI realisation of $G(s)$ is minimal, the TFM $G(s)$ is BIBO stable iff the LTI realisation is (internally) asymptotically stable.

Proof: LTI Asymptotic Stability $\iff \texttt{Real}[\lambda(A)] < 0$. For BIBO stability, we need the poles to have strict negative real parts. For minimal realisation, set of poles of $G(s) =$ the set of eigenvalues of A.

Minimality & McMillan Degree

LTI realisation is minimal iff McMillan Degree of $G(s) = n$. That is,

$$p_G(s) = \texttt{det}[SI - A], \quad z_G(s) = z_P(s).$$

State Estimation - Motivation

Consider the CT-LTI system with $\dot{x} = Ax + Bu, \quad y = Cx + Du$.

- When output $y = x$ (entire state can be measured exactly), we can design a control law $u = Kx$ to stabilise the above system.
- But when $y \neq x$, full state feedback control law is not possible.
- However, we also saw at a particular instant of time, the Gramian-based reconstruction of state on $[t_0, t_f]$.
- Need continuous estimate of state to implement control like $u = Kx$.

Open-Loop State Estimate

Simplest state estimate consists of copy of the original system

$$\dot{\hat{x}} = A\hat{x} + Bu.$$

State Estimate & Estimation Error

Define the state estimation error as $e = \hat{x} - x$, where $\hat{x}$ is the state estimate. If A is Hurwitz, $e \to 0$ exponentially fast $\forall u$ as follows

$$\dot{e} = \dot{\hat{x}} - \dot{x} = A\hat{x} + Bu - Ax - Bu = A(\hat{x} - x) = Ae.$$

Even if A is not Hurwitz, we can construct asymptotically correct state estimate using a **closed-loop** state estimator via **o/p injection**

$$\dot{\hat{x}} = A\hat{x} + Bu - L(\hat{y} - y), \quad \hat{y} = C\hat{x} + Du,$$

where $L \in \mathbb{R}^{n \times p}$ is called **o/p injection (or estimator) gain matrix**.

$$\implies \dot{e} = \dot{\hat{x}} - \dot{x} = A\hat{x} + Bu - L(\hat{y} - y) - Ax - Bu = (A - LC)e.$$

If L makes $A - LC$ Hurwitz, then $e \to 0$ exponentially fast $\forall u$.

Stabilisation By Output Feedback

Eigenvalue Assignment Theorem

When (A, C) is detectable, $\exists L \in \mathbb{R}^{n\times p}$ such that $A - LC$ is Hurwitz.

Then, given any set of $\Lambda = \{\lambda_i\}_{i=1}^n, \lambda_i \in \mathbb{C}, \exists L \in \mathbb{R}^{n\times p}$ such that $\texttt{eig}(A - LC) = \Lambda$. (Conditions are both iff)

The control law $u = K\hat{x}$ results in controller with the state space model

$$\dot{\hat{x}} = \underbrace{(A - LC - BK + LDK)}_{:=\tilde{A}}\hat{x} + Ly \iff C(s) = -K(sI - \tilde{A})^{-1}L.$$

Consider the closed loop system with states $\bar{x} = \begin{bmatrix} x^\top & e^\top \end{bmatrix}^\top$. Then,

$$\dot{x} = Ax + Bu = Ax - BK\hat{x} = Ax - BK(x + e) = (A - BK)x - BKe$$

$$\implies \begin{bmatrix} \dot{x} \\ \dot{e} \end{bmatrix} = A_{cl}\begin{bmatrix} x \\ e \end{bmatrix}, \quad \text{where} \quad A_{cl} = \begin{bmatrix} A - BK & -BK \\ 0 & A - LC \end{bmatrix}.$$

Separation Theorem

Separation Theorem

- The eigenvalues of A_{cl} with the output feedback controller $u = -K\hat{x}$ are the union of eigenvalues of state feedback closed loop matrix $A - BK$ with the eigenvalues of the state estimator matrix $A - LC$.
- This means that one can design the state feedback gain K and the estimator gain L matrices independently.

Questions: How can we

1. design the optimum $K_\star$ for a given CT-LTI? LQR.
2. design the optimum $L_\star$ for a given CT-LTI with noisy output? KF.
3. find the optimum $(K_\star, L_\star)$ instead of any (K, L)? LQG (LQR+KF).

Eigenvalue Problems for a Linear Operator

Eigenvalue Problem

Let $\mathcal{A}$ be a linear operator. Eigenvalue Problem for operator $\mathcal{A}$ involves finding non-trivial function-number pairs $(v(x), \lambda)$ that solves

$$\mathcal{A}v(x) = \lambda v(x) \quad \Longleftrightarrow \quad \mathcal{A}v(x) - \lambda v(x) = 0. \tag{1}$$

- The functions $v(x)$ that solve (1) are called **eigenfunctions** and each corresponds to an eigenvalue λ.
- Often the eigenfunctions form a basis for the underlying vector space, and the set is is called **complete**.
- Completeness property is useful for PDEs as it allows for solutions to be expressed as linear combinations of eigenfunctions.

Sylvester Equation

Spectrum of a Linear Operator

Let $\mathcal{A}$ be a linear operator. Then, the (eigen) discrete spectrum of the operator $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$ refers to the set of all eigenvalues of $\mathcal{A}$.

In control theory, we often solve Sylvester equations.

Sylvester Equation

Given $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{m\times m}$ and $C \in \mathbb{R}^{n\times m}$, the Sylvester equation

$$AX - XB = C. \tag{2}$$

Equation (2) has an unique solution $X \in \mathbb{R}^{n\times m}$ iff $\sigma(A) \cap \sigma(B) = \emptyset$.

- Sylvester equation is a linear equation in matrix variable X
- $AX + XA^\top = C$ (Lyapunov eqn) special case of (2) with $B = -A^\top$.

Sylvester Operator

Consider the Sylvester matrix equation in X of the form

$$AX + XB = Q. \tag{3}$$

Solutions & properties of (3) are determined by the Sylvester operator $\mathfrak{L}$

Sylvester operator

Given A, B, the Sylvester operator in variable X is $\mathfrak{L}(X) = AX + XB$.

Solvability of Sylvester Equations: $X \in \mathcal{N}(\mathfrak{L}) \iff X = VD$, where D is diagonal matrix & V corresponds to right eigenvector matrix of A.

$$AX - XB = C \quad \iff (I_m \otimes A + B^\top \otimes I_n)\,\mathtt{vec}(X) = \mathtt{vec}(C). \tag{4}$$

Cost of solving (2):

- Gaussian elimination is costly as it costs $O((mn)^3)$
- Bartels-Stewart efficient algorithm takes only $O(\max(m,n)^3)$

The Algebraic Riccati Equation

- The Riccati equation occurs in optimal control problems
- The algebraic Riccati equation is a nonlinear quadratic matrix equation that also can be expressed via an eigenvalue problem.

Continuous Algebraic Riccati Equation

Given A, B matrices and the cost matrices $Q \succeq 0, R \succ 0$, the CT optimal control problem poses the following algebraic Riccati equation with X denoting the symmetric solution matrix.

$$A^\top X + XA - XBR^{-1}B^\top X + Q = 0. \tag{5}$$

Rewriting LHS of Riccati equation as a pure quadratic, we get

$$\begin{bmatrix} I \\ X \end{bmatrix} \begin{bmatrix} Q & A \\ A^\top & -BR^{-1}B^\top \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = 0.$$

The Algebraic Riccati Equation

We can also characterise (5) by the relation

$$\begin{bmatrix} Q & A \\ A^\top & -BR^{-1}B^\top \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \iff \underbrace{\begin{bmatrix} A & -BR^{-1}B^\top \\ -Q & -A^\top \end{bmatrix}}_{:=Z} \begin{bmatrix} I \\ X \end{bmatrix} = 0.$$

We want a specific basis of the null space of a Hamiltonian matrix. That is, a matrix Z such that JZ is symmetric with $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

Hamiltonian Matrix

A matrix $M \in \mathbb{R}^{2n \times 2n}$ is called Hamiltonian if JM is symmetric. So,

$$JM = (JM)^\top \implies M^\top J + JM = 0.$$