Chapter 9

State Space Systems

"How can we represent mathematically a dynamical system accepting $m \geq 1$ input signals, and outputting $p \geq 1$ signals?"

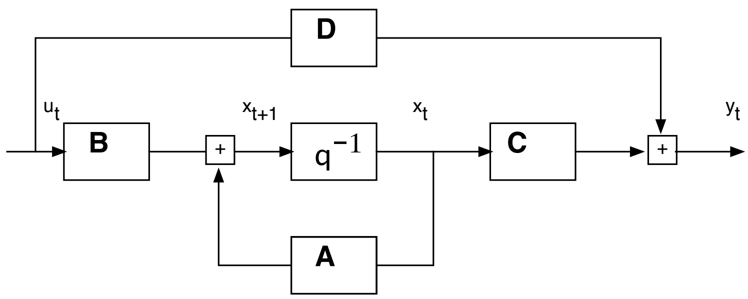


Figure 9.1: Schematic representation of a State Space System with matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$. The feedback is implemented by the system matrix $\mathbf{A}$, governing the dynamic behavior of the system.

9.1 State Space Model

A deterministic state-space model is given as

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t \\ \mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t, \end{cases} \quad \forall t = -\infty, \ldots, \infty. \tag{9.1}$$

where we have

- $\{\mathbf{x}_t\}_t \subset \mathbb{R}^n$ the state process.
- $\{\mathbf{u}_t\}_t \subset \mathbb{R}^p$ the input process.
- $\{\mathbf{y}_t\}_t \subset \mathbb{R}^q$ the output process.

- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the system matrix.
- $\mathbf{B} \in \mathbb{R}^{n \times p}$ the input matrix.
- $\mathbf{C} \in \mathbb{R}^{q \times n}$ the output matrix.
- $\mathbf{D} \in \mathbb{R}^{q \times p}$ the feed-through matrix.

The 'state variable' has different interpretations.

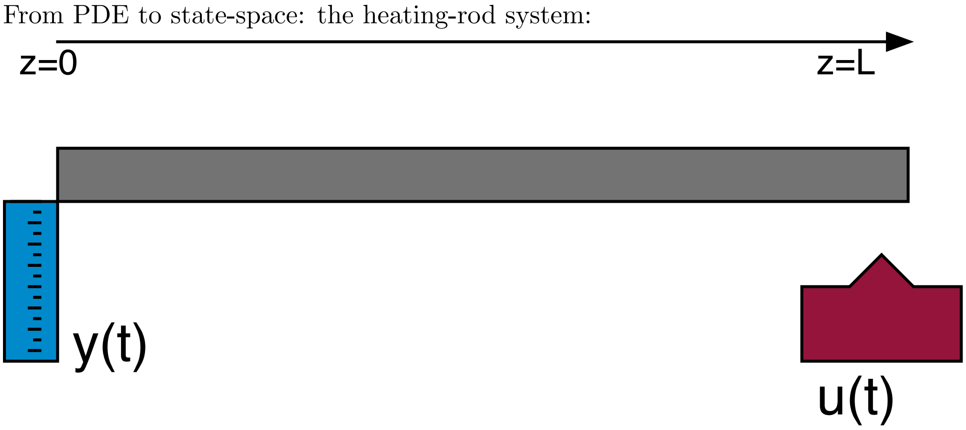
1. Representation of hidden 'state' of system (physical).
2. Summarization of what to remember from past.
3. Compact representation of information relevant to predict future.
4. Intersection of past and future.
5. Optimal estimate of the model parameters thus far (RLS).

Each interpretation leads to own algorithms. Advantages of the use of state-space systems over fractional polynomial models include

- Those models are typically closer to physical modeling.
- They are more appropriate for modeling MIMO systems.
- Such models make a clearer distinction between input noise, measurement noise and Innovations.
- Any LTI can be represented as a state-space model of sufficient order (Canonical representation).
- The study of the dynamics of the system concerns 'only' the $\mathbf{A}$ matrix. The matrices $\mathbf{B}, \mathbf{C}, \mathbf{D}$ 'make up' the system, but do not directly influence the qualitative dynamical behavior.
- Problems of identiafibility are readily phrased in this context.

9.1.1 State Space System - example 1

From PDE to state-space: the heating-rod system:



Let $x(t, z)$ denote temperature at time t, and location z on the rod.

$$\frac{\partial x(t,z)}{\partial t} = \kappa \frac{\partial^2 x(t,z)}{\partial z^2} \tag{9.2}$$

The heating at the far end mens that

$$\frac{\partial x(t,z)}{\partial z}\Big|_{z=L} = Ku(t), \tag{9.3}$$

The near-end is insulated such that

$$\frac{\partial x(t,z)}{\partial z}\Big|_{z=0} = 0. \tag{9.4}$$

The measurements are

$$y(t) = x(t,0) + v(t), \forall t = 1,2,\ldots \tag{9.5}$$

The unknown parameters are

$$\theta = \begin{bmatrix} \kappa \\ K \end{bmatrix} \tag{9.6}$$

This can be approximated as a system with n states

$$\mathbf{x}(t) = \Big(x(t,z_1), x(t,z_2), \ldots, x(t,z_n)\Big)^T \in \mathbb{R}^n \tag{9.7}$$

with $z_k = L(k-1)/(n-1)$.. Then we use the approximation that

$$\frac{\partial^2 x(t,z)}{\partial z^2} \approx \frac{x(t,z_{k+1}) - 2x(t,z_k) + x(t,z_{k-1})}{(L/(n-1))^2} \tag{9.8}$$

where $z_k = \text{argmin}_{z_1,,\ldots,z_n} \|z - z_k\|$. Hence the continuous state-space approximation becomes

$$\begin{cases} \dot{\mathbf{x}}(t) = \left(\frac{n-1}{L}\right)^2 \kappa \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ \vdots \\ K \end{bmatrix} u(t) \\ y(t) = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \mathbf{x}(t) + v(t) \end{cases} \tag{9.9}$$

and a discrete Euler approximation

$$\begin{cases} \mathbf{x}_{t+1} - \mathbf{x}_t & = \Delta' \left(\frac{n-1}{L}\right)^2 \kappa \\ & \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix} \mathbf{x}_t + \Delta' \begin{bmatrix} 0 \\ \vdots \\ K \end{bmatrix} \int_{\Delta'} u(t) \\ y_t & = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \mathbf{x}_t + \int_{\Delta'} v(t) \end{cases} \tag{9.10}$$

9.1.2 State Space System - example 2

Models for the future size of the population (UN, WWF).

Leslie model: key ideas: discretize population in n aging groups and

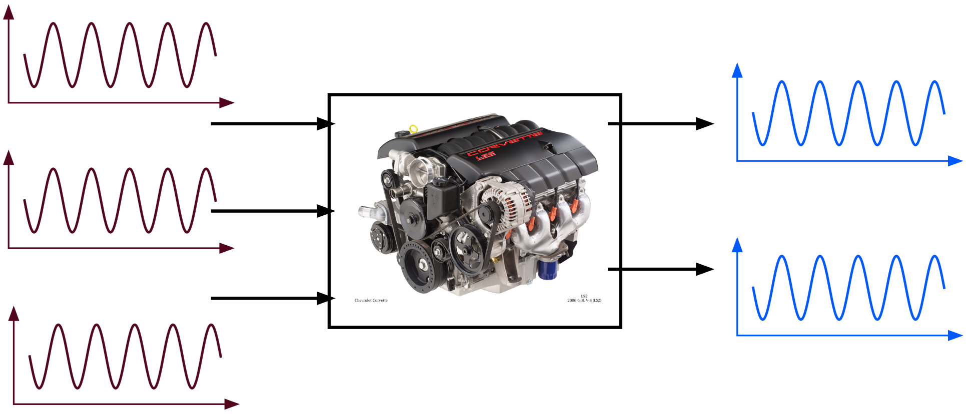
- Let $\mathbf{x}_{t,i} \in \mathbb{R}^+$ denote the size of the ith aging group at time t.
- Let $\mathbf{x}_{t+1,i+1} = s_i \mathbf{x}_{t,i}$ with $s_i \geq 0$ the 'survival' coefficient.
- Let $\mathbf{x}_{t+1,1} = s_0 \sum_{i=1}^n f_i \mathbf{x}_{t,i}$ with $f_i \geq 0$ the 'fertility' rate.

Hence, the dynamics of the population may be captured by the following discrete time model

$$\begin{cases} \mathbf{x}_{t+1} = \begin{bmatrix} s_0 f_1 & s_0 f_2 & \dots & s_0 f_n & \\ s_1 & 0 & & & \\ 0 & s_2 & 0 & & \\ & & \ddots & & \\ & & & s_{n-1} & 0 \end{bmatrix} \mathbf{x}_t + \mathbf{u}_t \\ y_t = \sum_{i=1}^n \mathbf{x}_{t,i} \end{cases} \tag{9.11}$$

9.1.3 Impulse Response to State Space System

What is now the relation of state-space machines, and the system theoretic tools seen in the previous Part?



Recall impulse response (SISO)

$$y_t = \sum_{\tau=0}^{\infty} h_\tau u_{t-\tau}, \tag{9.12}$$

and MIMO

$$\mathbf{y}_t = \sum_{\tau=0}^{\infty} \mathbf{H}_\tau \mathbf{u}_{t-\tau}, \tag{9.13}$$

where $\{\mathbf{H}_\tau\}_\tau \subset \mathbb{R}^{p\times q}$.

Recall: System identification studies method to build a model from observed input-output behaviors, i.e. $\{\mathbf{u}_t\}_t$ and $\{\mathbf{y}_t\}_t$.

Now it is a simple exercise to see which impulse response matrices $\{\mathbf{H}_\tau\}_\tau$ are implemented by a state-space model with matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$:

$$\mathbf{H}_\tau = \begin{cases} \mathbf{D} & \tau = 0 \\ \mathbf{C}\mathbf{A}^{\tau-1}\mathbf{B} & \tau = 1, 2, \dots \end{cases}, \quad \forall \tau = 0, 1, 2, \dots \tag{9.14}$$

Contrast with rational polynomials where typically

$$h_\tau \Leftrightarrow h(q^{-1}) = \frac{b_1 q^{-1} + b_2 q^{-2} + \dots}{1 + a_1 q^{-1} + a_2 q^{-2} + \dots} \tag{9.15}$$

Overlapping: consider FIR model

$$y_t = b_0 u_t + b_1 u_{t-1} + b_2 u_{t-2} + e_t \tag{9.16}$$

then equivalent state-space with states $\mathbf{x}_t = (u_t, u_{t-1}, u_{t-2})^T \in \mathbb{R}^3$ becomes

$$\begin{cases} \mathbf{x}_t &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}_{t-1} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u_t \\ y_t &= \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \mathbf{x}_t + e_t \end{cases} \tag{9.17}$$

and $\mathbf{x}_0 = (u_0, u_{-1}, u_{-2})^T$.

9.2 Realization Theory

9.2.1 Controllability and Observability

A state-space model is said to be *Controllable* iff for any terminal state $\mathbf{x} \in \mathbb{R}^n$ one has that for all initial state $\mathbf{x}_0 \in \mathbb{R}^n$, there exists an input process $\{\mathbf{u}_t\}_t$ which steers the model from state $\mathbf{x}_0$ to $\mathbf{x}$.

A state-space model is said to be *Reachable* iff for any initial state $\mathbf{x}_0 \in \mathbb{R}^n$ one has that for all terminal states $\mathbf{x} \in \mathbb{R}^n$ there exists an input process $\{\mathbf{u}_t\}_t$ which steers the model from state $\mathbf{x}_0$ to $\mathbf{x}$.

The mathematical definition goes as follows: Define the reachability matrix $\mathcal{C} \in \mathbb{R}^{n \times np}$ as

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{9.18}$$

The State space $(\mathbf{A}, \mathbf{B})$ is reachable (controllable) if

$$\text{rank}(\mathcal{C}) = n. \tag{9.19}$$

Intuition: if the matrix $\mathcal{C}$ is full rank, the image of $\mathcal{C}$ equals $\mathbb{R}^n$, and the superposition principle states that any linear combination of states can be reached by a linear combination of inputs.

A state-space model is Observable iff any two different initial states $\mathbf{x}_0 \neq \mathbf{x}'_0 \in \mathbb{R}^n$ lead to a different output $\{\mathbf{y}_s\}_{s \geq 0}$ of the state-space model in the future when the inputs are switched off henceforth (autonomous mode).

Define the Observability matrix $\mathcal{O} \in \mathbb{R}^{qn \times n}$ as

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{9.20}$$

Hence, a state-space model $(\mathbf{A}, \mathbf{C})$ is observable iff

$$\text{rank}(\mathcal{O}) = n \tag{9.21}$$

Intuition: if the (right) null space of $\mathcal{O}$ is empty, no two different $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ lead to the $\mathcal{O}\mathbf{x} = \mathcal{O}\mathbf{x}'$.

Let

$$\mathbf{u}_- = (\mathbf{u}_0, \mathbf{u}_{-1}, \mathbf{u}_{-2}, \dots)^T \tag{9.22}$$

And

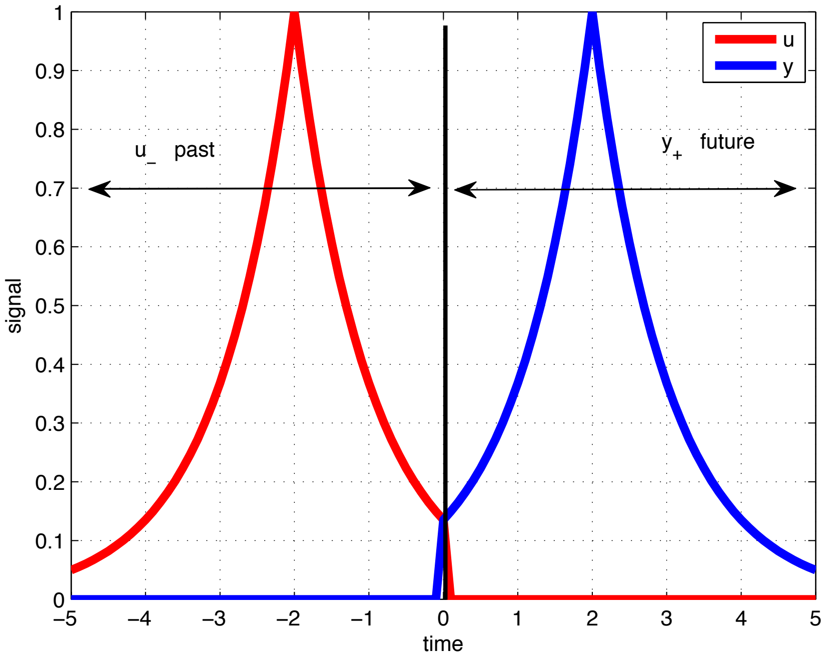
$$\mathbf{y}_+ = (\mathbf{y}_1, \mathbf{y}_2, \dots)^T \tag{9.23}$$

Then

$$\mathbf{x}_1 \propto \mathcal{C}\mathbf{u}_- \tag{9.24}$$

and

$$\mathbf{y}_+ \propto \mathcal{O}\mathbf{x}_1 \tag{9.25}$$



9.2.2 Kalman-Ho Realization

Problem: Given an impulse response sequence $\{\mathbf{H}_\tau\}_\tau$, can we recover $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$?

Def. Minimal Realization. A state-space model $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ is a minimal realization of order n iff the corresponding $\mathcal{C}$ and $\mathcal{O}$ are full rank, that is iff the model is reachable (observable) and controllable.

Thm. (Kalman) If $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ and $(\mathbf{A}', \mathbf{B}', \mathbf{C}', \mathbf{D}')$ are two minimal realizations of the same impulse response $\{\mathbf{H}_\tau\}$, then they are linearly related by a nonsingular matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$ such that

$$\begin{cases} \mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \\ \mathbf{B}' = \mathbf{T}^{-1}\mathbf{B} \\ \mathbf{C}' = \mathbf{C}\mathbf{T} \\ \mathbf{D}' = \mathbf{D} \end{cases} \tag{9.26}$$

Intuition: a linear transformation of the states does not alter input-output behavior; that is, the corresponding $\{\mathbf{H}_\tau\}_\tau$ is the same. The thm states that those are the only transformations for which this is valid.

Hence, it is only possible to reconstruct a minimal realization of a state-space model $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ from $\{\mathbf{H}_\tau\}_\tau$ *up to a linear transformation of the states.*

In case we only observe sequences $\{\mathbf{u}_t\}_{t\geq 1}$ and $\{\mathbf{y}_t\}_{t\geq 1}$, we have to account for the transient effects and need to estimate $\mathbf{x}_0 \in \mathbb{R}^n$ as well. This is in many situations crucial. The above thm. is extended to include $\mathbf{x}_0$ as well.

Now the celebrated Kalman-Ho realization algorithm goes as follows:

- Toeplitz-matrix

$$\mathbf{H}^n = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \mathbf{H}_3 & \dots & \mathbf{H}_n \\ \mathbf{H}_2 & \mathbf{H}_3 & \mathbf{H}_4 & & \\ & & \ddots & & \\ \mathbf{H}_n & & & & \mathbf{H}_{2n+1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{CB} & \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & \dots & \mathbf{CA}^{n-1}\mathbf{B} \\ \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & & & \\ & & & \ddots & \\ \mathbf{CA}^{n-1}\mathbf{B} & & & & \mathbf{CA}^{2n-1}\mathbf{B} \end{bmatrix} = \mathcal{O}\mathcal{C}$$

- The state space is identifiable up to a non-singular matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{H}^n = \mathcal{O}\mathcal{C} = \mathcal{O}\mathbf{T}\mathbf{T}^{-1}\mathcal{C} \tag{9.27}$$

.

- Then take the SVD of $\mathbf{H}^n$, such that

$$\mathbf{H}^n = \mathbf{U}\Sigma\mathbf{V}^T \tag{9.28}$$

with $\mathbf{U} \in \mathbb{R}^{pn\times n}, \mathbf{V} \in \mathbb{R}^{n\times nq}$ and $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n) \in \mathbb{R}^{n\times n}$.

- Hence a minimal realization is given as

$$\begin{cases} \mathcal{O}' = \mathbf{U}\sqrt{\Sigma} \\ \mathcal{C}' = \sqrt{\Sigma}\mathbf{V} \end{cases} \tag{9.29}$$

- From $\mathcal{O}', \mathcal{C}'$ it is not too difficult to extract $(\mathbf{A}, \mathbf{B}, \mathbf{C})$

9.2.3 An Example

Given an input

$$u = (1, 0, 0, 0, \dots)^T \tag{9.30}$$

and output signal

$$y = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots)^T \tag{9.31}$$

with system

$$y_t = y_{t-1} + y_{t-2},\ y_0 = 0, y_1 = u_1 \tag{9.32}$$

or SS with $\mathbf{x}_0 = (0,0)^T$ as

$$\begin{cases} \mathbf{x}_{t+1} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_t \\ y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_t \end{cases} \tag{9.33}$$

or transfer function

$$G(z) = \frac{z}{z^2 - z - 1} \tag{9.34}$$

Now realization

$$\mathbf{H}_5 = \begin{bmatrix} 1 & 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & 5 & 8 \\ 2 & 3 & 5 & 8 & 13 \\ 3 & 5 & 8 & 13 & 21 \\ 5 & 8 & 13 & 21 & 34 \end{bmatrix} \tag{9.35}$$

Then SVD of $\mathbf{H}_5$ gives $\sigma_1 = 54.5601$ and $\sigma_2 = 0.4399$, and a minimal realization is

$$\begin{cases} \mathbf{x}'_{t+1} = \begin{bmatrix} 1.6179 & 0.0185 \\ 0.0185 & -0.6179 \end{bmatrix} \mathbf{x}'_t + \begin{bmatrix} 0.8550 \\ -0.5187 \end{bmatrix} u_t \\ y_t = \begin{bmatrix} 0.8550 & -0.5187 \end{bmatrix} \mathbf{x}_t \end{cases} \tag{9.36}$$

9.3 Stochastic Systems

Now let us see how one can use a state-space system to describe a system which is driven by stochastic signals. Recall that we adopt the convention to indicate signals which are stochastically as capital letters. It is for example immediate that the states will be a stochastic process too with the rules of probability governing its behavior. A Stochastic State Space System takes the general form:

$$\begin{cases} X_{t+1} &= \mathbf{A} X_t + W_t \\ Y_t &= \mathbf{C} X_t + V_t \end{cases} \tag{9.37}$$

with

- $\{X_t\}_t$ the stochastic state process taking values in $\mathbb{R}^n$.
- $\{Y_t\}_t$ the stochastic output process, taking values in $\mathbb{R}^p$.
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the (deterministic) system matrix.
- $\mathbf{C} \in \mathbb{R}^{p \times n}$ the (deterministic) output matrix.
- $\{W_t\}_t$ the stochastic process disturbances taking values in $\mathbb{R}^n$.
- $\{V_t\}_t$ the stochastic measurement disturbances taking values in $\mathbb{R}^p$.

The stochastic vectors follow a probability law assumed to follow

- $\mathbb{E}[W_t] = 0_n$, and $\mathbb{E}[W_t W_s^T] = \delta_{s,t}\mathbf{Q} \in \mathbb{R}^{n\times n}$.
- $\mathbb{E}[V_t] = 0_p$, and $\mathbb{E}[V_t V_s^T] = \delta_{s,t}\mathbf{R} \in \mathbb{R}^{p\times p}$.
- $\mathbb{E}[W_t V_t^T] = \delta_{s,t}\mathbf{S} \in \mathbb{R}^{n\times p}$.
- W_t, V_t assumed independent of $\dots, X_t$.

Main questions:

- Covariance matrix states $\mathbb{E}[X_t X_t^T] = \Pi$:

$$\Pi = \mathbf{A}\Pi\mathbf{A}^T + \mathbf{Q} \tag{9.38}$$

  - Lyapunov, stable.
- Covariance matrix outputs $\mathbb{E}[Y_t Y_t^T]$.

This model can equivalently be described in its innovation form

$$\begin{cases} X'_{t+1} &= \mathbf{A}X'_t + \mathbf{K}D_t \\ Y_t &= \mathbf{C}X'_t + D_t \end{cases} \tag{9.39}$$

with $\mathbf{K} \in \mathbb{R}^{n\times p}$ the Kalman gain, such that $\mathbf{P}, \mathbf{K}$ solves

$$\begin{cases} \mathbf{P} = \mathbf{APA} + (\mathbf{G} - \mathbf{APC}^T)(\Lambda_0 - \mathbf{CPC}^T)^{-1}(\mathbf{G} - \mathbf{APC}^T)^T \\ \mathbf{K} = (\mathbf{G} - \mathbf{APC}^T)(\Lambda_0 - \mathbf{CPC}^T) \end{cases} \tag{9.40}$$

and

- $\mathbb{E}[D_t {D'_t}^T] = (\Lambda_0 - \mathbf{CPC}^T)$
- $\mathbf{P} = \mathbb{E}[X'_t {X'_t}^T]$

9.4 Conclusions

- State-space systems for MIMO - distributed parameter systems.
- Relation impulse response - state-space models.
- Controllability - Observability
- Kalman - Ho
- Stochastic Systems